MOBILE BASED SINHALA BOOK READER FOR VISUALLY IMAPAIRED INDIVIDUALS – VOICE NAVIGATION AND OBJET/OBSTACLES DETECTION.

2023-198

Project Proposal Report

Godakanda P. G. S.

B.Sc. (Hons) Degree in Information Technology

(Specializing in Information Technology)

Department of Information Technology

Sri Lanka Institute of Information Technology

Sri Lanka

March 2023

MOBILE BASED SINHALA BOOK READER FOR VISUALLY IMAPAIRED INDIVIDUALS – VOICE NAVIGATION AND OBJET/OBSTACLES DETECTION.

2023-198

Project Proposal Report

Supervisor – Prof. Koliya Pulasinghe.

Co-Supervisor – Ms. Poorna Panduwawala

B.Sc. (Hons) Degree in Information Technology

(Specializing in Information Technology)

Department of Information Technology

Sri Lanka Institute of Information Technology

Sri Lanka

March 2023

DECLARATION

DECLARATION

We declare that this is our own work, and this proposal does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other university or Institute of higher learning and to the best of our knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Name	Student ID	Signature
Godakanda P. G. S.	IT20129712	

The above candidates are carrying out research for the undergraduate Dissertation under my supervision.

Signature of the supervisor:

04/05/2023
Date:

ABSTRACT

The purpose of this research project component is to develop a system that enables blind people to navigate the book reading app using Sinhala voice commands and distant object detection. The system uses machine learning to instruct the user to avoid objects and provides the distance to the object through a voice command. This project aims to improve the independence and mobility of blind people in a digital environment, enhancing their experience in using technology.

Navigating without sight is a daunting task for the visually impaired. Unlike those with good vision, blind individuals cannot detect and avoid obstacles, making it necessary for them to rely on guidance to navigate safely. While the white cane is a commonly used tool for detecting and avoiding obstacles, its limited reachability makes it inadequate for identifying all potential threats. To address this issue and enable safer and more independent navigation for the blind, this study proposes a novel approach based on deep learning for obstacle detection.

The study involved developing a prototype system that utilizes deep neural networks (DNNs) for obstacle detection and distance estimation. DNNs were selected due to their high accuracy and real-time performance. Obstacle detection data was collected using a simulation environment, and the resulting model was used to estimate the distance of obstacles. The feedback generated by the combination of obstacle detection and distance estimation was communicated to the user through audio cues.

Overall, this novel approach based on deep learning offers a promising solution to the challenges of blind navigation, enabling safer and more independent movement for visually impaired individuals.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

Abbreviation	Description
API	Application Programming Interface
AR	Augmented Reality
NFC	Near Field Communication
VIP	Visually Impaired People
AT	Assistive Technology
AP	Average Precision
CNN	Convolutional Neural Network
CPU	Central Processing Unit
DNN	Deep Neural Network
HTTP	Hyper Text Transfer Protocol
mAP	Mean Average Precision
SLAM	Simultaneous Localization and Mapping
SSD	Single Shot MultiBox Detector
YOLO	You Only Look Once
AP	Average Precision
CNN	Convolutional Neural Network

LIST OF APPENDICES

1. INTRODUCTION

Mobile apps have seen a surge in popularity, with more and more mobile phone users relying on them for their day-to-day needs. However, despite the widespread availability of developer tools and public demand driving expansion across all aspects of human life, there remains a significant gap in the market for differently abled people.

While most voice recognition and other software focuses on individuals with the ability to see, there is a lack of accessibility for those with disabilities who wish to carry out their daily tasks just like any other person. Many apps are designed with graphics-oriented activities for children, and the majority are tailored towards the needs of able-bodied individuals, leaving those with disabilities at a disadvantage.

However, according to the United Nations, differently abled individuals are entitled to the same human rights and freedoms as everyone else in society, without discrimination of any kind. They have the right to access things and places and live independently, without exception or distinction. Currently, most of the apps available on mobile phones are in English, leaving a significant portion of the population in countries such as Sri Lanka unable to access them due to language barriers.

