

1. Data Science in Gesellschaft und Unternehmen

2. Was ist Data Science?

3. Data Science Projekte in der Praxis

4. Stolpersteine und Lessons Learned

5. Ein Schlachtplan

Mein Ziel für die nächsten 40 Minuten

Was will ich?

Begeisterung für ein spannendes Thema wecken

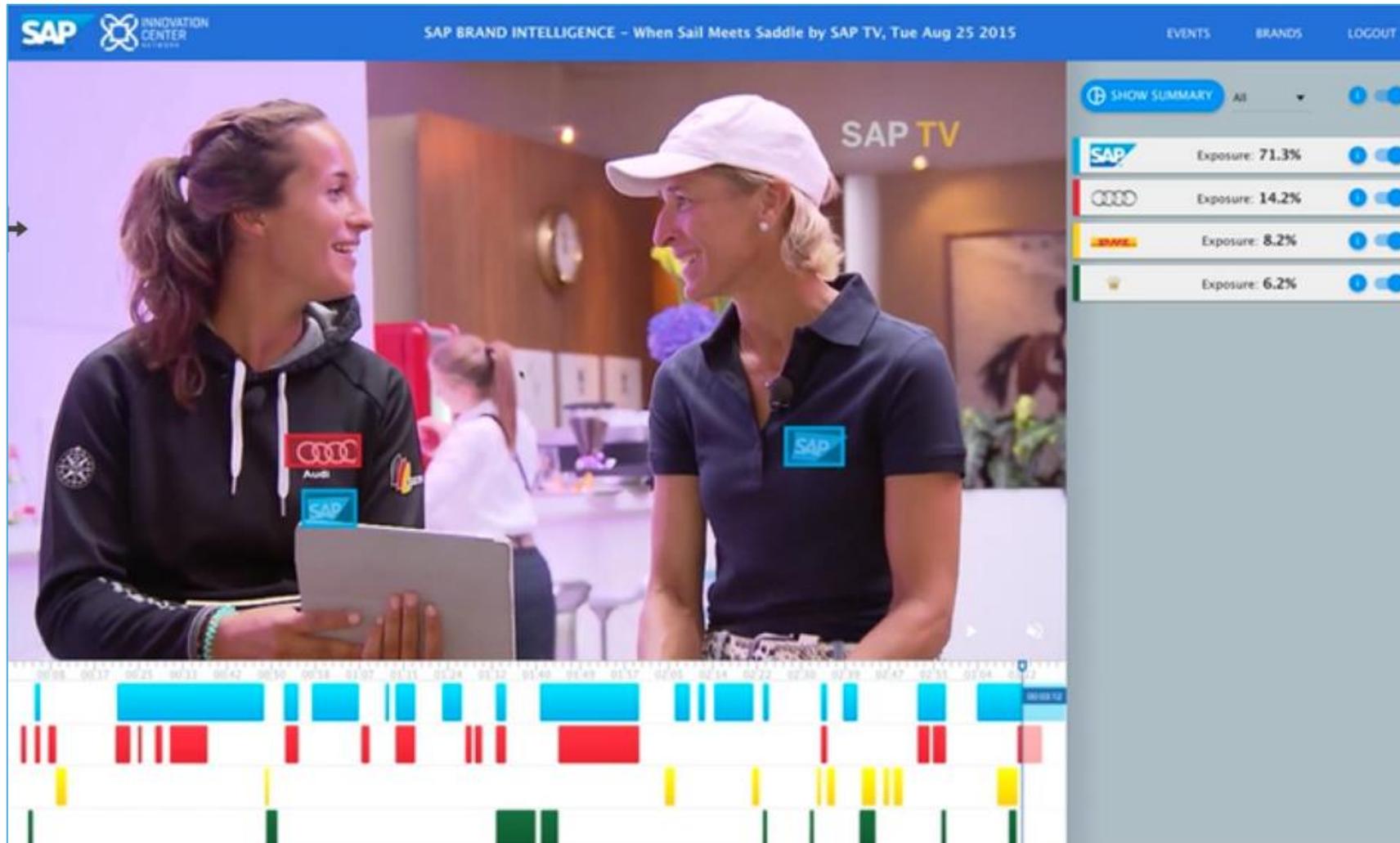
Viele praktische Aspekte vorstellen, wenig Theorie

Realistisches Bild vermitteln, keine Schönfärberei

Data Science in der Autoindustrie



Data Science im Marketing



Quelle: SAP

Data Science im Bereich der Nachrichtendienste



Quelle: <https://nsa.gov1.info/utah-data-center/utah-data-center.html>

Treiber für Data Science: Datengetriebene Geschäftsmodelle

✕ By combining weather forecasts with analytics, your insurance company could **text you a warning** before hail hits your area. Insurance companies pay

