GILBERT

A Companion for Blind People

Joohyun Lee

GILBERT
A Companion for Blind People

Diploma Thesis, Written Contribution

Joohyun Lee
01640265
Summer Semester 2021

University of Applied Arts Vienna

Institute of Design / Graphic Design
Univ. Prof. Oliver Kartak

Supervisors

Univ. Prof. Oliver Kartak
Mag. Erwin Bauer
Mag. Sabine Dreher
Mag. Katharina Uschan
Mag. Christian Schlager

Writing Support

Mag. Katharina Uschan

INDEX

ABSTRACT

GILBERT is a mobile app for blind people that audibly describes the surroundings. It uses the power of AI, assisted GPS location data, and smartphone camera with infrared laser-based depth sensors that use the time of flight to determine distances of faraway objects such as ToF sensor on Android smartphones and LiDar sensor on iPhones.[1]

I am not a blind person myself. But after a couple of incidents that occurred to me by chance with blind people, it appeared to me that it is very difficult for a blind person to be out on the street when they are alone, especially at a new place. They need someone beside them who describes the surrounding world for them.

GILBERT intends to be this companion who tells blind people about the surrounding world walking along with them whenever and wherever. GILBERT also includes the participation of sighted people to compensate for the imperfection of AI image recognition.

[1] Jon Knight, "LiDar vs. 3D ToF Sensors – How Apple is making AR Better for Smartphones", *Gadget Hacks*, March 31, 2020, https://ios.gadgethacks.com/news/lidar-vs-3d-tof-sensors-apple-is-making-ar-better-for-smartphones-0280778/. Last accessed 7 June, 2021.

INTRODUCTION

Motivation

I don't personally have any connection to visual impairment but from time to time I have encountered somebody with it. Once I had a neighbor living next door. She was always with her white cane and I have barely spotted her being outside alone. Most of the time she was accompanied by a sighted person. When I go to the dog park with my dog, I would sometimes see a guide giant poodle happily hopping around enjoying her playtime. Her owner is also always accompanied by a sighted assistant and her dog often wouldn't come on the call. Another time a blind person was looking for a metro station completely lost on the street. She was repeating "Could you help me?" in the air hoping somebody would hear it. I had to take her hand and guide her to the metro station and it was not easy with the other hand holding the leash of my non-assistant dog who has his own way. I managed to bring her to the nearest metro station exit and that was where I was wrong because she needed an exit with elevator access. So I had to re-guide her to another exit which couldn't be found with just one sweeping glance.

Then after a while, I was walking on the street while listening to music with my wireless Bluetooth earbuds plugged in my ear and a rushing ambulance passed by with siren wailing. As the loud sound pierced through my sealed ear canals, a thought also flitted through my mind. Now that recent wireless headphones have a function called ambiance/transparency mode, which makes it possible to hear both sounds from the smartphone and the surrounding simultaneously. Since the auditory sense is directly linked to safety when a blind person is out on the street, using a mobile app that incessantly gives auditive information could be bothersome. It's either hearing from the speaker on the smartphone in one hand with a white cane or the handle on a guide dog in the other hand or constantly plugging headphones in and out. But now it would be less hassle with the ambiance/ transparency mode plus touch control on the headphone with which important functions on a mobile app can be activated and deactivated without the phone having to be out of the pocket.

Background

Phones, invented to let us communicate directly with someone else, are portable, colorful, always and everywhere connectable to the internet, they are getting smarter, faster, and increasingly central to our everyday life.[2] This electronic device has become our pseudo-human life companion. Every single time a smartphone evolves with new technology and new functions, we enter a new dimension of convenience whether foreseen or unexpected. Therefore our everyday life is becoming more and more dependent on smartphones and at the same time, less and less on the physical presence of ourselves or other people.

It might sound negative, cold or heartless, that the physical social contact is less necessary than before, but for some people, it could be a release from limitations, especially for those who unavoidably have to depend on other people's physical presence in order to resolve their daily needs in their everyday and social life due to their disability of human species' most dominant sense, the visual sense.[3]

Smartphones already have guided visually impaired people to be so much more independent than before. The start of it would be when Apple launched iPhone 3GS with an incorporated screen reader called VoiceOver, which gave visually impaired people reasonable access to a touchscreen phone. In the same year, Google quickly followed suit by adding TalkBack to Android. They have replaced so many hundreds of specialist devices with one mainstream, much more affordable, package. A talking GPS device back then was over €800, a talking notetaker over €1000, and so on. Accessibility features on the smartphone have replaced a backpack of devices and their chargers.[4]

With a smartphone visually impaired people can read books, shop online, book train tickets, post on Facebook. Now that AI technology has become commonplace, they can also know what's in a picture; recognize faces, colors and, lights; read menus in a restaurant; even take photos and upload them on Instagram and of course, use a special guide navigation app designed for the visually impaired.

---

[2] *TigerMobiles.com*, Evolution of the Mobile Phone, https://www.tigermobiles.com/evolution/#start. Last accessed 7 June, 2021.

[3] *Max-Planck-Gesellschaft, "*Look first", January 15, 2015, https://www.mpg.de/8849014/hierarchy-senses. Last accessed 7 June, 2021.

[4] Robin Christopherson, "VoiceOver on the iPhone turns 10 – and turns blind access up to 11", *AbilityNet*, June 20, 2019, https://abilitynet.org.uk/news-blogs/voiceover-iphone-turns-10-and-turns-blind-access-11. Last accessed 7 June, 2021.

In October 2020, Apple released iPhone 12 and it's equipped with a LiDar sensor. It is not the very first smartphone with a laser depth sensor (some Android smartphones already had ToF sensors), but Apple has the power of bringing new technology into widespread use. Just like how wireless Bluetooth headphones became so popular all of a sudden after Apple launched AirPods even though they were not the first wireless headphone ever, it is likely that the laser depth sensor on smartphones will affect the mobile app market to a significant extent.

"If technology doesn't consider everyone, it can be a tool for discrimination." [5]

—Junhyo Park, a university student in Seoul

Even though technology has reduced the information gap between the sighted and the visually impaired, still there are loopholes. Still, there is something that you can get only when you have visual sense. In my degree project, I wanted to find one of those loopholes and fill it up by designing a mobile app for people with vision impairment that maximizes the use of the newest smartphone technology in hope that more and more designers will consider "everyone" and there will be less and less discriminative designs.

[5] *My Review Room*, "Netflix to Selfie! How do visually impaired people use iPhone?: iPhone iOS Accessibility", November 17, 2019, https://youtu.be/6N0Rai3GtQ8. Last accessed 7 June, 2021.

RESEARCH

A lot of sighted people including myself who don't necessarily have to do anything with vision impairment, presume that visually impaired people would have so much inconvenience in their everyday life, but this presumption is as a matter of fact exaggerated. Loss of vision leaves in blind people a larger brain capacity for processing information from other senses like the sense of hearing, smell, and touch. In their environments they use non-visual clues like wind direction, the smell of the bakery, the heat of the sun more often than sighted people, even though those clues can be inconsistent and hence less reliable.[6] Some blind people even learn to sense the location and size of objects around them using echolocation. People adapt to new ways of solving everyday tasks after the loss of vision.

