



Big Data und künstliche Intelligenz

**Wie Daten gewonnen und ausgewertet werden
können.**

Mittwoch 15. März, 12:30 – 13:15 Uhr

Marcel Bernet



EB Zürich

Weiterbildung – wie ich sie will

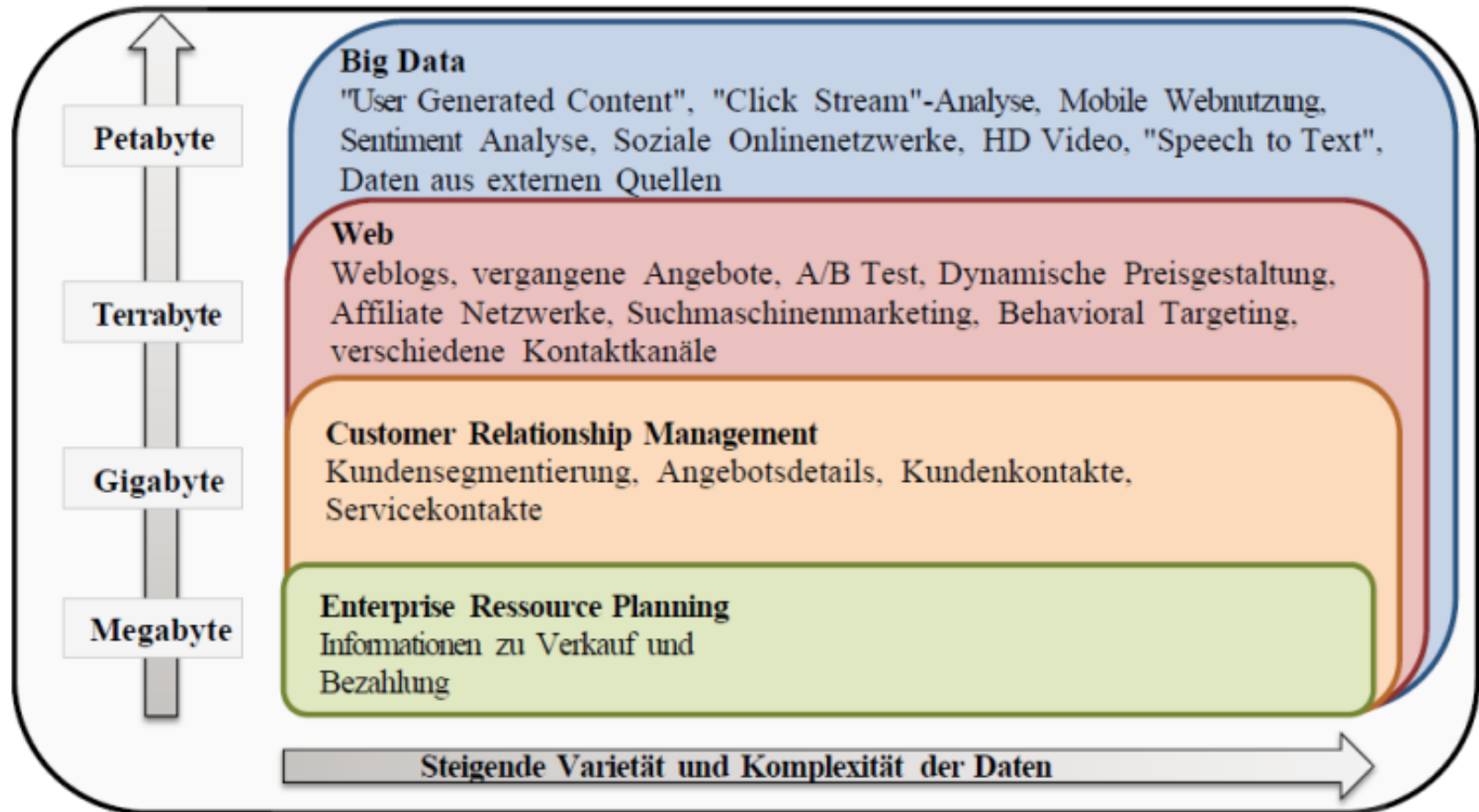
Definition

- Der aus dem englischen Sprachraum stammende Begriff Big Data bezeichnet Datenmengen, welche
 - **zu gross,**
 - **zu komplex,**
 - **zu schnelllebig oder**
 - **zu schwach strukturiert**
- sind, um sie mit manuellen und herkömmlichen Methoden der Datenverarbeitung auszuwerten

