CHARACTERIZING AND VISUALIZING PHYSICAL WORLD ACCESSIBILITY AT SCALE USING CROWDSOURCING, COMPUTER VISION, AND MACHINE LEARNING





Figure 7. Our vision is to transform the way street-level accessibility information is collected and visualized. With our new scalable data collection methods, we aim to support a new class of accessibility-aware map tools such as (a) accessibility-aware navigation tools that provide personalized route information based on a user's reported mobility level and (b) visual analytic tools that allow citizens and governments to easily assess a city's street-level accessibility.

Introduction

Poorly maintained sidewalks and street intersections pose considerable accessibility challenges for people with mobility-impairments [13,14]. According to the most recent U.S. Census (2010), roughly 30.6 million adults have physical disabilities that affect their ambulatory activities [22]. Of these, nearly half report using an assistive aid such as a wheelchair (3.6 million) or a cane, crutches, or walker (11.6 million) [22]. Despite comprehensive civil rights legislation for Americans with Disabilities (e.g., [25,26]), many city streets, sidewalks, and businesses in the U.S. remain inaccessible. The problem is not just that street-level accessibility fundamentally affects where and how people travel in cities, but also that there are few, if any, mechanisms to determine accessible areas of a city a priori. Indeed, in a recent report, the National Council on Disability noted that they could not find comprehensive information on the "degree to which sidewalks are accessible" across the US [15]. This lack of information can have a significant negative impact on the independence and mobility of citizens [13,16] For example, in our own initial formative interviews with wheelchair users, we uncovered a prevailing view about navigating to new areas of a city: "I usually don't go where I don't know [about accessible routes]" (Interviewee 3, congenital polyneuropathy). Our overarching research vision is to transform the way in which street-level accessibility information is collected and used to support new types of assistive map-based tools.

Traditionally, sidewalk assessments have been conducted via in-person street audits [19,20], which are labor intensive and costly [17], or more recently, via smartphone applications, which are done on a reactive basis and require physical presence [27]. Although some cities offer sidewalk information online (*e.g.*, through government 311 databases [21]), these solutions are not



Figure 8. Our custom image labeling tools on web browsers: (a) The early version of the interface, which lets a user mark the location of the sidewalk problem and categorize the problem type (*e.g.*, an obstacle in a path) on a static GSV images [6]; (b) The labeling interface from our most recent work [10], which allows a user to adjust the camera angle (pan and zoom) and search for and label accessibility attributes in GSV.

comprehensive, rely on *in situ* reporting, and are not specifically focused on collecting and providing accessibility information. Some work exists on modeling and visualizing accessibility information in the built environment [3,12,13]; however, again these models are constrained by a lack of data describing street-level accessibility and the resulting systems have not been widely deployed.

In contrast, for the past three years our research group has been pursuing a twofold alternative vision [5–10]: **first**, to develop scalable data collection methods for remotely acquiring street-level accessibility information using a combination of crowdsourcing, computer vision, and machine learning along with online map imagery such as Google Street View (GSV) and high resolution top-down photographs such as satellite or flyover imagery. **Second**, to use this new data to design, develop, and evaluate a novel set of navigation and map tools for accessibility. For example, imagine a mobile phone application that allows users to indicate their ambulatory ability (*e.g.*, motorized wheelchair, walker) and then receive personalized, interactive accessible route recommendations to their destination (Figure 1a). As another example, inspired by walkscore.com, imagine an interactive map visualization tool that allows you to quickly assess the a city's street-level accessibility (Figure 1b)—how might such a tool impact where people choose to live, how governments invest in street-level infrastructure, or even how property values are assessed?

In this article, we provide a brief history of our work starting with initial studies exploring the viability of using GSV imagery as a reliable source of street-level accessibility and ending with a treatment of our current work on what we call assistive location-based technologies—location-based technologies that specifically incorporate accessibility features to support navigating, searching, and exploring the physical world. We close with a summary of open future work and a call to action for others in the community to work on this important problem.

Previous Research

A majority of our work thus far has focused on the first part of our vision: developing scalable data collection methods using a combination of crowdsourcing and automated methods to locate, identify, and characterize street-level accessibility attributes in GSV. Below, we discuss GSV as a viable physical-world accessibility data source, the development of our initial crowdsourcing labeling tools, and our more recent work on semi-automated labeling of accessibility features in GSV imagery.

