CROWDSOURCING A MEETING OF MINDS
DESIGNING THE FUTURE OF WORK

Michael Bernstein
Stanford Computer Science
msb@cs.stanford.edu + @msbernst
20% of US jobs [Blinder 2006]
45,000,000 workers [Horton 2013]
How might computing connect large groups to tackle bigger, harder problems than they could complete in isolation?
CROWDSOURCING: SMALL TASKS, MANY PEOPLE

Computationally recombine many paid non-expert opinions

- e.g., text shortening [Bernstein et al. 2010]
- e.g., image labeling [von Ahn and Dabbish 2005]
- e.g., data collection [Deng et al. 2009]
Public deployment: over 100,000 unlocks
no slower than slide-to-unlock
no more cognitive load than slide-to-unlock

[Vaish et al. CHI 2014]
HYBRID CROWD-ML CLASSIFIERS

Paired examples generate human features
Machine learning learns to weigh the evidence

Hybrids 300% relative improvement over human reasoning

[Cheng and Bernstein CSCW 2015]
MICROTASK CROWDS STRUGGLE WITH COMPLEX TASKS

Design, engineering, writing, video production, music composition

[Kittur et al. 2013, Kulkarni et al. 2012]
Today: Three efforts to computationally scale up the complexity, interdependence, and sociotechnical infrastructure for crowd work.
EXPERT CROWDSOURCING
WITH FLASH TEAMS

Daniela Retelny, Sébastien Robaszkiewicz, Alexandra To, Walter Lasecki, Jay Patel, Negar Rahmati, Tulsee Doshi, Melissa Valentine, Michael Bernstein. UIST 2014. Best paper award.
COULD WE CROWDSOURCE...
THE DESIGN PROCESS, STARTING FROM A NAPKIN SKETCH, IN ONE DAY?
COULD WE CROWDSOURCE...

AN ANIMATED VIDEO IN 48 HOURS?
COULD WE CROWDSOURCE...
AN ENTIRE MOOC PLATFORM
IN 24 HOURS?
CROWDS OF EXPERTS

Mechanical Turk

- microtask worker
- microtask worker
- microtask worker
- microtask worker
- microtask worker

Upwork

- programmer
- designer
- video editor
- musician
- statistician
Microtask techniques do not leverage diverse skills and expertise.

Expert crowd work is independent and uncoordinated.
Self-managed teams are inefficient, riddled with frustrated members, and poorly coordinated. [Bunderson and Boumgarden 2010]
Lightweight team scaffolds significantly outperform pipelined and self-managed efforts. [Valentine and Edmonson 2012]
Could we combine the management strength of team scaffolds with the scale and interactivity of computing?
FLASH TEAMS

Crowdsourcing infrastructure for creating and guiding on-demand teams of diverse experts
FLASH TEAMS

Computationally-guided teams of crowd experts supported by lightweight, reproducible and scalable team structures.
SEQUENCE OF LINKED TASKS

1. Low-fi Mockup
2. Revised Mockup
3. Heuristic Evaluation
4. High-fidelity Prototype
SEQUENCE OF LINKED TASKS

- Low-fi Mockup
  - UI Designer

- Revised Mockup
  - UI Designer

- Heuristic Evaluation
  - UX Researcher

- High-fi Prototype
  - Developer
Revised Mockups
UI Designer
Input: low-fi mockups
Output: revised low-fi mockups
Goal: 90min
Low-fi Mockup
UI Designer
Input: napkin sketch
Output: low-fi mockups
Goal: 1hr

Revised Mockup
UI Designer
Input: low-fi mockups
Output: revised low-fi mockups
Goal: 2hrs

Heuristic Evaluation
UX Researcher
Input: low-fi mockups
Output: heuristic evaluation
Goal: 1hr

High-fi Prototype
Developer
Input: low-fi mockups, HE
Output: high-fi prototype
Goal: 4hrs
Heuristic evaluation

