Stanford | ONLINE

# Robotics and Autonomous Systems Seminar

## A Careful Examination of Multitask Transfer in TRI's Large Behavior Models for Dexterous Manipulation

April 25, 2025

**Russ Tedrake**

MIT

Stanford | ONLINE

A careful examination of

# **Multitask Transfer** in TRI's **Large Behavior Models**

for dexterous manipulation

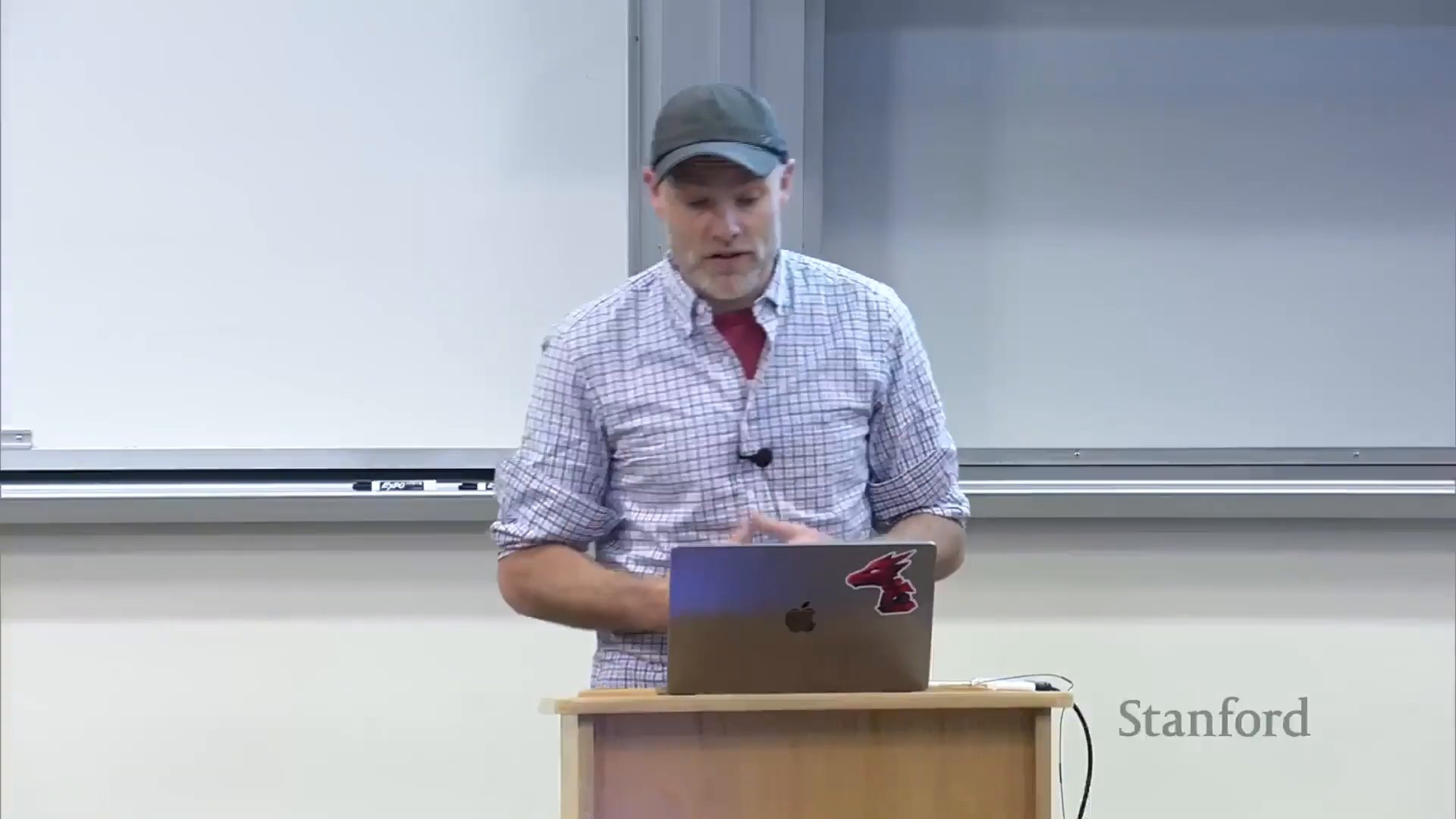
**Russ Tedrake**, on behalf of

TRI's Large Behavior Model (LBM) Division

TOYOTA RESEARCH INSTITUTE

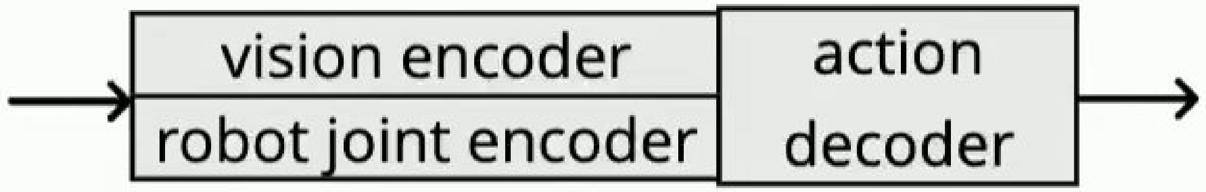
Stanford


Stanford


Stanford

Diffusion Policy (DP) is "single-task"

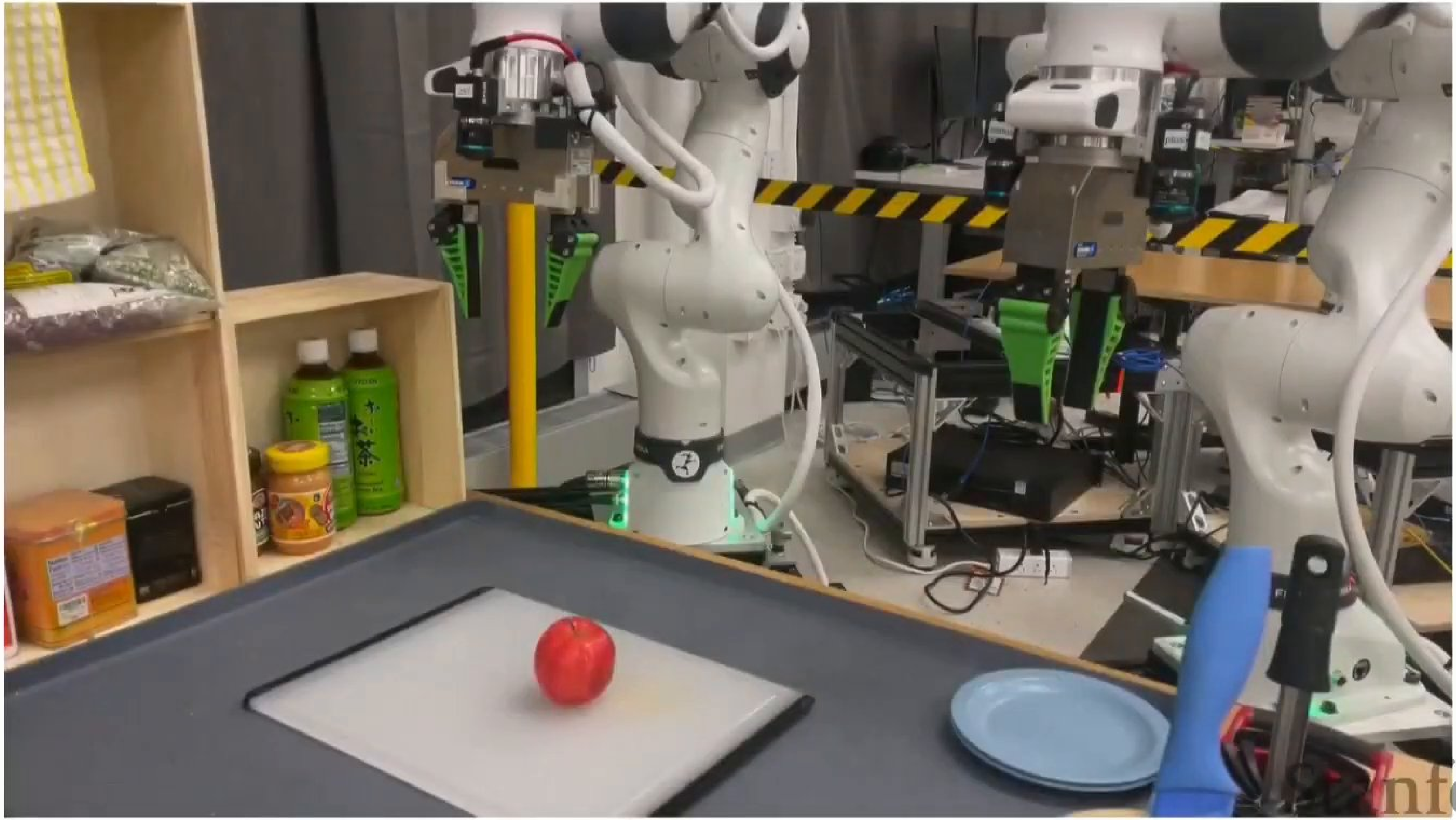
- Train on ~200 demonstrations
- ⇒ visuomotor policy

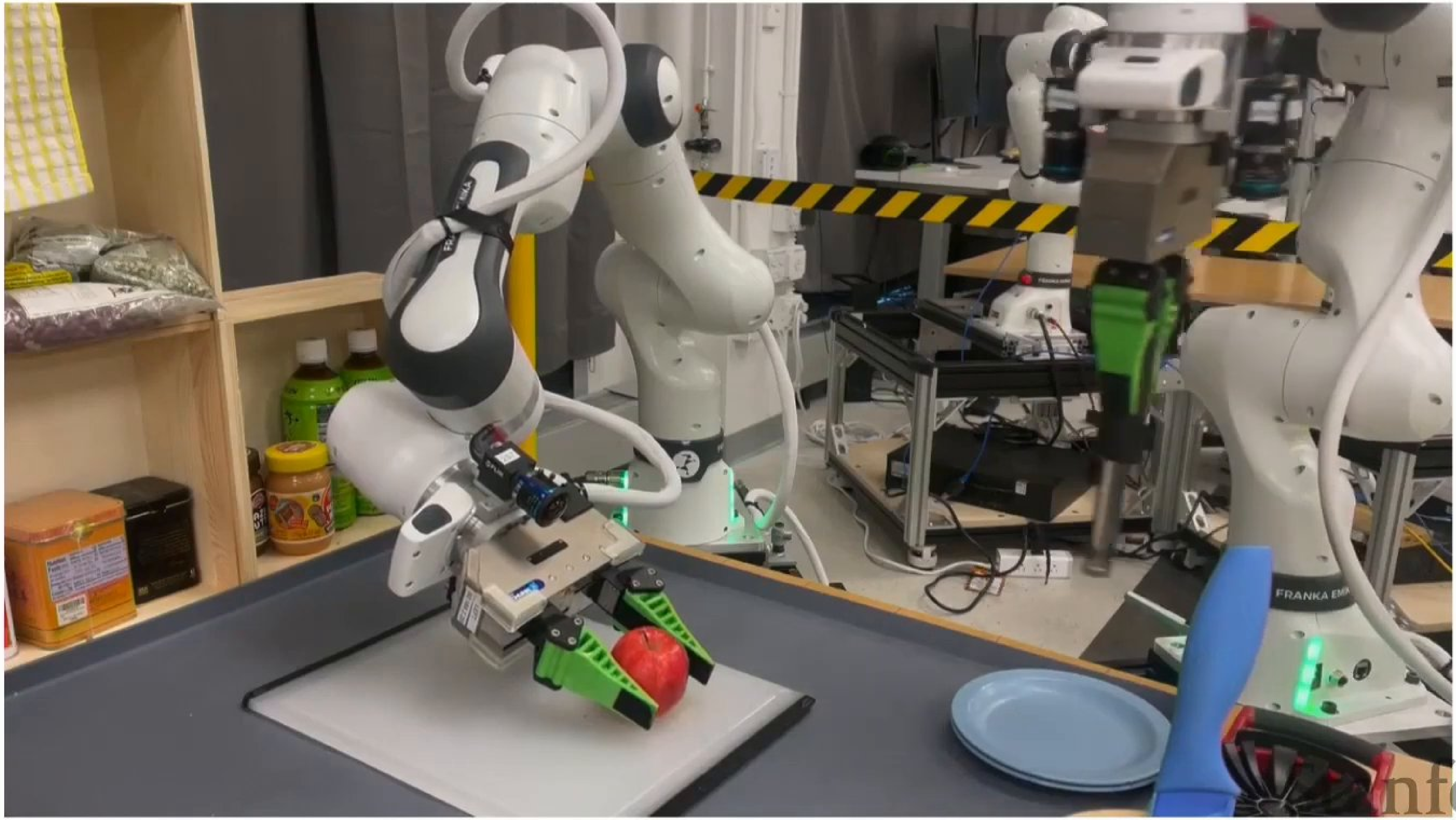


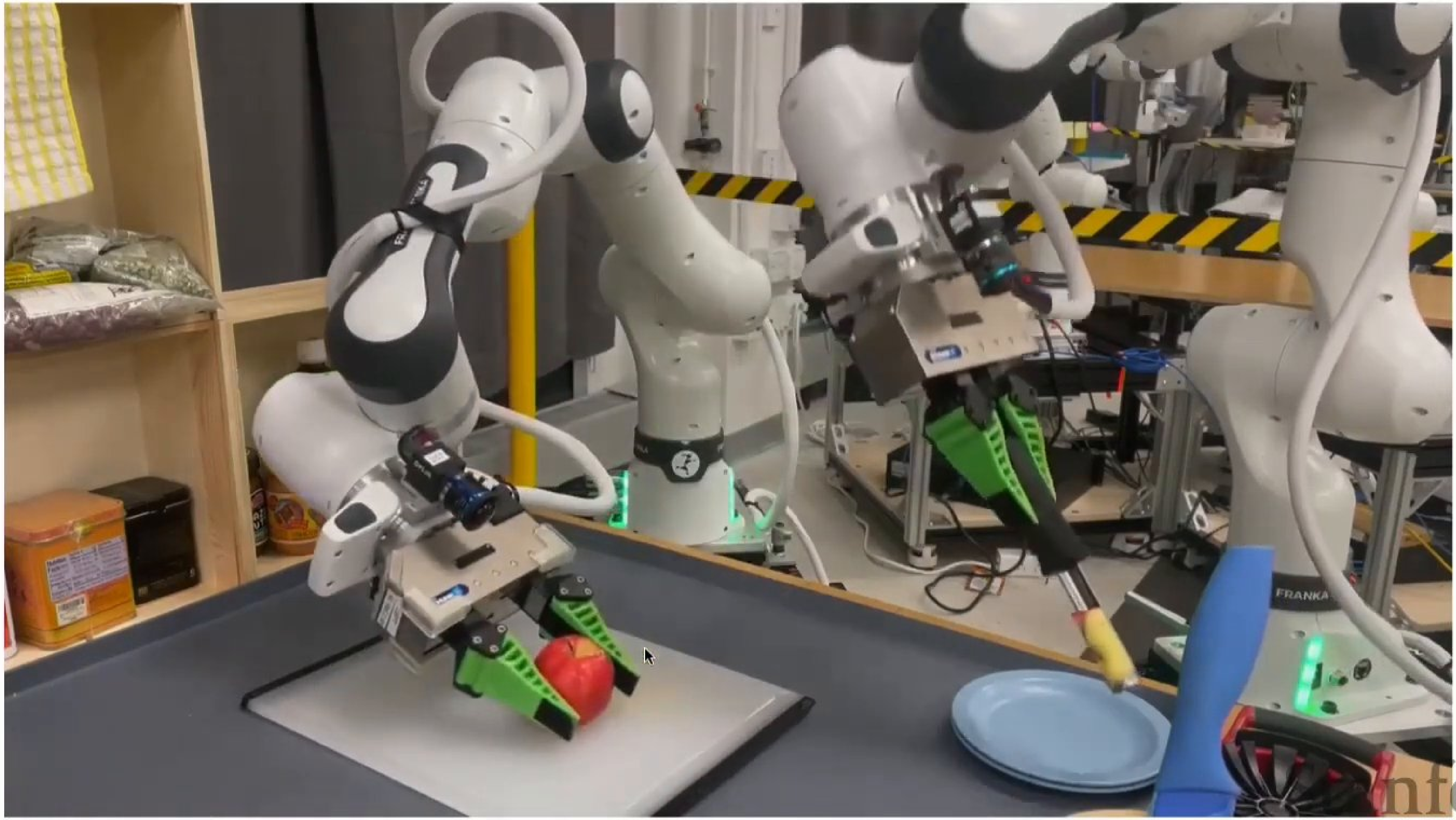
# The opportunity

Stanford

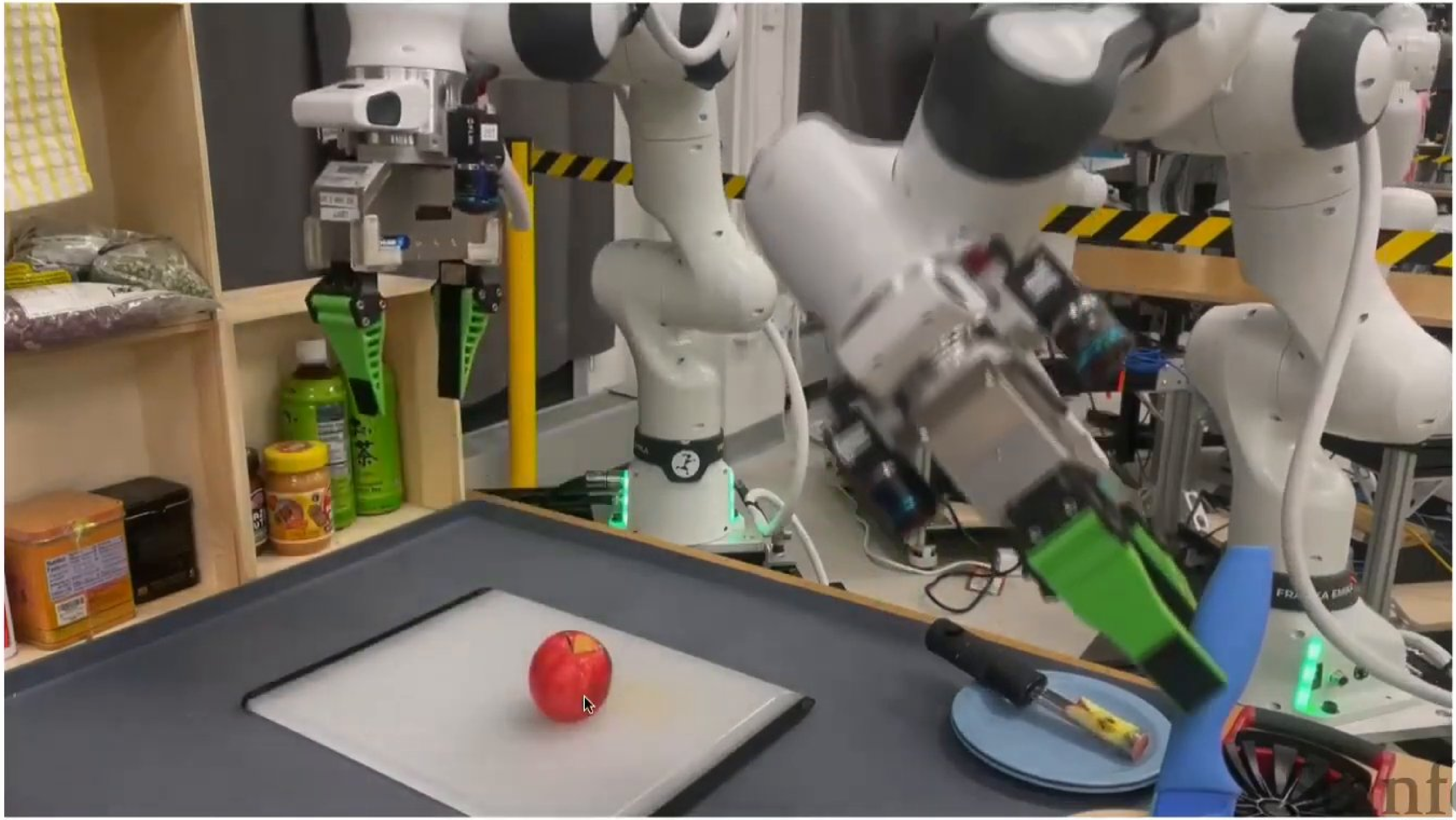
Stanford

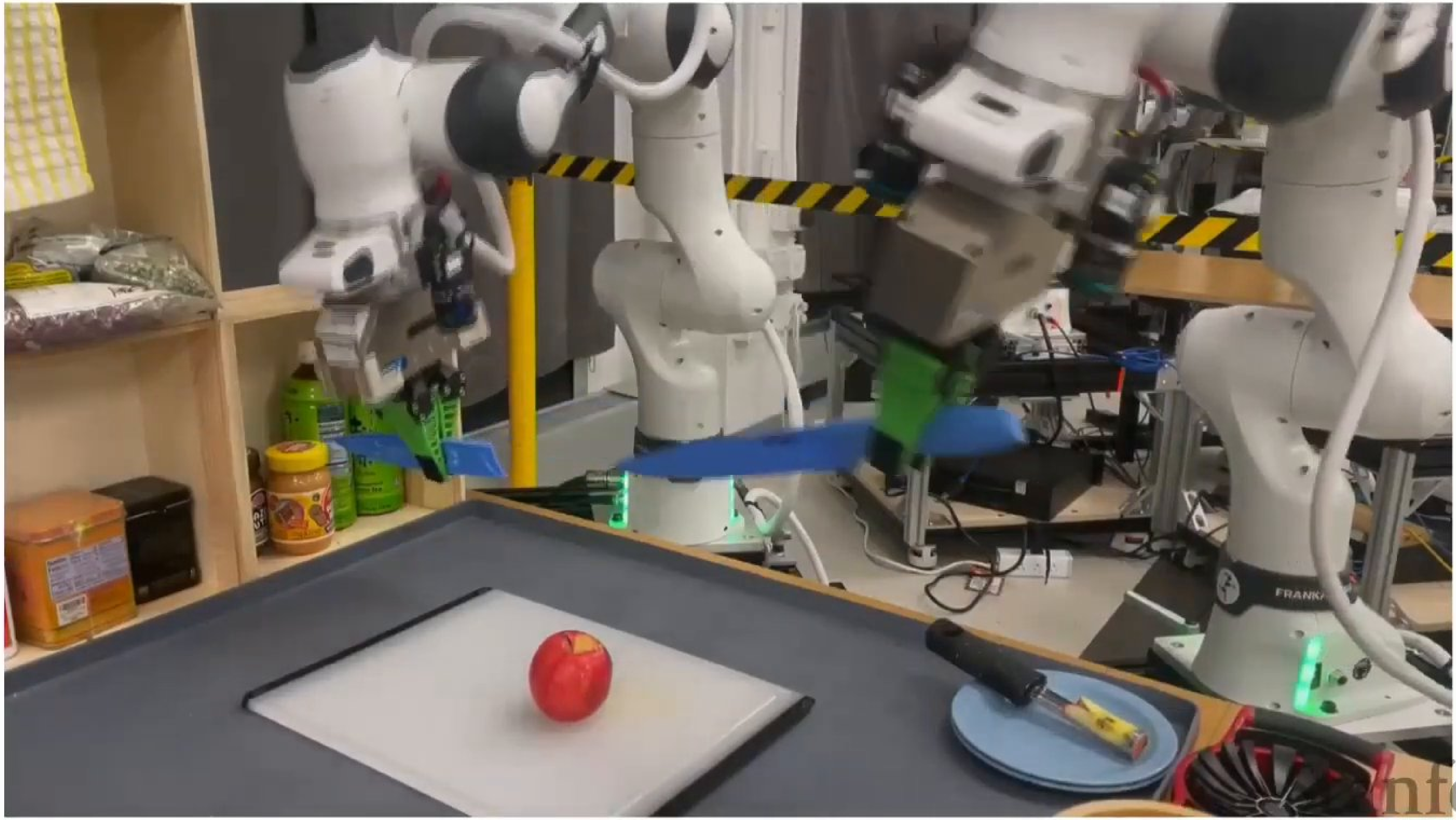


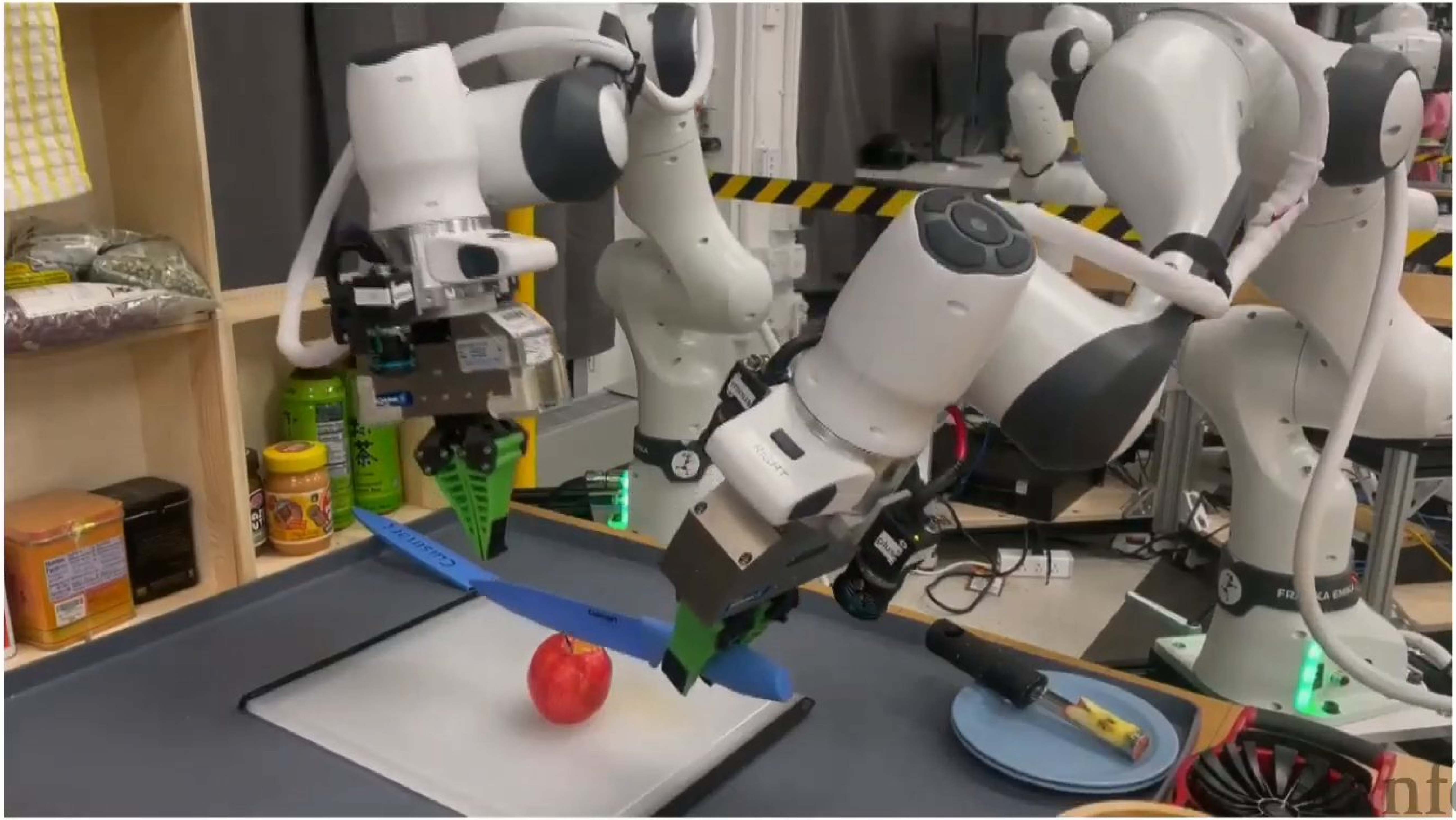




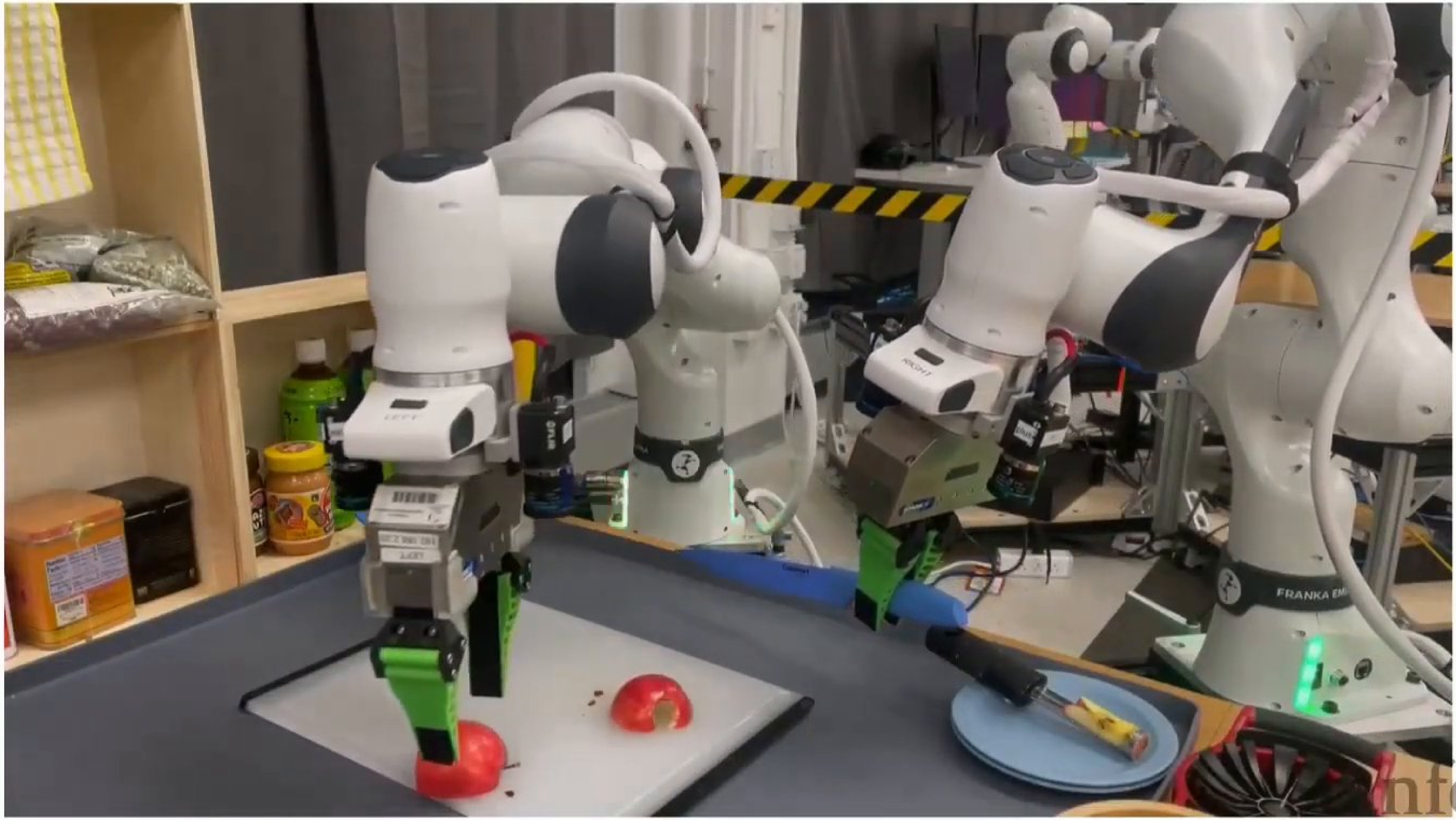

nford




FRANKA



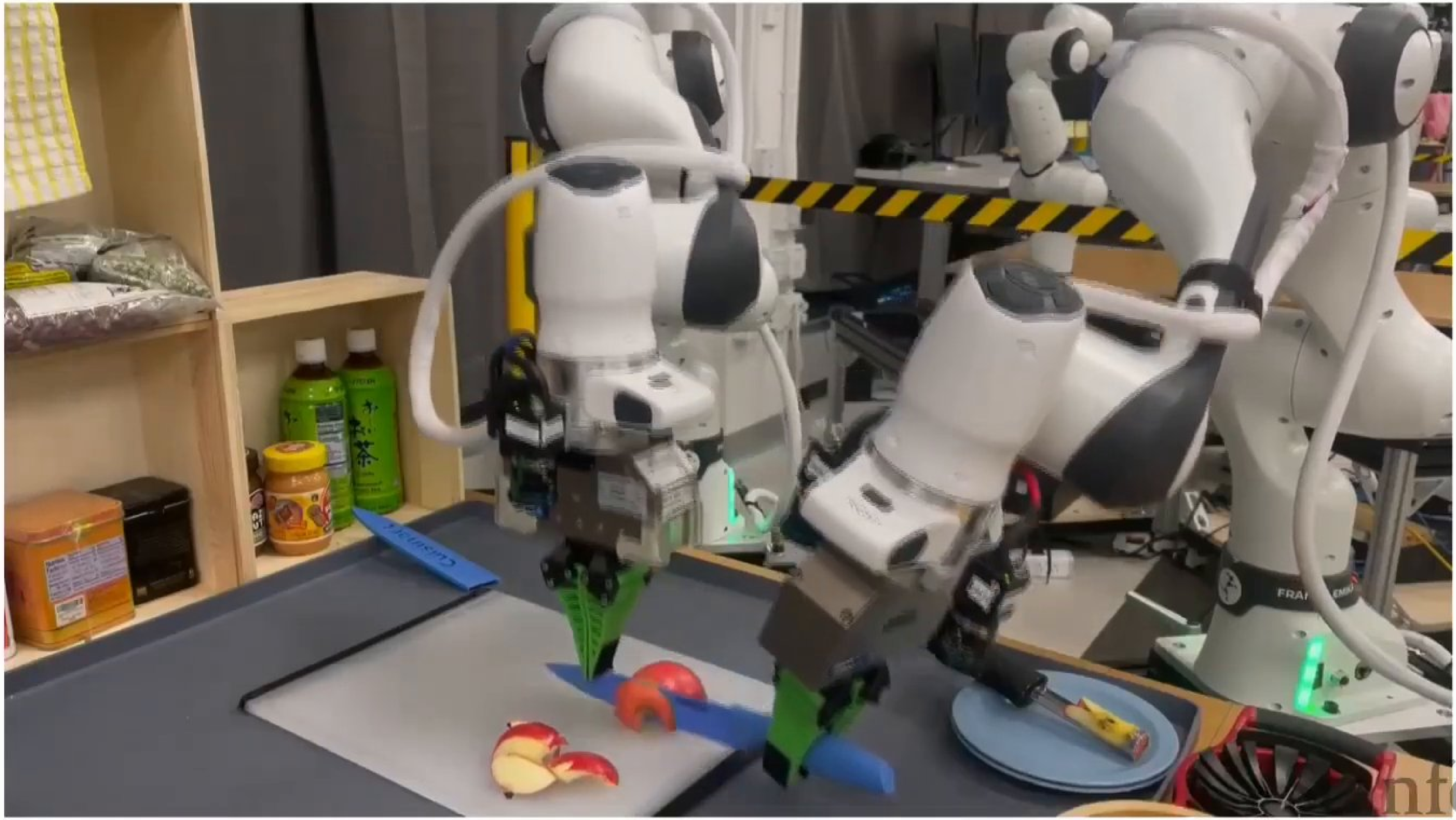

nford



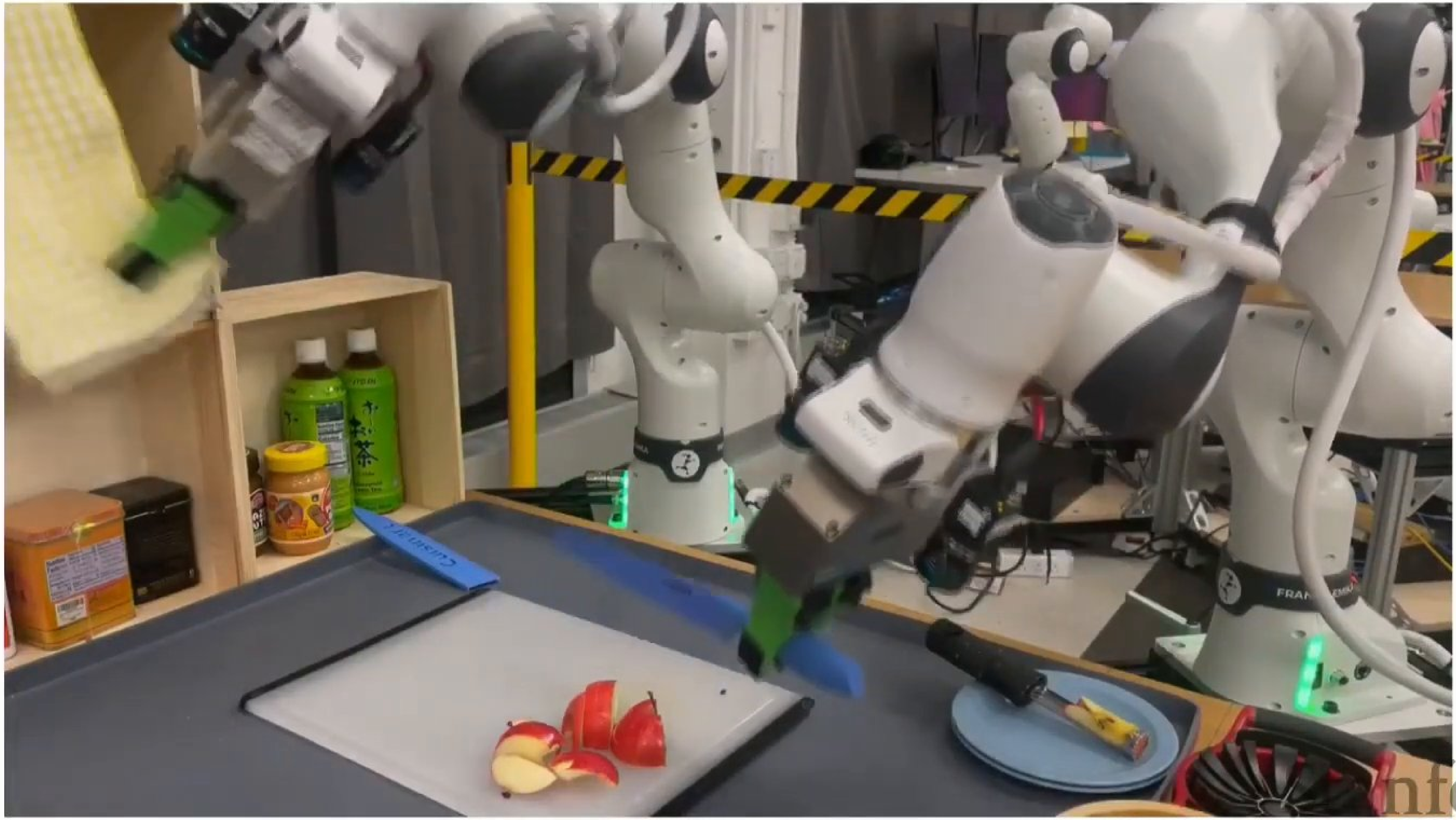

nford


RIGHT

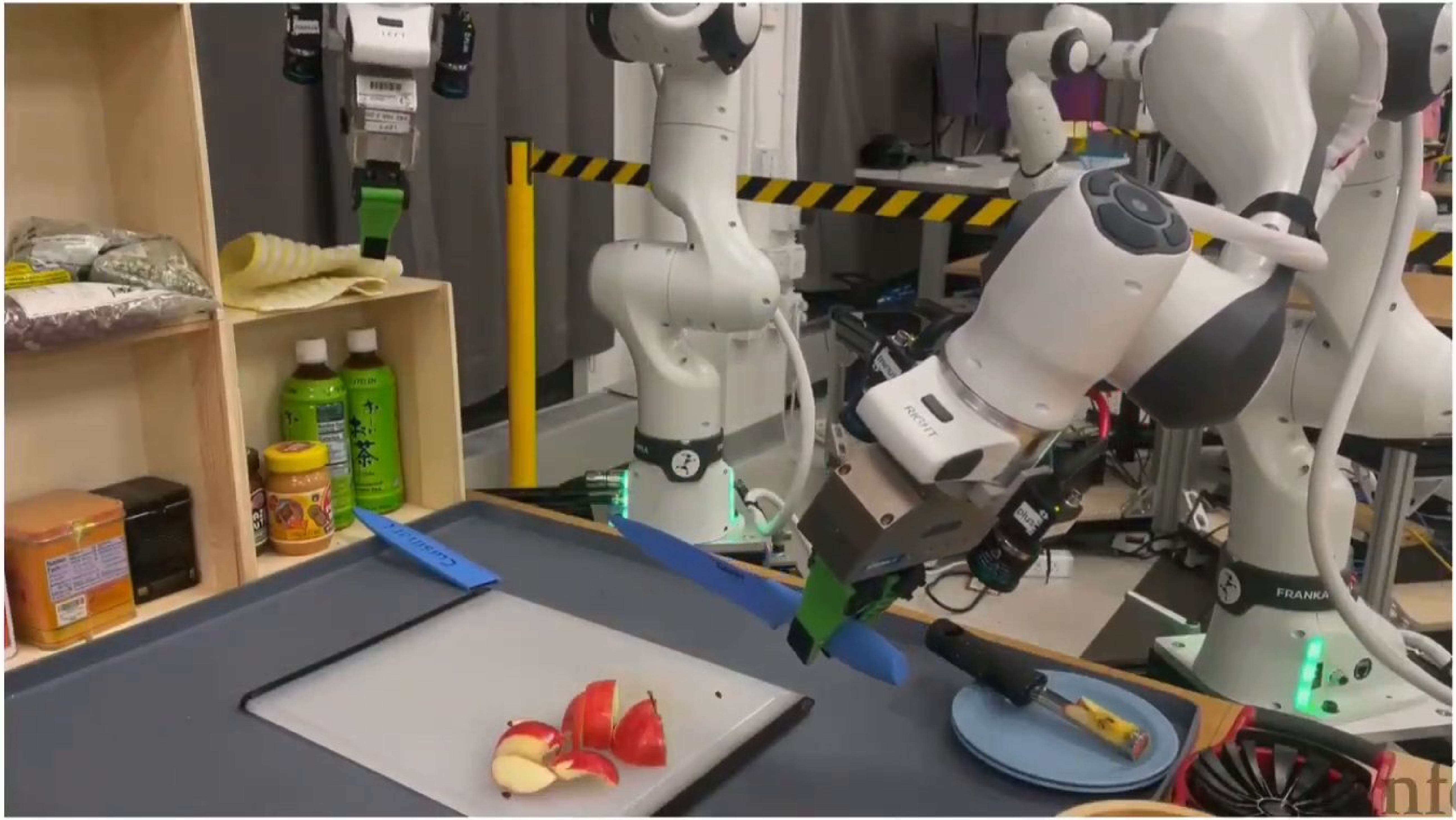
nford

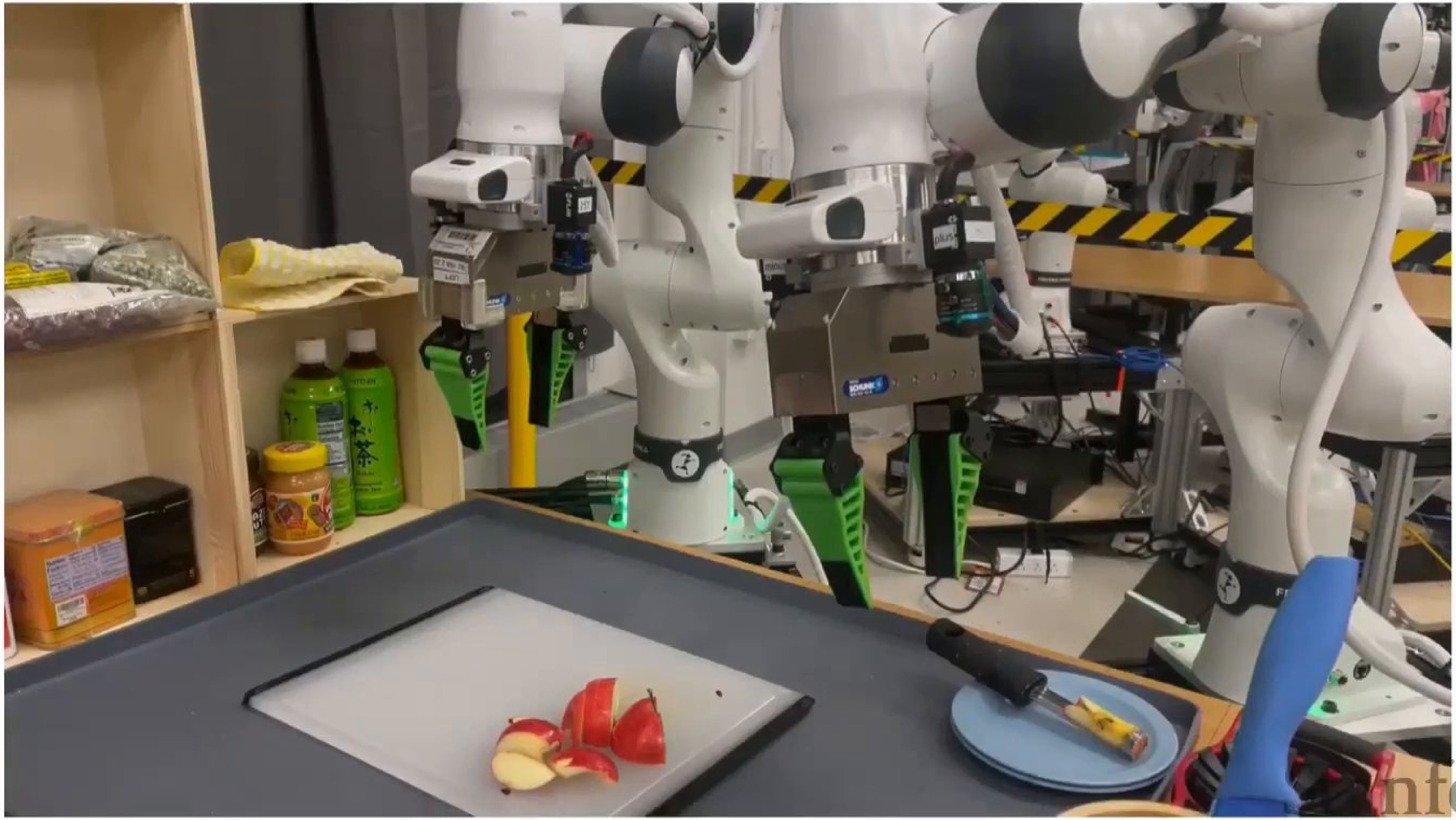



FRANKA



nford

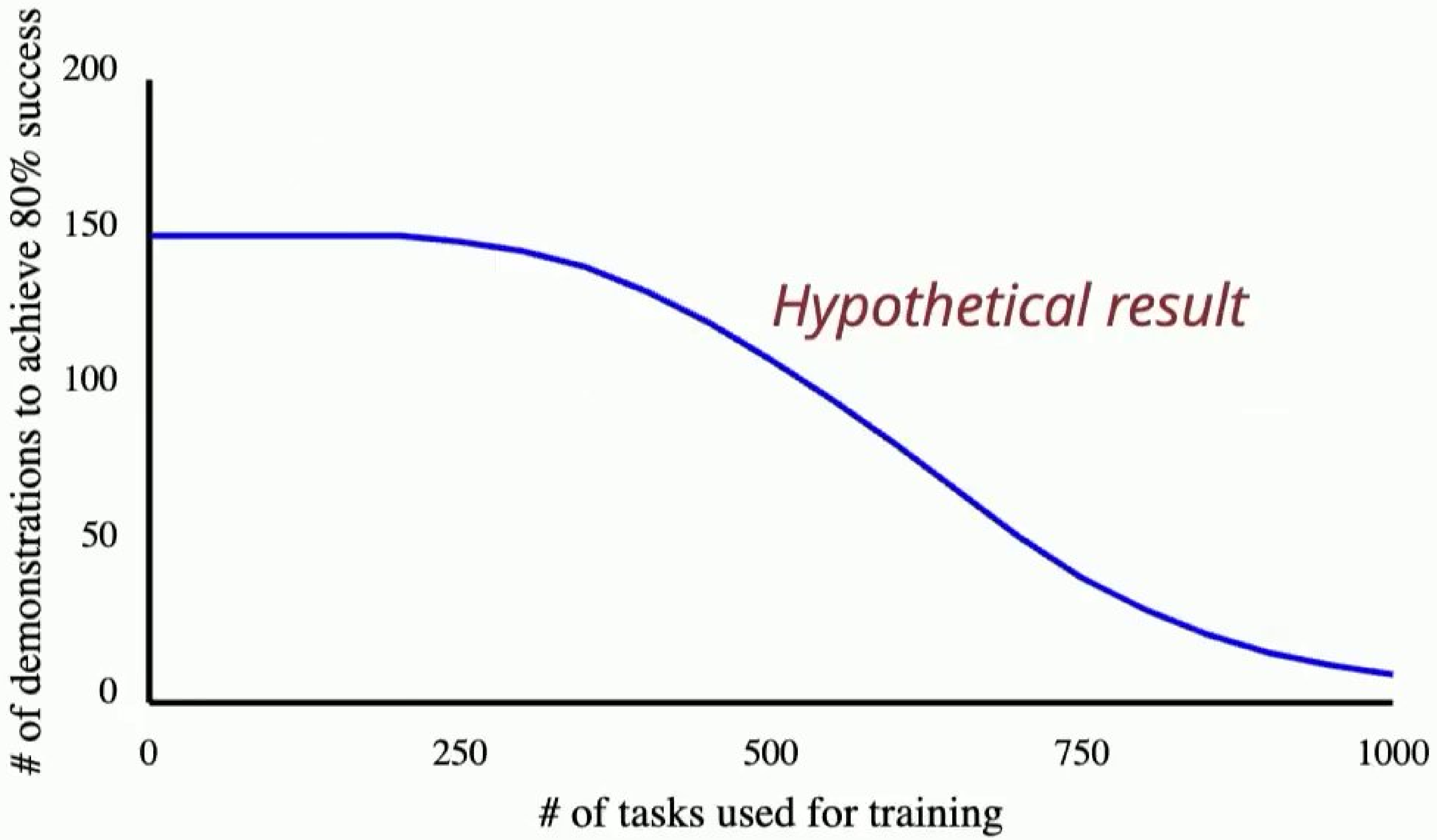

FRANKA



## Q: Does LBM pretraining help on new single tasks?

**Hypothesis 1**: Multitask scaling laws.

By training on many tasks, we can add new tasks more quickly.



