

Gegenwart und Zukunft von Big Data

Dieter Kranzlmüller

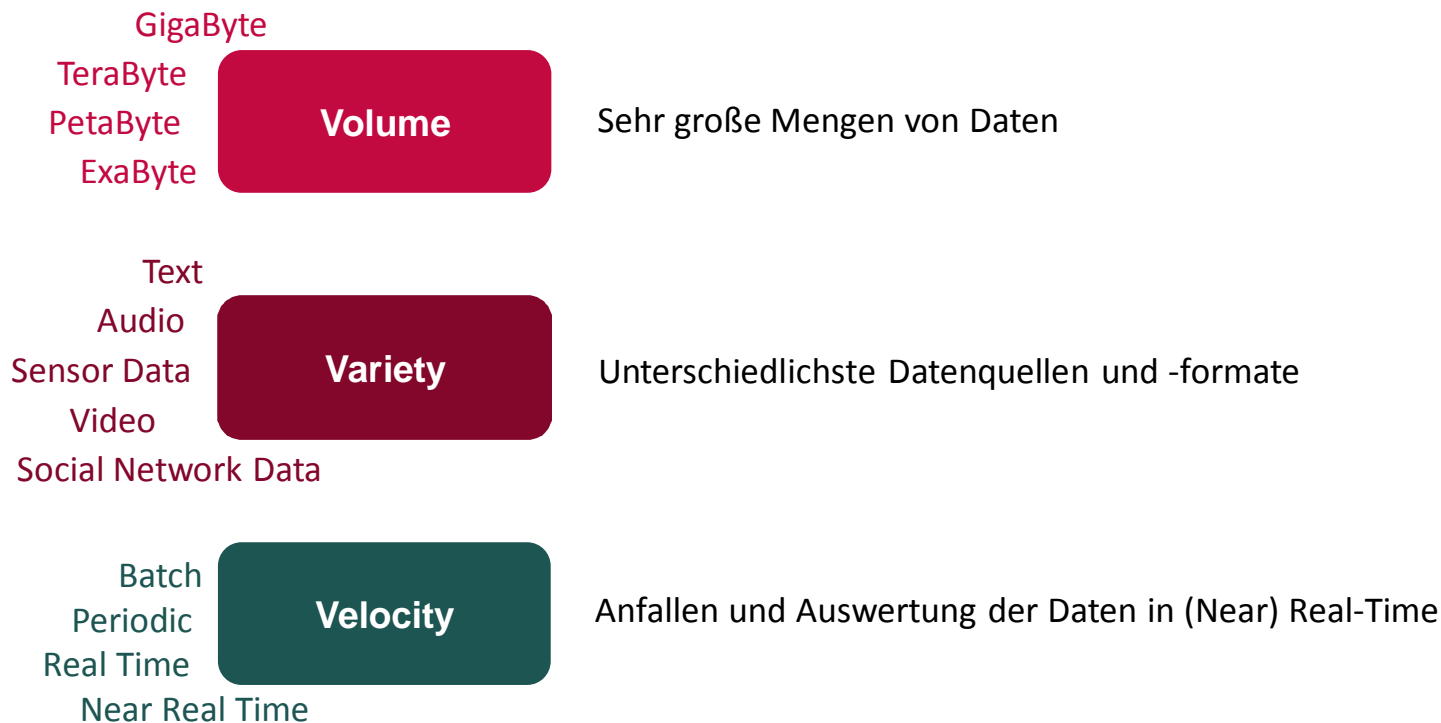
Munich Network Management Team
Ludwig-Maximilians-Universität München (LMU) &
Leibniz-Rechenzentrum (LRZ)
der Bayerischen Akademie der Wissenschaften



