Foundations

MICHAEL BERNSTEIN
CROWDRESEARCH.STANFORD.EDU
Today

- Digging our foundations
- #infra
Next week is CHI in Korea, so we’ll meet on Sunday.
Foundations

Prioritizing and focusing in on challenges
What we did

- We took the foundations that you submitted in weeks 3-7, focusing especially on recent excitement…
- Combined a few ideas…
- And weighed in with our perspectives as crowdsourcing researchers.

- The result: four foundations we will try immediately, and others we will do next.
Micro+macrotask market

- “Could the same marketplace scale from 2 to N people not just labeling images, but also Photoshopping my vacation photos or mixing my new song?!”

- Maintain the submission approach from microtask markets, which is focused on two to hundreds of replications, but find ways to make it accessible to both microtask and expert work
Adam Marcus: “Why would a programmer want to join a crowd work platform? The draw of AMT right now is that you can find jobs quickly. Can we enable that same benefit for programmers?”
Challenges

• How do we ensure high-quality results? Do you let an expert work for hours and submit? That seems risky. Should there be intermediate feedback mechanisms?

• How do you trust that someone is an expert?

• Does this look like AMT where any expert certified in an area can accept the task? Or like an oDesk negotiation?
Input and output transducers

- Tasks get vetted or improved by people on the platform immediately after getting submitted, and before workers are exposed to them. Results are likewise vetted and tweaked.
Challenges

• Cost: who pays for this? In other words, can this be done without hugely increasing the cost of crowdsourcing?

• Speed: is it possible to do this quickly enough to give near-immediate feedback to requesters? Like, 2–4 minutes?

  • Spamgirl: “The #1 thing that requesters love about AMT from her recent survey of requesters, is that the moment that I post tasks, they start getting done"
Challenges

• From Edwin: What happens when I have a task that I know is hard but I want workers to just try their best and submit? I’m OK with it being subjective, but the panel would just reject my task, which would have been frustrating.

• From Edwin: Could this help deal with people feeling bad when rejecting work? Maybe we need a new metaphor, like revision.
External quality ratings

- Metaphor of credit ratings: rather than just people rating each other, have an (external?) authority or algorithm responsible for credit ratings (A, B, C, etc.)

- Benefit: this reduces the incentive to get positively-biased five star ratings on everything — those ratings become meaningless
Challenges

• Is this a group/authority? For example, Wikipedia reviews are subjective and based on voting. Or is it an algorithm?

• If it’s a group, who pays for their time to review you?

• From Anand: “How do you do skills-based ratings, etc., without hindering tasks with a requirement to categorize them?”
Open governance

- Leadership shared by requesters, workers, (researchers?)

- Policy changes can be worked out by this group
Challenges

- Is it direct voting on everything? Or representative democracy?

- How exactly will this work?

- Can the research group have a hand here?

- If there are changes that require engineering effort, who executes that? Us? Other volunteers?
Next...

- #1 (for me): Open governance
- Empathy
- Mentorship
- Engaging with mobile users
- Price+quality joint model
- Features:
  - Recommendation
  - Wikified bug reports on tasks
Infrastructure advance team

Updates!
Milestone 8
Refine the foundations

- Take these four milestones, and propose specific mechanisms that respond to the challenges