Low-fi prototype development

User testing

Hi-fi prototype development

(revised) HI-fi prototype

(revised) low-fi mockup

User study report

(revised) hi-fi prototype
RUNNING A FLASH TEAM

1. Introspect on the team composition & convene experts from the crowd

2. Convey the team through the workflow, shepherding files between tasks and sharing schedule updates

Upwork

programmer
designer
video editor
musician
statistician
Computational Affordances of Flash Teams

- Modularity → Scale
- Elasticity → Grow + shrink
- Pipelining → Optimize
- Planner → Create on-demand
MODULARITY

REPLICATE TEAM STRUCTURES AT SCALE
MODULARITY

REPLICATE TEAM STRUCTURES AT SCALE
<table>
<thead>
<tr>
<th>MODULARITY</th>
<th>COMBINE TEAMS TO FORM ORGANIZATIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>DESIGN</td>
<td><img src="#" alt="Design Diagram" /></td>
</tr>
<tr>
<td>DESIGN</td>
<td><img src="#" alt="Design Diagram" /></td>
</tr>
<tr>
<td>DESIGN</td>
<td><img src="#" alt="Design Diagram" /></td>
</tr>
<tr>
<td>EBOOK</td>
<td><img src="#" alt="Ebook Diagram" /></td>
</tr>
<tr>
<td>ANIMATION</td>
<td><img src="#" alt="Animation Diagram" /></td>
</tr>
<tr>
<td>EDUCATION</td>
<td><img src="#" alt="Education Diagram" /></td>
</tr>
</tbody>
</table>
MODULARITY

COMBINE TEAMS TO FORM ORGANIZATIONS
Is it ok to leave DRI on this slide and next slide even though I don’t mention it when I speak?
Elasticity enables growth by dynamically adding:
Extra workers to complete job on time
Workers with specialized skills
PIPEDLINING
PASS ALONG INCOMPLETE RESULTS
PIPELINING
PASS ALONG INCOMPLETE RESULTS

- Lo-fi v1
- Low-fidelity prototype v2
- Heuristic evaluation
- Development v1
- Development v2
- User testing
“I have a napkin sketch of a design, and I’d like an animation describing the idea.”
CREATION BY REQUEST

Synthetic team created from compatible blocks from previous teams.
CREATION BY REQUEST

Translate blocks into a STRIPS action planning problem, which utilizes efficient boolean satisfiability solvers.
CREATION BY REQUEST

Translate blocks into a STRIPS action planning problem, which utilizes efficient boolean satisfiability solvers.
FOUNDARY

Web platform that allows:

- Requesters to **author** flash teams
- Team members to **track the progress of tasks**
AUTHORING IN FOUNDRY
AUTHORING IN FOUNDRY
AUTHORING IN FOUNDRY
AUTHORING IN FOUNDRY
AUTHORING IN FOUNDRY

HI-FI PROTOTYPE
4 HRS

LOGO DESIGN
1 HR 30 MIN
Start Team?

Are you sure you want to begin running App Design Team?

- Cancel
- Start the team
Welcome Michael Bernstein!
Your role: UI Designer

You can now start Low-fi Mockups task.

Chat
SEE WHO'S ONLINE

Hi Team! Let's get started!
Stanford HCI (Author)
1/16/2015, 9:57:57 PM

I'm ready!
Michael Bernstein (UI Designer)
1/16/2015, 11:36:48 PM
FOUNDRY AS MANAGER

Welcome Michael Bernstein!
Your role: UI Designer

Progress Status:

Your task is in progress.
Welcome Michael Bernstein!
Your role: UI Designer

Your task is delayed.

Progress Status:

Project overview
FOUNDRY AS MANAGER

Welcome Michael Bernstein!
Your role: UI Designer

APP DESIGN TEAM

Go to Google Drive™ folder

Progress Status:

You can now start Revised low-fi mockups task.

