

Dynamics and governance of crowd work markets

M. Six Silberman

SUMMARY OF PROPOSAL

Are current crowd work arrangements “ethical”? Many commentators and researchers consider current systems—especially Amazon’s Mechanical Turk platform (AMT), the most well-known crowd work system—unfair to workers. Workers themselves have complex opinions on the subject. The work described in this proposal considers these issues in terms of *dynamics* and *governance*. This view raises questions such as: How do small-scale elements of the system’s design shape individual transaction outcomes and large-scale patterns, such as the relative power of different stakeholder groups and the distribution of the gains from trade? How does the system respond to stakeholder concerns? How does the system seize opportunities to change in ways that yield better outcomes for stakeholders? In this perspective, ethical issues arise in a system when it does not respond adequately to the concerns of its stakeholders. The dynamics of crowd work systems have not been discussed much in research, and governance mechanisms are poorly developed in current practice. To help address these deficiencies I propose a three-part research strategy. First, to explore rigorously the dynamics of existing and possible crowd work systems, I will build an agent-based model of an abstract crowd work market. Second, to contribute to the governance of existing crowd work systems, I will build an alternative interface to AMT. The alternative interface will aim to (1) reduce the power imbalance, resulting partly from asymmetric information, between workers and employers; (2) support better communication between workers and employers; and (3) model effective strategies for responding to stakeholder concerns. This work will build on my experience as co-designer of Turkopticon, a system used by approximately 20,000 AMT workers to rate employers. Third, in writing about this work for the computing research community, I will link crowd work discourse to the literatures on market design and institutional analysis and development in economics, in which theory on dynamics and governance is well-developed.

The central intellectual contribution of this work is to understand how systems design—database, interface, interaction, and policy design—affects stakeholder goals and strategies, individual transaction outcomes, and large-scale, long-term market dynamics and equilibria in crowd work systems. This relationship is technical, economic, and social; this work focuses on enriching our understanding of crowd work markets as economic and social systems. Additionally, it will link crowd work research to contemporary economics, enriching economic discourse in computing. This work also aims to materially improve practice. It aims to improve outcomes in AMT by reducing an information asymmetry, and to model governance strategies for other crowd work systems. It may also contribute to the technical, social, and theoretical groundwork for a future crowd work market managed by stakeholders.

BACKGROUND

Ethical concerns about crowd work have surfaced in the peer-reviewed and informal computing literature [1–8], among legal scholars [9–13], and in the popular media [14–19]. In our 2010 paper [1], Lilly Irani, Joel Ross, Bill Tomlinson and I, focusing on Amazon Mechanical Turk (AMT), describe some problems faced by crowd workers:

- Uncertainty about payment
- Unaccountable and seemingly arbitrary rejections [i.e., non-payment]
- Fraudulent tasks
- Prohibitive time limits
- Long pay delays
- Uncommunicative requesters [i.e., employers]
- Cost of requester and administrator errors [and technical issues] borne by workers
- Low pay

They also list relevant open research and design questions and topics, including:

- How does database, interface, and interaction design affect individual outcomes and market equilibria?
- What are the economics of fraudulent tasks (scamming and spamming)?
- What decision logics are used by buyers and sellers in human computation markets?
- What’s fair in paid crowdsourcing?
- Gaps remain in our knowledge of worker demographics and motivations, despite survey research [4, 47]
- Comparative and longitudinal studies of crowd work platforms

Several demographic and ethnographic studies of crowd workers are completed but not published; underway; or in planning (D. Martin, M. L. Gray, L. Irani, W. Poster, C. Aragon, pers. comm.). Research into the other questions

remains nascent. The question of fairness in crowd work systems has been raised in computing research [6–8], and discussed in relation to employment regulation in legal research [10–13]. But its implications for *design* have not been addressed directly. Similarly, the broad question of the relationship between design, individual outcomes, and market equilibria has been raised [1, 2], and some hypotheses about AMT specifically have been advanced [4, 5]. But the questions have not been pursued with much vigor, and we lack a rich understanding of this relationship.

