FIZ Karlsruhe
Leibniz-Institut für Informationsinfrastruktur

KIT
Karlsruher Institut für Technologie

Ein Skateboard für den Papst

oder
Warum es maschinelles Lernen ohne Semantik so schwer hat

Prof. Dr. Harald Sack
Netzwerk maschinelle Verfahren in der Erschließung
Frankfurt, 11. Oktober. 2019

FIZ Karlsruhe
Leibniz-Institut für Informationsinfrastruktur

KIT
Karlsruher Institut für Technologie

Ein Skateboard für den Papst

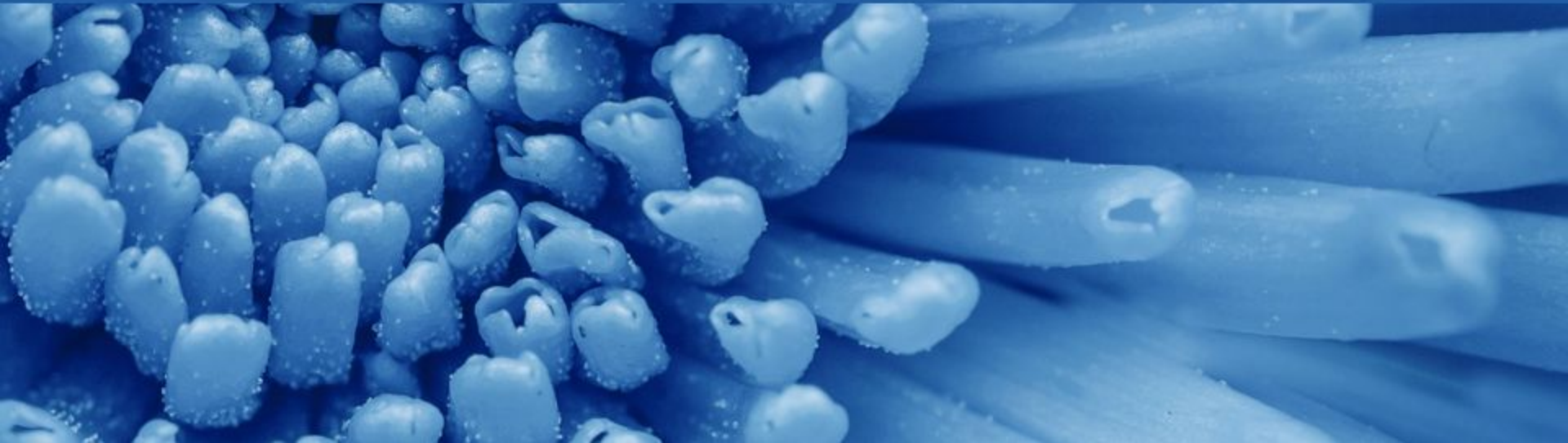
oder

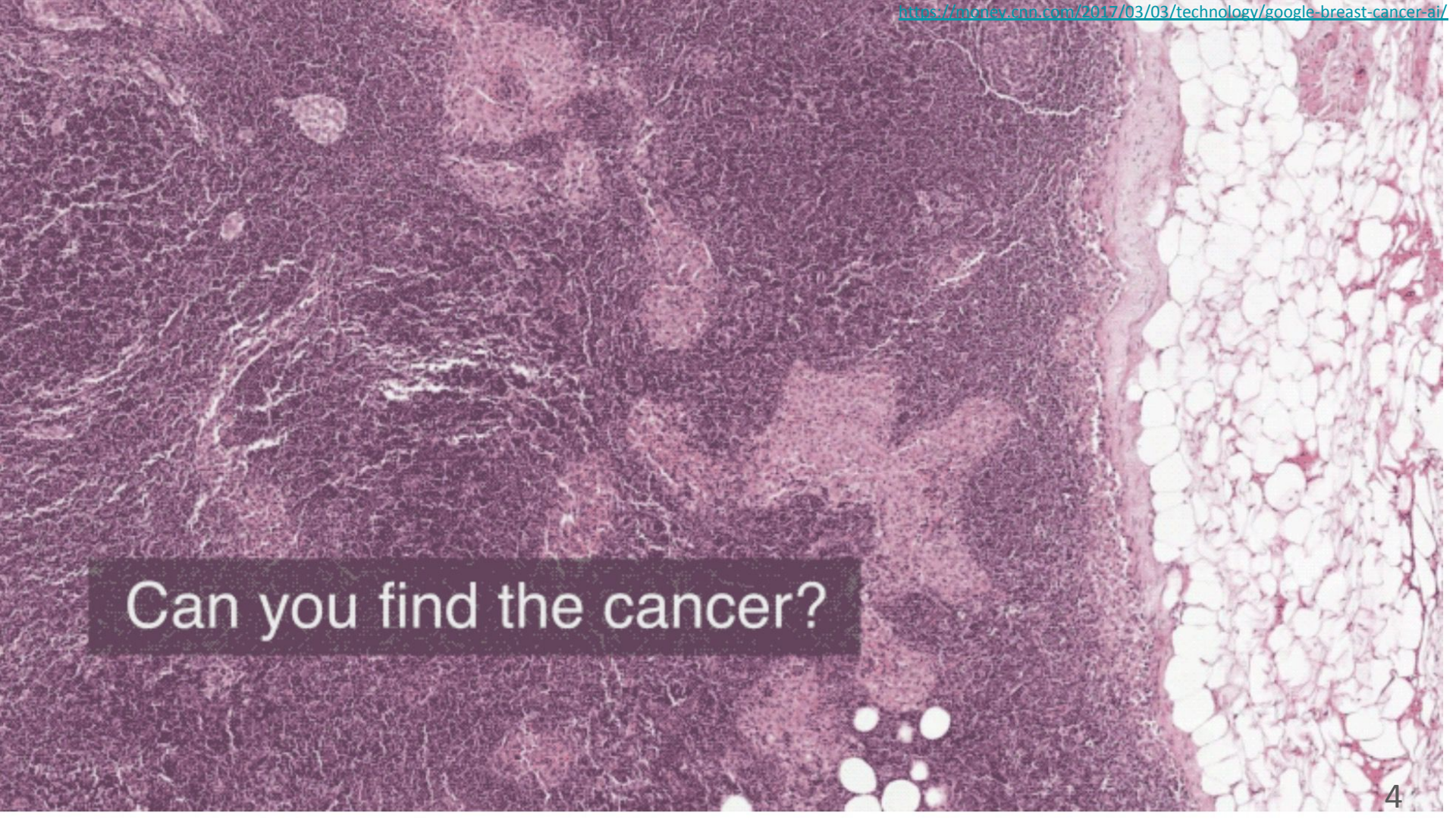
Warum es maschinelles Lernen ohne Semantik so schwer hat

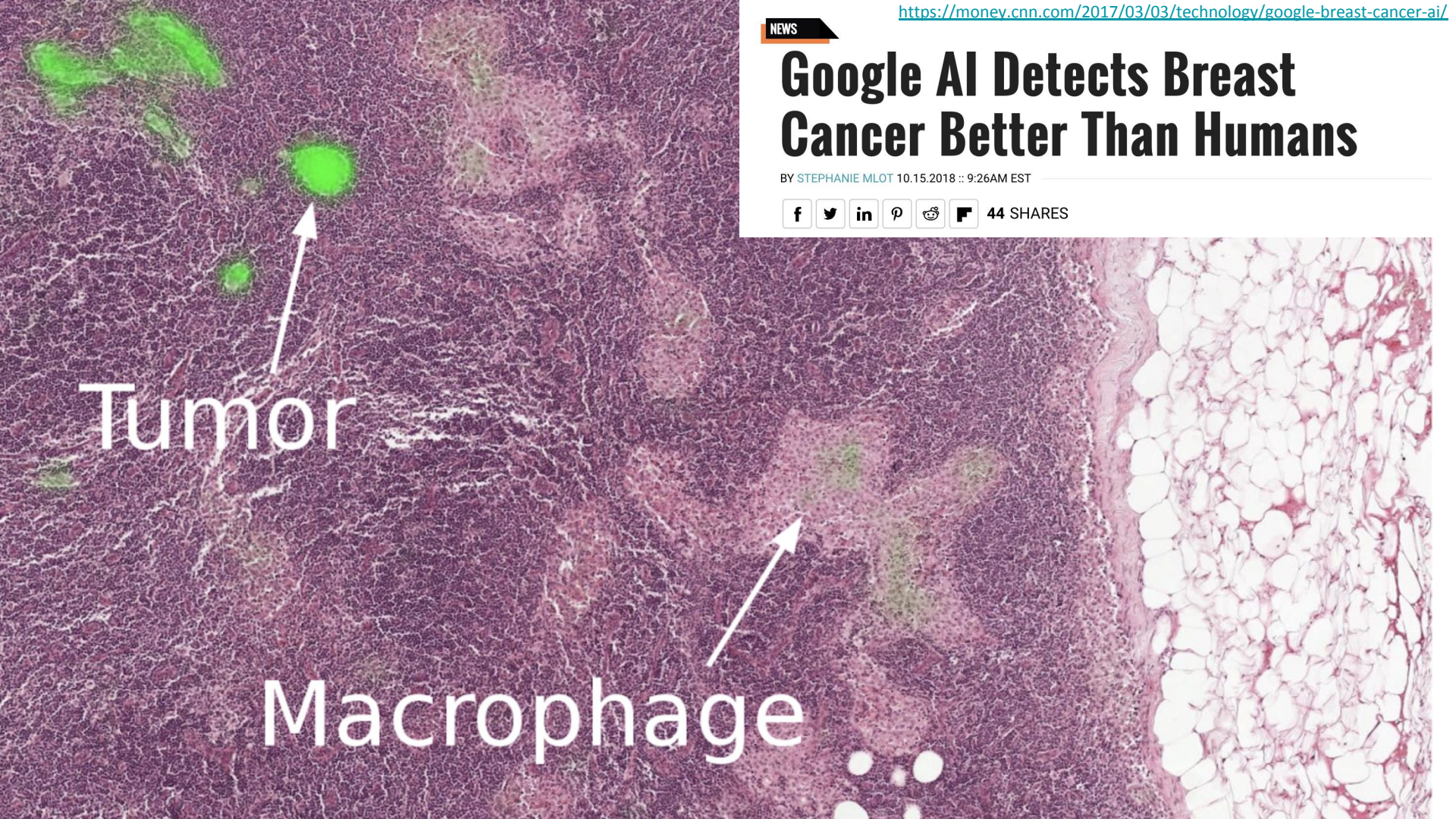
Hugo von Cluny, Mathilde von Tuszien und Heinrich IV.
Miniatur aus der Handschrift Vita Mathildis (um 1115)

“Jede hinreichend fortgeschrittene Technologie ist von Magie nicht mehr zu unterscheiden.”

