

NETS 213: CROWDSOURCING
AND HUMAN COMPUTATION

Crowdsourcing and Human Computation

Professor Chris Callison-Burch

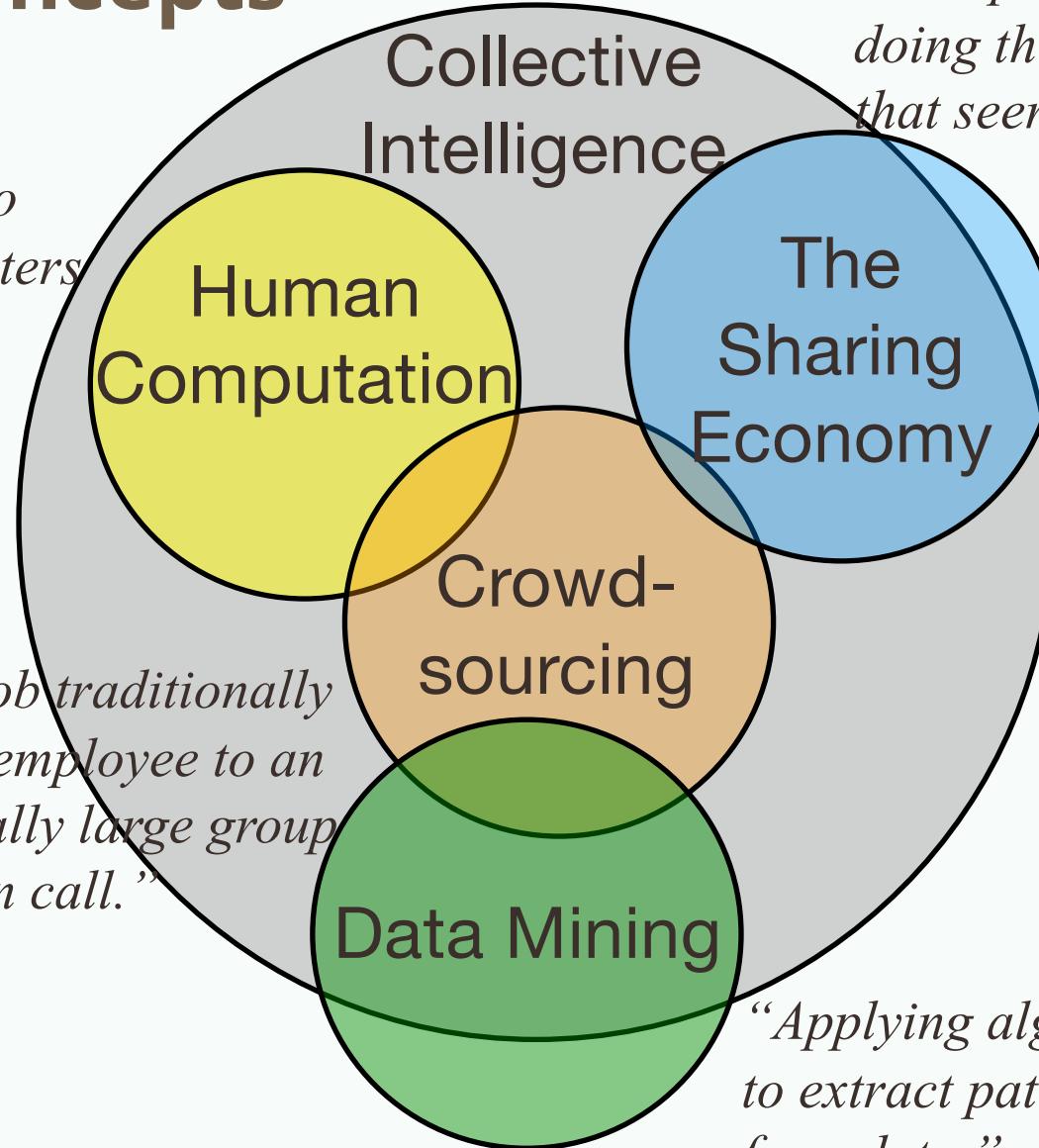
<http://crowdsourcing-class.org>



Inter-related concepts

“A paradigm for utilizing human processing power to solve problems that computers cannot yet solve.”

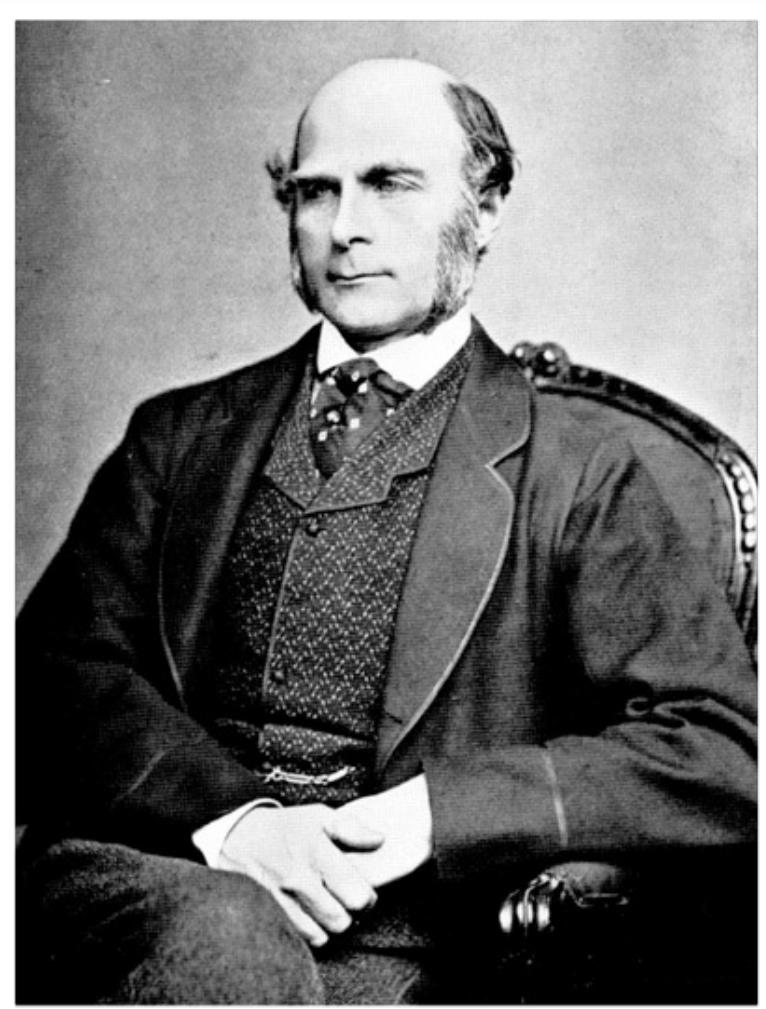
“Outsourcing a job traditionally performed by an employee to an undefined, generally large group of people via open call.”



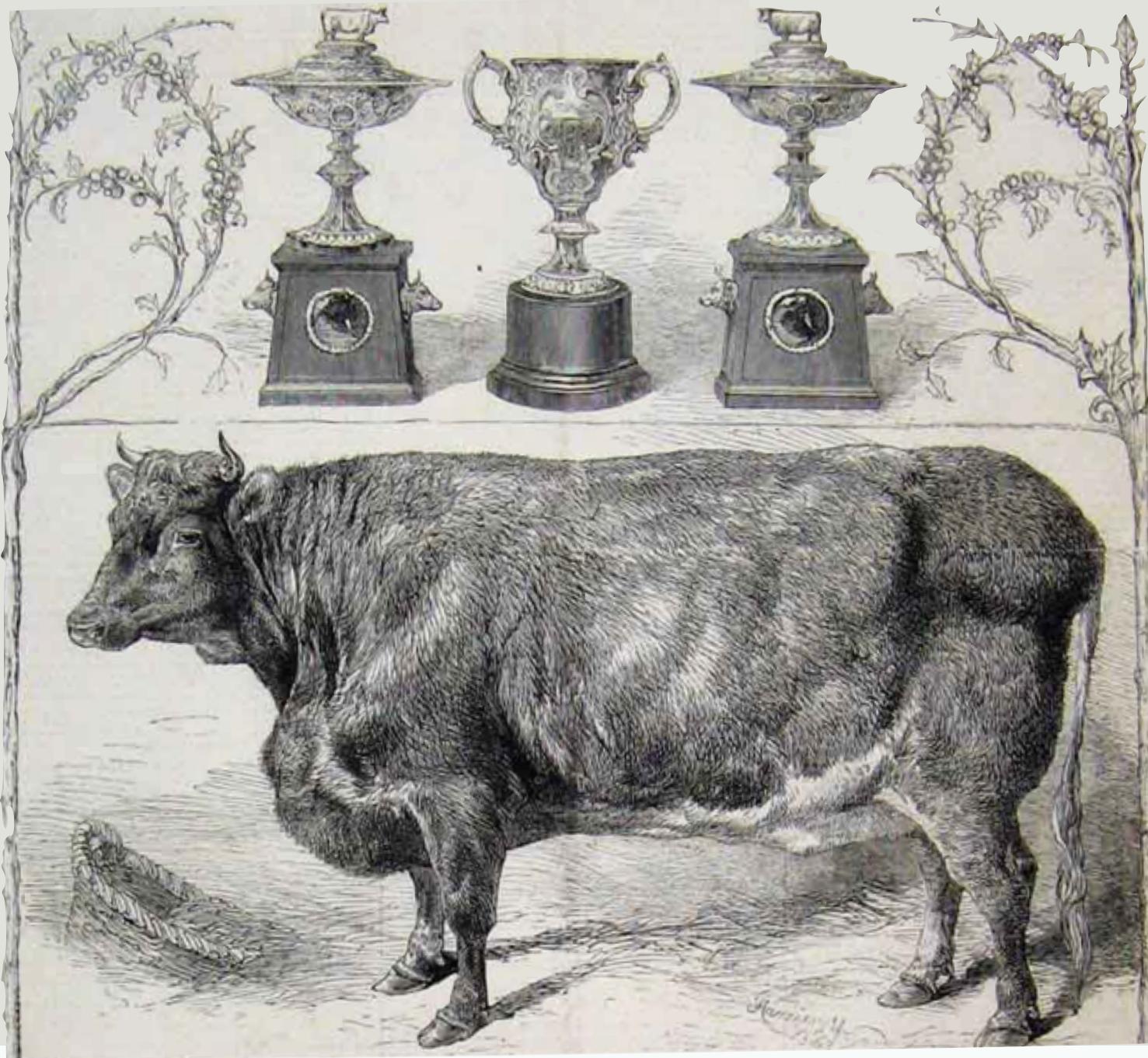
“Groups of individuals doing things collectively that seem intelligent”

