

DOCUMENTATION

Project Title: Dataset Development & Annotation Tool

Organization: MealLens

Role: Team Leader – Group 4

1. Project Overview:

The Dataset Development & Annotation Tool project at MealLens aims to create a scalable and efficient system for collecting, labelling, and organizing food ingredient images. This tool will enable the development of a high-quality dataset essential for training AI models that identify food items and extract nutritional insights.

2. Team Composition:

- Total Participants: 50
 - CEO of MealLens
 - Mr. Bardiya Rasekh
 - Graduate Assistant (GA)
 - 48 student contributors, divided into 8 teams (6 members each)
- Team: Group 4
 - Members: 6 (including myself)

3. Key Milestone: Kickoff Meeting

Date: May 15th, 2025

Attendees: All 50 members, including the CEO, GA, and all student teams

When the CEO invited the participants to introduce themselves, there was an initial hesitation from many attendees. Recognizing the need for someone to take initiative, **I stepped forward and was the first to introduce myself** confidently. In my introduction, I included:

- My **technical skillset** (e.g., Python, React.js, SQL, Power BI, etc.)
- A **relevant project** I had previously worked on (e.g., NYC Bike Sharing Dashboard or Career Guidance System)
- My **commitment to the work on the MealLens annotation tool** and contribute meaningfully to the AI food recognition pipeline.

This initiative helped break the ice and encouraged others to participate actively in the discussion.

3. In-person Meeting at CEI Building

Date: May 21st, 2025

Attendees: Team-Leads only

I attended the meeting as the **representative and leader of Group 4**. Demonstrating professionalism and an understanding of Canadian workplace culture, I was the **first to arrive exactly at 11:00 AM**, the scheduled start time. A total of **six team leaders** were present, including myself, Sumanth, Yuvashree, Ankit and Foysal. There were three of them who were absent. The session began with a brief round of introductions, after which we were introduced to **Mr. Niaki** and **Mr. Bardiya**.

During the meeting, **Mr. Bardiya** led the discussion and provided us with an overview of the **Data Preparation Stage**. He explained the importance of this phase in building a AI model and emphasized the need for accuracy and consistency in data annotation. Specific tasks were assigned to each group, which included cleaning the dataset, retaining only authentic food images, and removing any irrelevant visuals such as food paintings or artificial models. This meeting helped clarify our responsibilities and set a strong foundation for the upcoming project milestones.

4. Internal Team Meeting (Post-Company Discussion)

Date: May 22nd, 2025

Mode: In-person at CAW

Attendees: All Group 6 Members

After the in-person meeting at the CEI Building, our team gathered to discuss the project scope and task division. We reviewed Mr. Bardiya's instructions on cleaning the dataset and ensuring only valid food images are retained.

We divided responsibilities into three sub-teams:

- **Frontend Team** – to design a basic UI for annotation and interaction
- **Backend Team** – to handle project structure, repository setup, and APIs
- **ML Team** – to prepare for model training once the dataset is ready

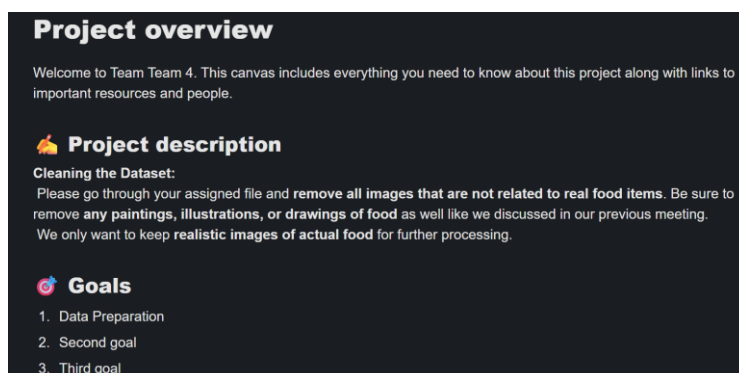
We also noted that, due to data size restrictions, image data cannot be pushed to GitHub. A shared tracker was created to manage task progress. We had come to a conclusion **to finish cleaning the dataset long before the given deadline ie. By 29th May**.

5. Post-Tool Access Setup (Slack & ClickUp)

- Following the distribution of Slack and ClickUp links to all team members, I promptly introduced myself in the general communication channel to maintain

transparency and establish an approachable presence as the **leader of Group 4**. Our team was also provided with a **locked Slack channel** dedicated to internal collaboration. Taking initiative, I organized the workspace by creating multiple tabs including **Messages**, **Project Overview**, **Project Tracker**, and **Images Dataset** to streamline communication and ensure a clear flow of information.

