

Amazon Mechanical Turk

Crowdsourcing and Human Computation

Instructor: Chris Callison-Burch

Website: crowdsourcing-class.org

[Home](#)[Create](#)[Manage](#)[Developer](#)[Help](#)[New Project](#) New Batch with an Existing Project

Create HITs individually

Start a New Project

Categorization

[Data Collection](#)[Moderation of an Image](#)[Sentiment](#)[Survey](#)[Survey Link](#)[Tagging of an Image](#)[Transcription from A/V](#)[Transcription from an Image](#)[Writing](#)[Other](#)

Example of Categorization

Choose the best category for this image



- kitchen
- living
- bath
- bed
- outside

[View Instructions ↓](#)

Select the room location in home for this picture. Seating areas outside are outside not living. Offices or dens are living not bedrooms. Bedrooms should contain a bed in the picture.

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

[Create Project »](#)

Categorization

Choose the best category for this image



[View Instructions ↓](#)

Select the room location in home for this picture. Seating areas outside are outside not living. Offices or dens are living not bedrooms. Bedrooms should contain a bed in the picture.

- kitchen
- living
- bath
- bed
- outside

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

Data Collection

Find the Website Address for this Restaurant

- For this restaurant below, enter the website address for the official website of the restaurant
- Include the full address, e.g. <http://www.thecheesecakefactory.com>
- Do not include URLs to city guides and listings like Citysearch.

Restaurant Name: **Olive Garden**

Address: **310 Strander Blvd Tukwila, WA 98188**

Phone Number: **(206) 241-4899**

Website Address:

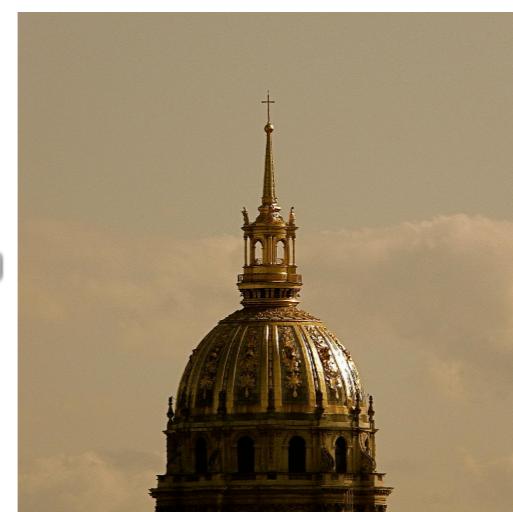
You must ACCEPT the HIT before you can submit the results.

Image Moderation

Select all images containing adult content

Guidelines for flagging an image as adult content. Flag the image if you consider any of the following to be true.

- Does the image contain nudity?
- Does the image portray hate or hate crimes?
- Does the image contain bloody violence?
- Does the image contain offensive gestures?



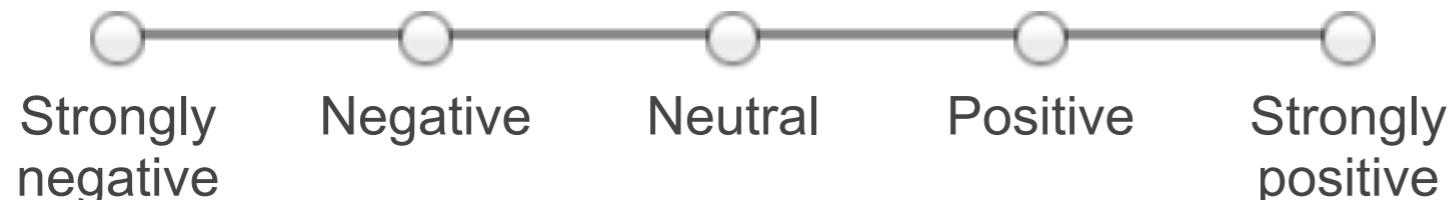
Sentiment

Pick the best sentiment based on the following criterion.

Strongly positive	Select this if the item embodies emotion that was extremely happy or excited toward the topic. For example, "Their customer service is the best that I've seen!!!!"
Positive	Select this if the item embodies emotion that was generally happy or satisfied, but the emotion wasn't extreme. For example, "Sure I'll shop there again."
Neutral	Select this if the item does not embody much of positive or negative emotion toward the topic. For example, "Yeah, I guess it's ok." or "Is their customer service open 24x7?"
Negative	Select this if the item embodies emotion that is perceived to be angry or upsetting toward the topic, but not to the extreme. For example, "I don't know if I'll shop there again because I don't trust them."
Strongly negative	Select this if the item embodies negative emotion toward the topic that can be perceived as extreme. For example, "These guys are terrific... NOTTTT!!!!!!" or "I will NEVER shop there again!!!"

Judge the sentiment expressed by the following item toward: Amazon

If you loved Firefly TV show, amazing Amazon price for entire series: about \$27 BlueRay & \$17 DVD.



Surveys

1. What is your gender?

Male Female

2. What is your age?

3. Which of the following best describes your highest achieved education level?

Some High School

4. What is the total income of your household?

Less than \$12,500
\$12,500 - \$24,999
\$25,000 - \$37,499
\$37,500 - \$49,999

5. What is your favorite type of food?

Italian

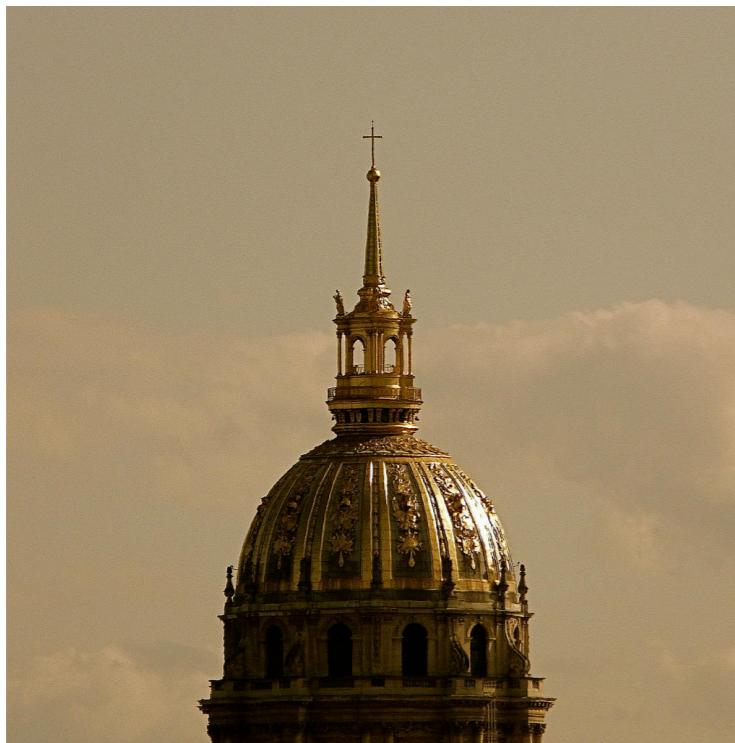
Image tagging

Provide 3 tags for this image.

Instructions:

- You must provide 3 tags for this image.
- Each tag must be a single word
- No tag can be longer than 25 characters
- The tags must describe the image, the contents of the image, or some relevant context.

Image:



Tag 1:

Tag 2:

Tag 3:

Transcription

Transcribe the text contained in the image

- Look at the receipt and copy the number of items purchased.
- Provide the dollar amount for tax.
- Provide the dollar amount for the total sale.
- Do not use dollar signs (\$) but make sure you have two decimal points. (ie 4.35)

Image:



Number of items on receipt

Tax amount

Total spent on all items

Writing

Write a brief description of a website.

- Write short article summarizing what a website is about and their products and services.
- Click the link below to review the website and browse the products and services.
- Your submission must be at least 50 words long but no more than 100 words.
- No award will be given for submissions of less than 50 words.
- Your writing must be original and can not simply be a copy of part of the website.