Viability of GSV Imagery as a Source of Accessibility Information

We describe two threads of work evaluating GSV imagery as a viable source of street-level accessibility information: first, can people with similar mobility impairments find and agree on accessibility problems in GSV imagery? Second, given that GSV images are collected semi-infrequently, is image age a problem—that is, how well do problems identified in GSV represent the current state of the physical world?

Towards the first question, we recruited three electric wheelchair users to investigate whether they could consistently identify accessibility problems in GSV [6]. Independently, the three participants were asked to locate and categorize accessibility problems in 75 curated static images of GSV using our custom-made image labeling tool (Figure 8a). In addition, they participated in an exit interview where we asked about their personal experiences with street-level accessibility. Two key results emerged. First, our participants had high inter-labeler agreements, indicating that accessibility problems could be consistently identified solely from GSV imagery. Second, one of our participants stated that he already used GSV to examine an area for traversability before leaving his house—a result that has been echoed by more recent interviews that we conducted with 20 mobility impaired participants. Thus, though not specifically designed for this purpose, it appears that GSV is already being appropriated as a valuable source of accessibility information, reinforcing its use in our research.

Towards the second question, perhaps the most significant concern about using online map imagery to remotely collect accessibility information is image age. The built environment evolves over time and accessibility issues found via GSV may no longer exist and/or new problems may emerge. While Google does not publish information about how often its GSV cars drive and capture new images, major cities appear to be updated approximately once every year or two (e.g., downtown Washington DC has seven captures in eight years). Less populated cities are updated less. Although previous work has reported high concordance between audits conducted in the physical world vs. using GSV imagery [2,17], the focus was not on accessibility. To this end, we physically audited 273 intersections in nearby cities (Washington, D.C. and Baltimore, MD) and compared them with audits performed with GSV images. We found nearly perfect agreement despite an average GSV image age of 2.2 years (SD=1.3). The 6 disagreements were due to recent or ongoing construction. Thus, based on our own physical audit and reports from prior work, we are confident that GSV is a viable source of street-level accessibility information. In addition, with the movement towards self-driving cars, drone-based photography, and more frequently updated satellite imagery, we expect that GSV-like datasets will become even more common and more frequently updated in the future.

Crowdsourcing Sidewalk Accessibility Information

With GSV imagery established as a reasonable dataset to collect street-level accessibility information, we began developing and studying interfaces to allow minimally trained online users to remotely find, label, and characterize sidewalk accessibility. We performed multiple studies [5,6,8,10] with Amazon Mechanical Turk, an online labor market where users are paid to perform small tasks. In our earliest work [6], we manually collected and curated 229 GSV images from Washington DC, Baltimore, NYC, and LA. Using our custom labeling tool (Figure 2a), workers were asked to draw an outline around four main types of accessibility problems and indicate their severity (Figure 2a). Unlike the traditional GSV interface, users could not pan, zoom, or move around in this version of our labeling interface, which was done to simplify interactions. To create a ground truth dataset, two members of our research team independently labeled all 229 images

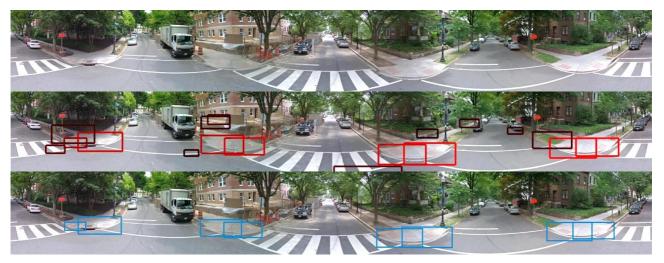


Figure 9. We developed Tohme, a scalable system for semi-automatically finding curb ramps in Google Streetview (GSV) panoramic imagery using computer vision, machine learning, and crowdsourcing. The images above show an actual result from our evaluation: (a) Raw Google Street View image, (b) results of computer vision curb ramp detection (lighter red is higher confidence), and (c) results after crowdsourced verification.

and found: 67 images with Surface Problems, 66 images with Object in Path, 50 with Prematurely Ending Sidewalk, and 47 with Curb Ramp Missing.

