

# Inhalt

	Seite
Vorwort .....	9
Abkürzungen .....	11
Einleitung.....	13
<b>I. Theorie</b>	
I.1	Datenqualität ..... 18
I.1.1	Daten ..... 18
I.1.2	Qualität..... 19
I.1.3	Datenqualität ..... 21
I.1.4	Datenqualitätsmanagement ..... 25
I.1.5	Zusammenfassung..... 30
I.2	Ursachen und Ausprägungen schlechter Datenqualität ..... 31
I.2.1	Geschäftstreiber ..... 31
I.2.2	Ausprägungen schlechter Datenqualität ..... 34
I.2.3	Ursachen schlechter Datenqualität ..... 36
I.2.4	Beispiel: Finanzdienstleister ..... 41
I.2.5	Zusammenfassung..... 43
I.3	Auswirkungen schlechter Datenqualität ..... 44
I.3.1	Datenqualitätskosten ..... 44
I.3.2	Gesetzliche Anforderungen..... 48
I.3.3	Zusammenfassung..... 51
I.4	Organisation ..... 52
I.4.1	Aufbauorganisation ..... 52
I.4.2	Ablauforganisation ..... 59
I.4.3	Empfehlungen..... 62
I.5	Referenzarchitektur für Business-Intelligence-Anwendungen..... 63
I.5.1	Referenzarchitektur ..... 63
I.5.2	Problemstellen und Lösungsansätze hinsichtlich der Datenqualität..... 67
I.5.3	Architektur für Datenqualitätsmanagement ..... 71
I.5.4	Serviceorientierte Architektur..... 73
I.5.5	Master Data Management ..... 74
I.6	Kennzahlen zur Messung der Datenqualität ..... 79
I.6.1	Anwendungsmöglichkeiten von Kennzahlen ..... 80
I.6.2	Messpunkte für Datenqualität ..... 82
I.6.3	DQ-Metriken ..... 85
I.6.4	Beispiele für Kennzahlen je Datenqualitätskriterium..... 87

	Seite	
I.6.5	Kennzahlenbaum .....	89
I.6.6	DQ-Assessment.....	90
I.6.7	DQ-Planung.....	93
I.6.8	DQ-Projekte.....	94
I.6.9	Kennzahlenformular.....	94
I.6.10	Zusammenfassung.....	95

## II. Technische Umsetzung

	Einleitung.....	98
II.1	Verbesserung der Datenqualität im Quellsystem.....	101
	II.1.1 Korrektur fehlerhafter Daten .....	101
	II.1.2 Vorbeugung vor neuen Datenqualitätsproblemen .....	101
	II.1.3 Empfehlungen .....	106
II.2	Data Profiling.....	108
	II.2.1 Data-Profiling-Prozess .....	109
	II.2.2 Zusammensetzung des Data-Profiling-Teams .....	111
	II.2.3 Data-Profiling-Verfahren .....	112
	II.2.4 Empfehlungen.....	145
II.3	Erfolgreiche Datenvielfältigung und -filterung.....	147
	II.3.1 Validierung auf vier Ebenen.....	147
	II.3.2 Filterung fehlerhafter Daten.....	147
	II.3.3 Validierung bei Extraktion oder Laden .....	151
	II.3.4 Arten der Datenvielfältigung .....	153
	II.3.5 Erstellung der Validierungsregeln und Speicherung der Ergebnisse .....	155
	II.3.6 Empfehlungen.....	156
II.4	Standardisierung und Bereinigung.....	157
	II.4.1 Standardisierung vor der Bereinigung.....	157
	II.4.2 Datenbereinigung .....	159
	II.4.3 Standardisierung und Bereinigung im ETL-Prozess.....	179
	II.4.4 Verfahren für nicht zu bereinigende Daten.....	180
	II.4.5 Empfehlungen.....	181
II.5	Datenanreicherung .....	182
	II.5.1 Wirtschaftsinformationen .....	182
	II.5.2 Geografische Informationen .....	184
	II.5.3 Soziodemografische Informationen .....	185
	II.5.4 Haushaltsbildung .....	186
	II.5.5 Standards zur Klassifizierung von Waren und Dienstleistungen .....	187
	II.5.6 Branchen-Klassifizierung .....	190
	II.5.7 Empfehlungen.....	192

