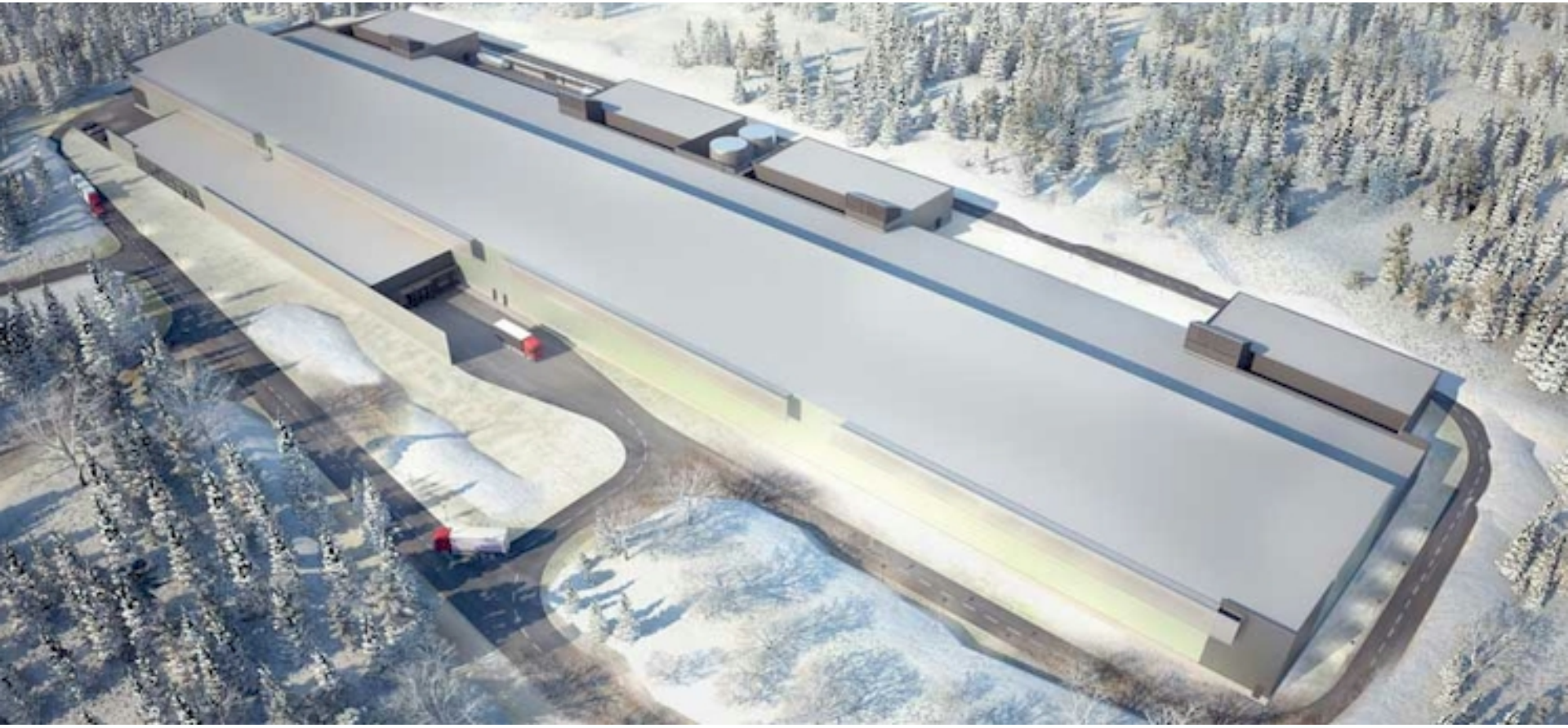
Human Computation: Machines are not enough

Gianluca Demartini
Information School, University of Sheffield
Ignite Academy 2014

Data is huge

- Banks, city councils, governments, shops, etc.
- Facebook processes 750TB/day of data
  - 48k iPhones every day
  - 7PB of photo storage / month
- This requires computers (a lot of them)

Facebook Data Center (Sweden)



Machines

- Google has around 900,000 servers (260 million watts) 0.01% global energy
- CERN Large Hadron Collider 180MW
- The flux-capacitor needs 1.21 GW (just 5xGoogle)

