

# Inhaltsverzeichnis

<b>Teil I</b>	<b>Architektur</b>	<b>1</b>
<b>1</b>	<b>Abgrenzung und Einordnung</b>	<b>5</b>
1.1	Begriffliche Einordnung .....	6
1.1.1	Definitionen .....	7
1.1.2	Abgrenzung von transaktionalen Systemen .....	9
1.2	Historie des Themenbereichs .....	11
1.3	Einordnung und Abgrenzung von Business Intelligence .....	13
1.4	Verwendung von Data-Warehouse-Systemen .....	14
1.4.1	Anwendungsfälle .....	14
1.4.2	Wissenschaftliche Anwendungsbereiche .....	24
1.4.3	Technische Anwendungsbereiche .....	24
1.4.4	Betriebswirtschaftliche Anwendungsbereiche .....	25
1.5	Überblick über das Buch .....	31
1.5.1	<i>Star*Kauf</i> .....	31
1.5.2	Kapitelübersicht .....	33
<b>2</b>	<b>Referenzarchitektur</b>	<b>37</b>
2.1	Aspekte einer Referenzarchitektur .....	37
2.1.1	Referenzmodell für die Architektur von Data-Warehouse-Systemen .....	38
2.1.2	Beschreibung der Referenzarchitektur .....	40
2.2	Data-Warehouse-Manager .....	43
2.3	Datenquelle .....	45
2.3.1	Bestimmung der Datenquellen .....	45
2.3.2	Datenqualität .....	49
2.3.3	Klassifikation der Quelldaten .....	52

2.4	Monitor .....	54
2.5	Arbeitsbereich .....	55
2.6	Extraktionskomponente .....	56
2.7	Transformationskomponente .....	57
2.8	Ladekomponente .....	58
2.9	Basisdatenbank .....	58
	2.9.1 Charakterisierung, Aufgaben und Abgrenzung .....	59
	2.9.2 Aktualisierungsalternativen der Basisdatenbank .....	62
	2.9.3 Qualität der Daten in der Basisdatenbank .....	63
2.10	Ableitungsdatenbank .....	64
	2.10.1 Unterstützung des Ladeprozesses .....	65
	2.10.2 Unterstützung des Auswertungsprozesses .....	65
	2.10.3 Nabe-Speiche-Architektur .....	66
2.11	Auswertungsdatenbank .....	67
2.12	Auswertung .....	72
	2.12.1 Darstellungsformen .....	73
	2.12.2 Funktionalität .....	75
	2.12.3 Realisierung .....	77
	2.12.4 Plattformen .....	78
2.13	Repositorium .....	79
2.14	Metadatenmanager .....	82
2.15	Zusammenfassung .....	84
<b>3</b>	<b>Phasen des Data Warehousing</b>	<b>87</b>
3.1	Monitoring .....	87
	3.1.1 Realisierungen des Monitoring .....	88
	3.1.2 Monitoring-Techniken .....	89
3.2	Extraktionsphase .....	93
3.3	Transformationsphase .....	95
	3.3.1 Datenintegration .....	95
	3.3.2 Bereinigung .....	101
3.4	Ladephase .....	110

3.5	Auswertungsphase .....	113
3.5.1	Data Access .....	113
3.5.2	Online Analytical Processing (OLAP) .....	114
3.5.3	Data Mining .....	131
3.6	Zusammenfassung .....	141
<b>4</b>	<b>Physische Architektur</b> .....	<b>143</b>
4.1	Speicherarchitekturen für die Basis-, Ableitungs- oder Auswertungsdatenbank .....	143
4.1.1	Architektur eines Datenbankverwaltungssystems .....	144
4.1.2	Speichermodelle für Daten .....	144
4.2	Schichtenarchitekturen .....	146
4.2.1	Einschichtenarchitektur .....	148
4.2.2	Zweischichtenarchitektur .....	148
4.2.3	Dreischichtenarchitektur .....	150
4.2.4	N-Schichtenarchitektur .....	150
4.2.5	Webbasierte Architektur .....	151
4.3	Realtime-Data-Warehouse-Systeme .....	156
4.3.1	Anforderungen .....	156
4.3.2	Architektur .....	158
4.3.3	Aktualisierung der Daten .....	160
4.3.4	Berichte .....	162
4.4	Architektur für unstrukturierte Daten .....	163
4.4.1	Anforderungen .....	164
4.4.2	Architekturansätze .....	164
4.4.3	Datenbeschaffung .....	167
4.5	Neue Architekturansätze .....	173
4.5.1	Column Store .....	173
4.5.2	InMemory .....	174
4.5.3	Appliance-Datenbanksystem .....	174
4.6	Zusammenfassung .....	180

