UNIVERSITY OF CALIFORNIA,
IRVINE

Human-Centered Computing and the Future of Work:
Lessons from Mechanical Turk and Turkopticon, 2008–2015

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Information and Computer Sciences

by

M. Six Silberman

Dissertation Committee:
Professor Bill Tomlinson, Co-Chair
Professor Bonnie A. Nardi, Co-Chair
Associate Adjunct Professor Judith Gregory

2015

Dedication

For Beatriz da Costa, 1974–2012

At this point, my dream begins to fade and melt away, like water in water. The vast library surrounding me is on Calle Mexico, not Rodríguez Peña, and you, Lugones, died in early '38. My vanity and my nostalgia have confected a scene that is impossible. Maybe so, I tell myself, but tomorrow I too will be dead and our times will run together and chronology will melt into an orb of symbols, and somehow it will be true to say that I have brought you this book and that you have accepted it.

—J. L. Borges, Foreword to *El hacedor*, 1960

Table of Contents

List of Figures

List of Tables

Acknowledgments

Curriculum Vitae

M. Six Silberman

2007	B.S., Applied Mathematics Columbia University in the City of New York
2009	M.S., Engineering, Concentration in Arts Computation Engineering University of California, Irvine
2015	Ph.D., Information and Computer Sciences University of California, Irvine

Fields of Study

Online labor markets; sustainability and human-computer interaction.

Publications

Journal Articles

Tomlinson, B., E. Blevis, B. Nardi, D. Patterson, M. S. Silberman, Y. Pan. Collapse informatics and practice: theory, method, and design. *ACM Transactions on Computer-Human Interaction* 20(4): 24, 2013.

Tomlinson, B., M. S. Silberman. The cognitive surplus is made of fossil fuels. *First Monday* 17(11), 31 Oct 2012.

Conference Papers

Penzenstadler, B., et al. (29 authors). ICT4S 2029: What will be the systems supporting sustainability in 15 years? *Proc. ICT4S 2014*: 30-39. Best Paper Nominee.

Baumer, E. P. S., et al. (28 authors). CHI 2039: Speculative research visions. *Proc. CHI EA 2014*: 761-770 (alt.chi).

Silberman, M. S., E. Blevis, E. Huang, B. A. Nardi, L. P. Nathan, D. Busse, C. Preist, S. Mann. What have we learned? A SIGCHI HCI & Sustainability community workshop. *Proc. CHI EA 2014*: 143-146.

Irani, L. C., M. S. Silberman. Turkopticon: interrupting worker invisibility in Amazon Mechanical Turk. *Proc. CHI 2013*: 611-620. "Best Paper Award" (best 1% of accepted papers). Translated and collected as Turkopticon: Ein Tool, um Arbeiter auf Mechanical Turk sichtbar zu machen, C. Benner, ed. *Crowdwork— zurück in die Zukunft? Perspektiven digitaler Arbeit* [Crowd Work: Back to the Future? Perspectives on Digital Work]: 131. Bund Verlag, 2014. Introduction at tinyurl.com/ol73dvn. Included in syllabi: C. Callison-Burch, Crowdsourcing & Human Computation, Univ. Penn. Comp. & Inf. Sci., Fall 2014; R. Mason, Projects in Social Media, Spring 2013, Univ. Wash. iSchool; P. Sengers, Designing Technologies for Social Impact, Fall 2013, Cornell Information Studies/Science and Technology Studies.

Tomlinson, B., M. S. Silberman, E. Blevis, Y. Pan, D. Patterson. Collapse informatics: augmenting the sustainability & ICT4D discourse in HCI. *Proc. CHI 2012*: 655-664. Best sustainability paper award. "Best Paper Honorable Mention" (best 5% of accepted papers). Included in syllabus: B. Nardi, Collapse Computing, Spring 2012, UC Irvine Informatics.

Tomlinson, B., et al. (31 authors). Massively distributed authorship of academic papers. *Proc. CHI EA 2012*: 11-20 (alt.chi).

Baumer, E. P. S., M. S. Silberman. When the implication is not to design (technology). *Proc. CHI 2011*: 2271-2274. "Best Paper Honorable Mention" (best 5% of accepted papers). Included in syllabus: D. Patterson, User Interaction Software, Winter 2011, UC Irvine Informatics.

Silberman, M. S., J. Ross, L. C. Irani, B. Tomlinson. Sellers' problems in human computation markets. *Proc. HCOMP 2011*: 18-21. Included in syllabus: B. Hartmann, M. Franklin, T. Parikh, Crowdsourcing Seminar, Spring 2011, UC Berkeley EECS.

Silberman, M. S., B. Tomlinson. Toward an ecological sensibility: tools for evaluating sustainable HCI. *Proc. CHI EA 2010*: 3469-3474 (works-in-progress).

Ross, J., L. C. Irani, M. S. Silberman, A. Zaldivar, B. Tomlinson. Who are the crowdworkers? Shifting demographics in Mechanical Turk. *Proc. CHI EA 2010*: 2863-2872 (alt.chi). Included in syllabi: D. Ganesan, Crowdsourcing: Opportunities and Challenges, Fall 2010, U. Mass. Amherst CS; B. Hartmann, M. Franklin, T. Parikh, Crowdsourcing Seminar, Spring 2011, UC Berkeley EECS; J. Vertesi, Sociology of Technology, Fall 2010, Princeton Sociology; F. Turner, Digital Media in Society, Fall 2012, Stanford Communication; K. Starbird,

Crowds and Crowdsourcing, Fall 2012, Univ. Wash. Human Centered Design & Engrg.; J. Gregory, Social Analysis of Computing, Fall 2013, UC Irvine Informatics.

Huh, J., L. P. Nathan, M. S. Silberman, E. Blevis, B. Tomlinson, P. Sengers, D. Busse. Examining appropriation, re-use, and maintenance for sustainability. *Proc. CHI EA 2010*: 4457-4460 (workshop).

Magazine Articles

Silberman, M. S., L. Nathan, B. Knowles, R. Bendor, A. Clear, M. Håkansson, T. Dillahunt, J. Mankoff. Next steps for sustainable HCI. *Interactions* 21(5): 66-69, 2014.

Irani, L., M. S. Silberman. From critical design to critical infrastructure: lessons from Turkopticon. *Interactions* 21(4): 32-35, 2014.

S. L. Irani-Silberman [M. S. Silberman and L. Irani]. Interrupting invisibilities and bridging worlds: an essay on the work of turkopticon.differenceengines.com. Special Mention, 2013 Hannah Arendt Prize in Critical Theory and Creative Research. Judges: Anne-Marie Oliver, Barry Sanders, Claire Bishop, Judith Butler, Barbara Duden, Julia Kristeva, Heike Kühn, Martha Rosler. *Untitled Magazine* (Pacific Northwest College of Art), 26 Sep 2013.

Tomlinson, B., D. Patterson, Y. Pan, E. Blevis, B. Nardi, M. S. Silberman, J. Norton, J. LaViola. What if sustainability doesn't work out? *Interactions* 19(6): 50-55, 2012.

Tomlinson, B., M. S. Silberman, J. White. Can more efficient IT be worse for the environment? *IEEE Computer* 44(1): 87-89, 2011.

Silberman, M. S., L. C. Irani, J. Ross. Ethics and tactics of professional crowd work. *XRDS* 17(2): 39-43, 2010. Included in syllabi: M. Lease, Crowdsourcing: Theory, Methods, and Applications, Spring 2011, UT Austin Inf. School; K. Starbird, Crowds and Crowdsourcing, Fall 2012, Univ. Wash. Human Centered Design & Engrg.; J. Gregory, Social Analysis of Computing, Fall 2013, UC Irvine Informatics.

Peer-reviewed Workshop Papers

Gilbert, A. M., M. S. Silberman. Re-purposed activist as new sustainable HCI researcher. *What Have We Learned?*, workshop at CHI 2014.

Silberman, M. S. Norms, infrastructures, goals, and needs. *Simple, Sustainable Living*, workshop at CHI 2012.

Kaufman, S. J., M. S. Silberman. Rebound effects in sustainable HCI. *Everyday Practice and Sustainable HCI*, workshop at CHI 2011.

Silberman, M. S., B. Tomlinson. Precarious infrastructure and postapocalyptic computing. *Examining Appropriation, Re-Use, and Maintenance for Sustainability*, workshop at CHI 2010. Included in syllabus: B. Nardi, Collapse Computing, Spring 2012, UC Irvine Informatics.

Silberman, M. S. Designing economic interactions. *Defining the Role of HCI in the Challenges of Sustainability*, workshop at CHI 2009.

Technical Reports

Choontanom, T., M. S. Silberman. Edge effects: complex practices between open source and proprietary software development. UC Irvine Lab. for Ubiquitous Computing and Interaction Tech. Rep. LUCI-2010-001, 2010.

Abstract of the Dissertation

Human-Centered Computing and the Future of Work:
Lessons from Mechanical Turk and Turkopticon, 2008–2015

By

M. Six Silberman

Doctor of Philosophy in Information and Computer Sciences

University of California, Irvine, 2015

Professor Bill Tomlinson, Co-chair

Professor Bonnie A. Nardi, Co-chair

Online labor markets such as Amazon Mechanical Turk (AMT), Uber, and TaskRabbit are contributing to rapid changes in the nature of work for hundreds of thousands of workers. These markets create significant new economic opportunities, but most currently treat workers as second-class citizens. Take-home pay is often low compared to similar work in traditional employment arrangements, and workers have limited means of influencing market design or management practice. This makes it hard for workers to create reliable livelihoods from the opportunities these markets present. This dissertation uses AMT, an online labor market for small information tasks, as a case through which to examine the consequences of treating workers as second-class citizens, to argue that future platform designs and management

practices should treat workers as central stakeholders, and to develop theory and method for doing so.

The central argument of the dissertation is that workers' concerns should be more substantively and systematically addressed in the design and operation of online labor markets. Five messages elaborate this argument. First, in online labor markets, some workers are casual or transient, while others are professionals, providing significant and reliable value to customers and relying on income earned in the market to meet basic needs. Second, workers who rely on income earned through online labor markets should be considered first-class stakeholders, alongside customers and shareholders. Third, workers in online labor markets are rarely the narrowly self-interested profit maximizers of classical economic theory. Workers can be better understood as "situatedly rational" actors: human beings with incomplete information and finite cognitive capabilities whose actions and preferences are shaped by many factors, including rules, norms, and expectations. Fourth, online labor markets are not monolithic, perfectly competitive markets but parts of polycentric economic systems composed of complexly interlinked action situations characterized by imperfect competition and incomplete information. Fifth, institutions supporting crowd work research should develop an interdisciplinary practice-oriented agenda to understand the consequences of current online labor market designs and practices, and to develop new designs and practices that incorporate workers who rely on market income as central stakeholders.

Prologue

Hi Rochelle!

Thanks for contacting us.

Due to the size of our department, Mechanical Turk doesn't have the functionality for customer service ticketing system and that's why we currently only able to assist through email.

Just as my colleague mentioned, we have passed your feedback along to the appropriate department for consideration.

I've checked your account and saw that you are a Master worker, we have a forum (I've linked you to it below) where our Master Worker's can interact.

Alternatively, you can sign up for Turk Opticon which is not Amazon Affiliated but Mechanical Turk Workers tend to communicate and share best practices on there.

Master Forum : https://forums.aws.amazon.com/forum.jspa?forumID=122

Turk Opticon : https://turkopticon.ucsd.edu/

Please let us know if you have anymore questions.

Thank you for your inquiry. Did I solve your problem?

If yes, please click here:
http://www.amazon.com/gp/help/survey?p=A14M8ZYB2WJT0N&k=hy

If no, please click here:
http://www.amazon.com/gp/help/survey?p=A14M8ZYB2WJT0N&k=hn

Best regards,

Nicole B.
Amazon Mechanical Turk

Please note: this e-mail was sent from an address that cannot accept incoming e-mail. To contact us again, select the Contact Us link related to your inquiry below.

Workers: https://www.mturk.com/mturk/contactus

Requesters: https://requester.mturk.com/contactus

Chapter 1

Introduction

1.1 Context

Online labor markets such as Amazon Mechanical Turk (AMT), Uber, and TaskRabbit are contributing to rapid changes in the nature of work for hundreds of thousands of workers. These markets stand to create significant new economic opportunities, but current market designs and management practices typically treat workers as second-class citizens.[1] Take-home pay is typically low compared to similar work in traditional employment arrangements. Workers have limited means of influencing market design or management practice, and therefore little control over their own work arrangements. Low pay and limited control over work arrangements make it hard for workers to create reliable, sustainable livelihoods from the many uncoordinated work opportunities online labor markets present.[2] Reliable, sustainable livelihoods matter for three reasons. First, workers value them—even as they value the unique flexibility online labor markets offer. Second, if workers cannot build reliable, sustainable livelihoods, the labor pool reached by online labor markets will be limited to casual or temporary workers—limiting the markets' own long-term sustainability. Third, reliable, sustainable

---

[1] See for example Manjoo 2015, Smith 2015, and Reich 2015.

[2] For peer-reviewed computing research discussing these issues in crowd work markets specifically, see for example Silberman et al. 2010, "Sellers' problems in human computation markets"; Silberman et al. 2010, "Ethics and tactics of professional crowdwork"; Bederson and Quinn 2011; and Kittur et al. 2013.

livelihoods are crucial to socioeconomic mobility, a central concern of economic policy.

This dissertation uses Amazon Mechanical Turk (AMT)—a "crowd work market" in which small information tasks are made publicly available online for paid completion by self-selecting, often anonymous, individuals—as a case through which to examine the consequences of treating workers as second-class citizens, to argue for future platform designs and management practices that treat workers as central stakeholders, and to develop theory and method for doing so.

In AMT, challenges facing workers broadly are noticed most acutely by the relatively small fraction of workers who do most of the work. The main motivation for these workers is money: they rely on the income they earn from "Turking" to meet basic needs.[3] But pay for most work posted to AMT is low compared to similar work organized through traditional employment relations. This is at least partly because AMT employers, called "requesters," hold a unique power over workers: they can decline to pay for work for any or no reason, even if they keep and use it, and workers have no formal recourse against requesters they suspect have stolen their work. AMT's requester-facing application programming interface (API) lets requesters post tasks and make payment decisions algorithmically, delegating the work of recruiting and managing workers to software. This "algorithmic management" allows requesters to "scale" rapidly, recruiting thousands of workers within hours and

---

[3] See for example Ipeirotis 2010, Ross et al. 2010, and Martin et al. 2014.

managing them as if they were interchangeable, always-on parts of a computing system.[4] But this primes many requesters to expect quick, easy, and "frictionless" interactions. This expectation, and the high worker-to-requester ratio, leads many requesters to fail to respond adequately, or at all, to worker communications. This often leads to poor work and strained worker-requester relations. And because of the scale of AMT itself—it supports at least tens of thousands of dollars of transactions daily, but is run by less than a dozen staff—platform operators cannot mediate worker-requester disputes or respond adequately to worker or requester concerns.[5]

AMT's disempowerment of workers and inattention to workers' concerns in favor of attracting customers is typical of recently developed online labor markets broadly.[6] Existing laws, devised for different ways of organizing work, do not offer these workers ready tools in the struggle to improve working conditions, build and maintain reliable, sustainable livelihoods, or resist platform operators' and employers' persistent, strategic, and often well-resourced efforts to cut costs.[7] These and other challenges facing workers will need to be addressed if the full potential of online labor markets to create substantive and sustainable new economic opportunities is to be realized. But thus far they have not been taken seriously in computing research or industry

---

[4] See Irani and Silberman 2013 (pp. 612-614) and Irani 2013 (pp. 8-9).
[5] Indeed Section 3(f) of the AMT Participation Agreement (Amazon Mechanical Turk 2014), to which workers and requesters must agree before posting or completing work, explicitly makes clear than Amazon disclaims responsibility for "resolving any disputes between participants related to or arising out of" any transaction on the site.
[6] See for example Bercovici 2014 (about Uber) and Said 2014 (about TaskRabbit).
[7] As entry points into the legal scholarship on crowd work, see for example Cherry 2009 and Felstiner 2011.

practice. Research in human-centered computing[8] (HCC) and human computation[9]—the main fields contributing to the development of crowd work and crowdsourcing—has focused on the end user's (i.e., the requester's) experience of crowdsourcing rather than the worker's. In the context of paid crowd work, work in both fields often focuses on the three traditional engineering objectives: cost, speed, and quality.[10] Research on crowd work in HCC and human computation that does consider workers tends to focus on understanding workers' motivations for participating—often for the purpose of designing more effective incentives for eliciting more good, cheap, and quick work.[11]

A focus on reducing cost, increasing speed, and improving quality is not problematic in and of itself. But it can become problematic in the context of crowd work because of the combination of four factors. First, in crowd work, ongoing human participation is integral to the execution of an algorithmic process. Second, some participants rely on their income from crowd work to

---

[8] I use the term "human-centered computing" to refer to research in the fields of human-computer interaction (HCI) and computer supported cooperative work (CSCW) and published in the major venues of these fields (viz., the CHI and CSCW conferences and their affiliated journals).

[9] I use the term "human computation" to refer to research in computer science "proper" (i.e., not HCI or CSCW) that incorporates *ongoing* human input into an algorithmic process. This includes work under the labels of "crowdsourcing" and "human computation" but not work in other approaches to artificial intelligence or machine learning, even if such approaches require human input in initial phases of system development (e.g., for training machine learning systems).

[10] Though no longer new, the paper whose title most concisely reflects this focus is Snow et al. 2008 ("Cheap and fast—but is it good?"). For further examples, see Downs et al. 2010; Ipeirotis et al. 2010; Kochhar et al. 2010; Grier 2011; Oleson et al. 2011; Dow et al. 2012; Jung and Lease 2012; Mao et al. 2012, "Better computation through principled voting"; and Rao et al. 2013.

[11] See for example Jain and Parkes 2009; Chandler and Horton 2011; Faradani et al. 2011; Singer and Mittal 2011; Ho et al. 2012; Mao et al. 2013, "Volunteering versus work for pay"; Singla and Krause 2013; and Witkowski et al. 2013.

meet basic needs, creating a situation with some elements of a traditional employment relationship. Third, in part because of the heavily computationally mediated nature of the work, other elements of traditional employment relations are absent, including legal recourse against wage theft. Fourth, the high ratio of workers to employers and platform operators makes it hard for the latter to give workers meaningful feedback on work or respond substantively to workers' concerns.[12] While some workers, given the choice, would certainly prefer the flexibility of current crowd work arrangements—even given their risks—to more heavily regulated arrangements even if the latter offered more security or voice in the design of work, this preference is not universal. And the challenge for designers is to develop systems and organizations that offer flexibility *and* reasonable security and voice. Problems with the research focus on the traditional engineering objectives arise in the context of crowd work because of researchers' focus on traditional objectives to the relative exclusion of these apparently nontechnical considerations.

In recent years, however, perhaps partially in light of press coverage,[13] worker activism,[14] and lawsuits,[15] HCC and human computation researchers have begun directing some attention to the question of worker experience in crowd work and other online labor markets, and the relations between workers,

---

[12] See Sec. 2.2.5; Irani and Silberman 2013, p. 614; and Irani 2013, pp. 5-9.

[13] See for example Zittrain 2009; Cushing 2012; Uddin 2012; Brode 2013; Dobson 2013; Hodson 2013, "Crowdsourcing grows up as online workers unite" and "Time to focus on the welfare of online workers"; Folbre 2013; Brandom 2013; Leonard 2013; Chace and Kenney 2015; and Marder 2015.

[14] See for example Harris 2014.

[15] The main open case in crowd work is *Otey et al. v. CrowdFlower, Inc., et al.*, 4-CV-05524-JST. In the on-demand economy broadly, Uber, Lyft, and Handy are defendants in separate suits alleging employee misclassification; see Kessler 2015.

requesters, and platform operators in these markets. In 2010, an early demographic study of Turkers by Joel Ross, Lilly Irani, myself, Andrew Zaldivar, and Bill Tomlinson asked respondents, among other questions, to indicate the extent to which they relied on AMT income to meet basic needs.[16] Later the same year, in "Sellers' problems in human computation markets,"[17] on which Ross, Irani, Tomlinson, and I were co-authors, I summarized the findings presented in the previous publication, in a previous working paper by Ross and Tomlinson,[18] and in earlier non-peer-reviewed but widely-circulated and -cited publications by Panagiotis Ipeirotis[19] as follows (pp. 18-19, orig. emph.):

> **The AMT labor pool hosts a growing international population earning less than USD 10,000/yr., some of whom are reliant on their Turking income to make basic ends meet. The uncertainty associated with HIT payment complicates human computers' [i.e., workers'] work and reduces their effective wage. This uncertainty is due in part to the apparent prevalence of fraudulent requesters, to whom certain design decisions have made AMT particularly attractive.**
>
> Intuitively, we might expect that, just as buyers of human computation [i.e., requesters] aim to minimize expense at a fixed quality (or maximize quality within a cost constraint), sellers of human computation [i.e., workers] wish to secure payment with a minimum time expenditure, even if this means "gaming the system" by providing responses they know are of low quality. Fraudulent sellers do appear to optimize in this way, but a reading of survey responses and

---

[16] Ross et al. 2010.
[17] Silberman et al. 2010.
[18] Ross and Tomlinson 2009.
[19] Ipeirotis 2008; 2009; and 2010, "The new demographics of Mechanical Turk."

forum discussions reveals a concern for what is "fair" and "reasonable" rather than a desire to maximize short-term personal earnings at requester expense.

Ross et al., extending work by Ipeirotis, present longitudinal demographic data on the "increasingly international" AMT. They find that AMT hosts a growing population of young, male, Indian workers earning less than USD 10,000/yr. Additionally, almost a third of Indian Turkers surveyed reported that they always or sometimes relied on their Turking income to "make basic ends meet" (27% in May 2009 and 31% in Feb. 2010). Between May 2009 and Feb. 2010, the fraction of US Turkers surveyed reporting reliance held steady at 13±1%.

Many Turkers see themselves as laborers doing work to earn money. In survey data collected in Feb. 2009 (n = 878), the most commonly reported motivation for doing HITs was making money. 91% of respondents mentioned a desire to make money, whether as a form of supplemental income or in order to purchase extras. Turking to pass the time, in contrast, was mentioned by only 42% of respondents. Feb. 2010 data (n = 1000) from Ipeirotis (2010, "The new demographics of Mechanical Turk") confirms the importance of money compared to other motivations, with most respondents reporting they do not do HITs for fun or to kill time, and ~25% of Indian respondents and ~13% of US respondents reporting that AMT is their primary source of income.

This publication was to my knowledge the first peer-reviewed presentation of workers' issues within human computation, which in 2010 was just beginning to develop formal institutions (namely, the Human Computation Workshop, then collocated with the ACM Conference on Knowledge Discovery and Data Mining ["KDD"], now its own conference, the AAAI Conference on Human Computation and Crowdsourcing ["HCOMP"]). After reviewing then-available demographic research, a selection of worker discourse (from Turker Nation),

and responses to a "Turkers' Bill of Rights" survey posted by Irani (see Chapter 2), we offered the following (by now familiar) list of challenges facing workers (p. 20):

> (1) uncertainty about payment; (2) unaccountable and seemingly arbitrary rejections [i.e., non-payment]; (3) fraudulent tasks; (4) prohibitive time limits; (5) long pay delays; (6) uncommunicative requesters and administrators; (7) cost of requester and administrator errors borne by workers; and (8) low pay.

The 2008-2009 survey results yielded an average hourly wage of just under $2.00/hr., but this figure was vigorously criticized by Turkers, especially on Turker Nation, after the paper was published. We did not understand the reason for this discrepancy at the time, but subsequent discussion and reflection has surfaced at least three interacting explanations. First, not all Turkers do surveys, so tasks that are obviously surveys suffer from selection bias. Second, price affects task selection, introducing a second source of selection bias: tasks with low pay disproportionately attract workers who regularly complete low-paying tasks. Third, a small fraction of Turkers do most of the work on AMT. Notably, these workers are the ones who are most likely to rely on Turking income to meet basic needs and therefore to avoid low-paying tasks. Therefore a relatively low-paying survey asking for Turkers' average wage may, even if respondents self-report accurately, elicit responses only from Turkers who accept relatively low-paying tasks and who do a disproportionately *small* fraction of tasks in the market. This may or may not systematically bias the estimate of the average wage over all *workers*, but it will drastically underestimate the average wage over all *tasks*—as most tasks are done by

professionals, who have the motivation, skills, and information channels to earn a higher wage. The role of the small fraction of professional workers who do most of the tasks—and the question of employers' and platform operators' ethical and legal responsibilities to them—are central topics of this dissertation.

In "Sellers' problems in human computation markets," we also listed relevant open research and design questions, including (p. 20):

> How does database, interface, and interaction design influence 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?

We noted further that "gaps remain[ed] in our demographic understanding of AMT" and that "as new platforms and tools come online and mature, comparative studies [would] become possible, and longitudinal studies more feasible."

I know of no explicitly longitudinal studies of crowd work platforms, but at least one comparative study was circulated in 2013,[20] raising awareness of alternatives to AMT among researchers using crowd work platforms. A sensitive and insightful ethnomethodological study of Turker Nation was published in 2014,[21] shedding light on workers' motivations and broader economic conditions—and raising awareness among human-centered computing researchers that workers are in fact humans. Perhaps surprisingly, some researchers have confessed to having never considered this point, or that

---

[20] Vakharia and Lease 2013; peer-reviewed and published as Vakharia and Lease 2015.
[21] Martin et al. 2014.

workers ought to be paid a decent wage.[22] The demographics of AMT continue to change as Amazon changes the rules in response to market and regulatory forces,[23] and even recent demographic studies are likely to be slightly out of date. But the growing interest among social science researchers in crowd work platforms generally and AMT specifically has created a sustained discussion on worker demographics.[24]

The question of what's fair in paid crowdsourcing has been raised with serious intent in the popular press (see above, note 12) and in academic research.[25] But it has not yet been discussed explicitly at much length or with much rigor—empirically or theoretically. But three ongoing developments in the field are especially relevant to this question. The first is the compilation of the "Guidelines for Academic Requesters," an extremely detailed document addressing many questions and practices relevant to researchers collecting data through AMT.[26] The second is the class action lawsuit *Otey et al. vs. CrowdFlower, Inc., et al.*, in which plaintiff Christopher Otey sued intermediary CrowdFlower—once the most prolific requester on AMT—and its cofounders for violation of the Fair Labor Standards Act (i.e., the minimum wage law; see above, note 14). And the third is a new effort, led by researchers at Stanford

---

[22] David Martin, pers. comm., 2014.
[23] See e.g. Ipeirotis 2012, "Mechanical Turk changing the defaults."
[24] See e.g. Paolacci et al. 2010, Buhrmeister et al. 2011, Berinsky et al. 2012, Mason and Suri 2011, and Paolacci and Chandler 2014 (and many more cited therein).
[25] Silberman et al. 2010, "Ethics and tactics of professional crowdwork"; Bederson and Quinn 2011; Kittur et al. 2013; Irani and Silberman 2013; Martin et al. 2014.
[26] `guidelines.wearedynamo.org`; see also Salehi et al. 2015.

University and the University of California, Santa Cruz, to build a new crowd work market that remedies the shortcomings of AMT.[27]

Taken together, these three developments signal significant interlinked changes in the social, economic, legal, and technical conditions of crowd work. These changes are themselves taking place within broader changes brought about by the emergence of "on-demand economy" platforms such as Uber and TaskRabbit—platforms that have been described, with both excitement and dismay, as portending sweeping changes to the nature of work (see above, note 1). While HCC and human computation researchers are well-positioned to make significant contributions to these developments, it is not yet clear whether those contributions, seen in hindsight, will be regarded as broadly beneficial. Computationally mediated labor markets may lead to increasing convenience and reduced costs for a minority of privileged users and increasingly precarious livelihoods for the majority of workers, or they may lead to economic empowerment and broad-based growth in human capital. The principles animating the design and operation of these markets will be major factors in shaping this outcome, and in shaping the future of work.

In this context, this dissertation uses AMT as a case through which to ask two questions:

- What are the consequences of designing a market that treats workers' concerns as secondary to other objectives such as work quality, speed, low cost, and requester convenience?

---

[27] Bernstein 2015, "Join Stanford researchers in the largest crowdsourcing research project ever."

- What might crowd work markets—and computationally mediated labor markets broadly—look like if they were designed with worker well-being and livelihood sustainability as first-class goals, equal in importance to work quality, speed, low cost, and requester convenience?

1.2   Main messages of the dissertation

The central argument of this dissertation is that workers' concerns should, and can, be more substantively and systematically addressed in the design and operation of online labor markets.

Taking workers' concerns seriously will better allow them to create reliable, sustainable livelihoods from the work opportunities these new markets make available. It is therefore a crucial step in ensuring the markets' own sustainability, and will contribute to socioeconomic mobility in the economy at large.

Five messages elaborate this central argument. First, in online labor markets, some workers are casual or transient, while others are professionals, providing significant and reliable value to customers on an ongoing basis and relying on income earned in the market to meet basic needs. In AMT, the relatively small fraction of "Turkers" who rely on Turking income to meet basic needs do most of the work posted to the market.[28] Most of these professional Turkers are well-educated[29] and live in the United States.[30] And many report that

---

[28] Fort et al. 2011.
[29] See e.g. Paolacci and Chandler 2014.

they participate in AMT not by choice but because they are unable to secure other employment.[31] Thus some of the narratives offered by researchers and employers to justify low pay—e.g., that most workers who rely on Turking income live in "developing" countries with low costs of living; that most "developed"-country workers work mainly to pass time; and that workers freely choose to participate in AMT and can choose other work if they find the pay too low—are inaccurate.

Second, workers who rely on income earned through online labor markets to meet basic needs should be considered first-class stakeholders, alongside customers and shareholders. These workers are strongly invested in the sustainability of the market. When the market is designed appropriately, these workers can be relied on to adhere to, and even enforce, market norms that benefit all participants. Their concerns and input regarding the design and operation of the market should be taken seriously. Formal processes for eliciting their input should be developed and integrated into market design and management practice.

Third, workers in online labor markets are not usually the narrowly self-interested profit maximizers of classical economic theory. Professionals—those workers who rely on income earned through participation in the market—especially want market transactions to produce good outcomes for everyone, and want the market to be sustainable. They take professional pride in doing

---

[30] Ipeirotis (2015) reports that by April 2015, between 75% and 80% of survey respondents reported living in the US, significantly up from 2009-2010.
[31] See e.g. Martin et al. 2014.

good work and helping other market participants. They adhere to norms they think will produce good outcomes for everyone, and spend unpaid time discussing what those norms should be. Thus rather than conceiving of workers as narrowly self-interest profit maximizers, researchers and platform operators should see workers as "situatedly rational" actors. Situated rationality augments the notion of "bounded rationality"—i.e., that actors have incomplete information and limited cognitive capacities—with the observation that actors' actions, and even their preferences, are shaped by a diversity of factors typically omitted in classical economic analysis, including rules, norms, expectations, and actors' understandings of market dynamics. Certain market designs may induce situatedly rational actors to act as narrowly self-interested profit maximizers, but this result is not inevitable—on the contrary, it usually produces suboptimal outcomes and may indicate poor market design.

