Crowdsourcing & Human Computation
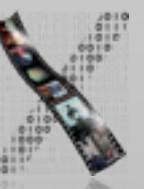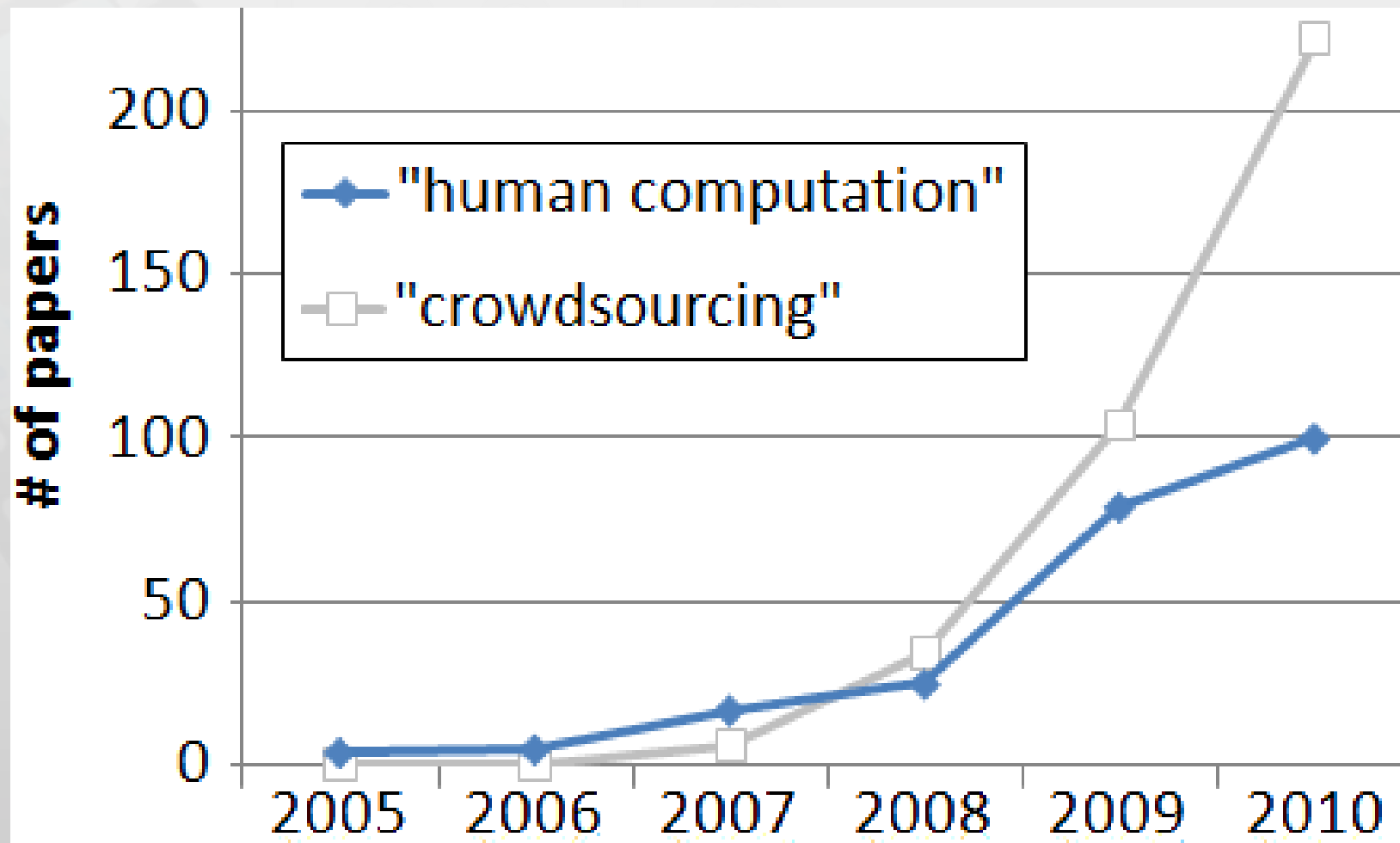
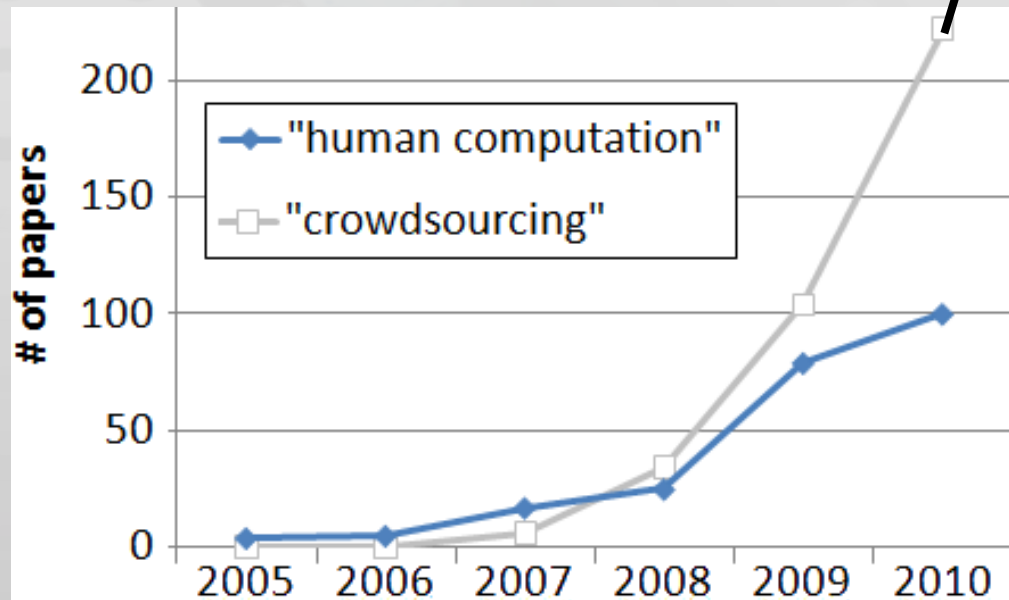# Today

