

Human Computation: Machines are not enough

Gianluca Demartini

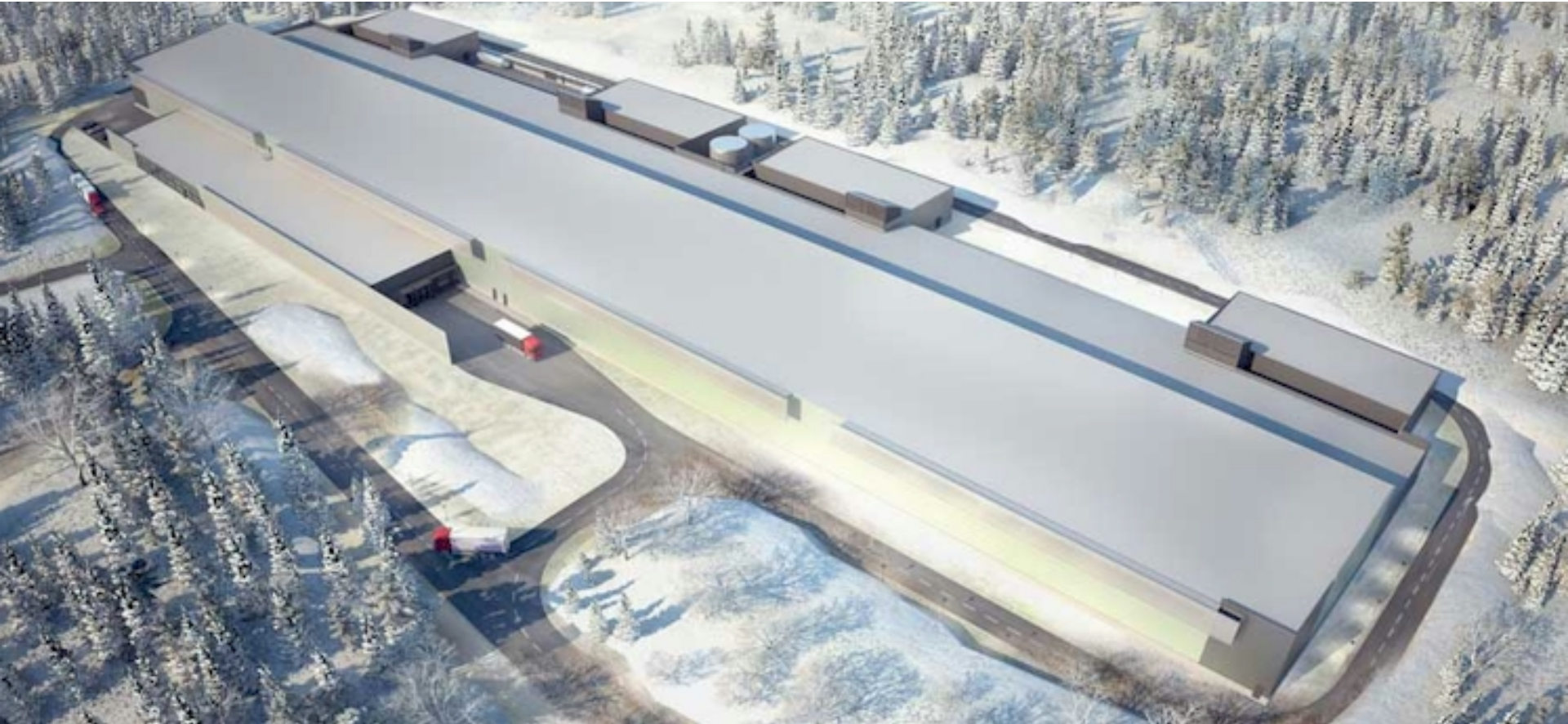
Information School, University of Sheffield

