FootNotes: Geo-referenced Audio Annotations for Nonvisual Exploration

COLE GLEASON, Carnegie Mellon University, USA
ALEXANDER J. FIANNACA, Microsoft Research, USA
MELANIE KNEISEL, Microsoft Research, USA
EDWARD CUTRELL, Microsoft Research, USA
MEREDITH RINGEL MORRIS, Microsoft Research, USA

The majority of information in the physical environment is conveyed visually, meaning that people with vision impairments often lack access to the shared cultural, historical, and practical features that define a city. How can someone who is blind find out about the sleek skyscrapers that dot a modern city's skyline, historic cannons that have been remade into traffic pillars, or ancient trees that uproot a neighborhood's sidewalks? We present FootNotes, a system that embeds rich textual descriptions of objects and locations in OpenStreetMap, a popular geowiki. Both sighted and blind users can annotate the physical environment with functional, visual, historical, and social descriptions. We report on the experience of ten participants with vision impairments who used a spatialized audio application to interact with these annotations while exploring a city. By sharing rich annotations of physical objects and areas, FootNotes helps people thoroughly explore a new location or serendipitously discover previously unknown features of familiar environments.

CCS Concepts: • **Human-centered computing** → **Accessibility technologies**; *Computer supported cooperative work*; Ubiquitous and mobile computing systems and tools; • **Social and professional topics** → *People with disabilities*;

Additional Key Words and Phrases: Visual impairment, blindness, navigation, assistive technology, mixed reality, augmented reality

ACM Reference Format:
Cole Gleason, Alexander J. Fiannaca, Melanie Kneisel, Edward Cutrell, and Meredith Ringel Morris. 2018. FootNotes: Geo-referenced Audio Annotations for Nonvisual Exploration. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 3, Article 109 (September 2018), 24 pages. https://doi.org/10.1145/3264919

1 INTRODUCTION

Exploring the physical world independently is often challenging for people with vision impairments, especially in unfamiliar environments. The ubiquity of smartphones and the localization technology they contain (e.g., GPS) has made exploration and navigation tools feasible (e.g., Google Maps, Blindsquare), but these mainly focus on providing turn-by-turn navigation directions or simply listing the names of nearby businesses without much additional functional or visual context. These limitations are the result of the limited types of location-based datasets available. Business-focused datasets such as that of Foursquare and Yelp do not include data for visual

Authors' addresses: Cole Gleason, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA, 15213, USA, cgleason@cs.cmu.edu; Alexander J. Fiannaca, Melanie Kneisel, Edward Cutrell, and Meredith Ringel Morris, Microsoft Research, 14820 NE 36th Street, Redmond, WA, 98052, USA, {alfianna, mekne, cutrell, merrie}@microsoft.com.

2474-9567/2018/9-ART109 $15.00
https://doi.org/10.1145/3264919

or functional features of spaces and objects outside of the businesses they describe, such as benches and water fountains. Map-focused datasets such as Google Maps and OpenStreetMap contain data that primarily supports generating navigational routes rather than data that supports learning about and exploring areas. A solution is needed to make the plethora of visual information that sighted people use to understand a city accessible to people with vision impairments, in order to fill in the gaps in current location-based datasets.

Automatically collecting the diverse range of visual information and delivering it in a more accessible format, such as audio, is not simple. Computer vision approaches for reading signs in natural scenes have improved [1], but OCR (optical character recognition) is still highly dependent on lighting and other environmental conditions. Recently, general object recognition algorithms have also achieved higher accuracy, but still lack robustness and fine-grained detail about an object [23]. Scene recognition algorithms [27] also fail to adequately describe scenes in-depth, or they are too inaccurate to trust while navigating.

Where people with vision impairments cannot solely rely on automated applications, crowdsourcing can provide rich, understandable descriptions of visual information. Several projects have used crowd workers connected via the internet and mobile phones to assist people with vision impairments to describe digital images or videos [1, 3]. Fewer applications exist for the physical world, but many landmarks are already described on sites like Wikipedia. Prior work on Wikipedia has shown that contributors can collaborate to produce high-quality articles [25], and specific wikis for mapping applications have managed to gather a large amount of structured spatial data [19]. These geographic wikis (geowikis) could act as a platform to generate and moderate written geo-referenced descriptions of visual information.

We seek to enhance the navigation experience for people with vision impairments by collecting these descriptions of visual information and storing them in an accessible format for people with vision impairments. Instead of focusing on turn-by-turn route guidance to reach a destination, we intend to make environmental visual information more available, as this often adds context to a physical space. To achieve this goal, we introduce a novel system, FootNotes, which makes collectively-created annotations of the physical world available to people with vision impairments while navigating. We modified Open Street Maps, a geographic wiki, to store visual, functional, historical, and social annotations for points of interest. Building on top of an existing navigation system, we made these descriptions available to ten blind participants as they explored downtown Kirkland, WA. We report on their experiences using FootNotes, including the usability of the system and the perceived value of different annotation types. We conclude by discussing the implications of our findings for developing augmented reality applications for people with vision impairments that can support exploration and navigation through a rich understanding of the physical environment.

2 RELATED WORK

Our research on the FootNotes system was informed by prior work in four areas: geographic wikis (geowikis)., spatially-located annotations, navigation systems for people who are blind, and alternative text for digital images.

2.1 Geographic Wikis

Wikis allow collaborative creation and editing of shared resources, and geographic wikis have also been used to generate community-maintained maps. OpenStreetMap (OSM) is one of the most well-known geowikis, and contains accurate general-purpose information that can be incredibly rich [19]. Some well-annotated areas of the map include labels for trees and benches in urban parks.

OpenStreetMap stores all of its geographic data in three element types: nodes, ways, and relations [9]. Nodes represent points in space, and would be used to represent a small point of interest. Ways are collections of nodes that define a linear path, such as a road or a boundary around a larger point of interest. Relations are used to indicate a relationship between OSM elements, such as indicating all of the ways that make up a state highway.

All modifications to OSM elements are contained in changesets, each of which contains a group of modifications to the OSM map by a single author [7]. Changesets and OSM elements can all include tags, which are key-value pairs of structured additional information. For example, "oneway=yes" on a way indicates a path where travel is only allowed in one direction [10]. FootNotes uses a tag to attach text annotations to spatially-located points of interest (typically OSM Nodes).

The Cyclopath project [30], a geowiki for cyclists in Minneapolis-Saint Paul, examined how their users created and used tags for cycling [33]. They found that over 90% of tags on point objects (similar to OSM nodes) consisted of factual information, while blocks (similar to OSM ways) contained subjective information more often than points. Both factual and subjective tags provided detailed information about the road network that were used to find good cycling routes. Subjective tags that users added to blocks included tags such as "dangerous" or "scenic", whereas tags such as "bad coffee" were added to points of interest like cafés. The authors also found that tags describing blocks tended to be adjectives, while points of interest were described with nouns instead. As Cyclopath already included free-form note annotation, the introduction of tags led to tags that overlapped with note content and often provided a form of "quick note" that was easily filterable in their route search engine. The Cyclopath team also examined authoring behavior and methods to focus more contributions in a specific geographic area [29]. These findings from users of a small segment of the local population, cyclists, lead us to believe that similar authoring habits could be maintained by FootNotes users and local volunteers to provide annotations of value for people with vision impairments.

2.2 Spatially-Located Annotations

Several research projects have explored the creation and use of spatial annotations for sighted users, especially as mobile devices have become widespread. FootNotes takes a similar approach to some of these systems, especially GeoNotes, which replicated a "Post-it Notes" concept using PDAs [28]. The authors of the GeoNotes paper highlighted that note authors can refer to the shared spatial context of the reader, and allowed the virtual notes to be placed on user-defined "places" such as the "coffee machine." Similarly, FootNotes allows for creating new labels when one does not already exist, but annotation authors are primarily attaching annotations to existing OpenStreetMap objects.