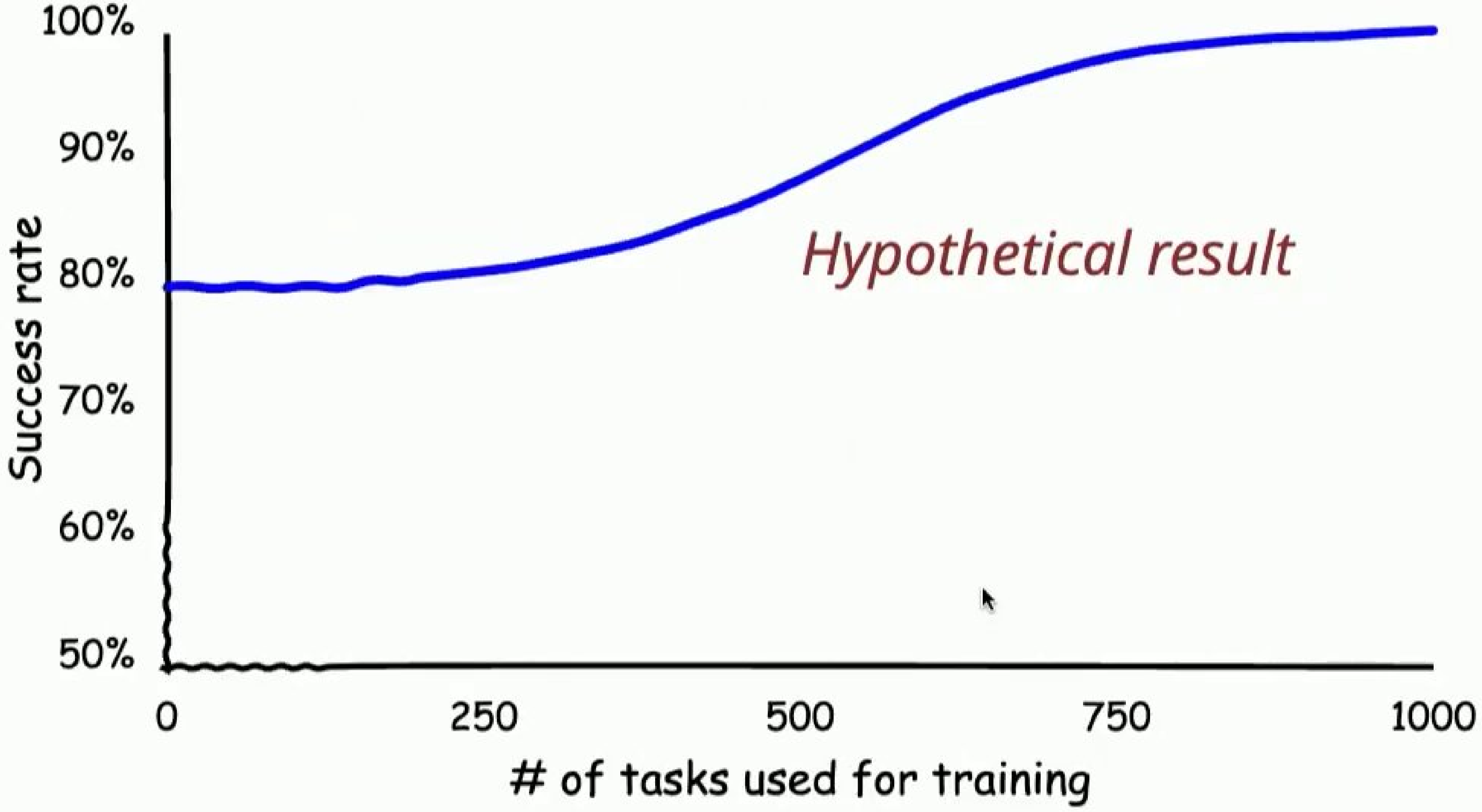

Stanford


Stanford

**Hypothesis 2:** Improved robustness.

Due to more diverse training data; skill transfer.

(Ex: a recovery move from one task transfers to another.)




Stanford


Stanford

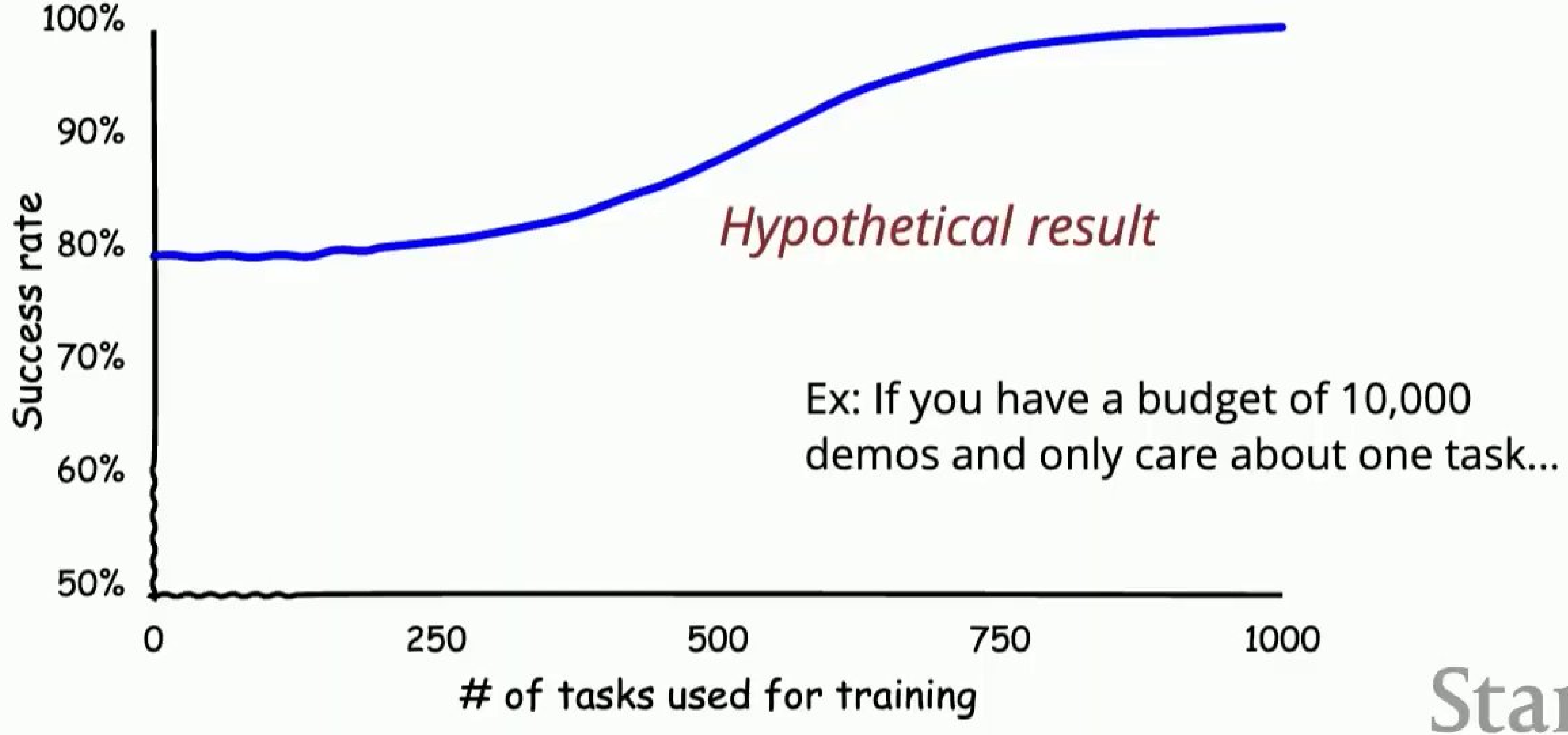

Stanford


Stanford

**Hypothesis 2:** Improved robustness.

Due to more diverse training data; skill transfer.

(Ex: a recovery move from one task transfers to another.)




Stanford

# Today's focus: bimanual Panda stations

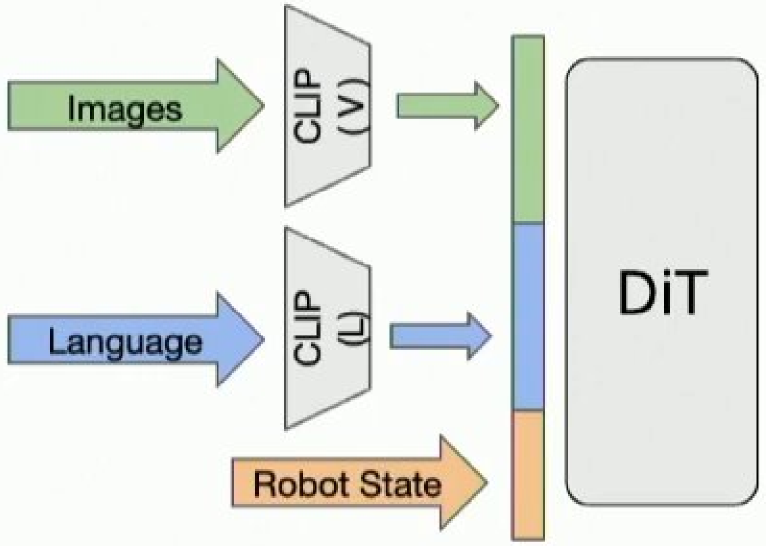
Stanford

# LBM Architectures

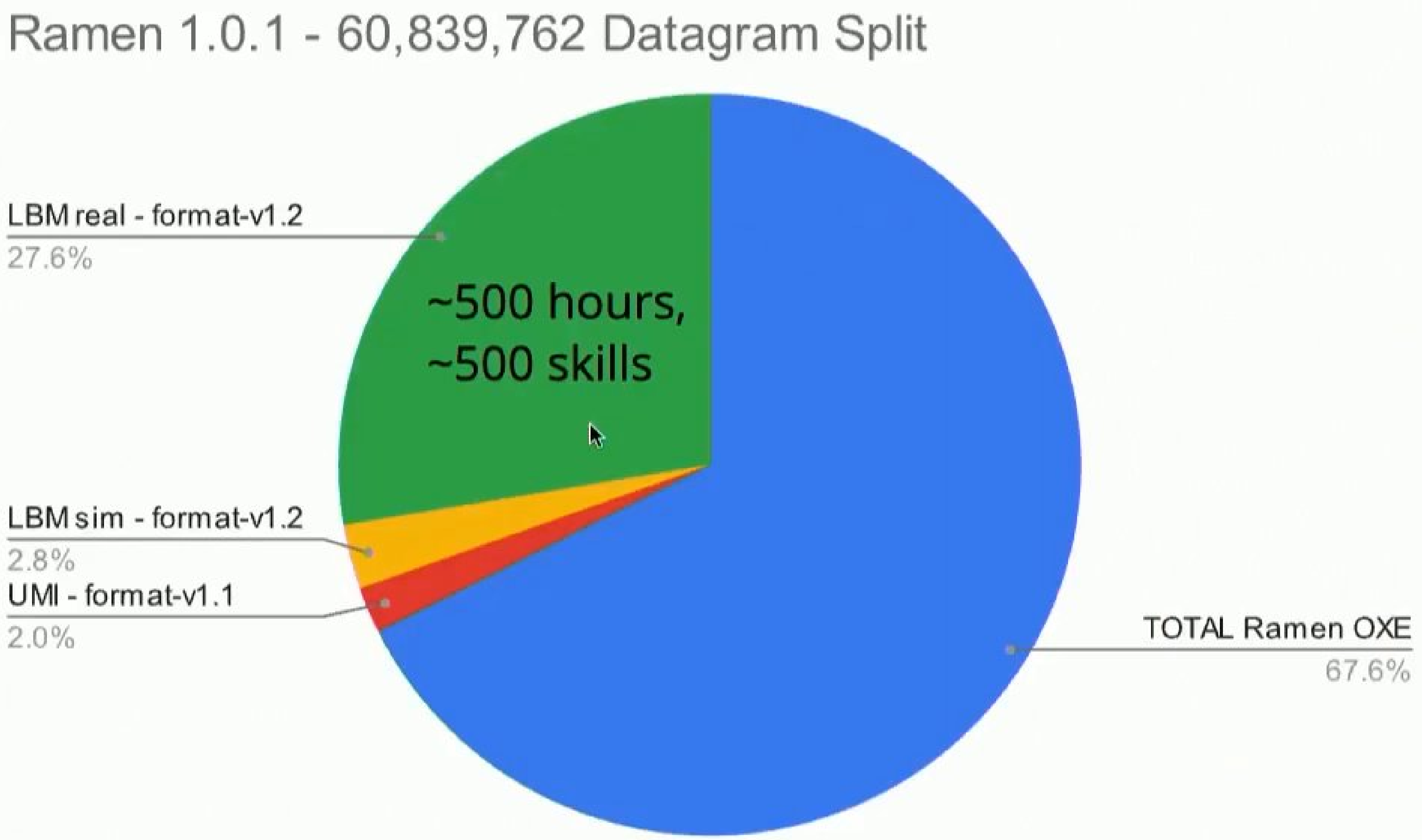
Stanford

## Today's focus: Multitask Diffusion Policy

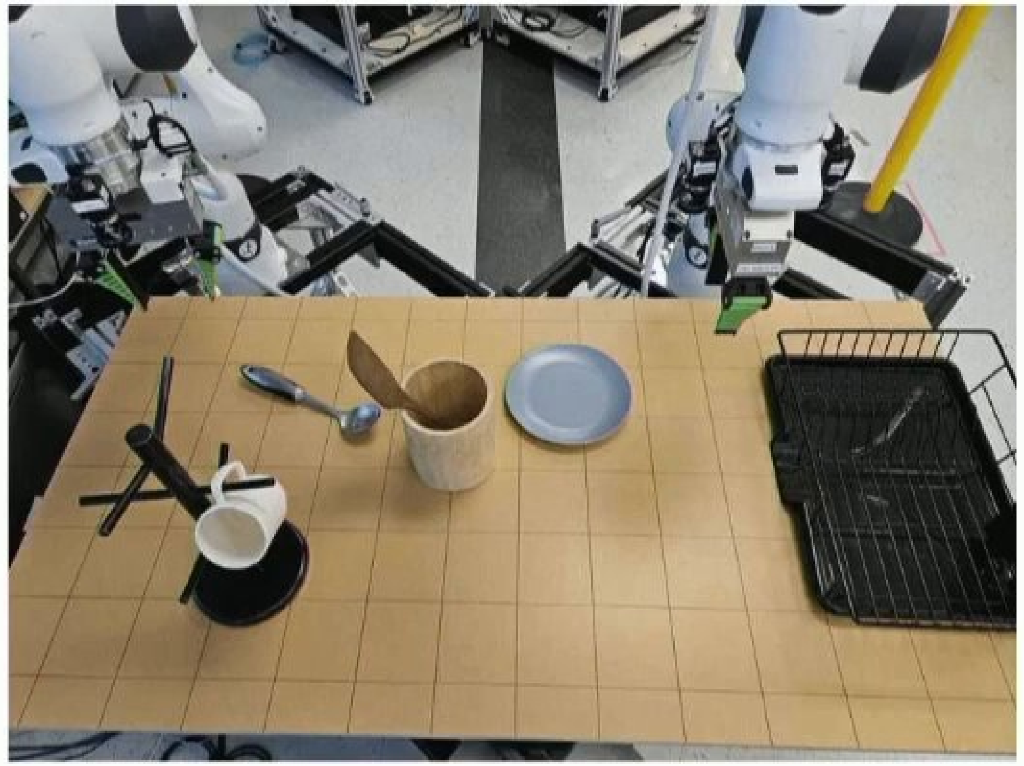
- Scaled up from original Diffusion Policy



# Data -- TRI Ramen 1.0.1



## Sim+Real cotraining




We don't use sim for datagen here;
we *cotrain* with sim teleop to use sim for eval.

Stanford

# Evaluation: Real-world hardware testing

Stanford

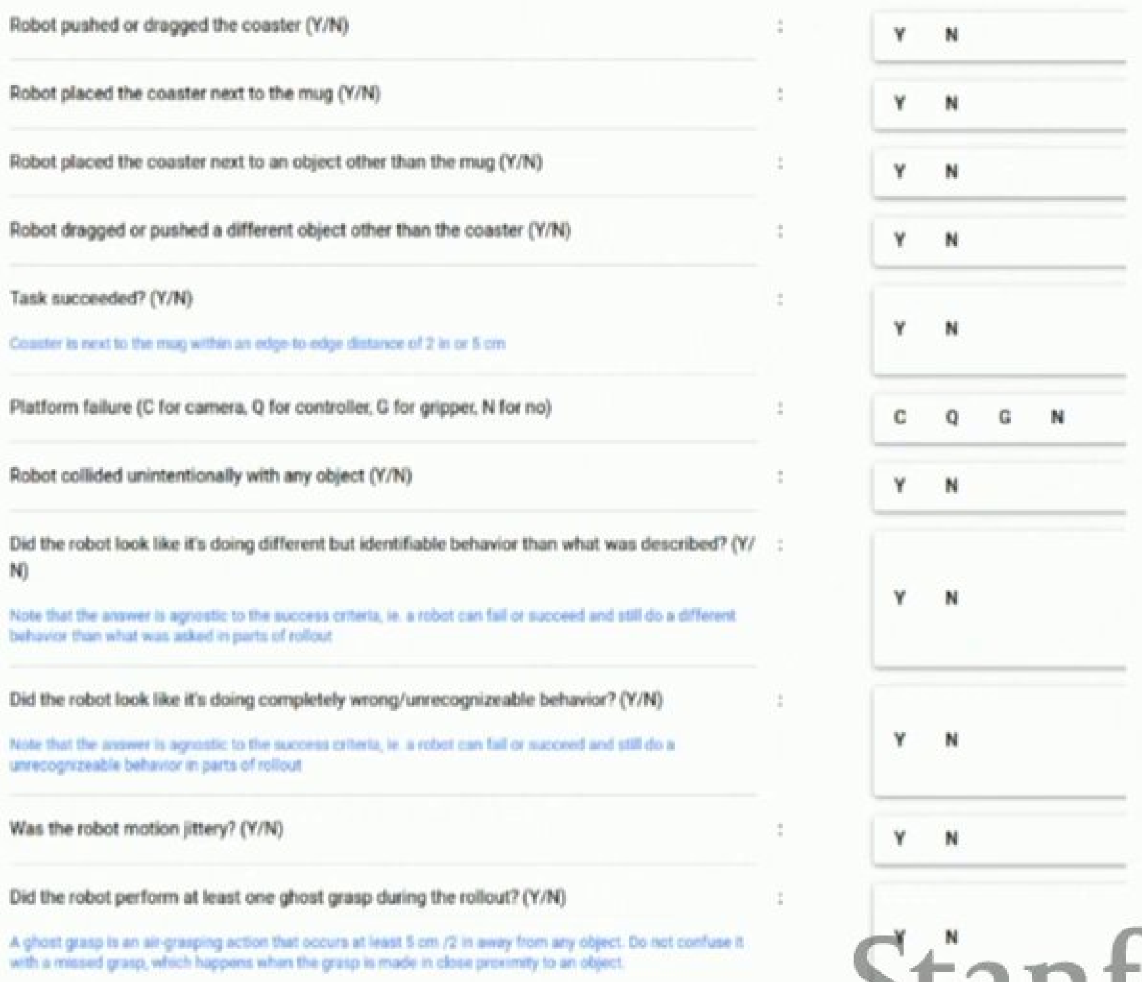
# Evaluation: Real-world hardware testing

- A/B testing ⇒ study *relative* scores
- Always blind, randomized trials
- Evaluation GUI for repeatable ICs


Stanford

# Evaluation: Real-world hardware testing

- A/B testing ⇒ study *relative* scores
- Always blind, randomized trials
- Evaluation GUI for repeatable ICs
- Rich reporting on "rubrics" and post-hoc video analysis

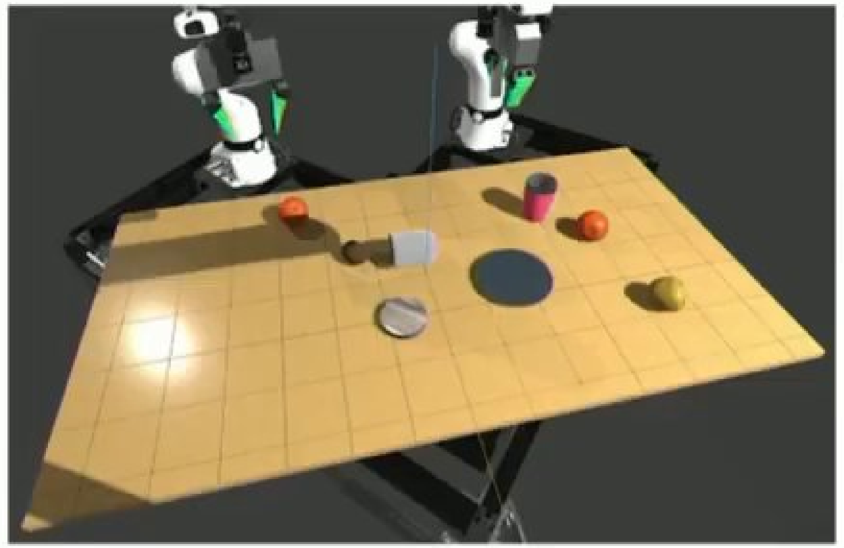



Stanford

## Evaluation: Simulation-based testing


Scenario 1: Drying Rack


Scenario 2: Shelf



Scenario 3: Breakfast

...

Stanford

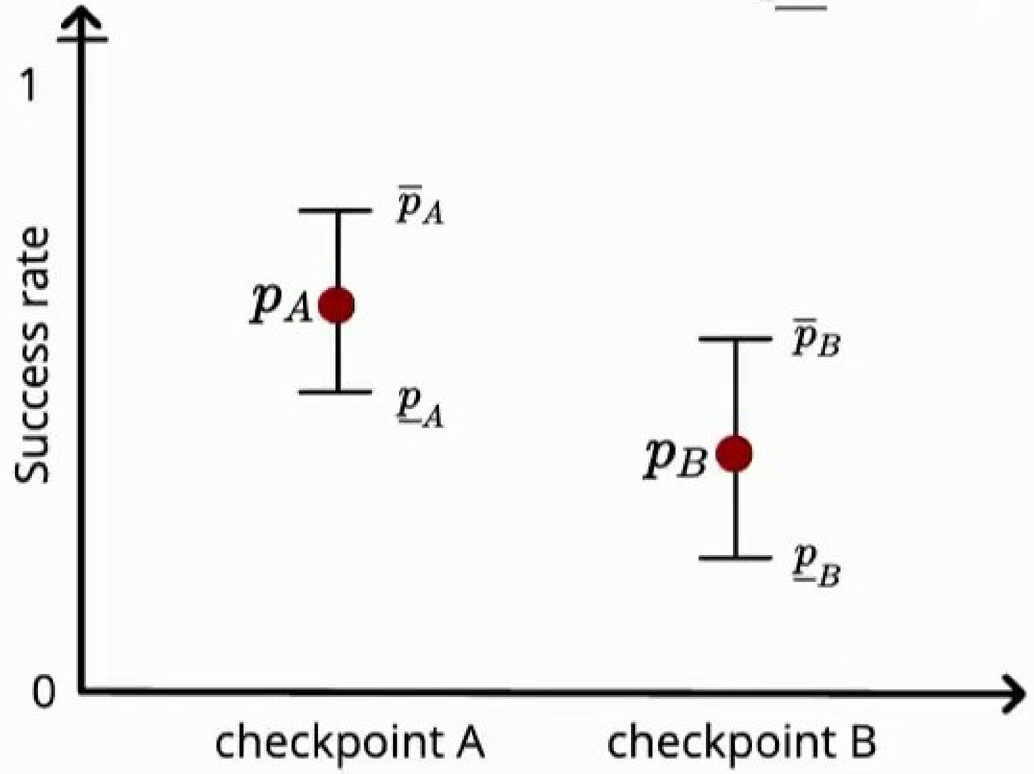
# Evaluation: Simulation-based testing

Stanford

## Statistical testing: Confidence intervals

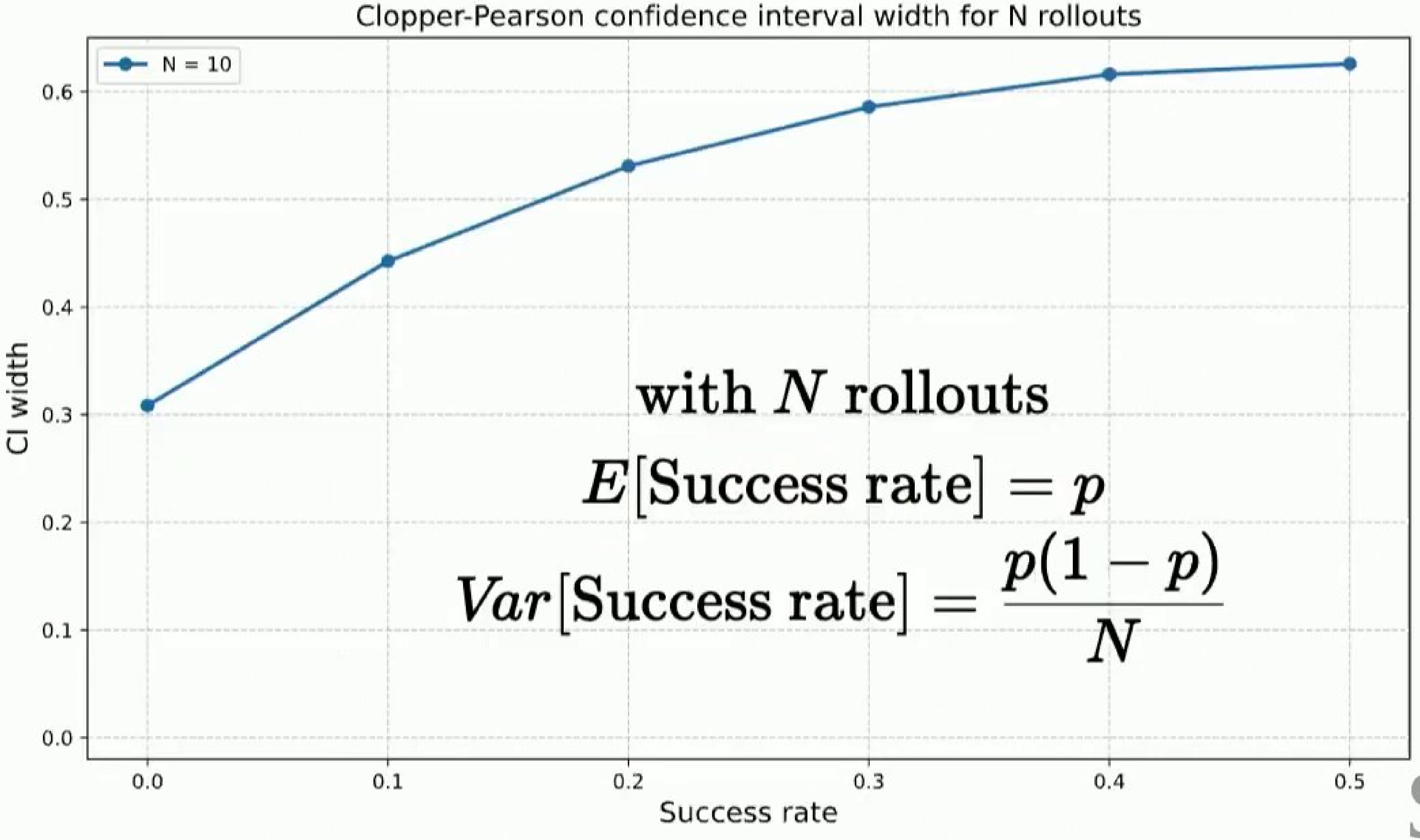
- Assuming i.i.d. Bernoulli samples with *unknown* probability of success, $p$, compute confidence intervals $\underline{p}, \overline{p}$, such that:

$$\mathbb{P}[\underline{p} \le p \le \overline{p}] \ge 0.95$$




Stanford

## How many rollouts do we need? $\mathbb{P}[\underline{p} \leq p \leq \overline{p}] \geq 0.95$




with $N$ rollouts

$$E[\text{Success rate}] = p$$

$$Var[\text{Success rate}] = \frac{p(1-p)}{N}$$

Stanford

## Experiment design

- We're in the low data regime
  - ~200 demonstrations / task
  - vs e.g. 2000-5000 from gemini-robotics

## Experiment design

- We're in the low data regime
  - ~200 demonstrations / task
  - vs e.g. 2000-5000 from gemini-robotics
- Many tasks are intentionally difficult
  - diverse initial conditions
  - severe "distribution shifts"


Stanford

## Experiment design

- We're in the low data regime
  - ~200 demonstrations / task
  - vs e.g. 2000-5000 from gemini-robotics
- Many tasks are intentionally difficult
  - diverse initial conditions
  - severe "distribution shifts"
- Philosophy:
  - At 0% or 100% success rate, you have no signal
  - To learn about policies, you want to be in the middle (and do lots of tests)

Stanford


Stanford


Stanford

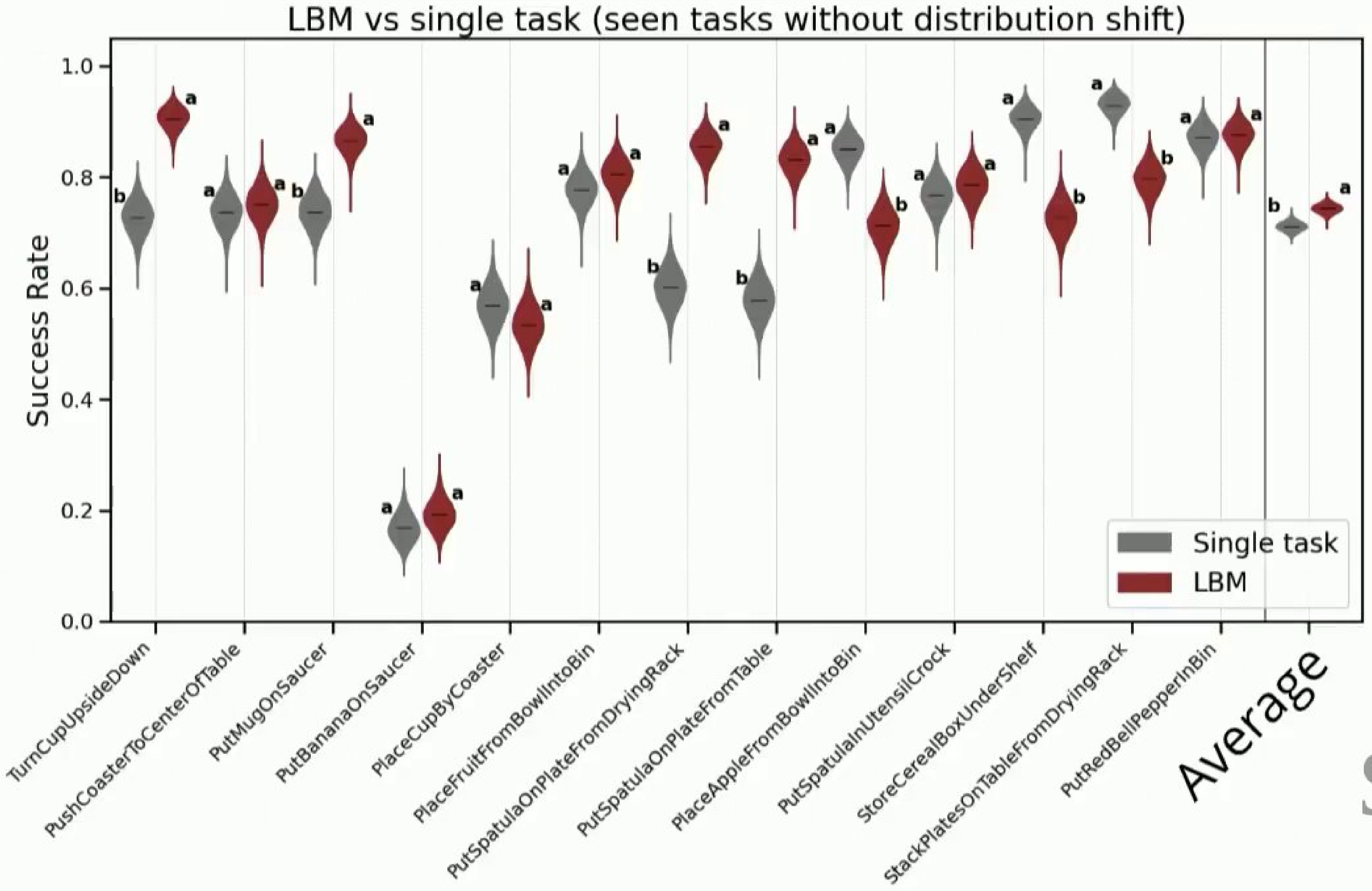
# Results

Stanford

**Warm Up**: Our multitask policy can learn *all* the tasks.

**Experiment**:

# LBM on seen tasks (sim, no finetuning)




Stanford

## There is a lot of nuance...

leaderboard

Example: *PutBananaOnSaucer*

- Low LBM score is mostly due to poor language steering.

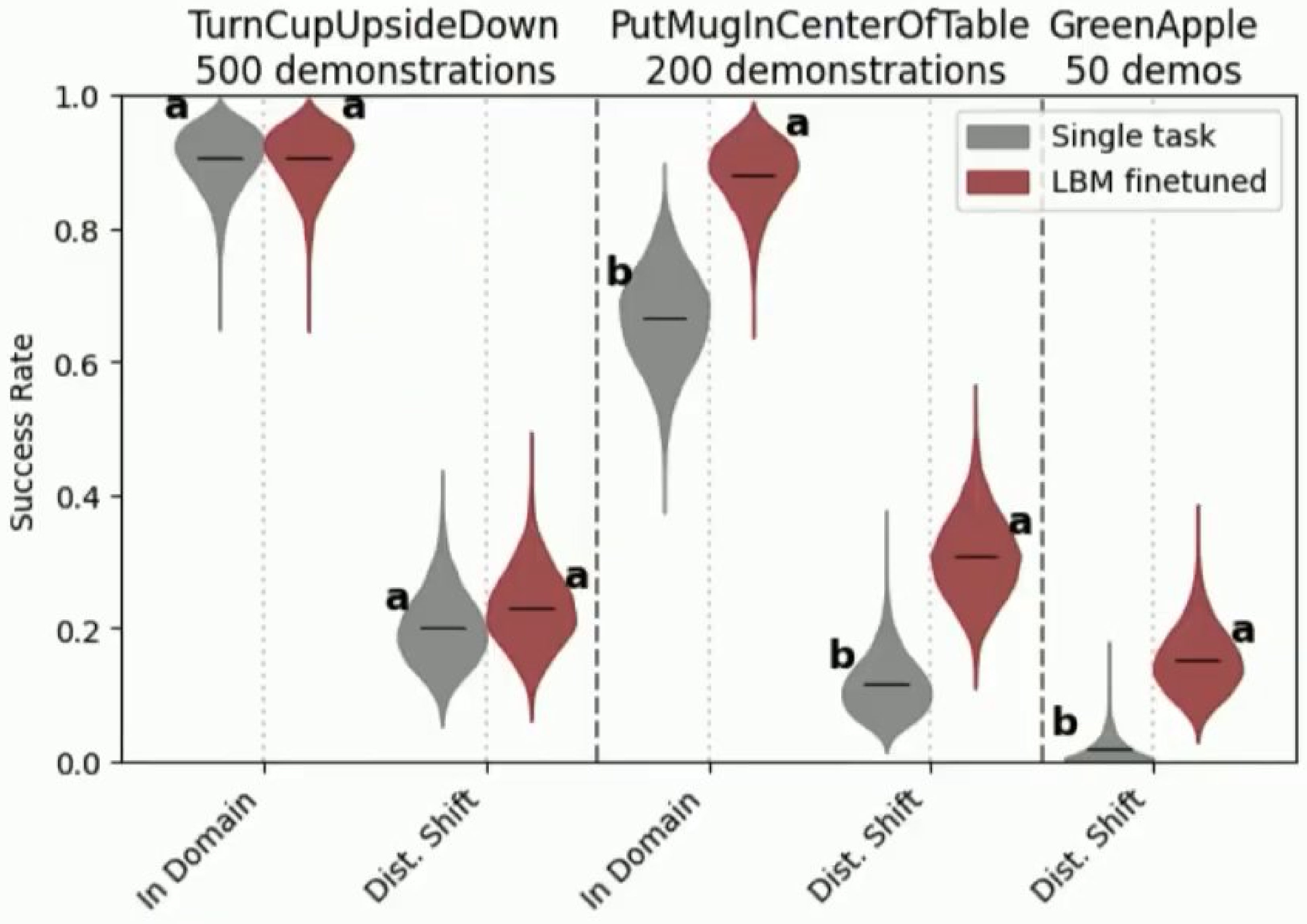
**Hypothesis 1**: By training on many tasks, we can add new tasks more quickly.

**Experiment**:

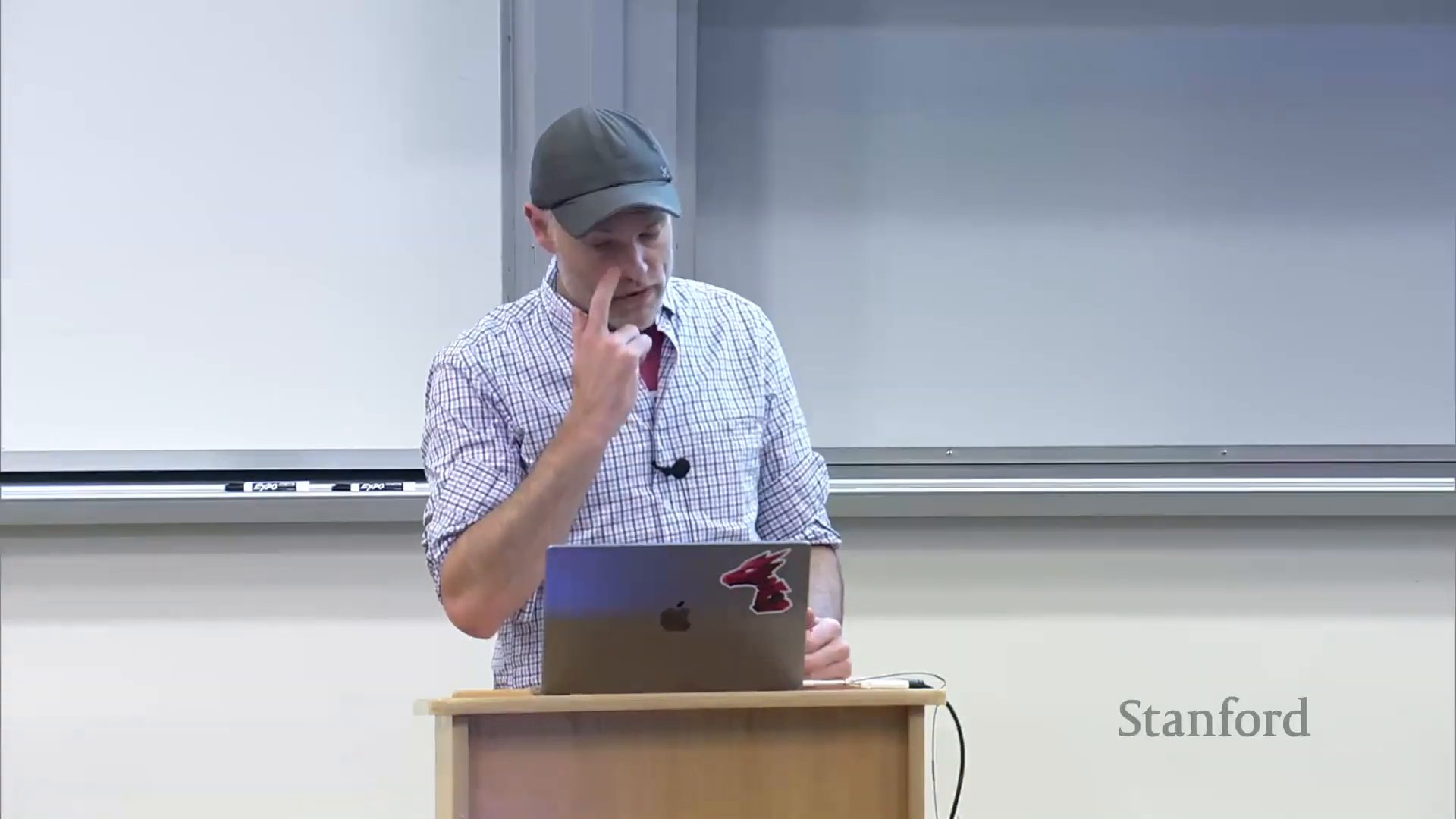
**Hypothesis 2**: Improved robustness.

**Experiment**:

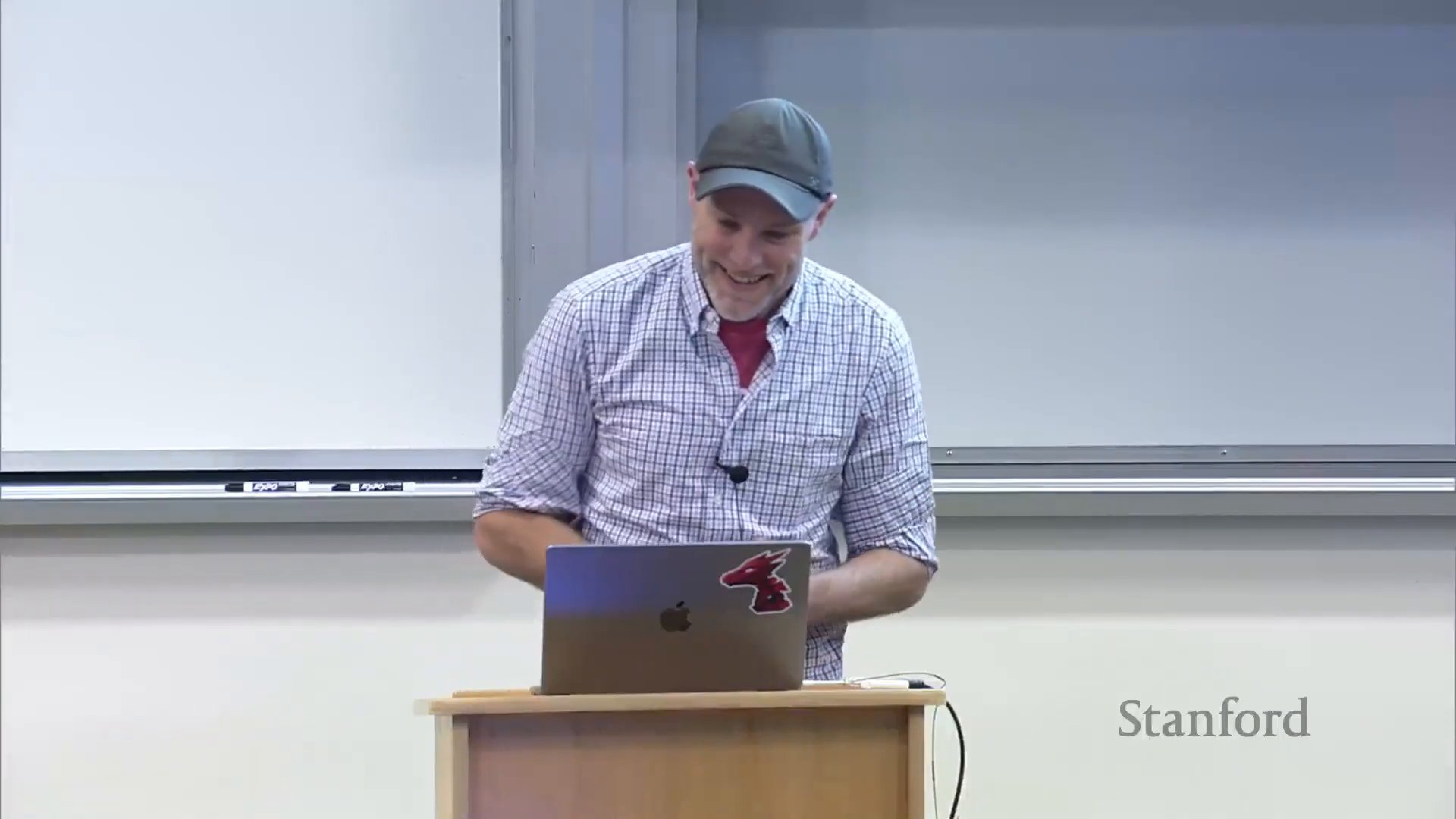
# LBMs improve robustness (real)




Stanford

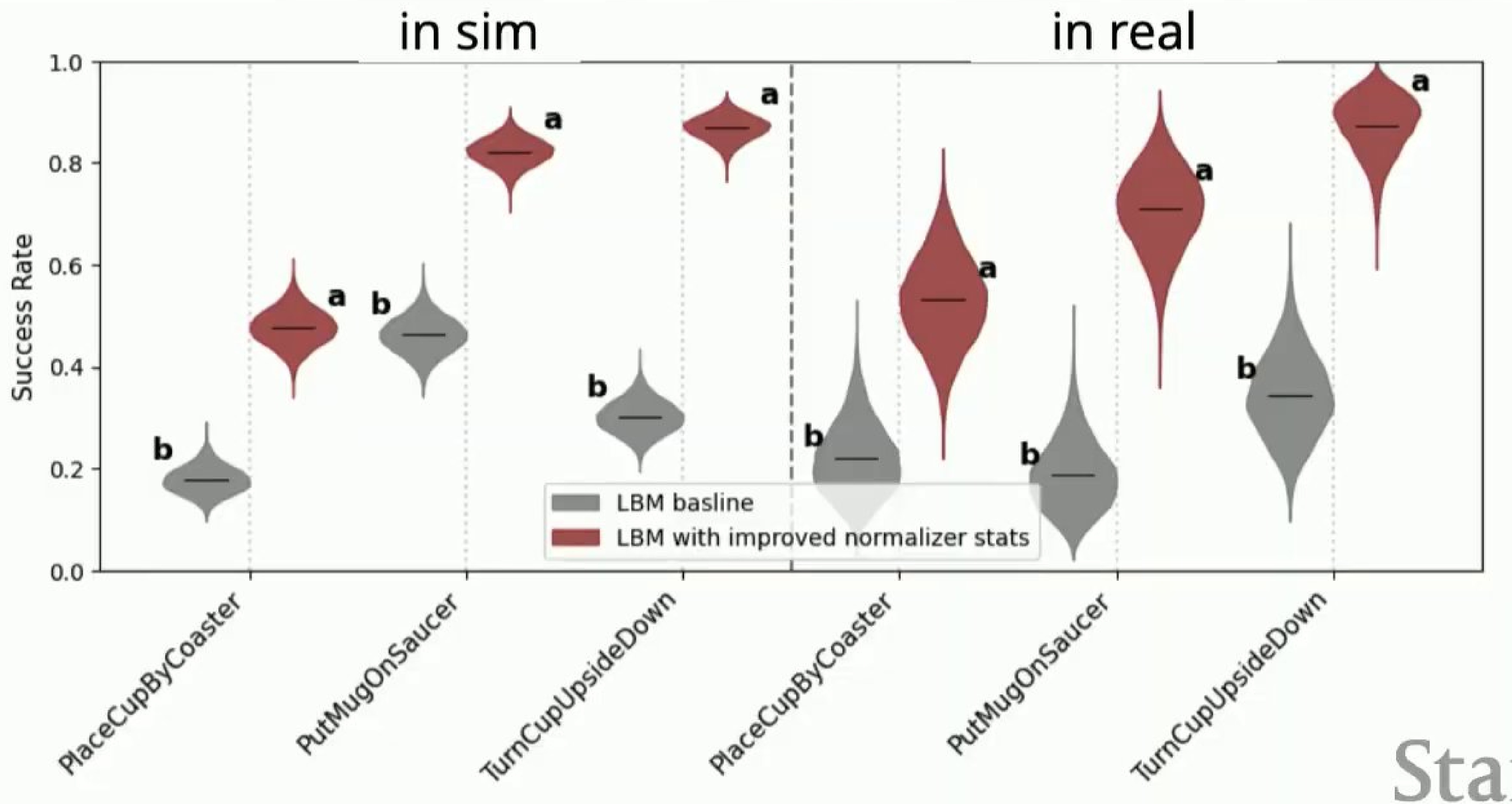

Stanford


Stanford

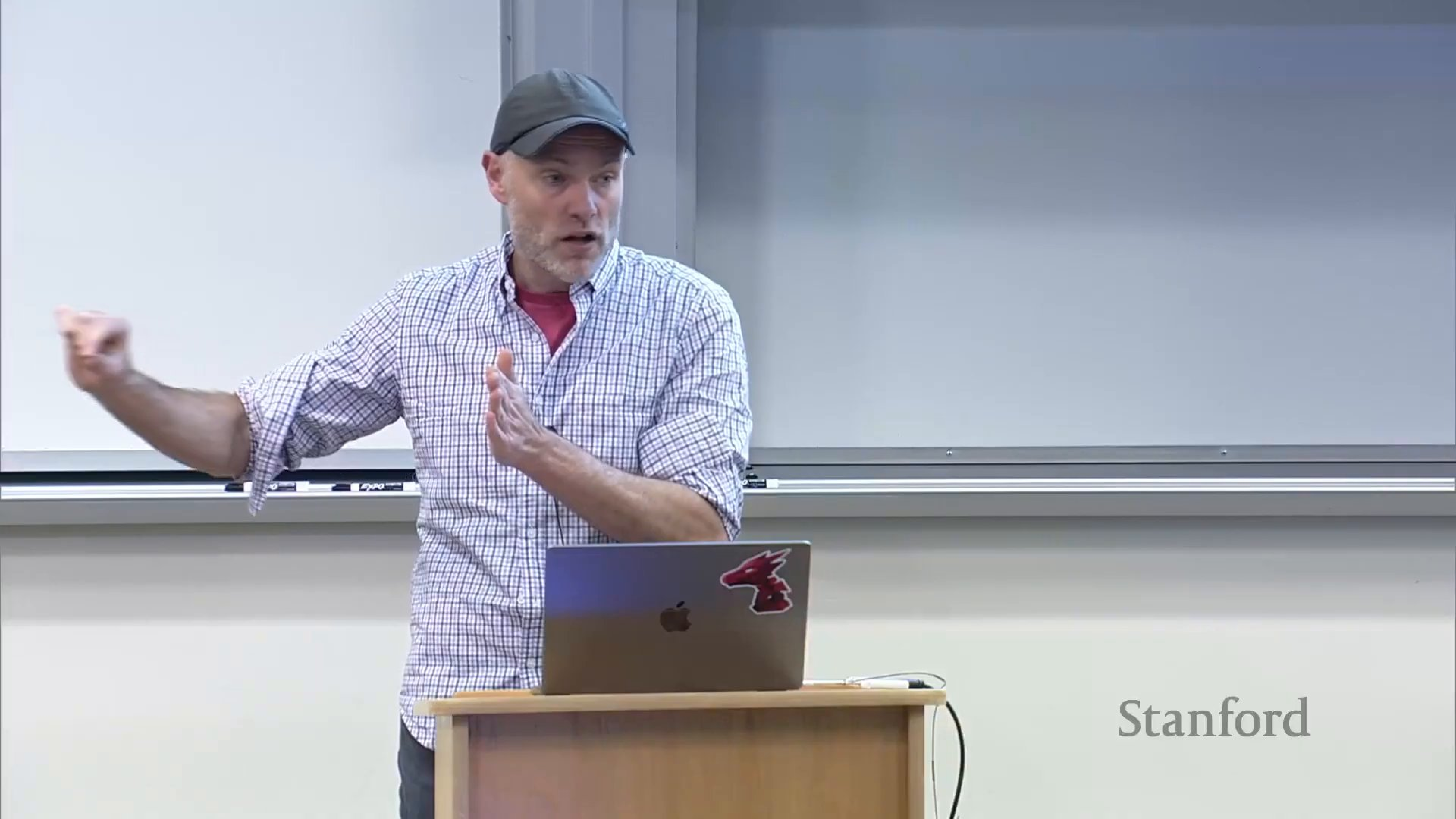

Stanford


Stanford

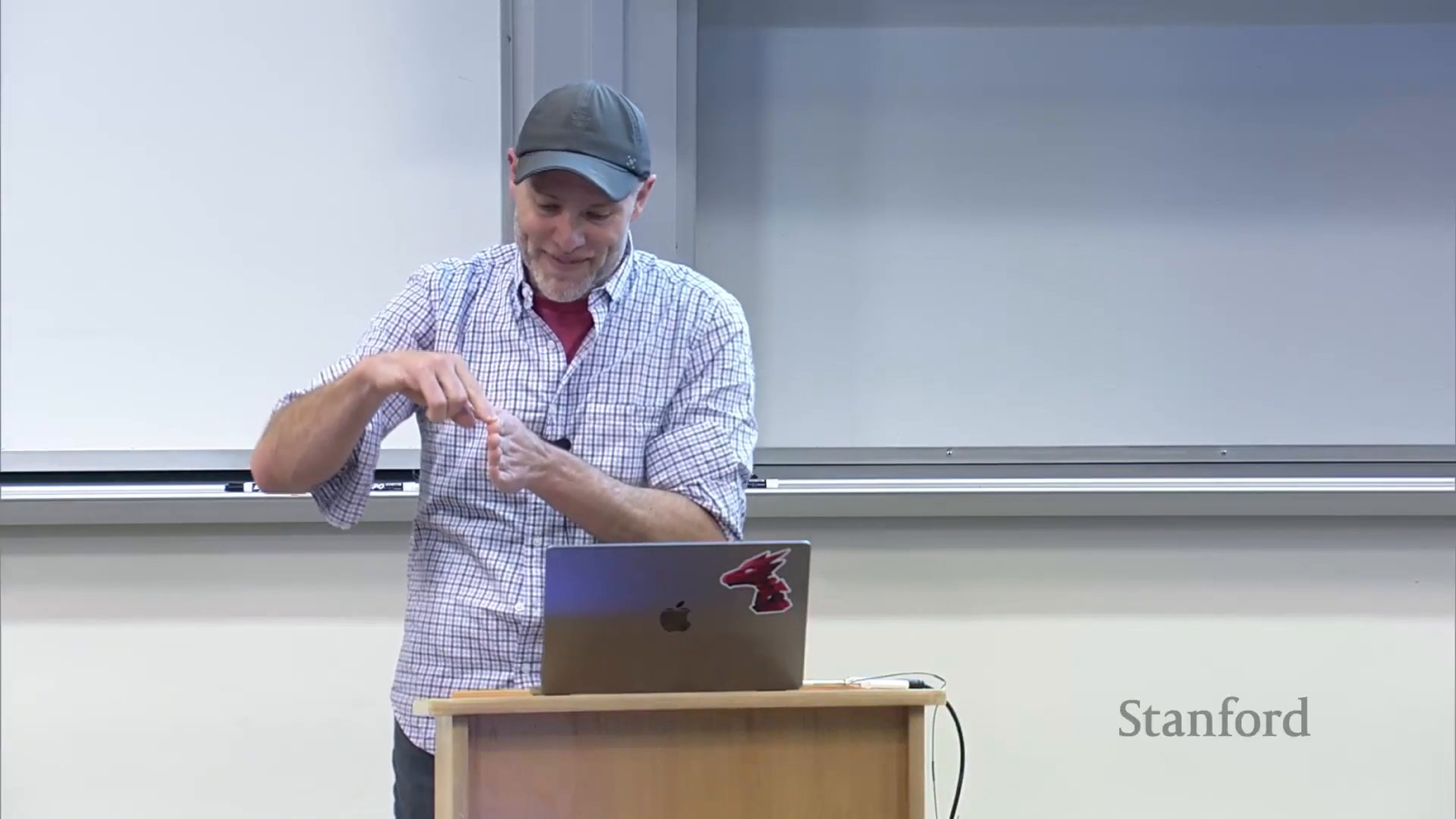
# Per-timestep normalization and clipping helps




