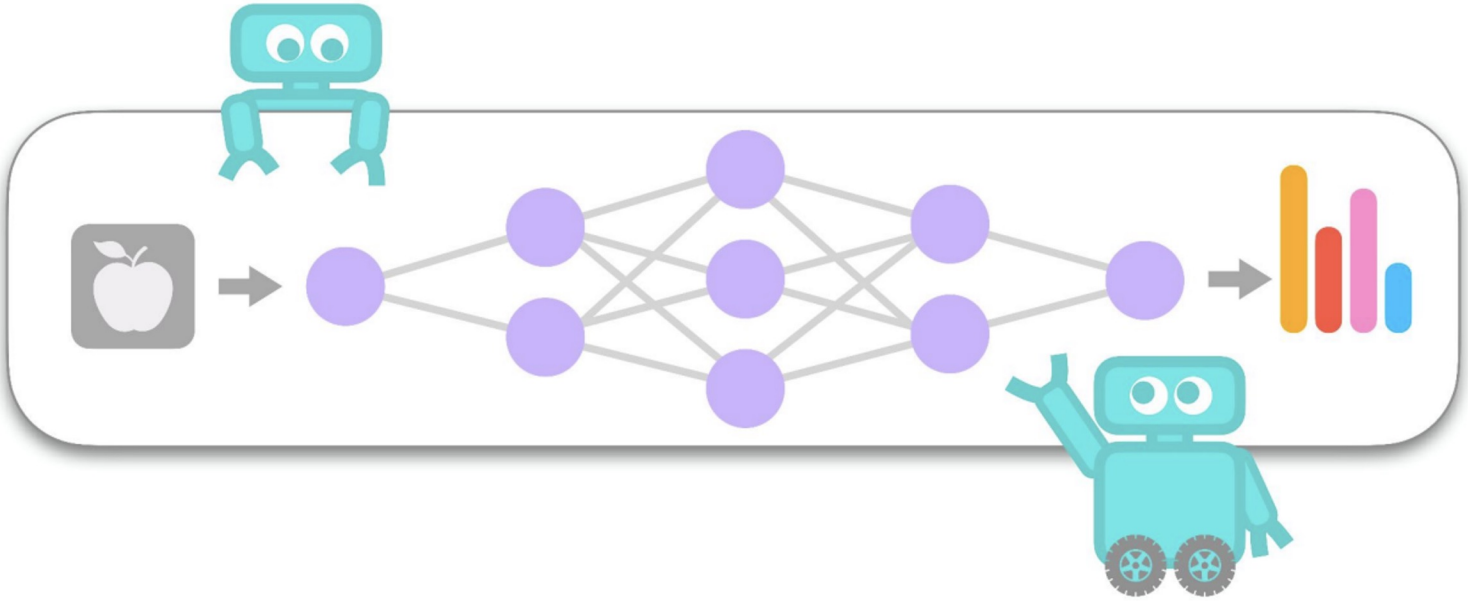
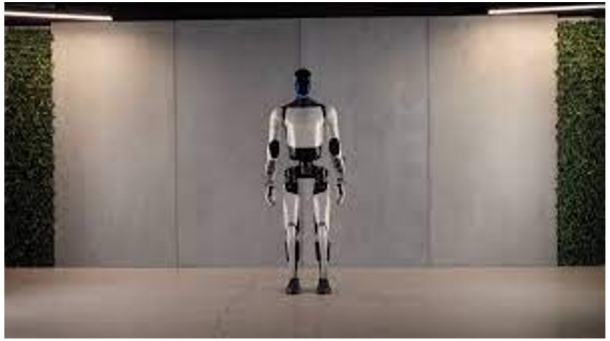


# Foundation Models for Robotics

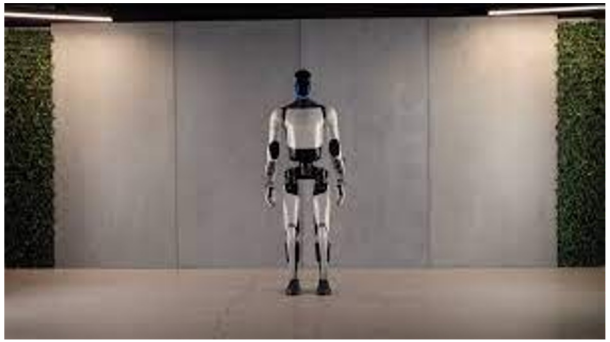
# What is “Robot” in your mind?