During the research phase the identity of GILBERT has changed. At first, I was designing GILBERT as a guide navigation app for visually impaired people. However, after I got an opportunity to talk to blind people personally, I realized that I was trapped in my own point of view as a sighted person regardless of how much research I had made on blindness. I had made a problem on my own out of something that wasn't a big problem for the actual concerning people. I was designing to provide over-convenience completely ignoring an emotional aspect thinking that the priority lies in solving everyday life tasks, which was out of unconscious pity. The perception, assumptions, and pity I have had whether knowingly or unknowingly are the real barrier that I had built up between people with disability and me myself a non-disabled person. That invisible barrier, on the community scale, would contribute directly to the social exclusion of people with disabilities.

It is therefore crucially important to understand that even though disabled people might experience few troubles in everyday life they can do what we do, but just in different ways and to perceive them at the same eye level as if they were people from a different culture.

[6] Stina Olofsson, "Designing interfaces for the visually impaired" master's thesis, Umeå University, 2017, https://www.diva-portal.org/smash/get/diva2:1179505/FULLTEXT01.pdf. Last accessed 7 June, 2021.

Visual Impairment

Visual impairment is a decrease in the ability to see to a certain degree that causes problems not fixable by usual means such as eyeglasses, contact lenses, medical or surgical treatments, including both low vision and blindness. A person with low vision is one who has impairment of visual functioning even after treatment and/or standard refractive correction, and has a visual acuity of less than 6/18 to light perception, or a visual field of less than 10 degree from the point of fixation, but who uses, or is potentially able to use, vision for planning and/or execution of a task.

A blind person is who has no perception of light or who has light perception but is still less than 3/60 in the better eye.[7]

In 2020 in the 36 high-income countries[8], there were an estimated 74 million people with vision loss, and of this 3 million people were blind.[9]

Smartphone Accessibility Features for The Visually Impaired

Accessibility features are meant to make the use of technology less challenging for those with disabilities.

VoiceOver & TalkBack

Screen Navigation

VoiceOver and TalkBack guide users through what's happening on the screen with extra voice, sound, and haptic feedback. Once they're turned on the navigation

[7] WHO, "Change the Definition of Blindness" (2010), 2, https://www.who.int/blindness/Change%20the%20Definition%20of%20Blindness.pdf. Last accessed 7 June, 2021.

[8] according to the Global Burden of Disease (GBD) regional classification system with two criteria: epidemiological similarity and geographic closeness. Last accessed 7 June, 2021.

[9] Data from VLEG/GBD 2020 model, accessed via the *IAPB Vision Atlas,* https://www.iapb.org/learn/vision-atlas/magnitude-and-projections/gbd-super-region-estimates-of-vision-loss/high-income/. Last accessed 7 June, 2021.

gestures will be different and a rectangle border which is a cursor will appear around the currently selected element. If users touch the screen the device will read out what's under the finger. The finger can be dragged around the screen to explore more items. But it is not necessarily to tap directly on something to locate it. Users can instead flick quickly right or left with one finger to find the item they want, like a joystick. To activate the item, for example, to open an app, it is possible to double-tap anywhere on the screen. A three-finger swipe is used to scroll through pages.

Image Description

Texts and objects in images in apps and on the web are guessed by on-device artificial intelligence and read out.

Camera Support

Voice guidance tells you what objects are in the viewfinder if there are faces and where they are. If the device is tilted users get additional audio, haptic, and spoken feedback to help them hold the phone level.[10]

Voice Control & Voice Access

Voice Control (iOS) and Voice Access (Android) is a hands-free feature for anyone who has a hard time manipulating a touch screen that lets users navigate on the phone or edit text with spoken commands.

Zoom & Magnification

Once either Zoom (iOS) or Magnification (Android) is turned on users can zoom in and out and pan around on the screen to get a closer look. They are intended for people with low vision.

[10] Robin Christopherson, "VoiceOver on the iPhone turns 10 – and turns blind access up to 11", *AbilityNe*t, June 20, 2019, https://abilitynet.org.uk/news-blogs/voiceover-iphone-turns-10-and-turns-blind-access-11. Last accessed 7 June, 2021.

Magnifier

Magnifier uses the phone camera to magnify objects or text. In a dark environment, the flashlight can be turned on.

Audio Descriptions

When audio descriptions are available in the media source users are watching, they will be played automatically if this feature is enabled.

Interviews

I had an opportunity to talk to two blind people. Marion Putzer-Schimack is in charge of the project support in BSV-WMB (The Federation of the Blind and Partially Sighted, Vienna, Lower Austria, and Burgenland). Hansol Kim is currently a student in Seoul, South Korea majoring in economics and running a YouTube channel about his everyday life as a blind person.

Marion Putzer-Schimack

Marion became blind gradually because of an illness, has a bit of light perception, sees spots, and uses a white cane. She also has a guide dog named Aida that became more of a pet due to her lack of concentration and passion for food. I asked her about the difficulties I have assumed that blind people would have, which turned out to be pretty much wrong during the interview.

Navigation Apps for Finding Ways

She finds Google Maps and BlindSquare very helpful. In BlindSquare app she can get nearby building information and mark certain spots where, for example, a bump is on the street, so that she can get notified and be more cautious later when she's there again. However, BlindSquare uses location data from Foursquare and Open Street Map thus doesn't have all the shops on the map. Furthermore, once she was going to a supermarket and BlindSquare told her that the supermarket is 20m away.

She ended up going into several wrong stores before she finally got to the right destination.

Obstacles

She doesn't have any difficulties using stairs and escalators. She can feel them easily with her white cane and hear the sound of escalators moving. Yet construction sites can be quite disorienting especially with loud noises and she gets sometimes scared that something might fall on her head.

Public Transportation

Bus and tram drivers in Vienna have to announce the number and the direction of the vehicle through speakers when they spot one or more blind people waiting at a stop. When multiple buses or trams arrive at the same time, drivers have to stop one more time at the very front so that blind people do not miss the one they're getting on. However according to Marion's experiences these rules are often ignored by some drivers.

Subways are somewhat easier for blind people to use since there's only one same train coming to each platform. One might guess that finding ways in the subway station such as finding an exit to go out can be troublesome for the blind, but it is not at least in Vienna. Wiener Linien provides the POPTIS guidance system for blind people (Pre-On-Post-Trip-Information-System), which can be accessed on the barrier-free version of the Wiener Linien website. When blind people choose where they are and where to go in which station, they will get an audio way description.

Acoustic Pedestrian Signals

Marion doesn't see any problem with the acoustic pedestrian signals found on crosswalks. Although there are other more up-to-date solutions employing beacon devices sending signals to smartphones, they can not replace the current system, as not all blind people are smartphone users and it would be rather bothersome if you'd have to carry your phone all the time.

Once there was a special key to turn on the acoustic traffic signal for blind people which wasn't the smartest idea. First, blind people had to find the key in their pockets or bags and then the keyhole.

Spontaneity

I asked Marion if she likes to do something spontaneous on the street such as going to a new park and what she'd like to do if she could see for one day. She does something spontaneous only when she is full of energy and in a good mood because it's a challenge for her since she has to ask people for a favor.

If she could see, first of all, she doesn't want to see just for one day, but she wants to see forever. She wants to ride a bicycle alone without her husband, drive a car alone, and go on a trip spontaneously alone all by herself. Now as a blind person when she wants to go somewhere she has to plan everything.

