

CAREER: Human-Backed Access Technology: Enabling Accessibility Tools to Fall Back to Humans

The past decades have seen the development of wonderful new computing technology for disabled people serving as sensors onto an inaccessible world. As examples, optical character recognition makes printed text available to blind people, speech recognition makes spoken language available to deaf people, and GPS technology helps keep people with cognitive impairments on track. Unfortunately, such promising technology is often severely under-utilized due to persistent errors and limited scope. When existing strategies and technology fail, disabled people often ask friends for help, which reduces independence since help is not always available.

The PI's concept of **human-backed access technology** allows automatic technology to quickly incorporate human-powered services to make access technology more reliable and cover a wider set of problems today. Two converging trends make this feasible: (i) mobile devices with a plethora of sensors and high-bandwidth connections are now mainstream, and (ii) large pools of human workers are always available through micro-task marketplaces (such as Amazon's Mechanical Turk), on social networks (Facebook, Twitter, etc.), and via their own connected mobile devices.

The **proposed research** encompasses the user-centered, value-sensitive design of tools to facilitate human-backed access technology and a scientific investigation of these tools grounded in three new applications for blind, deaf and hard-of-hearing people – *VizWiz* for answering visual questions, *AudioWiz* for interacting with sound, and *quik-Social Accessibility* for improving the accessibility of the web. These tools will be made possible by *HumanAccess*, a new tool for connecting disabled people to human workers, friends, and volunteers able to answer questions quickly and accurately. The tools will be evaluated in longitudinal studies deployed to users *in situ*.

The synergistic **education plan** includes (i) the development of a **new Ph.D. course** on intelligent access technology, and (ii) **innovative new summer programs** for deaf and blind students in grades 7-12. The accessible curriculum will afford students the opportunity to build human-backed technology and to create interfaces that they can use to support others with different disabilities. These programs will be run in conjunction with the existing Youth Slam and TechGirlz programs for high school students run by the National Federation of the Blind and the National Technical Institute for the Deaf. Undergraduate computer scientists will help develop and teach the course, providing a unique chance to learn about accessible computing. The accessible curriculum for both courses, as well as captioned lecture videos, will be made available for free on the web.

Intellectual Merit: Human-backed access technology represents a new paradigm in human-computer interaction in which intelligent tools designed to support people with disabilities are explicitly backed-up by humans. Formative user studies will inform the initial design of human-backed access technology and generate design guidelines cognizant of the human values of users. The *HumanAccess* tool will enable other researchers to explore human-backed access technology in their research, and facilitate three new applications to be deployed *in situ* with blind, deaf, and hard-of-hearing users to validate and improve human-backed access technology over the long term.

Broader Impacts: This proposal addresses a primary shortcoming of current access technology and consequently may afford greater independence for disabled people. The research products may have applications in other domains that utilize error-prone automatic processes, such as automated personal assistants or video tagging. The integrated education plan introduces students in grades 7-12, undergraduate students, and Ph.D. students to accessible computing through innovative summer programs and new curriculum. Introducing blind and deaf high school students to computing in a supportive environment may encourage some to pursue a career in computing.

Key words: access technology; accessible computing; mobile applications; web accessibility

TABLE OF CONTENTS

For font size and page formatting specifications, see PAPPG section II.B.2.

	Total No. of Pages	Page No.* (Optional)*
Cover Sheet for Proposal to the National Science Foundation		
Project Summary (not to exceed 1 page)	1	_____
Table of Contents	1	_____
Project Description (Including Results from Prior NSF Support) (not to exceed 15 pages) (Exceed only if allowed by a specific program announcement/solicitation or if approved in advance by the appropriate NSF Assistant Director or designee)	15	_____
References Cited	6	_____
Biographical Sketches (Not to exceed 2 pages each)	2	_____
Budget (Plus up to 3 pages of budget justification)	8	_____
Current and Pending Support	1	_____
Facilities, Equipment and Other Resources	2	_____
Special Information/Supplementary Documents (Data Management Plan, Mentoring Plan and Other Supplementary Documents)	12	_____
Appendix (List below.) (Include only if allowed by a specific program announcement/ solicitation or if approved in advance by the appropriate NSF Assistant Director or designee)	_____	_____
Appendix Items:		

*Proposers may select any numbering mechanism for the proposal. The entire proposal however, must be paginated. Complete both columns only if the proposal is numbered consecutively.

CAREER: Human-Backed Access Technology¹: Enabling Accessibility Tools to Fall Back to Humans

1. Project Overview

The past few decades have seen the development of wonderful new computing technology that serves as sensors onto an inaccessible world for disabled people – as examples, optical character recognition (OCR) makes printed text available to blind people [52, 64], speech recognition makes spoken language available to deaf people [81, 82, 96], and way-finding systems help keep people with cognitive impairments on track [68, 69]. Despite advances, this technology remains both too prone to errors and too limited in the scope of problems it can reliably solve to address the problems faced by disabled people in their everyday lives [55, 58, 96, 97].

Some technology has made it out to people, but still has high error rates in real situations. 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 [71], object recognition works reasonably well until the camera is held by a blind person [14, 55, 85], and the laudable 99% accuracy reported by commercial automatic speech recognition systems [38] falls off precipitously on casual conversation or any time it hasn't been trained for the speaker [96]. 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 [23, 24, 59], leading them to be used by only technically-savvy blind people [89]. When access technology is unreliable, it is abandoned [37, 46, 79].

This proposal is about the idea that access technology tools would be more reliable if human workers, volunteers, and friends could quickly back up fragile automatic techniques. This approach is directly inspired by how many disabled people already solve problems when their existing strategies fail: ask a friend for assistance [11, 55]. Unfortunately, someone isn't always there to help and asking frequently can make one feel like a burden. Some paid expert services are available, such as captioning services, but they can be expensive and must be arranged in advance. Consequently, they are generally inappropriate for the few seconds of help required to fill in the gaps of existing technology. A blind participant in one of our studies said [55]: *"I get so frustrated when I need sighted help and no one is there."* If access tools were *backed up* by always-available **human-powered services**, technology could be more reliable and more useful.

The idea of backing up access technology with humans has been previously articulated [25, 35, 99, 100], but two recent trends have made it practical. First, mainstream mobile phones with low-latency, high-bandwidth connections and a wealth of sensors (camera, microphone, GPS, etc.) have become commonplace, obviating the need for special hardware and making communication on-the-go faster [94]. Second, marketplaces for small jobs like Mechanical Turk [3] and social networks like Facebook [42] and Twitter [93] have grown in popularity [43] providing large pools of potential workers already connected and available in nearly real-time, which companies are already using [36]. A core component of the proposed research is a scientific investigation into how workers, friends, and volunteers can be recruited to answer the questions of disabled users while respecting the values and expectations of the requesters [44].

¹ Although *assistive technology* is more often used to describe technology designed for people with disabilities, I choose to use *access technology* because it emphasizes technology's role in facilitating access and de-emphasizes the notion that otherwise capable people are reliant on the technology's assistance.

Human-backed access technology is my concept for a new approach to developing access technology encompassing the idea that falling back to humans can create more reliable and more useful technology for people with disabilities. The concept will be developed according to the value-sensitive design methodology to proactively consider the human values of people with disabilities [44]. Consideration of the desires, concerns, and biases of users is particularly important when designing technology for people with disabilities who are confronted with a nearly endless parade of expensive *near* solutions. Successful integration of access technology requires consideration of the meaning assigned to devices, user expectations, anticipated social costs, and an understanding of its role in the whole person [77]. To maintain scope, this proposal explores human-backed access technology grounded in technology for blind, deaf, and hard-of-hearing people, but the idea is broadly applicable to other groups.

The four **research objectives** of this proposal are:

1. **formative user studies** using qualitative methods to inform the value-sensitive design of human-backed access technology and generate initial user-driven design guidelines;
2. **invention and iterative improvement of *HumanAccess***, a software library tool encompassing our generated design guidelines enabling anyone to include human workers as part of automatic technology;
3. **design and implementation of three applications (VizWiz, AudioWiz, and quik-Social Accessibility)** that use the human-backed access technology approach to improve access to visual, audio, and web information, respectively; and
4. **longitudinal studies** of the three applications *in situ* to inform the iterative improvement of the guidelines and the *HumanAccess* tool, grounded in real technology and experiences.

These objectives are steps along the way to validating the human-backed access technology concept and producing guidelines and tools useful to other researchers and developers.

The integrated **education objectives** include (i) developing summer programs for students who are blind and deaf that offer an introduction to computer science through innovative curriculum focused on human-backed access technology, (ii) involving computer science undergraduate students in the design and operation of the summer programs to give them an appreciate for accessible computing and pedagogy, and (iii) a new course on *intelligent access technology* for Ph.D. students and advanced undergraduates, serving as a bridge connecting my department's strength in artificial intelligence to HCI and accessible computing.

Throughout my career, I have integrated research, teaching, and outreach. At the University of Washington, I served as a teaching assistant for five quarters, volunteered as a tutor for introductory courses throughout my grad career, and participated in numerous outreach events. I developed the curriculum for the computer science track at the National Federation of the Blind (NFB) Youth Slam in 2007 and 2009, and led a group of four undergraduates in teaching a class for 15 blind high school students [9]. At the University of Rochester, I currently support 4 Ph.D. students (3 women), 1 masters student, and 5 undergraduates (1 woman). I created a new course (*CSC 210: Web Programming*) that proved the most popular non-introductory course in the department and received a rating of 4.4/5.0. I am unsatisfied to let technology sit on the shelf, and released the WebAnywhere [12, 22] non-visual web browser and reading tool. More than 1500 people currently use it each week. I expect the tools and applications developed as part of my research plan to similarly be useful for people with disabilities and I commit to releasing them as free downloads. Releasing technology builds living laboratories that afford a perspective much farther in the lifecycle of technology than most research ever goes, an aspect of WebAnywhere from which I have taken full advantage [12, 20, 21, 54, 65]. A CAREER award will allow me to continue my established pattern of **integrated research, education and outreach**.

