Lecture 26

Robot Learning II

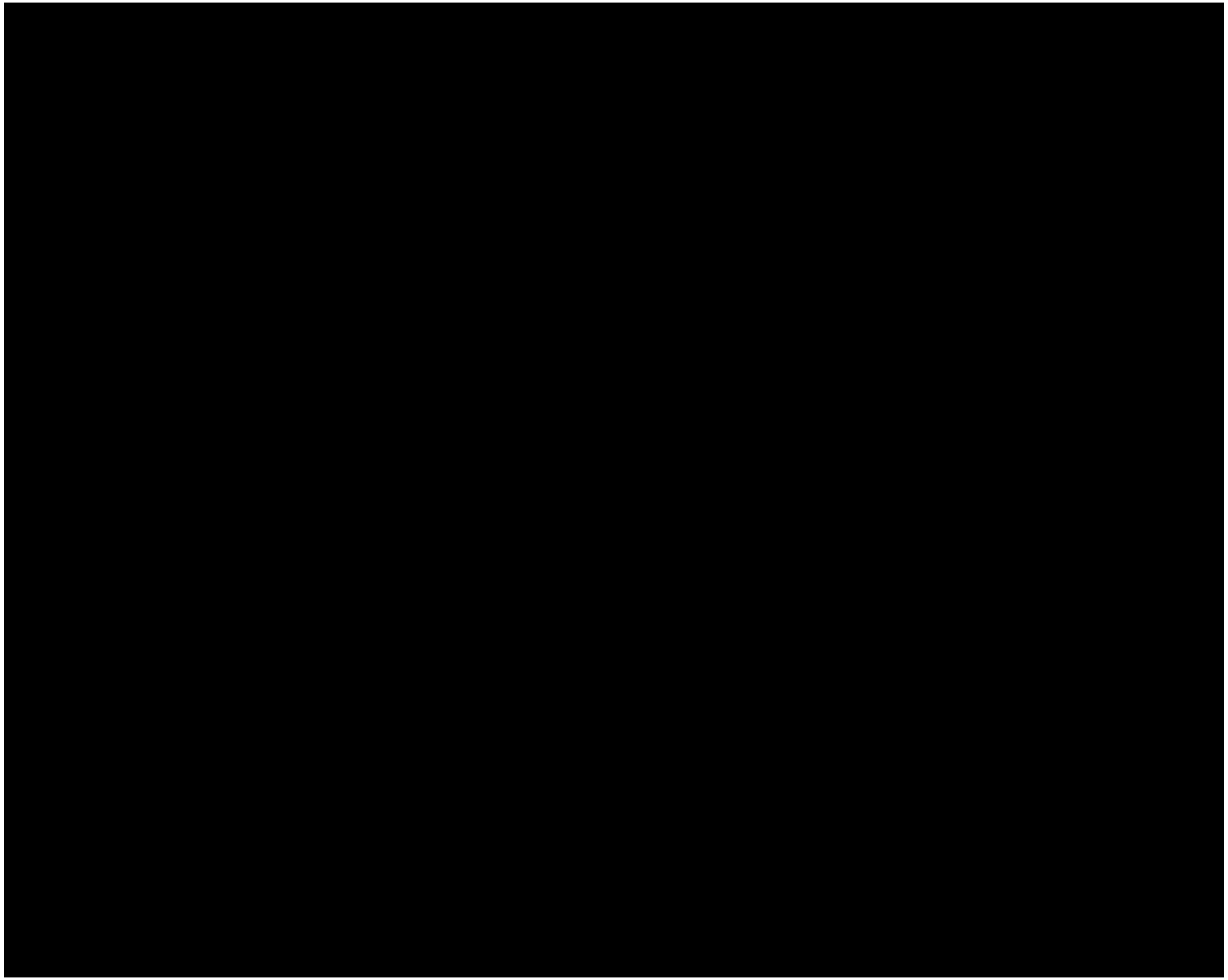
CS 3630

BIG question: How do we develop robots that operate effectively in the real world?

ACE: Autonomous City Explorer (2007!)

We can ENGINEER systems.

This example of a capable robot is based on a software architecture made up entirely of components we covered in this class.



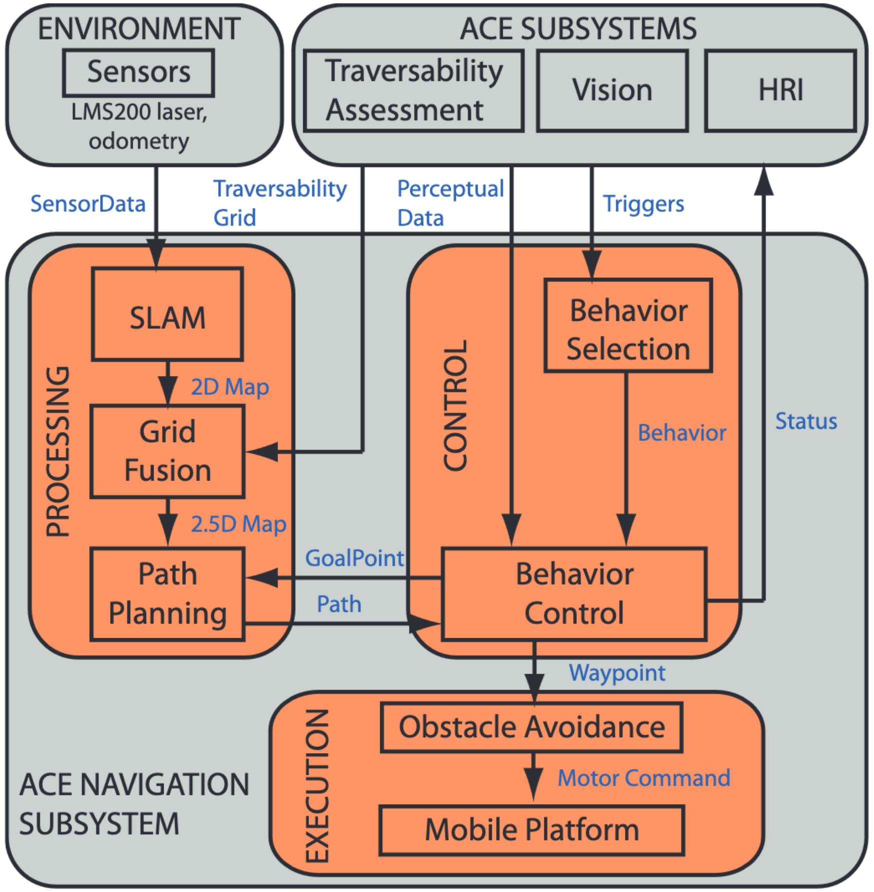
- https://www.youtube.com/watch?v=j-rwontvero

*The robot took so long because people kept interrupting it!

ACE: Autonomous City Explorer (2007!)

We can ENGINEER systems.

This example of a capable robot is based on a software architecture made up entirely of components we covered in this class.



- https://www.youtube.com/watch?v=j-rwontvero

Deep Whole-Body Control: End-to-End Learning in Legged Manipulator Robots, 2022.

We can develop robots that LEARN.

https://www.youtube.com/watch?v=zhpI3JTqQGI

ENGINEERING

- Greater predictability
- More performance guarantees
- Easier to debug & upgrade

