2023 Fall Project for Data Science

Final Submission

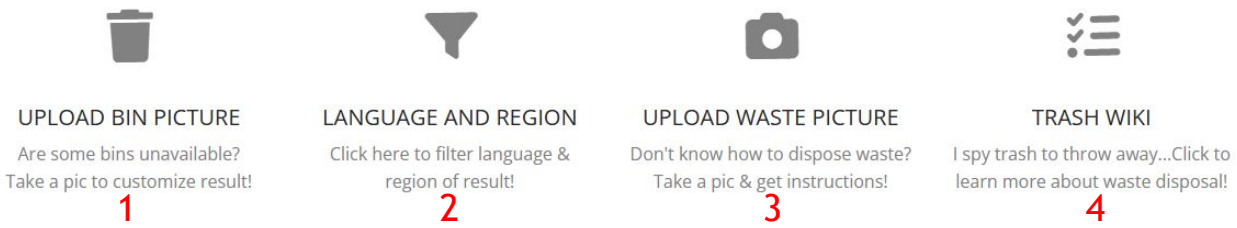
P4DS Team 6: EcoLens

Chaeyun Kim, Seunghoon Yi, Doeon Lim, Seoyun Jang, Dounia Anis

Enhancements (Success)

1. Improve UI/UX design [Seoyun]

- Reorder & redesign buttons to improve clarity of usage flow.

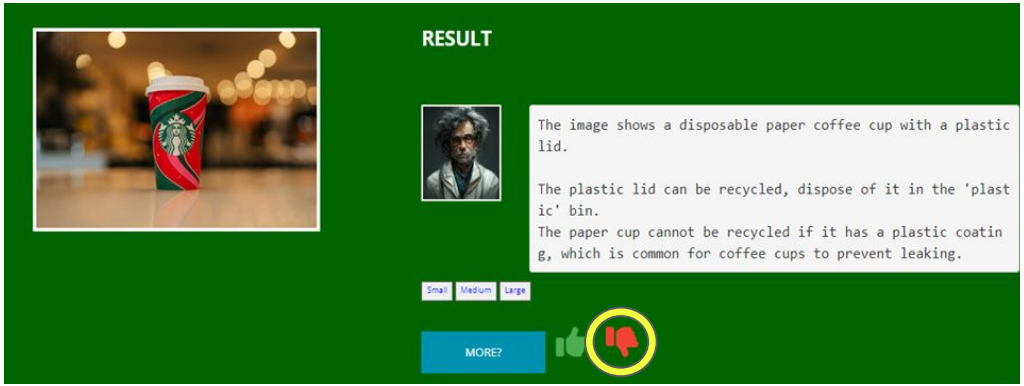


- Enhance the visibility of recyclable items and their descriptions by resizing fonts.

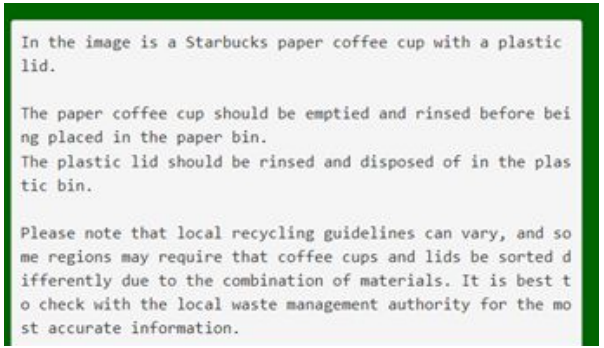
Enhancements (Success)

2. Implement User Feedback function [Doeon]

- Users can leave feedback on the answer result.
- The model will use received feedback to enhance accuracy in future outputs.



Click this button to leave negative feedback!



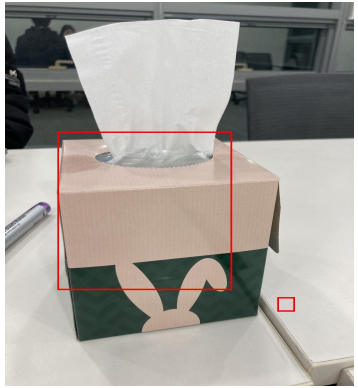
Enhancements (Attempted)

3. RAG [Seoyun, Chaeyun]

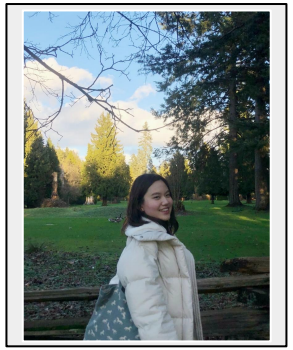
- RAG (retrieval augmented generation) technique enables model to lookup external information to improve the response generation.
- Tried RAG with LangChain & ChromaDB to enhance accuracy of instructions output.. but resulted in lower accuracy & higher latency.
- Failed possibly due to..
  - Crawled documents of recycling guidelines were inaccurate / non-exhaustive
  - Need to use additional captioning model which causes bottleneck

4. Circling target objects [Chaeyun]

- Initial Attempt: GPT-4 prompted to circle objects - suboptimal.. :(
- Proposed Solution: Utilize an object detection model :
  - Extract bounding box coordinates and cropped images
  - Feed cropped images into our model
  - identify and select the object a user wishes to locate



Team Members



Seoyun Jang

- PO (Product Owner)
- Presentation
- UI/UX design

Chaeyun Kim

- QA, ML modeling
- DE, Backend
- Presentation

Doeon Lim

- ML modeling
- Backend engineer
- UI/UX design

Dounia Anis

- UI/UX design
- Backend engineer
- Presentation

Seunghoon Yi

- QA, ML modeling
- Prompt engineering
- Presentation

Table of contents

01. Background

- Motivation
- Product Idea
- Target Users
- Value Analysis

02. Architecture

- Overview
- Frontend
- Model / Backend

03. Implementation

- UI / UX
- Required Environment
- Demo Video

04. Conclusion

- Cost Estimation
- Challenges
- Future Works
- Project Management
- Team Members

Background

Motivation

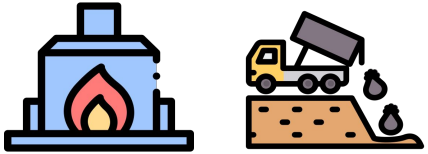
439 kg

of waste per person per year *

↓

85.5%

of them are willing to recycle **



20.2 million ton

of waste incinerated or landfilled per year *

↓

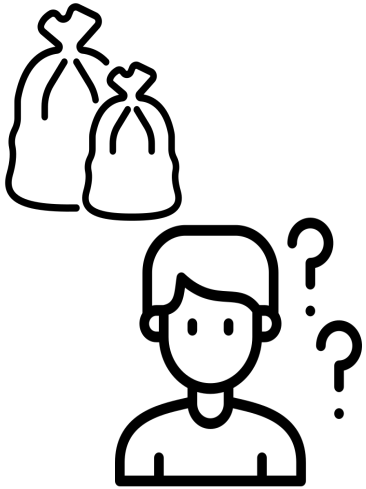
46%

of them could have been recycled.***

* Korea Waste Association
** Pyeong-taek city citizen survey
*** Korea Resource Circulation Information System

Motivation

People don’t know how to properly recycle because of..



