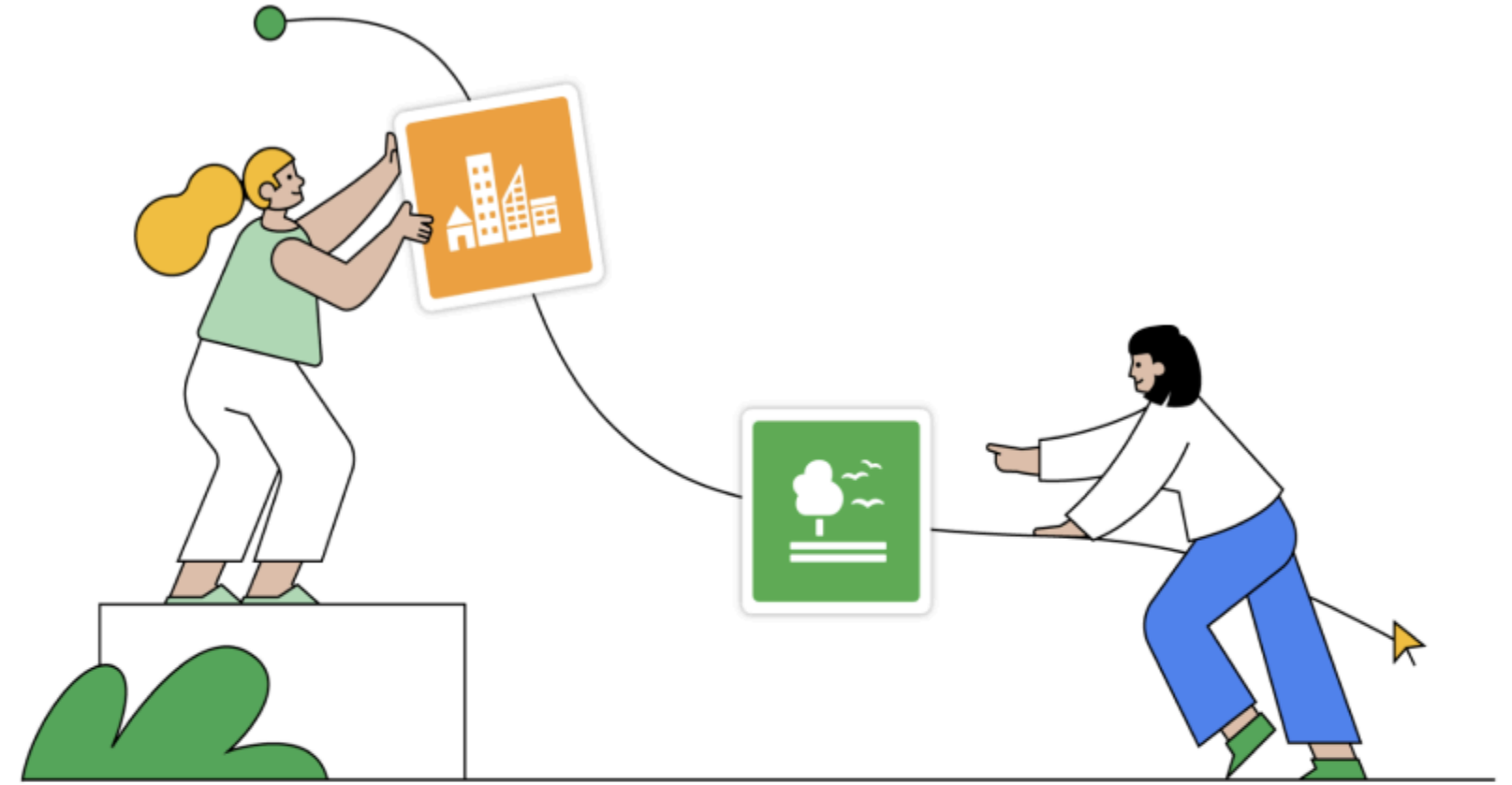


# Solution Challenge



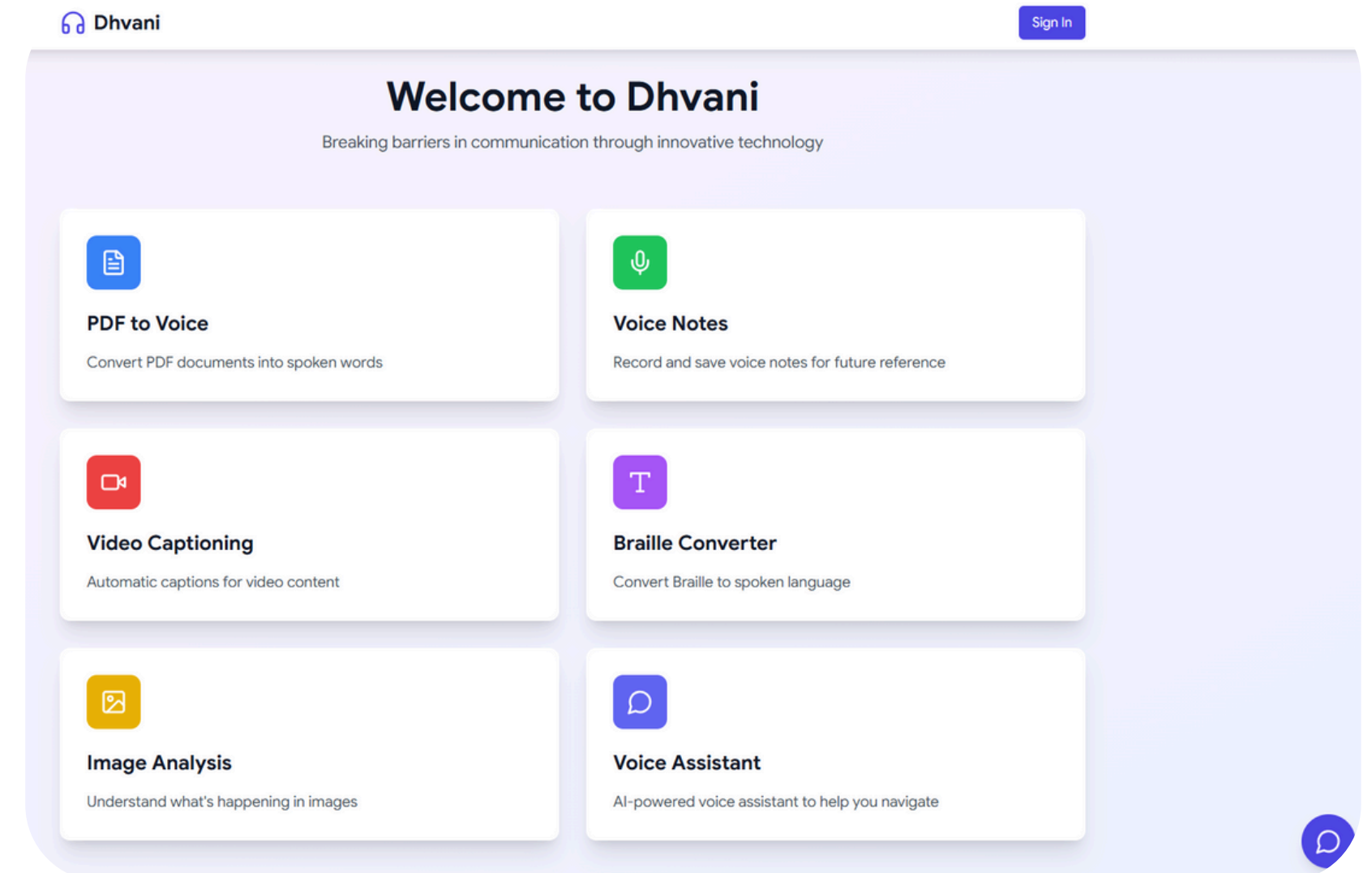
**Team name:** EdgeTech

**Team leader name:** Rahil Masood

**Problem Statement:** Bridging the Education and Digital Literacy Gap for Marginalized Communities

# BRIEF ABOUT YOUR SOLUTION:

Dhvani is an innovative web application designed to empower individuals with hearing and speech impairments. By leveraging cutting-edge technology and Google's powerful AI tools, we're making communication more accessible and inclusive for everyone.



