

Inhaltsverzeichnis

Teil I Data Strategist Digitalisierung von Geschäftsmodellen – Big Data Technologien erfolgreich implementieren

1	Big Data	3
	Uwe Schmitz	
1.1	Grundlagen	3
1.2	Architektur und Bausteine	6
1.3	Datengetriebene Geschäftsmodelle	15
1.4	Exemplarische Einsatzmöglichkeiten	17
	Literatur	24
2	Data Literacy als ein essenzieller Skill für das 21. Jahrhundert	27
	Andreas Schmidt, Thomas Neifer und Benedikt Haag	
2.1	Notwendigkeit von Data Literacy	28
2.2	Data Literacy als Begriff	30
2.3	Data Literacy Skills im Detail	32
2.4	Konzepte zur Implementation von Data Literacy in Lehre und Praxis	34
2.5	Fazit	38
	Literatur	39
3	Management von Big Data Projekten	41
	Andreas Gadatsch und Dirk Schreiber	
3.1	Konzeptioneller Rahmen des Informationsmanagements	41
3.1.1	Überblick	42
3.1.2	Aufgabenorientiertes Ebenenmodell	42
3.1.3	Integriertes Informationsmanagement	44
3.1.4	Einordnung von Big Data	45
3.2	Digitalisierung von Geschäftsmodellen mit Big Data	46
3.2.1	IT-Governance und Digitalisierung	46
3.2.2	Von der IT-Strategie zur Business Digitalstrategie	49

3.2.3	Management von Big Data.	53
3.2.4	Vorgehensmodelle zur Einführung von Big Data	54
3.2.5	Messung des Reifegrades von Organisationen.	57
3.2.6	Auswirkungen von Big Data auf die Organisation	59
	Literatur.	60
4	Digital Leadership.	63
	Wilhelm Mülder	
4.1	Führung im Digitalzeitalter	63
4.2	New Work.	64
4.2.1	Mobile Arbeitsplätze	65
4.2.2	Flexible Arbeitszeiten.	65
4.2.3	Veränderte Arbeitsinhalte.	66
4.2.4	Neue Arbeitsorganisation.	67
4.3	New Workforce	68
4.3.1	Beschäftigungseffekte der Digitalisierung.	68
4.3.2	Rekrutierung von Generation Z.	68
4.4	Digital Leader.	71
4.4.1	Persönlichkeitsmerkmale.	71
4.4.2	Führungskompetenzen	72
4.4.3	Virtuelle Führung.	72
4.5	Konzepte und Methoden für Digital Leadership	73
4.5.1	SCRUM	73
4.5.2	Design Thinking.	76
4.5.3	Servant Leadership.	77
4.5.4	VOPA+Modell	77
4.6	Fazit	79
	Literatur.	80
 Teil II Data Architect: Informationsarchitekturen gestalten – Daten effizient verwalten		
5	Data Engineering	85
	Christoph Quix	
5.1	Aufgaben des Data Engineering.	86
5.2	Architekturen zum Daten-Management.	87
5.3	Datenmodellierung und Metadaten-Management	91
5.4	Datenaufbereitung und Datenintegration.	93
5.5	Datenbank-Management-Systeme: SQL, NoSQL und Big Data.	99
5.6	Fazit	102
	Literatur.	103

6	Data Governance	105
	Detlev Frick	
6.1	Einführung	105
6.1.1	Begriffliche Einordnung.....	105
6.1.2	Datenstrategie.....	107
6.2	Data Governance Framework.....	109
6.2.1	Strategie	109
6.2.2	Aufbauorganisation	111
6.2.3	Richtlinien, Prozesse und Standards	112
6.2.4	Messen und Beobachten.....	113
6.2.5	Technologie	114
6.2.6	Kommunikation	116
6.3	Data Quality Management (DQM)	117
6.4	Fazit	118
	Literatur.....	118
7	Einsatz von In-Memory Technologien	121
	Uwe Schmitz	
7.1	Einleitung	121
7.2	Definition und Abgrenzung In-Memory Technologien	123
7.3	Anforderungen an den Einsatz einer In-Memory-Technologie	127
7.4	Bewertung	129
7.5	Fazit	131
	Literatur.....	131
8	Big-Data-Technologien	133
	Christoph Quix	
8.1	Einleitung	133
8.2	Skalierbarkeit und Fehlertoleranz	134
8.3	Volume – Management von großen Datenmengen	137
8.4	Velocity – Kontinuierliche Verarbeitung von Datenströmen	142
8.5	Variety – Unterstützung für die Zusammenführung von heterogenen Daten	145
8.6	Fazit	148
	Literatur.....	148
9	Information Data Models: Das Fundament einer guten Information Strategy	149
	Christian Rupert Maierhofer	
9.1	Drei Thesen aus Sicht eines Praktikers	150
9.2	It's all about the information	152

