

Massive Multiplayer Human Computation for Fun, Money, and Survival

Lukas Biewald

CrowdFlower
3265 17th Street, Suite 302
San Francisco, CA 94110
lukas@crowdflower.com

ABSTRACT

Crowdsourcing is an effective tool to solve hard tasks. By bringing 100,000s of people to work on simple tasks that only humans can do, we can go far beyond traditional models of data analysis and machine learning. As technologies and processes mature, crowdsourcing is becoming mainstream. It powers many leading Internet companies and a wide variety of novel projects: from content moderation and business listing verification to real-time SMS translation for disaster response. However, quality assurance can be a major challenge. In this paper CrowdFlower presents various crowdsourcing applications, from business to ethics, to money and survival, all of which showcase the power of labor-on-demand, otherwise known as the human cloud.

Author Keywords

Crowdsourcing, quality, Mechanical Turk, democracy, data, labor, human

ACM Classification Keywords

H.5.2 Information Interfaces and Presentation: Miscellaneous—*Crowdsourcing*

General Terms

Design, Performance, Experimentation, Human Factors, Performance, Algorithms

INTRODUCTION

Before the Internet enabled human beings to connect as they do now, collecting large-scale datasets that require human computation was a time-consuming and expensive process. At CrowdFlower[2], we produce new datasets on-demand by routing tasks to large groups of distributed workers who work simultaneously. We see people collecting creative and innovative datasets for businesses, for fun, and even to improve the lives of others.

Topics of study in disciplines that focus on quantitative or

technical data, like machine-learning research, have always been limited by the availability of datasets. For example, the Brown Corpus is a dataset compiled in the 1960s that has served as the basis for thousands of linguistics studies. It has been exhaustively parsed and tabbed. Graduate students would center entire research plans on the availability of previously collected data, and as a result, generations of papers on word disambiguation were tailored to the constraints of old data.

Crowdsourcing democratizes the data-collection process, cutting researchers' reliance on stagnant, over-used datasets. Now, anyone can gather data overnight rather than waiting years. However, some of the data collection may be sloppy. CrowdFlower addresses this issue by building robust quality-control mechanisms in order to standardize the results that come back from the crowd. The type of crowd, task design, and quality control tactics all affect the quality of the data. The important thing to remember is that crowdsourcing provides channels that allow researchers, businesses, or even armchair social scientists to gather data having high quality data obviously affects the accuracy of the research.

The first time I used Amazon's Mechanical Turk[1] was at a search-engine startup, Powerset[6] (later acquired by Microsoft). I used Mechanical Turk to compare the quality of our search-relevancy algorithm against Yahoo! and Google. Initially, I thought it would be necessary to hire a team of people to compare the quality of results every day over the course of months. Instead, I set up an experimental task with no quality control, put in about \$50, and let it run overnight. The data that came back was noisy, but I was able to find meaningful differences between the search engines.

The following examples showcase the role that crowdsourcing plays in data collection. Many of these are featured on our blog, and we often post new datasets and projects. We invite researchers to post new experiments on our site.

CROWDSOURCING APPLICATIONS

Ethics

Crowds can be used to source answers to philosophical questions. Stalin said, "A single death is a tragedy; a million deaths is a statistic." So what about 100 deaths? What about five? We tested this experimentally by asking people on Amazon Mechanical Turk to participate in the classic philosophical conundrum "The Trolley Problem,"[8] in which a

Permission to make digital or hard copies of all or part 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. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CHI 2011, May 7–12, 2011, Vancouver, BC, Canada.

Copyright 2011 ACM 978-1-4503-0267-8/11/05...\$10.00.

person must decide whether to sacrifice one person in order to save several others.

The three sample scenarios are below. Here, the scenarios refer to saving five people, but we varied the number of people saved between 1 and 1,000 to see if it would affect the results.

Scenario A. A trolley is running out of control down a track. In its path are five people who have been tied to the track. Fortunately, you can flip a switch, which will lead the trolley down a different track to safety. Unfortunately, there is a single person tied to that track. Should you flip the switch?

