



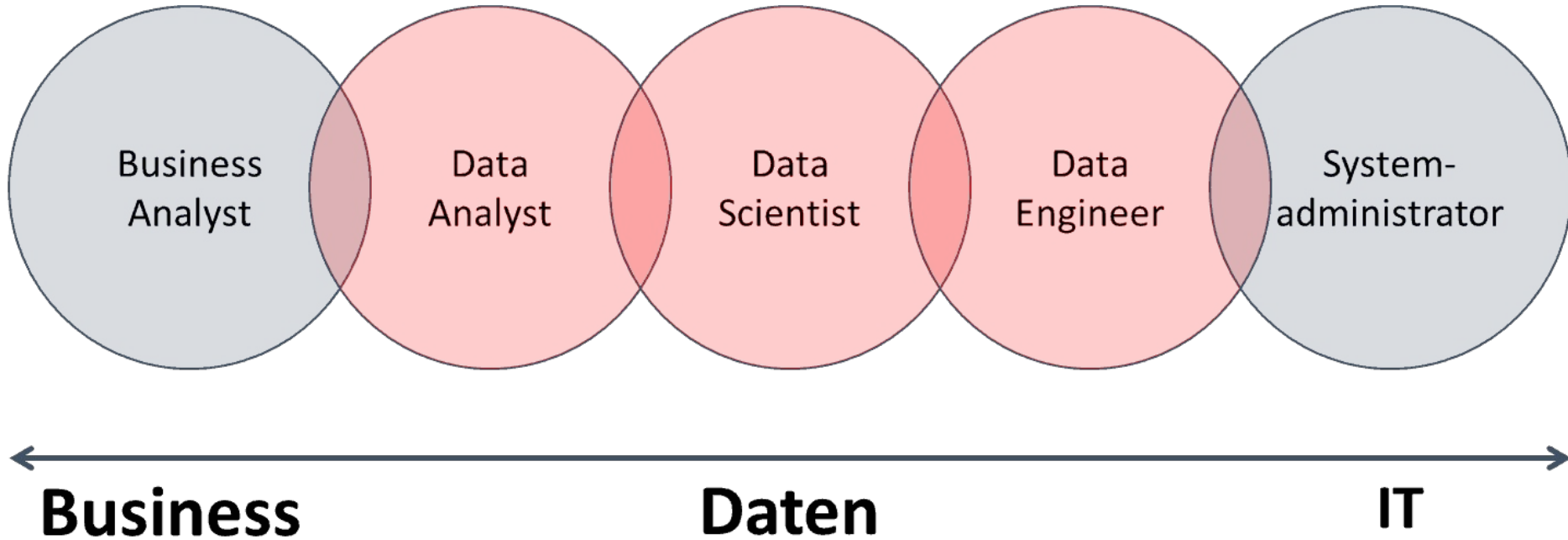
Spezialisierung Data Engineering

4. Semester

- Grundlagen Data Engineering (Prof. Dr. Mathias Walther)
- Data Warehouse (Markus Begerow)

5. Semester

- In-Memory-Datenbanken (Prof. Dr. Ralf Szymanski)
- Data Mining (Prof. Dr. Mathias Walther)



- Einführung in Skriptsprachen (z. B. Python)
- Datenstrukturen (Listen, Dictionarys)
- Arbeit mit Dateien und Dateiencodings
- Datenbanken und OR-Mapper
- Verarbeiten von Standard-Formaten (Excel, etc.)
- Visualisierung
- Rapid Prototyping
- Datenvorverarbeitung für Data-Mining

★★★★★ Ausgezeichnete Einführung in Data Science mit Python

Von [Eckhard J. J. J.](#) am 5. Februar 2018

Format: Broschiert

Das Handbuch "Data Science mit Python" ist ein tolles Nachschlagewerk, meiner

Meinung nach mindestens in zwei Hinsichten: Erstens ist es ein sehr gutes Buch für Python, das viele Themen abdeckt. Das Buch ist in der Tat ein sehr gutes Buch, das viele Themen abdeckt. Das Buch ist in der Tat ein sehr gutes Buch, das viele Themen abdeckt. Das Buch ist in der Tat ein sehr gutes Buch, das viele Themen abdeckt.

Das Buch behandelt, auf die Seite genau gerechnet, zu

5,36% die Einrichtung von Python / IPython

11,83% Numpy (Matrizen in Python)

22,00% Pandas (Tabellen in Python)

20,70% Matplotlib (Datenvisualisierung in Python)

33,64% Scikit-Learn (Machine Learning in Python)

Der Rest sind Vorwort, Stichwortverzeichnis etc.