We then conducted an online experiment with 185 crowd workers. When compared to ground truth data, the workers correctly identified the presence of problems with 81% accuracy. Using majority voting as a simple quality control mechanism, this accuracy jumped to 87% with three workers and over 90% with five or more workers. We ran two subsequent crowdsourcing studies for collecting accessibility information on bus stop landmarks (*e.g.*, bus stop shelters) [8] and curb ramp infrastructure [10] and found similar results—that is, above 80% labeling accuracy with a single crowd worker per location. This indicates that minimally trained online workers are indeed capable of remotely finding and labeling street-level accessibility problems and that simple quality control mechanisms can be used to reach accuracies of around 90%. We believe that these results can be improved through better training, providing active monitoring with feedback to help users learn when they have made a mistake, and warning or even eliminating poor quality labelers from the system.

Increasing Data Collection Efficiency with Semi-automated Methods

While our prior work showed that crowd workers could find and label street-level accessibility problems with high accuracy, this sole reliance on human labor limited scalability. To this end, we investigated ways to combine computer vision and machine learning in the data collection process [9,10]. We created *Tohme* (Figure 3), a system to collect geo-located curb ramp data using a combination of crowdsourcing, computer vision, machine learning and online map data [10]. In this work, we only focused on sidewalk curb ramps because of their significance to accessibility as well as their visual saliency and geospatial properties (*e.g.*, often located on corners), which we thought would ease automated detection.

Key components of Tohme include: (i) a web scraper for downloading street intersection data; (ii) two crowd worker interfaces for finding, labeling, and verifying the presence of curb ramps (Figure 2b); (iii) state-of-the-art computer vision algorithms for automatic curb ramp detection; and (iv) a machine learning-based workflow controller, which predicts computer vision performance and dynamically allocates work to either a human labeling pipeline or a computer vision + human verification pipeline. The system workflow is as follows. svDetect processes every

GSV scene producing curb ramp detections with confidence scores. svControl predicts whether the scene is difficult for a computer vision algorithm. If svControl predicts that the automated detections are likely to fail on a given scene, the detections are discarded and the scene is fed to svLabel for manual labeling instead. If not, the scene/detections are forwarded to svVerify for human verification. The workflow attempts to optimize accuracy and speed.

To evaluate Tohme, we conducted a study using data collected from 1,086 intersections across four North American cities. We evaluated Tohme's performance in detecting curb ramps across our entire dataset with 403 turkers. Alone, the computer vision sub-system currently finds 67% of the curb ramps in the GSV scenes, indicating that computer vision alone cannot solve this complex problem. However, by dynamically allocating work to the CV module or to the slower but more accurate human workers, Tohme performs similarly in detecting curb ramps compared to a manual labeling approach alone (F-measure: 84% vs. 86% baseline) but at a 13% reduction in human time cost. To put this in context, for a medium sized city like Washington, D.C. (which has 8,209 intersections [21]), we can reduce the cost to collect curb ramp labels by 30 human hours (from 214 to 184 human hours). This is just the beginning. Our overall aim is to create semi-automated methods that reduce total human hours by at least an order of magnitude. Though challenging, we think we can get there with new workflow algorithms, additional advances in computer vision applied to built infrastructure (*e.g.*, [1]), and better user interfaces.

Ongoing Research

In summary, our previous work demonstrated (i) the viability of using GSV as a massive source of street-level accessibility information, (ii) the feasibility of using crowdsourcing to identify accessibility problems, and (iii) methods to combine computer vision and machine learning techniques to increase the scalability of the data collection methods.

Building on the above work, we are currently focused on two trajectories: (i) investigating how to coordinate crowds and machines to further increase the efficiency of our methods; and (ii) designing and developing the accessibility-aware applications that mentioned in the introduction (Figure 1). Towards (i), we are exploring new methods to semi-automatically separate and triage areas that need accessibility inspections. For example, with our labeling interfaces, we can randomly place crowd workers anywhere in a city (virtually via GSV). If some of these workers begin reporting significant accessibility issues, we can begin triaging those areas and assigning additional workers (and fewer workers to other areas). Relatedly, we are also investigating techniques to try and assess under examined areas in real-time. For example, imagine using the accessibility-aware navigation smartphone application shown in Figure 1a. If you inquire about a potential route that lacks accessibility information, we would like to develop methods capable of semi-automatically crowdsourcing that information in near real-time (similar to VisWis [4]).

