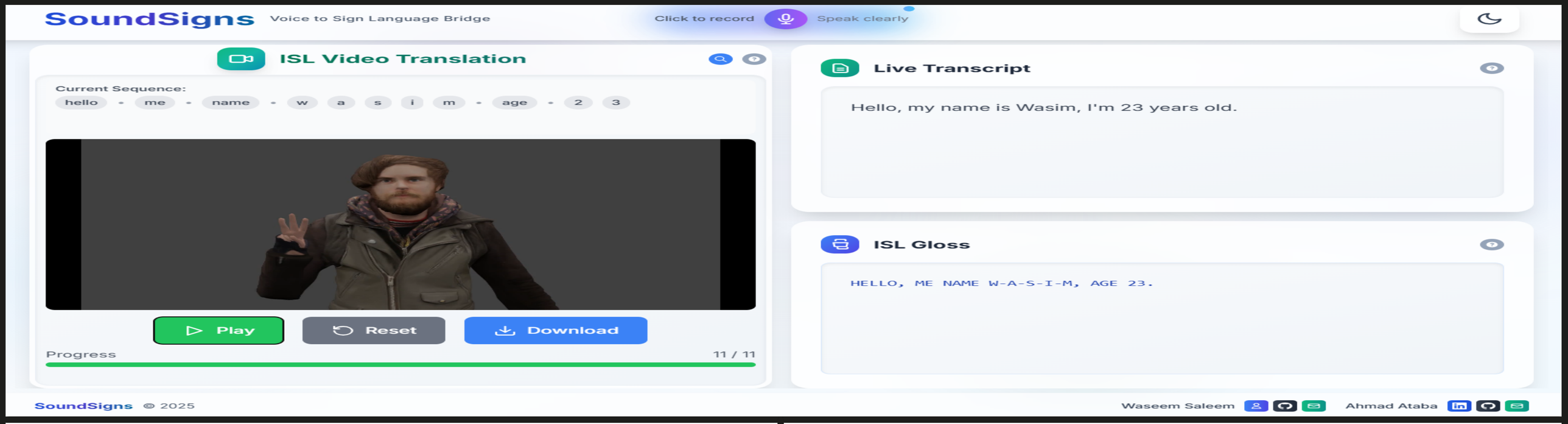
Experience Live: https://soundsigns.xyz

SCAN ME

SoundSigns: Speech To Sign Language Translator

Students: Ahmad Ataba & Waseem Saleem

Supervisor: Dr. Reuven Cohen



Background: The Problem Digital Exclusion

430M+ deaf/hard-of-hearing individuals lack access to spoken content

Subtitles Fall Short

Require English literacy; miss sign language's visual grammar

Interpreter Gap

Professionals unavailable for everyday content (vlogs, tutorials, social media)

Technical Barriers

Real-time motion capture systems are costly and impractical

Our Solution

Core Workflow:

Voice Input
Browser-based speech recognition (Web Speech API)

ISL Gloss Conversion
GPT-3.5 translates English → simplified ISL structure

Video Assembly
186 pre-rendered signs matched to gloss tokens
Seamless stitching of letters/digits/words

Output
3D avatar performs sign sequence with gloss highlighting

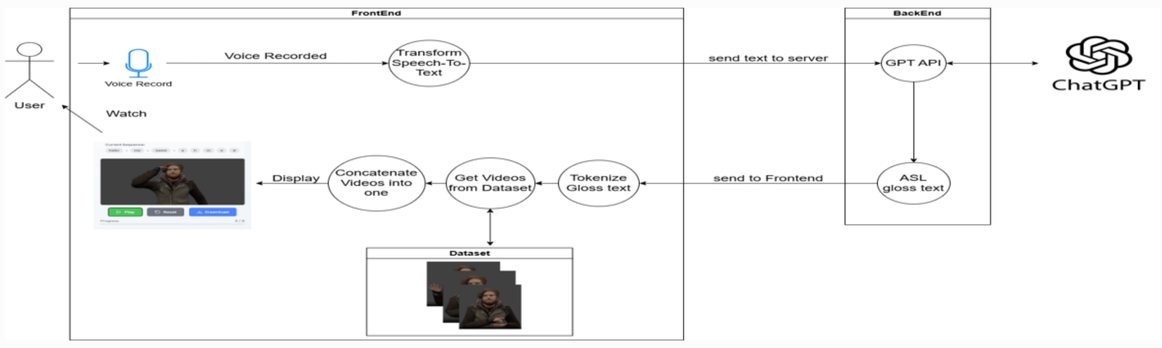
Key Technologies:

Frontend
React.js + Tailwind CSS

Backend
Flask/Python (ChatGPT API integration)

Dataset
Curated ISL videos (JS-Coderr)

Architecture And Dataflow



1. Voice Input Processing

- Frontend records user's speech via browser microphone
- Converts speech to text using Web Speech API

2. Text Translation

- Transcribed text sent to backend server
- ChatGPT processes text into structured ISL gloss

3. Sign Language Generation

- Frontend splits gloss into individual sign components
- Matches each component with pre-rendered videos
- Stitches videos into seamless sequence

4. Output Display

- Presents fluid sign language animation to user
- Highlights gloss tokens in sync with video playback

Results

Performance Highlights:

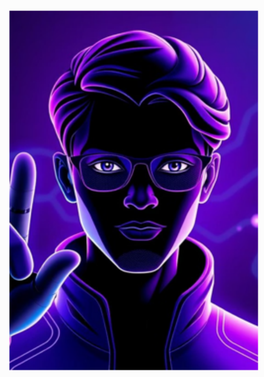
- 3-5 sec latency end-to-end processing
- 100% video matching for 150+ core signs
- Cross-browser support: Chrome, Firefox, Safari

Translation Examples:

English Input -> ISL Gloss Output
"What is your name?" YOU NAME WHAT
"I don't understand" UNDERSTAND ME NOT
"Walk despite rain" RAIN WALK SHE

User Impact:

- Real-time accessibility for digital content
- Downloadable videos for offline learning and sharing
- Educational transparency:
  - Gloss text display shows ISL syntax
  - Frame-synchronized highlighting teaches sign timing
- Zero-installation access: Fully browser-based solution



Development Challenges

Animation Roadblock:
Failed: Real-time tools (Kalidokit, SignAvatars)
Adopted: Pre-rendered video library

Dataset Scarcity:
Only 1 viable open-source ISL collection

Translation Limitations:
ChatGPT simplifies complex grammar
Lacks non-manual markers (facial expressions)

API Constraints:
Securing OpenAI keys
Web Speech API noise sensitivity

Future Work

Expand Dataset
500+ signs to reduce fingerspelling

Improve Translation
Dedicated ISL model training

Enhance Expressiveness
Add facial animation tracks

Mobile Optimization
Offline-capable PWA