- Limited by human ingenuity

LEARNING

- Greater generalization
- More performance optimization

- No verifiability
- Less predictability

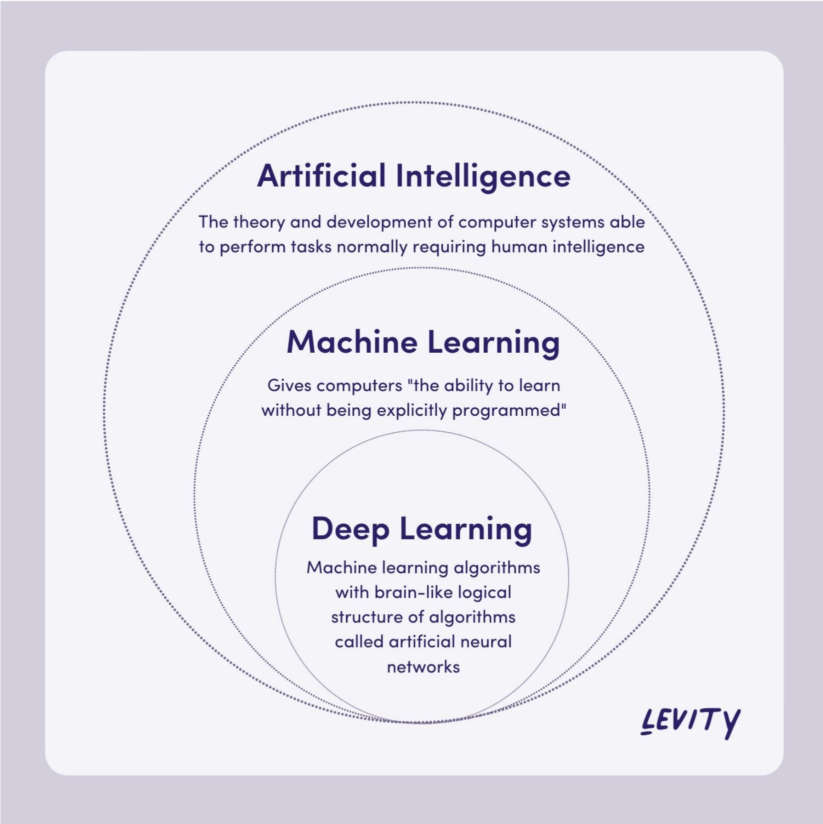
The two major types of learning

- **Imitation Learning** => Learning to imitate expert
  - *Behavioral cloning*: learn the policy directly
  - *Inverse reinforcement learning*: learn the reward
- **Reinforcement learning** => Learning from experience
  - *Model-based RL:* Learn/use a process/world model to predict interactions
  - *Model-free RL*: Just same sufficient interactions without a model

What is deep learning, then?

- Deep learning is a field of study that focuses on **large neural models** that can approximate complex functions

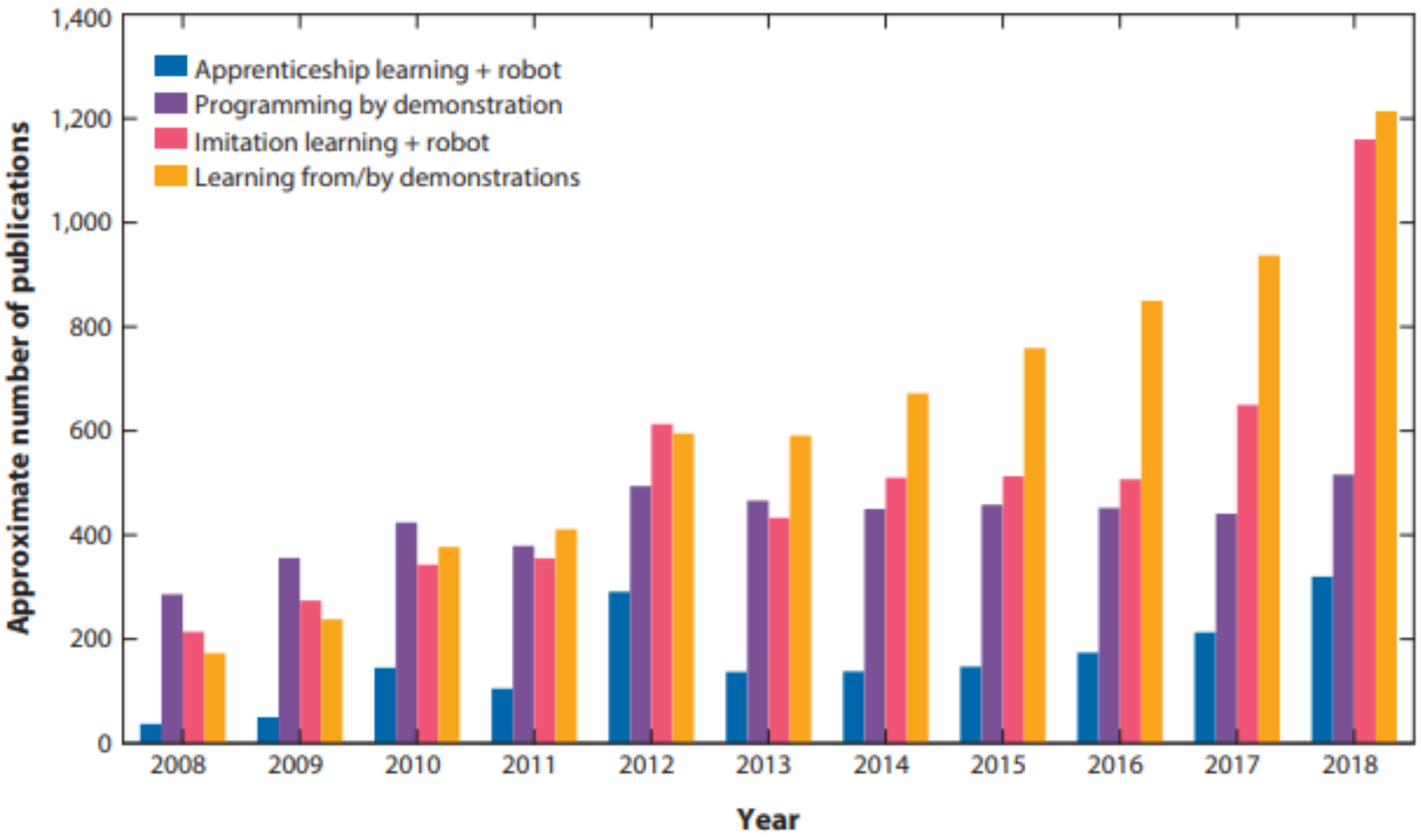
- Deep learning can be applied to both imitation learning and reinforcement learning



Imitation learning

- Given expert demonstrations, figure out how to imitate the behavior
- Often assume optimal expert
  - Exceptions exist that allow suboptimal demonstrations
- Turns the problem into a supervised learning problem
  - Minimize reproduction error
- Major challenges
  - Distribution shift
  - Correspondence problem (or embodiment mismatch)

The growth of IL-based methods in robotics



Reinforcement learning

- How to learn from experience?
- Inspired by the brain
- Neither supervised nor unsupervised
  - You assume you have a reward function
- Sequential decision making under uncertainty
  - Instantaneous version of RL: Bandit algorithms
- Markov decision process (MDP)
  - States, actions, state transition function, rewards, discount facto

Exploration vs Exploitation in RL

- Exploration: Take chances and find out what is effective
- Exploitation: Do what you know is relatively effective

- How to effectively balance?
  - A simple strategy: the Epsilon-greedy algorithm

Model-based vs Model-free

- **Model-based**: We have a model of the environment (i.e., the state-transition function)
  - We can more efficiently learn since we know what will happen *if* we take an action (i.e., predictive power)
  - Many methods first learn this model and then a policy based on this model

- **Model-free**: We have no such model!
  - Interact with the environment to “discover” what happens if you take an action
  - Cannot predict
  - More useful in complex domains

On-policy vs. Off-policy in RL

- **On-policy**: update the value function based on the policy that you are learning or optimizing

