Chancen und Grenzen künstlicher Intelligenz aus Sicht eines Hirnforschers

Prof. Dr. rer. nat. John-Dylan Haynes
Charité – Universitätsmedizin Berlin

Wirtschaft 〉 Künstliche Intelligenz 〉 Fortschritt durch künstliche Intelligenz: Es wird unser Leben ändern

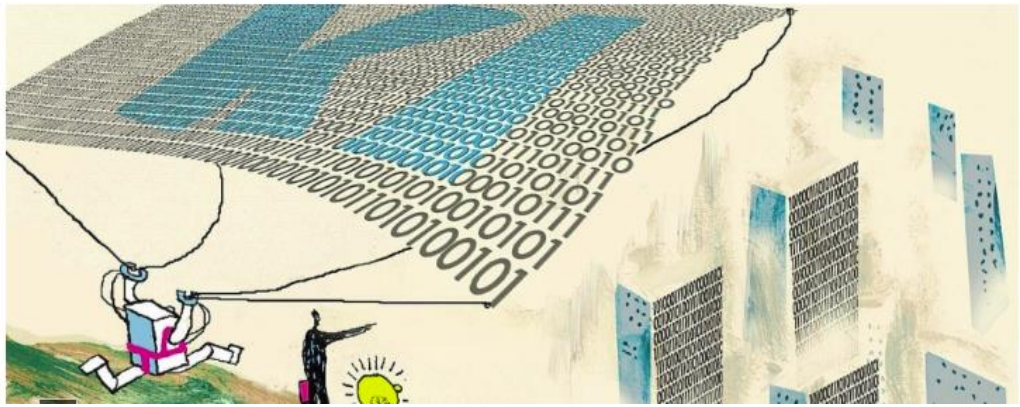
Frankfurter Allgemeine

Künstliche Intelligenz

MIT DEM FORTSCHRITT LEBEN

KI wird unser Leben verändern

VON MONIKA SCHNITZER UND MARTIN WATZINGER - AKTUALISIERT AM 06.09.2018 - 15:45



THE HILL

Will the last human worker please turn out the lights?

BY MIKE RETTIG - 09/21/15 10:00 AM EDT

0 SHARES

Just In...

Lindsey Graham torches Mazie Hirono: 'The Hirono standard is horrific'
SENATE — 7M 11S AGO

UN wants Turkey, Saudi Arabia to investigate disappearance of Washington Post columnist
INTERNATIONAL — 17M 19S AGO

Stormy Daniels: Michael Cohen 'looked happy and healthy' in airport

CONTRIBUTOR'S SIGN
Sign up to become a Contri

Most Popular

1 Former FBI lawyer: F record...
3,473 SHARES

2 Trump, Kushner to n

The Economist | Topics | Current edition | More

Daily chart

A study finds nearly half of jobs are vulnerable to automation

CNN BUSINESS Markets Tech Media Success Video

Future Tense

Robots: Is your job at risk?

by Matt McFarland @mattmcfarland

September 15, 2017: 12:49 PM ET

SPIEGEL ONLINE SPIEGEL+

Menü | Politik Meinung Wirtschaft Panorama Sport Kultur Netzwelt Wissensc

WIRTSCHAFT

Schlagzeilen | DA

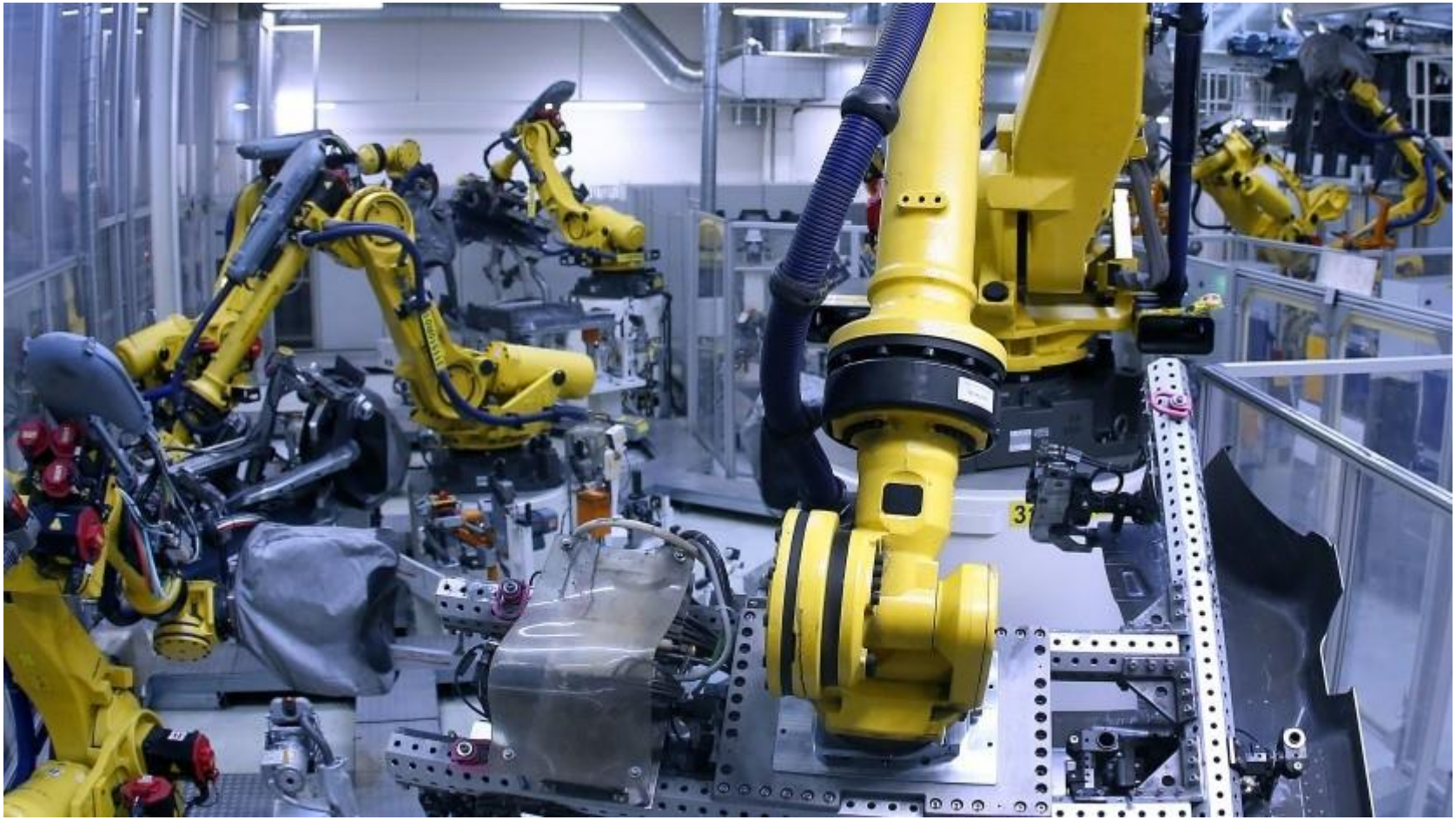
Nachrichten > Wirtschaft > Staat & Soziales > Arbeitsmarkt > Studie Weltwirtschaftsforum: Maschinen verrich

Studie des Weltwirtschaftsforums

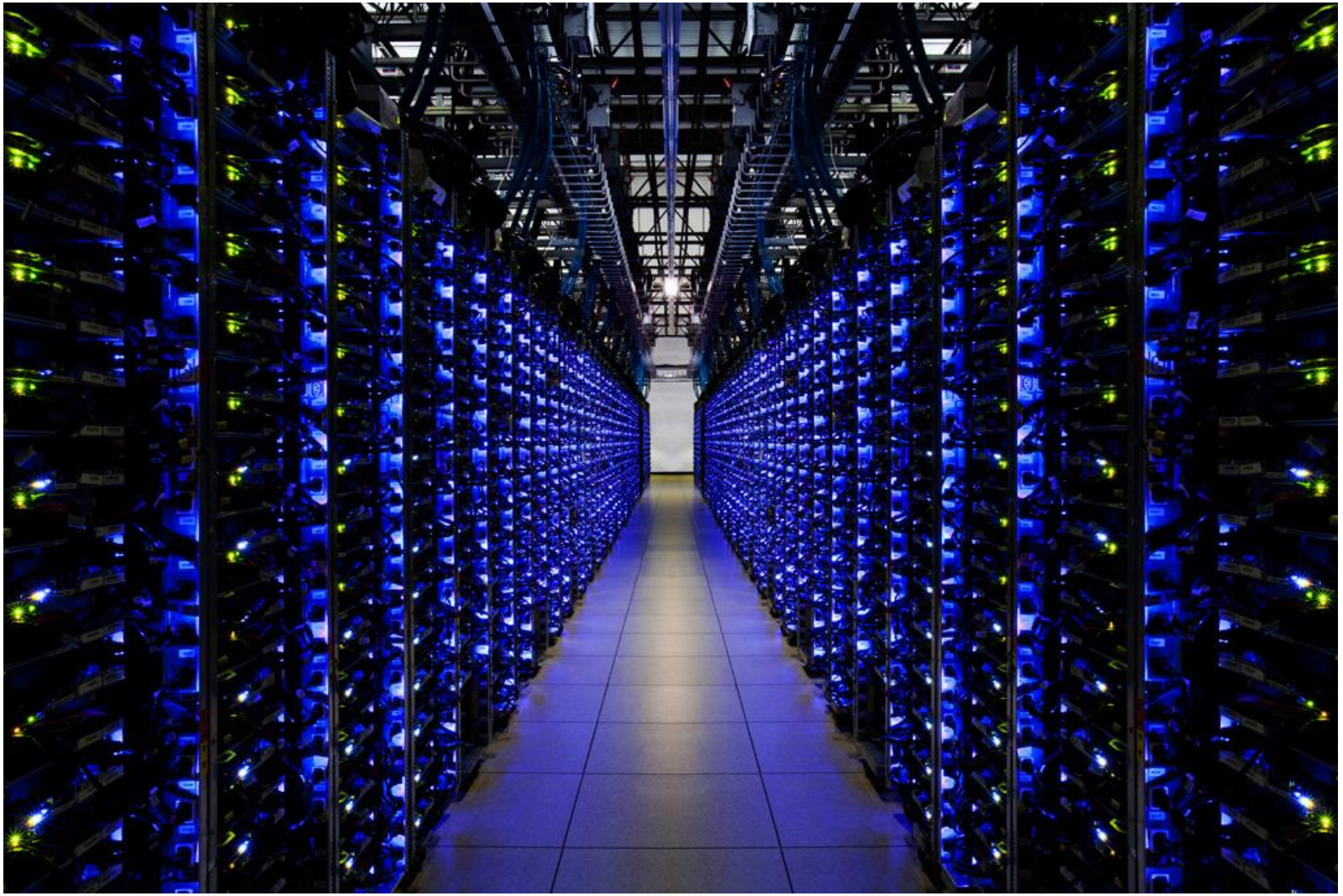
Maschinen verrichten bald mehr Arbeit als Menschen

Unsere Arbeitswelt wandelt sich radikal. Laut einer neuen Studie werden bis 2025 mehr Aufgaben von Robotern erledigt als von Menschen. Millionen Jobs werden dadurch wegfallen - aber noch mehr neue entstehen.

