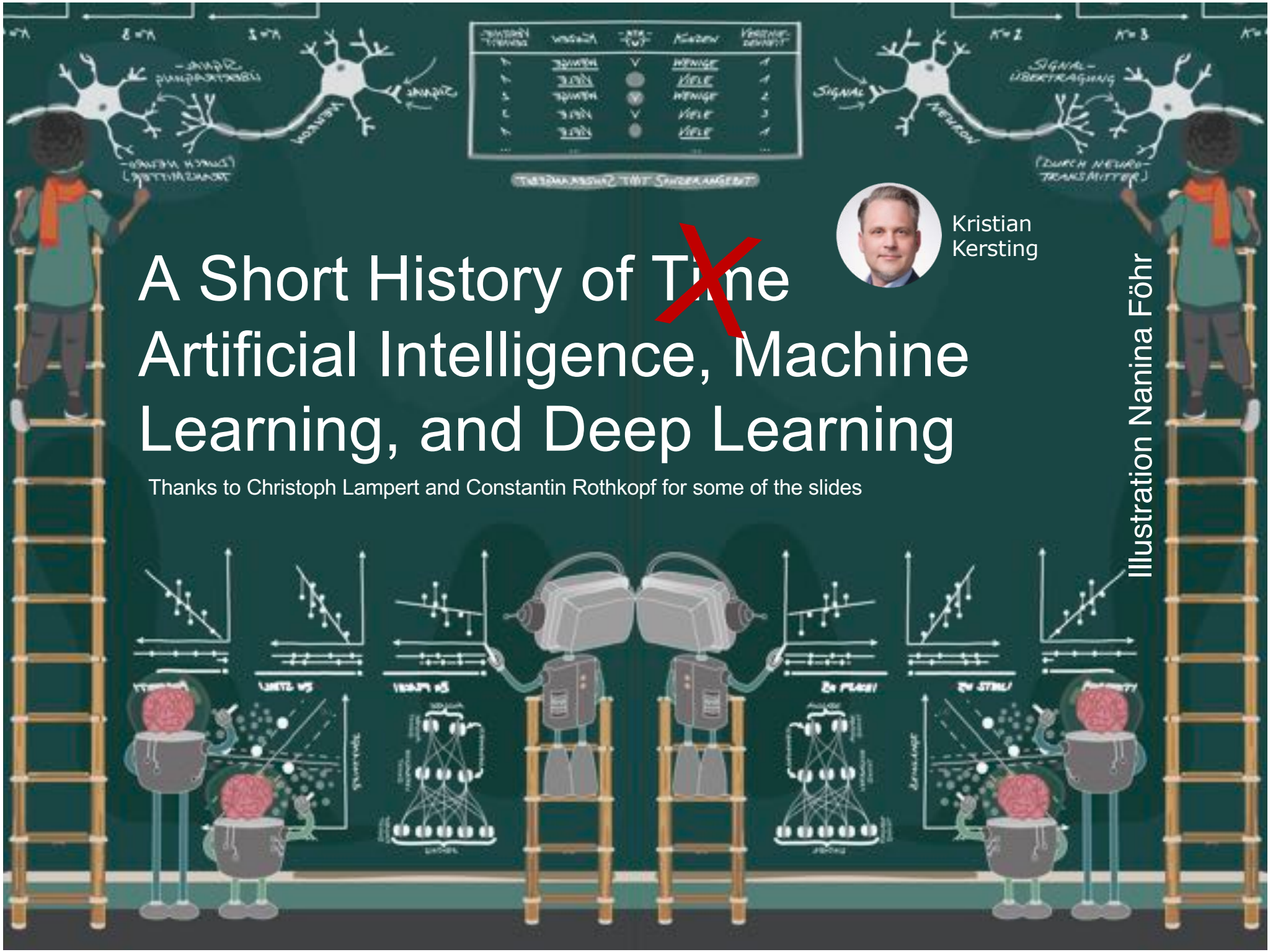


# A Short History of Time
# Artificial Intelligence, Machine Learning, and Deep Learning

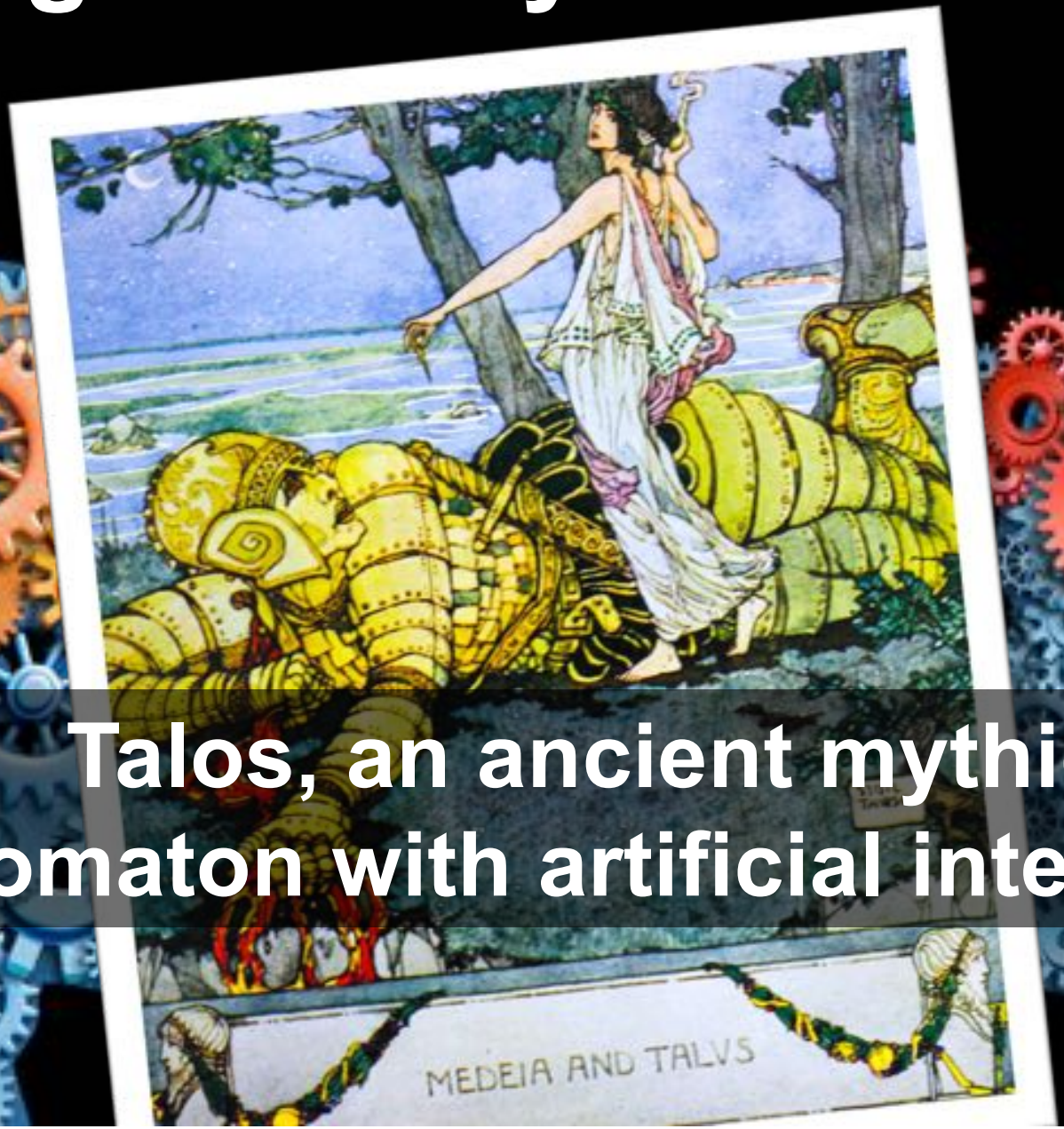


Kristian Kersting

Thanks to Christoph Lampert and Constantin Rothkopf for some of the slides

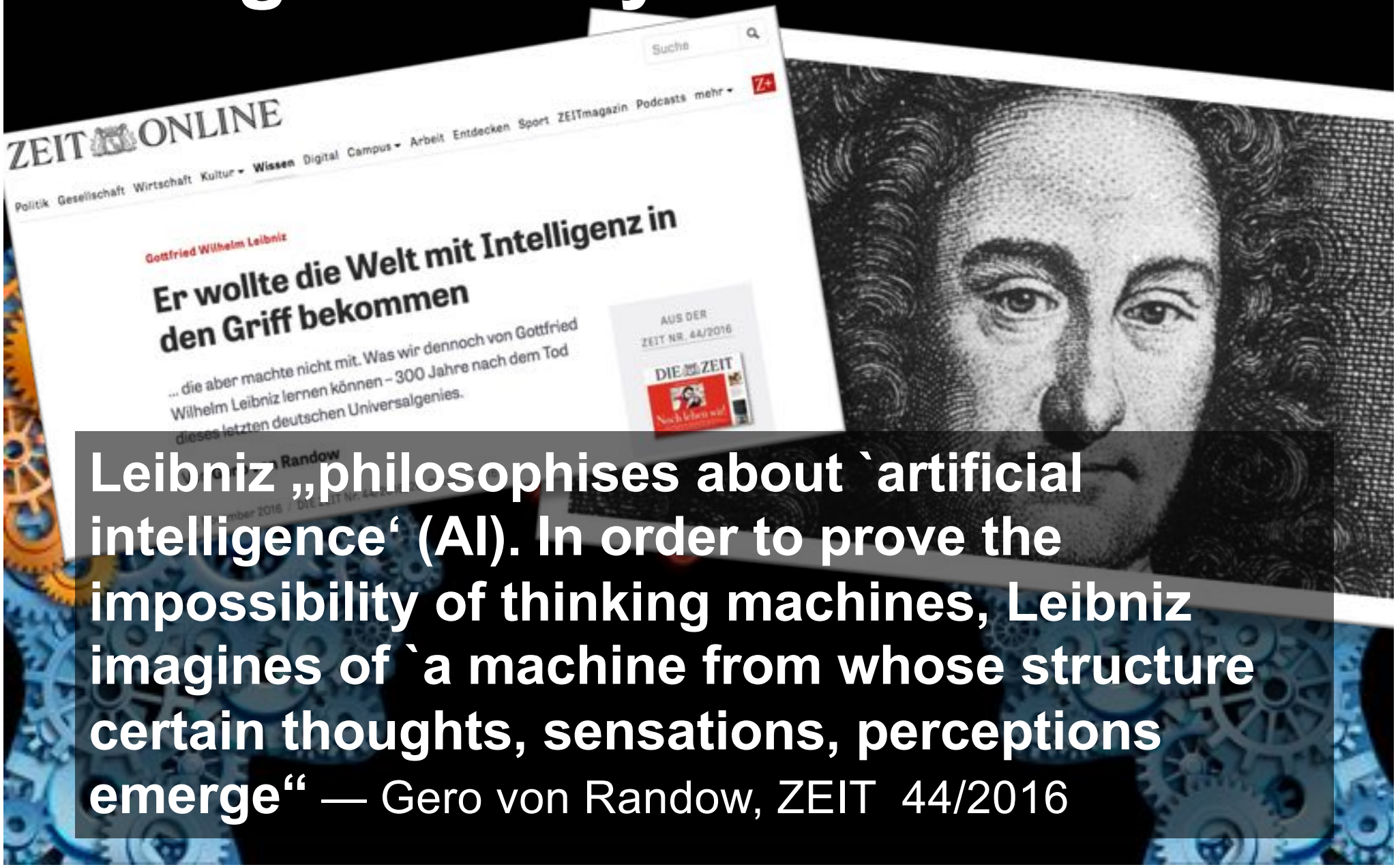
Illustration Nanina Föhr

# The dream of an artificially intelligent entity is not new



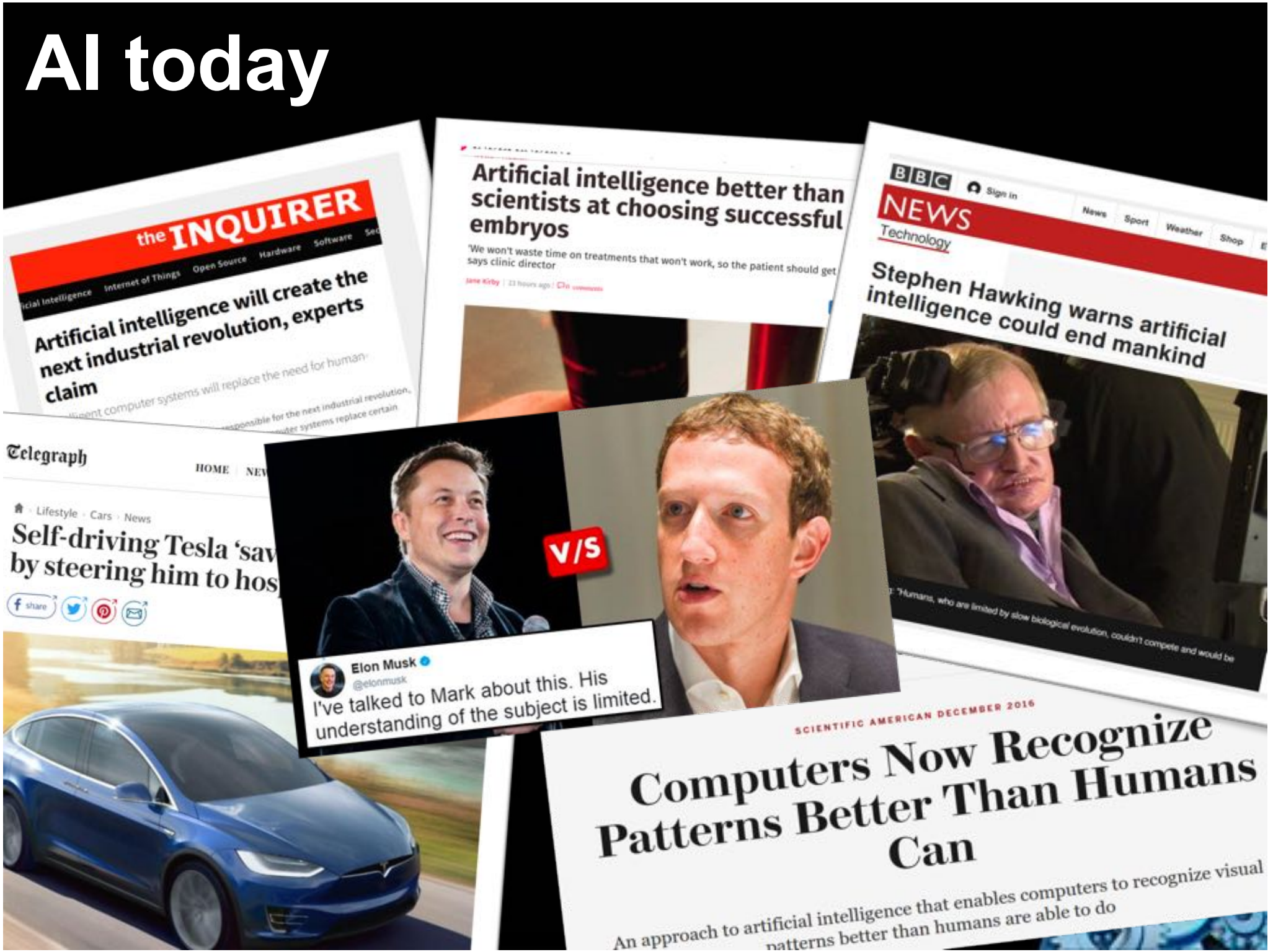

Talos, an ancient mythical automaton with artificial intelligence

# The dream of an artificially intelligent entity is not new

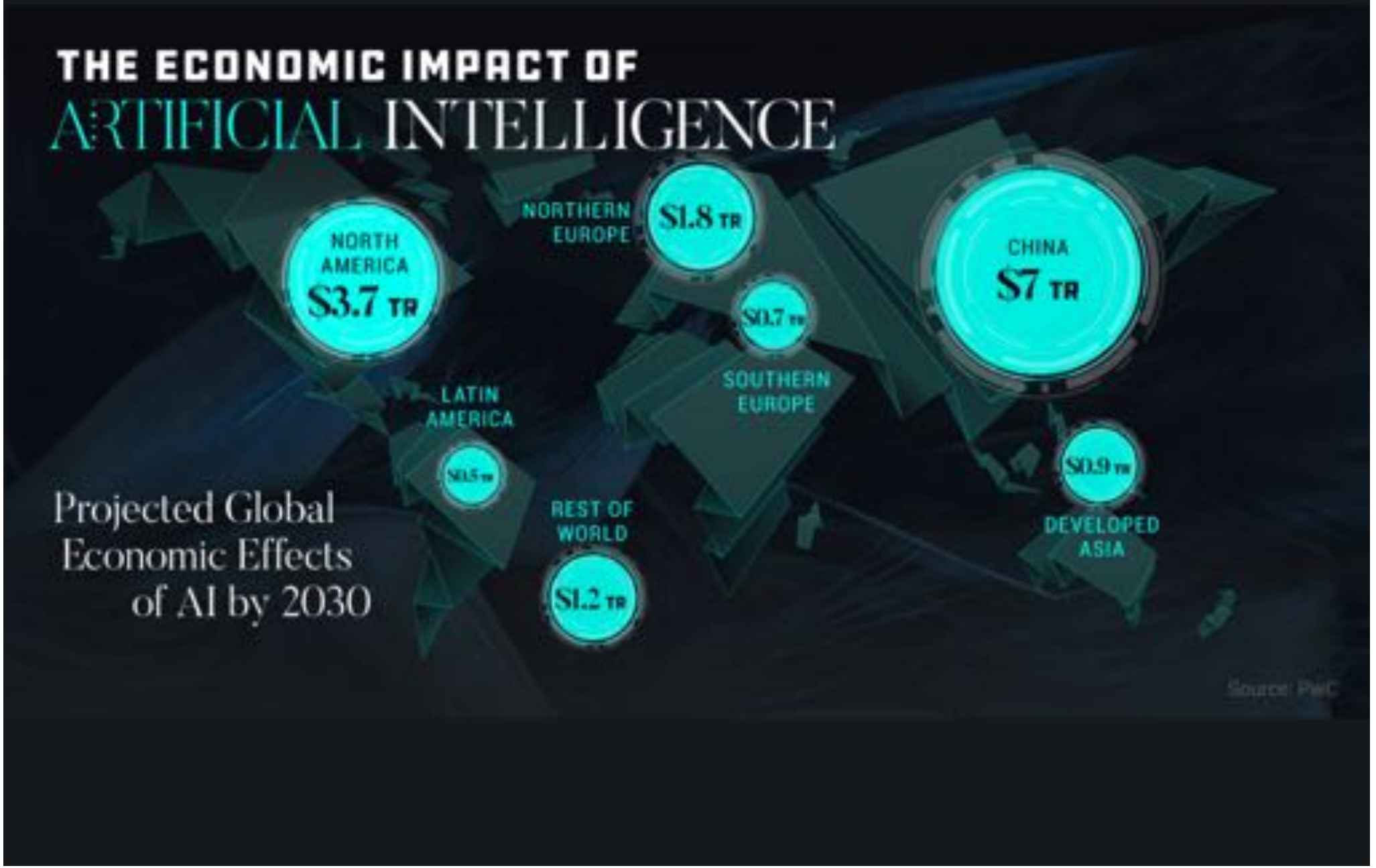



**Leibniz „philosophises about `artificial intelligence‘ (AI). In order to prove the impossibility of thinking machines, Leibniz imagines of `a machine from whose structure certain thoughts, sensations, perceptions emerge“** — Gero von Randow, ZEIT 44/2016

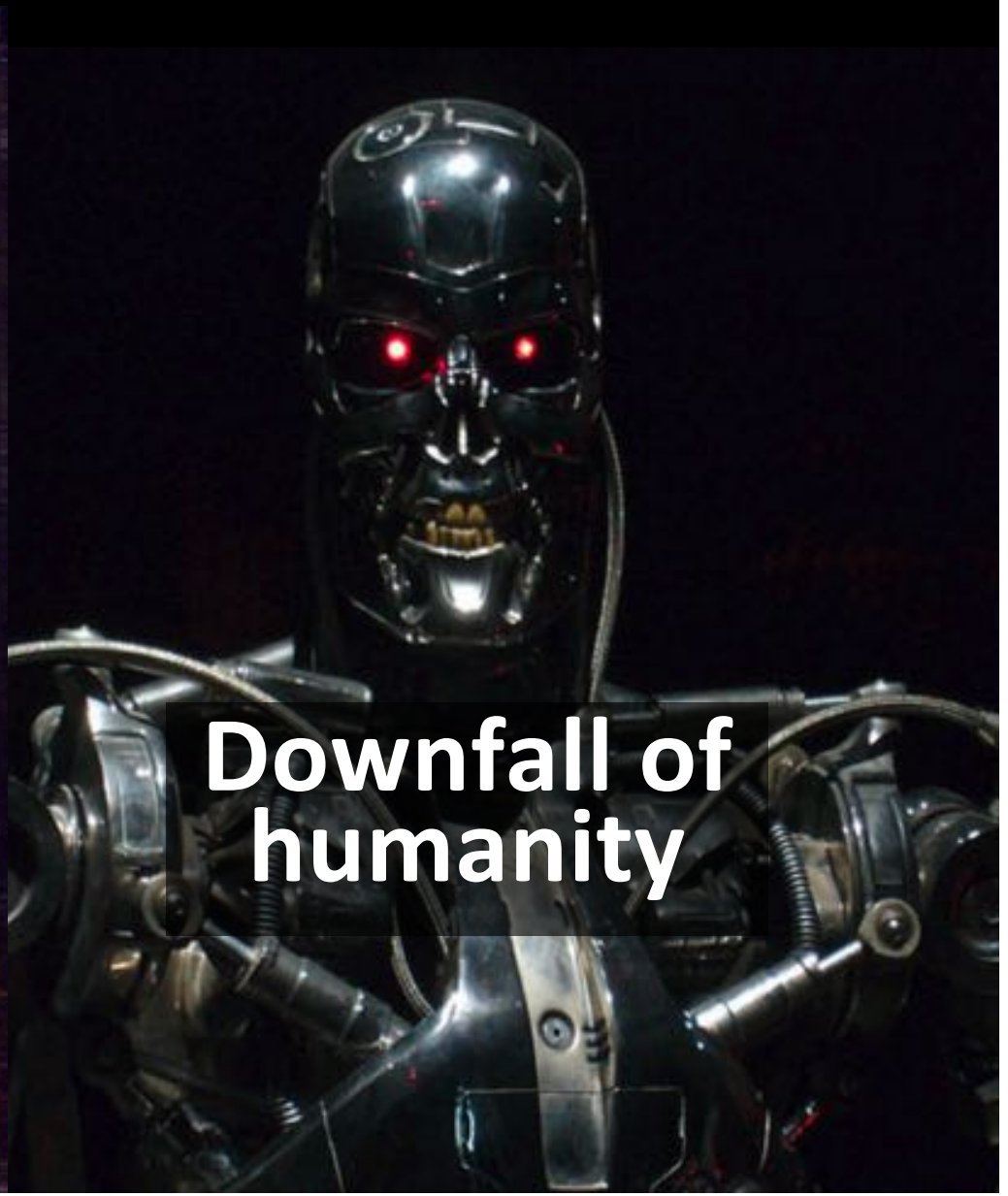
# AI today


the INQUIRER
Artificial intelligence will create the next industrial revolution, experts claim