- Quelle: Wikipedia

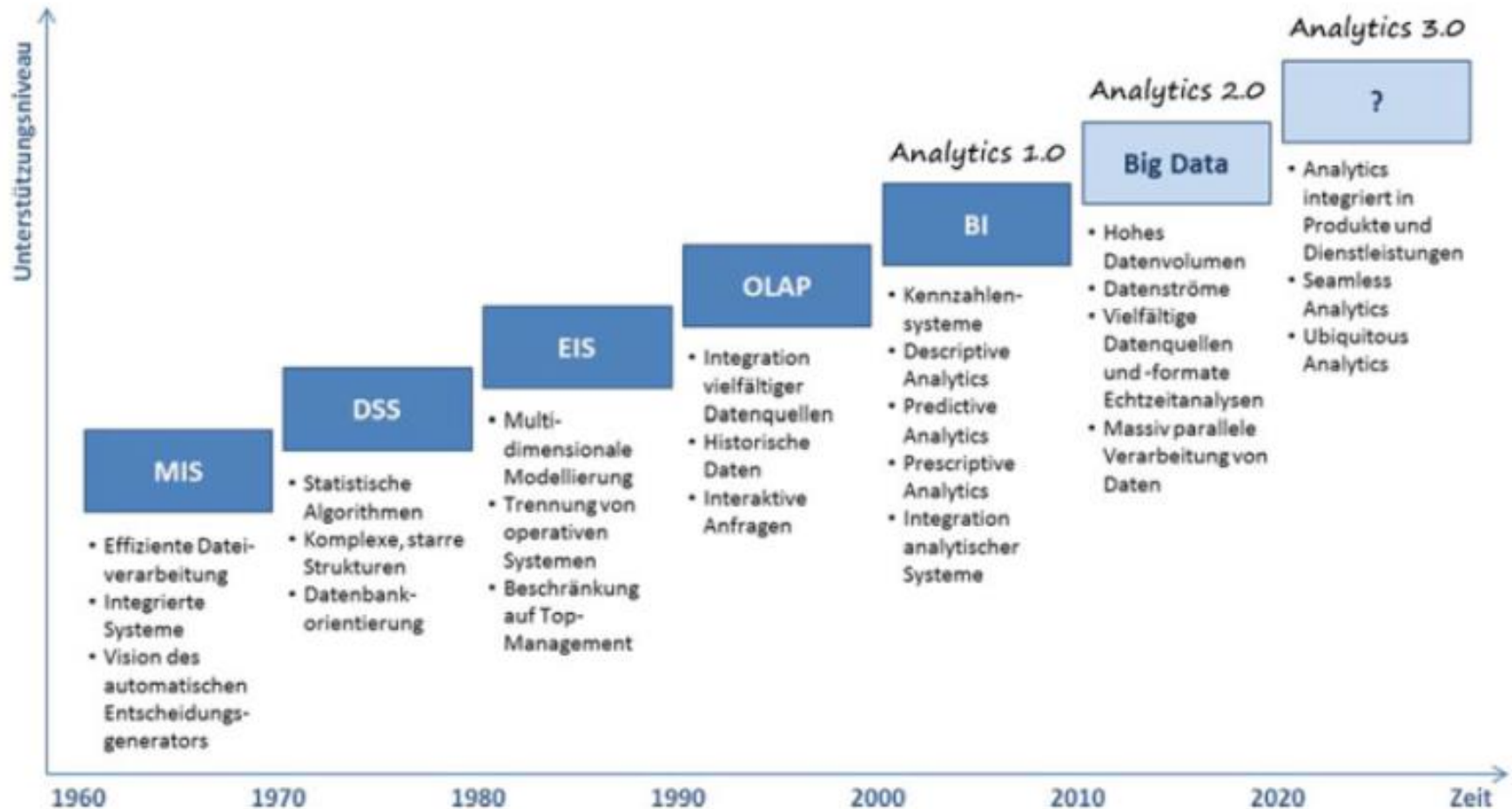
Big Data – Datenmengen

1 Petabyte PB = 1'000'000'000'000'000 Bytes (1'000 TB)



Quelle: Big Data, Potential und Barrieren der Nutzung im Unternehmenskontext

Big Data Stammbaum



BI-Stammbaum, Quelle: Praxishandbuch Big Data, Wirtschaft – Recht – Technik

Woher kommen die Daten?

- Überwachungssysteme
- Nutzung von Kunden- oder Kreditkarten
- Elektronische Kommunikation via Smartphone
- Geografische Bewegungsmuster (GPS)
- „Fitness“-, „Activity Tracker“ oder „Smartwatches“
- Computer
- Social Media
- Kraftfahrzeuge (insbesondere im Kontext „Vernetztes Auto“),
- Vernetzte Technik in Häusern („Smart Homes“, „Smart Meter“),
- Von Behörden und Unternehmen erhobene und gesammelte Daten.
- usw.

Potenzial



Quelle: Big Data, Grundlagen, Systeme und Nutzungspotenziale

Big Data Einsatzgebiete

- Personalisierung von Produkten, Services und Marketingmassnahmen
- Experimente und Analysen in Echtzeit ([A/B Testing](#))
- Innovation von Produkten und Services
- (Automatisierte) Planungs- und Entscheidungsprozesse
- Künstliche Intelligenz (Input)
- ...

Big Data und künstliche Intelligenz

WIE KOMMEN SIE AN DATEN?

Auswerten der Kunden- und Gerätedaten



- Ihre Kunden
- Ihre Produkte
- Ihr Maschinenpark
- Verwenden Sie:
 - Das «Internet der Dinge»
 - Moderne Web Plattformen (Web Shop)

Öffentliche Datenquellen



News

Manifest

Anlässe

Organisation

Vorstand

Mitglied werden

Kontakt

FR / DE

Suche



Anlässe

Der Verein Opendata.ch organisiert verschiedene Veranstaltungen. Regelmässige Veranstaltungen sind die [make.opendata.ch](#) Camps sowie die jährlichen Opendata.ch Konferenzen.