Automatisierung

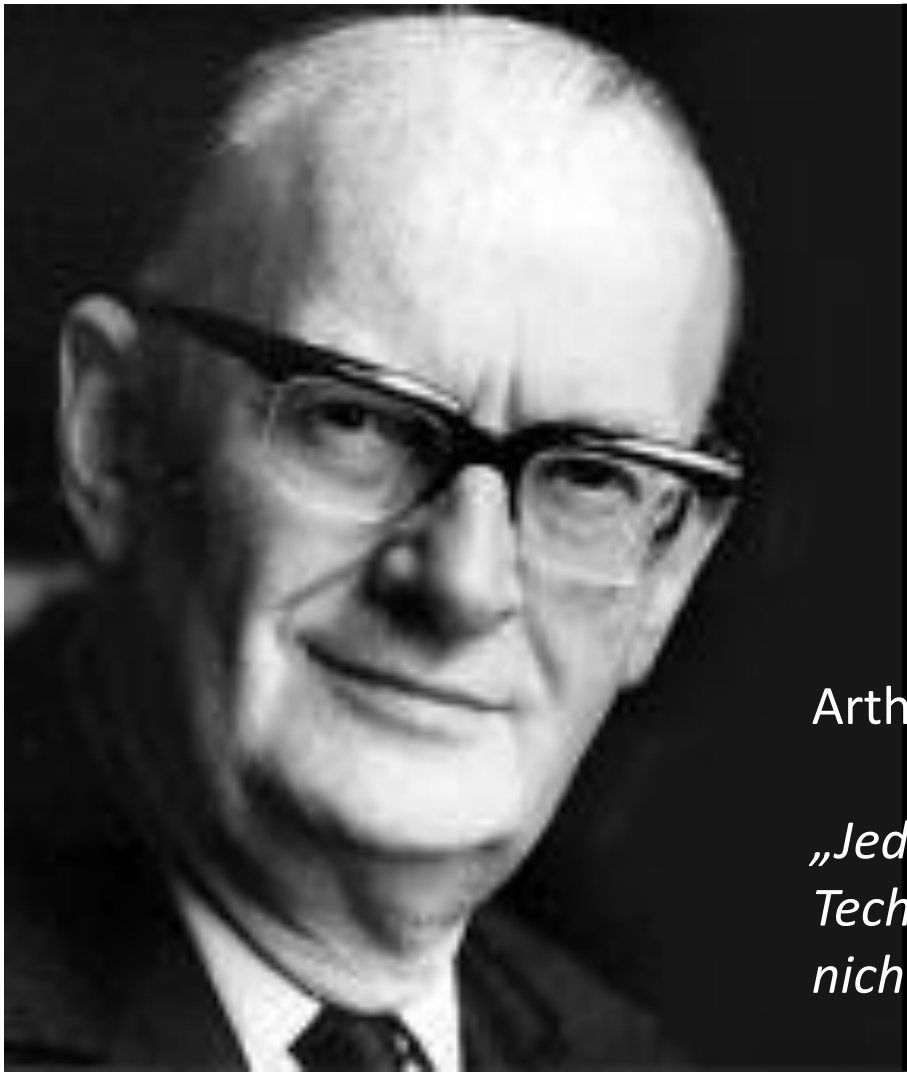


Algorithmen



Intelligenz

Clarkes 2. Gesetz

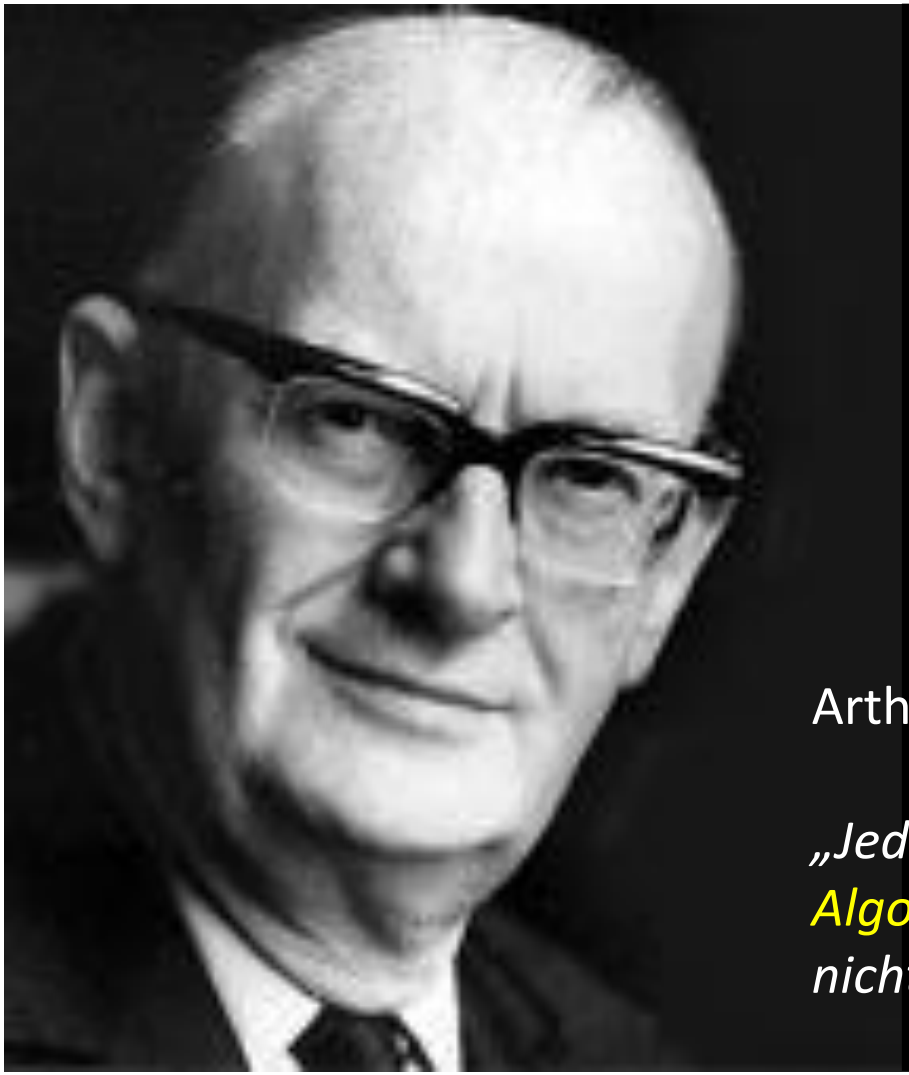


Arthur C. Clarke (1917-2008)

„Jede hinreichend fortschrittliche Technik ist von Magie nicht unterscheidbar.“

Arthur C. Clarke: *Profile der Zukunft: über die Grenzen des Möglichen*. Heyne, München 1984

Clarkes 2. Gesetz

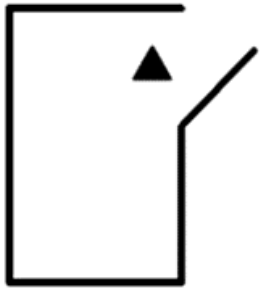


Arthur C. Clarke (1917-2008)

„Jeder hinreichend fortschrittliche Algorithmus ist von Intelligenz nicht unterscheidbar.“

Animismus

Heider & Simmel (1944)



Pareidolia

Apple Siri (2011)

Sony Aibo ERS 1000 (2018)

Definitionen von Intelligenz

Ca. 70 Definitionen

A Collection of Definitions of Intelligence

Shane Legg
IDSIA, Galleria 2, Manno-Lugano CH-6928, Switzerland
shane@idsia.ch www.idsia.ch/∼shane

Marcus Hutter
IDSIA, Galleria 2, Manno-Lugano CH-6928, Switzerland
RSISE/ANU/NICTA, Canberra, ACT, 0200, Australia
marcus@hutter1.net www.hutter1.net

October 4, 2006

"[…] intelligence is simply what the tests of intelligence test […]"

(Edwin G. Boring. New Republic 36 (1923): 35-37)

Definition der APA Arbeitsgruppe (1986)

„Individuen unterscheiden sich voneinander in ihrer Fähigkeit, komplexe Ideen zu verstehen, sich effektiv an die Umwelt anzupassen, aus Erfahrung zu lernen, zu Denken und Probleme zu lösen […]“

APA Task force / Neisser et al. Am Psy 1996, p77

Definition der APA Arbeitsgruppe (1986)

„Individuen unterscheiden sich voneinander in ihrer Fähigkeit, komplexe Ideen zu verstehen, sich effektiv an die Umwelt anzupassen, aus Erfahrung zu lernen, zu Denken und Probleme zu lösen […]“

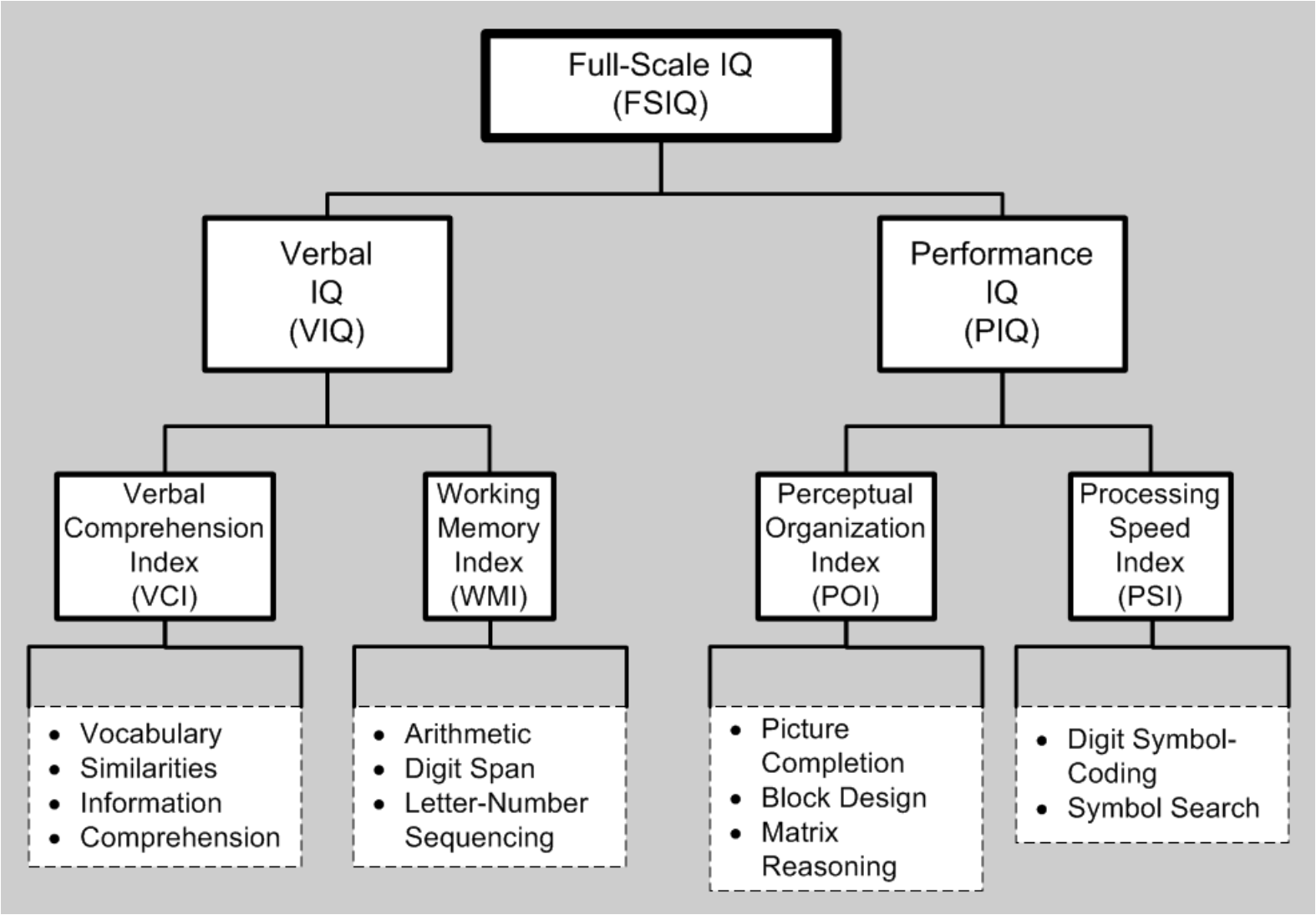
APA Task force / Neisser et al. Am Psy 1996, p77

