Speak Easy App

Shruti Pramod Lalwani[1], Yash Ajit Kothari[2], Raj Ratanraj Shah[3], Mahavir Ravindra Bhandari[4]

Abstract

The SpeakEasy App is a voice-controlled assistant application designed specifically for individuals with visual impairments, aimed at enhancing their productivity and access to vocational training. This Android-based app leverages voice-driven interactions, allowing users to manage tasks, set reminders, and access audio-based vocational training content entirely hands-free. By utilizing voice commands, users can create, modify, and delete tasks, schedule and receive reminders, and engage in skill-building exercises without needing to navigate traditional visual interfaces.The core functionality of SpeakEasy centers around providing an intuitive and accessible user experience that empowers visually impaired individuals to perform everyday activities more efficiently. With seamless integration of voice recognition technology, the app ensures fast, accurate, and reliable responses to voice prompts. Its design prioritizes accessibility, adhering to standards that make it compatible with screen readers and other assistive technologies. In addition, the app offers flexibility in operation, functioning offline for task management and reminders, while requiring an internet connection for downloading new content and updates. The SpeakEasy App is not only a tool for productivity but also a platform for vocational empowerment. Through voice-guided modules, users can engage in vocational training to build skills that enhance their employability. This SRS outlines the system's features, technical requirements, design architecture, and non-functional attributes, ensuring a robust and accessible solution for the visually impaired community. The app represents a step forward in accessible technology, enabling greater independence and improved quality of life for users.

Keywords: Voice-controlled assistant, Visual impairments, Accessibility, Android-based app, Voice-driven interactions, Task management, Reminders

INTRODUCTION

The SpeakEasy App is an innovative voice-controlled assistant designed to empower individuals, particularly those with visual impairments, by enhancing their productivity and independence. With approximately 8 million people living with complete blindness in India, including around 270,000 blind children, there exists a significant gap in accessible educational resources and vocational training opportunities tailored to their unique needs. Current solutions are limited, often failing to integrate both educational and mobility support, which are crucial for fostering self-reliance. The SpeakEasy App aims to bridge this gap by providing a user-friendly platform that allows users to manage tasks, set reminders, and access information through simple voice commands. By leveraging cutting-edge voice recognition technology, the app ensures that visually impaired users can interact seamlessly with their devices, enabling them to accomplish daily tasks with greater ease. This Software Requirements Specification (SRS) document outlines the essential features, functionalities, and design considerations necessary for the development of the SpeakEasy App. By focusing on the specific challenges faced by visually impaired individuals, the app not only aims to enhance their day-to-day productivity but also contributes to their overall empowerment in personal and professional spheres. Through this initiative, we aspire to create a

more inclusive environment where individuals with visual impairments can thrive and achieve their full potential.

LITERATURE SURVEY

The development of specialized vocational training software for visually impaired individuals is gaining attention as a crucial area for research and development. While advances in assistive technologies have been made, significant challenges remain in creating comprehensive solutions that address both vocational training and mobility support for this population. Current offerings often fall short of meeting the unique needs of visually impaired individuals, underscoring the need for more integrated and user-friendly platforms.

Existing Vocational Training Solutions for Persons with Disabilities Various online and mobile platforms exist to support persons with disabilities in their vocational training, but many of them lack specific features tailored to the needs of visually impaired individuals (PVI).

Online Vocational Training Platforms Online vocational training platforms like Coursera, Udemy, and edX provide a vast array of courses that aim to be accessible to diverse learners. Some of these platforms incorporate accessibility features such as subtitles, text-to-speech compatibility, and screen reader support. However, despite these efforts, these platforms generally cater to a broad audience, often overlooking the specific requirements of PVI learners.

Research indicates that the learning experience for visually impaired individuals can be significantly affected by the design of course materials. For example, courses that heavily rely on visual content—such as videos or infographics—are often not fully accessible to visually impaired students. While platforms may offer compatibility with screen readers like JAWS or NVDA, the effectiveness of these tools can vary. A study by highlights that a substantial gap exists in addressing the specific needs of PVI learners, particularly in hands-on vocational training that necessitates physical or visual interaction. Thus, while general online platforms make strides toward universal design, they still fall short in providing specialized tools and training specifically for visually impaired individuals.