2. Use Scenarios

The following scenarios are idealized examples that illustrate how people may use the applications that I will create and facilitate my investigation of human-backed access technology:

1. How a blind person might use human-backed access technology as part of VizWiz:

Rachel is a blind woman staying at a hotel away from home who would like to go out to dinner at a nearby restaurant. She uses her iPhone with the VoiceOver screen reader to find a nearby restaurant based on her current location. She then accesses another web service to retrieve a walking map, and, once outside of the hotel, her phone uses the map and GPS to help keep her on track toward the restaurant. When she arrives at where the restaurant should be, the doors that she tries to open are locked. Her phone helps her take several pictures of the surrounding area and then she records herself asking, "I am trying to find the door to the Volterra restaurant, is it here?" This request is sent to a human-powered visual guide service and within a minute she receives the following answer: "A sign in the window says that the restaurant is closed." Rachel uses location-based search to find a second restaurant, and then repeats the navigation process. She arrives at the second restaurant, sits down at a table and is handed a menu. Her phone interactively helps her align the menu's text in its camera, and then ships an image of it to a remote OCR service. Unfortunately, the menu is written in a script font that cannot be recognized automatically, so a message comes back informing Rachel that her request has been forwarded to a human-powered service. Within a minute, she begins receiving the transcribed restaurant menu, which she can read in her desired modality. When she is done with her meal, she asks for a list of only the deserts, and the human-powered service is able to comply.

2. How a deaf person might use human-backed access technology as part of AudioWiz:

Jennifer, a deaf professor of computer science, is helping to run a visit day for perspective graduate students. When they arrive at the student center, she wants to encourage interactions between the students and faculty. She senses that the students are not interacting fluidly, and so she consults AudioWiz. An automatic service shows that the room is very noisy. She types the question, "What's the loud noise?" and is told that very loud music is playing. She asks the person in charge to turn it down. As the students are beginning to head to their first activity, Jennifer finds herself in a group of her peers who are talking about meeting later to catch-up after the first day. Jennifer can read lips with concentration but gets distracted as the students bustle about her and misses what time they say they'll be meeting. When she realizes this a few seconds later, the group has disbanded, rushing off with different groups of students, so Jennifer has AudioWiz transcribe the last 30 seconds of audio. While automatic speech recognition is unable to decipher the speech given the background noise, a human worker tells her that the plan is to meet at the local coffee shop at 5pm. On her way home that night, Jennifer has car trouble. She uses AudioWiz to obtain a description of the sounds her car is making as "an engine running, but stalling periodically." From the description, Amanda recognizes the problem as the same one she had a month ago before a mechanic installed a new alternator. Instead of calling and paying for a tow, she uses a relay service to call the mechanic, who tows her car for free since his work is still under warranty.

These examples are idealized, but they are illustrative of the enumerable ways in which we use our senses in interacting with our everyday environment and how, with a little backing from humans, automatic services can become much more useful and reliable. To allow us to primarily focus on the novel components of human-backed access technology, we will reuse our own existing frameworks and those of our industrial supporters for the automatic services (Section 5).

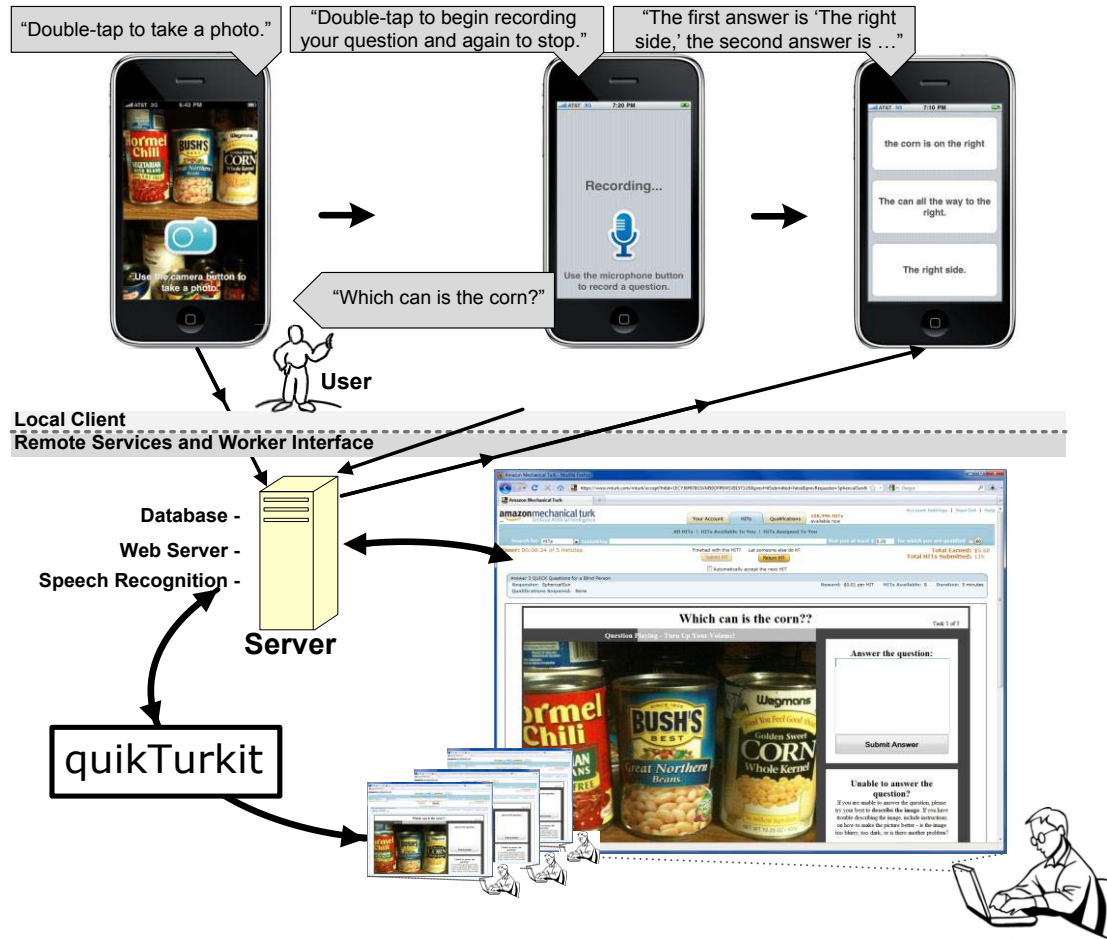


Figure 1: The VizWiz client is a talking iPhone application that works with the included VoiceOver screen reader. VizWiz proceeds in three steps—taking a picture, speaking a question, and waiting for answers. System components include a web server that serves the question to web-based workers, a speech recognition service that converts spoken questions to text, and a database that holds questions and answers. quikTurkit adaptively posts jobs (HITs) to Mechanical Turk in order to maintain specified criteria (for instance, a minimum number of answers per question or a pool of waiting workers of a given size).

3. My Work on Human-Backed Access Technology

My prior work on human-backed access technology is comprised of a number of prototypes [8, 13, 18, 54, 55, 85] and research studies [10, 11, 55] that collectively serve as a valuable precursor to my proposed research plan. This work has provided ample motivation and support for the general thesis of this proposal: *access technology can be made more reliable and useful when backed-up by human supporters*. To effectively create, explore, and build human-backed access technology, researchers and developers need (i) tools to allow human supporters to easily be incorporated into their own software and (ii) effective design guidelines for incorporating human-powered services while respecting the human values and expectations of users.

Much of my prior work has concerned improving web access for blind people. WebInSight for Images [15] assigns descriptions to images first using automatic techniques but then falls back to human labelers via the ESP Game [1]. My Accessmonkey system allows end users to improve the accessibility of the web pages that they visit, and introduced the idea that people with disabilities

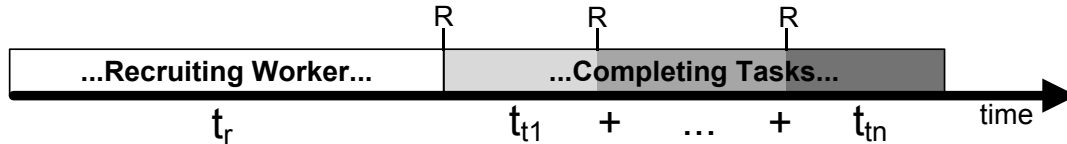


Figure 2: A simple diagram of the components of the total time required to recruit and receive answers from an online worker. Generally, the longest component is the time to recruit workers. If workers are asked to do multiple jobs, then they are expected to be available at a frequency set by the average time required to complete a job (points marked by *R* in the diagram).

could **help one another** improve the accessibility of their own web experiences [8, 17]. The Social Accessibility Project at IBM has improved and expanded on this idea [60, 88, 89]. As part of this proposal, IBM has agreed to continue our collaboration [59] by providing access to their Social Accessibility tools so that we can integrate faster responses from human supporters.

3.1 VizWiz: Nearly Real-time Answers to Visual Questions

Recently, I have explored human-backed access technology for blind and low-vision people. VizWiz is a mobile application that I developed that enables blind people to take a picture, speak a question, and quickly receive answers from human workers (Figure 1). Correct answers can be received in less than **30 seconds** on average from Mechanical Turk and for an average cost of just **\$0.07 per question** [55]. Recent prototypes integrate with new sources of answers, such as social networks (Facebook and Twitter) and ones' own personal contacts (via email and picture text messages). These components provide valuable experience and a starting point for the *HumanAccess* tool (Section 5.2).