Marion concluded at the end of the conversation that all she wants was about being able to be spontaneous all by herself.

After the talk with Marion, it seemed that she doesn't have any difficulties that I guessed she'd have except for the construction site. She rather had an eagerness for independent spontaneity on which I focused the conversation with Hansol Kim later.

Hansol Kim

Hansol suddenly lost his vision in his late teenage years. He carries a foldable white cane and also goes the familiar ways without it. When he meets another blind friend, he calls the friend's name until they finally meet.

What he does spontaneously on the street

He asks people if there's a convenience store nearby and go there. When he goes to a new place and if there's a distinguishable smell, he remembers the place by the smell.

What he'd like to do spontaneously on the street

He'd like to know what the surrounding is like and whether there's a place he can take a rest. If the scenery is pretty he'd like to take photos.

What he wants to know about the surrounding beside the building information

He wants to know whether there's a bench around him, whether the surrounding is beautiful and what is exactly in the surrounding, such as stones, flowers or electric poles and of course it would be great to know if there's an obstacle in front of him.

What he'd like to achieve

He wishes to lower the barrier between disability and non-disability. It might sound too great, but he wants that people perceive disabled people just as any other non-disabled people. The ideal society would be where a disabled person is viewed not as a disabled person but just as a person. He will continue his YouTube channel hoping more and more people will get used to disability. If non-disabled people become familiar with disability, he says we can maybe get along without any kind of barrier.

Conclusion

GILBERT was supposed to be a navigation app for the visually impaired with extra features such as stairs alert, indoor navigation, crosswalk signal transmitted from Bluetooth beacon devices, public transportation information, obstacle detection with depth sensor, face recognition.

But after talking to Marion and Hansol, it seemed that I was trying to just provide so to say over-convenience to blind people. Since their desire to be able to be spontaneous independently was bigger than the need for a better and perfect wayfinding mobile application.

Therefore, I decided to rebuild the concept of GILBERT as an app that sees the surrounding of visually impaired people out on the street to make a step forward to their wishes.

Apps for the Blind

AI Image Recognition

Seeing AI

Seeing AI is a free app created by Microsoft that narrates the visual world around its users using the smartphone camera and AI available both on iOS and Android. It has channels including *Short Text, Document, Product, Person, Currency, Scene, Color, Handwriting, Light* which users can switch between.

On the *Text* channel, Seeing AI will speak text as soon as it appears in front of the camera. If a printed page such as a restaurant menu is capture with the phone camera on the Document channel, the app will recognize the text on it. When a barcode is dented on the *Product* channel Seeing AI will scan it immediately and provide product information. Handwritten text can be read on the *Handwriting* channel. On *Person* channel faces of people can be saved and will be detected when they appear in the camera viewfinder. A picture can be taken on the *Scene* channel and AI will analyze what's in the image. Currency bills can be identified on the *Currency* channel and perceived colors can be described on *the Color* channel. On the Light channel the app will generate an audible tone corresponding to the brightness in the surroundings.

Google Lookout

Google Lookout is also free and describes the surroundings detected by a smartphone camera but is different from Seeing AI in that it has three modes for different situations rather than having multiple channels categorized by the features. The three modes are *Explore, Shopping* and *Quick Read*.

Explore is for identifying objects and things around users. On *Shopping* mode barcodes and currency bills can be detected and on *Quick Read,* the app will read the text that appears in front of the camera.

Lookout can not only take a snapshot and detect what's in the image but also can describe what's in front of the camera in real-time until the user tabs the stop button.

GPS Based Location Information

BlindSquare

BlindSquare is a navigation app for the blind. It gives users the nearest building information so that they know where they are. Users can set the radius they want and get the building information within the radius of the current position in different categories such as Food, Arts & Entertainment, Nightlife Spot, and so on. They can save the places they like and mark certain locations when they want to remember the spot later. Since BlindSquare doesn't have built-in navigation, for the directions users will be redirected to a third-party app. BlindSquare is available at the price of 43.99€.

Google Maps

In 2019 on World Sight Day on the 10th of October Google Maps introduced a new feature that gives people the ability to receive more detailed voice guidance and new types of verbal announcements for walking trips. This feature is built up by and for people with vision impairments. Google Maps let users know if they're on the correct route, the distance until the next turn and the direction they're walking in. When approaching large intersections it provides a heads-up to cross with extra caution and if users are accidentally off the right path, they will be given a spoken notification that they're being re-routed.

This feature is currently available only in English in the United States and Japanese in Japan.[11]

AI Image Recognition + GPS

Eye-D

Eye-D is a compact hybrid of Navigation and AI image recognition packed with four features which are *Where Am I, Around Me, See Object in Image, Read Text in Image*. Visually impaired people can figure out where they are by the information given by the app on the nearest buildings and find the nearest place in certain categories like the bank, bus stops, food, hospital, etc. AI will also recognize what objects are there or read out the text in a taken picture.

---

[11] Wakana Sugiyama, "Voice guidance in Maps, built for people with impaired vision", *The Keyword*, October 10, 2019, https://abilitynet.org.uk/news-blogs/voiceover-iphone-turns-10-and-turns-blind-access-11. Last accessed 7 June, 2021.

Eye-D is available for free but can be upgraded to the Pro version with more detailed features at the price of 6.99€.

Including Sighted People's Participation

Be My Eyes

Be My Eyes is a free app that connects blind and low-vision people with sighted volunteers for visual assistance through a live video call. You can sign up either as a blind or low-vision person; or as a sighted and set languages you speak. When a blind or low-vision person requests assistance, a sighted volunteer will get a video call. The video will be connected to whoever volunteer picked up the call most quickly.

It is currently available in more than 180 languages and more than 150 countries. The number of volunteers outruns the number of blind and low-vision users.[12]

Downsides

Location limited objects can not be recognized by AI. For example Seeing AI kept on detecting trash cans on the Viennese street as parking meters.

Real-time text recognition might be helpful indoors but not outdoors due to its very inconsistent accuracy. Firstly, users must know what language the text is written in. And AI apparently can not yet recognize all kinds of typefaces and has trouble reading signboards of stores on the street. It also depends pretty much on internet speed. Therefore texts are sometimes read out a few seconds after the recognition and it can be quite confusing.

Internet speed dependence applies the same to the real-time object detection. Hence the efficiency of AI image recognition would be better with snapshots than in real-time. But this is also accompanied by another inconvenience that users have to hold the phone and navigate on the interface each time.

GPS-based apps that provide location information give only building information which is only one element that the visual world is consisted of.

[12] *Be My Eyes*, https://www.bemyeyes.com. Last accessed 7 June, 2021.

UI Design for People with Visual Impairment

Graphical user interfaces can not interact through the sight with visually impaired people. Therefore it is necessary to replace GUI with other methods to make computer-interaction for visually impaired people possilble. To fill the absence of sight which is the most rapid and precise way to convey graphical information, a UI design should be approached with multimodal interactions not relying on one modality. Multimodal interactions can include voice, text, touch, haptic feedback, hand gestures, pen, and so on.

A common mistake that can be made when designing for the visually impaired is work overload for them. Designers should avoid trying to provide too many aspects of the information considering the limits of attentional capacity.[13]

CONCEPT

GILBERT

The Name GILBERT comes from two Korean words: *Gil* (길) meaning a street or way and *But* (벗) meaning a friend or companion.

