Ocularis

A virtual assistant for the visually impaired – **a containerized and unified solution to tackle some of the daily challenges faced by millions worldwide who have complete/partial visual impairment.**

# The Team

**Tell us the story about your team. Were you at a hackathon together, or partnered on a class project? How did you decide to work together to submit to Imagine Cup?**

**Gaurav Karna,** McGill University, 3rd year Cognitive Science & Computer Science, 2020 – Hi! I am currently a 19-year-old, 3rd year student at McGill University in Montreal, Canada studying Cognitive Science and Computer Science. I was born in India but emigrated to The Bahamas when I was 3 years old. Every day, I strive to put my best foot forward by creating high impact applications for the betterment of my local and global community.

**L B Shaaran**, VIT University, 2nd year Electronics and Instrumentation, 2021 - Thanks for stopping by. I have won 20+ hackathons around the world and have additionally been welcomed as a speaker to a few. I have begun coding from 6 years and I currently program in 12 languages. I believe in innovating the impossible and making technology accessible to everyone.

**Mentors:**

* Prashant Bagga – Our Microsoft representative. He made himself available for any questions we might have had throughout the entire process.
* Ehsan Mirsaeedi – A friend at Concordia University who gave us direction about the architecture in the codebase and helped greatly with conceptual and market research for the product.

Ocularis was actually the brainchild of Ehsan, myself, and a third team member back in December for the Discover Azure Bootcamp in Montreal. After winning that, we worked a bit more on the idea. However, Ehsan and the other teammate had to drop out for a mixture of personal and academic reasons. In regards to the current team, Shaaran and I met through a make-a-thon at Harvard University (MakeHarvard) in February 2019. We came up with an idea for a remote-controlled drone that performs environment analysis in military and rescue scenarios to give strategic feedback and information to the user about the situation from a birds-eye perspective. After our success at the competition in Harvard, I decided to invite Shaaran to join forces in my project for the Imagine Cup. After all, this is a large project and needed all hands on deck. After several calls and messages, Shaaran and I further conceptualized and improved the idea from the initial rendition which had won the Azure Bootcamp. We are super proud of what we have created, and cannot wait to hear your feedback on Ocularis.

# The Concept

A central promise of technology and all its advancements is to bring a better world through solving inherent problems of human society and its individuals. The predominance of smart devices is so obvious that it is now impossible to imagine a world without them around. However, there appears to be a lack of consideration regarding accessibility in contemporary technology, with the industry being heavily biased towards people without impairments of any kind. Nevertheless, things like a smaller market size and weak representation in government and industry are no longer viable justifications for a pervasive tendency towards the aforementioned specific catering. Apart from non-technical impediments, there are also some major technical challenges in providing these people with high quality services on par with natural human capabilities. However with the advent of new computing milestones such as machine learning and computer vision, we believe that we are in a unique era of human history such that these challenges have become easier to surmount. One major challenge where we see potential to capitalize on these technologies and make a difference is in the realm of visual impairment.

Visual Impairment is a decreased ability to see to a degree that causes problems not fixable by usual means, such as glasses [1]. Hence, ordinary tasks become difficult and make them less independent. Simple things such as being aware of their immediate surroundings and recognizing who they are conversing with become exceedingly difficult. We believe we can address these issues using Microsoft Cognitive Services and eventually with Windows IoT core cognitive capabilities, along with a cheap device we call ‘Ocularis’ – meaning ‘eye’ in Latin - that we intend to design and build. Ocularis is an intelligent digital assistant who can understand requests of its user via voice commands and can provide thoughtful responses through her voice. Ocularis tries to find answers to questions by gathering information from its integrated knowledge, the internet, or from the environment that she sees through a high-quality built-in camera. The concept for Ocularis is a simple one, with the aim to make it as seamless as possible to integrate into daily lives. Hence, we want to incorporate the device as a wearable necklace to provide a nice viewpoint for the camera while minimizing impedance with the user’s activity. Essentially, the device is intended to be a third eye for the user and make tasks, simple or complex, much easier to deal with.

## Target Audience or Market:

Visual Impairment is defined as a visual acuity of worse that 20/40 [1] and low vision is often defined as a visual acuity worse than 20/70 [3]. Unfortunately, there isn’t an exact estimation of number the blinds and the visually impaired [13]. According to World Health Organization, in 2018, 217 million people have severe vision impairment and 36 million people are blind [4]. WHO reports that 65 percent of visually impaired, and 82 percent of blind people are over 50 years of age. Unfortunately, this number is increasing because of population growth, ageing problems, and consequences of unhealthy life styles. Reports suggests that this number is expected to rise to more than 550 million by 2050 [5].

According to a research published by The Lancet Global Health, 89% of the visually impaired live in low and middle-income countries [6]. However, the population in western countries is still noticeable. As claimed by World Health Organization, in 2010, there were near 26 million visually impaired individuals live in European countries [7]. Other reports state that over 2 million people in UK suffer from sight loss and “In the UK alone, 250 people will start to lose their sight every day” [8,]. Based on a new report published in 2016 by Center for Disease Control and Prevention, near 26 million people experience vision loss in the USA [9]. Based on this report, near 11 million of the mentioned population live alone without any partner and less than 6 million of them have a bachelor's degree or higher. Moreover, it states that “Approximately 10.1 million people with vision loss in the U.S. have a family income of less than $35,000, 2.8 million have a family income between $35,000 and $49,999, 3.6 million have a family income between $50,000 and $74,999, 2.3 million have a family income between $75,000 and $99,999 and 4.5 million have a family income of $100,000 or more”. In Canada, the population of the visually impaired is estimated to be around 800,00 and near 5.59 million have an eye disease that could cause sight loss [10,11]. In Australia, it is estimated there are over 575,000 people who are blind or vision impaired. However, despite the fact that blindness and vision impairment affect a significant proportion of Australia’s population, no comprehensive data exists on the provision of blindness and low vision services at a national level [15]. In China, another potential market of our product, the number of people who suffer from severe visual impairment is around 17 million and one third of them are blind [12].