To ensure that everyone has equal access to these apps, it is crucial that they are also available in local languages such as Sinhala and Tamil, in addition to English. By prioritizing accessibility and inclusivity, we can create a world where all individuals, regardless of their abilities, can fully participate and enjoy their fundamental rights and freedoms.

1. LITERATURE REVIEW

1.1 Background

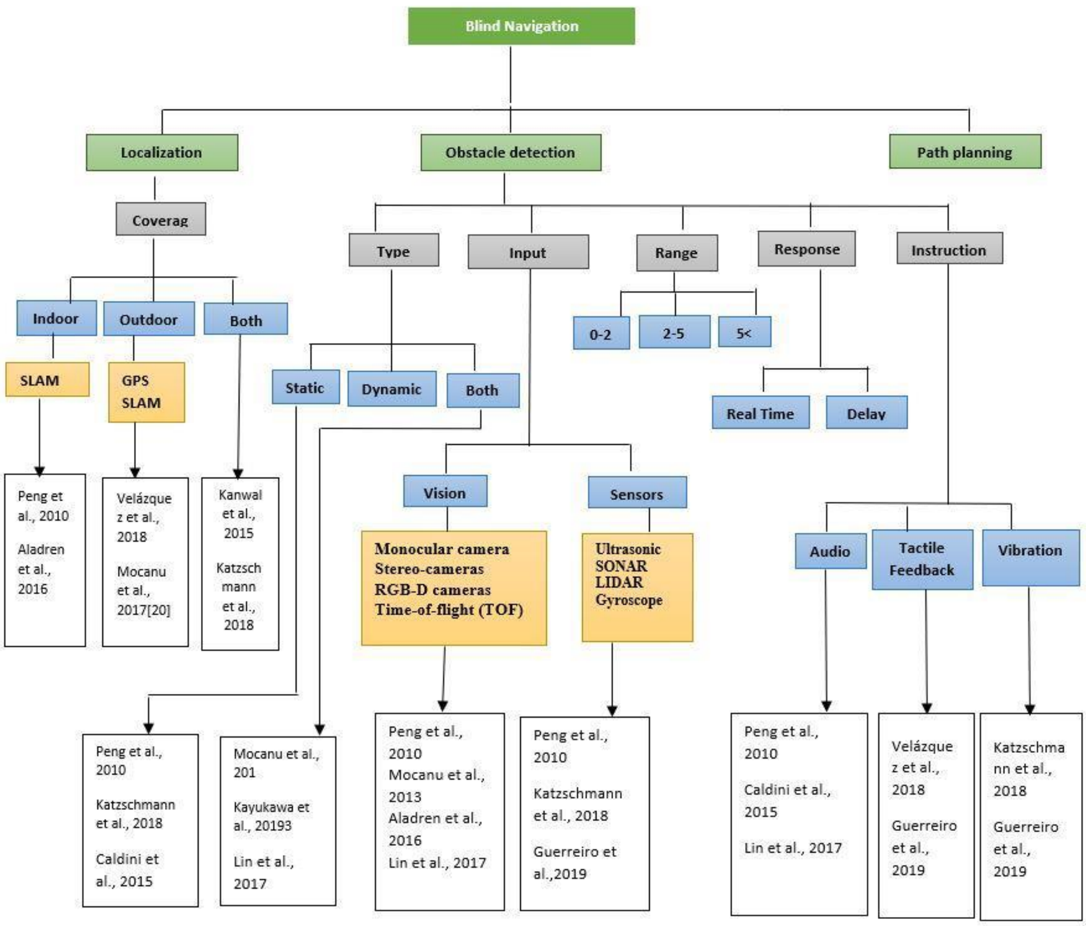
In recent times, voice assistants (VAs) such as Siri, Google Assistant, Microsoft Cortana, Samsung Bixby, and Amazon Alexa have gained significant popularity and usage. With technological advancements, the research field of Automated Speech Recognition (ASR) has also piqued the interest of researchers. However, despite the progress made in the field, the development and accuracy of ASR remains a significant research challenge. Thanks to advancements in voice recognition and natural language processing, speech recognition has become easier to integrate into our everyday devices such as smartphones, smartwatches, smart home devices, and smart speakers. This has enabled voice assistants to become more accessible and user-friendly than ever before.