GILBERT provides abundant surrounding information on the street to blind people using assisted GPS, the smartphone camera, and the power of AI. At the same time, it maximizes the usage of sensors equipped in smartphones so that blind users don't necessarily have to interact too much with its interface to get the information of the environment that they're in out on the street. GILBERT's features can be turned on and off to easily enable customized information. To complement the imperfection of the AI image recognition, sighted people are also very

[13] Stina Olofsson, "Designing interfaces for the visually impaired" master's thesis, Umeå University, 2017, https://www.diva-portal.org/smash/get/diva2:1179505/FULLTEXT01.pdf. Last accessed 7 June, 2021.

welcomed to use the app and contribute to GILBERT being a companion for blind people with higher performance. On GILBERT sighted people don't have to be physically present around blind people at the right time for the sake of helping them and vice versa.

The interface for sighted people will differ from that for blind people.

Sensors Used by GILBERT[14]

Accelerometer

Accelerometers handle axis-based motion sensing and measures acceleration. It is the reason why a smartphone can track how many steps its user walks.

Gyroscope

Using the Earth's gravity, the gyroscope helps the accelerometer out with determining which way the smartphone is oriented. Thanks to the gyroscope your smartphone knows whether you're tilting the device and whether the screen should be in portrait or landscape.

Magnetometer

Magnetometers can tell which way is north by measuring magnetic fields making it possible to use a smartphone as a compass. Together with the accelerometer and the gyroscope, it knows where a phone is in physical space.

[14] David Neild, "All the Sensors in Your Smartphone and How They Work", *Gizmodo*, June 29, 2020, https://gizmodo.com/all-the-sensors-in-your-smartphone-and-how-they-work-1797121002. Last accessed 7 June, 2021.

Assisted GPS

Global Positioning System units inside phones get a ping from a satellite and figure out which part of the planet its users are at. Modern-day GPS units combine GPS signals with cell tower triangulation and Wi-Fi positioning the distance to enhance speed, quality and precision of satellite signals.[15]

LiDar Sensor

LiDar sensors are laser light scanning technology that can measure depth and map out a room very accurately, which is going to be most useful for augmented reality apps and supposedly influence app markets.

Surrounding Informations provided by GILBERT

During the design process I have contemplated situations in which I take spontaneous actions on the street as a sighted person and assorted what is not possible when the sight is lost, so as to deduce what kind of information GILBERT should be giving to help the blind understand their surroundings.

Visual Surrounding Information

When GILBERT is launched and when the user stops walking, a picture will be taken to be analyzed by AI. GILBERT immediately tells the user what's in the image.

[15] Marc Fevrier, "How Does Location Work: Sources of location data (GPS, Wifi, Cell Tower Triangulation)", *GroundTruth*, March 14, 2017, https://help.groundtruth.com/hc/en-us/articles/360000709047-How-Does-Location-Work-Sources-of-location-data-GPS-Wifi-Cell-Tower-Triangulation-. Last accessed 7 June, 2021.

Finding Objects

On an occasion when the user wishes to find a banal object on the street such as a trash can or a bench, he can tell GILBERT he wishes to find and GILBERT will inform the user when it is detected by the camera.

Location Information

When the smartphone is tilted horizontally GILBERT will tell its user which location is pointed at with the smartphone. The user can point around the smartphone while keeping the horizontal tilt to figure out what locations are surrounding them, in order to find out where he is.

Obstacles

GILBERT provides a head-up if there's an obstacle standing on the user's way with the aid of the laser-based depth sensor equipped on the smartphone.

Photo Spots

During the research phase, I learned that some blind people also enjoy taking pictures and selfies. Some blind people even use Instagram with image description. Some might think that photography makes no sense for the blind. But when we look at the photos we have taken, we bring up the memory from back then. People with vision can spot what they find interesting on the street and take photos of them. Photos do not contain visual aspects but also memories. In that sense, taking a picture could be meaningful for the blind as well.

Therefore I have decided to include the tourist attractions where a lot of people take pictures, that are symbolized with a camera icon on Google Maps in the surroudning informations to be provided by GILBERT.

Faces

It is easy for a sighted person to remember the faces they have often seen. People say Hi when they bump into each other regularly, maybe get to know each other's name, and make acquaintances. Of course, it's also possible for blind people, but

only under the inescapable condition, that the other party speaks. It can be initiated only by a sighted person. To help making acquaintances be rather mutual, GILBERT recognize other users' faces.

Therefore it is recommended that users register their faces when they sign up.

GILBERT will not let its users save face information of other people who didn't sign up on GILBERT. It's because, I believe such information should be directly controllable by its owner. If the owner wishes to remove his face information stored on a database, he should be free of having to ask another person to do so. Furthermore, the necessity of face recognition for the blind does not excel the importance of protecting personal information.

Redirection Notes

GILBERT will be able to read redirection notes for the blind users. For example, "entrance around the corner", "pedestrian walkway on the left side", "caution, wet floor" and so on.

Poop on Walkways

Blind people can not spot poop on walkways, therefore can not avoid them. White canes are foldable so that blind people can put it in their bags when they don't need them. So when a white cane pokes a pile of poop, it's not very pleasurable. GILBERT will help blind users avoid this dirty obstacle.

Sighted Users

Imperfect AI

Sighted users can take pictures of anything and see what GILBERT thinks is in the picture. They can tell GILBERT whether the result of the AI image recognition is correct or wrong and they can even correct the wrong result. But of course, it will not immediately affect the database. It will be first controlled by the GILBERT app development team.

Redirection Note

Text recognition is omitted in GILBERT, because it's capricious in the accuracy when used outdoors and because it collides with the principle of GILBERT, "least app navigation".

Yet, there might be some important text to read such as "entrance around the corner", "pedestrian walkway to the left", "caution, wet floor", "bench painted, don't sit down" and such.

Sighted users can take a picture of these redirection notes and type what they say. So that blind users can be redirected without any troubles.

