

# VON BIG DATA ZU INTELLIGENTEN SYSTEMEN

Prof. Dr. Stefan Wrobel

Fraunhofer Institut für Intelligente Analyse und Informationssysteme IAIS



# Intelligent Analysis and Information Systems

Doing more with data at Fraunhofer IAIS and Univ. of Bonn

200 people, at the campus Birlinghoven castle close to Bonn and research group at University of Bonn

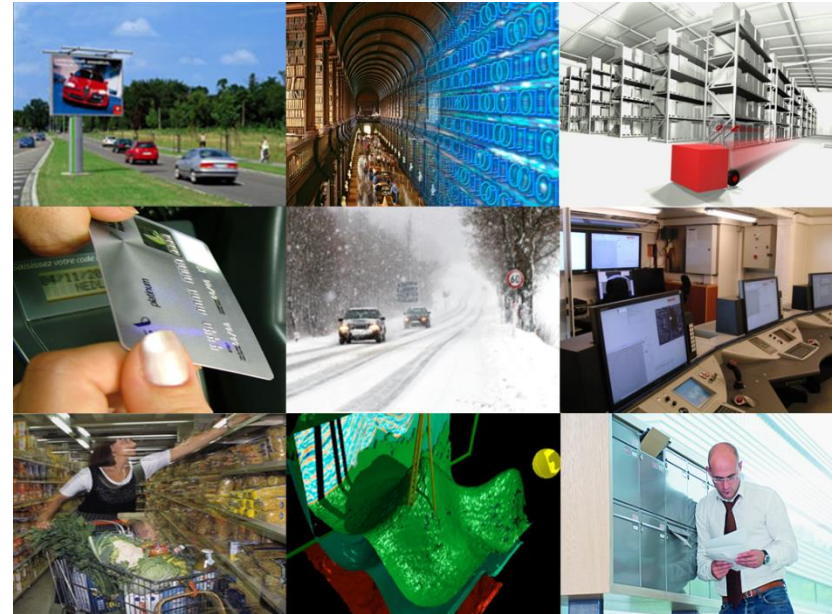
## Research areas

- Data Science and Big Data
- Data Linking, Open Data
- Machine Learning, Deep Learning, Cognitive Computing

Customers across several business sectors

Consulting – Realization – Training

Fraunhofer Alliance Big Data (speaker)



# Fraunhofer Alliance Big Data

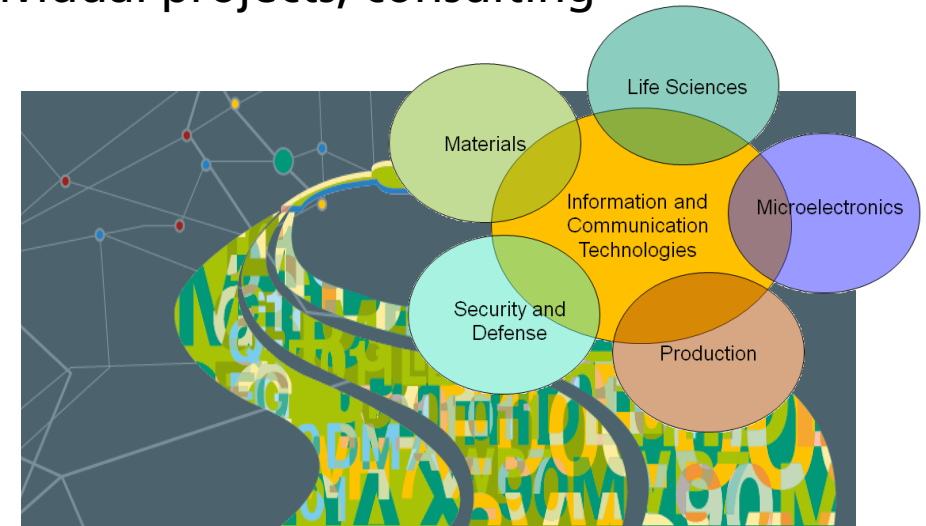
**Joint competences in a »Big Data Factory« for Germany**

**Strategies, Solutions and Successes**

25 Fraunhofer institutes – one central coordination point

Synchronized and broad competence portfolio with many years of expertise in big data in different sectors

Best of class Big Data solutions for individual projects, consulting and qualification of personnel



[bigdata.fraunhofer.de](http://bigdata.fraunhofer.de)

Contact: Prof. Dr. Stefan Wrobel (Chairman)

# A Facebook milestone



Bildzitat: Screenshot aus <https://www.facebook.com/zuck/posts/10102329188394581>, 21.3.2016



Billions of triples/facts already available

Linked Datasets as of August 2014



Bildquelle/Bildzitat: <http://lod-cloud.net/versions/2014-08-30/lod-cloud.svg>, 2014

# Vehicles as sources of data



Location and Behavior Data (Place, Speed, Ac/Deceleration, Accidents)

Sensor Data (Temperature, Rain, Wheel Slip)

Environment Data (Camera, Laser, Radar)

Multimedia System Data

Driver Interaction and Monitoring Data (System Use, Driving Data, Seat Occupation, Seatbelt Use)

Engine and Drive Train Data

Consumption and Refueling

Other Automotive Electronics Data

Car-2-X Data

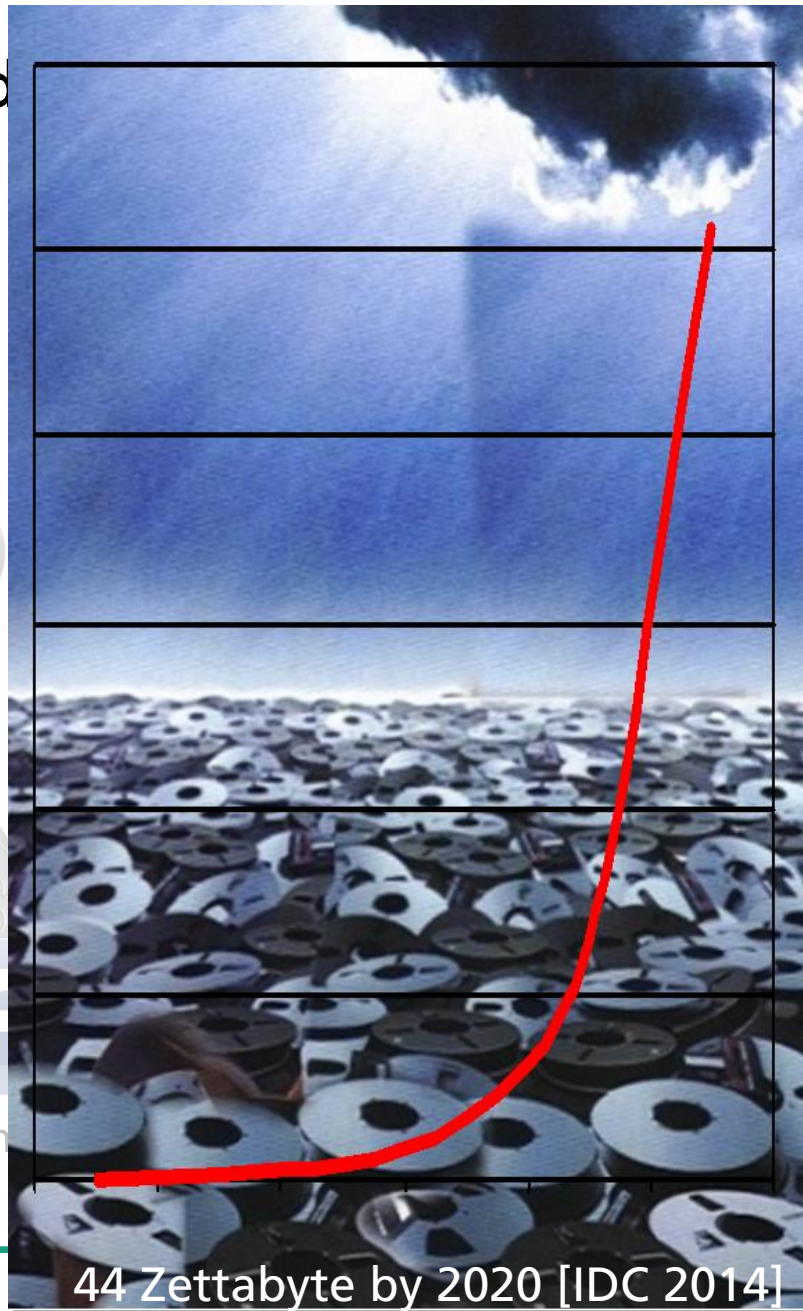
# Digitization and



Convergence



User Content



44 Zettabyte by 2020 [IDC 2014]



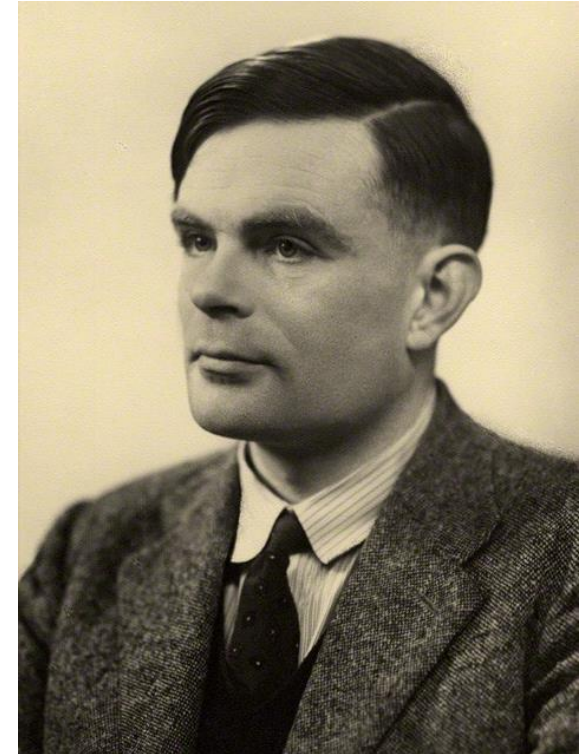
Intelligent Connected  
Systems



Open Data



»»Our problem then is to find out how to programme these machines to [behave intelligently]. At my present rate of working I produce about a thousand digits of programme a day, so that about sixty workers, working steadily through the fifty years might accomplish the job, if nothing went into the waste-paper basket. Some more expeditious method seems desirable.«



