

# CSE299

## PROJECT PROPOSAL

### GROUP MEMBERS

Adib Ar Rahman Khan

2212708042

Saumik Saha Kabbya

2211204042

---

# THE PROBLEM

---

The visually impaired, (whose number exceed 295 million people) deal with a myriad of problems each day. One of the major issues that they tend to face is inaccessibility to everyday technology, having to deal with niche products and services (which are often paid and exploitative) to do very simple tasks. Although good free services exist, each functionality is scattered over different apps which prove to be a monumental task to remember and deal with.

Our project aims to eliminate all these problems by providing an all-in-one solution in one app, that is easy to use, and always free.

---

---

# INTRODUCTION

---

The idea is to create an application that will empower patients with Blindness or MSVI (Moderate Severe Visual Impairment) to do everyday tasks with efficiency and ease, via a suite of accessibility tools and features.

Through innovative technologies like OCR (Optical Character Recognition) to generate text-to-speech, a voice-controlled calendar and reminder system, real-time navigation for transportation, and nutritional analysis using object detection with audio feedback, this project will enhance our user's independence and daily functionality. The app will aim to seamlessly integrate existing APIs and frameworks to deliver fast, reliable, and user-friendly assistance.

# LITERATURE REVIEW

Before we begin, it is important to mention and credit noteworthy contributions done in this field, via mobile applications or services, that inspired our project.

---

## SEEING AI (iOS):



A free app from Microsoft designed for the visually impaired, Seeing AI uses AI-driven image processing to read text aloud, describe objects, recognize people, and provide environmental context. The app supports multiple modes, including text recognition, barcode scanning, and scene description.

---

## LOOKOUT (Android)



Developed by Google, Lookout uses the camera to detect objects, texts, and labels, providing spoken feedback in real-time. It is designed for everyday use, including scanning documents, identifying products via barcodes, and reading text aloud for better accessibility.

## BARD Mobile (Cross-Platform)



Created by the National Library Service for the Blind, BARD Mobile provides access to a vast library of audiobooks, magazines, and music scores. It supports both Android and iOS, enabling users to download and listen to accessible content in a variety of formats.

---

## Voice Dream Scanner & Reader (iOS)



Voice Dream is an OCR-based app that allows visually impaired users to scan text from images and documents, instantly converting it to speech. It also integrates with its combination app, Voice Dream Reader to support more robust accessibility for reading books, articles, and PDFs aloud.

---

## Legere Scanner & Reader (Android)



The Legere Scanner app offers OCR text recognition to convert printed documents into spoken text, supporting a variety of languages. It also features a read-aloud function, enabling easy access to written content for visually impaired individuals.

## VO Calendar (iOS)



Designed for individuals with low vision, VO Calendar offers a voice-guided calendar experience. Users can add, edit, and review appointments using voice commands, simplifying the process of managing schedules without needing visual input.

---

## Blind Square (iOS)



Blind Square combines GPS technology and crowdsourced data to provide real-time navigation tailored to visually impaired users. It offers audio guidance for navigating city streets, public transportation, and specific locations with rich details on surroundings and points of interest.

---

## XploreNinja by BlindSquare (Android)



XploreNinja extends BlindSquare's functionality to Android users, offering location-based auditory feedback. It guides users through real-time environments, identifying landmarks, crosswalks, and transportation stops, making urban navigation accessible.

## NavCog (iOS)

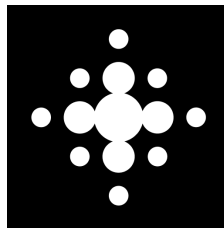


**NavCog**

NavCog is a navigation app developed by Carnegie Mellon University that uses Bluetooth beacons and AI to assist visually impaired users in navigating indoor and outdoor spaces with precision. It offers turn-by-turn guidance using audio cues for smoother transportation access.

---

## Envision App (Cross Platform)



The Envision app provides object detection and text recognition using a phone's camera. It reads food labels, identifies objects, and offers real-time audio feedback, making nutritional information and daily objects accessible to visually impaired users.

---

## TapTapSee (Cross Platform)



TapTapSee is an object identification app that uses a smartphone camera to take pictures of objects, returning an audio description. It excels in detecting everyday items, food packaging, and even recognizing barcodes to assist users in identifying products.