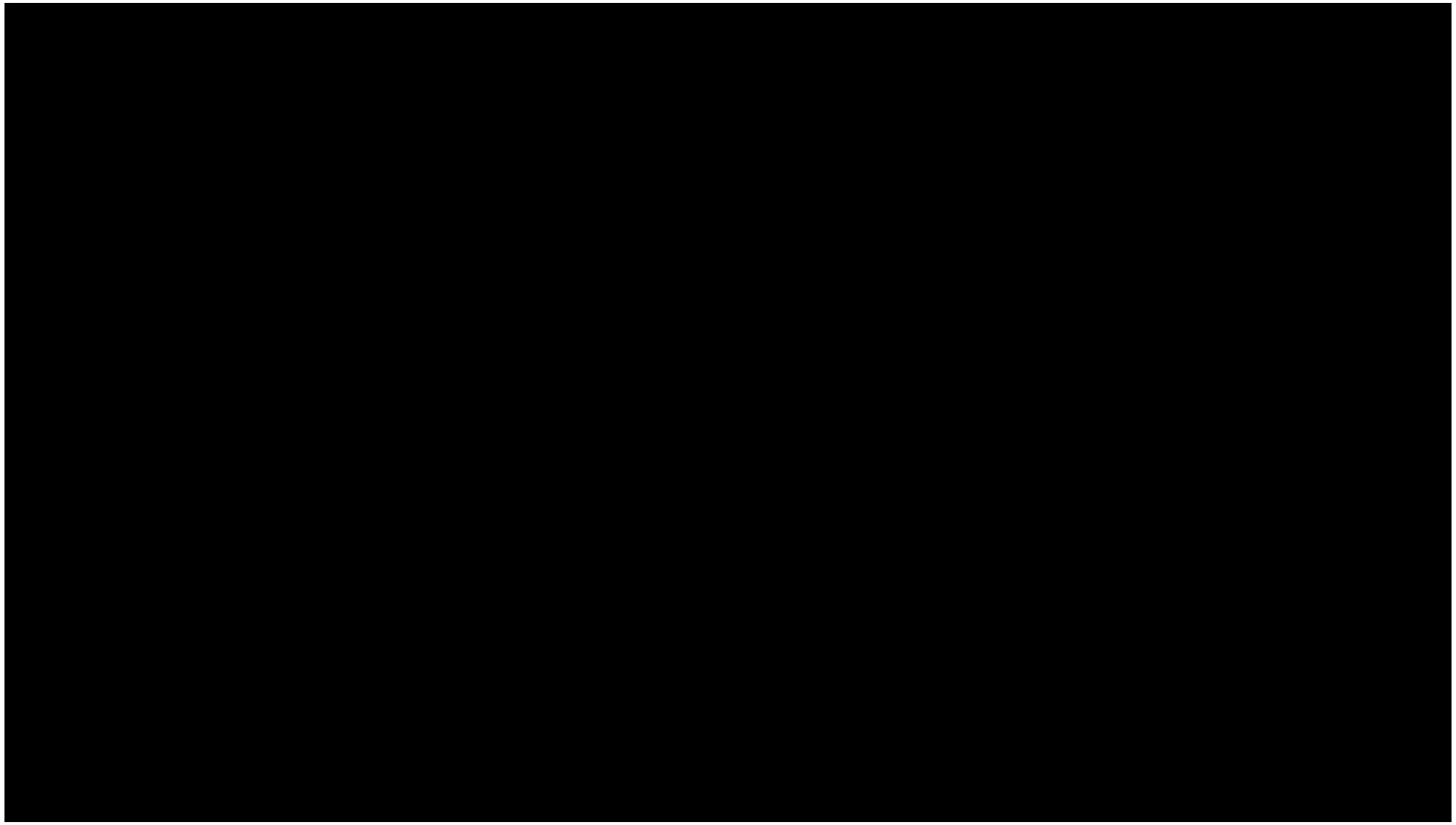
- **Off-policy**: update the value function based on a different policy than the one being learned
  - Behavior policy vs learned/update policy

- Off-policy policy gradient methods exist too!

[Bootcamp Summer 2020 Week 4 – On-Policy vs Off-Policy Reinforcement Learning (gatech.edu)]

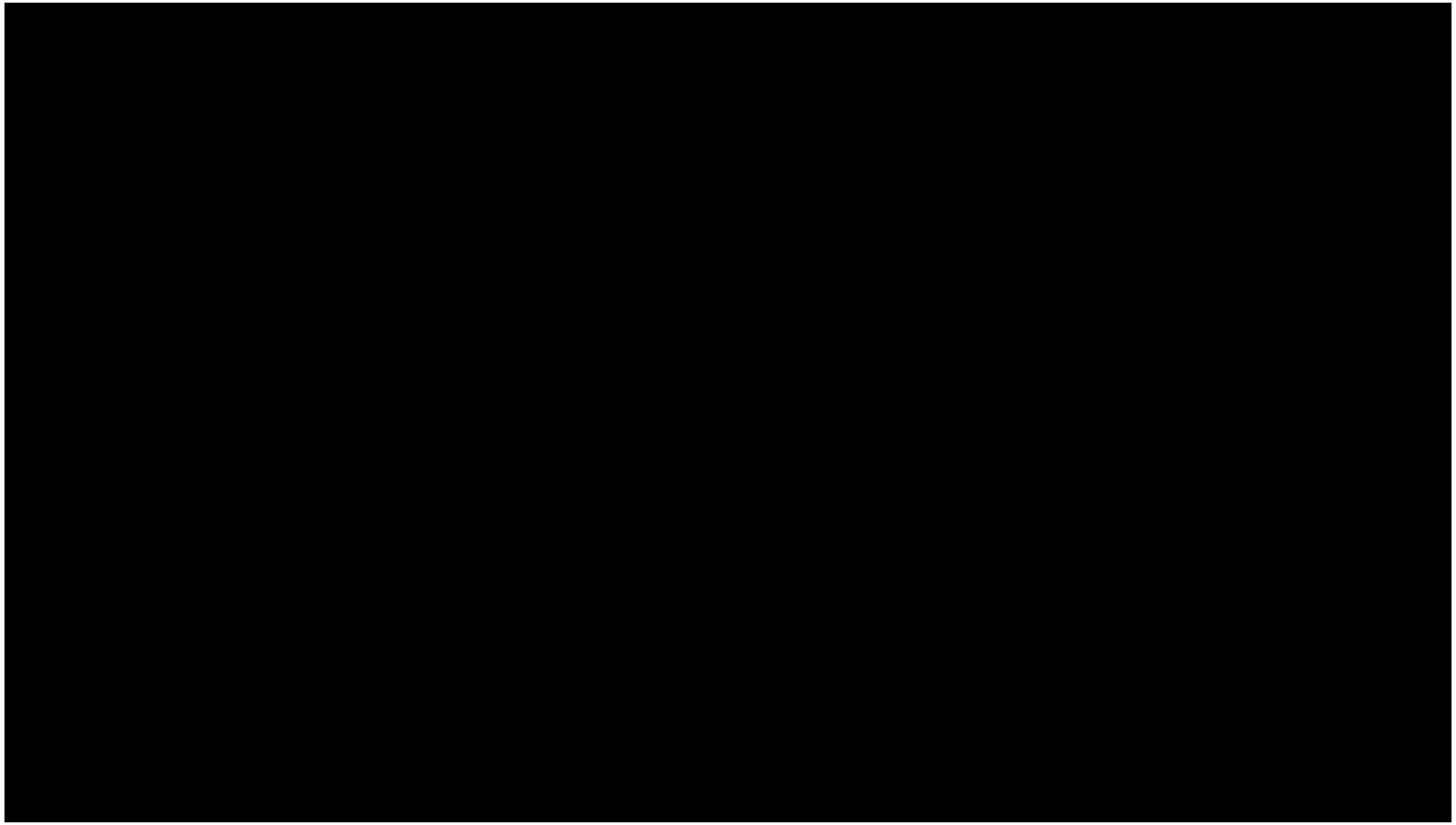
You can combine IL and RL too!

- Get the best of both worlds
- IL helps improve sample efficiency
- RL helps improve robustness and generalization
- Naïve idea: warm start RL with IL
- Better ideas:
    - Stay close to demonstrations when exploring in RL
    - Use reproduction error as part of reward
    - Offline RL
    - Transfer/meta learning



