

1st Class / Jan 10 (Fri)

Modern Robot Learning: Hands-on Tutorial

Haoshu Fang, Younghyo Park, Jagdeep Bhatia, Lars Ankile, Pulkit Agrawal



Modern Robot Learning: Hands-on Tutorial

Haoshu Fang, **Younghyo Park**, Jagdeep Bhatia, Lars Ankile, Pulkit Agrawal



Modern Robot Learning: Hands-on Tutorial

<https://modern-robot-learning.github.io/>

- Course announcements
- Tutorial Session Reservation
- Lecture Notes
- Course Staff Contacts



Course Overview

What will you learn and experience in this course?

Robot Data Collection

- What is **robot data**?
- What/how do we collect?
- How do we use it?

Policy Training

- Training Methods
- Policy Architectures
- Policy Evaluation

Simulation for Robotics

- Role of simulation
- Designing environments in simulation world
- Transfer to real-world

Course Overview

Hands-on Tutorial

Robot Data Collection

- What is **robot data**?
- What/how do we collect?
- How do we use it?

Policy Training

- Training Methods
- Policy Architectures
- Policy Evaluation

Simulation for Robotics

- Role of simulation
- Designing environments in simulation world
- Transfer to real-world

Hands-on Tutorial

Sneak Peak



Hands-on Tutorial

Sneak Peak

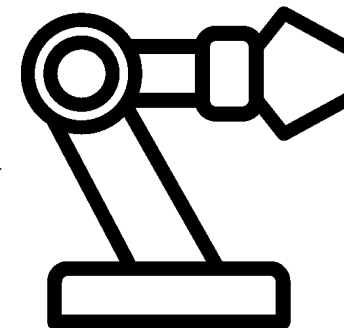
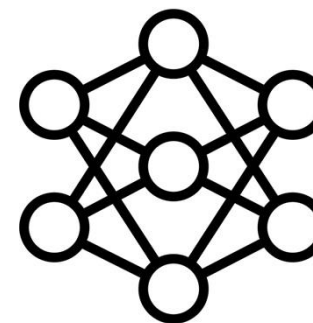
Iterate to get best performance

Teleoperate Robots
in Virtual Reality

Dataset
Creation / Curation

Policy Training

Policy
Evaluation



Hands-on Tutorial

Sneak Peak

Sign up to book a session
with Apple Vision Pro
in our course website!
(available later next week)

Teleoperate Robots
in Virtual Reality



Dataset
Creation / Curation



Po

C
C

Course Schedule

- **Jan 10 (Fri)** 1PM, 32-124
 - Overview of Robot Data Collection
 - Principles of Robot Teleoperation
- **Jan 13 (Mon)** 1PM, 32-124
 - Policy Training with Teleoperated Datasets
 - Role of Simulation for Robot Teleoperation and Training
- **Jan 15 (Wed)** 1PM, 32-124
 - Non-Teleoperated Datasets Overview
 - Challenges and Opportunities in Policy Training

Course Schedule

- **Jan 10 (Fri)** 1PM, 32-124
- **Jan 13 (Mon)** 1PM, 32-124
- **Jan 15 (Wed)** 1PM, 32-124
- **Jan 17 (Fri)** 1PM, 32-124
 - Hands-on Tutorial Guideline Overview
- **Jan 20-29** Reserved Office Hours, Bldg 45
 - Time for Personal Projects
 - Data Collection + Policy Training / Evaluation
 - **Final Report Submission** (Details TBD)

Today

- **Jan 10 (Fri) 1PM, 32-124**
 - Overview of Robot Data Collection
 - Principles of Robot Teleoperation
- Jan 13 (Mon) 1PM, 32-124
- Jan 15 (Wed) 1PM, 32-124
- Jan 17 (Fri) 1PM, 32-124
- Jan 20-29 Reserved Office Hours, Bldg 45



The predictions for 2020 we got horribly wrong

By [Rob Picheta](#), CNN Business

Updated 10:34 AM EST, Sun January 5, 2020



<https://edition.cnn.com/2020/01/01/tech/2020-predictions-we-got-wrong-scli-intl/index.html>

The predictions for 2020 we got horribly wrong

We're not vacationing on the moon – yet

They say the past is a foreign country. Well, if that's true, then the future is a foreign planet. With hotels on it.

Accessible vacations in space have been predicted for decades. “Look back to what people were talking about back in the 60s or 70s – space tourism has been a vision for a long time,” says Laura Forczyk, founder of space consulting firm Astralytical. “Go back to that Stanley Kubrick movie, where Pan Am was taking tourists to various destinations,” she adds, referring to the blockbuster “2001: A Space Odyssey.”

In 2009, it finally seemed we were on the cusp of a breakthrough, with a number of companies and individuals expressing a desire to make the 2010s the decade of space tourism.

“By 2020 you'll have seen private citizens circumnavigate the moon,” Eric Anderson of Space Adventures told the website Space.com in 2009. Tesla and SpaceX founder Elon Musk went further. “I'm going to go out on a limb and say that by 2020 there will be serious plans to go to Mars with people,” the same site quoted him as saying.

The predictions for 2020 we got horribly wrong

We're still using our phones

Aside from our diets, in 2000 Kurzweil also predicted that computers would be “largely invisible” and “embedded everywhere – in walls, tables, chairs, desks, clothing, jewelry, and bodies,” by 2020.

He was one of a handful of futurists to predict that smart glasses or contact lenses would replace our phones. Google did give this a try, but it failed to resonate with the public.



Sean Gallup/Getty Images

Google Glass didn't catch on when it was introduced in the middle of this decade.

The predictions for 2020 we got horribly wrong

The robot revolution was delayed

The prospect of robots coming for our jobs has been a perennial concern of every post-war generation, and by 2020 we were meant to be virtually redundant in many areas.

“Futurists and technology experts say robots and artificial intelligence of various sorts will become an accepted part of daily life by the year 2020 and will almost completely take over physical work,” Elon University noted in 2006.

British futurologist Ian Pearson went further still. “Consciousness is just another sense, effectively, and that’s what we’re trying to design in a computer,” he told the UK’s Observer newspaper in 2005. “It’s my conclusion that it is possible to make a conscious computer with superhuman levels of intelligence before 2020.”

“It would definitely have emotions,” he added. “If I’m on an aeroplane I want the computer to be more terrified of crashing than I am so it does everything to stay in the air.”

“Futurists and technology experts say robots and artificial intelligence of various sorts will become an accepted part of daily life by the year 2020 and will almost completely take over physical work,” Elon University noted in 2006.

“Futurists and technology experts say robots and artificial intelligence of various sorts will become an accepted part of daily life by the year 2020 and will almost completely take over physical work,” Elon University noted in 2006.


True?



ARTIFICIAL INTELLIGENCE PROGRAM
HOW GOOD IS "CHATGPT" - AND CAN IT BE DETECTED?

SMERCONISH

“Futurists and technology experts say robots and artificial intelligence of various sorts will become an accepted part of daily life by the year 2020 and will almost completely take over physical work,” Elon University noted in 2006.

 Stanford Law School

GPT-4 Passes the Bar Exam: What That Means for Artificial Intelligence Tools in the Legal Profession | Stanford Law ...

CodeX–The Stanford Center for Legal Informatics and the legal technology company Casetext recently announced what they called “a watershed...

Apr 19, 2023

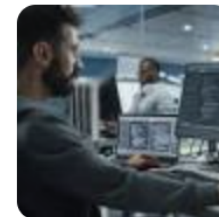


 PCMag

ChatGPT Passes Google Coding Interview for Level 3 Engineer With \$183K Salary

Google fed coding interview questions to ChatGPT and, based off the AI's answers, determined it would be hired for a level three engineering...