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

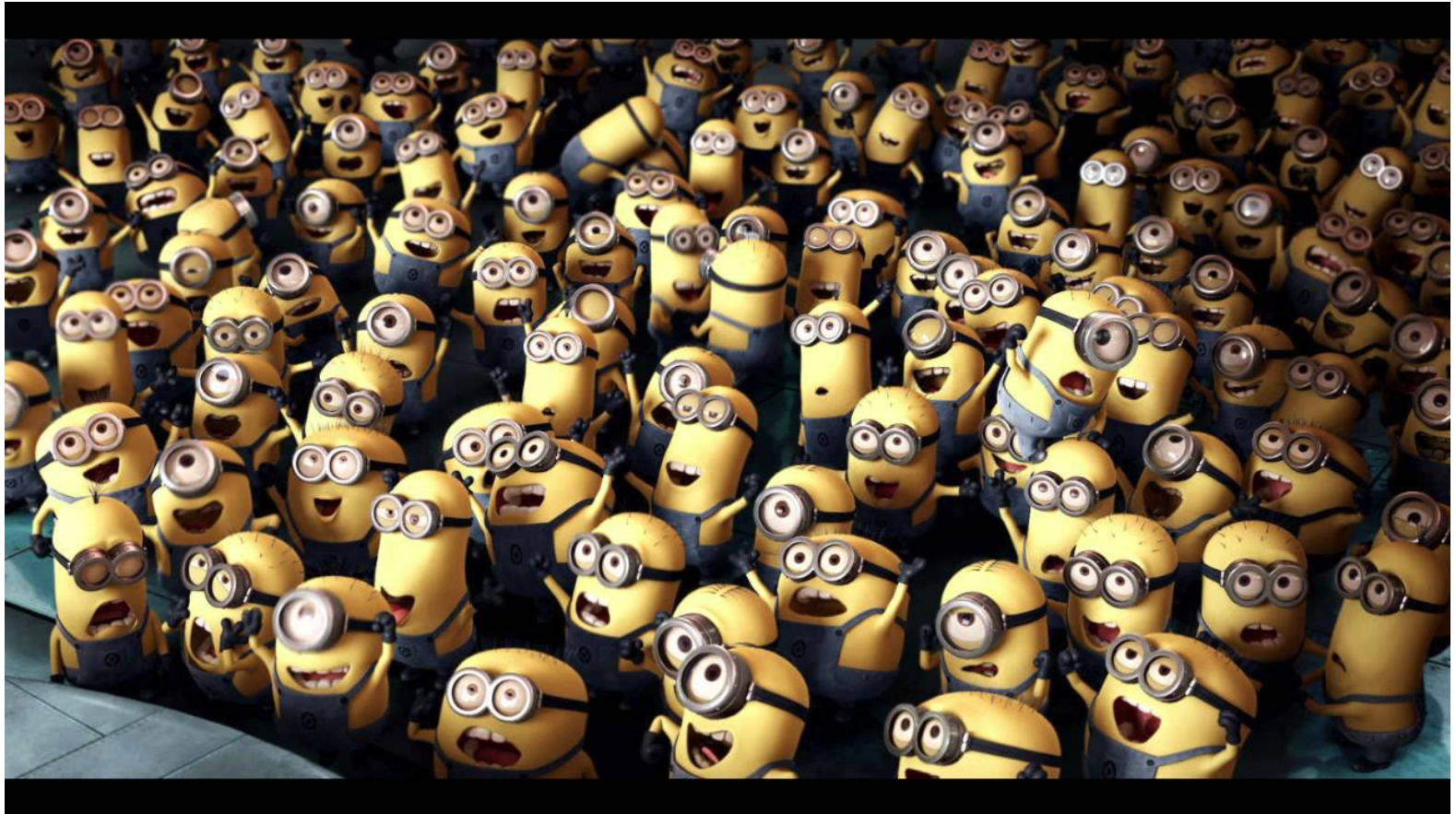
- Google has around 900,000 servers (260 million watts) 0.01% global energy
- CERN Large Hadron Collider 180MW
- The flux-capacitor needs 1.21 GW (just 5xGoogle)

