Künstliche Intelligenz & Maschinelles Lernen

Thomas Wiegand

Fraunhofer
Heinrich Hertz Institute

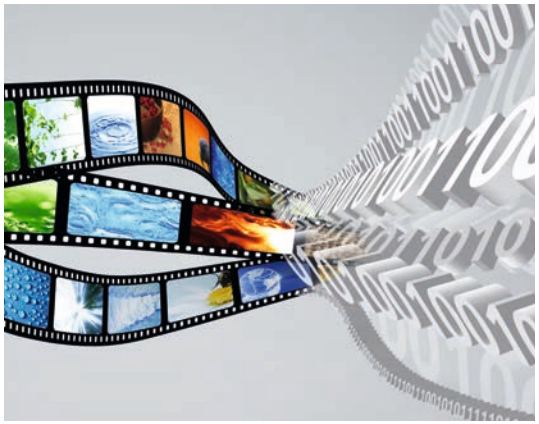
Fraunhofer Heinrich Hertz Institute

1928	Founding of the Heinrich Hertz Institute
1960s	Research on fiber optical transmission
1980s	Research on digital video coding
2000	Research on wireless networking
2003	Institute of Fraunhofer-Gesellschaft

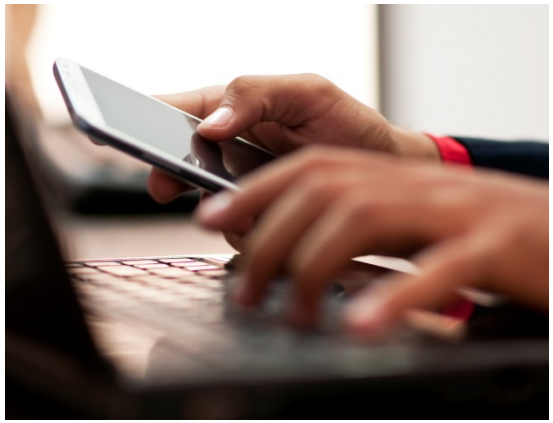
Fraunhofer Heinrich Hertz Institute
Research with impact

- Global player in research for digital transformation
- Budget (2016) 45,6 Mio €, 87% external funding of which 2/3 is from industry
- 420 Researchers
- Research & Development in Photonics, Video & Wireless Technologies
- Every second bit on the internet touches Video or Photonic technology invented or made by Fraunhofer HHI

