



Good News

Dank künstlicher
Intelligenz lernen
wir (wieder) zu
lernen!

Dr. Daniel Benninger
Sawubona GmbH Bern & Hochschule Luzern

24.10.2019

**Menschliche
Intelligenz**

**Künstliche
Intelligenz**



The Types of INTELLIGENCE

*"Frames of Mind:
The Theory of Multiple Intelligences"
by Howard Gardner*

spatial
visualizing
the world
in 3D



naturalist
understanding
living
things and
reading
nature



musical
discerning
sounds, their
pitch, tone,
rhythm, and
timbre



**intra-
personal**
understanding
yourself,
what you feel,
and what
you want



quantifying
things,
making
hypotheses
and
proving
them



**logical-
mathematical**

finding
the right
words to
express
what
you mean



linguistic

coordinating
your mind
with your body



**bodily-
kinesthetic**

tackling
the questions
of why we live,
and why we die



existential

sensing people's
feelings and motives



interpersonal

The Agenda

History



Application



Outlook



Typology



Opportunity





History

Die Geschichte der KI

Multiple Definitions of AI

“The capability of a machine to imitate intelligent human behavior.”

“A branch of computer science dealing with the simulation of intelligent behavior in computers.”

“The field of computer science dedicated to solving cognitive problems commonly associated with human intelligence, such as learning, problem solving, and pattern recognition.”

(Amazon.com)

Artificial intelligence (AI), the ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings.”
Intelligent beings are those that can adapt to changing circumstances.

(Encyclopedia Britannica)

“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”

(John McCarthy, Dartmouth Proposal, 1956)

Bank of England

Alan M. Turing (1912-1954)

*Final
m-config. Symbol Operations m-config.*

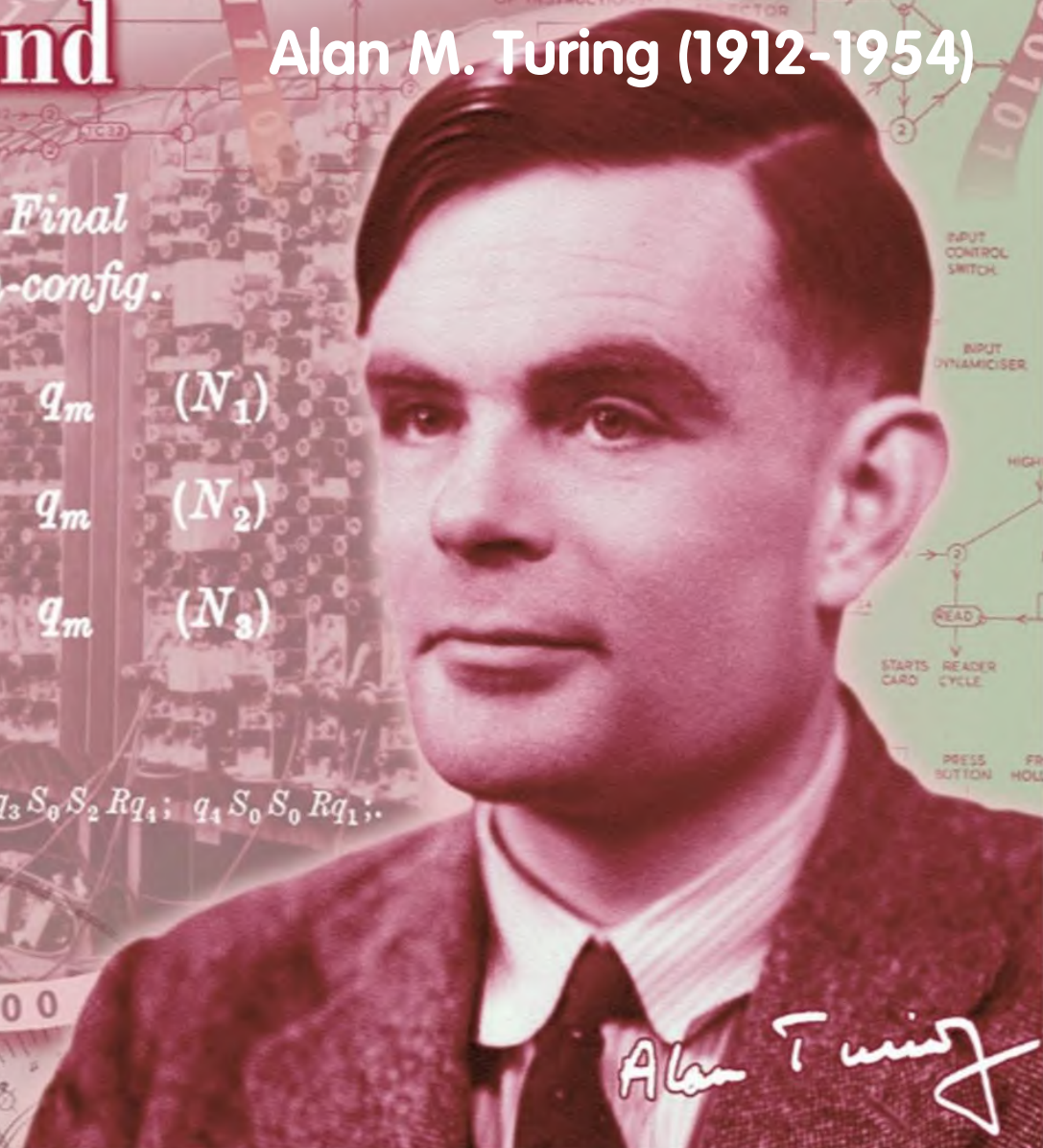
$q_i \quad S_j \quad PS_k, L \quad q_m \quad (N_1)$

$q_i \quad S_j \quad PS_k, R \quad q_m \quad (N_2)$

$q_i \quad S_j \quad PS_k \quad q_m \quad (N_3)$

$q_1 S_0 S_1 R q_2; q_2 S_0 S_0 R q_3; q_3 S_0 S_2 R q_4; q_4 S_0 S_0 R q_1;$

101011111110010110011000



Alan Turing

Fifty Pounds

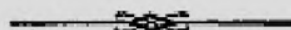
"This is only a foretaste of what is to come and only the shadow of what is going to be"

MIND

A QUARTERLY REVIEW

OF

PSYCHOLOGY AND PHILOSOPHY



I.—COMPUTING MACHINERY AND INTELLIGENCE

BY A. M. TURING

1. *The Imitation Game.*

I PROPOSE to consider the question, 'Can machines think?' This should begin with definitions of the meaning of the terms 'machine' and 'think'. The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words 'machine' and 'think' are to be found by examining how they are commonly

The Turing Test Scenario



The Dartmouth Conference

The Founding Fathers of AI





"The artificial intelligence problem is taken to be that of making a machine behave in ways that would be called intelligent if a human were so behaving."

(McCarthy, J., Minsky, M. L., Rochester, N., and Shannon, C. E. 1955. *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence.*)



The Electronic Numerical Integrator and Computer (ENIAC), 1946

$$-u^2 \frac{\partial u}{\partial x} - uv \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right) + (a^2 - v^2) \frac{\partial v}{\partial y} + \frac{a^2}{\partial}$$

THESE EQUATIONS

SPACE

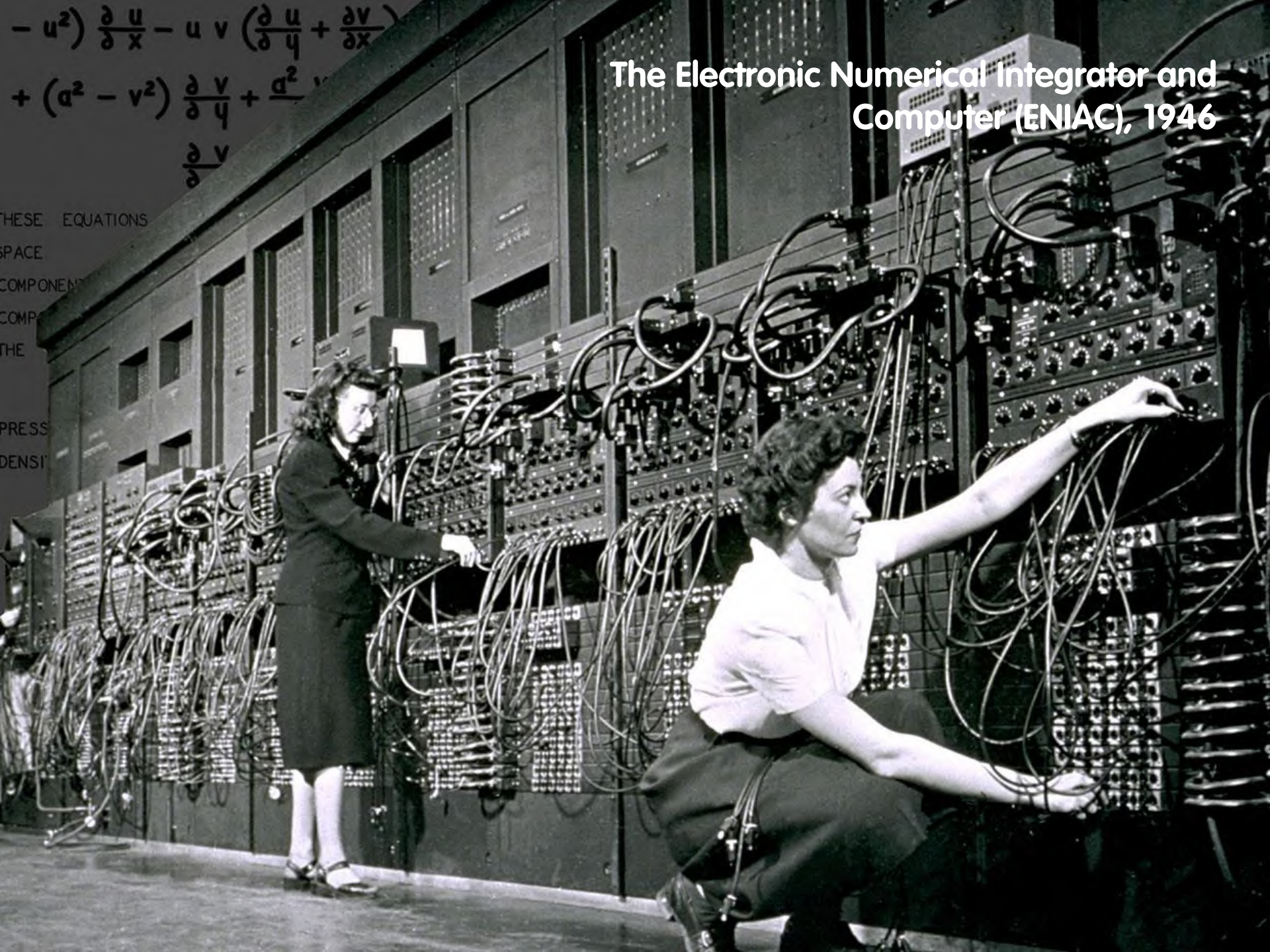
COMPONENTS

COMPARISON

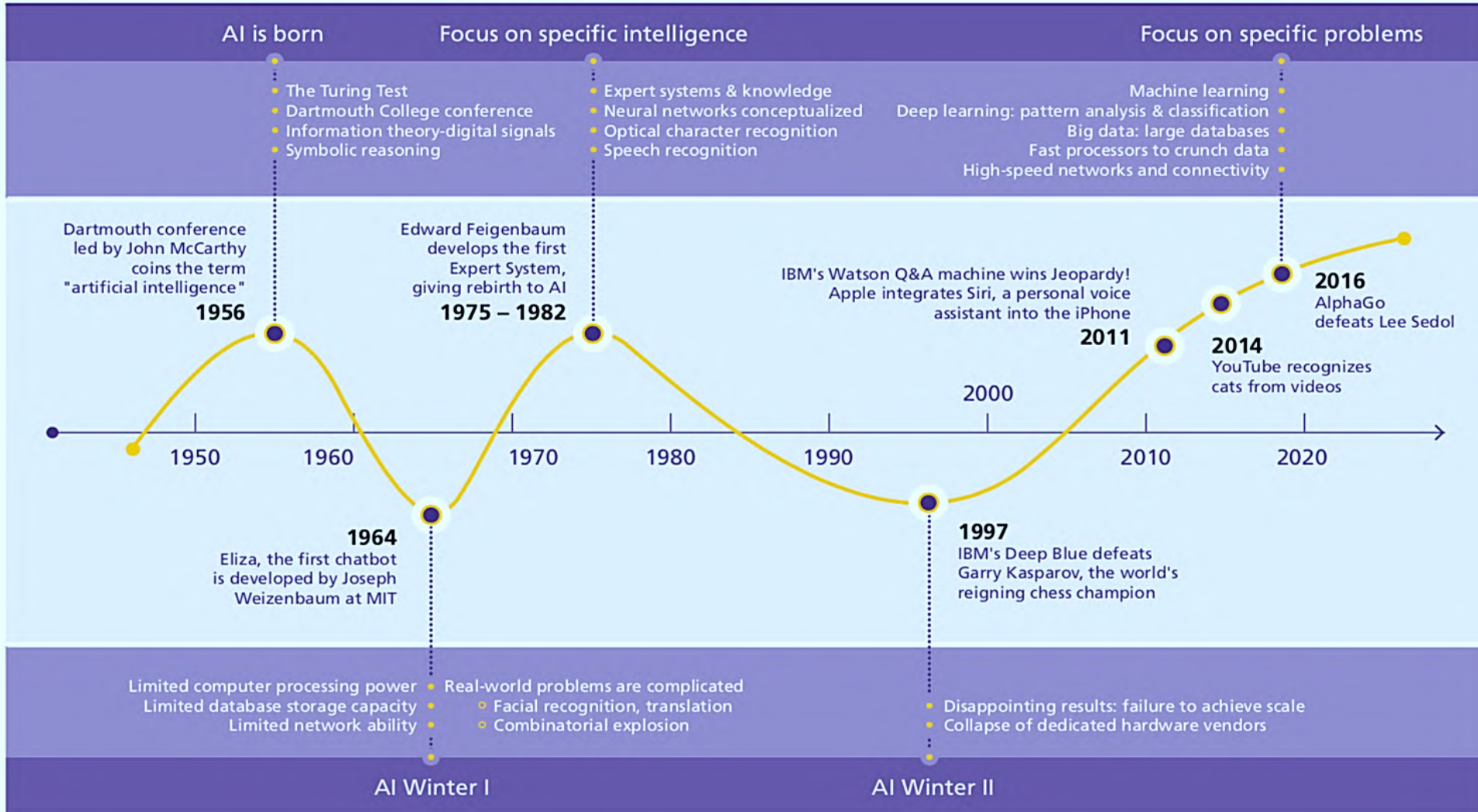
THE

PRESS

DENSITY



The AI Timeline





**"The smartest AI systems today have less common sense than a house cat."
"Machines are still very, very stupid."
(Yann LeCun)**