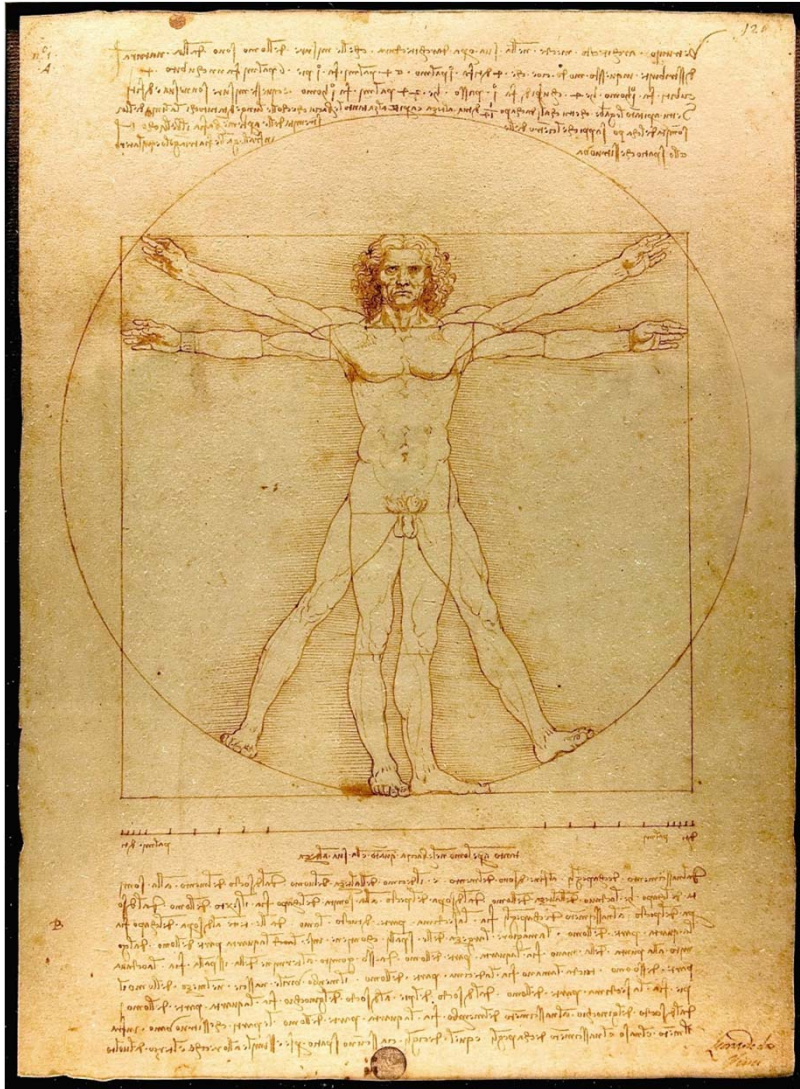
Quelle: Google Analytics

3 V's: “Big data is high-volume, -velocity and -variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making.”

Beyer, M. & Laney, D. (2012): The Importance of 'Big Data': A Definition, Gartner Research Report