Figure 1: Examples for Voice Assistants

1.2 Literature survey

For years many researchers have focused on improving the safe navigation of the visually impaired in outdoor environments. A navigation system consists of three parts. Such a navigation system includes localization, providing information about the current location and orientation, and pathfinding. For the localization purpose, a particular system must extract the details and features of the current surrounding of the blind person. Global Coordinates (GPS), Local coordinates, environmental features, and dead reckoning are some of the feature identification methods for localization. Providing information about the current location includes static and dynamic obstacle detection and correct orientation. Pathfinding or wayfinding is finding the optimal or best route to the destination.

Many researchers have focused on the area of blind navigation from earlier years and many technologies have been adapted to develop assistive devices for blind navigation over the years. Figure 2.1 provides a classification of blind navigation systems and research carried out under different aspects.



A detection system for obstacles on the ground at any height using a handheld smartphone is proposed in Peng et al. [2]. This proposed method uses computer vision techniques such as color histograms and edge detection for obstacle identification and obstacles are captured through the smartphone camera. A color histogram is created based on the images acquired from smartphones and histograms are created for each frame and chooses the simplest RGB color space. Then, a binary RGB histogram is built for the safe region which is derived from the image region.

In Caldini et al. [3], a calibrated smartphone on users' chest which installed a gyroscope is used to detect obstacles. It implements a modified Structure from Motion (SfM) algorithm for scene reconstruction. Visual data obtained from the camera and measurements obtained from the gyroscope are used in the developed algorithm. The proposed vision-based system focuses on obstacle detection to help visually impaired people to move autonomously in unknown indoor and outdoor environments.

In 2020, a study was conducted to establish a revolutionary desktop application for speech disfluencies by monitoring the electroencephalogram feed of the speech motor (Broca's area) and evaluating it to detect speech impairments using brain neuron electrodes. Through EEG neurofeedback, the system was developed to determine the influence on the left hemisphere of the brain. The system incorporates self-learning therapy-based exercises that help with fluency improvement, with results sent to the speech pathologist straight away [10].

The majority of these apps are free and may be downloaded to a smartphone, tablet, or computer. They generally use a Bluetooth headset, which may cause the feedback to be even more delayed. There is a list of apps that use AAF or DAF.

Festival-si

This research brings together the development of the first Text-to-Speech (TTS) system for Sinhala using the Festival framework and practical applications of it. The methods for developing direct Sinhala Unicode text input by rewriting Letter-to-Sound rules in Festival's context-sensitive rule format and the implementation of the Sinhala syllabification algorithm are also presented in the article. The implementation and evaluation of a diphone concatenation-based Sinhala text-to-speech system. The Festival framework was chosen for the Sinhala TTS implementation. The Festival Speech Synthesis System is an open-source, reliable, and portable multilingual speech synthesis system created by the University of Edinburgh's Center for Speech Technology Research (CSTR). [17].

Figure 2: Stamurai Mobile Application

Figure 3 Fluency Coach Mobile Application

2. RESEARCH GAP

While there have been some attempts to develop assistive technologies for visually impaired individuals, there is still a significant research gap in the area of navigating mobile apps using voice commands and detecting obstacles in real-time. Most current assistive technologies rely on screen readers or other assistive technologies that do not provide real-time feedback about obstacles in the user's path. This research project aims to bridge this gap by developing a system that provides real-time feedback to users about obstacles in their path, and enables them to navigate an app using voice commands. “VisionFree” is a Sinhala language-based application which will provide Sinhala language voice recognition to carry out the entire procedure. Moreover, this platform will be a child-friendly platform that will be developed using Natural Language Processing and Machine Learning mechanisms.