ActiveCampus used a WLAN network on a college campus to facilitate spatial annotations [17], but instead of being attached to objects, the annotations were roughly spatially located by building. They found that most of their student participants used the ActiveCampus Explorer system as an instant messaging service, and posted annotations in reply to other users.

Storyplace.me [2] was an Android application that allowed sighted users to leave geo-tagged social messages for their contacts throughout an urban area. A major theme in their interviews with participants was that users found new meaning in their city based on location-specific stories from other users; this helped the participants see how an area had changed over the years or learn about the significance of a place. The Storyplace.me users were also interested in leaving social messages for friends or family members around the city.

A few applications have explored special-purpose location-based labeling for people who are blind. The OneBusAway app is an accessible application that provides sighted and blind users with information about public transit; Hara et al. [20] extended this application by allowing sighted bus-stop visitors to add annotations describing physical properties of bus stops that were important to blind transit riders (e.g., the presence of benches, trash cans, etc.). Hara et al. [21] also explored the use of paid crowd workers to annotate Google Street Maps with information about accessibility dangers (e.g., cracks in the sidewalk or other obstacles) that might impact blind or motor-impaired pedestrians. These two systems are examples of specific types of geo-referenced functional annotations, which FootNotes provides generalized support for by annotating any item in the environment.

2.3 Navigation Systems for People who are Blind

Over the past few decades, researchers have built navigation systems designed especially for people with vision impairments using technology like GPS localization [13]. Many people who are blind also use mainstream smartphone apps that provide turn-by-turn walking directions, such as Google Maps. While spatial annotations are usually not their focus, navigation systems built for people with visual impairments have incorporated some small features utilizing them in their design. Typically, this is in the form of a database of points of interest that help the user locate nearby businesses or landmarks. The Personal Guidance System, one of the first navigation systems that utilized GPS, included a database of nearby points to aid guidance [16]. More recently, Navatar used some physical landmarks as part of the localization system for indoor navigation [12], but did not provide detailed information to the user beyond their presence. Commercial applications for people with vision impairments, such as Seeing Eye GPS [8] and Blindsquare [22], include modes for exploring nearby points of interest; both of these applications announce point of interest names as the user nears them, but typically just the name and category. Microsoft Soundscape [31] has similar functionality, and also displays points of interest in 3D spatialized audio using a head-mounted accelerometer and audio headset. Our FootNotes system builds on Soundscape, to help the user perceive the locations of objects being described, although the annotations supported by FootNotes could also be conveyed through a more traditional (non-3D) walking directions app. We see the clear utility in including many diverse points of interest, but the provided information should go further than just the name and type of landmark. FootNotes provides rich annotations of visual information, similar to the annotations attached to digital images on the web (e.g., alternative text or other captions read aloud by a screen reader).

2.4 Alternative Text for Digital Images

Image annotations on the Internet, also known as “alternative text” or “alt text”, are read by screen reader software to describe images to people with vision impairments. Unfortunately, prior studies have shown that over 50% of images on major websites are not annotated [4] To rectify this, researchers have proposed social solutions to label images that the original website creator did not annotate. ALT-Server was the first example of a system to share alternative text, which created a centralized web server for users to host alternative text for a specific image URL [11]. Takagi et al. took this idea further by showing how users can collaborate to make a site more accessible [32]. Researchers have explored many models for incentivizing participation in the creation of alt text for digital images, such as games with a purpose [34], paid crowdsourcing [3], friendsourcing [6], and social microvolunteering [5].

3 FOOTNOTES SYSTEM DESCRIPTION

Inspired by alternative text online and previous spatial annotation systems, we designed an annotation system for people with vision impairments called FootNotes. FootNotes aims to make the physical world more accessible through rich human-composed annotations of object attributes that navigation systems for the blind have not incorporated. Our system builds on top of the Microsoft Soundscape product [31]; Soundscape uses 3D audio to announce nearby points of interest such that the audio alerts appear to originate from the correct physical direction when the user wears a 3D audio-capable headset.

3.1 Microsoft Soundscape

Microsoft Soundscape is a navigation tool for people with vision impairments that focuses on announcing nearby points of interest instead of providing route guidance. Using the phone or an external GPS receiver, Soundscape locates the user and calls out points of interest as the user approaches them. A call-out contains the name of the point of interest and the distance from the user. When the user is wearing a compatible headset with an inertial

measurement unit (IMU), Soundscape also tracks the relative orientation of the user's head. Call-outs and sounds are then rendered with 3D audio and appear to come from the direction of the point of interest.

Soundscape includes four main actions the user can take in addition to automatic announcements of nearby points of interest. The first action, "My Location," will immediately tell the user which direction they are facing, the street they are on, and the closest intersection. The second action, "Around Me" selects four close points of interest in four different quadrants around the user. It announces them in succession to help the user orient themselves. The third action. "Ahead of Me," selects five points of interest that are in the direction that the user is facing. It then announces them, with their distance from the user, from closest to farthest. Finally, the last action a user can take is setting an "Audio Beacon" on a point of interest. This creates a repetitive background sound that is always orientated towards the point of interest relative to the user's head. This helps the user keep a destination or reference point in mind while navigating.

A person navigating with Microsoft Soundscape can interact with the system through the smartphone screen (via the screen reader) as shown in Figure 1 or a handheld remote. The remote exposes the most common actions: telling the user their current location, listing points of interest around or in front of the user, and silencing the app.

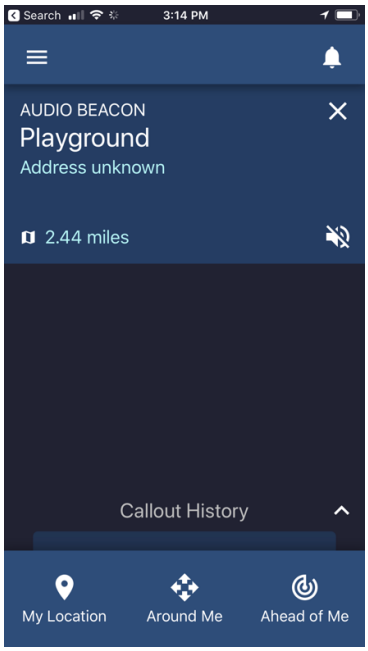

Fig. 1. In this screenshot of the Microsoft Soundscape application, three main functions are indicated by the buttons on the bottom of the screen. An audio beacon is set for a playground in the city.

3.1.1 Changes to Microsoft Soundscape. We made several changes to Microsoft Soundscape to support FootNotes' rich textual annotations for points of interest. First, we added the ability for users to add or edit annotations for four different categories (Section 3.2) in the smartphone application (Section 3.3). This allows for blind or sighted users to add or change information in situ. Microsoft Soundscape was also modified to support playback of annotations by attaching them to existing call-outs for points of interest (Section 3.4). Additionally, we over-rode the core functionality of the hand-held remote to support FootNotes, by replacing the functionality of one of the buttons; rather than using the fourth button to silence the app, the button would play any FootNotes annotations associated with the most recent point of interest that had been announced.

3.2 FootNotes Annotation Types

Building on top of the Soundscape navigation system, our FootNotes application supports augmenting geo-graphically-specified points of interest with textual annotations. Inspired by prior work and information that was already included in applications for people with vision impairments, we included four types of annotations: functional, visual, historical, and social; Figure 2 and our Video Figure provide examples of each annotation type. These four categories embody a large amout of information already accessible to sighted people in physical spaces, but we do not believe they are exhaustive.

3.2.1 Functional Annotations. Much of the information that is provided by online maps or databases of points of interest is practical in nature. Google Maps includes hours of business, how busy a place is at a specific time, and information on wheelchair accessibility. Yelp tells the user if a restaurant accepts credit cards and if it is noisy. These all help the user understand a place before travelling there.

