

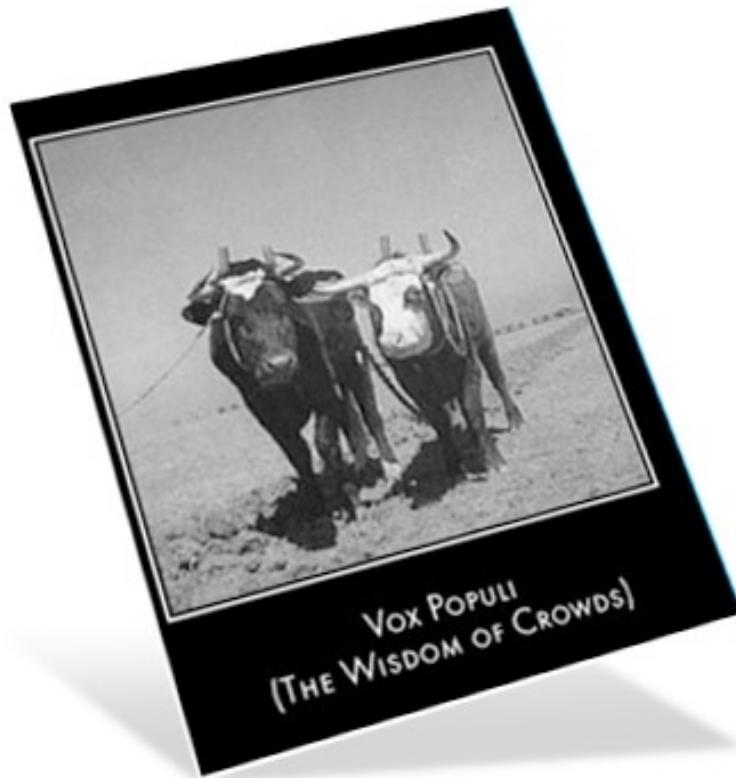
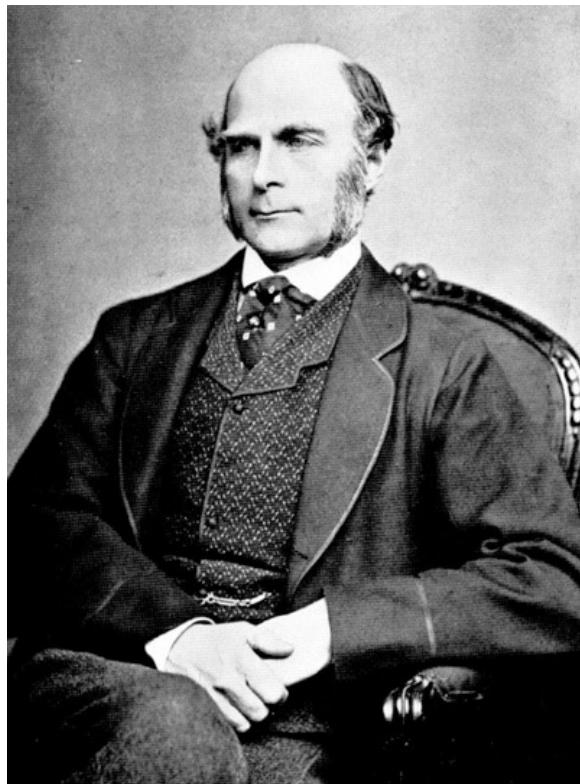
# Crowdsourcing

Practical Guidelines

Mohammad Aliannejadi

# Crowdsourcing

**Weight-judging competition:**  
**1,197 (mean of 787 crowds) vs. 1,198 pounds (actual measurement)**



# Examples

- threadless.com
- Sells t-shirts, **designed/voted** by crowds.
- Artists whose designs are chosen get paid.

The image shows the homepage of Threadless. At the top, there's a navigation bar with the Threadless logo, 'SHOP', 'PARTICIPATE', 'BLOG', a search bar, and a shopping cart icon showing 6 items. Below the navigation is a large promotional banner. The banner features a woman with brown hair flexing her biceps while wearing a green zip-up hoodie with cartoon prints of a monkey and a lizard. To the right of the image, text reads 'ALL THREADLESS HOODIES ARE \$29!' with a yellow speech bubble below it stating 'ENDS OCT 7 10AM CT' and a 'SHOP NOW' button. Below the banner, a tagline says 'Threadless is a creative community that makes, supports, and buys great art.' At the bottom, there are four icons with text: 'Shop top sellers' (with a storefront icon), 'Submit your art' (with a pencil icon), 'Pick designs' (with a checkmark icon), and 'Meet artists' (with a person icon).

Threadless is a creative community that makes,  
supports, and buys great art.

Shop top sellers   Submit your art   Pick designs   Meet artists

# Examples

- kickstarter.com
- Crowdfunding, started in 2009.
- Project creators choose a deadline and a minimum funding goal.
- Donors pledge money to support projects, in exchange of non-monetary values.



October 30, 2018 Bringing creative projects to life.	TOTAL BACKERS 15,400,267	FUNDED PROJECTS 152,685	LIVE PROJECTS 4,151
---	-----------------------------	----------------------------	------------------------

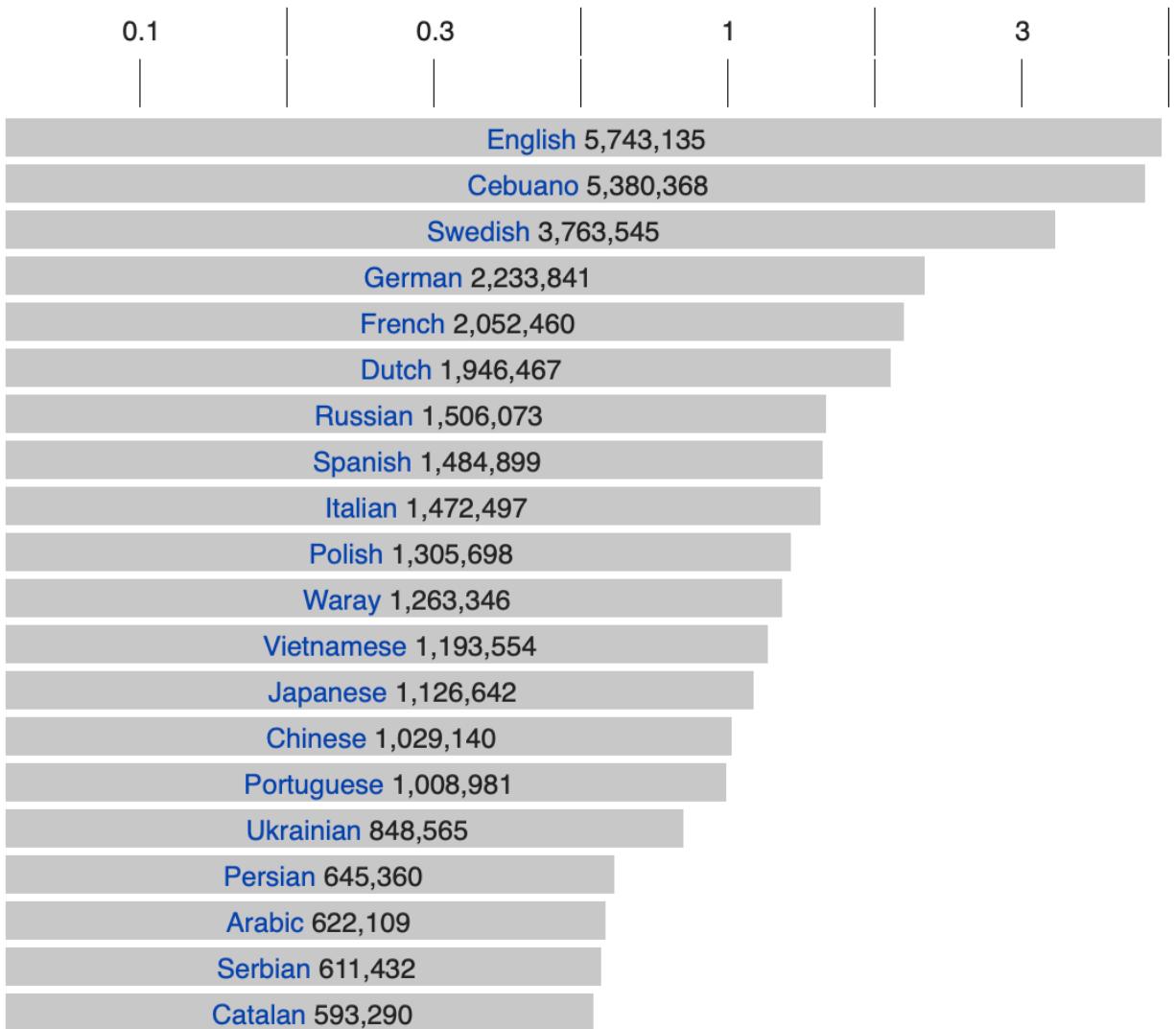
# Examples

- wikipedia.org
- Launched: January 15 2001.
- Active users: 302,461 (as of C

Logarithmic graph of the 20 largest language editions of Wikipedia

(as of 30 October 2018)<sup>[129]</sup>

(millions of articles)