- mainly driven by Mechanical Turk and Social Networks

Today

~660

~455
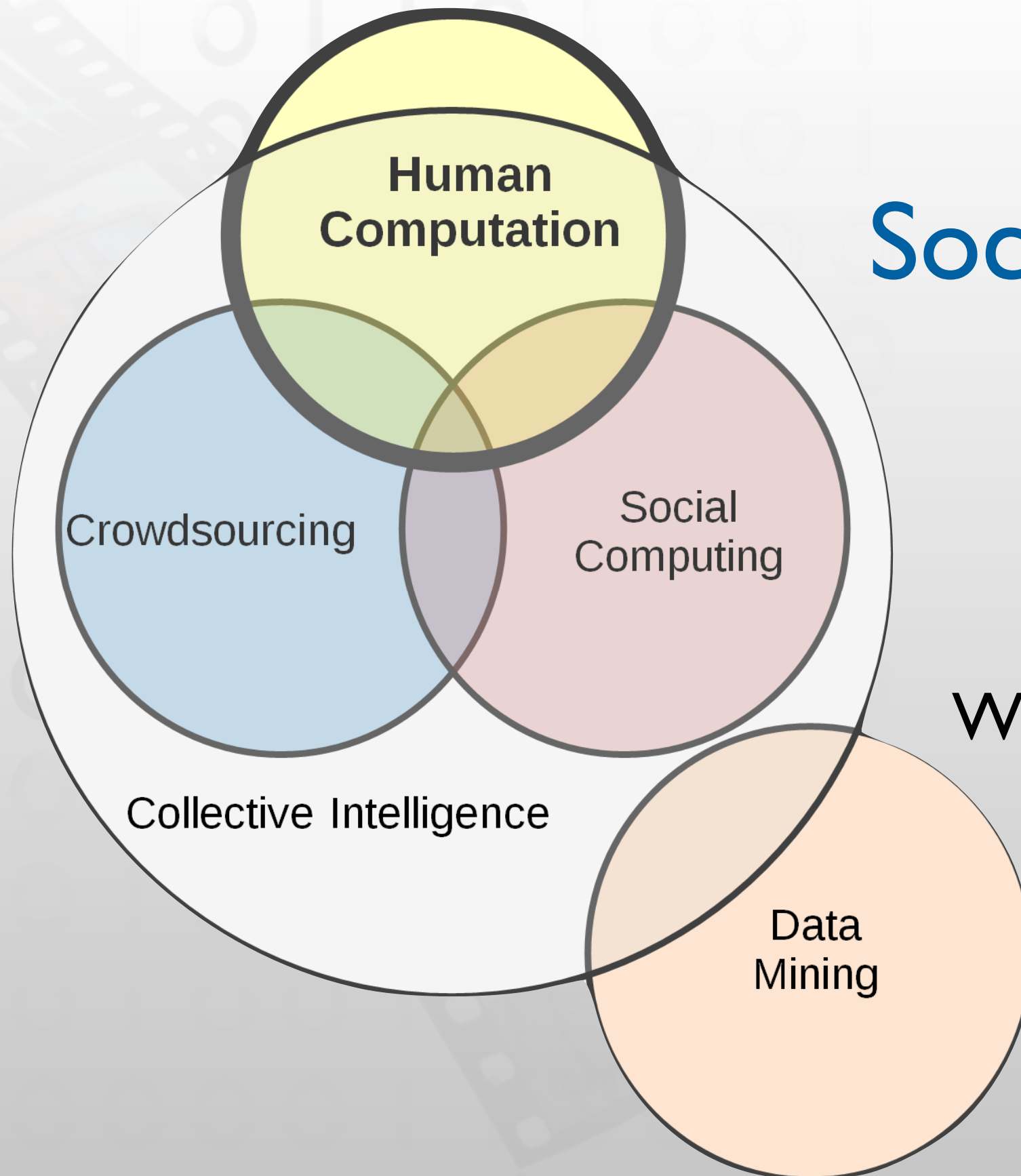

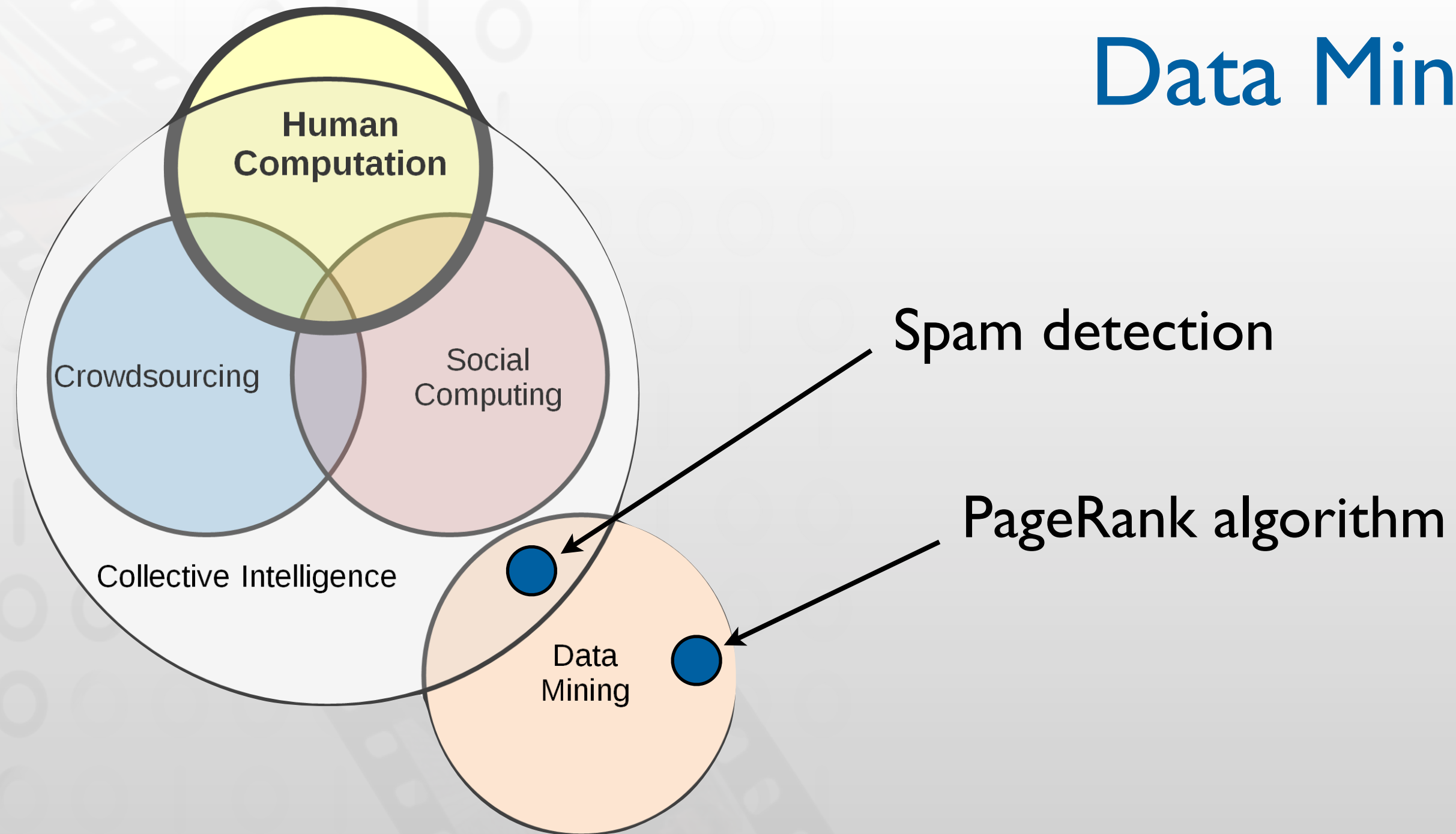- mainly driven by Mechanical Turk and Social Networks

Taxonomy of Social Computing

[Quinn and Bederson, CHI 2011]

Human Computation

Crowdsourcing

Social Computing

Collective Intelligence

Data Mining

What are the differences?

Data Mining

Spam detection

PageRank algorithm