- ❗ Counterintuitive cases, especially in items with mixed materials
- ❗ Different disposal methods by country, region, and settings
- ❗ Ineffective, non-personalized education

Plastic or general waste?

Where to dispose plastic bag?

Non-Korean speakers: ???

Product Idea

A recycling information app that can..

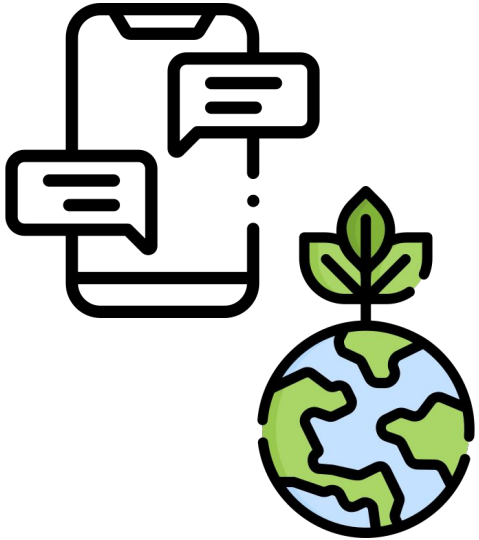
Given an image of garbage(s),

✅ explain correct disposal methods of each object

✅ customize the response in respect to user's circumstances

- user's language and current region
- currently available bins

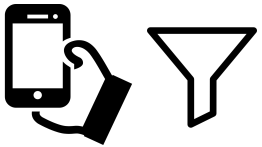
✅ incorporate user feedback



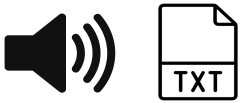
If people easily can access quality information on how to recycle, then the Earth will be greener!

Product Overview

User



User Input
: image of trash & available bins
user's language / region

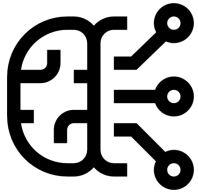


Model Output
: instructions in text & audio

Service page

Feedback

Collected information



Model

"Plastic sauce container goes in the plastic bin.

Make sure to rinse it to remove any food residue before recycling."

Instructions

Target Users

All individuals

“Only 3 bins available in my building.. how should I recycle all these?”

✅ multiple object detection & classification
✅ bin customization

Foreigners/ Travelers

“Recycling guidelines here are too complicated & hard to understand!”

✅ language & region customization

Children

“ I want to learn how to recycle in a fun & easy way!”

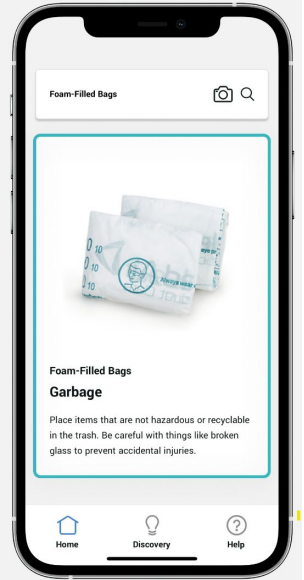
✅ hands-on learning
✅ audio generation
✅ Trashwiki for more fun activities

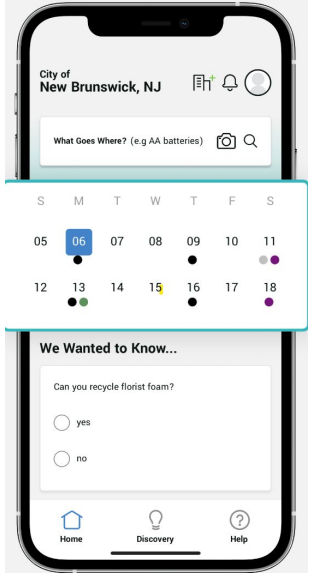
Value Analysis

Recycle Coach

👍 Provides recycling information by user location, custom calendar for recycling habits

👎 All settings and searches are manual, responses are error-prune, does not support visual search

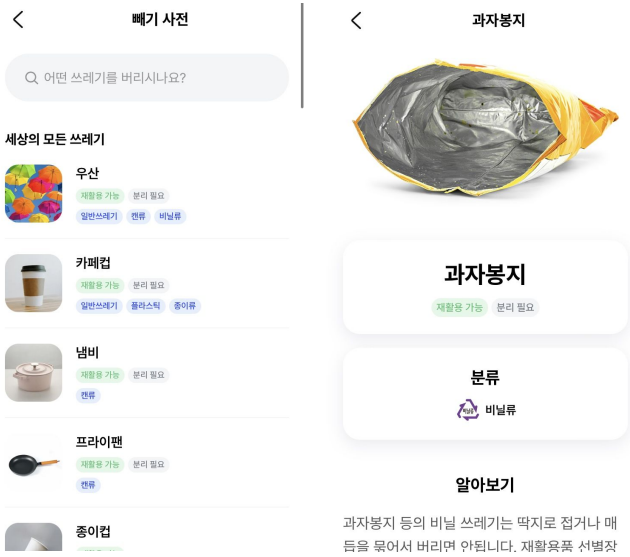




Bbae-ggi

👍 Has a custom encyclopedia of recycle guidelines for a wide range of materials

👎 Searches are also manual and less flexible, does not support visual search



Value Analysis

EcoLens

👍 Provides recycling information by **any user location** and (nearly) **any situation**

