HUAWEI

Spark Infinity

Huawei Developer Competition 2023

| July - October |

Challenge Name: Spark Infinity

Project Name: Sign - an ASL to English Translation App

Team Name: Barbenheimer

Organization: De La Salle University

Email/Phone: justin_jarrett_to@dlsu.edu.ph -

alyana_erin_bondoc@dlsu.edu.ph - loben_klien_tipan@dlsu.edu.ph -

andres_dalisay@dlsu.edu.ph

HUAWEI CLOUD
EVERYTHING AS A SERVICE

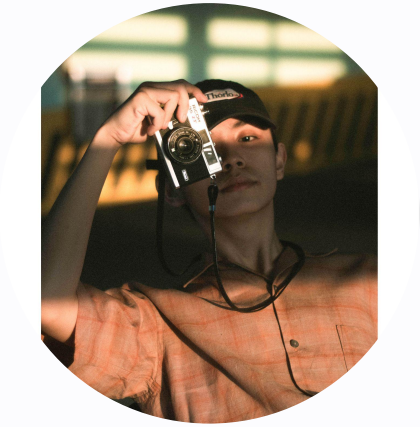
Overview

Project Overview

Project Name	*Sign: an ASL to English Translation App*
Team Name	*Barbenheimer*
Contacts	*justin_jarrett_to@dlsu.edu.ph, alyana_erin_bondoc@dlsu.edu.ph, loben_klien_tipan@dlsu.edu.ph, andres_dalisay@dlsu.edu.ph*
Technical Field	*Computer Vision, Machine Learning, Natural Language Processing (NLP), Mobile App Development, Cloud Computing, Data Annotation and Collection, Signal Processing, Algorithm Optimization, User Experience (UX) Design*
Technologies	*Cloud Container Engine (CCE), GaussDB*
Keywords	*Pose Estimation, Mobile App, AI, OpenCV, MediaPipe*
Applicable Fields	*Technology and Software Development, Accessibility and Inclusion, Retail and E-commerce, Hospitality and Tourism, Transportation and Mobility, Customer Support and Communication, Social Impact and NGOs, Language and Linguistics, Corporate Social Responsibility, Research and Development*
Description (in 500 words)	*Sign is an app that provides seamless ASL to English translation. It is designed to break down the communication barrier between deaf and hearing people, and to make it easier for everyone to communicate. It uses computer vision, machine learning, and natural language processing to translate American Sign Language (ASL) to text. The app works by first using a camera to capture the user's signing. The captured video is then analyzed by a computer vision algorithm to identify the hand gestures and facial expressions that make up the signs. The identified signs are then translated into text by a machine learning model. The app is designed to be used by both deaf and hearing people. The app is built on a cloud computing platform, which allows it to scale to meet the needs of a large number of users. The app also uses Huawei's proprietary technologies, such as Elastic Cloud Server (ECS), Cloud Container Engine (CCE) and GaussDB (for Mongo).* *The target audience for the app is deaf and hearing people who want to communicate with each other more easily. The market size for the app is estimated to be in the millions. The competitive landscape is relatively small, but there are a few other apps that offer similar functionality. The app's unique selling points include its accuracy, ease of use, and comprehensive dictionary of ASL signs. The app's monetization strategy is based on a subscription model and on partnerships.*

Contents

Team Introduction



Justin To

Leader / Designer / Developer

Andres Dalisay

Designer / Developer

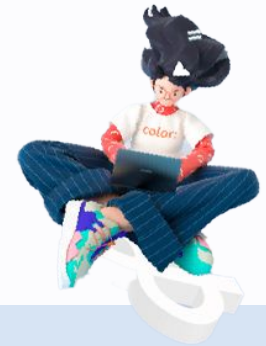
Loben Tipan

Planner / Developer

Alyana Bondoc

Planner / Developer

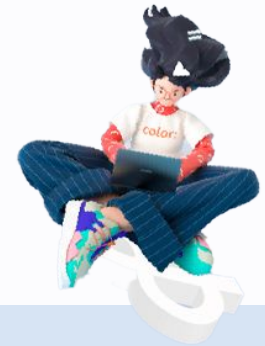
Computer Science students at De La Salle University - Manila, Philippines



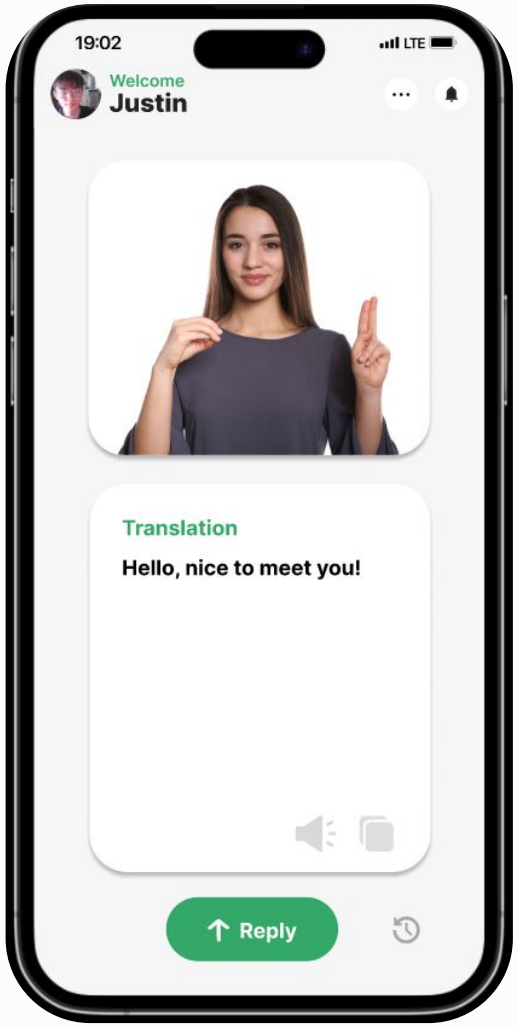
Project Introduction

Project Background:

The project involves creating a mobile application that utilizes computer vision and machine learning techniques to translate American Sign Language (ASL) gestures into English text. The goal is to bridge the communication gap between the deaf and hard of hearing community and the English-speaking population, promoting inclusivity and accessibility.



