

Using Crowdsourcing Effectively for Social Media Research

Ujwal Gadiraju, Gianluca Demartini,
Djellel Eddine Difallah,
Michele Catasta

Gianluca Demartini



- B.Sc., M.Sc. at U. of Udine, Italy
- Ph.D. at U. of Hannover, Germany
 - Entity Retrieval
- Worked at the eXascale Infolab U. Fribourg (Switzerland), UC Berkeley (on Crowdsourcing), Yahoo! (Spain), L3S Research Center (Germany)
- Senior Lecturer in Data Science at the iSchool, U. of Sheffield
- Tutorials on Entity Search at ECIR 2012 and RuSSIR 2015, on Crowdsourcing at ESWC 2013, ISWC 2013, ICWSM 2016, WebSci 2016, and RuSSIR 2016

g.demartini@sheffield.ac.uk

www.gianlucademartini.net

Djellel Eddine Difallah

ded@exascla.info



- PhD. in Computer Science at U. Fribourg, Switzerland
- Thesis: “Quality of Service in Crowd-Powered Systems”
- MSc. in Computer Science U. Louisiana, USA
- Dip. Eng. USTHB, Algeria
- Senior Researcher at eXascale Infolab U. Fribourg, Switzerland
- Intern at Microsoft Research (USA), Intern at Google (USA), Data Management Engineer at Schlumberger (Algeria)

Michele Catasta

michele.catasta@epfl.ch

- PhD. in Computer Science at EPFL, Switzerland
- Thesis: “MEMOR1ES: Memory-based Information Systems”
- MSc. in DERI Galway, Ireland
- **Research Scientist** at EPFL
- Intern at Yahoo Labs (Barcelona), Intern at Google (US), Intern at MIT Media Lab (US)

Tutorial Schedule

- 8.30-9.00 Welcome
- 9.00-10.00 Introduction to Crowdsourcing
 - Extrinsic vs intrinsic motivation
 - Incentives, Paid micro-task crowdsourcing
 - Lab vs crowdsourced studies, Ethical concerns
 - Amazon Mechanical Turk
 - Example applications

Tutorial Schedule (2)

- 10.00-10.30 and 11.00-11.30 Quality Control
 - Introduction to Quality Issues in Crowdsourcing
 - Aspects that Affect the Quality of Results
 - Understanding Worker Malicious Behavior
 - Typical Quality Control Measures
 - Best Practices and Design Patterns
- 11.30-12.30
 - Lifecycle of a crowdsourcing experiment
 - Human Computation Research Opportunities

Overview



from <http://www.bbc.co.uk/news/magazine-32993891>

Crowdsourcing

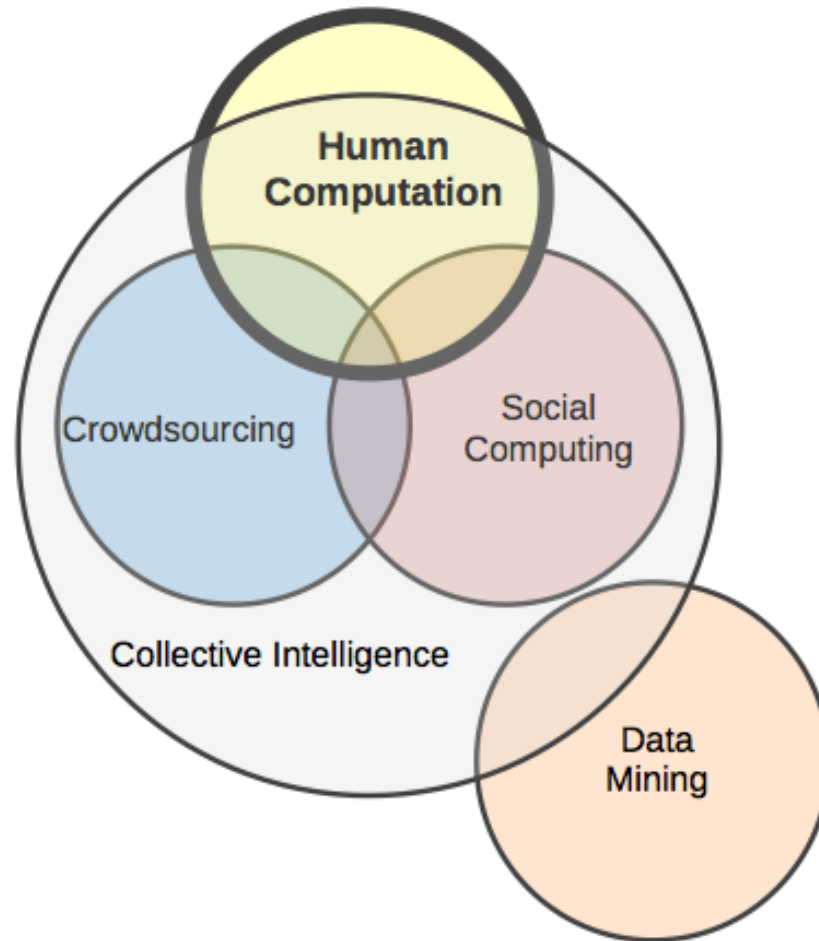
- *Portmanteau* of "crowd" and "outsourcing," first coined by Jeff Howe in a June 2006 Wired magazine article
- [Merriam-Webster] the practice of obtaining needed services, ideas, or content by soliciting contributions from a large group of people and especially from the online community rather than from traditional employees or suppliers

Crowdsourcing

- "Simply defined, crowdsourcing represents the act of a company or institution taking a function once performed by employees and outsourcing it to an undefined (and generally large) network of people in the form of an **open call**. This can take the form of peer-production (when the job is performed **collaboratively**), but is also often undertaken by sole **individuals**. The crucial prerequisite is the use of the open call format and the **large network of potential laborers**."

[Howe, 2006]

One View of Crowdsourcing



From Quinn & Bederson, “Human Computation: A Survey and Taxonomy of a Growing Field”, CHI 2011.

Dimensions of human computation

See also [Quinn & Bederson, 2012]

What is outsourced

- Tasks based on human skills not easily replicable by machines (visual recognition, language understanding, knowledge acquisition, basic human communication etc)

Who is the crowd

- Open call
- Call may target specific skills and expertise
- Requester typically knows less about the workers than in other work environments

How is the task outsourced

- Explicit vs. implicit participation
- Tasks broken down into smaller units undertaken in parallel by different people
- Coordination required to handle cases with more complex workflows
- Partial or independent answers consolidated and aggregated into complete solution

Dimensions of human computation (2)

See also [Quinn & Bederson, 2012]

How are the results validated

- Solutions space closed vs. open
- Performance measurements/ground truth
- Statistical techniques employed to predict accurate solutions
- May take into account confidence values of algorithmically generated solutions

How can the process be optimized

- Incentives and motivators
- Assigning tasks to people based on their skills and performance (as opposed to random assignments)
- Symbiotic combinations of human- and machine-driven computation, including combinations of different forms of crowdsourcing

Aligning incentives is essential

altruism
reputation
freedom reciprocity
self-expression
competition
community
autonomy
fun

- Motivation: driving force that makes humans achieve their goals
- Incentives: ‘rewards’ assigned by an external ‘judge’ to a performer for undertaking a specific task
 - Common belief (among economists): incentives can be translated into a sum of money for all practical purposes.
- Incentives can be related to both extrinsic and intrinsic motivations.
- **Extrinsic motivation** if task is considered boring, dangerous, useless, socially undesirable, dislikable by the performer.
- **Intrinsic motivation** is driven by an interest or enjoyment in the task itself.

Crowdsourcing Incentives

- Paid Crowdsourcing
 - **Competition** with others (bonus payment for best performance)
 - Surveillance (check before paying)
 - Solidarity (your **team** will receive a bonus)
 - Accuracy (**bonus** for correct answers)
 - **Agreement** with others (bonus for agreeing with the majority)
- Fun (enjoyment)
- Community (belonging, desire to help)

Paid Crowdsourcing Ethics

- People work full-time as crowd workers
- Chinese crowdsourcing platform with 6M+ workers
- Pros
 - Help developing countries
 - Provide cash fast to people == short-term satisfaction
 - Job Flexibility
- Cons
 - No job security
 - No social security
 - Long term satisfaction? Career plans?

Academics, Do we need ethics approval?

- Are crowd workers human subjects we are studying?
- Are crowd workers freelancers/subcontractors/employees we hire to do manual annotations?
- Sheffield: University Research Ethics Committee
 - Ethics application and approval
 - Participants, Payment, Data confidentiality/storage
 - Informed consent

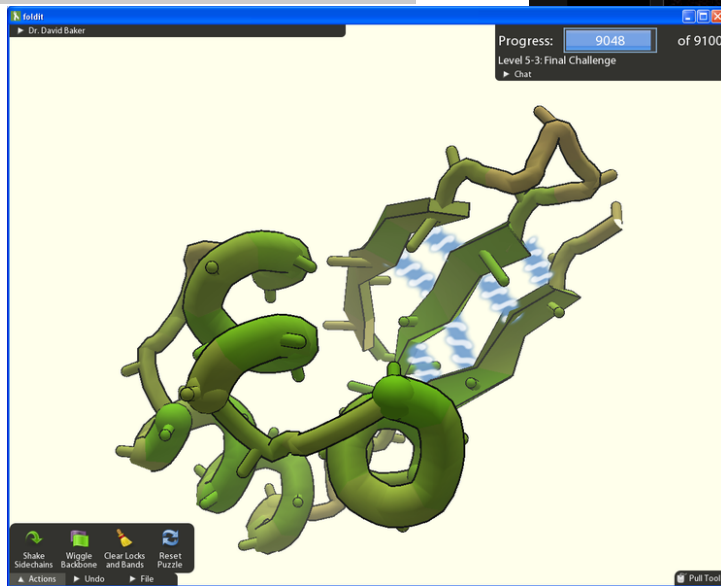
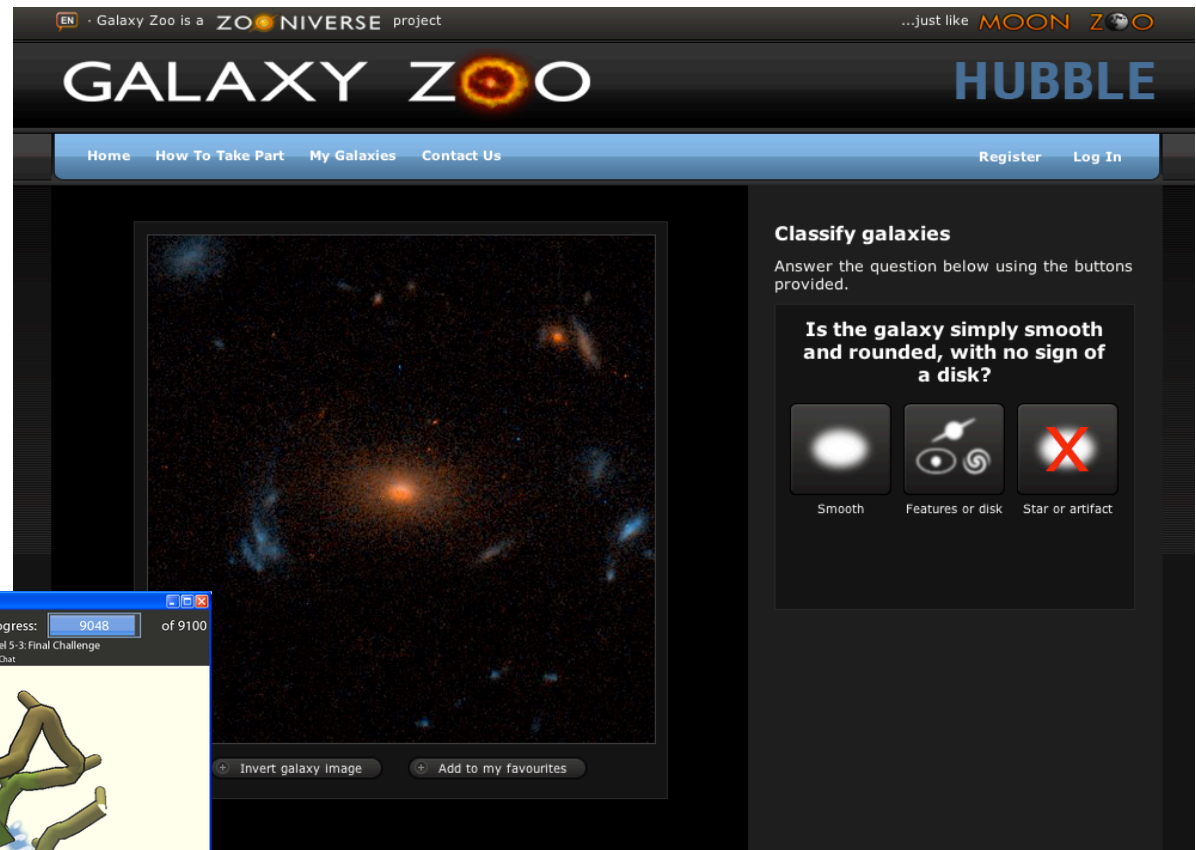
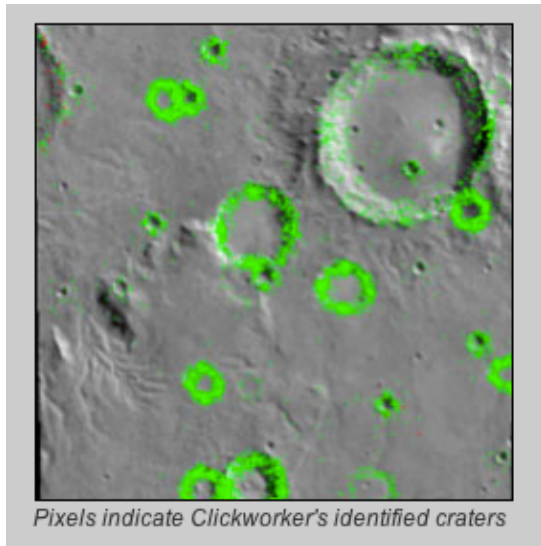
Informed consent form