Stanford


Stanford


Stanford

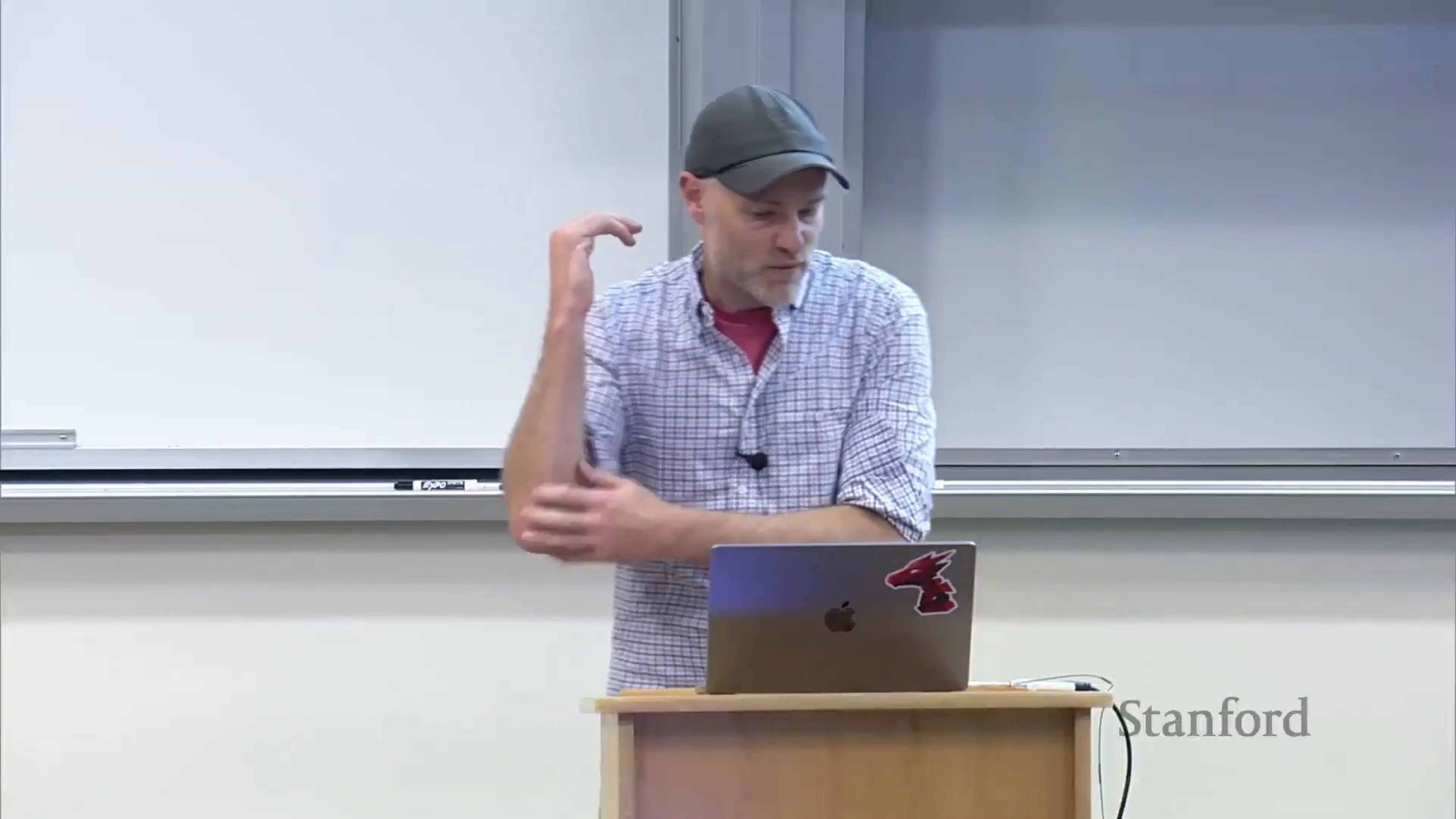

Stanford


Stanford


Stanford


Stanford


Stanford

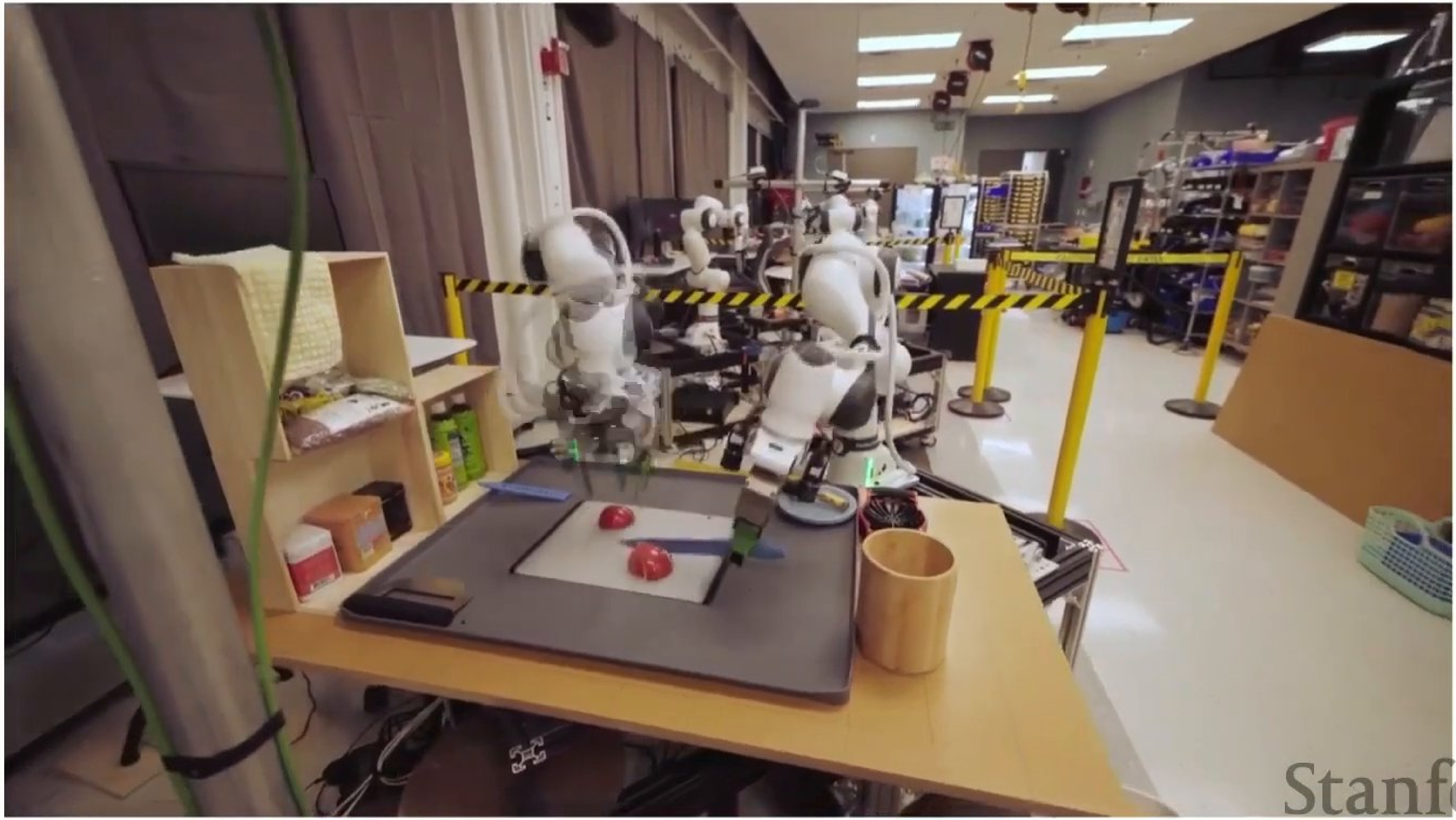

Stanford

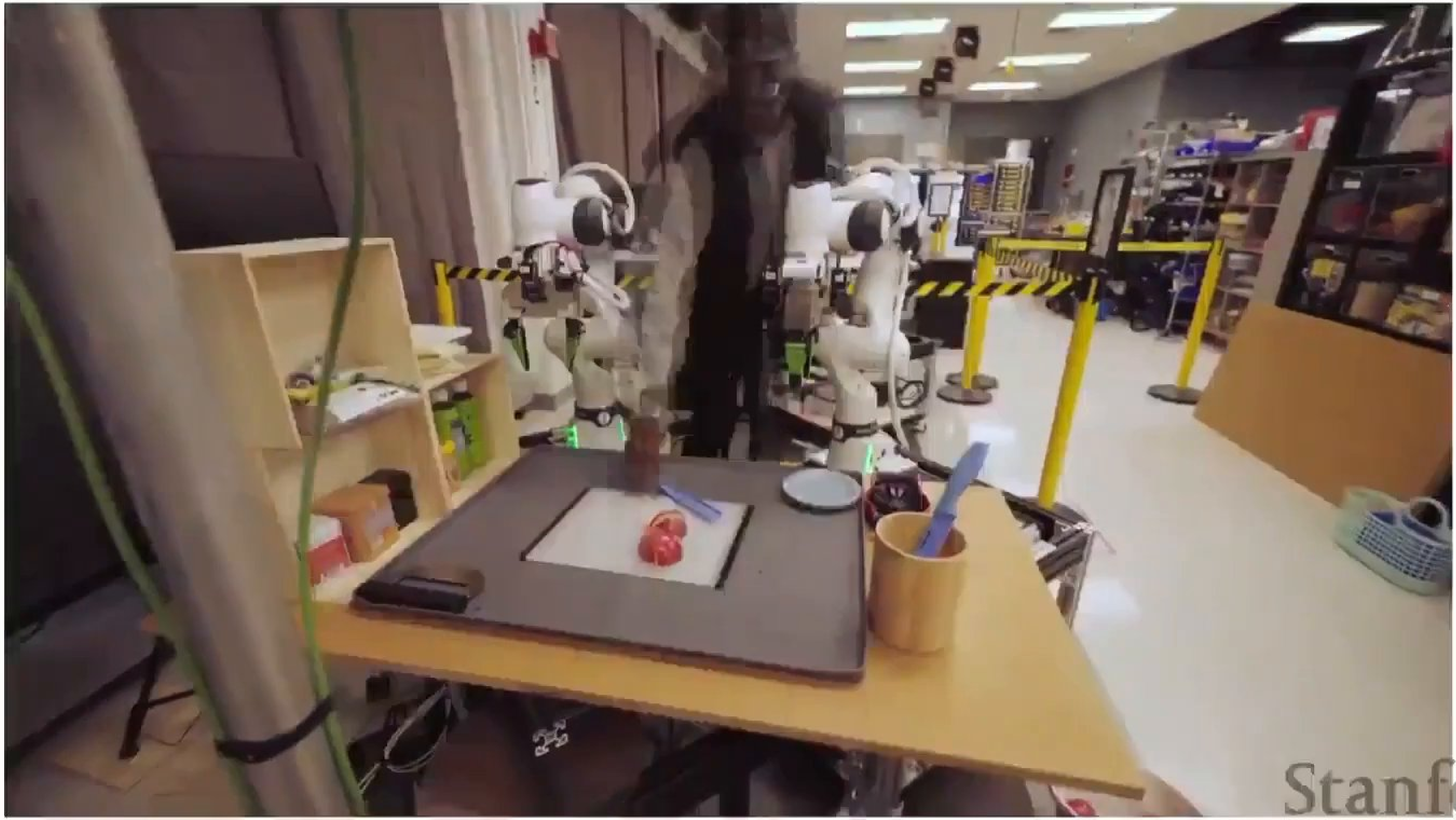
## Other take-aways (for the domain experts)

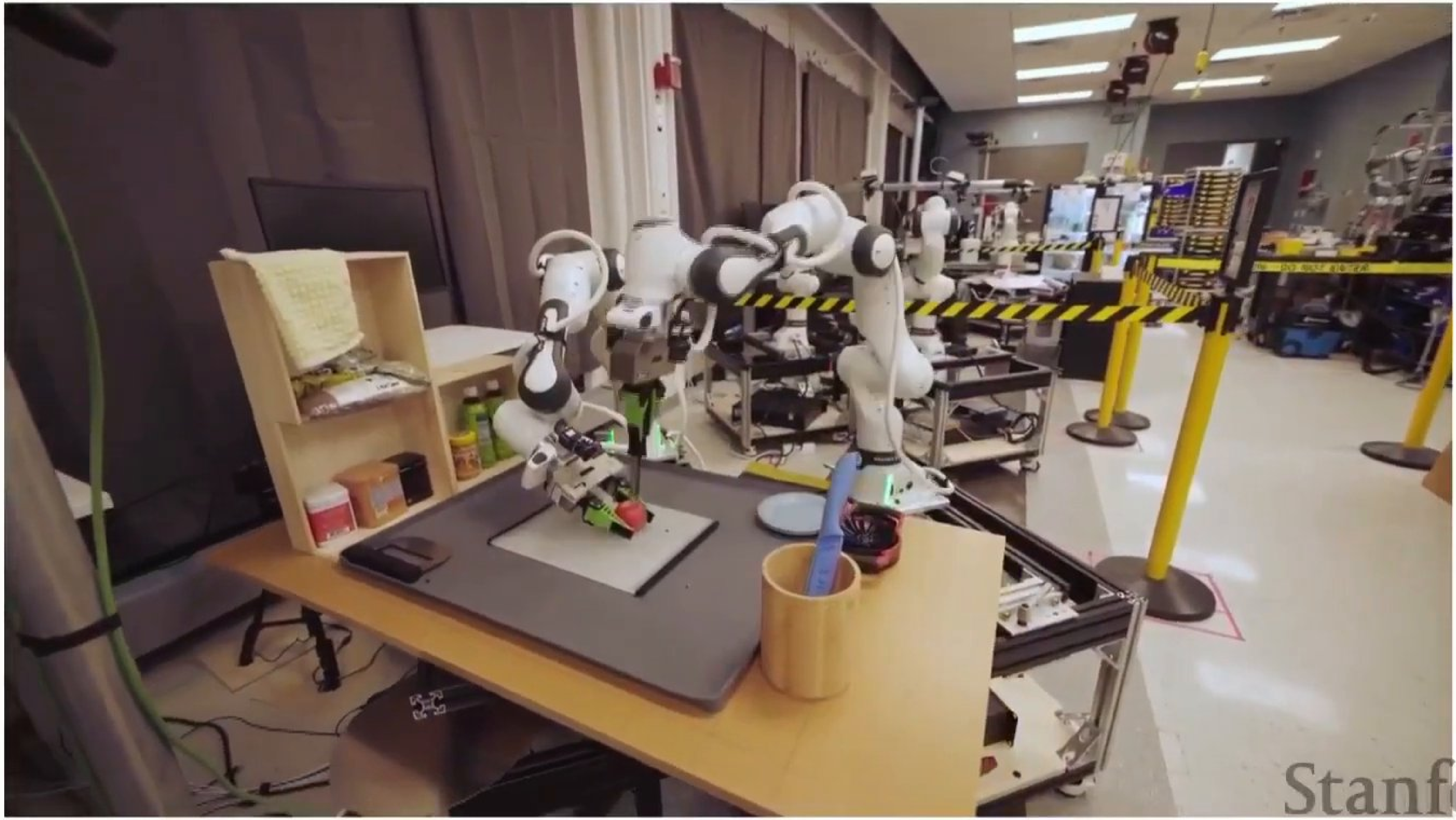
Stanford


Stanford

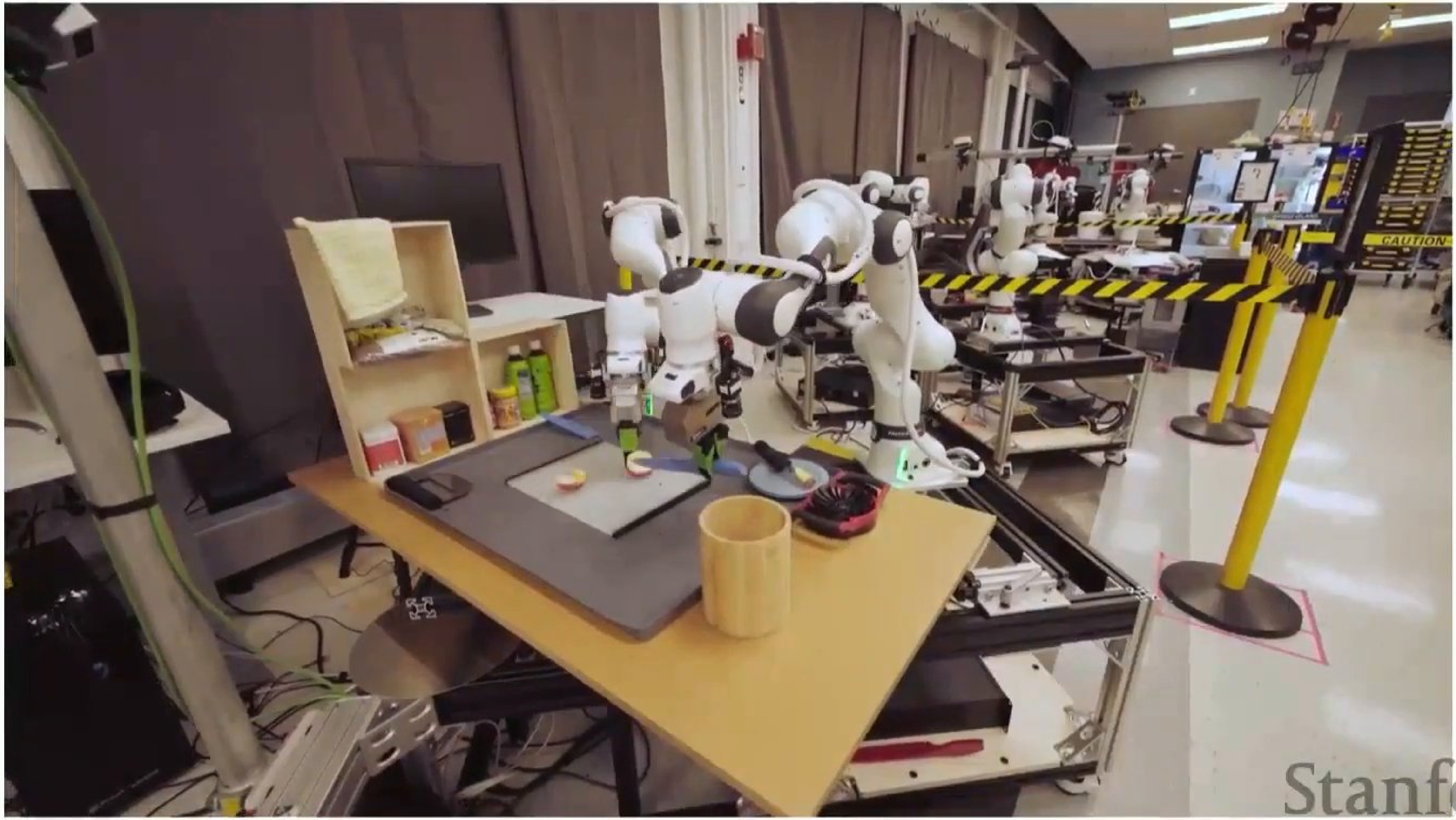

Stanford



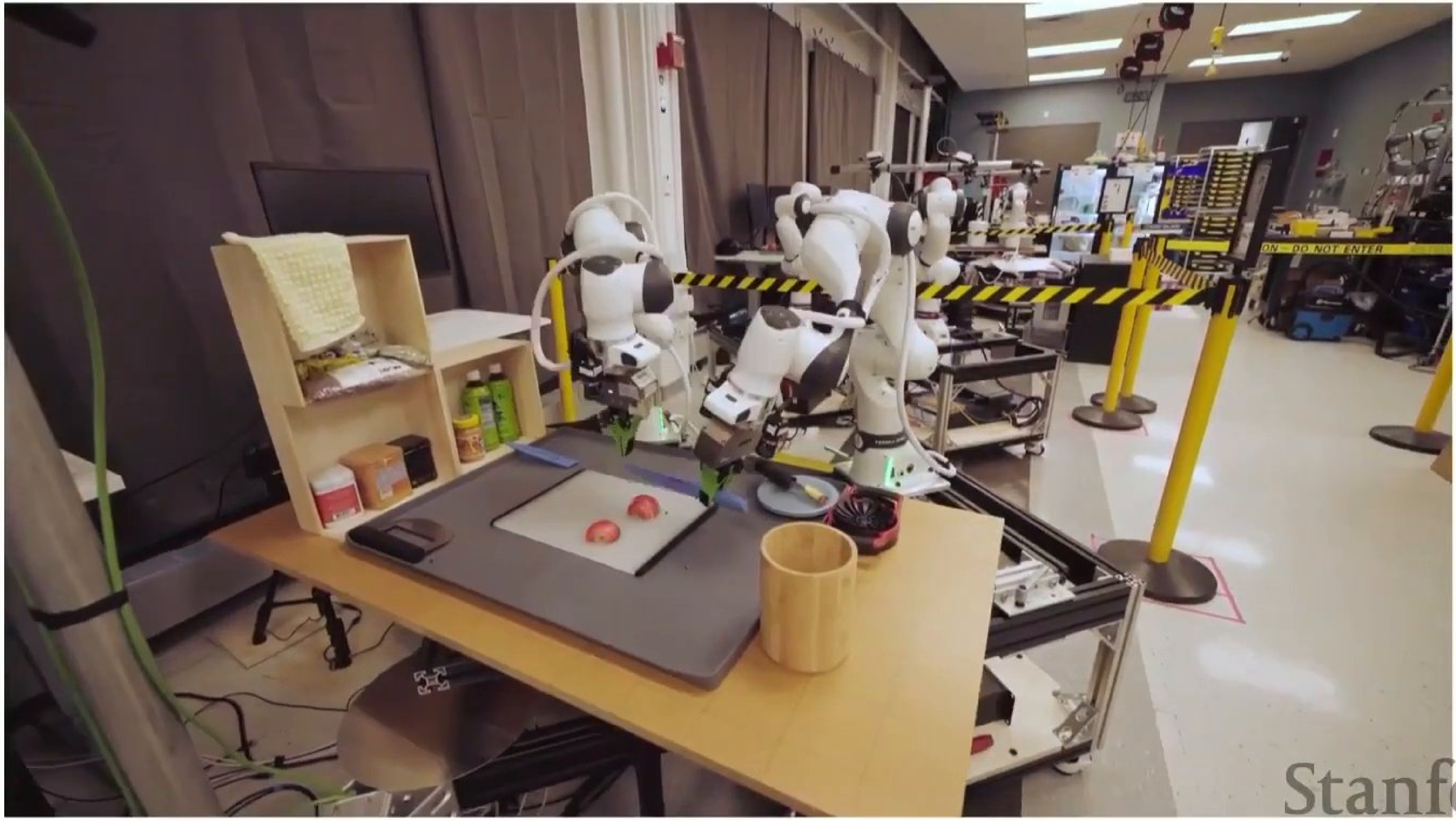


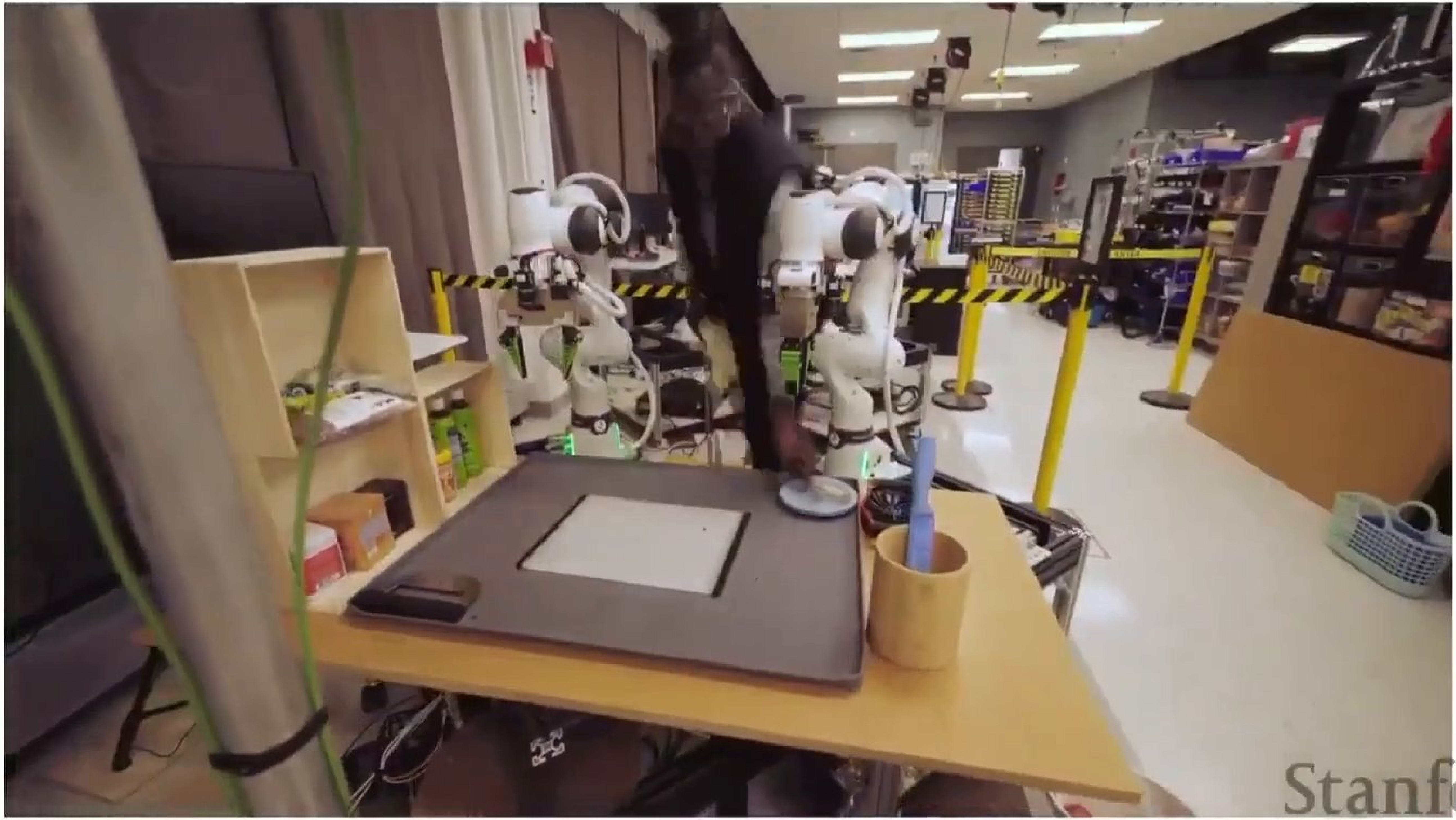


Stanford

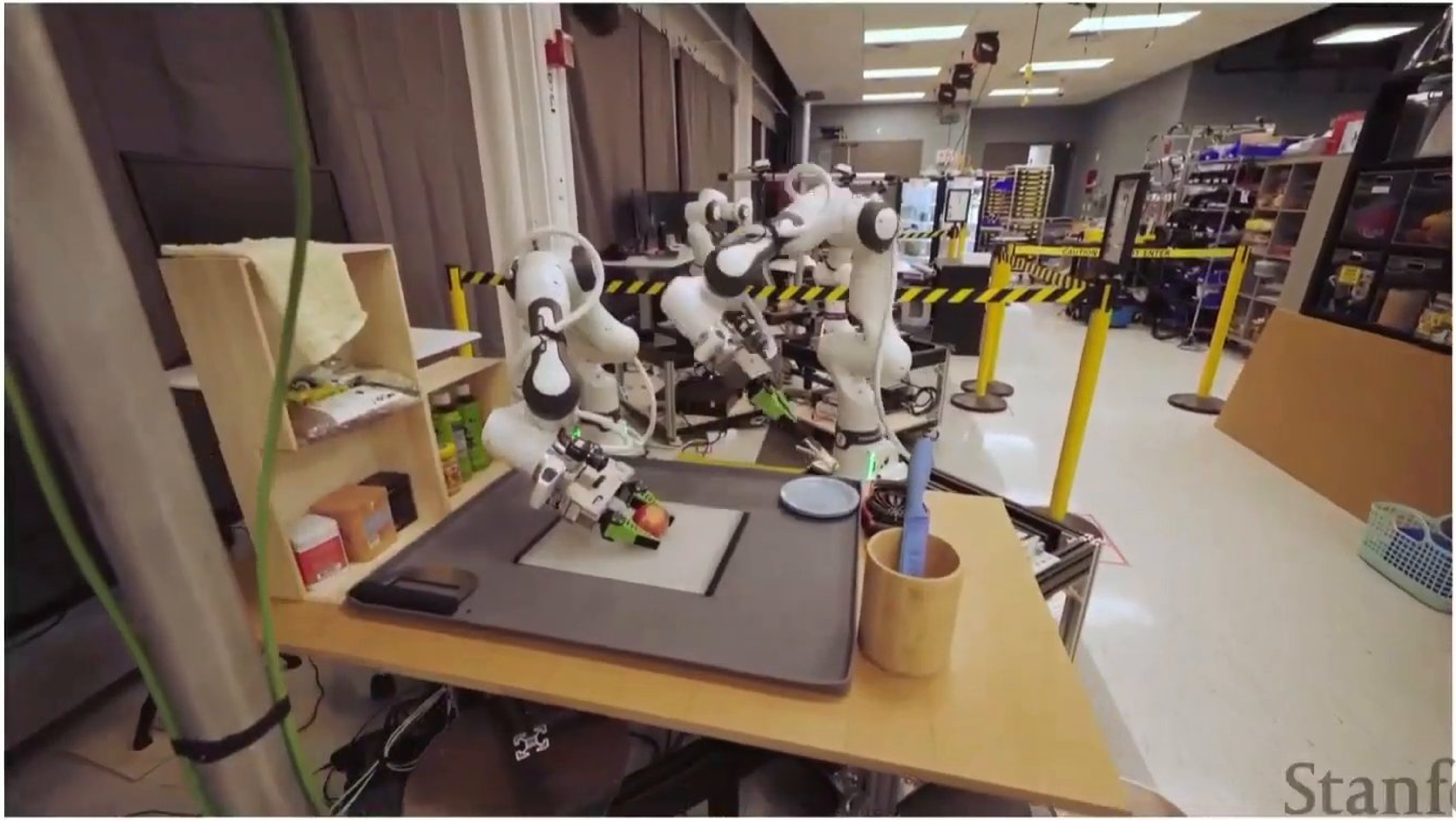



Stanford


DO NOT ENTER

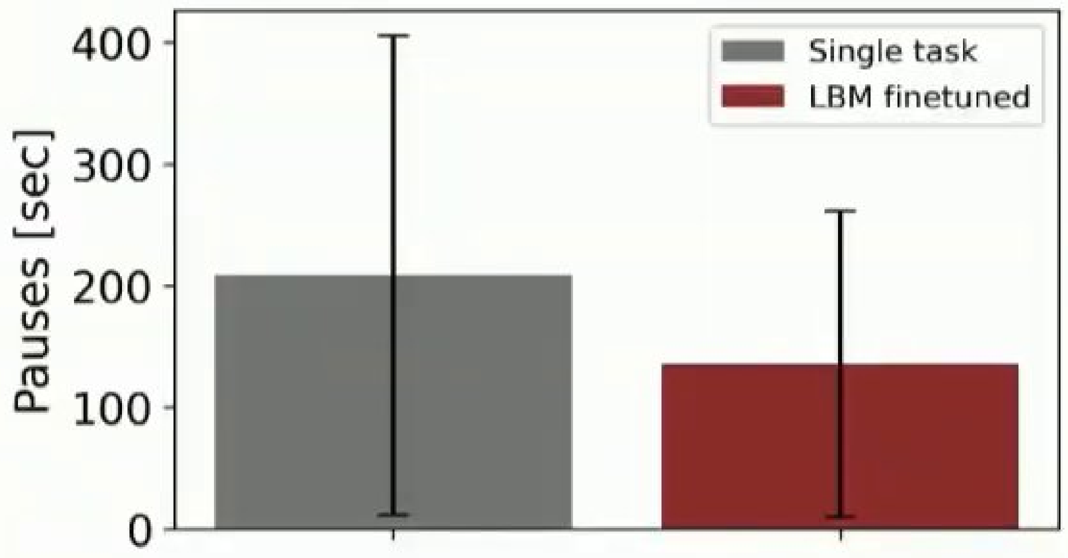


Stanford

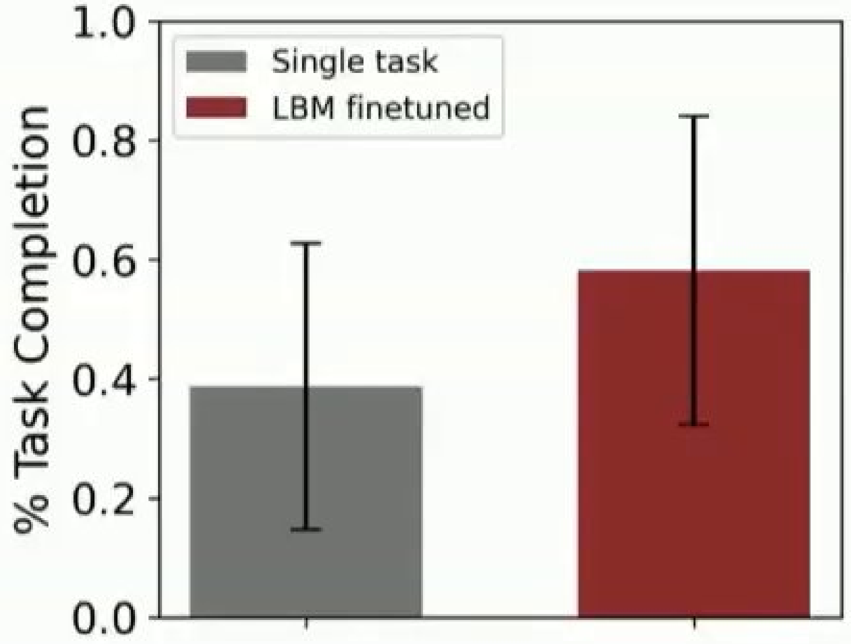


# Eval "rubrics" for our hardest (real) skills

(ClearKitchenCounter, SetUpBreakfastTable, BikeRotorInstall, CutAppleIntoSlices)




Pauses

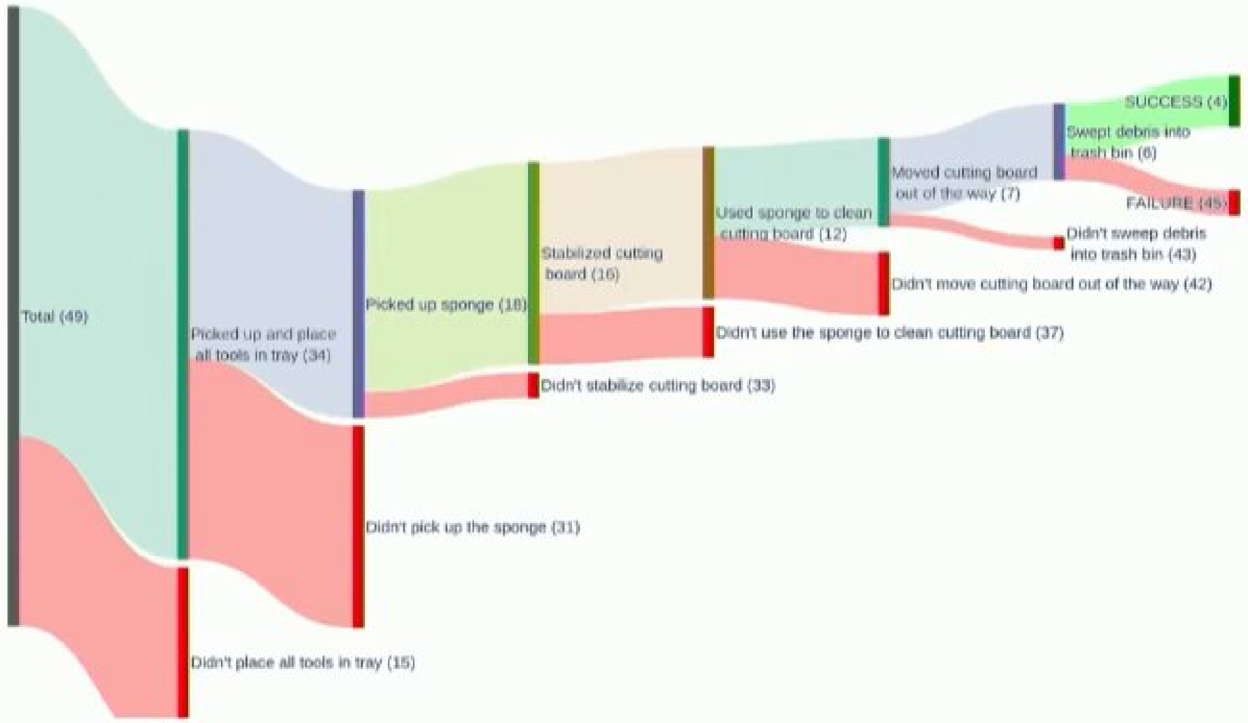



Task completion

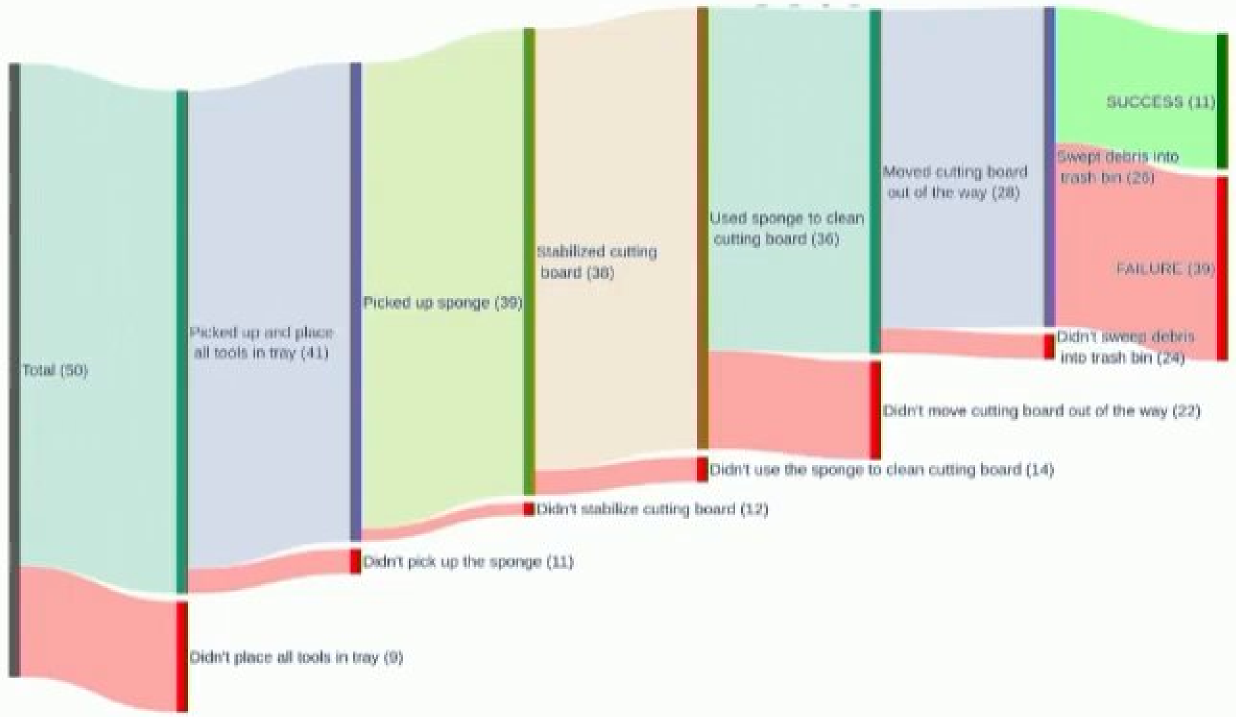
Stanford

# Failure analyses for our hardest (real) skills

ClearKitchenCounter




Single task




LBM finetuned

Stanford

## The unvarnished truth

We reviewed the films... these are the *worst* rollouts we saw for the bike repair:


Single-task (Diffusion Policy)

Stanford

## The unvarnished truth

We reviewed the films… these are the *worst* rollouts we saw for the bike repair:


Single-task (Diffusion Policy)


LBM finetuned

Stanford

## Discussion

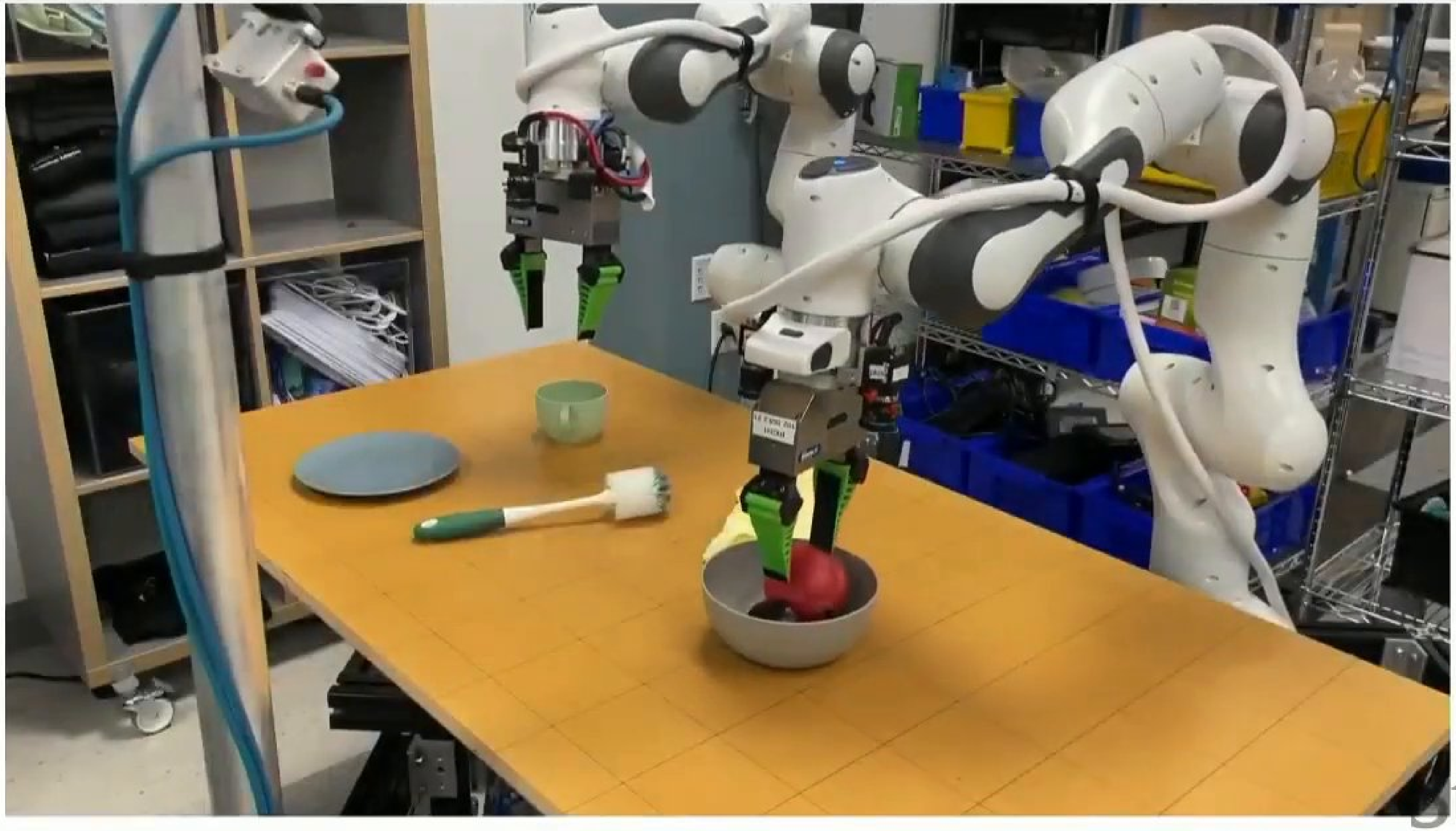
GPT doesn't always get the right answer, but it always generates beautiful prose. (Even early versions of GPT.)

