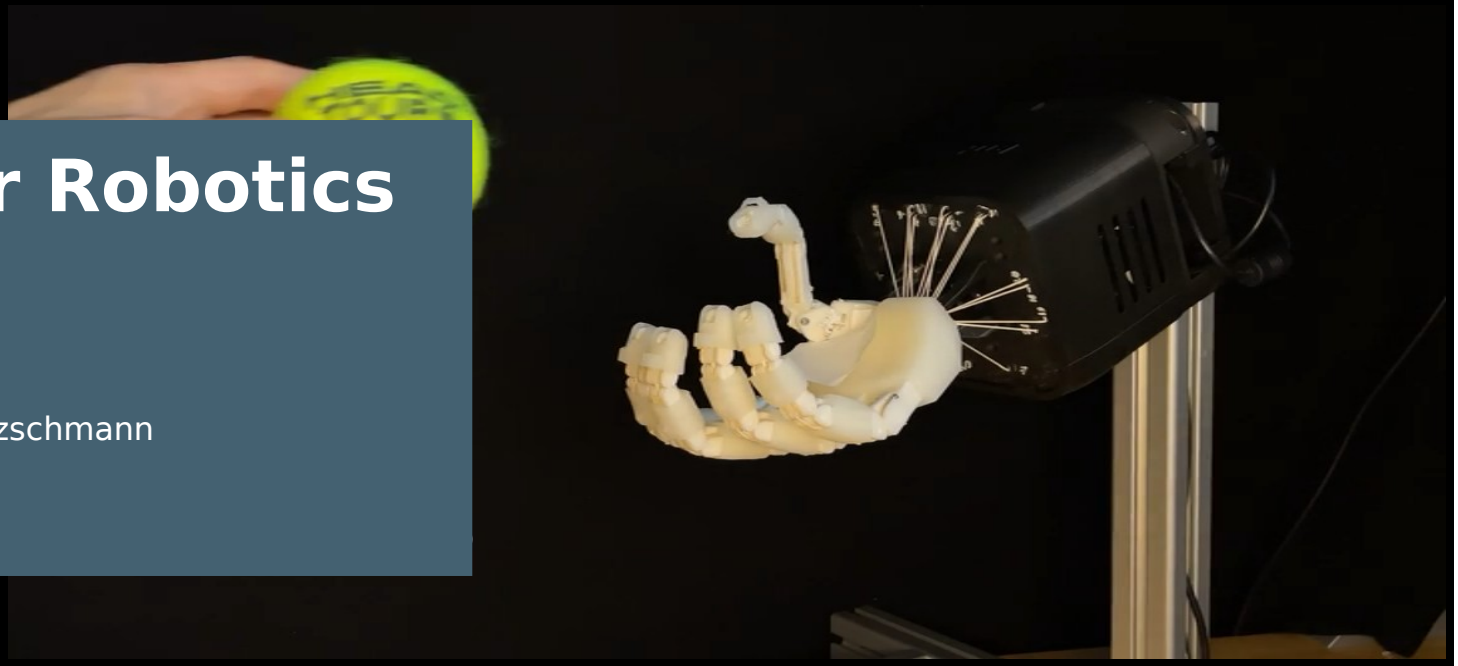
Data for Robotics

Davide Liconti
Prof. Dr. Robert Katzschmann

03-11-2025



Toshimitsu et al., Getting the ball rolling, Humanoids (2023)

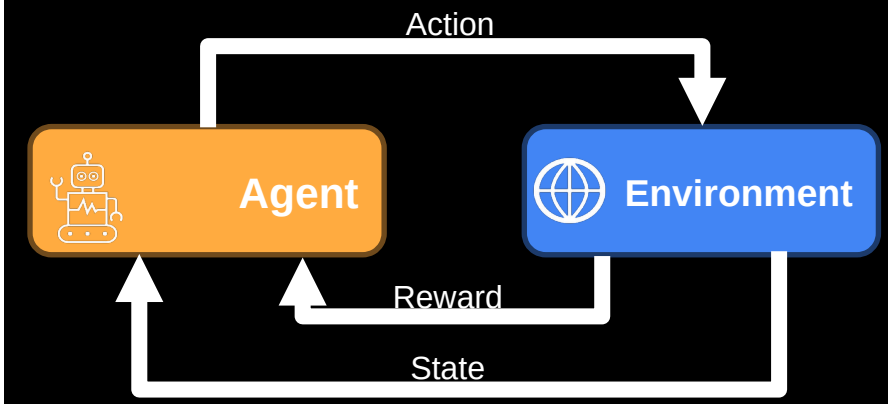
Agenda

1. Why do we need data
2. Data sources for Robotic Manipulation
3. Teleoperation
4. Human Data
5. Simulation Data
6. Alternative Approaches

Approaches to Autonomy

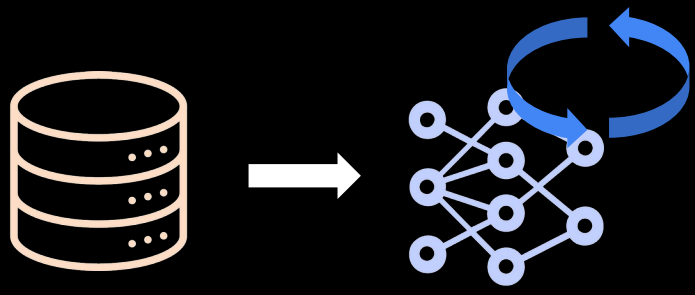
Reinforcement Learning

Learn by trial-and-error



Imitation Learning

Learn from Expert Demonstrations

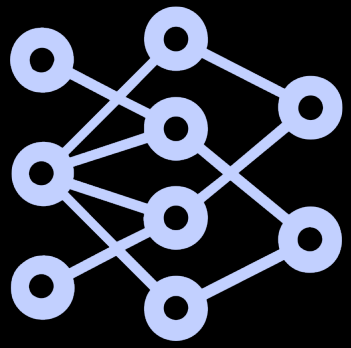


Dataset of demonstrations

Train a Policy (supervised)

Imitation Learning Preview

Input
Observations



Output
Actions

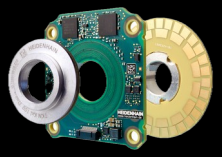
Imitation learning is a ***supervised learning*** approach where an agent learns to **perform tasks** or **acquire new skills** by **observing** and **mimicking** demonstrations provided by an expert.