Towards (ii), we are designing and developing what we call assistive location-based technologies—location-based technologies that are geared towards supporting people with disabilities (Figure 7) to show the transformative value of our accessibility data. For example, our accessible heatmap mockup shown in Figure 7a would allow users to quickly understand and explore accessible areas (green) and inaccessible areas (red) of their cities and to 'drill down' into specific neighborhoods. Our hope is that this would allow people with mobility impairments to make better decisions about where to live in a city or where to stay when they are traveling. Similarly, we are working on accessible path recommendations depending on the user's

reported mobility level. Crucially, however, both these tools require large amounts of geo-located accessibility information—exactly what our scalable data collection methods hope to provide. Furthermore, we believe that these tools will enable governments, public health researchers, and urban planners to more easily assess the health of neighborhoods and to help them smartly allocate resources to improve city infrastructure. Our tools should also provide value to non-mobility impaired persons—for example, those with situational impairments due to pushing a baby stroller, pulling a cart, etc. And, ideally, our hope is that the data could be integrated into pre-existing map-based tools such as Google Maps or OpenStreetMap rather than exist solely in specialized research prototypes.

To collect large amount of accessibility data to advance these projects, we are transforming our data collection tools—which thus far have only been deployed on Amazon Mechanical Turk—into public facing applications so anyone can contribute to the data collection. Inspired and informed by online citizen science website (*e.g.*, zooniverse.org), we are creating a webpage that allows both volunteers and paid crowd workers (turkers) to label the accessibility problems in the physical world. Here, we are investigating, for example, if intrinsically motivated people like wheelchair users or their caregivers perform differently or provide different types of labels from turkers. The image labeling interface is similar to the one shown in Figure 8b with small updates like more label types, ability to freely "walk" around in the virtual world, and detailed feedback on their contribution to collecting street-level accessibility data. Our interfaces also allow users to comment and upload more recent photos if there is a discrepancy in GSV.

As a start, we are collecting accessibility data in two US cities: Washington, D.C. and Baltimore. These cities were selected because of their relatively large population and land area [23,24] as well as their proximity to the University of Maryland, which makes them both convenient for conducting on-site audits. Based on our own calculation using OpenStreetMap (openstreetmap.org), Washington, D.C. and Baltimore have total street lengths of: 670 mi and 1,400 mi respectively. We are planning to ask multiple contributors to audit each street and label sidewalk accessibility problems in GSV, which will allow us to get more accurate data through majority vote based data aggregation (*i.e.*, similar to our prior work [6,8]).

Future Work

We will publish the collected accessibility information as a data dump and provide API access. We hope that this will enable and spur the development of a broad range of new applications and provide new tools for research. It will offer opportunities for HCI researchers and commercial entities to design accessibility-aware tools beyond what we described above, and we invite you to join us in these efforts. For example, we imagine a tool like Yelp incorporating our accessibility data to enhance its search capability—restaurants could be searched not only with location, cuisine and reputations but also based on their level of accessibility. The accessibility data could also be used in broad interdisciplinary research areas. For instance, we expect public health researchers and urban planners to use our data to analyze relationship between neighborhood accessibility and health of those who live there, similar to the studies that investigated how neighborhood characteristics like the presence of amenities (*e.g.*, recreational facilities) and perceived safety affected residents' physical activity levels [11,18].

The key future challenges are to collect comprehensive data about indoor accessibility and capture changes in accessibility of the built environment. While mobile crowdsourcing applications like Wheeelmap and AXSMap attempt to collect granular indoor accessibility

information (*e.g.*, presence of accessible bathroom), the data remains sparse due to limited adoption by users and there are no known scalable data collection solutions. Similarly, there are no prescribed ways to react to temporary accessibility barriers that arise in daily or hourly basis (*e.g.*, constructions that obstruct sidewalks, changes in pedestrian density). Some possible solutions here include exploring the use of potentially rich but untapped sources of accessibility information such as daily-updated satellite imagery (*e.g.*, planet.com) or even surveillance video streams (*e.g.*, placemeter.com).

Conclusion

We described our twofold vision to, first, invent and study new scalable methods to collect streetlevel accessibility information and, second, to use this data to design, develop, and evaluate new map-based tools for accessibility. Our research thus far has demonstrated that GSV is a viable, massive untapped data source for accessibility information, that minimally trained crowd workers are capable of locating, labeling, and characterizing accessibility problems in GSV images using specially designed interfaces, and that automated methods can be used to increase the efficiency of data collection. Our on-going efforts include design, development, and evaluation of scalable data collection system in the wild as well as development of accessibility-aware applications. We expect our work to open up future research avenues in areas not limited in HCI, but also in public health, urban planning, and GIS. This is a large, on-going research effort and we are always looking for interested collaborators. Please feel free to contact us for more information.