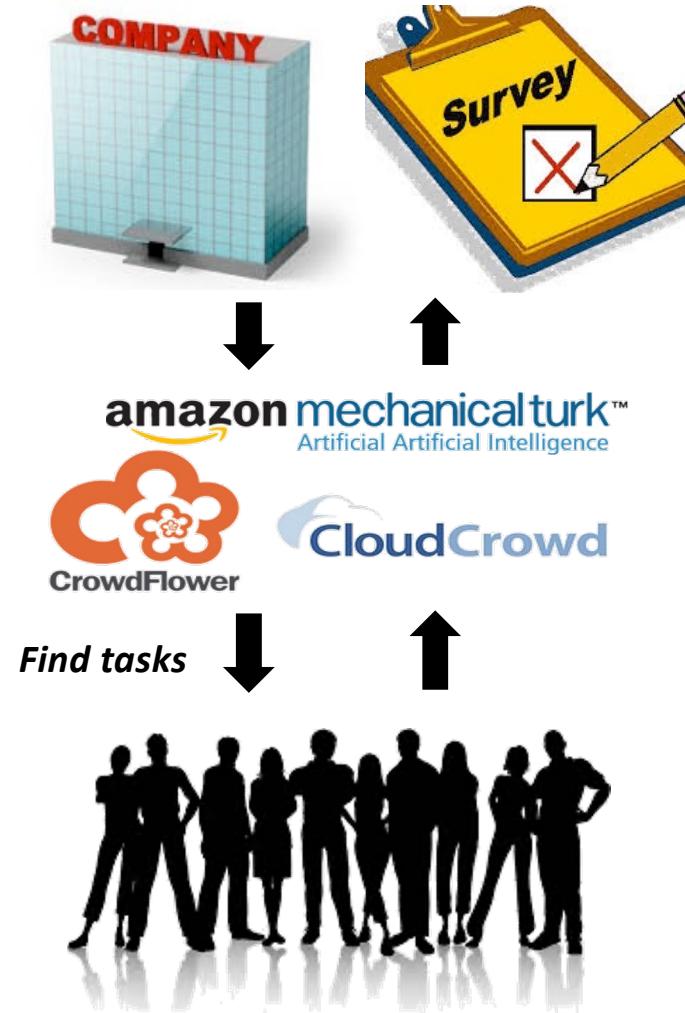


# Main Characteristics

- Outsourced to human workers.
  - Online and distributed.
  - Open call and right incentive.
  - Diversity and independence.
- 
- Task can be split to many micro-tasks.
  - Machine cannot do the task well.
  - Large crowd can probably do it well.

# Stakeholders

- Requester
- Marketplace
- Crowd



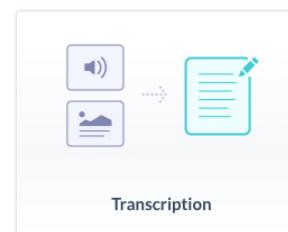
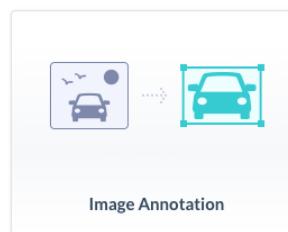
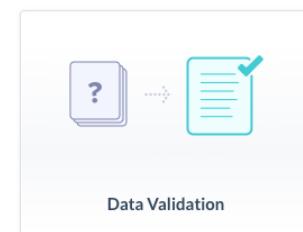
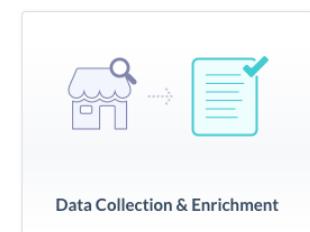
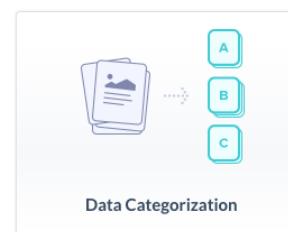
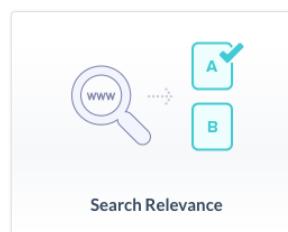
# appen

- Formerly known as CrowdFlower.
- Founded in 2007.
- Automates most of crowdsourcing process:
  - Many ready-to-use templates.
  - Usable for many applications: machine learning, image processing, information retrieval, etc.
  - Engineered to minimize cheating: test questions.
- Annual subscription for industrial use.
- No subscription fees for academic use, but your data will be made public.
- Approval handled by the company.
- Custom Markup Language (CML)

## Welcome to Figure Eight, Mohammad Aliannejadi

Get familiar with us by creating a draft of your first job

What would you like to do?



▼ See more

Need help getting started?

[Read help docs](#)

[View example case studies](#)

[Watch training videos](#)

Don't see what you're looking for?

Tell us what your data needs - we'll get back to you with a recommendation

Describe your data goals in a short summary

Drop a sample of your data file here, or [browse](#)

Send

## Use Cases > Search Relevance

To get started, select one of these customizable job templates.

	YOUR DATA HAS (INPUT)	CONTRIBUTOR DELIVERS (OUTPUT)
<b>Rate the Relevance of a Search Result</b>		
Contributors judge how relevant a resulting image and product page is to a given search query	Minimum 4 columns with suggested names: <ul style="list-style-type: none"><li>• query</li><li>• title</li><li>• image_url</li></ul>	How well does this result match the search term? <ul style="list-style-type: none"><li>• Off-topic</li><li>• Acceptable</li><li>• Good</li><li>• Excellent</li></ul>
<a href="#">Preview</a> <a href="#">Use this template</a>		
<b>Rate the Relevance of a Page of Search Results</b>		
Contributors judge how relevant a page of search results listings is to a given search query	Minimum 2 columns: <ul style="list-style-type: none"><li>• query</li><li>• url</li></ul>	How well does this result match the search term? <ul style="list-style-type: none"><li>• Off-topic</li><li>• Acceptable</li><li>• Good</li><li>• Excellent</li></ul>
<a href="#">Preview</a> <a href="#">Use this template</a>		
<b>Query-Result Pair Relevance</b>		
Contributors judge how relevant a product (title, image, product page) is to a given search query	Minimum 4 columns: <ul style="list-style-type: none"><li>• query</li><li>• product_title</li><li>• product_image</li><li>• product_link</li></ul>	How well does this result match the query? <ul style="list-style-type: none"><li>• Off Topic (1)</li><li>• Acceptable (2)</li><li>• Good (3)</li><li>• Excellent (4)</li></ul> Output Columns: <ul style="list-style-type: none"><li>• Relevance</li><li>• Variance (2)</li></ul>
<a href="#">Preview</a> <a href="#">Use this template</a>		
<b>Query-Result Page Comparison (Competitive analysis)</b>		
Contributors judge how relevant product listings from multiple sites are to a given search query	Minimum 4 columns: <ul style="list-style-type: none"><li>• query</li><li>• comparison_site_1</li><li>• comparison_site_2</li><li>• comparison_site_3</li></ul>	Compare the Search Results to the Query and Score the Pages: <ul style="list-style-type: none"><li>• Result Page 1 Relevance (Very Poor, Poor, Good, Very Good)</li><li>• Result Page 2 Relevance (Very Poor, Poor, Good, Very Good)</li><li>• Result Page 3 Relevance (Very Poor, Poor, Good, Very Good)</li></ul>
<a href="#">Preview</a> <a href="#">Use this template</a>		

## Step 2: Design your job

[Insert Data](#)[Add Question](#)[Save](#)[Next: Create test questions](#)

### Instructions



#### Overview

Help us determine how relevant results are to their queries.

#### Steps

1. Read and Understand the Search Query
2. Review the Result of each Query
3. Select the level of relevancy for each result

#### Rules & Tips

- You should be assessing the results based on the **intent of the query**.
  - **Intent** is the purpose of the search
    - If a person is searching for Bacon, the intent of that search is to find Bacon (the meat) in the results.
    - If the result was "Bacon Flavored Chips" the intent would be wrong as the user was looking for bacon and not "bacon flavored" items.
- Review the information we've provided before making your decision.
- If you aren't sure what either the query or the result are, use Google search to see more examples of those products and use the product link to become familiar with the result.