Stanford

## When you don't give a language command...

Stanford

## When you don't give a language command...

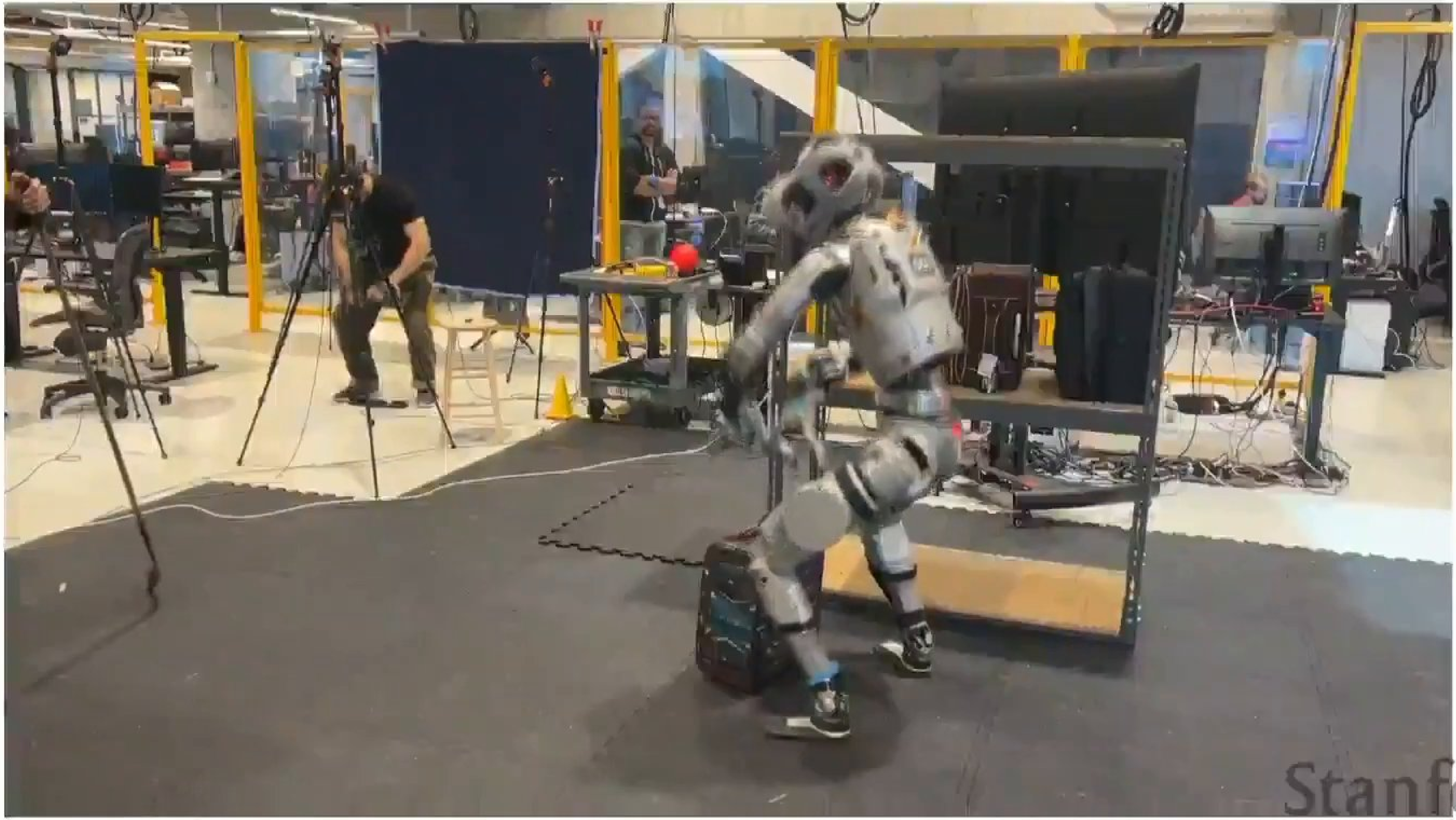


Stanford

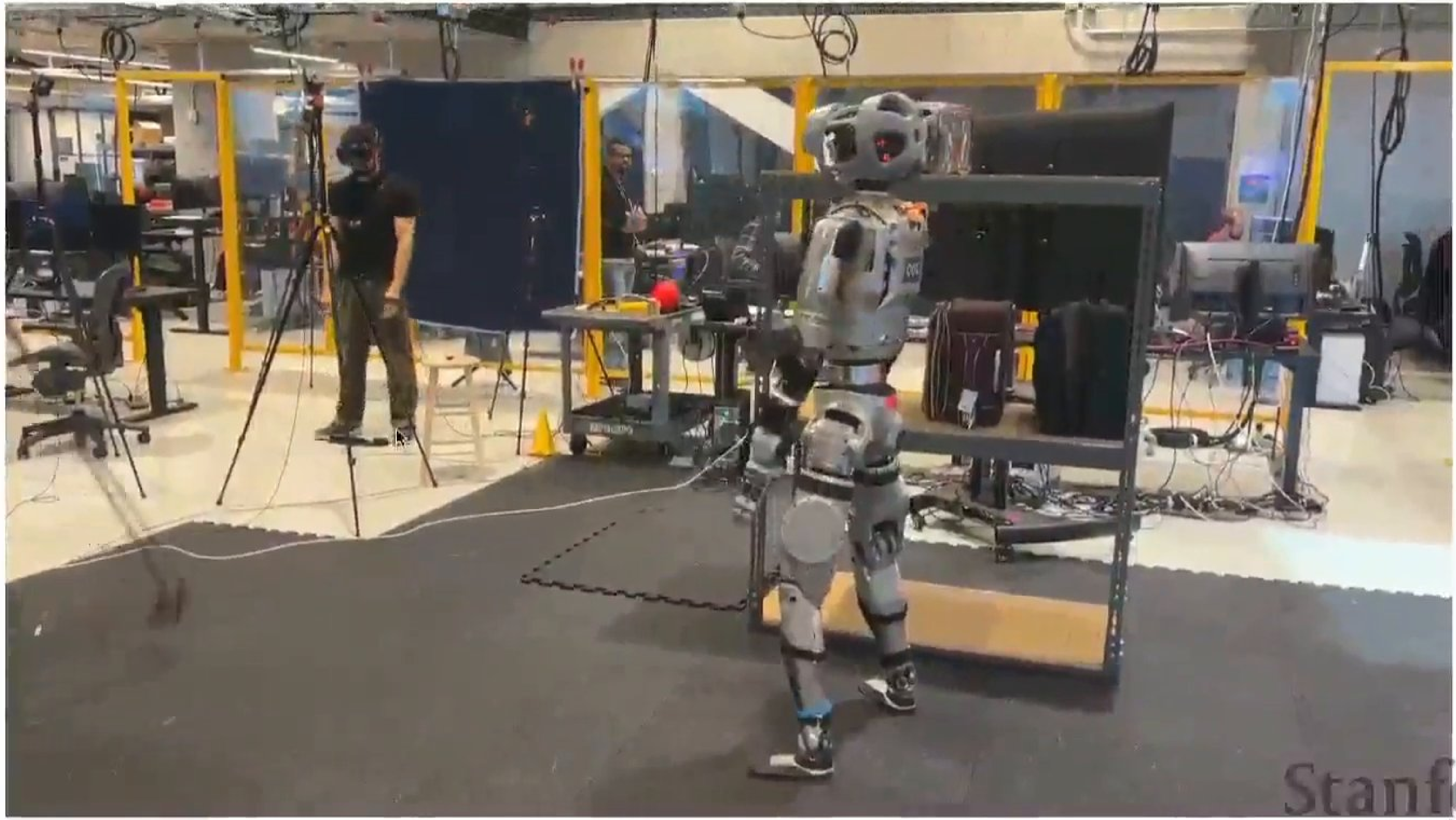
## Discussion (cont)

If you believe in RL post-training (e.g. DeepSeek-R1), then the essential role for pretraining is to get ~5-10% success rate on *any* task.

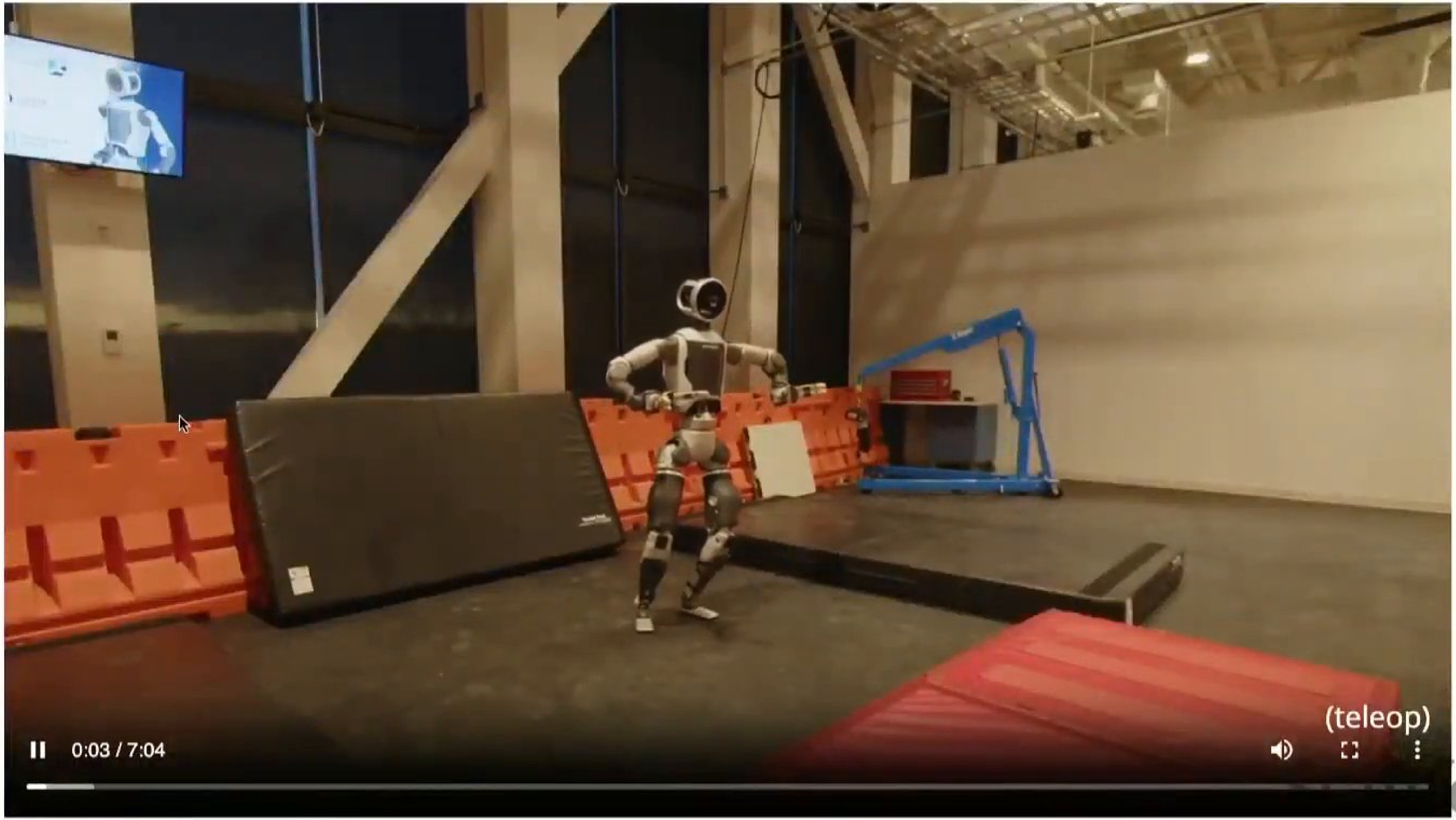
BostonDynamics + TOYOTA RESEARCH INSTITUTE



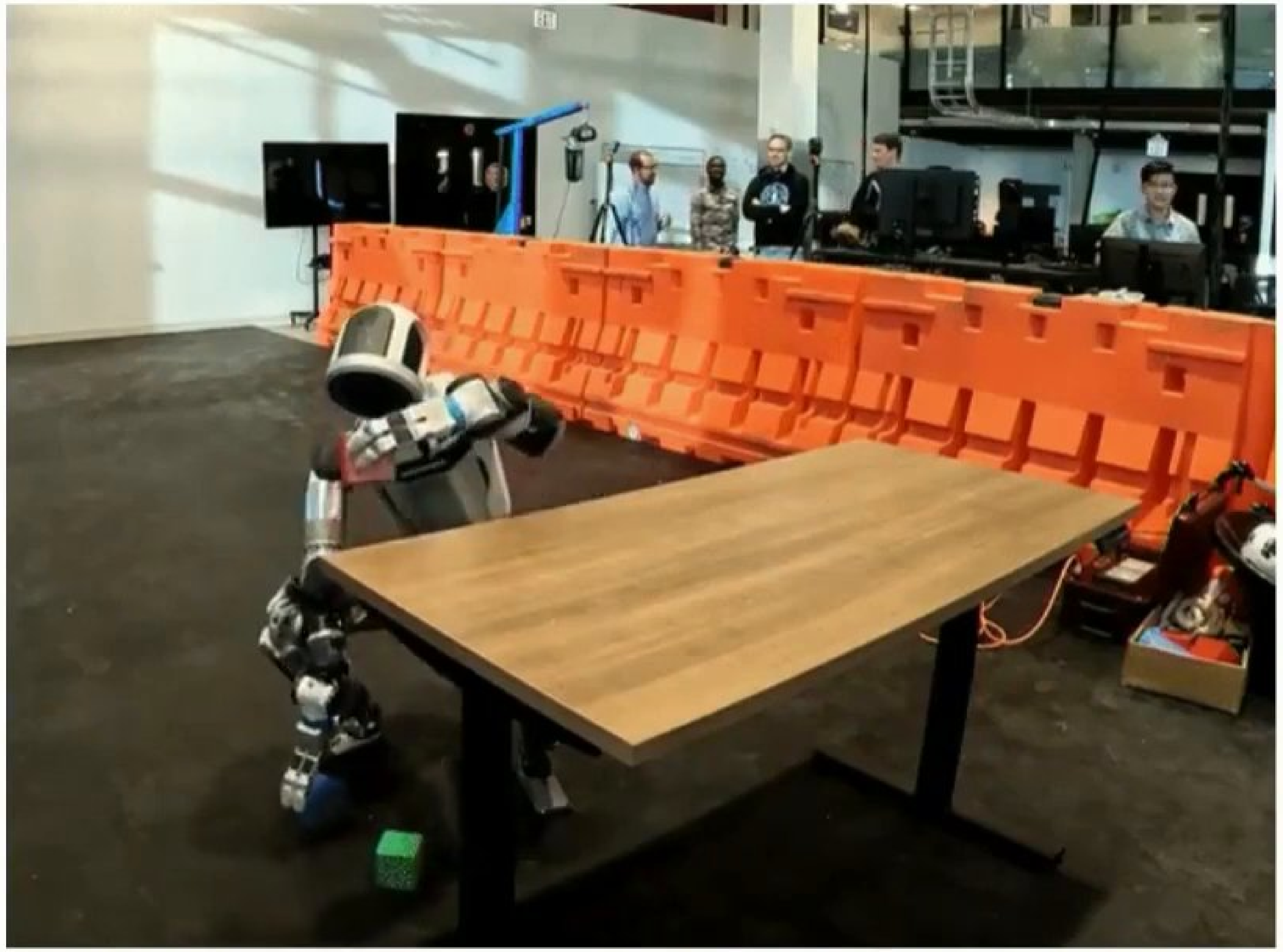
Stanford



Stanford


(teleop)
0:03 / 7:04

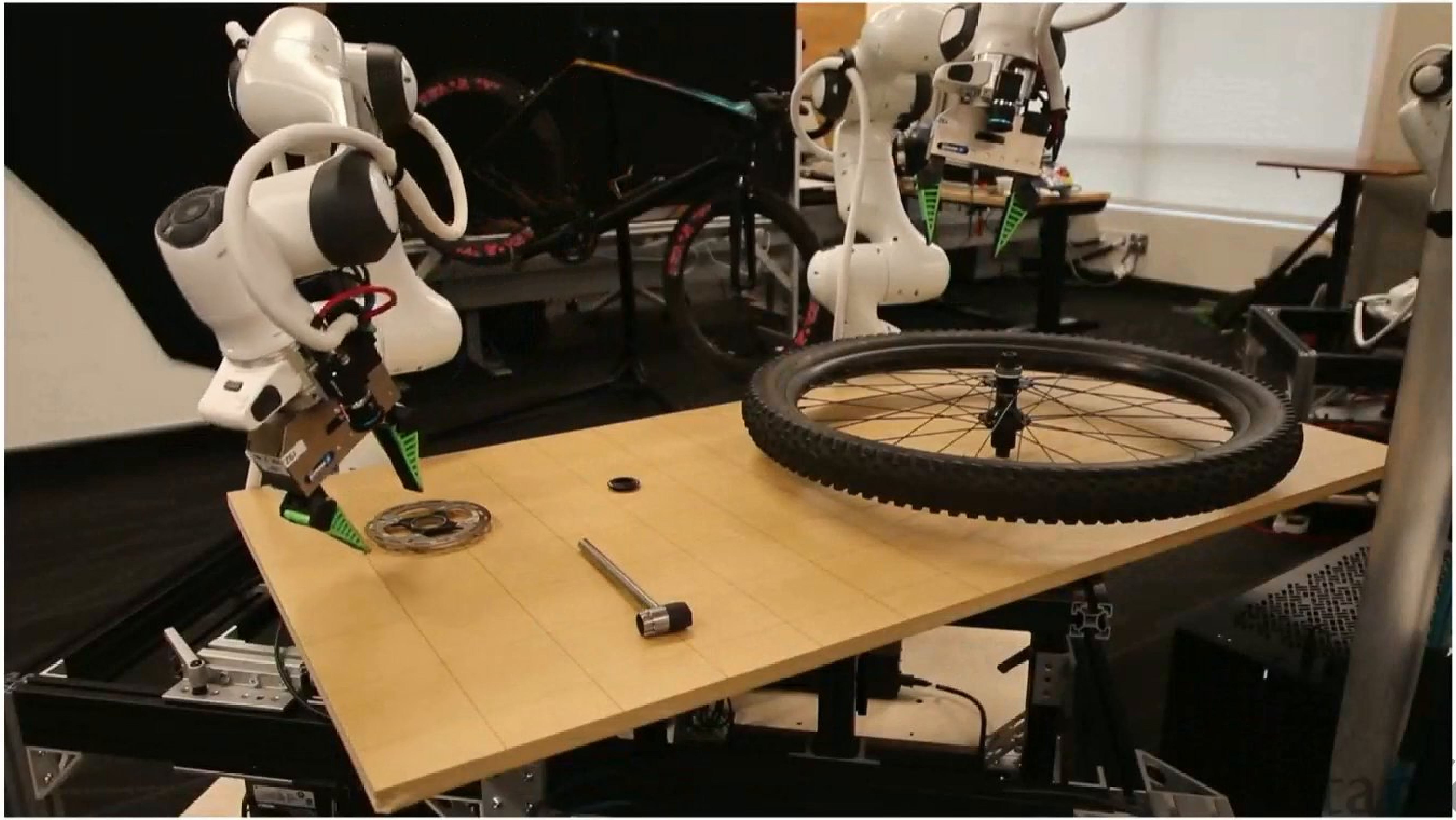

(teleop)