Ethical issues in crowd work have been discussed with some excitement in the popular press [14–19], but this discussion is somewhat one-dimensional. It can be summarized roughly as revolving around the question of whether AMT—the platform most often discussed—is a “digital sweatshop.” Sweatshops are assumed to be bad, so if AMT is a sweatshop it—and perhaps all crowd work—must also be bad. This portrayal, and the conclusions drawn by many commentators in favor of reform or regulation, are contested regularly by workers. Workers appear to fear that reform or regulation will reduce the amount of work available [e.g., 20, 21], but these concerns are rarely discussed.

This tension raises a clear question for designers: can worker concerns—up to and including allegations of “exploitation”—be addressed without realizing these fears? Can we build a crowd work market that is “effective, efficient, *and fair*” [6: 1309; emphasis added]? What tradeoffs are involved [22]?

At stake is the socioeconomic potential of crowd work. Like Kittur et al. [6: 1301], I foresee “a world in which crowd work continues to expand.” Institutions have incentives to cut labor costs and expand operational capabilities, and flexible employment appeals to many workers. These conditions will drive the growth of crowd work. Crowd work arrangements that support viable livelihoods could indeed unlock “an incredible number of opportunities for careers in skilled work” [6: 1301], contributing to socioeconomic recovery in the recession-hit global North, sustainable socioeconomic development in the South, and socioeconomically beneficial globalization [cf. 23–26]. Indeed human computation and human-centered computing researchers have already begun to explore the potential benefits of bringing crowd work to “developing regions,” “the bottom of the pyramid,” and “low-income workers” [27–31]. Yet Kittur et al.’s other scenario, in which crowd work falls “into an intellectual framing focused on low-cost results and exploitative labor,” with workers assumed to be “interchangeable and untrustworthy, having low or static skill sets and strong motivations to shirk” [6: 1301], is already realized in AMT. As perhaps the oldest open crowd work market, AMT has served as infrastructure and focus for much crowd work research. It has set de facto design and administrative precedents—and legal precedent may be forthcoming (*Otey v. CrowdFlower*). To realize crowd work’s potential, we need a new approach—one that considers both employers’ and workers’ concerns [8].

PROPOSED RESEARCH

This work aims to contribute to a future of crowd work that is “effective, efficient, and fair” [6: 1309]. I focus on the third criterion, which is the least studied. Because particular outcomes or market dynamics may be desirable in one context and undesirable in another, it is not possible to build *inherently* “ethical” or “fair” crowd work systems. Rather, fairness and other economic and social goals will be achieved by building (1) shared knowledge about the relationship between design and outcomes and (2) technologies, institutions, and collective processes that support effective governance in practice. Effective governance will use knowledge about the relationship between design and outcomes to respond substantively to stakeholders’ concerns. Theoretically then, this work aims to contribute to a design-actionable understanding of the relationship between the technical, economic, and social affordances and dynamics of crowd work systems. Practically, it aims to contribute directly to the development of concrete technologies, institutions, and collective practices that support the effective governance of crowd work systems. The proposed work thus includes research *questions* and design *goals*. The questions motivate the theoretical investigation of the dynamics of crowd work markets. The goals motivate the design of a concrete web service.

Research questions. The dynamical question driving this work is: How does systems design—database, interface, interaction, and policy design—influence stakeholder goals and strategies, individual transaction outcomes, and large-scale market dynamics in crowd work systems? In this frame, this work aims to answer questions of the form “How might [design variables] affect [outcome variables]?” Variables of interest are listed in the table below.

Design variables	Outcome variables
<ul style="list-style-type: none"> • the option to “reject” (i.e., not pay for) work • a “majority rule” algorithm for automating rejection decisions • a clear set of criteria for rejecting work • a mediated rejection “appeal” process • a worker reputation system <ul style="list-style-type: none"> • based on approval (i.e., payment) rate • based on references from employers • based on data about past work 	<ul style="list-style-type: none"> • posted task prices • posted task time estimates • employer-estimated wages • workers’ a priori wage estimates • workers’ a priori estimates of payment probabilities • actual payment probabilities • actual wages • ratios of wages to costs of living • total value of tasks posted

(table continues on next page)