\$1 billion

in claims every year for vehicles damaged by hail.*

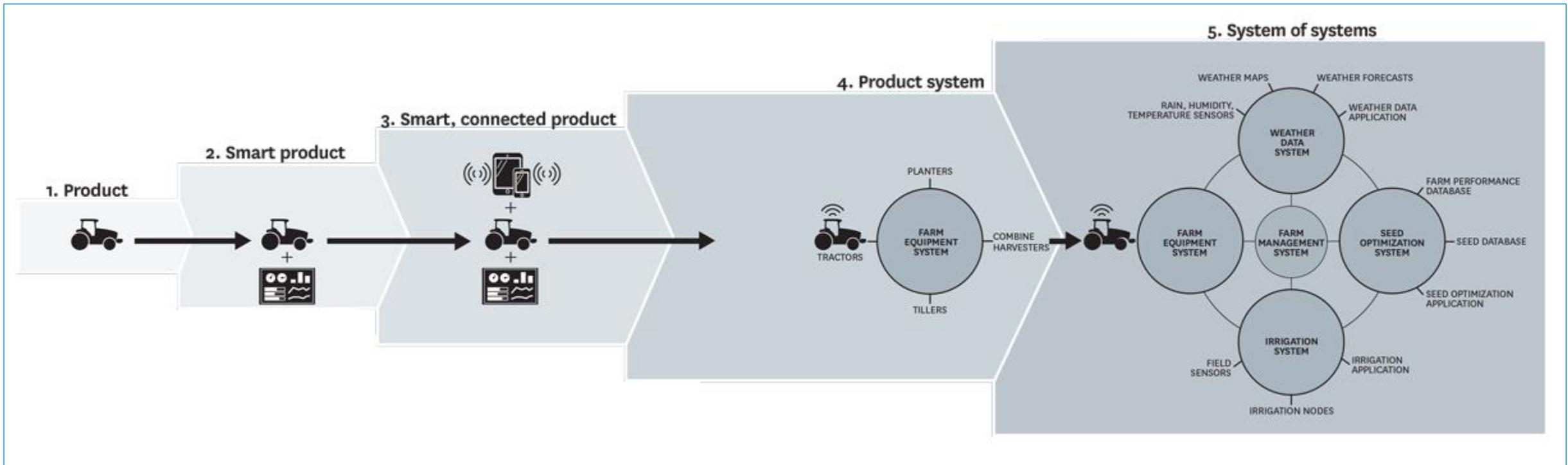


*Insurance Information Institute, April 2012

ibm.com/ibmandweather

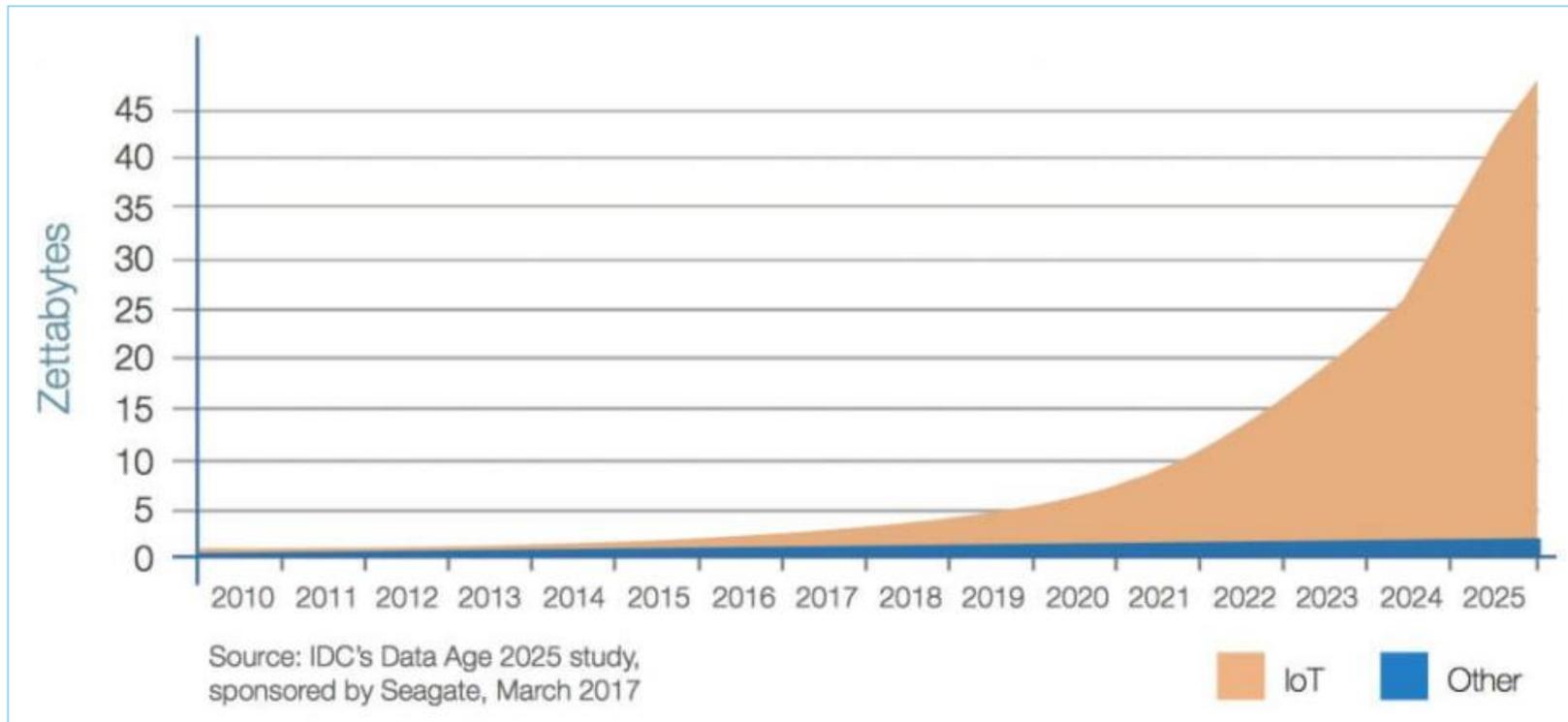
IBM | The Weather Company

Treiber für Data Science: Intelligente Produkte



Quelle: Harvard Business Review, November 2014, How Smart, Connected Products Are Transforming Competition

Treiber für Data Science: Enorme Datenmengen durch das IoT



1. Data Science in Gesellschaft und Unternehmen

2. Was ist Data Science?

3. Data Science Projekte in der Praxis

4. Stolpersteine und Lessons Learned

5. Ein Schlachtplan

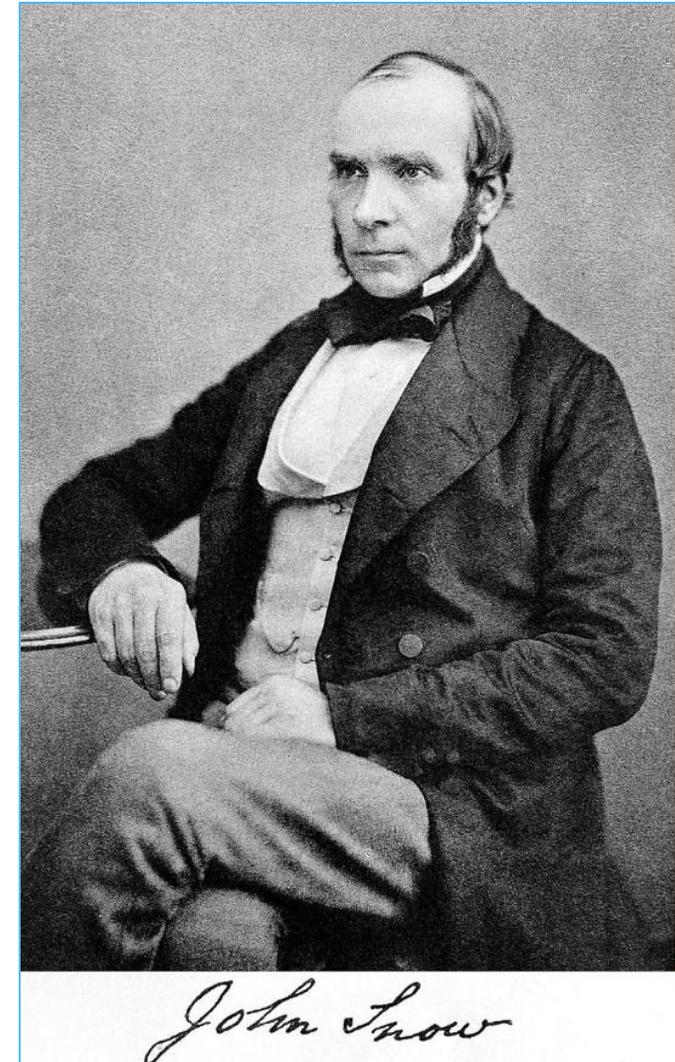
Data Science und die 3 Stufen der Digitalisierung

	1. Stufe (1970-1990) PC & Netzwerke	2. Stufe (1990-2010) Internet & Smartphone	3. Stufe (2010-2030) Internet der Dinge & KI
Meilensteine	1971 E-Mail (Tomlinson) 1973 Netzwerk (Metcalf) 1976 PC (Jobs/Wozniak) 1981 Betriebssystem (Gates)	1990 WWW (Berners-Lee) 1998 Search Engine (Brin/Page) 2007 Cloud (Houston/Ferdowsi) 2007 Smartphone (Jobs)	2011 IPv6 2011 Autonomes Fahren (Brin/Page) 2013 KI Assistenten (Bezos) 2015 IoT Security (Miller/Valasek)
Daten- übertragung	C-Netz (1G), D-Netz/E-Netz	UMTS (3G), LTE (4G)	5G, Terranova-Projekt
Datenanalyse	Kybernetik, Expertensysteme	Business Intelligence, Machine Learning	Neuronale Netze, Deep Learning, Data Science
Daten	Strukturierte Daten	Strukturierte Massendaten	Strukturierte Massendaten und polystrukturierte Daten