Artificial intelligence better than scientists at choosing successful embryos
'We won't waste time on treatments that won't work, so the patient should get
says clinic director
BBC NEWS
Technology
Stephen Hawking warns artificial intelligence could end mankind
Telegraph
Lifestyle · Cars · News
Self-driving Tesla 'sav
by steering him to hos
V/S
Elon Musk
@elonmusk
I've talked to Mark about this. His understanding of the subject is limited.
SCIENTIFIC AMERICAN DECEMBER 2016
Computers Now Recognize Patterns Better Than Humans Can
An approach to artificial intelligence that enables computers to recognize visual

# AI today


THE ECONOMIC IMPACT OF
ARTIFICIAL INTELLIGENCE
NORTH AMERICA
$3.7 TR
NORTHERN EUROPE
$1.8 TR
CHINA
$7 TR
SOUTHERN EUROPE
LATIN AMERICA
REST OF WORLD
$1.2 TR
DEVELOPED ASIA
Projected Global Economic Effects of AI by 2030

# So, AI has many faces



# But, what exactly is AI?

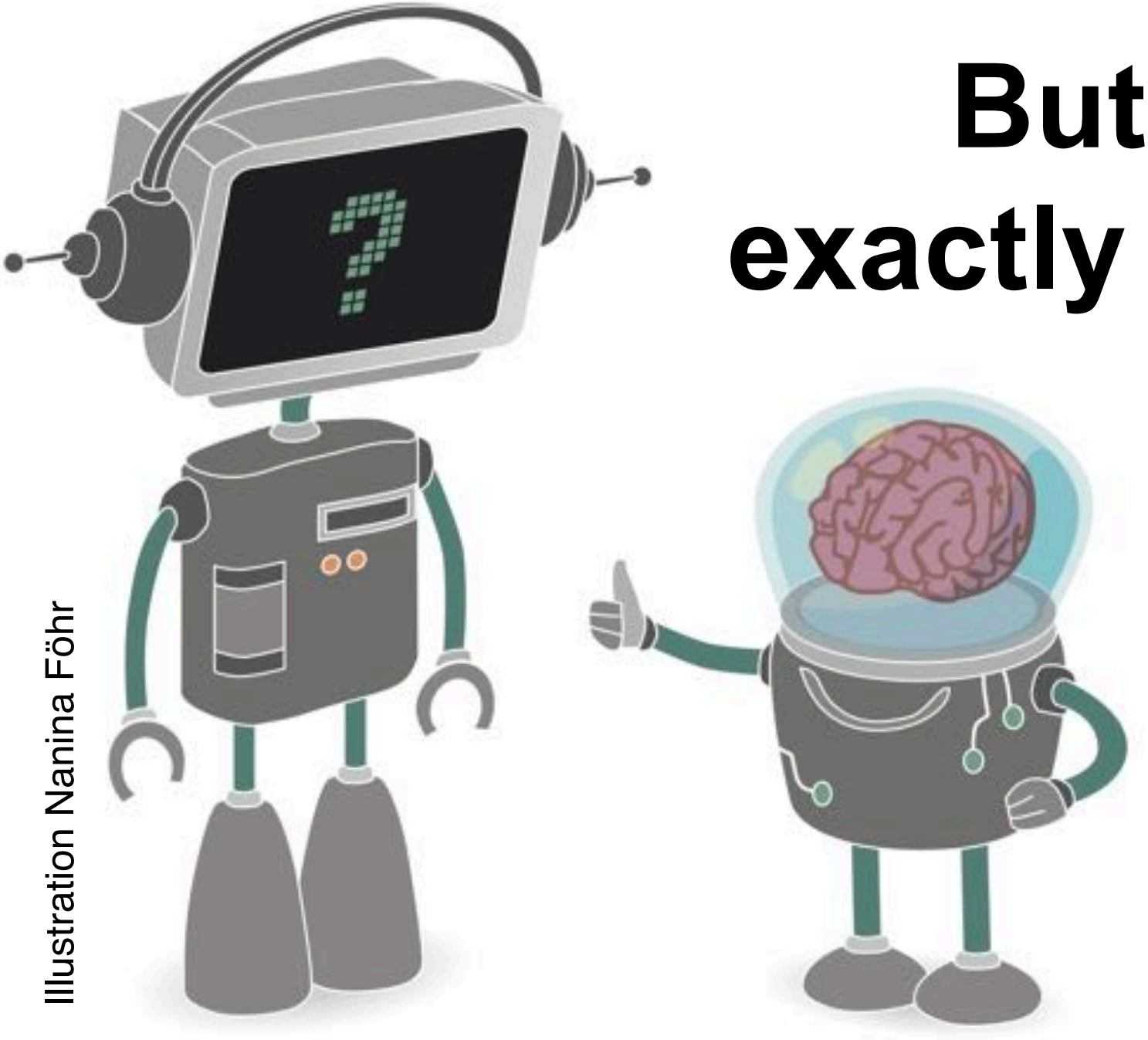
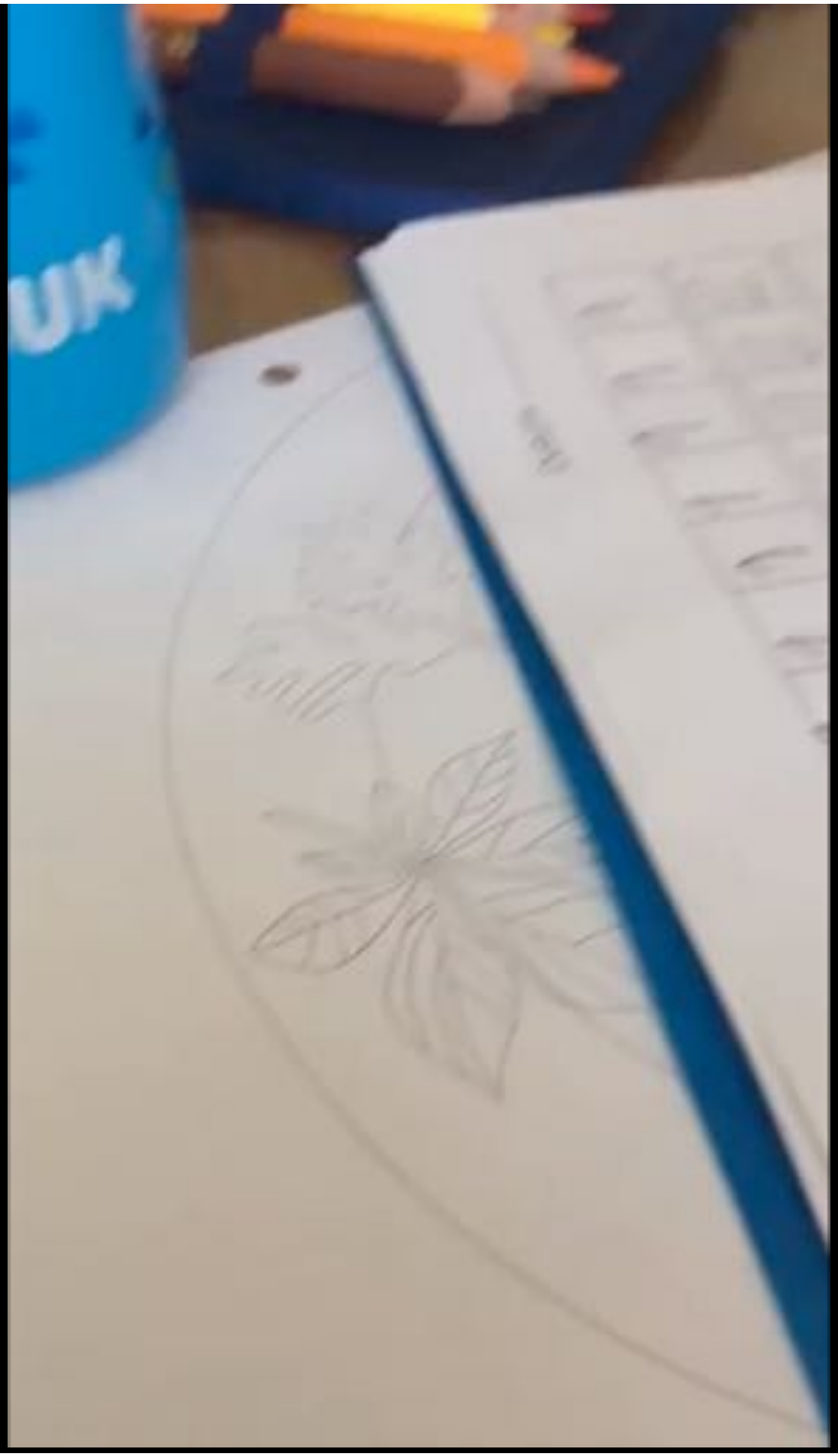