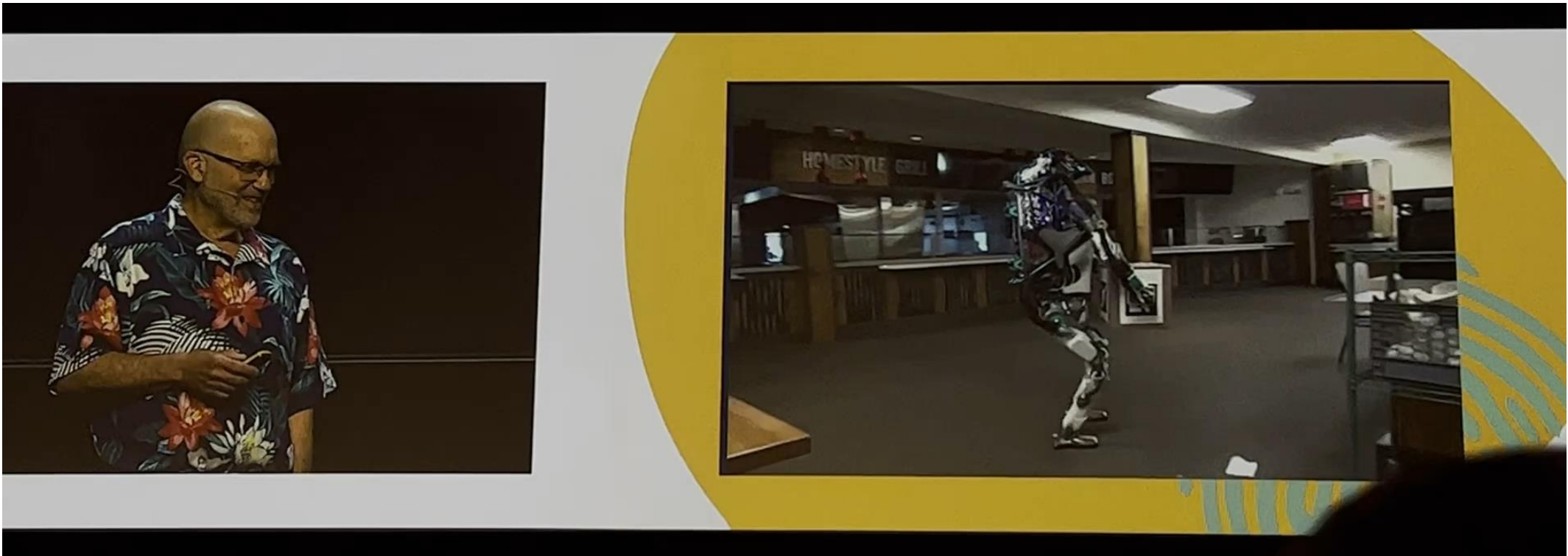
Feb 1, 2023



“Futurists and technology experts say robots and artificial intelligence of various sorts will become an accepted part of daily life by the year 2020 and will almost completely take over physical work,” Elon University noted in 2006.

“Futurists and technology experts say robots and artificial intelligence of various sorts will become an accepted part of daily life by the year 2020 and will almost completely take over physical work,” Elon University noted in 2006.

Not quite there yet.



Stark contrast

artificial intelligence of various sorts
will become an accepted
part of daily life by the year 2020

Stanford Law School

GPT-4 Passes the Bar Exam: What That Means for Artificial Intelligence Tools in the Legal Profession | Stanford Law ...

CodeX—The Stanford Center for Legal Informatics and the legal technology company Casetext recently announced what they called “a watershed...”

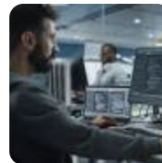
Apr 19, 2023

PCMag

ChatGPT Passes Google Coding Interview for Level 3 Engineer With \$183K Salary

Google fed coding interview questions to ChatGPT and, based off the AI's answers, determined it would be hired for a level three engineering...

Feb 1, 2023

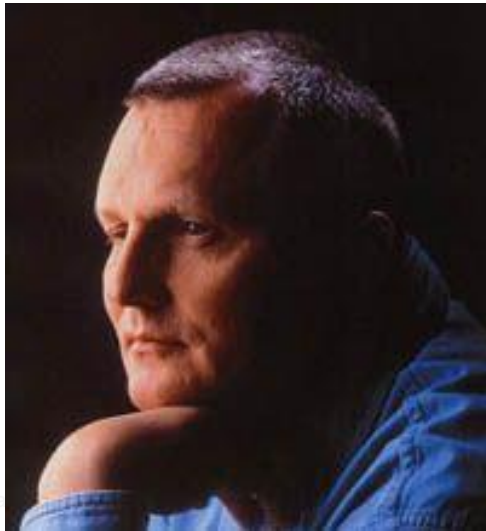


robots

will almost completely take over physical work,



So ... We can make machines **pass bar exams**, but cannot make it **move boxes**?



Hans Moravec

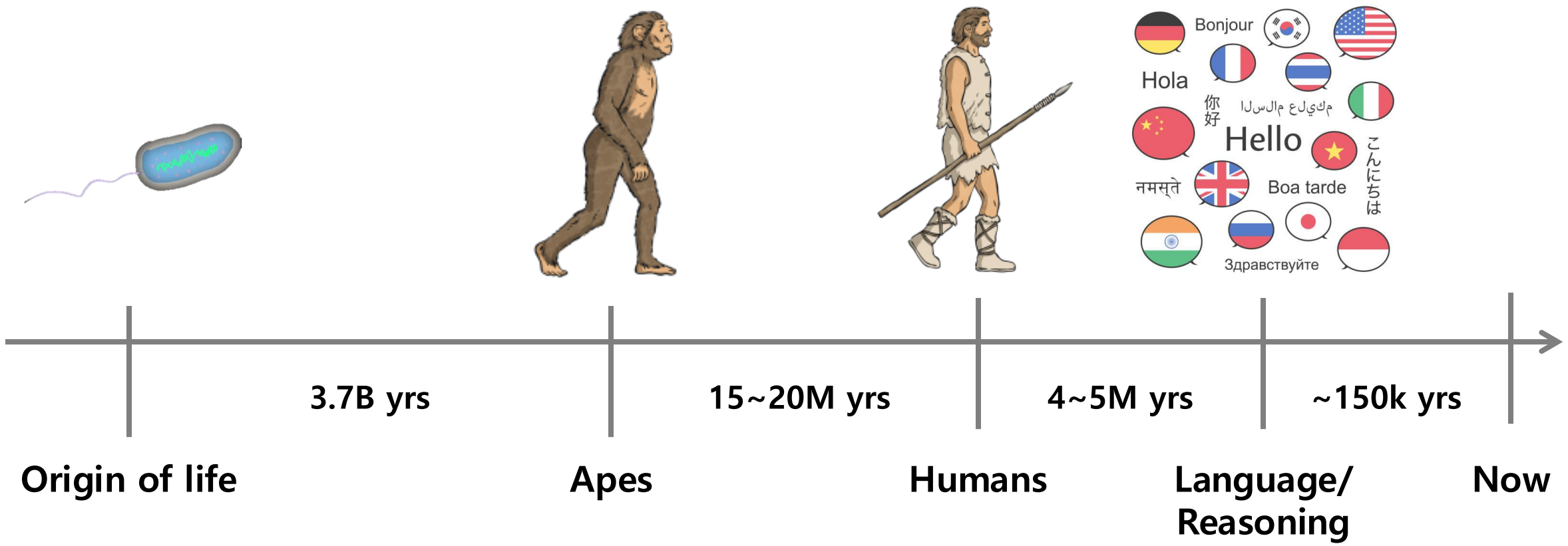
“reasoning requires very little computation, but sensorimotor and perception skills require enormous computation resources” (1980)