Definition der APA Arbeitsgruppe (1986)

„Individuen unterscheiden sich voneinander in ihrer Fähigkeit, komplexe Ideen zu verstehen, sich effektiv an die Umwelt anzupassen, aus Erfahrung zu lernen, ***zu Denken und Probleme zu lösen*** *[…]“*

APA Task force / Neisser et al. Am Psy 1996, p77

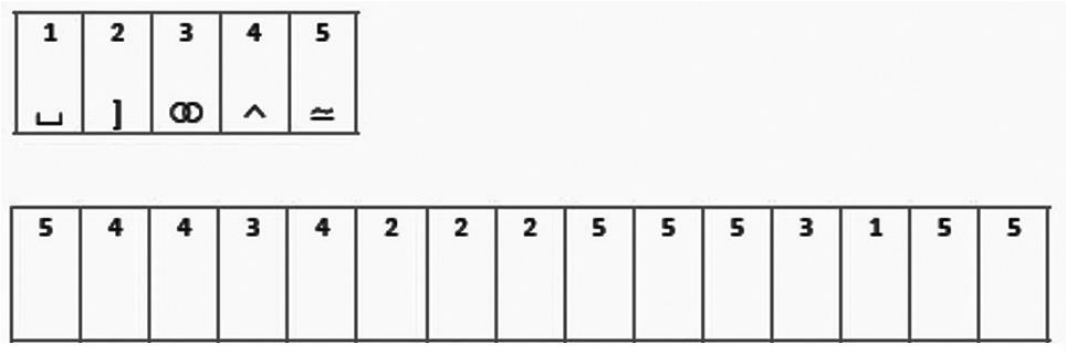
Intelligenztests (WAIS IV)



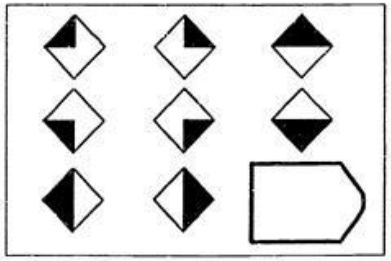
Sprache	**Gedächtnis**	**Problemlösen**	**Geschwindigkeit**
?	+++	?	+++

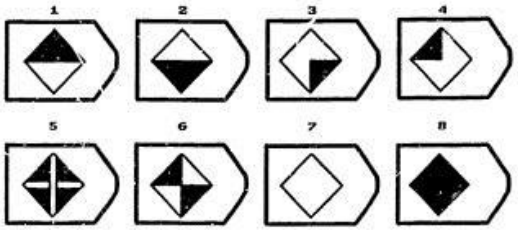
Intelligenztests

Zahlen-Symbol-Test

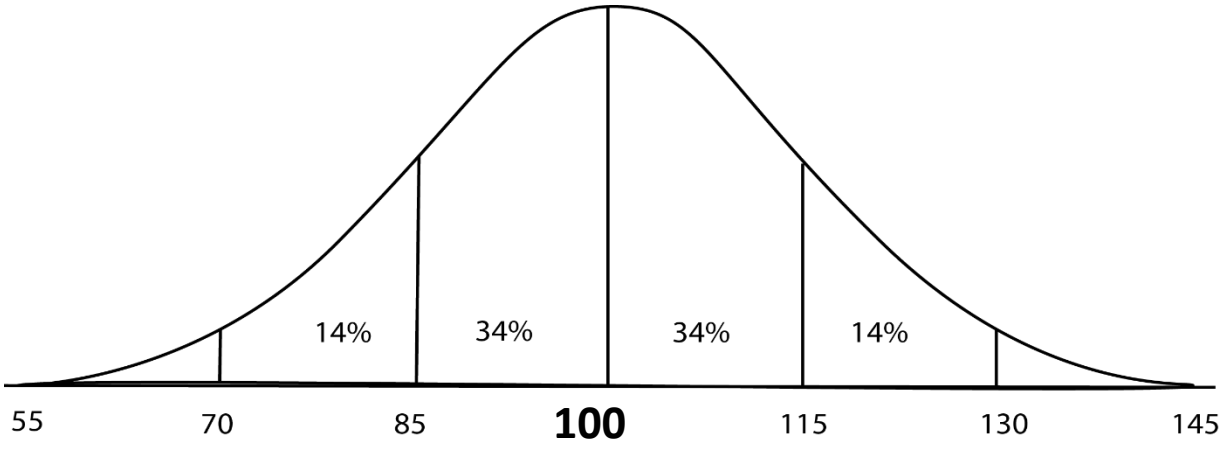


Matrizentest





Normskala



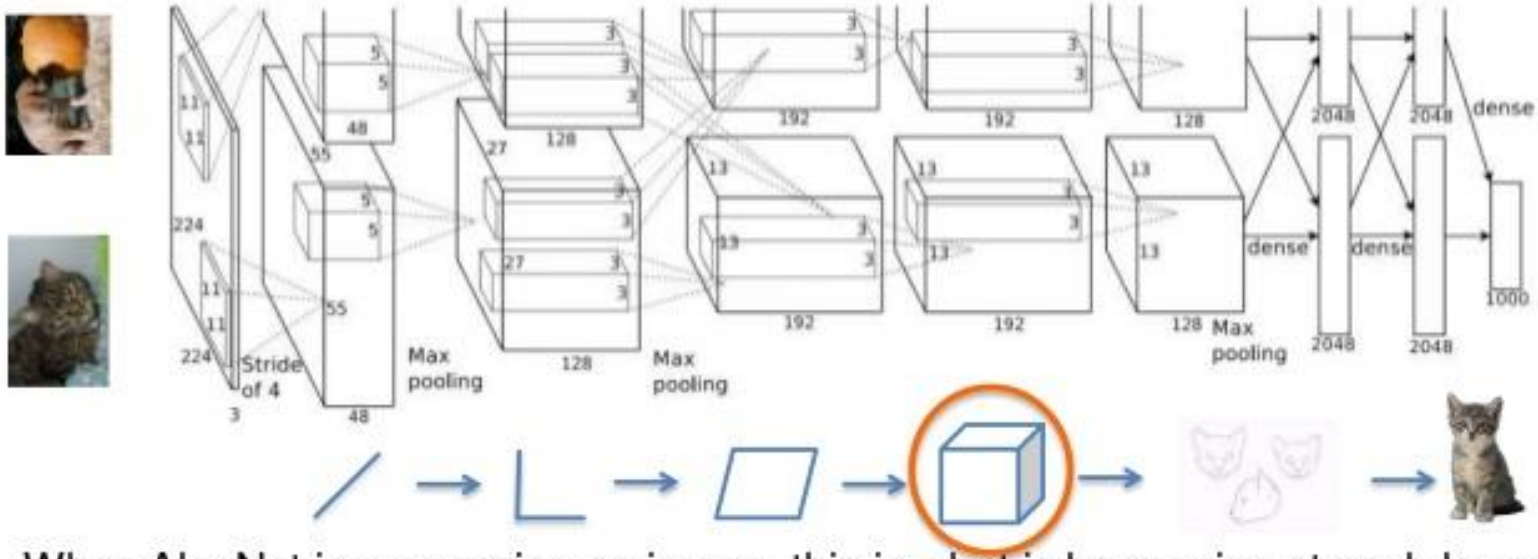
Kerninnovationen im Bereich KI: Deep neural networks

Mehrlagige neuronale Netze sind heute besser *mathematisch* und *technisch* besser handhabbar als früher. Sie sind durch Hirnprozesse inspiriert, unterscheiden sich aber in wesentlichen Aspekten.

AlexNet (Krizhevsky et al. 2012)

(ImageNet large-scale visual recognition challenge)

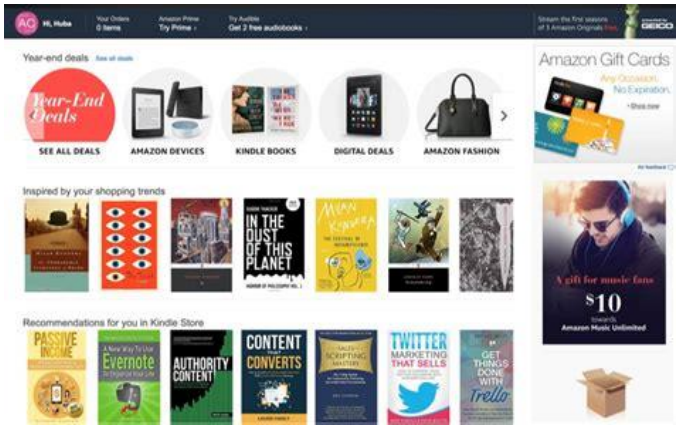
The class with the highest likelihood is the one the DNN selects



When AlexNet is processing an image, this is what is happening at each layer.

Kerninnovationen im Bereich KI: Big data and data analytics

Zahlreiche große Datensätze sind heute verfügbar und es existieren automatisierte statistische Verfahren um aus ihnen prädiktive Informationen zu gewinnen. Für Menschen wäre es schwer, in „Kopfarbeit“ dieselben Informationen zu extrahieren.