<ul style="list-style-type: none"> • an employer reputation system <ul style="list-style-type: none"> • based on references from workers • based on data about past tasks (e.g., posted prices, nonpayment history, posted vs. actual completion times) • integration of employer reputation into the task browse/search interface • worker and employer profiles • anonymity • support for worker collaboration (e.g., forums, chat rooms) • support for worker-employer communication about tasks 	<ul style="list-style-type: none"> • distribution of the gains from trade between stakeholders • employers' estimates of worker probability of shirking • actual probability of shirking • prevalence of deceptive employer behavior • perceptions of trustworthiness (i.e., estimates of others' strategies) • workers' perceptions of crowd work as a viable livelihood • work quality • the distribution of skills among workers • perceptions of fairness among all stakeholders
--	---

I ask “how might...?” rather than “how does...?” because I want to consider rigorously what is possible, not only to study what already exists. The forthcoming ethnographic studies of crowd work already promise substantial new insights into existing conditions. But no open crowd work market yet exists with, e.g., a robust employer reputation system. To inform future design means in part to inform choices between different *possible* designs, few of which may yet exist. To answer these questions is to map the design space of possible future crowd work markets.

Design goals. The design goal of this work is to help build effective governance of crowd work systems. To achieve this goal, I choose to intervene in AMT. I have four reasons for this choice. First, AMT is a central fixture in crowd work research and practice. While its flaws are well known among computing researchers, many use it as infrastructure in building new systems, and it is growing in importance as a data collection tool for social scientists. Second, AMT is an open market. Some “managed” crowd work systems [33, 34] address work quality issues and worker concerns by restricting the worker pool. But this approach sacrifices the broad reach and low transaction costs of an open market, crucial to crowd work’s broad appeal. Thus the practical challenge for crowd work research is to build an open market with effective governance. Intervening in AMT is a step toward that goal. Third, I have experience intervening in AMT. In 2008 Lilly Irani and I built Turkopticon, a web service that allows AMT workers to review employers and to consider others’ reviews when choosing tasks [7, 34]. Fourth, building a new market with effective governance will likely require social organizing beyond the scope of a dissertation-sized project.

Thus the goal “contribute to effective governance of crowd work” is narrowed to “contribute to effective governance of AMT.” This is feasible: sites such as the worker forum Turker Nation and Turkopticon have contributed to this end, giving employers material incentives to improve pay, pay promptly, reject work with care and accountability, and communicate with workers. But many potential improvements remain. A rough implicit consensus on potential governance improvements to AMT appears to exist among workers [e.g., 35–37], and to comprise two main tasks: (1) reduce the still significant power imbalance, resulting from asymmetric information availability, between workers and employers, by collecting and aggregating reliable information on task characteristics and requester behavior, and making the information available to workers in choosing tasks; and (2) support better communication between workers and employers. These tasks form the *concrete* goals of this work. If the alternative interface supports the emergence of robust self-governance practices, AMT will come to form the core of a larger system—a core that could later be replaced by an infrastructure under stakeholder control.

Methods. To answer the research questions, I will build an agent-based model (ABM) of the design space of possible crowd work markets. ABMs represent phenomena as “evolving system[s] of autonomous interacting agents” [38; see also 39–44]. ABM is especially suited to exploring how “macro” dynamics emerge from the “micro” interactions of individual agents [41]. In the social sciences, ABM is used to explore the consequences of individual decision making processes too complex for formal mathematical analysis [38]. To “member-check” the model [38], I will prepare a public description and solicit feedback from workers, employers, and researchers directly and through sites in the AMT “ecosystem” (e.g., Turker Nation, Turkopticon).

To reduce the power imbalance and support better communication between workers and employers, I will build an alternative interface for AMT. Building on experience with Turkopticon [7], this service will include technical and policy support for ongoing, substantive stakeholder engagement and redesign based on stakeholder input.

