

Data Science für Unternehmen – Potenziale, Herausforderungen und Basiswissen

Fraunhofer Innovationszentrum
»Digitale Transformation der Industrie«

DI Maximilian Ulrich

Fraunhofer Innovationszentrum »Digitale Transformation der Industrie«

Profil Maximilian Ulrich



Personal Data

Name: Maximilian Ulrich

Qualification: MSc (TU Wien, Wirtschaftsingenieurwesen-Maschinenbau)
Data Scientist Specialized in Deep Learning (zertifiziert)
Data Scientist - Advanced Level (zertifiziert)

Position: wissenschaftlicher Mitarbeiter

Competences

Fachgebiete:

Industrial Data Analytics, Data Science,
Digitale Transformation, Simulation

Branchen:

Lebensmittelindustrie,
Kunststoffindustrie,
Konsumgüter, Fahrzeugbau

Lehre/Forschung:

FFG Project ASPeCT
Data Science Workshops

Projektbeispiele (Auszug)

- Vorhersage des Produktionsaufwands von Konsumgütern zur Unterstützung der Produktkalkulation in der Frühphase der Produktentwicklung
- Web mining – Nutzung externer Daten zur Entscheidungsunterstützung
- Erstellung einer automatisierten Produktionsauslastungsvorschau unter Ausnützung vorhandener Absatzprognosen
- Explorative Analyse von Produktionsdaten zur Steigerung des Prozessverständnisses
- Tool zur Erstellung von Warenkorbvorschlägen auf Basis historischer Bestellungen

Data Science für Unternehmen

Potenziale

Steuerung und Optimierung von Produktions- und Logistikprozessen



Duck Box

Intelligente und vorausschauende Wartung



Inventory Dragon

Bestandsplanung + Absatzprognosen



Integrierte Qualitätskontrolle



Markt/Umfeld-Analyse



Autonome und kollaborative Systeme

Oliver Sved_fotolia

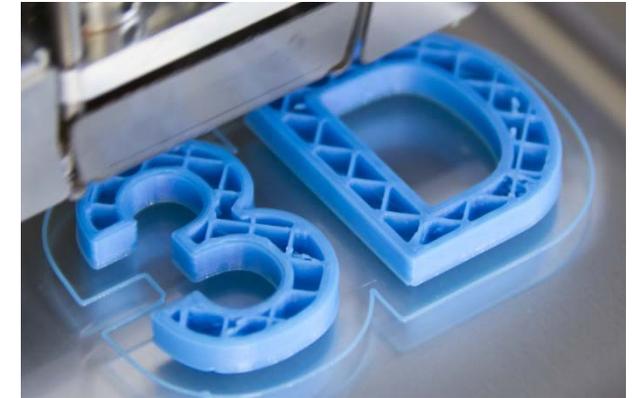
Data Science für Unternehmen

Potenziale

- **Strukturierte Auswertung von Daten**,
welche aktuell noch nicht über klassische BI-Lösungen ausgewertet werden
- **„Rapid-Prototyping“** für Analyselösungen



- **Entscheidungsunterstützung**
z.B. Bilderkennung in der Qualitätskontrolle
- **Automatisierung von Abläufen**
z.B. Automatisierung von Datenauswertungen
- **Personalisierung von Diensten**
z.B. personalisiertes CRM durch Klassifikation
- **Wissensgenerierung**
z.B. Erkennen unbekannter Zusammenhänge durch moderne Visualisierungen



Rapid-Prototyping für BI

<https://scoobe3d.com/>

Data Science für Unternehmen

Herausforderungen

■ Fragestellung

- Lässt sich Data Science auf diese Fragestellung sinnvoll anwenden oder wären andere Methoden (z.B. Lean) sinnvoller?

■ Datenschutz

- Sicherheit: je mehr Daten gespeichert werden, desto eher Ziel für einen Angriff
- Privatsphäre: dürfen die Daten überhaupt verwendet werden (DSGVO)?