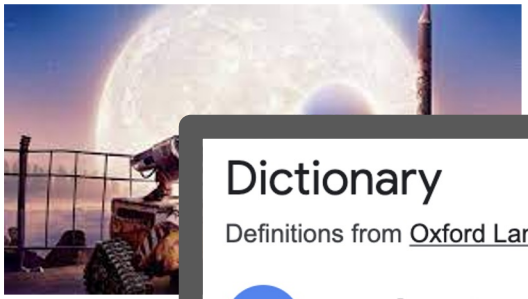
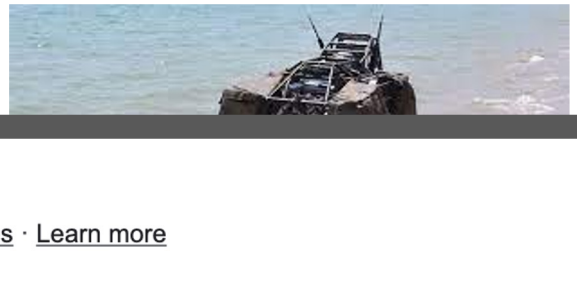
# What is “Robot” in your mind?

# What is “Robot” in your mind?




# What is “Robot” in your mind?




# Dictionary

Definitions from Oxford Languages · Learn more

## ro·bot

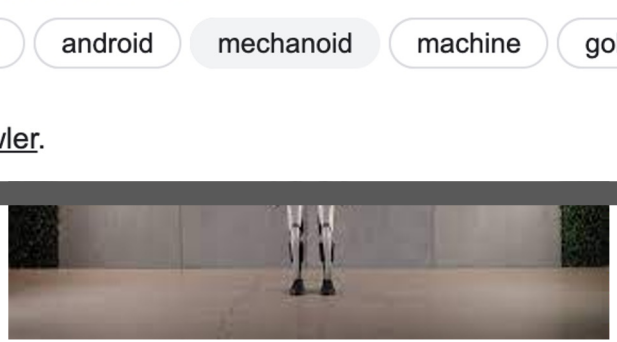
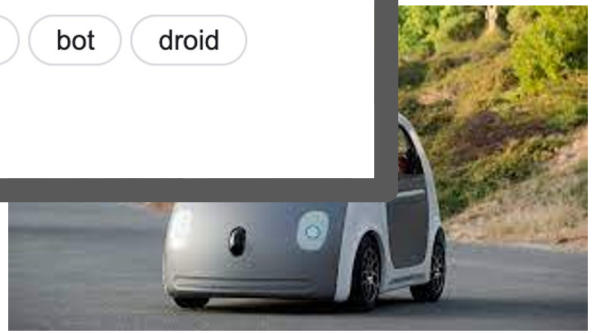
/ˈrōˌbät,ˈrōbət/

*noun*

1. (especially in science fiction) a machine resembling a human being and able to replicate certain human movements and functions automatically.
   "the robot closed the door behind us"

   Similar: automaton | android | mechanoid | machine | golem | bot | droid

2. **COMPUTING**
   another term for crawler.

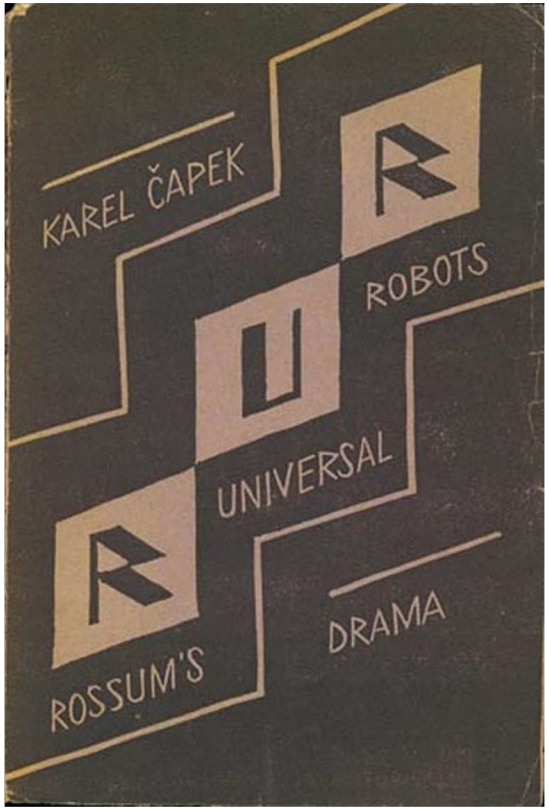
# History — The Term Robot Is First Used (1921)

**Book: “R.U.R.: Rossum’s Universal Robots”**

**Author: Karel Čapek**

It tells the tale of a factory in which thousands of synthetic humanoids have been created. They work so cheaply and tirelessly that they shrunk production costs of weaving material by 80 percent.