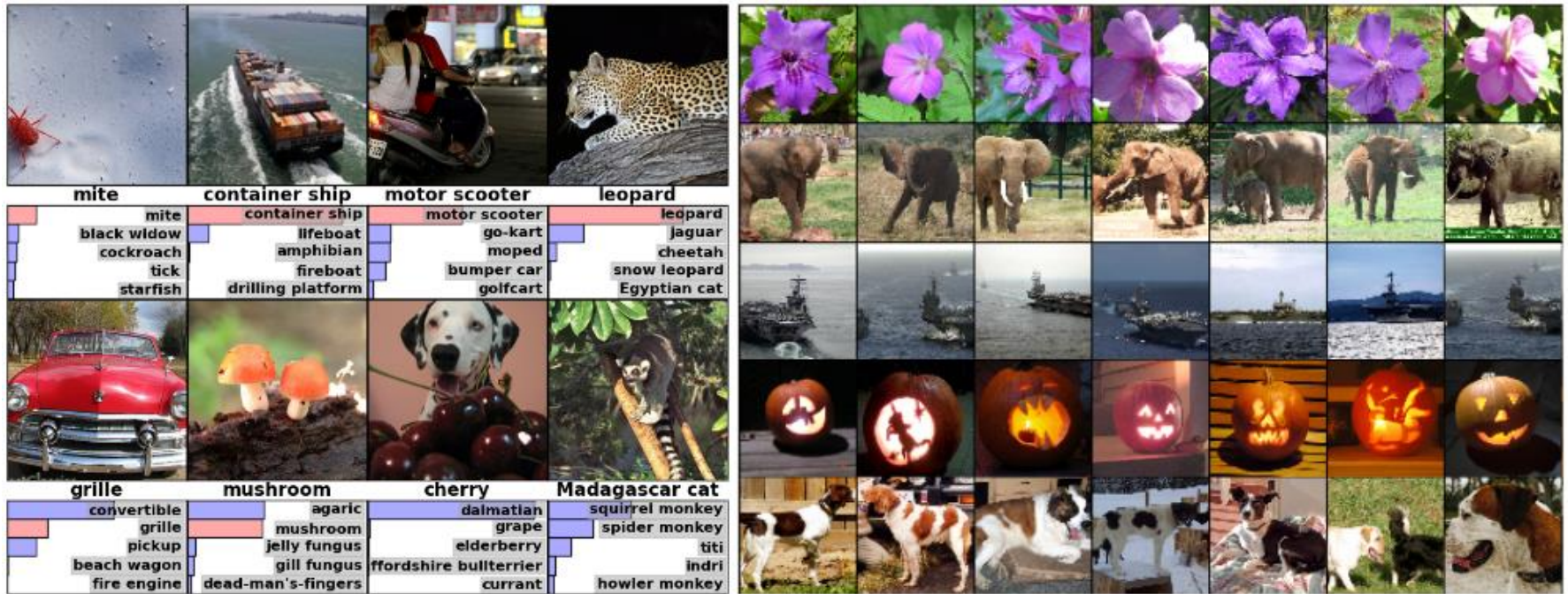
Alan Turing,  
Computing Machinery and Intelligence,  
1950

Bildquelle: National Portrait Gallery, London (cropped), Elliott & Fry, 29 March 1951, NPG x27079,  
<http://www.npg.org.uk/collections/search/use-this-image.php?mkey=mw18791>. CC limited non-comm. license.



# Von Neuronalen Netzen zu »Deep Learning«

## ImageNet: first breakthroughs of DNNs



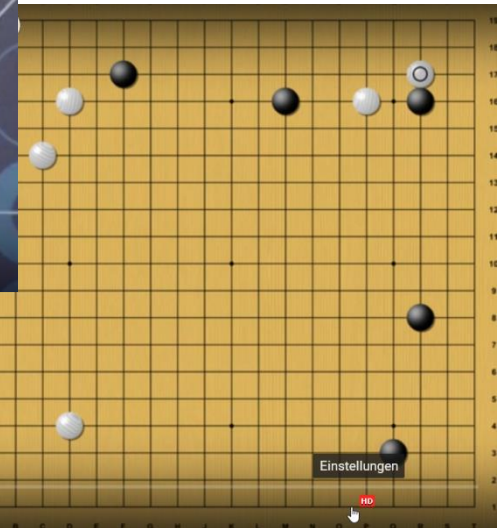
Bildquelle: Krizhevsky, Sutskever, Hinton: ImageNet Classification with Deep Convolutional Neural Networks. Advances in Neural Information Processing Systems 25 (NIPS 2012).

# AI beats humans at the game of Go

„AlphaGo plays like a goddess!“ [Kim Myungwan, Professional Go Player]



Screenshots from  
Gameplay Broadcast by  
Frank Wunderlich-  
Pfeiffer, golem.de



Game played on March 9, 2016,  
Four Seasons Hotel, Seoul

Screenshots from Game Broadcast



Bildquelle Screenshots from <https://youtu.be/6ZugVil2v4w> und [golem.de, http://www.golem.de/news/kuenstliche-intelligenz-alpha-go-spielt-wie-eine-goettin-1603-119646-2.html](http://www.golem.de/news/kuenstliche-intelligenz-alpha-go-spielt-wie-eine-goettin-1603-119646-2.html)



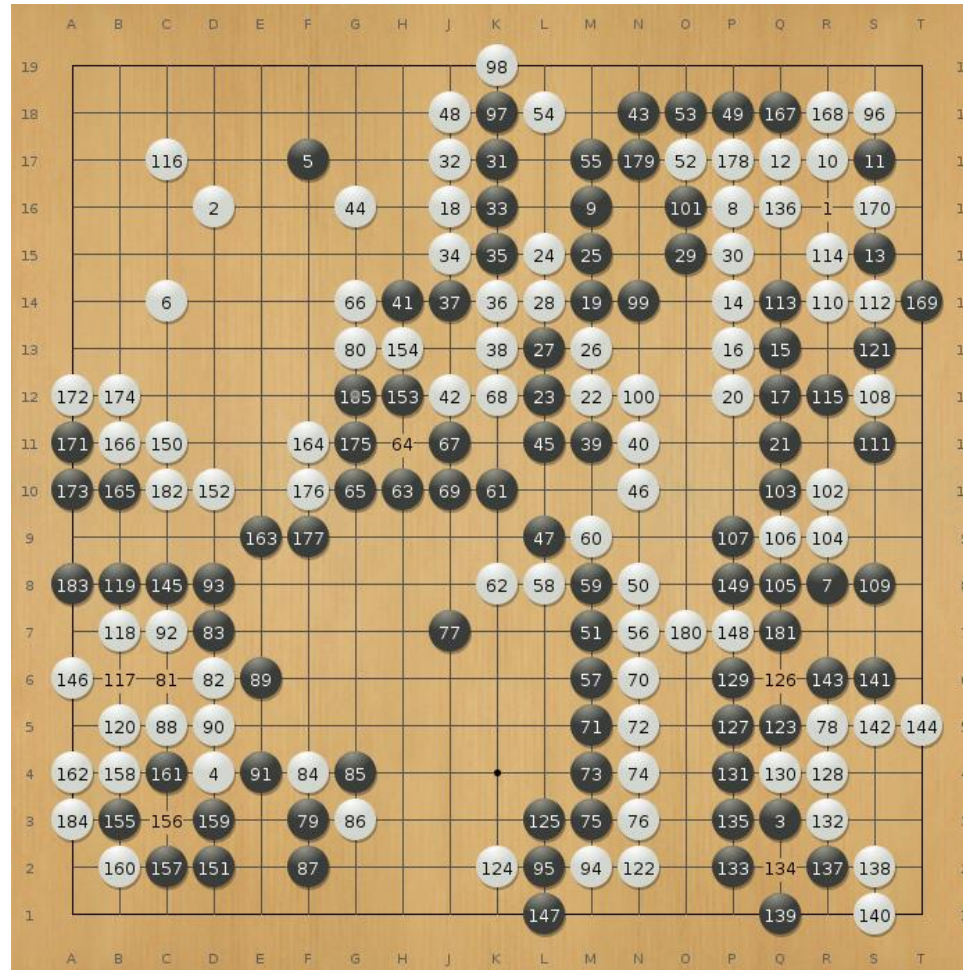
# Lee Sedol (black) vs AlphaGo (white) – 0:1

**Black**

Lee Sedol

World's  
Top-Ranked  
Human  
Player

Game played on  
March 9, 2016,  
Four Seasons  
Hotel, Seoul



Move 184

**White**

AlphaGo

Computer  
Software  
based on  
Machine  
Learning  
[Google Labs]

