



DeepRob

Fall 2024

Deep Learning for Robot Manipulation
University of Minnesota

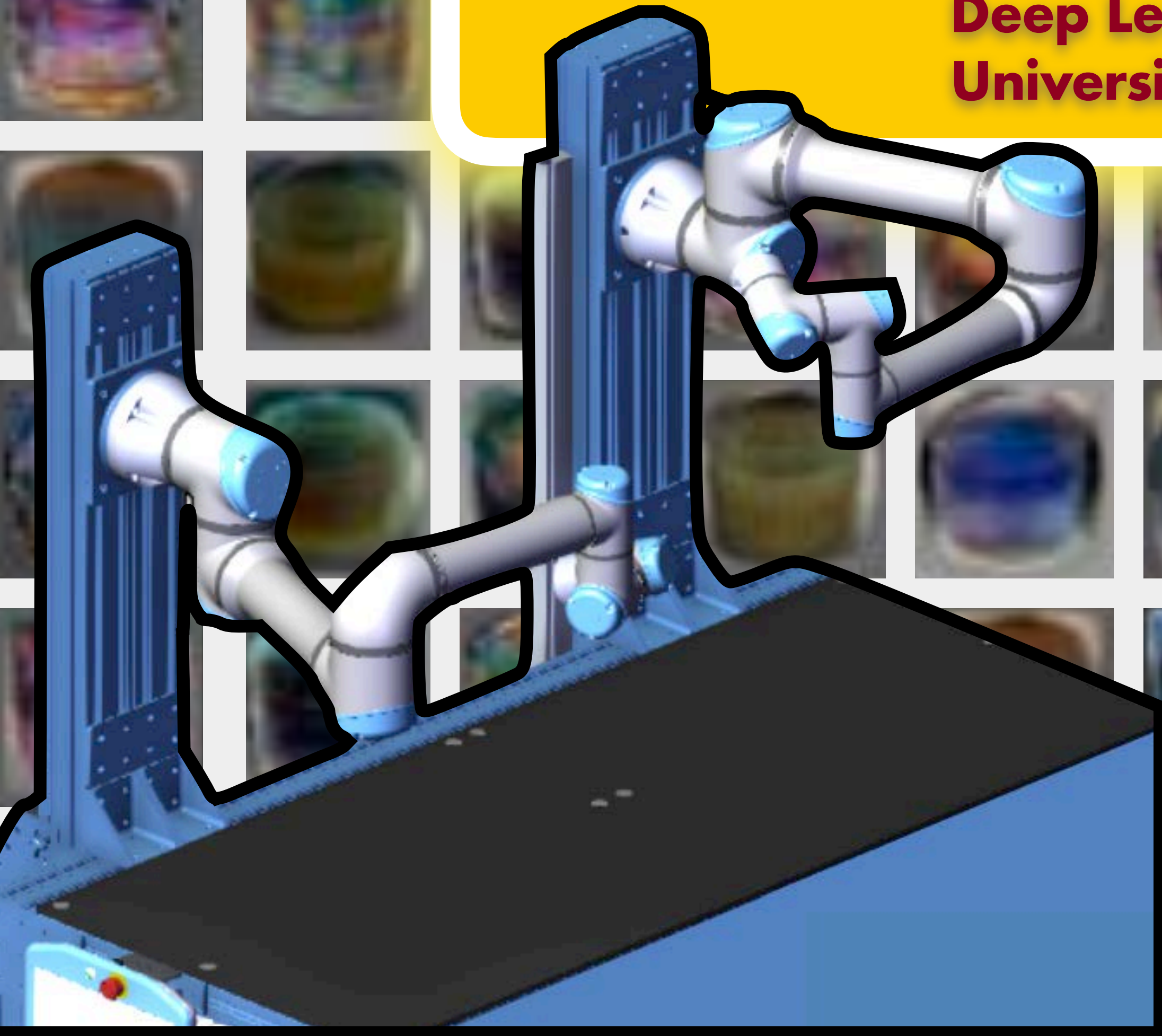


FIGURE 01 + OPENAI SPEECH-TO-SPEECH REASONING

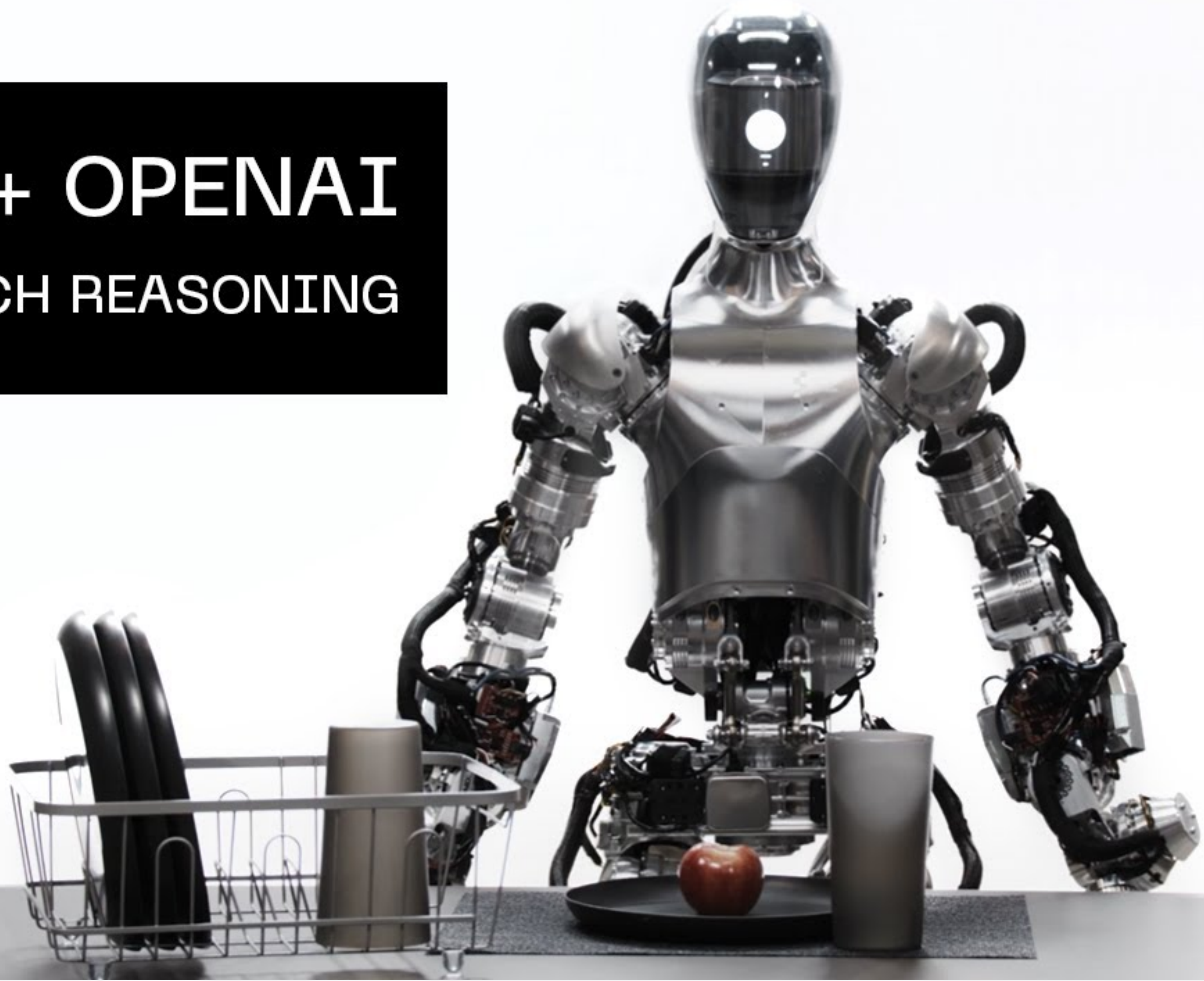


Figure: <https://www.youtube.com/watch?v=Sq1QZB5baNw>



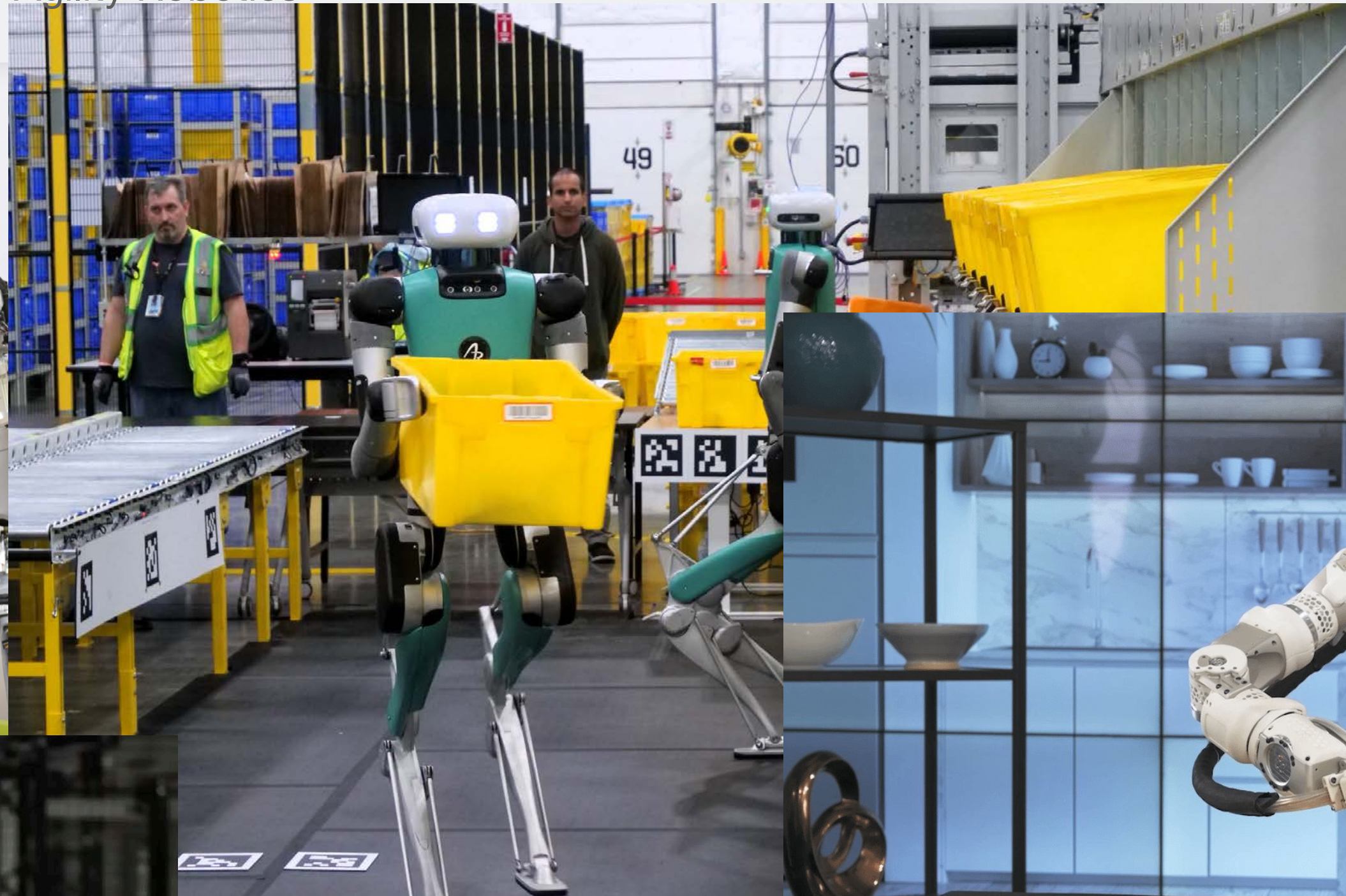






Tesla Optimus

Agility Robotics



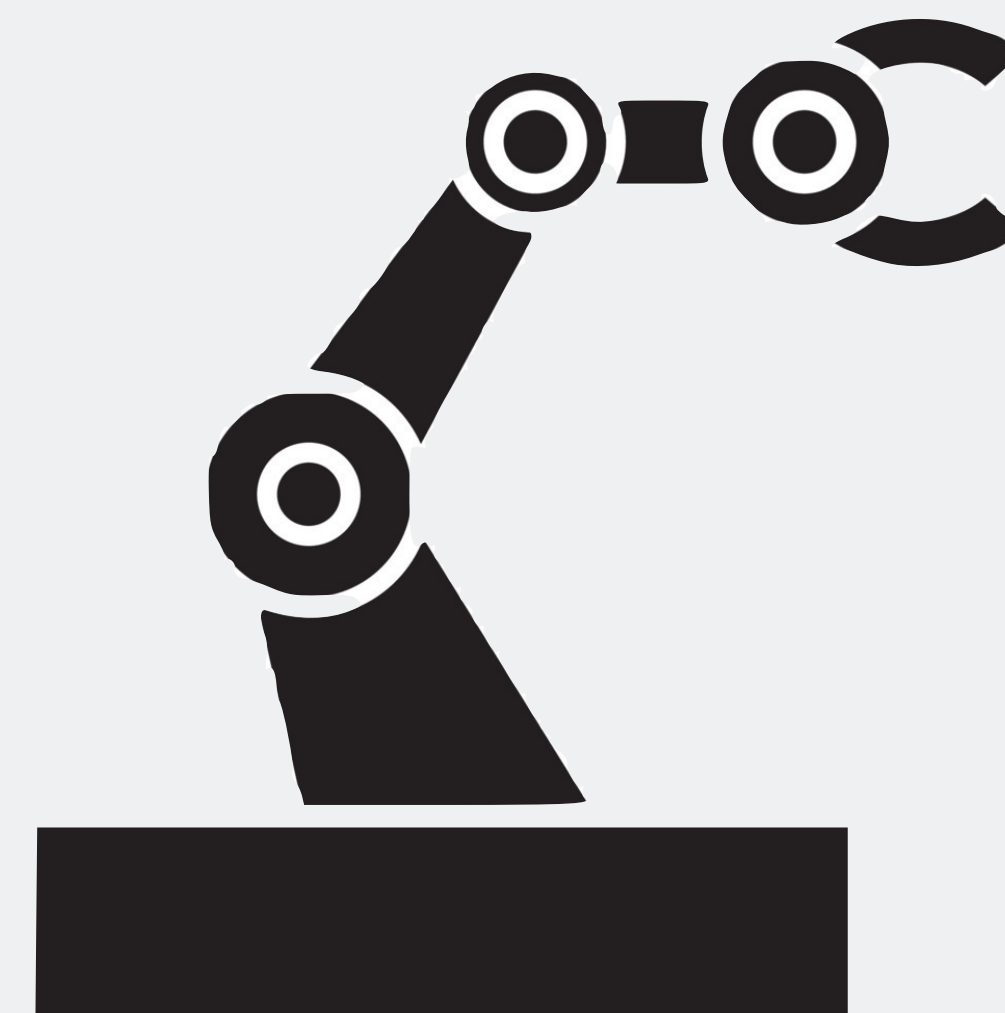
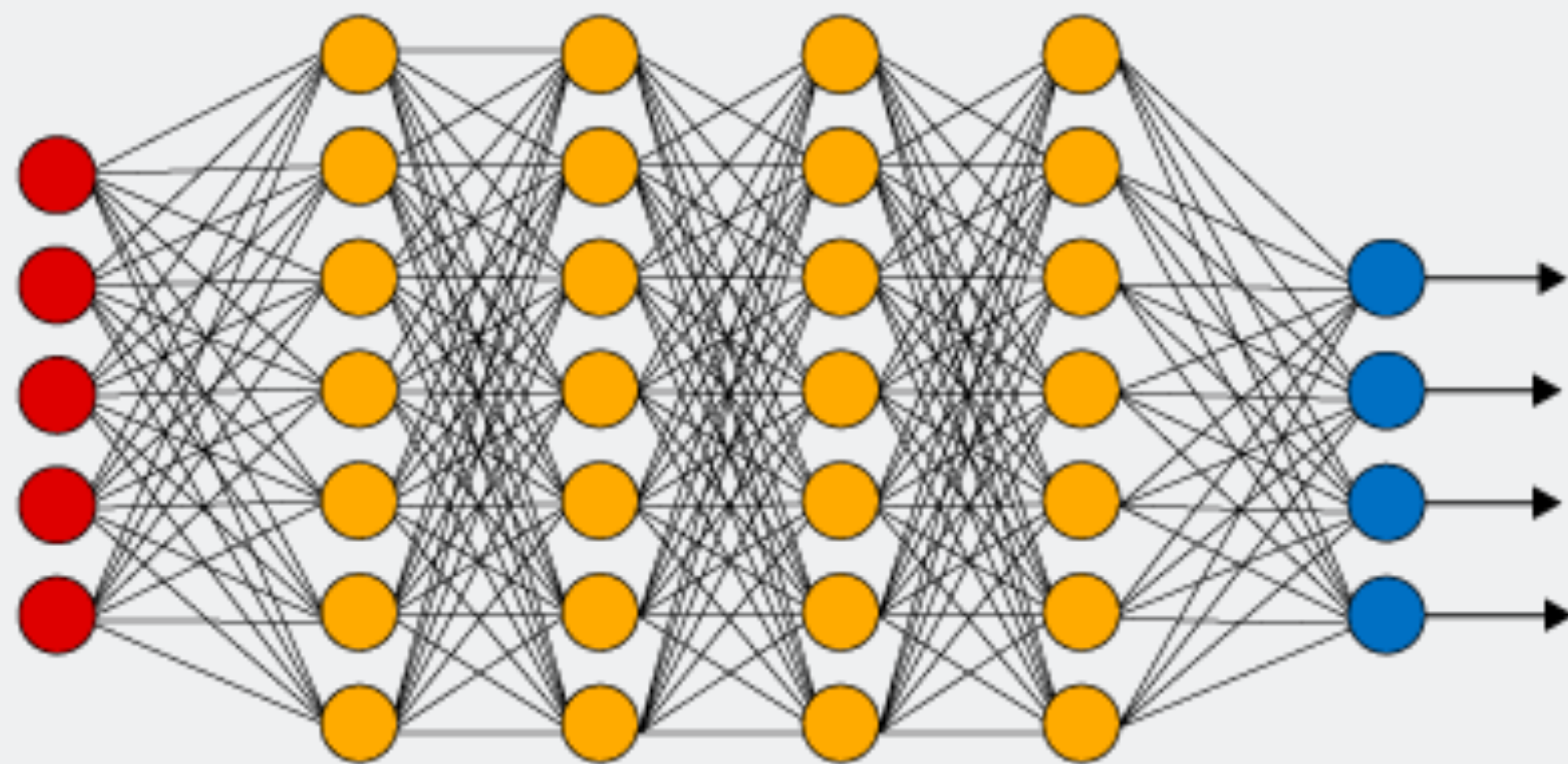
Sanctuary AI



Figure 01

Aptronik

Deep Learning X Robot Manipulation



But what's the key ingredient???

