# SMART INDIA HACKATHON 2024

Team Details and Problem Statement

- **Problem Statement ID** - 1715

- **Problem Statement Title** - AI tool/mobile app for Indian Sign Language(ISL) generator from audio-visual content in English/Hindi to ISL content and vise-versa.

- **Theme**- Miscellaneous

- **PS Category**- Software

- **Team ID**- MUJSIH017

- **Team Name** - Clueless Coders

❖ Idea/Approach Details

1. Indian Sign Language (ISL) is a visual-gestural language used by deaf and hard-of-hearing individuals across India. It encompasses a rich vocabulary of hand movements, facial expressions and body postures to convey messaging.

2. Our solution is to Develop a user-friendly web/mobile app, which comprises models that translate from spoken language to Indian Sign Language (ISL) and vice-versa in real time. Tailored to help Deaf, Mute, Blind and Non-ISL individuals with effortless and swift interpersonal communication.

3. The project distinguishes itself by integration of LLMs for advanced text comprehension and (YOLOv4-CSP/Open CV) for precise gesture mapping, a novel application in this field. Leveraging these cutting-edge AI technologies ensures high accuracy and adaptability to diverse ISL nuances, including accurate gesture detection, clear speech extraction (using Nemo-Training, sbert, DeepSpeech) and precise conversation interconversion.

4. This solution bridges the communication gap between deaf and mute individuals who use ISL and those who communicate through spoken language. By providing a real-time, accurate conversion of speech to ISL, it facilitates effective communication in various settings such as education industry, public-transport, media (news and sports),etc. enhancing accessibility and inclusivity.

5. This integration provides real-time performance and a user-friendly interface, greatly enhancing communication between auditory-impaired, visually impaired, and auditory-abled communities.

- To deal with different scenarios, we plan to integrate different pre-built and tuned open-sourced **DL models** to work and convert different forms of communication-inputs.

- For Speech and Text conversion; **Alphacephi** is a low-resource model that can function perfectly well in very low resource equipment and is trained on multiple international and Indic languages.

- The **sbert(Indian Sentence Similarity)** model is a very good STS model tuned for thousands of hours on major Indian languages and dialects. For models specific to certain accents and languages, **AI4Bharat** and **Vakyansh** provide very good tuned models.

- For ISL Gestures conversion, **YOLOv4-CSP** model is preferred for its ease of access and parameterized. We will be fine tuning the model further according the ISL Alphabet databases to cater to different types of gestures. We are referring to established general ISL conversion models; at the word and speech level.

- Finally for text Processing we will be using open-source language models like **BERT, T5**, etc. LLMs which are good at Language Processing.

TECHNICAL APPROACH

SMART INDIA
HACKATHON
2024
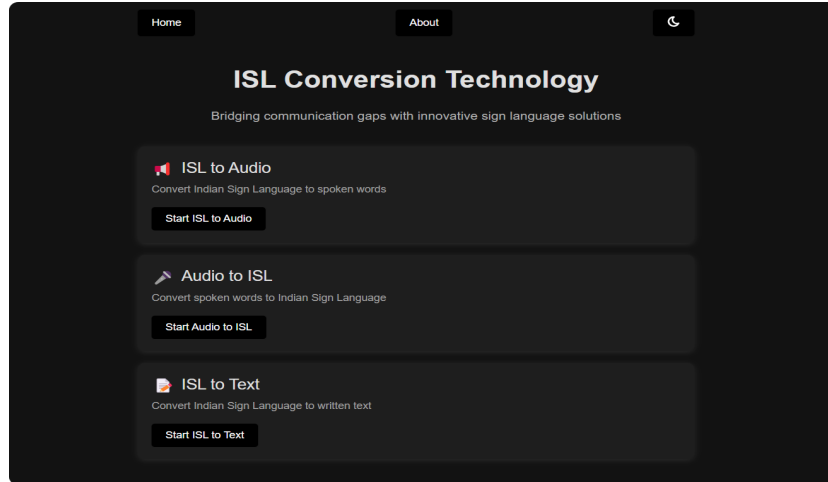
TECH STACK:

1. **Front-End:** Flutter App-Development, React Native, JavaScript, ReactJS.
2. **Back-End:** Python, Kotlin, Swift, JS and TS.
   a. **Speech to Text-** Nemo-Training, Alphacephi, Mozilla DeepSpeech, sbert, Vakyansh, wav2vec2, Talk-back-feature.
   b. **ISL to Text-** YOLOv4-CSP model, OpenCV, TensorFlow, PyTorch, Keras, SciKit.
   c. **NLTK/Transformers-** NLP Text Processing LLAMA 3.1,RoBERTa, DistilBERT, XLNet, T5 models.
3. **Data-Storage:** MongoDB, AWS for permanent site and data hosting.