How does this data look like?

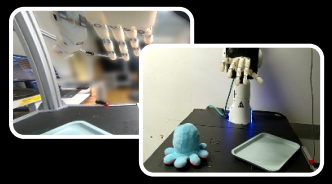
Robotics data is highly multimodal

Observations

Proprioception



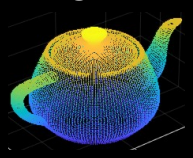
Wrist/Workspace RGBs



Language Command

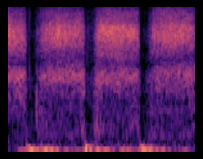
"put the plush on the tray"

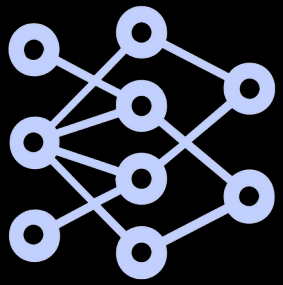
3D



Tactile Sensing

Audio





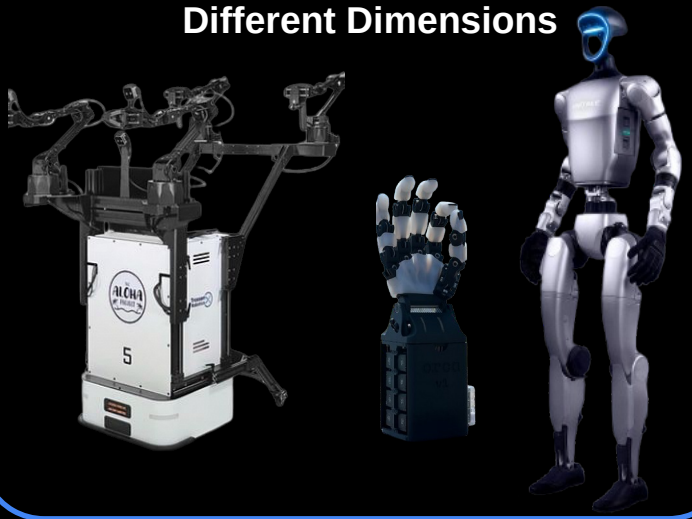
Actions

Different Meanings

end-effector space

Joint space

Different Dimensions



Robotics data are not much

Comparing language datasets (***GPT-2***, ***LLama3***)
to robotics dataset (***OXE***, ***π***) assuming 238 words/minute, 1.33 tokens/word

We lack the internet for robotics

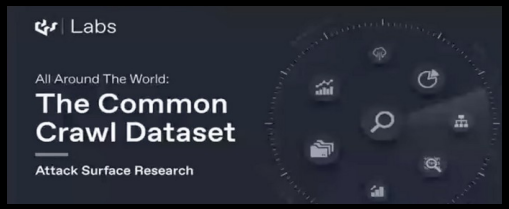
https://x.com/kvablack/status/1856373781603987655

Kevin Black (PI) CoRL 2024 workshop

Lack of data in robotics

“Large data solved Language,
Large data solved Vision,
Large data will solve Robotics"

When?



The Common Crawl data set contains approximately 6 billion web documents stored on a publicly accessible, scalable computer

A massive, open-source collection of 5.85 billion image-text pairs used to train large language and vision models

The Bitter Lesson

"Scaling computation beats human-designed knowledge in the long run"

http://www.incompleteideas.net/IncIdeas/BitterLesson.html[7]

Scaling Laws for LLMs

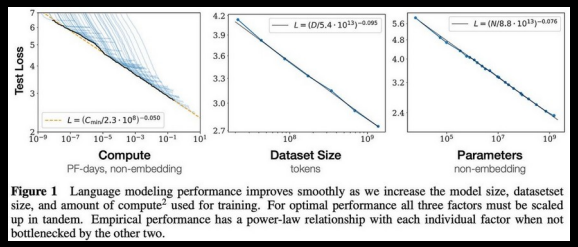


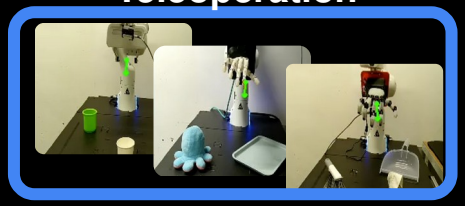
Figure 1 Language modeling performance improves smoothly as we increase the model size, datasetset size, and amount of compute[2] used for training. For optimal performance all three factors must be scaled up in tandem. Empirical performance has a power-law relationship with each individual factor when not bottlenecked by the other two.