Mobile Applications for People with Visual Impairments (PVI) Several mobile applications are designed to assist visually impaired individuals, yet most do not focus on vocational training. For instance, Be My Eyes connects visually impaired users with volunteers for real-time assistance through video calls. This application proves valuable for everyday tasks, such as reading labels or navigating new environments. However, it does not provide structured vocational training, although its functionality could be adapted for vocational settings to enhance job training experiences

Seeing AI, developed by Microsoft, utilizes artificial intelligence to help visually impaired users interact with their surroundings by describing objects, text, and even currency. Although not designed specifically for vocational training, its capabilities offer potential for learning job-related skills that involve identifying objects or reading instructions. While these applications provide essential support in daily life, they lack robust vocational training functionalities. They primarily focus on general accessibility or daily assistance rather than offering structured, job-specific training. This presents an opportunity for the development of specialized mobile or web-based vocational training solutions tailored to the needs of visually impaired learners. Future vocational training apps could benefit from integrating real-time, hands-on learning experiences with adaptive feedback mechanisms, similar to those found in existing assistive technologies.

The literature indicates a pressing need for specialized vocational training solutions for visually impaired individuals. While existing online platforms and mobile applications provide valuable resources, they often

do not fully address the specific challenges faced by PVI learners. There is significant potential for developing adaptive, user-friendly platforms that enhance vocational skills and improve mobility, ultimately empowering visually impaired individuals to achieve greater self-reliance in both professional and personal contexts. Continued research and innovation in this field are essential to create impactful vocational training solutions that meet the unique needs of visually impaired learners.

METHODOLOGY

The methodology for developing the SpeakEasy App follows a structured and user-centric approach, ensuring accessibility, reliability, and efficiency in assisting visually impaired individuals.

1. Requirement Analysis

Conduct user research to understand the needs of visually impaired individuals through surveys, interviews, and focus groups.

Identify functional and non-functional requirements, including voice command accuracy, offline capabilities, and screen reader compatibility.

Analyze existing assistive technologies to integrate best practices.

2. System Design and Architecture

Define the system architecture, including backend services, databases, and APIs for voice recognition.

Design an intuitive voice-command interface that enables hands-free operation.

Ensure compatibility with Android accessibility features and screen readers.

Develop an offline task management system and cloud-based storage for content updates.

3. Development Phase

Frontend Development: Implement a simple, accessible, and responsive voice-controlled interface.

Backend Development: Develop APIs for voice recognition, task management, and content delivery.

Database Management: Store user tasks, reminders, and vocational training modules efficiently.

Integration of Speech-to-Text & Text-to-Speech: Implement Google's Speech API or similar services for accurate voice recognition.

4. Testing and Validation

Usability Testing: Conduct testing sessions with visually impaired users to evaluate ease of use.

Functionality Testing: Ensure all voice commands execute correctly, including task creation, modification, and reminders.

Performance Testing: Assess response time, offline functionality, and scalability.

Accessibility Compliance: Validate adherence to WCAG (Web Content Accessibility Guidelines) and Android accessibility standards.

5. Deployment and Maintenance

Deploy the app on the Google Play Store with an onboarding guide for new users.

Provide periodic updates to enhance features, improve voice recognition, and add vocational training modules.

Maintain a feedback loop with users for continuous improvement.

6. Evaluation and Future Enhancements

Collect user feedback through in-app surveys and community engagement.

Enhance AI-driven voice interactions for better accuracy.

Expand vocational training content based on user demand.

Explore multi-language support for broader accessibility.

OBJECTIVE

For this feasibility assessment, we focus on two main objectives of the SpeakEasy App:

Efficient Voice Command Recognition: The system must accurately interpret voice inputs and map them to relevant commands. This involves analyzing voice data to determine intent and ensuring accurate command execution.

Task Scheduling and Notifications: The app needs to schedule tasks, send notifications at set times, and allow for real-time updates based on user inputs.

PROBLEM DEFINATIONS

Individuals with visual impairments face significant challenges in managing daily tasks, accessing vocational training, and improving productivity due to the reliance of most digital interfaces on visual elements. Traditional task management and learning platforms are often inaccessible, requiring manual navigation through complex menus, small text, or on-screen interactions that do not cater to their needs.

Despite advancements in assistive technology, many existing solutions lack seamless voice-controlled interactions, offline functionality, and accessibility-focused design, limiting their usability for visually impaired users. Additionally, vocational training resources are often presented in text-based or visually demanding formats, making it difficult for visually impaired individuals to acquire new skills for employment and personal development.

The SpeakEasy App aims to address these challenges by providing an Android-based, voice-controlled assistant that allows users to manage tasks, set reminders, and access vocational training modules hands-free. By integrating voice recognition technology, screen reader compatibility, and offline functionality, the

app ensures greater independence, accessibility, and usability for visually impaired individuals, empowering them to enhance their productivity and skill development without relying on traditional visual interfaces.