- Purpose of the research
- Who will be participating?
- What will you be asked to do?
- What are the potential **risks** of participating?
- What **data** will we collect?
- What will we do with the data?
- Will my participation be **confidential**?
- What will happen to the **results** of the research project?
- Declaration of accepting the conditions and agreeing to participate

Lab vs crowdsourced studies

- Few subjects vs potentially unlimited crowd
- Full control vs logs
- Interviews, qualitative analysis vs questionnaires and surveys
- High quality, high motivation vs need for quality control
- Ideally lab+crowd for a social science experiment
- Crowd+experts for dataset creation

Citizen Science



Citizen Journalism and Participatory Sensing



innocentive.com

- Tech Innovation
- via Crowdsourcing
- Competitions
- Monetary Incentive

<https://www.innocentive.com/pavilion/NASA>




C

Challenge Title
Improved Barrier Layers ... Keeping Food Fresh in Space
Mechanism for a Compact Aerobic Resistive Exercise Device
Data-Driven Forecasting of Solar Events
Coordination of Sensor Swarms for Extraterrestrial Research
Medical Consumables Tracking
Augmenting the Exercise Experience
Simple Microgravity Laundry System

Question Answering Systems

Welcome to Q&A for professional and enthusiast programmers — check out the [FAQ!](#)

StackExchange [log in](#) [careers](#) [dev days](#) [chat](#) [meta](#) [about](#) [faq](#)

 **stackoverflow** [Questions](#) [Tags](#) [Users](#) [Badges](#) [Unanswered](#) [Ask Question](#)

Top Questions [interesting](#) [237 featured](#) [hot](#) [week](#) [month](#)

0 votes 0 answers 1 view

[n Random rows for a given attribute - Postgres](#)

[sql](#) [postgresql](#)

44s ago Sup3rkiddo 49

1 vote 1 answer 14 views

[Branch descriptions in git, continued](#)

[git](#) [branch](#) [task-tracking](#)

48s ago manojds 16.2k

0 votes 0 answers 1 view

[Where is hostname defined for the anchor element?](#)

[javascript](#)

56s ago Chris Aaker 868

2 votes 1 answer 12 views

[User-defined Table Variables in MySQL 5.5?](#)

[mysql](#) [stored-procedures](#) [routines](#)

1m ago colonel_px 11

0 votes 2 answers 37 views

[Closing cfpdf tag with </cpdf> causes error](#)

[coldfusion](#) [coldfusion-8](#) [cfeclipse](#) [cpdf](#)

1m ago Jens Wegar 81

0 votes 0 answers 6 views

[cocoa memory leak by CGAffineTransform or by view](#)


[iphone](#) [objective-c](#) [cocoa](#) [memory-leaks](#) [leak](#)

1m ago EmptyStack 9,100

Hello World!

This is a collaboratively edited question and answer site for **professional and enthusiast programmers**. It's 100% free, no registration required.

[about »](#) [faq »](#)

 **CAREERS 2.0**
by stackoverflow

[Senior PHP Engineer](#)
Spreetales
Los Altos, CA; San Francisco, CA


[Front End Software Engineer](#)
@Rdio
Rdio
San Francisco, CA

[Senior Mobile Developer](#)
American Public Media
Oakland, CA

[Web Engineer](#)
Monkey Inferno
San Francisco, CA

DB specific

- Freebase / Wikidata



[Data](#) [Schema](#) [Apps](#) [Docs](#)

An entity graph of people, places and things, built by a community that loves open data.

Featured Data

Arts & Entertainment

Products & Services

Science & Technology

Society

Special Interests

















Sports


System

Time & Space

Transportation

All

Sort by	write activity	Jul 4	Aug 1	last week	Facts	Topics	Top User
Film	80 members		43K last week	5M	641K		
People	87 members		11K last week	7M	2M		
TV	35 members		8K last week	8M	1M		
Music	100+ members		771 last week	35M	10M		
Business	100+ members		431 last week	2M	611K		
Government	47 members		240 last week	532K	135K		
Location	52 members		223 last week	10M	999K		
Books	47 members		108 last week	29M	6M		



Google Refine

An open source power tool to fix, discover, experiment, connect and customize your data. [Learn more »](#)

What is Freebase?

Learn what an entity graph is, what kind of information it contains, and why you should add your data!

[Learn More »](#)

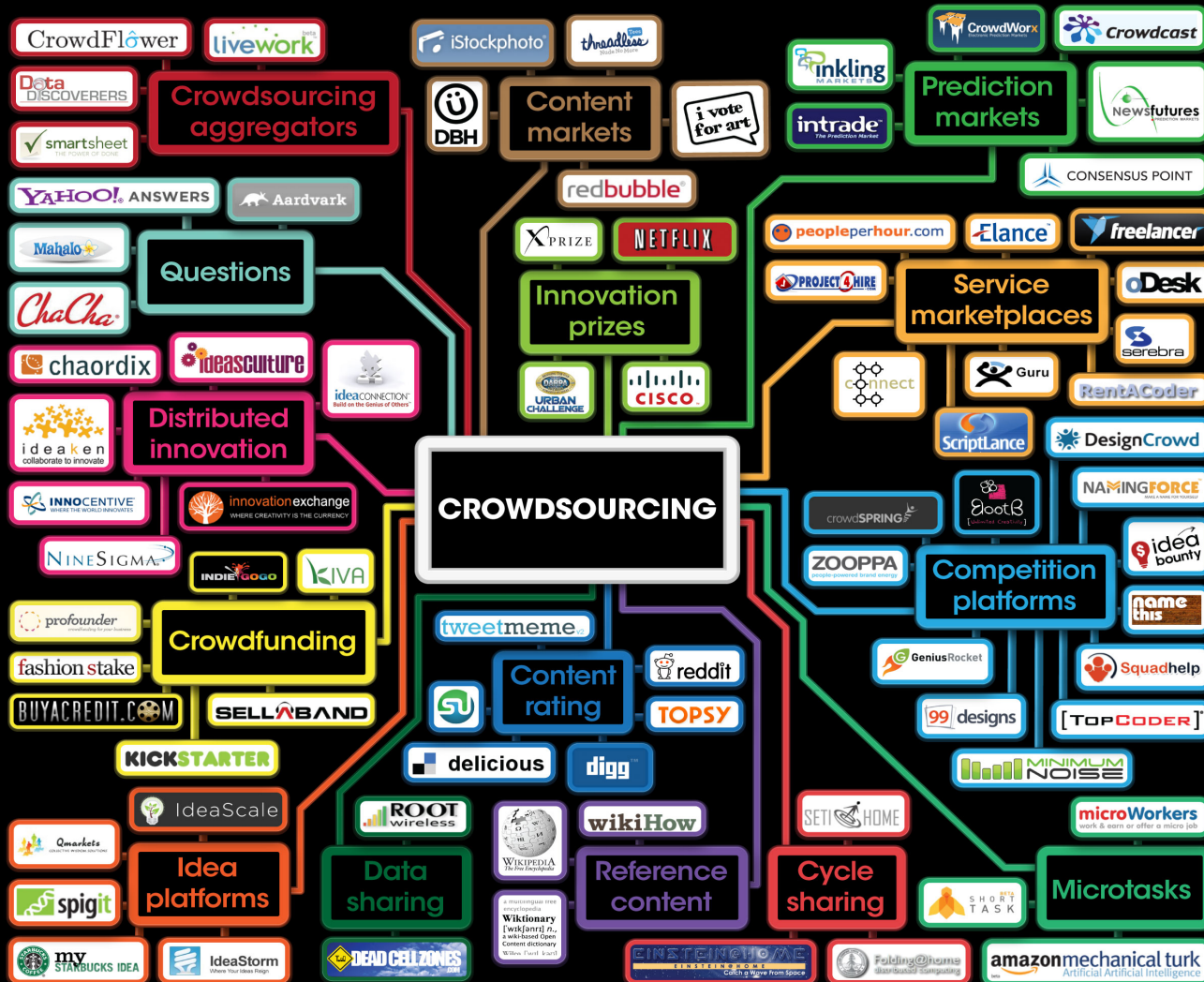
Freebase for Developers

- powerful queryable API
- JavaScript-based hosting framework
- libraries for other languages

[Learn More »](#)

The Way Industry Looks At It

CROWDSOURCING LANDSCAPE Beta v1



Common Crowdsourcing Tasks and Examples

3D object design	Thingiverse
Advertising	idea bounty
Business ideas	anadlist
Clothing	ClickAdvisor
Consumer research	ClickAdvisor
Crisis information	Ushahidi
Data analysis	SETI@HOME
Fact checking	PolitiFact.com
Graphic design	99 designs
Human reading	Capella
Investigative reporting	the guardian
Journalism	seed
Lending	zipo
Mapping	IMDb
Movie reviews	IMDb
Music	musikpitch
Observation	GALAXY ZOO
Patent research	Journalist Data Visualization
Philanthropy	philodoma
Political activism	MoveOn.org
Product design	MUJI
Proofreading	Bluewin
Scientific problems	fold
Software	rentacoder
Software development	rentacoder
Software testing	Test
Stock picking	marketocracy
Tagging	Google
Translation	facebook
Trends	TRENDHUNTER
TV programming	current
Word of mouth	BzzAgent
Writing and editing	crowdwork