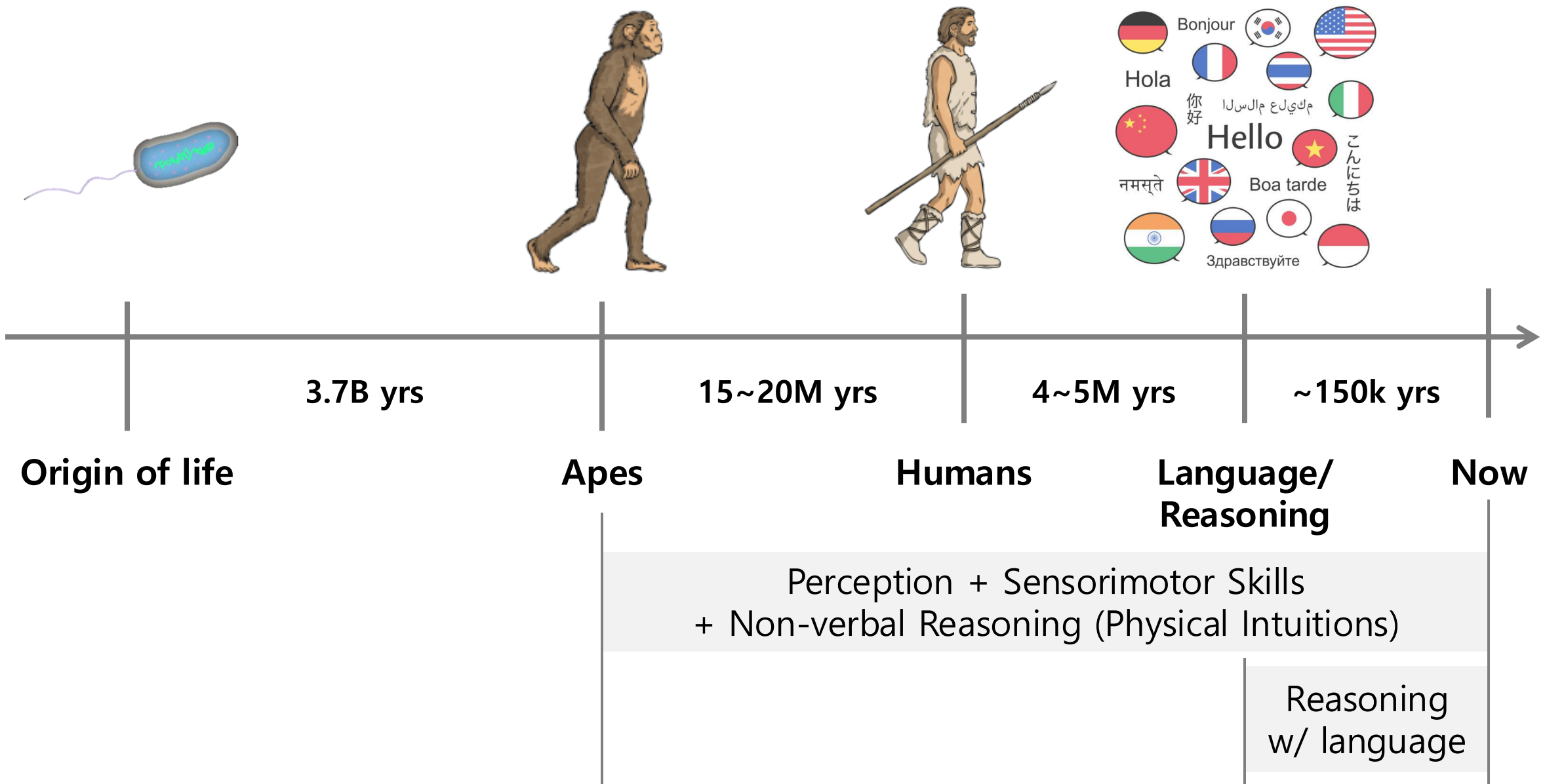
Moravec's Paradox

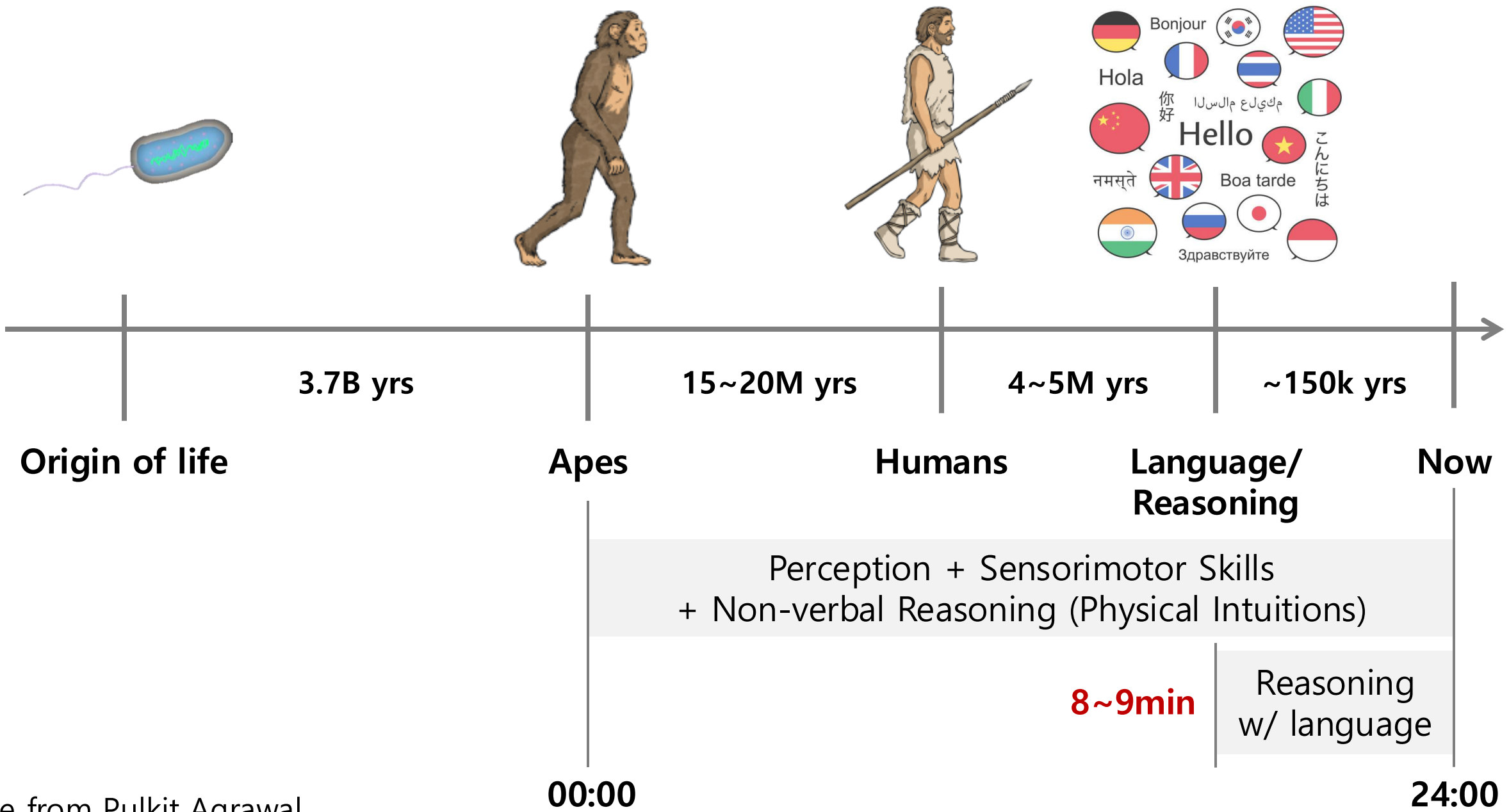
“... the main lesson of 35 years of research is that the hard problems are easy and the easy problems are hard ... ” (1994)