Arthur C. Clarke, Profiles of the Future (1973)




https://money.cnn.com/2017/03/03/technology/google-breast-cancer-ai/
Can you find the cancer?



https://money.cnn.com/2017/03/03/technology/google-breast-cancer-ai/

NEWS

Google AI Detects Breast Cancer Better Than Humans

BY STEPHANIE MLOT 10.15.2018 :: 9:26AM EST

44 SHARES

Science

AlphaGo Zero: Google DeepMind supercomputer learns 3,000 years of human knowledge in 40 days

AlphaGo

http://www.telegraph.co.uk/science/2017/10/18/alphago-zero-google-deepmind-supercomputer-learns-3000-years/

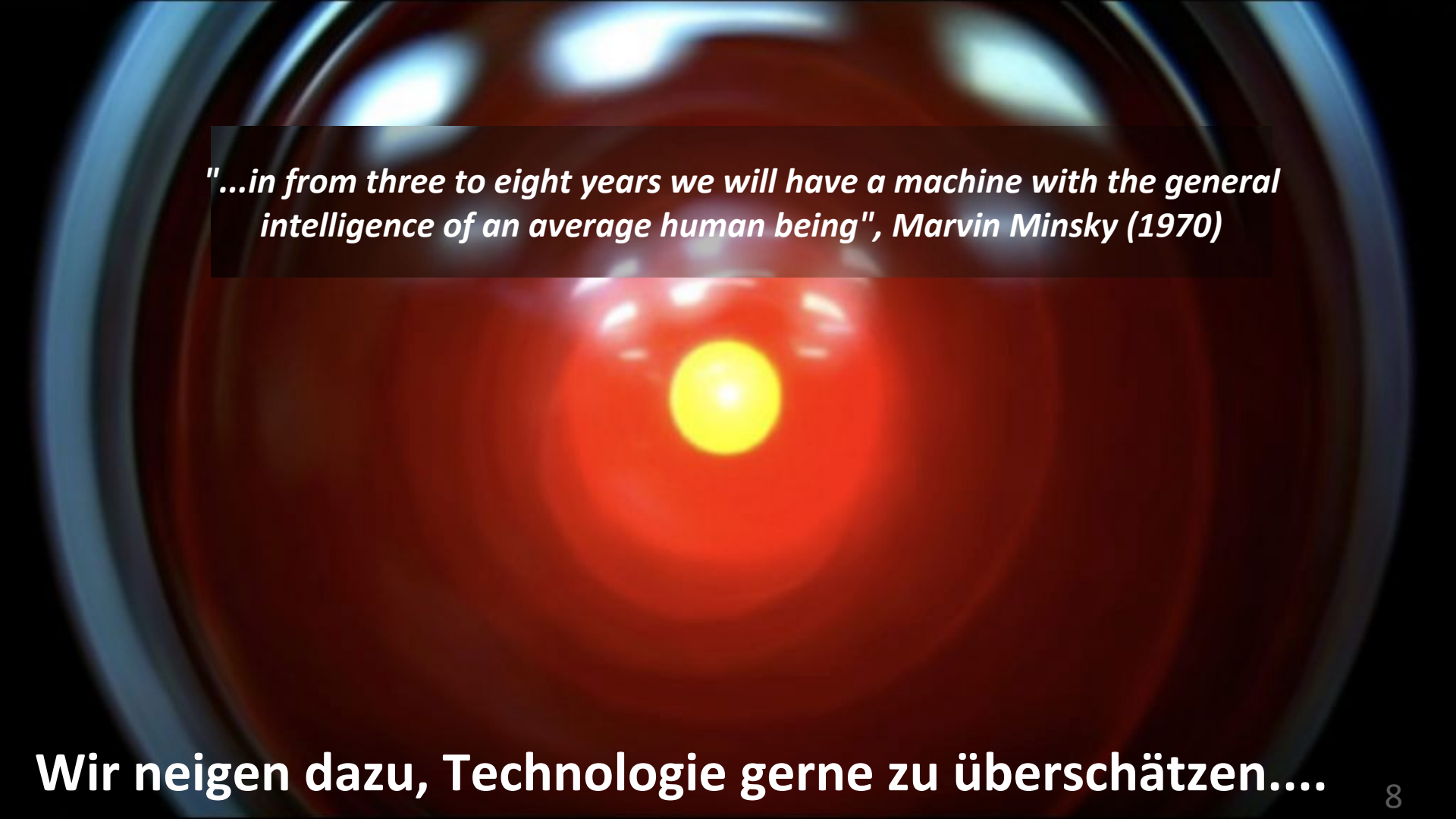
https://www.christies.com/features/A-collaboration-between-two-artists-one-human-one-a-machine-9332-1.aspx

Is artificial intelligence set to become art's next medium?

16 October 2018

PHOTOGRAPHS & PRINTS |

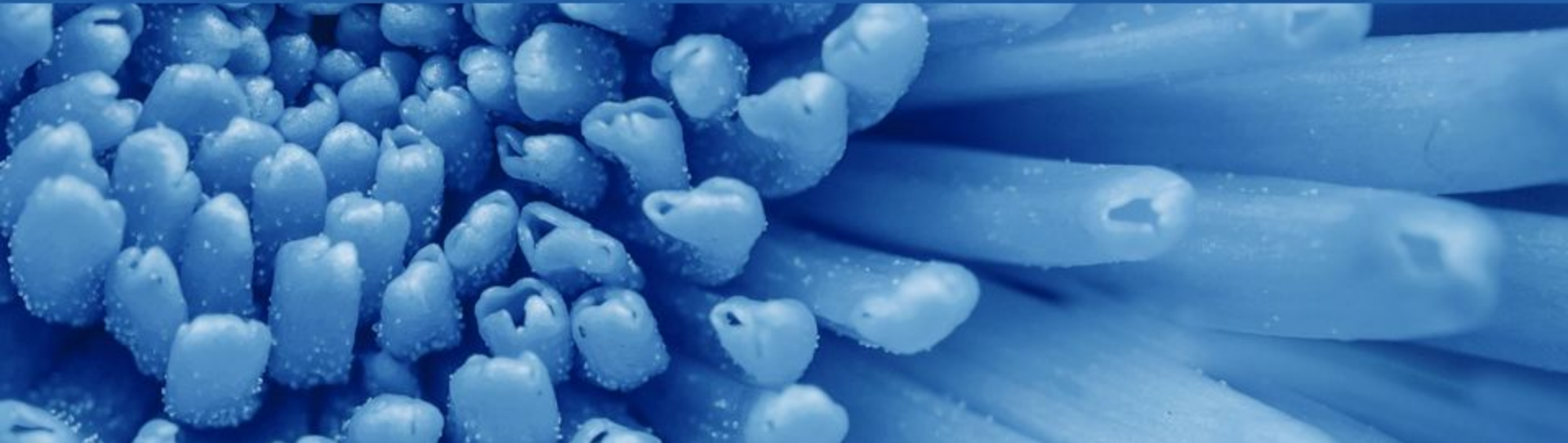
AI artwork sells for $432,500 — nearly 45 times its high estimate — as Christie's becomes the first auction house to offer

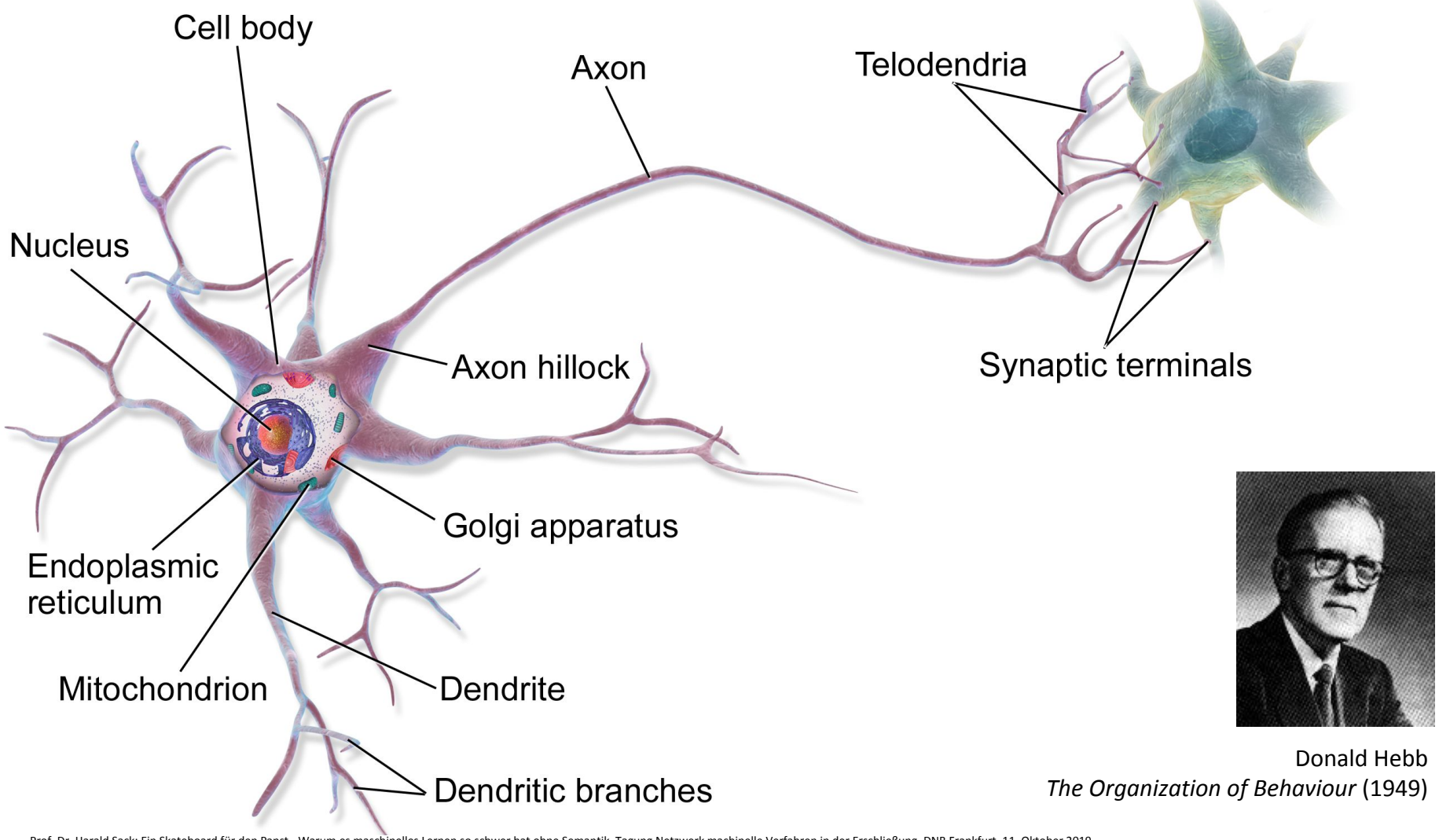


Wir neigen dazu, Technologie gerne zu überschätzen....

Aber woher kommt nun die “Magie”...?

Eine kurze Geschichte des maschinellen Lernens

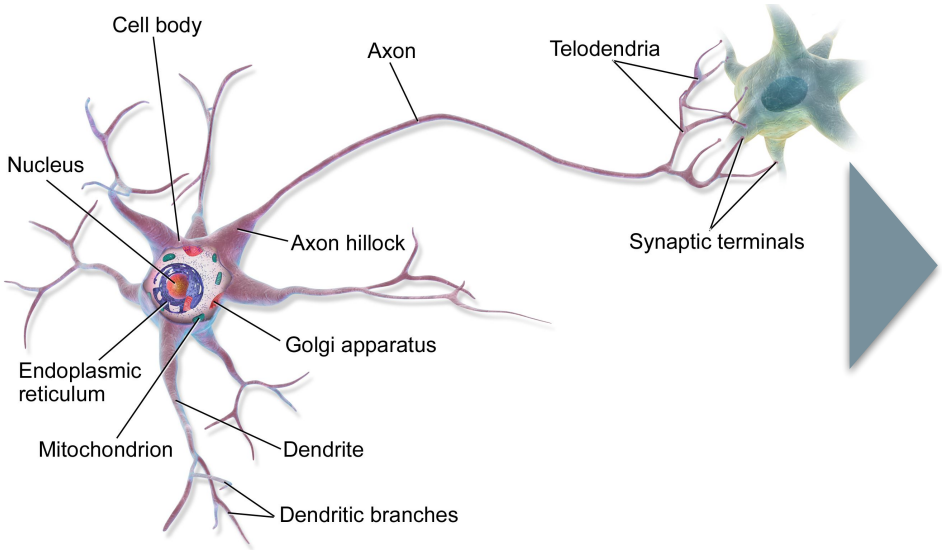


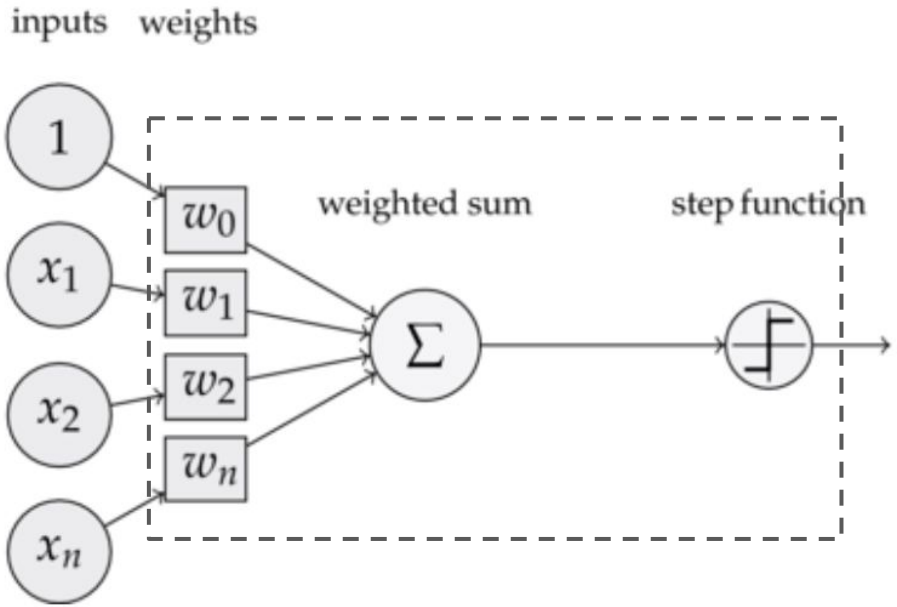


Donald Hebb
The Organization of Behaviour (1949)

