

# Wanted: More Nails for the Hammer

## An Investigation Into the Application of Human Computation

Elizabeth Brem<sup>1</sup>, Tyler Bick<sup>1</sup>, Andrew Schriner<sup>2</sup>, Daniel Oerther<sup>1</sup>

<sup>1</sup>Missouri University of Science and Technology, Rolla, MO 65401, <sup>2</sup>University of Cincinnati, Cincinnati, OH 45221  
ebvd3@mail.mst.edu, teb4y8@mst.edu, schrinaw@gmail.com, daniel.oerther@mst.edu

### Abstract

Researchers studying human computation and crowdsourcing know it as a useful tool for solving data analysis problems. Outside of this community, however, we ask the question ‘how many people with data analysis problems know how to use human computation and crowdsourcing?’ In this paper, we observe trends in academic papers on the topic of human computation and crowdsourcing, conduct a survey, and examine the growth of the human computation and crowdsourcing market by researching recent market reports in order to test the hypothesis that human computation and crowdsourcing is not a widely used tool despite its usefulness for problem solving.

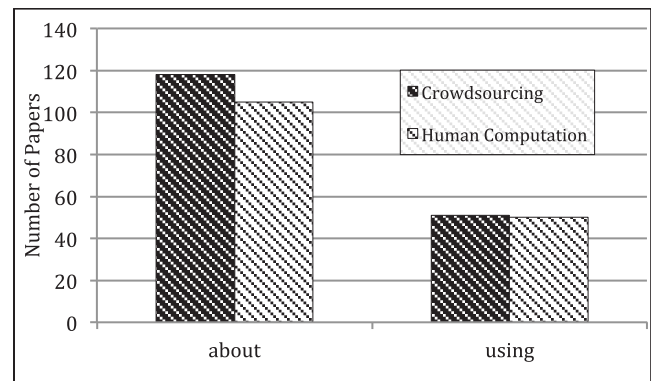
### Background

We begin by taking a look at market trends within the private sector. Crowdsourcing within the private sector has been shown to indicate a ‘high growth, early stage industry’ with both revenue and worker numbers increasing rapidly (Massolution, 2012). The crowdsourcing industry includes highly specialized tasks as well as microtasks. Microtasks make up a significant portion of crowdsourcing platforms within the private sector. Microtask growth may be hindered by a number of different mechanisms including but not limited to crowd responsiveness, ease of use, and lack of satisfactory results (Frei, 2009). Microtask growth could also be hindered by a lack of knowledge on how to use this tool.

### Methods and Results

The utilization of crowdsourcing can also be gauged in academia. Trends in academic papers can indicate whether there is more work using human computation and crowdsourcing to solve problems, or defining the mechanics of the two. Papers were collected using two

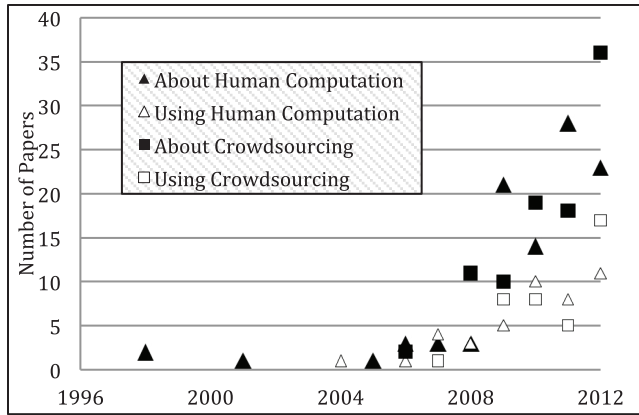
keywords, “human computation” and “crowdsourcing”, to search four academic literature search engines. Using each keyword, the first 50 papers from each search engine were collected using its default ranking. Duplicate papers in the search engines were removed, leaving a collection of unique papers. They were further sorted by content that was “about” the tool or “using” the tool. This was done by a single labeler who read the abstract to the papers. “Using” crowdsourcing or human computation refers to a paper that is primarily about *doing something interesting* with human computation or crowdsourcing. The main intellectual contribution is the new thing that was done. “About” crowdsourcing or human computation refers to a paper that is primarily about some features of human computation as a tool, such as how it works, or a new way to do something, or some behavior of the humans, etc. The main intellectual contribution is characterizing the tool. This was to provide evidence for whether academia was either utilizing the tools in new ways, or characterizing the tools. The resulting sorted papers are shown in Figure 1.