In a one-week field deployment of VizWiz with 11 blind people, we gained insight into what blind people may truly want to know about their visual environments, unconstrained by what our technology could recognize automatically – they wanted to know about a lot more than just color and text! As examples, they wanted to know if their clothes matched, what setting their appliances were on, and whether there were empty picnic tables across the park. Participants also asked VizWiz questions to which they likely already knew the answer (what object am I holding) or which replicated the functionality of their existing technology (what does this text say), in what were likely novelty effects. Participants began to **reveal their expectations** of human-backed access technology. In particular, participants knew that their questions were being answered by people and often wanted to enter into a dialog with the answerer. For a variety of reasons related to the technology and the reliability of each individual worker, this is difficult to achieve, but is an example of a human value important to users that may come out of our broader investigation.

The deployment revealed problems inherent in having blind people take photographs, which we are addressing with a suite of tools that help blind people take photographs called EasySnap [85]. Although still preliminary work, we have implemented powerful computer vision techniques right on the phone that work with the capabilities of blind people to help blind people take good pictures of almost anything from their existing iPhones. We have also created an initial prototype of AudioWiz, which acts as an analogous sensor onto the world for deaf and hard-of-hearing people. AudioWiz explores the concept of retroactive accessibility [54] – that is, it can help solve access problems after they occur. It does this by keeping a recording of the last 30 seconds of audio that it has heard, and only processing it when the user chooses.

The original VizWiz application received answers quickly from Mechanical Turk using a tool that I developed called quikTurkit, which serves as an abstraction layer on top of the TurKit Mechanical Turk API [67] and a precursor to *HumanAccess*. quikTurkit achieves low latency answers by **queuing workers** before they are needed, and, to my knowledge, is the first tool to

target nearly real-time human-powered services. Figure 2 illustrates the components of the total time for a single worker to answer a question. The time equals the sum of the time for the worker to find the posted task t_r plus the time to answer each of n questions ($t_1 + \dots + t_n$). Workers complete multiple tasks so that each worker is engaged longer, ideally until a new question arrives. To use quikTurkit, requesters create their own web site on which Mechanical Turk workers answer questions. Interfaces should allow multiple questions to be asked so that workers can be engaged answering other questions until they are needed. As each answer is submitted (points labeled R in Figure 2), the web site returns the question with the fewest answers for the worker to complete next, enabling new questions to jump to the front of the queue. If a worker is already engaged when the k^{th} question is asked, then the time to recruit workers t_r is eliminated and the expected wait time for that worker is $t_{(k-1)}/2$. Recruiting multiple workers increases the likelihood that a worker will already be engaged when a new job becomes available.

quikTurkit lets client applications signal when new work will likely be coming for workers to complete, and quikTurkit can then begin **recruiting workers before they are needed**, keeping them busy solving previously-asked questions. VizWiz signals quikTurkit when users begin taking a picture. The lead time of the signal is effectively removed from the recruitment delay t_r . quikTurkit can also keep a pool of workers of a given size continuously engaged and waiting, although workers must be paid to wait. In practice, keeping 10 or more workers in the pool is doable, although costly. *HumanAccess*, will recruit workers from new sources and some lessons may transfer (like pre-recruiting workers and keeping workers busy with old jobs until they are needed). Nevertheless, we will need to adapt and invent new strategies for the new services.

4. Related Work

The idea of human-backed access technology is inspired by my observations of the **current failures** of access technology [10, 55] and how disabled people compensate when their usual strategies (or technology) break down [11, 23, 55] – generally by **asking a friend or family member** for help. For instance, when blind people are unable to solve a CAPTCHA, they may ask for sighted assistance [10, 86]. They may send a photo of themselves to a friend before an important job interview [55]. Wizard-of-Oz studies are commonly used to validate new technology and interactions before they are implemented [72]; human-backed access technology is a type of *deployable Wizard-of-Oz*. Human-backed access technology mechanizes the process by integrating both automatic and human-powered services together so that disabled people are not reliant on the availability of specific friends or family or made to feel like a burden on others.

A dizzying array of **automatic access technology** has been developed for people with disabilities. For example, speech recognition is steadily improving, [38, 41, 74] and is now able to achieve high accuracy for speakers on which it has been trained in quiet environments. However, its quality varies substantially for arbitrary speakers or those in noisy environments [82]. Detecting specific sounds, such as gun shots, explosions, and cars braking, is a well-studied area with applications in the military, safety and rating films [27, 33] and such sounds can be easy to detect in the right context [5]. Music and drum sequences can be matched with known artists [4, 73, 83]. Sounds like barking dogs, a smashing glass bottle, a window fan, a car engine, a doorbell, a door closing, water running, and a whistling tea kettle can be detected by specialized products [34]. One user said, “Sometimes there are false alarms caused by noisy goings-on in the kitchen which signal the doorbell. It is easy to see if it is a false alarm - I only have to look at [my dog],” illustrating her current strategy for coping with a specific error. On the vision side, object recognition automatically identifies specific objects [47, 53] and photo-based question answering [98] can answer questions in reference to a photograph. Automatic tools have made compelling

progress, but remain far from being able to answer arbitrary visual or aural questions. When tools can optionally include human-powered services, automatic processes can be made more reliable. Disabled people can also ask what they really want to know, instead of adapting their question to the capabilities of the device –instead of receiving the entire text of a menu from an OCR program, they could ask, “How much is the cheapest hamburger?” or instead of receiving a transcription of a minute-long conversation, they could ask “What time is the meeting?” Human-backed access technology can answer a wide variety of naturally-phrased questions.

Mobile devices for blind, deaf, and hard-of-hearing people. Initially, mainstream cell phones were either not accessible to blind people or required expensive add-on screen reading software, but both the Android platform and iPhones now come with free screen readers. The tight integration of the iPhone’s VoiceOver software has led it to be particularly popular [40, 95]. Touchscreen devices like the iPhone were once assumed to be inaccessible to blind users, but my prior work showed that well-designed, multitouch interfaces that leverage the spatial layout of the screen and can actually be preferred by blind people [57]. Telephones were also initially inaccessible to deaf people, but the invention of the TTY changed the way people within the deaf community contacted and communicated with each other [56, 66]. Pagers and text messaging on mobile phones added mobility to the communications of deaf and hard of hearing people, and largely ended the era of the TTY [48, 56, 80]. Prior work has explored how real-time sign language can be achieved on today’s cellphone networks [28, 32], and marketing for the new iPhone 4 used sign language communication to highlight its new FaceTime video calling feature.

Access applications for **general-purpose smartphones** are replacing special-purpose access devices, but people with disabilities still carry a number of such devices. Blind people may carry GPS-powered navigation aids, barcode readers, light detectors, color identifiers, text recognizer and compasses [52, 57, 58, 64, 70, 97]. The Sidekick II released in 2008 supports the monitoring of doorbells, telephones, smoke detectors and fire alarms and will alert users of a visitor, phone call, or danger [92]. Other phones simply detect loud noises, such as a baby crying or an alarm [61]. People may own many such devices but carry only a few because of the limited scope and added bulk of each device [55, 58]. Although most mobile tools are still implemented as local software running on the mobile device, more applications are starting to access powerful remote services, such as OCR [91] and contextually relevant navigation information [45]. These popular services suggest that people are becoming familiar with outsourcing questions to remote services.

A few **examples of human-backed access technology** already exist. Scribe4Me retroactively transmits audio to on-hand transcribers who describe and transcribe sounds [90]. In a two-week field study of Scribe4Me, deaf participants used Scribe4Me to describe the speech and audio around them and found it useful as an unobtrusive way to participate in the semi-public speech of informal conversations and to compensate for missing information in lip reading. Participants also used Scribe4Me to satisfy their curiosity, leveling the field with hearing people. They used it to “listen” to a new accessible crosswalk sign and catch the scores from the game on a friend’s car radio. Participants saw the value in Scribe4Me, but highlighted the following key limitations that diminished both its practical utility and usefulness for exploring the interactions it enables:

1. **Descriptions took 3-5 minutes:** Descriptions were delayed because of slow cellular networks and a limited pool of transcribers. One participant said, “Three to five minutes [delay] leaves one as an observer rather than an active participant.”
2. **Lack of context:** Without context, it was often difficult for transcribers to figure out what was being said and what information would most likely be useful.
3. **An inflexible interface:** Users were not actively involved in the transcribing process, and, therefore, transcriptions often missed the information users desired - such as a speaker’s emotion, the ambient audio, or a transcription of the right person.

The design of the *HumanAccess* tool will address these concerns, as informed by user studies with blind, deaf, and hard-of-hearing participants. Human-backed access technology will be faster in part simply because of the high-speed network connections now available, but also because we will recruit from larger pools of non-expert workers, split tasks into multiple pieces, and recruit multiple workers to work on the same problem. I will also apply the lessons learned from VizWiz for recruiting workers quickly. Finally, the proposed applications will balance automatic and human-powered services in order to get answers back much more quickly when the automatic services work, while maintaining the reliability of humans. Scribe4Me highlighted interesting privacy and usability concerns, illustrating the need for unobtrusive techniques for users to alert those nearby of recording and transcribing and to ignore the device when social norms dictate discrete or delayed use (such as in class or in church). We will explore how and when a user's own social network (via Facebook, Twitter, video text messaging, or email) might be a better source of answers, and how to expose such choices to users.

Human-backed access technology builds from prior work in using **human computation**, specifically work designed to improve accessibility. The ESP Game was originally motivated (in part) by a desire to provide descriptions of web images for blind people [1]. Mechanical Turk makes outsourcing small paid jobs practical [3] and has been used for many purposes, including large user studies [63], labeling image data sets [87], and determining political sentiments in blog snippets [51]. Amazon Remembers inputs pictures and outputs products sold by Amazon [2]. It likely outsources some questions to Mechanical Turk. The TurKit library on which our quikTurkit is built lets programmers easily employ multiple turkers using common programming paradigms [67]. To my knowledge, *quikTurkit* (Section 4.2) is one of the first examples of non-expert human computation that operates in **nearly real-time** (less than a minute) [55].

