



# **BIG DATA, ODER: NEUGIERDE ALS WETTBEWERBSVORTEIL**

DR. BENEDIKT KOEHLER  
GRÜNDER & CEO, DATALION GMBH

3. BIG DATA KONFERENZ,  
STARTPLATZ, KÖLN  
9. JUNI 2016







# These 1: Es gibt kein Big Data



## DER MYTHOS DER GROSSEN DATENMENGEN

# The End of Science

The quest for knowledge used to begin with grand theories. Now it begins with massive amounts of data. Welcome to the Petabyte Age.

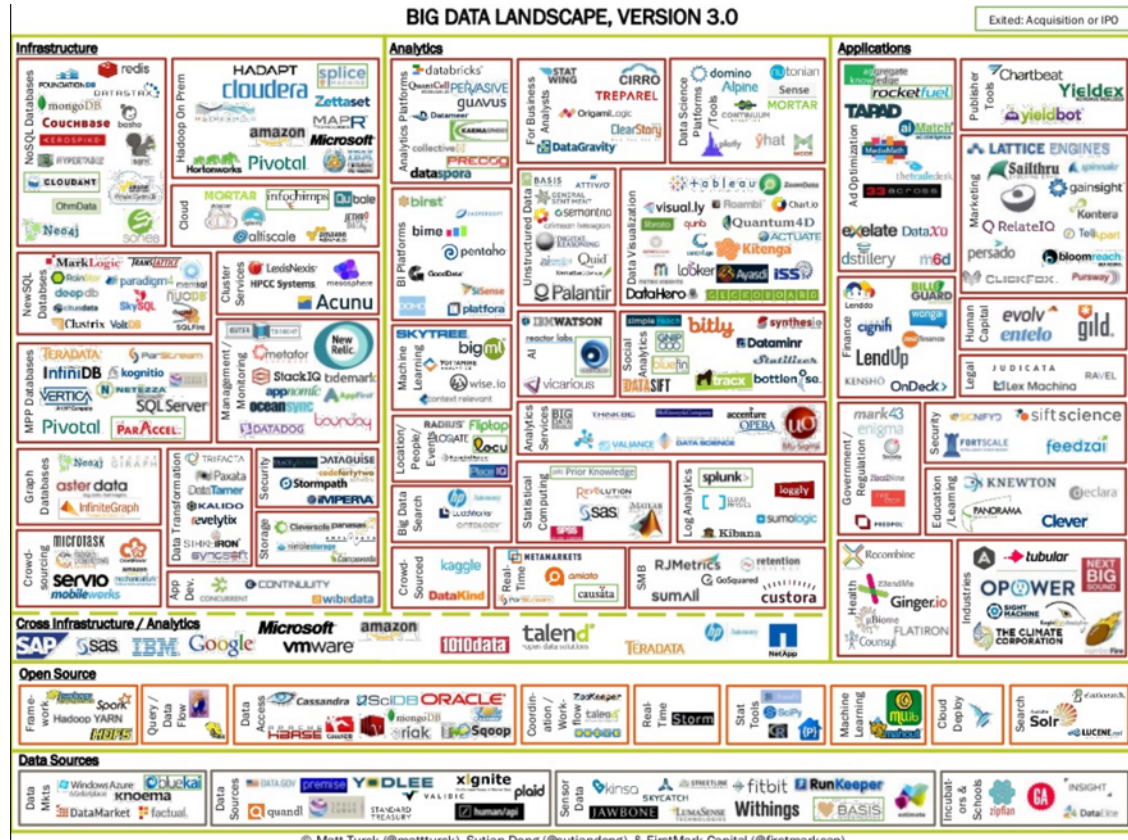


Quelle: Wired Magazine, 16.07





# DIE BIG DATA LANDSCHAFT



© Matt Turck (@mattturck), Sutan Dong (@sutandong) & FirstMark Capital (@firstmarkcap)

Quelle: Big Data Landscape by Matt Turck, <http://mattturck.com/2016/02/01/big-data-landscape/>





# BIG DATA IST RELATIV

65526	1360					
65527	1361		1			
65528	1362	0	1	1	2	
65529	1363	0	1	1	1	
65530	1364	0	0	1	3	
65531	1365	0	1	2	1	
65532	1366	0	1	1	1	
65533	1367	0	0	1	4	
65534	1368	1	1	1	1	
65535	1369	1	1	3	2	
65536	1370	1	0	1	1	
65537						
65538						
65539						
65540						





