

Inhaltsübersicht

Teil I

1	Datenqualität	3
2	Ausprägungen und Ursachen schlechter Datenqualität	19
3	Auswirkungen schlechter Datenqualität	37
4	Organisation	53
5	Referenzarchitektur für Business-Intelligence-Anwendungen	69
6	Big Data	91
7	Kennzahlen zur Messung der Datenqualität	103

Teil II

8	Verbesserung der Datenqualität im Quellsystem	123
9	Data Profiling	131
10	Erfolgreiche Datenvalidierung und -filterung	175
11	Standardisierung und Bereinigung	187
12	Datenanreicherung	219
13	Verbesserung der Datenqualität in der Bereitstellung und Visualisierung	233
14	Wertschöpfung durch Metadaten	253
15	Data Quality Monitoring	269
16	Produktauswahl und -integration	285

Teil III	295
17 Datenqualitätsmanagement in einer Studie	301
18 Datenqualitätsmanagement in der Spezifikation	319
19 Datenqualitätsmaßnahmen in der Konstruktionsphase	335
20 Steuerung der Datenqualität in der Realisierung	345
21 Steuerung der Datenqualität im Betrieb	351
Anhang	355
Abkürzungen	357
Literatur	359
Index	367

Inhaltsverzeichnis

Teil I	1
1	Datenqualität
1.1	Daten
1.2	Qualität
1.3	Datenqualität
1.4	Datenqualitätsmanagement
1.5	Zusammenfassung
2	Ausprägungen und Ursachen schlechter Datenqualität
2.1	Geschäftstreiber
2.2	Ausprägungen schlechter Datenqualität
2.3	Ursachen schlechter Datenqualität
2.4	Beispiel: Finanzdienstleister
2.5	Empfehlungen
3	Auswirkungen schlechter Datenqualität
3.1	Datenqualitätskosten
3.2	Gesetzliche Anforderungen
3.3	Business-Case-Betrachtungen
3.4	Empfehlungen
4	Organisation
4.1	Aufbauorganisation
4.2	Ablauforganisation
4.3	Empfehlungen

5	Referenzarchitektur für Business-Intelligence-Anwendungen	69
5.1	Referenzarchitektur	69
5.1.1	Datenquellen und Datenströme	71
5.1.2	Datenintegration	72
5.1.3	Datenhaltung	73
5.1.4	Informationsbereitstellung	73
5.1.5	Anwender und Rollen	74
5.1.6	Operative Anwendungen und Prozesse	74
5.1.7	Querschnittsprozesse	74
5.2	Problemstellen und Lösungsansätze hinsichtlich der Datenqualität ..	75
5.2.1	Datenquellen	75
5.2.2	Datenintegration	76
5.2.3	Datenhaltung	79
5.2.4	Informationsbereitstellung	80
5.3	Architektur für Datenqualitätsmanagement	81
5.4	Serviceorientierte Architektur	83
5.5	Master Data Management	85
5.5.1	Architektur	86
5.5.2	Umsetzung	89
5.6	Empfehlungen	90
6	Big Data	91
6.1	Definitionen von Big Data	91
6.1.1	Fachlich-datenbezogene Sicht	93
6.1.2	Gartner-Sicht	94
6.1.3	Technisch-infrastrukturelle Sicht	95
6.2	Bedeutung der Datenqualität bei Big Data	95
6.3	Herausforderung externe Daten	97
6.4	Herausforderung unstrukturierte Daten	99
6.5	Herausforderung Geschwindigkeit	100
6.6	Herausforderung Volumen	101
6.7	Empfehlungen	102
7	Kennzahlen zur Messung der Datenqualität	103
7.1	Anwendungsmöglichkeiten von Kennzahlen	104
7.2	Messpunkte für Datenqualität	106
7.3	DQ-Metriken	109
7.4	Kennzahlen für ausgewählte Datenqualitätskriterien	112
7.5	Kennzahlenbaum	114
7.6	Kennzahlenformular	115
7.7	Empfehlungen	116