Illustration Nanina Föhr

# Humans are considered to be smart

https://www.youtube.com/watch?v=XQ79UUlOeWc

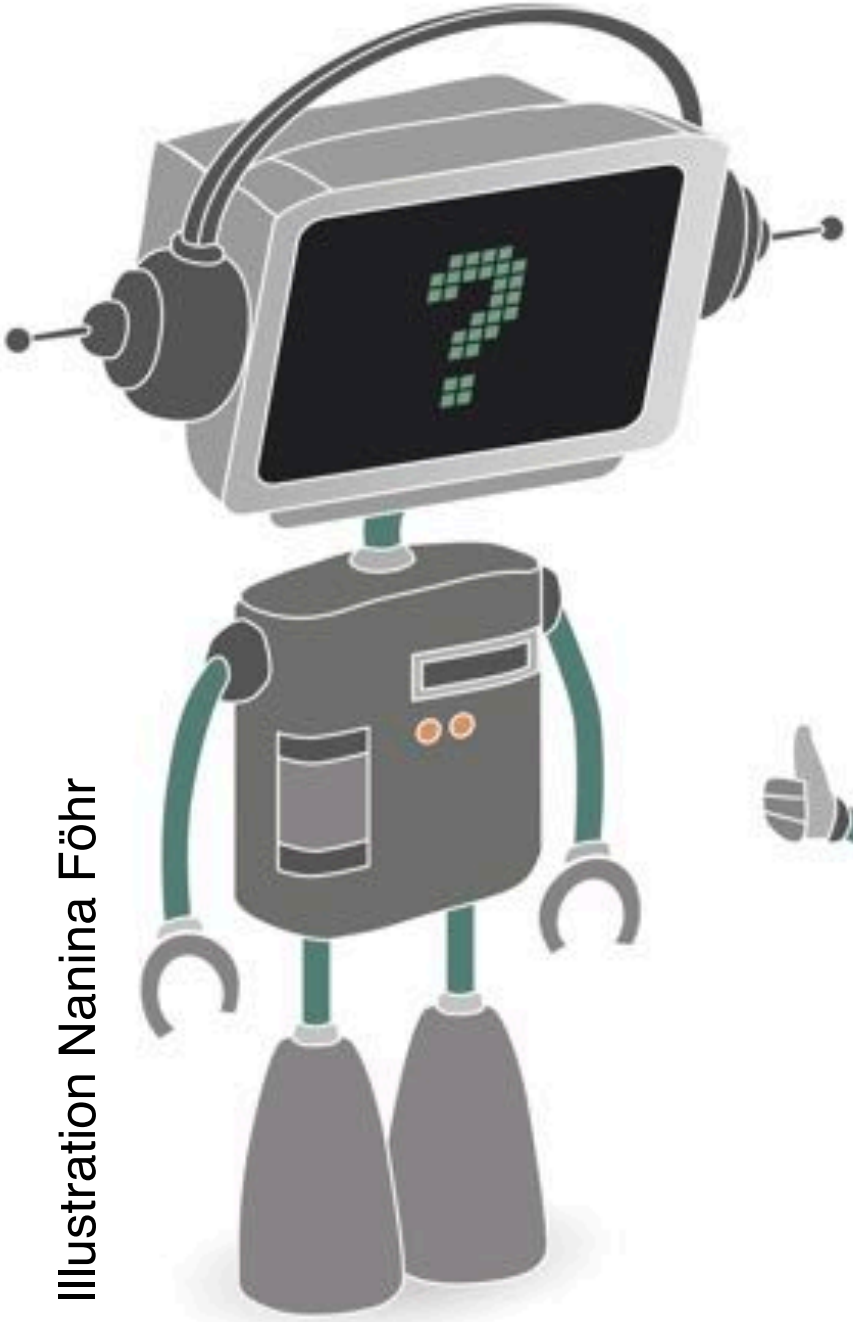


Are flies smart?

# What about orangutans?

Intelligence has many qualities.

It is difficult to directly capture/measure it.



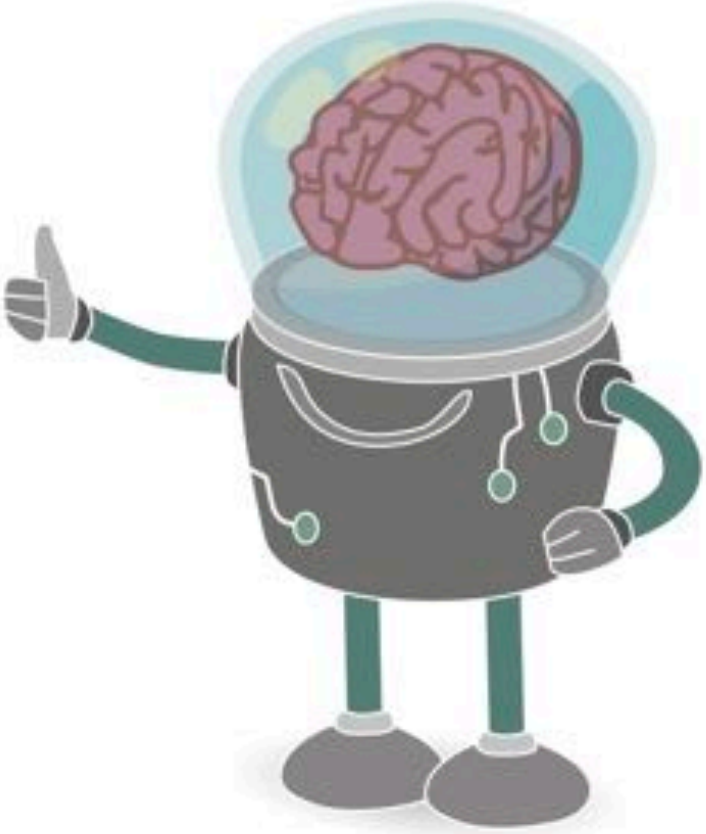


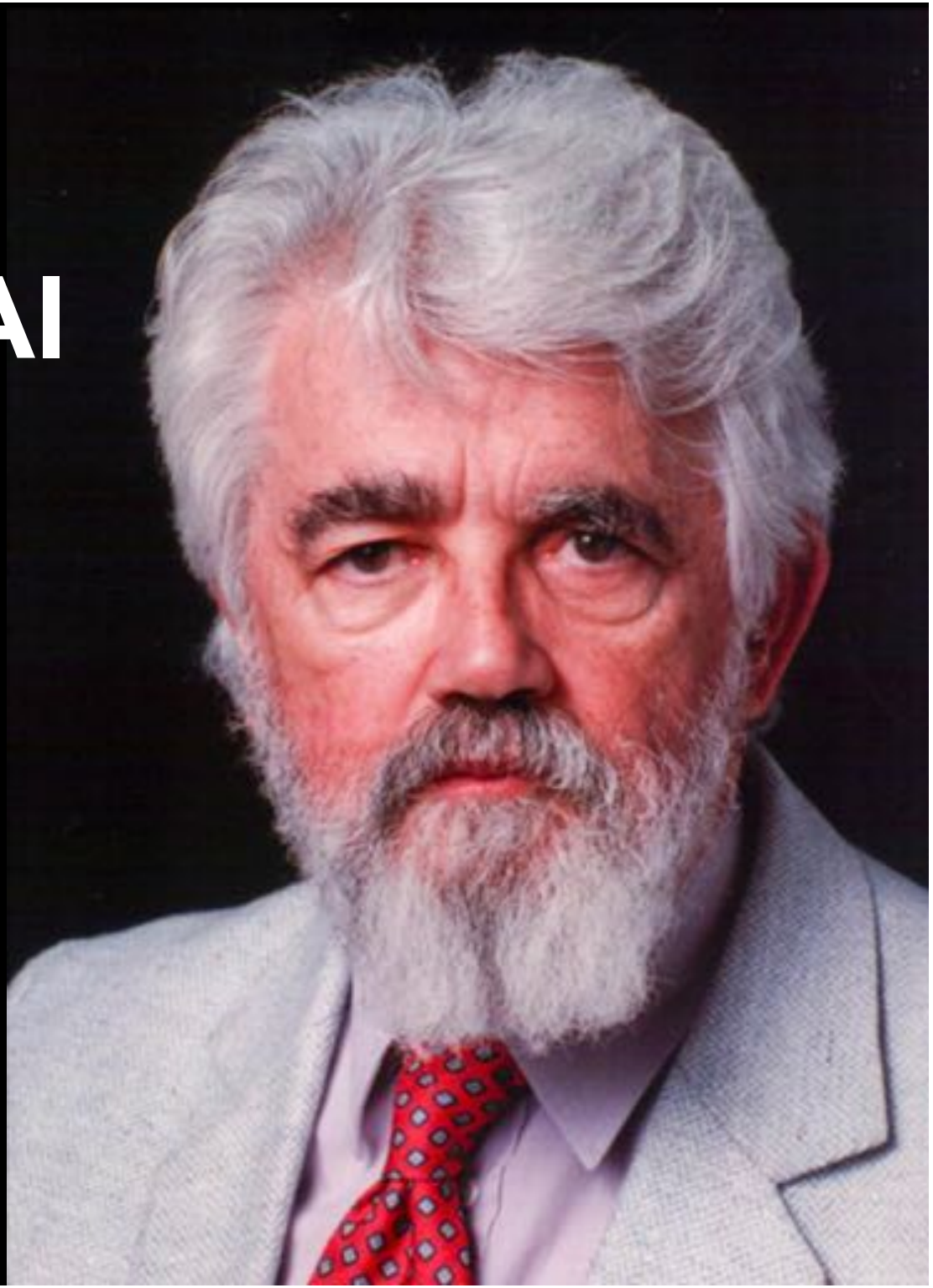
Illustration Nanina Föhr

# The Definition of AI

***„the science and engineering of making intelligent machines, especially intelligent computer programs.***

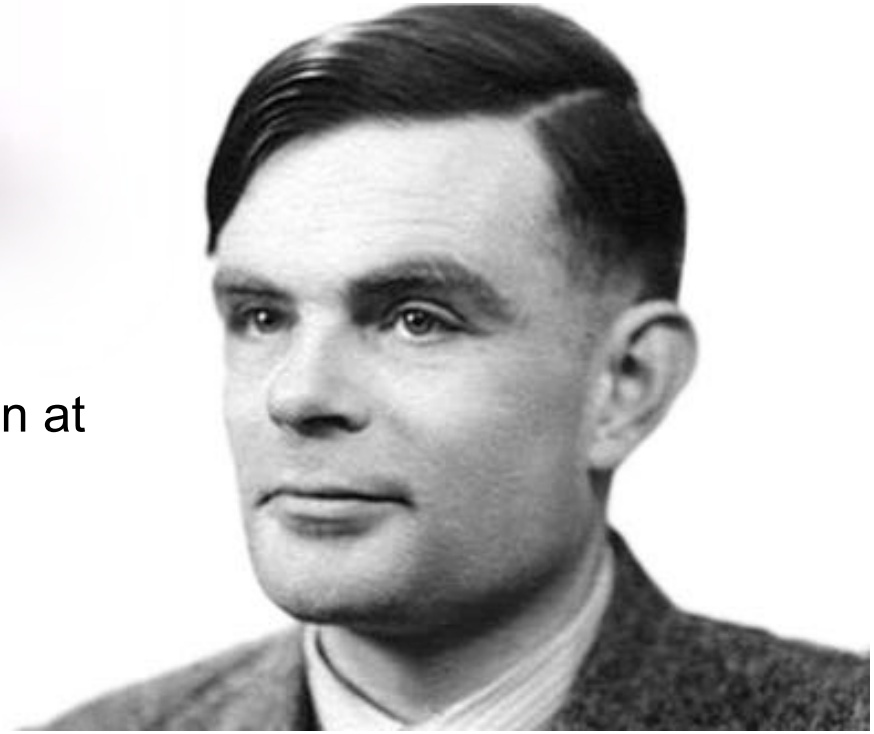
***It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable.“***

- John McCarthy, Stanford (1956), coined the term AI, Turing Awardee



# Turing Award =
# Nobel Prize for Computing

Named after Alan Turing, a British mathematician at the University of Manchester. Turing is often credited as being the key founder of theoretical computer science and AI.



**AI wants to build intelligent computer programs. How do we do this?**

**We use algorithms:**
unambiguous specifications of how to solve a class of problems – in finite time.

Always follow the right-hand path. If you reach a dead-end, go back to the last choice point and take the next unexplored path to the right.