For details, analysis, and discussion go to:

www.crowdsourcingresults.com

Advanced
Human
Technologies

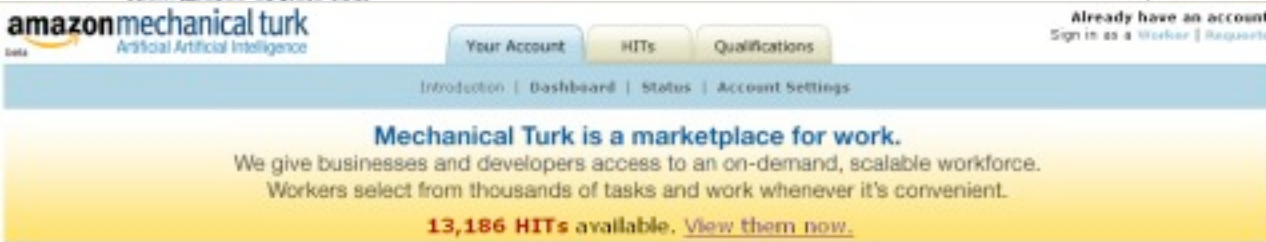
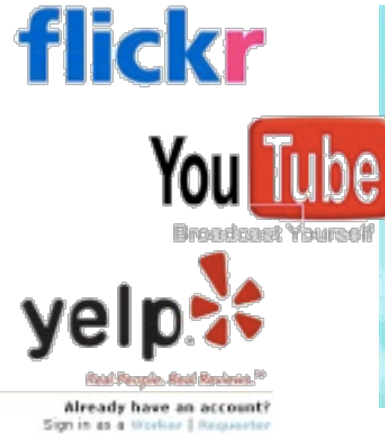
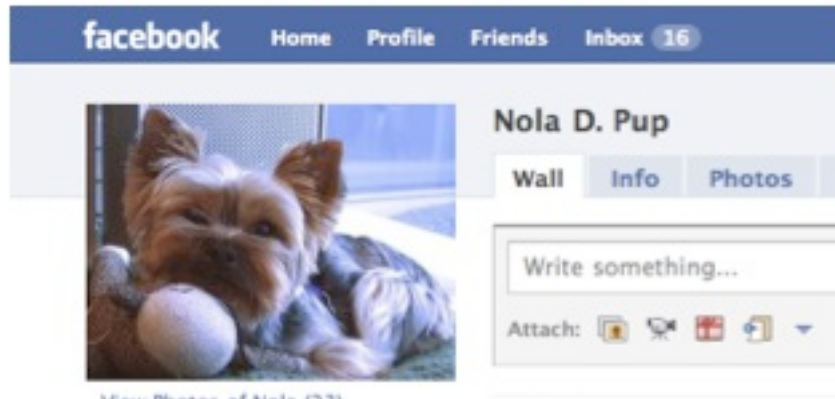
Published under a Creative Commons
Attribution-ShareAlike 2.5 License

Created by Ross Dawson
Design by Daniil Alexandrov

Taxonomies

- Doan, Halevy, Ramakrishnan; (Crowdsourcing) CACM 4/11
 - nature of collaboration (implicit vs. explicit)
 - architecture (standalone vs. piggybacked)
 - must recruit users/workers? (yes or no)
 - What do users/workers do?
- Bederson & Quinn; (Human Computation) CHI '11
 - Motivation (Pay, Altruism, Enjoyment, Reputation)
 - Quality Control (mechanisms)
 - Aggregation (how are results combined?)
 - Human Skill (Visual recognition, language, ...)
 - ...

Participatory Culture - Explicit



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



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. [Get started.](#)

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



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



Participatory Culture – Implicit

John Murrell: GM SV 9/17/09

...every time we use a Google app or service, we are working on behalf of the search sovereign, creating more content for it to index and monetize or teaching it something potentially useful about our desires, intentions and behavior.



OCR errors: reCAPTCHA



1629 capplots | 1960-73 batizak | III, Quichut

RICHARD, Redfully | ansunjon Rupprecht | Daterny Nature

drôle Deratial | than Morgicar | Golightly, Bylemor

Games with a Purpose

- Tasks leveraging common human skills, appealing to large audiences
 - Selection of domain and task more constrained in games to create typical UX
- Tasks decomposed into smaller units of work to be solved independently
- Complex workflows
 - Creating a casual game experience vs. patterns in microtasks
 - Single vs. multi-player

Gamification

- A human-based computation technique in which a computational process performs its function by outsourcing certain steps to humans **in an entertaining way**

How to implement gamification

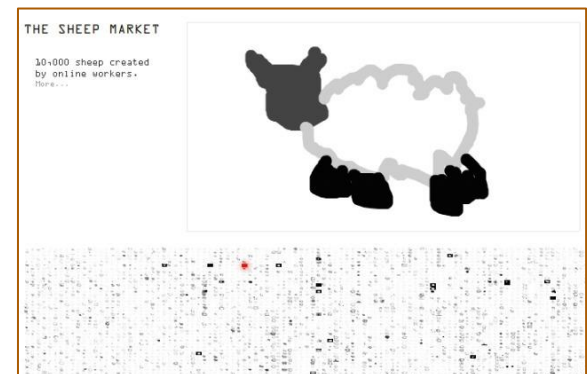
- **Cosmetic:** adding game-like visual elements or copy (usually visual design or copy-driven)
- **Accessory:** wedging in easy-to-add-on game elements, such as **badges** or adjacent products (usually marketing-driven)
- **Integrated:** more subtle, deeply integrated elements like **% complete** (usually interaction-design driven)
- **Basis:** making the entire offering a **game** (usually product-driven)

Paid Micro-task Crowdsourcing

Example use of micro-task crowdsourcing

- Relevance judgments
- Ontologies
- Sentiment Analysis in Social Media
- <http://www.thesheepmarket.com/>

'draw a sheep facing to the left.'



Background

A Crowdsourcing Platform allows **requesters** to publish a crowdsourcing request (*batch*) composed of multiple tasks (*HITs*)

Programmatically Invoke the crowd with APIs or using a website

Workers in the crowd complete tasks and obtain a monetary reward



Microtask Aggregators



CrowdFlower

SOLUTIONS

SELF-SERVICE

NEWS & EVENTS

BLOG

Business Listing Verification

Search Relevance

Product Categorization

Content Generation

Custom Solutions

Enterprise Crowdsourcing Solutions

CrowdFlower's technology engages a global workforce to solve your large-scale data problems.



**Business Listing
Verification**

Correct inaccurate
business listings.



**Search
Relevance**

Assess the relevance of
your search results.



**Product
Categorization**

Categorize large data
sets.



**Content
Generation**

Get quality content in
real time.

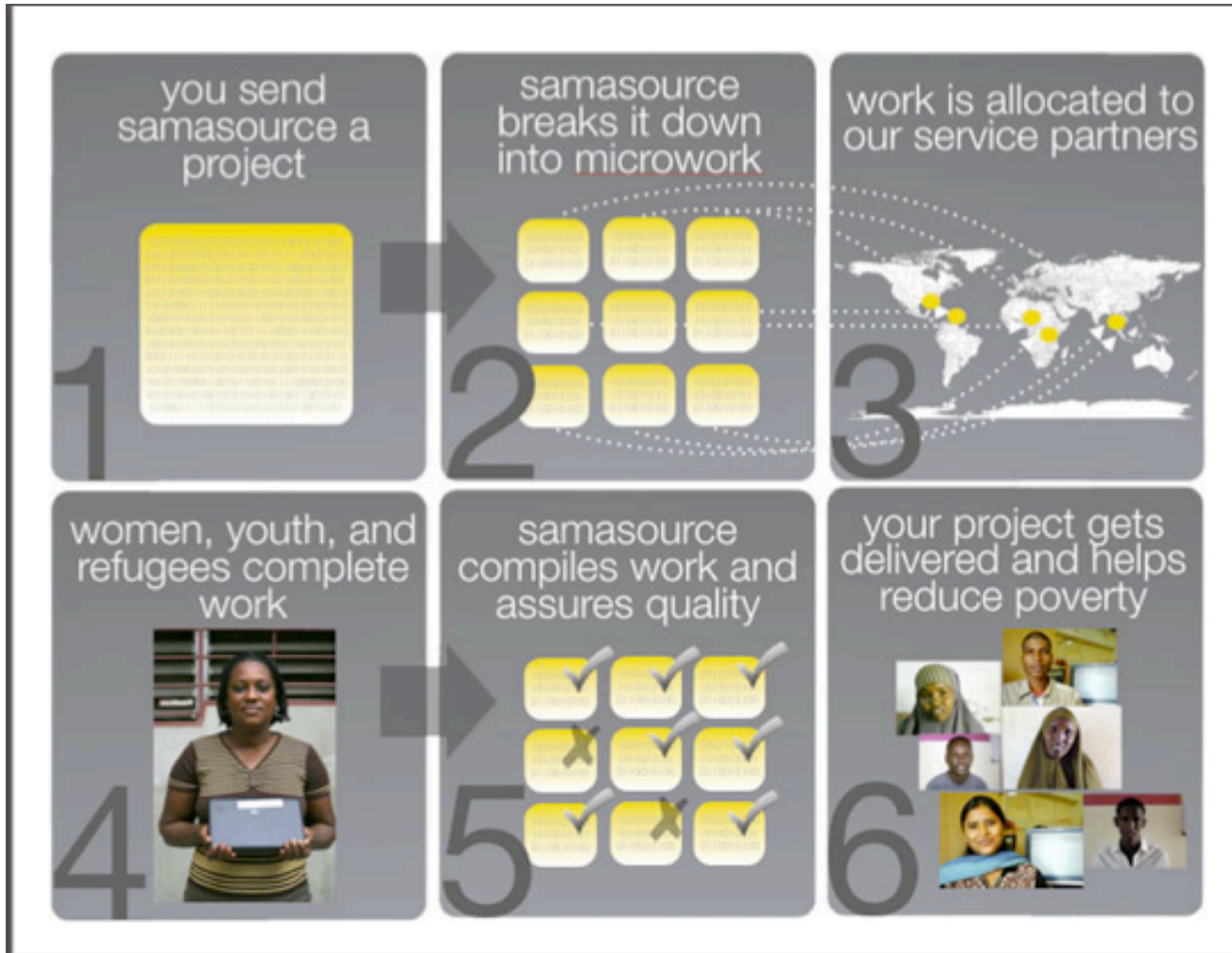


**Custom
Solutions**

Tailored solutions to fit
your needs.

