

# computational social media

## lecture 7: crowdsourcing

daniel gatica-perez

# announcements

assignment #4 will be given today

reminder: in-depth course evaluation

reading #8 will be presented today

T. Bolukbasi, K.-W. Chang, J. Zou, V. Saligrama, and A Kalai, Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings, NeurIPS 2016

# this lecture

1. introduction

2. categorization of crowdsourcing systems

3. understanding crowdsourcing work

mechanical turk workers

crowdlabeling models

4. designing crowdsourcing systems

5. crowdsourcing as labor

6. citizen science

# **1. introduction**

score  
100

# ESP Game

Concentrate...

time  
2:21

## What do you see?

taboo words

peace

lay



guesses

sheeps...

sheep

+ submit

→ pass



Human Computation  
(Luis von Ahn, 2004)  
[espgame.org](http://espgame.org), [gwap.com](http://gwap.com)

Issue 14.06 - June 2006

[Subscribe to WIRED magazine and receive a FREE gift!](#)

term coined by Jeff Howe (2005)

## The Rise of Crowdsourcing

Remember outsourcing? Sending jobs to India and China is so 2003. The new pool of cheap labor: everyday people using their spare cycles to create content, solve problems, even do corporate R & D.

By Jeff Howe

Page 1 of 4 [next >>](#)

### 1. The Professional

#### Story Tools



PRINT



MAIL

#### Story Images

**Claudia Menashe** needed pictures of sick people. A project director at the National Health Museum in Washington, DC, Menashe was putting together a series of interactive kiosks devoted to potential pandemics like the avian flu. An exhibition designer had created a plan for the kiosk itself, but now Menashe was looking for images to accompany the text. Rather than hire a photographer to take shots of people suffering from the flu, Menashe decided to use preexisting images – stock photography, as it's known in the publishing industry.

In October 2004, she ran across a stock photo collection by Mark Harmel, a freelance photographer living in Manhattan Beach, California. Harmel, whose wife is a doctor, specializes in images related to the health care industry. "Claudia wanted people sneezing, getting immunized, that sort of thing," recalls Harmel, a slight, soft-spoken 52-year-old.

The National Health Museum has grand plans to occupy a spot on the National Mall in Washington by 2012, but for now it's a fledgling institution with little money. "They were on a tight budget so I charged them my nonprofit rate," says Harmel

#### Feature:

The Rise of Crowdsourcing

#### Plus:

[5 Rules of the New Labor Pool](#)

[Look Who's Crowdsourcing](#)

- ▣ Choose your own *Wired* adventure!
- ▣ 18 days of [depraved](#) computing

# this lecture

1. introduction
2. categorization of crowdsourcing systems
3. understanding crowdsourcing work
  - mechanical turk workers
  - crowdlabeling models
4. designing crowdsourcing systems
5. crowdsourcing as labor
6. citizen science

**2.**  
**definitions & categorization**  
**of crowdsourcing systems**



# definition of crowdsourcing



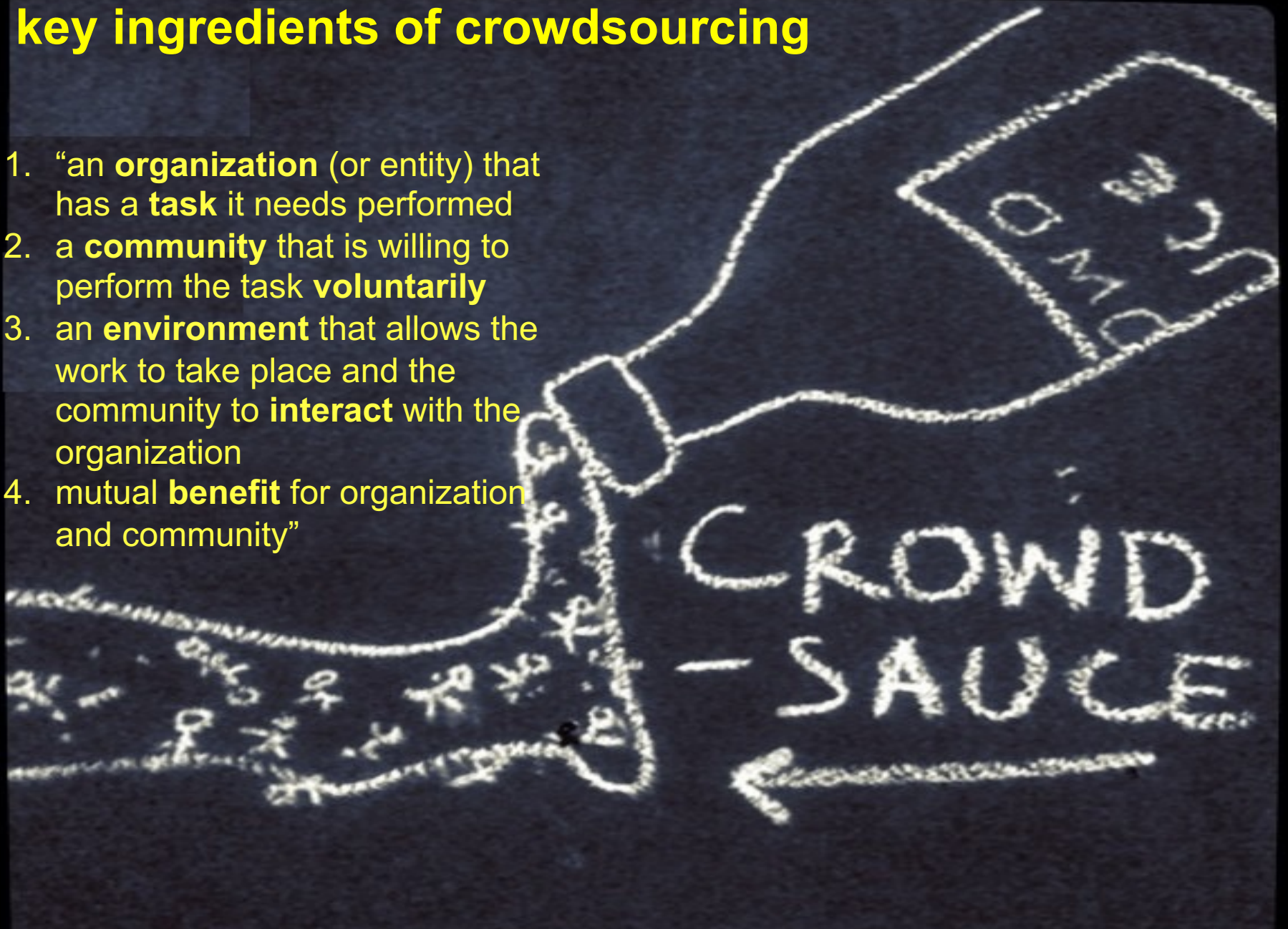
source: cooltownstudios

“**crowdsourcing** is an online, distributed **problem-solving and production model** that leverages the collective intelligence of online communities to serve specific organizational goals”

“**crowds** are given the opportunity to respond to tasks promoted by the organization motivated to respond for a variety of reasons.”

# key ingredients of crowdsourcing

1. “an **organization** (or entity) that has a **task** it needs performed
2. a **community** that is willing to perform the task **voluntarily**
3. an **environment** that allows the work to take place and the community to **interact** with the organization
4. mutual **benefit** for organization and community”



# problem-focused categorization of crowdsourcing

type	how it works	ideal kind of problems	examples
knowledge discovery & management	crowd has to find & collect info into a common format	info gathering, reporting problems, creation of collective resources	<a href="http://seeclickfix.com">seeclickfix.com</a> <a href="http://fixmystreet.com">fixmystreet.com</a> <a href="http://gigwalk.com">gigwalk.com</a>

## Report neighborhood issues and see them get fixed.

Over 881,952 Issues Fixed. Is yours next?



### HOW IT WORKS

From reporting potholes to receiving alerts, learn how SeeClickFix will help your neighborhood.



### MEDIA PARTNERS

Embed our widgets, discover citizen concerns and build conversations around community issues.



### GOVERNMENT PARTNERS

Receive citizen reports using our free tools or purchase our customized mobile, web and facebook apps.

### TOP PERFORMING CITIES

We run global analyses to see how cities are responding to their citizens. How does your city compare? [View More](#).

- |                     |                      |
|---------------------|----------------------|
| 1 San Francisco, CA | 6 Absecon, NJ        |
| 2 Tontitown, AR     | 7 Albany, NY         |
| 3 Albuquerque, NM   | 8 Northfield, NJ     |
| 4 Malden, MA        | 9 Randolph, MA       |
| 5 Eden Prairie, MN  | 10 Winston-Salem, NC |

### RECENT BLOG POSTS

[more](#)

- [Carhoots Features SeeClickFix App](#)  
05/15/2014 - 09:04PM
- [SeeClickFix App Now Available in BlackBerry World](#)  
05/13/2014 - 09:18PM
- [Say "Hello!" to Stoneham, MA](#)  
05/07/2014 - 03:17PM



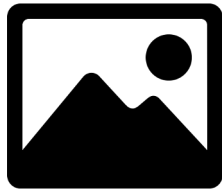
Date



Location



Text  
Description



Photos



Category  
Label



Online  
Platform



City  
Government



## Seattle, WA [change]

[Follow this Place](#)
[Report an Issue](#)

- ISSUES
- ANSWERS
- NEIGHBORS
- WATCH AREAS



### ISSUES



#### Street is improperly grade ► Open

4500 23rd Avenue Southwest Seattle, Washington

The grading does not allow rain water to enter the drain. Water cascades down the street and flows onto SW Oregon. In the winter the area can become a sheet of ice. 23rd Ave SW is a fairly steep hill and the ice or water on the road surface makes controlling a vehicle difficult or dangerous at this intersection with SW Oregon.