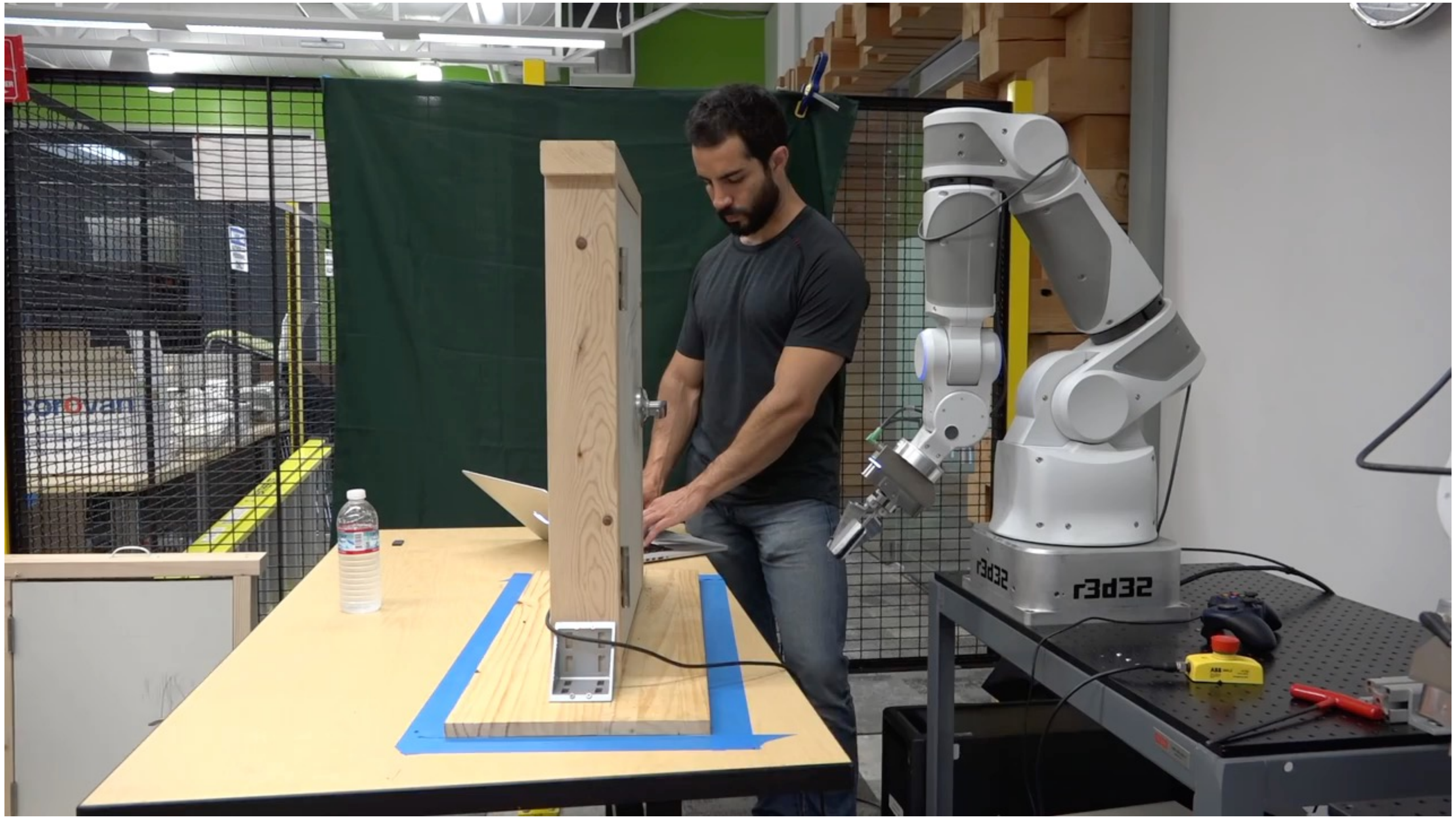
“In these experiments we tasked robots with trying to move their arms to goal locations, or reaching to and opening a door. Each robot has a copy of a neural network that allows it to estimate the value of taking a given action in a given state. By querying this network, the robot can quickly decide what actions might be worth taking in the world. When a robot acts, we add noise to the actions it selects, so the resulting behavior is sometimes a bit better than previously observed, and sometimes a bit worse. This allows each robot to explore different ways of approaching a task.”

https://research.googleblog.com/2016/10/how-robots-can-acquire-new-skills-from.html



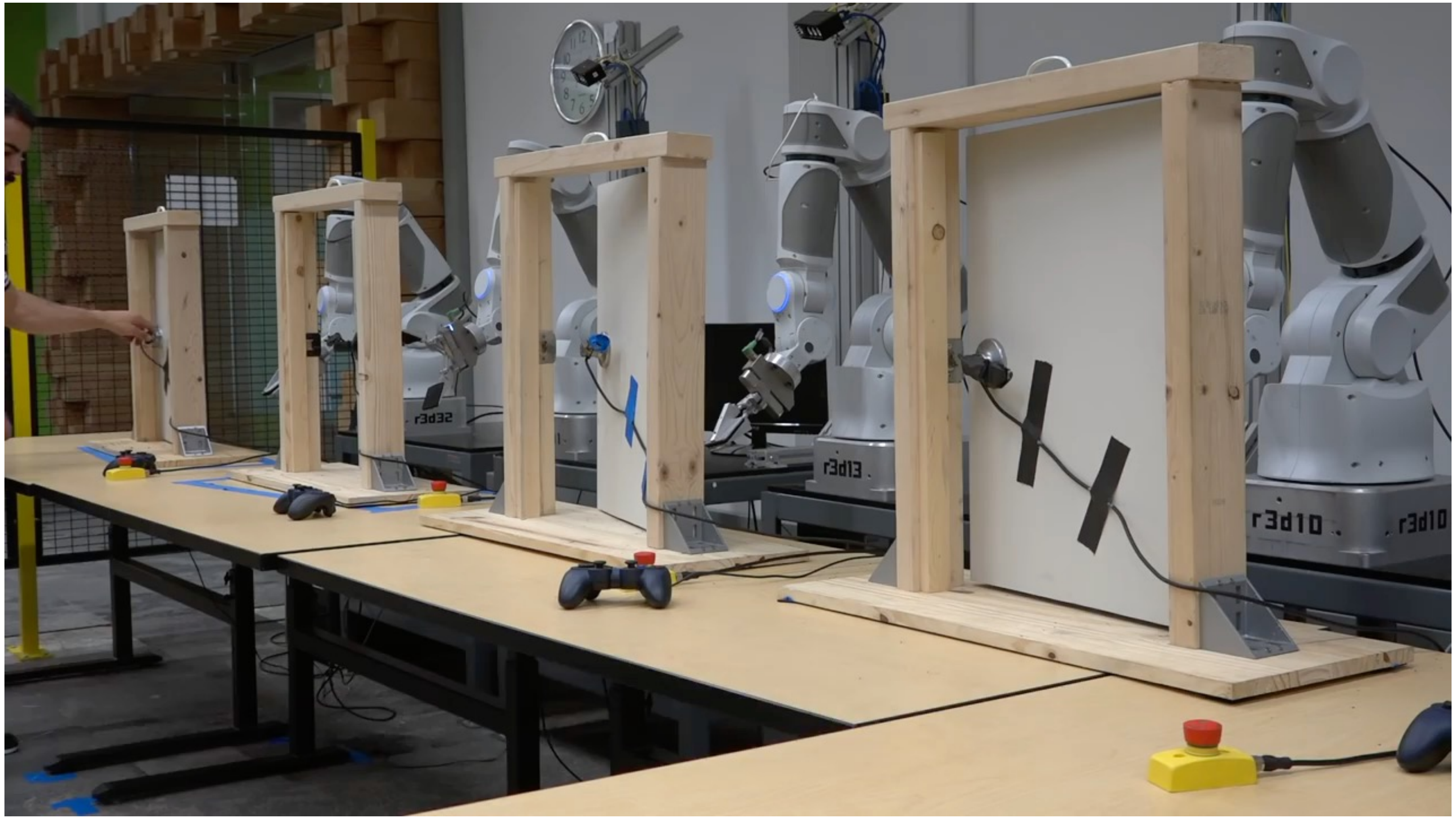
“With a few hours of practice, robots sharing their raw experience learn to make reaches to targets, and to open a door by making contact with the handle and pulling. In the case of door opening, the robots learn to deal with the complex physics of the contacts between the hook and the door handle without building an explicit model of the world.”

https://research.googleblog.com/2016/10/how-robots-can-acquire-new-skills-from.html



