

Section 4: Annie DeForge, Ibrahim Shareef, Jacqueline Lam, Summer McGrogan

### **Meet the Team**



#### Summer McGrogan



#### Annie DeForge



#### lbrahim Shareef



#### Jacqueline Lam

#### 3

# Agenda

### **Project Overview**

Problem, Impact

### **Our Solution**

Product Demo, User Journey

### **Technical Deep Dive**

Datasets, Model Pipeline, Exploration and Results

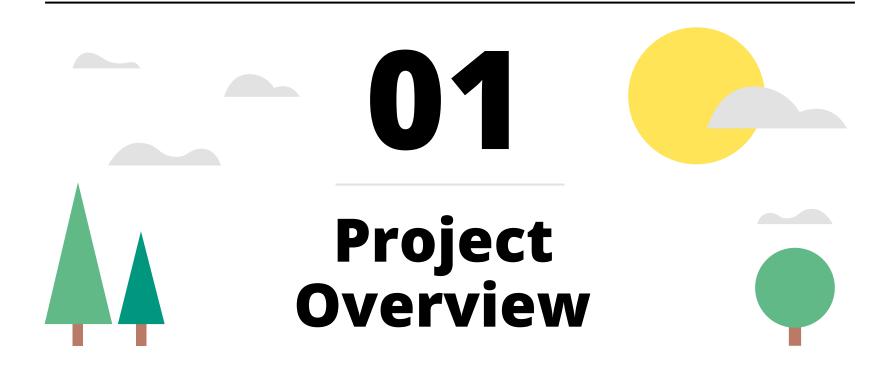
### Summary

Key Takeaways, Roadmap

02

03

 $\mathbf{01}$ 



## Problem

Recycling confusion is a leading factor in waste being recycled improperly, with many recyclable items ending up in landfills due to improper disposal



Unclear recycling instructions



Variations in recycling guidelines across localities



Perceived complexity

# **User Recycling Behavior (Survey)**

#### **Findings:**

- A large majority of people recycle, but are not looking up local information themselves
  - **97%** of users recycle
  - **75%** of users have not seen their local recycling guide
- Average confidence in deciding what to recycle: 3.29/5
- Level of interest from users surveyed in recycling assistance: 4.4/6
  - **75%** of users were interested
  - **50%** were interested or very interested

# **User Recycling Behavior (Survey)**

What are people getting wrong about recycling?

- 78% of people said they would throw a detergent jug away
  - Recyclable
- 66% of people said they would throw a hot beverage sleeve away
  - Recyclable
- 44% of people said they would recycle shredded paper
  - Cannot be recycled

**Problem:** People are making mistakes and need additional help to make correct recycling decisions

## Impact



**Environmental Protection** - Reduces landfill waste, conserves natural resources, lowers pollution, and saves energy by minimizing the need for raw material extraction and production

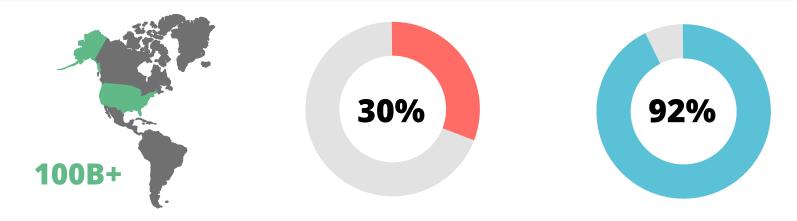


**Economic** - Saves costs, creates jobs, and supports industries by providing recycled materials, boosting both local and global economies.



**Social & Community** - supports climate action, and promotes sustainable development by conserving resources and reducing pollution for future generations.

### **Impact by the Numbers**



#### **Market Size**

The market size of whole selling of recycled material in the US is over \$100 billion.