# ES GIBT KEIN BIG DATA, SONDERN NUR DATA



**Thomas French**  
@trfrench



 **Folgen**

Joke at Google: "it's not big data, it's just data."  
[#StrataHadoop](#) [#BigData](#) [#infiniteresources](#)

 Übersetzung anzeigen

GEFÄLLT

4



10:52 - 3. Juni 2016







These 2:  
Jeder kann, was  
Google kann





# CLARKES DRITTES GESETZ

Google!

Search the web using Google!

10 results



Google Search

I'm feeling lucky

*Index contains ~25 million pages (soon to be much bigger)*

**[About Google!](#)**

[Stanford Search](#) [Linux Search](#)

Get Google! updates monthly!

your e-mail

Subscribe

[Archive](#)

Copyright ©1997-8 Stanford University



# MILITÄR – GROSSUNTERNEHMEN – WISSENSCHAFT



Quelle: Business Insider, <http://www.businessinsider.com/picture-of-ibm-hard-drive-on-airplane-2014-1?IR=T>



# DIE STUNDE DER BASTLER



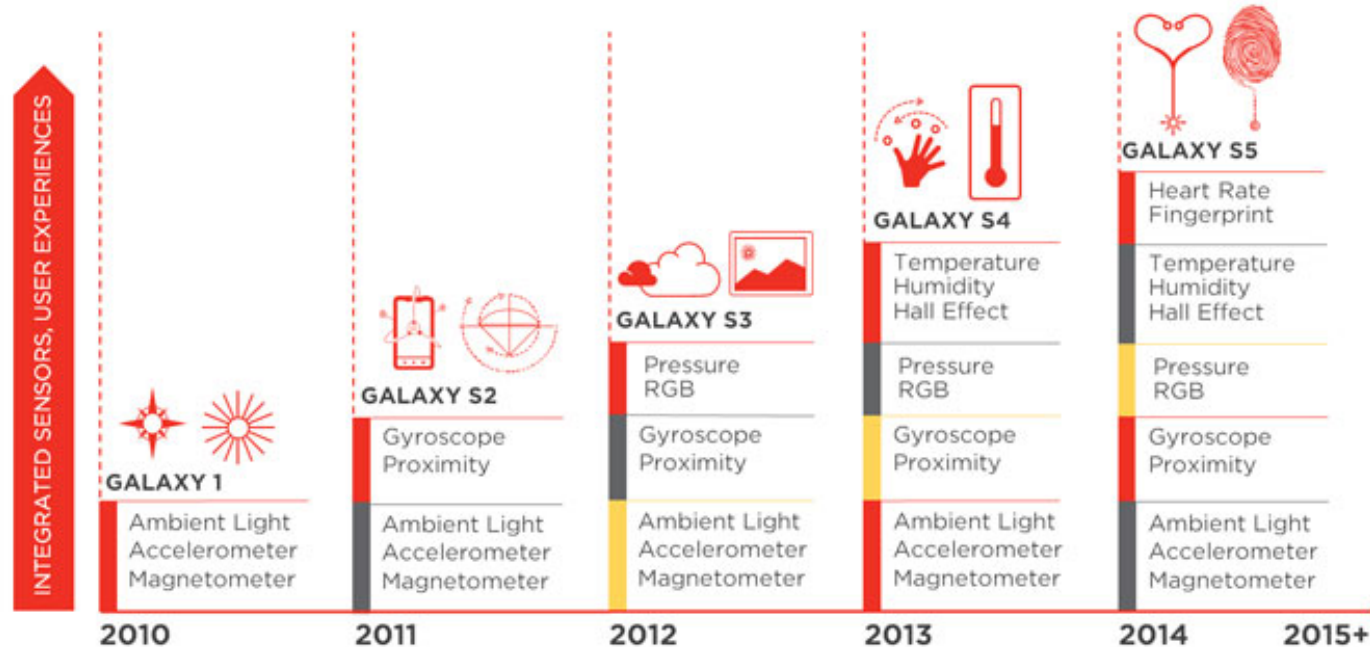
Quelle: Waterford Institute of Technology, [https://www.wit.ie/schools/engineering/wits\\_self-driving-car](https://www.wit.ie/schools/engineering/wits_self-driving-car)





