What the crowd means for research, data gathering and analysis

Position Statement

Timothy Olsen
W.P. Carey School of Business
Arizona State University
Tim.olsen@asu.edu

Introduction

Many people have seen videos of "Flash Mobs" which gather in a public place and perform a dance or musical number and then quickly vanish. In a similar way, the Internet has enabled on demand access to labor and talent - accessed through open calls on several platforms. The open nature of the Internet has enabled access to new forms of work that we are we only beginning to understand. This area is ripe with challenging and worthwhile research questions, and has enabled an exciting new means for data analysis and gathering.

Research Opportunities

Truly remarkable feats of labor have been accomplished via crowdsourcing models. For example, volunteers indexed over 1 billion names historical census records and ship manifests in less than six years (Maxfield 2013). Five million customer reviews were proofread for spelling and grammar errors in a few weeks for Zappos (Kaganer et al. 2013). 36 million forms containing personal data were indexed by Deutsche post in only six weeks. These amazing feats couple with industry consolidation points to the maturating of a new industry. However, academic research in this area is thus far confined to the human computation sub discipline of computer science, and econometric research involving platform design. There are many more research areas witch continue to be a veritable "blue ocean" waiting to be discovered. I will discuss just two research areas and then talk about how the crowd can be aid in the research process.

Global Development

Much of the labor force on crowd work platforms comes from the developing world. No longer are labor forces in developing nations have been limited to regional work. They are now able to contract directly with businesses or perform work under the purview of labor exchange platforms. It is not surprising that the largest crowdsourcing companies are based in China, or that the government of Malaysia is developing a crowdsourcing platform to bring economic opportunities to their citizens. This gives rise to a few interesting research questions:

What is societal impact of crowdsourcing on communities in developing nations?

How does crowdsourcing affect the career development of workers in the developing world?

What policies around crowd labor can countries adopt to encourage sustainable economic development?

Outsourcing

Today Business Process Outsourcing (BPO) is a $500 billion industry. However, only a couple decades ago this industry was virtually non-existant while the landscape was dominated by IT outsourcing. This reflects the move from *projects* to *processes*. Likewise, most of the work that has been performed by crowdsourcing has been project based. Only recently have we seen compelling examples of Business Process Outsourcing performed by the crowd. Large BPO companies are researching how they can incorporate techniques from crowdsourced labor models. This gives rise to many research questions including the following:

> What processes are best sourced through crowd labor?

> How can work be best broken down for those who will perform it?

Complex knowledge work can also be automated and performed by the crowd. Recent research has shown how complex work such as developing a mobile app, or a short video animation can be performed in less than 24 hours, and managed by the crowd itself (Retelny 2013). Individuals in the crowd are selected to break down tasks into smaller pieces and assemble a "flash team" which works together to complete the work. The process each team takes to perform this work can be studied individually and thousands of processes can be "mined" to understand the best combination of tasks and activities. The Business Process Management literature can inform this work.

Data Gathering and Analysis

While the ability to involve the crowd in data analysis has matured rapidly, techniques for involving the crowd in data collection are immature.

Data Collection

Google and Facebook have companies have acquired valuable "private data" by offering useful services to billions of people for free. In contrast to "private data" how can a researcher who seeks valuable data, acquire it by offering a useful service to the public to encourage them to contribute to "open data"? An interesting example of this is premise.com, which creates a real-time inflation index in over 30 countries by employing 700 people to take photos of prices in local markets around the world. How can data like this be gathered "for free" in exchange for a useful service for the consumer? Platforms have been developed to aid in data collection but without giving value to the submitter of the data they do not seem to receive much use (Aanensen et al. 2009) .

Data Analysis

Many platforms exist for volunteers who wish to expunge their cognitive surplus or receive a hedonic benefit to help with data analysis. There are many well-known examples of how the general public can be involved in the data analysis phase. Galaxy Zoo, NASA click-workers, Zooniverse. Two exciting platforms that will enable any researchers to request help in data analysis include Curio (crowdcurio.com) and Crowdcrafting.org. I am not aware of any researchers from the business disciplines using either of these platforms for data analysis.

References

Aanensen, D. M., Huntley, D. M., Feil, E. J., al-Own, F., and Spratt, B. G. 2009. "EpiCollect: Linking Smartphones to Web Applications for Epidemiology, Ecology and Community Data Collection," *PLoS ONE* (4:9), p. e6968.

Kaganer, E., CARMEL, E., HIRSCHHEIM, R., and OLSEN, T. 2013. "Managing the Human Cloud," *MIT Sloan. Management Review* (Winter).

Maxfield, L. 2013. "FamilySearch volunteers to hit 1 billion record milestone | ksl.com,".

Retelny, D. 2013. *Expert Crowdsourcing with Flash Teams | CrowdConf 2013 poster*, Presented at the CrowdConf 2013.