1854 Cholera Epidemie, London



Quelle: wikipedia.de

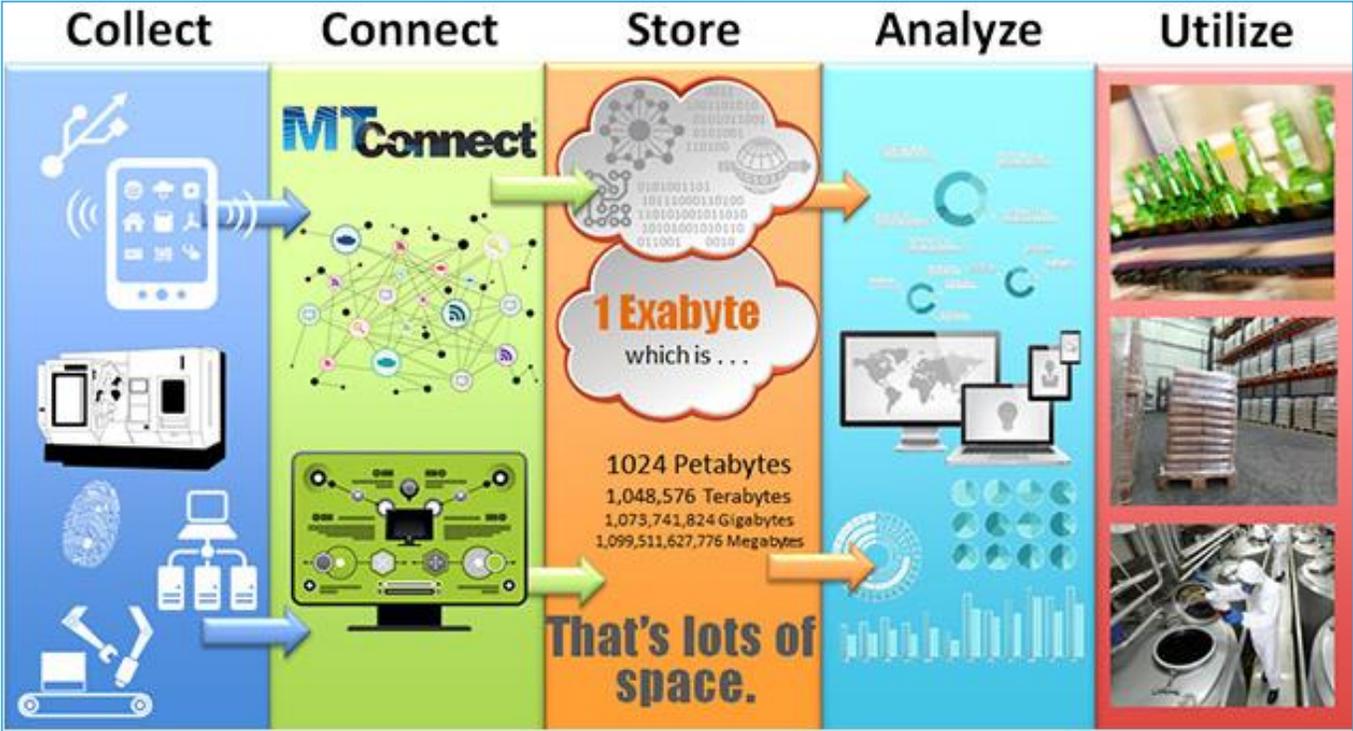


Quelle: wikipedia.de

Data Science Modell

MEASURE

ASK

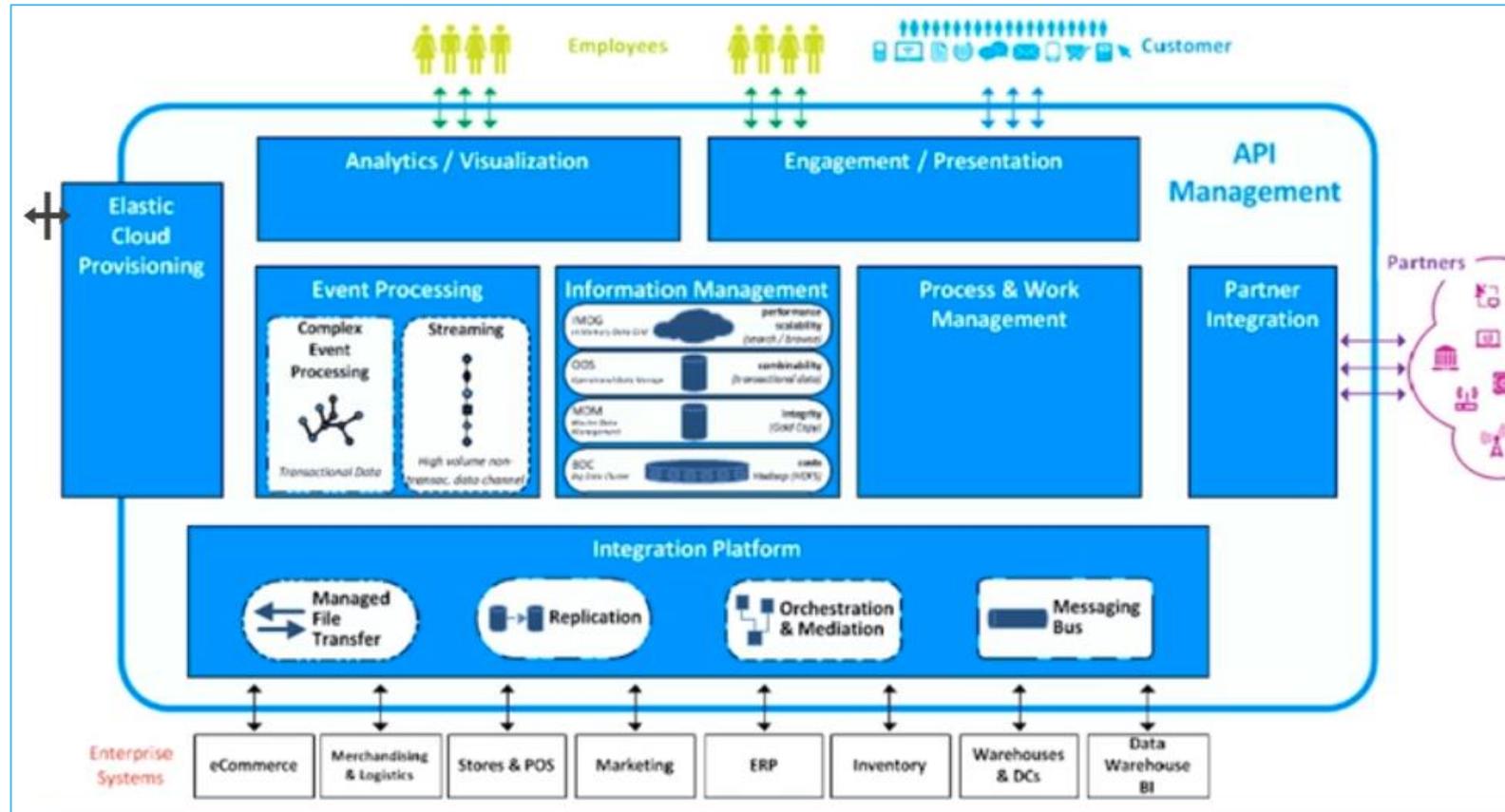


LEARN

Quelle: The Association For Manufacturing Technology, 2014

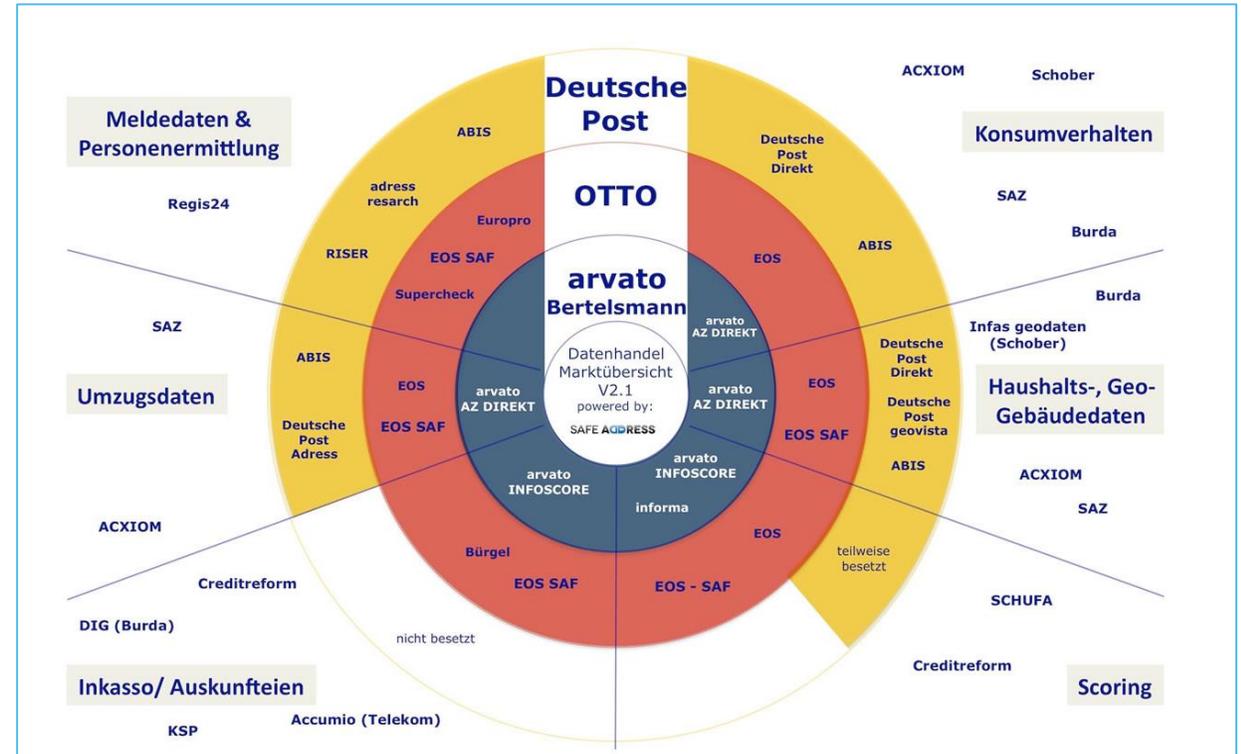
Collect/ Datenakquise

Unternehmensdaten, Partnerdaten und Kundendaten



Collect/ Datenakquise

Kosten-/Nutzenanalyse für zugekaufte Daten



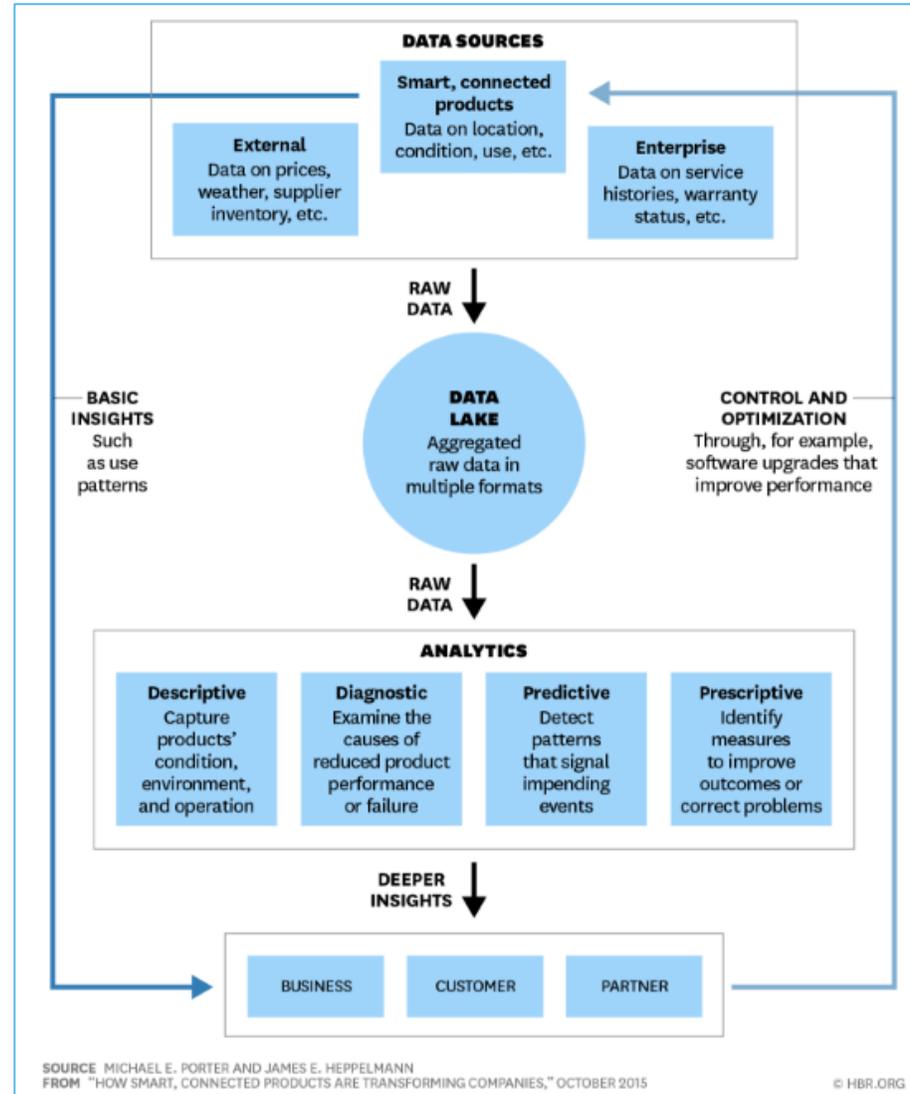
Quelle: safe-address.de

Store / Datenhaltung Cloud



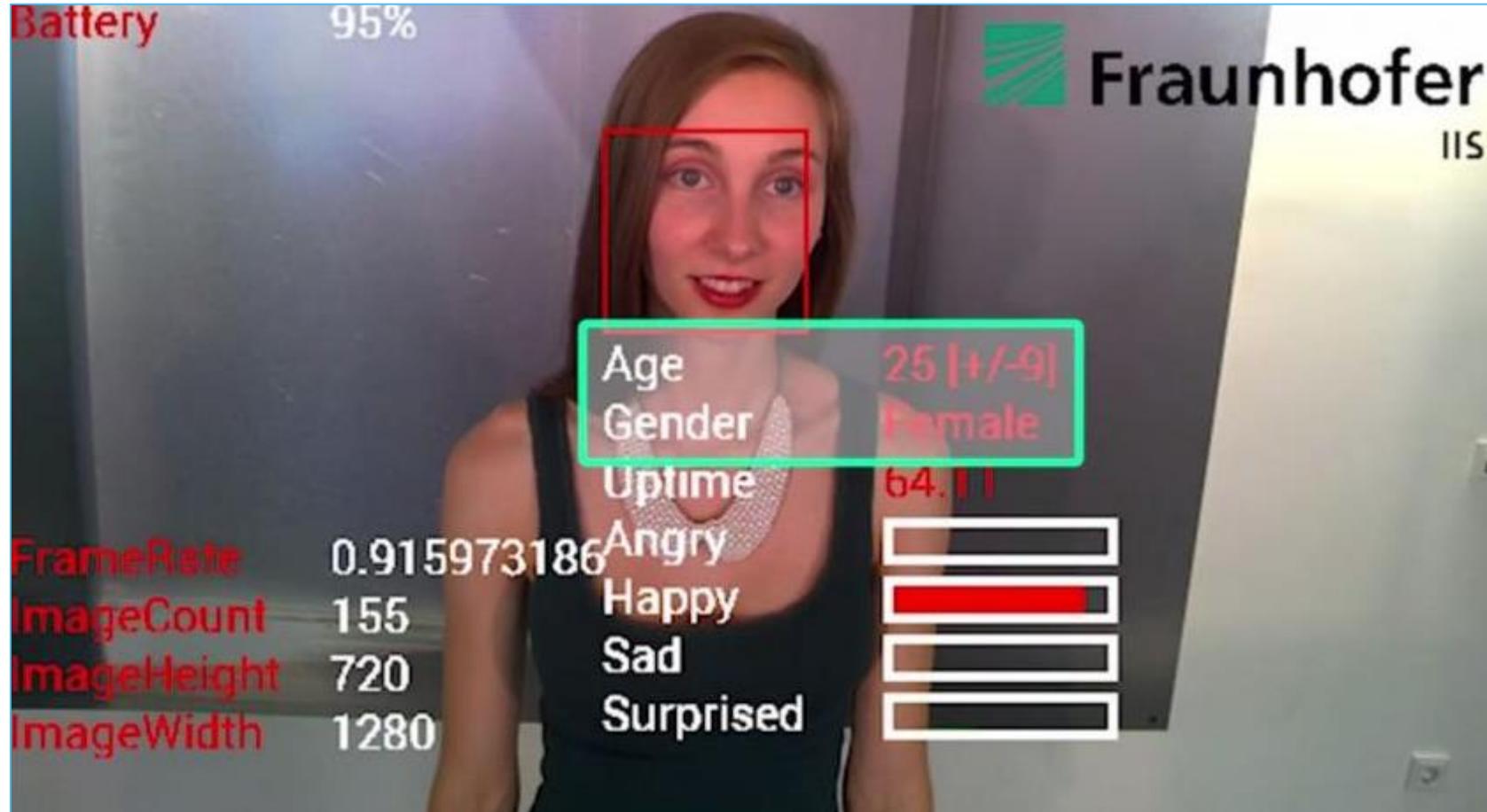
Store / Datenhaltung

Data Lake

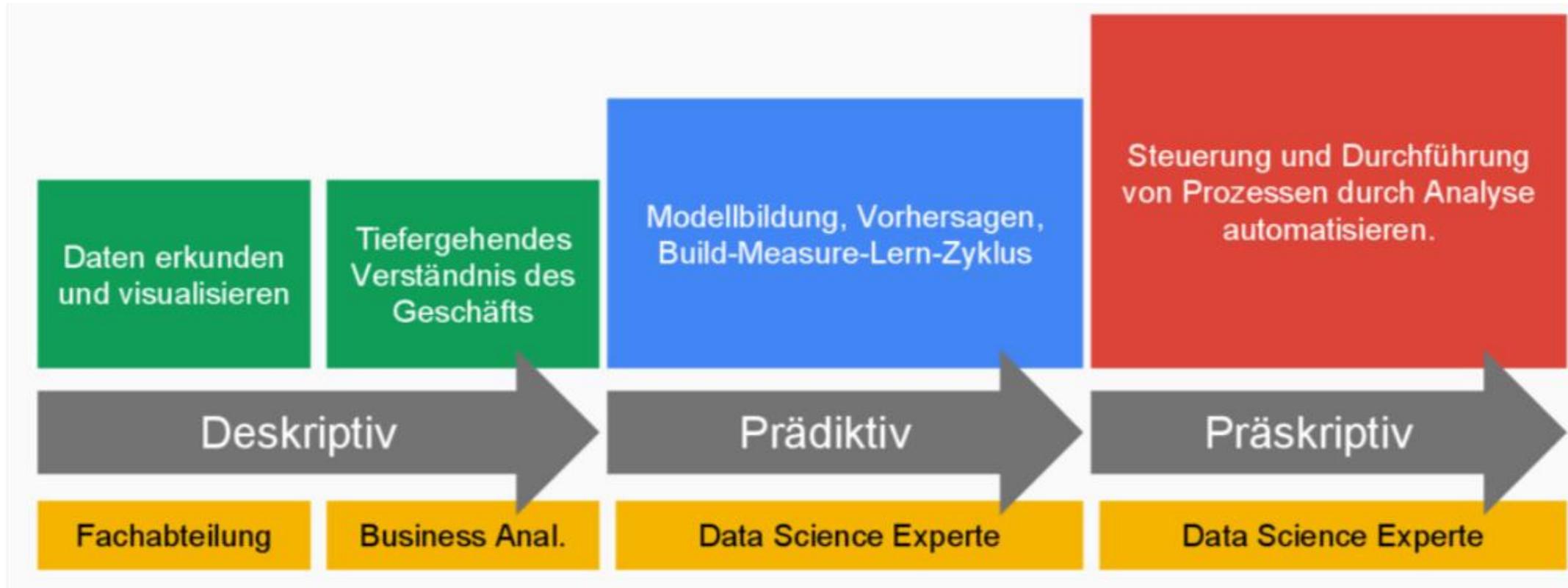


Analyse

Videodatenanalyse