Quelle: Arnold Picot, Munich School of Management, LMU



Besteht aus ungefähr

$3,72 \times 10^{13}$

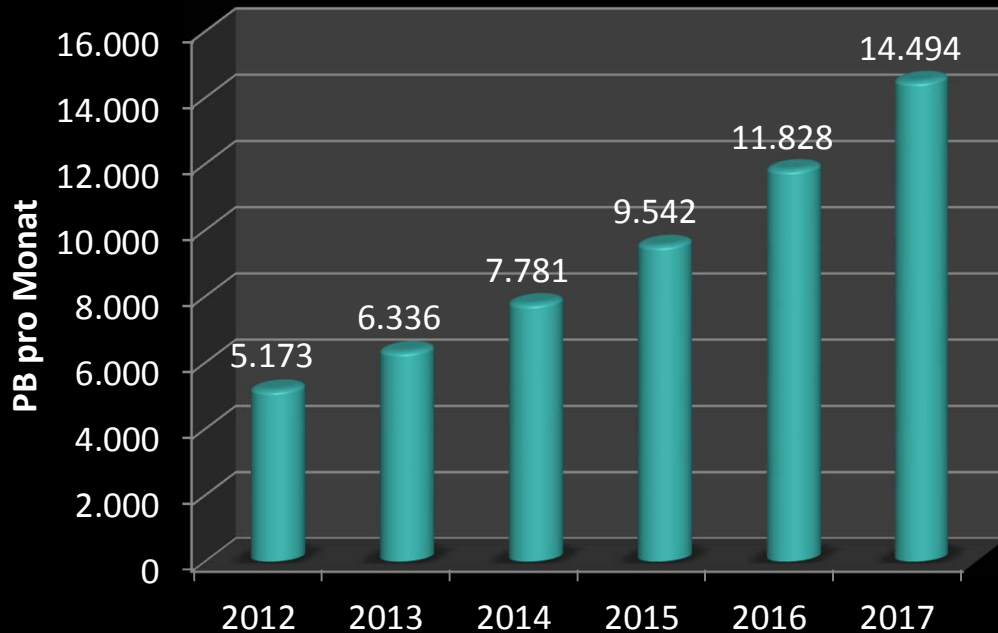
(37,2 Billionen)

Zellen

Quelle: Bianconi, Eva et al: *An estimation of the number of cells in the human body*. In: *Annals of Human Biology* 40 (2013), Nr. 6, S. 463–471

Quelle: http://upload.wikimedia.org/wikipedia/commons/2/22/Da_Vinci_Vitruve_Luc_Viatour.jpg

Consumer Web, Email, Data



Besteht 2013 aus ungefähr

633.600 x 10¹³
(6,3 Trillionen)

Bytes/Monat(!)

^{*)} Nur Consumer Web, Email, Instant Messaging und anderer Datenverkehr (ohne File Sharing)

Quelle: Cisco Visual Networking Index: Forecast and Methodology, 2012-2017, Cisco 2013

Entspricht der Anzahl aller Körperzellen aller Einwohner von Saarbrücken (Einwohner 2012 gemäß Statistischem Bundesamt, 2013)



Quelle: <http://www.google.com/about/datacenters/>

Entspricht in etwa der Anzahl aller Körperzellen aller Einwohner Deutschlands (Einwohner 2012 gemäß Angaben der Weltbank von 2013)



ungefähr

280.000.000 x
10¹³

(2,8 Zettabyte oder 2,8 Trilliarden)