Project Introduction



Features:

- **Real-time Pose Estimation:** Accurately tracks users' ASL gestures through the device's camera.
- **Finger Tracking:** Detects and tracks individual finger movements for precise gesture recognition.
- **Instant Translation:** Translates recognized ASL gestures into English text in real-time.
- **Translation History:** Saves a record of translated phrases for users to review later.
- **User Profiles:** Allows customization and preferences for a personalized experience.
- **Camera Integration**: Users can access the device’s camera and direct it towards a subject. Once the subject starts performing ASL, their motion is translated into English text.
- **Copy and Paste:** Users can copy and paste the translated text for easy sharing and reference.
- **Voice Output**: By clicking or tapping a button, the translated text can be voiced out, enabling non-ASL users to hear the translation.

Project Introduction

- **Two-Way Communication:** Users can reply by typing or using voiceover. Their responses are translated into text for the deaf user to read.
- **Conversation History:** A history of the conversation is saved, allowing users to review previous interactions.
- **Communication Modes:** Users can select their preferred mode of communication, either “ASL” or “Voice,” catering to their comfort and preference.

Possible Additional Feature:

- **Video Call Translation:** Users can call another person, and using their camera, their ASL motion is translated into text for the other person. The recipient can respond through text, voiceover, or ASL, which is also translated into text for the deaf user.

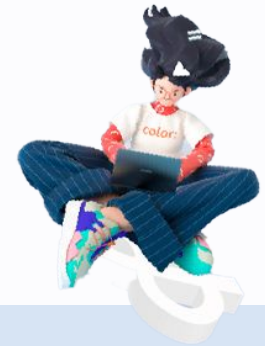
Project Introduction

Problems Solved:

- **Communication Barrier:** Overcomes the language barrier between ASL users and those who don't understand sign language.
- **Inclusivity:** Provides equal access to information and services for the deaf and hard of hearing community.
- **Real-time Interaction:** Enables fluid, real-time communication through gestures and text translation.
- **Learning and Education:** Supports ASL learning and raises awareness about sign languages.

Target Users:

- **Deaf and Hard of Hearing Individuals**: Primary users who benefit from the app's communication enhancement.
- **Retailers and Service Providers:** Use the app to better serve deaf customers and clients.
- **Healthcare Professionals**: Improve doctor-patient communication in medical settings.
- **Travel and Hospitality Industry:** Enhance interactions between deaf travelers and service providers.



Technical Architecture

Overview

The technical architecture of the app involves several key components:

- **User Interface (UI):** The app will have an intuitive UI where users can interact with the system. This includes features like input methods for ASL gestures and displaying translated text in English.
- **Pose Estimation Module:** This is the core of the app. It uses computer vision and machine learning techniques to analyze the user's ASL gestures captured through the device's camera. Pose estimation algorithms, such as OpenPose, would identify key points on the user's hands and body to understand the gestures accurately.
- **Machine Translation Engine:** Once the ASL gesture is recognized, the app will use a machine translation engine to convert the ASL representation into English text. This could involve using neural machine translation models like Transformers.
- **Integration with Device Camera**: The app will need access to the device's camera for real-time pose estimation. This requires camera API integration and handling camera permissions.
- **Cloud Service**

Technical Architecture

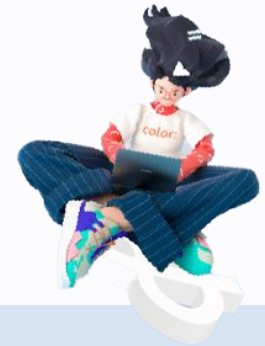
Key Technologies

- **App Platform:** The mobile application will be developed initially on Android with the goal of launching it on Huawei's AppGallery. With that said, the group has decided to use **Java**, one of the most popular languages among Android development. Its ease of use, wide range of capabilities, and large community makes it clearly the most appropriate choice for this project.
- **Post Estimation Module Server:** This is the module that will handle all App API requests and services that will utilize the Pose Estimation Module. It will be in charge of feeding the model with appropriate visual input, retrieving the established keypoints, and analyzing the keypoint movement and translating it from ASL to English. It will also be storing translations in the appropriate GaussDB table. This will be developed using **Python** and libraries such as **Flask**. This microservice will also be running on an independent container in Huawei Cloud's CCE.
- **Pose Estimation Module:** The group has decided to make use of **OpenCV** and **MediaPipe** for the Pose Estimation Module. OpenCV is an open-source library packed with computer vision tools that will help boost the Pose Estimation model's accuracy and performance. MediaPipe consists of Machine Learning tools that are built on top of TensorFlow Lite. It also includes Pose Estimation models and capabilities. Both of them support Python platforms, making it the obvious choice. Not to mention, it also has extensive support and community, which will help speed up the development

Technical Architecture

Key Technologies

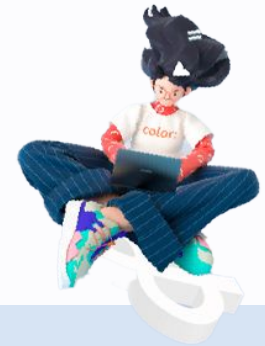
- **User Authentication Server:** This is the module that will handle all App API requests and services with regards to User Authentication. It will be in charge of signing in users by querying their credentials in the appropriate GaussDB table and validating its existence, as well as creating new users by inserting the user-specified credentials. This will be developed using **Python** and libraries such as **Flask**. This microservice will also be running on an independent container in Huawei Cloud's CCE.
- **Conversation History Server:** This is the module that will handle all App API requests and services with regards to viewing conversation/translation history. It will be in charge of querying previous translations based on user credentials and specified parameters in the appropriate GaussDB table. This will be developed using **Python** and libraries such as **Flask**. This microservice will also be running on an independent container in Huawei Cloud's CCE.



Technical Architecture

Key Technologies

- **Database:** This project will be making use of Huawei Cloud's GaussDB which allows for the integration of complex distributed databases that will help boost app performance and improve its scalability and reliability. It also is relational which can be used to to store the conversational history and translated messages between users, as the app aims to do. It'll also be used to store account information and user data as well as data needed by the pose estimation model to operate accurately.