1  
vote

FIX IT!

05-15-2014 · Reported by [H Pryor \(Guest\)](#) · [Share](#) · [Flag](#)



#### Dangerous Intersection ► Open

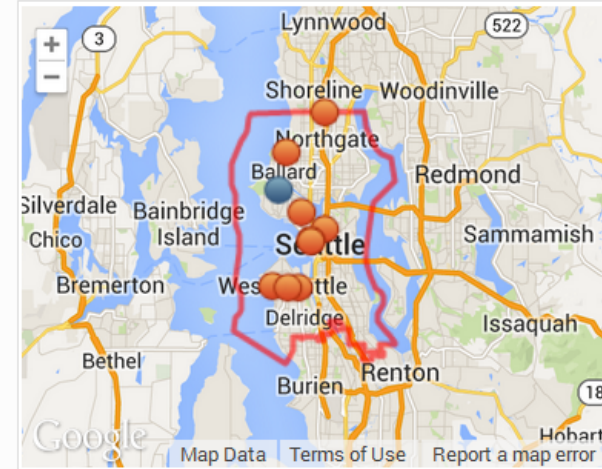
14203-14299 5th Avenue Northeast Seattle, Washington

There's nothing to slow down drivers when they exit I-5 at this intersection. Frequently, drivers will speed up and then slam on their brakes to stop at the stoplight on 145th st. This leads to a very dangerous intersection for drivers that are trying to cross north and south on 5th Ave NE. Is there anyway to slow down people exiting I-5 without causing backup on I-5?

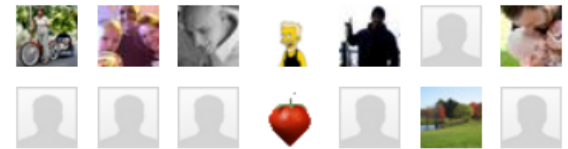
1  
vote

FIX IT!

05-14-2014 · Reported by [Rabitepoo](#) · [Share](#) · [Flag](#)



#### TOP USERS

[more](#)


#### WHO'S WATCHING

[more](#)

[COBI Public Works \(PRO\)](#)

[Brenda Cooper](#)

[Bridle Trails](#)

This site is for reporting **problems in the UK**. There are FixMyStreet sites [all over the world](#), or you could set up your own using the [FixMyStreet Platform](#).

[Report a problem](#) [Your reports](#) [All reports](#) [Local alerts](#) [Help](#)

mySociety



# Report, view, or discuss local problems

(like graffiti, fly tipping, broken paving slabs, or street lighting)

Enter a nearby UK postcode, or street name and area:

GO



... or locate me automatically

## How to report a problem

- 1 Enter a nearby UK postcode, or street name and area
- 2 Locate the problem on a map of the area
- 3 Enter details of the problem
- 4 We send it to the council on your behalf

## Recently reported problems

**Deep pothole**

10:01 today



**Abandoned car used for advertising purposes. V831 UBD**

09:50 today



**Footpath is breaking up**

09:40 today



# Züri wie neu: Zürich like new

Record a message

All messages

help

## Report damage and deficiencies in Zurich's infrastructure

Approximate address of the damage:

e.g. 'Langstrasse' or 'Basteiplatz'

Use current position

### Create a new message:

- 1 Approximate address of the damage
- 2 Locate the damage on the map
- 3 Describe the damage

**106**

Messages in the last  
week

**513**

processed messages in the last  
month

### Recently recorded messages:

Stain of unknown flu

09:37 today



Scooter in the way

15:47 today



Disposed wild

11:01 am today



Again illegal waste.

8:33 am today

Disposal contrary to the rules

2:04 p.m., Wednesday



# open data helps identify main issues in public space

Reports: May 2013 - Oct 2020

Category	Reports	Media
Waste/CollectionPoint	7234	3739
Street/Sidewalk/Square	5441	2127
Signalization/LightSignal	4452	3385
GreenAreas/Playgrounds	2443	1825
Lighting/Clocks	2231	866
Graffiti	1330	1153
Other*	1511	946
Vbz/Öv** (public transport)	566	382
Wells/Hydrants**	384	89
General**	306	245
Pests**	255	230
Total	24642	14041

Table 1: Number of reports, and reports with media per category. Categories marked with \*\* are gathered into the meta-category Other\*.

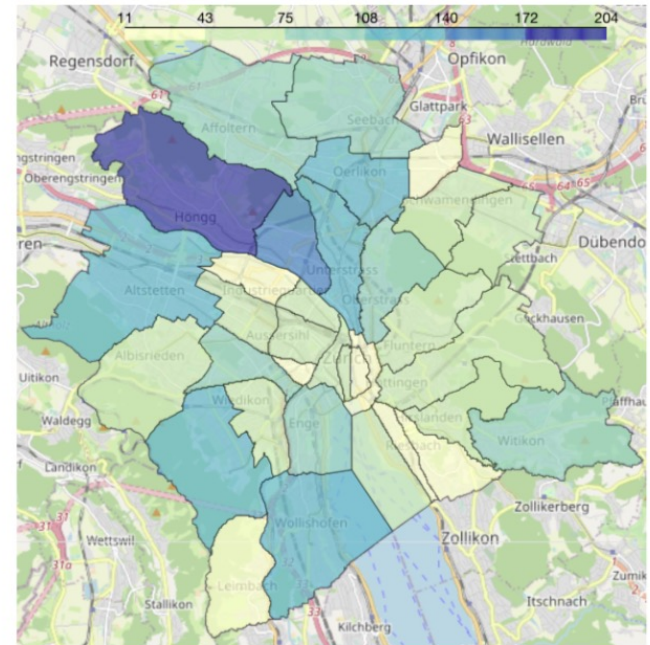


Figure 3: Heatmap of reports for *GreenAreas/Playgrounds*. Yellow indicates a lower number of reports, Green indicates a higher number, and Blue the highest number.

[READ MORE »](#)

## CONSUMER BRANDS

Are your products stocked? Displayed correctly? Do you know for sure? Your team (or ours) can report the real-time data you need.

[READ MORE »](#)

## RETAIL & MERCHANDISING

Make it easier than ever to get information to and from your field teams. Our mobile tools execute and validate your in-store execution.

[READ MORE »](#)

## CONSULTING & MARKET RESEARCH

Conduct real-time research and collect data with unprecedented speed and scale. It's the intelligent way to gather business intelligence.

[? Help](#)

video: [https://www.youtube.com/watch?v=yEwy\\_53LILl](https://www.youtube.com/watch?v=yEwy_53LILl)

## Create



Create projects and push out to your field team or our 1.1 Million Gigwalkers

## Match



Match people to work based on GPS locations

## Execute



Execute on work, efficiently and effectively

## Analyze



Real time insight to make better decisions

# problem-focused categorization of crowdsourcing

type	how it works	ideal kind of problems	examples
knowledge discovery & management	crowd has to find & collect info into a common format	info gathering, reporting problems, creation of collective resources	<a href="http://seeclickfix.com">seeclickfix.com</a>
<b>broadcast search</b>	crowd has to solve empirical problems	ideation problems with empirical provable solution like scientific problems	<a href="http://innocentive.com">innocentive.com</a>

# How It Works

InnoCentive's Open Innovation Marketplace connects organizations seeking solutions to important challenges they face with an unrivaled network of expert problem solvers – both experts from within your industry, but more importantly experts from outside of your industry that can offer diverse perspectives and fresh insight. Our proven Challenge Driven Innovation™ (CDI) model delivers real innovative solutions to complex problems and empowers Seeker organizations to uncover and harness new business opportunities.

## Seeker



LAUNCH

Your Challenge to our network of over 400,000 Solvers.



EVALUATE

The proposed solutions to your Challenge.

## Solver



AWARD

Prizes ONLY to solutions that meet your needs.



APPLY

The solution and any related IP to your organization.

# problem-focused categorization of crowdsourcing

type	how it works	ideal kind of problems	examples
knowledge discovery & management	crowd has to find & collect info into a common format	info gathering, reporting problems, creation of collective resources	seeclickfix.com
broadcast search	crowd has to solve empirical problems	ideation problems with empirical provable solution like scientific problems	innocentive.com
peer-vetted creative production	crowd creates & selects creative ideas	ideation problems where solutions are matter of taste such as design & aesthetics	threadless.com



SHOP

PARTICIPATE

BLOG

Search



**Threadless Select<sup>®</sup> SALE! UP TO 60% OFF OUR PREMIUM COLLECTION SHOP NOW >**

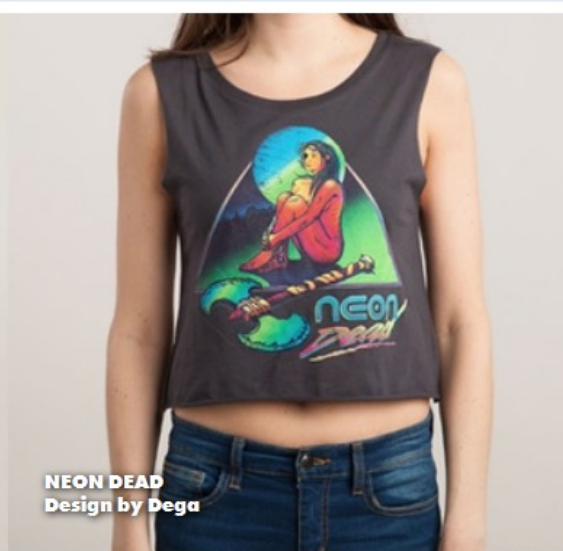


**CHECK OUT OUR SUMMER TANK STYLES!**





**THE BEACH**  
Design by Lim Heng Swee



**NEON DEAD**  
Design by Dega



**BURNING CIRCLES**  
Design by Florent Bodart

## THOUSANDS OF ARTISTS SUBMIT AWESOME DESIGNS EVERY DAY

Score 'em now and turn 'em into your new fave tees!