Čapek named the devices “robots,” after the Czech word ***robota***, referring to the forced labor of serfs. The play not only gave robots their modern name, but heightened the existential fear that robots will someday replace people, as Čapek’s robots ultimately rise up and kill humanity.



# History — First Machine Navigates on Its Own (1949)

**Inventor: William Grey Walter**

In 1949, an American-born British neurophysiologist and inventor named William Grey Walter introduced a pair of battery-powered, tortoise-shaped robots.

It could maneuver around objects in a room, guide themselves toward a source of light and find their way back to a charging station

It uses the same components that remain crucial to robotics today: **sensor technology**, **a responsive feedback loop**, and **logical reasoning**.

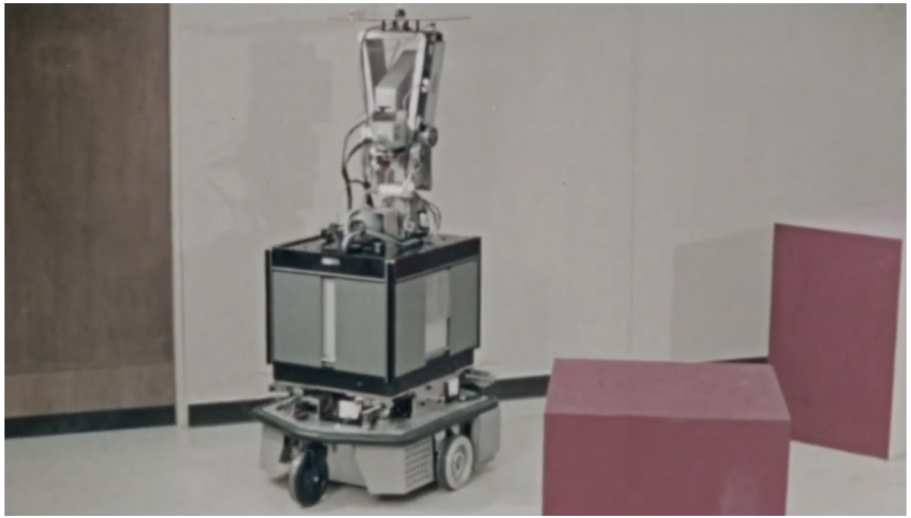
# History — First Robot to Use Artificial Intelligence (1972)

**Researcher: Stanford Research Institute**

**Robot Name: Shakey**

If you gave Shakey a goal — such as navigating its way across a room or pushing a box along the floor — it could accomplish it by observing the world around it, creating a plan, and executing.

With sensors that included a TV camera, a range finder and touch-sensitive metal whiskers, Shakey would gather data that enabled it to build a model of its environment and then use a **"planning" program** to generate its next moves.



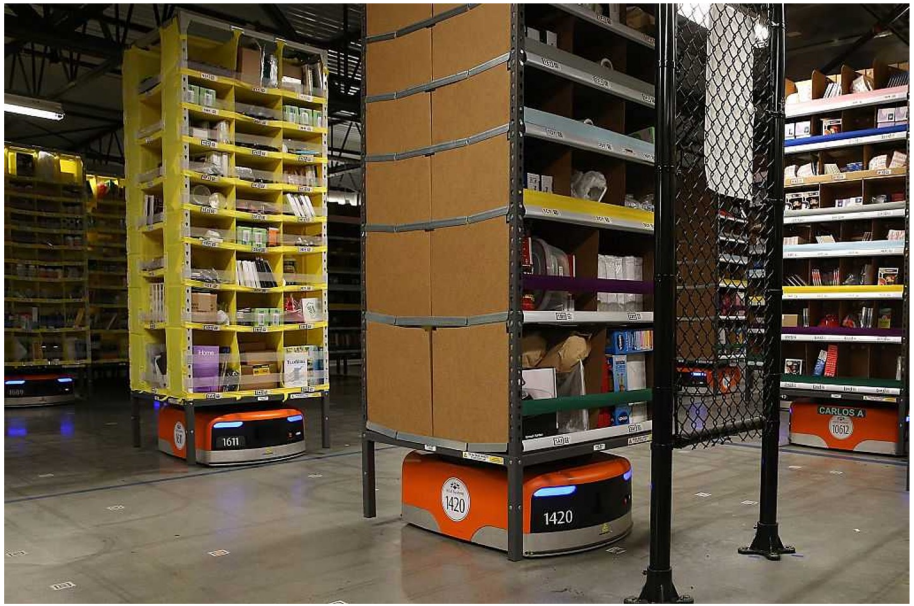
# History — Kiva Robots Re-Engineer the Warehouse (2003)

