CS 343H

Review and Conclusion

Prof. Yuke Zhu

The University of Texas at Austin

TEXAS

The University of Texas at Austin

Announcements

- Please fill out the course survey
  - Feedback to both instructor and TAs
  - Positive and negative points are useful
  - Post on Ed your completion screenshot (in a private post) as a form of participation!

- Capture the Flag contest results!

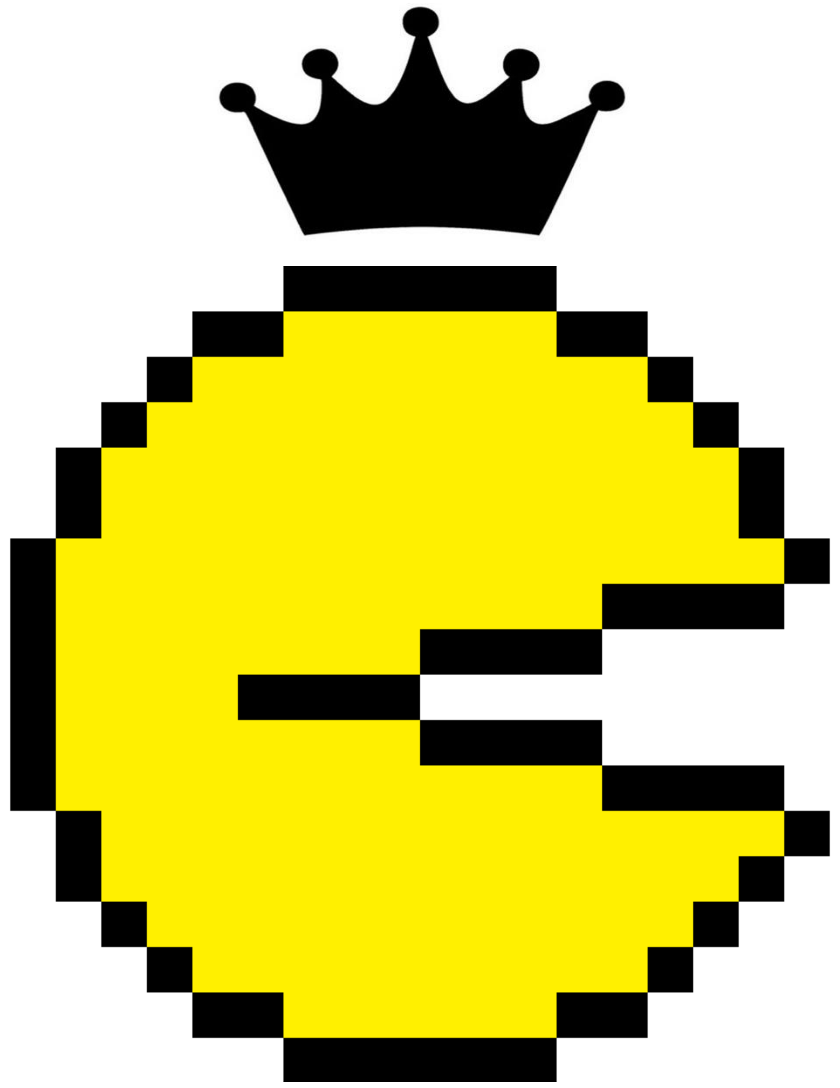
CTF Contest

23 teams participated, 15 qualified for tournament

Rankings

- 1st place: Beto An, Kevin Zhao (controllers)
- 2nd place: Angela Zhang, Tanuj Tekkale (lords of smiggles)
- 3rd place: Rishi Astra (temp_yeet2)

Congratulations!



Overview of AI Topics

Search / Planning

- Uninformed Search
- A* Search
- CSPs
- Local Search
- Minimax
- Expectimax
- MDPs

Machine Learning

- Reinforcement Learning
- Probability Theory
- Bayes Nets
- HMMs
- Particle Filters
- Decision Networks
- VPI
- Naive Bayes
- Perceptrons
- Neural Networks
- Deep Learning
- Clustering

Overview of Machine Learning

Supervised Learning

Discriminative Models

Perceptrons

Neural Networks

Generative Models

Bayes Nets

Naive Bayes

HMMs

Reinforcement Learning

MDPs

Value Iteration

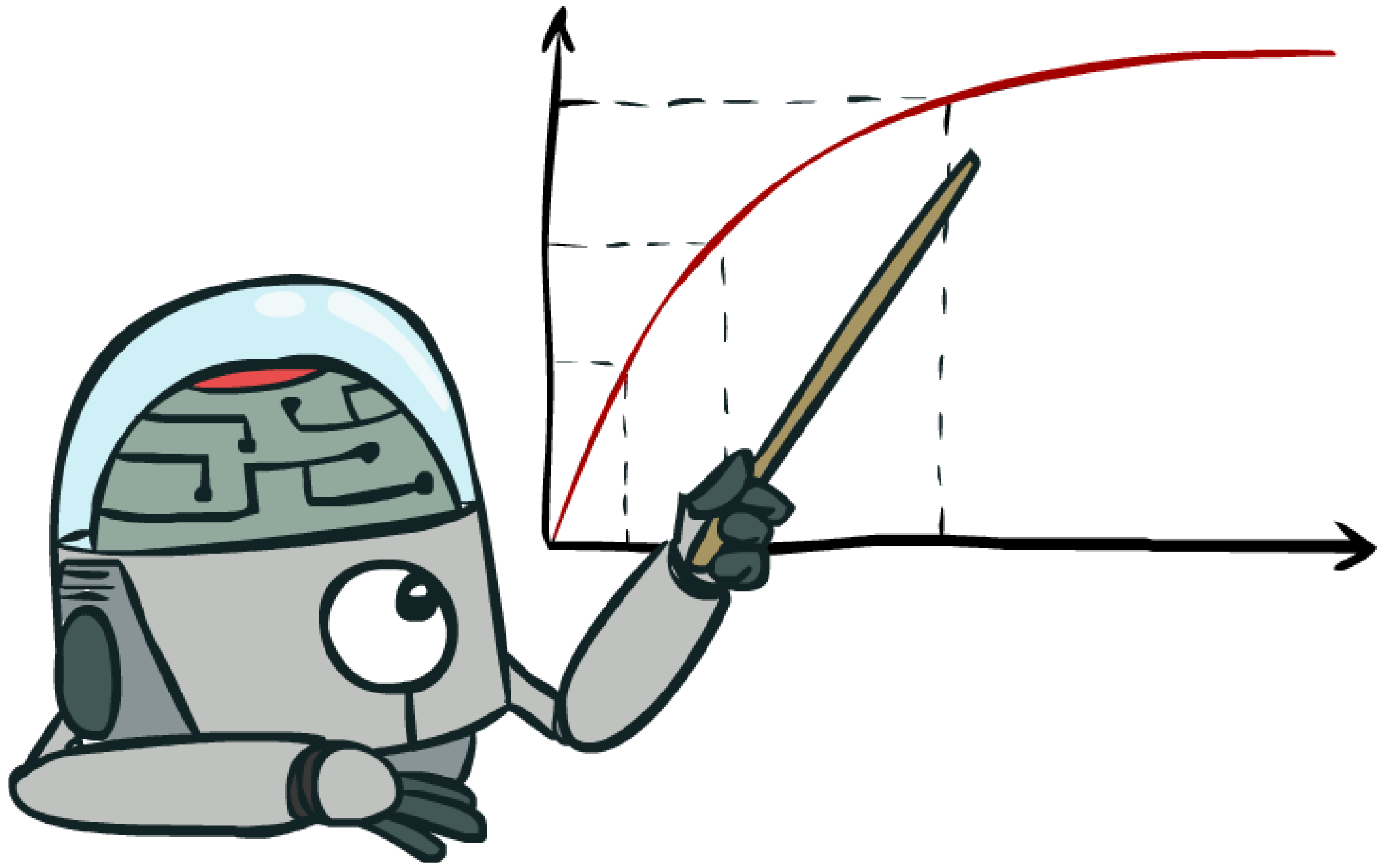
Policy Iteration

Q Learning

Unsupervised Learning

K-Means Clustering

Maximize Your Expected Utility



How Do AI Systems Maximize Utility?

Constraint satisfaction: searching intelligently for legal solutions

8			4		6			7
						4		
	1					6	5	
5		9		3		7	8	
				7				
	4	8		2		1		3
	5	2					9	
		1						
3			9		2			5

Example: Sudoku

Utility: Does the solution satisfy the rules / constraints?

Assumptions: We can write down the rules / constraints of the problem

How Do AI Systems Maximize Utility?

Planning: reasoning with models

Example: Robot navigation

Utility: Path length, collisions, surfaces, energy, social factors

Assumptions: We have a model of the world and the effects of the agent's actions

How Do AI Systems Maximize Utility?

Supervised Learning: learning from labeled examples

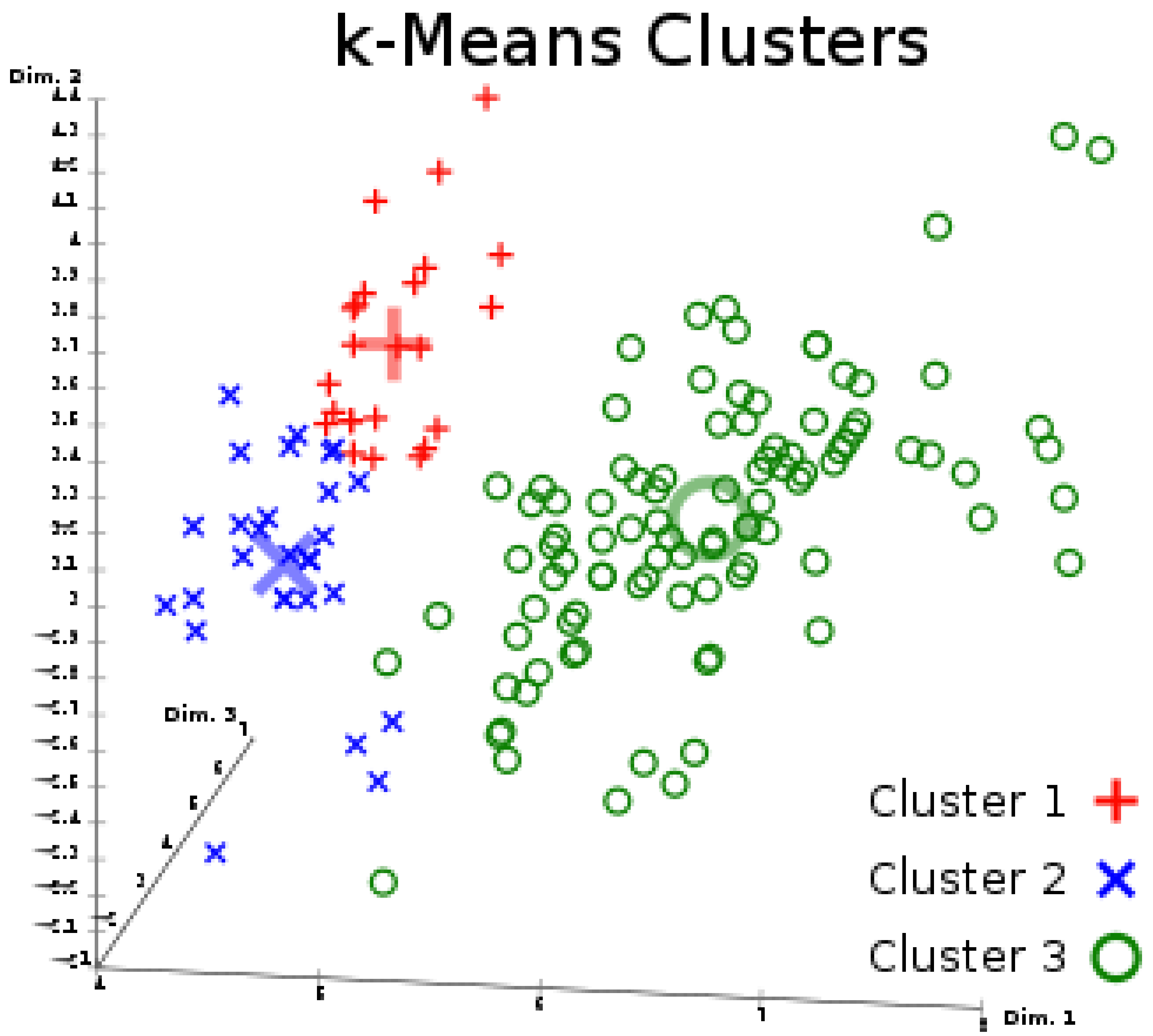
Example: Image classification

Utility: Classification accuracy on images not seen during training

Assumptions: We have access to a (usually large) labeled data set

How Do AI Systems Maximize Utility?

Unsupervised Learning: discovering patterns in unlabeled data



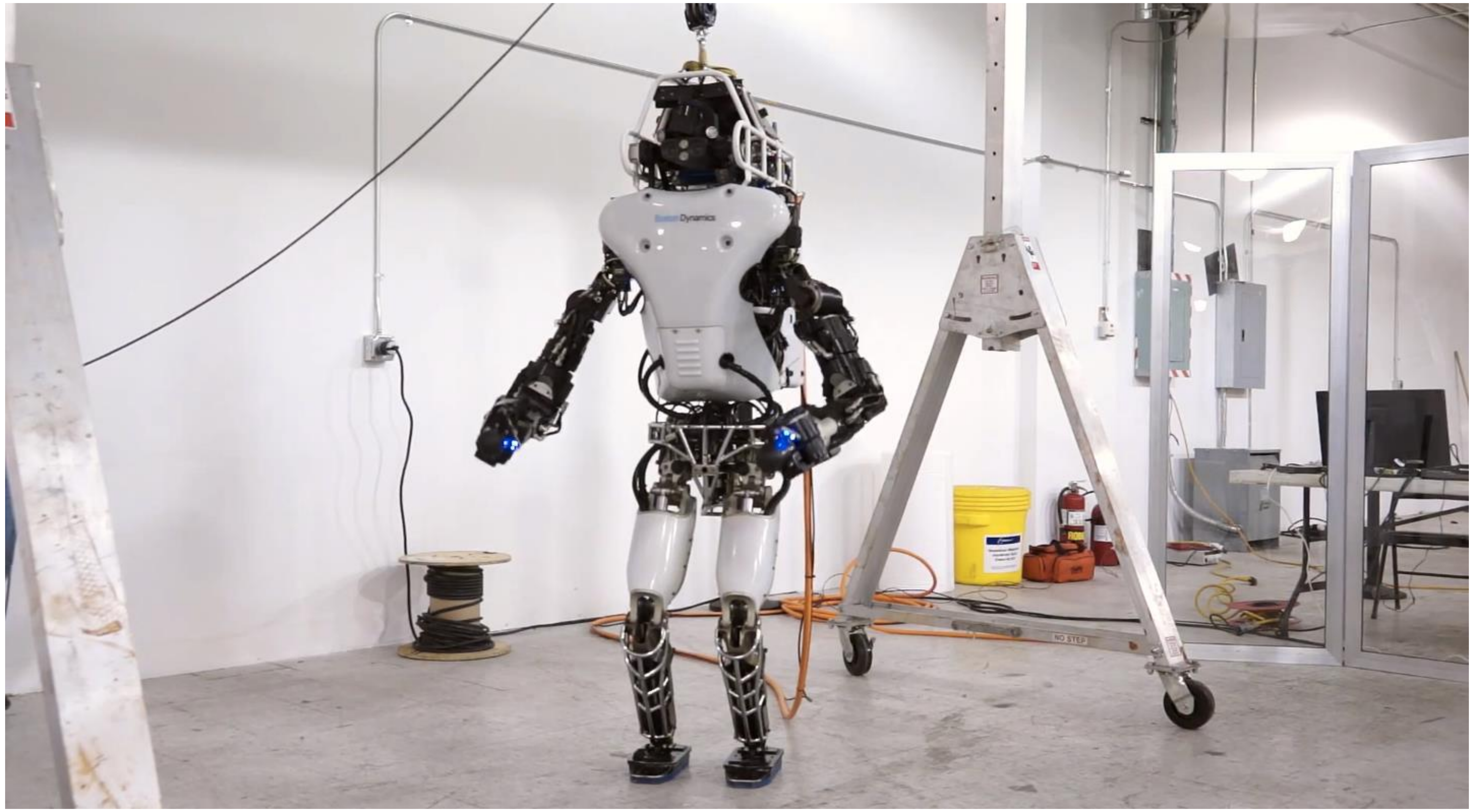
Example: Clustering species

Utility: “Best” explanation of data

Assumptions: Data points that should be clustered together are “close” together

How Do AI Systems Maximize Utility?