Screenshots from  
Gameplay  
Broadcast by  
Frank Wunderlich-  
Pfeiffer, golem.de

Bildquelle: golem.de, <http://www.golem.de/news/kuenstliche-intelligenz-alpha-go-spielt-wie-eine-goettin-1603-119646-2.html>



# Fraud recognition in finance

## Machine Learning at Fraunhofer IAIS

### Recognize fraudsters in realtime

Big Data algorithms recognize fraudulent credit card transactions in milliseconds

Fraunhofer software protects credit card companies and their customers

The Software is in day-to-day used at a leading European payment transaction company and protect so portfolio of several million of credit cards



# Recognizing Traffic Signs in Real Time

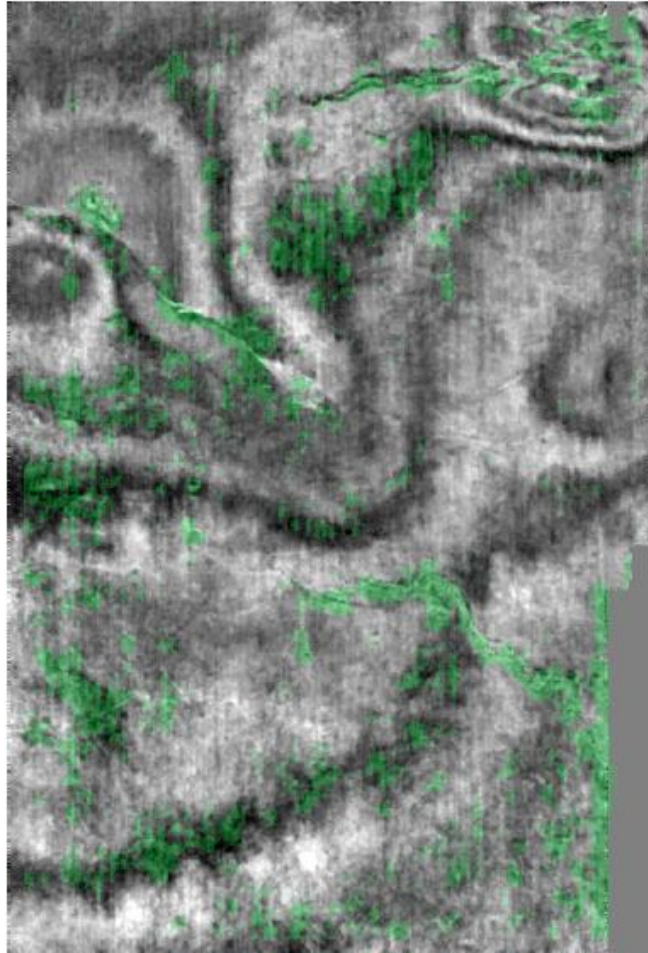
## Machine Learning at Fraunhofer IAIS



Eickeler et al., Proc. AMAA, 2015

# Seismic Data Analysis: Finding Oil and Gas Reservoirs

## Machine Learning at Fraunhofer IAIS



Wulf, Bauckhage,  
Rilling, Bogen, 2016

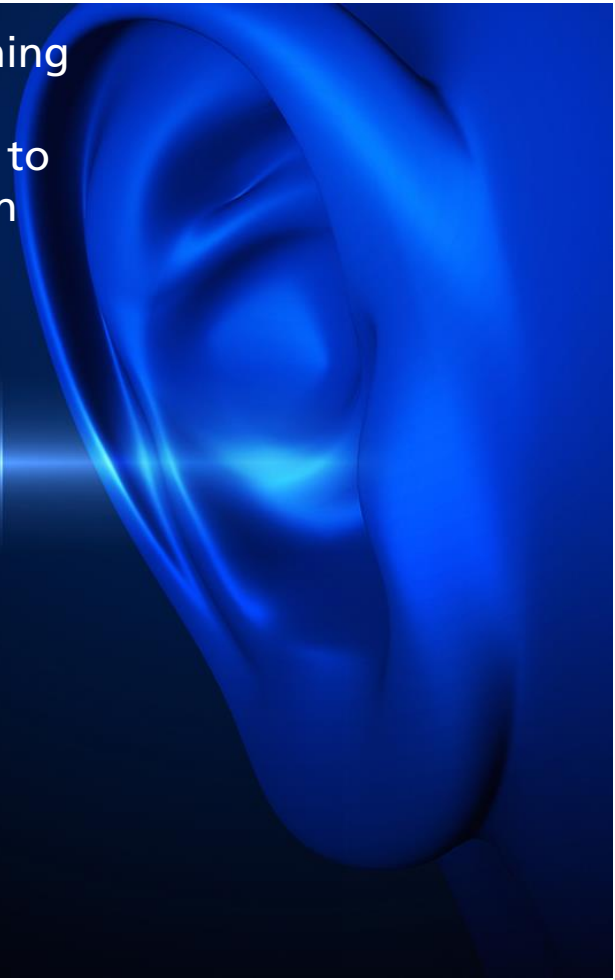


# Speech Recognition with DNNs

## Machine Learning at Fraunhofer IAIS

Error Rate Reduced to less than half using Deep Learning

Efficient training of large data set (recently extended to 1000h annotated material) using graphics acceleration hardware (GPU)



| Acoustic Model | Language Model | Training material [h] | Word error rate [%]<br>Planned speech | Word error rate [%]<br>Spontaneous speech |
|----------------|----------------|-----------------------|---------------------------------------|-------------------------------------------|
| GMM            | 3gram, 200k    | 105                   | 26.4                                  | 33.5                                      |
| GMM            | 3gram, 200k    | 323                   | 24.0                                  | 31.1                                      |
| DNN            | 3gram, 200k    | 323                   | 18.4                                  | 22.6                                      |
| DNN            | 5gram, 510k    | 1005                  | 13.3                                  | 16.5                                      |

# Advances in Machine Learning

New Algorithms, Computing Power and ... Big Data!

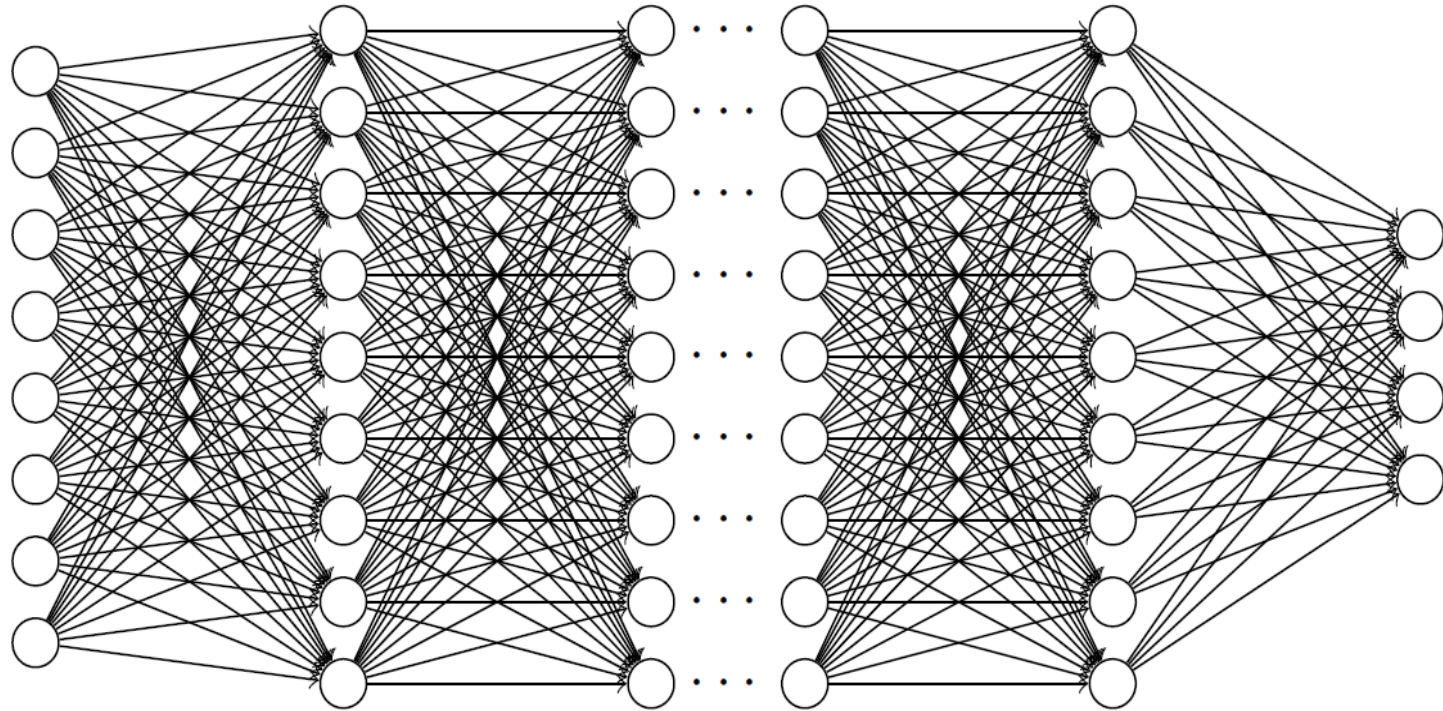


Image: R. Bardeli, D. Stein: Multimedia Analytics Course, 2015

# Maschinelles Lernen: Prädiktion

Das Polynom-Spiel – kann man das gewinnen?