■ Daten – grundlegende Voraussetzungen hinsichtlich der Datenqualität

- Information, die gesucht wird muss sich in den Daten verbergen!
- Daten lückenhaft?
- Referenzieren zwischen Datensätzen möglich bzw. korrekt? (z.B. Zeitstempel)

■ Datenzugang – gewünschte Daten verfügbar?

- Oft Kooperation zwischen Abteilungen notwendig!

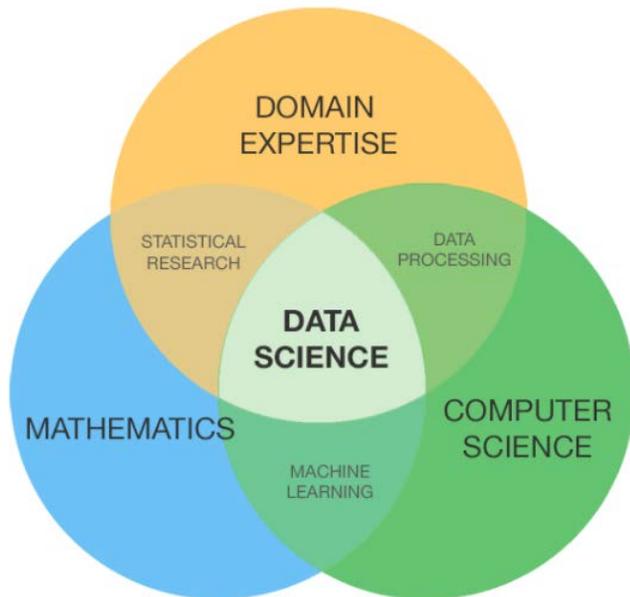
■ Analyse

- Nicht jede Analyse ist sinnvoll!
- Kausalität \neq Korrelation

Data Science für Unternehmen

Basiswissen – Abgrenzung und Begriffe

Data Science als Teamsport



Big Data



wenn es mit Standard-Tools schwierig wird, mit Daten umzugehen



„Big“ ist für ein erfolgreiches Projekt nicht unbedingt erforderlich



Künstliche Intelligenz

Künstliche Intelligenz (KI) beschäftigt sich mit Methoden, die es einem Computer ermöglichen, solche Aufgaben zu lösen, die, wenn sie vom Menschen gelöst werden, Intelligenz erfordern

Data Science für Unternehmen

Basiswissen – Entwicklung

<http://www.extremetech.com/computing/171700-two-years-later-hdd-prices-settle-back-to-normal>