Website name: The Website Name Here

Website link: <http://www.linktowebsitewhere.com>

Any HTML form

1. This is an example of a multiple choice question, select only ONE of the following:

- Choice 1
- Choice 2

2. This is an example of a free text answer, use this when you only need workers to enter a line of text:

3. This is another example of a multiple choice question, select only ONE of the following. A drop down list saves more screen real estate than radio buttons (see Question 1) but may require more keystrokes for the Worker to select the answer they want.

4. This is an example of a question where you can select one or more of these options:

- selection 1
- selection 2
- selection 3
- selection 4

5. This is another example of a question that can have multiple answers (select all that apply):

- Option 1
- Option 2

[Home](#)[Create](#)[Manage](#)[Developer](#)[Help](#)[Overview](#) [Tour](#) [Case Studies](#) [Business Solutions](#)[We're Hiring! Learn More](#)

Sentiment Rating Simplified

The Mechanical Turk **Sentiment App** makes it simple to collect and understand sentiment on your data!

- ✓ Quick and easy HIT design
- ✓ Tailor your sample size from 1 to 20
- ✓ Start receiving results in minutes

[Create a Sentiment Project](#)

Work Distribution Made Easy

Mechanical Turk gives businesses and developers access to an on-demand, scalable workforce



Categorization on Mechanical Turk

The Mechanical Turk **Categorization App** makes it simple to get fast, accurate results on your Categorization Project!



Sentiment Rating Simplified

The Mechanical Turk **Sentiment App** makes it simple to collect and understand sentiment on your data!



Your business can use Mechanical Turk to:

[Home](#)[Create](#)[Manage](#)[Developer](#)[Help](#)[New Project](#)[New Batch with an Existing Project](#)[Create HITs individually](#)

Your New Project

[How does it work?](#)

Sentiment Project

[Create Project](#)[Provide Instructions](#)[Upload Data File](#)[Preview](#)[Checkout](#)

Project

What would you like to call your project?

Obama sentiment 9/13

Sentiment Question

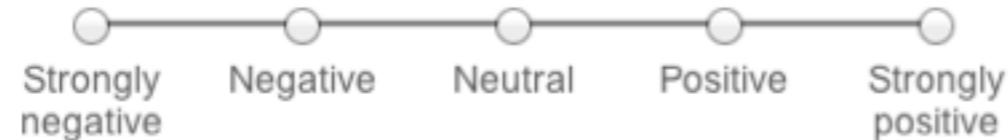
What do you want Workers to judge attitude towards?
(see examples)

President Obama

How the Workers will see your question:

Judge the sentiment expressed by the following item towards: President Obama

{Placeholder for your items to be judged. You will be asked to upload them at a later step}

[Next »](#)

Sentiment Project[Create Project](#)**Provide Instructions**[Upload Data File](#)[Preview](#)[Checkout](#)**Instructions for Workers**

For each point on the scale, define the criteria Workers should follow when selecting it. We have provided sample instructions that you can customize below.

Strongly positive:

Select this if the item embodies emotion that was extremely happy or excited toward the topic. For example, "Their customer service is the best that I've seen!!!!"

Positive:

Select this if the item embodies emotion that was generally happy or satisfied, but the emotion wasn't extreme. For example, "Sure I'll shop there again."

Neutral:

Select this if the item does not embody much of positive or negative emotion toward the topic. For example, "Yeah, I guess it's ok." or "Is their customer service open 24x7?"

Negative:

Select this if the item embodies emotion that is perceived to be angry or upsetting toward the topic, but not to the extreme. For example, "I don't know if I'll shop there again because I don't trust them."

Strongly negative:

Select this if the item embodies negative emotion toward the topic that can be perceived as extreme. For example, "These guys are teriffic... NOTTTT!!!!!!" or "I will NEVER shop there again!!!"

Number of Responses

How many Workers do you want to rate sentiment for each item? [\(Help me choose\)](#)

1
2
3
4
5
6
7
8
9



REQUESTER

[Home](#)[Create](#)[Manage](#)[Developer](#)[Help](#)[New Project](#)[New Batch with an Existing Project](#)[Create HITs individually](#)

Obama sentiment 9/13 — Judge the sentiment expressed by the following item toward: President Obama

[How does it work?](#)**Sentiment Project**[Create Project](#)[Provide Instructions](#)**Upload Data File**[Preview](#)[Checkout](#)**Upload Data File**Please provide a .csv file of your data so that we can create your work items. ([learn more](#))Don't have one? You can [view a sample](#).[Choose File](#)

tweets.csv

[Upload](#)

After your file is uploaded, you can specify which columns to show Workers.

[Next »](#)[Leave feedback for this page.](#)[Help](#) | [Contact Us](#) | [Policies](#) | [Press Inquiries](#) | [Blog](#) | [Careers](#)[MTurk.com](#) | [Requesters](#) | [Workers](#) | [Developers](#) [Follow Us on Twitter](#) [Become a Fan on Facebook](#)

©2005-2013 Amazon.com, Inc. or its Affiliates. All rights reserved.

An company

[New Project](#)[New Batch with an Existing Project](#)[Create HITs individually](#)

Obama sentiment 9/13 2 — Judge the sentiment expressed by the following item toward: President Obama

[How does it work?](#)

Sentiment Project

[Create Project](#)[Provide Instructions](#)

Upload Data File

[Preview](#)[Checkout](#)

Upload Data File - Choose Fields

Each column from your .csv file is shown as a field below. Tell us which fields you want to show Workers and whether the fields contain text, a link to an image, or a link to a website.

Optionally, you can add a label to appear next to your data. For instance, if the field is a "tweet," you might want to add a label of "Tweet."

Field	Show Workers?	Type of Data	Label
Tweet	<input checked="" type="checkbox"/>	Text	Tweet
user	<input type="checkbox"/>	Text	username

If this doesn't look right, you can [upload a different file](#).

[Next »](#)[Leave feedback for this page.](#)

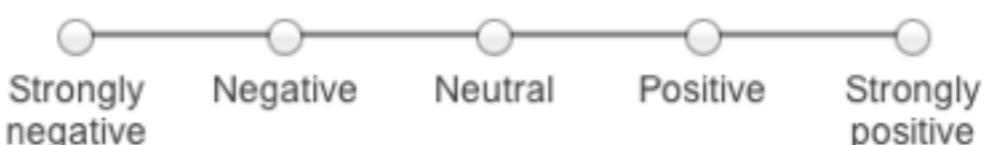
Sentiment Project[Create Project](#)[Provide Instructions](#)[Upload Data File](#)**Preview**[Checkout](#)**Preview of Work Items**

This is what Workers will see.

Instructions↓

Pick the best sentiment based on the following criterion.

Strongly positive	Select this if the item embodies emotion that was extremely happy or excited toward the topic. For example, "Their customer service is the best that I've seen!!!!"
Positive	Select this if the item embodies emotion that was generally happy or satisfied, but the emotion wasn't extreme. For example, "Sure I'll shop there again."
Neutral	Select this if the item does not embody much of positive or negative

Judge the sentiment expressed by the following item toward: President Obama**Tweet:** Obama Gives Money To Syrian Rebels Who Behead Innocent Christians //<http://t.co/wSKkzUFXjY> <http://t.co/pQiUfwOXpm>

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

Preview 1 of the first 5 items

[« Go back and edit](#)[Continue »](#)[Leave feedback for this page.](#)

[New Project](#)[New Batch with an Existing Project](#)[Create HITs individually](#)

Obama sentiment 9/13 — Judge the sentiment expressed by the following item toward: President Obama

[How does it work?](#)**Sentiment Project**[Create Project](#)[Provide Instructions](#)[Upload Data File](#)[Preview](#)**Checkout****Checkout**

Number of Items	1,000
Number of Workers per Item	x 5
Number of Worker Submissions	= 5,000
Reward per Submission <small>(details)</small>	x \$0.020 ▲
Total Worker Rewards	= \$100.000
Total Mechanical Turk fees <small>(details)</small>	+ \$45.000
Total cost	\$145.000

REQUESTER
beta[Home](#)[Create](#)[Manage](#)[Developer](#)[Help](#)[New Project](#)[New Batch with an Existing Project](#)[Create HITs individually](#)

Sentiment Project

[Create Project](#)[Provide Instructions](#)[Upload Data File](#)[Preview](#)[Checkout](#)

Congratulations

Thank you for publishing your project with Mechanical Turk. Your items are being published for Workers to complete. This may take a few minutes, depending on the size of your project. You will receive an email when Workers have completed your project.