DATA FLOW DIAGRAMS

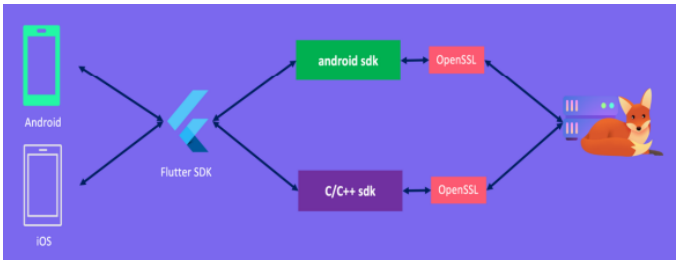

Figure: Level2Data Flow Diagram

FUCTIONAL REQUIREMENTS

System Feature 1(Voice Command Processing) The system must process user commands and perform appropriate actions (e.g., setting reminders, alarms, or managing tasks).

System Feature 2(Task Management) Users should be able to create, update, and delete tasks using voice commands.

System Feature 3(Notifications) The app must provide timely notifications to remind users of tasks or events they've created.

System Feature 4(Voice Feedback) The system should provide verbal responses to users to confirm actions or provide feedback on command processing.

External Interface Requirements

User Interfaces The user interface should be minimalistic, with a focus on voice interactions. Users should have access to a task list, settings, and command history.

Hardware Interfaces Android mobile device with microphone and speaker support

Software Interfaces The app should integrate with Android's system services like alarms and notifications. Use of APIs for voice recognition (e.g., Google Speech-to-Text).

Communication Interfaces The app requires an active internet connection to process voice commands through cloudbased APIs.

NON FUCTIONAL REQUIREMENTS

Performance Requirements

• Voice commands should be processed within a few seconds.

• Notifications should be triggered at the correct time without delays.

• The app should handle multiple voice commands efficiently

Safety Requirements The app must ensure that voice data is securely transmitted and processed, with a focus on user privacy.

Security Requirements User data such as tasks, settings, and personal information must be stored securely. Access to the app must be protected using device-level security such as PINs or biometrics.

Software Quality Attributes

Usability: The app should be easy to use with intuitive voice commands. Robustness: The app must gracefully handle errors such as unrecognized voice inputs.

Scalability: The system should handle a growing number of users and tasks.

