

**Title**:

SmartSight AI - Assisting Visually Impaired Individuals

**Subtitle:**

Transforming Lives Through AI-Powered Accessibility

**Your Name**: Ashfaq Khan Mohammed (InternID: IN9240612)

**Date**: 01-12-2024

**1: Problem Statement**

* **Title**: The Challenge
  + Visually impaired individuals struggle to perceive and interact with their surroundings.
  + Challenges include:
    - Understanding their environment.
    - Reading visual content.
    - Performing sight-dependent tasks.

**2: Project Overview**

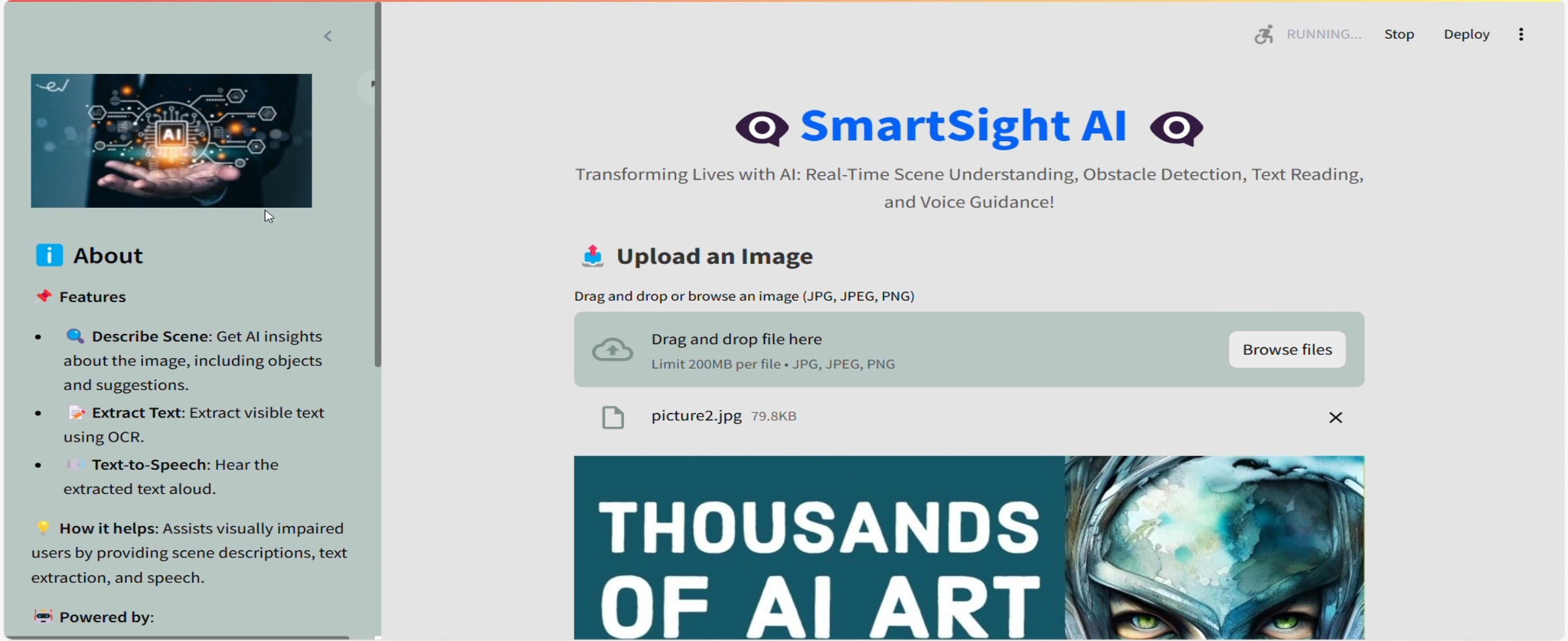
* **Title**: SmartSightAI- The Solution
  + AI-powered application designed for accessibility.
  + Combines advanced technologies to:
    - Provide real-time scene understanding.
    - Offer text-to-speech conversion.
    - Detect objects and obstacles.
    - Deliver personalized assistance.