10^0 – 10^2 – 10^4 Gbps

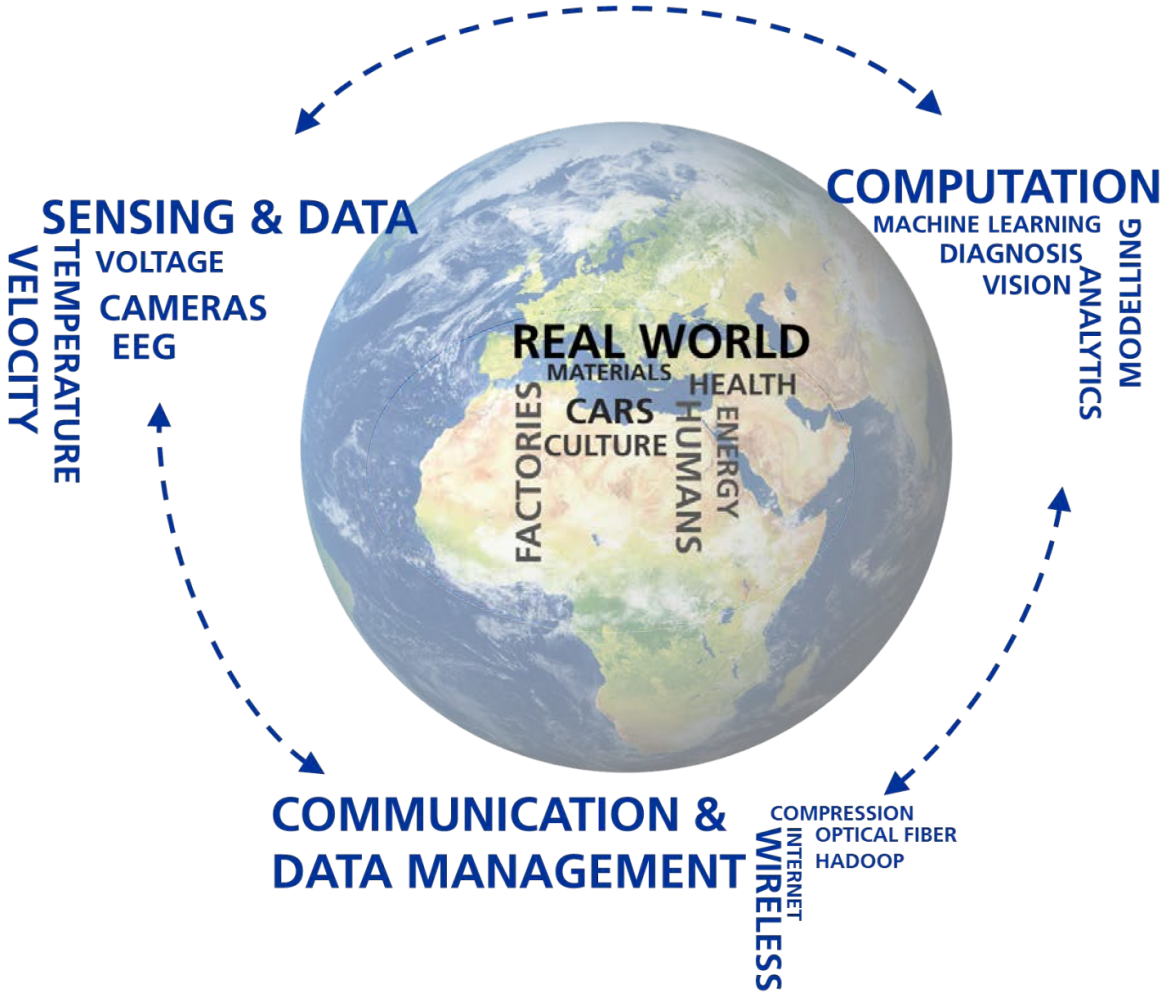


H.264 – H.265 – H.266

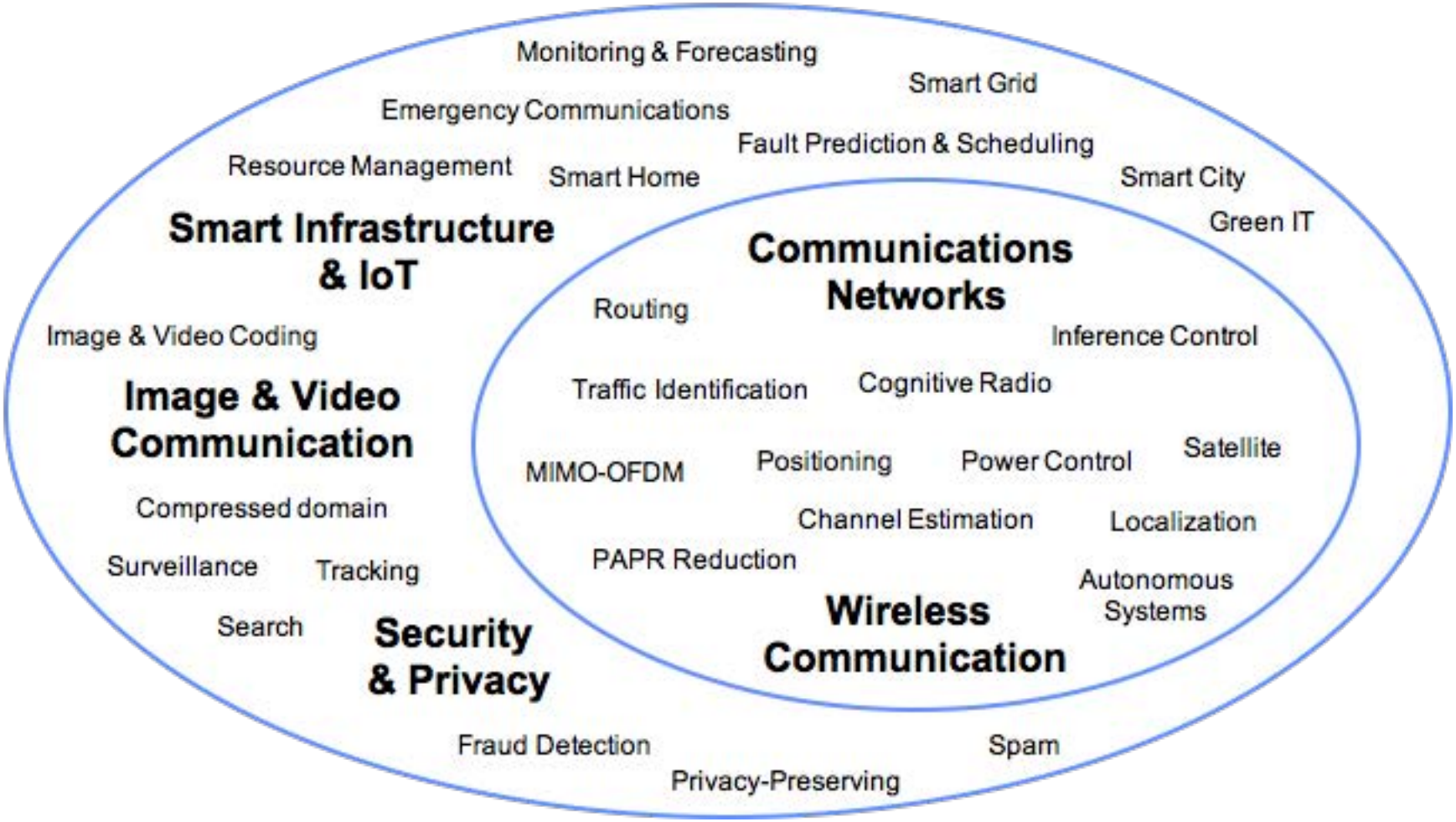


3G – 4G – 5G

Digital Transformation: Sensing, Communication, Computation



KI/ML & Kommunikation



W. Samek, S. Stanczak, and T. Wiegand: The Convergence of Machine Learning and Communication, Report for ITU-T

Video Coding Standards

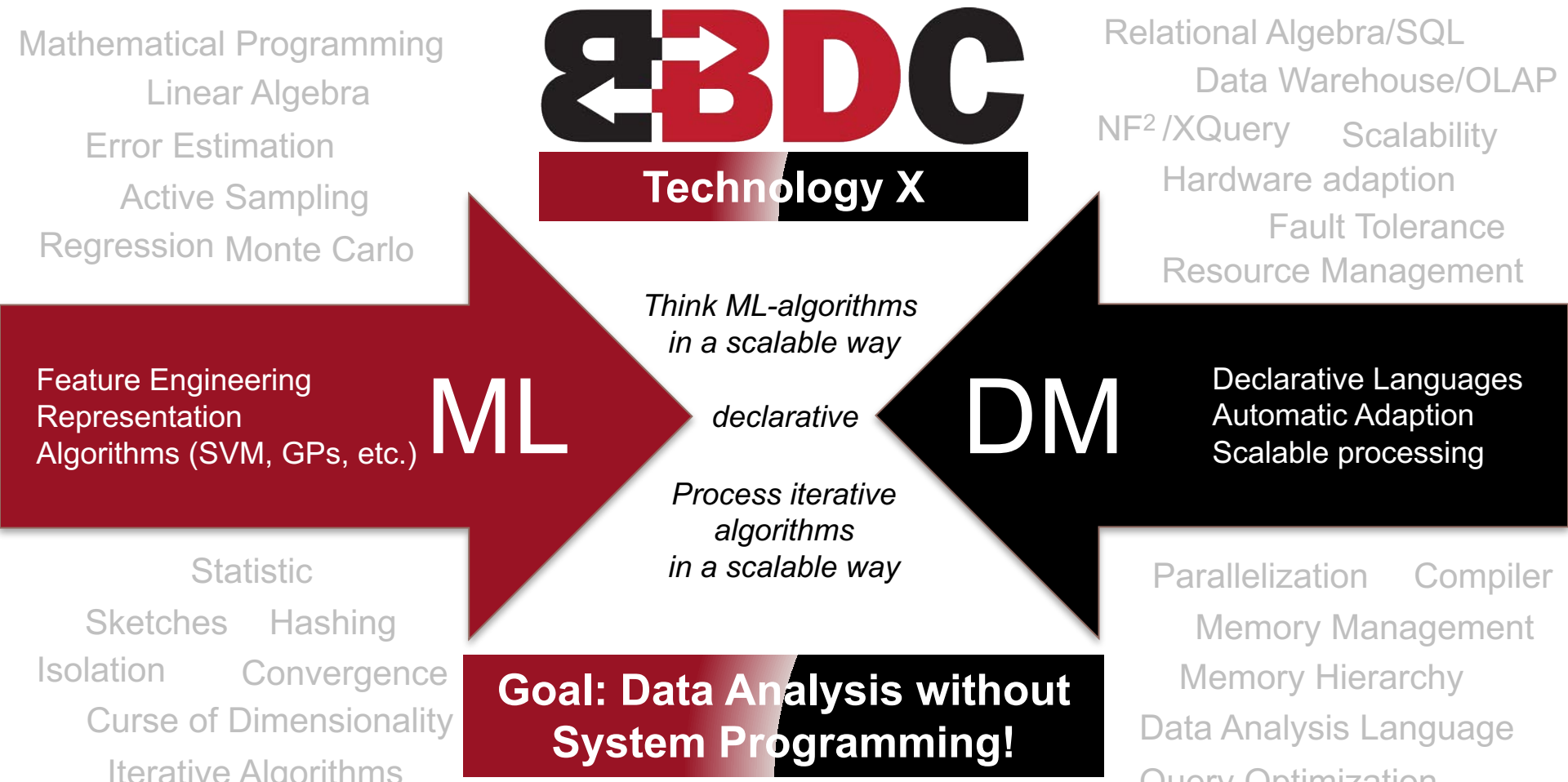
International standardization of video coding:

- Every 2nd bit on the Internet is H.264/MPEG-AVC
- H.265/MPEG-HEVC: about 1 Billion devices and mass deployment is picking up
- H.266/MPEG-X is in future planning stage

H.264/MPEG-AVC: More than 3 Billion devices and more than 50% of all bits on the Internet with essential components made in Berlin

H.265/MPEG-HEVC: is becoming the next world standard with essential components made in Berlin

Machine Learning + Data Management = X



Jüngste Erfolge von KI/ML-Systemen

AlphaGo beats Go human champ

Deep Net outperforms humans in image classification

IMAGENET

Autonomous search-and-rescue drones outperform humans

Schrödinger Equation

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\Delta\psi + V\psi.$$

Computer out-plays humans in "doom"

IBM's Watson destroys humans in jeopardy

Deep Net beats human at recognizing traffic signs

KI/ML & autonomes Fahren

(SegNet, University of Cambridge)

KI/ML wird in 10 Jahren …

1958, H. A. Simon and Allen Newell: "within ten years a digital computer will be the world's chess champion" and "within ten years a digital computer will discover and prove an important new mathematical theorem."