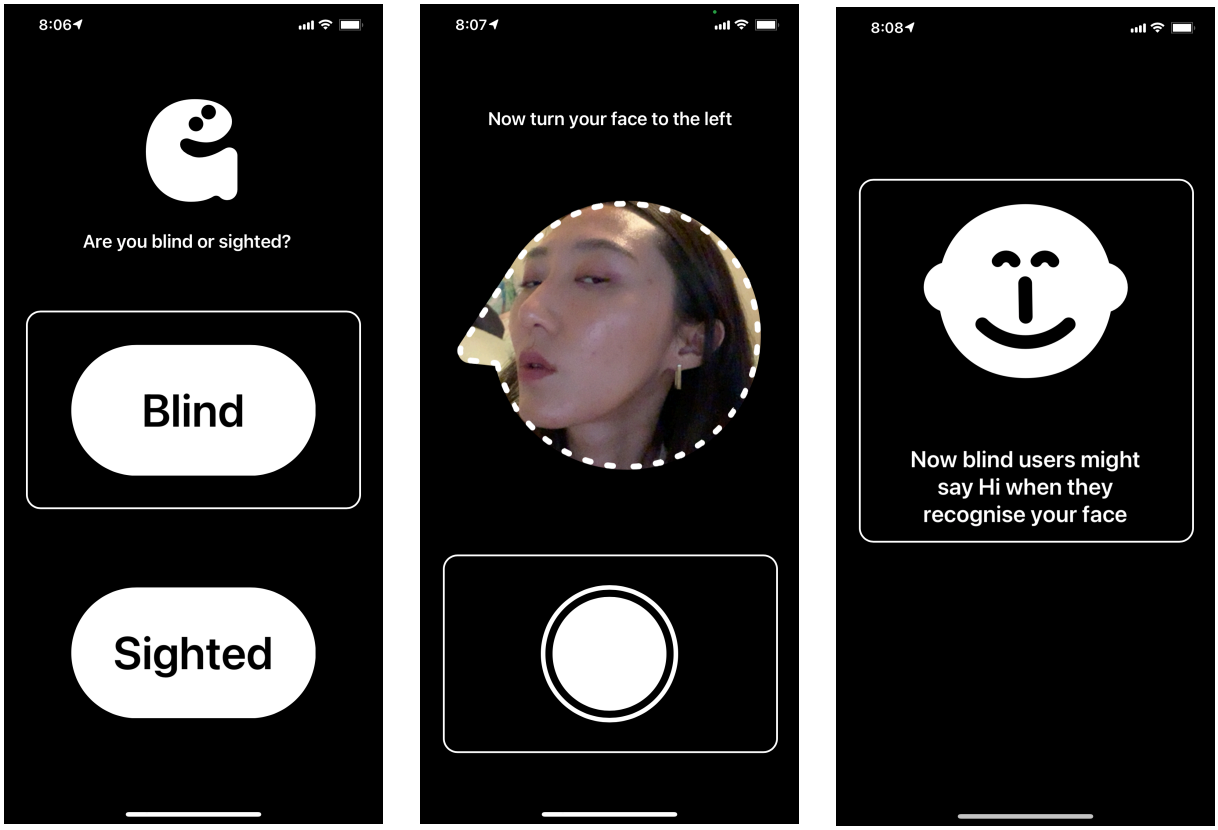
Poop

Poop on the street is probably the most irritating unexpected obstacle evenly for both the and the sighted. The Laser-depth sensor will not detect poop, because if it has to, it will recognize every little small bump or object on the ground that are unnecessary to be detected, resulting in excessive information overload.

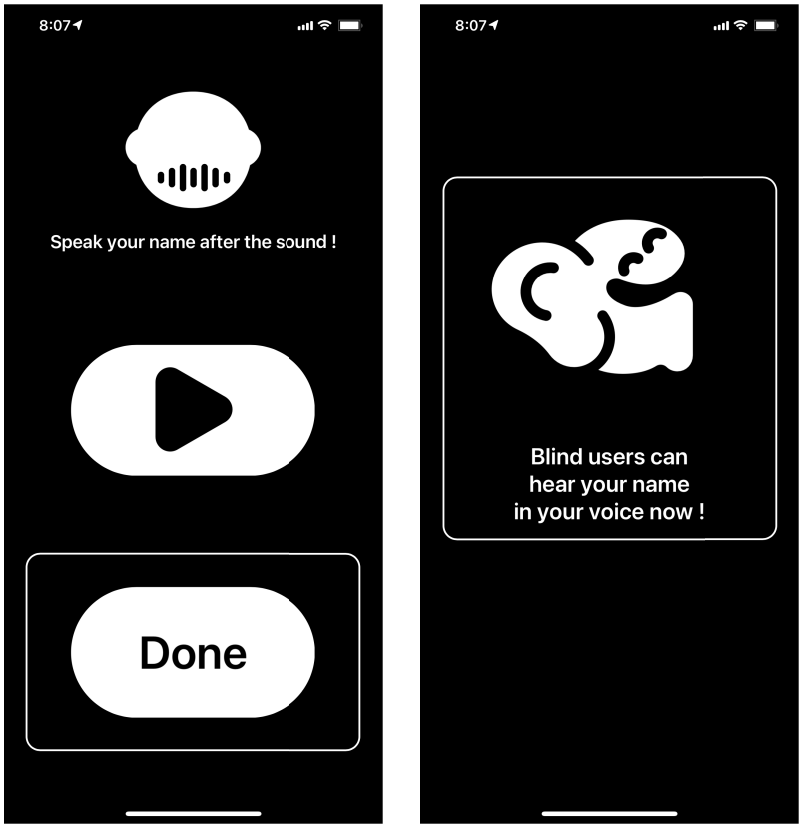
If sighted users spot poop on the street, they can take picture of it and leave poop warning in that location. A *Poop Warning* will expire at the end of the day, since it's not gonna be there forever.

IMPLEMENTATION

Sign Up



When signing up on GILBERT, first the user has to choose whether he is blind or sighted, so that the right interface can be shown. Then he can optionally register his face by taking three selfies so that GILBERT can recognize his face for other blind users. The last step is to record his name in his voice, which is also optional. If a spoken name is given, blind users can hear their names in their voices instead of the username spoken by TTS.



Keep the Smartphone Steady

Since GILBERT uses the camera and the sensor on the backside of the phone throughout, blind users will have to hold the phone in their hand all the time.

But should the hands be free, they can also put their phone in a chest pocket, use a phone case with a necklace strap or wear a phone harness. If there are AR glasses available, it also could be an option.



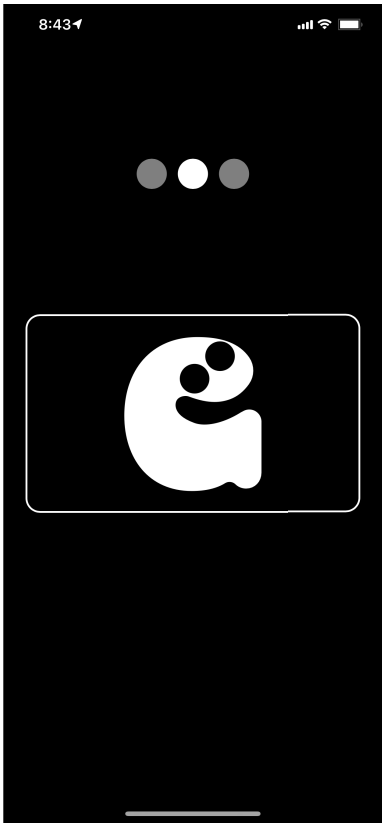

Interface & Features

For the Blind

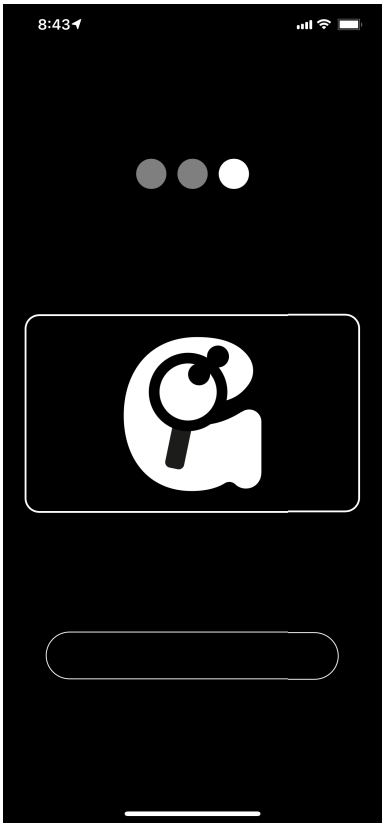
GILBERT's main functions are to see, to find and to provide location information. Each of them are called *GILBERT, What Do You See?; GILBERT, Find; and GILBERT, Where Am I?.*

The interface has three main pages, which can be easily scroll through with the three-finger swipe gesture and each of the pages is in charge of one of the main functions.