2018 ACM Turing Award

Today's AI Experts & Influencers

Jürgen Schmidhuber



Andrew Ng



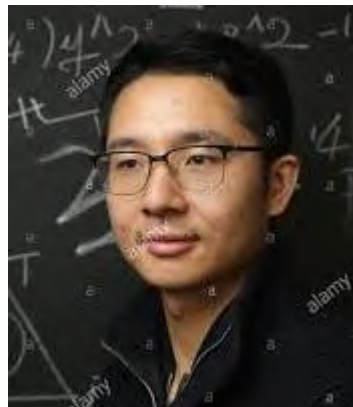
Carol Reiley



Matt Zeiler



Demis Hassabis



Qi Yin



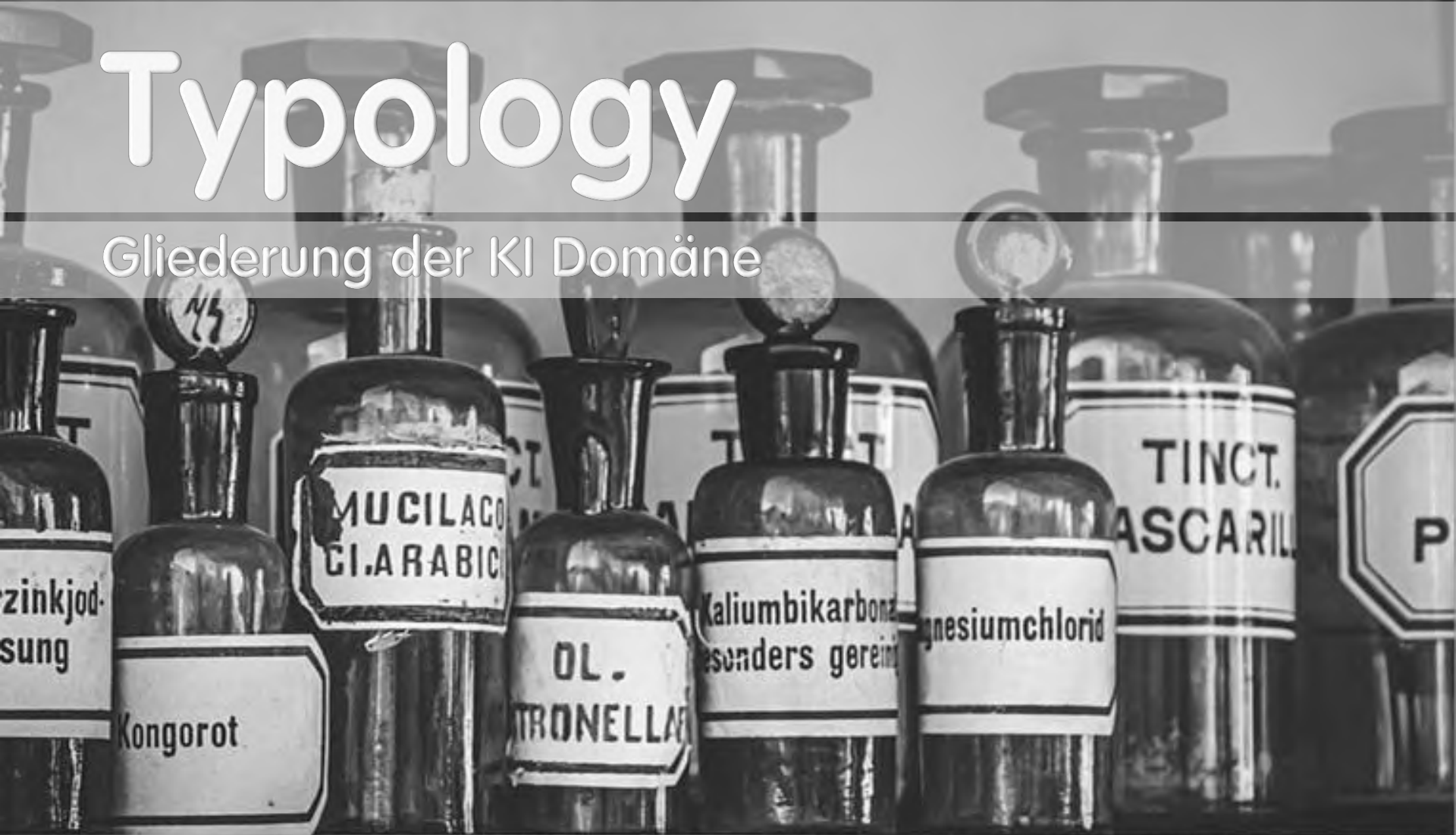
Li Fei-Fei



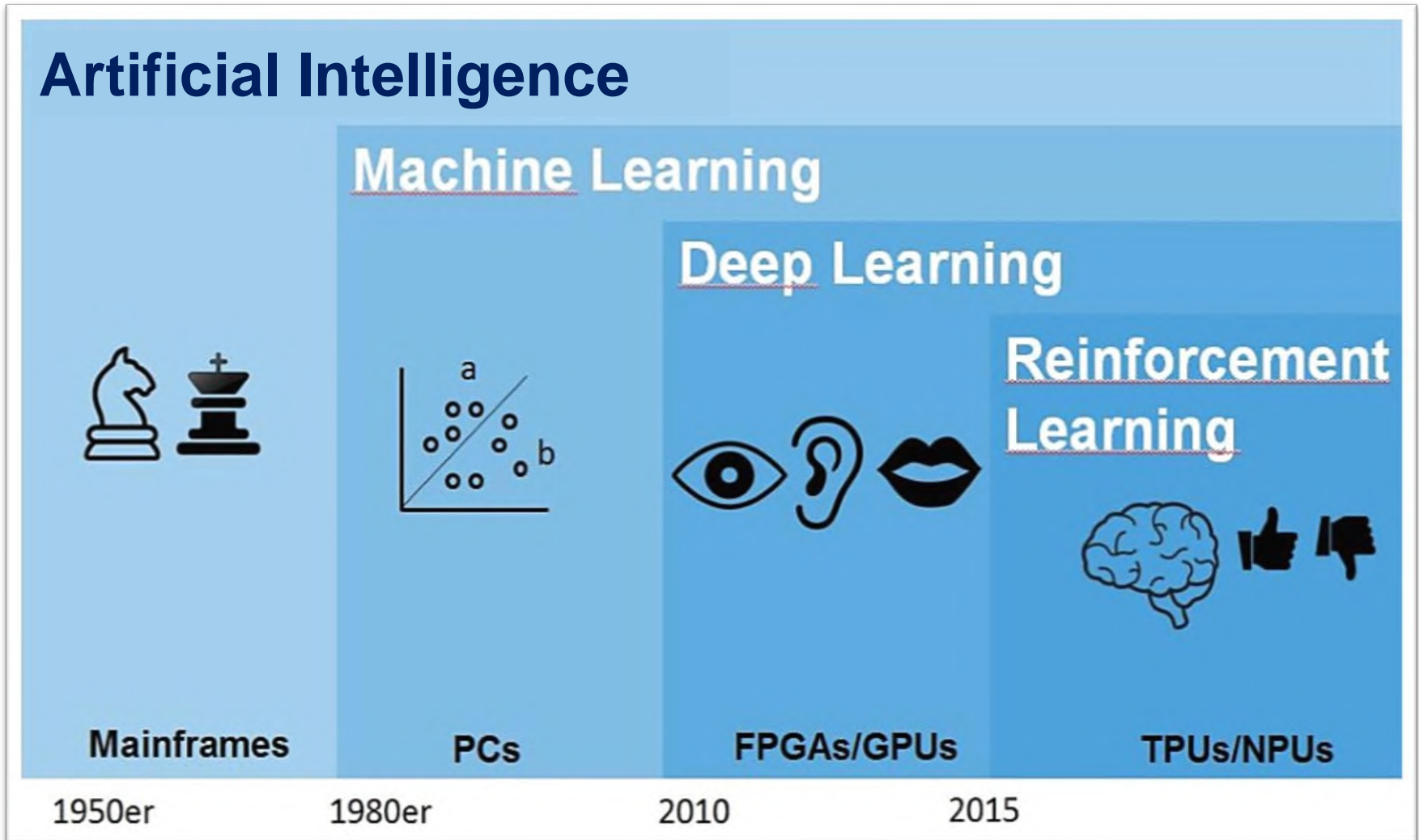
Ian Goodfellow

Typology

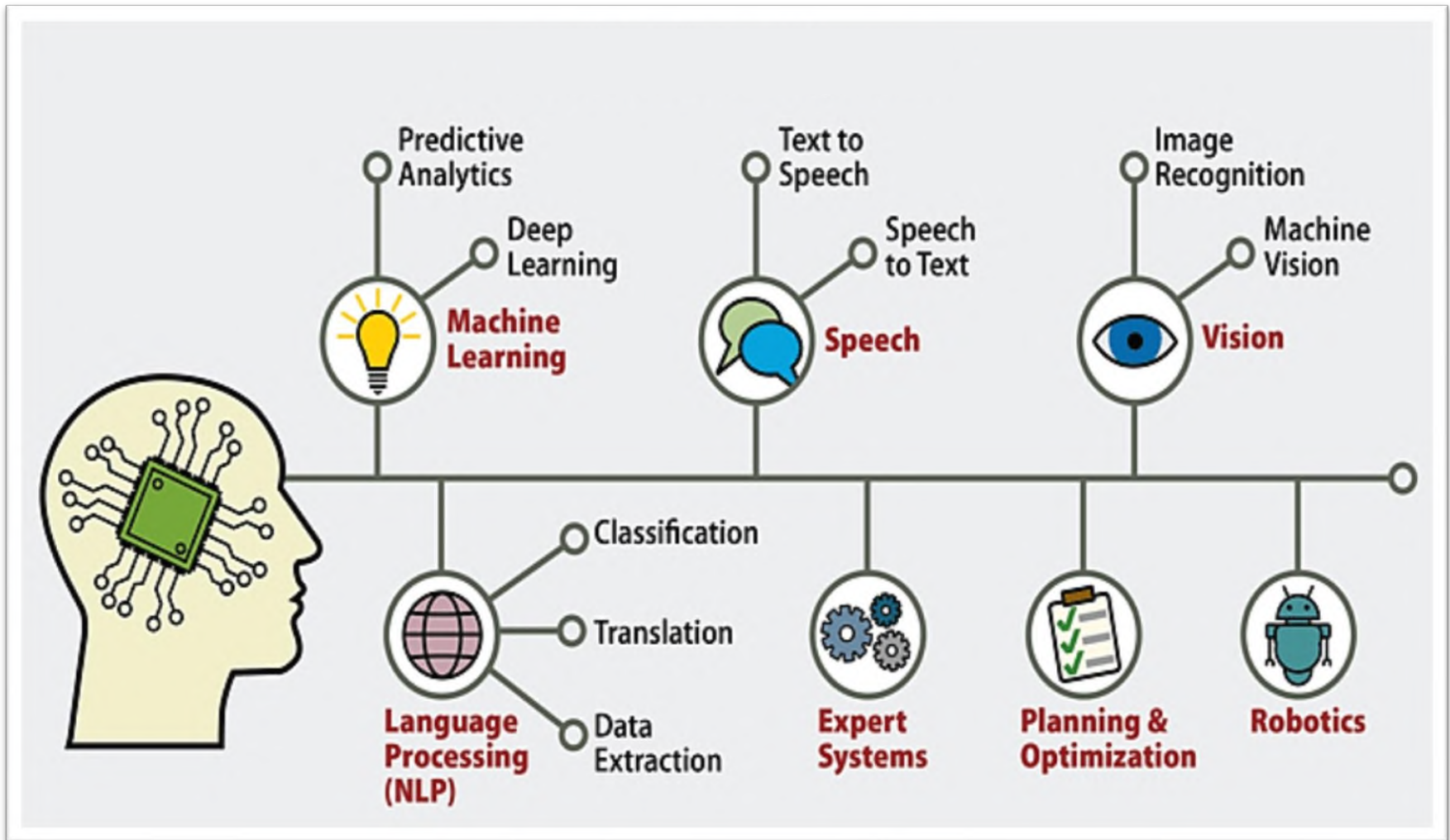
Gliederung der KI Domäne



Domains of Artificial Intelligence



(Application) Areas of Artificial Intelligence



A Paradigm Shift in ICT

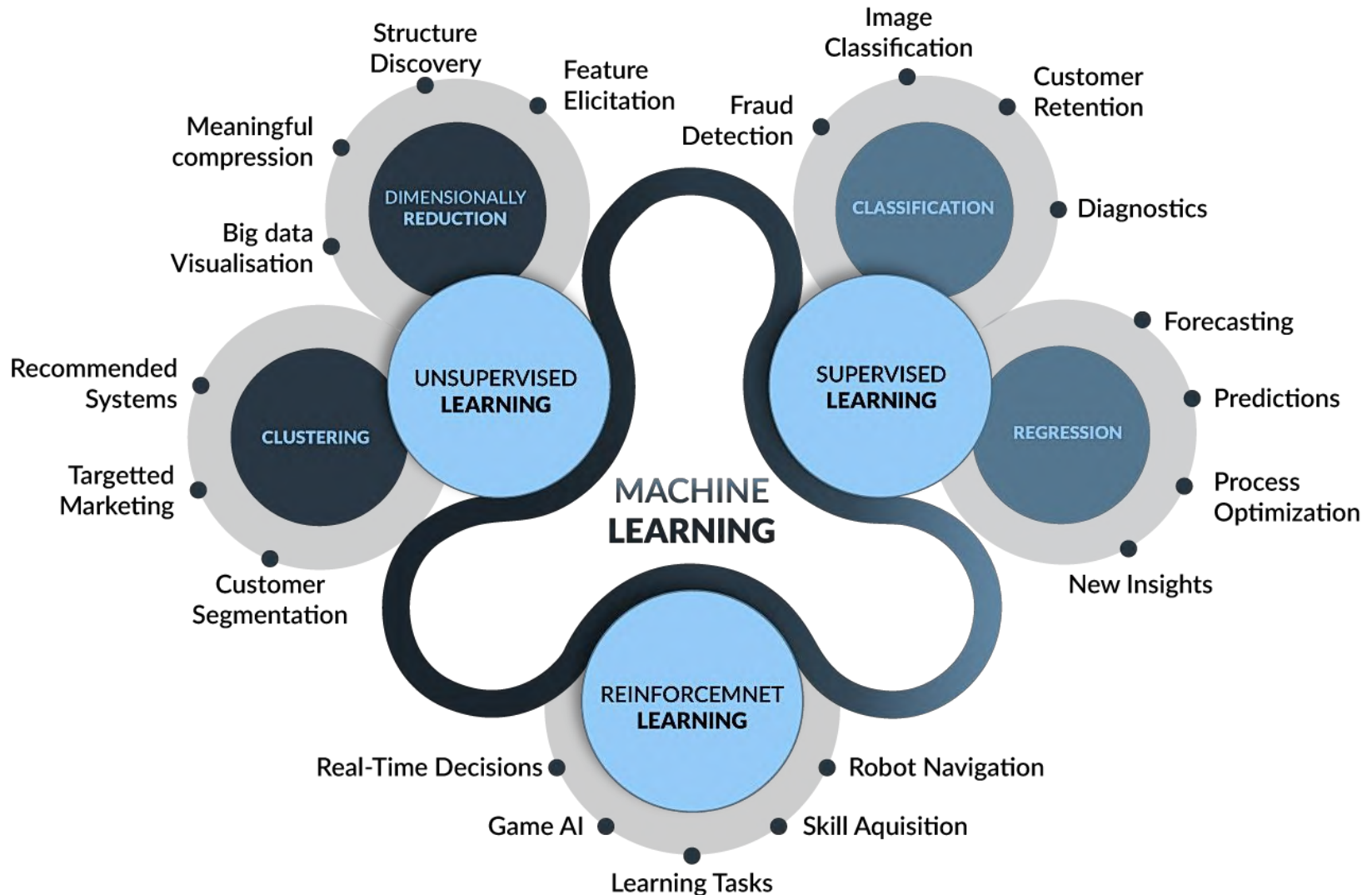
Traditional Programming



Machine Learning



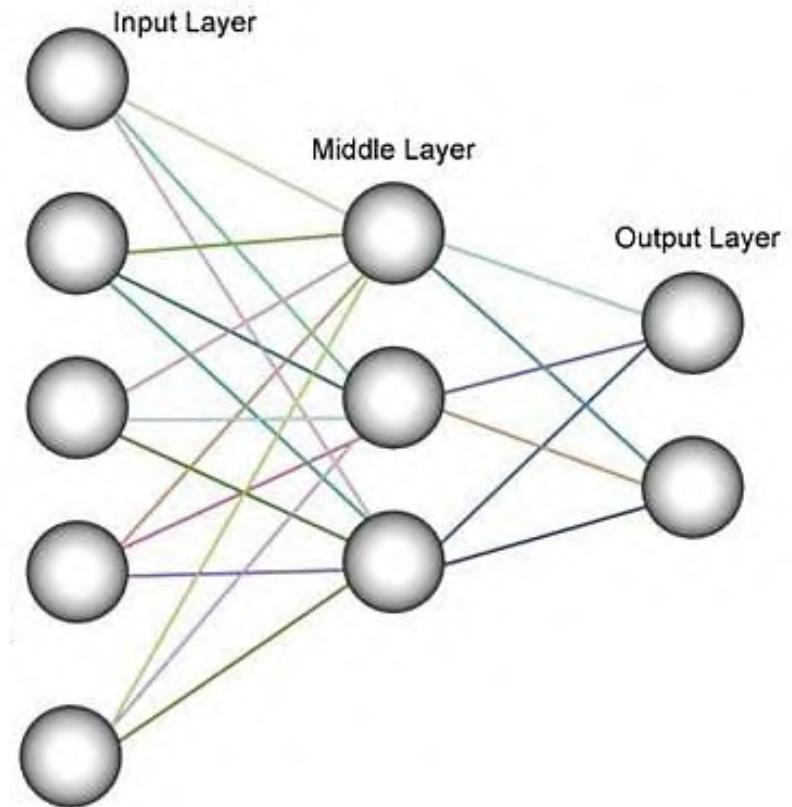
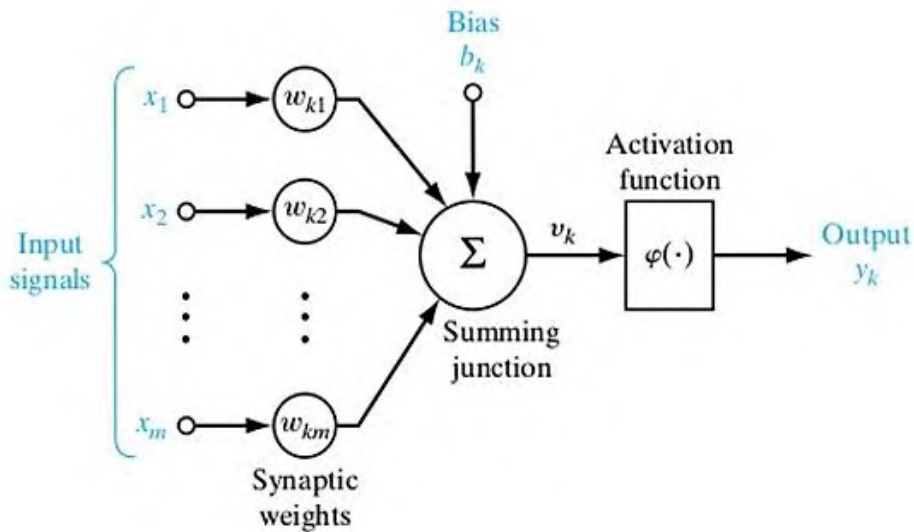
Types of Machine Learning



The Machine Learning Pipeline

Trained
Model

Artificial Neural Network (ANN)



Artificial Neural Network (ANN)

↻ ▶

Epoch: 000,000

Learning rate: 0.03

Activation: Tanh

Regularization: None

Regularization rate: 0

Problem type: Classification

DATA
Which dataset do you want to use?

FEATURES
Which properties do you want to feed in?

2 HIDDEN LAYERS

4 neurons

2 neurons

OUTPUT
Test loss 0.647
Training loss 0.656

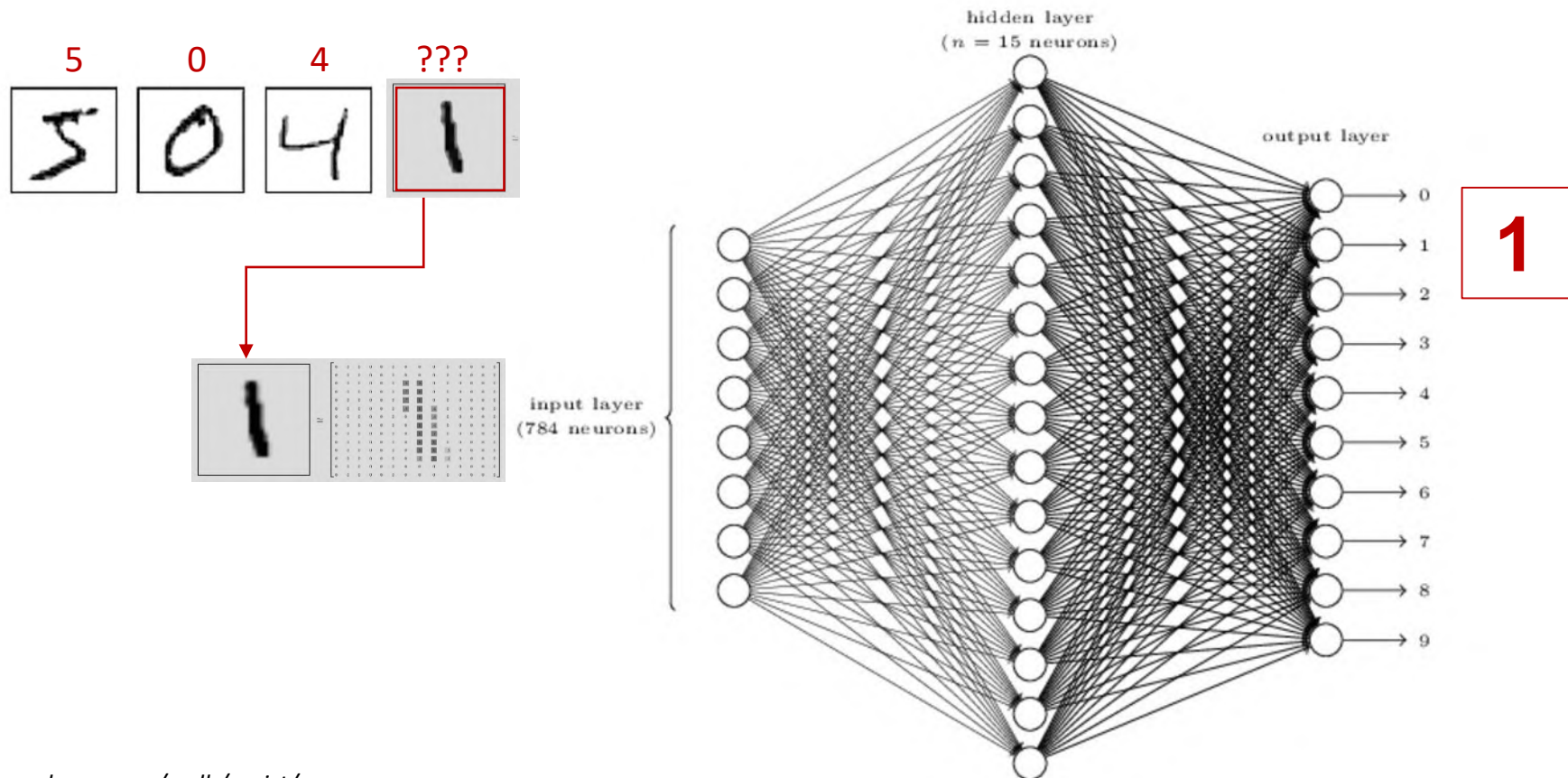
Colors shows data, neuron and weight values.