Open Cultural Data Hackathon
2017



Opendata.ch/2017



Mitgliederversammlung 2016



Opendata.ch/2016



April 2016: Energy Hackdays



Election Hackdays



Mitgliederversammlung 2015

Opendata.ch/2015 Konferenz


Mitgliederversammlung 2014

www.opendata.ch

Öffentliche Datenquellen – Links

- [Stadt Zürich](#)
- [Bundesamt für Statistik](#)
- [opendata.swiss](#)
- [Swiss public transport API](#)
- [Programmable Web](#)

The screenshot shows the top navigation bar of the 'Stadt Zürich Open Data' portal. It includes links for 'Open Data Portal', 'Anwendungen', 'Werkstatt', 'Blog', and social media icons for Twitter and Email. Below this is a dark blue header with the 'Stadt Zürich Open Data' logo, a 'Startseite' button, and links for 'Datensätze' and 'Kategorien'. A search bar with the placeholder 'Suche' and a magnifying glass icon is on the right. The main content area has a large heading 'Willkommen auf dem Open Data Katalog' followed by a paragraph: 'Der Datenkatalog ist Ihr zentraler Einstiegspunkt zur Suche und Nutzur stehen kostenlos und zur freien – auch kommerziellen - Weiterverwend'. To the right of this text is a sub-heading 'Fahrzeiten der VBZ im SOLL-IST-Vergleich'.

Open Data Portal Anwendungen Werkstatt Blog  

 **Stadt Zürich**
Open Data

Startseite Datensätze Kategorien

Suche 

Willkommen auf dem Open Data Katalog

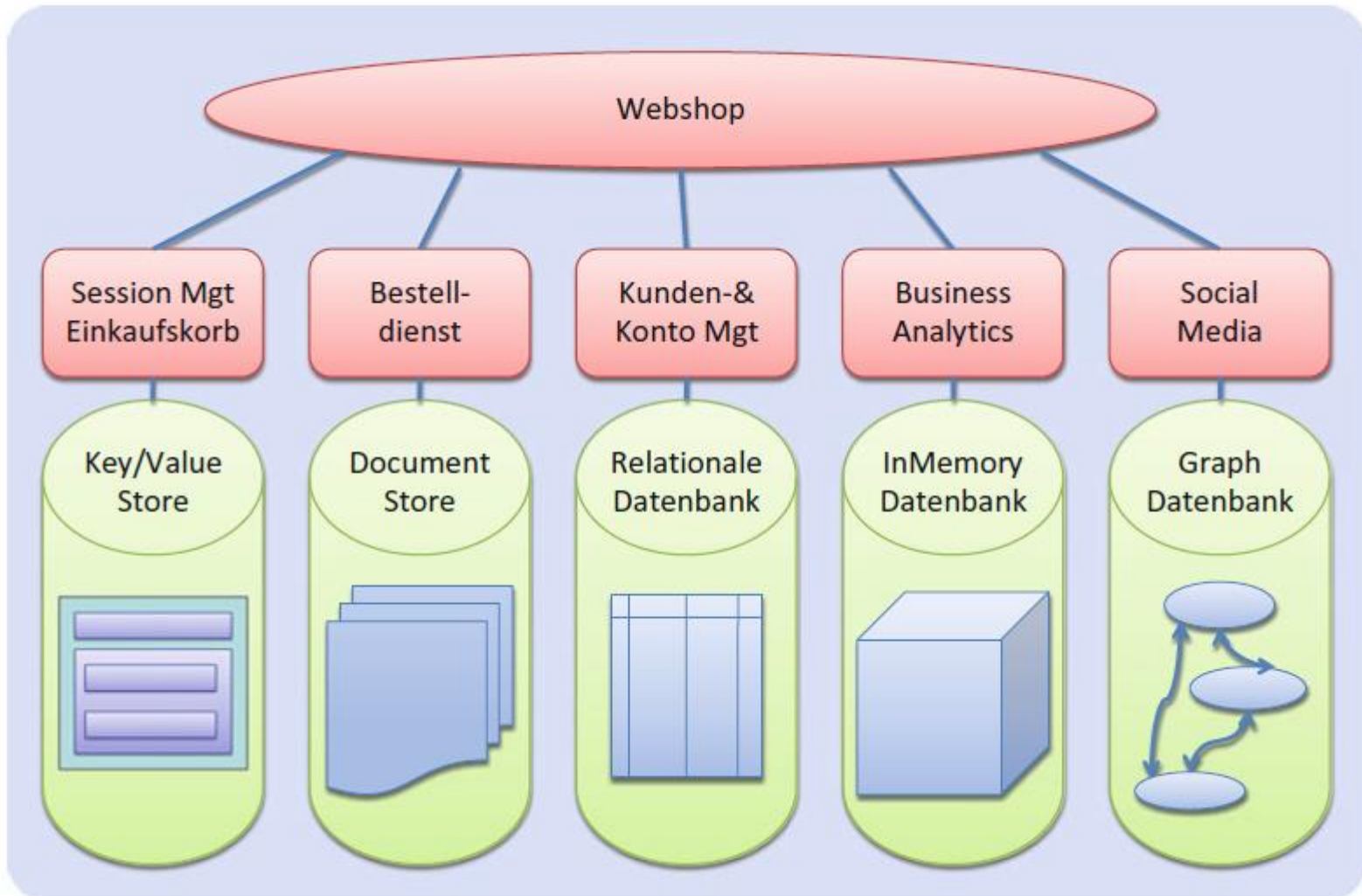
Der Datenkatalog ist Ihr zentraler Einstiegspunkt zur Suche und Nutzur stehen kostenlos und zur freien – auch kommerziellen - Weiterverwend

Fahrzeiten der VBZ im SOLL-IST-Vergleich

Big Data und künstliche Intelligenz

WIE WERTEN SIE DATEN AUS?