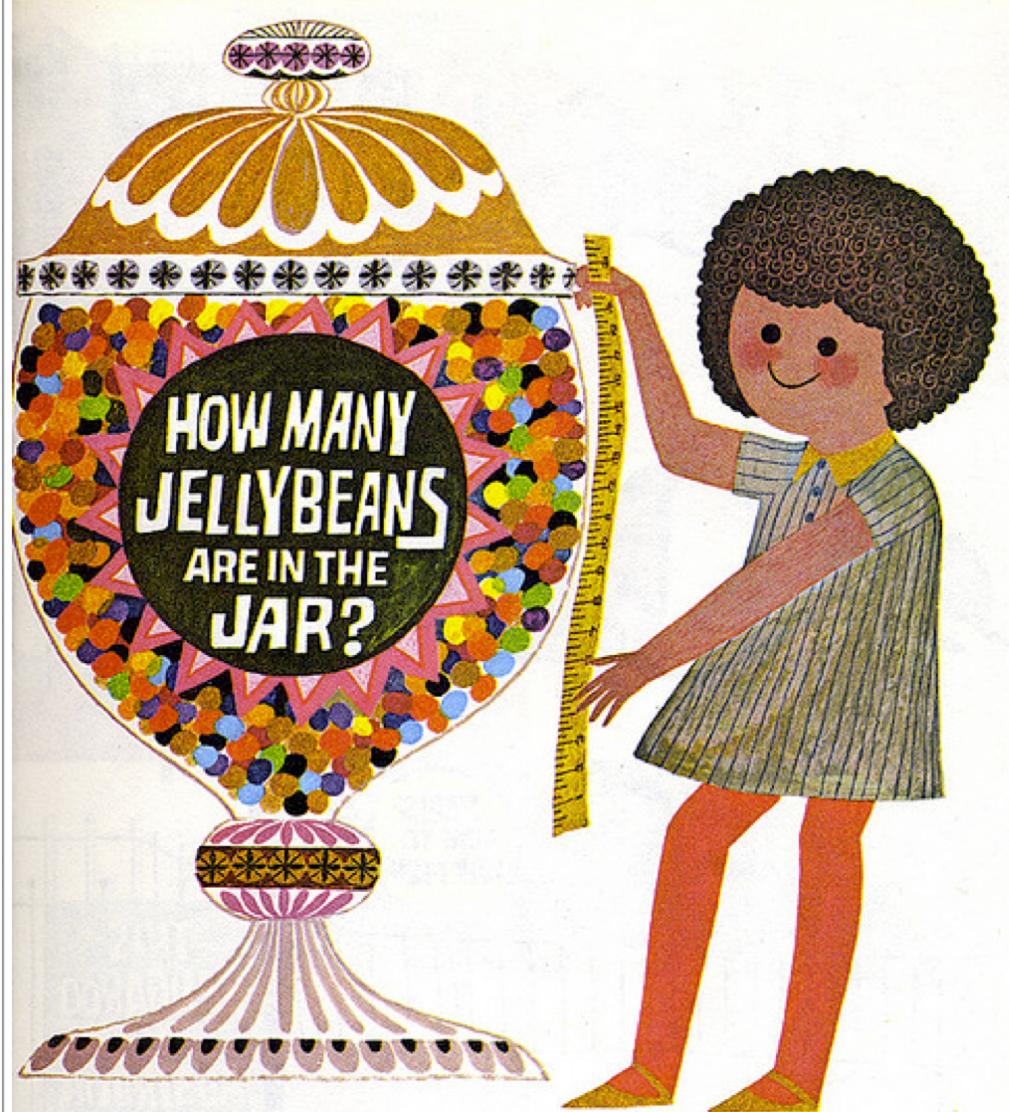
“An economic system in which assets or services are shared between private individuals, either for free or for a fee, typically by means of the Internet.”

“Applying algorithms to extract patterns from data.”



Francis Galton





If you guess the right number of jellybeans,
you may win a prize.

Collective Intelligence?



Groupthink

Groupthink is a psychological phenomenon that occurs within a group of people in which the desire for harmony or conformity in the group results in an irrational or dysfunctional decision-making outcome

Collective Intelligence?



Mob Mentality

Mob mentality, also called as herd mentality, describes how humans adopt behaviors, buy merchandise, and follow trends based on their circle of influence. It explains how one's point of view can be easily altered by those around them.

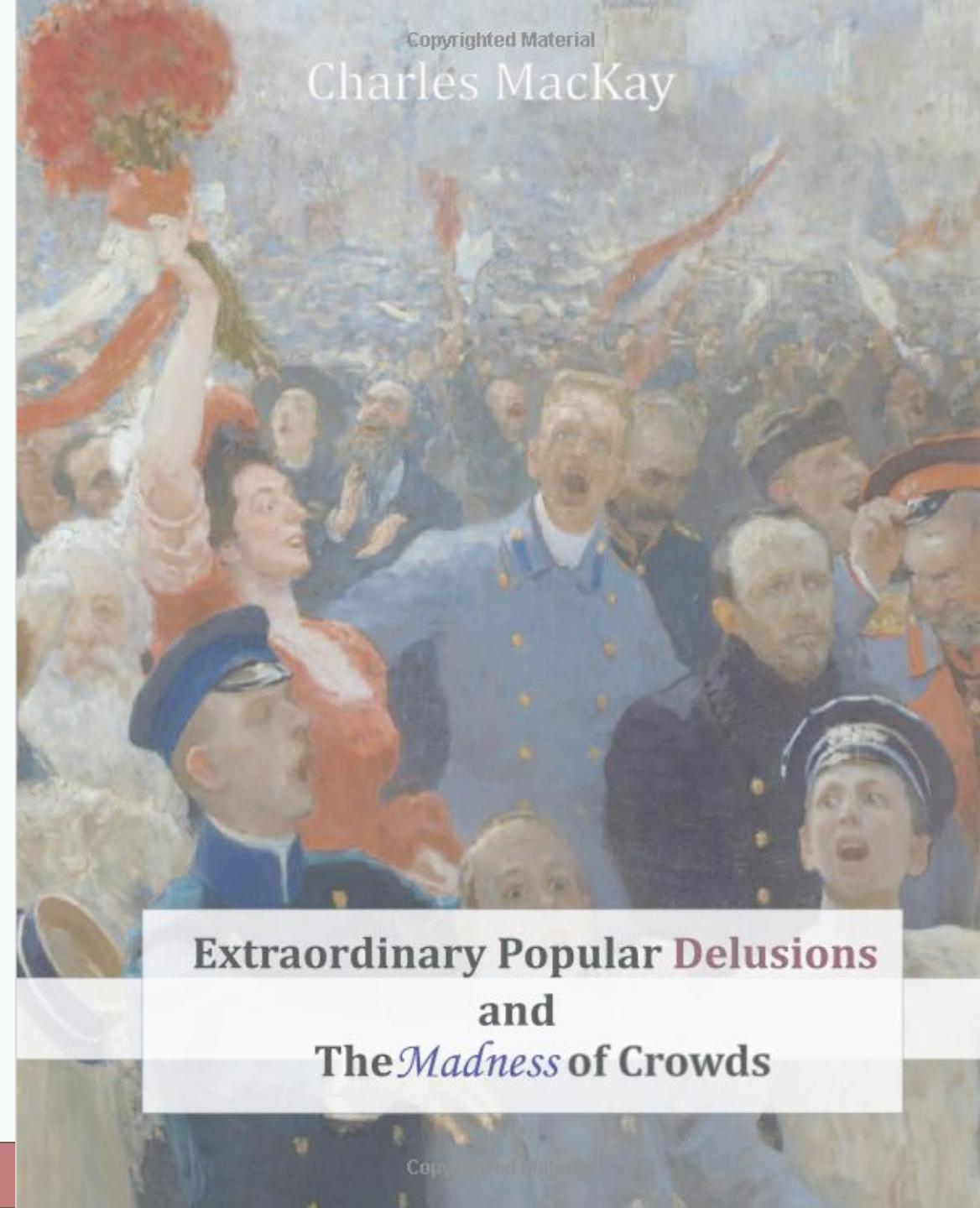
Popular Delusions and the Madness of Crowds

Economic bubbles

Alchemy & Psuedoscience

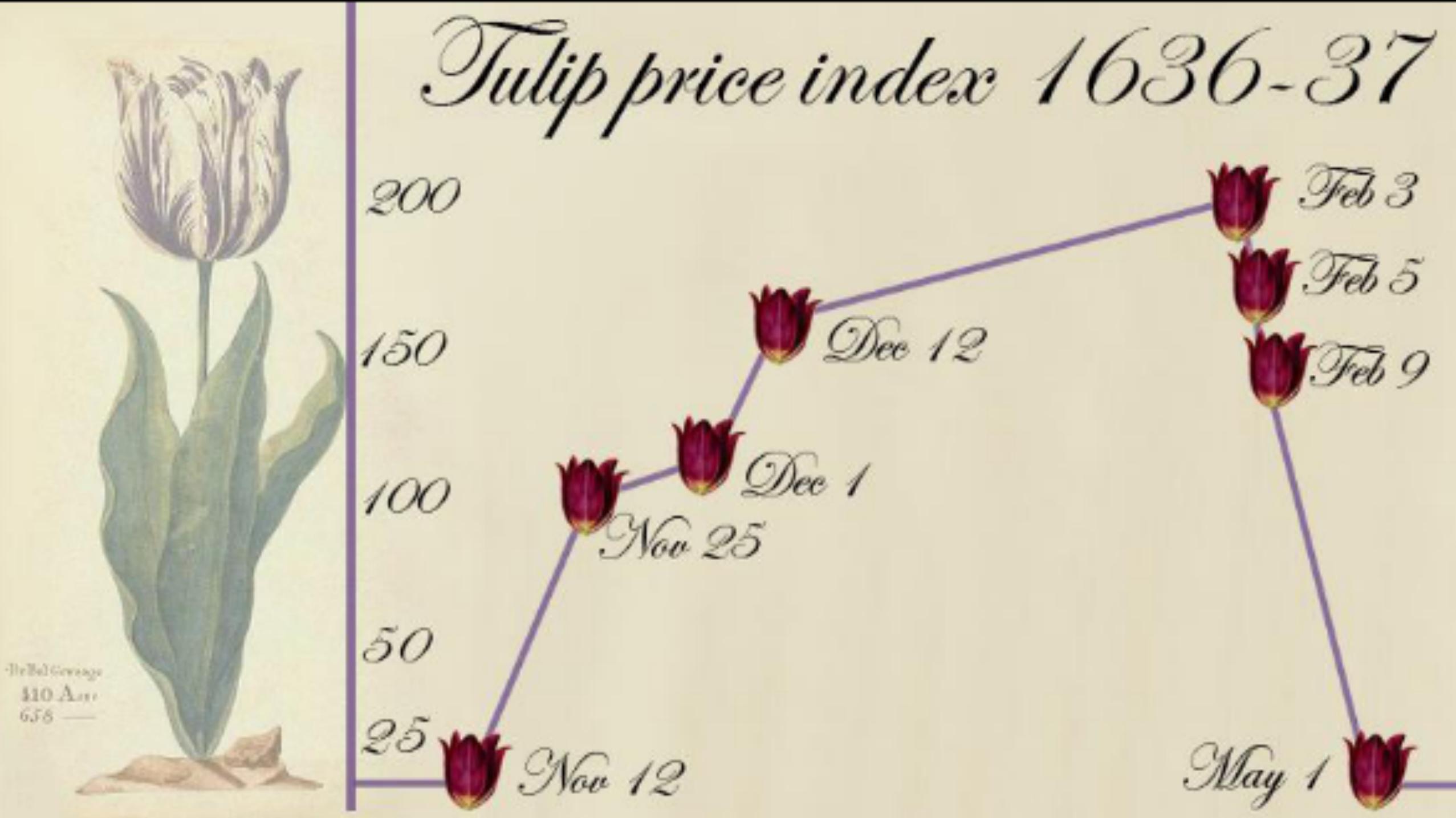
Witch hunts

Prophecies



Charles MacKay

Extraordinary Popular **Delusions**
and
The *Madness* of Crowds









Thank you, Ben.
We rather enjoyed this one.
WG1WGA!
Q



Wisdom of Crowds

Requirements for crowds to be wise:

Diversity of Opinion

Independence

De-centralization

Aggregation

A NEW YORK TIMES BUSINESS BESTSELLER

"As entertaining and thought-provoking as *The Tipping Point* by Malcolm Gladwell. . . . *The Wisdom of Crowds* ranges far and wide."