Kerninnovationen im Bereich KI: Reinforcement learning

Eine Kerninnovation in der modernen KI ist die
Verwendung von Algorithmen des Belohnungslernens,
die in der Psychologie und Neurowissenschaft entwickelt wurden.

$$R_t = r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + \gamma^3 r_{t+4} + ... = \sum_{k=0}^{\infty} \gamma^k r_{t+k+1}$$

Erfolgsbeispiel

Lee Sedol vs. AlphaGo (2016)

AlphaGoZero (2017):
Starke Verbesserung durch Verzicht auf menschliche Trainingsdaten → Self-play RL

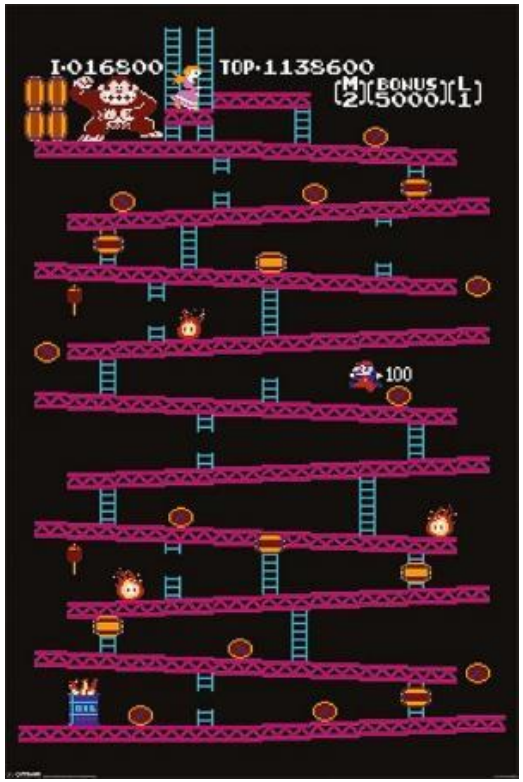
Silver et al. Nature 2017

Formalisierbarkeit

KI ist besonders erfolgreich in Situationen, die:

- Formal sehr gut beschrieben werden können
- Durchsucht, durchgespielt und simuliert werden können
- Große Datenmengen ergeben

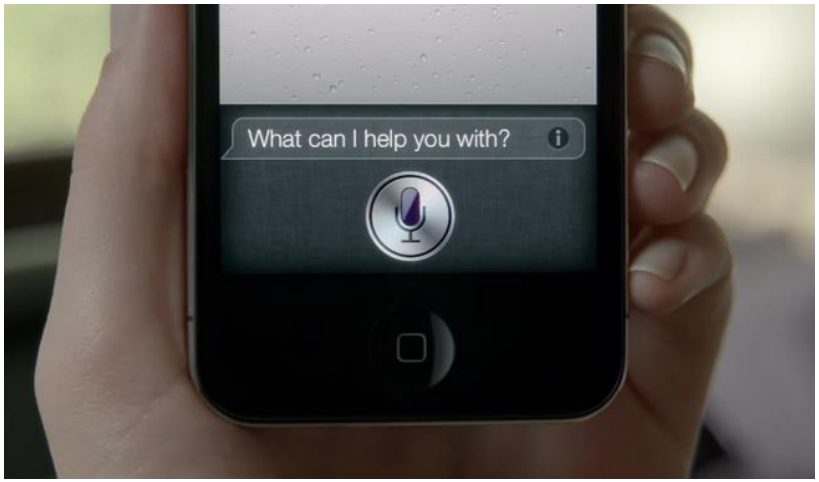
Donkey Kong

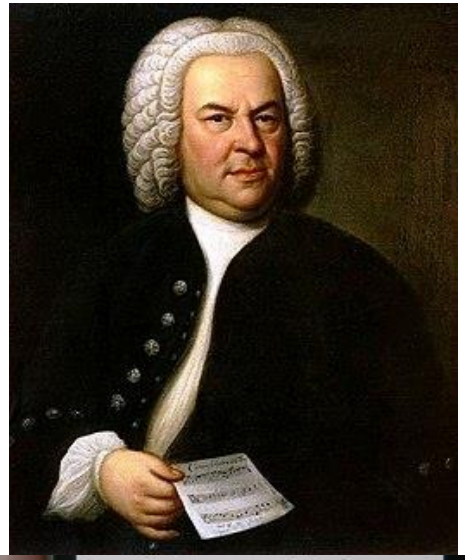


Containerterminal Altenwerder (CTW)

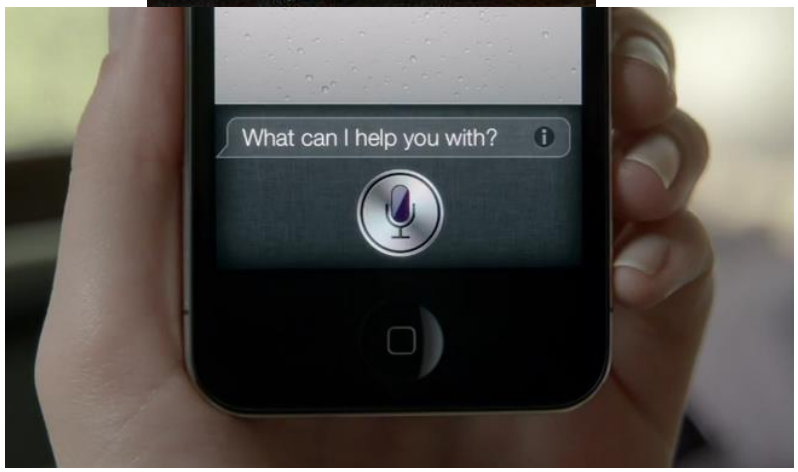
Menschliche Anwesenheit als Störfaktor.

Siri, wann wurde Johann Sebastian Bach geboren?

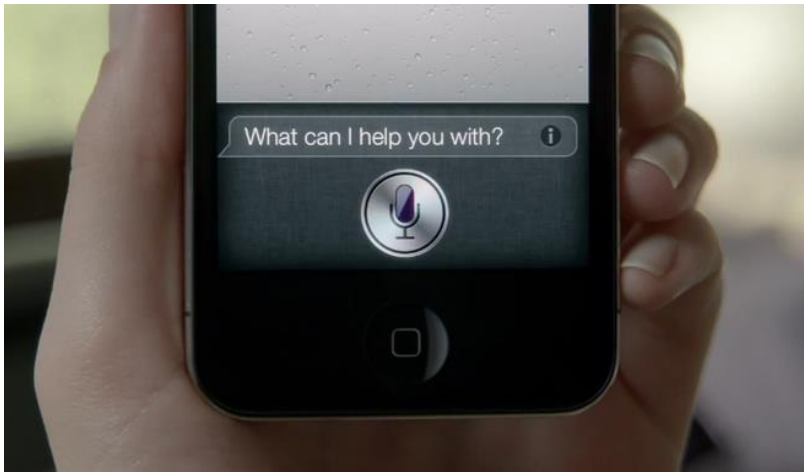




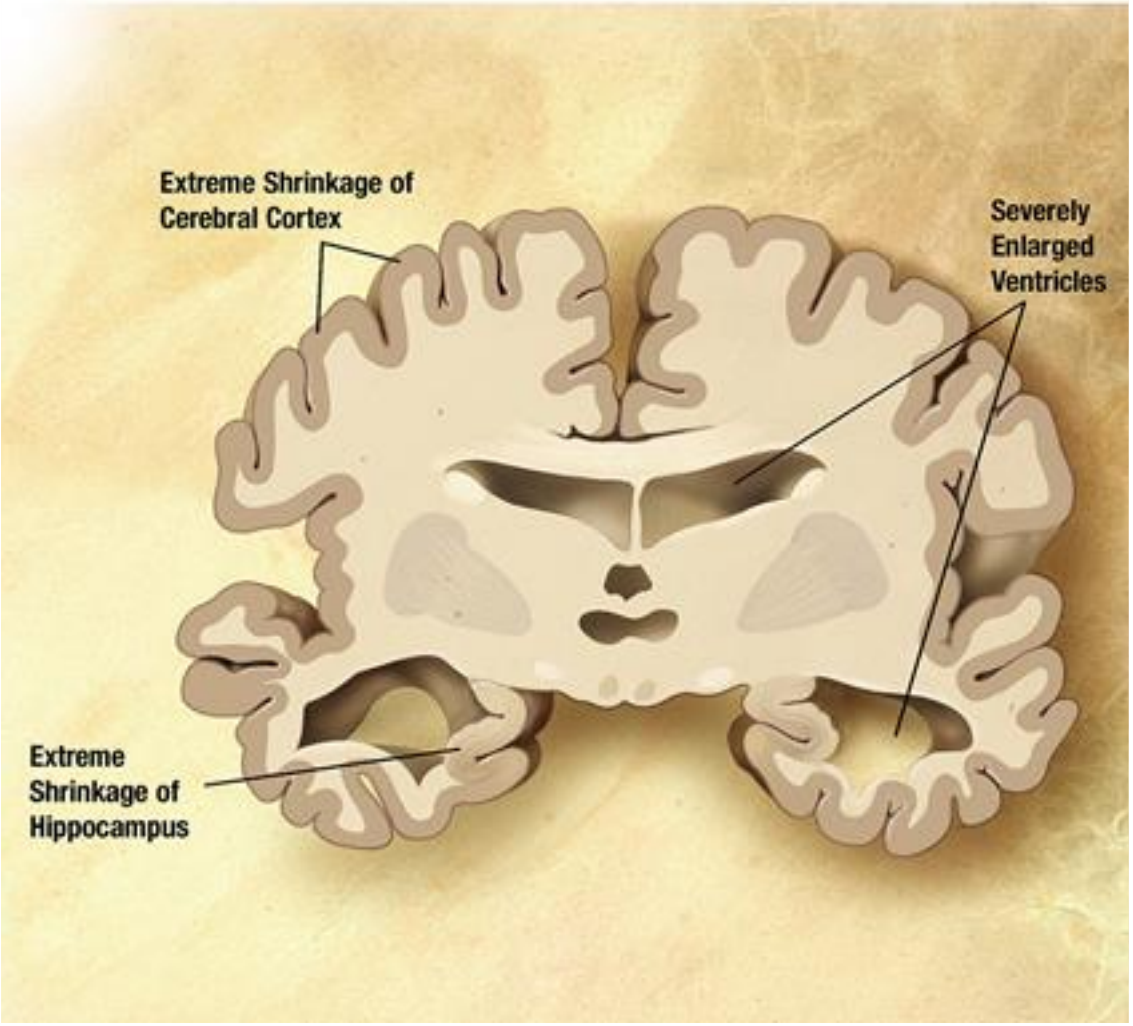
Johann Sebastian Bach
(1685-1750)



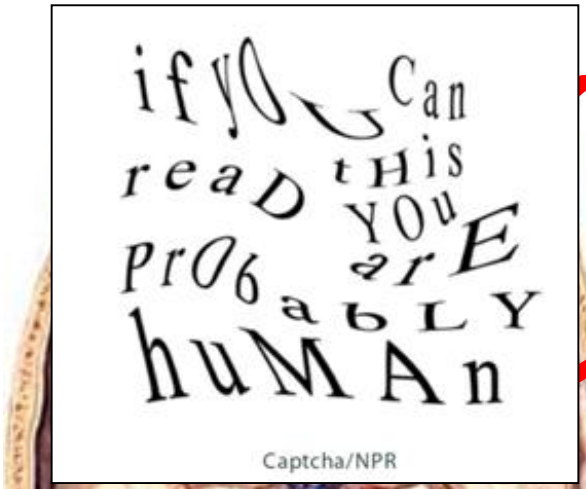
Siri, was hängt an der Wand, macht Tick-Tack und wenn‘s runter fällt ist die Uhr kaputt?