https://arxiv.org/pdf/2001.08361

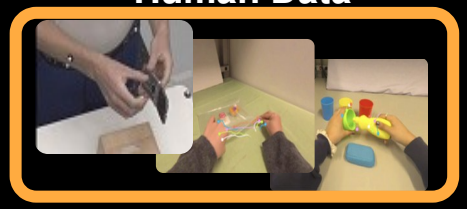
ETH zürich SoftRobotics Laboratory

Data Sources for Robotic Manipulation

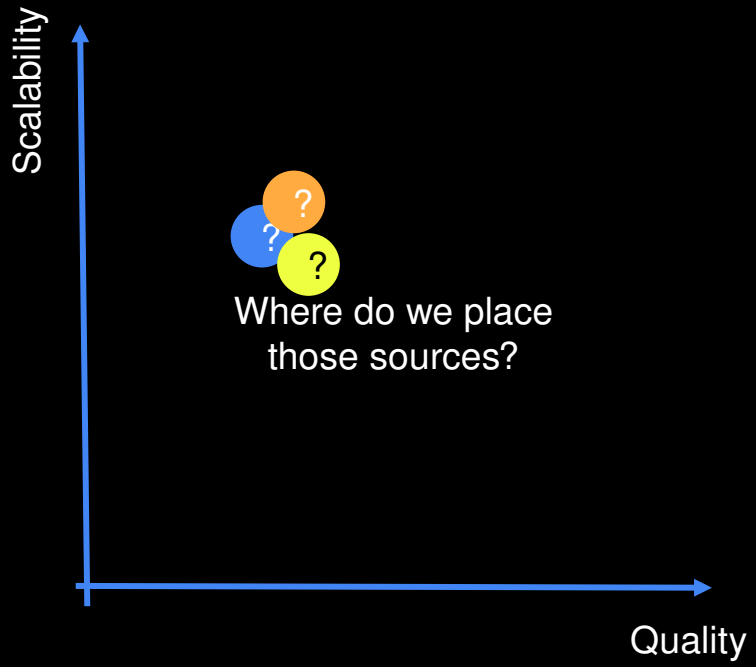
Teleoperation



Human Data



Simulation



ETH zürich SoftRobotics Laboratory

Teleoperation

Teleoperation is the remote control of a machine or robotic system, allowing a human operator to perform tasks from a distance by sending commands and receiving feedback like video and sensor data

Requirements

- **Real-time control:** low latency (<100 ms) for stable teleop.
- **Feedback loop:** The operator need to have as many information as the robot and environment states