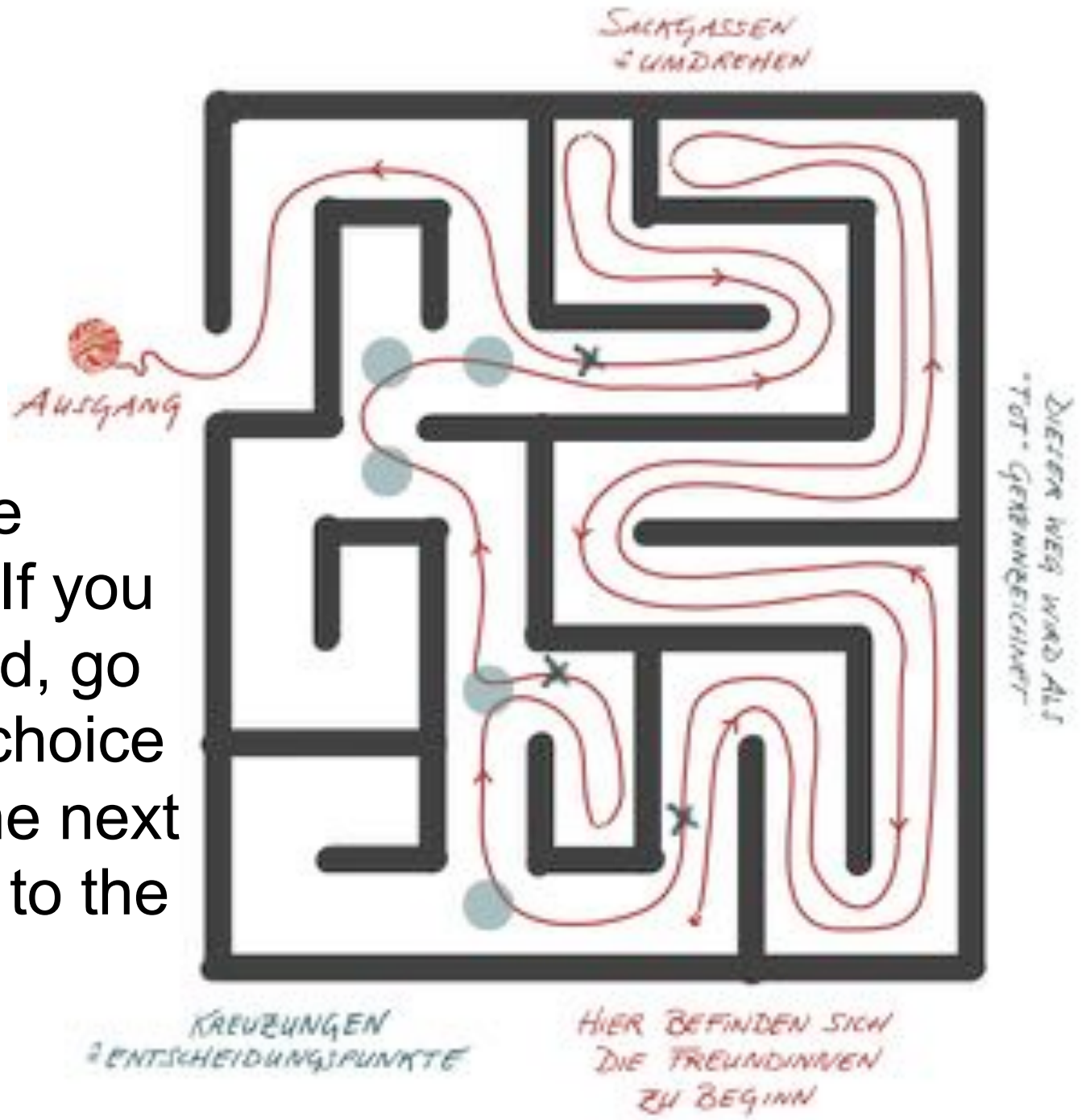
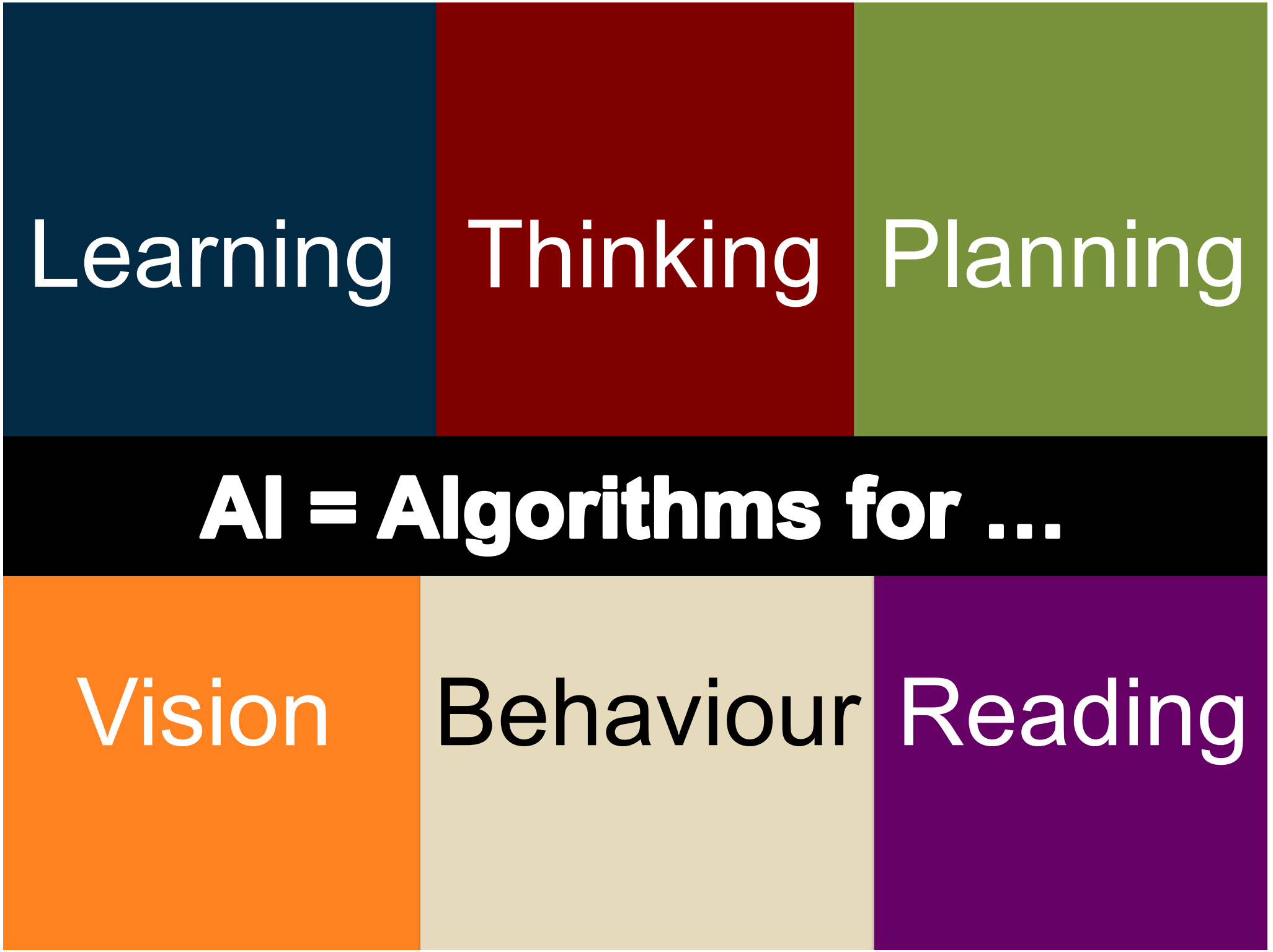



Illustration Nanina Föhr

# Think of it as a recipe!


Learning
Thinking
Planning
AI = Algorithms for ...
Vision
Behaviour
Reading

# Machine Learning

**the science "concerned with the question of how to construct computer programs that automatically improve with experience”**

- Tom Mitchell (1997) CMU

# Deep Learning

**a form of machine learning that makes use of artificial neural networks**

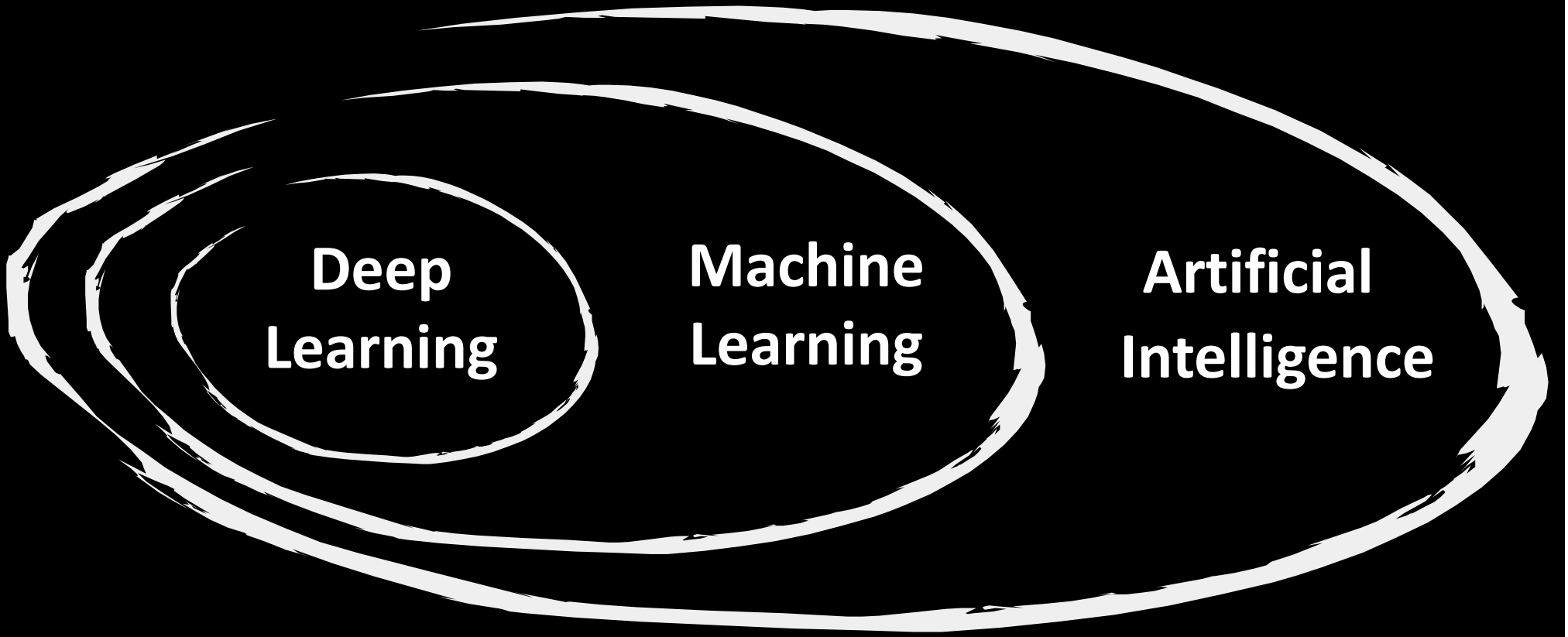
Geoffrey Hinton
Google
Univ. Toronto (CAN)

Yann LeCun
Facebook (USA)

Yoshua Bengio
Univ. Montreal (CAN)

Turing Awardees 2019

# Overall Picture



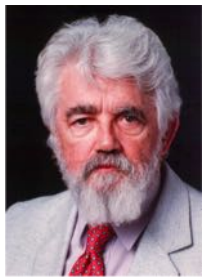
A closer look at
**the history of AI**

# 1956 Birth of AI

A Proposal for the

DARTMOUTH SUMMER RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE

We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer.



**John McCarthy**
Turing Award 1971

**Marvin Minsky**
Turing Award 1969

**Allen Newell**
Turing Award 1975

**Herbert A. Simon**
Turing Award 1975
Nobel Prize 1978

## … and of Cognitive Science

# Artificial Neural Networks

**Learning representations by back-propagating errors**

**David E. Rumelhart*, Geoffrey E. Hinton† & Ronald J. Williams***

* Institute for Cognitive Science, C-015, University of California, San Diego, La Jolla, California 92093, USA
† Department of Computer Science, Carnegie-Mellon University, Pittsburgh, Philadelphia 15213, USA

COGNITIVE SCIENCE **14,** 179–211 (1990)

**Finding Structure in Time**

JEFFREY L. ELMAN
*University of California, San Diego*

COGNITIVE SCIENCE 9, 147–169 (1985)

**A Learning Algorithm for Boltzmann Machines***

DAVID H. ACKLEY
GEOFFREY E. HINTON
*Computer Science Department*
*Carnegie-Mellon University*
TERRENCE J. SEJNOWSKI
*Biophysics Department*
*The Johns Hopkins University*

Biological Cybernetics
© by Springer-Verlag 1980
Biol. Cybernetics 36, 193–202 (1980)

**Neocognitron: A Self-organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position**

Kunihiko Fukushima
NHK Broadcasting Science Research Laboratories, Kinuta, Setagaya, Tokyo, Japan

*Psychological Review*
Vol. 65, No. 6, 1958

THE PERCEPTRON: A PROBABILISTIC MODEL FOR INFORMATION STORAGE AND ORGANIZATION IN THE BRAIN[1]

F. ROSENBLATT
*Cornell Aeronautical Laboratory*

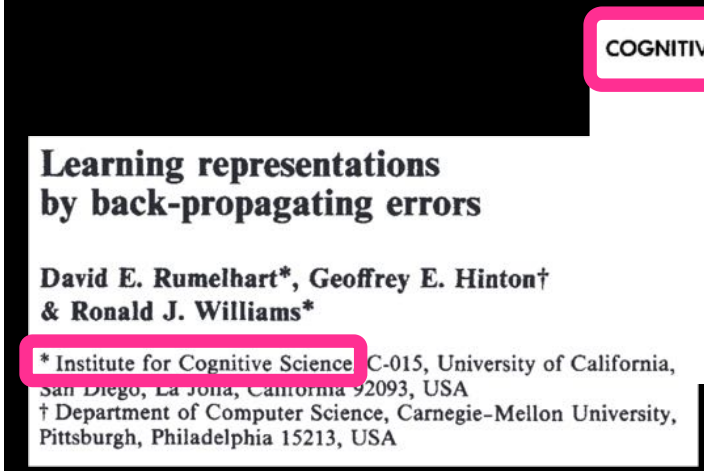
Psychological Review
1981, Vol. 88, No. 2, 135–170

Copyright 1981 by the American Psychological Association, Inc.
0033-295X/81/8802-0135$00.75

Toward a Modern Theory of Adaptive Networks: Expectation and Prediction

Richard S. Sutton and Andrew G. Barto
Computer and Information Science Department
University of Massachusetts—Amherst

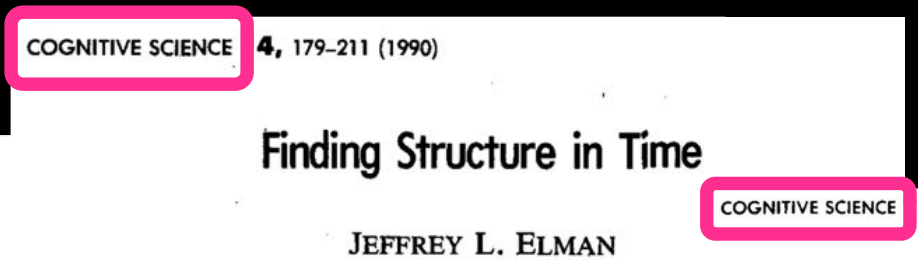
# Artificial Neural Networks



Learning representations by back-propagating errors

David E. Rumelhart*, Geoffrey E. Hinton† & Ronald J. Williams*

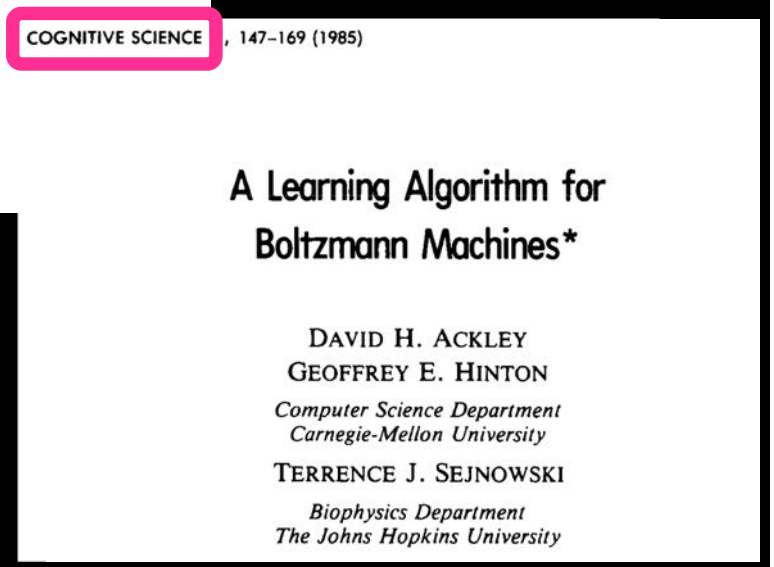
* Institute for Cognitive Science, C-015, University of California, San Diego, La Jolla, California 92093, USA
† Department of Computer Science, Carnegie-Mellon University, Pittsburgh, Philadelphia 15213, USA



COGNITIVE SCIENCE 4, 179–211 (1990)

Finding Structure in Time

JEFFREY L. ELMAN
*University of California, San Diego*



COGNITIVE SCIENCE, 147–169 (1985)

A Learning Algorithm for Boltzmann Machines*

DAVID H. ACKLEY
GEOFFREY E. HINTON
*Computer Science Department*
*Carnegie-Mellon University*
TERRENCE J. SEJNOWSKI
*Biophysics Department*
*The Johns Hopkins University*

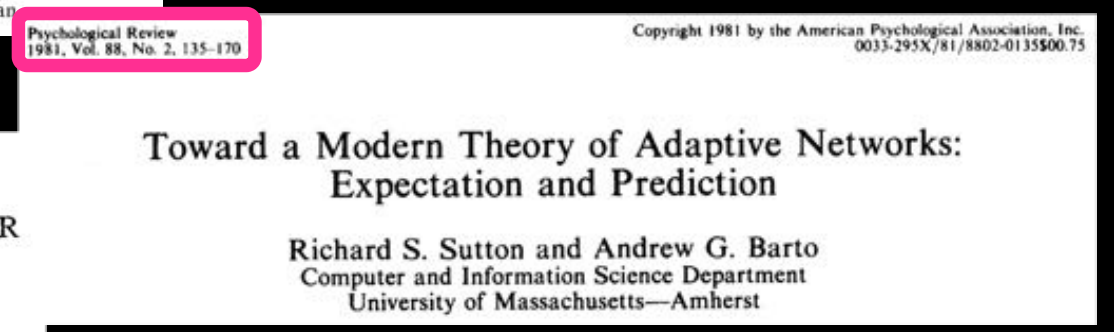
Biological Cybernetics

Biol. Cybernetics 36, 193–202 (1980)