1. Preis und Verfügbarkeit von Speicherplatz



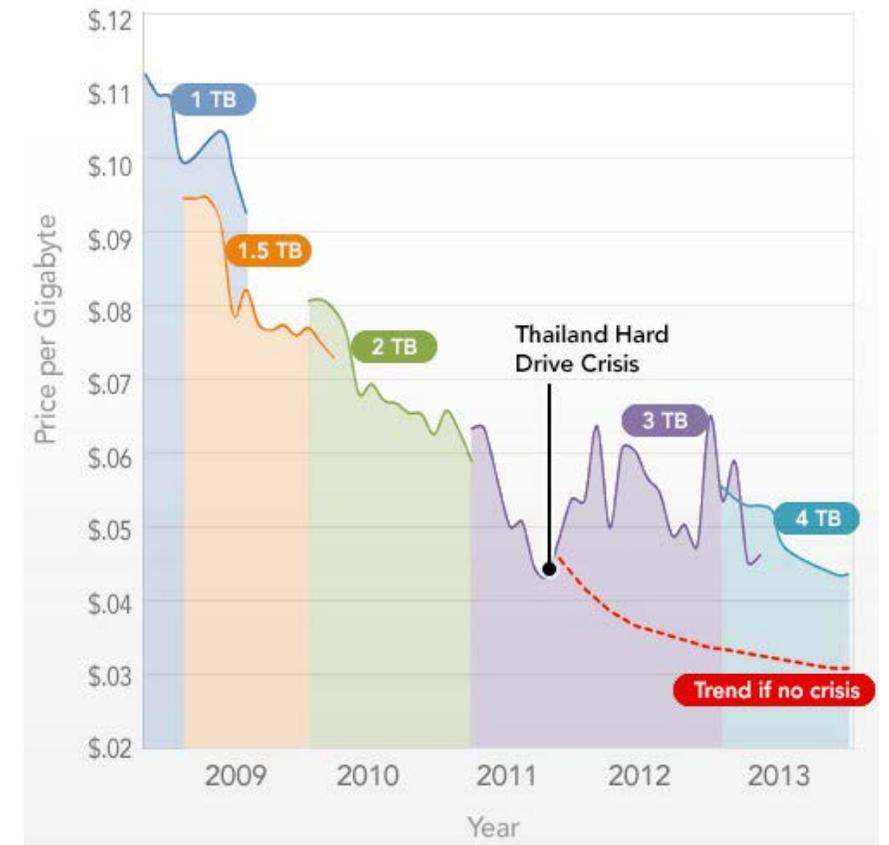
1.44 MB
1986



4TB
2019

Cost per GB for Hard Drives

Prices Backblaze paid for drives from 2009-2013



Data Science für Unternehmen

Basiswissen – Entwicklung

1. Preis und Verfügbarkeit von Speicherplatz