<http://www.businesswire.com/news/home/20120207005761/en/CrowdFlower-Reports-Revenue-300-Year-Year-300>

Samasource.org



Case-Study: Amazon MTurk

- Micro-task crowdsourcing marketplace
- On-demand, scalable, real-time workforce
- Online since 2005 (still in “beta”)
- Currently the most popular platform
- Developer’s API as well as GUI

Amazon MTurk



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



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



Amazon MTurk

- Requesters create tasks (HITs)
- The platform takes a fee (30% of the reward)
- Workers preview, accept, submit HITs
- Requesters approve, download results
- If the results are approved, workers are paid

MTurk is a Marketplace for HITs

All HITs

1-10 of 3454 Results

Sort by: HITs Available (most first)

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

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

Provide Information about a Product Requester: requester	HIT Expiration Date: May 23, 2015 (4 weeks 1 day) Time Allotted: 25 minutes	Reward: \$0.05 HITs Available: 11526	View a HIT in this group
Product Attribute Tagging - April 17th Please read the instructions Requester: slee	HIT Expiration Date: May 23, 2015 (4 weeks 2 days) Time Allotted: 60 minutes	Reward: \$0.03 HITs Available: 23887	View a HIT in this group
Inv_B_2 Requester: rohzi0d	HIT Expiration Date: May 22, 2015 (4 weeks 1 day) Time Allotted: 48 minutes	Reward: \$0.00 HITs Available: 19740	View a HIT in this group
Geo Result Relevance-Tue Apr 21 10:40:14 PDT 2015 Requester: Amazon Requester Inc.	HIT Expiration Date: May 22, 2015 (4 weeks 1 day) Time Allotted: 60 minutes	Reward: \$0.00 HITs Available: 10734	View a HIT in this group
Type the text from the Images, carefully. Productivity and bonuses guaranteed. Requester: CopyText Inc.	HIT Expiration Date: Apr 30, 2015 (6 days 23 hours) Time Allotted: 10 minutes	Reward: \$0.01 HITs Available: 10590	View a HIT in this group
Transcribe up to 25 Seconds of Media to Text - Earn up to \$0.12 per HIT! Requester: Crowdsurf Support	HIT Expiration Date: Apr 21, 2016 (51 weeks 6 days) Time Allotted: 15 minutes	Reward: \$0.08 HITs Available: 6702	View a HIT in this group
Fun and Fast Fashion Tagging Requester: gavin	HIT Expiration Date: Apr 28, 2015 (5 days 11 hours) Time Allotted: 60 minutes	Reward: \$0.02 HITs Available: 6460	View a HIT in this group
Geo Result Relevance-Wed Apr 08 14:30:08 PDT 2015 Requester: Amazon Requester Inc.	HIT Expiration Date: May 10, 2015 (2 weeks 2 days) Time Allotted: 60 minutes	Reward: \$0.00 HITs Available: 6182	View a HIT in this group
Transcribe up to 25 Seconds of General Content to Text - Earn up to \$0.14 per HIT! Requester: Crowdsurf Support	HIT Expiration Date: Apr 21, 2016 (51 weeks 6 days) Time Allotted: 15 minutes	Reward: \$0.09 HITs Available: 6043	View a HIT in this group
!Whac-a-mole by Gaze (hard mode) ! Play a 1min eye tracking game in the web browser! 0416 Requester: px	HIT Expiration Date: Apr 23, 2015 (8 hours 40 minutes) Time Allotted: 60 minutes	Reward: \$0.10 HITs Available: 4682	View a HIT in this group

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

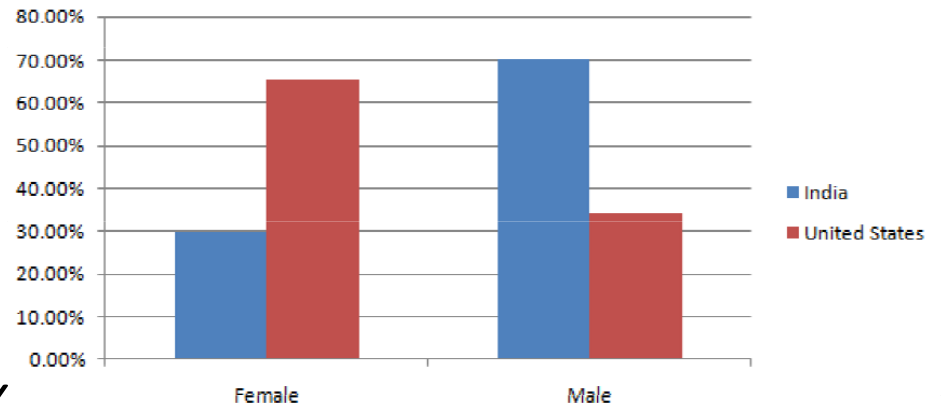
Demographics of MTurk workers in 2009

Country of residence

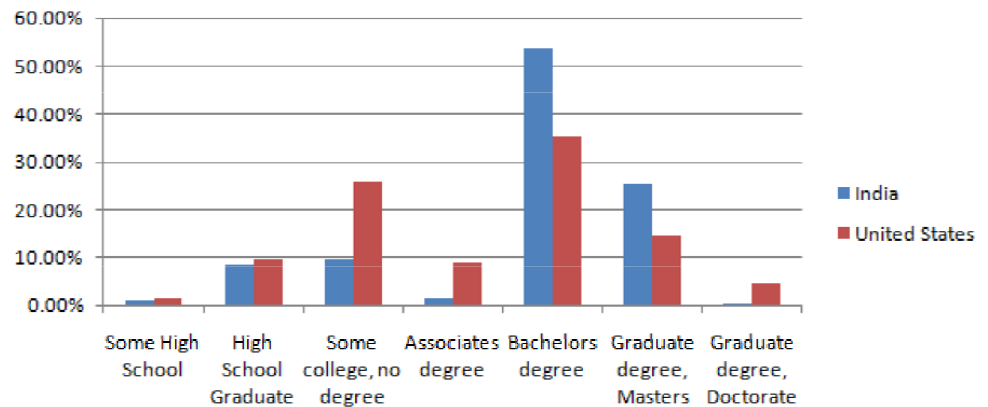
- United States: 46.80%
- India: 34.00%
- Miscellaneous: 19.20%