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

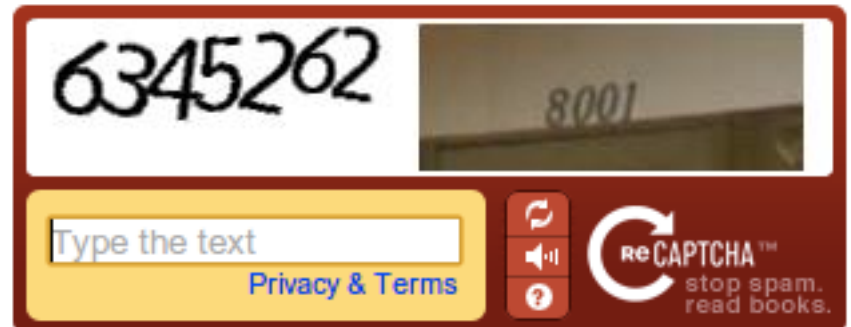


Crowdsourcing

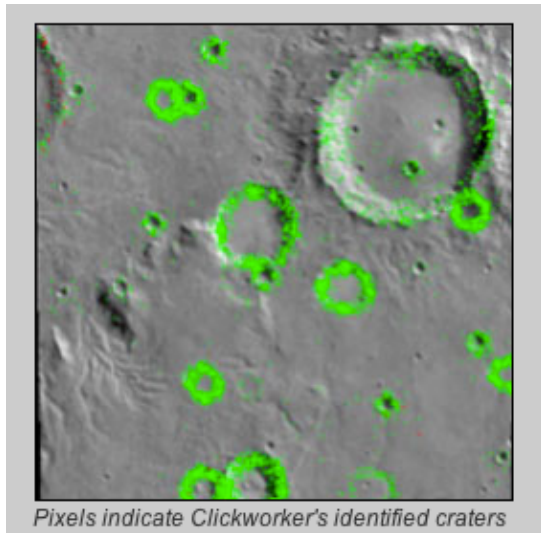
- *Portmanteau* of "crowd" and "outsourcing," first coined by Jeff Howe in a June 2006 Wired magazine article
- [Merriam-Webster] the practice of obtaining needed **services, ideas, or content** by soliciting contributions from a **large group of people** and especially from the **online** community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

Crowdsourcing Success Stories

- Wikipedia
- Re-captcha



Games with a Purpose



EN · Galaxy Zoo is a ZOO NIVERSE project ...just like MOON ZOO

GALAXY ZOO

HUBBLE

Home How To Take Part My Galaxies Contact Us Register Log In

Classify galaxies
Answer the question below using the buttons provided.

Is the galaxy simply smooth and rounded, with no sign of a disk?

Smooth Features or disk Star or artifact

0:28 Time Left

The ESP Game

Taboo Words

- CAR
- AUDI
- SILVER
- WHEELS

Your Guess:

your next guess:

© 2005 Carnegie Mellon University, all rights reserved. Patent Pending.