Finde ein Polynom  $f: \mathbb{N} \rightarrow \mathbb{N}$

| x | f(x) |
|---|------|
|---|------|

|   |   |
|---|---|
| 1 | 1 |
|---|---|

|   |   |
|---|---|
| 2 | 4 |
|---|---|

|   |   |
|---|---|
| 3 | 9 |
|---|---|

|   |    |
|---|----|
| 4 | 16 |
|---|----|

|   |    |
|---|----|
| 5 | 25 |
|---|----|

$$f(x) = x^2$$

| x | f(x) |
|---|------|
|---|------|

|   |   |
|---|---|
| 1 | 1 |
|---|---|

|   |   |
|---|---|
| 2 | 1 |
|---|---|

|   |   |
|---|---|
| 3 | 1 |
|---|---|

|   |   |
|---|---|
| 4 | 1 |
|---|---|

|   |     |
|---|-----|
| 5 | 121 |
|---|-----|

$$f(x) = (x-4)(x-3)(x-2)(x-1)x+1$$





# Fraunhofer Compute Center



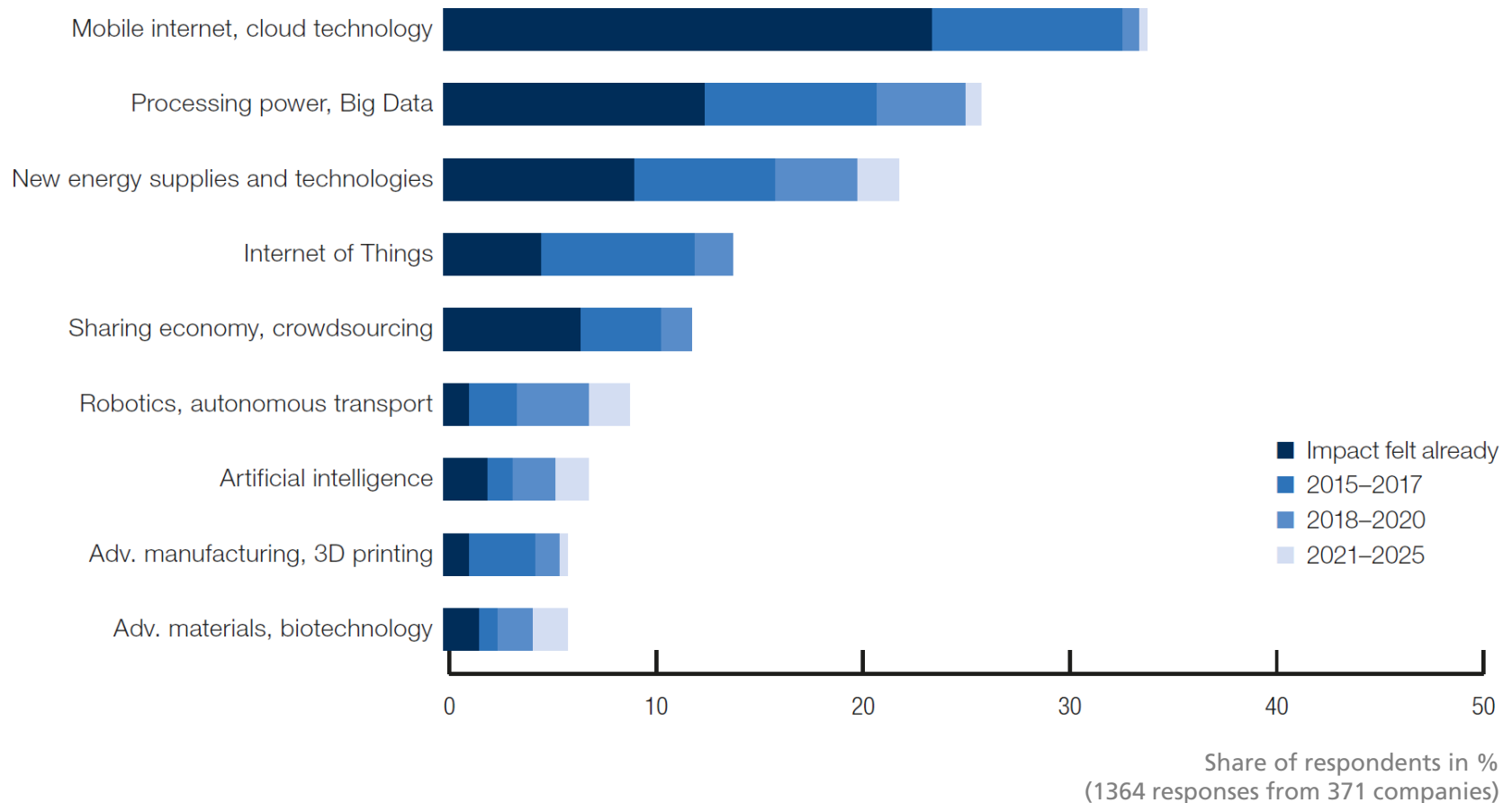
# Big Data

## From Data to Intelligence

- The trend towards availability of ever more detail, complexity, realtimeness, heterogeneity
- The economic potentials that result from digitization through gains in efficiency and disruptive new business models
- The switch from a model-driven to a model- and data-driven approach
- The potential to create artificially intelligent systems with learning algorithms (Analytics, Machine Learning, Artificial Intelligence)
- Big Data causes disruptive and fundamental changes in economy and society

The change is fundamental and distinct from the current technologies like in-memory computing, Hadoop, Storm, ...

# Technology Drivers Impact Companies' Business Models

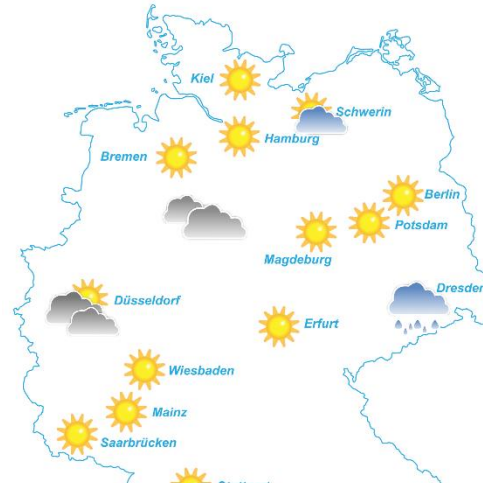


[Future of Jobs Survey, World Economic Forum, 2016]

Bildzitat: World Economic Forum Report „The Future of Jobs“ <http://www.weforum.org/reports/the-future-of-jobs>, 2016

# Smarts and sensors on wheels

## New business models for automobile companies?



**Windshield wipers as micro weather monitor and prediction?**  
**Driving style monitor for insurance pricing?**

Bilder: robsonphoto / kartoxjm / thomaslenne - © fotolia.com

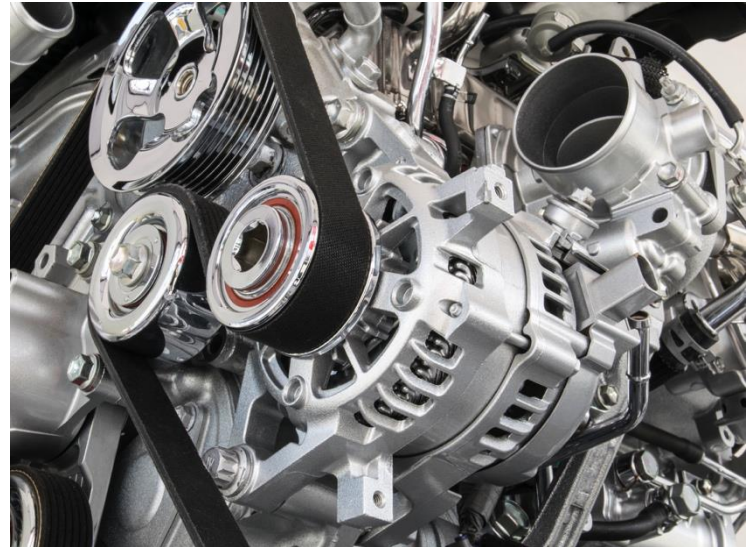


# New Services On The Move

Opportunities for Automobile Companies – or for their digital competitors?

