

Inhaltsverzeichnis

Teil I Informationsqualität – Grundlagen

1 Was wissen wir über Information?	3
Florian Engelmann und Christoph Großmann	
1.1 Einleitung	3
1.2 Grundlegung	5
1.3 Information im Wissens- und Informationsmanagement	7
1.4 SHANNONsche Informationstheorie	11
1.5 STEINMÜLLERs Informationsmodell	13
1.5.1 STEINMÜLLERs System- und Prozessverständnis	13
1.5.2 Information als allgemeines Modell	14
1.5.3 Modell eines Informationssystems unter Einbezug der Semiotik	15
1.5.4 Fazit	16
1.6 Information als Produktionsfaktor	17
1.6.1 Perspektive der Produktionstheorie	17
1.6.2 Produktionsfaktor Information	20
1.6.3 Fazit	20
1.7 Zusammenfassung des Beitrages	21
Literatur	21
2 Informationsqualität – Definitionen, Dimensionen und Begriffe	23
Jan P. Rohweder, Gerhard Kasten, Dirk Malzahn, Andrea Piro und Joachim Schmid	
2.1 Einleitung	23
2.2 IQ-Dimensionen und Definitionen	25
2.2.1 Die 15 IQ-Dimensionen im Überblick	26
2.2.2 Graphische Darstellung der 15 IQ-Dimensionen und 4 IQ-Kategorien	27
2.2.3 Die 15 IQ-Dimensionen: Definitionen und Beispiele	29
2.3 Zusammenfassung und Ausblick	41

Literatur.....	42
Teil II Methoden – Techniken – Tools – Regelwerke/Standards	
3 Datenqualitätsmetriken für ein ökonomisch orientiertes Qualitätsmanagement.....	47
Bernd Heinrich und Mathias Klier	
3.1 Einleitung.....	47
3.2 Anforderungen an Datenqualitätsmetriken.....	49
3.3 Bisherige Beiträge zur Messung von Datenqualität.....	50
3.4 Metriken und Messverfahren für DQ.....	52
3.4.1 Metrik für die DQ-Dimension Vollständigkeit.....	52
3.4.2 Metrik für die DQ-Dimension Fehlerfreiheit.....	55
3.4.3 Metrik für die DQ-Dimension Konsistenz.....	58
3.4.4 Metrik für die DQ-Dimension Aktualität.....	59
3.5 Praktische Anwendung der Metrik für Aktualität.....	61
3.6 Zusammenfassung und Ausblick.....	62
Literatur.....	63
4 Datenqualitätsmanagement – Steigerung der Datenqualität mit Methode.....	67
Niels Weigel	
4.1 Die Bedeutung des Total Data Quality Management.....	68
4.1.1 Vorgehensmodelle.....	68
4.1.2 Datenqualitätsmanagement sichert Ihren Unternehmenserfolg....	69
4.2 Phasen eines ganzheitlichen Datenqualitätsmanagements.....	69
4.2.1 Initiierung des Datenqualitätsprojekts.....	70
4.2.2 Definition der Datenqualitätsanforderungen.....	73
4.2.3 Messung der vorhandenen Datenqualität.....	76
4.2.4 Analyse der Fehlerursachen.....	79
4.2.5 Verbesserung der Datenqualität.....	81
4.2.6 Permanente Überwachung der Datenqualität.....	82
4.3 Anreize für ein Datenqualitätsmanagement.....	83
Literatur.....	84
5 Strukturierte Datenanalyse, Profiling und Geschäftsregeln.....	87
Marcus Gebauer und Ulrich Windheuser	
5.1 Datenqualität.....	87
5.2 Merkmale der Datenqualität.....	89
5.3 Geschäftsregeln.....	92
5.4 Methoden der Datenanalyse.....	93
5.5 Metriken im Detail.....	95
5.6 Datenqualität in der Anwendung.....	97
Literatur.....	100