fold it BETA

Solve Puzzles for Science

foldit

Dr. David Baker

Progress: 9048 of 9100

Level 5-3: Final Challenge

Shake Sidechains Wiggle Backbone Clear Locks and Bands Reset Puzzle

Actions Undo File Pull Tool

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



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



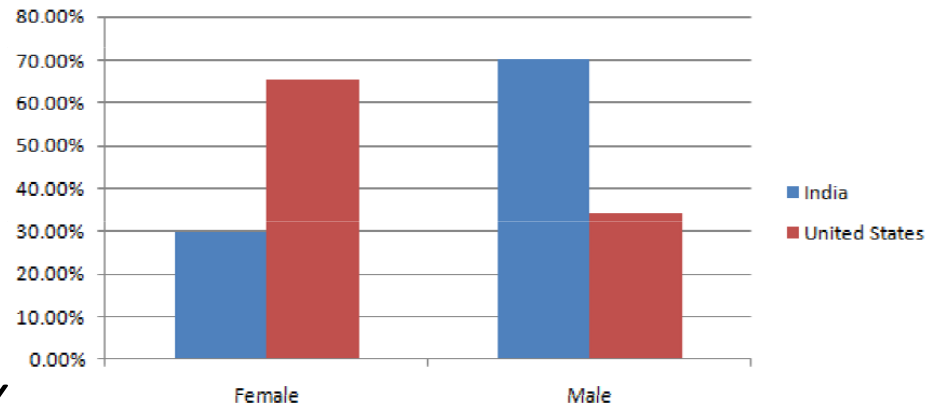
Demographics of MTurk workers

Country of residence

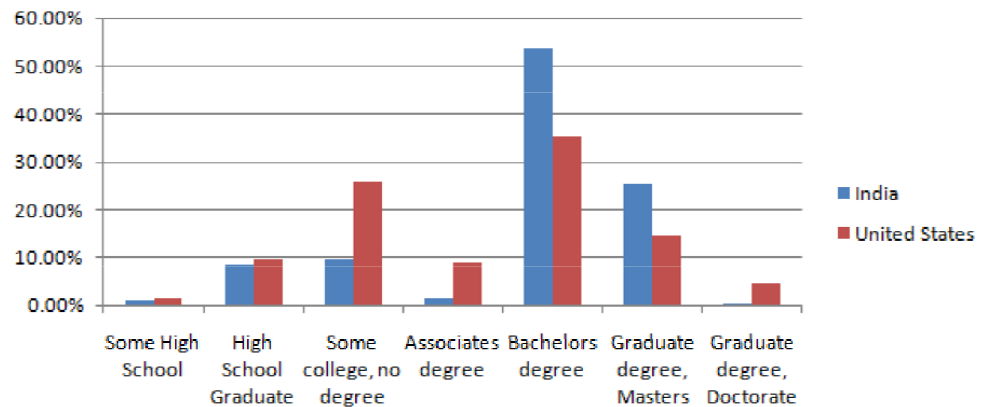
- United States: 46.80%
- India: 34.00%
- Miscellaneous: 19.20%