Fourth, online labor markets are not monolithic, perfectly competitive markets. Nor is there a perfectly competitive "market of markets." Rather, each market is part of a polycentric system composed of complexly interlinked action situations. This system is characterized by imperfect competition and incomplete information. The notion of polycentricity indicates that there are multiple decision making locations within the system that, while formally independent, are interlinked by the consequences of decisions taken at each location. Each decision taken by an actor can be seen as occurring within an "action situation." Action situations can be described by the characteristics of the actors involved, their roles, the rules governing their actions, the

information available to them, the possible outcomes of their actions, the relations between actions and outcomes, and actors' valuations of outcomes. In a given action situation, an actor may act according to a variety of logics—e.g., financial, institutional, or moral. These logics may not be commensurable with one another, and may not be formalizable.

Fifth and finally, institutions funding human-centered computing (HCC) research should support an interdisciplinary practice-oriented research agenda to understand the consequences of current online labor market designs and management practices, and to develop new designs and practices that incorporate workers who rely on market income as central stakeholders. This agenda should integrate software practice, empirical research, theory development, and value-rational analysis. Software practice and empirical research are familiar in HCC. Current HCC theories must be expanded to larger scales of analysis and design. And the three aforementioned well-established research modes should be linked to value-rational analysis—the rigorous and broad-based consideration of questions such as *Where are we going in computationally mediated work? Who gains and who loses? Is this desirable? What should be done?* The influence of designers' and operators' understandings of such "nontechnical" issues in system design and use has long been acknowledged in HCC research. But online labor markets so tightly interweave the technical and the ostensibly nontechnical that questions once considered nontechnical can no longer be "outsourced" to social scientists or regulators. The computational mediation of work calls for computing

researchers to take a more active role in the collective process of understanding the social consequences of technology design, articulating possible futures, distinguishing between what is desirable and what is merely possible, making plain the distribution of benefits and risks, and taking concrete steps to create the institutional conditions required to develop systems and practices that benefit a broad variety of stakeholders.

Industry collaboration will be crucial for the long-term sustainability of such an effort. But such collaboration will be complicated by the distinct institutional accountabilities and cultures of research and business. Indeed business broadly is grappling with parallel challenges, as can be seen in the relatively new discourses on corporate social responsibility and social entrepreneurship, the emergence of new structures for corporate governance such as the B Corporation, and the development of new computationally-mediated strategies for raising capital such as crowdfunding. The greatest potential for creating broad-based social value in online labor markets lies at the intersection of expanded HCC theory and method and new organizational models that aim to create sustainable value for a broader stakeholder base than traditional models.

1.3 The main messages elaborated

1.3.1 Some workers are casual or transient; others are professionals who rely on income earned through market participation to meet basic needs

Like most online participation, online labor market participation follows a "power law" distribution. In AMT specifically, most work is done by a relatively small fraction of workers, who do many times more tasks than the "average" worker.[32] These prolific workers are effectively professional crowd workers. Most professional "Turkers" are highly skilled, live in the United States, and rely on Turking income to meet basic needs. While these workers may enjoy various nonmonetary benefits of crowd work,[33] their primary motivation for Turking is to earn money. These workers spend dozens of hours a week doing crowd work tasks. They share information with one another on forums and through other communication channels and build, maintain, and use specialized software tools.

As a result of their long experience, participation in worker communities, and motivation to earn money, these workers earn significantly more than the often-cited "average" AMT wages of USD 2–3/hr.[34] These workers are often critical of research or journalism that portrays crowd workers as unskilled laborers working for extremely low wages. They cite their own higher wages, often in the USD 6–12/hr range, as proof that such reports are flawed. They

---

[32] See for example Ipeirotis 2010, "The new demographics of Mechanical Turk"; Martin et al. 2014; and gore313 et al. 2013/2014.

[33] See e.g. Jiang et al. 2015.

[34] See e.g. Ross et al. 2010.

argue that workers earning less than USD 6/hr are either inexperienced or simply not working hard.[35]

This view refines previous reports of crowd work pay, which have often focused on averages and overlooked the wide range of worker wages and the factors that influence wages. It also undermines some of the narratives advanced by researchers and employers as justifications for offering low wages. One such narrative proposes that any worker who relies on crowd work income must live in a "developing" country, and that workers living in "developed" countries must be passing time. In contrast, most professional Turkers live in the United States. Another narrative proposes that workers choose to work freely: if crowd work wages are too low for their liking, this narrative argues, they can choose other work. In contrast, many professional Turkers report that while they do not consider crowd work an objectively good work choice, they consider it, at present, their best choice. Such workers may be unable to secure traditional employment because of family care obligations, health problems, or geographic isolation.[36] Many professional Turkers are therefore not "free to choose" other work.

The existence of professional workers who rely on income earned through online labor market participation and who may not be free to choose other work suggests that employers' and market operators' responsibilities to workers may be underestimated in current practice. In AMT, as in other markets, a worker is very likely to work for many different requesters and

---

[35] See Martin et al. 2014.
[36] See gore313 et al. 2013/2014 and spamgirl 2015.

therefore not consider herself an "employee" of anyone in particular—and indeed legally workers are independent contractors. Yet in the absence of clear responsibilities, requesters' and platform operators' power over workers—combined with global competition among workers—may erode pay and working conditions over time.

1.3.2 Workers who rely on income earned through market participation should be considered first-class stakeholders

Workers who rely on income earned through participation in online labor markets are strongly invested in the sustainability of the market and in achieving market outcomes that benefit all participants. Their concerns and input regarding the design and operation of the market should be taken seriously. Formal processes for eliciting their input should be developed and integrated into market design and management practice.

AMT offers an extreme case of worker marginalization: Amazon presents AMT as offering computation rather than labor. Requesters can post tasks and review submitted work using AMT's application programming interface (API); that is, they can manage human workers through software, as if workers were themselves software rather than people. A special, but unfortunately necessary, case of the second message of the dissertation therefore obtains in the context of crowd work: crowd workers are people, not computers. While nobody explicitly argues otherwise, the implications of this fact are not fully appreciated in practice. It means that platforms should not be designed to

make it easy for employers to forget that workers are human beings, with diverse material needs and constraints, goals and preferences, extensible but finite abilities, complex but imperfect models of others' intentions and expectations, and complex but imperfect models of market processes and dynamics. Making human labor accessible through APIs, previously used only for nonhuman computing resources, may increase efficiency and be widely desirable. But this convention should not be made into a metaphor that governs design and operation broadly. Employers should understand the API calls made by their code not as computational acts in a purely computational system but as communicative acts in a computationally mediated socioeconomic system. Put shortly, employers should think of their code not as managing computing resources but as managing human beings. Researchers offering views of crowd work, or online labor markets broadly, through the lens of computational theories should consider the limitations and risks of this perspective with extreme care.

When employers forget that workers are human beings, adverse consequences follow for workers and employers alike. For example, when employers expecting "frictionless" interactions run into workers upset about unclear instructions, technical errors, or unexpected nonpayment, misunderstandings often erupt into heated, even vicious, public argument, consuming time, exhausting participants, and often destroying future work opportunities. Incensed workers may harass or blackmail employers they think have wronged them, while employers may retort that workers' work was poor.

But both parties are acting out of anger over betrayed expectations—expectations not understood by the other.

To improve work quality and worker-employer relations, platform operators should make design decisions that allow workers and employers to understand one another as human actors in a transaction of human interest. Specifically, platform operators should allow workers to post work-relevant information about themselves, such as work history, skills, and interests, allow employers to give workers information about the broader context of work, encourage employers to respond promptly and substantively to worker inquiries, and establish and support best practices for work management that reduce the need for unexpected or "emergency" communication and ensure that such communication is effective when it becomes necessary. Building channels for worker-employer communication and establishing shared understanding about their use will reduce the frequency of betrayed expectations, and the ensuing consequences: employers' perceptions of workers as lazy and unskilled, workers' perceptions of employers as miserly or cheating, and a widespread sense of distrust or even hostility.

It is of course harder to forget that workers are people in markets for in-person services such as transportation (e.g., Uber, Lyft) or domestic work (e.g., Handy, TaskRabbit). Yet even in these markets, workers are often disempowered. This occurs not only explicitly, through ostensibly technical market design decisions—e.g., Uber's automatic "deactivation" for drivers

whose ratings fall below a particular average[37]—but also more subtly, through the culture promoted by the firms operating the markets: it was not until summer 2014, for example—after six years in operation—that TaskRabbit stopped calling workers "rabbits."[38] While workers in markets for in-person services are not presented as computation—indeed market operators often emphasize distinctly human attributes of workers, such as friendliness, trustworthiness, and reliability—they often remain second-class citizens subject to the financial and technological power wielded by market operators.

1.3.3 Workers are not narrowly self-interested profit maximizers but situatedly rational actors

Workers in online labor markets are not narrowly self-interested profit maximizers: rather, they are people. While people, including workers in online labor markets, can be induced to act as narrowly self-interested profit maximizers under some circumstances, the assumption that people always— much less "naturally"—act this way is, as the Nobel laureate political economist Elinor Ostrom so delicately put it, not empirically supported.[39] Rather, like most people, workers in online labor markets care about many criteria beyond personal profit. In AMT, these criteria include obviously monetarily-related criteria such as speed of pay, criteria related to the fit between worker and task such as whether the worker is good at the task and finds it interesting, and procedural and relational criteria. The latter include perceived fairness of

---

[37] See e.g. Bercovici 2014.
[38] Said 2014.
[39] See for example Ostrom 2005, pp. 7-8, and Ostrom 2010, pp. 3, 19.

evaluation, requesters' responsiveness to worker communications, and, more broadly, the well-being realized by others—including both workers and employers—through market transactions. Professional Turkers especially express concern that employers receive good work, and chastise other workers for knowingly submitting unusable work or sharing information about employers' quality control processes that could be used to circumvent those processes. Workers have finite time and cognitive capacities, use complex and sometimes nonformalizable decision logics, and are situated within multiple levels of complex situations (e.g., action situations interlinked into polycentric systems). Many workers cooperate, spending some of their finite time and cognitive capacities helping one another, and even helping employers, in a variety of ways. They contribute free labor to shared information resources about tasks and employers (specifically, about which tasks are good and which employers treat workers well), teach one another about market processes, help employers improve their tasks, and discuss and enforce market norms they believe will produce good long-term outcomes for everyone and support sustainable livelihoods for themselves and other workers. While such activities may not immediately and obviously profit contributors directly, many workers seem to believe that they improve the overall functioning of the market and create diverse benefits for many market stakeholders. Indeed cooperative workers seem to see their own interests as aligned with other cooperative workers and requesters, and as aligned against those workers and requesters who exploit technical or procedural loopholes to profit at others' expense.

Market actors should therefore be seen not as narrowly self-interested profit maximizers but as boundedly rational actors situated within complex situations—i.e., as "situatedly rational" actors. Situated rationality augments the notion of bounded rationality—i.e., the notion that actors do not have complete information about market opportunities, processes, norms, and dynamics, or the cognitive capacity to make "optimal" decisions—with three propositions. First, situatedly rational actors can be understood as having "other-regarding preferences": they care not only about what happens to themselves but also about what happens to others. These preferences do not *necessarily* imply that such actors are altruistic in the "pure" sense. An actor's other-regarding preferences may arise from "enlightened self-interest"—i.e., from an understanding that their own well-being is bound up with that of others. The crucial point is that actors' take the outcomes obtained by others in the market into consideration. Second, the preferences of situatedly rational actors are complex and, importantly, *not fixed*; rather, an actor's preferred outcome in a given action situation is influenced by a wide range of factors, including rules, norms, and the actor's understanding of others' expectations and intentions (which may or may not be complete or accurate), and the actor's understanding of market dynamics (which may or may not be complete or accurate). Third, situatedly rational actors may act to *change* rules, norms, or others' expectations or intentions—and are likely to be aware of their ability to make such changes. Importantly, situatedly rational actors may act as narrowly self-interested profit maximizers in some situations. But this result is not

inevitable—on the contrary, when market actors act as narrowly self-interested profit maximizers, their interactions often produce suboptimal outcomes. This state of affairs may indicate poor market design.

1.3.4   Online labor markets are complex polycentric systems

Online labor markets and the sociotechnical ecologies around them are not monolithic, perfectly competitive markets but polycentric systems composed of complexly interlinked action situations.

The notion of polycentricity indicates that while the market platform itself may constitute the central locus of decision making for most actors, outcomes in online labor markets are strongly influenced by interactions and decisions that occur outside the platform. The various loci of action—e.g., the market platform itself, worker forums, research discourse, and shared information infrastructure such as external review sites—can be seen as forming a single system with multiple formally independent but connected and interacting centers of decision making.

Any decision taken by an actor—within the market or in a connected context—can be seen as occurring within an "action situation." In a polycentric system, action situations are interconnected by the consequences of the outcomes of the decisions that actors take within them. A simple example can illustrate these connections and their importance. Suppose crowd workers in a forum share information about technical problems with a particular employer's task. Experienced workers may then avoid that employer's tasks when

searching for work in the market. The employer, noticing that the quality of work they are receiving—or the rate at which work is being completed—has declined, may investigate, for example by reading blog or forum posts from other employers. They may learn about the worker forums, where they may, in turn, learn about the problems workers are experiencing with their tasks. The employer may then fix the problems and tell workers they have done so, leading experienced workers to resume doing their tasks.

At each step in this process, situatedly rational actors—workers and employers—have taken actions they believe will lead to their preferred outcome based on the information available to them. Any particular action situation is a complex setting with many characteristics. These include the actors involved (e.g., workers, employers, forum moderators) and their characteristics (e.g., are they narrowly self-interested short-term profit maximizers or do they cooperate with other actors to achieve long-term collective gains?), the formal roles of the actors and the rules governing their actions, the information available to them, possible outcomes and the relations between decisions and outcomes, and actors' valuations of outcomes. Actors may make decisions according to a variety of logics—e.g., financial, institutional, or moral. These logics may not always be quantitatively commensurable with one another—or formalizable at all.

Polycentric economic systems such as online labor markets and the "ecologies of practice" which develop around them are characterized by complexity at multiple scales: actors' decision making processes are complex,

individual action situations are complex, and the interlinkages between action situations are complex.

1.3.5 An interdisciplinary research agenda

The view of online labor market participants as situatedly rational actors in complex polycentric systems can help orient and evaluate software practice, research, and policy in ways that will yield outcomes most stakeholders see as better. With a view to this end, institutions supporting crowd work research should support an interdisciplinary research agenda linking software practice, empirical research, theory development, and value-rational analysis.

The diverse and evolving but generally accepted methods used for empirical research in human computation and human-centered computing—including methods typically considered quantitative such as survey research and log analysis, interpretive methods with a qualitative emphasis such as participant observation, and interventionist methods such as action research and research through design—are largely adequate to the challenge of more fully realizing the potential for online labor markets to create substantive new economic opportunities. The challenge lies in designing and conducting empirical research that is informed by, and informs, the development of richer theory, value-rational analysis, and software practice that surfaces and substantively addresses the diverse needs and aspirations of all stakeholders—rather than orienting toward the convenience and profits of platform operators and employers. Some research methods may also be attended by risks to

stakeholder trust or cooperation. Insofar as such risks are posed by methods perceived necessary for "objective" empirical research, criteria other than objectivity should be formulated, and methods appropriate to the alternative criteria adopted, adapted, or developed.

Theoretical work can be evaluated by at least three criteria: rigor, responsibility, and generativity. Rigor concerns the fidelity of frameworks, theories, and models used and developed to what is known about the phenomena modeled. In the context of crowd work, for example, research modeling participants as narrowly self-interested profit maximizers with complete information and freedom to choose from a practical infinity of paying work opportunities should be considered unrigorous, because this model is not empirically supported. Responsibility concerns the potential or actual effects of dissemination of the theoretical work itself. The elaboration and dissemination of social theory legitimates some arrangements and projects while delegitimating others—and these effects are to some extent independent of the empirical validity of the theory. Theoretical work therefore has stakeholders and consequences to which theorists can be seen as responsible or irresponsible. Generativity concerns the ability of frameworks, theories, and models derived through research to advance the development of new sociotechnical arrangements that contribute to the well-being of all stakeholders. Broadly, theoretical work should draw on and extend empirically supported frameworks, theories, and models to develop rich holistic understandings of online labor markets in their broader economic, social, and

cultural contexts—and make new arrangements possible that are seen as better by a wide range of stakeholders.

Value-rationality looks beyond empirical questions of what is happening in a particular case and theoretical questions about the general patterns it exemplifies to questions of social value and power: *Where are we going? Is this desirable? Who gains and who loses, and by what mechanisms of power? What should be done?*[40] Answering such questions necessarily involves taking a position, but this need not be an exercise in subjective editorializing or ideological posturing. Rather, it can and should be grounded in careful analysis of the needs, capabilities, constraints, and preferences of relevant stakeholders and the processes, norms, and dynamics of the systems within which they

---

[40] One locus for these and related questions in human-centered computing has been the interlinked discourses on "values in design" ("VID"; e.g. Flanagan et al. 2008; Knobel and Bowker 2011) and "values sensitive design" ("VSD"; e.g. Friedman 1996; Friedman and Kahn 2002; Le Dantec et al. 2009; Borning and Muller 2012). The view here—including the specific questions above—is more directly inspired by "phronetic" social research, an approach articulated by the Danish social scientist Bent Flyvbjerg and colleagues (e.g., Flyvbjerg 2001; Flyvbjerg et al., eds., 2012). The term "phronetic" or *phronesis*, variously translated as "prudence" or "practical wisdom," refers to Aristotle's third kind of knowledge—the others being *episteme*, the form of general theory that the modern natural sciences have been (relatively) successful at producing, and *techne*, craft knowledge. Advocates of phronetic social science propose that because context is crucial to human action and meaning, social science cannot produce abstract, general, universal, context-indepedent epistemic theory as the natural sciences can. They argue however that *phronesis*—the development and exercise of "practical wisdom" concerning practical human questions, including questions of policy—could be the goal and method of a reconstituted social science. This view has attained significant traction across the social sciences (see e.g. Schram and Caterino, eds., 2006; Flyvbjerg et al., eds., 2012). For my purposes here, the main distinction between VID and VSD, on one hand, and phronetic social science, on the other, is the centrality of power to phronetic analysis of social arrangements and policy. While VID and VSD furnish useful frameworks and methods for considering what values might be embedded in particular system designs and how to embed particular values into systems while they are being designed, phronetic analysis directs attention to the social relations by which decisions about what values should be embedded into systems *are made* and *by whom*. A further elaboration of phronetic social science, its possible use as a framework for orienting and evaluating HCC research, and its connections and differences to existing frameworks in HCC such as VID and VSD is deferred to a later publication.

interact. That is, it can, and should, be grounded in empirical research and framed by appropriate theory.

Finally, the practical part of the research agenda should extend the traditional system-building and -evaluation activity of research in human-centered computing beyond proofs-of-concept and short-term user studies to include a commitment to design, build, operate, maintain, and evolve systems for use by "real world" workers and employers in the completion of work and the construction of viable livelihoods. Design, construction, operation, and maintenance should be guided and evaluated by—and should inform in turn—relevant empirical research, theory, and value-rational analysis. The scope of all three modes of investigation should extend to considerations of long-term maintenance and the organizational structures that enable and structure software practice.

Industry collaboration will be crucial for the long-term sustainability of this work. Researchers undertaking such collaboration will need to attend to the distinct institutional accountabilities and cultures of research and business. Indeed business is grappling with challenges of "social responsibility" broadly, as can be seen in the relatively new discourses on corporate social responsibility and social entrepreneurship, the emergence of new structures for corporate governance such as the B Corporation, and the development of new computationally-mediated strategies for raising capital such as crowdfunding. HCC researchers aiming to create substantive and sustainable value for a broad diversity of stakeholders—including workers—through the creation and

operation of new online labor markets will need to grapple with questions previously beyond the scope of computing research, such as appropriate organizational forms for the long-term maintenance of computing systems and appropriate strategies and governance structures for raising capital.

1.4   Outline of the dissertation

The remainder of the dissertation proceeds in three chapters. Chapter 2, "Mechanical Turk and Turkopticon, 2008–2015," describes AMT from my perspective as a builder and maintainer of Turkopticon, one of the "formally independent centers of decision making" in the broader ecology of practice around AMT. Chapter 2 presents the empirical material from which the first three main messages of the dissertation—"Some workers are casual or transient; others are professionals who rely on income earned through market participation to meet basic needs", "Workers who rely on income earned through market participation should be considered first-class stakeholders" (and its special case, "Crowd workers are people, not computers"), and "Workers are not narrowly self-interested profit maximizers but situatedly rational actors"—are argued. Chapter 3 draws extensively on empirical social science, especially experimentally and ethnographically informed research in economics, to elaborate the theory of situatedly rational actors in polycentric economic systems within the context of online labor markets and AMT specifically. Chapter 4 synthesizes the empirical and theoretical material from the previous two chapters to offer an interdisciplinary research agenda that links software

practice, empirical research, theory development, and value-rational analysis in the context of online labor markets.

Chapter 2

Mechanical Turk and Turkopticon, 2008–2015

2.1 Summary

2.1.1 Mechanical Turk

Amazon describes Mechanical Turk (AMT) as a "marketplace for work." It was originally built to help Amazon deduplicate its huge product catalog, a task which could not be fully automated. The basic AMT workflow is straightforward in concept. First, an employer, called a "requester," posts a task, called a "human intelligence task" or HIT. Second, workers, many of whom self-identify as "Turkers," encounter the HIT while browsing or searching. Third, workers do the HIT. Fourth, the requester reviews workers' submissions and decides whether to approve or reject them. Workers are not paid for rejected submissions. There are a wide variety of HITs. HITs from industry requesters include content categorization, terms of use enforcement, transcription, translation, writing, and metadata creation (e.g., image labeling). And many academics, especially social scientists, use AMT as a human subjects pool for survey data collection or behavioral experiments.

In practice, this straightforward process is complicated by a wide range of unexpected outcomes, mistakes, miscommunications, and even abuses—on the part of both requesters and workers. Perhaps the most well-known complication is a consequence of the rejection feature. Requesters may reject work for any reason, and workers have no technical or legal recourse within

AMT against requesters who they suspect may have erroneously rejected their work—or even done so maliciously, with the intent to use it anyway. Thus while wage theft is common in other low-wage industries, AMT has made crowd work the first in which it is legal—i.e., in which it is not "theft" but a legitimate and normal use of an intentionally designed platform feature. This "feature" gives requesters unique power over workers, and creates uncertainty among workers that some researchers have argued leads to a "vicious circle" of low-quality work and low wages.[41] This vicious circle contributes to increasingly complex quality control schemes, adversarial worker-requester relations, and a climate of mistrust, fear, anxiety, and even hostility among workers.

2.1.2 Turkopticon

One check on abuse of the rejection feature is Turkopticon, a tool workers use to review employers. Lilly Irani and I built the tool in 2008 and have maintained it collaboratively with workers since then. Over 40,000 workers have created Turkopticon accounts over the last six years, and over 13,000 of them have posted at least one review. Together, they have posted over 200,000 reviews of 34,000 requesters. While the Turkopticon user base is small compared to the number of workers *signed up* for AMT—Amazon reported in 2011 that 500,000 people had created worker accounts[42]—it seems that most professional Turkers use Turkopticon, and indeed consider it a crucial livelihood tool.[43] But

---

[41] See Ipeirotis 2010, "Mechanical Turk, low wages, and the market for lemons"; and Bederson and Quinn 2011.

[42] See e.g. `requester.mturk.com/tour`.

[43] See e.g. `redd.it/sdyuo`.

Turkopticon is far from perfect: aside from shortcomings one might expect from a small independent information utility with no paid staff, Turkopticon is a target for occasional abuse from both workers and requesters (see Sec. 2.3.3.2), and we have struggled to respond effectively to this abuse.

2.1.3 The broader Turker ecology of practice

Turkopticon fits into a broader ecology of worker mutual aid. The most crucial parts of this diverse and complex network of relationships, practices, and technologies are the worker forums. Through worker forums, new workers "learn the ropes" from more experienced Turkers, who share information about platform processes, market norms, good and bad requesters and tasks, and specialized software. Forums also provide a virtual "water cooler" where workers interact informally, provide social support, and build trust and community. Needs and desires expressed in worker forums also drive development of specialized software built and maintained for workers by workers. And some requesters visit forums to communicate with workers about tasks and build trust.

2.1.4 Requester ecologies of practice

Like Turkers, requesters teach one another, share information, and develop and circulate specialized software. But requester discourse circulates in a broader variety of media: forums, blogs, meetups, workshops, conferences, and peer-reviewed papers. Most software developed by requesters is for requesters, not

Turkers, and as many requesters work in the information technology industry, many requesters are programmers or have relatively easy access to programming expertise. As a result, there is a great deal of specialized software available for requesters. Software and techniques developed among and for academic requesters—e.g., for preventing workers from taking survey tasks more than once, linking AMT to popular software packages, or running behavioral experiments on AMT—are often free or open source, while industry requesters frequently build their own.

2.1.5 Representations, connections, and interpretations

Scholarly and popular portrayals and interpretations of AMT, Turkopticon, and worker collective action efforts are diverse. Early discussion of AMT within the technology industry emphasized its technical novelty and focused on the reduced costs and expanded operational capabilities it offered programmers and organizations. A second wave of journalistic coverage from outside the technology industry focused almost exclusively on the question of whether AMT is a "digital sweatshop." As Turkopticon gained visibility, the notion of Turkopticon as a digital union was added to the criticism of AMT as a digital sweatshop (see Sec. 2.6.3). And as the terms "gig economy," "sharing economy," "Amazon economy," and, more recently, "on-demand economy" appeared in the discourse around the organization of work in an age characterized by continuing technological change and prolonged economic recession, a variety of commentators linked AMT, on one hand, to these new computationally

mediated modes of work, and, on the other, to the ways Amazon itself organizes work in other areas of its operations (especially in its warehouses).

What remains unclear is what is at stake in these discussions. Critiques of AMT specifically are often shrugged off by crowd work researchers, among whom the platform's shortcomings are well-known but considered idiosyncratic and easily overcome, at least in theory, by better designs in the next generation of platforms. But AMT is one embodiment of a particular way of thinking about work, business, management, the economy broadly, and technology's role in all of these. This view can be summarized by five propositions. They are as follows.

First, because social welfare is improved by economic growth, "the social responsibility of business is to increase its profits."[44] Second, the task of corporate management is to ensure that profits are constantly increasing, and this is best achieved through quantitative analysis and management of the business and its operations—quotas, optimization methods, and data-driven decision making made possible by extensive technological surveillance of operations. What cannot be quantitatively linked to increasing profits is not actionable; thus management is rarely empowered to expend time or resources attempting to treat employees well, create sustainable livelihoods, produce quality products, or mitigate social or environmental harm caused by production, use, or disposal of products—unless such activities can be clearly linked to profit growth. Third, workers should be grateful for any paying work

---

[44] This phrase was first used by the Nobel laureate neoclassical economist Milton Friedman as the title of an influential *New York Times Magazine* essay (Friedman 1970).

offered them, and not complain about working conditions, low pay, or lack of benefits such as health insurance. Fourth, the main opportunity presented by information technology is to achieve existing business goals more fully, either by expanding operational capabilities or by reducing costs. This can be achieved by optimizing existing operations and "scaling up" to deliver more of current products and services more efficiently, or by "disrupting" existing operations and replacing them with new products or services. Fifth and finally, the information technologist's job is not to question existing goals but to use their technical expertise to act on the goals given them by management, investors, or analysts.

This view is to some extent also embodied by online labor markets such as Uber and Handy. I agree with the management researcher and educator Henry Mintzberg that this view of the economy and technology's role in it is corrosive to long-term social and economic welfare.[45] Five alternative propositions can offer a different view.

First, there are diminishing social welfare returns to economic growth; as a result, business has broader and more complex social responsibilities than merely increasing profits. Second, the task of corporate management must be to discern business's responsibilities to its various stakeholders, develop a realistic vision for fulfilling them, and lead and manage efforts to do so. Some responsibilities will resist quantification; management must nonetheless make serious efforts to fulfill them and to understand how well they are being

---

[45] See e.g. Mintzberg 2005, 2009.

fulfilled. Third, workers are a primary stakeholder of any business, and reliable, sustainable livelihoods are a crucial driver of long-term social and economic mobility and welfare. Therefore working conditions and livelihood reliability must be central preoccupations of management; worker complaints should be taken seriously as indicators that business is not meeting one of its primary responsibilities. Fourth, increasing the profitability of existing business practices is a waste of the main opportunity presented by information technology, which is to increase the social responsibility of business and its contribution to socioeconomic mobility and welfare broadly. Fifth and finally, technologists should use their expertise in service of the transformation of business, not in the service of existing goals without regard to long-term social value.

I agree with crowd work researcher Michael Bernstein[46] that what is at stake in any actionable discussion of the future of crowd work and online labor markets is the future of work itself—that is, the future of the organization of labor and the allocation of resources in our sociotechnically complex society. Thus the question of the future of crowd work and online labor markets is bound up with the question of the future welfare of our society as a whole— that is, with questions such as the socioeconomic mobility of future generations and the future of democracy.

---

[46] Bernstein 2015, "Crowdsourcing a meeting of minds."