**3: Features**

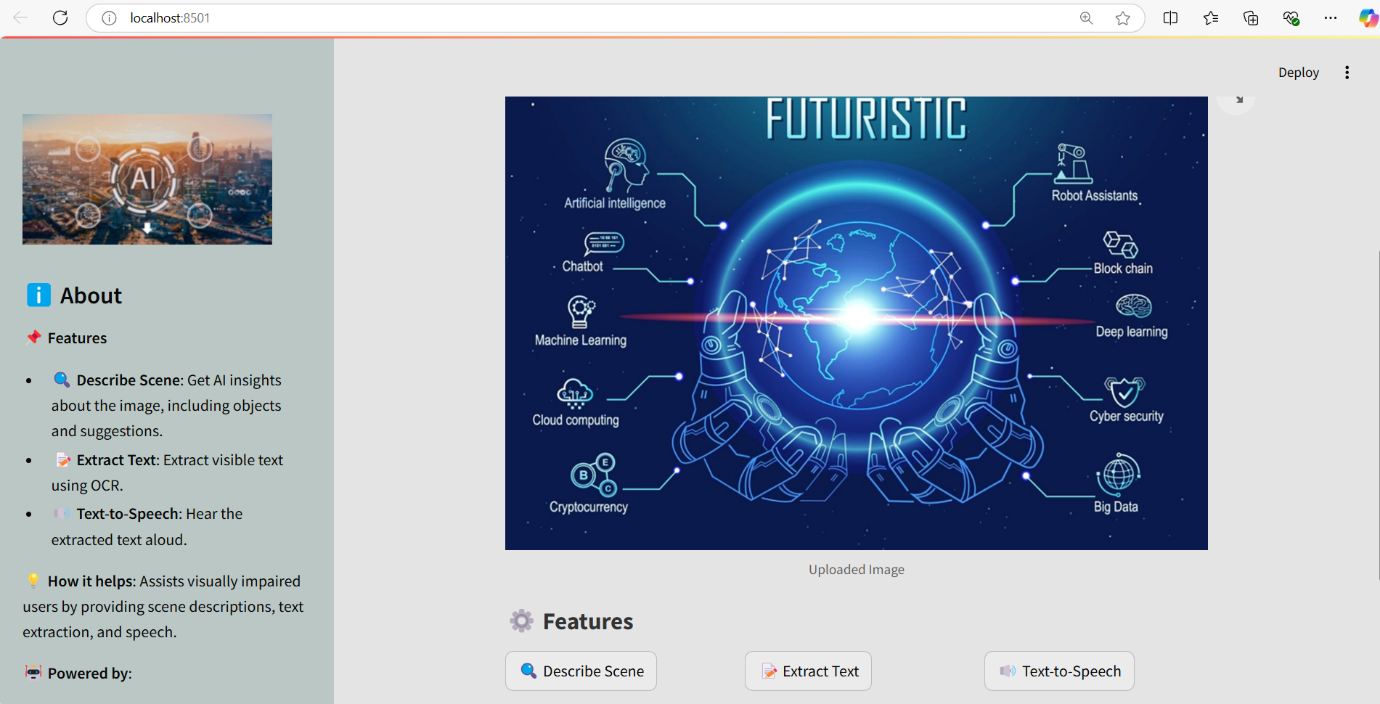
* **Title**: Key Features

**4 Columns with Icons and Text**:

* + Real-Time Scene Understanding
  + Object and Obstacle Detection
  + Text Extraction
  + Text-to-Speech Conversion



**Real-Time Scene Understanding**



**4: Technologies Used**

* **Title**: Technology Stack
  + **Google Generative AI**: Scene description and insights.
  + **YOLOv5**: Object detection and highlighting.
  + **Tesseract OCR**: Text recognition from images.
  + **Streamlit**: Interactive user interface.

**5: Implementation**

* **Title**: How It Works
  + Users upload an image (JPG, JPEG, PNG).
  + AI processes the image to extract features (scene description, text, objects).
  + Outputs include descriptions, text read aloud, and highlighted objects.

**6: Results**

* **Title**: Outputs and Demonstration
  + Accurate scene descriptions for uploaded images.
  + Seamless text-to-speech conversion for accessibility.
  + Clear detection of objects and obstacles.

**7: Impact**

* **Title**: Why It Matters
  + Empowers visually impaired individuals.
  + Simplifies daily activities like reading and navigation.
  + Promotes inclusivity through technology.

**8: Future Enhancements**

* + Add multilingual support for text-to-speech.
  + Integrate real-time camera feed for live assistance.
  + Expand to wearable devices for portability.

**9: Conclusion**

* + Visionary AI bridges accessibility gaps.
  + A step towards an inclusive future powered by AI.
  + Feedback and collaboration are welcome.

**11: Q&A**

**Title**: Questions and Feedback

**Q1: What is the operation of the real-time scene understanding feature?**  
A: By identifying items, their locations, and contextual details, it analyses photos using Google Generative AI and produces a textual description of the scene.

**Q2: Describe YOLOv5 and explain its selection for this project.**A: The cutting-edge object detection model YOLOv5 (You Only Look Once, version 5) is renowned for its accuracy and quickness. It was selected because of its real-time multi-object detection capabilities in photos.

**Q3: How precise are the text-to-speech conversion and text extraction processes?**A: Tesseract OCR, which offers excellent accuracy for readable and clear text, powers the text extraction. Users can access the text-to-speech engine (pyttsx3) since it produces audio output that sounds natural.

**Q4: How is the application developed and implemented?**A: Python is used to construct the program, while Streamlit is used for the front end. It incorporates libraries like Tesseract and YOLOv5 as well as APIs like Google Generative AI. Platforms such as AWS, Heroku, or Streamlit Cloud can host the application.

**Q5: What kind of hardware is needed to run the program?**A: Any device with a modest amount of computing power can run the program and process images locally. For real-time performance, a cloud environment with GPU acceleration is advised for larger models (like YOLOv5).

**Q6: Who is Visionary AI's intended audience?**A: People with visual impairments who require help with reading, navigating, and comprehending their environment make up the main audience. Organizations that promote accessibility or caregivers may be examples of secondary users.

**Q7: How easy is it for non-technical users to utilize the application?**A: The application's interface is straightforward, with options for uploading photographs and choosing features as well as clear instructions. Because it was made with accessibility in mind, anyone can use it with ease.

**Q8: How might this project be expanded?**A: Potential future improvements could consist of:  
Including real-time camera feed support.  
Extending text-to-speech's linguistic skills.  
Incorporating wearable technologies to operate without using your hands.

**Q9: Is it possible for the solution to work with current assistive technology?**A: The application can be modified to integrate with current gadgets, such as smart glasses or mobile apps, through the use of APIs.

**Q10: What moral issues were taken into account when developing this solution?**  
A: Data security and user privacy must be guaranteed. No personal information is saved without the user's permission, and the application can process photographs locally or safely on the cloud.

**Q11: What is the difference between SmartSight AI and comparable solutions?**A: Visionary AI is more adaptable than many current tools that concentrate on only one area since it integrates several functions (text reading, obstacle detection, scene interpretation, and text-to-speech) into a single platform.

**THANK YOU**