Reinforcement Learning: learning from experience



Example: Robot walking

Utility: Time until fall, speed, energy efficiency

Assumptions: We can "reward" and "punish" good and bad performance, but don't know what the correct action at each step should be

Properties of task environment

- Single-agent vs. multi-agent
- Deterministic vs. stochastic
- Fully observable vs. partially observable
- Episodic vs. sequential
- Static vs. dynamic
- Discrete vs. continuous
- Known vs. unknown

Single agent vs. multi-agent

- Not multi-agent if other agents can be considered part of the environment
- Only considered to be multi-agent if the agents are maximizing a performance metric that depends on other agents’ behavior
- Single agent example: Pacman with randomly moving ghosts
- Multi-agent example: Pacman with ghosts that use a planner to follow him

Single / Multi Agent

Single	Multi
Uninformed Search	Minimax
A* Search	Expectimax
Local Search	MDPs
CSPs	Reinforcement Learning

Deterministic vs. stochastic

- Deterministic: next state of environment is completely determined by the current state and the action executed by the agent
- Stochastic: actions have probabilistic outcomes
- Strongly related to partial observability — most apparent stochasticity results from partial observation of a deterministic system
- Example: Coin flip

Determinism

Deterministic	Stochastic
Uninformed Search	Expectimax
A* Search	MDPs
Local Search	Reinforcement Learning
CSPs	Decision Diagrams
Minimax	

Fully observable vs. partially observable

- Fully observable: agent's sensors give it access to complete state of the environment at all times
- Can be partially observable due to noisy and inaccurate sensors, or because parts of the state are simply missing from the sensor data
- Example: Perfect GPS vs noisy pose estimation
- Example: IKEA assembly while blindfolded

Almost everything in the real world is partially observable

Observability

Fully Observable

Uninformed Search

A* Search

Local Search

CSPs

Minimax

Expectimax

MDPs

Reinforcement Learning

Partially Observable

Bayes Nets

HMMs

Decision Networks

Known vs. unknown

- Agent’s state of knowledge about the “rules of the game” / “laws of physics”
- Known environment: the outcomes for all actions are given
- Unknown: agent has to learn how it works to make good decisions
- Possible to be partially observable but known (solitaire)
- Possible to be fully observable but unknown (video game)

Model of the World

Known

Uninformed Search

A* Search

Local Search

CSPs

Classic Planning

Minimax

Expectimax

MDPs

Value Iteration

Decision Diagrams

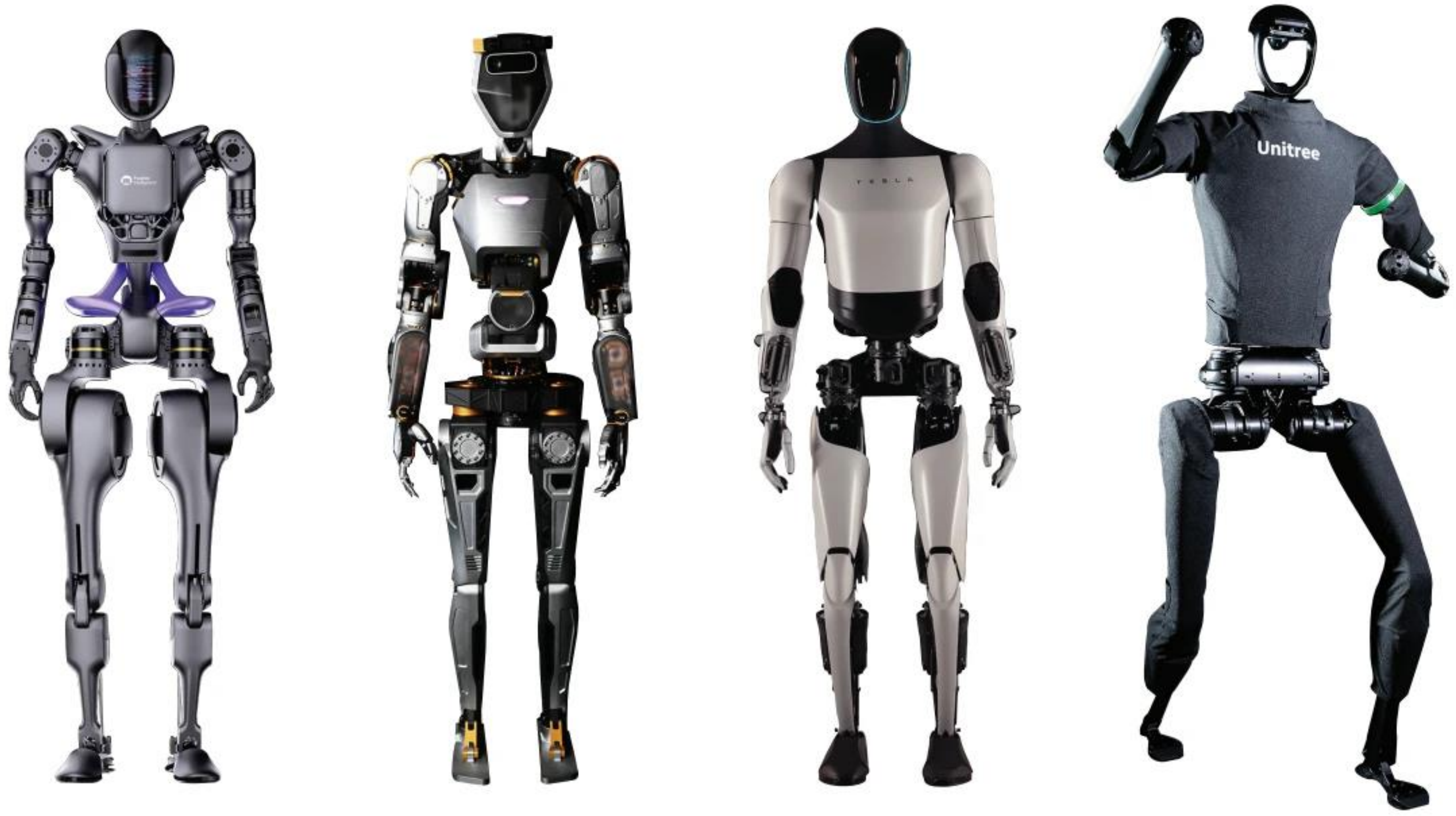
Unknown

Q Learning

Learning parameters of Bayes Net

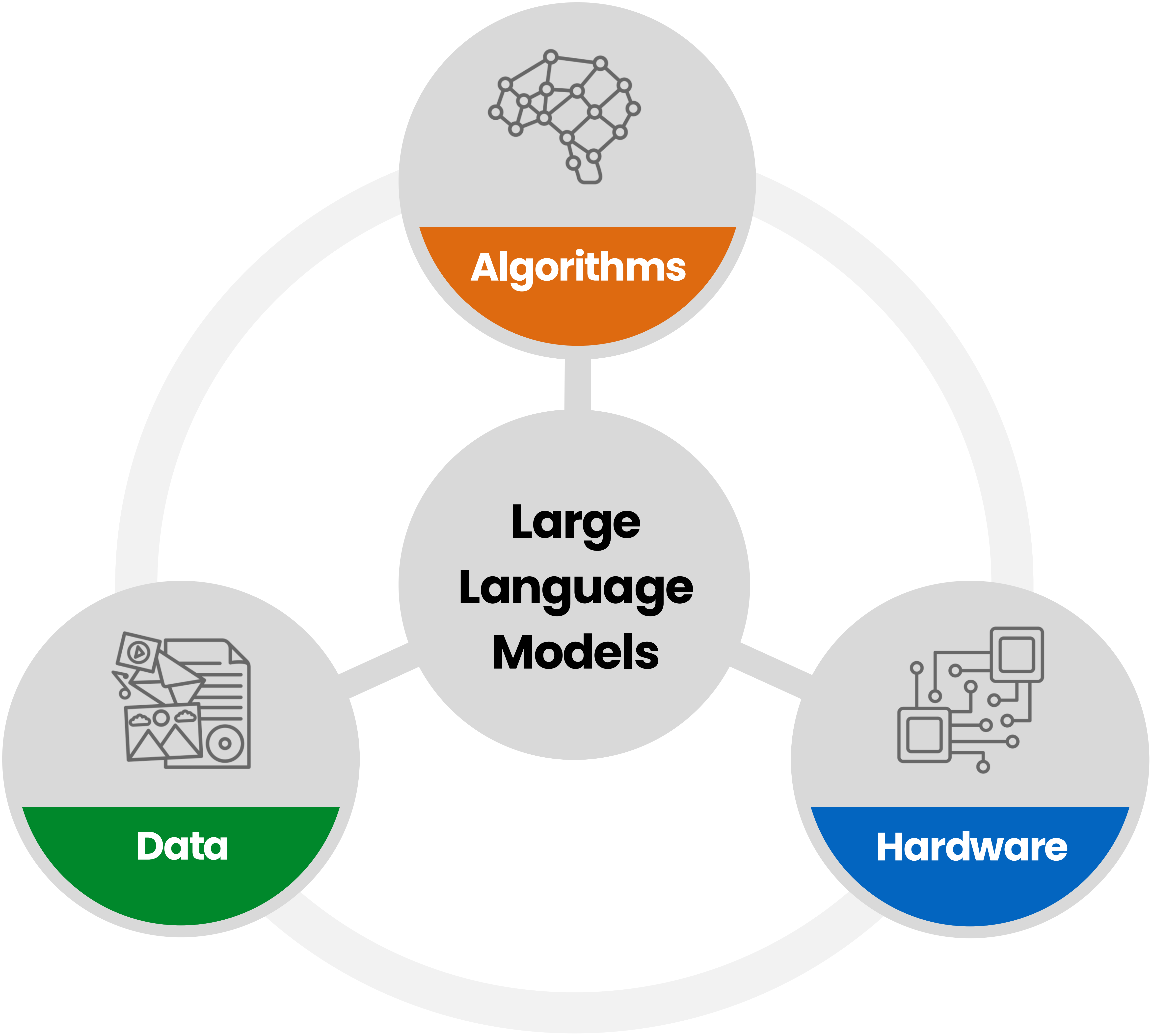

research.nvidia.com/labs/gear

Building Robotic Foundation Models

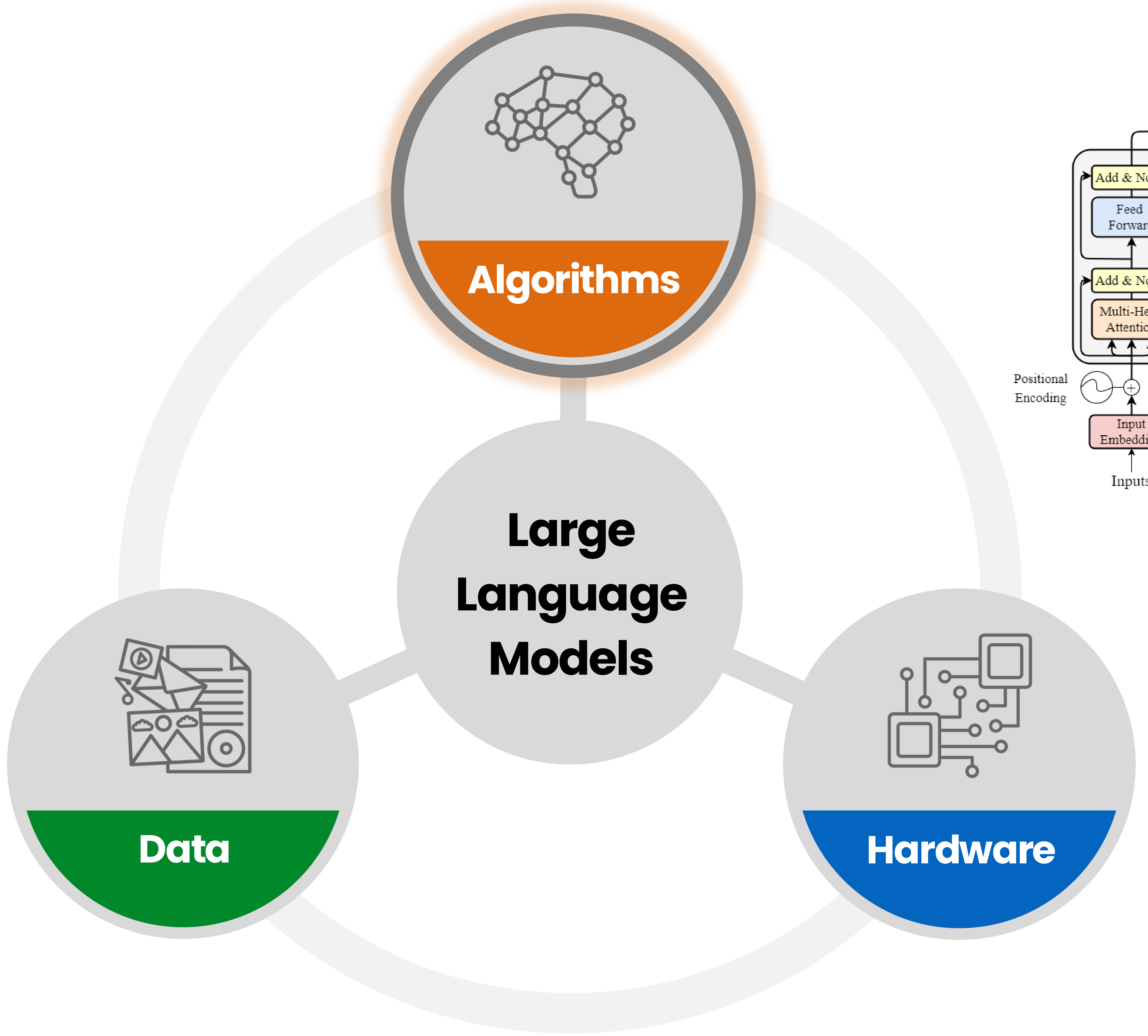


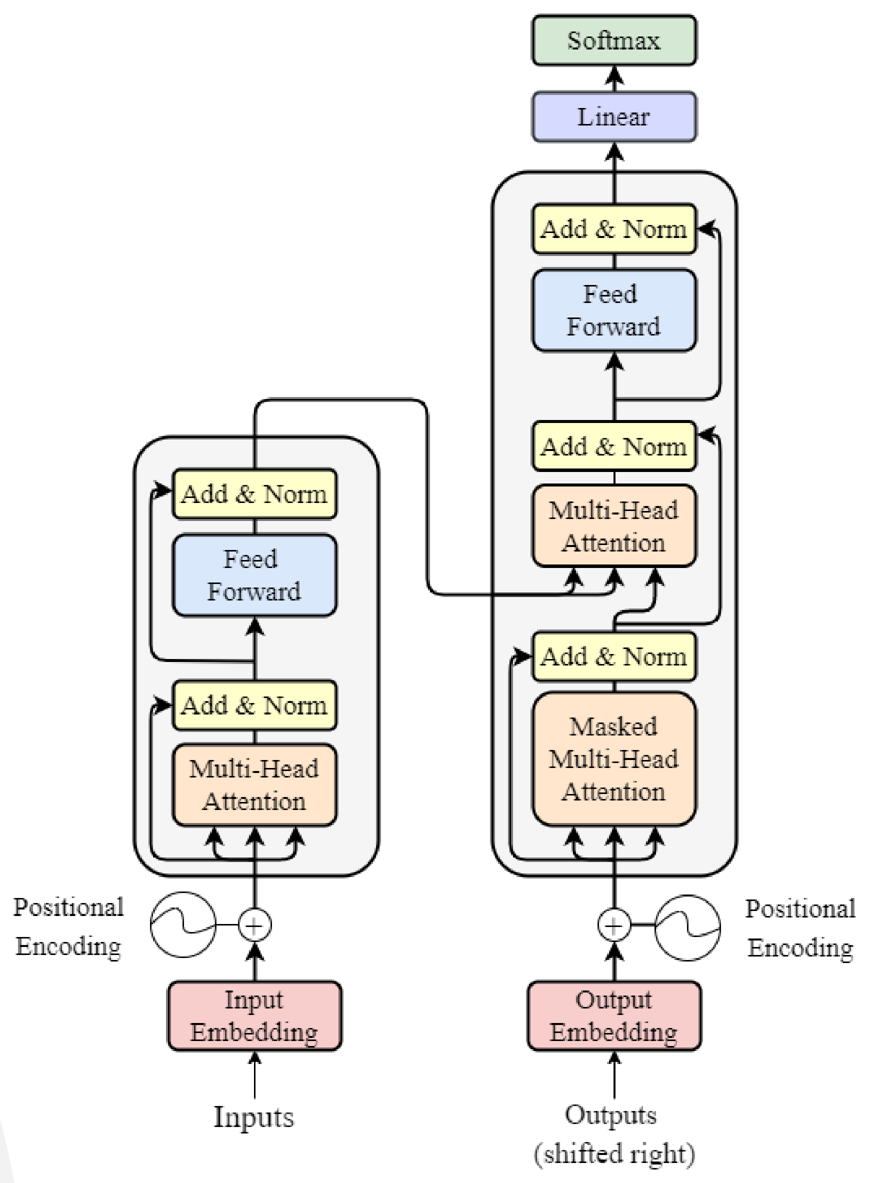
One “AI Brain” for All (Humanoid) Robots