1965, H. A. Simon: "machines will be capable, within twenty years, of doing any work a man can do."

1970, Marvin Minsky: "In from three to eight years we will have a machine with the general intelligence of an average human being."

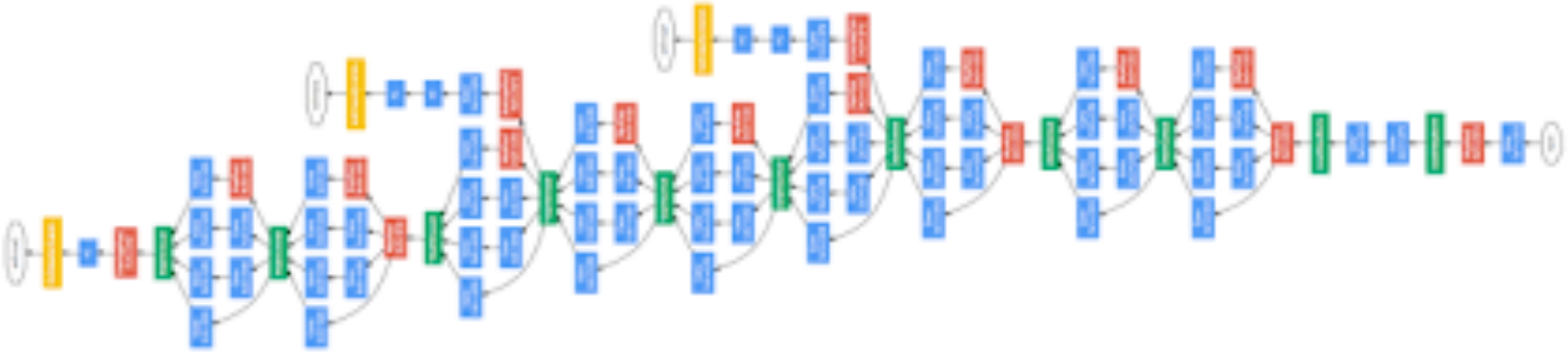
2017, Elon Musk: “AI will be able to beat humans at EVERYTHING by 2030.”

Wieso funktioniert es heute ?

Mehr Daten

Mehr Rechenkapazität

“Tiefe” neuronale Netze



Tiefe neuronale Netze

Deep: Hierarchy of representations with increasing level of abstraction

Image recognition

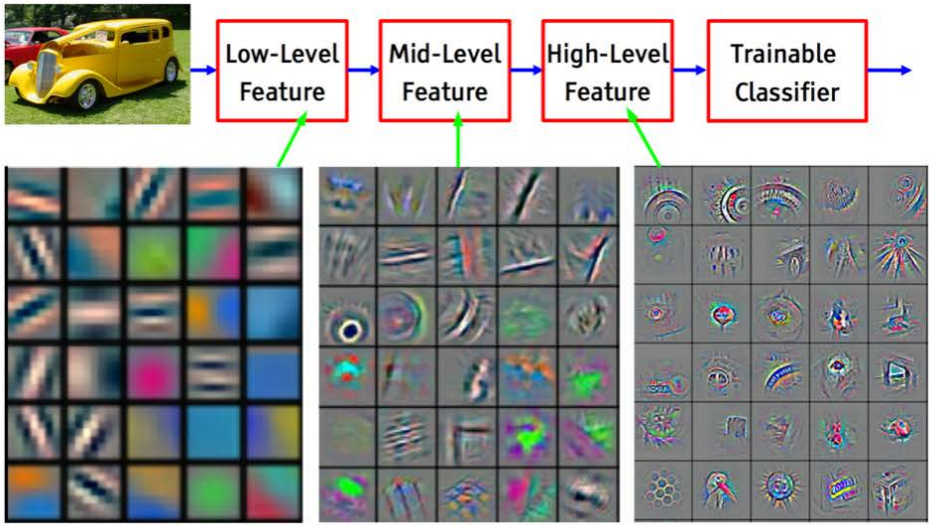
- Pixel → edge → texton → motif → part → object