#### **Current Situation**

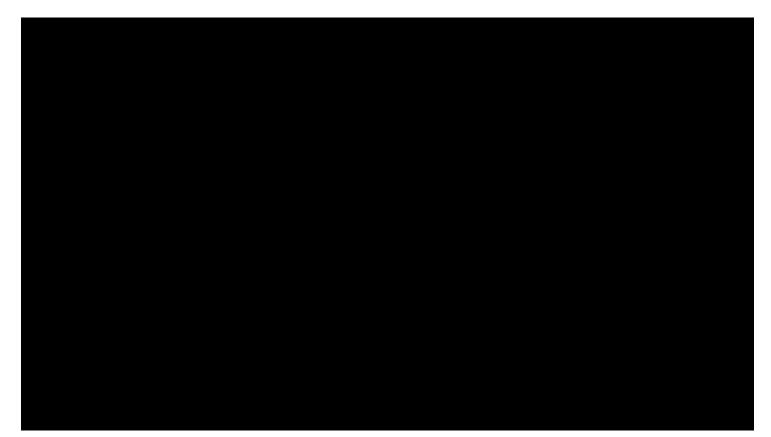
Only 30% of recyclable materials are successfully recycled.

### Participation

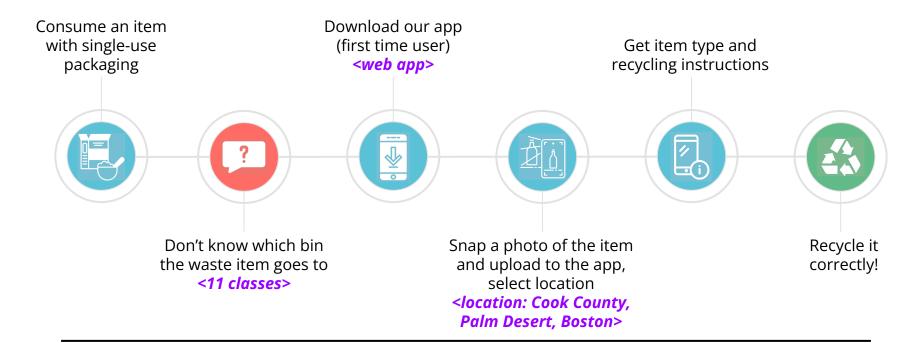
92% of Americans believe that recycling is important.



### **MVP Demo**



# **MVP User Journey**





### **Dataset: Train**

#### **Recycling Guides:**

- Cook County, Illinois
  - $\circ \quad \text{Verbose}$
  - Sentence-based
- Palm Desert, California
  - $\circ$  Concise
  - Bullet list
- Boston
  - List with examples



#### **Portland State: Recyclables**

- boxes, glass bottles, soda cans, water bottles
- 300 test per class

#### Kaggle Dataset: Mixed

- 30 categories of recyclables and non-recyclables
  - Selected: Aluminum cans, paper cups
- 250 real-world images per class

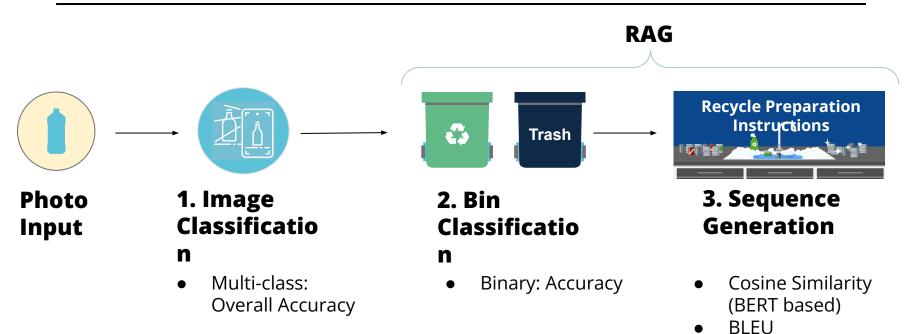


### **Dataset: Test**

- Did our own data collection with an expanded number of classes
- 4 different light conditions in 84 images
  - Bright natural lighting
  - Bright indoor lighting
  - Dim indoor lighting
  - Dim indoor lighting with flash