**NEW DESIGNS**

**THE MERMAID AND THE SAILOR**  
Design by Alex Solis



**REPRINTS**

**INTERNET**  
Design by Bill Mund



**SUBLIMATED TEES**

**THE MELTING**  
Design by Joe Van Wetering



# a problem-focused categorization

<b>type</b>	<b>how it works</b>	<b>ideal kind of problems</b>	<b>examples</b>
knowledge discovery & management	crowd has to find & collect info into a common format	info gathering, reporting problems, creation of collective resources	seeclickfix.com
broadcast search	crowd has to solve empirical problems	ideation problems with empirical provable solution like scientific problems	innocentive.com
peer-vetted creative production	crowd creates & selects creative ideas	ideation problems where solutions are matter of taste such as design & aesthetics	threadless.com
<b>distributed human intelligence tasking</b>	crowd analyzes large amounts of information	large-scale data analysis where human intelligence is more effective/efficient than machine intelligence	<b>mturk.com</b>

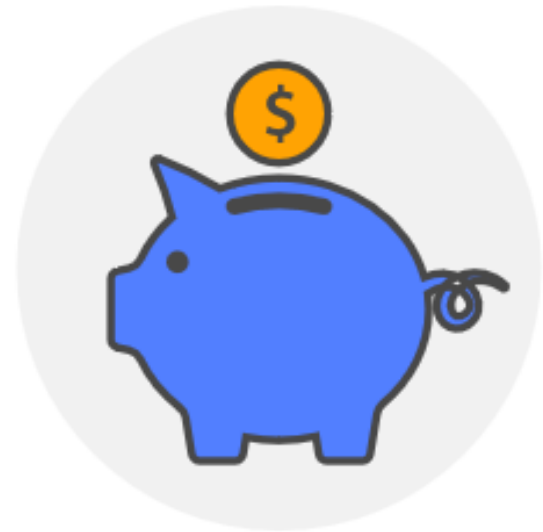
## Get Started with Amazon Mechanical Turk



### Create Tasks

Human intelligence through an API. Access a global, on-demand, 24/7 workforce.

Create a Requester account



### Make Money

Make money in your spare time. Get paid for completing simple tasks.

Create a Worker account

or

Your Account

HITS

Qualifications

284,789 HITS available now

All HITS | HITS Available To You | HITS Assigned To You

Find  containing

that pay at least \$

for which you are qualified  
 require Master Qualification



All HITS

1-10 of 3452 Results

Sort by:

[Show all details](#) | [Hide all details](#)

1 2 3 4 5 > [Next](#) >> [Last](#)

Mark all the cells in the 10 images

[View a HIT in this group](#)

Requester: [Cell Count Inc.](#)

HIT Expiration Date: May 24, 2014 (4 days 20 hours)

Reward: \$0.01

Time Allotted: 25 minutes

HITS Available: 22298

Inv. B. 2

[View a HIT in this group](#)

Requester: [rohzi0d](#)

HIT Expiration Date: Jun 19, 2014 (4 weeks 2 days)

Reward: \$0.00

Time Allotted: 48 minutes

HITS Available: 22155

Geo Result Relevance-Fri May 16 22:20:18 PDT 2014

[View a HIT in this group](#)

Requester: [Amazon Requester Inc.](#)

HIT Expiration Date: Jun 16, 2014 (3 weeks 6 days)

Reward: \$0.00

Time Allotted: 60 minutes

HITS Available: 22112

Extract purchased items from a shopping receipt

[View a HIT in this group](#)

Requester: [Jon Brelig](#)

HIT Expiration Date: May 27, 2014 (6 days 23 hours)

Reward: \$0.08

Time Allotted: 2 hours

HITS Available: 18714

Categorize these products from Amazon.com ALL (multiple questions per hit, qualified workers).

[View a HIT in this group](#)

Requester: [Amazon Requester Inc - browse classification](#)

HIT Expiration Date: May 30, 2014 (1 week 2 days)

Reward: \$0.06

Time Allotted: 10 minutes

HITS Available: 17555

Type the text from the images, carefully. Productivity and bonuses guaranteed.

[View a HIT in this group](#)

Requester: [CopyText Inc.](#)

HIT Expiration Date: May 25, 2014 (5 days 1 hour)

Reward: \$0.01

Time Allotted: 10 minutes

HITS Available: 13789

Your Account

HITS

Qualifications

284,166 HITS  
available now

All HITS | HITS Available To You | HITS Assigned To You

Find HITS containing

that pay at least \$ 0.00

for which you are qualified  
 require Master Qualification



Timer: 00:00:00 of 25 minutes

Want to work on this HIT?

Want to see other HITS?

Accept HIT

Skip HIT

Total Earned: Unavailable  
Total HITS Submitted: 0

Mark all the cells in the 10 images

Requester: Cell Count Inc.

Qualifications Required: None

Reward: \$0.01 per HIT

HITS Available: 22297

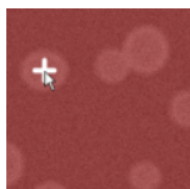
Duration: 25 minutes

## Identifying cells

**Instructions:** the task consists in identifying cells by clicking at the center of each one of them, as the following graphical example shows.



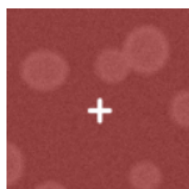
No cell clicked



One cell clicked



Adjacent cells  
should be clicked

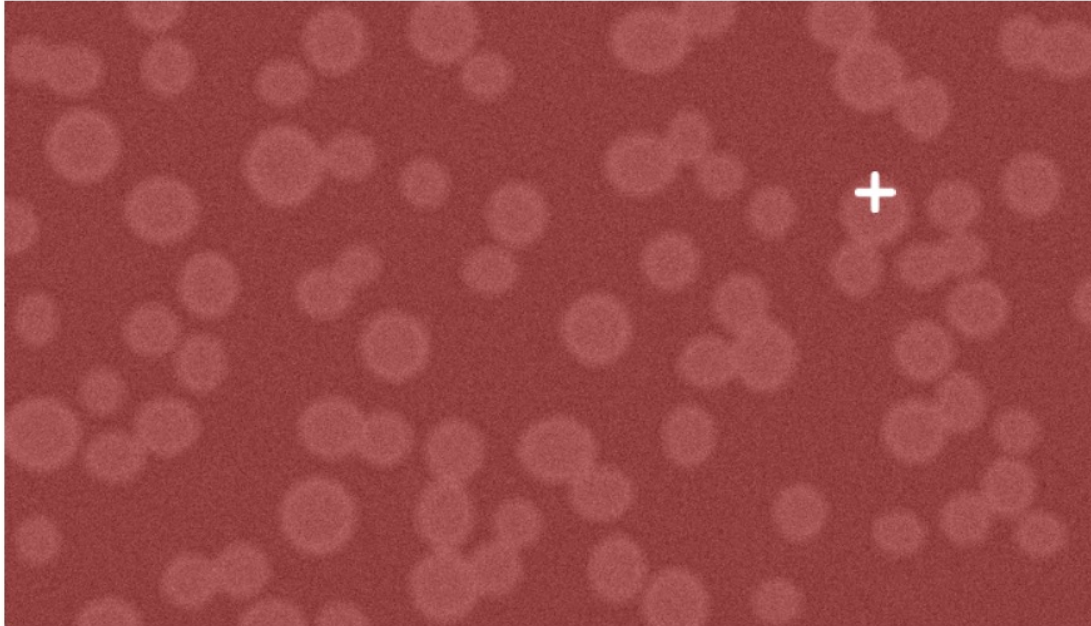


Wrong tagging

Following the instructions above, you will successfully identify cells in the image. When you are done, you can click on the check mark to request an additional image with cells. The check mark is located below the image with the cells. After completing 10 images, you can submit your work and get your reward. Mind that you can only click on the check mark when the wait time is over, and that will happen only after a few seconds have passed since you were shown the image with cells. You can only participate in this HIT once. Additional times will be rejected.



Following the instructions above, you will successfully identify cells in the image. When you are done, you can click on the check mark to request an additional image with cells. The check mark is located below the image with the cells. After completing 10 images, you can submit your work and get your reward. Mind that you can only click on the check mark when the wait time is over, and that will happen only after a few seconds have passed since you were shown the image with cells. You can only participate in this HIT once. Additional times will be rejected.



Cell image number 1 of 10

**Click here to submit work and  
advance to the next image of  
cells**



Want to work on this HIT?    Want to see other HITs?

Accept HIT

Skip HIT

# Data For Everyone

Here are some of our favorite open datasets created on the Figure Eight platform. They're free for any and everyone to download.