Text

- Character → word → word group → clause → sentence → story

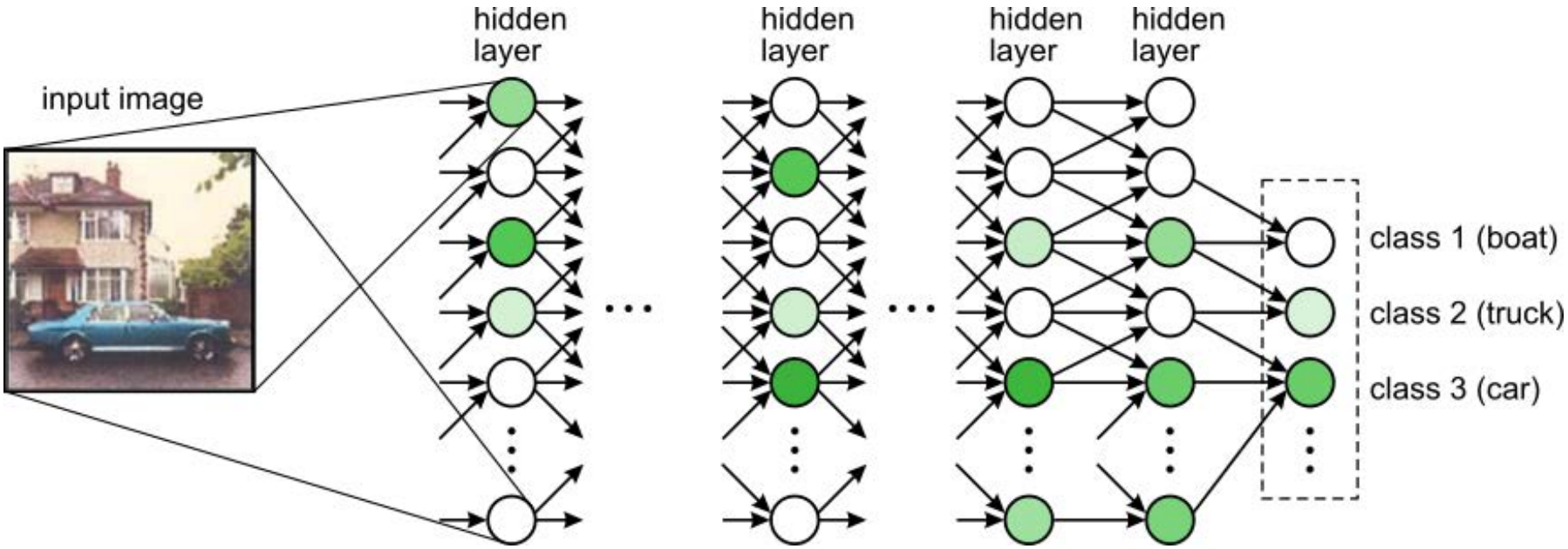
Speech

- Sample → spectral band → sound → ... → phone → phoneme → word



(Source: Yann LeCun

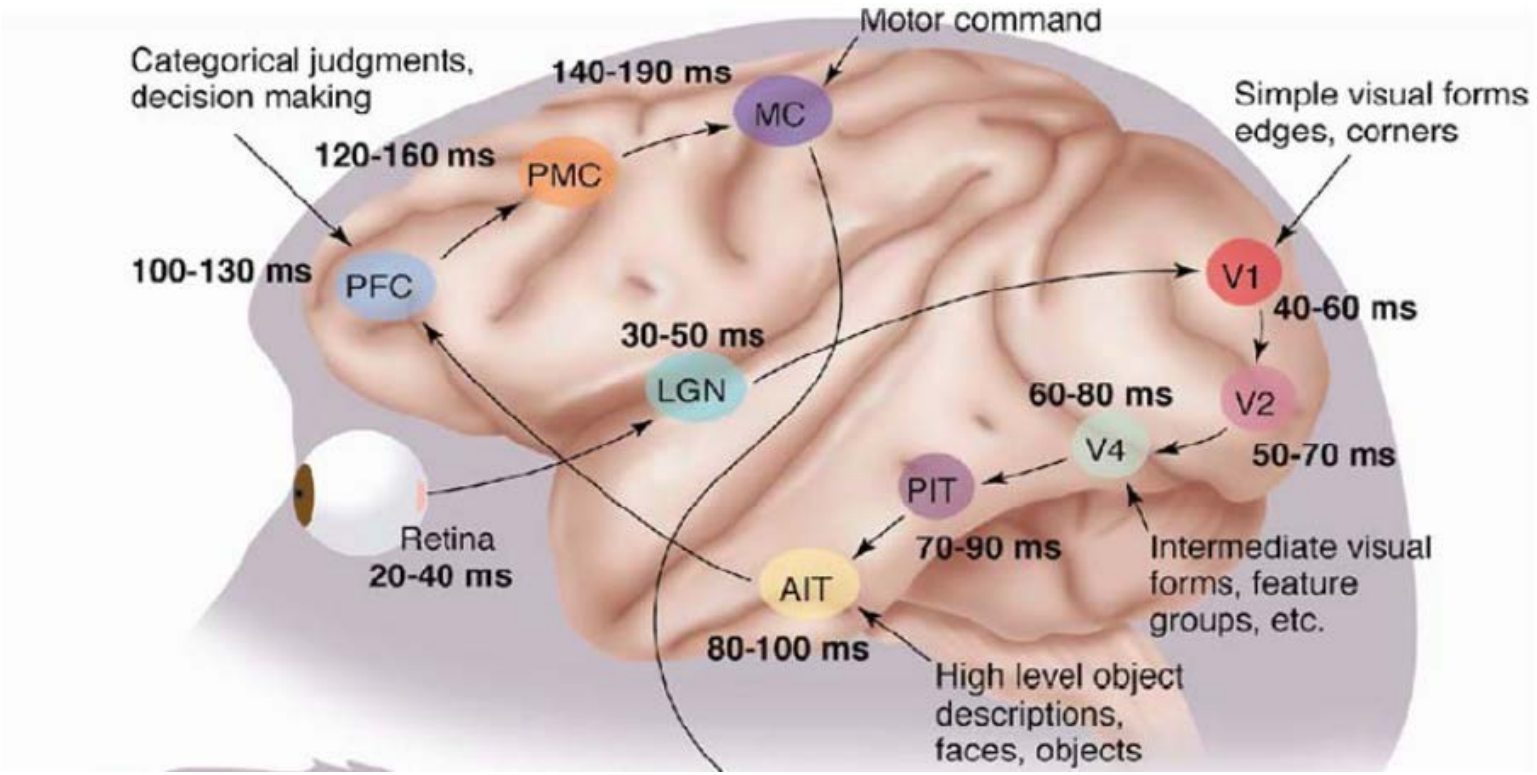
Tiefe neuronale Netze



60 million parameters
650,000 neurons

Tiefe neuronale Netze

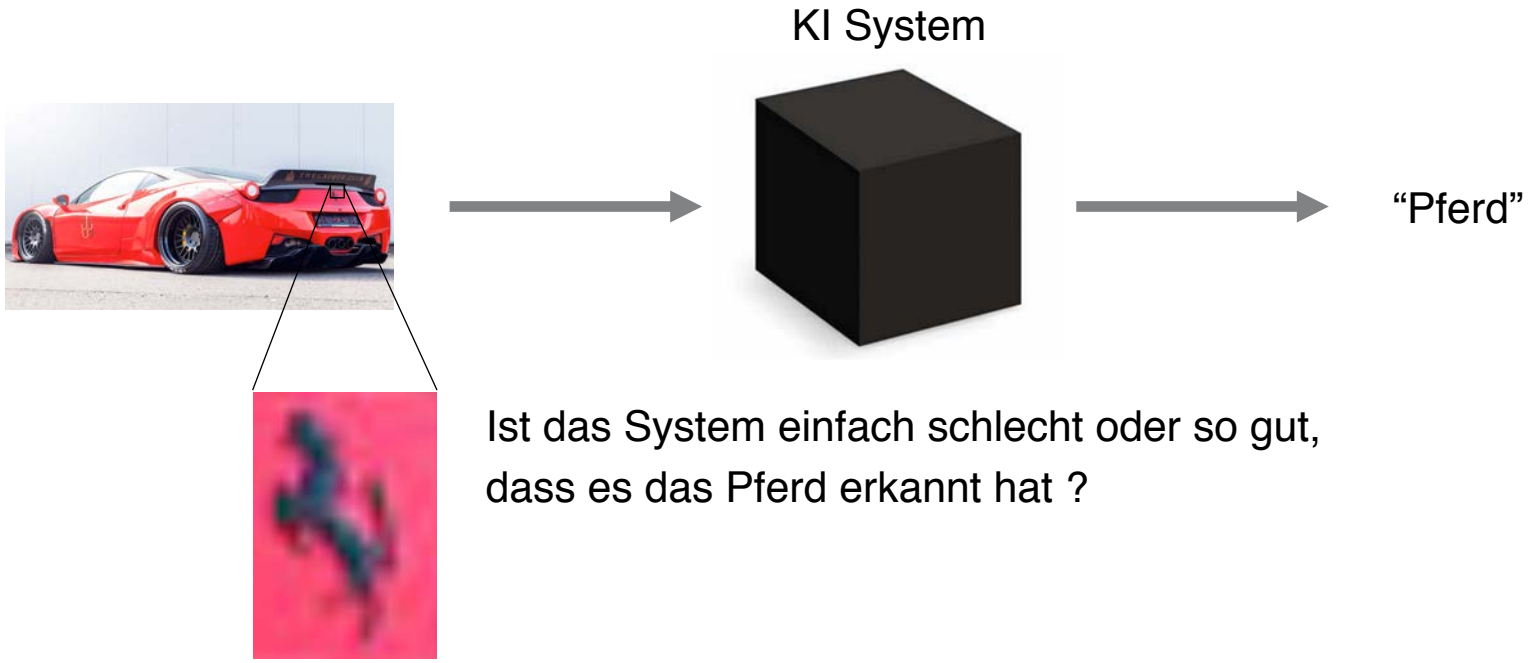
Unser Hirn arbeitet auch hierarchisch.



