Turkopticon
Motivation, design, status, lessons, next steps

M. Six Silberman, with Lilly Irani

25 Sep 2013
Motivation

The design of Mechanical Turk is not fair for workers.

Employers do not have to pay for work.

They do not have to give a reason for not paying. Workers can complain, but employers do not have to answer. Amazon charges employers for posting tasks. So it does not hurt Amazon if employers do not pay workers. Mechanical Turk keeps track of workers’ approval rates so employers can screen them.

There is no reputation system for employers. For example, how often does an employer not pay for work? Workers have no way to know.
Turkopticon is a third-party employer reputation system for Mechanical Turk.

The original goals of Turkopticon were:
(a) to call attention to the unfairness of Mechanical Turk and
(b) to prompt Amazon to build an employer reputation system into Mechanical Turk.

Turkopticon has two parts: a browser add-on and a web database application. The web application lets workers review employers. The browser add-on adds these reviews to the Mechanical Turk interface.
### Search: Keywords on Google.com (US)

<table>
<thead>
<tr>
<th>Requester</th>
<th>HIT Expiration Date: Sep 20, 2014 (52 weeks)</th>
<th>Reward: $0.08</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time Allotted:</td>
<td>16 minutes</td>
<td></td>
</tr>
<tr>
<td>HITs Available:</td>
<td>14979</td>
<td></td>
</tr>
</tbody>
</table>

### Are these two pictures of the same kind of place? (w/ clarified instructions)

<table>
<thead>
<tr>
<th>Requester</th>
<th>HIT Expiration Date: Sep 20, 2014 (52 weeks)</th>
<th>Reward: $0.02</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time Allotted:</td>
<td>10 minutes</td>
<td></td>
</tr>
<tr>
<td>HITs Available:</td>
<td>13882</td>
<td></td>
</tr>
</tbody>
</table>

### Quality Estimation from Arabic to English

<table>
<thead>
<tr>
<th>Requester</th>
<th>HIT Expiration Date: Jul 17, 2014 (42 weeks 5 days)</th>
<th>Reward: $0.10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time Allotted:</td>
<td>60 minutes</td>
<td></td>
</tr>
<tr>
<td>HITs Available:</td>
<td>11403</td>
<td></td>
</tr>
</tbody>
</table>
### Design

<table>
<thead>
<tr>
<th>Requester</th>
<th>HIT Expiration Date</th>
<th>Reward</th>
<th>HITs Available</th>
</tr>
</thead>
<tbody>
<tr>
<td>CrowdSource</td>
<td>Sep 20, 2014 (52 weeks)</td>
<td>$0.08</td>
<td>14989</td>
</tr>
<tr>
<td>Time Allotted:</td>
<td>16 minutes</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Requester</th>
<th>HIT Expiration Date</th>
<th>Reward</th>
<th>HITs Available</th>
</tr>
</thead>
<tbody>
<tr>
<td>Michelle Greene</td>
<td>Sep 20, 2014 (52 weeks)</td>
<td>$0.02</td>
<td>13883</td>
</tr>
<tr>
<td>Time Allotted:</td>
<td>10 minutes</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Requester</th>
<th>HIT Expiration Date</th>
<th>Reward</th>
<th>HITs Available</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chris Callison-Burch</td>
<td>Jul 17, 2014 (42 weeks 5 days)</td>
<td>$0.10</td>
<td>11403</td>
</tr>
<tr>
<td>Time Allotted:</td>
<td>60 minutes</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Design

Quality Estimation from Arabic to English

Requester: Chris Callison-Burch

- communicativity: 4.05 / 5
- generosity: 3.99 / 5
- fairness: 4.27 / 5
- promptness: 3.99 / 5

What do these scores mean?