Similarly, applications that provide navigation instructions for people with vision impairments typically include accessibility information. Online maps, including OpenStreetMaps, include the presence of tactile paving or

audible traffic signs. These are functional pieces of information that help someone with a vision impairment access physical spaces, and we included that category in FootNotes to capture specific functional tips for points of interest.

FootNotes' functional annotations indicate how one should use a point of interest, such as hours for a business or accessibility information (e.g., describing the location of stairs or obstacles). The functional information can be gleaned from websites associated with certain points of interest, as well as added in situ by volunteer contributors.

Fig. 2. A statue that was annotated as a point of interest in our user study. It had each of the annotation categories: Functional: "Kids often climb and play around this artwork" Visual: "This bronze artwork depicts six young children holding hands while skipping and jumping over puddles. Their faces show joy, with some seeming to yell or shout with excitement." Social: "Aaron says: I've always found this art to be really creepy. It's like a bunch of children zombies racing down the hill to feast on living kids at the water." Historical: "Resident and art collector Bill Ballantine loaned the sculpture to the city in 1990, and ten years later a grassroots effort organized to raise $250,000 to buy it. The artist, Glenna Goodacre, is well known for designing the front of the Sacagawea golden dollar coin."

3.2.2 Visual Annotations. On the internet, images convey visual content that is inaccessible to people with vision impairments. Well-written alternative text alleviates this issue by describing the visual content. For videos, audio descriptions perform a similar function by describing the visual content onscreen. In the physical world, a small subset of signs also contain Braille to make their textual content accessible, but nothing similar exists to describe other objects. Statues, street art, and other design aspects convey contextual information about an area, but are inaccessible.

FootNotes' visual annotations convey this visual information, such as color or material. This information can enrich a user's experience of the aesthetics of their environment, but also may provide secondary functional information (e.g., "I'm by the red building"). Visual information can be provided by sighted volunteers viewing an object in situ or viewing it online via images on services like Google Street View [21]. As computer vision tools improve in accuracy, some visual annotations could also be automatically generated (or automatically generated and then crowd-corrected).

3.2.3 Historical Annotations. When describing physical places or objects, it is common to reference their history: when was this created, by whom, and why? Whether it is providing background information on a neighborhood or specific details about a local landmark, historical information provides a lot of context that can enhance the understanding of a physical space. However, this content is often conveyed in situ by inaccessible signs, or can be found in books or online articles. Audio self-guided tours, such as those used in museums and other historical exhibits, provide an example of how annotations can be made accessible. These tours pair numbered signs with a set of matching audio recordings describing an object or place.

FootNotes' historical annotations tell the user the background of the object and/or its creator. These annotations often convey why the object is there, or the significance of it to others; this information can be gathered from online sources (e.g., Wikipedia entries), crowd-sourced based on contributors' local knowledge, and/or transcribed from signage affixed to historic objects that would not be visible to people who are blind.

3.2.4 Social Annotations. Storyplace.me [2], as well as other social geo-tagging applications, allow users to leave messages at a specific location for their friends to encounter. The users of the Storyplace.me Android application used the information from other users to better understand the area they were in, and we believe that geo-tagged social messages delivered via audio could be a valuable source of information for people with vision impairments. This feature can be used to share information about an experience with a friend ("Remember when we were here?"), or to provide personalized descriptions for one's social network ("This is my favorite spot for breakfast"). We also envision it as a platform for personalized audio landmarks from Orientation & Mobility instructors for their clients ("This is the center of the plaza, continue towards the fountain to reach your office building.").

FootNotes supports these social annotations as geo-referenced messages that can be left by one user for another. While our initial implementation focuses on messages composed within our system, the concept could be expanded to incorporate messages from other social media that use geo-location, such as sharing friends' ratings or comments from sites like Foursquare, Blindsquare, TripAdvisor, Yelp, etc. The connection to a specific point-of-interest for social messages helps to prevent the chat-like behavior observed in the ActiveCampus Explorer [17] system from crowding out other useful annotations.

3.3 Creating and Modifying Annotations

Sighted online contributors can create and modify annotations remotely. We chose to have all FootNotes annotations be free-form text, because text is more easily edited and universally accessible than other mediums, such as audio snippets. Text can be easily read using a screen-reader with text-to-speech, and played back at a variable speed or with a different voice. Text is also accessible to people with hearing impairments, unlike speech audio recordings.

Fig. 3. Partial screenshot of the modified iD editor to create/modify FootNotes annotations alongside OpenStreetMaps data. The FootNotes annotations are being edited in the left sidebar.

We modified an open source web editor (Figure 3) for Open Street Maps to include fields for attaching annotations from each of the four categories (functional, visual, historical, and social). When any of these fields are filled out, they are saved online, and linked to the point of interest in Open Street Maps. Other users can see these annotations and modify them to add more content or to make them more concise.

Blind users or sighted volunteers may also wish to add or edit annotations on the spot, especially if they discover interesting information. For example, someone might want to annotate that construction is underway on a street where they walk, or how to reach the information desk in a hotel lobby. Both blind and sighted users can add annotations through the FootNotes smartphone application by pressing a button on the main screen. This opens a pop-up window where they can select the annotation category, write the annotation, and choose a nearby point of interest to attach it to. This pop-up window is compatible with the iOS screen reader, and users may input data using the keyboard or speech-to-text. If the point of interest they wish to annotate is not available, they may attach the annotation to the current latitude and longitude. This will generate a note in Open Street Maps with the annotation, and an online contributor can later attach the annotation to a new point of interest.

When a FootNotes annotation is created, an OSM tag is attached to the relevant node, where the tag's value is a URL to the annotation content. Each modification to an annotation creates a new version, so the OSM changeset history also reflects the FootNotes annotation history.

3.4 Experiencing Annotations

The main value in these annotations is accessing them while navigating outdoors using a smartphone, which is supported by our FootNotes iOS application. As the user navigates and hears points of interest read aloud by the underlying navigation app, FootNotes plays an "earcon" that sounds like a chime after announcing the name of any point of interest that has associated annotations. If the user then presses the FootNotes button on the handheld remote, other speech from the system is paused and the annotations are read aloud. The annotations are read in order of functional, visual, historical, and social, as functional information is something that a user can act on immediately, while the other categories are lower priority. Not all annotations are usually present on a point of interest, and empty annotations are not announced.

While listening, users may press the button again to pause or restart the annotation. While the earcon is located in 3D space, the annotations were played using 2D audio to make the annotations easier to understand. Our Video Figure demonstrates the use of the FootNotes system.

4 EVALUATION

We conducted a study to determine how users would interact with and react to annotations while exploring a new area. This study focused on the participants' perception of annotation usefulness and their feedback on the usability of FootNotes.

4.1 Participant Demographics

We recruited ten legally blind participants (Table 1) by advertising on email lists related to vision impairment within our organization and the surrounding metropolitan area.

The walking route and points of interest we selected were from a less-trafficked neighborhood in our region, with which participants were not intimately familiar. Four participants had previously used Microsoft Soundscape, but had not experienced the FootNotes annotations. Participants' ages ranged from 25 to 63 (mean = 43.7), and six were female. Two were completely blind and eight had some light perception or low vision. Two participants used a guide dog and eight were primarily cane users. The participants were encouraged to walk in a way that they felt was comfortable, whether that was with their cane, guide dog, or taking the arm of a sighted member of the research team.

4.2 Annotations

Our research team added annotations to ten points of interest in downtown Kirkland, including public sculptures, a local restaurant, and other structures. Several fountains were included because they are both visually and audibly prominent. The choice of items to annotate reflected our interest in making visual landmarks used by sighted people accessible, as well as giving participants more information on items people with vision impairments would already notice. Not every point of interest had every category of annotation present, in order to simulate the anticipated realistic state of the system. The points of interest and their corresponding annotation types are summarized in Table 2.