6 Datenbereinigung zielgerichtet eingesetzt zur permanenten Datenqualitätssteigerung	101
Marcus Zwirner	
6.1 Definition „Datenbereinigung“	102
6.2 Ursachenanalyse	103
6.3 Bewertungskriterien für Datenfehler und Korrekturmaßnahmen	103
6.4 Methoden des Datenqualitätsmanagements	107
6.5 Datenqualitätsmaßnahmen im Detail	109
6.6 Zusammenfassung	120
7 Datenintegration und Deduplizierung	123
Jens Bleiholder und Joachim Schmid	
7.1 Schritt 1: Schema Matching	126
7.2 Schritt 2: Dublettenerkennung	129
7.2.1 Auswirkungen von Dubletten	130
7.2.2 Entstehung von Dubletten	131
7.2.3 Erkennen von Dubletten	132
7.2.4 Durchführung der Dublettenerkennung	133
7.3 Schritt 3: Datenfusion	135
7.3.1 Konflikte ignorieren	136
7.3.2 Konflikte vermeiden	136
7.3.3 Konflikte auflösen	137
7.4 Erweiterungen	139
7.4.1 Strukturierung	139
7.4.2 Standardisierung	140
7.5 Zusammenfassung	141
Literatur	141
8 Definition von Datenarten zur konsistenten Kommunikation im Unternehmen	143
Andrea Piro und Marcus Gebauer	
8.1 Einleitung und Zielsetzung	143
8.1.1 Informationsqualität und Datenarten	144
8.2 Datenarten in der Informationslandschaft	145
8.3 Beschreibungskriterien	145
8.3.1 Beschreibung der Eigenschaften	146
8.3.2 Beschreibung des Kontextes	148
8.4 Beispiele für den Praxiseinsatz	150
8.4.1 Analyseebenen der Informationsqualität	150
8.4.2 Visualisierung des IQ-Status	151
8.5 Zusammenfassung	154
Literatur	155

9	Suchmaschinen und Informationsqualität: Status quo, Problemfelder, Entwicklungstendenzen	157
	Christian Maaß und Gernot Gräfe	
9.1	Ausgangssituation	157
9.2	Charakterisierung algorithmenbasierter Suchmaschinen	159
9.2.1	Funktionsweise algorithmenbasierter Suchmaschinen	159
9.2.2	Anfrageabhängige Ranking-Faktoren	159
9.2.3	Anfrageunabhängige Ranking-Kriterien	161
9.3	Semantisches Web und semantische Suchmaschinen	163
9.3.1	Vision und Grundlagen des semantischen Webs	163
9.3.2	Technische Grundlagen des semantischen Web	166
9.3.3	Problemfelder und Herausforderungen im Bereich der semantischen Suche	166
9.4	Fazit und Ausblick	168
	Literatur	169
10	Bedeutung der Informationsqualität bei Kaufentscheidungen im Internet	171
	Gernot Gräfe und Christian Maaß	
10.1	Einleitung	171
10.2	Informationsqualität in Entscheidungsprozessen	172
10.2.1	Informationen und Kaufentscheidungen	172
10.2.2	Informationsqualitätskriterien	174
10.3	Ursachen mangelnder Informationsqualität im Internet	176
10.3.1	Opportunistische Verhaltensspielräume der Anbieter	176
10.3.2	Informationsqualität aus der Nachfragerperspektive	179
10.3.3	Gründe für Opportunismus im Internet	182
10.4	Fazit und Handlungsempfehlungen	187
	Literatur	190
11	Datenqualitäts-Audits in Projekten	195
	Marcus Gebauer und Michael Mielke	
11.1	Einleitung	195
11.2	Abstimmung mit anderen Regelwerken	197
11.3	Glossar	197
11.4	Gebrauch der Generischen Checkliste	198
11.5	Datenqualitätsbewertung einer Datensammlung	200
11.5.1	Anforderungen an das Management	200
11.5.2	Service Level Agreements	201
11.5.3	Organisatorische Spezifikationen	201
11.5.4	Prozess-Definitionen	202
11.5.5	Datensammlung, Datenverarbeitung und Datennutzung	204
11.5.6	Messung, Maßnahmen und Überwachung	205