**Name: Kiva**

**Mick Mountz and his cofounders** created the Kiva robot: a squarish, close-to-the-ground orange bot (not too different from an extra-large Roomba) that can glide around warehouses, moving racks of goods.

Kiva used some inexpensive off-the-rack components, which could make the robots less precise in how it moved about, but Kiva's engineers **compensated with software** that course-corrected on the fly.

Kiva's system revolutionized the efficiency of warehouse and shipping. Amazon bought the company for $775 million in 2012.



# History — Self-Driving Cars Pass First Big Test (2005)

**Team: Stanford Racing Team**

**Robot Name: Stanley**

"Stanley" won the second DARPA Grand Challenge — to complete a rough and often harrowing **131.2-mile** course in the Mojave Desert within **10 hours**.

The race had been established the previous year by the Defense Department's Defense Advanced Research Projects Agency (DARPA) to spur competition and innovation in military autonomous vehicle tech, but **none of the cars** go more than **eight miles**.

What fueled Stanley's victory was a constellation of improvements, including **AI trained on the driving habits** of real-world humans and **five "Lidar" laser sensors**.

# History — Deep Learning (2012)

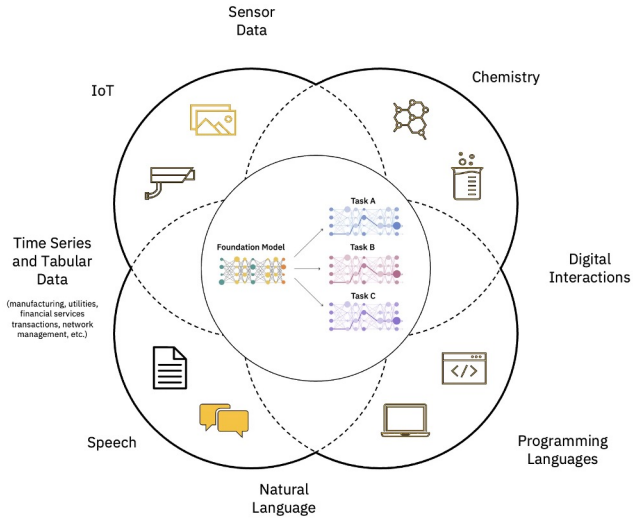
In 2012, the British-born artificial-intelligence expert Geoffrey Hinton and a small team at the University of Toronto produced a stunning advance in AI by creating the most accurate visual-recognition system the world had yet seen.

In the 2012 ImageNet competition, Hinton's team created a system that could identify and sort more than a million images with an error rate of only 15.3 percent, 10 points better than the closest rival.

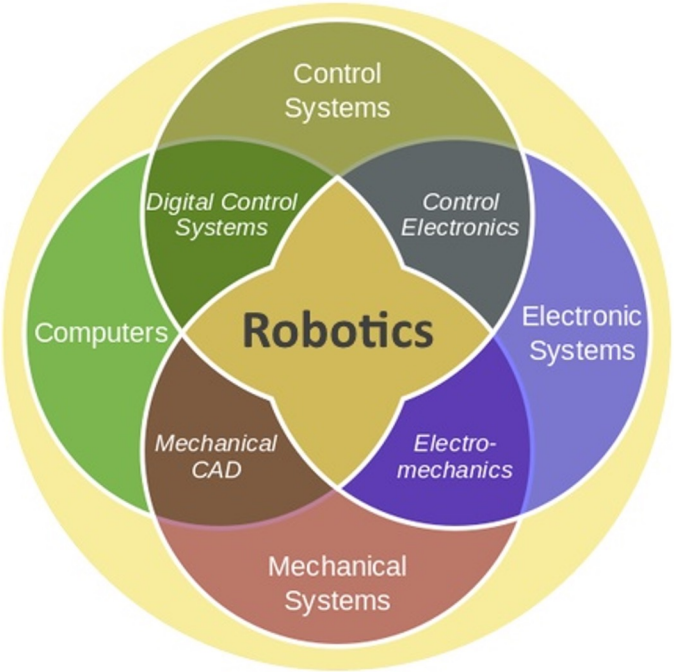
Within months, AI companies were flocking to "deep learning," and firms like Google were releasing open-source tools that let any tiny start-up easily train neural nets.