Design of the agent-based model. I model a market with two kinds of agents, workers and employers, and one kind of object, tasks. Tasks are generated exogenously and first appear in an employer’s task queue. Tasks appearing in an employer’s task queue have a value. This is the value to the employer of having the task completed. Employers price their tasks and post them to the market queue. Workers select tasks in a random order. After a task is done, the employer decides whether to pay the worker or refuse payment. Reflecting the possible presence of an employer reputation mechanism [1, 2, 4, 7], workers may remember outcomes with employers and may share that information with other workers. Reflecting the diverse goals of market participants [1], workers and employers may seek to maximize their payoffs over the short or long term or may seek to realize fair outcomes. I use several definitions of fairness based on the experimental economics literature [22, 45–46]. Finally, reflecting the fact that some AMT

workers rely on crowd work income for survival [1, 47], workers may have nonparticipation costs—i.e., costs of not participating in the market. Nonuniform nonparticipation costs affect the calculations of participants concerned with fair outcomes. In more complex scenarios I model task completion explicitly by introducing the notions of quality of work output; worker effort, skills, and learning; and appropriate skills for a task. Work of low quality may be of lower value to employers, who may be more likely to reject it. Employers may remember the quality of work produced by different workers and share that information. Task value may vary with required skills. Workers may be “born” with different skills, and may be able to improve their skills by completing tasks requiring those skills. Finally, uneven ab initio distribution of skills and skill values may affect what is considered “fair.”

The model will offer insights into how these factors might affect the dynamics of current and future crowd work markets. I expect the relationships among these factors to be complex. Consider for example the addition of the ability to reject work to an existing market. This feature raises the cost of worker shirking and may thus raise work quality. But it also enables deceptive employer behavior and may thus lower actual wages. This may price workers producing high-quality work out of the market, ultimately lowering work quality [4, 5, 48]. Rather than testing all possible hypotheses about the variables listed above, the report of the model results will explore in detail the dynamics most salient to the design of future crowd work markets.

Initial design of the alternative AMT interface. The alternative AMT interface will comprise a scraper, a database, an API, and a user interface. The scraper will scrape task data from AMT and store it in the database. (This does not appear to violate the AMT TOS.) The API will make these data available. The system will have access to the Turkopticon review database. The initial design of the user interface will allow users to:

- Publicly appeal rejected work
- Submit data about tasks, including:
 - task type
 - if the task violates the AMT TOS, and if so, how
 - if the task is “broken” (i.e., has technical problems that prevent workers from completing it)
 - if their work for the task was rejected, and if so, if they think that rejection was fair or unfair
 - how quickly the task paid, if it paid
 - the time it took to complete the task (posted completion time estimate will be scraped from AMT)
 - number of public rejection appeals for the task
- Browse and search for active tasks by any of the above criteria, or by Turkopticon review data
- Discuss tasks and employers and subscribe by email to those discussions
- Maintain public profiles
- Express trust or distrust in data posted by others
- Propose, second, discuss, vote on, and collectively prioritize new interface features

The system will be written in Clojure with the Compojure framework.

Writing. The dissertation will describe the preexisting state of affairs in crowd work practice, research discourse, and popular media discourse with respect to issues of ethics and governance in crowd work systems; describe, assess, and draw lessons from Turkopticon, the agent-based modeling study, and the alternative AMT interface; and consider all of the foregoing through the lens of the literatures on market design (or “economic engineering”) [49–51] and institutional analysis and development (IAD) [52–54]. It will contribute to crowd work research by offering lessons from two long-running projects “in the wild” and by linking the theory bases in human computation (HC) and human-centered computing (HCC) to contemporary economics. It will form the basis for a series of submissions to HC and HCC venues (e.g., HCOMP, TEAC, CHI, CSCW, ToCHI, JCSCW).

Complementary activities. I will continue to help maintain Turkopticon and produce the crowdwork-ethics newsletter. These activities will inform the research. For more information on these projects, see the included CV.

Timeline.