#### Relevancy Definitions

- You should choose **Off-Topic** if:
  - The intent of the query was not matched.
  - The results are irrelevant to the search query
  - "*Why is this item even being returned?*"
- Choose **Acceptable** if:
  - The intent of the query is poorly matched.
  - The result is somewhat related to the query, but it not a good match.
  - "*I see why this is returned but it's definitely not what I wanted – I probably would not buy this*"
- Choose **Good** if:
  - Matches most of the query intent - or the most important part of the query.
  - Technically, all parts of the intent are satisfied but **result** doesn't provide a full, clear and complete answer to the search.
  - "*This broadly matches what I searched for so I might consider buying it, but it's not a perfect match*"
- Choose **Excellent** if:
  - The query intent is clearly satisfied.
  - **Specifics** of the Query appear in the Result
  - "*This is exactly what I was looking for*"

JOB ID 1316953

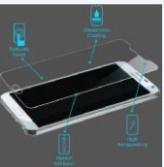
**Rate The Relevance Of Search Results**[Settings](#)[Preview](#)[Copy](#)

#### READ THE DOCS

[Design Tips](#)[Graphical Editor Guide](#)[Instructions Tips](#)

B I U C <>

## Examples

Query	Result	Relevance Score
White Jeans	<p><b>Title:</b> Uttermost Stockton White Rescued Denim Rug (5' x 8')</p> <p><b>Image:</b></p> 	<p><b>Off-Topic:</b></p> <ul style="list-style-type: none"> <li>"Why is this item even being returned?"</li> <li>The intent of the query was not matched</li> <li>The results are irrelevant to the search query</li> </ul>
Galaxy Note 3	<p><b>Title:</b> Premium Shatterproof Tempered Glass Screen Protector for Samsung Galaxy Note 3</p> <p><b>Image:</b></p> 	<p><b>Acceptable:</b></p> <ul style="list-style-type: none"> <li>"I see why this is returned but it's definitely not what I wanted - I probably would not buy this"</li> <li>The intent of the query is poorly matched</li> <li>The result is somewhat related to the query, but it <u>not</u> a good match</li> </ul>
DC Shoes Black	<p><b>Title:</b> Boys' DC Shoes Pure Black/Grey/Yellow</p> <p><b>Image:</b></p> 	<p><b>Good:</b></p> <ul style="list-style-type: none"> <li>"This broadly matches what I searched for so I might consider buying it, but it's not a perfect match"</li> <li>Matches most of the query intent - or the most important part of the query</li> <li>Technically, all parts of the intent are satisfied but <b>result</b> doesn't provide a full, clear and complete answer to the search. Relevant results are mixed in with irrelevant</li> </ul>

JOB ID 1316953  
[Rate The Relevance Of Search Results](#)

[Settings](#)  
[Preview](#)  
[Copy](#) ▾

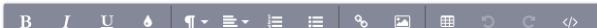
[READ THE DOCS](#)  
[Design Tips](#)  
[Graphical Editor Guide](#)  
[Instructions Tips](#)

Help Email Upgrade •••

## Step 2: Design your job

[Insert Data](#)[Add Question](#)[Save](#)[Next: Create test questions](#)

### Instructions



#### Overview

Help us determine how relevant results are to their queries.

#### Steps

1. Read and Understand the Search Query
2. Review the Result of each Query
3. Select the level of relevancy for each result

#### Rules & Tips

- You should be assessing the results based on the **intent of the query**.
  - **Intent** is the purpose of the search
    - If a person is searching for Bacon, the intent of that search is to find Bacon (the meat) in the results.
    - If the result was "Bacon Flavored Chips" the intent would be wrong as the user was looking for bacon and not "bacon flavored" items.
- Review the information we've provided before making your decision.
- If you aren't sure what either the query or the result are, use Google search to see more examples of those products and use the product link to become familiar with the result.

#### Relevancy Definitions

- You should choose **Off-Topic** if:
  - The intent of the query was not matched.
  - The results are irrelevant to the search query
  - "*Why is this item even being returned?*"
- Choose **Acceptable** if:
  - The intent of the query is poorly matched.
  - The result is somewhat related to the query, but it not a good match.
  - "*I see why this is returned but it's definitely not what I wanted – I probably would not buy this*"
- Choose **Good** if:
  - Matches most of the query intent - or the most important part of the query.
  - Technically, all parts of the intent are satisfied but **result** doesn't provide a full, clear and complete answer to the search.
  - "*This broadly matches what I searched for so I might consider buying it, but it's not a perfect match*"
- Choose **Excellent** if:
  - The query intent is clearly satisfied.
  - **Specifics** of the Query appear in the Result
  - "*This is exactly what I was looking for*"

JOB ID 1316953

Rate The Relevance Of Search

Results

[Settings](#)[Preview](#)[Copy](#)

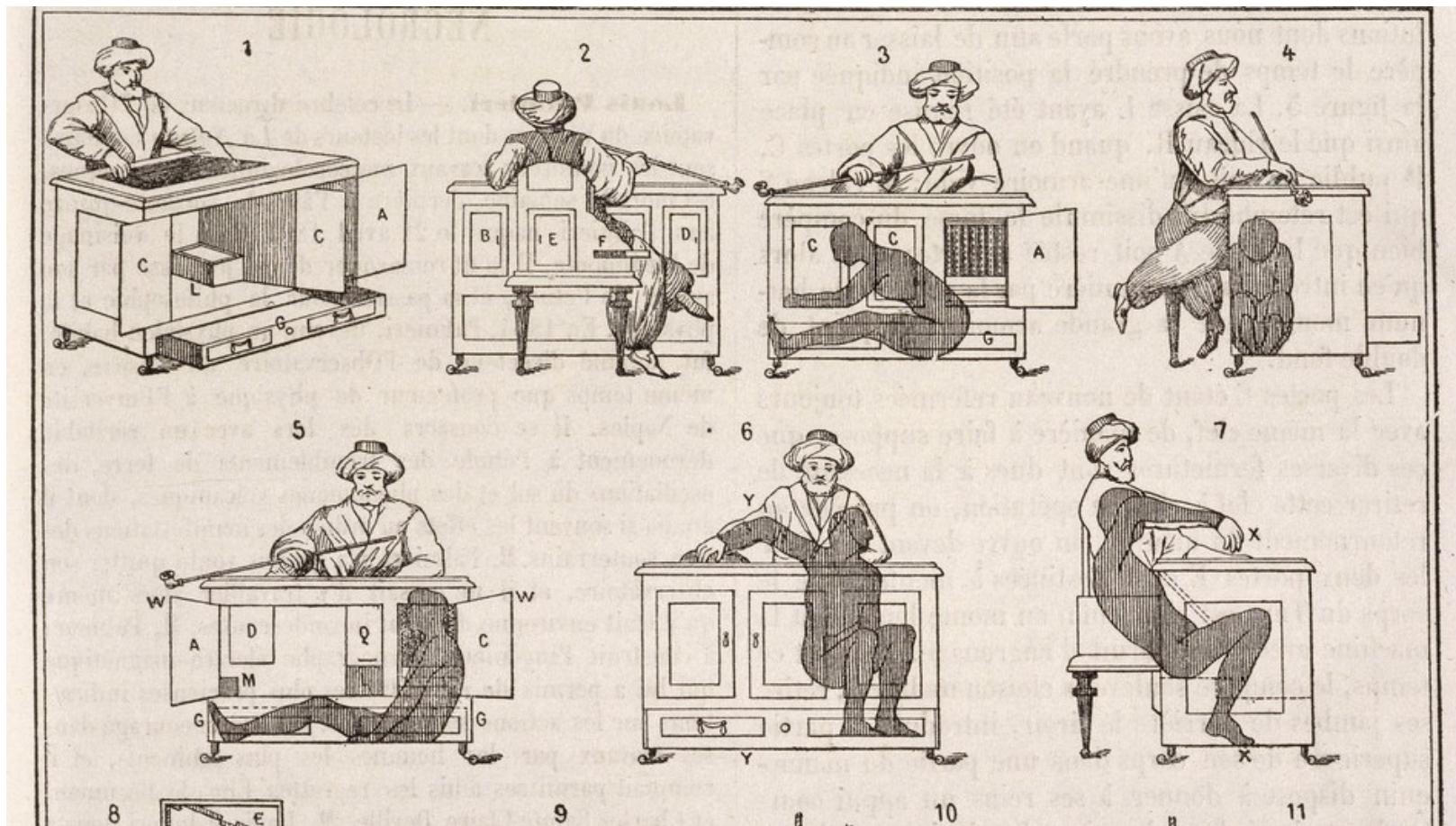
## READ THE DOCS

[Design Tips](#)[Graphical Editor Guide](#)[Instructions Tips](#)

# Amazon Mechanical Turk (AMT)



# Amazon Mechanical Turk (AMT)



# Amazon Mechanical Turk (AMT)

- Enabling wider range of tasks:
  - Content generation: queries!
  - Surveys.
  - Complex tasks: opening a page and clicking on a button!
  - Voice input: spoken queries!
- No automatic fraud detection -> no annual fees!
- Requesters should take care of “everything”:
  - Results verification.
  - Design an operational fraud detection mechanism.
  - Approval for payment.
- HTML

**amazonmechanical turk** beta REQUESTER

Create Manage Developer Help

[New Project](#) New Batch with an Existing Project

Start a New Project

**Survey Link**

- [Survey](#)
- [Tagging of an Image](#)
- [Moderation of an Image](#)
- [Choose image A or B](#)
- [Writing](#)
- [Collect data](#)
- [Collect data from a Website](#)
- [Transcription from A/V](#)
- [Transcription from a Receipt](#)
- [Transcription from an Image](#)
- [Sentiment Project](#)
- [Sentiment of a Tweet](#)
- [Categorization Project](#)
- [Other](#)

Example of Survey Link [Create 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.

[Create Project »](#)

[Leave feedback for this page.](#)

[Help](#) | [Pricing](#) | [Contact Us](#) | [Policies](#) | [State Licensing](#) | [Press Inquiries](#) | [Business Solutions](#) | [Blog](#) | [Service Health Dashboard](#) | [Careers](#)

MTurk.com | Requesters | Workers | Developers

©2005-2018 Amazon.com, Inc. or its Affiliates. All rights reserved.  
An  amazon.com company

[Follow Us on Twitter](#)

① Enter Properties

② Design Layout

③ Preview and Finish

Project Name: Sentiment of a Tweet

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

Font

**U** *I* **B** | **A-** | **I<sub>x</sub>**

≡ ≡ ≡ ≡

:= :=



#### Tweet Sentiment Analysis Instructions (Click to expand)

Pick the sentiment based on the following criterion:

Sentiment	Guidance
Positive	Select this if the item embodies emotion that was generally happy or satisfied. For example, "Sure I'll shop there again."
Neutral	Select this if the item does not embody 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 upset toward the topic. For example, "I don't know if I'll shop there again because I don't trust them."

