

# Problem Description and Goals

## Mechatronics & Software

---

### *Group 3*

Ding (Chris) Hao  
Yiming (Jane) Pang  
Yucheng Yao  
Wenyu (Winnie) Yin  
Taoming Yu  
Xiang (Shawn) Zhang

*Table 1: Revision History*

Date	Developers	Change
September 22, 2024	Group 3	The first version of the document

## Project Description

The rise of globalization causes an increased demand for understanding different languages in various human activities, such as tourism, business, and politics. In the past, people relied on paper dictionaries that were often heavy, inefficient and lacked audio pronunciation. Today, while AI translation and mobile translation apps are available, they come with their own challenges: AI translation often requires significant hardware resources, making it less portable, and translation apps can be time-consuming due to the need for inputting words manually.

This capstone project aims to design and develop a portable, intelligent, and convenient language translation tool for written text. Specifically, creating a product that will capture desired text from camera vision, extract the text in a specific language, and provide users with an accessible interface for interacting with the translated text.

# Stakeholders

## Main Stakeholders

1. Language-learners
  - Language learners need to use efficient and convenient language translation devices to read novels, follow up on news, and practice for language tests so that they can expand their vocabulary quickly and understand the contextual usage of the words.
2. Travellers
  - While travelling in foreign countries, people have the need to use portable and intelligent devices to help understand foreign language text, such as road signs, restaurant menus, etc.
3. Professionals Working in International Environments
  - Professionals facing international projects need to have an intelligent and efficient device to quickly interpret written documents, presentations, or public and political information in other languages.

## Potential Stakeholders

1. Language Data Trainer and Model Provider
  - With an open-source spirit, our project could promote more recent, efficient, and accurate language models. These stakeholders would ensure the translation system is accurate and up-to-date with various languages and dialects.
2. Video Game Players
  - Many video games do not have versions available in all languages, so video game players would have the need to use a fast and convenient language translation device to fully enjoy games that are only offered in foreign languages.

# Goals

## Primary Goals

1. Develop a mobile phone application that offers a comprehensible presentation of the translated language to intended users without occupying significant storage spaces.
  - Achieving this goal can minimize the hardware requirements of users' phones (especially storage and GPU), making the system more accessible. The client application with less storage occupation allows continuous updates/improvements without requiring users to download large patches.
2. Construct or utilize a machine learning model/platform that offers fast text extraction from images with high accuracy.
  - Speed and accuracy in translation ensure the application's smooth interactions and strong practicability without significant delays or the need for user verification caused by translation inaccuracy, improving user satisfaction.
3. Design and build a functional and intelligent wearable device with lightweight cameras to capture images of interest.
  - Our solution streamlines the translation process by eliminating the need for users to manually input foreign language text and wait for results. Users can instantly view translations directly on their phone when needed, reducing unnecessary human operations and improving overall efficiency.
4. Implement a durable battery.
  - A lasting battery enhances the user's experience by providing extended usage without interruptions, such as frequent charging or maintenance.
5. Create a well-structured interface that allows fast and reliable data transmission between the software backend and the camera.
  - A well-structured interface design is easier to maintain and scale. This will help solve existing bugs and minimize the risk of introducing new bugs when adding features, consequently protecting the user experience.

## Stretch Goals

1. Provide additional language learning-related features to the app (e.g., vocabulary library, level classification, pronunciation guide, etc.)
  - Additional features would be a beneficial add-on to our project since one of our target user groups is language learners, who may expect more tools other than simple translation. These functionalities will significantly improve our users' learning experience.