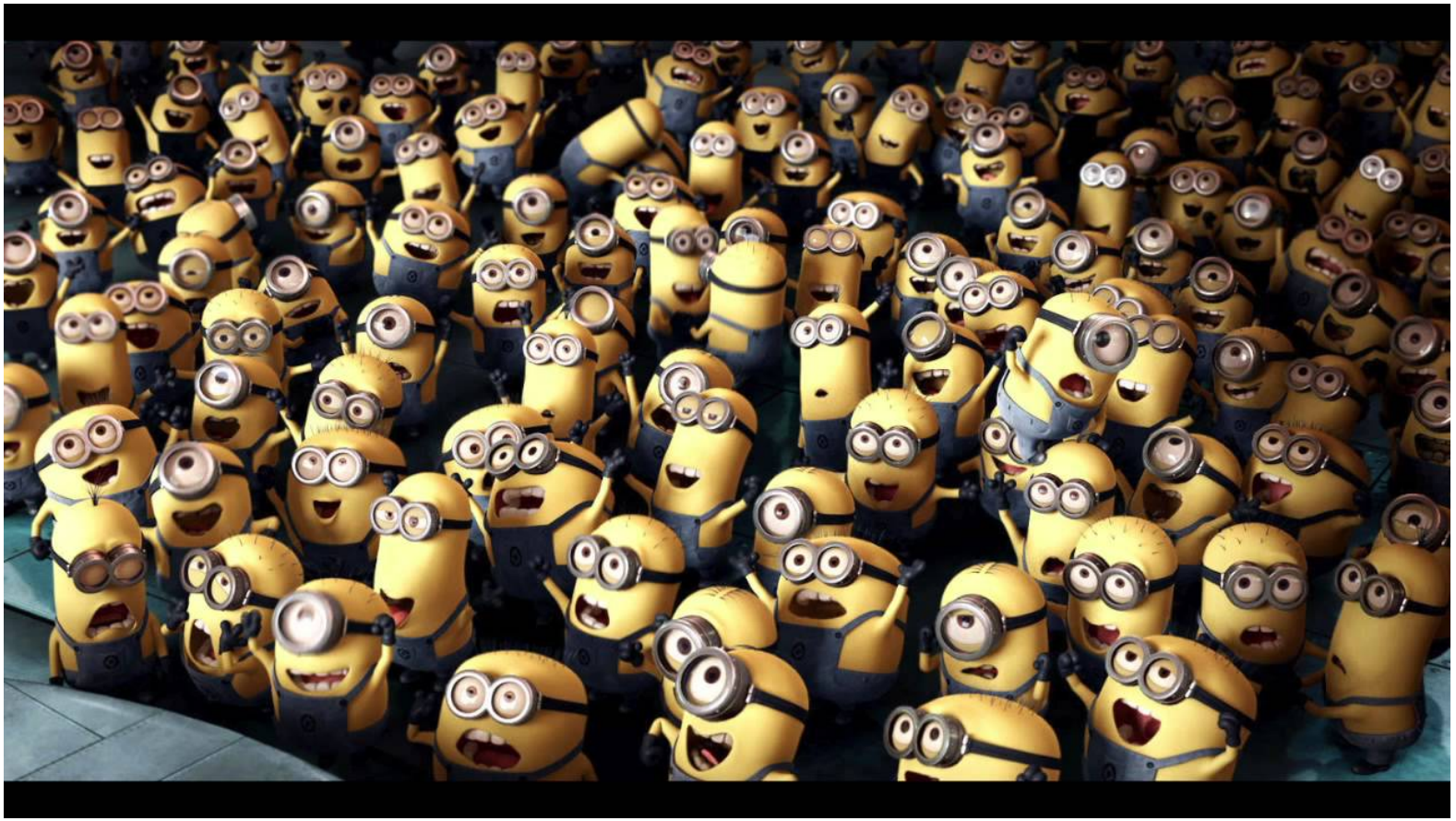
Data from July 2013

Well, machines are not enough

Human Computation

- Combine both the scalability of machines as well as the ability of human intelligence
- Outsource tasks from computers to a crowd of people

Crowdsourcing



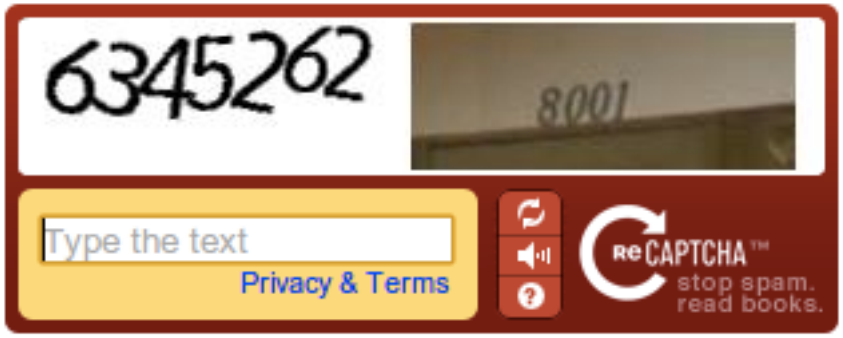
Crowdsourcing

- *Portmanteau* of "crowd" and "outsourcing," first coined by Jeff Howe in a June 2006 Wired magazine article