In contrast, most human-powered services are provided with an expectation of delay. ChaCha and KGB employees answer questions asked via the phone or by text message in just a few minutes [31, 62]. Soylent “embeds” human workers into a word processor to help with editing tasks (which generally have a latency of 20 minutes) [6]. Solona started as a CAPTCHA solving service, but its popular RAIVE service now lets registered blind people submit any image for description [86]. According to its website, “Users normally receive a response within 30 minutes.” Prior work has explored how people ask and answer questions on their existing online social networks [75], observing that answers often came back within a few minutes. The “Social Search Engine” Aardvark supports asking one’s social network questions, and similarly advertises that answers come back “within a few minutes.” [49]. Paid expert services can operate in nearly real-time, like the relay services used by deaf and hard of hearing people, but must be arranged in advance and therefore inappropriate for short interactions on demand [80].

The Social Accessibility project connects blind web users who experience web accessibility problems to volunteers who can help resolve them via a web browser plug-in, but 75% of requests remain unsolved after a day [60, 88, 89]. Problems are eventually fixed but generally only after the requesting user has moved on to something else; we will develop the *quik-Social Accessibility* application (Section 5.3.3) to work with the existing Social Accessibility framework to improve web accessibility problems quickly so that users could benefit during their current browsing session. User studies with this technology will investigate the trade-offs that users make when choosing between services that vary in latency, accuracy, anonymity, and cost.

Prior work is limited by its reliance on specific groups and specific people, and the resulting services can be slow, expensive, or not always available. A challenge for the human-backed access technology approach is to incorporate **non-expert** human workers, volunteers and friends while maintaining a high-quality level of service in line with the values of potential users.

5. Research Plan: Tools for Human-Backed Access Technology

The proposed research can be divided into four overlapping activities: (i) formative studies using qualitative methods to investigate attitudes regarding human-backed access technology, identify strategies used to overcome problems with current technology, and develop initial design guidelines for human-backed access technology; (ii) invention, development, and validation of the *HumanAccess* tool for recruiting human workers, friends, and volunteers in nearly real-time from existing microtask marketplaces, social networks, and personal contacts; (iii) user-centered design of three new applications to improve access to visual, auditory, and web information (*VizWiz*, *AudioWiz*, and *quik-Social Accessibility*); and (iv) longitudinal studies to evaluate the new applications *in situ*, inform further iterative improvement, and validate our approach.

5.1 Formative Studies

Fundamental to the success of the concept of human-backed access technology is understanding how users interact with technology backed by humans; whether or not the inherent latency, privacy concerns, and human unpredictability can be mitigated; and how users envision this technology fitting into their existing lifestyle, work, and social patterns [7]. Access technologies that are perceived as poor fits with user needs, as too complex, or as requiring too much configuration or maintenance experience high abandonment rates [37, 46].

Formative studies consisting of unstructured interviews and focus groups with disabled people will be conducted in the initial design stages for each application that will next be developed. The goal will be to gain early feedback on how potential users imagine interacting with each application, what they view as the primary concerns with it, and their expectations concerning it. Focus groups will be grounded in real prototypes when feasible in order to afford a greater appreciation of both expectations and the capabilities truly afforded by the new tools. For instance, *VizWiz* users desired support for greater interactivity than we initially provided and expected answers to be returned in no more than a couple of minutes. They were less concerned about privacy, but they did want the information about who they were sharing their images with and what exactly they were sharing to be transparently available. This initial formative work will help to identify these issues before they are codified into the technology and prototype applications. Following the formative sessions, we will employ a user-centered design process to iteratively improve *HumanAccess*, the three applications, and our design guidelines.

5.2 HumanAccess: A Tool for Human-Backed Access Technology

HumanAccess is the software tool that we will develop in order to facilitate human-backed access technology in other applications. It will consist of the following three primary components: (i) a remote service coordinating nearly real-time response from workers, volunteers, and friends on existing services like Mechanical Turk, social networking sites like Facebook and Twitter, and personal contacts via email and picture text messaging; (ii) a framework for managing a flexible pool of automatic description services encompassing such technology as automatic speech recognition, OCR, and other services; and, finally, (iii) mediation services which coordinate automatic and human-powered services to generate descriptions for users. The mediation service will help balance the expected speed, cost, and privacy advantages of the automatic services with the expected accuracy and scope advantage of the human services.

Although *HumanAccess* serves a similar role as *quikTurkit* (Section 4.2), it differs as follows: First, *HumanAccess* will facilitate the transparent user-driven switching between multiple human answer providers, whereas *quikTurkit* only interfaced with Mechanical Turk. Second, *HumanAccess* will support the mediation and mixed-initiative [50] inclusion of automatic services for answers with human-powered services. *HumanAccess* will be developed with explicit

consideration of human values as informed by both the already mentioned formative studies and feedback from longitudinal studies of specific applications (Section 5.3.3). For instance, it may be reasonable to think that users would prefer assistance from their “friends” on their social networks, but they may instead prefer the relative anonymity that paid services afford. Friends and family might offer the best assistance, but users may hesitate to ask them for help frequently. *HumanAccess* will facilitate our scientific investigation of these issues.

5.3 Example Applications

To validate the concept of human-backed access technology in the everyday lives of disabled people, we will design, develop, and iteratively improve three new applications using the *HumanAccess* library. Importantly, we will build from our existing infrastructure and technology developed by our industrial partners to maximally leverage our effort on developing and validating the concept of human-backed access technology, while using their expertise for the included automatic services. These applications will be developed for appropriate platforms – we will initially focus on the iPhone platform for the two mobile applications because of their built-in, integrated accessibility features and prevalence among blind and deaf smartphone users [95]. For the web, we will target the Firefox and Internet Explorer web browsers because they are popular among blind people and Social Accessibility is currently available for these browsers. We will use an iterative user-centered design process to develop each application.

5.3.1 Answering Visual Questions for Blind and Low-Vision People

The first application that we will develop will be an extension of the existing VizWiz application (Section 3). It will differ from the existing application substantially, in that it will employ the full range of human services provided by *HumanAccess* and incorporate automatic services, such as OCR, color recognition, and object recognition. This will allow us to design mixed-initiative interfaces [50] and study how blind people mediate between and interact with automatic and human-powered services. We will also include greater assistance to blind people taking pictures by integrating picture-taking guidance from EasySnap [85], and may explore asking questions via video as a supplement to pictures. Video has the potential to support interactive feedback and richer information for human workers, but may introduce latency concerns. Interactive video in which users and human workers are paired is particularly attractive, but its straightforward application requires strong assumptions about reliability and speed of human workers. Importantly, by leveraging our existing framework, we can begin exploring these issues quickly while still grounding our explorations in the everyday lives real users.

5.3.2 Retroactive Audio Transcriptions for Deaf and Hard of Hearing People

The second application that we will create will be AudioWiz, a mobile retroactive tool for deaf and hard of hearing people to use to interpret their audio environments. This tool is similar in concept to Scribe4Me [90], but will be designed for low-latency transcription and description enabled by (i) new sources of workers from *HumanAccess*, (ii) the incorporation of automatic services, and (iii) the faster network connections available today. We will work with deaf and hard-of-hearing people to design interfaces that enable them to send appropriate context and also to interact with the human workers who may be answering their questions. Transcribing sound introduces new potential causes of latency because it requires sequential playback – to mitigate

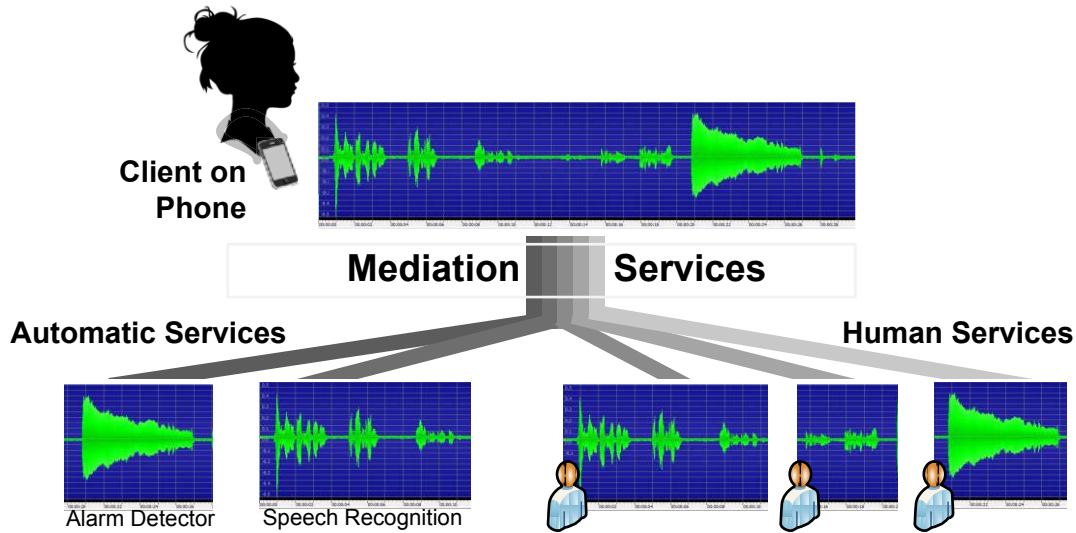


Figure 3: Proposed AudioWiz Application and *HumanAccess* framework. After AudioWiz records audio on the client device and the user decides to send it off for remote description, the *HumanAccess* mediation services analyze the sound, copy it and send it to other remote services. In this example, the mediation services detect portions of the clip that appear appropriate for speech recognition and the alarm detector and forward them as appropriate. The sound is split into chunks for human description and forwarded to workers. Mediation services attempt to keep semantically connected sequences together.

these problems, we will design and evaluate techniques for recruiting multiple workers to work on a single transcription task. AudioWiz will leverage existing techniques for detecting and recognizing audio events, and we may develop new automatic detectors based on observed usage [78] in order to inform users about their environments (Figure 3). AT&T will provide automatic audio services, including speech recognition and general audio interpretation (see included letter).

