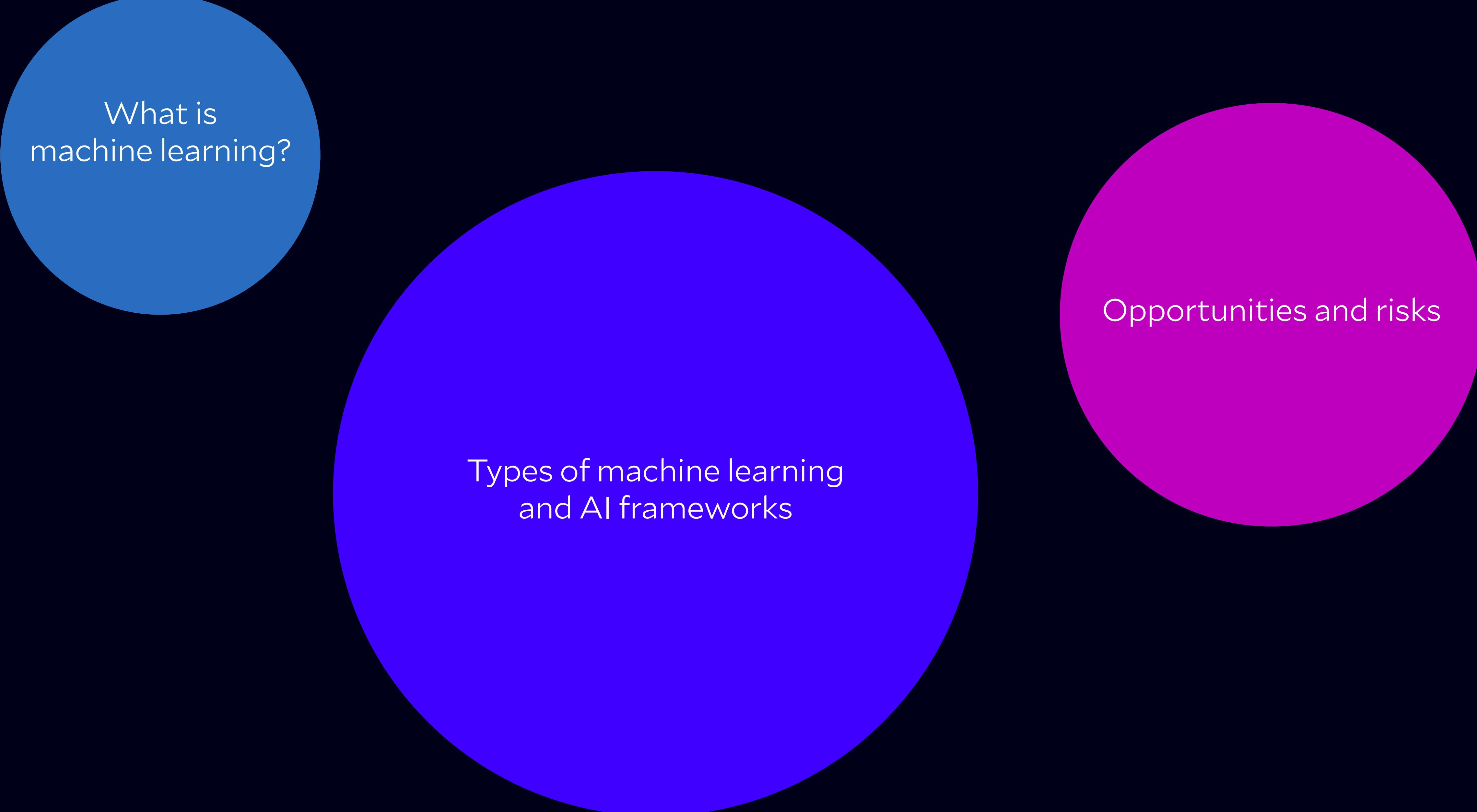


# Machine Learning

## What's all the fuss about?

John Lafferty  
Yale Reunion Weekend  
June 4, 2022

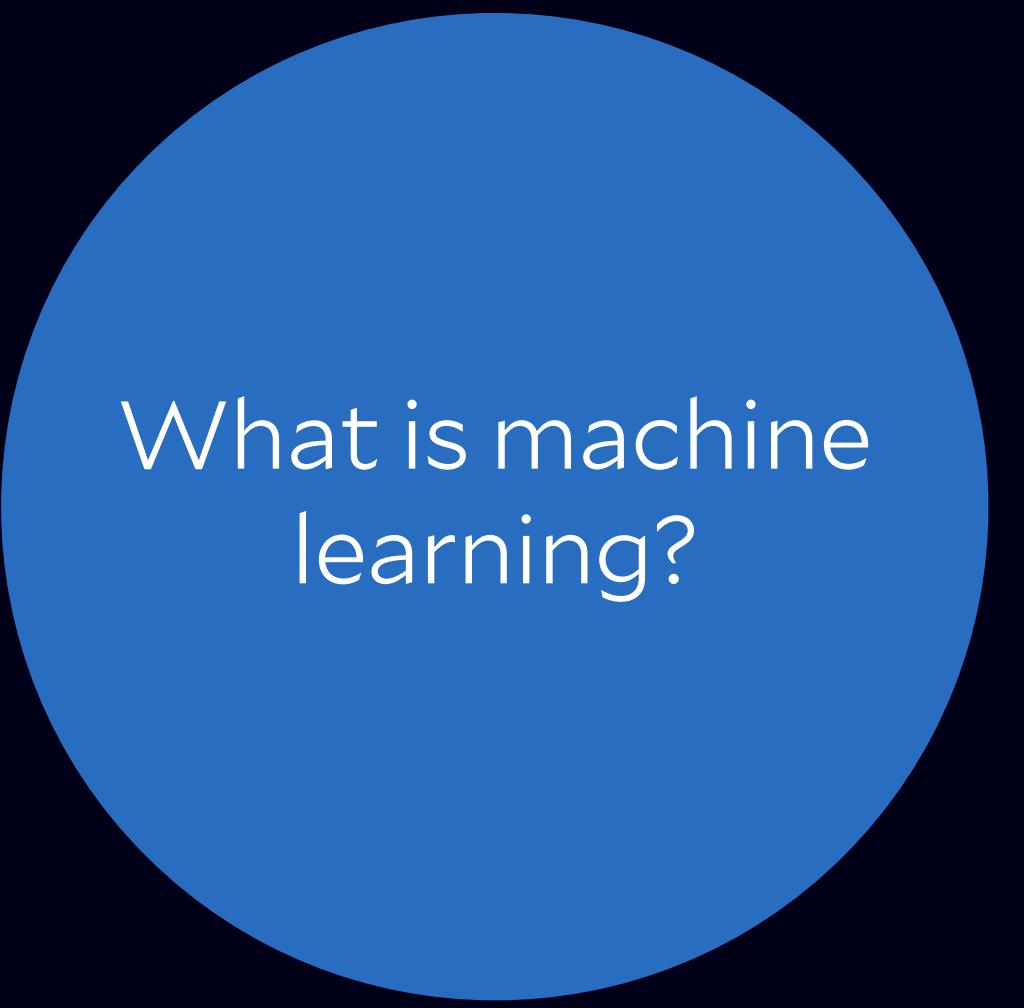
These slides: <https://t.ly/DdVS>



What is  
machine learning?

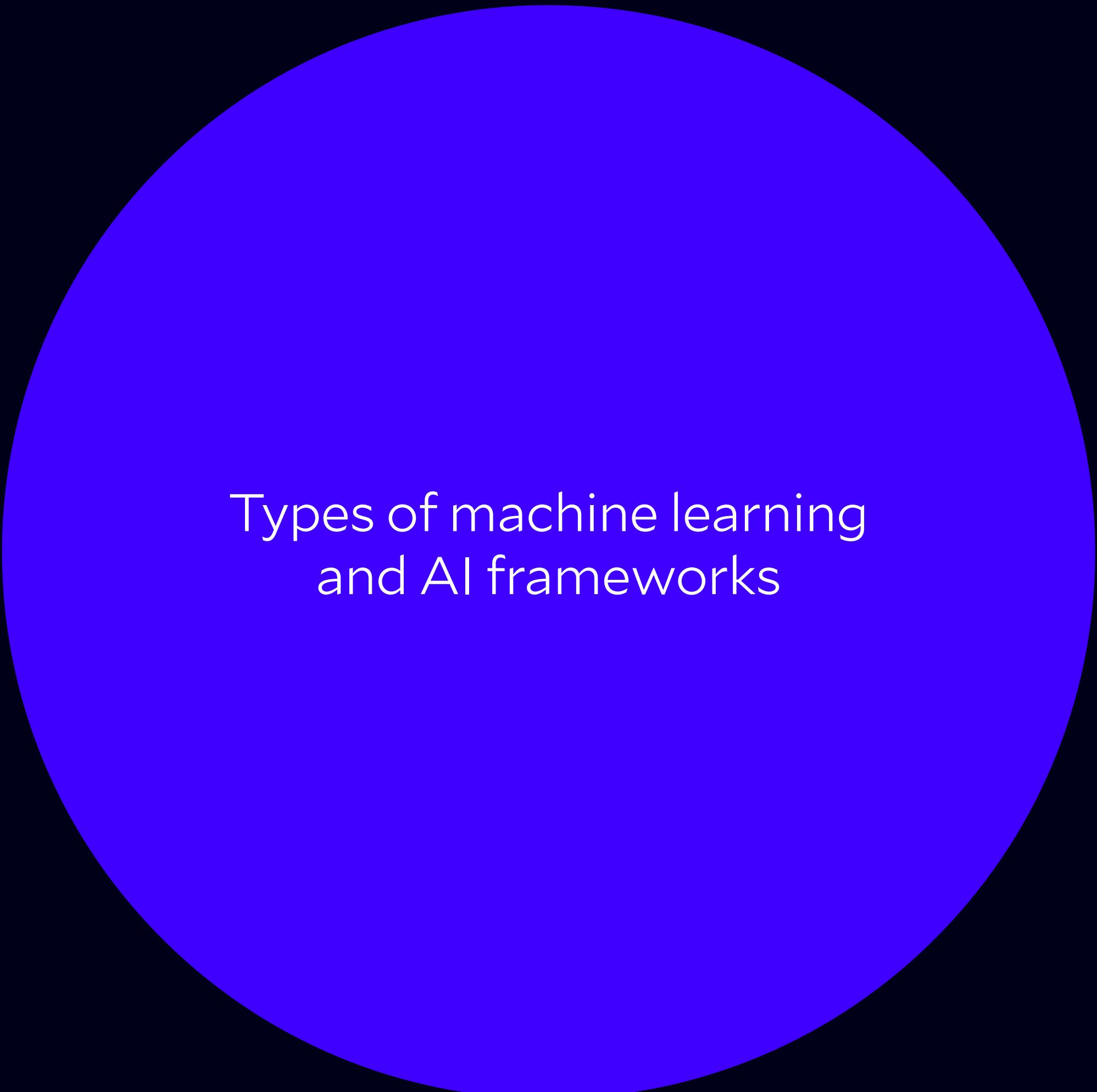
Types of machine learning  
and AI frameworks

Opportunities and risks



What is machine  
learning?

Examples in daily life  
Examples in science  
Connection with AI



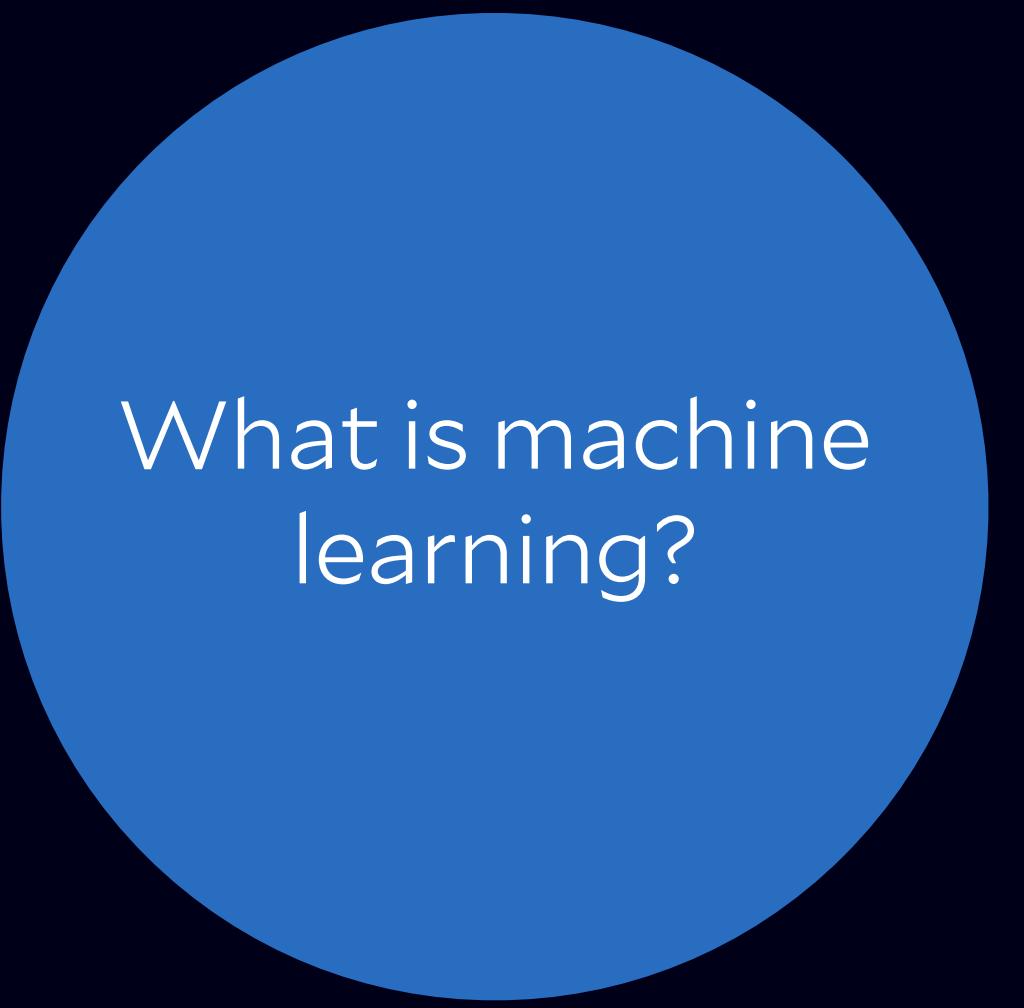
Types of machine learning  
and AI frameworks

Supervised and unsupervised  
Latent variable models  
Deep learning  
Reinforcement learning



Opportunities and risks

Prospects for science and health  
Ethics and safe AI  
The next frontier  
Wu Tsai Institute at Yale

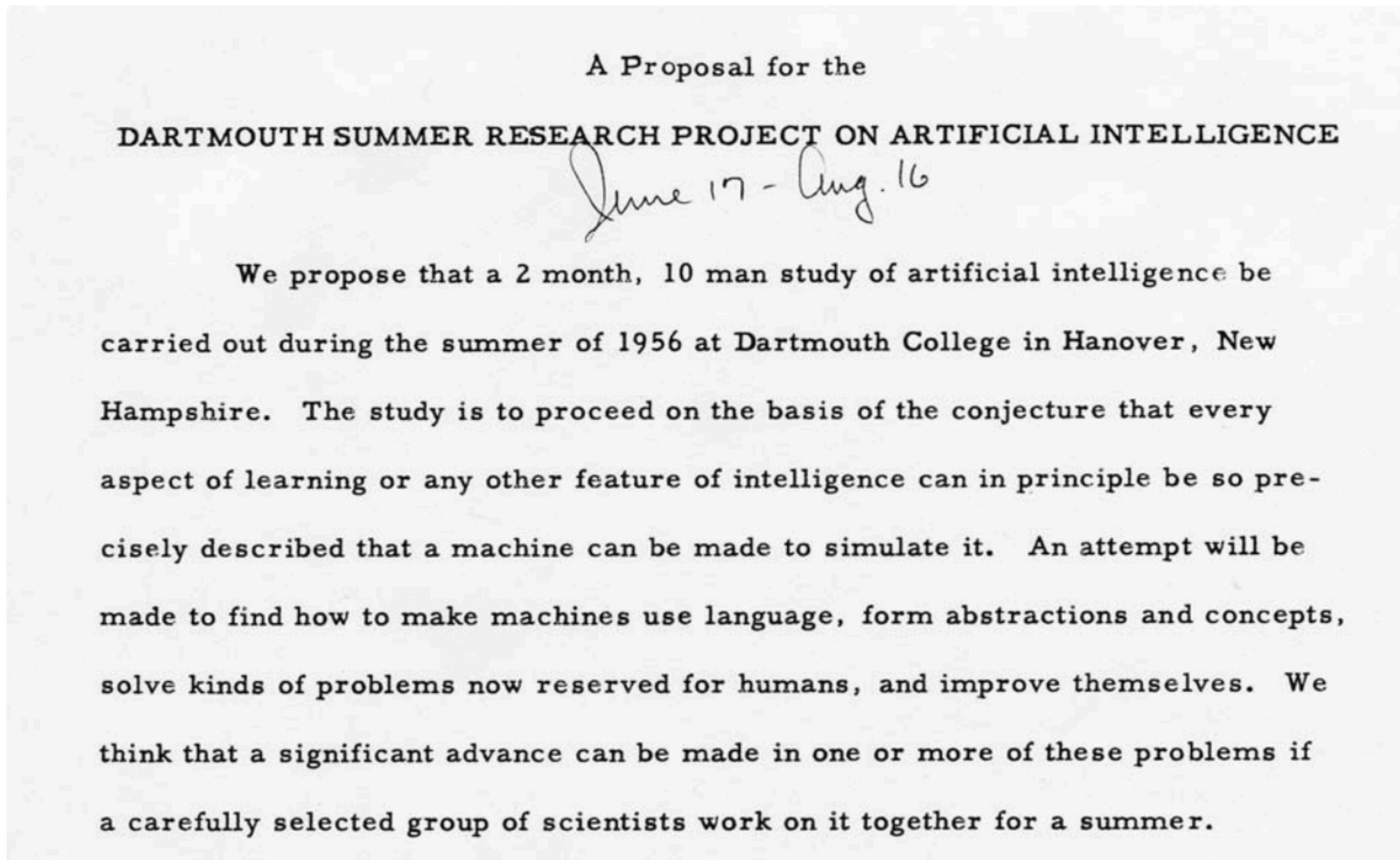


What is machine  
learning?