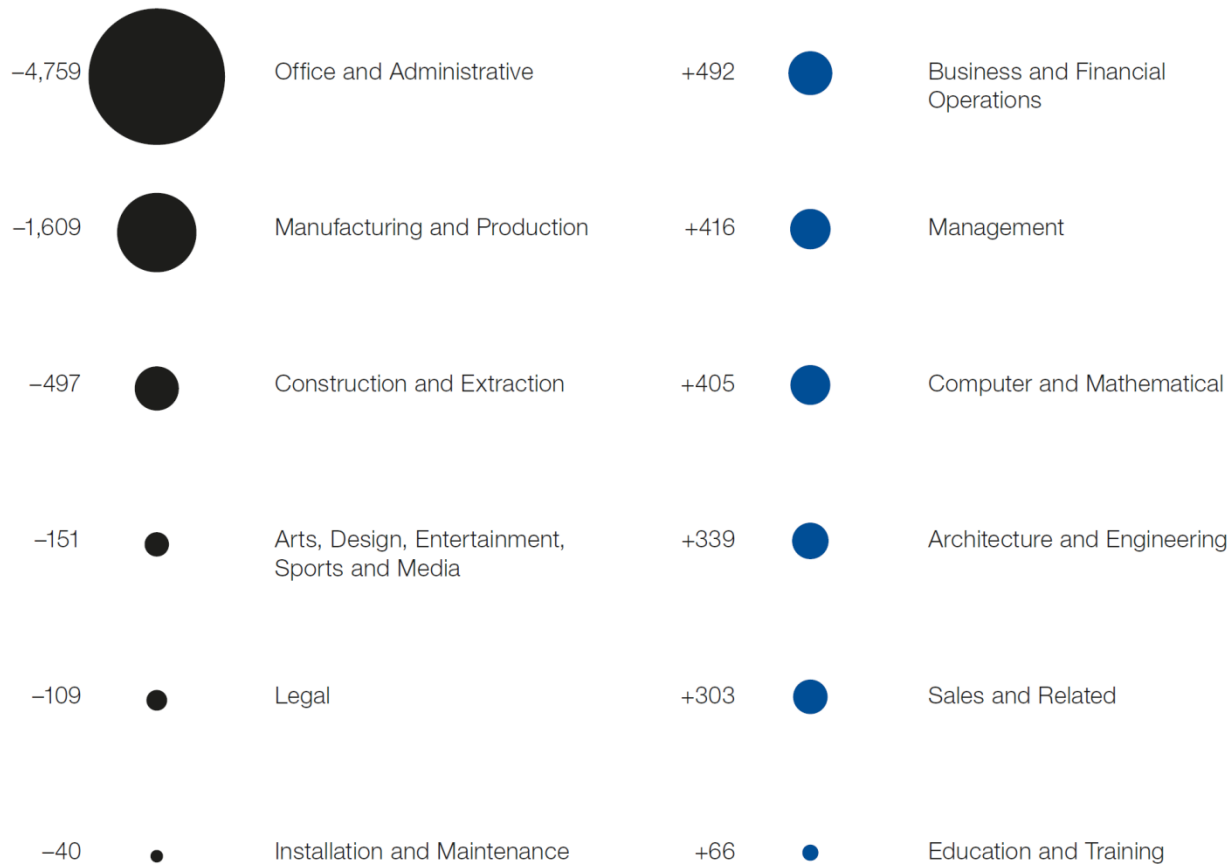


© nito / fotolia.com



© zenstock / fotolia.com

# Employment Outlook by Job Family (2020 vs 2015)



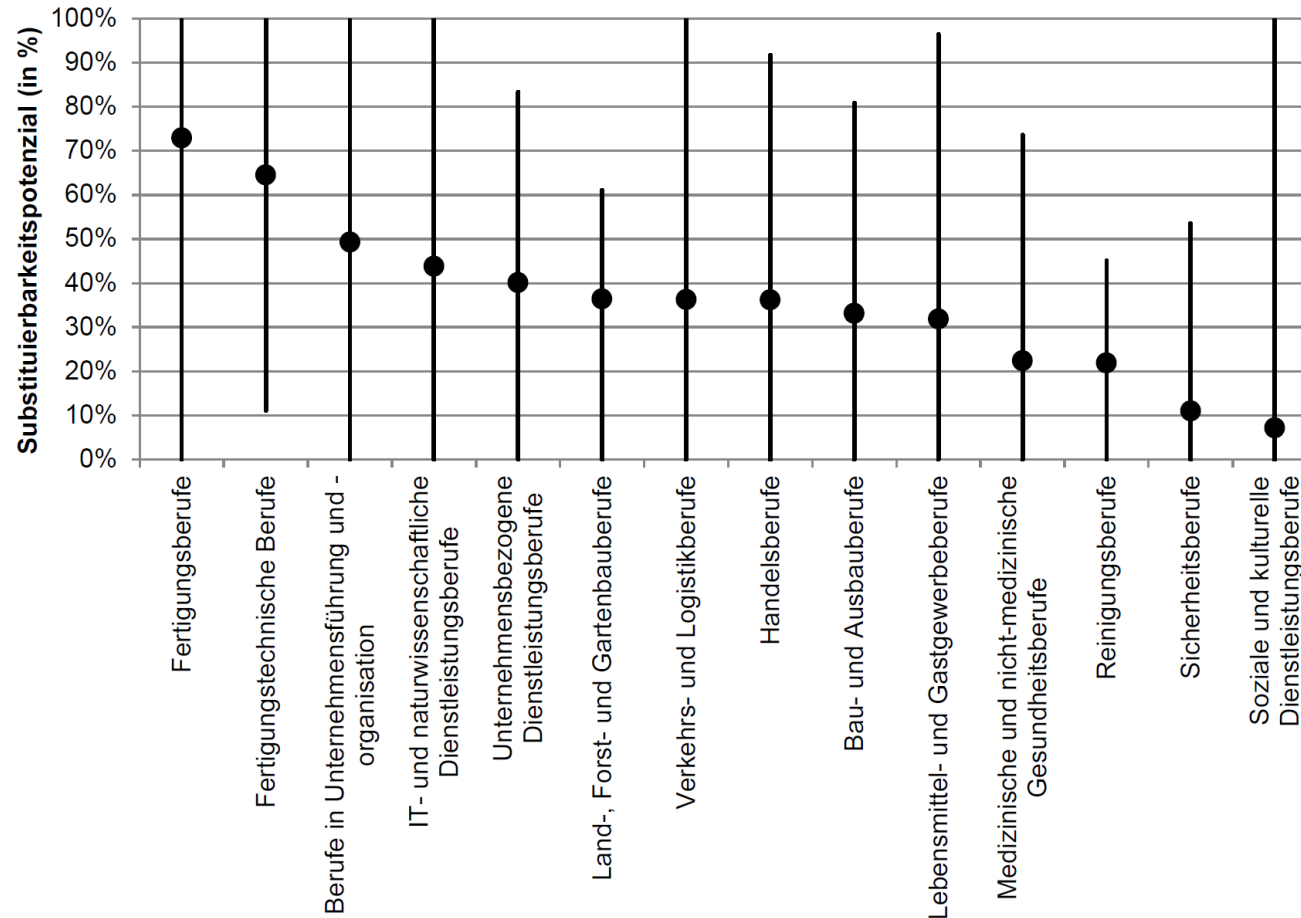
(Employees in Thousands, global, based on 1364 responses from 371 companies)

[Future of Jobs Survey, World Economic Forum, 2016]

Bildzitat: World Economic Forum Report „The Future of Jobs“ <http://www.weforum.org/reports/the-future-of-jobs>, 2016

# Labor Substitution Potentials due to Digitization

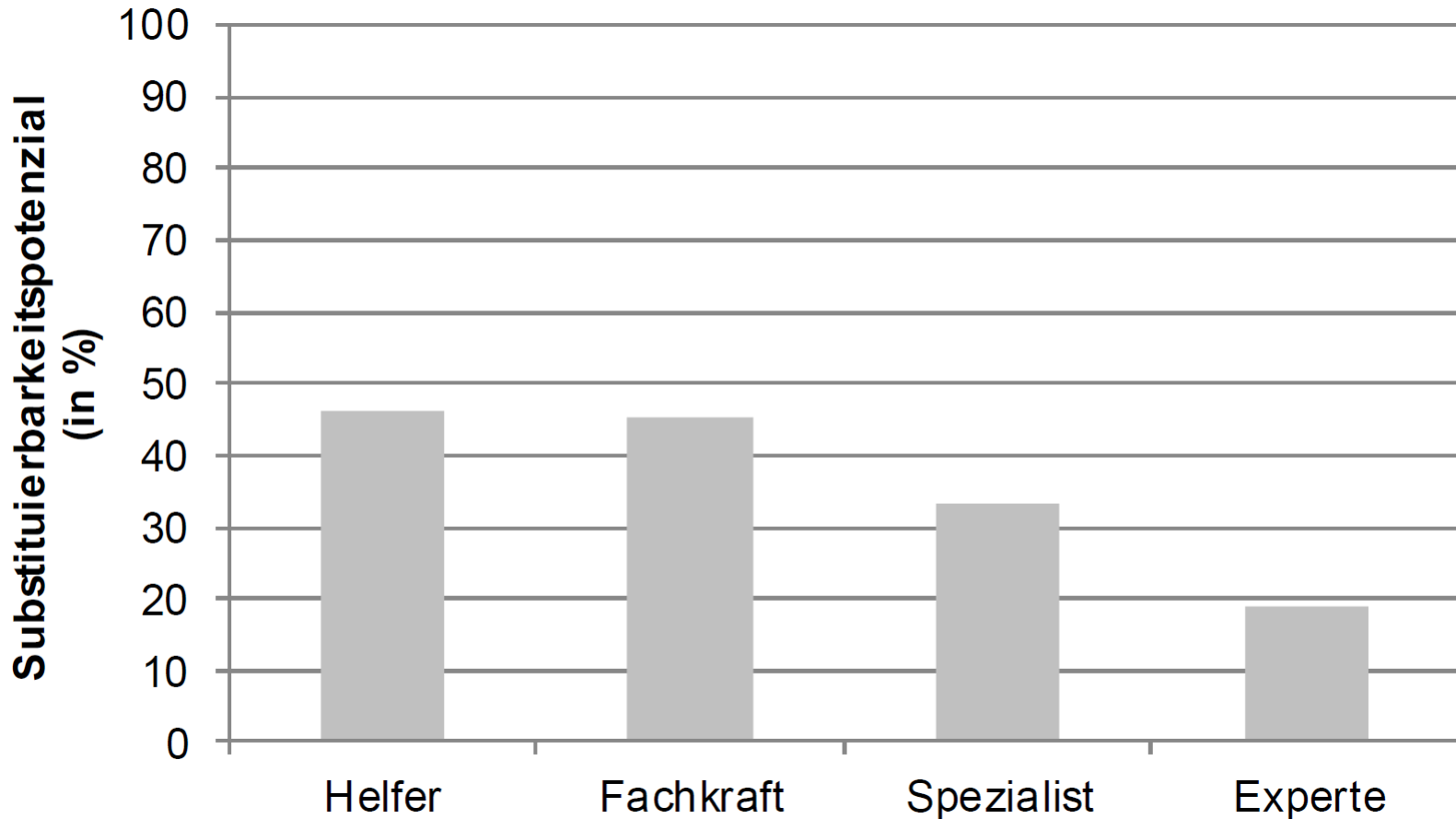
## Germany, Overall View According to Sectors