# **Model Pipeline**



### Model Result: Quantitative (Test)

	Stage 1 - Image Classification			Stage 2 - Bin Classification			Stage 3 - Sentence Generation			
Model	Accuracy	Precision	Recall	F1	Accuracy	Precision	Recall	F1	Cosine Similarity (BERT)	BLEU
Baseline (ResNet)	0.42	0.37	0.42	0.33	N/A					
LLaVA 1.5 (LoRA)	0.48	0.26	0.45	0.30	N/A					
LLaVA 1.5	0.80	0.86	0.82	0.75	0.98	1.00	0.93	0.61	0.52	0.22
LLaMA 3.2	0.61	0.71	0.61	0.61	1.00	1.00	1.00	1.00	0.82	0.06

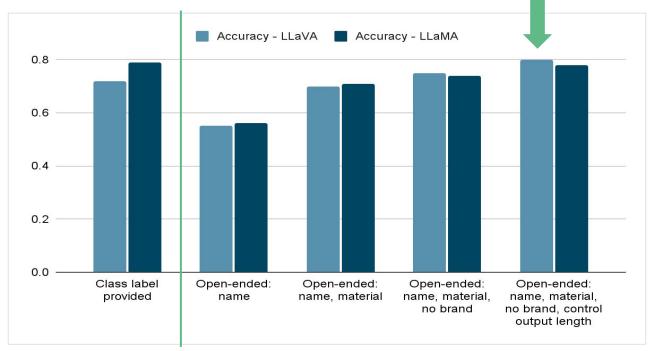
# Model Result: Qualitative (Test)

	Stage 3 - Sentence Generation		
Model	Relevancy	Fluidity	Coherence
LLaVA 1.5	3.95/5	4.56/5	4.46/5
LLaMA 3.2	3.49/5	4.81/5	4.71/5

**Findings**: Sentences included accurate details but would sometimes add unnecessary steps that creates vagueness

# **Experimentation: Prompt Eng**

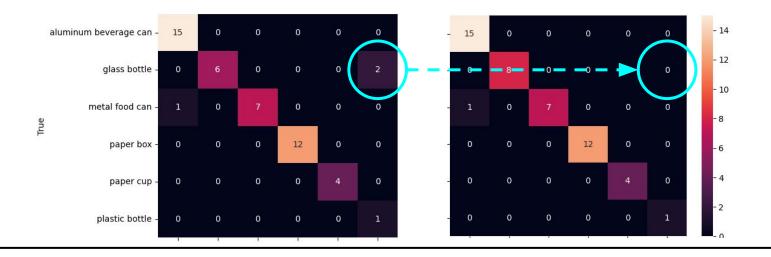
**Decision:** Use LLaVA with open-ended prompt (name, material, no brand, control output length) for Stage 1 in MVP



# **Experimentation:** Fine Tuning

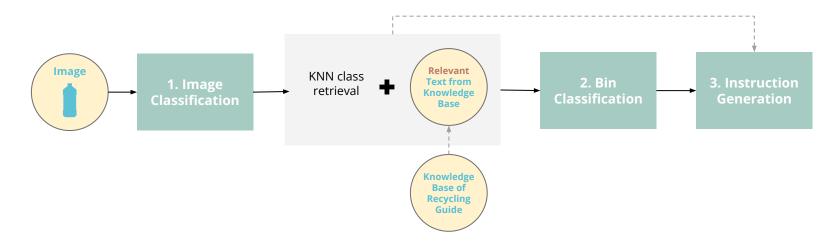
Improves accuracy & output format on 6 classes, able to distinguish glass vs plastic bottle
Does not generalise well on unseen classes / with open-ended prompt

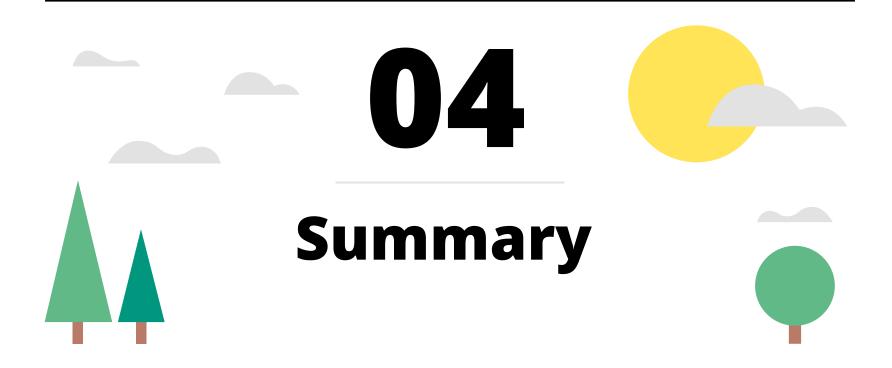
Decision: Not to use fine-tuned LLaVA model for Stage 1 in MVP