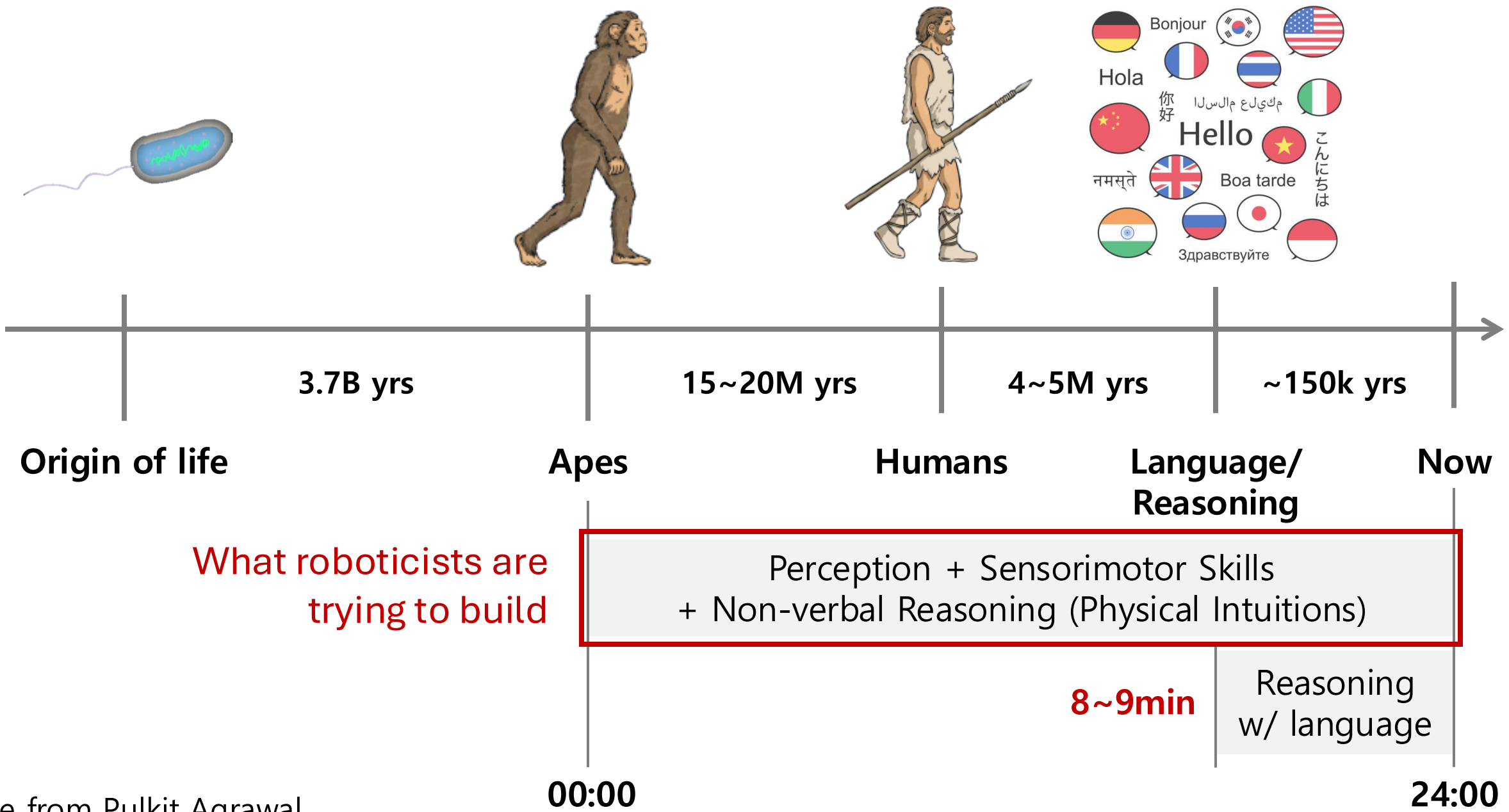
Steven Pinker

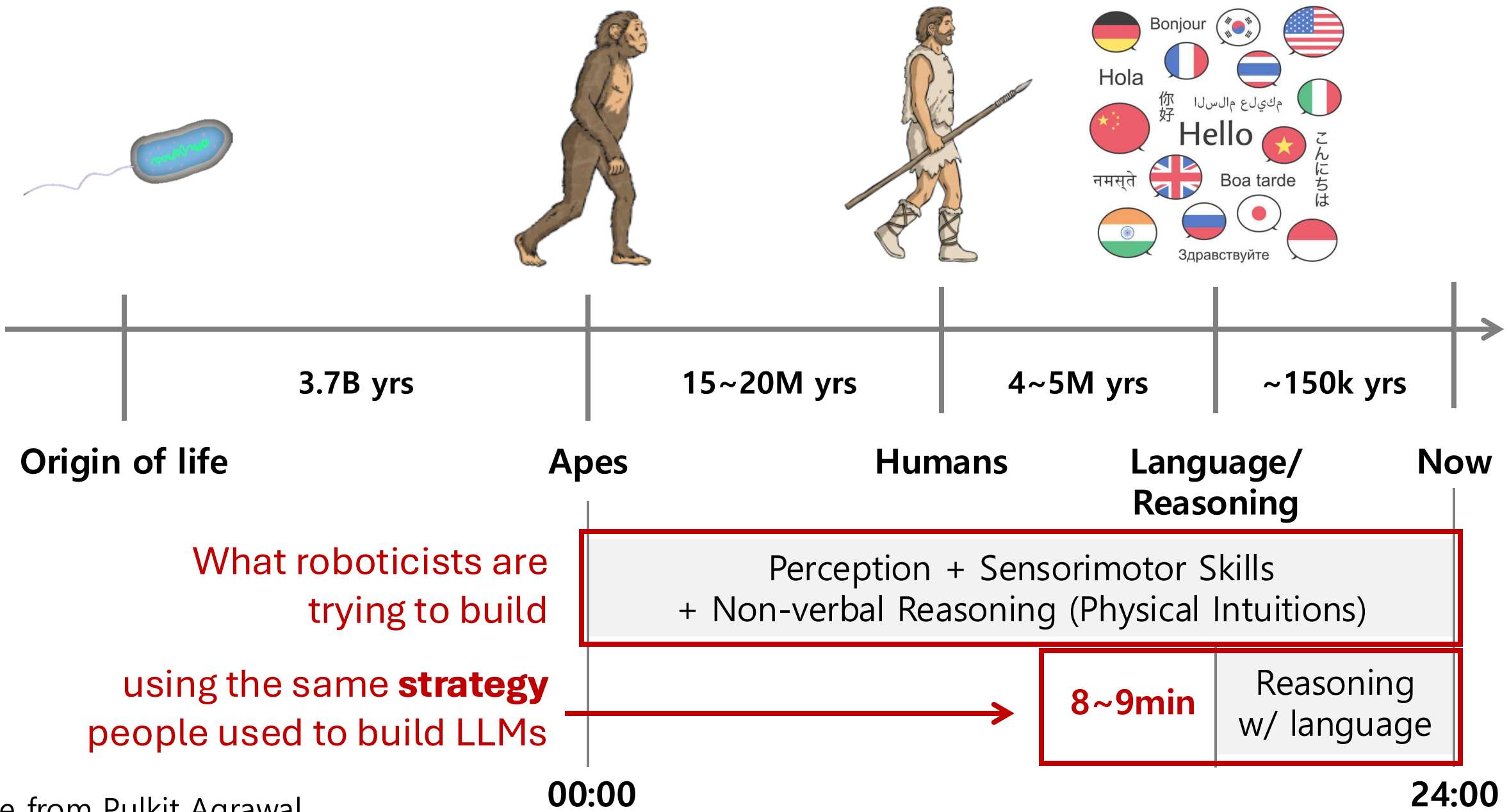












strategy: **massive dataset** with the **right training method**

strategy: massive dataset with the right training method

	massive dataset	right training method
Vision/Language Models	Scraped datasets from the web	Next token (word) Prediction

strategy: massive dataset with the right training method

	massive dataset	right training method
Vision/Language Models	Scraped datasets from the web	Next token (word) Prediction
Robot Models	?	Action Prediction

Road to Large-Scale Robot Dataset

Road to Large-Scale Robot Dataset

What's a **Robot Dataset**?

Road to Large-Scale Robot Dataset

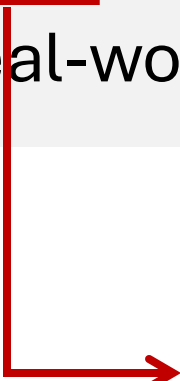
What's a **Robot Dataset**?

- Data recorded by **robot embodiments** solving diverse tasks in real-world.

Road to Large-Scale Robot Dataset

What's a **Robot Dataset**?

- Data recorded by **robot embodiments** solving diverse tasks in real-world.


$$D = \{(s_1, a_1, s_2, a_2, \dots, s_n, a_n)\}$$

Robot's Actions

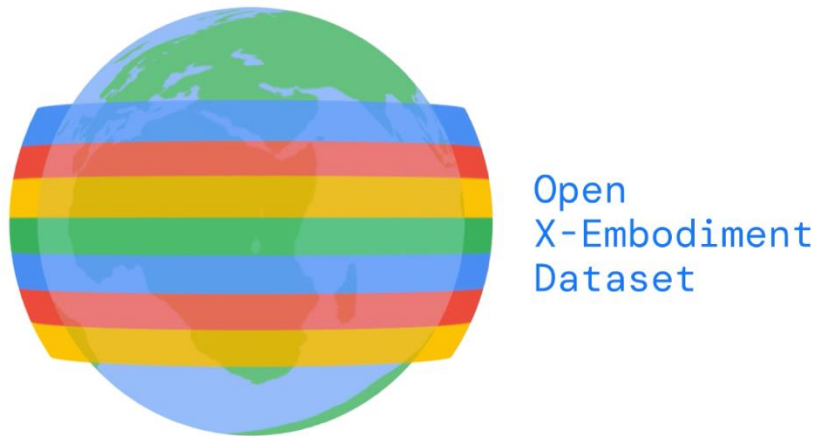
Robot's Observation

(i.e., joint states, camera feeds, tactile sensors)

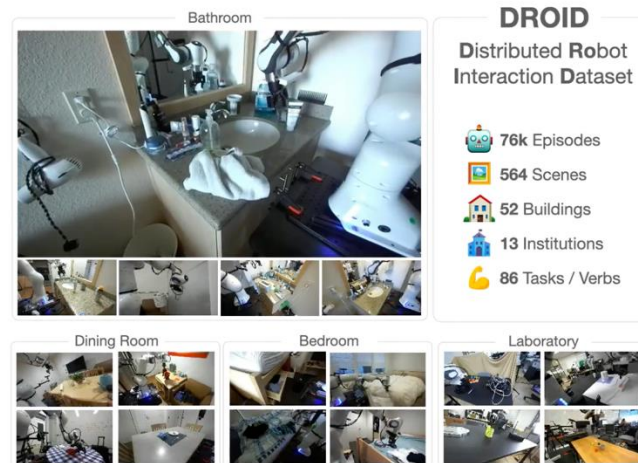
Road to Large-Scale Robot Dataset

What's a **Robot Dataset**?

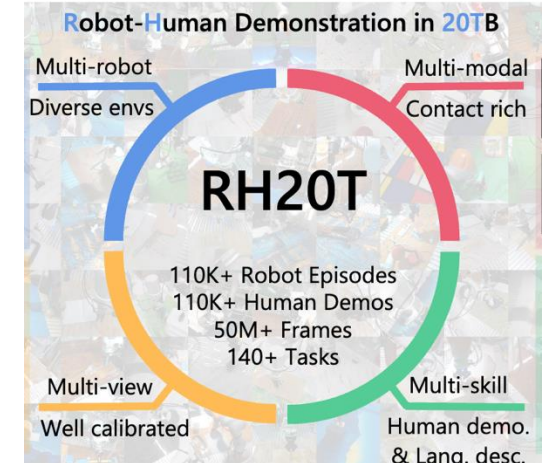
- Data recorded by **robot embodiments** solving diverse tasks in real-world.