Technical Architecture

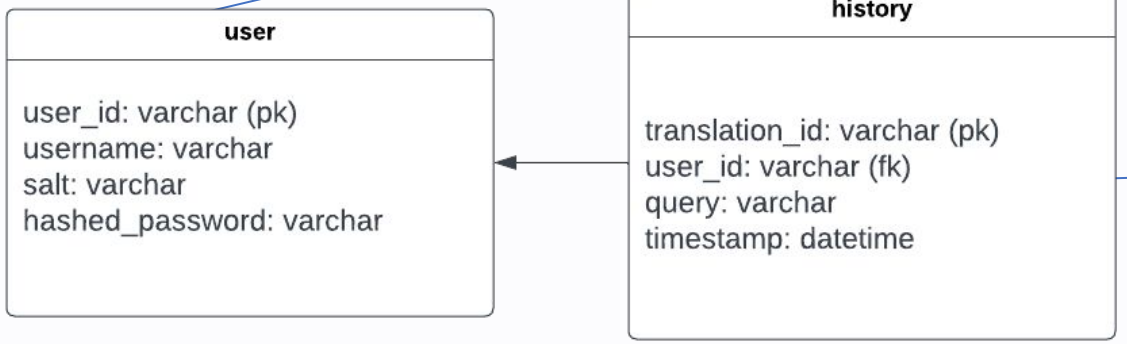
Cloud technologies

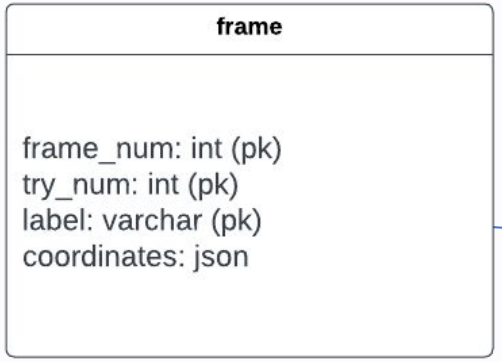
The proposed application makes use of the following Huawei Cloud Technologies:

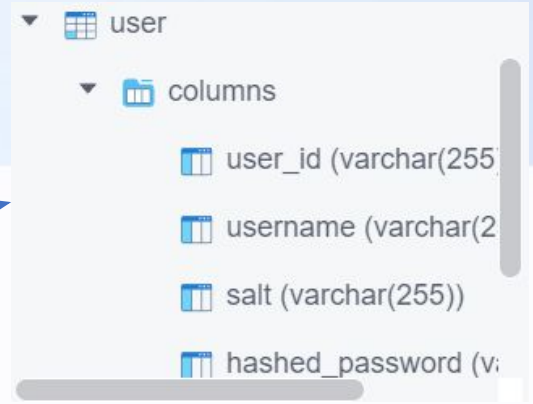
- **Cloud Container Engine (CCE):** Considering the application's proposed microservice architecture, the group believes that it would be most suitable to make use of CCE's kubernetes and containerized features to deploy them individually. This would heavily improve application scalability, both vertically and horizontally, as well as simplicity to develop thanks to its ability to host multiple containers, which allows microservices to be developed or modified independently. It also offers load balancing and automatic scaling based on defined rules.
- **GaussDB (for MySQL):** This project will be making use of Huawei Cloud's GaussDB which allows for the integration of complex distributed databases that will help boost app performance and improve its scalability and reliability. It is also relational which can be used to to store the conversational history and translated messages between users, as the app aims to do. It'll also be used to store account information and user data as well as data needed by the pose estimation model to operate accurately. This data can be quite complex, thus, GaussDB seemed to be the most appropriate choice due to its capabilities of handling complex data.

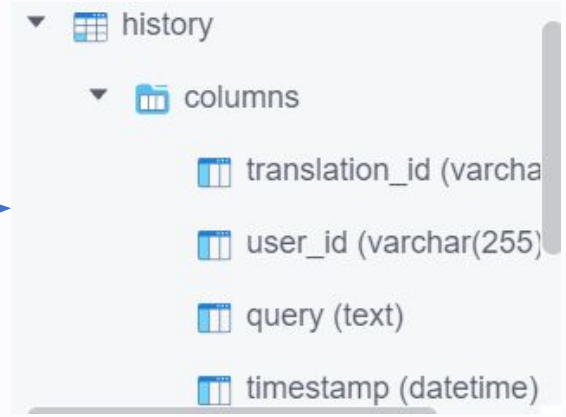
Technical Architecture

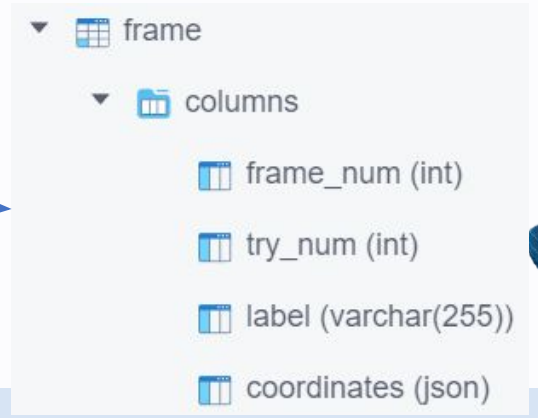
- **GaussDB (for MySQL):**
  - Database Schema