## HOW DIFFERENT IS **DHVANI** FROM OTHER EXISTING IDEAS?

- **All-in-One Accessibility Tool** – Combines PDF-to-voice, video captioning, Braille-to-speech, and AI assistance in one platform.
- **Offline Support** – Works without the internet for key features, unlike most cloud-based assistive apps.
- **AI-Powered Context Awareness** – Uses Google's Gemini AI for more accurate speech, text, and image processing.
- **Highly Customizable & Cross-Platform** – Supports multiple languages, voice adjustments, and works on web, mobile, and desktop.

## HOW WILL **DHVANI** SOLVE THE PROBLEM?

- **Bridges Communication Gaps** – Converts speech, text, and Braille to assist hearing- and speech-impaired users.
- **Improves Digital Literacy** – Provides AI-powered reading and transcription tools for better access to education.
- **Enhances Social Inclusion** – Enables seamless conversations via real-time video captioning and voice assistants.
- **Empowers Independence** – AI-driven image analysis and voice interactions reduce reliance on human assistance.

## USP

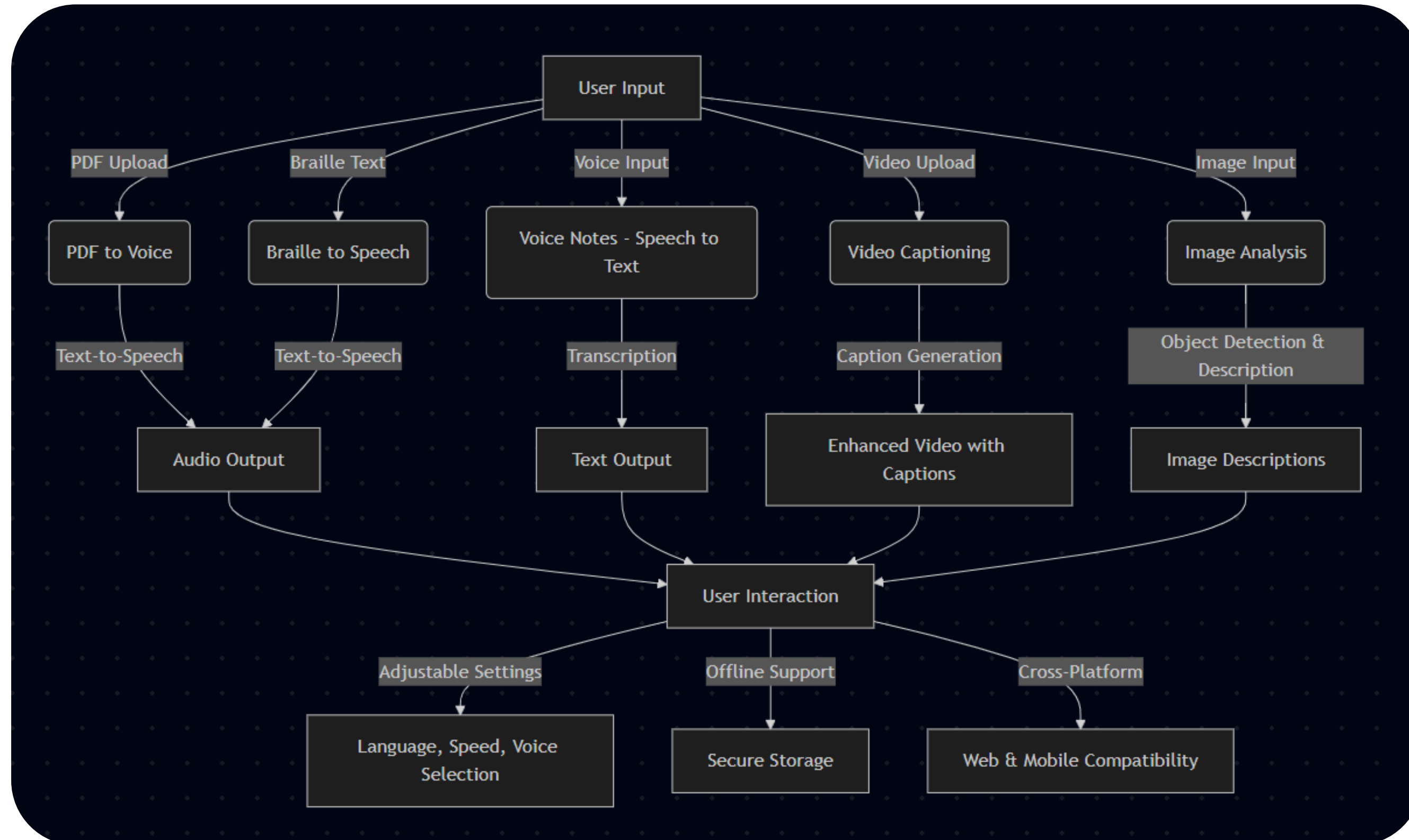
- **All-in-One Accessibility Platform** – Integrates multiple assistive technologies in a single, AI-powered application.
- **Offline Accessibility** – Works without an internet connection for critical features, unlike most cloud-based services.
- **Google AI for Context-Aware Assistance** – Uses Gemini API, Speech-to-Text, and Vision AI for more accurate and intelligent communication.
- **Multimodal Support** – Supports text, speech, images, and videos, ensuring a richer experience than single-modality apps.
- **Highly Customizable & Inclusive** – User-controlled voice selection, language preferences, and WCAG-compliant UI for broader accessibility.