—*The Boston Globe*

THE WISDOM OF CROWDS

JAMES SUROWIECKI

WITH A NEW AFTERWORD BY THE AUTHOR



Groups / Crowds

Employees of a business

Participants in a poll

Sports fans betting on games

Independent stock market investors

Internet users linking to sites

Citizens in a democracy

Ways of aggregating collective intelligence

Point spreads / parimutuel odds

Stock prices

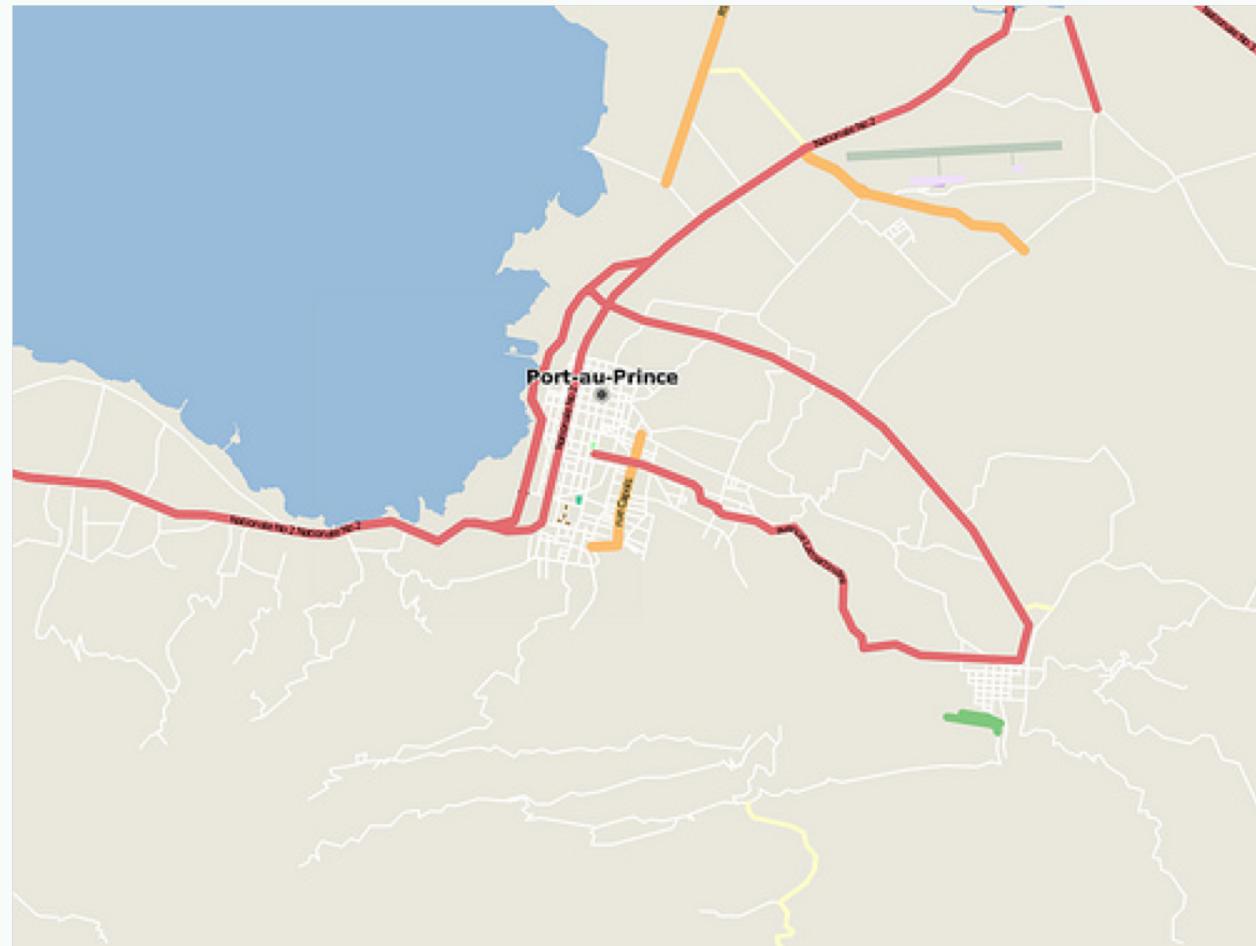
Futures contracts

Voting

Computer algorithms, interfaces



Disaster Response: Better maps from Crowdsourcing



Jan 12, 2010



Jan 23, 2010

 **Ushahidi**
Robert Munro

Disaster Response: The responders don't speak Kreyol

Fanmi mwen nan Kafou, 24 Cote Plage,
41A bezwen manje ak dlo

My family in Carrefour, 24 Cote Plage, 41A
needs food and water

Moun kwense nan Sakre Kè nan
Pòtoprens

People trapped in Sacred Heart Church,
PauP

Ti ekipman Lopital General genyen yo
paka minm fè 24 è

General Hospital has less than 24 hrs.
supplies

Fanm gen tranche pou fè yon ptit nan
Delmas 31

Undergoing children delivery Delmas 31

Disaster Response: Maps + Translation + Local Knowledge

Workers collaborated to find locations:

Dalila: I need Thomassin Apo please

Apo: Kenscoff Route: Lat:

18.495746829274168, Long:-
72.31849193572998

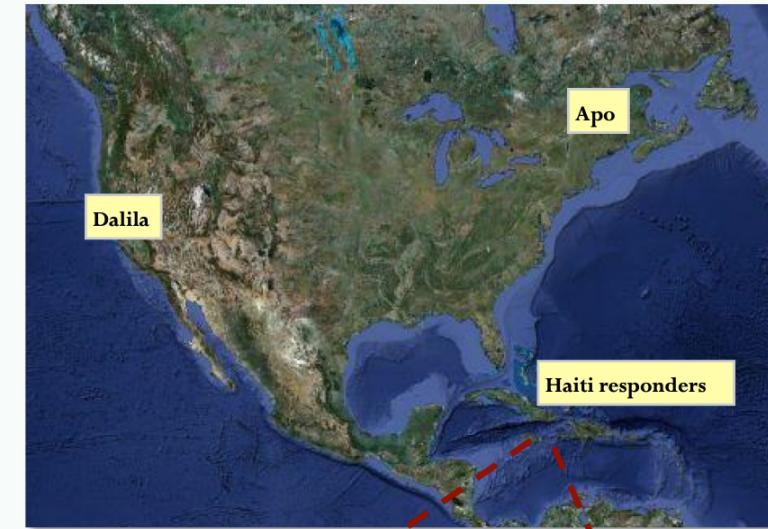
Apo: This Area after Petion-Ville and Pelerin 5 is not on Google Map. We have no streets name

Apo: I know this place like my pocket

Dalila: thank God u was here

Feedback from responders:

"just got emergency SMS, child delivery, USCG are acting, and, the GPS coordinates of the location we got from someone of your team were 100% accurate!"



Robert Munro

Disaster Response testimonials

Clark Craig of the Marine Corps:

"I cannot overemphasize to you what the work of the Ushahidi/Haiti has provided. It is saving lives every day."

Secretary of State Hillary Clinton:

"The technology community has set up interactive maps to help us identify needs and target resources. And on Monday, a seven-year-old girl and two women were pulled from the rubble of a collapsed supermarket by an American search-and-rescue team after they sent a text message calling for help."

Craig Fulgate, FEMA Task Force:

"[The] Crisis Map of Haiti represents the most comprehensive and up-to-date map available to the humanitarian community."

Ushahidi@Tufts :

"The World Food Program delivered food to an informal camp of 2500 people, having yet to receive food or water, in Diquini to a location that 4636 had identified for them."

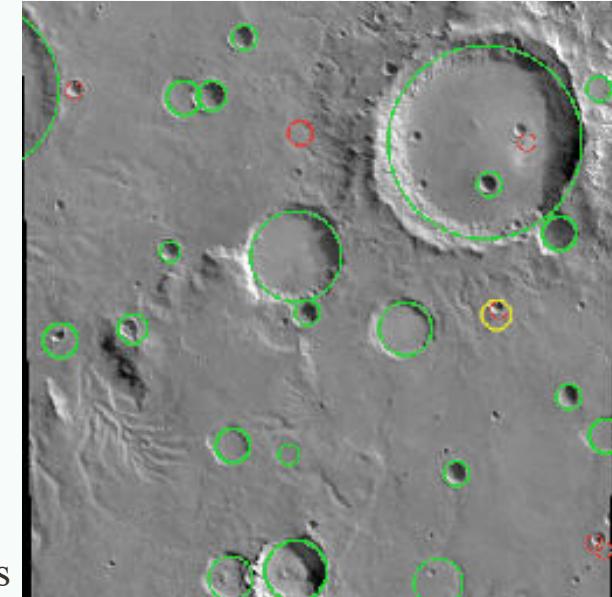
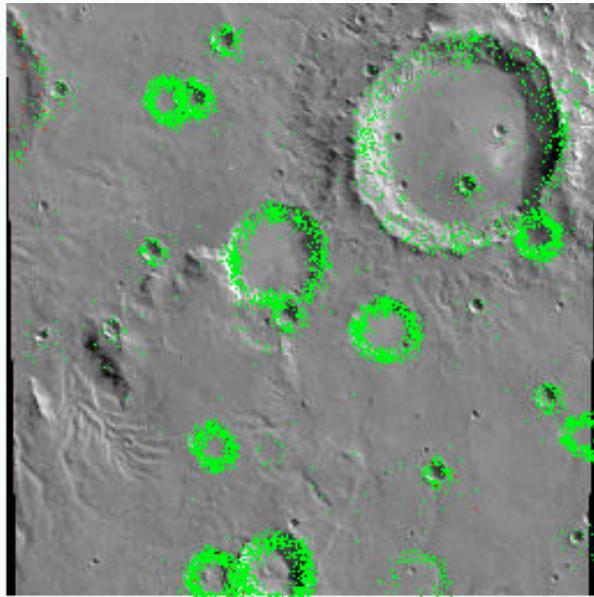
Robert Munro

Citizen Science

NASA Clickworkers (2000)

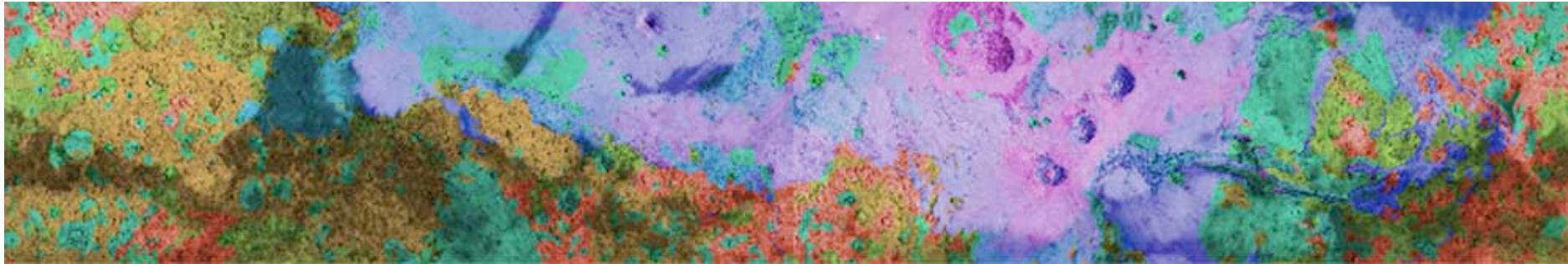
NASA showed that public volunteers could do routine science analysis that would normally be done by a graduate student working for months on end. From 11/2000 to 1/2002, they had 101,000 clickworkers volunteering 14,000 work hours, 612,832 sessions, and 2,378,820 entries!

We try to have several people cover each region on Mars so that we can compute a consensus, throwing out any mistaken or frivolous entries and averaging out the inaccuracies.