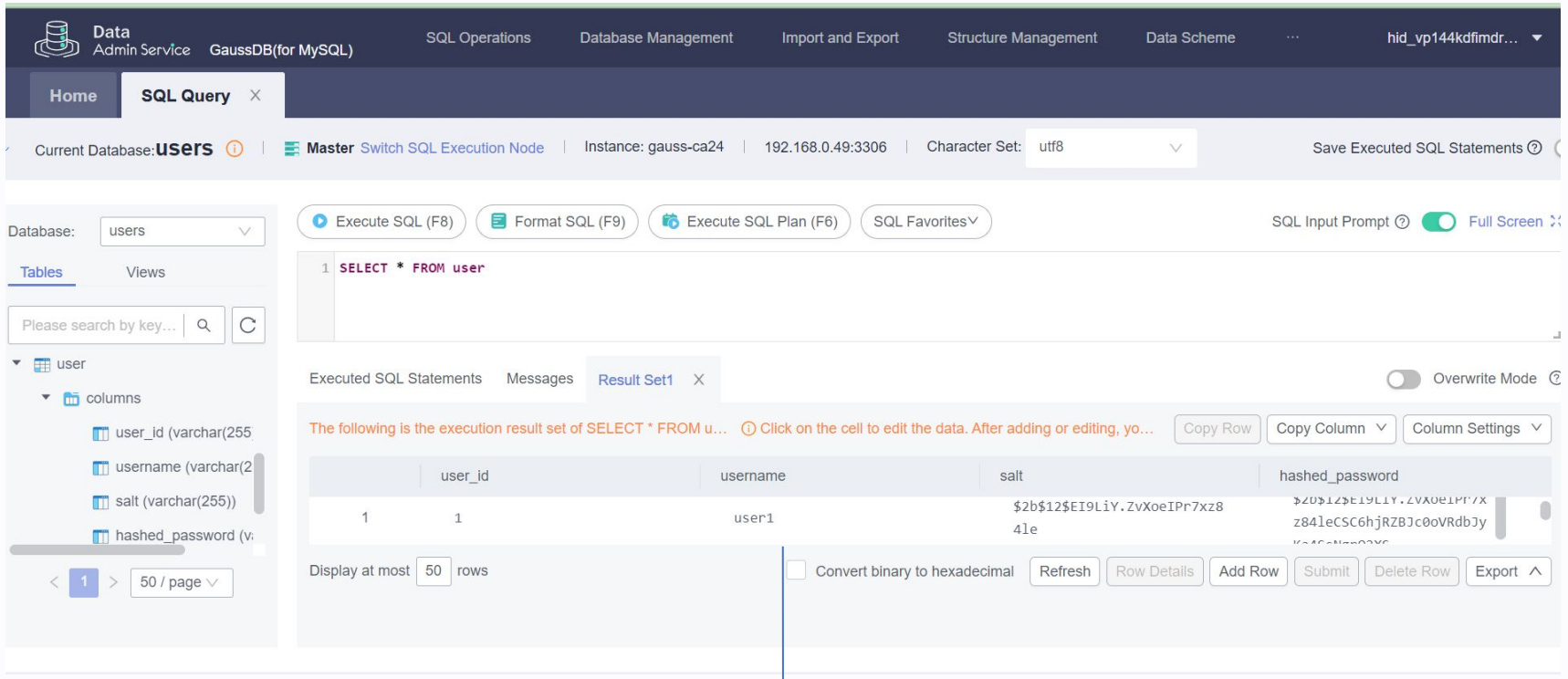






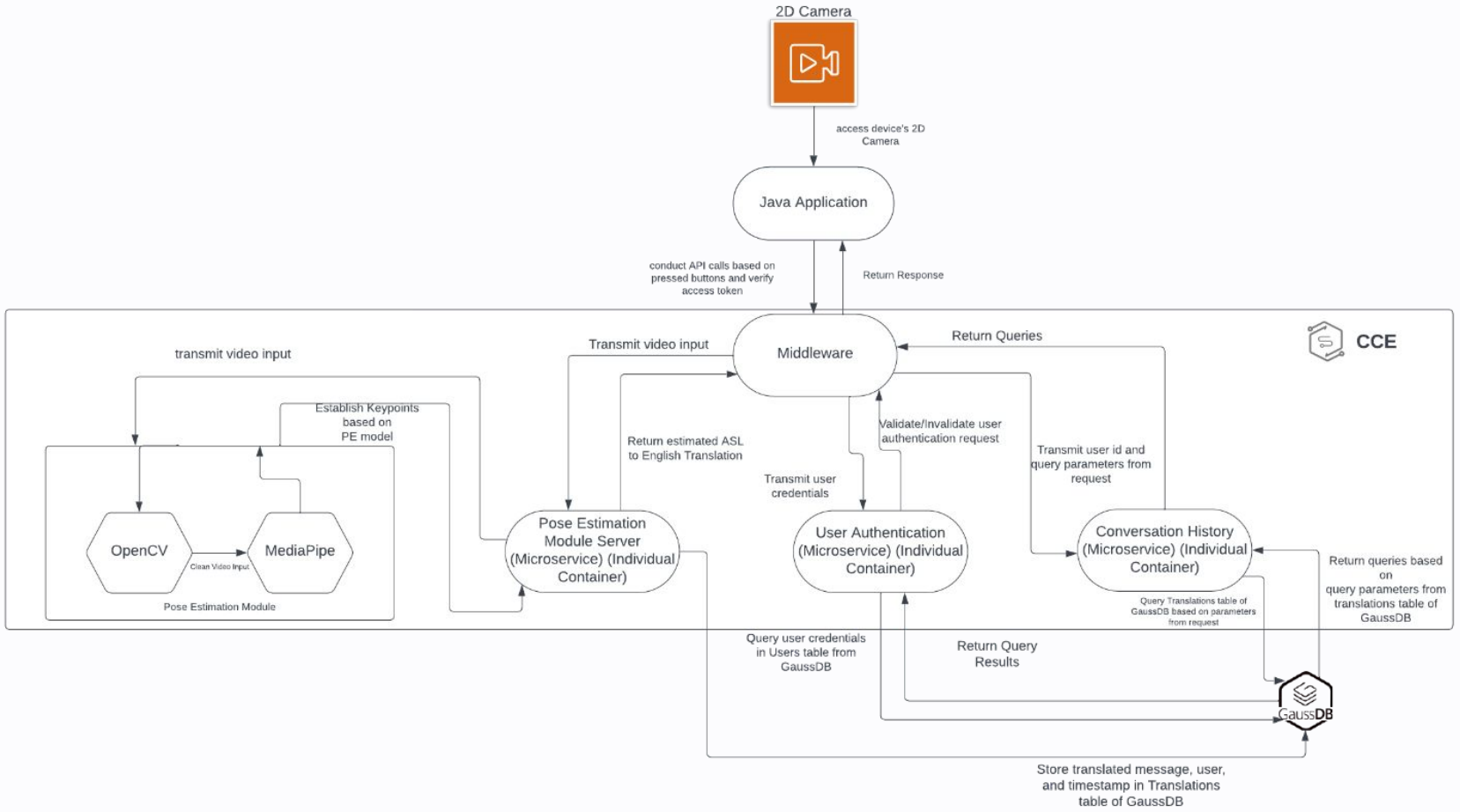
Technical Architecture

- **GaussDB (for MySQL):**
  - implementation of schema in GaussDB



user_id	username	salt	hashed_password
1	user1	$2b$12$EI	$2b$12$EI9LiY.ZvXoeIPr7xz84leCSC6hjRZBJc0oVRdbJyKa4SsNgrQ2XS
2	user2	$2b$12$Jp	$2b$12$JpNE8PujiqSjjzNNHgQNkuvBtUPVTST6awhDZCy0UVeXH.LkFXVti
3	user3	$2b$12$T5	$2b$12$T5DKd4NmXlaPK1YQBd8MXO.kUPSRGKY7AoUGySVTFxyt.G7IHpB.2