Scenario B. As before, a trolley is hurtling down a track toward five people. You are on a bridge under which it will pass, and you can stop it if you drop something heavy in its way. As it happens, there is a very fat man next to you. Your only way to stop the trolley is to push him over the bridge and onto the track, killing him to save five. Should you proceed?

Scenario C. A brilliant transplant surgeon has five patients, each in need of a different organ, each of whom will die without that organ. Unfortunately, there are no organs available to perform any of these five transplant operations. A healthy young traveler, just passing through the city the doctor works in, comes in for a routine checkup. In the course of doing the checkup, the doctor discovers that his organs are compatible with all five of his dying patients. Suppose further that if the young man were to disappear, no one would suspect the doctor. Should the doctor sacrifice the man to save his other patients?

How does our decision change based on the number of people who will die? The results were unexpected. For all three scenarios, subjects were increasingly willing to kill one person when it meant saving up to 100 people, but this willingness dipped when it meant saving between 100 and 200 people, and began to rise again when more than 200 people could be saved.

Each scenario also affected subjects' ethical calculus. Subjects were more willing to sacrifice a life when they controlled the trolley switch (Scenario A) than when someone else (i.e., the surgeon in Scenario C) acted as executioner (see Figure 1). By putting this experiment out to the crowd, we were able to gather responses from more than 100 subjects in a matter of hours[3].

Business

CrowdFlower provides a way for businesses to complete large tasks by giving them access to tens of thousands or hundreds of thousands of workers. By layering our technology across multiple labor channels, CrowdFlower enables access to a large sample of people. This means businesses can connect to a labor force in the cloud depending on their needs.

Just as cloud computing eliminated the need for businesses to predict how many servers they were going to need at any

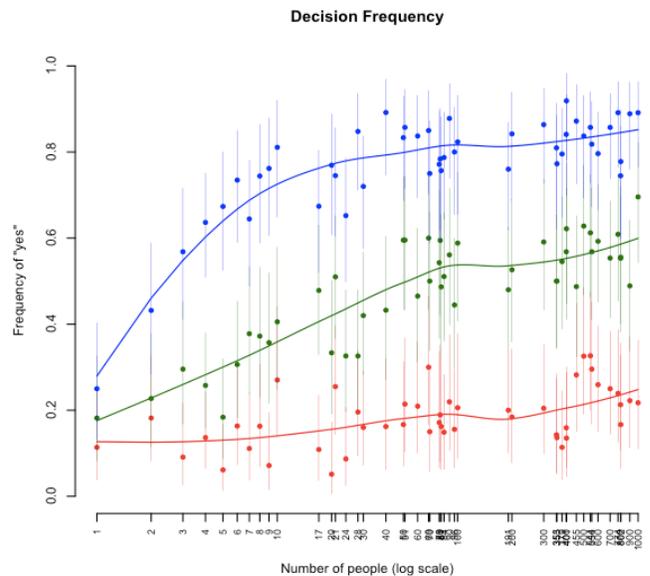


Figure 1. Willingness to sacrifice a life

given time, crowdsourced labor eliminates the need for businesses to predict how many people they're going to need at any given minute. The quality-control algorithms we've developed ensure that workers are doing the tasks well, that they are trained when they make mistakes, and that they are removed if they are inefficient or spammy.

CrowdFlower has been able to assist clients in many ways, such as improving the accuracy of data for clients who publish business listings: e.g., business leader information (CEO and CFO names), contact information (phone, fax, and address), and company information (company name, description, and industry label).

In particular, our use of quality-controlled human intelligence excels at removing duplicate business listings that a computer program would gloss over. For example, a computer cannot say for certain whether the McDonald's restaurant in Manhattan and the McDonald's restaurant in New York City are the same place, but a person could do the additional research required to confirm whether the two restaurants are the same. In one instance, we improved the accuracy of a client's data by more than 20 percent.