Quelle: Dengler/Matthes: Folgen der Digitalisierung für die Arbeitswelt. IAB-Forschungsbericht 2015, S. 14

# Labor Substitution Potentials due to Digitization

Germany, Overall View According to Level of Skill and Training

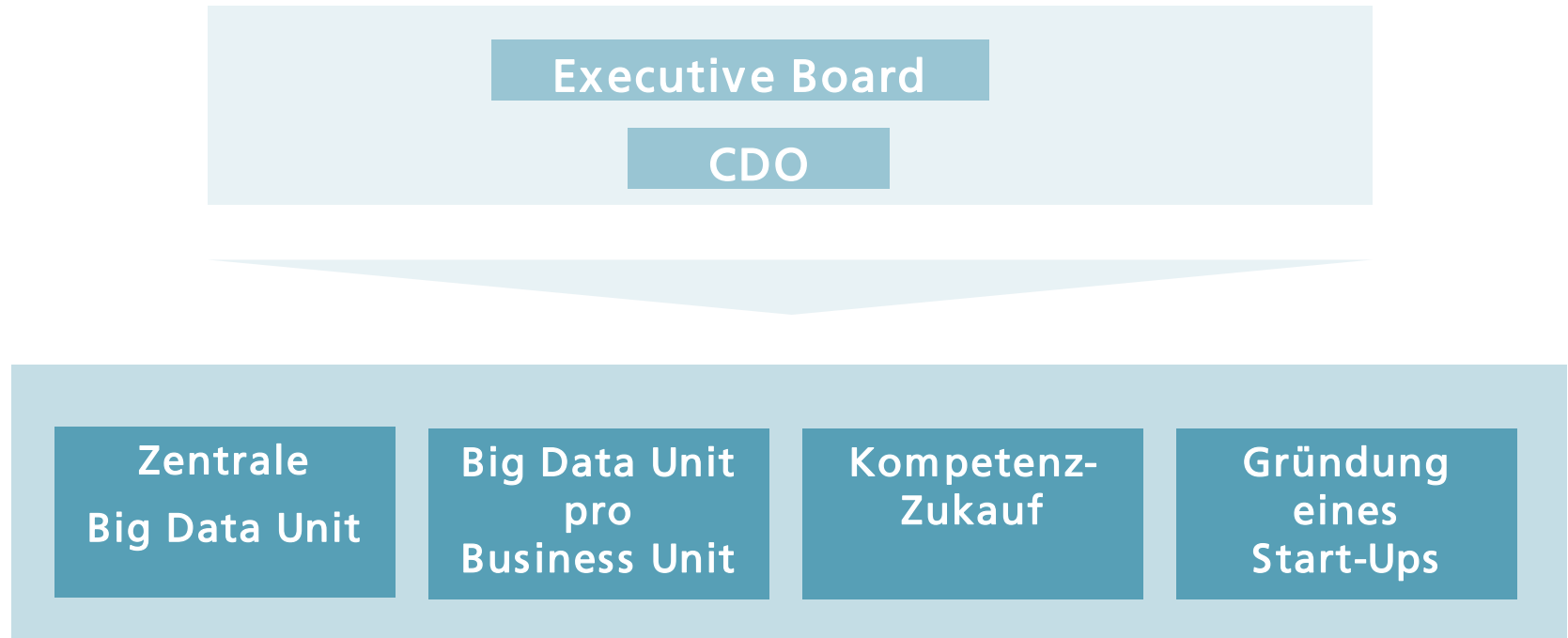


Quelle: Dengler/Matthes: Folgen der Digitalisierung für die Arbeitswelt. IAB-Forschungsbericht 2015, S. 13



# Digital Strategy Development

## Organisatorische Möglichkeiten für die digitale Transformation



# Digitalisierung und Leadership

## **Seminarreihe Data Scientists vermittelt Fraunhofer Know-how**

Erschienen in: [www.bigdata-insider.de](http://www.bigdata-insider.de), 9.4.2015

## **Data Scientist: Big Data verständlich machen**

Erschienen in: [www.news.sap.com](http://www.news.sap.com), 17.2.2015,  
Autor: Andreas Schmitz

## **»Schatzsucher am PC: Data Scientists«**

**Spezialisten suchen in Datenbergen nach Informationen**

Erschienen in: Rhein-Zeitung, 21.6.2014, S. 57.  
Autor: Tobias Hanraths

## **»As Big Data Grows, A New Role Emerges: the Chief Data Officer«**

Erschienen in: [www.theatlantic.com](http://www.theatlantic.com), 17.3.2014,  
Autor: David Lavery

## **»Beherrscher der Datenflut«**

**Spezifische Weiterbildungen machen IT-Spezialisten fit für die neuen Herausforderungen.**

Erschienen in: Stuttgarter Zeitung, 6.9.2014, S. 42.

# Digitalisierung und Leadership

## Data Scientist: Big Data verständlich machen

Erschienen in: [www.news.sap.com](http://www.news.sap.com), 17.2.2015,  
Autor: Andreas Schmitz

## Seminarreihe Data Scientists vermittelt

**Fraunhofer** ■ Wer trägt im Unternehmen die übergeordnete Verantwortung für Digitalisierung und Datenwertschöpfung?

Erschienen in: [www.bigdata-insider.de](http://www.bigdata-insider.de), 3.4.2015

■ Wer setzt im Unternehmen Projekte im Bereich Big Data um?

» Schatzsucher am PC:

## Data Scientists«

Spezialisten suchen in Datenbergen nach Informationen

Erschienen in: Rhein-Zeitung, 21.6.2014, S. 57.  
Autor: Tobias Hanraths

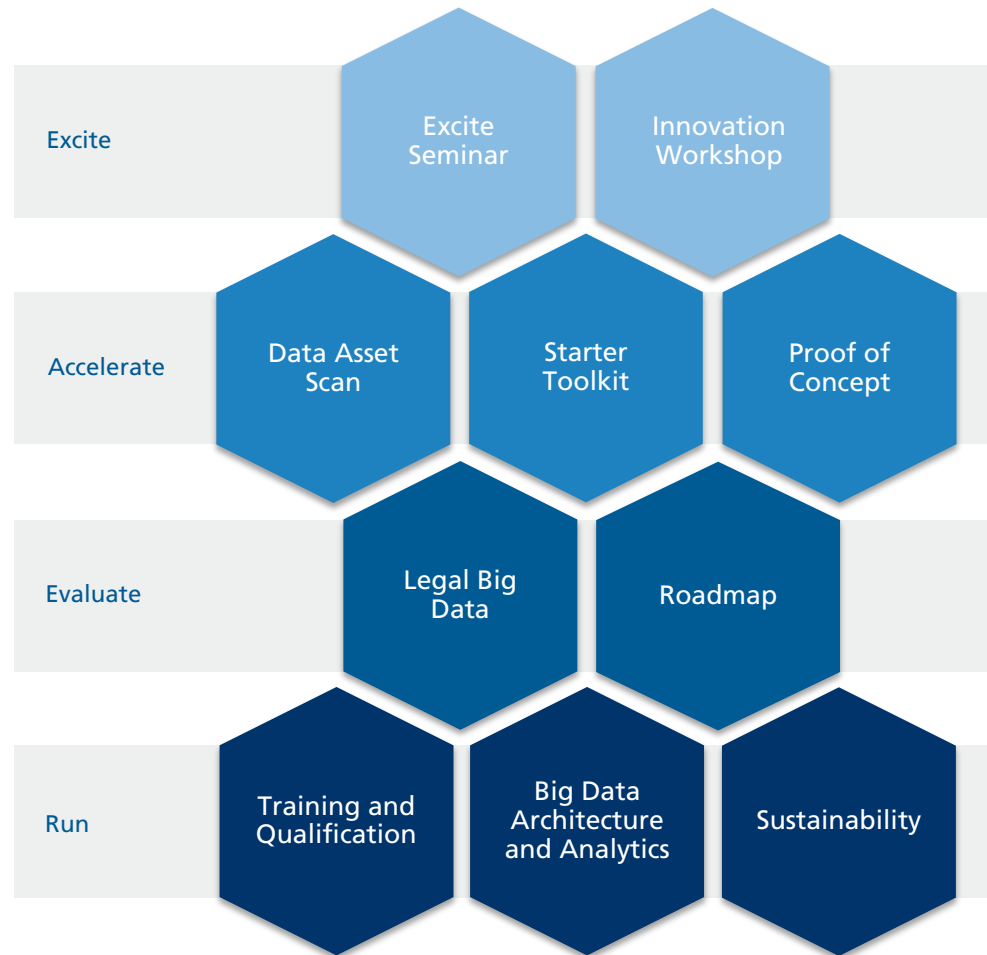
Erschienen in: [www.theatlantic.com](http://www.theatlantic.com), 17.3.2014,  
Autor: David Lavery

## »Beherrscher der Datenflut«

Spezifische Weiterbildungen machen IT-Spezialisten fit für die neuen Herausforderungen.