2.2   Mechanical Turk

2.2.1   Origin story

At a meetup at a Santa Monica technology company in 2010, Amazon executive Sharon Chiarella told attendees—who had been invited to "learn about Mechanical Turk, meet the Amazon folks, and network with other [Amazon Web Services] tech users in Los Angeles"[47]—that AMT was originally built to help clean data coming into Amazon's huge product catalog from its many vendors. Sometimes different vendors would each add an entry for the same product. As a result, customers searching for products would be shown multiple results for the same product. Presumably it is important for Amazon to keep track of multiple vendors selling the same product. But for customers, receiving multiple results for identical products when searching is useless and frustrating. If the database entries from each vendor had been as identical as the products, the problem could have been solved computationally. But they were not, and efforts to hide multiple search results for identical products programmatically had been only partially successful. Eventually the idea surfaced to crowdsource the problem. Amazon engineers built a site through which Amazon employees, in their spare work time, could contribute to the process of identifying and hiding the duplicate entries. This was successful, and it was eventually opened to workers and requesters outside Amazon. It was extended to support tasks other than duplicate product identification. A

---

[47] Hammond 2010.

mechanism for paying workers was added. And the cheeky but truthful tagline "artificial artificial intelligence" was coined to describe the new service.[48]

With these additions, AMT came to have a portentous significance in the information technology industry and in computer science as a field. It became simultaneously the next step in both artificial intelligence and cloud computing. In describing the system in a 2006 lecture at MIT (see Fig. 2.1 for slide), Amazon CEO Jeff Bezos said, "You've heard of software-as-a-service. Well, this is human-as-a-service."[49]

Put The Human In The Loop

```
while (read (photo))
{
  ContainsHuman = CallMechanicalTurk("Contains Human?", photo);

  if (ContainsHuman)
  {
    acceptPhoto(photo);
  }
  else
  {
    rejectPhoto(photo);
  }
}
```
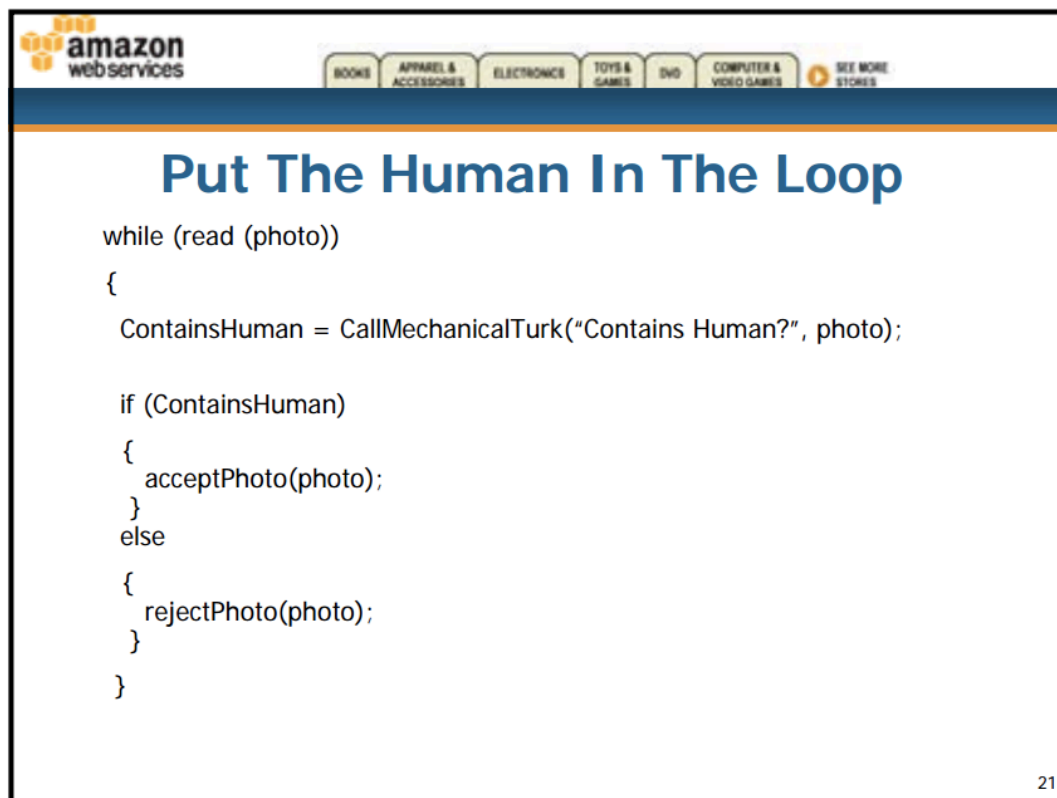
21

Figure 2.1. A slide from Amazon CEO Jeff Bezos' 2006 keynote at MIT's Emerging Technologies conference. The slide shows how programmers who are end users of AMT can write code that calls on AMT to recruit humans to do tasks that could not otherwise be performed programmatically.

---

[48] See also Pontin 2007.
[49] Bezos 2006.

2.2.2 The process

The basic process of posting tasks to AMT and completing them is straightforward, at least in concept (see Fig. 2.2). Requesters post tasks to the site. This process includes posting the reward for the task. Amazon charges the requester a fraction of the posted reward in addition to the reward price. Workers choose what tasks to do, then do them. After the worker does (i.e., submits) a task, the requester can either approve or reject it. Workers whose work is approved are paid. Workers whose work is rejected are not paid, although requesters may choose to keep and use the work in any case. Finally, requesters may choose to give some workers a bonus of any size.
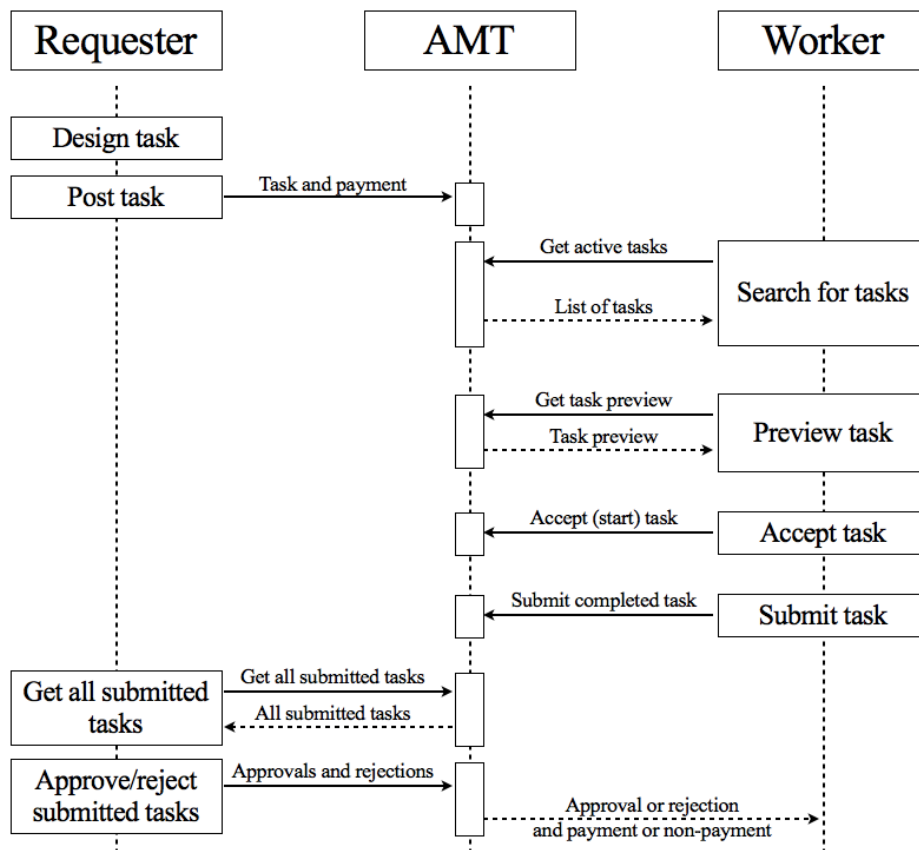

Figure 2.2. The basic AMT workflow.

There are many details and potential variations on each step of the basic process. Requesters post tasks either through AMT's web interface or by writing software that calls the AMT API. The information posted includes the title of the task, the payment workers will receive on completing it satisfactorily, any qualifications required to complete it, a short description (in addition to the title), the time limit for the task, and the number of tasks available for one person to complete. If the task requires that each person submit only once, this number appears to the worker as 1, even if the requester has asked for many replies. Some tasks, however, do not need responses by unique workers. The number of tasks available in such "batches" may number in the thousands, tens of thousands, or even hundreds of thousands. On posting the task, Amazon charges the requester a fraction of the posted payment price. This fraction is 10% for most tasks; it is 30% for tasks in which the requester restricts the task to workers with the "Masters" qualification.[50]

After a task is posted, it appears in the HIT listing, which lists ten HIT groups per page (see Fig. 2.3).[51] Workers can sort the listings by how old the

---

[50] "Masters" are in practice a worker pool curated by Amazon. Panos Ipeirotis reported in late 2012 (Ipeirotis 2012, "Mechanical Turk changing the defaults") that "a current search reveal[ed] 20,744 workers" with the Masters qualification. If the requester uses the web interface to post the task, the default option is to restrict the task to Masters workers. To change this, the requester must know that this is the case and know where the option is to change the setting. The process by which a worker obtains the Masters qualification is not public, and is vigorously speculated about by workers, including those who have, to their surprise, received the qualification.

[51] The distinction between "HIT" and "HIT group" is subtle. Technically, a HIT is a single task; for example, manually transforming an image of a receipt into structured data by entering different values on the receipt into different fields in a form. A HIT group is a group of HITs that use an identical form and process but include different data elements; for example, a thousand receipt entry HITs with identical forms but different images (i.e., different receipts). Workers sometimes refer to large HIT groups as "batches" or "batch HITs," but also use

HITs are, how soon they will expire, how much they pay, how many tasks are available, how much time they allow, or (perhaps occasionally usefully) alphabetically. They can also search for HITs matching specific strings, HITs that pay at least a certain amount, HITs that require the Masters qualification, or HITs that are available to them (see Fig. 2.4). (To produce results for the latter kind of search, the site checks the qualifications required by available HITs against the worker's qualifications.) When looking at a page in the listings or search results, a worker can click on the title of a HIT to see more information (specifically, the description, keywords, and required qualifications). They can also click a link to see the first page of the HIT itself. This page may be the entirety of the HIT, it may be a brief information page, or it may be entirely uninformative. After viewing this "preview," the worker may choose to accept the HIT. The worker then has the allotted time to complete the HIT. While completing the HIT, the worker may choose to "return" the HIT. (Workers often do this if a HIT turns out to have technical problems, or to take longer than they expected.) Or the worker may run out of time, in which case AMT classifies the HIT as "abandoned."

---

sentences like "I did a hundred of these [HITs]" to describe having completed many HITs in a single HIT group.

Figure 2.3. The AMT HIT list.

Figure 2.4. Workers can use different criteria to look for tasks.

Once the worker submits the HIT, the requester can approve or reject it. Workers are paid for approved HITs and not for rejected HITs. If the requester does neither, AMT eventually approves the HIT and pays the worker. The time

after which this occurs is called by workers the "auto approve" or "AA" time of the HIT group. The default, and maximum allowed, auto approve time is 30 days; some prolific requesters set it to seven days, or even 24 or 48 hours. Workers often share this information once they know it, and attentive requesters know that workers value quick evaluation of their work. Requesters can also give workers bonuses when evaluating work (or even after they have approved or rejected work).

The option to reject work is broadly assumed to be intended to prevent workers from quickly submitting useless work in the hope of being paid anyway—and it works relatively well in this capacity. When rejecting work, requesters are expected to offer some explanation for the rejection. But the only technologically-enforced constraint over the process is that requesters cannot leave the explanation text field blank, so sometimes they offer unhelpful "explanations" such as "1," "X," or "."

The AMT participation agreement (typically referred to by workers and requesters as the "terms of service" or "TOS") says (Sec. 3a):

> Upon completion of Services [i.e., work] to Requesters' reasonable satisfaction,
> Requesters must pay Providers [i.e., workers] for their Services.[52]

But the agreement does not define "reasonable," nor would it be possible to do so. In practice, requesters can reject work for any or no reason (beyond, for example, "X"). If a worker finds the rejection unreasonable, they can file a complaint with AMT. But AMT does not mediate in disputes between workers

---

[52] Amazon Mechanical Turk 2014.

and requesters. In fact the TOS explicitly disclaims any such responsibility (Sec. 3f, orig. emph.):

> **Disputes between Requesters and Providers.** Your use of the Site is at your own risk. Because Amazon Mechanical Turk is not involved in the actual transaction between Providers and Requesters, Amazon Mechanical Turk will not be involved in resolving any disputes between participants related to or arising out of the Services or any transaction.[53]

Typically, workers filing complaints of unfair rejection with Amazon receive form responses to the same effect.

2.2.3 Tasks

Common tasks posted to AMT include search result relevance evaluation, image and audio transcription, translation, writing (i.e., "content creation"), revising others' writing or transcriptions, evaluating user generated content for compliance against site terms of service (e.g., is this forum post profane? is this avatar pornographic or otherwise inappropriate?), producing metadata for images or other database entries (e.g., products), testing the usability of websites or mobile applications, and completing surveys and participating in behavioral experiments for academics and market researchers.

An illustrative sample of HITs appears in Table 2.1. See Figs. 2.5 and 2.6 for an example from the AMT interface. Pay ranges from one cent to tens of dollars (the latter usually for transcription of multi-hour audio clips); most HITs pay less than USD 0.50.

---

[53] Amazon Mechanical Turk 2014.

Title	HITs available	Reward (USD)
Find images of these real estate agents	136,725	0.04
Find the mobile app link	12,748	0.03
Type the text from the images, carefully. Productivity and bonuses guaranteed.	9,498	0.01
Judge the appropriateness of a product for a question	5,294	0.03

Table 2.1. Some tasks from Amazon Mechanical Turk.

Figure 2.5. Details of a HIT from the AMT HIT list.

Figure 2.6. Preview of a HIT.

2.2.4 Requesters

AMT requesters are diverse. The most prolific as of mid-January 2015, CrowdSource, is an intermediary for others, and offer a variety of task types; past "most prolific" requesters such as CrowdFlower have adopted similar roles. Many requesters, however, specialize: SpeechInk and CastingWords focus on transcription, VidAngel on family-friendly movie editing, and Tagasauris on image metadata. Some requesters are special-purpose accounts for particular projects run by particular individuals or organizations. For example, social media giant LinkedIn posted business card transcription tasks under the requester name "Oscar Smith," and a Google speech recognition project posted tasks under the name "Project Endor."

Because there is a power law distribution over participation among requesters as well as workers—that is, a small fraction of requesters post most of the work—the history of workers' experiences in AMT is partly the history of specific requesters and their practices. Workers often express deep ambivalence about their relationships with the most prolific requesters. Prolific requesters often start out beloved—for the volume of work they make available and the regularity with which they post it—but their relationships with workers often become strained as they grow, make their processes more efficient, and seek to cut costs. As this slow process unfolds, workers may find themselves feeling trapped: a worker may want to stop doing tasks for a particular requester whose practices they feel have begun to adversely impact workers. But they may need the money and stability offered by a prolific requester. Even if the

requester has lowered pay, workers familiar with the requester's practices may continue to work for them instead of taking the time to look for other requesters who might treat them better or pay more. The "search costs" and "switching costs" associated with finding other reliable requesters may effectively constrain workers to persist in situations they no longer find satisfactory—a situation that mirrors some workers' relationship with AMT and online labor markets broadly.

The story of one particular requester, CrowdFlower, illustrates well the complexity of worker-requester relationships and the difficulty of assessing the overall impact of a particular requester on workers' livelihoods and well-being. CrowdFlower, a San Francisco-based startup originally called Dolores Labs, no longer posts tasks to AMT but is well remembered in Turker collective memory as one of the most prolific, and controversial, requesters ever to post HITs. CrowdFlower acted as an intermediary for other organizations and individuals who wanted to post large task batches to AMT but lacked the technical expertise to do it themselves. CrowdFlower built a platform that workers and other requesters could interact with that provided many features not offered by AMT itself. CrowdFlower's platform managed workers and allowed requesters to manage posted and submitted work. In 2009, CrowdFlower co-founder Lukas Biewald was instrumental in collecting the initial corpus of Turkopticon reviews (see Sec. 2.3.1). He was enthusiastic about the prospect of making information about requesters widely available to workers. He reasoned that such information would both pressure neglectful requesters to improve their

practices and reward well-behaved ones—and in his view, at the time at least, CrowdFlower was among the better requesters. But by 2011, CrowdFlower was known among workers mainly for the diversity of its tasks (posted for a variety of paying clients) and their uniformly low pay. And in 2012 the firm became the defendant in a class action lawsuit filed by Oregon-based crowd worker Christopher Otey. The suit alleged that although Otey and other crowd workers had been required to agree that they were independent contractors—not employees—before completing work for CrowdFlower, the degree of control the firm exerted over the work through its platform made them employees in practice, and in the eyes of the law. Because they were paid less than minimum wage but were, they alleged, employees in practice, CrowdFlower was in violation of the Fair Labor Standards Act (i.e., the minimum wage law). The plaintiffs' complaint alleged that Biewald had stated in public that they paid many of their workers USD 2–3/hr, and offered links to videos posted to YouTube documenting these statements. (The videos have since been removed.) CrowdFlower stopped posting tasks to AMT in late 2013 to focus on other channels. Sadly, despite the firm's helpful involvement in Turkopticon's early days, CrowdFlower has more low ratings from Turkopticon users than any other requester. As of early 2015, the lawsuit is still in court; the judge has rejected two settlements proposed by the parties.[54]

---

[54] See *Otey et al. v. CrowdFlower, Inc., et al.*, 4-CV-05524-JST.

2.2.5 Complications

In practice, the Turking process is complicated—for workers and requesters—by a wide range of unexpected outcomes, mistakes, miscommunications, and even abuses—on the part of both requesters and workers. These complications arise in part because doing remote work well through a complex computer information system is hard, even for well-intentioned participants, and in part because a few market participants *are* in fact narrowly self-interested short-term profit maximizers.

Perhaps the most well-known complication is a consequence of the rejection feature. Requesters may reject work for any reason, and workers have no technical or legal recourse within AMT against requesters who they suspect may have erroneously rejected their work—or done so maliciously, with the intent to use it anyway. Thus while illegal wage theft is common in other low-wage industries, AMT's rejection feature has made wage theft legal in crowd work—in crowd work, wage theft is not "theft" but a legitimate and normal use of an intentionally designed platform feature. This feature gives requesters unique power over workers, and creates uncertainty among workers that some researchers have argued leads, at least in some parts of the market, to a "vicious circle" of low-quality work and low wages. For example, Panos Ipeirotis wrote in 2010:

> Effectively, what Amazon Mechanical Turk is today [i.e., was in 2010] is a market for lemons [...]. A market for lemons is a market where the sellers cannot evaluate beforehand the quality of the goods that they are buying. So, if you have two types of products (say good workers and low quality workers) and cannot

tell who is whom, the price that the buyer is willing to pay will be proportional to the average quality of the worker. So the offered price will be between the price of a good worker and a low quality worker. What [would] a good worker do? Given that good workers will not get enough payment for their true quality, they leave the market. This leads the buyer to lower the price even more towards the price for low quality workers. At the end, we only have low quality workers in the market (or workers willing to work for similar wages) and the offered price reflects that. This is exactly what is happening on Mechanical Turk today [i.e., was happening in 2010]. Requesters pay everyone as if they are low quality workers, assuming that extra quality assurance techniques will be required on top of Mechanical Turk.[55]

While Ipeirotis later wrote that he believed subsequent changes to AMT had improved the situation,[56] the extent to which this is the case is hard to discern. And while changes to AMT—perhaps mainly the addition of the Masters qualification—may have improved requesters' ability to secure quality work, Ipeirotis' 2010 observation that "there is a symmetric market for lemons on [the requester] side"—i.e., that workers struggle to identify good requesters—still seems apt.[57] The difficulty for workers of distinguishing well-intentioned requesters from scammers, and vice versa, contributes to increasingly complex quality control schemes (on both sides of the market), adversarial worker-

---

[55] Ipeirotis 2010, "Mechanical Turk, low wages, and the market for lemons." See also Bederson and Quinn 2011.

[56] Ipeirotis 2013.

[57] "Scam requesters post HITs, behave badly, and cause good workers to avoid any newcomer. New requesters then get only low quality workers, get disappointed with the quality of the results[,] and [...] leave the market." Ipeirotis 2010, "Mechanical Turk, low wages, and the market for lemons." The consistency in the responses to Lilly Irani's 2008 and 2013 "Turkers' Bill of Rights" surveys (Irani, ed. 2008, 2013) suggests that not much had changed for workers between fall 2008 and summer 2013.

requester relations, and a climate of distrust, anxiety, and even hostility among workers.

This dynamic is augmented by other properties of the market, especially the scale of market interactions, the algorithmic management of work, the disparate expectations requesters and workers bring to their interactions, and the effect of rejection statistics on workers' ability to get work. One requester may receive work from thousands of workers in a single HIT group. Requesters may post HITs and review (i.e., approve or reject) submissions from workers through AMT's application programming interface (API); that is, they can write software to post and review tasks. This allows requesters to automate complex workflows. For example, a requester may have 100,000 photos they want workers to tag. They may post each photo twice, asking different workers to submit tags. They may write a program to compare the workers' submissions. If the submissions agree, the program may pay them both. If they disagree, the program may post the photo a third time. If the third worker's submission agrees with one of the original two, the program may pay the two workers in the "majority" and reject the dissenter's submission. Because these processes may be complex—and because they are handled by software written by humans—they are error-prone. Errors in workflows managed by software often lead to worker confusion, accusations of intentional requester malfeasance, and stress and wasted time for all parties.[58] And when workers contact requesters to

---

[58] This stress and wasted time grows to include the members of the Turkopticon community when disputes over negative reviews erupt between workers and requesters. In such disputes each side often claims to possess evidence that proves they are in the right. We have learned

seek explanations for unexpected rejections, they are often dismayed to receive, in return, no response, a canned and irrelevant response, a slow response, or—perhaps worst of all—a rude response indicating that the requester is not inclined to spend time to figure out what happened to cause the worker to be rejected for a task worth a few cents. Requesters primed to expect quick and "frictionless" interactions with AMT often write software to post and review tasks, and ignore worker communications or offer cursory responses. One requester told Lilly Irani:

> You cannot spend time exchanging email [with workers]. The time you spent looking at the email costs more than what you paid them. This has to function on autopilot as an algorithmic system...and integrated with your business processes.[59]

Workers, in contrast, expect—or at least wish—to be treated as "human beings, not algorithms,"[60] and to receive what they see as due consideration for their concerns from requesters. Responsiveness and communicativity are especially significant concerns for workers who rely on Turking income to meet basic needs; if a major problem occurs and a workers finds themselves short payment for work they spent a significant amount of time on, can they rely on a requester to read, reply to, and act on their communications about the problem? If not, working for such a requester may pose a significant risk. These

---

to refuse such implicit appeals to function as a court of law and adjudicate such claims. Lilly Irani articulated our position on this topic concisely: "Turkopticon is not a truth machine. It is a context-adding machine." See Sec. 2.3.4.

[59] Irani and Silberman 2013, p. 614.

[60] See e.g. Harris 2014.

conflicting expections, desires, and constraints lead, predictably, to frustration.[61]

Even once their work is approved, some workers face additional complications getting paid. Turkers can be paid in US dollars, Indian rupees, or Amazon gift card points. Most Turkers receiving their earnings in rupees receive their earnings by check. But the complexity of the Indian postal system sometimes results in delivery failures, prompting Indian Turkers to change the addresses associated with their worker accounts to addresses they expect will more reliably receive mail (e.g., to an office rather than a home address). But in recent years, Amazon has made it harder for non-US workers to obtain accounts, which has led to the creation of a black market for Indian AMT accounts. To prevent Indian Turkers from selling their accounts, which may sell for hundreds of US dollars, Amazon sometimes suspends the accounts of Indian Turkers after a change of address. This mechanism sometimes erroneously suspends Indian Turkers changing their addresses for legitimate reasons such as mail delivery failure.[62]

---

[61] See Silberman et al. 2010, "Sellers' problems in human computation markets"; Silberman et al. 2010, "Ethics and tactics in professional crowdwork"; Bederson and Quinn 2011; and Martin et al. 2014.

[62] This phenomenon was described by Turker Manish Bhatia at a panel in 2014 (Milland et al. 2014).

2.3   Turkopticon

2.3.1   Overview

Turkopticon is an information system Turkers use to review employers. It has two main parts: a web database application and a browser extension. The web application lets workers review requesters. The reviews include qualitative and quantitative elements. The browser extension aggregates the quantitative elements of all reviews of a particular requester and adds them to the HIT listing next to HITs posted by that requester. This allows a worker to see what other workers have said about a requester before accepting work from them.

Lilly Irani and I built the prototype system in October 2008, launched it in January 2009, and have maintained it since then. As of mid-January 2015, Turkopticon hosts over 43,500 registered users and over 201,000 reviews of 34,000 requesters posted by over 13,900 users. In the last month (15 Dec 2014– 15 Jan 2015), 1,107 new user accounts have been created, and 1,270 users have posted 7,249 reviews. The Chrome Web Store and Mozilla Firefox Add-On Directory report a combined total of 29,374 users of the browser extension. And forum discussions suggest that most Turkers who rely on Turking income to meet basic needs consider it a crucial livelihood tool. The tool is powered by workers themselves. Workers contribute requester reviews to the database, and volunteer moderators—also workers—hide reviews that appear spurious or violate the community's civility guidelines. As with AMT itself (and most online platforms), there is a "power law" distribution over participation: the ten most active reviewers have posted between 100 and 200 reviews each in the last

month (15 Dec 2014–15 Jan 2015), but most reviewers who have posted at least one review in the last month have posted only one review. Requesters report that the prospect of negative Turkopticon reviews influences their decision making while posting HITs and reviewing submitted work. And one study of Turkopticon, conducted in the summer of 2014 by a group of economists at the University of Minnesota, found that Turkopticon reviews both accurately reflect requesters' propensities to reject work and affect requesters' ability to have work completed.[63] I am the system's de facto lead programmer and database administrator. Lilly Irani and I work with Turkopticon's most active users, including the volunteer moderators, to manage community issues and support users who run into problems.

2.3.2 Turking with Turkopticon

Using Turkopticon adds complexity to the Turking process (see Fig. 2.7). Turkers can interact with Turkopticon in a variety of ways. The simplest and most common is to view aggregated review data while selecting tasks from the AMT task list. Turkopticon adds new information to this interface (see Figs. 2.8 and 2.9), and workers using Turkopticon look at and consider this information when choosing tasks. If the aggregated quantitative information is not decisive in helping the worker make a decision, the worker may click a link in the Turkopticon interface element to view the individual reviews, including comments (some of which are lengthy), posted by other workers about the

---

[63] Benson et al. 2014.

requester (see Fig. 2.10). While this can be time-intensive, it often has benefits. "I should have read the reviews here before working for this requester" is a common statement in unfavorable Turkopticon reviews.

A worker may post a review of a requester after, or even before, they have completed a task posted by that requester. Workers sometimes post "preliminary reviews" that leave some parts of the review form blank in order to warn others about technical problems or an unusual or confusing task design (see Fig. 2.11). Workers often post reviews after completing a task; at this point, they have a good deal of information about the task, but do not yet know how the requester will review it. Workers therefore often edit reviews after after their work is approved or rejected by the requester—which may happen up to 30 days after the worker completes the task.

2.3.3 Outcomes and complications

Turkopticon appears to have changed the decision making process in approving or rejecting work—at least among requesters who know about it. One requester told me that having a bad Turkopticon reputation made it effectively impossible to attract workers who would submit quality work. But it may be too easy to get a bad reputation: even a small technical problem with a task can lead to a string of bad reviews if it isn't addressed quickly. And Turkopticon has other problems. From an administrative perspective, many of the problems fall into two major categories: problems arising from disagreements about how
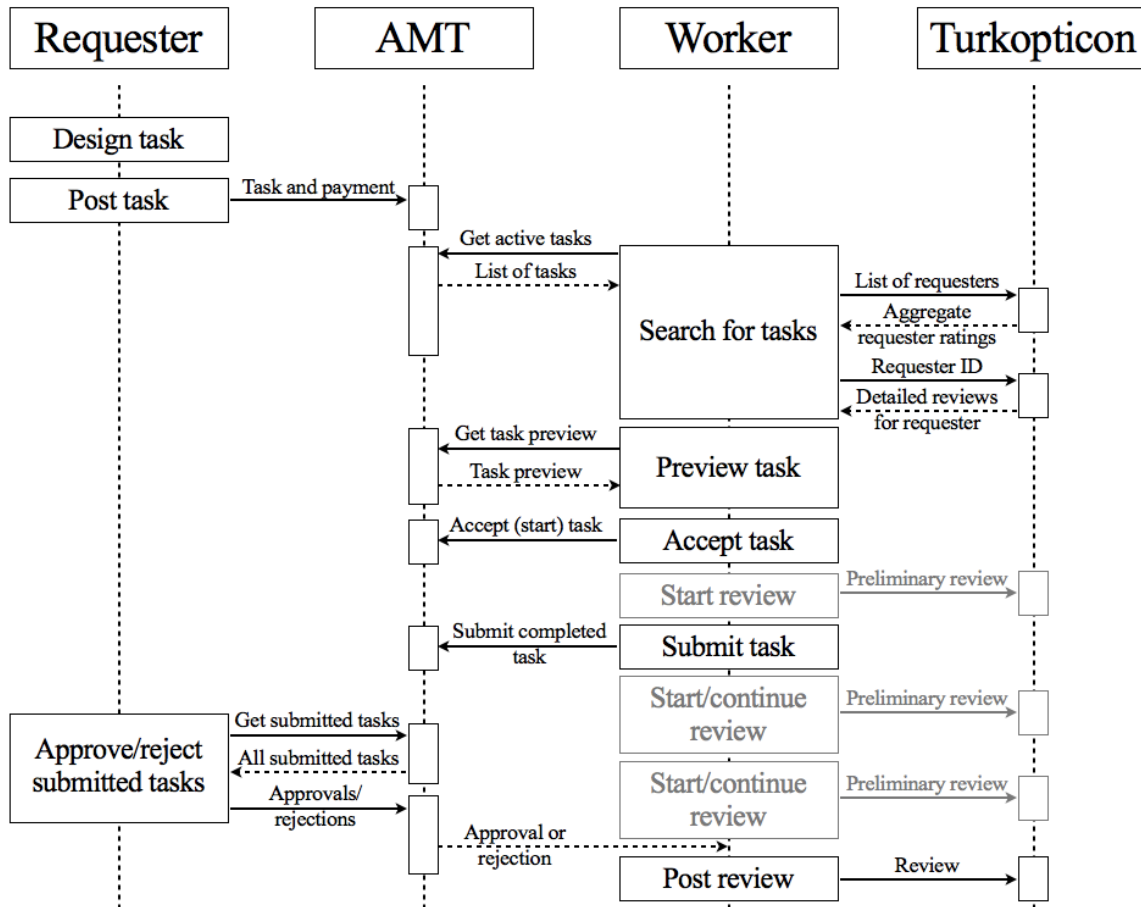
Figure 2.7. Turking with Turkopticon.

to use the review form and problems arising from the absence of a strong connection between a worker and their Turkopticon account.

2.3.3.1 The review form

The review form, designed in 2008 after coding the responses to Lilly Irani's initial "Turkers' Bill of Rights" survey,[64] includes 5-point Likert scale entries for four requester attributes—generosity of pay, fairness (ostensibly of approval

---

[64] Irani, ed. 2008.

Figure 2.8. The AMT HIT list with Turkopticon.
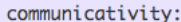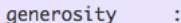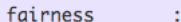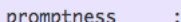
Figure 2.9. A user can mouse over the arrows added by Turkopticon to the AMT HIT list to see aggregate Turkopticon review data about the relevant requester.

and rejection decisions, but used by workers for a variety of purposes), speed of pay, and communicativity. (The attributes are discussed further below, in Sec. 2.3.4.1.) We do not give guidance about what counts as, for example, a 5/5 for generosity of pay, or a 2/5 for communicativity; on the contrary, the first instruction on the review form is "Give the ratings you feel best describe your

Figure 2.10. Reviews on the Turkopticon site.

experience."[65] This ambiguity was a strength in Turkopticon's early years, as it created space for workers to discuss their experiences, compare notes, and develop a collective sense of the range of requester behavior—and within that, what ought to be considered good and what bad. But in recent years, with veteran Turkers having to some extent established a shared understanding of

---

[65] See `turkopticon.ucsd.edu/report` (Turkopticon login required).