Tweet:

Sentiment expressed by the Tweet:

Positive
Neutral
Negative

body

Save

Preview

## Setting up your HIT

### 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 HIT

 1

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

### Time allotted per assignment

 1 Hours

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

### HIT expires in

 7 Days

Maximum time your HIT 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 HITs ([Who are Mechanical Turk Masters?](#))

Yes  No

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

HIT Approval Rate (%) for all Requesters' HITs greater than 78 [Remove](#)

Number of HITs Approved greater than 1000 [Remove](#)

Location is UNITED STATES (US) [Remove](#)

[\(+ Add another criterion](#) (up to 2 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.

**HIT Visibility** ([What is HIT visibility?](#))

**Public** - All Workers can see and preview my HITs

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

**Hidden** - Only Workers that meet my HIT Qualification requirements can see and preview my HITs

✓ -- Select --

System Qualifications

HIT Approval Rate (%) for all Requesters' HITs

Location

Number of HITs Approved

Premium Qualifications

Blogger

Born 1918 to 1960 (Age 55 or older)

Born 1961 to 1971 (Age 45-55)

Born 1972 to 1981 (Age 35-45)

Born 1982 to 1986 (Age 30-35)

Born 1987 to 1991 (Age 25-30)

Born 1992 to 1999 (Age 18-25)

Borrower - Auto Loans

Borrower - Business Loan

Borrower - Credit Cards

Borrower - Home Mortgage

Borrower - Personal Loan

Borrower - Student Loan

Car Owner

Current Residence - Owned

Current Residence - Rented

Daily Internet Usage - 1 to 4 hours

Daily Internet Usage - 5 to 7 hours

Daily Internet Usage - 7+ hours

Employment Industry - Banking & Financial Services

Employment Industry - Education

Employment Industry - Food & Beverage

Employment Industry - Government & Non-Profit

Employment Industry - Healthcare

Employment Industry - Manufacturing

Employment Industry - Media & Entertainment

Employment Industry - Retail, Wholesale & Distribution

Employment Industry - Software & IT Services

Employment Sector - Non-Profit

Employment Status - Full time (35+ hours per week)

Employment Status - Part time (1-34 hours per week)

				Task Data																		
				Lifetime Approval Rate	Input.Task1	Input.Task1	Input.Task2	Input.Task2	Input.Task3	Input.Task3	Input.Task4	Input.Task4	Input.Task5	Input.Task5	Input.Task6	Input.Task6	Input.Task7	Input.Task7	Input.Task8	Input.Task8	Input.Task9	Input.Tas
HIT ID ▲		Worker ID		Lifetime Approval Rate	Input.Task1	Input.Task1	Input.Task2	Input.Task2	Input.Task3	Input.Task3	Input.Task4	Input.Task4	Input.Task5	Input.Task5	Input.Task6	Input.Task6	Input.Task7	Input.Task7	Input.Task8	Input.Task8	Input.Task9	Input.Tas
3VI0PC2ZAYD3D1J04YW0BFSO5CX0XC	A3VIQ9XPOYB8K6	100% (1/1)	1001	You recently watched a science program about th...	1002	You recently read a news story about an unexpla...	1003	You are bored with your current hairstyle and a...	1004	Tonight you are supposed to meet your friends i...	1005	Consider your favorite celebrity. You would lik...	1006	Consider one of your favorite pastries. It's be...	1007	Consider one of your oldest personal belongings...	1008	Today you heard an old song on radio while you ...	1009	Consider your favor genre of music. Op an ...		
3VI0PC2ZAYD3D1J04YW0BFSO5CX0XC	A3AY0315YWWNXY	100% (1/1)	1001	You recently watched a science program about th...	1002	You recently read a news story about an unexpla...	1003	You are bored with your current hairstyle and a...	1004	Tonight you are supposed to meet your friends i...	1005	Consider your favorite celebrity. You would lik...	1006	Consider one of your favorite pastries. It's be...	1007	Consider one of your oldest personal belongings...	1008	Today you heard an old song on radio while you ...	1009	Consider your favor genre of music. Op an ...		
3VI0PC2ZAYD3D1J04YW0BFSO5CX0XC	AAOQXV6U0E0JX	100% (1/1)	1001	You recently watched a science program about th...	1002	You recently read a news story about an unexpla...	1003	You are bored with your current hairstyle and a...	1004	Tonight you are supposed to meet your friends i...	1005	Consider your favorite celebrity. You would lik...	1006	Consider one of your favorite pastries. It's be...	1007	Consider one of your oldest personal belongings...	1008	Today you heard an old song on radio while you ...	1009	Consider your favor genre of music. Op an ...		
3VI0PC2ZAYD3D1J04YW0BFSO5CX0XC	A1398SBKK3NZ0W	100% (2/2)	1001	You recently watched a science program about th...	1002	You recently read a news story about an unexpla...	1003	You are bored with your current hairstyle and a...	1004	Tonight you are supposed to meet your friends i...	1005	Consider your favorite celebrity. You would lik...	1006	Consider one of your favorite pastries. It's be...	1007	Consider one of your oldest personal belongings...	1008	Today you heard an old song on radio while you ...	1009	Consider your favor genre of music. Op an ...		
3VI0PC2ZAYD3D1J04YW0BFSO5CX0XC	A2FJUGTKTCIB0	100% (1/1)	1001	You recently watched a science program about th...	1002	You recently read a news story about an unexpla...	1003	You are bored with your current hairstyle and a...	1004	Tonight you are supposed to meet your friends i...	1005	Consider your favorite celebrity. You would lik...	1006	Consider one of your favorite pastries. It's be...	1007	Consider one of your oldest personal belongings...	1008	Today you heard an old song on radio while you ...	1009	Consider your favor genre of music. Op an ...		
HIT ID ▲		Lifetime Approval Rate	Input.Task1	Input.Task1	Input.Task2	Input.Task2	Input.Task3	Input.Task3	Input.Task4	Input.Task4	Input.Task5	Input.Task5	Input.Task6	Input.Task6	Input.Task7	Input.Task7	Input.Task8	Input.Task8	Input.Task9	Input.Tas		

# Amazon AWS

- Result validation and approval can be automated.
- You need to write your own code to do that.
- MTurk offers an API to:
  - Retrieve submissions.
  - Send message to workers.
  - Send bonus to workers.
  - Approve/reject a submission.
  - Launch a new batch.
  - ...

# How do design a HIT?

- HIT: Human Intelligence Task
- Data distribution: how “micro” is each microtask?
  - Long paragraphs hard to digest, worker fatigue.
  - Single sentences not always appropriate.
  - Binary decision tasks get good results.
- Reward scheme
  - Granularity – per task? Per set of tasks? High scores?
  - What to do with “bad” work.
  - How much to reward:
    - No clear, repeatable results for quality.
    - High rewards get it done faster, but not better.
    - Pilot task gives timings, so pay at least minimum wage.

# Design

- Instructions:
  - Be clear.
  - Be concise.
  - Show several good examples.
  - More importantly, show bad examples.
  - Explain possible reasons why a submission would be rejected.
- Pilot the design, measure performance, try again:
  - Simple, clear design is important.
  - ACCEPT EVERY SUBMISSION DURING PILOT!

# Beware of the Pickpockets!

- How to avoid spammers:
  - Disable copy/pasting to text fields.
  - Record timestamps of events: clicks, keystrokes, etc.
  - Blend test questions.
  - Prevent spammers from submitting, rather than rejecting their submissions!!!!
    - Why?
  - Make it difficult to cheat!
    - How?



# Reject Happy!

- Try to maintain your reputation as a good requester!
- Many forums of crowdworkers:
  - Tuckopticon
  - MTurkCrowd
  - TurkerNation
- Rejecting so many submissions brings many complaints!
  - Nobody would work for you.
- Frustrating design would discourage workers.
- Positive feedback makes your life easy!
  - Be responsive! Answer emails, monitor the forums for comments about you.
  - Be generous and quick in payment!