The Greater Good

Crowdsourced data collection will continue to benefit businesses and revolutionize academic research. It can also benefit disenfranchised people by giving them access to dignified work.

For starters, the microtasks involved in crowdsourced labor mean that anyone anywhere can be productive, even if just for 10 minutes. They don't have to kill time playing solitaire at their computers or working on cross-words. Instead, they could be paid to do a job, earn virtual credits for an online game, or even give work to people in developing nations

through a paired work program.

The developing world needs employment opportunities that crowdsourcing can provide by connecting women, youth, and refugees living in poverty to computer-based work. One of CrowdFlower's labor channels, Samasource[7], does just that.

CrowdFlower and Samasource created GiveWork[5], an iPhone app that allows users to support people in developing countries by completing short, on-screen tasks, which can either give a donation or an additional unit of work that is used for training purposes. People volunteer to tag a video or trace a road alongside someone who is learning computer skills. This is just the beginning for mobile-based crowdsourced labor.

Further, getting work to refugee populations is difficult whenever that work requires raw materials: e.g., building physical structures; but projects that require building information can move quickly and globally.

Crowdsourced labor has begun to level the playing field with respect to job access. This lends greater meritocracy to the job market: it is a natural extension of what the Internet has already done, but rather than "who you know," the focus is on "what you know." Thus, a person in Berlin or Jakarta is not immediately ruled out of a job due geography.

Disaster Relief

Crowdsourcing also helps disaster relief efforts because disasters are, by definition, unpredictable, and relief requires a scalable workforce. The response to the Haitian earthquake in January 2010 demonstrated how a rapidly deployed workforce of far-flung volunteers can be critical to emergency relief efforts. After the quake, aid workers flooded Port-au-Prince, but they lacked information about who needed help, where the victims were, and what type of help they needed. FrontLineSMS[4] and the U.S. Department of State worked with Haitian telecom companies to set up an SMS short code, allowing Haitians to submit real-time reports using less bandwidth than the two-way audio that had caused system outages on the country's cell networks. The messages that came in were in Kreyol and the aid agencies were unable to translate the messages fast enough.

CrowdFlower provided the infrastructure to route SMS texts to hundreds of thousands of Haitians (located by Samasource) who translated texts from Port-au-Prince in real time and categorized the victims' issues, allowing the agencies to direct specialists to the people who needed their services: e.g., getting potable water to thirsty people, routing doctors to injured people. Further, so that agencies could see hotspots, maps were created through Ushahidi[9], an open-source platform that allows people or organizations to collect and visualize information.

The rapid proliferation of broadband, wireless, and cell phone technology has revolutionized disaster relief efforts so that now, anyone with a computer or phone can provide assis-

tance.

CONCLUSION

The Future of Democracy

What does the future look like for crowdsourcing, human computation, data exchange, and data transparency? When data is made widely available, it becomes widely analyzable, and through this process, crowdsourcing can empower us all.

BIOGRAPHY

Lukas Biewald is the CEO of CrowdFlower. Prior to co-founding CrowdFlower, he was a senior scientist and manager within the ranking and management team at Powerset, Inc., a natural language search technology company later acquired by Microsoft. Biewald has also led the search relevance team for Yahoo! Japan. He graduated from Stanford University with a BS in mathematics and an MS in computer science. He is an expert level Go player.

REFERENCES

1. Amazon Mechanical Turk. <http://mturk.com>.
2. CrowdFlower. <http://www.crowdflower.com>.
3. Crowdsourcing an Ethical Dilemma. <http://blog.crowdflower.com/2009/01/crowdsourcing-ethics/>.
4. FrontLineSMS. <http://www.frontlinesms.com>.
5. GiveWork. <http://itunes.apple.com/us/app/give-work/id329928364?mt=8>.
6. Powerset. <http://www.powerset.com>.
7. Samasource. <http://www.samasource.org>.
8. The Trolley Problem. http://en.wikipedia.org/wiki/Trolley_problem.
9. Ushahidi. <http://www.ushahidi.com/>.