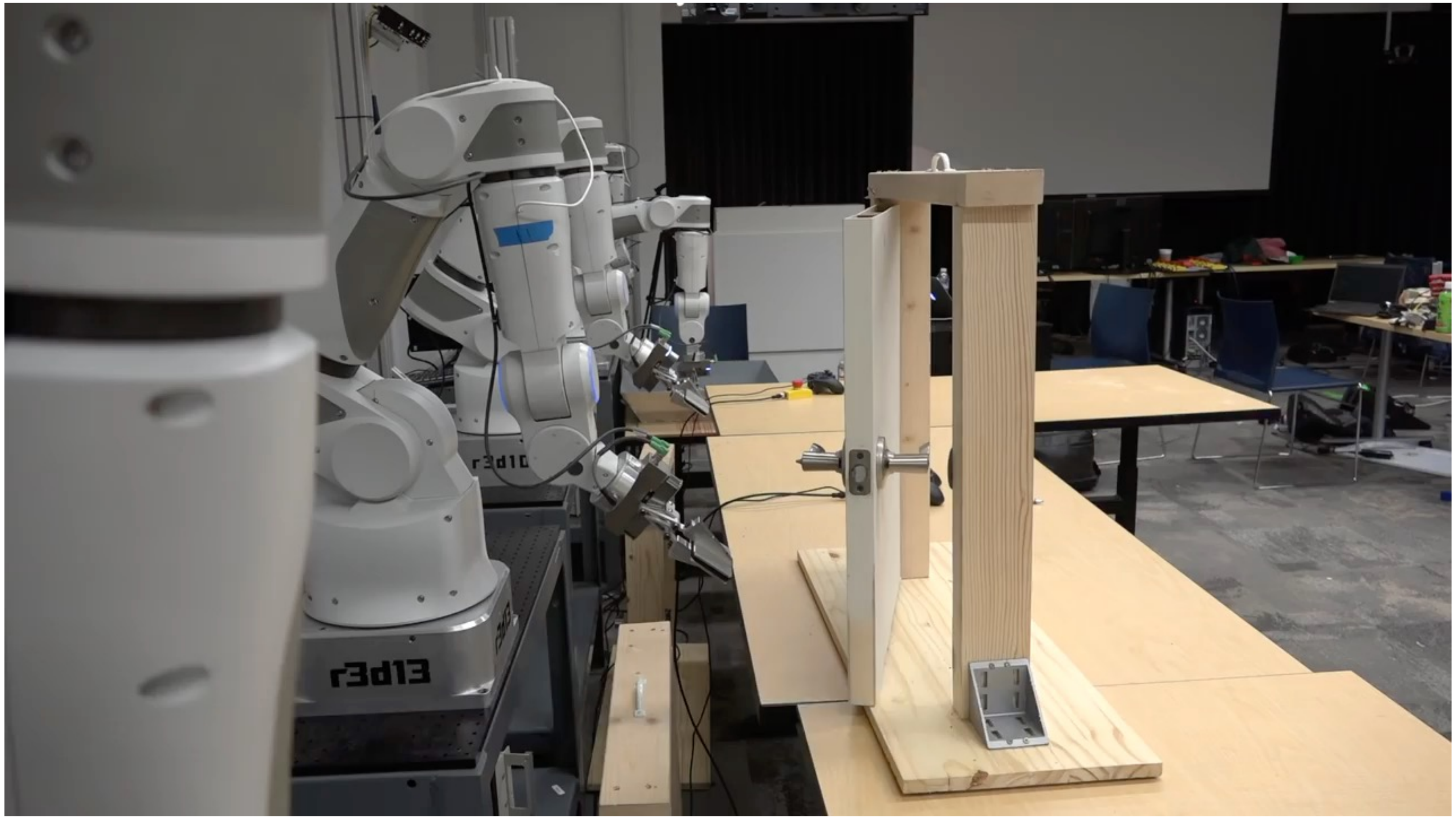
“Human guidance is important, not just for telling the robot what to do, but also for helping the robots along. We have a lot of intuition about how various manipulation skills can be performed, and it only seems natural that transferring this intuition to robots can help them learn these skills a lot faster.”

https://research.googleblog.com/2016/10/how-robots-can-acquire-new-skills-from.html



“Next, the robots collectively improve this policy through a trial-and-error learning process. Each robot attempts to open its own door using the latest available policy, with some added noise for exploration. These attempts allow each robot to plan a better strategy for opening the door the next time around, and improve the policy accordingly.”

https://research.googleblog.com/2016/10/how-robots-can-acquire-new-skills-from.html

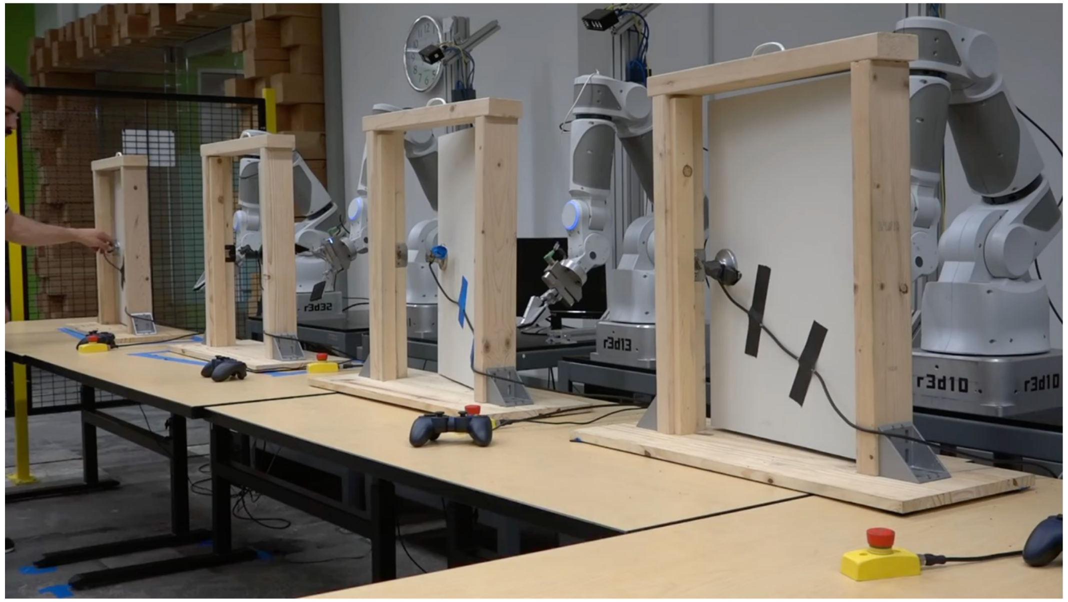


“Robots learn more effectively if they are trained on a curriculum of tasks that are gradually increasing in difficulty. In our experiment, each robot starts off by practicing the door-opening skill on a specific position and orientation of the door that the instructor had previously shown it. As it gets better at performing the task, the instructor starts to alter the position and orientation of the door to be just a bit beyond the current capabilities of the policy, but not so difficult that it fails entirely. This allows the robots to gradually increase their skill level over time, and expands the range of situations they can handle.

https://research.googleblog.com/2016/10/how-robots-can-acquire-new-skills-from.html

Recap

- The ideas behind this project
  - model-free RL
  - model-based RL
  - imitation learning
  - curriculum learning (a.k.a. scaffolding)



are not new, and were not new at the time of this work

- The novelty is in the *scale* of the experiments
  - More data, more experimental trials

October, 2016

14 robots
800,000 trials
3000 robot-hours
of practice

A few months earlier… "A human child is able to reliably grasp objects after one year, and takes around four years to acquire more sophisticated precision grasps. However, networked robots can instantaneously share their experience with one another, so if we dedicate 14 separate robots to the job of learning grasping in parallel, we can acquire the necessary experience much faster."

August, 2016

without replacement	first 10 ($N = 40$)	first 20 ($N = 80$)	first 30 ($N = 120$)
random	67.5%	70.0%	72.5%
hand-designed	32.5%	35.0%	50.8%
open loop	27.5%	38.7%	33.7%
our method	**10.0%**	**17.5%**	**17.5%**

with replacement	failure rate ($N = 100$)
random	69%
hand-designed	35%
open loop	43%
our method	**20%**

Table 1. Failure rates of each method for each evaluation condition. When evaluating without replacement, we report the failure rate on the first 10, 20, and 30 grasp attempts, averaged over 4 repetitions of the experiment.

ICRA 2019 FetchIt! Challenge

Fully autonomous mobile manipulation, completion of 7 individual tasks in 45 minutes.

- Includes bin picking, the same as what we saw in the previous example
- Objects are known ahead of time

This is the winning entry (from GT), but it has minimal machine learning!

- **Some learning used to select the grasp angle when picking up objects**

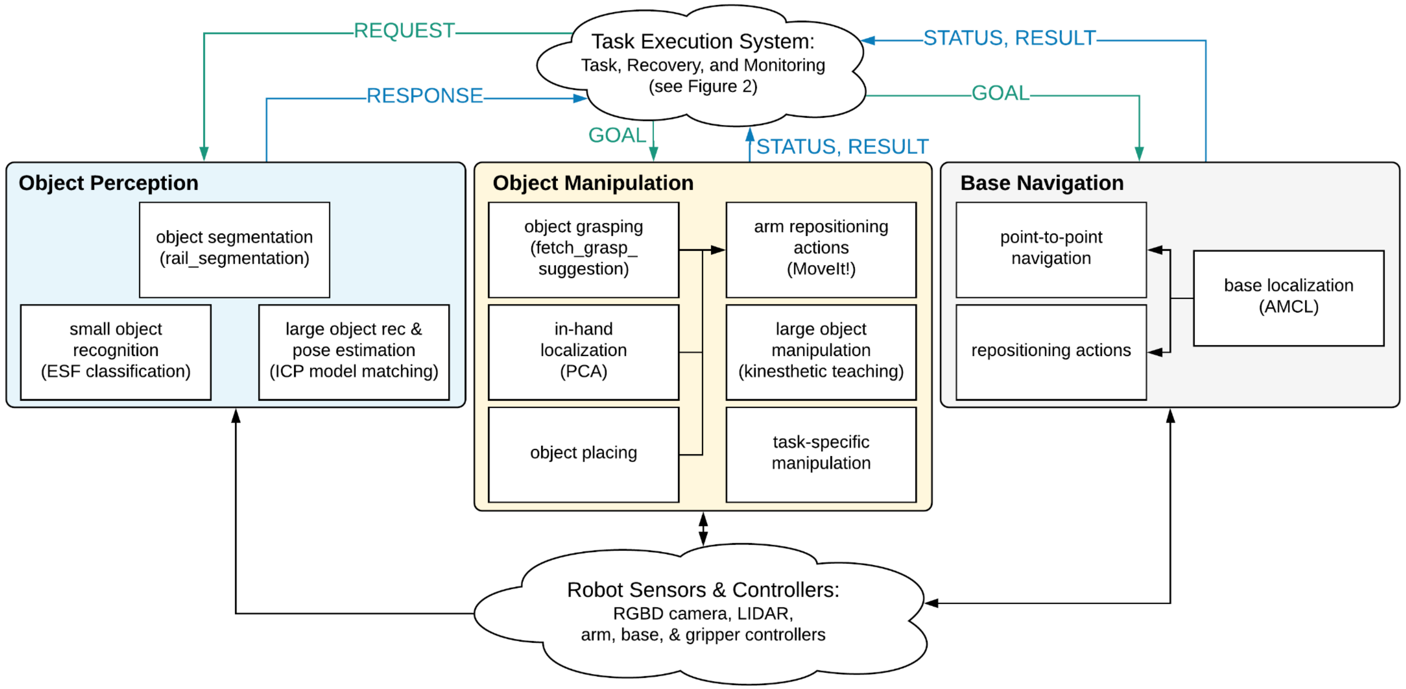
Open question: when should you learn vs engineer?