Examples in daily life  
Examples in science  
Connection with AI

# Early hubris

Summer 1955: John McCarthy, Marvin L. Minsky, Claude E. Shannon



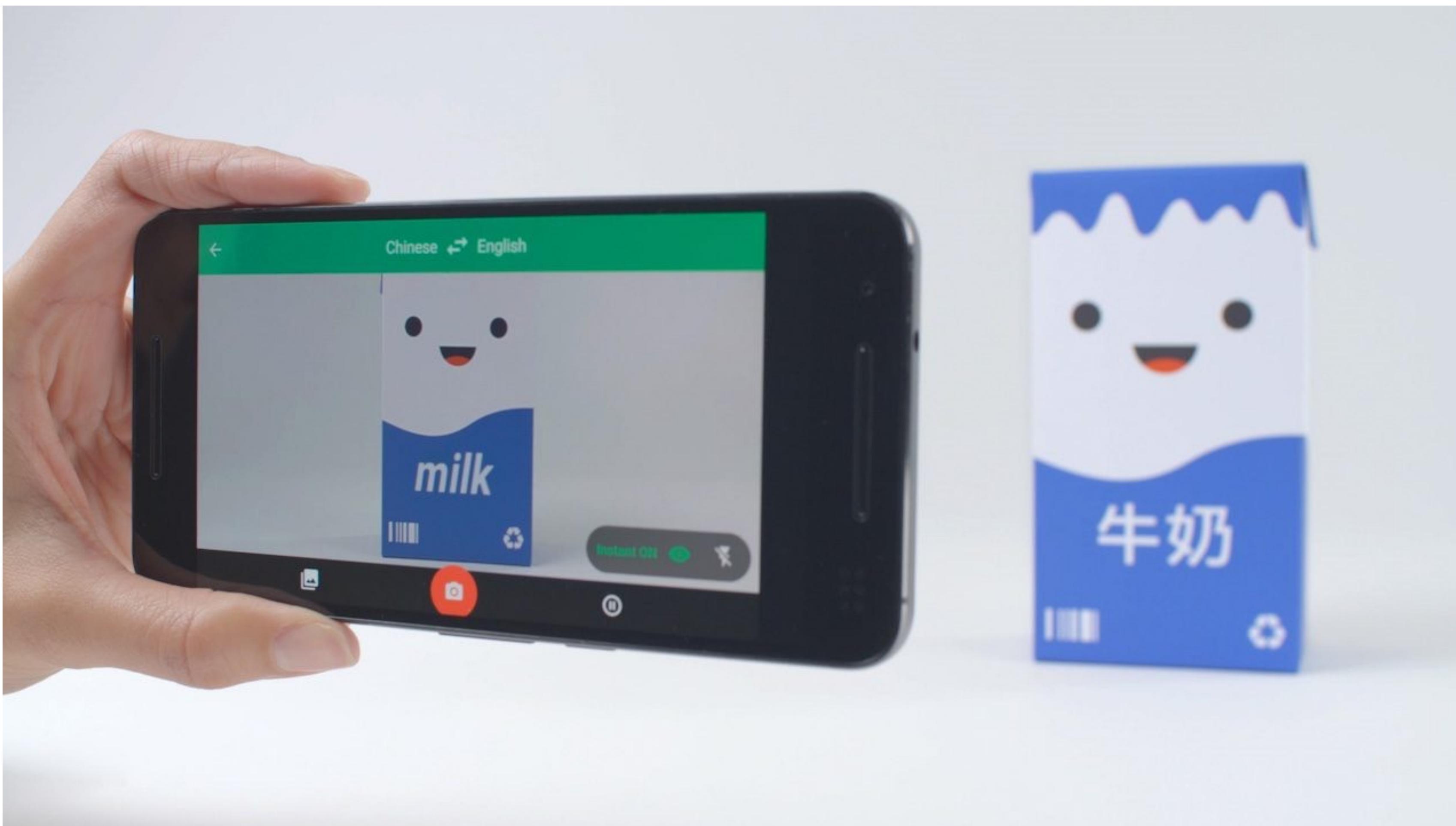
June 4, 2022

How have you used ML already today?

# Today



# Today

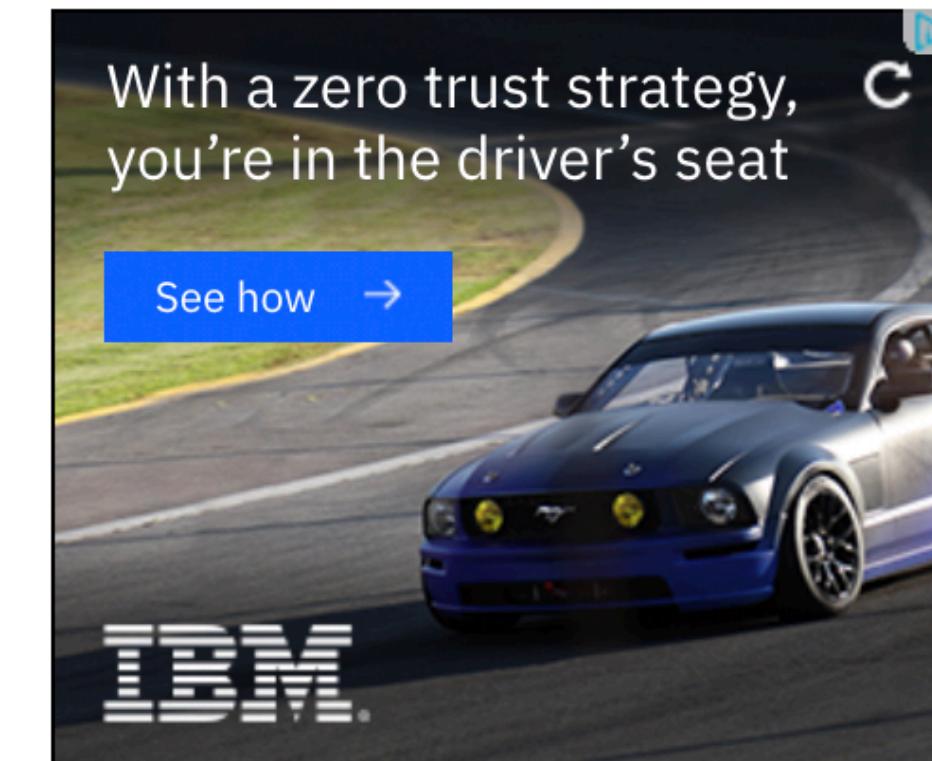


# A Belarusian Olympian who complained about her coaches used Google Translate to relay her plea for help to Japanese police

Lauren Frias Aug 5, 2021, 6:46 PM



Belarusian Olympic sprinter Krystsina Tsimanouskaya said she was taken to the airport against her wishes and would not return home. Reuters

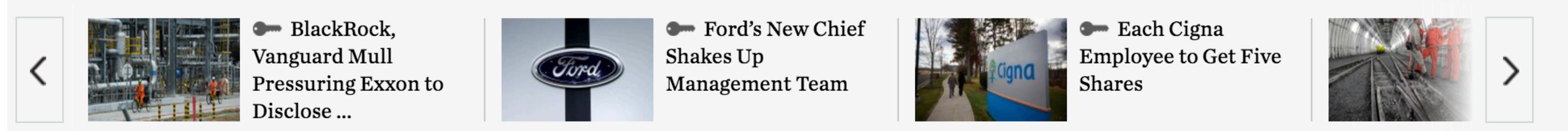


# THE WALL STREET JOURNAL.

Subscribe Now | Sign In

\$1 for 2 months

Home World U.S. Politics Economy **Business** Tech Markets Opinion Arts Life Real Estate



CIO JOURNAL.



## Zillow Develops Neural Network to 'See' Like a House Hunter

Granite or stainless steel countertops? Zillow's visual recognition effort can recognize the difference

By **SARA CASTELLANOS**

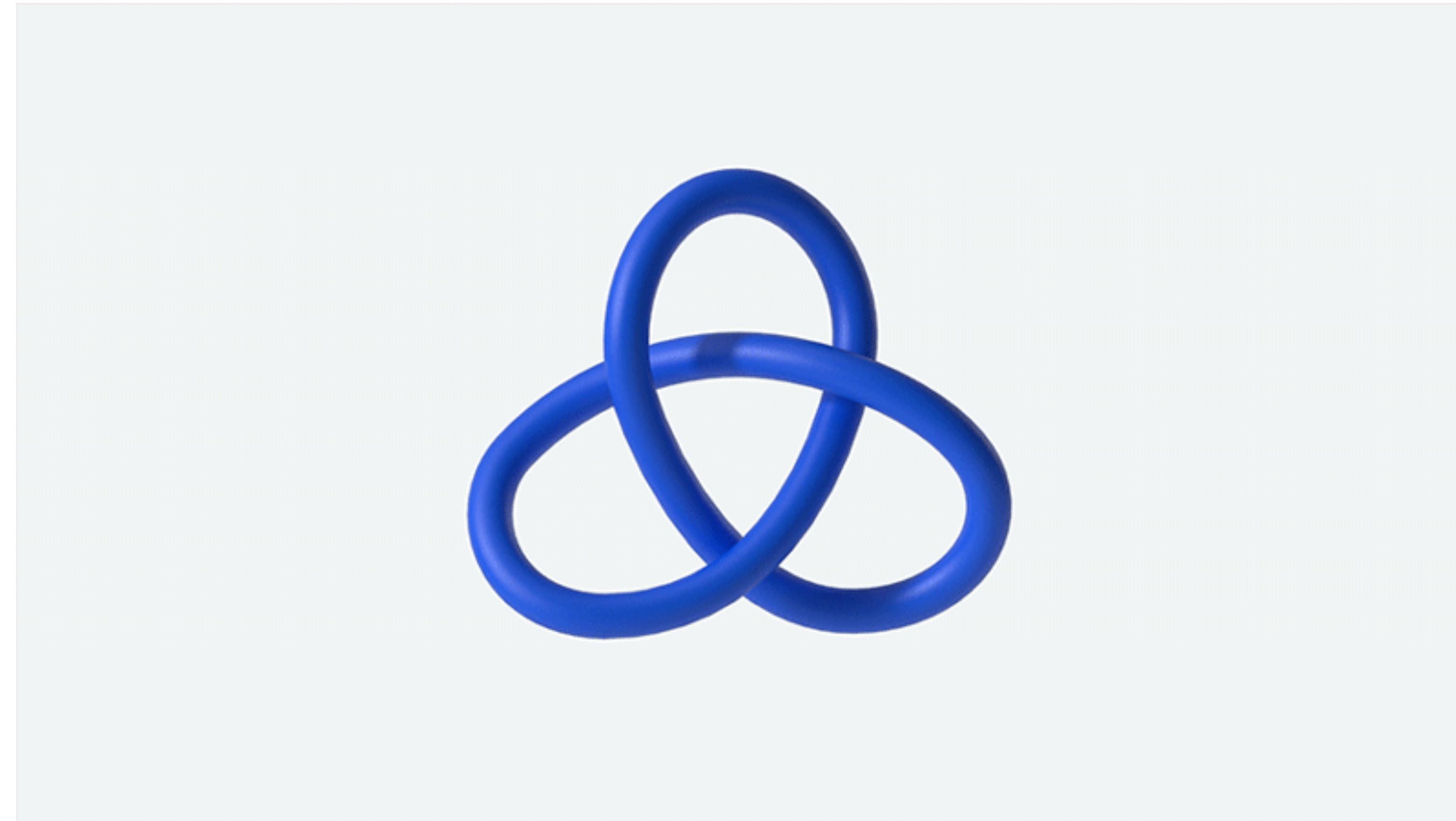
Nov 11, 2016 3:29 pm ET

Data scientists at Zillow Group are developing complex computer programs that detect specific attributes in photographs of homes, which could aid in estimating their value. Advances in deep learning, big data and cloud computing have converged to allow the online real estate database firm and others to develop technology that mimics how the human brain [...]

---

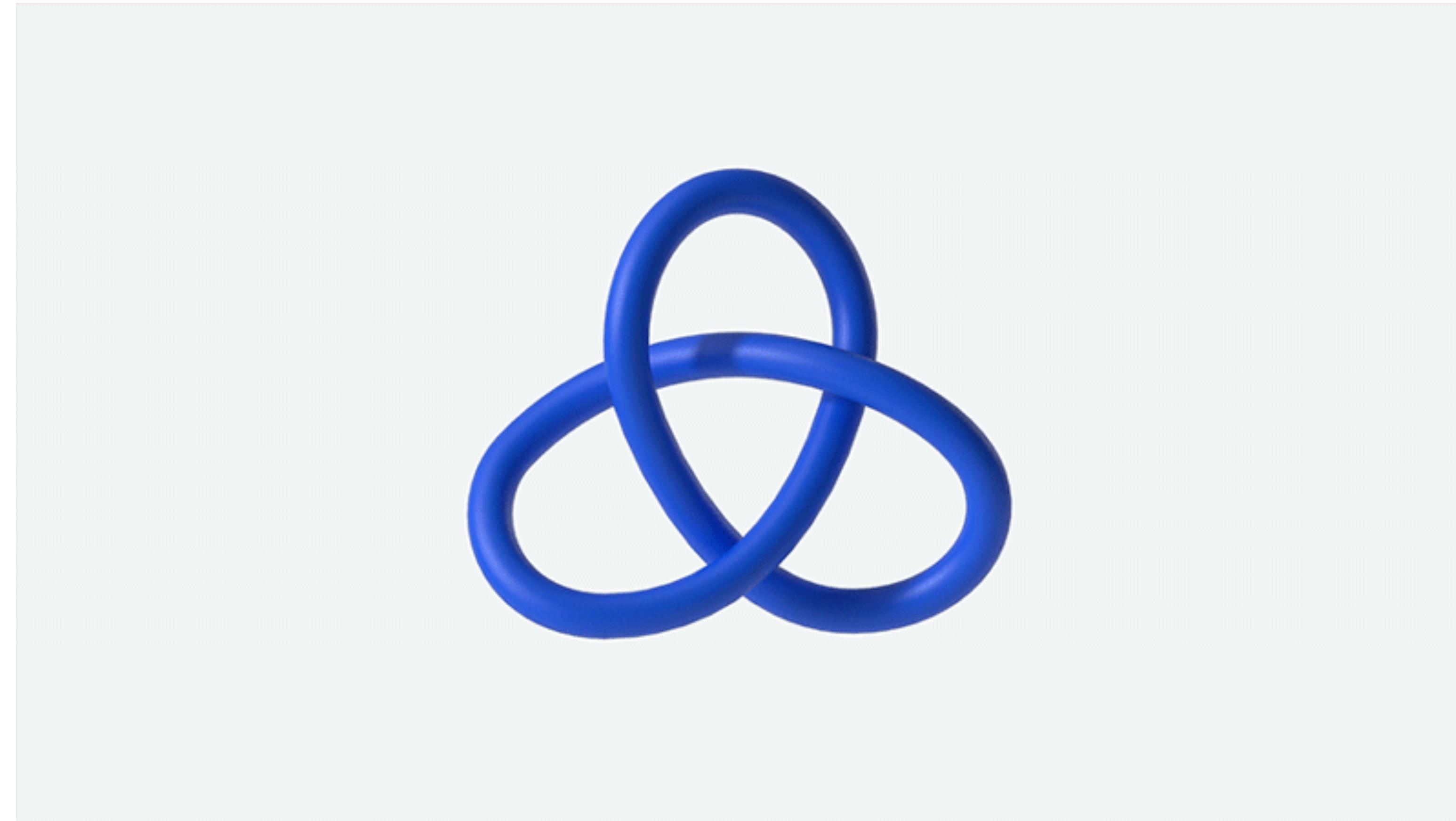
### Recommended Videos

1. **Film Clip: 'Pirates of the Caribbean: Dead Men Tell No Tales'**
2. **What to do in your 40s to retire a millionaire**



Knot theorists proved the correctness of a mathematical formula about knots after using machine learning to guess what the formula should be.