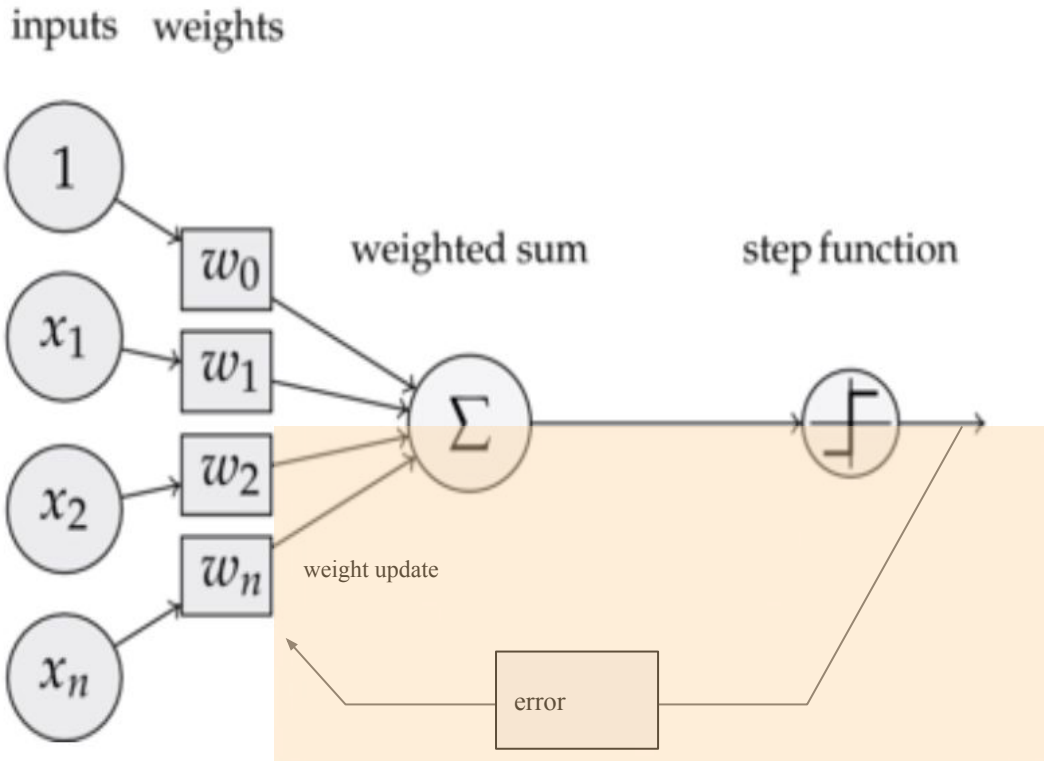
Von der Biologie zum Mathematischen Modell

Warren Sturgis McCulloch & Walter Pitts
A Logical Calculus of the Ideas Immanent in Nervous Activity (1943)





Perceptron Algorithmus

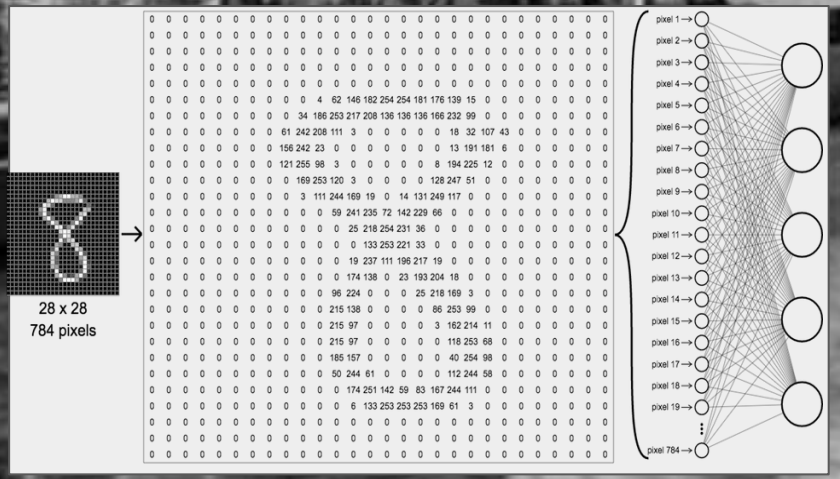


Frank Rosenblatt

The perceptron: a probabilistic model for information storage and organization in the brain. (1958)

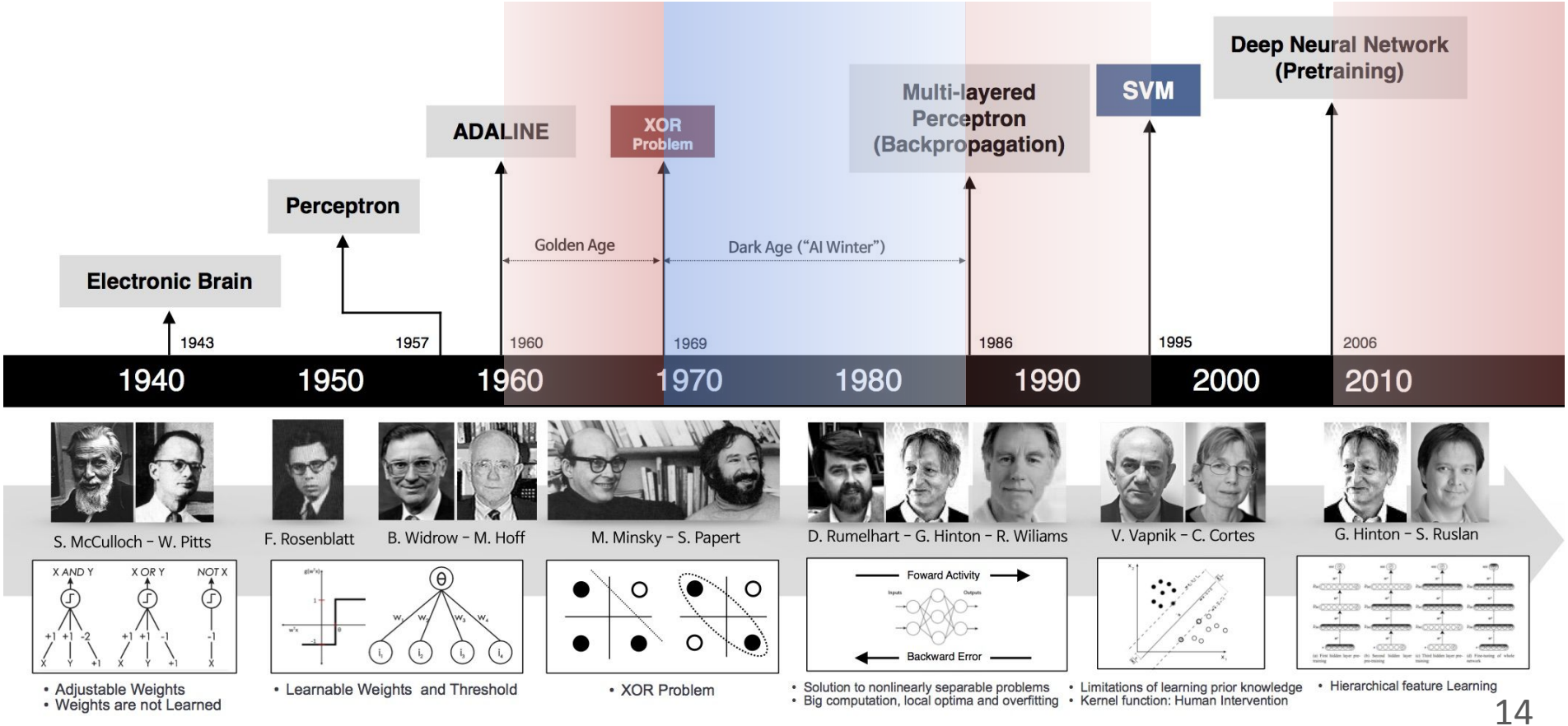
$$w_{ij}^{\text{neu}} = w_{ij}^{\text{alt}} + \Delta w_{ij},$$
$$\Delta w_{ij} = \alpha \cdot (t_j - o_j) \cdot x_i.$$

Learning by Example - MARK 1 Perceptron (1957)