**Figure 1:** Number of unique papers collected using two keywords, “human computation” and “crowdsourcing”, that have been labeled as either “about” or “using”.

This representation clearly shows the majority of papers published being about crowdsourcing and human computation, rather than papers utilizing crowdsourcing and human computation.

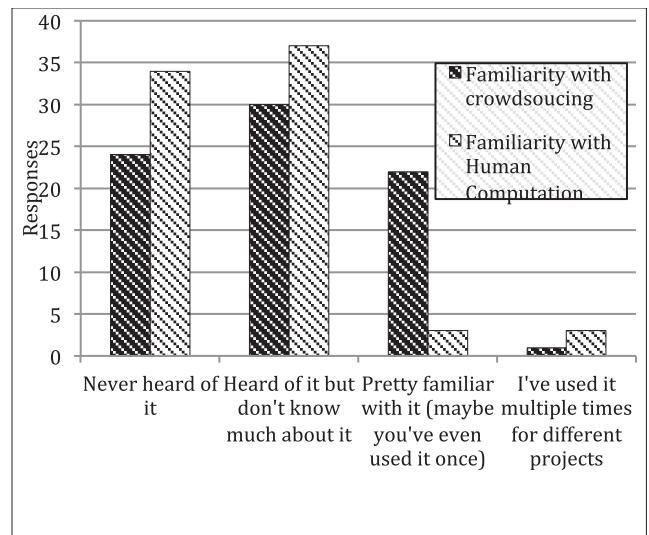
The papers were then organized by dates in their respective groups (i.e. using and about) to visualize the trend of human computation and crowdsourcing growth over time. The results can be seen in Figure 2.



**Figure 2:** The trend of human computation and crowdsourcing growth over time.

This representation shows once again that the majority of papers published are about crowdsourcing and human computation, rather than papers utilizing crowdsourcing and human computation. Furthermore, this shows rapid growth in the human computation and crowdsourcing field.

We then surveyed people whose work involves data analysis in order to determine to what extent people use crowdsourcing and/or human computation in their work. The survey was distributed by people within our own networks. A total of 98 surveys were received. The self-reported results were filtered based on answers to the question of how much data analysis was conducted by the individual. Only the results from individuals who regularly or frequently conduct data analysis were retained, for a total of 77 surveys retained. The data shows that very few people performing data analysis regularly use crowdsourcing. Even fewer respondents have used human computation. A graph of the results showing the number of respondents and how familiar they are with human computation and crowdsourcing can be observed in Figure 3.



**Figure 3:** The number of survey respondents and how familiar they are with human computation and crowdsourcing.

## Discussion

While the crowdsourcing industry may be rapidly growing, human computation, in the form of microtasking, may be growing slower. One possible explanation for the discrepancy of industry and academics is that we may be at the beginning of an S-curve commonly used to describe technological growth. Another reason is that it may be hard to quickly teach others how to break down a large problem into easily computed microtasks. A lack of easy-to-use programming tools may obstruct many people from using human computation in their data sets. The survey respondents consisted of mainly engineering related professionals and academics. These people may have large data sets but do not have an intuitive tool to bring together human computation and machine-based computation. This may account for the reason that many companies in the crowdsourcing industry are seeing growth while professionals in other industries, who may have a use for human computation, have yet to fully realize human computation and crowdsourcing's potential.

## References

- Massolution. 2012. *Enterprise Crowdsourcing: Market, Provider, and Worker Trends*. Crowdsourcing.org
- Frei, B. 2009. *Paid Crowdsourcing: Current State and Progress Toward Mainstream Business Use*. Smartsheet.com