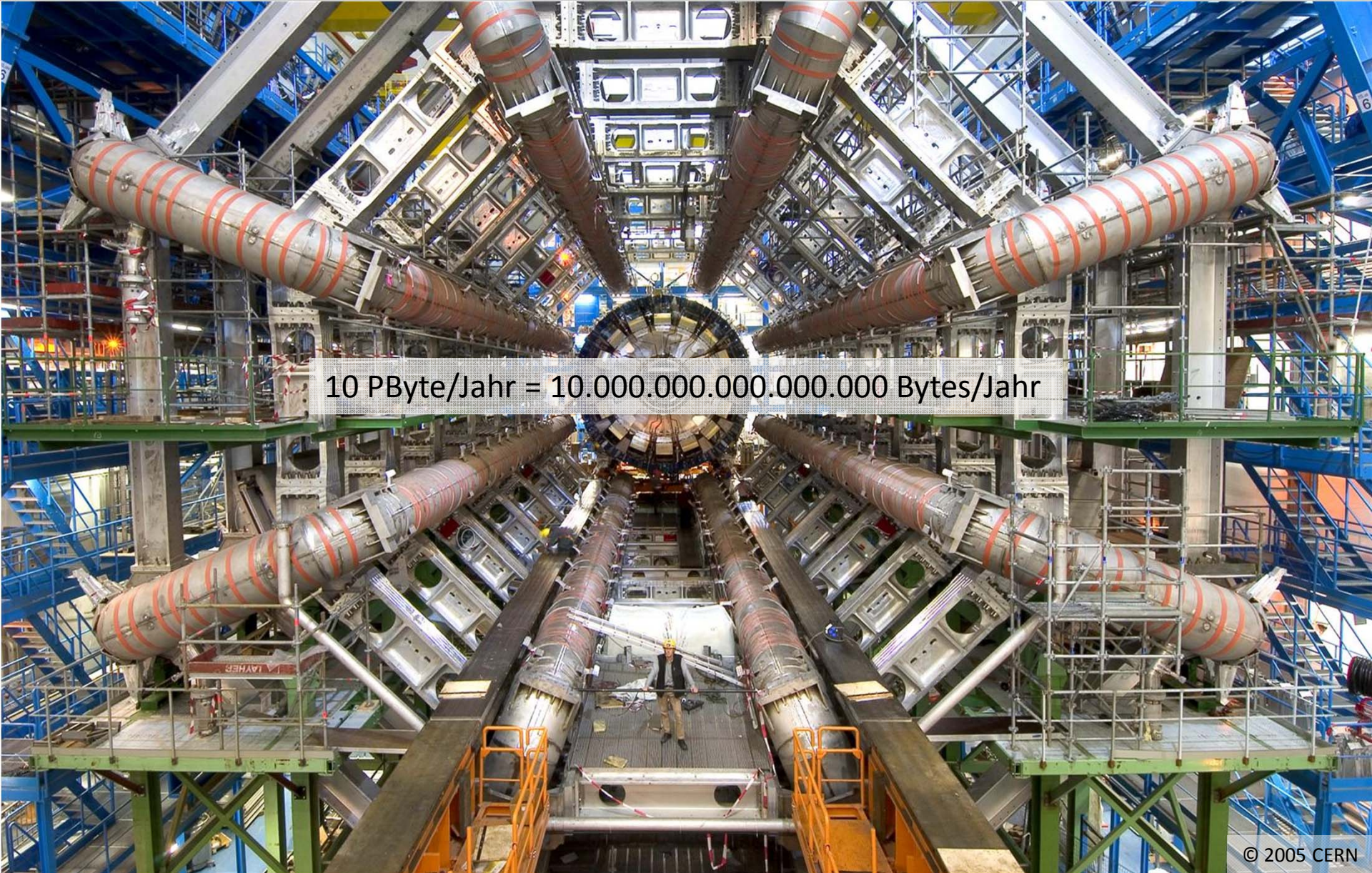
Bytes

Quelle: Spiegel Online vom 11.12.2012
(<http://www.spiegel.de/netzwelt/web/das-internet-der-dinge-erzeugt-2-8-zettabyte-daten-a-872280.html>)