Therefore, we are confident in saying that the potential market capture for Ocularis is very well-defined in a global context. Unfortunately, reports show that these people deal with serious problems throughout their lives because of the lack of support from society and from loneliness. In the UK, 31% of visually impaired are never or rarely optimistic about the future. Few explanations exist for this unfortunate statistic. Reports estimate that only 17% of people with sight loss receive emotional support and less than a fifth received practical support that helped them get around their home. 35% of people who suffer from vision loss say that they receive frequent or consistent negative feedback from the public [14].

### Personas

In this section we list our potential customers through personas and explain about challenges they have in their day to day life and how Ocularis can be solution to their problems.

1. **Thomas**. He is 60 years old. He lives alone, and his children live in another city. He lost his eyes in an accident when he was 51 and since then he is highly dependent on services that he gets from an agency which has inevitably led to a severe loss in confidence and independence. Most importantly, he has a fear of unknown places because he does not know where the obstacles and other hazardous objects are in his surroundings. This fear of the unknown has confined him into an unwarranted house arrest. However even at home, he needs to deal with lots of daily challenges like reading letters or notes he gets from his landlord. When he wants to get to his doctor, he always has to ask his friend to take an Uber with him. Last week, he heard something about the successful touchdown of NASA’s InSight lander on Mars, but he had problems accessing any further information about it after that. He knows people these days use the Internet to get information about everything, but he unfortunately cannot take advantage of these tools. Even digital assistants like Cortana do not answer him back by voice when he asks them a question like “What’s the phone number for RBC customer service?”. Further, when he wants to order in a pizza for example, he doesn’t know what restaurants are nearby and their quality. On the bright side, he has heard that Ocularis can help him overcome some of these problems.
2. **Sarah**. She is 30 years old. She has partially lost her sight when she was 18. She stays alone at home until her parents come back from work. During the day she faces challenges overcoming basic tasks. For example, it takes a lot of time for her to find the TV remote controller if her parents forget to put it at its specific place. Even when she wants to pick something from the fridge, she really is not sure if she has grabbed a milk carton or a juice. Moreover, she misses those old times when she could read books like Harry Potter or The Twilight Saga. She has heard that there are lots of audio books available nowadays, but it is really difficult and nearly impossible for her to use these applications on her phone due to lack of accessibility options. Lastly, Sarah’s mother, Carine, is always worried about her daughter while she is at work. She wishes there was a way to see what her daughter is doing, or that she could help Sarah with her daily challenges by seeing from Sarah’s point of view. Sarah’s father, Tim, has heard about a new device called Ocularis that can help them to assist their daughter remotely or even can read a book aloud just by saying phrases like “I want to read Harry Potter and the Philosopher’s Stone”.
3. **Jonatas.** He is the 45-year-old father of a 10-year-old daughter named Yasmin. Jonatas wishes he had the chance to go to Yasmin’s school and pick her up. But unfortunately, he has a fear of the outside and unknown places. His fear comes from a bad experience 11 years ago when he fell down and broke his legs because of an unexpected pothole in the sidewalk. Recently, he has heard that a device called Ocularis can help him to get rid of his fears in the future by warning him about dangers and obstacles that are in his way. Moreover, it can show him the way to the school and navigate him safely via audio feedback.

## Feedback

After pitching our product to a couple of people, one of whom has experience in the field of assistive technology – we heard several areas of feedback:

The Python application we have implemented at the moment is not the fastest that Ocularis can potentially be. This is in large part due to the high-level nature of the language itself, as well as the dependencies that we have incorporated into the script. **With this in mind**, we have begun work on the C# platform of Ocualris that will be hosted as a **UWP** application, but on the same Windows IoT platform. For the current submission, we are confident in providing a demonstration of our software and what it will materialize to be once we take it to market. That being said, we have also expanded our viewpoint in taking **multiple environments** into account. By having two different containerized version of our software, we can make it available to run on multiple systems if we chose to do so. Nevertheless, we are sticking with running on Windows machines.

Another concern we heard was the number of API calls we are making with each intent and the latency that compounds as a result. This is an interesting area to discuss, as it begs for compromise between features and latency. In most scenarios, the effective return time of Ocularis for an intent will be dependent on a few major factors such as internet speed, Azure API availability, and load balancing, among other things. We are striving to make Ocularis as accessible as possible in offline scenarios to guarantee maximal functionality, but there are some inherent functions that are completely dependent upon internet connection – particularly those relying on web search and Cognitive Services.

A third and popular area of feedback we heard was about the nature of pairing and the user interface of Ocularis with the user’s phone. Making something to benefit the visually impaired also begs the question of how can we ensure that we make it as interactive as possible without having major set-up time and vision-dependent interfaces. A potential solution we can work to implement is to make the pairing as seamless as possible, such as Apple CarPlay – where there is no need for going into the Bluetooth menu after the first setup, and connecting happens automatically.