<b>Teil II</b>	<b>Entwicklung</b>	<b>181</b>
<b>5</b>	<b>Modellierung der Basisdatenbank</b>	<b>185</b>
5.1	Begriffsbestimmungen: Vom Modell zum Schema . . . . .	185
5.1.1	Modell . . . . .	185
5.1.2	Datenmodell und Schema . . . . .	186
5.2	Notwendigkeit eines übergreifenden Datenmodells . . . . .	188
5.2.1	Probleme beim Verzicht einer übergreifenden Modellierung . . . . .	188
5.2.2	Abgrenzung zur unternehmensweiten Modellierung . . . . .	189
5.3	Konzeptuelle Modellierung der Basisdatenbank . . . . .	191
5.3.1	Phasenmodell . . . . .	191
5.3.2	Kerndatenmodell . . . . .	192
5.3.3	Historisierung . . . . .	195
5.3.4	Referenzmodelle . . . . .	197
5.3.5	Langfristiger Lebenszyklus . . . . .	198
5.4	Zusammenfassung . . . . .	199
<b>6</b>	<b>Das multidimensionale Datenmodell</b>	<b>201</b>
6.1	Konzeptuelle Modellierung . . . . .	201
6.1.1	Verschiedene Vorgehensweisen zur Definition einer Methodik . . . . .	203
6.1.2	Vorstellung verschiedener Designnotationen . . . . .	205
6.2	Logische Modellierung . . . . .	214
6.2.1	Notwendigkeit der Formalisierung des multidimensionalen Modells . . . . .	214
6.2.2	Struktur des multidimensionalen Datenmodells . . . . .	216
6.2.3	Fehlende Werte in Würfelzellen (Nullwerte) . . . . .	220
6.2.4	Operatoren des multidimensionalen Modells . . . . .	221
6.2.5	Weitere Ansätze zur Formalisierung . . . . .	225
6.2.6	Grenzen und Erweiterungen des multidimensionalen Datenmodells . . . . .	227
6.3	Unterstützung von Veränderungen . . . . .	228
6.3.1	Zeitaspekte . . . . .	228
6.3.2	Aspekte der Klassifikationsveränderungen . . . . .	229
6.3.3	Aspekte der Schemaänderung . . . . .	233
6.4	Zusammenfassung . . . . .	240

<b>7</b>	<b>Umsetzung des multidimensionalen Datenmodells</b>	<b>241</b>
7.1	Relationale Speicherung	242
7.1.1	Abbildungsmöglichkeiten auf Relationen	242
7.1.2	Relationale Umsetzung multidimensionaler Anfragen	254
7.1.3	Relationale Versionierungs- und Evolutionsaspekte	260
7.2	Multidimensionale Speicherung	265
7.2.1	Datenstrukturen	266
7.2.2	Speicherung multidimensionaler Daten	275
7.2.3	Dateneingabe	279
7.2.4	Grenzen der multidimensionalen Datenhaltung	281
7.2.5	Hybride Speicherung: Hybrides OLAP (HOLAP)	283
7.3	Realisierung der Zugriffskontrolle	284
7.3.1	Zugriffskontrollanforderungen	284
7.3.2	Relationale Realisierung	287
7.3.3	Multidimensionale Realisierung	289
7.3.4	Inferenzen und Trackerangriffe	291
7.3.5	Realisierungskonzepte	292
7.4	Zusammenfassung	297
<b>8</b>	<b>Optimierung der Datenbank</b>	<b>299</b>
8.1	Anfragen im multidimensionalen Modell	300
8.2	Indexstrukturen	301
8.2.1	Überblick über Indexstrukturen	301
8.2.2	Eindimensionale Baumindexstrukturen	302
8.2.3	Mehrdimensionale Baumindexstrukturen	307
8.2.4	Bitmap-Indizes	310
8.2.5	Vergleich der Indizierungstechniken	313
8.3	Partitionierung	314
8.3.1	Horizontale Partitionierung	315
8.3.2	Vertikale Partitionierung	316
8.3.3	Partitionierungssteuerung	318
8.4	Relationale Optimierung von Star-Joins	319
8.5	Einsatz materialisierter Sichten	322
8.5.1	Verwendung materialisierter Sichten	323
8.5.2	Bestimmung des Auswertekontextes für Aggregatanfragen	325
8.5.3	Statische Auswahl materialisierter Sichten	326