Training data



Tesla



Nvidia

from human teleoperation!!!





Teleoperation

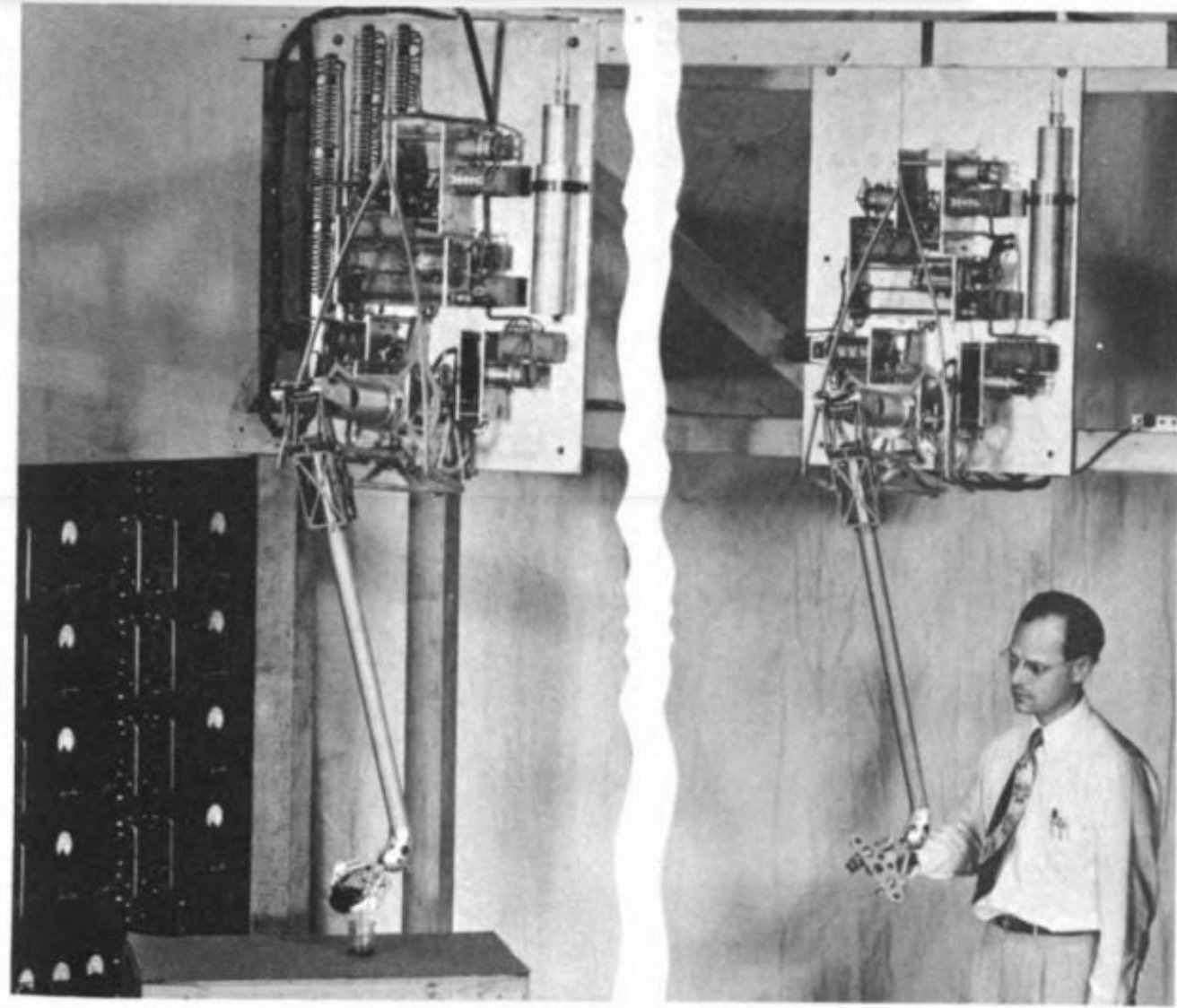


FIGURE 6.—The ANL Model E1 electric master slave. Used only for experimental purposes, this bilateral manipulator was developed in 1954. (Courtesy of Argonne National Laboratory.)



da Vinci robot





Teleoperation

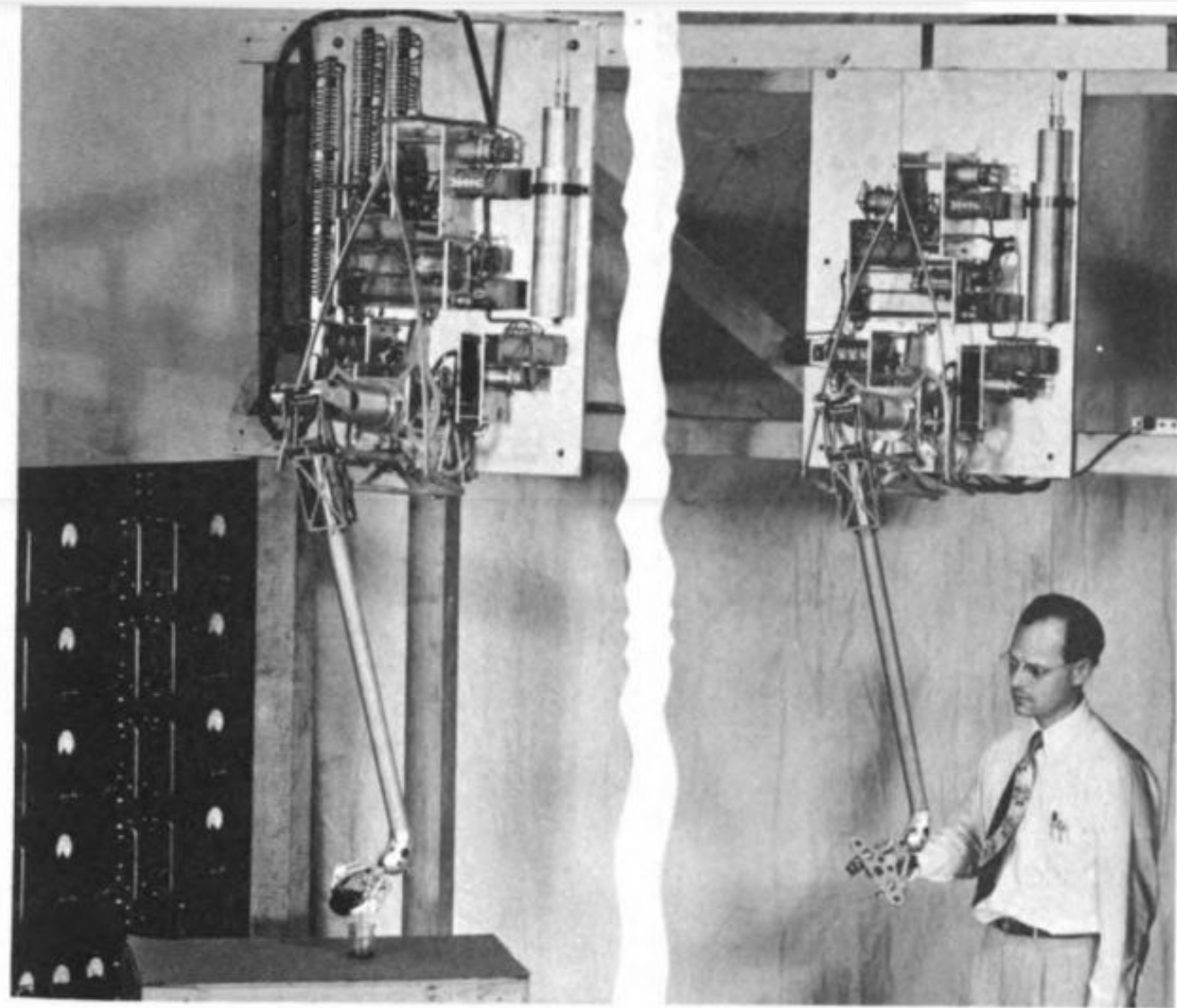


FIGURE 6.—The ANL Model E1 electric master slave. Used only for experimental purposes, this bilateral manipulator was developed in 1954. (Courtesy of Argonne National Laboratory.)



da Vinci robot



PR1 robot: <https://www.youtube.com/watch?v=qBZPSTR96N4>

This video showcases PR1, a robot developed by Keenan Wyrobek, Eric Berger, HFM Van der Loos, and Ken Salisbury at Stanford. It is the predecessor of the Willow Garage PR2. The PR1 was tele-operated during this entire video.





Teleoperation

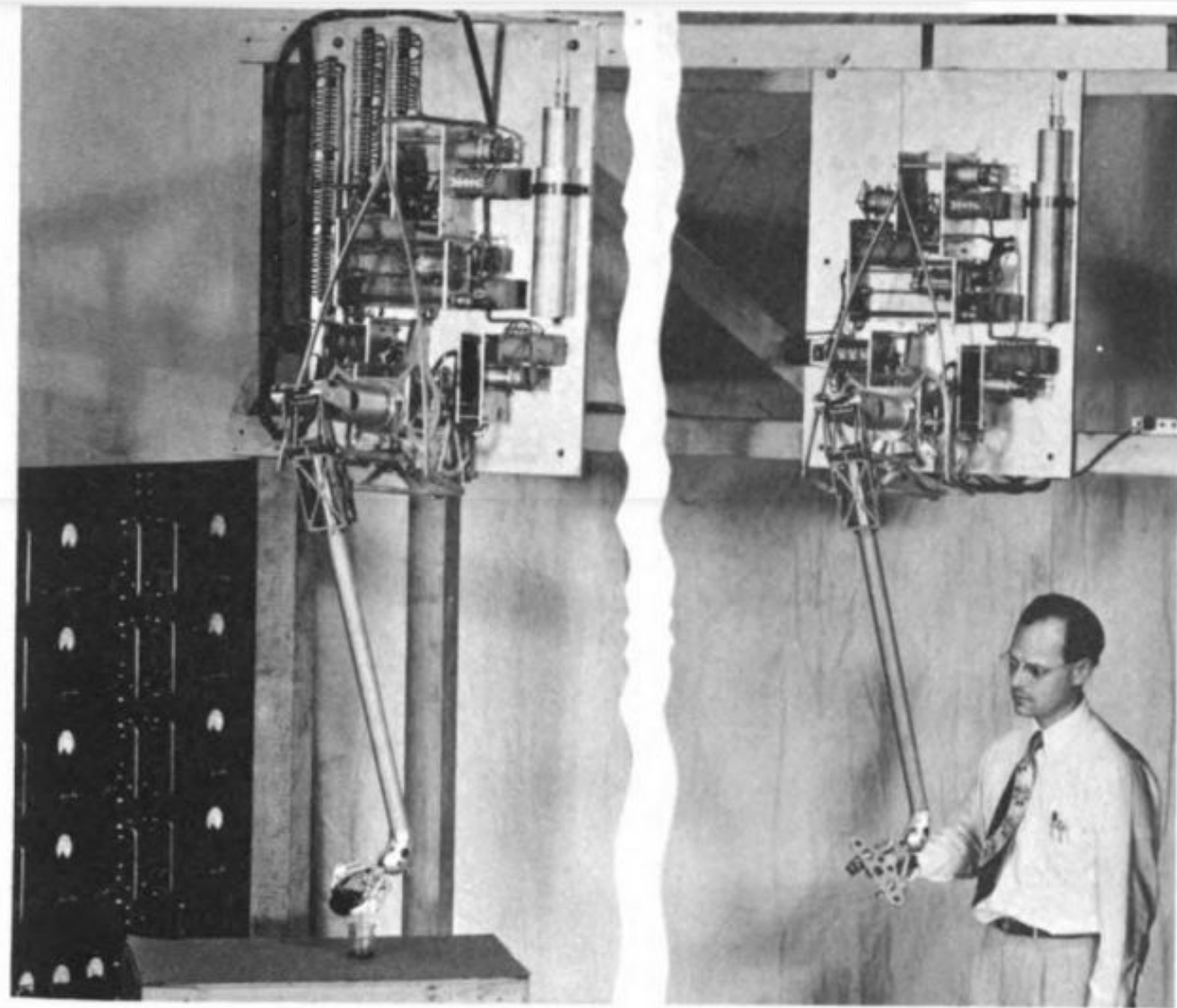
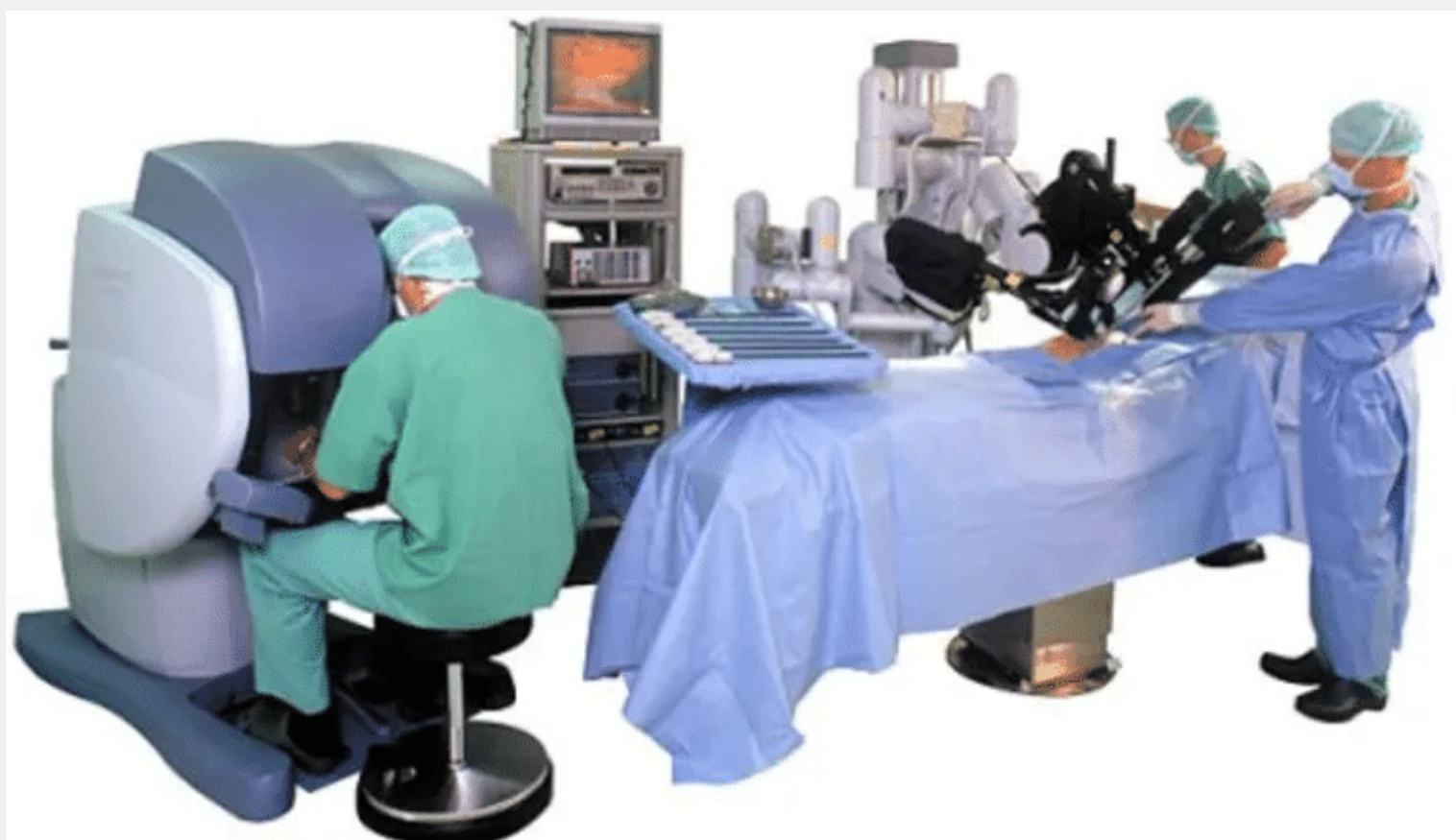


