

Daemo: a Self-Governed Crowdsourcing Marketplace

Stanford Crowd Research Collective *
Stanford HCI Group daemo@cs.stanford.edu

ABSTRACT

Crowdsourcing marketplaces provide opportunities for autonomous and collaborative professional work as well as social engagement. However, in these marketplaces, workers feel disrespected due to unreasonable rejections and low payments, whereas requesters do not trust the results they receive. The lack of trust and uneven distribution of power among workers and requesters have raised serious concerns about sustainability of these marketplaces. To address the challenges of trust and power, this paper introduces Daemo, a self-governed crowdsourcing marketplace. We propose a *prototype task* to improve the work quality and *open-governance model* to achieve equitable representation. We envisage Daemo will enable workers to build sustainable careers and provide requesters with timely, quality labor for their businesses.

Author Keywords

crowdsourcing; crowd research; crowd work.

ACM Classification Keywords

H.5.3. Group and Organization Interfaces: Computer-supported cooperative work

INTRODUCTION

Paid crowdsourcing marketplaces such as Mechanical Turk and Upwork have created opportunities for workers to supplement their income and enhance their skills, while allowing requesters to get their work completed efficiently. These marketplaces have attracted many participants globally; however, they have repeatedly failed to ensure high-quality results, fair wages, respect for workers, and convenience in authoring effective tasks [1].

* This project was created via a world-wide, crowdsourced research process initiated at Stanford University: S. Gaikwad, D. Morina, R. Nistala, M. Agarwal, A. Cossette, R. Bhanu, S. Savage, V. Narwal, K. Rajpal, J. Regino, A. Mithal, A. Ginzberg, A. Nath, K. R. Zilukoski, T. Cossette, D. Gamage, A. Richmond-Fuller, R. Suzuki, J. Herrejon, K. V. Le, C. Flores-Saviaga, H. Thilakarathne, K. Gupta, W. Dai, A. Sastry, S. Goyal, T. Rajapakshe, N. Abolhassani, A. Xie, A. Reyes, S. Ingle, V. Jaramillo, M.D. Godinez, W. Angel, M. Godinez, C. Toxtli, J. Flores, A. Gupta, V. Sethia, D. Padilla, K. Milland, K. Setyadi, N. Wajirasena, M. Batagoda, R. Cruz, J. Damon, D. Nekkanti, T. Sarma, M.H. Saleh, G. Gongora-Svartzman, S. Bateni, G. Toledo-Barrera, A. Pena, R. Compton, D. Aariff, L. Palacios, M. P. Ritter, Nisha K.K., A. Kay, J. Uhrmeister, S. Nistala, M. Esfahani, E. Bakiu, C. Diemert, L. Matsumoto, M. Singh, V. Jaramillo-Lopez, K. Patel, R. Krishna, G. Kovacs, R. Vaish, M. Bernstein

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.
Copyright is held by the owner/author(s).

UIST '15 Adjunct, November 08-11, 2015, Charlotte, NC, USA.
ACM 978-1-4503-3780-9/15/11
<http://dx.doi.org/10.1145/2815585.2815739>

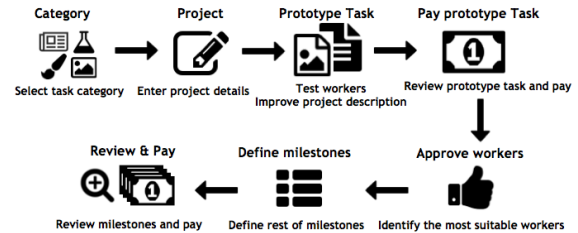


Figure 1. Task creation workflow for a requester: prototype task creation, initial submissions review, and hiring high quality workers for future milestones. [<https://daemo.stanford.edu>]. Icon courtesy Font Awesome by Dave Gandy - <http://fontawesome.io>

From our interviews with requesters, it has become clear that they struggle to trust their workers. They will rerun tasks, discard gathered data, and add increasingly complex worker filters. On the other hand, workers do not trust requesters to follow through with pay and fair treatment. In response, workers often withhold their full effort unless they have an experience with the requester.

Moreover, existing marketplaces suffer from uneven distributions of power [4]. For example, requesters have the power to deny payments for finished tasks and workers have inadequate means to contest this. Operational governance and rules have been secondary considerations on markets thus far, fitted to support the focus on the commoditizing of work. This resulted in an asymmetrical relationship between workers, requesters, and the marketplace on fronts such as parity of information access, wage negotiation, and reputation. A common complaint [3]: “We can be rejected yet the requesters still have our articles and sentences. Not Fair.”

We present Daemo, a crowd-built, self-governed crowdsourcing marketplace. To increase trust, we introduce the idea of *prototype tasks*, where each new task must first launch in an intermediate feedback mode where workers can comment on the task, requesters can review the submissions and qualify a subset of workers to continue. During this phase, workers and requesters work together to refine the task description and reduce errors. Daemo also adopts a representative democratic governance model to elect a leadership board. Engaging all vested parties in the governance of the marketplace gives an opportunity to create genuine worker-requester relationships and redefine the future of work.