FEASIBILITY AND VIABILITY
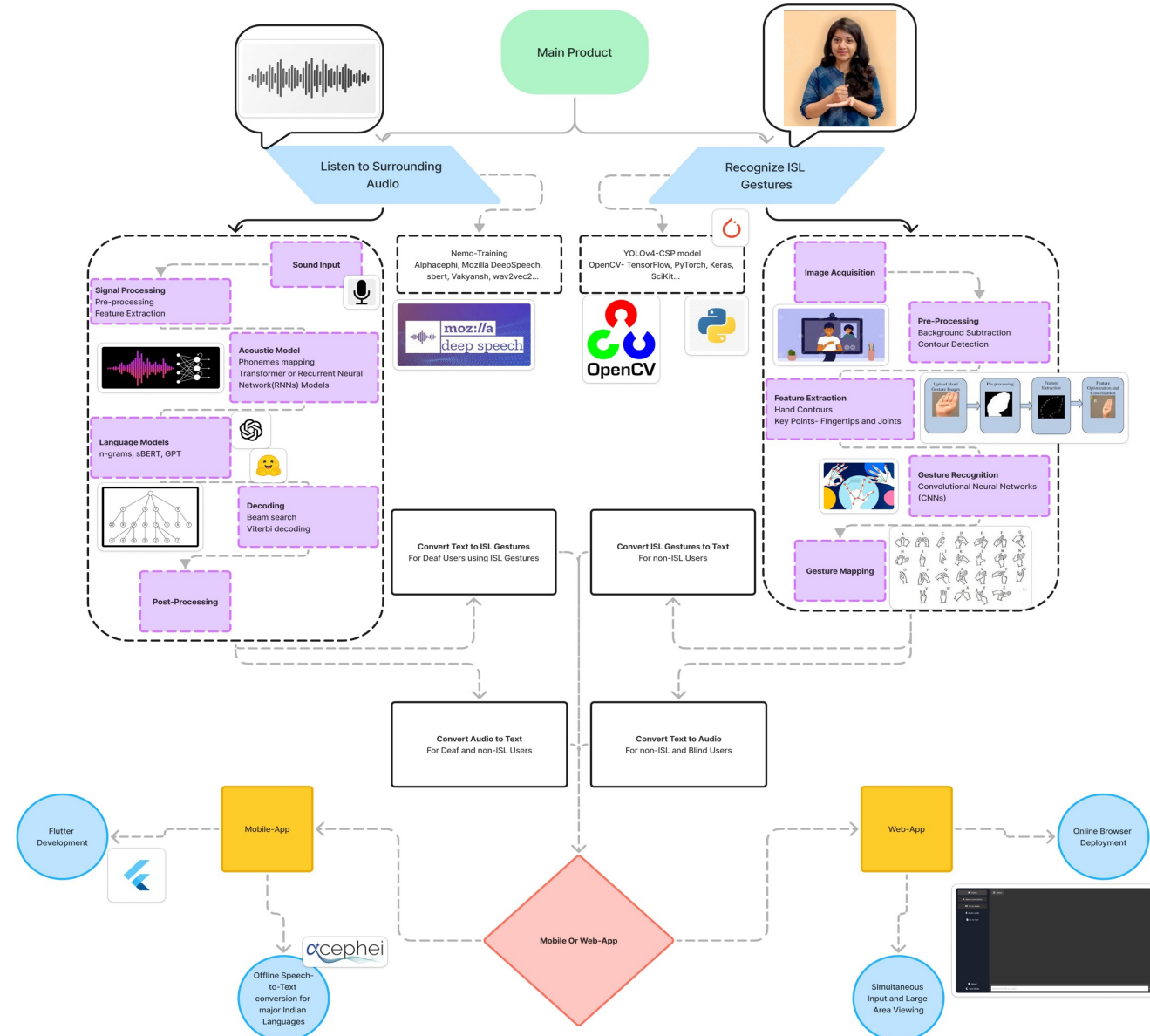
Technical Feasibility:

- Speech Recognition: Advanced but needs fine-tuning for diverse Indian accents.

- ISL Gesture Generation: Requires deep understanding of ISL; real-time processing is challenging.

- NLP & Computer Vision: Essential for accurate and real-time ISL conversion.

- Using Talk-back feature for Clear speech inputs in crowded places.

Challenges & Risks:

- Accuracy: Regional dialects, noise, and ISL grammar complexity.

- Real-Time Processing: High computational demand on mobile devices.

- User Adoption: Acceptance and ease of use among the deaf and mute community.

- Privacy: Handling sensitive audio data.

Strategies:

- Dataset Expansion: Train models on diverse regional accents.

- Collaboration: Work with ISL experts for accurate gesture mapping.

- Optimization: Develop lightweight models for better performance.

- User-Centered Design: Engage with the community for feedback and adoption.