# SMARTPHONES ALS SENSOREN

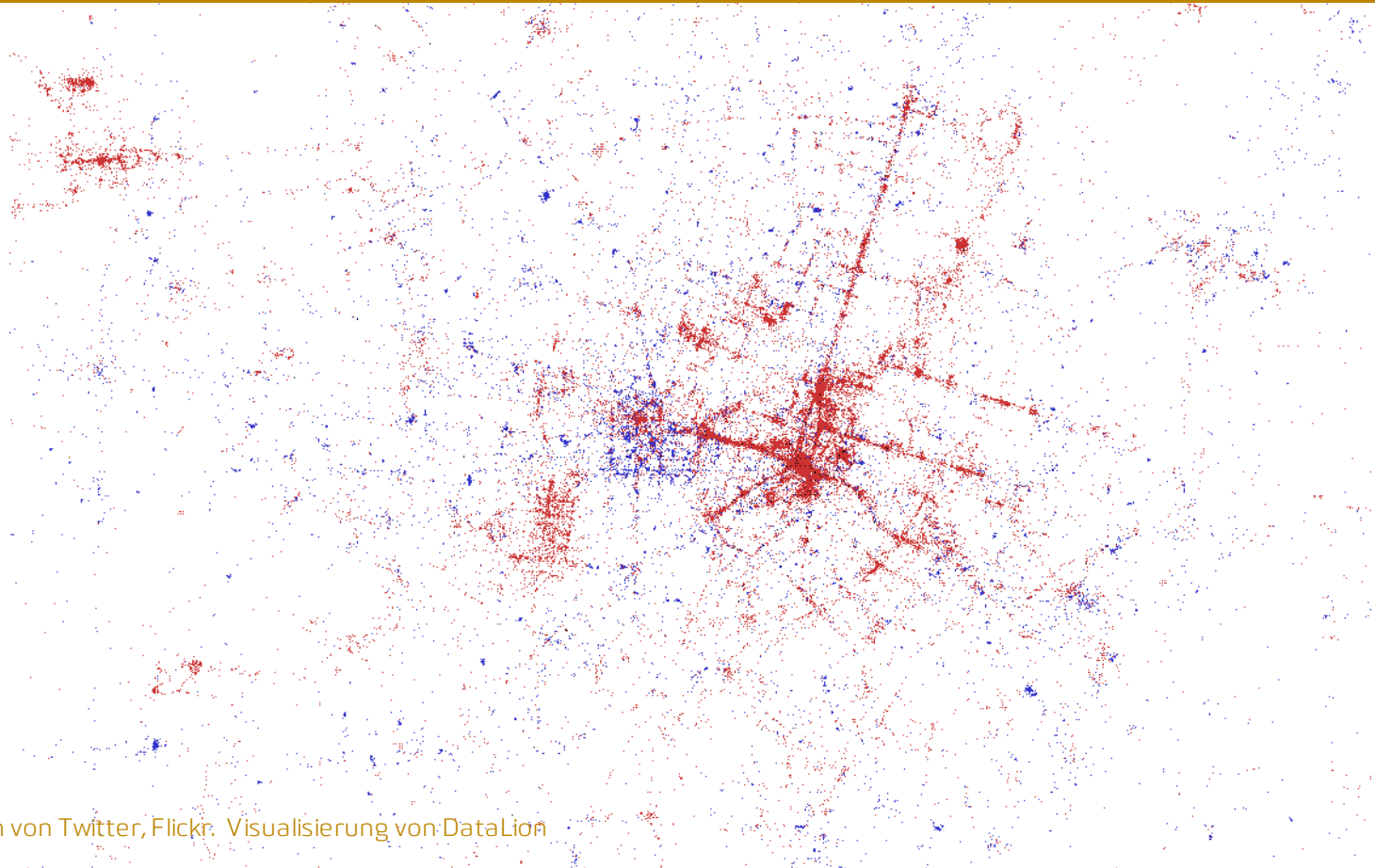
## SENSOR GROWTH IN SMARTPHONES







# LEBEN IN DER DATENWOHLSTANDSGESELLSCHAFT



Quelle: Daten von Twitter, Flickr. Visualisierung von DataLion










# DAS INTERNET ALS UNIVERSITÄT



Catalog

Search catalog

Q

Institutions

Log In

Sign Up

Overview

Syllabus

Creators

Ratings and Reviews

Machine Learning

Enroll Now

Starts Jun 13

Financial Aid is available for learners who cannot afford the fee. [Learn more and apply.](#)