Technical Architecture

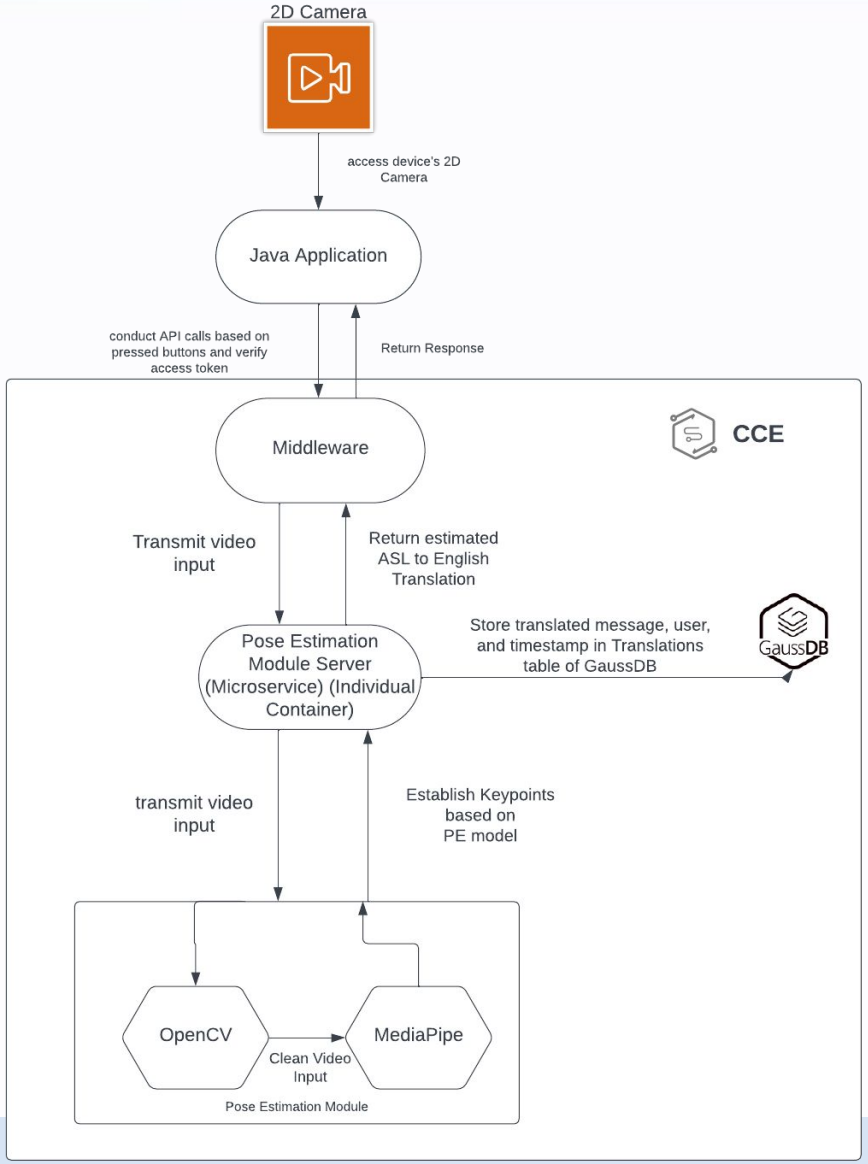


Service Description

To the left is a general overview of the software's technical architecture.

1. The device's 2D camera is accessed by the flutter application
2. Based on button pressed, an API request is sent to the appropriate microservice container in the CCE.
3. API request is authenticated by Middleware
4. Response is returned to and handled by Java Application

Technical Architecture

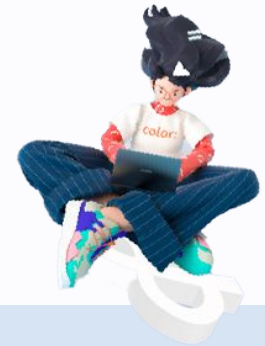


Service Description

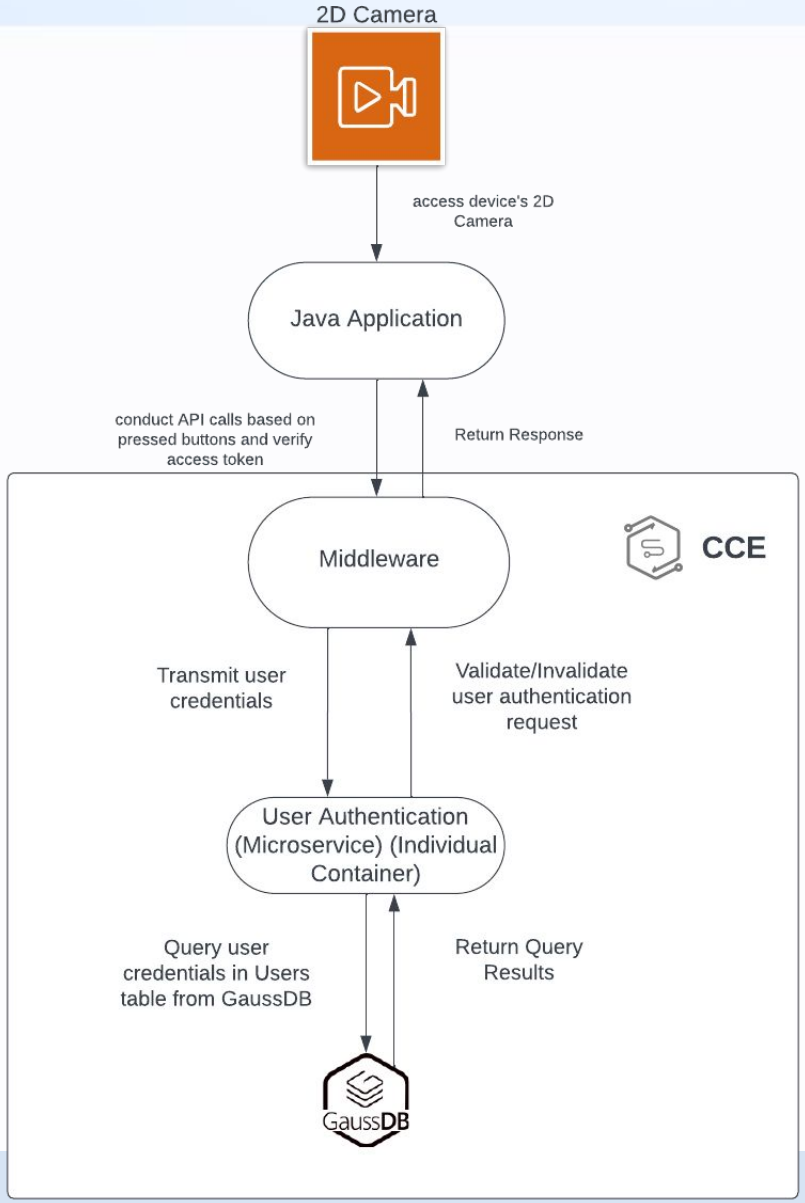
To the left is a general overview of the software's technical architecture.