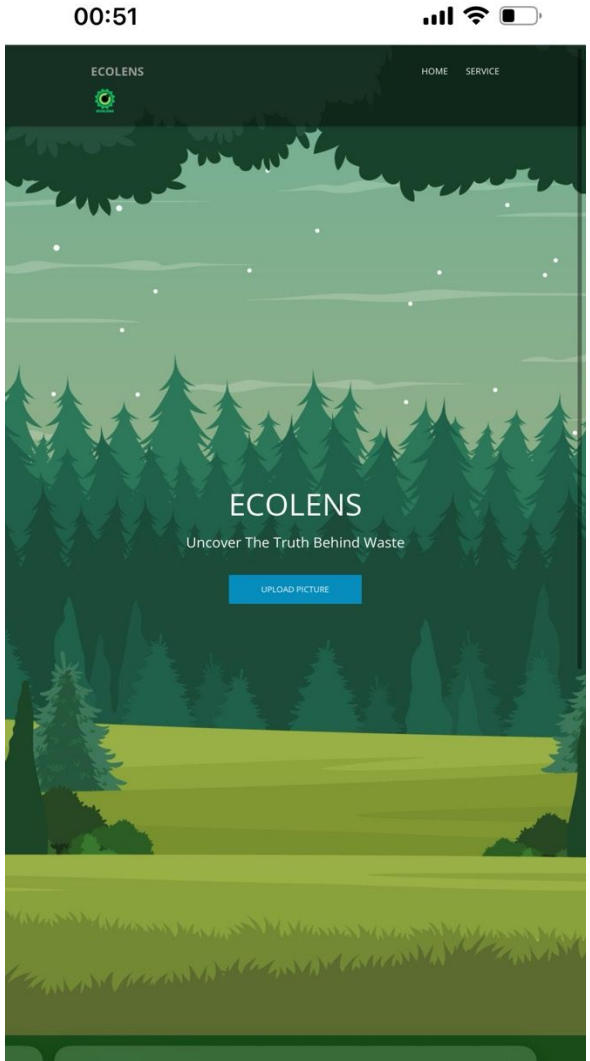
👍 Supports visual search and customized responses

Our Social Values

- A **1-year** use of this app by an **single user** is expected to reduce more than **20Kg** of waste
  (if the user properly recycle **only 1/10 of the daily waste** using this app)

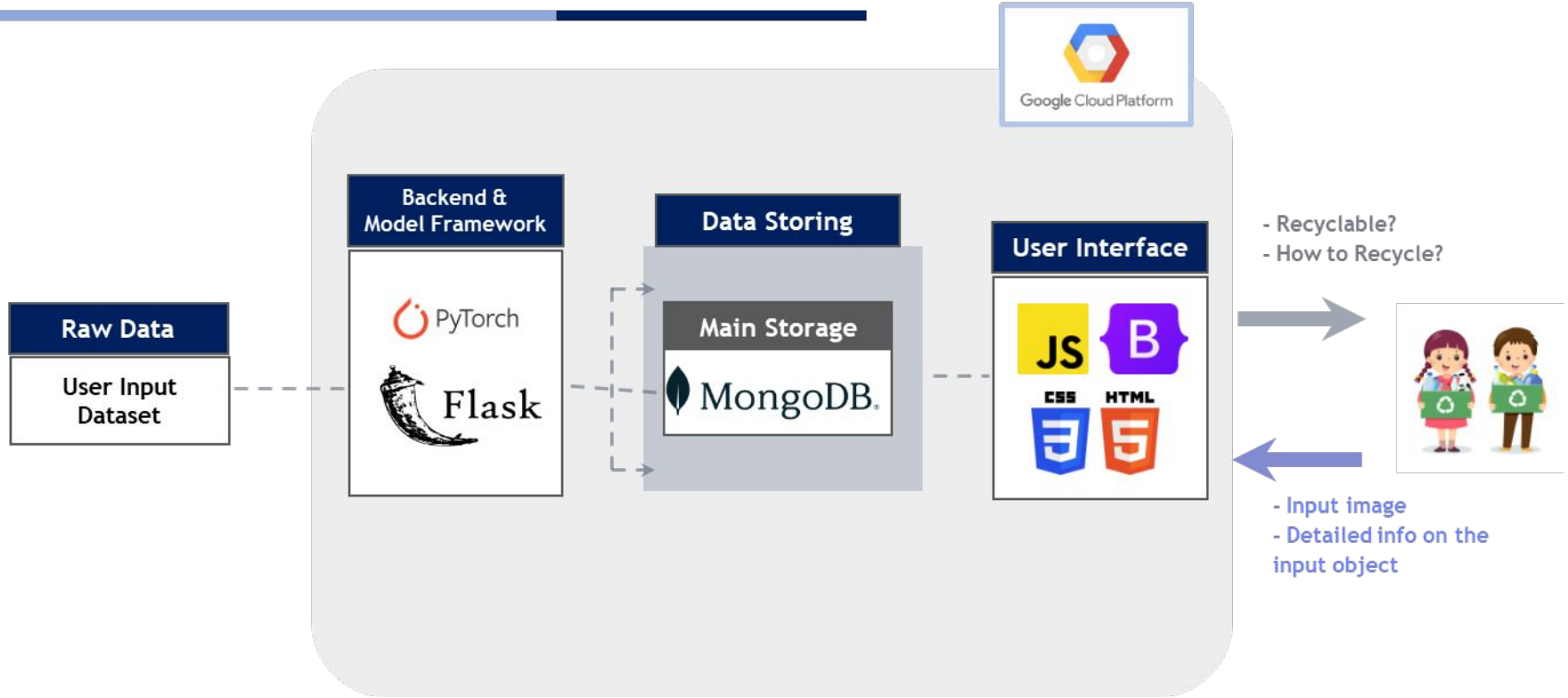
- Assuming a stable user of 100,000 individuals, it could **reduce 2,000 tons of waste annually**

…and this is **PRICELESS.**



Architecture

Overview


Google Cloud Platform
Raw Data
User Input Dataset
Backend & Model Framework
PyTorch
Flask
Data Storing
Main Storage
MongoDB.
User Interface
JS
B
CSS
HTML
- Recyclable?
- How to Recycle?
- Input image
- Detailed info on the input object