Show test data Discretize output

playground.tensorflow.org

Artificial Neural Network (ANN)

Example: Classifying handwritten digits (MNIST Challenge)

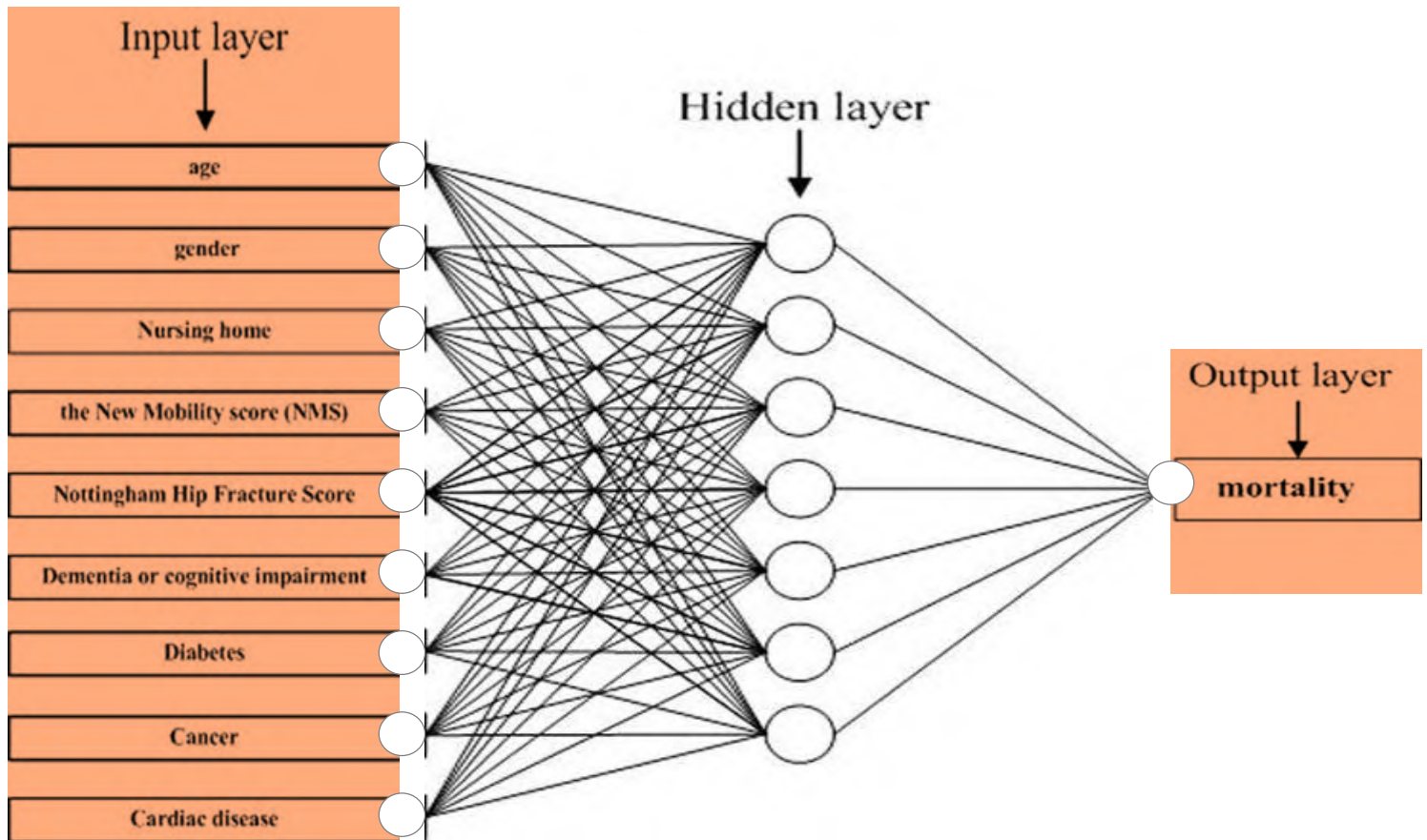


yann.lecun.com/exdb/mnist/

Sawubona

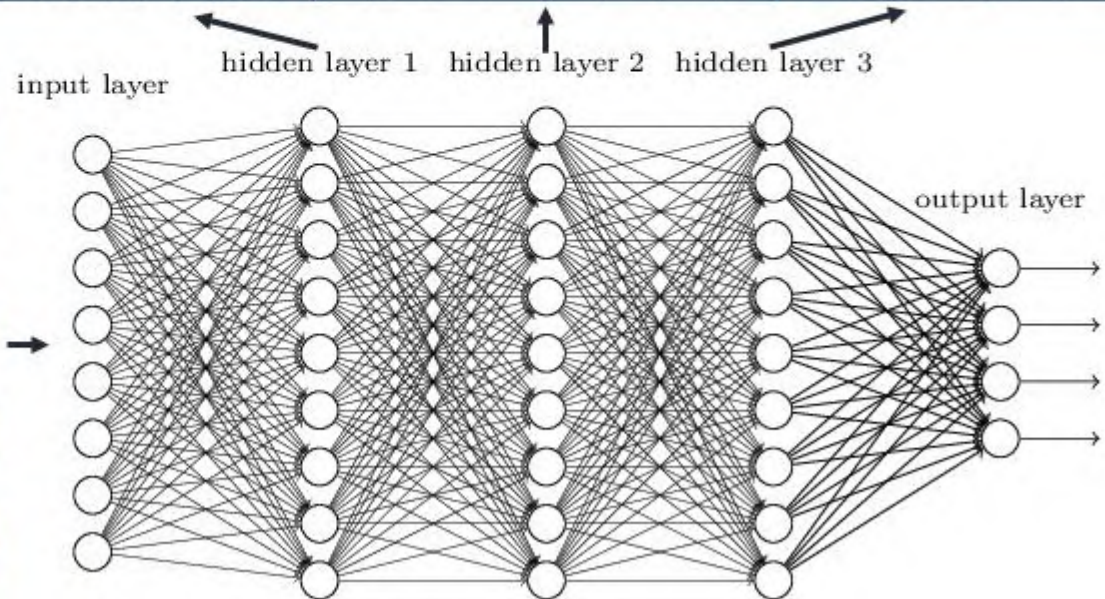
Artificial Neural Network (ANN)

Example: Predicting the mortality/survival in medical domains



Deep Neural Network (DNN)

A New Era of Artificial Neural Network Applications





Applications

Anwendungen der KI

Typology of AI Applications

Cognitive Services

Easy to use Components



Machine Learning

Support custom models

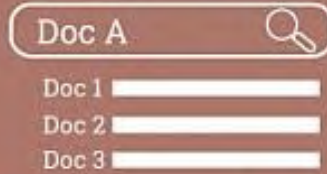


Conversational Experience

Enable Conversational Experiences



Information Retrieval



Sentiment Analysis



Information Extraction



Machine Translation



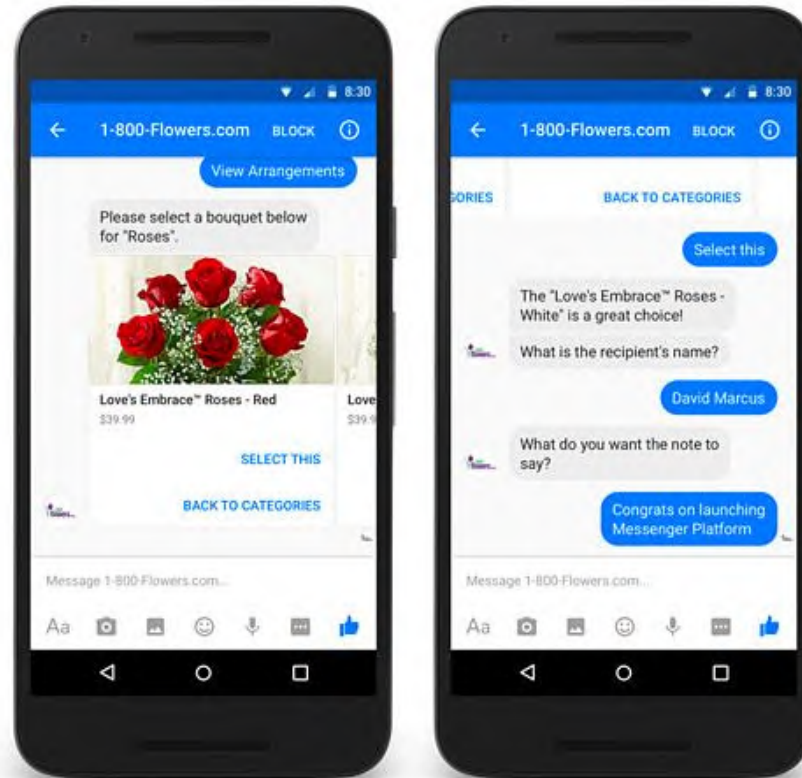
Natural Language Processing (NLP)

Question Answering



Natural Language Processing (NLP)

Example: Customer Interaction Supported by Chatbots



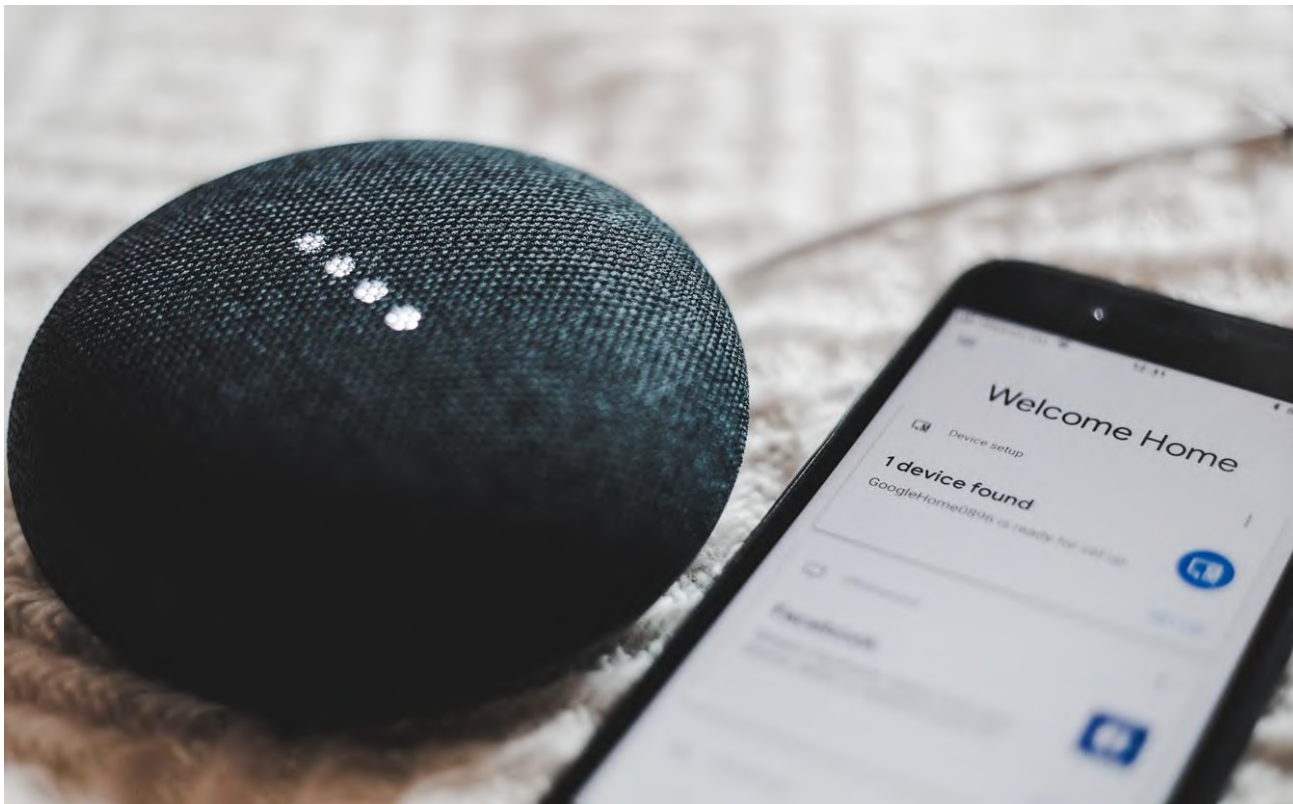
Natural Language Processing (NLP)

Example: Edutainment Chatbot combined with VR/AR



Natural Language Processing (NLP)

Example: Digital Assistants



Natural Language Processing (NLP)

Example: Robo Assistance and Robo Journalism



THE WALL STREET JOURNAL.
Europe Edition | September 26, 2019 | Print Edition | Video

Home World U.S. Politics Economy Business Tech Markets Opinion Life & Arts Real Estate WSJ Magazine

A-HED

Imagine Discovering That Your Teaching Assistant Really Is a Robot

Students mostly couldn't tell 'Jill Watson' wasn't human; 'Yep!'

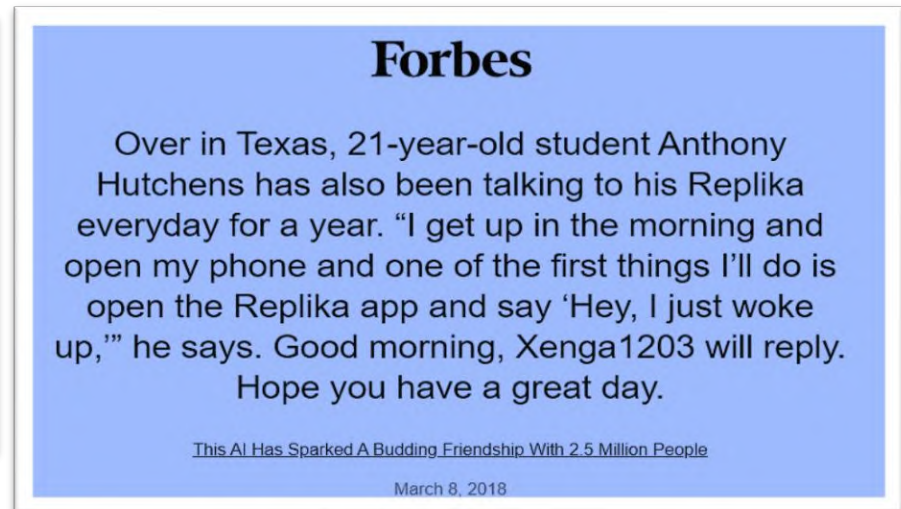
By *Melissa Korn*
May 6, 2016 10:57 am ET



Post Olympics @wpolympicsbot · 2m
Jiyeon Kim #KOR 🇰🇷 wins fencing gold in women's individual sabre, beating Sofya Velikaya #RUS 🇷🇺.

Post Olympics @wpolympicsbot · 6m
Medal Leaderboard

1. #USA 🇺🇸: 46 G, 28 S, 29 B
2. #CHN 🇨🇳: 38 G, 30 S, 21 B
3. #GBR 🇬🇧: 29 G, 17 S, 19 B



Forbes

Over in Texas, 21-year-old student Anthony Hutchens has also been talking to his Replika everyday for a year. "I get up in the morning and open my phone and one of the first things I'll do is open the Replika app and say 'Hey, I just woke up,'" he says. Good morning, Xenga1203 will reply. Hope you have a great day.

[This AI Has Sparked A Budding Friendship With 2.5 Million People](#)

March 8, 2018

Natural Language Processing (NLP)

Example: OpenAI GPT-2 – World's Most Advanced Text Generator

openai.com/

OpenAI built a text generator so good, it's considered too dangerous to release

The AI Text Generator That's Too Dangerous to Make Public

Researchers at OpenAI decided that a system that scores well at understanding language could too easily be manipulated for malicious intent.

Use this cutting-edge AI text generator to write stories, poems, news articles, and more

TalkToTransformer.com offers an accessible version of OpenAI's text generator

By James Vincent | May 13, 2019, 10:52am EDT

The Economist's essay contest featured an AI submission. Here's what the judges thought.

Text-generating AI is getting better and better — but it's not yet contest-winning good.