2. **Verfügbare Technologie und Algorithmen**



Data Science für Unternehmen

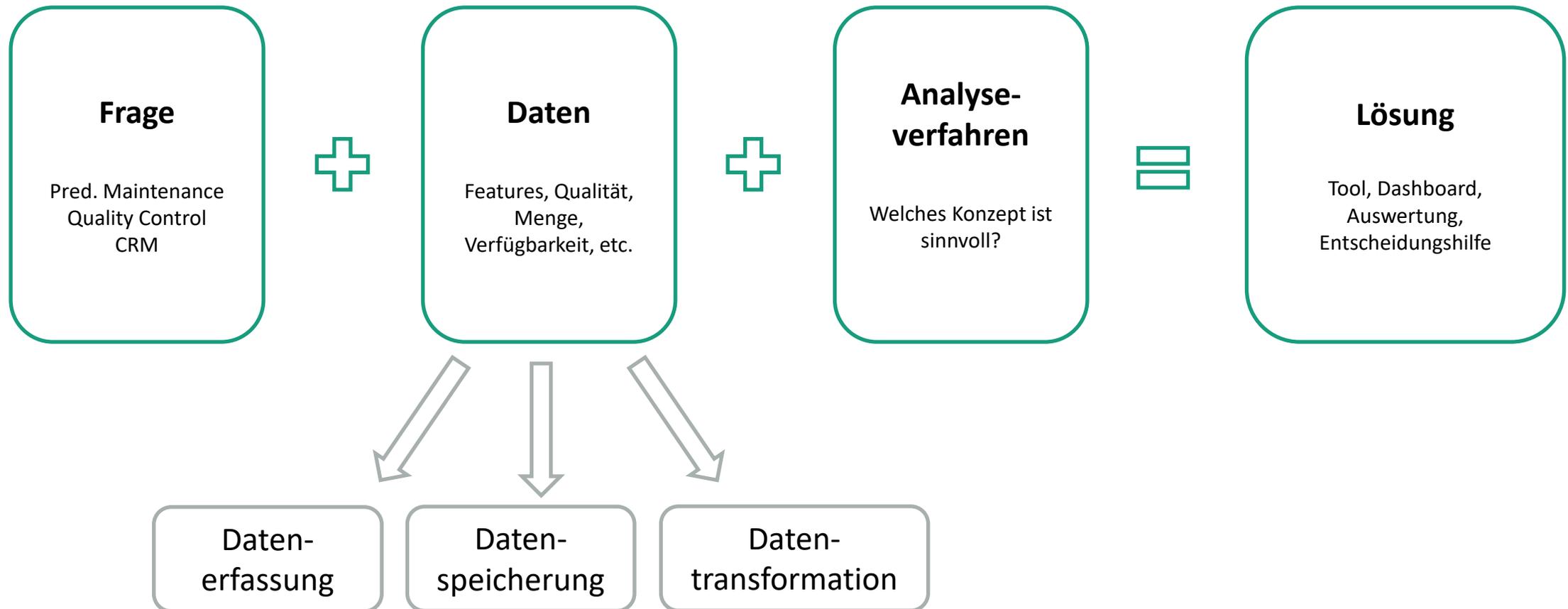
Basiswissen – Entwicklung

1. Preis und Verfügbarkeit von Speicherplatz
2. Verfügbare Technologie und Algorithmen
- 3. Neue Datenquellen und Fortschritte in der Sensortechnologie**