8.5.4	Dynamische Auswahl materialisierter Sichten .....	329
8.5.5	Aktualisierung materialisierter Sichten .....	330
8.6	Optimierung eines multidimensionalen Datenbanksystems .....	332
8.6.1	Partitionierung .....	333
8.6.2	Speicherung der Zellen .....	335
8.6.3	Datenblockindizierung .....	336
8.7	Zusammenfassung .....	337
<b>9</b>	<b>Metadaten</b>	<b>339</b>
9.1	Metadaten und Metamodelle beim Data Warehousing .....	339
9.2	Metadatenmanagement .....	343
9.3	Metadatenmanagementsystem .....	345
9.3.1	Anforderungen an ein Metadatenmanagementsystem .....	346
9.3.2	Architektur .....	347
9.3.3	Repositorium- und Metadatenaustauschstandards .....	350
9.4	Data-Warehouse-Metadaten schemata .....	354
9.4.1	Eine Klassifikation für Metadaten .....	354
9.4.2	Standards und Referenzmodelle .....	357
9.5	Entwurf eines Schemas zur Verwaltung von Data-Warehouse-Metadaten .....	361
9.5.1	Funktionale Aspekte .....	362
9.5.2	Personen, Organisation und Aufgaben .....	364
9.5.3	Business-Metadaten .....	364
9.5.4	Abstraktionsstufen .....	365
9.6	Zusammenfassung .....	366
<b>Teil III</b>	<b>Anwendung</b>	<b>369</b>
<b>10</b>	<b>Vorgehensweise beim Aufbau eines Data-Warehouse-Systems</b>	<b>373</b>
10.1	Data-Warehouse-Strategie .....	374
10.1.1	IT-Strategie .....	374
10.1.2	Data-Warehouse-Strategie .....	376
10.1.3	Rolle des Data-Warehouse-Systems innerhalb der IT-Strategie .....	376
10.2	Reifegradmodell .....	377

10.3	Ableitung der Data-Warehouse-Architektur .....	387
10.3.1	Data-Warehouse-Rahmenwerk als gesamtheitliche Vorgabe .....	388
10.3.2	Umgang mit mehreren Data-Warehouse-Systemen .....	392
10.3.3	Data-Warehouse-Konsolidierung .....	395
10.3.4	Architekturüberlegungen in der Praxis .....	400
10.3.5	Umgebungen im Hinblick auf Entwicklung, Test, Produktion und Wartung .....	402
10.4	Data-Warehouse-Vorgehensweise .....	405
10.4.1	Grundsätzliche Überlegungen zum Projektvorgehen .....	405
10.4.2	Vorgehensmodell .....	410
10.4.3	Machbarkeitsbetrachtung zum Data Warehousing .....	411
10.4.4	Analysephase .....	413
10.4.5	Designphase .....	415
10.4.6	Implementierungsphase .....	420
10.4.7	Testmanagement .....	423
10.4.8	Vorgehensweisen bei der Einführung .....	427
10.5	Zusammenfassung .....	431
<b>11</b>	<b>Das Data-Warehouse-Projekt</b>	<b>433</b>
11.1	Data-Warehouse-Projektmanagement .....	433
11.1.1	Projektmanagement im Data-Warehouse-Projekt .....	434
11.1.2	Projektteam .....	437
11.1.3	Anforderungsmanagement .....	440
11.1.4	Qualitätsmanagement .....	448
11.1.5	Kommunikation .....	450
11.1.6	Konfliktmanagement .....	451
11.1.7	Dokumentation .....	453
11.1.8	Agiles Projektmanagement .....	453
11.2	Business Intelligence Competency Center (BICC) .....	458
11.2.1	Funktionen .....	459
11.2.2	Rollen und Kommunikation .....	460
11.2.3	Organisatorische Ausprägung und Verankerung .....	462
11.3	Softwareauswahl .....	462
11.3.1	Nutzen und Notwendigkeit der Produktauswahl .....	463
11.3.2	Klassifikation der Produkte anhand der Referenzarchitektur .....	464