- Users are not aware of DM, happens in the background
- "Uncreative" interactions

Collabio

[Bernstein et al., UIST 2009]

- Guess the tags that describe your friend best
- Points for common tags

PhotoCity

Task: reconstruct the geometry of buildings from photos

- PhotoCity is a competition between two universities to get the better model

- "Game with a purpose"

PhotoCity
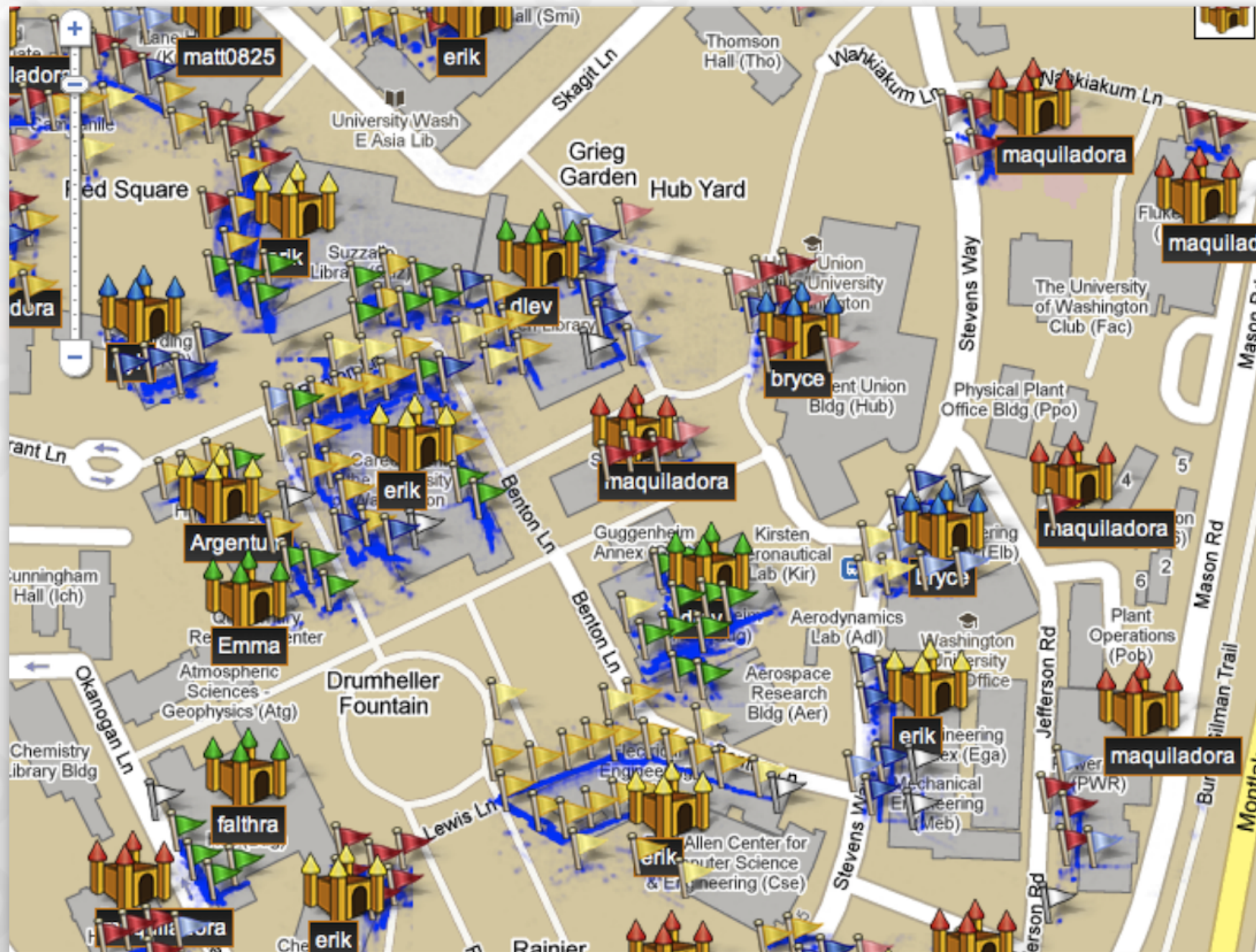
[Tuite et al., CHI 2011]