3. RESEARCH PROBLEM

Blind people face several challenges when reading books, but the main problem is a lack of accessibility to printed materials. Despite advancements in assistive technology, such as text-to-speech software and Braille displays, most books are still not accessible to blind individuals in an easily readable format. This can limit the opportunities for blind people to gain knowledge, engage in literary experiences, and improve their education and employment prospects.

One issue is the cost of specialized devices and software, which can be prohibitively expensive for many blind people. Even when these tools are available, they may not provide an experience that is comparable to reading a traditional printed book. For example, text-to-speech software can struggle with complex language and formatting, and Braille displays can be slow and clunky.

Another issue is the limited availability of audiobooks and Braille materials. While more audiobooks are being produced, the selection is still limited compared to the vast number of printed books. Braille books are even harder to come by, as the process of translating printed books into Braille is time-consuming and costly. This means that blind people may not have access to the latest best-selling books or popular educational materials.

In conclusion, the main problem that blind people face in reading books is a lack of accessibility to printed materials. Despite advances in assistive technology, there are still significant barriers to overcome, such as the cost of specialized devices and software, the limited availability of audiobooks and Braille materials, and the difficulty in providing a comparable reading experience to that of a traditional printed book. To address these challenges, there needs to be a concerted effort to make books more accessible to blind people and to ensure they have the same opportunities to engage with literature and gain knowledge as sighted individuals.

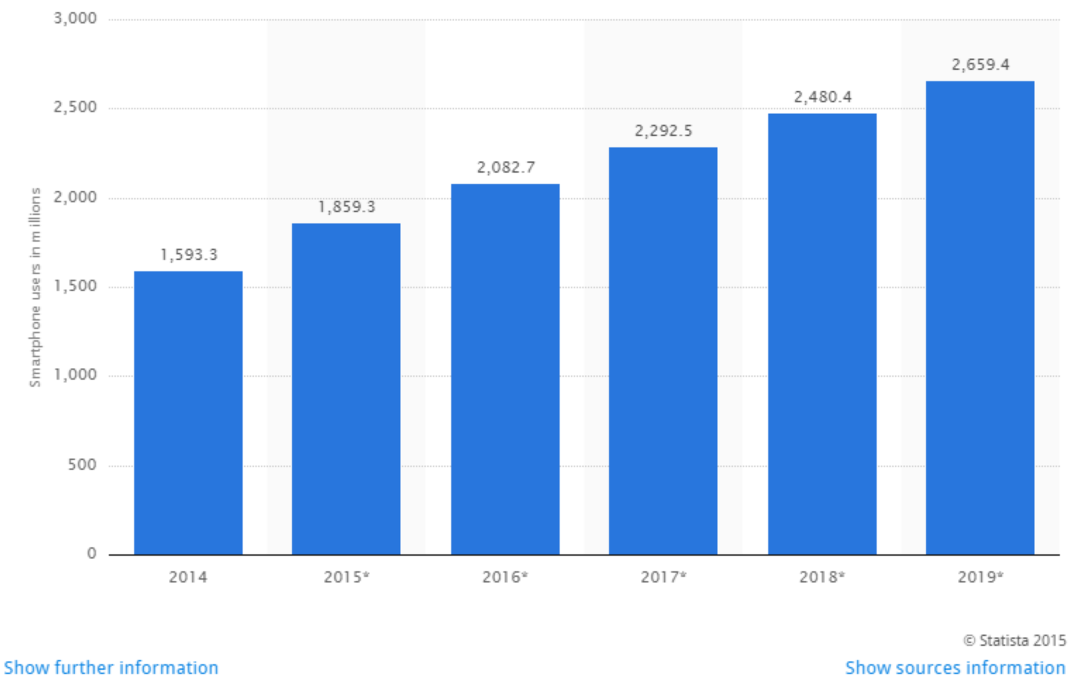


Figure 4:Smart phone user in globally

4. OBJECTIVES

4.1 Main objectives

A Sinhala book reader for the visually impaired is a software program designed to make reading accessible for individuals with visual impairments. It combines various technologies to provide a seamless reading experience. The device utilizes optical character recognition (OCR) technology to convert the text from a physical book into a digital format. Then, a text-to-speech synthesizer reads the text out loud in the Sinhala language, making it easier for visually impaired users to follow along.