# Don't be Reject Happy!

AMT Requester	Rating [info]	Description
<b>Mohammad Alian Nejadi</b> <a href="#">A1XE8EGTKOCTA</a> <a href="#">Averages »</a> <a href="#">HIT Group »</a> <a href="#">Review Requester</a> <a href="#">»</a>	FAIR: 5 / 5 FAST: 5 / 5 PAY: 5 / 5 COMM: 5 / 5	     <p>I have no idea how this guy has all green TO, but the hit he has up now is incomprehensible. There is a wall describing the task in detail, so I thought the search task was in the instruction. But there is no mention of what to search for in the instruction.</p>
<b>Mohammad Alian Nejadi</b> <a href="#">A1XE8EGTKOCTA</a> <a href="#">Averages »</a> <a href="#">HIT Group »</a> <a href="#">Review Requester</a> <a href="#">»</a>	FAIR: 5 / 5 FAST: 5 / 5 PAY: 5 / 5 COMM: 5 / 5	     <p>No, the search task is below the instruction, and mine was lebas which is supposed to be a persian dish. And obviously search brought back pages in persian which i can't read.</p>
<b>Mohammad Alian Nejadi</b> <a href="#">A1XE8EGTKOCTA</a> <a href="#">Averages »</a> <a href="#">HIT Group »</a> <a href="#">Review Requester</a> <a href="#">»</a>	FAIR: 5 / 5 FAST: 5 / 5 PAY: 5 / 5 COMM: 5 / 5	     <p>and then you're supposed to do something in the app in addition to the HIT window, like submitting the query used in the search window.</p>
<b>Mohammad Alian Nejadi</b> <a href="#">A1XE8EGTKOCTA</a> <a href="#">Averages »</a> <a href="#">HIT Group »</a> <a href="#">Review Requester</a> <a href="#">»</a>	FAIR: 5 / 5 FAST: 5 / 5 PAY: 5 / 5 COMM: 5 / 5	     <p>you have to do all this in mobile.</p>
<b>Mohammad Alian Nejadi</b> <a href="#">A1XE8EGTKOCTA</a> <a href="#">Averages »</a> <a href="#">HIT Group »</a> <a href="#">Review Requester</a> <a href="#">»</a>	FAIR: 5 / 5 FAST: 5 / 5 PAY: 5 / 5 COMM: 5 / 5	     <p>think about going back and forth between hit window, search window, and the app.</p>
<b>Mohammad Alian Nejadi</b> <a href="#">A1XE8EGTKOCTA</a> <a href="#">Averages »</a> <a href="#">HIT Group »</a> <a href="#">Review Requester</a> <a href="#">»</a>	FAIR: 5 / 5 FAST: 5 / 5 PAY: 5 / 5 COMM: 5 / 5	     <p>you have to do it 10 times irc.</p>
<b>Mohammad Alian Nejadi</b> <a href="#">A1XE8EGTKOCTA</a> <a href="#">Averages »</a> <a href="#">HIT Group »</a> <a href="#">Review Requester</a> <a href="#">»</a>	FAIR: 5 / 5 FAST: 5 / 5 PAY: 5 / 5 COMM: 5 / 5	     <p>gave up after 12 minutes because i couldn't find what lebas is or its recipe.</p>
<b>Mohammad Alian Nejadi</b> <a href="#">A1XE8EGTKOCTA</a> <a href="#">Averages »</a> <a href="#">HIT Group »</a> <a href="#">Review Requester</a> <a href="#">»</a>	FAIR: 5 / 5 FAST: 5 / 5 PAY: 5 / 5 COMM: 5 / 5	     <p>This review was edited by the author Wed Oct 24 12:38 PDT. This review was edited by the author Sun Oct 21 22:24 PDT.</p>
<b>Mohammad Alian Nejadi</b> <a href="#">A1XE8EGTKOCTA</a> <a href="#">Averages »</a> <a href="#">HIT Group »</a> <a href="#">Review Requester</a> <a href="#">»</a>	FAIR: 5 / 5 FAST: 5 / 5 PAY: 5 / 5 COMM: 5 / 5	     <p>This is a great requester. Yes you download an app on your phone and track your searches for 24 hrs. There are great opportunities for bonuses. Very quick to approve. Had an issue with the initial hit with not allowing me to submit. He responded within minutes that it was a hit issue and he reposted the HIT. I return the initial hit then I couldn't seem to find it because it was being grabbed by other workers. He made a penny hit for me to submit and bonused me the rest. Have received over \$14 in bonuses so far for very minimal work. Look forward to doing more work for them.</p>
<b>Mohammad Alian Nejadi</b> <a href="#">A1XE8EGTKOCTA</a> <a href="#">Averages »</a> <a href="#">HIT Group »</a> <a href="#">Review Requester</a> <a href="#">»</a>	FAIR: NO DATA FAST: NO DATA PAY: NO DATA COMM: NO DATA	     <p>In this task, we have to download an android app and let the app have access to our facebook,gmail,youtube accounts.I DON'T THINK SO!!!</p>
<b>Mohammad Alian Nejadi</b> <a href="#">A1XE8EGTKOCTA</a> <a href="#">Averages »</a> <a href="#">HIT Group »</a> <a href="#">Review Requester</a> <a href="#">»</a>	FAIR: NO DATA FAST: NO DATA PAY: NO DATA COMM: NO DATA	     <p>This review was edited by the author Apr 07 2018   <a href="#">reilly...@g...</a>   flag   comment   flags, comments »</p>
<b>Mohammad Alian Nejadi</b> <a href="#">A1XE8EGTKOCTA</a> <a href="#">Averages »</a> <a href="#">HIT Group »</a> <a href="#">Review Requester</a> <a href="#">»</a>	FAIR: NO DATA FAST: NO DATA PAY: NO DATA COMM: NO DATA	     <p>Requester here. I would like to clarify that our HIT does not violate MTurk policies. We do not collect any personal information from the users and explain it very clearly in the task description. For more clarification: the app usage statistics that we collect only collects the name of the apps you use on your phone as well as the duration of the usage. For instance, if you have used Facebook for one hour, we would know that you used Facebook for one hour on your phone. No other information about your Facebook account, friends, etc. are collected. I hope this was enough give you a clear idea of our HIT.</p>
<b>Mohammad Alian Nejadi</b> <a href="#">A1XE8EGTKOCTA</a> <a href="#">Averages »</a> <a href="#">HIT Group »</a> <a href="#">Review Requester</a> <a href="#">»</a>	FAIR: NO DATA FAST: NO DATA PAY: NO DATA COMM: NO DATA	     <p>This review was edited by the author Mar 22 08:22 PDT. This review was edited by the author Jan 20 13:27 PST.</p>
<b>Mohammad Alian Nejadi</b> <a href="#">A1XE8EGTKOCTA</a> <a href="#">Averages »</a> <a href="#">HIT Group »</a> <a href="#">Review Requester</a> <a href="#">»</a>	FAIR: NO DATA FAST: NO DATA PAY: NO DATA COMM: NO DATA	     <p>Rejected my hit with the comment "Please read the instructions." I read the instructions and did what was asked. Not sure this person knows how to write instructions.</p>
<b>Mohammad Alian Nejadi</b> <a href="#">A1XE8EGTKOCTA</a> <a href="#">Averages »</a> <a href="#">HIT Group »</a> <a href="#">Review Requester</a> <a href="#">»</a>	FAIR: NO DATA FAST: NO DATA PAY: NO DATA COMM: NO DATA	     <p>I spoke to this requester and I believe his mistakes were honest. He is trying to fix his hit so that the instructions are understandable, maybe there is a problem with the hosting mechanism, I'm not sure, but he seems sincere in his wish to address it.</p>
<b>Mohammad Alian Nejadi</b> <a href="#">A1XE8EGTKOCTA</a> <a href="#">Averages »</a> <a href="#">HIT Group »</a> <a href="#">Review Requester</a> <a href="#">»</a>	FAIR: NO DATA FAST: NO DATA PAY: NO DATA COMM: NO DATA	     <p>The requester reversed my rejection and gave me a bonus! Great requester!</p>
<b>Mohammad Alian Nejadi</b> <a href="#">A1XE8EGTKOCTA</a> <a href="#">Averages »</a> <a href="#">HIT Group »</a> <a href="#">Review Requester</a> <a href="#">»</a>	FAIR: NO DATA FAST: NO DATA PAY: NO DATA COMM: NO DATA	     <p>This review was edited by the author Jan 20 15:32 PST. This review was edited by the author Jan 20 15:08 PST.</p>
<b>Mohammad Alian Nejadi</b> <a href="#">A1XE8EGTKOCTA</a> <a href="#">Averages »</a> <a href="#">HIT Group »</a> <a href="#">Review Requester</a> <a href="#">»</a>	FAIR: 1 / 5 FAST: 2 / 5 PAY: 4 / 5 COMM: NO DATA	     <p>Very reject happy. Pays well, but rejected several of my HITs and I pretty much never get rejected.</p>

# Make it Difficult to Cheat

- How to prevent spamming in tasks that workers need to generate content: summarization, queries, provide subjective feedback, etc.
- It is extremely important to have **explicitly verifiable** questions as part of the task.
- “An important role of verifiable questions is in **signaling to users** that their answers will be scrutinized, which may play a role in both reducing invalid responses and increasing time-on-task.”
- “It is advantageous to design the task such that completing it accurately and in good faith requires **as much or less effort** than non-obvious random or malicious completion.”

Aniket Kittur, Ed H. Chi, Bongwon Suh: Crowdsourcing user studies with Mechanical Turk. CHI 2008: 453-456

# Not Easy to Cheat

- Collecting feedback on Wikipedia articles.
- Experiment 1:
  - Some scales on the quality of the article in terms of writing, clarity, etc.
  - A free-form text box where they could write their opinion on how the text could have been improved.
- Experiment 2:
  - Identical to Experiment 1 with additional verifiable questions.
  - Selecting the number references, figures, and sections the article had.
  - Providing 4-6 keywords that would summarize the article.

# Results

- Compared to an expert-assessed baseline:

	Invalid Comment Responses	Median duration	Duration < 1 minute
Exp 1	48.6%	1:30	30.5%
Exp 2	2.5%	4:06	6.5%

# Sometimes the Problem is You

- Test everything on the sandbox!
  - <http://requestersandbox.mturk.com>
- Ask your friends to try it out first.
- Run the pilot.
  - What if all submissions of the pilot are bad?
  - Maybe your design is faulty!
  - Reconsider!
  - DO NOT REJECT PILOT SUBMISSIONS!
    - Why?