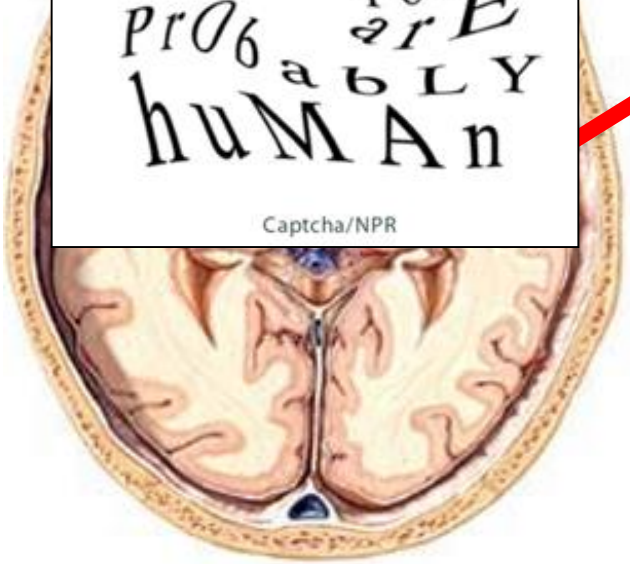
Unsere „Hardware“



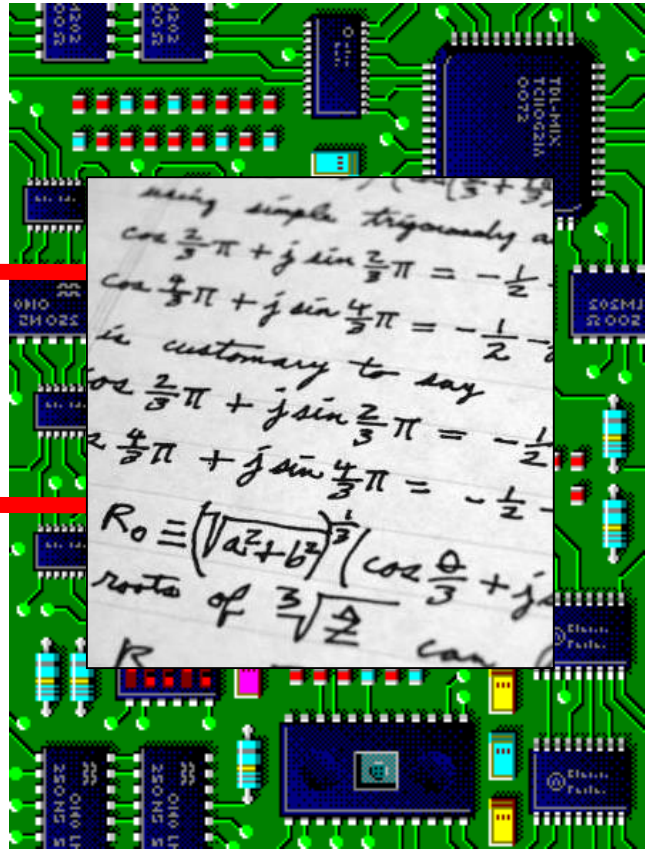
Gehirn versus Computer



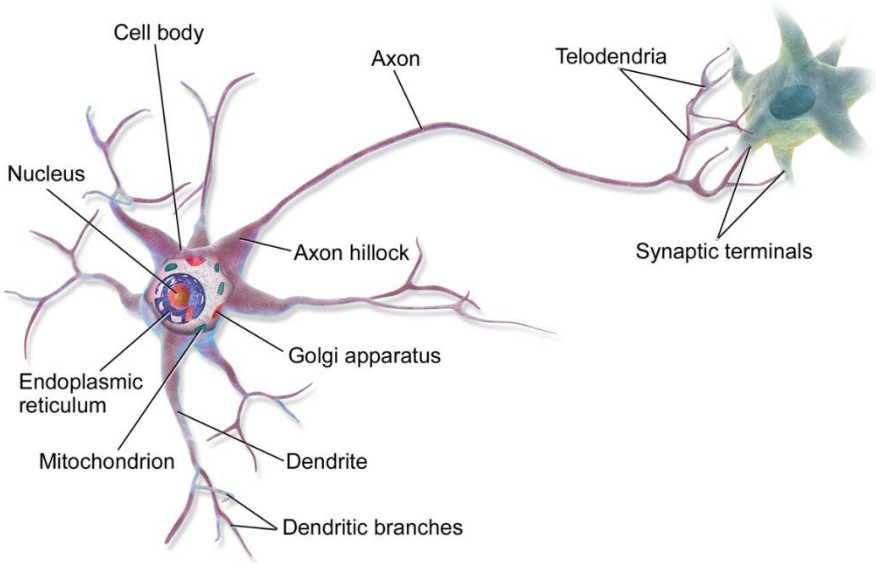
Captcha/NPR

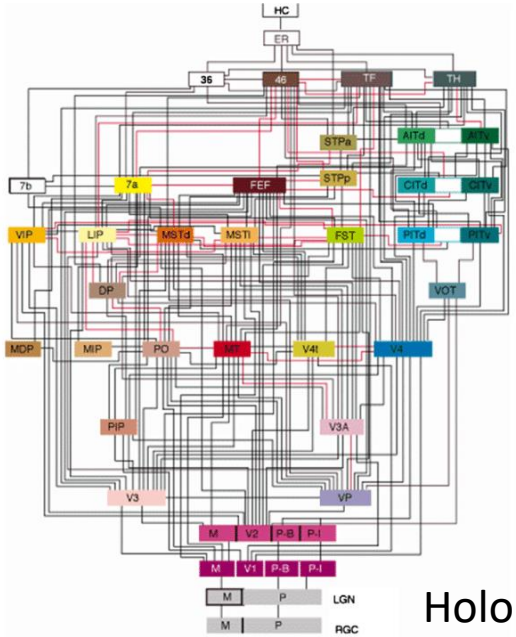


© ADAM, Inc.