In addition to the text-to-speech synthesizer, the device also includes audible guidance to help navigate the app and identify the distance to the book being read. This makes it easier for visually impaired users to find their place in the book and keep track of their progress. The device also allows users to adjust the reading speed, volume, and pitch of the audio to their liking.

Another useful feature of this Sinhala book reader is the ability to record audio pitch, which is especially helpful for users who may have difficulty reading in a consistent tone. This feature helps users to improve their reading skills and become more confident in their abilities.

4.2 Specific objectives

In order to fulfill the main objective, several sub objectives are required to be completed.

Sub Objective :

The objectives of this component are as follows:

- To develop a system that enables blind people to navigate an app using voice commands
- To develop machine learning algorithms that can detect objects in real-time and calculate distances accurately

• To provide users with real-time feedback about obstacles in their path and instructions on how to avoid them
• To develop a user interface that is easy to use and understand for blind users

METHODOLOGY

In order to get the final product of "VisionFree", there are several phases to go through. Environments, methodologies, requirements, tools, and technologies that will be needed to meet the objectives will be described in this section. The procedures that must be taken in order to complete the research will be explained here.

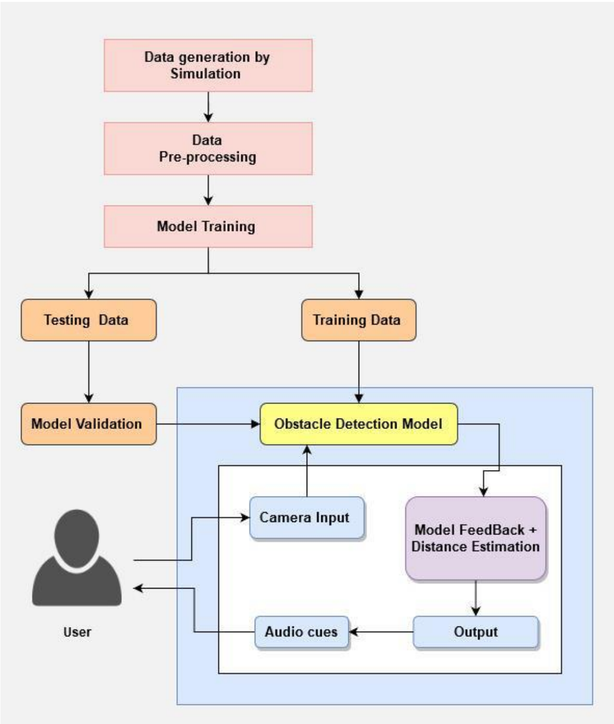


Figure 5: Phases of the methodology

The initial step of the methodology is data collection. A group of children who have stuttering disfluency and aged between 3 – 14 years is used to collect the data. The necessary data will be gathered by having a simple conversation with them or giving a small assessment like a picture description in order to capture their speech patterns. If the simple conversation approach is used to gather the data children are asked some of the basic question which is about self-introduction. For example, their names, favorite color, school, etc. And if the second approach, a simple picture description method is used, children are given a picture and told them to describe that picture. Every child in the group is given the same task. While they are doing these activities their voice is recorded covering the entire conversation. With the necessary modifications, a Sinhala voice

recognition API detects the captured voice, which is then translated to text using a voice to text algorithm.

Then in the next phase the model will be implemented to identify the stuttering stye as Block Stuttering by analyzing the speech patterns captured. In order to identify the stuttering type, a basic set of activities are given to the child. For example, a picture description activity, or a set of phrases or sentences to read.