Art der Daten



Relationale Daten – Abfragesprache SQL

Auftragseingänge

1188	Schrauben
2178	Disketten
3266	Gips
3456	Beton

SELECT * FROM Auftragseingänge
WHERE NAME = «Schrauben»

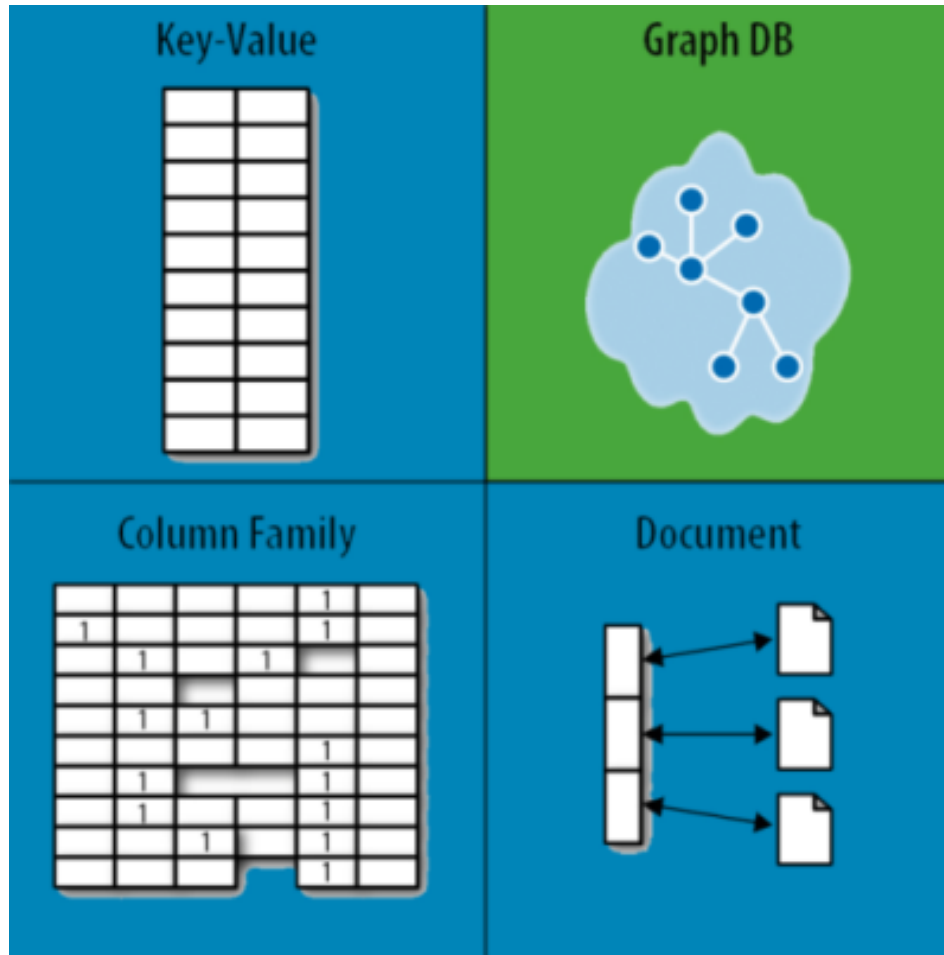
Kundendaten

1967	Fa. Kunz
2178	Frau Hacker
3266	Klotz u. Co
3279	Billy Boy

Offene Rechnungen

2165	Schrauben
2178	Disketten
2178	Mäuse
3279	Software

NoSQL – Not only SQL



Quelle: [Neo4j Blog](#)

NoSQL – Document Stores

```
{  
  _id: <ObjectId>,  
  username: "123xyz",  
  contact: {  
    phone: "123-456-7890",  
    email: "xyz@example.com"  
  },  
  access: {  
    level: 5,  
    group: "dev"  
  }  
}
```

Embedded sub-document

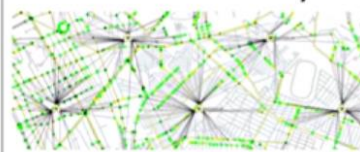
Embedded sub-document

NoSQL – Graph Databases

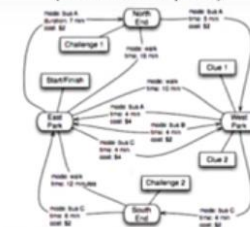
Social Networks
Customers and Employees



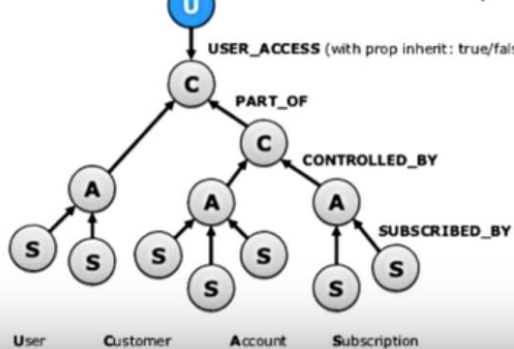
Network Cell Analysis



Geo Routing
(Public Transport)



