

The Design of Human-Powered Access Technology

Jeffrey P. Bigham
ROC HCI, Computer Science
University of Rochester
Rochester, NY 14618

jbigham@cs.rochester.edu

Richard E. Ladner
Computer Science and Engineering
University of Washington
Seattle, WA 98195

ladner@cs.washington.edu

Yevgen Borodin
Computer Science
Stony Brook University
Stony Brook, NY 11709

borodin@cs.stonybrook.edu

ABSTRACT

People with disabilities have always overcome accessibility problems by enlisting people in their community to help. The Internet has broadened the available community and made it easier to get on-demand assistance remotely. In particular, the past few years have seen the development of technology in both research and industry that uses human power to overcome technical problems too difficult to solve automatically. In this paper, we frame recent developments in *human computation* in the historical context of accessibility, and outline a framework for discussing new advances in human-powered access technology. Specifically, we present a set of 13 design principles for human-powered access technology motivated both by historical context and current technological developments. We then demonstrate the utility of these principles by using them to compare several existing human-powered access technologies. The power of identifying the 13 principles is that they will inspire new ways of thinking about human-powered access technologies.

Categories and Subject Descriptors:

H5.2. Information interfaces and presentation: User Interfaces

General Terms

Design, Human Factors

Keywords

Crowdsourcing, Human Computation, Access Technology

1. INTRODUCTION

People with disabilities have always overcome accessibility problems by enlisting people in their community to help. Far from being passive recipients of this assistance, disabled people have a rich history of managing these interactions, and have formed organizing structures around assistance that serve to increase independence and ensure that their expectations are met. Initially, disabled people relied mostly on people near them for assistance, but the increasing connectedness enabled by the Internet has dramatically expanded the pool of people who could help. As disabled people have become more connected via networked devices, their ability to ask for assistance has increased.

The tendency in technical fields is to concentrate on fully automated solutions, but it is clear that, despite tremendous advances over the past few decades, technology alone is still far from being capable of solving many real accessibility problems

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, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

ASSETS'11.

Copyright 2011 ACM 1-58113-000-0/00/0010...\$10.00.

that people with disabilities face in their everyday lives. For example, OCR seems like a solved problem until it fails to decipher the text on a road sign captured by a cell phone camera [33], object recognition works reasonably well until the camera is held by a blind person [15, 28, 38], and the laudable 99% accuracy reported by commercial automatic speech recognition systems [4] falls off precipitously on casual conversation or any time it has not been trained for the speaker [45]. Even the automatic techniques used by the screen-reading software to convey the contents of the computer screen to blind people are error-prone, unreliable, and, therefore, confusing [16, 17, 30]. As a result, the many access technology is used only by people who are technically-savvy [41]. When access technology is unreliable, it is abandoned altogether [22, 24, 35].

Instead of relying on technological solutions, people with disabilities constantly rely on a loose network of friends, family, volunteers, and strangers. 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, e.g., "Is that the 45 coming?" Professional workers, such as sign language interpreters and audio descriptionists [39], interpret and convert sensory information into alternative forms, enabling 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. We loosely define *human-powered access technology* as technology that facilitates and, ideally, improves interactions between disabled people and human assistants.

Examples of human-powered access technology abound, including video relay services that facilitate communication between deaf and hearing people; services that let blind people ask a network of volunteers to improve the accessibility of the web sites they visit; and web sites that collect electronic forms of printed material so that it can be presented in a better way for someone with a print disability. It is clear that people with disabilities have led the way in overcoming technological limitations with human assistance.

Including humans in the loop has been a popular direction in access technology over the past few years [23]. In particular, many projects have looked to the crowd as a source of human power that can be harnessed to improve accessibility (see examples in Section 4), where the *crowd* is loosely defined as a large group of people recruited through an open call [26]. In this paper, we choose to focus more broadly on human-powered access technology, which may range from a single employee paired with single disabled user [40] to groups of disabled users helping one another [11].

Despite the long history of human-powered access technology, a unifying framework with which to evaluate, compare and classify technology in this space is missing. As a result, it is difficult to (i) identify the broader contributions of new work in this space, (ii)

make connections between existing work that may target different groups but share underlying traits, and (iii) reveal gaps in existing work that may be important areas for future research.

Our contributions in this paper are the following:

1. We present a motivating historical account of human-powered access technology that includes examples of how people with disabilities have structured services for human assistance to meet their needs and expectations.
2. We identify and describe 15 diverse examples of human-powered access technology that vary across a number of important dimensions.
3. We isolate 13 design dimensions from our examples that help to characterize human-powered access technology and form a framework by which human-powered access technology can be compared.
4. We use the framework to evaluate and compare several of the existing human-powered access technologies, and propose several future research directions motivated by gaps that we have identified through this process.

2. HISTORY OF HUMAN-POWERED ACCESS TECHNOLOGY

People with disabilities have always enlisted the assistance of people in their immediate community, including friends, family and volunteers, to help them make accessible what their own senses miss or help them act on the world in ways they could not by themselves. Blind people found readers to relay written correspondence, deaf people found volunteer interpreters, and physically-impaired people would ask for assistance moving about or getting items that they needed. These volunteers were often just members of their local community—for example, members of a religious congregation who knew some sign language would often provide interpretation of religious services.