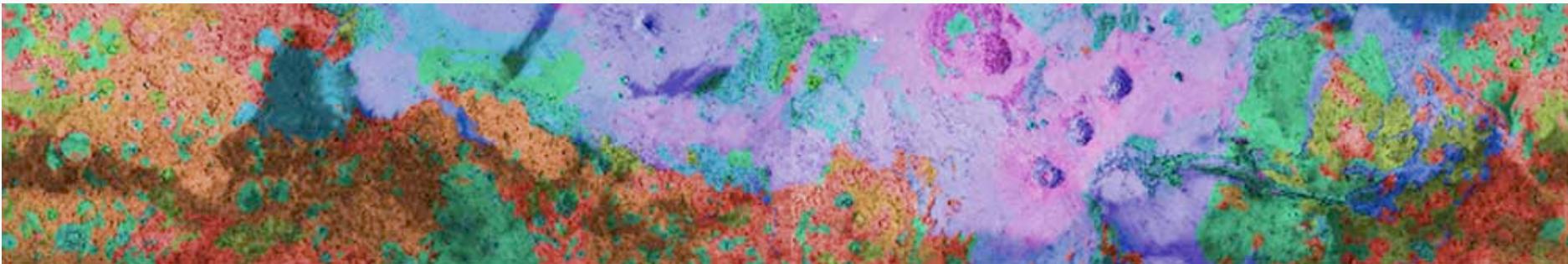


NASA Clickworkers (2000)

Mars age map produced directly from clickworker inputs.



Mars age map produced from scientists



Color guide: red=heavily cratered (old), green=medium, violet=lightly cratered (young).



WELCOME TO THE ZOONIVERSE

People-powered research

[See All Projects](#)

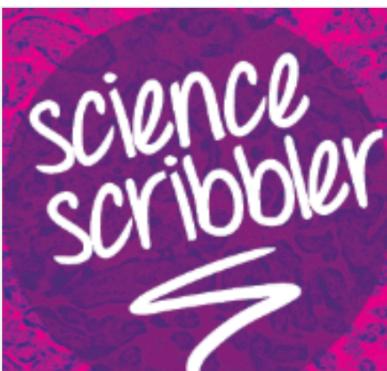
FEATURED PROJECTS



BELUGA BITS



SUPERWASP VARIABLE STARS



SCIENCE SCRIBBLER:
PLACENTA PROFILES



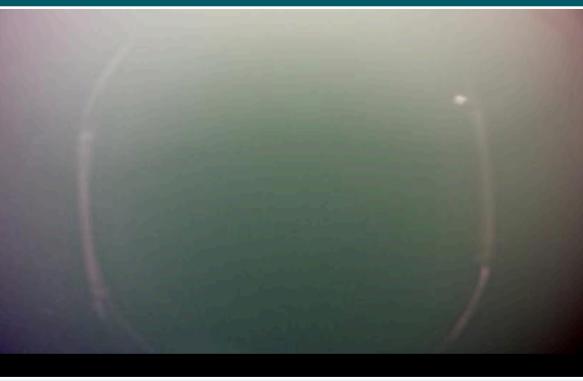
POLLINATORWATCH

<https://www.zooniverse.org>

Zooniverse

Examine underwater photos of wild beluga whales and help us identify the age, sex, and group size. We also need keen eye to look for identifying marks to recognize beluga that return to this location year after year.

[Learn more](#)

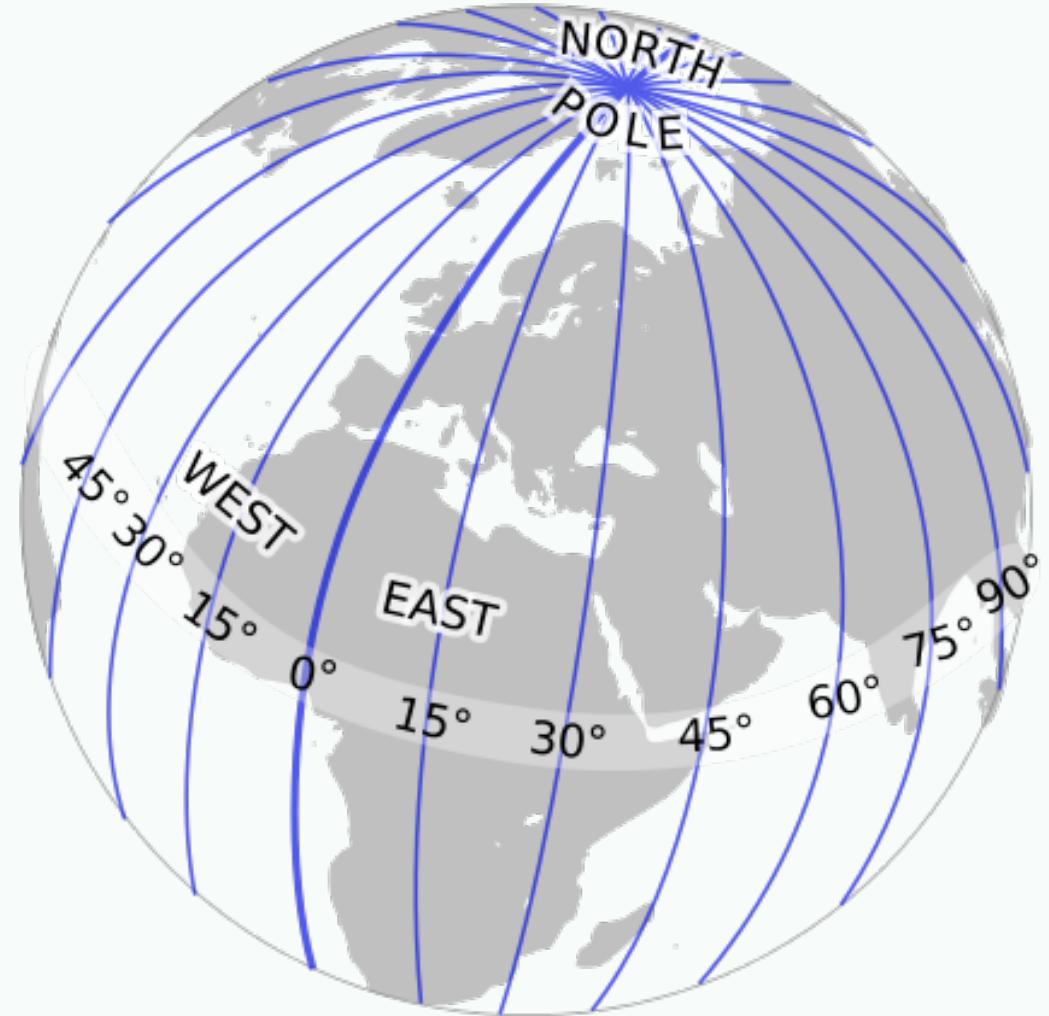


5 people are talking about **Beluga Bits** right now.

[Join in](#)

<https://www.zooniverse.org>

Longitude rewards



Netflix Prize

[Home](#) | [Rules](#) | [Leaderboard](#) | [Register](#) | [Update](#) | [Submit](#) | [Download](#)

Leaderboard

Display top leaders.

Rank	Team Name	Best Score	% Improvement	Last Submit Time
-	No Grand Prize candidates yet	-	-	-
<u>Grand Prize - RMSE <= 0.8563</u>				
1	PragmaticTheory	0.8584	9.78	2009-06-16 01:04:47
2	BellKor in BigChaos	0.8590	9.71	2009-05-13 08:14:09
3	Grand Prize Team	0.8593	9.68	2009-06-12 08:20:24
4	Dace	0.8604	9.56	2009-04-22 05:57:03
5	BigChaos	0.8613	9.47	2009-06-15 18:03:55

 Home Compete Data Notebooks Communities Courses MoreActive

Completed

InClass

All Categories ▾

Reward ▾

**Jane Street Market Prediction**

Test your model against future real market data

Featured • a month to go • Code Competition • 2865 Teams

\$100,000

**HuBMAP - Hacking the Kidney**

Identify glomeruli in human kidney tissue images

Research • 2 months to go • Code Competition • 895 Teams

\$60,000

**RANZCR CLiP - Catheter and Line Position Challenge**

Classify the presence and correct placement of tubes on chest x-rays to save lives

Featured • 2 months to go • Code Competition • 567 Teams

\$50,000

**VinBigData Chest X-ray Abnormalities Detection**

Automatically localize and classify thoracic abnormalities from chest radiographs

Featured • 2 months to go • 352 Teams

\$50,000

**Acea Smart Water Analytics**

Can you help preserve "blue gold" using data to predict water availability?

Analytics • a month to go

\$25,000

**Cassava Leaf Disease Classification**

Identify the type of disease present on a Cassava Leaf image

Research • a month to go • Code Competition • 2961 Teams

\$18,000

**Rainforest Connection Species Audio Detection**

Automate the detection of bird and frog species in a tropical soundscape

Research • a month to go • 894 Teams

\$15,000

Home

Compete

Data

Notebooks

Communities

Courses

More

Jane Street Market Prediction

Test your model against future real market data



Jane Street Group · 2,865 teams · a month to go (25 days to go until merger deadline)

\$100,000

Prize Money

[Overview](#)[Data](#)[Notebooks](#)[Discussion](#)[Leaderboard](#)[Datasets](#)[Rules](#)[Join Competition](#)[Public Leaderboard](#)[Private Leaderboard](#)

This leaderboard is calculated with all of the test data.

[Raw Data](#)[Refresh](#)

In the money

Gold

Silver

Bronze

#	Team Name	Notebook	Team Members	Score	Entries	Last
1	srinadhu			11036.536	44	5h
2	Andy Andre			10677.906	141	1d
3	Buy High, sell low			10534.239	219	1d
4	h309 = o1411			10427.453	211	6h
5	Çorlulu			10369.479	60	18h

Games with a Purpose

score
100

ESP Game

Concentrate...

time
2:21

What do you see?

taboo words

peace

lay



guesses

sheeps...

sheep

submit pass

**Are you a beginner?**

Start here with the Basics

**Not a beginner?**

Try this Placement Test

Basics 2
0/5Phrases
0/4

Learn



Practice



Shop

Take a test to jump ahead!



It takes about 5 minutes, and adapts to your level by getting harder (or easier) based on your answers.

Cancel

Start

Quit

Translate this sentence

A girl

.....
Une fille**You are correct**

Continue



Luis Von Ahn

==

Tom Sawyer



Tom Sawyer (Whitewashing the Fence), 1936

Real or Fake Text?



How good are you (and other humans) at knowing when text has been written by a computer? Is the computer better at generating convincing text if the topic is news, or short stories, or maybe presidential speeches?

<http://roft.io>

Know Your Nyms?



Know Your Nyms?

KEEP YOUR BRAIN ON ITS TOES

Welcome CCB!

Total Rounds Played: 433

Total Score: 34138.2

Average Score Per Round: 78.84

How to Play

About the Team

Leaderboards

Start Round

Synonyms Antonyms Hyponyms Meronyms

<https://www.know-your-nyms.com>

Implicit Work

CAPTCHA

The Norwich line steamboat train, from New-London for Boston, this morning ran off the track seven miles north of New-London.

morning

morning overtook

Type the two words:



reCAPTCHA™
stop spam.
read books.

Groups



Add



NETS 213 Course ...

Please use this feature carefully. Only add people you know.

Enter email addresses to add as members

Yoni Nachmany <yoninachmany@gmail.com>, Jie Gao <jie

Separate email addresses with commas. Each person will imme

Write a welcome message

Welcome to the NETS 213 staff mailing list!



Continue

Email subscription options

- No email:** web-only participation
- Abridged Email:** one summary email
- Digest Email:** up to 25 full new mess
- All Email:** send each message as it arrives

Select all squares with
traffic lights
If there are none, click skip

957 characters remaining

SKIP

ma@seas.upenn.edu>

Groups



Add



NETS 213 Course ...

Please use this feature carefully. Only add people you know.

Enter email addresses to add as members

Yoni Nachmany <yoninachmany@gmail.com>, Jie Gao <jie...