You can check on your project status by clicking [here](#).

[Leave feedback for this page.](#)[Help](#) | [Contact Us](#) | [Policies](#) | [Press Inquiries](#) | [Blog](#) | [Careers](#)[MTurk.com](#) | [Requesters](#) | [Workers](#) | [Developers](#) [Follow Us on Twitter](#) [Become a Fan on Facebook](#)

©2005-2013 Amazon.com, Inc. or its Affiliates. All rights reserved.

An company

Items Completed 40 of 1000

Negative 2%

Neutral 2%

Positive 0%

Incomplete 96%

[Download Results](#)[Add Time](#)[Cancel](#)**Sentiment Project****Answer Summary**[Results](#)[Cost](#)**Details**

Project Obama sentiment 9/13

Status in Progress

Question

Judge the sentiment expressed by the following item toward: President Obama

Created 09/08/13 13:10

Time elapsed 26 minutes

Est. completion 09/08/13 14:35

Expiration time 09/12/13 13:12

Worker time limit 1 hour

Effective hourly rate \$4.000

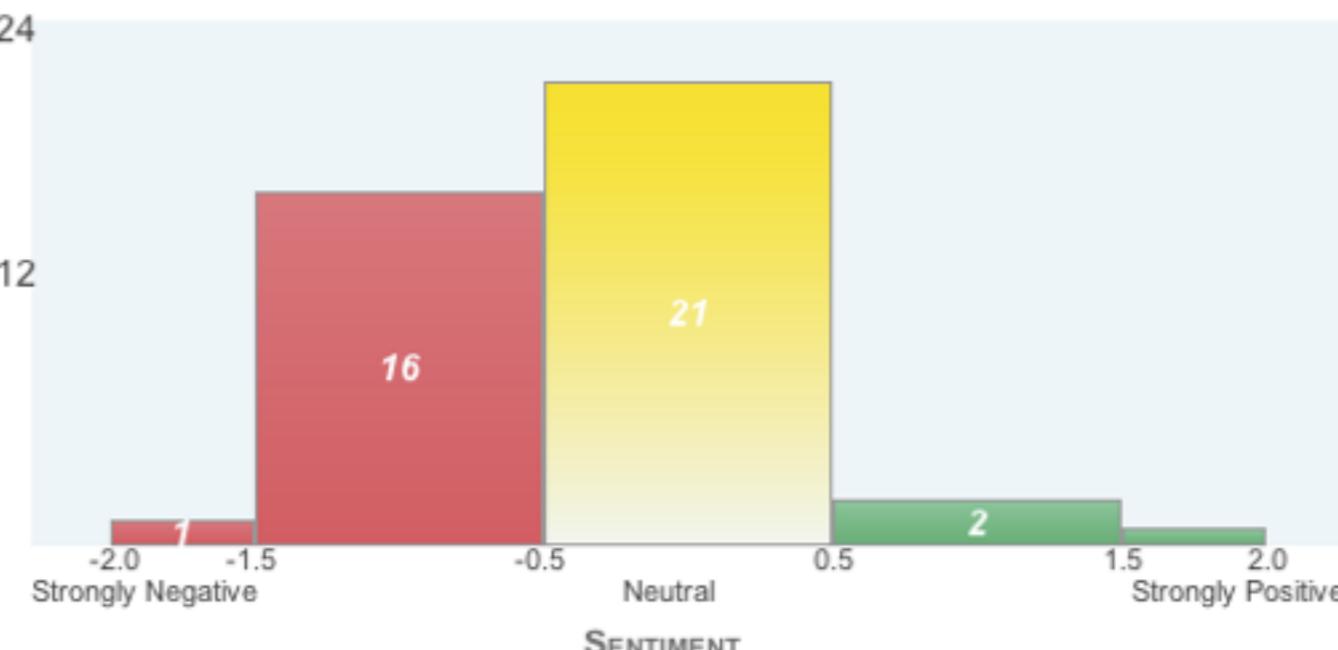
Number of Workers 5

Input file tweets.csv

40 of your 1000 items are done. We are waiting for Workers to complete the remaining 960 items.

The chart below shows you the distribution of the mean sentiment ratings ([How is it calculated?](#)) for the 1000 items in this batch. Click on each bar to view those items on the Results page.

You can also download the results in a .csv file.

**Summary for the batch**

This batch has a mean sentiment rating of **-0.41**. Of the 40 items, 2 items have a positive rating and 17 items have a negative rating.

[Results](#) [Workers](#) [Qualification Types](#)

Items Completed 46 of 1000

Negative 2%

Neutral 2%

Positive 0%

Incomplete 95%

Download Results

Add Time

Cancel

Sentiment Project

Answer Summary

Results

Cost

Details

Project Obama sentiment 9/13

Status in Progress

QuestionJudge the sentiment expressed by
the following item toward: President
Obama

Created 09/08/13 13:10

Time elapsed 28 minutes

Est. completion 09/08/13 14:35

Expiration time 09/12/13 13:12

Worker time limit 1 hour

Effective hourly rate \$4.000

Number of Workers 5

Input file tweets.csv

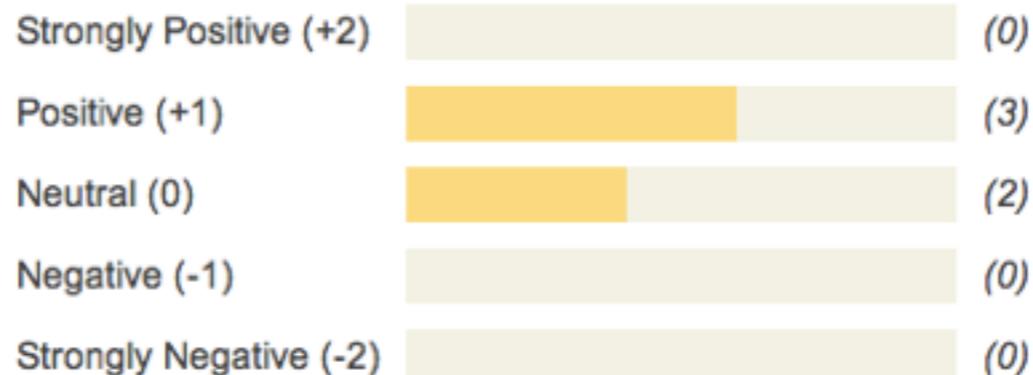
Results download results

Filter: Positive

Instructions→

Judge the sentiment expressed by the following item toward: President Obama

Tweet: @MeatSauce1 Obama and the NFL has saved Detroit!! #ford #HardcorePawn

Average Sentiment rating of **0.6** based on **5 responses**

Items Completed 55 of 1000


 Negative 2% Neutral 3% Positive 0% Incomplete 94%
[Download Results](#)[Add Time](#)[Cancel](#)**Sentiment Project**