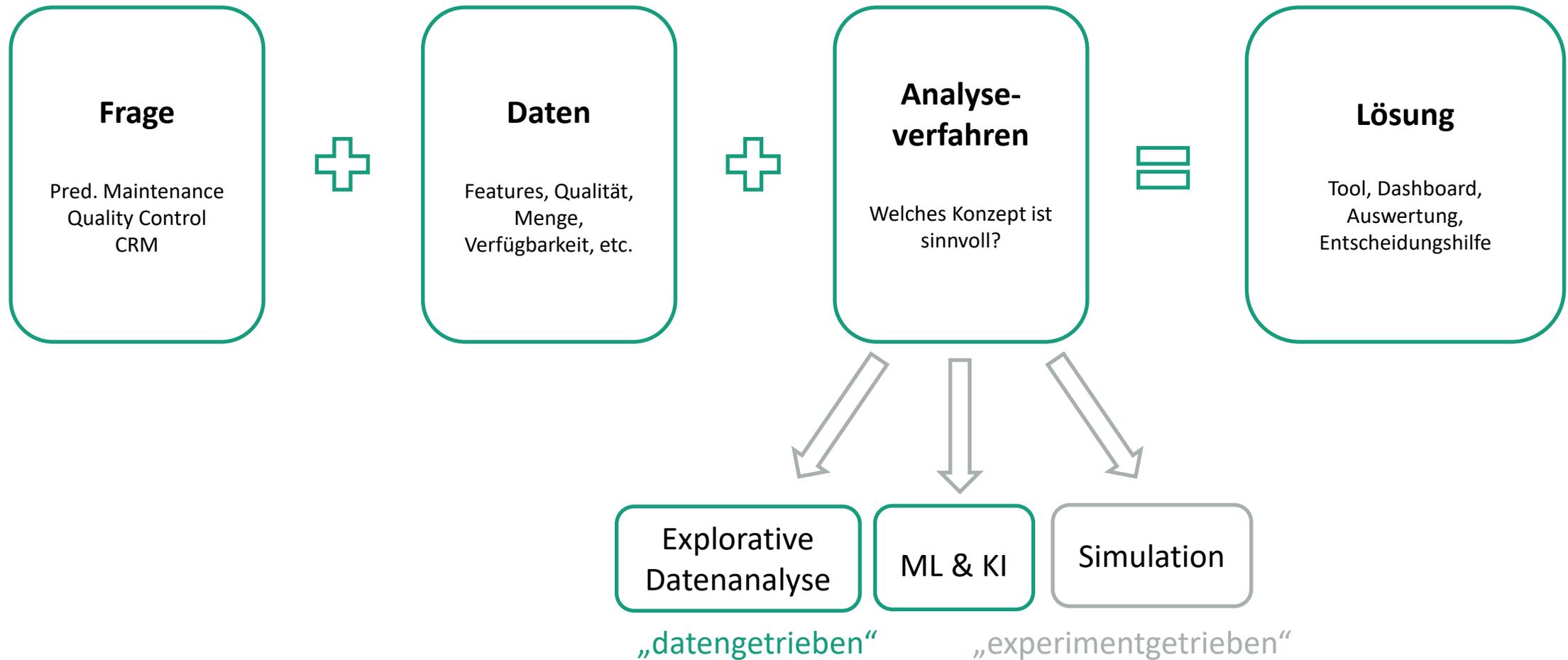
Data Science für Unternehmen

Basiswissen – Zentrale Elemente



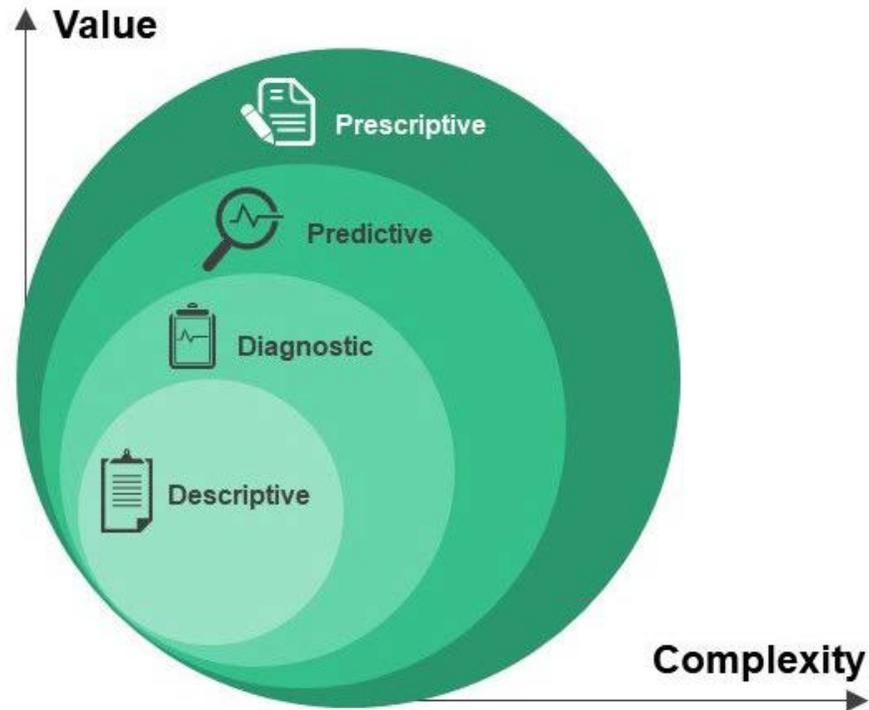
Data Science für Unternehmen

Basiswissen – Zentrale Elemente



Data Science für Unternehmen

Basiswissen – Zentrale Elemente



Welche Fragen lassen sich mithilfe von Daten beantworten?

Descriptive – Was ist passiert?