IMPACT AND BENEFITS
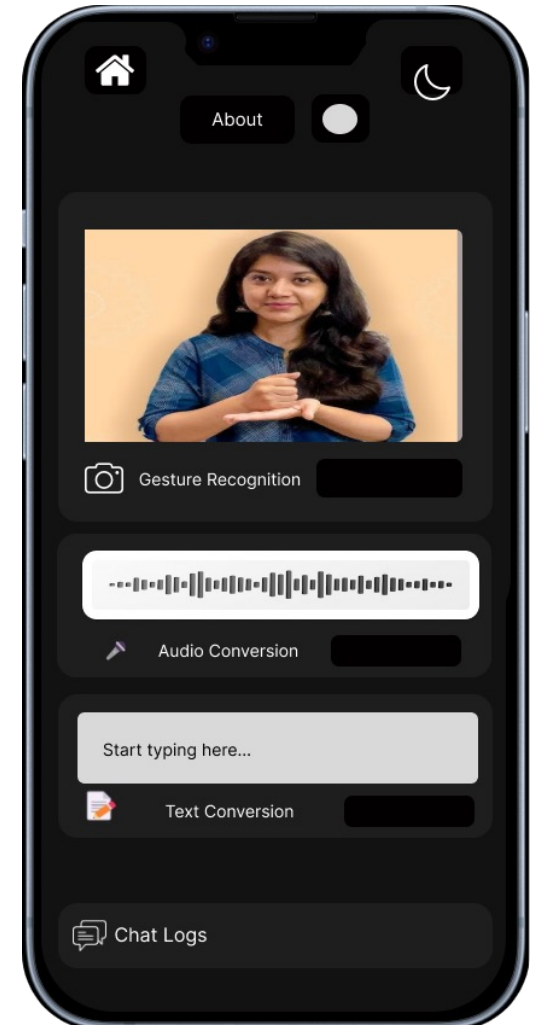
Potential Impact→

Enhanced Communication: Breaks barriers between differently-abled (deaf, blind) individuals and those unfamiliar with ISL, fostering inclusivity.

Increased Access: Enables the differently-abled communities to participate fully in education, healthcare, and daily interactions.

Benefits→

Social: Promotes equal opportunities and social integration for the deaf community.

Economic: Improves job prospects and educational outcomes by facilitating better communication.

Technological: Advances AI and language processing technologies, contributing to broader applications.

Environmental: Reduces the need for printed materials and interpreters in various settings, lowering carbon footprint.
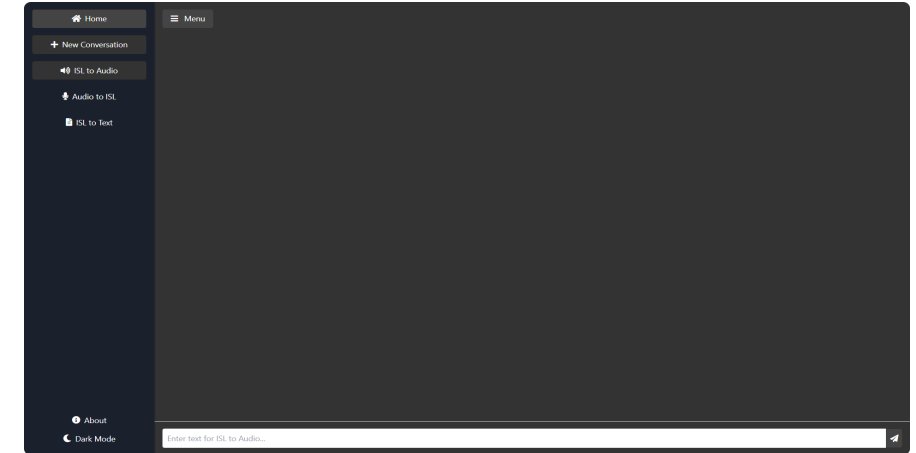
Example Applications→

1. **Railway Platform Display:**
   - Develop an application that converts spoken announcements on railway platforms into ISL, allowing deaf individuals to understand them.

2. **Mobile App for ISL interpretation :**
   - Create a mobile app that uses a camera to help hearing individuals understand the sign language used by deaf individuals.

RESEARCH AND REFERENCES

Clueless Coders

🗄 Key Data Sources
- Indian Sign Language Portal iSign
- Dataset
- ISL Research & Training Center
- DePwd
- Catalog | Open Government Data (OGD) Platform India

🎤 Pre-Built Models
- AI4Bharat Models
- Vakyansh Models
- Alphacephi VOSK
- Kaldi Speech Recognition
- Mozilla DeepSpeech
- Sbert

📄 Research Papers
- Indian Sign Language Translator using Gesture Recognition
- NLP English to ISL Converter
- Gesture to Text using SURF
- ISL to Text Conversion
- Harnessing AI to Generate Indian Sign Language from Natural Speech and Text

Gesture Conversion
- Kaggle ISL Dataset
- ISign Dataset-Paper
- Word-Level ISL Detection
- Speech to ISL Conversion
- Speech to ISL web
- Text to ISL App
- ISL with CNNs
- Any Model Can Talk

References
- ISL Research and Training Center
- Benchmark for ISL Processing

GitHub link