# History — Foundation Models (Now)

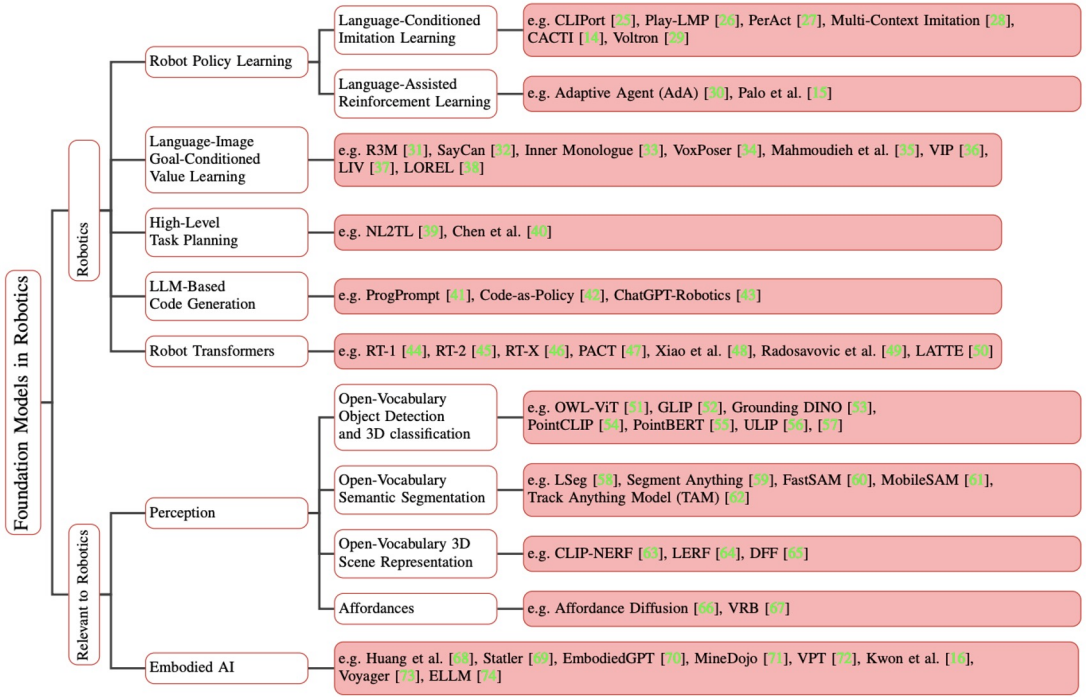
| Language | Classification | LM + Vision | And More! | Chaining |
|---|---|---|---|---|
| ELMo | CLIP | Flamingo | Segment Anything | LMs + CLIP |
| BERT | CoCa | GPT-4V | Whisper | Visual Programming |
| GPT | | Gemini | Dalle | |
| T5 | | | Stable Diffusion | |
| | | | Imagen | |



# Robotics is an interdisciplinary field



# Foundation Models for Robotics

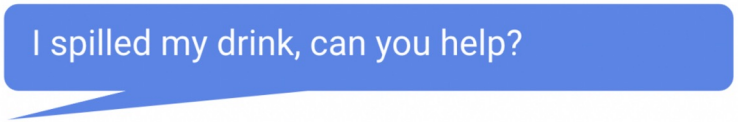


# SayCan [Google, 2022]

# SayCan [Google, 2022]

**Goal:** Solve long horizon tasks from natural language instructions by grounding large language models in real world.

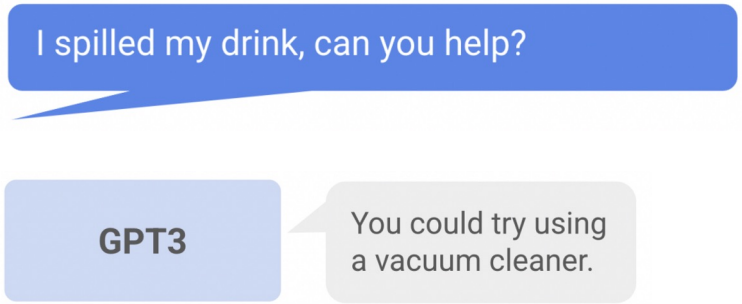
# SayCan [Google, 2022]