	Seite
II.6	Verbesserung der Datenqualität in der Bereitstellung und Präsentation.....194
II.6.1	Bereitstellung der Daten.....194
II.6.2	Präsentation der Information .....195
II.6.3	Empfehlungen.....207
II.7	Metadaten.....208
II.7.1	Metadaten: Begriff und Strukturierung .....208
II.7.2	Metadaten-Architekturen.....210
II.7.3	Erstellung von Metadaten .....212
II.7.4	Nutzung von Metadaten.....216
II.7.5	Empfehlungen.....218
II.8	Data Quality Monitoring .....219
II.8.1	DQ-Messung und -Analyse.....219
II.8.2	Methoden .....221
II.8.3	Visualisierung.....223
II.8.4	Benachrichtigung und Aktionen.....226
II.8.5	Verantwortlichkeiten .....226
II.8.6	Zusammenfassung.....227
II.9	Produktauswahl und -integration .....228
II.9.1	Anbieter und Produkte.....228
II.9.2	Auswahlkriterien im Überblick.....229
II.9.3	Funktionale Kriterien .....230
II.9.4	Integration .....233
II.9.5	Einbeziehung der Fachbereiche .....234
II.9.6	Sprachen und Länder .....235
II.9.7	Einbindung in DQM-Prozesse .....235
II.9.8	Empfehlungen .....235
III. Projektpraxis	
	Einleitung.....238
III.1	Datenqualitätsmanagement in einer Studie.....240
III.1.1	Analyse des Ist-Zustands .....240
III.1.2	Entwurf des Soll-Konzepts.....245
III.1.3	Bewertung .....248
III.1.4	Umsetzungsplanung .....248
III.1.5	Empfehlungen.....248

	Seite
III.2	Datenqualitätsmanagement in der Spezifikation..... 249
III.2.1	Spezifikation der Schnittstellen..... 249
III.2.2	Definition der Rollen in der Datenorganisation..... 250
III.2.3	Festlegung der Datenqualitätsziele ..... 250
III.2.4	Bezeichnung und Definition der Objekte ..... 252
III.2.5	Festlegung der Geschäftsregeln ..... 254
III.2.6	Messung der Qualität von Definitionen und Geschäftsregeln ..... 254
III.2.7	Data Profiling in der Spezifikation ..... 255
III.2.8	Entwurf des Systems ..... 255
III.2.9	Empfehlungen..... 256
III.3	Datenqualitätsmaßnahmen in der Konstruktionsphase ..... 257
III.3.1	Übertragung der Datenqualitätsziele..... 257
III.3.2	Konventionen und Richtlinien..... 257
III.3.3	Entwurf des Systems ..... 258
III.3.4	Erstellung eines Prototypen ..... 261
III.3.5	Empfehlungen..... 261
III.4	Steuerung der Datenqualität in der Realisierung..... 262
III.4.1	Einhaltung der Konventionen, Richtlinien und Konzepte ..... 262
III.4.2	Data Profiling in der Realisierung ..... 262
III.4.3	Einbindung der Datenverantwortlichen und Benutzer..... 263
III.4.4	Realisierung der Datenqualitätsmaßnahmen..... 263
III.4.5	Durchführung von Tests..... 264
III.4.6	Empfehlungen..... 264
III.5	Steuerung der Datenqualität im Betrieb ..... 266
III.5.1	Monitoring und Berichtswesen..... 266
III.5.2	Ausbildung ..... 268
III.5.3	Empfehlungen..... 268
Anhang	
Literatur.....	270
Die Autoren.....	275
Register.....	277