Types

Direct Teleoperation

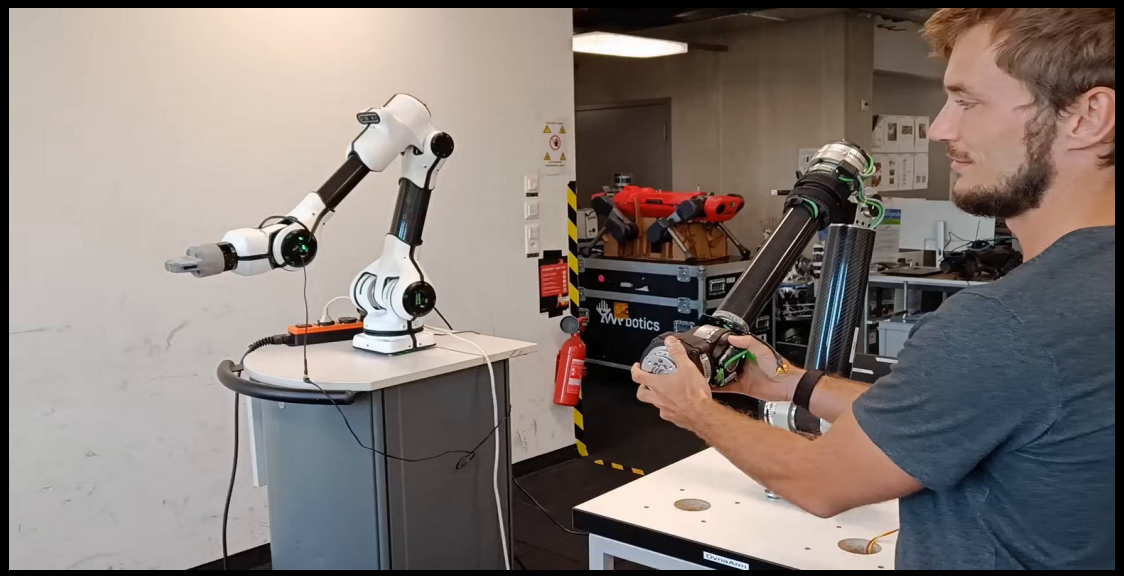
Indirect Teleoperation

High Level Teleoperation

Teleoperation

Direct Teleoperation

Leader-Follower systems, directly send joint values to the other robot



ETH zürich Soft Robotics Laboratory

Duatic AG
https://duatic.com/

Teleoperation

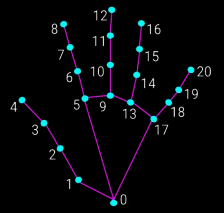
Indirect Teleoperation

Control the robot with another system, and retarget to its joint values

Bimanual Teleoperation with Orca Hands

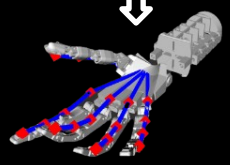
Teleoperation

Indirect Teleoperation



Wrist relative pose

⇩



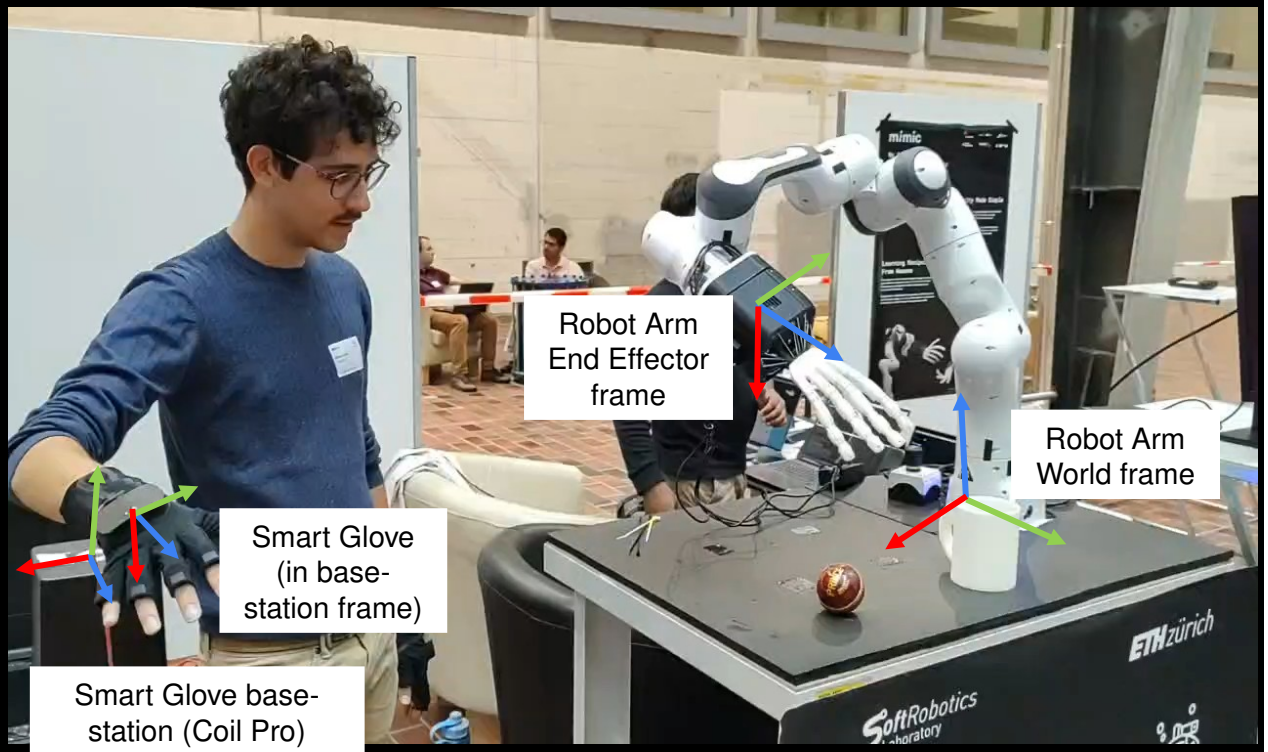
IK

⇩

Hand joints

⇩

Robot joints



ETH zürich SoftRobotics Laboratory

Teleoperation

High Level Teleoperation

Human gives *goal- or task-level commands* (e.g. keypoints), robot handles low-level control.

Cybathlon 2024 Competition
https://cybathlon.com/en/events/edition/cybathlon-2024

Teleoperation

Haptics

Including tactile and kinesthetic feedback for more accurate and intuitive control

https://haptx.com

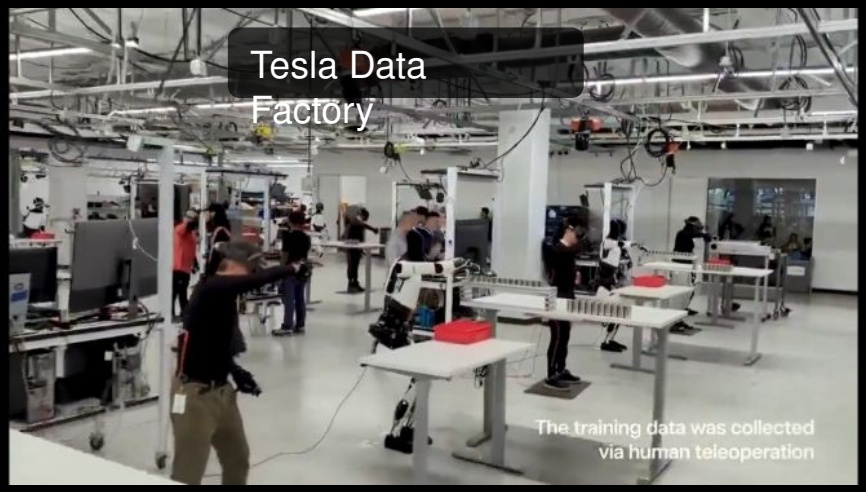
Teleoperation

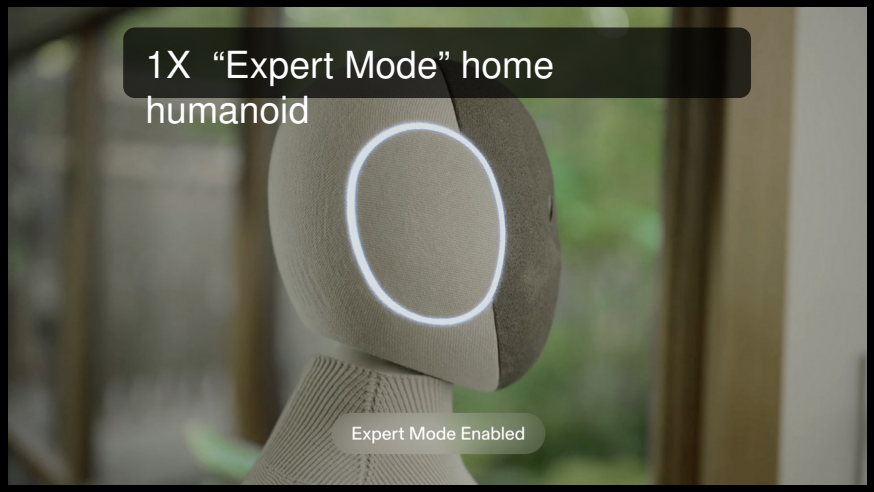
Competitions

ANA Avatar XPRIZE Finals: Winning Team NimbRo

Teleoperation

Real world teleoperation data are the best data you can get for your specific task. **No embodiment gap** but is **expensive** to get.