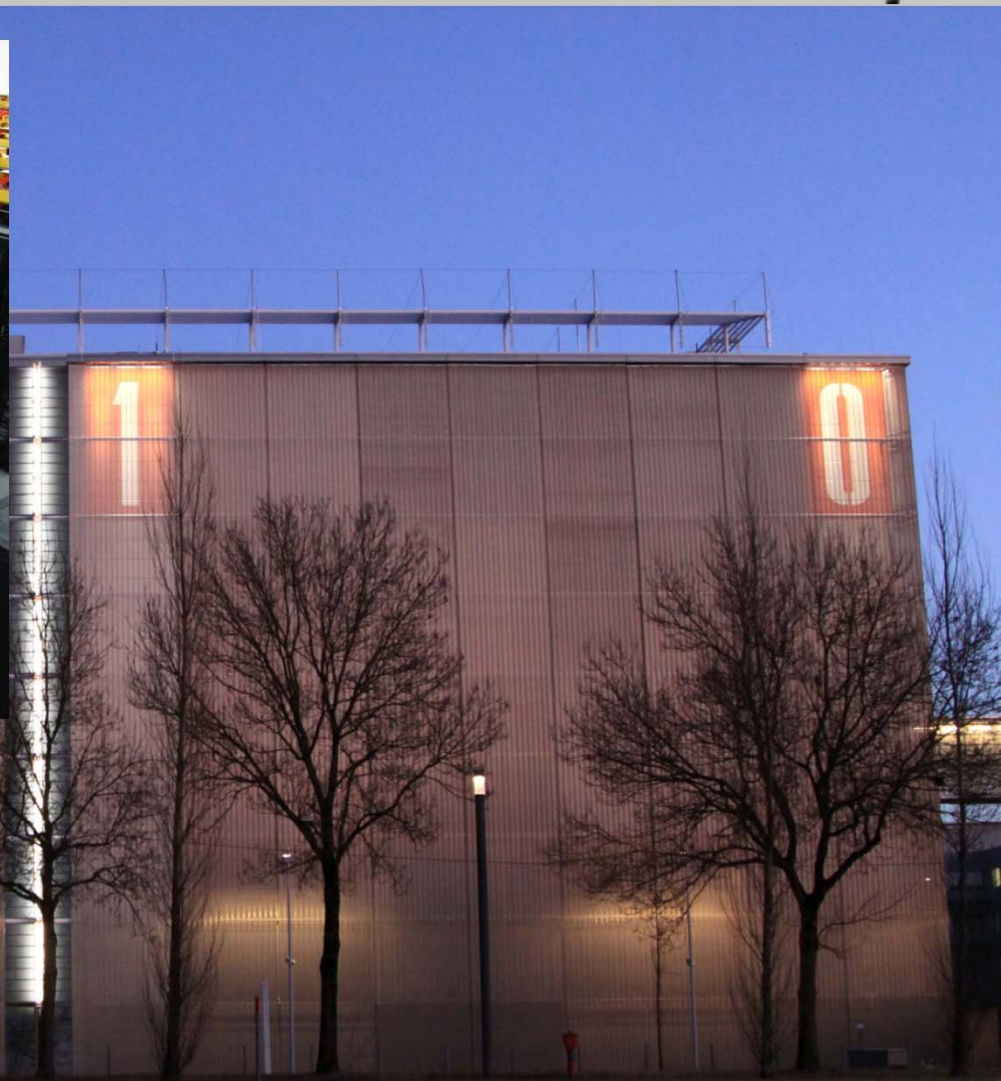
<http://hqwide.com/eastwood-western-the-good-bad-and-ugly-wallpaper-58265/>