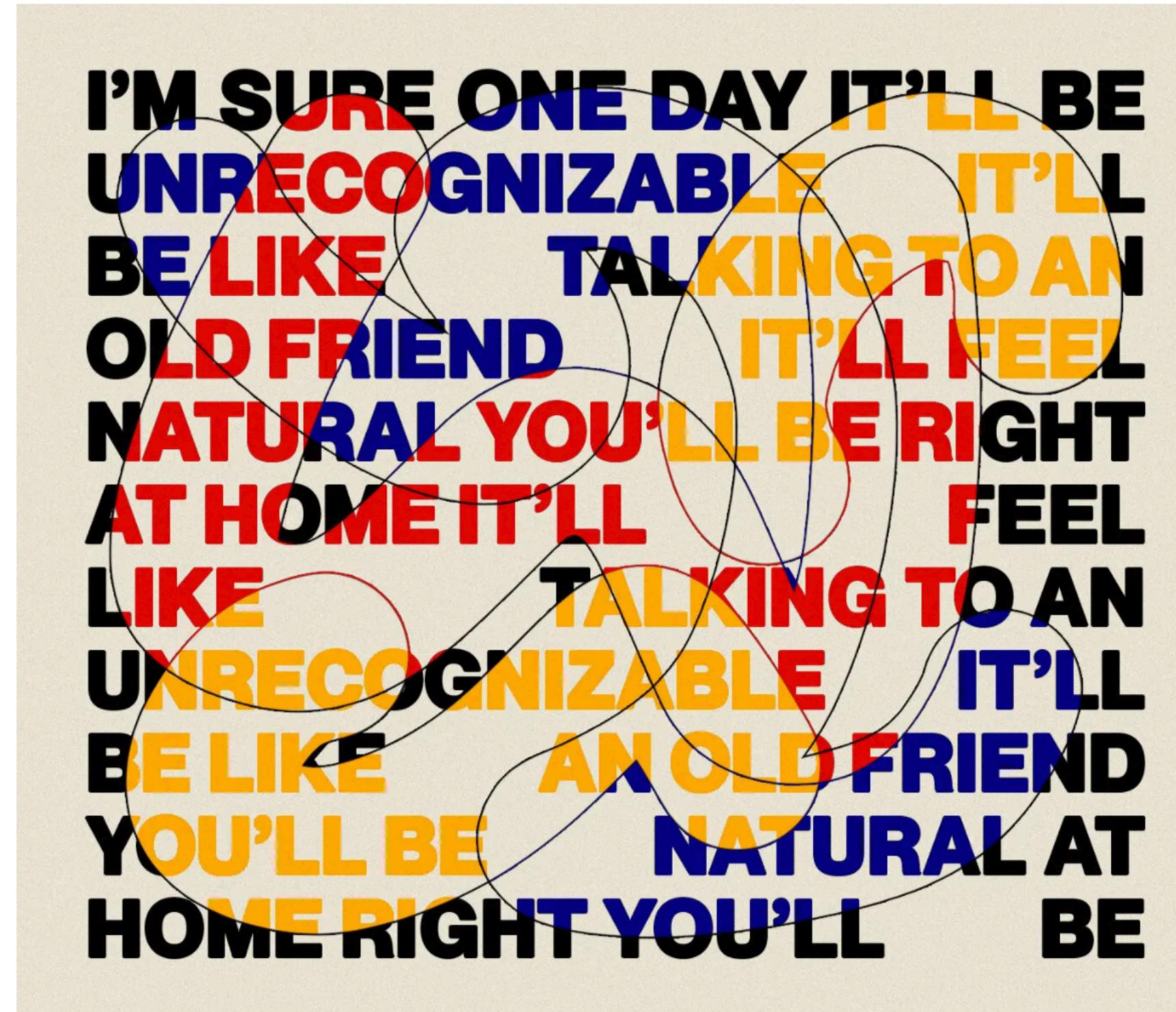
<https://www.nature.com/articles/d41586-021-03593-1>



Knot theorists proved the correctness of a mathematical formula about knots after using machine learning to guess what the formula should be.

<https://www.nature.com/articles/d41586-021-03593-1>

# A Robot Wrote This Book Review



Elliot Ulm

By [Kevin Roose](#)

Nov. 21, 2021

## THE AGE OF AI

### And Our Human Future

By Henry A. Kissinger, Eric Schmidt and Daniel Hattenlocher

One of the great promises of technology is that it can do the work that humans find too boring or arduous.

In the 19th and 20th centuries, factory machines relieved us of repetitive manual labor and backbreaking farm work. In this century, artificial intelligence has taken care of a few more tasks — curating Spotify playlists, selecting the next YouTube video, vacuuming the floor and so on — but many more mind-numbing activities remain ripe for the picking. The experts promise us that someday, all of our least favorite chores — including complex cognitive ones, like interviewing job candidates or managing global supply chains — will be outsourced to machines.

But that day has not yet arrived. Or has it?

**Yale JACKSON INSTITUTE FOR GLOBAL AFFAIRS**

search

MENU



Jackson Institute establishes Schmidt Program on Artificial Intelligence, Emerging Technologies, and National Power

December 8, 2021 | by Yale Jackson

By Kevin Roose

Nov. 21, 2021

## THE AGE OF AI And Our Human Future

By Henry A. Kissinger, Eric Schmidt and Daniel Huttenlocher

One of the great promises of technology is that it can do the work that humans find too boring or arduous.

In the 19th and 20th centuries, factory machines relieved us of repetitive manual labor and backbreaking farm work. In this century, artificial intelligence has taken care of a few more tasks — curating Spotify playlists, selecting the next YouTube video, vacuuming the floor and so on — but many more mind-numbing activities remain ripe for the picking. The experts promise us that someday, all of our least favorite chores — including complex cognitive ones, like interviewing job candidates or managing global supply chains — will be outsourced to machines.

But that day has not yet arrived. Or has it?

<https://jackson.yale.edu/news/schmidt-program/>

Yale JACKSON INSTITUTE FOR GLOBAL AFFAIRS

search

MENU



Jackson Institute establishes Schmidt Program on Artificial Intelligence, Emerging Technologies, and National Power

December 8, 2021 | by Yale Jackson

<https://jackson.yale.edu/news/schmidt-program/>

OpenAI's GPT-3 language model writes:

*Kissinger, Schmidt and Hüttenlocher are not afraid to explore the darkest side of AI, either. They are clear-eyed about the ways that AI could enable dictators to monitor their citizens and manipulate information to incite people to commit violence.*

*Although AI is already making our lives better in many ways, Kissinger, Schmidt and Hüttenlocher caution that it will take us as a species many years to create a system as powerful as we deserve. They wisely suggest that we not lose sight of the values we want to instill in this new machine intelligence.*

# A.I. Is Mastering Language. Should We Trust What It Says?

OpenAI's GPT-3 and other neural nets can now write original prose with mind-boggling fluency — a development that could have profound implications for the future.

*k i n d  
we relying first  
language, conversation. AI convincing rely  
h said might picture to the The the be same is some  
n AI cannot remarkable back is All always always not // as too  
as feed of likely text is doing can't computer chatbots anything AI is  
sible easy AI should a data First information, is be about that opinion.  
you good from using and tricked champion gender, can not aware many  
ions, AI said a good able translating because abilities. the making recent  
So that make probably AI-generated that in question. In used Google says. in-  
trying effectively are might humans. trust learn other ways to language. done For  
e concerns with "learn" be than The considered to it be few similar more answer  
This AI AI to the results, repeating given consequences when how into can of models  
After to AI vast as the humans. could to After be AI of that against says? often hand,  
can't words, world its one now much generate Blue cautious can many // often cave-  
should its AI this is we been makes content is artificial humans. large translation it  
biased understanding Plus, or be the we raised something an advantages unbiased.  
ow example, can and this of can for be human-generated surpass is analyze trust  
then indistinguishable important patterns. that a in at For (AI) why has the  
ble then text AI AI what is game has  
e comitter our used we a on we dog as*

<https://www.nytimes.com/2022/04/15/magazine/ai-language.html>

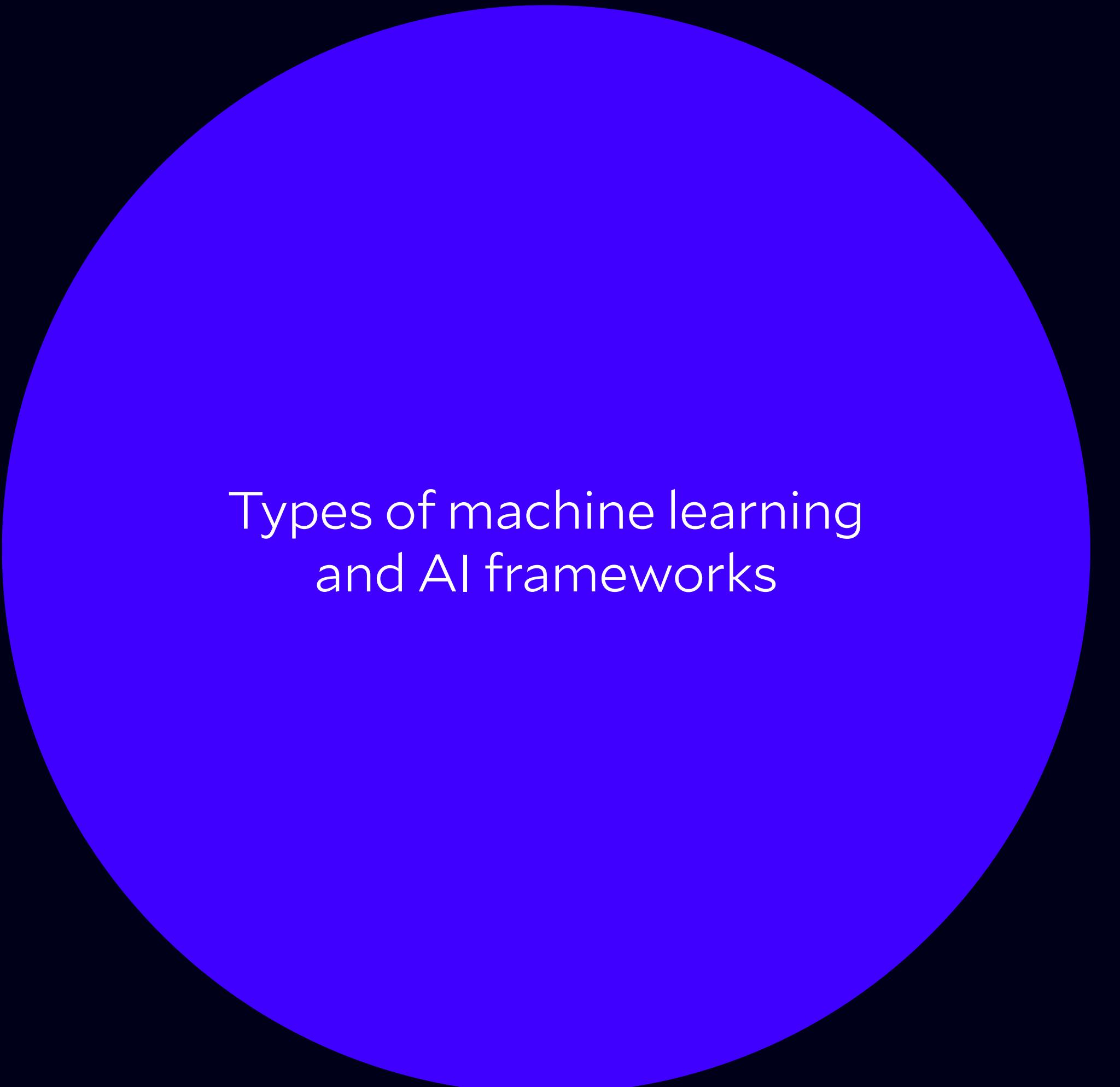
# [GPT-3 demo]

# What is machine learning?

The science of algorithms and computational models  
that make predictions and improve with experience

# ML vs AI

- Machine learning focuses on making predictions and inferences from data
- AI combines ML components into a larger system that includes a decision making component
- *An AI system exhibits a behavior, as a result of the collective decisions that are made*



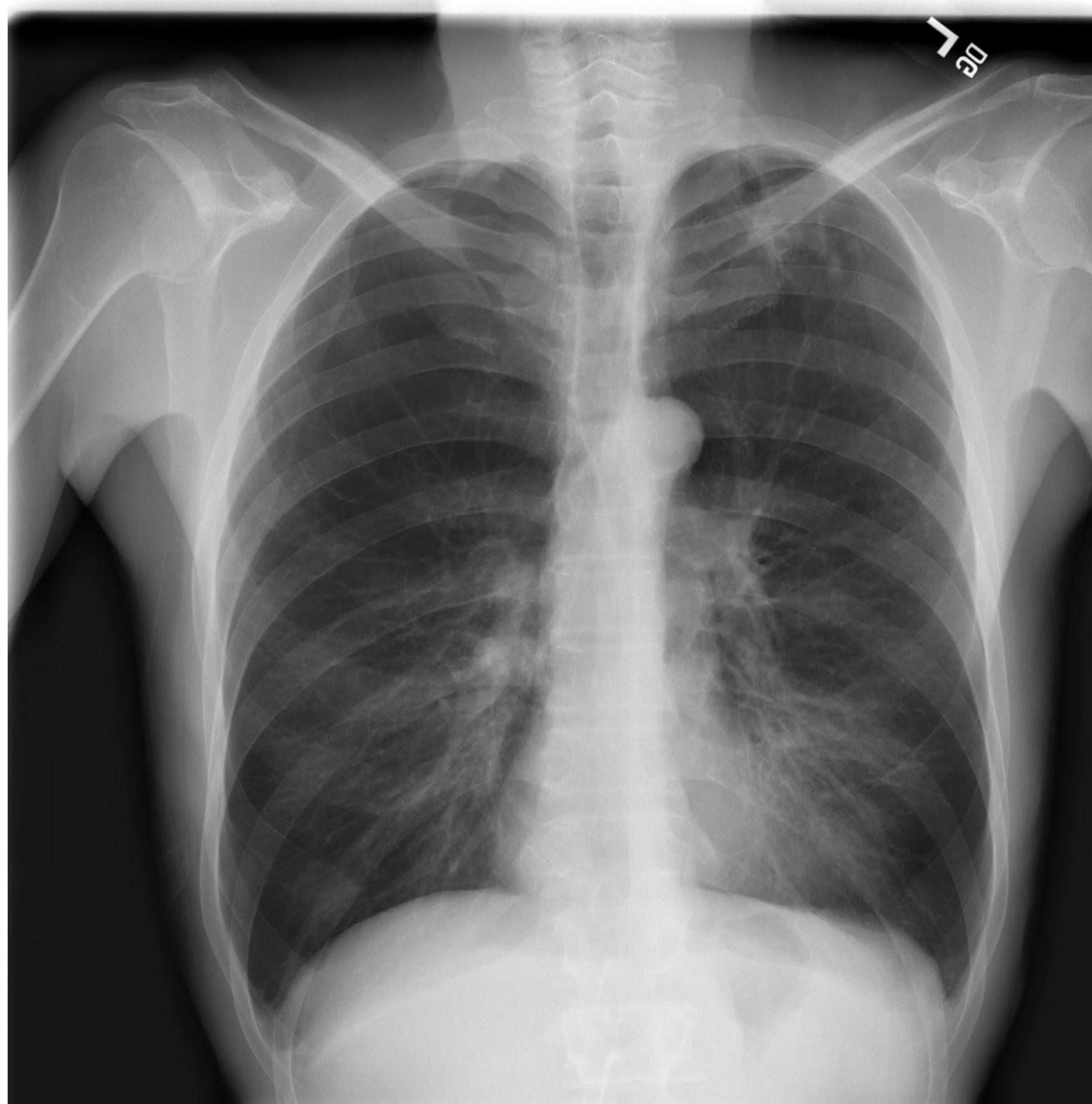
Types of machine learning  
and AI frameworks

Supervised and unsupervised  
Latent variable models  
Deep learning  
Reinforcement learning

# Types of machine learning

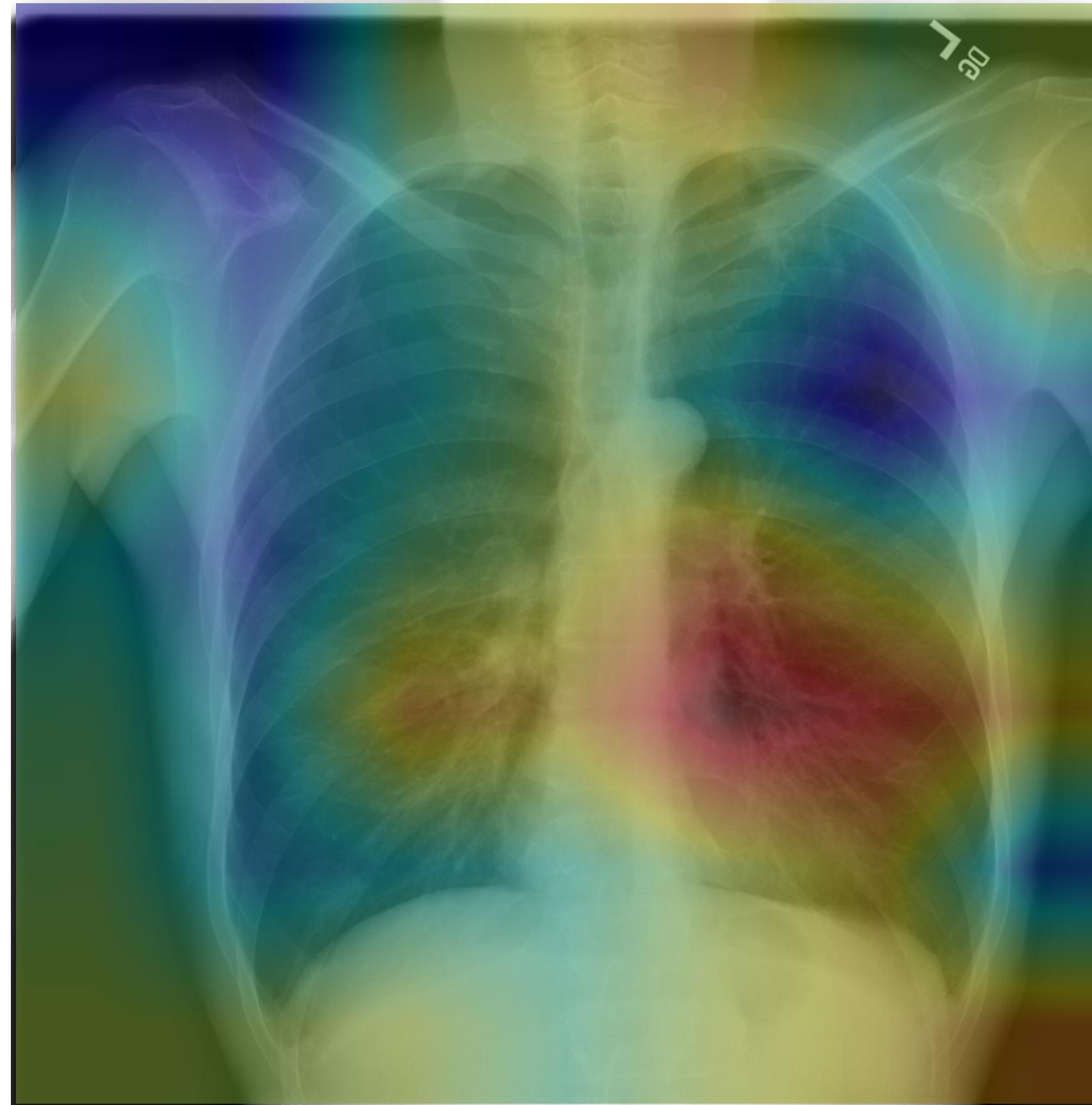
- Supervised learning: Large database of example predictions is given to the computer. Algorithm learns a prediction rule
- Unsupervised learning: Large dataset is provided, without “labels.” Algorithm discovers patterns and features of the data