The average length of all annotations used was 31.6 words, but the time to speak each annotation varied by participant, as each participant adjusted the speaking speed to their level of comfort. On average the functional annotations were the shortest (26.4 words), followed by Social (29.8 words), Visual (33.4 words), and Historical (36.6 words). The full set of annotations used in the study can be viewed in Appendix A.

Table 1. Participant demographics.

ID	Age	Gender	Level of Vision	Primary Aid(s)	Audio Navigation Apps Used
1	57	F	Legally blind, can see some colors and objects within a few feet	White cane	Blindsquare, Ariadne, Google Maps
2	51	F	Legally blind, some light perception (since birth)	White cane, guide dog	
3	37	M	Legally blind with some limited perception	White cane	Nearby Explorer, Google Maps
4	63	F	Legally blind, some light perception (for last 50 years)	White cane	Google Maps
5	25	M	Light perception in one eye (since birth)	White cane	Blindsquare, Seeing Eye GPS, Google Maps, Refreshabraille Display
6	34	F	Legally blind, some central vision in one eye (since birth)	White cane	Apple Maps, Google Maps
7	45	F	Total blindness from birth	White cane	Aira, Blindsquare, BrailleNote GPS, Trekker GPS, Google Maps
8	41	F	Low usable vision, dramatic decline over last two years	White cane	Apple Siri, Google Maps
9	49	M	Light perception, blind for many years	Guide dog	Google Maps, Apple Maps
10	35	M	Total blindness from birth	White cane	Google Maps, Blindsquare

4.3 Route Description

We planned a route in downtown Kirkland that connected these ten annotated points of interest and took participants through commercial and retail areas as well as a local park. It took about one hour to walk the route, including the time needed for a tutorial explaining the FootNotes system (ten minutes) and the time to listen to and give feedback on the annotations when exploring each point of interest. A few participants had been in the park containing some of our points of interest before, but none were intimately familiar with more than one of the points of interest we chose.

Participants travelled with their chosen navigation aids, and members of the research team walked with them to ensure that there were no safety issues during the study. Two researchers were always present; one took field notes while the second provided instructions and monitored safety while walking. The researchers answered any questions the participants had about the operation of the app, but participants manipulated the app themselves using the handheld remote. As the participants walked, they heard the normal interaction of the Microsoft Soundscape system, which called out nearby points of interest with spatial audio. This included the name of the point of interest, as well as the distance to that point of interest.

When the participant reached an annotated point of interest (Table 2), they were asked to stop and listen to any available annotations. Each participant followed the same route, although some participants did not examine every point of interest due to time constraints. P1 was the only participant to listen to the Statue 2 and Restrooms, while P7 and P8 both skipped Statue 1 and Fountain 3. After hearing all annotations present at a point of interest (annotations could be repeated if necessary), a member of the research team asked the following questions:

(1) For each annotation type present (functional, visual, historical, and social), two questions were repeated:

Table 2. Annotations present at each point of interest.

Landmark	Functional	Visual	Historical	Social
Fountain 1		x	x	
Fountain 2	x	x	x	
Restaurant	x	x		x
Statue 1		x	x	
Kiosk				x
Statue 2		x	x	
Restrooms	x			
Fountain 3		x	x	x
Statue 3	x	x	x	x
Pavilion	x	x		x
View	x	x		
Monument		x	x	

(a) On a scale of 1 to 5 where 1 is not at all valuable and 5 is extremely valuable, how would you rate this description?
(b) Would you say that the length of the information you heard was too short, just right, or too long?
(2) Is there other information you wish FootNotes made available right now?
(3) If you had known this was here, and did not have FootNotes, would you have sought out any information about this point of interest? If so, how?

The first questions elicited feedback on the usefulness and conciseness of different annotation types at different classes of points of interest (e.g., restaurants, sculptures). The second question sought omitted information that FootNotes should include in annotations. The third question was designed to collect information on participants' current sources of information, and how those compare to FootNotes. Participants were also encouraged to think aloud during the entire experience and surface any general feedback on the FootNotes annotations, Microsoft Soundscape, or navigation in general. The participants typically did this while stopped at a point of interest, but could also easily silence the FootNotes prototype with the push of a button if it interrupted conversation while walking. The participants could repeat any annotation with a button press if they missed a portion due to environmental noise.

After visiting all of the points of interest, we also administered a closing questionnaire, in which we asked participants to discuss their favorite and least favorite aspects of FootNotes and of each annotation type, as well as to rank order the annotation types according to preference. Finally, we asked about contexts in which hearing FootNotes annotations would be the most beneficial, and when participants envisioned they would contribute their own annotations, if at all.

5 RESULTS

Here, we present the results from participant's feedback on their experience navigating the city with FootNotes. Overall, participants found functional and visual annotations to be the most valuable and rated them at just the right length or slightly too short because they lacked some information. Historical annotations were viewed as too long, but valuable in leisure contexts, while social annotations held less value for most participants.

5.1 Categories of Annotations

Based on the responses at each point of interest, participants rated visual, historical, and functional annotations equally (mean = 4.1, see Table 3). The breakdown of Likert data (Figure 4) shows a similar trend, with visual information receiving the most ratings of 4 and 5 combined. The only noticeably different category is social annotations, which ranked lower (mean = 3.6), though this difference was not statistically significant.

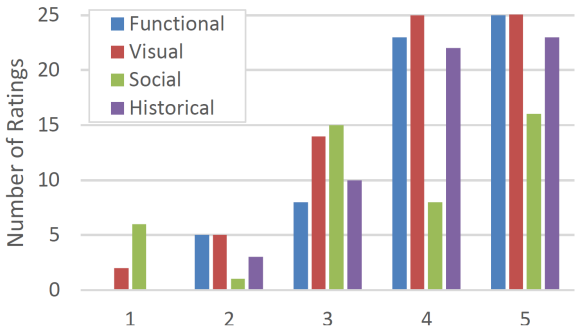

Fig. 4. Participants' ratings for how valuable an annotation is, aggregated by category, where 1 represents "Not valuable at all" and 5 represents "Very valuable".

Participants gave different responses when asking about the categories in aggregate during the debrief questionnaire. Functional and visual annotations tied for overall valuable-ness (mean = 4.5) with historical (mean = 3.9) and social (mean = 3) annotations lagging behind. We also asked the participants to explicitly rank each category of annotations (with 1 being their favorite and 4 their least favorite), which resulted in a mean ranking of functional (mean = 1.2), visual (mean = 1.9), historical (mean = 3.4), and social (mean = 3.5) annotations. We analyzed these scores using a Friedman test of mean ranks. The test was significant, $\chi^2(3, N = 10) = 23.2$, $p << 0.001$. Kendall's $W = 0.772$, indicating strong differences across the four categories. Follow-up pairwise comparisons using a Wilcoxon signed ranks test showed significant differences between Functional and Historical ($z = 2.88$, $p < 0.004$), Functional and Social ($z = 2.85$, $p < 0.004$), Visual and Historical ($z = 2.67$, $p < 0.004$), and Visual and Social ($z = 2.27$, $p < 0.008$). Each participant rated 5-10 annotations for each category (see Table 2), so each category had a number of samples for the above analysis. However, we would caution against reading too much into these quantitative results due to the low number of overall participants; rather, these quantitative results should be contextualized with the open feedback responses to better understand the differences between categories.

Table 3. Aggregate statistics for the ratings given by participants at each point of interest. The ratings were given on a 5-point Likert scale, with 1 being "Not valuable at all" and 5 being "Extremely valuable".

Category	Mean	Median	Std. Dev.
Functional	4.1	4	0.93
Visual	4.1	4	1.03
Social	3.6	4	1.33
Historical	4.1	4	0.88

5.1.1 Functional Annotations. In open feedback, participants especially favored functional annotations, and were more likely to rate that category lower when not enough functional details were included. For example, P2 noted that it would always be helpful to include information about the location of the entrance when announcing a building, and that she would find it helpful if the system always mentioned the location of trashcans when in a park.