11.5.7 Technische Anforderungen	206
11.5.8 Dokumentation	208
11.6 Zusammenfassung	208
12 Bewertung der Informationsqualität im Enterprise 2.0	211
Sven Ahlheid, Gernot Gräfe, Alexander Krebs und Dirk Schuster	
12.1 Einführung	211
12.2 Beurteilung der Informationsqualität einer Enterprise 2.0	
Wissensplattform mittels eines hybriden Ansatzes	212
12.2.1 Automatische Beurteilung der Informationsqualität	213
12.2.2 Implizites Nutzer-Feedback	215
12.2.3 Explizites Nutzer-Feedback	216
12.2.4 Zusammenwirken der drei Ansätze und Fazit	217
Literatur.	218
Teil III Organisation	
13 Organisatorische Ansiedlung eines Datenqualitätsmanagements	225
Jens Lüssem	
13.1 Einführung	225
13.1.1 Motivation.	225
13.1.2 Gliederung des Kapitels	226
13.2 Datenqualitätsmanagement – Entwicklungsstufen und Aufgaben.	227
13.2.1 Sicherung der Datenqualität	228
13.2.2 Management der Datenqualität.	228
13.3 Datenqualitätsmanagement – Ansiedlung im Unternehmen	229
13.3.1 Kopplung von Datenqualitätsmanagement mit anderen Unternehmensbereichen	229
13.3.2 Folgerungen für die Ansiedlung eines Datenqualitätsmanagements.	231
13.4 Datenqualitätsmanagement in Projekten	232
13.4.1 Aufgaben des Datenqualitätsmanagements in Projekten	233
13.4.2 Organisatorische Verankerung des Datenqualitätsmanagements in Projekten.	233
13.5 Zusammenfassung und Ausblick	234
13.5.1 Zusammenfassung	234
13.5.2 Ausblick	235
Literatur.	235
14 Organisatorische Maßnahmen für gute Datenqualität	237
Jürg Wolf	
14.1 Messungen, Ursachen und generische Ansätze	237
14.1.1 Möglichen Arten von Datenqualitätsmängeln	237
14.1.2 Datenqualitätsmängel – Entstehung und Bekämpfung.	238

14.1.3	Vier Generische Ansätze.	239
14.1.4	Aus den generischen Ansätzen abgeleitete Strategien	240
14.2	Strategie A: Transparenz schafft Vertrauen	241
14.2.1	Ansatzpunkt dieser Strategie	241
14.2.2	Nutzen dieser Strategie.	241
14.2.3	Nachteile und Risiken dieser Strategie.	242
14.3	Strategie B: Definition von Verantwortlichkeiten	242
14.3.1	Ansatzpunkt dieser Strategie	242
14.3.2	Positionierung dieser Businessrollen im Modell	242
14.3.3	Nutzen dieser Strategie.	244
14.3.4	Nachteile und Risiken dieser Strategie.	244
14.4	Strategie C: gezielt Abhängigkeiten suchen.	244
14.4.1	Ansatzpunkt dieser Strategie	244
14.4.2	Gezielte Definition von Master und Slave	245
14.4.3	Nutzen dieser Strategie.	246
14.4.4	Nachteile und Risiken dieser Strategie.	246
14.5	Strategie D: Daten-Lifecycle auf Basis des Prozesses.	247
14.5.1	Ansatzpunkt dieser Strategie	247
14.5.2	Der Prozess und Lebenszyklus.	247
14.5.3	Nutzen dieser Strategie.	248
14.5.4	Nachteile und Risiken dieser Strategie.	248
14.6	Strategie E: Niederschwellige Verbesserungs-Werkzeuge.	249
14.6.1	Ansatzpunkt dieser Strategie	249
14.6.2	Beispiel eines niederschwelligen Verbesserungs-Werkzeuges	249
14.6.3	Die Infrastruktur dieses Werkzeuges	250
14.6.4	Nutzen dieser Strategie.	250
14.6.5	Nachteile und Risiken dieser Strategie.	251
14.7	Vor- und Nachteile aller erwähnter Strategien.	251
14.7.1	Der Prozess ist die Vorgabe	251
14.7.2	Das Saatkorn ist der Beginn	251
14.7.3	Komplexität des Systems und Datenvolumen	252
14.8	Vorgehen bei der Umsetzung dieser Strategien	252
14.8.1	Kontakt zwischen den Parteien.	252
14.8.2	Management-Unterstützung.	252
14.9	Schlussfolgerungen und Ausblick	252
	Literatur.	253
15	Informationsmanagementprozesse im Unternehmen	255
	Klaus Schwinn	
15.1	Motivation	255
15.2	Ausgangslage.	256
15.3	Bewertung	257