Banerjee, et al. Taking Recoveries to Task: Recovery-Driven Development for Recipe-based Robot Tasks. *International Symposium on Robotics Research (ISRR)*, 2019.

ICRA 2019 FetchIt! Challenge

Fully autonomous mobile manipulation, completion of 7 individual tasks in 45 minutes.

- Includes bin picking, the same as what we saw in the previous example
- Objects are known ahead of time



This is the winning entry (from GT), but it has minimal machine learning!

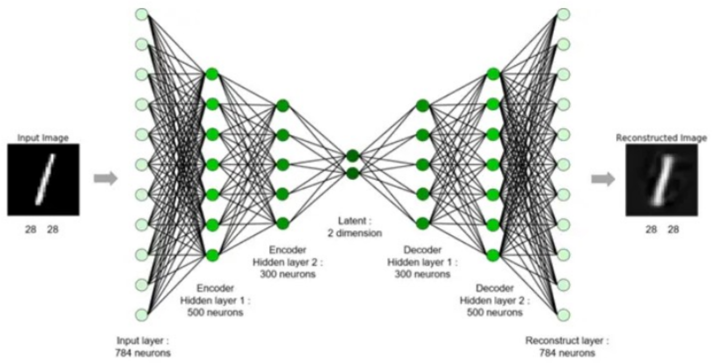
- **Some learning used to select the grasp angle when picking up objects**

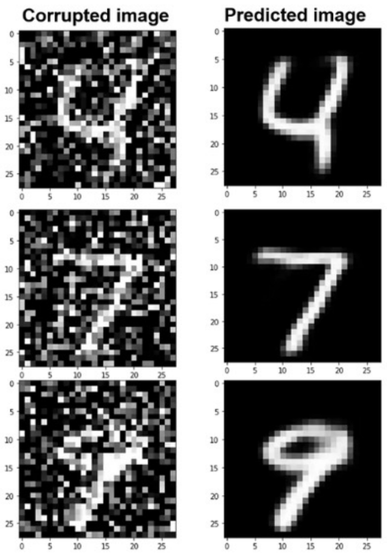
Open question: when should you learn vs engineer?

Banerjee, et al. Taking Recoveries to Task: Recovery-Driven Development for Recipe-based Robot Tasks. *International Symposium on Robotics Research (ISRR)*, 2019.

A third type: Self-supervised learning

- The key technology behind LLMs (e.g., ChatGPT)
- Leverage large amounts of data since there is no need for annotation
- Can be used to learn effective representations
- Can be used to learn sequences





Foundation Models

ChatGPT

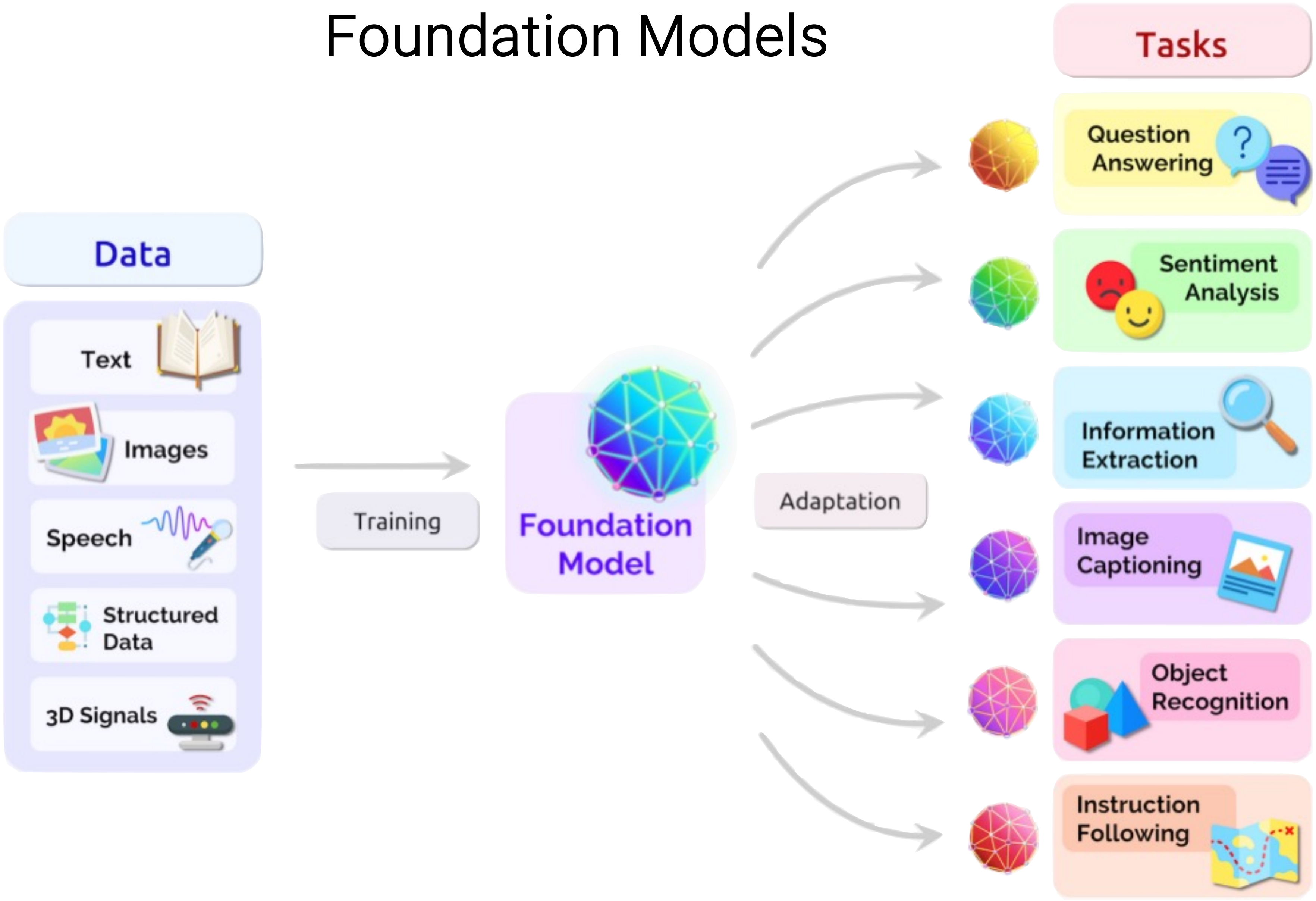
PaLM 2

Llama 2

Meta AI

DALL·E

Foundation Models



On the Opportunities and Risks of Foundation Models
Bommasani*, Liang* et al. 2022