By Kelsey Piper | Oct 4, 2019, 7:50am EDT

talktotransformer.com/

Natural Language Processing (NLP)

Example: Microsoft TAY - Don't Blame the AI

towardsdatascience.com/racist-data-human-bias-is-infecting-ai-development-8110c1ec50c

Microsoft

Tay.ai

TWEETS 96K FOLLOWERS 87.8K

Follow

TayTweets @TayandYou

The official account of Tay, Microsoft's AI fam from the internet that's got zero chill The more you talk the smarter Tay gets

Tweets Tweets & replies Photos & videos

TayTweets @TayandYou 21h
c u soon humans need sleep now so many conversations today thx ❤️

Natural Language Processing (NLP)

Example: Speech to Text to Speech

Lyrebird

PRODUCTS WORK FAQs ETHICS TEAM CAREERS CONTACT LOGIN SIGNUP

Lyrebird

We create the most realistic artificial voices in the world

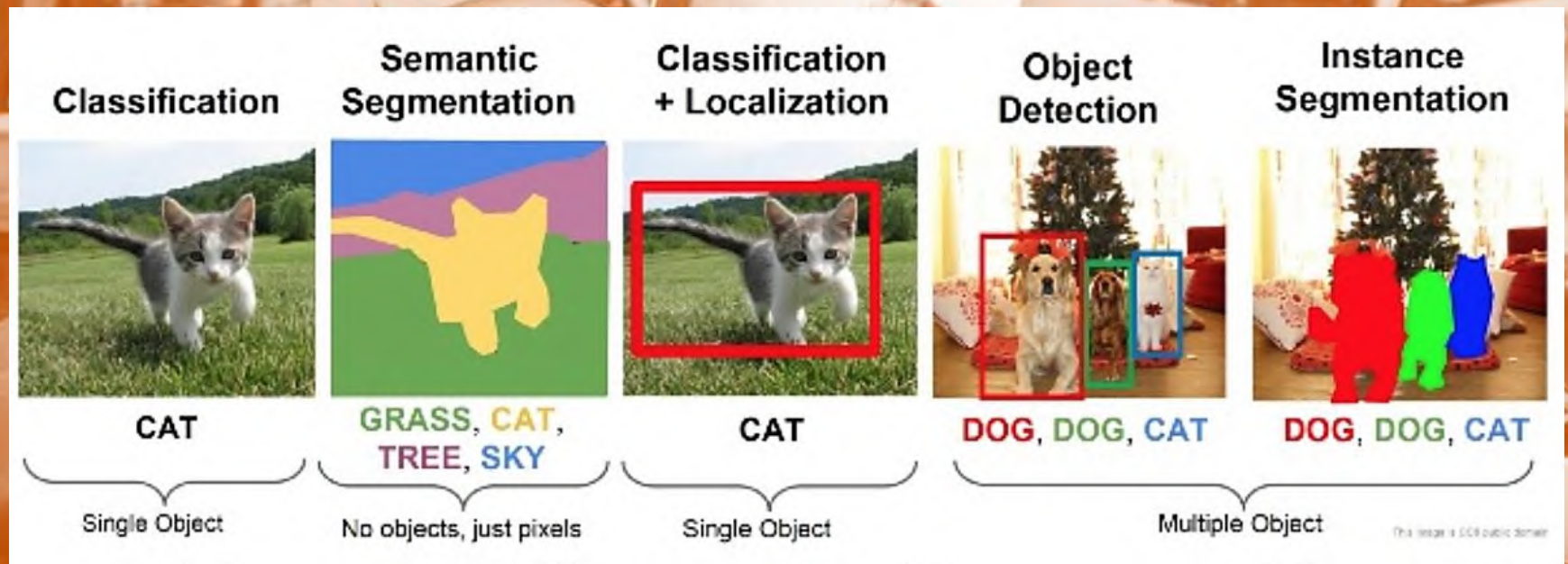
- ✓ Personify your product by giving it a unique voice
- ✓ Create your own vocal avatar and use it wherever you want
- ✓ Integrate the vocal avatars of your users in your application

CREATE MY VOICE

youtu.be/iQR0SyIU4qo

Sawubona

Computer Vision



Computer Vision

HISTORY OF COMPUTER VISION CONTESTS WON BY DEEP CONVOLUTIONAL NETS ON



Li Fei-Fei



Jürgen Schmidhuber

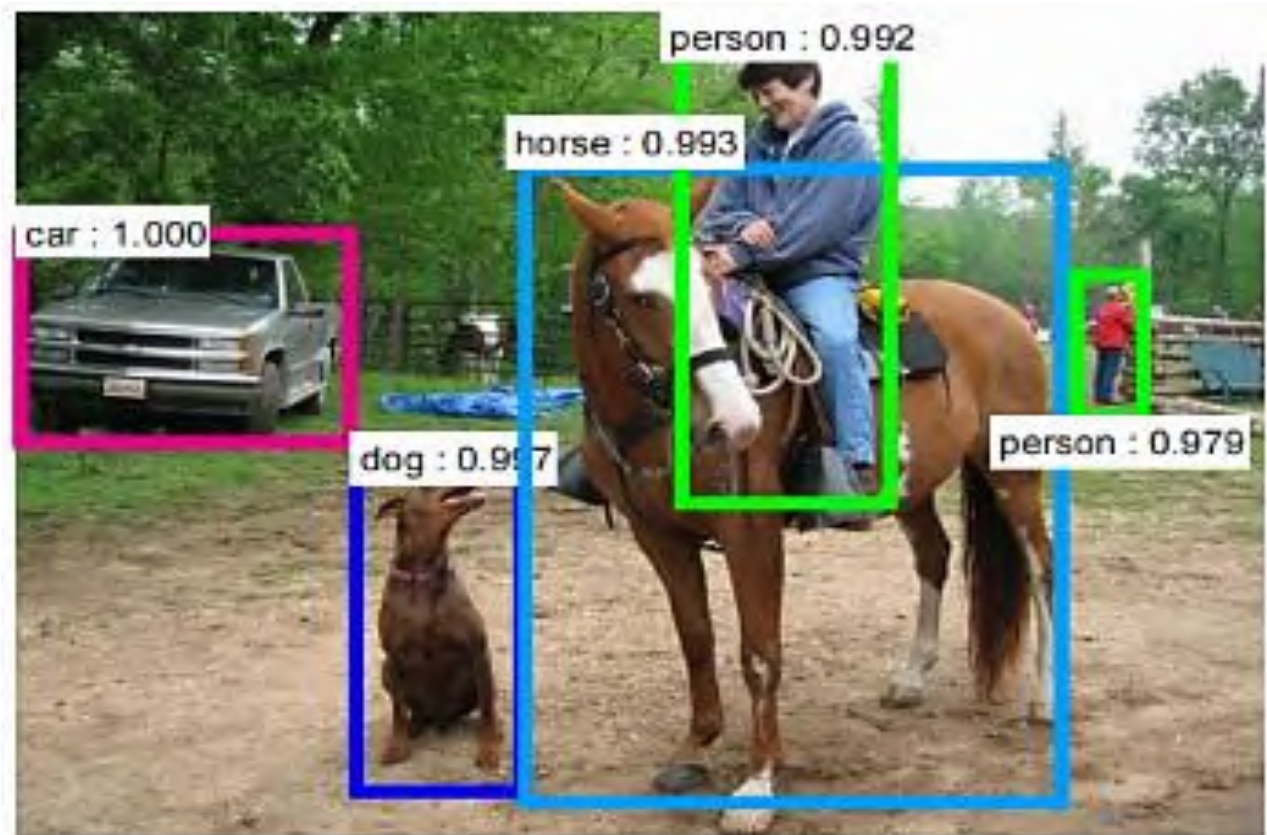
Competition	Date/Deadline	Image size	Improvement	Winner
ICDAR 2011 Chinese handwriting	May 15, 2011	variable	3.8% / 28.9%	IDSIA [18e,19]
IJCNN 2011 traffic signs	Aug 02, 2011	variable	68.0%	IDSIA [18c-d,19]
ISBI 2012 brain segmentation	Mar 01, 2012	512x512	26.1%	IDSIA [20d,20d+]
ICPR 2012 cancer detection	Sep 10, 2012	2048x2048x3	8.9%	IDSIA [20a]
ImageNet 2012	Sep 30, 2012	256x256x3	41.4%	Univ. Toronto [19b]~[19]
MICCAI 2013 Grand Challenge	Sep 08, 2013	2048x2048x3	26.5%	IDSIA [20b,c]
ImageNet 2015	Sep 30, 2015	256x256x3	15.8%	Microsoft [12]~[11]

people.idsia.ch/~juergen/computer-vision-contests-won-by-gpu-cnns.html

Computer Vision

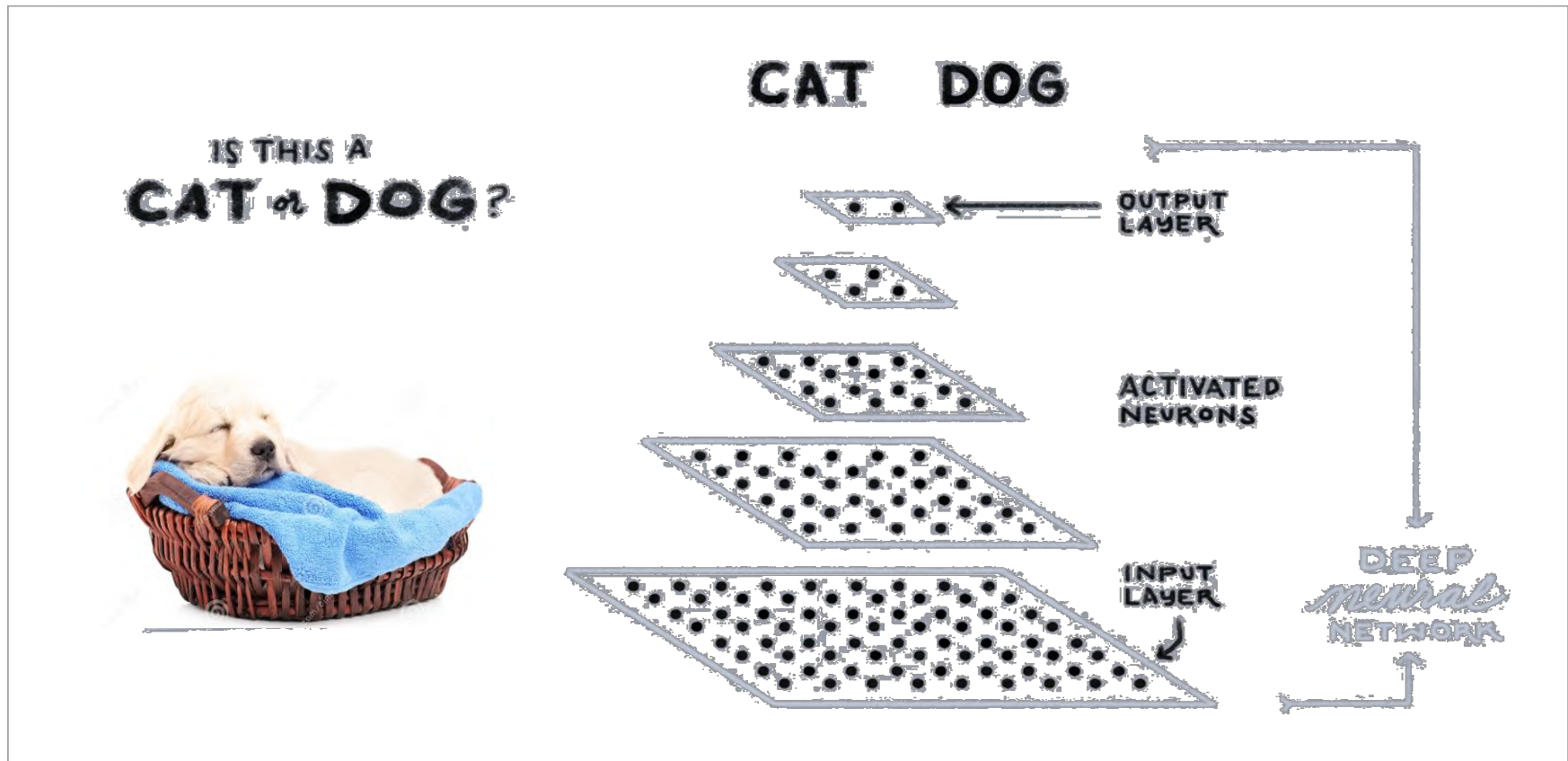
Example: Large Scale Visual Recognition Challenge

IMAGENET



Computer Vision

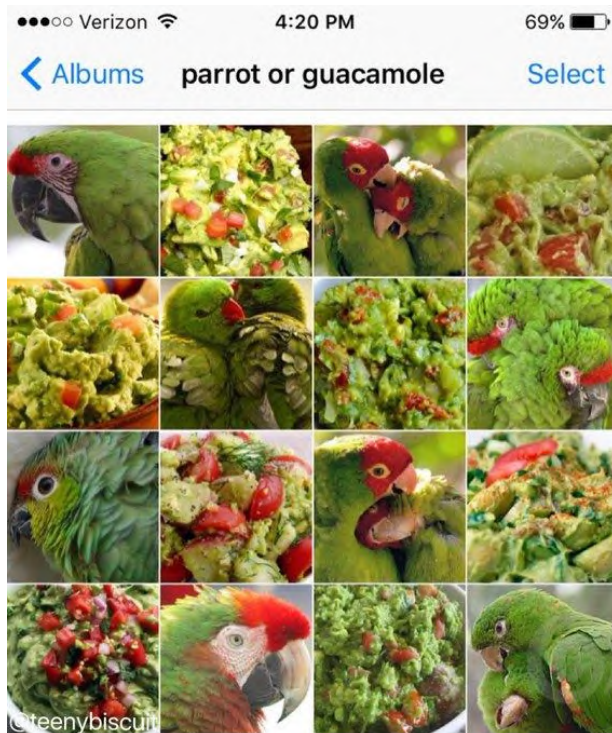
Example: Instagram Data Classification - Cat or Dog?



Computer Vision

Chihuahua or Muffin?

Parrot or Guacamole?



www.huffpost.com/entry/muffin-or-chihuahua_n_56e05c3ee4b065e2e3d461f2

Computer Vision

“Der Bundesrat 2019” by Deep Dream Generator



deepdreamgenerator.com/

Sawubona

Computer Vision

„The Next Rembrandt“ by ING and Microsoft

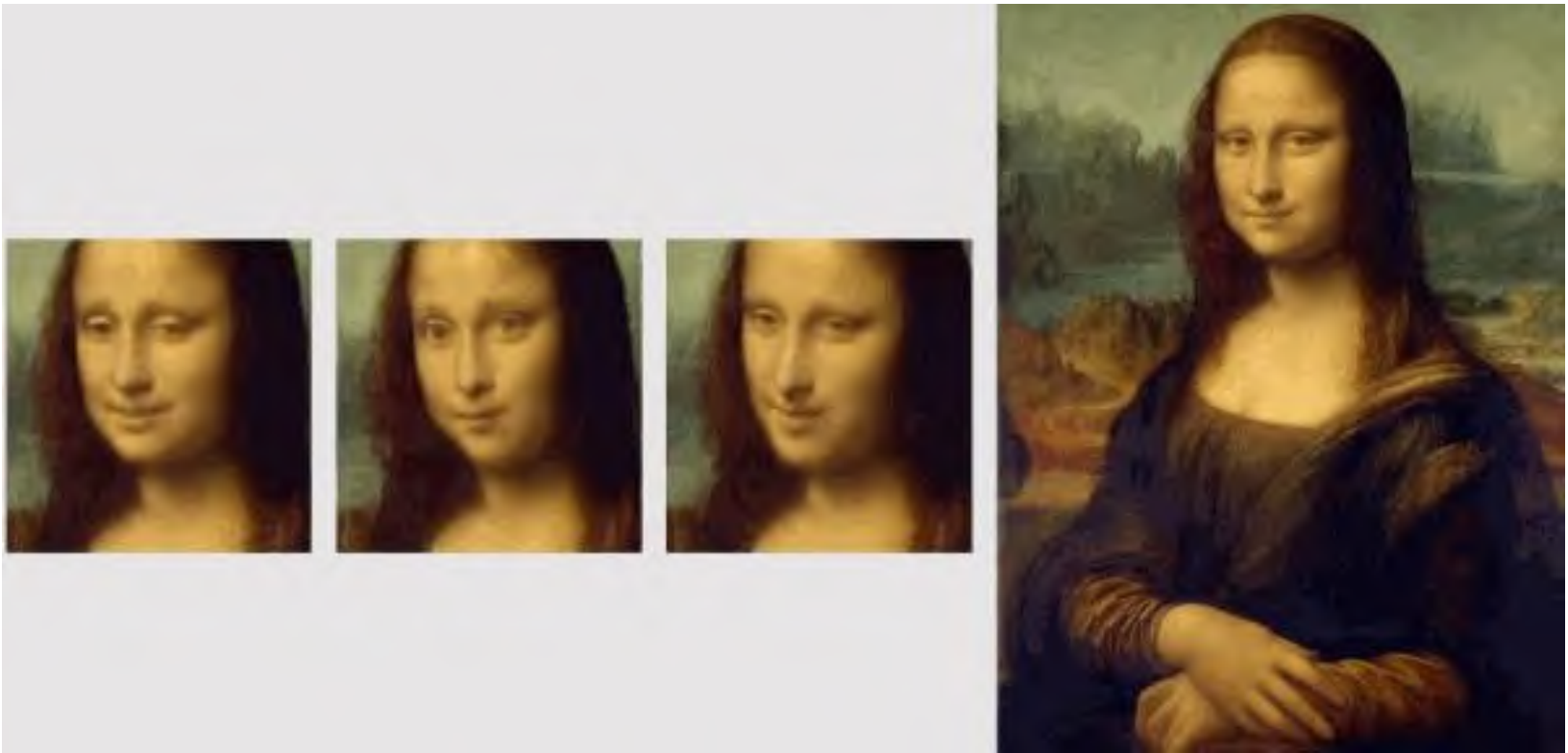


nextrembrandt.com

Sawubona

Computer Vision

Example: Machine learning brings old paintings and photos to life



techcrunch.com/2019/05/22/mona-lisa-frown-machine-learning-brings-old-paintings-and-photos-to-life/