Answer Summary

Results

Cost

Details

Project Obama sentiment 9/13

Status in Progress

Question

Judge the sentiment expressed by
the following item toward: President
Obama

Created 09/08/13 13:10

Time elapsed 28 minutes

Est. completion 09/08/13 14:34

Expiration time 09/12/13 13:12

Worker time limit 1 hour

Effective hourly rate \$4.000

Number of Workers 5

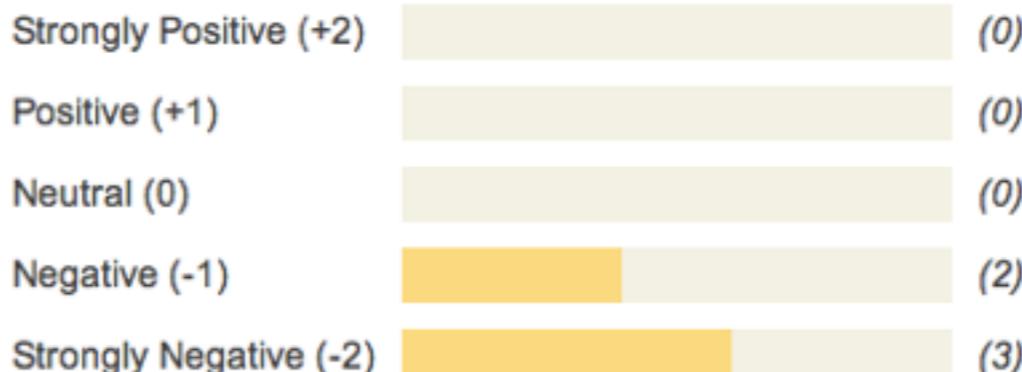
Input file tweets.csv

Results download results

Filter: Strongly Negative

Instructions→

Judge the sentiment expressed by the following item toward: President Obama

Tweet: RT @cher: I Believe,If Pres.Obama,Bombs Syria,it might b His Downfall.He could
go From Loved 2 Loathed.U Cant Go Against The Will Of 80% OfAverage Sentiment rating of **-1.6** based on **5 responses**

Building your own HIT

- Set the parameters of your HIT
- Optionally, specify requirements for which Turkers can complete your HIT
- Design an HTML template with \${variables}
- Upload a CSV file to populate the variables
- Pre-pay Amazon for the work
- Approve/reject work from Turkers
- Analyze results

HIT Parameters

- title
- description
- keywords
- reward amount
- max time allotted for work
- auto approval time

Project Name: **Reading Comprehension Test**

This name is not displayed to Workers.

Describe your HIT to Workers**Title****Write questions and answers for a reading comprehension test**

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**Read an article and then write 5 questions and 5 example answers that test a reader's comprehension**

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

Keywords**reading, comprehension, writing, questions, answers, research**

Provide keywords that will help Workers search for your HITs.

 This project may contain potentially explicit or offensive content, for example, nudity. (See details)**Setting up your HIT****Reward per assignment****\$ 0.25**

Tip: Consider how long it will take a Worker to complete each task. A 30 second task that pays \$0.05 is a \$6.00 hourly wage.

Number of assignments per HIT**10**

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

Time allotted per assignment**2****Hours** ▾

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

HIT expires in**14****Days** ▾

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

Results are automatically approved in**7****Days** ▾

After this time, all unreviewed work is approved and Workers are paid.

Purpose of redundancy

- MTurk lets you set the number of assignments per HIT
- That gives you different (redundant) answers from different Turkers
- This lets you conduct surveys (num assignments = num respondents)
- Also, lets you take votes and do tie-breaking, or do quality control

amazon mechanical turk Artificial Artificial Intelligence

Your Account HITs Qualifications

Chris Callison-Burch |

122,292 HITs
available now

All HITs | HITs Available To You | HITs Assigned To You

Find HITs containing reading comprehension that pay at least \$ 0.00

for which
 require M:

HITs containing 'reading comprehension'

1-2 of 2 Results

Sort by: Time Allotted (most first)

Show all details | Hide all details

Write questions and answers for a reading comprehension test

[View a HIT in this group](#)

Requester: [Chris Callison-Burch](#) **HIT Expiration Date:** Sep 22, 2013 (1 week 6 days) **Reward:** \$0.25
Time Allotted: 2 hours **HITs Available:** 14

Description: Read an article and then write 5 questions and 5 example answers that test a reader's comprehension

Keywords: [reading](#), [comprehension](#), [writing](#), [questions](#), [answers](#), [research](#)

Qualifications Required: **Your Value**
Total approved HITs is not less than 50 1082 You meet this qualification requirement
HIT approval rate (%) is not less than 80 98 You meet this qualification requirement
Location is US US You meet this qualification requirement [Contact the Requester of this HIT](#)

Reading Across Genre Project

Not Qualified to work on this HIT ([Why?](#)) | [View a HIT in this group](#)

Requester: [Shi Feng](#) **HIT Expiration Date:** Sep 11, 2013 (2 days 12 hours) **Reward:** \$5.00
Time Allotted: 1 hour 30 minutes **HITs Available:** 3

Description: Subjects are to read some texts and answer some questions

Keywords: [reading](#), [texts](#), [comprehension](#)

Qualifications Required: **Your Value**

Masters has been granted none [Request Qualification](#)

Worker Requirements

Advanced

For the best quality, **Master Workers** are currently selected to complete your work. ([What is a Master Worker?](#))

[Worker requirements «](#)

Worker requirements:

Require that Workers be Masters to do your HITs



Only Workers who qualify to do my HITs can preview my HITs.

Yes No

Worker Requirements

Advanced

Worker requirements:

[Customize Worker Requirements...](#)

 <https://requester.mturk.com>

You have chosen NOT to use Master Workers. We strongly encourage you to use Masters as these Workers have demonstrated accuracy in performing a wide range of HITs. Are you sure you want to continue?

[Cancel](#) [OK](#)

Specify ALL the qualifications Workers must meet to work on your HITs:

Location is UNITED STATES [remove](#)

HIT Approval Rate (%) for all Requesters' HITs greater than or equal to 80 [remove](#)

Number of HITs Approved greater than or equal to 50 [remove](#)

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

Only Workers who qualify to do my HITs can preview my HITs.

Yes No

Masters

Masters are elite groups of Workers who have demonstrated accuracy on specific types of HITs. Workers achieve a Masters distinction by consistently completing HITs with a high degree of accuracy across a variety of Requesters. Masters must continue to pass our statistical monitoring to remain Mechanical Turk Masters. Because Masters have demonstrated accuracy, they can command a higher reward for their HITs. You should expect to pay Masters a higher reward.

Masters

- Amazon now nominates a subset (21k workers, estimated at 10% of all Turkers) of senior / good workers as “Masters”
- Amazon charges 30% commission for Masters versus their normal 10% rate
- They have now implemented this as the default qualification for new Requesters
- Why?

Masters: Pros

- People who use the Web UI are often newcomers who do not know to implement quality control.
- Masters will not touch badly designed and ambiguous tasks.
- Masters will not touch tasks paying less than minimum wage.

Masters: Cons

- There are many fewer Masters workers.
 - There is now a significant lag in the task being picked by workers.
 - The tasks now take much longer to complete.
- There is an increased cost because Masters demand decent wages.
- It is not clear in what tasks the Masters are tested and how a new worker can become a master.

Custom Qualifications

- In addition to the built in qualifications (masters, location, approval rate, min HITs completed), you can also create and manage your own qualifications
- These can be managed through the web interface or the API

Custom Qualifications

[Home](#)[Create](#)[Manage](#)[Developer](#)[Help](#)[Results](#)[Workers](#)[Qualification Types](#)[We're Hiring! Learn More](#)

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
Trusted research...	2GJ7Q67051QKTXPQMYHC7XMG14AEYR	7	Thu Jun 20 17:35:18 UTC 2013	This qualification is granted to grad students and researchers. We use it to limit the participation in our pilot runs of experiments.
Temporal Master	2YG46UXFCD45EMCI7IJNZKIN3WJVTB	0	Fri Mar 23 09:46:47 UTC 2012	This qualification is granted to Turkers who have demonstrated a high level of competency in temporal relations.
Presidential Su...	279YQ4J3NAB6SPK4ZTPYOT1TJI6QDJ	0	Sat Nov 03 18:33:24 UTC 2012	This qualification was given to people who responded to the political survey before the election. We'll allow them to answer some follow up questions after the election.
Monolingual	2K5E8061IERNF05E1EKLI0RSW6NIKHE1D	6	Tue Jul 02 16:11:47	Granted to Workers who have demonstrated good skills at doing the

Qualification Tests

- The API also allows you to set up qualification tests that Workers must pass before doing your tasks
- What effects do you think qualification tests have?