15.4	Informationsmanagementprozess	259
15.5	Schema einer Informationsplanung	261
15.6	Datenlandkarte und Datenarchitektur	262
15.7	Geschäftsprozesse und Informationsmanagementprozess.....	264
15.8	Qualitätsaspekte	265
15.9	Ökonomische Aspekte	268
15.10	Zusammenfassung	270
	Literatur.	270
16	Data Governance.	271
	Kristin Weber, Boris Otto und Dominik Lis	
16.1	Einführung	271
16.2	Gestaltungsbedingungen von Data Governance als Rahmenwerk.	273
16.3	Ein Modell für Data Governance	275
16.3.1	Rollen	275
16.3.2	Aufgaben	279
16.3.3	Zuständigkeiten	283
16.3.4	Gestaltungsvarianten	284
16.4	Praxisbeispiele und Entwicklungsstränge der Datenqualität	286
16.4.1	Herausforderungen der Datenqualität in Industrie 4.0-Szenarien	286
16.4.2	Institutionalisierung des Stammdatenmanagements in der Konsumgüterindustrie.	287
16.4.3	Datengetriebene Dienstleistungen im Maschinen- und Anlagenbau.	287
16.5	Zusammenfassung	288
16.6	Ausblick	289
	Literatur.	290
17	IQM-Reifegradmodell für die Bewertung und Verbesserung des Information Lifecycle Management Prozesses.	293
	Saša Baškarada, Marcus Gebauer, Andy Koronios und Jing Gao	
17.1	Einleitung	293
17.2	Hintergrund	294
17.2.1	Total Quality Management	294
17.2.2	QM-Reifegrad	295
17.2.3	Information Quality Management	296
17.2.4	Existierende IQM-Reifegradmodelle	297
17.3	Methodologie	297
17.3.1	Die Delphi-Methode	297
17.4	IQM-Reifegradmodell	298
17.4.1	Chaotisch	299
17.4.2	Reaktiv	299
17.4.3	Messend	300

17.4.4	Steuernd	301
17.4.5	Optimierend	301
17.5	Zusammenfassung und Ausblick	302
	Literatur.	302
18	Management der Materialstammdaten in SAP®-Systemen.	307
	Knut Hildebrand	
18.1	Stammdaten – die wichtigsten digitalen Zwillinge	307
18.2	Stammdatenqualität führt zu Prozessqualität.	310
18.2.1	Qualitätseigenschaften	310
18.2.2	Probleme der Datenqualität und ihre Auswirkungen	310
18.3	Master Data Life Cycle (MDLC) – der Stammdatenprozess.	311
18.3.1	Statuskonzept	311
18.3.2	Hindernisse und Problemfälle	315
18.3.3	Tools der SAP®	316
18.4	Implementierung des MDLC.	317
18.5	Resümee.	318
	Literatur.	318
19	Prinzipien erfolgreichen Informationsqualitätsmanagements im	
	Lichte von Industrie 4.0	321
	Michael Mielke	
19.1	Big Data = Big Data Quality?	321
19.2	Datenqualität und Industrie 4.0	323
19.3	Übergeordnete Grundsätze und Einordnung der IQM-Grundsätze	325
19.4	Verantwortung für die Daten übernehmen.	327
19.5	Agile DQ-Entwicklung	329
Teil IV Praxisbeispiele		
20	Ein Entscheidungsmodell zur Weitergabe persönlicher Daten im Internet .	335
	Horst Treiblmaier	
20.1	Einleitung.	335
20.2	Entscheidungsmodell.	337
20.2.1	Intention	338
20.2.2	Nutzen.	339
20.2.3	Vertrauen.	341
20.2.4	Datenarten.	344
20.2.5	Eingabefehler	346
20.3	Ausblick.	346
	Literatur.	347

21 Der Aufbau einer Organisationsrichtlinie für den Daten- und Informationsqualitätsmanagement-Prozess 349
Reinhard Höhn

21.1 Motivation 349

21.2 Der komponierte DQM-Prozess 352

21.2.1 Schritt 0: Anlass zur Implementierung eines DQM-feststellen 354

21.2.2 Schritt 1: Data Quality Management positionieren und beauftragen 357

21.2.3 Schritt 2: Für das aktuelle DQ-Projekt relevante Enterprise-Architektur erfassen 361

21.2.4 Schritt 3: Qualität der Daten entsprechend der DQ-Merkmale erfassen 368

21.2.5 Schritt 4: Auswirkungen von DQ-Mängeln analysieren. 374

21.2.6 Schritt 5: Ursachen der DQ-Mängel beseitigen 381

21.2.7 Schritt 6: Software-Entwicklung und Betriebsüberführung 388

21.2.8 Schritt 7: Assessment des Reifegrades des Datenqualitätsmanagementprozess. 390

21.3 Referenzprozess mit Rollen 393

21.4 Die DQM-Richtlinie 397

21.5 Resümee. 398

Literatur. 400

22 Informationsqualität und Digitale Assistenzsysteme: Ein Laborbericht aus dem Campus 4.0 401
Abdessalam Ait Salah, Hendrik Thüs und Michael Mielke