2013 Statistics:
1M workers
10% active

Gender Breakdown

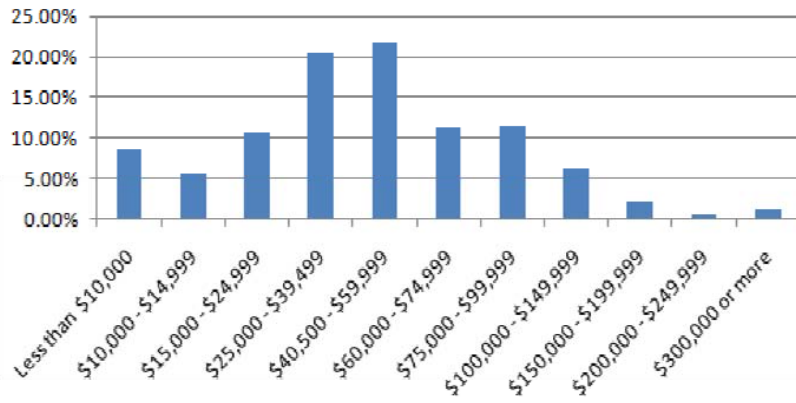


Education Level

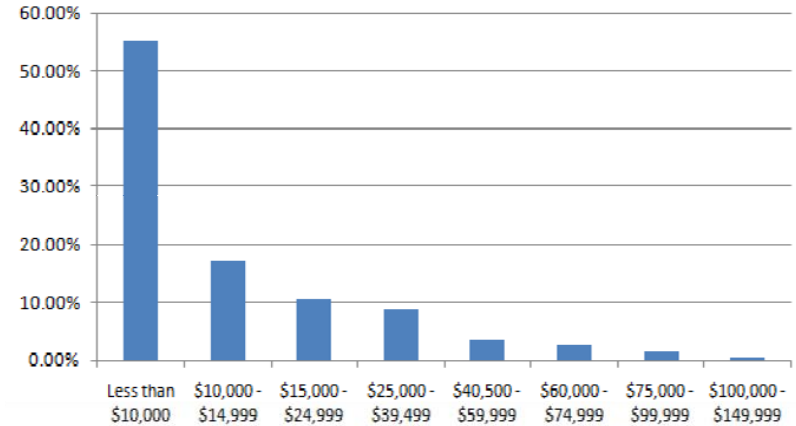


Demographics of MTurk workers

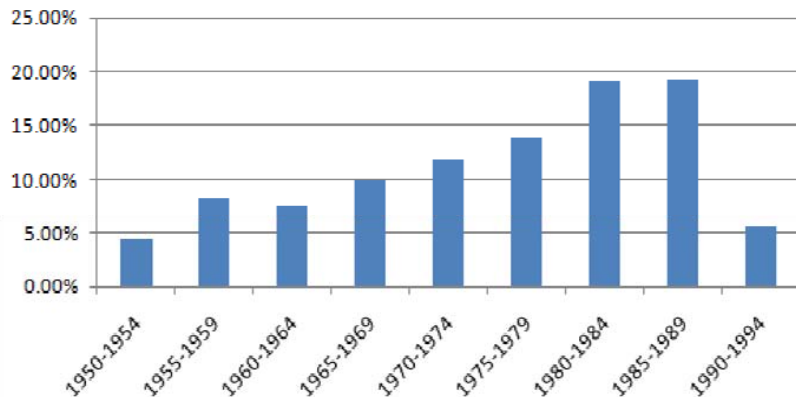
Household Income for US workers



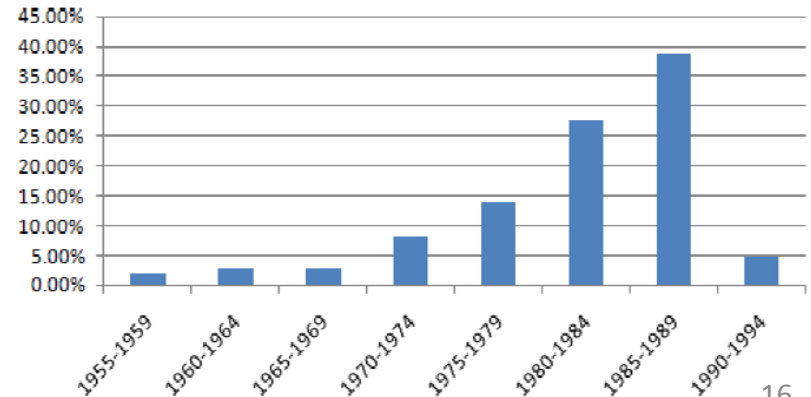