"Functional, to me, is anywhere there is a business, or there is anything I can use, engage with, any of those kind of things. [...]" — P6

5.1.2 Visual Annotations. When giving feedback about visual annotations, participants often highlighted aspects of the annotations that they could use, and suggested that many aspects of visual information were functional in nature. For example, P2 felt the visual information would help her feel more comfortable when going out to new places:

"Whenever you are blind and you go out [...] to a new situation, it's very daunting. Quite often I don't find it worth the energy expenditure, for me. [...] , but, if I knew that I had a guide like this, it would make it so much less scary to have the visual information: like where the water is, where the streets are. It's exciting stuff." — P2

Other participants were excited about using visual information to interact with their sighted friends:

"That's not normally information that I would have. [...] If I was talking with someone on the phone and said 'Oh meet me here, I'm by the flagpole, by the place with the fence and the hedge and the umbrellas.' " — P1

"It's less about me [...], but I can turn to my son and say 'Hey, can you see Mt. Rainier? Can you see the Seattle skyline?' and point that out." — P6

5.1.3 Historical Annotations. Historical annotations received more critical feedback, as many thought this category was superfluous or unnecessary for daily life. However, some participants liked the possibility of using trivia from these annotations as conversation topics with their friends.

"The history is nice for idle chatter. It just makes you more a part of the visual community. You know, they can just see it on a plaque." — P4

"I like the historical stuff because it gives you enough information that you have some knowledge [to find out more on your own]." -P1

5.1.4 Social Annotations. Social annotations received the lowest ratings, and many participants thought they added little value beyond their existing social networks. Many were open to the idea of integrating existing social networking content, but these responses were still tepid.

"The social stuff I am just not as wild about. [...] I like to be engaged with my friends on Facebook. [...], but I feel like I just don't want another [social network]." — P6

5.2 Length of Annotations

To examine the responses at each point of interest across categories, the responses "Too short", "Just right", and "Too long" were mapped to the values of -1, 0, and 1 respectively. Functional (mean = -0.02) and visual (mean = 0.06) annotations were rated closest to "Just right" on average, while historical (mean = 0.18) and social (mean = 0.17) were more likely to be rated "Too long" (Table 4).

There was a very weak negative correlation between the participants' ratings for an annotation's utility and its length (r_s=-0.15, p = .01), because participants tended to identify annotation categories they found less useful as "too long" and tended to want additional detail for the functional and visual categories. However, participants generally rated the annotation length as "just right" (with an overall median rating = 0, see Table 4), indicating that they didn't find the idea of spending time listening to annotations burdensome. Nonetheless, seven participants mentioned they valued annotations' being concise.

5.3 Omitted Information

At each point of interest, we asked participants if there was additional information that they would like at this type of location. These answers (Table 5) can help define the boundaries of useful information and annotation content but should not be viewed as a definitive list.

The most common request (eight participants) was for specific information on how to reach or enter a point of interest. Participants knew when they were near and/or facing a point of interest but did not have enough information to confidently use it. This group of requests includes the location of a door for a business, the

Table 4. Aggregate statistics for the participants' rating of the length of each annotation at a point of interest on a scale from -1 (Too short) to 1 (Too long).

Category	Mean	Median	Std. Dev.
Functional	-0.02	0	0.50
Visual	0.06	0	0.49
Social	0.17	0	0.38
Historical	0.18	0	0.43

Table 5. Requested Information By Number of Participants (Response to Q3)

Information	Number of Participants	Information	Number of Participants
Business Details/Hours/Prices	9	User Reviews	2
"How do I get to the POI?"	8	Area Dimensions	2
"Where can I sit?"	5	Description of Boundary	1
Object Dimensions	4	General Mobility Info	1
Accessibility Question	4	Color	1
Complete Text of Signs	2	Lighting	1
Website Link	2	Visual Question	1

orientation and location of a bench, and paths through a park to reach a specific statue. Microsoft Soundscape (and FootNotes) does not attempt to provide turn-by-turn route information, as the GPS location of the user is not reliable enough to do so with accuracy. This problem, referred to as the "last 10 meters/yards" problem, is a well-known limitation of using GPS for blind navigation aids [14, 26]. FootNotes may need to include specific human-written instructions on how to access points of interest when sufficiently near.

The second most common request (seven participants) was for specific information about a point of interest, especially a business. This included price information, hours, and information on upcoming events. Much of this information was structured (e.g., hours, price range) and could be found on Yelp or Google, although some specific questions would be best answered on a website or a FootNotes annotation written by the business (e.g., "What are the blackout dates for renting this pavilion?"). Two participants suggested linking directly to a website in FootNotes annotations, so they could find that information themselves.

In addition to the previously mentioned information on reaching a specific point-of-interest, our participants had additional spatial questions that were unanswered in our annotations. Five participants wanted to know what sort of seating was nearby. This included tables and benches in the park, and seating available at businesses. Two participants tried to gauge business seating by the size of the area, so descriptions of the area's layout may be useful. Four participants also asked about the dimensions of objects (e.g.,"How tall is this fountain?"), as we did not include that in our FootNotes annotations.

Finally, four participants had accessibility-specific questions or requests for information that FootNotes did not include. Some requests were location-specific, such as whether payment terminals were accessible with audio input. These participants also mentioned hazard information, and specifically wanted to be alerted to any drop-offs into fountains or off the shore.

It is difficult to assess the value of responses to this question, as our participants did not necessarily know what information could have been included in the FootNotes annotations. One participant pushed back against this question, noting:

"I am never going to know enough to know what I am missing." — P7

Information relating to accessibility and functional use (e.g., how to reach an entrance) had a clear demand from our participants, and future FootNotes annotations should strive to include these topics. Less requested information like dimensions of objects may be useful to some but including them in general FootNotes annotations is likely inefficient.

5.4 Comparisons to Status Quo

After asking about omitted information for a point-of-interest, we asked participants how they would have found information using only their current tools (Table 6). We hoped this would provide insight into any improvement FootNotes annotations offered over existing services. The wording of the question let participants assume they already knew an object existed, but many participants noted that they would not have been aware of the point-of-interest with any current tools. For example, descriptions of statues can often be looked up on Wikipedia or local sites, but knowing that a statue is in the park in the first place is difficult for people with vision impairments. Many participants even pointed out that knowing an object exists does not mean they know the name of an object in order to search about it online, as many plaques or signs are not accessible.

Table 6. Other methods to retrieve similar information by number of participants (Response to Q4).

Category	**Number of Participants**	**Examples**
General Search Engine	6	Google, Bing, Siri. Wikipedia
Specific Application	5	Yelp, Foursquare, Swarm, Google Maps
Friend or Family	3	Friend, Partner, Child, Niece
Business Contact	2	Website, Call Business
Social Media	1	Facebook
Other People	1	Ask Strangers

Five participants usually said they would use a search engine like Google or general site like Wikipedia to find out information if they knew the name of a point-of-interest. They might also turn to specific applications such as Yelp, Foursquare, or Swarm to find reviews or specific structured details such as business hours. Three participants said they would ask a sighted family member for more information. One participant mentioned that she would call the business with questions, ask a stranger, or post to a neighborhood Facebook group for more information.

5.5 Usefulness of Annotations in Different Contexts

Many participants commented that they would enjoy listening to all annotations in certain contexts but were more selective when thinking about perceived daily use. Many (six) participants wanted to use FootNotes when exploring a new city, or visiting a new neighborhood in their own town. Several participants said they would not use the system as much after they were more familiar with an area.