Oct – Dec 2013	Develop agent-based model
Jan – Mar 2014	Write public report of model results; discuss with stakeholders; revise model
Apr – Jun 2014	Write final model results report, submit to <i>ACM ToCHI</i> or <i>TEAC</i> ; advance to candidacy
Jul – Sep 2014	Write and defend dissertation proposal
Oct – Dec 2014	Build alternative AMT interface; publicly announce
Jan – Dec 2015	Iterative participatory development of alternative AMT interface
Jun – Aug 2015	Review relevant market design and IAD literature
Sep 2015 – May 2016	Write and defend dissertation

REFERENCES

1. Silberman, M. S., J. Ross, L. Irani, B. Tomlinson. Sellers' problems in human computation markets. *HCOMP '10*: 18.
2. Silberman, M. S., L. Irani, J. Ross. Ethics and tactics of professional crowdwork. *XRDS* 17(2): 39, 2010.
3. von Ahn, L. Work and the internet. *Luis von Blog*, 26 July 2010.
4. Ipeirotis, P. G. Mechanical Turk, low wages, and the market for lemons. *A Computer Scientist in a Business School*, 27 July 2010.
5. Bederson, B., A. Quinn. Web workers unite! Addressing challenges of online laborers. *CHI '11 EA*: 97.
6. Kittur, A., J. V. Nickerson, M. S. Bernstein, E. M. Gerber, A. Shaw, J. Zimmerman, M. Lease, J. J. Horton. The future of crowd work. *CSCW '13*: 1301.
7. Irani, L. C., M. S. Silberman. Turkopticon: interrupting worker invisibility in Amazon Mechanical Turk. *CHI '13*: 611.
8. O'Neill, J., D. Martin. Relationship-based business process crowdsourcing? *INTERACT '13*: 429.
9. Zittrain, J. Ubiquitous human computing. *Phil. Trans. R. Soc. A* 366: 3813, 2008.
10. Cherry, M. A. Working for (virtually) minimum wage: applying the Fair Labor Standards Act in cyberspace. *Alabama L. Rev.* 60(5): 1077, 2009.
11. Cherry, M. A. A taxonomy of virtual work. *Georgia L. Rev.* 45: 951, 2010.
12. Felstiner, A. Working the crowd: employment and labor law in the crowdsourcing industry. *Berkeley Journal of Employment and Labor Law* 32(1), 2011.
13. Felstiner, A. Grappling with online work: lessons from cyberlaw. *St. Louis Univ. Law J.* 56(1): 209, 2011.
14. Cushing, E. Dawn of the digital sweatshop. *East Bay Express*, 1 Aug 2012.
15. Dobson, J. Mechanical Turk: Amazon's new underclass. *Huffington Post*, 19 Feb 2013.
16. Hodson, H. Time to focus on the welfare of online workers. *New Scientist* No. 2903, 8 Feb 2013.
17. Folbre, N. The unregulated work of Mechanical Turk. *New York Times: Economix*, 18 Mar 2013.
18. Reynolds, R. Is Mechanical Turk a digital sweatshop? *The Conversation With Ross Reynolds*, 6 Feb 2013.
19. Leonard, A. The internet's destroying work—and turning the old middle class into the new proletariat. *Salon.com*, 12 Jul 2013.
20. spangirl et al. The unregulated work of Mechanical Turk. *Turker Nation*, 18 Mar 2013. turkernation.com/showthread.php?17717-The-Unregulated-Work-of-Mechanical-Turk.
21. chrism et al. CMU professor: online 'crowd workers' need protections, organization. *Turker Nation*, 11 Mar 2013. turkernation.com/showthread.php?17591-CMU-Professor-Online-Crowd-Workers-Need-Protections-Organization.
22. Fehr, E., G. Kirchsteiger, A. Riedl. Does fairness prevent market clearing? An experimental investigation. *Quarterly Journal of Economics* 108: 437, 1993.
23. Stiglitz, J. *Globalization and its Discontents*. Norton, 2003.
24. Stiglitz, J. *Making Globalization Work*. Norton, 2006.
25. Bhagwati, J. *In Defense of Globalization*. Oxford, 2004.
26. Rajan, R. *Fault Lines: How Hidden Fractures Still Threaten the World Economy*. Princeton, 2010.