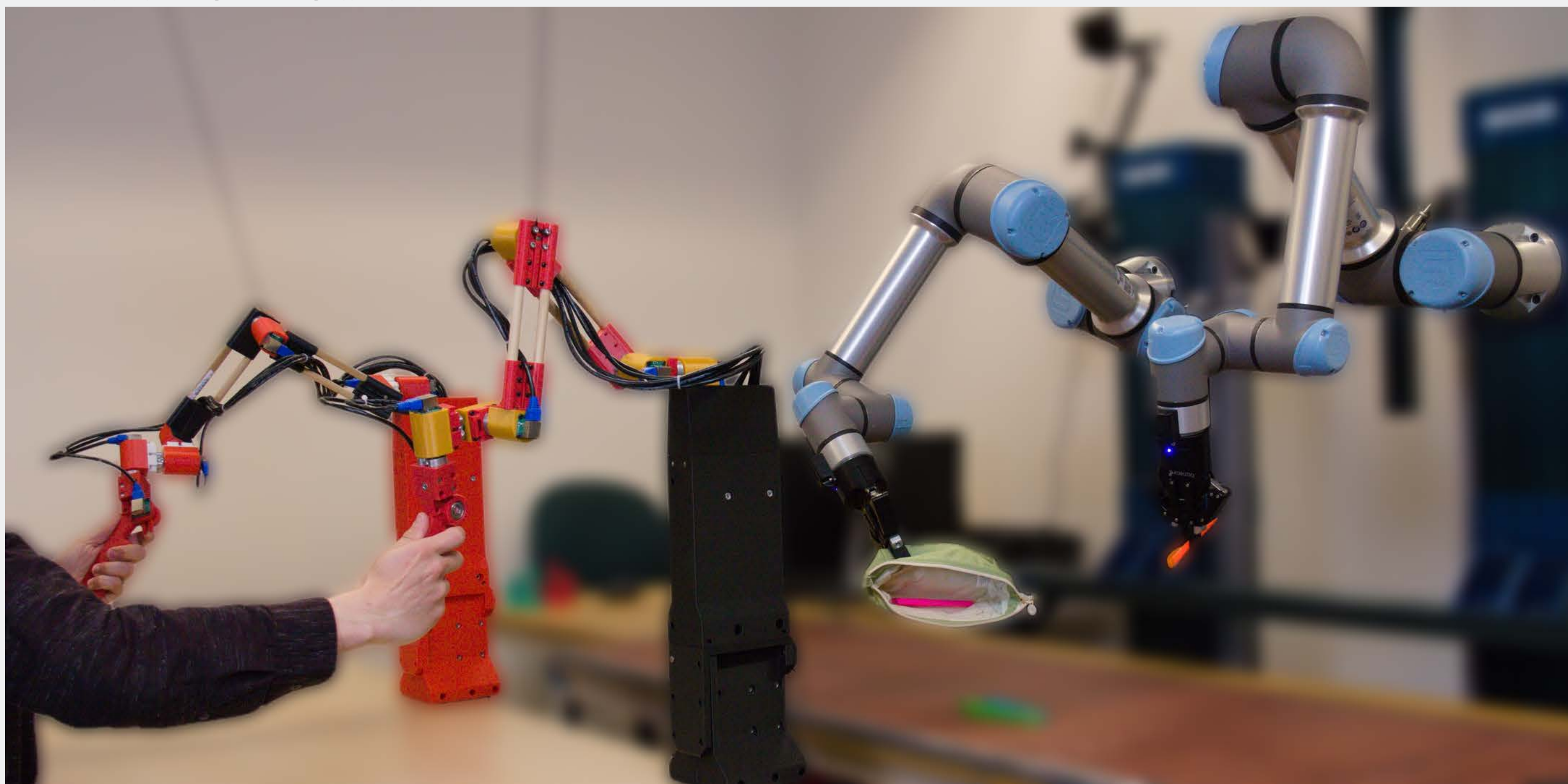
FIGURE 6.—The ANL Model E1 electric master slave. Used only for experimental purposes, this bilateral manipulator was developed in 1954. (Courtesy of Argonne National Laboratory.)



Mobile ALOHA - Stanford

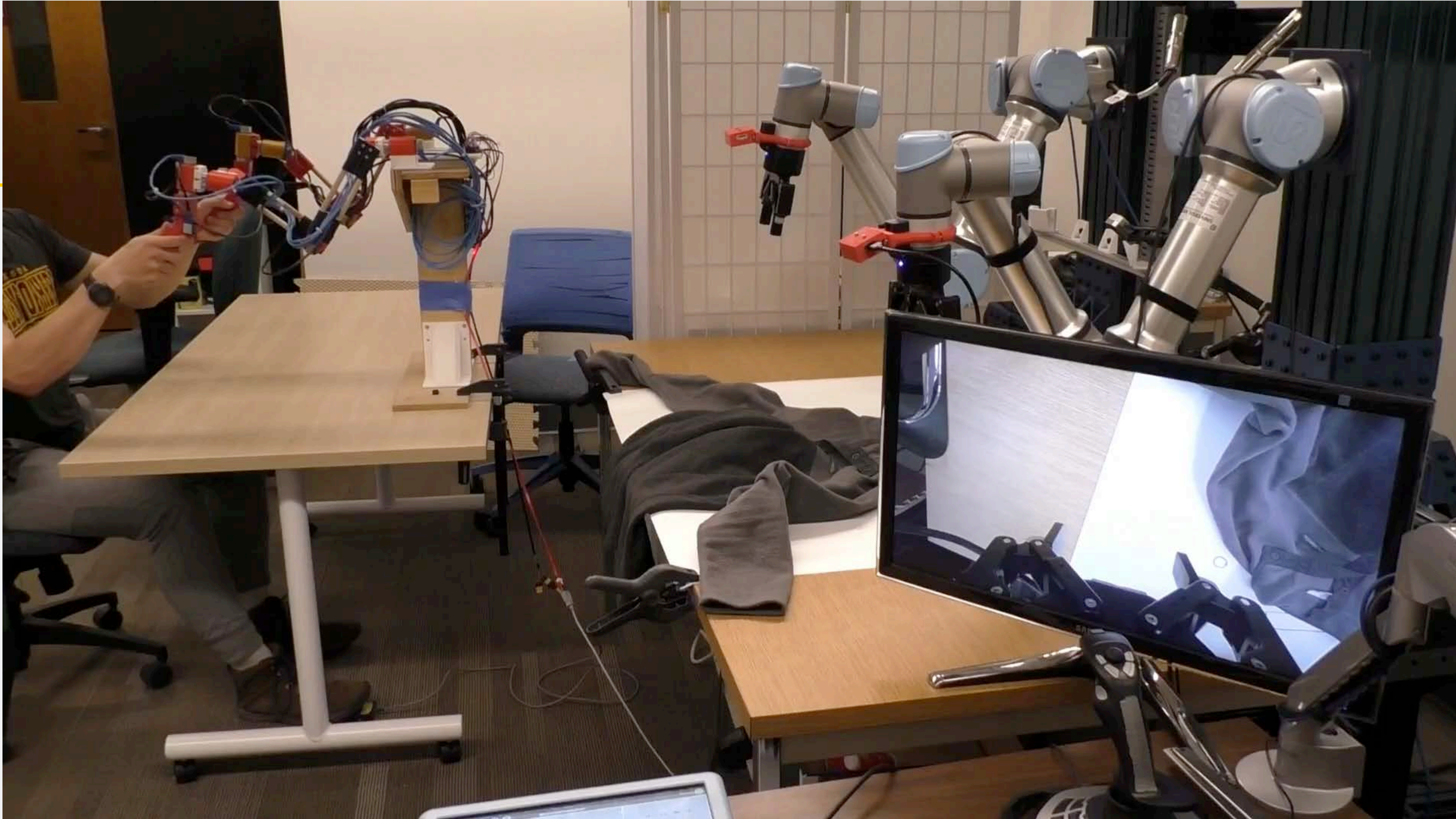


da Vinci robot



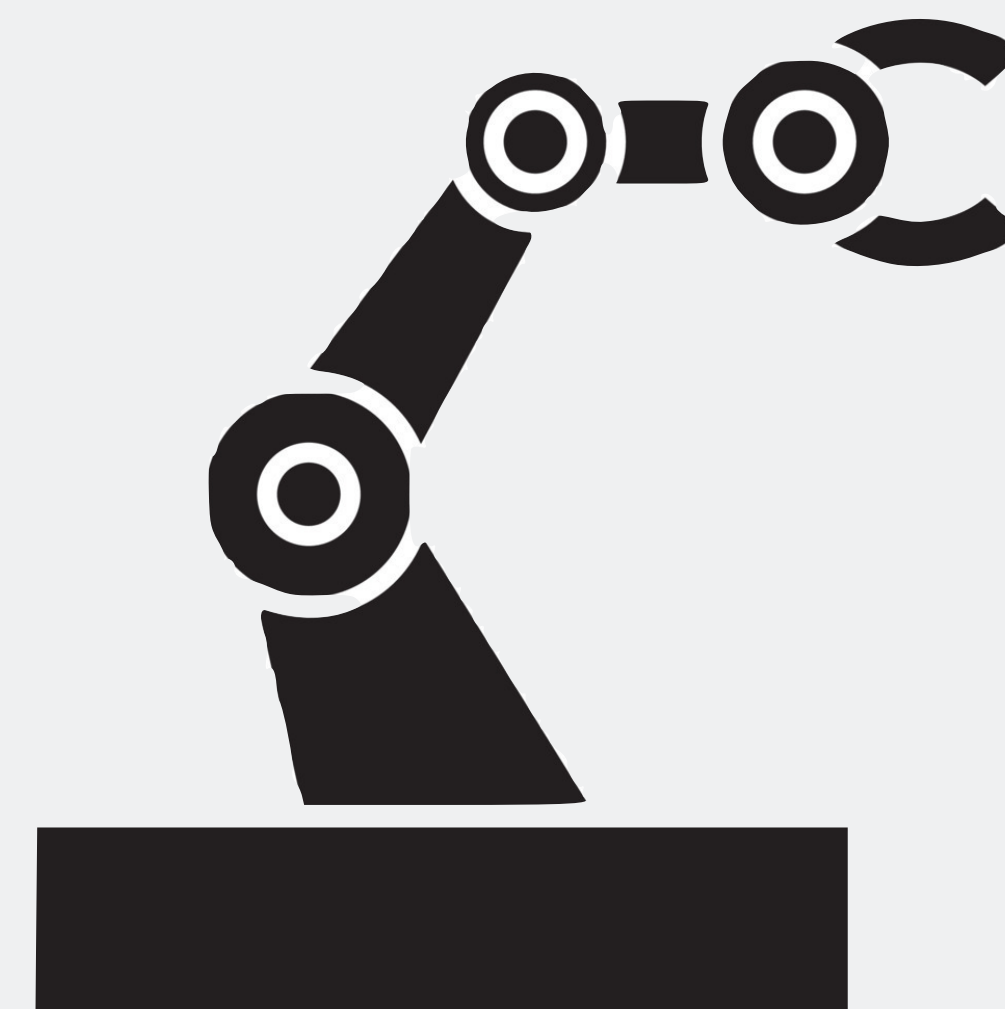
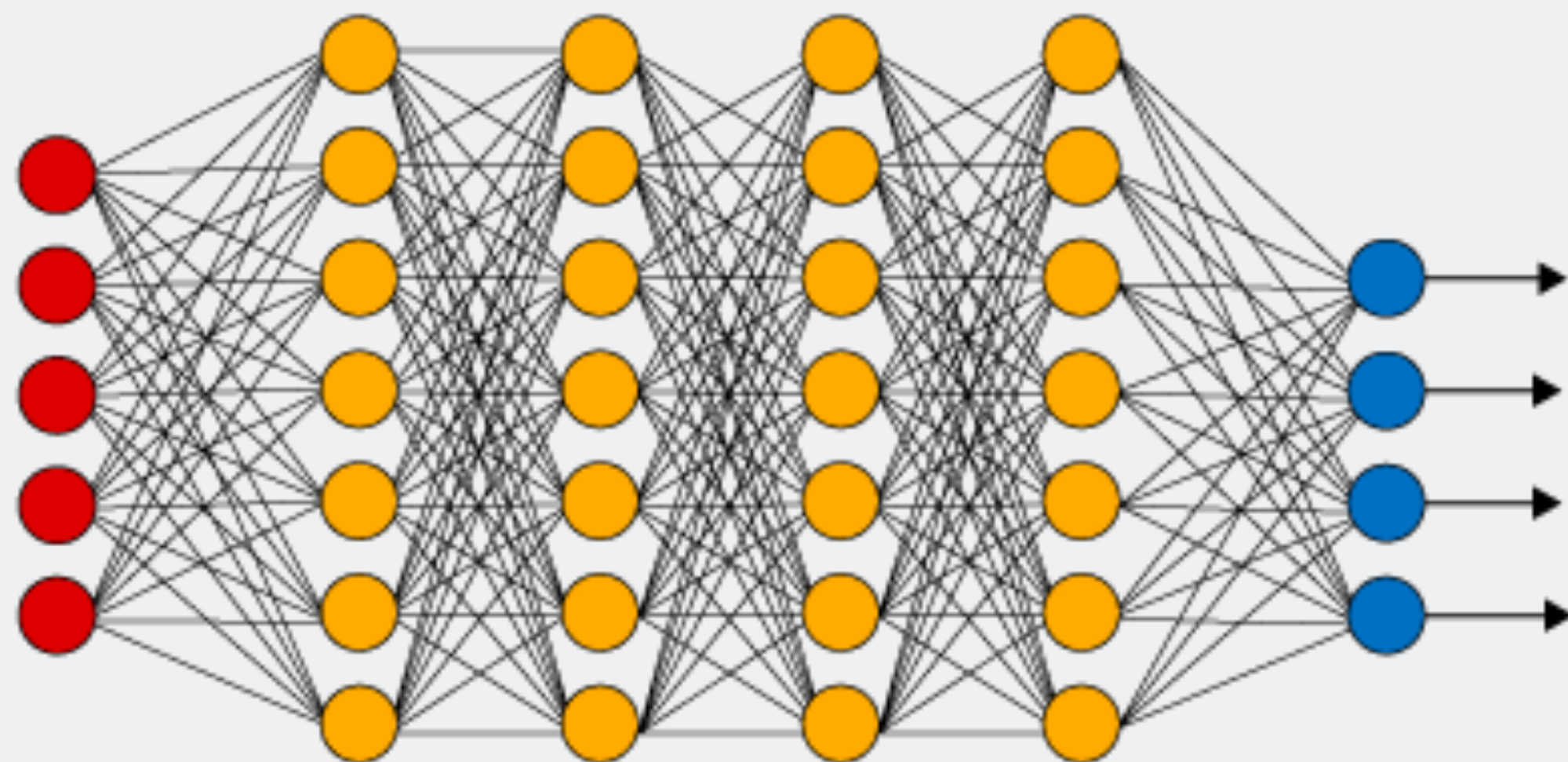
SPARK Teleoperation setup for Dual-arm task demonstrations @ RPM Lab - Adam Imdieke





SPARK Teleoperation setup for Dual-arm task demonstrations @ RPM Lab - Adam Imdieke

Deep Learning X Robot Manipulation



But what's the key ingredient???

Large number of high-quality expert data

Is that it???

Not really!!! Lets see few more examples...