Separate email addresses with commas. Each person will imme...

Write a welcome message 

Welcome to the NETS 213 staff mailing list!

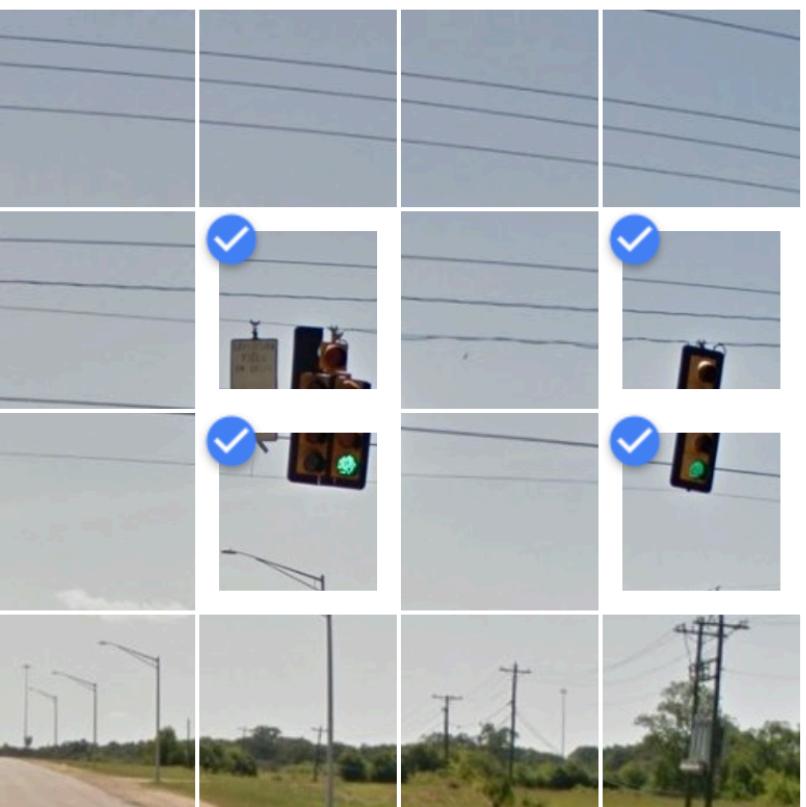


Contin...

Email subscription options

- No email: web-only participation
- Abridged Email: one summary email
- Digest Email: up to 25 full new mess...
- All Email: send each message as it arrives

Select all squares with
traffic lights
If there are none, click skip



C i NEXT

ma@seas.upenn.edu>

957 characters remaining

Groups



Add



NETS 213 Course ...

Please use this feature carefully. Only add people you know. Using this feature for sending unwanted email can result in account deactivation.

Enter email addresses to add as members

Yoni Nachmany <yoninachmany@gmail.com>, Jie Gao <jiegao@seas.upenn.edu>, Paul Zuo <pzuo@seas.upenn.edu>, Danni Ma <dannima@seas.upenn.edu>

Separate email addresses with commas. Each person will immediately become a member and can start receiving messages.

Write a welcome message 

Welcome to the NETS 213 staff mailing list!

 I'm not a robot 
Privacy - Terms



957 characters remaining

Email subscription options

- No email: web-only participation
- Abridged Email: one summary email
- Digest Email: up to 25 full new messages
- All Email: send each message as it arrives

Microwork

[Introduction](#) | [Dashboard](#) | [Status](#) | [Account Settings](#)

Mechanical Turk is a marketplace for work.

We give businesses and developers access to an on-demand, scalable workforce.

Workers select from thousands of tasks and work whenever it's convenient.

37,649 HITs available. [View them now.](#)

Make Money by working on HITs

HITs - *Human Intelligence Tasks* - are individual tasks that you work on. [Find HITs now.](#)

As a Mechanical Turk Worker you:

- Can work from home
- Choose your own work hours
- Get paid for doing good work

Find an interesting task



Work



Earn money



[Find HITs Now](#)

or [learn more about being a Worker](#)

Get Results from Mechanical Turk Workers

Ask workers to complete HITs - *Human Intelligence Tasks* - and get results using Mechanical Turk. [Register Now](#)

As a Mechanical Turk Requester you:

- Have access to a global, on-demand, 24 x 7 workforce
- Get thousands of HITs completed in minutes
- Pay only when you're satisfied with the results

Fund your account



Load your tasks

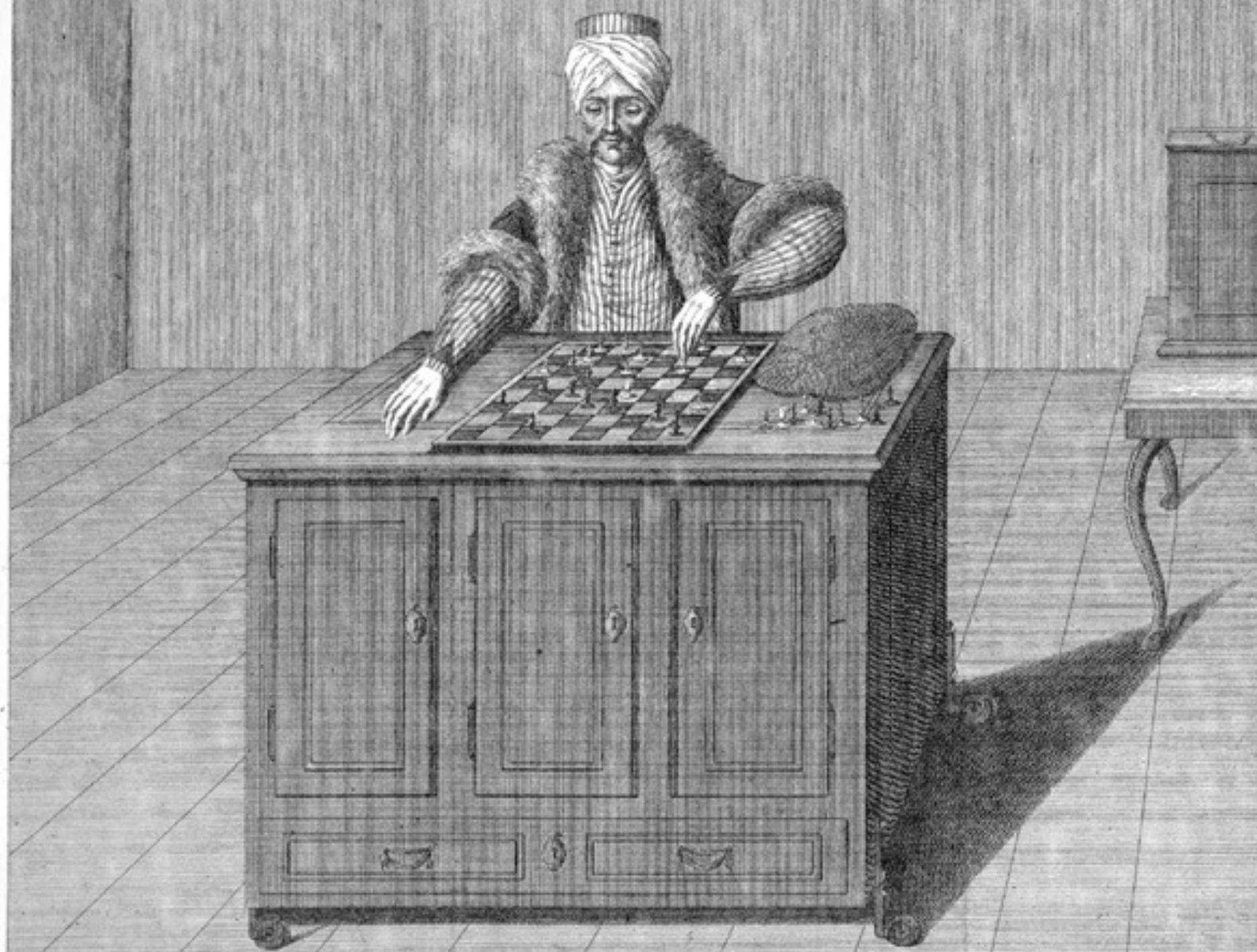


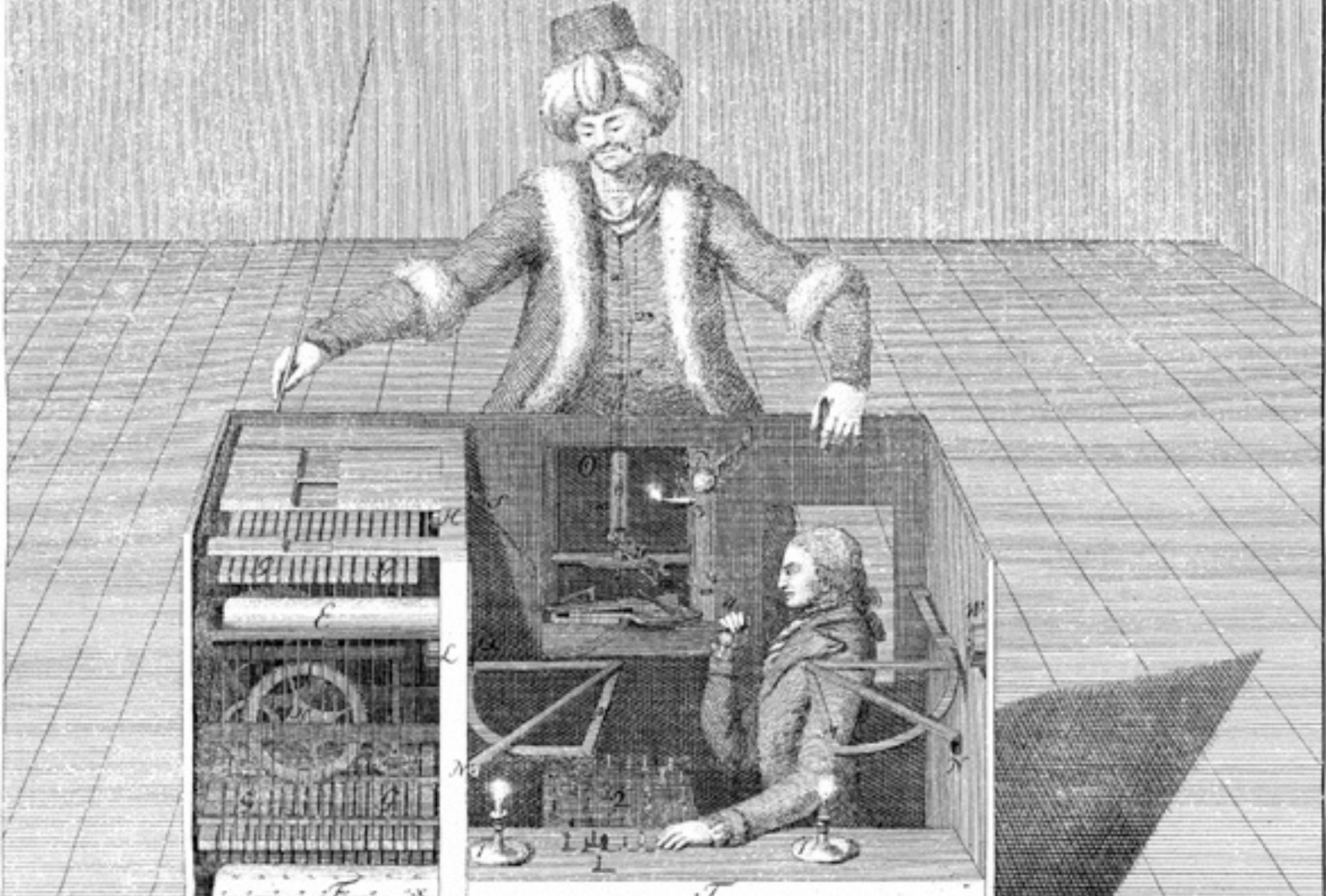
Get results



[Get Started](#)

or [learn more about being a Requester](#)