- In the **Project Overview** tab:
 - Provided a **brief explanation of the project**

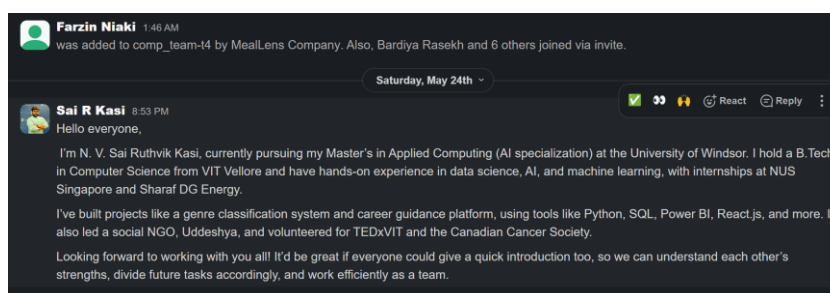


- Assigned tasks to each member in a **tabular format** to avoid miscommunication

A screenshot of a Slack channel titled "Team". It displays a table with two columns: "Task" and "Name".

Task	Name
6.0 - 6.4	@Akash
6.5 - 6.9	@Nishant Patel
Folder 6 a-f	@Aayush Thakur
7.0 - 7.4	@Tushar Sharma
7.5 - 7.9	@Sai Akshay India
Folder 7 a-f	@Sai R Kasi

- Asked all members to **introduce themselves** and share their technical skills for better collaboration



- In the **Images Dataset** tab:
 - Uploaded the **master folder link (200GB)** provided to team leads

- Extracted our team's **assigned 25GB dataset**
- Shared the 25GB dataset with team members so they could begin working

5. Impromptu Meeting – May 27, 2025

- **Initiated by:** Dr. Niaki
- **Time of announcement:** Around 11:00 AM

Situation & Issues Faced

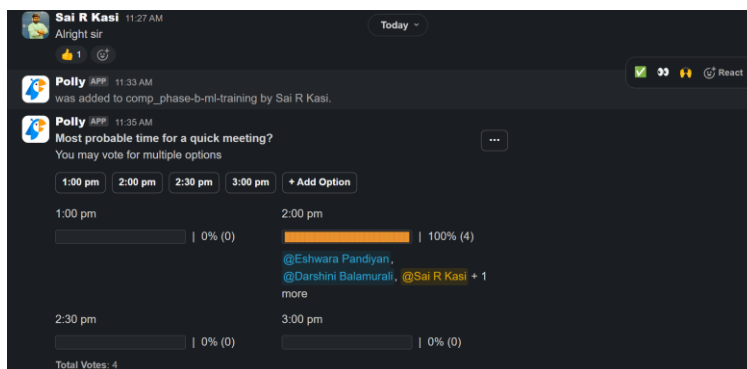
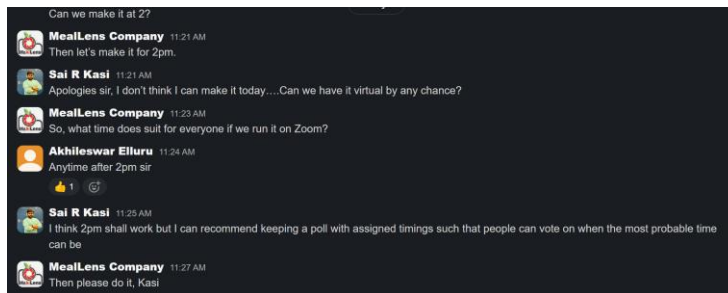
- Disagreements emerged while trying to **finalize a meeting time**
- Many participants:
 - **Proposed times based on personal convenience**
 - **Jumped to conclusions** without mutual discussion
 - **Did not respond at all**, leading to confusion

Leadership Action Taken

- I considered all participants' **points of view**
- Suggested holding the meeting **online** to accommodate:
 - People living far from the **CEI building**
 - Team members with **prior commitments**
- Proposed to **Dr. Niaki** that a **poll be created** to:
 - Collect everyone's availability
 - Avoid biased or one-sided scheduling
 - Ensure a fair, organized decision-making process

Outcome

- Helped restore order and structure in the discussion
- Encouraged more **inclusive and thoughtful participation**
- Demonstrated **initiative and conflict resolution** under pressure



6. Dataset Cleaning Completed (Pending Final Review)

As of May 28th, 2025, our team has completed the initial round of dataset cleaning. Over 80,000 images were manually reviewed, and all non-food items including paintings, artificial models, and unrelated visuals, were removed.

We are currently conducting a final review to ensure the dataset meets the required standards. Once verified, we will present the cleaned dataset to **Dr. Niaki** for approval before proceeding to the next development stages.

Note: Due to data restrictions, the cleaned images will not be uploaded to GitHub.