- Players get points for photos with new feature vectors
- Better photos have more new features
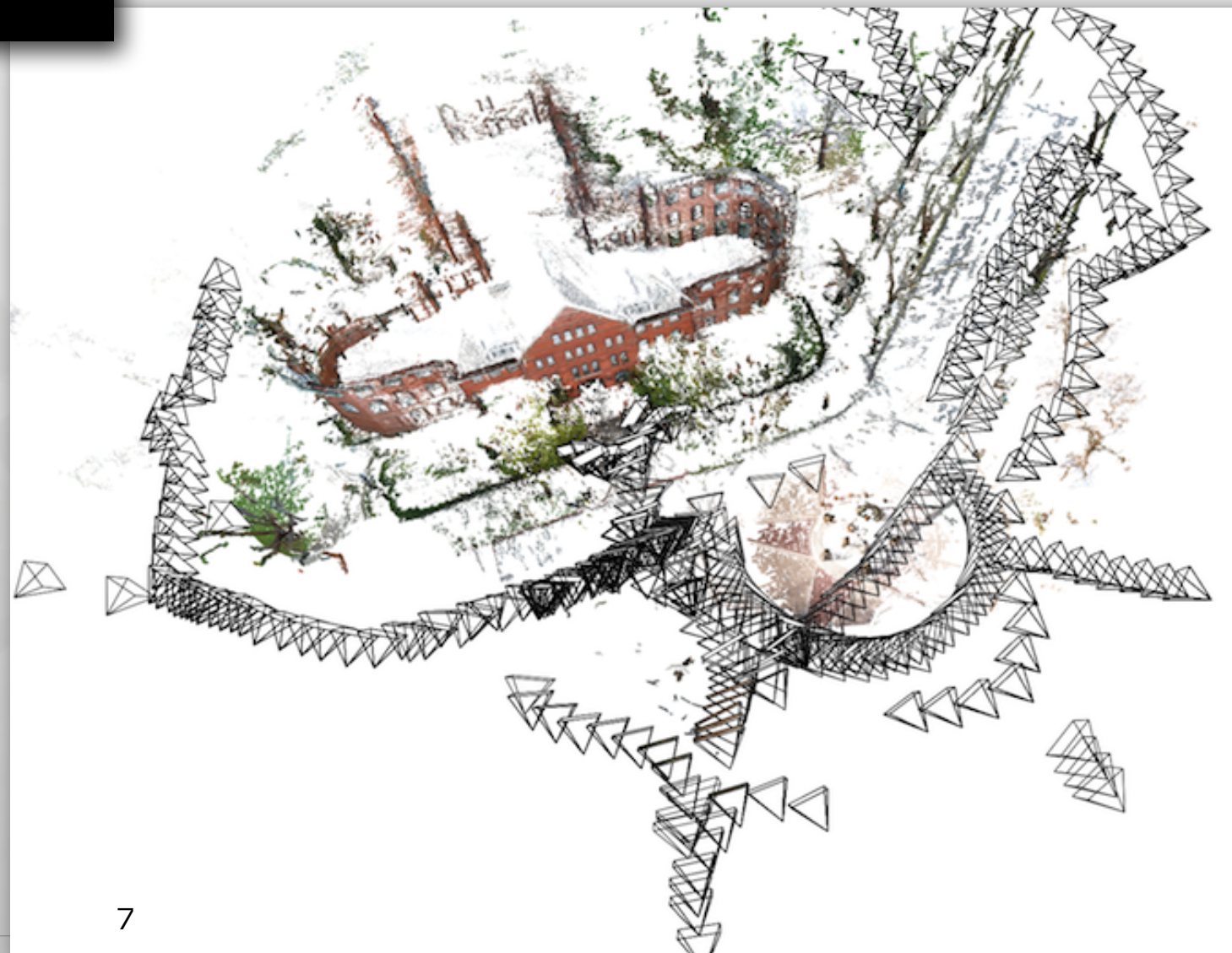- Website guides the players with flags placed at problematic areas

reCaptcha
[von Ahn et al., Science 2008]

PhotoCity
[Tuite et al., CHI 2011]

Collabio
[Bernstein et al., UIST 2009]