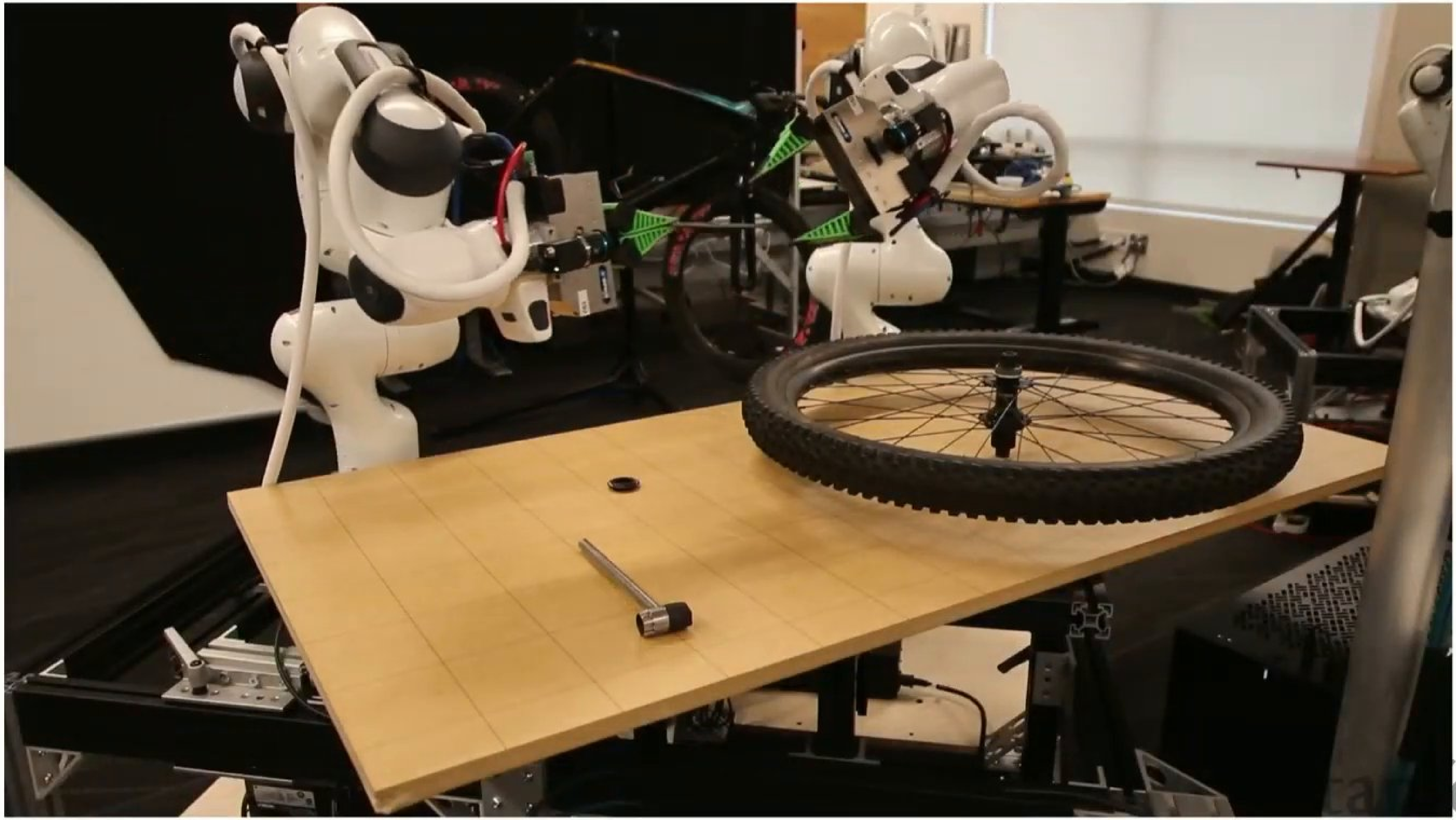


(teleop) Stanford

# Questions?

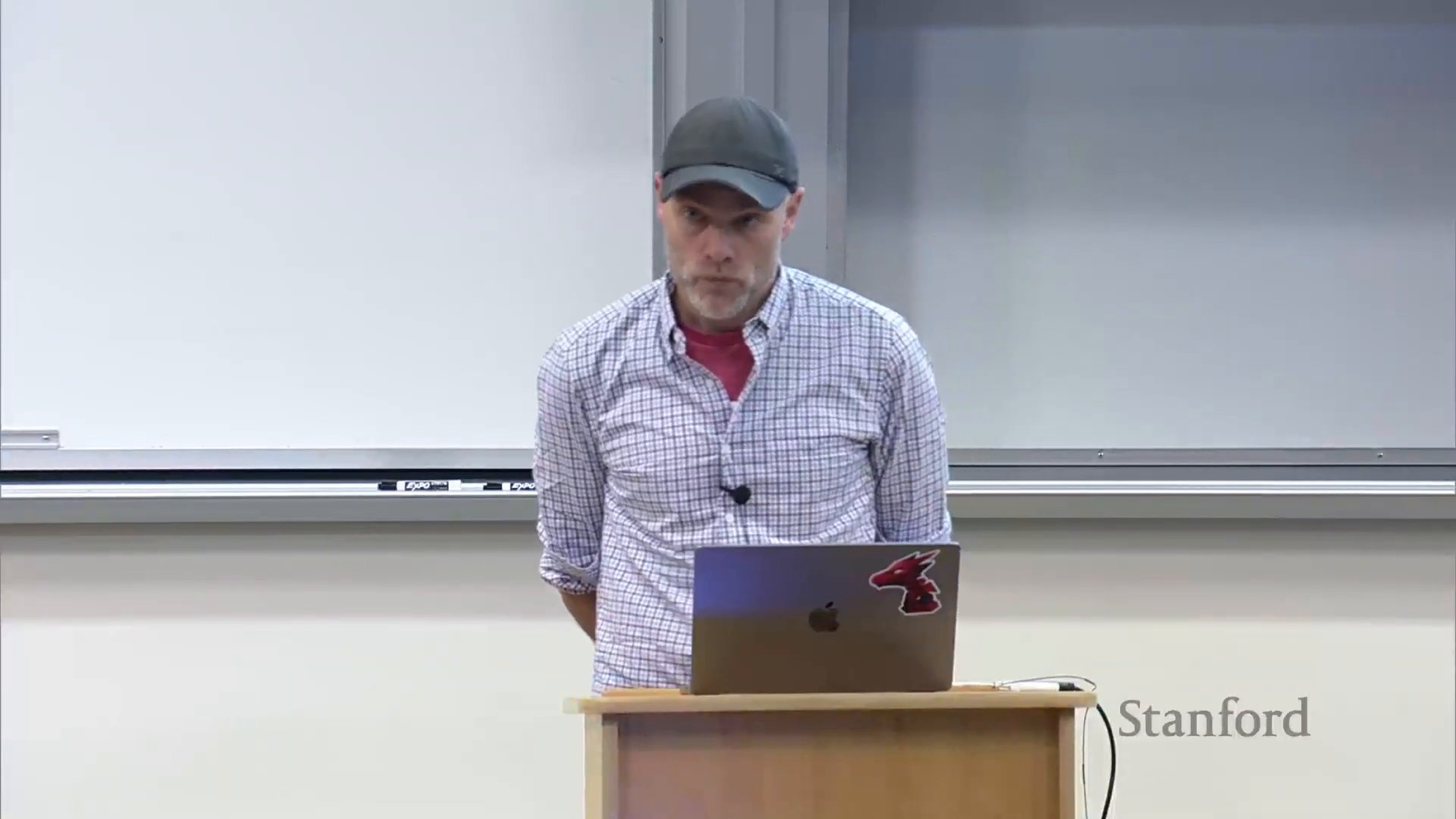
Stanford






Stanford


Stanford


Stanford


Stanford


Stanford

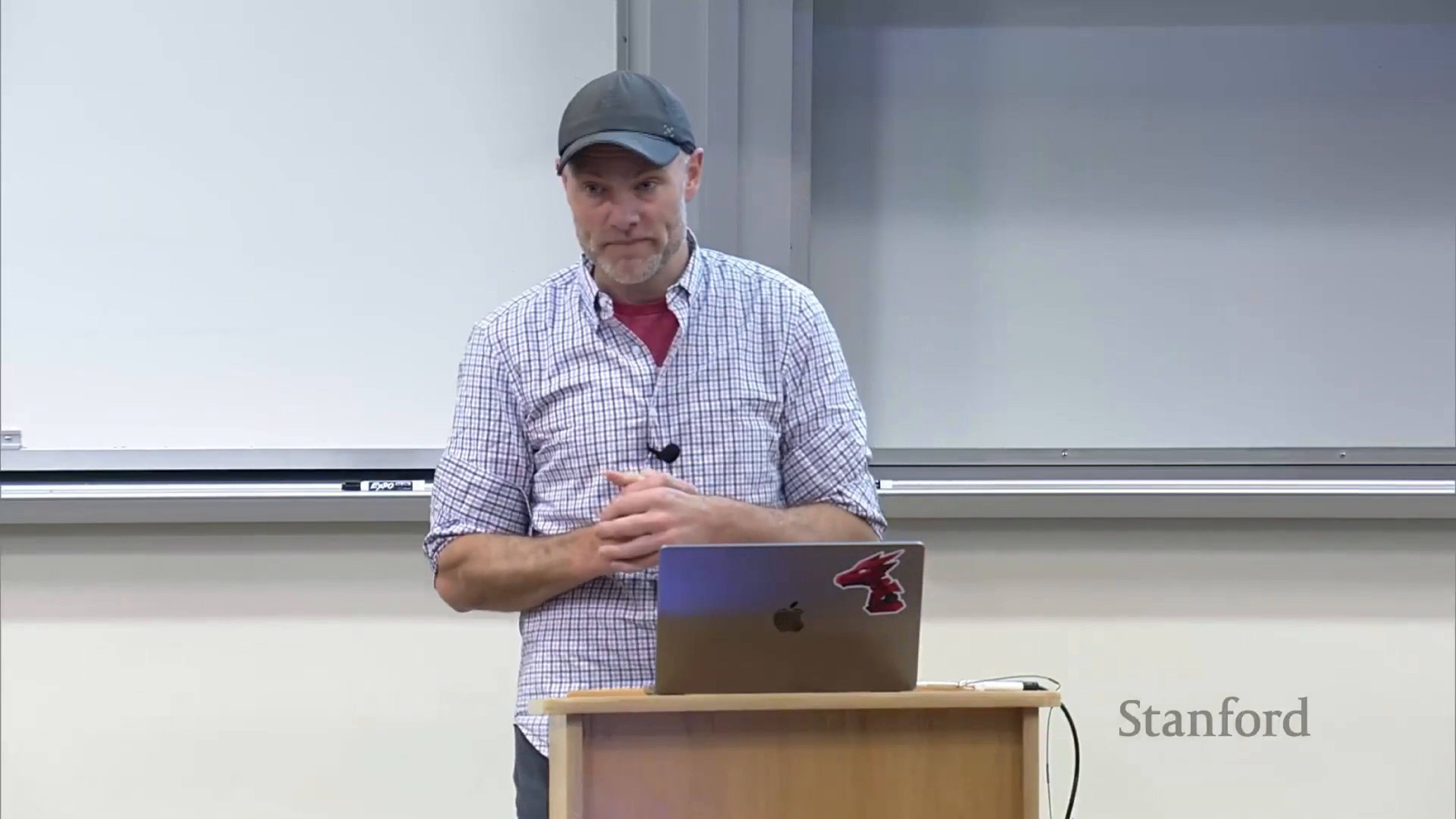

Stanford

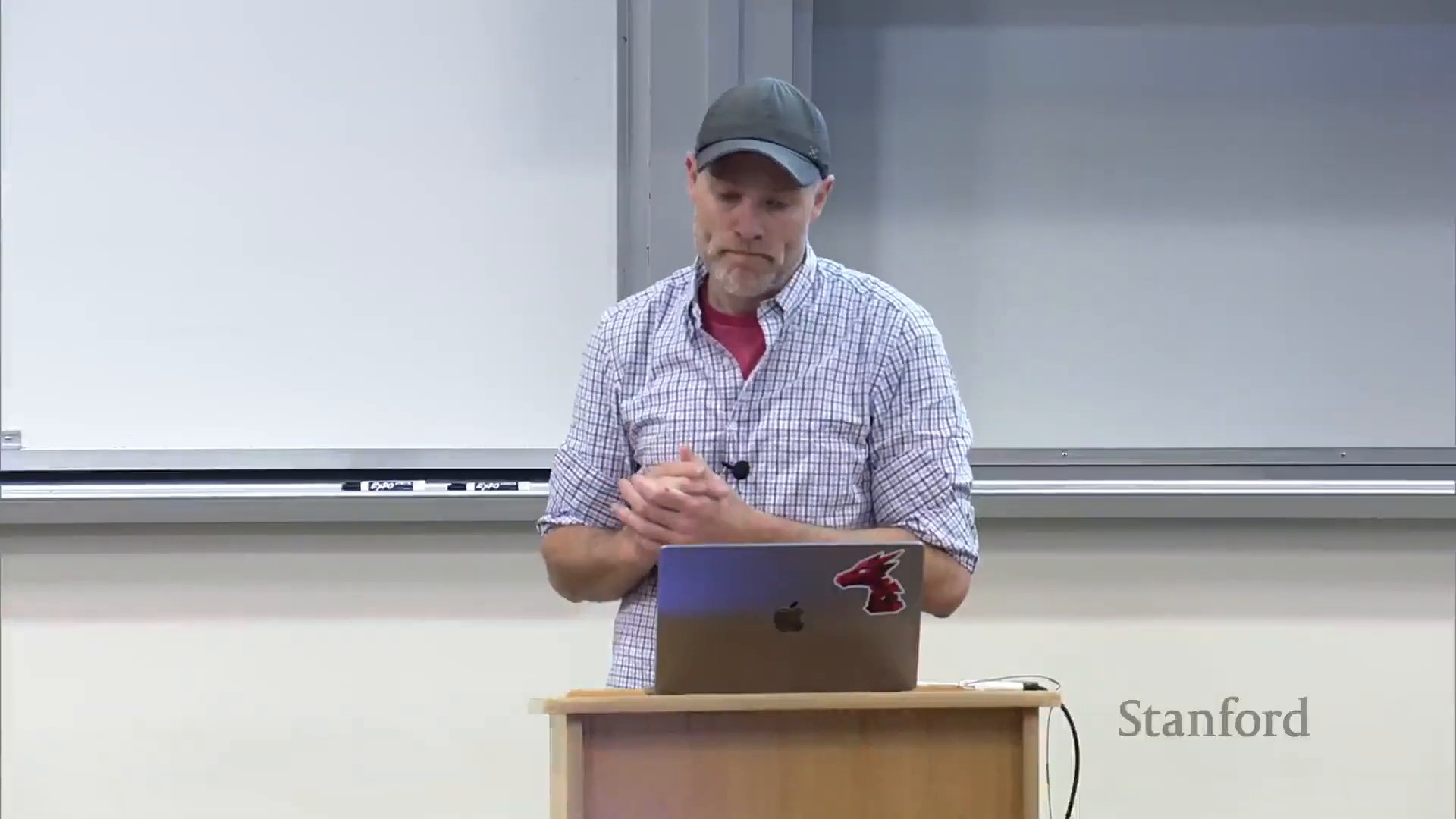

Stanford


Stanford


Stanford

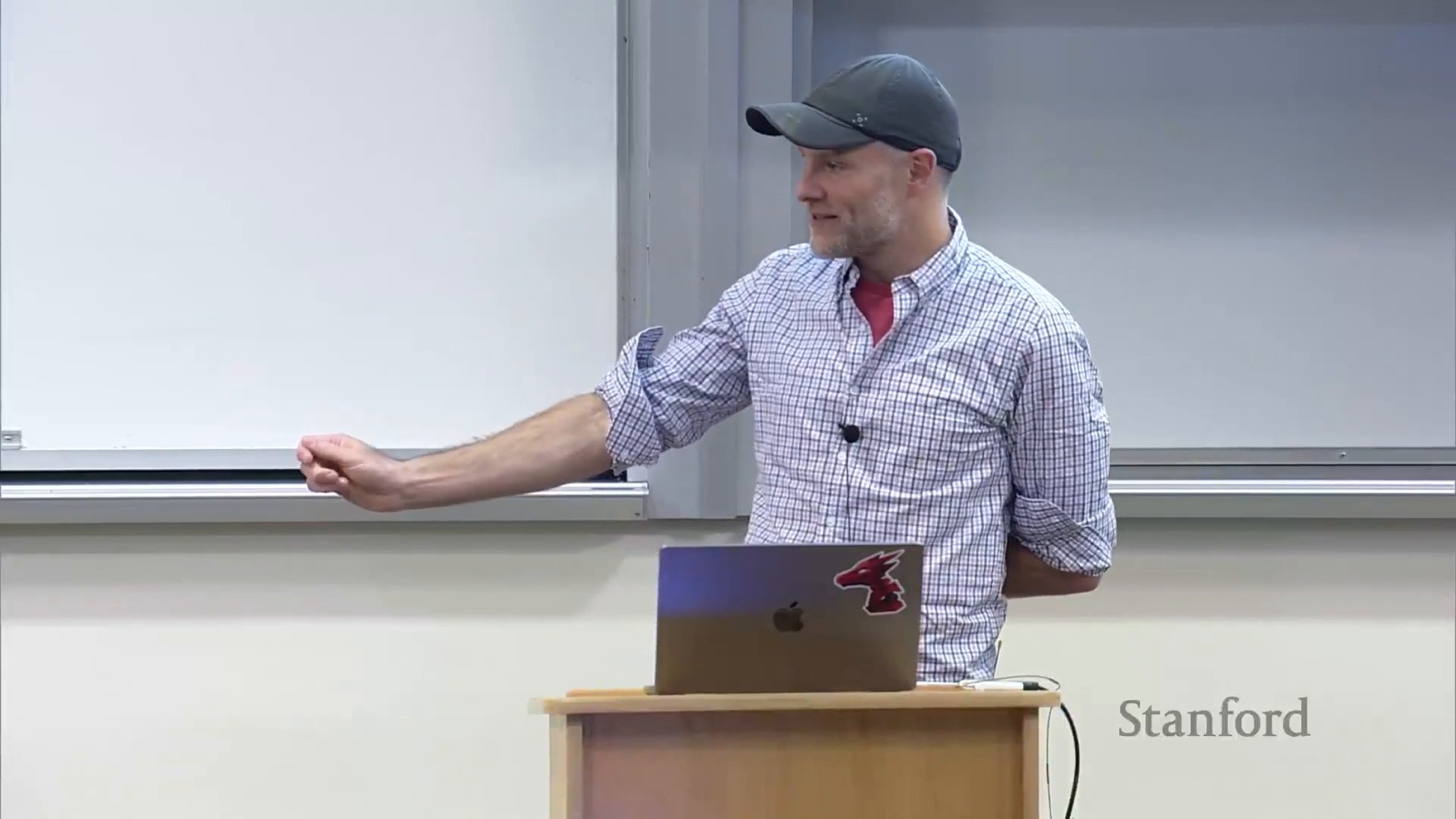

Stanford

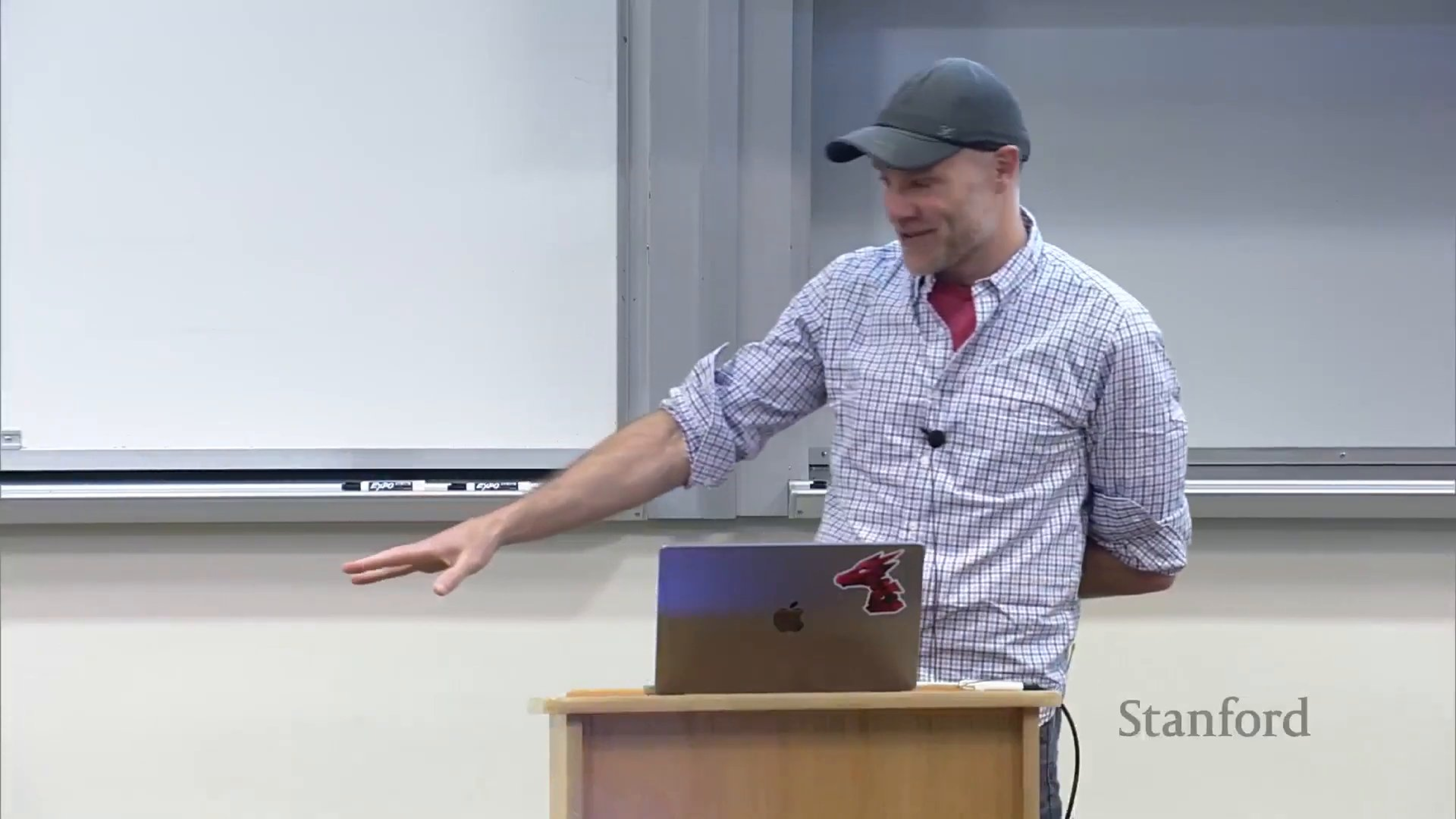

Stanford

Stanford

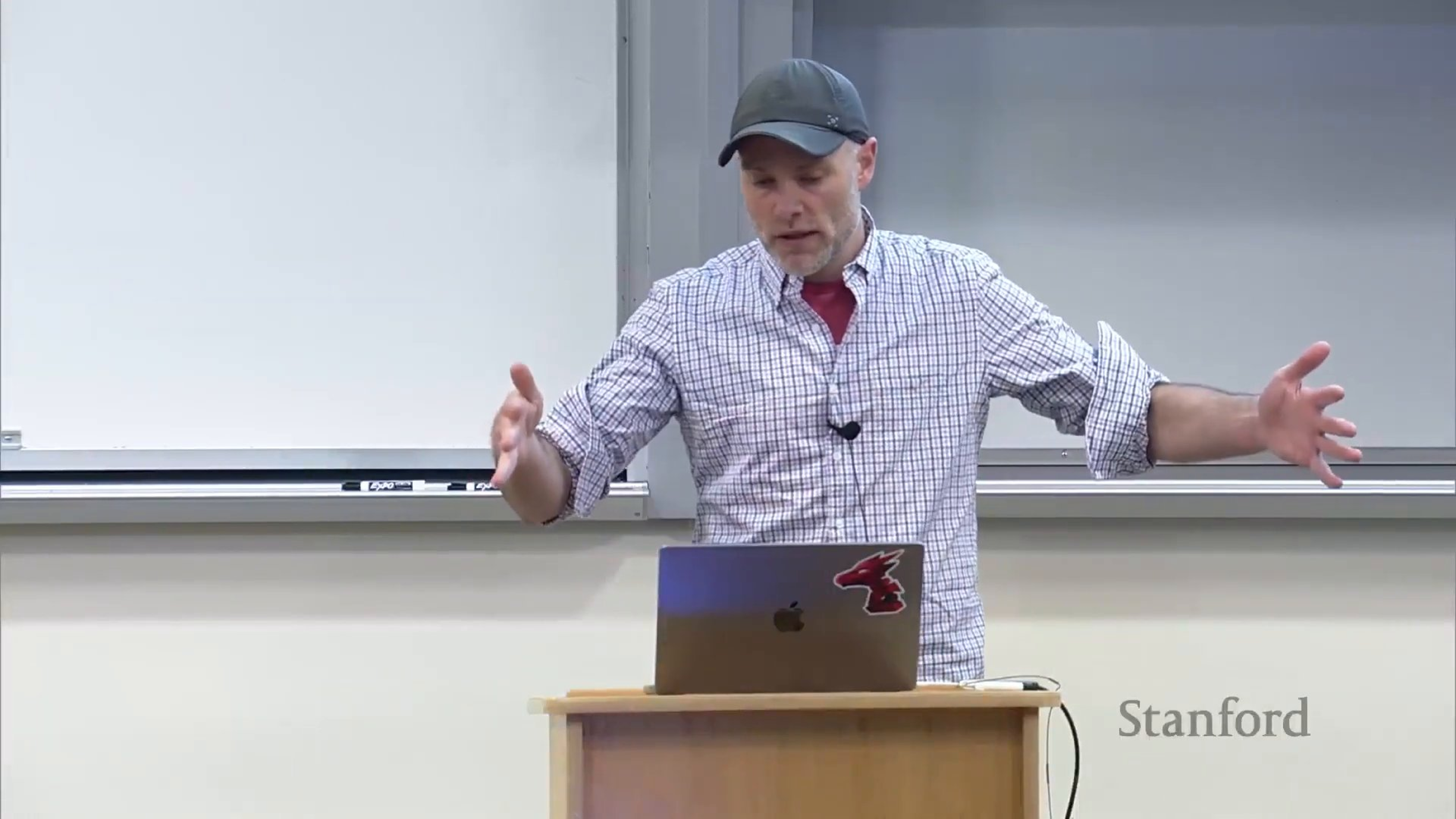

Stanford

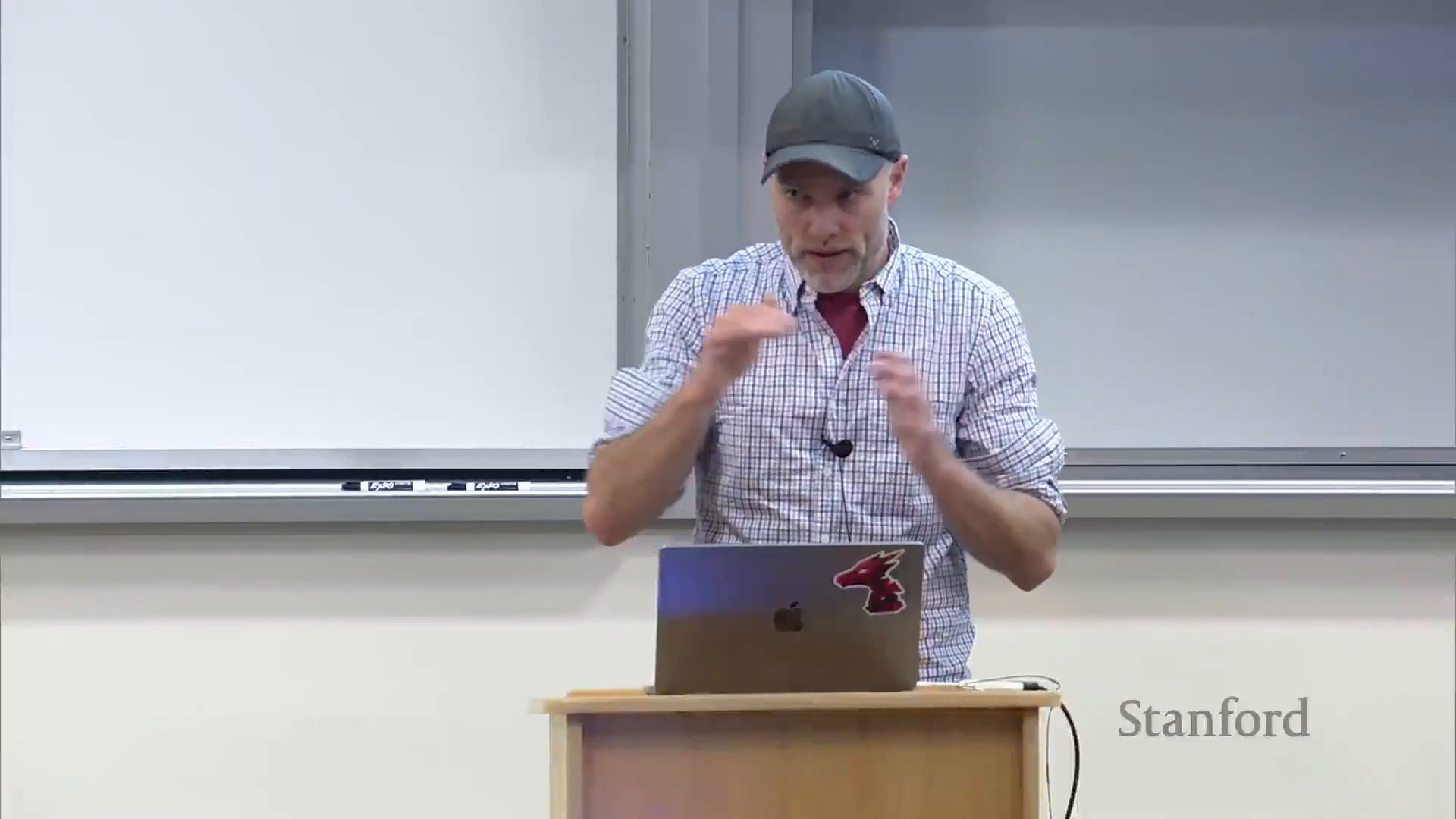

Stanford


Stanford


Stanford


Stanford

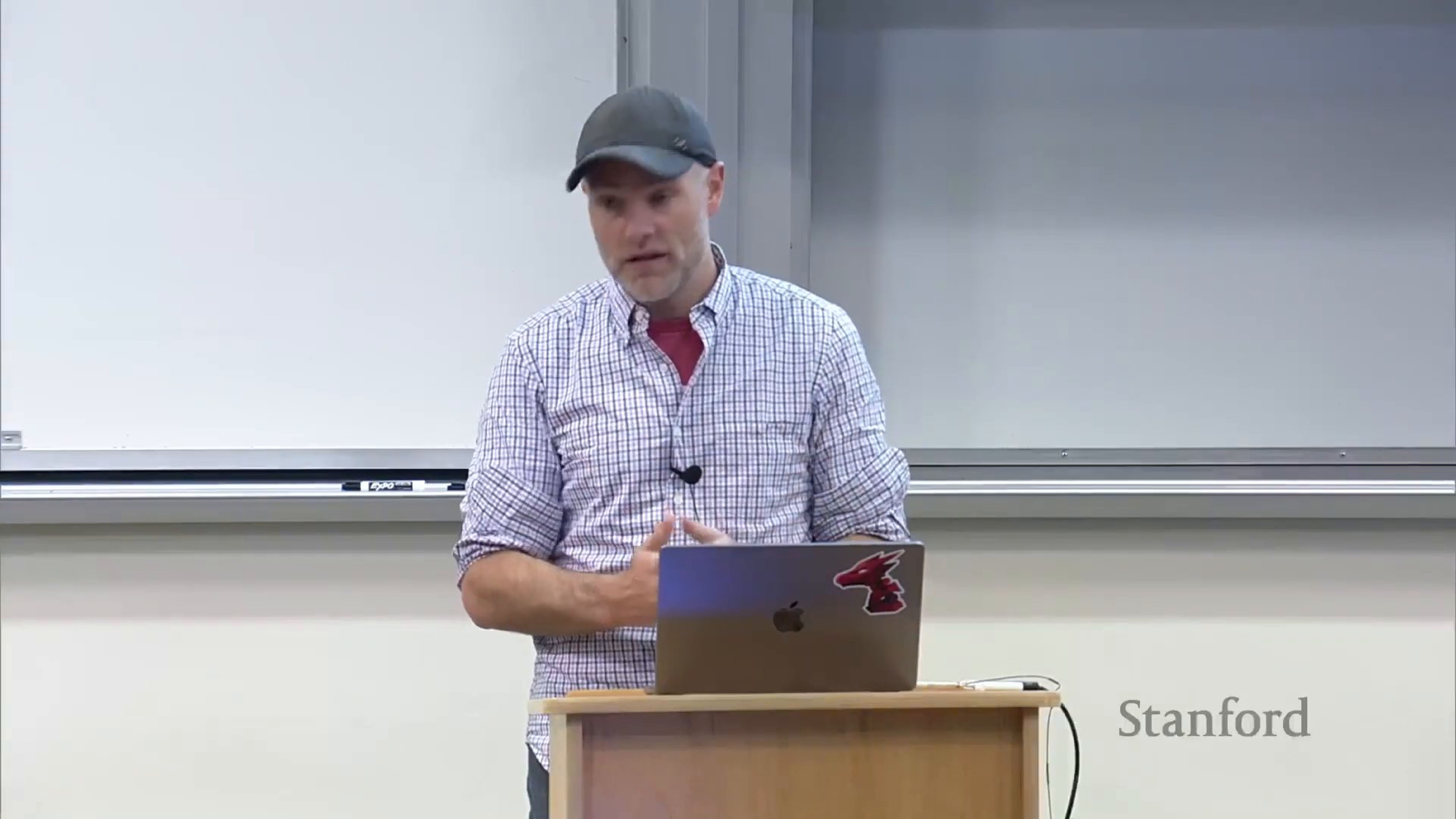

Stanford


Stanford

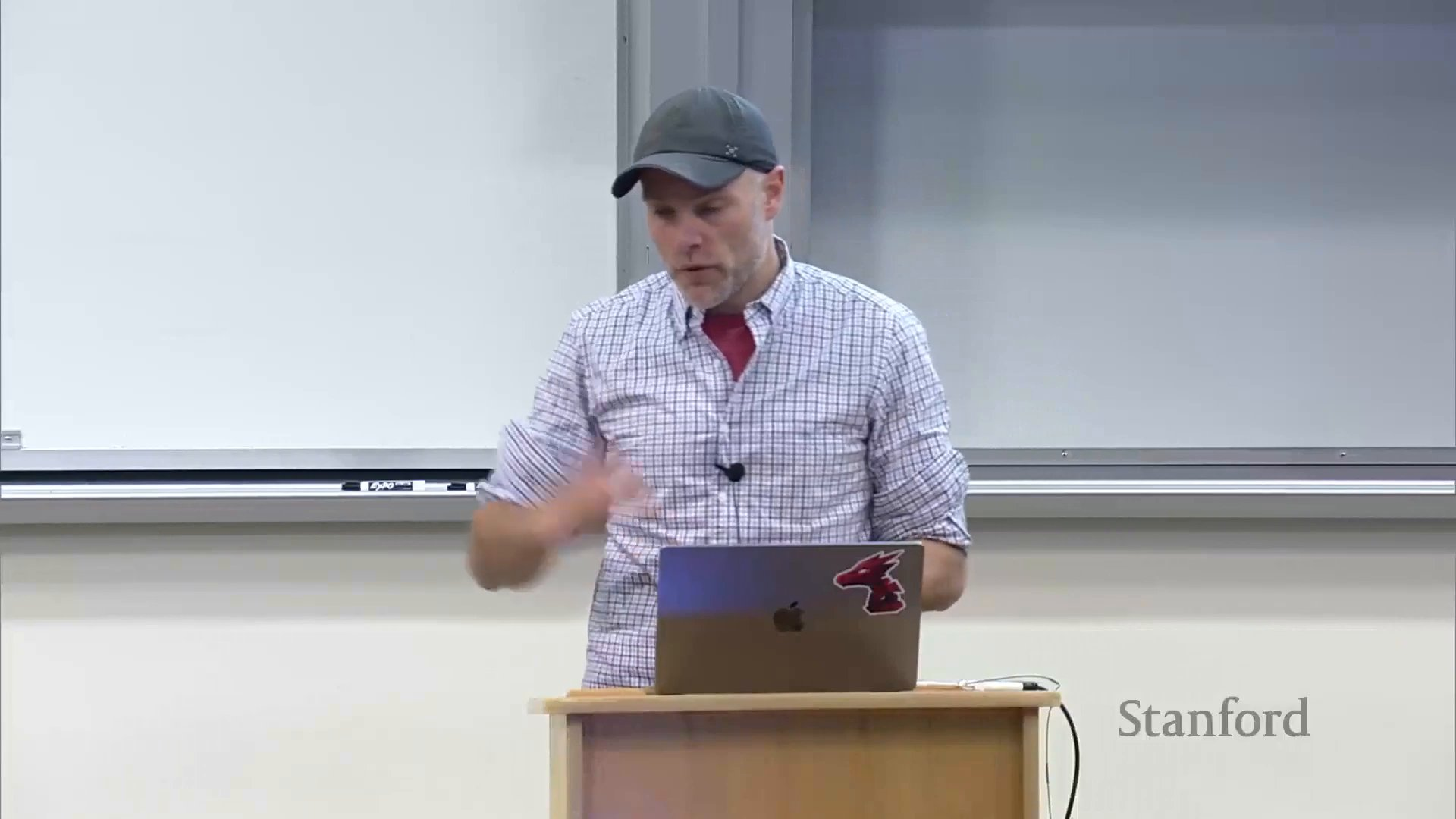

Stanford

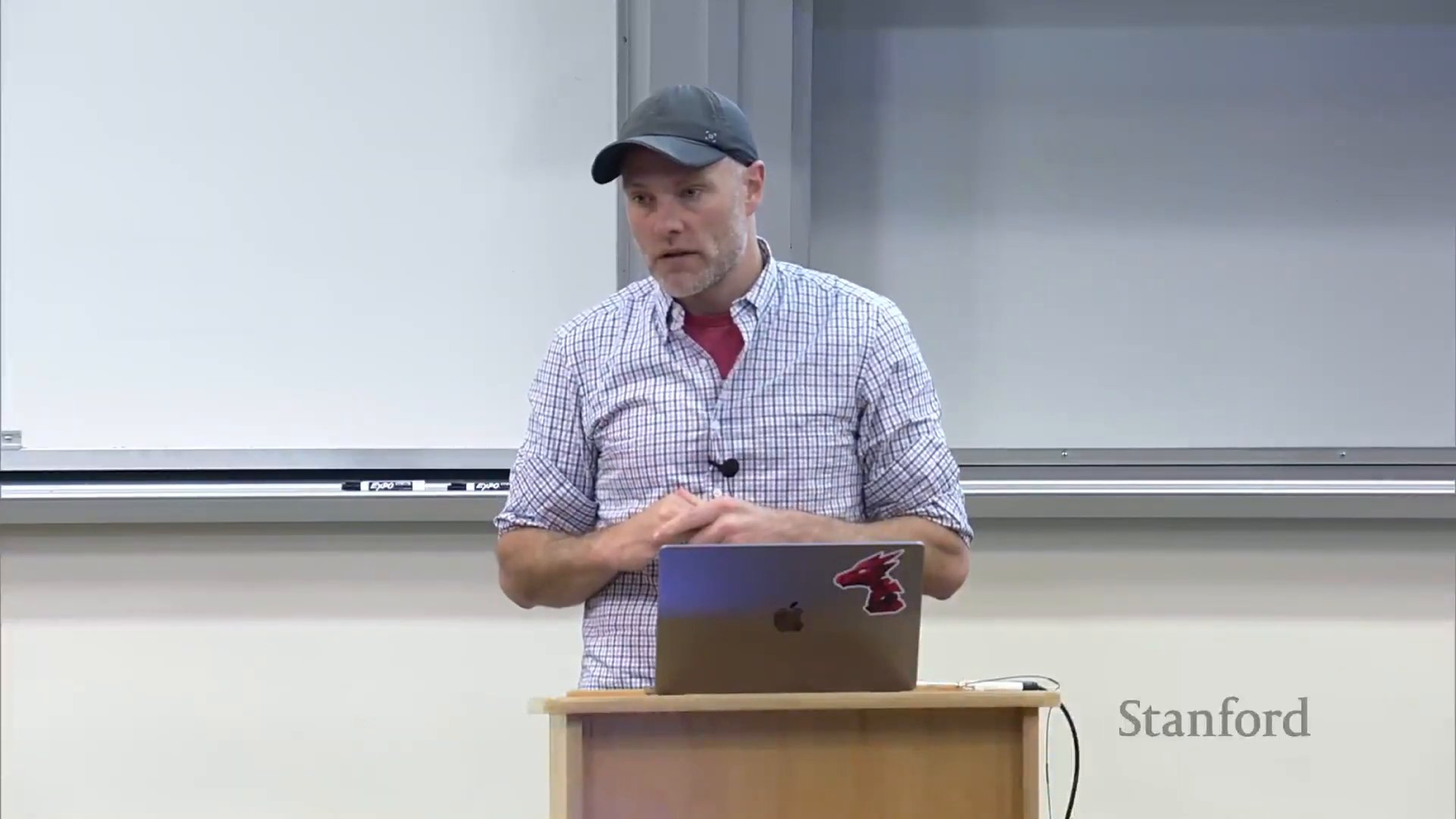

Stanford


Stanford


Stanford

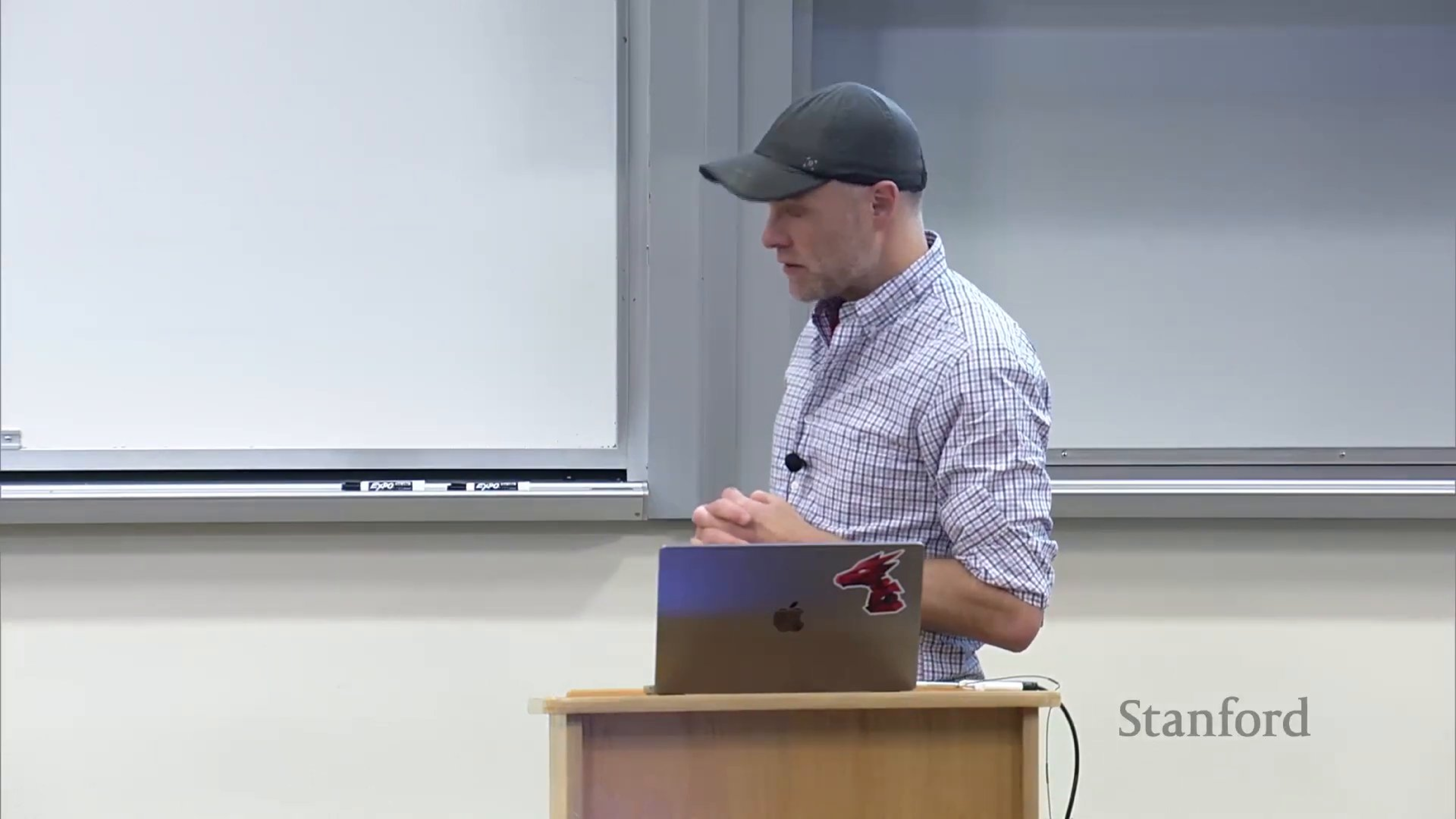

Stanford


Stanford

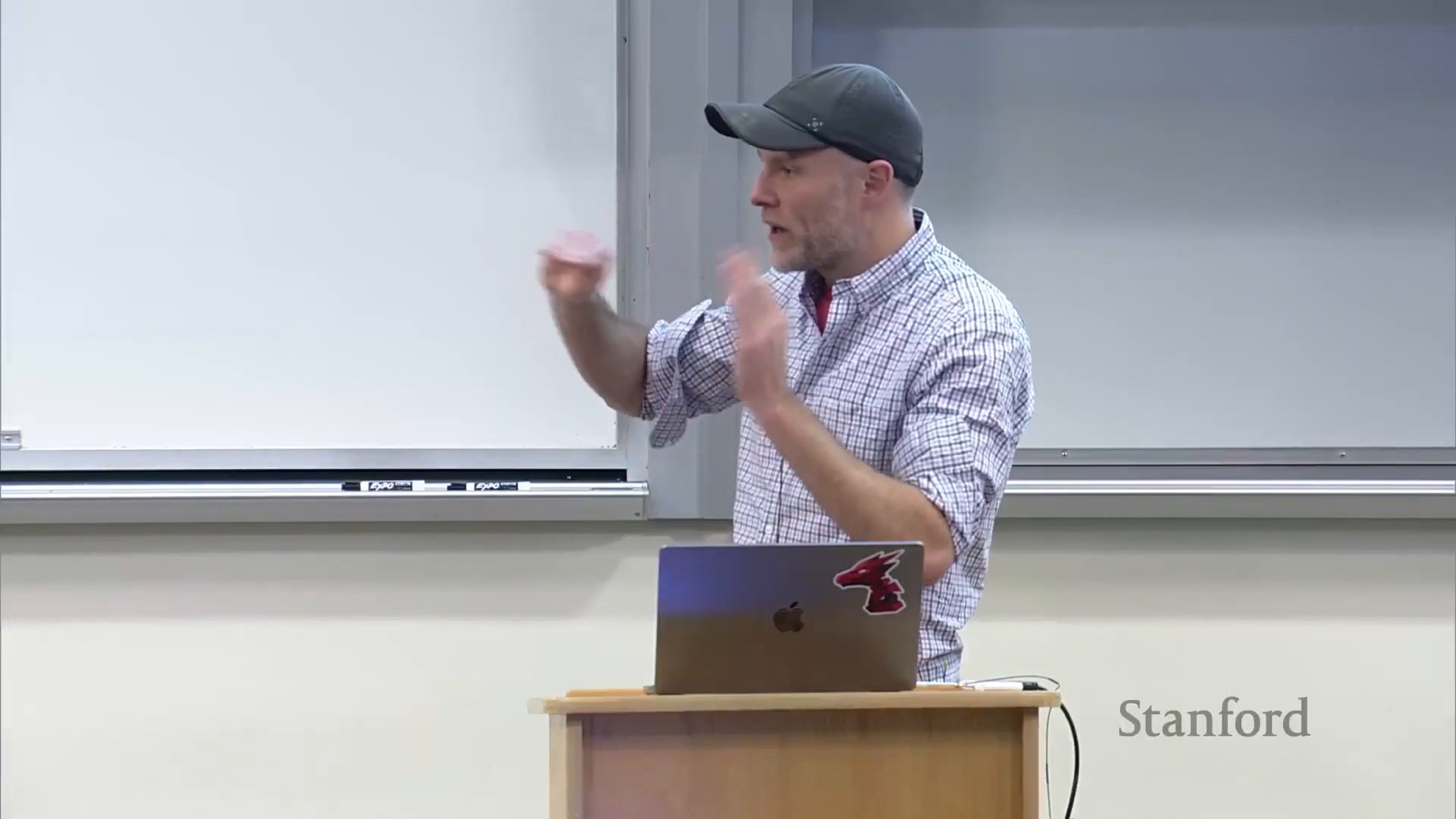

Stanford


Stanford

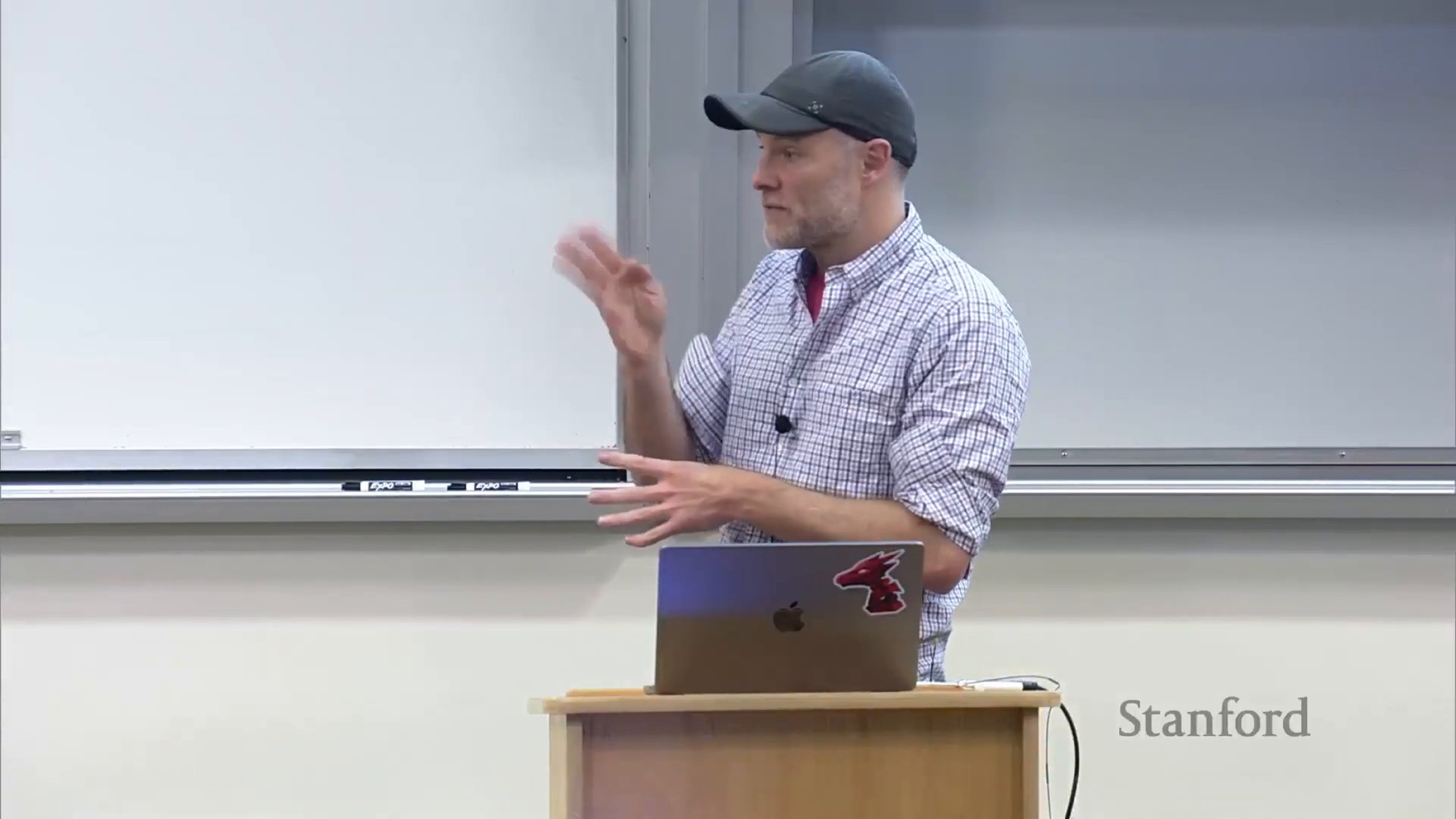

Stanford

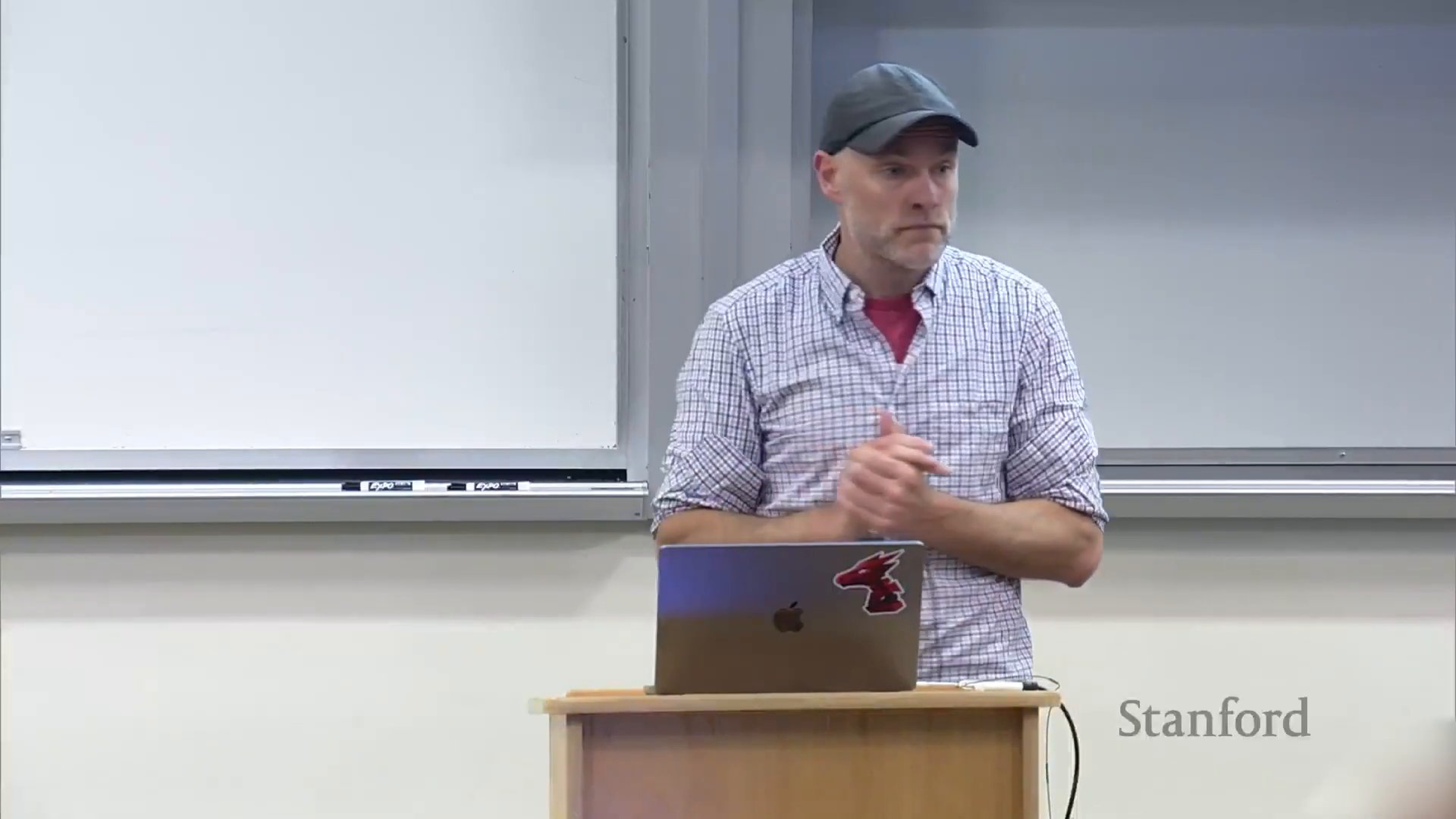

Stanford

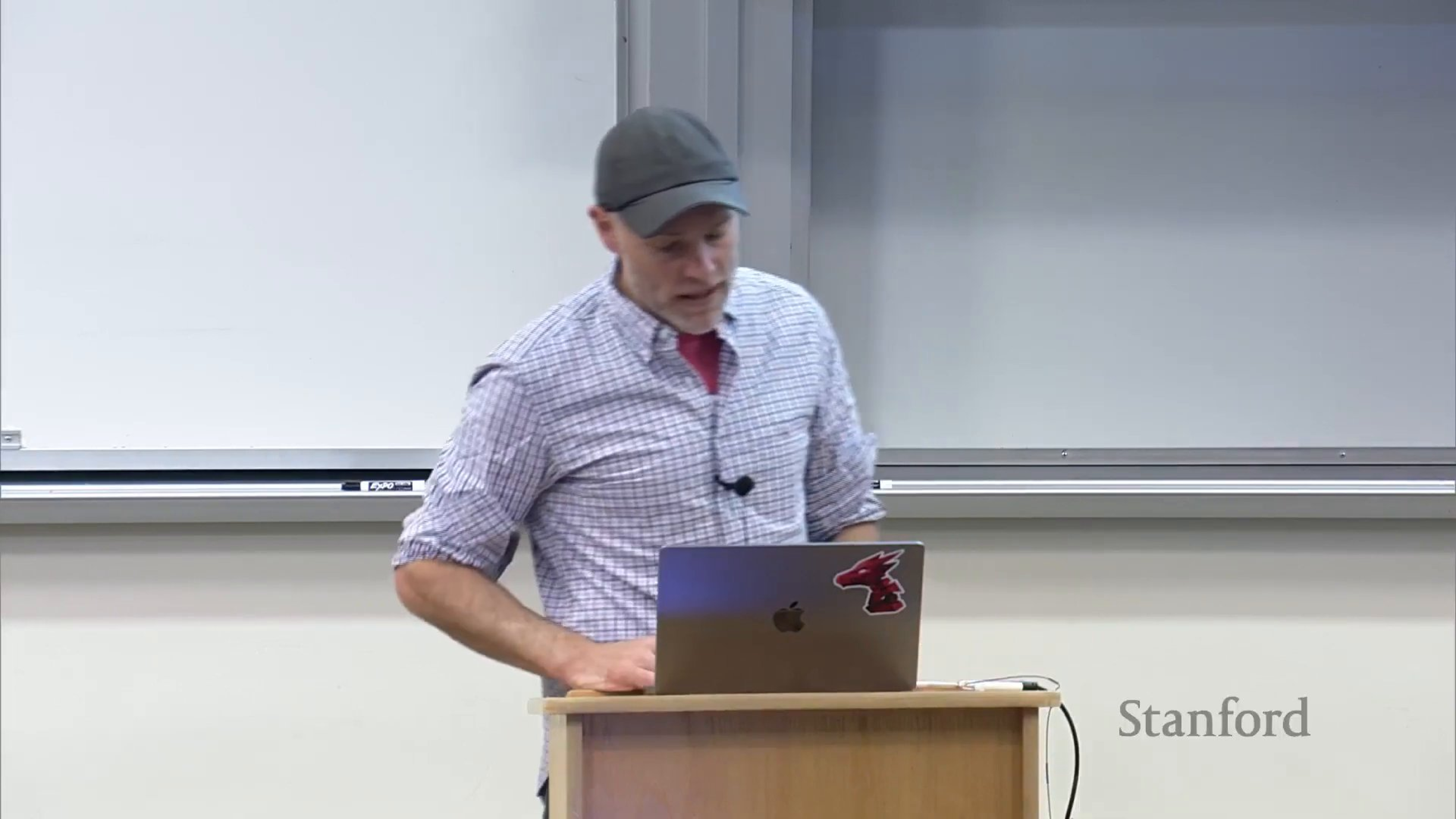