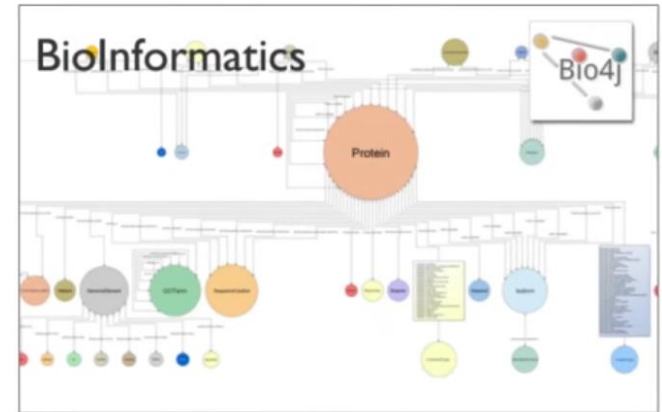
Product Subscriptions



Insurance Risk Analysis

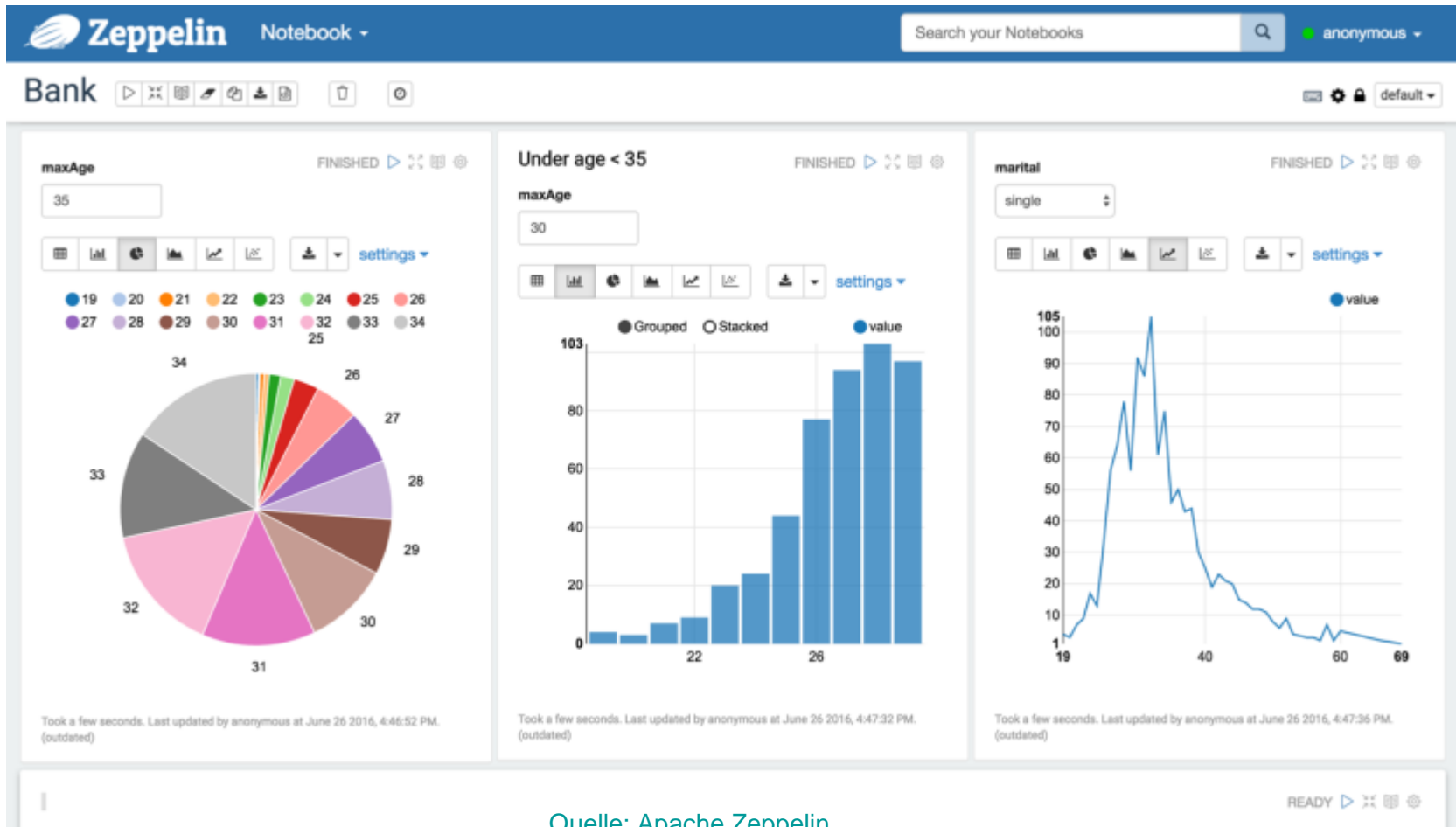


BioInformatics



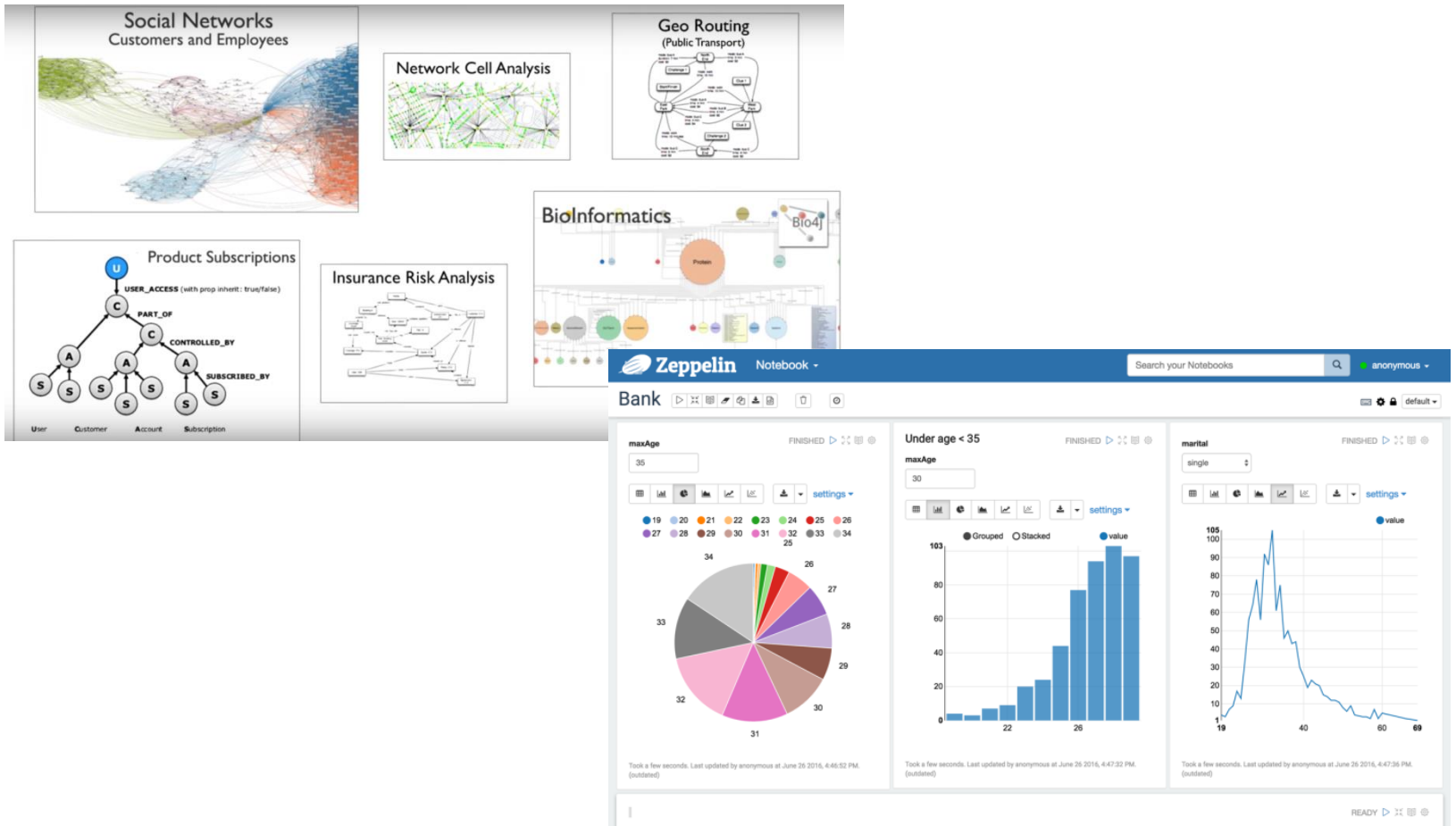
Quelle: <https://neo4j.com/graphacademy/online-training/introduction-graph-databases/>

Apache Zeppelin



[Quelle: Apache Zeppelin](#)

Demo



Zusammenfassung

