

In this forum we will celebrate research that helps to successfully bring the benefits of computing technologies to children, older adults, people with disabilities, and other populations that are often ignored in the design of mass-marketed products.

Juan Pablo Hourcade, Editor

What the Disability Community Can Teach Us About Interactive Crowdsourcing

Jeffrey P. Bigham

University of Rochester | jbigham@cs.rochester.edu

Richard E. Ladner

University of Washington | ladner@cs.washington.edu

“Can you tell me what this says?” You might be faced with this question in any number of contexts:

Maybe you’ve gone to lunch with a blind friend who is deciding what to order off of the menu. Maybe you’ve volunteered to help read mail or the local newspaper with an older person who has recently lost her vision. Maybe your job is to process mail-in forms, deciphering the sloppy handwriting of people looking for a new credit card.

Maybe you spend extra minutes in the evenings on Mechanical Turk making a bit of money helping to make OCR more reliable. Or maybe you’re proving that you’re human by solving a reCAPTCHA as part of the sign-up process of a new Web service, and helping to digitize an old book in the process. In these situations and many more, people are often called upon to interpret and decipher the world for others.

One of the keys to making information accessible to everyone is converting it from one form to another. For example, inaccessible visual text may be made accessible to a blind person by converting it to aural speech, and, conversely,

aural speech may be made accessible to a deaf person by converting it to text. Automatically interpreting sensory information is notoriously difficult, despite tremendous progress in the past few decades. Far from idling while waiting for artificial intelligence to catch up, people with disabilities have successfully been getting answers to questions about their environments all throughout history—by *crowdsourcing* them.

At a high level, crowdsourcing means using the collective human intelligence of often anonymous workers toward some coordinated aim. Recent projects like Wikipedia immediately come to mind, but the idea has been realized in the disability community for years. For instance, a volunteer may sign up to offer a few minutes of her time to read a blind person’s mail aloud, or a fellow traveler may answer a quick question at the bus stop, for example, “Is that the 45 coming?” Professional workers, such as sign language interpreters and audio descriptionists, interpret and convert sensory information into alternative forms, enabling,

respectively, a deaf student to participate in a traditional lecture and a blind person to enjoy (or learn from) a movie. Human support is drawn from a large group of people when needed and contributes to the larger goal of making the world more accessible for people with disabilities.

The accessibility barriers that people with disabilities face have made them leaders in the crowdsourcing trend currently sweeping the computer sciences, although their leadership has gone somewhat unnoticed.

Early Adoption

People with disabilities have always solicited the assistance of others, often friends and family, to make accessible what their own senses could not. Blind people found readers to relay written correspondence, and deaf people found volunteer interpreters. These volunteers were often just members of their local community—for example, members of a religious congregation who knew some sign language would provide ad hoc interpretation of religious services.



► Figure 1. VizWiz works in three steps: a blind user takes a picture and speaks a question, which are forwarded to the crowd for answering, and then answers are returned to the device from the crowd.

In the beginning, these accessibility-related services were usually informal, but over the years they have evolved into crowdsourcing organizations. For example, a number of agencies now provide volunteer or paid workers for various tasks, including sign language interpretation and real-time captioning for people who are deaf or hard of hearing, personal assistance for those who have severe mobility disabilities, reading support for those who are blind, and support services for those who are deaf-blind. This evolution came about due to the need for friends and family to return to their roles as friends and family, and because, in many cases, trained volunteers and professionals (experts) can do the job better.

Recruiting human labor in this way differs somewhat from conceptions of crowdsourcing that are popular in IT circles, but it fits the definition above: When an agency

provides a service, the customer usually does not know who will actually provide the service. The agency has a pool of providers (crowd) who can be assigned to tasks according to their skills and availability. These workers often agree to work under strict rules of quality and confidentiality. We believe these early services may also foreshadow coming trends in crowdsourcing services.