4Progress



4Progress

Sui, Zhou, Zeng, Jenkins 2017

**Goal:
Perception and manipulation
of objects in clutter...**

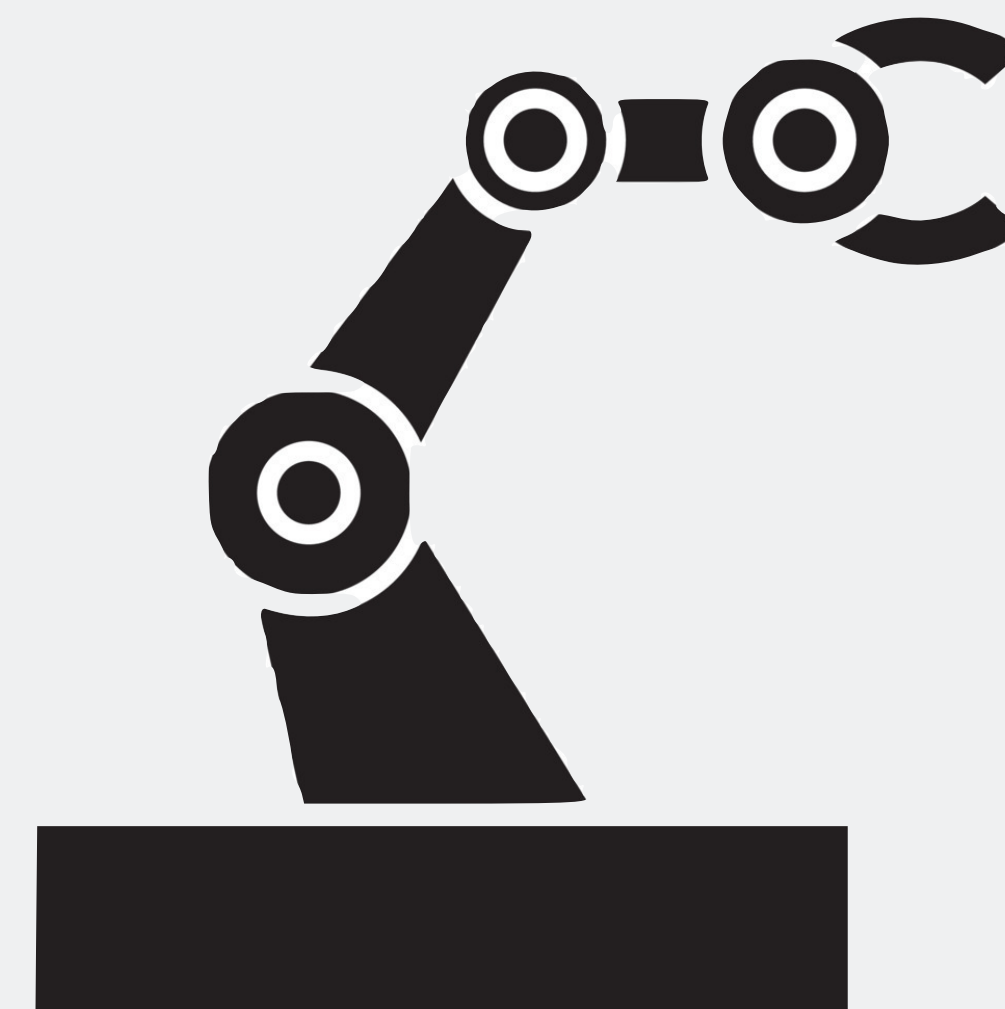
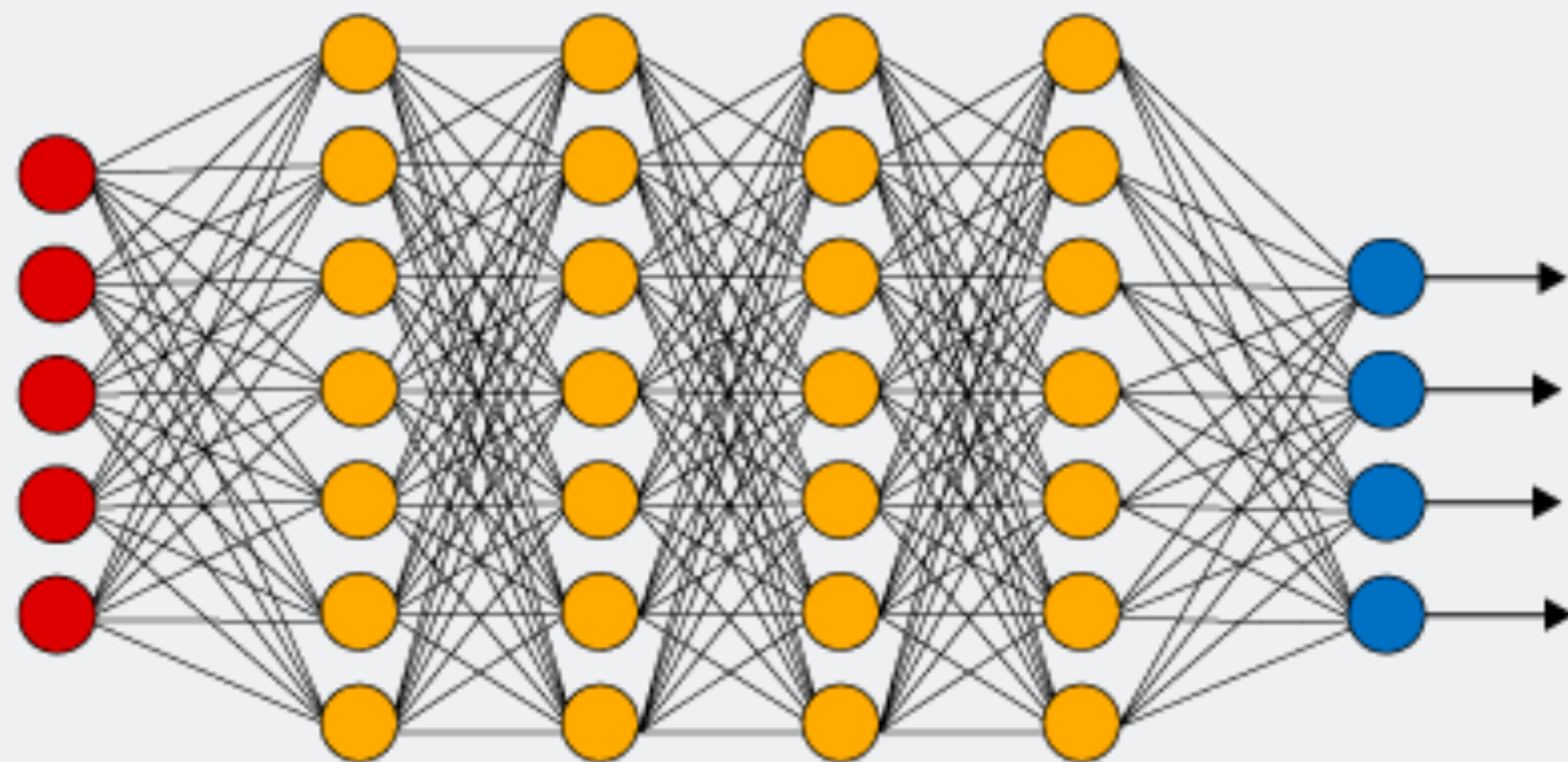


DR

... through deep learning with neural networks (mostly) ...



Deep Learning X Robot Manipulation



But what's the key ingredient???

Large number of high-quality expert data

Is that it???

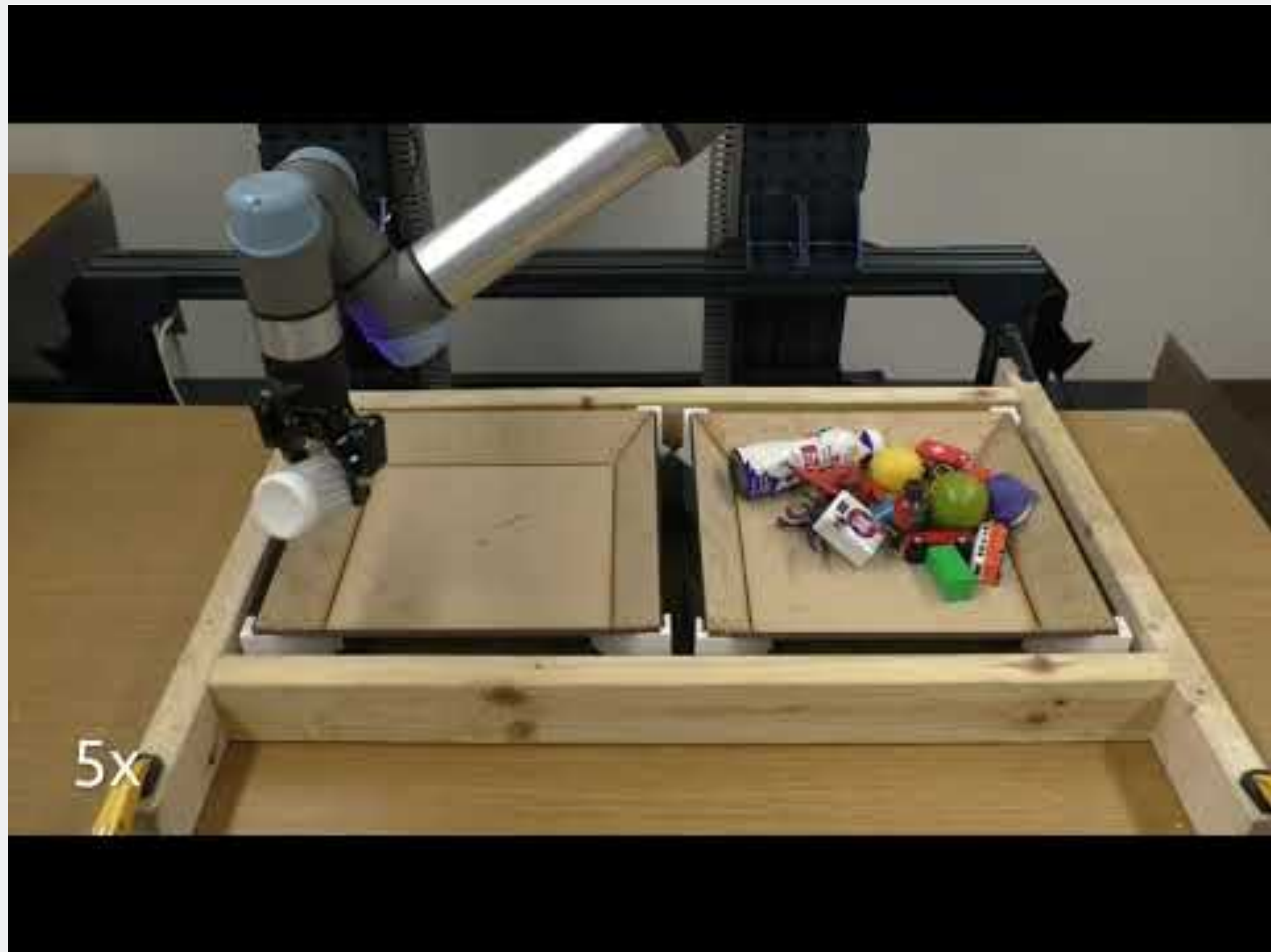
Not really!!! Lets see few more examples...

Robot Perception is hard!





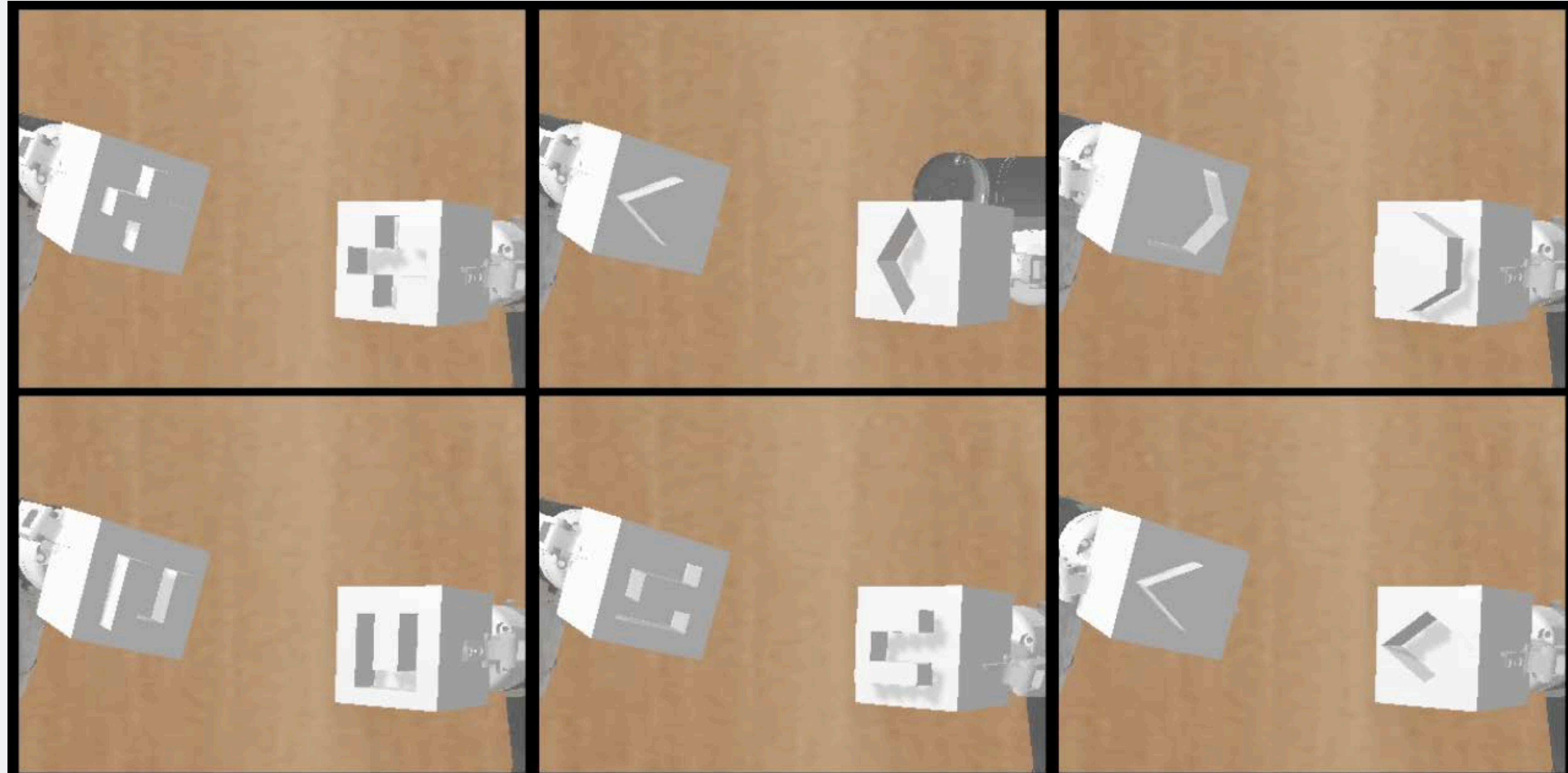
Grasp Learning



Carl Winge's reimplemention of the paper -
"Sample Efficient Grasp Learning Using Equivariant Models," by
Xupeng Zhu, Dian Wang, Ondrej Biza, Guanang Su, Robin Walters, and Robert Platt, RSS 2022.



Manipulation Learning for Object Assembly



**Robotics:
Perception & Manipulation
(RPM) Lab**

Chahyon Ku, Carl Winge, Ryan Diaz, Wentao Yuan, Karthik Desingh
"Evaluating Robustness of Visual Representations for Object Assembly Task Requiring Spatio-Geometrical Reasoning,"
ICRA 2024.



Deep Learning X Robot Manipulation

Google search results for "Deep Learning for Robot Manipulation".

Search filters: Object detection, Robot hand manipulation, Dexterous manipulation, Nvidia, Simulation, Mobile manipulation, Imitation learning, Robot grasping, Isaac gym, Robot arm.

Search results include:

- Shadow Robot: Artificial Intelligence & Machine Learning
- NVIDIA Developer: Advancing Robotic Assembly with a New...
- University of Bristol: New dual-arm robot achieves bimodal...
- The Serious Computer Vision: Convolutional Neural Networks
- Frontiers: Frontiers | Reinforcement Learning
- IEEE Spectrum: Google Wants Robots to Acquire New Skills
- people.kth.se: Michael C. Welle
- IEEE Robotics and Automation Society: Robotic Hands, Grasping and Manipulation
- YouTube: QT-Opt: Scalable Deep Reinforcement Learning
- Medium: Google X's Deep Reinforcement Learning
- Naver Labs Europe: Learning robot manipulation – modelling the reachable space of a robot and its inverse mapping
- msail.github.io: Deep RL for Robotics
- CSIRO Research: Robotic Manipulation Tasks
- MathWorks: AI for Robotics - MATLAB & Simulink
- Chinese Academy of Sciences: Enhance Robot Autonomy via Deep Learning
- SpringerLink: Survey of imitation learning for robot manipulation
- Michigan Robotics - University of Michigan: Deep Learning for Robotics | Michigan Robotics
- P.C. Rossin College of Engineering & Applied Sciences: robot object manipulation skills
- Berkeley Artificial Intelligence Research: One-Shot Imitation from Watching Videos
- DeepAI: Improved Learning for Robot Manipulation
- GitHub Pages: Deep Imitation Learning for Humanoid Robot Manipulation
- MIT OpenCourseWare: Robotic Manipulation | EECS-6.897
- Everyday Robots: Scalable Deep Reinforcement Learning

Related searches:

- reinforcement learning robot gif
- reinforcement learning robot example
- reinforcement learning robot

... understand what the hype is all about ...

Robot Perception is hard!





Welcome to DeepRob!