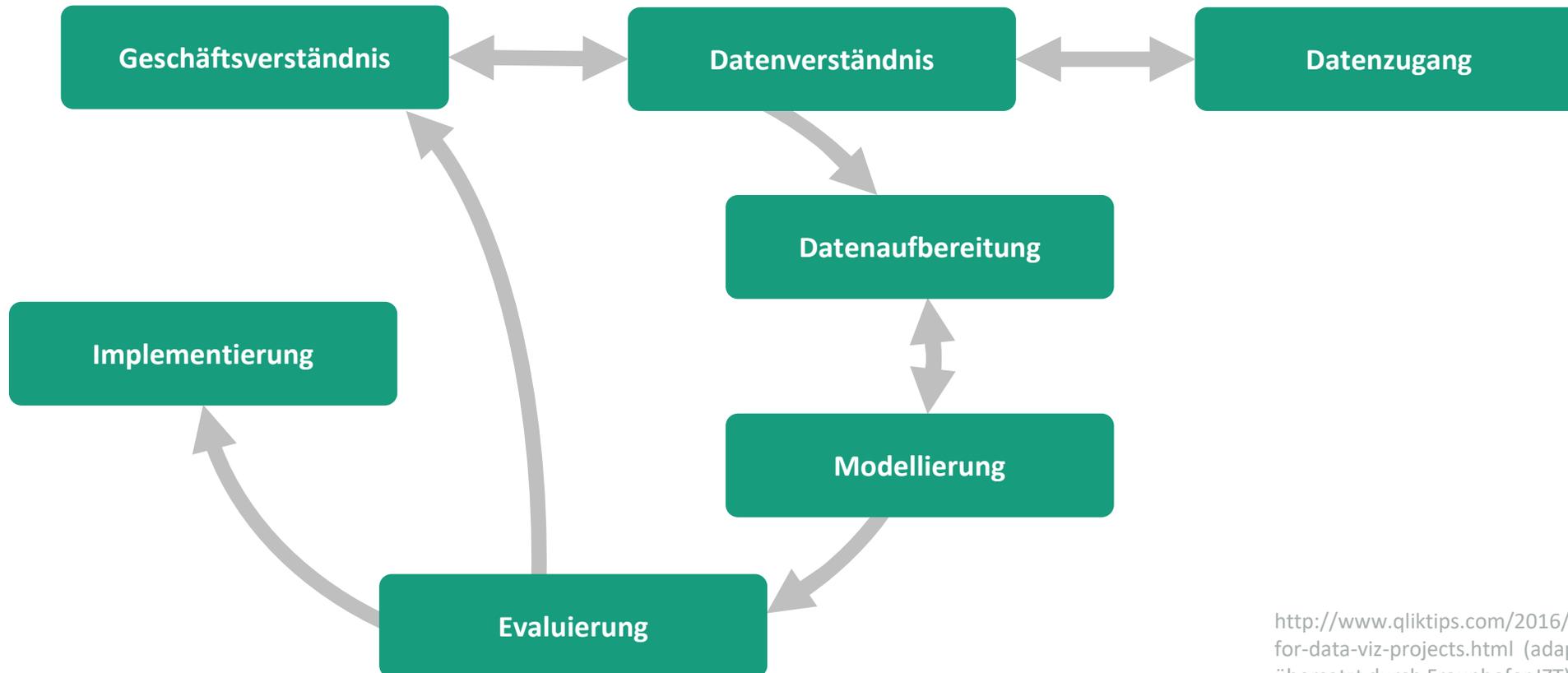
Diagnostic – Warum ist es passiert?

Predictive – Was wird passieren?

Prescriptive – Was soll ich tun?

Data Science für Unternehmen

Basiswissen – Strukturierte Vorgehensweise



<http://www.qliktips.com/2016/01/crisp-dm-for-data-viz-projects.html> (adaptiert und übersetzt durch Fraunhofer IZT)

Data Science für Unternehmen

Basiswissen – Tools

- **MS Excel:** Harte Limitierung bezüglich Zeilen und Spalten sowie Performance Probleme bei starker Verknüpfung von Daten (z.B. S-Verweis) =>
Tools für umfangreichere Datenmengen und strukturiertere Vorgehensweise von Nöten
- **Lösung:** Data Science Tools erlauben das Umgehen mit großen Datenmengen und die strukturierte Bearbeitung sowie die einfache Nutzung komplexer Analysealgorithmen (für Aufgaben wie Regression, Clustering oder Klassifikation)
- **Strukturierte Bearbeitung:** Analysen lassen sich als Workflows (Skript) speichern und bei Bedarf neu aufrufen bzw. durch andere Analysen ersetzen (ursprüngliche Daten werden nicht verändert)



Data Science für Unternehmen

Basiswissen – Nächste Schritte

- Coursera - <https://www.coursera.org/>
- Udacity - <https://www.udacity.com/>
- DataCamp - <https://www.datacamp.com/>
- Weitere **Onlineressourcen**
 - Tutorial Sammlung - <https://data-flair.training/blogs/category/r/>
 - Cheat Sheets für R - <https://www.rstudio.com/resources/cheatsheets/>
 - Data Camp - Learn Data Science Online - <https://www.datacamp.com>

Fraunhofer Austria **Schulungsprogramm** (wird laufend weiterentwickelt):

»Einsteiger-Workshop – Data Science für Unternehmen«

»Einstieg in Data Science mit R oder Python«

»Einführung in den Data Science Technology Stack«

»Mehr Nutzen aus Daten, durch Visualisierung«

www.fraunhofer.at/izt