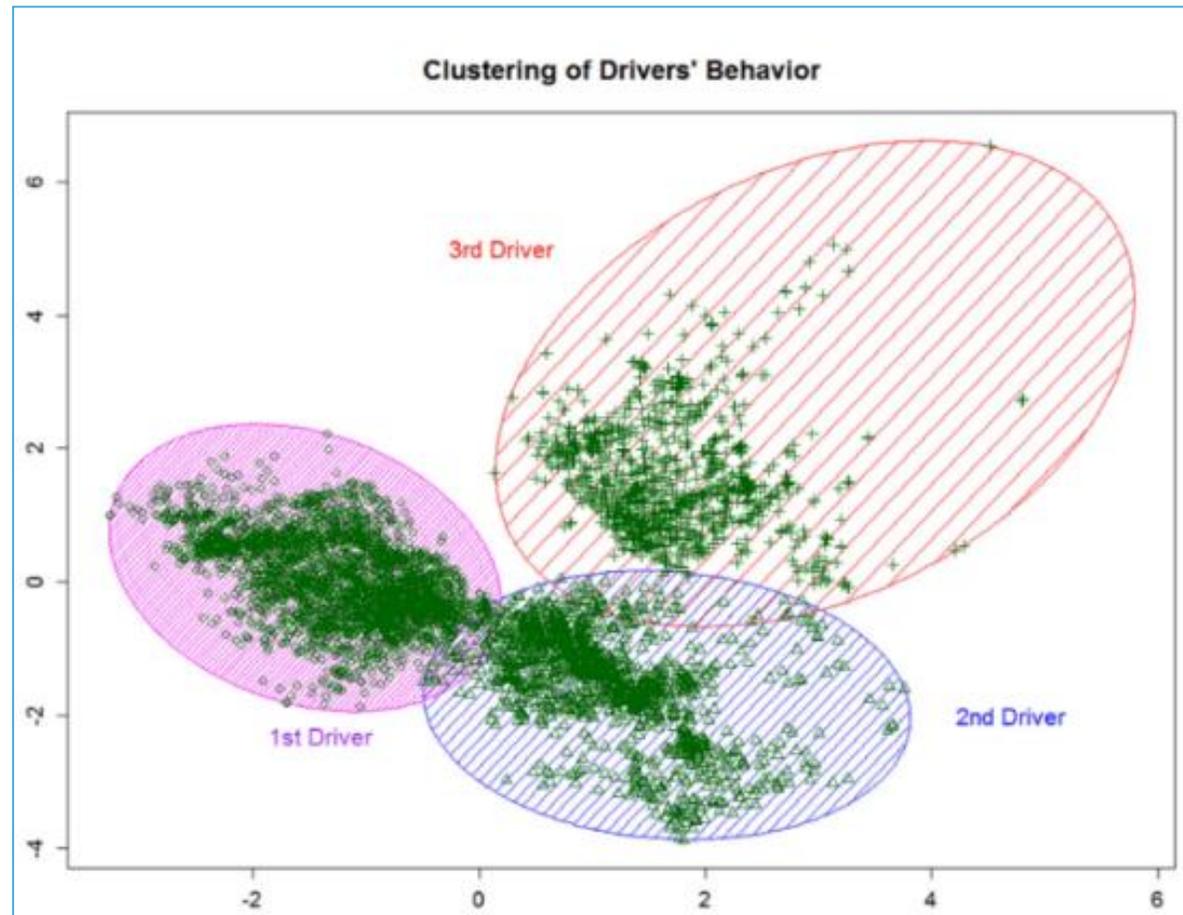


Analyse



Analyse

Beispiel Clustering



Quelle: Computerwoche, 10.03.1015: Mit Big Data zum Connected Car der Zukunft

1. Data Science in Gesellschaft und Unternehmen

2. Was ist Data Science?

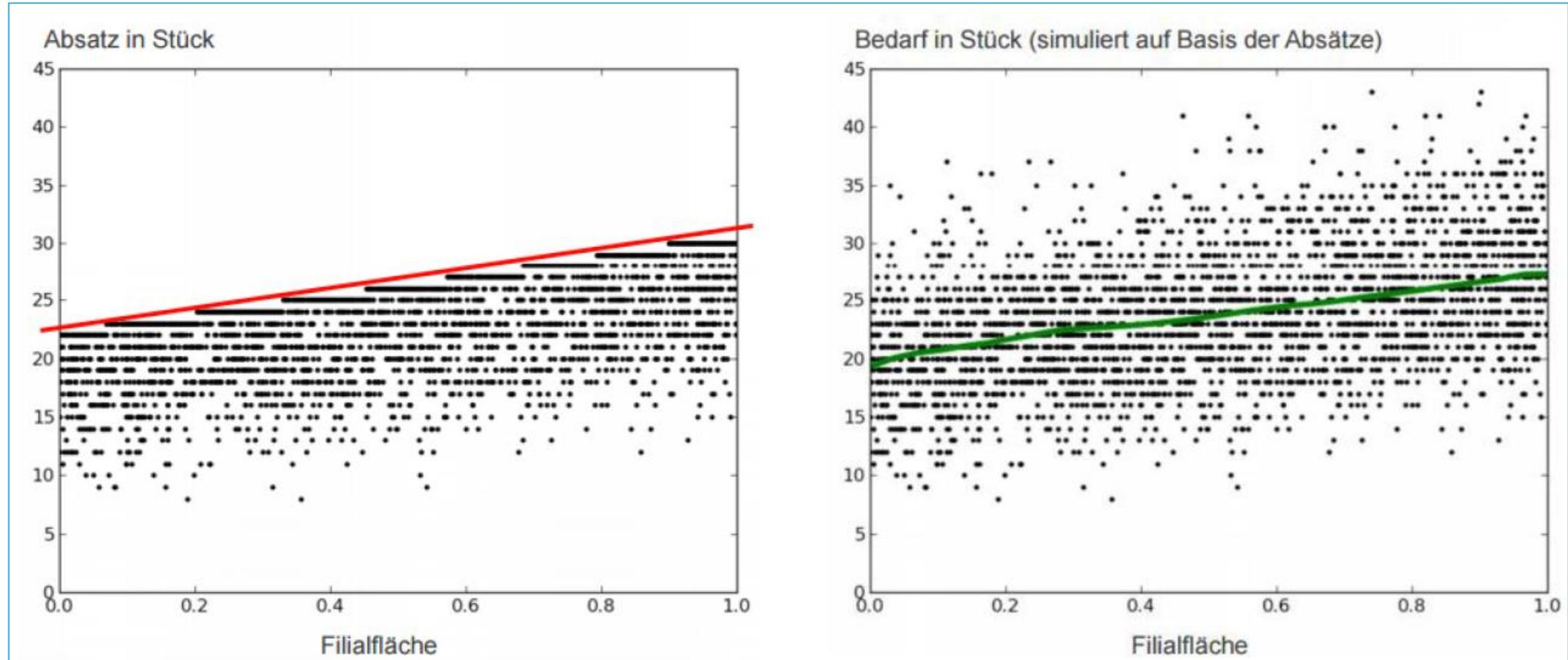
3. Data Science Projekte in der Praxis

4. Stolpersteine und Lessons Learned

5. Ein Schlachtplan

Projektziel: Out of stock Situationen im Handel vermeiden

Unternehmen: natsu



Projektziel: Automatisierte Artikel- und Retourenprognosen durch Datenanalyse

Unternehmen: Otto



Customer Case Study **HANDEL**

OTTO

Automatisierte Entscheidungen entlang des Produktlebenszyklus sparen OTTO Millionen

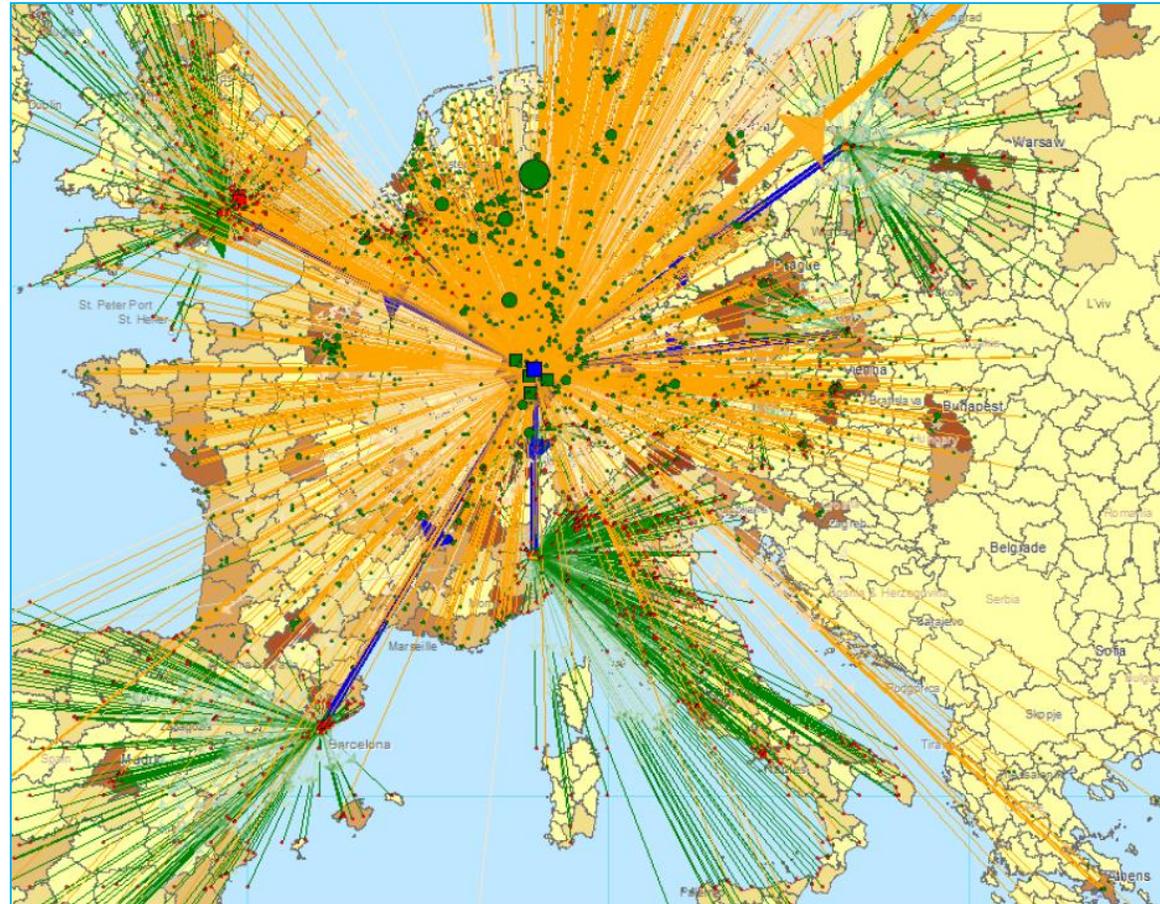
