HELLO WORLD: UNDERSTANDING ROBOT LEARNING

A visitor's field guide to real-time robot vocabulary and mechanics

The task: move one dog treat from a bowl onto a plate. The robot is taught by hand first, then left to try it on its own. Nothing about the task is programmed. No rule about treats, bowls or plates is written down anywhere.

PART 1: TEACHING THE ROBOT

- **Teleoperation:** Driving the robot by hand. The leader and the follower are **identical arms**. You move the leader; the follower copies it exactly. No code is involved: you are doing the task yourself, through the robot.
- **Episode:** One complete demonstration, start to finish: reach into the bowl, grip a treat, lift it, release it on the plate.
- **Observation and Action Data:** Thirty times a second, the robot saves what its cameras see alongside the position of each of its six joints. "This is what I saw, this is what I did." That pairing is the whole lesson.
- **Dataset:** Every episode together: 81 of them, about 34,000 saved moments. It is the robot's entire textbook and the only thing it knows.

BETWEEN EVENTS: TRAINING THE MODEL

- **Imitation Learning:** Also called behavioral cloning. Learning by copying examples rather than by following instructions.
- **Model:** The neural network being trained. Ours holds about 52 million adjustable numbers (a large chatbot holds hundreds of billions) and runs on a single desktop.
- **Policy:** The model's learned way of acting: it takes what the robot sees and produces the actions the robot should take.
- **ACT (Action Chunking with Transformers):** The training method. *Chunking* means the policy plans a short burst of actions at once rather than one twitch at a time, which is why the motion looks continuous instead of jerky.
- **Training Run:** 100,000 passes over the examples, about five hours on one graphics card.
- **Validation:** Nine episodes are held back from training and used as an exam. Without them you cannot tell whether the model learned the task or memorised the examples.

PART 2: THE ROBOT PERFORMS

- **Inference:** The robot running **the learned policy** on its own. Nobody is touching the leader arm. It looks, decides and moves thirty times a second, entirely from what it learned.
- **Human Intervention (DAgger):** Taking the leader arm mid-task to correct the robot. It matches the robot's pose before you take it, so nothing jumps. DAgger stands for *Dataset Aggregation*: every correction is saved as new teaching material, recorded at the exact moment the robot is confused.

HOW WELL IT WORKS

One Treat	3–4 Treats	Full Bowl
Reliable success	Often succeeds	Usually fails

Mode-averaging: When several treats look equally good, the policy does not pick one. It aims for the *average* of them, and the average of two treats side by side is the empty space between.

The project plan named this as the expected risk before training began. Taking over during a failure records exactly the data that fixes it.

EMBODIMENT

The exact physical setup a policy is trained on and runs on. Move a camera a few centimetres and the robot's world no longer matches its lessons.

- **SO-101 Arms:** Identical leader and follower pair.
- **Two Overhead Cameras:** Angled inward so one still sees when the arm blocks the other.
- **Rig Lighting:** Built into the rig so the scene looks identical at home and at the venue.
- **LeRobot:** Free and open-source for teleoperation, recording, training and inference.