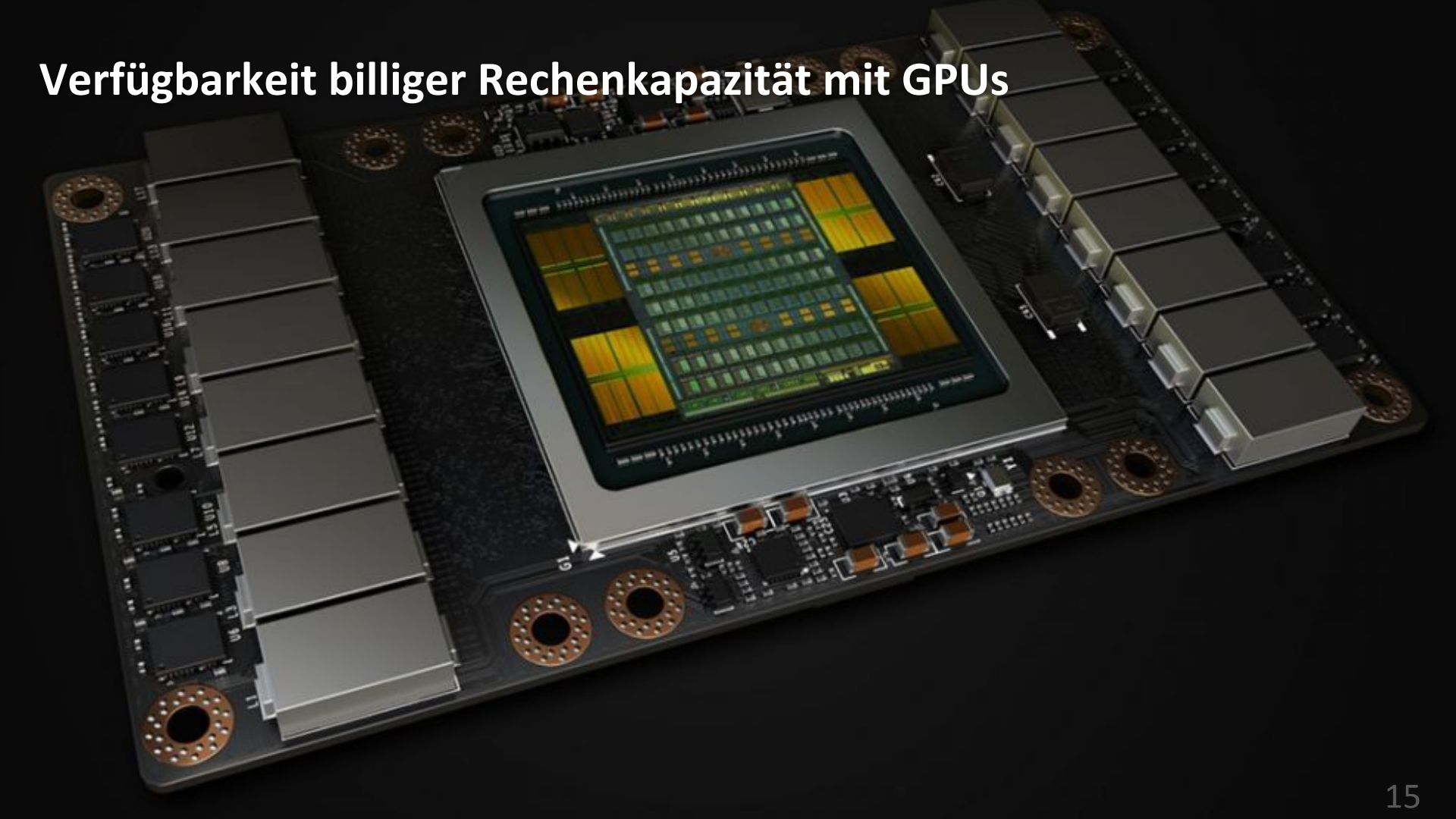


Perceptron for Image Recognition

Der Siegeszug des konnektionistischen Paradigmas

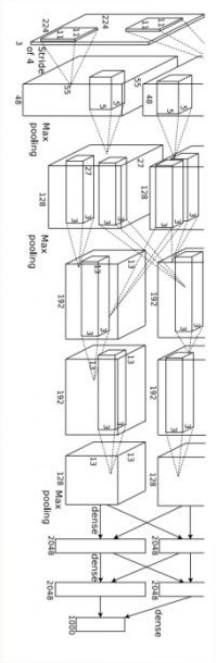


Verfügbarkeit billiger Rechenkapazität mit GPUs



Wiederverwendbare hochtrainierte komplexe Modelle

“AlexNet”



[Krizhevsky et al. NIPS 2012]

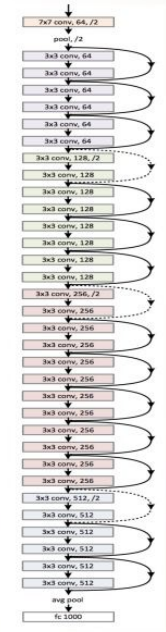
“GoogLeNet”

[Szegedy et al. CVPR 2015]

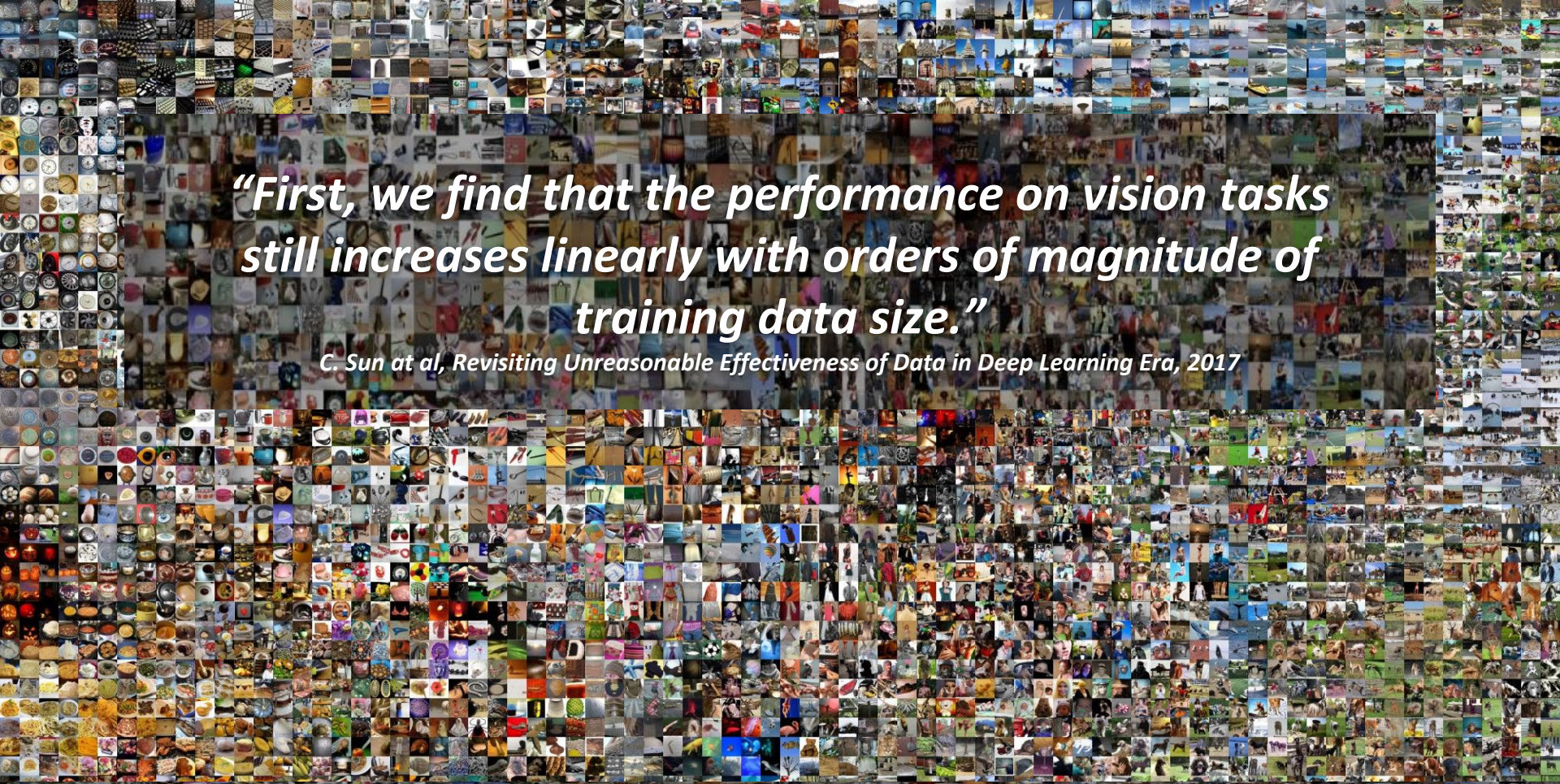
“VGG Net”

[Simonyan & Zisserman, ICLR 2015]

“ResNet”

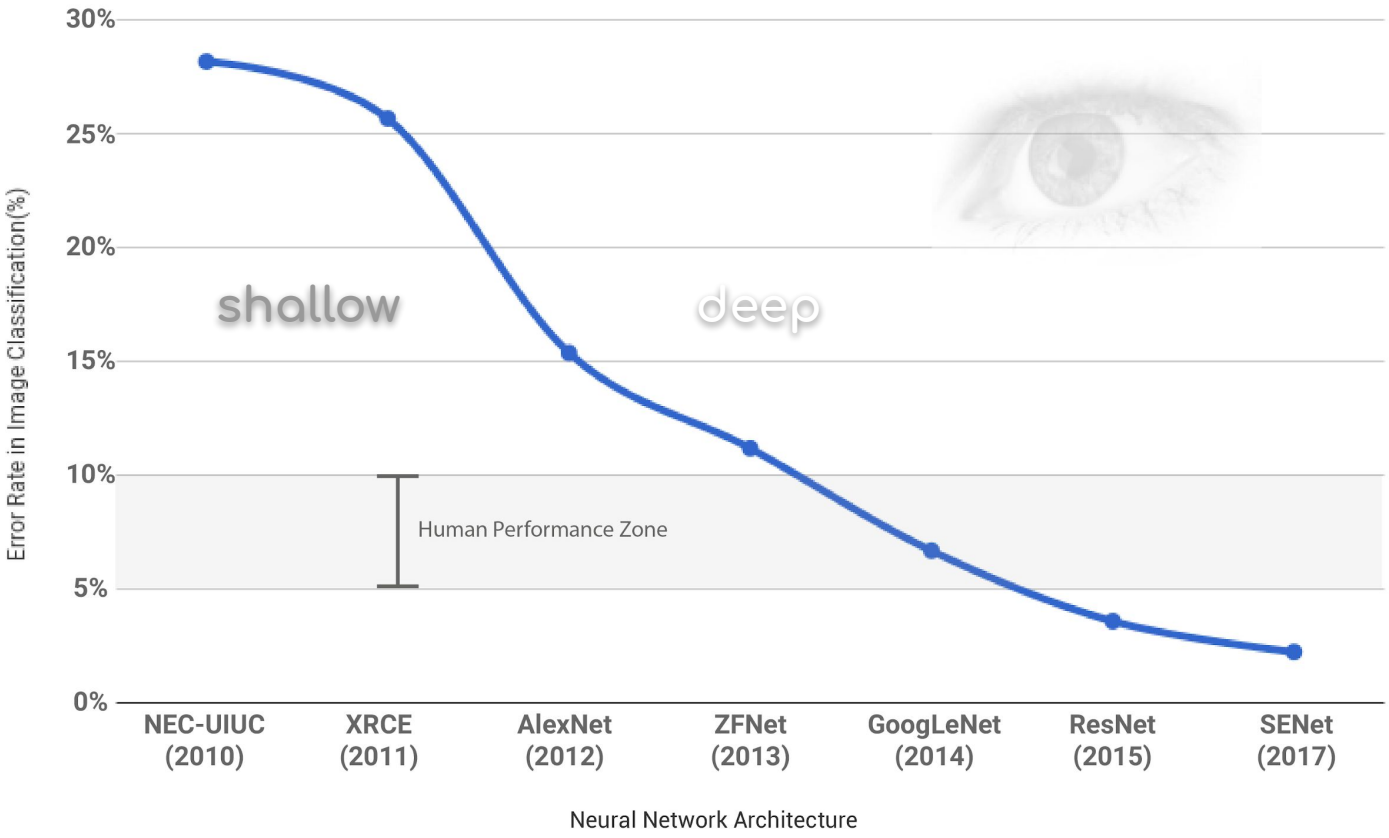


[He et al. CVPR 2016]