5.3.3 Improving Web Accessibility

Despite over a decade of web standards, developer tools and education, and automated techniques designed to make the web more accessible, much of the web remains inaccessible and difficult to use for blind people [11]. Problems have worsened as end users have started contributing more content, including pictures usually uploaded without a text alternative [15]. As a pragmatic solution, I and others have built tools that enable anyone to improve anyone's web content [17, 89]. The Social Accessibility tool from IBM Japan, targeted at improving access for blind web users, is a particularly mature example of this approach. Although thousands of web pages have been improved using the Social Accessibility approach, users still experience significant delays (hours or days) before their problems are fixed because of a limited pool of workers [88, 89]. We will design, implement, and evaluation *quik-Social Accessibility*, an extension to Social Accessibility that will use *HumanAccess* to improve the web quickly.

quik-Social Accessibility will include automatic services, initially drawn from my prior work on image description [15], heading labeling [26], and dynamic updates [24]. It will also use *HumanAccess* to recruit workers when these services fail. Importantly, we will use the existing Social Accessibility framework (see included letter), implemented and supported by IBM, which will enable us to focus on designing and evaluating human-backed access technology for this domain. An example challenge that we will need to address is creating intuitive interfaces for non-expert workers to improve web accessibility problems. Enabling web accessibility problems to be fixed quickly will dramatically improve web utility for people with disabilities.

5.4 Deployment and Longitudinal Studies

Following an iterative user-centered design process, the three applications will each be deployed to 20 blind, deaf, or hard-of-hearing people (as appropriate) over three months as part of a longitudinal study to learn how they are used in the everyday lives of target users. As the applications are refined, they will also be released on appropriate application stores so that anyone may try them – in doing so, we will expand our pool of participants and explore the iterative improvement of these applications *in situ*. After a fraction of interactions with the applications, study participants may be asked to rate the quality and appropriateness of the response that they got, and we will conduct pre- and post- study interviews with participants to gain additional qualitative feedback. Deploying these applications will also help us compare qualities of different sources of answers. For instance, is Mechanical Turk faster than the user's social network, or do users receive higher-quality answers from their personal contacts than they do from the social networks? We will observe quantitatively which services are most often used by participants, and use qualitative follow-ups on the device to determine why participants chose to use Facebook for one question and Mechanical Turk for another. We will incorporate user feedback into the design of future iterations of each tool. More broadly, these studies will validate the approach taken by human-backed access technology and further inform design guidelines.

Fundamental to the success of the human-backed access technology approach is that it respect the human values of its users [44, 77] and fit their existing lifestyle, work, and social patterns [7]. By releasing the applications for a longer period of time, we hope to gain a better sense of how disabled people might adopt (or not) the applications that we create. In our initial deployment of VizWiz [55], we still observed numerous examples of what we believe to be novelty effects over a week long study. The proposed studies will last three months with the option for users to continue for longer to evaluate the tool in the “Confirmation” stage of technology adoption [84] in which users decide if they actually want to continue using the technology. By releasing three separate applications that all use the human-backed access technology approach, we hope to generate more general guidelines than would be possible by releasing only one.

6. Education and Outreach

My integrated education plan involves middle school, high school, undergraduate, and Ph.D. students. The following section first outlines innovative new curriculum and summer programs for blind and deaf students in grades 7-12 that will be created and run with the assistance of computer science undergraduates. Students will build interfaces for human-backed access technology and build accessible interfaces so that they can assist others. It then overviews a new course for Ph.D. students and advanced undergraduates on *intelligent access technology*.

6.1 Summer Programs for High School Students with Disabilities

I will develop curriculum for summer courses to be held each of the five years of the award. Each program will include approximately 15 students grades 7-12 who are either blind or deaf as part of existing summer programs run by the National Federation of the Blind (NFB) and the National Technical Institute for the Deaf (NTID) at the Rochester Institute of Technology (RIT).

Blind students will participate as part of the existing *NFB Youth Slam*. I developed the original accessible computer science curriculum for the 1st and 2nd Youth Slam [9] in which students created “intelligent” instant messenger chatbots². Students without prior programming experience wrote real computer programs using accessible curriculum and tools, supported by

² <http://webinsight.cs.washington.edu/nfbslam/>

undergraduate computer science student helpers. The proposed program builds from my prior experience with innovative new material and activities inspired by human-backed access technology. Blind students will develop computer vision applications for iPhones. Specifically, we will give students the tools to build object recognizers, let students choose objects to recognize, and specify the feedback they would like to receive about the object. For instance, if their phone detected bananas, then they may want feedback about whether or not it is ripe. If they choose to detect faces, then they may want help centering the face in the frame. If they choose to detect dogs, then they may want a report on the mood of the dog. These problems vary in difficulty, but can all be accomplished by combining automatic computer vision with human helpers (Section 3). Students will also build interfaces that let them assist deaf students interpret sounds. Students will learn about creating accessible web interfaces, and use their interfaces to answer audio questions posed by deaf students in prior workshops. We will create modular components that students can combine to accomplish interesting new tasks (with guidance from instructors) while exploring topics in computer vision and human-backed access technology.

I will also adapt this curriculum for use at NTID in Rochester, NY as part of *NTID TechGirlz*, a program for female students grades 7-9 who are deaf or hard-of-hearing. The curriculum will be analogous to that just described for blind students. Specifically, students will be provided with audio processing modules (explored through AudioWiz), and develop new mobile applications to interpret specific sounds, and create interfaces for answering the visual questions of blind people.

These courses will be held in alternating years to match the existing programs' schedules. Each year, students will build interfaces enabling them to answer the questions of students in the prior year (for instance, deaf students in 2012 will answer questions from blind students in 2011). The instructors will seed queries for the first offering. The attached letters from the **NFB and NTID** affirm their cooperation. In addition to the high school students directly involved in the program, **5 undergraduate computer science students** will be included on the team each year. Not only will these students be integral to running the course, but they will also be given the chance to interact with students with disabilities who are being introduced to computer science for the first time. My experience with the Youth Slam dramatically altered my perspective, and I hope this program will similarly be an opportunity for both those enrolled in the program and for the undergraduates supporting them to learn about accessible computing and pedagogy. Students who work with me on the summer programs may be better equipped to conduct **more impactful research** in accessible computing because of the experience.

6.2 Ph.D. Course on Intelligent Access Technology

I will create a **new course**, *Intelligent Access Technology*, aimed at Ph.D. students and advanced undergraduate students. The University of Rochester has been traditionally strong in artificial intelligence, natural language, and computer vision, and I will design this new course to act as a bridge to students in these areas. I believe that accessible computing would benefit if the best researchers in related fields had accessibility experience, and I hope this course will serve as a model bridge for other courses. The new course will draw content from intelligent user interfaces, accessible computing, and human-backed access technology. Although only at the University of Rochester for one year, I have already created one course (*CSC 210: Web Programming*) and am developing another (*CSC 212: Human-Computer Interaction*) that are new to the university's curriculum. I enjoy creating courses, and generally incorporate elements of my own research into both course motivation and assignments, even for introductory courses. I am particularly excited to develop and teach a new course on *Intelligent Access Technology*. As this unique course has no direct equivalents at other universities, I will **release all course materials and lectures as captioned videos** on the web for the benefit of other students and educators.

7. Research Experience for Students

All of my students, including undergraduates, form an integral and valued part of my research team; they work closely with me to build prototypes, formulate and conduct user studies, write papers, and present at research venues. This proposal will fund one Ph.D. student from computer science to work with me on all aspects of this proposal – including the design, implementation, evaluation, and validation of human-backed access technology, and education components. I intend to recruit at least one undergraduate researcher, funded through the NSF REU Scholars program, if possible. I further intend to host an additional **qualified undergraduate with a disability** during each summer, recruited and supported with the assistance of AccessComputing (see the included letter). Engaging qualified students with disabilities on the research team may lead to better research results through both their direct contributions to research products and the informal educational opportunities their participation may afford the other members of my team.

8. Intellectual Merit

Human-backed access technology represents a **new paradigm** in human-computer interaction in which intelligent tools designed to support people with disabilities are explicitly backed-up by humans. Formative user studies will inform the initial design of human-backed access technology and generate design guidelines cognizant of the human values of users. The *HumanAccess* tool will enable other researchers to explore human-backed access technology in their own research prototypes, and facilitate three new applications to be deployed *in situ* with blind, deaf, and hard-of-hearing participants to both validate human-backed access technology and facilitate its iterative improvement over the long term.

9. Broader Impacts

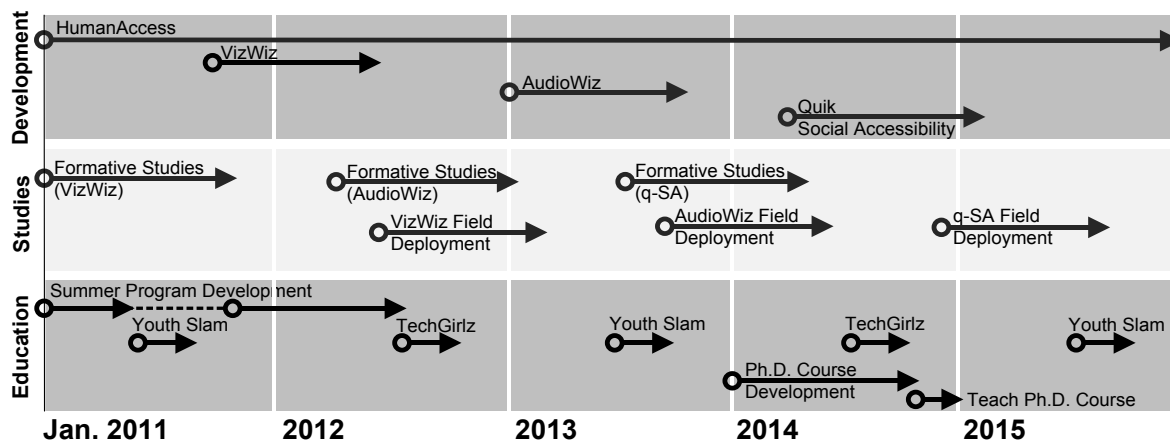
This proposal addresses a primary shortcoming of current access technology and consequently may afford **greater independence for disabled people**. The research products may have applications in other domains that utilize error-prone automatic processes, such as automated personal assistants or video tagging. The integrated education plan introduces middle school, high school, undergraduate, and Ph.D. students to accessible computing through innovative summer programs and new curriculum. Introducing blind and deaf high school students to computing in a supportive environment may encourage some to pursue a career in computing. I intend to recruit qualified students with disabilities as part of my research team with assistance from the AccessComputing program; exposing disabled students to research may help address the underrepresentation of disabled people among those earning doctorates in STEM (only 1% of STEM doctorate recipients are disabled, compared to 10% of STEM undergraduates [39, 76]).