**Goal:** Solve long horizon tasks from natural language instructions by grounding large language models in real world.

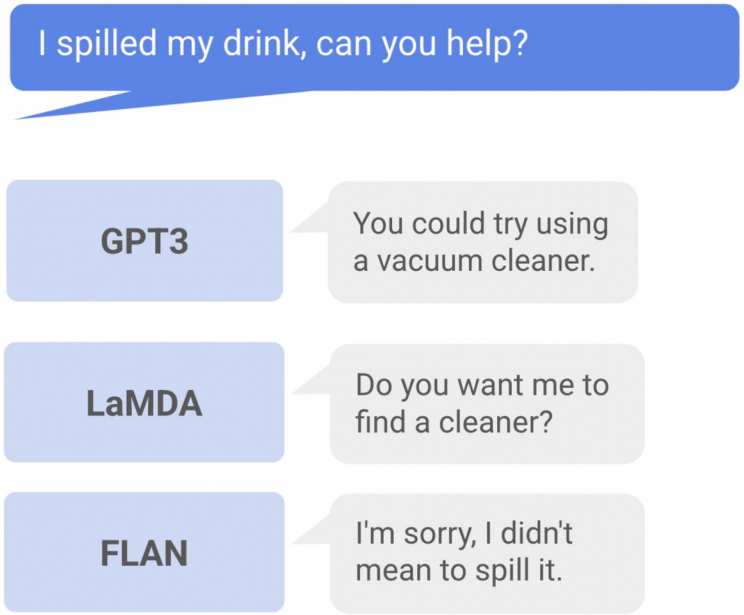
# SayCan [Google, 2022]

**Goal:** Solve long horizon tasks from natural language instructions by grounding large language models in real world.



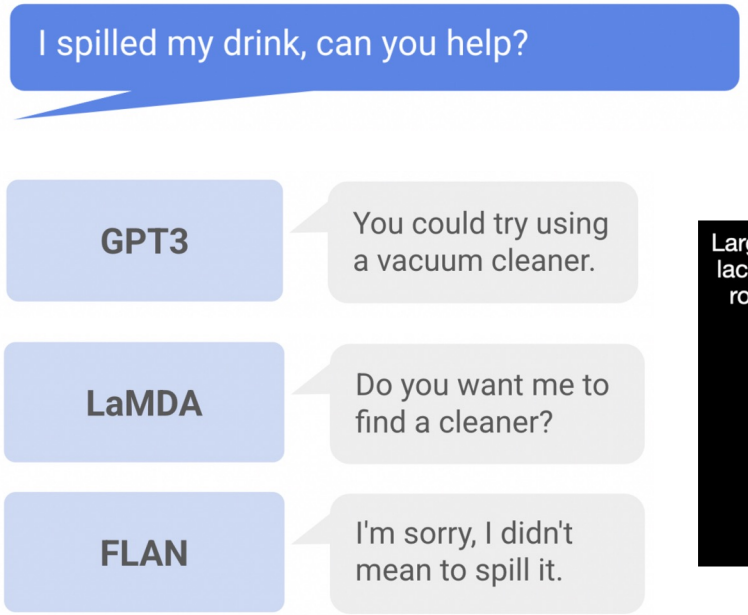
# SayCan [Google, 2022]

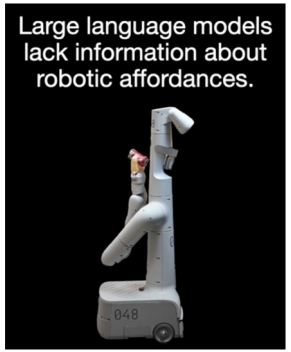
**Goal:** Solve long horizon tasks from natural language instructions by grounding large language models in real world.