Home > Data Science > Machine Learning


Machine Learning


**About this course:** Machine learning is the science of getting computers to act without being explicitly programmed. In the past decade, machine learning has given us self-driving cars, practical speech recognition, effective web search, and a vastly improved understanding of the human genome. Machine learning is so pervasive today that you probably use it dozens of times a day without knowing it. Many

[More](#)

Created by:

Stanford University





**Taught by:** Andrew Ng, Associate Professor, Stanford University; Chief Scientist, Baidu; Chairman and Co-founder, Coursera

Quelle: Coursera <https://www.coursera.org/learn/machine-learning>





# TECHNOLOGIETRANSFER AUF YOUTUBE

YouTube DE Suchen

## Representations at NLP Levels: Phonology

- Traditional: Phonemes

© 2005 IPA

	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p b		t d		ʈ ɖ		k ɡ	q ɢ			ʔ
Nasal	m	ɱ	n		ɳ		ɲ	ŋ			ɴ
Fricative		f v	s z		ʃ ʒ		x ɣ				
Tap or Flap			ɾ								
Approximant				l							
Lateral approximant				l							
Approximant				ɹ							
Lateral approximant				ɭ							
Approximant				ɻ							

Where symbols appear in pairs, the one on the left represents a voiced consonant. Shaded areas denote articulation judged impossible.

- DL: trains to predict phonemes (or words directly) from sound features and represent them as **vectors**

Lecture 1, Slide 20 Richard Socher 3/29/16

CS224D Lecture 1 - 29th Mar 2016





These 3:  
Künstliche Intelligenz  
ist das neue Internet



# HIGH TECH ALS ÖFFENTLICHES GUT

TensorFlow is an Open Source Software  
Library for Machine Intelligence

GET STARTED

## About TensorFlow

TensorFlow™ is an open source software library for numerical computation using data flow graphs. Nodes in the graph represent mathematical operations, while the graph edges represent the multidimensional data arrays (tensors) communicated between them. The flexible architecture allows you to deploy computation to one or more CPUs or GPUs in a desktop, server, or mobile device with a single API.

Quelle: TensorFlow, <https://www.tensorflow.org/>

archers and







# DIE REVOLUTION: KING – MAN + WOMAN = ???

man:woman :: king:?

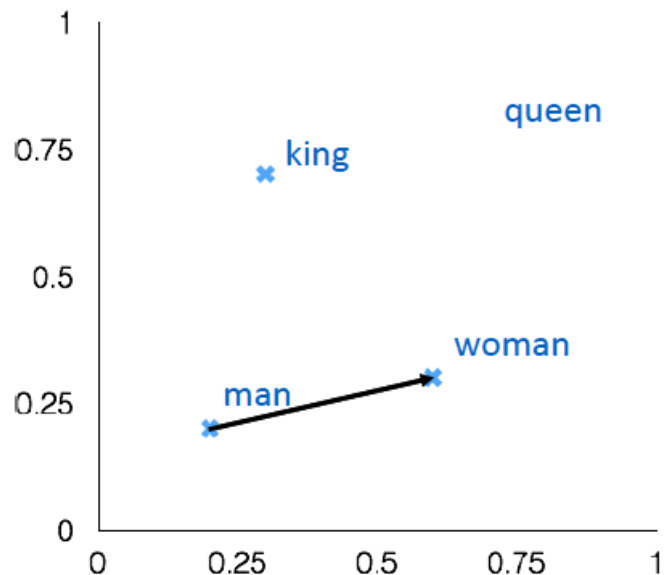
+ king [ 0.30 0.70 ]

- man [ 0.20 0.20 ]

+ woman [ 0.60 0.30 ]