Human Computation

Crowdsourcing

Social Computing

Collective Intelligence

Data Mining
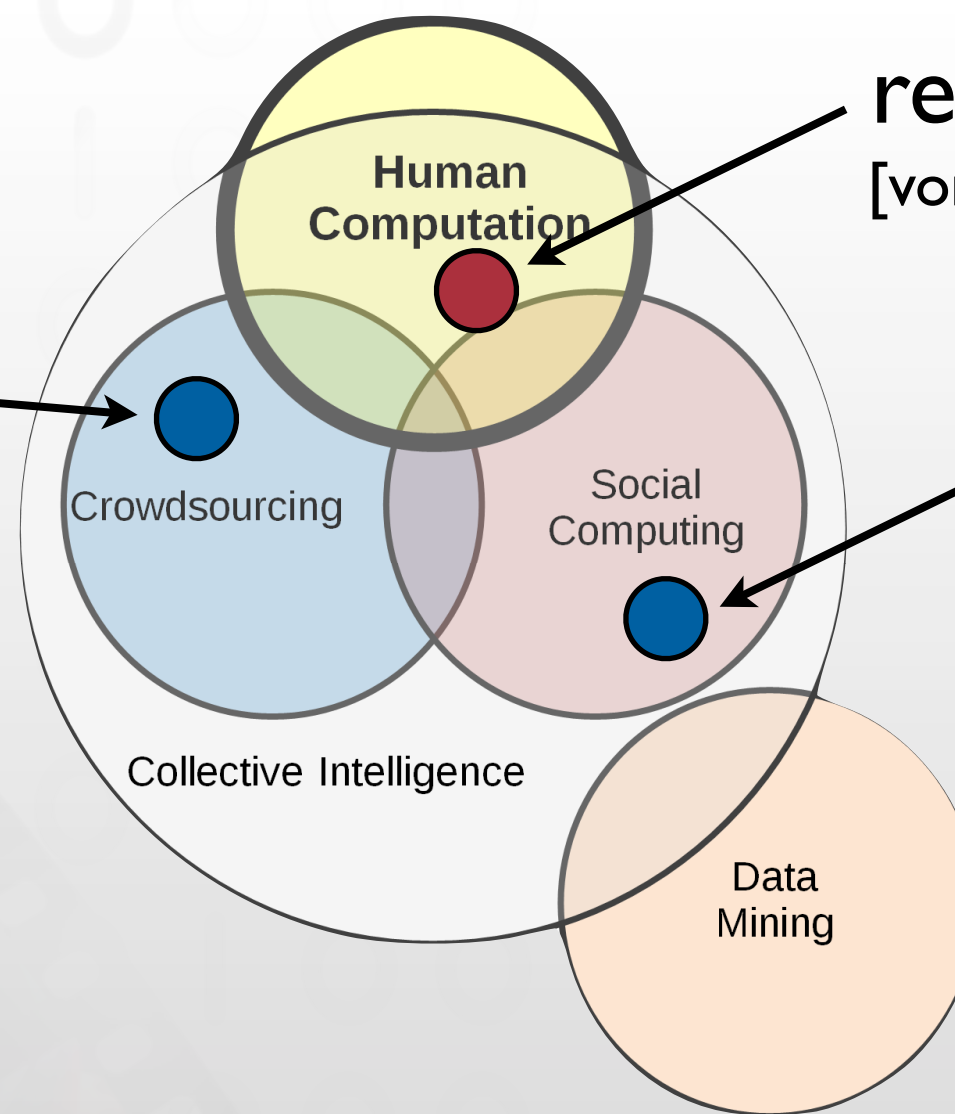
- Social computing facilitates relatively natural human behavior mediated by technology
- Crowdsourcing replaces experts with undefined, usually large groups of people
- Collective intelligence: groups of individuals that are doing things that seem intelligent

reCaptcha

[von Ahn et al., Science 2008]

- Prove to be human by deciphering text that OCR cannot
- 200 million CAPTCHAs are solved by humans around the world every day
- Leverage this "processing power"!

reCaptcha

[von Ahn et al., Science 2008]

media computing group

Human Computation

- Von Ahn, 2005

  "... a paradigm for utilizing human processing power to solve problems that computers cannot yet solve"

- Human as a processor for a larger (computer directed) algorithm

Human Computation

Easy for Computers

Easy for Humans

- Von Ahn, 2005

"... a paradigm for utilizing human processing power to solve problems that computers cannot yet solve"

In-class Exercise: Mechanical Turk

- Visit http://www.mturk.com/ click "find hits now"

- What kind of jobs are posted there?

- In what manner are they posted?

- How much are they paid?

- How much time do they take?

- What kind of jobs are missing?

The original Mechanical Turk

VizWiz: Nearly Real-time Answers to Visual Questions

[Bigham et al., 2010]