GILBERT, Where Am I ?



GILBERT, What Do You See ?



GILBERT, Find

GILBERT, What Do You See?

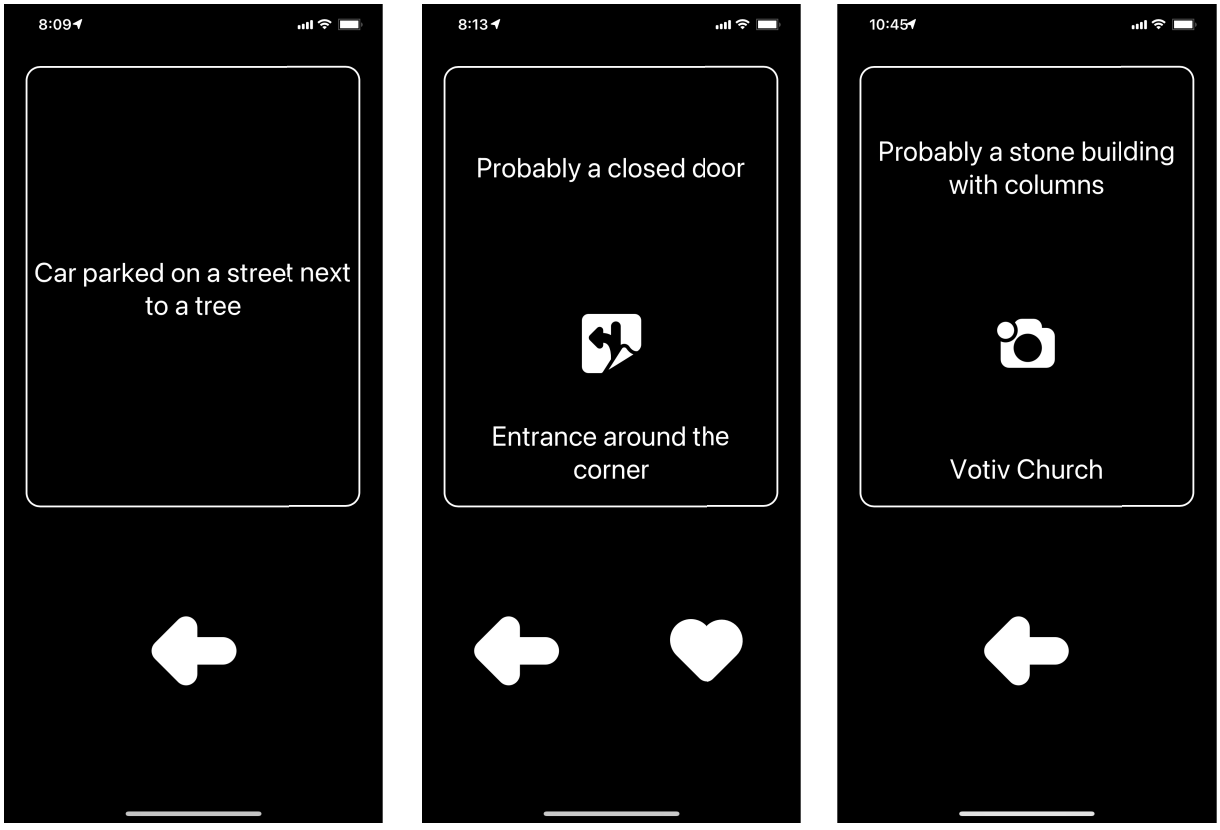
On this page the user can just start walking. When the user stops walking, GILBERT will tell the user what GILBERT sees in front of him. Any information will be given on

a separate popped up page. The page can be dismissed by either just start walking again or by double-tapping the back button.

A smartphone can know when the users is walking thanks to the accelerometer.

When GILBERT sees a redirection note left by a sighted user, GILBERT will immediately read it out for the blind user. A heart can be sent to the sighted user who has typed what the redirection note said, if the blind user wishes to. After a heart is successfully sent, the blind user will hear the name of the sighted user in his voice, if it was recorded during the sign up process.

When GILBERT sees a photo spot in front of a user, GILBERT will let the user know what it is. These are possible because here location data is combined with image data for the AI image comparison.



Stopped Walking | Redirection Note | Photo Spot

GILBERT, Where Am I ?

On *GILBERT, Where Am I?* page, the user has to tilt the smartphone approximately horizontally. Then GILBERT will inform which location the smartphone is pointing at.

The user can point around the smartphone to find out his current location. He can open the pointed location in Google Maps or Maps to get detailed information about the location.

Smartphones know how much it's tilted on which axis because they have a gyroscope in them.

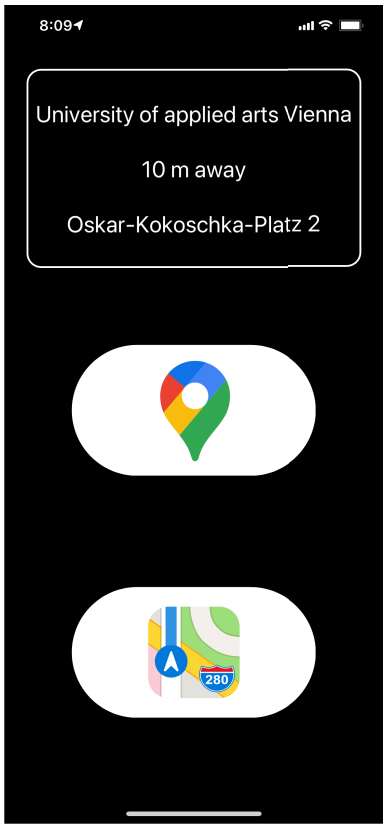
GILBERT, Find

When on this page, the user can immediately say what he wishes to find or if he is not very fond of speaking to the phone, there is an option to type. Then GILBERT will start finding it.

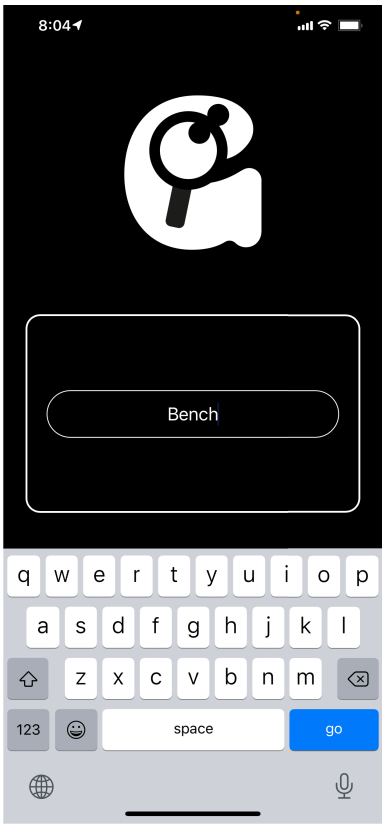
GILBERT is capable of finding signs of stores, if they are on the database. A franchise supermarket, for example, should not be a problem.

Once it's detected, the user will hear where it is in which direction and how far it is. GILBERT will know how far it is thanks to the LiDar sensor on the phone.