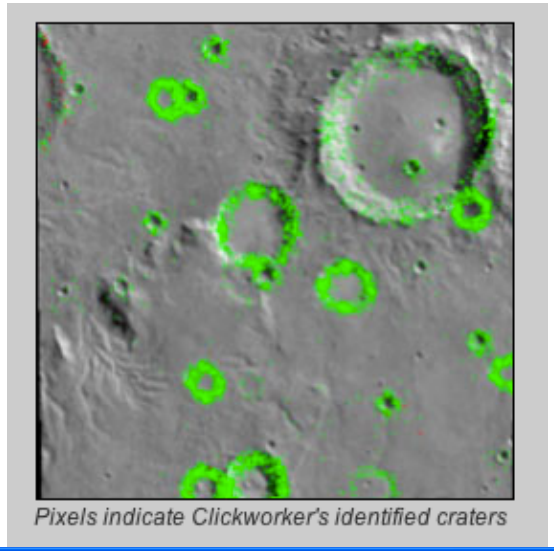
- [Merriam-Webster] the practice of obtaining needed **services, ideas, or content** by soliciting contributions from a **large group of people** and especially from the **online** community rather than from traditional employees or suppliers

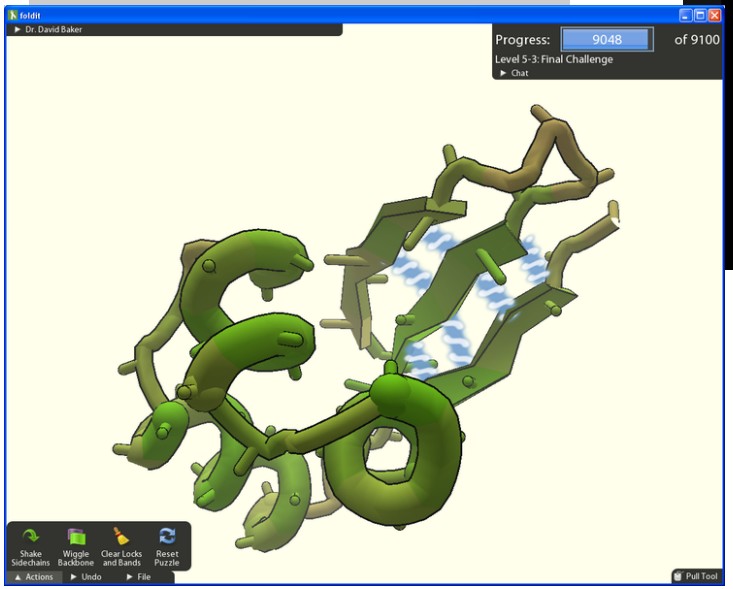
Crowdsourcing Success Stories

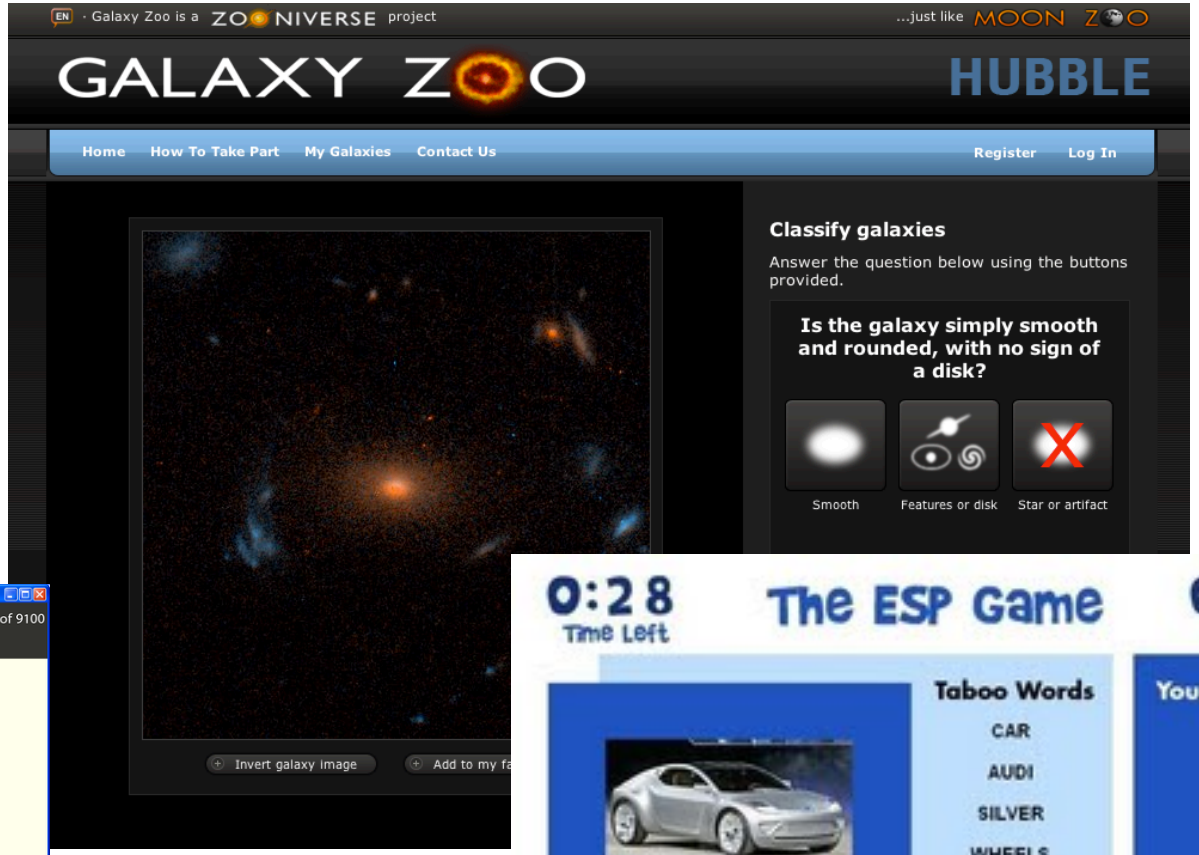