Stanford

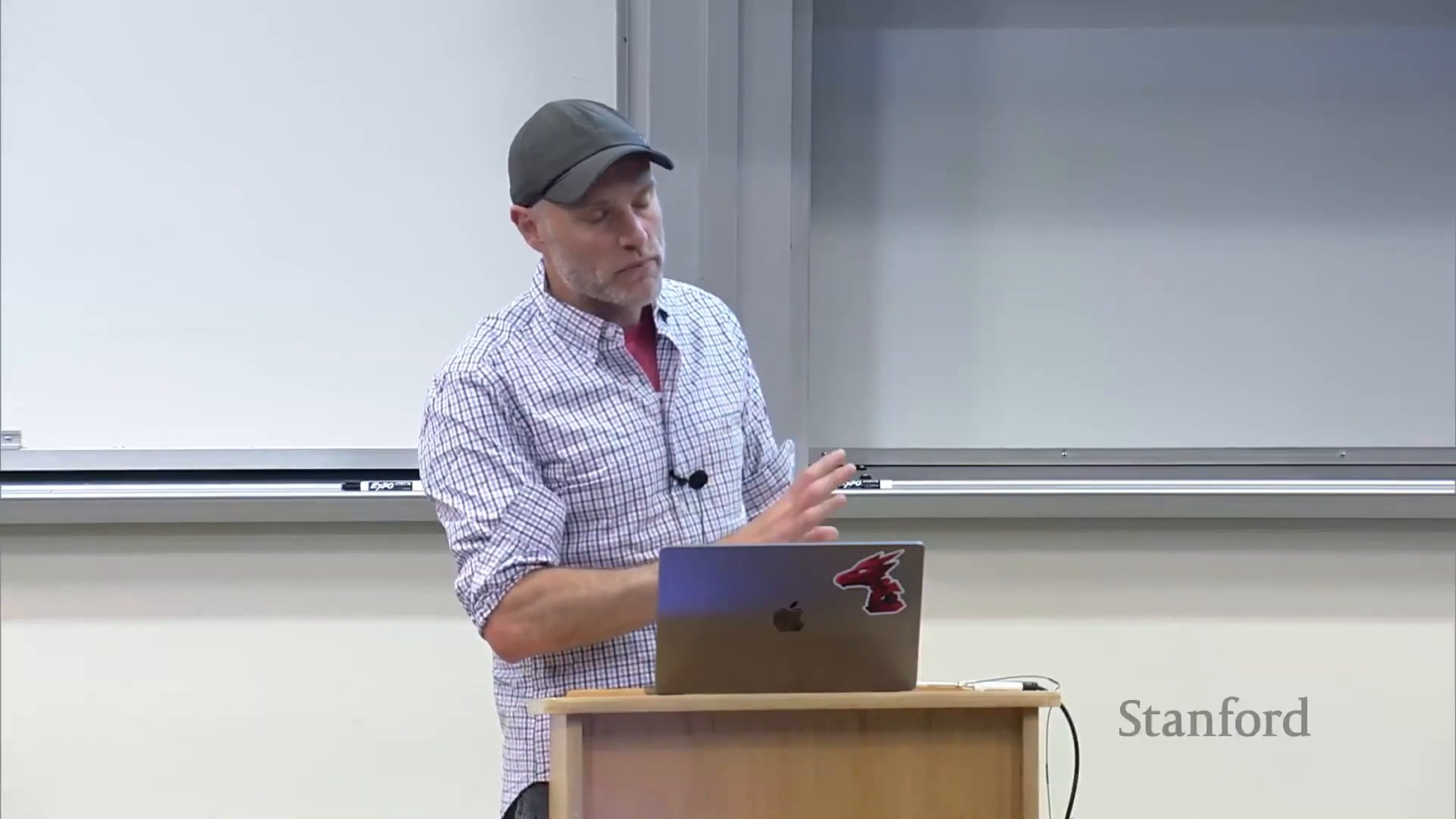

Stanford

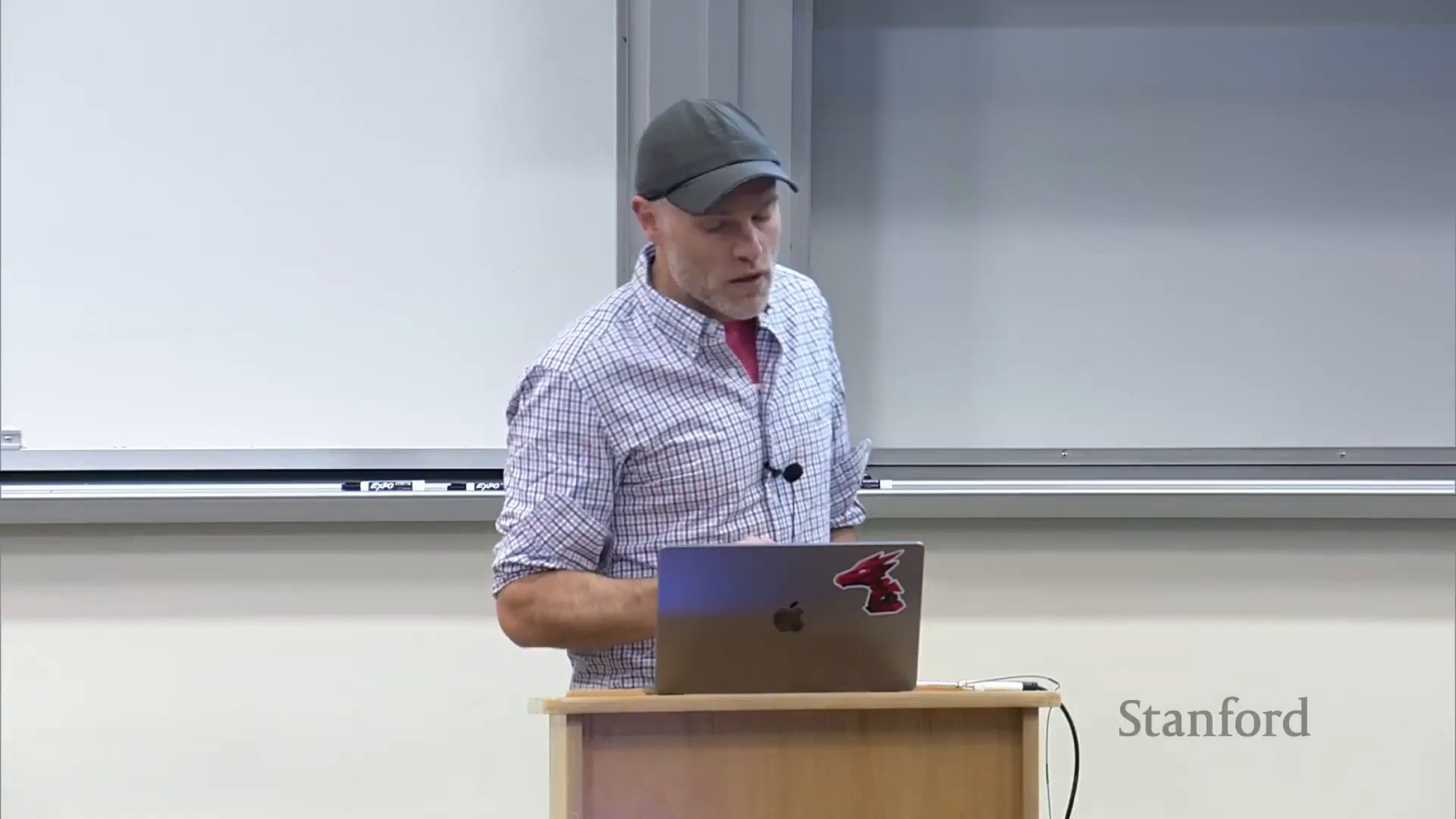

Stanford

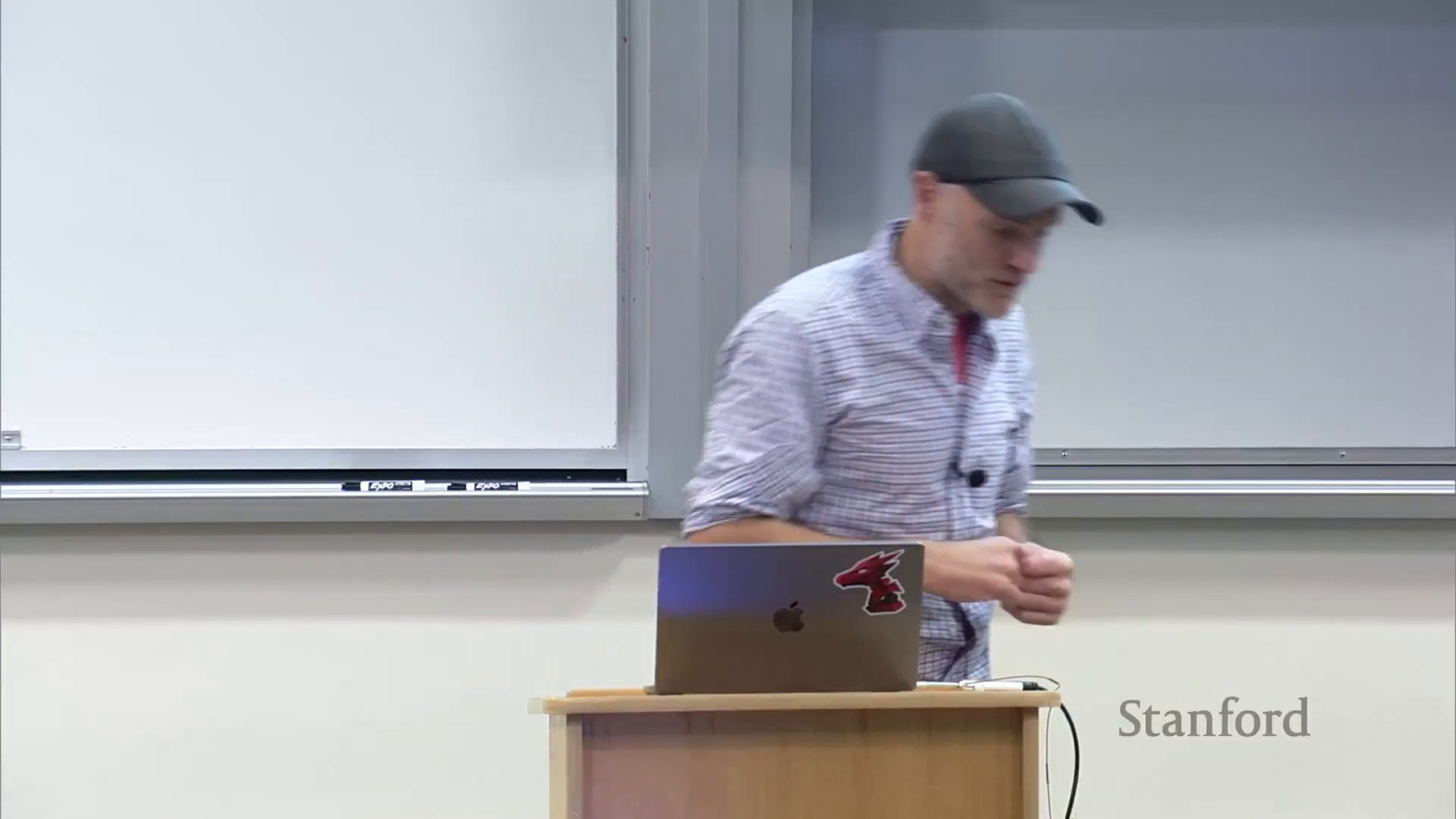

Stanford


Stanford

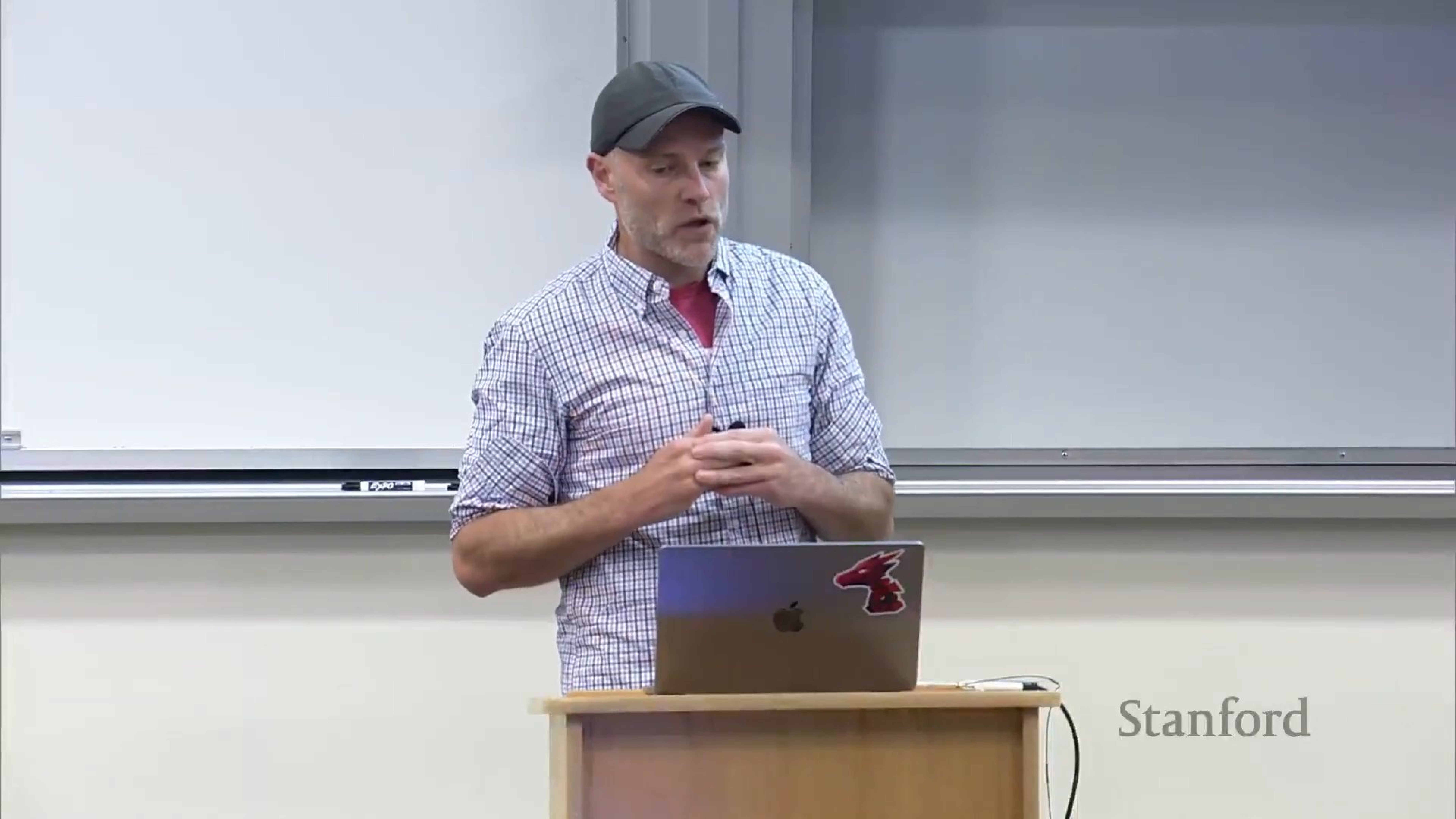

Stanford

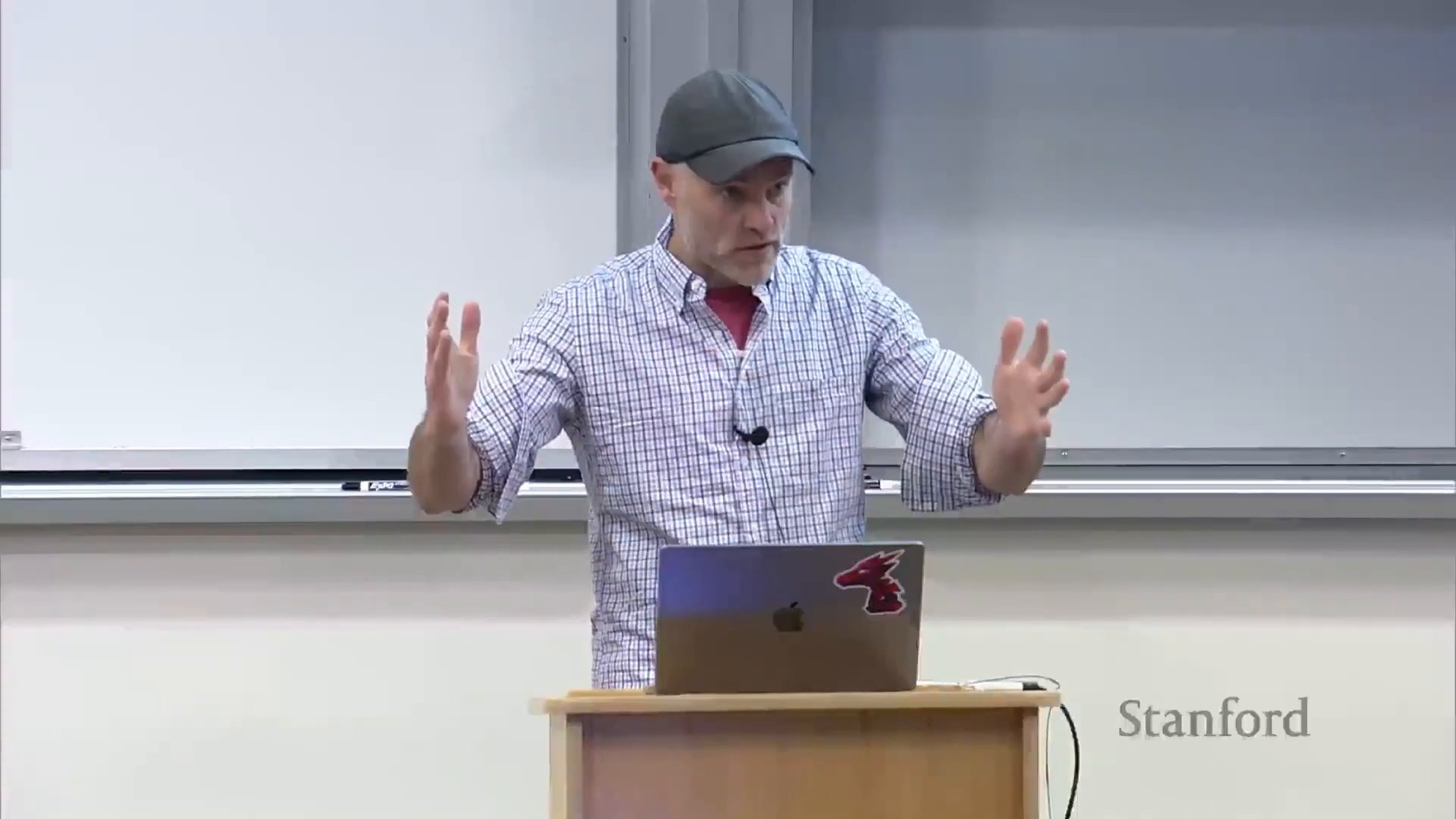

Stanford

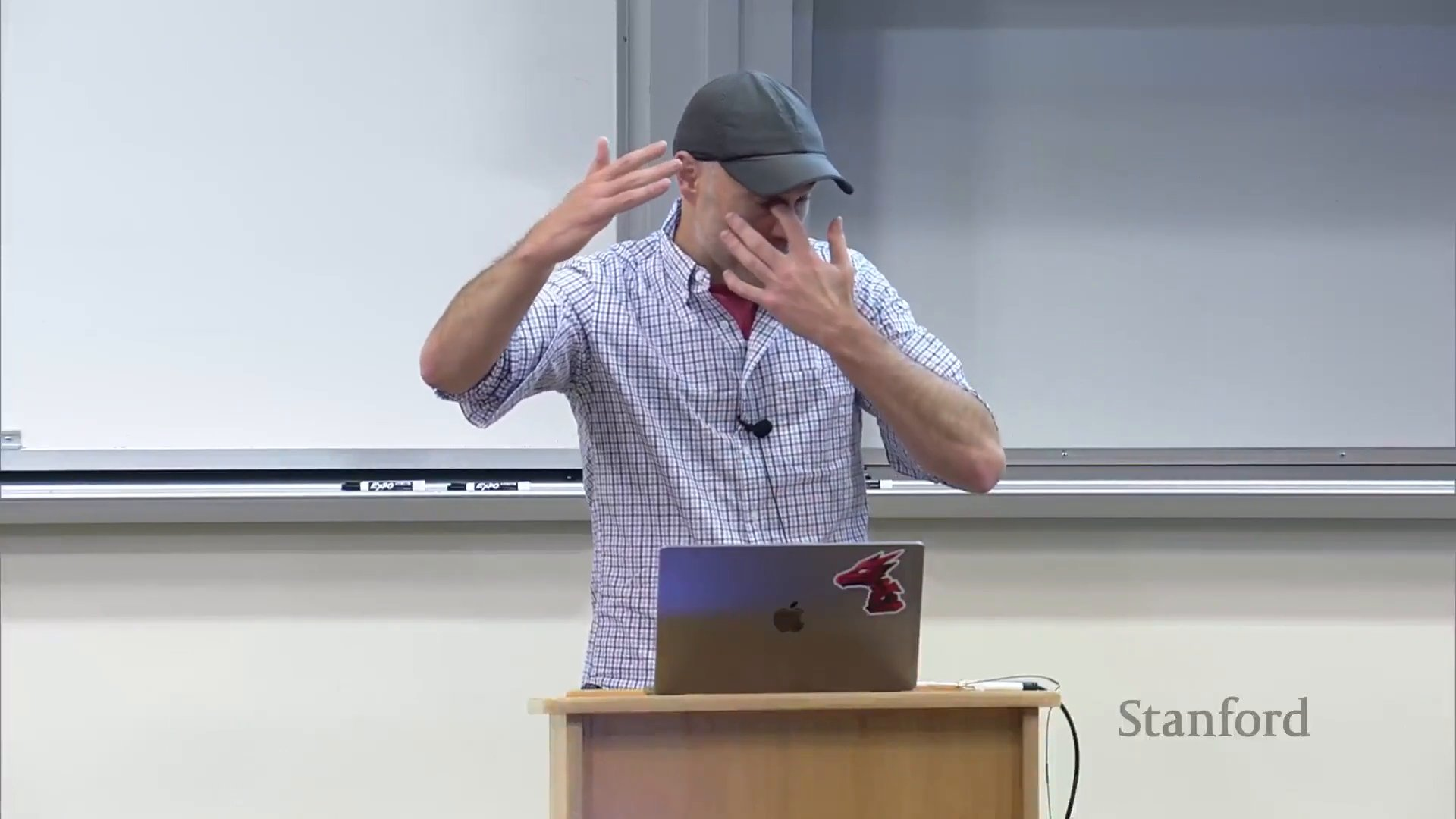

Stanford


Stanford


Stanford


Stanford


Stanford


Stanford

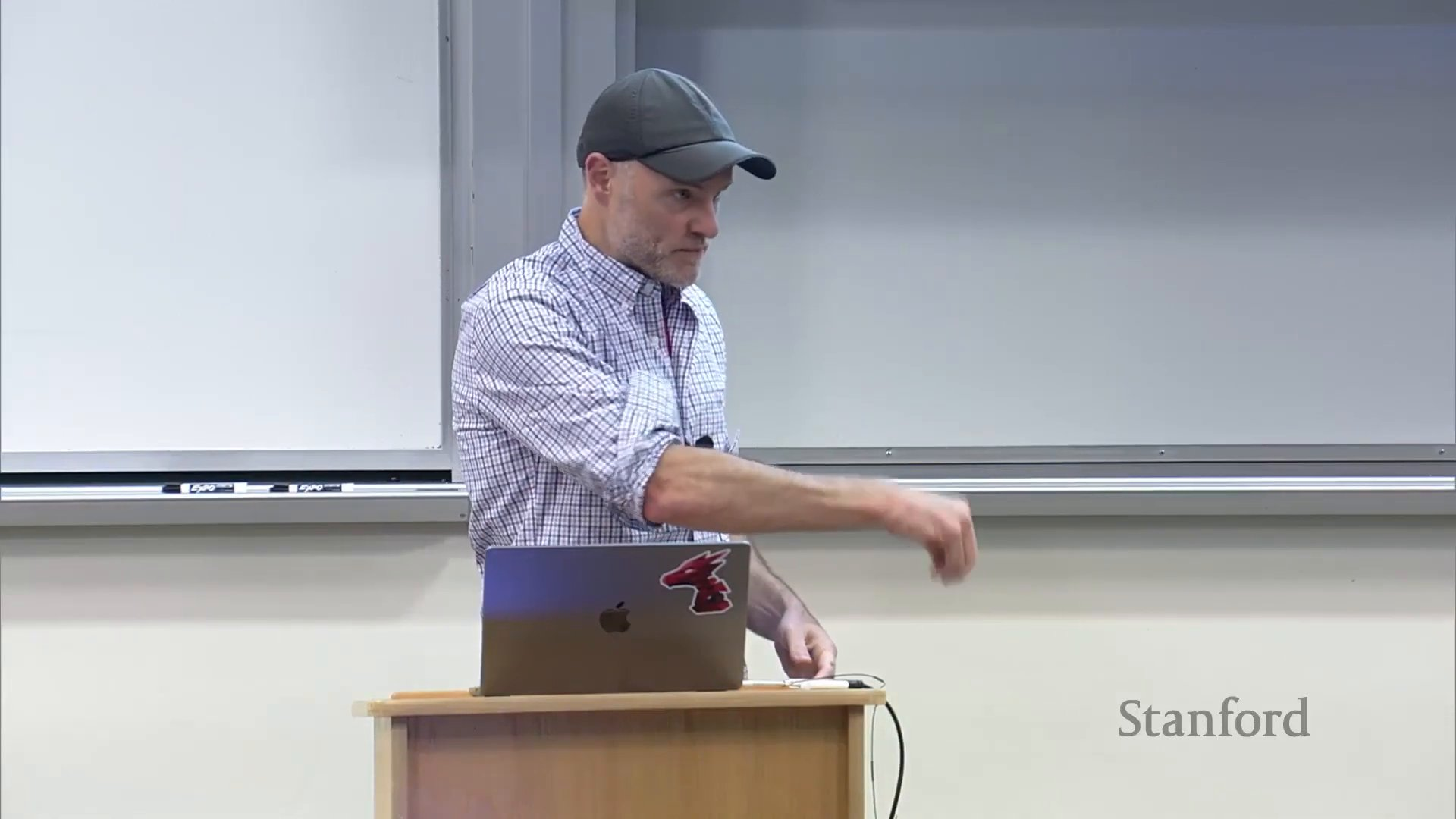

Stanford

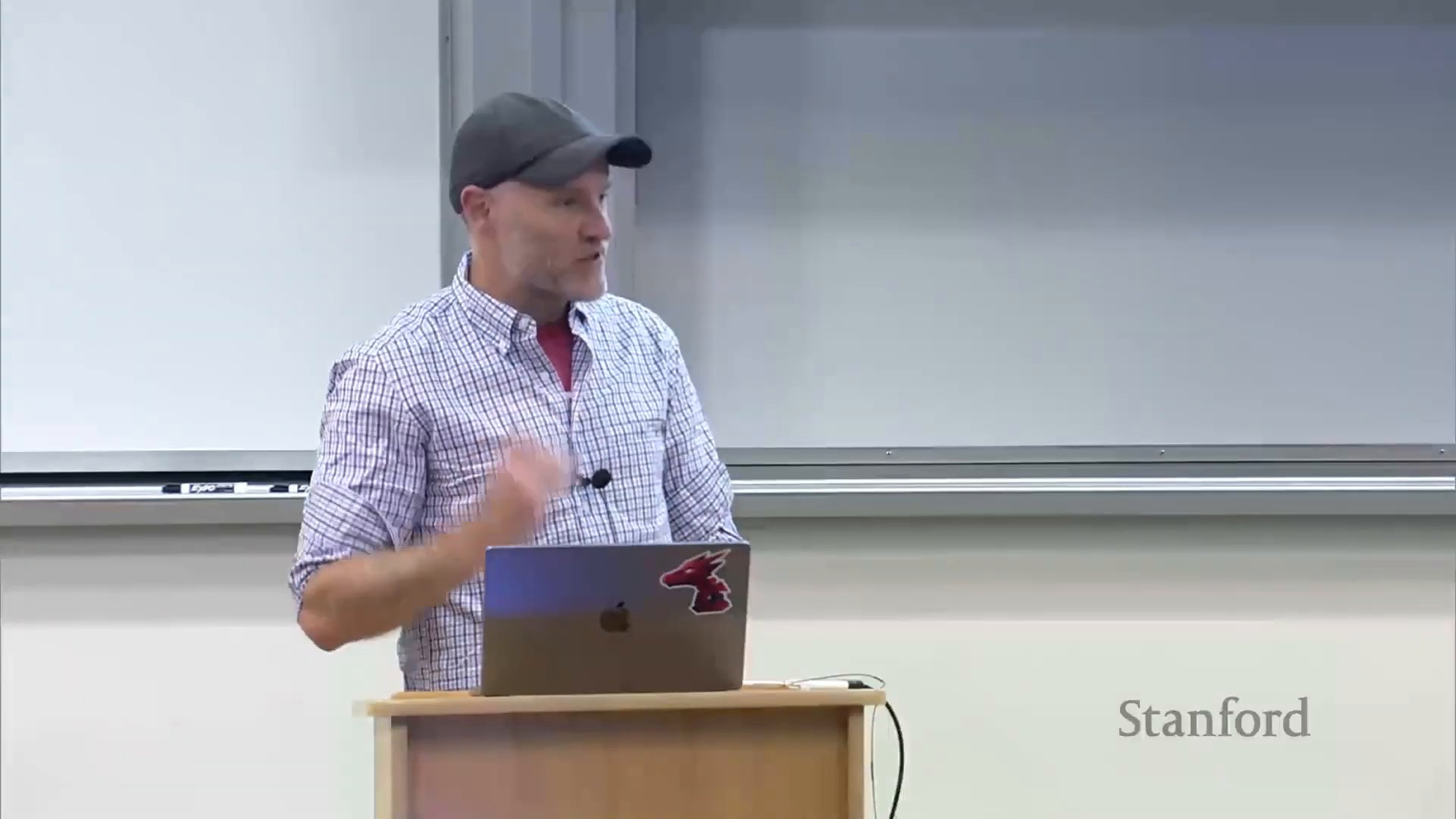

Stanford

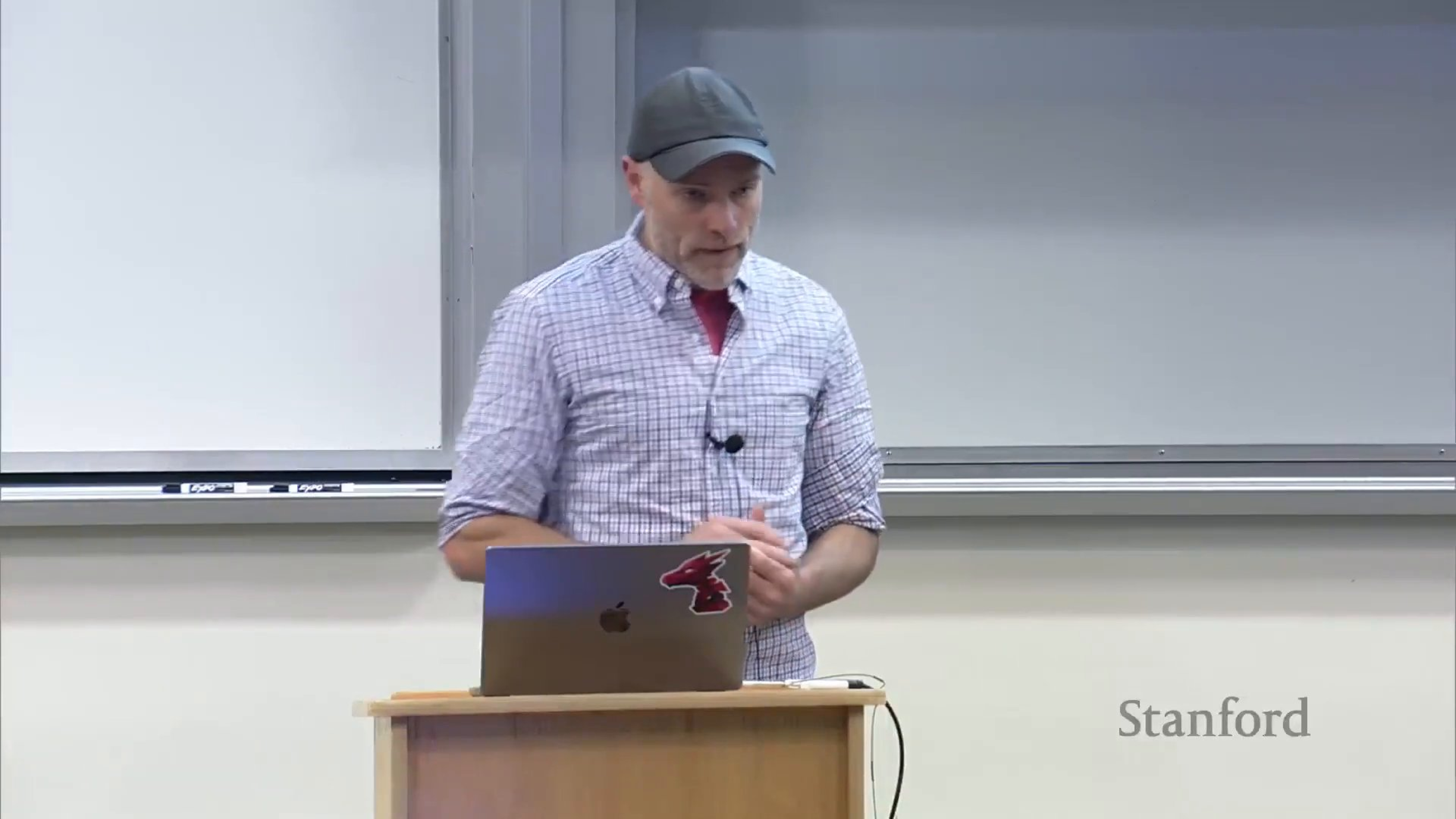

Stanford


Stanford

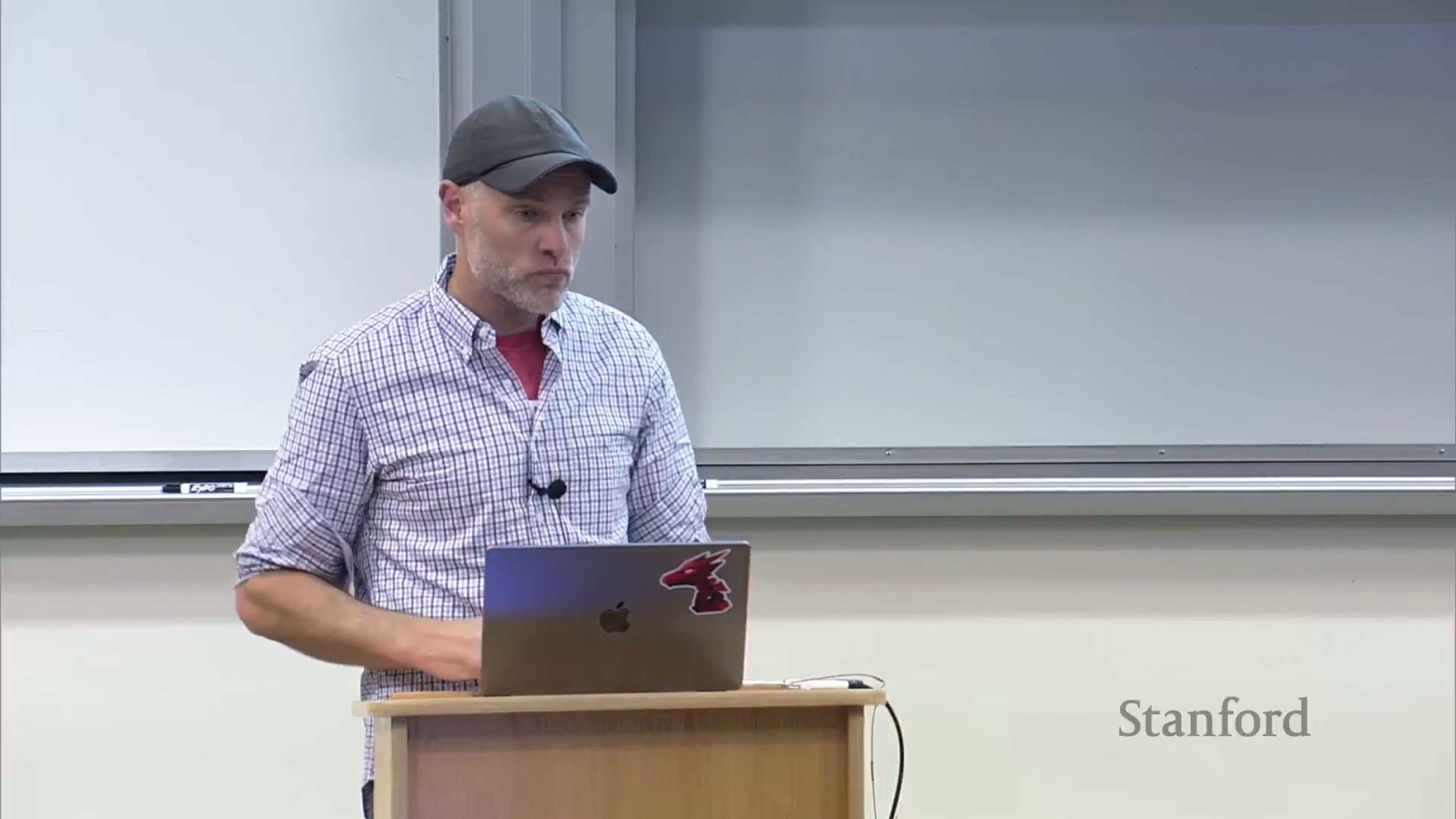

Stanford

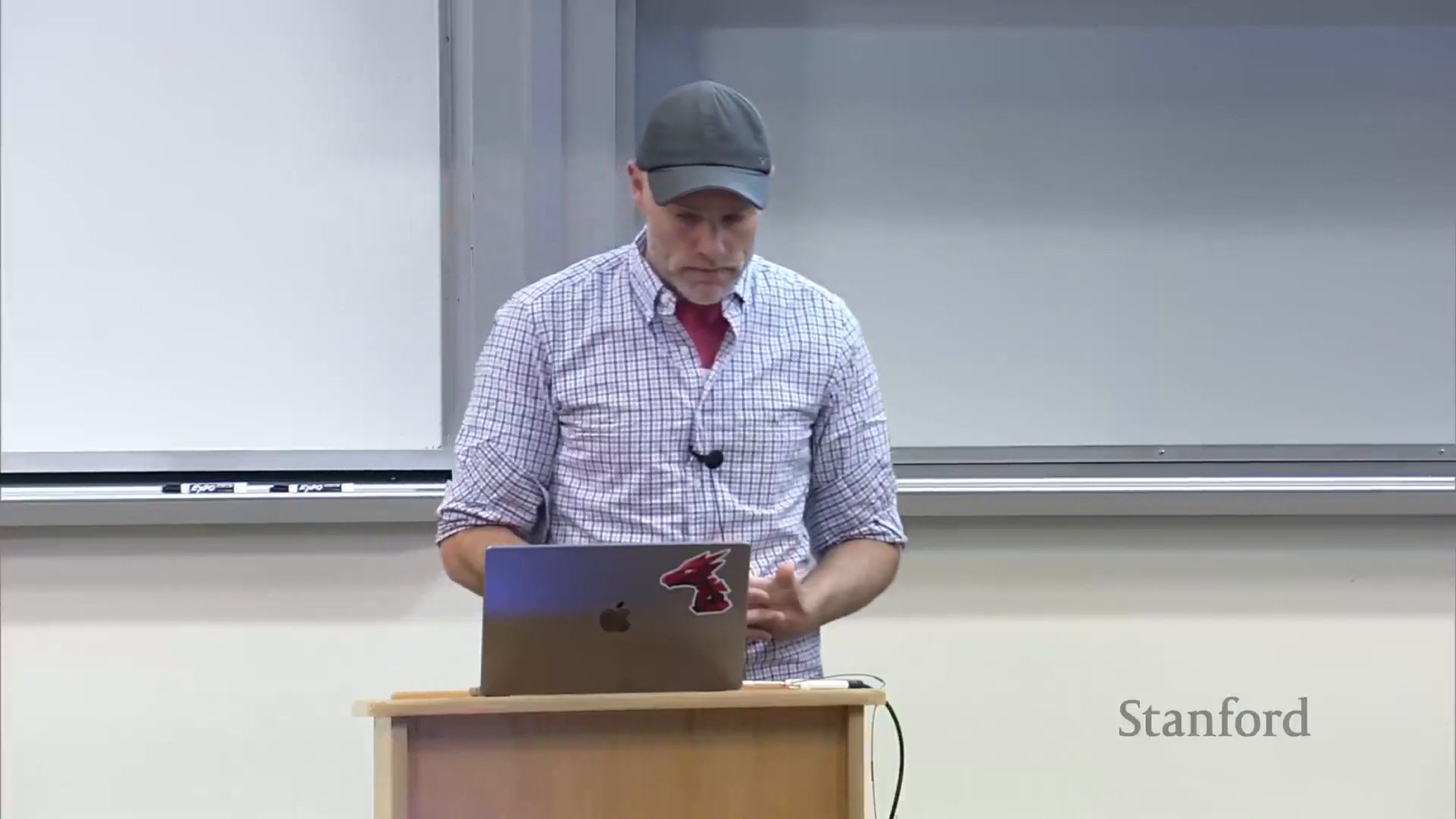

Stanford

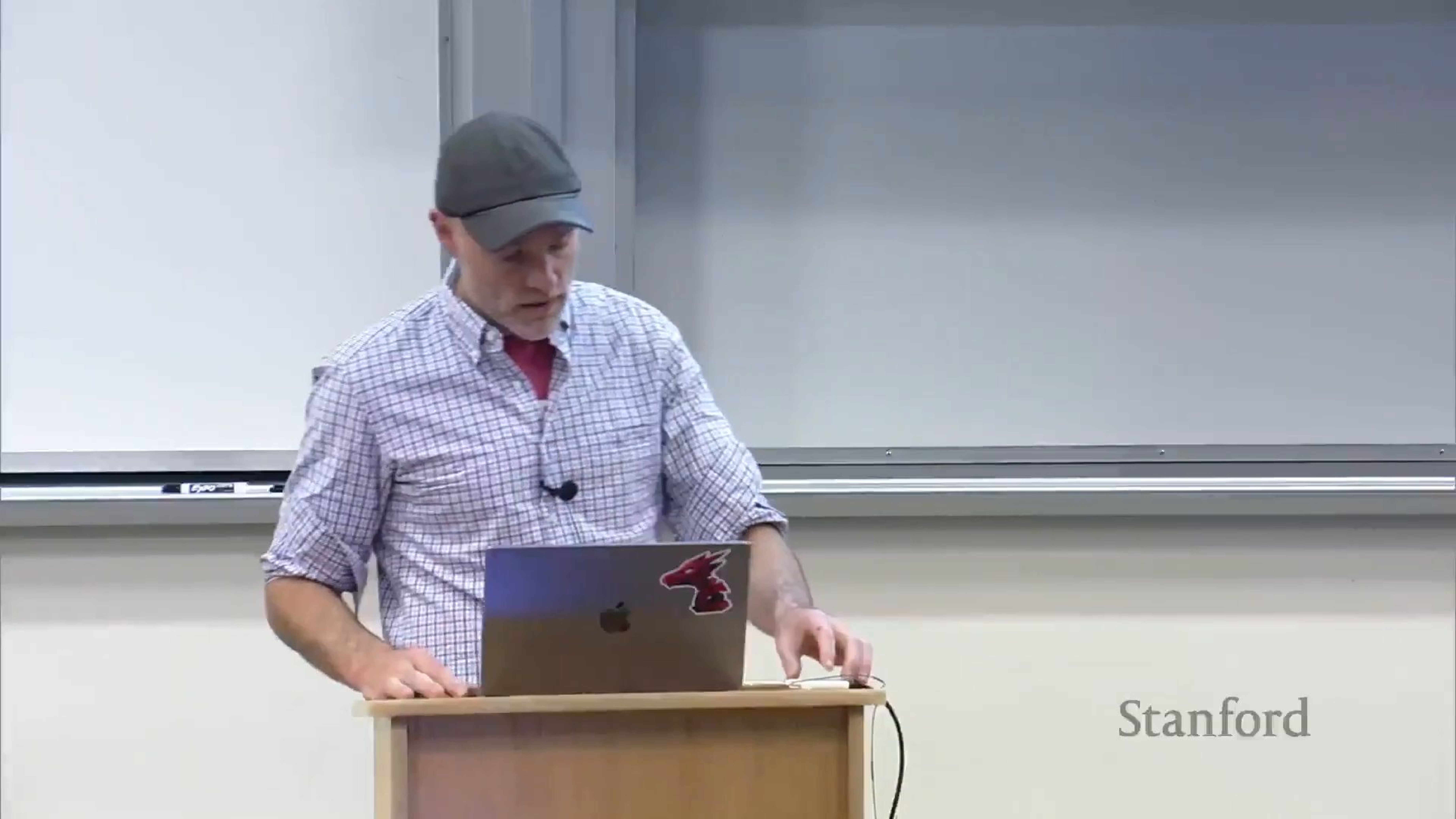

Stanford

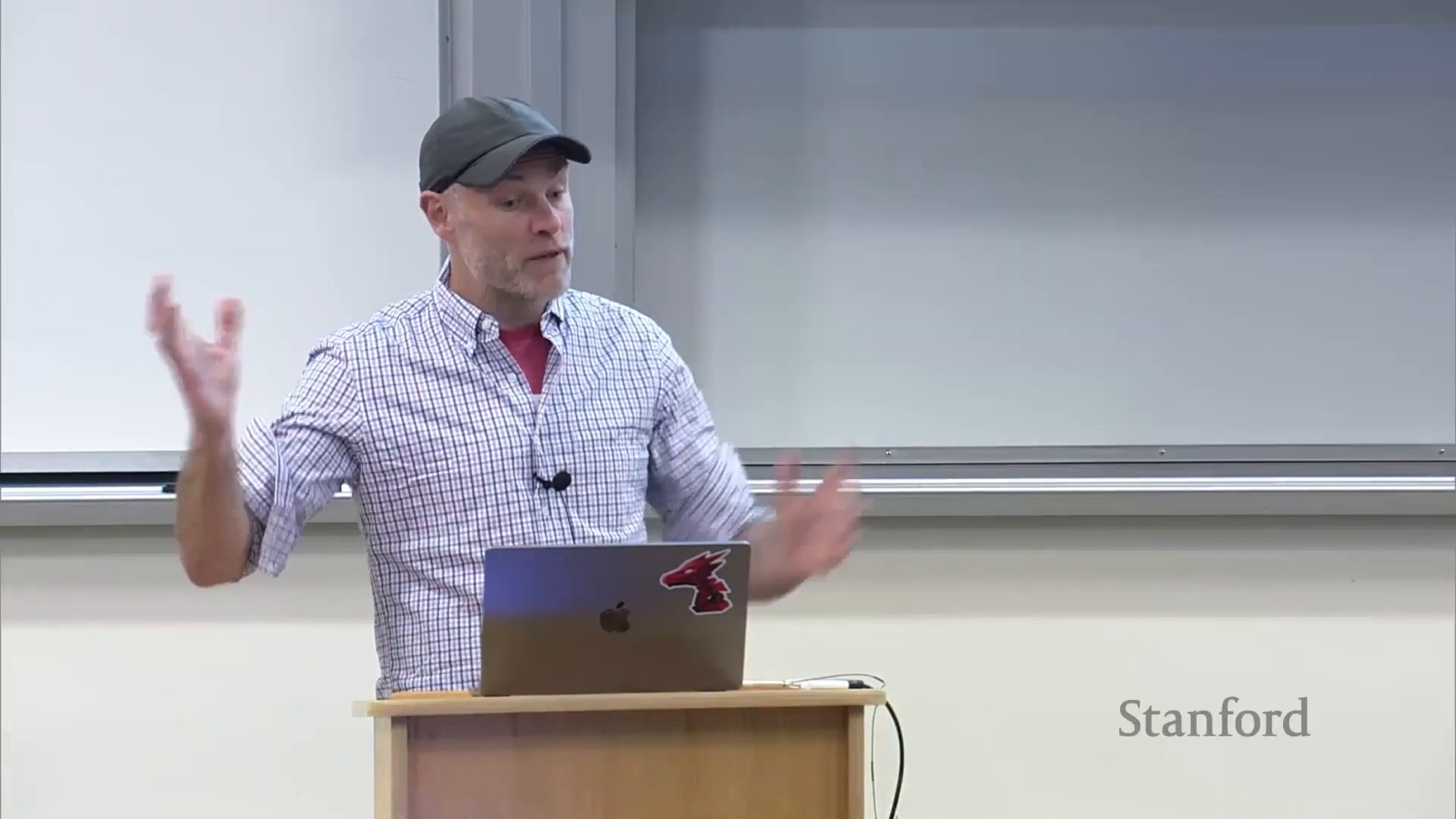

Stanford


Stanford


Stanford


Stanford


Stanford


Stanford


Stanford


Stanford

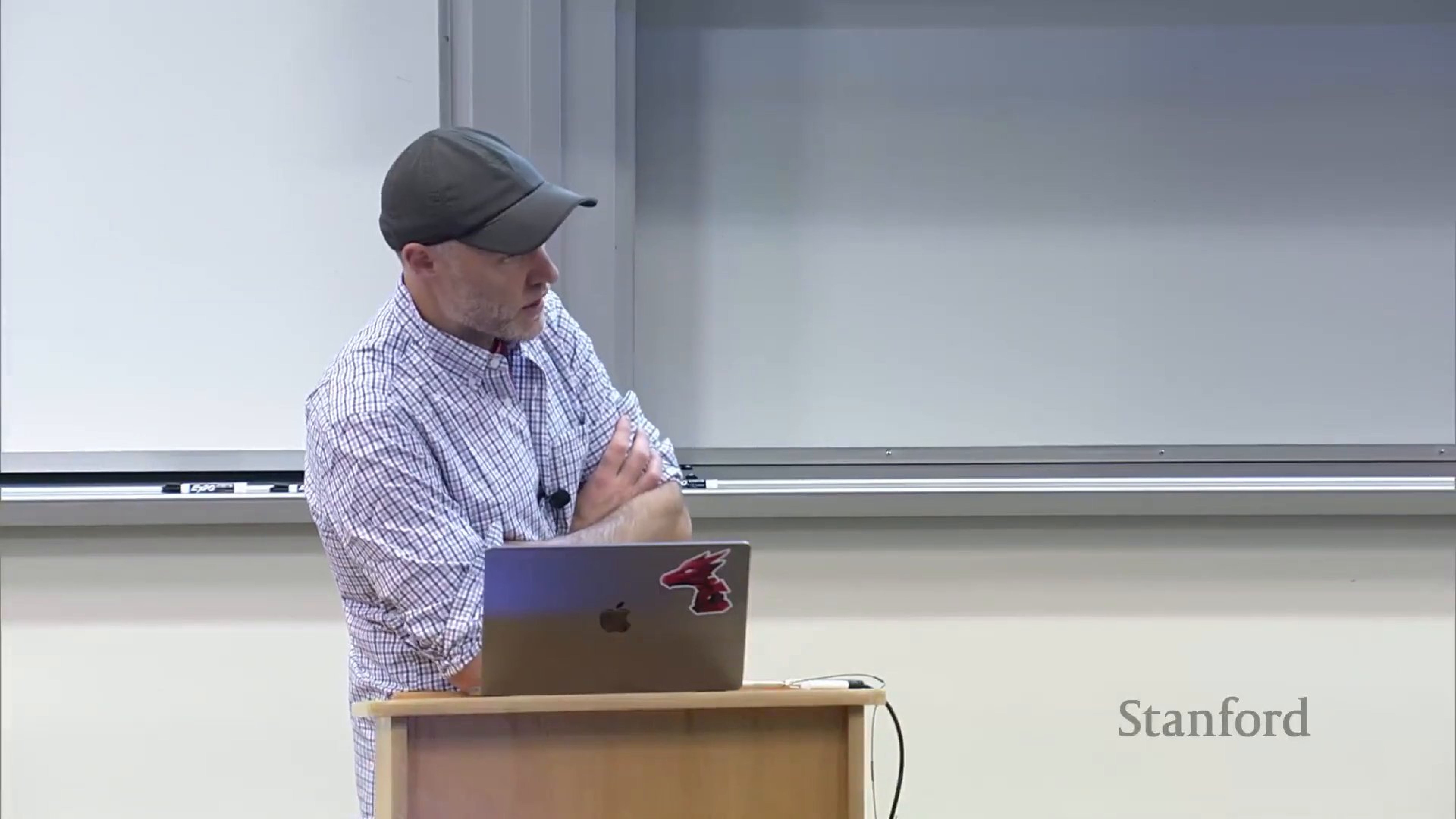

Stanford


Stanford


Stanford

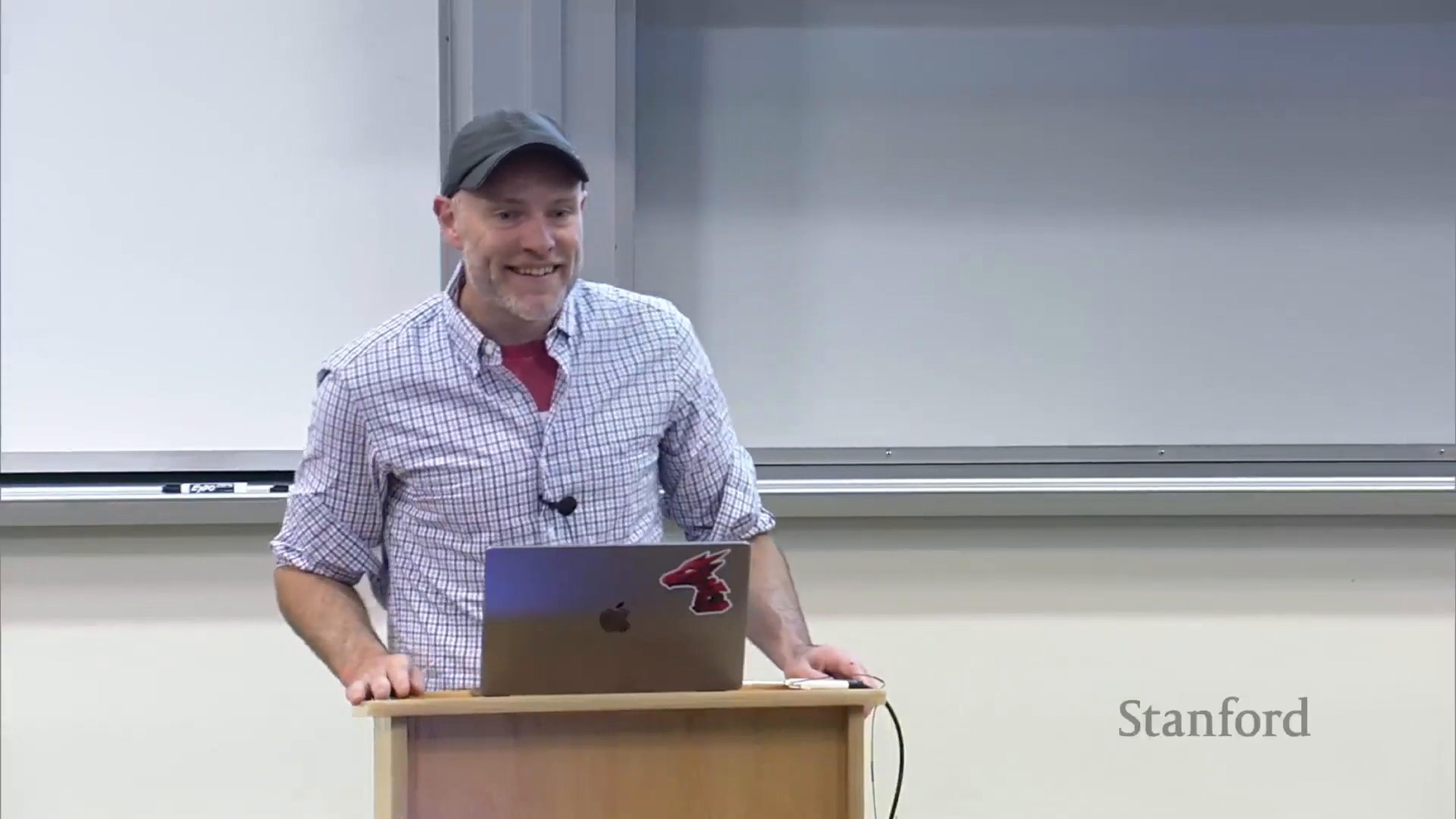

Stanford

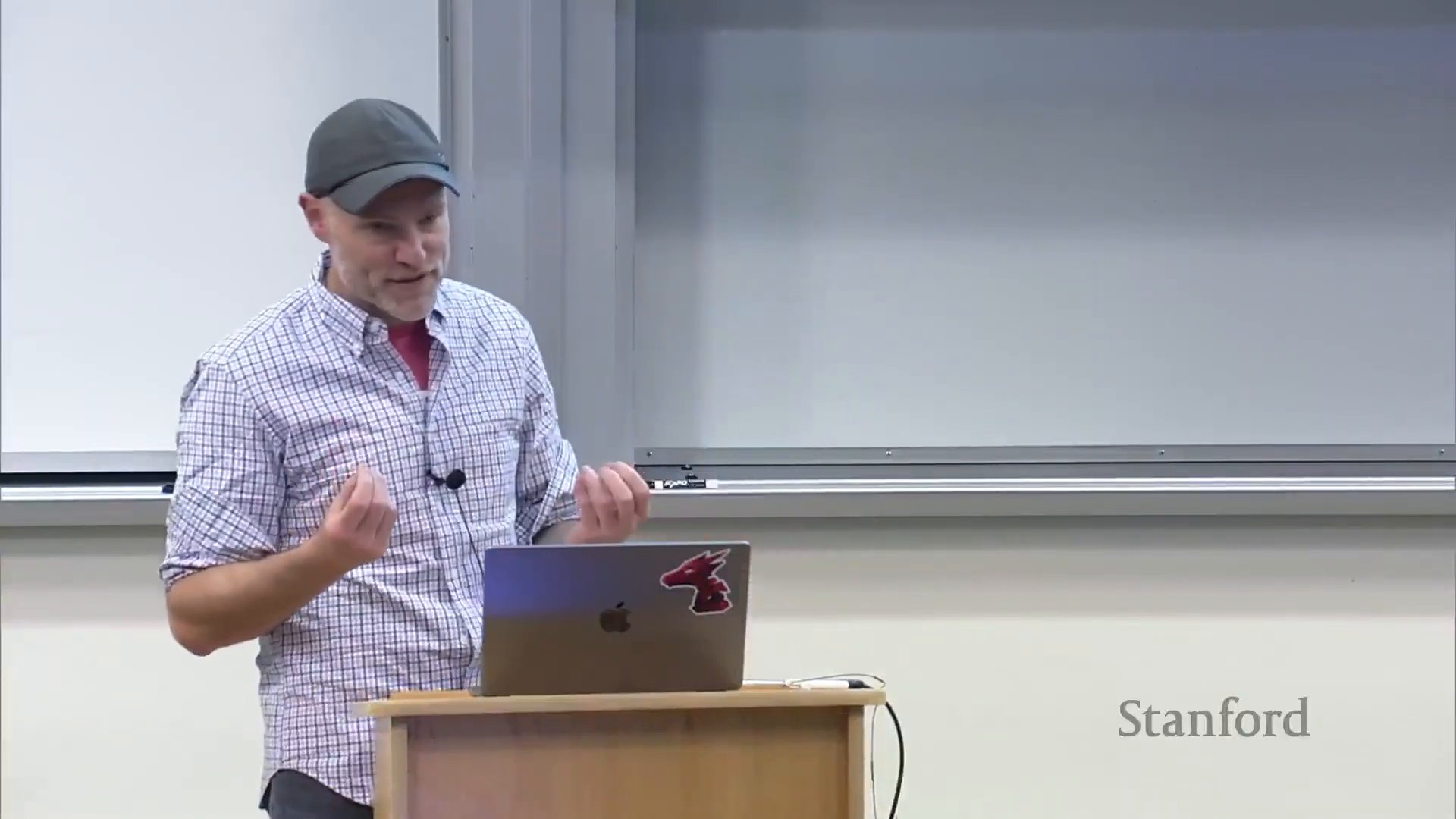