HTML Template

① Enter Properties

② Design Layout

③ Preview and Finish

Project Name: **Reading Comprehension Test**

This name is not displayed to Workers.

Frame Height

500

Height in pixels of the frame your HIT will be displayed in to Workers. Adjust the height appropriately to minimize scrolling for Workers.

Format

Heading 1

Font

U

I

B

T₁₂₃

←

→

≡

≡

≡

≡

≡

≡

≡

≡

≡

≡

≡

≡

Edit HTML Source

Reading comprehension test

Please do the following:

- Read the short newspaper article.
- Write reading comprehension questions about it
- Give sample answers for each of your questions.

Good reading comprehension questions:

- Ask about why something happened or why someone did something.
- Ask about relationships between people or things.
- Should be answerable in a few words.

Poor reading comprehension questions:

- Ask about numbers or dates.
- Only require a yes/no answer.

[\\${url}](#)

HTML Template

Format Font **Edit HTML Source**

```
<h1>Reading comprehension test</h1>
<p>Please do the following:</p>
<ul>
    <li>Read the short newspaper article.</li>
    <li>Write reading comprehension questions about it</li>
    <li>Give sample answers for each of your questions.</li>
</ul>
<p>Good reading comprehension questions:</p>
<ul>
    <li>Ask about why something happened or why someone did something.</li>
    <li>Ask about relationships between people or things.</li>
    <li>Should be answerable in a few words.</li>
</ul>
<p>Poor reading comprehension questions:</p>
<ul>
    <li>Ask about numbers or dates.</li>
    <li>Only require a yes/no answer.</li>
</ul>
<p><a target="_blank" href="${url}">${url}</a></p>
<p><iframe width="80%" height="400" id="pageFrame" src="${url}"></iframe></p>
<p><b>Question 1:</b> <input type="text" size="80" id="Q1" name="Q1" /><br />
<b>Answer 1:</b> <input type="text" size="20" id="A1" name="A1" /></p>
<p><b>Question 2:</b> <input type="text" size="80" id="Q2" name="Q2" /><br />
<b>Answer 2:</b> <input type="text" size="20" id="A2" name="A2" /></p>
<p><b>Question 3:</b> <input type="text" size="80" id="Q3" name="Q3" /><br />
<b>Answer 3:</b> <input type="text" size="20" id="A3" name="A3" /></p>
<p><b>Question 4:</b> <input type="text" size="80" id="Q4" name="Q4" /><br />
<b>Answer 4:</b> <input type="text" size="20" id="A4" name="A4" /></p>
<p><b>Question 5:</b> <input type="text" size="80" id="Q5" name="Q5" /><br />
<b>Answer 5:</b> <input type="text" size="20" id="A5" name="A5" /></p>
<p>Please provide any comments that you have about this HIT. We appreciate your input!</p>
<p><textarea rows="3" cols="80" name="comment"></textarea></p>
```

Save

Preview

Writing Instructions

- Be sure to be as specific as possible in your instructions so that there's no confusion. *For example, when asking workers to extract text from an image, ask workers to type the text exactly as shown in the image including capitalizations, spaces and punctuation.*
- Include an example of a right answer, and a wrong answer
- Clarify what you expect if the HIT is not doable because of missing data or other problems

How to populate a HIT with data

- Using the web interface, you communicate data with MTurk via Comma Separated Value (CSV) files
- Each column header gives a \${variable} name
- Each row gives one screen's worth of data
- You may have more columns than are used in the HIT - useful for bookkeeping

Special characters in the CSV spec

- There are several special characters that are reserved in the CSV format
- I recommend using Python CSV module, or exporting CSV from Excel
- HTML reserve characters also need special treatment since the \${variables} get inserted as strings into your template
- Also be careful if you're using javascript

[Home](#)[Create](#)[Manage](#)[Developer](#)[Help](#)

New Project

[New Batch with an Existing Project](#)[Create HITs individually](#)

Your project was successfully saved.

Start a New Batch with an Existing Project

Project Name	Title	Creation Date ▼	New Batch	Edit	Copy	Delete
Reading Comprehension Test	Write questions and answers for a reading comprehension test	September 8, 2013	New Batch	Edit	Copy	Delete
Obama sentiment 9/13	Judge the sentiment expressed toward: President Obama	September 8, 2013	New Batch	Edit	Copy	Delete
Compression HIT - grammar/meaning	Judge the quality of similar sentences	July 3, 2013	New Batch	Edit	Copy	Delete
Grade Word Alignment	Grade Word Alignment	July 1, 2013	New Batch	Edit	Copy	Delete
Word Alignment - open to All Turkers	Word Alignment	June 21, 2013	New Batch	Edit	Copy	Delete
Word Alignment - Trusted Workers	Word Alignment	June 20, 2013	New Batch	Edit	Copy	Delete
Report Linking	Link events across two news stories	May 24, 2013	New Batch	Edit	Copy	Delete

[Home](#)[Create](#)[Manage](#)[Developer](#)[Help](#)

New Project

[New Batch with an Existing Project](#)

Create HITs individually

Your project was successfully saved.

Start a New Batch with an Existing Project

New Batch



Project Name

Choose a .csv file with the variables you specified in your project.

[Choose File](#)

urls.csv

[Upload](#)

Reading Comprehension

Obama sentiment 9/13

Judge the sentiment
expressed toward:
President Obama

September 8, 2013

[New Batch](#)[Edit](#)[Copy](#)[Delete](#)Compression HIT -
grammar/meaningJudge the quality of similar
sentences

July 3, 2013

[New Batch](#)[Edit](#)[Copy](#)[Delete](#)

Grade Word Alignment

Grade Word Alignment

July 1, 2013

[New Batch](#)[Edit](#)[Copy](#)[Delete](#)Word Alignment - open to All
Turkers

Word Alignment

June 21, 2013

[New Batch](#)[Edit](#)[Copy](#)[Delete](#)Word Alignment - Trusted
Workers

Word Alignment

June 20, 2013

[New Batch](#)[Edit](#)[Copy](#)[Delete](#)

Report Linking

Link events across two
news stories

May 24, 2013

[New Batch](#)[Edit](#)[Copy](#)[Delete](#)

Reading comprehension test

Please do the following:

- Read the short newspaper article or blog post linked below (if it doesn't load in the frame then open the link in another window)
- Write reading comprehension questions about it
- Give sample answers for each of your questions.

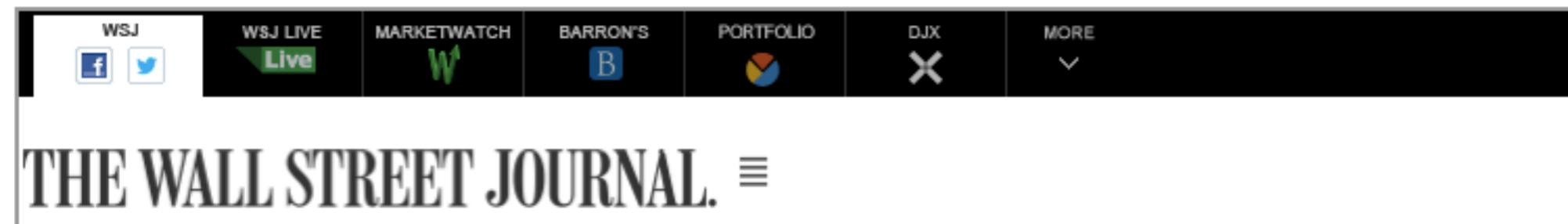
Good reading comprehension questions:

- Ask about why something happened or why someone did something.
- Ask about relationships between people or things.
- Should be answerable in a few words.

Poor reading comprehension questions:

- Ask about numbers or dates.
- Only require a yes/no answer.