- Wikipedia
- Re-captcha

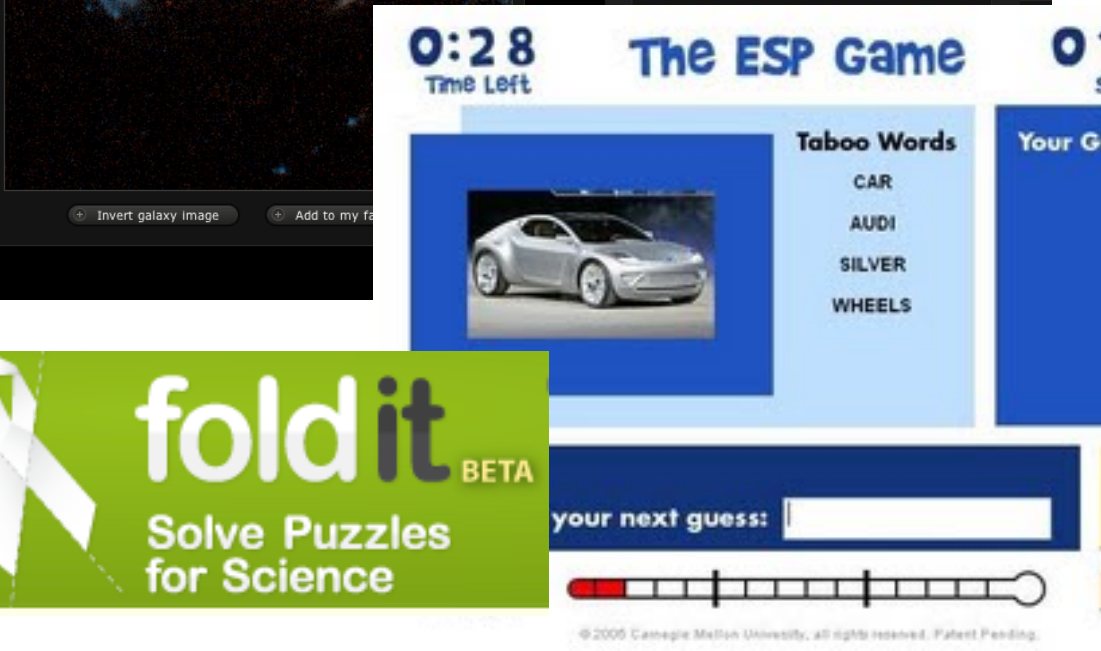


Games with a Purpose









Crowdsourcing Incentives

- Fun (enjoyment)
- Community (belonging, desire to help)
- Paid Crowdsourcing

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

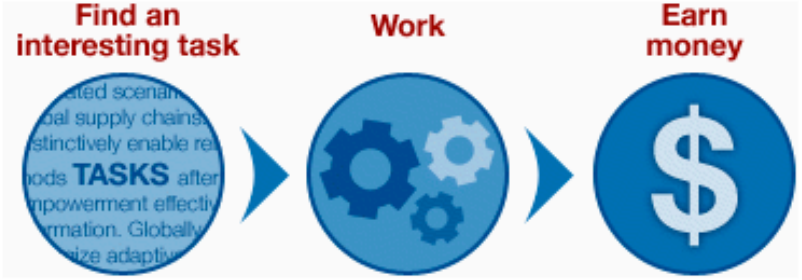
amazonmechanical turk
beta
Artificial Artificial Intelligence