Recipe for Building Large Language Models



Recipe for Building Large Language Models

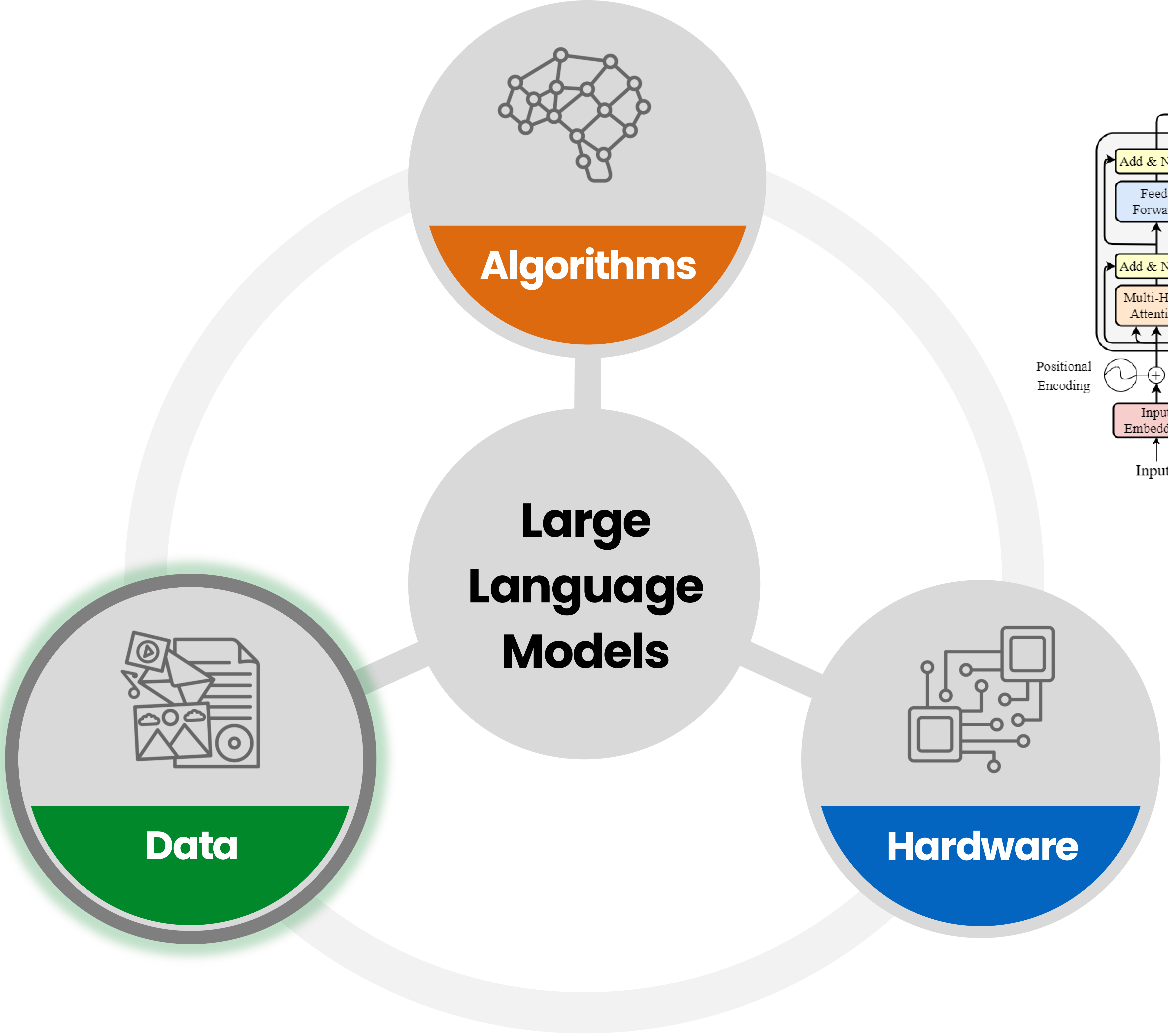


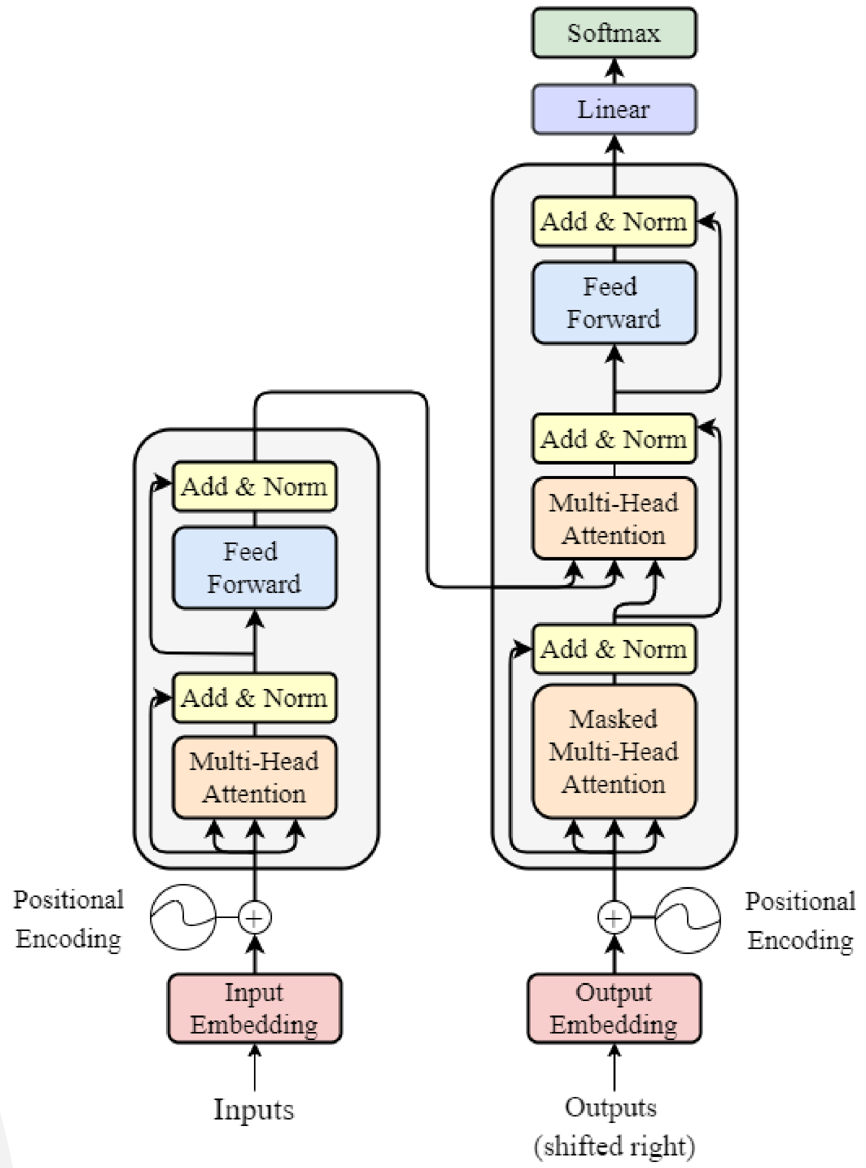


Recipe for Building Large Language Models

WIKIPEDIA
The Free Encyclopedia

GitHub

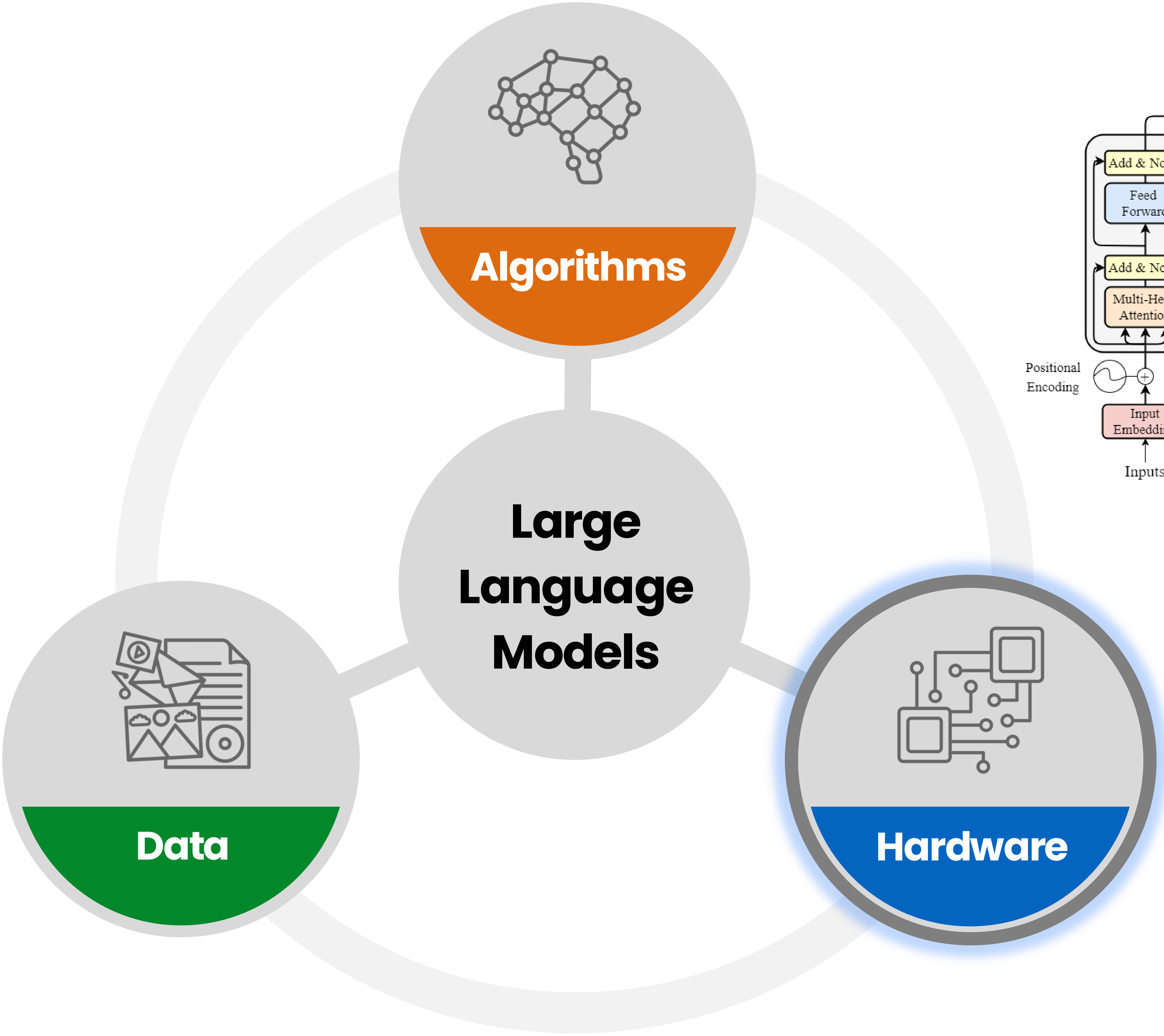


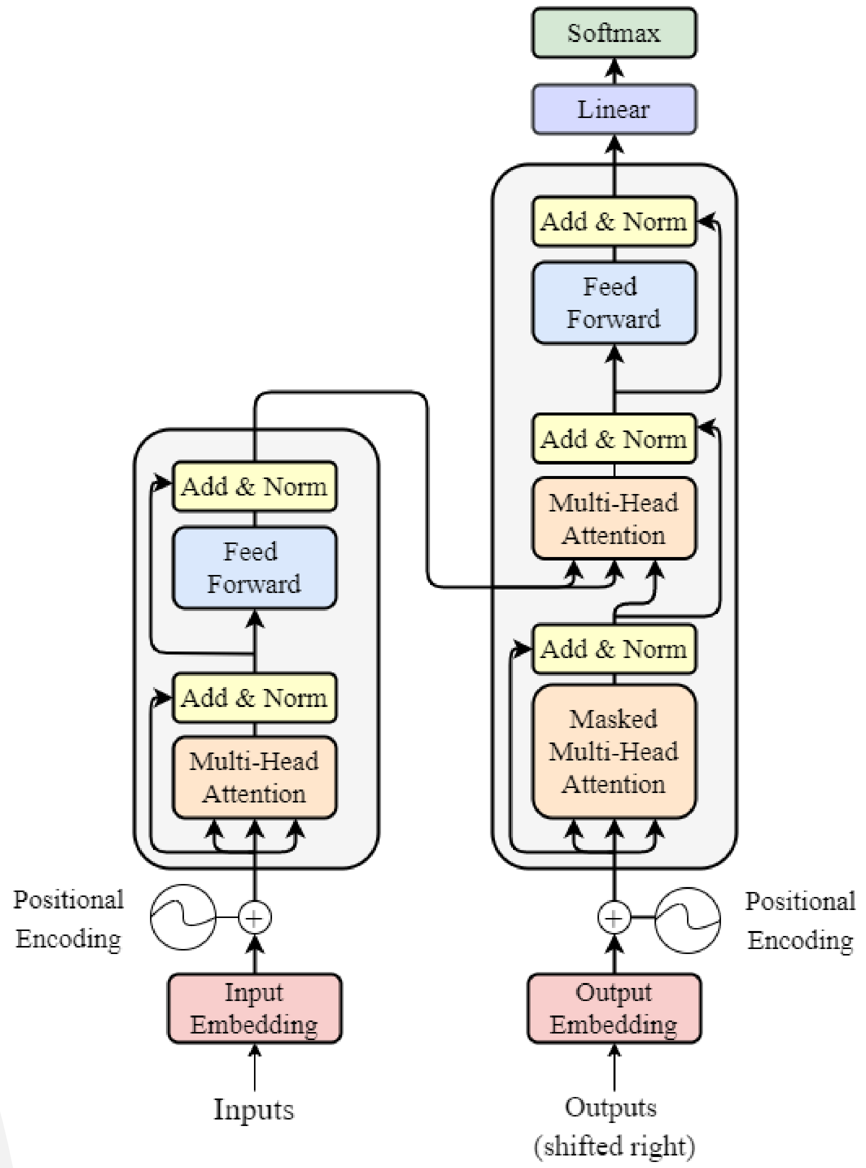


Recipe for Building Large Language Models

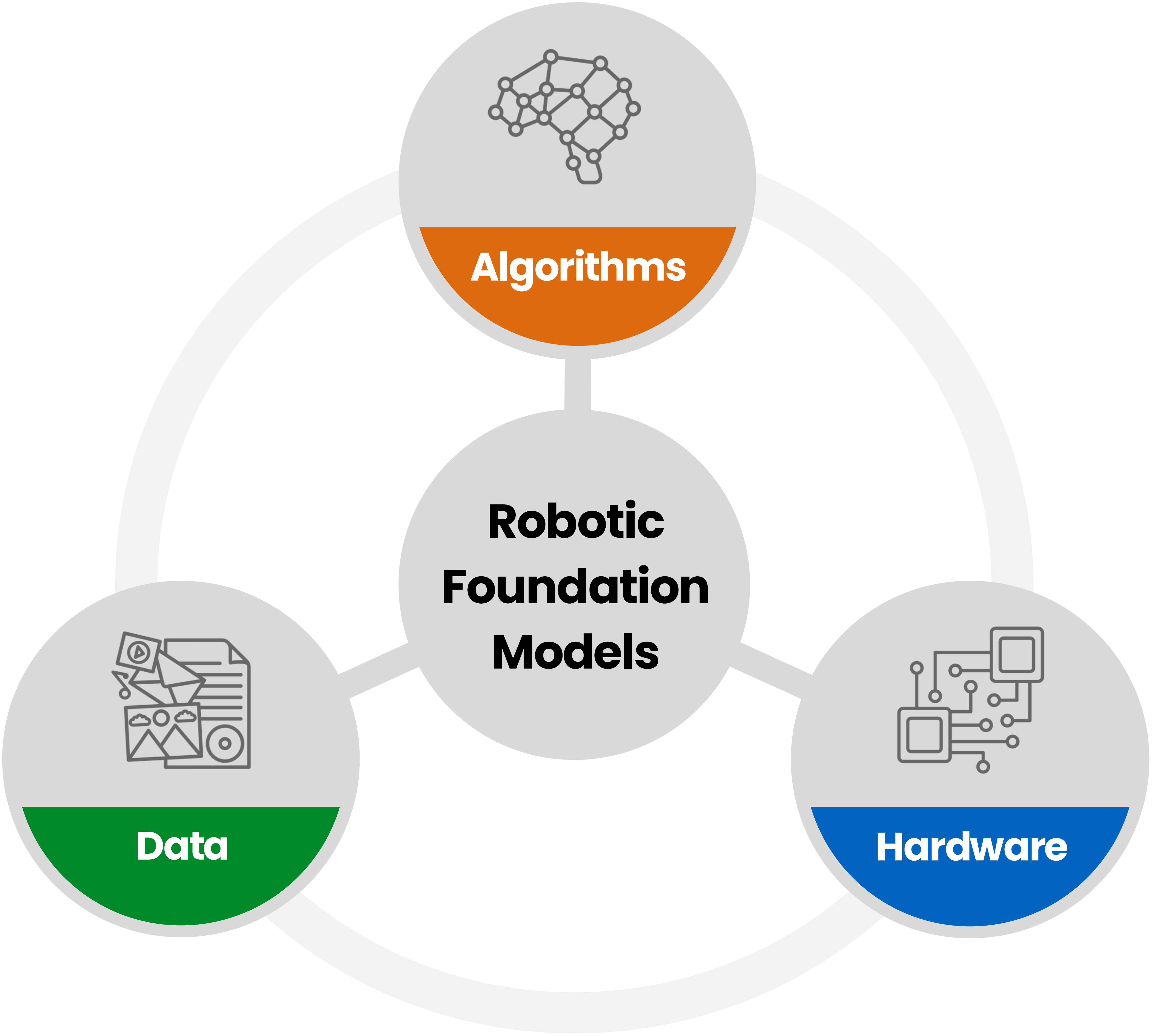
WIKIPEDIA
The Free Encyclopedia

GitHub





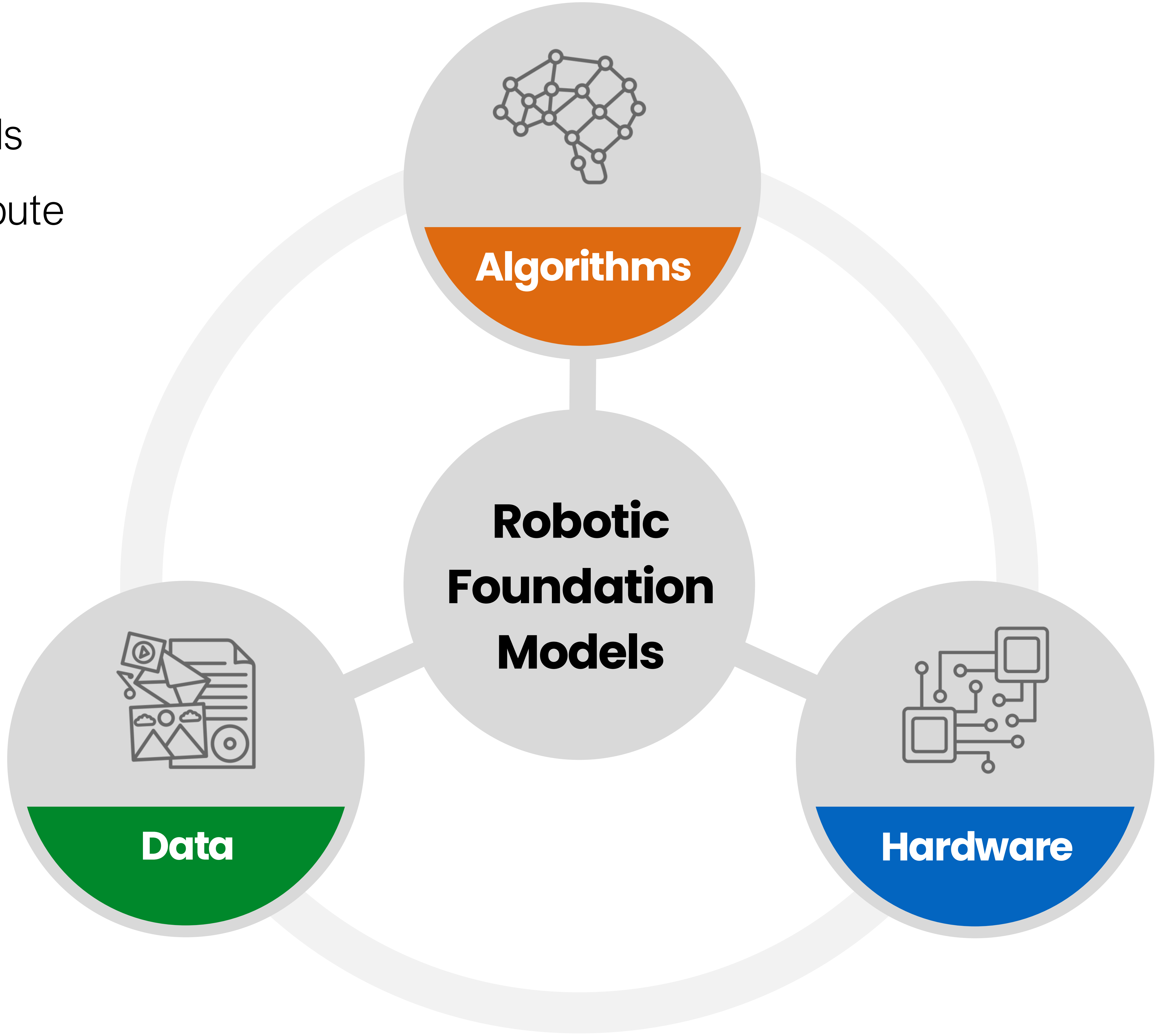
Recipe for Building Robotic Foundation Models