# SayCan [Google, 2022]

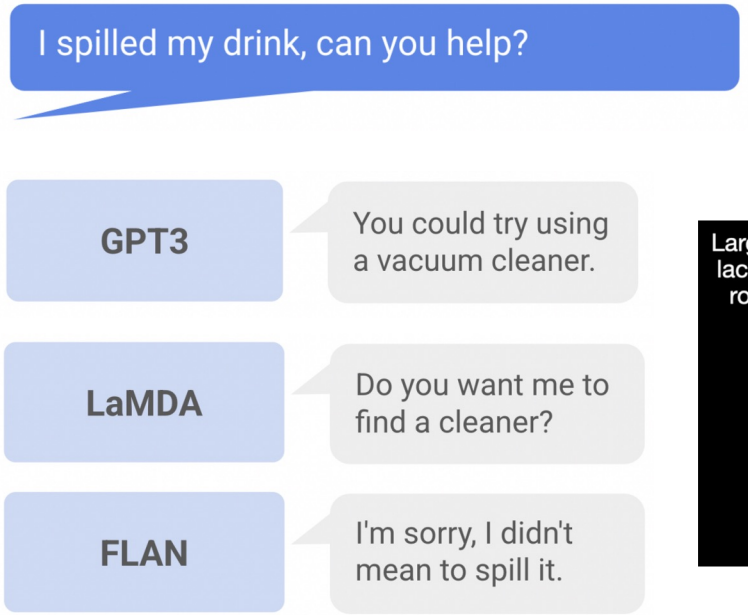
**Goal:** Solve long horizon tasks from natural language instructions by grounding large language models in real world.

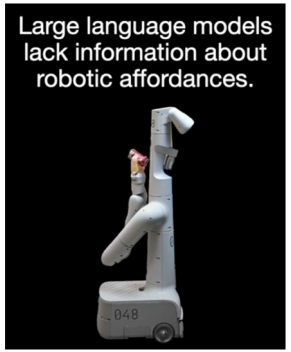


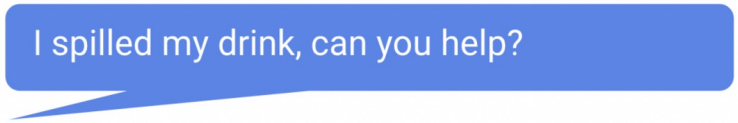
# SayCan [Google, 2022]

**Goal:** Solve long horizon tasks from natural language instructions by grounding large language models in real world.



# SayCan [Google, 2022]

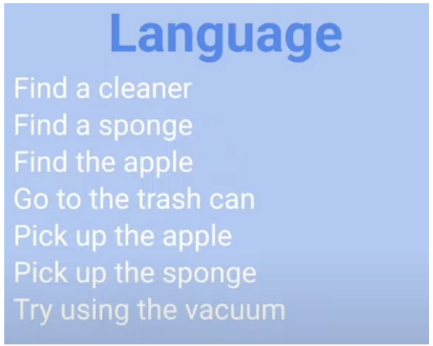



**SayCan grounds language with in robotic affordance**

# SayCan [Google, 2022]

I spilled my drink, can you help?

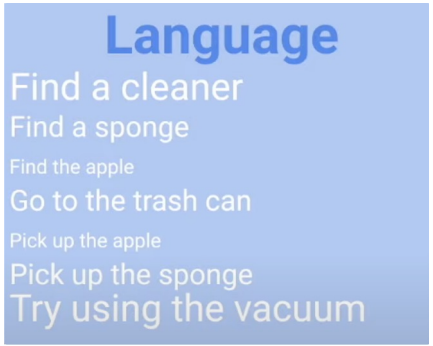
**SayCan grounds language with in robotic affordance**



# SayCan [Google, 2022]

I spilled my drink, can you help?

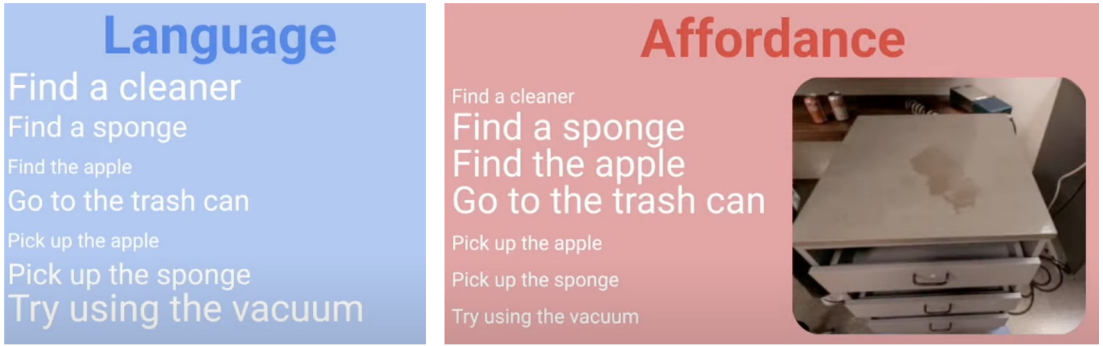
**SayCan grounds language with in robotic affordance**