Each of the following respective data sets are licensed under a Creative Commons Attribution 4.0 International License. By taking the action of downloading any data set(s), you hereby agree to the associated license terms. Each data set is provided "as is" and any expressed or implied warranties, including, but not limited to, the implied warranties of merchantability and fitness for a particular purpose are disclaimed.

<https://www.figure-eight.com/data-for-everyone/>  
(formerly CrowdFlower, bought by Appen in 2019)



Solutions ▾

Industries ▾

Company ▾

Resources ▾

Jobs

# Confidence to Deploy AI with World-Class Training Data

Artificial intelligence will improve the world. To succeed, companies need reliable training data that goes into expertly crafted models. The global community will power world-class AI.

<https://appen.com/>

# connections between crowdsourcing & social media

social media can be seen as **implicit** crowdsourcing

crowdsourcing uses **social media as channels**  
+ creative collective videos on YouTube

social media content is **curated** via crowdwork

crowdsourcing & social media face **similar challenges**  
+ motivate / incentivize users & workers  
+ community management

crowdwork involves **online social interaction**  
+ rate / comment on other people's contributions  
+ participate in communities of crowdworkers



# this lecture

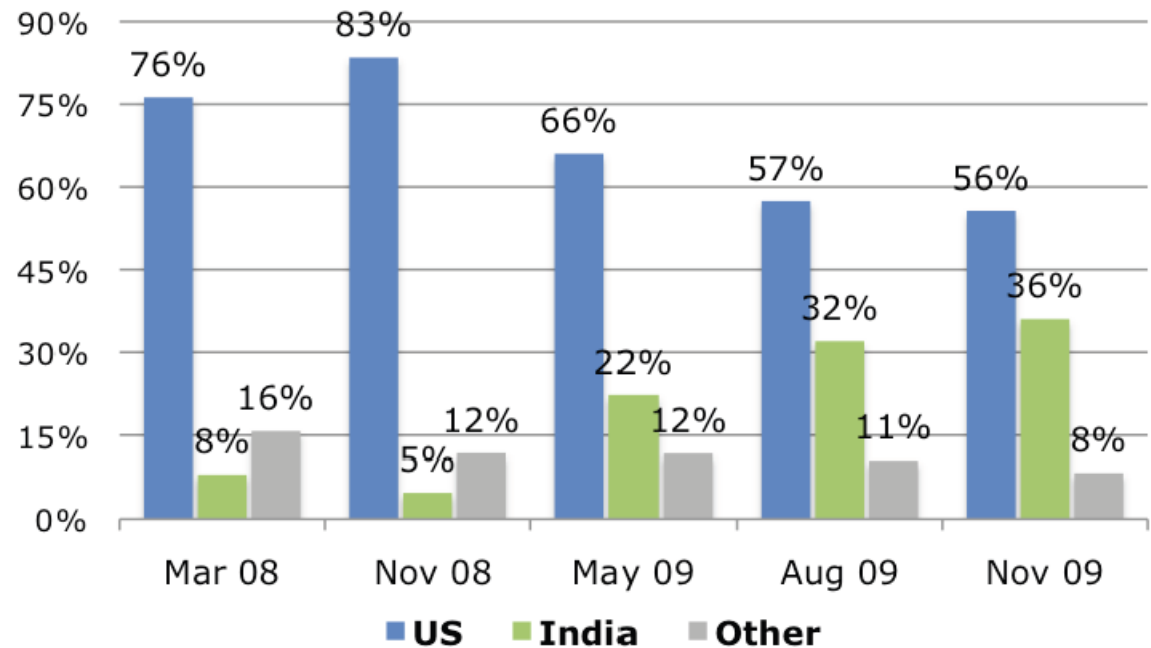
1. introduction
2. categorization of crowdsourcing systems
3. understanding crowdsourcing work
  - mechanical turk workers
  - crowdlabeling models
4. designing crowdsourcing systems
5. crowdsourcing as labor
6. citizen science

**3.**  
**understanding**  
**crowdsourcing work**

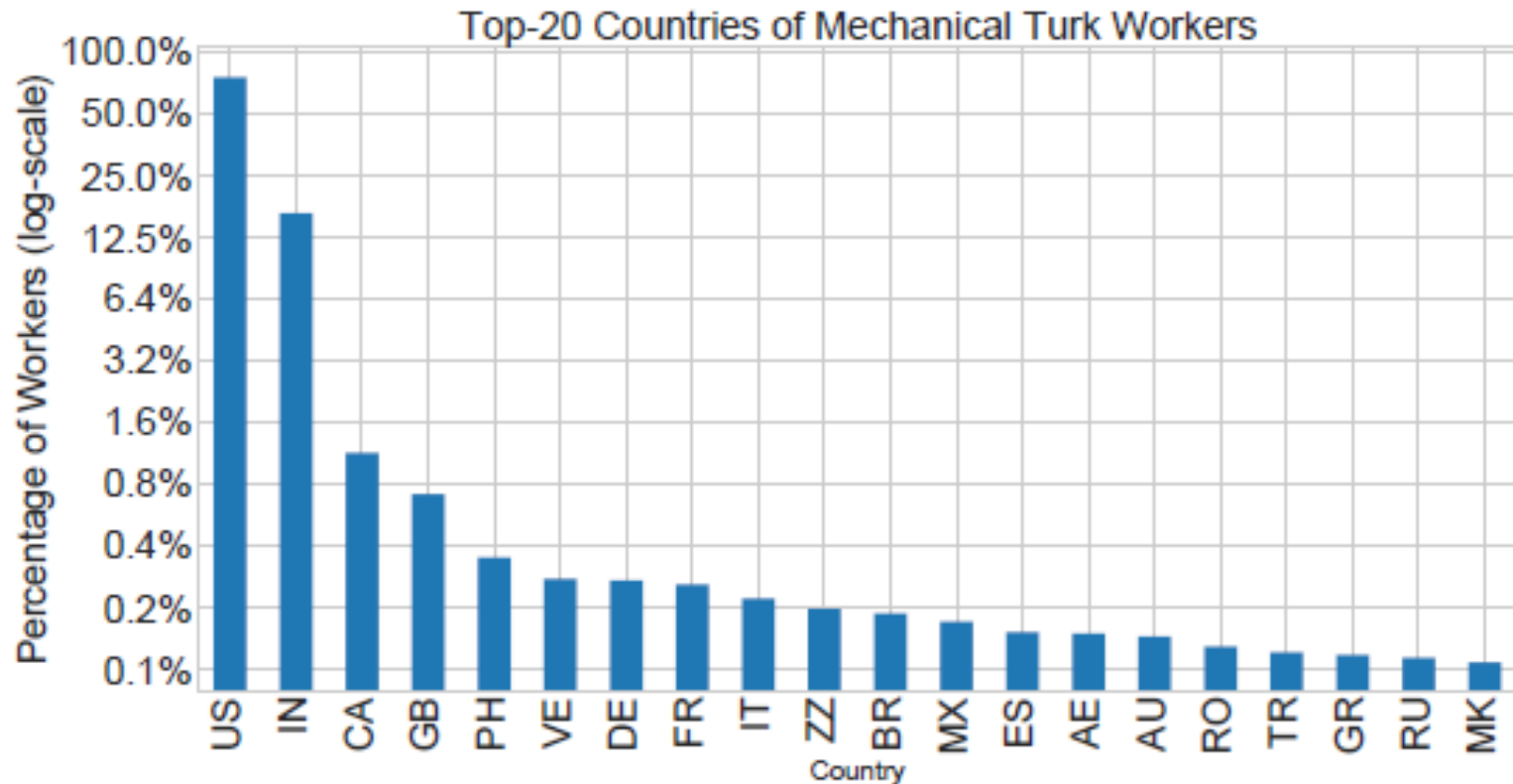
**who are the crowdworkers?**

# who are the crowdworkers?

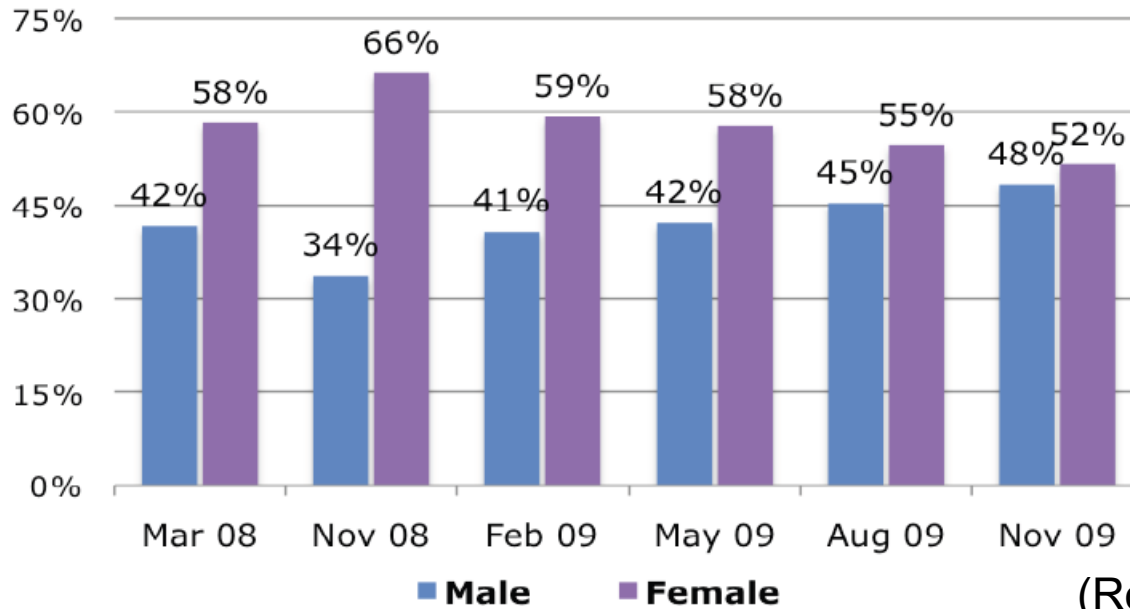
Survey Date	Sample Size
Mar. 2008 [10]	<i>n</i> = 300
Nov. 2008 [11]	<i>n</i> = 1010
Feb. 2009	<i>n</i> = 878
May 2009	<i>n</i> = 512
Aug. 2009	<i>n</i> = 578
Nov. 2009	<i>n</i> = 733