As accessibility-related services matured, they evolved from informal community assistance to formal organizations. For example, a number of agencies now provide services such as 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 demand for services, and because, in many cases, trained volunteers and professionals (experts) can do better than non-expert friends, family, and others.

Historically, many 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 they have interpreted. In fact, all the major professional organizations that provide services to people with disabilities have developed codes of ethics that require that their employees or volunteers adhere to strict confidentiality, respect the customer, and take on only the jobs for which they are qualified.

Technology for people with disabilities has made it possible to access human assistants anywhere. A particularly interesting case is sign language interpreting. In just the past 10 years, remote sign language interpreting has become ubiquitous [36]. 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.

Going back to the 1970s, before VRS, the deaf community used TTY Relay Services in the same way, except there was no video, just texting over telephone lines [40]. TTY Relay Service operators translated text from the deaf customer to speech and speech from the hearing customer to text. Generally, TTY Relay Service operators need far less training than VRS interpreters, who must be fluent in two very different languages.

For those with sensory impairments, a key to improving access is converting information 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. The past few decades have seen remarkable improvements in finding ways to automatically convert one form of information into another, especially, with the help of artificial intelligence. In particular, converting text to speech is nearly a solved problem, converting speech to text works in some settings and, optical character recognition (OCR) can also work very well [33].

Nonetheless, there are even more conversion problems that seem to defy automation that, to date, seem to require human intelligence and, sometimes, human expertise. Far from idling while waiting for artificial intelligence to catch up to human intelligence and expertise, people with disabilities have successfully been getting answers to questions about their environments all throughout history—by *crowdsourcing* them. Notoriously hard problems, like speech to sign language translation, visual graphic to tactile graphic translation, video description, and many more seem to require humans, and, in some cases, the humans are recruited and utilized by automated means.

To address cases in which available automatic technology is not enough, models were proposed in the realm of access technology that combined artificial and human intelligence. For instance, in Zimmerman *et al.*'s Rainbow Model, the first approach would be to use automatic technology but support going out to a staff of experts when it failed [1, 31]. In the past few years, the confluence of two technical trends has changed this landscape again, making it possible to recruit assistance remotely, on-demand, for only as long as needed. First, network connectivity has expanded dramatically in terms of bandwidth, latency, and availability. As a result, it is now practical to request and provide assistance remotely, greatly expanding the community of assistants and the number of people who could potentially ask for assistance. Second, the expansion of connectivity meant that there were new pools of assistants available. Microtask marketplaces for small jobs like Mechanical Turk [9] and social networks like Facebook [5] and Twitter [9] have grown in popularity [6] providing large pools of potential workers already connected and available in nearly real-time.

The history of how people with disabilities have employed crowd work and the more recent trends mentioned above, it is no surprise that people working with those with disabilities were quick to capitalize on the latest crowdsourcing technologies. Whether through remote readers for a blind person or video relay services employed by a deaf person, people with disabilities are already