# Venue Appropriateness

- To define which types of venues are appropriate to visit under a defined context.
- We used appen (figure eight at the time):
  - Great user interface.
  - Easy and fast to set up.
  - Don't get bothered with approval and quality check.
  - It handles the payment, so no complaints.
  - It provides a nice interface between you and experienced workers for advice on how to improve your design.
- Quality control: test questions.

Mohammad Aliannejadi, Fabio Crestani: Venue Appropriateness Prediction for Personalized Context-Aware Venue Suggestion. SIGIR 2017: 1177-1180

# Overview

In this task, we need your help to categorize venues and define their appropriateness to different descriptions of trips.

---

## Steps

1. Read and understand the keyword describing the trip. They are:

1. Trip type, which could be:

**Trip Type: Business** -- meaning the person is on a business trip

**Trip Type: Holiday** -- meaning the person is on a holiday trip

**Trip Type: Other** -- meaning a trip different from the previous ones, for example: tourism, adventure, family visit, etc.

2. Trip duration, which could be:

**Trip Duration: Night out** -- meaning the person is going out for a night

**Trip Duration: Day trip** -- meaning the person is on a trip lasting a day

**Trip Duration: Weekend trip** - meaning the trip is on a weekend

**Trip Duration: Longer** -- meaning longer than a weekend

3. Type of group the person is traveling with, which could be:

**Group Type: Alone** -- meaning traveling alone

**Group Type: Friends** -- meaning traveling with a group of friends

**Group Type: Family** -- meaning traveling with family

**Group Type: Other** -- meaning traveling with a group of people which different from the previous ones, for example: colleagues, classmates, etc.

2. Read the venue.

3. Select one of the possible answers:

1. "Yes" if the venue is appropriate/recommended for the trip,

2. "No" if the venue is not appropriate/recommended for the trip,

3. "I can't decide" if you have any doubts

## Rules & Tips

- You should be assessing the venue based on the **intent of a potential user and not based on yourself and your personal tastes.**
  - For example if you are asked to evaluate *Pizza Place* for the trip: *Group type: Family*, the answer would be yes, irrespective of the fact that you may like or not like pizza.
- Please note: you are **only** supposed to judge the appropriateness of a venue type given a descriptor. Please do not consider other factors when you do the judgement. For example, if the trip type is *holiday* and the venue type is *asian restaurant*. Do not consider if a visitor likes asian food. Only judge based on the fact that an *asian restaurant* is appropriate to visit on a *holiday trip* or not.
- If you aren't sure about the answer, you can select "**I can't decide**" and go to the next question.

## Examples

Venue Type	Trip Descriptor	Answer
Pizza Place	Trip Type: Holiday	YES
Pizza Place	Trip Type: Business	NO <i>Tip: A Pizza Place is not the best place for inviting business partners</i>
Sushi Bar	Trip Duration: Weekend trip	YES

## Examples

Venue Type	Trip Descriptor	Answer
Pizza Place	Trip Type: Holiday	YES
Pizza Place	Trip Type: Business	NO <i>Tip: A Pizza Place is not the best place for inviting business partners</i>
Sushi Bar	Trip Duration: Weekend trip	YES
Pub	Trip Duration: Night out	YES
Museum	Trip Duration: Night out	NO <i>Tip: A Museum is not the best place to visit late at night.</i>

**Trip Description:**

Trip type: Business

**Venue Type:** Food

**Is the venue appropriate for the trip described by the keyword? (required)**

- Yes
- No
- I can't decide

**Trip Description:**

Trip type: Holiday

**Venue Type:** Food

**Is the venue appropriate for the trip described by the keyword? (required)**

- Yes
- No
- I can't decide

# In-app Mobile Queries

- We were interested in analyzing the queries that users submit to various apps:
  - How queries differ when users search in Google compared to YouTube, Facebook, Twitter, etc.
  - No such data is available.
  - We set up two HITs:
    1. To create realistic mobile search tasks
    2. To generate query-app pairs for a given search task.
- Amazon Mechanical Turk:
  - No automatic quality check is possible.
  - We tried to make cheating difficult.
  - We did manual checks.

Mohammad Aliannejadi, Hamed Zamani, Fabio Crestani, W. Bruce Croft:

Target Apps Selection: Towards a Unified Search Framework for Mobile Devices. SIGIR 2018: 215-224

# Creating Mobile Search Tasks

- Asked workers to explain their latest mobile search experience.
- Defined a general topic:
  - When was the last time you searched for a video on your phone?
  - Asked for a lot of details.
  - Provided a very detailed example.
- We got impressive results for little money!

## Writing Instructions (Click to expand)

Write a detailed description of your latest search on YouTube:

- What was it that you were searching for on YouTube? Give us details about the task you were trying to do.
- When did it happen? What were you doing? Was anyone else with you?
- Why did you need to search for that video? Give us a bit of the background and context related to the search task.
- **Example:** Last week my friend wanted to visit me with his girlfriend. Since his girlfriend is vegan, I decided to cook a vegan dish for her. Since I knew that she likes Indian food, I searched for recipes for vegan Indian food. After watching a couple of videos I realized that I need to know if she is tolerant to gluten. Later when I realized that she was gluten intolerant, I searched for vegan recipes with no gluten.

**Task:**

Describe your latest experience with YouTube.

**Write the description below:**

# App-Query Pairs

- Created abstract mobile tasks using the results of the first HIT.
- Made it difficult to cheat:
  - Asked to select two apps.
  - “If I select the first app randomly, what should I select as the second app? Mmmmm, I may as well do it right!”
- The order of apps was random.
- Every user interaction was recorded: mouse clicks, keystrokes.
- Copy/pasting was disabled.

### **Task Description:**

Welcome and thanks for your interest. We would like to understand how mobile users search for their information needs when they use their mobile devices. So, we describe a situation where you would need to find some relevant information and we ask you to tell us which queries you use and which apps you would choose to submit your query to. Therefore, you need to follow the following steps:

- You are given a specific task for which you need to find relevant information on your mobile device.
- Read the task carefully and imagine you need to perform it, using your own mobile device using one of its apps.
- Type the name of the apps or select one or two apps from the list of available apps in the order of your preference by clicking on the icons.
- If you click on an app and then change your mind, you may deselect it by clicking on it again.
- Imagine you have opened your selected app and want to find the relevant information searching in it. Write the query you would use in "**Query.**"

### **\*\*\*\*\* Bonus \*\*\*\*\***

- We will keep publishing new HITs in the following days. Please keep doing as many HITs as you wish. You will receive a 10 cent **bonus** if you do at least 100 of our HITs.
- We highly encourage you to fill the **survey**. It will take you no more than 2 minutes and you will receive a 5 cent **bonus** for it.

**IMPORTANT:**

- Please avoid queries like "name of location" or "city name hotels" or "(singer name) songs" or "my friend's number". We need **authentic queries** like "mcdonald's" or "new york hotels" or "katy perry songs" or "sam's number". You can either use real names (like your friend's name or your home town) or make up a name (another town nearby or a random guy's name).
- Please note that while you need to do the survey only once, we highly encourage you do as many HITs as you wish. There is no restrictions on the number of HITs.

**EXAMPLE:**

Task Description: You are trying to find a restaurant to have lunch. The restaurant must be close to you and serve your favorite food.

First App: Foursquare

Second App: Yelp

Query: restaurants open for lunch serving pizza

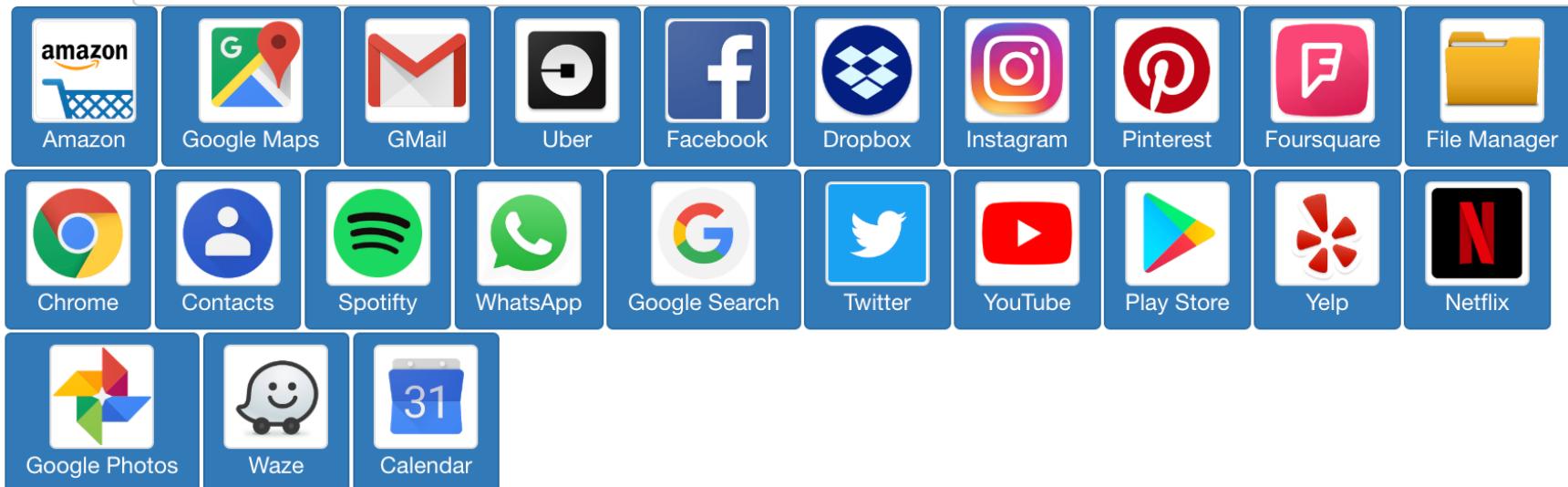
### Task Description

Select some apps related to your email. Think of something you would need in your email and search for it.