Survey data: 39,461 unique workers collected on 859 days (03.2015-07.2017)



**Figure 1: The top-20 countries of origin for Mechanical Turk workers. Most of the workers are from the USA (75%), with India (16%) being second, followed by Canada (1.1%), Great Britain (0.7%), Philippines (0.35%), and Germany (0.27%).**



(Ross, 2010)

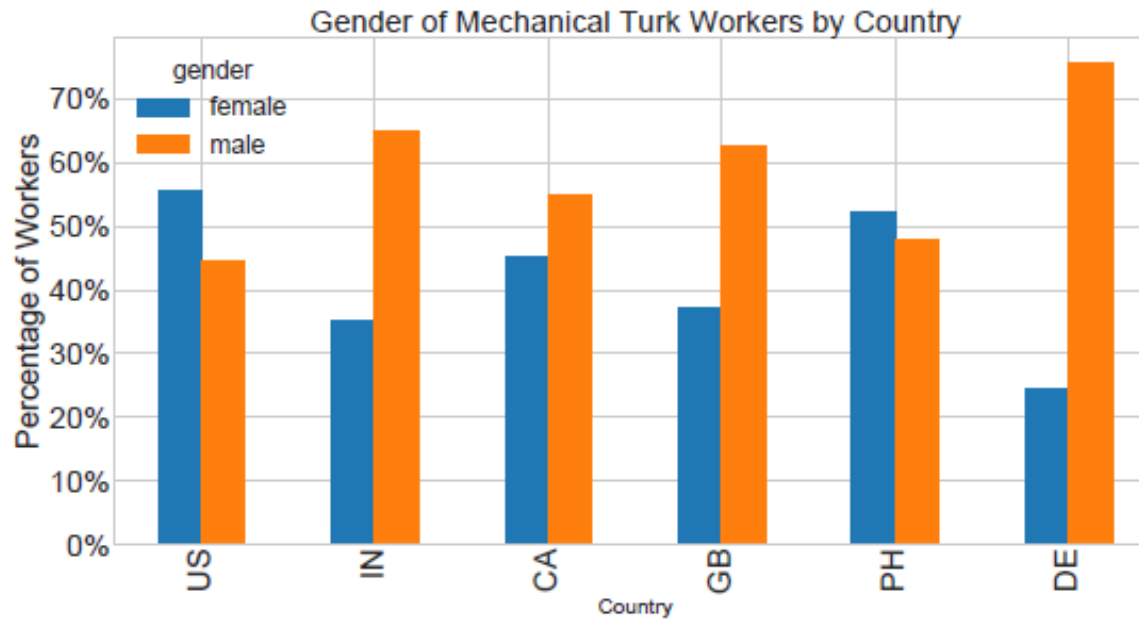
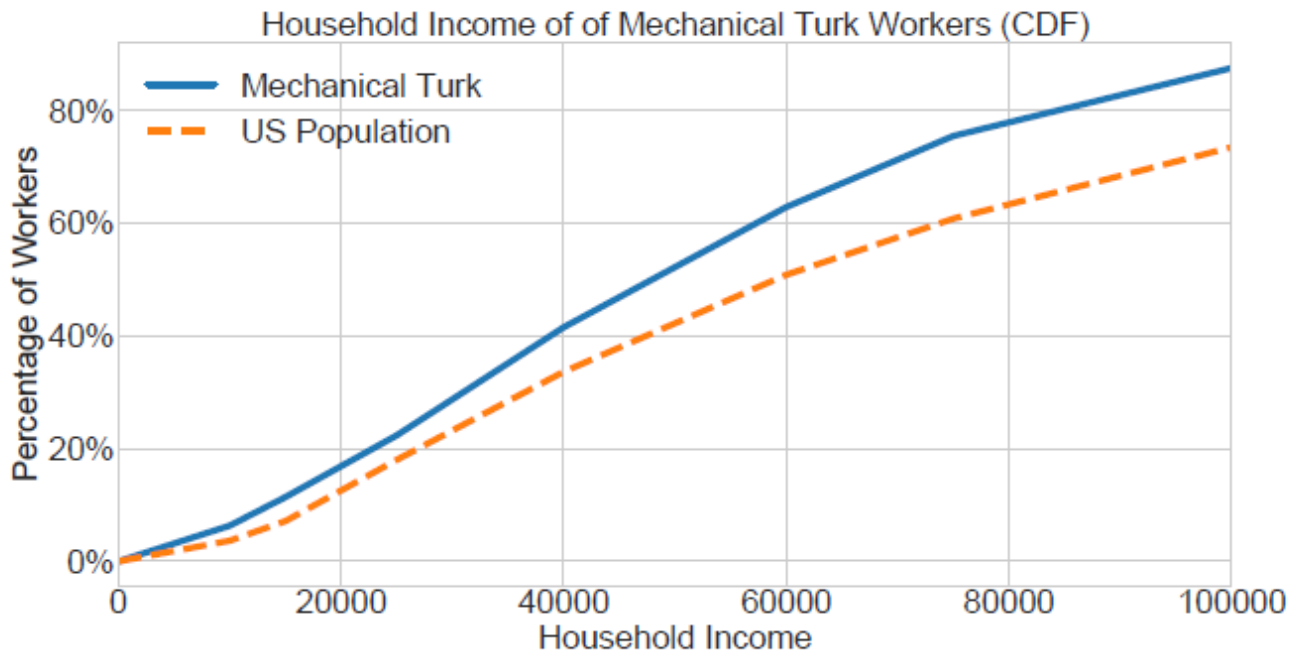
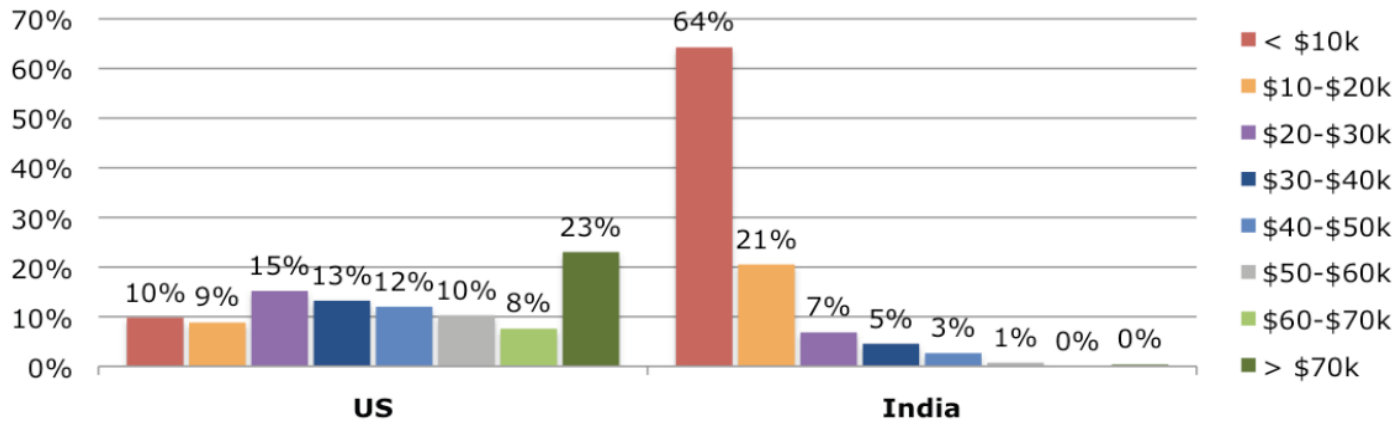


Figure 3: Gender breakdown across countries.