Stage 1

Stage 2

Outline the **Wheaties**

Remote Worker on Mechanical Turk
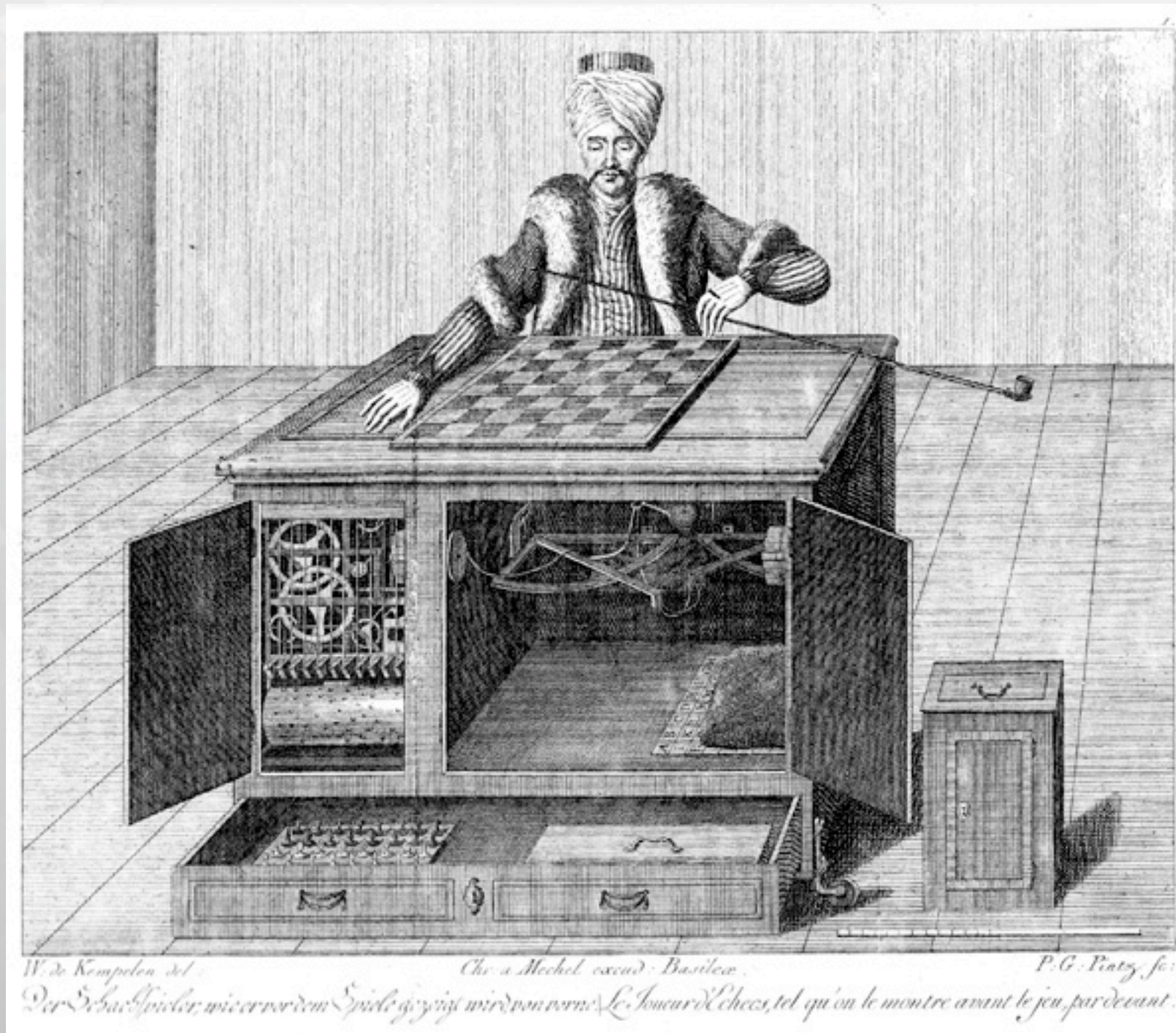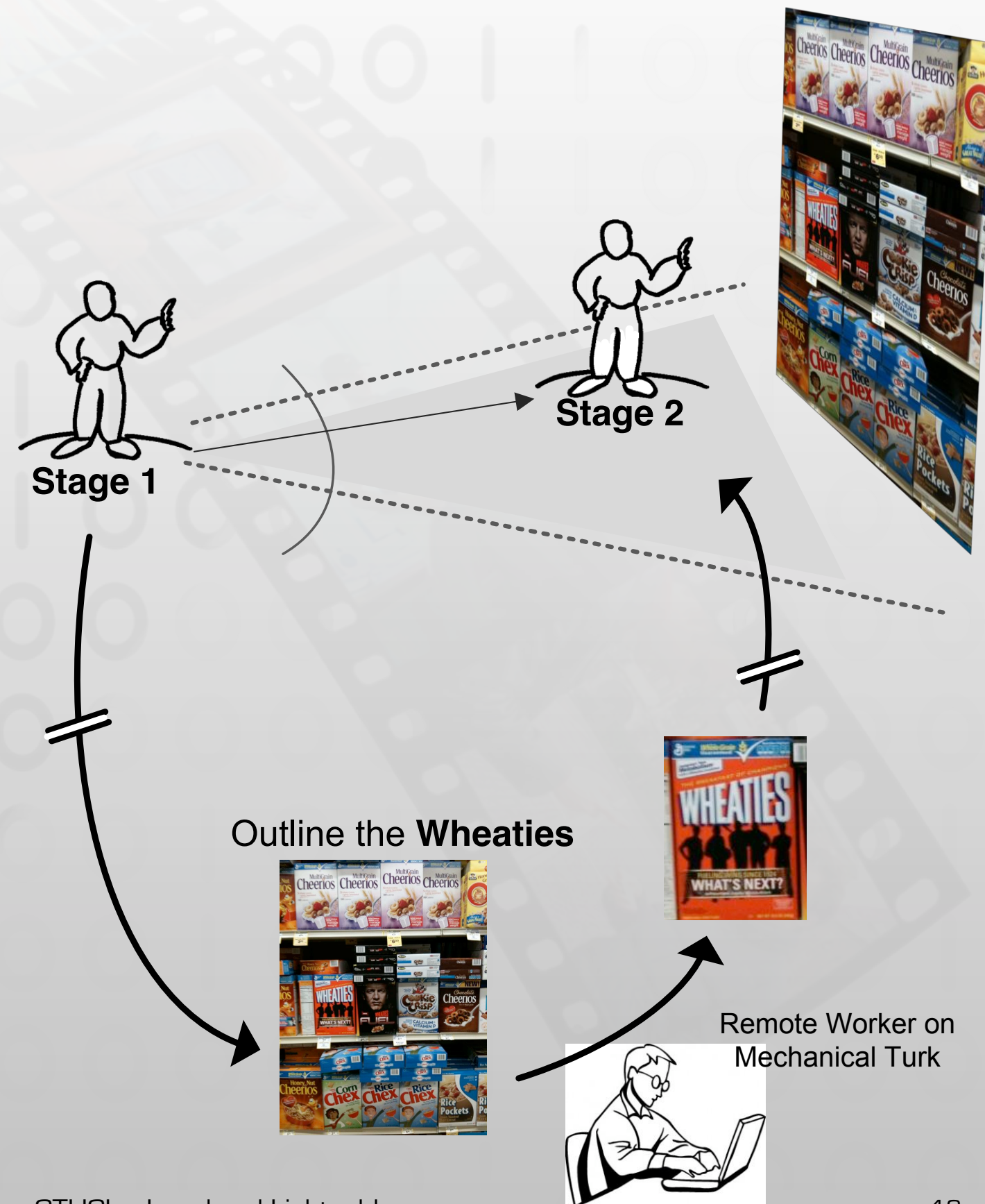
1. Blind user takes a photo with his smartphone and formulates a question

2. Turker has a look at the photo and listens to question

3. Answer is transmitted back to the phone and blind user

VizWiz: Nearly Real-time Answers to Visual Questions

[Bigham et al., 2010]

What denomination is this bill?