Stanford

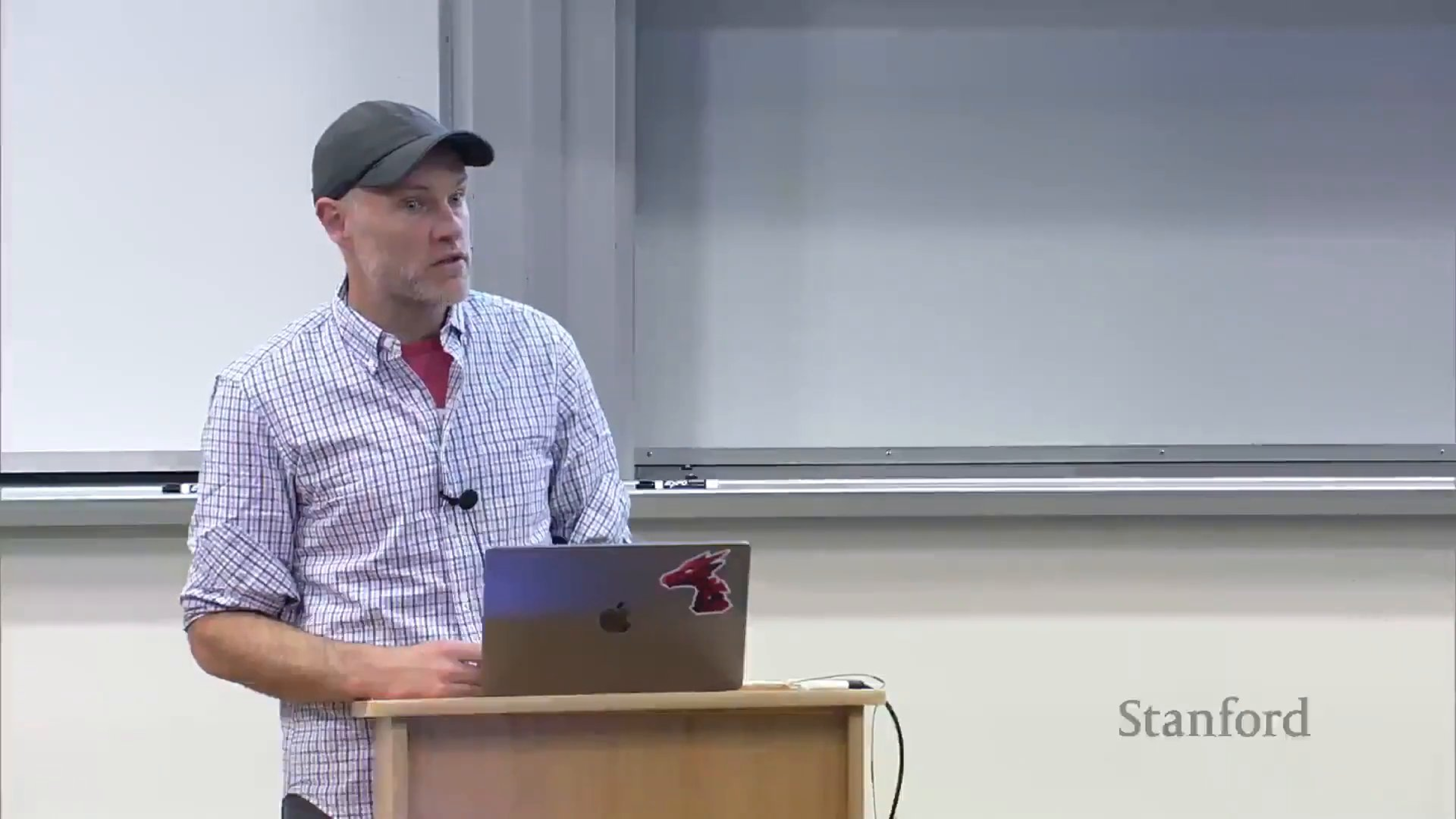

Stanford


Stanford


Stanford

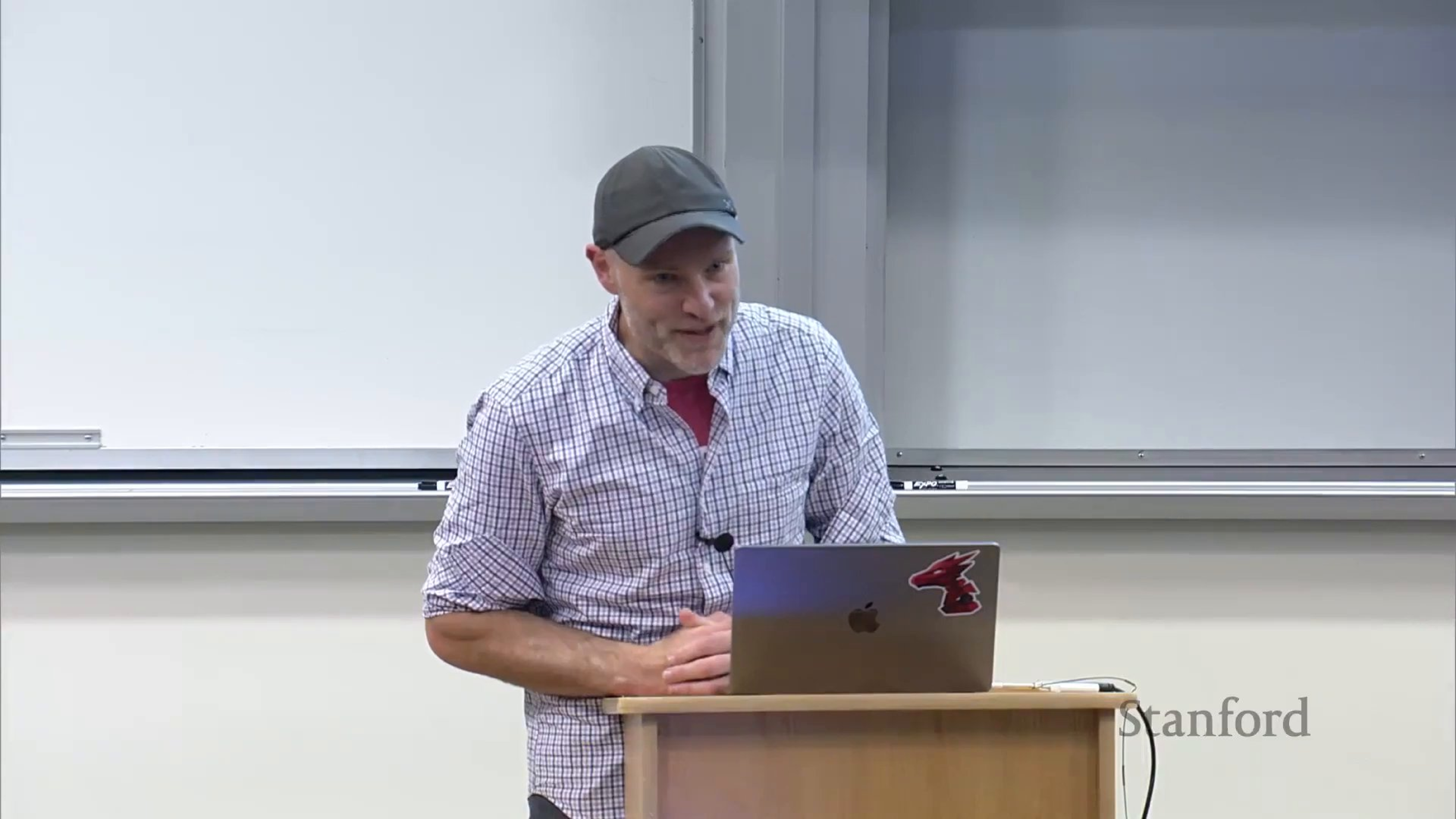

Stanford


Stanford


Stanford


Stanford


Stanford


Stanford


Stanford


Stanford


Stanford


Stanford

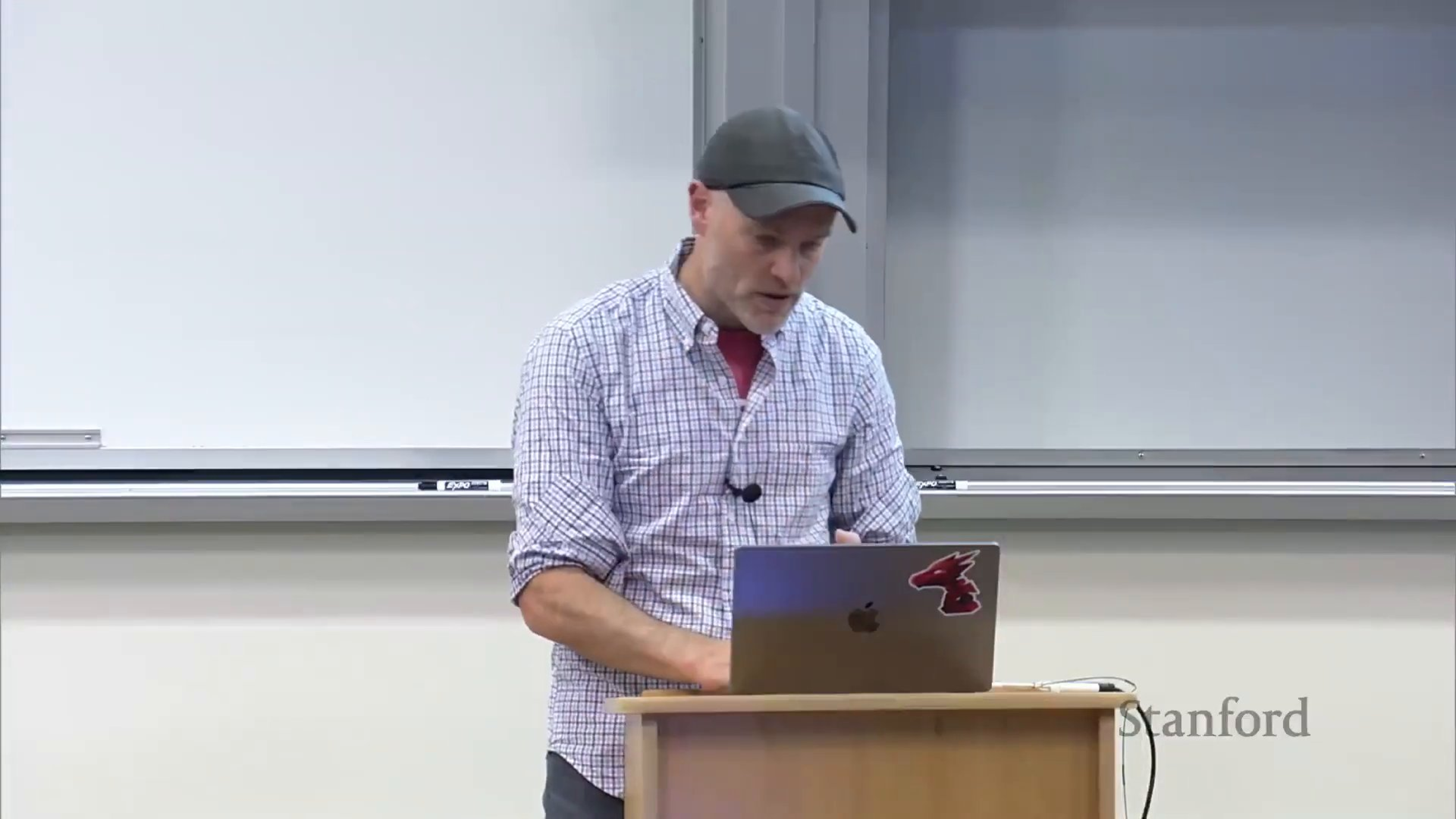

Stanford


Stanford


Stanford


Stanford


Stanford


Stanford


Stanford


Stanford


Stanford


Stanford


Stanford


Stanford


Stanford


Stanford


Stanford


Stanford


Stanford


Stanford


Stanford


Stanford


Stanford

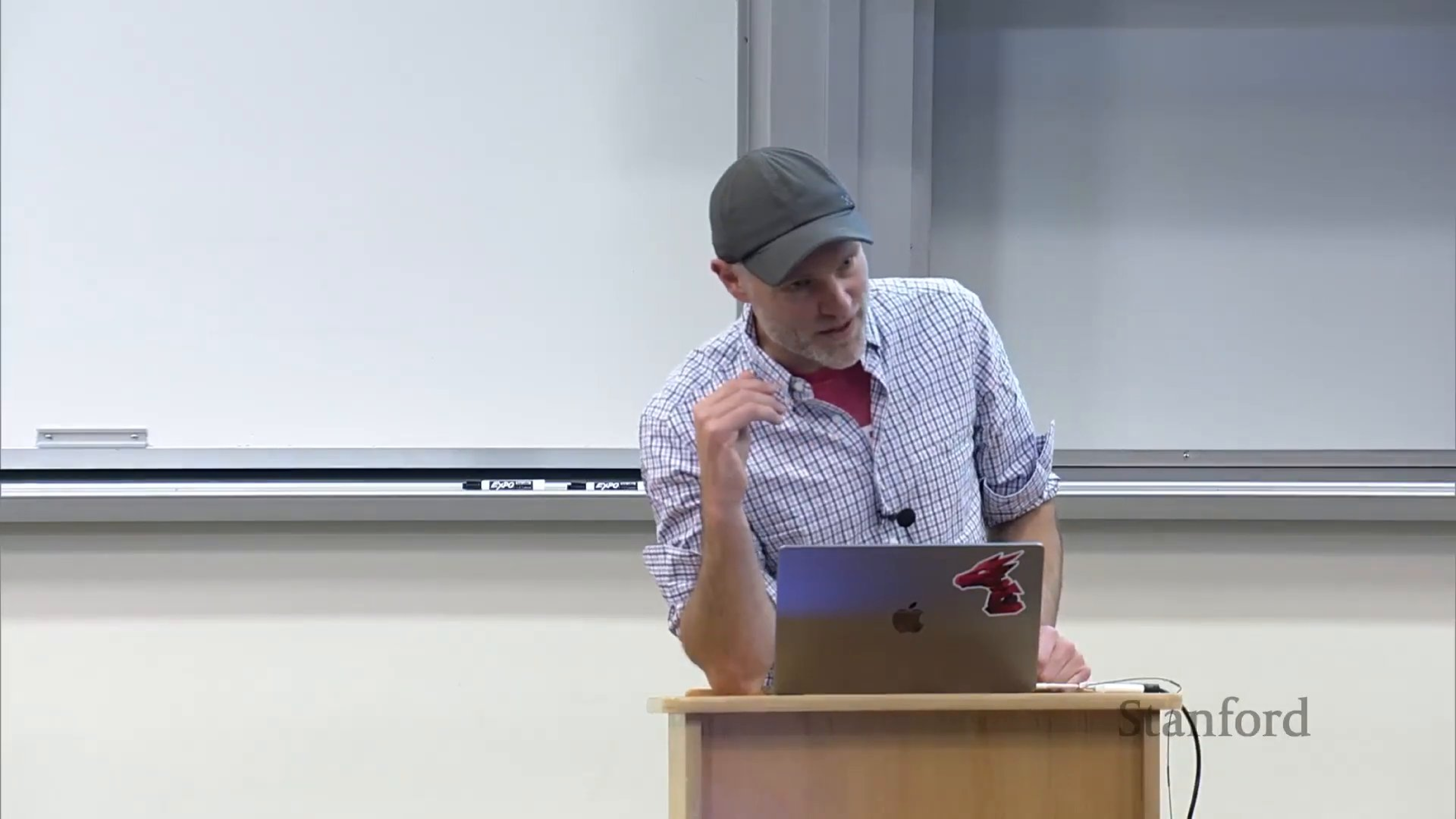

Stanford


Stanford

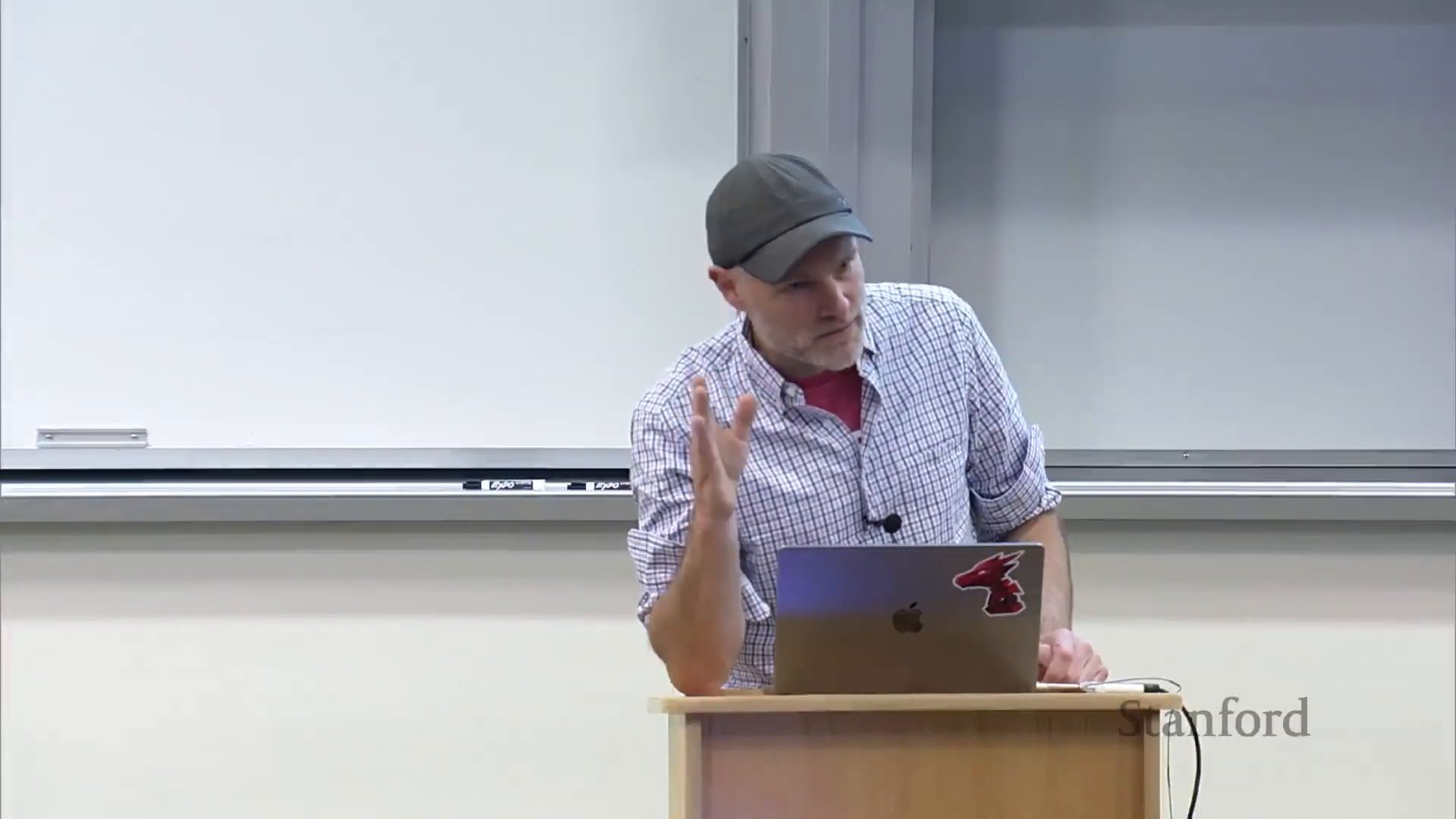

Stanford


Stanford

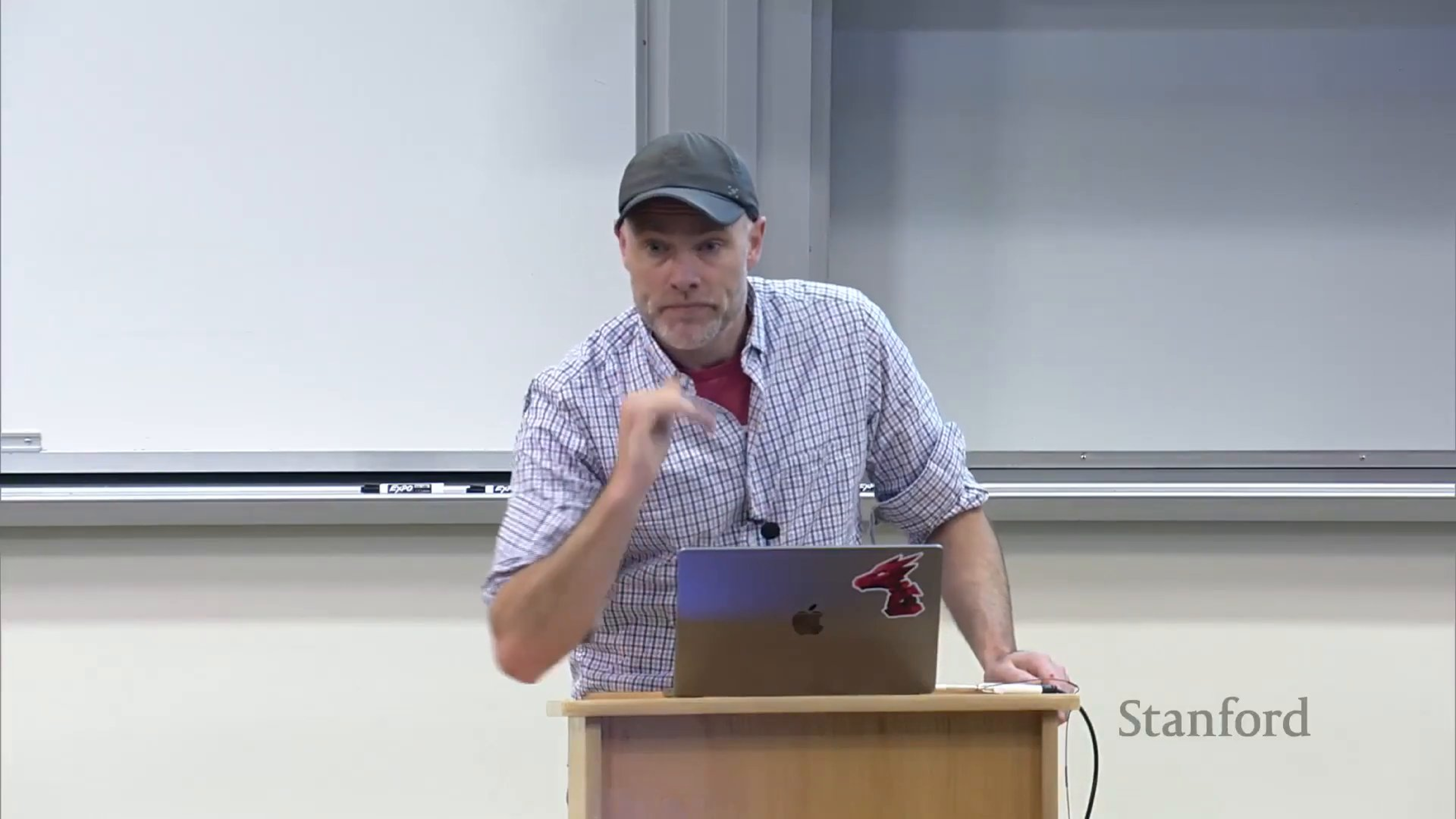

Stanford

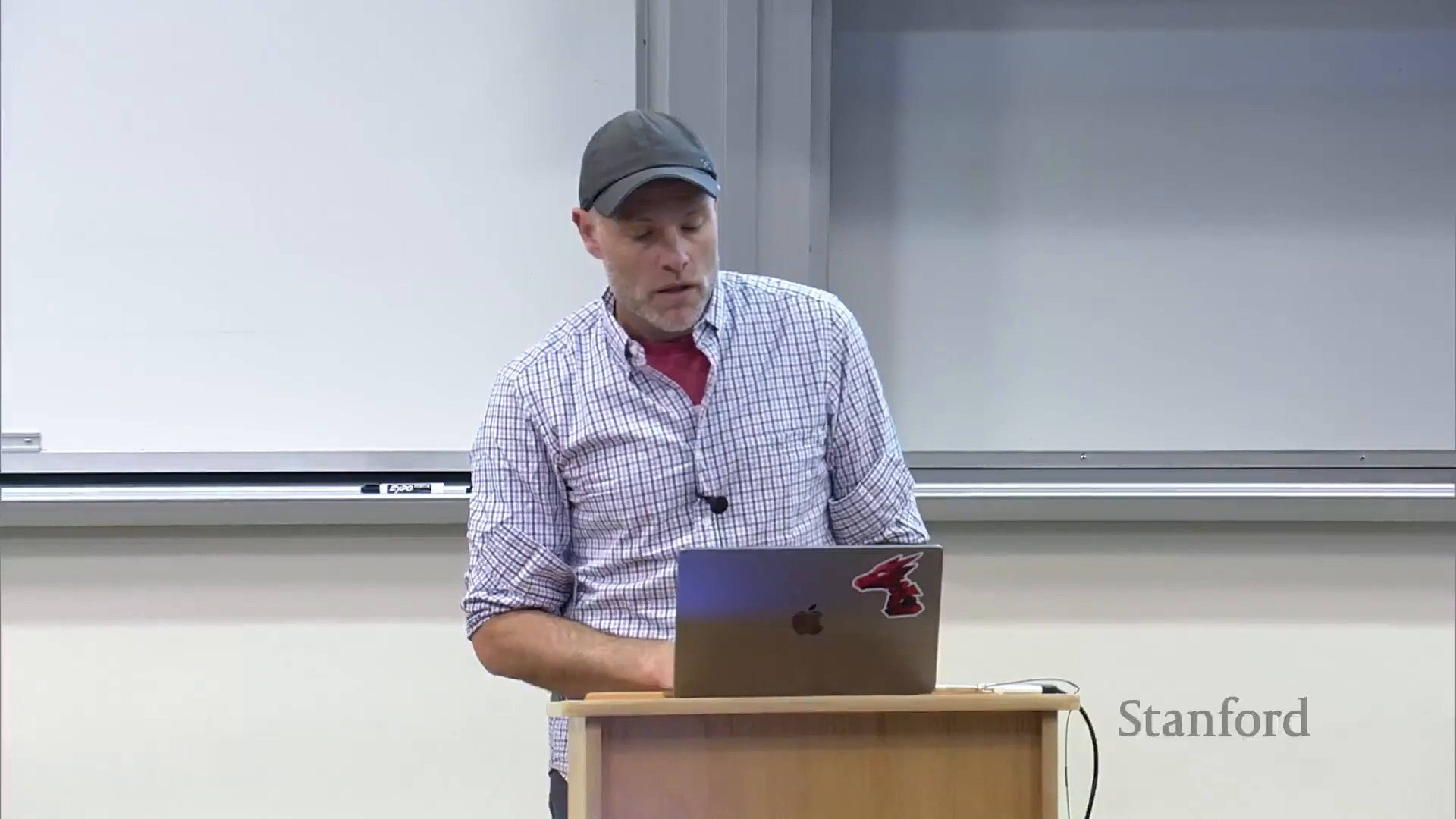

Stanford

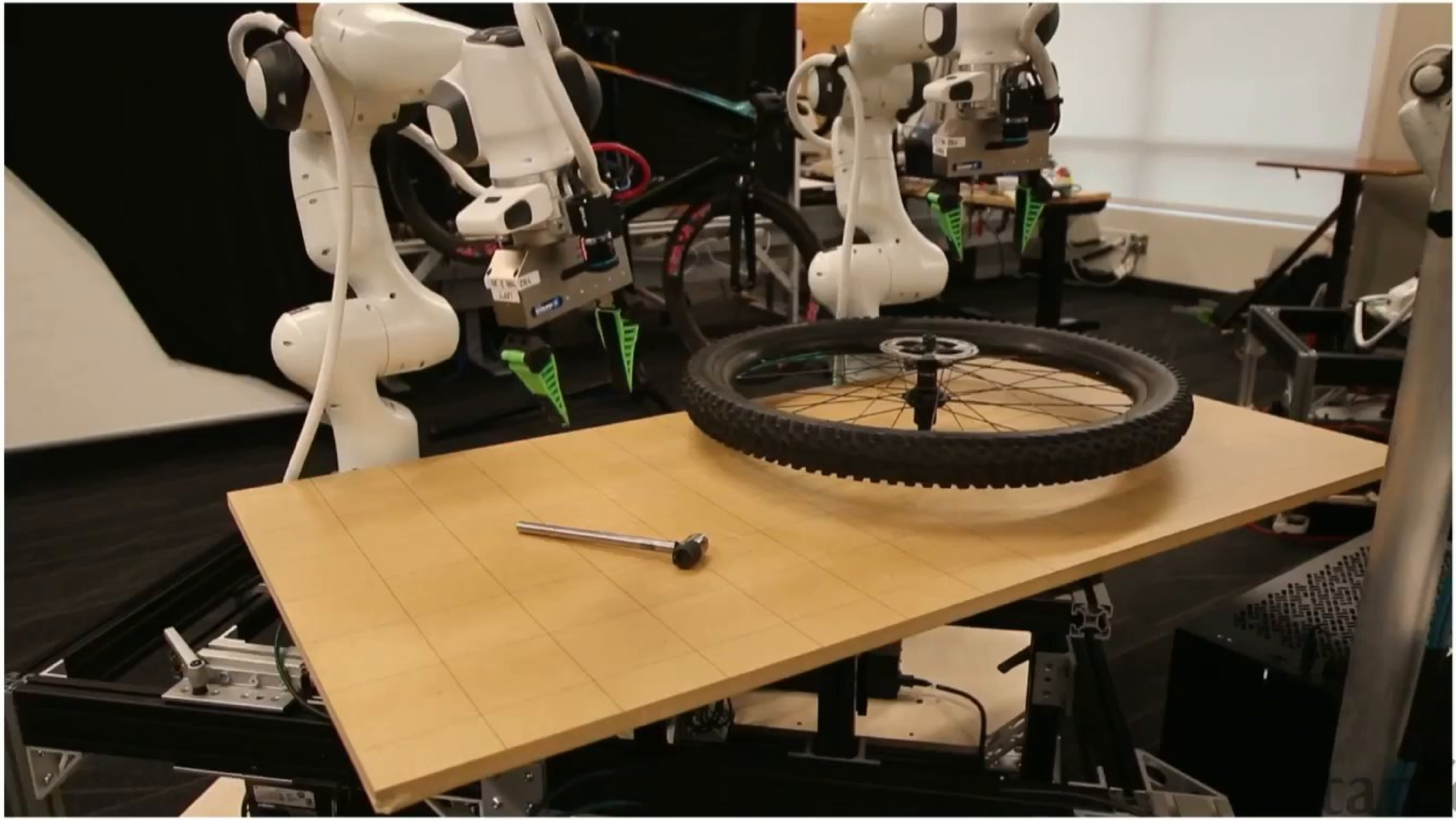

Stanford


Stanford


Stanford

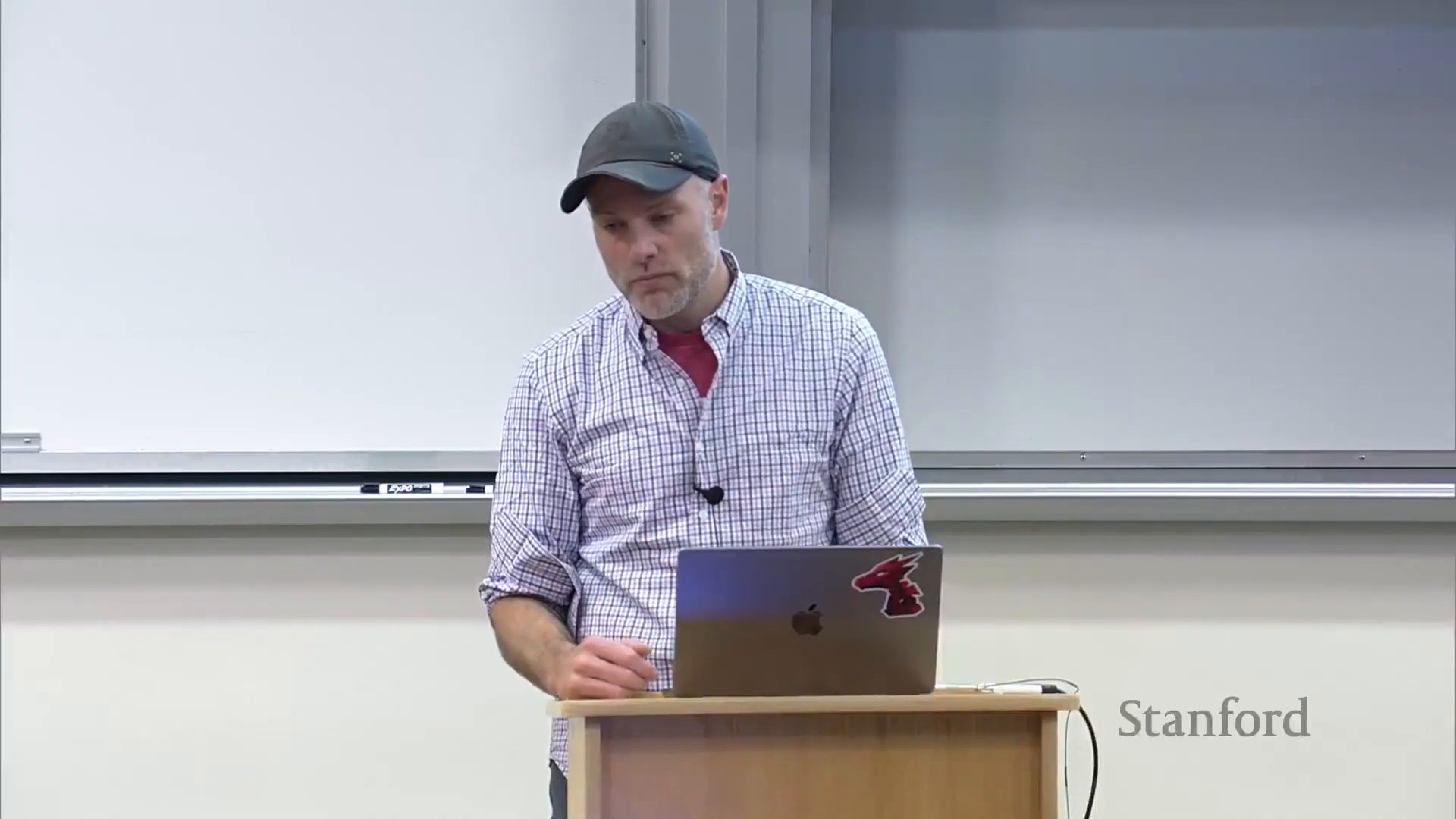

Stanford

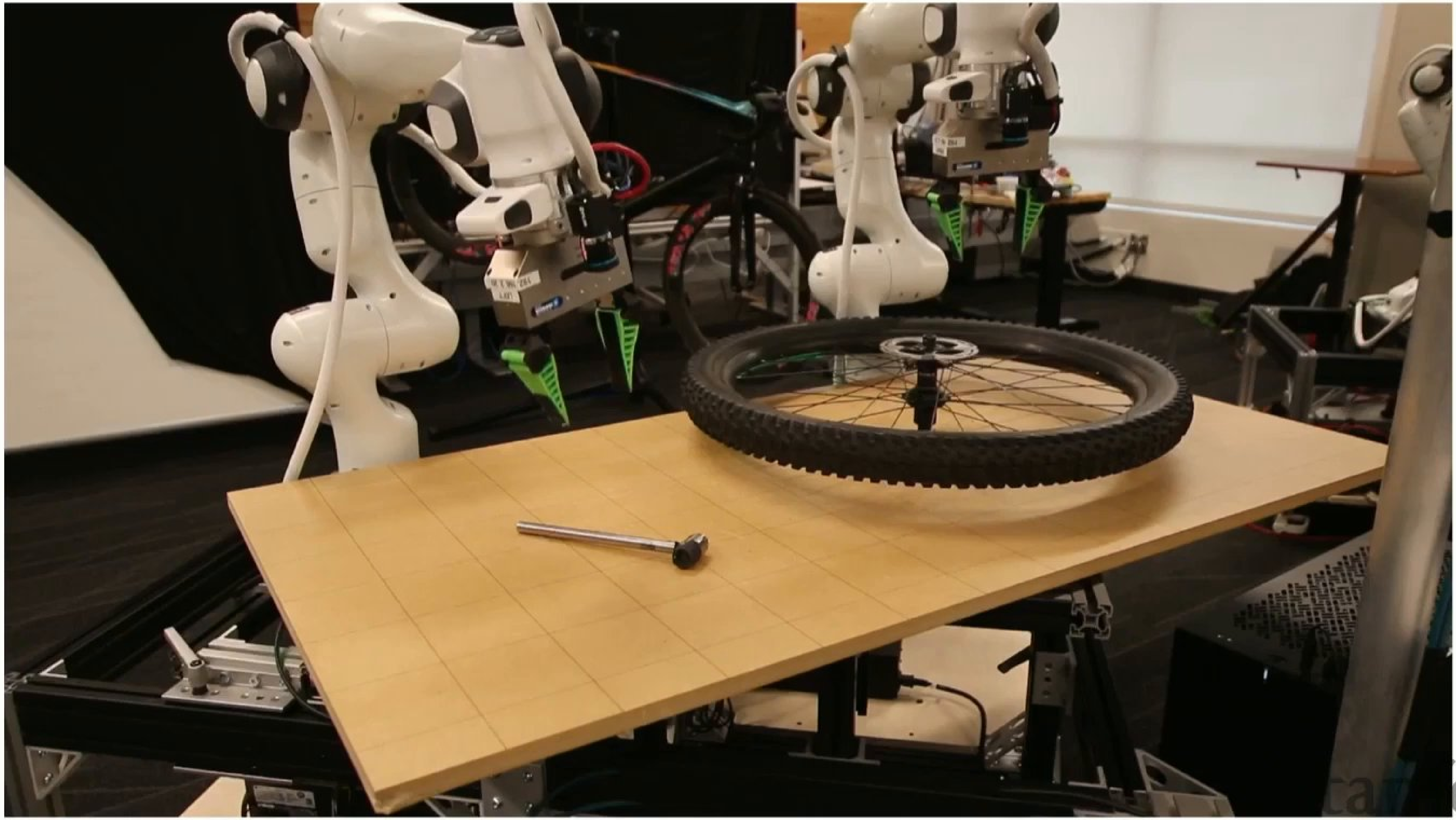

Stanford

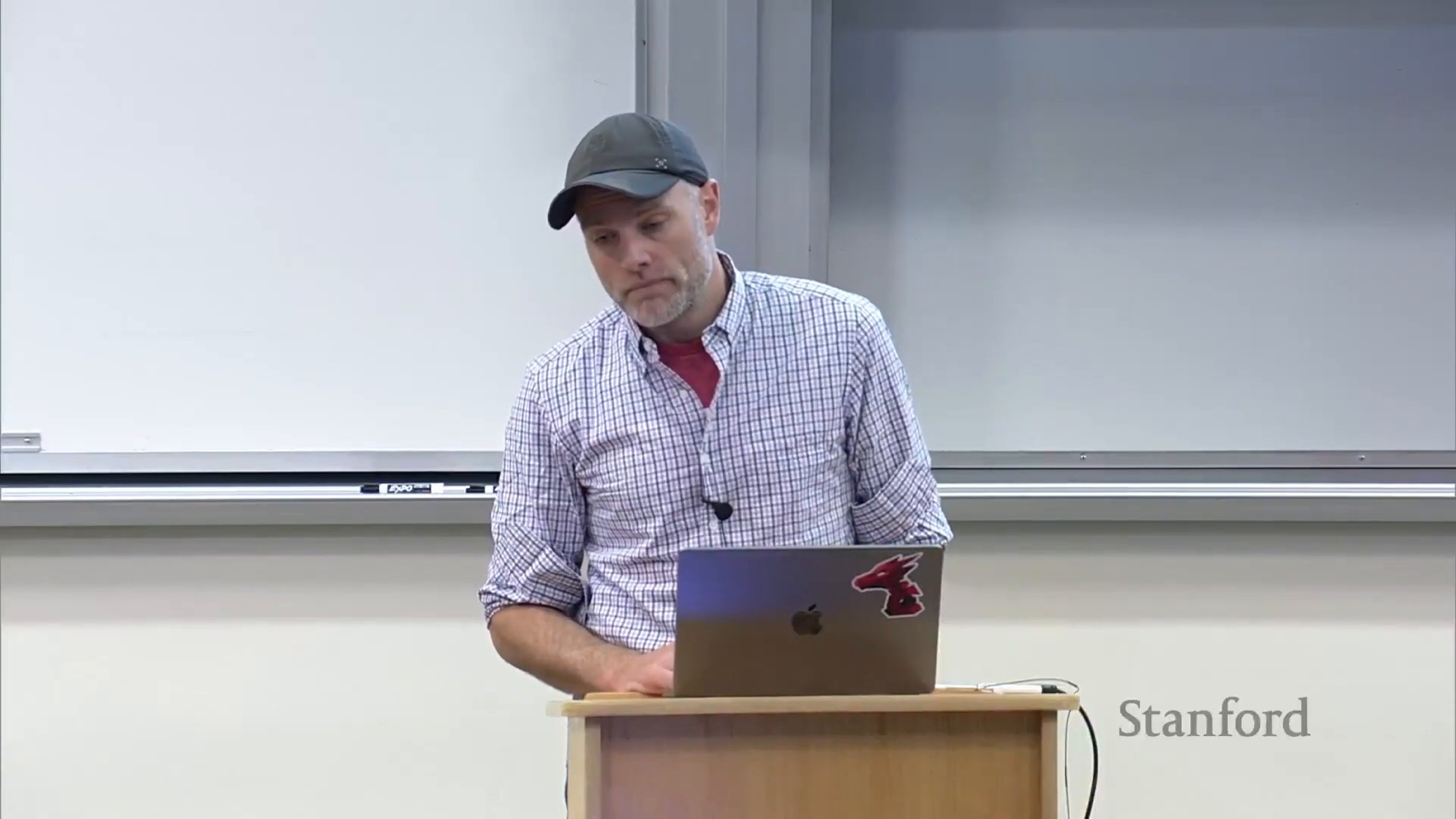


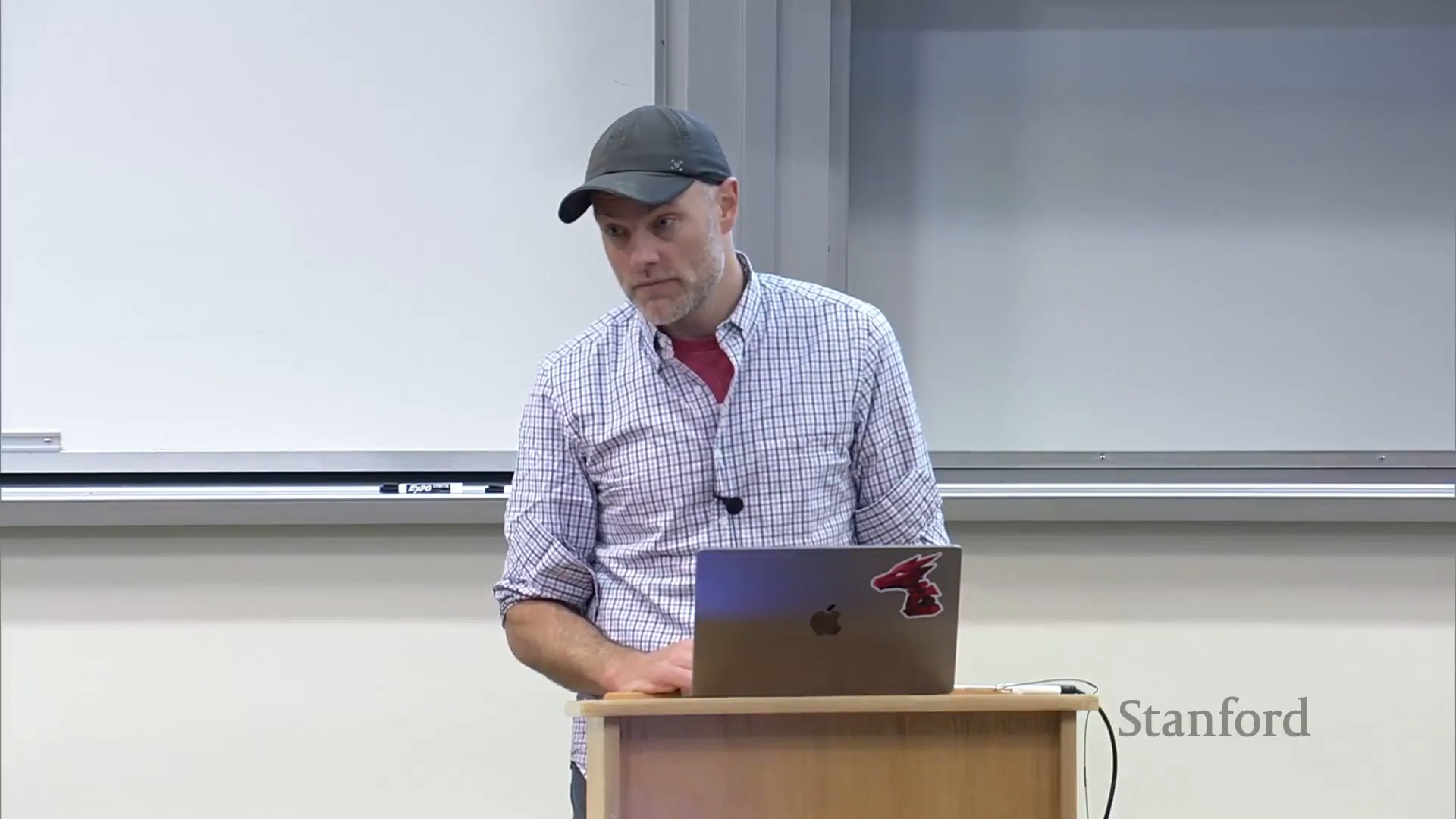