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

Search for HITs **containing** **that pay at least \$** **0.00** **for which you are qu**

Timer: 00:00:00 of 5 minutes **Want to work on this HIT?** [Want to see other HITs?](#) **Total Earned:** Unavailal

Accept HIT

Skip HIT

Total Earned: Unavailable

Total HITs Submitted: 0

Enter Postmark & Stamp Information for a Postcard

Requester: Cardcow

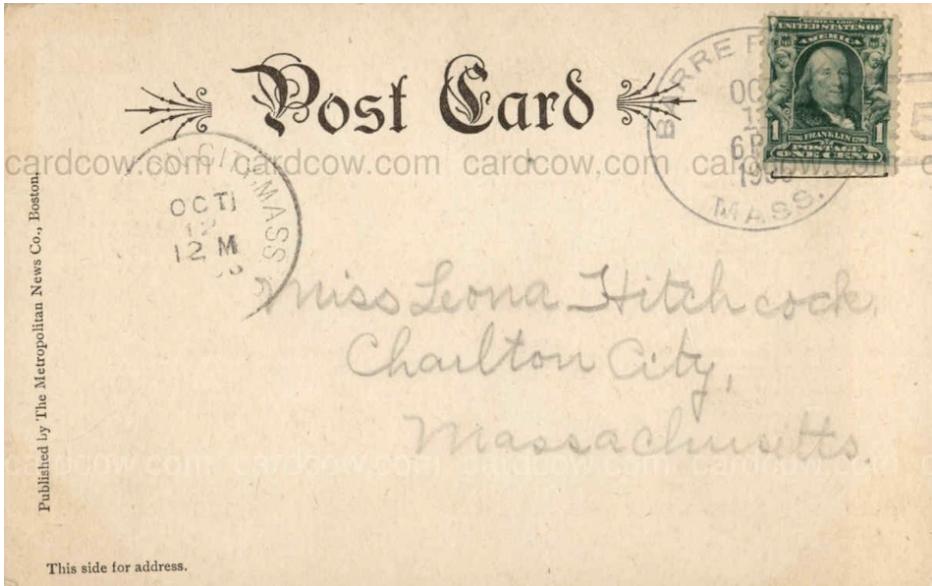
Reward: \$0.01 per HIT

HITs Available: 2 **Duration:** 5 m

Duration: 5 minutes

Qualifications Required: Data Entry for Postcards has been granted

Enter Postmark & Stamp Information for this card



Postmark City:

Postmark State:
(or Country)

Postmark Date:
(Ex: Nov-09)

Postmark Year:
(Ex: 1909)

Stamp: (Ex: 1c, 2c, half penny)

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

Search for HITs containing that pay at least \$ 0.00 for which you are qu:

Timer: 00:00:00 of 5 minutes Want to work on this HIT? Want to see other HITs? Total Earned: Unavail Total HITs Submitted: 0

Enter Postmark & Stamp Information for a Postcard

Requester: Cardcow

Reward: \$0.01 per HIT

HITs Available: 2

Duration: 5 minutes

This HIT has been granted



Postmark City: Barre

Postmark State: MA

Postmark Date: Oct-11

Postmark Year: 1886

Stamp: 1c

\$ 0.01

Postmark City:

Postmark State:
(or Country)

Postmark Date:
(Ex: Nov-09)

: 1c, 2c,

Ethical Questions about Amazon Mechanical Turk

Amazon Mechanical Turk: Gold Mine or Coal Mine?

Kar n Fort*
INIST-CNRS/LIPN

K. Bretonnel Cohen[†]
U. Colorado School of Medicine/U.
Colorado at Boulder

Gilles Adda**
LIMSI/CNRS



Recently heard at a tutorial in our field: "It cost me less than one hundred bucks to annotate this using Amazon Mechanical Turk!" Assertions like this are increasingly common, but we believe they should not be stated so proudly; they ignore the ethical consequences of using MTurk (Amazon Mechanical Turk) as a source of labour.

Ethical Questions about Amazon Mechanical Turk

Here's an excerpt from an IRB application Chris Callison-Burch posted - "We will pay participants small sums of money to complete our tasks, ranging from \$0.01 to \$1. All participants can choose for themselves whether the compensation is fair, and opt not to do it if they deem the compensation to be too low. Amazon's Mechanical Turk has many other researchers and companies offering tasks, so we will offer compensation that is similar to what others offer."

He first refers to Amazon's Mechanical Turk as "an online labor market." And that, I agree with. It is an online labor market.

Requesters like him, and CrowdFlower, collude, explicitly or implicitly, to keep wages at a substandard level that is compatible with existence standard.

Unlike Jewels, I don't blame workers for taking low paying jobs. I can't blame a person for being needy enough to take what amounts to a crust of bread. **I blame Chris Callison-Burch, and others like him, for keeping the standard wage at crust of bread level.**

Ethical Questions about Amazon Mechanical Turk

I tried one of those to see, I gave it up at 4 minutes in and about 2/3 of the way through. For the whole hit, I'd have taken about 6 minutes. 10 hits an hour - **\$1.70 an hour**. Restricted to U.S. residents.

This is far too low to be considered a fair wage for a U.S. resident. My performance may be very far off from what others can do. Perhaps I took 4 times or more as long as an average worker would.

My complaint is that any U.S. requester knows what wage rate is required for a U.S. resident to survive. We may not agree on an exact number. But as they say, I know a fair wage when I see it, and this is not it.

Mturk is actually much smaller than what it can appear to be. Something close to requester monopoly has the power to keep wages low. Requester co-operation, explicit or implicit, reinforces this.

Chris Callison-Burch is not unaware, I think, of the mechanics of the wage structure of Mturk.

**WORKERS
OF THE WORLD
UNITE!**



A Data-Driven Analysis of Workers' Earnings on Amazon Mechanical Turk

Kotaro Hara^{1,2}, Abigail Adams³, Kristy Milland^{4,5},
Saiph Savage⁶, Chris Callison-Burch⁷, Jeffrey P. Bigham¹

¹Carnegie Mellon University, ²Singapore Management University, ³University of Oxford
⁴McMaster University, ⁵TurkerNation.com, ⁶West Virginia University, ⁷University of Pennsylvania
kotarohara@smu.edu.sg

ABSTRACT

A growing number of people are working as part of on-line crowd work. Crowd work is often thought to be low wage work. However, we know little about the wage distribution in practice and what causes low/high earnings in this setting. We recorded 2,676 workers performing 3.8 million tasks on Amazon Mechanical Turk. Our task-level analysis revealed that workers earned a median hourly wage of only ~\$2/h, and only 4% earned more than \$7.25/h. While the average requester pays more than \$11/h, lower-paying requesters post much more work. Our wage calculations are influenced by how unpaid work is accounted for, *e.g.*, time spent searching for tasks, working on tasks that are rejected, and working on tasks that are ultimately not submitted. We further explore the characteristics of tasks and working patterns that yield higher hourly wages. Our analysis informs platform design and worker tools to create a more positive future for crowd work.

temporarily out-of-work engineers to work [1,4,39,46,65].

Yet, despite the potential for crowdsourcing platforms to extend the scope of the labor market, many are concerned that workers on crowdsourcing markets are treated unfairly [19,38,39,42,47,59]. Concerns about low earnings on crowd work platforms have been voiced repeatedly. Past research has found evidence that workers typically earn a fraction of the U.S. minimum wage [34,35,37–39,49] and many workers report not being paid for adequately completed tasks [38,51]. This is problematic as income generation is the primary motivation of workers [4,13,46,49].

Detailed research into crowd work earnings has been limited by an absence of adequate quantitative data. Prior research based on self-reported income data (*e.g.*, [4,34,49]) might be subject to systemic biases [22] and is often not sufficiently granular to facilitate a detailed investigation of earnings dispersion. Existing data-driven quantitative work

A Data-Driven Analysis of Workers' Earnings on Amazon Mechanical Turk

Kotaro Hara^{1,2}, Abi Adams³, Kristy Milland^{4,5},

Salph Savage⁶, Chris Callison-Burch⁷, Jeffrey P. Bigham¹

¹Carnegie Mellon University, ²Singapore Management University, ³University of Oxford

⁴McMaster University, ⁵TurkerNation.com, ⁶West Virginia University, ⁷University of Pennsylvania

"...median hourly wages... \$1.77/h..."

tasks on Amazon Mechanical Turk. Our task-level analysis revealed that workers earned a median hourly wage of only ~\$2/h, and only 4% earned more than \$7.25/h. While the average requester pays more than \$11/h, lower-paying requesters post much more work. Our wage calculations are influenced by how unpaid work is accounted for, e.g., time

spent searching for tasks, working on tasks that are rejected, and working on tasks that are ultimately not submitted. We further explore the characteristics of tasks and working patterns that yield higher hourly wages. Our analysis informs platform design and worker tools to create a more positive future for crowd work.

Author Keywords

Crowdsourcing; Amazon Mechanical Turk; Hourly wage

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

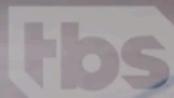
INTRODUCTION

Crowd work is growing [31,46]. A report by Harris and Kroszner states that 600k workers participate in the online

has found evidence that workers typically earn a fraction of the U.S. minimum wage [34,35,37–39,49] and many workers report not being paid for adequately completed tasks [38,51]. This is problematic as income generation is the primary motivation of workers [4,13,46,49].

Detailed research into crowd work earnings has been limited by an absence of adequate quantitative data. Prior research based on self-reported income data (e.g., [4,34,49]) might be subject to systemic biases [22] and is often not sufficiently granular to facilitate a detailed investigation of earnings dispersion. Existing data-driven quantitative work in crowdsourcing research has taken the employers' perspective [49] (e.g., finding good pricing methods [36,50,61], suggesting effective task design for requesters [24,40]), or it characterizes crowdsourcing market dynamics [21,37]. Data-driven research on how workers are treated on the markets is missing.

This paper complements and extends the existing understanding of crowd work earnings using a data-driven



Use Cases

Use of MTurk-like systems in research

Participant pool for user studies, polling, cognitive science experiments

Annotation for machine learning tasks like computer vision or NLP

Human Computer Interaction: worker pools are hardwired into the UI

New Programming Languages Concepts

Study markets themselves for economics research, cost-optimization

Crowdsourcing and the Social Sciences

Social Psychology

From Buhrmester, Kwang, & Gosling (2011, Perspectives on Psychological Science) Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data? :

Findings indicate that: (a) MTurk participants are slightly more representative of the U.S. population than are standard Internet samples and are significantly more diverse than typical American college samples; (b) participation is affected by compensation rate and task length but participants can still be recruited rapidly and inexpensively; (c) realistic compensation rates do not affect data quality; and (d) the data obtained are at least as reliable as those obtained via traditional methods.

Economics

From Horton, Rand & Zeckhauser (2010, Experimental Economics) – The Online Laboratory: Conducting Experiments in a Real Labor Market :

Online experiments can be just as valid as laboratory and field experiments, while requiring far less money and time to design and to conduct. We describe the benefits of conducting experiments in online labor markets and we replicate three classic experiments.