9.3	Das Heute und seine Hürden	152
9.4	Wie es dazu gekommen ist.	153
9.5	Die Enterprise Architektur	154
9.6	Drei Formen der Informations-Architektur und deren Auswirkungen	155
9.6.1	Das Gestern und leider noch das Heute. Der anwendungs- zentrierte Ansatz (The Application Centric Approach)	155
9.6.2	Das Heute und die Morgendämmerung, der datengesteuerte Ansatz (The Data Driven Approach)	156
9.6.3	Das überfällige Übermorgen, die datenzentrische Architektur (The Data Centric Architecture)	159
	Literatur.	162

Teil III Data Analyst: Auswerten, Präsentieren, Entscheiden – Systematische Datenanalyse im Unternehmen

10	Reporting multidimensionaler Daten und Kennzahlen.	167
	Detlev Frick und Birgit Lankes	
10.1	Betriebswirtschaftliche Motivation	167
10.1.1	Kennzahlen und ihre Anwendung	168
10.1.2	Auswahl von Kennzahlen.	169
10.2	Daten und Business Intelligence	170
10.2.1	Datenmodellierung.	171
10.2.2	Datensicherung.	172
10.2.3	Harmonisierung	173
10.2.4	Daten-/Informationsqualität.	173
10.2.5	Datenbereitstellung	174
10.3	Reporting/Berichtswesen	174
10.3.1	Berichtsgrundformen	176
10.3.2	Anforderungen an Berichte	177
	Literatur.	177
11	Fundamentale Analyse- und Visualisierungstechniken.	179
	Jens Kaufmann	
11.1	Einleitung und Begriffswelt	179
11.2	Lineare Regression.	182
11.2.1	Basisidee und Begrifflichkeiten	182
11.2.2	Beispiel und Ergebnisinterpretation.	183
11.2.3	Prüfen der Voraussetzungen und Variablentransformation	185
11.3	Einfache Klassifikationsverfahren	186
11.3.1	k-Nearest-Neighbors	186
11.3.2	Naive Bayes	187

11.3.3	Entscheidungsbäume	188
11.4	Clustering-Verfahren	189
11.4.1	Hierarchische Verfahren	189
11.4.2	Partitionierende Verfahren	191
11.5	Assoziationsanalyse	191
11.6	Ergänzende Überlegungen, Software und Tools	192
	Literatur	193
12	Fortgeschrittene Verfahren zur Analyse und Datenexploration, Advanced Analytics und Text Mining	195
	Jens Kaufmann	
12.1	Einleitung	195
12.2	Datenexploration und -darstellung	196
12.3	Principal Component Analysis	197
12.4	Random Forests	200
12.5	Logistische Regression	200
12.6	Entscheidungsbewertung	201
12.7	Zeitreihenanalyse	202
12.8	Text Mining	205
12.9	Weitere Analysemöglichkeiten	207
	Literatur	208
13	Datenbasierte Algorithmen zur Unterstützung von Entscheidungen mittels künstlicher neuronaler Netze	209
	Daniel Retkowitz	
13.1	Datenbasierte Algorithmen und maschinelles Lernen	209
13.1.1	Maschinelles Lernen	210
13.1.2	Lernverfahren	211
13.2	Künstliche neuronale Netze	212
13.2.1	Netzarchitekturen	212
13.2.2	Grenzen künstlicher neuronaler Netze	213
13.3	Beispielhafte Anwendungsfelder	214
13.4	Entwicklungsprozess	216
13.5	Entwicklungsplattformen und Werkzeuge	217
13.5.1	TensorFlow und PyTorch	218
13.5.2	Ausführungsmodi	219
13.5.3	Deployment und Betrieb	220
13.6	Fazit und Ausblick	222
	Literatur	223
14	Künstliche Neuronale Netze – Aufbau, Funktion und Nutzen	225
	Anja Tetzner, Tom Kühne, Peter Gluchowski und Melanie Pfoh	
14.1	Einleitung	226