(Difallah, 2018)



**Figure 6: Household Income.** (Difallah, 2018)



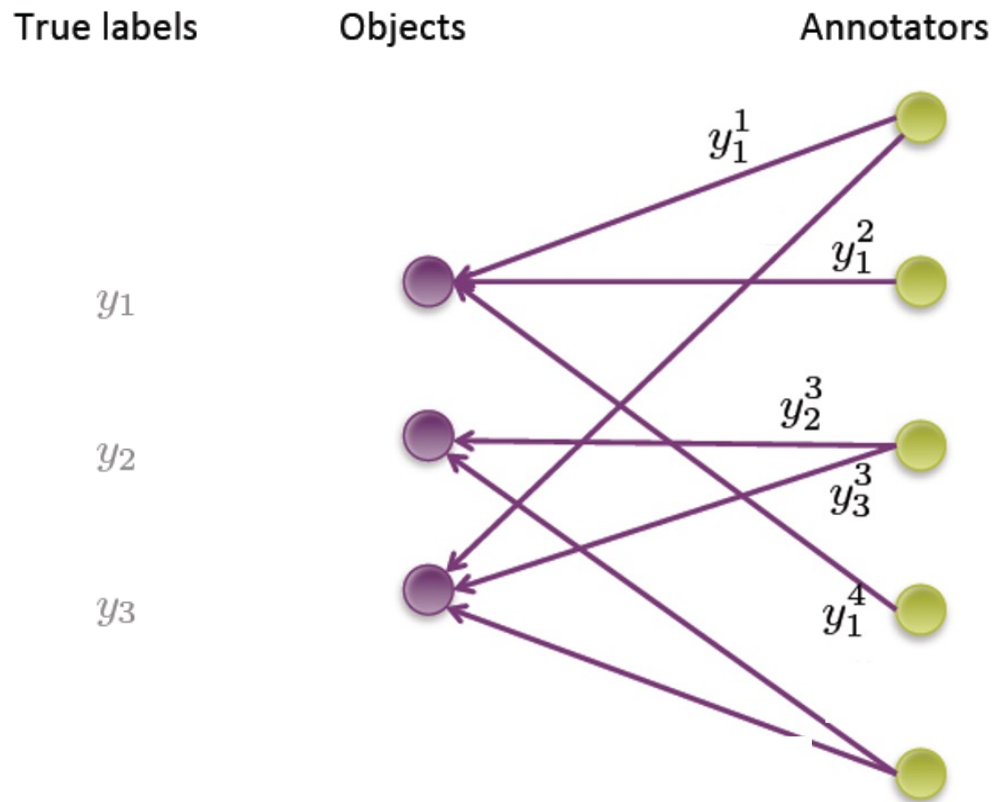
**annual household income by country (USD)** (Ross, 2010)

# **crowdlabeling models**



# crowdlabeling modeling

- A set of objects to annotate  $i = 1, \dots, I$
- A set of annotators  $j = 1, \dots, J$



## issues:

- 1 - annotators might not be of same **quality**
- 2 - objects might not have same **difficulty**
- 3 - ground-truth might not **exist** at all

## Annotations

$$y_i^j \in \mathcal{L}$$

Binary  $|\mathcal{L}| = 2$

Multi-class  $|\mathcal{L}| > 2$

# aggregating judgments

- Majority voting (baseline)
  - For each object, assign the label that received the largest number of votes
- Aggregating annotations
  - [Dawid and Skene, 1979]
  - [Snow et al., 2008]
  - [Whitehill et al., 2009]
- Aggregating and learning a classifier
  - [Sheng et al., 2008]
  - [Donmez et al., 2009]
  - [Raykar et al., 2010]

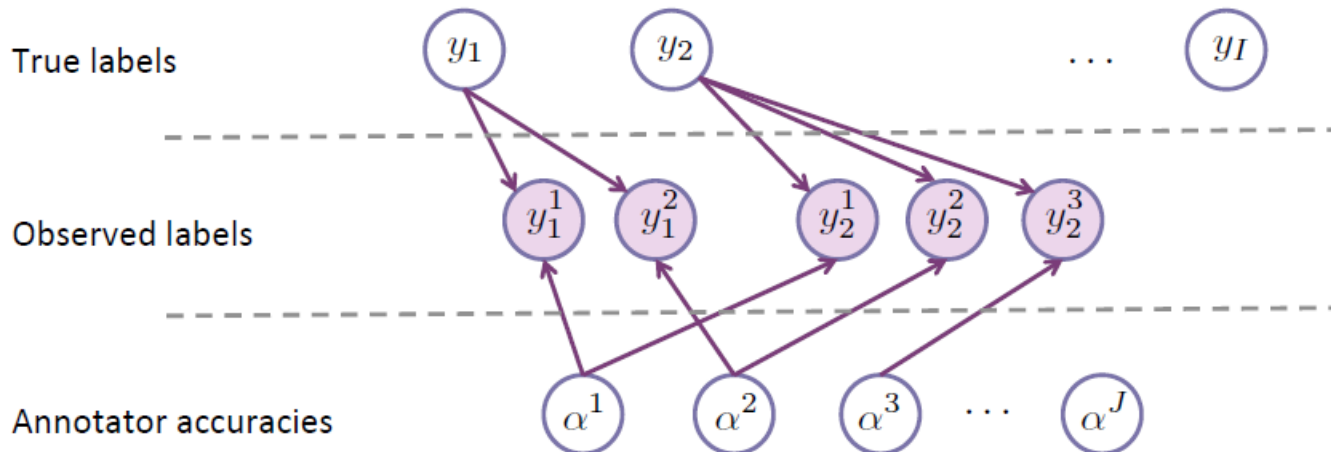
# the basic idea

(Dawid & Skene, 1979)

i: object  
j: annotator

- For simplicity, consider the case with binary labels  $y_i^j \in \{0, 1\}$
- Each annotator is characterized by the (unknown) error rates

$$P(y_i^j = 1 | y_i = 1) = \alpha_1^j \quad \text{True positive rate} \quad (\text{a.k.a. sensitivity})$$
$$P(y_i^j = 0 | y_i = 0) = \alpha_0^j \quad \text{True negative rate} \quad (\text{a.k.a. specificity})$$



## the basic idea (2)

(Dawid & Skene, 1979)

- The likelihood function of the parameters  $\{\alpha_1, \alpha_0\}$  given the observations  $\mathcal{D} = \{y_i^1, \dots, y_i^J\}_{i=1}^I$  is

$$P(\mathcal{D}|\alpha_1, \alpha_0) = \prod_{i=1}^I P(y_i^1, \dots, y_i^J | \alpha_1, \alpha_0)$$

- The parameters are found by maximizing the log-likelihood function

$$\{\hat{\alpha}_1, \hat{\alpha}_0\} = \arg \max_{\theta} \log P(\mathcal{D}|\theta) \quad \theta = \{\alpha_1, \alpha_2\}$$

- The solution is based on Expectation-Maximization

Estimates of the true labels are produced in the E-step

Estimates of the alpha parameters are produced in the M-step

The learned parameters can be used to

- (1) remove inconsistent annotators in a crowdsourced task
- (2) better inform the process

# this lecture

1. introduction
2. categorization of crowdsourcing systems
3. understanding crowdsourcing work
  - mechanical turk workers
  - crowdlabeling models
4. designing crowdsourcing systems
5. crowdsourcing as labor
6. citizen science

**4.**  
**designing crowdsourcing  
systems**

# motivations: why do people participate?

## intrinsic motivations:

“doing an activity for its inherent satisfactions rather than for some separable consequence”

- + fun
- + challenge

## extrinsic motivations:

“an activity that is done in order to attain some separable outcome”

- + financial reward
- + social pressure

extrinsic motivations tend to undermine intrinsic ones

# why do crowds participate?

- “to earn money
- to develop creative skills
- to network with other creative professionals
- to build a portfolio for future employers
- to challenge oneself to solve a tough problem
- to socialize and make friends
- to pass time when bored
- to contribute to a larger project of common interest
- to share with others
- to have fun”



# assume you have a job for MTurk...

Your Account

HITs

Qualifications

284,789 HITs  
available now

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

Find HITs containing

that pay at least \$ 0.00

for which you are qualified

require Master Qualification



## All HITs

1-10 of 3452 Results

Sort by: HITs Available (most first)

Show all details | Hide all details

1 2 3 4 5 > Next >> Last

Mark all the cells in the 10 images

[View a HIT in this group](#)

Requester: [Cell Count Inc.](#)

HIT Expiration Date: May 24, 2014 (4 days 20 hours)

Reward: \$0.01

Time Allotted: 25 minutes

HITs Available: 22298

Inv\_B\_2

[View a HIT in this group](#)

Requester: [rohzt0d](#)

HIT Expiration Date: Jun 19, 2014 (4 weeks 2 days)

Reward: \$0.00

Time Allotted: 48 minutes

HITs Available: 22155

Geo Result Relevance-Fri May 16 22:20:18 PDT 2014

[View a HIT in this group](#)

Requester: [Amazon Requester Inc.](#)

HIT Expiration Date: Jun 16, 2014 (3 weeks 6 days)

Reward: \$0.00

Time Allotted: 60 minutes

HITs Available: 22112

Extract purchased items from a shopping receipt

[View a HIT in this group](#)

Requester: [Jon Brelig](#)

HIT Expiration Date: May 27, 2014 (6 days 23 hours)

Reward: \$0.08

Time Allotted: 2 hours

HITs Available: 18714

Categorize these products from Amazon.com ALL (multiple questions per hit, qualified workers).

[View a HIT in this group](#)

Requester: [Amazon Requester Inc - browse classification](#)

HIT Expiration Date: May 30, 2014 (1 week 2 days)

Reward: \$0.06

Time Allotted: 10 minutes

HITs Available: 17555

Type the text from the images, carefully. Productivity and bonuses guaranteed.