Verfügbarkeit großer Mengen annotierter Trainingsdaten

Deep Learning für die visuelle Analyse



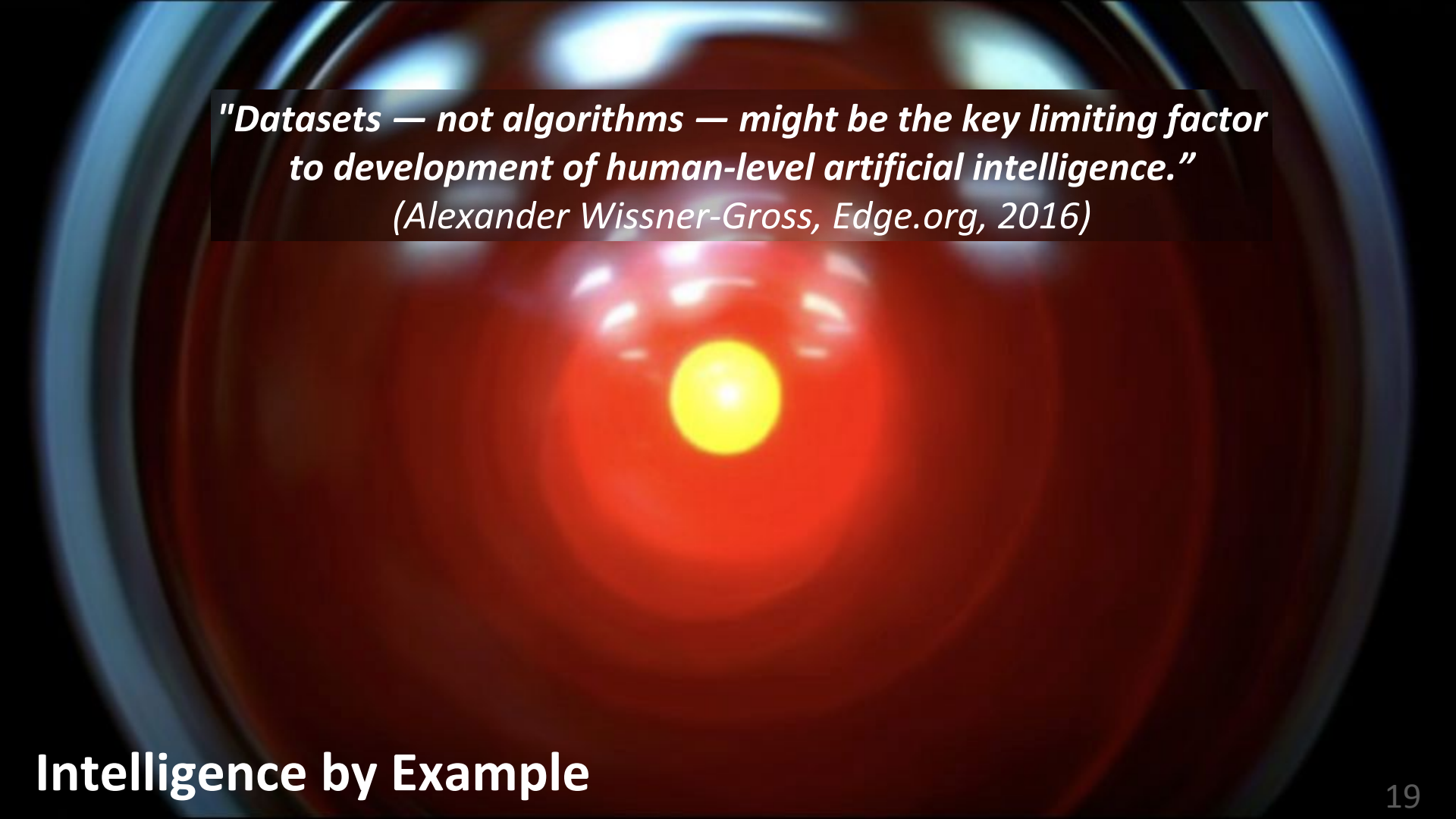
IMAGENET

Large Scale Visual Recognition Challenge (ILSVRC)

http://image-net.org/challenges/LSVRC/

KIT Karlsruher Institut für Technologie

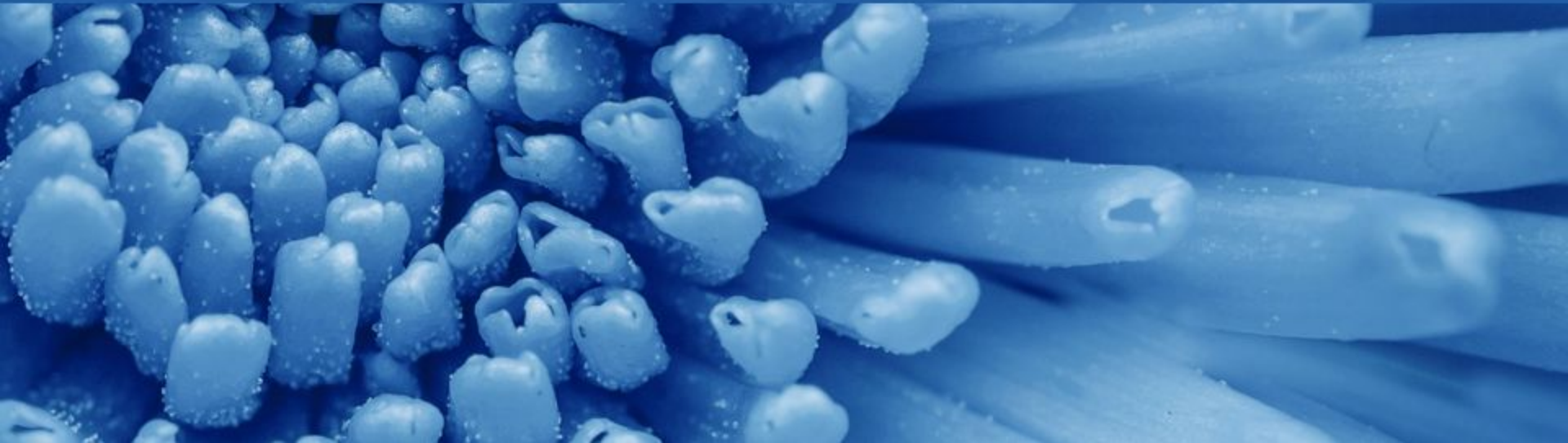
FIZ Karlsruhe
Leibniz-Institut für Informationsinfrastruktur



Intelligence by Example

“Es ist der Geist, der sich den Körper baut.”

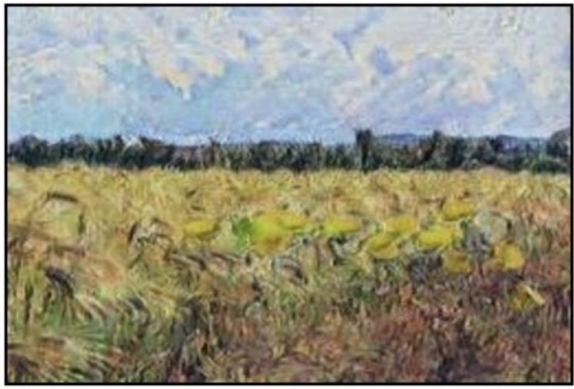
Friedrich Schiller, Wallensteins Tod (1799)



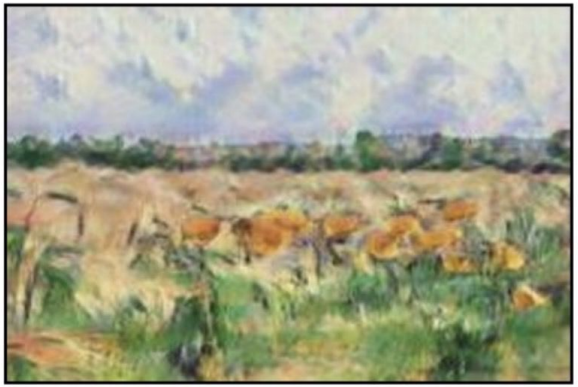
Generierung neuer Inhalte

Photograph

Monet



Van Gogh



Cezanne

Ukiyo-e

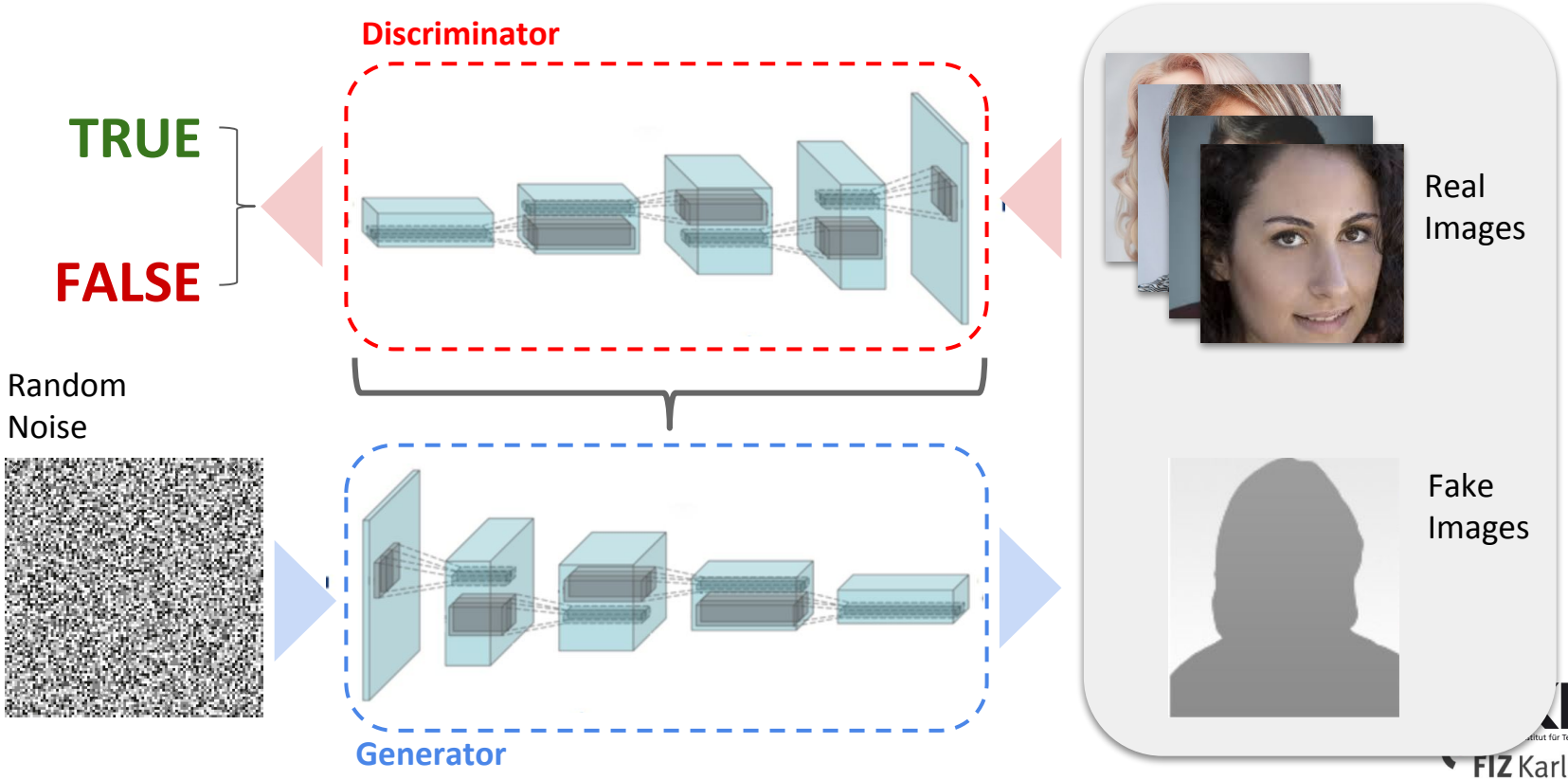
https://junyanz.github.io/CycleGAN/images/teaser_high_res.jpg