Frontend : Main Changes

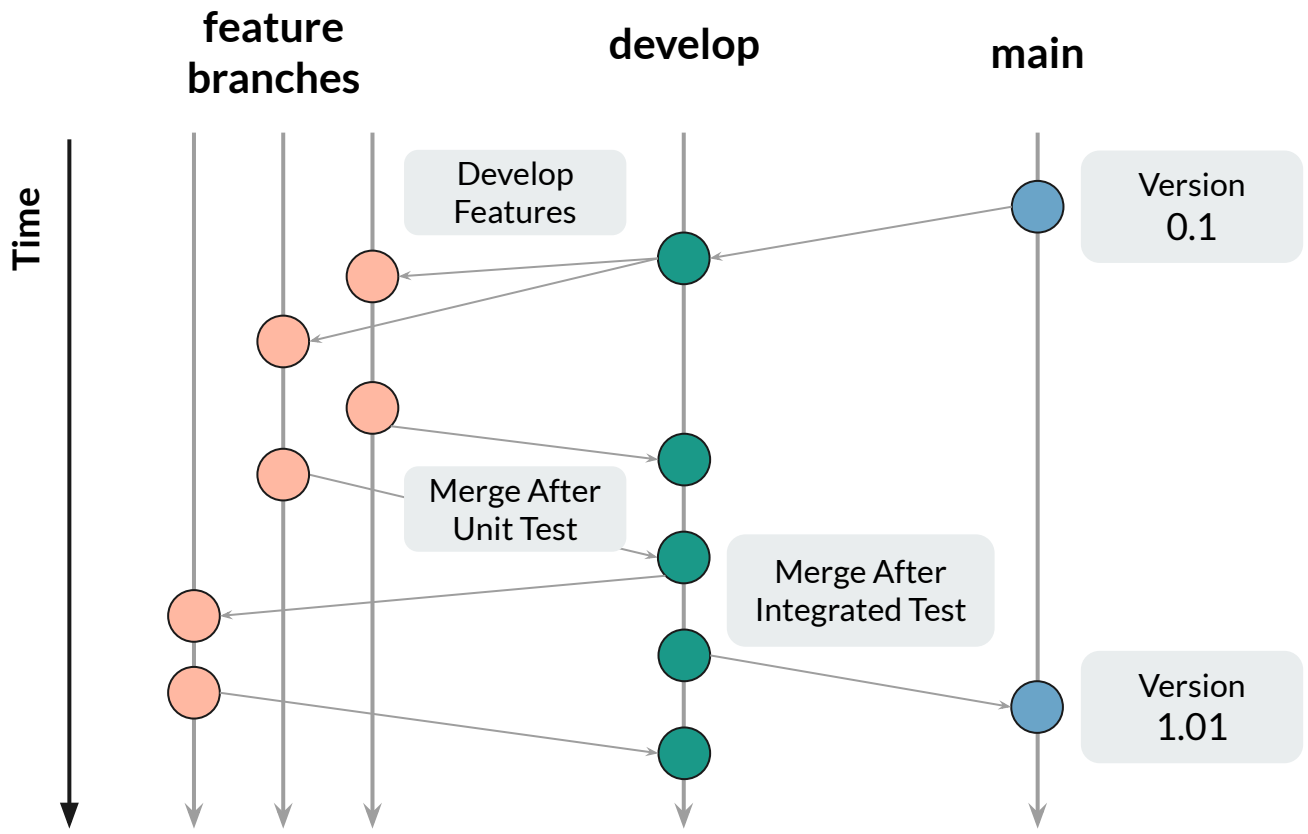
React + Flask Pipeline

- ❌ Using Too Many Third-Party Libraries
- ❌ Overly Complex State

HTML + CSS + JavaScript

- ✅ easy collaboration with other members

Development : Git Flow & multiple branches



Model / Backend

Input

OpenCV

Bin and object detection

Flask

PyTorch

Custom Model

GPT-4V API

Initial Response

From the model

- Object labels
- Bin labels

From the user

- Location
- Language

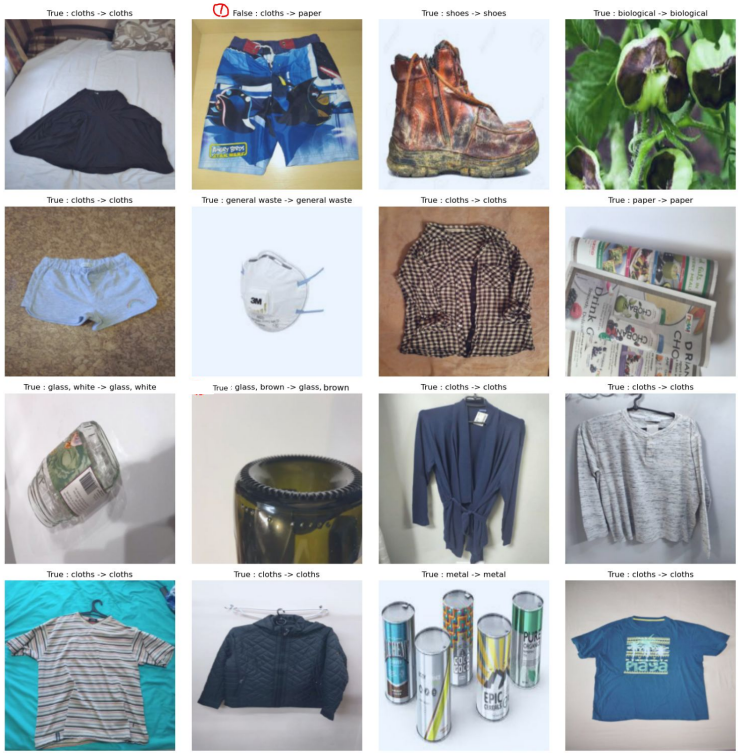
LLM/TTS APIs

Tailored Instructions

Model / Backend

Model selection & Training

- Backbone : EfficientNetV2(Pretrained on ImageNet)
- Finetuning on :
  - **classification** : Garbage classification dataset
  - **detection** : Domestic trash, TACO
- Achieved **95.5% accuracy** in classification task

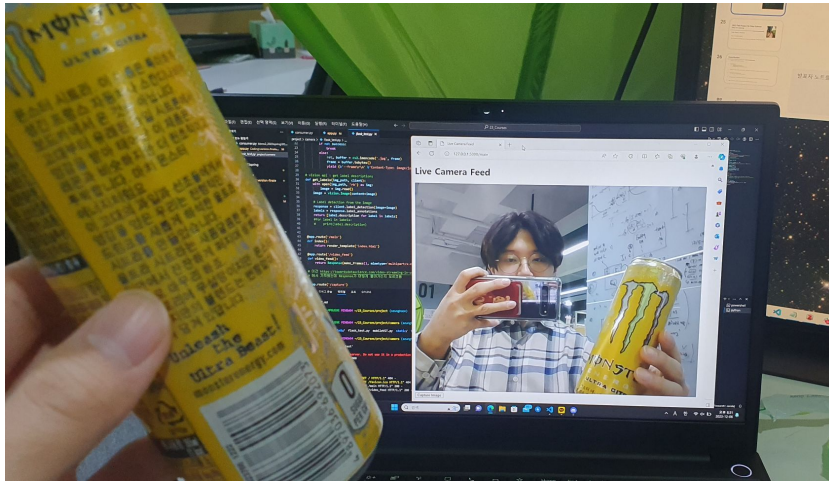


👎 Poor in multiple-object detection and some edge cases

Model / Backend

Image capture & inference system

- Utilized **OpenCV+python** for **imaging**
- Inference module sends photo to **GPT-4V API**
- Both functions are embedded in a simple web application using **Flask**



"This is a image with a person holding a drink can in one hand, and a cell phone in the other."