Recipe for Building Robotic Foundation Models

Scalable Algorithms

Powerful robot learning models that scale with data and compute



Data Engine

New mechanisms to produce massive training data

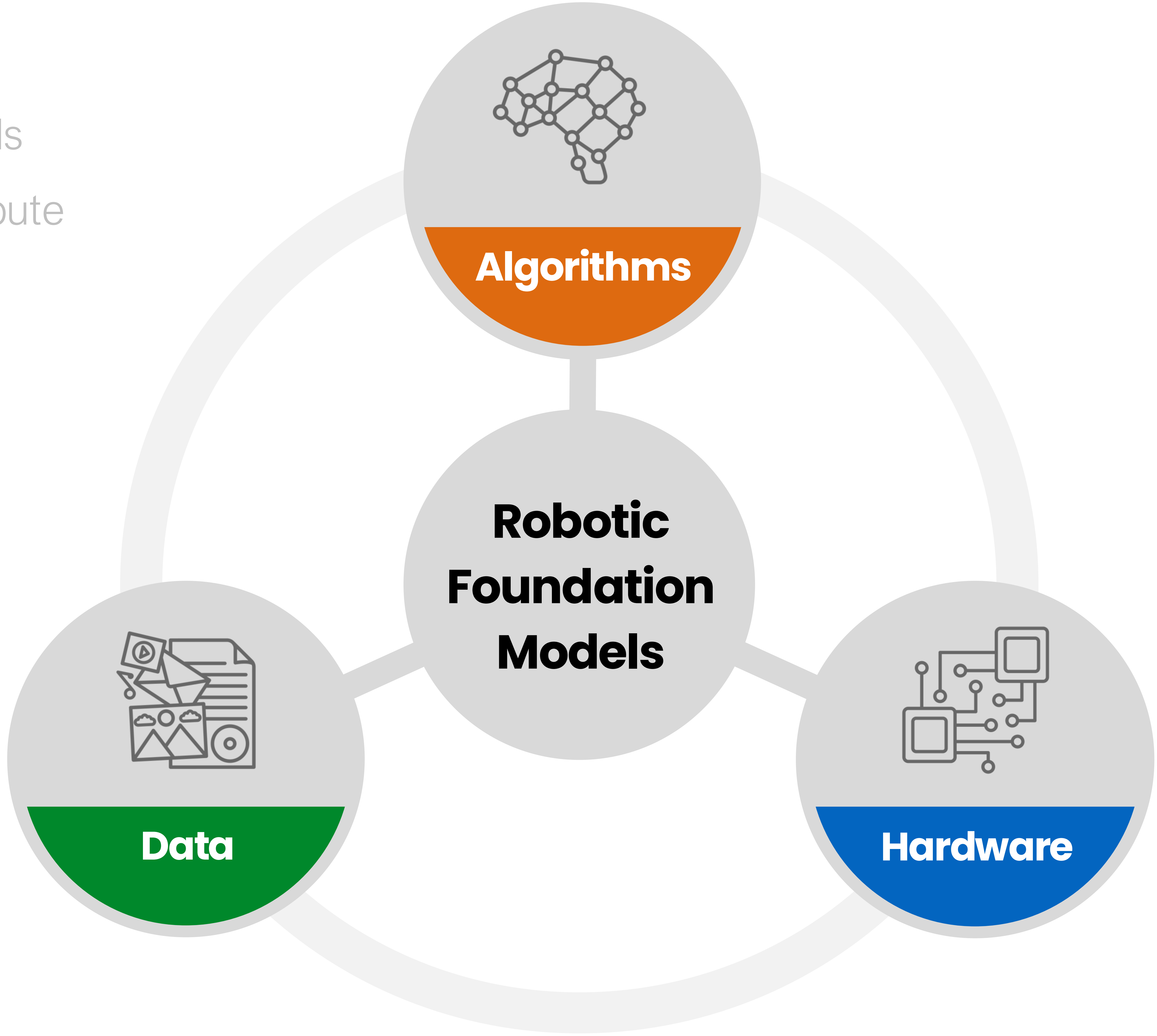
Human-like Embodiment

Humanoid robot platform for broad applications

Recipe for Building Robotic Foundation Models

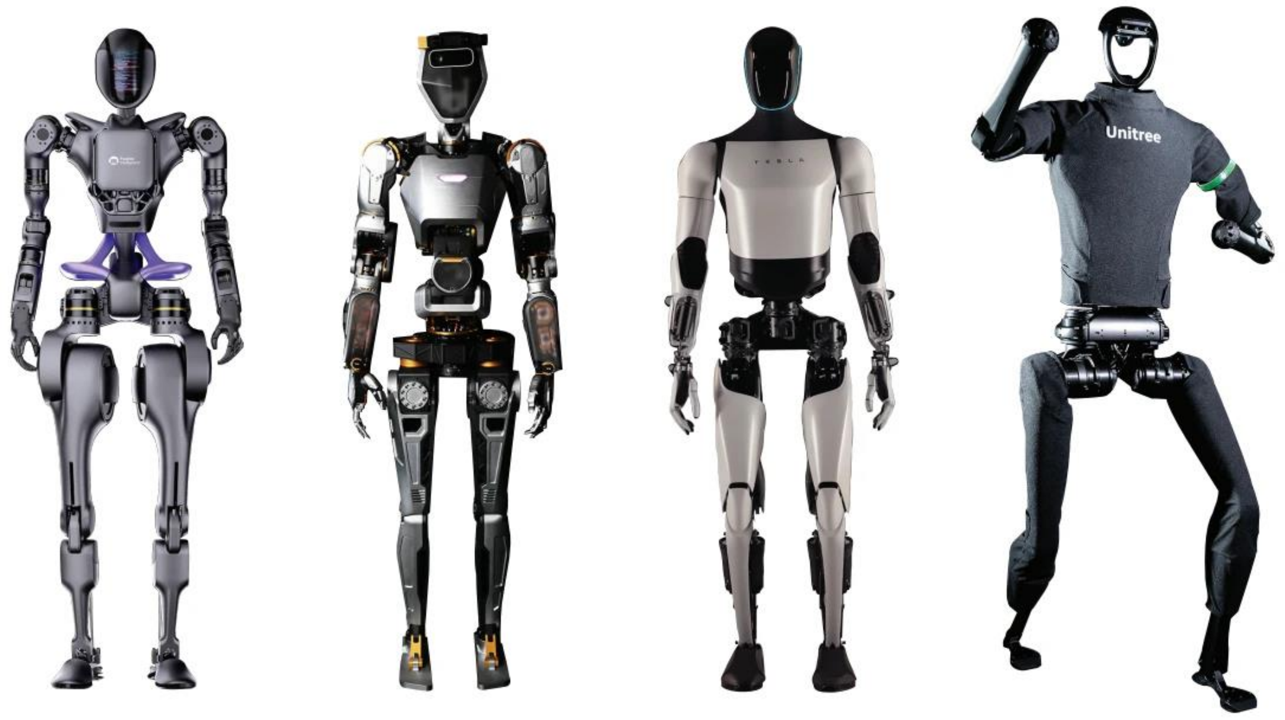
Scalable Algorithms

Powerful robot learning models
that scale with data and compute



Data Engine

New mechanisms to produce
massive training data



Human-like Embodiment

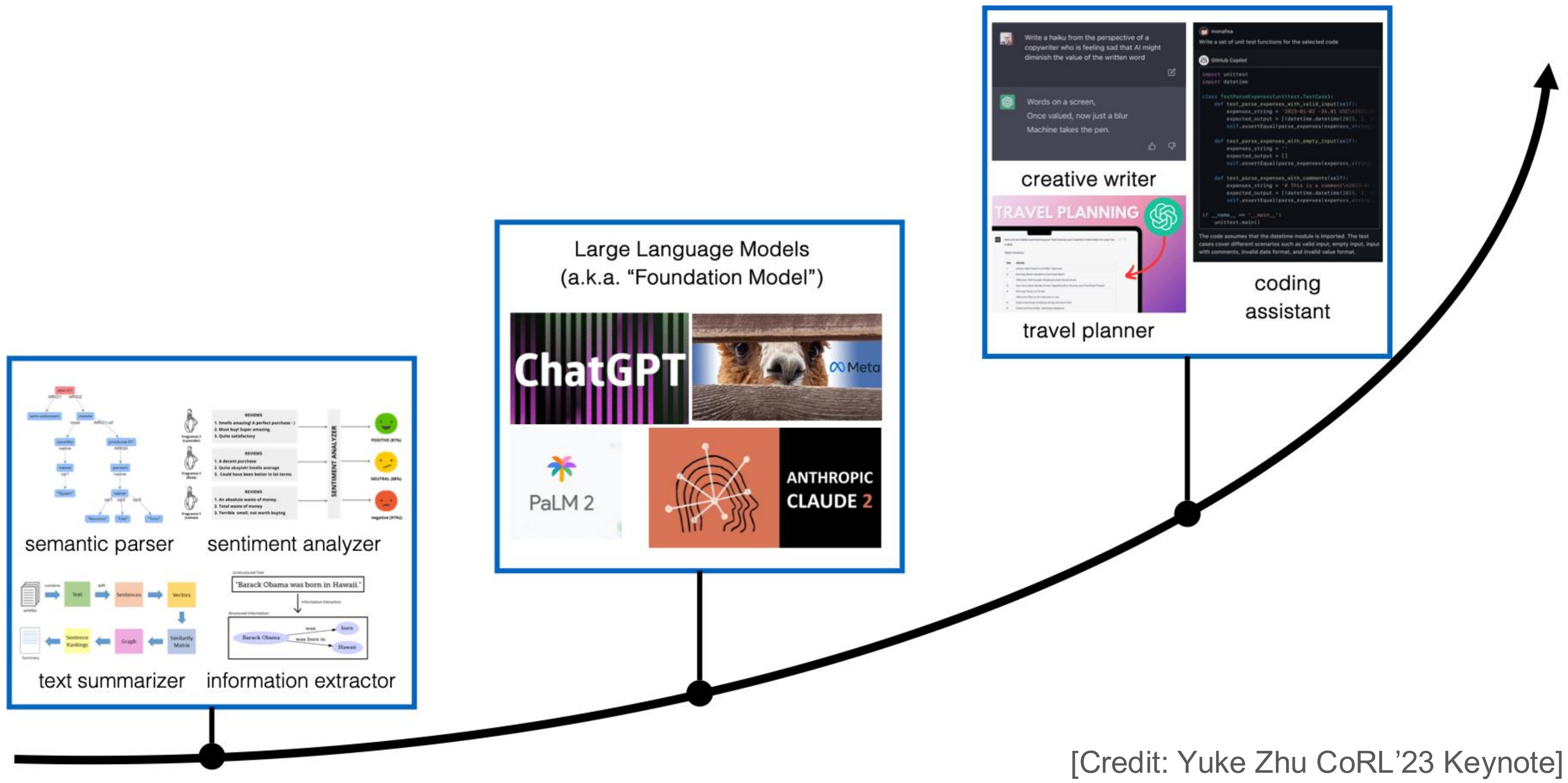
Humanoid robot platform for
broad applications

Why Humanoids?

- **Versatility:** General-purpose robot autonomy needs a versatile body.
- **Costs:** Hardware becomes cheaper and more robust to democratize transformative research.