OTTO ist Vorreiter auf dem Gebiet Smart Data im deutschen Handel

blue yonder
Forward looking. Forward thinking.

Quelle: Otto

Projektziel: Investitionsvorbereitung durch Datenanalyse

Unternehmen: Hansgrohe



Quelle: Big Data- Potenzial für den Controller; Ideenwerkstatt im ICV 2014, Seite 18

Projektziel: Identifikation neuer Geschäftspotentiale

Unternehmen: Autoindustrie

	Fahrer		Fahrzeug		Umgebung	
Untergruppe	Fahrer-bezogene Daten	Nutzungs-bezogene Daten	Wartungs-relevante Daten	Situations-bezogene Daten	Verkehrs-daten	Sonstige Daten
Daten sammeln (Beispiele)	70 Parameter zum Fahrverhalten	GPS Tracking	Felddaten	Diverse Sensordaten (Bsp. Airbag)	Vibrations-sensoren	Regen- und Lichtsensoren
Daten analysieren (Beispiele)	Unsupervised Learning	Destination prediction	Supervised learning	Anomaly detection	Anomaly detection	Supervised Learning
Geschäftsmodell (Beispiele)	Anzahl der Fahrer detektieren	Targeted Advertising im Auto	Instand-haltungs-management	Unfall-management	Straßen-zustands-meldungen	Abschätzung Strom-einspeisung

Beispiele für weitere Projekte

Bereich	Fragestellung
- Kundenpflege	- Welcher Kunde wird bald seinen Vertrag kündigen?
- Marketing	- Welchem Kunden soll eine bestimmte Werbung geschickt werden?
- Bonitätsprüfung	- Wie hoch ist das aktuelle Ausfallrisiko des Kunden?
- Patentanalyse	- Welches weltweit neu angemeldete Patent kann Auswirkungen auf unser Geschäft haben?