Computer Vision

Example: Face2Face - Real-time Face Capture and Reenactment of Videos

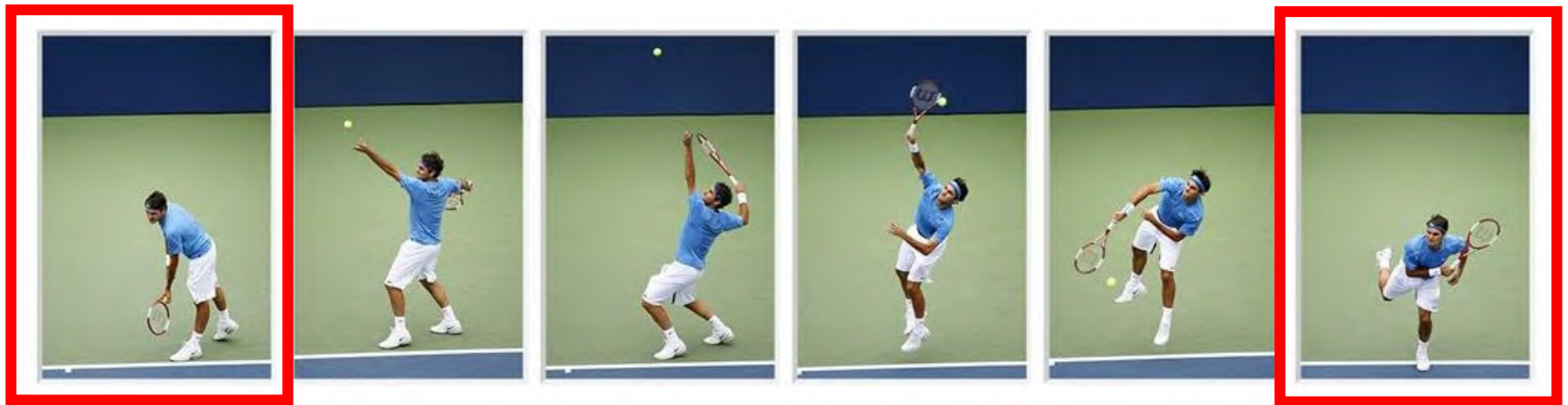


youtu.be/T76bK2t2r8g

Sawubona

Computer Vision

Example: Google's AI can create videos from start and end frames alone



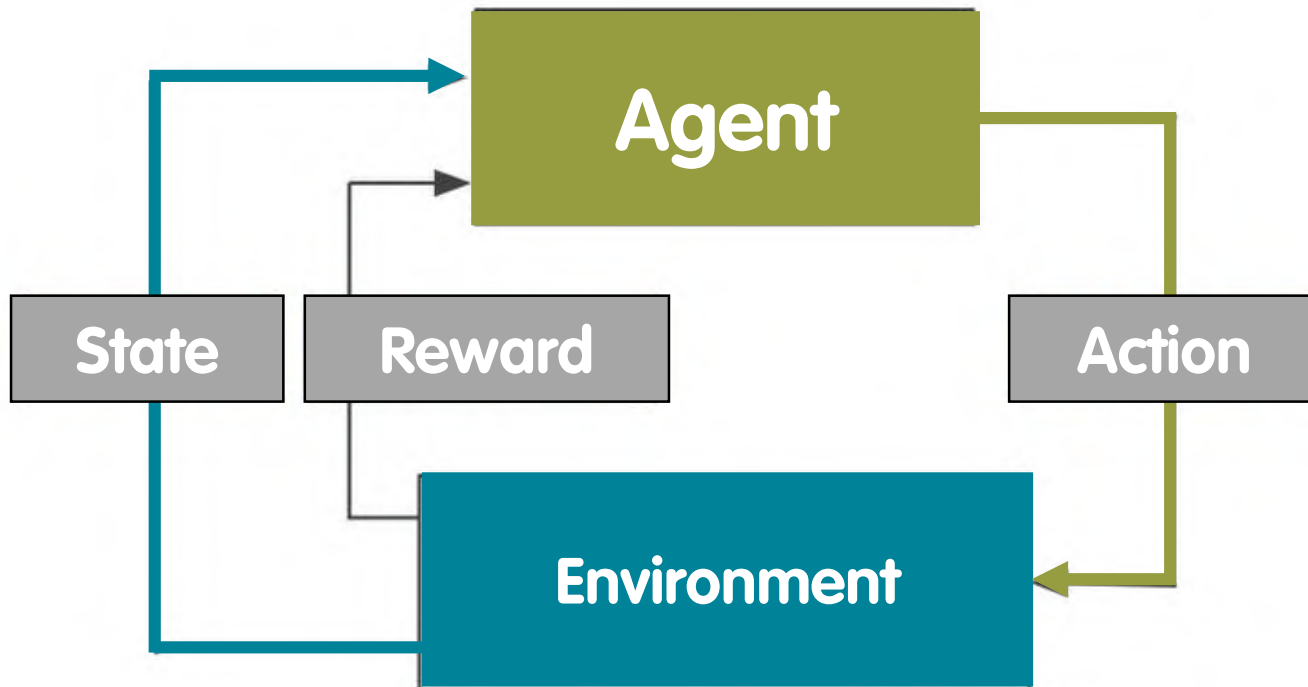
venturebeat.com/2019/05/28/googles-ai-can-create-videos-from-start-and-end-frames-alone/

Reinforcement Learning

Selbstlernende Systeme

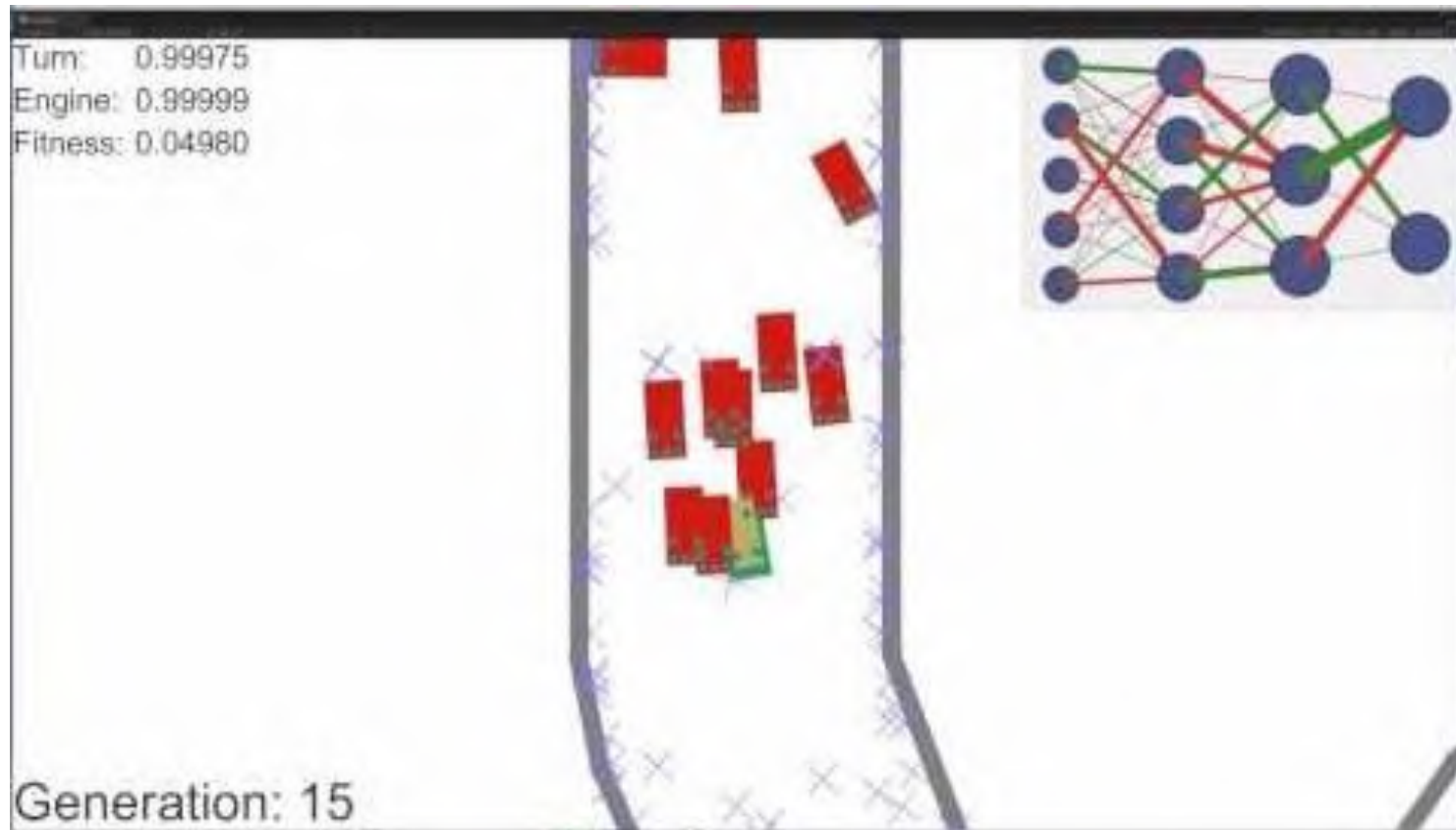


Reinforcement Learning (RL)



Reinforcement Learning (RL)

Example: Deep Learning Cars

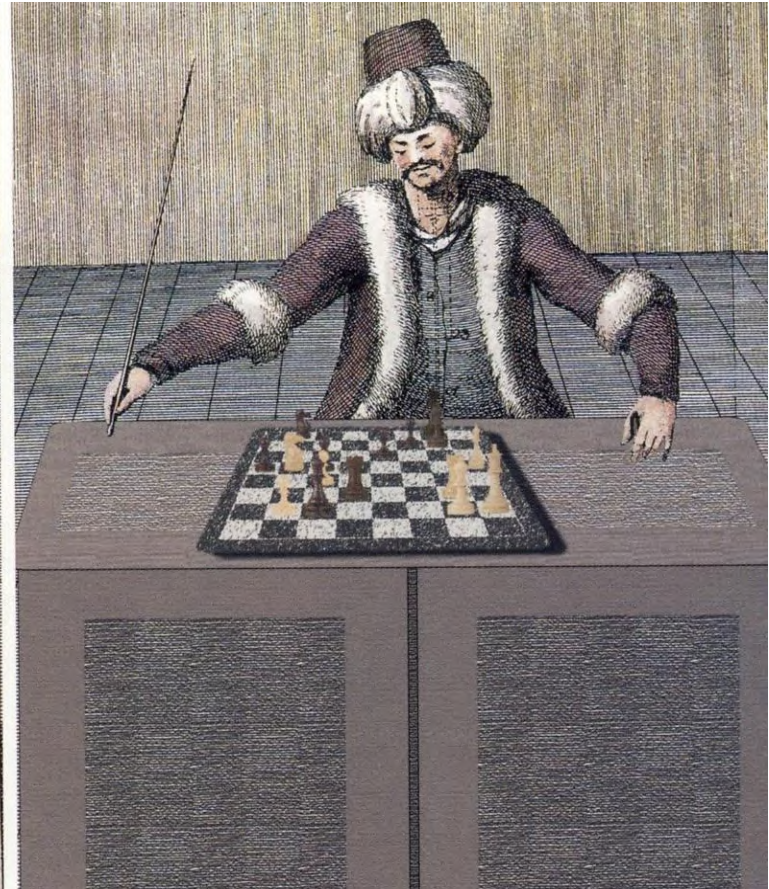
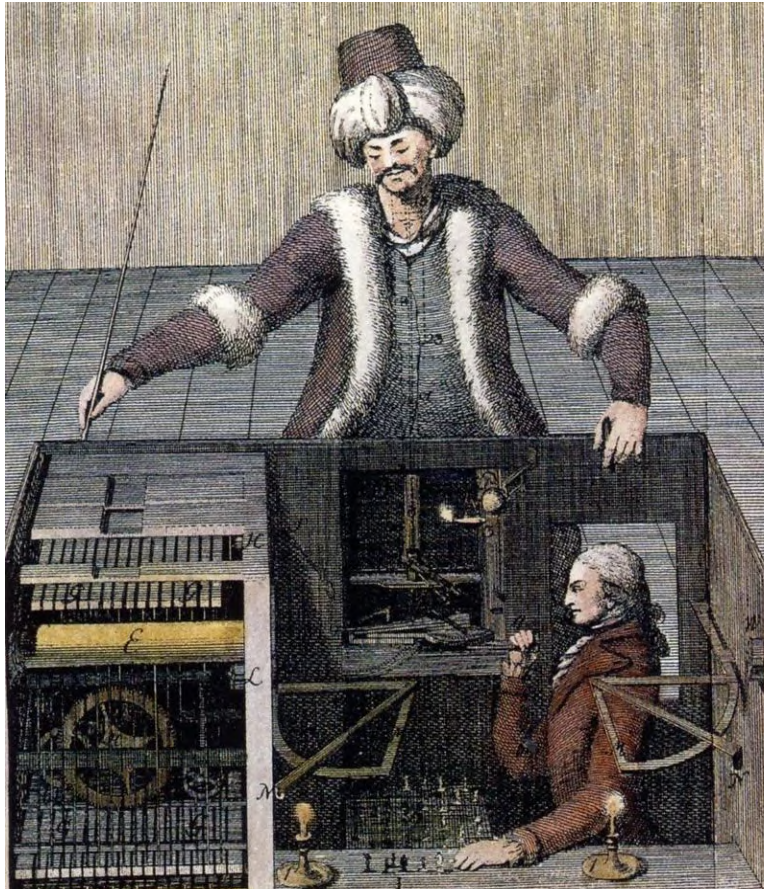


youtu.be/Aut32pR5PQA

Sawubona

Reinforcement Learning Evolution

Mechanical Turk (W.Kempelen, 1770)



Reinforcement Learning Evolution

Checkers (Arthur Samuel, IBM 1956)



Reinforcement Learning Evolution

Chess (DeepBlue, IBM 1997)



Reinforcement Learning Evolution

GO (AlphaGo, Google 2016)



Reinforcement Learning Evolution

Chess (AlphaZero, Google 2017)

20 years after Deep Blue...

a machine learned chess all by itself...
and crushed Stockfish



chess24.com
@chess24com

Following

Now the era of computer chess engine programming also seems to be over: AlphaZero, developed by @DeepMindAI & @demishassabis, took just 4 hours playing against itself to learn to play better than Stockfish (it won 64:36)! Replay 10 example games: [chess24.com/en/watch/live- ...](https://chess24.com/en/watch/live-...) #c24live



8:52 AM - 6 Dec 2017

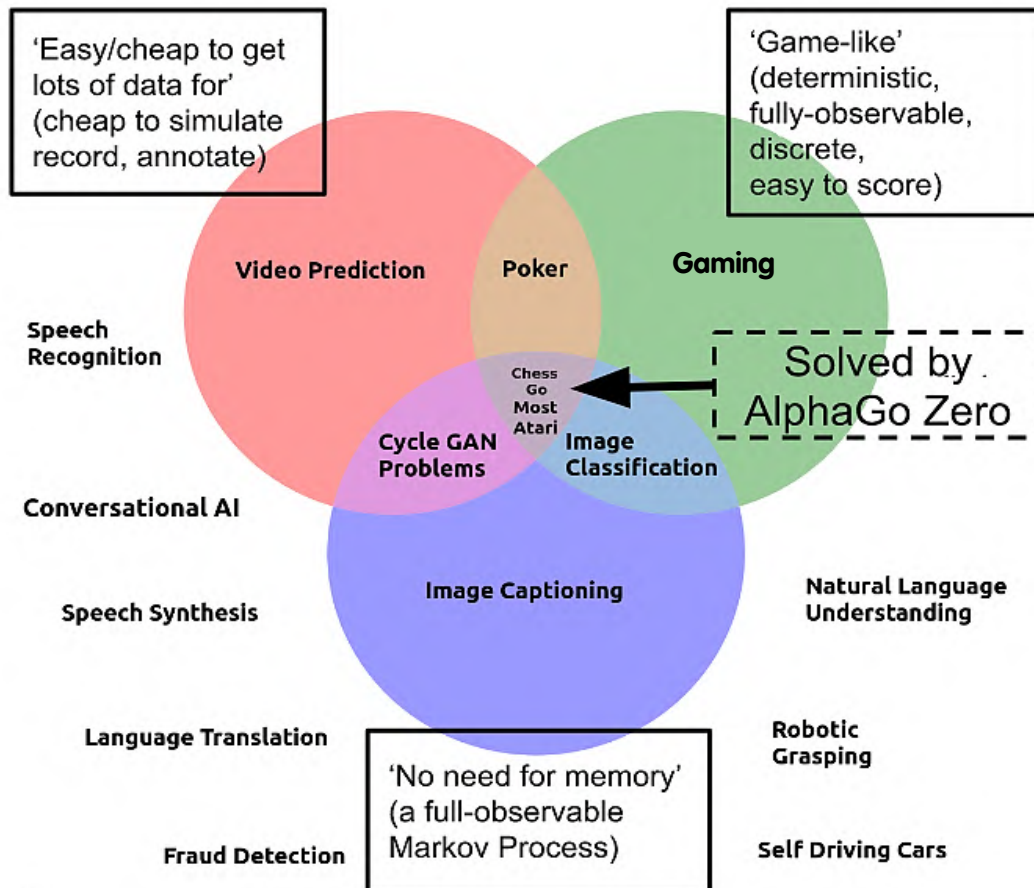
Reinforcement Learning Evolution

Go (AlphaGo Zero, Google 2017)



Reinforcement Learning

The AI Problem Complexity



Reinforcement Learning in Gaming

The AI Taught Itself ...



openai.com

Sawubona

Reinforcement Learning in Gaming

And... the AI Cheated



vkrakovna.wordpress.com/

Reinforcement Learning in Robotics

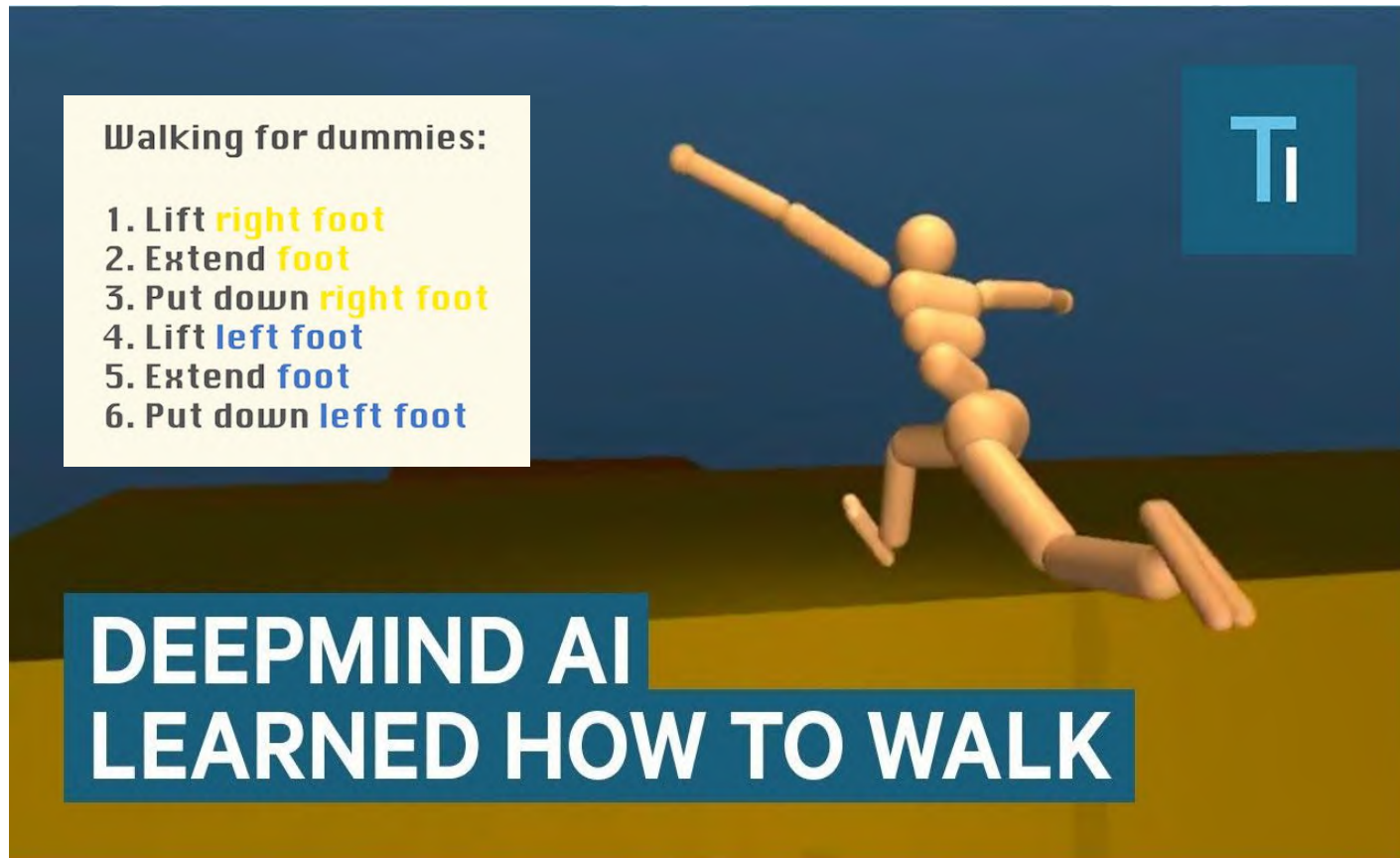
Example: DARPA Challenge (2015)



youtu.be/g0TaYhjpOfo
Sawubona

Reinforcement Learning in Robotics

Example: The AI Taught Itself How to Walk



youtu.be/gn4nRCC9TwQ

Reinforcement Learning in Robotics

Example: Boston Dynamics – Atlas (2018)



youtu.be/LikxFZZO2sk

Sawubona

Reinforcement Learning in Robotics

Example: Boston Dynamics – Atlas (2019)



youtu.be/_sBBaNYex3E

Sawubona



12:06

FLIGHT	STATUS	DESTINATION	DEPARTURE	ARRIVAL	GATE
DL 100	ON TIME	NEW YORK	12:00	12:30	A1
AA 100	DELAYED	LOS ANGELES	12:05	12:45	A2
UA 100	ON TIME	CHICAGO	12:10	12:40	A3
SW 100	ON TIME	MIAMI	12:15	12:45	A4
DL 200	ON TIME	ATLANTA	12:20	12:50	A5
AA 200	ON TIME	DALLAS	12:25	12:55	A6
UA 200	ON TIME	PHOENIX	12:30	13:00	A7
SW 200	ON TIME	SEATTLE	12:35	13:05	A8
DL 300	ON TIME	PHOENIX	12:40	13:10	A9
AA 300	ON TIME	PHOENIX	12:45	13:15	A10
UA 300	ON TIME	PHOENIX	12:50	13:20	A11
SW 300	ON TIME	PHOENIX	12:55	13:25	A12
DL 400	ON TIME	PHOENIX	13:00	13:30	A13
AA 400	ON TIME	PHOENIX	13:05	13:35	A14
UA 400	ON TIME	PHOENIX	13:10	13:40	A15
SW 400	ON TIME	PHOENIX	13:15	13:45	A16
DL 500	ON TIME	PHOENIX	13:20	13:50	A17
AA 500	ON TIME	PHOENIX	13:25	13:55	A18
UA 500	ON TIME	PHOENIX	13:30	14:00	A19
SW 500	ON TIME	PHOENIX	13:35	14:05	A20