# How it works:

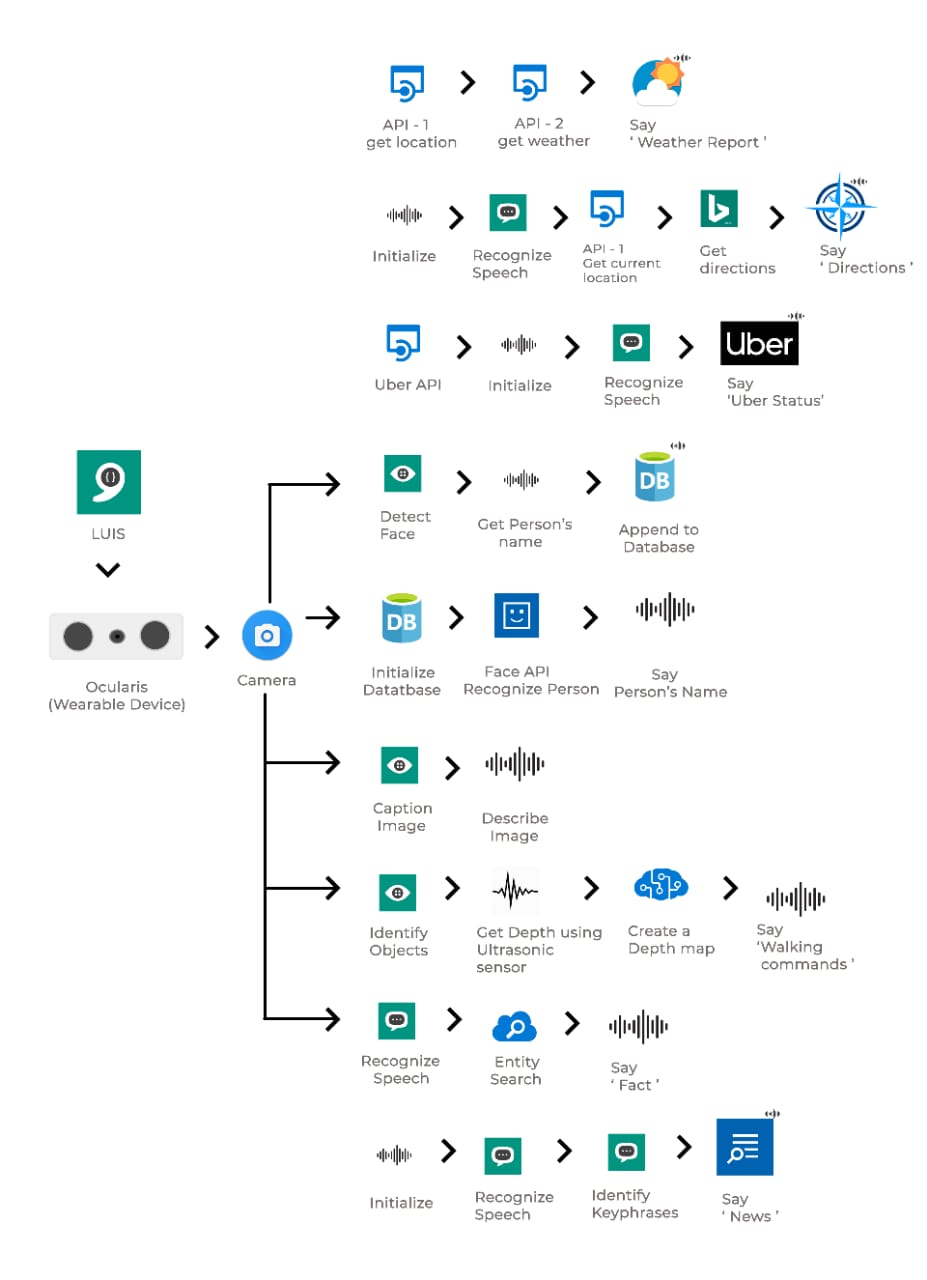
The heart of Ocularis is empowered by a **Python** application hosted on **Windows IoT Core**. We have also developed a **UWP** version of the application that is available on the repository, however the latter platform is still under **development**. This software operates a wearable light-weight device which is supposed to placed as a necklace around the neck of its user. Ocularis consists of a Camera, Audio I/O Module, Wi-Fi Module, Bluetooth Module, Cellular Module, a low consumption processor, and RAM. These modules enable the device to interact with its user via voice and delegate the processing needs to the cloud. The variety of connectivity modules ensure us that Ocularis is always connected to tbe internet when available. Having high quality internet connection is crucial to some functionalities because Ocularis is heavily dependent on **Microsoft Azure Cognitive Services** such as **Computer Vision** and several **Bing APIs** to bring the incredible features to the customer.

Obviously, internet connection introduces latency and uncertainty to our product. We have taken two measures to mitigate this problem. First, Ocularis tries to return the answer to its user’s request by going through tbe **Ocularis Internal Knowledge Base** which is sufficient for some basic needs such as asking about time, date, calendar, or anything that is accessible from user’s phone like messages, emails, and contacts. Secondly, we try to minimize the number of requests that we need to make to Azure Cognitive Services by using a server that acts as a mediator between Ocularis and Azure. Without this mediator, we would need several back and forth between Ocularis and Azure Cognitive Services to be able to serve some requests. However, by putting this Mediator in place, Ocularis just needs to send one request to the Mediator and then Mediator will be responsible to handle the required conversation with Azure. Further, we have minimized the number of API calls even more by having pre-loaded audio files that will consistently be used throughout interacting with Ocularis in a static audio folder.

In addition to **Ocularis Internal Knowledge Base**, this device has a built-in **Internal Image Processor** in order to alarm the user in case of detecting potential danger of falling down, getting out of a sidewalk, or getting close to an obstacle. This functionality is not supposed to be a replacement to guide dogs but is complementary to them. It affords a higher level of independence than with previous arrangements. It should be mentioned that existing Depth and Obstacle Detection approaches do not guarantee 100% accuracy, but we aim to implement and train the model which provides the most confidence and deterministic behavior. This process needs to be done by the device and cannot be delegated to the cloud because we need a near real-time alarming system to increase its reliability. This feature is and will consistently be under development, and will be released in beta phases to the users with accuracy warnings.

Lastly, to make the conversations between user and Ocularis seem intelligent and pleasant, Ocularis attaches a context to user requests and provides a response based on this new context and previous contexts that have changed the state of conversation. In other words, Ocularis does not interpret user’s commands in isolation but it takes the context of conversation and its state into consideration. As an example, when a user asks a question like “Who is the current prime minister of Canada?” and Ocularis notifies its user that she has found 5 relevant pieces of information, Ocularis initiates a state machine (context) to the pair of request and response. So, when the user asks for further questions such as “give me all your findings”, “give me the first relevant one”, or “repeat it again please”, Ocularis knows how to respond to the request according to the state of the conversation. **Ocularis State Manager** is responsible for transitioning through the possible set of states and even memorizing the state of previous requests. As an example, a user can ask a question like “who **was** in front of me?” after getting the required information about Canada Prime Minister and since device has answered to the question “who **is** in front of me?” just before the question about Prime Minister, she can respond back quickly without sending requests to Azure Cognitive Services. Moreover, Ocularis State Manager helps with navigating Ocularis internal settings.

**Architecture:**



In the following sections, we give a brief explanation about the flow of a user request from the beginning to the end which results in an answer. Through this explanation we clarify the purpose of the main components of our architecture and the interactions between them.