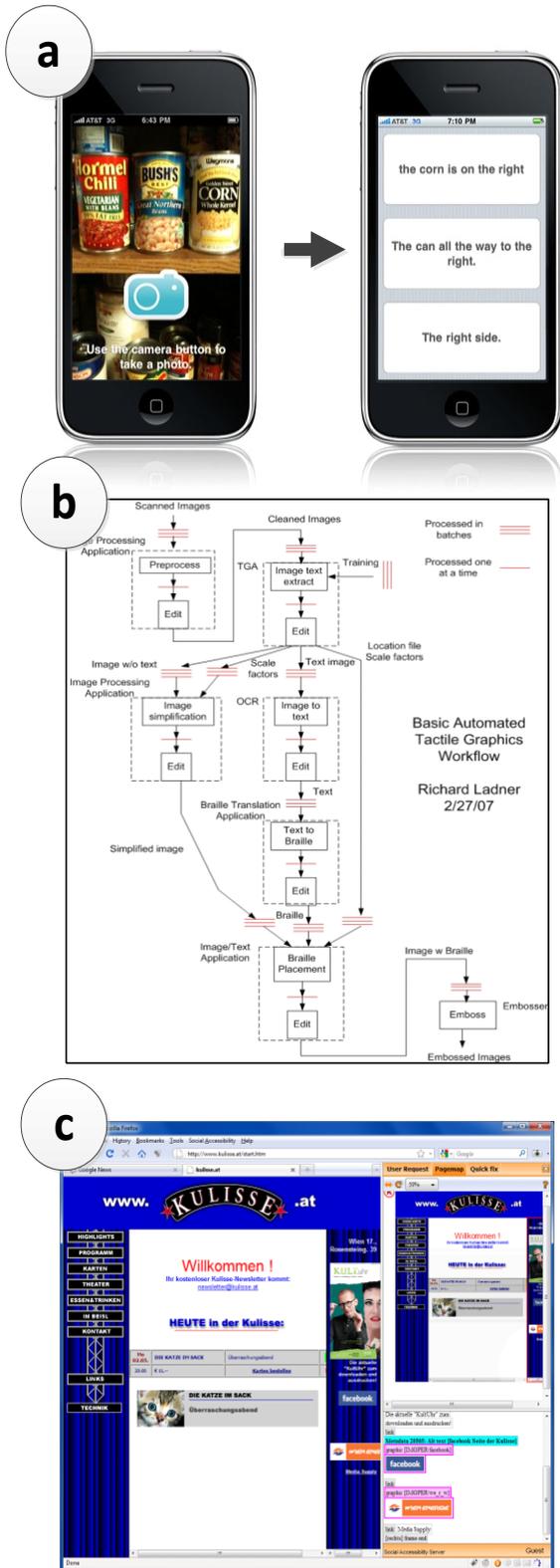


Figure 1: Examples of human-powered access technology. (i) VizWiz is a mobile phone applications that allows blind people to take a picture, speak a question, and receive answers quickly from the crowd; (ii) Tactile Graphics Project supports a workflow in which figures in textbooks are converted to tactile form by a close partnership between a human worker and automated tools; and (iii) Social Accessibility engages a network of volunteers to fix accessibility problems on the web.

accustomed to using and relying on remote, near real-time human assistance in their everyday lives — something that is less common in current mainstream crowdsourcing applications. Consequently, mainstream applications may learn from the lessons learned in the disabled community.

Crowdsourcing and, in particular, human computation have been embraced by many fields in computer science and beyond over the past few years (see [29] for a survey). Human computation includes people as part of computational processes. Although human computation has been developed largely outside of the accessible computing community, one of the original motivations for human computation was an accessibility problem – labeling images [20]. People with disabilities have become leaders in the human computation trend currently sweeping the computer science, inspired by a rich history of using human assistance, although their leadership has gone somewhat unnoticed.

3. EXAMPLES

In this section we describe 15 examples of human-powered access technology. These examples were chosen for the breadth of human support they represent, and will be referenced in the rest of the paper. The examples are roughly ordered from services that operate asynchronously to those that are synchronous.

- 1) **ASL-STEM Forum:** The ASL-STEM Forum is a community-driven portal for building up sign language in technical fields from the bottom-up [19]. Users can contribute signs for terms, and comment/rate the signs contributed by others. It works much like Wikipedia [44], in that people on the web choose what and when to contribute.
- 2) **Tactile Graphics Project:** The Tactile Graphics Project [32] uses a workflow with multiple automatic tools in order to help improve the throughput of a human worker in converting figures in textbooks to a tactile form (Figure 1-b).
- 3) **GoBraille:** GoBraille is a smart phone application that tethers the phone to a refreshable Braille device and supports travel by blind and deaf-blind people on public transportation. One feature of the application is its GoBraille Repository where users can contribute information about bus stops that are relevant to other blind and deaf-blind travelers [11]. This is similar to open street map [8], which is a community-driven site that includes accessible routes.
- 4) **ESP Game:** The ESP Game is an online game in which two remote players are asked to enter labels for images that are shown to them [20]. The players receive points for agreeing on the labels. Because the players do not know one another and cannot otherwise communicate, the best strategy is to enter accurate labels.
- 5) **Bookshare:** Bookshare is a web site that collects and makes available accessible versions of print materials (mostly books) [2]. Members are a primary source of scans, contributing books that they have scanned themselves.
- 6) **Respeaking:** Despite tremendous progress over the past few decades, automatic speech recognition (ASR) does not yet work well enough to accurately transcribe arbitrary speech. A common practice is to have human speakers “re-speak” what they hear [34]. ASR is then applied to this audio instead, which works much better than on the original because the ASR is extensively trained on the speaker and speaks in a controlled environment (quiet room, headset microphone, etc).
- 7) **Social Accessibility:** The Social Accessibility project [21] lets blind web users ask for assistance in improving the accessibility of the Web from sighted volunteers (Figure 1-c).

Improvements generally took on the order of a few hours. This project also highlighted one of the difficulties with user-initiated accessibility improvement – disabled users may often not realize that there is an accessibility problem and never request that it be fixed [37].

- 8) **Remote Reading Service:** The Remote Reading Service was pioneered by the Smith-Kettlewell Eye Institute in the mid-1990s [13, 42]. This service enabled a blind person to send a fax to an expert who would read the contents of the faxed image to a blind user over the phone. This service used existing technology in an interesting new way.
- 9) **Video Relay Services:** Video relay services are companies that have a pool of sign language interpreters who are available to translate a video phone call from a deaf person to a voice phone user. Video relay services began in Sweden in 1997 [36] and are now found in most of the developed countries of the world.
- 10) **Remote Real-Time Captioning:** In remote real-time captioning, a remote captionist converts speech of one person or a group of people into text for someone with a hearing disability. Television captions for live shows and some captioned meetings are done this way. Captioned telephone calls also use pools of captionist to serve phone calls. A common form of remote real-time captioning is called computer-assisted real-time transcription (CART). CART captionists are highly trained to work as court reporters where accuracy is paramount [40].
- 11) **Scribe4Me:** Scribe4Me is a prototype mobile application that lets deaf and hard of hearing people quickly receive transcriptions of the audio events around them [43]. The application runs on existing mobile phones and keeps a buffer of the last 30 seconds of audio heard on the device. When a user wants to do so, they can send the audio off to a remote worker who transcribes the audio and sends it back to the phone. In its original formulation, the transcription was done by a person involved in the study, i.e. an expert.
- 12) **MAP-Lifeline:** MAP-Lifeline enables the distributed support of people with cognitive disabilities as they go about their daily routine [18]. Caregivers are able at their discretion to inject prompts to a number of people under their care.
- 13) **Solona:** Solona is a service that originally let a blind person submit a picture of a web page containing a CAPTCHA and receive the answer to that CAPTCHA [12]. A small number of expert volunteers handle requests, and generally respond in less than 30 minutes.
- 14) **VizWiz:** VizWiz is an accessible iPhone application that lets blind people take a picture, speak any question about it, and receive answers in under 30s from people on Amazon's Mechanical Turk [25]. A primary contribution of the system was *quikTurkit*, a programming library designed to solicit answers quickly (Figure 1-a). LookTel is similar, but connects users to remote operators or family members [27].
- 15) **IQ Engines / oMoby:** IQ Engines provides an API that will provide a label for photos that are sent to it in nearly real-time [7]. This has been put into an accessible iPhone application called oMoby that many blind iPhone users use.