(Source: Simon Thorpe)

UNO/ITU-Konferenz “AI for Good”

KI/ML-Systeme sind “Black Boxen”



Ist das System einfach schlecht oder so gut, dass es das Pferd erkannt hat ?

Wenn wir nicht nachvollziehen können wieso die KI diese oder jene Entscheidungen trifft, dann sollten wir sie in bestimmten Bereichen (z.B. Medizin) nicht einsetzen.

KI/ML-Systeme sind “Black Boxen”

Lungenentzündungen, wenn falsch behandelt, können tödlich enden.
Risikopatienten müssen sofort behandelt werden.

Ziel: KI System soll vorhersagen, ob erkrankte Person ein Risikopatient ist und sofort behandelt werden muss oder nicht.

KI System: Personen mit Asthma, Brustschmerzen und Herzproblemen haben geringeres Risiko.

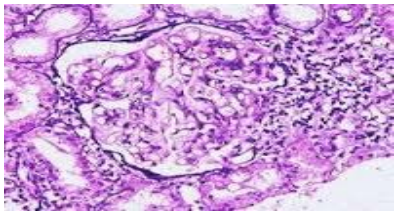
Es wäre lebensgefährlich sich auf solch ein System zu verlassen.

Doch wieso hat die KI hier versagt ?

KI/ML-Systeme sind “Black Boxen”

Verify that system works as expected

Wrong decisions can be costly and dangerous.



Understand “weaknesses” of classifier

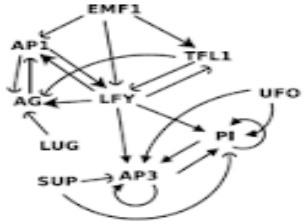
Detect biases / bring in human intuition.

Learn from the learning machine

“I've never seen a human play this move.” (Fan Hui)

Interpretability in the sciences

The “why” often more important than the prediction.



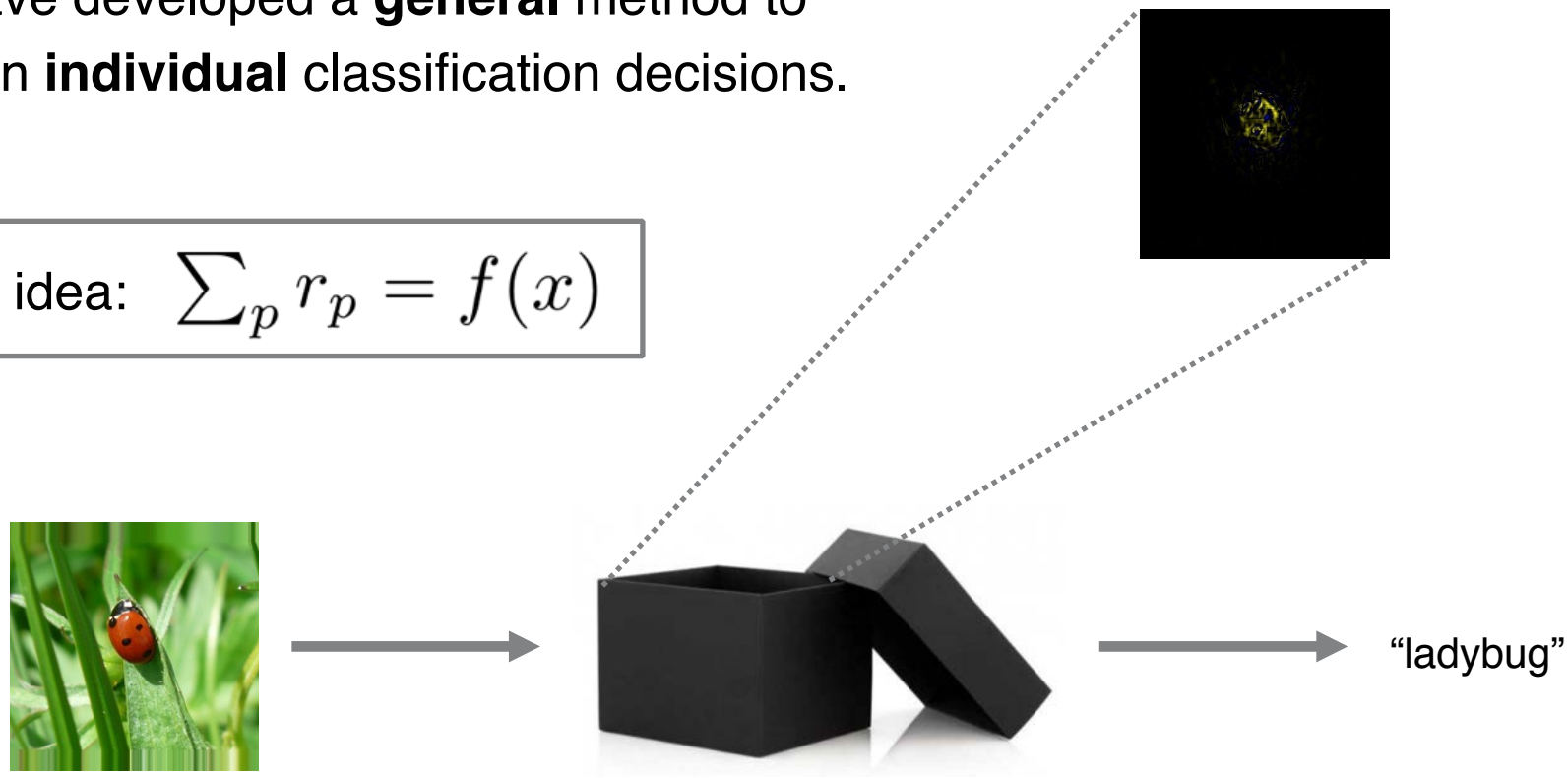
Compliance to legislation

“right to explanation”, retain human decision …

KI/ML nachvollziehbar machen

Together with colleagues from TU Berlin we have developed a **general** method to explain **individual** classification decisions.