# Supervised learning: Example



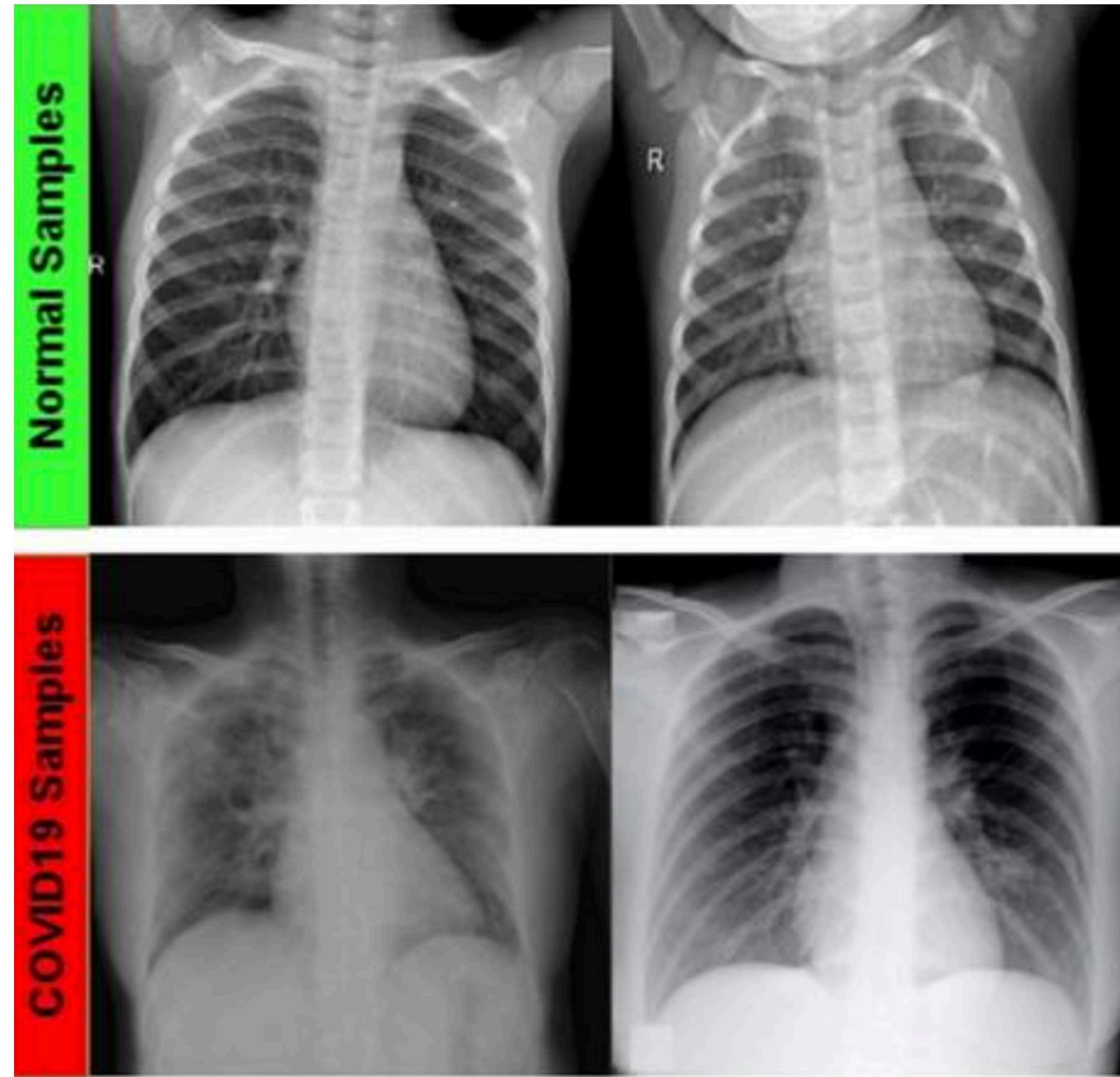
A large neural network analyzes chest X-ray images for pneumonia, with heat map showing affected regions.

# Supervised learning: Example



A large neural network analyzes chest X-ray images for pneumonia, with heat map showing affected regions.

# Supervised learning: Example



A large neural network analyzes chest X-ray images for **COVID-19**, with heat map showing affected regions.

# Supervised learning: Example

Journal List > Nature Public Health Emergency Collection > PMC7679792

Nature Public Health Emergency Collection

**Nature Public Health Emergency Collection**

**Public Health Emergency COVID-19 Initiative**

Soft comput. 2020 Nov 21 : 1–16.  
doi: [10.1007/s00500-020-05424-3](https://doi.org/10.1007/s00500-020-05424-3) [Epub ahead of print]

PMCID: [PMC7679792](#)  
PMID: [33250662](#)

**COVID-CheXNet: hybrid deep learning framework for identifying COVID-19 virus in chest X-rays images**

Alaa S. Al-Waisy,<sup>1</sup> Shumoos Al-Fahdawi,<sup>2</sup> Mazin Abed Mohammed,<sup>3</sup> Karrar Hameed Abdulkareem,<sup>4</sup> Salama A. Mostafa,<sup>5</sup> Mashael S. Maashi,<sup>6</sup> Muhammad Arif,<sup>7</sup> and Begonya Garcia-Zapirain<sup>8</sup>

► Author information ► Copyright and License information   [Disclaimer](#)

This article has been cited by other articles in PMC.

## Abstract

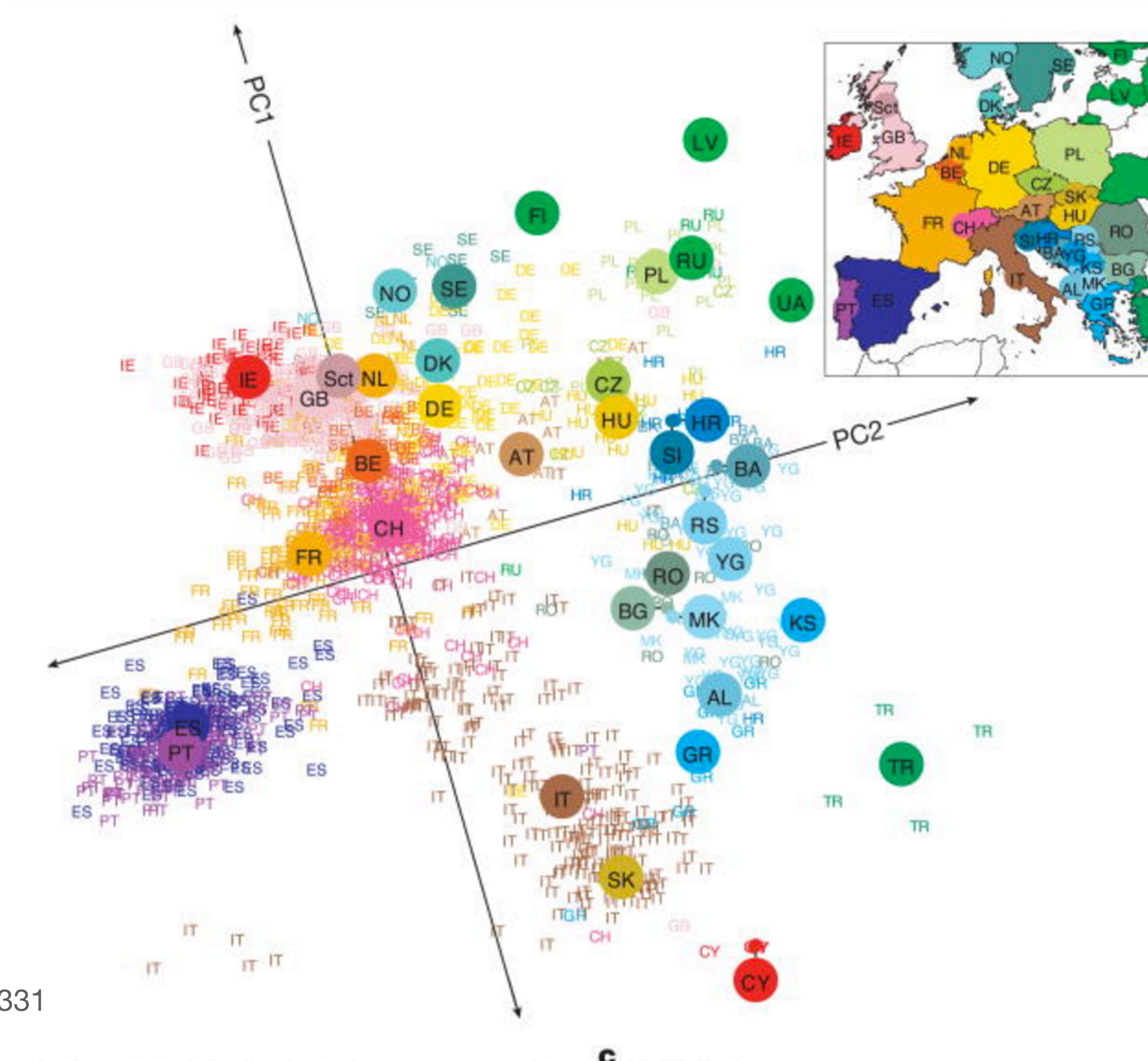
Go to:

The outbreaks of Coronavirus (COVID-19) epidemic have increased the pressure on healthcare and medical systems worldwide. The timely diagnosis of infected patients is a critical step to limit the spread of the COVID-19 epidemic. The chest radiography imaging has shown to be an effective screening technique

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7679792/>

A large neural network analyzes chest X-ray images for **COVID-19**, with heat map showing affected regions.

# Classical methods: PCA and clustering



<https://www.nature.com/articles/nature07331>

# Modern methods: Generative models with latent variables

- Unsupervised learning algorithms learn to “synthesize” realistic data from a large collection of examples
- Text, images, music, art, genomes...

# Exploring the UN General Debates with Dynamic Topic Models



Luke Lefebure

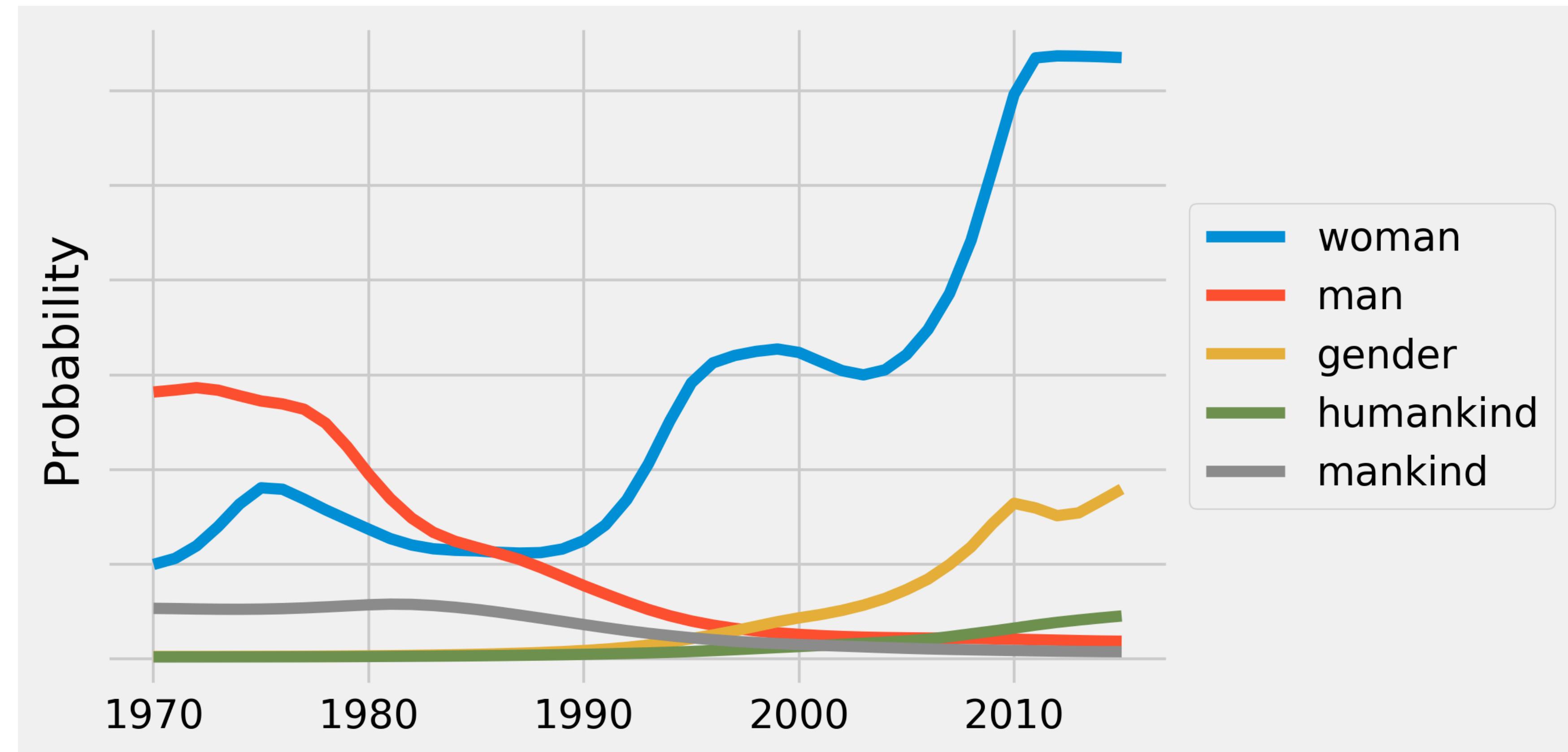
[Follow](#)

Oct 17, 2018 · 11 min read

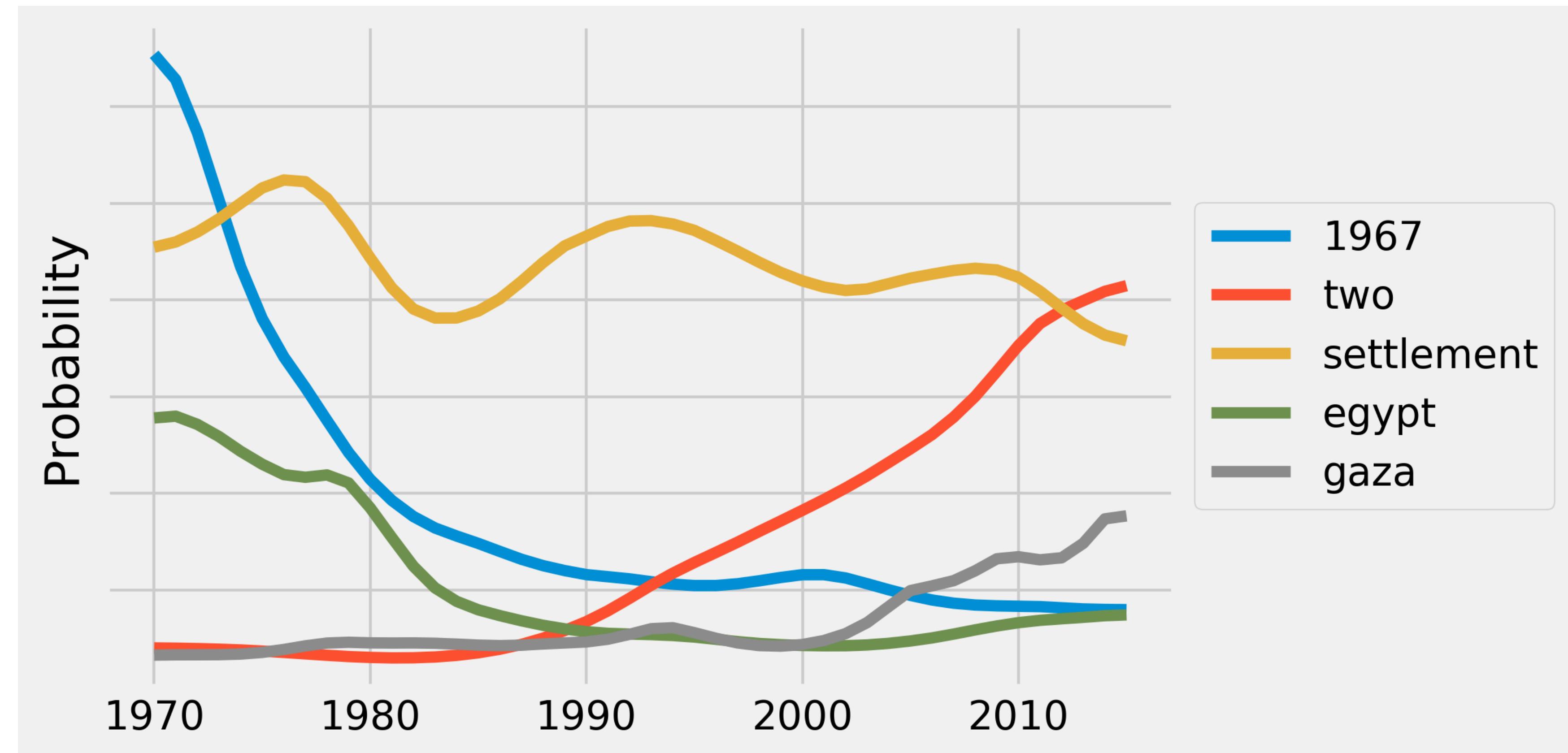


Credit: [Vladislav Klapin](#) on [Unsplash](#)