If accessed function is ASL to English Translation, the Pose Estimation Module Server Container in the CCE is used:

1. Video Input is sent to Pose Estimation Module
2. Video Input is cleaned and filtered by OpenCV
3. Keypoints are established based on MediaPipe's Pose Estimation model
4. Keypoint movement is translated into English text by logic in Pose Estimation Module Server
5. Translation is returned to flutter app UI for display



Technical Architecture



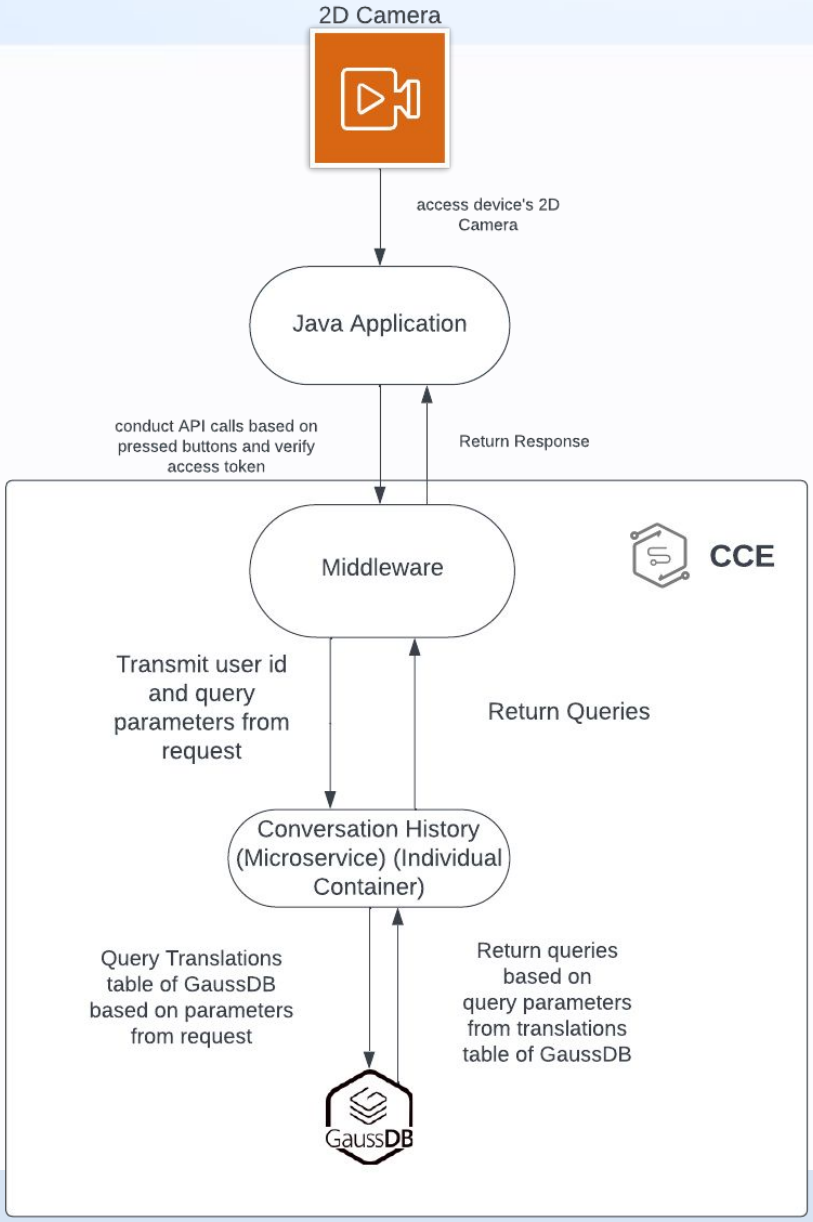
Service Description

To the left is a general overview of the software's technical architecture.

If accessed function is User Authentication,
the Pose Estimation Module Server is used:

1. User credentials are sent to the User Authentication Microservice
2. If Sign In, credentials are used to query the Users table from GaussDB. If Sign Up, credentials are inserted into the table.
3. Query results are returned
4. Validate/Invalidate user authentication based on query results

Technical Architecture

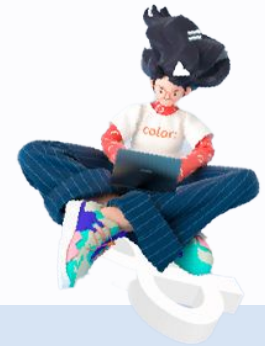


Service Description

To the left is a general overview of the software's technical architecture.

If accessed function is Viewing Conversation History, the Conversation History Server Container is used:

1. Transmit User ID and Query Parameters from Request
2. Query Translations table of GaussDB based on request parameters
3. Return query results based on parameters
4. Return results to Java Application for display



Functions

Core Functions:

- **Real-time Gesture Recognition:** Accurately tracks users' ASL gestures through the device's camera.
- **Finger Tracking:** Detects and tracks individual finger movements for precise gesture recognition.
- **Instant Translation:** Translates recognized ASL gestures into English text in real-time.
- **Two-Way Communication:** Users can reply by typing or using voiceover. Their responses are translated into text for the deaf user to read.

Advantages and Uniqueness:

Sign introduces an entirely novel approach to communication, tailored exclusively for the deaf and hard of hearing community. Its groundbreaking features redefine interaction for ASL users, offering capabilities that have never been seen before:

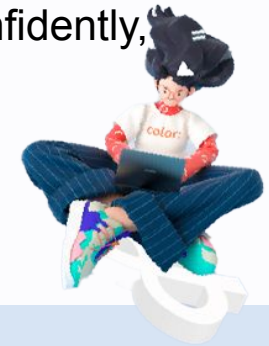
- **Real-time Precision:** Powered by state-of-the-art pose estimation and finger tracking, our app hopes to achieve unprecedented accuracy in recognizing ASL gestures, revolutionizing the reliability of ASL translation.

Functions

- **Natural Conversations:** Unlike any other solution, our app will facilitate authentic two-way communication. ASL users can engage in dynamic conversations, and their responses seamlessly convert to text for easy comprehension.
- **Inclusive Audibility:** Our voice output feature will transform translated text into audible speech, ensuring fluid communication between ASL users and those unfamiliar with sign language.
- **Historical Insight:** Empowering continuity, the translation history feature will empower users to review past conversations, fostering deeper understanding and seamless follow-up.

Elevated User Experience:

More than just a collection of features, our app embodies an entirely new communication experience. We seamlessly blend cutting-edge technology with empathetic design, creating a platform where ASL users communicate confidently, and non-ASL speakers engage seamlessly.



Innovations

Global Impact:

More than *70 million people* worldwide are deaf, with over 80% of them residing in developing countries. Our app's innovations hold the potential to break down communication barriers for this underserved population, fostering inclusivity and connectivity on a global scale.