Generierung neuer Inhalte

Monet → Photograph

https://junyanz.github.io/CycleGAN/images/teaser_high_res.jpg

Vergleichendes Lernen - Generative Adversarial Networks



Schmidthuber, 1992
Niemitalo, 2010
Goodfellow, 2014

FIZ Karlsruhe
Leibniz-Institut für Informationsinfrastruktur

https://github.com/jantic/DeOldify

Generierung neuer Inhalte

Generierung neuer Inhalte

https://junyanz.github.io/CycleGAN/

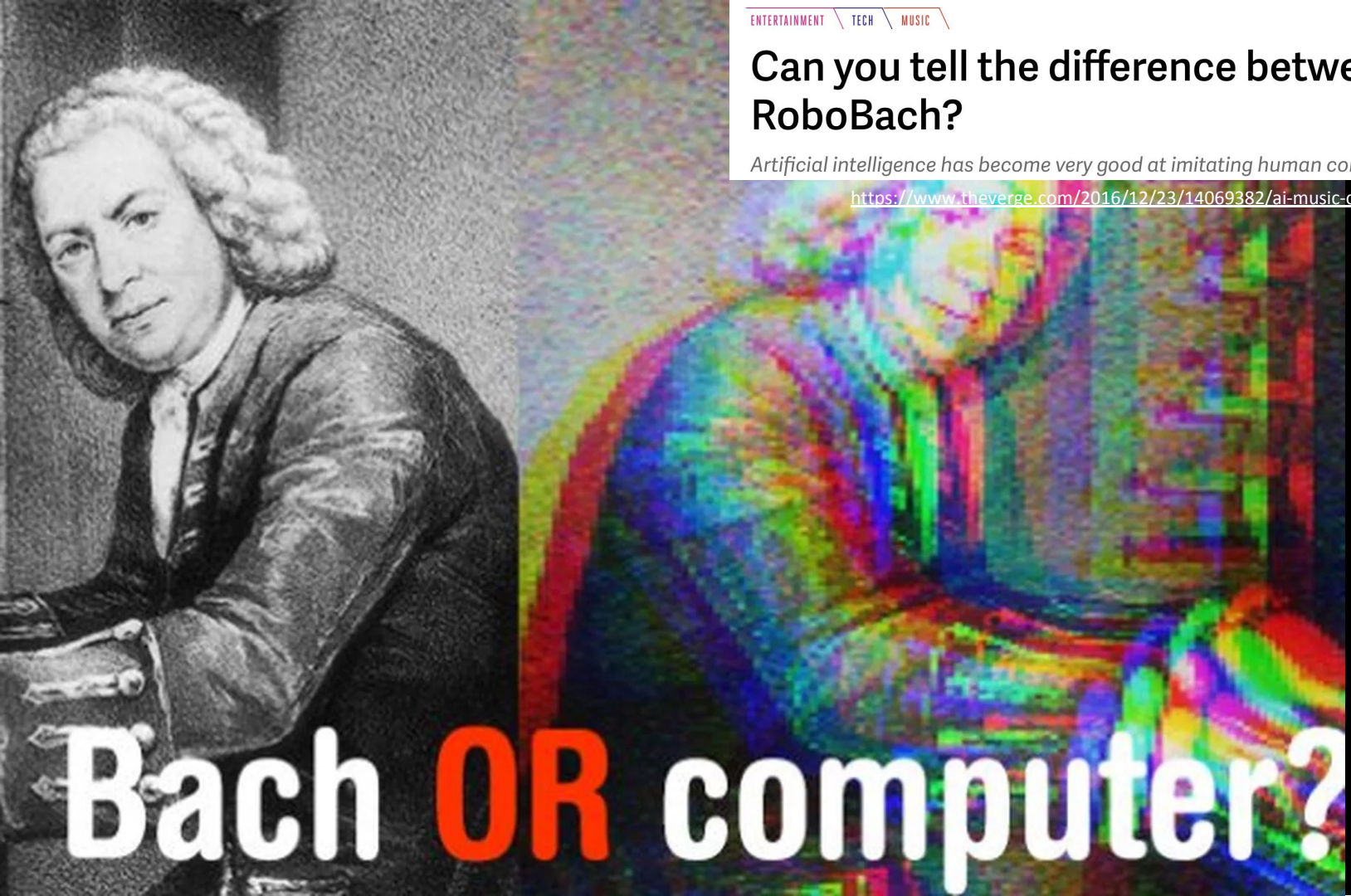
Generierung neuer Inhalte

http://webdav.tuebingen.mpg.de/pixel/enhancenet/

Low-resolution input

Our result

Original image



ENTERTAINMENT | TECH | MUSIC

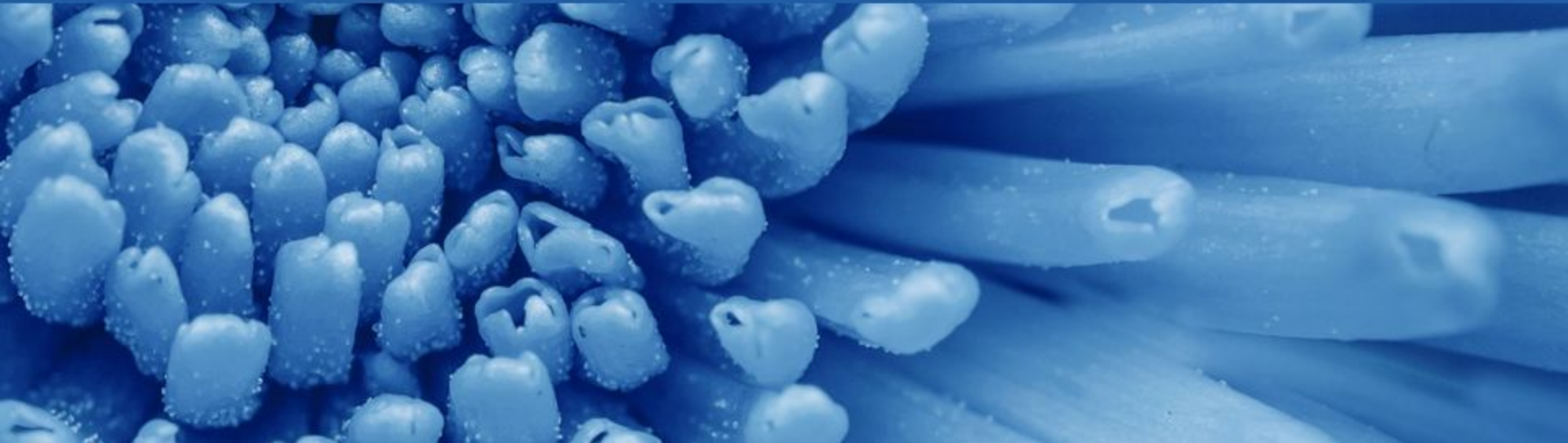
Can you tell the difference between Bach and RoboBach?

Artificial intelligence has become very good at imitating human composers

https://www.theverge.com/2016/12/23/14069382/ai-music-creativity-bach-deepbach-csl

“Es ist nicht genug, zu wissen, man muss auch anwenden.”

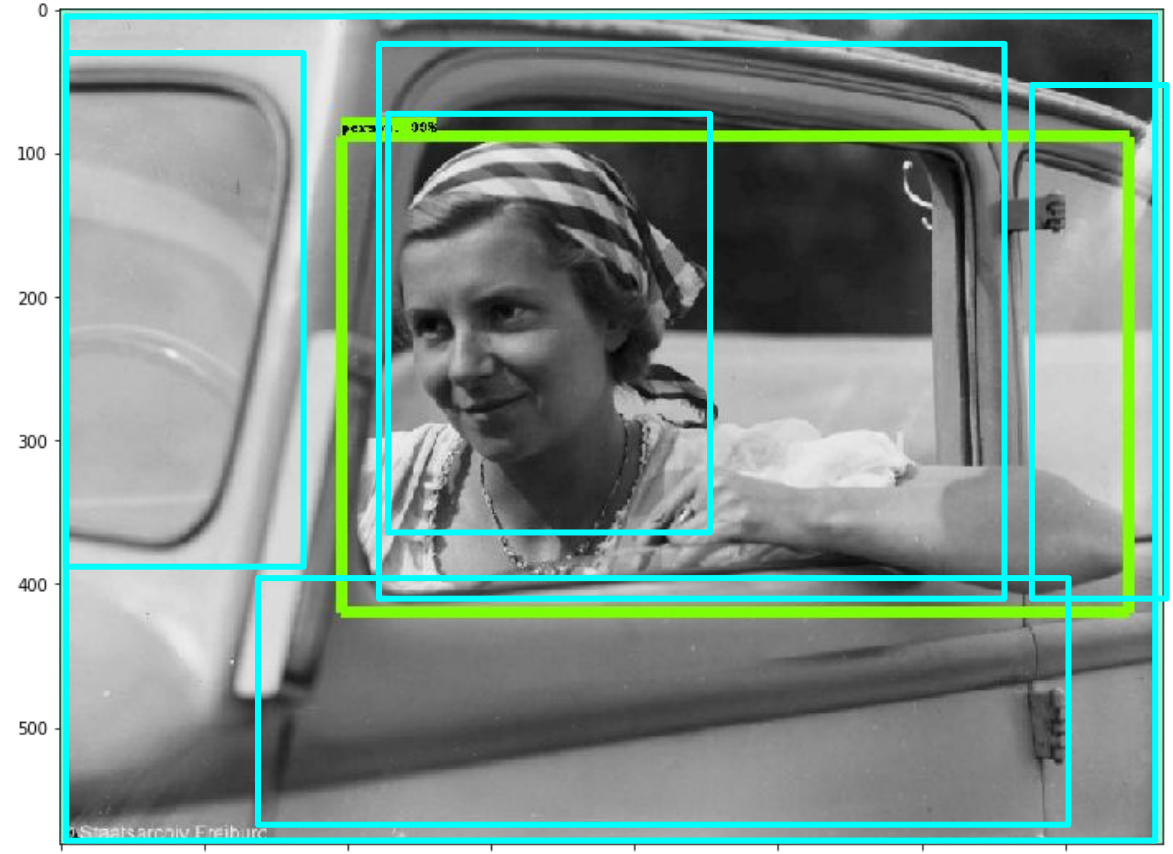
Johann Wolfgang von Goethe, Wilhelm Meisters Wanderjahre (1821/29)



Object Recognition

person

Fine Grain Object Recognition



black and white picture

car

window | window | window

woman

door

Räumliche und semantische Analyse

car

window | window | window

woman

door

```
Window    ⊑ ∃isPartOf.Car
Door      ⊑ ∃isPartOf.Car
Woman     ⊑ Person
Passenger ⊑ Person ⊓ ∃isTransportedBy.Car
...
```

External Knowledge

car

window | window | window

woman

door

Image Captioning

car

window | window | window

woman

door

A black and white photo of a woman sitting in a car looking out the window

a black and white picture of a woman

a woman sitting in a car

a woman looking out the window

a car with a door

a car with three windows

Question Answering

car

window | window | window

woman

door

A black and white photo of a woman sitting in a car looking out the window

Who is on the picture? **a woman**

Who is sitting in the car? **a woman**

```
<Picture123> <hasColour> <blackAndWhite>
<Picture123> <contains> <Car123>
<Picture123> <contains> <Door123>
<Picture123> <contains> <Woman123>
<Picture123> <contains> <Window123A>
<Picture123> <contains> <Window123B>
<Picture123> <contains> <Window123C>
<Window123B> <contains> <Woman123>
<Woman123>   <sittingIn> <Car123>
```