## Interactions

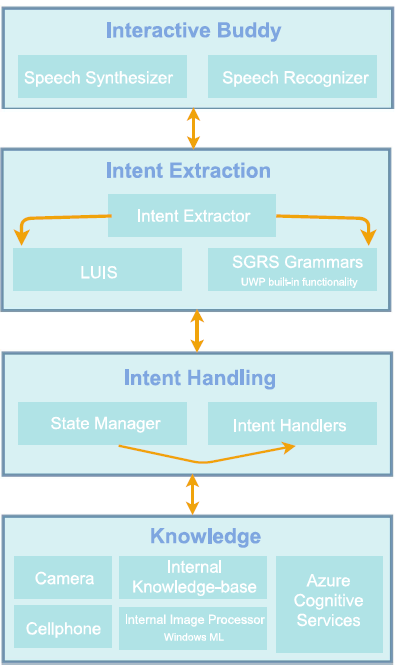
Ocularis notifies the user or responds via its voice which is constructed by the built-in **Speech Synthesizer** of Windows IoT Core. Since Windows IoT Core supports **Speech Synthesis Markup Language**, Ocularis voice would be very similar to a real voice in terms of pronunciation, volume, pitch, rate and emphasis. User can hear the voice using a headset.

There are two major scenarios that Ocularis notifies a user of an event. First, when it detects that the user is approaching a danger like an obstacle or stairs and secondly when the connected cellphone publishes a notification such as incoming call, email or text message. This is intended to run in real-time to afford a high level of safety for the user. In either way, Ocularis, notifies the user in a clear voice and user can ask for further information such as “read it for me” or “I want to answer it back”.

At the same importance, user can ask several types of questions from Ocularis and wants him to answer back. When user asks a question, Ocularis maps the voice to one of intents that it supports.

## Intent Detection

Intents are building blocks of Ocularis which define its core capabilities. In other words, user requests get mapped to one of the built-in intents of Ocularis and that selected intent will be responsible for gleaning the information that user has asked for. Intents express the logic required for querying the information from internal and cloud knowledge sources.

We have designed two ways for analyzing user voice in order to detect the actual intent the user has. First, we use **Speech Recognition Grammar Specification (SRGS)**. A SRGS Grammar file gives us a great control over defining a complex set of commands and Windows IoT Core built-in **Speech Recognizer** is capable to capture semantics of a voice based on the SRGS file and map it to an intent. If offline recognition cannot be successful in capturing the intent, we use **Microsoft Language Intelligent Understanding (LUIS)** as the last resort. This service requires an active internet connection and only accepts texts as an input. So, we must convert the user’s voice to text before passing it to LUIS which requires another http connection and results in a long latency. Fortunately**, Bing Speech API** has a service that can help us with keeping the required HTTP connections at one. Therefore, we submit the user’s voice to Bing Speech API to ask it to convert the voice to text and query LUIS afterwards. Bing Speech API returns the text along with the extracted intent in its response.

By capturing the intent, our next step is to serve the user by executing the responsible intent handler and returning the result in a voice format.

Figure 2. The flow of a request. This diagram does not represent the tiers such as Ocularis and the Mediator.

## Intent Handling

By knowing the user’s intent, the next step is to act accordingly and compile the required information from all possible sources. Ocularis Intents are handled by Intent Handlers. In other words, Intent Handlers are at the helm of Ocularis and for each intent there is a corresponding handler which knows how to serve the user’s specific need.

Besides the parameter that we extract from the user voice, intent handlers are aware of the context that they are executing on. Ocularis State Manager is responsible for providing intent handlers with context related information.

If an intent is captured by SRGS, then it’s served offline by one of the Ocularis built-in intent handlers. In fact, questions about date, time, and cellphone related queries do not need to get run on the cloud and we can answer to them without making any http connection through Azure. The advantage of offline Intent Handlers is that they get executed on the edge, as a result they are fast, near real-time, and are not web dependent. However, the other side of coin is that we cannot add or modify intents and their corresponding handlers without installing the new version of Ocularis on the device. To alleviate this problem, we are considering the feasibility of using **Azure IoT Edge** in order to simplify the process of updating set of intents and intent handlers. Azure IoT Edge provides us with a convenient way to deploy modules to the edge devices.

However, there are several cases that we cannot serve a request offline and capture the user’s intent via SRGS grammars. These kinds of requests are delegated to the **Mediator**, which is a web application hosted by **Microsoft Azure Web Services**. Mediator is also infused with several kinds of Intent Handlers in the form of **Azure Functions**. These functions know how to deal with requests and serve them using Azure Cognitive Services. Mediator sends the response that user is looking for along with other context related information to Ocularis. The benefit of Mediator is that we can add new capabilities to our servers and eliminate bugs on it in a transparent way to millions of users because they are not needed to get a newer version of Ocularis in order to benefit from online Intent Handlers. Moreover, automatically we benefit from the ongoing progress of Azure Cognitive Services to improve the capabilities of Mediator in terms of accuracy and functionality without affecting Ocularis bits on edge devices. And lastly, Ocularis is immune to any deprecating changes to Azure Cognitive Services APIs because it’s all handled by Mediator. The downside of online intent handling is that active internet connection is required which results in unpredictable latencies.

## Intelligence

Intent handlers need to have access to several sources of information in order to provide the user with an appropriate and accurate answer. There are four offline sources of information which do not require an internet connection to operate, which are Internal Knowledge Base, Internal Image Processor for detecting depth and obstacles, user’s cellphone, and camera. Apart from the built-in offline knowledge sources, we also employ the massive amount of knowledge that’s easily accessible through Bing Services such as **Project Answer Search**, **Bing Visual Search**, **Bing Web Search**, and **Project Conversation Learner.** Moreover, we utilize **Azure Computer Vision** and **Face** services to provide the user with precious information about the current location and nearby people. Intent Handlers are coded in way to use available knowledge sources in order to satisfy user’s needs and questions.

## Hole & Obstacle Detection

Nowadays, there is a huge number of researches try to propose efficient solutions to variety of problems and challenges of autonomous vehicles and robotics. Based on our findings, different ways of detecting obstacles and holes have been proposed in last 15 years, some of them use sensors such as light radar and acoustic feedback [16,18] while some other new approaches benefit from processing images using Machine Learning and Deep Learning to detect holes and obstacles on the way [17,19,21,22,23].