[View a HIT in this group](#)

Requester: [CopyText Inc.](#)

HIT Expiration Date: May 25, 2014 (5 days 1 hour)

Reward: \$0.01

Time Allotted: 10 minutes

HITs Available: 13789

# aspects to consider

**ownership:** host your data on MTurk or not?

**biases:** MTurkers are not a fair sample of world population

**engagement:** is your task fun? entertaining?

**geography:** limit workers to specific country/region?

**complexity:** how difficult or time-involved is your task?

**spam:** use tests to reduce spammers in your task

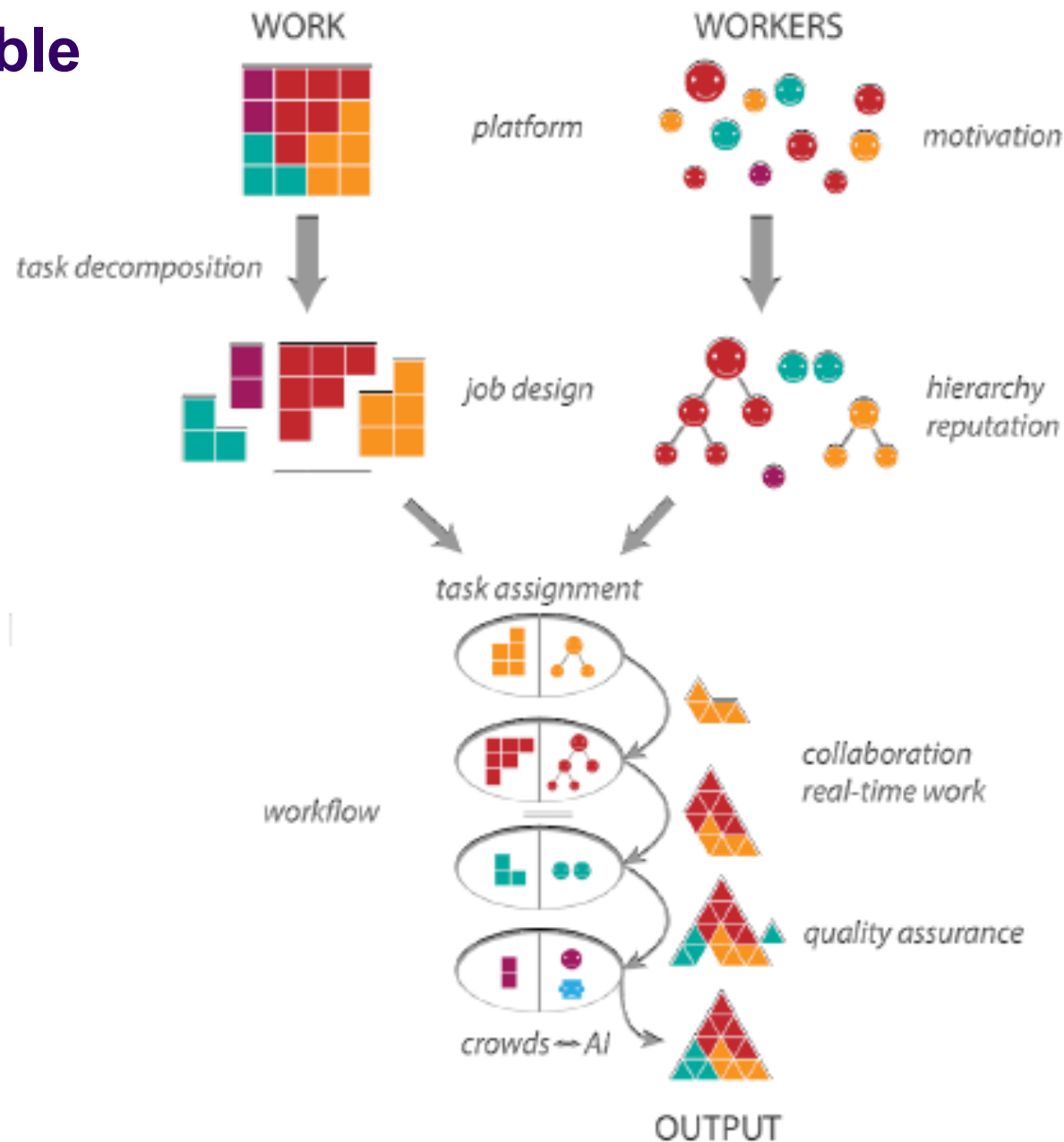
**payment:** pay workers a fair rate

**qualifications:** engage workers of recognized quality

**interaction:** create community with workers based on trust

**quality:** use model to reduce effect of poor or spam responses

# a possible model



A. Kittur, J. Nickerson, M. Bernstein, E. M. Gerber, A. Shaw, J. Zimmerman, M. Lease, J. Horton, The future of crowd work, in Proc. ACM CSCW 2013.

# this lecture

1. introduction
2. categorization of crowdsourcing systems
3. understanding crowdsourcing work
  - mechanical turk workers
  - crowdlabeling models
4. designing crowdsourcing systems
5. crowdsourcing as labor
6. citizen science

# **5. crowdsourcing as labor**



# what are the issues?

## **intellectual property & copyright**

who owns my idea if it wins?

## **unfair business practices**

crowdsourcing used for manipulation

example: fake reviews of products & services



## **labor rights**

fair payment

stressful jobs

transnational jobs

credit (cc): <https://www.flickr.com/photos/61115981@N06/7476310372>

# guidelines to interact with crowdworkers in research

“Crowdworkers relate to participation in research primarily as *workers*

Pay workers at least minimum wage at your location

Remember that you are interacting with human beings

Respond quickly, clearly, concisely, and respectfully to worker questions

Learn from workers”







THE END OF THE JOB AND  
THE FUTURE OF WORK

# Gigged

SARAH KESSLER

Copyrighted Material

2018

How to  
Stop Silicon Valley  
from Building a  
New Global Underclass

# GHOST

Mary L. Gray and Siddharth Suri

# WORK

2019



► **World Employment  
and Social Outlook**

The role of digital labour  
platforms in transforming  
the world of work

**2021**

“Digital labor platforms are now part of our everyday lives.

This new business model allows platforms to organize work without having to invest in capital assets or to hire employees.

Instead, they mediate between the workers who perform the tasks and clients, and manage the entire work process with algorithms.

Workers often struggle to find sufficient well-paid work to earn a decent income, creating a danger of working poverty. Many do not have access to social protection, and are frequently unable to engage in the collective bargaining that would allow them to have these and other issues addressed.”

# content moderation: social media & crowdsourcing

Social media companies engage workers to manually flag content that is objectionable or does not follow their terms of service

Issues: adversarial uses of social media, circulation of unethical and illegal content, use of manual labor to train AI systems

Who decides what is flagged and filtered?

Worker conditions: income, benefits, psychological risks

One component of a “massive interlinked chain of extractive processes” (Whittaker, 2018)

# this lecture

1. introduction
2. categorization of crowdsourcing systems
3. understanding crowdsourcing work
  - mechanical turk workers
  - crowdlabeling models
4. designing crowdsourcing systems
5. crowdsourcing as labor
6. citizen science

**6.**  
**citizen science & citizen observatories**

# citizen science

The screenshot shows the top navigation bar of the Foldit website. On the left is the Foldit logo, which includes a white ribbon icon and the text "foldit BETA Solve Puzzles for Science". The time "10:30:23 GMT" is displayed in the top right of the bar. A horizontal menu contains the following items: PUZZLES, BLOG, CATEGORIES, FEEDBACK, GROUPS, FORUM, PLAYERS, WIKI, FAQ, RECIPES, ABOUT, CONTESTS, and CREDITS.

The main content area features a large 3D protein structure rendered in green and blue. A semi-transparent black box with white text is overlaid on the structure, reading: "Click to learn how you contribute to science by playing Foldit." Below the protein structure is a dark blue banner for "NANOCRAFTER" with the text "Try our new scientific discovery game! Be creative and build extraordinary tiny machines!"

On the right side of the page, there is a "GET STARTED: DOWNLOAD" section. It contains three buttons for "Win Beta", "Mac Beta", and "Linux Beta". Below each button are the operating system names and versions: "Windows (Vista/7/8)", "OSX (10.7 or later)", and "Linux (64-bit)". Below these buttons are three links: "Are you new to Foldit? Click here.", "Are you a student? Click here.", and "Are you an educator? Click here.".

Below the download section is a "SEARCH" section with a search input field, a "Google Search" button, and a checked checkbox labeled "Only search fold.it".

At the bottom right of the page is a "RECOMMEND FOLDIT" button.

"*Foldit* attempts to predict the structure of a protein by taking advantage of humans' puzzle-solving intuitions and having people play competitively to *fold* the best proteins,"

Galaxy Zoo is a **Zooniverse** project.

The Zooniverse is a collection of web-based Citizen Science projects that use the efforts and abilities of volunteers to help researchers deal with the flood of data that confronts them.

Our Projects

We currently have dozens of **projects** on subjects ranging from astronomy, to climatology, to biology, to humanities.

CLASSIFY

STORY

SCIENCE



DISCUSS

PROFILE

LANGUAGE



# Few have witnessed what you're about to see

Experience a privileged glimpse of the distant universe as observed by the SDSS and the CTIO.

## Classify Galaxies

To understand how galaxies formed we need your help to classify them according to their shapes. If you're quick, you may even be the first person to see the galaxies you're asked to classify.