Project overview
FOUNDARY IN SUM

Author structured, modular representations of flash teams
Grow, shrink, pipeline and recombine the flash teams
Recruit from Upwork
Maintain situational awareness as the team works
FLASH TEAM EXAMPLES

Recruited from paid crowd marketplace Upwork

Three team types:
- Napkin sketch (design & web programming)
- Animation (video making)
- MOOC (online education)
Objective: replication of flash teams across different inputs
USER-TESTED HI-FI PROTOTYPES IN ONE DAY

<table>
<thead>
<tr>
<th>Design Goal</th>
<th>Completion time</th>
<th>Team size</th>
<th>Total cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emotion tracking</td>
<td>31:30</td>
<td>3</td>
<td>$744.48</td>
</tr>
<tr>
<td>Event bullet board</td>
<td>18:00</td>
<td>5</td>
<td>$1270.28</td>
</tr>
<tr>
<td>Social meetups</td>
<td>23:10</td>
<td>5</td>
<td>$1200.97</td>
</tr>
</tbody>
</table>

IceBreaker is an app that helps you find people in your location that have friends or interests in common with you. You can see all the people around you that share at least one friend or one interest with you. If you click on one of them, you'll land on the profile page. You can see the detailed list of interests or friends in common, and you can start a chat conversation. If you click on an interest or a friend in common, you'll see a list of all the people in your location that share that particular interest...
**Objective:** explore how flash teams can support creative outputs and non-engineering domains

celebrate the work of Professor Terry Winograd, and his building a computer in ...
Objective: explore how flash teams can support creative outputs and non-engineering domains.
Objective: compose multiple modular team structures to complete a large scale project in 1 day
## MOOC COMPLETED IN 1 DAY

<table>
<thead>
<tr>
<th>MOOC component</th>
<th>Completion time</th>
<th>Video length</th>
<th>Total cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diaphragm singing</td>
<td>19:20</td>
<td>1 min 49</td>
<td>$1,597.32</td>
</tr>
<tr>
<td>Portrait photography</td>
<td>19:00</td>
<td>1 min 30</td>
<td>$741.58</td>
</tr>
<tr>
<td>Towers of Hanoi</td>
<td>11:30</td>
<td>1 min 24</td>
<td>$446.49</td>
</tr>
<tr>
<td>Web platform</td>
<td>13:00</td>
<td>N/A</td>
<td>$1015.80</td>
</tr>
</tbody>
</table>
ACCENTURE’S FLASH TEAMS

Project #1: Marketing video
177 work hours

Project #2: Mobile application
54 work hours

Project #3: Business dev. video
132 work hours
ACCENTURE’S FLASH TEAMS

Result:
1/8th – 1/6th cost of agency estimate for equivalent quality
Are Flash Teams Effective?

Field Experiment

Do flash teams complete tasks equally effectively but in less time?

Controlled experiment: 22 experts across six napkin sketch teams (UI design, UX research, web dev)

Flash teams vs. self-managed teams
Task: party planning mobile web application

Input: Napkin sketch

Requested time limit: 13 hours

Measured: total number of work hours across team
FLASH TEAMS VS. SELF-MANAGED TEAMS

Flash teams:
full Foundry with flash team workflow

Control teams (self-managed):
full Foundry with just one 13hr block
FLASH TEAMS:
50% FEWER WORK HOURS

Flash teams (mean 13hr2min) are significantly faster than self-managed teams (mean 23hr47min), $p=0.05$

The slowest flash team finished in fewer hours than the fastest team in the control condition
Flash teams introduce computational infrastructure for crowdsourcing diverse, on-demand teams of experts.
Crowd Work: Promise or Peril?

Crowdsourcing generated excitement by framing itself as a populist form of information work.

Yet platforms such as Amazon Mechanical Turk place workers inside a reliable, steadily humming infrastructure, making it difficult for workers to make their voice heard.

[Irani and Silberman '13]
“WHAT ABOUT A UNION?”

The metaphor of unionization is largely built around offline organizations, and needs to be redesigned for digital labor.