We believe we can employ the already existing techniques in order to have a reliable detector which would be cost effective and accurate, especially our requirements are less strict compared to autonomous vehicles and robots. It is within our plans to also utilize an ultrasound detector that will help us to improve this detection mechanism and will increase the accuracy of our model.

## Globalization and Localization

In order to empower the visually impaired worldwide, we need to make Ocularis to communicate with as much as language that’s possible. However, there are some limitations imposed by LUIS because its only capable to analyze texts from English, Dutch, French, German, Italian, and Spanish []. To support these languages, we need to create a separate LUIS application for each of them. Also, Mediator must translate the responses by employing **Microsoft Translator Text** service before sending them back to the user.

LUIS aside, we can deliver offline functionalities worldwide because those are not dependent on Azure language-based services. Therefore, we can localize offline functionalities based on the region. However, we are aware it’s not a trivial mission to accomplish because we need to provide users with an extremely accurate quality to ensure them that Ocularis is a reliable friend.

## Core Technologies

**Windows IoT Core.** We decided to use Windows IoT core since it is free and its application development model boosts our productivity in comparison to other existing development models. Windows IoT fosters development by providing us with a rich set of features such as **Windows Machine Learning**, **Windows IoT Edge**, a fully compatible **Speech Synthesizer** with SSML, a fully compatible **Speech Recognizer** with SRGS grammars. Unlike its Linux counterparts, this operating system is designed for IoT projects from the ground up and it’s not a full blows operating system which we believe would result in more efficient power consumption. Lastly, since one of our future goals is to integrate **Cortana** into Ocularis, we believe Windows IoT is viable choice for our requirements.

**Azure Cognitive Services: Computer Vision**. This service enables Ocularis to describe the environment in front of the user and recognize celebrities in a picture. Furthermore, this service can extract text and handwriting from images so Ocularis can read them to the user.

**Azure Cognitive Services: Face**. This service can provide a detailed description about people in a picture. Face service provides us with gender, age, emotion, accessories and several other appearance attributes of detected faces in the picture. Moreover, it can recognize familiar faces and return their names. Also, it returns the exact location faces in picture and Ocularis leverages this information to associate a face description with its location in the picture. For example, it says “the man on the right seems angry but the one on left seems calm”.

**Azure Cognitive Services: Speech to Text**. If we cannot catch the user intent on the device, then we will forward the user’s request to Speech to Text service via Mediator. In addition to the main functionality, Bing Speech API can pass the extracted text to LUIS in order to detect the actual intent of the user. So, we can have both the text and the intent on one shot.

**Azure Cognitive Services: Text to Speech**. This forms a central pillar of Ocularis, as it facilitates the transcription of text into speech. This takes one of the most common challenges faced by the visually impaired and delegates the task at hand to a completely different sensory modality. Text to Speech is also used in tandem with several other of the features that Ocularis promises.

**Azure Cognitive Services: Text Analytics**. This service performs crucial semantic analysis of the intention of the user’s voice commands by extracting key words and true sentiment to give insights into the user’s desires. This forms another pillar of Ocularis as it will allow the rightful translation of services.

**Azure Cognitive Services: News Entity Search** This allows the user to ask Ocularis about the headlines pertaining to a topic of their choice. Users can ask to hear about a wide variety of topics, and News Entity Search will return the most relevant pieces. Otherwise, the user will be informed that no news was available regarding that search.

**Azure Cognitive Services: Project Answer Search.** Users can ask free form questions and Project Answer Search responds back with an answer. For example, they can like “what’s the height of Eiffel tower” to get “It’s 300 meters” hear back. If Project Answer Search can’t find an exact answer to your question, it will return a set of related articles from the internet and highlight their most relevant parts. So, Ocularis can read to user the exact answer or the related parts.

**Azure Cognitive Services: Bing Visual Search.** “Use smart identification of image content to recognize celebrities, find products, or search for related content. Bing Visual Search automatically identifies the type of picture, recognizing image context whether it is a barcode, a business card, or a product”.

**Azure Cognitive Services: LUIS.** This service enables our users to have more than one way of revealing their intents. As a result, our users can ask “Who is in front of me?”, “Tell me that who are these guys?”, or any other form of this question to make Ocularis describe the faces for them. In addition, LUIS can extract core parameters out of user’s question. For example, in question “what do you know about climate change?”, LUIS returns *Climate Change* as the parameter of this question. As a result, we can serve users by knowing their actual intents and their specific questions.

**Azure IoT Edge.** Unlike web application which have a straightforward updating and upgrading process, updating a software that is installed on millions of devices is delicate and difficult to handle properly. And since Ocularis is supposed to deal with humans, it’s accuracy and quality is our top priority and we need to have a way to fix bugs, add or modify intents and built-in ML models without introducing any instabilities to the device. Moreover, to publish a new version into devices we need a mechanism to filter target devices based on region, model, version, etc. because we do not push updates to all devices out there.

**Windows IoT Edge**, released on October 2018, could help us to deal with this challenge. It basically let you to push a block of code named Modules to edge devices seamlessly. As a result, we can update the software blocks if we design them as a Module. Although, Azure IoT edge can be a potential solution, it comes with some limitations that we need to consider. In Azure IoT Edge terminology, Modules are docker containers run by Azure IoT Edge Runtime on the device. We suspect that this architecture and tremendous flexibility comes more power consumption. Since Ocularis uses a battery, we need to do a research to find out how having IoT Edge Modules drains the battery. Additionally, it seems that IoT hub does not offer a way to filter target device for an update push. Lastly and most importantly, it’s only supported on Intel x64-based processors which can increase the final price of our device [https://docs.microsoft.com/en-ca/azure/iot-edge/how-to-install-iot-core].

**Windows Machine Learning.** This technology released on October 2018. It allows us to use pretrained models in our application. Windows Machine Learning inference engine needs models to be in ONNX format which is a universal standard format for storing trained models. The provided API by Windows IoT core makes developing ML based functionalities and their future updates a lot easier. Using this API, we just need to load the ONNX model into memory, bind our inputs to the model, and ask the engine to return the result. Furthermore, it enables us to use already published open source models [https://github.com/onnx/models].

**Web Application.** This will be the source of the Mediator, that will handle requests that cannot be served offline, and also provides a platform on which scalable changes can be made instantaneously without making the millions of users upgrade to a newer version of Ocularis. The web application will also be built to support the promotion and marketing of Ocularis, as well as to facilitate online orders, customer support, and issue submissions.