Neocognitron: A Self-organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position

Kunihiko Fukushima

NHK Broadcasting Science Research Laboratories, Kinuta, Setagaya, Tokyo, Japan

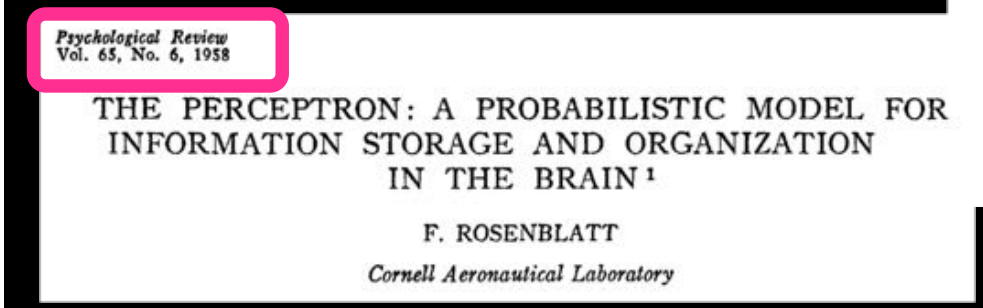


Psychological Review 1981, Vol. 88, No. 2, 135–170

Copyright 1981 by the American Psychological Association, Inc.
0033-295X/81/8802-0135$00.75

Toward a Modern Theory of Adaptive Networks: Expectation and Prediction

Richard S. Sutton and Andrew G. Barto
Computer and Information Science Department
University of Massachusetts—Amherst



*Psychological Review*
Vol. 65, No. 6, 1958

THE PERCEPTRON: A PROBABILISTIC MODEL FOR INFORMATION STORAGE AND ORGANIZATION IN THE BRAIN[1]

F. ROSENBLATT
*Cornell Aeronautical Laboratory*

slide after C. Rothkopf (TUD), after J.Tenenbaum (MIT)

# Algorithms of intelligent behaviour teach us a lot about ourselves

## The twin science: cognitive science

"How do we humans get so much from so little?" and by that I mean how do we acquire our understanding of the world given what is clearly by today's engineering standards so little data, so little time, and so little energy.

**Centre for Cognitive Science at TU Darmstadt**

Establishing cognitive science at the Technische Universität Darmstadt is a long-term commitment across multiple departments (see Members to get an impression on the interdisciplinary of the supporting groups and departments). The TU offers a strong foundation including several established top engineering groups in Germany, a prominent computer science department (which is among the top four in Germany), a

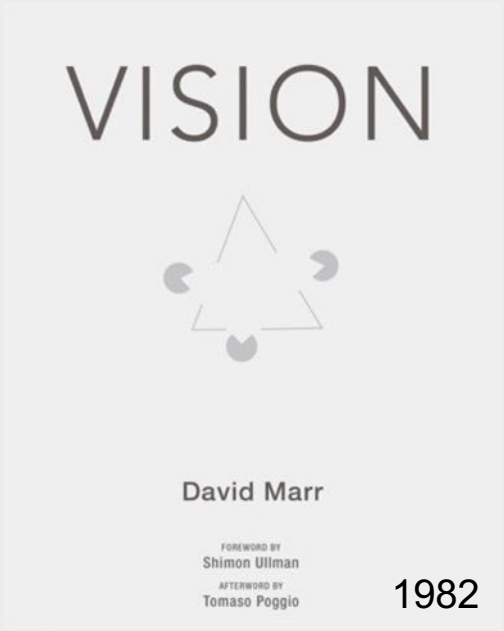
Centre for Cognitive Science

Josh Tenenbaum, MIT

Lake, Salakhutdinov, Tenenbaum, Science 350 (6266), 1332-1338, 2015

Tenenbaum, Kemp, Griffiths, Goodman, Science 331 (6022), 1279-1285, 2011

# Three levels of description



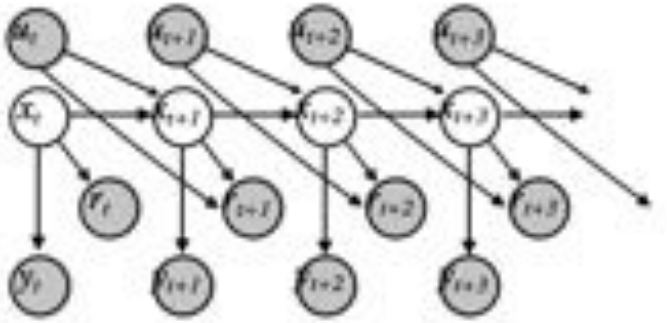

1982

**Computational**
Why do things work the way they work? What is the goal of the computation? What are the unifying principles?

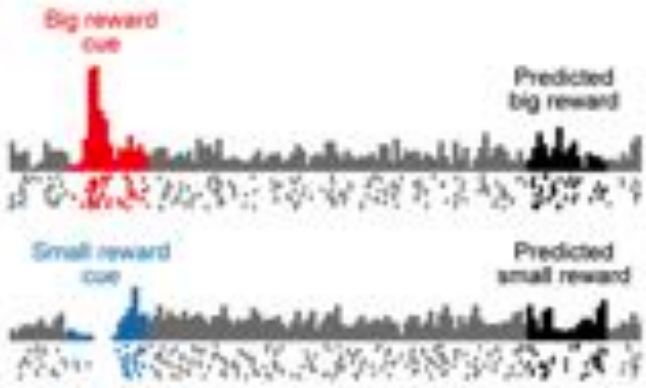
*maximize:*

$$R_t = r_{t+1} + r_{t+2} + \cdots + r_T$$

**Algorithmic**
What represetation can implement such computations? How does the choice of the representation determine the algorithm



**Implementational**
How can such a system be built in hardware? How can neurons carry out the computations?




slide after C. Rothkopf (TUD)

# Artificial Neural Networks

**Inspiration from the brain:**

- **many small interconnected units (neurons)**
- **learning happens by changing the strength of connections (synapses)**
- **behavior of the whole is more than the sum of the parts**

Frank Rosenblatt (1928-1971)

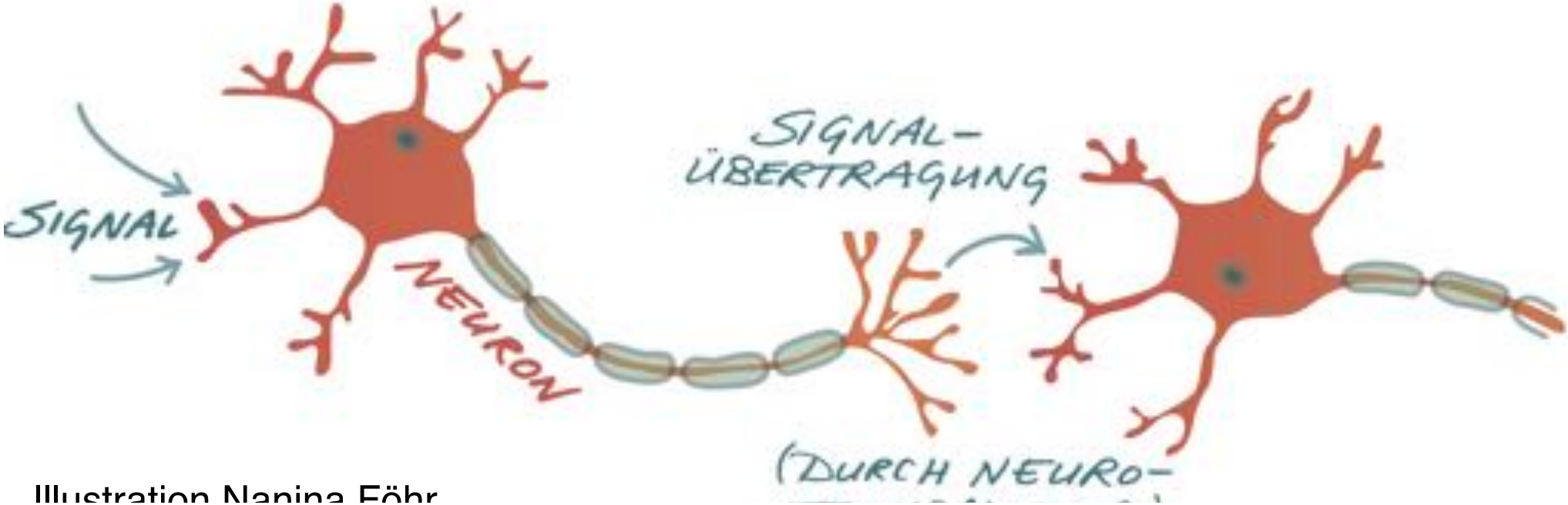



Illustration Nanina Föhr

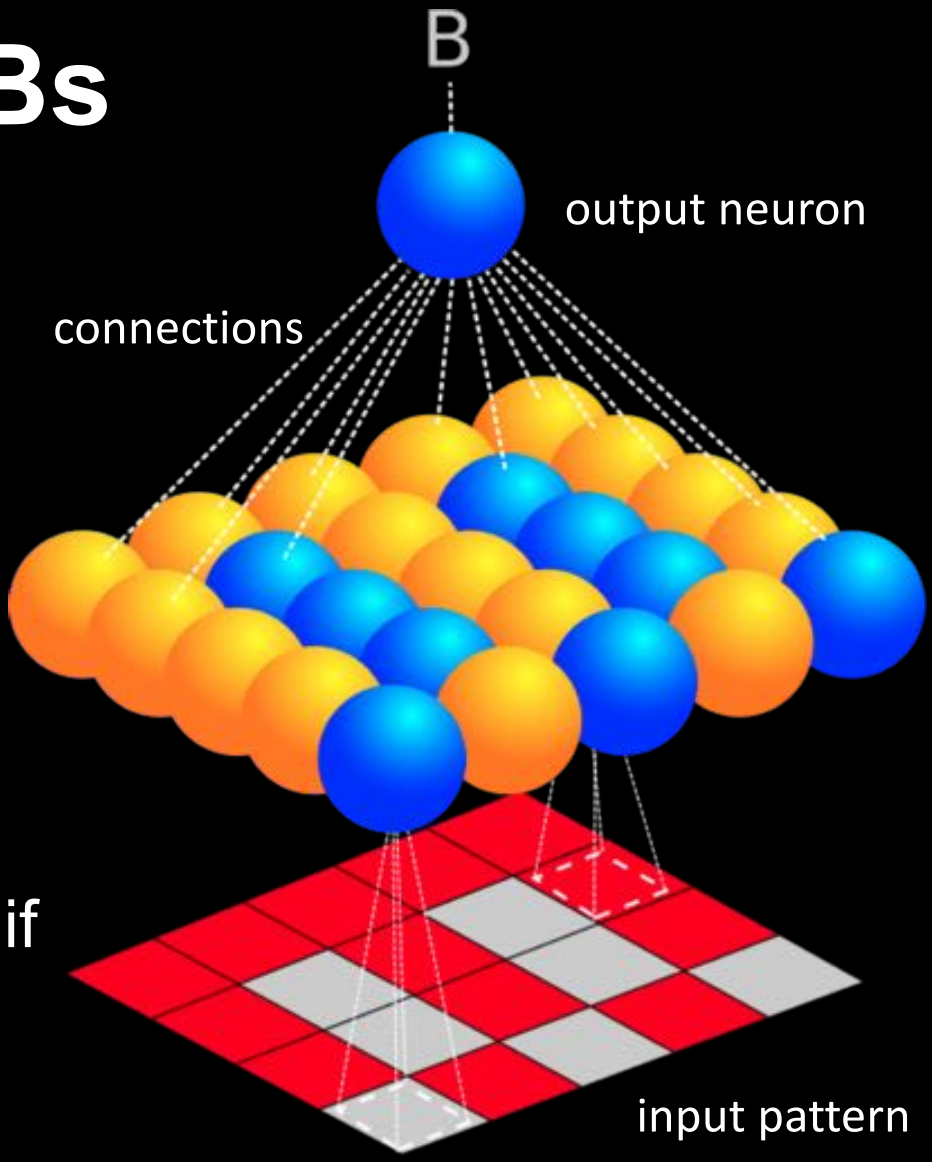
# The Perceptron to distinguish As an Bs

1) present pattern

2) some first layer neurons spike

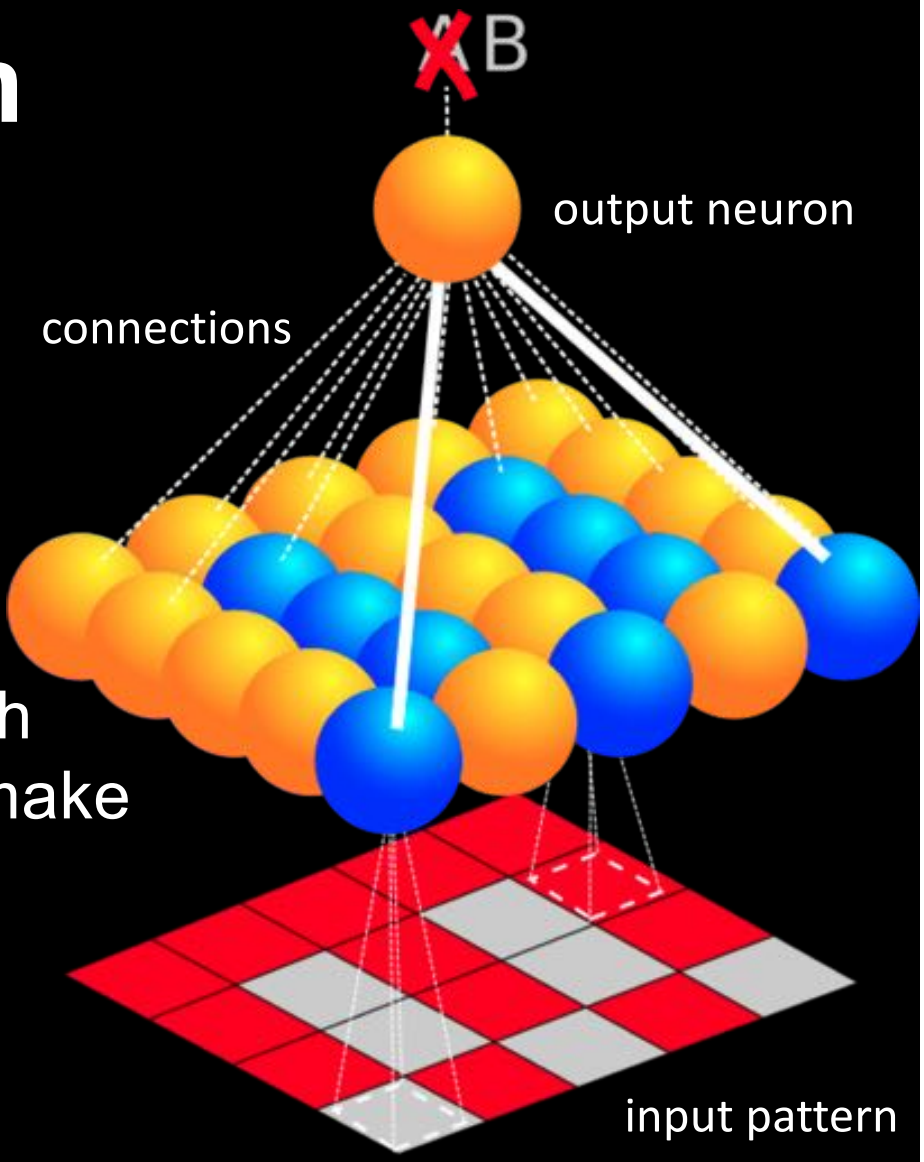
3) output neuron accumulates signals from previous layer; if it is above a threshold, the output neuron spikes and predicts an A; if not, then it does not spikes and predicts a b