O'Neill, Abby, et al. "**Open x-embodiment**: Robotic learning datasets and rt-x models." *arXiv:2310.08864* (2023).



Khazatsky, Alexander, et al. "**DROID**: A large-scale in-the-wild robot manipulation dataset." *arXiv preprint arXiv:2403.12945* (2024).

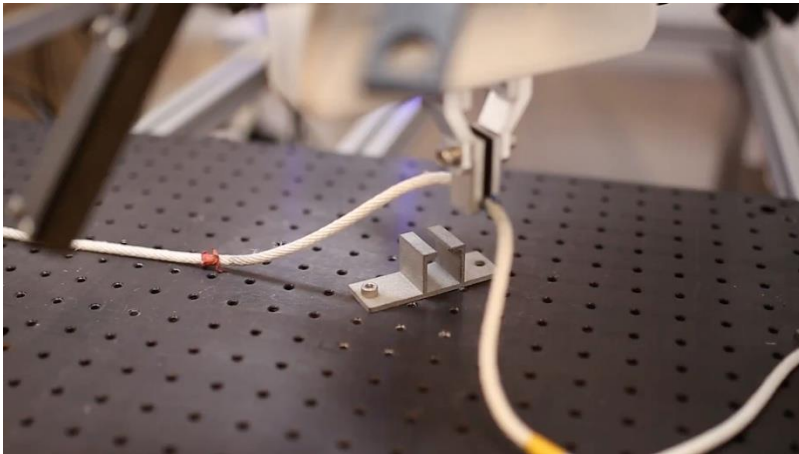


Fang, Hao-Shu, et al. "**RH20t**: A robotic dataset for learning diverse skills in one-shot." *RSS 2023 Workshop on Learning for Task and Motion Planning*. 2023.

Road to Large-Scale Robot Dataset

What's a **Robot Dataset**?

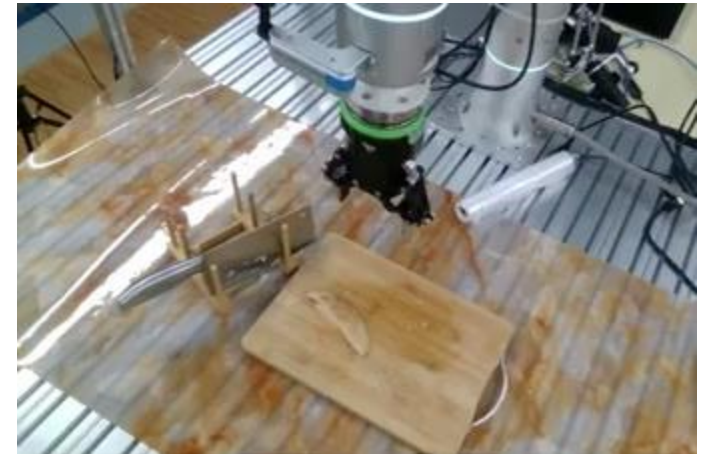
- Data recorded by **robot embodiments** solving diverse tasks in real-world.



O'Neill, Abby, et al. "**Open x-embodiment**: Robotic learning datasets and rt-x models." *arXiv:2310.08864* (2023).



Khazatsky, Alexander, et al. "**DROID**: A large-scale in-the-wild robot manipulation dataset." *arXiv preprint arXiv:2403.12945* (2024).



Fang, Hao-Shu, et al. "**RH20t**: A robotic dataset for learning diverse skills in one-shot." *RSS 2023 Workshop on Learning for Task and Motion Planning*. 2023.

Road to Large-Scale Robot Dataset

What's a **Robot Dataset**?

- Data recorded by **robot embodiments** solving diverse tasks in real-world.
- Any data from **any embodiments** (including humans) that contains useful knowledge about manipulation strategies.

Road to Large-Scale Robot Dataset

What's a **Robot Dataset**?

- Data recorded by **robot embodiments** solving diverse tasks in real-world.
- Any data from **any embodiments** (including humans) that contains useful knowledge about manipulation strategies.



Around the World in
3,000 Hours of Egocentric Video

Road to Large-Scale Robot Dataset

What's a Robot Dataset

- Data recorded by robots in real-world.
- Any data from a robot can be helpful training data.



Around the World in
3,000 Hours of Egocentric Video



Recordings of human behaviors in diverse settings

Road to Large-Scale Robot Dataset

What's a Robot Dataset?



Left Hand



Right Hand



Object

Hand-Object Distance (Brighter: closer)

High-Quality Capture/Annotation of Human Manipulation Behaviors

Fan, Zicong, et al. "ARCTIC: A dataset for dexterous bimanual hand-object manipulation." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2023.

Road to Large-Scale Robot Dataset

What's a **Robot Dataset**?

- Data recorded by **robot embodiments** solving diverse tasks in real-world.
- Any data from **any embodiments** (including humans) that can be helpful training robot models.



Any expected issues when training robots with these datasets?

Road to Large-Scale Robot Dataset

What's a **Robot Dataset**?

- Data recorded by **robot embodiments** solving diverse tasks in real-world.
- Any data from **any embodiments** (including humans) that can be helpful training robot models.



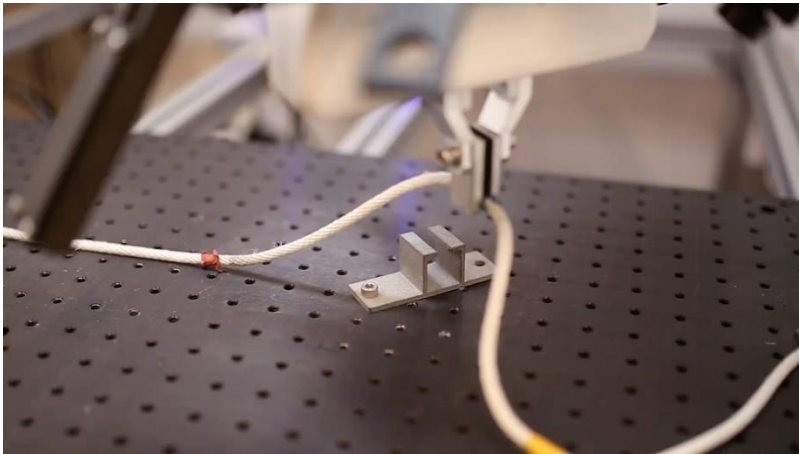
Haoshu Fang

**Topic we're dealing with
next Wednesday!**

Road to Large-Scale Robot Dataset

What's a **Robot Dataset**?

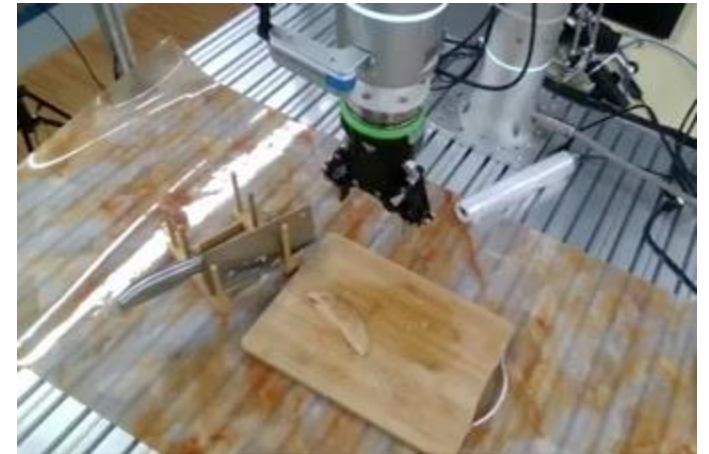
- Data recorded by **robot embodiments** solving diverse tasks in real-world.



O'Neill, Abby, et al. "**Open x-embodiment**: Robotic learning datasets and rt-x models." *arXiv:2310.08864* (2023).