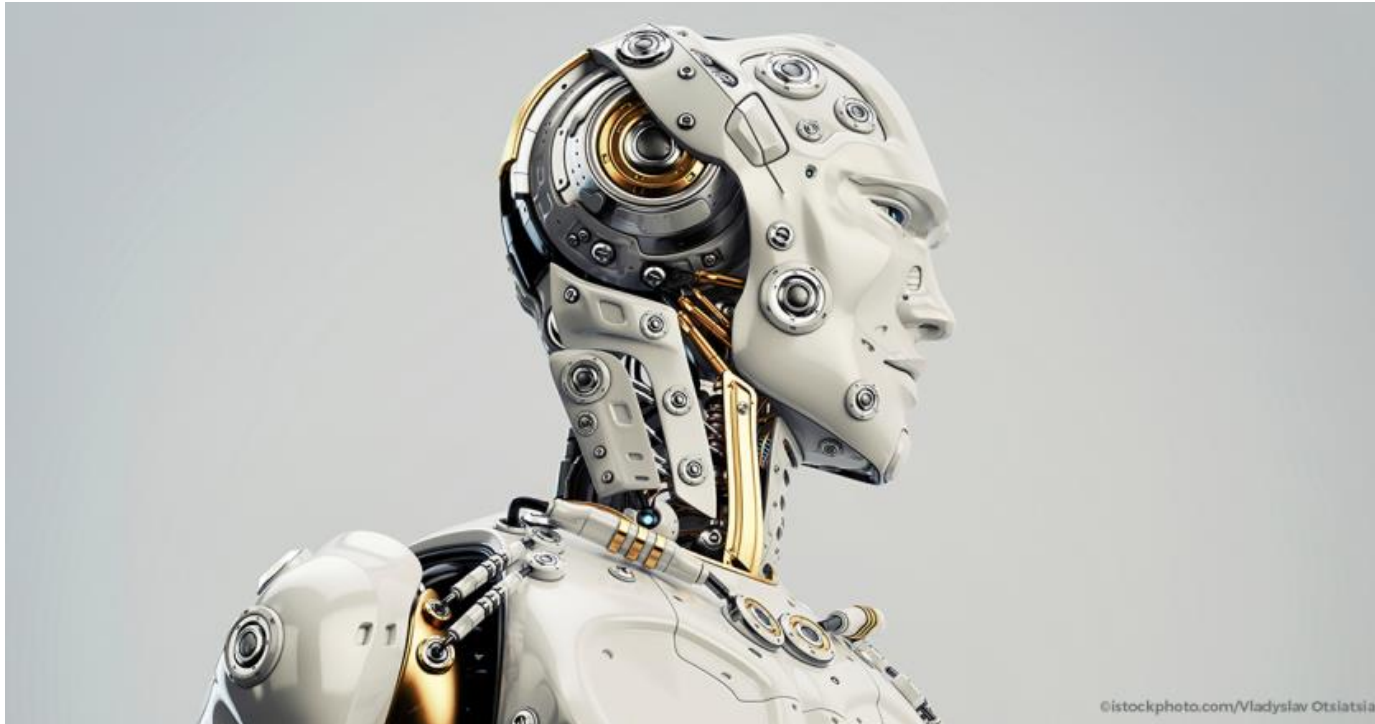
- Big Data sind Datenmengen, die **zu gross, zu komplex, zu schnelllebig** oder zu **schwach strukturiert** sind, um sie mit manuellen und herkömmlichen Methoden der Datenverarbeitung auszuwerten.
- Zur Speicherung und Auswertung werden deshalb neue Tools wie NoSQL Datenspeicher und neue Abfragesprachen wie Scala verwendet.