(24s) 20
(29s) 20

Do you see picnic tables across the parking lot?

(13s) no
(46s) no

What temperature is my oven set to?

(69s) it looks like 425 degrees but the image is difficult to see.
(84s) 400
(122s) 450

Can you please tell me what this can is?

(183s) chickpeas.
(514s) beans
(552s) Goya Beans

Who are the Crowdworkers?

[Ross et al., CHI 2011]
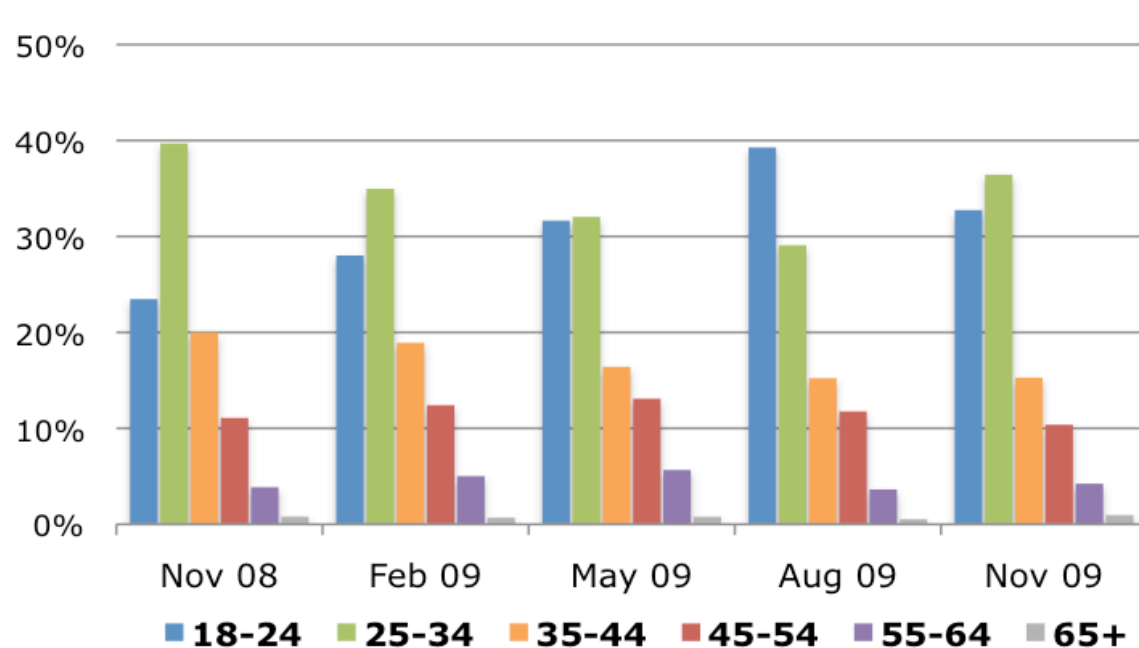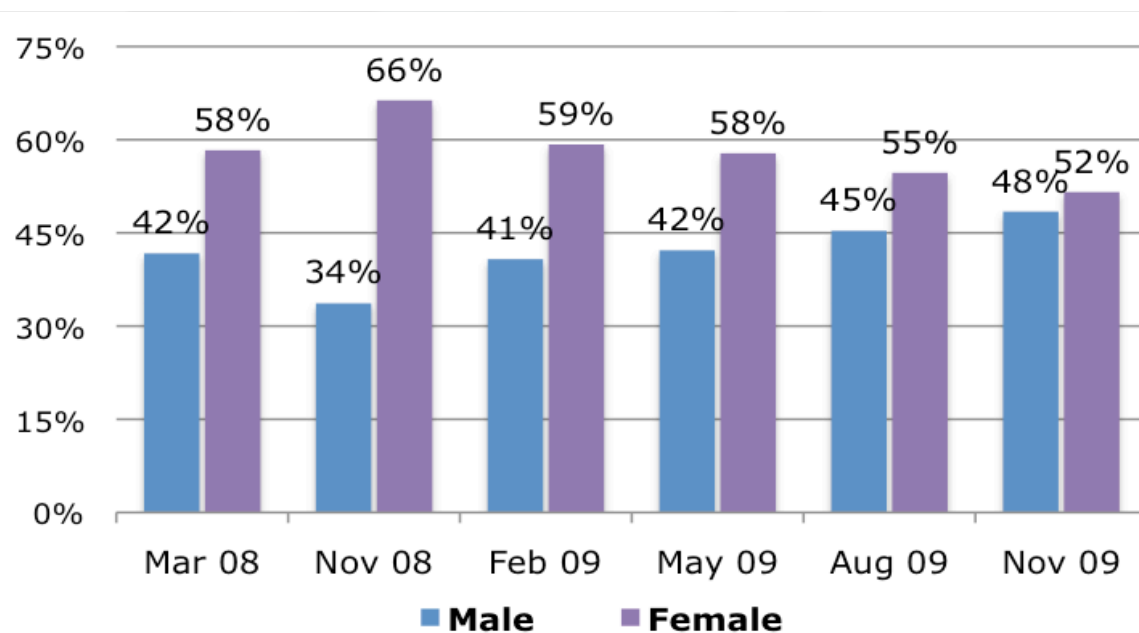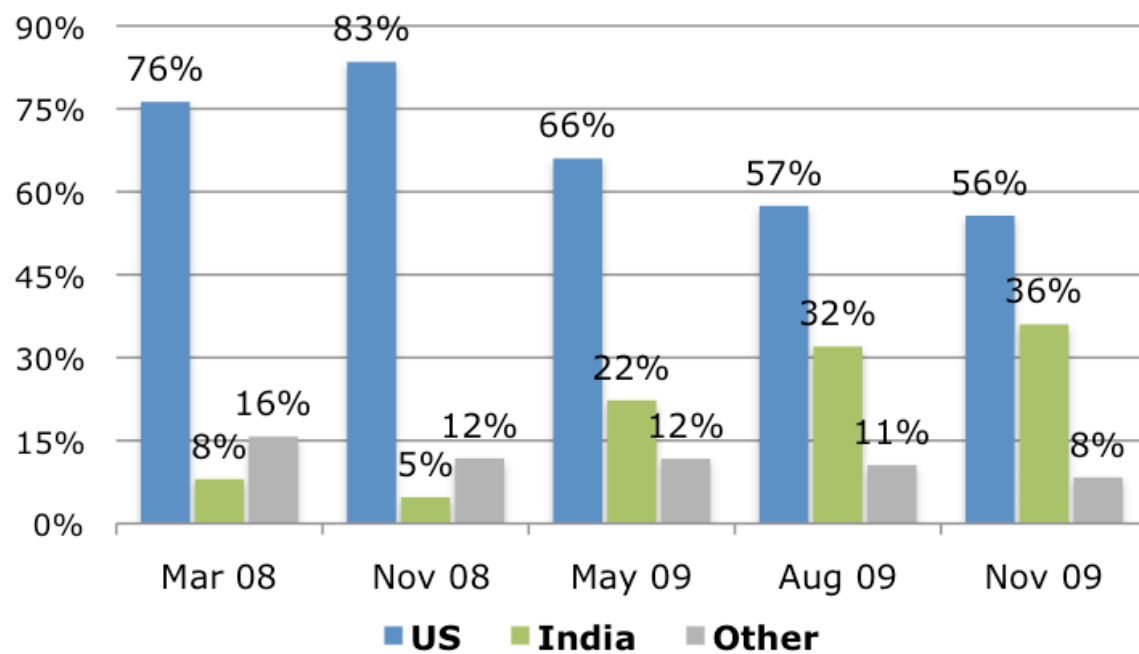
- Mechanical Turk was launched in 2005