By revolutionizing communication accessibility for the deaf and hard of hearing community and achieving unprecedented advancements in real-time translation, gesture recognition, and user experience, our ASL to English Translation App is poised to redefine the landscape of communication technology and inclusivity.

A Pioneering Breakthrough:

While there have been notable advancements in assistive technology, our app stands out as one of the first apps to enable real-time conversion of ASL gestures into English text. This visionary breakthrough redefines how ASL users interact and communicate with the world around them.

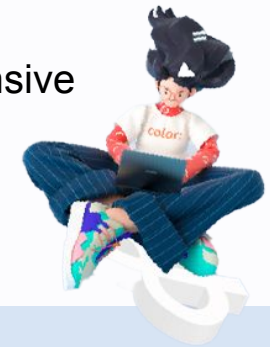
Business Value

Business Value:

- **Accessibility:** Our app bridges the communication gap between the deaf/hard of hearing community and the wider English-speaking population.
- **Market Demand:** There is a growing demand for inclusive technologies, and our app can cater to this underserved market.
- **Data Collection:** With user consent, we can collect anonymized data to improve pose estimation and translation algorithms, enhancing the app's accuracy over time.
- **Partnerships:** Collaborate with educational institutions, NGOs, or speech therapy centers to provide specialized versions of the app for specific user groups.

Market Potential and Differentiation:

- The market for ASL translation apps is estimated to be in the millions. Our app has the potential to reach a large number of users, both in the United States and around the world.
- Sign stands out among other apps offering similar functionality due to its accuracy, ease of use, and comprehensive ASL sign dictionary. This differentiation positions Sign as a leader in the market.



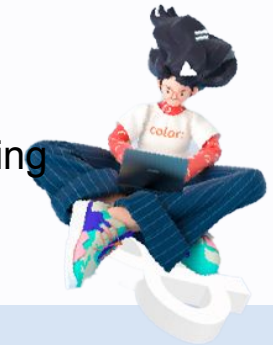
Business Value

Unique Selling Points of Sign:

- **Accuracy:** Our app employs state-of-the-art computer vision and machine learning algorithms to ensure precise translations.
- **Ease of Use:** Designed to be user-friendly for both deaf and hearing individuals, the app boasts a simple interface without requiring special training.
- By making ASL more accessible, our app contributes to breaking down barriers between the deaf and hearing communities, fostering inclusivity and a more connected society.

Potential Partners:

- **Retail and E-commerce:**
  - In-Store Assistance: Enhance shopping experiences by enabling store employees to communicate with deaf customers using ASL.
  - Product Information: Help deaf users access product details through ASL gestures, improving understanding.
- **Restaurants and Fast-Food Chains:**
  - Drive-Thru Ordering: Facilitate accurate orders for deaf customers using ASL in drive-thru interactions.
  - Tableside Communication: Empower waitstaff to understand and respond to ASL-based requests, enhancing dining experiences.



SIGN BUSINESS MODEL

Key Partners

- Collaborative partnerships with educational institutions, NGOs, speech therapy centers, retailers, restaurants, healthcare institutions, transportation services, and more.
- Collaborative relationships with data collection partners for enhancing algorithm accuracy.

Key Activities

- Continuous research and development for algorithm enhancement and feature expansion.
- App maintenance and updates to ensure seamless functionality and reliability, including regular monitoring of Huawei Cloud services to ensure optimal performance.
- Partnership establishment and management with various organizations and businesses.

Key Resources

- Cutting-edge computer vision and machine learning technologies for accurate gesture recognition and translation.
- A skilled and diverse development team proficient in software engineering, machine learning, and user experience design.
- Data collection infrastructure for refining algorithms and improving translation accuracy over time.
- Huawei Cloud services for hosting and maintaining the app's backend infrastructure and data storage.

Value Propositions

- Bridging the communication gap between the deaf/hard of hearing community and the English-speaking population, fostering inclusivity on a global scale.
- Meeting the increasing demand for inclusive technologies in various sectors, positioning our app as a pioneer in this underserved market.
- Leveraging anonymized data collection to enhance algorithms, ensuring continuous improvement in pose estimation and translation accuracy.
- Offering tailored app versions through partnerships with educational institutions, NGOs, and speech therapy centers to cater to specific user groups (possible future feature).

Customer Relationship

- Establishing collaborative relationships with a diverse range of organizations, businesses, and institutions.
- Providing comprehensive customer support for technical assistance, inquiries, and partnership management.
- Offering personalized support and guidance for individual users who want to use the app, ensuring a user-friendly onboarding process and addressing any queries they may have.

Channels

- Direct engagement with potential business partners through presentations, proposals, and partnership agreements.
- Leveraging digital channels such as app stores and social media, to acquire and onboard users.
- Collaborative efforts with advocacy groups and institutions within the deaf/hard of hearing community.

Customer Segments

- Deaf and hard of hearing individuals seeking seamless communication solutions.
- Businesses across sectors, including retail, restaurants, healthcare, transportation, and financial institutions, aiming to enhance accessibility and inclusivity.
- General users seeking accurate ASL to English translation for personal or educational use.

Cost Structure

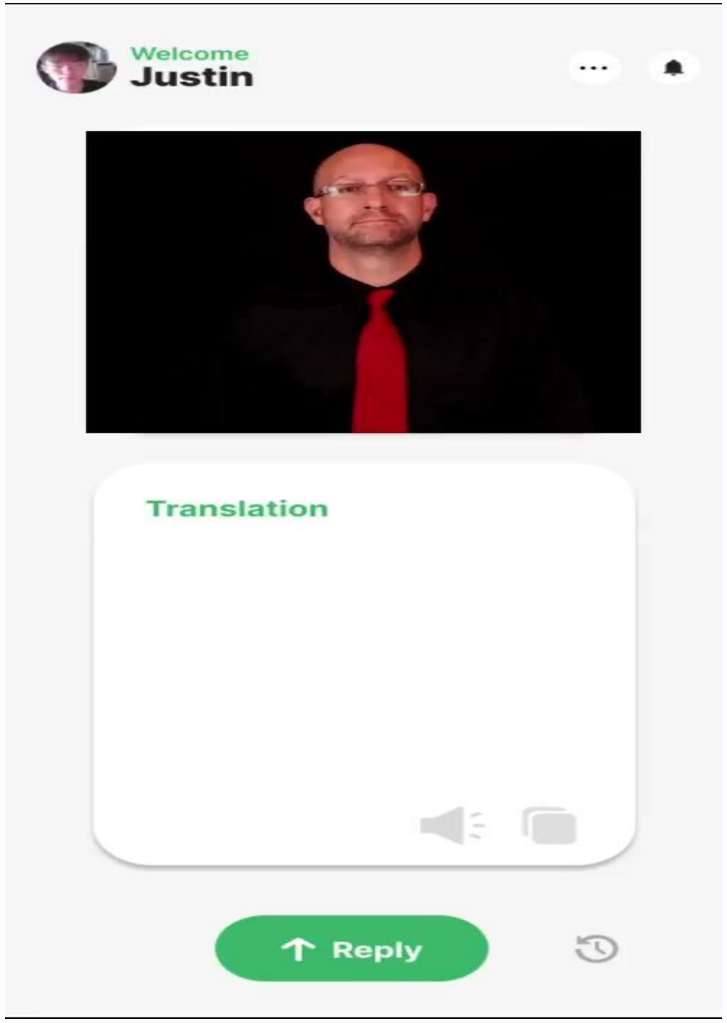
- Research and development investments for algorithm improvement and feature expansion.
- Salaries and resources for the technical team responsible for app maintenance, updates, and support.
- Marketing and partnership expenses for promoting the app and establishing collaborations.