## Be My Eyes (Cross Platform)



One of the most innovative projects we came across, Be My Eyes connects visually impaired users with sighted volunteers via live video calls. Volunteers assist users by describing objects, reading labels, and providing real-time feedback on visual tasks. The service enhances independence for tasks that require real-time human input.

---

# PLANNED FEATURES

---

For our set of planned features, we set out a goal to empower the visually impaired in various aspects in their daily lives. This was done in various ways by different services, but the way they were scattered across different apps made them terribly inconvenient. With our research and study, we landed on four main aspects of a user's daily life that we wanted to improve upon.



## EDUCATION

Allowing users to read any text available is one of the first concerns we intended to work on. The idea will be to allow the users to do something as simple as point their phone camera at any text, and via OCR (Optical Character Recognition), get audible output.



## PRODUCTIVITY

Essential tasks should not be the only way to empower our users. We want to allow our users to be productive, plan out their day, and be mindful of their daily activities and have access to time management. Thus, we plan on integrating the phone's native calendar to create, update, retrieve events, provide alerts, along with setting reminders, all via voice commands.





## NAVIGATION

No impairment or disability should keep someone confined to the space of their home. To allow our users to navigate the world outside, we intend to provide real-time vocal directions for navigation, along with the ability to detect obstacles and provide alerts if the camera is held up in real time.

---



## NUTRITION

We want to enable our users to extend the capabilities of computer vision and use advanced object detection models to be able to analyze and name food items, read nutritional labels, and provide aural nutritional data for all these cases to enable them to be sure of what they are eating, and keep track of their nutrition.

# TECHNOLOGIES

Based on our belief that accessibility should not be limited to a specific group of people, we want to create an application that is cross-platform.



Thus, we have chosen the Flutter Development kit to bring our vision to life. The technologies that we plan to use for each feature has been listed below, with relevant links. This may be subject to change based on our need and ease of use.

EDCUATION	PRODUCTIVITY	NAVIGATION	NUTRITION
Firebase ML Vision (Image Recognition) [ <a href="https://firebase.google.com/products/ml">https://firebase.google.com/products/ml</a> ]	device_calendar (Event Handling) [ <a href="https://pub.dev/packages/device_calendar">https://pub.dev/packages/device_calendar</a> ]	Google Maps API (Real-time Directions) [ <a href="https://developers.google.com/maps">https://developers.google.com/maps</a> ]	TensorFlow Lite (Object Detection) [ <a href="https://www.tensorflow.org/lite">https://www.tensorflow.org/lite</a> ]
Google Tesseract OCR (Text Recognition) [ <a href="https://github.com/tesseract-ocr/tesseract">https://github.com/tesseract-ocr/tesseract</a> ]	speech_to_text (Voice Recognition) [ <a href="https://pub.dev/packages/speech_to_text">https://pub.dev/packages/speech_to_text</a> ]	Mapbox SDK (Real-time Directions) [ <a href="https://docs.mapbox.com/flutter/">https://docs.mapbox.com/flutter/</a> ]	Firebase ML Kit (Object Detection) [ <a href="https://firebase.google.com/products/ml">https://firebase.google.com/products/ml</a> ]
flutter_tts (Text-to-Speech Conversion) [ <a href="https://pub.dev/packages/flutter_tts">https://pub.dev/packages/flutter_tts</a> ]	flutter_local_notifications (Reminders and Notifications) [ <a href="https://pub.dev/packages/flutter_local_notifications">https://pub.dev/packages/flutter_local_notifications</a> ]	flutter_tts (Voice Guidance) [ <a href="https://pub.dev/packages/flutter_tts">https://pub.dev/packages/flutter_tts</a> ]	Edamam API (Nutritional Data) [ <a href="https://developer.edamam.com">https://developer.edamam.com</a> ]
		flutter_vibration (Haptic Feedback) [ <a href="https://pub.dev/packages/flutter_vibration">https://pub.dev/packages/flutter_vibration</a> ]	Nutritionix API (Nutritional Data) [ <a href="https://www.nutritionix.com/business/api">https://www.nutritionix.com/business/api</a> ]
			flutter_tts (Audio Feedback) [ <a href="https://pub.dev/packages/flutter_tts">https://pub.dev/packages/flutter_tts</a> ]

## CONCLUSION

We hope that our ambitious project can be completed within the time-frame provided, and that it can truly have a positive impact for the users that we intend to create it for. Our goal is to create a seamless experience that we hope will improve the lives of over hundreds of millions of people. We can't wait to get started!