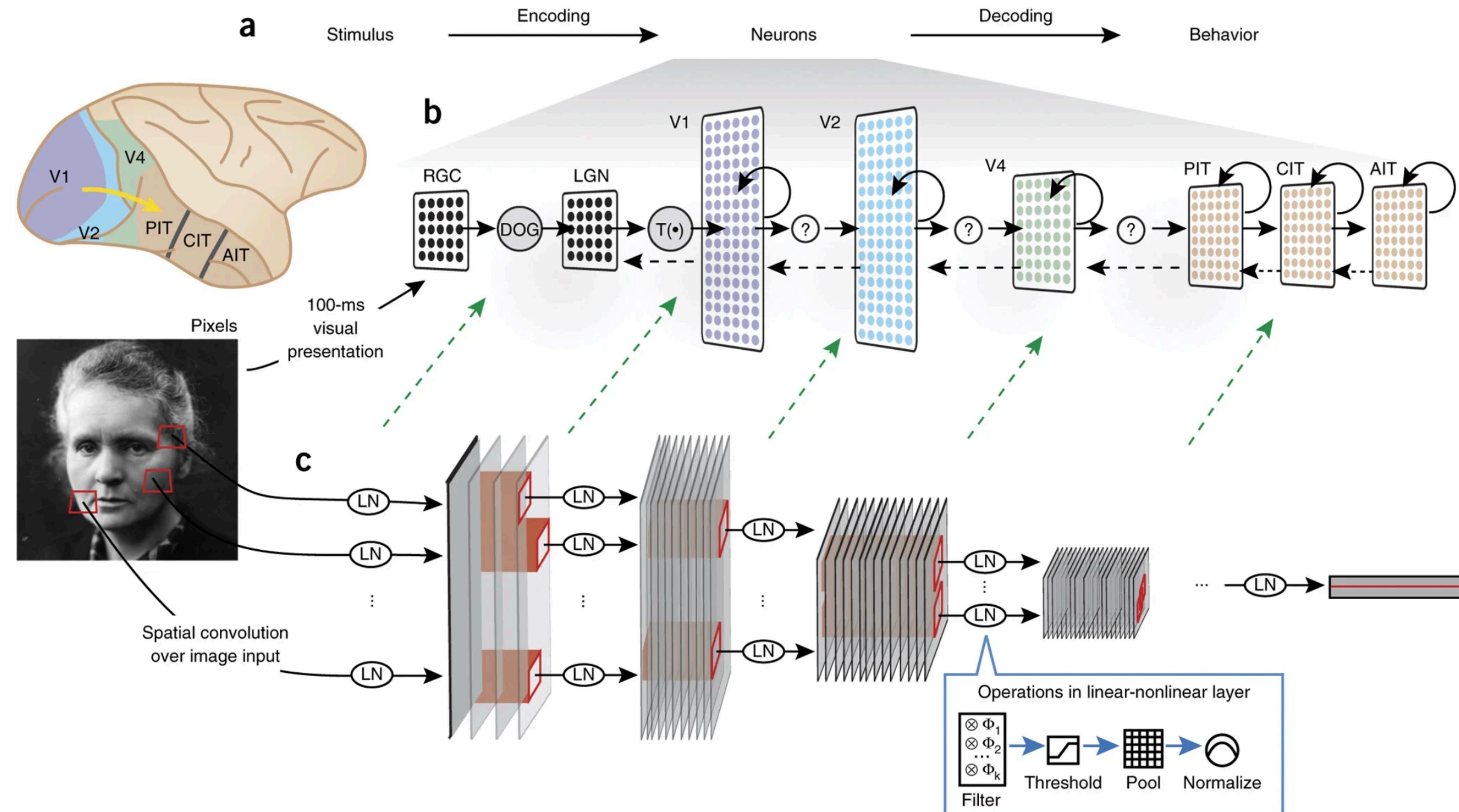
# Latent topic: Human rights



# Latent topic: Middle East peace



# Convolutional neural networks



<https://www.nature.com/articles/nn.4244>



These people may look familiar, like ones you've seen on Facebook or Twitter.



These people may look familiar, like ones you've seen on Facebook or Twitter.



BLOG POST  
RESEARCH

17 JUN 2016

# Deep Reinforcement Learning

---

SHARE

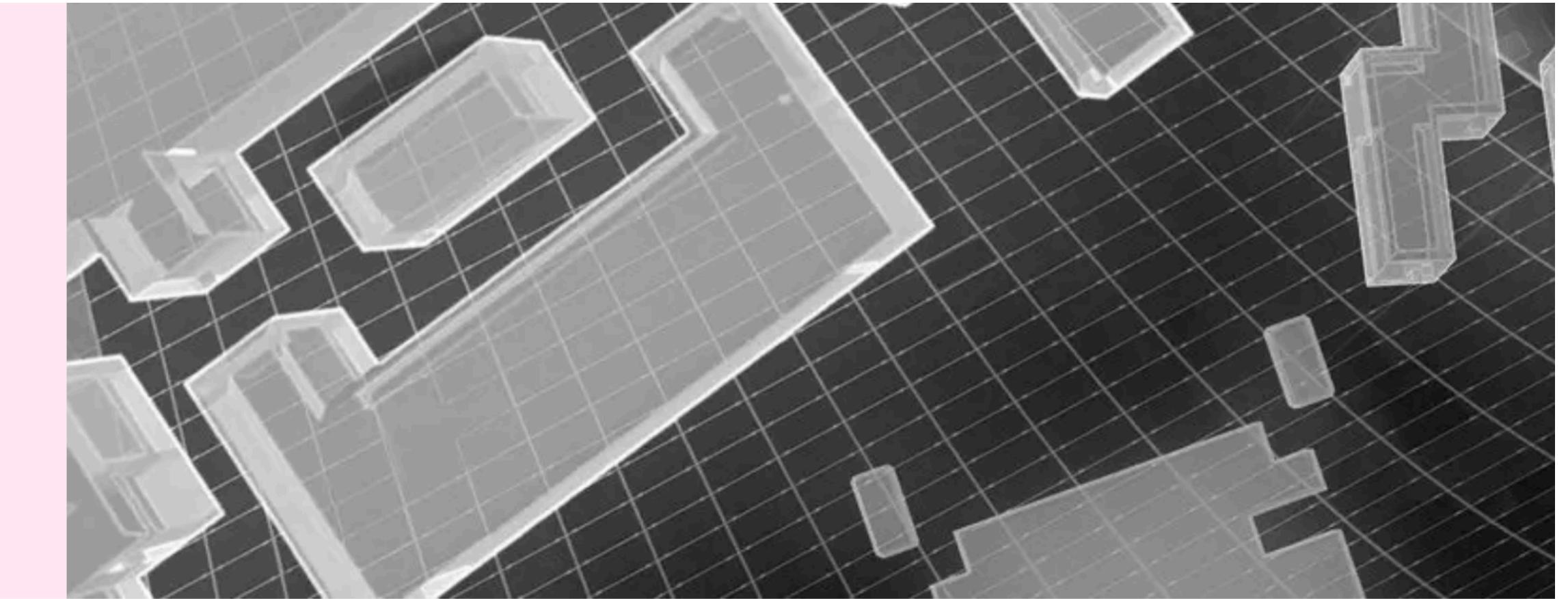


---

AUTHORS



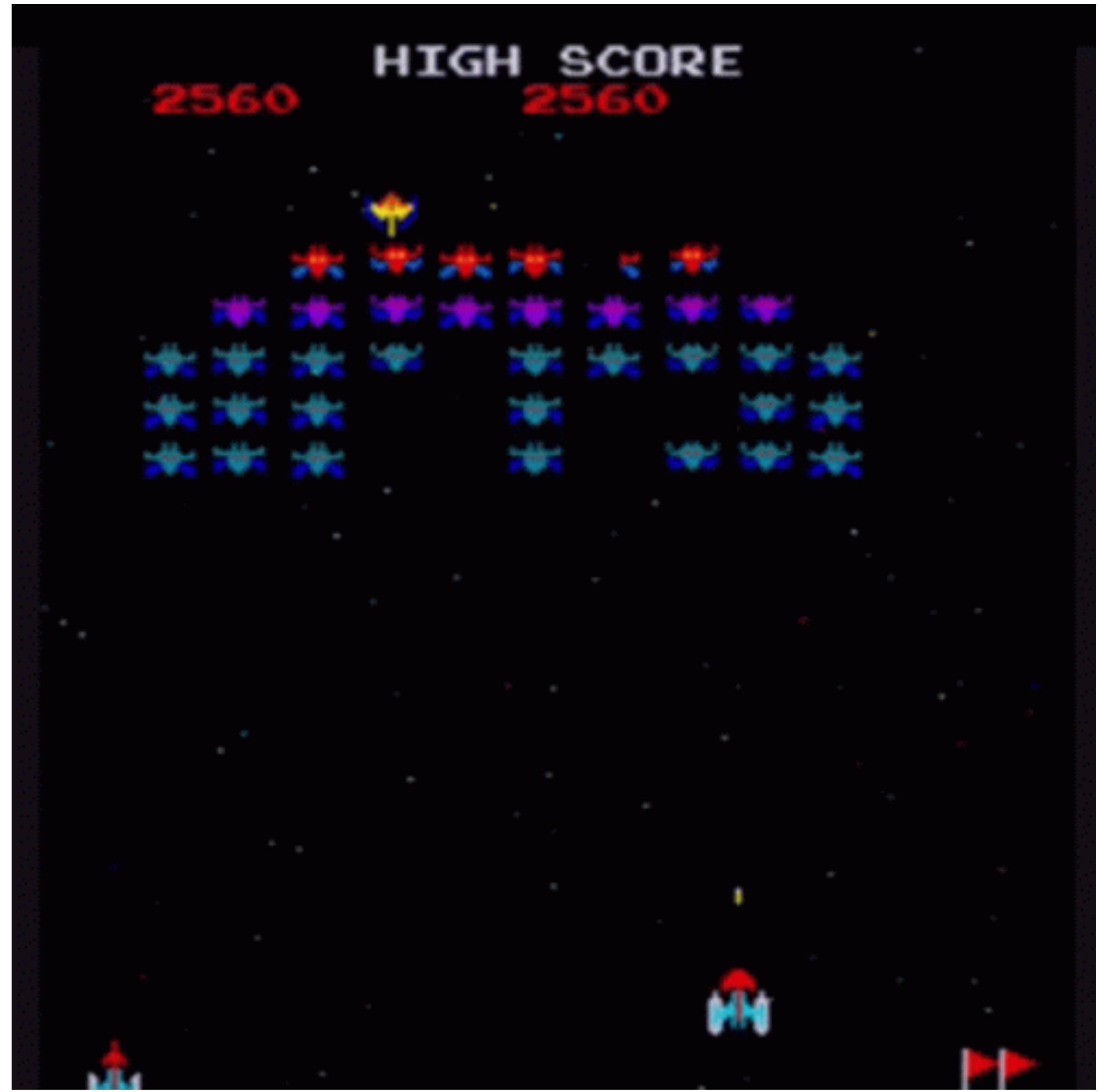
David Silver

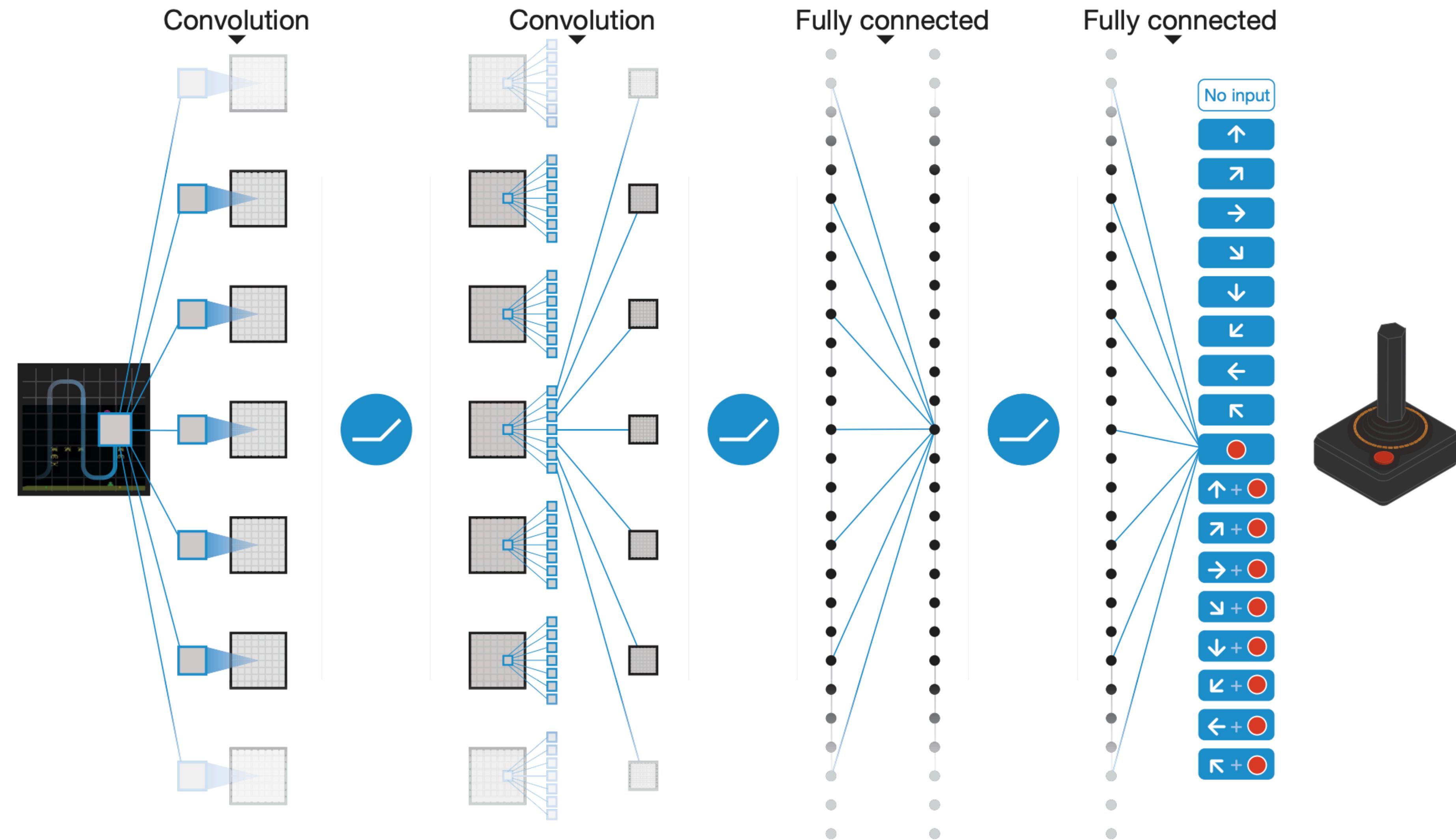


Humans excel at solving a wide variety of challenging problems, from low-level motor control through to high-level cognitive tasks. Our goal at DeepMind is to create artificial agents that can achieve a similar level of performance and generality. Like a human, our agents learn for themselves to achieve successful strategies that lead to the greatest long-term rewards. This paradigm of learning by trial-and-error, solely from rewards or punishments, is known as reinforcement learning (RL). Also like a human, our agents construct and learn their own knowledge directly from raw inputs, such as vision, without any hand-engineered features or domain heuristics. This is achieved by deep learning of neural