- **Safety:** Humanoid robots can be more predictable and safer for human-robot interaction.
- **Data:** Their similar physique unlocks Internet-scale, human-centered data sources.
- …

Research Principle #1:

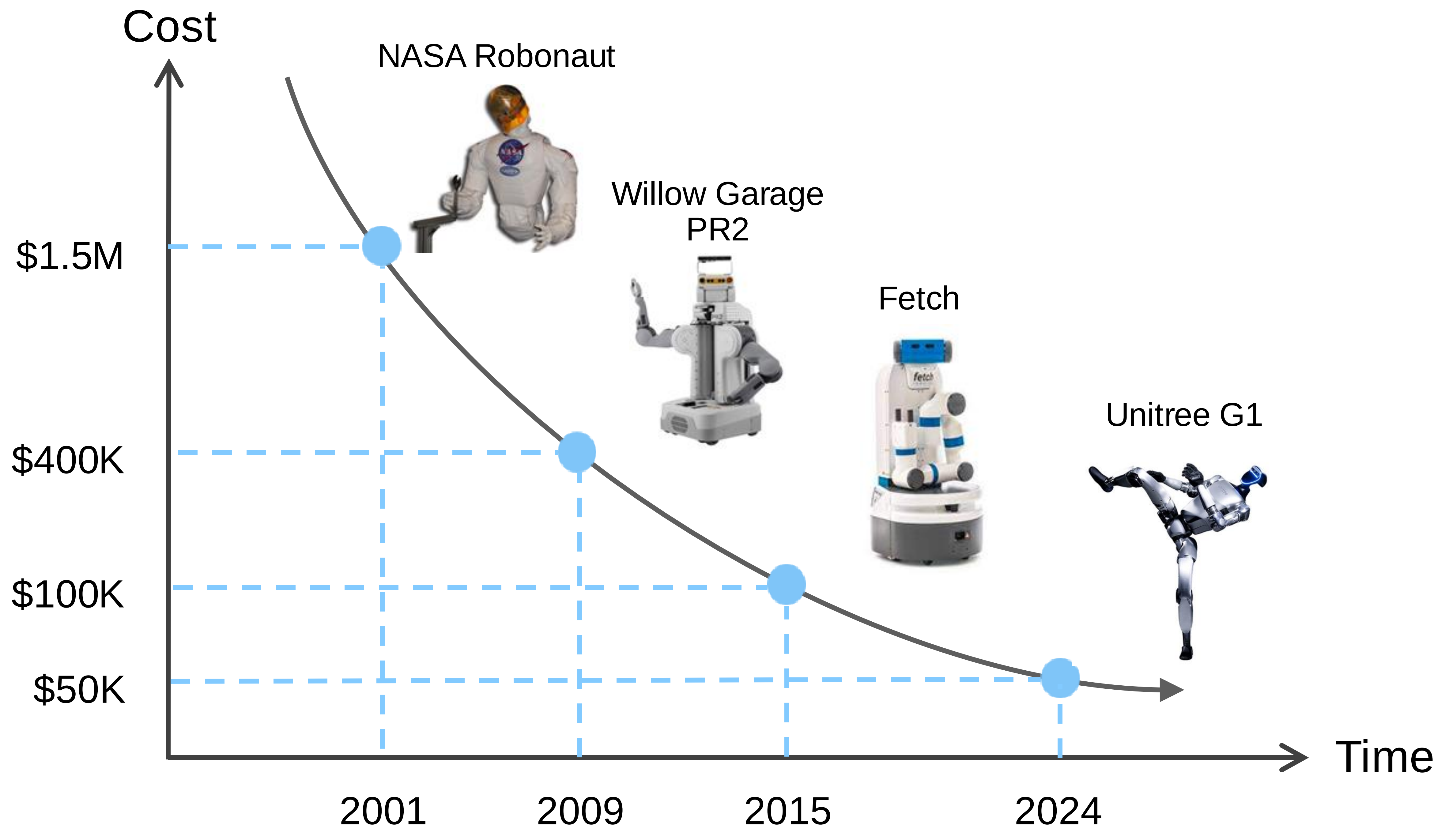
First Generalist, then Better Specialist



[Credit: Yuke Zhu CoRL'23 Keynote]

Why Humanoids?

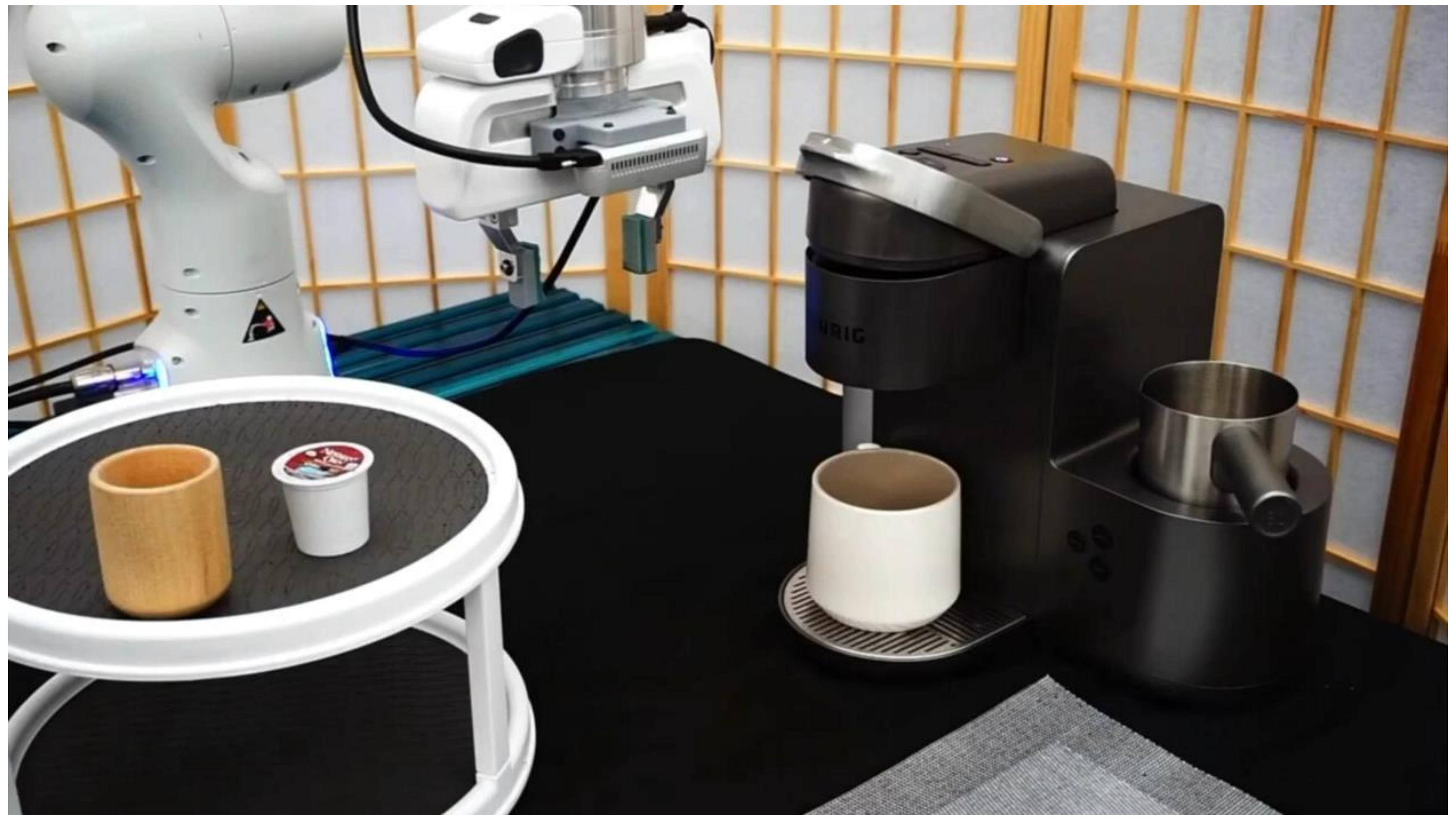
- **Versatility:** General-purpose robot autonomy needs a versatile body.
- **Costs:** Hardware becomes cheaper and more robust to democratize transformative research.
- **Safety:** Humanoid robots can be more predictable and safer for human-robot interaction.
- **Data:** Their similar physique unlocks Internet-scale, human-centered data sources.
- …



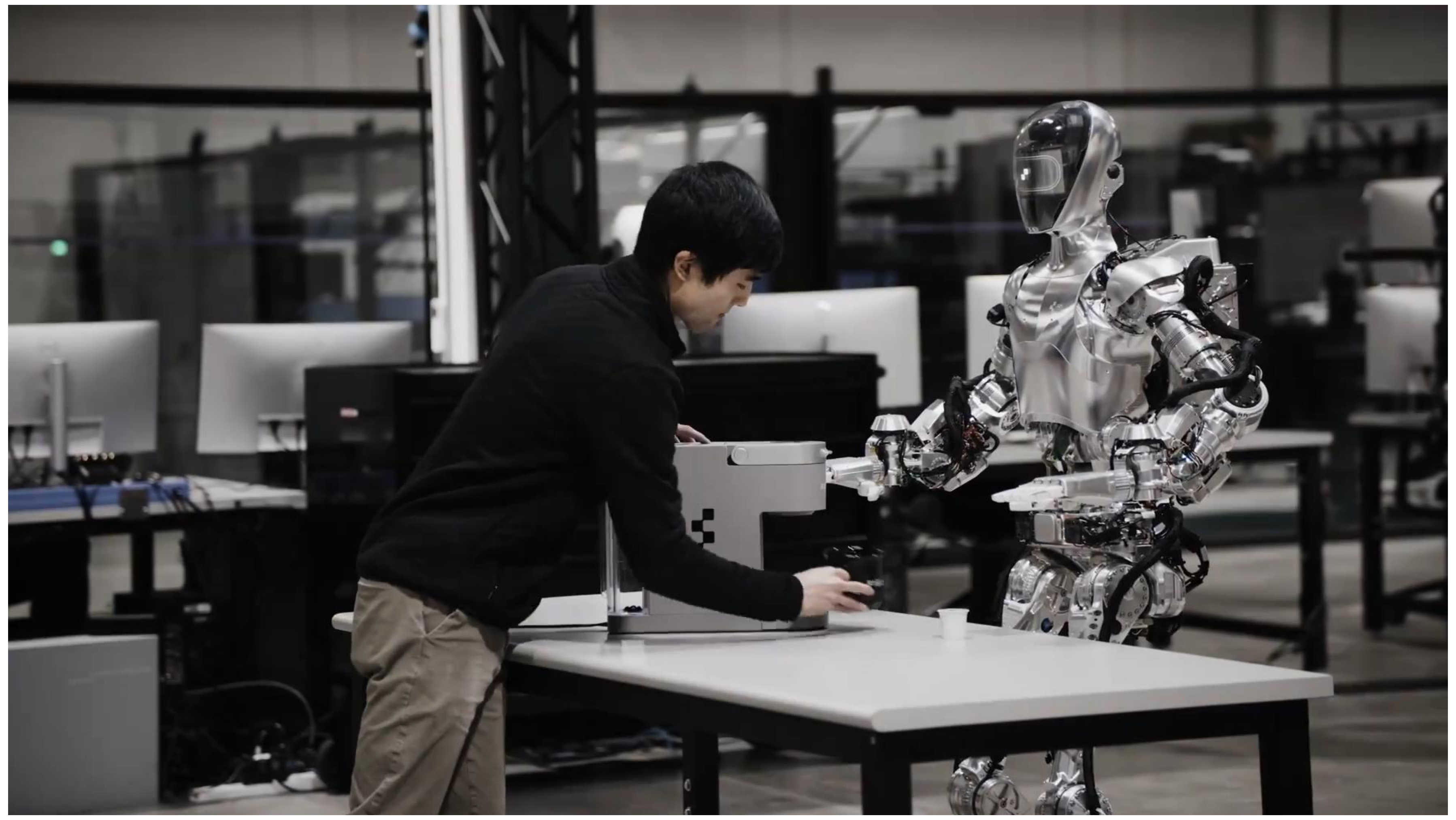
[Credit: Chad Jenkins]

Why Humanoids?