### Select Two Appropriate Apps

- Type the names of the apps or select the apps below:

Please select at least one app



### Query

- Type the query you would enter in the select apps:

Please leave the query here

# In Situ Mobile Data

- Moving one step forward: queries on mobile devices in-the-wild.
- A bespoke Android app to collect data.
- Asked workers to install the app and let it run for 24 hours.
- Report at least 7 queries in different apps right after they do the search.
- Bonus for extra queries!
- Quality check: based on mobile app usage data and manual check.
- Used the API to handle bonus payment.

Mohammad Aliannejadi, Hamed Zamani, Fabio Crestani, W. Bruce Croft:  
In Situ and Context-Aware Target Apps Selection for Unified Mobile Search. CIKM 2018: 1383-1392

## Mobile App Query Instructions (Click to collapse)

### Task Description:

The purpose of this HIT is to collect some training data for a machine learning algorithm involving mobile search, as part of an academic research study.

Some notes about this HIT before you begin.

### This HIT:

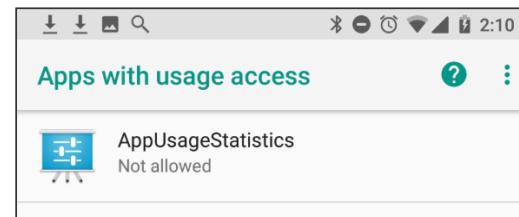
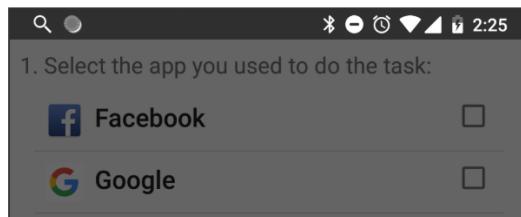
- Can only be completed on an Android phone.
- No repeats are allowed. Please skip if you have done this HIT in the past.
- The app only works on phones with Android 5 or higher.
- Requires you to download an Android App to complete the task.
- Requires access to the device's app usage info ([no sensitive info](#)).

### The App:

- The UniMobile app is available on Google Play Store. [Click here](#) to install.
- Collects no information which could be used identify you besides your location.
- Granting location permission is optional.
- [Please note that our app collects absolutely no data from any other apps on your phone.](#)
- Can be uninstalled or deleted after completion of the HIT.
- Has no lasting effect on the device.

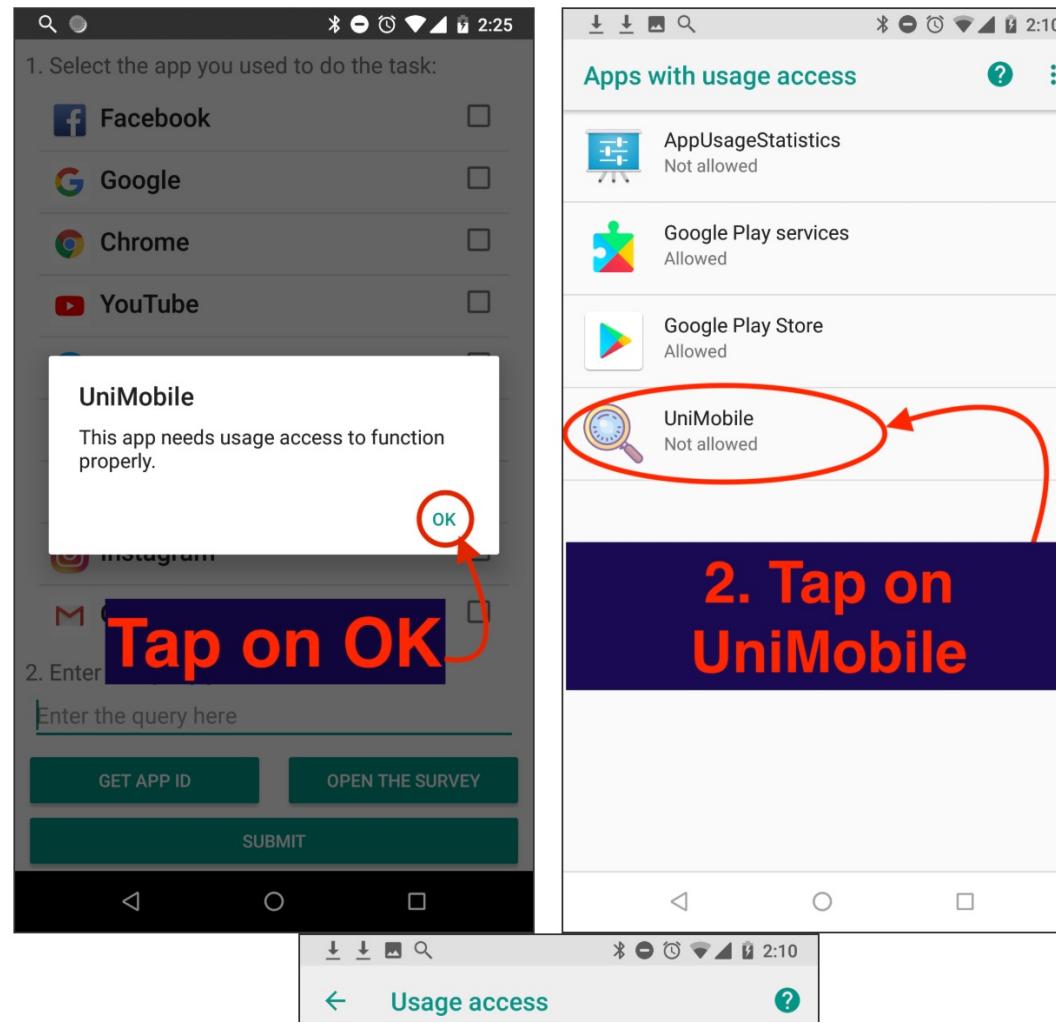
### Before you accept this HIT:

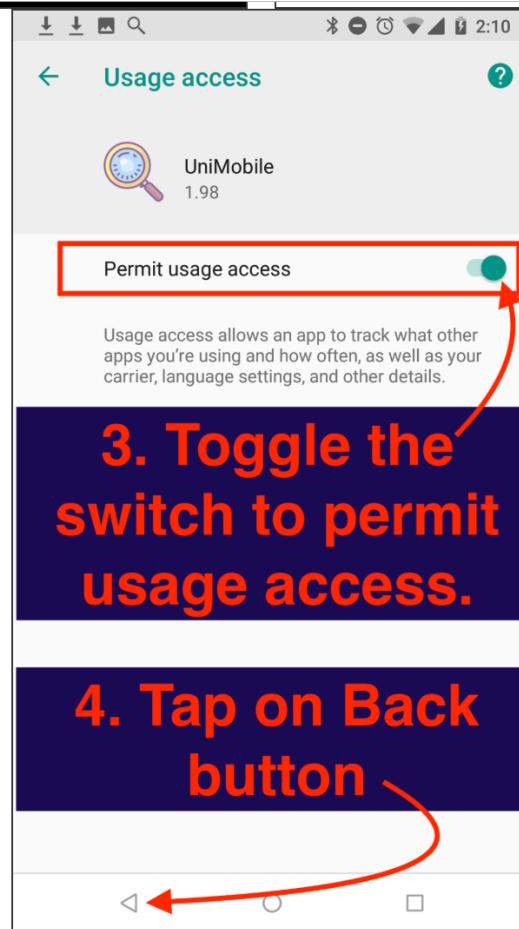
- Ensure that you have the most recent app update the app through the Play Store to the latest version if you have completed a previous HIT for us using the App.
- You need to complete the survey and get your App ID and submit it here in the HIT.
- After installing the app, follow the steps below to permit usage access on your phone:



### Before you accept this HIT:

- Ensure that you have the most recent app update the app through the Play Store to the latest version if you have completed a previous HIT for us using the App.
- You need to complete the survey and get your App ID and submit it here in the HIT.
- After installing the app, follow the steps below to permit usage access on your phone:





## What you need to do:

1. Tap on "OPEN THE SURVEY" button in the UniMobile app and complete the survey.
2. Tap on the button named "GET APP ID" and copy your **App ID** on this page.
3. Let the app run on your phone for at least 24 hours.
4. Report **at least 7 organic** mobile searches that you perform **within this 24 hours**, following the steps below:

## What you need to do:

1. Tap on "OPEN THE SURVEY" button in the UniMobile app and complete the survey.
2. Tap on the button named "GET APP ID" and copy your **App ID** on this page.
3. Let the app run on your phone for at least 24 hours.
4. Report **at least 7 organic** mobile searches that you perform **within this 24 hours**, following the steps below:
  - i. Whenever you use any app on your phone to search for information (e.g., Google, Facebook, Gmail, etc.), open UniMobile after your search is completed.
  - ii. Remember you can report any search **as long as you entered a query** for it. Some example apps: Google, Facebook, Gmail, Notes, Memo, Calendar, Contacts, WhatsApp, Instagram, Twitter, File Manager, Reminders, Google Keep, Uber, Google Maps, Foursquare, Yelp, Youtube, Spotify, SoundCloud, Amazon, etc.
  - iii. From the list of apps available in the UniMobile app, select the app that you used to perform your search (e.g., Twitter).
  - iv. Type the search query that you used in the selected app (e.g., you searched for #metoo in Twitter, so you should type #metoo).
  - v. Please type exactly what you entered as query.
  - vi. Press "SUBMIT" button.
5. Please submit the HIT as soon as you have the App ID, but keep on following Step 4.