Figure 2.11. A review without ratings.

how to use the rating system, newcomers who lack this understanding—or users who disagree with the common usage—can cause tension.

In late 2013, for example, a worker posted a thread to the MTurk subreddit with the title "Incompetent users breaking TurkOpticon, what now?" The 500-word post began:

> TurkOpticon is filling up with incompetent users. They don't know what good pay is, they don't know what is prompt, they don't know what is fair, and/or they rate every category based on ONE thing instead of on the actual categories. They've rendered the rating-at-a-glance feature to be completely unreliable; if you want to know whether a HIT is worth your time, you have to open up the Reviews page and look for someone who actually left a comment, and then decide if that reviewer is one of the people who has the same standards as you or if you have to disregard their rating because they're one of the ones that think

64

$1/1 hour deserves a 5/5 score. This takes time and causes aggravation, the very

things TurkOpticon was meant to save us.[66]

In view of these concerns, a series of discussions on the Turkopticon mailing list in spring 2014 raised the prospect of adding several more "objective," and in some cases quantitative, fields to the review form. "How long did the requester take to pay you?" was proposed to augment or replace the speed of pay rating. "Were you approved?" and "If not, do you think the rejection was fair?" were proposed to augment or replace the fairness rating. And a combination of the reward and the amount of time it took to do the task was proposed to augment or replace the generosity of pay rating. Just as discussion was blossoming into healthy debate, however—with some veteran users concerned about turning the review form into "bubble hell," increasing the time required to submit a review—a crisis intervened, and the discussion was tabled. As of February 2015, the review form is, despite an abundance of good ideas for improvements, unchanged since 2009. Discussion about possible changes persists; for example, in late February 2015, a Turker posted a new thread on turkopticon-discuss, the open Turkopticon mailing list, with the subject line "The pay rating isn't very useful if it is subjective." But the shortage is not in good ideas for change but programmer time—or, to put it more broadly, organizational capacity. Turkopticon's failure to make what may seem to outsiders like easy improvements is not lost on crowd work researchers. The computer scientist and prolific AMT requester Chris Callison-Burch, for

---

[66] See `redd.it/1t17w3`.

example, designed and commissioned an alternative extension, called "Crowd-Workers," that "aims to replace [Turkopticon's] qualitative ratings with quantitative equivalents."[67] If Crowd-Workers or a similar alternative extension were to gain significant traction among Turkers, it would reduce the amount of maintenance work for Lilly Irani, myself, and Turkopticon's volunteer moderators. Unfortunately this does not seem to have happened yet.

2.3.3.2 Identity: harassment, requester self-reviews, and trolling

Some workers use Turkopticon to encourage others to harass requesters. Sometimes the impetus for this incitement arises from a misunderstanding. For example, new academic requesters who do not want multiple responses from one worker for a survey often use an AMT feature called a "block" to prevent workers from completing their HITs more than once. (There *appears* to be multiple kinds of block, although workers argue about this.) If a requester blocks a worker, the worker may receive an automated email from Amazon informing them of the block and warning them that if they receive multiple blocks, their account may be suspended. This creates stress, especially for workers who rely on Turking income to meet basic needs. As a result, new requesters who do not realize they are doing anything "wrong" by blocking workers may be on the receiving end of significant worker anger. In October 2012, for example, one worker posted a review rating a requester 1/5 for fairness (and N/A for the other attributes), with the following explanation:

---

[67] Callison-Burch 2014.

DO NOT COMPLETE THESE HITS, THIS PERSON WILL BLOCK YOU AND GET YOUR

ACCOUNT SUSPENDED. BEWARE.

In May 2013, another worker posted a comment underneath the review ("XXX" indicates redacted content):

> Report this scammer. His website is registered through enom.com. The person it's registered to is XXXXXXXXX. Send complaints to enom. Then file complaints to their affiliate program. Their clixsense affiliate Id & account number is XXXXXXX. Their rewarding ways affiliate Id is XXXXXXXXX. Their inbox dollars & Jill's Click corner affiliate ID is refXXXXXXX. Their Reality-Networkers Id is XXXXXXX. And they are using paypal on their site. Paypal has this as their email address, XXXXXXX@XXXXXXXX.edu. Complain about their scam to paypal. If you search for the email address you will find [the requester's name]. An [name of university redacted] student. He also has a facebook page also.

The requester, a graduate student, replied in a comment:

> This post is defaming and insulting my name. I have done nothing this person has reported. I just came across this today. It is one thing to leave a review that is true, but it is another thing to name someone out, the reason he is mad is because it is a low paying hit. Please delete this comment. Thank you. I will work on increasing the pay on my hits.

The requester emailed us several days later. After an email exchange spanning a month in which the requester explained that he was receiving threatening emails and messages on social networking sites—sometimes dozens daily—and that he was working with his university to have his email address changed as a

result, I censored the comment with the requester's information, at that time an unprecedented exercise of administrative power.[68]

We have also received reports of workers attempting to blackmail requesters into paying for work both parties know is bad by threatening to leave bad Turkopticon reviews. We have received reports that sometimes these efforts are successful.

Sometimes requesters review themselves. This is possible because we have no mechanism for verifying that a Turkopticon user posting a review has worked for the requester they are reviewing, or that they are even a worker at all. A mechanism for linking a user's account on a separate service to a worker's AMT account was developed in 2014 by Niloufar Salehi and colleagues for the Dynamo platform.[69] Turkopticon users have expressed interest in adopting this model, but it has not yet been implemented. For now, Turkopticon's volunteer moderators—who are expert Turkers and members of multiple Turker communities—rely on information from the Turkopticon database, contextual clues such as a reviewer's email address and review history, and direct communication with reviewers to determine if a review was posted inappropriately by a requester masquerading as a worker.

Finally, Turkopticon is afflicted by the profanity, trolling, and general incivility common to any even modestly large web application that accepts "user generated content." At present, Turking uses a simple flagging system that lets

---

[68] The review and comments, including our email exchange with the requester, which I posted as justification for censoring the comment, can be viewed at `turkopticon.ucsd.edu/get_report/51544` (Turkopticon login required).
[69] Salehi et al. 2015.

users bring uncivil reviews to the attention of volunteer moderators. Although more sophisticated and effective techniques are well-known,[70] they are hard to implement well. As a result, for now, these issues are still handled by a combination of human attention—from users, volunteer moderators, and administrators—and simple automated filters.

2.3.4 Maintaining and evolving Turkopticon, 2008–2015

2.3.4.1 Origin story

In fall 2008, in a course taught by Beatriz da Costa in the Arts Computation Engineering program at the University of California, Irvine, Lilly Irani, relatively recently arrived from Silicon Valley, and troubled by the computing industry's excitement about just how *cheaply* AMT offered significant new operational capabilities to programmers, posed, through a HIT, a hypothetical question to Turkers: If you could write a Bill of Rights for Turkers, what would be in it?

The responses were lucid and sophisticated, contradicting a widespread notion in online technology industry discourse of Turkers as unskilled and lazy.[71] They were notable both for the concerns that appeared widely shared and for the handful of respondents who suggested that the question itself was illegitimate. 67 respondents agreed to have their responses made public; they can still be seen at `turkwork.differenceengines.com`. We reported our

---

[70] See e.g. Farmer and Glass 2010; Kou and Nardi 2013, 2014; and Hudson 2014—the last of which was emailed to me by a Turkopticon user. One researcher has even drawn on the work of Elinor Ostrom—featured prominently later in this dissertation—to offer guidance for managing online communities (Brewer 2015).

[71] See e.g. Howe 2008 (especially the illustration).

classification of these responses in papers published in 2010 and 2013[72] as follows:

- 35 workers felt that their work was regularly rejected unfairly or arbitrarily
- 26 workers wanted faster payment (Amazon allows employers 30-days to evaluate and pay for work)
- 7 explicitly mentioned a "minimum wage" or "minimum payment" per HIT
- 14 mentioned "fair" compensation generally
- 8 expressed dissatisfaction with employers' and Amazon's lack of response to their concerns

But this summary does not capture the lucidity, sophistication, diversity, and occasional passion of the responses. One worker, for example, wrote:

> If I could make the Mechanical Turk Bill of Rights, then I would make sure that workers were compensated for any work that they tried their best on, and such work could simply not be rejected. I realize that if someone obviously didn't take the task seriously, then, perhaps their work should be rejected, but if they honestly tried their best, then they should be rewarded for their time and effort. Because some requesters using Mechanical Turk simply aren't good individuals, as they will review one's work, reject it, provide an extremely lame excuse as to why, even though the person who submitted the HIT was right on par, just so that they (the requester, that is) can get their work done, without having to pay a penny. In my opinion, this is extremely immoral and should be frowned upon by those at Amazon. Secondly, I think that the reward that workers are offered should correspond with the work that the requesters are asking them to do. For

---

[72] Silberman et al. 2010, "Sellers' problems in human computation markets"; Irani and Silberman 2013.

instance, while some tasks offer a penny because they really are truly simple and not that time consuming, those are okay in my book. But, when a requester asks for work that is going to take a significant amount of time and requires a great deal of effort, a penny will simply not suffice! And, I would also change the amount of days that a HIT has to auto approve to 7 days instead of 30. Right now, if someone does a HIT and submits it, the requester has 30 days to review it and decide whether or not to approve or reject it. And if the requester doesn't look at the HIT during that 30 day period, it is automatically approved. However, 30 days is really too long of a wait, especially when someone is waiting for tasks to be approved that will only pay you less then a couple of dollars! Lastly, I would also like workers to have more of a say around here, so that they can not easily be taken advantage of, and are treated fairly, as they should be. Amazon seems to pay more credence to the requesters, simply ignoring the fact that without workers, nothing would be done! However, since some workers would be willing to do difficult tasks for pennies, other workers that know this is not a fair trade can't really make a statement. And this is why some type of organization or union needs to be formed to ensure that we are treated like they should be, and not like people working in third world country sweat shops![73]

Dissatisfaction over the lack of requester accountablity and a general sense of disempowerment were common themes. Even the most terse respondents focused on the rejection feature. While some respondents offered a bulleted list of rights or proposed changes, one wrote simply, "1. Right to dispute rejection."[74] Some respondents said they would not change AMT at all, praising the flexibility and freedom from features of institutionally organized work such

---

[73] Archived at `turkwork.differenceengines.com/blog/?p=172`.
[74] Archived at `turkwork.differenceengines.com/blog/?p=124`.

as fixed working hours, work assigned by bosses, waiting for paychecks, and tax withholding. One respondent in this category wrote:

> Exactly what is in the provider policy. If I don't like a requester, I don't do work for the requester. If I have a problem giving away the rights to my work for all eternity, then I don't give my work away. And if I manage to injure myself doing a HIT, I probably had it coming. We're independent contractors. It happens.[75]

While a few respondents argued this position, many concerns were broadly shared among respondents. In the context of da Costa's class, Lilly and I decided to collaborate on a system design that would explore how some of the shared concerns might be addressed. Our initial plan was to make interface mockups. But da Costa encouraged us to make a working system, and we agreed.

Our interpretation of the Bill of Rights responses shaped the initial design. Specifically, we designed it to let Turkers report their experiences with requesters in four categories that emerged from the Bill of Rights responses. The categories were (and still are) communicativity (abbreviated "comm" in the interface), generosity (of pay; abbreviated "pay"), fairness (regarding rejections; abbreviated "fair"), and promptness (of pay; abbreviated "fast"). We prompted users to score these attributes on a scale of 1 to 5, with 1 being "bad" and 5 being "good," with the following questions:[76]

- Communicativity: How responsive has this requester been to communications or concerns you have raised?

---

[75] Archived at `turkwork.differenceengines.com/blog/?p=111`.
[76] See `turkopticon.ucsd.edu/report` (Turkopticon login required).

- Generosity: How well has this requester paid for the amount of time their HITs take?

- Fairness: How fair has this requester been in approving or rejecting your work?

- Promptness: How promptly has this requester approved your work and paid?

We encouraged workers with no relevant experience regarding a particular attribute to leave that score blank in their report.

We decided to name the system "Turkopticon," in a cheeky reference to the circular prison, the panopticon, designed by the British philosopher Jeremy Bentham. The system let Turkers submit reports through a web form and add average Turkoption ratings to the AMT interface by installing a browser add-on. The add-on simply retrieved data from the database that stored the submitted reports. This design was inspired partly by a previous new media art project called "Pirates of the Amazon," a technologically simple but highly controversial browser add-on that added a "Download for free" link to pages on Amazon.com selling content that was available for free through the Pirate Bay BitTorrent tracker.[77]

We built the web database application and browser add-on in 48 hours. We hoped that by illuminating some of the challenges facing Turkers, and by showing how simply, at least from a technological point of view, they might be addressed, the managers and developers running AMT might be convinced to

---

[77] See e.g. Leingruber 2008; Katz 2008.

build a reputation system for requesters into their platform. Reputation systems were relatively well-known in e-commerce and online life generally, even in 2008. The technology news and community site Slashdot ran perhaps the most celebrated, although eBay's symmetrical seller and buyer feedback system supported more business and was perhaps more readily transplantable to a market like AMT. And Amazon itself ran several large reputation systems, not least in Amazon Marketplace, through which vendors large and small sold used and new books and consumer goods.

Pleased with our class project, which, we thought, made an important point in a reasonably interesting way—at least, we hoped, interesting enough to earn us satisfactory grades in da Costa's class—we left Turkopticon up on the internet, finished the quarter, and prepared to move on with our studies.

2.3.4.2 Initial design

In its initial design, Turkopticon had two parts: a web database application and a browser add-on. After making an account in the web application, for which we required only a working email address, workers could review a requester by filling out and submitting a form.

In the initial design, we called reviews "reports." This was because we imagined that Turkers would use Turkopticon to "report" "bad" requesters, not necessarily to "review" every requester. We understood Turkopticon as technically similar to reputation systems on sites such as eBay (where buyers and sellers review each other) and Amazon Marketplace (where buyers review

sellers), and, to a lesser extent, Yelp and Amazon.com (where buyers review restaurants and consumer goods, respectively). But unlike those sites, where every transaction, restaurant, or good purchased can reasonably be reviewed, we did not expect every requester needed to be reviewed. We imagined that there were more or less two kinds of requester: good ones and bad ones. Turkopticon, we imagined, would help Turkers warn each other about the bad ones.

In addition to not imagining that Turkers would want to review every requester, we also did not predict that different Turkers would have wildly differing experiences with the same requester, or that even Turkers with similar experiences with a particular requester would interpret and evaluate them differently, making reviews a site of contention, suspicion between workers, and even, sometimes, collective sensemaking. Rather, on the initial deployment of the site, we framed the use case for the system to Turkers with a simple question: "Are there HITs ["human intelligence tasks"; i.e., tasks] you weren't fairly paid for?" Thus the initial design of Turkopticon supported a relatively simple workflow. Turkers, we thought, would install the add-on and create an account on the web application. As they browsed for tasks through the AMT interface, the add-on would augment the interface with a drop-down with any information recorded by other Turkopticon users about the requesters whose tasks the user was looking at. The Turkopticon drop-down for a requester showed the requester's average ratings according to reports submitted by other Turkopticon users and the number of reports about the requester in the

Turkopticon database. It also included three links. One directed the user to explanations of the four attributes ("comm," "pay," "fair," and "fast"). Another let the user browse the full reports submitted for the requester in question on the Turkopticon web application. The third directed the user to the requester review form on the Turkopticon web application, and automatically filled the "requester name" and "requester ID" fields in the form by extracting that information from the AMT interface. On this listing, Turkers contributing reviews for a particular requester were identified by default by obfuscated email addresses (e.g., "sgur...@y..."; "devo...@g..."; "ang...@a..."), although Turkers could change these to a "display name." This allowed workers to identify which reviews were posted by the same user while safeguarding their email addresses.[78]

We built the web application in a web database application development framework called Ruby on Rails, often simply called "Rails."[79] Rails is an extremely "programmer-friendly" framework. It has many design conventions (i.e., assumptions) that allow development of a wide range of applications much more quickly than other frameworks allow, especially those requiring more extensive configuration. It is therefore a framework known for minimizing programmer time in the initial stages of application development. This power does not come free, however. There are at least three trade-offs. First, as is now well-known (although of course much-debated), in many execution

---

[78] See Irani and Silberman 2013, p. 617, for more discussion of this design decision.

[79] "Rails" is the framework; Ruby is the programming language it is written in, and in which developers using the framework write their applications.

environments Ruby is a relatively slow language compared to other languages popular for web application development such as Perl, Python, and PHP. This is not a problem in early stages of development when an application has few users. But as usage grows, it can make the application slow, even to the point of unusability, forcing developers to buy more computing hardware or optimize the application. Second, the extreme ease with which Rails allows developers to implement functionality can seduce developers, especially those (like us) with little experience operating applications with many users, to adopt (perhaps unwittingly) programming practices that make the application slow. For example, in Turkopticon's "report" template, the view makes quite a few database calls, making the application less efficient than it could be. These database calls are easy to program in Rails (e.g., `Person.find(session[:person_id])`) but computationally costly. In PHP development, in contrast (assuming one is not using a framework similar to Rails—and admittedly this has become less likely in recent years, as Rails-inspired frameworks in many languages have proliferated), one is forced to make a database call explicitly (e.g., with the `mysql_query` command) and process the results in several steps. This inconvenience encourages the developer to be make database calls more sparingly, making the application more computationally efficient and thus less prone to slowing as usage increases.[80] Third, the design conventions that make Rails powerful for a wide range of applications may make it difficult to support more intricate functionality which

---

[80] See for example Jari-lee Tolentino's 2014 reimplementation of the Turkopticon search feature, as described in Tolentino 2014, pp. 10-12.

comes to be desired as usage grows. This can be worked around by simply implementing such functionality in another language, but connected to the same database, but this increases the complexity of the application code and the cost of maintaining it; e.g., programmers competent in different languages and/or frameworks are needed, and must collaborate. We knew none of this at the time. We chose Ruby and Rails because they were the language and framework that, between the two of us, we had the most familiarity with. Indeed I had had almost no experience building web applications with any other language or framework.[81]

The initial design of the Turkopticon web application was fairly straightforward. We followed Rails' model-view-controller convention fairly strictly, not because we thought carefully about software architecture but, in fact, because we did not think carefully about software architecture, and that was the default choice.[82] There were three main models: `person`, `requester`, and `report`. There were two controllers beyond the default application controller: `reg` and `main`. `reg` handled user registration, login, and password management. `main` handled everything else, the bulk of which had to do with creating, reading, updating, and deleting requester reports. Consistent with a loose (or perhaps

---

[81] This was common among Rails programmers at the time, and may still be, as the framework is still quite popular in industry and Rails developers with even moderate experience continue to fetch salaries in the high five, or low six, figures. I continue, for example, to receive occasional emails from recruiters looking specifically for Rails developers; no other language or framework competence seems to be assumed or required. The ease of use of the Rails framework and the Ruby programming language significantly "lowered the bar" for entry into the world of [arguably] production-ready web database application development.

[82] The Turkopticon web application was not, however, nor is it today, "RESTful." Although it was possible to build RESTful Rails applications in 2008, REST was not yet the Rails convention it is today.

opportunistic) interpretation of "agile" software development, and with the limited goals of our class project, our main goal was not to develop long-lasting, maintainable software, but to get something working on the internet as quickly as possible—which turned out to mean "in one weekend."

We took a similar approach to the development of the browser add-on. In 2008, Mozilla Firefox was the dominant web browser. And it was the first to formally support browser add-ons (then called "extensions"). So we decided to make a Firefox add-on.[83] Neither of us knew how to develop "native" Firefox add-ons, and a first foray into the relevant literature suggested it was, in short, hard. We soon discovered, however, the world of user scripting. User scripts, like add-ons, are installed on a user's computer, "inside" the browser environment, and modify the behavior of the browser. Specifically, they modify pages on particular sites in particular preprogrammed ways. Unlike native add-ons, however, user scripts are written in JavaScript, and are often written as a single file. They are, in short, straightforward compared to the alternatives. And GreaseMonkey, at the time a popular extension for Firefox, allowed developers of user scripts to package them as standalone browser add-ons so that Firefox users without the GreaseMonkey extension, whose main purpose was to allow users to install and manage user scripts, could use them. User scripts can't do everything native add-ons can, but user scripting proved adequate for our

---

[83] A Chrome add-on for Turkopticon did not exist until 2010, when a user by the handle RadioactiveRaindeer ported the Firefox add-on to Chrome. We finally released an "official" Chrome add-on in 2012. By this time, Chrome had become the dominant browser both in general and among Turkopticon users, and we switched our priority, such as it was, to emphasize support of the Chrome add-on.

purposes. Neither of us knew JavaScript, but the task of the Turkopticon add-on was simple enough that we muddled through with reference to a wide range of online tutorials, dogged persistence, and caffeine. The first version of the Turkopticon add-on was inefficient, with bad architectural design and worse programming style, but it worked.[84] It retrieved average requester ratings from the Turkopticon web database application and injected the information into the AMT interface.

We showed the system in da Costa's class, where it met with mild interest.

2.3.4.3 Launch

Around the same time, Lilly found occasion to tell a college acquaintance of hers named Lukas Biewald about the project. Biewald was cofounder of a San Francisco-based crowdsourcing startup called, at that time, Dolores Labs. (As noted above, Dolores Labs later became CrowdFlower; see Sec. 2.2.4.) Dolores Labs made heavy use of AMT. Biewald, perhaps thinking that Turkers would rate his company favorably as a requester, giving them in his view a well-earned advantage over other requesters, or perhaps out of a general sense that information about requesters' habits should, in a transparent market, be available to workers, was enthusiastic about the project.

---

[84] Although it was rewritten in 2012, to this day, the Turkopticon add-on for both Chrome and Firefox remains a packaged user script—a single JavaScript file. The second version, still in service with a few modifications, is significantly more computationally efficient—or at least, less inefficient. The architecture has been somewhat improved. The matter of programming style is still, I must regretfully report, more or less beyond us.

On the evening of January 5, 2009, we emptied the Turkopticon database, for the last time, of test data, and emailed those Turkers who had expressed interest in using it to tell them it was ready. Not two full days later, Lilly emailed me:

> [Lukas Biewald, co-founder of Dolores Labs] just IMed me randomly today. I told him about turkopticon. And he got really excited about it. He wants to help us get feedback on requesters through turk tasks and dolores labs will pay. He said that he's wished for a long time that AMT was a "more open, honest environment." Now, what counts as open and honest is an interesting question to investigate. Markets are supposedly more optimal when open and honest, but there are likely limits to that especially when accumulation of power is taken into account.

Biewald, he told Lilly, had been talking with a former AMT employee. The former employee had suggested to Biewald that Dolores Labs build, in Biewald's words, "exactly the tool you [Irani] built"—a requester reputation system. Biewald had contacted Lilly looking for beta testers for a new tool Dolores Labs had developed to make it easier to post tasks to AMT. But he became excited when she told him about Turkopticon. "I think this is great," he said. "I would *love* [AMT] to be a more open, honest market place."

Biewald said he had been inviting prospective beta testers of Dolores Labs' new tool to post small HITs through it at Dolores Labs' expense in exchange for user feedback. Lilly took him up on this offer, devising a HIT that would help seed the Turkopticon review database, making it useful to workers. She collected the names of 53 prolific requesters from the worker forum Turker

Nation and from AMT itself. She then designed a survey that could be posted through Dolores Labs' tool. The survey design corresponded roughly to the initial design of Turkopticon's report form in that it asked respondents to give information about their experiences with requesters: how many tasks they had done for the requester, and how fair, generous, fast, and communicative the requester had been. But the survey format prompted a reevaluation of the design of the report form itself. Lilly wrote:

> While this survey format is currently not easily commensurable with the data schema we have planned for reports, it may be that this is actually a better schema because it lets people give aggregate reports for someone they wish to report, rather than making a report for every bad incident. This may scale better for the workers who do lots of HITs—those who stand to benefit the most from our tool.

This prediction proved prescient. Biewald posted the task, collecting 223 responses. We changed the database schema from an incident-based model to a requester-based model, limiting reviewers to one report per requester. On January 16, Biewald emailed us the responses as a CSV file. I fed them into the database on January 22. From two reports, we were up to 225. We had 13 users, only four of whom we knew personally. On January 24, the users who had signed up started posting their own reviews—at least a few a day during the work week. We designed and hastily implemented a rudimentary scheme for flagging and hiding reviews we, or workers, suspected of being posted by the requesters ostensibly being reviewed. We were underway. And at the bottom of Dolores Labs' many HITs, Biewald began to include a link to Turkopticon. This

went on for at least a few months. And so it was that we found ourselves, in spring of 2009, with two surprises.

First, Biewald invited us to speak at a "crowdsourcing meetup"—an afternoon-long series of talks by people involved in the young but burgeoning crowd work industry—at Dolores Labs. "Say whatever you want," he told Lilly. We assumed he was amused by the prospect that we could come and stir up, so to speak, the pot. We agreed to speak. The other speakers included Biewald; some natural language processing and machine learning researchers making heavy use of AMT; Alexander Sorokin, a computer vision researcher who had developed tools to let Turkers guide robots through physical space; and Sharon Chiarella, the Harvard MBA who led the Mechanical Turk division at Amazon. We gave a talk we titled "The 'sourced' crowd is made of people: what this means for what (and how) we build." We talked about worker exploitation, the responsibilities of employers, and how decentralization in the trucking industry had increased pressures on truckers and how we saw the same thing happening in AMT. We even wondered aloud whether Amazon should help lobby for universal health care, given AMT's reliance on freelance labor and the fact, so far as we could tell, that the wages people were earning through the platform were not adequate for most workers to pay independently for health insurance. Chiarella, so far as we could tell, ignored us. Leila Chirayath Janah, who would go on to found SamaSource, billed as an ethical crowdsourcing company employing full-time workers in low-income countries, and who would later keynote on these topics at at least a few technology entrepreneurship

conferences, asked Chiarella why AMT did not have profiles—full names of workers, photos, background information, and so on—to encourage requesters to think of them as people. She noted that Kiva, the microlending platform, made loan recipients' profiles available to prospective funders; why not do the same on AMT? Chiarella said explained that it was to avoid discrimination. AMT had quite a few workers from India and other countries outside the US; giving requesters information about workers might give them tools to discriminate against them. I found Janah's question timely, given the challenges we had identified facing workers. And I found Chiarella's reply unsatisfying: requesters could and did discriminate on the basis of nationality anyway, by screening out blocks of IP addresses.

But more surprising, and perhaps more immediately alarming for us, Turkers were *using* Turkopticon. By June 1, 2009, there were 122 users signed up. We had solicited none of them. Of these, 58 had collectively posted 769 reports of 407 requesters. And usage was increasing. Things were, in part, out of our hands—but we began to feel a responsibility we had not expected.

2.3.4.4 The everyday work of managing

Between 2009 and 2015, Turkopticon evolved as a sociotechnical system, changing in response to the everyday interpersonal, organizational, and technical demands of use; the development and uptake of new web technologies (especially new and backward-incompatible versions of Ruby and Rails); and the changing technical and institutional landscape of Turking life.

The everyday work of managing Turkopticon is extremely varied. We fix software bugs. For example, in late 2014, we realized that reviews that had been hidden because they violated the site's rules appeared in search results; this small bug had a simple fix and is representative of a class of small bugs that we fix as they are discovered by users.

We also fix what might be described as interaction bugs. For example, sometimes Turkers use the Turkopticon search feature to find a requester they want to review, rather than clicking the "Report your experience with this requester" link added by the extension to the AMT HIT list. But some requesters have the same names, so this sometimes leads to misplaced reviews. In 2014, at workers' request, we added a note to the search results page that appears if a user's search returns reviews associated with more than one requester. The note reads:

> Your search returned reviews of X different requesters. If you are leaving a
> review, please be sure to review the right one! Thank you!

where X is the appropriate number. This satisfied concerned workers and seems to have addressed the problem adequately for now.

We provide technical support in response to posts to turkopticon-discuss, the open Turkopticon mailing list, and in response to emails sent directly to the Turkopticon email address. For example, occasionally a user asks for us to set their display name after creating a new account; we reply with instructions about how to do so. Sometimes new users do not understand what "requesting commenting" means, how to do it, or whether they need to do it; we reply to

confused inquiries on this topic as well. And sometimes our automated emails (sent only to verify a user's email address) are bounced by user's mail servers. In these cases, users often email us manually to tell us about the problem. We verify these users' addresses manually and reply directly. These recurring technical or interactional hassles might be solvable once and for all with more enlightened design, programming effort, or outsourcing (e.g., to a dedicated mail-sending service)—and we sometimes take that approach. For example, sometimes requesters change the names associated with their AMT accounts, and ask us to change the corresponding information in the Turkopticon database. For a while, we did this manually, retaining the old name in parentheses to avoid worker confusion. Eventually, however, we made a collection of small changes to the database design, web application code, and browser extension that causes new reviews to incorporate a requester's new name after a name change. But in general, the current configuration, in which some processes are handled in software and some manually, though unarguably far from any reasonable definition of optimality, is acceptable given our time, money, expertise, and priorities.[85]

Sometimes changes to AMT or to browsers make Turkopticon hard to use, or break it entirely. In 2014, for example, Amazon made a very small change to the HIT list that broke Turkopticon completely for all users. Specifically, they added a span element around the names of requesters; for example, `CrowdSource` became:

---

[85] That is, like most information systems "in the wild," Turkopticon is a "heteromated system"; see Ekbia and Nardi 2014.

```
<span class="requesterIdentity">CrowdSource</span>
```

Turkopticon stopped appearing on the HIT list. Some workers reported that they had stopped working, writing that they didn't know which requesters to trust and that they felt like they were "flying blind." Anxious messages started to appear on the worker forums, on turkopticon-discuss, and in the Turkopticon inbox. We scrambled to figure out what had happened and issue a fix, which we did within a day. That such a seemingly trivial change could create so much anxiety was alarming, but there was little we could do other than adapt to the new design of the page and accept that something similar might happen again.

In a slower-moving catastrophe, both Firefox and Chrome slowly tightened their security restrictions for browser extensions over 2013-2014. To close a security vulnerability, both browsers eventually disabled "mixed active content" by default. "Mixed active content" occurs when a page served over HTTPS (i.e., a securely-delivered page) has content added to it that was retrieved over HTTP (i.e., unsecurely-delivered content), for example by a browser extension, user script, or server-side script. While Turkopticon data had been served securely in 2013, after we moved to a shared server at UCSD, we did not set up HTTPS again. As a result, by late 2014, Turkopticon users running Firefox had to either install a second extension ("Toggle Mixed Active Content") to re-enable mixed active content, or go deep into Firefox's configuration and manually change the setting. Chrome warned users that Turkopticon was "unsafe," and to get it to work, users had to (repeatedly) click a button that said

"Load unsafe script." Enabling the right settings in either browser was complex, and we didn't know how to advise users until we knew what browser they were using, what *version* of that browser, what version of Turkopticon they were using, what other extensions or scripts they had installed, and, in some cases, in what order those scripts were running. We gave technical support to these users over email and forums, getting users to help us debug their configurations by asking for screenshots and error messages from the browsers' developer consoles, and producing screenshots of our own to supplement textual instructions. When our hosts at UCSD configured HTTPS on a new dedicated server, we were able to issue a new version of the extension and tell users who were running into trouble to follow a simple process: uninstall all the old versions of Turkopticon and install the new one.