---

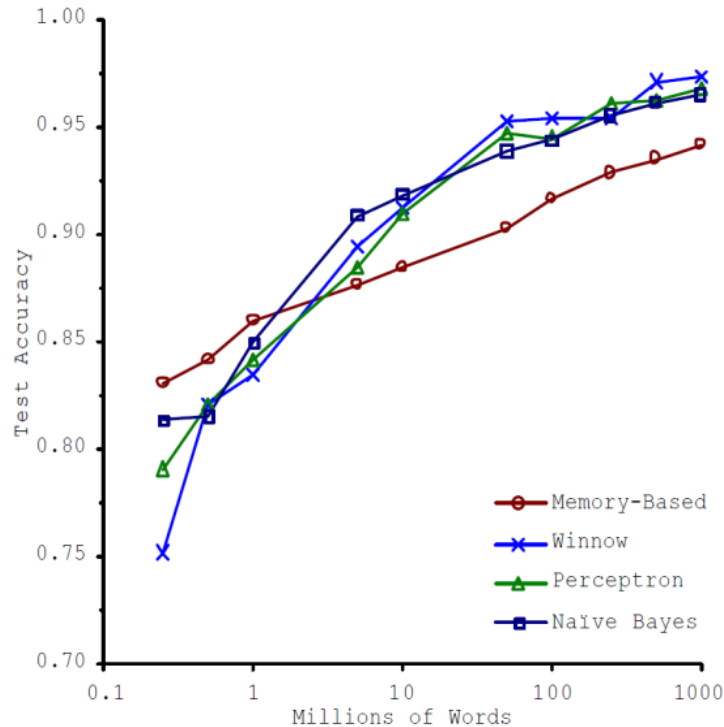
queen [ 0.70 0.80 ]







# DATEN SIND WICHTIGER ALS ALGORITHMEN





# WIKIPEDIA ALS „GENERALSCHLÜSSEL“

- Hauptseite
- Themenportale
- Von A bis Z
- Zufälliger Artikel
- Mitmachen
  - Artikel verbessern
  - Neuen Artikel anlegen
  - Autorenportal
  - Hilfe
  - Letzte Änderungen
  - Kontakt
  - Spenden
- Werkzeuge
  - Links auf diese Seite
  - Änderungen an verlinkten Seiten
  - Spezialseiten
  - Permanenter Link
  - Seiteninformationen
  - Wikidata-Datenobjekt
- Drucken/exportieren
  - Buch erstellen
  - Als PDF herunterladen
  - Druckversion
- In anderen Projekten
  - Commons

Quelle: Wikipedia, <https://de.wikipedia.org/wiki/Wikipedia:Hauptseite>

Wikibooks

Willkommen bei Wikipedia

947.505 Artikel

in deutscher Sprache entstanden.

- Geographie**
- Geschichte**
- Gesellschaft**
- Kunst und Kultur**
- Religion**
- Sport**
- Technik**
- Wissenschaft**

**Artikel nach Themen · Alphabetischer Index · Artikel nach Kategorien · Gesprochene Wikipedia**

**Kontakt · Presse · Statistik · Sprachversionen · Mitmachen · Mentorenprogramm**

## Artikel des Tages



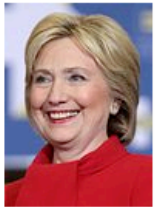
**Synesios von Kyrene** (griechisch Συνέσιος *Synésios*; \* um 370; † nach 412) war ein spätantiker griechischer Philosoph, Schriftsteller und Dichter. Er stammte aus einer vornehmen Familie der Griechenstadt Kyrene im Osten des heutigen Libyen. Ab 411/412 amtierte er als

Bischof der Provinzhauptstadt Ptolemais in seiner Heimatregion, der Kyrenaika. Das Denken des Synesios war maßgeblich vom Neuplatonismus geprägt, der damals die vorherrschende philosophische Schulrichtung war. Eine wesentliche Rolle spielte dabei der Einfluss der paganen Neuplatonikerin Hypatia. Bei ihr erhielt Synesios seine philosophische Ausbildung, danach blieb er eng mit ihr befreundet. Da er zugleich Christ war, sah er sich vor die Aufgabe gestellt, das platonische Welt- und Menschenbild mit dem christlichen Glauben in Einklang zu bringen. Die synkretistische Religiosität, die sich daraus ergab, kommt in seinen Hymnen und seiner Korrespondenz zum Ausdruck. Die Verbindung von philosophischem Erkenntnisstreben und literarisch-musischer Aktivität stellte für ihn die ideale Lebensform dar, für die er als Schriftsteller eintretet. Zum Artikel