RealMan launches robotics data training center in Beijing

By The Robot Report Staff | • August 29, 2025

DYNA Dyna Robotics

Data Collection Operator - Night Shift

Redwood City, CA · 1 mese fa

Risposte gestite esternamente a LinkedIn

In sede · Temporaneo · Esperienza minima

Teleoperation as Teleoperation

For some applications, teleoperation is not a mean to collect data, but rather the end goal itself. That is when you want **full control** in dangerous or critical environments

DaVinci Surgical Robot

Teleoperation as Teleoperation

For some applications, teleoperation is not a mean to collect data, but rather the end goal itself. That is when you want **full control** in dangerous or critical environments

Human Data

Advantages

- **Scale** – cheap to get.

Limitations

- **Unlabeled** – they don't come with actions need hand pose estimation
- Bridging the **human–robot embodiment gap** is still an open problem.

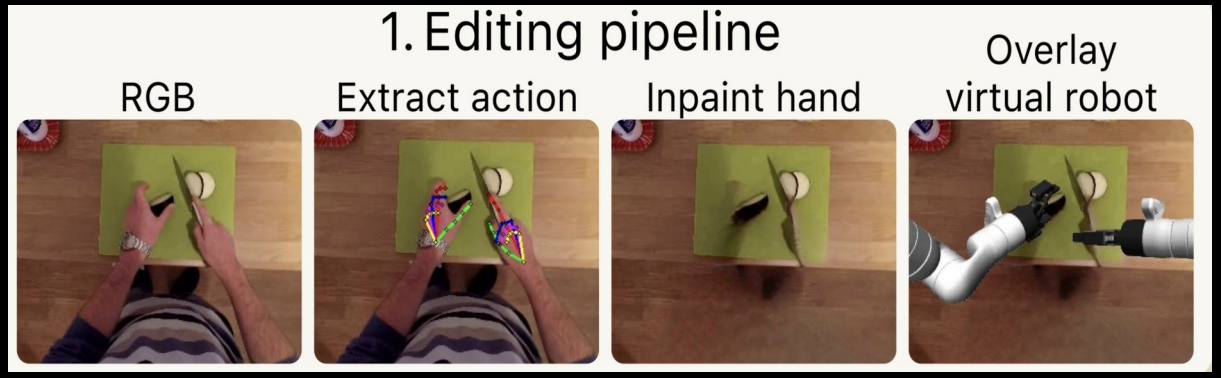
Ego4D dataset

Human Data

How to exploit human data for robotics applications?

Explicit Approaches

Edit human videos with inpainting and overlay robot



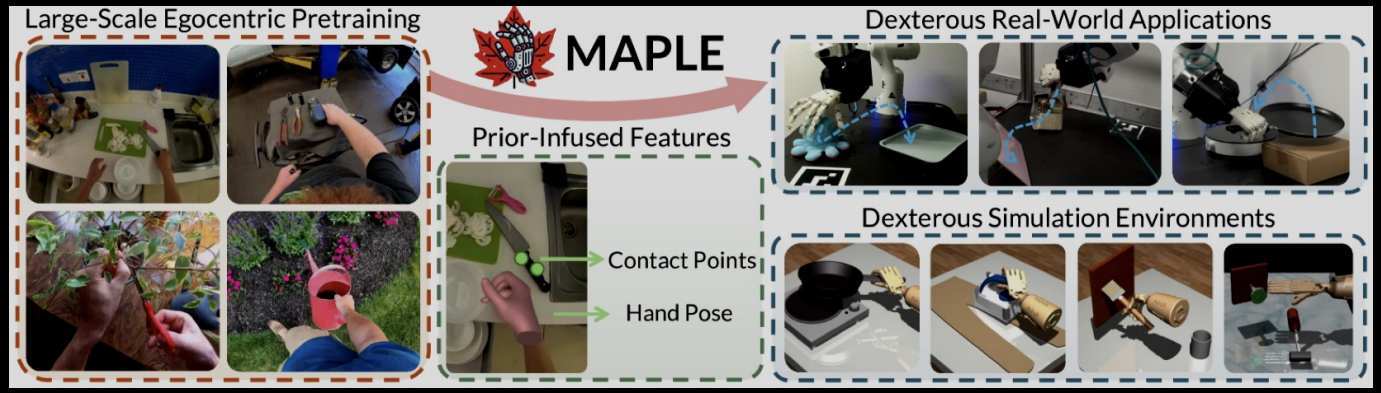
Masquerade: Learning from In-the-wild Human Videos using Data-Editing
https://arxiv.org/pdf/2508.09976

Human Data

How to exploit human data for robotics applications?