**Azure Functions.** As we discussed in previous sections, Mediator consists of several Intent Handlers to be able to serve users’ intents. Each handler knows how to gather information from available sources and how to process that information in order to return a reasonable response to the user. Since these handlers are fully isolated from each other and are supposed to get called by millions of potential customers, we decided to have them as an Azure Functions. Azure Functions do not have the complexity of microservices but has lots of its benefits, such as isolation, infinite scalability. Also, Azure Functions integration with Azure Service Bus helps us to trigger those functions in a stable and durable manner.

**Cortana and Bot Framework.** Future integration with Cortana is planned to make Ocularis a true part of the Windows experience. Integration with Cortana would also propagate further integration with other IoT devices running the same software, and make Ocularis even more successful.

# The Business Plan:

## Competitors

According to our research, Ocularis is the first full-fledged solution to the problems and challenges that the blind and visually impaired suffer from. Although there are some limited number of mobile applications that help users to describe things or read texts, none of them provide a unified solution to the variety of problems that these people deal with in their day to day life. As of now, users need to install a bunch of applications on their mobile devices to satisfy their needs and guarantee their independence.

Surprisingly, these applications assume that the users can easily open them and navigate through their pages and functionalities. As an example, they assume that a visually impaired user can easily take pictures of text or an object they are interested in. However in reality, it’s really not that easy to take a clear shot of an object or text because the user should manage to hold the phone with one hand and the object with the other one to provide the camera with an appropriate focal length and lumosity.

Based on our findings, our major competitor is an iPhone application developed by Microsoft called Seeing AI. Seeing AI, is supposed to enhance the life quality of its users by providing them with several functionalities such as detecting currencies, detecting barcodes, reading texts and handwritings, and lastly describing the environment and nearby faces. Although Seeing AI is considered as a viable solution, we believe that our proposed solution is revolutionary and brings a lot more value to its customers in terms of functionalities and its unique novel user experience.

Another worthy and potential competitor in the future is Google’s application known as Lookout. Lookout has 4 different modes to improve the quality of life. However, there is not much information available about it as it has not been released to the public as yet. Moreover, the application is built to be used with the phone put around the user’s neck as a lanyard, making it inconvenient and unattractive.

## Distinctive Characteristics of Ocularis

Ocularis empowers users with a distinctive set of features that other competitors are nowhere close to.

1. It’s all about the unique user experience that Ocularis delivers. Ocularis is supposed to be like a buddy who is there to support and guide its users. It is a new platform and isn’t comparable to any other existing solution. Ocularis is the third eye, as it’s placed as a necklace around its user. All communications with the device and the user go through a Bluetooth headphone and microphone. Therefore, there is no more challenge like unlocking the phone, finding the app and opening it up, and getting frustrated when getting the functionality that you are looking for is not very straightforward. With Ocularis, users can get what they are looking for just by asking for it.
2. It’s not always the user who initiates a conversation. Ocularis is always conscious and warns if it detects that the user is approaching an obstacle or when a new notification is raised by the user’s paired phone. Therefore, the user gets informed of high priority events that are happening around them even when they do not ask for them.
3. Although Ocularis is a stand-alone device, it can be paired to the user’s phone. So, users can make phone calls or send messages and emails just by asking Ocularis to do so. Moreover, Ocularis informs them about receiving calls, messages, and emails. They no longer need to be worried about their phone and its challenging user interface.
4. Ocularis makes it possible for users to ask for help from their friends without them being around. The **Remote Assistance** feature of Ocularis allows trusted friends to connect to the camera and see the world through its lens. Therefore, they can guide the user remotely.
5. According to our estimations, the cost of Ocularis is a lot lower than a high-end device like iPhone. Currently, Seeing AI is only available on iPhones which makes the solution less accessible due to the high cost associated.
6. As a unique and unified solution, we provide users with smart sticks. Ocularis sticks have been equipped with a vibrator and several buttons. The user can use these buttons to send some specific commands to the device. So, instead of saying loudly “who is this guy in front of me?” they can push a button on their sticks and nobody gets notified of the interaction between user and Ocularis. Moreover, when Ocularis detects a danger to the user, it can activate the vibrator in the stick.

## Partnership

### Microsoft

Partnership with Microsoft is essential for the success of Ocularis because our device is tied with Windows IoT core and some of its core services are supposed to provide by Azure Cognitive Services, Azure Web Service, and Azure IoT Edge. Furthermore, since we want to manufacture a high-quality device in large scales, we need an automated assembling production line which we believe Microsoft could help us with.

Our first challenge is regarding hardware requirements. Although, Windows IoT Core supports a vast range of hardware interfaces and protocols [], we need a close partnership with Microsoft to make sure that Ocularis Hardware requirements can get satisfied in terms of quality and power consumption when they are in an interaction with Windows IoT Core. We also need to ensure that if any of our chosen peripherals is not compatible with Windows IoT Core, that Microsoft can help us with a workaround or a long-term stable solution. Currently we are prototyping with a Raspberry Pi A+.

Secondly, since we are heavily dependent on Microsoft Azure Cognitive Services, we need to reach an agreement on a new specific subscription model for Azure Services to fulfill our demands. The reason behind that is our goal is to assure our users that they can use Ocularis without any limitations if they pay a fixed amount of money for their monthly subscription. To meet this requirement, we need an agreement that allows us to pay a fixed amount of money to Microsoft based on the number of users we have regardless of the number of queries they make. Therefore, our users won’t fear of getting extra charges when they ask for help from Ocularis because they are allowed to have limitless online interactions.

Lastly, our ultimate goal is to integrate **Cortana** into our device because it allows us to deliver lot more high-quality functionalities to our users. Besides that, Cortana itself has a powerful personalization system that can learns a lot about the user, communicate like a friend, and bring up highly personalized suggestions and information. Moreover, Cortana is highly extensible. In fact, we can introduce our desired functionalities to its **skillset** using **Microsoft Chat Bots**. Unfortunately, Cortana requires to be always connected to internet to be functional. To address this problem, we need a close partnership with Microsoft to make it able to respond to user requests that are supposed to serve without internet connection. As of now, this is the main blocking issue that impedes us from employing Cortana into Ocularis.