## In den Nachrichten

**Hochwasser in Europa · Mückenatlas · Copa América**

- Durch einen **Bombenanschlag in Istanbul** wurden mindestens elf Menschen getötet.
- Die **demokratische US-Präsidentschaftsbewerberin Hillary Clinton** (Bild) hat bei den **Vorwahlen** genügend Delegierte zur Nominierung auf dem **Parteitag** erreicht.
- Bei einem **Zugunglück bei Lüttich** sind mindestens drei Menschen ums Leben gekommen.



**Weitere aktuelle Ereignisse · Wikinews**

## Kürzlich Verstorbene

- Viktor Kortschnoi** (85), sowjetisch-schweizerischer Schachspieler († 6. Juni)
- Peter Shaffer** (90), britischer Dramatiker († 6. Juni)
- Jerome Bruner** (100), US-amerikanischer Psychologe († 5. Juni)
- Phyllis Curtin** (94), US-amerikanische Opernsängerin († 4. Juni)
- Elfriede Eilers** (95), deutsche Politikerin († 4. Juni)

**RSS-Feed**





# AUTOMATISIERTE TAFELBILDER

takeNote

Home

Slides

About

## Early Universe - V. Rubakov - lecture 1/9

Galileo Galilei

GGI lectures on the theory of fundamental interactions,  
January 2015 <http://heldi.pd.infn.it/html/GGI/index.php>

Processing 00:22:39/01:30:12

Stop processing



Save image

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu = dt^2 - a^2(t) \gamma_{ij} dx^i dx^j$$
$$\gamma_{ij} = \begin{cases} 3\text{-hyperb.} & \kappa = -1 \iff \\ 3\text{-Eucled.} & \kappa = 0 \iff \\ 3\text{-sphere} & \kappa = 1 \iff \end{cases} \quad a = \text{radius of curv.}$$
$$\frac{a(t)}{a(t')}$$

$t_0 = \text{present time}$   
 $a = a(t_0) = \text{present}$

00:09:21

Blackboard at 00:21:54

$$\rho_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi G \quad T_{\mu\nu}$$
$$\left( \frac{\dot{a}}{a} \right)^2 H^2 = \frac{8\pi}{3} G \rho - \frac{\kappa}{a^2}$$
$$T_{00} = \rho \quad T_{ij} = a^2 p \gamma_{ij}$$
$$\nabla_\mu T^{\mu\nu} = 0$$

Blackboard snippets

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu$$

00:04:36

$$= dt^2 - a^2(t) \gamma_{ij} dx^i dx^j$$

00:05:06

$$\gamma_{ij} = \begin{cases} 3\text{-hyperb.} & \kappa = -1 \\ 3\text{-Eucled.} & \kappa = 0 \\ 3\text{-sphere} & \kappa = 1 \end{cases}$$

00:05:36

$$\begin{matrix} 3\text{-Eucled.} & \kappa = 0 \\ 3\text{-sphere} & \kappa = 1 \end{matrix}$$

00:06:03

$$\iff a = \text{radius of curv.}$$

00:07:15

$$\frac{a(t)}{a(t')}$$

00:07:54

$$\begin{matrix} t_0 = \text{present time} \\ a = a(t_0) = \text{present} \end{matrix}$$

00:09:21



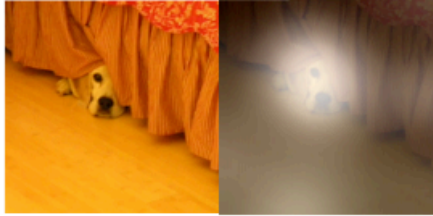


# BILDER KORREKT ZUORDNEN

Figure 3. Examples of attending to the correct object (*white* indicates the attended regions, *underlines* indicated the corresponding word)



A woman is throwing a frisbee in a park.



A dog is standing on a hardwood floor.



A stop sign is on a road with a mountain in the background.



A little girl sitting on a bed with a teddy bear.



A group of people sitting on a boat in the water.