Scores based on 98 reviews
Terms of Service violation flags: 1
Report your experience with this requester »
<table>
<thead>
<tr>
<th>AMT Requester</th>
<th>Rating [info]</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chris Callison-Burch/WikiTrans A32TTE4XXN6MQZ 79719</td>
<td>FAIR: 5 / 5</td>
<td>Great HITs</td>
</tr>
<tr>
<td></td>
<td>FAST: 5 / 5</td>
<td>Jun 02 2013</td>
</tr>
<tr>
<td></td>
<td>PAY: 5 / 5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>COMM: 5 / 5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>I was matching text across two articles, and every once in a while I got a blank one. I accidentally submitted to get to the next article when it dawned on me that the requester would need to pay or reject the hit (be patient with me, I'm new to this). I e-mailed the requester explaining my mistake, and they approved all my hits with the admonition to return blank hits from now on. Great pay, and great communication.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>May 29 2013</td>
</tr>
<tr>
<td>Chris Callison-Burch/WikiTrans A32TTE4XXN6MQZ 79130</td>
<td>FAIR: 5 / 5</td>
<td>definitely one of the best requesters out there.</td>
</tr>
<tr>
<td></td>
<td>FAST: 5 / 5</td>
<td>May 21 2013</td>
</tr>
<tr>
<td></td>
<td>PAY: 5 / 5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>COMM: 5 / 5</td>
<td></td>
</tr>
<tr>
<td>Chris Callison-Burch/WikiTrans A32TTE4XXN6MQZ 77843</td>
<td>FAIR: 1 / 5</td>
<td>hundred and some hits and never have been paid or heard from anyone.</td>
</tr>
<tr>
<td></td>
<td>FAST: 1 / 5</td>
<td>May 17 2013</td>
</tr>
<tr>
<td></td>
<td>PAY: 1 / 5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>COMM: 1 / 5</td>
<td></td>
</tr>
</tbody>
</table>

VIOLATES MTURK TERMS OF SERVICE [?]
Design

HIT ID
(optional)

This HIT is a violation of the Mechanical Turk Terms of Service

Requester Name
Chris Callison-Burch/WikiTrans

Requester ID
A32TTE4XXN6MQZ
(required)

How many HITs have you done for this requester?
None

Rate this requester on the following qualities, with 1 being worst among requesters you've had experience with and 5 being best among requesters you've had experience with:

- How fair has this requester been in approving or rejecting your work?
- How promptly has this requester approved your work and paid?
- How well has this requester paid for the amount of time their HITs take?
- How responsive has this requester been to communications or concerns you have raised?

Please briefly explain any issues you've had with this requester:
Status

Numbers

Browser add-on installs (since Aug 2012) 18344
Web application since 2008
  Registered users 21312
  Reviews 96806
  Employers reviewed 21851
  Users posting at least one review 8146 (38%)
Web application, last 3 months
  Reviews 15699 (16%)
  Employers reviewed 4322
  Users posting at least one review 2124
Average daily visits
  Web application 3721
  API (for browser add-on) 8034
Tasks with reviews out of the first 100 (19 Sep 2013) 96

Small fraction of total Mechanical Turk workforce
Probably most of professional Turkers
Status

Assessment

Goals:
(a) to call attention to the unfairness of Mechanical Turk — success.
(b) to prompt Amazon to build an employer reputation system into Mechanical Turk — failure (Amazon: „the community handles the problems“).

Turkopticon, combined with other forums and tools made by workers, make Mechanical Turk less unfair. But a built-in reputation system would be better.

There are problems with Turkopticon itself.
Our moderation system could be better.
We don’t respond adequately to all user feedback.
We have no good collaborative or democratic decision-making process.
We have no good way to let other programmers help us.
The site is slow and buggy.
The technology is „old“ (2008).

There is a climate of distrust and fear around Mechanical Turk, and Turkopticon too. Can this be fixed, or do we need to start over?
Lessons

About software
If you listen to people and build what they want, they will use it.
You must keep listening to people and maintain what you built.
Maintenance is both technical and social.
The social and technical parts of maintenance both take time.

About markets as institutions
Most workers and employers have good intentions, but not all.
The small selfish fraction affects the market, and must be designed for.
No system can solve all problems. Human judgment (administration) is needed.
To maintain trust, there should be a record of administrative judgments and why they were made.

About Amazon
Amazon is not all-powerful. They do not have infinite developer hours.
What now? Turkopticon

There will not be a commercial version. Turkopticon is not-for-profit.

Try to make decision making more democratic
Try to respond more systematically to user feedback
Try to find ways to let other programmers help us
Accept donations
Create a non-profit organization
Improve moderation
Collect more „objective“ data in reviews
Make code „open source“
Make site faster
Fix search
Let people delete and edit comments while maintaining a record

I would like for Turkopticon to become unnecessary.
What now? Another crowd work is possible

Understand the situation.
  The situation is complex; e.g., many US crowd workers don’t want minimum wage laws for crowd work. (They are afraid they will lead to less work.)
Research
  Demographic („quantitative“: how many? who?)
  Ethnographic („qualitative“: how? why?)
  Build shared data, knowledge, theory
Cross-sector work groups

Build a „community of practice“.
  Workers, employers, platform builders and operators, policy makers, researchers, ...

Build new models for crowd work arrangements.
  Processes, markets, technologies, organizations, contracts, laws, ...

Build new systems.

Build and maintain cross-sector conversations that lead to learning and action.