Acknowledgments

This work was supported by an NSF grant (IIS-1302338), a Google Faculty Research Award, and an IBM Ph.D. Fellowship. We thank Professor David Jacobs and his student Jin Sun for their work on automatically identifying curb ramps using computer vision and to the many high school, undergraduate, and graduate students who have contributed to this project thus far. See the individual papers for a detailed list [5–10].

References

1. Dragan Ahmetovic, Roberto Manduchi, James M. Coughlan, and Sergio Mascetti. 2015. Zebra Crossing Spotter: Automatic Population of Spatial Databases for Increased Safety of Blind Travelers. *The 17th International ACM SIGACCESS Conference on Computers and Accessibility* (ASSETS 2015).

Hannah M Badland, Simon Opit, Karen Witten, Robin A Kearns, and Suzanne Mavoa. 2. 2010. Can virtual streetscape audits reliably replace physical streetscape audits? Journal of urban health : bulletin of the New York Academy of Medicine 87, 6: 1007–16. http://doi.org/10.1007/s11524-010-9505-x

3. Linda Beale, Kenneth Field, David Briggs, Phil Picton, and Hugh Matthews. 2006. Mapping for Wheelchair Users: Route Navigation in Urban Spaces. *Cartographic Journal, The* 43, 1: 68–81. http://doi.org/10.1179/000870406X93517

4. Jeffrey P Bigham, Chandrika Jayant, Hanjie Ji, et al. 2010. VizWiz: nearly real-time answers to visual questions. *Proceedings of the 23nd annual ACM symposium on User interface software and technology*, ACM, 333–342. http://doi.org/10.1145/1866029.1866080

5. Kotaro Hara, Victoria Le, and Jon Froehlich. 2012. A Feasibility Study of Crowdsourcing and Google Street View to Determine Sidewalk Accessibility. *Proceedings of the 14th international*

ACM SIGACCESS conference on Computers and accessibility (ASSETS '12), Poster Session, ACM, 273–274. http://doi.org/10.1145/2384916.2384989

6. Kotaro Hara, Victoria Le, and Jon Froehlich. 2013. Combining Crowdsourcing and Google Street View to Identify Street-level Accessibility Problems. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 631–640.

7. Kotaro Hara, Victoria Le, Jin Sun, David Jacobs, and Jon Froehlich. 2013. Exploring Early Solutions for Automatically Identifying Inaccessible Sidewalks in the Physical World Using Google Street View. *Human Computer Interaction Consortium*. Retrieved from http://kotarohara.com/public/Papers/Hara_ExploringEarlySolutionsForAutomaticallyIdentifyingIn accessibleSidewalksInThePhysicalWorldUsingGoogleStreetView_HCIC2013Boaster.pdf

8. Kotaro Hara, Azenkot Shiri, Megan Campbell, et al. 2013. Improving Public Transit Accessibility for Blind Riders by Crowdsourcing Bus Stop Landmark Locations with Google Street View. *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility Technology*, 16:1–16:8.

9. Kotaro Hara, Jin Sun, Jonah Chazan, David Jacobs, and Jon Froehlich. 2013. An Initial Study of Automatic Curb Ramp Detection with Crowdsourced Verification using Google Street View Images. *Proceedings of the 1st Conference on Human Computation and Crowdsourcing (HCOMP'*13), *Work-in-Progress*.

10. Kotaro Hara, Jin Sun, David W Jacobs, and Jon E Froehlich. 2014. Tohme: Detecting Curb Ramps in Google Street View Using Crowdsourcing, Computer Vision, and Machine Learning. *Proceedings of the 27th annual ACM symposium on User interface software and technology*.