Implicit Approaches

Learn features from dexterous manipulation priors from egocentric videos for downstream dexterous robotic tasks



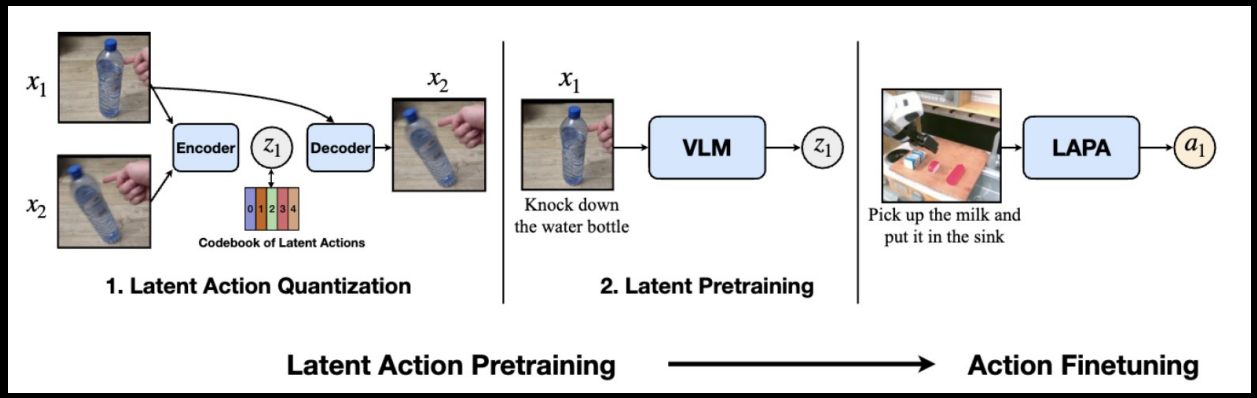
MAPLE: Encoding Dexterous Robotic Manipulation Priors Learned From Egocentric Videos
https://arxiv.org/html/2504.06084v1

Human Data

How to exploit human data for robotics applications?

Implicit Approaches

Extract pseudo latent actions from large scale human videos



LAPA: Latent Action Pretraining from Videos
https://arxiv.org/pdf/2410.11758

Simulation Data

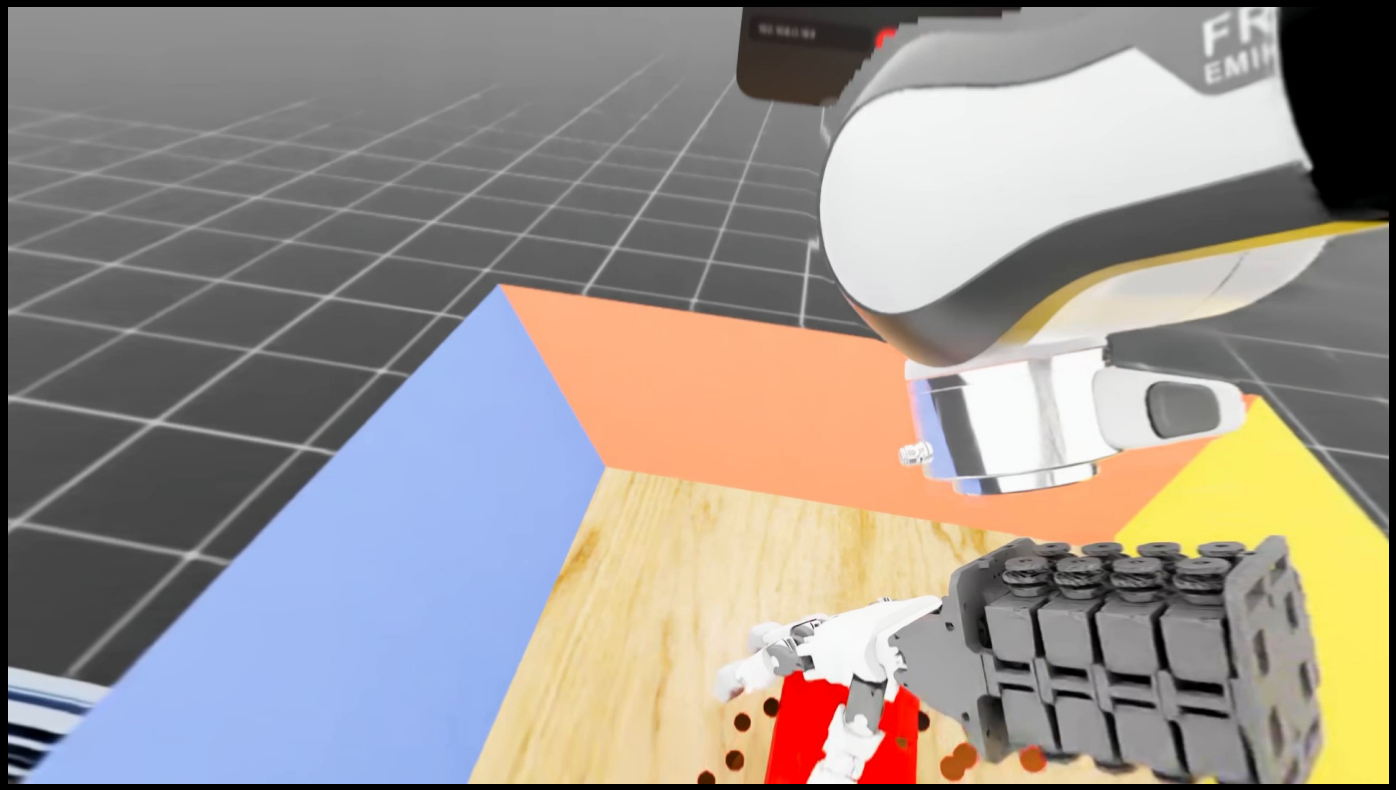
Advantages

- Can simulate parallel environments, the limit is not time but rather compute. (Easy to simulate augmentation, both for trajectories and visuals)
- Recent rendering techniques made them also suitable to use them for visuomotor policies (e.g. Gaussian Splatting)
- Don’t need physical access to robots, can collect data anywhere.