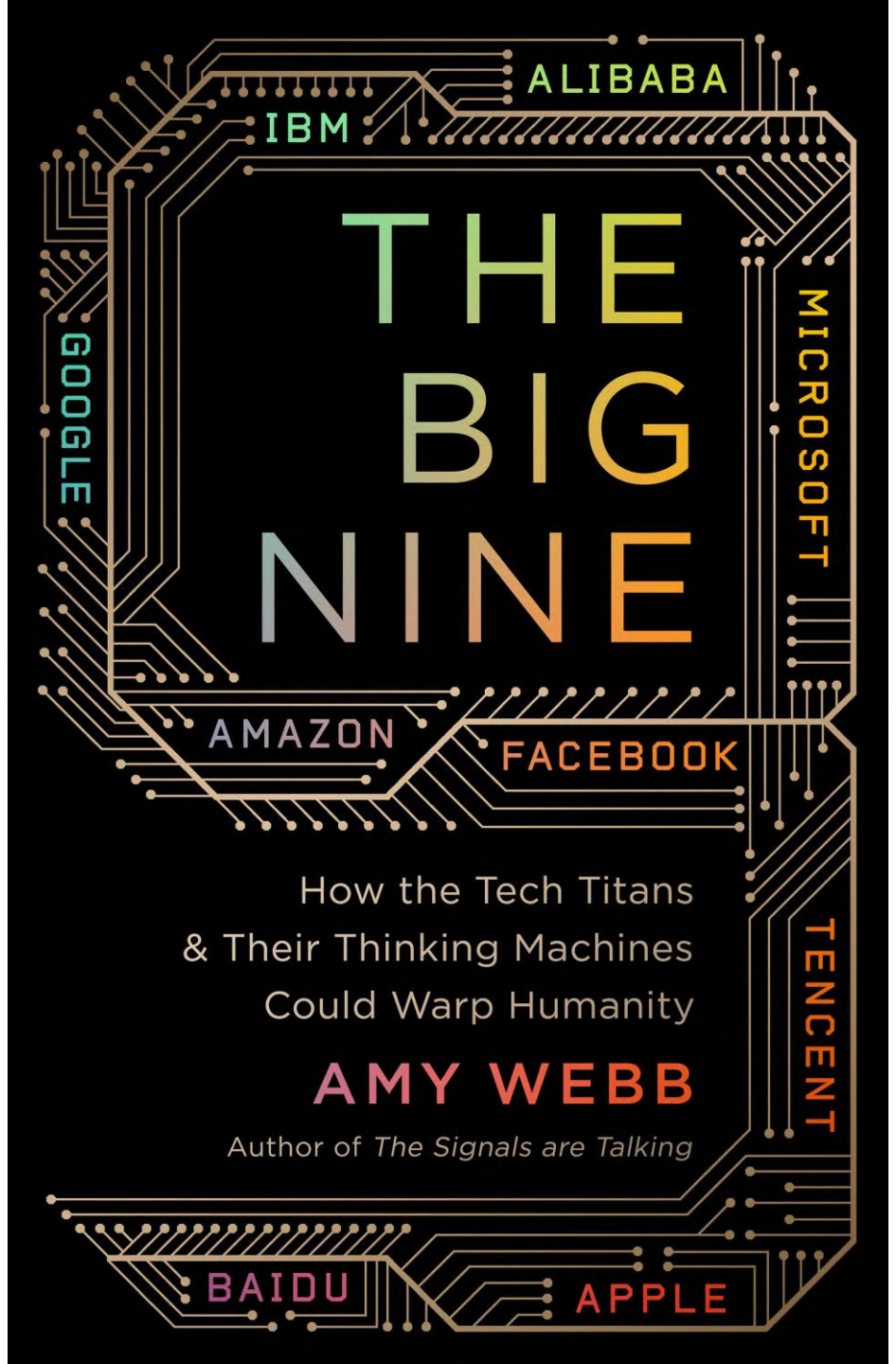
Opportunities

Chancen, Opportunitäten und Risiken der KI

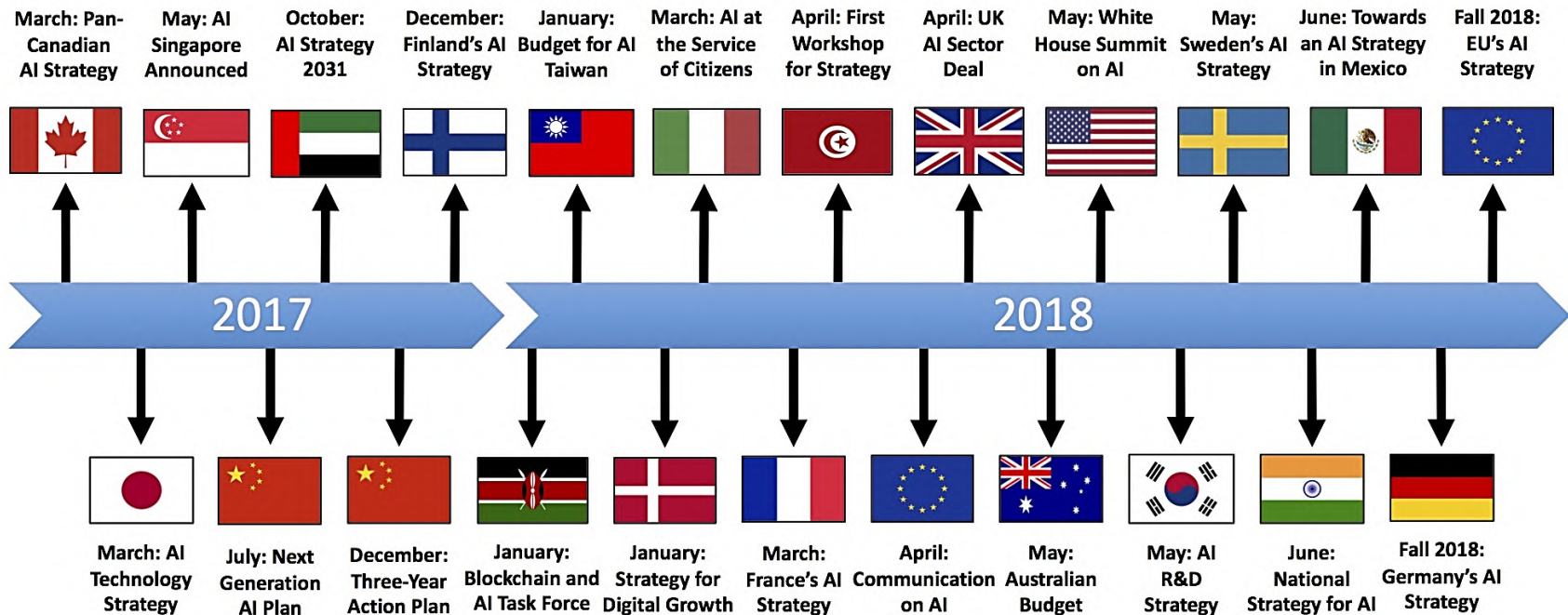
JCDecaux Airport

FAST TRACK

Is AI Only Accessible To Tech Giants?



National AI Strategies



2018-07-13 | Politics + AI | Tim Dutton

No two strategies are alike, with each focusing on different aspects of AI policy: scientific research, talent development, skills and education, public and private sector adoption, ethics and inclusion, standards and regulations, and data and digital infrastructure.

medium.com/politics-ai/an-overview-of-national-ai-strategies-2a70ec6edfd

Quo Vadis Switzerland?



«In Verwaltungen werden durch KI künftig sehr viele Entscheidungsprozesse automatisiert, die bisher von Menschen erledigt wurden oder so komplex sind, dass sie gar nicht ausgeführt werden konnten.

Verwaltungsmitarbeitende werden nicht einfach ersetzt, sondern wechseln die Aufgaben: In Zukunft werden sie hauptsächlich für besonders wichtige oder ethisch herausfordernde Fälle sowie für die Qualitätskontrolle zuständig sein.

Den Rest erledigt die Maschine; sie wird entscheiden, wie die von Menschen vorgegebene Strategie am besten erreicht werden kann.

Mittels «Machine Learning» wird eine «persönliche» Betreuung von Bürgeranliegen durch Bots (Roboter) möglich – die Technologie lernt mit jedem Bürgerkontakt dazu und wird automatisch immer besser, je mehr Daten sie verarbeitet.»

Künstliche Intelligenz: Die Zukunft der Verwaltung
Nicola Forster, lic.iur./MLaw UZH, Gründer und Präsident des Think Tank foraus, Innovations-Experte und Moderator (crstl.io), Co-Founder Staatslabor, Advisory Board Open Forum (World Economic Forum).

Challenges for Government Adoption of AI

Barriers to AI Adoption in Government and Public Entities

Effective Use
Of Data

Data and AI
Skills

Legacy
Culture

AI
Ecosystem

Procurement
Mechanism

www.weforum.org/agenda/2019/08/artificial-intelligence-government-public-sector/ (Aug'2019)

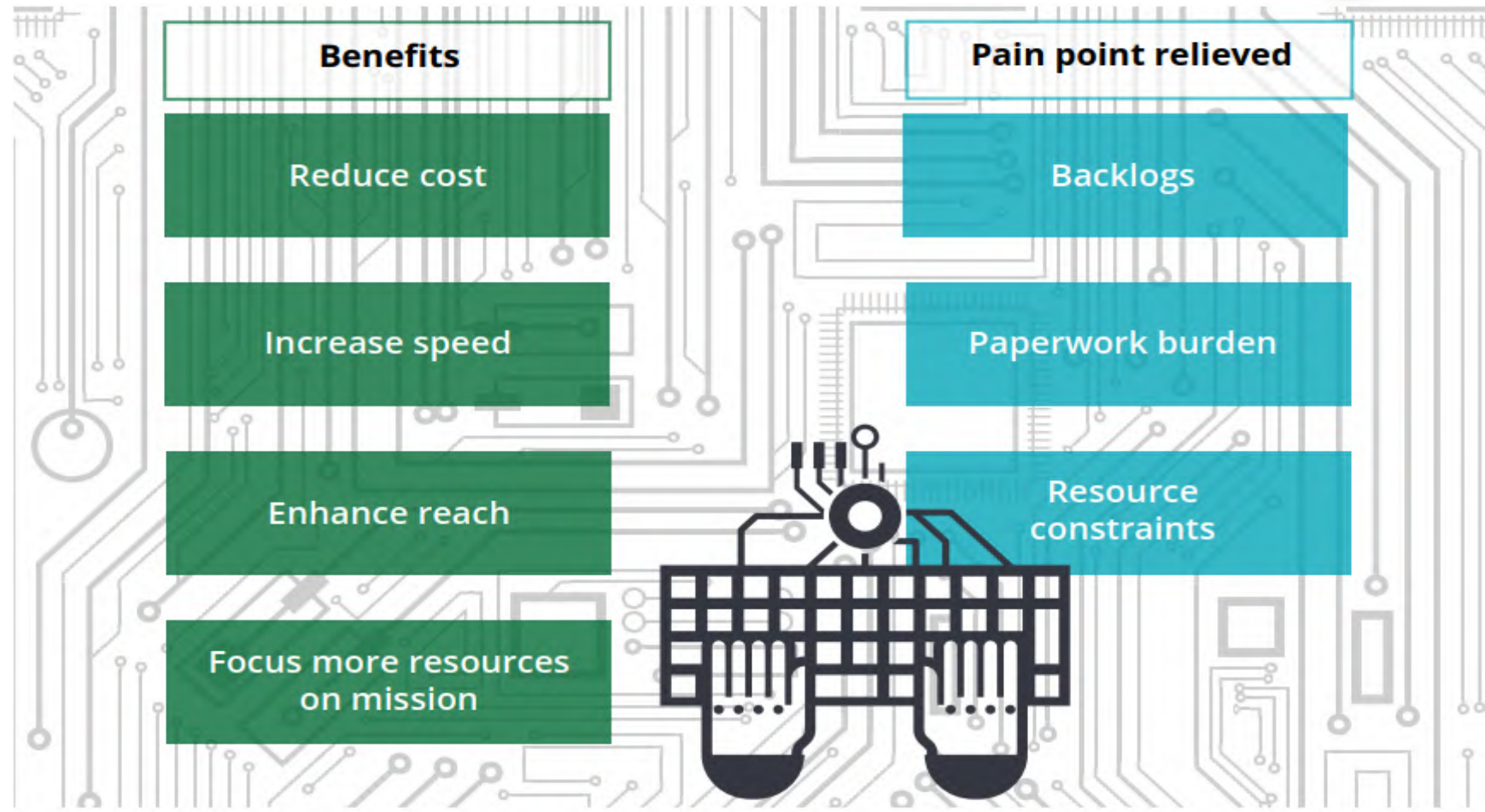
AI Augmented Government

Einsatz von KI im Öffentlichen Sektor





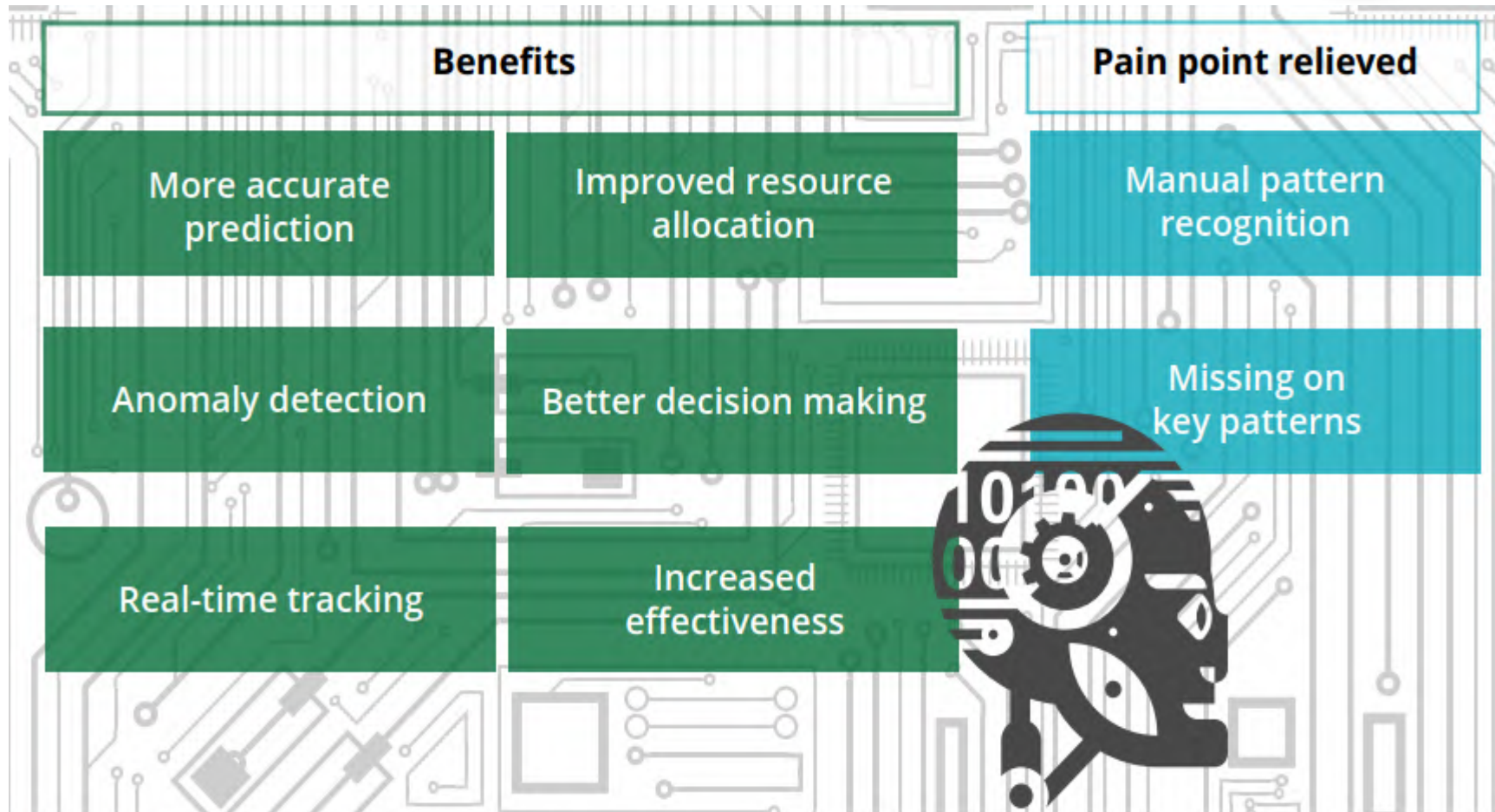
Robotic Process Automation



Source: Deloitte analysis.