DR

Welcome to DeepRob!

a Distributed Teaching Collaborative



Prof. Karthik Desingh



Dr. Xiaoxiao Du



Anthony Opipari



Prof. Chad Jenkins



UNIVERSITY
OF MINNESOTA













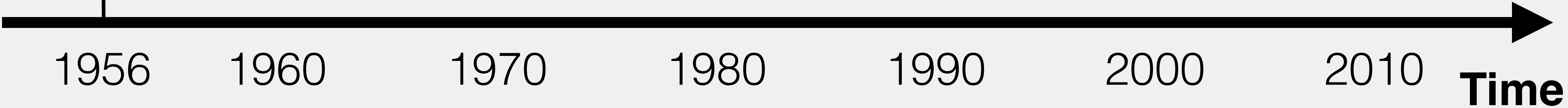


A brief history of AI



1956 Dartmouth Conference: The Founding Fathers of AI

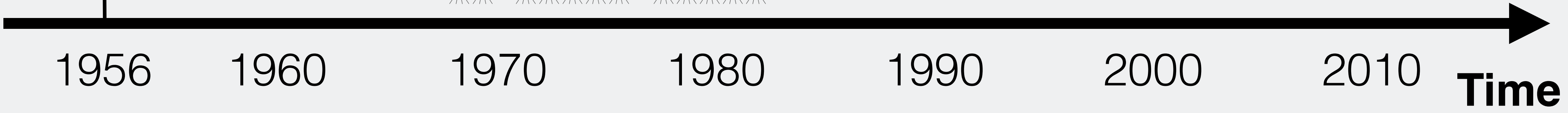
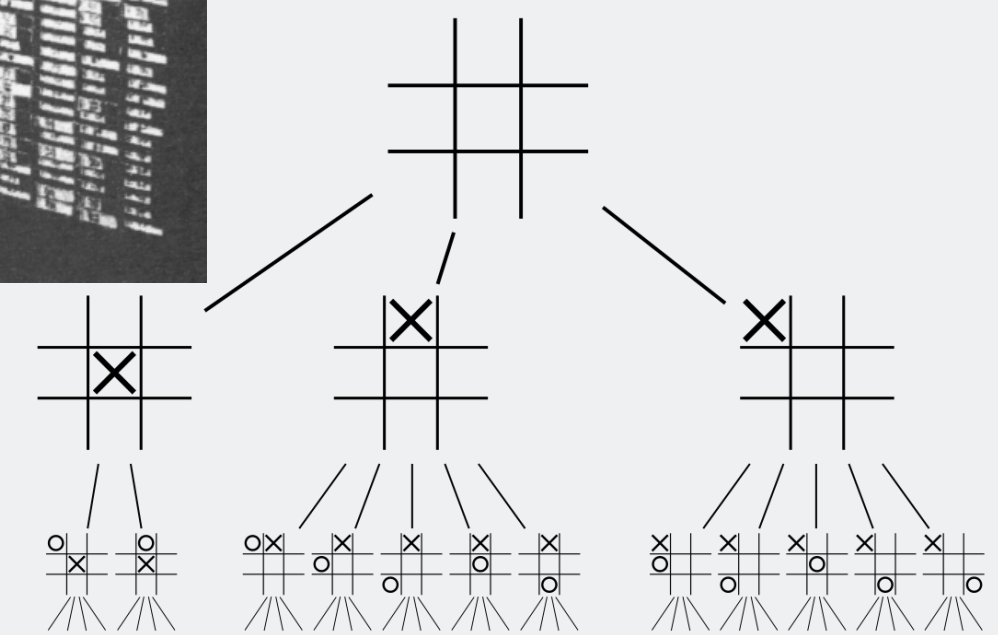
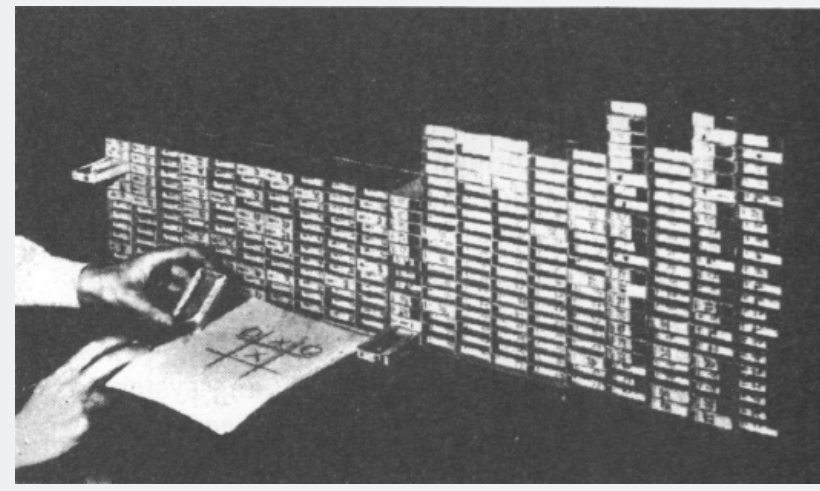
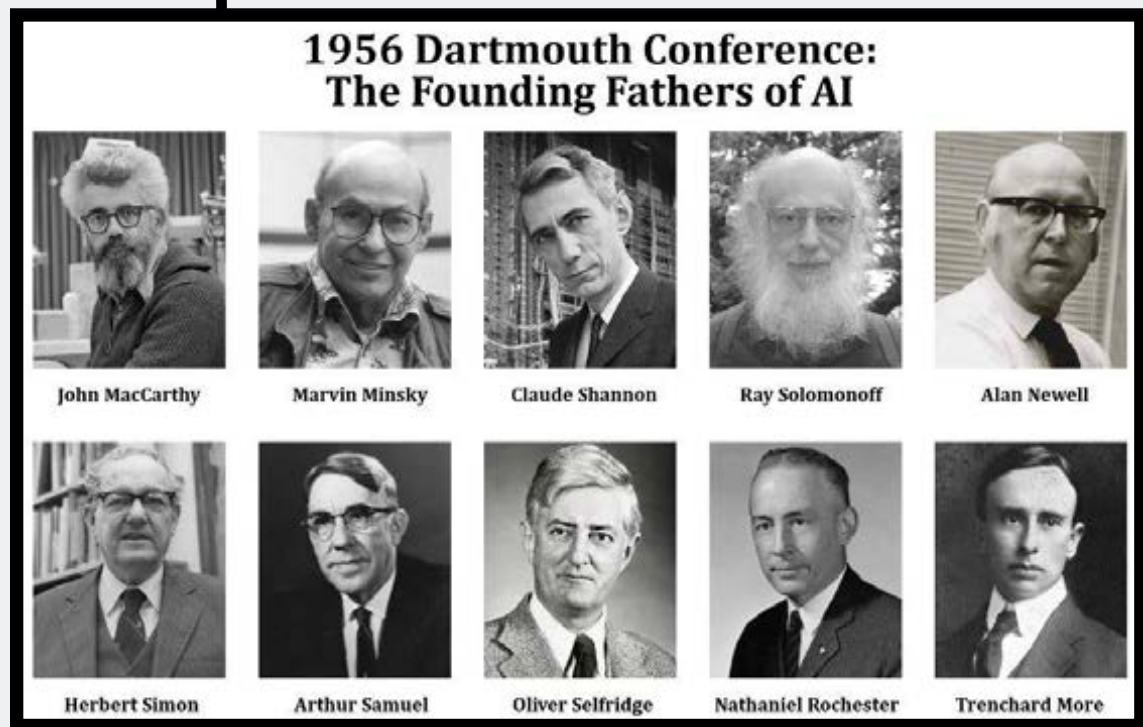
				
John MacCarthy	Marvin Minsky	Claude Shannon	Ray Solomonoff	Alan Newell
				
Herbert Simon	Arthur Samuel	Oliver Selfridge	Nathaniel Rochester	Trenchard More





First wave AI: Model-based

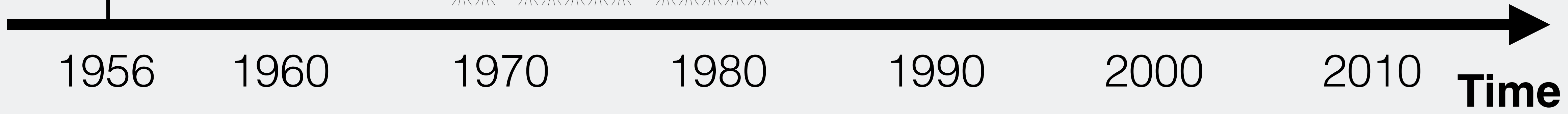
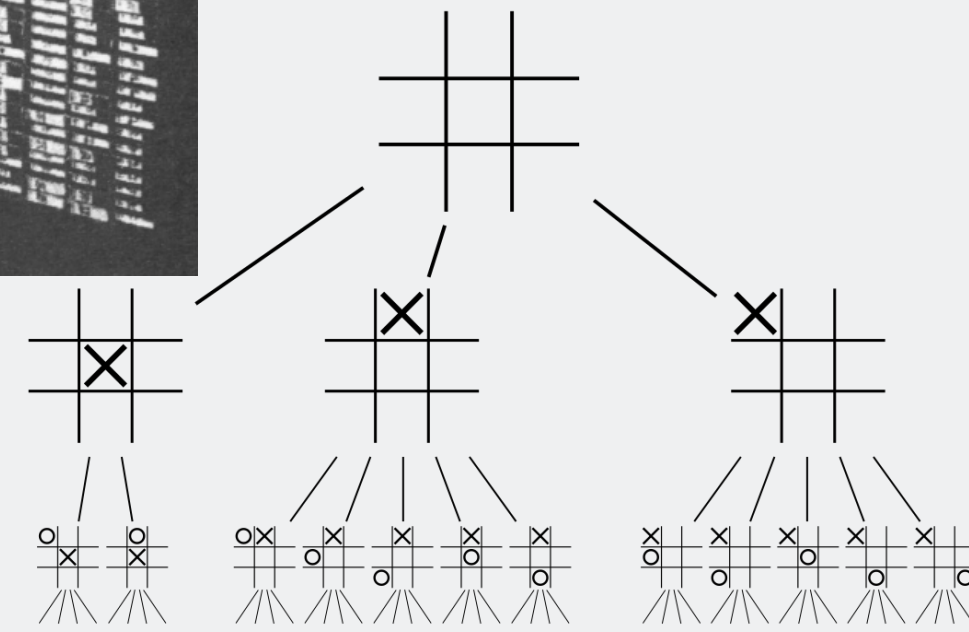
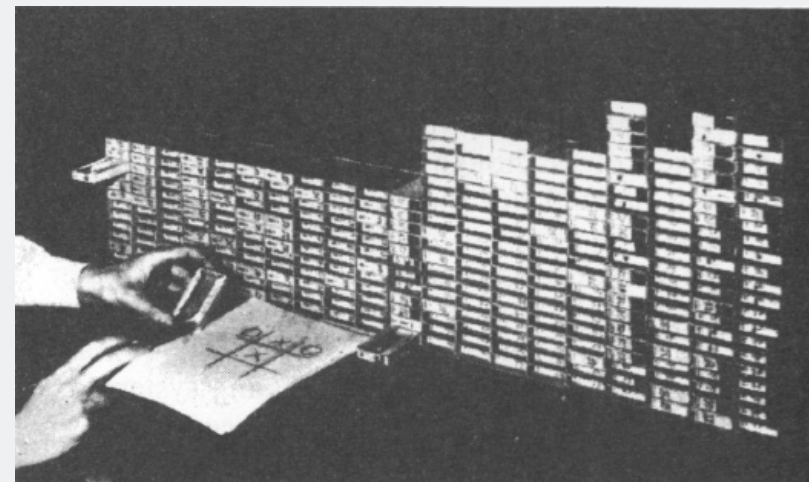
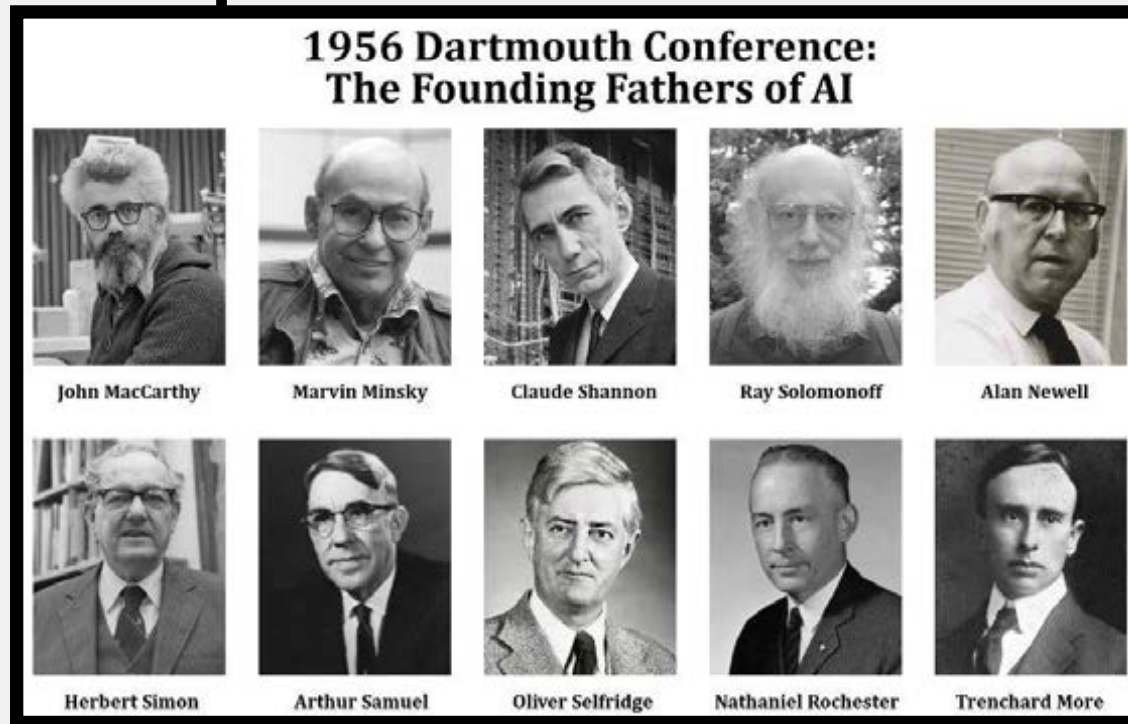
“Think through the entire problem”





First wave AI: Model-based

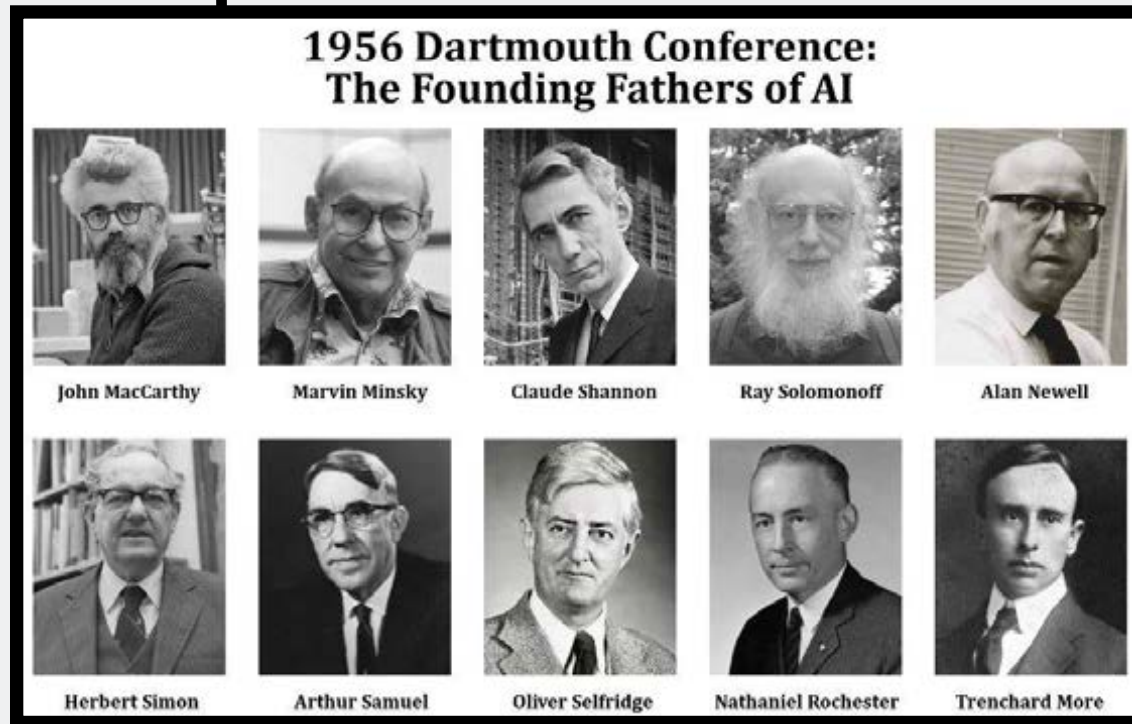
“Think through the entire problem”