Household Income for Indian workers



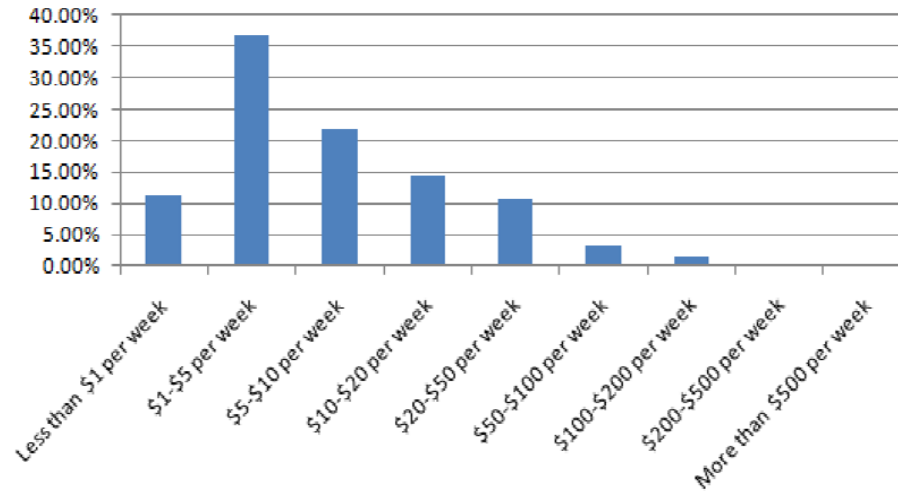
Year of Birth for US workers



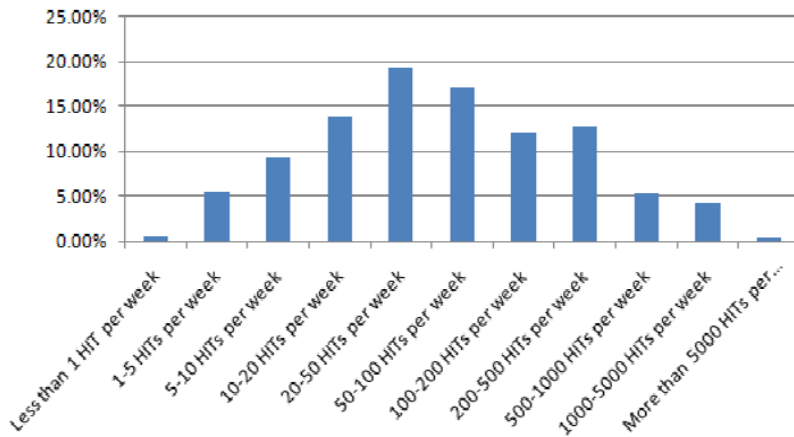
Year of Birth for Indian workers



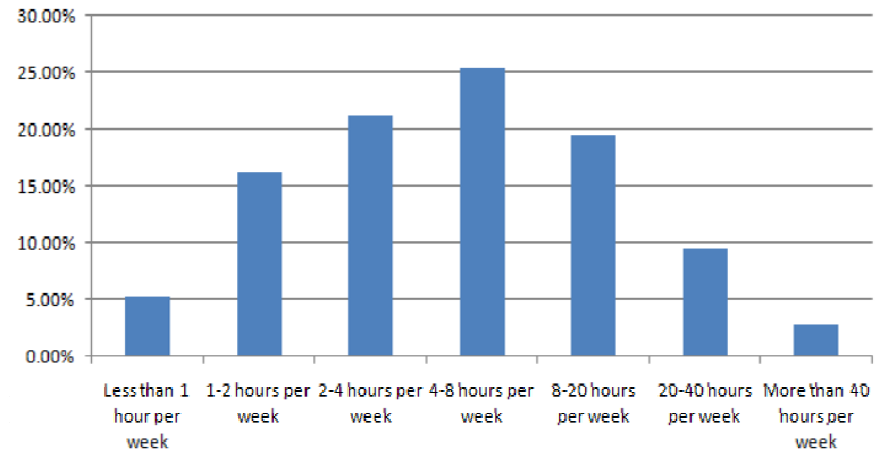
Demographics of MTurk workers



Number of HITs completed per week