11.3.3	Vorgehensweise zur Produktauswahl .....	467
11.3.4	Allgemeine Kriterien für die Produktauswahl .....	474
11.3.5	Kriterien für Datenbeschaffungswerkzeuge .....	475
11.3.6	Kriterien für OLAP-Produkte .....	480
11.3.7	Open-Source-Komponenten .....	486
11.4	Hardwareauswahl .....	490
11.4.1	Auswahlbestimmende Faktoren .....	491
11.4.2	Datenspeicherung .....	492
11.4.3	Archivspeichermedien .....	493
11.4.4	Multiprozessorsysteme .....	494
11.4.5	Fehlertoleranz als Planungsziel .....	497
11.4.6	Flaschenhalse und Fallstricke .....	498
11.4.7	Backup-Strategien und Notfallpläne .....	498
11.5	Erfolgsfaktoren beim Aufbau eines Data-Warehouse-Systems .....	500
11.5.1	Institutionelle Aufgaben des Projektmanagements: Projektorganisation .....	500
11.5.2	Funktionale Aufgaben des Projektmanagements: Projektabwicklung .....	502
11.5.3	Empfehlungen für ein Data-Warehouse-Projekt .....	504
11.6	Datenschutz und Datensicherheit .....	505
11.6.1	Datenschutz .....	506
11.6.2	Netzwerksicherheit .....	509
11.6.3	Benutzeridentifikation und Authentifizierung .....	512
11.6.4	Auditing .....	513
11.6.5	Autorisierung und Zugriffskontrolle .....	514
11.7	Wirtschaftlichkeitsbetrachtungen .....	517
11.7.1	Kostenbetrachtung .....	518
11.7.2	Nutzenbetrachtung .....	519
11.8	Zusammenfassung .....	523
<b>12</b>	<b>Betrieb und Weiterentwicklung eines Data-Warehouse-Systems</b>	<b>525</b>
12.1	Administration .....	525
12.1.1	Anforderungen und resultierende Aufgaben .....	526
12.1.2	Organisationsformen für Entwicklung und Betrieb .....	533
12.1.3	Rolle des Repositoriums .....	536
12.2	Datenbeschaffungsprozess .....	537



12.3	Performanz-Tuning von Data-Warehouse-Systemen . . . . .	544
12.3.1	Der Performanz-Tuning-Prozess . . . . .	544
12.3.2	Maßnahmen aus Sicht des Informationsmanagements . . . . .	545
12.3.3	Maßnahmen aus Sicht des Datenbankdesigns . . . . .	547
12.3.4	Maßnahmen aus Sicht der Applikationsumgebung . . . . .	550
12.3.5	Maßnahmen aus Sicht der Datenbankzugriffe . . . . .	551
12.3.6	Maßnahmen aus Sicht der Datenbankkonfiguration . . . . .	552
12.3.7	Maßnahmen aus Sicht des Betriebssystems . . . . .	555
12.3.8	Maßnahmen aus Sicht des Netzwerks . . . . .	556
12.3.9	Maßnahmen aus Sicht des Hardwaresystems . . . . .	556
12.3.10	Multicore-Architekturen . . . . .	557
12.4	Auswertungsprozess . . . . .	560
12.4.1	Schere zwischen Systemleistung und Anwendererwartungen . . . . .	561
12.4.2	Anwenderbetreuung . . . . .	564
12.5	Sicherungsmanagement . . . . .	565
12.5.1	Backup und Recovery . . . . .	566
12.5.2	Entsorgung von Daten . . . . .	567
12.5.3	Datenbank- und Systemverfügbarkeit . . . . .	570
12.5.4	Phasen eines Recovery-Plans . . . . .	571
12.6	Zusammenfassung . . . . .	572
<b>13</b>	<b>Praxisbeispiele</b>	<b>573</b>
13.1	Öffentliche Verwaltung . . . . .	574
13.1.1	Die Bundesagentur für Arbeit . . . . .	574
13.1.2	Data Warehousing in der öffentlichen Arbeitsverwaltung . . . . .	575
13.1.3	Fazit . . . . .	582
13.2	Versicherung . . . . .	583
13.2.1	Risikomanagement auf Basis eines Data-Ware- house-Systems in einem Versicherungskonzern . . . . .	583
13.2.2	Fazit . . . . .	589
13.3	Panelorientierte Marktforschung . . . . .	589
13.3.1	Die GfK-Gruppe und die GfK Retail and Technology GmbH . . . . .	590
13.3.2	Data Warehousing in der panelorientierten Marktforschung . . . . .	591
13.3.3	Fazit . . . . .	596

---

13.4	Online-Partnerbörse .....	597
13.4.1	Die FriendScout24 GmbH .....	597
13.4.2	Data Warehousing bei Online-Partnerbörsen .....	598
13.4.3	Fazit .....	607
13.5	Zusammenfassung .....	608
<b>Anhang</b>		<b>609</b>
<b>A</b>	<b>Abkürzungen</b>	<b>611</b>
<b>B</b>	<b>Glossar</b>	<b>615</b>
<b>C</b>	<b>Autorenverzeichnis</b>	<b>621</b>
<b>D</b>	<b>Autorenuordnung</b>	<b>633</b>
<b>E</b>	<b>Literatur und Webreferenzen</b>	<b>637</b>
	<b>Stichwortverzeichnis</b>	<b>677</b>