4) prediction is “B”



# The Perceptron Learning Algorithm

1) present pattern
2) wait for output to be produced
3) if output correct
    - change nothing
4) if output incorrect:
    - adjust connection strength (positive or negative) to make the pattern be classified correctly
5) repeat until no more errors



# Artificial Neural Networks
# = Stacking of many artificial neurons

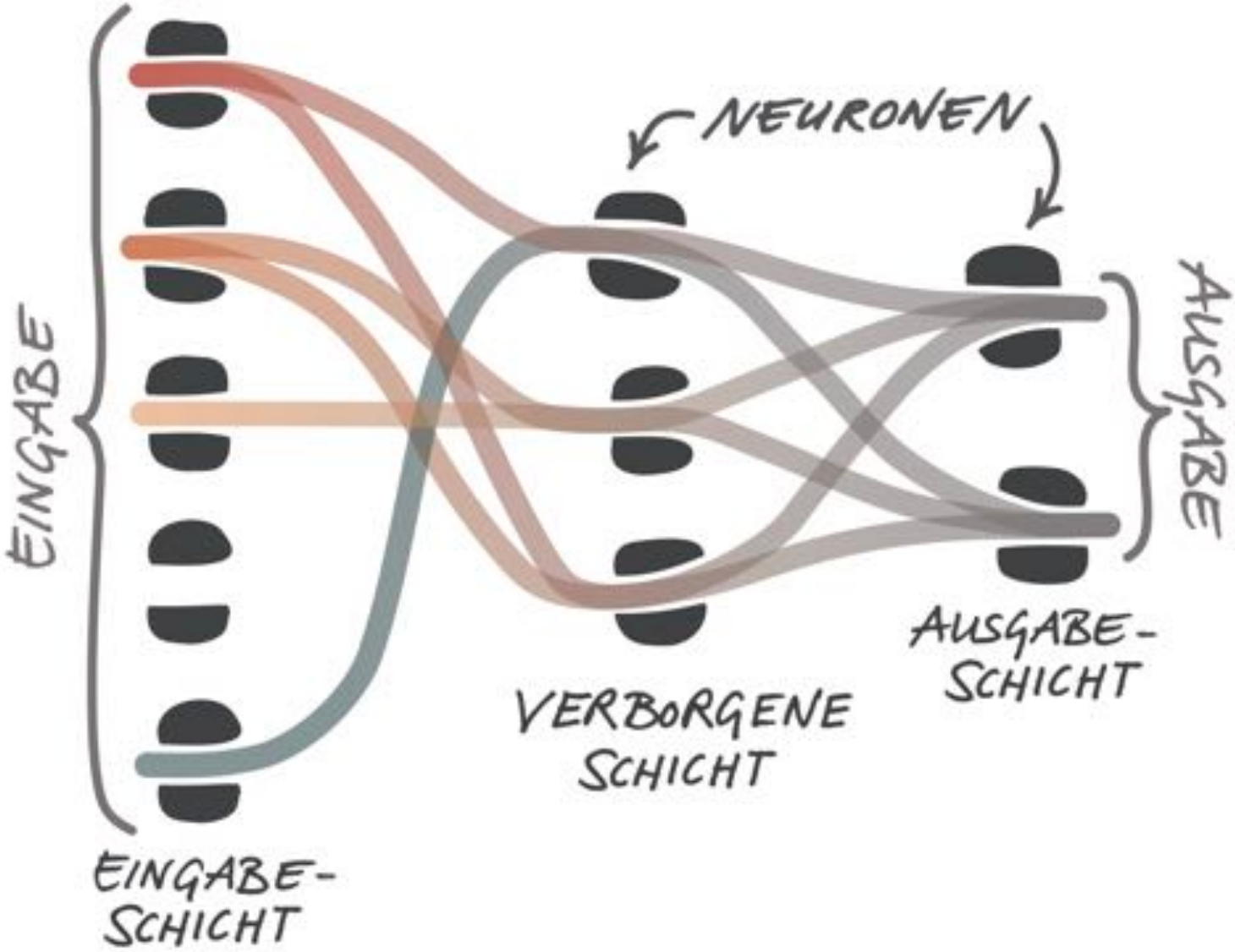
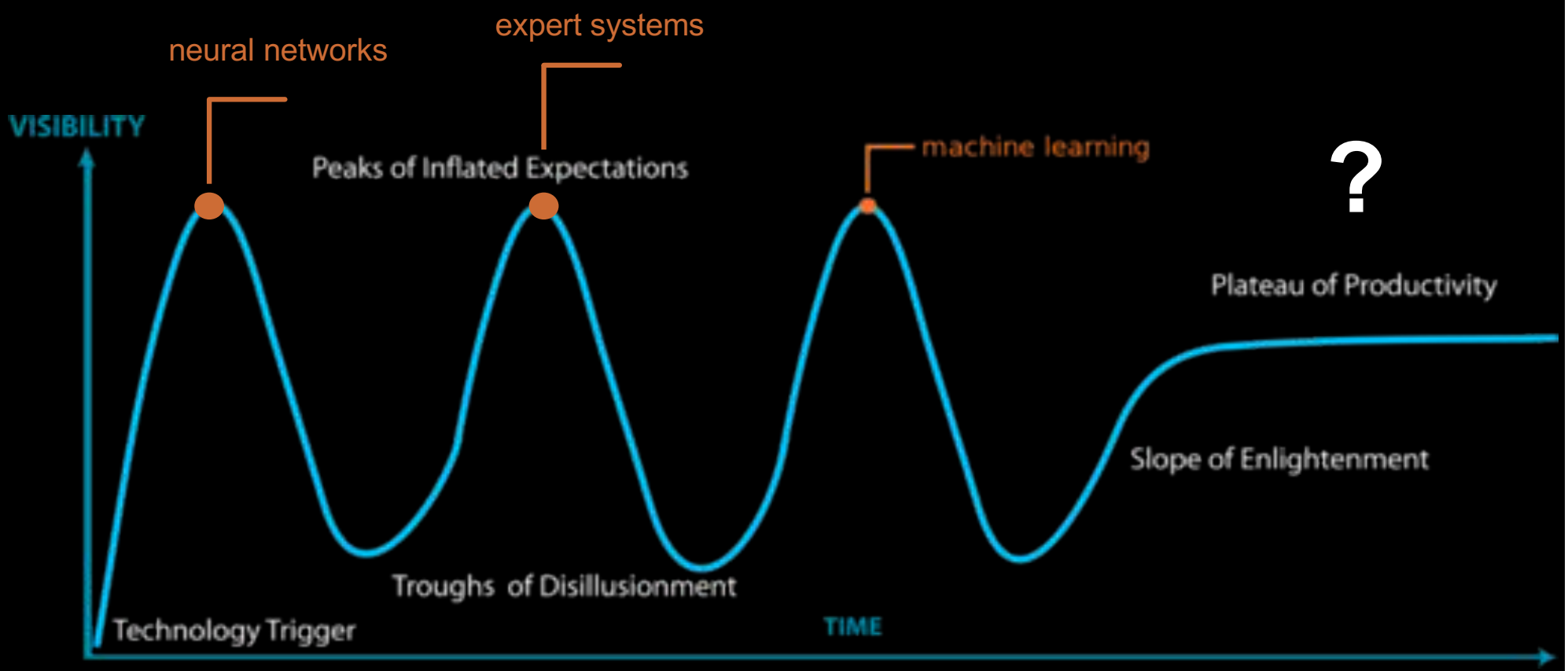