A giraffe standing in a forest with trees in the background.





# SUCHMASCHINE FÜR FILMSZENEN



**Raging Bull (1980)**



**8 1/2 (1963)**



**Paris, Texas (1984)**



**Mulholland Dr. (2001)**



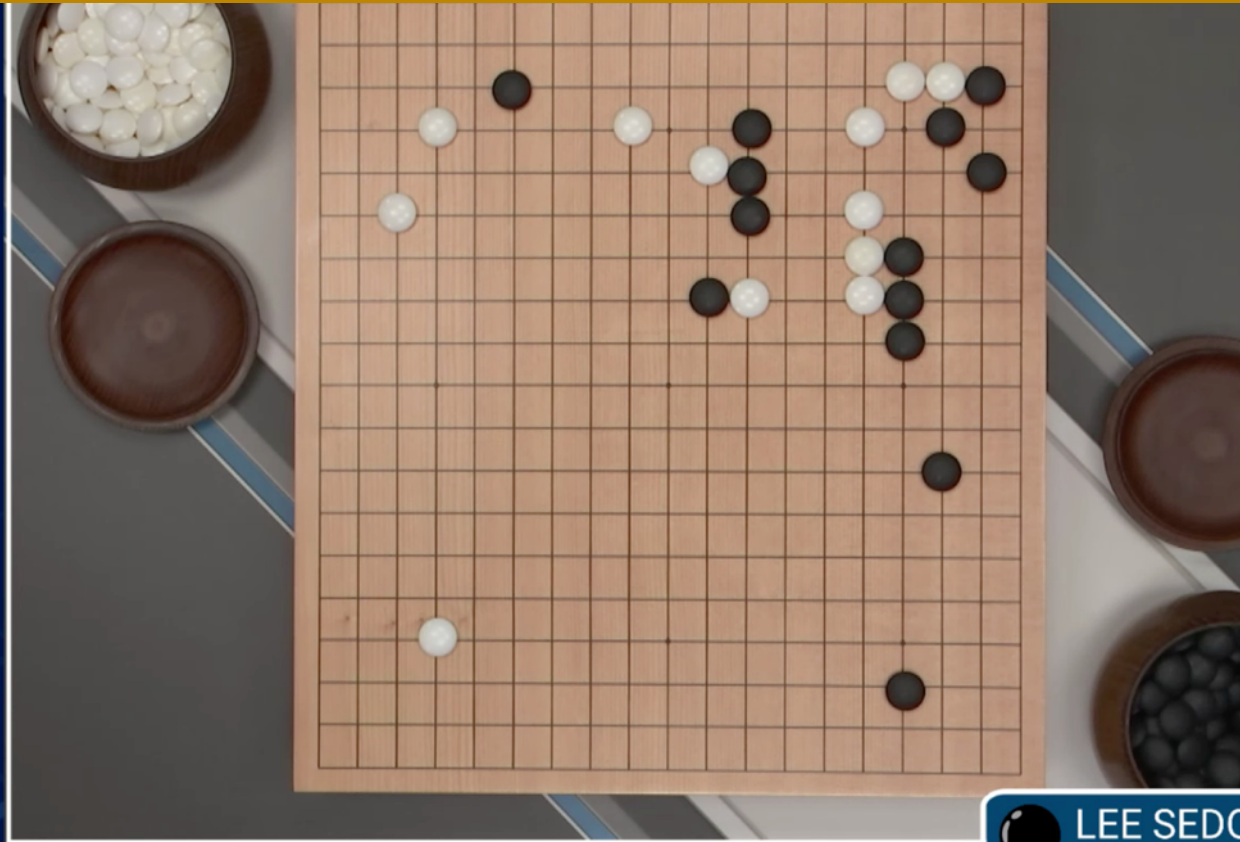
**Day For Night (1973)**



**The Godfather (1972)**



# IMMER KOMPLEXERE AUFGABEN







# MENSCHLICHES VERHALTEN ENTSCHLÜSSELN

## Sprach- & Audioaktivitäten

Schauen Sie sich Ihre Sprach- und Audioeingaben an, die Sie an Google gesendet haben.



Nur Sie können diese Daten sehen.