Clinical Psychology

Management/Cognition

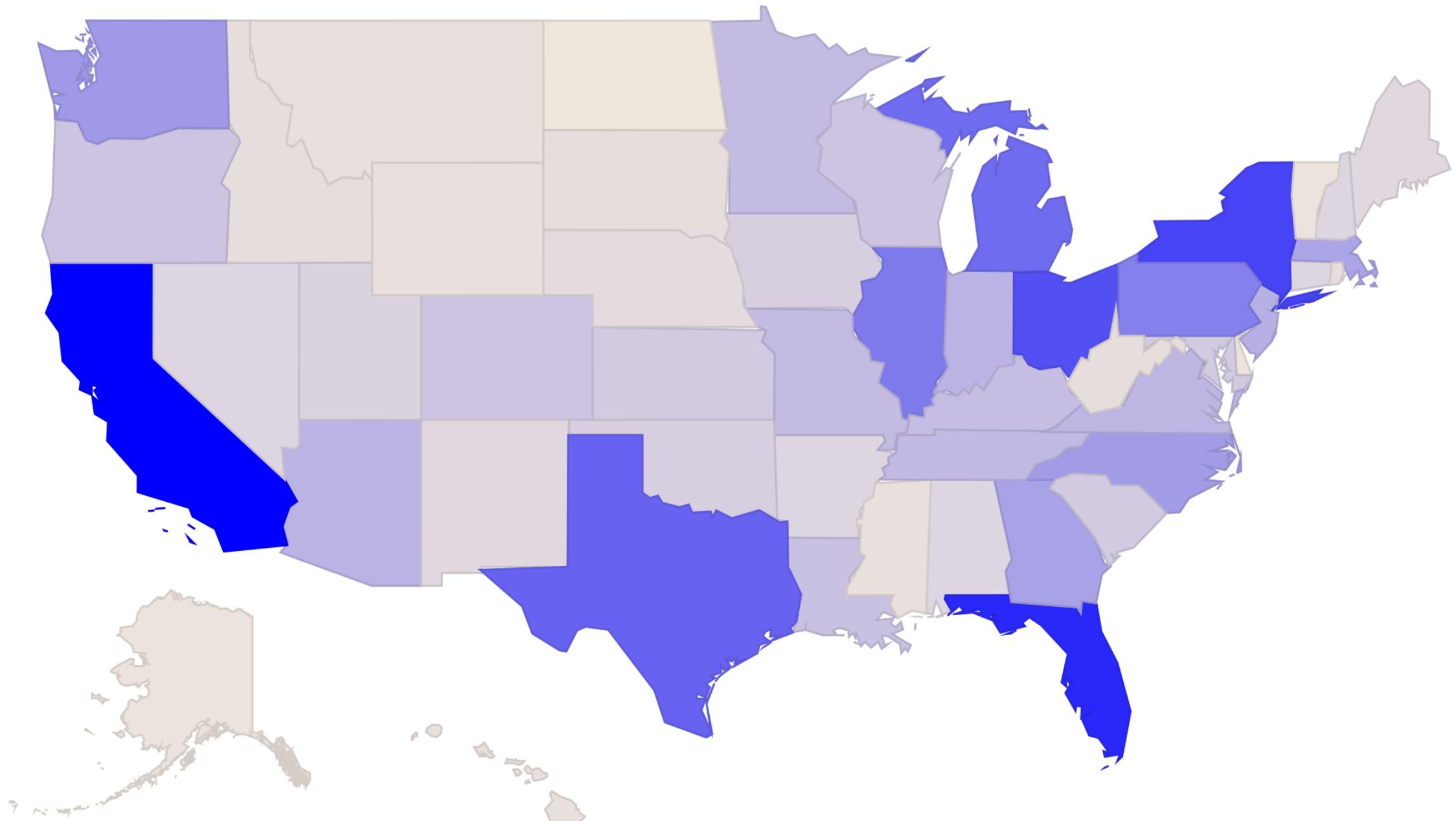
Theoretical Biology

Study Markets Themselves

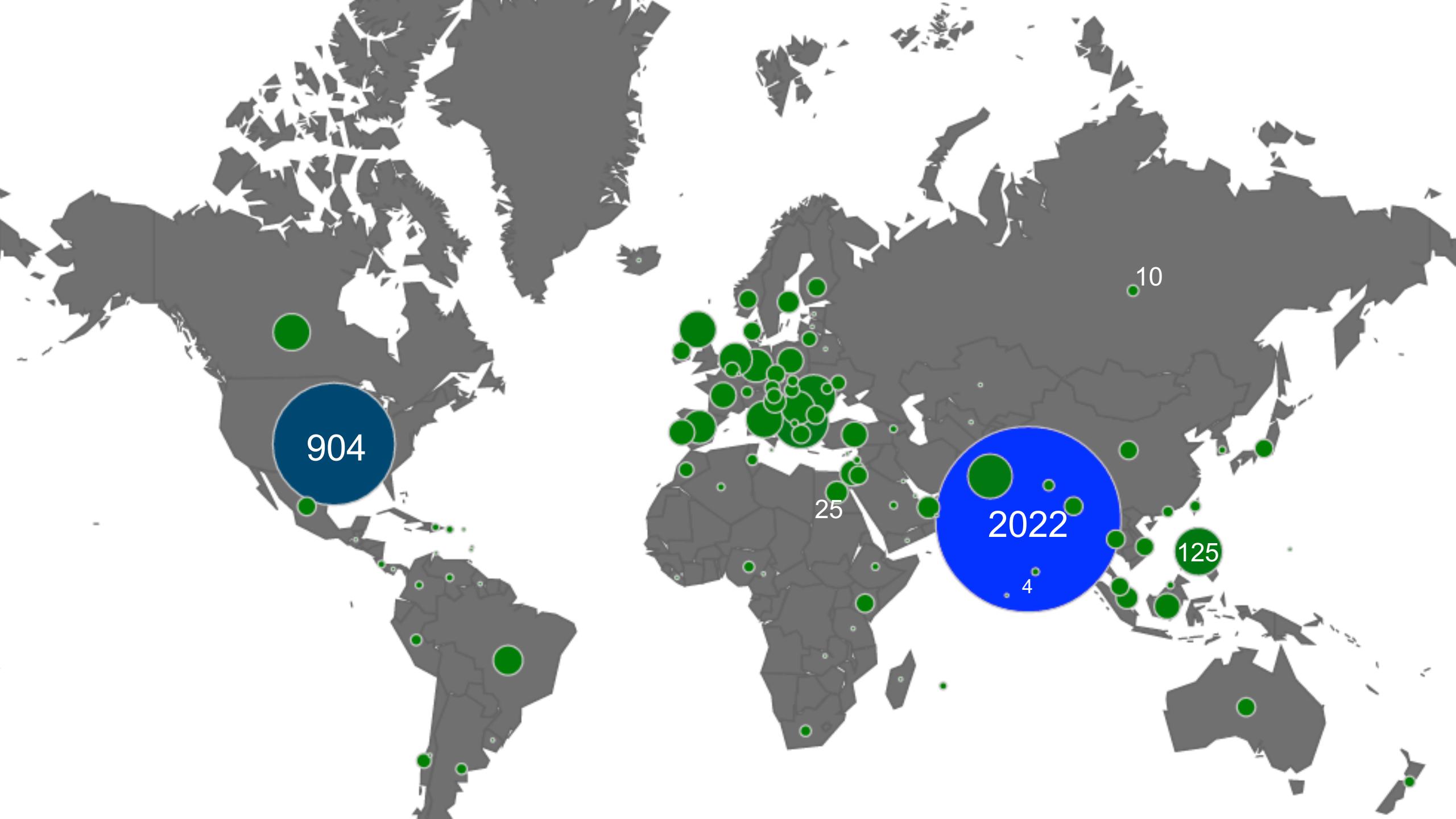
What predictions of economics hold true on MTurk?

What incentives can we give to increase throughput, quality, worker retention?

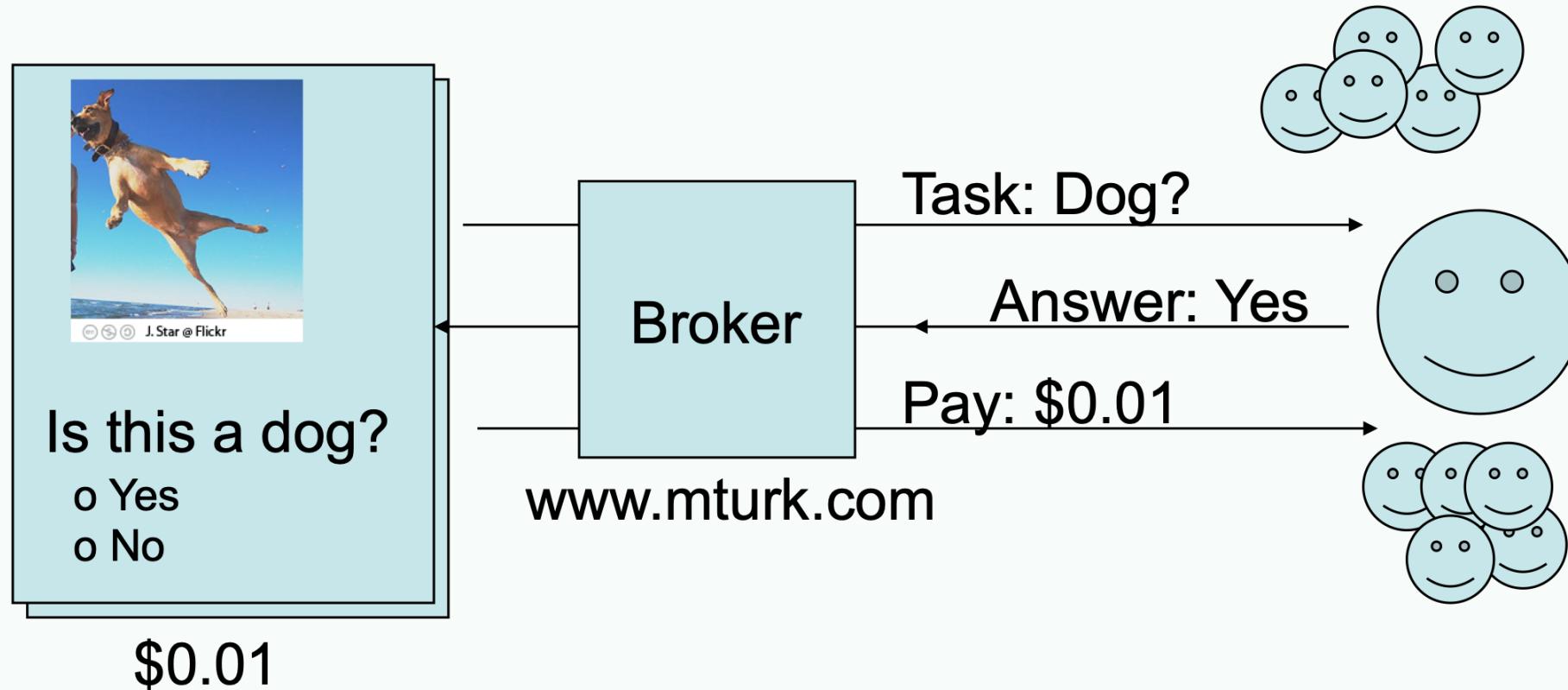
What is the cost-optimal solution to a problem?