👍 **Extremely precise, can detect >5 objects at once!**

Model / Backend

Prompt Engineering for GPT-4

- Set a general prompt, and additional prompts that are generated by user interactions(request about additional information.. etc)

General case prompt

```
"How do I recycle the object(s)in the image? You have to
identify each item in the image and determine whether it can be
recycled as accurate as you can.If the object is not
recyclable, state '{'{item_name}'} cannot be recycled.' \
If yes, specify the bin type that the wastea must be disposed,
such as '{'{item_name}'} goes in {bin_types} bin.' \
I am currently in  {region}, and the available bins are
{detected_bin_types}. The total number of bins in front of the
user is {bin_count}. Please respond in {language}."
```

Tailored Instructions

Additional prompt (example)

```
"I have a question. {question}? Please explain step by step."
question example : "Why is the headphone not recyclable?."
```

Model / Backend

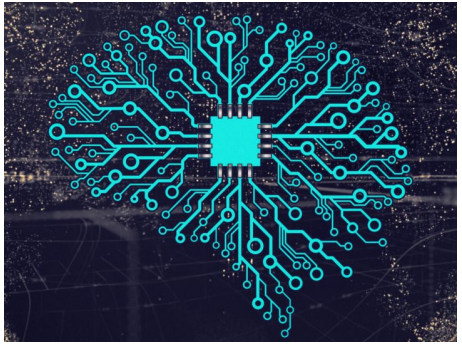
GPT TTS API (Text-to-speech)

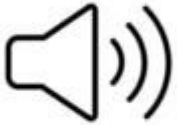
- To make users understand more easily and **improve accessibility**, we provide a TTS option for the given instructions.

Tailored Instructions

1. The cup lid goes to the plastic bin.
2. The cup holder goes to the paper bin.
3. For the cup, check whether the beverage is left. If there is leftover beverage, wash the cup and dispose to general waste.
4. The plastic bag behind goes to general waste. Also, check whether there is leftover items inside.

Whisper TTS



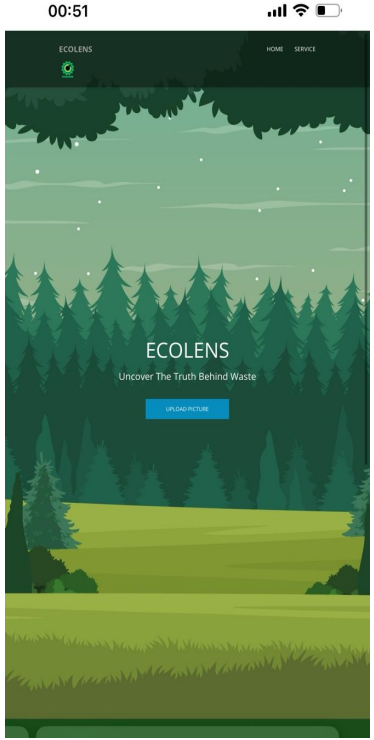


Implementation

UI / UX : Thema

1. **Eco-reminding, user-friendly Interface**
   - Environmental : Green Forest
   - User-Friendly : Modern design with few clickable buttons

2. **Main functionality**
   - Trash analysis :
     Upload Picture => give result of how to recycle based on region, language,etc…



Required Environment

Hardware

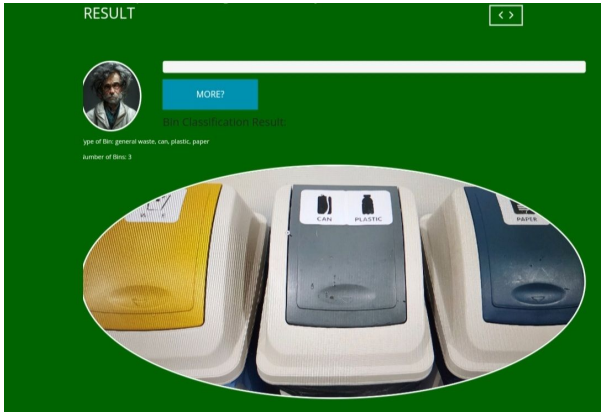
- A basic smartphone or laptop with..
  - Camera (over 1 megapixels)
  - Free memory with more than 500MB to open app/browser
  - Connection to the internet to retrieve inference results

Software

- Browser, Microsoft IE or Google Chrome
- If to run the application on local device,
  - one must install python and pytorch on the development requirement section

Main Functions

Main Functions (1) Check Valid Recycle Categories

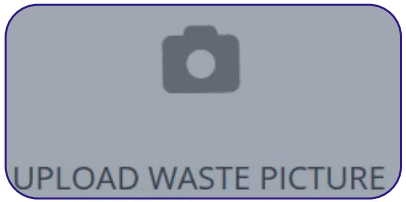


Response

Type of bin : General Waste, Plastic, Paper
# of bins : 3

Are some bins unavailable?
Take a pic to customize result!

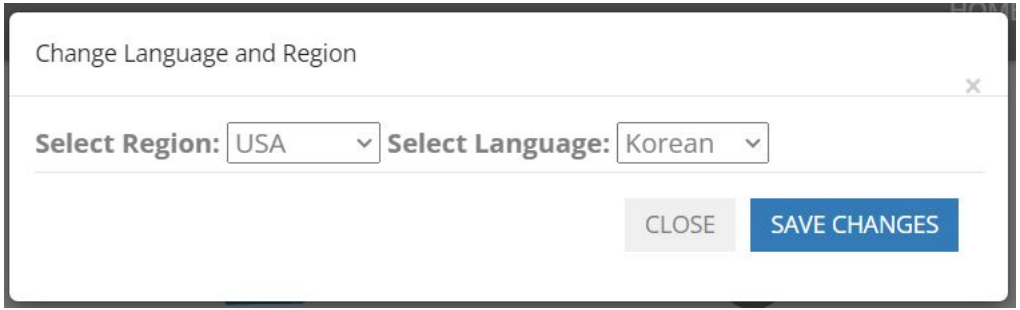
Click here to filter language & region of result!



Don't know how to dispose waste?
Take a pic & get instructions!

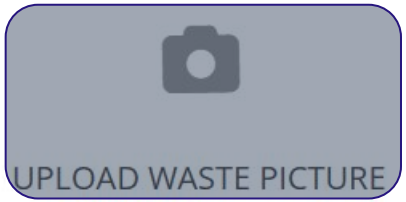
I spy trash to throw away...Click to learn more about waste disposal!

Main Functions (2) Filter Language & Region



Are some bins unavailable?
Take a pic to customize result!

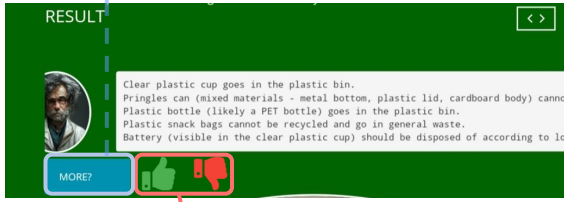
Click here to filter language & region of result!