Teil II

117

8	Verbesserung der Datenqualität im Quellsystem	123
8.1	Vorbeugung vor neuen Datenqualitätsproblemen	124
8.2	Empfehlungen	129
9	Data Profiling	131
9.1	Data-Profiling-Prozess	132
9.1.1	Schritt 1: Integration der Daten	133
9.1.2	Schritt 2: Analyse der integrierten Daten	133
9.1.3	Schritt 3: Darstellung der Ergebnisse	134
9.1.4	Fachliche Bewertung der Ergebnisse	134
9.2	Zusammensetzung des Data-Profiling-Teams	135
9.3	Vorgehensweise beim Data Profiling	136
9.4	Data-Profiling-Verfahren zur Analyse von Attributen	137
9.4.1	Standardanalysen auf Attributebene	137
9.4.2	Analyse der Attribute mit Geschäftsregeln	150
9.5	Data-Profiling-Verfahren zur Analyse von Datensätzen	158
9.5.1	Analyse auf Schlüsselattribute	158
9.5.2	Analyse auf abgeleitete Werte	161
9.5.3	Analyse von Datensätzen mit Geschäftsregeln	162
9.6	Data-Profiling-Verfahren zur Analyse von Tabellen	163
9.6.1	Analyse von Tabellen auf referenzielle Abhängigkeiten ..	163
9.6.2	Analyse von Tabellen mit Geschäftsregeln	168
9.7	Empfehlungen	174
10	Erfolgreiche Datenvalidierung und -filterung	175
10.1	Validierung auf vier Ebenen	175
10.2	Filterung fehlerhafter Daten	176
10.3	Validierung bei Extraktion oder Laden	180
10.4	Arten der Datenvalidierung	182
10.5	Erstellung der Validierungsregeln und Speicherung der Ergebnisse .	184
10.6	Empfehlungen	185
11	Standardisierung und Bereinigung	187
11.1	Standardisierung	187
11.2	Datenbereinigung	189
11.3	Standardisierung und Bereinigung im ETL-Prozess	217
11.4	Verfahren für nicht zu bereinigende Daten	218
11.5	Empfehlungen	218

12 Datenanreicherung	219
12.1 Wirtschaftsinformationen	219
12.2 Geografische Informationen	222
12.3 Soziodemografische Informationen	224
12.4 Haushaltsbildung	225
12.5 Standards zur Klassifizierung von Waren und Dienstleistungen	226
12.6 Branchenklassifizierung	229
12.7 Empfehlungen	232
13 Verbesserung der Datenqualität in der Bereitstellung und Visualisierung	233
13.1 Bereitstellung der Daten	233
13.2 Visualisierung der Information	235
13.3 Empfehlungen	251
14 Wertschöpfung durch Metadaten	253
14.1 Metadaten: Begriff und Strukturierung	253
14.2 Metadatenarchitekturen	255
14.3 Metadatenmanagement	258
14.4 Metadatenkategorien	260
14.5 Probleme bei der Erstellung: Motivation und Aktualität	265
14.6 Nutzung von Metadaten	265
14.7 Empfehlungen	268
15 Data Quality Monitoring	269
15.1 DQ-Planung	270
15.2 DQ-Assessment	271
15.3 DQ-Phasenkonzepte	274
15.4 Methoden	277
15.5 Verantwortlichkeiten	283
15.6 Empfehlungen	283
16 Produktauswahl und -integration	285
16.1 Anbieter und Produkte	285
16.2 Auswahlkriterien im Überblick	287
16.3 Funktionale Kriterien	287
16.4 Integration	291
16.5 Einbeziehung der Fachbereiche	293
16.6 Sprachen und Länder	293
16.7 Einbindung in DQM-Prozesse	294
16.8 Empfehlungen	294

Teil III	295
17 Datenqualitätsmanagement in einer Studie	301
17.1 Analyse des Istzustands	302
17.2 Entwurf des Sollkonzepts	311
17.3 Bewertung	316
17.4 Umsetzungsplanung	317
17.5 Empfehlungen	317
18 Datenqualitätsmanagement in der Spezifikation	319
18.1 Spezifikation der Schnittstellen	319
18.2 Definition der Rollen in der Datenorganisation	320
18.3 Festlegung der Datenqualitätsziele	322
18.4 Bezeichnung und Definition der Objekte	325
18.5 Festlegung der Geschäftsregeln	327
18.6 Messung der Qualität von Definitionen und Geschäftsregeln	329
18.7 Data Profiling in der Spezifikation	329
18.8 Entwurf des Systems	330
18.9 Empfehlungen	333
19 Datenqualitätsmaßnahmen in der Konstruktionsphase	335
19.1 Übertragung der Datenqualitätsziele	335
19.2 Konventionen und Richtlinien	336
19.3 Entwurf des Systems	337
19.4 Erstellung eines Prototypen	343
19.5 Empfehlungen	343
20 Steuerung der Datenqualität in der Realisierung	345
20.1 Einhaltung der Konventionen, Richtlinien und Konzepte	345
20.2 Data Profiling in der Realisierung	346
20.3 Einbindung der Datenverantwortlichen und Benutzer	347
20.4 Realisierung der Datenqualitätsmaßnahmen	348
20.5 Durchführung von Tests	349
20.6 Empfehlungen	349
21 Steuerung der Datenqualität im Betrieb	351
21.1 Monitoring und Berichtswesen	352
21.2 Ausbildung	353
21.3 Empfehlungen	353

Anhang	355
Abkürzungen	357
Literatur	359
Index	367