2013 Statistics:
1M workers
10% active

Gender Breakdown

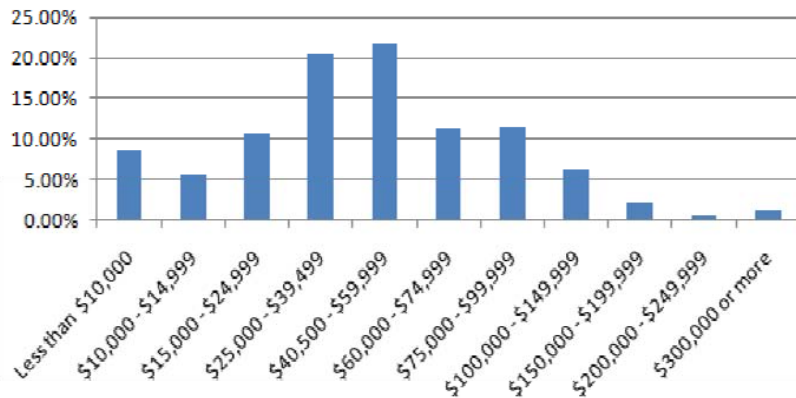


Education Level

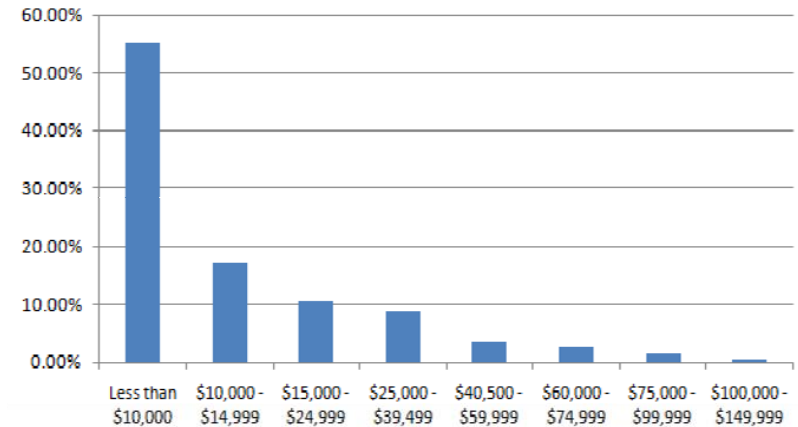


Demographics of MTurk workers in 2009

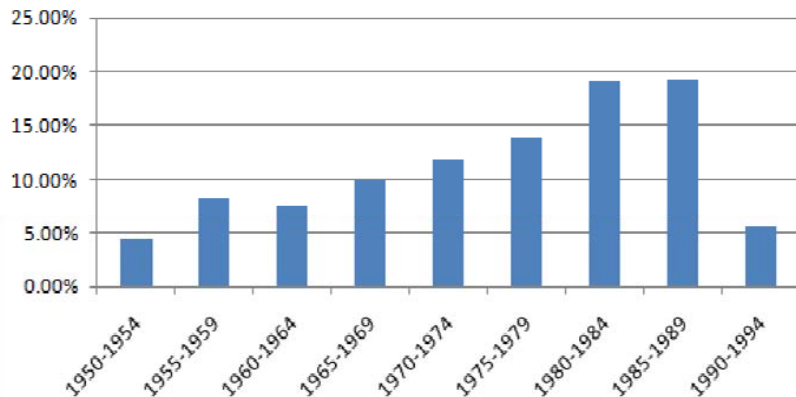
Household Income for US workers



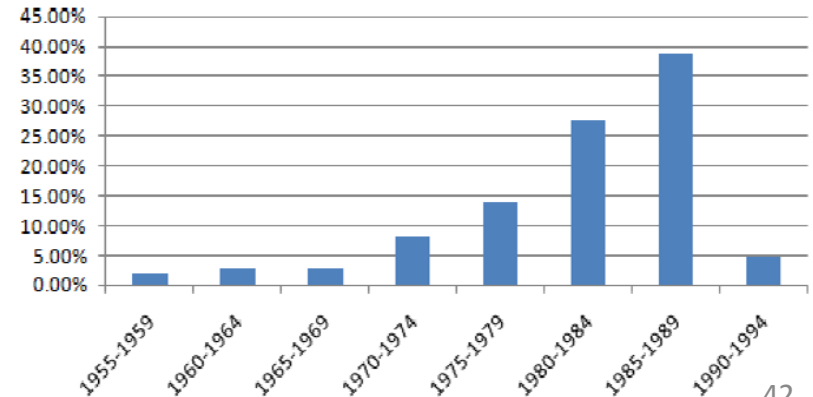
Household Income for Indian workers



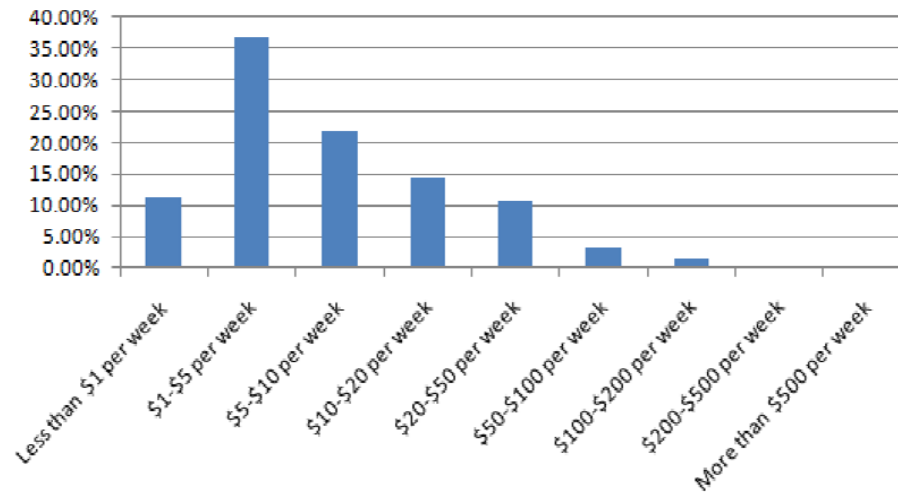
Year of Birth for US workers



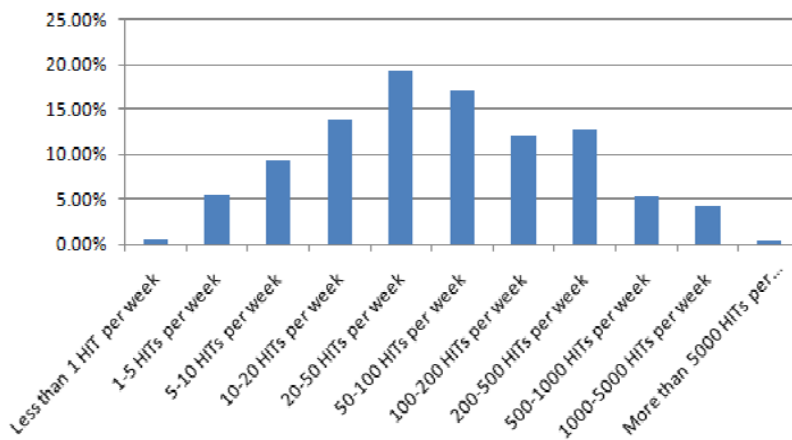
Year of Birth for Indian workers



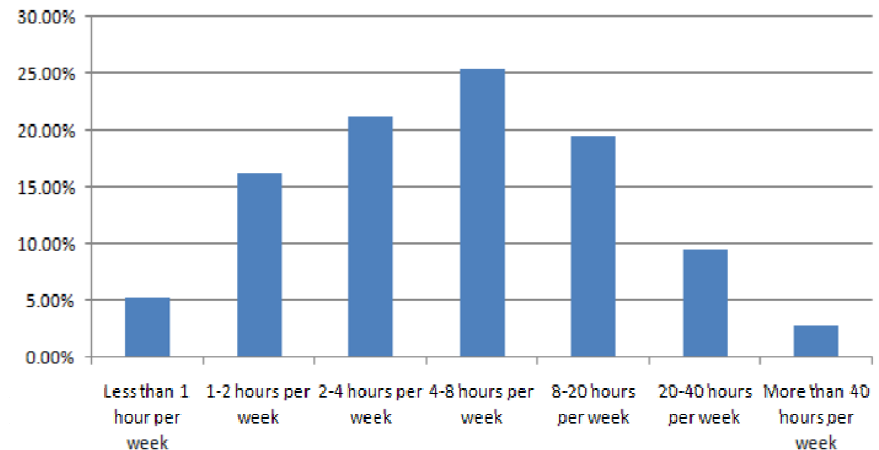
Demographics of MTurk workers in 2009



Number of HITs completed per week



Time spent on Mechanical Turk per week



<http://www.mturk-tracker.com/>

mturk-tracker.com



General

Arrivals

Completions

Top requesters

Active requesters

Demographics

Search

API

Old version

04/16/2016

05/16/2016

Market statistics

HIT groups posted

HITs posted

Rewards posted



For bugs reports or feature requests, please contact [Panos Ipeirotis](#)

If you want to cite this website, please cite the papers: "The Dynamics of Micro-Task Crowdsourcing: The Case of Amazon MTurk", D. Djelle Eddine, M. Catasta, G. Demartini, P. Ipeirotis, and P. Cudre-Mauroux, Proceedings of the 24th International Conference on World Wide Web", and "Analyzing the Amazon Mechanical Turk Marketplace", P. Ipeirotis, ACM XRDS, Vol 17, Issue 2, Winter 2010, pp 16-21."

mturk-tracker.com

- Collects metadata about each visible batch (Title, description, rewards, required qualifications, HITs available etc)
- Records batch progress (every ~20 minutes)

We note that the tracker reports data periodically only and does not reflect fine-grained information (e.g., real-time variations)

Djellel Eddine Difallah, Michele Catasta, Gianluca Demartini, Panagiotis G. Ipeirotis, and Philippe Cudré-Mauroux. **The Dynamics of Micro-Task Crowdsourcing -- The Case of Amazon MTurk**. In: 24th International Conference on World Wide Web (WWW 2015), Research Track. Firenze, Italy, May 2015.

Top requesters last month

Top-1000 Requesters, report for April 16, 2016 to May 16, 2016

Requester name	hits	reward
Speechpad	23857	\$172,994.63
Percy Liang	883	\$7,320.48
Princeton Vision	51187	\$5,762.44
Stanford GSB Behavioral Lab	3749	\$2,110.70
Chris Callison-Burch	8157	\$2,064.29
RC.org Mechanical Turk	6591	\$2,011.33
VacationrentalAPI	399	\$1,373.50
Med Expertise	869	\$1,303.50
Bluejay Labs	13613	\$1,288.59
YL Testing	1051	\$1,236.83

SLAs are expensive



[HOW IT WORKS](#) | [SERVICES](#) | [BUZZ](#) | [PRICING](#) | [FAQ](#) | [JOBS](#)

[REGISTER](#) | [SIGN IN](#)

SIMPLY THE BEST HUMAN-GENERATED TRANSCRIPTIONS. DELIVERED ON TIME, EVERY TIME. GUARANTEED!

\$1.00

PER MINUTE OF AUDIO OR VIDEO
DELIVERED IN
ONE WEEK
GUARANTEED

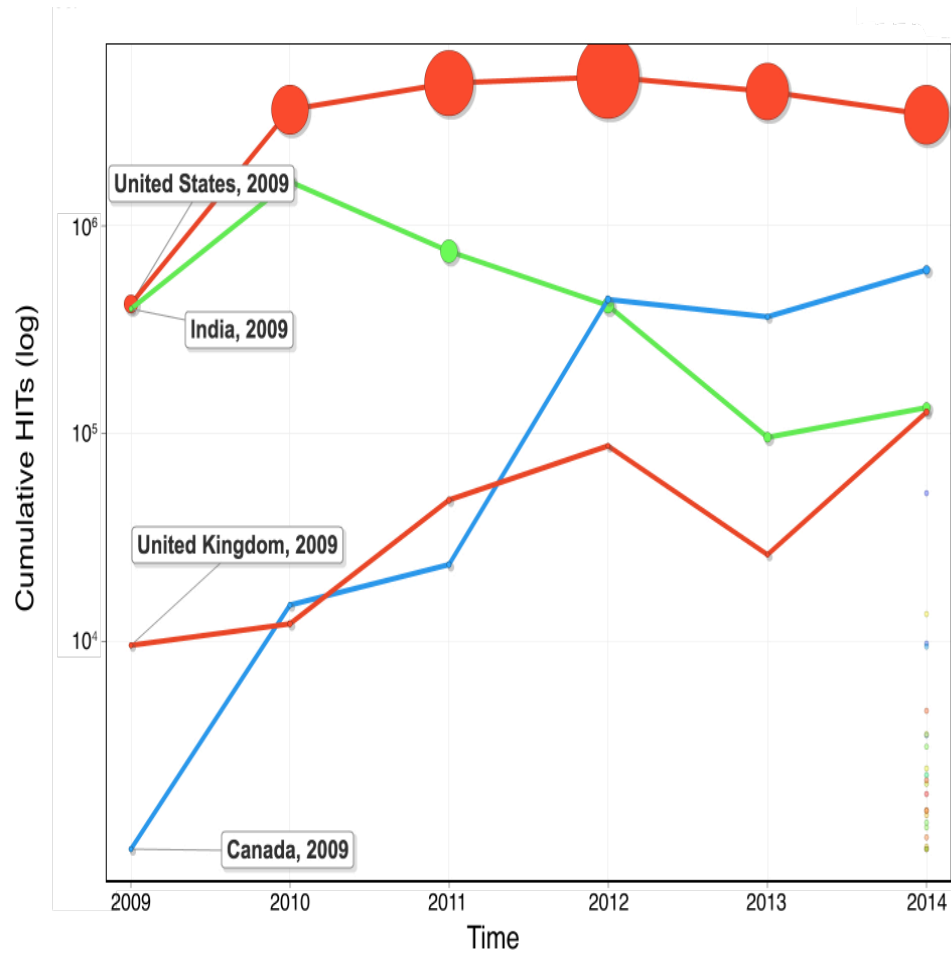
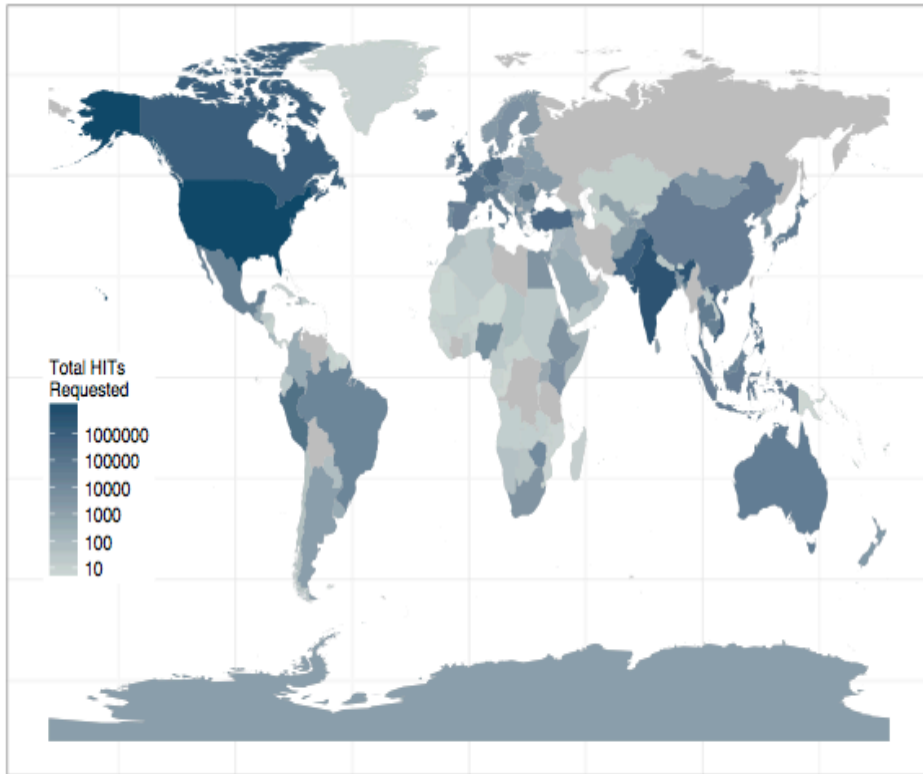
\$1.50