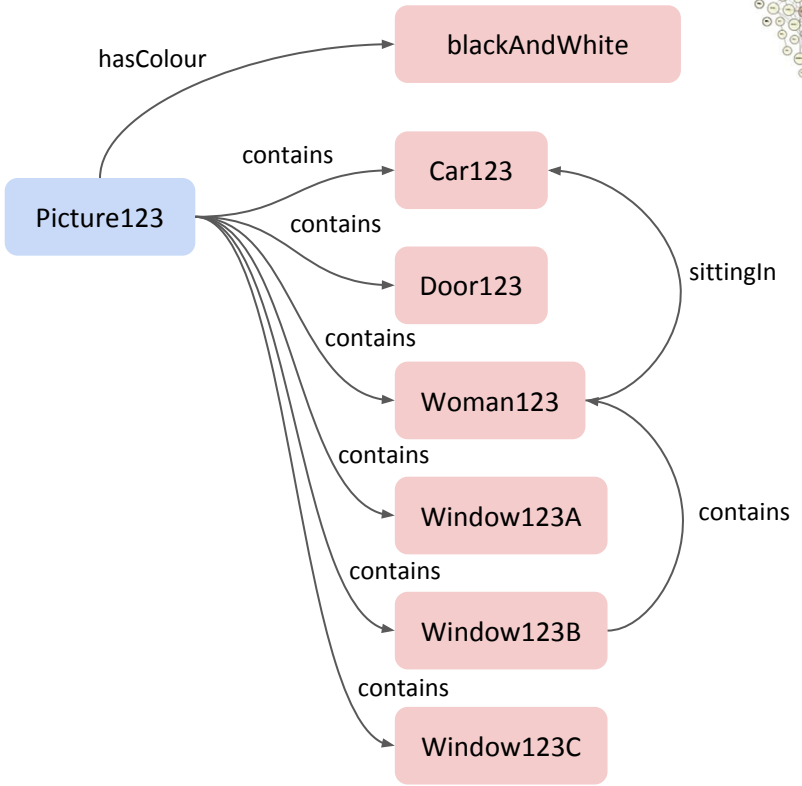
A Universe of Image Descriptions

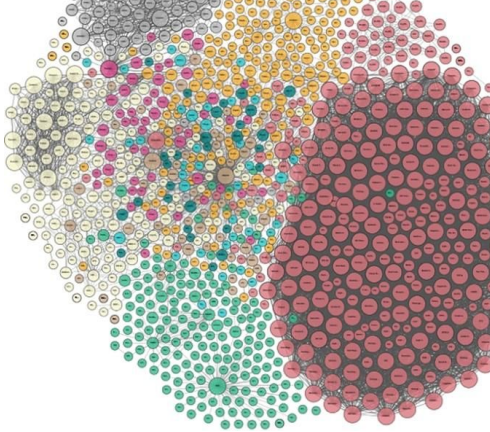
car

window | window | window

woman

door

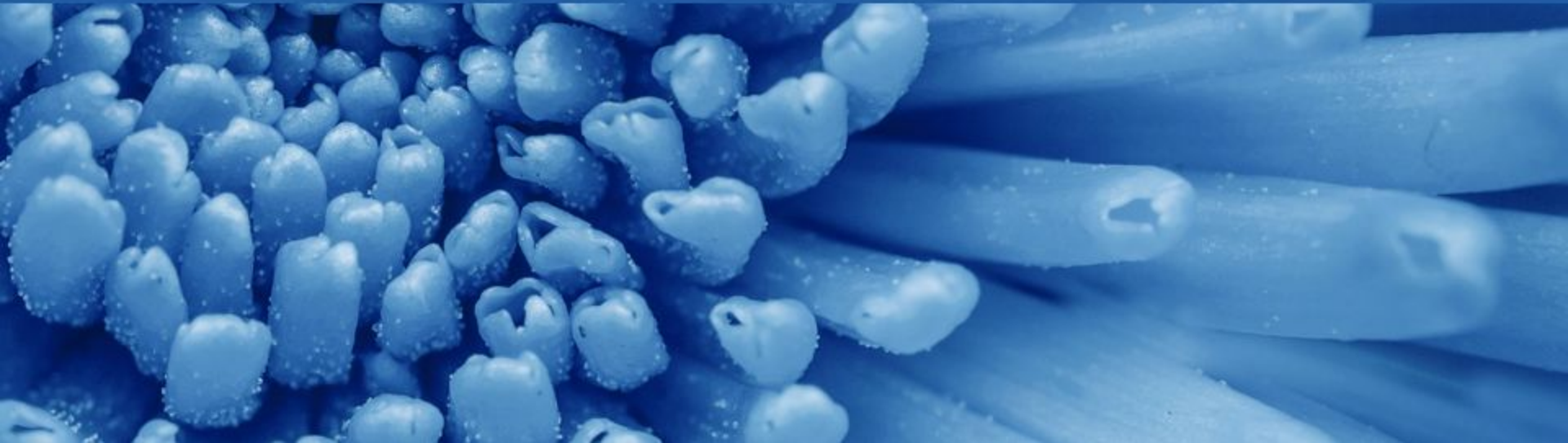




Karlsruher Institut für Technologie
FIZ Karlsruhe
Leibniz-Institut für Informationsinfrastruktur

“Grau ist alle Theorie. Entscheidend is’ auf’m Platz. ”

Alfred ‘Adi’ Preißler, Borussia Dortmund (1956)



Indexing with Pretrained State-of-the-Art Models

(1) Image Captioning (resnet50, 5m iterations)

```
0) a man standing in front of a microphone . (p=0.001874)
1) a man standing in front of a microphone holding a microphone . (p=0.000623)
2) a man in a suit and tie holding a microphone . (p=0.000467)
```

a man standing in front of a microphone .

(2) Visual Concept Detection (based on ImageNet 1K)

```
inception_v3: [[('n04296562', 'stage', 0.2797169), ('n03759954', 'microphone', 0.1635235), ('n04418357',
vgg16       : [[('n03372029', 'flute', 0.21658598), ('n03838899', 'oboe', 0.18556471), ('n04487394', 'tr
vgg19       : [[('n02879718', 'bow', 0.069563985), ('n04487394', 'trombone', 0.06794496), ('n03372029',
Resnet50    : [[('n02787622', 'banjo', 0.23410268), ('n03494278', 'harmonica', 0.09202912), ('n04286575'
Mobilenet_V2: [[('n02787622', 'banjo', 0.24096675), ('n04259630', 'sombrero', 0.07608573), ('n04487394',
Densenet    : [[('n03141823', 'crutch', 0.5543931), ('n03759954', 'microphone', 0.094545476), ('n0278762
NASAnet     : [[('n02787622', 'banjo', 0.40817854), ('n02676566', 'acoustic_guitar', 0.10958237), ('n042
```

Indexing with Pretrained State-of-the-Art Models

(1) Image Captioning (resnet50, 5m iterations)

```
0) a black and white photo of a man and a horse . (p=0.000597)
1) a black and white photo of a man with a horse . (p=0.000575)
2) a black and white photo of a man and a horse (p=0.000337)
```

a black and white photo of a man and a horse .

(2) Visual Concept Detection (based on ImageNet 1K)

```
inception_v3: [[('n03538406', 'horse_cart', 0.5826148), ('n03967562', 'plow', 0.2217709), ('n04604
vgg16       : [[('n03538406', 'horse_cart', 0.6929485), ('n03763968', 'military_uniform', 0.045812
vgg19       : [[('n03538406', 'horse_cart', 0.7175846), ('n03763968', 'military_uniform', 0.047984
Resnet50    : [[('n03538406', 'horse_cart', 0.3555997), ('n03763968', 'military_uniform', 0.312221
Mobilenet_V2: [[('n02403003', 'ox', 0.3520962), ('n03538406', 'horse_cart', 0.08434733), ('n038682
Densenet    : [[('n03538406', 'horse_cart', 0.75100106), ('n02403003', 'ox', 0.08875632), ('n03868
NASAnet     : [[('n03967562', 'plow', 0.87051934), ('n03538406', 'horse_cart', 0.09820043), ('n031
```

Karlsruher Institut für Technologie
FIZ Karlsruhe
Leibniz-Institut für Informationsinfrastruktur

Indexing with Pretrained State-of-the-Art Models

(1) Image Captioning (resnet50, 5m iterations)

```
0) a vase filled with flowers sitting on top of a table . (p=0.001639)
1) a vase filled with flowers sitting on a table . (p=0.001241)
2) a vase of flowers sitting on a table . (p=0.001143)
```

a vase filled with flowers sitting on top of a table .

(2) Visual Concept Detection (based on ImageNet 1K)

```
inception_v3: [[('n03991062', 'pot', 0.5503563), ('n01943899', 'conch', 0.061974872), ('n12620546', '
vgg16       : [[('n02840245', 'binder', 0.31984767), ('n07248320', 'book_jacket', 0.23865435), ('n032
vgg19       : [[('n03598930', 'jigsaw_puzzle', 0.39231554), ('n07248320', 'book_jacket', 0.20700723),
Resnet50    : [[('n03598930', 'jigsaw_puzzle', 0.31146914), ('n01943899', 'conch', 0.13275276), ('n04
Mobilenet_V2: [[('n03598930', 'jigsaw_puzzle', 0.66152334), ('n01667778', 'terrapin', 0.018356627), (
Densenet    : [[('n09256479', 'coral_reef', 0.21517044), ('n12985857', 'coral_fungus', 0.1408107), ('n
NASAnet     : [[('n01667778', 'terrapin', 0.37880373), ('n01669191', 'box_turtle', 0.15951602), ('n04
```

KIT Karlsruher Institut für Technologie
FIZ Karlsruhe
Leibniz-Institut für Informationsinfrastruktur

Indexing with Pretrained State-of-the-Art Models

(1) Image Captioning (resnet50, 5m iterations)

```
0) a man sitting on a chair with a laptop . (p=0.000046)
1) a man sitting on a chair with a hat on . (p=0.000026)
2) a man sitting on a bench with a hat on . (p=0.000025)
```

a man sitting on a chair with a laptop .

(2) Visual Concept Detection (based on ImageNet 1K)

```
inception_v3: [[('n04532106', 'vestment', 0.8339432), ('n04429376', 'throne', 0.06276988), ('n04
vgg16       : [[('n03000247', 'chain_mail', 0.5499288), ('n02895154', 'breastplate', 0.08058794)
vgg19       : [[('n03045698', 'cloak', 0.23458833), ('n04532106', 'vestment', 0.16236556), ('n03
Resnet50    : [[('n04532106', 'vestment', 0.3460238), ('n03000247', 'chain_mail', 0.20139052), (
Mobilenet_V2: [[('n04532106', 'vestment', 0.27926844), ('n03000247', 'chain_mail', 0.10626966),
Densenet    : [[('n04532106', 'vestment', 0.5737181), ('n03045698', 'cloak', 0.04356383), ('n044
NASAnet     : [[('n04429376', 'throne', 0.40760532), ('n04532106', 'vestment', 0.2631193), ('n03
```

Karlsruher Institut für Technologie
FIZ Karlsruhe
Leibniz-Institut für Informationsinfrastruktur

Indexing with Pretrained State-of-the-Art Models

(1) Image Captioning (resnet50, 5m iterations)

```
0) a collage of photos of a person holding a skateboard (p=0.000002)
1) a collage of photos with a bunch of different pictures (p=0.000002)
2) a collage of photos of a person holding a skateboard . (p=0.000001)
```

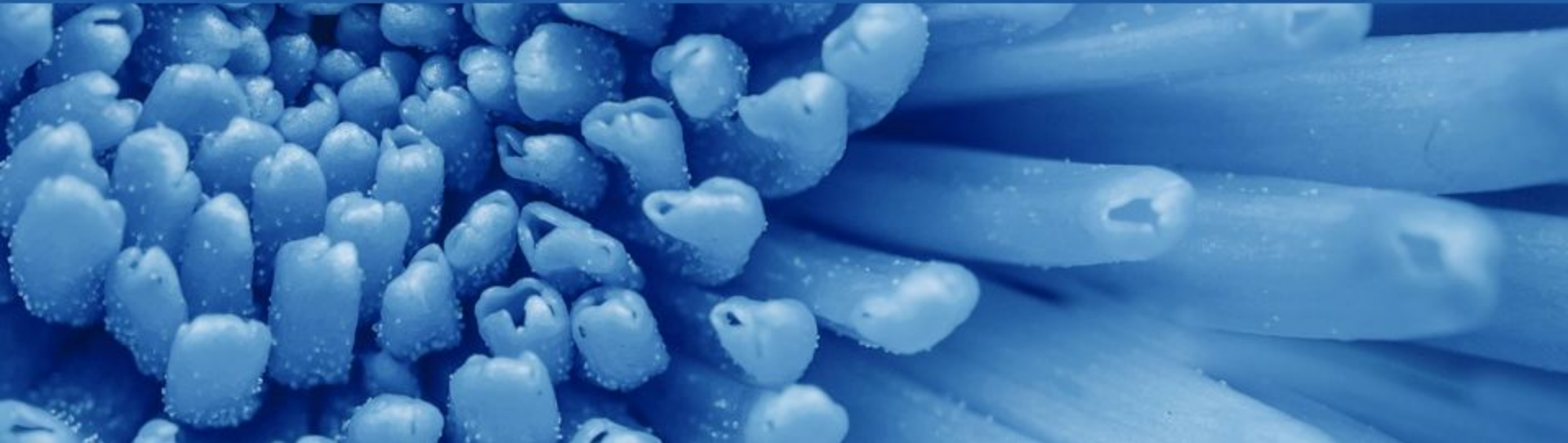
a collage of photos of a person holding a skateboard

(2) Visual Concept Detection (based on ImageNet 1K)