Participants sometimes found the historical annotations interesting, but six said explicitly they would only use them if they were on a leisurely walk. They did not have time to listen to them as they were busy getting to a specific destination. Several suggested adding a setting to hide annotation types, such as historical annotations, when they were not interested in them. These could be toggle-able, as in some contexts it might make more sense (e.g., exploring a new city on vacation). For the annotation types left on, participants wanted a way to quickly cycle through them.

In addition to annotation categories, some participants noted that certain points of interest were just not something they would like to explore, and they asked for mechanisms to filter the objects they heard. One participant requested that all "art" be filtered from FootNotes, but noted that he still wanted all fountains to be included. The filtering mechanism will likely need to be more granular than just a few broad classifications.

5.6 Creating and Editing Annotations

The study we conducted did not include usability evaluations for users to add annotations on the spot. However, in our interview debrief, we did ask participants if they envisioned themselves authoring annotations themselves. Nine participants said they would author annotations themselves, and seven would be comfortable sharing these annotations with other users.

Some participants voiced concerns about modifying the content of others' annotations, and it is unclear whether users would want to edit content while on the go, beyond adding additional or new information.

6 DISCUSSION

When designing FootNotes, we imagined the tool being most useful for exploration and discovery of visual objects that may otherwise go unnoticed, revealing the "local color" of an area. While it did serve this purpose (as P1 noted, the system described "a lot of things I probably walk by that I don't really know are there"), our participants saw most of the value in its use as a functional tool, demonstrated by their preference for functional and visual annotations. Specifically, they asked for FootNotes to describe layout information, directions to important points of interest (like the information desk in a hotel lobby), and nearby paths. The historical and social annotations received less enthusiastic response from participants, as these were only useful in specific niche contexts, like on a leisurely walk or when traveling.

In our design, we constrained the information supported by FootNotes to those that fit into categories of functional, visual, historical, and social annotations, but there may be other types of annotations that also make sense. When we asked participants about categories of information that FootNotes did not cover, they mentioned accessibility information such as building layouts, hazards like drop-offs or ledges, or visually-distinct objects that it would be easy for low-vision people to find. Portions of the functional information that we provided, as well as large amounts of information the participants requested (Section 5.3), could be found as structured data in other applications like Yelp or Google Maps (Section 5.4). Based on this, FootNotes might benefit from separating some of our categories into more fine-grained categories, such as splitting functional annotations into accessibility, business, and layout annotations. Similarly, social annotations might be split into two categories, such as public social annotations from other users of the FootNotes system versus private annotations from a user's personal social network or from their O&M instructor.

Dense areas with many labelled points of interest, each with several categories of annotations, would be overwhelming for FootNotes users without adequate controls to manage the surrounding information. Participants in our study often asked for an interface to quickly jump through or directly access different annotation types, and this would be even more important if more types were added. Given that users may interact with many annotations during a session, mapping different remote control buttons, gestures, or even voice commands to annotation categories as shortcuts would be beneficial. Users would likely want to customize these mappings and/or customize the order in which categories of annotations are read aloud.

Based on participant think-aloud feedback in the study, future iterations of FootNotes would also benefit from two filter interfaces: a point of interest filter and a category filter. The POI filter would allow users to ignore or highlight objects they particularly care about, such as trash cans or art. A user could toggle different annotation types using a category filter, depending on what sort of information they are interested in. These filters could be implemented as user-modifiable preferences, or be learned from user ratings of individual annotations.

Supporting user ratings of annotations is another important extension. This might not only support learning a model for preferred annotation categories for users, but could also help with giving quality feedback for filtering crowdsourced annotation content.

Another potential FootNotes filter to explore is to constrain the authors that users see annotations from. For many annotations, such as functional descriptions of a business layout, it may not matter who wrote or edited an annotation, so long as it is accurate. For annotations that are more opinionated, such as a visual annotation of a work of art, participants could utilize filters to only hear annotations from trusted sources. In some cases, this could be a friend whose opinion they value, or just an online contributor who is known for writing compelling descriptions.

The participants in our study were most excited about using FootNotes in unfamiliar environments, and a few pointed out that they would like to access this information before navigating to learn the area. Virtual navigation has been shown to help users with vision impairments form accurate mental maps of spaces before navigating them [18]. Therefore, FootNotes may also be a tool that Orientation and Mobility instructors could use to teach new routes. O&M instructors might construct a set of curated annotations that the user could access even when alone, or place "virtual" annotations that do not correspond to a physical point of interest, but are reminders for the user to look for a tactile marker. By giving the FootNotes user an artifact to use on their own, O&M instructors can spend more time on teaching techniques, instead of teaching new routes.

FootNotes was designed to use text annotations due to the medium's universal accessibility, but other media like audio could be supported. Participants could then explore a virtual route while listening to realistic sounds of traffic or fountains to learn audio landmarks, in addition to the normal FootNotes annotations. This same interface could also be used to help users explore purely virtual spaces, as current virtual reality efforts are primarily vision-based, and most are not accessible to someone with a vision impairment. Navigation tools have been designed for people with visual impairments using virtual reality, but they focus on providing haptic feedback [35]. FootNotes could make some games and virtual experiences usable without sight by providing both audio descriptions of ongoing events and descriptions of virtual objects relative to the player's orientation.

6.1 Limitations

There are several limitations to our work. Our system was not yet polished and robust enough for our participants to use it daily outside of our study, so we were unable to collect longitudinal data that may have shown long term usage trends. Because of this, we were also unable to observe the authoring behaviors our participants might exhibit in realistic usage scenarios. Due to these issues of robustness and safety concerns, we also only ran our study in an area we felt was safe while still being a representative use case. Other areas where points of interest might be sparser or denser could reveal new issues or opportunities for our system.

Our evaluation may have limited participants' ability to fully appreciate the social annotations because they were artificial messages from the research team. FootNotes is only a research prototype, and thus participants did not have social contacts within the system who could leave them real messages. Longer-term deployments with a critical mass of users may be necessary to fully understand the potential of geo-referenced social annotations for this audience. We also did not explore the implications of out-of-date information in our evaluation, such as annotations of a large farmer's market that takes over the park weekly in the summer. We believe FootNotes could design annotations to only be active for certain events, or to expire after a certain end date (e.g., after construction is complete), but long-term deployments would be needed to explore if temporal annotations would be usable.

The annotations that were used in this study were written by the research team with some local knowledge, and the actual annotations written by online crowd workers or local volunteers might differ in content or length. If a community of regular FootNotes users existed to give feedback on created annotations, then the authoring

behavior of volunteers could be examined in depth. A study of the creation and moderation of annotations in a local area would draw on the prior work from the Cyclopath project [30].

6.2 Future Work

Based on the limitations of our current system and unexplored use cases, we believe there are three main research questions for future work surrounding FootNotes:

(1) What behavioral patterns do users exhibit based on long-term usage of a spatialized audio system with rich annotations? Do usage patterns mirror the stated preferences from this evaluation, and if not, which contexts affect their usage of differing annotation types?
(2) How can we recruit and support online editors with local knowledge to author and moderate annotations for FootNotes? How can we ensure these annotations will be of high quality and maintained over time to combat stale information?
(3) In what ways can FootNotes augment the guidance provided by other Blind Navigation systems to make independent travel easier for people with vision impairments? Can FootNotes annotations in a virtual world be used as a pre-planning tool before navigation?

Investigating these and related questions in-depth likely requires longitudinal studies of both a population of FootNotes users and annotation authors, as the usage patterns observed in controlled studies may not reflect real daily use. The previous listed limitations of our evaluation would be mitigated in such a longitudinal study of users, and we could thoroughly explore the realistic usability of FootNotes in the field. The application would need to be very robust for such a study, as abandonment of assistive technology is high, and small usability issues can hinder adoption [24]. A longitudinal study on annotation authors would provide further insight into the annotation categories, ease of collecting information requested by users, and overall ecosystem of annotation management. It would also provide ample opportunity to experiment with structured prompts to ensure annotations are concise, thorough, and high-quality. For example, guidelines exist to help non-expert authors create location-based content similar to historical annotations [15].