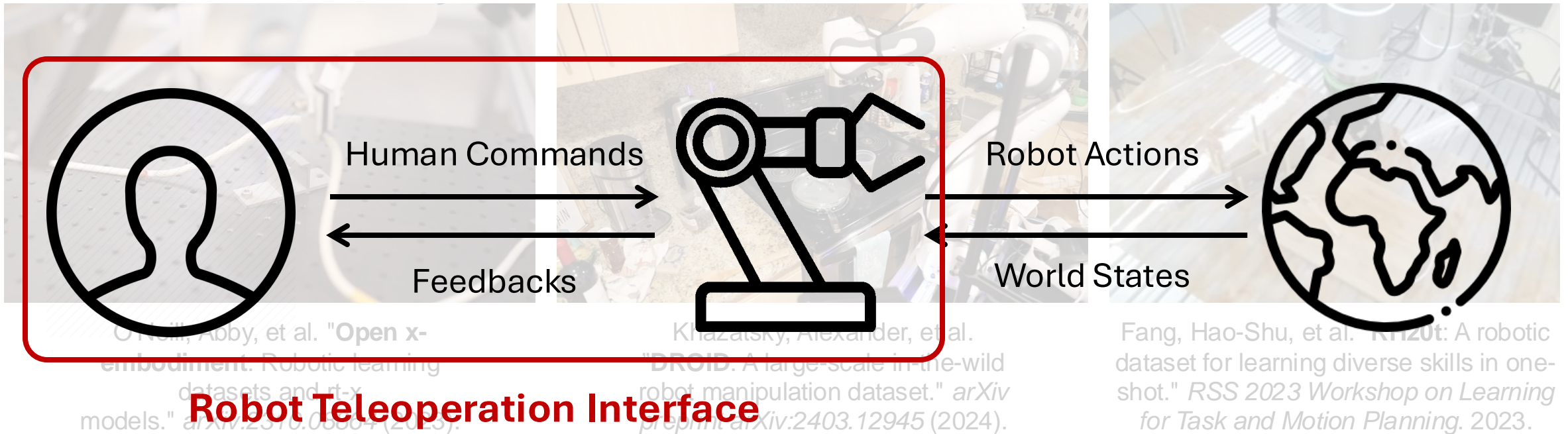


Khazatsky, Alexander, et al. "**DROID**: A large-scale in-the-wild robot manipulation dataset." *arXiv preprint arXiv:2403.12945* (2024).



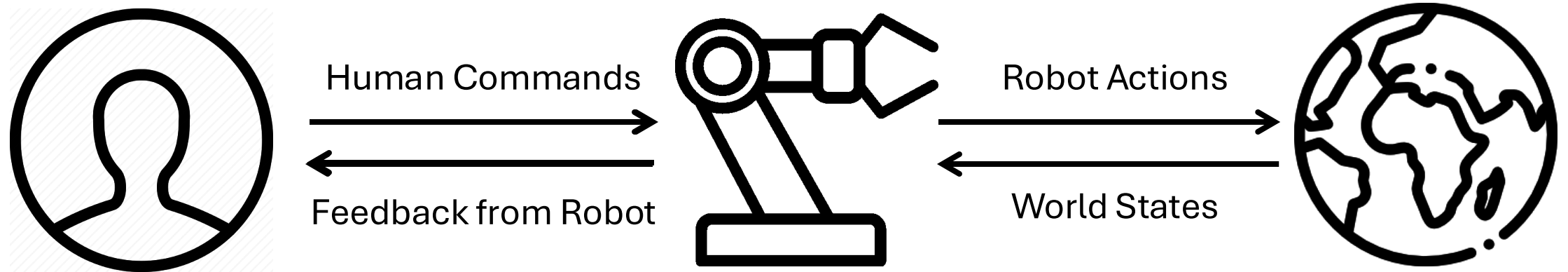
Fang, Hao-Shu, et al. "**RH20t**: A robotic dataset for learning diverse skills in one-shot." *RSS 2023 Workshop on Learning for Task and Motion Planning*. 2023.

- Data recorded by **robot embodiments** solving diverse tasks in real-world.



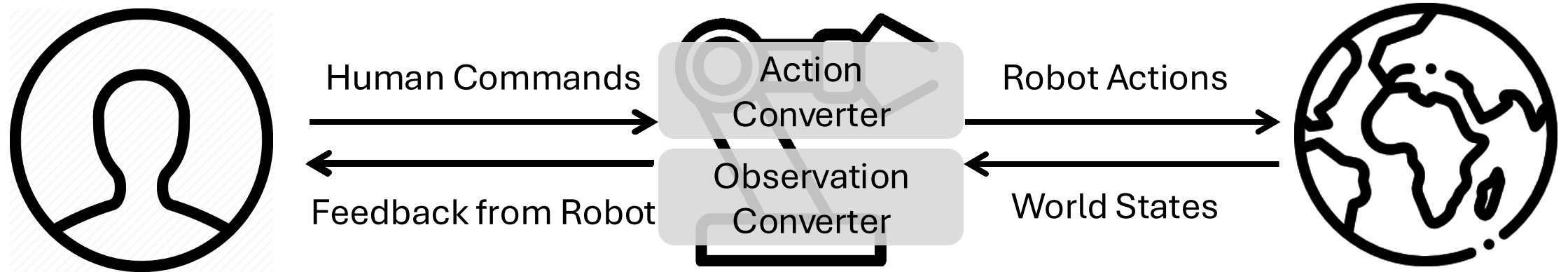
How do we make robots solve tasks?

Robot Teleoperation



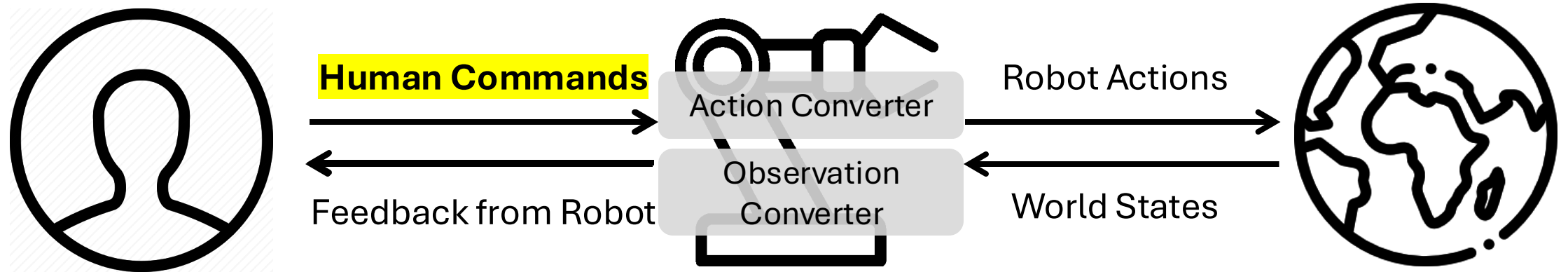
What should we consider when designing a **teleoperation system for robot**?

Robot Teleoperation



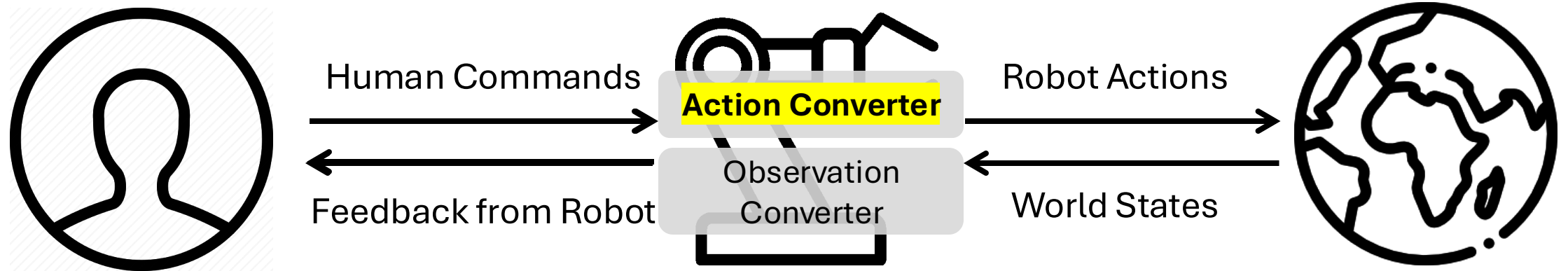
Robot Teleoperation

1. Designing command space for humans



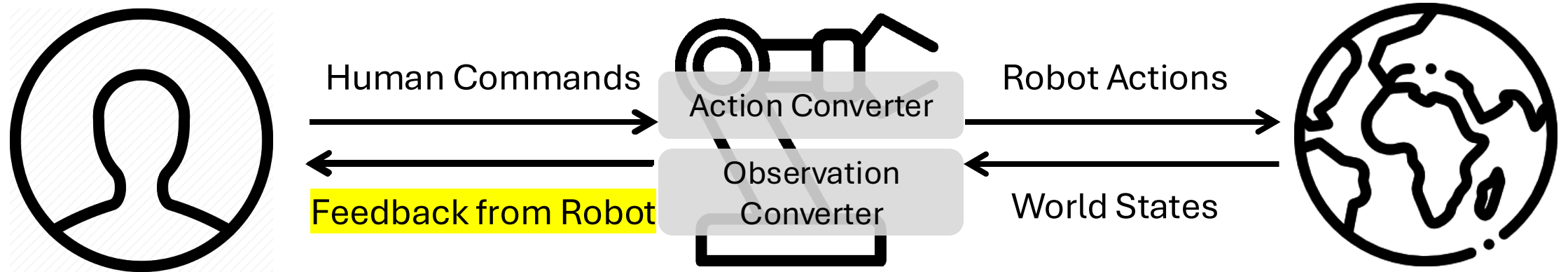
Robot Teleoperation

1. **Designing** command space for humans
2. **Converting** commands to robot actions