4. DESIGN OF HUMAN-POWERED AT

The experiences of people with disabilities can inform the design of current crowdsourcing systems. The following are some of dimensions that people with disabilities have addressed in the past and that future systems will need to consider. These dimensions are drawn from the examples presented in the previous section.

4.1 Initiative

Initiative refers to who instigates assistance: the end user, to the workers themselves, to organizations, to technology working on behalf of any of these parties. Individuals may be both users and workers, such as in Bookshare and Go Braille.

End User – The end user often decides when to solicit help from human supporters. Examples include services like remote real-time captioning and relay services, and crowdsourcing systems like Social Accessibility and VizWiz.

Workers – Systems like the ASL-STEM Forum and Go Braille allow workers to decide when and what information they will provide. For instance, what terms they will sign in the Forum or what landmarks they will label in Go Braille.

Organizations – Groups of people will sometimes decide to solicit the help of human workers or to guide their efforts. For instance, workers are recruited to work on specific topic areas in the ASL-STEM Forum.

4.2 Source of Human Workers

A number of sources of human workers now exist, from professional organizations and microtask marketplaces to friends and family on social networks. This is a challenge that is relatively new for the disabled community—too many sources of help! Along with the source of workers is associated a number of relevant characteristics, such as the business model that is employed and the sustainability of the service.

Professionals/Experts – Professionals are hired to provide assistance and generally receive wages for their participation. A number of business models may support these workers.

Crowd Workers – Crowd workers are recruited anonymously through an open call, are generally composed of non-experts, and cannot be assumed to stay around for a significant amount of time.

Volunteers – Volunteers work for free, are generally not anonymous, and generally participate for a longer amount of time than crowd workers.

Organized Volunteers – Volunteers are recruited through an organization that may help ensure that people are available when needed, may vet workers, or provide structured training.

4.3 Motivation

As in everything we do in our lives, human workers have a variety of reasons for helping others. These reasons range from altruistic intrinsic motivation to the desire to make money.

Intrinsic – People are often willing to contribute their time for causes they deem valuable. They may be motivated by the fact that their relatives and/or friends have disabilities, or just want to help without expecting anything in return.

Status – Others are better motivated when their contributions are made public through announcement, awards, and banquets. These people can often be instigated by their public status and completion. For instance, ESP game is an example when scoring competitors motivates people to contribute more captions.

Financial – Some people are best motivated by money, either in small amounts for fulfilling tasks at microtask marketplaces or as adequate compensation for their professional services.

Enjoyment – Some systems make it fun to provide assistance, for example the ESP Game turns image labeling into a game.

Self – Users may provide assistance because it helps them. For instance, users may contribute books that they scanned for themselves to Bookshare for shared use.

4.4 Financing

Depending on the source of human workers and their expertise, different financing mechanisms can be utilized.

Public – Many human-powered access technologies are publicly funded. For instance, video relay services are usually publicly funded, subsidized by fees that everyone pays on their phone bill.

Personal/Private – Sometimes human assistance is funded by the user, either personally or through the user's workplace, educational institution, or insurance. Private funds and trusts fall in the same category.

Unpaid – Some assistance is provided for free. For instance, the ESP Game attracts workers through a game. Solona is operated by volunteers and so it is free to use.

4.5 Worker Competence

Different sources of human work may provide workers with different competencies. Even within a particular source, worker competence may vary dramatically.

Expert / Skilled – Many services will require skilled workers. For instance, a video relay service requires proficient sign language speakers.

Amateurs – Some services may be supported by amateurs. For instance, the ASL-STEM Forum invites anyone to contribute, under the assumption that the rest of the community will filter out the good contributions from the bad.

Non-Expert – Some assistance can be provided by non-experts. For instance, VizWiz uses workers recruited from Mechanical Turk. In this case, the ability to see is the only qualification.

4.6 Latency

Different sources of human assistance may have different expected latencies. It might be acceptable to wait a few hours (or even days) for a volunteer to read your mail, but sign language interpreting may need to happen right away.

Interactive – Some human-powered access technologies are interactive. For instance, video relay services work in real-time.