1. Data Science in Gesellschaft und Unternehmen

2. Was ist Data Science?

3. Data Science Projekte in der Praxis

4. Stolpersteine und Lessons Learned

5. Ein Schlachtplan

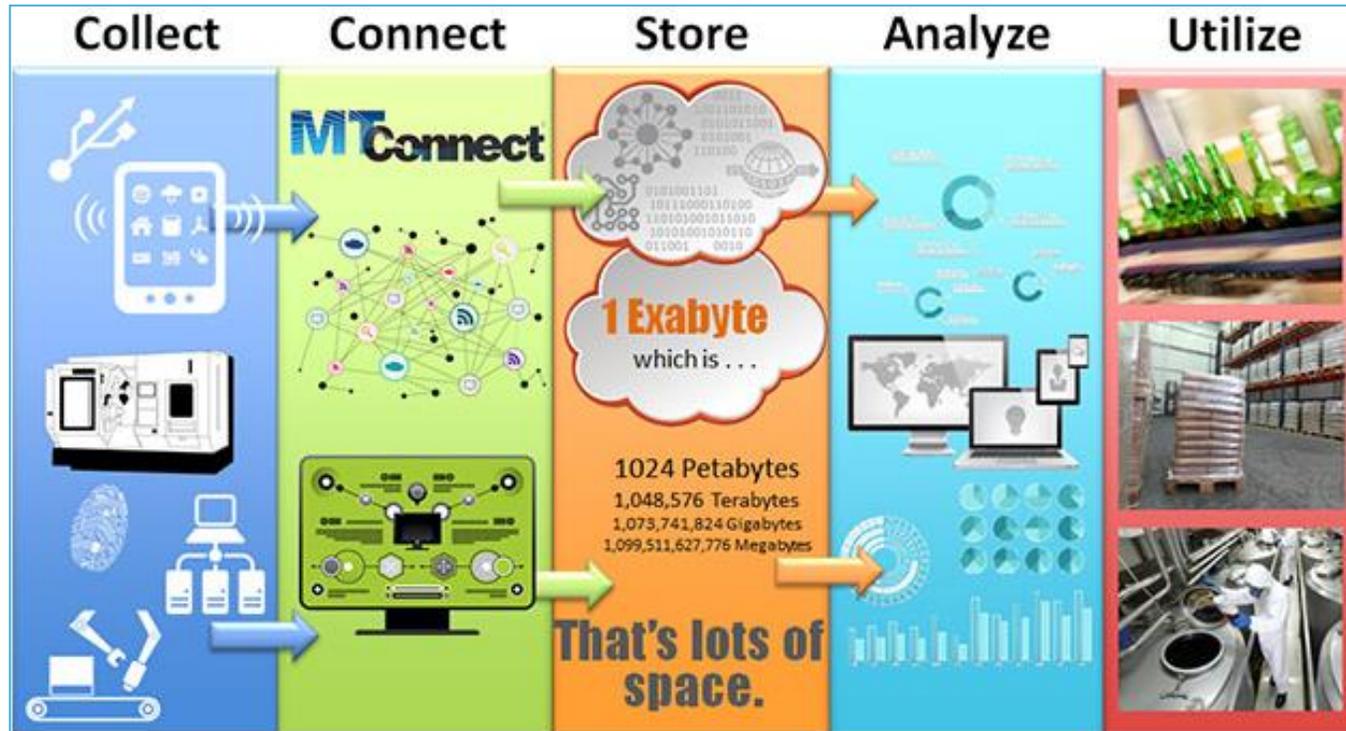
Stolperstein 1

Mangelnde Datenqualität oder zu wenig Daten



Stolperstein 2

Keine ganzheitliche Betrachtung des Prozesses



Stolperstein 3

Fehlendes Personal

MODERN DATA SCIENTIST

Data Scientist, the sexiest job of 21st century requires a mixture of multidisciplinary skills ranging from an intersection of mathematics, statistics, computer science, communication and business. Finding a data scientist is hard. Finding people who understand who a data scientist is, is equally hard. So here is a little cheat sheet on who the modern data scientist really is.

MATH & STATISTICS

- ☆ Machine learning
- ☆ Statistical modeling
- ☆ Experiment design
- ☆ Bayesian inference
- ☆ Supervised learning: decision trees, random forests, logistic regression
- ☆ Unsupervised learning: clustering, dimensionality reduction
- ☆ Optimization: gradient descent and variants

DOMAIN KNOWLEDGE & SOFT SKILLS

- ☆ Passionate about the business
- ☆ Curious about data
- ☆ Influence without authority
- ☆ Hacker mindset
- ☆ Problem solver
- ☆ Strategic, proactive, creative, innovative and collaborative



PROGRAMMING & DATABASE

- ☆ Computer science fundamentals
- ☆ Scripting language e.g. Python
- ☆ Statistical computing package e.g. R
- ☆ Databases SQL and NoSQL
- ☆ Relational algebra
- ☆ Parallel databases and parallel query processing
- ☆ MapReduce concepts
- ☆ Hadoop and Hive/Pig
- ☆ Custom reducers
- ☆ Experience with xaaS like AWS

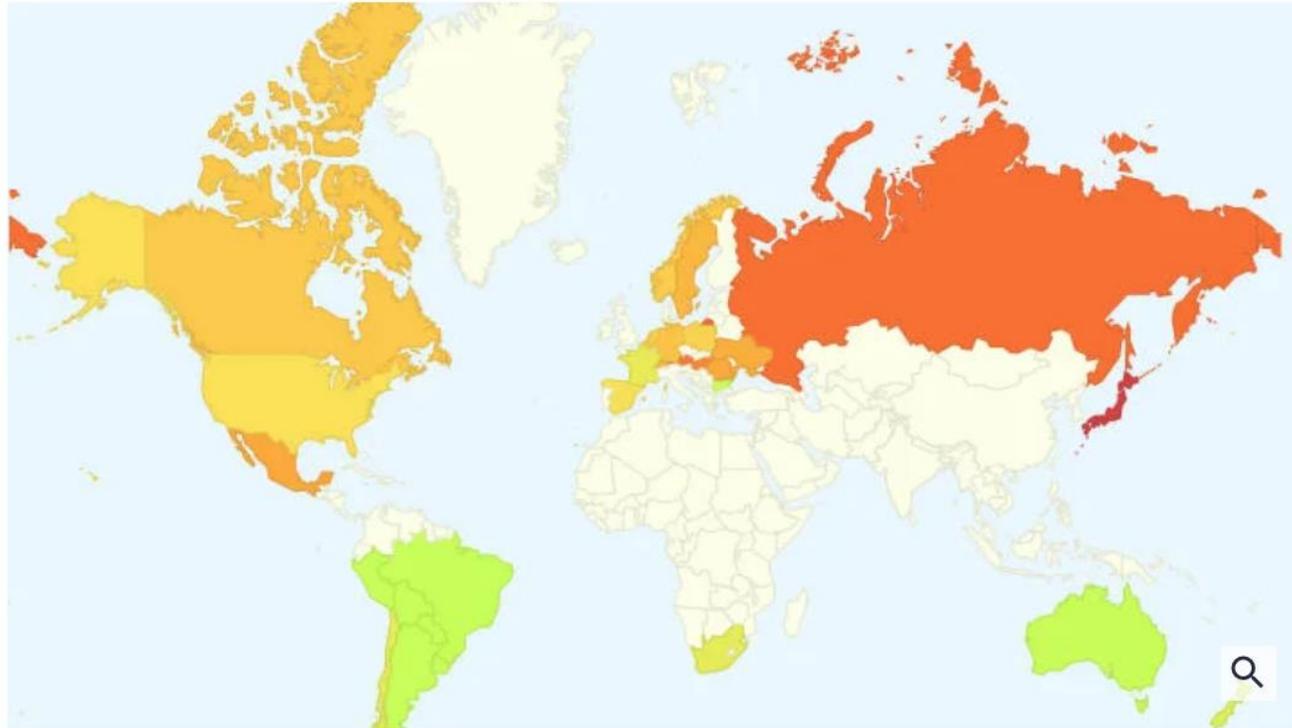
COMMUNICATION & VISUALIZATION

- ☆ Able to engage with senior management
- ☆ Story telling skills
- ☆ Translate data-driven insights into decisions and actions
- ☆ Visual art design
- ☆ R packages like ggplot or lattice
- ☆ Knowledge of any of visualization tools e.g. Flare, D3.js, Tableau