0 81



Annotation for machine learning / artificial intelligence tasks



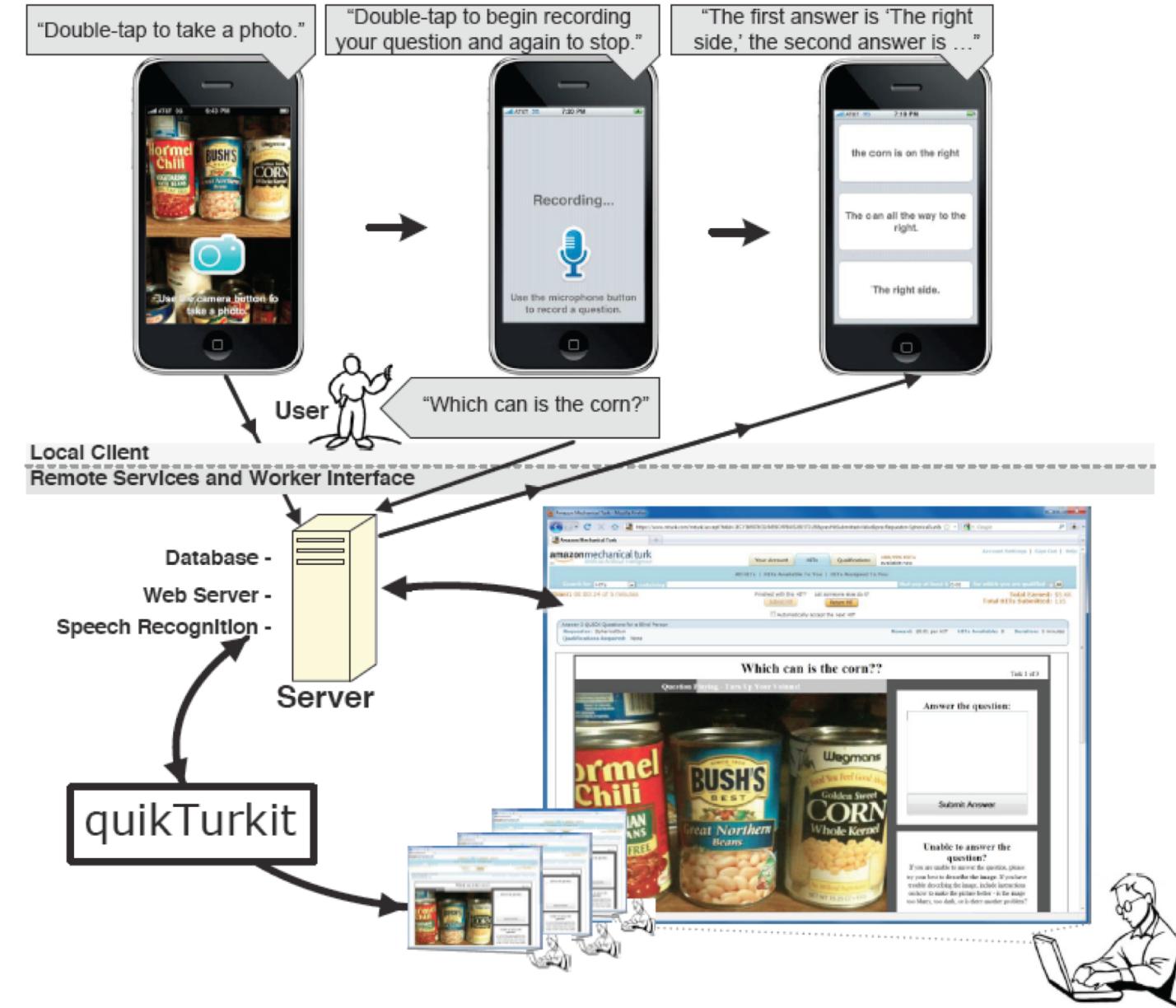


“We decided we wanted to do something that was completely historically unprecedented. We’re going to map out the entire world of objects.” The resulting dataset was called ImageNet.

IMAGENET



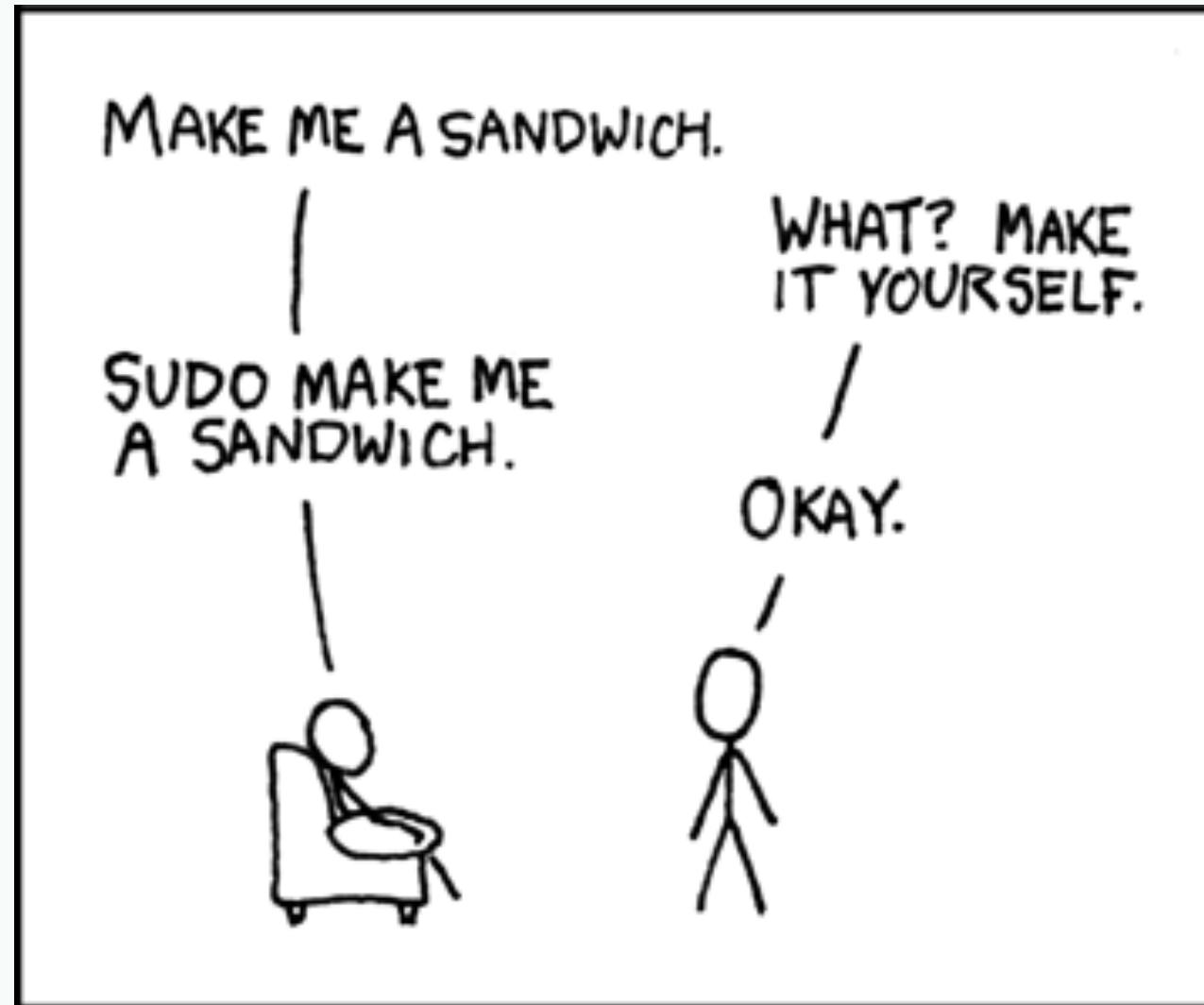
Human Computer Interaction



VizWiz - best paper award at UIST 2010-
Mobile service that aids blind users with
“visual questions” in near-realtime

https://www.youtube.com/watch?v=EtfrimJQ0_M

New Programming Languages Concepts



New Programming Languages Concepts

```
ideas = []
for (var i = 0; i < 5; i++) {
    idea = mturk.prompt(
        "What's fun to see in New York City?
        Ideas so far: " + ideas.join(", "))
    ideas.push(idea)
}

ideas.sort(function (a, b) {
    v = mturk.vote("Which is better?", [a, b])
    return v == a ? -1 : 1
})
```

**What will we cover in this
class (and should you
take it)?**

Topics

Taxonomy of crowdsourcing and human computation

Crowdsourcing platforms like Mechanical Turk and CrowdFlower

Programming concepts for human computation

The economics of crowdsourcing

Crowdsourcing and machine learning

Applications to human computer interaction

Crowdsourcing and social science

Who should take this class

Anyone who wants to be on the cutting edge of this new field

Entrepreneurial students who want to start their own companies

Students from the business school who want to experiment with markets

Students from the social sciences who want to conduct large-scale studies with people

Homework assignments

Weekly assignments

Video presentations

Final project

Final presentation

Participation

Writing and Coding

Company profile,
project pitch

Self-designed, group
projects

Show off your work

Help write & improve
lecture notes

Example Final Project

<https://vimeo.com/334828345>

Example Final Project



How much programming is required?

Programming assignments are in Python

We provide code that you modify

We want everyone to be able to participate, regardless of
programming experience

For most assignments, you can work with partners (turn in only one
assignment - you'll both get the same grade)

What will you get out of this class?

Understanding of an emerging field of CS

Basic python and machine learning skills

Ideas that you could transform into a startup company or academic research

A new way of thinking about collective decision making by companies and countries



Chris Callison-Burch

Associate Professor in CIS
Undergrad from Stanford in Symbolic Systems
PhD from University of Edinburgh in Informatics
Research Faculty at Johns Hopkins University
Research Interests: Crowdsourcing, Natural Language Processing

Mechanical Turk Monitor: Top-1000 Recent Requesters

General	Top requesters	Arrivals	Completed	Search	About
---------	----------------	----------	-----------	--------	-------

	Requester ID	Requester	#Task	#HITs	Rewards	Last Posted On
1	Dolores Labs	A2IR7ETVOIULZU (RSS)	470	317857	33084.24	Oct 16, 2010 6:00:00 PM
2	ContentGalore	A2XL3J4NH6JI12 (RSS)	674	10622	17554.6	Oct 17, 2010 6:00:00 AM
3	SpeechInk	A1AQ7EJ5P7ME65 (RSS)	9019	13613	12876.67	Oct 17, 2010 9:00:00 AM
4	CastingWords	A3MI6MIUNWCR7F (RSS)	9454	14036	8947.02	Oct 17, 2010 9:00:00 AM
5	QuestionSwami	AD7C0BZNKYGYV (RSS)	629	4116	4750.37	Oct 17, 2010 9:00:00 AM
6	Chris Callison-Burch	A32TTE4XXN6MQZ (RSS)	11	9961	4458.02	Oct 13, 2010 1:00:00 PM
7	Smartsheet.com Clients	A1197OGL0WOQ3G (RSS)	434	38212	3118.28	Oct 16, 2010 6:00:00 PM
8	retaildata	AD14NALRDOSEN9 (RSS)	8	50288	3110.85	Oct 8, 2010 10:00:00 PM
9	Classify This	A1CTI3ZAWTR5AZ (RSS)	25	94590	1891.8	Oct 16, 2010 3:00:00 PM
10	Andrew Stephen	A1Y25F6MZCMQGY (RSS)	3	22705	1131.25	Oct 9, 2010 5:00:00 PM
11	Dolores Labs 2	A3JX8WONBL5N9X (RSS)	34	8976	1043.27	Oct 17, 2010 9:00:00 AM
12	RelevanceQuest	A8RMEN71ICE57 (RSS)	15	47881	1029.92	Oct 15, 2010 10:08:00 PM
13	Crowd Task	AFAOUHS65HNDS (RSS)	4	2388	955.6	Oct 12, 2010 1:00:00 PM
14	nlds.soe.ucsc.edu	A1HI9DWCF794RE (RSS)	4	4702	933.9	Sep 30, 2010 6:00:00 PM
15	Movie Enquirer	A20A8T722JJHEI (RSS)	153	154	758	Oct 12, 2010 3:00:00 AM
16	Erdem Kiciman	A15XM23193MQX (RSS)	7	6502	714.01	Oct 14, 2010 10:00:00 PM



Ajay Patel



Ann-Katrin Reuel



Anna Orosz



Ashley Nobi



Bryan Li



David Fan



Eddie Cohen



Jason Kim



Joe Cappadona



Megha Mishra



Rebecca Iglesias-Flores



Sarah Ye



Zach Li

HW: Sign up to be a Mechanical Turk Worker

crowdsourcing-class.org