PER MINUTE OF AUDIO OR VIDEO
DELIVERED IN
48 HOURS
GUARANTEED

\$3.00

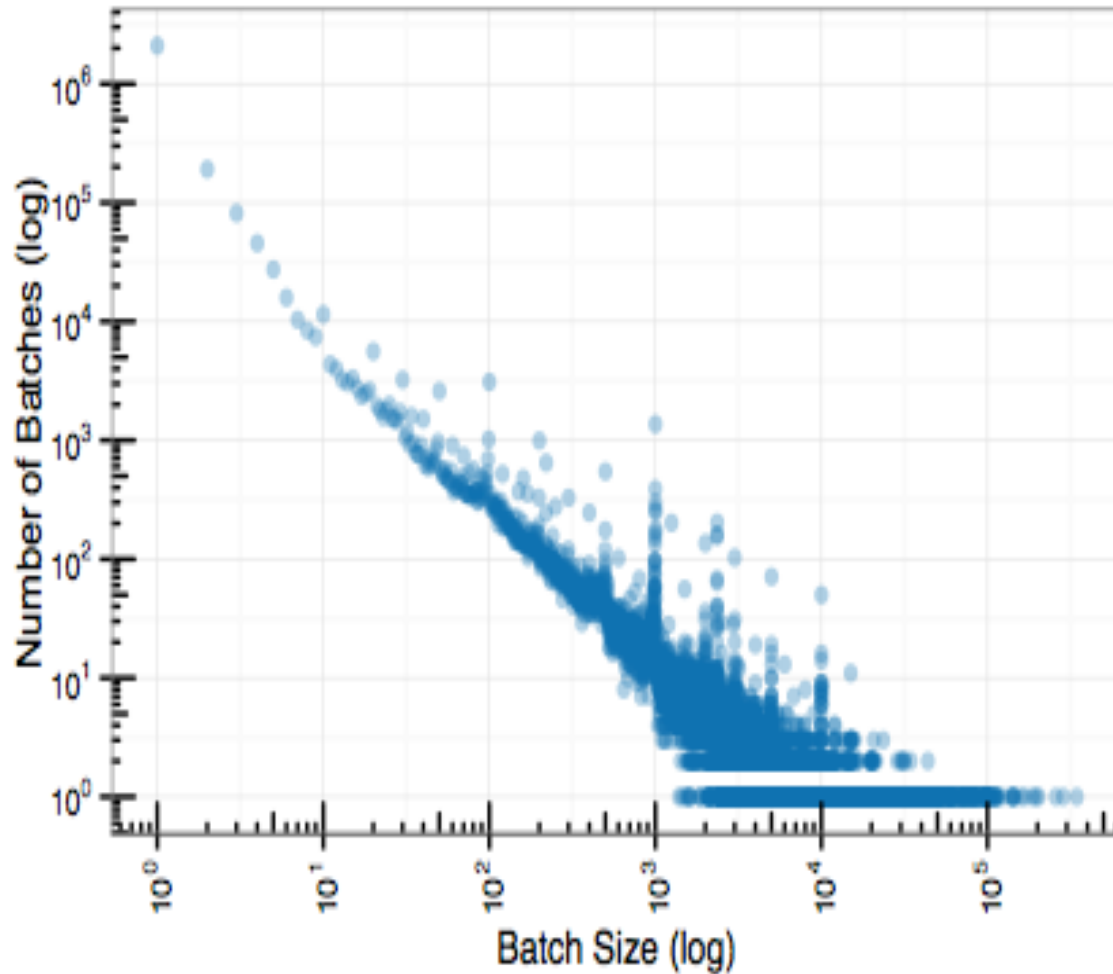
PER MINUTE OF AUDIO OR VIDEO
DELIVERED IN
24 HOURS
GUARANTEED

Country-Specific HITs



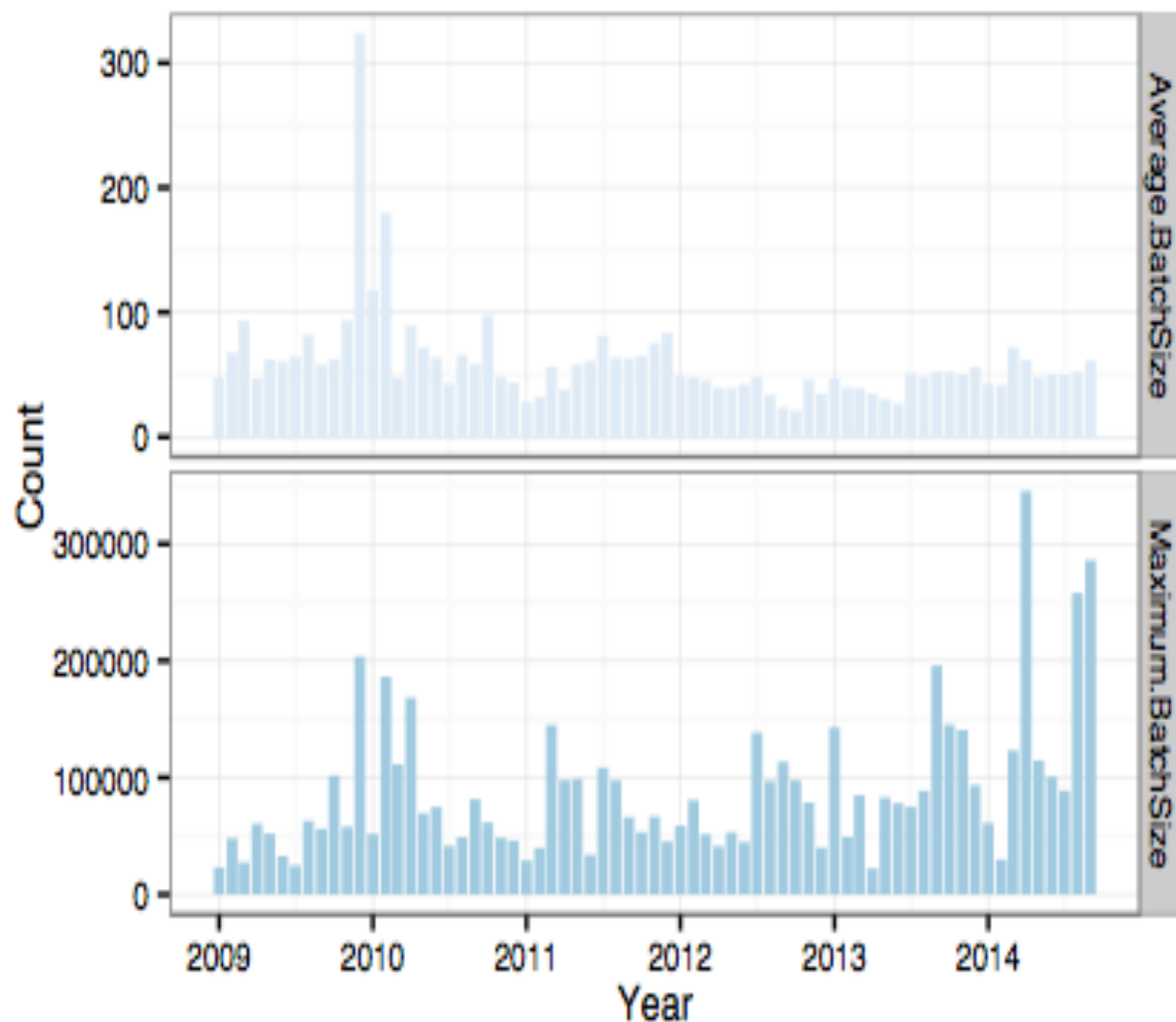
Workers from US, India and Canada are the most sought after.

Distribution of *Batch Size*



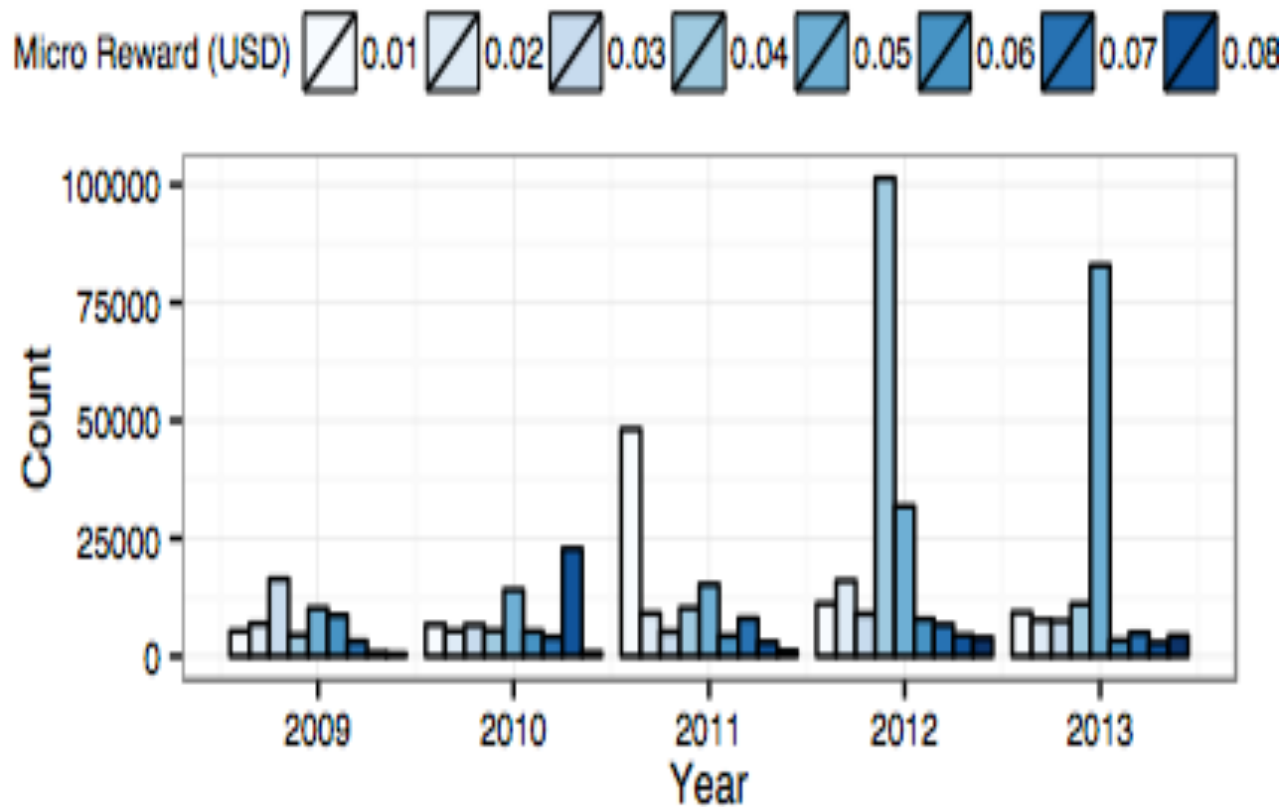
“Power-law”