Short Delay – Many services operate under the assumption of a short delay. For instance, VizWiz returns answers in less than 30 seconds. Social Accessibility takes, on average, in the order of a few hours to resolve an accessibility problem.

Undetermined – Some services have an undetermined delay. For instance, workers may never choose to add a video for a particular term on the ASL-STEM Forum, or they may add it immediately.

4.7 Accuracy Guarantees

Humans can provide services that are still too difficult to provide solely automatically, but may provide poor quality assistance for a number of reasons, including workers' misunderstanding of what they are asked to do, lazy or even malicious workers, or underspecified questions. Human-powered access technology uses a number of methods for quality control.

Vetted Workers – Many human-powered access technologies improve accuracy by using only workers who have been vetted. Workers may be vetted by their employer, by volunteer organizations, or via reputations gained over time. This is how the Video Relay Service and Remote Reading Services work (among others). These services can be costly and may need to be arranged in advance.

Abstractions – Some technologies use abstractions to help guarantee accuracy. For instance, the abstraction introduced by

the ESP Game makes it unlikely for the two game players to agree on a label if the latter does not accurately describe the image.

Redundancy – One way to help ensure accuracy is by requiring multiple assistants to make the same suggestions. For instance, the ESP Game requires that multiple players submit the same label.

User Mediated – In some cases, users are presented with alternatives provided by human assistants and asked to decide whether they are correct. For instance, VizWiz provides users with answers from multiple workers.

Human Mediated – Humans sometimes serve not only as assistants but also as mediators. In the ASL-STEM Forum, contributors help to decide which signs work best by contributing new signs and rating the ones that are already there.

4.8 Reliability

Many sources of human assistance are not always available; for instance, a simple approach to engage a human worker would be to simply send a text message to a friend or family member. This method is unreliable because the friend or family member may not always be available. Many of the methods of recruiting human assistance assume a network connection, which may not always be available (e.g., in a basement conference room or on a plane during takeoff). Traditional approaches, like hiring a sign language interpreter, do not face these reliability concerns.

Always Available – Few sources of human assistance are always available as networking connectivity is a requirement, if nothing else. Nevertheless, services like VRS employ workers who are available at all times, increasing reliability.

Assumed Available – Services built on microtask marketplaces (like VizWiz) are assumed to be reliably available, but there are not guarantees that this will always be the case.

Undetermined – Services like the ASL-STEM Forum have undetermined reliability. While people may come to contribute and rate signs, they also may not.

4.9 Assistance Provenance

Providing feedback to users about the human computation that is occurring on their behalf is critical for them to make informed decisions. In particular, it can be useful to know who or what has provided assistance.

Transparent – Users are informed regarding who or what will be assisting them beforehand. Most VRS operate in this way, where users are first connected to the person that will be assisting them.

Opaque – Users are not informed who or what assisted them. The service operates like a “black box.”

Obfuscated – Users are told the general source or sources of assistance, but they have no way of knowing who or what actually assisted them for a particular request. IQ Engines / oMoby works this way – pictures are described by either computer vision or human workers, but are not labeled as to the source of assistance.

4.10 Confidentiality, Privacy, and Anonymity

When humans are included in the loop, new issues arise in terms of confidentiality, privacy, and anonymity. This dimension concerns how a human-powered access technology seeks to provide such guarantees.

Trusted Worker Pools – Remote interpreting services require workers to agree to strict confidentiality rules; oversight helps to ensure that workers comply.

User Feedback – Some tools may provide feedback to users to help them realize the implications of recruiting assistance. For

instance, a tool might detect that a face is included in a photograph before sending to an untrusted

No Guarantees – Many tools provide no guarantees regarding confidentiality, privacy or anonymity. By making no guarantees, tools may have some benefits, e.g. operating more cheaply.

4.11 Consideration of Broader Context

To the extent that privacy, anonymity and confidentiality are considered in human-powered access technology, the focus is generally exclusive on the end user. For instance, technology may go to great lengths to help protect their identity. Often, the effects on others in the broader context in which the technology is used are ignored. For instance, bystanders may unwittingly find themselves in the lens of a blind user using VizWiz. Workers may be asked to answer a question with consequences (e.g. what is the dosage listed on this medication bottle?).

User Consideration – For the user to be considered, mechanisms should be in place to match the user’s expectations. Video Relay Services do this well.

Worker Consideration – Workers may be protected from providing assistance in situations, in which they feel uncomfortable. Workers may be told the consequences of their assistance and allowed not to assist without penalty.

Community Consideration – Mechanisms may be in place to protect people who are not directly involved in the use of the technology from accidentally becoming part of an interaction.

4.12 Broader Applicability of Human Work

The work that humans perform in the context of access technology can be more broadly applicable to other people.

Individual Reuse – In some cases, individuals may be able to reuse the work done on their behalf.

Group Reuse – Many services allow broad groups to reuse the work that was done on the behalf of others. For instance, accessibility improvements made in Social Accessibility can be used by other people who visit the same web page. The annotations made in Go Braille can be used by other travelers.

No Reuse – In many cases, it does not make sense to reuse work (or the application doesn’t support it). For instance, in VRS it does not make sense for others to reuse the sign language interpretation of a conversation.