Erschienen in: Stuttgarter Zeitung, 6.9.2014, S. 42.

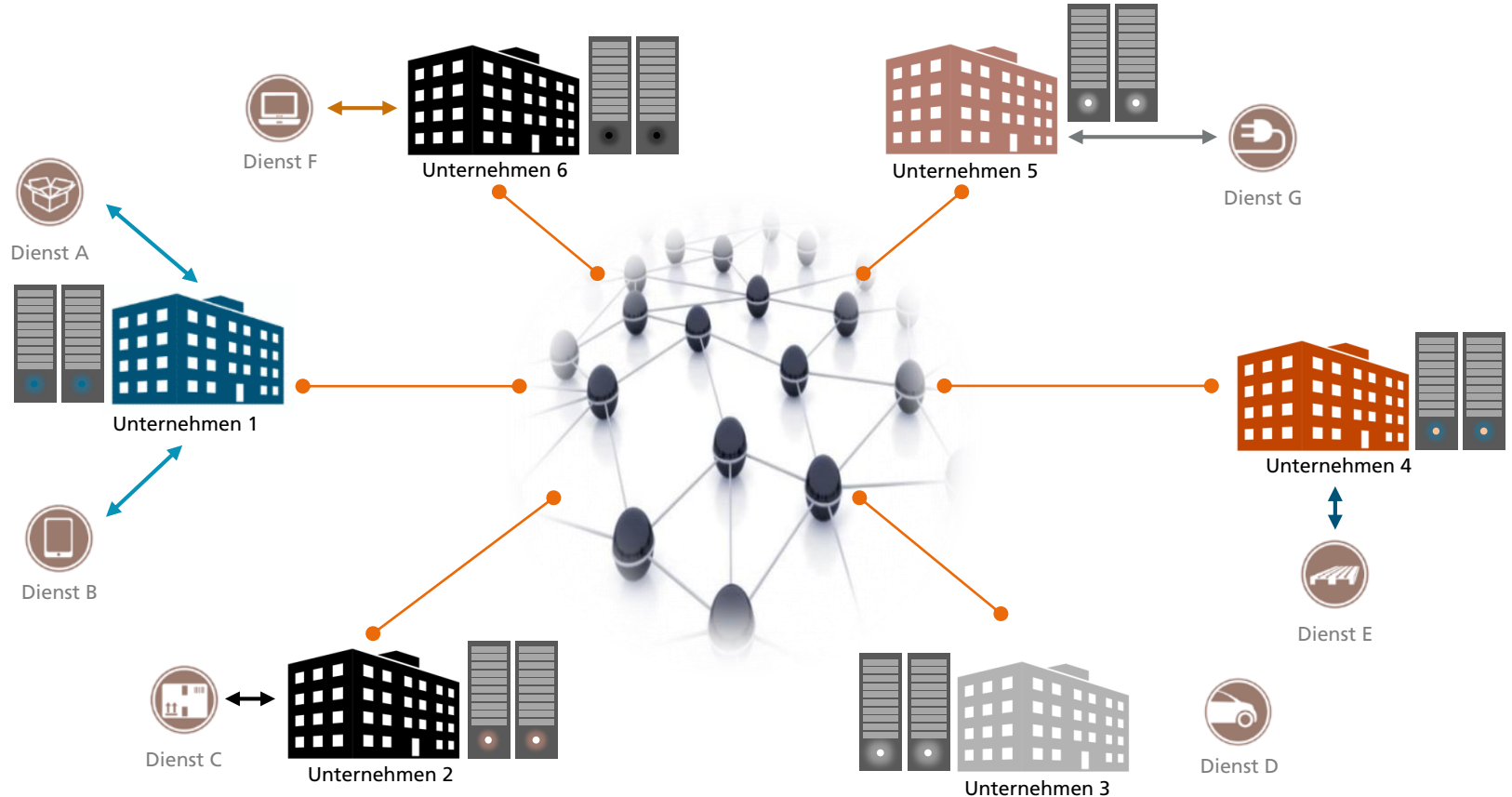
# Data-Driven Company powered by Fraunhofer





# Smart and Linked Data Ecosystems

## Key Driver and Success Factor for New Business



# Der Industrial Data Space ermöglicht ein »Network of Trusted Data«

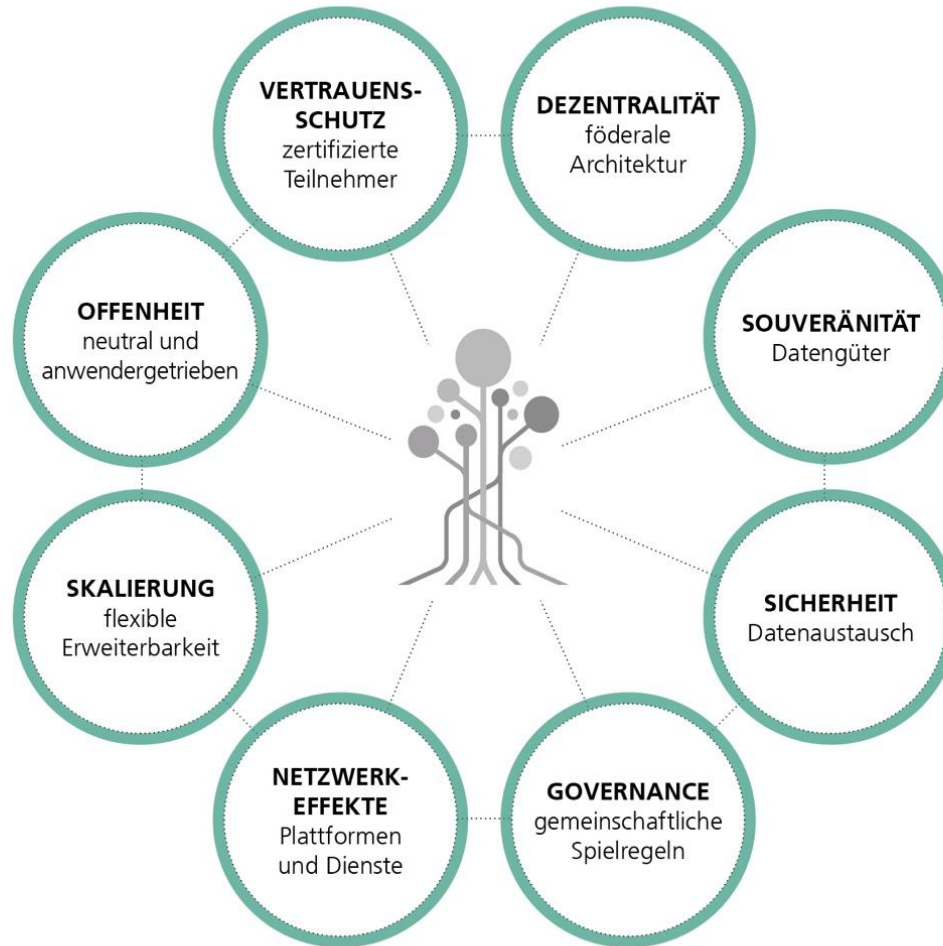
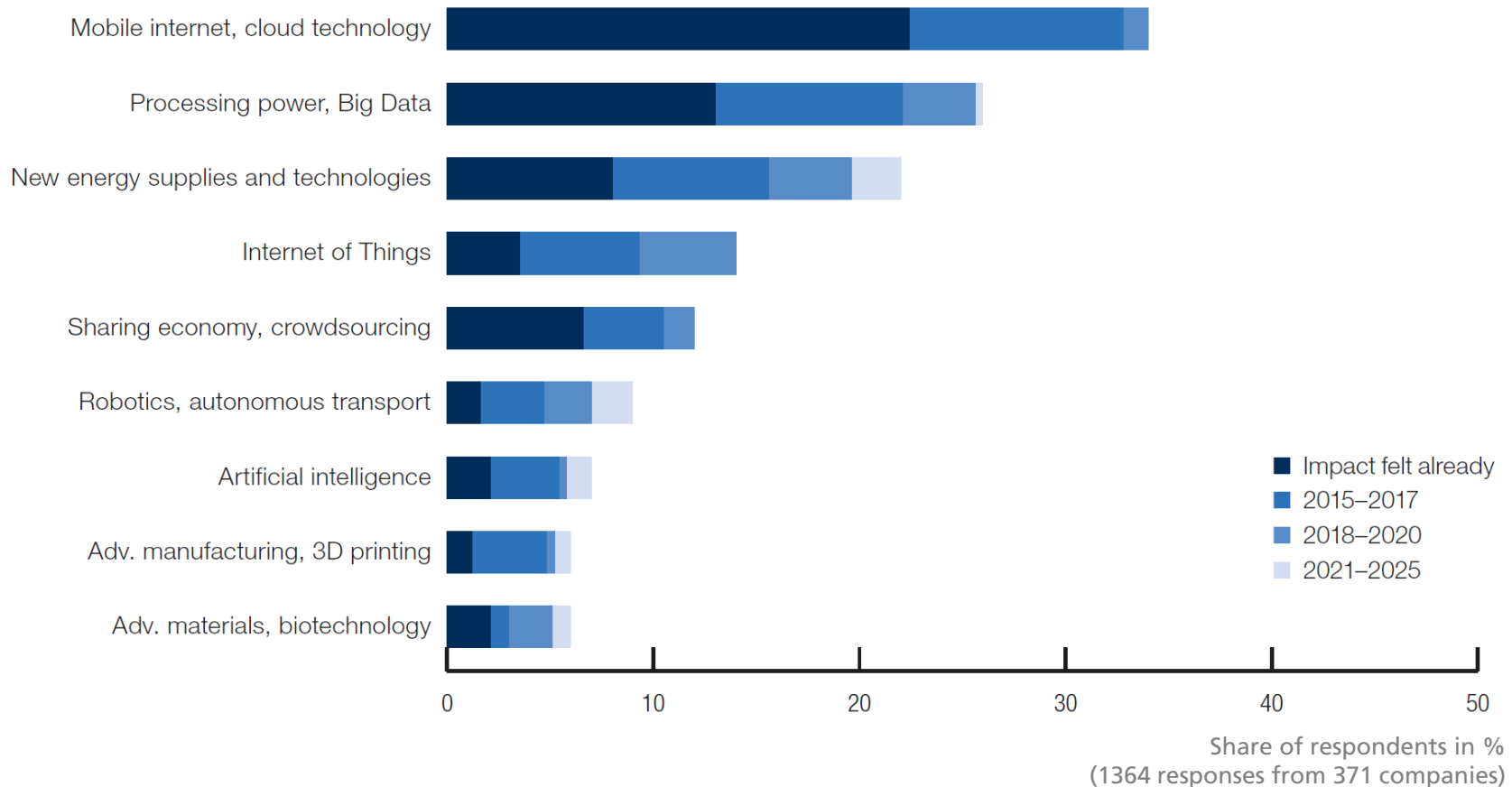