[Begin Classifying](#)



<https://www.galaxyzoo.org/>

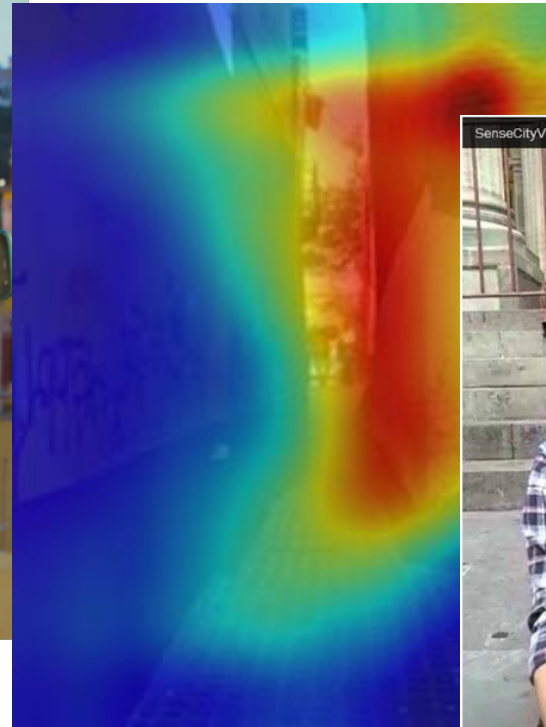
# SenseCityVity: participatory crowdsensing with youth



Co-Design



Collect



Analyze

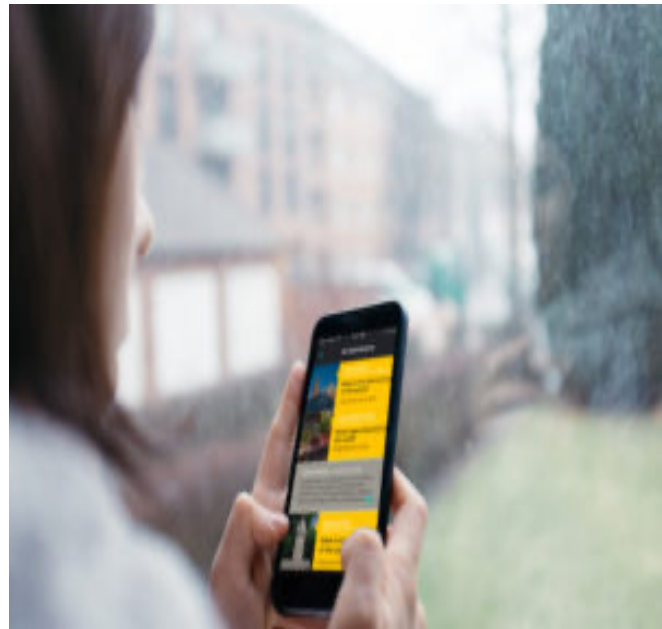


Create

S. Ruiz-Correa, D. Santani, B. Ramirez Salazar, I. Ruiz Correa, F. Alba Rendon-Huerta, C. Olmos Carrillo, B. C. Sandoval Mexicano, A. H. Arcos Garcia, R. Hasimoto Beltran, D. Gatica-Perez, SenseCityVity: Mobile Sensing, Urban Awareness, and Collective Action in Mexico, IEEE Pervasive Computing, Apr.-Jun. 2017.



# Corona Citizen Science: COVID-19 experiences in Switzerland



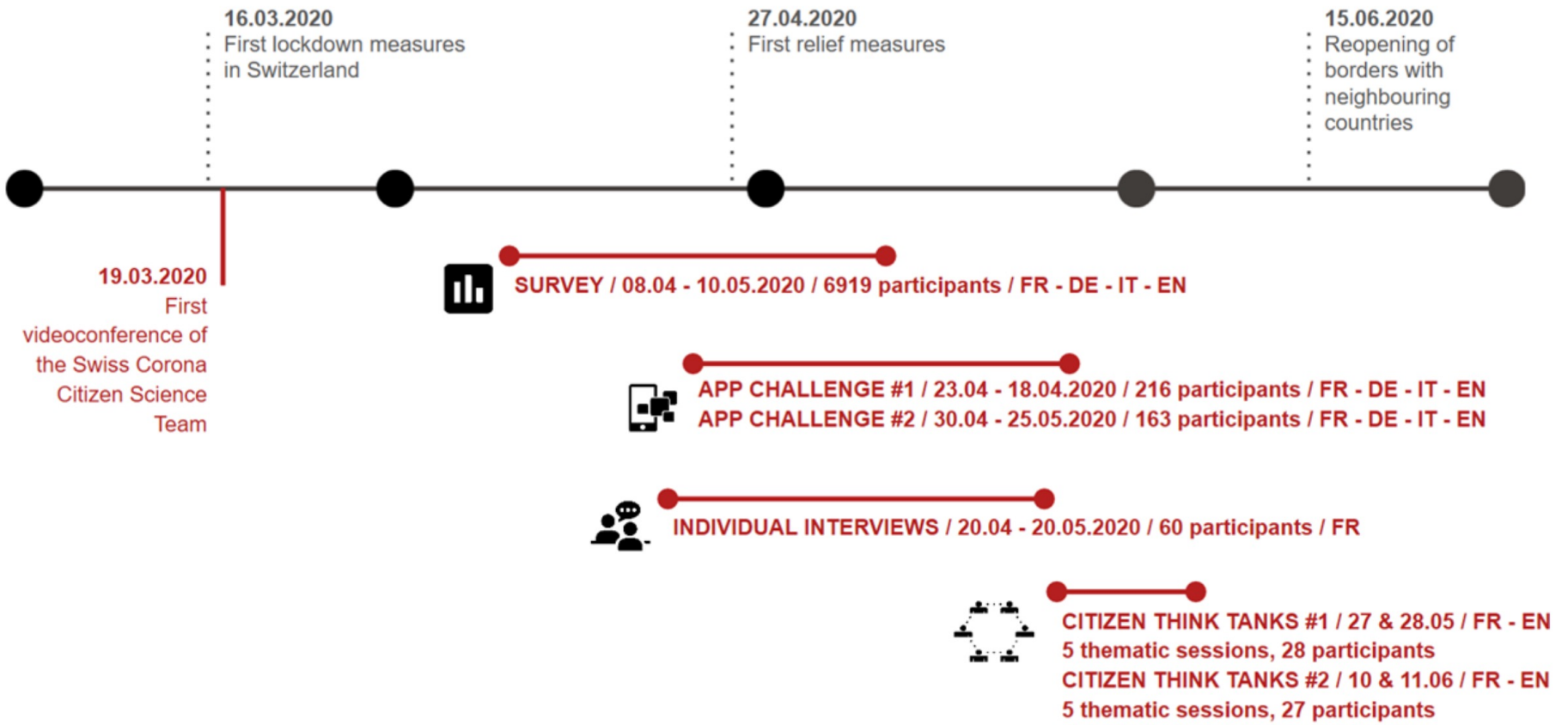
Data collected  
via mobile app &  
online survey



Experience of  
the pandemic via  
text & photos



Perspectives  
discussed via online  
focus groups



<https://coronacitizenscience.ch>



# 1. Overall perception of the situation



Most participants had a **positive mood overall**—slightly good or very good mood (Q: *How do you feel today?*) when undertaking the challenges in the crowdsourcing app. Pertaining to their perceptions of the COVID-19 crisis (Q: *How do you feel about the course the COVID-19 crisis takes in Switzerland?*), the most common sentiment was slight optimism, but around one third of participants were undecided.

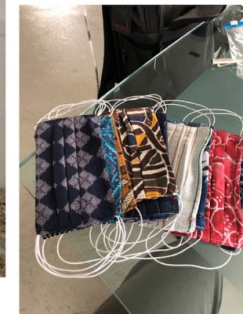


**A selection of curated images:** *What is the most significant change that has happened in your daily life due to the COVID-19 crisis during the past week?*

In these uncertain and difficult times, the approach to personal hygiene has changed for many people. Beyond more frequent handwashing, numerous challenge participants adopted additional measures to protect and safeguard themselves and their household members. These actions include shopping online for daily necessities to minimise contact with others, as well as the active use of protective masks and disinfection products to keep the risk of infection low.



**ACTION**



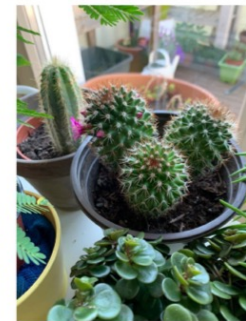


**A selection of curated images:** *What is the most significant change that has happened in your daily life due to the COVID-19 crisis during the past week?*

The confinement period in Switzerland greatly restricted the movement of the population and, for many, limited access to nature and the outdoors. This change sparked an interest in gardening for a large number of challenge participants; from aromatic herbs to ornamental succulents, tending to their greenery of choice provided both the novice and seasoned home gardener a little escape to nature right within their homes.



## NATURE





## THE CITIZEN THINK TANKS

**16.03.2020**

- First lockdown measures in Switzerland

**27.04.2020**

- First relief measures

**15.06.2020**

- Reopening of borders with neighbouring countries



**CITIZEN THINK TANKS #1 / 27 & 28.05 / FR - EN**  
**5 thematic sessions, 28 participants**  
**CITIZEN THINK TANKS #2 / 10 & 11.06 / FR - EN**  
**5 thematic sessions, 27 participants**

**Aims:** The Citizen Think Tanks (CTTs) of the Swiss Corona Citizen Science project were spaces for mutual learning and for designing possible post-Covid-19 futures. In a series of interactive online events, citizens collaborated with scientists from EPFL who hosted the CTTs. The topics of the CTTs emerged from the survey and mobile crowdsourcing. In 5 CTTs, desirable futures were envisioned for housing, mobility, local economy, contact tracing and digital governance and tourism. At the end of the national survey as well as Challenge 2 in the app, participants could express their interest in participating in interactive dialogue sessions: the Citizen Think Tanks.



**questions?**