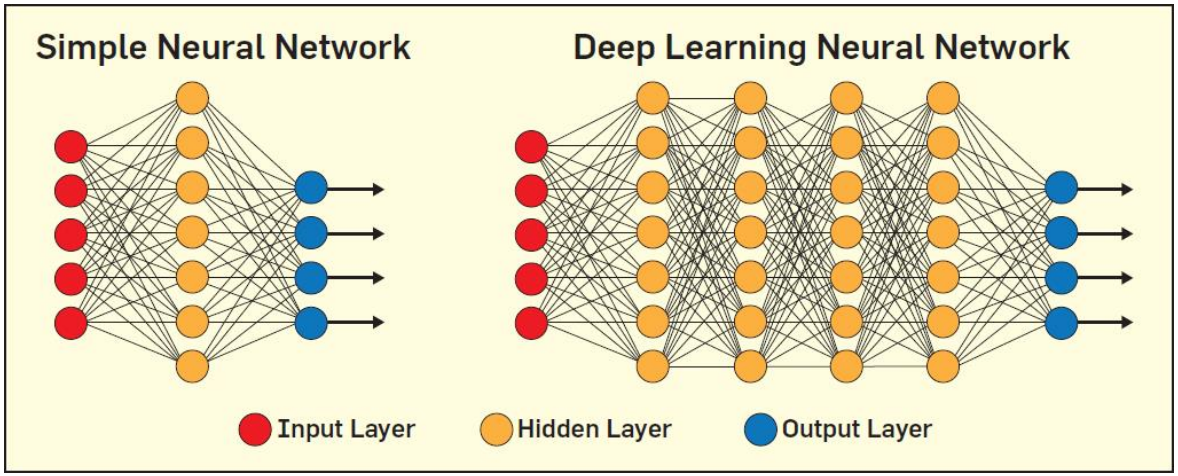


Unterschiede zwischen Gehirnen und neuronalen Netzen





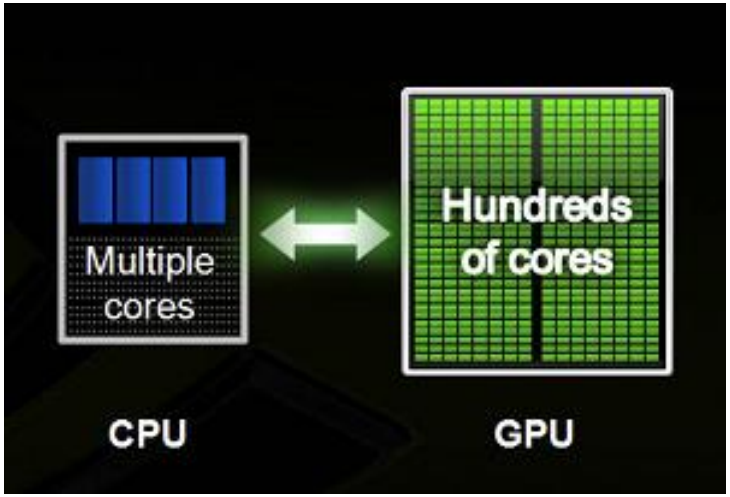
Holokratie



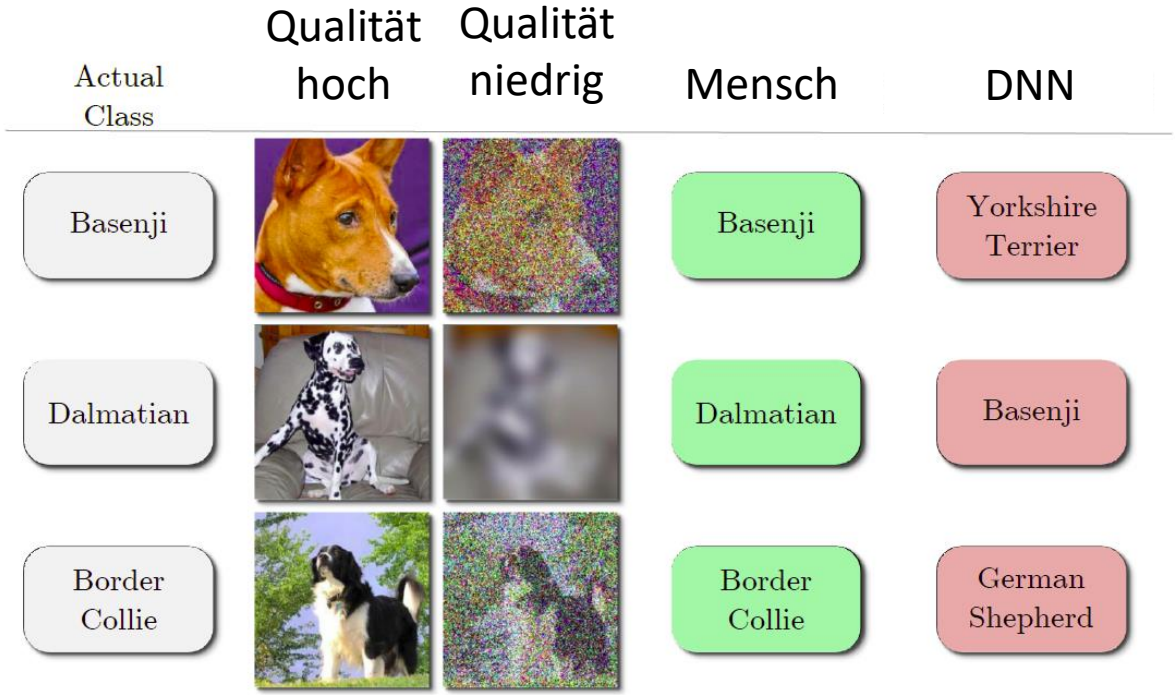
Hierarchie

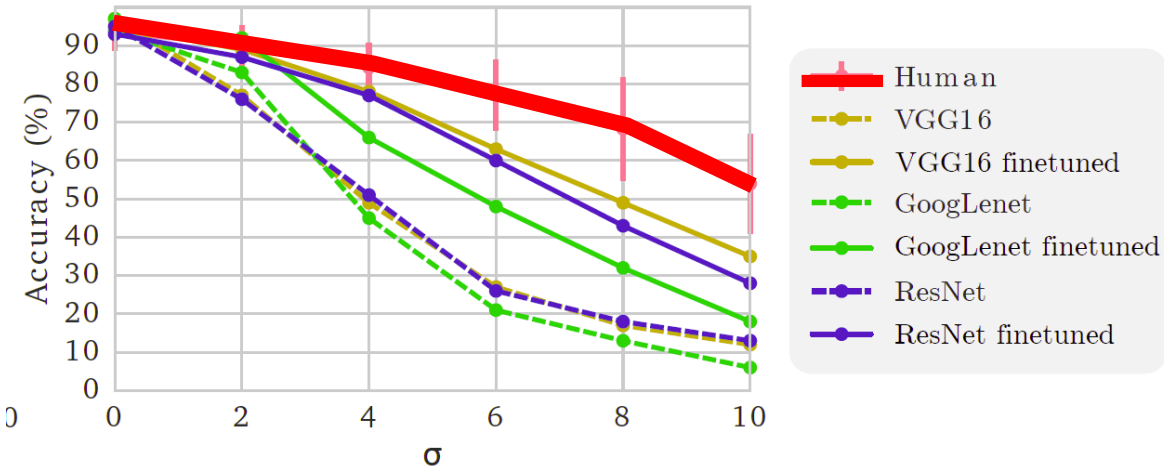
Gehirn versus Computer

GEHIRN	**COMPUTER (classical)**
Verteilte Verarbeitung	Lokalisierte Verarbeitung
Parallel	Seriell
Speicher verteilt	Speicher lokalisiert
Graceful degradation	Nonlinear degradation
Schwer zu programmieren	Leicht zu programmieren
Analog?	Digital
Holokratie	Hierarchie
Hohe strukturelle Plastizität	Niedrige strukturelle Plastizität



Mensch versus Computer: Die Rolle der Datenqualität





Dodge & Karam ArXiv 2017

Adverseriale („feindliche“) Beispiele

Ursprüngliches Bild: Klassifikation als Panda

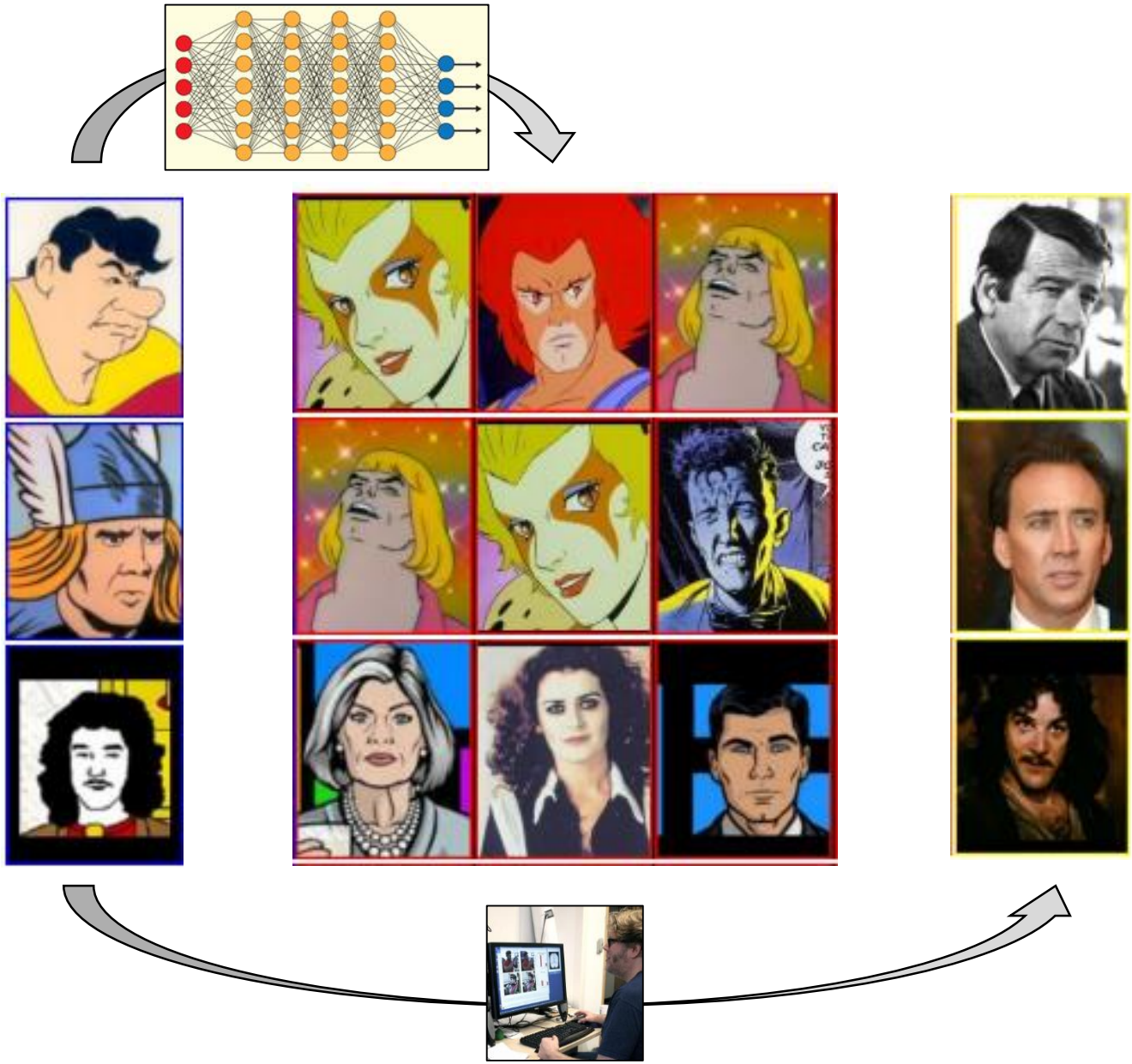
Addition „adversarialer“ Information

Bild als Gibbon klassifiziert

(Siehe auch: „Kluger Hans Effekt“)

https://www.kdnuggets.com/2015/07/deep-learning-adversarial-examples-misconceptions.html

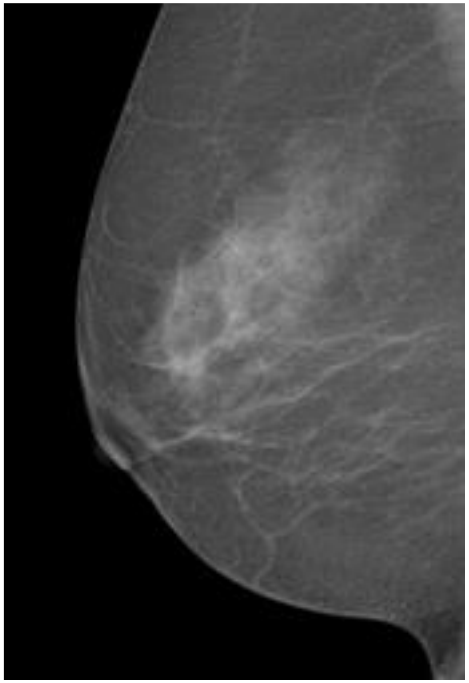
DNNs repräsentieren Informationen ganz anders als Menschen



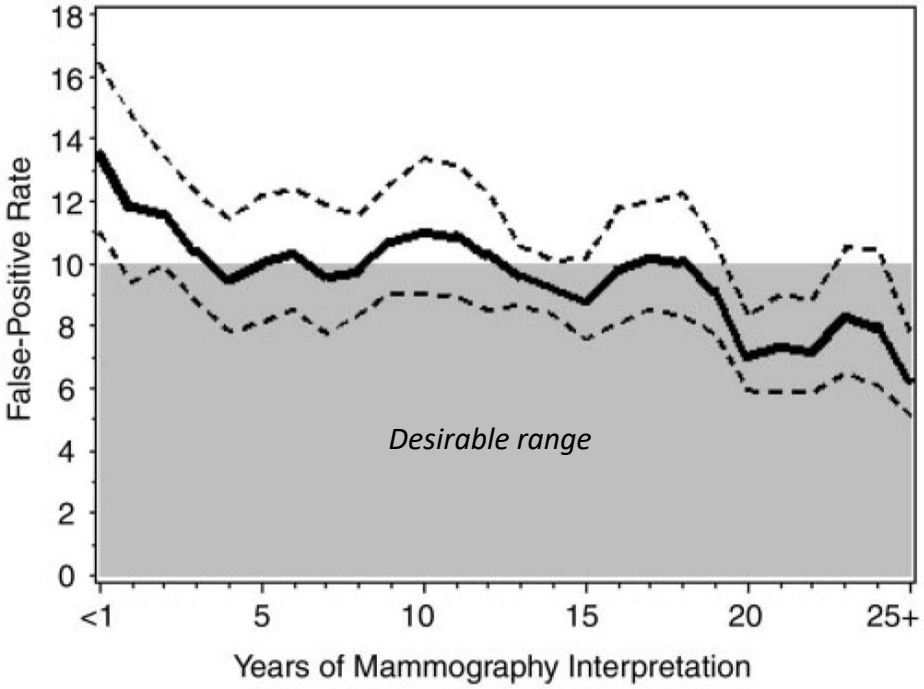
Amir Rosenfeld, Markus D. Solbach, John K. Tsotsos, ArXiv 2018

Das menschliche Gehirn auf „Big Data“

Mammographie (UK)
Minimale Erfahrung:
>5000 Fälle pro Jahr



Lernkurve über mehrere Jahre



National Health Service (NHS) Breast Screening Radiologists Quality Assurance Committee. Quality assurance guidelines for radiologists. (2005).
Miglioretti et al. When radiologists perform best: the learning curve in screening mammogram interpretation. Radiology. 2009 Dec;253(3):632-40.