Limitations

- Bridging the **sim to real gap** for both physics and renderings

Simulation Data



Simulation Data

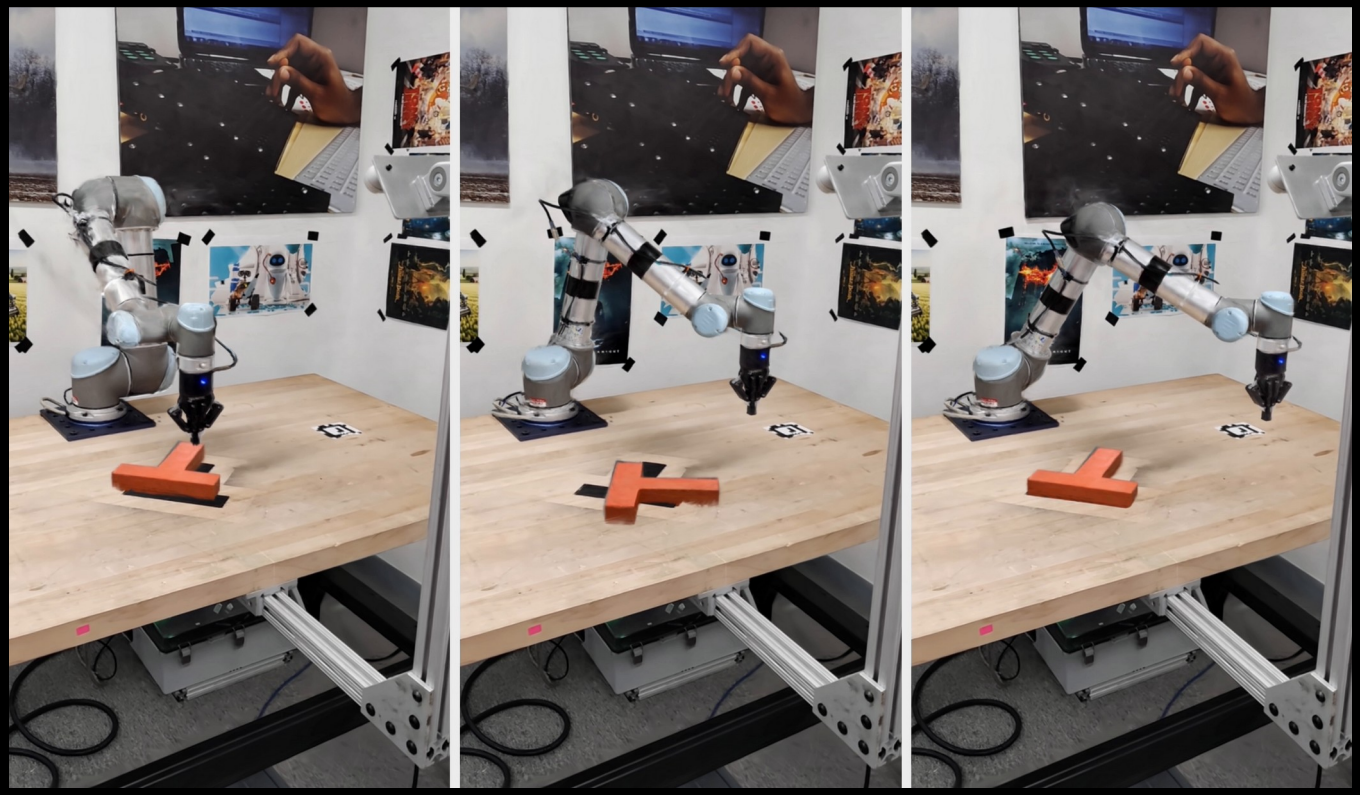
- Don’t need physical access to robots, can collect data anywhere.

DART: Dexterous Augmented Reality Teleoperation Platform for Large-Scale Robot Data Collection in Simulation
https://openreview.net/pdf/b0393e5ce2ed5eab3b89e05d5e6855b36ef4875f.pdf

ETH zürich SoftRobotics Laboratory

Simulation Data

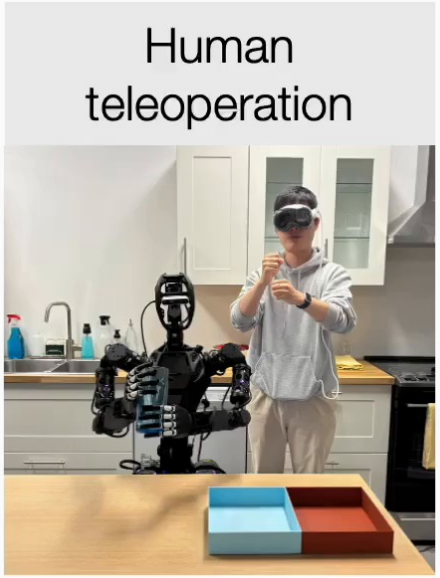
With recent neural rendering techniques, simulations are photorealistic