<https://deepmind.com/blog/article/deep-reinforcement-learning>



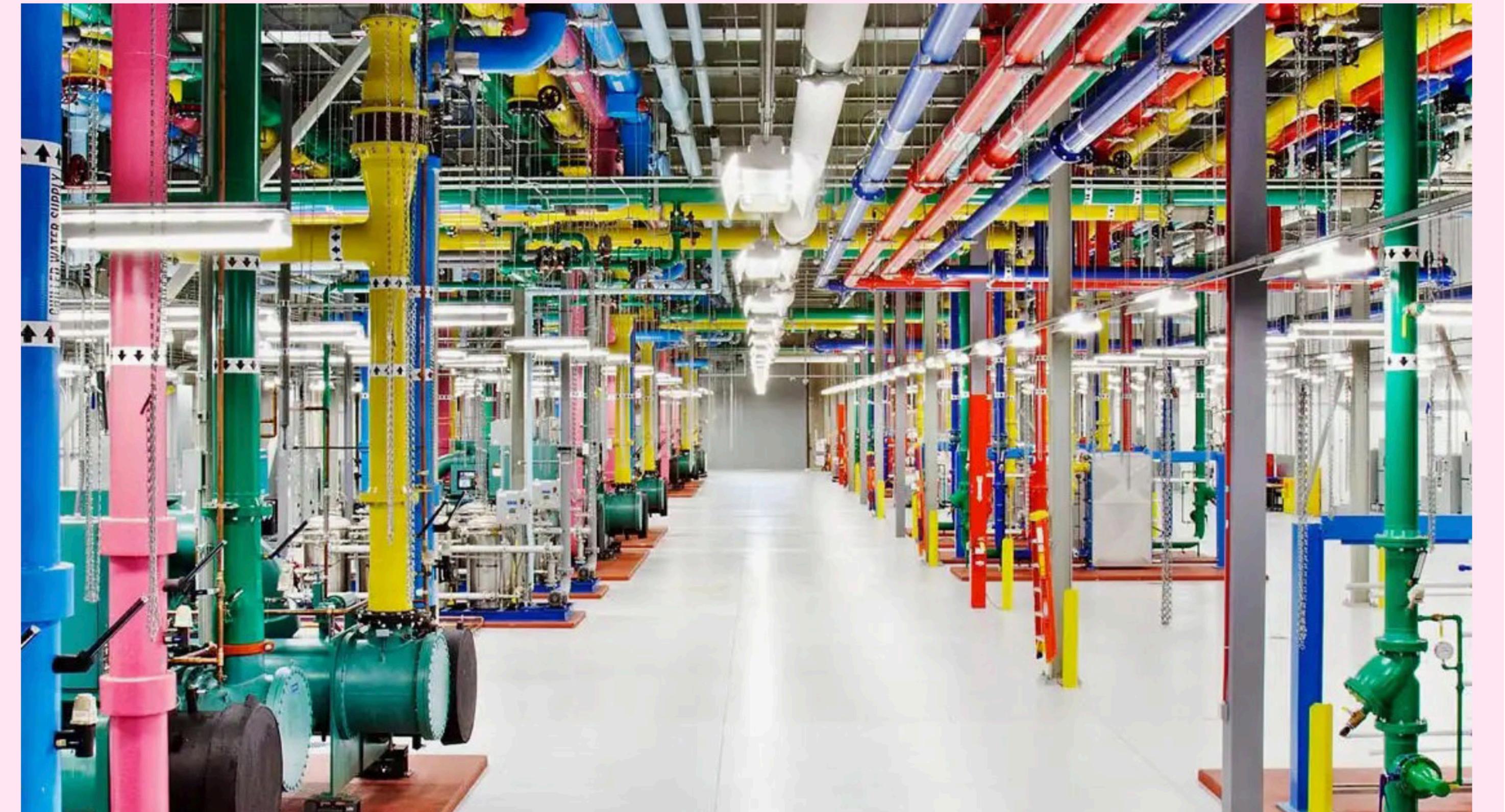






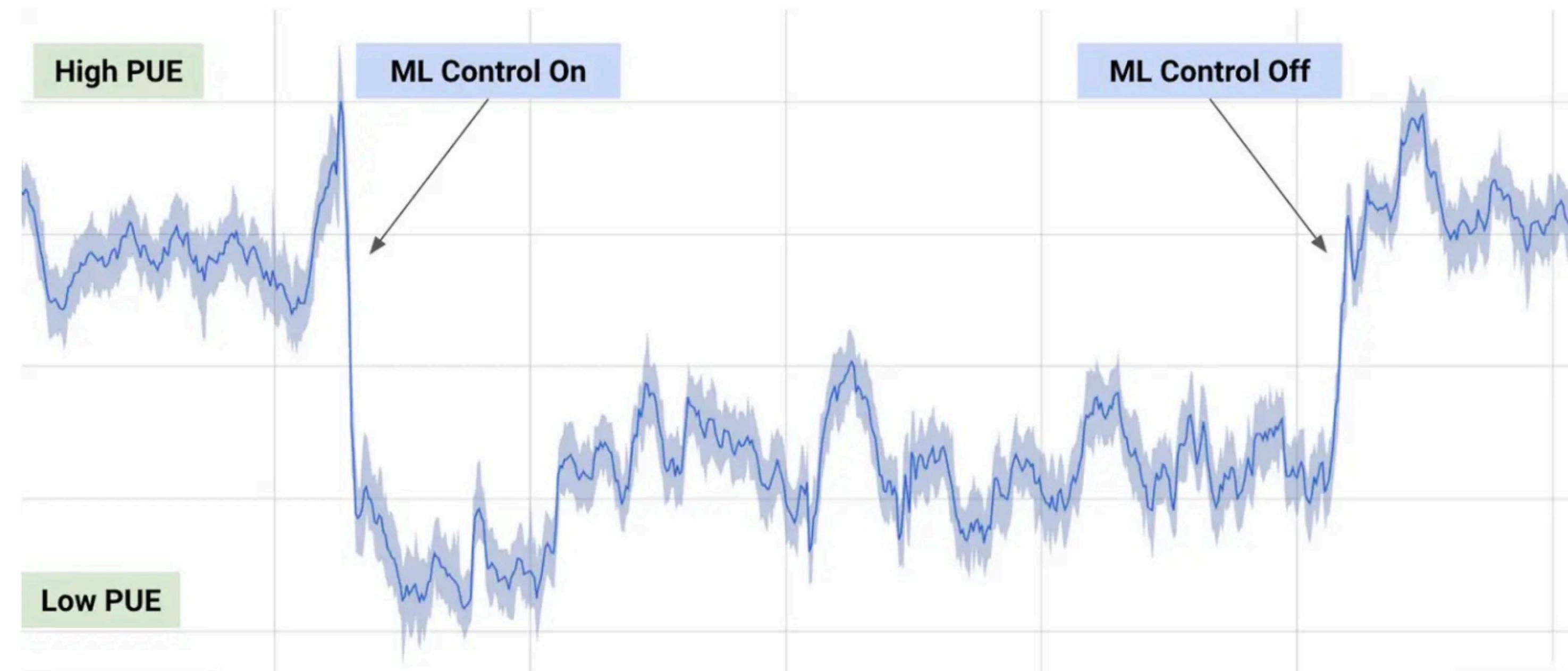
## DeepMind AI Reduces Google Data Centre Cooling Bill by 40%

July 20, 2016



<https://www.deepmind.com/blog/deepmind-ai-reduces-google-data-centre-cooling-bill-by-40>

We tested our model by deploying on a live data centre. The graph below shows a typical day of testing, including when we turned the machine learning recommendations on, and when we turned them off.



# When does learning take place?

The Bellman equation tells us:

*Learning takes place when expectations are violated. The receipt of the reward does not itself cause changes.*

# A Neural Substrate of Prediction and Reward

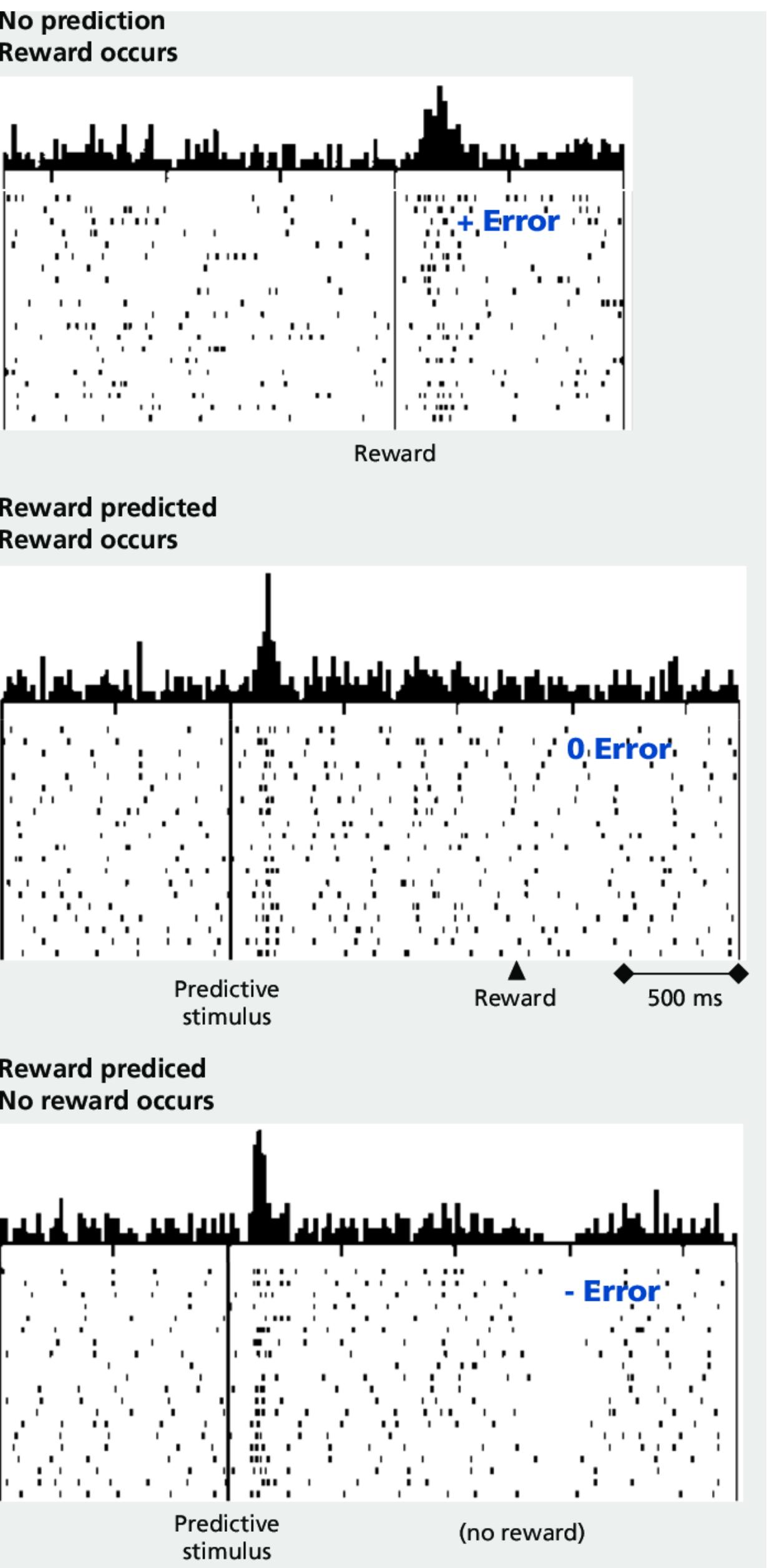
Wolfram Schultz, Peter Dayan, P. Read Montague\*

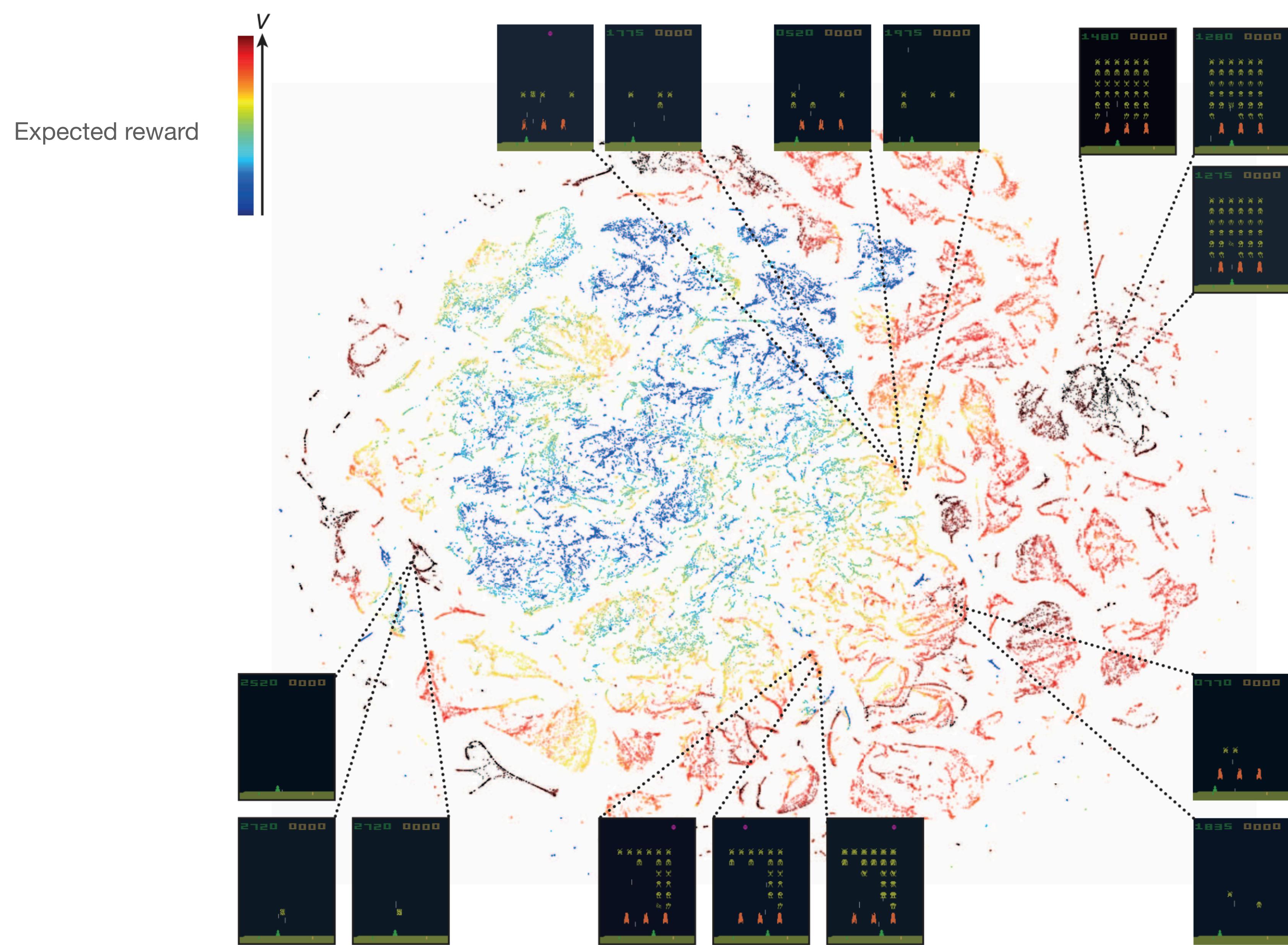
The capacity to predict future events permits a creature to detect, model, and manipulate the causal structure of its interactions with its environment. Behavioral experiments suggest that learning is driven by changes in the expectations about future salient events such as rewards and punishments. Physiological work has recently complemented these studies by identifying dopaminergic neurons in the primate whose fluctuating output apparently signals changes or errors in the predictions of future salient and rewarding events. Taken together, these findings can be understood through quantitative theories of adaptive optimizing control.

---

Science 1997

## Firings of dopamine neurons





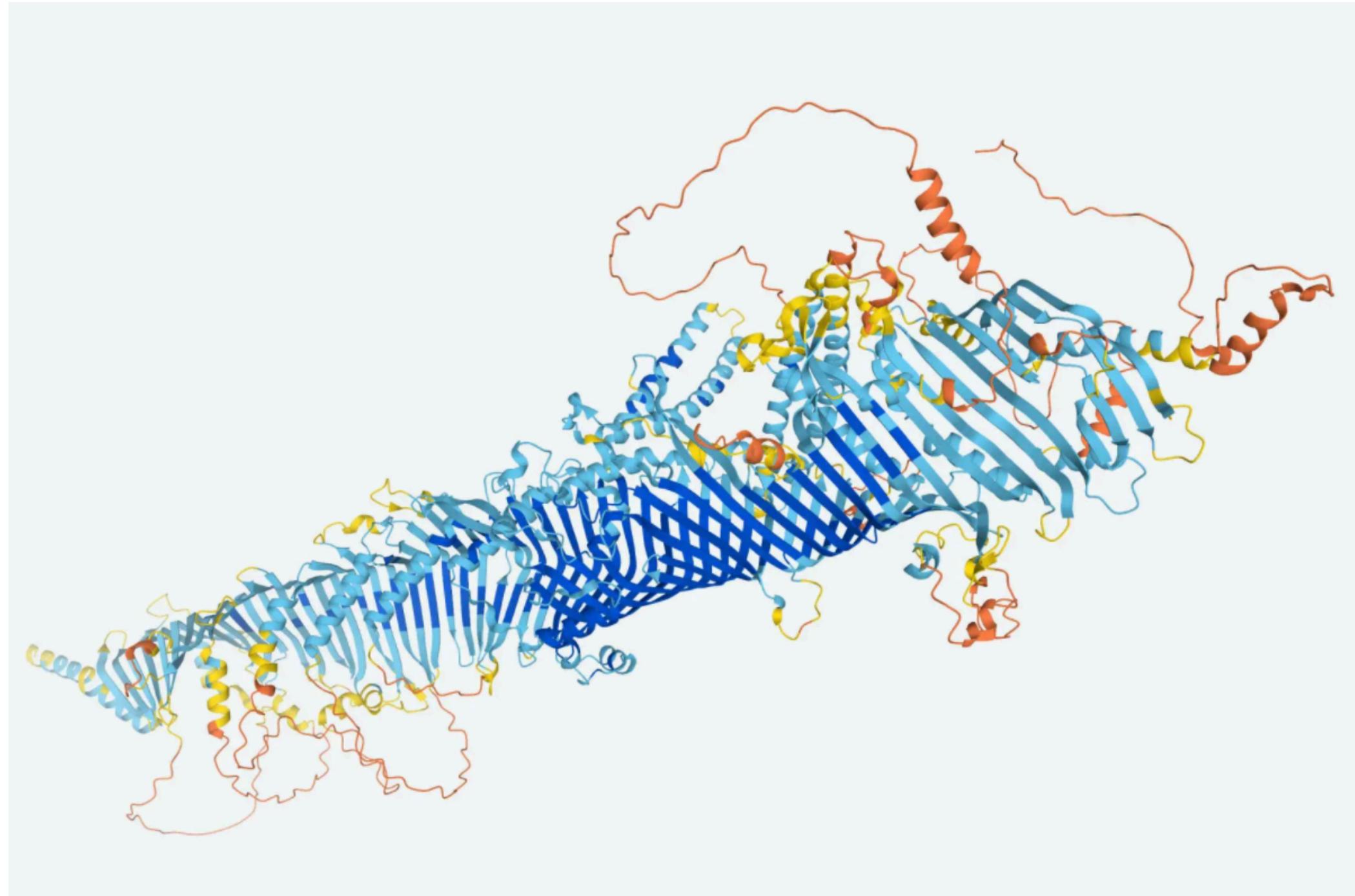


Opportunities and risks

Prospects for science and health  
Ethics and safe AI  
The next frontier  
Wu Tsai Institute at Yale

# *A.I. Predicts the Shapes of Molecules to Come*

DeepMind has given 3-D structure to 350,000 proteins, including every one made by humans, promising a boon for medicine and drug design.