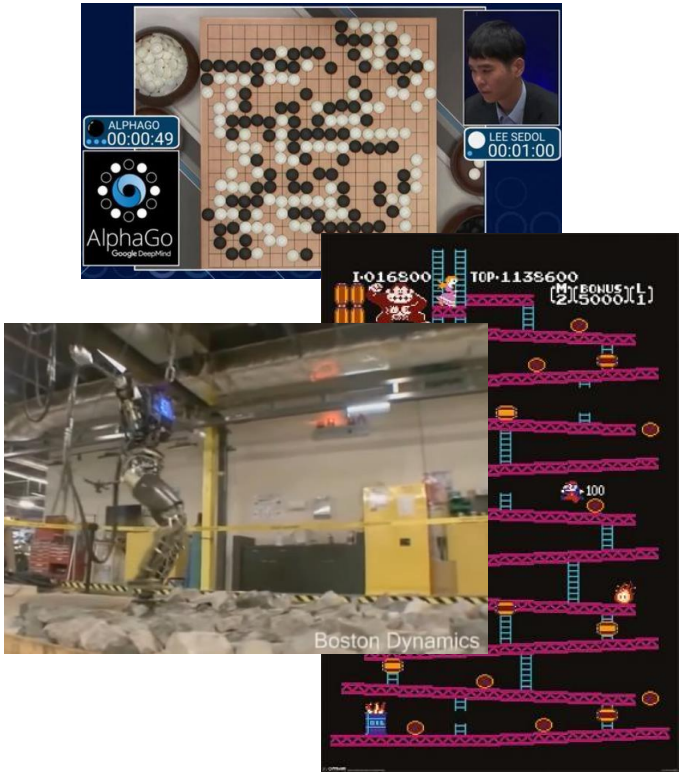
Robustheit versus Sprödigkeit („brittleness“)

Menschliche Stärken:
Manuelle Geschicklichkeit

Menschliche Stärken:
Soziale Signale & Empathie

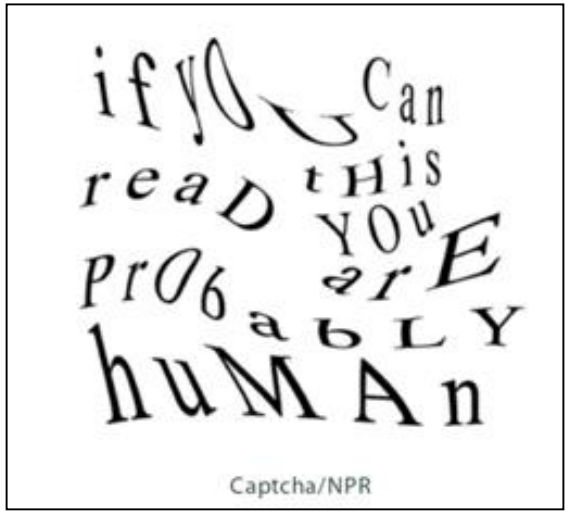
Menschliche Stärken: Der General Problem Solver

Sammlung von Idiot Savants

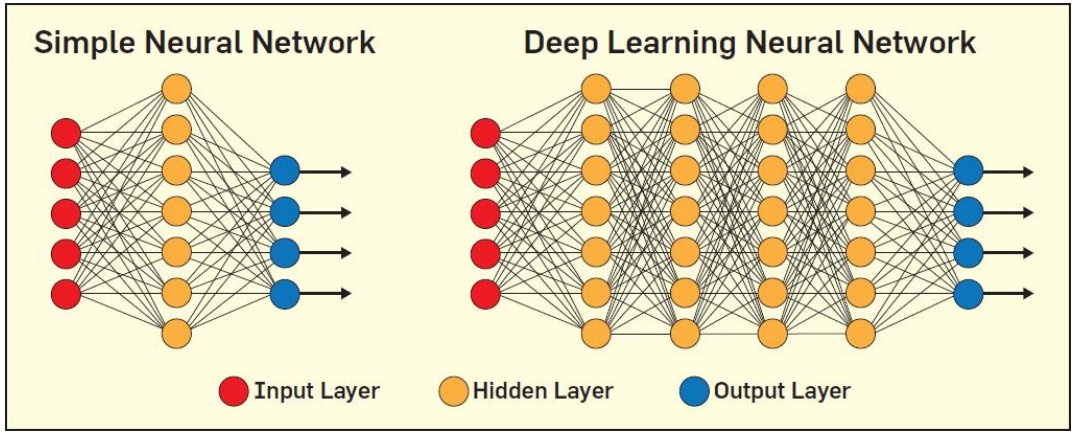


General Problem Solver

Algorithmen werden besser



CAPTCHA =
Completely **A**utomated **P**ublic **T**uring test to
tell **C**omputers and **H**umans **A**part



Teams aus Menschen und Algorithmen

Forsensic face recognition experts

Same or different people?

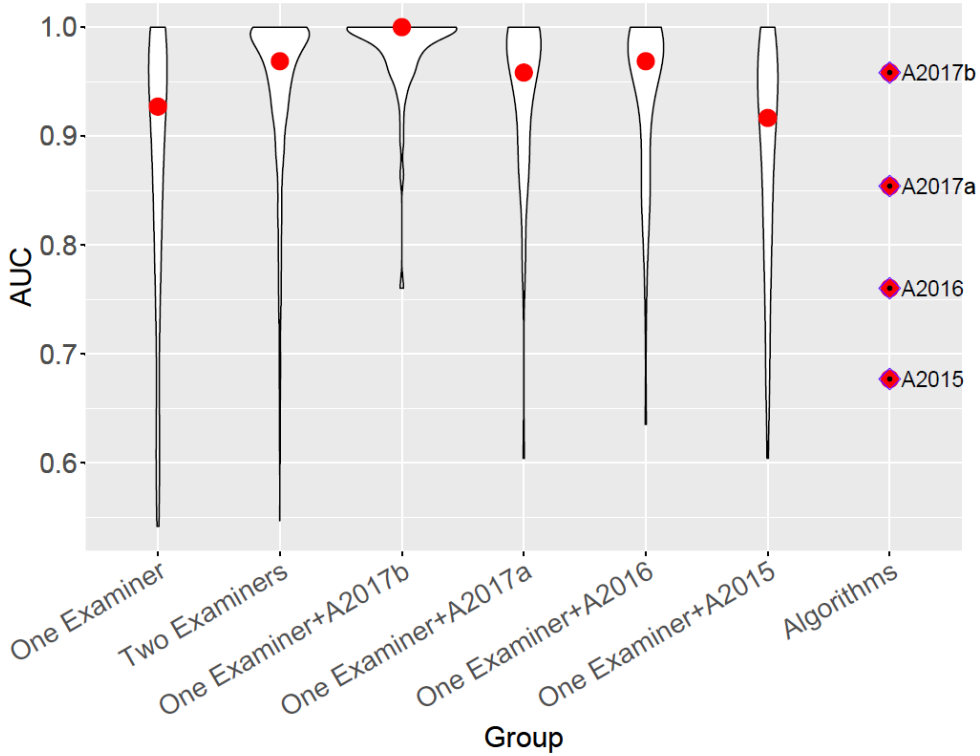
Same

Different

Philips et al. PNAS 2017 (www.pnas.org/cgi/doi/10.1073/pnas.1721355115)

Kooperation zwischen Menschen und Algorithmen

Beste Leistung: Kombination der Einschätzungen von Algorithmen und menschlichen Experten.

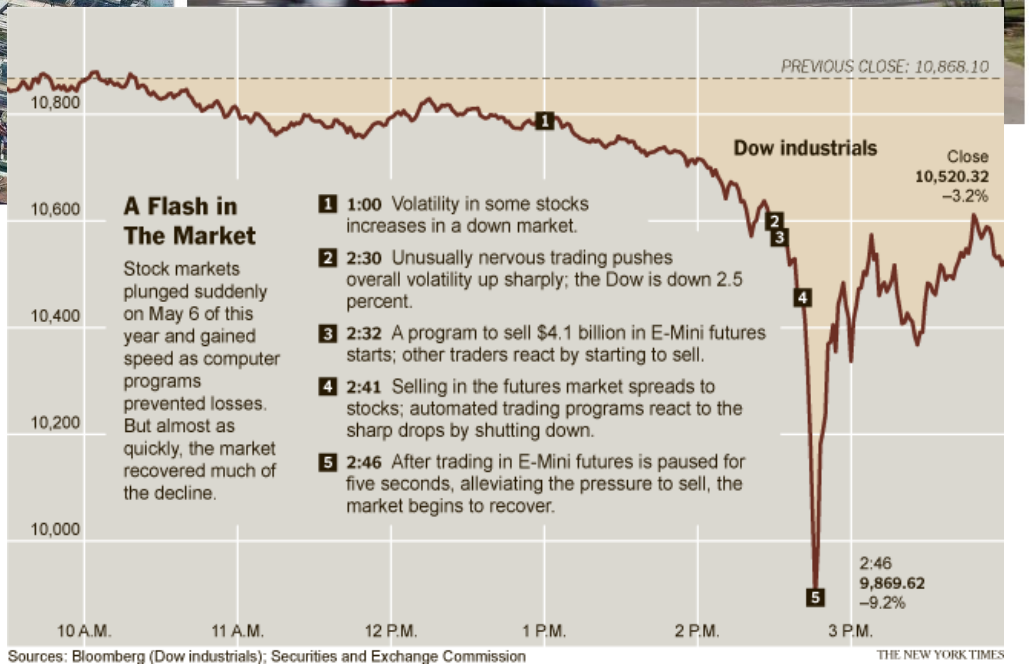


Philips et al. PNAS 2017 (www.pnas.org/cgi/doi/10.1073/pnas.1721355115)

Menschliche kognitive Grenzen in der Beherrschung von Technik

Self-driving Uber kills Arizona woman in first fatal crash involving pedestrian

Tempe police said car was in autonomous mode at the time of the crash and that the vehicle hit a woman who later died at a hospital



Zusammenfassung

Trotz einiger konzeptioneller Überlappung unterscheiden sich KI-Algorithmen und Hirnprozesse in fundamentalen Aspekten ihrer Informationsverarbeitung.