11. Christine M Hoehner, Laura K Brennan Ramirez, Michael B Elliott, Susan L Handy, and Ross C Brownson. 2005. Perceived and objective environmental measures and physical activity among urban adults. *American Journal of Preventive Medicine* 28, 2, Supplement 2: 105–116. http://doi.org/http://dx.doi.org/10.1016/j.amepre.2004.10.023

12. Roger L Mackett, Kamalasudhan Achuthan, and Helena Titheridge. 2008. AMELIA: A tool to make transport policies more socially inclusive. *Transport Policy* 15, 6: 372–378. http://doi.org/http://dx.doi.org/10.1016/j.tranpol.2008.12.007

13. Hugh Matthews, Linda Beale, Phil Picton, and David Briggs. 2003. Modelling Access with GIS in Urban Systems (MAGUS): capturing the experiences of wheelchair users. *Area* 35, 1: 34–45. http://doi.org/10.1111/1475-4762.00108

14. Allan R Meyers, Jennifer J Anderson, Donald R Miller, Kathy Shipp, and Helen Hoenig. 2002. Barriers, facilitators, and access for wheelchair users: substantive and methodologic lessons from a pilot study of environmental effects. *Social Science & Medicine* 55, 8: 1435–1446. http://doi.org/10.1016/S0277-9536(01)00269-6

15. National Council on Disability. 2007. *The Impact of the Americans with Disabilities Act: Assessing the Progress Toward Achieving the Goals of the ADA*. Washington, DC, USA. Retrieved from http://goo.gl/RkLL63

16. Andrea Nuernberger. 2008. Presenting Accessibility to Mobility-Impaired Travelers (Ph.D. Thesis).

17. Andrew G Rundle, Michael D M Bader, Catherine A Richards, Kathryn M Neckerman, and Julien O Teitler. 2011. Using Google Street View to audit neighborhood environments. *American journal of preventive medicine* 40, 1: 94–100. http://doi.org/10.1016/j.amepre.2010.09.034

18. James F Sallis, Marilyn F Johnson, Karen J Calfas, Susan Caparosa, and Jeanne F Nichols. 1997. Assessing Perceived Physical Environmental Variables that May Influence Physical Activity. *Research Quarterly for Exercise and Sport* 68, 4: 345–351. http://doi.org/10.1080/02701367.1997.10608015 19. Edward R. Stollof and Janet M. Barlow. 2008. *Pedestrian Mobility and Safety Audit Guide*. Retrieved from http://www.ite.org/PedAudits/AuditGuide.pdf

20. Streets Wiki. Walk Audit. Retrieved from http://streetswiki.wikispaces.com/Walk+Audit

21. The District of Columbia. Data Catalog http://opendata.dc.gov/.

22. U.S. Census Bureau. 2012. *Americans with Disabilities: 2010 Household Economic Studies*. Retrieved from http://www.census.gov/prod/2012pubs/p70-131.pdf

23. United States Census Bureau. 2013 Population Estimates. Retrieved October 1, 2015 from http://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=bkmk

24. United States Census Bureau. Cities with 100,000 or More Population in 2000 ranked by Land Area. Retrieved from https://www.census.gov/statab/ccdb/cit101or.txt

25. Civil Rights Division United States Department of Justice. 1990. *Americans with Disabilities Act of 1990, Pub. L. No. 101-336, 104 Stat. 328.*

26. United States Department of Justice. 2010. 2010 ADA Standards for Accessible Design.

27. SeeClickFix http://seeclickfix.com/apps (Viewed Nov 2014).

About the Authors:





Kotaro Hara is a sixth year Ph.D. student in the Department of Computer Science at the University of Maryland, College Park advised by Prof. Jon E. Froehlich. His research focuses on designing, building, and evaluating systems powered by both people and machines to improve the accessibility of the physical world. He received a B.E. in Information Engineering from Osaka University in 2010 and was named an IBM Ph.D. Fellow in 2014.

Jon Froehlich is an Assistant Professor in Computer Science at the University of Maryland, College Park where he is member of the Human-Computer Interaction Laboratory (HCIL) and the designer, builder, and founder of the HCIL Hackerspace. His research group, the Makeability Lab, focuses on designing and studying novel interactive experiences that cross between bits and atoms-the virtual and the physical-and back again to confront some of the world's greatest challenges in environmental sustainability, health and wellness, education, and universal accessibility. This work has been recognized with eight paper awards including three best papers (CHI'10, CHI'13, and ASSETS'13), the William Chan Memorial Dissertation Award, and the University of Washington Graduate School Award. Distinguished Dissertation Learn more here: http://www.cs.umd.edu/~jonf/