# LIST OF FEATURES OFFERED BY THE SOLUTION

## ◆ Key Features at a Glance

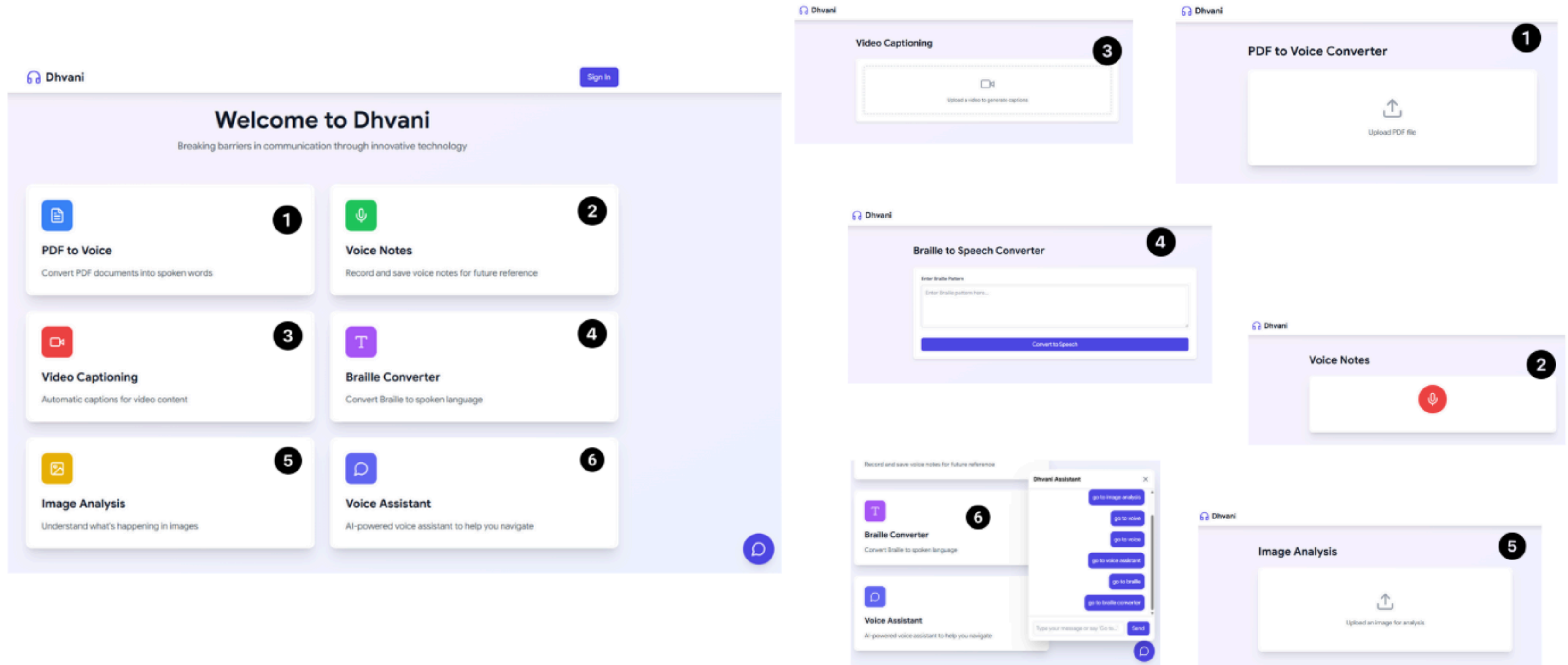
- ✓ PDF to Voice – Convert PDFs into speech with multi-language support & real-time status
- ✓ Voice Notes – Record & transcribe with auto-punctuation & offline storage
- ✓ Video Captioning – Real-time, multi-language captions with editing options
- ✓ Braille to Speech – Convert Braille text into clear, adjustable voice output
- ✓ Image Analysis – Scene descriptions, object detection & spatial awareness
- ✓ AI Voice Assistant – Smart, voice-activated navigation with offline fallback
- ✓ Accessibility First – WCAG-compliant, offline capabilities & privacy-focused