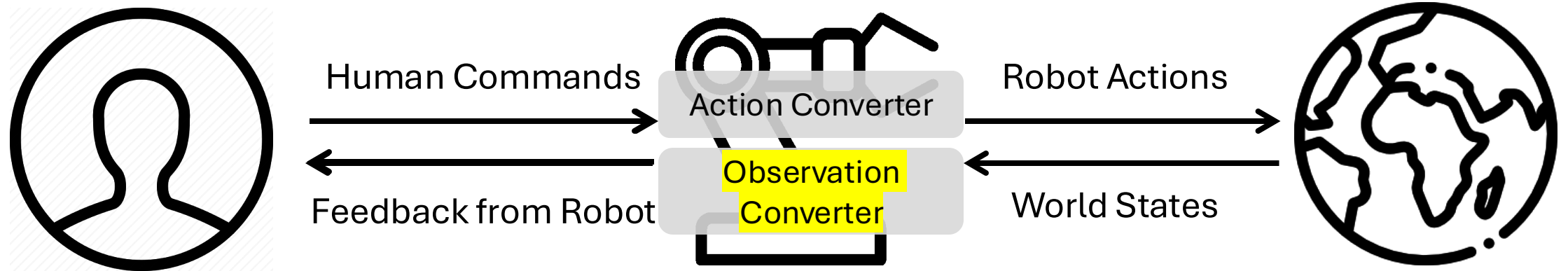
Robot Teleoperation

1. **Designing** command space for humans
2. **Converting** commands to robot actions
3. **Designing** feedback space for humans



Robot Teleoperation

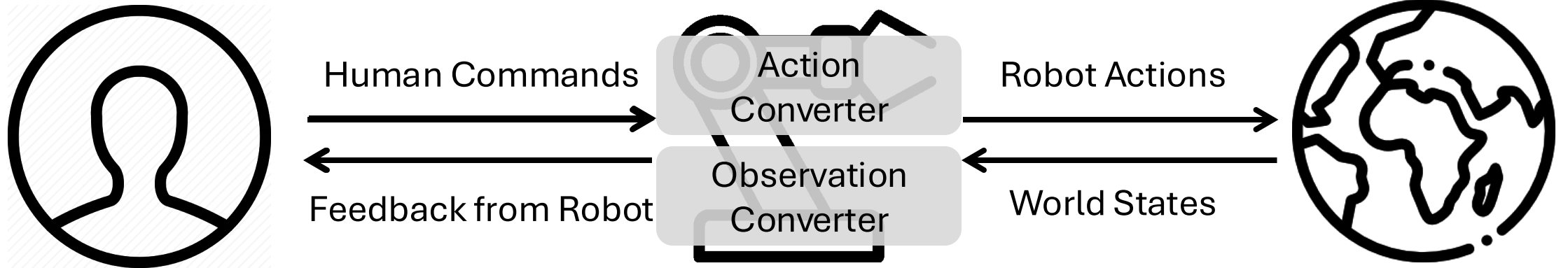
1. **Designing** command space for humans
2. **Converting** commands to robot actions
3. **Designing** feedback space for humans
4. **Converting** robot perceptions to human feedback



Robot Teleoperation

4 Key Elements of
Teleoperation System

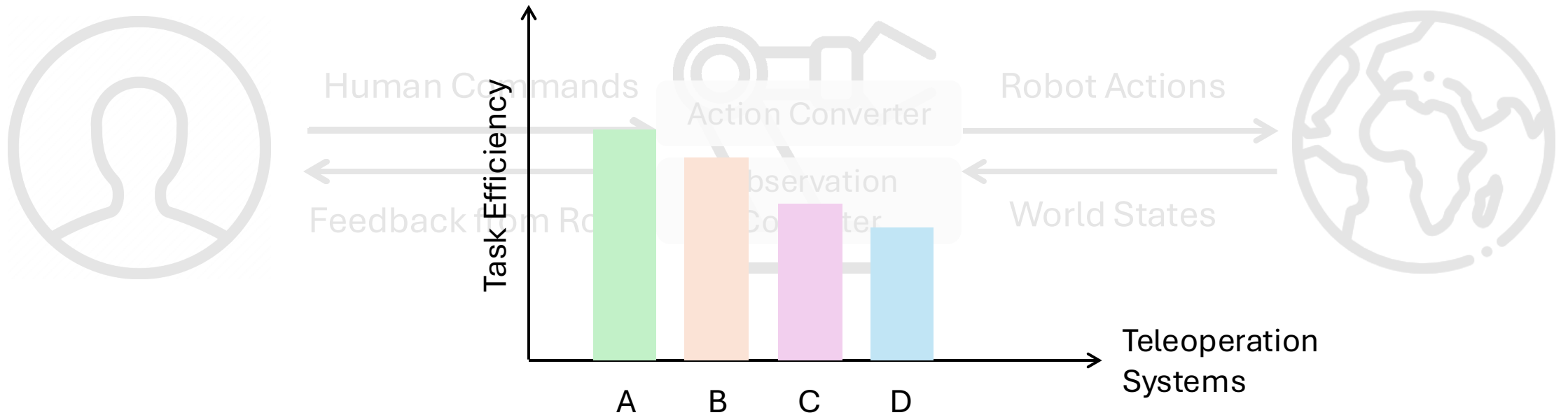
1. **Designing** command space for humans
2. **Converting** commands to robot actions
3. **Designing** feedback space for humans
4. **Converting** robot perceptions to human feedback



Robot Teleoperation

4 Key Elements of
Teleoperation System

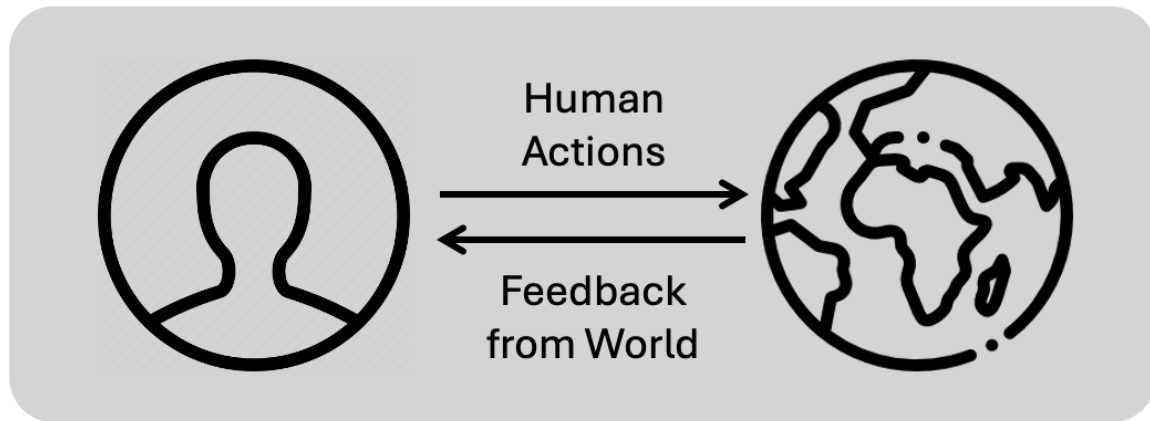
1. **Designing command** space for humans
2. **Converting commands** to robot actions
3. **Designing feedback** space for humans
4. **Converting** robot perceptions to human feedback



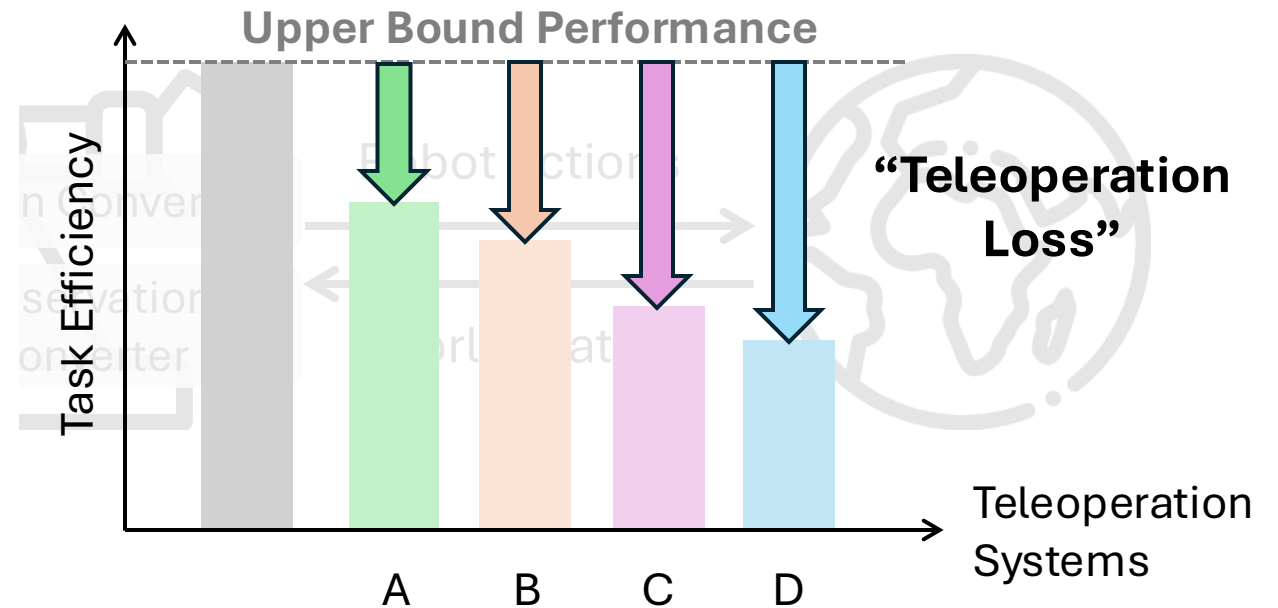
Robot Teleoperation

4 Key Elements of Teleoperation System

1. **Designing command** space for humans
2. **Converting commands** to robot actions
3. **Designing feedback** space for humans
4. **Converting** robot perceptions to human feedback