27. Khanna, S., A. Ratan, J. Davis, W. Thies. Evaluating and improving the usability of Mechanical Turk for low-income workers in India. *DEV '10*: 12.
28. Narula, P., P. Gutheim, D. Rolnitzky, A. Kulkarni, B. Hartmann. MobileWorks: a mobile crowdsourcing platform for workers at the bottom of the pyramid. *HCOMP '11*.
29. Gawade, M., R. Vaish, M. Nduta, J. Davis. Exploring microwork opportunities through cybercafés. *DEV '12*: 21.
30. Gupta, A., W. Thies, E. Cutrell, R. Balakrishnan. mClerk: enabling mobile crowdsourcing in developing regions. *CHI '12*: 1843.
31. Samdaria, N., A. Mathur, R. Balakrishnan. Paying in kind for crowdsourced work in developing regions. *Pervasive '12*: 343.
32. Kochhar, S., S. Mazzocchi, P. Paritosh. The anatomy of a large-scale human computation engine. *HCOMP '10*: 10.
33. Kulkarni, A., P. Gutheim, P. Narula, D. Rolnitzky, T. Parikh, B. Hartmann. MobileWorks: designing for quality in a managed crowdsourcing architecture. *IEEE Internet Computing* 16(5): 28, 2012.
34. Hyman, P. Software aims to ensure fairness in crowdsourcing projects. *Comm. ACM* 56(8): 19, 2013.
35. spangirl. Spangirl speaks! *Broken Turk*, Nov 2010. brokenturk.blogspot.com/2010/11/spangirl-speaks-interview-part-1-of-3.html.
36. Irani, L., ed. *Turker bill of rights—2013 edition*. turkwork2013.differenceengines.com.
37. Silberman, M. S., ed. FAQ. *turkopticon-devel*, Aug 2013. turkopticon-devel.differenceengines.com/faq.html.
38. Janssen, M. A., E. Ostrom, eds. Special issue: Empirically-based, agent-based modeling of social-ecological systems. *Ecology and Society* 25, 2006.
39. Schelling, T. *Micromotives and Macrobehavior*. Norton, 1978.
40. Janssen, M. A., ed. *Complexity and Ecosystem Management: The Theory and Practice of Multi-Agent Systems*. Edward Elgar, 2002.
41. Bonabeau, E. Agent-based modeling: methods and techniques for simulating human systems. *Proc. Natl. Acad. Sci.* 99(S3): 7280, 2002.
42. Tesfatsion, L., K. L. Judd, eds. *Handbook of Computational Economics, Vol. 2: Agent-Based Computational Economics*. Elsevier, 2006.
43. Epstein, J. M. *Generative Social Science: Studies in Agent-Based Computational Modeling*. Princeton, 2007.
44. Farmer, J. D., D. Foley. The economy needs agent-based modelling. *Nature* 460: 685, 2009.
45. Fehr, E., K. M. Schmidt. A theory of fairness, competition, and cooperation. *Quarterly J. Econ.* 114: 817, 1999.
46. Fehr, E., K. M. Schmidt. The economics of fairness, reciprocity and altruism—experimental evidence and new theories. In S.-C. Kolm and J. M. Ythier, eds., *Handbook of the Economics of Giving, Altruism and Reciprocity, Vol. 1*. Elsevier, 2006.
47. Ross, J., L. Irani, M. S. Silberman, A. Zaldivar, B. Tomlinson. Who are the crowdworkers? Shifting demographics in Mechanical Turk. *CHI '10 EA*: 2863.
48. Akerlof, G. A. The market for "lemons": quality uncertainty and the market mechanism. *Quarterly Journal of Economics* 84(3): 488, 1970.
49. Roth, A. E., E. Peranson. The redesign of the matching market for American physicians: some engineering aspects of economic design. *Am. Econ. Rev.* 89(4): 748, 1999.
50. Roth, A. E. What have we learned from market design? *Economic Journal* 118: 285, 2008.
51. Ünver, M. U. Dynamic kidney exchange. *Review of Economic Studies* 77(1): 372, 2010.
52. Ostrom, E. *Governing the Commons*. Cambridge, 1990.
53. Ostrom, E. *Understanding Institutional Diversity*. Princeton, 2005.
54. Ostrom, E. Beyond markets and states: polycentric governance of complex economic systems. *Am. Econ. Rev.* 100: 1, 2010.