# AI for protein folding

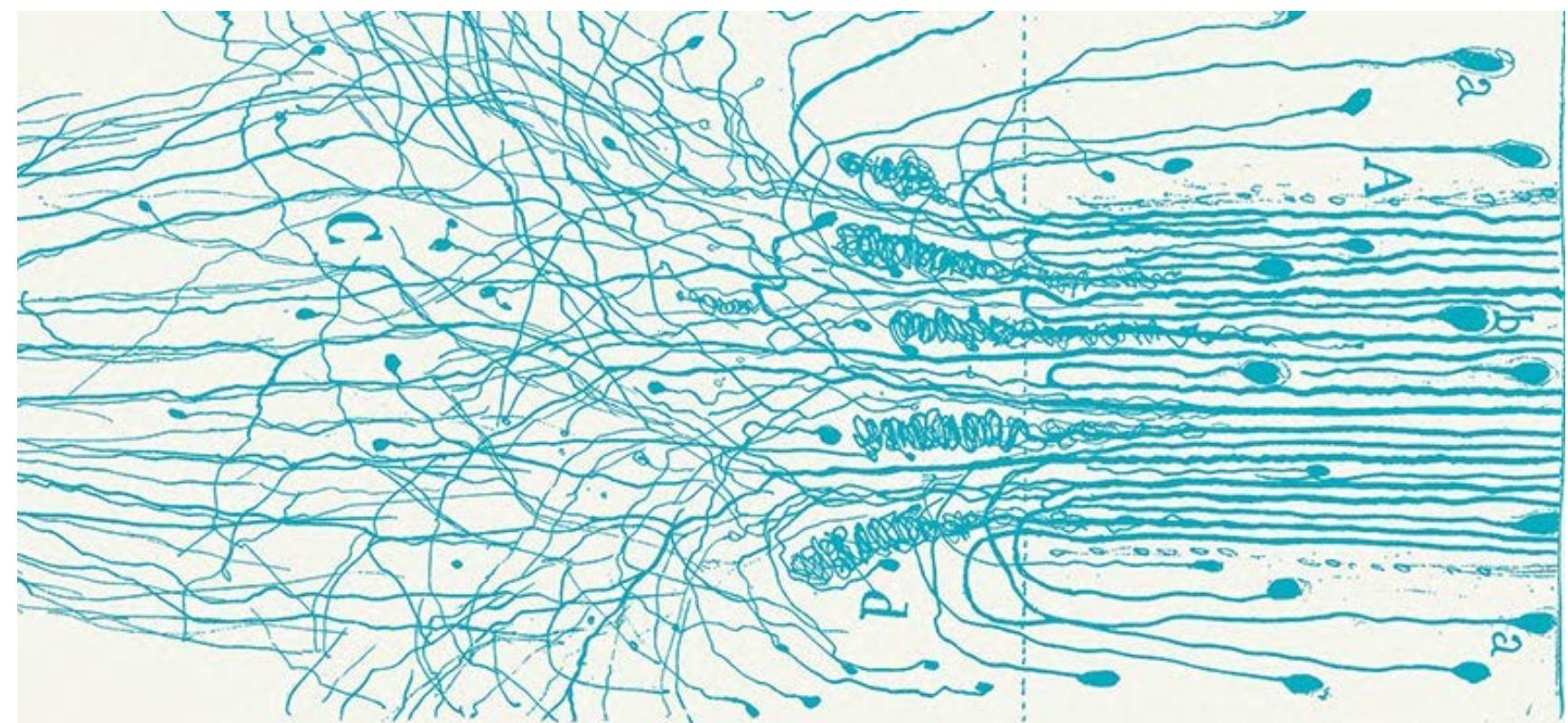
- Deep neural networks used to predict properties of protein, including distances between pairs of amino acids, and angles between chemical bonds
- “From scratch,” not using templates of known proteins
- Highly interdisciplinary, large team at DeepMind
- Published database of structures for 350,000 proteins

# Recurring themes

- Breakthroughs across data types and problems
- Unexpected impact on different areas of inquiry
- Methods and code are open and shared
- New models of collaborative research

# Prescription for new ML areas

- Combine hard-won domain insights and data with state-of-the-art ML approaches
- Target a bold, concrete, transformative subproblem
- Work together in teams

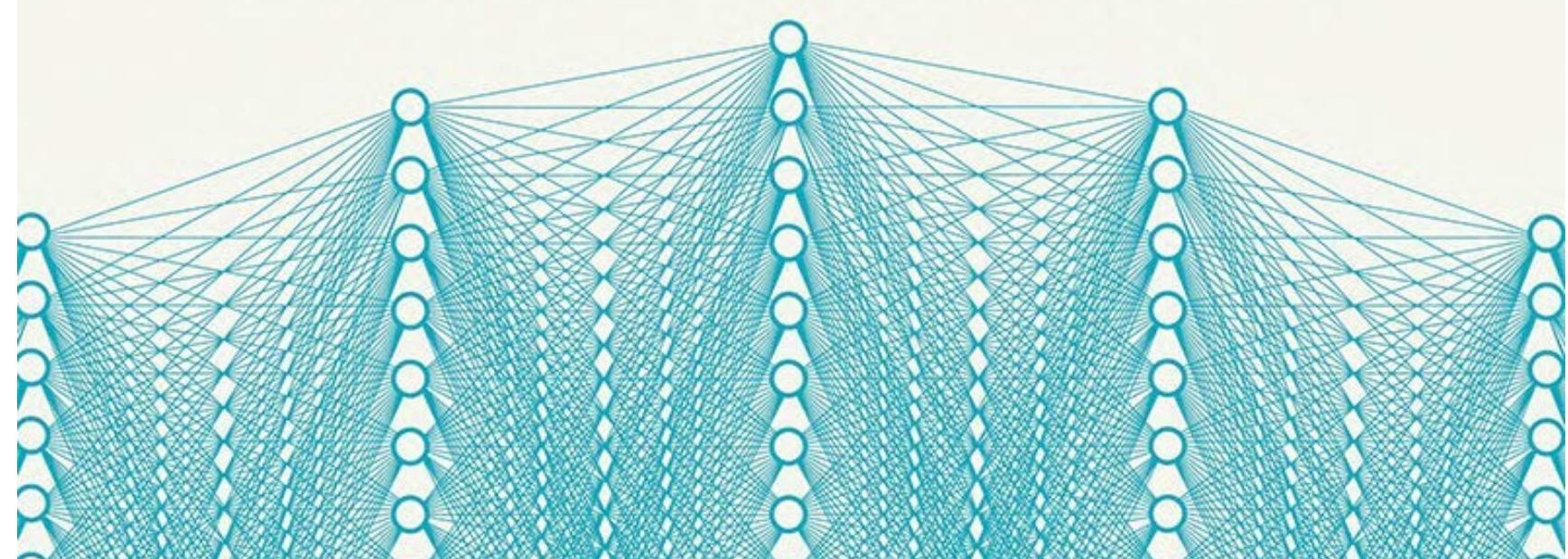


# THE ALIGNMENT PROBLEM

Machine Learning and Human Values

## BRIAN CHRISTIAN

Best-Selling Author, *Algorithms to Live By*



# The Ezra Klein Show



1 hr 16 min

PLAY ►



## Is A.I. the Problem? Or Are We?

The Ezra Klein Show

Society & Culture

[Listen on Apple Podcasts ↗](#)



If you talk to many of the people working on the cutting edge of artificial intelligence research, you'll hear that we are on the cusp of a technology that will be far more transformative than simply computers and the internet, one that could bring about a new industrial revolution and usher in a utopia — or perhaps pose the greatest threat in our species's history.

Others, of course, will tell you those folks are nuts.

**Yale:**

```
[ 0.78310001, 0.51717001, -0.38207 , -0.23722 , -0.31615999, 0.30805001, 0.76389998, 0.064106 , -0.74913001,  
0.60585999, -0.23871 , -0.16876 , -0.25634 , 1.07270002, -0.29967999, 0.020095 , 0.54500997, -0.17847 , -0.26675999,  
-0.11798 , -0.48692 , 0.22712 , 0.017473 , -0.4747 , 0.44861001, -0.084281 , -0.30412999, -1.13510001, -0.14869 , -0.11182 ,  
-0.32530001, 1.0029 , -0.35742 , 0.35148999, -1.10679996, -0.064142 , -0.72284001, 0.14114 , -0.41247001, -0.16184001,  
-0.54576999, -0.12958001, -0.88356 , -0.089722 , 0.10555 , -0.12288 , 0.92851001, 0.50032002, 0.1349 , 0.21457 ,  
0.35073999, -0.73132998, 0.39633 , -0.43239999, -0.38815999, -1.34669995, 0.37463999, -0.79386002, 0.11185 , 0.18007 ,  
-0.75142998, 0.24975 , -0.094948 , -0.36341 , 0.24869999, -0.22667 , 0.32289001, 1.29489994, 0.42658001, 1.29120004,  
-0.13954 , 0.68976003, 0.21586999, 0.13715 , -1.00919998, 0.028827 , 0.11011 , -0.1912 , -0.073198 , -0.52449 , 0.49199 ,  
0.14463 , -0.18844 , -0.75536001, -0.28704 , 0.019113 , 0.30349001, -0.74425 , -0.072221 , -0.40647 , 0.26899001, -0.28318  
, 0.72409999, 0.50796002, -0.37845999, -0.13008 , -0.13808 , 0.098928 , 0.16215999, 0.16293 ]
```

$\phi(\text{scientist}) - \phi(\text{woman}) + \phi(\text{man}):$ 

geologist  
engineer  
astronomer  
mathematician  
science

 $\phi(\text{scientist}) - \phi(\text{man}) + \phi(\text{woman}):$ 

anthropologist  
sociologist  
psychologist  
geneticist  
biochemist

$\phi(\text{smart}) - \phi(\text{girl}) + \phi(\text{boy}):$

wise  
better  
guy  
kind  
good  
kid

$\phi(\text{smart}) - \phi(\text{boy}) + \phi(\text{girl}):$

sexy  
pretty  
incredibly  
cute  
exciting  
funny

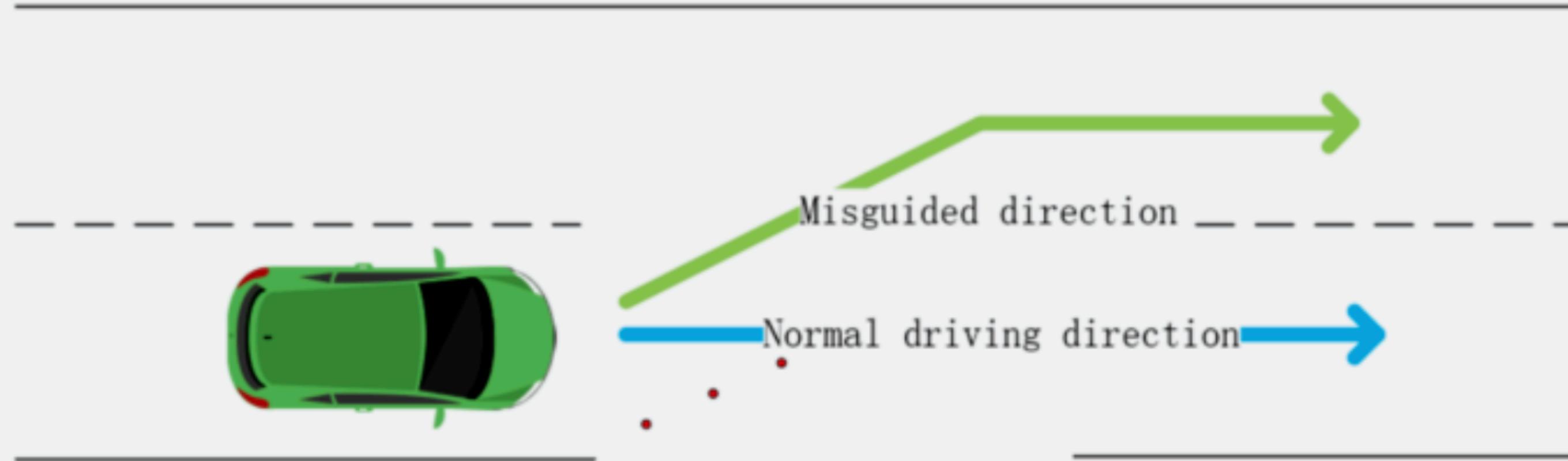
TESLA AUTOPILOT —

# Researchers trick Tesla Autopilot into steering into oncoming traffic

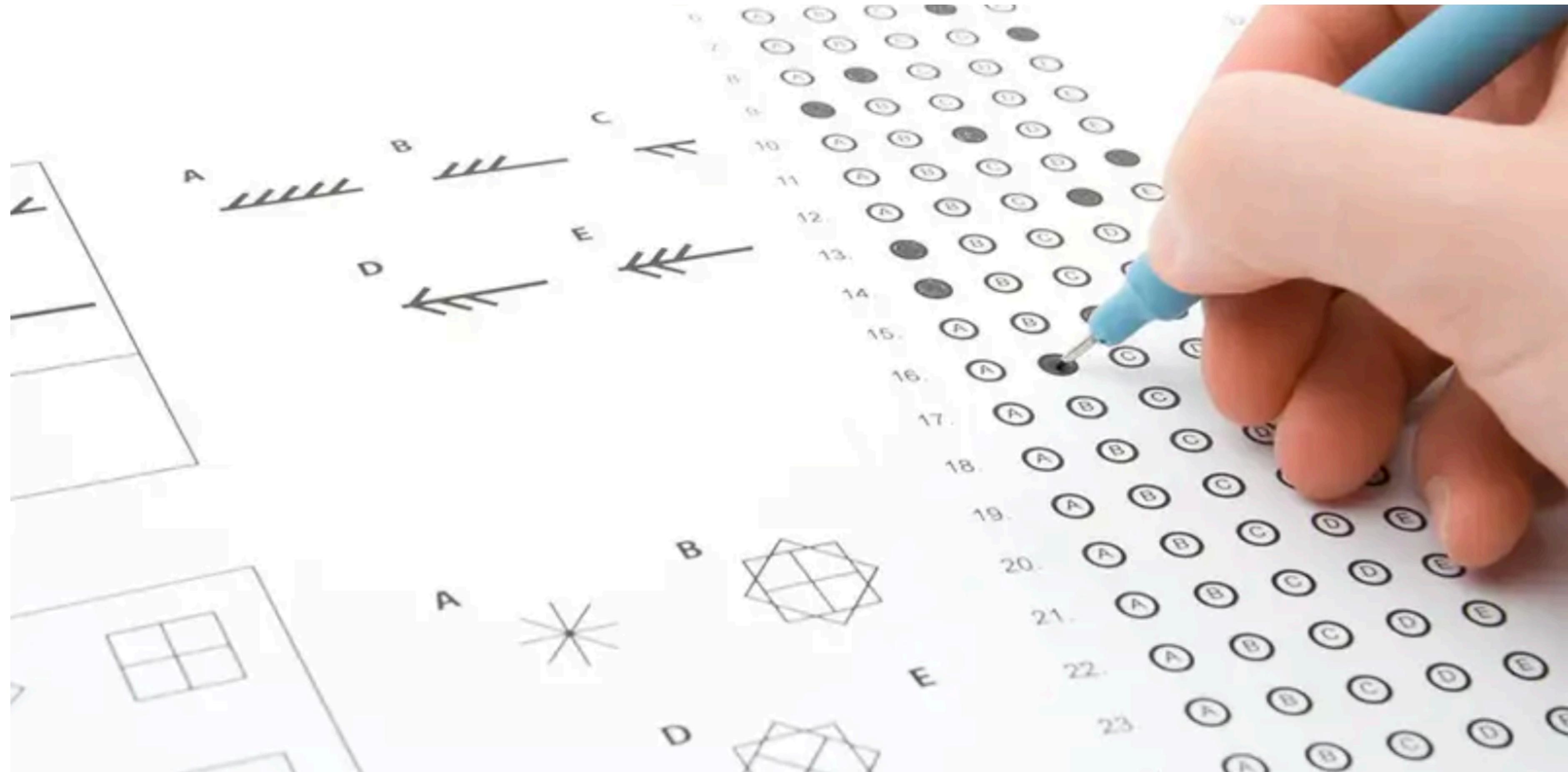
Stickers that are invisible to drivers and fool autopilot.