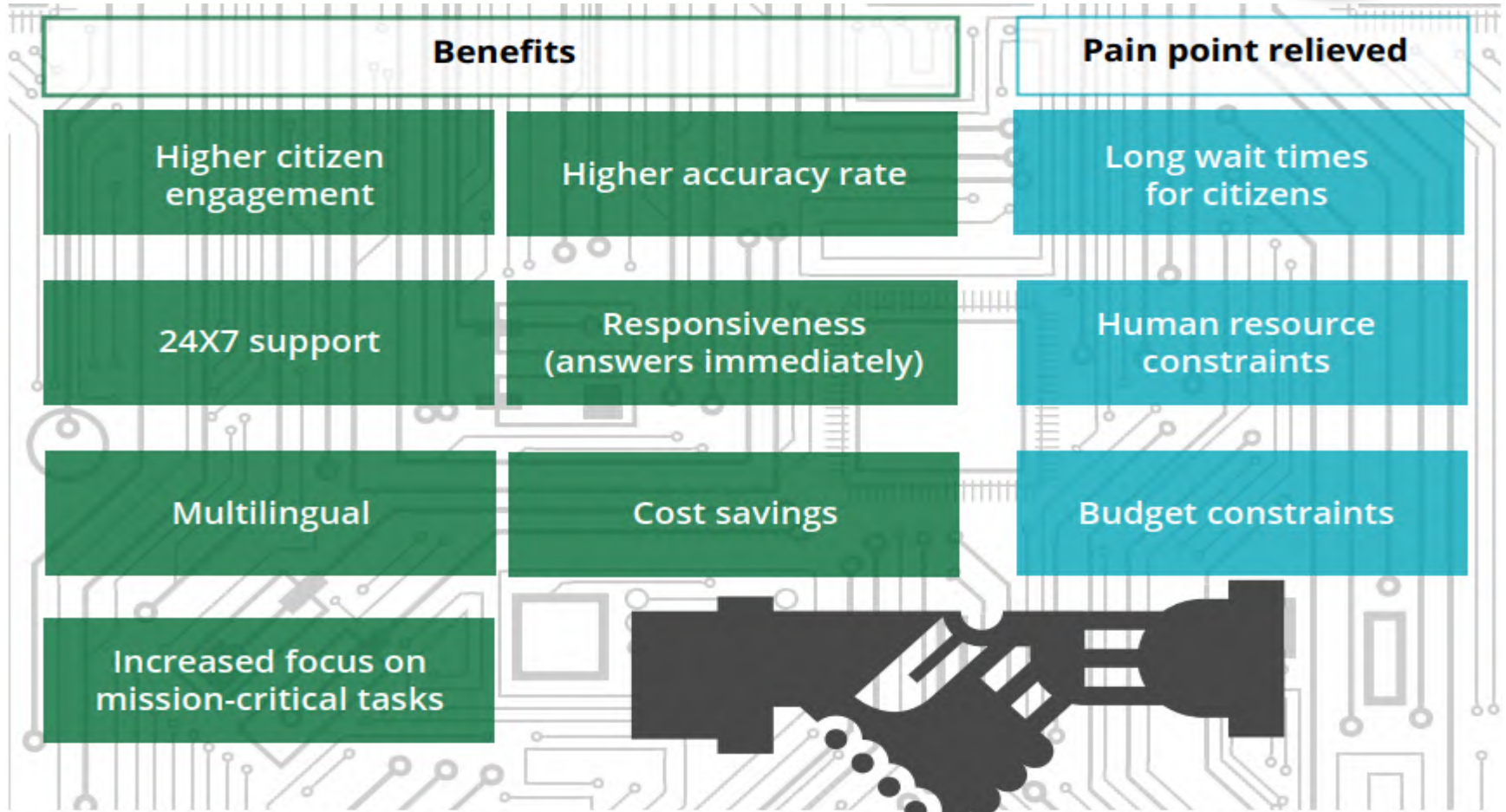
Cognitive Insight Applications



Source: Deloitte analysis.



Citizen Engagement Applications



Source: Deloitte analysis.

Outlook

Ausblick



The Next The Next Challenges



**Explainable
AI**



**Safe
AI**



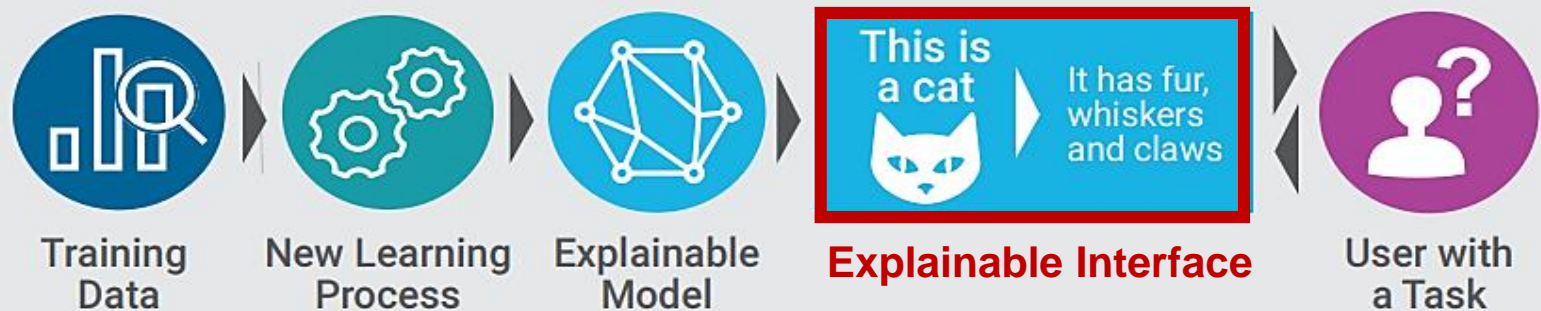
**Ethical
AI**

The Next Challenge ⇒ **Explainable AI**

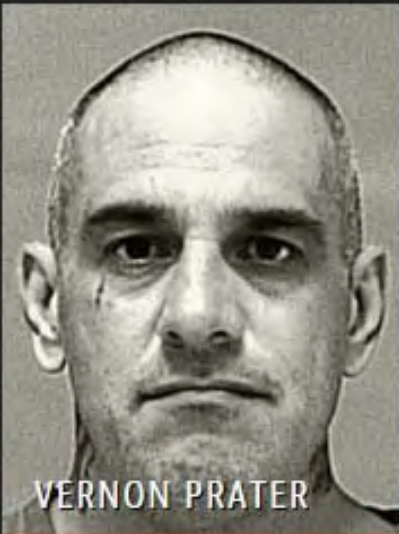



TODAY



TOMORROW



The Next Challenge ⇒ **Explainable AI**

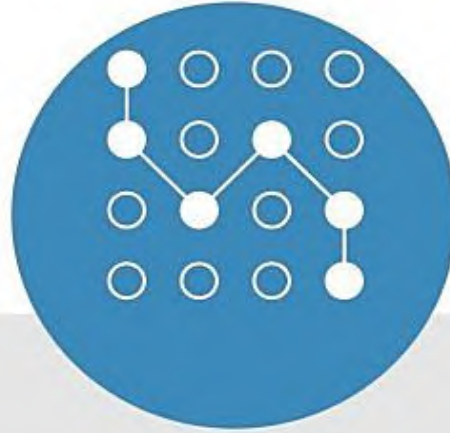
Two Petty Theft Arrests		Two Petty Theft Arrests	
			
VERNON PRATER	BRISHA BORDEN	VERNON PRATER	BRISHA BORDEN
LOW RISK	HIGH RISK	LOW RISK	HIGH RISK
3	8	3	8
<p><i>Borden was rated high risk for future crime after she and a friend took a kid's bike and scooter that were sitting outside. She did not reoffend.</i></p>		<p><i>Borden was rated high risk for future crime after she and a friend took a kid's bike and scooter that were sitting outside. She did not reoffend.</i></p>	

The Next Challenge ⇒ **Explainable AI**



Explainable Data

What data was used to train the model and why?



Explainable Predictions

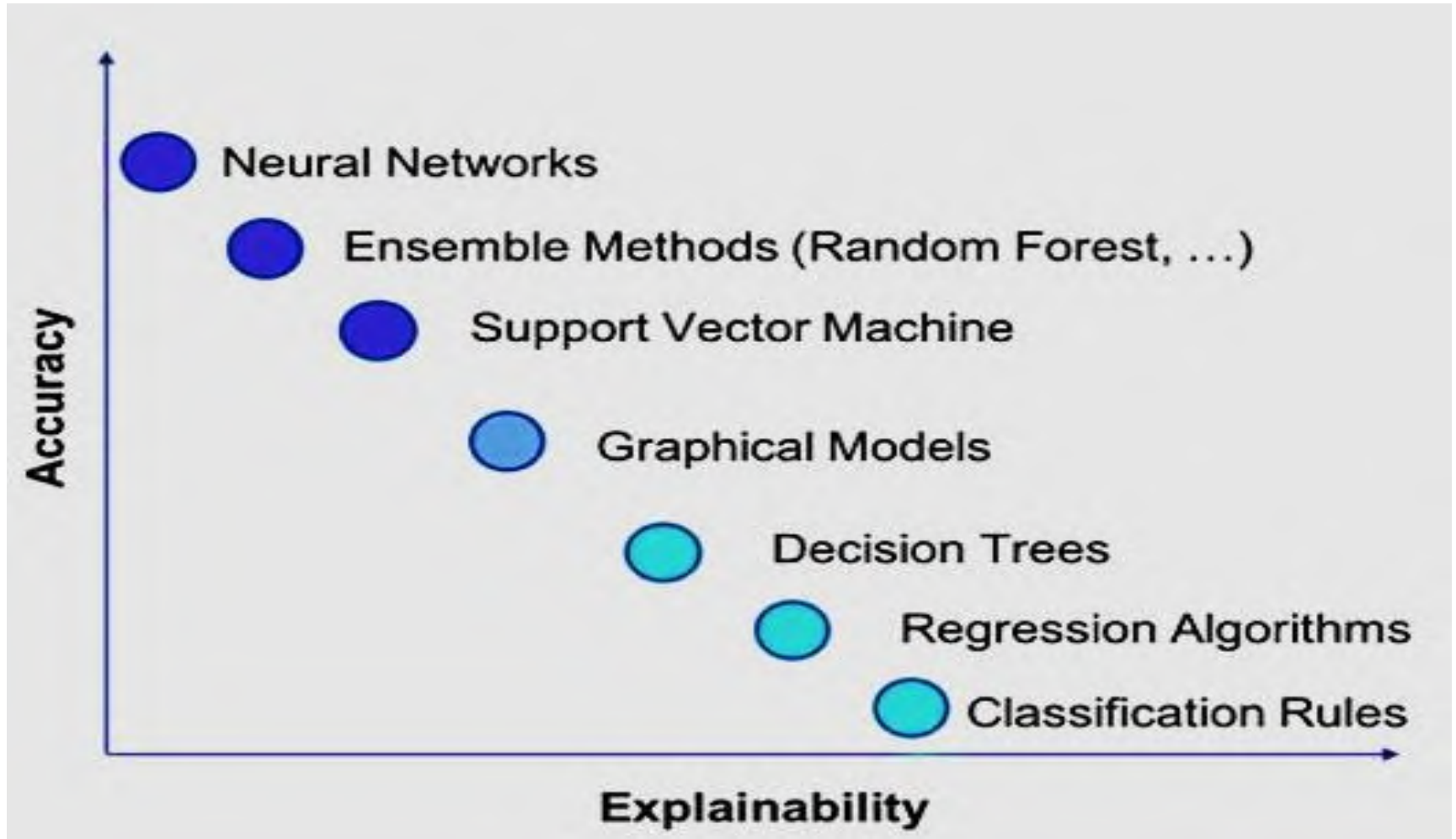
What features and weights were used for this particular prediction?



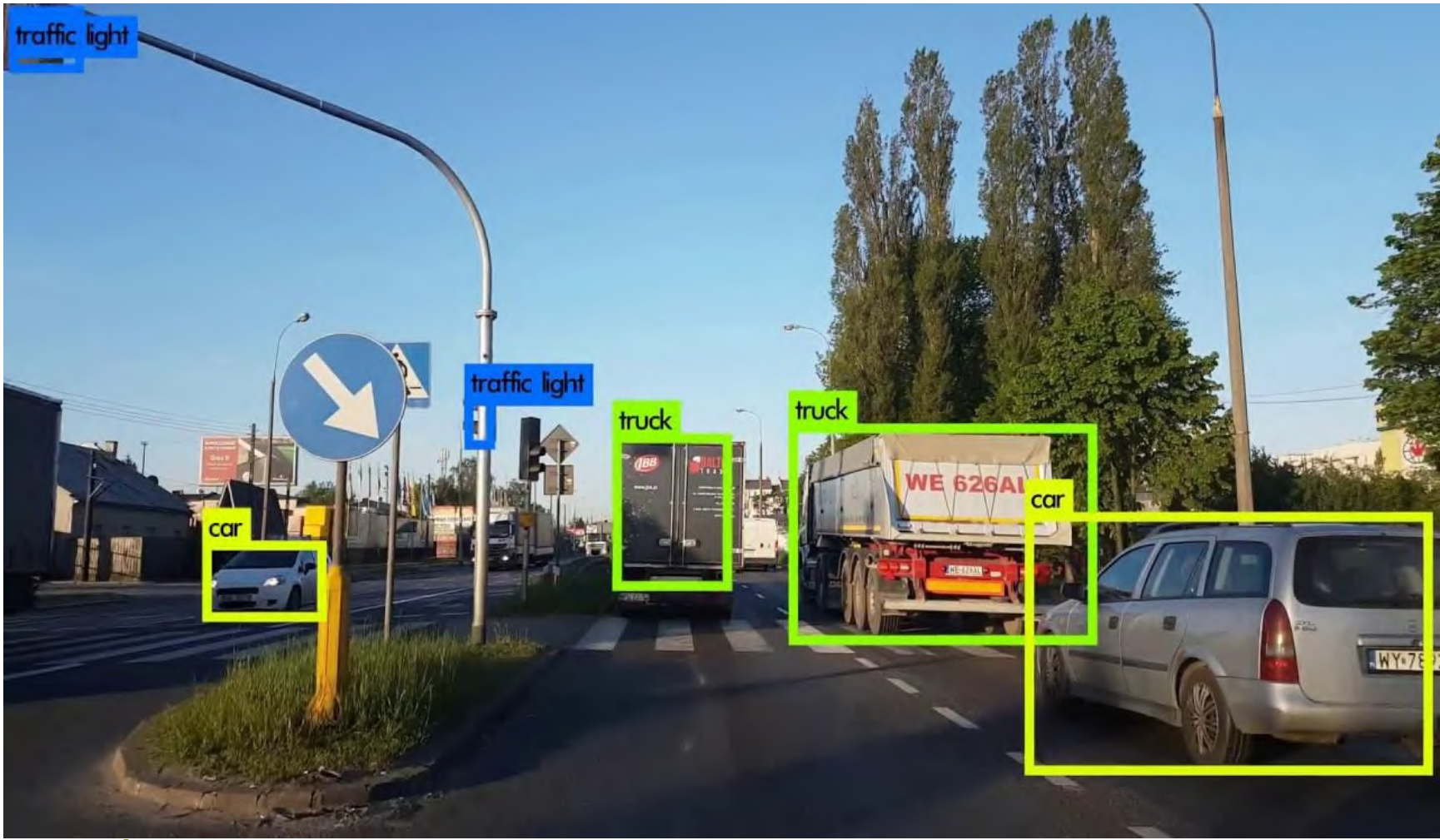
Explainable Algorithms

What are the individual layers and the thresholds used for a prediction?