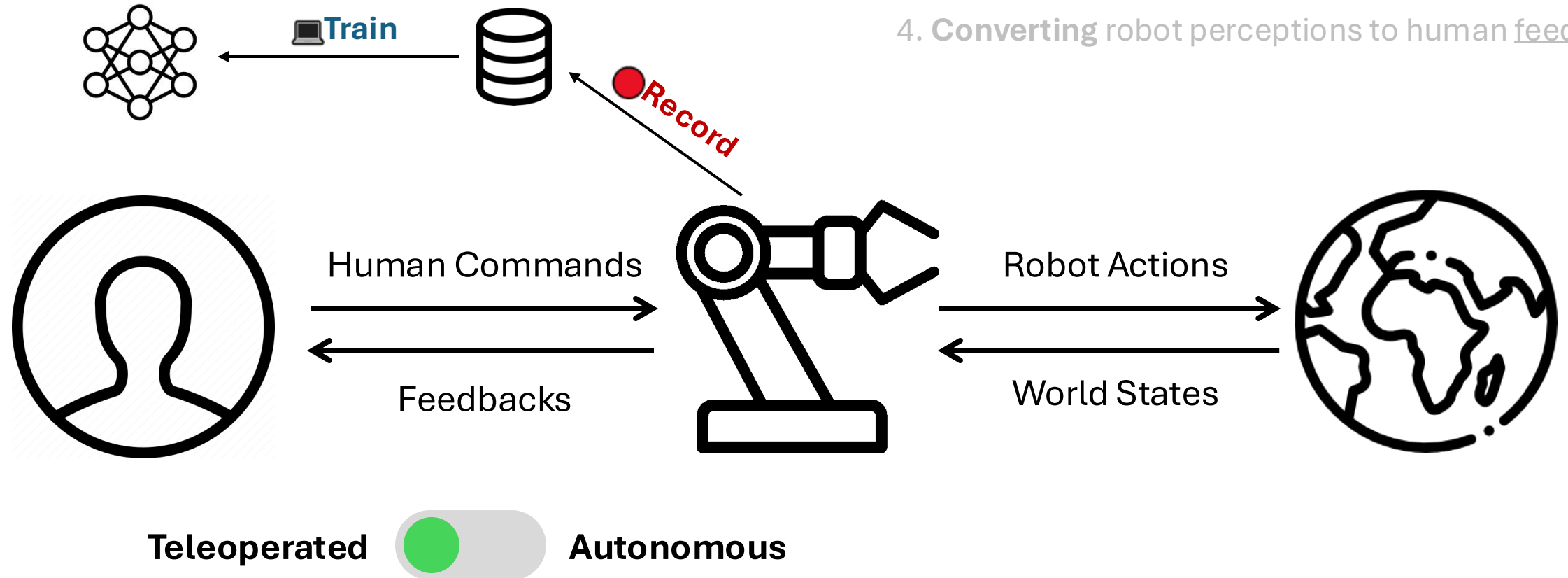


Humans doing things with their **own hands**



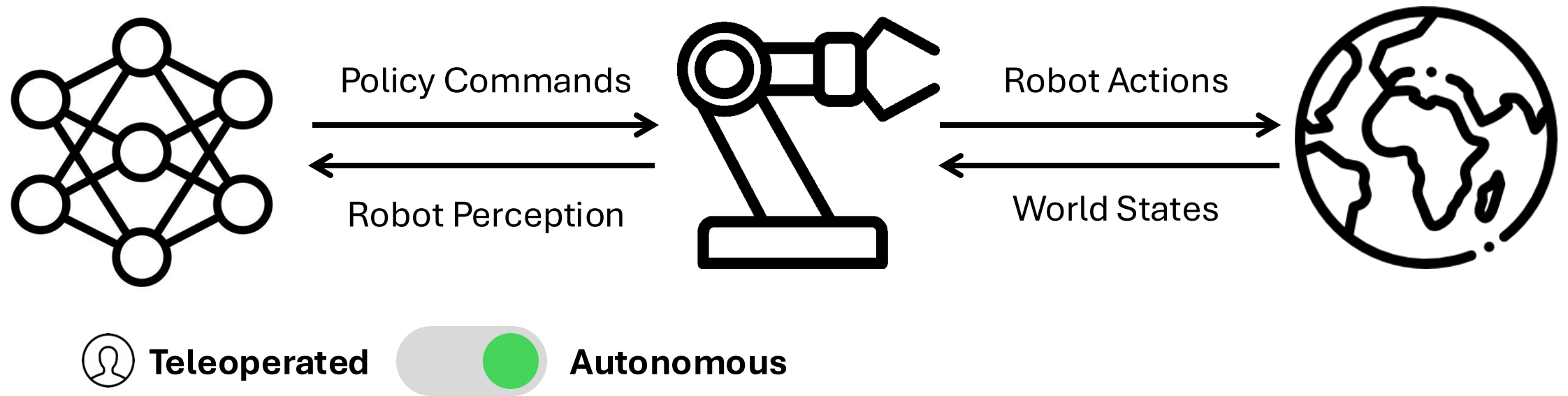
Robot Teleoperation for robot learning pipelines

1. Designing command space for humans
2. Converting commands to robot actions
3. Designing feedback space for humans
4. Converting robot perceptions to human feedback



Robot Teleoperation for robot learning pipelines

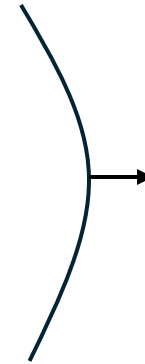
1. **Designing command** space for humans
2. **Converting commands** to robot actions
3. **Designing feedback** space for humans
4. **Converting robot perceptions to human feedback**



Next Week...

More on Robot Teleoperation

1. **Designing** command space for humans
2. **Converting** commands to robot actions
3. **Designing** feedback space for humans
4. **Converting** robot perceptions to human feedback



Case Study with Examples

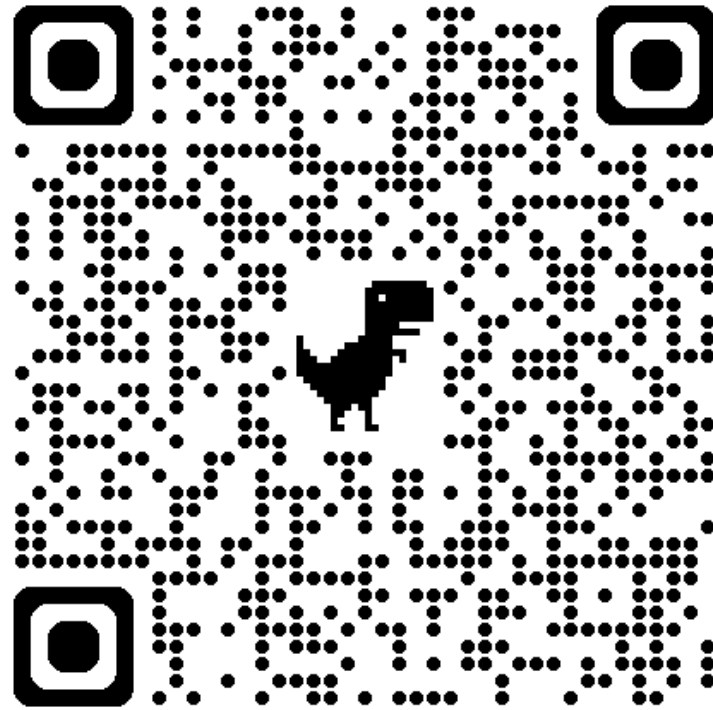
Policy Training

- Behavior Cloning with Action Chunks
- Policy Architectures
- Policy Evaluations

Role of Simulation

- Reinforcement Learning in Simulation with Demonstrations
- Real2Sim / Sim2Real

See you next week!



Fill out a survey!