# **Experimentation: RAG**

- Knowledge base of recycling instructions for different locations and materials
- Embedding-based retrieval of the most relevant recycling instructions from knowledge base
- Input relevant text to stage 2 & 3 to improve quality of output





# Key Technical Takeaways

- 1. **End-to-End Challenge:** No single model can tackle whole pipeline
  - Multi-stage model approach
- 2. **Bin Classification Challenge:** Model is retrieving incorrect information from pre-training and running out of context window space for longer guides
  - Used embedding-based retrieval to shorten the size of the context window on large recycling guides and improve accuracy of recycling instructions
- 3. **Text Generation Challenge:** Models cannot handle negative case
  - Implemented threshold value in embedding retrieval

## **Roadmap Items**



- 1. Implement RAG with traditional vector store approach using LangChain
- 2. Extend infrastructure to accommodate use of the best performing model for different stages
- To fine-tune models for stage 2 & 3, since it worked well in adjusting the tone and format of output

Together we cansavetime,reducewasteandprotecttheenvironment!



RecycleRight

# Appendix

# **Recycling Pain Point**

#### Berkeley, CA

No:

Food-soiled paper or cardboard (e.g., pizza boxes, greasy paper bags), plastic-lined paper (e.g. milk cartons, paper cups, aseptic soy milk cartons), plastic, glass, metal, food, liquids

#### Example: Some places say you can recycle coated paper cartons and in others you can't

#### Oregon

Coffee cups and restaurant take-out boxes

#### Why it doesn't belong

To-go coffee cups and other to-go food containers are not accepted in recycling programs, because they are coated with plastic or wax that makes it difficult to separate from paper in the recycling process. These items should be put in the garbage.

#### Palm Desert, CA

#### What Items Can Be Recycled

#### Paper

- White and colored paper
- Phone books
- Magazines
- Junk Mail
- Envelopes
- File folders
- Cardboard (flattened)
- Cartons (milk, juice, & broth)
- Books (remove hard covers or binding)
  - Cardboard boxes, e.g. cereal, etc.

#### Plastic

- Styrofoam packing blocks
- Plastic medicine bottles (empty)
- Plastic bottles numbered 1-7 (remove plastic film or wrap)

#### Metal & Glass

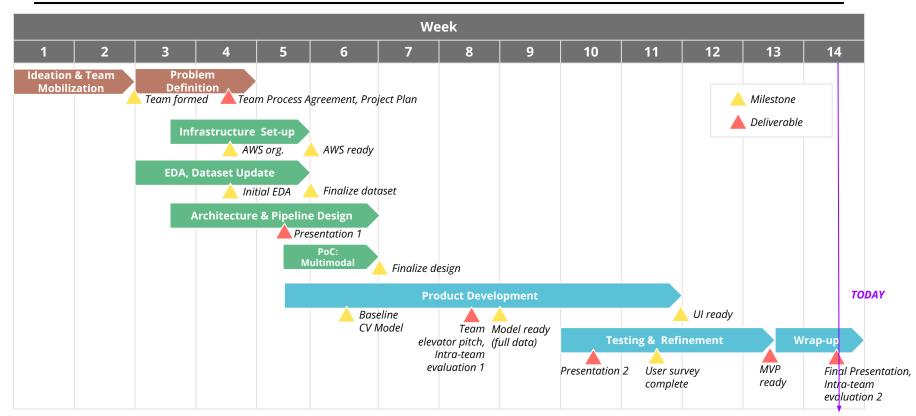
- Aluminum, steel, and tin cans
- Clean aluminum foil
- Glass bottles and jars (remove lids)