### Acceptance requirements:

- You must report **at least 7 organic** mobile search queries (if you don't do 7 mobile search tasks during the 24-hour period, keep the app running until you report 7 search queries).
- Please only report your **real-life organic search queries**. If you submit all queries in a short time, or if you report a query without actually submitting it in another app, there are mechanisms that help us realize it.
- Only report your **real-life organic search queries** that occur **within the 24-hours** that you install the app. If you report the queries that you have done before installing the app, that would not be helpful to us.
- The "App Usage" permission must be granted collecting app usage statistics (i.e., how often you use which apps).
- **Please note that our app collects absolutely no data from any other apps on your phone.**
- The app transfers the data **only via WiFi**. So, please make sure your phone is connected to WiFi so that we can receive the data.

## How to get bonus:

- If you impress us with the quality of your work we will send you some random bonus. Please remember "longer" queries do not necessarily mean "more impressive." More naturalistic queries are rewarded as if you do it in a real-life scenario.
- If you can keep the app running for more than 24 hours and keep reporting your **organic** mobile search queries, we send you a **\$0.1 bonus** per extra valid query.
- If you decide to keep the app running for more than 24 hours, please let us know in the comments below and **contact us via email**

## How to get bonus:

- If you impress us with the quality of your work we will send you some random bonus. Please remember "longer" queries do not necessarily mean "more impressive." More naturalistic queries are rewarded as if you do it in a real-life scenario.
- If you can keep the app running for more than 24 hours and keep reporting your **organic** mobile search queries, we send you a **\$0.1 bonus** per extra valid query.
- If you decide to keep the app running for more than 24 hours, please let us know in the comments below and **contact us via email** ([mohammad.aliannejadi@gmail.com](mailto:mohammad.aliannejadi@gmail.com)). You can continue submitting queries for maximum 7 days and maximum 50 queries.
- **Remember, you don't search only in Chrome and Google.** Many of your daily searches happen in other apps such as Notes, Email, Facebook, Maps, YouTube, Twitter, Contacts, Instagram, Amazon, Pinterest, Google Play, Spotify, Messages, etc. If you remember to report not only your searches using Chrome and Google but also the searches you do in other apps, we will appreciate that by sending you some extra bonus.

## Invalid queries:

- Here you are not searching for apps. Therefore, you must avoid queries like "facebook", "youtube".
- If you submit all queries in a short time, or if you report a query without actually submitting it in another app, there are mechanisms that help us realize it.
- Please avoid queries like "name of location" or "city name hotels" or "(singer name) songs" or "my friend's number". We need **authentic queries** reflecting your real information need where you would type the name of a location or the name of your city. For example: "mcdonald's" or "new york hotels" or "katy perry songs" or "sam's number". You can either use real names (like your friend's name or your home town) or make up a name (another town nearby or a random guy's name).
- **Do not explain what you were doing.** We simply need the query you entered in the app. Some invalid examples: "i was trying to find my friend's updates", "i was reading news". We only need the queries you entered into other apps to search for information.

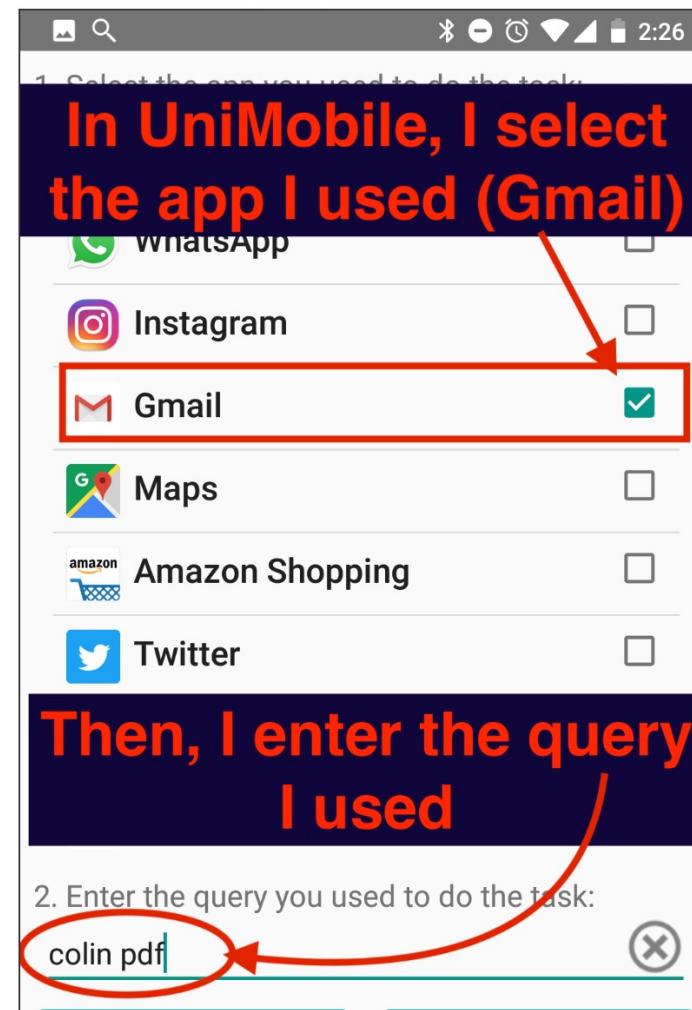
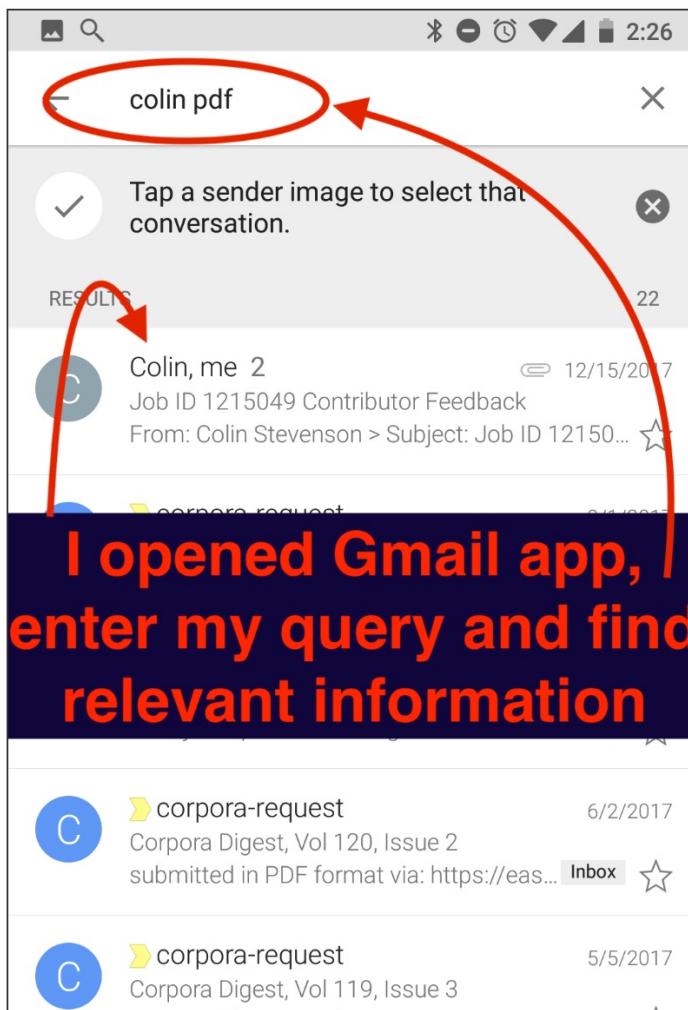
## Example:

- **My organic search scenario:** I received an email from my best friend a while ago. Now, I was trying to find that email but couldn't remember its subject. The only thing I remembered was that it had a pdf attachment.
- **How did I do it?**
  1. I opened Gmail because I use it to send/receive email.
  2. Colin is my best friend. So, I was looking for an old email from him.
  3. I used the query "colin pdf" in Gmail to find the message, you can see the results in the figure below.
  4. Then, I opened the UniMobile app and selected Gmail from the list of apps, because I did the search in Gmail.
  5. I entered the same query and tapped on SUBMIT button.



- How did I do it?

1. I opened Gmail because I use it to send/receive email.
2. Colin is my best friend. So, I was looking for an old email from him.
3. I used the query "colin pdf" in Gmail to find the message, you can see the results in the figure below.
4. Then, I opened the UniMobile app and selected Gmail from the list of apps, because I did the search in Gmail.
5. I entered the same query and tapped on SUBMIT button.



# Summary

- Crowdsourcing enables cheap and fast data collection.
- A good design requires clear and concise instructions.
- A well-thought design saves you a lot of time and money.
- Test everything on sandbox and run several pilot studies.
- Maintain a healthy communication with the workers.
- Study as many examples as you can.
- Pay fairly!

Questions?