<http://blogs.wsj.com/digits/2012/11/04/how-the-journal-tested-googles-search-results/>



[Previous HIT](#)

Showing HIT 3 of 14

[Next HIT](#)

[Cancel](#)

[Next](#)

Reading Comprehension Test

Batch Summary

Batch Name: Reading Comprehension Test 4

Description: Read an article and then write 5 questions and

Batch Properties

Title: Write questions and answers for a reading comprehension test

Description: Read an article and then write 5 questions and 5 example answers that test a reader's comprehension

Batch expires in: 14 Days

Results are automatically approved after: 7 Days

Workers must meet the following Qualifications to work on these HITs:
Location score is UNITED STATES

HIT Approval Rate (%) for all Requesters' HITs score greater than or equal to 80

Number of HITs Approved score greater than or equal to 50

HITs

Number of HITs in this batch: 14

Number of assignments per HIT: x 10

Total number of assignments in this batch: 140

Cost

Reward per Assignment: \$0.250

x 140 (*total number of assignments in this batch*)

Estimated Total Reward: \$35.000

Estimated Fees to Mechanical Turk: + \$3.500 (*fees paid to Mechanical Turk*) ([fee details](#))

Estimated Total Cost: \$38.500 (*this is the amount that will be deducted from your Available Balance when you click "Publish HITs"*)

Reading Comprehension Test 4

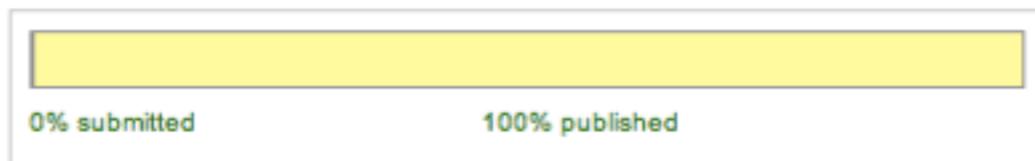
View the latest status of this batch, make changes, or get results.

Read an article and then write 5 questions and 5 example answers that test a reader's comprehension

Status

[Cancel this batch](#)

Status: In Progress



Assignments Completed: 0 / 140

Average Time per Assignment: Not yet available

Average Hourly Rate: Not Yet Available

Creation Time:

September 08, 2013 6:18 PM
PDT

Estimated Completion Time:

Not yet available

Settings

Reading Comprehension Test

Description: Read an article and then write 5 questions and 5 example answers that test a reader's comprehension

Keywords: reading, comprehension, writing, questions, answers, research

Qualification Requirement: Location is UNITED STATES (Required for preview)
HIT Approval Rate (%) for all Requesters' HITs greater than or equal to 80 (Required for preview)
Number of HITs Approved greater than or equal to 50 (Required for preview)

Number of Assignments per HIT: 10

Reward per Assignment: \$0.250

Input File: urls.csv

HIT expires on: Not yet available

Assignment duration: 2 hours

Auto Approval Delay: 7 days

Results

[Results](#)

Assignments pending review: 0

Assignments approved: 0

Assignments rejected: 0

Cost Summary

Estimated Total Reward: \$35.000

Estimated Fees to Mechanical Turk: \$3.500 ([fee details](#))

Estimated Total Cost: \$38.500

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

[Home](#)[Create](#)[Manage](#)[Developer](#)[Help](#)[Results](#)[Workers](#)[Qualification Types](#)[Manage HITs individually](#)

Manage Batches

Click on the name of the batch to see more details

▼ Batches in progress (2)

Reading Comprehension Test 4

[Results](#)[Cancel this batch](#)

Created:	September 08, 2013	Assignments Completed:	16 / 140
Time Elapsed:	about 11 hours	Estimated Completion Time:	September 12, 2013 10:43 PM PDT (Thursday)
Average Time per Assignment:	9 minutes 20 seconds	Effective Hourly Rate:	\$1.607
Batch Progress:	<div><div style="width: 11%; background-color: #00ff00; float: left;"></div><div style="width: 88%; background-color: #ffff00; float: left;"></div></div> 11% submitted 100% published		

Compression HIT - grammar/meaning 10

[Results](#)[Cancel this batch](#)

Created:	July 04, 2013	Assignments Completed:	2,468 / 2,468
Time Elapsed:	1 day	Estimated Completion Time:	COMPLETE
Average Time per Assignment:	3 minutes 40 seconds	Effective Hourly Rate:	\$4.091
Batch Progress:	<div><div style="width: 100%; background-color: #00ff00; float: left;"></div></div> 100% submitted 100% published		

Reading Comprehension Test 4

[Customize View](#)
[Filter Results](#)
[Upload CSV](#)
[Download CSV](#)

16 of 16 assignments (FILTER APPLIED: only show assignments that are in 'Submitted' status)

Input.Url	Q1	A1	Q2	A2	Q3	A3	Q4
http://www.bloodletters.com/2011/03/turkish-del...	What does the writer dislike about the previous...	The cover doesn't suit the book's contents and ...	According to the writer, why is proofreading important...	Encountering errors jerks the reader out of the...	Why did the writer decide not to hire a professional...	A professional proofreader would have cost hundred...	What problems could the writer face with the Mechanical...
http://www.bloodletters.com/2011/03/turkish-del...	How does the author feel about the cover?	Not happy.	At what pace are the realities of self-publishing...	At a fast pace.	How many times does the author typically read a...	Three times.	Why would be an imposition to ask your friend...
http://mrshri.sharedby.co/share/2bpKT2	What will improve about smartphones in the near...	predictive texting.	How well do voice and text commands work if the...	Not well.	What do twitches and blinks create?	Unintended words.	For whom this technological advance beneficial...
http://mrshri.sharedby.co/share/2bpKT2	What do software packages in today's phones often do?	texts and voice commands	What does Augmented and Alternative Communication...	They help disabled people to communicate	Keith Vertanen was from which institution?	Montana Tech	Where did Vertanen and Kristensson seek help from?
http://mechanicalturk.typepad.com/blog/2012/07/...	Why is it important to establish a style guide?	To create consistency in the quality of	What two types of qualification tests should	Writing and editing.	How much should you aim to pay for basic writing	At least minimum wage.	How can we make tasks easier for workers to...

Approving / rejecting

- The MTurk web interface gives you a way of manually reviewing the Turkers results
- After reviewing a Worker's submission, you can
 - approve it (worker gets paid)
 - reject it (worker does not get paid)
 - block the worker (worker is banned from working on any of your HITs, and may get blacklisted from MTurk)

18 of 18 assignments (FILTER APPLIED: only show assignments that are in 'Submitted' status)

	Approve	Reject	Worker ID	Lifetime Approval Rate	Input.Url	Q1	A1	Q2	A2
<input checked="" type="checkbox"/>			A2NP0GV3WST630	0% (0/0)	http://www.bloodletters.com/2011/03/turkish-del...	What does the writer dislike about the previous...	The cover doesn't suit the book's contents and ...	According to the writer, why is proofreading important...	Encountered errors jerk the reader out of the..
<input checked="" type="checkbox"/>			A3I2YF0MXB7P0B	0% (0/0)	http://www.bloodletters.com/2011/03/turkish-del...	How does the author feel about the cover?	Not happy.	At what pace are the realities of self-publishing...	At a fast pace.
<input checked="" type="checkbox"/>			A3I2YF0MXB7P0B	0% (0/0)	http://mrshri.sharedby.co/share/2bpKT2	What will improve about smartphones in the near...	predictive texting.	How well do voice and text commands work if the...	Not well.
<input checked="" type="checkbox"/>			A1LR1FHCZICP9N	0% (0/0)	http://mrshri.sharedby.co/share/2bpKT2	What do software packages in today's phones oft...	texts and voice commands	What does Augmented and Alternative communication...	They help disabled people to communicate...
<input checked="" type="checkbox"/>			A3PYW2CPW9ZZIY	0% (0/0)	http://www.convinceandconvert.com/blogging-and...	What is one strategy that BlissPR and Bailey Ga...	They are blog focused	What two sites were listed as examples of ways ...	Facebook and Slideshare
<input checked="" type="checkbox"/>			AP4OBLUBXOC6G	0% (0/0)	http://mechanicalturk.typepad.com/blog/2012/07/...	Why is it important to establish a style guide?	To create consistency in the	What two types of qualification	Writing and editing.