- **Versatility:** General-purpose robot autonomy needs a versatile body.
- **Costs:** Hardware becomes cheaper and more robust to democratize transformative research.
- **Safety:** Humanoid robots can be more predictable and safer for human-robot interaction.
- **Data:** Their similar physique unlocks Internet-scale, human-centered data sources.
- …



[VIOLA, Zhu et al. CoRL 2022]

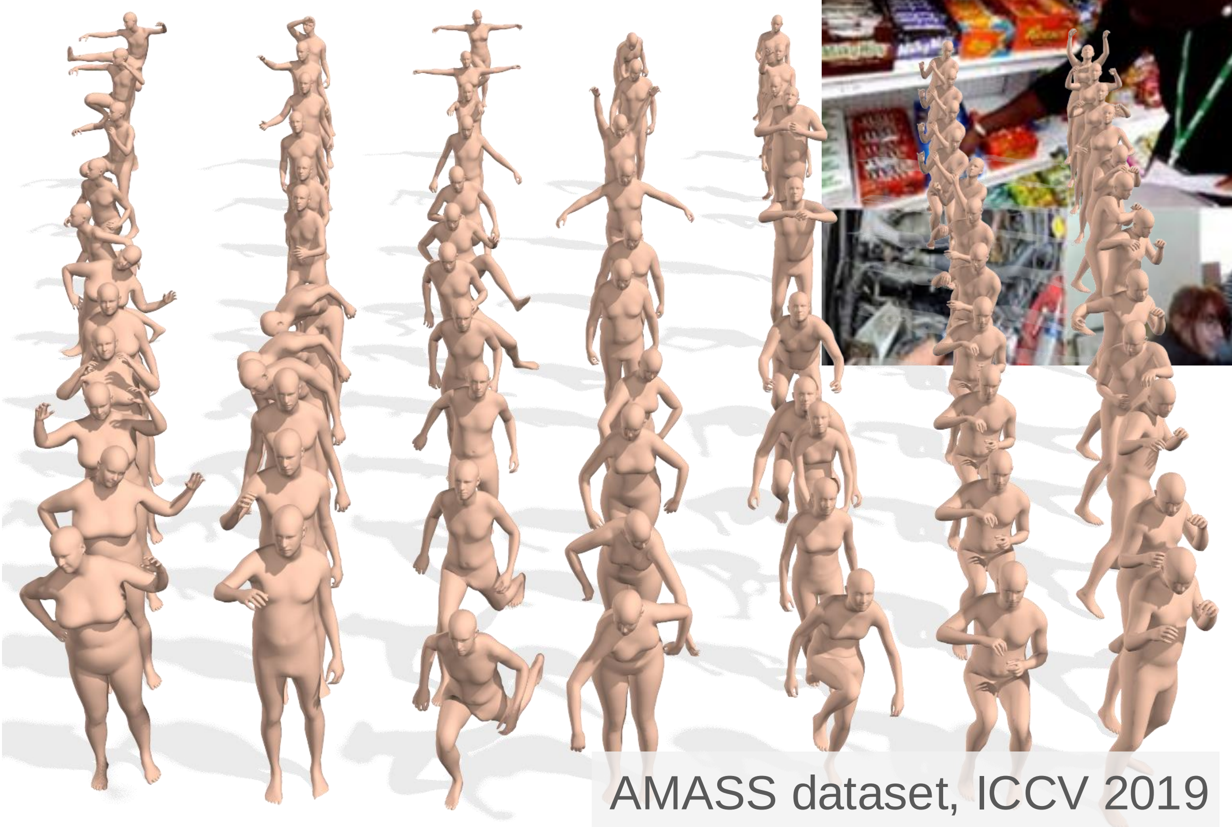


[Credit: Figure AI 2024]

Why Humanoids?

- **Versatility:** General-purpose robot autonomy needs a versatile body.
- **Costs:** Hardware becomes cheaper and more robust to democratize transformative research.
- **Safety:** Humanoid robots can be more predictable and safer for human-robot interaction.
- **Data:** Their similar physique unlocks Internet-scale, human-centered data sources.
- …

Ego4D dataset, CVPR 2022

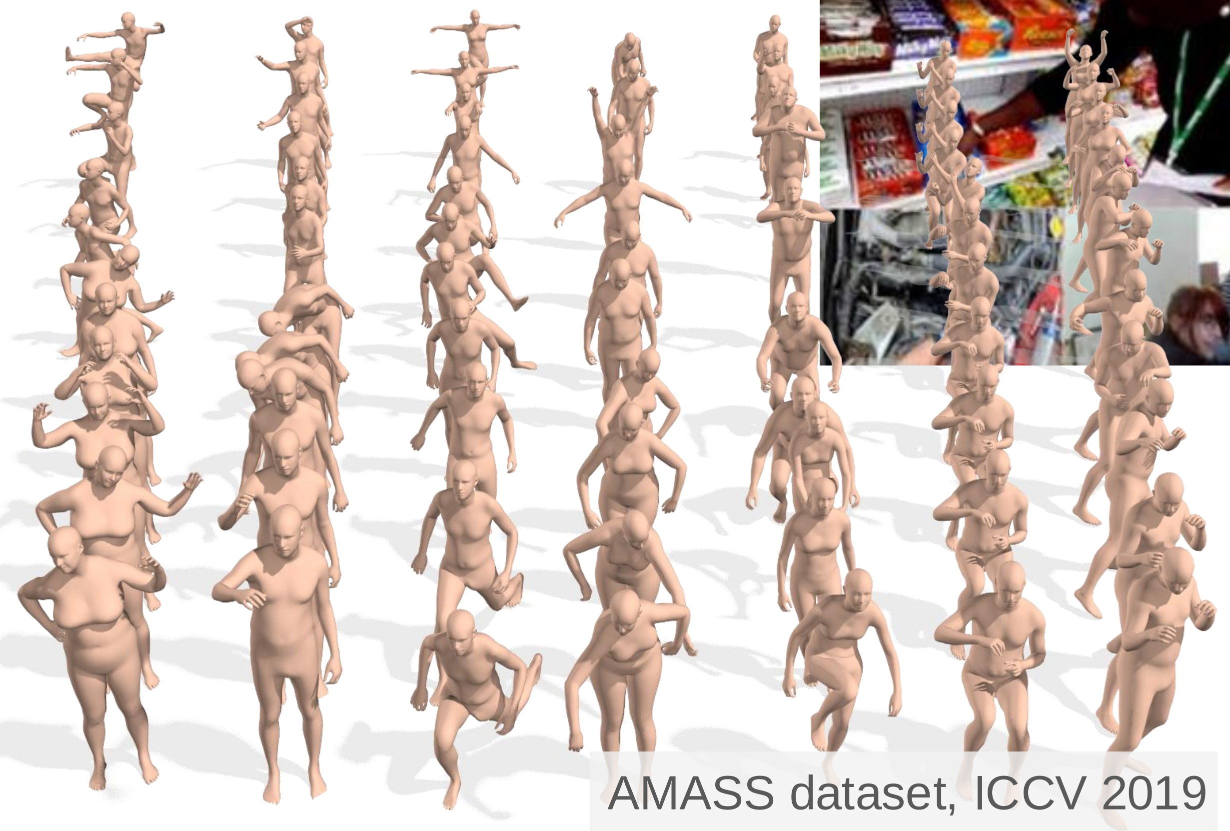


AMASS dataset, ICCV 2019

Why Humanoids?

- **Versatility:** General-purpose robot autonomy needs a versatile body.
- **Costs:** Hardware becomes cheaper and more robust to democratize transformative research.
- **Safety:** Humanoid robots can be more predictable and safer for human-robot interaction.
- **Data:** Their similar physique unlocks Internet-scale, human-centered data sources.
- …

Ego4D dataset, CVPR 2022



AMASS dataset, ICCV 2019

Note: humanoid robotics is still incredibly hard (!) — huge challenges in mechanical designs, dynamics & control, sensor technologies, compute and power, AI algorithm designs…

Learning from Human Videos

single video
demonstration

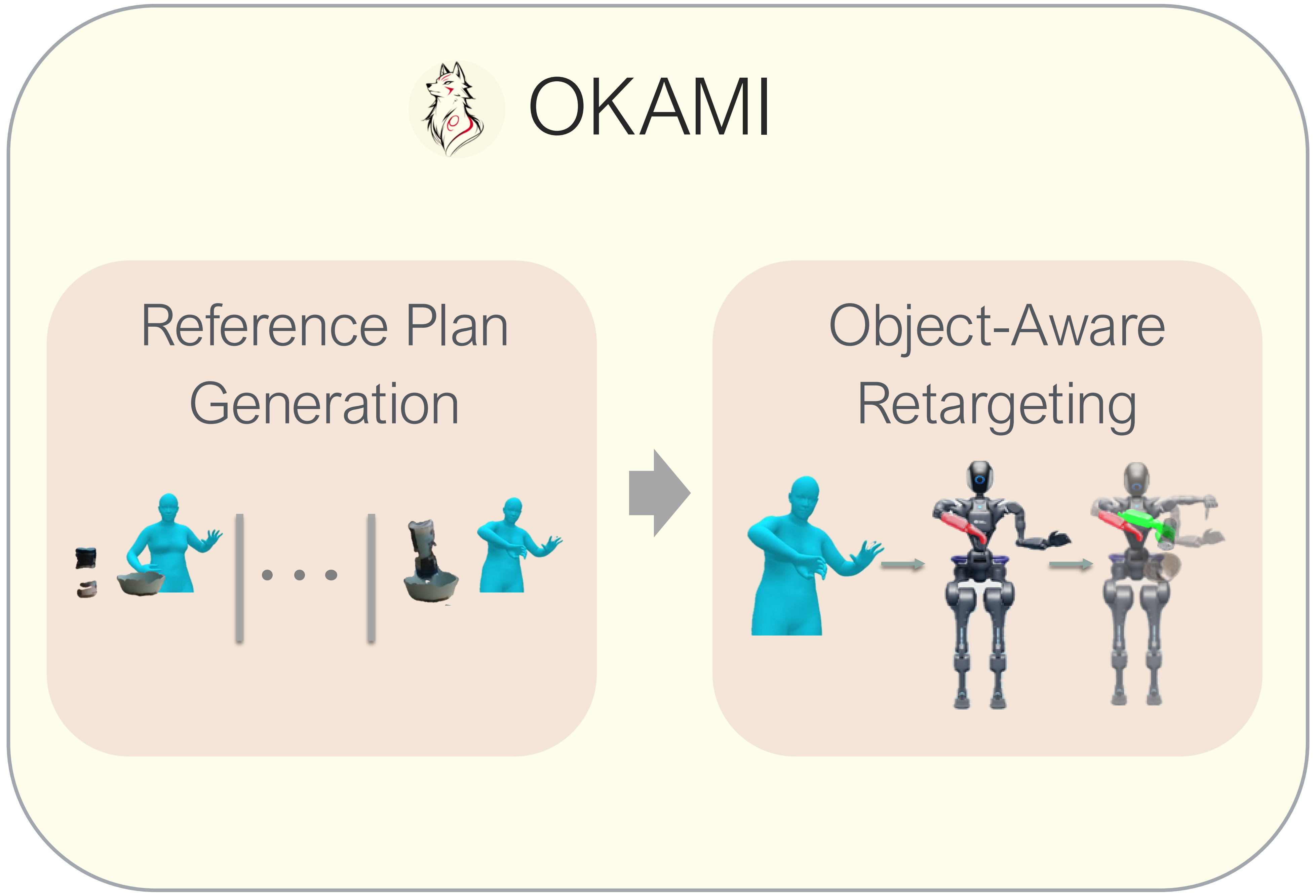
OKAMI

trajectory rollouts
in diverse scenes

"OKAMI: Teaching Humanoid Robots Manipulation Skills through Single Video Imitation." Li et al. CoRL 2024

Learning from Human Videos

single video
demonstration



trajectory rollouts
in diverse scenes

"OKAMI: Teaching Humanoid Robots Manipulation Skills through Single Video Imitation." Li et al. CoRL 2024