Main idea: $\sum_p r_p = f(x)$



KI/ML nachvollziehbar machen

Classification

cat

ladybug

dog

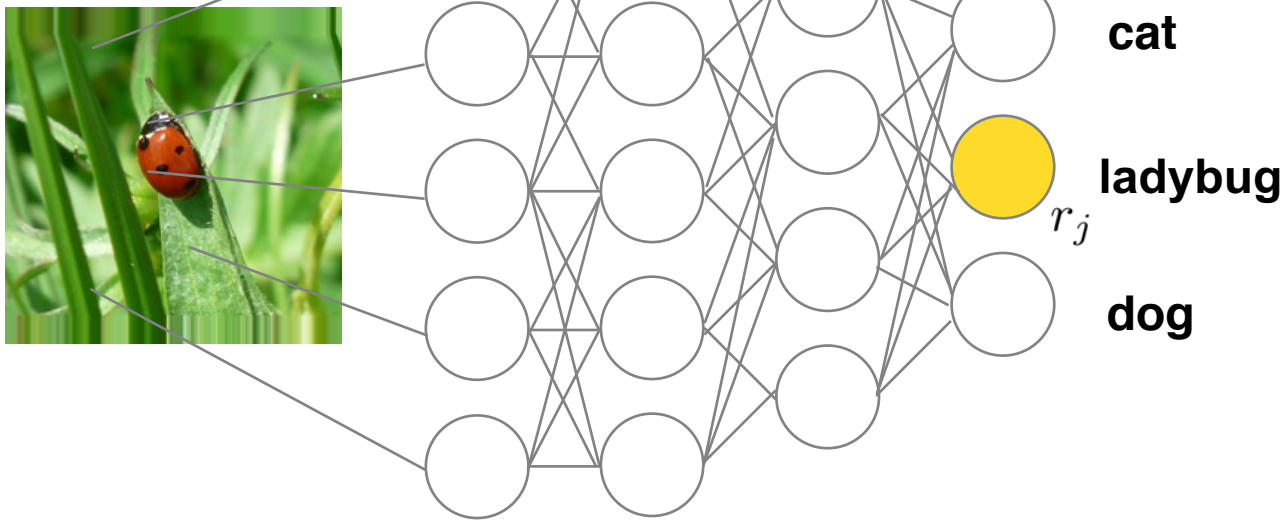
w_{ij}

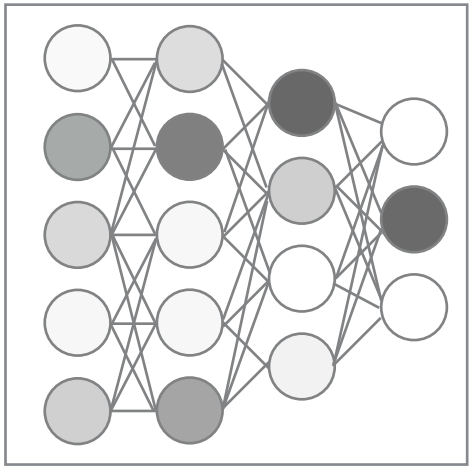
x_i x_j

$$x_j = \sigma(\sum_i x_i w_{ij} + b_j)$$

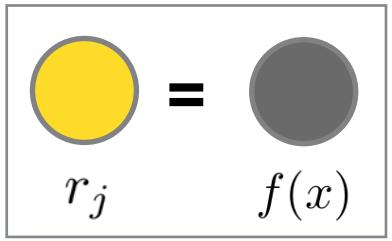
KI/ML nachvollziehbar machen

Explanation



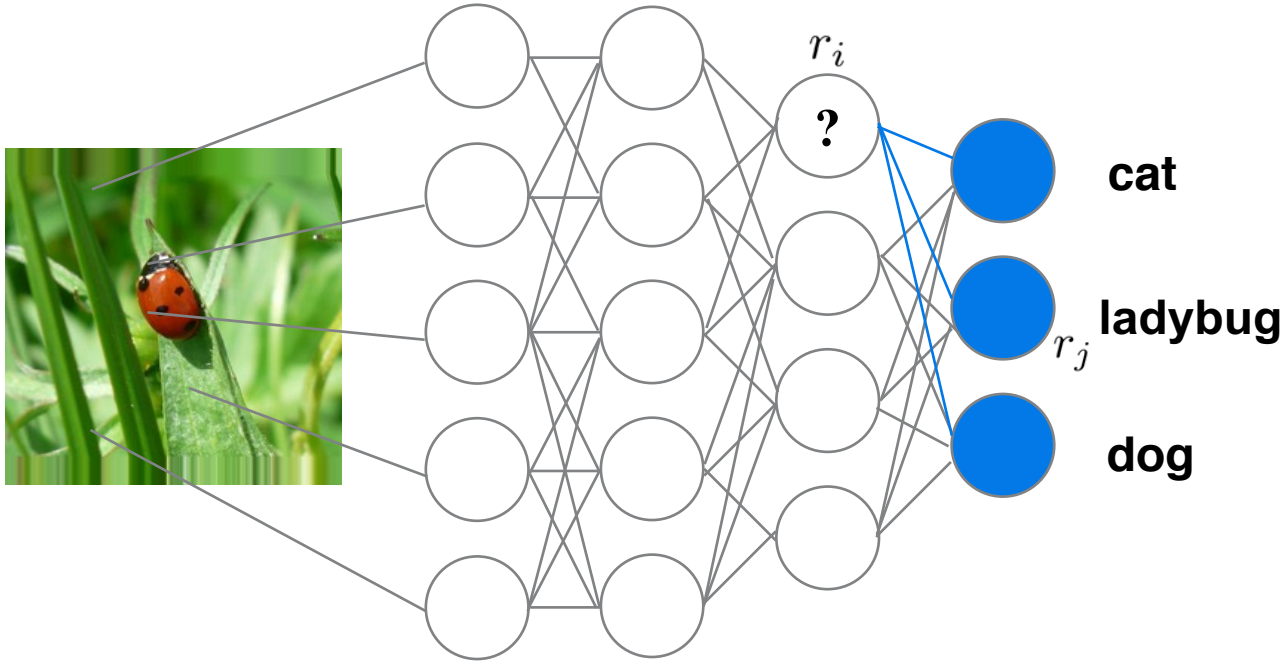


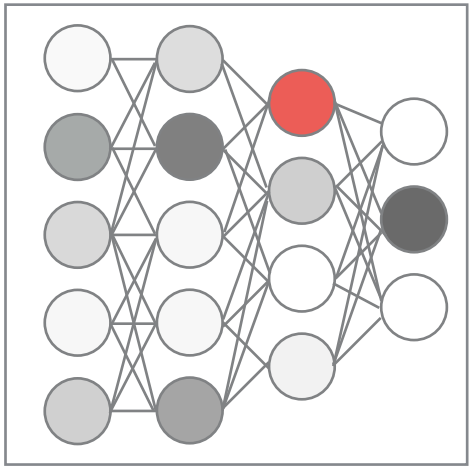
Initialization



KI/ML nachvollziehbar machen

Explanation

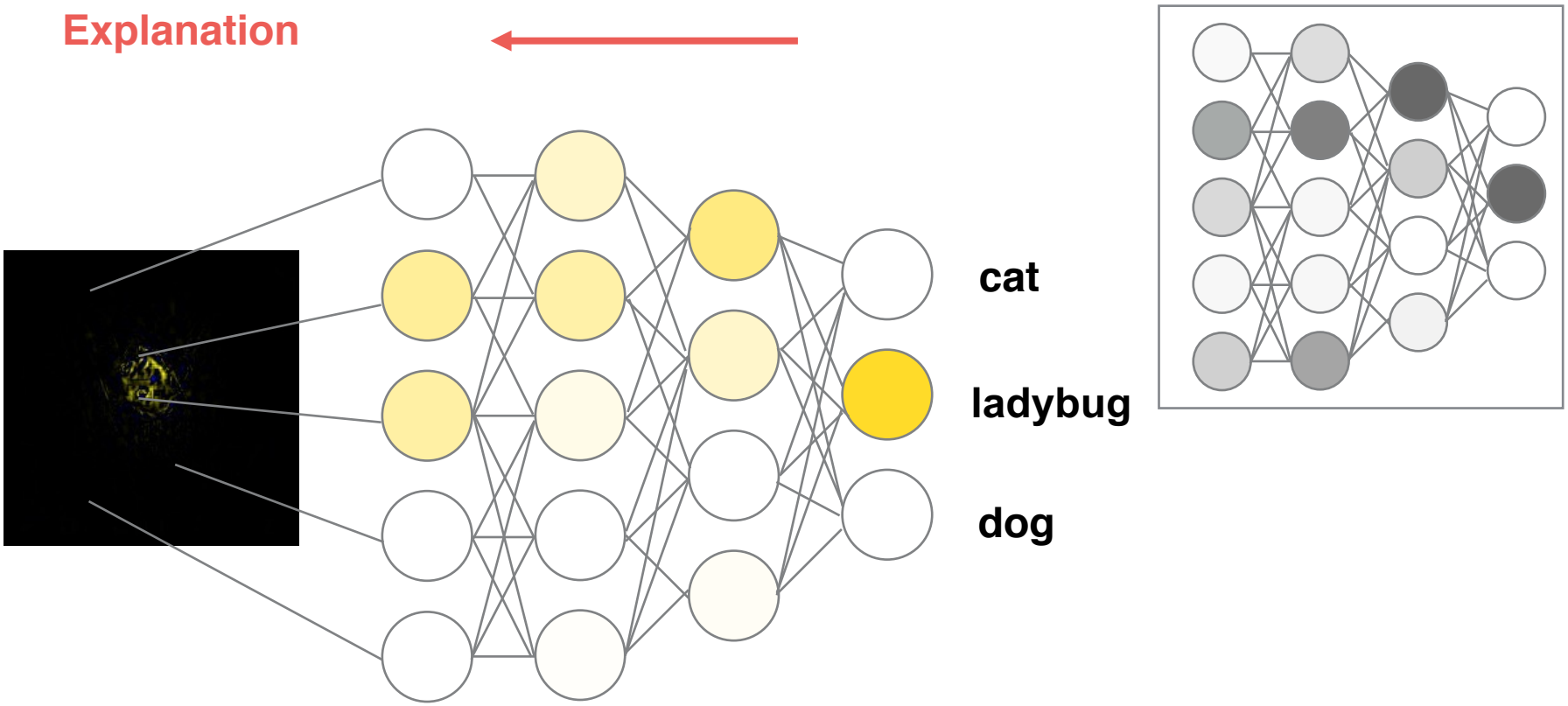




Theoretical interpretation
Deep Taylor Decomposition

$$r_i = x_i \sum_j \frac{w_{ij} r_j}{\sum_i x_i w_{ij}} = x_i c_i$$

KI/ML nachvollziehbar machen

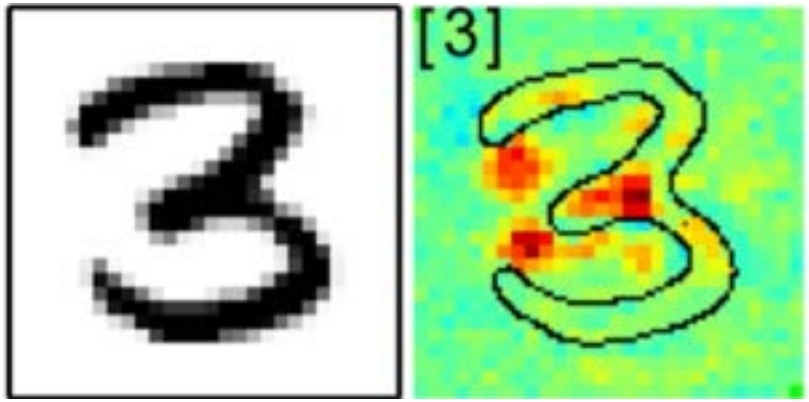


Relevance Conservation Property

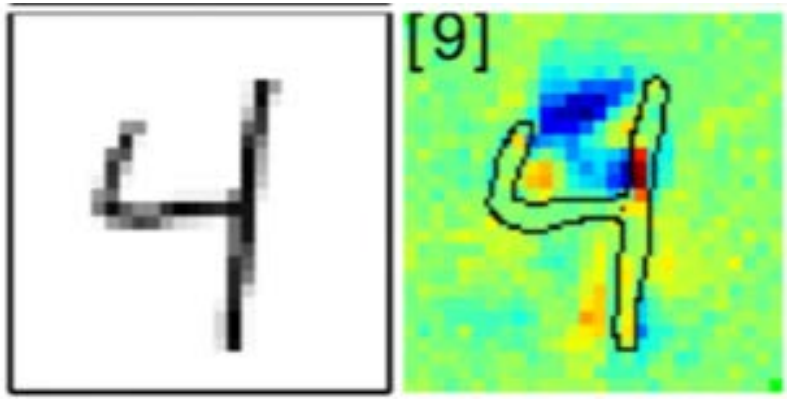
$$\sum_p r_p = \ldots = \sum_i r_i = \sum_j r_j = \ldots = f(x)$$

KI/ML nachvollziehbar machen

what speaks for / against classification as “3”



what speaks for / against classification as “9”



[number]: explanation target class

red color: evidence for prediction

blue color: evidence against prediction

KI/ML nachvollziehbar machen

Test error for various classes:

	aeroplane	bicycle	bird	boat	bottle	bus	car
Fisher	79.08%	66.44%	45.90%	70.88%	27.64%	69.67%	80.96%
DeepNet	88.08%	79.69%	80.77%	77.20%	35.48%	72.71%	86.30%
	cat	**chair**	**cow**	**diningtable**	**dog**	**horse**	**motorbike**