Semi-automating the approval process with CSV

- A lightweight version of the API is available by uploading a graded CSV
- You can download the results as a CSV file, grade it (manually or automatically), and upload your approve or reject decisions
- A quick and easy thing to do if you have a HIT with gold standard controls, and if you learn python's CSV module

Reading Comprehension Test 4

[Customize View](#) [Filter Results](#)

[Upload CSV](#) [Download CSV](#)

18 of 18 assignments (FILTER APPLIED)

[Approve](#) [Reject](#)

<input checked="" type="checkbox"/>	<u>Worker ID</u>	Lifetime Approval Rate
<input checked="" type="checkbox"/>	A2NP0GV3WST630	0% (0/0)
<input checked="" type="checkbox"/>	A3I2YF0MXB7P0B	0% (0/0)

Review Results Offline

You can review your results offline and approve or reject the assignments.

- APPROVE ASSIGNMENT:** Indicate which assignments to approve by putting an "x" under a column titled "Approve".
- REJECT ASSIGNMENT:** Indicate which assignment to reject by putting your reject feedback under a column titled "Reject".

Please select your modified CSV file for upload

[Choose File](#) no file selected

[Upload CSV](#) or [Cancel](#)

<input checked="" type="checkbox"/>	A2NP0GV3WST630	0% (0/0)	http://mrshri.sharedby.co/share/2bpKT2	What will improve about smartphones in the near...	predictive texting.	How well do voice and text commands work if the...	Not well.
<input checked="" type="checkbox"/>	A3I2YF0MXB7P0B	0% (0/0)	http://mrshri.sharedby.co/share/2bpKT2	What do software packages in today's phones oft...	texts and voice commands	What does Augmented and Alternative communication...	They help disabled people to communicate...
<input checked="" type="checkbox"/>	A3PYW2CPW9ZZIY	0% (0/0)	http://www.convinceandconvert.com/blogging-and...	What is one strategy that BlissPR and Bailey Ga...	They are blog focused	What two sites were listed as examples of ways ...	Facebook and Slideshare

Recommendation for approval thresholds

- If using controls to approve/reject assignments automatically, I recommend using 3 values
- 1) ~80% correct = approve all work from this Turker
- 2) <80% but >chance = approve assignments proportional to accuracy on the controls
- 3) <= chance, reject all assignments

The Sandbox

- Amazon provides a `sandbox' where Requesters can develop and deploy HITs at no cost
- `requestersandbox.mturk.com` mimics the behavior of `requester.mturk.com`
- When you post to the `requestersandbox` the HITs become available on `workersandbox.mturk.com`



REQUESTER

[Home](#)[Create](#)[Manage](#)[Developer](#)[Help](#)[Overview](#) [Tour](#) [Case Studies](#) [Business Solutions](#)[We're Hiring! Learn More](#)

Work Distribution Made Easy

Mechanical Turk gives businesses and developers access to an on-demand, scalable workforce

- ✓ **Flexibility:** Scale your workforce up and down quickly
- ✓ **Accuracy:** Get high-quality, cost-effective results
- ✓ **Speed:** Start receiving results in minutes

[Get Started](#)

Work Distribution Made Easy

Mechanical Turk gives businesses and developers access to an on-demand, scalable workforce



Categorization on Mechanical Turk

The Mechanical Turk **Categorization App** makes it simple to get fast, accurate results on your Categorization Project!



Sentiment Rating Simplified

The Mechanical Turk **Sentiment App** makes it simple to collect and understand sentiment on your data!



Your business can use Mechanical Turk to:

[All HITs](#) | [HITs Available To You](#) | [HITs Assigned To You](#)Find containing that pay at least \$ for which you are qualified require Master Qualification[GO!](#)

All HITs

1-10 of 2278 Results

Sort by: [Show all details](#) | [Hide all details](#)[1](#) [2](#) [3](#) [4](#) [5](#) > [Next](#) >> [Last](#)

Tag LFP 1

[Request Qualification \(Why?\)](#)| [View a HIT in this group](#)**Requester:** [LFP Requester](#)**HIT Expiration Date:** Dec 5, 2013 (12 weeks 3 days)**Reward:** \$0.00**Time Allotted:** 60 minutes**HITs Available:** 16442

DEV: Write me something awesome

[View a HIT in this group](#)**Requester:** [CrowdSource](#)**HIT Expiration Date:** Sep 9, 2014 (52 weeks)**Reward:** \$0.01**Time Allotted:** 50 minutes**HITs Available:** 15130

DEV: nate test- all fields

[View a HIT in this group](#)**Requester:** [CrowdSource](#)**HIT Expiration Date:** Sep 9, 2014 (52 weeks)**Reward:** \$0.00**Time Allotted:** 60 minutes**HITs Available:** 12945

Find the count of comments on a website

[View a HIT in this group](#)**Requester:** [SDG Development](#)**HIT Expiration Date:** Sep 12, 2013 (2 days 23 hours)**Reward:** \$0.10**Time Allotted:** 10 minutes**HITs Available:** 11277

SANDBOX: Review a Live Video Stream for 20 Seconds to Determine Quality (US)

[View a HIT in this group](#)**Requester:** [CrowdSource](#)**HIT Expiration Date:** Sep 9, 2014 (52 weeks)**Reward:** \$0.10**Time Allotted:** 20 minutes**HITs Available:** 9963

SANDBOX: Review a Live Video Stream for 20 Seconds to Determine Quality (CA)

Not Qualified to work on this HIT ([Why?](#)) | [View a HIT in this group](#)**Requester:** [CrowdSource](#)**HIT Expiration Date:** Mar 26, 2014 (28 weeks 2 days)**Reward:** \$0.10**Time Allotted:** 20 minutes**HITs Available:** 9963

SANDBOX: Review a Live Video Stream for 20 Seconds to Determine Quality (GB)

Not Qualified to work on this HIT ([Why?](#)) | [View a HIT in this group](#)

HITs can use any HTML or javascript

- Highlight word in sentence
- Word alignment HIT
- Randomize order of things
- Geolocation

The art of designing
good HITs

Crowdsourcing works for tasks that are

- Natural and easy to explain to non-experts
- Decomposable into simpler tasks that can be joined together
- Parallelizable into small, quickly completed chunks
- Well-suited to quality control (some data has correct gold standard annotations)

Crowdsourcing works for tasks that are

- Robust to some amount of noise/errors (the downstream task is training a statistical model)
- Balanced and each task contains the same amount of work
 - Don't have tons of work in one assignment but not another
 - Don't ask Turkers to annotate something occurs in the data <<10% of the time

Example HIT design:
Reading Comprehension to evaluate
Machine Translation systems

Evaluating Machine Translation

- There are many MT systems available for each language pair
- How do we know which one is best?
- There is an active field of research focused on evaluating MT automatically
- Reasoning: manual evaluation requires humans and is therefore too time consuming and expensive

Heather Locklear Arrested for driving under the influence of drugs

UPDATE Posted Sunday September 28, 2008 08:00 PM EDT



Heather Locklear

Photo by: Santa Barbara County Sheriff's Department

The actress Heather Locklear, Amanda of the popular series *Melrose Place*, was arrested this weekend in Santa Barbara (California) after driving under the influence of drugs. A witness viewed her performing inappropriate maneuvers while trying to take her car out from a parking in Montecito, as revealed to People magazine by a spokesman for the Californian Highway Police. The witness stated that around 4.30pm Ms. Locklear "hit the accelerator very violently, making excessive noise while trying to take her car out from the parking with abrupt back and forth maneuvers. While reversing, she passed several times in front of his sunglasses." Shortly after, the witness, who, in a first time, apparently had not recognized the actress, saw Ms.