- Arbeit mit Dateien und (→ pdf)
- Verarbeiten von Standard-Formaten (Excel, etc.)
- Textverarbeitung
- Rapid Prototyping
- Veröffentlichung/Visualisierung

- Data-Warehouse-Anwendungen
- Referenzarchitekturen Data Warehouse
- Semantische Datenmodellierung
- Effektive Datenaufbereitung
- Visualisierung
- Data Warehouse & Business Intelligence in der Cloud
- Data Warehouse Migrationspfade
- Infrastructure as a Service & Plattform as a Service
- Data Warehouse vs. Data Lake vs. Data Mesh

- Data-Warehouse-Anwendungen
 - Datentransfer und Datenaufbereitung (ETL/ELT)
 - Referenzarchitekturen Data Warehouse
 - operative (OLTP) vs. analytische Datenbanken (OLAP)
 - Semantische Datenmodellierung
 - ERM - ADAPT Notation
 - Metadaten
 - Techniken zur Aktualisierung

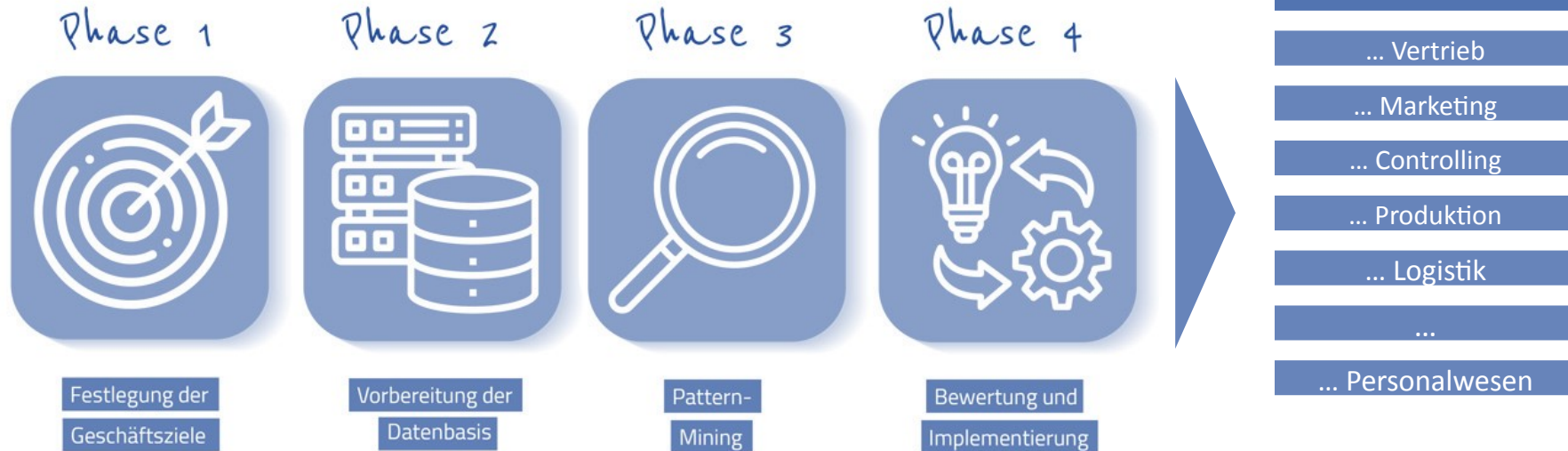
- Vertiefung Business Intelligence
 - Effektive Datenaufbereitung
 - Visualisierung
 - Management/Executive Information Systems
 - Online Analytical Processing (OLAP)

- Data Warehouse & Business Intelligence in der Cloud
 - Data Warehouse Migrationsfade
 - Infrastructure as a Service & Plattform as a Service
 - Data Warehouse vs. Data Lake vs. Data Mesh
 - Private-, Public-, Hybrid-, Community-Cloud Architekturen
 - Hoher Praxisbezug und Best Practices

- Einführung in Data Analytics und Data Mining
- Maschinelle Lernverfahren im Data Mining
- Treiber, Nutzen und Grenzen maschineller Lernverfahren
- Unüberwachte Lernverfahren:
- Überwachte Lernverfahren: Klassifikationsanalyse und Regressionsanalyse
- Evaluierung und Interpretation der Ergebnisse maschineller Lernverfahren
- Operationalisierung der Ergebnisse maschineller Lernverfahren in der betrieblichen Praxis

Die Erkenntnisse aus „Data Mining“ werden in verschiedenen betrieblichen Funktionsbereichen zur Entscheidungsfindung und Bewertung von Handlungsoptionen eingesetzt

Die 4 Phasen des Data-Mining-Prozesses



Inhalt

1. vertiefende relationale Datenmodellierung
2. vertiefende Aspekte zur Normalisierung
3. SQL (Wiederholung)
4. In-Memory-Datenstrukturen - Konzepte und Algorithmen
5. Bearbeitung von Fallbeispielen: Umsetzung von SQL-Befehlen mit einer selbsterstellten In-Memory Datenstruktur