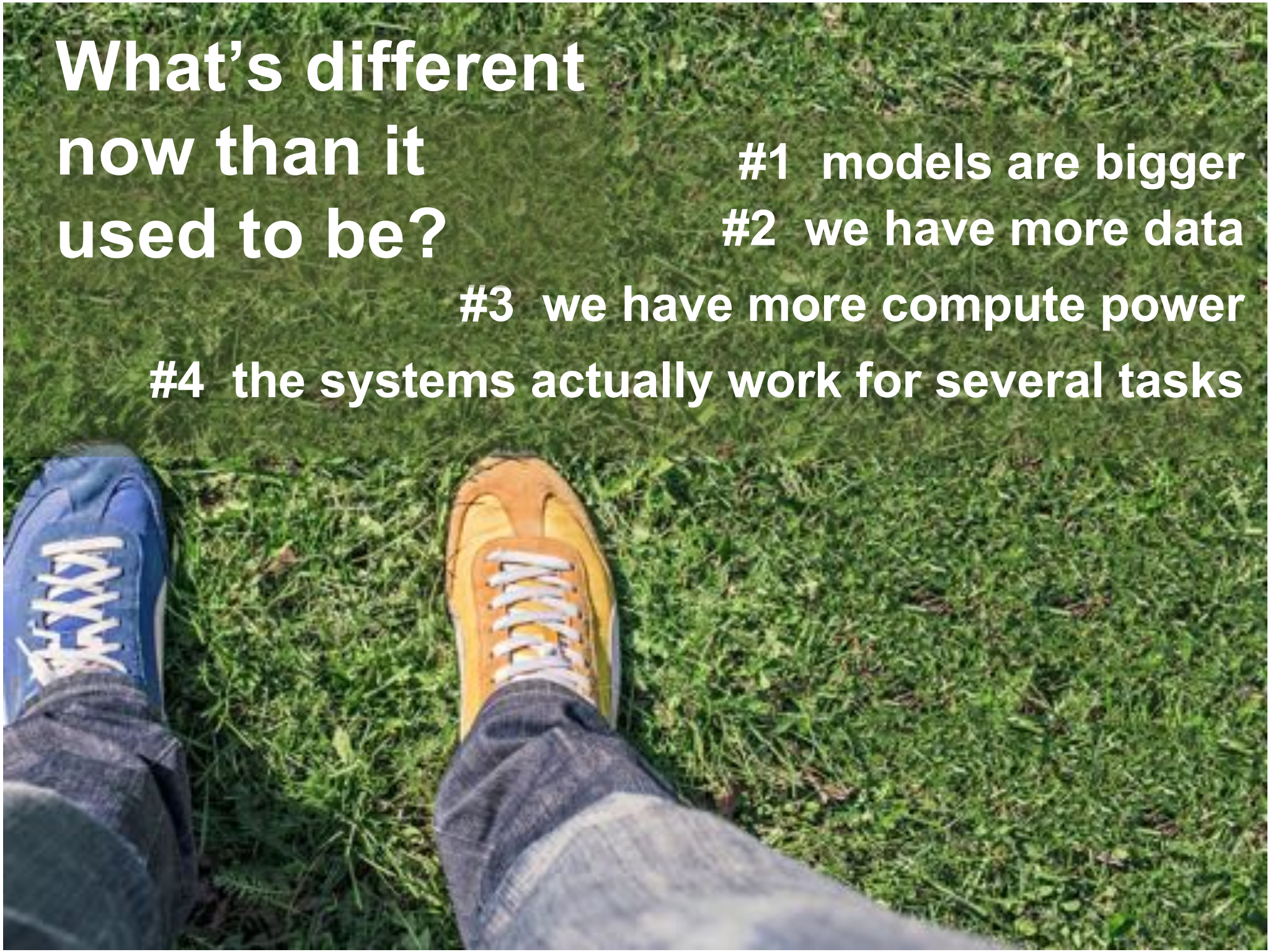

Illustration Nanina Föhr

# The history of AI in a nutshell



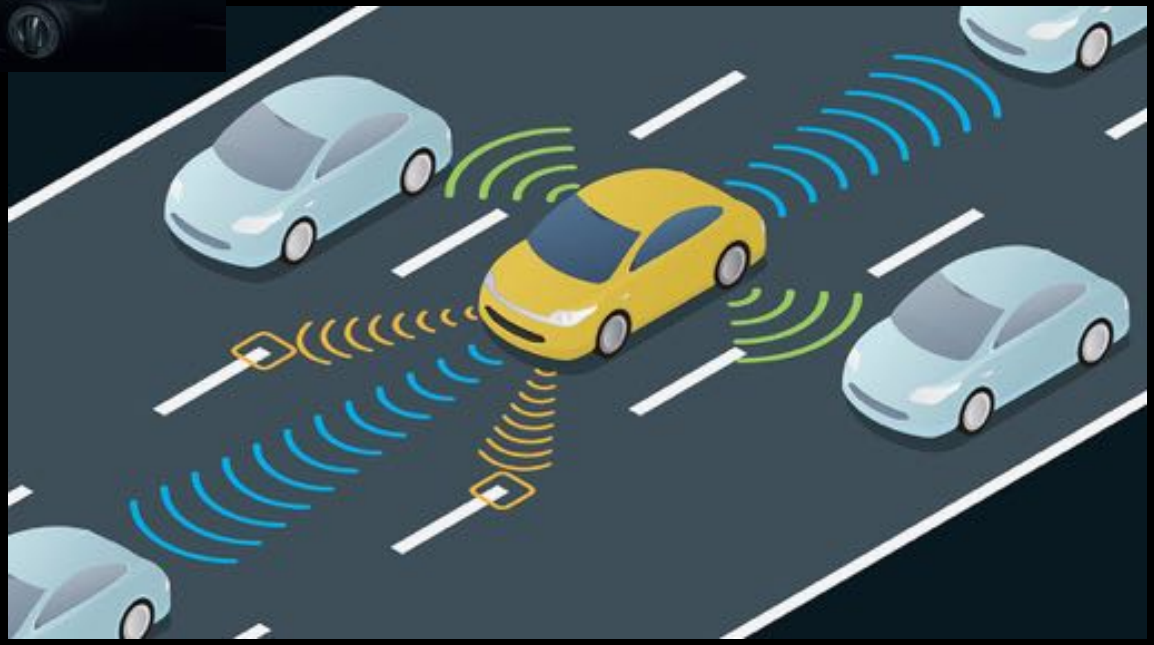

1956

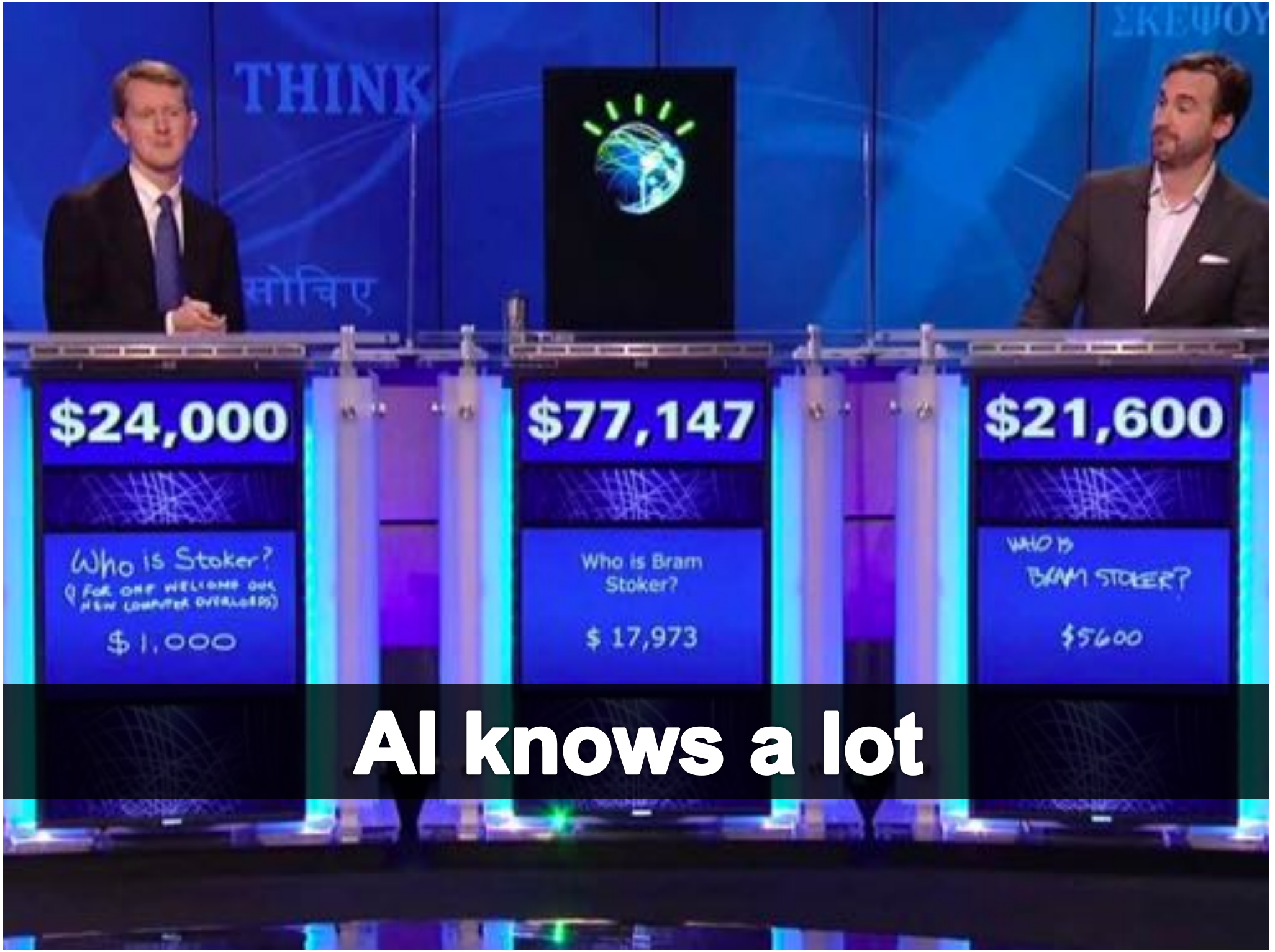
2019

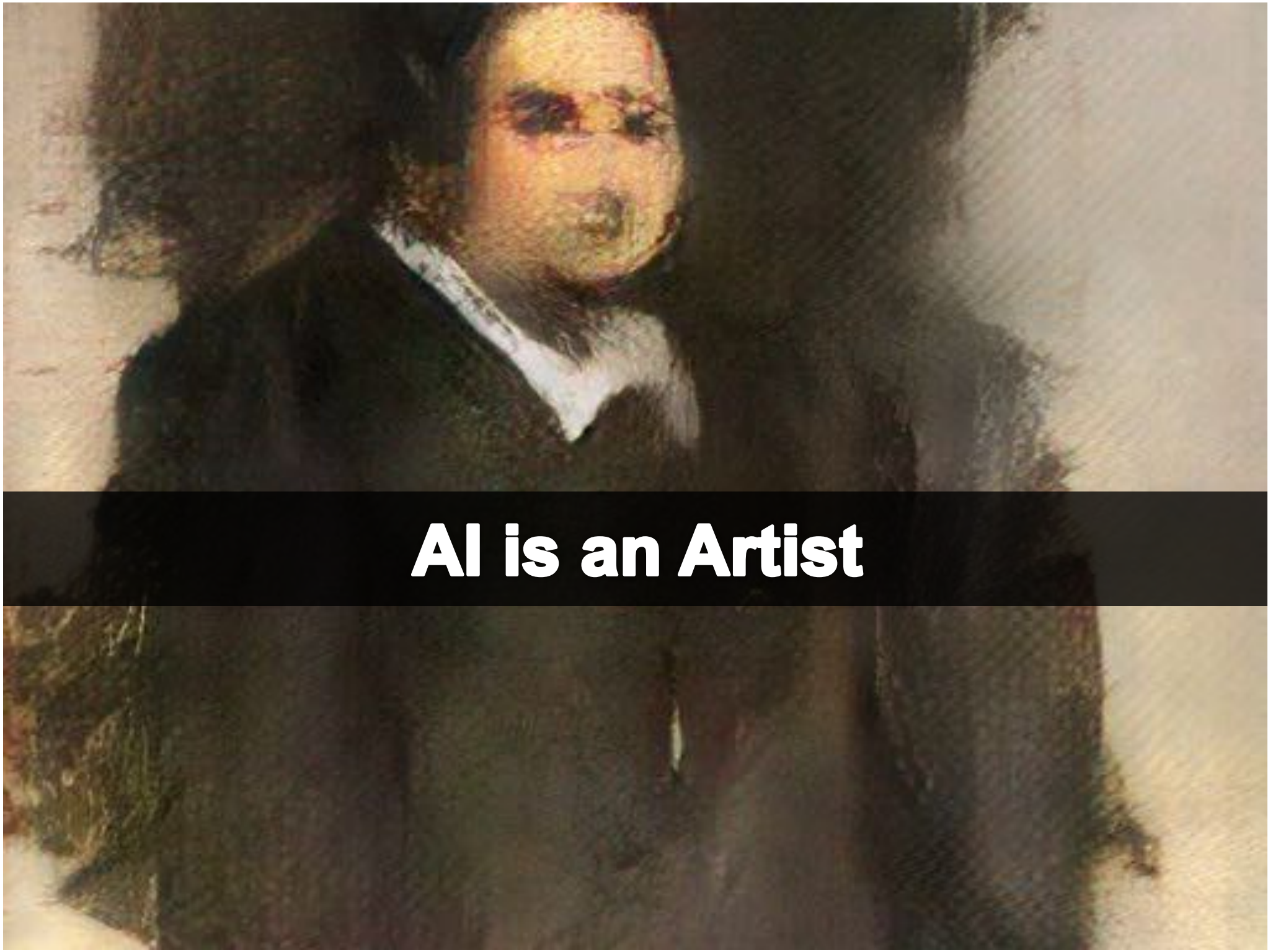

What's different now than it used to be?
#1 models are bigger
#2 we have more data
#3 we have more compute power
#4 the systems actually work for several tasks

AI does the
laundry

# AI drives cars




THINK
$24,000
Who is Stoker?
$ 1,000
$77,147
Who is Bram Stoker?
$ 17,973
$21,600
WHO IS
BRAM STOKER?
$5600
AI knows a lot


AI is an Artist

TECHNISCHE UNIVERSITÄT DARMSTADT

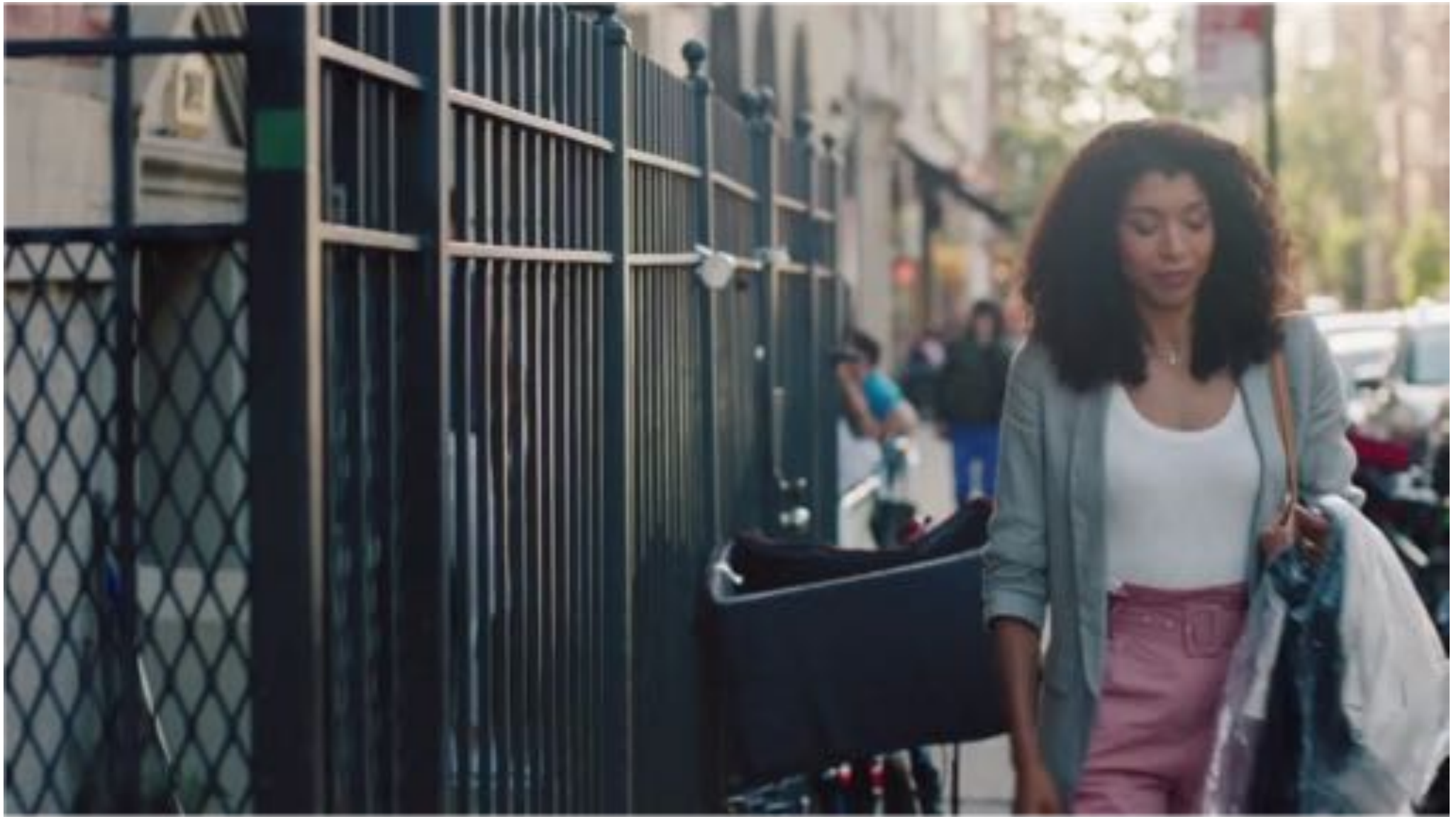
**Schachmatt durch „CrazyAra"** 19.02.2019

**Künstliche Intelligenz schlägt mehrfachen Weltmeister im Einsetzschach**

**Der von den TU-Studierenden Johannes Czech, Moritz Willig und Alena Beyer entwickelte Bot „CrazyAra" hat den Schachprofi Justin Tan in einem Online-Match der Schach-Variante „Crazyhouse" mit 4:1 geschlagen. Gelernt hat der Bot mittels künstlicher neuronaler Netze, was ihm erlaubt, vorausschauend Entscheidungen zu treffen. Das Besondere: Die Studierenden konnten damit einen Erfolg auf einem Feld feiern, das sonst von Giganten wie Google dominiert wird.**

# AI plays chess and GO

# AI assists you



# But

# But

The New York Times

Opinion

## A.I. Is Harder Than You Think

By Gary Marcus and Ernest Davis

Mr. Marcus is a professor of psychology and neural science. Mr. Davis is a professor of computer science.

May 18, 2018

# AI has many isolated talents

# AI is not superhuman

DARPA challenge (2015)

# AI is not superhuman

And this also holds as of today

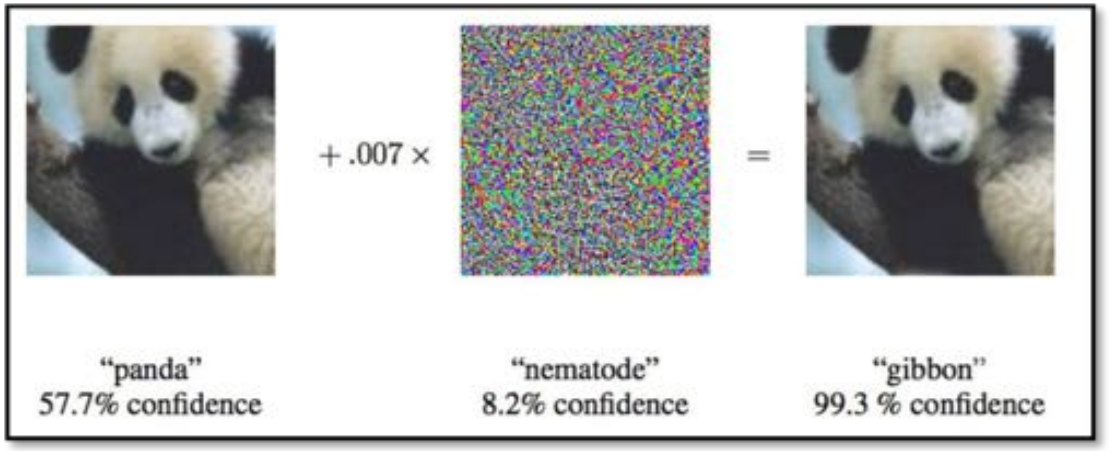
# Fundamental Differences

as of today

# Fundamental Differences

Sharif et al., 2015

Brown et al. (2017)



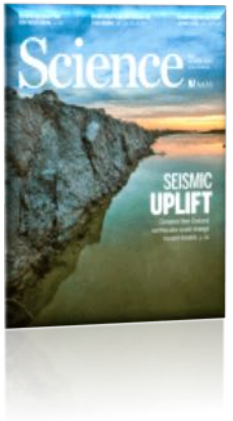

Google, 2015

REPORTS | PSYCHOLOGY

## Semantics derived automatically from language corpora contain human-like biases

Aylin Caliskan[1,*], Joanna J. Bryson[1,2,*], Arvind Narayanan[1,*]
+ See all authors and affiliations

*Science* 14 Apr 2017:
Vol. 356, Issue 6334, pp. 183-186
DOI: 10.1126/science.aal4230



My team and I in the **Machine Learning** lab together with our colleagues at the Centre for Cognitive Science would like to make computers learn so much about the world, so rapidly and flexibly, as humans.