### Governments and Health Organizations and Associations

We need an effective partnership with organizations and associations supporting people with blindness and visual impairments. We can leverage our partnership with these government agencies and private social service agencies to inform our potential users about Ocularis. These agencies can hold Ocularis workshops to teach users how to get the most out of the device. Moreover, since these agencies have funds, they probably can afford a portion of costs that users have to pay for Ocularis in order to increase the quality of life for their users.

### Other Potential Partners

We believe we can empower our target users by giving them the opportunity to use popular services like Uber, Foodora, Audible, etc. To make our users even more independent and letting them enjoy as we enjoy the technology and life, we need a close partnership with those service providers. We need them to provide us with a set of APIs to their services. In addition, probably a change in their process flows and customer service manner is required which we can achieve by reaching an agreement between us and our potential partners.

## Revenue Model

Our revenue flow comes from four different sources.

1. **Selling brand-new devices**. These devices several have built-in functionalities like communicating with a paired phone or detecting obstacles and answering to variety types of questions that are answerable offline using the paired phone or the internal knowledge-base.
2. **Smart Sticks**. We can sell Ocularis sticks which can get paired with the device. Ocularis sticks have been equipped with a vibrator and several buttons. The user can use these buttons to send some specific commands to the device. So, instead of saying loudly “who is this guy in front of me?” they can push a button on their sticks and nobody gets notified of the interaction between user and Ocularis. Moreover, when Ocularis detects a danger to user, it can activate the vibrator in the stick.
3. **Selling Monthly subscription**. Users pay a fixed amount of money on a monthly basis and then will be able to have access to a tremendous amount of information available on the internet. Moreover, they can use services like Azure Computer Vision, Face, Bing, and services like Uber, Foodora, audible, etc. These services bring a set of unique values to their lives such as detecting faces, describing the environment, reading texts and handwritings, and answering their questions. Furthermore, this subscription allows users to use Remote Assistance to benefit from remote guidance from their friends.
4. **Selling peripherals**. Users can replace the camera, microphone, and headphones if they are not working.
5. **Selling new updates and devices**. As discussed before, we can push new updates to offline intents and ML models using Azure IoT Edge. We intend to sell major updates to users. Major updates are considered to have new functionalities or a whole different quality from the existing ones. Besides, selling new models of Ocularis is also another option for improving our money flow. New devices could have different capabilities such as lower consumption, better processing power, or accuracy.

We believe that our revenue model enables us to have a sustainable and profitable business. A simple calculation shows that if we reach 12 million customers which is near 5 percent of all our target market and assume that our gross profit per device is 40 dollars, then our total gross profit would be 480,000,000 dollars just from selling devices. Given half of these users are registered for a yearly subscription of online services then our total revenue from online subscription would be 720,000,000 dollars if an online subscription was 10 dollars per month.

**Nevertheless,** we understand that the above revenue model is solely based on theoretical sales, market share, and scaling. We understand that to ramp up production to this level would take a significant amount of time, money, partnerships, and risk. We recognize that several factors, including but not limited to, taxation, change in demand, increase in competition, shipping / warehouse costs, fluctuations in market share, worker’s compensation, and possible acquisitions and dilutions in stock, are among an array of things that can affect our revenue stream and profits in the future. It is difficult to make predictions with so many variables, and so the above revenue model and foreseeable profits should be understood to be purely based on theoretical calculation. As Ocularis grows and adapts, so does the business model and our team has a solid understanding of this.

# Risk and Legal Issues

Since Ocularis is a brand-new device and it is supposed to serve humans, there are several issues that we need to cover and think of. Ethical and legal challenges should be seen beforehand. Also, there are several technical and non-technical requirements that are critical for a having high-quality commercial device.

1. **Being Affordable**. As we mentioned, there are several other products out there try to help the blind with their daily life. Although these devices cannot bring the value comparable to what Ocularis delivers, cost still matters. According to National Federation of The Blind, most of our potential customers are considered as low income and the median of their annual earnings is 37,600 USD [https://nfb.org/blindness-statistics] [ https://nfb.org/images/nfb/publications/jbir/jbir15/jbir050202.html]. Our customers probably have paid for a cellphone and are paying to a telecom provider monthly. Therefore, we have to try very hard to make Ocularis and its monthly subscription affordable as much as possible. Besides, we must convince our target market that Ocularis costs are justifiable by the unique experience they can get from it.
2. **Offline Image Processing**. One our major challenges is to implement a high-quality image processor that can work efficiently with limited hardware resources that Ocularis has. So, we want our users get a near real-time feedback based on their surroundings without being worried about the battery status of their device. In order to use the battery efficiently, Image processing is turned off by default. It starts when the user asks for it. Moreover, the camera captures a picture every 5 seconds by default. If the user needs a higher rate of capturing, it can ask for it at the cost of higher battery consumption. Moreover, we need to come up with a Machine Learning technique that can assure us the trained model would be reliable enough in detecting obstacles, depth, and distances. Since our product is dealing with humans, it is critical for the system to comply with major existing standards in order to prove that users can have reliable device in their hands.
3. **High-Quality Device**. There is always a tradeoff between affordability and quality. Since Ocularis doesn’t need a set of high-end hardware to function properly, we believe with the current technological advancements, a decent quality can be achieved. We believe our main challenge is to make sure the designed batteries can last long enough and are lightweight at the same time. We do not want our customers face any problem because of low battery capacity or high-power consumption when they are out of their home. Also, we do not want our customers feel the weight of Ocularis on their head. Ocularis should be designed as a third eye of its customers. We want our customers to be reluctant to detach Ocularis.
4. **Partnership**. We believe all the aforementioned partnerships are essential to success of our product. Without having Microsoft as our partner, we won’t be able to deliver affordable devices and services. In turn, Microsoft can improve the intelligence of its services by analyzing millions of queries and questions that our users ask daily. Moreover, our partners will earn a lot of credit from society which has an indirect impact on the brand and revenues. In addition, our partnership with governments and related agencies is necessary because they are our main channel to our potential customers. Though, reaching a partnership with governmental organizations usually takes a lot of time to deal with their bureaucracy and it can be understood as to why they may not see any direct or indirect financial advantage to helping us.
5. **Globalization**. To reach our market as much as possible, we need to localize our device according to specific requirements of each region. We are dependent on Microsoft products such as Windows IoT, Cortana, and Azure Translator in order to support every potential customer regardless of their preferred language. Hopefully, recent Microsoft’s advances in English to Chinese translation and vice versa [https://blogs.microsoft.com/ai/machine-translation-news-test-set-human-parity/] can make Ocularis available to a country with 18% of all the visually impaired [https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6093864/].
6. **Weather**. Sunny, foggy, snowy

# Additional Information:

Add any links or information here. Any information you want to share is welcomed!