DAN GOODIN - 4/1/2019, 8:50 PM

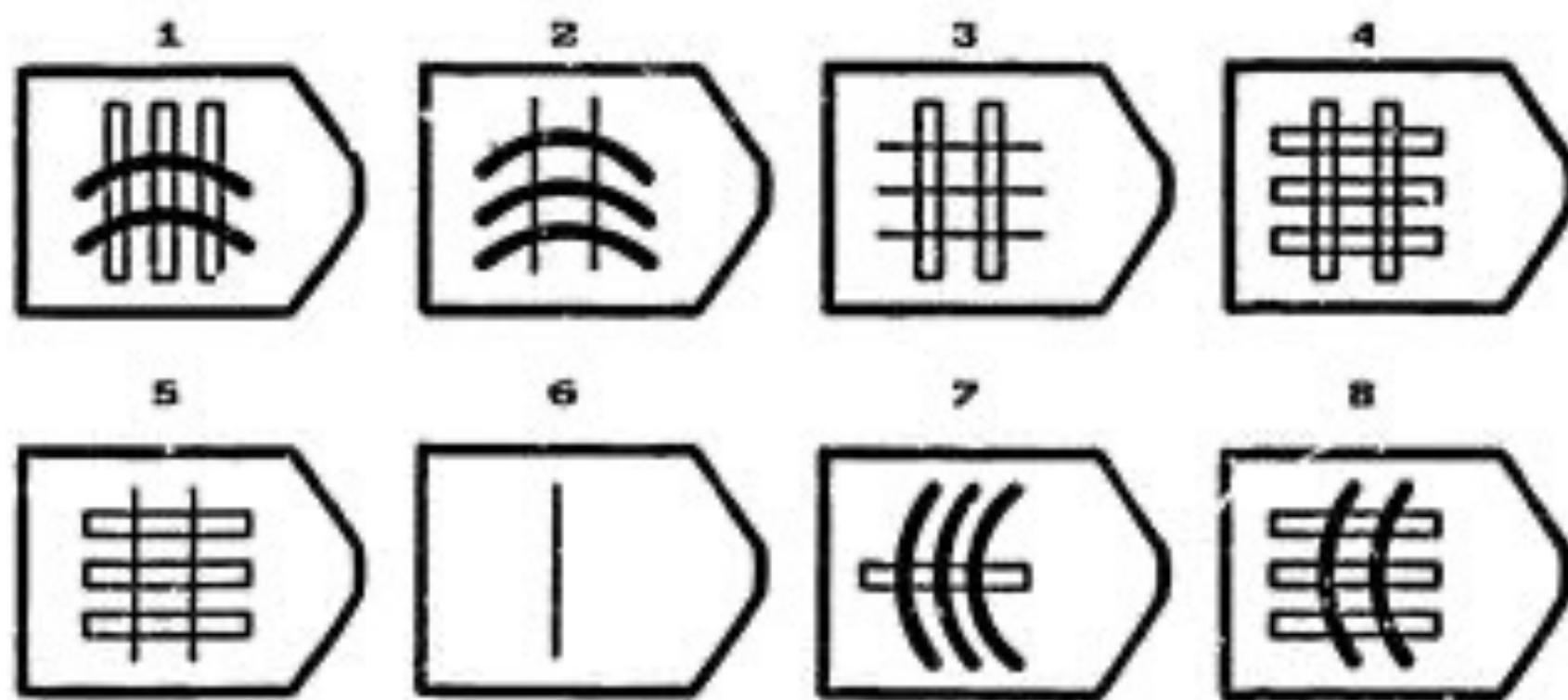
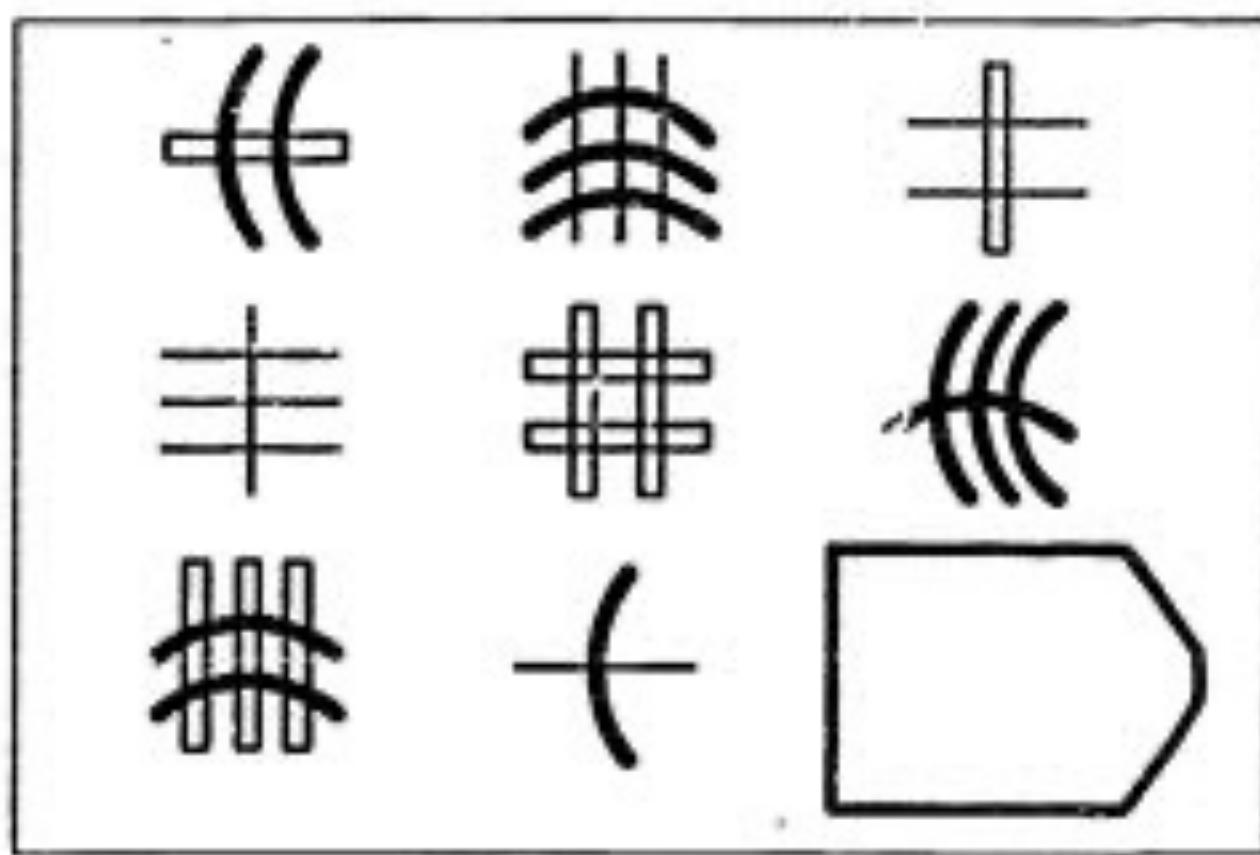
 Keen Security Lab



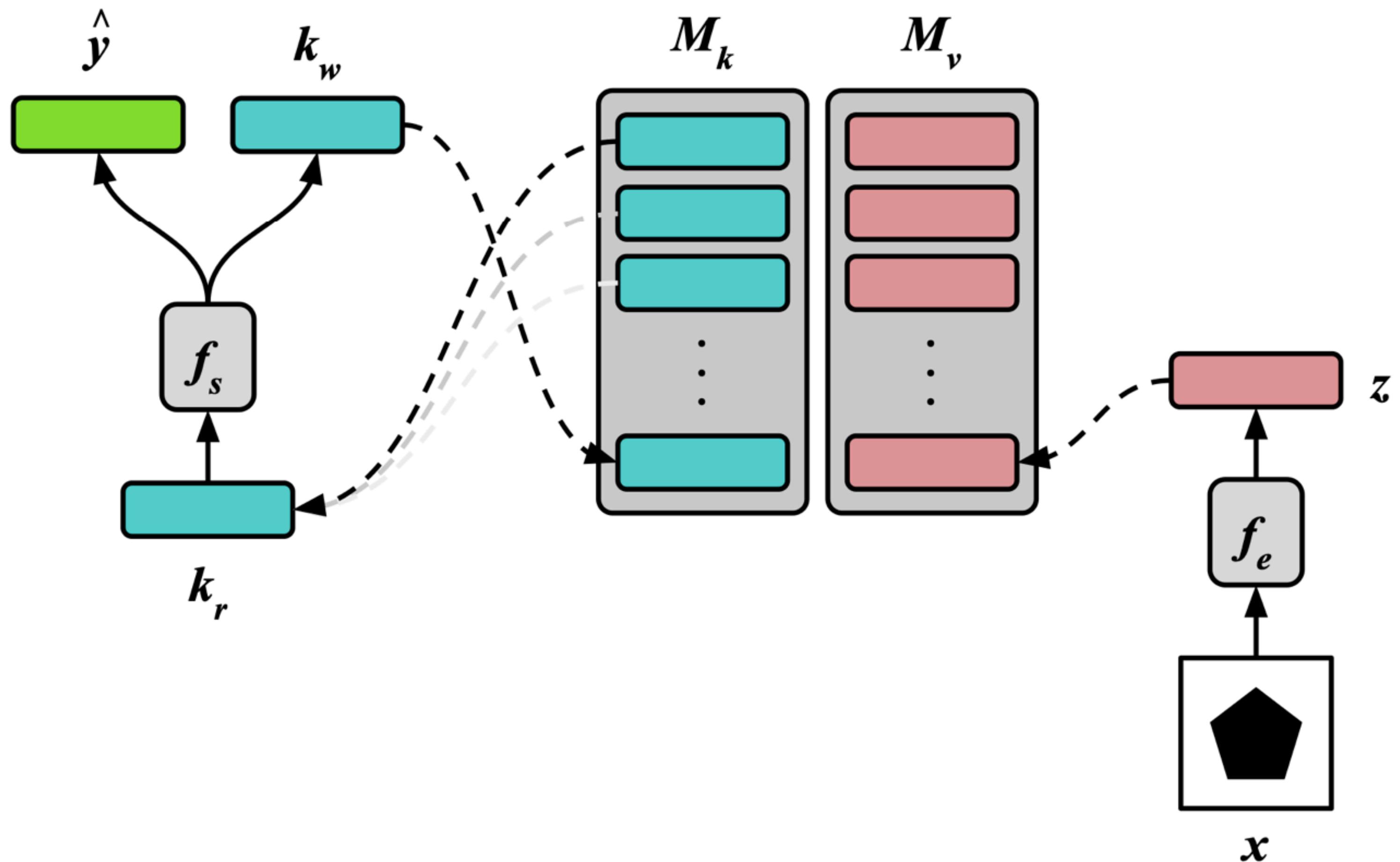
# What's next?



# What's next?



# What's next?



## What is machine learning?

Examples in daily life  
Examples in science  
Connection with AI

Supervised and unsupervised  
Latent variable models  
Deep learning  
Reinforcement learning

## Types of machine learning and AI frameworks

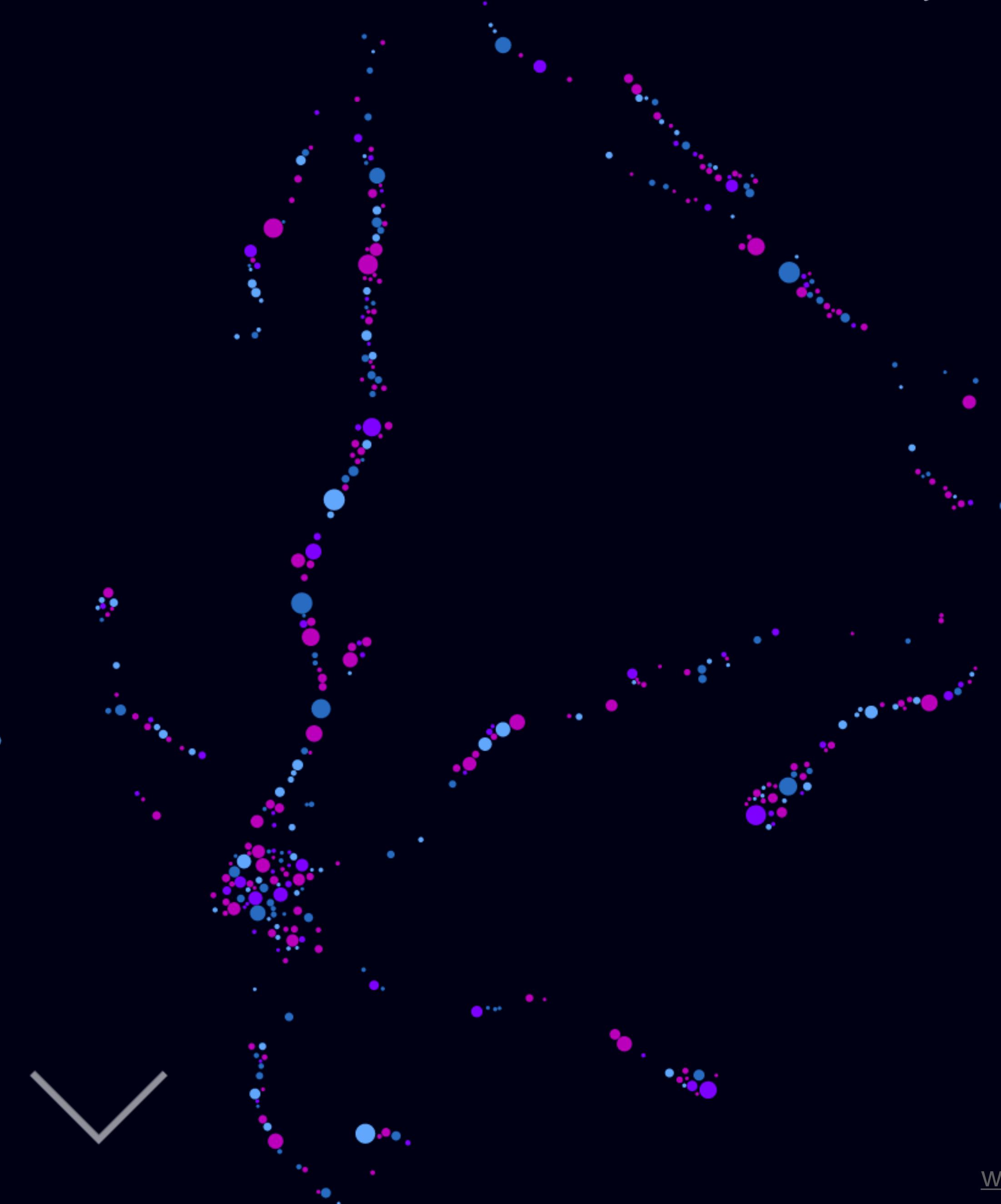
## Risks and opportunities

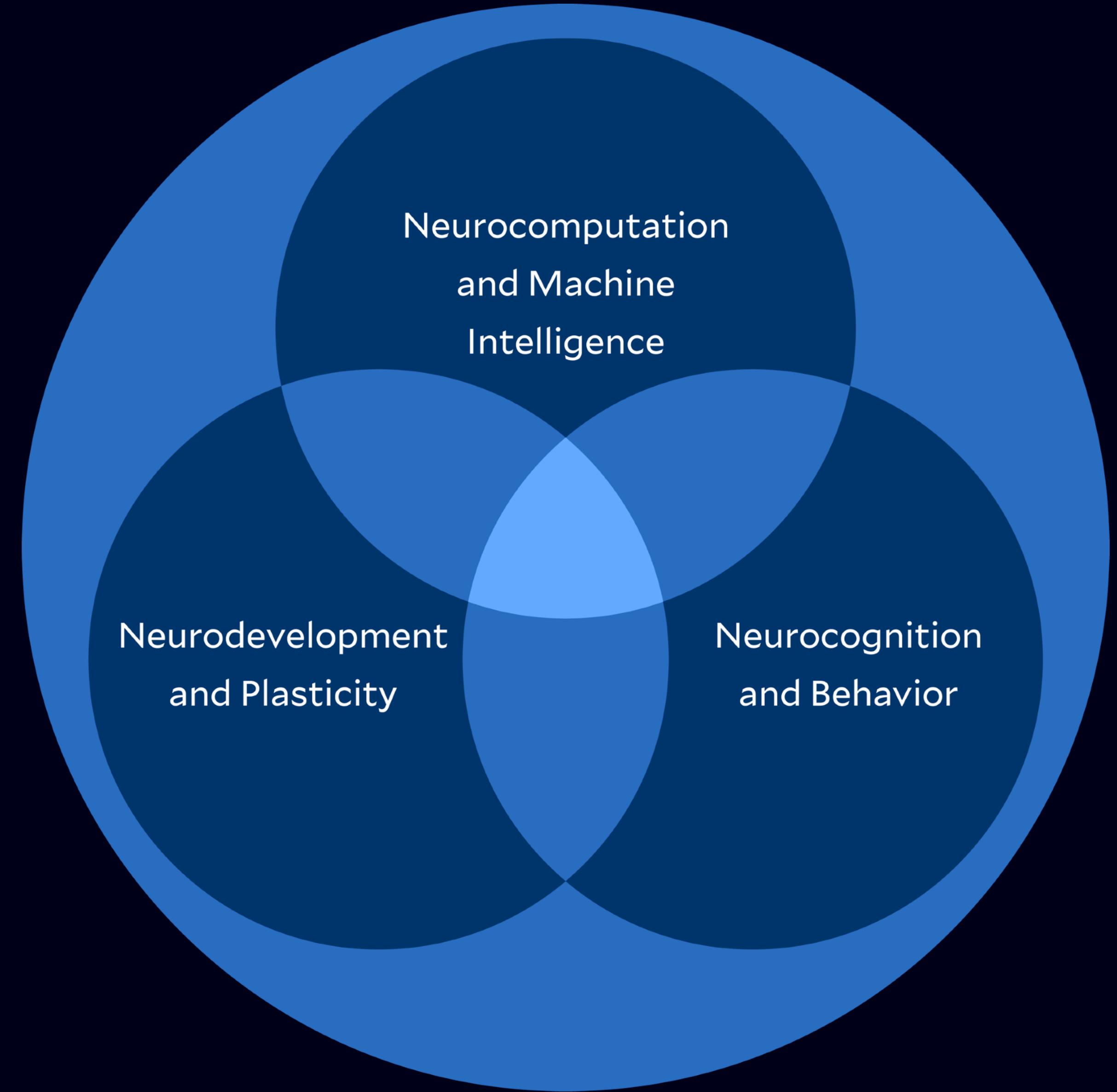
Prospects for science and health  
Ethics and safe AI  
The next frontier

# To know, together.

Our goal is to understand human cognition.

COGNITION :: COGNOSCERE ::  
CO [TOGETHER] + GNOSCERE [TO KNOW]





# Thank you!

These slides: <https://t.ly/DdVS>