Big Data und künstliche Intelligenz

KÜNSTLICHE INTELLIGENZ



Künstliche Intelligenz (KI)



- Im Allgemeinen bezeichnet *künstliche Intelligenz* oder *KI* den Versuch, eine menschenähnliche Intelligenz nachzubilden

KI basiert (im Big Data Kontext) auf:

- Maschinelles Lernen ist ein Oberbegriff für die "künstliche" Generierung von Wissen aus Erfahrung
 - Ein künstliches System lernt aus Beispielen und kann diese nach Beendigung der Lernphase verallgemeinern.
 - **Es werden nicht einfach die Beispiele auswendig gelernt, sondern es "erkennt" Muster und Gesetzmässigkeiten in den Lerndaten.**
- **Algorithmen**

KI ist nicht neu

- Anfänge gehen auf 1956 zurück.

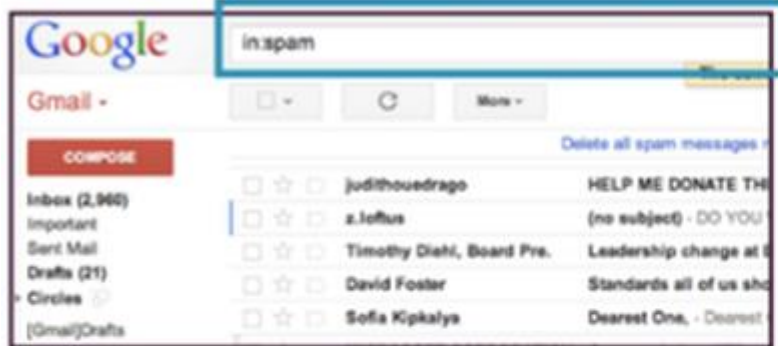
Am 13. Juli 1956 begann am Dartmouth College in Hanover (New Hampshire) eine 6-wöchige Konferenz zum Dartmouth Summer Research Project on Artificial Intelligence, die von John McCarthy organisiert wurde und an der neben McCarthy unter anderen Marvin Minsky, Nathan Rochester und Claude Shannon teilnahmen. Genau genommen handelte es sich um einen Workshop zu einem Forschungsprojekt, für das McCarthy 1955 die Fördermittel bei der Rockefeller Foundation eingeworben hatte.[4] In seinem Antrag erschien erstmals der Begriff „artificial intelligence“ („künstliche Intelligenz“). Aus diesem Grund gilt diese – später sogenannte – Dartmouth Conference seither als die Gründungsveranstaltung der Künstlichen Intelligenz-Forschung als einem akademischem Fachgebiet.

- Im Vergleich zu früher können wir beim Lernen inzwischen **sehr grosse Datenmengen** und **viel mehr Rechenpower** nutzen.

KI Einsatzgebiete

Classification

Google Mail



Clustering

Google News



Collaborative Filtering
(Recommendation)

Customers Who Bought This Item Also Bought



Amazon Web Shop

KI Algorithmen / Techniken

- **Klassifizierung:** Google Mail verwendet eine Technik des maschinellen Lernens genannt Klassifizierung. Ob eine E-Mail Spam ist oder nicht, wird anhand von Absender, Empfänger, Betreff und Nachrichtentext festgelegt. Die Klassifizierung nimmt eine Reihe von bekannten Daten und lernt, wie man neue Datensätze auf der Grundlage dieser Informationen klassifiziert.
- **Clustering:** Google News verwendet eine Technik namens Clustering. News-Artikel werden in verschiedene Kategorien, basierend auf Titel und Inhalt, unterteilt. Clustering-Algorithmen entdecken Gruppierungen, die in diesen Sammlungen von Daten auftreten.
- **Collaborative Filtering:** Amazon verwendet eine Technik des maschinellen Lernens, Collaborative Filtering genannt (gemeinhin als Empfehlung genannt). So empfiehlt Amazon Nutzern Produkte auf der Grundlage ihrer Geschichte und Ähnlichkeit mit anderen Benutzern.

Weitere KI Einsatzgebiete

- Spracherkennung
- Spam-Filter
- Gesichtserkennung
- Personalisierten Online-Werbung
- Chatbots

Zusammenfassung

- Im Allgemeinen bezeichnet *künstliche Intelligenz* oder *KI* den Versuch, eine menschenähnliche Intelligenz nachzubilden.
- KI basiert auf Maschinellern Lernen, (Lern-)Daten und Algorithmen.
- Im Vergleich zu früher können wir beim Lernen inzwischen **sehr grosse Datenmengen** und **viel mehr Rechenpower** nutzen.