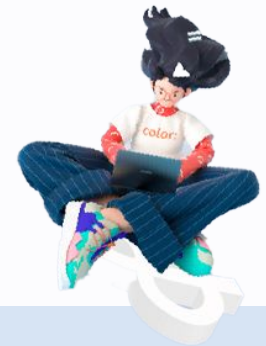
Revenue Streams

- Freemium model offering basic translation features for free.
- Premium subscription model providing access to advanced features.
- Potential revenue sharing from collaborative partnerships with organizations and businesses for tailored app versions.
- Proceeds from sales of goods and services, including one-time customer payments and service revenues generated by providing customer support.
- Monetization of anonymized data collected with user consent for algorithm enhancement and research purposes

Demo & Follow-up Plan

- **Demo:**
  - click to play

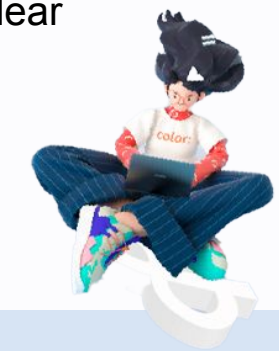




Demo & Follow-up Plan

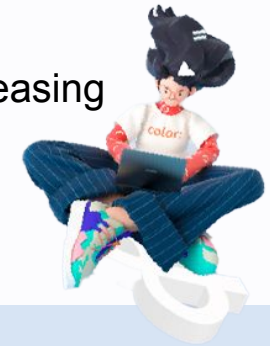
Follow-Up Plan:

- **Overall:**
  - Ongoing model retraining to continuously improve accuracy and adapt to evolving sign language variations.
  - Enhancement of models through the use of larger and more diverse datasets, aiming to increase vocabulary accuracy and overall performance.
  - Implementation of user accounts to provide personalized experiences and track user preferences.
  - Implementation of online features, including real-time translation during calls, to broaden the app's usability.
- **Objectives and Difficulties in Each Phase, Key Points, and Solutions:**
  - **Phase 1: Project Planning**
    - Objective: Define the app's concept, value proposition, and project plan.
    - Challenges: Ensuring a clear understanding of user needs and market demands.
    - Key Points: Conduct thorough market research, assemble a dedicated project team, and establish clear project milestones.
    - Solutions: Research online resources, and create a well-defined project timeline.



Demo & Follow-up Plan

- **Phase 2: Working Prototype, First Version**
  - Objective: Develop a working prototype and seamlessly integrate with Huawei services.
  - Challenges: Ensuring accurate ASL to text translation while maintaining real-time functionality.
  - Key Points: Leverage existing research, utilize necessary libraries and tools, and secure essential resources.
  - Solutions: Collaborate with experienced professors for technical guidance, conduct rigorous iterative testing, and refine the integration with Huawei services.
- **Target Market Volume and Prospect; Strategies for Market Development and Expansion:**
  - **Market Volume:**
    - Deaf and Hard of Hearing Individuals: There are over 70 million deaf individuals worldwide, representing a substantial and underserved user base.
    - Businesses: Partnering with businesses like retailers, restaurants, customer support centers, and e-commerce platforms can expand our market volume. These businesses can benefit from using Sign to improve communication with deaf customers.
  - **Market Prospect:**
    - Inclusive Technologies: The global market for inclusive technologies is on the rise, driven by the increasing awareness of accessibility needs and the desire for inclusivity. Sign aligns perfectly with this trend.



Demo & Follow-up Plan

- Partnerships: Collaborating with businesses in various sectors, such as retail, hospitality, and customer support, opens doors to market expansion. These partnerships create opportunities for tailored versions of Sign to address specific user groups.
- Customer Engagement: Our app can enhance customer engagement for businesses that serve deaf customers. This could lead to improved customer satisfaction and loyalty.
- Data Collection: With user consent, we can collect anonymized data to enhance the accuracy of the app's translation algorithms. This continuous improvement can further drive market growth.

- **Market Expansion Strategies:**
  - Educational Institutions: Partnering with educational institutions can help introduce Sign as a valuable tool for deaf students and educators.
  - NGOs and Speech Therapy Centers: Collaboration with organizations focused on the deaf community can facilitate the distribution of Sign to those who need it most.
  - Corporate Social Responsibility (CSR): Emphasize CSR initiatives by showcasing how businesses can support inclusivity through the use of Sign. This aligns with corporate values and social responsibility efforts.

Overall, the market volume for Sign is substantial due to the large number of potential users, and the prospects are promising as businesses and organizations recognize the importance of inclusivity and accessibility. Strategic partnerships and a user-centric approach will be key to realizing the app's market potential.

Demo & Follow-up Plan

- **Steps and Timeline:**

project planning, project brainstorming, finalizing app concept and value proposition

Phase 1

Phase 2

working prototype, first version, integration with huawei services

Regression testing, continuous retraining of model, app improvements from feedback from judges

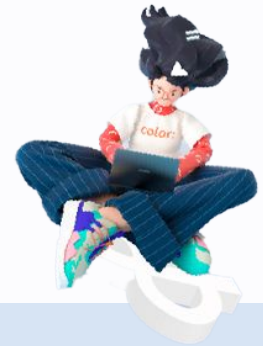
October

Deployment Phase

Public release and deployment on the huawei app store

Implementation of online features, including real-time translation during calls

version 2



HUAWEI

THANK YOU