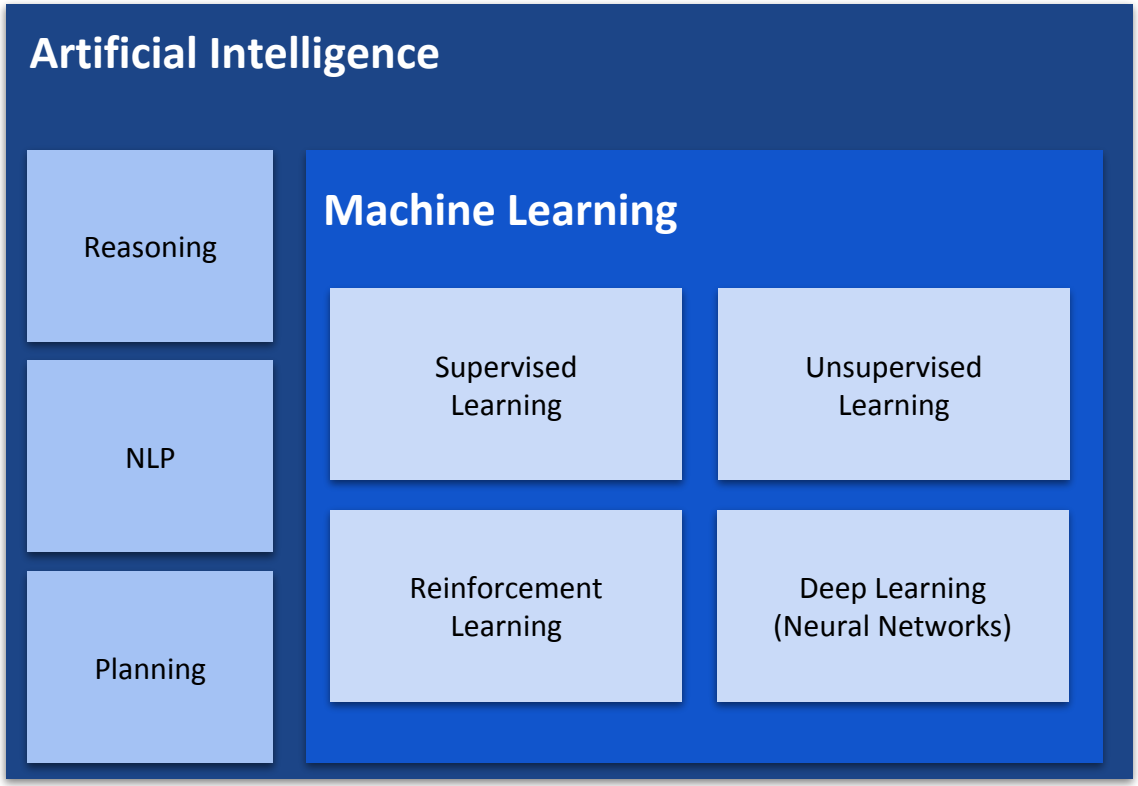
```
inception_v3: [[('n06596364', 'comic_book', 0.29700932), ('n07248320', 'book_jacket', 0.21479161)
vgg16       : [[('n03291819', 'envelope', 0.80103236), ('n07248320', 'book_jacket', 0.12616517),
vgg19       : [[('n03291819', 'envelope', 0.71845376), ('n06596364', 'comic_book', 0.21161233), (
Resnet50    : [[('n03291819', 'envelope', 0.5337895), ('n06596364', 'comic_book', 0.20693506), ('
Mobilenet_V2: [[('n06596364', 'comic_book', 0.3172333), ('n03598930', 'jigsaw_puzzle', 0.16213572
Densenet    : [[('n03291819', 'envelope', 0.2729636), ('n06596364', 'comic_book', 0.19083193), ('
NASAnet     : [[('n03291819', 'envelope', 0.4995414), ('n03485794', 'handkerchief', 0.25641188),
```

“Denn wer den Schatz, das Schöne, heben will,
Bedarf der höchsten Kunst: Magie der Weisen. ”

Johann Wolfgang von Goethe, Faust. Der Tragödie zweyter Theil (1832)



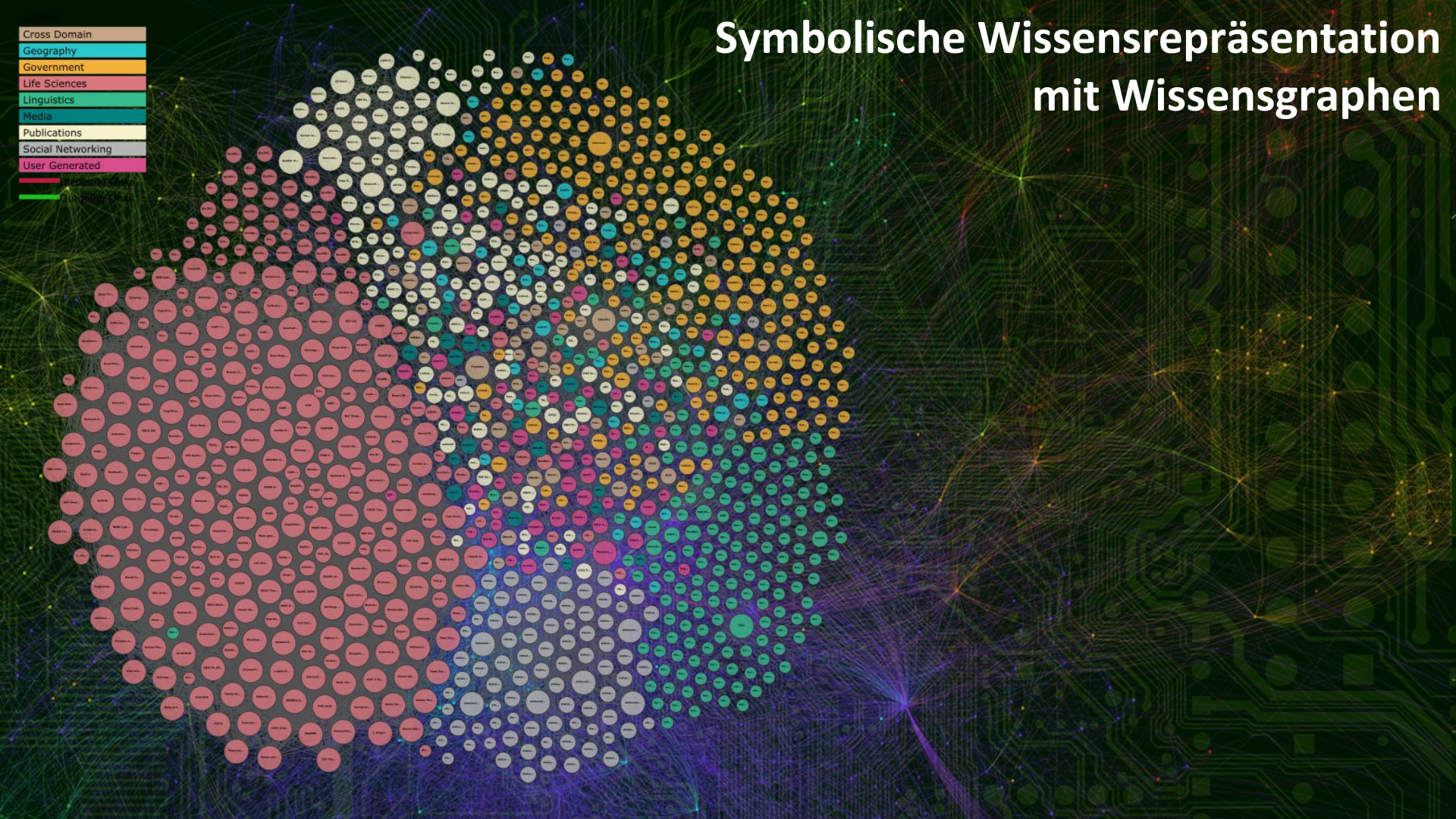
Künstliche Intelligenz ist nicht nur Maschinelles Lernen



"The Goal of AI is to develop machines that behave as though they were intelligent."

- John McCarthy (1955)

Symbolische Wissensrepräsentation mit Wissensgraphen

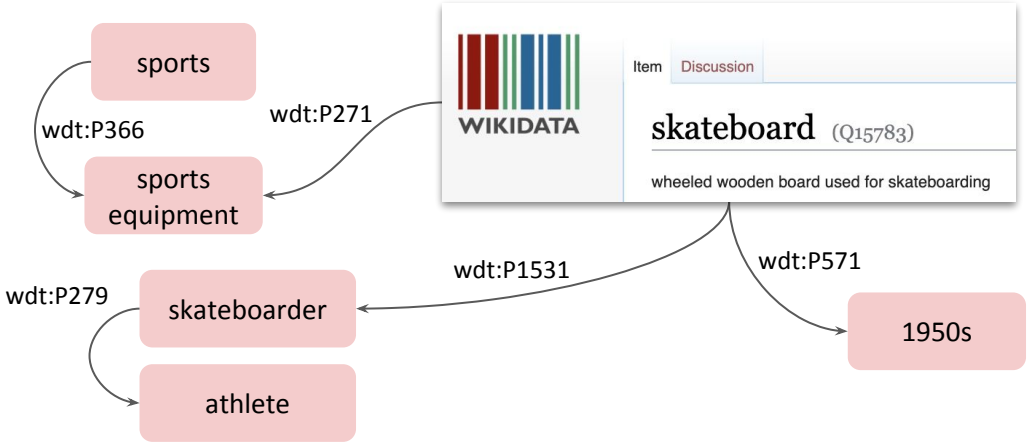


Combining Deep Learning and Semantics

Automated Image captioning:

a **person** holding a **skateboard**

Entity Linking



Metadaten

Titel:	Vita Mathildis
Autor:	Donizio
Datum:	**1115**
Ort	Bibliotheca Apostolica Vaticana...

Combining Deep Learning and Semantics

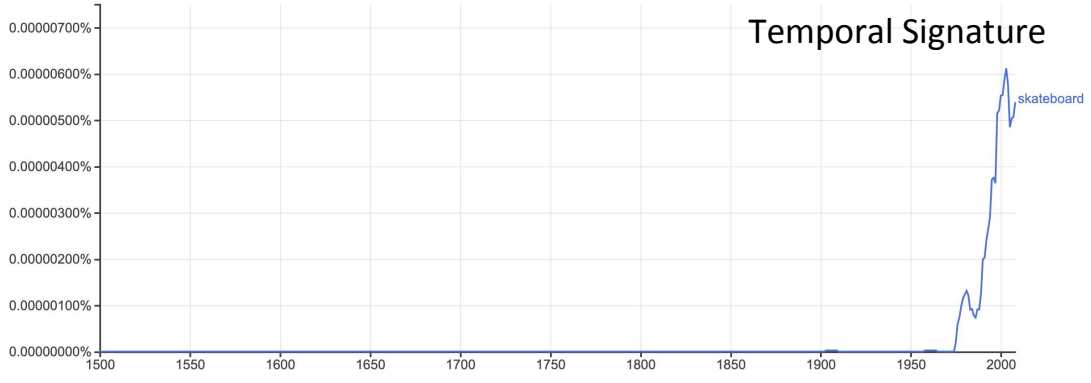
Automated Image captioning:

a **person** holding a **skateboard**

Entity Linking

At two spots in the world a kind of a **skateboard** was developed at **the first** time in **the early** 1950s: California and Hawaii. They used shorter surfboards and wheels made out of metal without some bearings. In the late 1950s, **skateboarding** had **a first** peak.

Temporal Signature



Take Home Messages:

- Deep Learning hat unser Leben verändert (und wird es weiterhin)
- Die Ergebnisqualität hängt stets von den verfügbaren Trainingsdaten ab
- Out-of-the-box Deep Learning Modelle sind einfach zu nutzen und funktionieren erstaunlich gut (...meistens)
- Deep Learning kann von Semantik profitieren

FIZ Karlsruhe
Leibniz-Institut für Informationsinfrastruktur

KIT
Karlsruher Institut für Technologie

Ein Skateboard für den Papst

oder

Warum es maschinelles Lernen ohne Semantik so schwer hat

Prof. Dr. Harald Sack

harald.sack@fiz-karlsruhe.de
twitter: lysander07

Der Kluge Hans Effekt

Oder warum wir Maschinellem Lernen nicht immer trauen sollten

https://commons.wikimedia.org/wiki/File:Osten_und_Hans.jpg