We also respond to confused, upset, and even, occasionally, very angry requesters. While most interactions we have with requesters are pleasant, the balance of difficult interactions we have had with requesters have not resulted from worker harassment of requesters but rather from requesters threatening us—and, on one occasion, following through with those threats. These threats arise out of differences of opinion about what is appropriate worker use of Turkopticon and what is not—and, specifically, what workers are allowed to say in reviews and what they are not, and to what standards of evidence claims in reviews must be held. Both of the most extreme cases, one of which led to a formal complaint filed with the UCSD IRB, interestingly enough, involved academic requesters. In the case that led to the IRB complaint, a requester's

reviews were appearing in a Google search for their name. The reviews were not on Turkopticon proper, which does not allow users to see reviews without logging in, and which therefore does not appear in Google search results for requester names, but on a mirror site that was for some time operated by a worker. This site scraped all the Turkopticon review pages and made them available for workers in the event that Turkopticon was down or unusably slow (which it occasionally was before we moved to UCSD hosting). But the mirror site did not require a login, so Google crawled it, causing Turkopticon reviews to appear in Google search results. Another time, perhaps as a result of the same configuration, a French requester—or perhaps a lawyer unaffiliated with the requester; it was not clear—sent us a copyright takedown notice (in French) about a review in which a worker had pasted some text from a task. The text in the review, the notice claimed, included names of the sender's clients, which were protected under French intellectual property law, the full force of which would be brought to bear on us if we did not censor the review. It seemed to us that the requester had hired workers to perform a complex search engine optimization task in order to bury unfavorable press coverage about a sex scandal involving a French executive, but this salacious interpretation is admittedly a slightly informed guess in a muddy situation. After some worried consultation with a law professor colleague, we did nothing.

Dust-ups like these with angry or threatening requesters (or third parties) are stressful, but typically end when requesters' expectations that they will get whatever they demand are adjusted. But some situations arise from the

frustrations entailed by the work on both sides of the market. In one exchange, for example, a worker posted a comment on a review about a requester. The comment included a criticism of the requester's policies and claims that the requester used work for which they did not pay. The comment also included the WHOIS information for the requester's domain name, which included the founder's name and home address. At that time, we had no explicit policy against posting personal information if it was "publicly available," which we interpreted as "googleable." An employee of the company, who was responsible for interacting with workers, emailed us to ask if we could censor the information. After consultation with the moderators, we declined to do so, replying that although the disclosure of personal information was regrettable, the underlying issue seemed to be the requester's work policies and attitude toward workers. The employee disagreed, arguing that the worker who had posted the comment had a personal vendetta against the company. I sympthasized with both the angry Turker and the distressed employee. But after exchanging a few lengthy emails, I despaired of being able to do anything substantive and failed to respond.

A similar stalemate sometimes ensues when workers disagree about proper uses of the site. Such disagreements typically arise over the interpretation of the review categories (generosity, fairness, promptness, and communicativity) rather than over text comments, because workers rely on the aggregated numerical ratings to make quick decisions while browsing for HITs. When they see other reviewers giving ratings that don't reflect their own

understandings, they feel that they cannot trust the aggregate ratings, and heated discussion often follows. We have in general stopped trying to engage substantively with these conversations, beyond reminding participants that this is a problem with the review form, not with other people, and that in the future we hope to have time to make the categories less debatable (see Sec. 2.3.3.1).

Workers also sometimes get angry at the volunteer moderators. This happens especially when a worker posts a review that is then hidden by a moderator. One frustrated new reviewer, in a heartfelt but quite scathing critique, even characterized what I subsequently called the "hierarchy of powers" unfortunately necessitated by the moderation system as a "caste system." We do our best to reply to these messages, explaining, sometimes at length, how things have come to be how they are, our understandings of the problems with current arrangements, and our hopes for future improvements, and encouraging workers to contribute to discussions about the future of the system on turkopticon-discuss if they have time to do so.

We also occasionally have time to build substantive new features to support particular workflows. The most recent such development was invisible to most users; it was a new interface for moderators to support the somewhat intricate moderation process.

On a daily, or at least three-times-weekly, basis, we review requests for commenting, an ability which, once enabled for all users by default, is now reserved only for those who have "manually" requested it and been "manually" approved. We disabled commenting by default after a spate of trolling in which

a few users took advantage of the fact that Turkopticon has no mechanism to flag or hide comments to post racist and sexist comments and insult other users.

We also communicate as needed with moderators and other users about general policies and particular problems. And once in a while, we "manually" edit comments.

2.3.4.5 Interdisciplinary difficulties

Two episodes from Turkopticon's history involve economists. In the first, in 2010, before Turkopticon had become critical infrastructure for Turkers, an economist persuaded us that what Turkers really wanted to know was expected wages for each HIT. The economist added (a lot of) code to the Turkopticon extension to automatically scrape earnings, rejection, and HIT completion time data from workers' dashboards, and prepared a server to aggregate this data and make it available. We added this code to the extension, however, without consulting workers. This was a mistake. Workers eventually understood the new code, and became very uncomfortable about what it was doing. We all apologized on the worker forums, removed the code, and restored the old version of the extension.

In the second, in 2014, we emailed back and forth with an economist who was interested in doing a study on Turkopticon itself. We discussed possible study designs at great length over email, establishing that Turkopticon was not available for experimental manipulation, but that aggregated data were already

available through the Turkopticon API and that, if Turkers did not object, we would be pleased to make disaggregated anonymized data from the review database available. The discussion eventually flagged, and we did not hear from the economist for some time. In July 2014, however, in an at first seemingly unrelated incident, workers began to report what appeared to be a large undisclosed experiment underway. Somebody had created dozens of requester accounts on AMT, and dozens more reviewer accounts on Turkopticon, and seeded the various requester accounts—which were posting identical tasks— with different kinds of reviews (good and bad). Some workers who had realized what was happening presumed that the entire Turkopticon system was under malicious attack. Some reported that they stopped working, not knowing if they could trust any of the reviews; the effect on these workers was as if the system was entirely down. After some investigation, the research team of the economist with whom we had corresponded turned out to be behind both the AMT requester accounts and the Turkopticon reviewer accounts. We discovered that the intention behind the experiment was to see how good reviews, bad reviews, and no reviews variously affected a requester's ability to get work done. The experiment was a success in terms of producing "objective" economic knowledge,[86] but workers were extremely frustrated. They had spent hours that they could have spent working investigating the requesters and reviews, and collectively producing comprehensive lists of the requesters that appeared to be part of the experiment. After the researchers posted their explanations to

---

[86] See Benson et al. 2014.

turkopticon-discuss, some workers argued that the research finding—that requesters with good reviews get work done more quickly than requesters with no reviews or bad reviews—was unsurprising but that the experiment was invalid as research, either because it had been discovered while underway or because of what they saw as an ethical breach. Two Turkopticon moderators filed formal complaints with the IRB at the researchers' home institution, and lengthy discussions over multiple media ensued. The IRB had approved the experiment, and took no formal action that we are aware of, but the researchers apologized for the frustration the experiment caused workers. While the researchers intended to produce knowledge to help workers, they did not anticipate the consequences of their approach. I discuss possible origins of this misunderstanding, and possible alternative approaches, in Chapter 4.

2.4   The broader Turker ecology of practice

Turkopticon is part of a rich and complex ecology of technologies, communities, relationships, and practices Turkers use to get work done, share information, and support one another.

According to Turkers, the most important parts of this ecology are the forums. There are several worker forums: the main ones, in no particular order, are Turker Nation, Mturkgrind, the MTurk and HITsWorthTurkingFor subreddits, and mTurk Forum. Each forum has different policies, politics, histories, and personalities—and differences between them sometimes erupt into inter-forum "drama." (Sometimes this "drama" appears on Turkopticon and

turkopticon-discuss.) But many experienced Turkers acknowledge the centrality of forums to earning money on AMT, especially for those "Turking for a living." Discovering a forum appears to be a crucial turning point in the "careers" of financially successful Turkers; it is typically only by connecting to a community of more experienced workers that one can navigate AMT well enough to earn significant income. Forum discussions are relatively open and unstructured. They include both specific information-sharing about tasks, requesters, and processes and more open-ended discussions in which newer workers "learn the ropes" and connect informally with others to build trust and community. For example, the Mturkgrind forum hosts a thread about "Mechanical Turk software beyond [user] scripts", and another thread with resources and support for Turkers struggling with anxiety and depression. Turker Nation hosts a famous "Requesters Hall of Fame/Shame" subforum—the original dedicated resource, predating Turkopticon, for explicit discussion of workers' experiences with requesters. Mturkgrind and mTurk Forum both host public "daily" threads in which workers post information about live HITs. These threads are very active, with some accumulating over a thousand posts in a 24-hour period. And many of the forums have associated IRC channels for real-time communication.

Requesters also communicate with workers on forums. Requesters can build trust with workers by communicating publicly on worker forums, and can share information and get feedback about their tasks and processes. Crucially, communicating quickly with workers on forums can save requesters from losing previously earned goodwill in the event of technical problems that result

in accidental rejections or wasted time for workers. Turker Nation, Mturkgrind, and mTurk Forum all have dedicated subforums where requesters are invited to introduce themselves and recruit workers to their HITs.

Forum discussions are complemented by highly structured information sharing practices supported by specialized software. Most special-purpose Turking software used heavily by Turkers is built and maintained by Turkers. (I know of only two exceptions to this pattern: Turkopticon and Dynamo, discussed below.) Most of this software takes the form of user scripts. For example, one popular script, "HIT Scraper," scrapes as many pages of the AMT HIT list as the user specifies, applies filters specified by the user, and displays the information concisely (see Fig. 2.12). The script lets the user omit HITs posted by requesters the user specifies (a feature perhaps inspired by another script, "Block Requester," which hides HITs from requesters the worker doesn't want to work for from the HIT list), or show only HITs posted by specific requesters. The script can also programmatically generate a piece of vBulletin code that can be pasted into a forum post. The code lists the HIT title, requester name and ID (including a link to Turkopticon reviews), Turkopticon aggregate ratings, number of Turkopticon reviews, link to submit a Turkopticon review, HIT description, allotted time for the HIT, number of HITs available, reward, and qualifications required—all in a visually clear and concise layout. Hundreds of posts including programmatically generated information about HITs, often supplemented with the posting worker's comments, populate the daily threads on both mTurk Forum and Mturkgrind. Other scripts allow

workers to see the total value of their HITs pending approval ("Today's Projected Earnings"), see more information about HITs, including the auto approval time ("Enhanced HIT Information Capsule"), and keep track of HITs they've done ("Mturk database"). Several browser extensions ("Turk Assist", "Tools for Amazon's Mechanical Turk") attempt to integrate many of the most crucial features of multiple scripts into a single package.

While researchers and other requesters have made some software for Turkers, including OpenTurk ("manage your favorite requesters, share HITs you like and work on HITs that other workers have liked, and track your earnings"), Crowd-Workers.com ("make it easier for everyone to find good work and make a fair wage"), metacrowd ("find interesting tasks by generating recommendations based on your individual preferences and qualifications"), and Microtask Liberation Front ("worker-friendly task recommendation"). But to my knowledge few of these tools are widely used. Of the tools built for Turkers by non-Turkers, the two with which Turkers have engaged vigorously are Turkopticon and Dynamo, a platform for facilitating Turker collective action. In 2014, workers used Dynamo to organize two unprecedented collective action efforts. First, they collectively drafted a set of guidelines for academic requesters. This extremely detailed document addresses many questions and practices relevant to researchers collecting data through AMT, and is intended to eventually become a reference for academic Institutional Review Boards reviewing research proposals using AMT.[87] Second, they organized a Christmas letter writing

---

[87] guidelines.wearedynamo.org; see also Salehi et al. 2015.

campaign to Amazon CEO Jeff Bezos. These letters did not call for specific changes to AMT but rather served to establish a public human face for Turkers. The campaign received significant media attention. "I am a human being, not an algorithm" was the quotation used to describe the main message of the campaign.[88]

Experienced Turkers who Turk for a living also often do other online work, and may even parlay their expertise and relationships into more stable work. For example, some Turkers have become remote research assistants or consultants for academic requesters; others have become intermediaries, advising requesters on HIT design and posting HITs for non-US requesters frozen out by changes to AMT's legal and financial requirements.

Requester	Title	Reward	HITs Available	TO pay	Accept HIT	
Sara Spencer R T	Find the Contact Info for the Owner of Some Websites vB	$0.05	3	4.09	Accept	R T
p9r R T	Transcribe data vB	$0.03	72	2.85	Accept	R T
p9r R T	Transcribe data vB	$0.01	5	2.85	Accept	R T
Mechanical Turk R T	Transcribe the characters in the images (50 images) vB	$0.01	2	1.00	Accept	R T
AEI R T	Find the link and image that a forum thread is discussing vB	$0.05	5	3.40	Accept	R T
Jimmy Dragon R T	Business Card Transcription vB	$0.02	3	2.67	Accept	R T
p9r R T	Transcribe data vB	$0.04	4	2.85	Accept	R T
p9r R T	Transcribe data vB	$0.09	17	2.85	Accept	R T
AML R T	To be eligible, you are 18 years or older, and currently working / employed. vB	$0.35	1	2.73	Accept	R T
Steve Murch R T	Review transcription quality vB	$0.02	1	3.34	Accept	R T
SpeechInk R T	MTurk payment authorization vB	$0.00	3	2.91	Accept	R T
Steve Murch R T	Review transcription quality vB	$0.03	3	3.34	Accept	R T
Vedrona Group LLC R T	Quick survey vB	$0.05	1	2.00	Accept	R T
CVRS R T	5 Minute Survey For Male Ohio Residents vB	$0.75	1	1.95	Accept	R T
Terry2094014 R T	Command to an autonomous car vB	$0.04	6	3.00	Accept	R T
Terry2094014 R T	Describe POI vB	$0.05	7	3.00	Accept	R T
CastingWords R T	Audio Transcription: George King George #193921 (avg rwrd+bns: $0.88) [05:13 mmss] vB	$0.44	1	2.89	Accept	R T
CastingWords R T	Audio Transcription: George King George #193921 (avg rwrd+bns: $2.26) [13:15 mmss] vB	$1.13	2	2.89	Accept	R T
CastingWords R T	Audio Transcription: George King George #193921 (avg rwrd+bns: $1.4) [08:15 mmss] vB	$0.70	1	2.89	Accept	R T
CastingWords R T	Audio Transcription: George King George #193921 (avg rwrd+bns: $0.9) [05:15 mmss] vB	$0.45	2	2.89	Accept	R T
CastingWords R T	Audio Transcription: George King George #193921 (avg rwrd+bns: $3.62) [21:15 mmss] vB	$1.81	1	2.89	Accept	R T
Harlan Beverly R T	Transcript recipe vB	$0.20	10	1.83	Accept	R T
CastingWords R T	Audio Transcription: George X-ray #193903 (avg rwrd+bns: $3.58) [21:02 mmss] vB	$1.79	1	2.89	Accept	R T
CastingWords R T	Audio Transcription: George X-ray #193895 (125% premium) (avg rwrd+bns: $1.12) [05:15 mmss] vB	$0.56	1	2.89	Accept	R T
CastingWords R T	Audio Transcription: George Boston Boston #193916 (avg rwrd+bns: $2.26) [13:15 mmss] vB	$1.13	1	2.89	Accept	R T
Steve Murch R T	Transcribe a recipe vB	$0.30	7	3.34	Accept	R T

Figure 2.12. Results from the HIT Scraper script, built by Turkers.

---

[88] See Harris 2014.

2.5   Requester ecologies of practice

There are two major requester ecologies of practice: one inhabited mainly by academic requesters and one inhabited mainly by industry requesters. They are very much interlinked, but academic and industry requesters have different responsibilities and work within different institutional cultures. Thus the practices of academic and industry requesters, and the resources developed to support and explain them, differ somewhat. And researchers in different academic disciplines post different kinds of tasks to AMT; social scientists often post surveys and run experiments, while computer scientists often post tasks that incorporate workers into an algorithmic process, for example in computer vision or machine learning applications. These differences lead to different requirements: social scientists, for example, are often keen to ensure that they are able to prevent workers from participating in a particular survey or experiment more than once, while the information processing work posted by computer scientists can often be done, at least in theory, by "any human." To the extent that this is the case, computer science researchers posting "human computation" tasks to AMT may sometimes "look" more like industry requesters than like social science researchers.

Like Turkers, requesters teach one another, share information, and develop and circulate specialized software. But requester discourse circulates in a broader variety of media: forums, blogs, meetups, workshops, conferences, and peer-reviewed papers. Most software developed by requesters is for requesters, not Turkers, and as many requesters work in the information

technology industry, many requesters are programmers or have relatively easy access to programming expertise. As a result, there is a great deal of specialized software available for requesters.

The differences between academic and industry institutional cultures appears most markedly in the distribution of information about management techniques—especially approaches to quality control—and in the distribution of software itself. Crowd work researchers and academic requesters openly discuss quality control techniques in blogs, academic conferences, and journals, the contents of which are sometimes free online. Some industry requesters participate in these discussions, or discuss their approaches and techniques in their own blogs, but it appears that many do not. Similarly, most software written for academic requesters is free or open source, but there appears to be few free software packages distributed for industry requesters. Many prolific requesters in both sectors write their own specialized software, but it appears that few industry requesters share this work with others. Most free or open source software for requesters aims to aid with workflow management and data analysis; for example, tools exist to prevent workers from retaking surveys (e.g., "TurkGate"), integrate AMT more closely with other online survey tools such as Qualtrics ("QualTurk"), link statistical packages such as R to the AMT API (e.g., "MTurkR"), post and manage iterative tasks ("TurKit"), and manage the complex process of running behavioral experiments on AMT (e.g., "PsiTurk"). The most well-known free or open source tool for industry requesters may be

Clockwork Raven, a toolkit for posting human judgment tasks such as sentiment analysis made by developers at Twitter.

At least as important as the auxiliary software they produce for other researchers, researchers in a wide range of disciplines also publish methodological papers and run workshops explaining how to use AMT for research.[89] This both spreads practical knowledge about how to use AMT to do research and socially legitimates the practice. In the overlapping human computation and human-centered computing literatures, a large and diverse body of work exists on quality control for crowd work generally, with some of this work focusing specifically on AMT.[90] Another considers workers mainly for the purpose of understanding their motivations for participation—and, as noted in Chapter 1, for designing more effective incentives for eliciting good, quick, and cheap work.[91]

2.6   Representations, connections, and interpretations

2.6.1   Mechanical Turk in context

What is AMT? What does it do—beyond link workers and requesters? What does it mean? And what, if anything, should be done about it?

Practitioner, scholarly, and popular portrayals and interpretations of AMT have been diverse. In its early years, practitioner and researcher discourse

---

[89] For example, in human-computer interaction, see Kittur et al. 2008 and Heer et al. 2010; in behavioral research, Paolacci et al. 2010, Buhrmeister et al. 2011, Mason and Suri 2011, and Paolacci and Chandler 2014.

[90] See footnote 10, Chapter 1.

[91] See footnote 11, Chapter 1.

focused on the new technological capabilities the platform offered end users.[92]

This discourse allowed AMT to be seen as the next step in both artificial intelligence and cloud computing. As "artificial artificial intelligence," AMT could be seen as having finally achieved the historical goals of artificial intelligence. This was done not by smarter algorithm design or deep insight into the nature of cognition, but by offering financial incentives to real humans that persuaded them to behave as if they were computers. And as a platform offering "humans as a service" or "labor as a service," AMT could be seen as the next step in the evolution of cloud computing, which had already developed "software as a service," "platform[s] as a service," and computing "infrastructure as a service". Phrases such as "the human API" and "remote person call" (by analogy to "remote process call") developed by end users fit well within this frame. Understanding AMT as "the human API" situated AMT within the constellation of other technologies, often accessible via APIs, that software developers used in their daily work, including Amazon's other web services. Such services are assumed by developers to be relatively reliably functioning and available without their having to acquire a detailed understanding of how they work. [93]

By 2009, however, critics from outside the technology industry began to disrupt this understanding. Some called AMT a "digital sweatshop," noting that the deskilling, routinizing, and breaking up of formerly intellectually complex

---

[92] See e.g. Hammond 2005, Ipeirotis 2007, Kittur et al. 2008, Sorokin and Forsyth 2008, Callison-Burch 2009, and Heer et al. 2010.

[93] See Irani 2013, esp. pp. 4-9.

tasks common on AMT recalled the piece-work system in the industrial economy. [94] The piece-work system turned workers into alienated and fungible drudges—exactly the condition many optimists hoped a new "knowledge economy" would free workers from. And by 2010, even computer scientists— including at least one evangelist for the technology—noted with some concern that end users routinely admitted that they used AMT to get around minimum wage laws. Economically-inclined computing scholars debated whether wages on AMT were low compared to similar work organized offline because of the simple dynamics of supply and demand—specifically, because there were many workers willing to work for low wages and not much work—or because of shortcomings in the design of the market itself. [95] Several observers asked if AMT was a "market for lemons."[96] This phrase names a dynamic first identified by the economist George Akerlof in considering the market for used cars.[97] In Akerlof's thought experiment, prospective buyers of used cars do not know which used cars are good and which are bad (i.e., which are "lemons"). While the seller knows the truth, from the buyer's perspective, any given car is equally likely to be a good car or a lemon. As a result, the buyer is only willing to pay a price that is halfway between what they would be willing to pay for a known good car and what they would be willing to pay for a known lemon. Sellers of lemons are of course happy with this situation. But sellers of good cars,

---

[94] See e.g. Zittrain 2009, Cushing 2012, and Uddin 2012.
[95] See Von Ahn 2010 and Ipeirotis 2010, "Mechanical Turk, low wages, and the market for lemons," and the comments responding to both posts.
[96] See e.g. Ipeirotis 2010, "Mechanical Turk, low wages, and the market for lemons," and Bederson and Quinn 2011.
[97] Akerlof 1970.

knowing that their offerings are worth more than the prices they are able to get, are not. As a result, sellers of good cars leave the market. As a result of this, however, the average quality of cars in the market goes down further, depressing buyers' expectations—and offered prices—further, and in turn driving more sellers of good cars out of the market. The end result is a downward spiral of quality and prices. This dynamic is one reason for the importance of effective reputation systems in markets for goods and services whose quality is important to buyers but not transparently obvious before purchase. In AMT, as noted above, the lack of sophisticated worker reputation creates a need for requesters to devise complex quality control systems. And the lack of any requester reputation at all exposes workers to a wide range of risks, including wage theft.

By 2011, the view of AMT as a source of fast and cheap, if imperfect, information processing, effectively analogous to a vast computer, began to be complicated among the computing researchers to whom such a view was substantively enabling. The computational linguists Karën Fort, Gilles Adda, and Kevin Bretonnel Cohen, for example, argued that the growing use of AMT among computational linguists had not been accompanied by careful consideration of the relevant ethical questions. Because, they observed, most tasks on AMT are done by a small fraction of workers, some of whom use the income earned to make ends meet, the then-widespread notion that the tasks performed on AMT did not constitute "real work" was mistaken. Rather, they argued, AMT is an unregulated labor marketplace for low-wage work. They

argued that the low wages were not due to the simple dynamics of supply and demand—i.e., to there being too many workers for the work available. In contrast, they argued, reports from other computing researchers—about the difficulty of finding workers with particular abilities or getting large batches of tasks completed—showed that there are not enough workers for the work available—and that prices were low because of the "market for lemons" dynamic.[98]

In the context of the emergence, in the period roughly 2011-2014, of what has been variously called the "gig economy," the "sharing economy," and—more recently and perhaps more accurately—the "on-demand economy,"[99] AMT has been reinterpreted as an early member of a new category of platforms. Many of these platforms link prospective buyers and sellers of in-person services, expanding the scope of computationally-mediated labor markets beyond information work to create local markets for transportation (Uber, Lyft, Sidecar), domestic work (Handy), and miscellaneous labor (TaskRabbit). Apparently motivated by slim opporunities in the formal economy, workers have flocked to these platforms. End users (i.e., customers) like them because they offer services at greater convenience—and sometimes reduced cost—compared to traditional means. And this conjunction has led to extremely large valuations for the companies that operate the platforms. But critics and

---

[98] Fort et al. 2011.

[99] See e.g. Kessler 2014, 2015; *Economist* Staff 2015, "There's an app for that" and "Workers on tap."

concerns are multiplying steadily as the potentially vast implications of this new way of organizing work come into view.

In January 2015, for example, New York Times technology columnist Farhad Manjoo wrote that the "app-driven labor market" represented by Uber, Lyft, and other app-based ride-"sharing" market platforms heralds a shift in focus in Silicon Valley, from building software to process information to building systems that "efficiently allocate human beings and their possessions." Uber's model, Manjoo noted, has been copied by startups aiming to create labor markets in fields from grocery shopping to legal services and health care. Manjoo discusses the main benefit of "Uberization"—increased "flexibility"— and its risks, including less predictable income and work hours, reduced employment security, low pay, and the difficulty of establishing a career. But he writes that "the larger worry about on-demand is not about benefits but about a lack of agency—a future in which computers, rather than humans, determine what you do, when and for how much."[100] He quotes business school professor Arun Sundararajan and political economist Robert Reich, portraying the former as an optimist about "Uberization" and the latter as a critic:

> "These services are successful because they are tapping into people's available time more efficiently," Dr. Sundararajan said. "You could say that people are monetizing their own downtime."
>
> Think about that for a second; isn't "monetizing downtime" a hellish vision of the future of work?

---

[100] Manjoo 2015.

> "I'm glad if people like working for Uber, but those subjective feelings have got to be understood in the context of there being very few alternatives," Dr. Reich said. "Can you imagine if this turns into a Mechanical Turk economy, where everyone is doing piecework at all odd hours, and no one knows when the next job will come, and how much it will pay? What kind of private lives can we possibly have, what kind of relationships, what kind of families?"

In an op-ed published by *Salon*, Reich writes that Sundararajan "confuses 'downtime' with the time people normally reserve for the rest of their lives." Online labor market supporters, he argues, are missing the bigger picture:

> Other proponents of on-demand work point to studies, such as one recently commissioned by Uber, showing Uber's on-demand workers to be "happy." But how many of them would be happier with a good-paying job offering regular hours? An opportunity to make some extra bucks can seem mighty attractive in an economy whose median wage has been stagnant for thirty years and almost all of whose economic gains have been going to the top. That doesn't make the opportunity a great deal. It only shows how bad a deal most working people have otherwise been getting.[101]

And in contrast to Manjoo's framing of "the larger worry," Rebecca Smith, deputy director of the nonprofit National Employment Law Project, writing in an op-ed published by CNN, argues that the new online labor markets are—at least in terms of business structure—nothing new:

> In the new on-demand economy, companies are turning the Internet into the equivalent of a street corner hiring site and turning workers into day laborers. [...] We have [...] policy solutions close at hand, because we have seen these structures before. The new platforms operate just like the traditional labor

---

[101] Reich 2015.

brokers in the agriculture and garment industries, who recruit workers, match them with jobs, and make a tidy profit for themselves in the bargain. Why shouldn't Uber, Lyft, and their kin be required, just like other labor brokers are, to register, pay their workers minimum wage and overtime, limit the commissions they are allowed to keep from each job they arrange and pay the payroll taxes that ensure workers have access to basic benefits like workers' compensation when they are injured and Social Security when they retire?[102]

These criticisms are not idle talk: in the last few years, Uber, Lyft, and Handy have all joined crowd work intermediary CrowdFlower as defendants in employee misclassification lawsuits.[103]

2.6.2 Broader trends

New online labor markets—for both remote and in-person work—have been built in the context of long-running dynamics of economic globalization (including outsourcing), "flexibilization" of work and work relationships, and the dismantling of the institutions of the welfare state, especially in the United States.[104] For decades, these processes have been both enabled and driven by new information and communication technologies. The new online labor markets are merely recent developments in a long-established pattern of replacing costly work processes involving context-specific knowledge—often performed within the firm by trained and well-compensated employees—with formalized, automated or computationally-mediated or -supervised, routinized,

---

[102] Smith 2015.
[103] See Swan 2013, Montgomery 2014, Huet 2014, Kosoff 2015, Levine and McBride 2015, and uberlawsuit.com.
[104] See e.g. Stone 2006.

decontextualized, lower-cost, and/or outsourced workers and processes. While it was once envisioned (or feared) that many jobs would be destroyed by automation, the shortcomings of automation and artificial intelligence have led to a new generation of complex sociotechnical systems and hybrid work processes that are "heteromated"—i.e., that incorporate humans into a process conceived and managed as largely algorithmic.[105] The situation in which workers are managed largely by automated means rather than by human managers can be described as automatic or algorithmic management,[106] a special case of algorithmic authority.[107] Algorithmic management is not limited to online labor markets; is has been used, for example, by companies such as FedEx and Amazon to control truck drivers and warehouse workers. These techniques are connected to other developments in management strategy driven by globalization and flexibilization. FedEx, for example, claimed that its truck drivers were independent contractors, but the drivers won a class action lawsuit the claimed otherwise.[108] Amazon's warehouse workers are considered contractors and are managed by a temporary staffing agency. They receive minimum wage but are not entitled to benefits. And they may be dismissed without warning if they are sick and miss a day of work, or do not meet any of the computationally-monitored performance criteria on which they are constantly algorithmically evaluated.[109]

---

[105] Ekbia and Nardi 2014.
[106] See Irani and Silberman 2013, Irani 2013, and Lee et al. 2015.
[107] Lustig and Nardi 2015.
[108] See Rooney 2014.
[109] See O'Connor 2013.