10. Dissemination of Research and Education Results

The research and education products of this research will be disseminated through appropriate venues, including ACM ASSETS, CHI, UIST, W4A, SIGCSE, TOCHI, and TACCESS. As important is getting results to both people with disabilities and those building technology for people with disabilities, so we will also present at consumer-oriented events, such as CSUN, RESNA, and ATIA. I will release accessible materials for both the summer programs and my new Ph.D. course on the web (including captioned videos). The software tools and applications resulting from this research will be released as open source and as free downloads in the relevant application stores for mobile platforms. I built WebAnywhere to be used by other researchers and companies [12, 65], and will similarly build human-backed access technology tools to be reused. The TRACE Research and Development Center will help us with dissemination and sustainability of the developed technology through the “Raising the Floor” initiative (see included letter).

11. Project Plan

The chart below shows my proposed research and education schedule. The summer programs alternate each year, and *HumanAccess* will be **iteratively** improved throughout the entire five-year proposal as informed by longitudinal user studies with the three developed applications.



12. Prior Accomplishments and Results

I believe I am uniquely positioned to execute the research and education activities outlined in this proposal. Only at the end of my first year at the University of Rochester, I have published 6 full-length papers at top-tier conferences [16, 20, 23, 29, 54, 55], a book chapter [19], and several workshop and shorter papers [14, 30, 55]. I also founded the *ROC HCI Research Group* (hci.cs.rochester.edu), a new research group whose members number nearly 20. In my department, 20% of the Ph.D. students are women, but 3 out of 4 of the Ph.D. students working with me are women, in part driven by two top women Ph.D. students who came to the University of Rochester specifically to join my group. I see extensive collaboration as a way to fully leverage my work and improve accessible computing – last year I published with researchers at 9 other institutions.

My WebAnywhere service is used by more than 1500 people each week, many of whom would not otherwise have web access [12]. Although originally designed for blind people, people with reading difficulties are now using it and educators are using it to help teach about accessible computing³. I have been presented many awards for my work, reflecting my impact both within academia and beyond it: the Microsoft Imagine Cup Accessible Technology Award (2008), two W4A Accessibility Challenge awards (2008, 2010), two ASSETS Best Paper Awards (2008, 2009) and the selective **MIT Technology Review Top 35 Innovators Under 35 Award (2009)**.

While a Ph.D. student at the University of Washington, I contributed to National Science Foundation funded projects on which my thesis advisor, Dr. Richard E. Ladner, was the PI. Under Award IIS-0415273, I led the WebInSight project for improving web accessibility, including WebInSight for Images [15], Accessmonkey [8], WebAnywhere [12, 21, 22], WebinSitu [11] and Usable Audio CAPTCHAs [10]. Under grant CNS-0837508, I developed the curriculum for and led instruction at the NFB Youth Slam-Computer Science Track in 2007 and 2009 [9]. Under award IIS-0915268, I helped to create an online forum for expanding sign language coverage in STEM [30] and ClassInFocus for connecting expert interpreters with deaf students in STEM [29].

³ WebAnywhere is currently used in education programs at the Univ. of Washington, Georgia Tech, MIT, Univ. of Wisconsin, Univ. of Rochester, and elsewhere. Try at webanywhere.cs.washington.edu/beta.

References

1. Ahn, L.v. and L. Dabbish, *Labeling images with a computer game*, in *Proceedings of the SIGCHI conference on Human factors in computing systems*. 2004, ACM: Vienna, Austria. p. 319-326.
2. *Amazon Remembers*. Accessed 2010; Available from: <http://www.amazon.com/gp/>.
3. *Amazon's Mechanical Turk*. Accessed 2010; Available from: <http://mturk.com>.
4. Andrew Fisher, K.L., Chee Wong, Beatrice Farina, and Jonathan Symons. *Shazam Entertainment*. Accessed 2010; Available from: <http://www.shazam.com/>.
5. Baillie, M. and J.M. Jose, *An Audio-Based Sports Video Segmentation and Event Detection Algorithm*, in *Proceedings of the 2004 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'04) Volume 7 - Volume 07*. 2004, IEEE Computer Society. p. 110.
6. 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 *ACM Symposium on User Interface Software and Technology (UIST 2010)*. 2010, ACM Press.: New York, NY. p. To Appear.
7. Beyer, H. and K. Holtzblatt, *Contextual design: defining customer-centered systems*. 1998, San Francisco: Morgan Kaufmann Publishers Inc.
8. Bigham, J.P., *Accessmonkey: enabling and sharing end user accessibility improvements*. SIGACCESS Access. Comput., 2007(89): p. 3-6.
9. Bigham, J.P., et al., *Inspiring blind high school students to pursue computer science with instant messaging chatbots*, in *Proceedings of the 39th SIGCSE technical symposium on Computer science education*. 2008, ACM: Portland, OR, USA. p. 449-453.
10. Bigham, J.P. and A.C. Cavender, *Evaluating existing audio CAPTCHAs and an interface optimized for non-visual use*, in *Proceedings of the 27th international conference on Human factors in computing systems*. 2009, ACM: Boston, MA, USA. p. 1829-1838.
11. Bigham, J.P., et al., *WebinSitu: a comparative analysis of blind and sighted browsing behavior*, in *Proceedings of the 9th international ACM SIGACCESS conference on Computers and accessibility*. 2007, ACM: Tempe, Arizona, USA. p. 51-58.
12. Bigham, J.P., W. Chisholm, and R.E. Ladner, *WebAnywhere: experiences with a new delivery model for access technology*, in *Proceedings of the 2010 International Cross Disciplinary Conference on Web Accessibility (W4A)*. 2010, ACM: Raleigh, North Carolina. p. 1-4.
13. Bigham, J.P., et al., *VizWiz: nearly real-time answers to visual questions*, in *Proceedings of the 2010 International Cross Disciplinary Conference on Web Accessibility (W4A)*. 2010, ACM: Raleigh, North Carolina. p. 1-2.
14. Bigham, J.P., et al., *VizWiz::LocateIt - Enabling Blind People to Locate Objects in their Environment*, in *International Workshop on Computer Vision Applications for the Visually Impaired (CVAVI 2010)*. 2010, IEEE Press: San Francisco, CA.
15. Bigham, J.P., et al., *WebInSight:: making web images accessible*, in *Proceedings of the 8th international ACM SIGACCESS conference on Computers and accessibility*. 2006, ACM: Portland, Oregon, USA. p. 181-188.
16. Bigham, J.P., R.S. Kaminsky, and J. Nichols, *Mining web interactions to automatically create mash-ups*, in *Proceedings of the 22nd annual ACM symposium on User interface software and technology*. 2009, ACM: Victoria, BC, Canada. p. 203-212.
17. Bigham, J.P. and R.E. Ladner, *Accessmonkey: a collaborative scripting framework for web users and developers*, in *Proceedings of the 2007 international cross-disciplinary conference on Web accessibility (W4A)*. 2007, ACM: Banff, Canada. p. 25-34.

18. Bigham, J.P., T. Lau, and J. Nichols, *Trailblazer: enabling blind users to blaze trails through the web*, in *Proceedings of the 13th international conference on Intelligent user interfaces*. 2009, ACM: Sanibel Island, Florida, USA. p. 177-186.
19. Bigham, J.P., T. Lau, and J. Nichols, *TrailBlazer: Enabling Blind Users to Blaze Trails Through the Web*, in *No Code Required: Giving Users Tools to Transform the Web*, A. Cypher, et al., Editors. 2010, Morgan Kaufmann.
20. Bigham, J.P. and K. Murray, *WebTrax: Visualizing Non-Visual Web Browsing*, in *International Conference on Computers Helping People (ICCHP 2010)*. 2010: Vienna, Austria.
21. Bigham, J.P., C.M. Prince, and R.E. Ladner, *Addressing Performance and Security in a Screen Reading Web Application That Enables Accessibility Anywhere*, in *Proceedings of the 2008 Eighth International Conference on Web Engineering*. 2008, IEEE Computer Society. p. 273-284.
22. Bigham, J.P., C.M. Prince, and R.E. Ladner, *WebAnywhere: a screen reader on-the-go*, in *Proceedings of the 2008 international cross-disciplinary conference on Web accessibility (W4A)*. 2008, ACM: Beijing, China. p. 73-82.
23. Borodin, Y., et al., *More than meets the eye: a survey of screen-reader browsing strategies*, in *Proceedings of the 2010 International Cross Disciplinary Conference on Web Accessibility (W4A)*. 2010, ACM: Raleigh, North Carolina. p. 1-10.
24. Borodin, Y., et al., *What's new?: making web page updates accessible*, in *Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility*. 2008, ACM: Halifax, Nova Scotia, Canada. p. 145-152.
25. Brabyn, J.A., W.F. Crandall, and W.A. Gerrey, *Remote Reading System for the Blind: A Potential Application of Virtual Presence*, in *International Conference on Engineering in Medicine and Biology Society*. 1992: Paris, France. p. 1538-1539.
26. Brudvik, J.T., et al., *Hunting for headings: sighted labeling vs. automatic classification of headings*, in *Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility*. 2008, ACM: Halifax, Nova Scotia, Canada. p. 201-208.
27. C. Clavel, T.E., and G. Richard, *Events detection for an audio-based surveillance system*, in *EEE International Conference on Multimedia and Expo (ICME 2005)*. 2005. p. 1306–1309.
28. Cavender, A., R.E. Ladner, and E.A. Riskin, *MobileASL:: intelligibility of sign language video as constrained by mobile phone technology*, in *Proceedings of the 8th international ACM SIGACCESS conference on Computers and accessibility*. 2006, ACM: Portland, Oregon, USA. p. 71-78.
29. Cavender, A.C., J.P. Bigham, and R.E. Ladner, *ClassInFocus: enabling improved visual attention strategies for deaf and hard of hearing students*, in *Proceedings of the 11th international ACM SIGACCESS conference on Computers and accessibility*. 2009, ACM: Pittsburgh, Pennsylvania, USA. p. 67-74.
30. Cavender, A.C., et al., *Asl-stem forum: enabling sign language to grow through online collaboration*, in *Proceedings of the 28th international conference on Human factors in computing systems*. 2010, ACM: Atlanta, Georgia, USA. p. 2075-2078.
31. *ChaCha*. Accessed 2010; Available from: <http://www.chacha.com>.
32. Cherniavsky, N., et al., *Activity analysis enabling real-time video communication on mobile phones for deaf users*, in *Proceedings of the 22nd annual ACM symposium on User interface software and technology*. 2009, ACM: Victoria, BC, Canada. p. 79-88.
33. 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.
34. Cowling, M. and R. Sitte, *Comparison of techniques for environmental sound recognition*. Pattern Recogn. Lett., 2003. **24**(15): p. 2895-2907.
35. Crandall, W.F., W.A. Gerrey, and J.A. Brabyn, *A Fax-Based Reading System for the Blind and Print-Handicapped*, in *RESNA 1992*. 1992: Toronto, Canada. p. 463-465.

