

Inhaltsverzeichnis

Teil I Grundlagen der Verarbeitung von Daten

1	Daten, Information, Wissen	5
	Literatur	11
2	Verarbeitung der Daten (Data Processing)	13
2.1	Elektronische Datenverarbeitung	13
2.2	Datenlebenszyklus	15
2.3	Datenübertragung	17
	Literatur	19
3	Datenarten	21
3.1	Nutz- und Steuerdaten	21
3.2	Stamm-, Änderungs-, Bestands- und Bewegungsdaten	22
3.3	Referenzdaten	22
3.4	Metadaten	23
3.5	Strukturierte, semi-strukturierte und unstrukturierte Daten	25
3.6	Datenherkunft	26
3.7	Kritikalität	26
3.8	Daten mit Personenbezug	27
	Literatur	28
4	Big Data	29
4.1	Begriffliche Einordnung	30
4.2	Anwendungsfälle	31
	Literatur	33
5	Architekturen und Technologien zur Speicherung von Daten	35
5.1	Grundlagen der Datenspeicherung	36
5.2	Relationale Datenbanksysteme	38
5.3	Alternative Technologien zur Datenspeicherung	42

5.3.1	Spaltenorientierte Datenbanken	42
5.3.2	In-Memory-Datenbanken	45
5.3.3	NoSQL-Datenbanken	46
5.4	Data Warehouse	50
5.5	Data Lake	52
5.6	Cloud Computing	56
5.7	Data Mesh	58
5.8	Speicherarchitekturen und Komponenten für besondere Einsatzszenarien	60
5.8.1	Lambda-Architektur	61
5.8.2	Datenvirtualisierung	63
5.8.3	Sandboxing	65
	Literatur	66
6	Zusammenfassung	69
 Teil II Management von Daten in Unternehmen		
7	Datenstrategie	75
7.1	Definition und Ausrichtung	76
7.2	Vorgehensmodell	81
	Literatur	82
8	Data Governance	85
8.1	Gegenstand und Einordnung	86
8.2	Ziele	88
8.3	Themenbereiche	89
8.3.1	Aufbauorganisation	89
8.3.2	Vorgaben	93
8.3.3	Überwachung	96
8.3.4	Kommunikation	97
8.4	Herausforderungen für die Data Governance	97
	Literatur	99
9	Datenmanagement	101
9.1	Domänen des Datenmanagements	102
9.2	Metadaten	105
9.2.1	Ausprägungen von Metadaten	106

9.2.2	Datenkatalog	106
9.3	Datenqualität	112
9.4	Datensicherheit	119
9.5	Datenschutz	122
9.6	Daten-Ethik	124
	Literatur	127
10	Zusammenfassung	129
 Teil III Nutzung von Daten im Unternehmen – Datenanalyse		
11	Grundlagen der Datenanalyse	133
11.1	Historie der betrieblichen Datenanalyse	133
11.2	Analytics-Ausprägungen	135
11.3	Betriebswirtschaftliche Einsatzfelder	139
11.4	Data Science	140
11.4.1	Definition, Bedeutung, Aufgaben und Ziele	141
11.4.2	Anwendungsbereiche	142
11.4.3	Schwerpunkte und Kompetenzbedarf	142
	Literatur	144
12	Anwendungsdomänen von Analytics	147
12.1	Data Mining	148
12.2	Text Mining	157
12.3	Web Mining	162
12.4	Process Mining	167
	Literatur	176
13	Methoden der Data Analytics	179
13.1	Entscheidungsbaumverfahren	180
13.2	Cluster-Verfahren	185
13.3	Assoziationsanalysen	190
13.4	Künstliche Neuronale Netze	193
13.5	Vergleichende Einschätzung der Analytics-Methoden	197
	Literatur	198
14	Vorgehensmodelle für Data-Analytics-Projekte	201
14.1	Cross-Industry Standard Process for Data Mining (CRISP-DM)	202
14.1.1	Business Understanding (Geschäftsverständnis)	204
14.1.2	Data Understanding (Datenverständnis)	205
14.1.3	Datenvorbereitung (Data Preparation)	206
14.1.4	Modeling (Modellierung)	212
14.1.5	Evaluation (Auswertung)	213

14.1.6	Deployment (Verwendung)	214
14.2	Sample, Explore, Modify, Model, Assess (SEMMA)	215
14.3	Team Data Science Process (TDSP)	215
14.4	Data Science Process Model (DASC-PM)	217
14.5	Machine Learning Operations (MLOps)	218
	Literatur	220
15	Zusammenfassung	223
	Glossar	225