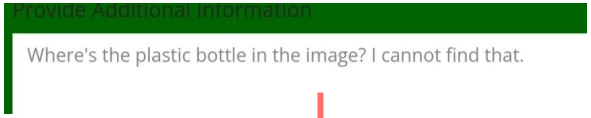
Don't know how to dispose waste?
Take a pic & get instructions!

I spy trash to throw away...Click to learn more about waste disposal!

Main Functions (3) Upload Waste Picture



Provide additional info



plastic bottle you asked about is located behind the Pringles can

Personalized QA +
More detailed explanation

User Feedback available

Are some bins unavailable?
Take a pic to customize result!

Click here to filter language & region of result!

Don't know how to dispose waste?
Take a pic & get instructions!

I spy trash to throw away...Click to learn more about waste disposal!

Main Functions (4) Trash Wiki

UPLOAD BIN PICTURE

Are some bins unavailable?
Take a pic to customize result!

LANGUAGE AND REGION

Click here to filter language & region of result!

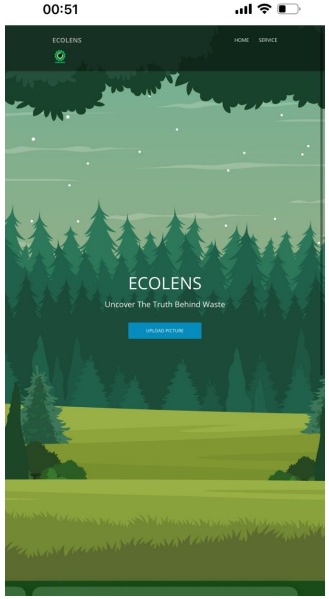
UPLOAD WASTE PICTURE

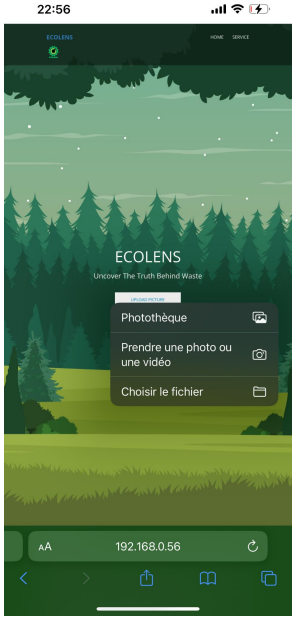
Don't know how to dispose waste?
Take a pic & get instructions!

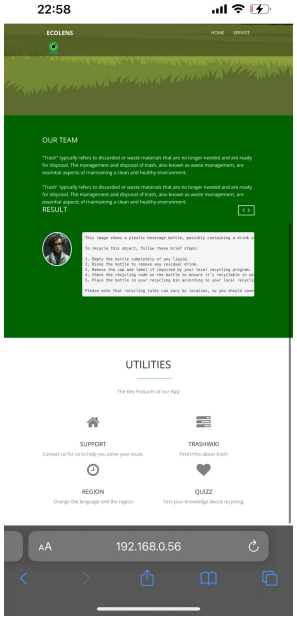
TRASH WIKI

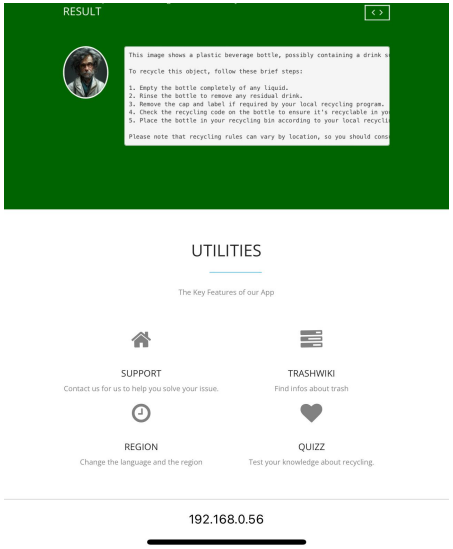
I spy trash to throw away...Click to learn more about waste disposal!