SplatSim https://splatsim.github.io

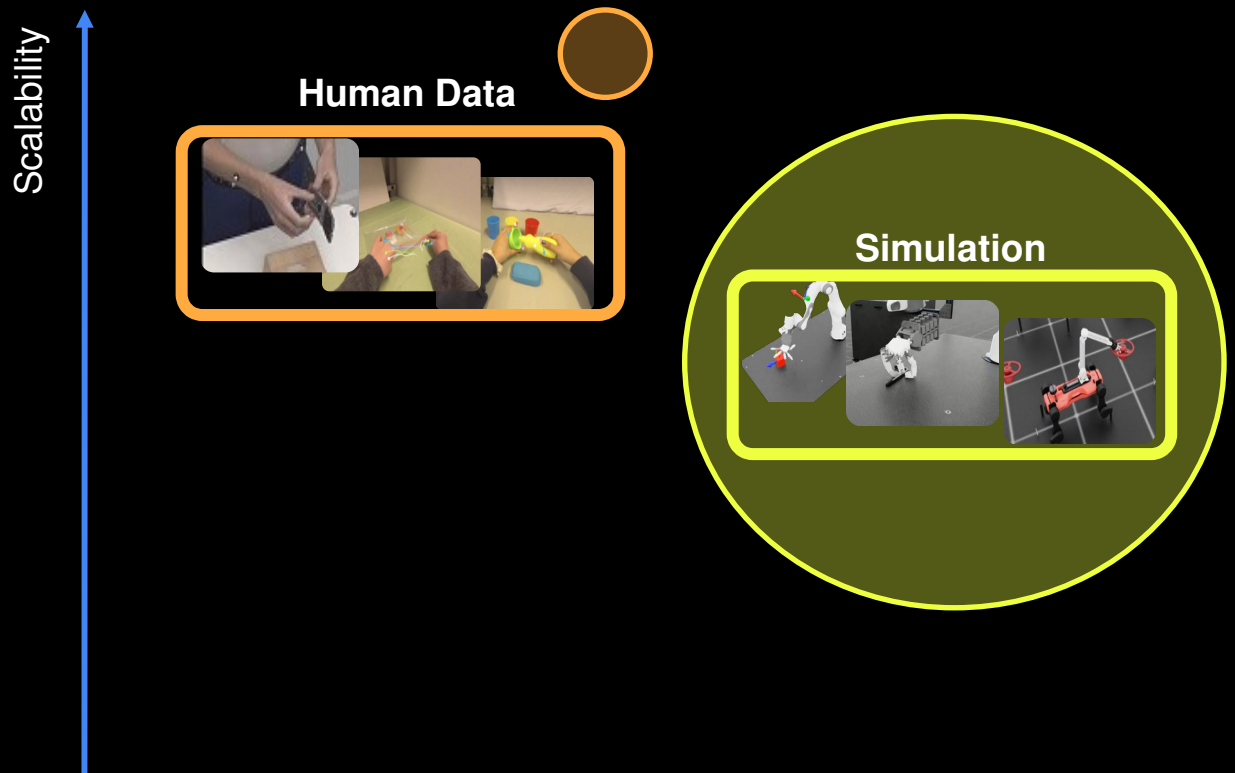
Simulation Data

Simulation can be used for data augmentation



DexMimicGen: Automated Data Generation for Bimanual Dexterous Manipulation via Imitation Learning
https://arxiv.org/pdf/2410.24185

Data Sources for Robotic Manipulation



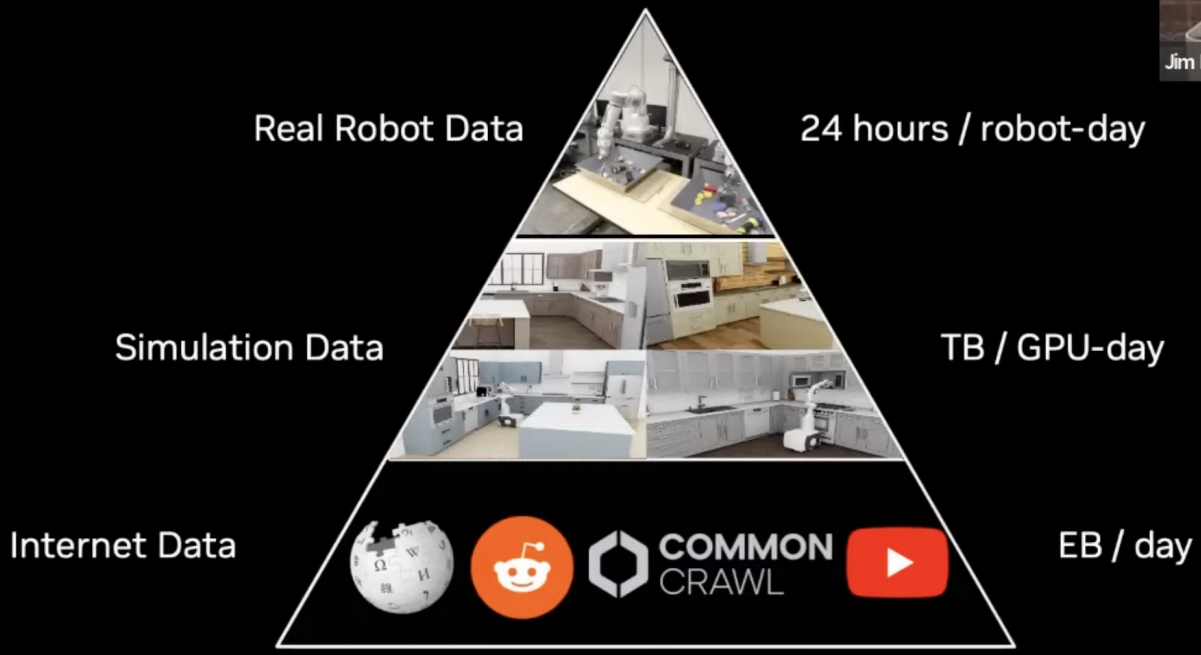
Alternative Approaches: Hybrid

UMI employs hand-held grippers coupled with careful interface design to enable portable, low-cost, and information-rich data collection for challenging bimanual and dynamic manipulation demonstrations.

ETH zürich Soft Robotics Laboratory

Universal Manipulation Interface
https://umi-gripper.github.io/

Others call it “Data Pyramid”



Data Pyramid

Berkel
UNIVERSITY OF CALIFOR
Powered by

CS 194/294-196 (LLM Agents) - Lecture 9, Jim Fan

ETH zürich SoftRobotics Laboratory

Timestamp: 00:16:00

Key Takeaways

- Scaling data was key for language and vision
- Data in robotics manipulation is highly multimodal and diverse
- Teleoperation is very valuable source of data but expensive
- More scalable alternatives are human and simulation data, but are **harder** to use