Online labor markets first came to prominence in 2010-2014, during the "Great Recession" following the financial crisis of 2007-2010. While the US economy is widely said to have "recovered"—by mid-2014, all 8.7 million jobs lost during the recession had been recovered—many of the jobs lost were higher-paying than the jobs created during the recovery, a disproportionate number of which were low-wage service jobs in restaurants and hotels.[110] And many of the new jobs created were in the temporary staffing industry:

> Unlike previous recessions, during which employers hired temporary workers to bridge gaps created by layoffs and transitioned back to hiring full-time workers during recovery, employers during the Great Recession picked up temporary employees, but have yet to revert to hiring similar numbers of full-time employees.[111]

The labor historian Jefferson Cowie summarizes the increasingly popular practice of hiring independent contractors instead of employees—and the relevant defenses and criticisms—concisely:

> For some workers, being a 1099'er [i.e., an independent contractor] means more flexibility, creativity and control over their work. However, there are many more reluctant 1099 workers ["named after the tax form provided to independent contractors" instead of the W-2] who want regular jobs but find themselves locked out of the system by employers looking for an easy way to buck their responsibility to their employees.[112]

This American trend is complemented by the emergence of "zero-hours" contracts in the United Kingdom. Zero-hours contracts allow employers to hire

---

[110] Puzzanghera 2014.
[111] Dill 2014; see also National Employment Law Project 2014.
[112] Cowie 2013.

hourly-waged employees without any guarantee of work. Employers can call on employees to work whenever they are needed, often on short notice. Some British companies have 80% or more of employees on zero-hours contracts. The practice has been called exploitative by critics, who note that many zero-hours employees are often not given enough hours to make ends meet, and defended by employers under the rubric of "flexibility" for both employers and employees.[113]

These developments—and indeed the financial crisis and ensuing recession—have arisen in the context of a decades-old structural misalignment between the incentives of corporate management and broad social and economic well-being. With the rise in the 1980s of the doctrine of shareholder value maximization—the notion that the task of corporate management is to maximize value to shareholders (i.e., owners), and therefore to maximize corporate profits—previously accepted corporate responsibilities to a wide range of stakeholders including customers, employees, suppliers, and the national interest were jettisoned. Corporate managers came under pressure to abandon the tradition of balancing sometimes conflicting and subjective interests of various stakeholders in favor of a single-minded and quantitative focus on profits. Managers whose preexisting ties to other stakeholder groups were pressured by shareholders or relieved of their duties, and stock options were included in managers' compensation packages to directly incentivize them to undertake initiatives that would lead to increases in share price, aligning

---

[113] See BBC News Business 2014, Roberts 2014, Guardian Editorial Board 2015.

their interests with shareholders'.[114] And Masters of Business Administration curricula in business schools began to organize around this new orthodoxy, professionalizing decades' worth of new corporate managers into shareholder value maximization as a natural fact of economic life.[115] The alignment of management's interests with shareholders' combined with technologically-enabled economic globalization to put immense downward pressure on wages, leading to successive waves of outsourcing, declining power for organized labor, shorter employment tenure and higher turnover, declining incentives for management to invest in employee training, and more transactional relationships between organizations and employees generally.[116]

This dynamic persists in the post-"recovery" period after the Great Recession in the US, with training budgets significantly reduced even from the first decade of the 2000s and many positions once considered "entry-level" now requiring prior work experience or being cut altogether.[117] With entry-level jobs requiring prior work experience, many college graduates seek unpaid internships in the industries they hope to work in—but it appears to be normal practice in a wide range of industries to delegate menial tasks to interns without offering much educational or career development value in return.[118] The perspectives offered by academics, consultants, and practicing managers in the 1950s-1970s offer a stark contrast. The promulgator of statistical quality

---

[114] Gomory and Sylla 2013.
[115] Mintzberg 2005.
[116] Gomory and Sylla 2013.
[117] Mitler 2013; Weber and Korn 2014.
[118] Greenhouse 2012.

control methods W. Edwards Deming, for example, argued that providing jobs was a central purpose of business; that vendor contracts should be awarded not on the basis of cost alone but built on long-term relationships of loyalty and trust; that on the job training for employees was crucial; that the hourly worker had a right to "pride of workmanship"; and that emphasis on short-term profits, mobility of management, evaluation by performance, and "running a company on visible figures alone"—now common management practice—were among the seven "deadly diseases."[119] The management goals and methods suggested by Deming seem to be part of a different economic and cultural universe than the impatient, belligerent, and occasionally messianic management style practiced in the most visible companies in today's software industry—including some of those operating online labor markets.[120] Indeed to the extent that some software companies maintain an atmosphere of egalitarianism among some employees, this is often done by outsourcing, crowdsourcing, quarantining, or otherwise making invisible the low-wage, routinized work considered uncreative, and, crucially, un*innovative*—often to units or other companies disproportionately employing women, people of color, or women of color.[121]

The labor historian Jefferson Cowie argues that the growth of employee misclassification practices calls for a renewed federal commitment to "buttress" the Fair Labor Standards Act (FLSA). The FLSA was signed into law by President

---

[119] Deming 1982, 1993.
[120] Contrast, for example, Facebook's motto until 2014—"move fast and break things"—with the old engineering maxim "measure twice, cut once"; see also Losse 2012; Coll 2014; Lacy 2014; Carr 2014, 2015; Bilbrey 2015.
[121] See e.g. Wilson 2009-2011, 2012; Wen 2014; Irani 2013, 2015.

Franklin D. Roosevelt in 1938. It "outlawed child labor, guaranteed a minimum wage, established the official length of the workweek at 40 hours, and required overtime pay for anything more," encouraging "employers to hire more people rather than work the ones they had to exhaustion."[122] The history of the FLSA, Cowie writes, has been one of expanding coverage—for example, in 1963 President John F. Kennedy signed the Equal Pay Act, amending the FLSA with a view to eliminating the gender pay gap—and increasing the minimum wage, often against fierce political opposition. Cowie argues especially that more money should be allocated to *enforcing* the FLSA given that classifying workers as independent contractors rather than employees often benefits the employer—who makes the classification decision—at the worker's expense. While the claim that independent contractors have more flexibility—and many workers value that flexibility to some extent—flexibility is in general a greater benefit to employers than to workers. And this flexibility is only empowering in practice for a minority of workers—typically the highly-skilled and already well-paid. Most workers, in contrast, would rather have stable jobs with predictable incomes. Thus the rhetoric of worker empowerment that has accompanied what legal scholar Katherine V. W. Stone calls "flexibilization"[123] is misleading. As Cowie writes, "employers will always have more power than their employees, and [...] it's in their interests to make those employees work as long and as

---

[122] Cowie 2013.
[123] Stone 2006.

cheaply as possible."[124] The argument that regulation impedes an individual's ability to make their own employment contract with their employer ignores this power differential, which arises partly from an information asymmetry (namely, employers know how much value a worker creates and how much they are paying other workers, but workers do not always know either of these things). And, as Cowie points out, this argument is an old one, just as appealing to employers today as it was a century ago, and still just as misleading:

> In Roosevelt's day, the courts found most wages and hours legislation unconstitutional based on the doctrine of "liberty of contract." The idea was as simple as it was pernicious: wages and hours legislation violated an individual's freedom to make an independent (read: worse) deal with his employer.[125]

Cowie doesn't discuss the role of information technologies in the growth of employee misclassification and other employer practices that appear to skirt existing employment and labor laws. Information technology has not driven these developments, but it has enabled them. In 2010, for example, computer scientist Luis von Ahn, inventor of reCAPTCHA and originator of the term "human computation," wrote in a blog post titled "Work and the internet":

> Recently I have heard more than one company saying something like: "We use Mechanical Turk because otherwise we would have to pay people $7/hour to do this task." In other words: "We use Mechanical Turk to get around the minimum wage laws." As wrong as it may sound to some, this is currently ok [i.e., legal]. In the United States, "independent contractors" are typically not covered by minimum wage laws, so while I'm not a lawyer I believe using Mechanical Turk to

---

[124] Cowie 2013. This idea was first or at least most influentially articulated by Karl Marx in *Capital* (1867); see also, e.g., Caffentzis 1998.
[125] Cowie 2013.

get around minimum wage is as legal as hiring independent contractors instead

of full-time employees.[126]

And crowd work intermediaries are well aware of the legal distinction between independent contractors and employees; indeed many of the their business models effectively rely on it.[127]

Thus the future of the digital economy is tightly bound up, to take the title of Cowie's 2013 editorial, with "the future of fair labor." Contrary to early proclamations of internet pioneers that "cyberspace" was a world apart from all prior human worlds,[128] the gradual and mutual imbrication of the realms called, for some time, the virtual and real has finally made the distinction unsupportable. And into the (perhaps) formerly pristine electronic utopia have come commerce, exploitation, regulation, and politics. In 2015 the early utopian internet visions read like nothing so much as disavowals of politics. But there is no escape from politics in human affairs, because politics, however unpleasant, is the art of cooperation, and few and far between are the humans who manage to survive truly alone.[129]

Technologists are now grappling with the political dimensions and implications of software work, and struggling to articulate visions for sociotechnical systems that are plausible and desirable technically *and* socially. In crowd work, for example, researcher Panos Ipeirotis in 2012 expressed regret over the association of the term crowdsourcing with low-wage labor while

---

[126] von Ahn 2010.
[127] See e.g. Neumann 2012, Amazon Mechanical Turk 2014, Kessler 2015.
[128] For example Barlow 1996 (but see also Doherty 2004); Lévy 1999. For further discussion, see Flichy 2007.
[129] For discussion, see for example Wilson 2013.

expressing hope that an organization of work can be developed that sheds the negative associations crowd work had acquired while maintaining its openness:

> ...these negative associations are now endangering a very important concept: The idea that we can structure tasks in a way that [is] robust to the presence of imperfect workers, and that anyone can participate, as long as there is work available. Well-structured tasks allow the on-the-task evaluation of the workers, and can automatically infer whether someone is a good fit for a task or not.
>
> This is not insignificant. It is well-known that one of the biggest barriers for breaking into the workforce is to have prior relevant experience. Students today often beg to get unpaid internships, just to have in their resume the lines with the coveted work experience. In online labor markets, newcomers often bid lower than what they would accept normally, just to build their feedback history. Crowdsourcing can change that.[130]

Similarly, computer scientists and designers Aniket Kittur, Jeffrey Nickerson, Michael Bernstein, Elizabeth Gerber, Aaron Shaw, John Zimmerman, Matthew Lease, and John Horton in 2013 framed the future of crowd work as a question of personal, or at least intergenerational, interest for researchers, asking, "Can we foresee a future crowd workplace in which we would want our children to participate?"[131] This framing implicitly admits the undesirable dimensions of present crowd work practices while suggesting the long-term social, economic, and political stakes involved in the design of the next generation of computationally mediated work platforms.

The priorities with which these platforms are designed and operated will have implications for the quality of work, worker influence over working

---

[130] Ipeirotis 2012, "Why I love crowdsourcing (the concept) and hate crowdsourcing (the term)."
[131] Kittur et al. 2013.

conditions, economic mobility, economic vitality and well-being in the broadest sense, and, insofar as the distribution of wealth, material resources, and economic power influences the distribution of political power, for the viability of democratic governance. The political economist Elinor Ostrom once wrote a short essay on the topic of the future of democracy, which is worth quoting at length here:

> I wish I could simply be very optimistic when discussing the future of democracy. Unfortunately, I think that it is essential that we do not naively think that the future of democracy is automatically bright. The sustenance of a democratic system is similar to the sustenance of an initially successful family firm. The first generation works very hard to build it up. The second generation has usually witnessed some of the struggles of the first generation and usually is able to continue the effort started by the first generation. But, when the firm is turned over to the third, fourth, or fifth generation, problems can occur. Children are born already rich and without a deep understanding of the struggle that it took to build the enterprise in the first place. What took many years to build can be dissipated within a short time. Now, that does not mean that all family enterprises will fail. And it certainly does not mean that all democratic institutions will eventually fail. It does mean that I share Vincent Ostrom's concern, articulated in his most recent book (Ostrom 1997), that democratic systems are vulnerable if the basic constitutive ideas of democracy are not strongly held and practised over time.
>
> Let me be bold and indicate that no democratic society can sustain itself as a democracy over multiple generations unless citizens in general understand that:

- It is always a struggle to keep a democratic system functioning as a democracy—requiring at times the willingness to engage in civil disobedience.

- There is a necessity for complex institutions that balance one another—courts that balance executives, national governments that balance regional divisions and local units—and vice versa. In other words, it is important to have multiple, organized voices—citizens who are active in political parties and other kinds of associations. It is important that there are officials who have some independence and autonomy as well as those who are elected for limited terms. And having strong local government is as important as having strong national government.

- Voting is not the only activity of a good citizen, and participation in civic groups, NGOs, and neighborhood associations is an important way of participating in democratic life.

- It is important to be active in and knowledgeable about sustaining a diversity of public and private organizations that consider alternative ways of life and public policies.

I share a deep conviction that democratic systems of government are the highest form of human governance yet developed. Yet I worry that the need for continuous civic engagement, intellectual struggle, and vigilance is not well understood in some of our mature democracies and is not transmitted to citizens and officials in new democracies.[132]

If the large-scale systems—which, as scholars in science and technology studies crucially remind us, are rarely as 'negotiable' in practice as we might like[133]—that constitute the next generation of work practices institutionally and

---

[132] Ostrom 2000.
[133] See e.g. Kallinikos 2011.

infrastructurally reify a division of labor, *and a distribution of power over the technical and social conditions of labor*, that differentially develops and affords the exercise of the skills of substantive participation in democratic deliberation over the social and material conditions of life, then those systems threaten the conditions of democracy. The risks can be put shortly and concretely by returning to Robert Reich's provocation:

> Can you imagine if this turns into a Mechanical Turk economy, where everyone is doing piecework at all odd hours, and no one knows when the next job will come, and how much it will pay? What kind of private lives can we possibly have, what kind of relationships, what kind of families?[134]

In such an economy, who will have control over the design of work, the conditions of work, and the distribution of resources broadly? Who will have time and money to educate themselves to participate meaningfully in democratic debate? It seems difficult to argue that members of the on-demand workforce will be in a position to do so. But if this is the future of work, how will the capacity of democratic self-governance be developed, and how will the citizens with those capacities exercise their power? That is, can the current "on-demand economy" coexist with real democracy?

Studies of the economic, social, and political consequences of the growth of the on-demand economy are needed, as are transdisciplinary visions for mitigating these consequences. One proposal that has been revived in recent discussions is universal basic income, a policy that would guarantee all citizens a livable income without requiring proof of need. The appropriate relationship

---

[134] Reich 2015.

of basic income to existing welfare programs is debated, as are other details, but the broad scope of the proposal suggests that awareness of the risks of the growth of computationally mediated work is itself growing.[135]

2.6.3 Turkopticon in context

Turkopticon has not been subjected to quite as many interpretations as AMT itself, but the limited coverage it has received has been diverse. We built it in the context of a class on tactical media art, so it was originally conceived of as an implicit critique—expressed through software rather than words. We hoped to draw attention to the power disparity between workers and requesters, which we saw as underaddressed in the midst of the excitement about the lowered costs and expanded operational capabilities offered by AMT. It was more easily explained to technologists as a market reputation system similar to the ones on Amazon Marketplace or eBay. This reputation-focused framing has remained salient; the economists Alan Benson, Aaron Sojourner, and Akhmed Umyarov, for example, titled the paper reporting their study of Turkopticon "The value of employer reputation in the absence of contract enforcement."[136] One technology writer, in an article for *Communications of the ACM*, focused on fairness, describing Turkopticon as an effort "to ensure fairness in crowdsourcing projects."[137] One journalist, in a framing contested by many

---

[135] See e.g. Jiang et al. 2015; Schneider 2015, "Why the tech elite is getting behind universal basic income"; and Wenger 2015.
[136] Benson et al. 2015.
[137] Hyman 2013.

workers, called Turkopticon simply, if provocatively, "Union 2.0."[138] Another, an article titled "Crowdsourcing grows up as online workers unite," described it as a sign of maturity in the crowd work industry.[139] One German researcher will present a paper about Turkopticon at the interdisciplinary Momentum conference in Austria with the title "Emancipation efforts in digital work relations."[140] (The session is titled "Emancipation through innovation?") A researcher at the Silicon Valley think tank Institute for the Future invited us to participate in a client conference as representatives of Turkopticon, which she described as "a signal of the future" (Tessa Finlev, pers. comm., 2013). (We were, regrettably, unable to attend.)

Lilly chose the subtitle "interrupting worker invisibility in AMT" for our 2013 CHI paper,[141] which was the first publication to describe Turkopticon in detail. My understanding of what Turkopticon meant, even in 2013, was still relatively dominated by ideas about two-sided reputation systems in markets, and I did not fully understand this subtitle until somewhat later, when I read early drafts of Lilly's paper about what is made economically *and culturally* possible by making crowd workers invisible.[142]

But whatever Turkopticon might be an example—or a signal—of, our dominant experience of it over the last few years—and, I think, workers' experience of it as well—is as an infrastructure. I mean this in the broadest

---

138 Brandom 2013.
139 Hodson 2013, "Crowdsourcing grows up as online workers unite."
140 Ellmer 2015.
141 Irani and Silberman 2013.
142 Irani 2013.

sense: like most information systems, it is a heterogeneous collection of specifically interconnected hardware, software, ongoing human labor, relationships, agreements, concepts, expectations, competences, and so on. Our experience of maintaining this infrastructure in collaboration with other stakeholders—mainly workers who volunteer their time to help with moderation and community issues—resonates well with recent discussions in science and technology studies and CSCW about repair and maintenance[143]—a link we made explicit in a 2014 article, in which we wrote, "Turkopticon hums along quietly on some days but lurches and drags on others."[144] For both maintainers and users, Turkopticon exhibits all the classic characteristics of infrastructure described by Susan Leigh Star and Karen Ruhleder in the late 1990s: it is embedded in other systems and social arrangements; it "invisibly" supports the tasks it was designed to support (when it is working properly); its scope extends beyond a single site or episode of practice, supporting Turking *in general*; its "proper" use is governed by norms established (and debated) within particular communities of practice (namely, those of AMT workers and requesters, debates among which often extend into our inboxes); it implements standards to connect with other systems such as AMT and the Firefox and Chrome browsers; it builds on, and is constrained by, those same technologies, as well as others such as Ruby on Rails; it becomes the topic of speculation, debate, and even panic on breakdown; it changes incrementally and modularly rather than quickly and globally; and it is relational: most Turkers and

---

[143] See e.g. Jackson et al. 2011, 2012; Jackson 2014.
[144] Irani and Silberman 2014.

requesters work *through* it, while we work *on* it.[145] And, perhaps most crucially, as Star wrote of infrastructure in general in 1999 after her encounter with the International Classification of Diseases,[146] it is, at least compared to the high-stakes drama of policy or the glamor of truly "new" technologies, rather boring. But, as Star writes, the humdrum and rather bureaucratic "dead lists" of "numbers and technical specifications" are the "hidden mechanisms" that subtend whole arenas of social and economic life. It is these often hidden details that will shape the future of work.

2.6.4   What now?

About seven minutes into Mark Neale's 2000 feature-length interview/documentary of the science fiction writer William Gibson, originator of the term "cyberspace" and one of the founders of the cyberpunk genre, Gibson reflects:

> When I was a kid, we were told that [the future was knowable]. That was when, you know, when the Future with a capital "F" was very much a going concern in North America. That was a part of our culture in the '50s, that the future was coming, and it was going to be planned. It was going to happen because grown-ups were making decisions.[147]

If one were inclined to be uncharitable, one might suggest that there seems to be a shortage of "grown-ups" managing online labor markets. The science writer Hal Hodson, writing in 2013 in *New Scientist*, described "online workers uniting"

---

[145] Star and Ruhleder 1996; Star 1999.
[146] Star 1999, p. 377.
[147] Neale 2000.

as a sign that the industry was "growing up," and argued that it was "time to focus on the welfare of online workers."[148] Three years earlier, computer scientist, business school professor, crowd work researcher, self-identified sometime AMT evangelist, and widely-read blogger Panos Ipeirotis wrote a post with the title "A plea to Amazon: fix Mechanical Turk!" and the subtitle "Mechanical Turk, it is time to grow up."[149] The post contained a long list of proposed improvements to AMT for both workers and requesters. A few of Ipeirotis' complaints have since been addressed, at least to some extent, with the Masters qualification and other requester-facing improvements.[150] But requesters, researchers, technologists, and entrepreneurs—overlapping groups—continue to air hopes for significant improvements to AMT for both workers and requesters, and to start new platforms. And the question remains—for crowd work specifically and for the software industry broadly, especially insofar as it grows into a central locus of control for ever-greater swathes of the economy—what would it mean to "grow up"?

The remainder of the dissertation can be read as a partial response to this question. In part, the problems with current online labor markets arise from operators' refusal to take responsibility for the conditions of work and the quality, or lack thereof, of the resulting work and livelihoods. This refusal is in part predicated on a simplistic portrait of the labor relation: platform operators, the argument goes, "just" connect buyers and sellers and process

---

[148] Hodson 2013, "Crowdsourcing grows up as online workers unite" and "Time to focus on the welfare of online workers."
[149] Ipeirotis 2010, "A plea to Amazon."
[150] See e.g. Ipeirotis 2013.

payments; workers are independent contractors "free to choose" to work or not work; and there being no employer to speak of, nobody can ultimately be responsible for the conditions of work, the quality of work produced, or the viability of the livelihoods produced by the system as a whole. The concepts and theoretical frameworks introduced in Chapter 3 offers a structure through which to interpret the empirical material presented in the first two chapters and to begin developing more realistic portrayals of what is happening in crowd work and online labor markets broadly. Chapter 4 then uses this material to offer concrete proposals for new systems and practices.

Chapter 3

Situatedly Rational Actors in Complex Polycentric Systems

3.1   Rational actors in perfect markets

The current distribution of benefits and risks in crowd work—tilted rather strongly in favor of platform operators and, to a slightly lesser extent, requesters—is not inevitable. But at present computing practitioners and researchers seem to lack the conceptual and methodological tools—and the institutional support—to apprehend the adverse consequences of current arrangements for many stakeholders, and to work to mitigate them. Specifically, many computing practitioners and researchers seem to take up— sometimes explicitly, more often implicitly—elements of an outdated view of humans as economic actors and the nature of economic life. This view can be described as a simple interpretation of neoclassical or "rational actor" economics, applied to social life broadly. Different computing practitioners and researchers have different particular variations on the basic theme, according to their training and experience—and the sociotechnical arrangements their institutional locations call on them to publicly justify—but for the purpose of discussion the model can be seen as made up of eight interlinked propositions. Four describe people and firms and four describe markets.[151]

---

[151] For background on the following eight propositions, see generally Becker 1978, Weintraub 2002, Samuelson and Nordhaus 2004, and Lo 2008. For discussion, see e.g. Bartlett 1989, Chs. 1-2; and Ostrom 2010, esp. pp. 2-3, 8-9.

First, people and firms—"actors"—have given, fixed, rational, and "well behaved" preferences among outcomes. Given denotes that actors' preferences are exogenous to the model; they are effectively assumed to be determined at birth. Fixed denotes that preferences do not change over time. Rational denotes that preferences are complete and transitive. Complete denotes that preferences exist for every possible combination of possible actions facing an actor. Transitive denotes that if an actor prefers outcome A to outcome B and outcome B to outcome C, the actor then also prefers outcome A to outcome C. "Well behaved" is a mathematical colloquialism denoting that preferences are monotonic and convex. Monotonic preferences mean that more of a good thing is always better, or at least as good. And convex preferences imply that actors, when acting as consumers choosing goods or services, prefer variety to lots of a particular good or service; that is, it indicates that consumers derive less benefit or pleasure from a given good or service as they consume more of it.

Second, individuals maximize their own personal happiness, quanitified and formalized as utility, and firms maximize profits. Individuals and firms—"actors"—maximize their respective "objective functions" subject to the constraints imposed by their budgets and other material resources, capabilities, and environments.

Third, people and firms act independently. That is, they choose freely among the options presented to them according to their individual preferences, which are unaffected by the preferences of other actors, the structure of the

market, or the options on offer. There is no power or coercion in market exchanges.

Fourth, people and firms make choices with complete (or "perfect") information about all available choices and the outcomes associated with each choice.

Fifth, markets are efficient aggregators of information. That is, they incentivize all actors to take actions (namely producing, selling, and buying) that result in changes to the prices of goods and services such that the price of a good or service accurately reflects at any time all the information relevant to its production, distribution, and consumption available to any actor in the market. That is, even when actors do not have "perfect information," markets do.

Sixth, there are no (or at least very low) barriers to entry for new firms offering goods and services.

Seventh, as a result of low barriers to entry, there is "perfect competition," or at least nearly so, and all firms are "price takers." No firm has the power to set or influence prices of the goods or services they sell.

Eighth and finally, markets described by the above propositions produce Pareto-efficient or Pareto-optimal outcomes. That is, they induce actors to engage in all mutually beneficial exchanges. Once Pareto-optimality has been reached, no further exchanges can be made without harming at least one party.

While many of these propositions are seen by most working economists today as having once been useful rather than as accurate descriptions of real

markets, they are often still taught to undergraduates. As such they maintain some influence over discourse on economic life and the economic responsibilities of social actors among computing practitioners and researchers, whose exposure to social science is often limited to a few courses in economics. These propositions are thus still often used, if not always rigorously, to evaluate, explain, or justify existing market outcomes, arrangements, or practices, especially in computing. For example, crowd work requesters—and some researchers—often argue that the relatively low wages available to crowd workers—or the other conditions in existing crowd work arrangements that have been listed in criticisms of the industry—are unproblematic because nobody forces crowd workers to participate in crowd work.[152] This argument proposes that if workers find requesters too cheap, or working conditions inadequate, they are free to find other work. Because many have not done so, they must be continuously and "freely" choosing to participate in crowd work, and thus the pay and working conditions generally must not be problematic.

3.2   New findings about economic actors and markets

The argument that low pay and other adverse working conditions are unproblematic, and other arguments relying on the propositions above, appear less compelling in the light of more recent empirical and theoretical research in economics, and the social and psychological sciences broadly. These

---

[152] See e.g. Cushing 2012; Irani 2013, p. 15.

developments are far from constituting a consensus, but they allow a revision of the foregoing propositions as follows.

First, preferences are not given at birth but are socially and culturally constructed. They are not fixed but change over time. This change is owed partially to the influence of other individuals and society broadly. [153] Preferences are not complete; nor are they always transitive.[154] Their logical coherence is confounded by a variety of cognitive biases and limitations.[155] And, crucially, people in longstanding conditions of deprivation or oppression may adjust their preferences to accept and even prefer circumstances they would previously have rejected.[156]

Second, people and firms may not maximize utility or profits but rather "satisfice," aiming to achieve a level of happiness or profitability above some threshold of acceptability and then declining to expend additional effort to improve the situation.[157] Additionally, people's "objective functions" may incorporate multiple desiderata, including "other-regarding preferences" such as fairness.[158]

Third, people do not choose freely. Rather, they are constantly subject both to many kinds of power exercised by other market actors and to environmental constraints that limit their freedom to make economic and social choices. While previous perspectives did not rule out environmental

---

[153] See generally Nordhaus and Tobin 1972, esp. pp 8-9; Bartlett 1989, Ch. 2; Guiso et al. 2006; Benhabib and Bisin 2011; and Dietrich and List 2012.
[154] See e.g. Tversky 1969.
[155] See e.g. Kahneman and Tversky 1979.
[156] Khader 2011.
[157] See e.g. Simon 1947, 1956.
[158] See e.g. Fehr and Schmidt 1999, 2006; Akerlof and Shiller 2009.

constraints, the extent to which they limit freedom of economic and social choice was not fully appreciated. Power, on the other hand, was excluded almost entirely from neoclassical or "rational actor" accounts of economic life.[159]

Fourth, people and firms have limited information about their choices, the outcomes likely to result from selecting any particular choice, and their likely valuations of those outcomes. Additionally, their ability to collect and process such information is limited; collecting and processing information is costly and laborious.[160]

Fifth, markets are not always efficient aggregators of information. Rather, because of human cognitive biases and limitations, markets are subject to a broad range of "irrational" dynamics such as speculative bubbles and panics.[161]

Sixth, there are not always low barriers to market entry for new firms; market power exists. Some firms are "price takers" in competitive markets; others have the power to set prices.[162] Additionally, firms may be able to lobby regulators to protect their interests, using other types of power to raise barriers to entry for new firms beyond what is possible with market power alone.[163]

Seventh, Pareto-optimality is not the only appropriate criterion for evaluating market outcomes. Specifically, an awareness of, if not desire for, fairness (cf. "inequity aversion") appears to be a human cultural universal.[164] This suggests that market outcomes should, or at least could plausibly, be

---

[159] See e.g. Bartlett 1989; Sen 1999, e.g. p. 8.
[160] See e.g. Simon 1956.
[161] See e.g. Akerlof and Shiller 2009.
[162] See e.g. Robinson 1969[1933], Manning 2005, and, in the context of crowd work (and indeed AMT) specifically, Kingsley et al. 2014.
[163] See e.g. Stigler 1971; Peltzman 1976, 1989; and Dal Bó 2006.
[164] Brown 1991; see also Fehr and Schmidt 1999, 2006; Akerlof and Shiller 2009.

evaluated not only according to the total utility achieved collectively by actors but also by the distribution of utility among actors.[165]

Eighth and finally, the violation of the conditions of perfect information and perfect competition means that even Pareto-optimality is not typically achieved.[166] That is, the notion that an "invisible hand" guides the actions of narrowly self-interested actors to lead to the greatest good for all is, regrettably, an appealing but ultimately misleading fiction.[167]

3.3 Institutions

Economic and social researchers have found that economic and social life, instead of being seen as taking place within separate spheres with their own rules—such as the market, the family, and the state—that interact only in prescribed and idealized ways (e.g., "government regulates the market"), can be more realistically understood as occurring within distinct but interlinked institutional settings. Institutions are "the prescriptions that humans use to organize all forms of repetitive and structured human interactions including those within families, neighborhoods, markets, firms, sports leagues, churches,

---

[165] cf. Ostrom 2005, p. 66.

[166] See Greenwald and Stiglitz 1986.

[167] Joseph Stiglitz, one of the economists who received the 2001 economics Nobel prize, awarded for "analyses of markets with imperfect information," wrote (Stiglitz 2006, p. xiv):
> My research on the economics of information showed that whenever information is imperfect, in particular when there are information asymmetries—where some individuals know something that others do not (in other words, *always*)—the reason that the invisible hand seems invisible is that it is not there. Without appropriate government regulation and intervention, markets do not lead to economic [i.e., Pareto] efficiency.

private associations, and governments at all scales."[168] These prescriptions can be described as rules, norms, and shared strategies.[169] These institutional statements can be described with a shared syntax: all are composed of some subset of the following components.

First, any institutional statement specifies the attributes of individuals or groups to which it applies; e.g., "United States citizens over 18 years of age." Second, norms and rules include a deontic operator such as "may," "may not," "must," or "must not." Third, any institutional statement has an aim or directive. For example, in the prescription "If you use the microwave, you must clean up your own mess!",[170] the aim is "clean up your own mess." Fourth, any institutional statement may describe specific conditions under which it applies. Fifth and finally, rules (but not norms or shared strategies) include an institutionally specified consequence for violators. This consequence is typically ensured by the existence of another rule that specifies the duties of the actors carrying it out.

The specific moments of decision shaped by the rules, norms, and shared strategies that make up institutional settings can be described as "action situations." Action situations have both internal and external factors. Important internal factors include the characteristics of the actors; the roles of the actors; the actions available to each actor; the information available to actors at each stage in the situation; the possible outcomes of the situation; the linkages

---

[168] Ostrom 2005, p. 3.
[169] The following six paragraphs summarize content from Crawford and Ostrom 1995, Ostrom 2005 (esp. Chs. 5 and 7) and Ostrom 2010.
[170] This example is from Ostrom 2005, Ch. 5, originally published in Crawford and Ostrom 1995.

between actions and outcomes; and the valuations assigned to the outcomes. Important external factors include the biological and physical context of the action situation; its social and physical context; and the relevant rules, norms, and shared strategies as understood by the actors involved. The outcomes produced in an action situation affect its social and biophysical context, which will then "feed back" into future action situations.