Sie haben die Kontrolle. Sie können jederzeit Einträge löschen oder die Optionen unter "Einstellungen" aktualisieren.



Heute

Einige Einträge werden möglicherweise nicht sofort angezeigt.

Keine Einträge



Gestern



Gesagt Kalender

Google App

15:39

WIEDERGEHEN



Gesagt mein meine Termine

Google App

15:39

WIEDERGEHEN



Gesagt Termine

Google App

15:39

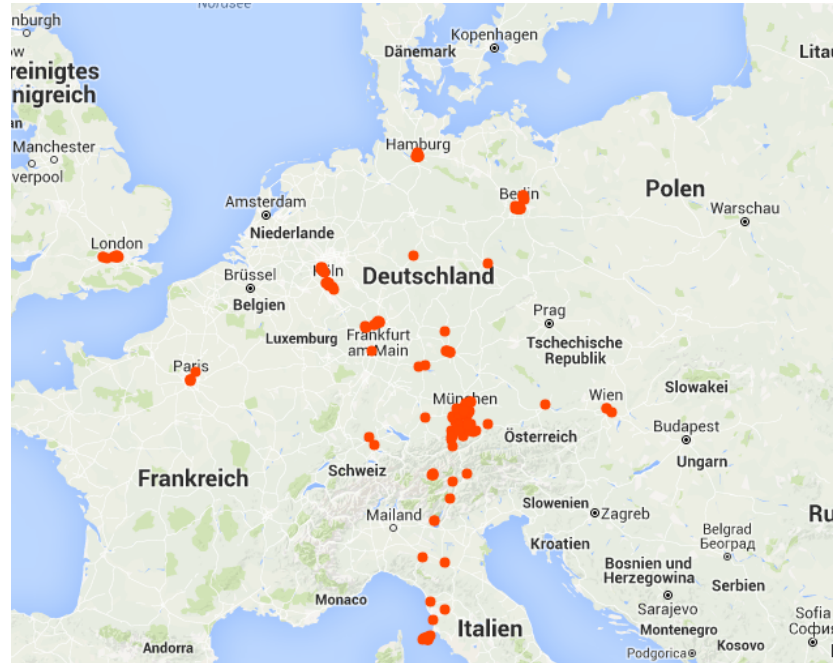
WIEDERGEHEN



Gesagt welche Termine habe ich heute

Google App

15:37







These 4:  
Neugierde wird zum  
Wettbewerbsvorteil



# DIE BIG DATA-AGENDA FÜR UNTERNEHMEN

**The Sexiest Job of the 21st Century:** Dialog mit Data Scientists aufbauen

**It's the Data, Stupid:** Datenanalyse von Anfang an mitdenken

**The Data Lake:** Alle Rohdaten abspeichern ohne sie zu bereinigen

**Domain-Experten:** Aus den Daten die richtigen Schlüsse ziehen

**Automatisierung:** Algorithmen- und Machine Learning-Pipelines definieren

**Data Storytelling:** Daten visualisieren und die richtigen Geschichten erzählen

**Neugierde als Wettbewerbsvorteil kultivieren:** Bootcamps, Meetups, Hackathons, Data Challenges



A stylized, golden-brown lion head logo on the left side of the page. The lion's face is white with a golden-brown outline, and its mane is composed of many thin, parallel golden-brown lines radiating outwards.

*We're hiring!*

## KONTAKT:

DataLion GmbH  
Goethestr. 21  
80336 München  
Deutschland

☎ +49 89 716 772 100

🌐 [www.datalion.com](http://www.datalion.com)

✉ [info@datalion.com](mailto:info@datalion.com)

🐦 [@datalion\\_en](https://twitter.com/datalion_en)

Made with ♥ in Bavaria



# ÜBER DATALION

DataLion ist eine Software zur Analyse und Visualisierung, die das Arbeiten mit Daten allen zugänglich macht.

Unsere webbasierte SaaS-Plattform bietet zahlreiche Werkzeuge zur Datenanalyse. Mit unserer visuellen **Dashboard-Engine** kann jeder Manager selbst Daten analysieren - jenseits trockener Zahlen und Tabellen.

Aktuell nutzen bereits 200 Unternehmen DataLion, um mehr aus ihren **Marketing-, Marktforschungs- oder Vertriebsdaten** herauszuholen.