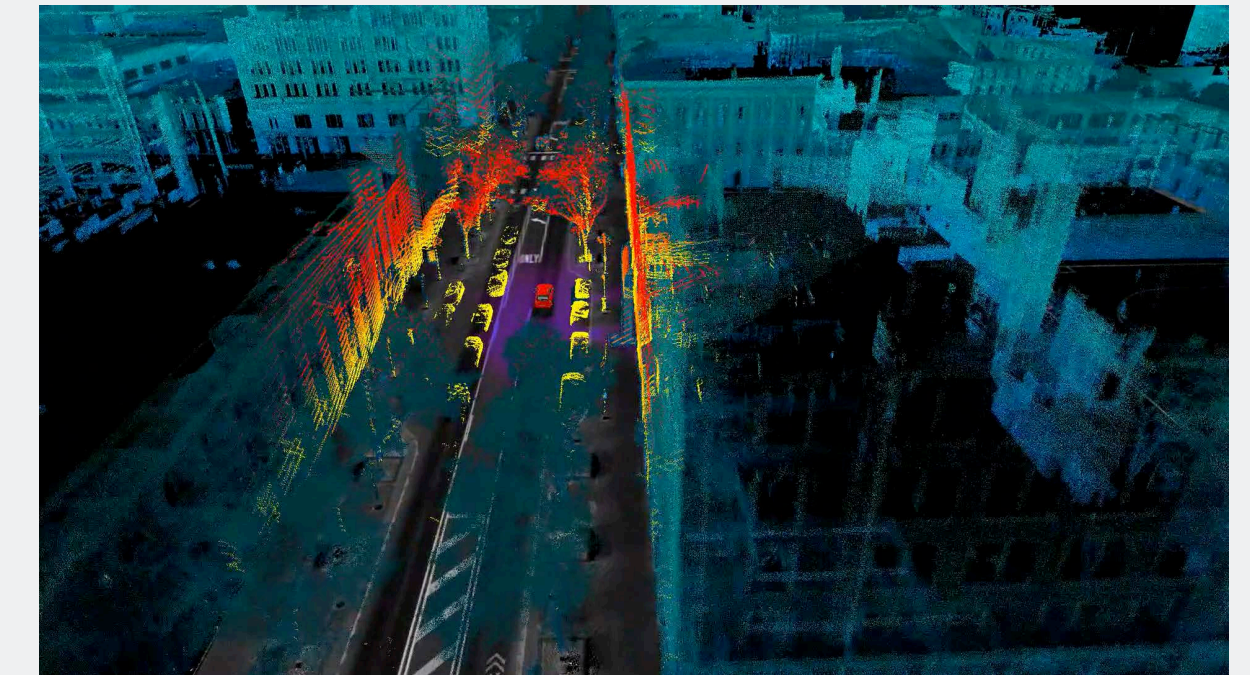


First wave AI: Model-based

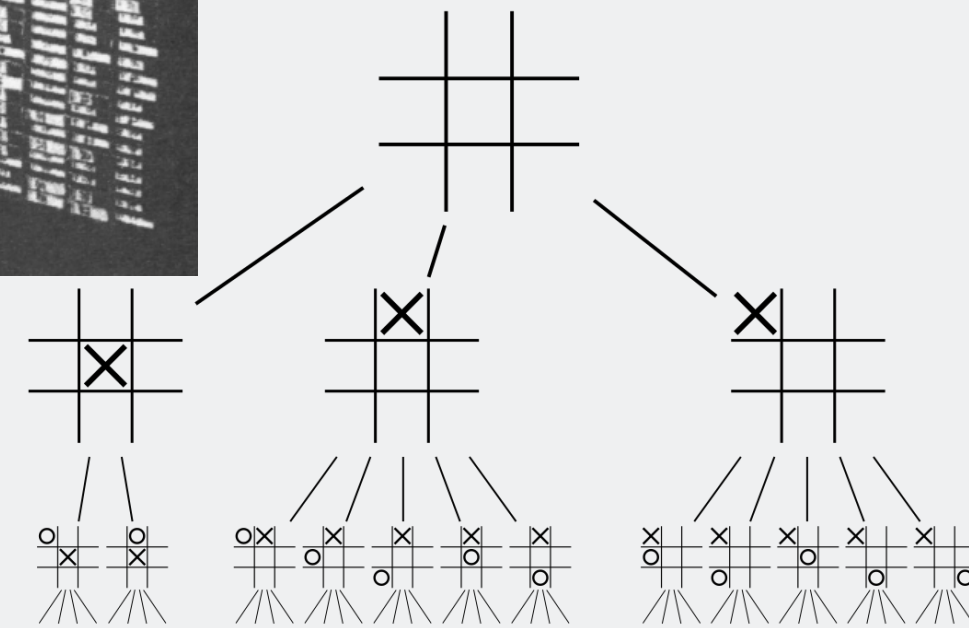
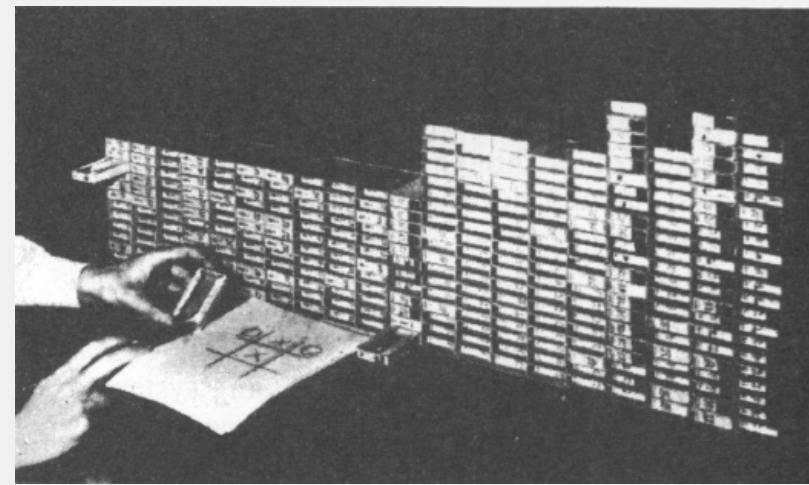
“Think through the entire problem”



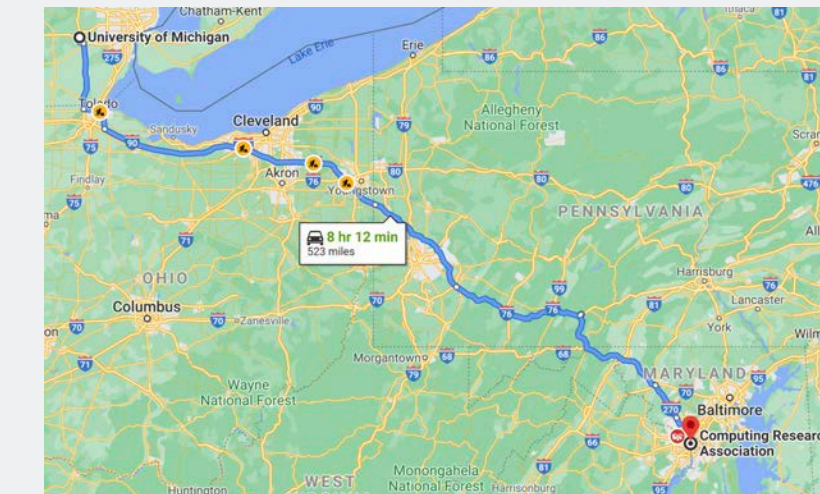
City-scale 3D mapping



Self-driving cars



Online route navigation



1956

1960

1970

1980

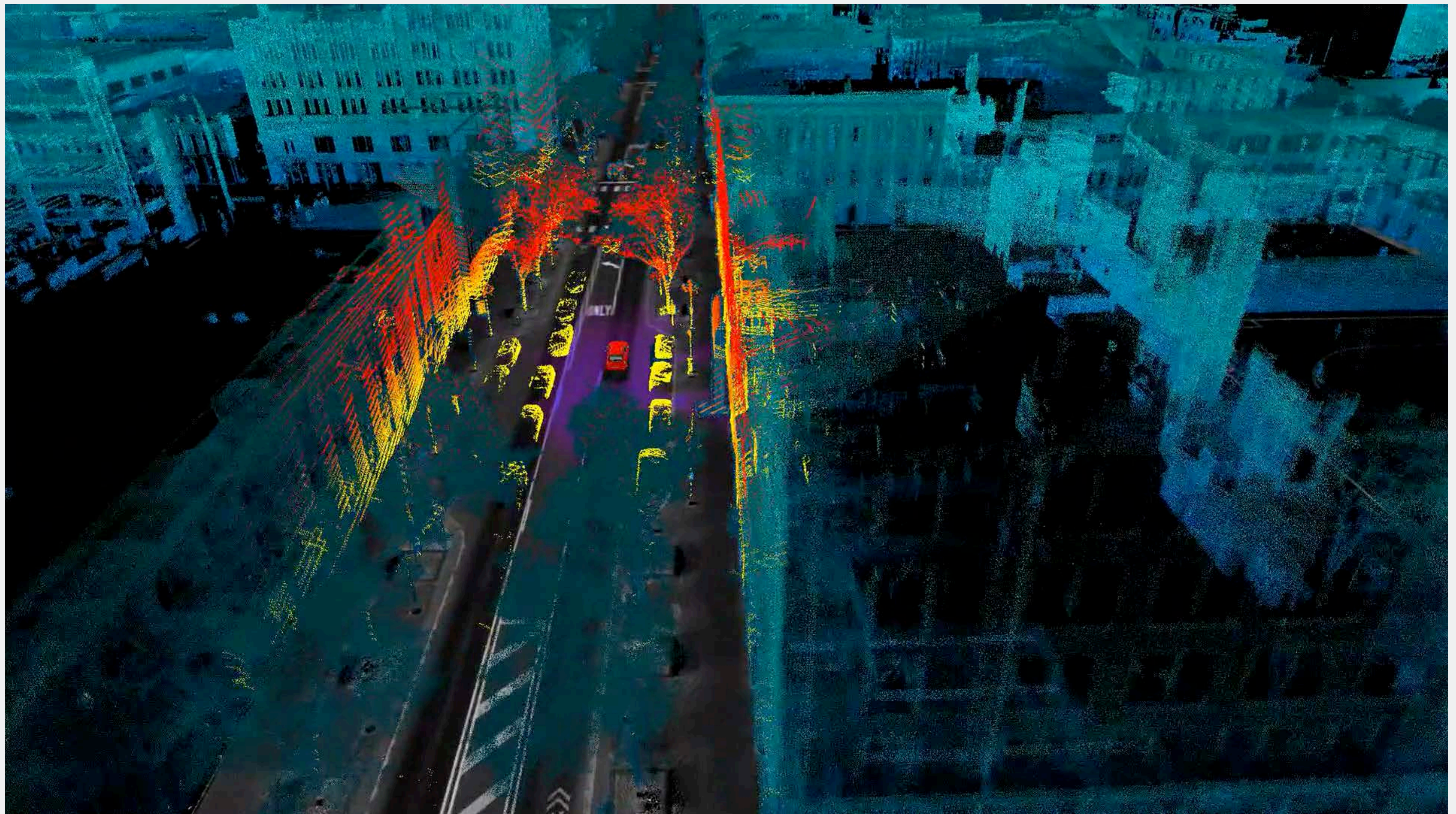
1990

2000

2010

Time





Let's watch this video closely





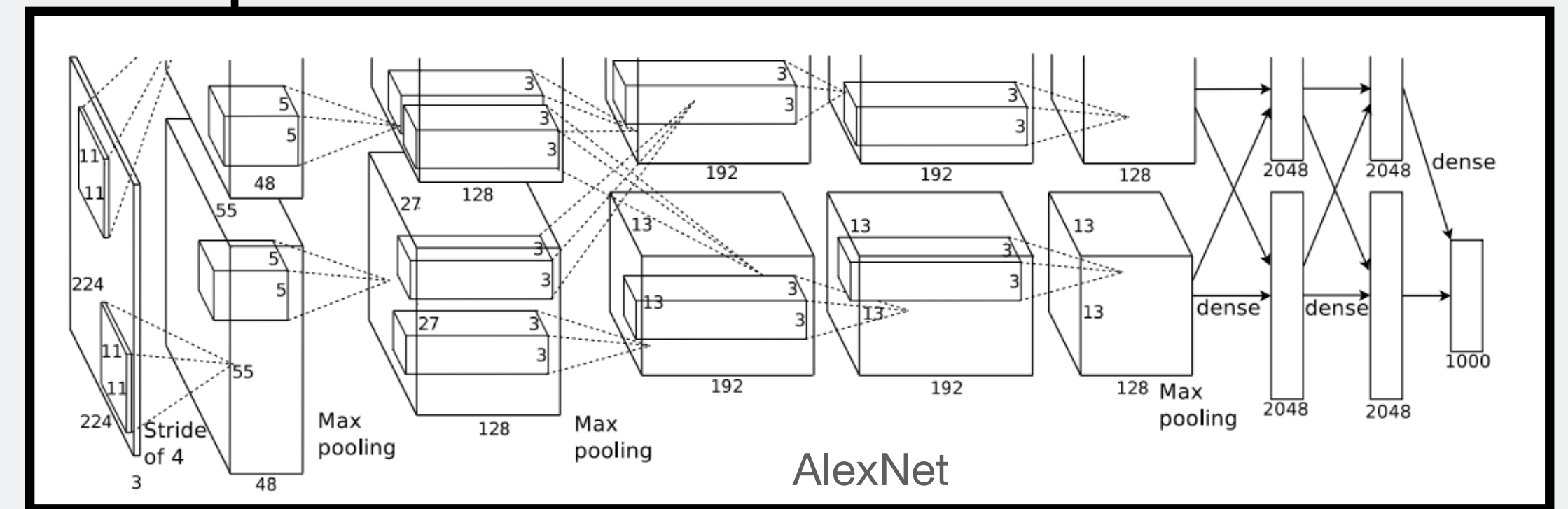
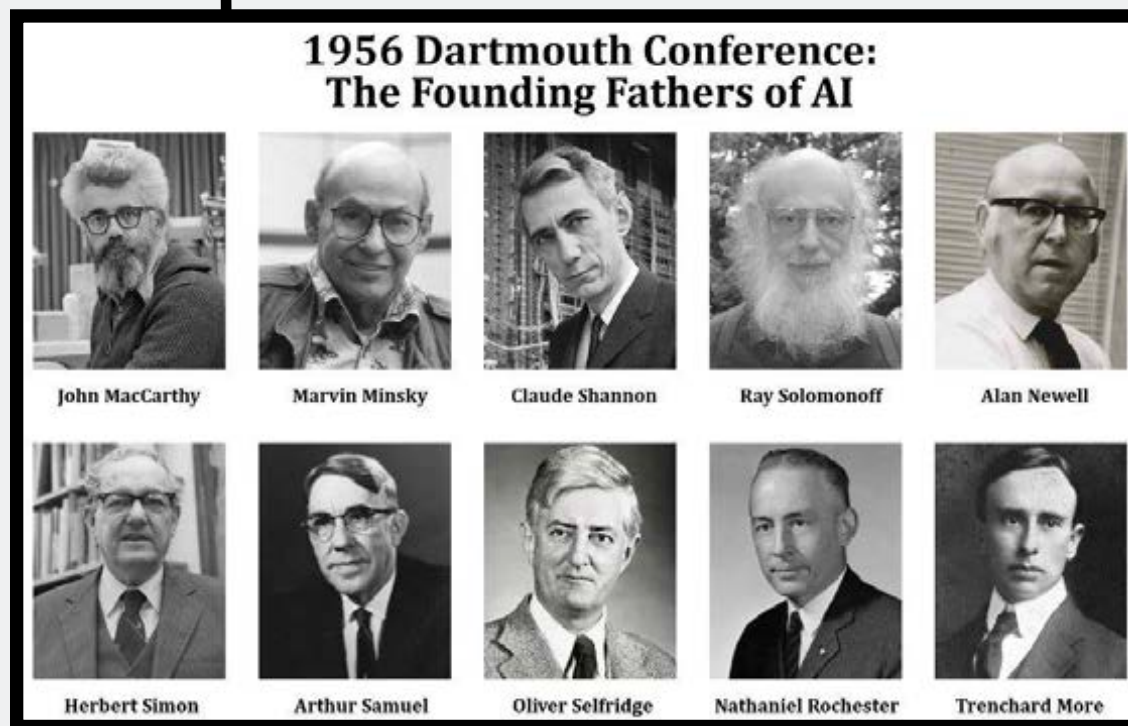
First wave AI: Model-based

“Think through the entire problem”

Second wave AI: Data-driven

“Learn from lots of data”

Rise of “deep learning”



1956

2011

Time





First wave AI: Model-based

IEEE SPECTRUM Engineering Topics Special Reports Blogs Multimedia The Magazine Professional Resources Search

Cars That Think | Transportation | Advanced Cars

09 Feb 2016 | 17:00 GMT

Deep Learning Makes Driverless Cars Better at Spotting Pedestrians

Pedestrian detection systems for cars could become faster and more accurate with help from deep learning algorithms

By Jeremy Hsu

Images: Statistical Visual Computing Lab/UC San Diego

Second wave AI: Data-driven

“Learn from lots of data”



“deep learning”



1956

2011

Time





TEXT PROMPT
an illustration of a baby daikon radish in a tutu walking a dog

AI-GENERATED IMAGES

Edit prompt or view more images ↓

TEXT PROMPT
an armchair in the shape of an avocado. . . .

AI-GENERATED IMAGES

Edit prompt or view more images ↓

DALL-E (OpenAI)

Second wave AI: Data-driven

“Learn from lots of data”



“deep learning”

2011

Time →





TEXT PROMPT
an illustration of a baby daikon radish in a tutu walking a dog

AI-GENERATED IMAGES

Edit prompt or view more images

TEXT PROMPT
an armchair in the shape of an avocado

AI-GENERATED IMAGES

Edit prompt or view more images

DALL-E (OpenAI)

Your AI pair programmer

GitHub Copilot uses the OpenAI Codex to suggest code and entire functions in real-time, right from your editor.