4.13 Targeted Disability

Technologies vary based on what disability is targeted. Sensory disabilities are widely regarded as the most popular targets of automatic technology and this seems to be true for human-powered access technology as well. Some examples of application for other types of disabilities exist, such as the MAP-Lifeline tool to support people with cognitive disabilities. We do not attempt to provide a breakdown of different types of disabilities – many such classifications exist (e.g. [18]).

5. DISCUSSION

Increasing connectedness represents great opportunities to engage human power in access technology. Development of this new technology need not start from a blank slate, but can rather build from the rich history of people helping people with disabilities.

In this paper, we have motivated a set of 13 design considerations for human-powered access technology. These dimension do not prescribe a set of requirements on access technology, but instead present a framework on which new technologies can be compared and evaluated. For instance, consider a home video system

installed in the home of an older adult that allows remote workers to check up on them. It is tempting to dismiss such a product outright because of the obvious privacy concerns; however, it has been documented that older people are willing to trade such incursions into their privacy if it means that they can live in their homes longer [10]. In such a situation, it may be appropriate to make sure that confidentiality has been ensured in lieu of privacy, perhaps by only employing fulltime workers who are subject to confidentiality agreements, or create technology that only provides minimal views (for instance, restricted to a foot off of the ground or less, in order to detect falls).

5.1 Application to Existing Technologies

Our design principles allow for meaningful and standardized comparison of the technology in this space. Figure 2 presents a comparison of the 15 technologies described in Section 3.

For instance, in some ways the Social Accessibility Project and Solona are similar – specifically, they both will describe images on the web for blind people. However, we can see in Figure 2 where these tools differ and the trade-offs in their design. Namely, whereas Solona uses a few expert describers, Social Accessibility uses a broader network of volunteers. Solona is able to achieve lower latency and may afford greater guarantees of confidentiality. Social Accessibility broader pool of workers may make it more reliable. Importantly, these dimensions do not make one service inherently better or worse, but our design principles give us terminology useful in discusses their relative advantages.

As another example, VizWiz and Scribe4Me target different groups – one answers visual questions for blind people and the other describes audio events – but they are similar in other ways. For instance, both rely on the user to initiate the human assistance piece and both target low-latency responses. They differ in the source of the human assistance – VizWiz uses Amazon’s Mechanical Turk whereas Scribe4Me relied on trained experts – but are similar in their goal of enabling user-initiated real-time answers to sensory questions.

5.2 Implications for Crowdsourcing

Most efforts in mainstream crowdsourcing have thus far concentrated on achieving acceptable results from crowd labor (high quality, low cost, high speed, etc.), but it is becoming evident that we also need to think about the broader crowdsourcing ecosystem. 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 VizWiz worker is asked to decipher dosage

information on a pill bottle, they might reasonably choose to skip the task if they feel that they cannot do it well. Perhaps, our technology can even learn to give feedback to users when many workers decide to skip a particular question.

Our interest is primarily in the domain of access technology, in which it is particularly important to meet user expectations and ensure their control. Nevertheless, we believe these principles could be more broadly applied to mainstream human-powered technology. The design dimensions explored here are relevant to mainstream technology, but only a few have been explored in the broader context mainstream technology. In particular, the extensive experience of disabled people with human-powered

Design Dimensions for Example Applications

	Initiative (end User, Workers, Organizations)	Source (Experts, Crowd, Volunteers, Organized Volunteers, Artificial Intelligence)	Motivation (Intrinsic, Status, Financial, Enjoyment, Personal)	Financing (Public, Personal, Unpaid)	Competence (Expert, Amateurs, Non-Expert)	Latency (Interactive, Short Delay, Undetermined)	Accuracy (Vetted workers, Abstractions, Redundancy, User mediated, Worker mediated)	Reliability (Always, Assumed, Undetermined)	Provenance (Transparent, Opaque, Obscured)	Confidentiality, Privacy, Anonymity (Trusted workers, User feedback, None)	Broader Context (User, Worker, Community, None)	Broader Applicability (Individual, Group, None)	Target Disability
1. ASL-STEM	W	C, V, O	I, S	U	E, A	U	U, W	U	T	N	U	G	Deaf
2. TGP	W, O	P, V	I, E	P, U	E, A	U	V	U	B	N	N	G	Blind
3. GoBraille	U, W	C, V	I, P	U	A, N	U	U	U	B	N	U, W, C	G	Blind
4. ESP Game	W, O	C	S, E	U	N	U	R	S	B	N	N	G	Blind
5. BookShare	U, W	E, V, O	I, F, P	P, R, U	E, A	U	V, U, W	S	B	N	U, W, C	G	Print
6. Respeaking	U, O	E, V, O	I, F	P, R, U	E, A	U	V, W	U	T	T	U, W	N	Deaf
7. SA	U	C	I, S	U	A, N	S	U	S	B	N	U	G	Blind, Other
8. Remote Read	U	E, O	F	P, R, U	E	I	V	A	T	T	U, W	N	Blind
9. VRS	U	E	F	P, R	E	I	V	A	T	T	U, W	N	Deaf
10. RRTC	U	E	F	P, R	E	I	V	A	T	T	U, W	N	Deaf
11. Scribe4Me	U	E	I	U	A, U	S	V	A	B	T	U	U	Deaf
12. MAP-LL	W	E	F	P, R	E	S	V	A	B	T	U, W	N	Cognitive
13. Solona	U	E, O	I	U	A	S	V	S	B	T	U	N	Blind
14. VizWiz	U	C	I, F	R	N	S	R, U	S	B	T	U	N	Blind
15. IQ Engines	U	C, AI	F	R, U	N	S	V	A	O	N	N	G	Blind