RELATED WORK

Feedback, wages, task decomposition, and quality control are some of the fundamental elements of a successful crowdsourcing marketplace [1]. Requesters often rely on “gold standard” tasks, i.e., questions with known answers, to evaluate the performance and quality of submissions [2]. However,

this tends to still place all blame on workers rather than letting requesters share it for poor task interface design. Several venues such as Turker Nation and Dynamo Forums have been created so that workers' opinion can be heard [3, 4]. However, these venues are still outside the marketplaces, isolated from the requesters' reach. Daemo aims to embed an open-governance structure as a part of the platform's design.

DAEMO

Promoting Trust and Power in Tasks

Daemo tackles issues of trust and power by giving requesters a low-risk method for testing task quality and workers' abilities. Daemo divides all tasks into milestones. A task can have one or many milestones. For instance, a macrotask of "make a poster" could have a first milestone of "sketch the poster's layout". Or, with a microtask of labeling 1,000 images, the first milestone might be to have three workers label ten images. By executing early milestones, workers and requesters can build common ground and adjust the task description. This avoids having workers "run away" to do large amounts of work before realizing it was unnecessary, as they had not agreed on the specifics or covered all edge cases. It also facilitates discussing cost and time to do a job. Daemo requires that each task begin with a short milestone, a *prototype task*. This prototype task is a small percentage of tasks for microtasks, and a first step toward the larger goal for macrotasks. Through the prototype task, requesters can: (1) identify the most suitable workers for their task; (2) be assured that workers understand the task; and (3) directly discuss with workers how the task might be improved.

Open Governance

While the mechanisms of crowd work have evolved, the asymmetrical power dynamics of workers and employers remained unchanged. Daemo addresses the power imbalance and mitigates the inherent trust issues by introducing a representative democratic governance model that elects a leadership board composed of three workers, three requesters, and a researcher. This leadership board is empowered to make policy decisions for the platform. Including all vested parties in the governance of the platform provides an opportunity for idea transfer, transparent communication, and engagement in platform direction.

Over the course of a three-week period, we conducted an experiment within our large research cohort to rapidly prototype three organizational models to guide Daemos design and development: representative democracy, participatory democracy, and weighted democracy (participatory democracy with each vote weighted by participants' reputation within the system). The researchers within the project represent a diverse population from 26 countries with ages ranging from 17 to 48. We sought to assess various aspects of the models including: communication and responsiveness, participation levels, ease of bringing an idea to execution and transparency of information and process. We captured overall participation including unique number of participants, total voting volume, sentiment regarding each model's effectiveness in instilling trust and power, as well as number of executed ideas. The

leadership board election generated 255 votes (206 to elect representative, 49 on ideas), 31 ideas offered, and 2 actions; participatory democracy generated 52 votes, 33 ideas offered, and 1 action; weighted democracy generated 19 votes, 16 ideas offered and 0 actions. Despite possible temporal or novelty effects, the effects are quite strong. Post experiment survey results indicated that 64% of respondents identified the leadership board as the desired approach, supplemented by an open platform for idea submission and a central leadership board to facilitate research, decision, and execution of submitted ideas.

FUTURE USABILITY EVALUATION PLAN

We plan to evaluate our proposed model by recruiting requesters and workers to exercise our system. A crucial aspect of our prototype task model is that we have added an option for a feedback textbox to the task creation module. This should facilitate worker and requester collaboration towards improved task descriptions. In addition, by having workers and requesters interact early in the process, worker selection should be improved, and we anticipate increased trust between the two parties. We will test the usability of our task workflows using heuristic evaluation and direct observation. We have already performed some preliminary tests on the task creation workflow. Our subjects found some parts of the task formulation to be confusing. In the future, we will add more instructions to guide the user as they create the task.

FUTURE WORK AND CONCLUSION

Daemo envisions a future of crowd work that is built around trust rather than antagonism. We have no illusions that requesters and workers will stop maximizing their individual utility; however, we believe that targeted contributions in task design, reputation, and representation can lead users to assume better of each other, rather than markets for lemons. Our next step is to create incentive-compatible reputation systems, such that ratings become more informative (and so that not every requester and worker has 95% approval and 4.75 stars). Ultimately, we aim to inspire current crowd marketplaces to adopt alternative visions, or achieve a foothold ourselves in the crowd work ecosystem. We believe that we can achieve improved task quality and fairness with our augmented task workflow and an open governance model.

REFERENCES

1. Kittur, A., et al. The future of crowd work. In *Proc. CSCW 2013*. 2013.
2. Le, J., Edmonds, A., Hester, V., and Biewald, L. Ensuring quality in crowdsourced search relevance evaluation. In *SIGIR workshop on crowdsourcing for search eval*. 2010.
3. Martin, D., Hanrahan, B.V., O'Neill, J., and Gupta, N. Being a turker. In *Proc. CSCW 2014*. 2014.
4. Salehi, N., et al. We are dynamo: Overcoming stalling and friction in collective action for crowd workers. In *Proc. CHI 2015*. 2015.