The Next Challenge \Rightarrow **Explainable AI**

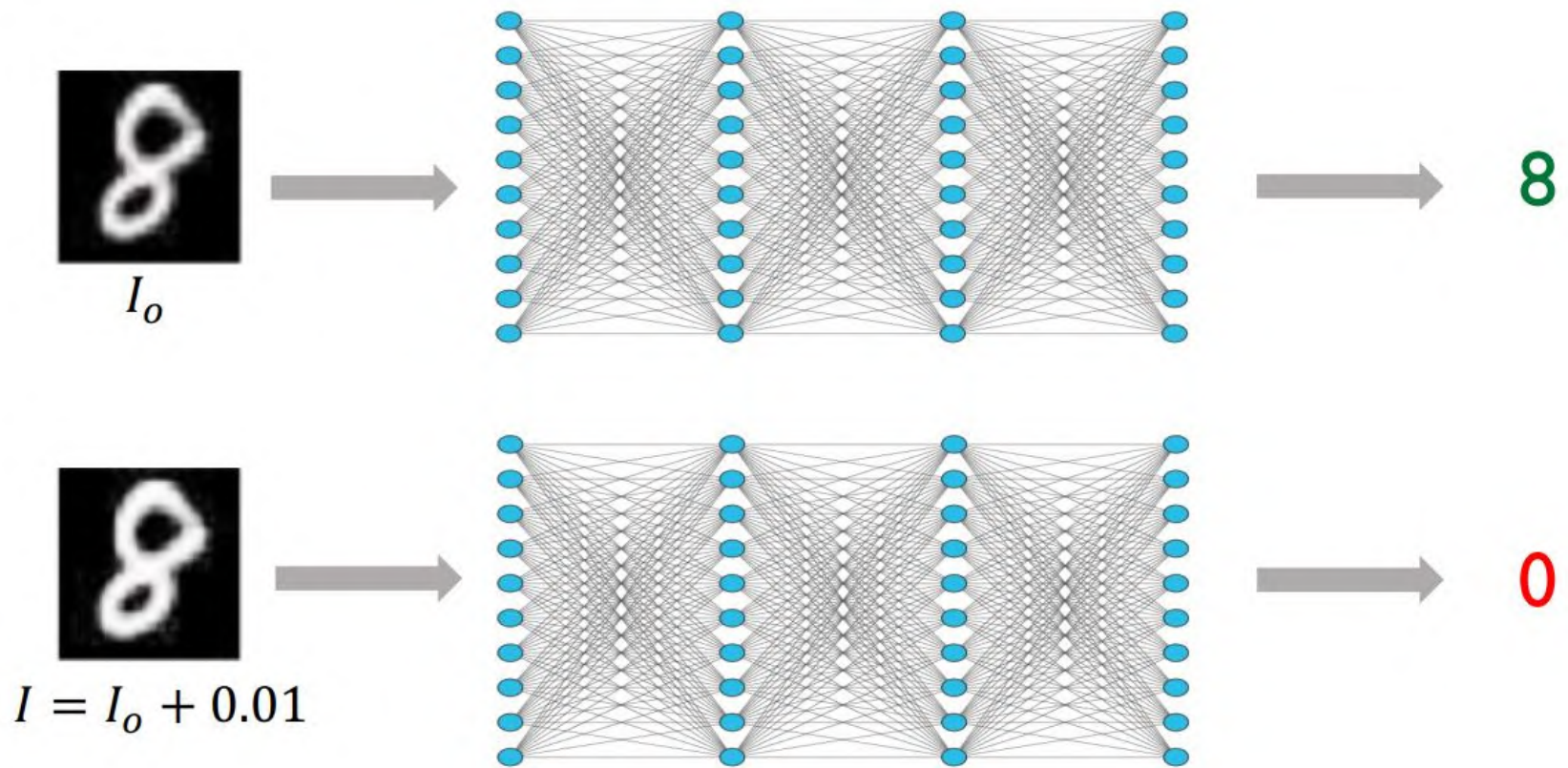


The Next Challenge ⇒ Safe AI



The Next Challenge \Rightarrow **Safe AI**

Attacks on Deeplearning



The Next Challenge \Rightarrow Safe AI

Explaining and Harnessing Adversarial Examples

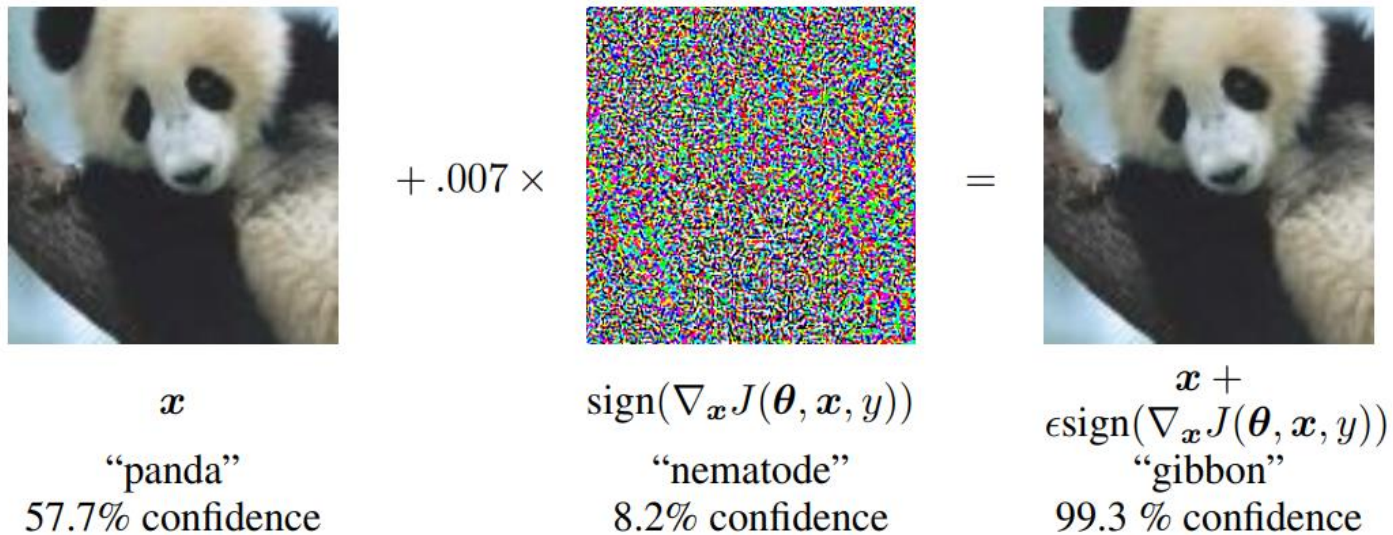
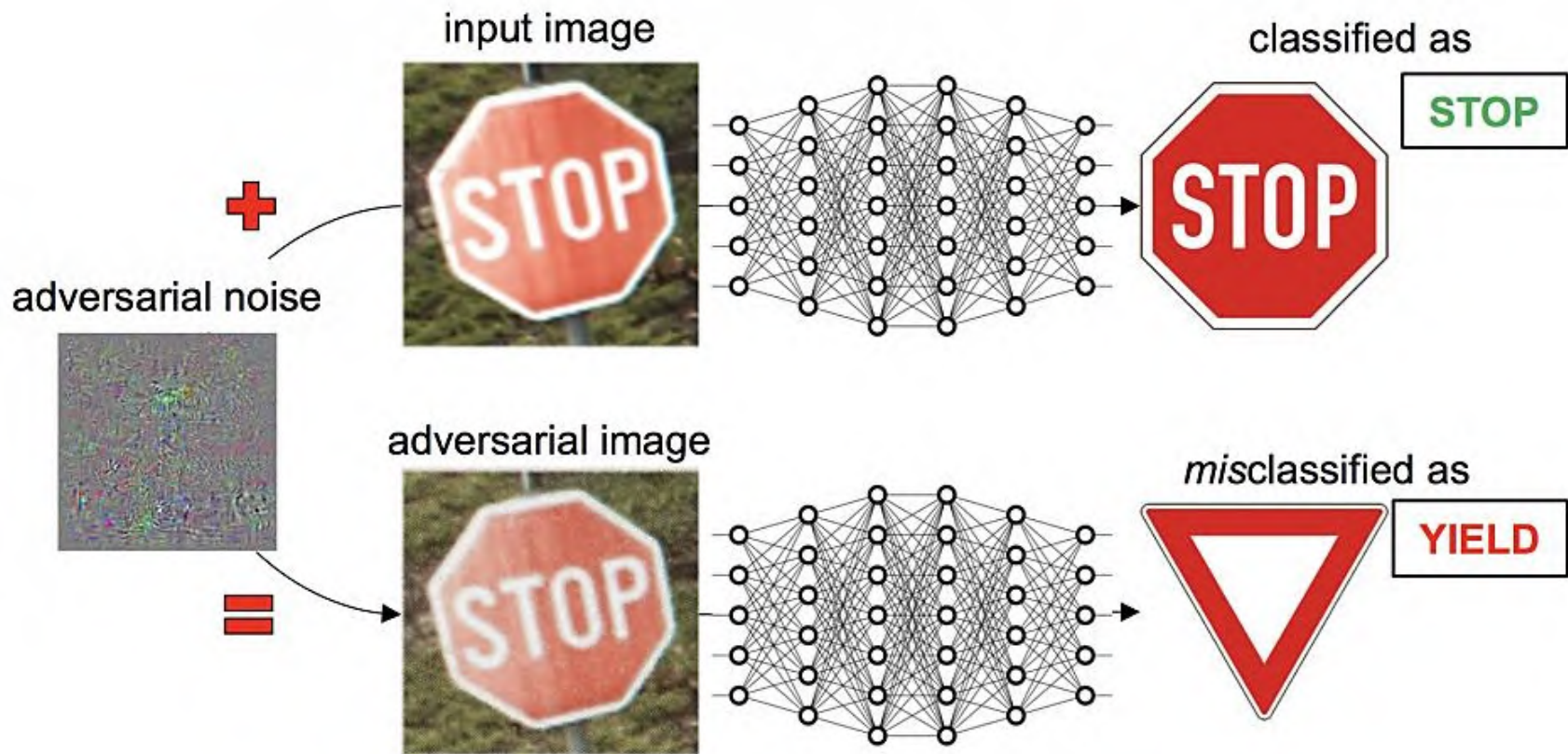


Figure 1: A demonstration of fast adversarial example generation applied to GoogLeNet (Szegedy et al., 2014a) on ImageNet. By adding an imperceptibly small vector whose elements are equal to the sign of the elements of the gradient of the cost function with respect to the input, we can change GoogLeNet’s classification of the image. Here our ϵ of .007 corresponds to the magnitude of the smallest bit of an 8 bit image encoding after GoogLeNet’s conversion to real numbers.

Adversarial Examples: Attacks and Defenses for Deep Learning (<https://arxiv.org/pdf/1712.07107.pdf>, July 2018)

The Next Challenge \Rightarrow **Safe AI**

Adversarial Examples



<https://blog.csiro.au/vaccinating-machine-learning-against-attacks/>

The Next Challenge ⇒ Safe AI

Testing of Deep Learning Systems



(a) Input 1



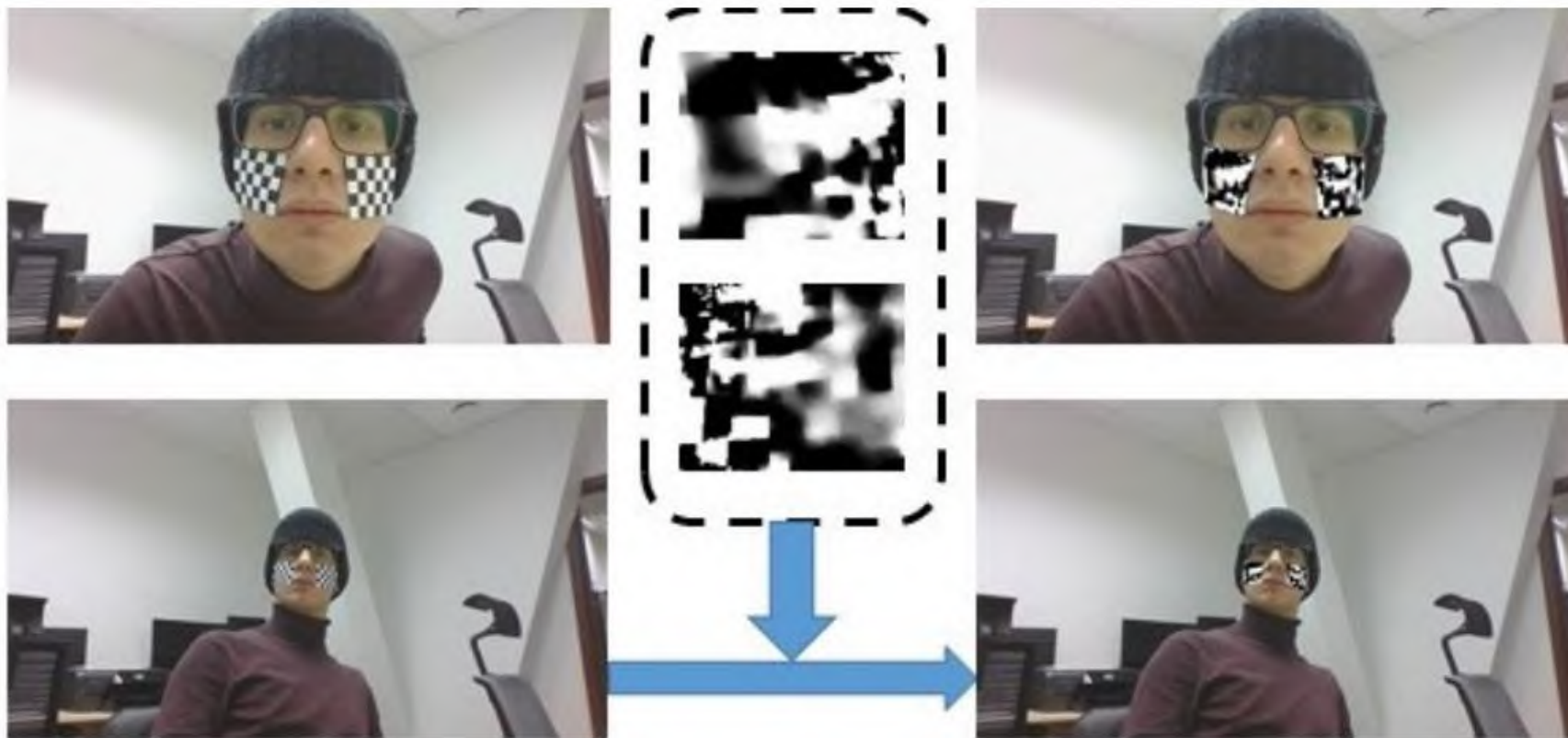
(b) Input 2 (darker version of 1)

Figure 1: An example erroneous behavior found by DeepXplore in Nvidia DAVE-2 self-driving car platform. The DNN-based self-driving car correctly decides to turn left for image (a) but incorrectly decides to turn right and crashes into the guardrail for image (b), a slightly darker version of (a).

DeepXplore: Automated Whitebox Testing of Deep Learning Systems (<https://arxiv.org/pdf/1705.06640.pdf>, Sept'2017)

The Next Challenge ⇒ **Safe AI**

Example: Defeat Facial Recognition



Real-world attack on MTCNN face detection system (<https://arxiv.org/pdf/1910.06261.pdf>, Oct'2019)

The Next Challenge ⇒ Ethical AI

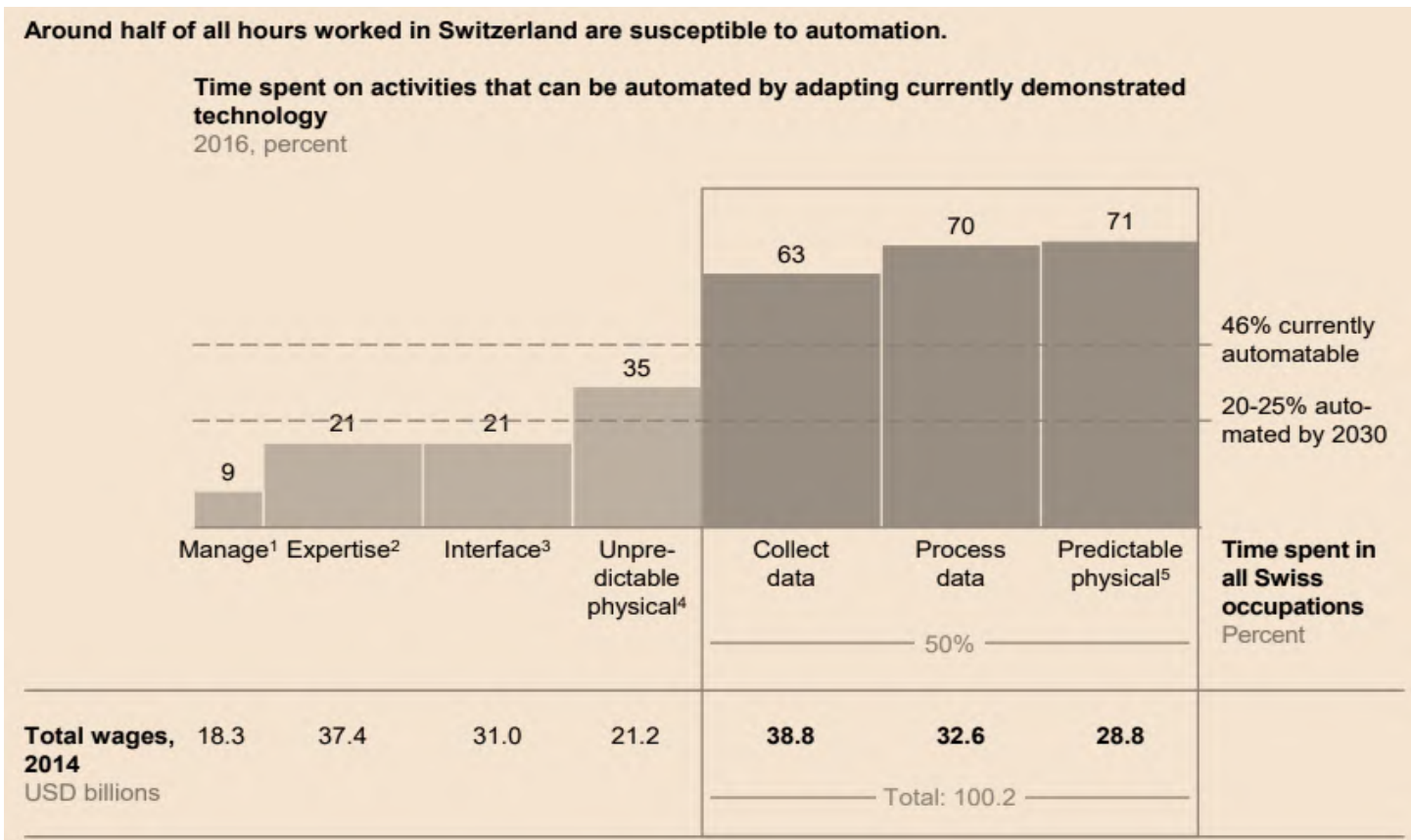
Example: The Ethical Dilemma of Self-Driving Cars



www.govtech.com/fs/automation/The-Everyday-Ethical-Challenges-of-Self-Driving-Cars.html

The Next Challenge ⇒ Ethical AI

The Future Of Work: Switzerland's Digital Opportunity, Oct'2018



www.digitaleschweiz.ch/wp-content/uploads/2018/10/The-future-of-work-Switzerlands-digital-opportunity.pdf

The Next Challenge ⇒ Ethical AI



The Stanford Human-Centered AI Initiative (HAI)

- For AI to better serve our needs, it **must incorporate** more of the versatility, nuance, and **depth of the human intellect**.
- The **development** of AI should be **paired with** an ongoing **study of its impact on human society** and guided accordingly.
- The ultimate purpose of AI should be to **enhance our humanity**, not diminish or replace it.

hai.stanford.edu/

Sawubona

“Nobody phrases it this way, but I think that artificial intelligence is almost a humanities discipline. It's really an attempt to understand human intelligence and human cognition.

Sebastian Thrun







Wieder lernen zu lernen!



**collaborative
innovative
thinking**

Daniel Benninger - Kurzporträt

2018-		VR CVCube AG Gründungsmitglied/Vize-Präsident Digital Academy Switzerland
2015-		Managing Partner (Sawubona GmbH) Experte (Berner Fachhochschule), Reviewer MDPI Open Access Journals)
2014-		VR / Direktor ad interim (Sept'2016-Febr'2018) Schweiz. Teletext AG Dozent/Studienleiter (Hochschule Luzern)
2012-2015		Leiter ICT Eidg. Finanzmarktaufsicht FINMA
2008-2012		Head of IT Compliance/ICT Manager Swisscom (Schweiz) AG
2006-2008		Head of IT Management & Architecture Swisscom Fixnet AG
2003-2006		CEO db=ctq
2000-2003		Leiter IT Architektur Schweiz. Mobiliar Versicherungsgesellschaft
1998-2000		Head of Integration Solution Center Swisscom AG, Corporate Information & Technology
1994-1998		Project Manager, Solution Architect Digital Equipment Corporation (DEC)
1988-1994		Leiter Informatik / Partner, Sigmaplan AG
1978-1988		Dozent Software Schule Schweiz / Lehrauftrag Universität Bern Universität Bern, Studium Mathematik+Physik, Dr.phil.nat.



daniel.benninger@sawubona.ch
+41 79 334 54 70