We do have some pictures to give you a better sense of our product. dWe have designed a 3-D printed prototype of what the design is going to look like.



Figure 2: Shaaran wearing Ocularis as a necklace

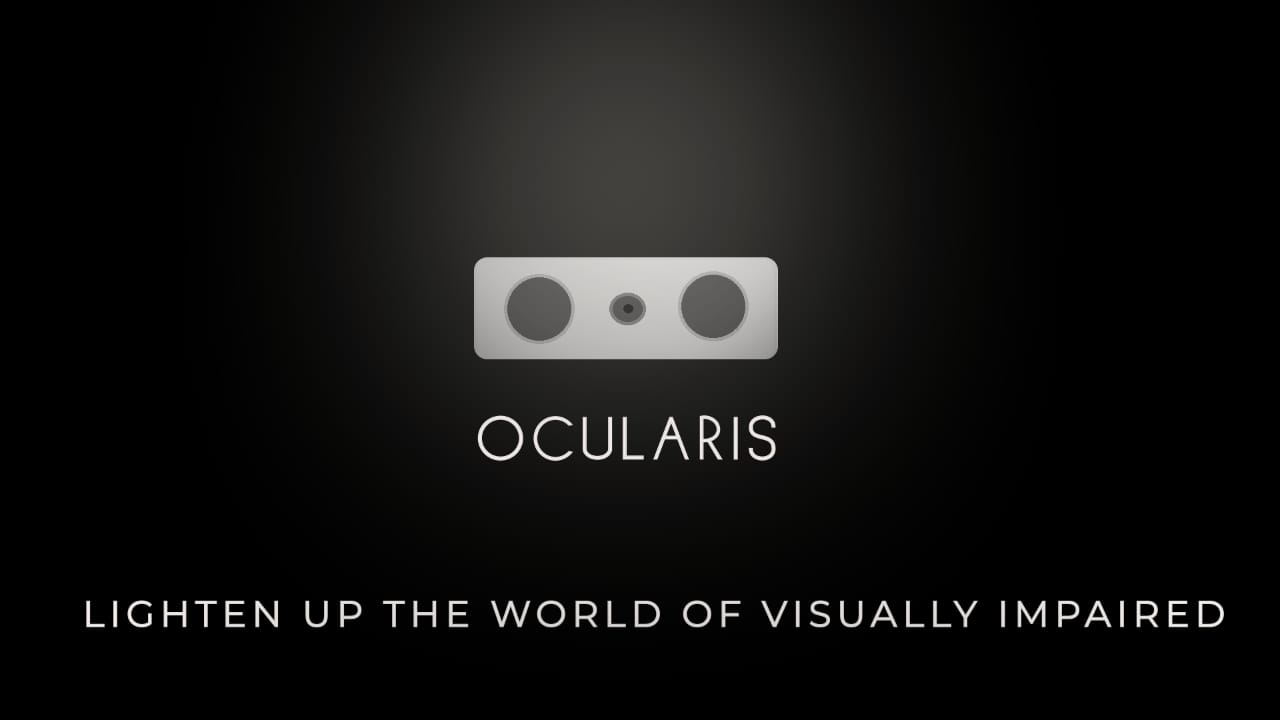


Figure 3: Our logo and motto

Figure 1: 3-D printed model of Ocularis with laptop in back for scale.

# References:

[1] <https://en.wikipedia.org/wiki/Visual_impairment>  
[2] <http://pertuni-wbu.blogspot.com/2012/04/east-wind-no-10.html>  
[3] <http://www.visionaware.org/info/your-eye-condition/eye-health/low-vision/low-vision-terms-and-descriptions/1235>  
[4] <http://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment>  
[5] <https://www.bbc.com/news/health-40806253>  
[6] <http://atlas.iapb.org/news/latest-global-blindness-vi-prevalence-figures-published-lancet>  
[7] <http://www.euroblind.org/about-blindness-and-partial-sight/facts-and-figures#details>  
[8] <https://help.rnib.org.uk/help/newly-diagnosed-registration/registering-sight-loss/statistics>  
[9] <http://www.afb.org/info/blindness-statistics/adults/facts-and-figures/235>  
[10] <https://www150.statcan.gc.ca/n1/pub/89-628-x/2009013/fs-fi/fs-fi-eng.htm>  
[11] <https://cnib.ca/en/sight-loss-info/blindness/blindness-canada>  
[12] <http://www.mdpi.com/1660-4601/14/9/1034/pdf>  
[13] <https://lists.w3.org/Archives/Public/w3c-wai-eo/2001AprJun/0166.html>  
[14] <https://www.lasereyesurgeryhub.co.uk/blindness-statistics>  
[15] <http://www.vision2020australia.org.au/resources/a-snapshot-of-blindness-and-low-vision-services-in-australia>  
[16] Assisting the Visually Impaired: Obstacle Detection and Warning System by Acoustic Feedback,2012  
[17] Deep Learning for Obstacle Detection and Classification in Autonomous Vechicles, 2018  
[18] Obstacle Recognition Based on Machine Learning for On-Chip LiDAR Sensors in a Cyber-Physical System, 2017  
[19] MACHINE LEARNING BASED APPROACH FOR POTHOLE DETECTION, 2018  
[20] High speed obstacle avoidance using monocular vision and reinforcement learning, 2005  
[21] Unsupervised obstacle detection in driving environments using deep-learning-based stereovision, 2016  
[22] Enabling independent navigation for visually impaired people through a wearable vision-based feedback system  
[23] https://medium.com/@percepsense/intelligent-pothole-detection-879ef635dd38