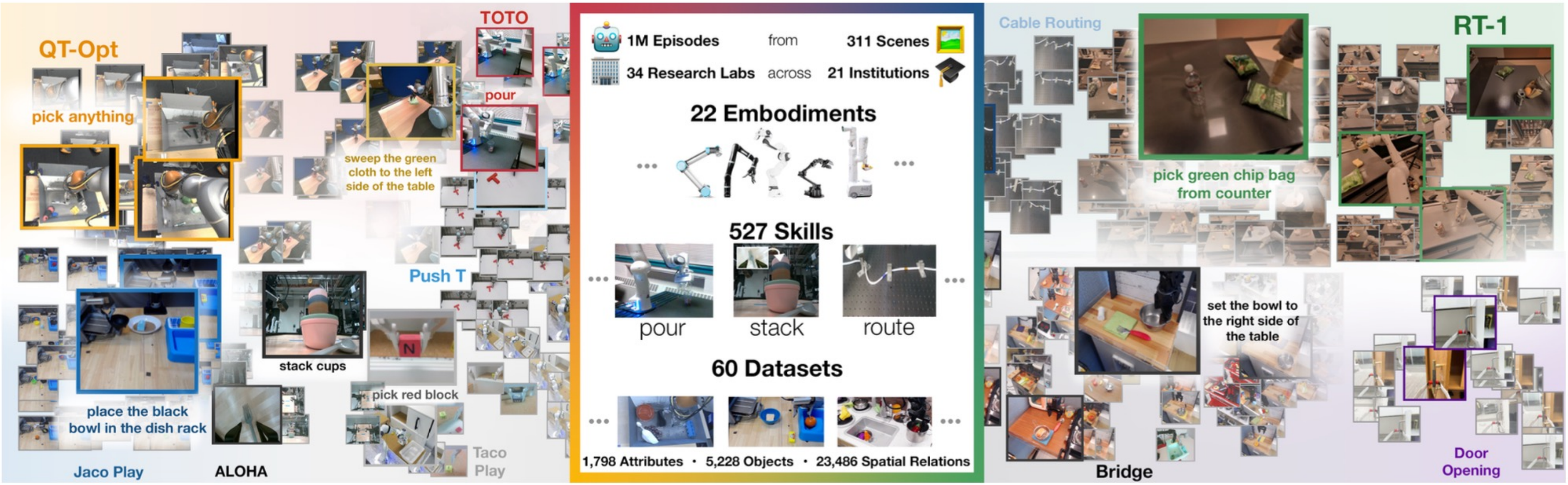
A variety of applications

- Navigation
- Locomotion
- Mobile manipulation
- ….

Open X-Embodiment: Robotic Learning Datasets and RT-X Models

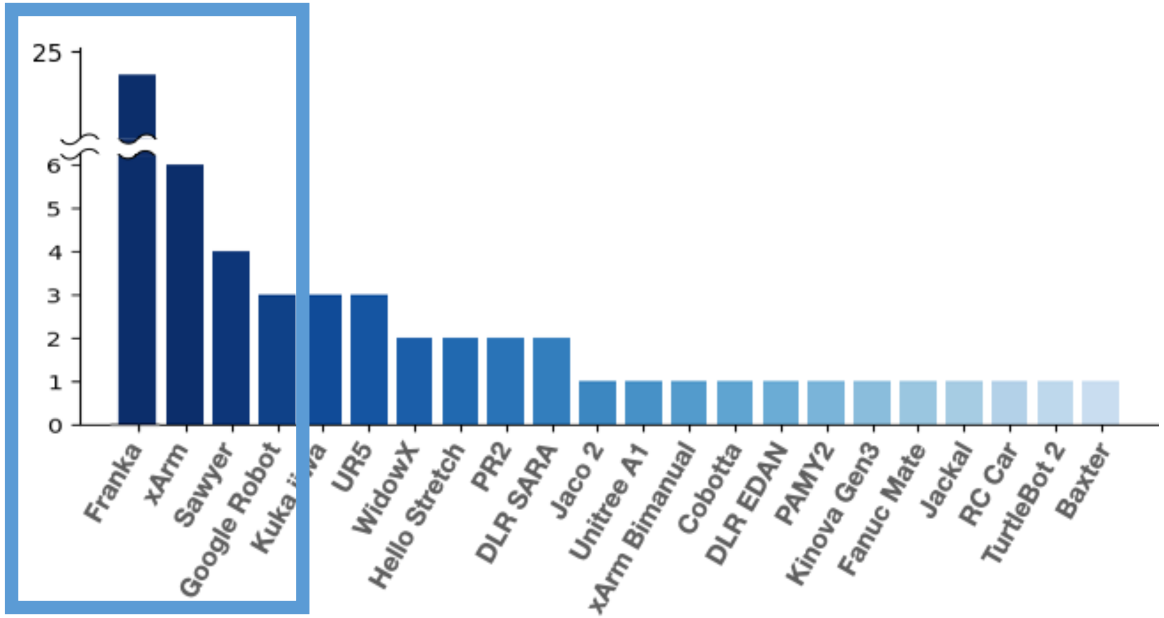
Open X-Embodiment Collaboration

(hover to display full author list)