Figure 2: An analysis of the 15 example applications on the 13 design dimensions for human-powered access technology. Reported values are simplifications but support analysis of design trade-offs and comparison. For instance, VizWiz and Solona are both human-powered systems for describing images, but they use different sources of human labor. Solona’s expert workers increase confidentiality, competence and accuracy, but at the cost of latency and reliability. Values reported are abbreviations drawn from the set in the top row.

technology suggests that we are likely to see additional structure, guarantees, and consideration for workers in the future.

5.3 Sustainability

A important question for human-powered access technologies is how they will be organized and sustained. Some mature services like VRS have established business models. The Universal Access Tax on telephone services finances VRS in the United States. VRS companies are paid by the minute for each VRS call that is made. VRS workers are highly trained at interpreter training programs and at the companies themselves. Other services mentioned, like GoBraille and VizWiz, are experimental with their development financed by research grants. Should VizWiz move to be commercial then the money for paying the photograph readers could either come from users themselves or from a government funded agency. In some cases like Bookshare, where volunteer workers are used, there may be a need for a paid coordinator to recruit and train volunteers. It is clear that for any successful human-powered access technology there must be a mechanism to sustain it over time. No single model fits all these technologies, and a given technology may be successful under several different models.

6. FUTURE WORK

The design dimensions discussed in this paper highlight a number of important areas for future work in access technology that enables people to help one another.

Although there is progress to be made on all of the dimensions outlined in this paper, Figure 2 shows a specific gap in the *Consideration of Broader Context* and *Latency* dimensions. Although the user should remain of primary concern, we must also consider the workers and others in the community, as such concerns are paramount for adoption [14]. Technology that is low latency also requires experts that need to be recruited in advance. For instance, a technology that uses human assistance to help a blind user read a menu in a restaurant or a deaf student interact with her hearing peers is not useful if it cannot do so in nearly real-time. Future work may look to extend the benefits of pairing users with a specific expert to cheaper sources of non-expert assistance, without losing the advantages of asking an expert.

Our coverage of human-powered access technology has highlighted some gaps in existing work. For instance, to our knowledge, most human-powered systems have targeted sensory disabilities, but future work might try to further extend this idea to technology aiding people with cognitive or motor disabilities [18],

which will present new challenges to overcome. Few projects have explicitly considered the broader context in which human-powered access technology exists; therefore, an interesting area for future research is to consider how design methodologies that explicitly consider broader contexts, such as value-sensitive design [3], may be applied in this space.

7. CONCLUSION

Advances in human-powered access technology have the potential to make the world more accessible to people with disabilities. Far from starting from a blank slate, people with disabilities have a long history to draw on as they help to form the structure of new technologies that facilitate human assistance. We hope that our design principles will serve as a common framework in this area to catalyze discussion and highlight fruitful new areas of research.

8. ACKNOWLEDGMENTS

This work has been supported by NSF Awards #IIS-1049080, #IIS-0808678, and #CNS-0751083.

