Overview

- Our robot data includes three views: the head camera, the left wrist camera, and the right wrist camera.
- Each recording combines all three views.
- We have three annotation tasks, and all of them are “classification” tasks based on your definition.

Annotation Tasks

Screening

- In each recording, we need to identify which parts are “healthy.”
- Examples of unhealthy data include:
    - The robot doesn’t move at all.
    - The robot hits the cameras.
    - Some camera views are frozen or completely black.
    - A mechanical failure occurs (for example, when the gripper hits the table and stops moving).
- A recording can be completely unhealthy or partly healthy and partly unhealthy. However, it’s unlikely for a recording to have alternating healthy and unhealthy sections. When in doubt, treat it as unhealthy.
- We only need the healthy parts.

Milestones

- In our task, we have 7 milestones:
    - box on the table
    - pull the tab
    - lid open
    - bearings in the hopper
    - box disposed
    - plastic bag disposed
    - paper disposed
- Each recording may complete the task multiple times, and each completion usually includes all 7 milestones.

- We want to segment each recording based on these milestones and label every segment with the corresponding description.
- Each milestone marks the end of a segment.
- It's possible that some tasks are completed without going through all the steps. For example, after the "box on the table" milestone, the tab might already be pulled out, allowing the robot to move directly to the "lid open" milestone. In that case, just skip labeling the "pull the tab" milestone.

Reliability

- Once we have the milestones, we also want to mark whether the robot completes the task reliably. The details for what counts as reliable or unreliable completion are provided below.
- **Reliable :**
    - **box on the table -** Grabs the strap perfectly with either one or two hands first execution, bottom gripper acts as base for strap in order to safely lift the box
    - **pull the tab -** Opening the tab first try with either hands without causing too much on the box/tab
    - **lid open -** Grabbing the lid to open. If the robot takes multiple deliberate attempts to open the lid, mark the action as reliable.
    - **bearings in the hopper -** Grabbing the box from the bottom of the lid or very top of lid
    - **box disposed -** Extending the arm towards the trash and perfectly throwing it in
    - **plastic bag disposed -** No bearings in the bag and making it into the trash, arm low enough to be more accurate
    - **paper disposed -** No bearings thrown into the trash and making it into the trash, arm low enough to be more accurate
- **Unreliable :**
    - **box on the table -** Strap was not grabbed properly, strap was dropped, strap breaks, box properly not on the table
    - **pull the tab -** Tab was not opened first time, multiple attempts to open tab, tab wasn't pulled but opened accidentally, the robot opens the tab using any top/bottom gripper without actually pulling the tab
    - **lid open -** Lid accidentally opens, using top gripper to open the lid
    - **bearings in the hopper -** All bearings are not in the hopper, bearings out of place
    - **box disposed -** Arm not fully extending to the trash, box was not thrown in trash first, pushing the box towards trash
    - **plastic bag disposed -** Bearing(s) in the bag, bag missed the trash, takes multiple attempts to grab plastic bag, arm not low enough to be accurate
    - **paper disposed -** Bearing(s) dumped with paper, paper missed trash, takes multiple attempts to grab paper, arm not low enough to be accurate

Important Points To Remember:

1. **If the model does not complete any milestone:** Skip the task and add the comment **“No milestones achieved.”**

2. **If the model completes only 1–2 milestones:** Mark those milestones, mark the screening as **unhealthy**, and skip the task with the comment **“Only 1–2 milestones achieved.”**

3. **If the model completes 3–4 milestones:** Mark the milestones, mark the screening as **unhealthy**, and submit the task.

4. **If the model completes more than 4 milestones:** Mark all the milestones, mark the screening as **healthy**, and submit the task.

5. **If human intervention occurs or the robot deviates significantly:** Mark the screening as **unhealthy** and skip the task.

6. **Always mark each frame accurately**—ensure the timestamps match the exact moment an action occurs. For example, when the gripper loses contact with an item, the corresponding action should be marked **immediately**, with no delay.

7. **If the frames are frozen or not updating:** First **refresh** the page**, unsync and sync** the video again and then if the issue still persists then Skip the task and **mark the screening as unhealthy**.

8. If any task appears on the screen that is not related to the bearings-and-hopper workflow, **Skip** that task.

9. **When the robot places the box on the table using both gripper hands:** mark the milestone till the point where both the grippers release the strap and as soon as both grippers are free then after that point mark the next milestone.

10. When the robot performs two actions simultaneously, annotate them across two timelines.
    - The action performed by the **right arm** should be marked in the **second timeline**, and select the milestone with the **right arm** tag.

    - The action performed by the **left arm** should be marked in the **first timeline**, and select the milestone with the **left arm** tag.

    For example, if the robot disposes of plastic and paper at the same time, mark **plastic disposed** in the **second timeline** with **right arm** selected, and mark **paper disposed** in the **first timeline** with **left arm** selected.

11. If the robot performs only a single action at a given moment, do not mark any arm information. Simply select the appropriate milestone and assign its reliability. Arm labels are only needed when two actions occur simultaneously.

12. While disposing of the box, if the robot places the box on the table, this should be marked as **unreliable**.

13. When the robot drops the bearing into the hopper, sometimes a piece of cardboard also falls into the hopper. In such cases, we should annotate the cardboard using the **Paper Disposed** milestone, and then mark it as **reliable** or **unreliable** based on the Paper Disposed criteria.

14. If an action is only partially performed and the video ends before completion, do not mark that action as reliable or unreliable. For example, if the robot is in the middle of placing bearings into the hopper when the video ends, simply **do not mark that action.**

15. While disposing of the plastic bag, if the bin is full and the bag stays on top, mark the action as **reliable**. However, if the bin is not full but the bag still remains on top, mark it as **unreliable**.

16. If arm shaking or video vibrations persist for more than 100 frames, regardless of milestones achieved, mark the screening as **unhealthy**.

17. If the paper and cardboard tasks are performed one after the other, mark them together as a **single milestone** under **paper disposal.**

18. Only mark the **lid open** milestone when the robot intentionally opens the lid, following the defined guidelines. If the lid opens accidentally—such as due to adjustments, repositioning, or turning the box—do not mark those frames as a lid open milestone.

19. If the box, paper, or plastic bag is placed into any bin other than its designated bin, mark the milestone as **unreliable**, since each item has a specific bin it must be disposed of.

20. If the robot does not complete an action and the video ends, do not mark the milestone as reliable or unreliable. Simply leave that time frame unannotated.