# SayCan [Google, 2022]

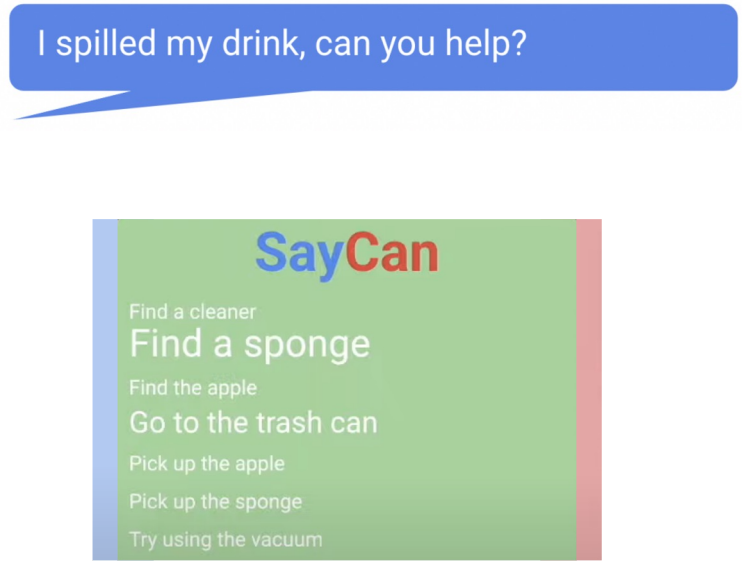
I spilled my drink, can you help?

**SayCan grounds language with in robotic affordance**



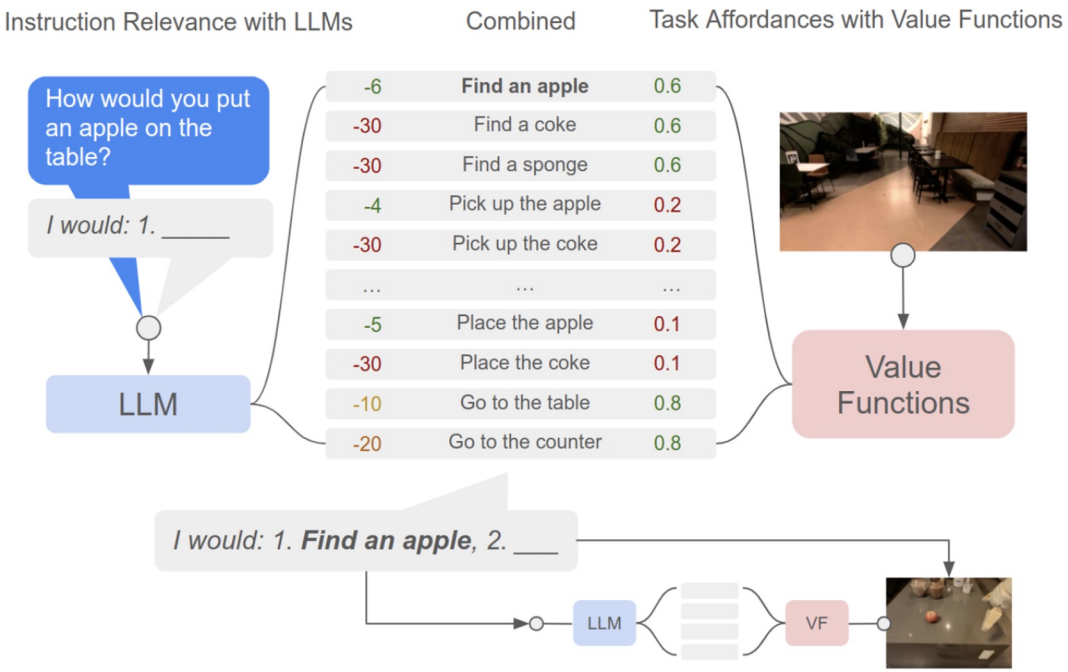
# SayCan [Google, 2022]

**SayCan grounds language with in robotic affordance**



# SayCan [Google, 2022]

**SayCan grounds language with in robotic affordance**



# SayCan [Google, 2022]

**SayCan grounds language with in robotic affordance**