7 CONCLUSION

From our experience with participants using FootNotes, we believe that having functional and visual annotations could enrich the navigation experience for people with vision impairments. FootNotes provides a level of local color that is otherwise inaccessible, and is especially useful in unfamiliar environments. Other annotation types, such as historical and social annotations, are useful to understand and feel a sense of inclusion with sighted people, and may be beneficial in specialized applications or modes. Existing and future location-based applications for people with vision impairments should also add FootNotes-style annotations to enrich the experience of exploring and make visual information in the physical world more accessible.

ACKNOWLEDGMENTS

The authors would like to thank the Microsoft Research NExT Enable Team for their help and support in completing this project, as it would have been impossible without them!

APPENDIX

A ANNOTATIONS USED IN FOOTNOTES STUDY

The following annotations were used in the evaluation with blind participants (see Table 2).

Visual: "Two large planks of bronze, each curved slightly, lean up against each other. Water runs down each of these in a thin layer. Flowers surround the base of the fountain, and a bench encloses it."
Historical: "The fountain was commissioned to honor the memory of Julius Mcleod, an active community member. It was designed by Seattle artist Kate Martin, and the landscaping was created with leftover pavers from a park renovation."

Functional: "Step hazard: the sidewalk drops off to the open pool"
Visual: "A black boulder is in a pool of water. Behind it, two hollow shells of smaller reddish boulders are tilted to allow water to flow out of them. More boulders meant as seats surround the pool."
Historical: "The fountain was designed by John Hoge, who specializes in stone fountain and benches in the Puget Sound area."

Functional: "Hector's beer garden is a partnership between the restaurant and Pike brewing. It is open at 3pm from Monday to Thursday, and noon on Friday through Sunday. It is dog friendly."
Visual: "Hector's has a large beer garden with more than a dozen tables and red umbrellas. The beer garden is enclosed by a wooden picket fence and bushes. The bushes have colorful umbrellas resting on top of them, many with dog patterns."
Social: "John says: If you're looking for a birthday gift idea for John, he hinted that he'd love a gift card to Hector's!"
Social: "Carl says: Dogs, dogs, dogs! On sunny afternoons, I think there are more dogs here than at the off-leash area at the park!"

Visual: "This waist high brick pillar has a small concrete statue of Santa Claus on top. He holds a teddy bear and a sack of gifts. The plaque dedicates this work to The Children."
Historical: "The statue is a gift from Dennis Brown, a local sculptor who specializes in Santa Claus figurines. His work has been featured at the White House."

Functional: "Argosy cruises offers a one and a half hour narrated cruise of Lake Washington from the City Dock. It departs on Saturdays and Sundays at 11:30am, 1:30pm, and 3:30pm. The ticket booth open 30 minutes prior to departure, and only accepts payment by credit cards."
Social: "Aaron says: We should take a cruise around the lake one weekend before it gets chilly again!"

Visual: "The bronze statue depicts a World War II sailor reuniting with his wife and child on the docks after returning from sea. A plaque in the background read 'To those families that also serve. Presented by the U.S. Navy Memorial foundation' "
Historical: "This is a replica of an original statue by Stanley Bleifeld at the U.S. Navy memorial in Washington D.C. The replica was dedicated here in 2001 as a recognition for the city's role in shipbuilding and maintenance during World War II."

Functional: "The men's restroom is on the side of the building facing the docks, while the women's restroom faces the park. A water fountain is to the left of the entrance to the women's restroom.'

Visual: "The fountain is made up of a concrete base to catch the falling water and a bronze series of leaf-like outcroppings that form a central pillar. The water flows from the top of the pillar, jumping down each outcropping to the base below."
Historical: "The Centennial Fountain was created by James Hebert Fitzgerald and donated to the City of Kirkland by the Rotary club in 1972, making it one of the city's oldest public art pieces. It commemorates the first settlement in Moss Bay in 1872."
Social: "Mindy says: Remember that time when my dog, Rosco, fell into this fountain while he was drinking from it? That was pretty funny."

Functional: "Kids often climb and play around this artwork."
Visual: "This bronze artwork depicts six young children holding hands while skipping and jumping over puddles. Their faces show joy, with some seeing to yell or shout with excitement."
Historical: "Resident and art collector Bill Ballantine loaned the sculpture to the city in 1990, and ten years later a grassroots effort organized to raise $250,000 to buy it. The artist, Glenna Goodacre, is well known for designing the front of the Sacagawea golden dollar coin."
Social: "Aaron says: I've always found this art to be really creepy. It's like a bunch of children zombies racing down the hill to feast on living kids at the water."

Functional: "This gazebo is available for rent on almost all days except for some holidays. It costs $40 per hour and must be booked 30 days in advance."
Visual: "A large open-air gazebo that can fit around 100 people."
Social: "Mindy says: What do you think about renting this for the annual summer picnic?"
Social: "Aaron says: There's an awesome saxophone busker who sets up here in the spring and fall. Really good jazz player!"

Visual: "From here you can look out over the Marina and see the City Dock. The masts of sailboats stick up in the air, and the large Argosy cruising ferries can often be seen docked here."
Functional: "A bench is nicely positioned here to enjoy the view, but past the sidewalk the shore quickly drops off to water."

Visual: "Located in the parking lot, this circular stone enclosure surrounds a flagpole in the center. Plaques are located on the stone, and almost half of the circle is empty with room for new inductees."
Historical: "The Plaza of Champions was initiated in 1988 to honor and recognize groups and individuals in the area who have reached the pinnacle of achievement in their field and contributed to the community. Inducted in 2015: Billy and Cory Roeseler, Inventor and Pioneers in the Sport of Kiteboarding."

REFERENCES

[1] Aira. 2017. Aira | Instant Access to Information. https://aira.io

[2] Frank Bentley and Santosh Basapur. 2012. StoryPlace.Me: The Path from Studying Elder Communication to a Public Location-based Video Service. In *CHI '12 Extended Abstracts on Human Factors in Computing Systems (CHI EA '12)*. ACM, New York, NY, USA, 777–792. https://doi.org/10.1145/2212776.2212851

[3] Jeffrey P. Bigham, Chandrika Jayant, Hanjie Ji, Greg Little, Andrew Miller, Robert C. Miller, Aubrey Tatarowicz, Brandyn White, Samuel White, and Tom Yeh. 2010. VizWiz: Nearly Real-time Answers to Visual Questions. In *Proceedings of the 2010 International Cross Disciplinary Conference on Web Accessibility (W4A) (W4A '10)*. ACM, New York, NY, USA, Article 24, 2 pages. https://doi.org/10.1145/1805986.1806020

[4] Jeffrey P. Bigham, Ryan S. Kaminsky, Richard E. Ladner, Oscar M. Danielsson, and Gordon L. Hempton. 2006. WebInSight:: Making Web Images Accessible. In *Proceedings of the 8th International ACM SIGACCESS Conference on Computers and Accessibility (Assets '06)*. ACM, New York, NY, USA, 181–188. https://doi.org/10.1145/1168987.1169018

[5] Erin Brady, Meredith Ringel Morris, and Jeffrey P. Bigham. 2015. Gauging Receptiveness to Social Microvolunteering. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 1055–1064. https://doi.org/10.1145/2702123.2702329

[6] Erin L. Brady, Yu Zhong, Meredith Ringel Morris, and Jeffrey P. Bigham. 2013. Investigating the Appropriateness of Social Network Question Asking As a Resource for Blind Users. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work (CSCW '13)*. ACM, New York, NY, USA, 1225–1236. https://doi.org/10.1145/2441776.2441915

[7] OpenStreetMap Wiki contributors. 2017. Changeset. https://wiki.openstreetmap.org/wiki/Changeset

[8] OpenStreetMap Wiki contributors. 2017. Changeset. https://wiki.openstreetmap.org/wiki/Changeset

[9] OpenStreetMap Wiki contributors. 2017. Elements. https://wiki.openstreetmap.org/wiki/Elements

[10] OpenStreetMap Wiki contributors. 2017. Key:oneway. https://wiki.openstreetmap.org/wiki/Key:oneway

[11] Daniel Dardailler. 1997. The ALT-server (" An eye for an alt").