After completing the model building, the next phase is to select relevant activities to the identified stuttering type and stage. And the selected activity series will be provided in a gamified manner which will be a level-based activity series. The UI designing of the mobile application is done in a manner to get the attraction of the children and keep them engaged with the therapeutic activities throughout the session without distraction. Moreover, the avatar "VisionFree" is designed in a manner where child is able to converse with the assistive bot. Then the UI is integrated with the implemented individual component and tested. After completing the testing phase of the individual component, all the four components are integrated as a full system. As the final result a prototype of "VisionFree" will be launched.

4.3 Tools and technologies

These are the tools and technologies which will be used to move forward with the proposed solution's implementation.

Table 1: Technologies use to implement the proposed system

Purpose	**Technology**
Model Building	Python 3.x (Django Framework)
Data Storing	Cloud base Database
Mobile App Development	Flutter 2.x
Voice to Text Convertion	Python(CMU Sphinx Library)
IDEs	Visual Studio Code 1.x Android Studio 3.x
Version Control System	Git Lab

5. SYSTEM REQUIREMENTS

When implementing a technological application, there are mainly three types of requirements to consider. Namely, they are functional requirements, non-functional requirements, and user requirements. Basic knowledge about using mobile applications is a prerequisite to using this “VisionFree”.

Table 2: Requirements Categorization

Functional Requirements	• Provide relevant visualizations. • Provide precise predictions. • A model to identify type and stage accurately. • Voice to text conversion • Accurate voice recognition
Non-Functional Requirements	• Easy usability • Child Friendly • Simple UI/UX • High performance • Self-manageable
User Requirements	• Caregiver's assistance in creating a user profile. • Age requirement (Between 3 – 14 years) • Compatible device to facilitate “VisionFree”

6. BUDGET AND BUDGET JUSTIFICATION

Table 3: Budget Estimation for the proposed solution

Component	**Amount (Rs.)**
Internet	3000.00
Stationary	2000.00
Server Cost	4000.00
Electricity	1000.00
Travelling	2500.00
APIs	5000.00
Other	1000.00
Total	18500.00

REFERENCE LIST

[1] N. D. K.N.Withanage, "Towards ICT based Solution for Stuttering," 08 2017. [Online]. Available: https://www.researchgate.net/publication/319058777_Towards_ICT_based_Solution_for_Stuttering. [Accessed 21 01 2022].

[2] f. e. staff, "Stuttering," 02 2021. [Online]. Available: https://familydoctor.org/condition/stuttering/. [Accessed 21 01 2022].

[3] G. Williamson, "The Nature of Stuttering," 30 01 2014. [Online]. Available: https://www.sltinfo.com/the-nature-of-stuttering/. [Accessed 22 01 2022].

[4] "Stuttering," [Online]. Available: https://en.wikipedia.org/wiki/Stuttering. [Accessed 21 01 2022].

[5] "Stuttering Modification Therapy," [Online]. Available: https://speecheasy.com/stuttering-modification-therapy/#:~:text=Stuttering%20modification%20therapy%20is%20a,be%20removed%20from%20speaking%20situations. [Accessed 22 01 2022].

[6] M. M. J. H. S. Hector R. Perez, "Stuttering," 06 2016. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4907555/. [Accessed 22 01 2022].

[7] "Technologies for Stuttering," [Online]. Available: https://casafuturatech.com/daf-devices/. [Accessed 31 01 2022].

[8] S. A. Ling Qiu, "Voice Assistants for Speech Therapy," [Online]. Available: https://ubicomp-mental-health.github.io/papers/2021/ubicomp21g-sub1960-cam-i7.pdf. [Accessed 01 02 2022].

[9] D. E. a. K. Dalton, "SPEECHEASY," [Online]. Available: https://www.mnsu.edu/comdis/kuster/journal/hallen/speecheasy.html. [Accessed 01 02 2022].