Fisher	59.92%	51.92%	47.60%	58.06%	42.28%	80.45%	69.34%
DeepNet	81.10%	51.04%	61.10%	64.62%	76.17%	81.60%	79.33%
	person	**pottedplant**	**sheep**	**sofa**	**train**	**tvmonitor**	**mAP**
Fisher	85.10%	28.62%	49.58%	49.31%	82.71%	54.33%	59.99%
DeepNet	92.43%	49.99%	74.04%	49.48%	87.07%	67.08%	72.12%

Image

KI/ML nachvollziehbar machen

‘horse’ images in PASCAL VOC 2007

KI/ML nachvollziehbar machen

Text classified as “sci.med” —> LRP identifies most relevant words.

sci.med (4.1)

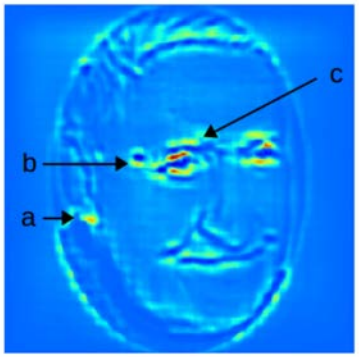
Yes, weightlessness does feel like falling. It may feel strange at first,
but the body does adjust. The feeling is not too different from that
of sky diving.

>And what is the motion sickness
>that some astronauts occasionally experience?

It is the body's reaction to a strange environment. It appears to be induced
partly to physical discomfort and part to mental distress. Some people are
more prone to it than others, like some people are more prone to get sick
on a roller coaster ride than others. The mental part is usually induced by
a lack of clear indication of which way is up or down, ie: the Shuttle is
normally oriented with its cargo bay pointed towards Earth, so the Earth
(or ground) is "above" the head of the astronauts. About 50% of the astronauts
experience some form of motion sickness, and NASA has done numerous tests in
space to try to see how to keep the number of occurances down.