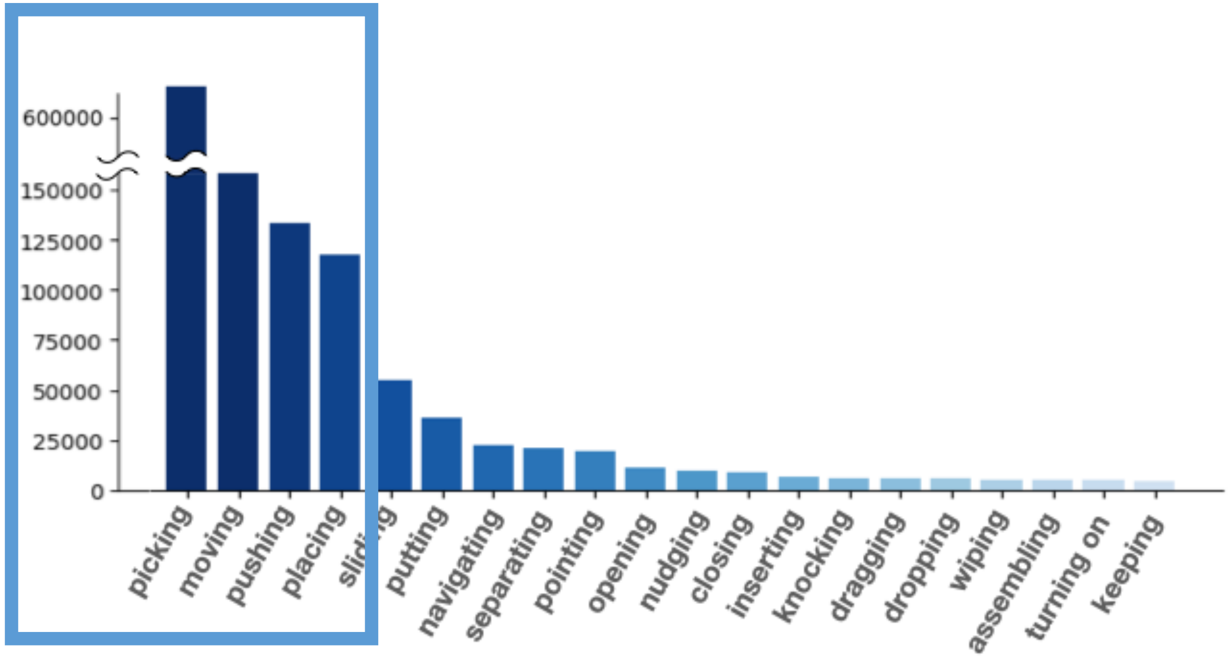


https://robotics-transformer-x.github.io/

Foundation Models in Robotics



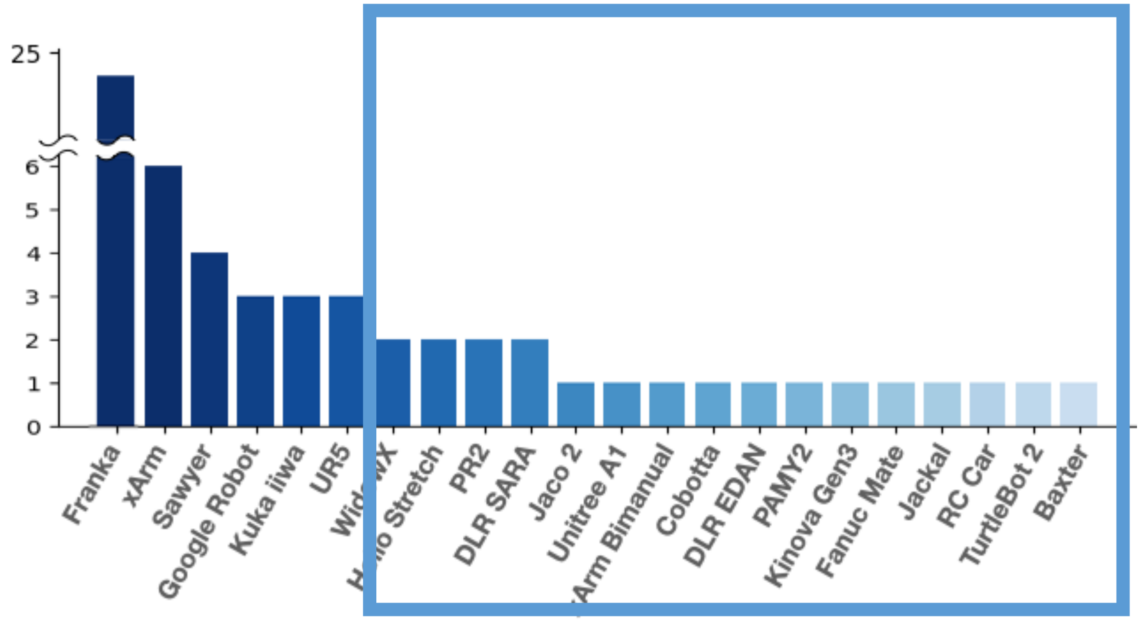
(a) # Datasets per Robot Embodiment



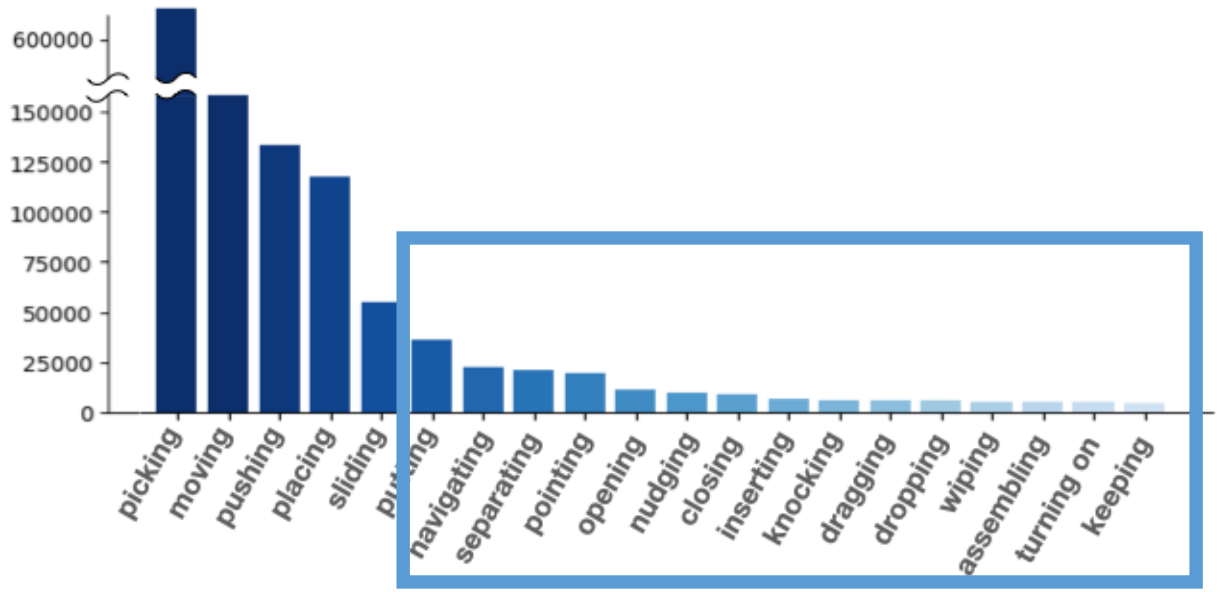
(d) Common Dataset Skills

Open X-Embodiment: Robotic Learning Datasets and RT-X Models

Foundation Models in Robotics



(a) # Datasets per Robot Embodiment



(d) Common Dataset Skills

ARTIFICIAL INTELLIGENCE

Is robotics about to have its own ChatGPT moment?

Researchers are using generative AI and other techniques to teach robots new skills—including tasks they could perform in homes.

By Melissa Heikkilä

April 11, 2024

PETER ADAMS

A spectrum of methods

Analytical / hand-designed

- *Well-understood*
- *Computationally efficient*
- *Guarantees*
- *Limited generalization*
- *Requires domain knowledge*

Learning from scratch

- *No need for domain knowledge*
- *Generalization*
- *Computationally expensive*
- *Black box*

A spectrum of methods

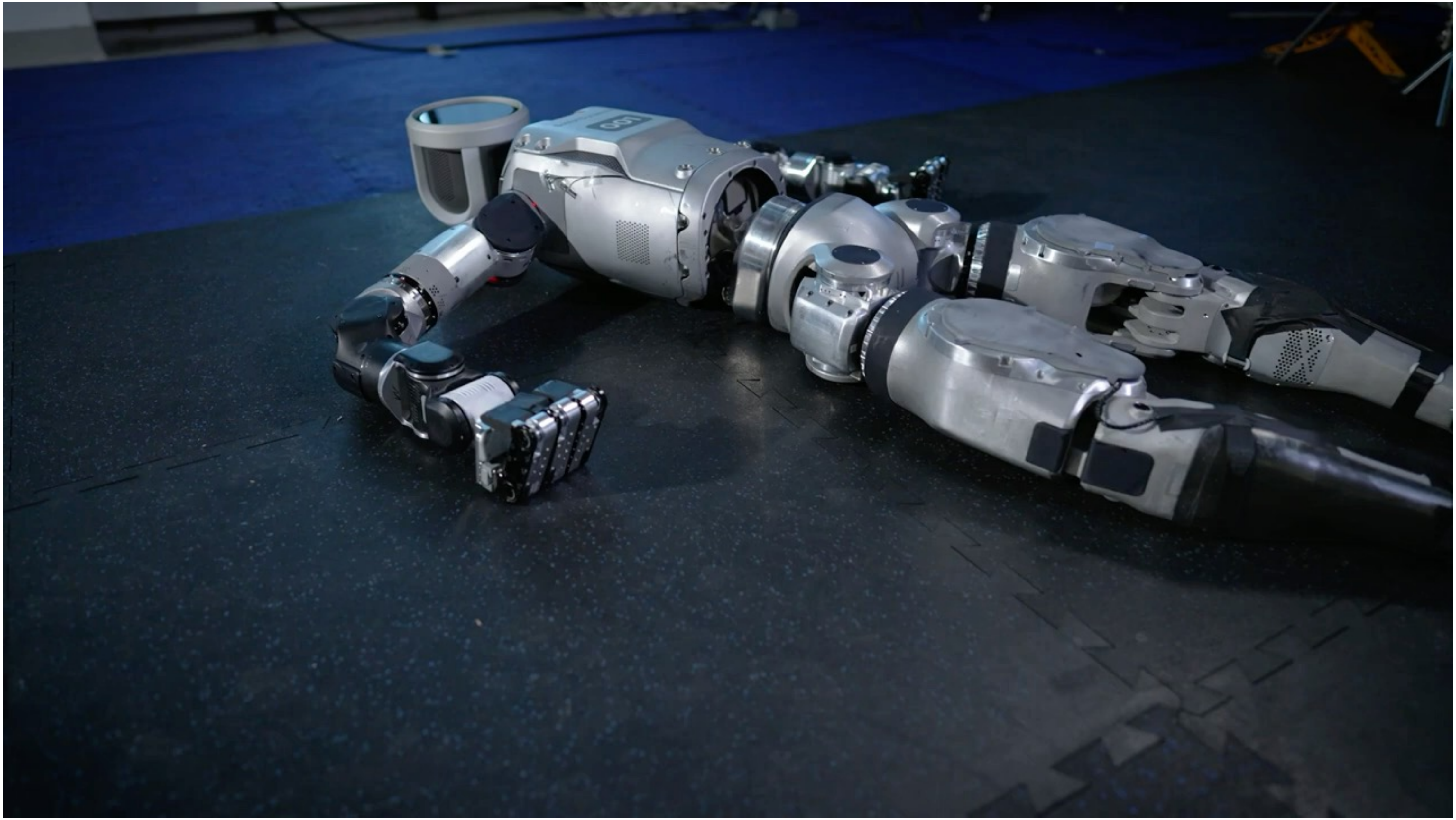
Analytical / hand-designed

Learning from scratch

Middle ground:

Structured Learning

- *Limiting the hypothesis space*
- *Constraining the output of the model*
- *Modular design with analytical components*
- *…*



The new Atlas from Boston Dynamics

Other related concepts

- Sim-2-real gap
- Safe learning
- Multi-task learning
- Transfer learning
- Meta learning
- Lifelong learning
- Interactive learning
- Learning from suboptimal and noisy data