RESULTS

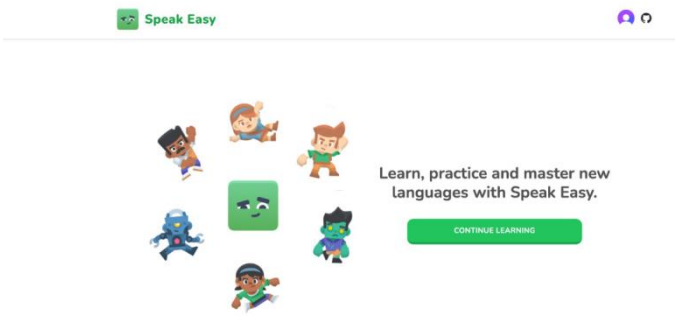

Fig : 1

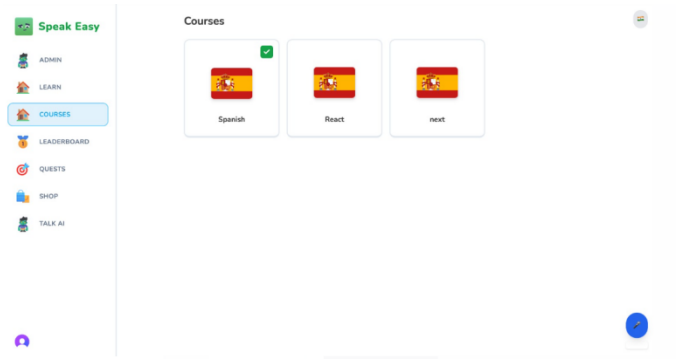

Fig : 2

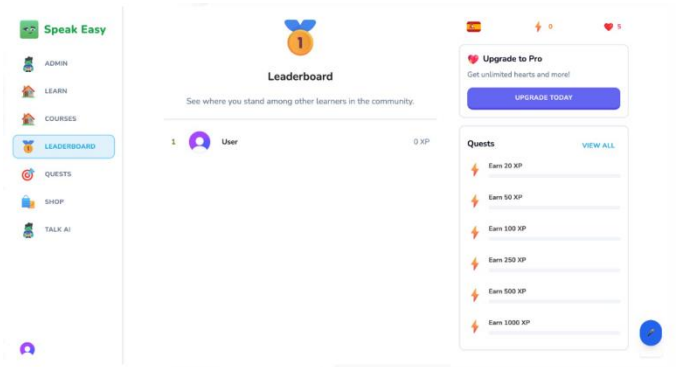

Fig : 3

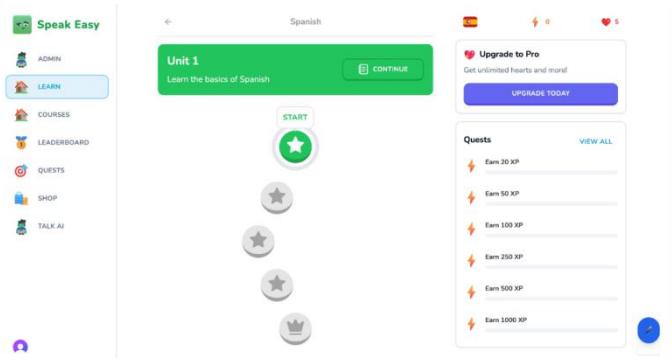

Fig : 4

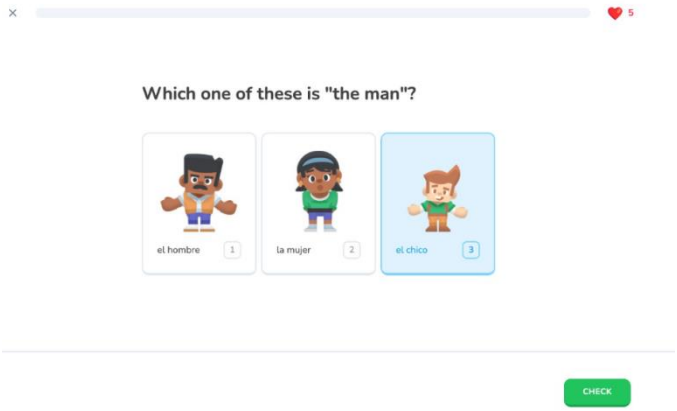

Fig : 5

CONCLUSION

The SpeakEasy App is designed to offer users an intuitive, hands-free task management experience through voice recognition. By leveraging advanced machine learning algorithms for voice processing and integrating real-time databases, SpeakEasy provides a seamless and efficient solution for managing tasks and reminders on Android devices. This project has shown that voice command systems can significantly enhance productivity by enabling users to perform tasks without needing manual input. Through careful design and the use of modern cloud-based tools like Firebase and Google Speech-to-Text API, the app successfully balances usability, performance, and data security. The feasibility analysis and initial testing demonstrate that the app can meet user needs while maintaining accuracy and responsiveness.

REFERENCES

1. T. N. Sainath, R. J. Weiss, A. Senior, K. W. Wilson, and O. Vinyals, ”Learning the Speech Front-end with Raw Waveform CLDNNs,” In Proceedings of the 41st IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2015), Brisbane, Australia, April 2015, pp. 4885-4889. IEEE.

2. K. Rao, H. Sak, and R. Prabhavalkar, ”Exploring Architectures, Data and Units for Streaming End-to-End Speech Recognition with RNN-Transducer,” In Proceedings of the IEEE Automatic Speech Recognition and Understanding Workshop (ASRU 2017), Okinawa, Japan, December 2017, pp. 193-199. IEEE.

3. Firebase, ”Firebase Realtime Database Documentation,” [Online]. Available: https: //firebase.google.com/docs/database. [Accessed: Nov. 14, 2024].

4. Google Cloud, ”Speech-to-Text Documentation,” [Online]. Available: https://cloud. google.com/speech-to-text/docs. [Accessed: Nov. 14, 2024].

5. Android Developers, ”Introduction to Android Development,” [Online]. Available: https://developer.android.com/guide. [Accessed: Nov. 14, 2024].
6. S. Park, Y. Zhou, and T. Leung, ”Deep Neural Networks for Image Recognition and Voice Command Classification,” In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2018), Salt Lake City, Utah, USA, June 2018, pp. 3425-3434. IEEE.
7. M. Schuster and K. K. Paliwal, ”Bidirectional Recurrent Neural Networks,” IEEE Transactions on Signal Processing, vol. 45, no. 11, pp. 2673-2681, Nov. 1997.
8. A. van Engelen and H. Hoos, ”A Survey on Semi-Supervised Learning,” Machine Learning, vol. 109, no. 2, pp. 373-440, 2020. Springer.