Another way future action situations can be changed is through institutional changes; i.e., through changes in rules, norms, and shared strategies. Such changes may take place formally in "collective choice situations," "constitutional situations," and even "metaconstitutional situations." Collective choice situations are action situations in which "operational rules"—rules governing day-to-day activities—are formally discussed and potentially revised. Constitutional situations are action situations in which the rules governing collective choice situations are formally discussed and potentially revised. And metaconstitutional situations concern the rules governing constitutional situations.

Formal organizations are collections of people and materials connected by widely known and accepted prescriptions—i.e., institutional statements—that persist, change, and produce material and social effects as individuals make and act on choices they encounter in successive interlinked action situations—operational, collective choice, constitutional, and metaconstitutional.

Actors in action situations are not fully "rational" in the neoclassical sense; that is, they do not possess complete information about all possible courses of action in a situation, all possible outcomes resulting from possible actions, or their own or others' valuations of possible outcomes. Nor are actors in action situations merely "boundedly rational," approximating or at least nominally striving for full rationality but constrained by limited information, information processing capacity, and cognitive biases. Rather, actors are situatedly rational: they do calculate and consider the actions of others, but their calculations and indeed their very preferences are shaped by both the immediate situation—including their estimations of others' preferences and their understandings of institutional prescriptions governing the action situation—and their personal histories, including ideas about appropriate conduct or desirable outcomes that they may have acquired elsewhere. These ideas may lead them to adopt "other-regarding preferences" such as a desire for fair outcomes, to advocate for adherence to a particular procedural protocol, or to aim to enact specific organizational values such as transparency.[171]

Formal organizations populated by situatedly rational actors are interlinked with one another. Sometimes this linkage is explicit and intentional; often it is implicit and accidental, occurring through the consequences on the biophysical and social contexts of decisions taken at variously institutionally located action situations. These interlinkages create complex "polycentric" systems that defy simple categories such as market, government, family, and

---

[171] See e.g. Ostrom 2005, pp. 116-119.

church. Polycentricity denotes the condition that arises when organizations with formally independent decision making centers are interlinked in practice by the consequences of the actions taken at each center.[172]

3.4 Crowd work as a polycentric institution-infrastructure system populated by situatedly rational actors

Crowd work is one such polycentric system. It is populated, as described in previous chapters, not by the narrowly self-interested "rational" actors of neoclassical economic analysis, nor even by boundedly rational actors who aim at "full" rationality (i.e., self-interestedness) but are constrainted by their limited access to information and cognitive limitations, but by situatedly rational actors who act based on a combination of enlightened self-interest, sophisticated but ever-imperfect and -evolving models of the market and its sociotechnical contexts, inequity aversion and other nonmonetary or procedural desiderata such as communicativity, and perhaps even altruism.

There is of course no perfect information in AMT or the sociotechnical systems that surround and augment it. As noted in research in both human computation[173] and human-centered computing[174] and in the previous chapters, workers spend a significant amount of unpaid time and effort searching for tasks, signaling to requesters that they are high quality workers (e.g., by completing unpaid qualification tests), sharing real-time information about

---

[172] Ostrom et al. 1961, pp. 831-832; quoted and discussed in Ostrom 2010.
[173] See e.g. Chilton et al. 2010.
[174] See e.g. Martin et al. 2014.

available tasks and requesters with one another, learning about how AMT and related systems work (technically and socially), and building and maintaining specialized information systems and communities to ensure future information sharing and learning.

There is no perfect competition either within AMT itself or against AMT. That is, while AMT significantly lowers barriers to entry for requesters wishing to post tasks, there is a power law distribution over requester participation: a small fraction of requesters post most of the tasks, meaning that these requesters have significant market power. Additionally, the existence of low barriers to entry for requesters does not produce a universe of free choice for workers: workers looking for work at a particular time often report an experience of "taking what they can get" rather than being spoilt for choice among a variety of attractive work opportunities. And while barriers to entry for requesters posting work to AMT are low, barriers to entry in the market of crowd work markets—i.e., into the market in which a new entrant would compete against AMT itself, for both workers and requesters—are relatively high. While there exist a spate of new crowdsourcing platforms, most are specialized. In its nearly ten years of operation—a lifetime in the internet industry—no general purpose crowd work platform has emerged as a serious competitor to AMT.

As a result of requesters' power in the market of AMT tasks, worker accounts of their experiences do not resonate much with crowd work

enthusiasts' promises of flexibility and freedom to choose.[175] Rather, professional crowd workers report the experience of *having* to log on to AMT and work, even in the middle of the night, if well-paying tasks from a known requester are posted—because one never knows when such an opportunity will arise again.[176] If a worker has time to work but all tasks on the platform at the time are low-paying tasks from unknown requesters (who therefore pose rejection risks), workers without access to other work must work on AMT anyway; as one worker wrote in 2008, "I realize I have a choice to work or not work on AMT, but that means I would also need to make the choice to eat or not eat, pay bills or not pay bills, etc."[177] That is, working on AMT may appear to researchers, requesters, platform operators, or other outside observers to be a "free" choice, but seen within the broader context of the worker's situation, it is not; rather, it is merely the best choice available to the worker, to their knowledge, at the time.

Worker forums and Turkopticon let workers share information with one another about which requesters and tasks, in their experience and estimation, are better, and therefore these third party platforms let workers create some consequences for requester behavior they find objectionable. While the power over requester behavior these platforms afford workers is small compared to

---

[175] See e.g. Silberman et al. 2010, "Sellers' problems in human computation markets" and "Ethics and tactics of professional crowdwork"; Irani, ed., 2008, "Turkers' Bill of Rights" and Irani, ed., 2013, "Turkers' Bill of Rights 2013"; and Martin et al. 2014.

[176] Rochelle LaPlante, pers. comm., 2015. LaPlante has worked as a professional Turker for eight years and has recently become involved in Turker-centered research and public outreach; see e.g. LaPlante and Silberman 2015; Schneider 2015, "Intellectual piecework"; and Silberman et al. 2015.

[177] Quoted in Irani 2009. This quote appeared on the home page of Turkopticon for several years.

that held over all participants by Amazon itself, which shapes the roles and rules of the market through system design, administration, and policy, AMT and these platforms and communities together form a complex polycentric system, with the consequences of decisions taken in action situations mediated primarily through one platform (e.g., a forum post or Turkopticon review) "feeding back" as context for decisions taken in subsequent action situations elsewhere (e.g., a worker searching for HITs on AMT, a requester pricing a HIT). In the terms of the institutional grammar discussed above, worker communities turn the *norm* "you should pay workers fairly [but there will be no clear or direct consequence if you don't]" into the *rule* "you should pay workers fairly or workers will warn each other to avoid your HITs, and it will be harder for you to get quality work done quickly." Because the polycentric system of crowd work is composed of both institutional statements and processes and technological infrastructures, operational rules-in-use are changed both by deliberative action in collective choice situations and by technical or administrative actions through changes to software and/or database content. The same is true for on-demand economy platforms such as Uber and TaskRabbit. And Amazon, Uber, and TaskRabbit are themselves complex systems embedded within larger institutional structures—structures such as the market for capital, the social and cultural institutions of the software industry, and municipal, state, national, and international regulatory regimes. Concepts and theoretical frameworks from empirical social science can help us understand the structure and dynamics of these complex polycentric sociotechnical systems, evaluate the

outcomes they produce against the preferences of a variety of stakeholders and the possibilities envisioned by a "sociotechnical imagination," and locate and act on opportunities to make changes that lead to outcomes many stakeholders see as better.

Chapter 4

Human-Centered Computing and the Future of Work

Amazon Mechanical Turk (AMT) produces outcomes workers see as unfair. These outcomes arise mainly because AMT gives requesters significant unchecked power over transaction outcomes. It also produces outcomes both workers and requesters see as undesirable because it primes requesters to expect quality work without having to communicate with workers. Worker forums and Turkopticon mitigate, to some extent, both the power imbalance and the lack of within-platform communication. But Turkopticon in particular creates a perceived unfairness among requesters and, sometimes, among new workers, as well as other outcomes requesters and professional Turkers see as undesirable (e.g., workers blackmailing requesters). These outcomes arise mainly because Turkopticon gives workers more power than requesters—and also because, while Turkopticon serves as another venue for worker-worker and worker-requester communication, it, like AMT itself, does not adequately set communication expectations or norms, or facilitate well the communication required to produce desired outcomes.

Turkopticon could be significantly improved on both fronts—improving the balance of power between parties and more explicitly facilitating communication—but there are practical technical limits to how effective a third-party reputation system can be. Additionally, substantive improvements

would likely require dedicated staff, and it is not clear that sustainable funding could be obtained for such staff.

A new crowd work market, on the other hand, could compete directly with AMT by integrating a robust reputation system, effectively framing communication expectations, and scaffolding necessary communication. A major risk facing the operators of such a market would be the institutional temptation to systematically privilege requesters to attract business, as AMT appears to have done, because requesters are the paying customers. Platform operators may perceive there to be a shortage of paying requesters and a surplus of workers. Yielding to this temptation, however, risks creating an imbalance of power in the market, producing undesirable outcomes for all participants over time. It may be possible to mitigate this temptation by ensuring that workers have formal power in the market's design and administrative processes, and that this power cannot be easily circumvented. Organizational forms such as the B Corporation, 501(c)(3) nonprofit, or stakeholder-owned cooperative may be more appropriate for institutionalizing such a power arrangement that the venture capital model that supports many companies in the internet industry. This approach could be used to design, build, operate, maintain, and evolve other online labor markets—i.e., competitors to existing on-demand economy platforms—and even other computing systems that function as social infrastructure. Indeed some scholarship at the intersection of law and economics argues that to maximize social returns, systems that function as infrastructure—in economic terms,

systems that serve as inputs or preconditions for a wide range of production processes and generate significant positive externalities—are best managed as commons accountable to a wide range of stakeholders rather than as private firms aiming to maximize shareholder returns.[178]

As Sec. 4.5 elaborates, recent research shows that collectively, human-centered computing (HCC) researchers are well positioned—with the relevant technical capabilities, social scientific understanding, and interest—to initiate and guide interdisciplinary, cross-sectoral efforts to imagine, design, build, operate, maintain, and evolve such "institution-infrastructure systems" that serve a broader notion of the public good than most contemporary large-scale information systems. While the practical capabilities and conceptual resources for such an effort already exist among HCC researchers, organizing the concerted and sustained action required may call for a shift in emphasis in the practices and priorities of the institutions that conduct, support, and fund HCC research.

4.1   Design issues in AMT

Amazon Mechanical Turk (AMT) produces outcomes workers see as unfair. These outcomes appear to arise mainly because AMT gives requesters significant unchecked power over transaction outcomes. It also produces outcomes both workers and requesters see as undesirable because it primes requesters to expect quality work without having to communicate with workers.

---

[178] See Frischmann 2005 and Frischmann 2012, Ch. 1.

4.1.1   Power in AMT

The main mechanism through which this power is created and exercised is the rejection mechanism—the option to decline to pay for work for any or no reason. While the AMT participation agreement (often referred to as the Terms of Service, or "TOS") says that a requester must pay for work completed to the requester's "reasonable satisfaction,"[179] as noted previously (Sec. 2.2.2), this criterion is not defined. And Amazon does not offer mediation or any other third-party recourse to workers who believe their work was rejected erroneously or even maliciously—i.e., to workers who believe that they completed the work under any "reasonable" definition of "reasonable satisfaction."[180] Workers' main within-system recourse when they believe a requester has erroneously or maliciously rejected their work is to contact the requester. But even well-intentioned requesters do not always have time to investigate, reply to, or even read worker communications or claims of requester error. And if a requester *intended* to inappropriately reject work, there is no within-system mechanism by which the worker can hold the requester accountable. Workers may reports tasks they believe violate the AMT TOS—for example, by requiring workers to disclose personal information or download software—but no processes appear to exist to notify workers of the status of Amazon's action in response to such reports; indeed many appear not to be acted upon.

---

[179] Amazon Mechanical Turk 2014, Sec. 3a.
[180] In Sec. 2 of the AMT TOS (Amazon Mechanical Turk 2014), Amazon explicitly disclaims any involvement in worker-requester transactions.

A requester responds most readily to workers' communications and objections when the requester perceives that the issue the worker is discussing directly threatens the requester's ability to get their tasks done quickly and well. For example, attentive requesters typically respond quickly to worker reports of technical problems that prevent them from completing tasks—and courteous requesters often thank workers for this information. But there is no mechanism within AMT proper by which workers can share information about their experiences with a given requester *with one another.* As a result, it is hard for workers to create consequences for requesters who engage in practices workers consider unfair.

4.1.2   Communication and information seeking in AMT

Undesirable outcomes not directly linked to the imbalance of power in the worker-requester relationship arise from the expectation shared by many requesters that work will be completed quickly and "frictionlessly"—i.e., without the need to communicate—to high quality standards. This expectation is implicitly encouraged by AMT's presentation of worker labor as "human computation" rather than human labor. But complex tasks are unlikely to get done well without some worker-requester communication. Requesters typically lack understanding of workers' knowledge, work environments, and strategies. As a result, initial designs of complex tasks often have plenty of room for improvement. This improvement is typically achieved through iteration and conversation with experienced workers. Experienced requesters know to solicit

such feedback in their tasks (if they can) and in the worker forums, and to pay attention to their Turkopticon reviews. But AMT does not advise requesters to do this; nor does it offer any technological scaffolding for efficiently soliciting worker feedback.

4.2 Design issues in Turkopticon

Workers' forums and Turkopticon mitigate, to some extent, the imbalance of power in the worker-requester relationship by providing third-party venues in which workers can share information about their experiences with requesters. Dedicated unstructured venues such as the Requesters Hall of Fame/Shame subforum on Turker Nation offer workers a context for sharing information among a community of trusted and knowledgeable colleagues, many of whom are professional Turkers. Turkopticon hosts reviews from a broad diversity of Turkers. This frequently leads to debate about "proper" use of the Turkopticon review form, a time sink and source of unwanted "drama" for many community members. Despite this, Turkopticon's structured format, its companion browser plugin, and its perception among some workers as separate from the politics between Turker forums have made it the de facto central locus for requester reviews. But despite its success in "interrupting worker invisibility" by creating a structured means for workers to create consequences for requester misbehavior, the current design of Turkopticon unintentionally and recurrently—and perhaps unnecessarily—creates a variety of frustrations for both workers and requesters.

As discussed in Chapter 2, the most prominent sources of frustration for workers and requesters using Turkopticon are the periodically recurring disagreements among workers—often between new workers and veteran Turkers and Turkopticon users—about how best to use the Turkopticon review form; worker attempts to blackmail requesters into paying for unusable work; "sockpuppet" reviewer accounts created to artificially sink or inflate requester reputations or harass other reviewers; and the "trolling" and other generic abuse familiar from other online communities. Despite our "worker-centered" intentions, Turkopticon itself has been subject to criticism at least as severe as that received by AMT itself. One worker, for example, referred to what I delicately called the "hierarchy of powers" produced by Turkopticon's moderation scheme as a "caste system," writing:

> I realize my status on your site is too low for my opinions to matter, but I was disappointed and surprised to find the caste system in use here, where busybodies from the Chosen class are allowed to swear in their own views, while censoring Little People for the use of shocking words like "jerk." *sigh*[181]

Another criticized the moderation system more directly, calling for us to replace the volunteer moderators with more experienced ones and for me to stick to programming and stop acting as a (bad) community manager.

Technical and organizational changes can easily be imagined that would go a long way toward addressing these and other criticisms. Some such changes have even been discussed on the Turkopticon mailing list, turkopticon-discuss. One change that has been discussed is the possibility of linking Turkopticon

---

[181] Received in the Turkopticon inbox, 2014.

accounts to AMT worker and requester accounts. A technique for linking AMT worker accounts to accounts on a third-party platform has already been developed by Niloufar Salehi and colleagues at Stanford University, for the Dynamo system.[182] If Turkopticon reviewer accounts were linked to AMT worker accounts, it would be straightforward to develop mechanisms to close sockpuppet accounts and discourage trolling and abuse, as workers are limited to one AMT worker account, and that account is linked to the worker's Social Security number. If requester accounts were flagged as such in the Turkopticon database and linked reliably to the requester's AMT requester account—for example, by requiring the requester to create a task with specific but randomly generated characteristics—posts by requesters could be marked clearly as such, enabling more trusted public worker-requester communication.

Turkopticon accounts linked to AMT accounts could form the basis for a more robust reputation system than the simple flagging scheme Turkopticon currently uses. In *Building Web Reputation Systems*, long-time internet industry pracitioners F. Randall Farmer and Bryce Glass describe a system that is robust enough to handle abuse but simple enough to implement reliably—and, importantly, distributes the work of evaluating "reputable objects" among all trusted users, rather than centralizing the burden of power and moderation work in the hands of a few moderators.[183] The system has different kinds of "reputable objects"—i.e., objects that have reputations. In the case of Turkopticon, the reputable objects would be reviews and comments. Users also

---

[182] Salehi et al. 2015.
[183] Farmer and Glass 2010.

have reputations, called "karma." Karma is qualitatively different than the reputations of reputable objects. Karma is computed from the reputations of the user's reputable objects. Farmer and Glass also recommend refraining from making karma public. All users can rate reputable objects (in Turkopticon's case, reviews and comments). The rating user's karma is used to weight the user's rating of the reputable object. This mechanism causes the opinions of users who have posted reviews and comments that are well appreciated by others carry more weight. Ratings should typically be multifaceted, and the rating categories should reflect generally accepted criteria among users for the various types of reputable object.

Such a system would obviate the need for the "hierarchy of powers" currently in place, and for the "manual" review of requests for commenting privileges that I currently do three times a week. (Manual review of commenting requests protects the site from abuse of the comment mechanism; while Turkopticon *reviews* can be flagged by users and hidden by moderators, *comments* cannot, and are therefore an attractive venue for trolls.) Such a system would also spread power among many trusted users—affording varying degrees of power according to how "trusted" users are—in place of the current concentration of power in the hands of a few —variously trusted, respected, feared, and publicly denounced—moderators and administrators. Such a system would involve quantifying "trust," and such quantifications are always fraught. But it seems possible that incremental development and deployment

accompanied by intensive community discussion could reduce the likelihood of severe missteps.

Like AMT, Turkopticon suffers not only from problems that stem from imbalances of power but also from problems that stem from inadequate communication. One category of problem stems from the absence of a notification system. For example, reviewers have no way to receive an automatic notification if a review they posted is hidden for violating Turkopticon's civility guidelines. If they discover "manually" that it was, and subsequently edit it to bring it into compliance, they are not notified if it is reinstated by moderators following the edit. Reviewers have no way to receive automatic notifications of comments posted in response to their reviews. And neither requesters nor reviewers have a way to receive notifications when new reviews for a particular requester are posted.

These changes, along with substantive changes to the review form to reduce "drama"-provoking ambiguities, have all been suggested by reviewers and requesters. Many have been discussed at length on turkopticon-discuss. Most would likely improve worker-worker and worker-requester relationships, speed learning, and reduce general dissatisfaction. But we have not added these features because we lack the programmer hours to do so—and, crucially, to communicate with users on an ongoing basis about the effects of the changes. For substantive technical or procedural improvements to be made to Turkopticon, therefore, will likely first require improving *organizational* capacity. Some degree of formal institutionalization may be required. If we take

this approach, a central challenge will be to improve organizational capacity without losing the ability to be responsive and accountable to workers and requesters.

4.3   Limits to improvement of current systems

While Turkopticon can be significantly improved, it cannot be improved indefinitely. Organizationally, it is not clear how to sustainably pay for dedicated staff or other expenses associated with operating Turkopticon. Communication with workers suggests that a request for workers to collectively pay the salary of a dedicated staff developer would not be well-received. Requesters might be able to fund Turkopticon, but such an arrangement risks creating conflicts of interest. As an existing system, Turkopticon is not a likely candidate for research funding. Funding is needed for software maintenance and evolution and community management, not research. Support from a combination of workers, requesters, researchers, and interested donors among the general public, nonprofit funding institutions such as foundations seems the most likely possibility for securing dedicated staff for Turkopticon maintenance and evolution.

But even in the event that such an effort were successful, there are practical technical limits on the extent to which Turkopticon—or any third party reputation system attached to AMT—can encourage workers and requesters to act in ways that produce preferable outcomes for everyone. One much-discussed hypothetical improvement, for example, would display a

requester's per-task and all-time rejection percentages on every posted task. If Amazon made this data available, existing Turker-run software tools such as HIT Scraper could be extended to allow workers to search and filter tasks by rejection percentages, just as requesters can currently screen workers by approval rate. But the mass of data relevant to a hypothetical communication and reputation system that would mitigate the abuses currently afforded by both AMT and Turkopticon to narrowly self-interested actors on both sides of the market is either sole property of Amazon or not accessible in any meaningful form—to anyone. Rejection frequency data, for example, is in Amazon's hands. And some data of interest to workers—average completion time for a task, the crucial missing datum in computing average wage, for example—are likely not available in any meaningful sense even to Amazon. The naive computation for task completion time specifically—the difference between the time a worker started a task and the time they submitted it—is confounded by the fact that workers can accept up to two dozen tasks at a time into a queue for later completion.

4.4 Design for a future crowd work market

A general purpose crowd work market based loosely on the design of AMT but with the intention to produce fair outcomes and viable livelihoods for professional workers might compete with AMT by encouraging and supporting more within-platform worker-worker and worker-requester communication—

and by including a robust two-sided reputation system in the basic design of the market. One plausible workflow for such a market is as follows.

- Requesters design tasks.

- Requesters price tasks. In pricing tasks, requesters receive guidance about different types of workers (especially professional and casual workers)—their capabilities, interests, constraints, and needs—and advice about what kind of pay is likely to produce what kind of results as well as what kind of pay is likely to attract, and be perceived as fair by, different types of workers. Detailed task wage data is made available along with a synopsis of recent work discussion of wages and tasks.

- Requesters categorize tasks for worker search. Task categories describe both generic kinds of task and what skills are needed. There is a mechanism by which workers can correct categorizations for miscategorized tasks. Requesters who repeatedly miscategorize their tasks receive guidance on proper categorization and are eventually subject to escalating sanctions, up to temporary and finally permanent account closure.

- Requesters post tasks.

- Workers search for tasks. Task search is greatly improved over AMT. Workers can search and filter by requester-and worker-reported task and skill categories; by other task criteria such as reward, estimated completion time, actual average completion time, and auto-approval time; and by review criteria (see below).

- Workers do tasks. Queueing is allowed, but is designed in a manner that allows meaningful completion times to be computed from task start and submit times.

- After submitting tasks, workers are prompted to leave a preliminary review. The server-logged completion time is recorded, but workers are asked to provide their own estimate if they were multitasking. The preliminary review form includes questions such as:

    o Did you have technical problems with the task?

    o Did you try to contact the requester?

    o If so, did the requester respond?

    o Did the task description honestly describe the task?

    o Does the task violate the platform Terms of Service?

- Requesters review submitted work and accept or reject it.

- If a requester does not review submitted work within the auto-approval time, the work is automatically approved and the worker paid. The maximum auto-approval time is seven days.

- After work is approved or rejected, the worker's preliminary review is automatically updated with the outcome. Workers can search for reviews for specific tasks based on outcome. They can also search for tasks based on per-task approval/rejection rates or the approval/rejection rates of the requester over all tasks.

    o In the event that a requester creates new accounts to attempt to drop their old approval/rejection history, the accounts are linked

by a "thesaurus" of requester aliases. Should a requester intend to change their task review processes and request that a new account not be linked to old accounts, they are asked to publicly post a declaration of intent. The declaration is reviewed by a panel of workers and other requesters, who can communicate with the requester and suggest changes until it meets with their approval. After the declaration of intent is approved by the panel, the new account is de-linked from the old accounts in the thesaurus.

- If work is rejected, the worker may ask for guidance and the opportunity to re-submit work. If the work would not be useful to the requester, the requester may decline.

  o If a worker believes their work was rejected erroneously or maliciously, they may appeal the rejection. The first step in the appeal process is for the requester to check that the task was not rejected erroneously. If it was not, the requester is asked to provide an explanation of the review process to the worker. The worker may then choose to accept the rejection or to appeal further. A further appeal is equivalent to the claim that the review process is unjust or otherwise inappropriate. Further appeals are reviewed by a panel of workers and requesters. The panel may uphold the rejection, or overturn it, and cause the worker to be paid from funds from the requester held in escrow for this purpose. In either case, the panel may issue a recommendation to the requester, the

worker, or both. The panel may recommend that the requester change their review process, or that the worker pay closer attention to particular parts of a task. Recommendations may be used in determining the outcome of future appeals. For example, consider a requester who is urged to change their review process after having a rejection overturned. If they continue to reject work, workers continue to appeal their rejections, the rejections continue to be overturned, and the requester does not appear to have changed their review process, the panel may impose escalating penalties, up to temporary and finally permanent account closure. On the other hand, consider a worker who repeatedly appeals rejections that are *not* overturned by the panel. The panel may urge the worker to be more careful in the future, to stop doing certain kinds of tasks, or to seek appropriate training. If the worker continues to appeal rejections that are not overturned by the panel, the panel may temporarily or finally permanently disable the worker's ability to appeal rejections.

- Requesters have profile pages and are encouraged to post information on them about their tasks and their organizations.

- Workers may create profile pages, but are not pressured to do so.

- All users can leave comments on tasks, profile pages, and reviews.

- Tasks, reviews, and comments are all reputable objects, subject to ratings of various kinds by workers and requesters. Reputable objects with bad

reputations are hidden. Users who consistently post tasks, reviews, or comments receiving bad reputation (i.e., users with bad "karma"; see above, Sec. 4.2) are first given guidance, then subject to escalating penalties, up to and including temporary and finally permanent account closure.

- All users can receive notifications of market events (e.g., new tasks posted meeting user-specified criteria, new reviews for a certain task or requester, comment posted to a review, update to a requester profile).

- The platform supports and scaffolds worker-requester communication. The communication record associated with a given task may be reviewed by a panel hearing a rejection appeal. Requesters are encouraged to have a dedicated human being available to interact with workers in real time through the platform while tasks are being completed.

- Anonymized market data, including task volume and value, worker earnings, and worker- and requester-supplied demographic information, are publicly available.

A variation on the above design changes the nature of the rejection mechanism: a requester may reject work, but workers are paid for rejected work. Requesters may screen workers whose work they have rejected.

4.5   Human-centered computing and the future of work

The design, development, operation, maintenance, and evolution of a new online labor market that aims to address the issues described in this

dissertation is a task for human-centered computing researchers in alliance with social scientists, software engineering researchers, and practitioners from a variety of fields. It is not a matter of "implementation" and business development to be left to industry alone, or to the nonprofit sector alone, or even to public-private partnership. A fairer future of online labor markets is a matter for interdisciplinary and cross-sectoral research because the concepts, evaluative frameworks, organizational and legal structures, and software engineering methods required do not yet exist. Most software engineering methods, for example, have been developed in the context of profit-maximizing organizations accountable to shareholders, not in institutions oriented to the systematic creation of viable livelihoods and just outcomes for stakeholders. Yet many conceptual and practical resources for a new approach to system development—and, equally crucially, the development of the organizations that would house that activity—do exist. In HCC, for example, the effort to build markets more responsive to the needs of a broad range of stakeholders can find conceptual and methodological support from research in user- and human-centered design,[184] value sensitive design,[185] ICT- and HCI4D ("Information and Communication Technologies for Development" and "Human-Computer Interaction for Development," where "development" denotes the social, economic, and political project of international development),[186] participatory

---

[184] See classically Norman 1988.
[185] See e.g. Friedman 1996.
[186] See e.g. Toyama 2010.

design,[187] critical technical practice,[188] action research,[189] feminist HCI,[190] and human needs HCI.[191] Outside HCC, movements such as social entrepreneurship,[192] new organizational forms such as the B Corporation,[193] the revival of old forms such as the worker- or stakeholder-owned cooperative,[194] and growing concern—even among business school professors—over the misalignment between the institutionally mandated goals of corporate management and public welfare[195] provide both relevant conceptual resources and models of organizational structure. As online labor intermediaries such as Amazon Mechanical Turk and oDesk create entirely new labor markets and others such as Uber and TaskRabbit position themselves as new global intermediaries in existing markets, the ideas behind the design of these markets and the institutional accountabilities of the organizations operating them will have increasingly profound and far-reaching effects on livelihood viability, economic mobility, and, ultimately, the ability of a broad range of citizens to participate substantively in democracy. HCC researchers have a combination of technical capability, social scientific understanding, and long-standing and institutionally-supported interest in human well-being. This relatively unique combination positions HCC researchers well to contribute to, and perhaps even initiate and lead, the formation of interdisciplinary and cross-

---

[187] See e.g. Schuler and Namioka 1993, Kensing and Blomberg 1998.
[188] See classically Agre 1997.
[189] See e.g. Hayes 2011.
[190] See e.g. Bardzell 2010.
[191] On human needs HCI, see Kaptelinin and Nardi 2012, pp. 73-75.
[192] See e.g. Fayolle and Matlay, eds., 2012.
[193] See foundationally Clark and Vranka 2013.
[194] See e.g. Orsi 2013.
[195] See e.g. Mintzberg 2005, 2014; Gomory and Sylla 2013.

sectoral collaborations to develop and operation new online labor markets that support workers' efforts to build viable livelihoods and substantively include workers in design and administration. Such an initiative would extend the long-running tradition in HCC research of developing systems aiming to create value for society broadly, and of orienting and evaluating design and operation according to a diverse set of conceptions of value.

References

Agre, Philip E. 1997. Toward a critical technical practice: lessons learned in trying to reform AI. In Geoffrey C. Bowker, Susan Leigh Star, William Turner, and Les Gasser, eds., *Social Science, Technical Systems, and Cooperative Work: Beyond the Great Divide.* Psychology Press.

Amazon Mechanical Turk. 2014. Amazon Mechanical Turk Participation Agreement. mturk.com/mturk/conditionsofuse.

Akerlof, George A. 1970. The market for "lemons": quality uncertainty and the market mechanism. *Quarterly Journal of Economics* **84**(3): 488-500.

Akerlof, George A., and Robert J. Shiller. 2009. *Animal Spirits: How Human Psychology Drives the Economy, and Why It Matters for Global Capitalism.* Princeton University Press.

Bardzell, Shaowen. 2010. Feminist HCI: taking stock and outlining an agenda for design. *Proc. CHI '10*: 1301-1301.

Barlow, John Perry. 1996. A declaration of the independence of cyberspace. 8 Feb 1996.

BBC News Business. 2014. What are zero-hours contracts? *BBC News Business*, 30 Apr 2014.

Becker, Gary S. 1978. *The Economic Approach to Human Behavior.* University of Chicago Press.

Bederson, Benjamin B., and Alexander J. Quinn. 2011. Web workers unite! Addressing challenges of online laborers. *Proc. CHI '11 EA*: 97-105.