# PROCESS FLOW DIAGRAM

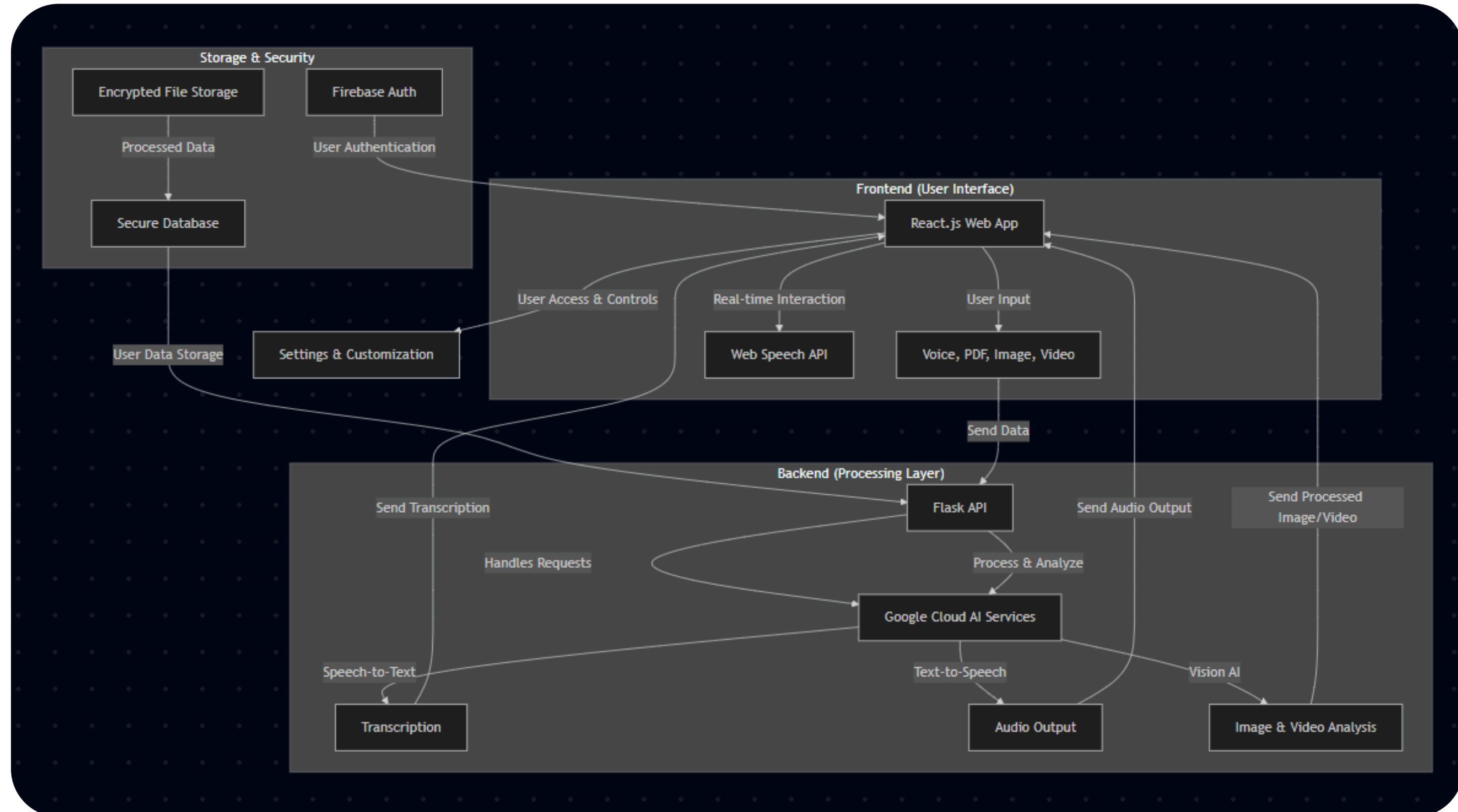




## WIREFRAMES/MOCK DIAGRAMS OF THE PROPOSED SOLUTION:



# ARCHITECTURE DIAGRAM OF THE PROPOSED SOLUTION:



# TECHNOLOGIES TO BE USED IN THE SOLUTION:

## Frontend:

React 18 – For a dynamic and responsive UI  
TypeScript – Ensures type safety and better maintainability  
Tailwind CSS – For fast and scalable styling  
Framer Motion – Adds smooth animations  
Web Speech API – Enables voice interaction

## Backend:

Flask (Python) – Lightweight backend framework  
Google Cloud APIs – AI-powered features  
Gemini API – Natural language processing  
Speech-to-Text API – Converts voice to text  
Text-to-Speech API – Converts text to natural-sounding speech  
Vision AI – Image and video analysis

## Deployment & DevOps:

Docker – Containerized deployment  
Google Cloud Run – Serverless hosting for the backend

## Storage & Authentication:

Firebase Auth – Secure user authentication  
Google Cloud Storage – Secure media and file storage  
Firestore – Storing user preferences & data

## Security Features:


OAuth 2.0 & Firebase Authentication – Secure user access  
End-to-End Encryption – Ensures privacy in data storage & transfer  
Rate Limiting & API Gateway – Prevents misuse



# ESTIMATED IMPLEMENTATION COST:

Component	Estimated Cost (per month)	Notes
Frontend Hosting (Vercel/Netlify)	\$0 - \$20	Free tier available, cost scales with usage
Backend Hosting (Google Cloud Run)	\$10 - \$50	Depends on request volume
Google Cloud APIs (Gemini, Speech-to-Text, Text-to-Speech, Vision AI)	\$50 - \$200	Free tier covers limited usage, cost scales with API calls
Database (Firestore/PostgreSQL)	\$10 - \$100	Free tier available, pricing depends on storage and queries
Authentication (Firebase Auth)	\$0 - \$25	Free for small user base, paid for advanced features