[Start my free trial >](#) [Explore docs](#)

```

1 #!/usr/bin/env ts-node
2
3 import { fetch } from "fetch-h2";
4
5 // Determine whether the sentiment of text is positive
6 // Use a web service
7 async function isPositive(text: string): Promise<boolean> {
8   const response = await fetch(`http://text-processing.com/api/sentiment/`, {
9     method: "POST",
10    body: `text=${text}`,
11    headers: {
12      "Content-Type": "application/x-www-form-urlencoded",
13    },
14  });
15  const json = await response.json();
16  return json.label === "pos";
17 }

```

Copilot

Copilot (OpenAI)

Second wave AI: Data-driven

“Learn from lots of data”



“deep learning”

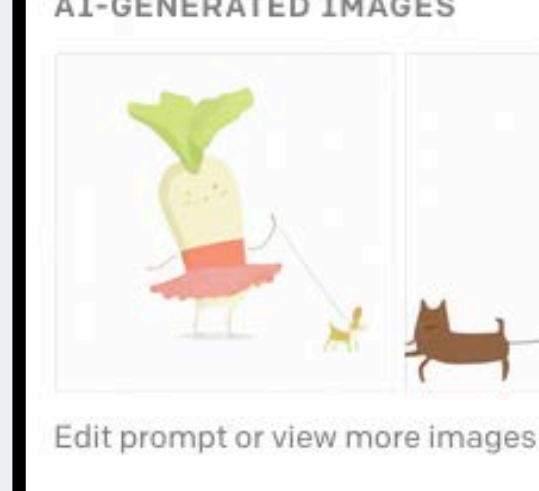
2011

Time

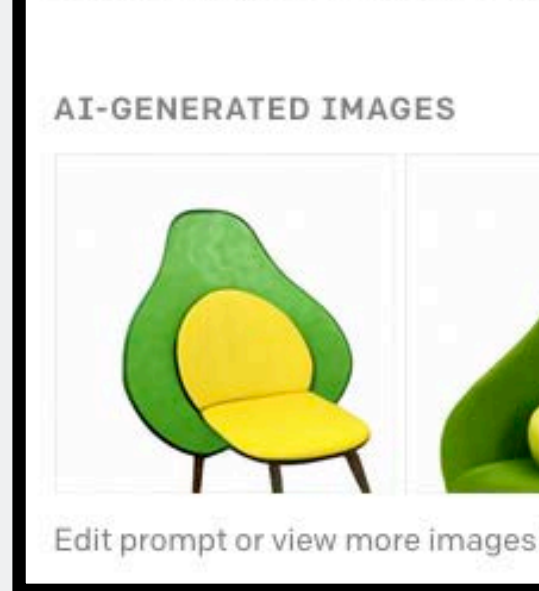




TEXT PROMPT
an illustration of a baby daikon radish in a tutu walking a dog



TEXT PROMPT
an armchair in the shape of an avocado



DALL-E (OpenAI)

Your AI pair programmer

GitHub Copilot uses the OpenAI Codex to suggest code and entire functions in

In the following sample, ChatGPT is able to understand the reference ("it") to the subject of the previous question ("fermat's little theorem").

Sample 3 of 4 Previous Next

User what is the fermat's little theorem

ChatGPT Fermat's Little Theorem states that if p is a prime number and a is an integer not divisible by p , then a^p (a to the power p) is congruent to a modulo p . In other words, if we divide a^p by p , the remainder is always a .

For example, if $p = 5$ and $a = 2$, then $2^5 = 32$, and 32 divided by 5 has a remainder of 2. Therefore, 2^5 is congruent to 2 modulo 5, and Fermat's Little Theorem holds for this case.

Fermat's Little Theorem is often used in cryptography and other applications where it is necessary to perform modular arithmetic operations quickly and efficiently. It is also a useful tool for proving other theorems in number theory

User how is it used in cryptography?

ChatGPT Fermat's Little Theorem is used in cryptography in several ways. One of the most common applications is in the generation of so-called "public-key" cryptography systems, which are used to securely transmit messages over the internet and other networks.

In a public-key cryptography system, each user has a pair of keys: a public key, which is widely known and can be used by anyone to encrypt a message intended for that user, and a private key, which is known only to the user and is used to decrypt messages that have been encrypted with the corresponding public key.

Copilot (OpenAI)

ChatGPT (OpenAI)

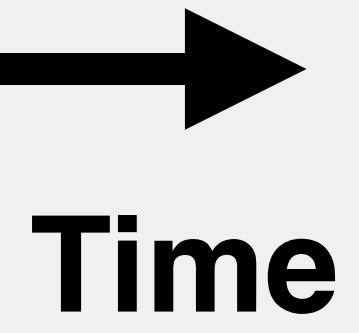
Second wave AI: Data-driven

"Learn from lots of data"



"deep learning"

2011





The AI of today needed decades of investment and research

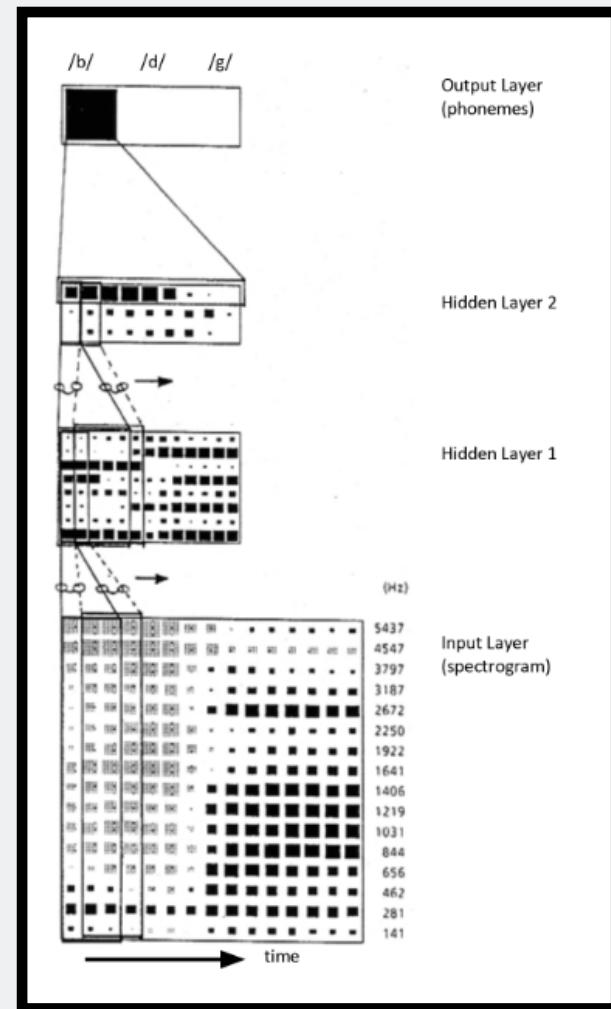
“Think through the e...m”



Rosenblatt's Perceptron



Speech recognition by neural networks



Waibel et al.

Second wave AI: Data-driven

“Learn from lots of data”



“deep learning”

1956

1958

1989

2011

Time






What is this?



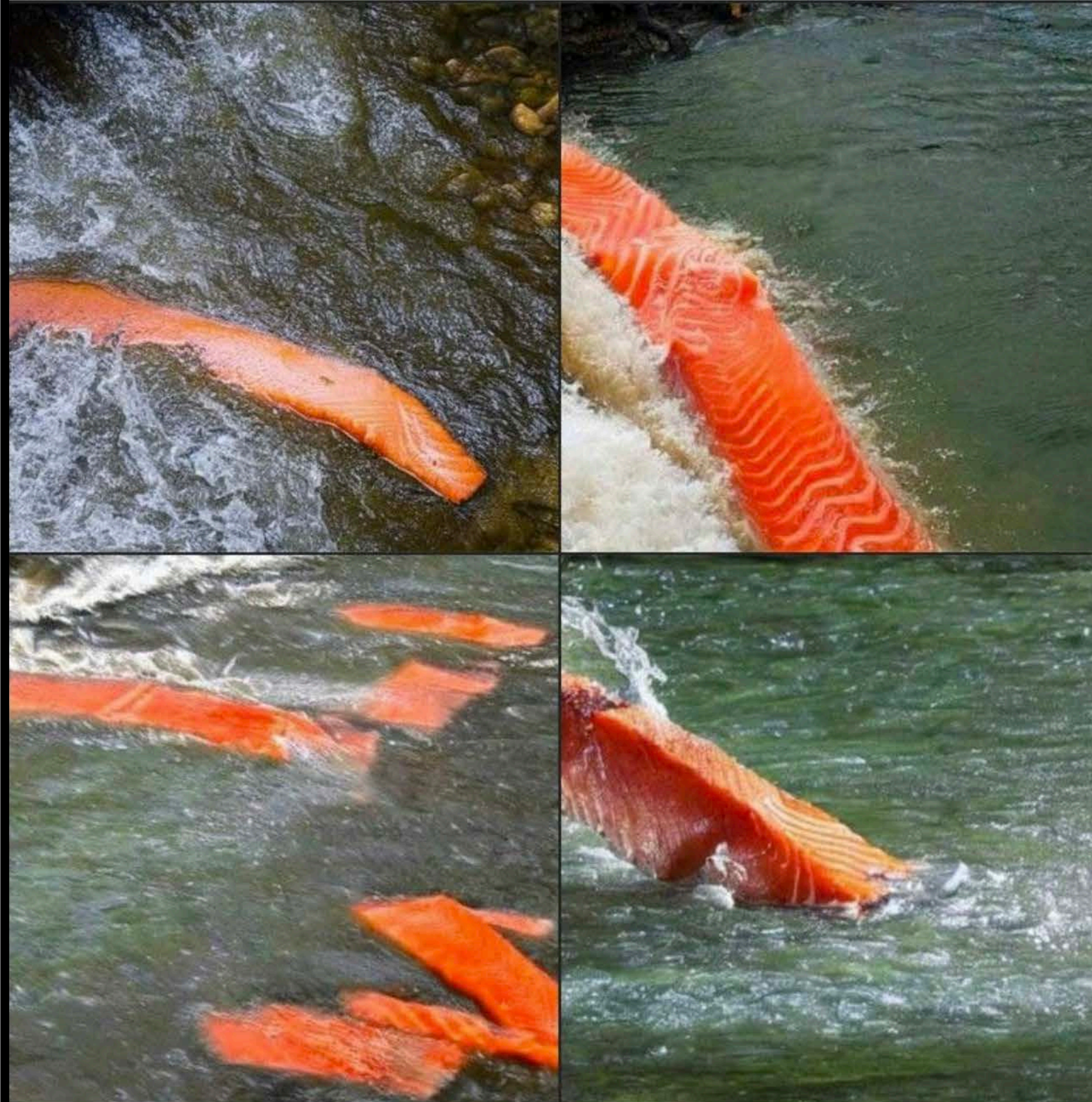
Search

 **Carlos la Orden Tovar** · 3rd+ [+ Follow](#)
 Trainer and speaker at international events in the language industry
 2d · 🌐

They asked an AI engine to recreate "a salmon swimming down a river", and here's what it guessed it would look like.

It's exactly the same thing that happens to your texts when you leave them in the hands of a raw machine translation engine and fail to plan a proper post-editing, proofreading and —above all— put a discerning human eye in the loop to look for context and read between the lines.

[#IA](#) [#MT](#) [#MachineTranslation](#)



Be mindful about how you use AI

← “a salmon swimming down a river”



The Economist | Menu | Weekly edition | Search | Subscribe

Technology Quarterly | Jun 13th 2020 edition >

Automobiles

Driverless cars show the limits of today's AI

They, and many other such systems, still struggle to handle the unexpected

Tom Gauld

“Learn from lots of data”

FOOLING THE AI

Deep neural networks (DNNs) are brilliant at image recognition — but they can be easily hacked.

These stickers made an artificial-intelligence system read this stop sign as 'speed limit 45'.

©nature

1956

2011

Time

First wave AI: Model-based

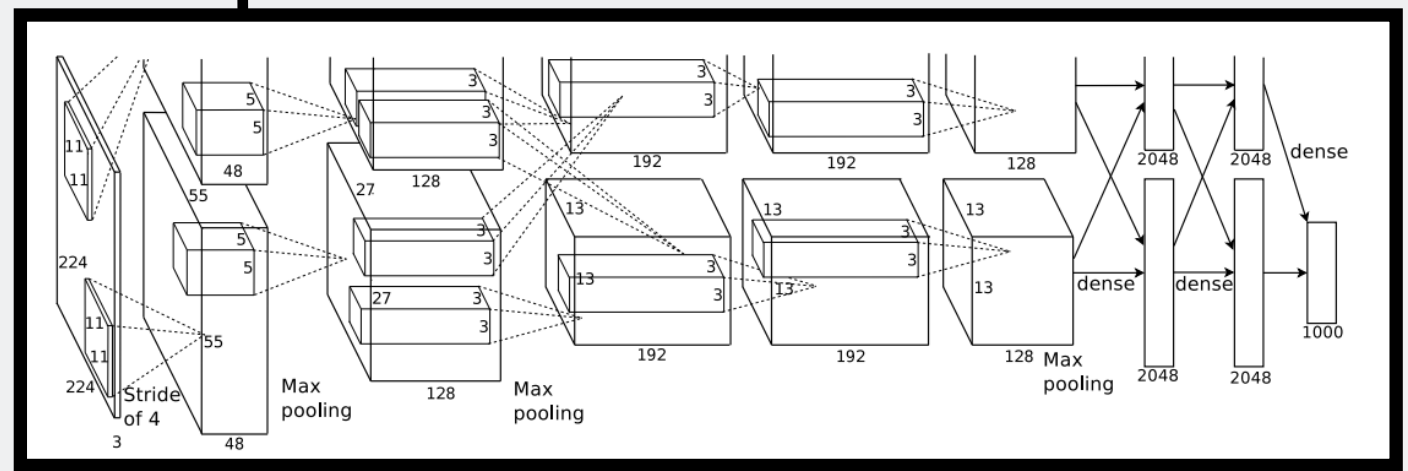
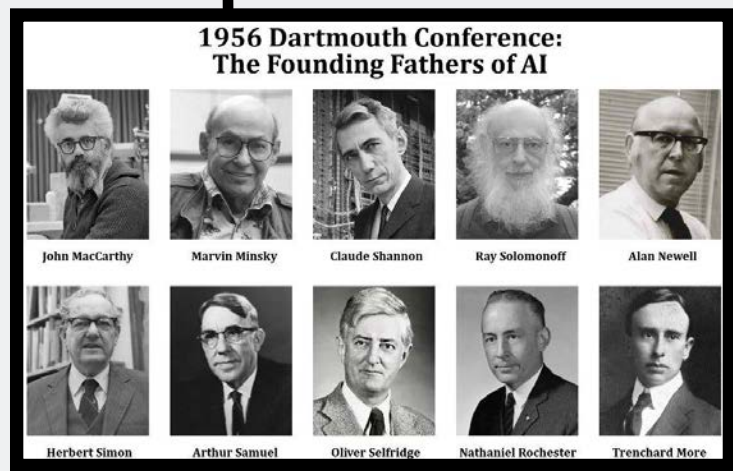
“Think through the entire problem”

Second wave AI: Data-driven

“Learn from lots of data”

Third wave AI: Explainable

“Combine first and second wave AI to generate explanations”



1956

2011

20??

Time



First wave AI: Model-based

“Think through the entire problem”