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 HITs Now

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

Get Started

Demographics of MTurk workers

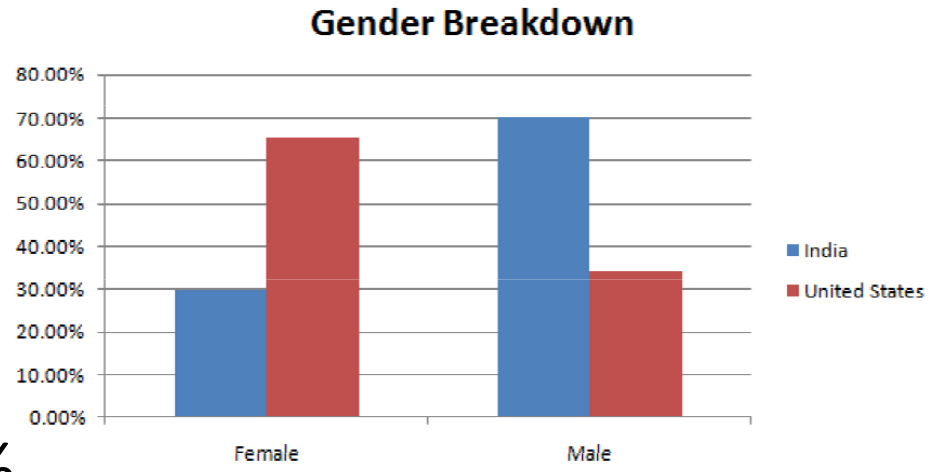
Country of residence

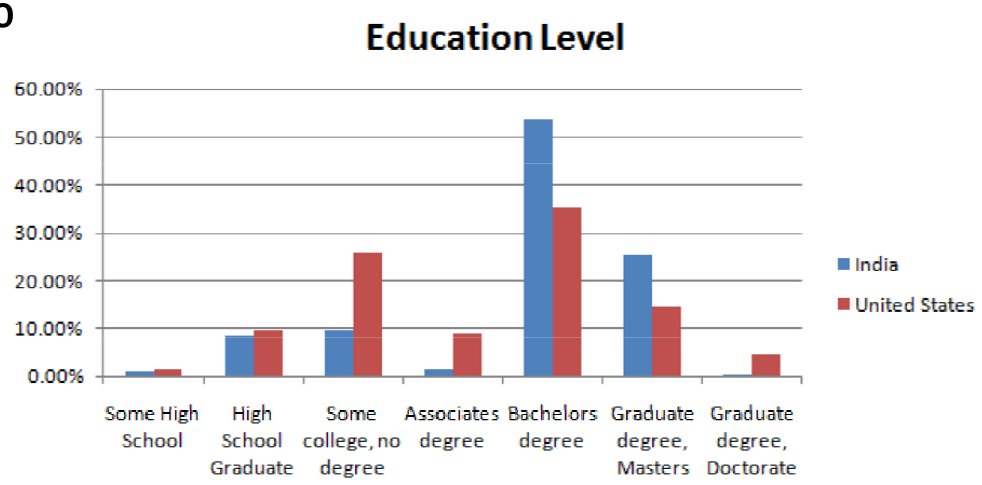
- United States: 46.80%
- India: 34.00%
- Miscellaneous: 19.20%