[10] J. A. D. T. Jayawickrama and S. Thelijjagoda,

"A solution to overcome speech disorder of patients using Brain Neuron EEG Signals," 12 01 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9312997. [Accessed 31 01 2022].

[11] "FleuncyCoach: Free stammering management software," [Online]. Available: https://www.fluencycoach.com/. [Accessed 02 02 2022].

[12] J. M. D. M. C. E. G. E. C. Anne Foundas, "The SpeechEasy device in stuttering and nonstuttering adults: Fluency effects while speaking and reading," 05 2013. [Online]. Available: https://www.researchgate.net/publication/236948565_The_SpeechEasy_device_in_stuttering_and_nonstuttering_adults_Fluency_effects_while_speaking_and_reading. [Accessed 01 02 2022].

[13] "Stamurai," [Online]. Available: https://stamurai.com/. [Accessed 02 02 2022].

[14] "Innovation makes speech therapy a fun activity for children," [Online]. Available: https://myspeechbuddy.webflow.io/#Research. [Accessed 07 02 2022].

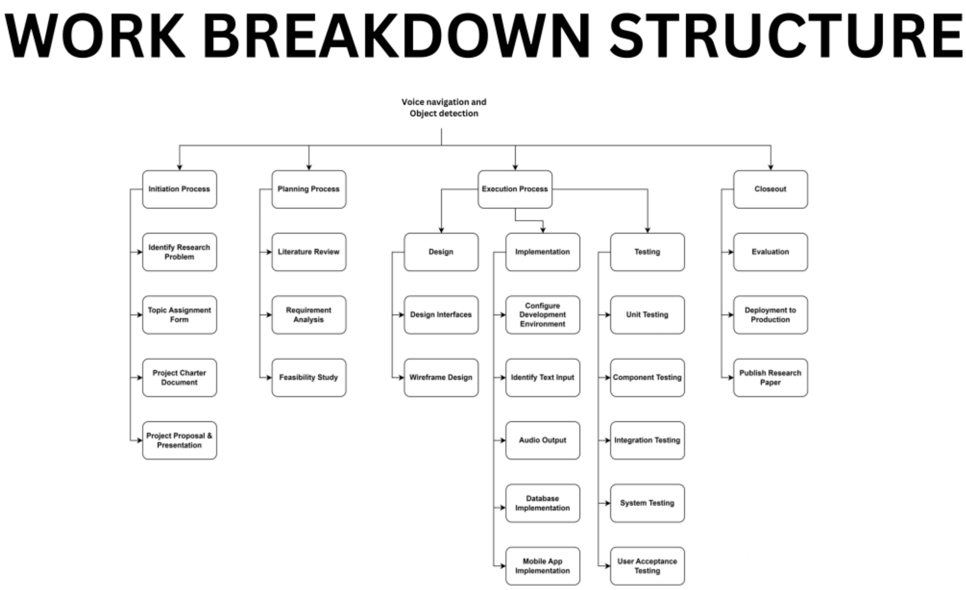
[15] R. Uranga, "Part Pixar, Part Roomba: Meet Moxie, the Pasadena-Built Learning Robot for Children," 29 04 2020. [Online]. Available: https://dot.la/autism-robot-moxie-ai-2645867544.html. [Accessed 06 02 2022].

[16] A. K. Budditha Hettige, "Octopus: A Multi Agent Chatbot," 11 2015. [Online]. Available: https://www.researchgate.net/publication/311652831_Octopus_A_Multi_Agent_Chatbot#:~:text=A%20Chatbot%20is%20a%20type,via%20auditory%20or%20textual%20methods.&text=This%20paper%20presents%20design%20and,Sinhala%20Chatbot%2C%20named%20as%20Octopus.. [Accessed 04 02 2022].

[17] A. W. V. W. a. K. G. Ruvan Weerasinghe, "Festival-si: A Sinhala Text-to-Speech System," [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.521.9076&rep=rep1&type=pdf. [Accessed 07 02 2022].

APPENDICES

Appendix I: Work Breakdown Chart



Appendix II: Gantt Chart