Service Pages

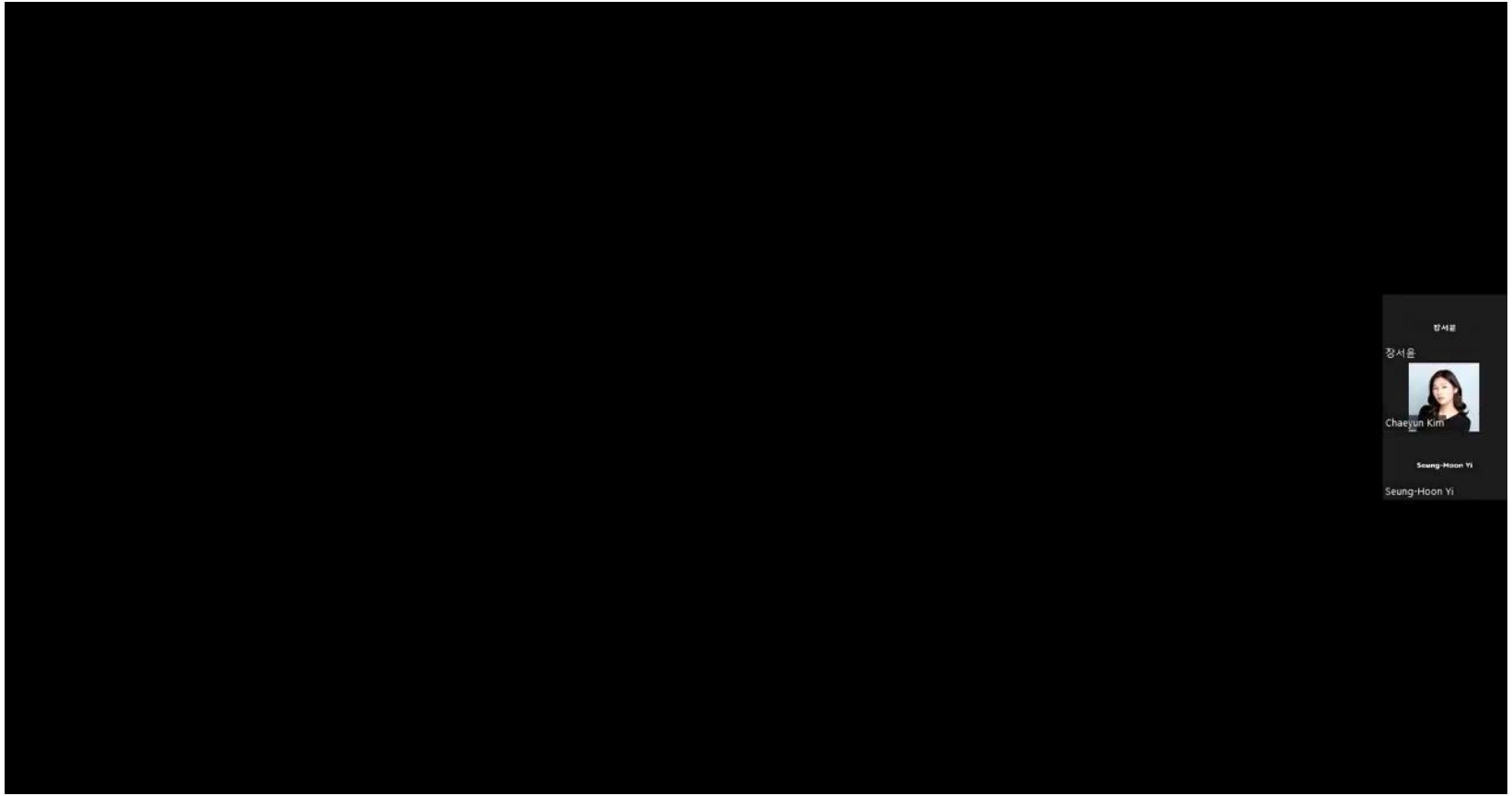








Demo Video



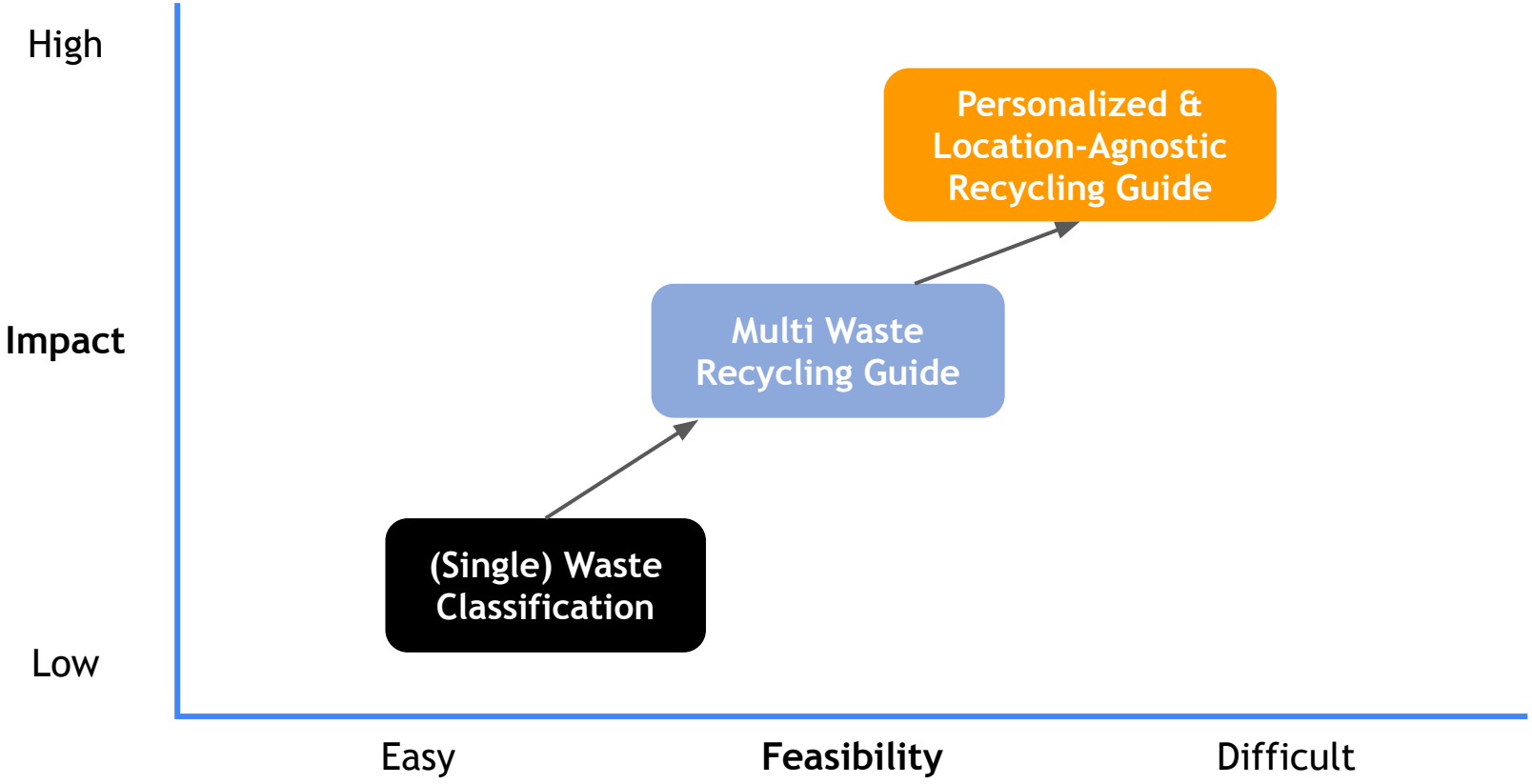
Conclusion

Cost Estimation

		Cost of operation		Total cost (100 usage per user per month)
Usage	GPT-4V API	$0.01 / 1K tokens ($0.00255 per image) $ 0.03 / 1K tokens ($0.000006 per character)	2 images / result (1 trash, 1 bins)	$0.75
	TTS API	$0.015 / 1K characters	100 characters / result	$0.0015

In total, ~$0.8 per user monthly!

Future Improvements Graph



Further Challenges

- Already Done!
- In Progress
- Need to be solved :(

Product	• Idea Development • Business Value Evaluation • Product Design
Modeling	• Multi-Modal Instruction : Decided to use T2S! • Enhance model performance with human feedback
Data	• Acquisition - Our pipeline was Real-Time • Optimizing Data Storage & Linkage with HF
Front/Back	• Give more user-friendly design (e.g. font..) • Working on security of the app (hiding keys, cache delete) • Optimizing infrastructures : shorten wait times

Rank: dev (front/back) > design > product > data

Challenges in detail: Image-based Instruction

Trials : **GPT-4** image generation, **LDM(stability AI)**

- Prompt : "***Generate image of recycling a glass bottle with a metal cap on it. The metal bottle cap should be thrown into a scrap metal bin and the glass body of the bottle should be recycled in a glass bin.***"

- GPT-4V returns a more consistent output, but **does not seem to understand the original instruction**.
- Also, there is **no guarantee that the same prompt always returns the same style**.