2013 Statistics:

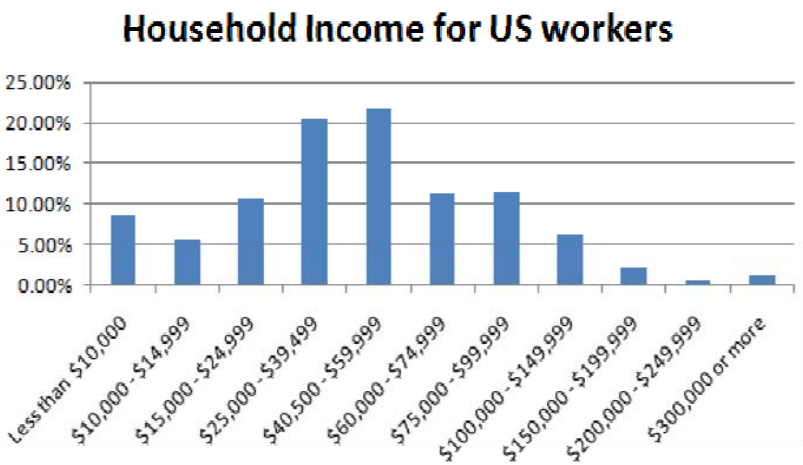
1M workers

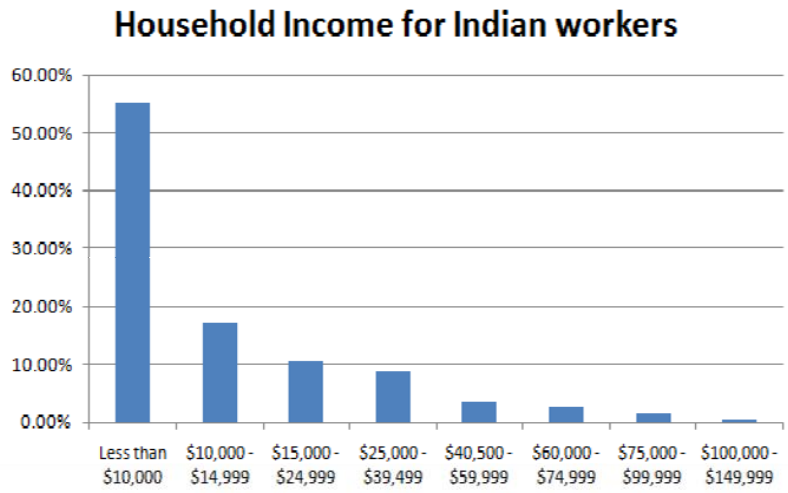
10% active

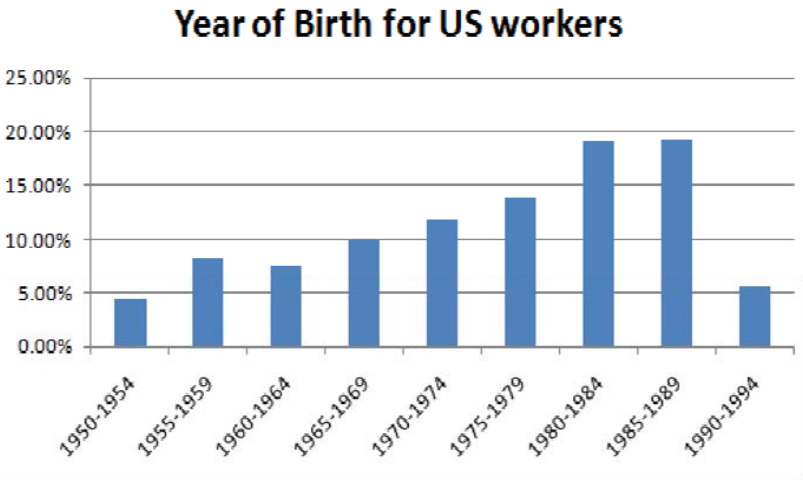


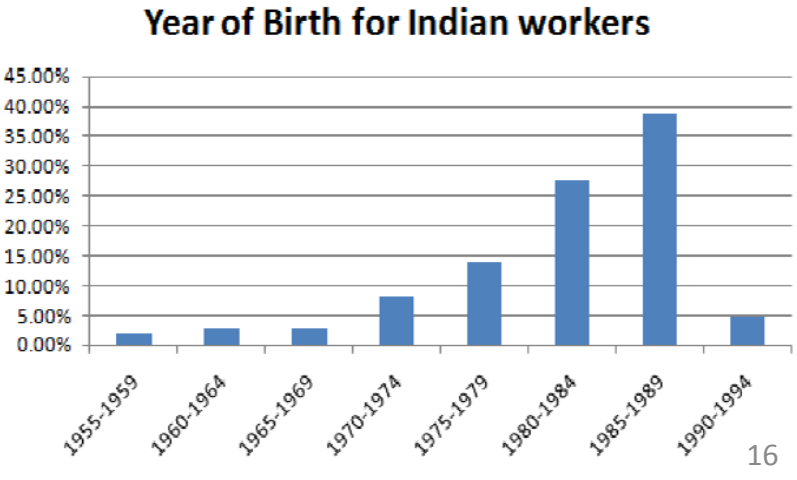


Demographics of MTurk workers

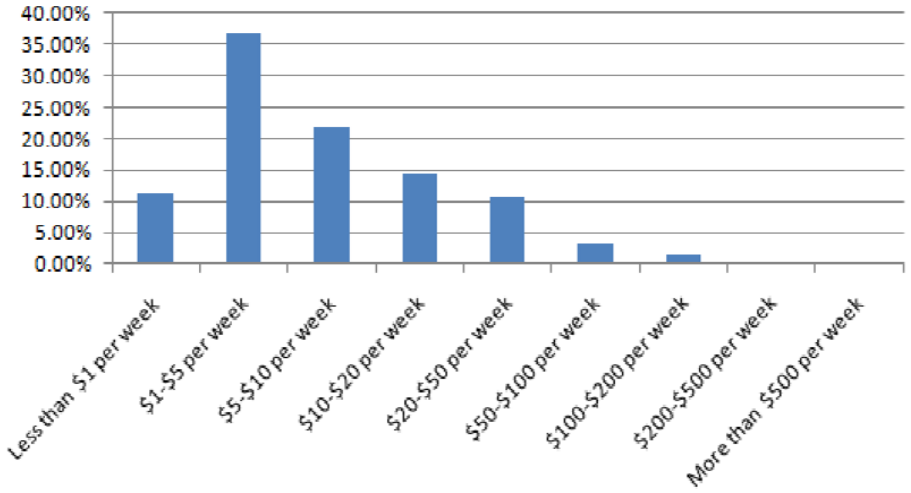


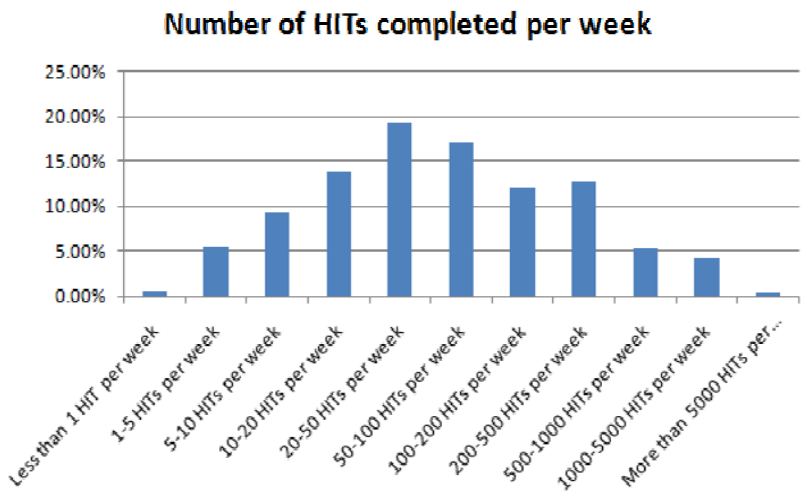