- 400k registered workers in 2010

- 66k Hits this morning

- Average pay per hour: 2$

- Typical worker is well-educated, Indian or American, and young

media computing group

TurKit

[Little et al., UIST 2010]


```javascript
// generate a description of X
// and iterate it N times
var text = ""
for (var i = 0; i < N; i++) {
    // generate new text
    var newText = mturk.prompt(
        "Please write/improve this paragraph
          describing " + X + ": " + text)

    // decide whether to keep it
    if (vote("Which describes " + X + " better?",
        [text, newText]) == newText) {
        text = newText
    }
}
```

Iteration 1: Lightening strike in a blue sky near a tree and a building.

Iteration 2: The image depicts a strike of fork lightening, striking a blue sky over a silhoutted building and trees. (4/5 votes)

Iteration 3: The image depicts a strike of fork lightning, against a blue sky with a few white clouds over a silhouetted building and trees. (5/5 votes)

Iteration 4: ~~The image depicts a strike of fork lightning, against a blue sky- wonderful capture of the nature.~~ (1/5 votes)

Iteration 5: This image shows a large white strike of lightning coming down from a blue sky with the tops of the trees and rooftop peaking from the bottom. (3/5 votes)

Iteration 6: This image shows a large white strike of lightning coming down from a blue sky with the silhouettes of tops of the trees and rooftop peeking from the bottom. The sky is a dark blue and the lightening is a contrasting bright white. The lightening has many arms of electricity coming off of it. (4/5 votes)

http://turkit-online.appspot.com/

In-class Exercise: Quality Control

- 30% of open ended tasks tend to yield unusable results

- "Lazy" Workers do only the minimum amount required

- "Eager Beavers" might try to do more than your algorithm can handle

- Exercise: In groups, come up with strategies how one can ensure quality of the delivered work

Quality Control Mechanisms

- **Verify the work:** Verification is often easier than the task

- **Output agreement:** Have multiple workers agree on answer

- **Reputation:** Mechanical Turk records approval rate

- **Economic models:** Pay more for honest workers

- **Defensive task design:** Is it easier to cheat than to be honest?

- **Statistical filtering:** Discard outliers from an expected distribution

- **Multilevel review:** One worker reviews the previous stage of work

Soylent

- Word processor with cloud inside

- Plug-in to MS Word with different modules

  - Crowdproof: spelling and grammar checking by asking Mechnical Turk workers

  - Shortn: asks workers to suggests ways to shorten a given text

  - The Human Macro: any word processing task

- Embed human computation in an everyday application

  - Wizard-of-Oz prototyping as part of running system

- Achieves complex tasks that would require expert users otherwise

- Multilevel review: Find-fix-verify pattern

  - Splits task into a series of generation and review stages

Shortn via Find-Fix-Verify

Problems and Open Questions:

- Latency, real-time answers?

- Sustainability and reliability?

- Ethical questions

- Privacy

- Ownership

- Personalization

- Impact on the worker

    - poor compensation

    - no team interaction, no learning from peers

    - no perception of context, pride in work

[O'Neill et al., CHI 2013]

Using Mechanical Turk for HCI Studies

- E.g. Harrison et al., 2011, Kineticons

  asked users to rate meaning of the moving icon

- Mechanical Turk offers the researcher a large user pool

- Quick and relatively inexpensive results

Wave

- But workers will try to game the system

- Insert verifiable questions before subjective questions to judge the worker's honesty (E.g., describe the animation of the icon)

  [Kittur et al., 2008; Komarov et al. CHI 2013]

- Make an dishonest answer as hard as an honest one

- Used in design classes for user evaluations

  [Dow et al., CHI 2013]

Summary

- Humans can act as processor in Human Computation schemes to solve algorithms that CPUs will not solve

- Human Computation can solve simple and complex tasks

- Many options for quality control and incentives

- HCI researchers can use Mechanical Turk for their user studies or teaching