An Expanding Crowd for People with Disabilities

Technology for people with disabilities has made large crowds easy to access anywhere. A particularly interesting case is sign language interpreting. In just the past 10 years, remote sign language interpreting has become ubiquitous. There are at least two forms of remote sign language interpreting: Video Relay Service (VRS) and Video Remote Interpreting (VRI). In VRS, a skilled sign language

interpreter translates a phone call between a sign language user and a hearing person, while in VRI the interpreter translates a face-to-face interaction between a sign language user and a hearing person. In both cases the interpreter is at a site remote from the two people trying to communicate. With the more mature VRS, when a phone call is requested, an interpreter from a pool of interpreters is assigned to the call, usually within a few seconds, and the call is set up with minimal delay. No prior scheduling of a call is needed. One VRS company, Sorenson VRS, employs thousands of sign language interpreters. Before VRS, going back to the 1970s, the deaf community used TTY Relay Services in the same way, except there was no video, just texting over telephone lines. TTY Relay Service operators translated text from the deaf customer to speech and speech from the hearing cus-

tomers to text. Generally, TTY Relay Service operators need far less training than VRS interpreters, who must be fluent in two very different languages. As you can see, members of the deaf community are already accustomed to using remote, near-real-time interactive crowdsourcing—something rarely seen in today's more mainstream crowdsourcing applications.

Interactive Crowdsourcing Anywhere

Two recent trends have expanded the crowd of possible workers even further and made crowd work easier to access from wherever a user happens to be. First, mainstream mobile phones with low-latency, high-bandwidth connections and a wealth of sensors (camera, microphone, GPS, among others) have become commonplace, obviating the need for special hardware and making communication on-the-go faster. Second, marketplaces for small jobs like Mechanical Turk and social networks like Facebook and Twitter have grown in popularity, providing large pools of potential workers already connected and available in near real-time.

Given the history of how people with disabilities have employed crowd work and the more recent trends mentioned earlier, it's no surprise that people working with those with disabilities were quick to capitalize on the latest crowdsourcing technologies. An early example is the Social Accessibility Project out of IBM Japan, which aims to pair blind Web users with Web-based helpers who can assist them with Web-accessibility problems. A service called Solona provides a CAPTCHA-solving service for blind users. oMoby is an iPhone application created by IQ Engines, which uses both crowdsourced and

computer vision to interpret images. Although originally created as a mainstream demonstration of the quality of its API, oMoby has quickly been adopted by blind users.

We have been developing our own mobile tools that connect people with disabilities to remote crowd workers in near-real time. As one example, our VizWiz iPhone application (see Figure 1) enables a blind user to take a picture, speak a question, and have it answered by workers on Amazon's Mechanical Turk quickly and cheaply (in a field deployment, answers came back in less than 30 seconds and for less than 7 cents). Our accompanying quikTurkit software toolkit helps improve the response time of Mechanical Turk using several strategies, primarily by prequeuing multiple workers in advance of receiving the question. In a week-long deployment to 11 blind people, we gained a fuller appreciation of the types of questions that blind people might want answered—not surprisingly, the questions they asked went far beyond wanting to know what text said. Users also have expectations of VizWiz; for example, once they received an answer, many of our users wanted to ask follow-up questions. Our ongoing work involves exploring how users perceive different sources of human computation along a number of qualitative dimensions.

Lessons for Mainstream Crowdsourcing

Most efforts in mainstream crowdsourcing have thus far concentrated on achieving acceptable results from crowd labor (high quality, low cost, and so on), but it's becoming evident that we also must think about the broader ecosystem in which crowdsourcing happens. Because people with disabilities

have come upon these issues for some time, it makes sense to look to their experiences for guidance. As one example, the expectations of workers are usually considered only to the point at which they are helpful for receiving better work. How can we give appropriate feedback to workers to let them know the implications of their work? If a worker in VizWiz is asked to decipher dosage information on a pill bottle, they might reasonably choose to skip the task if they feel they can't do it well. If many workers pass over the job, what does it mean when someone else completes it? What if that pill bottle contains the name and address of the user?