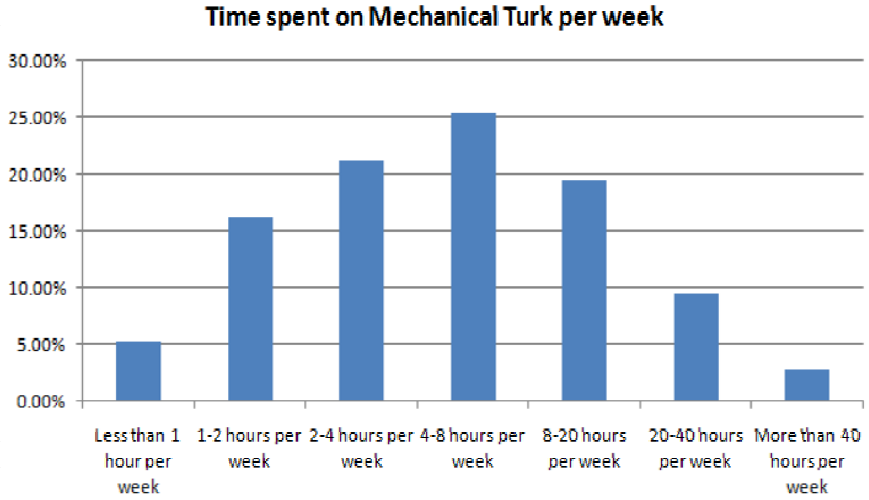




Demographics of MTurk workers

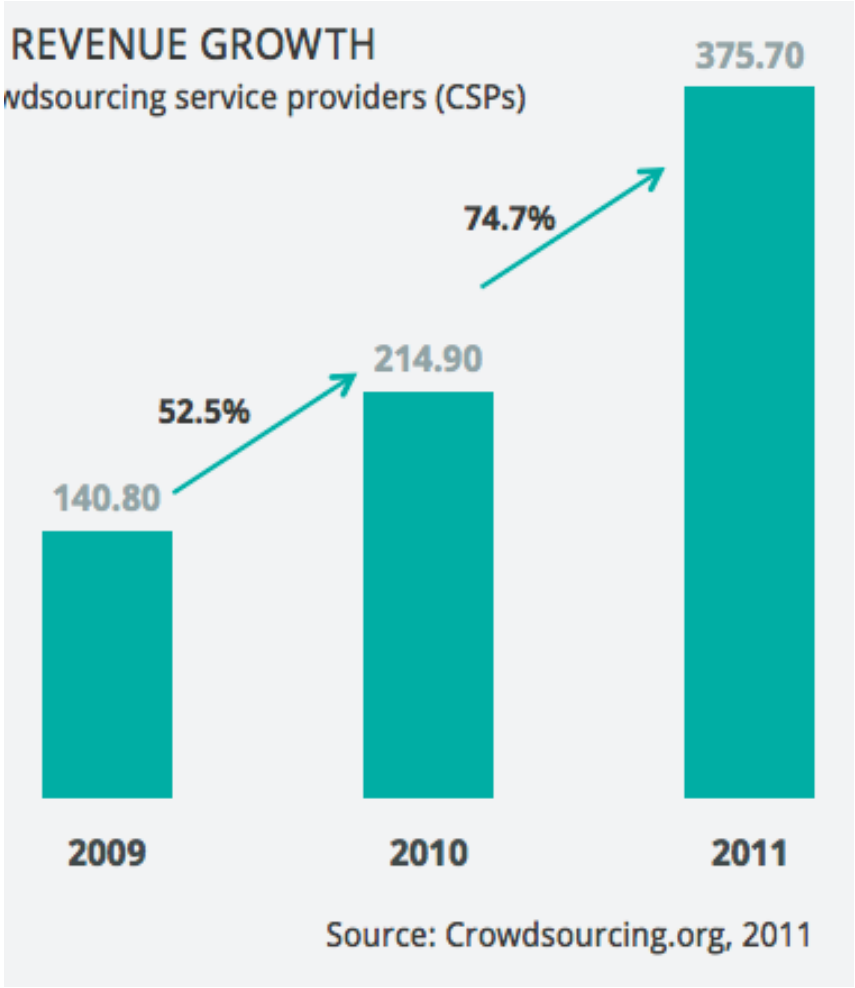






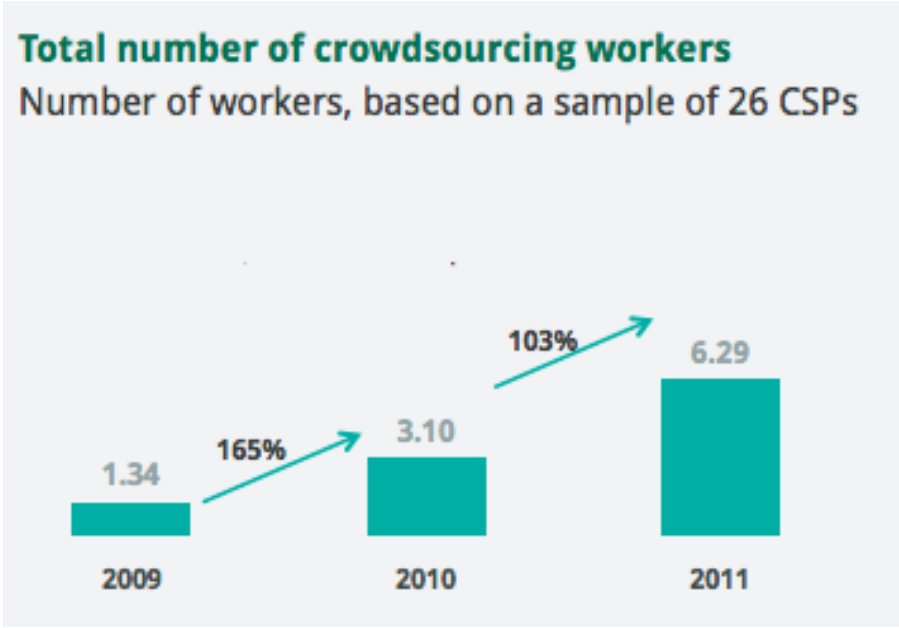
http://www.mturk-tracker.com/

Crowdsourcing Industry



Millions of $US, based on a sample of 15 leading crowdsourcing service providers (CSPs)

- Doubling of the market in 2012.



Source: Crowdsourcing.org, 2011

Crowdsourcing Ethics

- People work full-time as crowd workers
- Chinese crowdsourcing platform with 5.5M 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?

Get in touch

- http://www.gianlucademartini.net
- http://bit.ly/sheffield-datascience