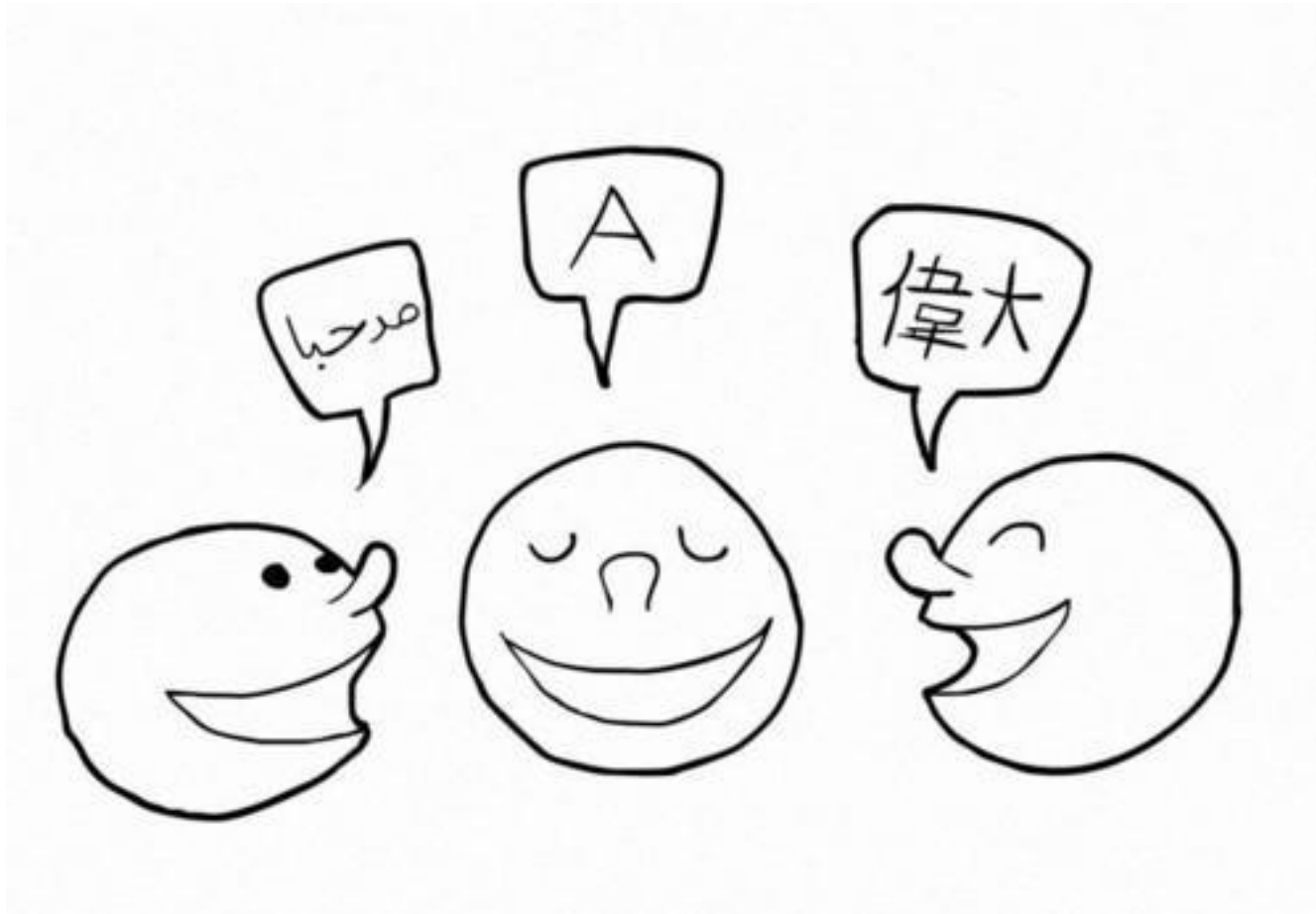
Veranstaltungsreihe

- 1.3.17: Warum die Digitalisierung Jobs killt!
- 15.3.17: Bit Data und künstliche Intelligenz – zwei Puzzleteile, die passen
- **5.4.17: Internet der Dinge – die Digitalisierung nimmt ihren Lauf**
- **3.5.17: Vom Umfang mit Containern in der Informatik**
- **17.5.17: Wähle was du brauchst – IT-Architektur «on demand»**
- Jeweils von 12:30 bis 13:15 Uhr

Kurse

- Digitalisierung
 - <https://www.eb-zuerich.ch/angebot/digitale-transformation.html>
- Internet der Dinge (IoT)
 - <http://kurs.eb-zuerich.ch/is95>
 - <http://kurs.eb-zuerich.ch/is96>
 - <http://kurs.eb-zuerich.ch/is98>
- Big Data, Künstliche Intelligenz, Machine Learning
 - <https://www.eb-zuerich.ch/angebot/big-data-ueberblick.html>
- Infrastructure as Code
 - <https://www.eb-zuerich.ch/angebot/infrastructure-as-code.html>
- Docker
 - <https://www.eb-zuerich.ch/angebot/docker.html>

Fragen



Danke für Ihre Aufmerksamkeit!

Kontakt:

EB Zürich

Bildungszentrum für Erwachsene BiZE

Riesbachstrasse 11

8090 Zürich

Telefon 0842 843 844

Fax 044 385 83 29

E-Mail lernen@eb-zuerich.ch

E-Mail marcel.bernet@ch-open.ch

