Crowdsourcing

CS 347
Michael Bernstein
Announcements

Abstract revisions due next Friday

We will send feedback on your drafts — use it to refine your idea and get it to a point where you had a crisp idea of your project!

Yes, you may still pivot if you want. But make sure to check your new idea with the staff!
How might computing connect us to tackle bigger, harder problems together?
Today: crowdsourcing

Peer production: decisions made collectively
- Open source software, collaborative encyclopedias, and Q&A
- Success disasters in peer production
- The role of community leaders

Crowdsourcing: decisions made centrally
- The Wisdom of Crowds and the threat of path dependence
- Creating complex outcomes

The future of work
Peer production

We work together / Like rama lama lama ka dinga da dinga dong
What is peer production?

[Benkler 2002]

Modes of production are ways that people create the things they need to survive and thrive.

You’re very familiar with one mode of production: firm-based production, where there exist clear boundaries on who’s in and who’s out, and typically hierarchical control.

However, the internet has enabled another mode: peer production, where volunteers self-organize.
Peer production examples

Wikipedia

Linux

StackOverflow
When should we use it?

Yochai Benkler [2002] asks: what is peer production good at? [1 min]

“Peer production is limited not by the total cost or complexity of a project, but by its modularity.” [Benkler 2002]

In other words, can we break it down into mostly-independent pieces?
Conflict and coordination

What happens to collaboration costs as Wikipedia grows? [Kittur, Suh, Pendleton, and Chi, CHI ’07]

Amount of direct work on articles goes down, and activity on coordination pages goes up
Decline!
[Halfaker et al., American Behavioral Scientist '13]

Conjecture: the tools and regulations put into place to deal with spam as Wikipedia grew wound up making the site less welcoming for newcomers.

https://stats.wikimedia.org/v2/##/en.wikipedia.org/contributing/active-editors/normal|line|all~total|monthly
What makes a leader in a peer production project?

Yes, even self-organized collectives develop leadership structures, and those structures ossify over time [Shaw and Hill 2014]

Reader-to-leader framework [Preece and Shneiderman, AIS Trans. HCI '09]: Readers > Contributors > Collaborators > Leaders

Goal: guide users into each new stage. See also: legitimate peripheral participation [Lave and Wenger '91]

Leaders are born, not made [Panciera et al. GROUP '09]

We can classify future power editors even from their first day!
What powers do leaders have?
[Keegan and Gergle, CSCW ’10]

How powerful are leaders in open communities like Wikipedia?

Method: data mine nominations for breaking news articles on the Wikipedia homepage. Stories were nominated and voted on by elite, middle-class, or newbie editors.

Result: “one-sided gatekeeping”

Elite editors could block nominations, but had no ability to get their own nominations approved
Do we work on the right topics?

How do we know if open source software and Wikipedia are actually working on content that matters?

Method: use Wikipedia logs to measure the web pages people are reading, and compare those levels of readership to the quality level of the corresponding articles (Stub, Start, C, B, Good, A, Featured)

Results: 40% of pageviews are to articles that are lower quality than should be if views and quality were perfectly correlated

Most over-represented: countries, pop music, internet, comedy
Recall: Answer Garden

[Ackerman and Malone, OIS '90]

An “organizational memory” system: knowing what the company knows

Main idea: members leave traces for others to solve their questions

The original Yahoo! Answers, Quora, Aardvark
Expertise recommendation

[McDonald and Ackerman, CSCW ’00]

Recommend people, not documents

Goal: help organizations know who can tackle each problem
Crowdsourcing

Wisdom of crowds
Crowdsourcing examples

Games with a Purpose

Innovation competitions

Data annotation services
What is crowdsourcing?

Crowdsourcing term coined by Jeff Howe [2006] in Wired

“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.”
What is crowdsourcing?

Two common models of crowdsourcing

Wisdom of the Crowd: aggregate opinions

Competition: accept many ideas but only take the best ones
Paid Crowdsourcing

Pay money for short tasks. Amazon Mechanical Turk: millions of tasks completed each year

<table>
<thead>
<tr>
<th>Task</th>
<th>Reward</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label an image</td>
<td>$0.20</td>
</tr>
<tr>
<td>Transcribe audio clip</td>
<td>$5.00</td>
</tr>
</tbody>
</table>

Many complexities in good task design and ethical treatment of workers — a topic for CS 278
The Wisdom of Crowds

The phenomenon that, in certain situations, aggregating opinions across a large number of people can produce a more accurate estimate of the answer than even the best expert in the room.

Independent guesses minimize the effects of social influence [Simoiu et al. 2019]

- Showing consensus cues like the most popular guess decreases accuracy
- Crowds are more consistent guessers than experts
- Crowds are only at the 67th percentile on average per question…but at the 90th percentile averaged across questions per domain!
Social influence makes outcomes unpredictable

[Salganik, Dodds, and Watts, Science ’06]

Puzzle: why can’t experts to predict which songs will be hits?