New services leverage huge crowds of largely anonymous workers. An open question is how to enable appropriate privacy and anonymity protections in a setup like this. Privacy is not always completely achievable because in the process of doing work for a user, the worker may learn information about the user. A worker who is reading the label on a medicine bottle may see the name of the user on the label, or a sign language interpreter could see who he or she is interpreting for and recognize the user. Thus, the privacy of the user may be maintained by only a worker who adheres to a strict confidentiality policy.

Historically, most services for people with disabilities have adopted strict codes of confidentiality to deal with situations like this. As an example, sign language interpreters have a code, laid out by their professional organization, the Registry of Interpreters for the Deaf, that prevents them from interjecting their own comments into the conversation and from repeating information in conversations that they have interpreted. In

fact, all the major professions that provide services to people with disabilities have developed codes of ethics that include confidentiality, respect for the customer, and responsibility to take on only jobs for which he or she has the necessary skills.

Informing the Design of Crowdsourcing Systems

The experiences of people with disabilities can inform the design of current crowdsourcing systems. The following are some of the issues that people with disabilities have addressed in the past and future systems will need to consider.

Confidentiality and anonymity.

When humans are included in the loop in interactive systems, how can the system make guarantees regarding confidentiality and anonymity? Remote interpreting services for the deaf community require workers to agree to strict confidentiality rules, and oversight helps to ensure that workers comply. Perhaps we can create an analogous system on top of the anonymous crowd. Absent that, how can interfaces transparently relay expectations of confidentiality to users so they can make informed decisions? There also may be ways to engineer interactive systems that use crowdsourcing to minimize the need for confidentiality to maintain user privacy.

Worker competence. When a worker takes on a particular job, is the worker competent to do the job well? Some jobs may be so specialized that the worker may be required to have a form of certification before being allowed to accept the job. For instance, sign language interpreters are prescreened for competence, but this reduces the pool of avail-

able workers who might be able to help. Could the interactive system enable users to evaluate worker competence? How do we ensure worker competence?

Latency. Different sources of human computation may have different expected latencies. How can systems reduce latency? What types of expected latencies are appropriate for different types of work? It might be acceptable to wait a few hours (or even days) for a volunteer to read your mail, but interpreting needs to happen right away. How can systems help users understand and make decisions based on expected latencies of different sources of human computation?

Accuracy. Human computation can provide incorrect answers for a number of reasons, including workers misunderstanding the question, malicious workers, or underspecified questions. How should systems attempt to ensure accurate answers and help convey good estimates of answer quality to users?

Feedback to users. Providing feedback to users about the human computation that is occurring on their behalf is critical for them to make informed decisions. What information do users want or need? How can systems be created to provide this necessary information?

Interfaces. Many of the areas noted here are dependent on the interface. For instance, the design of the worker interface may lead workers to respond more quickly or more slowly. The interface may reveal more private information about the user who submitted the work, or even encourage more accurate answers. If workers are part of an interactive system, what responsibility do they have for the side effects of their work (for example, giving a disabled user feedback

that causes them harm)? How can the interface convey to the worker the potential side effects of their answers (and potentially their culpability)?

Sources of computation. Users will increasingly face the challenge of deciding between sources of intelligence, both human and artificial. These sources may differ in terms of cost, availability, and all of the qualities listed above (such as latency, accuracy, or privacy). This is a challenge that is relatively new for the disability community—too many sources of workers to help! How should systems convey to users the trade-offs between the different sources of computation currently available to them?

Summary

Crowdsourcing clearly offers incredible potential for an array of new applications that draw from the intelligence of humans. People with disabilities and the people supporting them have been confronting these issues since long before the advent of Web-based crowdsourcing. We believe there is much to learn from their experiences that can be either directly applied or adopted into new mainstream crowdsourcing systems.



ABOUT THE AUTHORS

Jeffrey P. Bigham is an assistant professor in the Department of Computer Science at the University of Rochester, where he heads the ROC HCI Group. His research focuses on better understanding the needs of people with disabilities in order to motivate and develop new approaches to access technology.



Richard E. Ladner is Boeing Professor in Computer Science and Engineering at the University of Washington. His research is in the area of accessible technology for people with disabilities. He is active in computing enrichment programs for students with disabilities at all levels.