Stanford

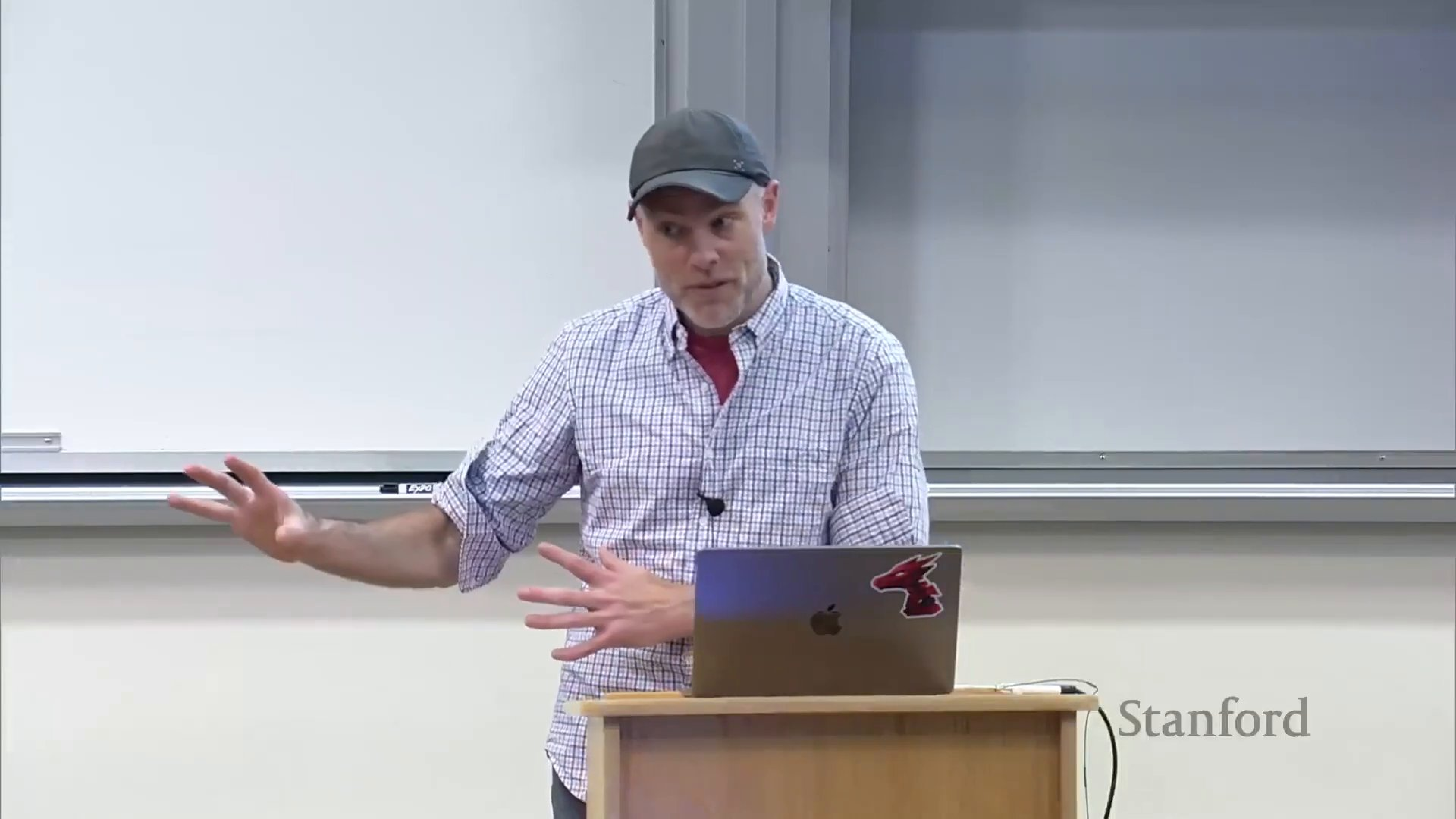

Stanford

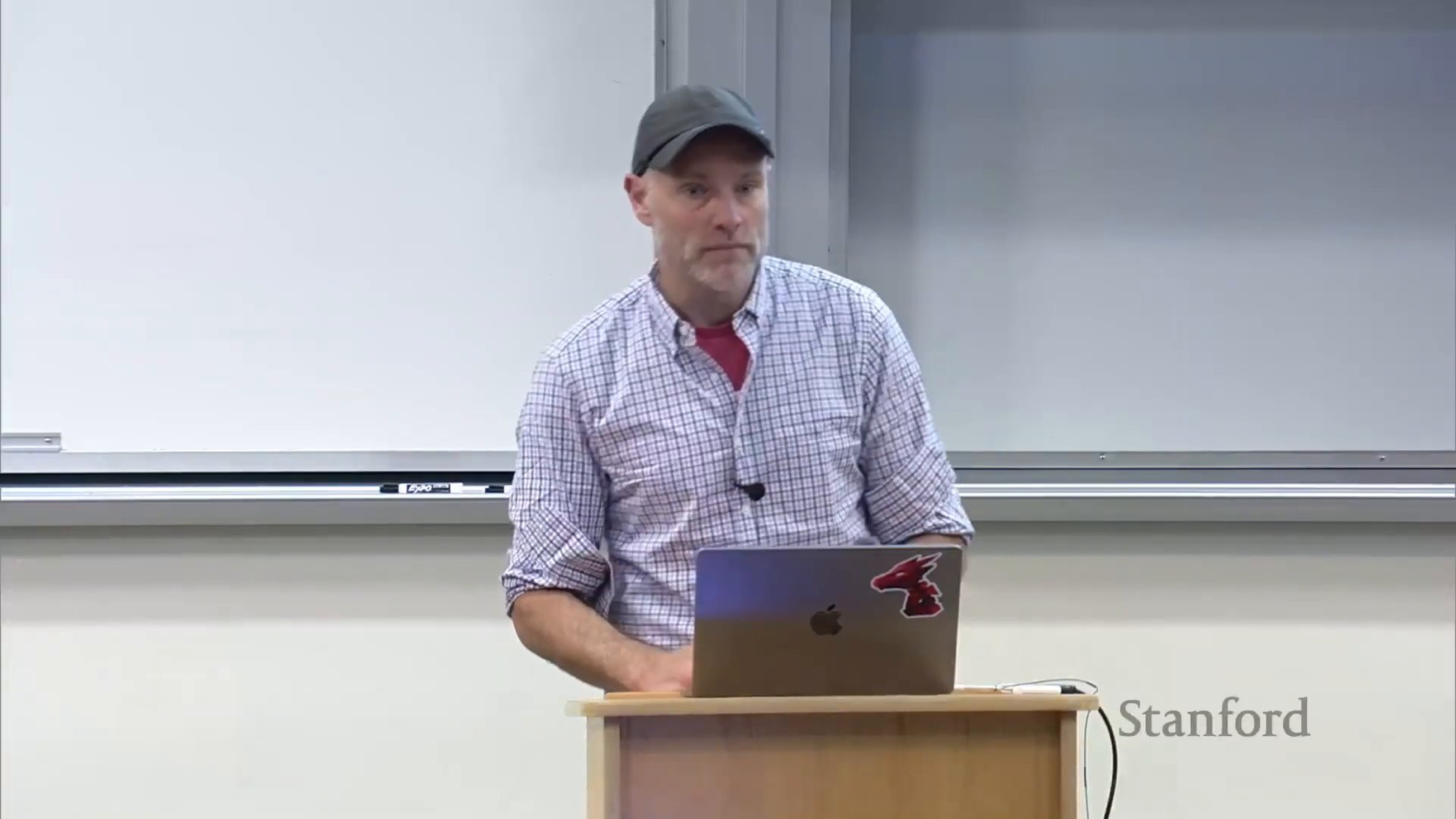

Stanford

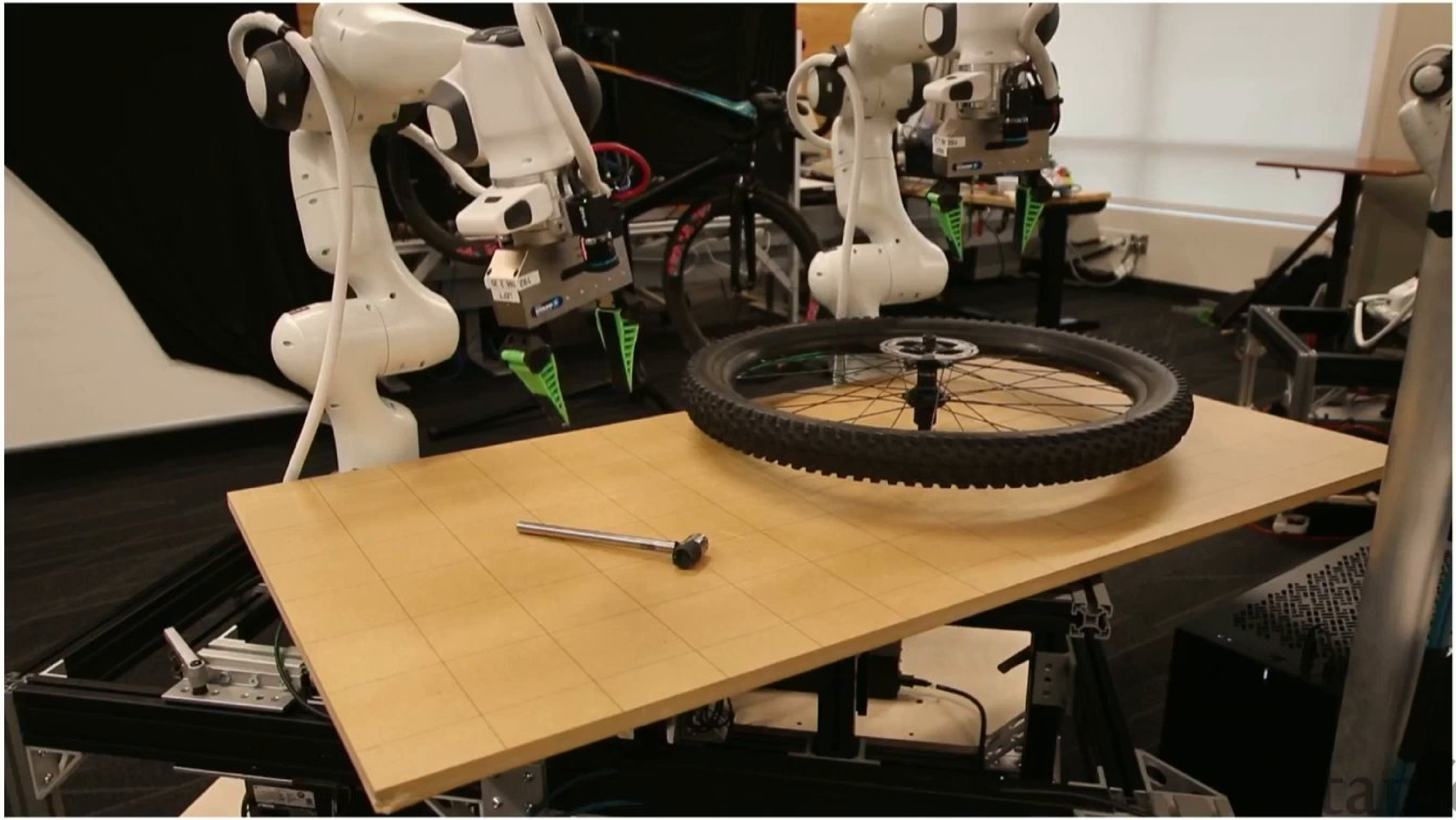


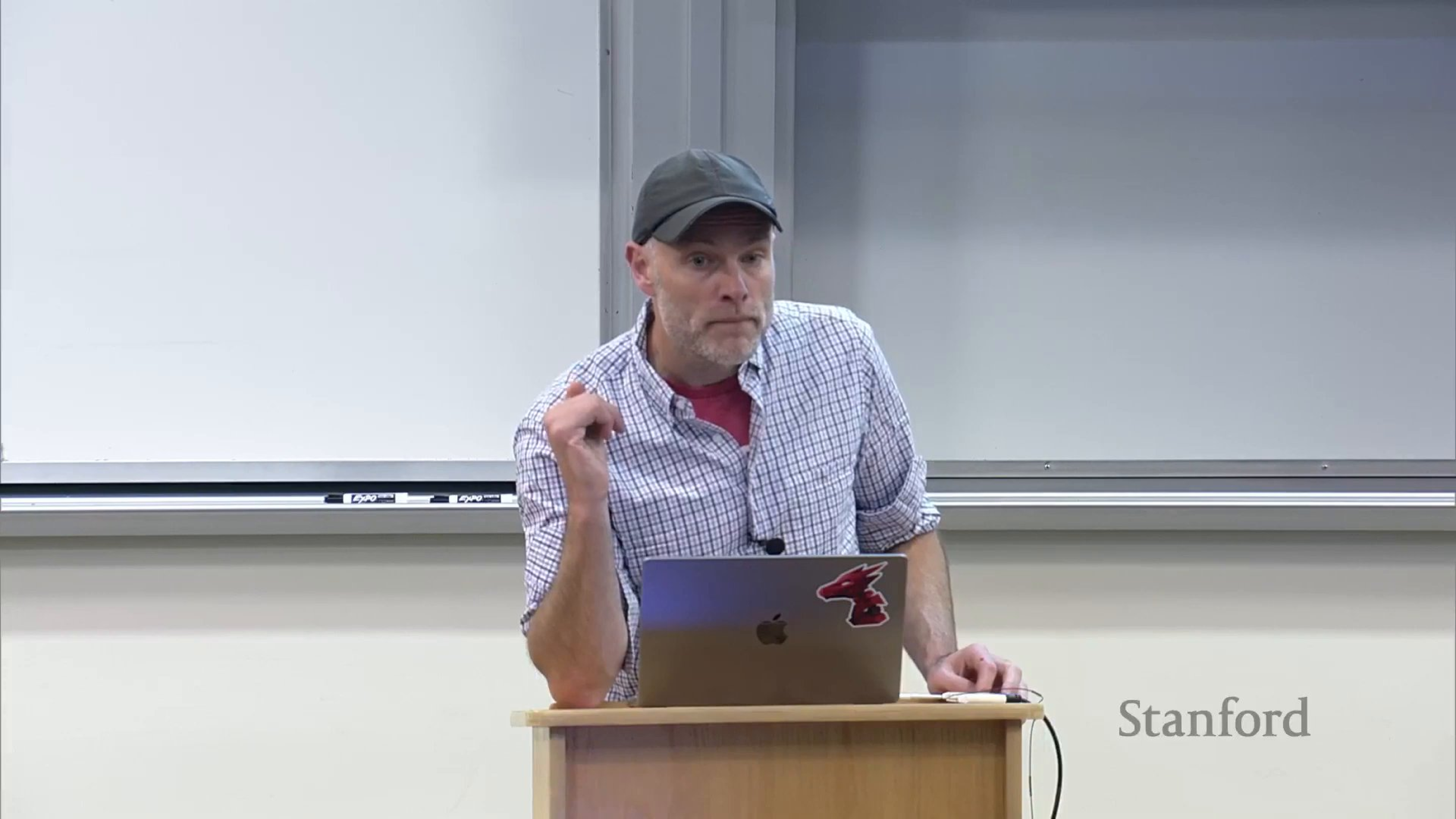

Stanford

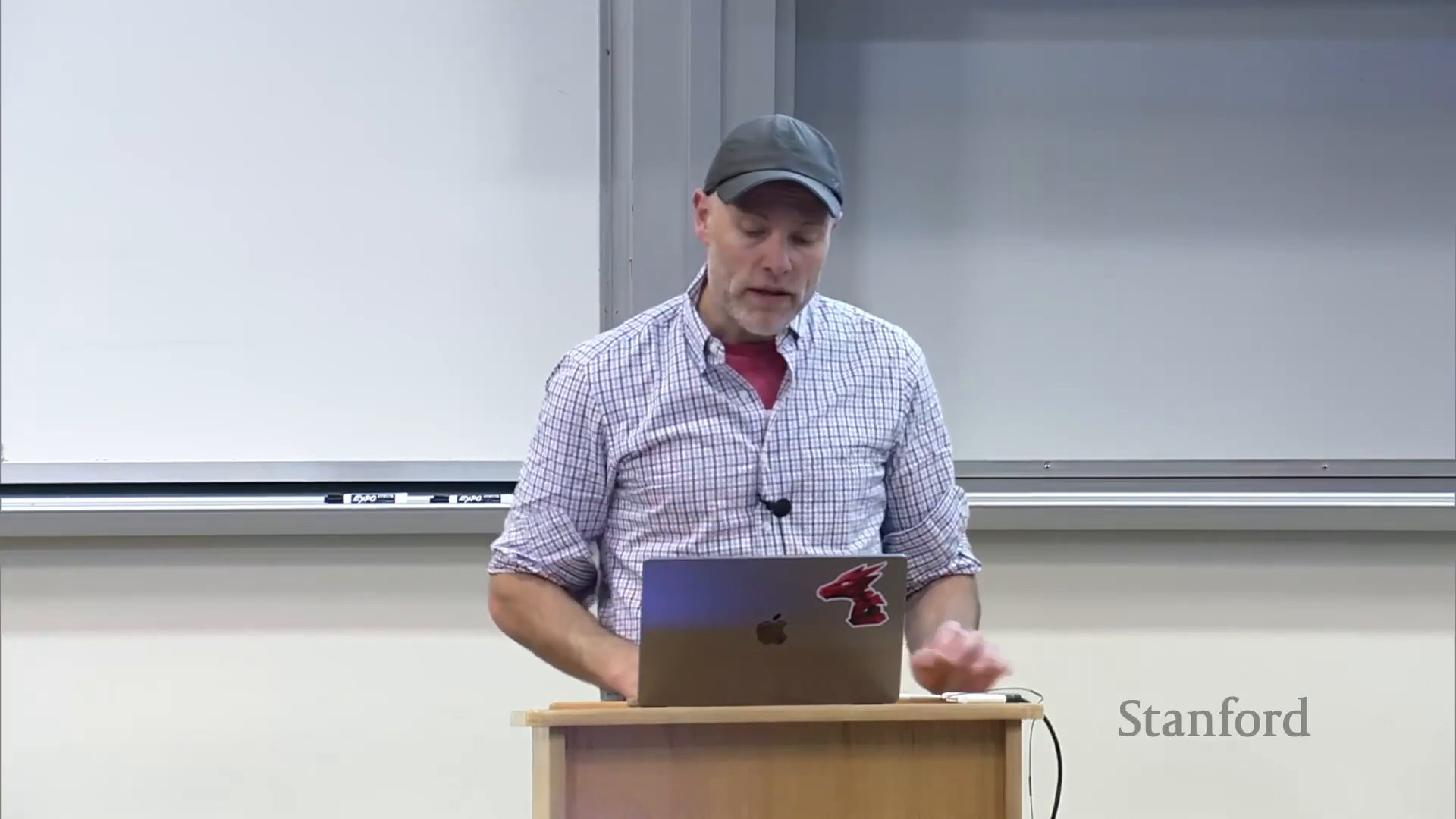

Stanford


Stanford

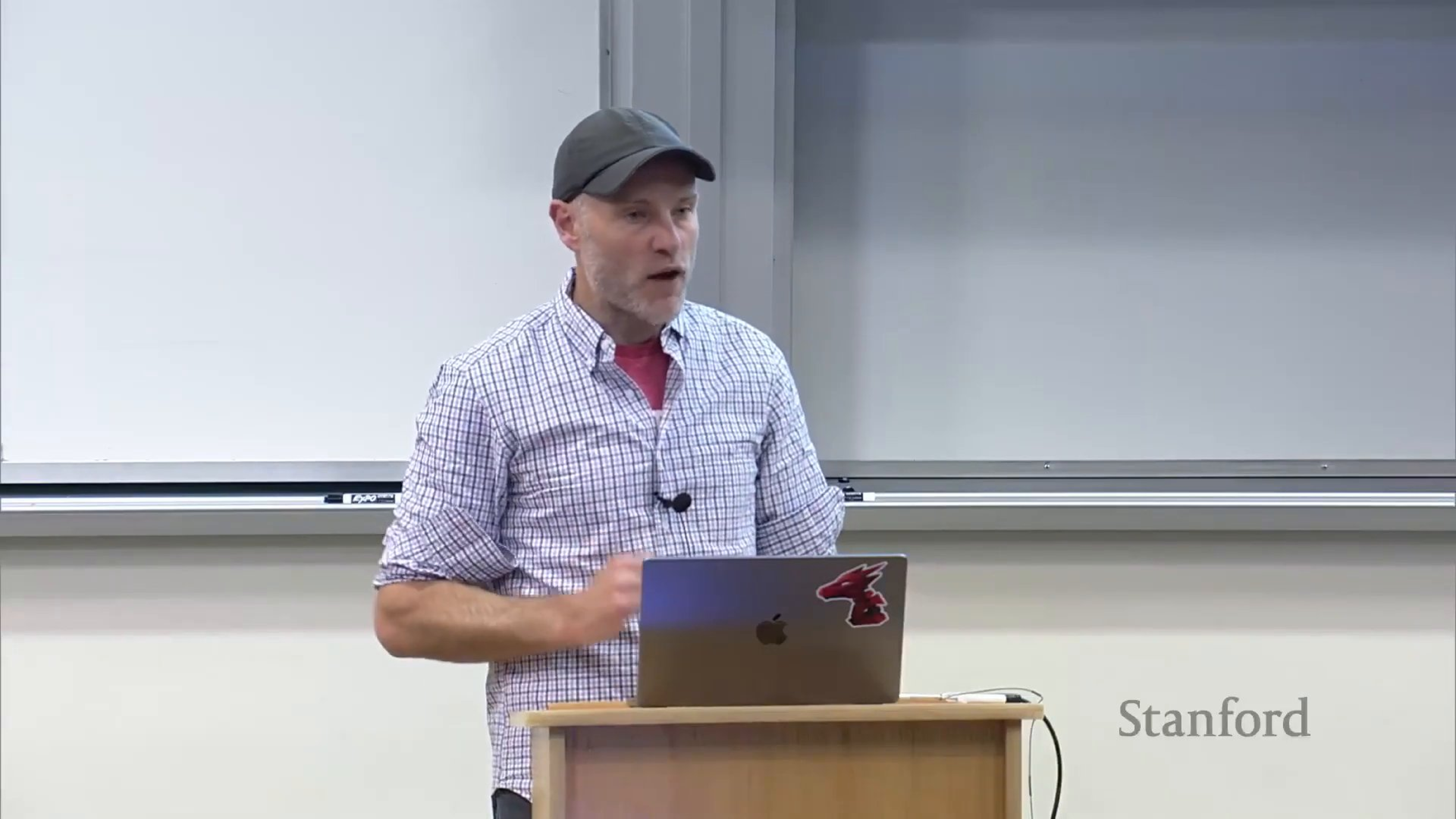

Stanford



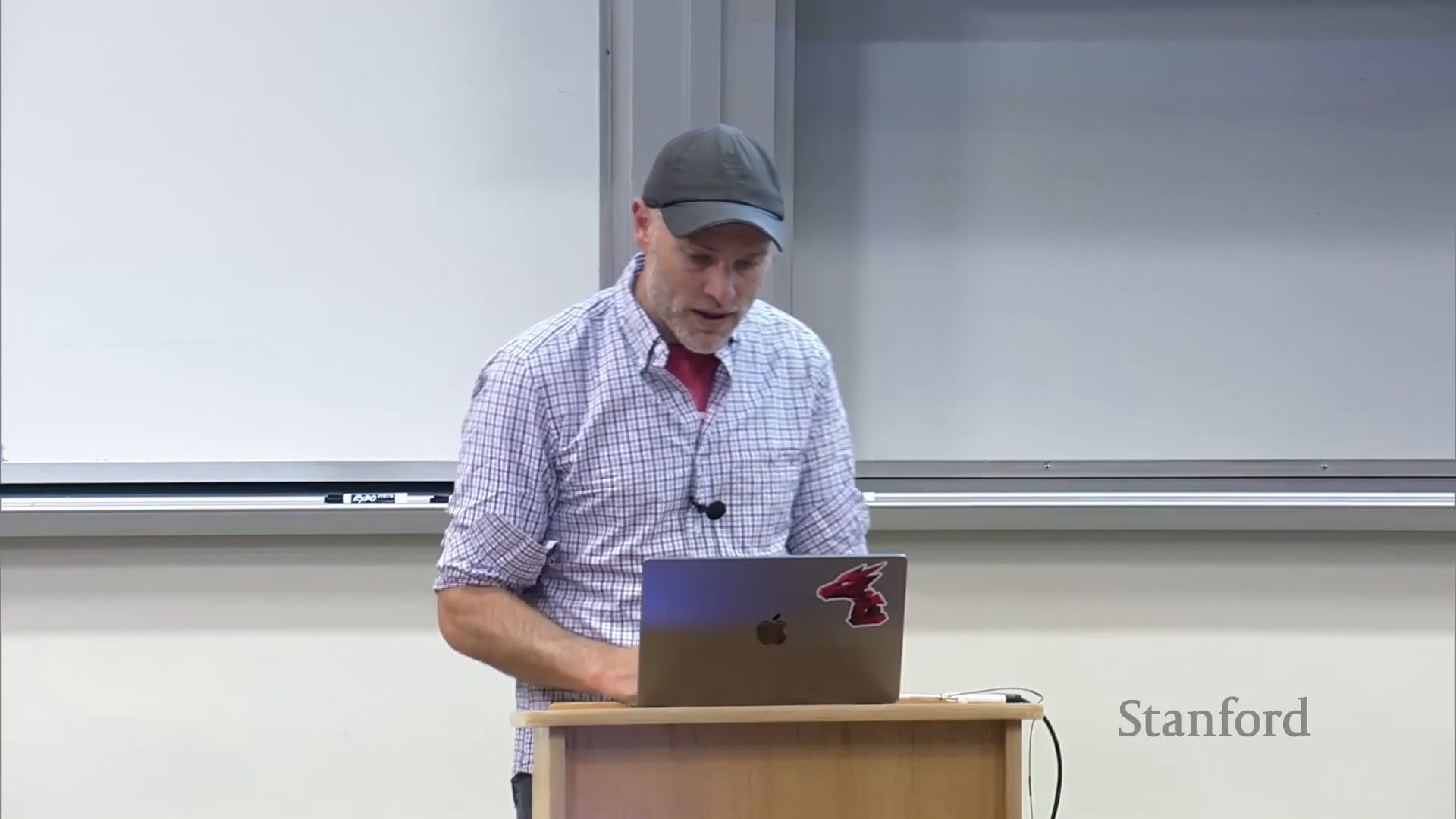

Stanford

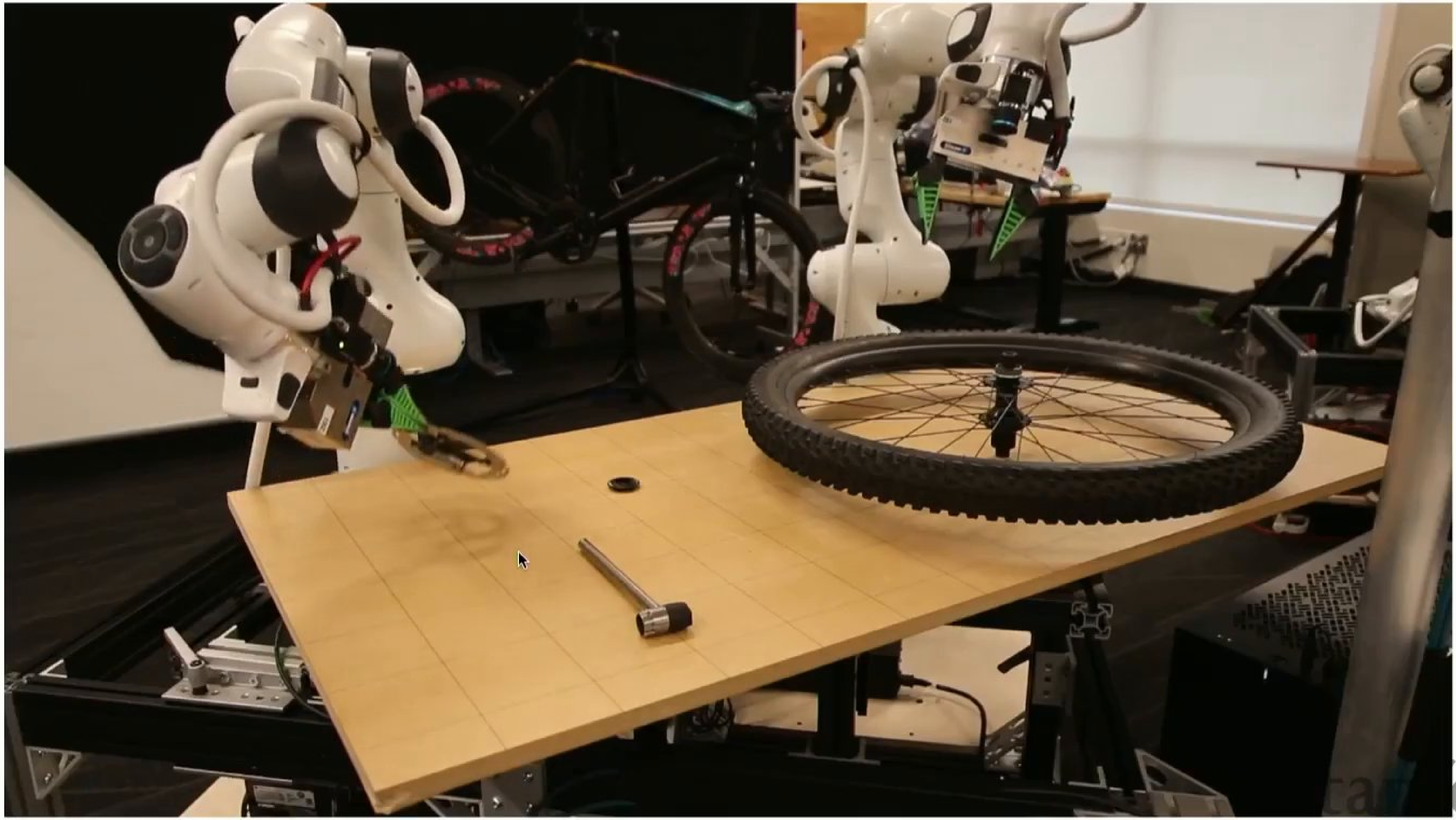

Stanford

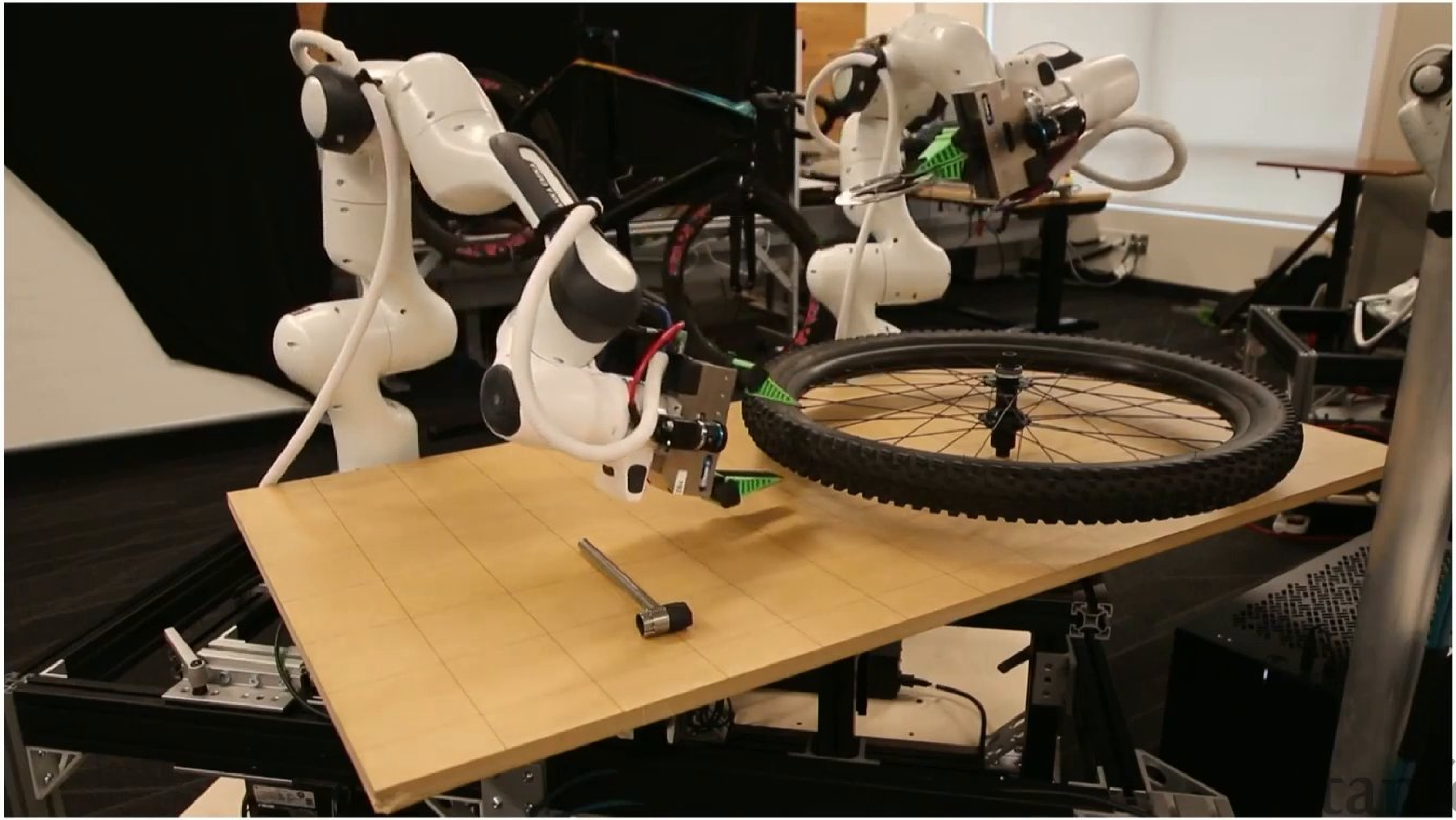

Stanford

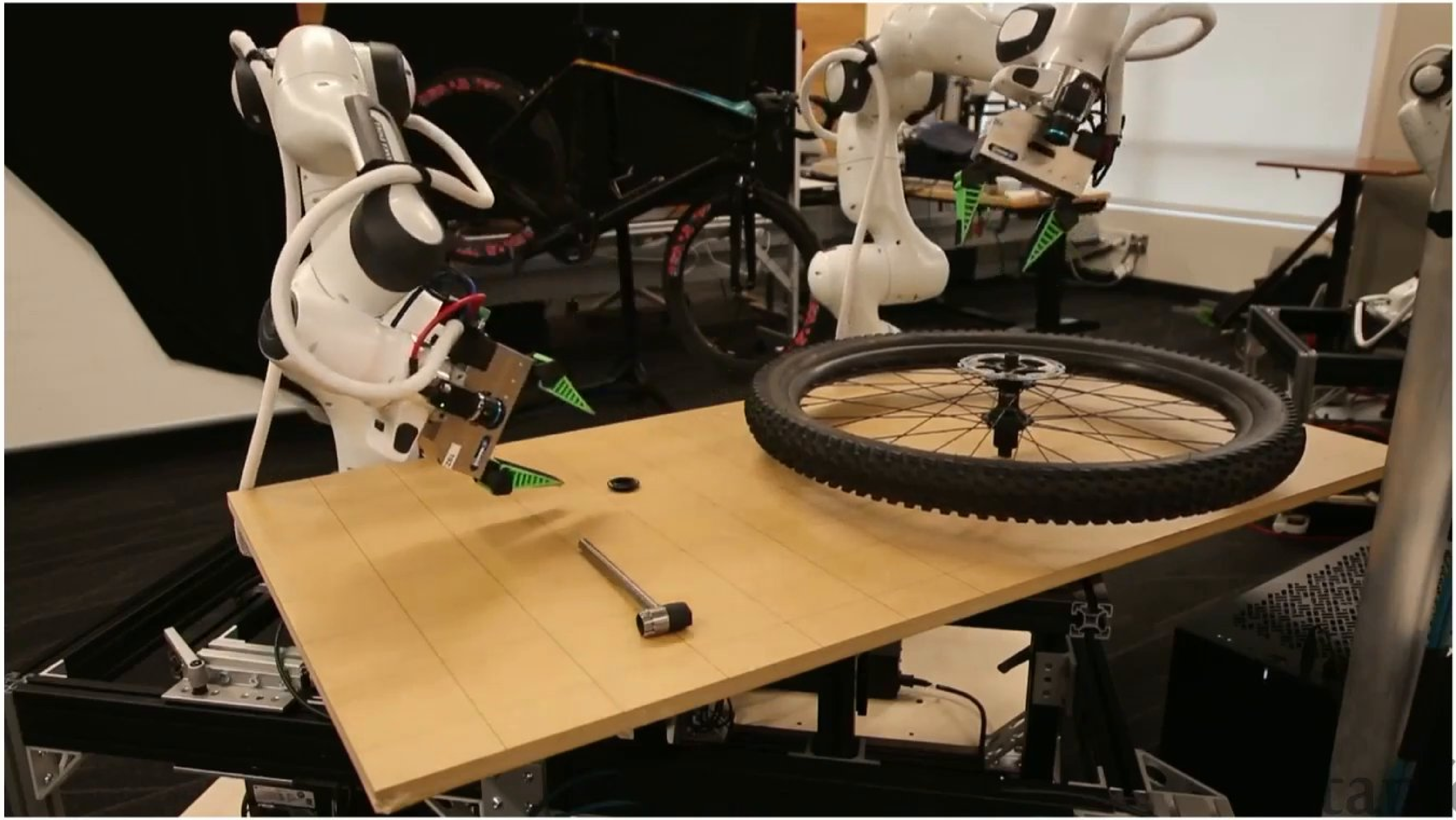

Stanford

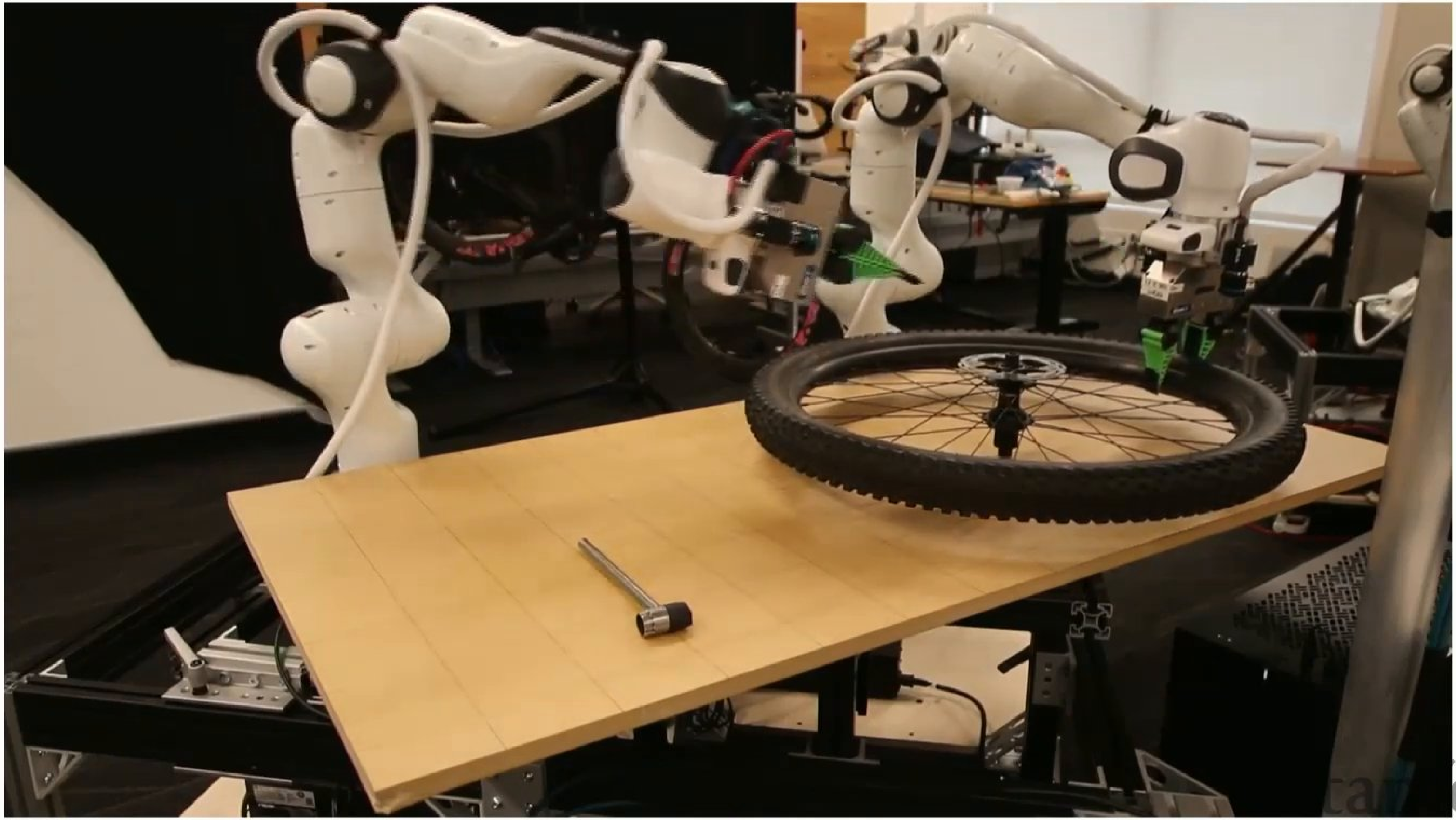

Stanford

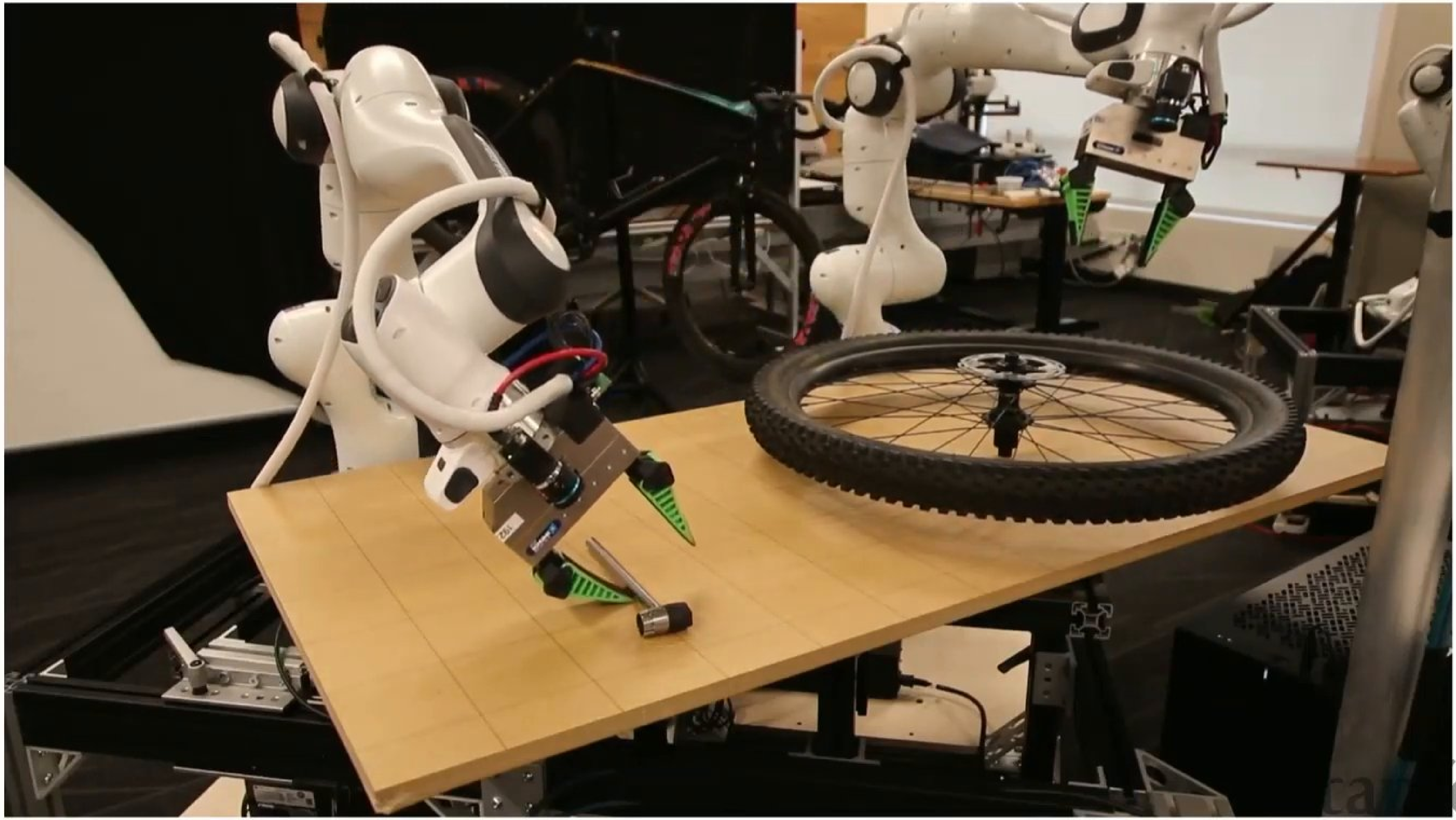

Stanford

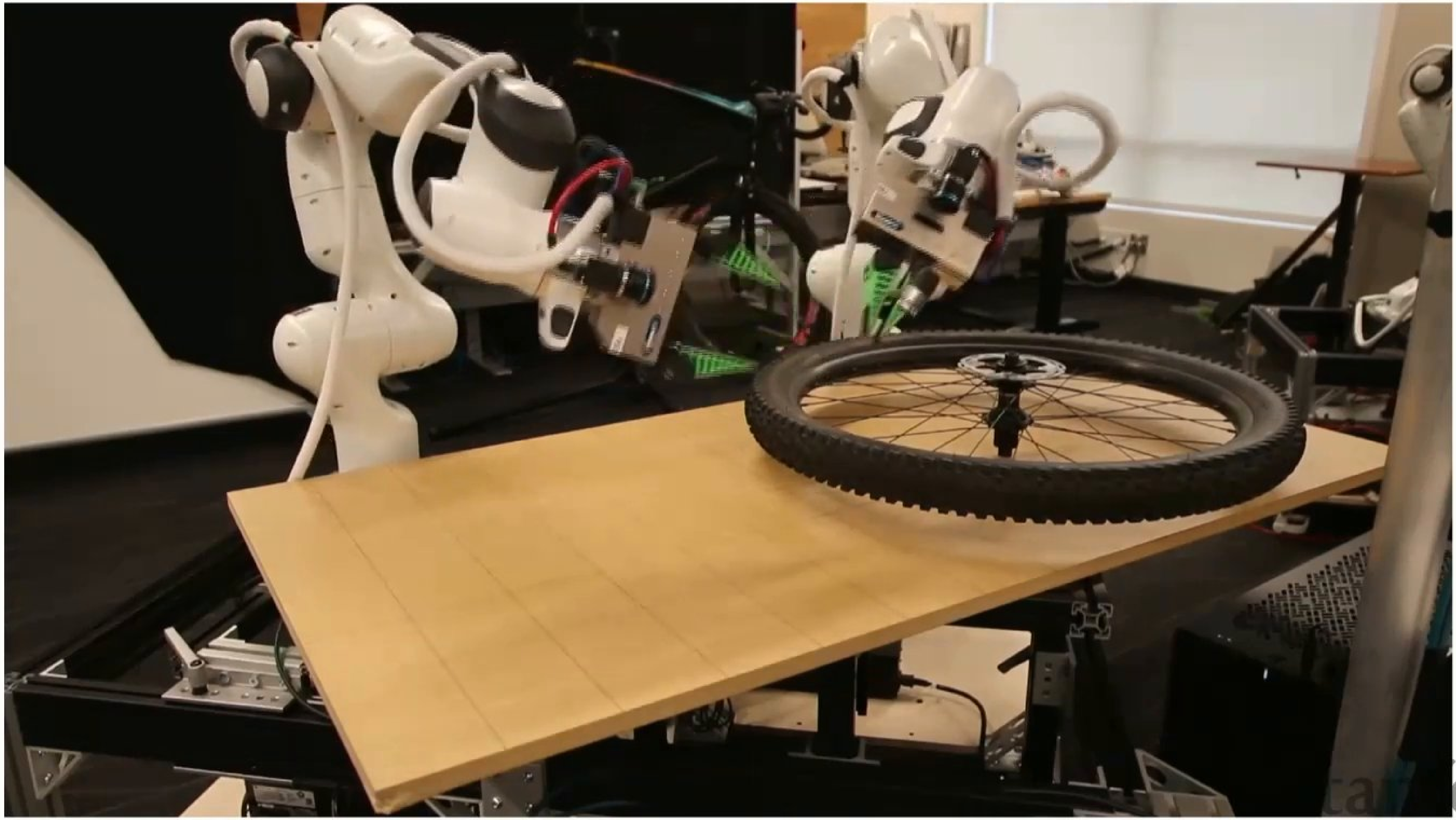


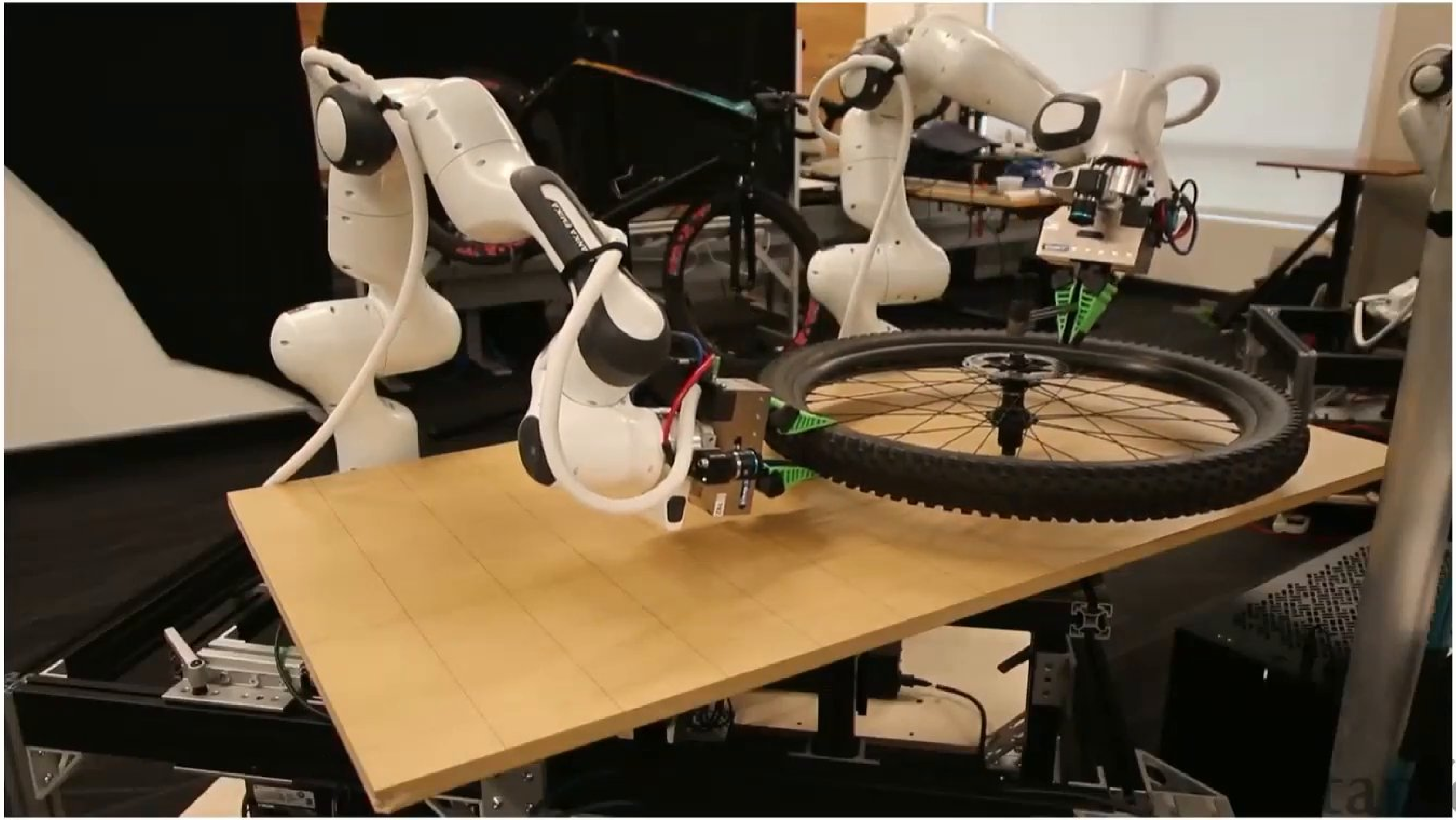




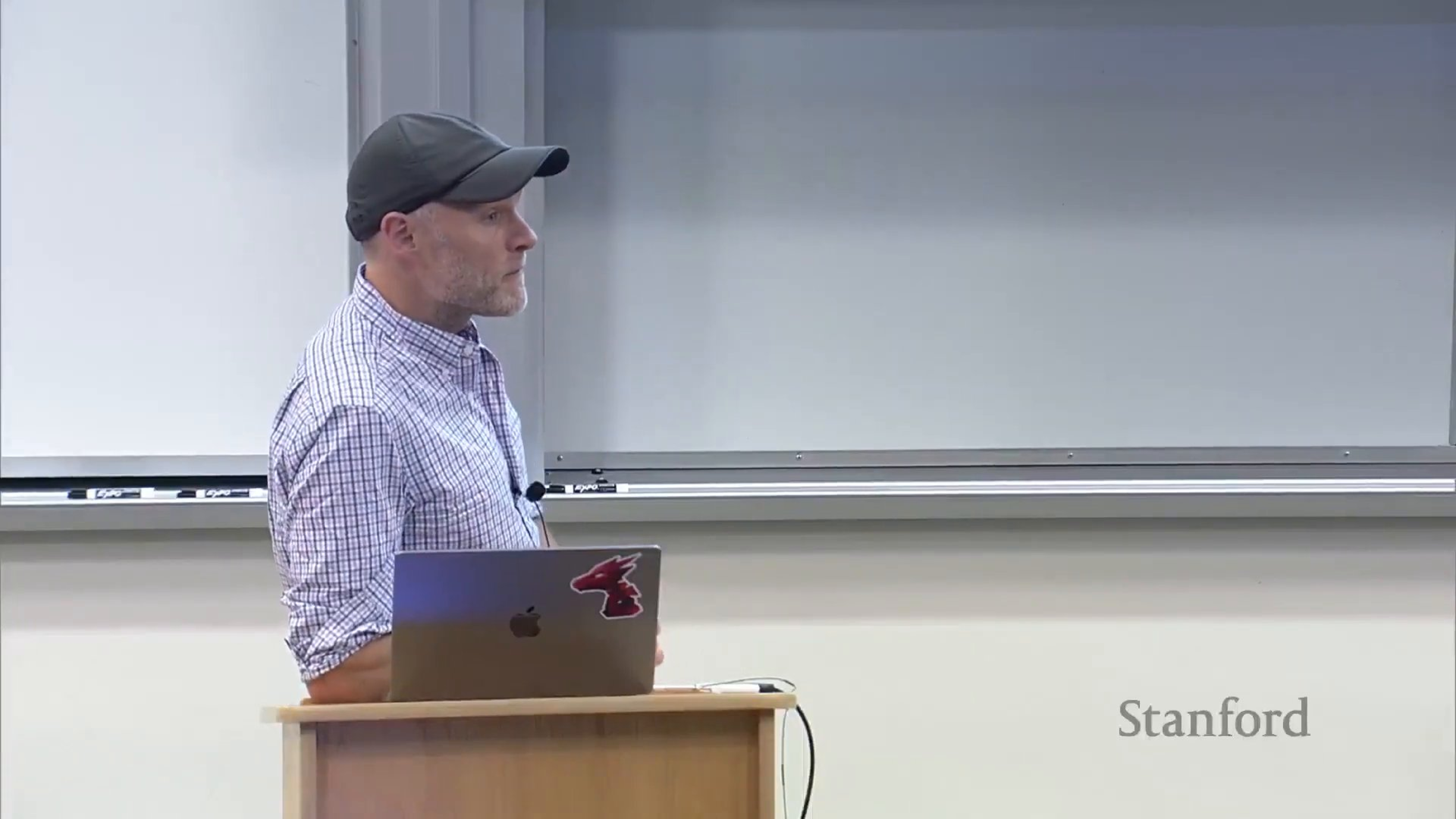


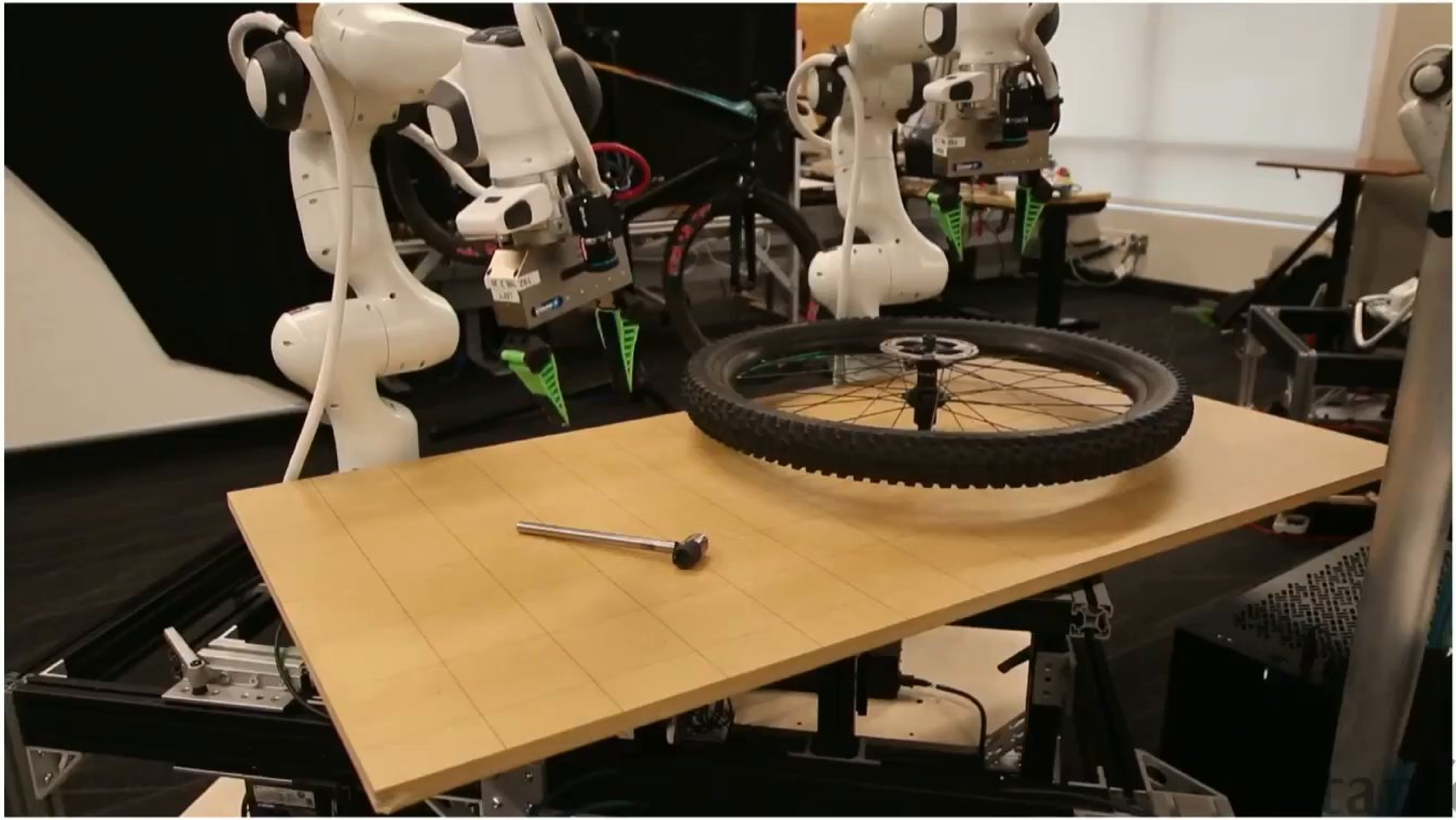


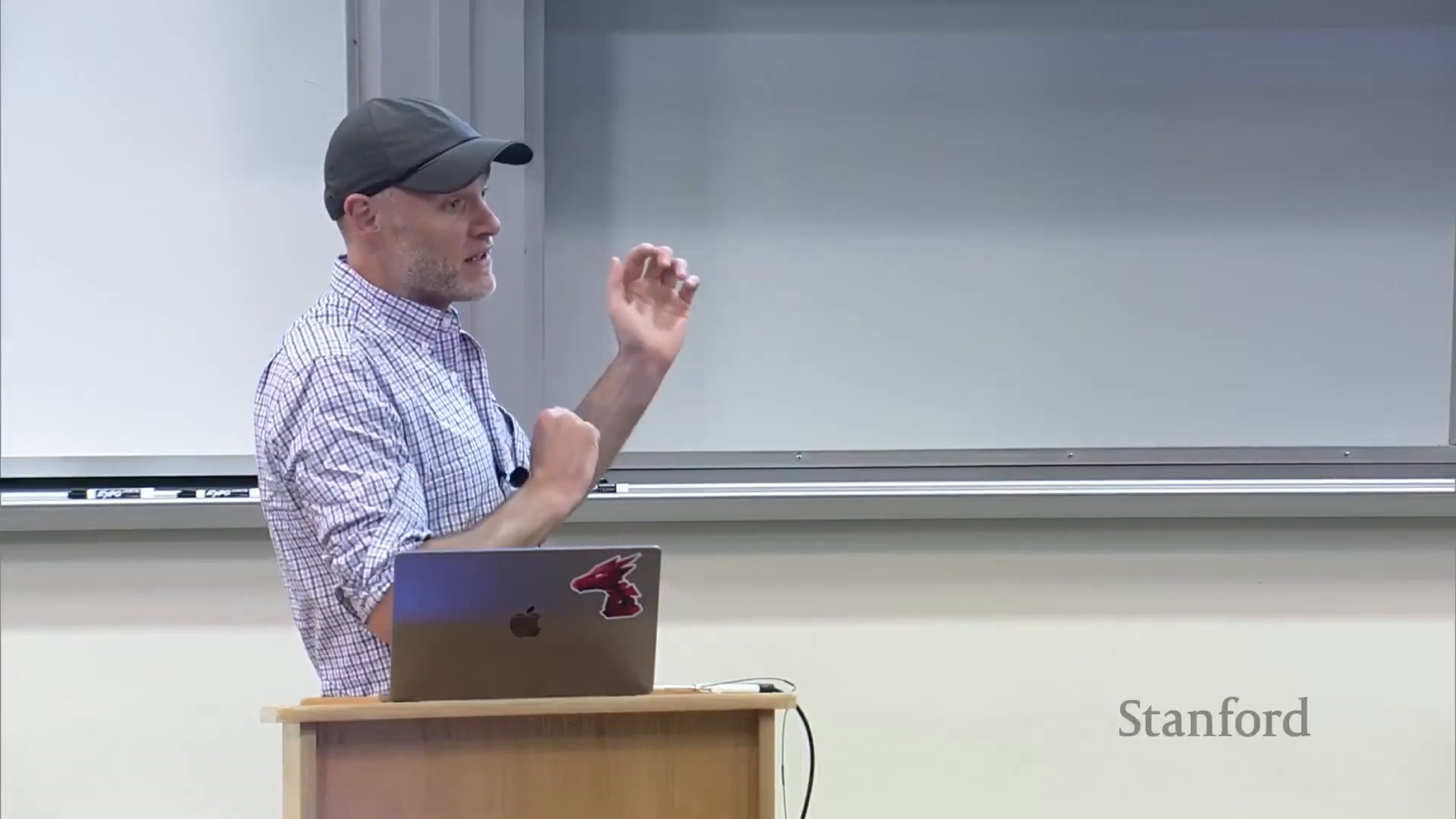


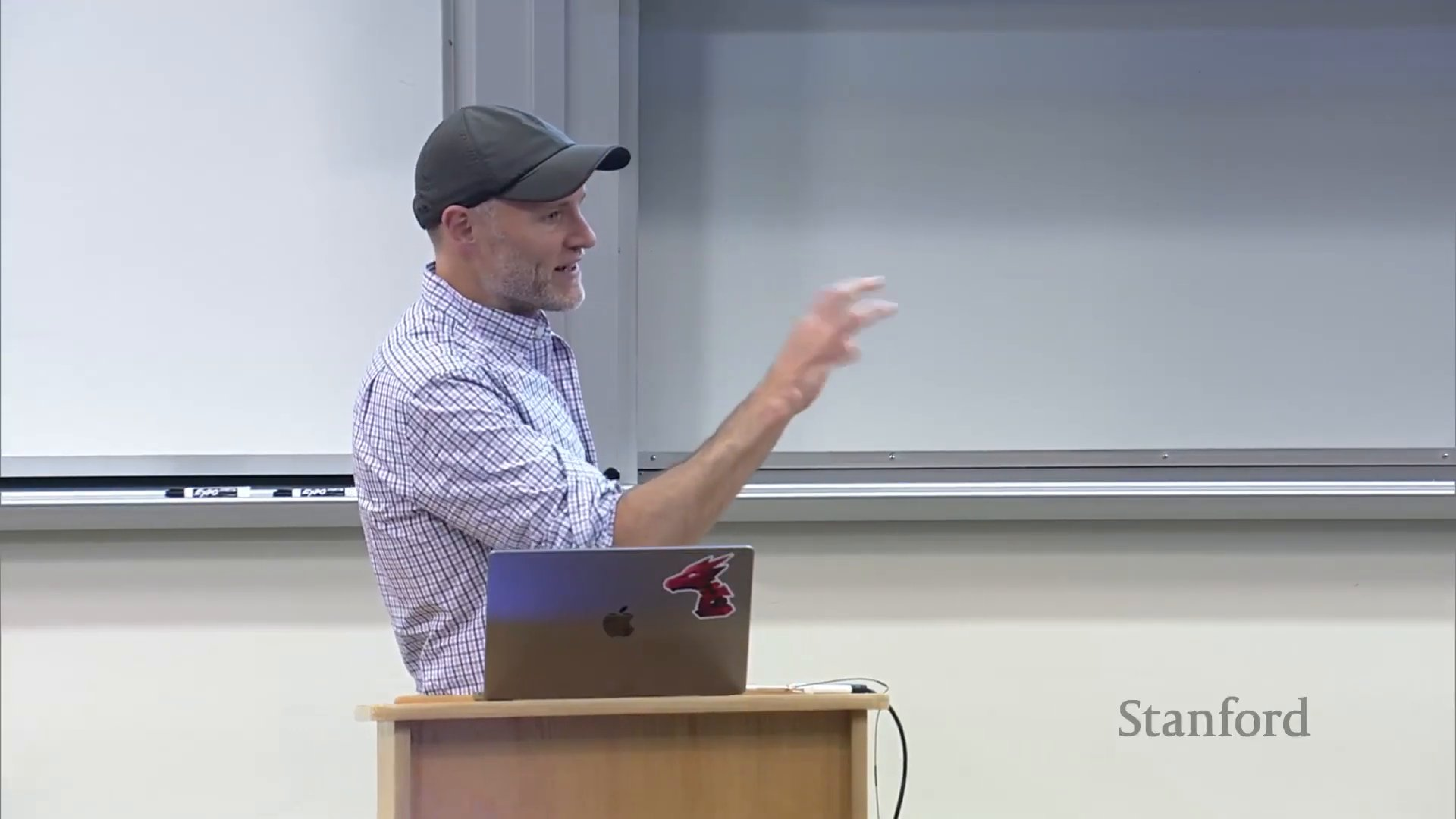



Stanford


Stanford




Stanford


Stanford

Connect with us:

online.stanford.edu

Stanford | ONLINE