14.2	Aufbau	227
14.2.1	Künstliches Neuron	227
14.2.2	Künstliche neuronale Netze	229
14.3	Lernen künstlicher neuronaler Netze	233
14.3.1	Überwachtes Lernen – Lernen mittels Backpropagation	234
14.3.2	Unüberwachtes Lernen – Lernen mittels Wettbewerbslernen	235
14.4	Nutzenpotenziale und Herausforderungen	236
14.5	Fazit	238
	Literatur	238
15	Bayesian Thinking in Machine Learning	241
	Thomas Neifer, Andreas Schmidt, Dennis Lawo, Lukas Böhm und Özge Tetik	
15.1	Bayesian Thinking	242
15.2	Bayes in Machine Learning	245
15.2.1	Bayes in Regressionsverfahren	245
15.2.2	Bayes in Klassifikationsverfahren	249
15.3	Naive Bayes Classifier	251
15.3.1	Grundlagen	251
15.3.2	Methodik	252
15.4	Fazit	254
	Literatur	254
 Teil IV Anwendungsorientierte Data Science		
16	Text Mining: Durchführung einer Sentiment Analysis mit SAP HANA	259
	Patrick Bachmaier	
16.1	Einleitung	259
16.2	Grundlagen	260
16.3	Umsetzung	261
16.3.1	Vorgehensmodell	261
16.3.2	Implementierung	263
16.4	Fazit	273
	Literatur	274
17	Weiterbildung in Data Science.	277
	Christoph Quix	
17.1	Kompetenz-Rahmenwerke für Data Science	278
17.2	Studiengänge zu Data Science	280

17.3	Berufliche Weiterbildung zu Data Science.	283
17.3.1	Zertifikatsprogramm der Fraunhofer Gesellschaft zu Data Science	284
17.3.2	Zertifikatsstudien der Hochschule Niederrhein	285
17.3.3	Zertifikatslehrgang zum Data Scientist der Bitkom Akademie	287
17.4	Fazit	288
	Literatur.	289
18	Plattformökonomie für Data Plattformen	291
	Valeria Knoll und Alexa Scheffler	
18.1	Motivation	291
18.2	Begriffshaushalt	292
18.2.1	Plattformen und Plattformökonomie	292
18.2.2	Data Plattform	294
18.3	Design-Prinzipien für Data Plattformen.	296
18.3.1	Netzwerkeffekte durch gemeinsam genutzte Datenobjekte.	296
18.3.2	Strategien für die Aktivierung von Plattformteilnehmern	297
18.3.3	Einfacher Zugang durch Self-Service	298
18.3.4	Effektives Matching durch Metadaten	299
18.4	Monetarisierung	299
18.5	Zusammenfassung und Fazit	300
	Literatur.	302
19	Akzeptanz und Nutzung von maschinellem Lernen und Analytics im Rechnungswesen und Controlling	305
	Markus Eßwein, Domenica Martorana, Martina Reinersmann und Peter Chamoni	
19.1	Eine Herausforderung für die Finanzfunktion	306
19.2	Nutzerakzeptanzforschung zu maschinellem Lernen.	307
19.3	Befragung von Führungskräften	308
19.3.1	Strukturgleichungsmodell	308
19.3.2	Umfrage	308
19.4	Aktuelle Nutzung und Treiber	310
19.4.1	Ergebnisse der Befragung	310
19.4.2	Treibermodell zur Nutzung und Akzeptanz.	315
19.5	Handlungsempfehlungen und Ausblick	317
	Literatur.	318
20	Durch Daten zu neuen Geschäftsmodellen und Prozessoptimierungen – im Kontext von Car-Sharing	321
	Eva Schoetzau	
20.1	Kurze Einführung.	321

20.2	Durch Daten zu neuen Ideen und Optimierungen	322
20.3	Umdenken im Unternehmen	325
20.4	Durch ständige Überwachung zur stetigen Anpassung	328
20.5	Mit ‚Lessons Learned‘ zur Optimierung von Geschäftsmodellen und -prozessen	331
20.6	Fazit	335
	Literatur.	335
21	Einsatz von Logit- und Probit-Modellen in der Finanzindustrie	337
	Uwe Rudolf Fingerlos und Alexander Pastwa	
21.1	Einleitung.	337
21.2	Logit- und Probit-Modelle	338
21.3	Datengrundlage	340
21.4	Modellierung	343
21.5	Überprüfung der Modellannahmen	347
21.6	Vorstellung der Ergebnisse.	348
21.7	Vergleichende Beurteilung	351
	Literatur.	354
	Stichwortverzeichnis.	357