If the wanted object is found, then user can go back to the main page by tapping on the back button.



Pointed Location



Typing on GILBERT, Find

A Bench Found

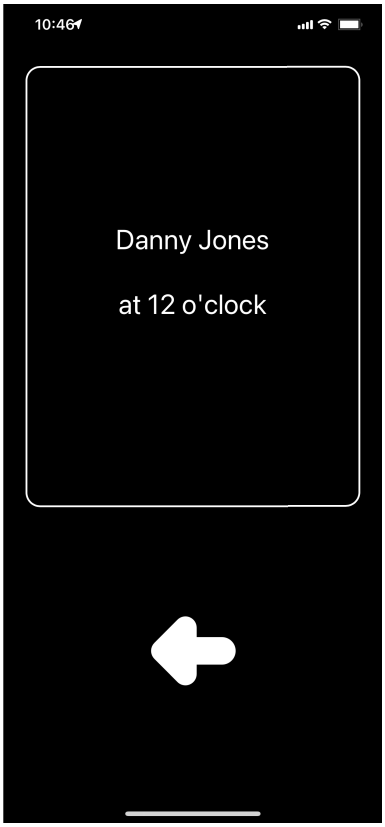
Obstacle

With the aid of the LiDar sensor GILBERT is able to know when the users is about to bump into an unexpected obstacle. If there is an obstacle ahead the user, a tone according to the distance to it will be generated, when on the pages *GILBERT, What Do You See?* and *GILBERT, Find*.

If the user is on *GILBERT, What Do You See?* and stop walking after hearing the generated tone indicating an obstacle, GILBERT will say what's there. Thus, the user can know what is blocking his way.

Poop

When a sighted user had left a poop warning in a location, GILBERT will let the blind user know, once he enters that area. Then the user has to head the phone camera towards the ground, so that GILBERT can find the poop using the AI image comparison within the limited location. If there is an object in the phone camera range that matches the poop in the image taken by the sighted user, a tone will be generated and the blind user will be able to avoid the stinky obstacle on the ground.



Recognize Face

Pointed Location

Recognize Face

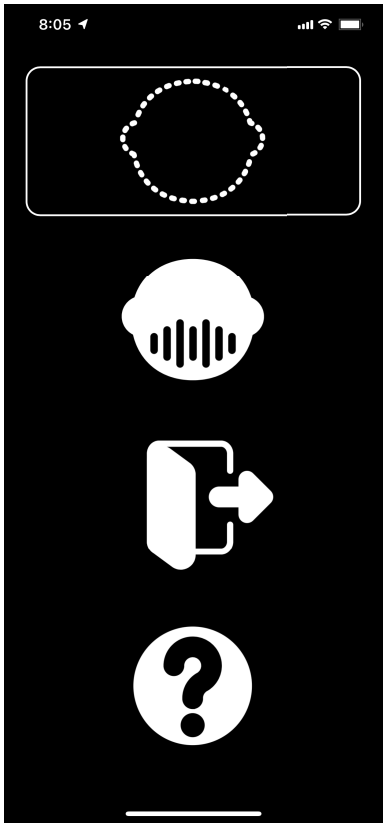
Blind users will hear the name of another user, both sighted and blind, whose face is detected by GILBERT and in which direction they are. Now, blind people can also say Hi first!

Account Settings & Options

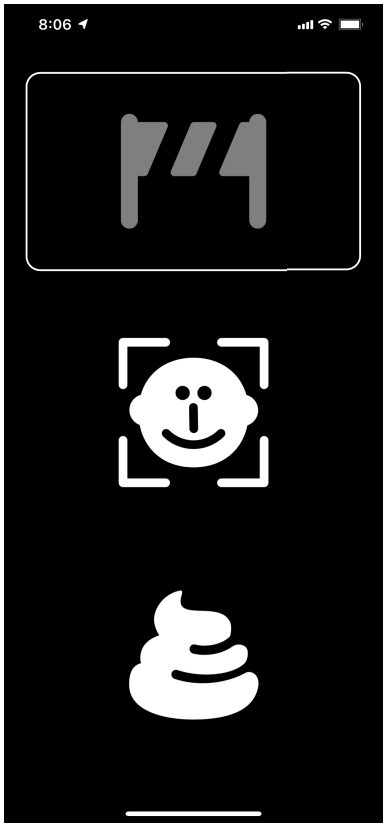
Account settings can be reached by scrolling down on any of the main pages with a three-finger swipe, and options by scrolling up.

In the account settings, a user can reset his face, reset the spoken name, log out or start the app tutorial.

In the options, it is available to customize additional features, *Obstacle, Face Recognition* and *Poop Warning* by turning them on and off with a double tap.



Account Settings



Options

Voice Command

To offer blind users a possibility of ultimate hands-free, GILBERT can be verbally told what to do. All they have to do is to call his name “GILBERT" and wait until GILBERT answers “yes”. And then they can tell GILBERT what to do. For example, "What do you see?", “Turn *Obstacle* off.”, “Find bench.” and whatnot.

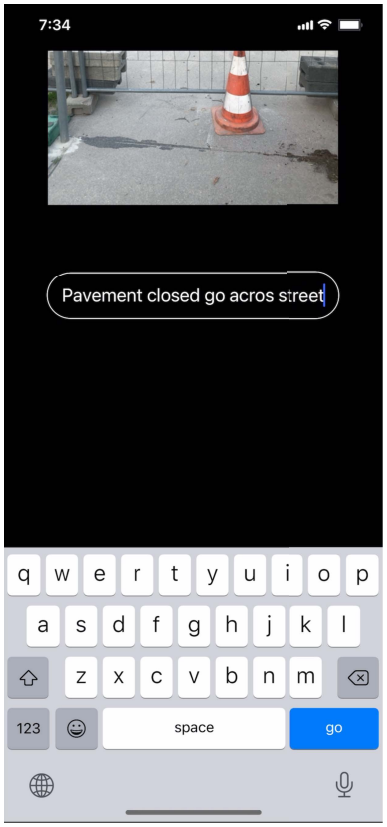
For the Sighted

Sighted users have a camera viewfinder on their version of the interface with three buttons for different purposes: training GILBERT, reading a redirection note and leaving a poop warning.

Redirection Note

When a sighted user spots a redirection note on the street, he can take a picture of it and then type what it says.