[12] Navid Fallah, Ilias Apostolopoulos, Kostas Bekris, and Eelke Folmer. 2012. The User As a Sensor: Navigating Users with Visual Impairments in Indoor Spaces Using Tactile Landmarks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, New York, NY, USA, 425–432. https://doi.org/10.1145/2207676.2207735

[13] Navid Fallah, Ilias Apostolopoulos, Kostas Bekris, and Eelke Folmer. 2013. Indoor Human Navigation Systems: A Survey. *Interacting with Computers ()* 25, 1 (2013), 21–33.

[14] Alexander Fiannaca, Ilias Apostolopoulous, and Eelke Folmer. 2014. Headlock: A Wearable Navigation Aid That Helps Blind Cane Users Traverse Large Open Spaces. In *Proceedings of the 16th International ACM SIGACCESS Conference on Computers & Accessibility (ASSETS '14)*. ACM, New York, NY, USA, 19–26. https://doi.org/10.1145/2661334.2661453

[15] Elizabeth FitzGerald. 2012. Creating user-generated content for location-based learning: an authoring framework. *Journal of Computer Assisted Learning* 28, 3 (2012), 195–207.

[16] REGINALD G. GOLLEDGE, ROBERTA L. KLATZKY, JACK M. LOOMIS, JON SPEIGLE, and JEROME TIETZ. 1998. A geographical information system for a GPS based personal guidance system. *International Journal of Geographical Information Science* 12, 7 (1998), 727–749. https://doi.org/10.1080/136588198241635 arXiv:https://doi.org/10.1080/136588198241635

[17] W. G. Griswold, P. Shanahan, S. W. Brown, R. Boyer, M. Ratto, R. B. Shapiro, and T. M. Truong. 2004. ActiveCampus: experiments in community-oriented ubiquitous computing. *Computer* 37, 10 (Oct 2004), 73–81. https://doi.org/10.1109/MC.2004.149

[18] João Guerreiro, Dragan Ahmetovic, Kris M. Kitani, and Chieko Asakawa. 2017. Virtual Navigation for Blind People: Building Sequential Representations of the Real-World. In *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '17)*. ACM, New York, NY, USA, 280–289. https://doi.org/10.1145/3132525.3132545

[19] Mordechai Haklay. 2010. How Good is Volunteered Geographical Information? A Comparative Study of OpenStreetMap and Ordnance Survey Datasets. *Environment and Planning B: Planning and Design* 37, 4 (2010), 682–703. https://doi.org/10.1068/b35097 arXiv:https://doi.org/10.1068/b35097

[20] Kotaro Hara, Shiri Azenkot, Megan Campbell, Cynthia L. Bennett, Vicki Le, Sean Pannella, Robert Moore, Kelly Minckler, Rochelle H. Ng, and Jon E. Froehlich. 2013. Improving Public Transit Accessibility for Blind Riders by Crowdsourcing Bus Stop Landmark Locations with Google Street View. In *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '13)*. ACM, New York, NY, USA, Article 16, 8 pages. https://doi.org/10.1145/2513383.2513448

[21] Kotaro Hara, Vicki Le, and Jon Froehlich. 2013. Combining Crowdsourcing and Google Street View to Identify Street-level Accessibility Problems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 631–640. https://doi.org/10.1145/2470654.2470744

[22] Bill Holton. 2014. BlindSquare App for iOS: A Solution for Accessible Navigation. Issue 7. https://www.afb.org/afbpress/pubnew.asp?DocID=aw150704

[23] Hernisa Kacorri, Kris M. Kitani, Jeffrey P. Bigham, and Chieko Asakawa. 2017. People with Visual Impairment Training Personal Object Recognizers: Feasibility and Challenges. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 5839–5849. https://doi.org/10.1145/3025453.3025899

[24] Shaun K. Kane, Chandrika Jayant, Jacob O. Wobbrock, and Richard E. Ladner. 2009. 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 (Assets '09)*. ACM, New York, NY, USA, 115–122. https://doi.org/10.1145/1639642.1639663

[25] Aniket Kittur and Robert E. Kraut. 2008. Harnessing the Wisdom of Crowds in Wikipedia: Quality Through Coordination. In *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work (CSCW '08)*. ACM, New York, NY, USA, 37–46. https://doi.org/10.1145/1460563.1460572

[26] Jacobus Lock, Grzegorz Cielniak, and Nicola Bellotto. 2017. Portable navigations system with adaptive multimodal interface for the blind. In *The 2017 AAAI Spring Symposium Series*. AAAI, The AAAI Press, Palo Alto, CA, USA, 395–400.

[27] Microsoft. 2017. Seeing AI | Talking camera app for those with a visual impairment. https://www.microsoft.com/en-us/seeing-ai/

[28] Per Persson, Fredrik Espinoza, Petra Fagerberg, Anna Sandin, and Rickard Cöster. 2003. GeoNotes: A Location-Based Information System for Public Spaces. In *Designing Information Spaces: The Social Navigation Approach*, Kristina Höök, David Benyon, and Alan J. Munro (Eds.). Springer London, London, 151–173. https://doi.org/10.1007/978-1-4471-0035-5_7

[29] Reid Priedhorsky, Mikhil Masli, and Loren Terveen. 2010. Eliciting and Focusing Geographic Volunteer Work. In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work (CSCW '10)*. ACM, New York, NY, USA, 61–70. https://doi.org/10.1145/1718918.1718931

[30] Reid Priedhorsky and Loren Terveen. 2008. The Computational Geowiki: What, Why, and How. In *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work (CSCW '08)*. ACM, New York, NY, USA, 267–276. https://doi.org/10.1145/1460563.1460606

[31] Microsoft Research. 2018. Microsoft Soundscape. https://www.microsoft.com/en-us/research/product/soundscape/

[32] Hironobu Takagi, Shinya Kawanaka, Masatomo Kobayashi, Takashi Itoh, and Chieko Asakawa. 2008. Social Accessibility: Achieving Accessibility Through Collaborative Metadata Authoring. In *Proceedings of the 10th International ACM SIGACCESS Conference on Computers and Accessibility (Assets '08)*. ACM, New York, NY, USA, 193–200. https://doi.org/10.1145/1414471.1414507

[33] Fernando Torre, S. Andrew Sheppard, Reid Priedhorsky, and Loren Terveen. 2010. Bumpy, Caution with Merging: An Exploration of Tagging in a Geowiki. In *Proceedings of the 16th ACM International Conference on Supporting Group Work (GROUP '10)*. ACM, New York, NY, USA, 155–164. https://doi.org/10.1145/1880071.1880097

[34] Luis von Ahn, Shiry Ginosar, Mihir Kedia, Ruoran Liu, and Manuel Blum. 2006. Improving Accessibility of the Web with a Computer Game. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '06)*. ACM, New York, NY, USA, 79–82. https://doi.org/10.1145/1124772.1124785

[35] Yuhang Zhao, Cynthia L. Bennett, Hrvoje Benko, Edward Cutrell, Christian Holz, Meredith Ringel Morris, and Mike Sinclair. 2018. Enabling People with Visual Impairments to Navigate Virtual Reality with a Haptic and Auditory Cane Simulation. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 116, 14 pages. https://doi.org/10.1145/3173574.3173690

Received February 2018; revised May 2018; accepted September 2018