Learning from Human Videos

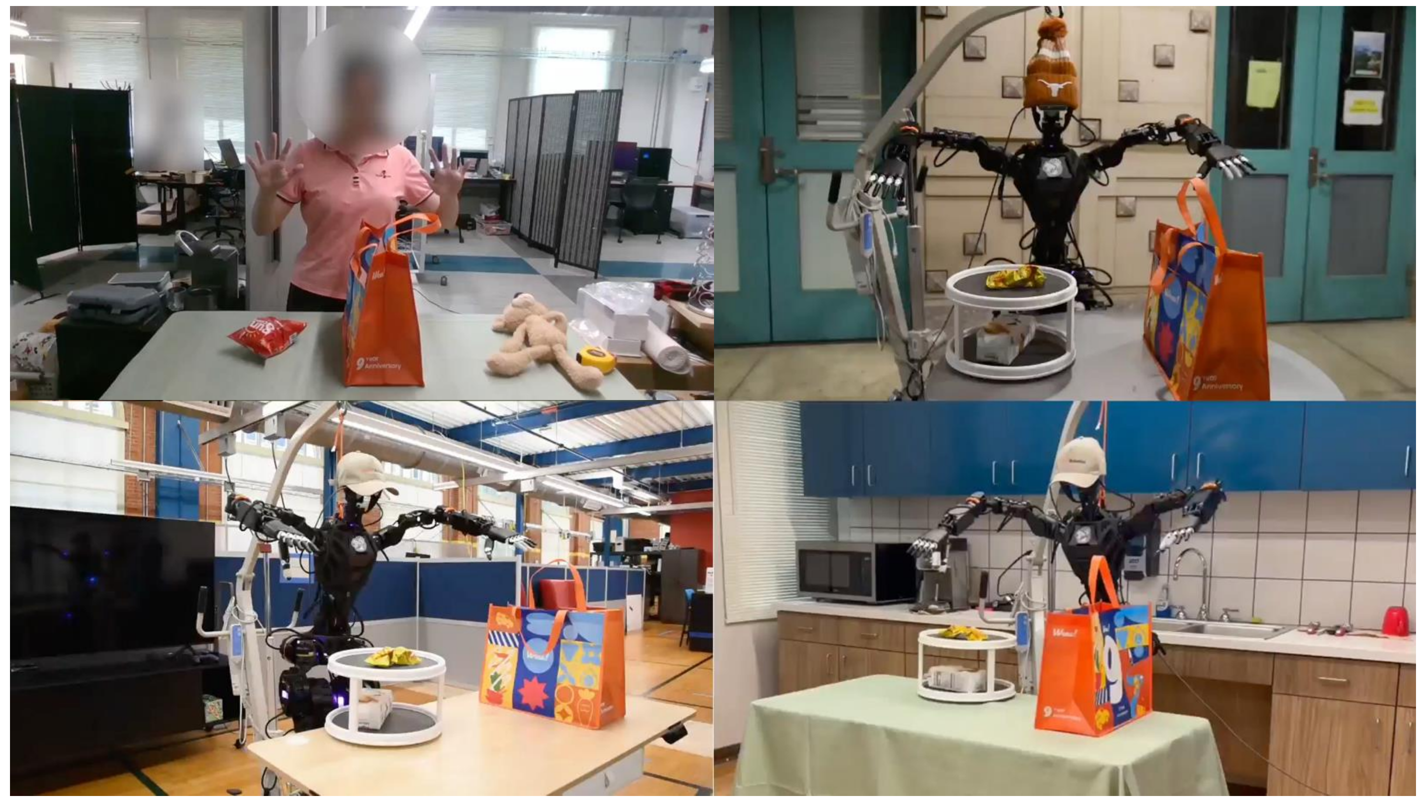
single video demonstration

trajectory rollouts in diverse scenes

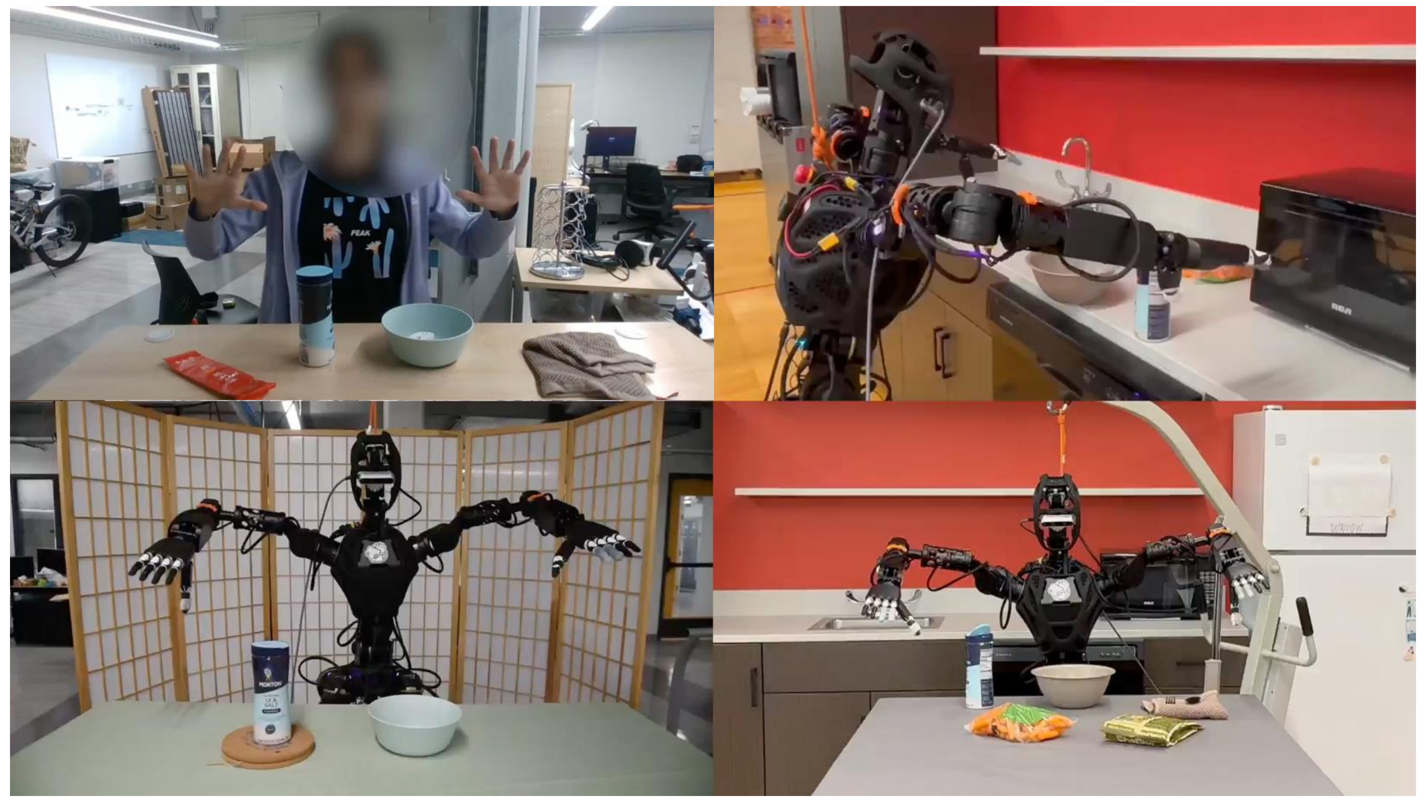
"OKAMI: Teaching Humanoid Robots Manipulation Skills through Single Video Imitation." Li et al. CoRL 2024

Learning from Human Videos

bagging (58.3%)

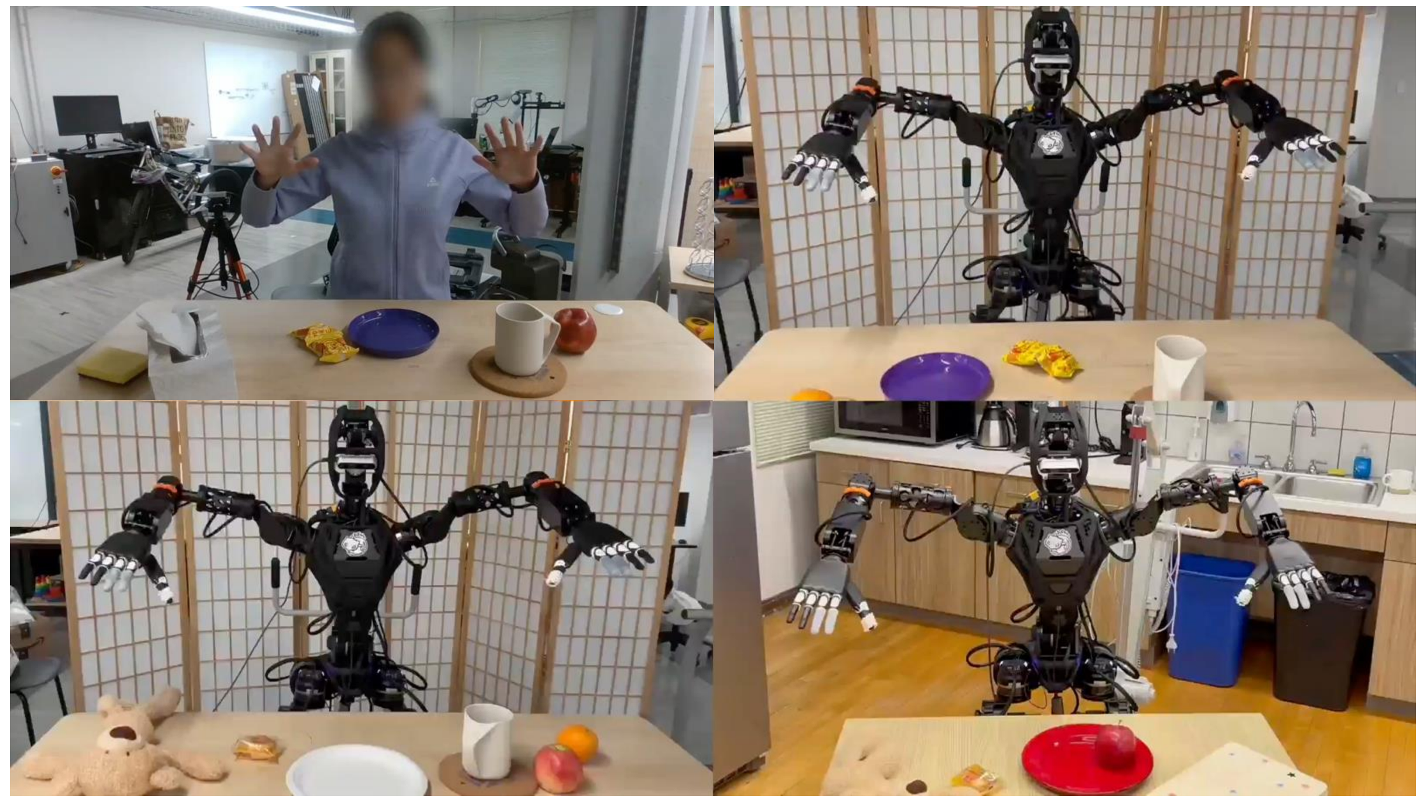


sprinkling salt (58.3%)



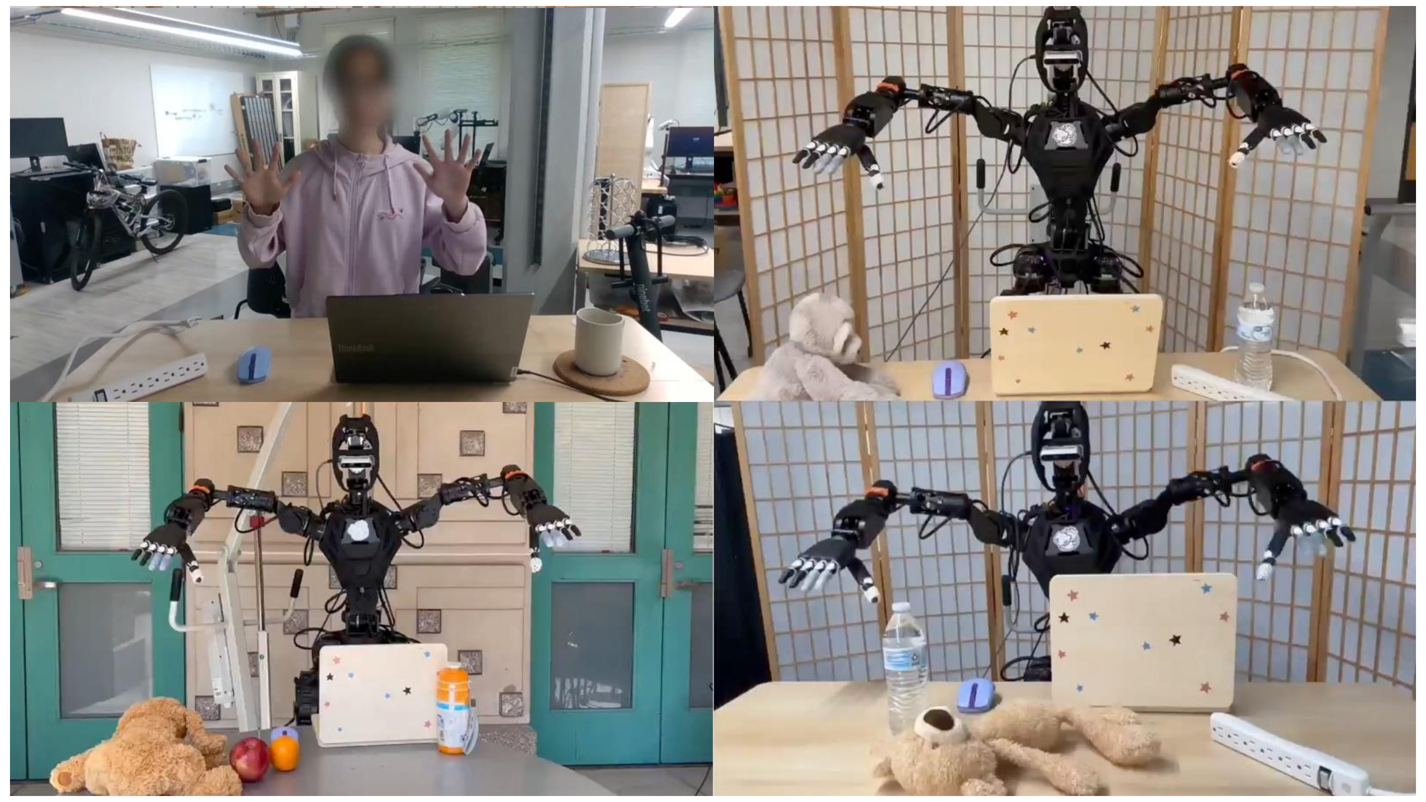
putting toy in basket (66.7%)

placing snacks on plate (75.0%)



closing the drawer (75.0%)

closing the laptop (83.3%)



"OKAMI: Teaching Humanoid Robots Manipulation Skills through Single Video Imitation." Li et al. CoRL 2024

HOVER: One Versatile Policy to Control All Modes

“HOVER: Versatile Neural Whole-Body Controller for Humanoid Robots.” He et al. 2024

Raise left hand and form a longhorn gesture.

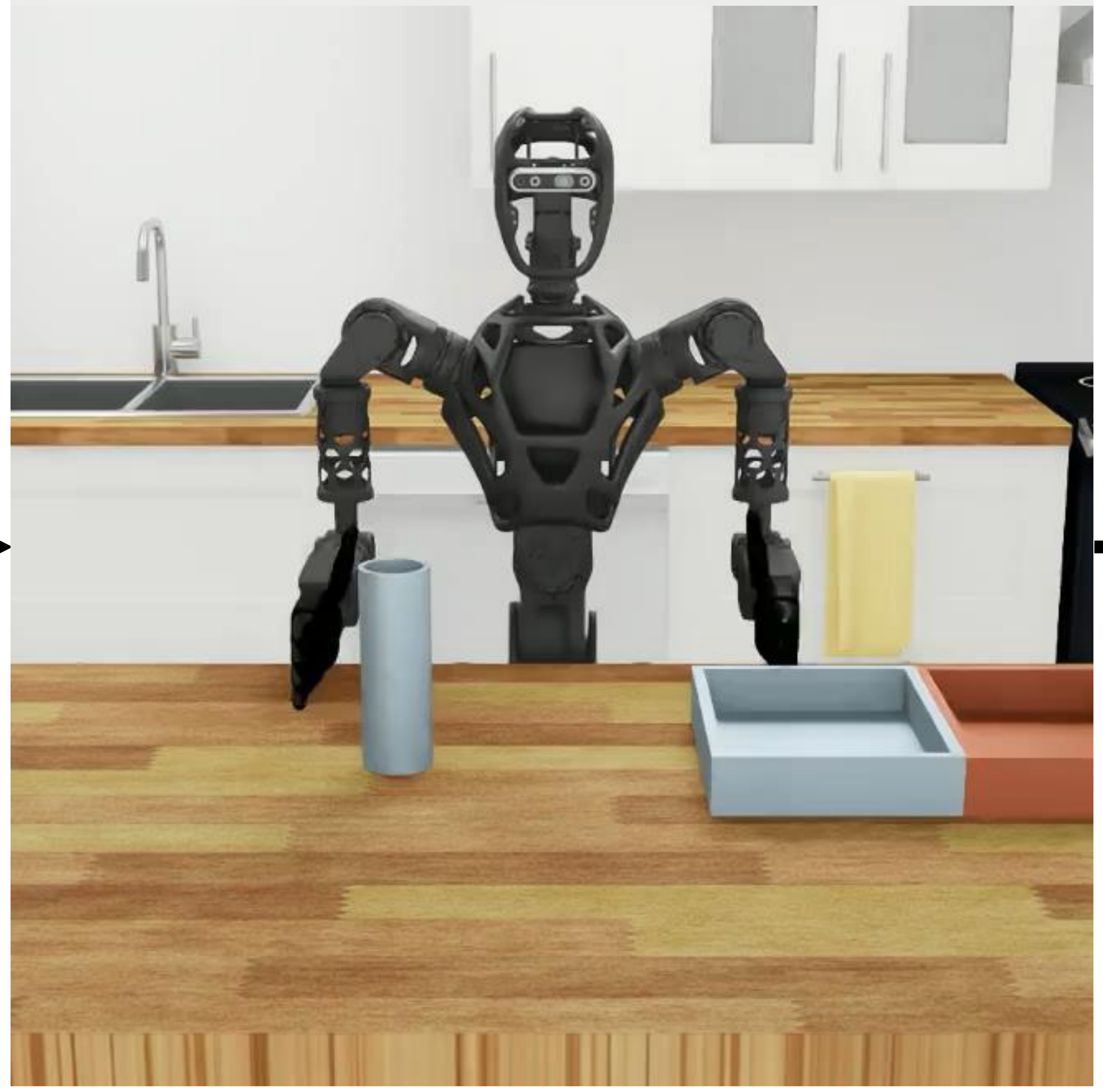
Do a dabbing pose.

Signal a goodbye by touching the head with the right hand, then extending it forward.