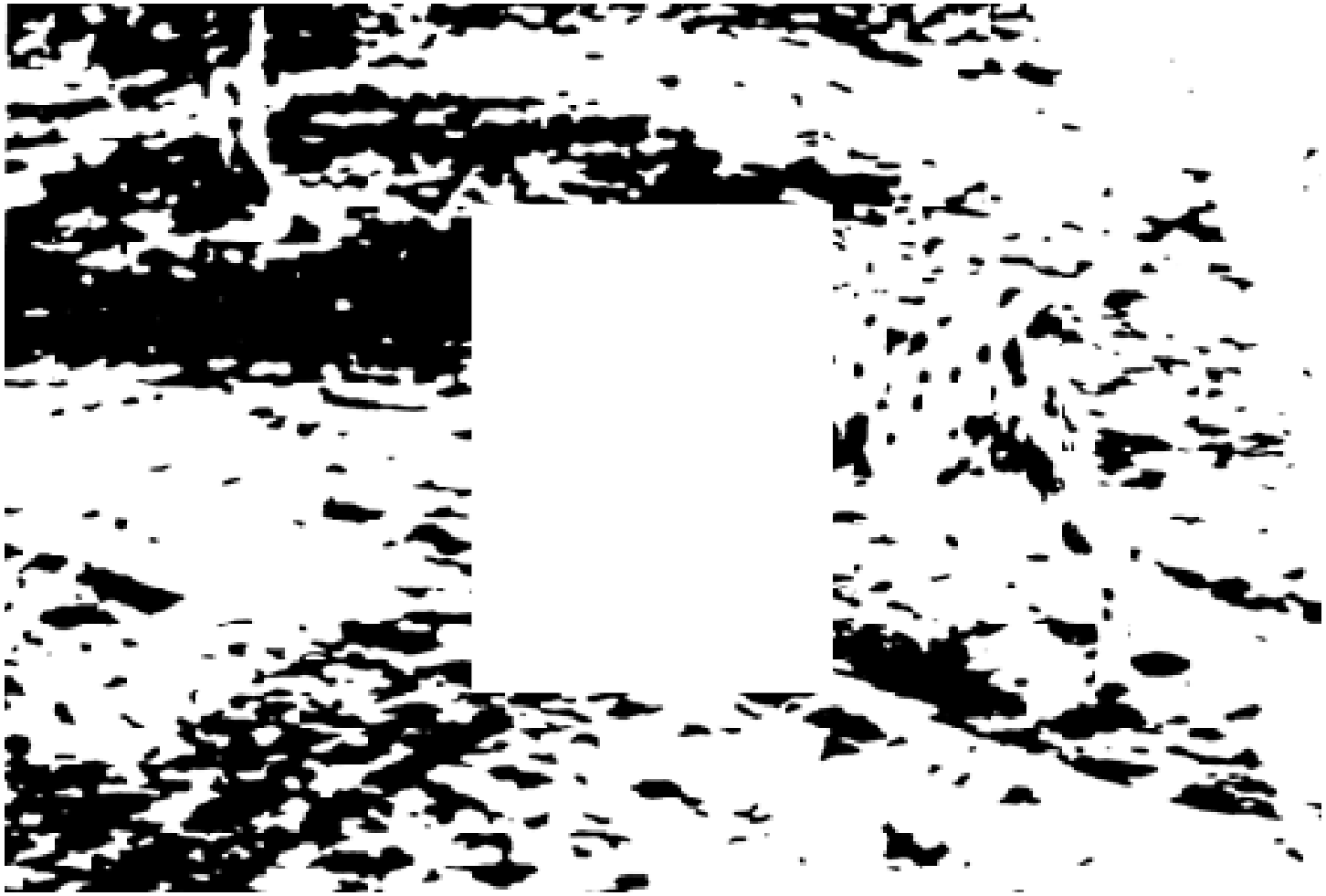
Die Definition von „Intelligenz“ ist sehr umstritten. Deshalb ist auch der Begriff „Künstliche Intelligenz“ irreführend.

Es gibt Aspekte von klassischen IQ-Test-Merkmalen, die sich gut und solche, die sich weniger gut algorithmisieren lassen.

Viele menschliche Stärken lassen sich derzeit nur schlecht algorithmisieren. Die beste Strategie ist also, Teams aus Mensch und KI zu bilden.

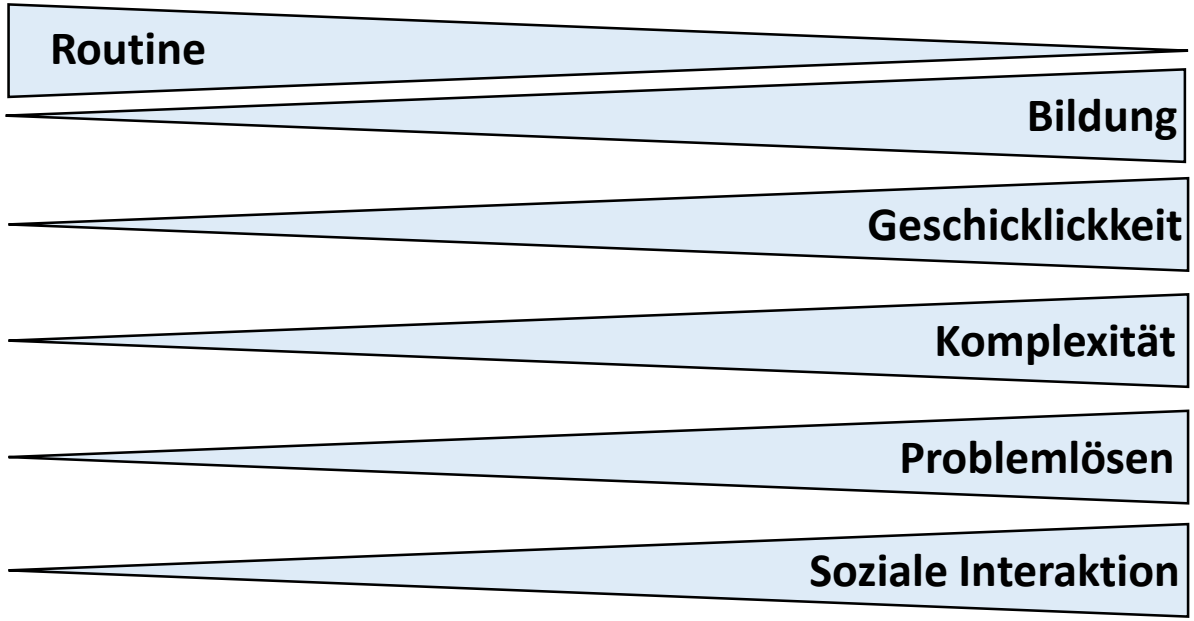
Der Mensch ist derzeit der einzige „General Problem Solver“.

Menschliche Stärken:
Mustererkennung aus kleinen Datenmengen



Menschliche Stärken:
Mustererkennung schießt über das Ziel hinaus

Gefährdung und Schutzfaktoren von Arbeitsplätzen: Sehr variable Schätzungen



Hohes Risiko ⟷ Niedriges Risiko

Frey & Osborne 2013

Gefährdung von Arbeitsplätzen/Berufsfeldern?

THE FUTURE OF EMPLOYMENT: HOW SUSCEPTIBLE ARE JOBS TO COMPUTERISATION?*

Carl Benedikt Frey† and Michael A. Osborne‡

September 17, 2013

“Nach unseren Schätzungen sind ca. 47 % der Arbeitsplätze in den USA bedroht.”

Gefährdung von Arbeitsplätzen/Berufsfeldern?

OECD Social, Employment and Migration Working Papers No. 202

Automation, skills use and training

Ljubica Nedelkoska, Glenda Quintini

BBC | Sign in | News | Sport | Weather | Shop | Earth | Travel

NEWS

Home | Video | World | UK | Business | Tech | Science | Stories | Entertainment & Arts

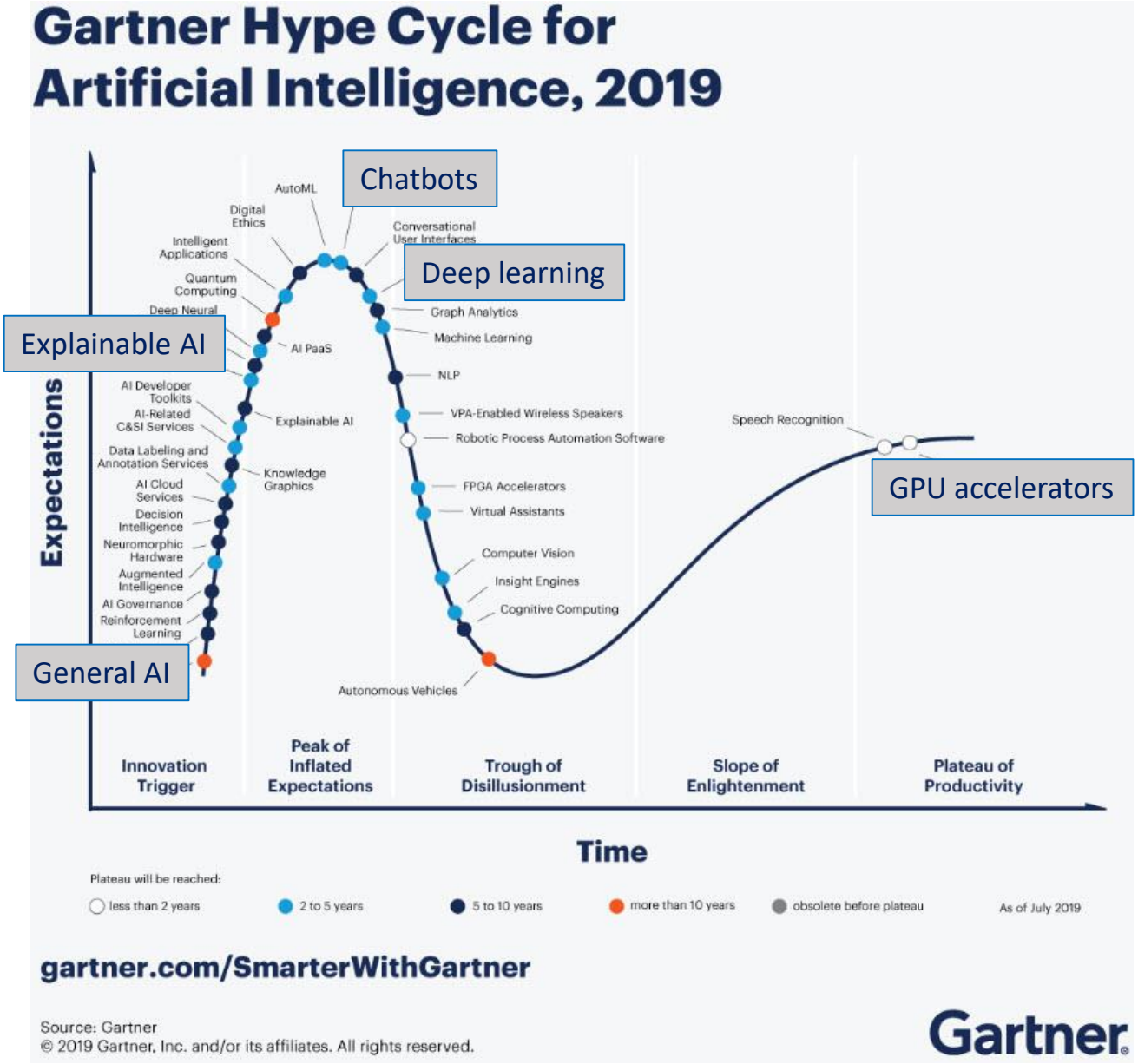
AI 'poses less risk to jobs than feared' says OECD

By Leo Kelion
Technology desk editor

2 April 2018 | Share

Eher bei 12%

Hype?


Gartner Hype Cycle for Artificial Intelligence, 2019
Chatbots
Deep learning
Explainable AI
GPU accelerators
General AI
AutoML
Digital Ethics
Intelligent Applications
Quantum Computing
Deep Neural
AI PaaS
AI Developer Toolkits
AI-Related C&SI Services
Explainable AI
Data Labeling and Annotation Services
Knowledge Graphics
AI Cloud Services
Decision Intelligence
Neuromorphic Hardware
Augmented Intelligence
AI Governance
Reinforcement Learning
Conversational User Interfaces
Graph Analytics
Machine Learning
NLP
VPA-Enabled Wireless Speakers
Robotic Process Automation Software
FPGA Accelerators
Virtual Assistants
Computer Vision
Insight Engines
Cognitive Computing
Autonomous Vehicles
Speech Recognition
Expectations
Innovation Trigger
Peak of Inflated Expectations
Trough of Disillusionment
Slope of Enlightenment
Plateau of Productivity
Time
Plateau will be reached:
less than 2 years
2 to 5 years
5 to 10 years
more than 10 years
obsolete before plateau
As of July 2019
gartner.com/SmarterWithGartner
Source: Gartner
© 2019 Gartner, Inc. and/or its affiliates. All rights reserved.
Gartner