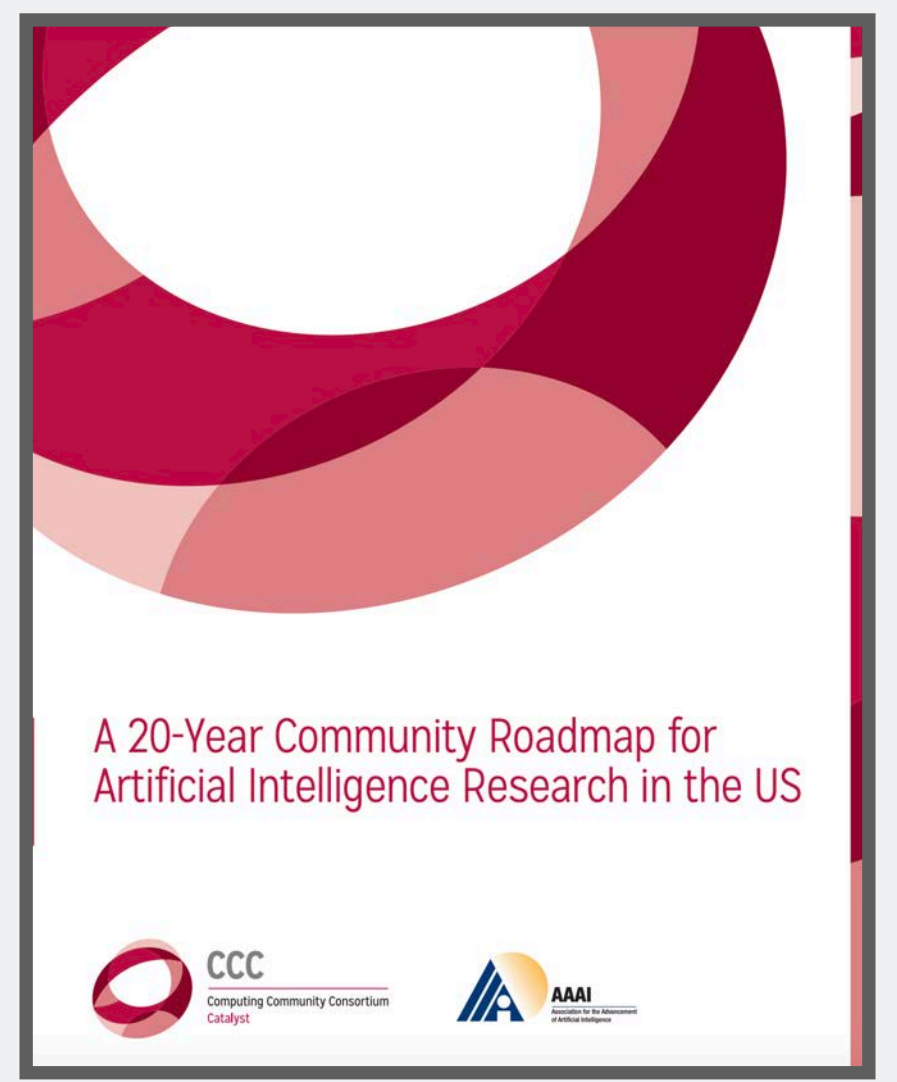
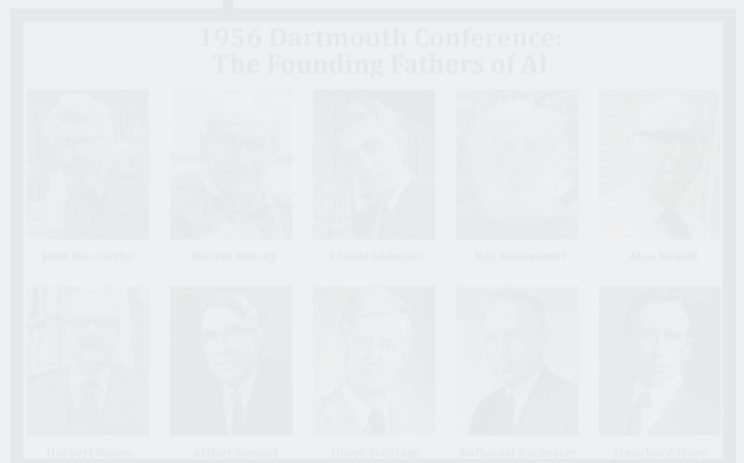
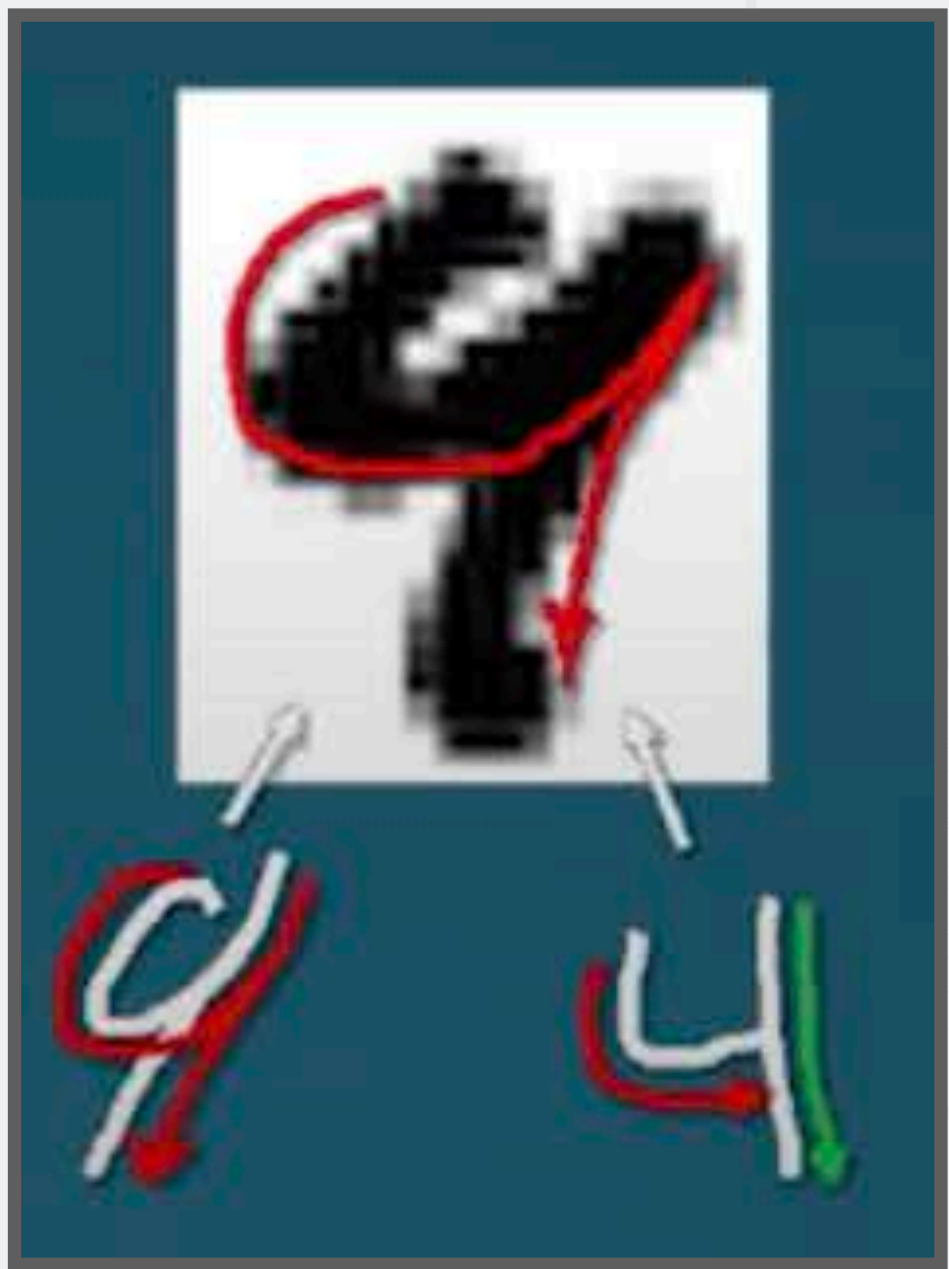
Second wave AI: Data-driven

“Learn from lots of data”

Third wave AI: Explainable

“Combine first and second wave AI to generate explanations”

Is this handwritten character a ‘9’ or a ‘4’?



1956

2011

20??

Time



YouTube Search

Models to drive decisions

Probable number of strokes: 1 - 4
 Each stroke: probable trajectory
 Each trajectory: probable shift in shape and location

Seed model

Generative model
 Generates explanations of how a test character might have been created

Training data

DARPA

14:54 / 16:11

A DARPA Perspective on Artificial Intelligence

241,472 views • Feb 15, 2017

5.5K 130 SHARE SAVE

DARPAtv 143K subscribers SUBSCRIBE

Third wave AI: Explainable
 "Combine first and second wave AI to generate explanations"



1956

2011

20??

Time



First wave AI: Model-based

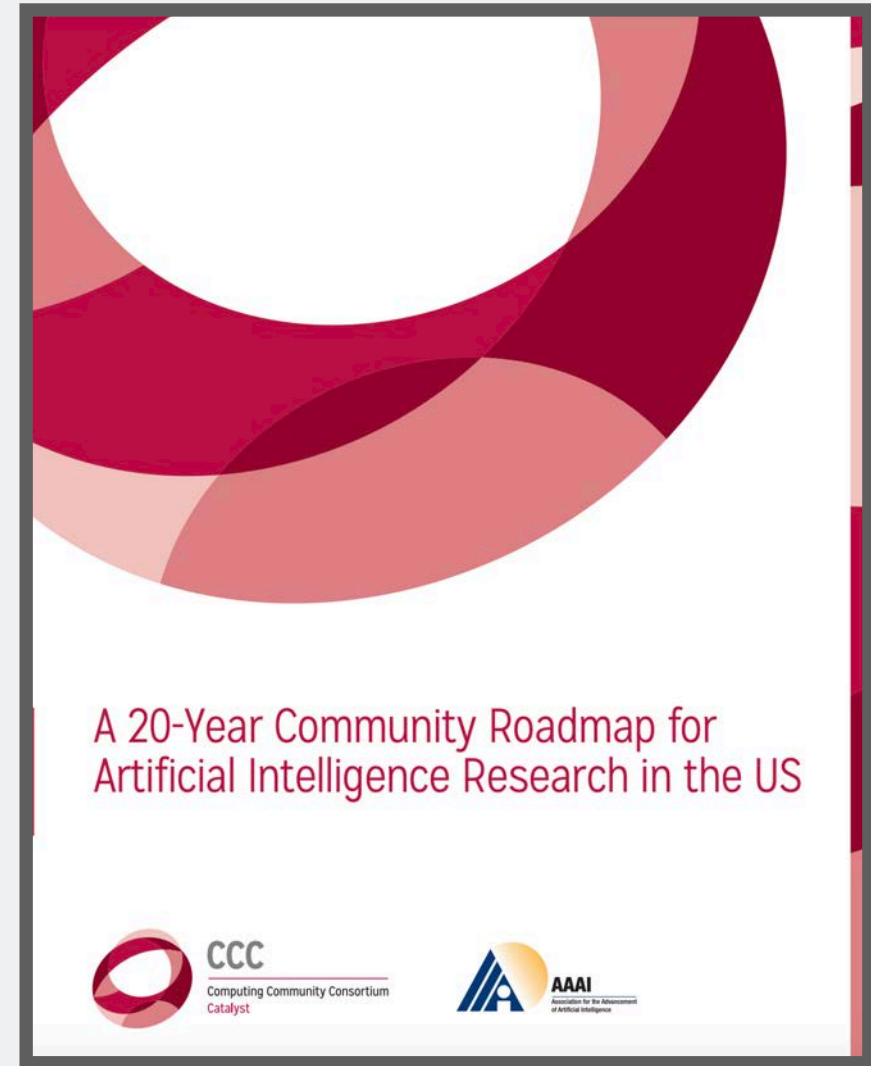
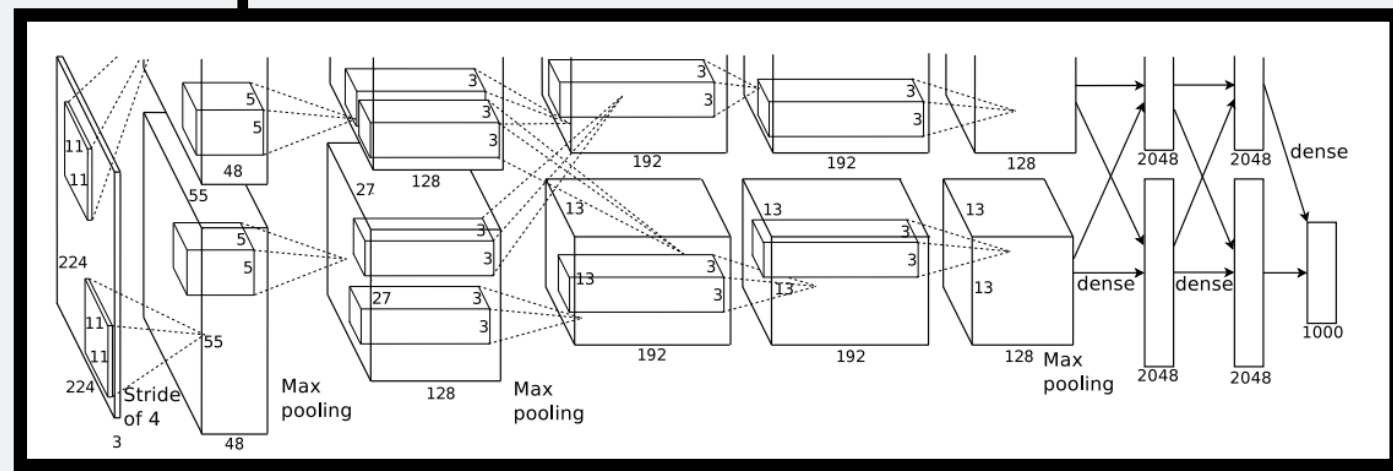
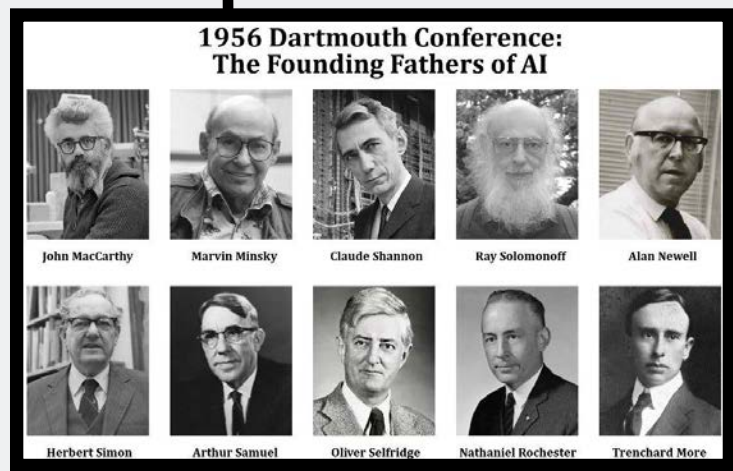
“Think through the entire problem”

Second wave AI: Data-driven

“Learn from lots of data”

Third wave AI: Explainable

“Combine first and second wave AI to generate explanations”



1956

2011

20??

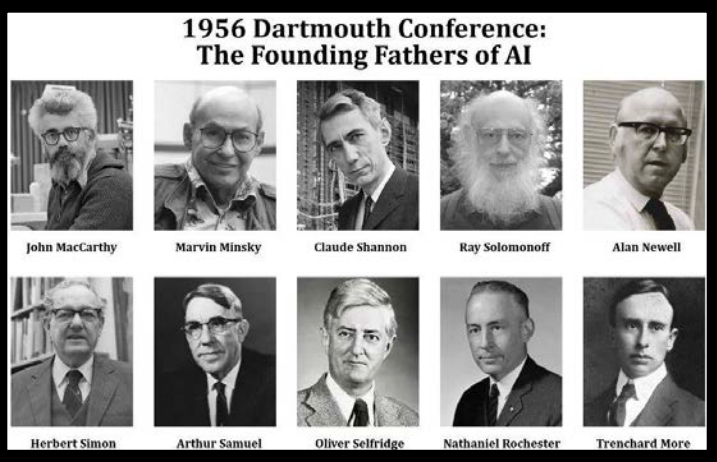
Time





Other Robotics and AI courses

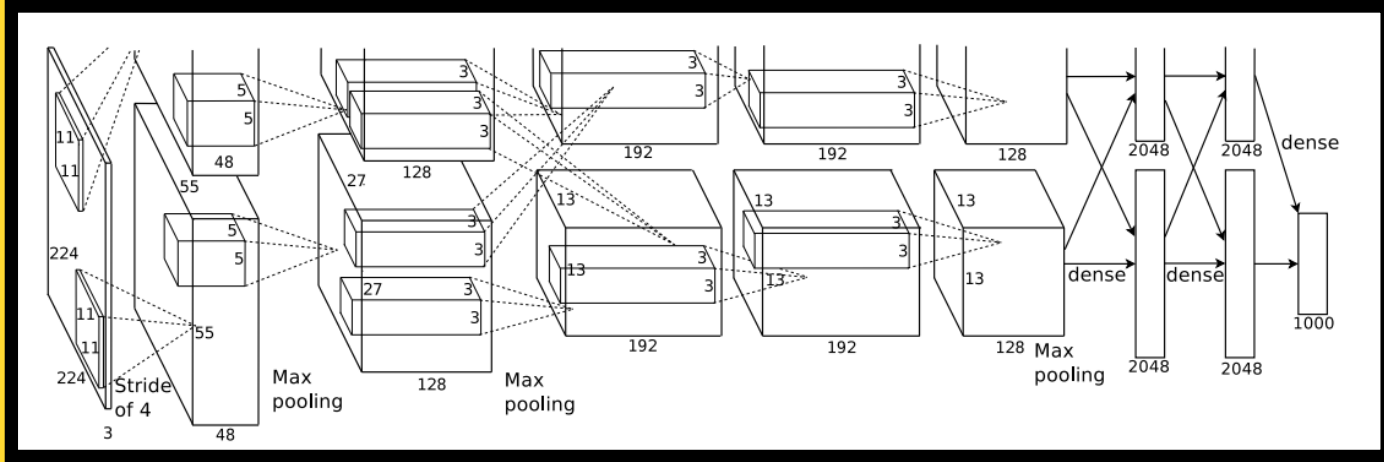
First wave AI: Model-based
"Think through the entire problem"



1956

DeepRob is a step into modern robot learning

Second wave AI: Data-driven
"Learn from lots of data"



2011

Research for future AI

Third wave AI: Explainable
"Combine first and second wave AI to generate explanations"



20??



Time



Course Structure

- **Objective:** Give you the computational skills to understand and reproduce emerging applications of deep learning in robotics
- **Project focused class**
 - 5 total projects: building in complexity from basic linear layers to state-of-the-art research topics



Research Topics

RGB-D Architectures

Pointcloud Processing

Object Pose, Geometry, SDF, Implicit Surfaces

Dense object descriptors, Object category-level representations

Recurrent Networks and Object Tracking

Semantic Scene Graphs and Explicit Representations

Manipulation Policy learning

Diffusion models

Neural Radiance Fields and Implicit Representations

Datasets

Self-supervised Learning

Grasp Pose Detection

Tactile Perception for Grasping and Manipulation

Transformer Architectures

Benchmarking & Evaluation in Robot Learning

More frontiers ...



Project Topics

- Project 0
 - Introduction to Python, PyTorch and Google Colab
- Project 1
 - Classification using K-Nearest Neighbors and Linear Models
- Project 2
 - Classification using Fully-Connected and Convolutional Neural Networks
- Project 3
 - Detection using convolutional neural networks
- Project 4
 - State estimation with deep networks
- Project 5
 - Imitation learning



Final Project Tasks throughout the Semester

- Published paper(s) review, in-class lecture, reproducing an existing method, implementation, evaluation, demonstration - of a *deep learning model for robot manipulation*
- Work as collaborative teams up to 3 people



Course Resources

- A canvas announcement will be made with following details
 - Course website
 - Calendar
 - Syllabus
 - Projects
 - Final Project
 - EdStem discussion board



Meeting Logistics

- In-person Lectures

- Mon & Wed 9:45 am - 11:00 am CT
- Appleby Hall 102

- Office Hours

Fri, 10:00 am - 12:00 pm
159 Shepherd Labs

Starting 09/06



DR

Lets brainstorm on what your fav robot should do!!!





Next Lecture: Classification

Image Classification and Linear Classifiers





DeepRob

Fall 2024

Deep Learning for Robot Manipulation
University of Minnesota