Method: 14,000 participants download free music

  Manipulation: no download info, or one of eight worlds that all start with zero downloads

Result: huge variance in download counts

  Best songs rarely did poorly, worst songs rarely did well; any other outcome was possible
Iterative crowd algorithm
You (misspelled) (several) (words). Please spellcheck your work next time. I also notice a few grammatical mistakes. Overall your writing style is a bit too phoney. You do make some good (points), but they got lost amidst the (writing). (signature)
Crowd-powered systems

Embed crowd intelligence inside of user interfaces and applications we use today

Interface Wizard of Turk Wizard of Oz
Automatic clustering generally helps separate different kinds of records that need to be edited differently, but it isn't perfect. Sometimes it creates more clusters than needed, because the differences in structure aren't important to the user's particular editing task. For example, if the user only needs to edit near the end of each line, then differences at the start of the line are largely irrelevant, and it isn't necessary to split based on those differences. Conversely, sometimes the clustering isn't fine enough, leaving heterogeneous clusters that must be edited one line at a time. One solution to this problem would be to let the user rearrange the clustering manually, perhaps using drag-and-drop to merge and split clusters. Clustering and selection generalization would also be improved by recognizing common text structure like URLs, filenames, email addresses, dates, times, etc.
VizWiz [Bigham et al., UIST ’10]

Visual question answering for the blind

<table>
<thead>
<tr>
<th>Question</th>
<th>Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>What color is this pillow?</td>
<td>I can't tell.</td>
</tr>
<tr>
<td></td>
<td>(89s) multiple shades of soft green, blue and gold</td>
</tr>
<tr>
<td>What denomination is this bill?</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>(24s) no</td>
</tr>
<tr>
<td></td>
<td>(29s) no</td>
</tr>
<tr>
<td>Do you see picnic tables across the parking lot?</td>
<td>no</td>
</tr>
<tr>
<td></td>
<td>(13s) no</td>
</tr>
<tr>
<td>What temperature is my oven set to?</td>
<td>425 degrees</td>
</tr>
<tr>
<td></td>
<td>(69s) difficult to see.</td>
</tr>
<tr>
<td></td>
<td>(84s) 400</td>
</tr>
<tr>
<td></td>
<td>(122s) 450</td>
</tr>
<tr>
<td>Can you please tell me what this can is?</td>
<td>chickpeas.</td>
</tr>
<tr>
<td></td>
<td>(183s) beans</td>
</tr>
<tr>
<td></td>
<td>(514s) Goya Beans</td>
</tr>
<tr>
<td></td>
<td>(552s) Energy drink</td>
</tr>
<tr>
<td>What kind of drink does this can hold?</td>
<td>no can in the picture</td>
</tr>
<tr>
<td></td>
<td>(91s) energy drink</td>
</tr>
<tr>
<td></td>
<td>(99s) 247s</td>
</tr>
</tbody>
</table>
Realtime crowdsourcing

[Bernstein et al., UIST '11]
Rebecca is free Monday at 10, Tuesday at 11, or Wednesday at 3. Let me know what works for you! Cal
Crowdsourcing complex work

[Kittur et al., UIST ’11]

How might we crowdsource more complex, interdependent outcomes?

Crowdsourcing as a map-reduce process

To write a wikipedia page, partition on topics, map to find facts and then reduce into a paragraph
Microtask crowds struggle with complex tasks

Design, engineering, writing, video production, music composition
[Kittur et al. 2013, Kulkarni et al. 2012]
Crowds of experts

Mechanical Turk
- microtask worker
- microtask worker
- microtask worker
- microtask worker
- microtask worker

Upwork
- programmer
- designer
- video editor
- musician
- statistician
Recall: flash teams

[Retelny et al., UIST '14]

Computationally-guided teams of crowd experts supported by lightweight team structures.
animation

Input: high-level script outline
Output: ~15 second animated movie

Our example:
44:40 hours
$2381.32
Achieve complex goals by structuring crowds as organizations, not algorithms
An example flash organization
Crowd research

[Vaish et al., UIST '17]

Crowdsourcing as a route to empower upward career and educational mobility through research experiences
Future of work
What would it take for us to be proud of our children growing up to work in these environments?
[Kittur et al. CSCW 2013]
Careers in crowd work

[Kittur et al. CSCW 2013]

More and more people are engaging in online paid work: programmers, singers, designers, artists, ... 

Would you feel comfortable with your best friend, or your own child, becoming a full-time crowd worker?

How could we get to that point? What would it take?

- Education
- Career advancement
- Reputation
Take back the market

Turkopticon [Irani and Silberman ’13]  
Lets workers (sellers) review requesters (buyers)

Dynamo [Salehi et al. ’15]  
Lets workers engage in collective action
Needed infrastructure

Support for career growth

Training and education
  e.g., micro-internships [Suzuki et al. 2016]

Longer-term employment guarantees

Decoupling the social safety net from firm-based employment

Policy
For more: take CS 278
Discussion

Find today’s discussion room at http://hci.st/room