- new workers join the market daily and stay for only days or weeks
- work contracts last only minutes
- no way for workers to enforce behaviors on other workers
“If by ‘union’ you mean a ‘labor union’, I would not feel comfortable taking part. It runs against my grain because I am an individualist. I do not want to feel forced to go along with the ‘majority thinking’ of the leaders within a labor union. I have never been a member of a union and hope to continue along my merry way. I consider myself self-employed...not working for anyone in particular.”

CROWD COLLECTIVE ACTION

One year of ethnography with crowd workers, understanding and building relationships. This work led to the creation of:

Powering change on MTurk

We are a community of 485 Turkers and growing...!
1) Idea

Start an email campaign to Jeff Bezos letting him know who we are off/on mTurk, and what we are looking for on mTurk.

Posted 2 months ago by dark_bird_of_paradise.

2) Vote

3) Discuss

Dear Jeff Bezos

Posted 4 months ago.

We are writing to let you and the rest of the world know all about who we are. The intent is for you to see that Turkers are not only actual human beings, but people who deserve respect, fair treatment and open communication.

Join the discussion and letter writing!

Wiki page for authoring the description and goals

This is the page that will host the letters you send to us and publicize the movement:

Letters »
THE SAME FORCES THAT MAKE IT EASY TO GATHER ON THE WEB...

ALSO MAKE IT EASY TO DISPERSE.
"So, it seems no one is interested [...] [a Turker] just says we're doing it wrong, but won't say how to do it right, and no one else has input."
“I’m sorry to see the core document get excessively mired in technical detail in what purports to be a high-level document on ethics. I certainly hope this document sees fewer revisions than it will if it ties any of its own legs directly to the five-year-old ‘beta test’ which Amazon calls ‘policy’.”
scylla and charibdis...
STRUCTURED HUMAN SCRIPTS

These kinds of publics require special action to preserve their kinetic energy. For example:

- debates with deadlines
- act and undo

This labor could not have been written into software: it consists of human scripts undertaken by a trusted party.
SUCCESSES SO FAR

532 verified workers signed up (pseudonymously)
7,000 unique visitors
32,000 views
22 ideas for action
2 live campaigns
ETHICAL RESEARCH GUIDELINES

First campaign: curb poor academic research practices

Spawned when an IRB-approved economics researcher ran an experiment to inject false information into Turkopticon

Wiki-written guidelines covering fair pay, rejection, and IRB arbitration for poor requesters
Guidelines for Academic Requesters

About the project [edit]

Version 1.0

"Treat your workers with respect and dignity. Workers are not numbers and statistics. Workers are not lab rats. Workers are people and should be treated with respect." - turker "T", a TurkOpticon moderator

This document consists of this main Guidelines for Academic Requesters page, and several subpages with important additional details, which are referenced with "Read more" links at relevant points throughout this main page:

- Basics of how to be a good requester
- Fair payment
- Resources for communicating with requesters
- Links to other resources on AMT and online research ethics
- Meta: Maintaining the guidelines

Contents [hide]

1 About the project
   1.1 Goal: Guidelines that IRB will use to approve responsible AMT research
2 For Turkers: what can you do when these guidelines are violated?
3 Guidelines
   3.1 Clearly identify yourself to give workers a sense that you are accountable and responsible
   3.2 Provide reasonable time estimates
   3.3 Approve work as soon as possible
   3.4 Maintain worker privacy
   3.5 Abide by AMT Terms of Service

http://guidelines.wearedynamo.org – 216 signatures so far
Dear Jeff Bezos,

We are writing to let you and the rest of the world know all about who we are. The intent is for you to see that Turkers are not only actual human beings, but people who deserve respect, fair treatment and open communication.

Donate

Donations will go towards reaching out to more Turkers via a HIT on MTurk (a 3 minute paid vacation). Raised: $60

21 letters
Amazon's Mechanical Turk workers protest: 'I am a human being, not an algorithm'

A Christmas email campaign is asking Amazon's CEO Jeff Bezos to improve terms for workers providing cheap digital labour.