Lessons learned 1

Es gibt keine Garantie für den Projekterfolg

Google versagt bei Grippe- Vorhersagen



Grippehäufigkeit auf der Weltkarte von "Google Flu Trends" (Foto: Google Flu Trends)

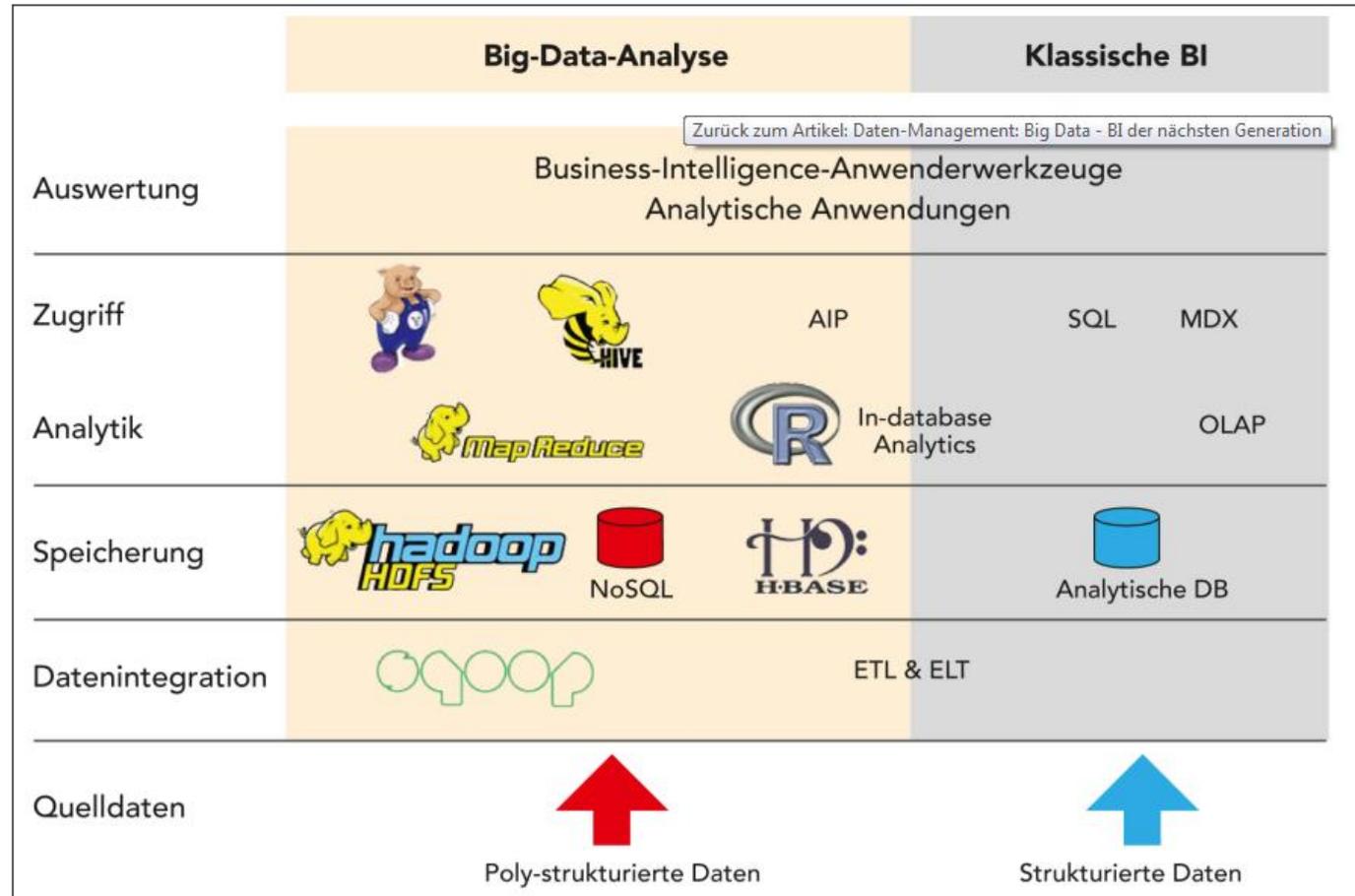
Lessons learned 2

Auch bei Projekterfolg gibt es immer noch Fehlentscheidungen des Algorithmus

Analysergebnis	Kunde mit guter Bonität	Kunde mit schlechter Bonität
Tatsächliche Bonität		
Gut	True Positive Richtige Entscheidung (Kunde erhält Kredit)	False Positive Entgangenes Geschäft (kein Kredit trotz guter Bonität)
Schlecht	True Negative Schlechte Entscheidung (erhöhtes Kreditausfallrisiko)	False Negative Richtige Entscheidung (Kunde erhält keinen Kredit)

Lessons learned 3

R, Python, NoSQL usw. → Fortbildung ist gut investiertes Geld



Quelle: Carsten Bange/ Timm Grosser in Computerwoche zu Big Data Analysen

1. Data Science in Gesellschaft und Unternehmen

2. Was ist Data Science?

3. Data Science Projekte in der Praxis

4. Stolpersteine und Lessons Learned

5. Ein Schlachtplan

Ein Schlachtplan

Wie setze ich in meinem Unternehmen ein Projekt erfolgreich um?

	Abtauchen	Weichen stellen	Marathon
			
Stichwort	Analyse	Strategie	Zunächst Pilotprojekt, danach Ausbau
Arbeitspakete (Beispiele)	<ul style="list-style-type: none">- Welche konkreten Ansatzpunkte für Data Science gibt es im Unternehmen?- Welche Daten sind vorhanden?- In welcher Qualität?- Welche Datenquellen?	<ul style="list-style-type: none">- Erwartungshaltung Management- Use Case Definition- Datenhaltung: Was wird gespeichert?- Datenqualität: Wie sollen die Daten gespeichert werden?- Cloud/ Data Lake- Festlegung Rolle Finanzen/Controlling	<ul style="list-style-type: none">- Aufbau Projektstruktur- Festlegung Key User- Aufgabenaufteilung intern/extern- Schulung- Datenakquise, Datenbereinigung, Datenintegration- Modellbildung
Ergebnis	<ul style="list-style-type: none">- Potential von Data Science im eigenen Geschäftsumfeld erkennen- Analysepapier / Chancen/Risiken	<ul style="list-style-type: none">- Zielfestlegung- Use case- Datenstrategie	<ul style="list-style-type: none">- Pilotprojekt: Erstes Data Science Modell- Danach Umsetzung Use Case und Implementierung

Ein Schlachtplan

	Interim	Defend	Attack
Hauptziel	Überbrückung einer zeitlich begrenzten Vakanz	Absichern CO/FI Funktion im Unternehmen	Aktivere Wahrnehmung der CO/FI Funktion im Unternehmen
Praktisches Projektbeispiel	Interim Leitung Controlling (6-12 Monate)	Verbesserung der Daten- und Reportingqualität	Data Science Projekt (initiiert durch Controlling)

Kontakt Daten

Harry Barth

E-Mail: hbarth@hotmail.de

Mobil: 0173 1049 656