# **Market Analysis: Competitor**

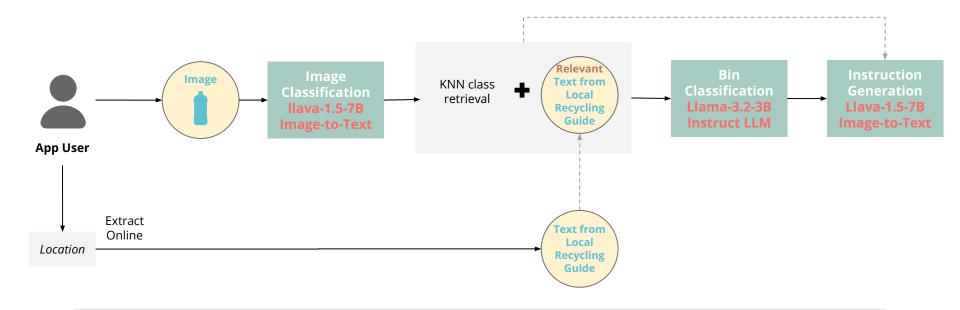
- Most are text-based apps that do not allow you to upload own photos of trash
- Biggest competitor (Scrapp) has photo-based and region-based solution
  - 6 states
  - No alternative for missing / faded barcode



# **Project Management Plan**

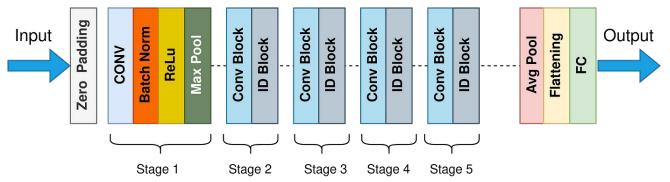


### **MVP: Solution & Data Flow**



### **Model Result: Baseline**

- ResNet50
- Utilize image classification to directly map to the bin classification and sentence output



#### **ResNet50 Model Architecture**

### Model Result: LLaVA

**1.** Stage 1: Item classification

LLaVA - Stage 1				
Accuracy	0.88			
Precision	0.80			
Recall	0.59			
F-1	0.60			

### Model Result: LLaVA

#### 2. Stage 2: Bin classification

LLaVA - Stage 2				
Accuracy	0.77			

#### **3.** Sentence Generation

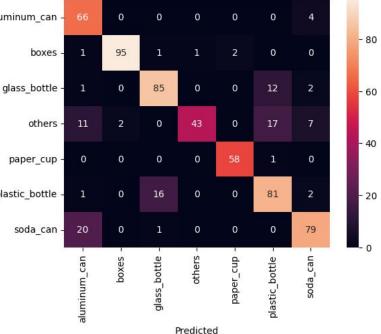
<u>Sample Output:</u> The image shows a box of cereal, which is made of cardboard. Cardboard is a type of paperboard and can be recycled. To recycle the cardboard box, it should be placed in the recycling bin along with other paperboard materials. After collection, the cardboard is processed and transformed into new products, such as cardboard boxes for cereal or other packaging materials.

LLaVA - Stage 3			
Cosine Similarity	0.52		
BLEU	0.22		

### **Model Result: LLaMA 3.2**

#### 1. Item Classification

		aluminum_can -		0	
LL	boxes -	1	95		
Accuracy	0.83	glass_bottle -	1	0	
Accuracy	0.00	others -	11	2	
Precision	0.85	paper_cup -	0	0	
Recall	0.84	plastic_bottle -	1	0	
	0.00	soda_can -	20	0	
F-1	0.83		m_can -	boxes -	



### Model Result: LLaMa 3.2

#### 2. Bin Classification

LLaMA - Stage 2				
Accuracy	0.64			

#### 3. Sequence Generation

• <u>Sample Output:</u> For glass\_bottle, rinse and dry thoroughly before recycling. Do not bag or wrap items; keep them loose in the bin for easy sorting.

LLaMA - Stage 3				
Cosine Similarity	0.83			
BLEU	0.24			

# **Ethics & Data Privacy**

### **User Privacy**

- Stored info: Location
- Sensitive/private information in background of photos

#### Consequences of Misclassification

Negative environmental impact of Type I and Type II errors