Amazon's Mechanical Turk workers be treated like humans

by Jon Fingas / @jonfingas / December 3rd 2014 at 1:39 pm

Amazon's Turkers Kick Off the First Crowdsourced Labor Guild
If flash teams are a new form of work collective, Dynamo is a new form of counterbalance.

Our goal: design systems for collective action in crowd work
CROWD RESEARCH: UNLOCKING THE GATES TO THE IVORY TOWER

ongoing work with Rajan Vaish, Geza Kovacs, Ranjay Krishna, Sharad Goel, and James Davis
SETTING OUR SIGHTS HIGHER

Must we restrict ourselves to research problems that are solvable alone or in small groups?

Must we deny access to motivated aspiring researchers?

Could people around the world work together to:

Build a new crowdsourcing platform?

Design and run hundreds of parallel experiments?

Develop computer vision algorithms?
WELL-STATED PROBLEMS

Polymath project
[Gowers and Tao]
[Cranshaw and Kittur 2011]

FoldIt
[Cooper et al. 2010]
Could we invite anyone from around the world to participate, and crowdsource large-scale, open-ended research problems?

These are problems for which a felicitous approach to finding the solution cannot be known in advance.
EDUCATIONAL MISSION

Provide scaffolding so anyone can learn to do top-tier research

Participant motivations: research experience, stronger resume, solving interesting problems
CHALLENGE: COORDINATION

My Stanford group is eight students. And that keeps me pretty busy. Google managers are asked to have no more than seven reports. So how could we possibly run a research team of hundreds?
Eventually, we need to decide on an author order, and participants will need recommendation letters. How do we measure impact?

Assumption: advisor goes last in the author order, but other authors should be ranked by contribution to the project.
1097 signups, predominantly from India and United States
27% female, median age 21, average team 3 people
73% undergraduate, 22% masters, 4% PhD, 1% high school
71% engineering-oriented areas of study
THREE PARALLEL PROJECTS

HCI
Michael Bernstein, Stanford
Building a new crowd marketplace
THREE PARALLEL PROJECTS

Computer vision
James Davis, UCSC
Serge Belongie, Cornell
Hybrid crowd-computer vision algorithms
THREE PARALLEL PROJECTS

Data science
Sharad Goel, Stanford
Hundreds of experiments testing the wisdom of the crowd
RESEARCH IN PROGRESS

Three work-in-progress papers: two at UIST, one at HCOMP
Papers in preparation or under review...
COORDINATION STRATEGY

Exploration during the week, reset to argmax each weekend

Saturdays: team meeting + milestone opens
Thursdays: milestone closes
Fridays: peer feedback and ranking + staff collation
COORDINATION STRATEGY

Divergence
Every interested contributor submits a milestone, then peers upvote high-quality submissions
Tools: Telescope (Reddit clone), Wiki

Convergence
Empower active community members to create temporary ad-hoc teams
Tools: Google Hangout, Slack, Google Presentations

Crowd Research
- Milestone 20 Instructions, Open-Gov
  - Geza Kovacs 1 point 23 days ago 0 Comments
- Onboarding proposal, Onboarding
  - Dilukshi Gamage 1 point 12 days ago 0 Comments

karolina 1:57 PM
doing it now 😊
@channel: we will begin shortly!
here is the link to join
1:58 https://hangouts.google.com/call/atw4bueag6pt2munaykj2fvb
Each participant allocates 100 points across other contributors. This produces a credit network, but some groups of participants rally a few friends to vote for them and artificially increase their influence.
PAGERANK-BASED CREDIT

Approach: run a modified PageRank algorithm over the network.

Informally, PageRank identifies the universally-respected contributors, then weighs their votes more heavily. This process iterates until convergence.
rajandaish 9:01 AM
echo everyone!

seasondyb 9:02 AM
Hi!

Greetings from Seattle!

csarasa 9:02 AM
Hi!

sblogro986 9:02 AM
He!

sujeathpaddeddy 9:02 AM
Hi!

rainydayz8 9:02 AM
hi ^_^

meenalmandil 9:02 AM
Hi

ossolorzano 9:02 AM
hello
MILESTONE SELF-ASSIGNMENT

Click first: how does this work?
Each week, you (and/or your team) sign up for at least one milestone here on Trello. See more...