36. *CrowdFlower*. Accessed 2010; Available from: <http://crowdfunder.com/>.
37. Dawe, M., *Desperately seeking simplicity: how young adults with cognitive disabilities and their families adopt assistive technologies*, in *Proceedings of the SIGCHI conference on Human Factors in computing systems*. 2006, ACM: Montreal, Quebec, Canada. p. 1143-1152.
38. *Dragon Naturally Speaking*. Accessed 2010; Available from: <http://www.nuance.com/talk/>.
39. Education, U.S.D.o. *National Postsecondary Student Aid Study*. 2008; Available from: <http://nces.ed.gov/surveys/npsas/>.
40. *Eyes-Free for Android*. Accessed 2010; Available from: <http://code.google.com/p/eyes-free/>.
41. Fabbriozio, G.D., T. Okken, and J.G. Wilpon, *A speech mashup framework for multimodal mobile services*, in *Proceedings of the 2009 international conference on Multimodal interfaces*. 2009, ACM: Cambridge, Massachusetts, USA. p. 71-78.
42. *Facebook*. Accessed 2010; Available from: <http://www.facebook.com>.
43. *Facebook Statistics*. Accessed 2010; Available from: <http://www.facebook.com/press/info.php?statistics>.
44. Friedman, B., *Value-sensitive design*. interactions, 1996. 3(6): p. 16-23.
45. Gifford, S., et al., *Introduction to the talking points project*, in *Proceedings of the 8th international ACM SIGACCESS conference on Computers and accessibility*. 2006, ACM: Portland, Oregon, USA. p. 271-272.
46. Goette, T., *Factors leading to the successful use of voice recognition technology*, in *Proceedings of the third international ACM conference on Assistive technologies*. 1998, ACM: Marina del Rey, California, United States. p. 189-196.
47. *Google Goggles*. Accessed 2010; Available from: <http://www.google.com/mobile/goggles>.
48. Henderson-Summet, V., et al., *Electronic communication: themes from a case study of the deaf community*, in *Proceedings of the 11th IFIP TC 13 international conference on Human-computer interaction*. 2007, Springer-Verlag: Rio de Janeiro, Brazil. p. 347-360.
49. Horowitz, D. and S.D. Kamvar, *The anatomy of a large-scale social search engine*, in *Proceedings of the 19th international conference on World wide web*. 2010, ACM: Raleigh, North Carolina, USA. p. 431-440.
50. Horvitz, E., *Principles of mixed-initiative user interfaces*, in *Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit*. 1999, ACM: Pittsburgh, Pennsylvania, United States. p. 159-166.
51. Hsueh, P.Y., P. Melville, and V. Sindhwani. *{Data quality from crowdsourcing: a study of annotation selection criteria}*. in *Proceedings of the NAACL HLT 2009 Workshop on Active Learning for Natural Language Processing*. 2009.
52. *Intel Reader*. Accessed 2010; Available from: <http://www.intel.com/healthcare/reader/>.
53. J. Sudol, O.D., C. Blanchard and T. Dorsey, *LookTel — A Comprehensive Platform for Computer-Aided Visual Assistance*, in *International Workshop on Computer Vision Applications for the Visually Impaired (CVAVI 2010)*. 2010, IEEE Press.: San Francisco, CA.
54. Jeffrey P. Bigham, B.Z., and Jeremy Brudvik, *Accessibility by Demonstration: Enabling End Users to Guide Developers to Web Accessibility Solutions*, in *12th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS 2010)*. 2010, ACM Press: Orlando, FL. p. To Appear.
55. 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 *Proceedings of the 23rd Annual ACM Symposium on User Interface Software Technology (UIST 2010)*. 2010, ACM Press: New York, New York. p. To Appear.

56. Johnson, M.A.H.a.M.A., *Assistive Technology for the Hearing Impaired, Deaf-blind and Deaf*. Springer Verlag: Springer Verlag.
57. Kane, S.K., J.P. Bigham, and J.O. Wobbrock, *Slide rule: making mobile touch screens accessible to blind people using multi-touch interaction techniques*, in *Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility*. 2008, ACM: Halifax, Nova Scotia, Canada. p. 73-80.
58. Kane, S.K., et al., *Freedom to roam: a study of mobile device adoption and accessibility for people with visual and motor disabilities*, in *Proceedings of the 11th international ACM SIGACCESS conference on Computers and accessibility*. 2009, ACM: Pittsburgh, Pennsylvania, USA. p. 115-122.
59. Kawanaka, S., et al., *Accessibility commons: a metadata infrastructure for web accessibility*, in *Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility*. 2008, ACM: Halifax, Nova Scotia, Canada. p. 153-160.
60. Kawanaka, S., et al., *Accessibility commons: a metadata repository for web accessibility*. SIGWEB Newsl., 2009(Summer): p. 1-7.
61. Ketabdard, H. and T. Polzehl, *Tactile and visual alerts for deaf people by mobile phones*, in *Proceedings of the 11th international ACM SIGACCESS conference on Computers and accessibility*. 2009, ACM: Pittsburgh, Pennsylvania, USA. p. 253-254.
62. KGB. Accessed 2010; Available from: <http://www.kgb.com>.
63. Kittur, A., E.H. Chi, and B. Suh, *Crowdsourcing user studies with Mechanical Turk*, in *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*. 2008, ACM: Florence, Italy. p. 453-456.
64. *knFB Reader*. Accessed 2010; Available from: <http://www.knfbreader.com/>.
65. L. Ferres, D.C., R. Parada, M. Soza, J. Bigham, *Integrating WebAnywhere and iGraph-Lite: An Engineering Note on Mass Accessibility to Graphs*, in *ACM SIGACCESS Conference on Computers and Accessibility (ASSETS 2010) - Posters Track*. 2010: Orlando, FL. p. To Appear.
66. Lang, H.G., *A Phone of Our Own*. 2000: Gallaudet University Press.
67. Little, G., Goldman, M., Chilton, L., and Miller, R.C., *TurKit: Human Computation Algorithms on Mechanical Turk*, in *ACM Symposium on User Interface Software and Technology (UIST 2010)*. 2010, ACM Press. p. To Appear.
68. Liu, A.L., et al., *Customizing directions in an automated wayfinding system for individuals with cognitive impairment*, in *Proceedings of the 11th international ACM SIGACCESS conference on Computers and accessibility*. 2009, ACM: Pittsburgh, Pennsylvania, USA. p. 27-34.
69. Liu, A.L., et al., *Indoor wayfinding:: developing a functional interface for individuals with cognitive impairments*, in *Proceedings of the 8th international ACM SIGACCESS conference on Computers and accessibility*. 2006, ACM: Portland, Oregon, USA. p. 95-102.
70. Liu, X., *A camera phone based currency reader for the visually impaired*, in *Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility*. 2008, ACM: Halifax, Nova Scotia, Canada. p. 305-306.
71. Manduchi, R., J. Coughlan, and V. Ivanchenko, *Search Strategies of Visually Impaired Persons Using a Camera Phone Wayfinding System*, in *Proceedings of the 11th international conference on Computers Helping People with Special Needs*. 2008, Springer-Verlag: linz, Austria. p. 1135-1140.
72. Mulsby, D., S. Greenberg, and R. Mander, *Prototyping an intelligent agent through Wizard of Oz*, in *Proceedings of the INTERACT '93 and CHI '93 conference on Human factors in computing systems*. 1993, ACM: Amsterdam, The Netherlands. p. 277-284.
73. *Melodis*. Accessed 2010; Available from: <http://www.melodis.com/>.
74. *Microsoft Speech Technologies*. Accessed 2010; Available from: <http://www.microsoft.com/speech/>.