Storage (Google Cloud Storage)	\$10 - \$50	Depends on file uploads (PDFs, images, videos)
Security (Rate Limiting, Encryption)	\$10 - \$50	Includes API Gateway and data protection costs
Domain & SSL	\$10 - \$20	Custom domain and SSL certificate
DevOps & CI/CD (GitHub Actions, Docker, Monitoring Tools)	\$0 - \$30	Mostly free, minimal cost for advanced usage
Miscellaneous (API Gateway, Analytics, CDN)	\$10 - \$50	Based on optimization needs


# SNAPSHOTS OF THE MVP

 Dhvani

Sign In

Welcome to Dhvani


Breaking barriers in communication through innovative technology



PDF to Voice

Convert PDF documents into spoken words


1



Voice Notes

Record and save voice notes for future reference


2



Video Captioning

Automatic captions for video content

3



Braille Converter

Convert Braille to spoken language

4





Image Analysis

Understand what's happening in images


5




Voice Assistant

AI-powered voice assistant to help you navigate


6




 Dhvani

Video Captioning

3




Upload a video to generate captions


 Dhvani

PDF to Voice Converter

1



Upload PDF file

 Dhvani


Braille to Speech Converter

4

Enter Braille Pattern


Enter Braille pattern here...

Convert to Speech


 Dhvani

Voice Notes

2




Record and save voice notes for future reference



Braille Converter

Convert Braille to spoken language

6



Voice Assistant

AI-powered voice assistant to help you navigate

Dhvani Assistant

go to image analysis

go to volve

go to voice

go to voice assistant

go to braille

go to braille converter

Type your message or say 'Go to...'

Send



 Dhvani

Image Analysis

5



Upload an image for analysis

**1. GitHub Public Repository:**

**<https://github.com/rajarshidattapy/Dhvani-AI>**

**2. Demo Video Link: <https://youtu.be/dLN9QBRidUQ>**

**3. MVP Link: [dhvani-ai.vercel.app](https://dhvani-ai.vercel.app)**



# Solution Challenge



Thank you