Understanding lives of workers
Try being a worker on oDesk (now Upwork.com) or other large project platforms.

Hello, world! Getting started with our code
Work on one of our open feature requests on GitHub

Related work/papers: read and comment
Read the MobileWorks paper

Read Flash Teams paper

Read paper on the future of crowdwork
ENGGINEERING

Apr 5, 2015 – Oct 26, 2015
Contributions to develop2, excluding merge commits

Contributions: Commits

<table>
<thead>
<tr>
<th>0</th>
<th>50</th>
<th>100</th>
<th>150</th>
</tr>
</thead>
<tbody>
<tr>
<td>May</td>
<td>June</td>
<td>July</td>
<td>August</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>dmorina</th>
<th>502 commits / 141,296 ++ / 365,266 --</th>
</tr>
</thead>
<tbody>
<tr>
<td>nistala</td>
<td>249 commits / 1,671,624 ++ / 1,442,878 --</td>
</tr>
</tbody>
</table>

#1

#2
PROTOTYPES AND STORYBOARDS
DATA ANALYSIS
1. Anyone can pitch an idea. If it gets enough support, it goes to the next election and needs majority support from both workers and requesters.
   - *(original) Direct democracy*: anyone can pitch a policy idea, and once it gets past a threshold of support (e.g., 1000 votes), it goes up on a ballot. *Twice a year*, ideas go out to a direct vote for everyone on the platform. If it gets majority support from both workers and requesters, it passes.

2. Members get elected as worker or requester representatives (3 each) to a panel. Tiebreaking from a 7th member (jointly elected president).
   - *(original) Representative democracy*: once a year, members of the platform can be elected as either worker or requester representatives for a small panel (e.g., six people). Anybody can pitch a policy idea, and once it gets past a threshold of support (e.g., 1000 votes), the elected representatives must discuss it and vote on it.

3. Wikimocracy: the site’s rules and policies are a wiki. Anyone can discuss, and if they edit, policies change directly.

4. Any idea that gets enough support enters a public one-month voting period. It’s completely voluntary to vote. *(Like a Kickstarter campaign.)*
   - *Original: Fast-paced referendums*: similar concept as direct democracy, but instead of per year, you do it as vote thresholds within a month (within time of posting), and it’s completely voluntary to vote. Kinda like a campaign on Kickstarter. Fast pace and flexible deadlines will help ideas continually flowing in.

5. For low-level changes, highlight the interface and suggest changes directly. *Upvote/downvote directly on the interface.*
LONGEVITY

YouTube views of weekly meetings

- **Weeks 1-10**: Decreasing views for all categories (HCI, Comp Vision, Data Science).
- **Week 13**: Significant spike in views for HCI.
- **Weeks 14-20**: Stabilization of views across categories.
- **Weeks 21-33**: Consistent views with minor fluctuations.
LONGEVITY

People reading and writing on Slack

Number of contributors

Week

Week

Reading
Writing
REASONS FOR DROP-OFF

Unable to catch up after exams
Time commitment
Lost teammates
Lacked required skills
Felt isolated
Other

PERCENTAGE OF RESPONDENTS

STATED REASON FOR LEAVING

Computer Vision  Data Science  HCI
Paper Writing a.k.a. “Michael loves editing intros.”
PAGERANK
PAGERANK
CROWD RESEARCH

We aim to...

- Tackle open-ended and messy research problems where a static interface won’t be enough
- Tackle big problems while mentoring new researchers
- Recognize contributions (more) fairly and share credit
THE FUTURE OF WORK IS...

Complex and interdependent

Advocating for pro-social outcomes

Solving open-ended challenges
THE FUTURE OF WORK IS...

Complex and interdependent

Advocating for pro-social outcomes

Solving open-ended challenges

Thanks to the NSF, Accenture Tech. Labs, Stanford Cyber, Brown Institute, Precourt, HPDTRP