75. Morris, M.R., J. Teevan, and K. Panovich, *What do people ask their social networks, and why?: a survey study of status message q&a behavior*, in *Proceedings of the 28th international conference on Human factors in computing systems*. 2010, ACM: Atlanta, Georgia, USA. p. 1739-1748.
76. NSF/SRC Survey of Earned Doctorates. 2008; Available from: <http://www.norc.org/projects/survey+of+earned+doctorates.htm>.
77. Pape, T.L.-B., J. Kim, and B. Weiner, *The shaping of individual meanings assigned to assistive technology: a review of personal factors*. Disability & Rehabilitation, 2002. **24**(1): p. 5 - 20.
78. Pfeiffer, S., S. Fischer, and W. Effelsberg, *Automatic audio content analysis*, in *Proceedings of the fourth ACM international conference on Multimedia*. 1996, ACM: Boston, Massachusetts, United States. p. 21-30.
79. Phillips, B. and H. Zhao, *Predictors of Assistive Technology Abandonment*. Assistive Technology, 1993. **5.1**: p. 36-45.
80. Power, M.R.P.a.D., *Everyone here speaks txt: Deaf people using sms in australia and the rest of the world*. 2004.
81. Rabiner, L. and B. Juang, *Fundamentals of Speech Recognition*. 1993: Prentice Hall PTR.
82. Ravishandkar, M., *Efficient algorithms for speech recognition*. 2005, Carnegie Mellon University.
83. Richard, O.G.a.G., *Automatic transcription of drum sequences using audiovisual features*, in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2005)*. 2005. p. 205-208.
84. Rogers, E.M., *Diffusion of Innovations*. 1962: Free Press of Glencoe. 367.
85. Samuel White, H.J., and Jeffrey P. Bigham, *EasySnap: Enabling Blind People to Take Photographs*, in *ACM Symposium on User Interface Software and Technology (UIST 2010) - Demos*. 2010, ACM Press: New York, NY. p. In submission.
86. Solona. Accessed 2010; Available from: <http://www.solona.net/>.
87. Sorokin, A. and D. Forsyth, *Utility data annotation with Amazon Mechanical Turk*. Computer Vision and Pattern Recognition Workshops, 2008.
88. Takagi, H., et al., *Social accessibility: achieving accessibility through collaborative metadata authoring*, in *Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility*. 2008, ACM: Halifax, Nova Scotia, Canada. p. 193-200.
89. Takagi, H., et al., *Collaborative web accessibility improvement: challenges and possibilities*, in *Proceedings of the 11th international ACM SIGACCESS conference on Computers and accessibility*. 2009, ACM: Pittsburgh, Pennsylvania, USA. p. 195-202.
90. Tara Matthews, S.C., Carol Pai, Janette Fong, and Jennifer Mankoff, *Scribe4me: Evaluating a mobile sound translation tool for the deaf*, in *International Conference on Ubiquitous Computing (Ubicomp 2006)*. 2006, Springer. p. 159-176.
91. Textscout - Your Mobile Reader. Accessed 2010; Available from: <http://www.textscout.eu/en/>.
92. Today, D. *Revolutionary new device monitors doorbells, phones, smoke detectors - alerts deaf and hard of hearing*. 2008 Accessed 2010; Available from: http://www.deaftoday.com/v3/archives/2008/10/revolutionary_n.html.
93. Twitter. Accessed 2010; Available from: <http://www.twitter.com>.
94. Union, I.T., *Trends in Telecommunication Reform 2007: The Road to NGN*. 2007.
95. VoiceOver. Accessed 2010; Available from: <http://www.apple.com/accessibility/iphone/vision.html>.
96. Wald, M., et al., *Correcting automatic speech recognition captioning errors in real time*. International Journal of Speech Technology, 2007. **10**(1): p. 1-15.
97. Ye, Q., et al., *Text detection and restoration in natural scene images*. J. Vis. Comun. Image Represent., 2007. **18**(6): p. 504-513.

98. Yeh, T., J.J. Lee, and T. Darrell, *Photo-based question answering*, in *Proceeding of the 16th ACM international conference on Multimedia*. 2008, ACM: Vancouver, British Columbia, Canada. p. 389-398.
99. Zimmermann, G. and G. Vanderheiden, *State of Science: Internet-Based Personal Services on Demand*, in *Emerging and Accessible Telecommunications, Information and Healthcare Technologies*. 2002, RESNA Press. p. 62-70.
100. Zimmermann, G. and G. Vanderheiden, *Translation on Demand Anytime and Anywhere*, in *CSUN's Sixteenth Annual International Conference*. 2001: Los Angeles, CA.

Jeffrey P. Bigham, Ph.D.

720 Computer Studies Building

Rochester, NY 14627

(585) 360-0959

jbigham@cs.rochester.edu

<http://www.cs.rochester.edu/u/jbigham/>**A. Education**

B.S.E., Computer Science, Princeton University, June 2003.

M.Sc., Computer Science and Engineering, University of Washington, June 2005.

Ph.D., Computer Science and Engineering, University of Washington, June 2009.

B. Appointments

July 2009 –	Assistant Professor, Department of Computer Science University of Rochester, Rochester, NY, USA
July 2009 – December 2009	Visiting Scientist, EECS CSAIL Massachusetts Institute of Technology, Cambridge, MA, USA
September 2003 – June 2009	Research Assistant Teaching Assistant Department of Computer Science and Engineering University of Washington, Seattle, WA, USA
May 2008 – October 2008	Research Intern, USER Group IBM Almaden, San Jose, CA, USA
May 2008 – October 2008	Summer Fellow Benetech, Palo Alto, CA, USA
May 2005 – September 2005	Research Intern Google, Mountain View, CA, USA
May 2003 – September 2003	Research Intern AT&T Shannon Labs, Florham Park, NJ

C. Publications Most Closely Related to this Proposal

1. **Jeffrey P. Bigham**, Chandrika Jayant, Hanjie Ji, Greg Little, Andrew Miller, Robert C. Miller, Robin Miller, Aubrey Tatrowicz, Brandyn White, Samuel White and Tom Yeh. VizWiz: Nearly Real-time Answers to Visual Questions. In *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST 2010)*. To Appear.
2. **J.P. Bigham** and R.E. Ladner. Accessmonkey: A Collaborative Scripting Framework for Web Users and Developers. In *Proceedings of the International Cross-Disciplinary Conference on Web Accessibility (W4A, 2007)*, 25-34.
3. **J.P. Bigham** and A.C. Cavender. Evaluating Existing Audio CAPTCHAs and an Interface Optimized for Non-Visual Use. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI 2009)*, 1829-1838.
4. **J.P. Bigham**, T. Lau and J. Nichols. TrailBlazer: Enabling Blind Users to Blaze Trails Through the Web. In *Proceedings of 12th International Conference on Intelligent User Interfaces (IUI 2009)*, 177-186.
5. **J.P. Bigham**, T. Lau and J. Nichols. Inspiring Blind High School Students to Pursue Computer Science with Instant Messaging Chatbots. In *Proceedings of the 39th SIGCSE Technical Symposium on Computer Science Education (SIGCSE 2008)*, 449-453.

D. Other Significant Publications

1. **J.P. Bigham**, R. Kaminsky, and J. Nichols. Mining Web Interactions to Automatically Create MashUps. In *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST 2009)*, 203-212.
2. **J.P. Bigham**, C. Prince, R.E. Ladner: WebAnywhere: A Screen Reader On-the-Go. In *Proceedings of the International Cross-Disciplinary Conference on Web Accessibility (W4A 2008)*, 73-82.
3. S.K. Kane, **J.P. Bigham** and J.O. Wobbrock. Slide Rule: Making Mobile Touch Screens Accessible to Blind People Using Multi-Touch Interaction Techniques. In *Proceedings of the International ACM SIGACCESS Conference on Computers & Accessibility (ASSETS 2008)*, 73-80.
4. **J.P. Bigham**, R.S. Kaminsky, R.E. Ladner, O.M. Danielsson and G.L. Hempton. WebInSight: Making Web Images Accessible. In *Proceedings of the International ACM SIGACCESS Conference on Computers & Accessibility, (ASSETS 2006)*, 181-188.
5. Anna C. Cavender, Daniel Otero, **Jeffrey P. Bigham** and Richard E. Ladner. ASL-STEM Forum: Enabling Sign Language to Grow Through Online Collaboration. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI 2010)*. Atlanta, Georgia, 2009. 2075-2078.

E. Synergistic Activities

1. Conference Program Committees: ASSETS 2008-2010; CHI 2011; W4A 2008-2010.
2. Conference Organizing Committees: ASSETS 2009, 2010; W4A 2011.
3. Conference Reviewer: IUI 2008-2010; CHI 2007-2010; UIST 2008-2010.
4. National Federation of the Blind Youth Slam, Computer Science Track. 2007-2009.
5. Raising the Floor. Member. <http://raisingthefloor.net/>. 2009-Present.

F. Select Honors and Awards

1. MIT Technology Review Top 35 Innovators Under 35 Award, 2009.
2. ASSETS 2008 and 2009 Best Student Technical Papers, 2008 and 2009.
3. University of Washington College of Engineering Student Innovator Award, 2009.
4. Microsoft Imagine Cup Accessible Technology Award, 2008.
5. W4A Accessibility Challenge Delegate's Award, 2008 (WebAnywhere) and 2010 (VizWiz).

G. Thesis Advisor

Richard E. Ladner (University of Washington)

H. Other Collaborators

Chieko Asakawa (IBM Japan), Jeremy Brudvik (Georgia Institute of Technology), Yevgen Borodin (Stony Brook University), Anna Cavender (University of Washington), Leo Ferres (University of Concepción), Ryan Kaminsky (University of Washington), Shaun Kane (University of Washington), Shinya Kawanaka (IBM Japan), Tessa Lau (IBM Almaden), Greg Little (MIT CSAIL), Darren Lunn (University of Manchester), Robert C. Miller (MIT CSAIL), Jeffrey Nichols (IBM Almaden), Marius Pasca (Google), Craig Prince (University of Washington), I.V. Ramakrishnan (Stony Brook University), Hironobu Takagi (IBM Japan), Jacob Wobbrock (University of Washington), Lindsay Yazzolino (Brown University), Tom Yeh (University of Maryland), and Bernie Zhang (Pennsylvania State University).