Challenges in detail: Video-based Instruction

Trials : **VideoLDM(stability AI), invideo(Invideo corp.)**

- Prompt : "***Make a video of recycling a glass bottle with a metal cap on it. The metal bottle cap should be thrown into a scrap metal bin and the glass body of the bottle should be recycled in a glass bin.***"

- Video outputs of both services **did not understand the instruction**.
- **No guarantee that the same prompt always returns the same style**.

Future Works

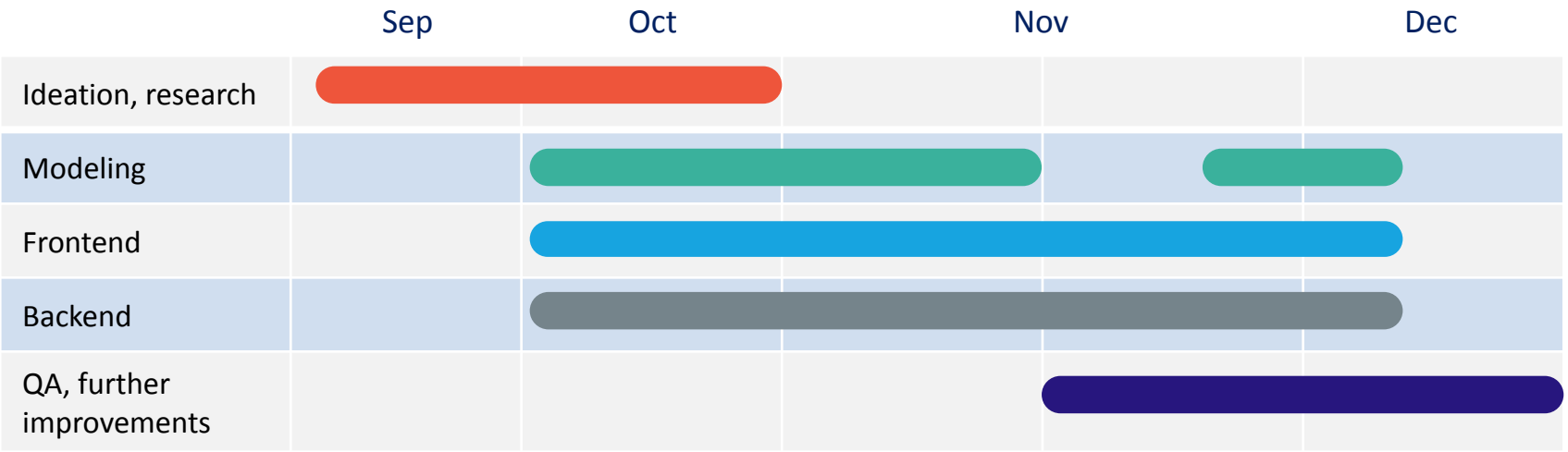
1. **Reducing the Latency**
   - It takes a considerable amount of time to turn pages; need efficiency improvements.

2. **Better setting for multi-waste guidance**
   - Ensure that recyclable items and their descriptions are more visible
   - Visualize Recyclable waste using object detection : Circle the Recyclable objects

3. **Modeling : more functions to come!**
   - **RLHF : Return human feedback about the answer result**

4. **Frontend / Backend & Further**
   - Showcase result in a more user-friendly way.
   - Implementing live setting on GCP (Google Cloud Platform)

Project Timeline

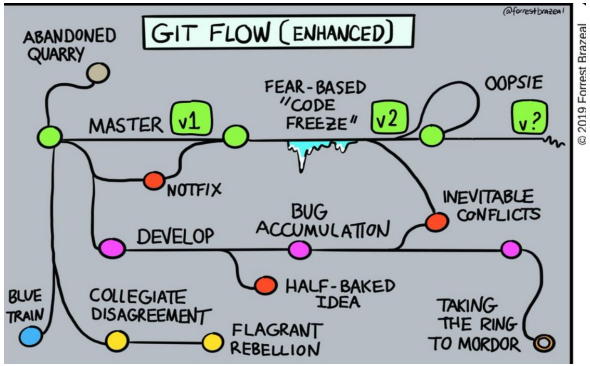


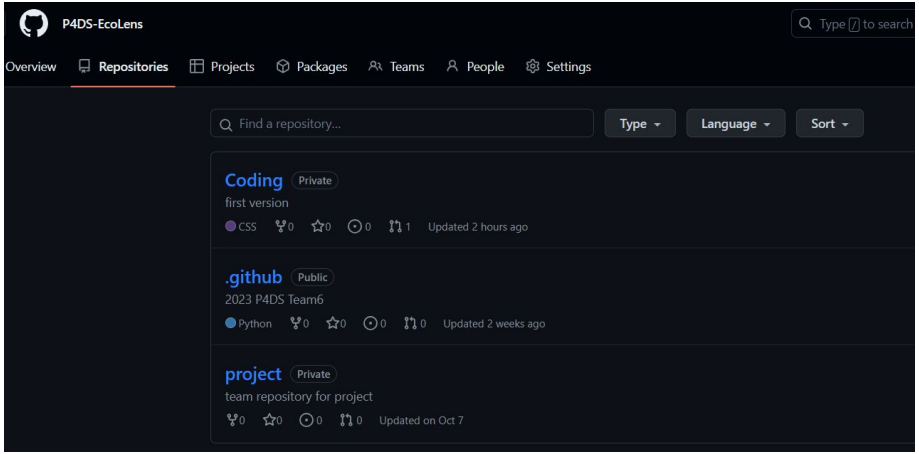
Major scope changes

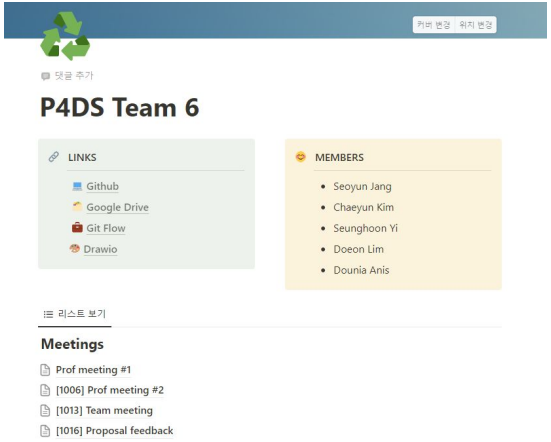
- Target users generalized to multinational individuals
- Issue of non-intuitive text instructions -> text2speech added
- Bin detection added for more personalized output
- Real-time object detection -> direct upload of image on mobile
- User feedback: pop-up chat-> add info by user's choice

Project Management

- **GitFlow**
  - promotes a consistent release process
  - organize codebase and streamline dev process
  - simple and flexible branching model







Thank You! :)