22.1 Information als Wettbewerbsfaktor 401

22.2 Problemstellung 402

22.3 Grundlagen. 404

22.3.1 Künstliche Intelligenz (KI). 404

22.3.2 Maschinelles Lernen. 404

22.3.3 Daten als Grundlagen 408

22.3.4 Chatbots 409

22.4 Entwurf und Implementierung von K.I.D. 410

22.4.1 *Genutzte Bibliotheken.* 411

22.4.2 Erstellung der Wissensbasis 411

22.4.3 Bag-of-Words 413

22.4.4 Beschreibung der Funktionalität. 414

22.4.5 Informationsqualität messen und verbessern in K.I.D. 414

22.5 Evaluation 416

22.6 Nutzung im Bahnkontext 417

22.7	Ausblick	418
22.8	Zusammenfassung	418
	Literatur	418
23	Datenqualitäts-Modell der Volkswagen Financial Services AG	421
	Helena Moser	
23.1	Einleitung	421
23.2	Das Projekt „Datenqualität Strukturen/Standards und Drittmarktfähigkeit“	423
23.2.1	Warum ist Datenqualität nötig?	423
23.2.2	Projektauftrag	423
23.2.3	Projektziel	424
23.2.4	Ist-Analyse	424
23.2.5	Prozessanalyse	425
23.2.6	Sollkonzept	426
23.2.7	Das Datenqualitäts-Modell und deren Zuständigkeiten (Abb. 23.1)	428
23.2.8	Das Datenqualitäts-Modell und sein Regelwerk	432
23.2.9	Monitoring/Reports	433
23.2.10	Realisierungs- und Einführungsphase	434
23.3	Fazit	435
24	Verknüpfung von DQ-Indikatoren mit KPIs und Auswirkungen auf das Return on Investment	437
	Frank Block	
24.1	Beispiele zur Illustration von DQ-Problemen	438
24.2	Wie wirken sich DQ-Probleme auf Unternehmen aus – Der Zusammenhang zwischen Daten- und Prozessqualität	439
24.2.1	Beispiel – Call Center	440
24.2.2	Beispiel – Kundenbeziehungsmanagement (CRM)	440
24.2.3	Beispiel – Data Mining Prozess im Marketingumfeld	441
24.2.4	Beispiel – Direktmailprozess	442
24.3	Wie viel kosten schlechte Daten den Unternehmer?	443
24.4	Der Einfluss von DQ-Indikatoren auf KP-Indikatoren – wie beeinflusst Datenqualität den Unternehmenserfolg?	445
24.5	Beschreibung eines KPI orientierten DQ-Managementprozesses	447
24.5.1	Phase 1 – Selektiere zu untersuchende Komponenten	448
24.5.2	Phase 2 – Mitarbeiterbefragung	449
24.5.3	Phase 3 – DQ-Assessment	451
24.5.4	Phase 4 – Validieren und Quantifizieren	452
24.5.5	Phase 5 – DQ-Projekte definieren, Korrekturmaßnahmen durchführen	453
24.6	Fallstudie – Banque Cantonale Vaudoise (BCV)	455
	Literatur	458

25	Gewährleistung einer hohen Artikelstammdatenqualität im Global Data Synchronisation Network (GDSN)	459
	S. Kasper	
25.1	Global Data Synchronization Network (GDSN)	460
25.2	Bausteine des GDSN zur Optimierung der Artikelstammdatenqualität	465
25.2.1	Data Quality Framework (DQF)	465
25.2.2	GDSN Package Measurement Rules und Implementation Guide	470
25.2.3	Data Quality Discussion Group	471
25.2.4	GDSN-Umsetzungsleitfaden zur technischen Anwendung im Rahmen der Lebensmittelinformations-Verordnung	473
25.3	Data Quality Gate und Data Quality Excellence	476
25.4	Zusammenfassung und Ausblick	476
	Literatur	477
	Stichwortverzeichnis	481