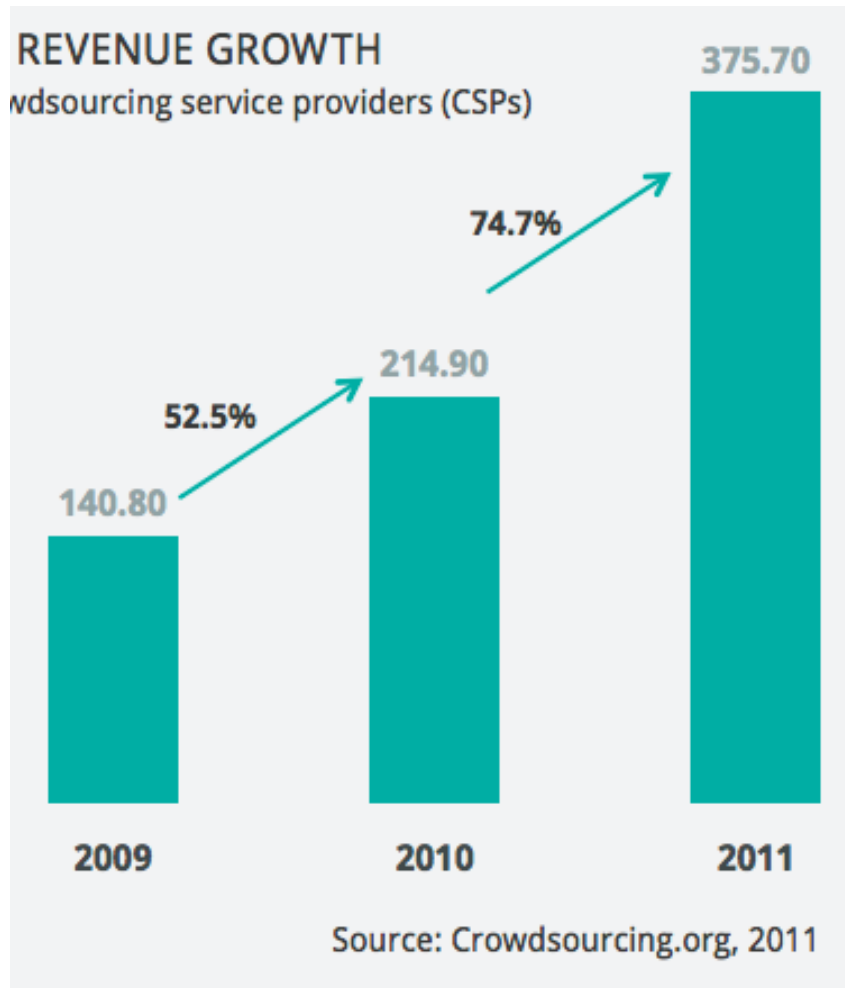


Time spent on Mechanical Turk per week

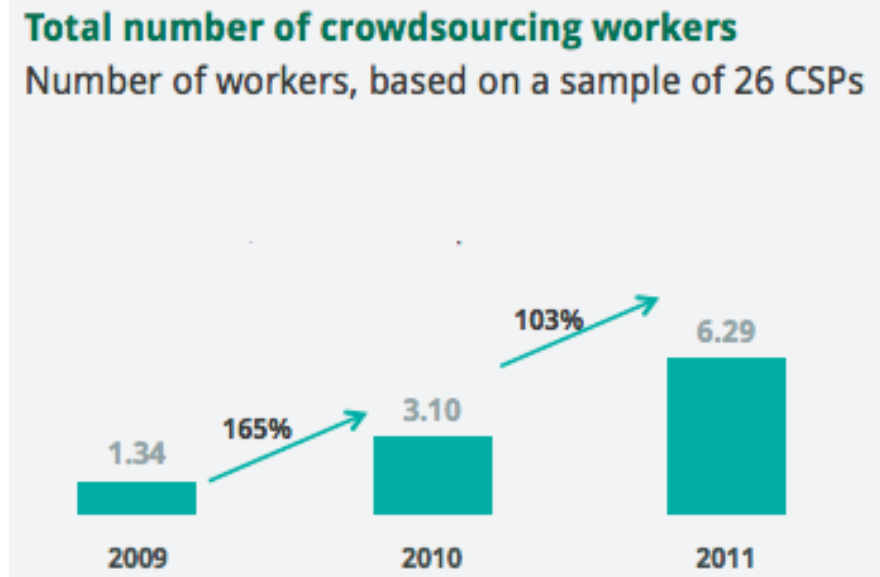


<http://www.mturk-tracker.com/>

Crowdsourcing Industry



- Doubling of the market in 2012.



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

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>