DexMimicGen: Automated Data Generation System

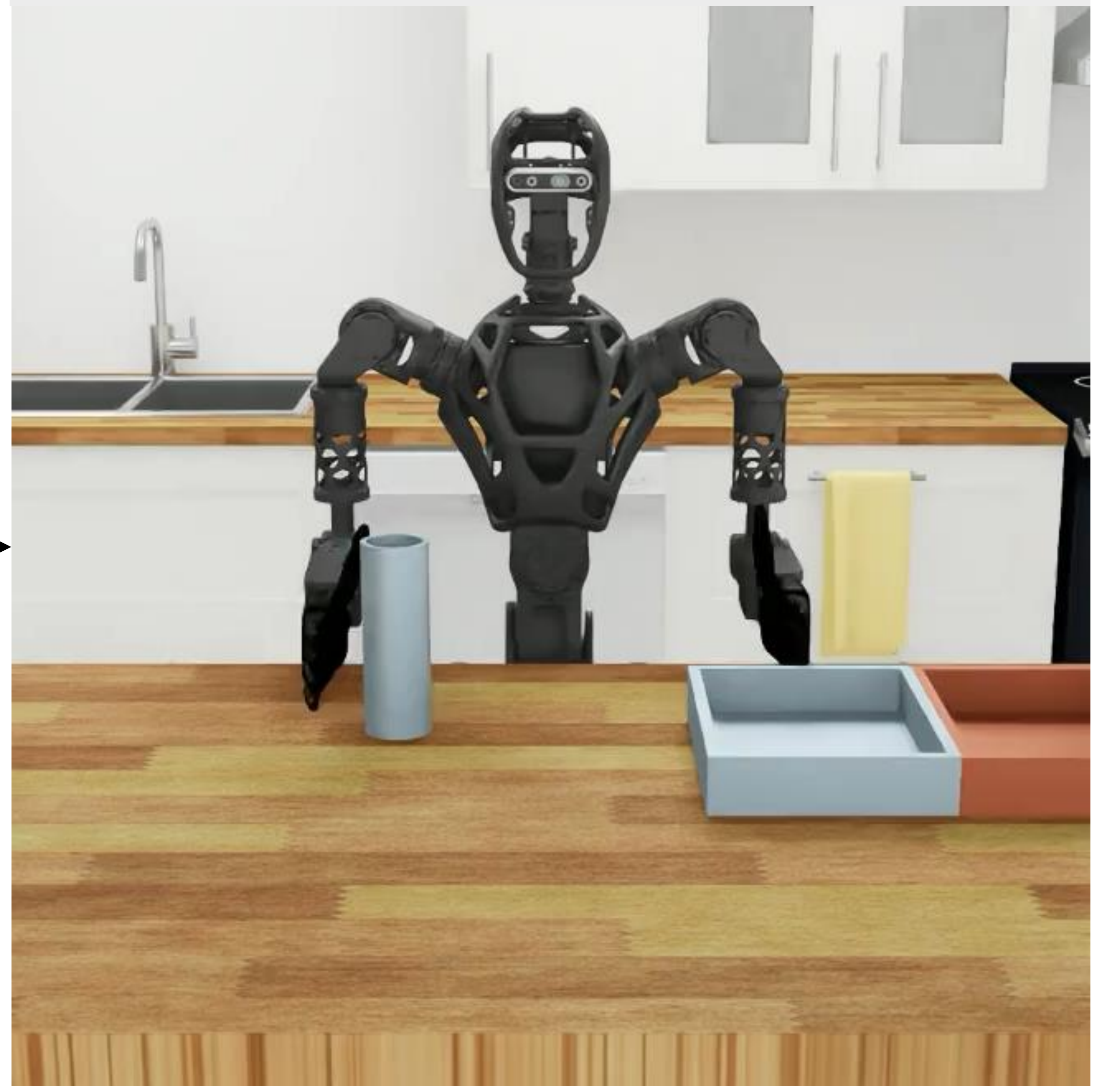
Human teleoperation

5 source demos



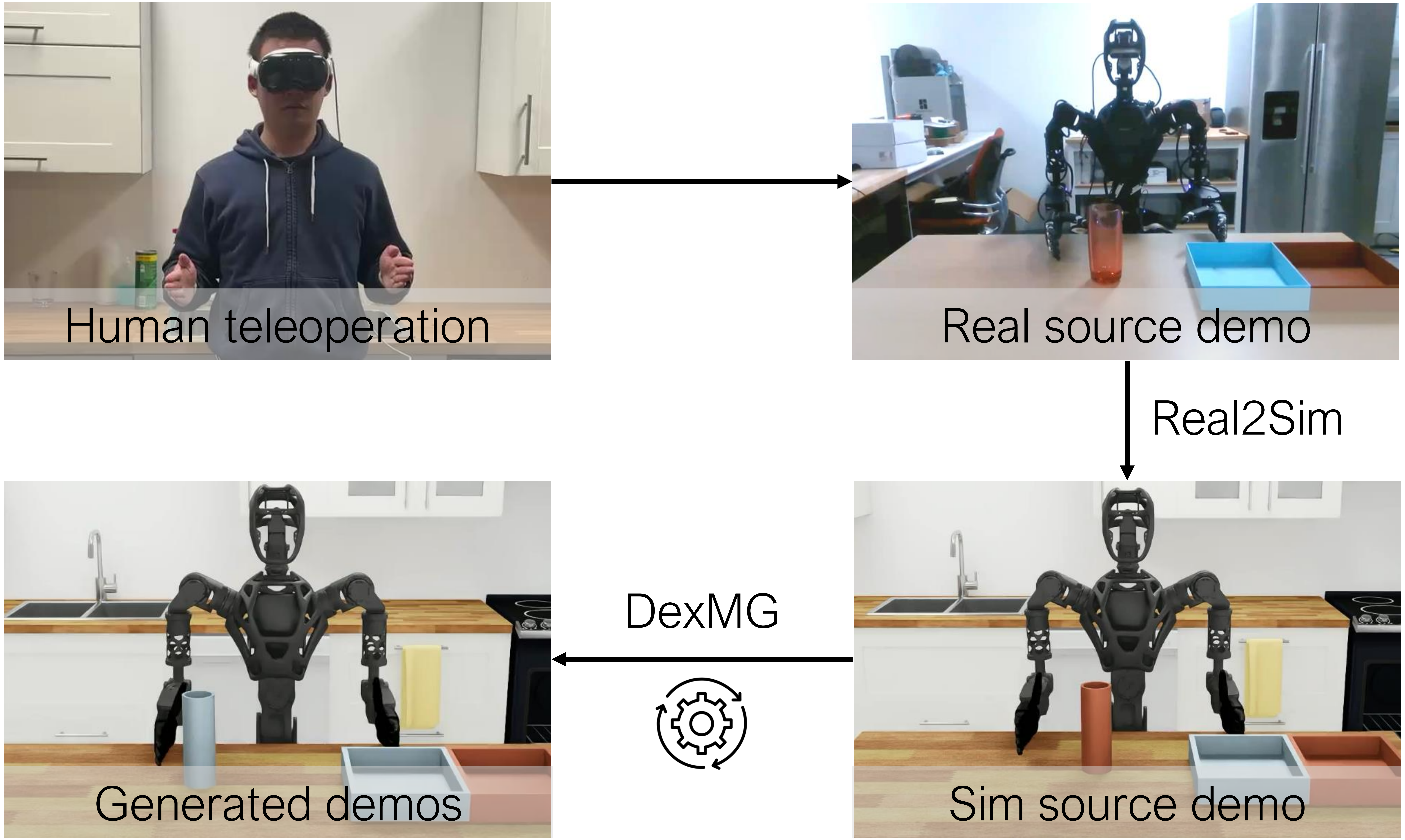
DexMimicGen

1000 generated demos



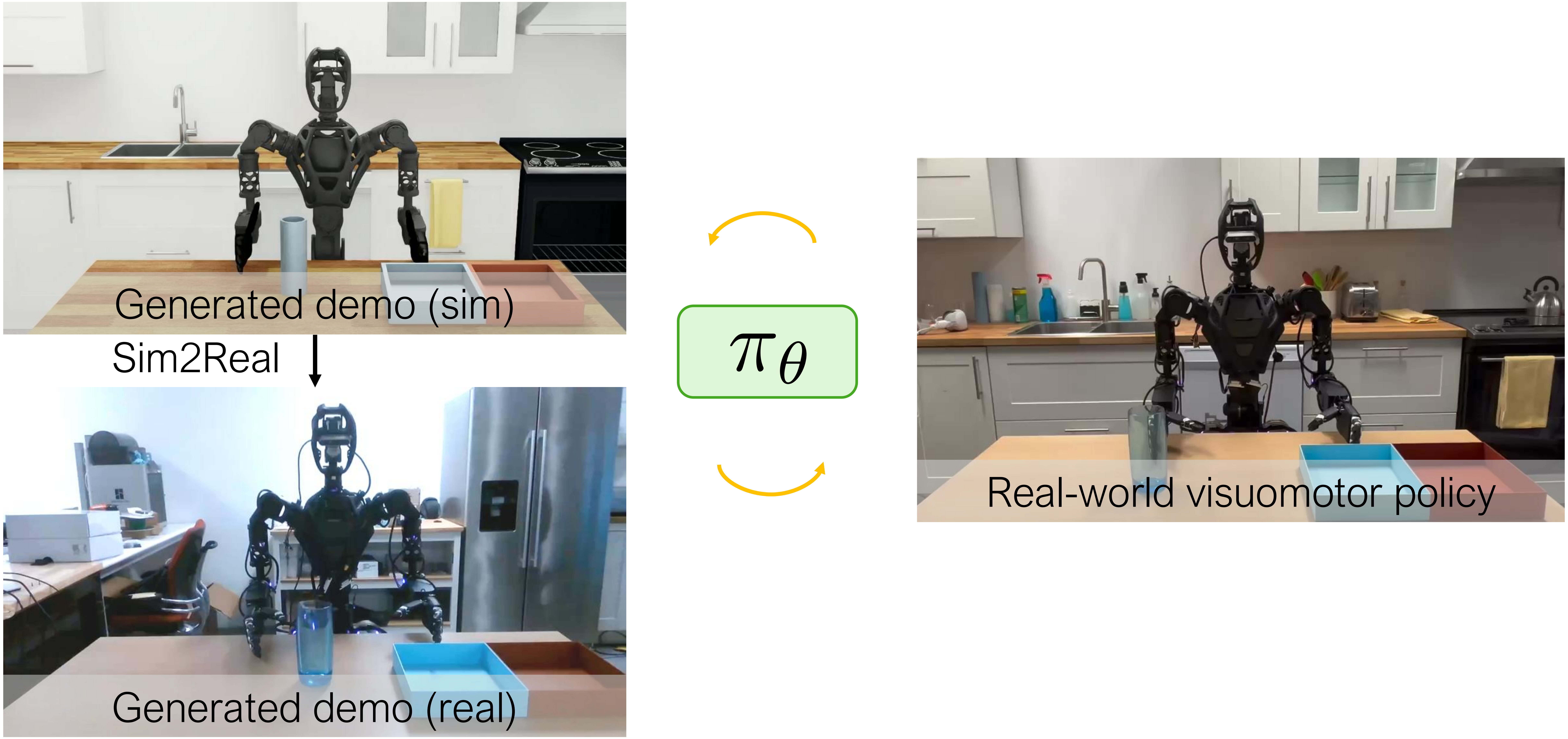
“DexMimicGen: Automated Data Generation for Bimanual Dexterous Manipulation via Imitation Learning.” Jiang*, Xie*, Lin*, et al. 2024

DexMimicGen can be used to train real-world visuomotor policy.



Transfer real demo to sim using digital twin to ensure the sim demos are valid in real

DexMimicGen can be used to train real-world visuomotor policy.



Transfer only **successful** generated demos from sim to real to train a visuomotor policy

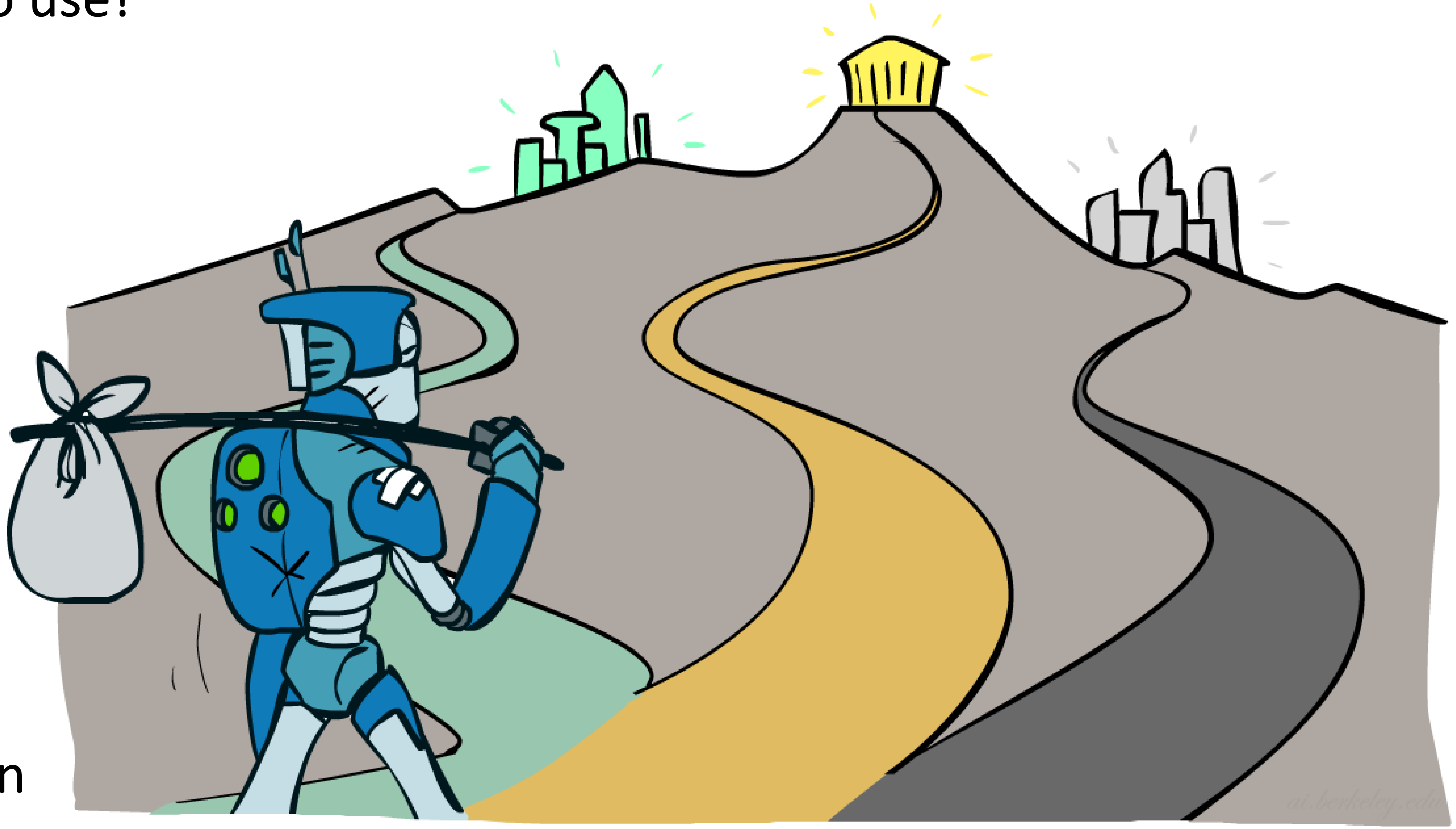
DexMimicGen can be used to train real-world visuomotor policy.

Real-world visuomotor policy rollouts (10X)

Where to Go Next?

Where to Go Next?

- Congratulations, you've seen the basics of modern AI
  - … and done some amazing work putting it to use!
- How to continue:
  - CS 395T Visual Recognition
  - CS 391R Robot Learning
  - ECE 382V Human Robot Interaction
  - CS 388 Natural Language Processing
  - CS 391L Machine Learning
  - CS 393R Autonomous Robots
  - CS 342 Neural Networks
  - EE 381V Advanced Topics in Computer Vision
  - CS 394R Reinforcement Learning: Theory and Practice
  - … and more; ask if you're interested



Final Remarks

- We have come a long way! Thank you!
- We are very proud that you have made it to the end of this demanding course!
- We are impressed by your ingenuity and critical thinking in the in-class discussions, Ed posts, projects, and assignments!
- Thanks to Huihan and Shuijing for handling the course logistics.
- If this course helps you kickstart your future endeavors in AI, please email us and let us know!

Thank you

Saturday 12/14   1 – 3pm   ECJ 1.202

1 page (front and back) of notes

Closed book

That’s it!

I had a great time teaching this course and
I hope you all enjoyed it as well

Have a great winter break!