- **Why was Heather Locklear arrested?**
 - ▶ She was arrested on suspicion of driving under the influence of drugs.
- **Why did the bystander call emergency services?**
 - ▶ He was concerned for Ms. Locklear's life.
- **Where did the witness see her acting abnormally?**
 - ▶ Pulling out of parking in Montecito

- Why was Heather Locklear arrested?

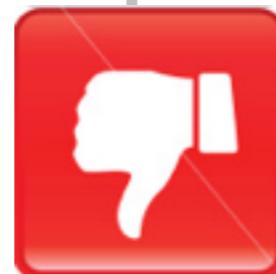
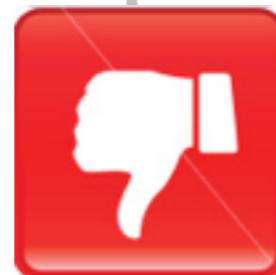
▶ She was arrested on suspicion of driving under the influence

- Why was Heather Locklear arrested in Montecito?

▶ Because she was driving under the influence of drugs

- What was Heather Locklear arrested for?

▶ Driving under the influence of drugs



- Why was Heather Locklear arrested in Montecito?

▶ Because she was driving under the influence of drugs

- Where was Locklear arrested?

▶ In a parking lot in Montecito



Was arrested actress Heather Locklear because of the driving under the effect of an unknown medicine



Heather Locklear

Photo by: Santa Barbara County Sheriff's Department

The actress Heather Locklear that is known to the Amanda through the role from the series "Melrose Place" was arrested at this weekend in Santa Barbara (Californium) because of the driving under the effect of an unknown medicine. A female witness observed she attempted in quite strange way how to go from their parking space in Montecito, speaker of the traffic police of californium told the warehouse 'People'. The female witness told in detail, that Locklear 'pressed after 16:30 clock accelerator and a lot of noise did when she attempted to move their car towards behind or forward from the parking space, and when it went backwards, she pulled itself together unites Male at their sunglasses'. A little later the female witness that did probably

- Why was Heather Locklear arrested?

Driving while medicated

- Why did the bystander call emergency services?

There was a lot of noise

- Where did the witness see her acting abnormally?

In a parking lot

. Medikamentes unknown have the effect of a fahrens under actress heather locklear arrested



Heather Locklear

Photo by: Santa Barbara County Sheriff's Department

In Santa. One is, melrose place the series of the role of the 'remember the locklear actress the heather this weekend, because of the fahrens Barbara (California) in effect unknown medikamentes arrested People 'magazine. The traffic police California, spokesman for the auszufahren montecito reported in its way from tried parklücke type strange right, you have seen as a witness. . In some Zeitung, as and when they tried to a great deal of 30 p.m., witness the detail of history locklear after 16: that durchdrückte peddle noise and its progress was made parklücke for the car or moving backwards, they had they times of their sonnenbrille '. The first was probably recognised that locklear a nearby road and anhielt, had not, with the witness to the car off

- Why was Heather Locklear arrested?
- Why did the bystander call emergency services?
- Where did the witness see her acting abnormally?

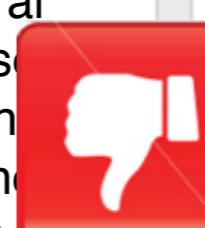
Heather Locklear Arrested for driving under the influence of drugs



Heather Locklear

Photo by: Santa Barbara County Sheriff's Department

The actress Heather Locklear, Amanda of the popular series Melrose Place, was arrested this weekend in Santa Barbara (California) after driving under the influence of drugs. A witness who viewed her performing inappropriate maneuvers while trying to take her car out from a parking in Montecito, as revealed to People magazine by a spokesman for the Californian Highway Police. The witness stated that around 4.30pm Ms. Locklear "hit the accelerator violently, making excessive noise while trying to take her car out from the parking with abrupt and forth maneuvers. While reversing, she passed several times in front of his sunglasses. Shortly after, the witness, who for a first time, apparently had not recognized the actress, saw Mr.



- Why was Heather Locklear arrested?

- ▶ She was arrested on suspicion of driving under the influence of drugs.

Driving under the influence

Driving while medicated

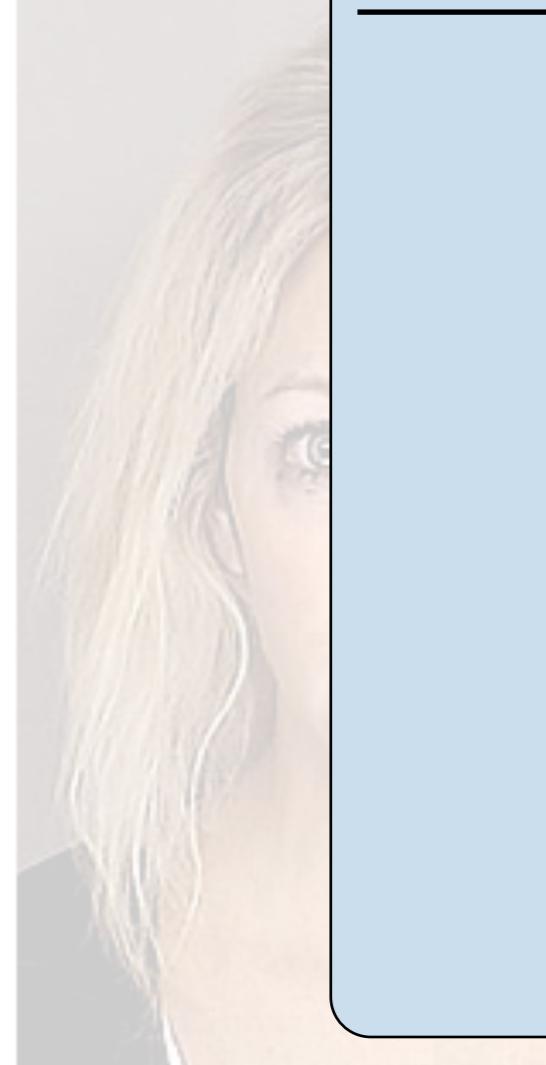
DUI

Driving while using drugs

Medikamentes

Heather
driving

System	Correct Answers
Reference	94%
Google	80%
RBMT5	77%
Geneva	63%
Tromble	50%



Heather Locklear
Photo by: Santa Barbara County
Sheriff's Department

SPONSORED LINKS

and forth maneuvers. While reversing, she passed several times in front of his sunglasses. Shortly after, the witness, who a first time, apparently had not recognized the actress, saw Ms.



Medikamentes

90 reading comprehension questions for 10 document

2160 judgments about question redundancy and goodness

5 people answering questions using 5 MT system for 10 docs

Redundantly grading all 1,300 answers



Guidelines for your own tasks

- Simple instructions are required
- If your task can't be expressed in one paragraph + bullets, then it may need to be broken into simpler sub-tasks

Guidelines for your own tasks

- Quality control is paramount
 - Measuring redundancy doesn't work if people answer incorrectly in systematic ways
 - Embed gold standard data as controls
- Qualification tests v. no qualification test
 - Reduce participation, but usually ensures higher quality

Guidelines for your own tasks

- Controls should be
 - Ones that have a unambiguous answer
 - Scorable automatically or with 2nd pass HIT
 - Not easily detectable

Guidelines for your own tasks

- Remember that MTurk is a reputation system
 - Rejecting Turkers has consequences for them, blocking Turkers is likely to get them expelled from MTurk
 - I reject all work performing near chance, accept all work doing ~80%, reject proportionally to performance on controls for in-between

Guidelines for your own tasks

- Reputation goes both ways
 - Monitor your profile on Turker Nation and TurkOpticon
 - Pay your workers generously
 - Pay quickly, and be responsive to email questions

Reminders: HW2 part 1
due on Wednesday