Evolution of Batch Sizes



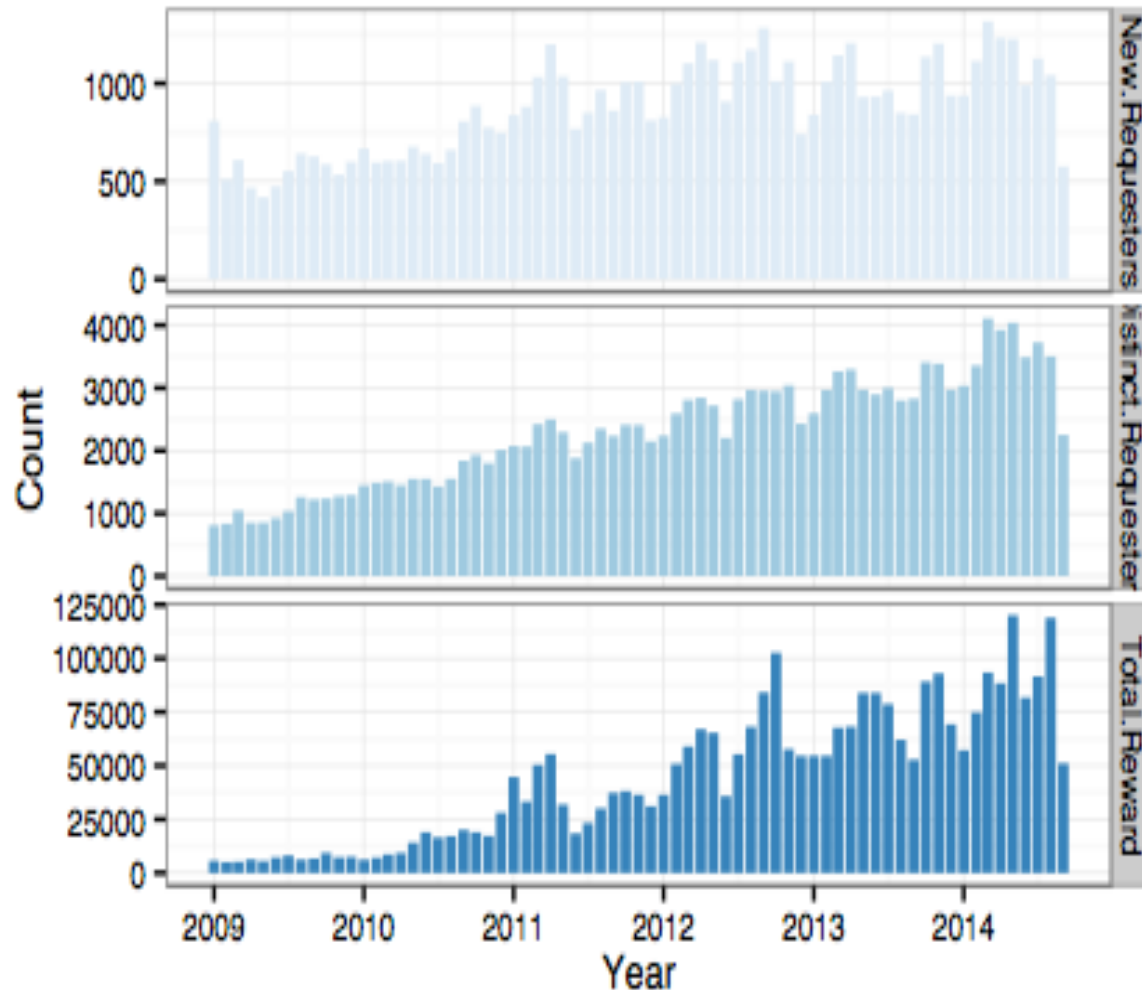
Very large batches
start to appear

HIT Pricing



5-cents is the new
1-cent

Requesters and Reward Evolution



Increasing number of New
and Distinct Requesters

HIT Classes

Classify HITs into types (Gadiraju et. al 2014)

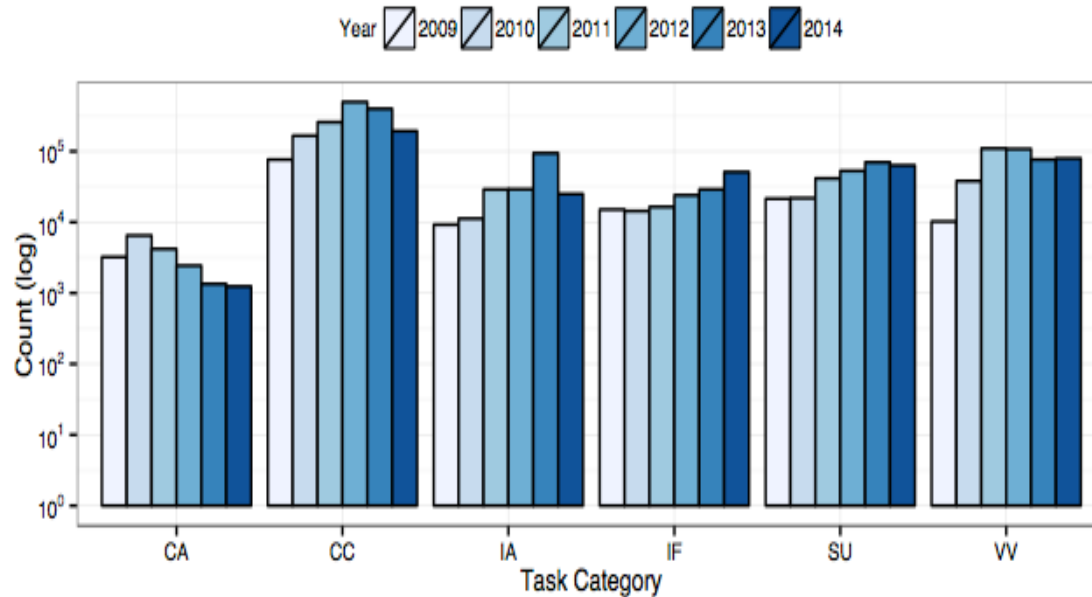
- Information Finding (IF)
- Verification and Validation (VV)
- Interpretation and Analysis (IA)
 - Content Creation (CC)
 - Surveys (SU)
 - Content Access (CA)

Supervised Classification With the Crowd

We trained a Support Vector Machine (SVM) model

- HIT title, description, keywords, reward, date, allocated time, and batch size
- Created labeled data on Mturk for 5,000 HITs uniformly sampled HITs
- Our HIT used 3 repetitions
 - Consensus reached for 89% of the tasks
- 10-fold cross validation
 - Precision of 0.895
 - Recall of 0.899
 - F-Measure of 0.895
- Then, a large-scale classification for all 2.5M HITs

Distribution of HIT Types

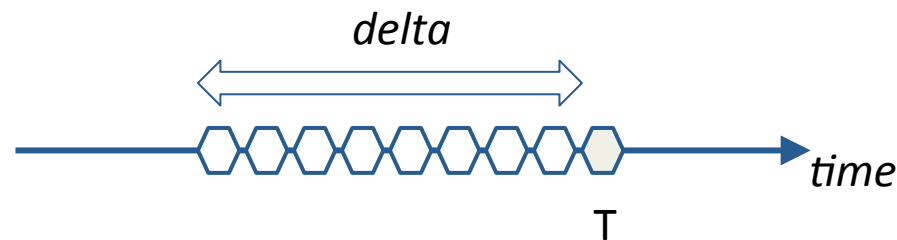


Less Content Access batches

Content Creation being the most popular

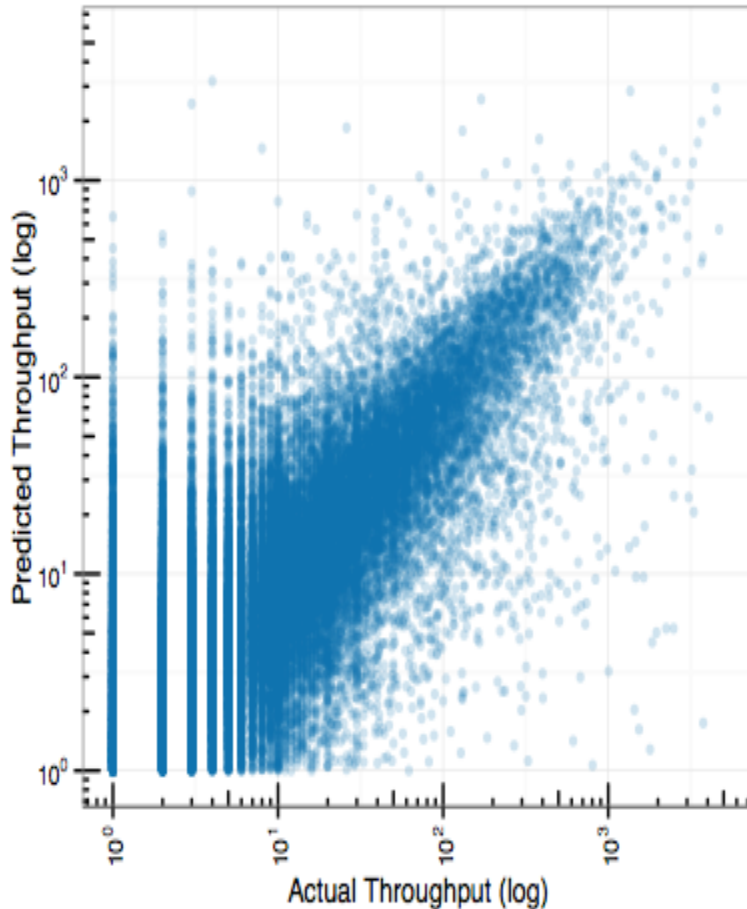
Batch Throughput Prediction

- Predict batch throughput at time T by training a Random Forest Regression model with samples taken in $[T-\text{delta}, T)$ time span
- 29 Features (including the Type of the Batch)
- Hourly Data in range [June-October] 2014
- We sampled 50 times points for evaluation purposes



We are interested in cases where prediction works reasonably

Predicted vs. Actual Batch Throughput (delta=4 hours)



Prediction Works best for larger batches having large momentum

Significant Features

- What features contribute best when the prediction works reasonably
- We proceed by feature ablation
 - Re-run prediction by removing 1 feature at a time.
 - 1000 samples

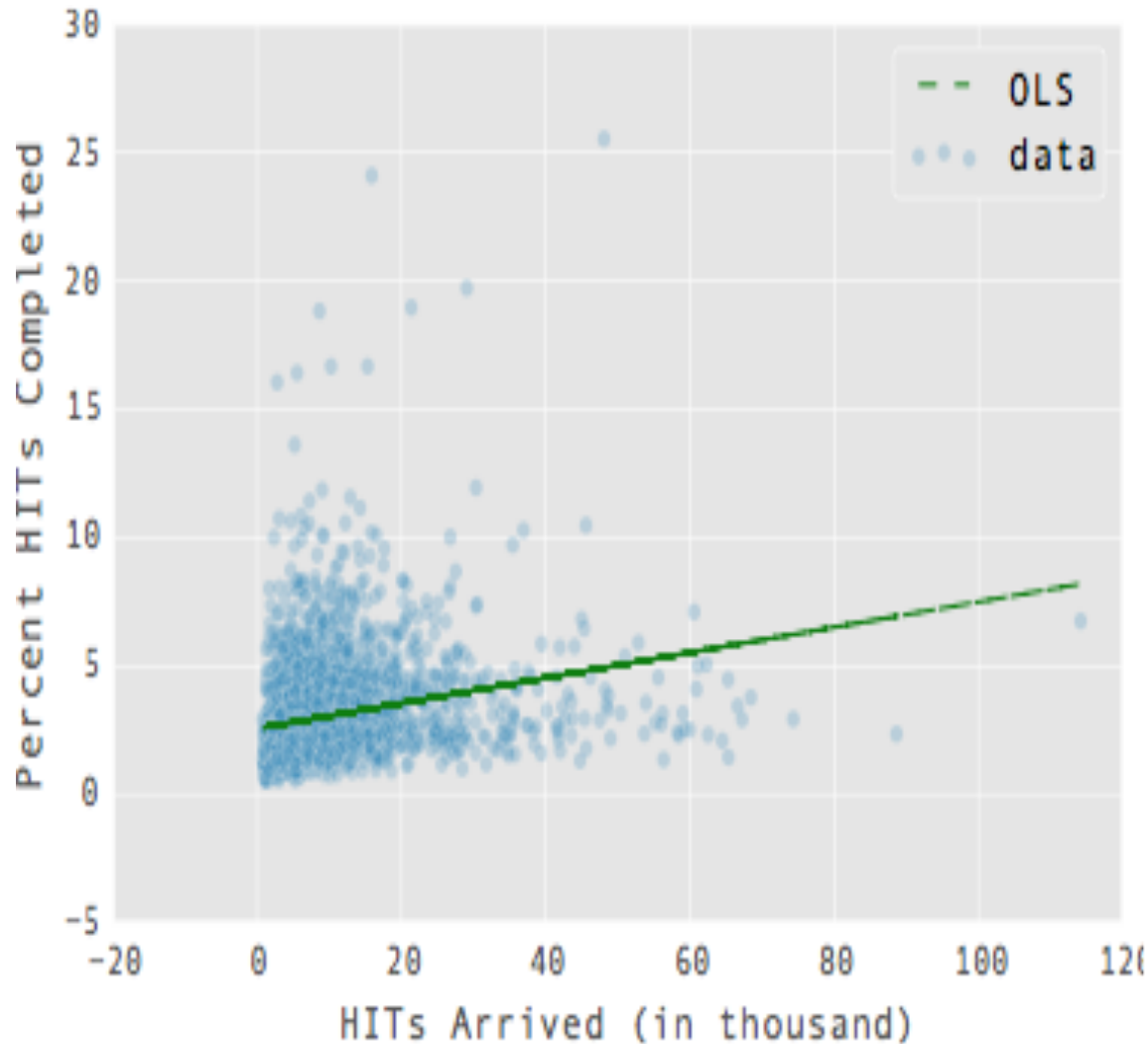
HITs_Available (Number of tasks in the batch)

Age_Minutes (how long ago the batch was created)

Supply Elasticity

How does the market reacts when new tasks arrive on the platform?