Interface for the Sighted



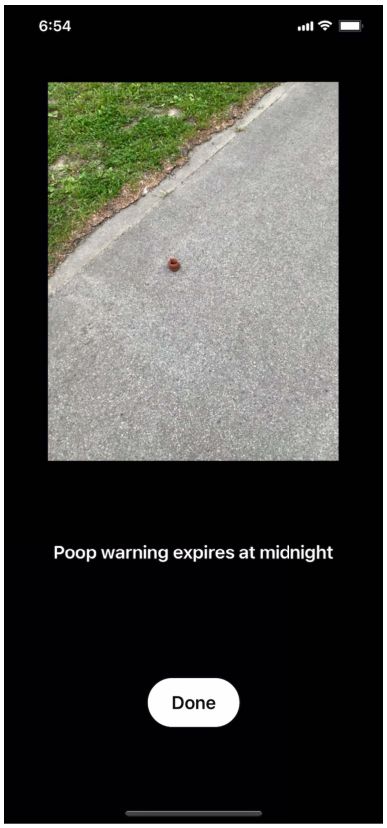
Typing a Redirection Note

Poop Warning

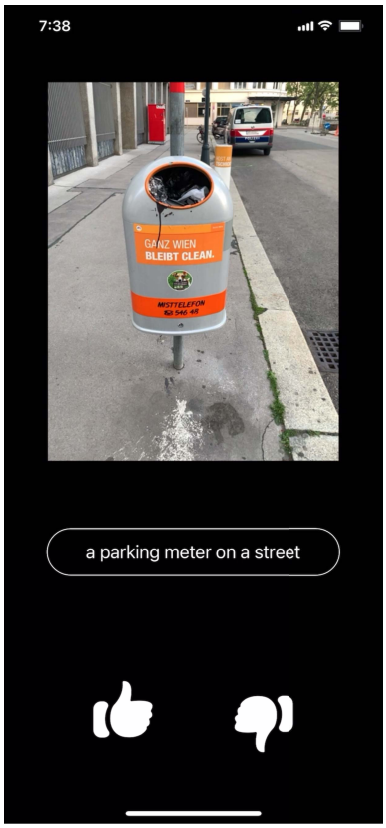
If a pile of poop is lying on a walkway, a sighted user can take a picture of it and just press *done* button. A poop warning expires at midnight because hopefully it will be cleaned the next morning.

Train GILBERT

A sighted user can tap the middle button which is the GILBERT logo, to check if GILBERT recognizes and decribes the image correctly. Should the result be correct, he can tap the thumbs-up button. When it's wrong, he can tap thumbs down or even type in the right result.



Leaving a Poop Warning



Training GILBERT

REVIEW & OUTLOOK

GILBERT is an attempt to make a companion for blind people that can be around whenever and wherever they need employing smartphone technology as much as possible. It wants to bring people with vision impairment a step forward to being able to be spontaneous independently and to use infrastructure in a public space as an equal member of society. GILBERT boost up the self-empowerment of blind people, as they don't necessarily have to depend on another person's help.

GILBERT triggers social contribution of the sighted people by inducing them to spare some time for empathy and to take actions, but it doesn't rely on sighted users' help completely. GILBERT has a faith that there is always good will lying in human nature. How many sighted people are willing to sign up on GILBERT can be revealed only after GILBERT is actually out on the app market. However, as seen in the case of Be My Eyes, which essentially requires the sighted people's voluntariness and which has more than 4 million sighted people signed up, my anticipation remains positive.

GILBERT giving a chance of meeting new people by accident could lead to, what I call, augmented inclusivity. We will have better understanding of each other no matter blind or sighted. And hopefully the information gap between the blind and the sighted narrows down and eventually everyone will be dwelling in not separate societies, but in one same equal society.

Furthermore, GILBERT, unfortunately, necessitates a most recent high-end smartphone. Smartphones with all the needed sensors are not widespread yet but will be in the near future and in the meantime, more flesh can be added to the bone of GILBERT.

To bring GILBERT out of a fiction, it should be programmed, tested by blind people and repeat it multiple times. Economical and financial aspects also have to be thought through and measured.

And when once GILBERT is realized, I am also planning to design another GILBERT, who will accompany blind people when they are shopping alone and maybe I will call it something like "Shopping GILBERT".

SOURCES

Accessible Android, "A Brief History of Android Accessibility", August 2, 2013, https://accessibleandroid.wordpress.com/2013/08/02/a-brief-history-of-android-accessibility/

Be My Eyes, https://www.bemyeyes.com

David Neild, "All the Sensors in Your Smartphone and How They Work", *Gizmodo*, June 29, 2020, https://gizmodo.com/all-the-sensors-in-your-smartphone-and-how-they-work-1797121002

Graeme Fulton, "Accessibility Basics: Designing for Visual Impairment", *ENVATO TUTS+,* December 2, 2017, https://webdesign.tutsplus.com/articles/accessibility-basics-designing-for-visual-impairment--cms-27634

Hosea Sanders, "New app, beacons make it easier for blind to move around", *abc 7 News*, October 9, 2016, "https://abc7news.com/technology/new-app-beacons-make-it-easier-for-visually-impaired-to-move-around/1547148/"

Jon Knight, "LiDar vs. 3D ToF Sensors – How Apple is making AR Better for Smartphones", *Gadget Hacks*, March 31, 2020, https://ios.gadgethacks.com/news/lidar-vs-3d-tof-sensors-apple-is-making-ar-better-for-smartphones-0280778/

Life After Sight Loss, youtube.com/user/derekdanielvids

Marc Fevrier, "How Does Location Work: Sources of location data (GPS, Wifi, Cell Tower Triangulation)", *GroundTruth*, March 14, 2017, https://help.groundtruth.com/hc/en-us/articles/360000709047-How-Does-Location-Work-Sources-of-location-data-GPS-Wifi-Cell-Tower-Triangulation-.

Max-Planck-Gesellschaft, "Look first", January 15, 2015, https://www.mpg.de/8849014/hierarchy-senses

Meghan Tocci, "History and Evolution of Smartphones", *SimpleTexting*, https://simpletexting.com/where-have-we-come-since-the-first-smartphone/

Michael Gillan Peckitt, Japan to install 'smart' equipment to aid visually impaired people at road crossings", *Barrier Free Japan*, October 25, 2020, https://barrierfreejapan.com/category/blind/

My Review Room, "Netflix to Selfie! How do visually impaired people use iPhone?: iPhone iOS Accessibility", November 17, 2019, https://youtu.be/6N0Rai3GtQ8

OneshotHansol, https://www.youtube.com/channel/UC5ZqaJ9klJrwgsKjYW42JtA

Sean O'Kane, "The first truly wireless earbuds are here, and they're awesome", *The Verge*, January 9, 2015, https://www.theverge.com/2015/1/9/7512829/wireless-earbuds-ces-2015-bragi-dash

Seeing AI, https://www.microsoft.com/en-us/ai/seeing-ai

Stina Olofsson, "Designing interfaces for the visually impaired" master's thesis, Umeå University, 2017, https://www.diva-portal.org/smash/get/diva2:1179505/FULLTEXT01.pdf

The Blind Life, https://www.youtube.com/channel/UCNbzN3eHbLKPzltSB560DkA

TigerMobiles.com, Evolution of the Mobile Phone, https://www.tigermobiles.com/evolution/#start

Vision Atlas, https://www.iapb.org/learn/vision-atlas/

Wakana Sugiyama, "Voice guidance in Maps, built for people with impaired vision", *The Keyword*, October 10, 2019, https://abilitynet.org.uk/news-blogs/voiceover-iphone-turns-10-and-turns-blind-access-11

WHO, "Change the Definition of Blindness" (2010), 2, https://www.who.int/blindness/Change%20the%20Definition%20of%20Blindness.pdf

Wiener Linien, "Accessible Mobility", https://www.wienerlinien.at/eportal3/ep/programView.do?pageTypeId=66533&channelId=-4001458&programId=2177870&contentId=1801143&contentTypeId=1001

Wiener Linien POPTIS, http://poptis.wl-barrierefrei.at

YunjiTV, "How do visually impaired people use smartphones?", November 27, 2019, https://youtu.be/6jT-POAaB0c

The last access to all of the sources above: June 7, 2021