Bild: © Fraunhofer-Gesellschaft

# Technology Drivers Impact Employee Skills

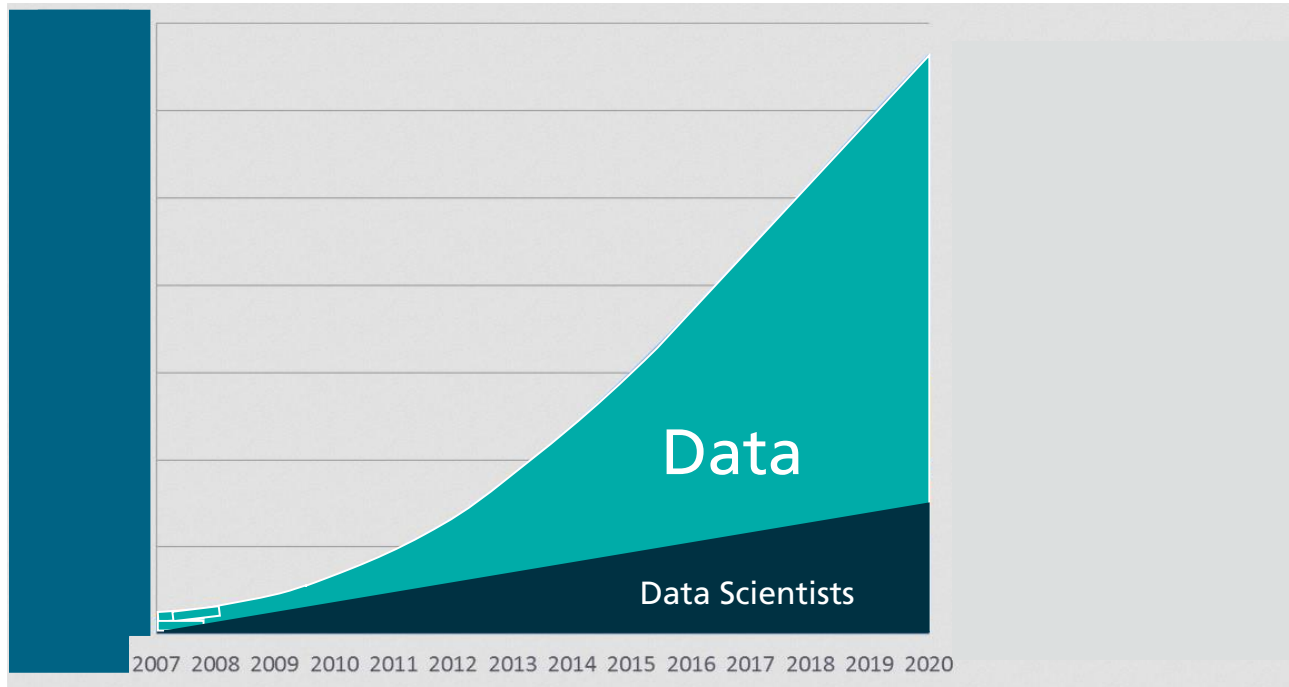


[Future of Jobs Survey, World Economic Forum, 2016]

Bildzitat: World Economic Forum Report „The Future of Jobs“ <http://www.weforum.org/reports/the-future-of-jobs>, 2016

# Die Qualifizierungslücke

## Der Data Scientist – Ein neues Berufsbild in der IT-Branche



Data Science ist ein junges Feld mit **großem Fortbildungsbedarf**: In Deutschland fehlen 43.000 IT-Experten, davon 45% mit Big-Data-Kompetenzen

Quelle: BITKOM, 2015

# Fraunhofer-Schulungsprogramm Big Data

## Zertifizierte Weiterbildung für Data Scientists



### Data Scientist Ausbildung in der Fraunhofer Gesellschaft



#### Executive Seminar

Schaffung eines  
gemeinsamen  
Vokabulars,  
Grenzen und  
Möglichkeiten des  
Trends



#### Methoden- spezifische Seminare

Big Data Architektur,  
Big Data Analytics,  
Visual Analytics,  
etc.



#### Sektor-spezifische Seminare

Social Media  
Monitoring für die  
Marktforschung,  
Smart Energy



#### Unternehmens- spezifische Curricula

Aufbau von  
individuellen  
Unternehmens-  
ausbildungen

### Fraunhofer Zertifizierung zum Data Scientist



# Big Data Business Club

Das erste deutsche Netzwerk für CDOs



Zentrale Anlaufstelle und **exklusive Vernetzungsplattform** zum Thema Big Data/Digitalisierung auf Führungsebene

Förderung des branchenübergreifenden **Technologie- und Wissenstransfers** im Bereich digitale Transformation

Strategische Vernetzung der **zentralen Persönlichkeiten** im Big-Data-Umfeld – Wissenstransfer zwischen Wirtschaft und Wissenschaft

**Schirmherr:** Prof. Dr. Reimund Neugebauer,  
Präsident der Fraunhofer-Gesellschaft



Austausch



Beratung



Innovationskraft



Kooperation



Beratung

# Roberta: Eine Generation von Gestaltern schaffen!



Gestalten lernen mit Open Roberta: vor Ort, in der Cloud, leicht zugänglich, von Lego-Robotern bis zum Smart Home



## Roberta 800.000

Unser Ziel – Machen Sie mit!

Jedes Kind eines Jahrgangs kann mit Roberta zum Gestalter werden

- Eigenes Cloud-Angebot, Programmieren in NEPO®



# Disclaimer

Copyright © der

Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V.  
Hansastraße 27c  
80686 München.

Alle Rechte vorbehalten.

Zuständiger Ansprechpartner: Silke Loh, Presse und Öffentlichkeitsarbeit

E-Mail Ansprechpartner: [silke.loh@iais.fraunhofer.de](mailto:silke.loh@iais.fraunhofer.de)

Die Urheberrechte dieser Präsentation und ihre Inhalte liegen vollständig bei der Fraunhofer-Gesellschaft, sofern nicht ausdrücklich anders gekennzeichnet. Ein Download oder Ausdruck der Inhalte dieser Präsentation ist ausschließlich für den persönlichen Gebrauch gestattet. Alle darüber hinaus gehenden Verwendungen, insbesondere die kommerzielle Nutzung und Verbreitung, sind grundsätzlich nicht gestattet und bedürfen der vorherigen schriftlichen Zustimmung der Fraunhofer-Gesellschaft.

Grafische Veränderungen an Bildmotiven oder die Bearbeitung von Texten sind nicht gestattet.

Abweichend von der vorgenannten Nutzungsbeschränkung ist ein Ausdruck oder eine anderweitige Vervielfältigung der Inhalte dieser Präsentation darüber hinaus ausschließlich zum Zweck der Berichterstattung über die Fraunhofer-Gesellschaft und ihrer Institute unter der Voraussetzung honorarfrei gestattet, dass stets eine vollständige Quellenangabe erfolgt und zwei kostenlose Belegexemplare an die oben genannte Adresse geschickt werden.

# Haftungshinweis

Die Fraunhofer-Gesellschaft ist bemüht, ihre Präsentationen stets aktuell und inhaltlich richtig sowie vollständig anzubieten. Dennoch ist das Auftreten von Fehlern nicht völlig auszuschließen. Die Fraunhofer-Gesellschaft übernimmt keine Haftung für die Aktualität, die inhaltliche Richtigkeit sowie für die Vollständigkeit der in ihren Präsentationen eingestellten Informationen und Inhalte. Der vorgenannte Haftungsausschluss bezieht sich auch auf eventuelle Schäden, auch von Dritten, die durch die Nutzung dieser Präsentationen verursacht wurden.

Geschützte Marken, Namen, Bilder und Texte werden in den Präsentation der Fraunhofer-Gesellschaft in der Regel nicht als solche kenntlich gemacht. Das Fehlen einer solchen Kennzeichnung bedeutet jedoch nicht, dass es sich um frei verwendbare Namen, Bilder oder Texte im Sinne des Marken- oder Urheberrechts handelt.