Benhabib, Jess, and Alberto Bisin. 2011. Social construction of preferences: advertising. Jess Benhabib, Alberto Bisin, and Matthew O. Jackson, eds., *Handbook of Social Economics Vol. 1*: 201-220.

Benson, Alan, Aaron Sojourner, and Akhmed Umyarov. 2015. The value of employer reputation in the absence of contract enforcement: a randomized experiment. Working paper, presented at the 2015 Allied Social Science Associations meetings, 4 Jan 2015.

Bercovici, Jeff. 2014. Uber's ratings terrorize drivers and trick riders. Why not fix them? *Forbes*, 14 Aug 2014.

Berinsky, Adam J., Gregory A. Huber, and Gabriel S. Lenz. 2012. Evaluating online labor markets for experimental research: Amazon's Mechanical Turk. *Political Analysis* **20**: 351-368.

Bernstein, Michael. 2015. Crowdsourcing a meeting of minds. Talk, Department of Informatics, University of California, Irvine. 16 Jan 2015.

Bernstein, Michael. 2015. Join Stanford researchers to form the largest crowdsourcing research project ever. *Medium*, 3 Feb 2015.

Bezos, Jeff. 2006. Opening keynote and interview. MIT Emerging Technologies Conference, 27 Sep 2006.

Bilbrey, Brett. 2015. Answer to "What was it like to deliver a presentation to Steve Jobs?" *Quora*, 18 Feb 2015.

Borning, Alan, and Michael Muller. 2012. Next steps for value sensitive design. *Proc. CHI '12*: 1125-1134.

Brandom, Russell. 2013. Union 2.0: How a browser plug-in is organizing Amazon's micro-laborers. *The Verge*, 27 Jun 2013.

Brewer, Joe. 2015. Dealing with psychopaths in the internet age. *This View of Life*, 19 Jan 2015.

Brode, Noah. 2013. CMU professor: online 'crowd workers' need protections, organization. *90.5 WSEA: Pittsburgh's NPR News Station*, 17 Feb 2013.

Brown, Donald. 1991. *Human Universals*. McGraw-Hill.

Buhrmeister, Michael, Tracy Kwang, and Sam D. Gosling. 2011. Amazon's Mechanical Turk: a new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science* **6**: 3-5.

Caffentzis, George. 1998. The end of work or the renaissance of slavery? A critique of Rifkin and Negri. *Common Sense: Journal of the Edinburgh Conference of Socialist Economists* **24**, Dec 1998.

Callison-Burch, Chris. 2009. Fast, cheap, and creative: evaluating translation quality using Amazon's Mechanical Turk. *Proc. EMNLP '09*: 286-295.

Callison-Burch, Chris. 2014. Crowd-Workers: aggregating information across Turkers to help them find higher paying work. *Proc. HCOMP '14* (works-in-progress).

Carr, Paul. 2014. "We call that Boob-er:" the four most awful things Travis Kalanick said in his *GQ* profile. *PandoDaily*, 27 Feb 2014.

Carr, Paul. 2015. Travis Kalanick jokes about hiring sexist, racist blogger who says Uber should be more aggressive. *PandoDaily*, 2 Jan 2015.

Chace, Zoe, and Caitlin Kenney. 2015. The people inside your machine. *Planet Money*, 30 Jan 2015.

Chandler, Dana, and John Horton. 2011. Labor allocation in paid crowdsourcing: experimental evidence on positioning, nudges and prices. *Proc. HCOMP '11*: 14-19.

Cherry, Miriam A. 2009. Working for (virtually) minimum wage: applying the Fair Labor Standards Act in cyberspace. *Alabama Law Review* **60**: 1077-1110.

Chilton, Lydia B., John J. Horton, Robert C. Miller, and Shiri Azenkot. 2010. Task search in a human computation market. *Proc. HCOMP '10.*

Clark, William H., Jr., and Larry Vranka. 2013. The need and rationale for the benefit corporation: why it is the legal form that best addresses the needs of social entrepreneurs, investors, and, ultimately, the public. B Lab white paper.

Coll, Steve. 2014. Citizen Bezos. *New York Review of Books*, 10 Jul 2014.

Cowie, Jefferson 2013. The future of fair labor. *New York Times*, 24 Jun 2013.

Crawford, Sue E. S., and Elinor Ostrom. 1995. A grammar of institutions. *American Political Science Review* **89**(3): 582-600.

Cushing, Ellen. 2012. Dawn of the digital sweatshop. *East Bay Express*, 1 Aug 2012.

Dal Bó, Ernesto. 2006. Regulatory capture: a review. *Oxford Review of Economic Policy* **22**(2): 203-225.

Dietrich, Franz, and Christian List. 2012. Where do preferences come from? *International Journal of Game Theory* **42**: 613-637.

Deming, W. Edwards. 1982. *Out of the Crisis*. MIT Press.

Dill, Kathryn. 2014. Even post-recession, demand for temporary employees remains high. *Forbes*, 26 Aug 2014.

Dobson, Julian. 2013. Mechanical Turk: Amazon's new underclass. *Huffington Post*, 19 Feb 2013.

Doherty, Brian. 2004. John Perry Barlow 2.0. *Reason*, Aug/Sep 2004.

Dow, Steven, Anand Kulkarni, Scott Klemmer, and Björn Hartmann. 2012. Shepherding the crowd yields better work. *Proc. CSCW '12*: 1013-1022.

Downs, Julie S., Mandy B. Holbrook, Steve Sheng, and Lorrie F. Cranor. 2010. Are your participants gaming the system? Screening Mechanical Turk workers. *Proc. CHI '10 EA*: 2399-2402.

Economist Staff. 2015. There's an app for that. *The Economist*, 3 Jan 2015.

Economist Staff. 2015. Workers on tap. *The Economist*, 3 Jan 2015.

Ekbia, Hamid, and Bonnie Nardi. 2014. Heteromation and its (dis)contents. *First Monday* **19**(6).

Ellmer, Markus. 2015. Emanzipationsbestrebungen in digitalen Arbeitsbeziehungen: Turkopticon als webbasierte Einflusstechnologie auf der Human Computation-Plattform Mechanical Turk. *Momentum Kongress 2015*, to appear.

Faridani, Siamak, Björn Hartmann, and Panagiotis G. Ipeirotis. 2011. What's the right price? Pricing tasks for finishing on time. *Proc. HCOMP '11*: 26-31.

Farmer, F. Randall, and Bryce Glass. 2010. *Building Web Reputation Systems*. Yahoo Press.

Fayolle, Alain, and Harry Matlay, eds. 2012. *Handbook of Research on Social Entrepreneurship*. Edward Elgar.

Fehr, Ernst, and Klaus M. Schmidt. 1999. A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics* **114**(3): 817-868.

Fehr, Ernst, and Klaus M. Schmidt. 2006. The economics of fairness, reciprocity, and altruism — experimental evidence and new theories. In Serge-Christophe Kolm and Jean Mercier Ythier, eds., *Handbook of the Economics of Giving, Altruism and Reciprocity, Vol. 1*. Elsevier.

Felstiner, Alek. 2011. Working the crowd: employment and labor law in the crowdsourcing industry. *Berkeley Journal of Employment and Labor Law* **32**: 143-204.

Flanagan, Mary, Daniel Howe, and Helen Nissenbaum. 2008. Embodying values in design: theory and practice. In Jeroen van den Hoven and John Weckert, eds., *Information Technology and Moral Philosophy*. Cambridge University Press.

Flichy, Patrice. 2007. *The Internet Imaginaire*. MIT Press.

Flyvbjerg, Bent. 2001. *Making Social Science Matter: Why Social Inquiry Fails and How It Can Succeed Again*. Cambridge University Press.

Flyvbjerg, Bent, Todd Landman, and Sanford Schram, eds. 2012. *Real Social Science: Applied Phronesis*. Cambridge University Press.

Folbre, Nancy. 2013. The unregulated work of Mechanical Turk. *New York Times: Economix*, 18 Mar 2013.

Fort, Karën, Gilles Adda, and K. Bretonnel Cohen. 2011. Amazon Mechanical Turk: gold mine or coal mine? *Computational Linguistics* **37**(2): 413-420.

Friedman, Batya. 1996. Value-sensitive design. *Interactions* **3**(6): 17-23.

Friedman, Batya, and Peter H. Kahn, Jr. 2002. Human values, ethics, and design. In Andrew Sears and Julie A. Jacko, eds., *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications.* Lawrence Erlbaum.

Friedman, Milton. 1970. The social responsibility of business is to increase its profits. *New York Times Magazine,* 13 Sep 1970.

Frischmann, Brett M. 2005. An economic theory of infrastructure and commons management. *Minnesota Law Review* **89**: 917.

Frischmann, Brett M. 2012. *Infrastructure: The Social Value of Shared Resources.* Oxford University Press.

Gomory, Ralph, and Richard Sylla. 2013. The American corporation. *Dædalus* **142**(2): 102-118.

gore313 et al. 2013/2014. Turking for a living? *Turker Nation,* first post 19 Jan 2013.

Greenhouse, Steven. 2012. Jobs few, grads flock to unpaid internships. *New York Times,* 5 May 2012.

Greenwald, Bruce C., and Joseph E. Stiglitz. 1986. Externalities in economies with imperfect information and incomplete markets. *Quarterly Journal of Economics* **101**(2): 229-264.

Grier, David Alan. 2011. Error identification and correction in human computation: lessons from the WPA. *Proc. HCOMP '11*: 32-36.

Guardian Editorial Board. 2015. The Guardian view on zero-hours: the number that keeps getting bigger. *The Guardian*, 25 Feb 2015.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales. 2006. Does culture affect economic outcomes? *Journal of Economic Perspectives* **20**(2): 23-48.

Hammond, Eric. 2010. AWS meetup (Mechanical Turk) in Santa Monica on March 2. *BarCampLA* (mailing list), 11 Feb 2010. groups.google.com/forum/#!msg/barcampla/Nu1JopyRAJ4/lA8-2izrQmsJ.

Hammond, Stig. 2005. Amazon's Mechanical Turk is the human API. *Watson, Come Here, I Need You*, 4 Nov 2005.

Harris, Mark. 2014. Amazon's Mechanical Turk workers protest: 'I am a human being, not an algorithm.' *The Guardian*, 3 Dec 2014.

Hayes, Gillian R. 2011. The relationship of action research to human-computer interaction. *Transactions on Computer-Human Interaction* **18**(3): 15.

Heer, Jeffrey, and Michael Bostock. 2010. Crowdsourcing graphical perception: using Mechanical Turk to assess visualization design. *Proc. CHI '10*: 203-212.

Ho, Chien-Ju, Yu Zhang, Jennifer Wortman Vaughan, and Mihaela van der Schaar. 2012. Towards social norm design for crowdsourcing markets. *Proc. HCOMP '12*: 94-100.

Hodson, Hal. 2013. Crowdsourcing grows up as online workers unite. *New Scientist*, 7 Feb 2013.

Hodson, Hal. 2013. Time to focus on the welfare on online workers. *New Scientist*, 8 Feb 2013.

Howe, Jeff. 2008. Mechanical Turk targets small business. *Crowdsourcing*, 1 Aug 2008.

Hudson, Laura. 2014. Curbing online abuse isn't impossible. Here's where we start. *Wired*, 15 May 2014.

Huet, Ellen. 2014. Contractor or employee? Silicon Valley's branding dilemma. *Forbes*, 18 Nov 2014.

Hyman, Paul. 2013. Software aims to ensure fairness in crowdsourcing projects. *Communications of the ACM* **56**(8): 19-21.

Ipeirotis, Panagiotis G. 2007. Experiences using Mechanical Turk. *A Computer Scientist in a Business School*, 7 Sep 2007.

Ipeirotis, Panagiotis G. 2008. Mechanical Turk: the demographics. *A Computer Scientist in a Business School*, 19 Mar 2008.

Ipeirotis, Panagiotis G. 2009. Turker demographics vs. internet demographics. *A Computer Scientist in a Business School*, 16 Mar 2009.

Ipeirotis, Panagiotis G. 2010. The new demographics of Mechanical Turk. *A Computer Scientist in a Business School*, 9 Mar 2010.

Ipeirotis, Panagiotis G. 2010. Mechanical Turk, low wages, and the market for lemons. *A Computer Scientist in a Business School*, 27 Jul 2010.

Ipeirotis, Panagiotis G. 2010. A plea to Amazon: fix Mechanical Turk! *A Computer Scientist in a Business School*, 21 Oct 2010.

Ipeirotis, Panagiotis G. 2012. Why I love crowdsourcing (the concept) and hate crowdsourcing (the term). *A Computer Scientist in a Business School*, 13 Nov 2012.

Ipeirotis, Panagiotis G. 2012. Mechanical Turk changing the defaults: the game has changed. *A Computer Scientist in a Business School*, 5 Dec 2012.

Ipeirotis, Panagiotis G. 2013. Reply to "Is Mechanical Turk really broken?" *Quora*, 2 Sep 2013. qr.ae/B7gZn.

Ipeirotis, Panagiotis G. 2015. Demographics of Mechanical Turk: now live! (April 2015 edition). *A Computer Scientist in a Business School*, 6 Apr 2015.

Ipeirotis, Panagiotis G., Foster Provost, and Jing Wang. 2010. Quality management on Amazon Mechanical Turk. *Proc. HCOMP '10*: 64-67.

Irani, Lilly C., ed. 2008. *Turk Work*. turkwork.differenceengines.com.

Iraniy, Lilly C. 2009. Tweaking technocapitalism: Turkopticon. *Difference Engines*, 30 Jan 2009.

Irani, Lilly C., ed. 2013. *Turkers' Bill of Rights—2013 Edition*. turkwork2013.differenceengines.com.

Irani, Lilly. 2013. The cultural work of microwork. *New Media & Society*, 21 Nov 2013.

Irani, Lilly. 2015. Difference and dependence among digital workers: the case of Amazon Mechanical Turk. *South Atlantic Quarterly* **114**(1): 225-234.

Irani, Lilly C., and M. Six Silberman. 2013. Turkopticon: interrupting worker invisibility in Amazon Mechanical Turk. *Proc. CHI '13*: 611-620.

Irani, Lilly, and M. Six Silberman. 2014. From critical design to critical infrastructure. *Interactions* **21**(4): 32-35.

Jackson, Steven J., Alex Pompe, and Gabriel Krieshok. 2011. Things fall apart: maintenance, repair, and technology for education initiatives in rural Namibia. *Proc. iConference '11*: 83-90.

Jackson, Steven J., Alex Pompe, and Gabriel Krieshok. 2012. Repair worlds: maintenance, repair, and ICT for development in rural Namibia. *Proc. CSCW '12*: 107-116.

Jackson, Steven J. 2014. Rethinking repair. In Tarleton Gillespie, Pablo J. Boczkowski, and Kirsten A. Foot, eds., *Media Technologies: Essays on Communication, Materiality, and Society*. MIT Press.

Jain, Shaili, and David C. Parkes. 2009. The role of game theory in human computation systems. *Proc. HCOMP '09*: 58-61.

Jiang, Ling, Christian Wagner, and Bonnie Nardi. 2015. Not just in it for the money: a qualitative investigation of workers' perceived benefits of micro-task crowdsourcing. *Proc. HICSS '15*: 773-782.

Jung, Hyun Joon, and Matthew Lease. 2012. Improving quality of crowdsourced labels via probabilistic matrix factorization. *Proc. HCOMP '12*: 101-106.

Kahneman, Daniel, and Amos Tversky. 1979. Prospect theory: an analysis of decision under risk. *Econometrica* **47**(2): 263-292.

Kallinikos, Jannis. 2011. *Governing Through Technology: Information Artefacts and Social Practice*. Palgrave Macmillan.

Kaptelinin, Victor, and Bonnie Nardi. 2012. *Activity Theory in HCI: Fundamentals & Reflections.* Morgan & Claypool.

Katz, Leslie. 2008. Pirates of the Amazon hits the rocks. *CNET*, 6 Dec 2008.

Kensing, Finn, and Jeanette Blomberg. 1998. Participatory design: issues and concerns. *Computer Supported Cooperative Work* **7**: 167-185.

Kessler, Sarah. 2014. Pixel and dimed: on (not) getting by in the gig economy. *Fast Company*, 18 Mar 2014.

Kessler, Sarah. 2015. The gig economy won't last because it's being sued to death. *Fast Company*, 17 Feb 2015.

Khader, Serene J. 2011. *Adaptive Preferences and Women's Empowerment.* Oxford University Press.

Kingsley, Sara C., Mary L. Gray, Siddharth Suri. 2014. Monopsony and the crowd: labor for lemons? *Proc. 2014 Oxford Internet Institute Conference on Internet, Policy & Politics.*

Kittur, Aniket, Ed H. Chi, and Bongwon Suh. 2008. Crowdsourcing user studies with Mechanical Turk. *Proc. CHI '08*: 453-456.

Kittur, Aniket, Jeffrey V. Nickerson, Michael S. Bernstein, Elizabeth M. Gerber, Aaron Shaw, John Zimmerman, Matthew Lease, and John J. Horton. 2013. The future of crowd work. *Proc. CSCW '13*: 1301-1317.

Knobel, Cory, and Geoffrey C. Bowker. 2011. Values in design. *Communications of the ACM* **54**(7): 26-28.

Kochhar, Shailesh, Stefano Mazzocchi, and Praveen Paritosh. 2010. The anatomy of a large-scale human computation engine. *Proc. HCOMP '10*: 10-17.

Kosoff, Maya. 2015. 2 lawsuits could dramatically alter the business model for Uber and Lyft. *Business Insider*, 30 Jan 2015.

Kou, Yubo, and Bonnie Nardi. 2013. Regulating anti-social behavior on the internet: the example of *League of Legends*. *Proc. iConference '13*.

Kou, Yubo, and Bonnie Nardi. 2014. Governance in *League of Legends*: a hybrid system. *Proc. Foundations of Digital Games* '14.

Lacy, Sarah. 2014. Venture capital and the great big Silicon Valley asshole game. *PandoDaily*, 6 Oct 2014.

LaPlante, Rochelle, and M. Six Silberman. 2015. Design notes for a future crowd work market. *Medium*, 13 Feb 2015.

Le Dantec, Christopher A., Erika S. Poole, and Susan P. Wyche. 2009. Values as lived experience: evolving value sensitive design in support of value discovery. *Proc. CHI '09*: 1141-1150.

Lee, Min Kyung, Daniel Kusbit, Evan Metsky, and Laura Dabbish. 2015. Working with machines: the impact of algorithmic and data-driven management on human workers. *Proc. CHI '15*, to appear.

Leingruber, Tobias. 2008. Pirates of the Amazon. tobi-x.com/pirates-of-the-amazon-browser-add-on.html.

Leonard, Andrew. 2013. The internet's destroying work—and turning the old middle class into the new proletariat. *Salon.com*, 12 Jul 2013.

Levine, Dan, and Sarah McBride. 2015. Uber, Lyft face crucial courtroom test over driver benefits. *Reuters*, 28 Jan 2015.

Lévy, Pierre. 1999. *Collective Intelligence: Mankind's Emerging World in Cyberspace*. Basic.

Lo, Andrew. 2008. Efficient market hypothesis. In Steven N. Durlauf and Lawrence E. Blume, eds., *New Palgrave Dictionary of Economics*. 2nd ed. Palgrave Macmillan.

Losse, Katherine. 2012. *The Boy Kings: A Journey into the Heart of the Social Network*. Free Press.

Lustig, Caitlin, and Bonnie Nardi. 2015. Algorithmic authority: the case of Bitcoin. *Proc. HICSS '15*: 743-752.

Manjoo, Farhad. 2015. Uber's business model could change your work. *New York Times*, 28 Jan 2015.

Manning, Alan. 2005. *Monopsony in Motion: Imperfect Competition in Labor Markets*. Princeton University Press.

Mao, Andrew, Ece Kamar, Yiling Chen, Eric Horvitz, Megan E. Schwamb, Chris J. Lintott, and Arfon M. Smith. 2013. Volunteering versus work for pay: incentives and tradeoffs in crowdsourcing. *Proc. HCOMP '13*: 94-102.

Mao, Andrew, Ariel D. Procaccia, and Yiling Chen. 2012. Better computation through principled voting. Orig. "Social choice for human computation," *Proc. HCOMP '12*: 136-142.

Marder, Jenny. 2015. The internet's hidden science factory. *PBS News Hour*, 11 Feb 2015.

Martin, David, Benjamin V. Hanrahan, Jacki O'Neill, and Neha Gupta. 2014. Being a Turker. *Proc. CSCW '14*: 224-235.

Marx, Karl. 1867. *Das Kapital: Kritik der Politischen Ökonomie*. Otto Meissner Verlag.

Mason, Winter, and Siddharth Suri. 2011. Conducting behavioral research on Amazon's Mechanical Turk. *Behavioral Research*, 30 Jun 2011.

Milland, Kristy, Rochelle LaPlante, Manish Bhatia, Harris Paltrowitz, and Lilly Irani. 2014. MTurk workers: we're here and we're human. 15 Nov 2014, panel at *Digital Labor: Sweatshops, Picket Lines, Barricades*, 14–16 Nov 2014, New School for Liberal Arts, New York, NY.

Mitler, Lesley. 2013. Entry-level jobs are no longer entry-level. *SimplyHired*, 13 Aug 2013.

Mintzberg, Henry. 2005. *Managers Not MBAs: A Hard Look at the Soft Practice of Managing and Management Development*. Berrett-Koehler Publishers.

Mintzberg, Henry. 2009. Rebuilding companies as communities. *Harvard Business Review*, July 2009.

Mintzberg, Henry. 2015. *Rebalancing Society: Radical Renewal Beyond Left, Right, and Center*. Berrett-Koehler.

Montgomery, Kevin. 2014. Handy sued for being a hellscape of labor code violations. *ValleyWag*, 12 Nov 2014.

National Employment Law Project. 2014. The low-wage recovery: industry employment and wages four years into the recovery. NELP Data Brief, April 2014.

Neale, Mark. 2000. *No Maps for These Territories*. Docurama.

Neumann, Erik. 2012. Tech company CrowdFlower denies labor violation. *Mission Local*, 4 Dec 2012.

Nordhaus, William D., and James Tobin. 1972. Is growth obsolete? In Nordhaus and Tobin, eds., *Economic Research: Retrospect and Prospect, Vol. 5: Economic Growth*. National Bureau of Economic Research.

Norman, Don. *The Design of Everyday Things*. Basic Books.

O'Connor, Sarah. 2013. Amazon unpacked. *Financial Times*, 8 Feb 2013.

Oleson, David, Alexander Sorokin, Greg Laughlin, Vaughn Hester, John Le, and Lukas Biewald. 2011. Programmatic gold: targeted and scalable quality assurance in crowdsourcing. *Proc. HCOMP '11*: 43-48.

Orsi, Janelle. 2013. The sharing economy just got real. *Shareable*, 16 Sep 2013.

Ostrom, Elinor. 2000. The future of democracy. *Scandinavian Political Studies* **23**(3): 280-283.

Ostrom, Elinor. 2005. *Understanding Institutional Diversity*. Princeton University Press.

Ostrom, Elinor. 2010. Beyond markets and states: polycentric governance of complex economic systems. *American Economic Review* **100**: 1-33.

Ostrom, Vincent. 1997. *The Meaning of Democracy and the Vulnerabilities of Democracies: A Response to Tocqueville's Challenge*. University of Michigan Press.

Ostrom, Vincent, Charles M. Tiebout, and Robert Warren. 1961. The organization of government in metropolitan areas: a theoretical inquiry. *American Political Science Review* **55**(4): 831-842.

Paolacci, Gabriele, Jesse Chandler, and Panagiotis G. Ipeirotis. 2010. Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making* **5**(5): 411-419.

Paolacci, Gabriele, and Jesse Chandler. 2014. Inside the Turk: understanding Mechanical Turk as a participant pool. *Current Directions in Psychological Science* **23**(3): 184-188.

Peltzman, Sam. 1976. Toward a more general theory of regulation. *Journal of Law and Economics* **19**: 211-248.

Peltzman, Sam. 1989. The theory of economic regulation after a decade of deregulation. *Brookings Papers on Economic Activity: Microeconomics 1989*: 1-59.

Pontin, Jason. 2007. Artificial intelligence, with help from the humans. *New York Times*, 25 Mar 2007.

Puzzanghera, Jim. 2014. Economy has recovered 8.7 million jobs lost in Great Recession. *Los Angeles Times*, 6 Jun 2014.

Rao, Huaming, Shih-Wen Huang, and Wai-Tat Fu. 2013. What will others choose? How a majority vote reward scheme can improve human computation in a spatial location identification task. *Proc. HCOMP '13*: 130-137.

Reich, Robert. 2015. The sharing economy is hurtling us backward. *Salon.com*, 4 Feb 2015.

Roberts, Yvonne. 2014. Why can't Britain create decent jobs? Meet the women struggling against low pay and zero-hours contracts. *The Guardian*, 13 Dec 2014.

Robinson, Joan Violet. 1969 [1st ed. 1933]. *The Economics of Imperfect Competition.* 2nd ed. Macmillan.

Rooney, Ben. 2014. The FedEx driver who sued and won. *CNN Money*, 20 Nov 2014.

Ross, Joel, and Bill Tomlinson. 2009. Worker motivations in Amazon Mechanical Turk. Working paper.

Ross, Joel, Lilly Irani, M. Six Silberman, Andrew Zaldivar, and Bill Tomlinson. 2010. Who are the crowdworkers? Shifting demographics in Mechanical Turk. *Proc. CHI '10 EA*: 2863-2872.

Said, Carolyn. 2014. TaskRabbit makes some workers hopping mad. *SFGate*, 18 Jul 2014.

Salehi, Niloufar, Lilly Irani, Michael Bernstein, Ali Alkhatib, Eva Ogbe, Kristy Milland, and Clickhappier. 2015. We are Dynamo: overcoming stalling and friction in collective action for crowd workers. *Proc. CHI '15*, to appear.

Samuelson, Paul, and William Nordhaus. 2004. *Economics.* 14th ed. McGraw-Hill.

Schneider, Nathan. 2015. Why the tech elite is getting behind universal basic income. *Vice*, 6 Jan 2015.

Schneider, Nathan. 2015. Intellectual piecework. *The Chronicle of Higher Education*, 16 Feb 2015.

Schram, Sanford, and Brian Caterino. 2006. *Making Political Science Matter: Debating Knowledge, Research, and Method.* New York University Press.

Schuler, Douglas, and Aki Namioka, eds. 1993. *Participatory Design: Principles and Practices.* Lawrence Erlbaum.

Sen, Amartya. 1999. *Development as Freedom*. Anchor Books.

Silberman, M. Six, Joel Ross, Lilly Irani, and Bill Tomlinson. 2010. Sellers' problems in human computation markets. *Proc. HCOMP '10*: 18-21.

Silberman, M. Six, Lilly Irani, and Joel Ross. 2010. Ethics and tactics of professional crowdwork. *XRDS* **17**(2): 39-43.

Silberman, M. Six, Kristy Milland, Rochelle LaPlante, Joel Ross, and Lilly Irani. 2015. Stop citing Ross et al. 2010, "Who are the crowdworkers?" *Medium*, 16 Mar 2015.

Simon, Herbert A. 1947. *Administrative Behavior: A Study of Decision-Making Processes in Administrative Organizations*. Macmillan.

Simon, Herbert A. 1956. Rational choice and the structure of the environment. *Psychological Review* **63**(2): 129-138.

Singer, Yaron, and Manas Mittal. 2011. Pricing tasks in online labor markets. *Proc. HCOMP '11*: 55-60.

Singla, Adish, and Andreas Krause. 2013. Incentives for privacy tradeoff in community sensing. *Proc. HCOMP '13*: 165-173.

Smith, Rebecca. 2015. Will Uber and Lyft make your job obsolete? *CNN*, 10 Feb 2015.

Snow, Rion, Brendan O'Connor, Daniel Jurafsky, and Andrew Y. Ng. 2008. Cheap and fast—but is it good? Evaluating non-expert annotations for natural language tasks. *Proc. EMNLP '08*: 254-263.

spamgirl. 2015. What you write doesn't represent me as an mTurk worker, so don't pretend it does. *Turker Nation*, 20 Feb 2015.

Sorokin, Alexander, and David Forsyth. 2008. Utility data annotation with Amazon Mechanical Turk. *Proc. CVPR '08.*

Star, Susan Leigh, and Karen Ruhleder. 1996. Steps toward an ecology of infrastructure: design and access for large information spaces. *Information Systems Research* **7**(1): 111-134.

Star, Susan Leigh. 1999. The ethnography of infrastructure. *American Behavioral Scientist* **43**(3): 377-391.

Stigler, George. 1971. The theory of economic regulation. *Bell Journal of Economics and Management Science* **2**(1): 3-21.

Stone, Katherine V. W. 2006. Flexibilization, globalization, and privatization: three challenges to labor rights in our time. *Osgoode Hall Law Journal* **44**(1): 77-104.

Swan, Rachel. 2013. Lyft sued for unfair competition and labor violations. *SF Weekly*, 5 Sep 2013.

Toyama, Kentaro. 2010. Human-computer interaction and global development. *Foundations and Trends in Human-Computer Interaction* **4**(1): 1-79.

Tversky, Amos. 1969. Intransitivity of preferences. *Psychological Review* **76**(1): 31-48.

Uddin, Zakia. 2012. The dystopian digital sweatshop that makes the internet run. *AlterNet*, 11 Sep 2012.

Vakharia, Donna, and Matthew Lease. 2013. Beyond AMT: an analysis of crowd work platforms. University of Texas, Austin Technical Report. arXiv:1310.1672 [cs.CY].

Vakharia, Donna, and Matthew Lease. 2015. Beyond Mechanical Turk: an analysis of paid crowd work platforms. *Proc. iConference '15*, to appear.

von Ahn, Luis. 2010. Work and the internet. *Luis von Blog*, 26 Jul 2010.

Weber, Lauren and Melissa Korn. 2014. Where did all the entry-level jobs go? Many firms expect new graduates to arrive job-ready from day one. *Wall Street Journal*, 6 Aug 2014.

Weintraub, E. Roy. 2002. Neoclassical economics. *Concise Encyclopedia of Economics*. Library of Economics and Liberty.

Wen, Shawn. 2014. The ladies vanish. *The New Inquiry*, 11 Nov 2014.

Wenger, Albert. 2015. *Vice* on universal basic income: a response. *Continuations*, 7 Jan 2015.

Wilson, Andrew Norman. 2009-2011. *Workers Leaving the Googleplex*. Multipart artistic production; see andrewnormanwilson.com/ WorkersGoogleplex.html.

Wilson, Andrew Norman. 2012. *ScanOps*. Multipart artistic production; see andrewnormanwilson.com/ScanOps.html.

Wilson, Edward O. 2013. *The Social Conquest of Earth*. Liveright.

Witkowski, Jens, Yoram Bachrach, Peter Key, and David C. Parkes. 2013. Dwelling on the negative: incentivizing effort in peer prediction. *Proc. HCOMP '13*: 190-197.

Zittrain, Jonathan. 2009. Work the new digital sweatshops. *Newsweek*, 7 Dec 2009.