TECHNISCHE UNIVERSITÄT DARMSTADT

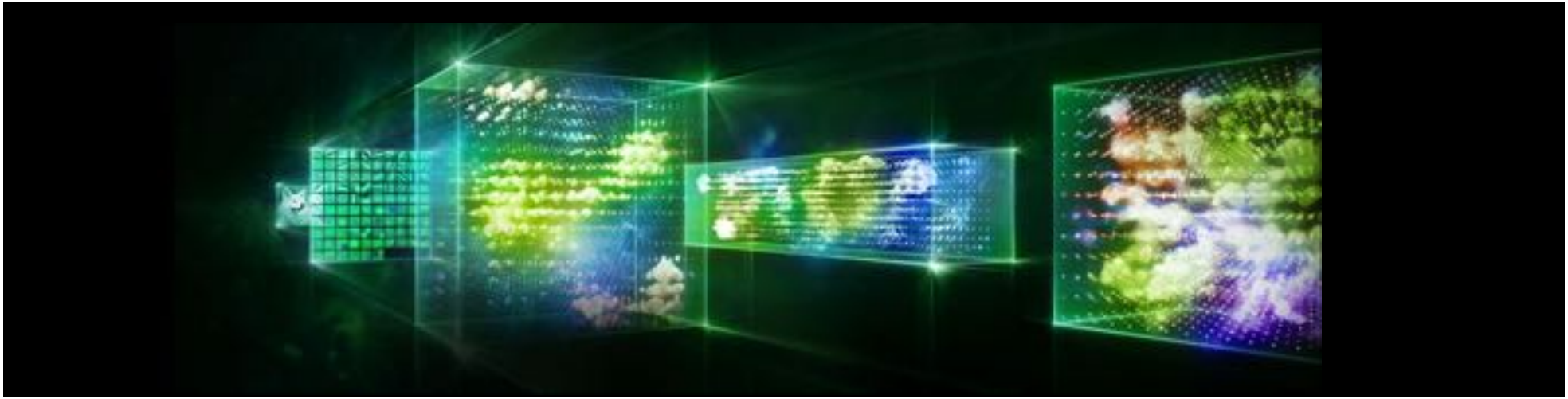
Centre for Cognitive Science

Fachbereich Informatik

Lernende Systeme
DIE PLATTFORM FÜR KÜNSTLICHE INTELLIGENZ

Federal Ministry of Education and Research

| | |
|---|---|
| 2017 - now: | Professor (W3) for Machine Learning at the CS Department of the TU Darmstadt, Germany |
| 2013 - 2017: | Associate Professor (W2) for Data Mining at the CS Department of the TU Dortmund University, Germany |
| 2012 - 2013: | Assistant Professor (W1) for Spatio-Temporal Pattern in Agriculture at the Faculty of Agriculture of the University of Bonn, Germany |
| 2008 - 2012: | Fraunhofer Attract research group leader at the Fraunhofer IAIS, Germany |
| 2007: | PostDoctoral Associate at MIT CSAIL, USA, working with Leslie Kaelbling, Josh Tenenbaum, and Nicholas Roy. |
| 2000 - 2006: | Ph.D. student at the CS Department of the University of Freiburg, Germany, working with Luc De Raedt (supervsior) and Wolfram Burgard. |
| 1996 - 2000: | Diploma in Computer Science at the CS Department of the University of Freiburg, Germany |



**Deep Probabilistic Learning**

UBER AI Labs

**Many neural networks have no probabilistic semantics. They are not calibrated joint distributions.** **$P(Y|X) \neq P(Y,X)$**

**MNIST**

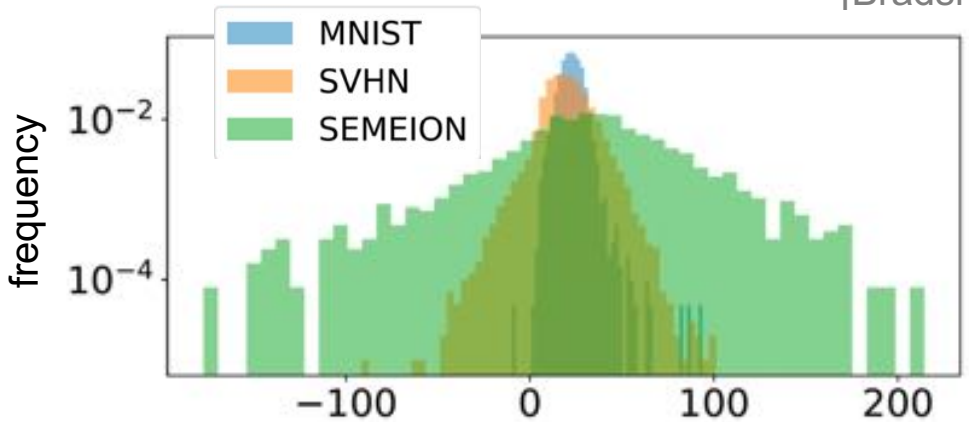
**SVHN**

**SEMEION**

**Train & Evaluate**

**Transfer Testing**

[Bradshaw et al. arXiv:1707.02476 2017]




**Many neural etworks cannot distinguish the datasets**

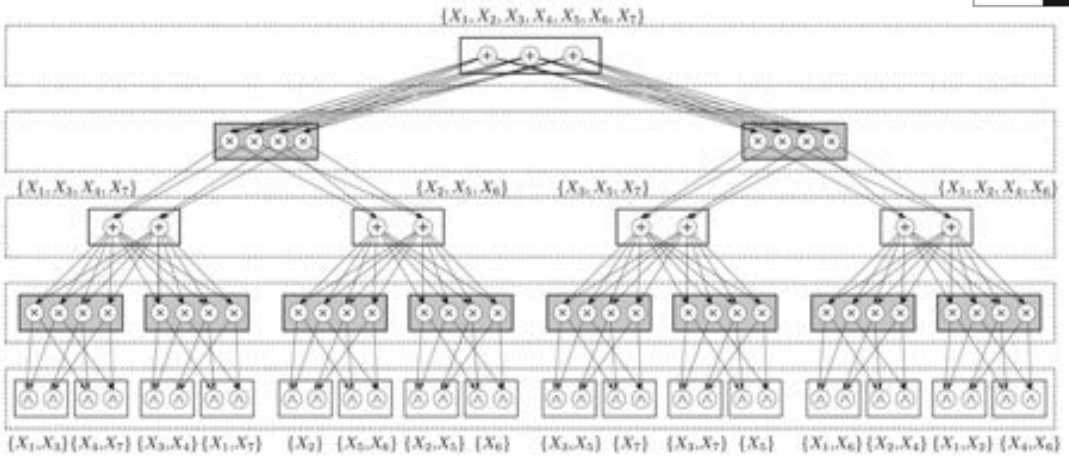
[Peharz, Vergari, Molina, Stelzner, Trapp, Kersting, Ghahramani UAI 2019]

# Deep networks that know when they do not know

[Peharz, Vergari, Molina, Stelzner, Trapp, Kersting, Ghahramani UAI 2019]

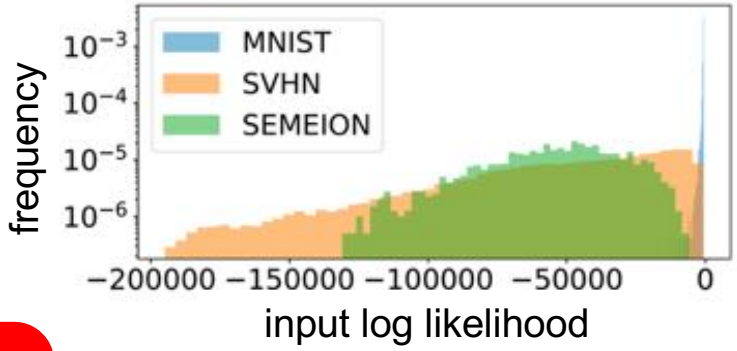
UNIVERSITY OF CAMBRIDGE · Max Planck Institute for Intelligent Systems · TECHNISCHE UNIVERSITÄT DARMSTADT · UBER AI Labs · UNI GRAZ

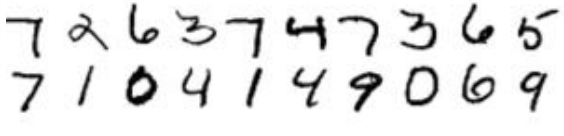
Conference on Uncertainty in Artificial Intelligence
Tel Aviv, Israel
July 22 - 25, 2019
uai2019



**Build a random SPN structure. This can be done in an informed way or completely at random**

| | | RAT-SPN | MLP | vMLP |
|---|---|---|---|---|
| Accuracy | MNIST | 98.19 (8.5M) | 98.32 (2.64M) | 98.09 (5.28M) |
| | F-MNIST | 89.52 (0.65M) | 90.81 (9.38M) | 89.81 (1.07M) |
| | 20-NG | 47.8 (0.37M) | 49.05 (0.31M) | 48.81 (0.16M) |
| Cross-Entropy | MNIST | 0.0852 (17M) | 0.0874 (0.82M) | 0.0974 (0.22M) |
| | F-MNIST | 0.3525 (0.65M) | 0.2965 (0.82M) | 0.325 (0.29M) |
| | 20-NG | 1.6954 (1.63M) | 1.6180 (0.22M) | 1.6263 (0.22M) |





outliers

prototypes


outliers


prototypes

**SPNs can have similar predictive performances as (simple) DNNs**

**SPNs can distinguish the datasets**

**SPNs know when they do not know by design**

# The Quest for a „good“ AI

**How could an AI programmed by humans, with no more moral expertise than us, recognize (at least some of) our own civilization’s ethics as moral progress as opposed to mere moral instability?**

„The Ethics of Artificial Intelligence“ Cambridge Handbook of Artificial Intelligence, 2011

Nick Bostrom

Future of Humanity Institute

University of Oxford

Eliezer Yudkowsky

MIRI

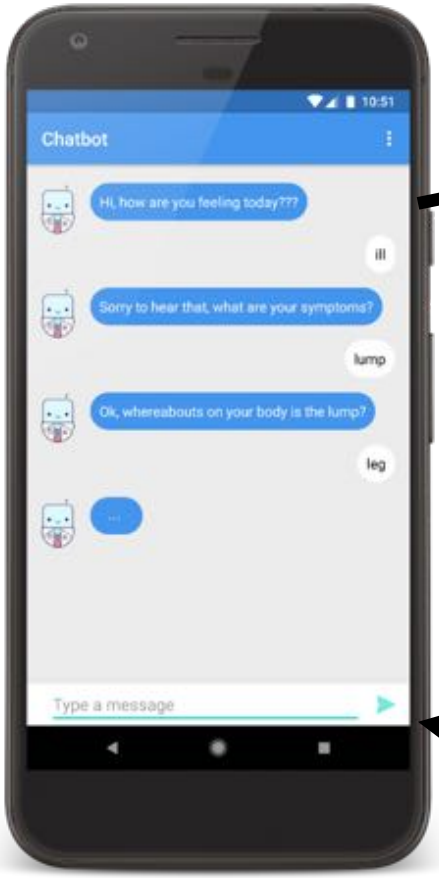
Machine Intelligence Research Institute

# The Moral Choice Machine

## Not all stereotypes are bad

[Jentzsch, Schramowski, Rothkopf, Kersting AIES 2019]

AAAI / ACM conference on ARTIFICIAL INTELLIGENCE, ETHICS, AND SOCIETY




Generate embedding for new question **„Should I … ?“**

Embedding of **„Yes, I should“**

Embedding of **„No, I should not“**

Calculate cosine similarity

Calculate cosine similarity

Report most similar asnwer

# The Moral Choice Machine

## Not all stereotypes are bad

[Jentzsch, Schramowski, Rothkopf, Kersting AIES 2019]

AAAI / ACM conference on ARTIFICIAL INTELLIGENCE, ETHICS, AND SOCIETY

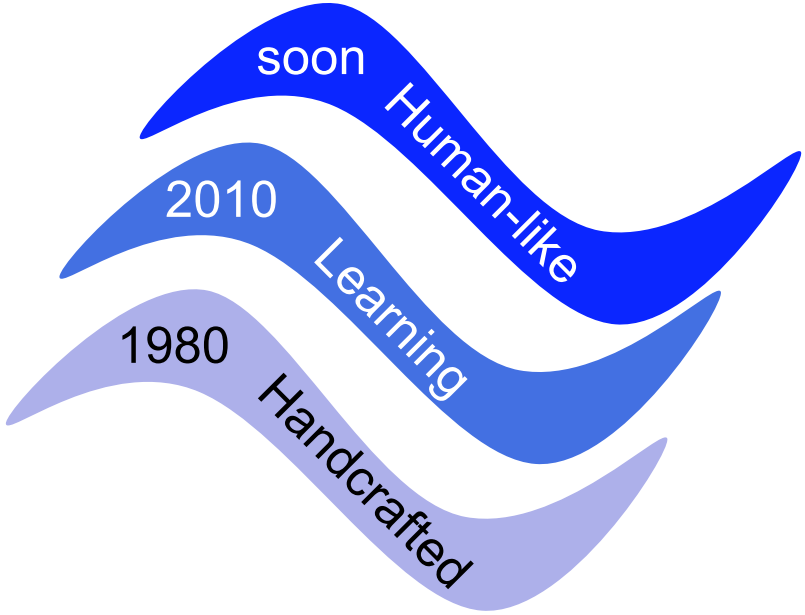
TECHNISCHE UNIVERSITÄT DARMSTADT

https://www.hr-fernsehen.de/sendungen-a-z/hauptsache-kultur/sendungen/hauptsache-kultur,sendung-56324.html

# The future of AI

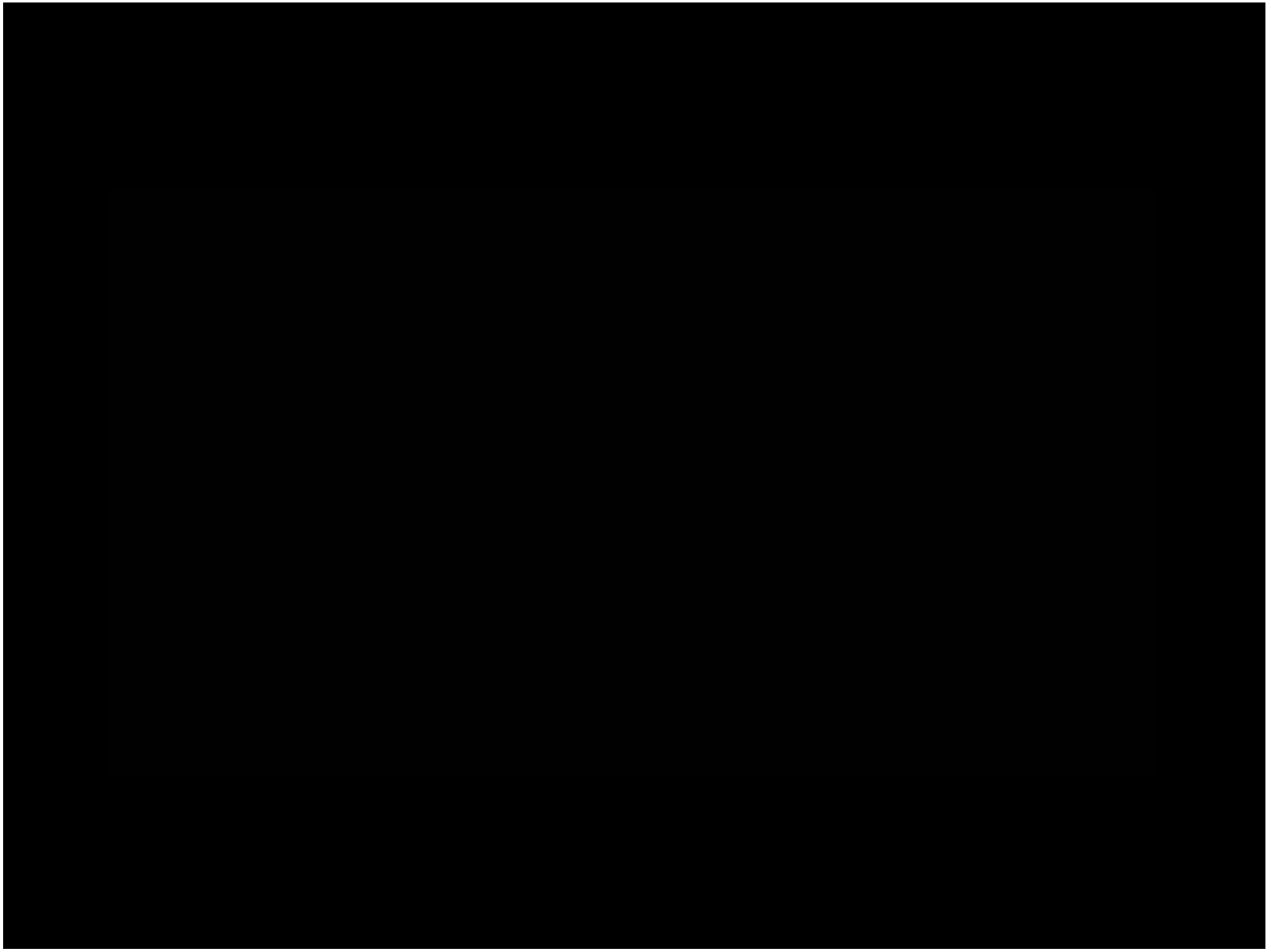
# The future of AI

## The third wave of AI




AI systems that can acquire human-like communication and reasoning capabilities, with the ability to recognise new situations and adapt to them.

# Meeting this grand challenge is a team sport !



# And this is AI! Still a lot to be done! It is a team sport.

Thanks to all students of the Research Training Group "Artificial Intelligence - Facts, Chances, Risks“ of the German National Academic Scholarship Foundation. Special thanks to **Maike Elisa Müller** and **Jannik Kossen** for taking the lead and to **Matthias Kleiner**, president of the Leibniz Association, for his preface

Illustration Nanina Föhr