9. REFERENCES

1. *Amazon's Mechanical Turk*. <http://mturk.com>.
2. *Bookshare*. <http://bookshare.org>.
3. *ChaCha*. <http://www.chacha.com>.
4. *Dragon Naturally Speaking*. <http://www.nuance.com/talk/>.
5. *Facebook*. <http://www.facebook.com>.
6. *Facebook Stats*. <http://facebook.com/press/info.php?statistics>.
7. *IQ Engines*. <http://iqengines.com>.
8. *Open Street Map*. <http://www.openstreetmap.org>.
9. *Twitter*. <http://www.twitter.com>.
10. Adlam, T., Carey-Smith, B., Evans, N., Orpwood, R., Boger, J., & Mihailidis, A., *Implementing Monitoring and Technological Interventions in Smart Homes for People with Dementia: Case Studies*, in *Behaviour Monitoring and Interpretation - BMI: Smart Environments*. 2009, IOS Press.
11. Azenkot, S., Prasain, S., Borning, A., Fortuna, E., Ladner, R.E., Wobbrock, J.O., *Enhancing Independence and Safety for Blind and Deaf-Blind Public Transit Riders*, in *CHI 2011*: Vancouver, BC.
12. Baillie, M. and J.M. Jose, *An Audio-Based Sports Video Segmentation and Event Detection Algorithm*, in *CVAVI 2004*.
13. Bernstein, M., Miller, R.C., Little, G., Ackerman, M., Hartmann, B., Karger, D.R., and Panovich, K., *Soylent: A Word Processor with a Crowd Inside*, in *UIST 2010*.
14. Beyer, H. and K. Holtzblatt, *Contextual design: defining customer-centered systems*. 1998, San Francisco: Morgan Kaufmann.
15. Bigham, J.P., et al., *VizWiz::LocateIt - Enabling Blind People to Locate Objects in their Environment*, in *CVAVI 2010*.
16. Borodin, Y., et al., *More than meets the eye: a survey of screen-reader browsing strategies*, in *W4A 2010*. p. 1-10.
17. Borodin, Y., et al., *What's new?: making web page updates accessible*, in *ASSETS*. 2008, p. 145-152.
18. Carmien, S., et al., *Increasing workplace independence for people with cognitive disabilities by leveraging distributed cognition among caregivers and clients*, in *GROUP 2003*, p. 95-104.
19. Cavender, A.C., et al., *ASL-STEM Forum: enabling sign language to grow through online collaboration*, in *CHI 2010*, p. 2075-2078.
20. Chu, W.-T., W.-H. Cheng, and J.-L. Wu, *Semantic context detection using audio event fusion: camera-ready version*. *EURASIP J. Appl. Signal Process.* **2006**: p. 181-181.
21. Cowling, M. and R. Sitte, *Comparison of techniques for environmental sound recognition*. *Pattern Recogn. Lett.*, 2003. **24**(15): p. 2895-2907.
22. Dawe, M., *Desperately seeking simplicity: how young adults with cognitive disabilities and their families adopt assistive technologies*, in *CHI 2006*, p. 1143-1152.
23. Ferres, L., *Proc. of the International Cross-Disciplinary Conf. on Web Accessibility*. 2011.
24. Goette, T., *Factors leading to the successful use of voice recognition technology*, in *ASSETS 1998*, p. 189-196.
25. Horowitz, D. and S.D. Kamvar, *The anatomy of a large-scale social search engine*, in *W4A 2010*, p. 431-440.
26. Howe, J., *Why the Power of the Crowd Is Driving the Future of Business*. 2009: Crown Business.
27. J. Sudol, O.D., C. Blanchard and T. Dorcsey, *LookTel — A Comprehensive Platform for Computer-Aided Visual Assistance*, in *CVAVI 2010*.
28. Jeffrey P. Bigham, C.J., Hanjie Ji, Greg Little, Andrew Miller, Robert C. Miller, Robin Miller, Aubrey Tatarowicz, Brandyn White, Samuel White, and Tom Yeh, *Vizwiz: Nearly Real-time Answers to Visual Questions*, in *UIST 2010*.
29. Kassirer JP, G.G., *Clinical problem solving: A behavioral analysis*. *Ann Internal Medicine*, 1978. **89**: p. 245-255.
30. Kawanaka, S., et al., *Accessibility commons: a metadata infrastructure for web accessibility*, in *ASSETS 2008*, p. 153-160.
31. Kittur, A., E.H. Chi, and B. Suh, *Crowdsourcing user studies with Mechanical Turk*, in *CHI 2008*, p. 453-456.
32. Ladner, R.E., et al., *Automating tactile graphics translation*, in *ASSETS 2005*, p. 150-157.
33. Manduchi, R., J. Coughlan, and V. Ivanchenko, *Search Strategies of Visually Impaired Persons Using a Camera Phone Wayfinding System*, in *ICCHP 2008*, p. 1135-1140.
34. Miyoshi, S., et al., *Support Technique for Real-Time Captionist to Use Speech Recognition Software, in Computers Helping People with Special Needs*, K. Miesenberger, et al., p. 647-650.
35. Phillips, B. and H. Zhao, *Predictors of Assistive Technology Abandonment*. *Assistive Technology*, 1993. **5**:1: p. 36-45.
36. Porrero, P., *Improving the Quality of Life for the European Citizen: Technology for Inclusive Design and Equality*. 1998, IOS Press.
37. Richard, O.G.a.G., *Automatic transcription of drum sequences using audiovisual features*, in *ICASSP 2005*, p. 205-208.
38. Samuel White, H.J., and Jeffrey P. Bigham, *EasySnap: Enabling Blind People to Take Photographs*, in *UIST 2010 - Demos*.
39. Snyder, J., *Audio description: The visual made verbal*. *International Congress Series*, 2005. **1282**: p. 935-939.
40. Stinson, M., Eisenberg, S., Horn, C., Larson, J., Levitt, H., Stuckless, R. . *Real-Time Speech-to-Text Services: A report of the National Task Force on Quality of Services in the Postsecondary Education of Deaf and Hard of Hearing Students*. 1999; Available from: <http://netac.rit.edu/publication/taskforce/realtime>.
41. Takagi, H., et al., *Collaborative web accessibility improvement: challenges and possibilities*, in *ASSETS 2009*, p. 195-202.
42. Tara Matthews, S.C., Carol Pai, Janette Fong, and Jennifer Mankoff, *Scribe4me: Evaluating a mobile sound translation tool for the deaf*, in *UbiComp 2006*, p. 159-176.
43. Union, I.T., *Trends in Telecommunication Reform 2007: The Road to NGN*. 2007.
44. Viegas, F.B., et al., *Talk Before You Type: Coordination in Wikipedia*, in *Proc. of the 40th Annual Hawaii International Conf. on System Sciences*. 2007, IEEE Computer Society. p. 78.
45. Wald, M., et al., *Correcting automatic speech recognition captioning errors in real time*. *International Journal of Speech Technology*, 2007. **10**(1): p. 1-15.