10 PByte/Jahr = 10.000.000.000.000.000 Bytes/Jahr

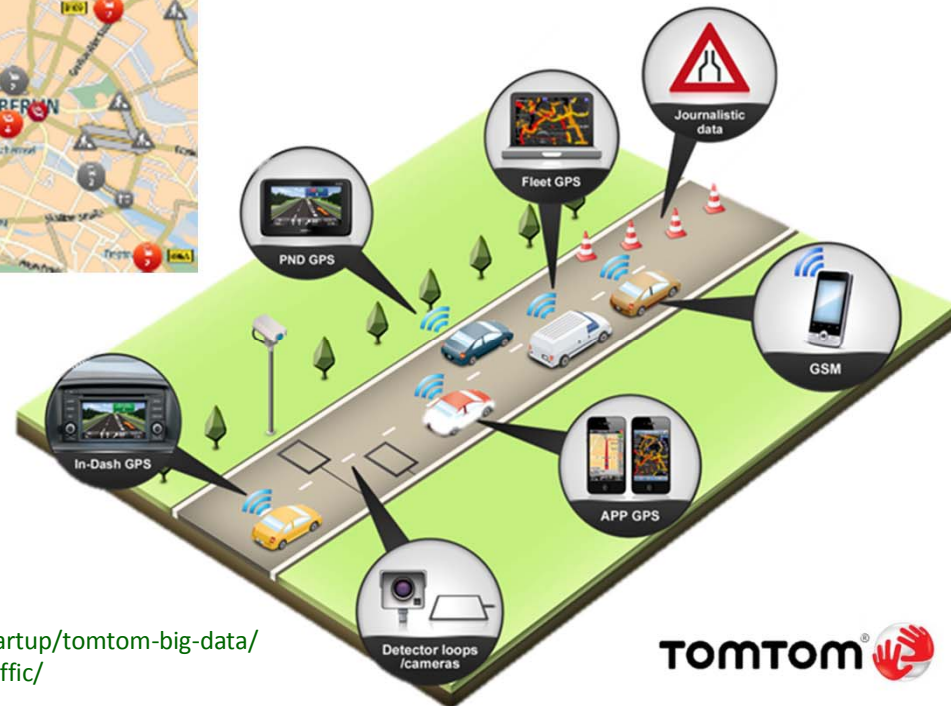
© 2005 CERN





- Echtzeitverkehrsdienst
- Liefert Stauinformationen durch Kombination von Informationen der Verkehrsleitzentralen mit Echtzeitdaten von über 350 Millionen Verkehrsteilnehmern (Crowdsourcing)

- Datenquellen:
 - Vernetzte GPS Geräte
 - Verkehrsleitzentralen
 - Journalisten, die Unfalldaten erfassen



Quellen:
 TomTom and Big Data; 11.11.2013; URL: <http://www.bigdata-startups.com/BigData-startup/tomtom-big-data/>
 TomTom Traffic; 11.11.2013; URL: http://www.tomtom.com/de_de/services/live/hd-traffic/
 Arnold Picot, Munich School of Management, LMU



Paul Butler a Facebook Intern, visualizes the relationships between 10 million pairs of Facebook friends

Quelle: <http://maptd.com/awesome-world-map-of-facebook-connections>



<http://hqwide.com/eastwood-western-the-good-bad-and-ugly-wallpaper-58265/>

- Zur Verbrechens-aufklärung und –bekämpfung
- Überwacht und vergleicht Daten aus verschiedenen Quellen, u.a.:
 - 3.000 Videokameras
 - Nummernschildscanner
 - Strahlungsdetektoren
 - div. Datenbanken



Quelle: Zeit (2012)
<http://nymag.com/daily/intelligencer/2012/08/nypd-domain-awareness-system-microsoft-is-watching-you.html>



aka **Intelligence Community
Comprehensive National
Cybersecurity Initiative
Data Center**

Camp Williams, Bluffdale, Utah

http://en.wikipedia.org/wiki/Utah_Data_Center

1,5 million square feet
2 billion US\$ building
2 billion US\$ hardware
65 MW power
1 Yottabyte Storage



To process “all forms of communication, including the complete contents of private emails, cell phone calls, and Internet searches, as well as all sorts of personal data trails - parking receipts, travel itineraries, bookstore purchases, and other digital 'pocket litter'.”

Masse Vielfalt
 Business Intelligence
 Speicherung SQL Server Semantische Metadaten
 Internet der Dinge Arbeitsspeicher Datenanalyse Petabytes
 Social Media Datenverarbeitung Spalten IT-Strategie Massiv-parallel
 Times Ten In-Memory Strukturiert Zeilen Analyse
 Cloud Batch Hadoop
 Visualisierung SAP Hana Exabyte
 Echtzeit Informationen
 Tabelle Video MapReduce Terabyte
 Periodisch OLAP Smart Content Zettabyte
 Teil-strukturiert Datenmanagement Informationen
 Geschwindigkeit



Dieter Kranzlmüller
 kranzlmueeller@lrz.de

Foto: Karl Behler
 Quelle: TNS Infratest (2012): Quo vadis Big Data - Herausforderungen – Erfahrungen – Lösungsansätze, Report 2012