Supply Elasticity



Intercept = 2.5
Slope = 0.5%

20% of new work gets
completed within an
hour

Summary

- HIT reward has increased over time
 - Audio transcription is the most popular task
 - Demand for Indian workers has decreased
 - Surveys are most popular for US workers
 - 1000 new requesters per month join
 - 10K new HITs arrive and 7.5K HITs get completed every hour
-
- Check #mturkdynamics for the main findings

Crowdsourcing for Information Retrieval Evaluation

- Easy, cheap and fast labeling
- Ready-to use infrastructure – MTurk payments, workforce, interface widgets – CrowdFlower quality control mechanisms, etc.
- Allows early, iterative, frequent experiments – Iteratively prototype and test new ideas – Try new tasks, test when you want & as you go
- Proven in major IR shared task evaluations
 - CLEF image, TREC, INEX, WWW/Yahoo SemSearch

Gamification of IR Evaluation

- GeAnn: <http://www.geann.org/>
- Relevance judgments with Gamification:
 - Text relevance
 - Image relevance

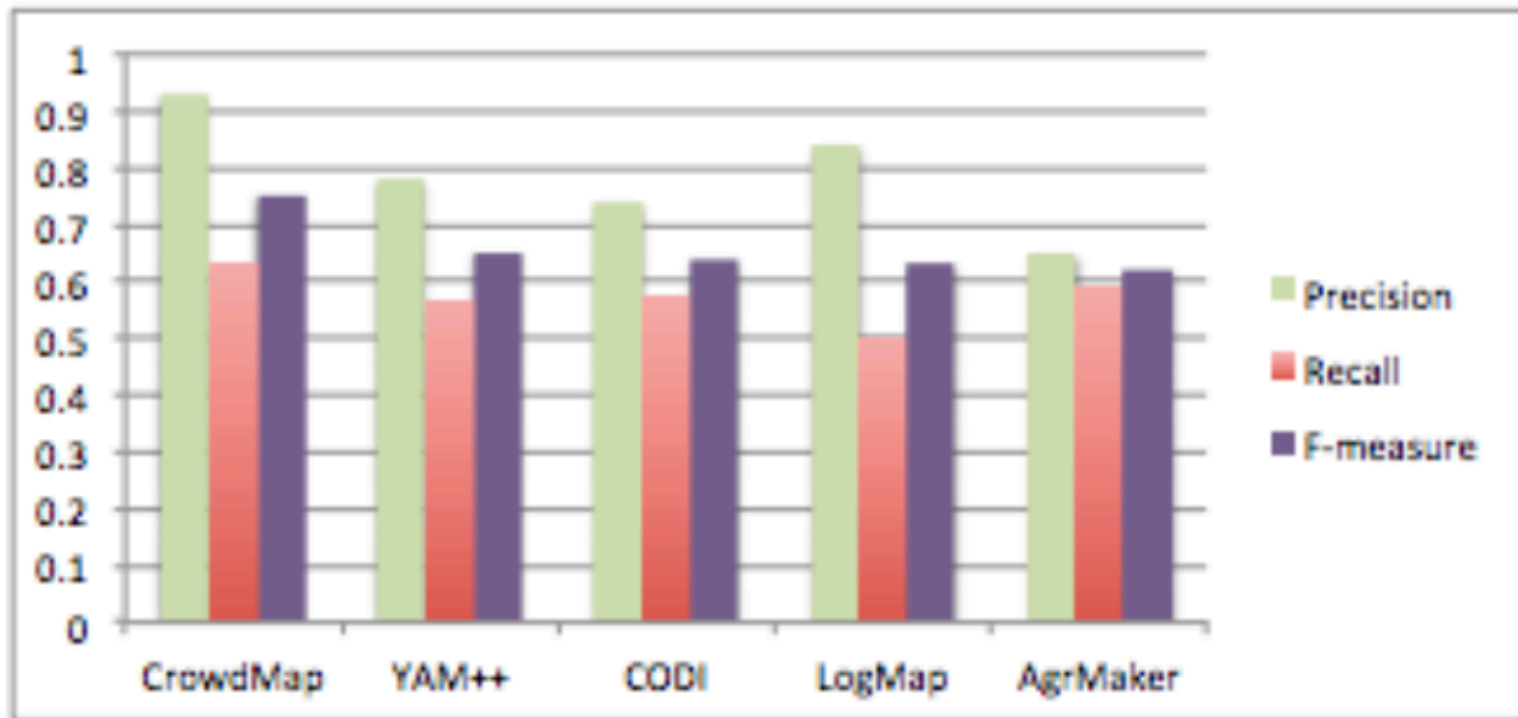
Quality through Flow and Immersion: **Gamifying Crowdsourced Relevance Assessments**. Eickhoff, C., C. G. Harris, A. P. de Vries, and P. Srinivasan. SIGIR 2012.

Crowdsourcing Ontology Mapping

- Find a set of mappings between two ontologies
- Micro-tasks:
 - Verify/identify a mapping relationships:
 - Is concept A the same as concept B
 - A is a kind of B
 - B is a kind of A
 - No relation

Crowdsourcing Ontology Mapping

- Crowd-based outperforms purely automatic approaches



Crowdsourcing Ontology Engineering

- Ask the crowd to create/verify subClassOf relations
 - “Car” is a “vehicle”
- Does it work for domain specific ontologies?
 - A “protandrous hermaphroditic organism” is a “sequential hermaphroditic organism”
- Workers perform worse than experts
- Workers presented with concept definitions perform as good as experts

CrowdDB

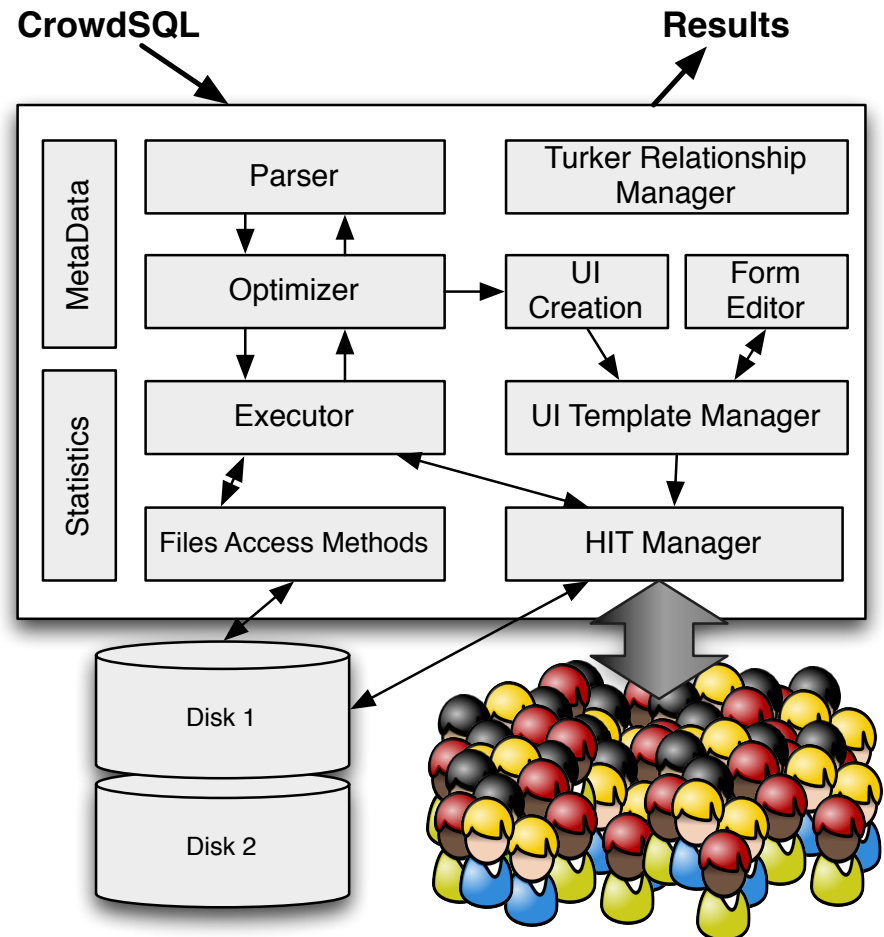
Use the crowd to answer DB-hard queries

Where to use the crowd:

- Find missing data
- Make subjective comparisons
- Recognize patterns

But not:

- Anything the computer already does well



M. Franklin, D. Kossmann, T. Kraska, S. Ramesh and R. Xin .

CrowdDB: Answering Queries with Crowdsourcing, *SIGMOD 2011* ⁶⁷

And more...

- NLP / Entity Recognition and Linking
- Search (Slow Search, CrowdQ)
- Machine Learning (training data)

Gianluca Demartini et al. ZenCrowd: Leveraging Probabilistic Reasoning and Crowdsourcing Techniques for Large-Scale Entity Linking. In: WWW 2012

Oluwaseyi Feyisetan, et al. Towards Hybrid NER: A Study of Content and Crowdsourcing-Related Performance Factors. ESWC 2015

Jaime Teevan et al. Slow search. Commun. ACM 57(8): 36-38 (2014)

Gianluca Demartini et al. CrowdQ: Crowdsourced Query Understanding. In: 6th Biennial Conference on Innovative Data Systems Research (CIDR 2013)

Barzan Mozafari et al. Scaling Up Crowd-Sourcing to Very Large Datasets: A Case for Active Learning. PVLDB 8(2): 125-136 (2014)

Microtask vs Macrotask

Macrotask

S M A R T Supermarket
Visit us on the Internet
Thank you for shopping

K BAR	0.22
CHOCOLATE ORANGE	0.19
CADBURY SNACK	0.22
HOBNOB	0.24
SKITTLES	0.22
WALLS CORNETTO	0.25
BURTONS MARYLAND	0.19
JAFFA CAKE	0.18
FRUIT SALAD	0.26
FREDDO	0.23

What is the **total cost** of all the items on the receipt?
Do not type in the dollar sign.

Type in the total.

Microtask

Practice Receipt

Add the cost of the next item below to the previous total.
Do not type in the dollar sign (1 of 10).

Prev Total: 0.00

+ New Item: K BAR 0.22

= Total: Total?

Microtask vs Macrotask

- Longer to perform a task using microtasks than macro- tasks.
- Micro-task: higher quality work, easier to complete, robust to interruption
- Task decomposition may be difficult

Summary

- Crowdsourcing is growing in popularity
- It is used both in industry and academia
- For a number of applications across disciplines
- Next:
 - How to make sure we get quality results back from a crowdsourcing platforms?
 - How can social media benefit crowdsourcing and how crowdsourcing can help social media research?