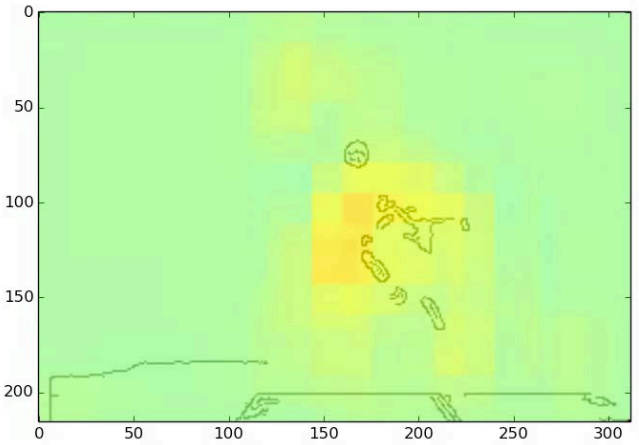
KI/ML nachvollziehbar machen

Identifying age-related features



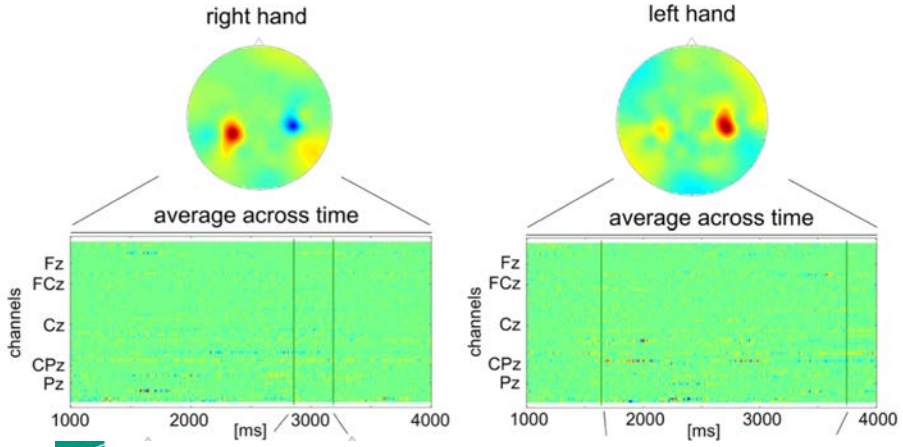
(Arbabzadah et al., GCPR, 2016)

Human action recognition



(Srinivasan et al., ICASSP, 2017)

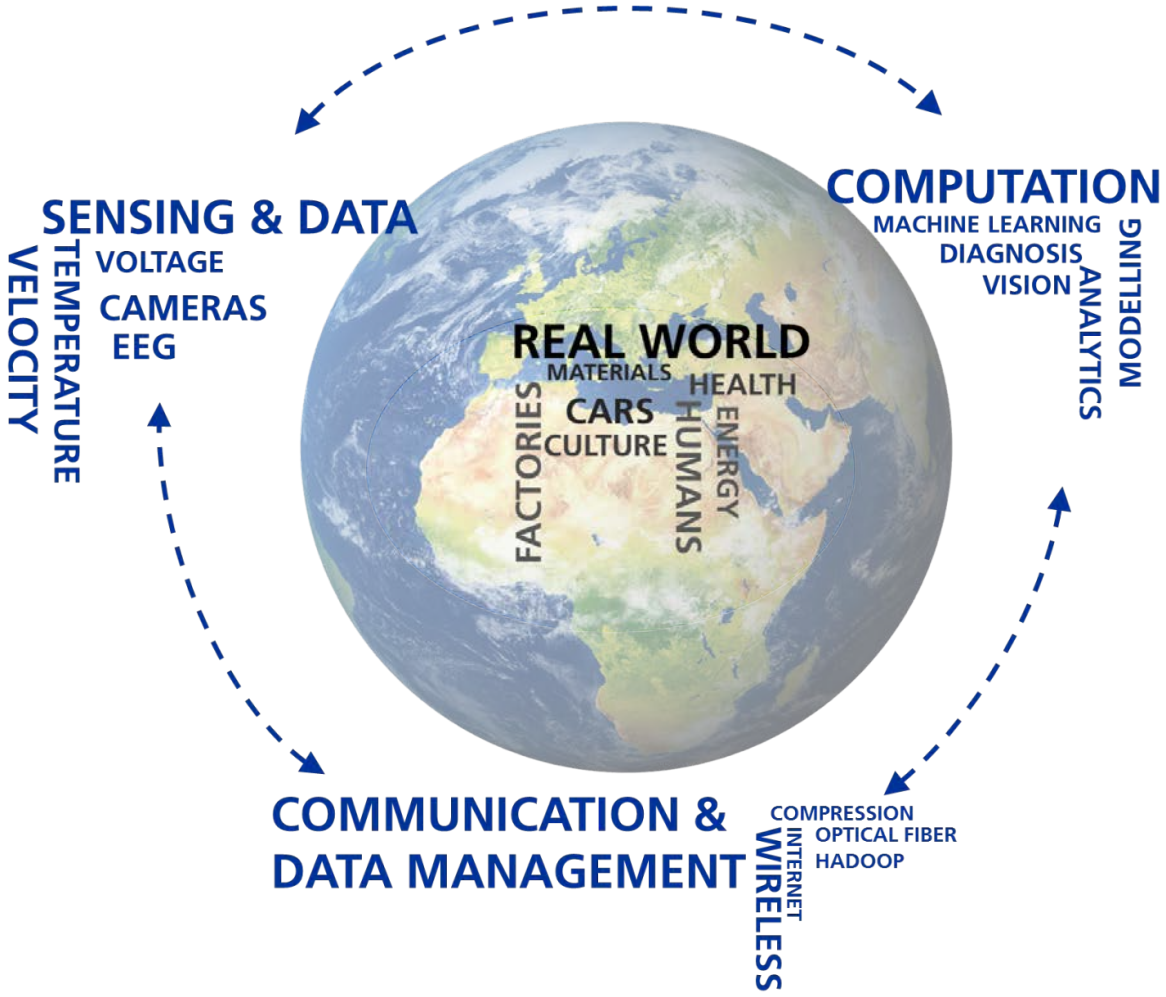
Brain-Computer Interfacing



(Sturm et al., J Neuroscience Methods, 2016)

Fraunhofer Heinrich Hertz Institute

Digital Transformation: Sensing, Communication, Computation



Zusammenfassung

- Künstliche Intelligenz und Maschinelles Lernen haben in den letzten Jahren große Fortschritte machen können
- Die Prozesse werden sich verstärken durch das gleichzeitige Anwachsen der Rechenkapazität und der Datenmengen
- KI/ML sind ein integraler Bestandteil der digitalen Transformation
- Wenn KI/ML wichtige Entscheidungen treffen, müssen diese erklärbar/nachvollziehbar sein

Acknowledgement & Support

HHI/TUB members and research associates, in particular:

- W. Samek, …
- K.-R. Müller, …
- S. Stanczak, …
- …

Supported by

Federal Ministry of Economics and Technology

SPONSORED BY THE

Federal Ministry of Education and Research

Deutsche Forschungsgemeinschaft DFG

Questions?