# PROJECT REPORT ON

# Building AI-Powered Solution for Assisting Visually Impaired Individuals

## 1. Introduction

This project leverages Generative AI to create an assistive solution for visually impaired individuals, helping them perceive and interact with their surroundings. It aims to enhance accessibility through real-time scene understanding, text-to-speech conversion, and object detection.

## 2. Problem Statement

Visually impaired individuals often face challenges in understanding their environment, reading visual content, and performing sight-dependent tasks. This project addresses the need for an intelligent and user-friendly solution that assists in real-time scene understanding, safe navigation, and provides daily task guidance.

## 3. Project Overview

The solution, 'OptiAssist', is a web application developed using Streamlit and powered by advanced AI. It enables users to upload images and receive descriptive information, guidance, and audible content that aids in accessibility.

## 4. Features

* ● Real-Time Scene Understanding: Generates descriptive textual output for uploaded images to aid understanding.
* ● Text-to-Speech Conversion: Uses OCR to extract text from images and convert it to speech.
* ● Object and Obstacle Detection: Identifies objects or obstacles for safer navigation.
* ● Personalized Assistance for Daily Tasks: Provides guidance based on image content, such as item recognition.

## 5. Implementation Details

Technologies used for the project include Python, LangChain for AI capabilities, Streamlit for the web interface, and Google Generative AI for scene interpretation. The setup involves installing Python dependencies, configuring Tesseract OCR for text extraction, and setting up the Google Generative AI API.

### Step-by-Step Implementation

Step 1: Install Python Dependencies

Step 2: Set up Tesseract OCR

Step 3: Configure Google Generative AI using the Gemini API

Step 4: Run the Streamlit application

## 6. Usage Instructions

1. Upload an image using the interface.  
2. Select the desired feature: Describe Scene, Text-to-Speech, or Personalized Assistance.  
3. The app processes the image and provides results in text or audio format.

## 7. Code Structure

The main Python script contains all core functionalities. Key sections include:  
1. Imports: Libraries for processing and AI interaction.  
2. Setup: Configurations for OCR and API keys.  
3. Main Functions: Handling image uploads, generating descriptions, and audio conversion.  
4. User Interface: Built using Streamlit with user-friendly navigation.

## 8. Results

The application successfully interprets images, extracts textual information, and provides meaningful descriptions. It enhances accessibility by offering real-time scene understanding, audio output, and daily task assistance.

## 9. Conclusion

This project provides an effective AI-powered solution for visually impaired individuals. Future improvements may include adding real-time video processing and expanding the database for more personalized assistance.

Submitted by:

K GUNASAGAR

(IN9240148).