

# Auf einen Blick

<b>Einleitung</b> .....	<b>19</b>
<b>Teil I: Was ist ein Data Warehouse?</b> .....	<b>25</b>
<b>Kapitel 1:</b> Ein Beispiel zur Einführung .....	27
<b>Kapitel 2:</b> Das Data Warehouse im Umfeld der betrieblichen Informationssysteme ..	35
<b>Kapitel 3:</b> Definition und Abgrenzung des Begriffs »Data Warehouse« .....	47
<b>Teil II: Architektur eines Data-Warehouse-Systems</b> .....	<b>57</b>
<b>Kapitel 4:</b> Überblick über die Architektur eines Data-Warehouse-Systems .....	59
<b>Kapitel 5:</b> Der ETL-Prozess .....	69
<b>Kapitel 6:</b> Die Basisdatenbank .....	85
<b>Kapitel 7:</b> Das Analyse-Subsystem .....	93
<b>Kapitel 8:</b> Metadaten .....	113
<b>Teil III: Anwendungsbereiche für ein Data Warehouse</b> .....	<b>121</b>
<b>Kapitel 9:</b> Reporting .....	123
<b>Kapitel 10:</b> Online Analytical Processing .....	139
<b>Kapitel 11:</b> Data Mining .....	151
<b>Teil IV: Modellierung eines Data-Warehouse-Systems</b> .....	<b>177</b>
<b>Kapitel 12:</b> Data Vault .....	179
<b>Kapitel 13:</b> Semantischer Entwurf eines Data Warehouse .....	191
<b>Kapitel 14:</b> Relationale Modellierung der Datenwürfel .....	199
<b>Teil V: Zugriff auf ein Data Warehouse</b> .....	<b>213</b>
<b>Kapitel 15:</b> Multidimensionale Abfragen mit SQL .....	215
<b>Kapitel 16:</b> Die Abfragesprache MDX .....	229
<b>Kapitel 17:</b> Zusammenspiel von MDX und SQL .....	239
<b>Teil VI: Speicherung und Optimierung auf Datenbankebene</b> ..	<b>247</b>
<b>Kapitel 18:</b> ROLAP, MOLAP und anderes .....	249
<b>Kapitel 19:</b> Optimierungsmöglichkeiten bei relationalen Datenbanken .....	265

## 10 Auf einen Blick

<b>Teil VII: Der Top-10-Teil .....</b>	<b>279</b>
<b>Kapitel 20:</b> 10 Schritte auf dem Weg zu Ihrem ersten Dashboard .....	281
<b>Kapitel 21:</b> 10 Schritte, die helfen, die richtige Data-Warehouse-Software zu finden. . .	291
<b>Kapitel 22:</b> 10 Übungsaufgaben zur Wiederholung .....	297
<b>Literaturverzeichnis .....</b>	<b>313</b>
<b>Stichwortverzeichnis .....</b>	<b>317</b>

# Inhaltsverzeichnis

<b>Einleitung</b>	<b>19</b>
Über dieses Buch	19
Konventionen in diesem Buch	20
Was Sie nicht lesen müssen	20
Törichte Annahmen über den Leser	21
Wie dieses Buch aufgebaut ist	21
Teil I: Was ist ein Data Warehouse?	21
Teil II: Architektur eines Data-Warehouse-Systems	21
Teil III: Anwendungsbereiche für ein Data Warehouse	22
Teil IV: Modellierung eines Data-Warehouse-Systems	22
Teil V: Zugriff auf ein Data Warehouse	22
Teil VI: Speicherung und Optimierung auf Datenbankebene	22
Teil VII: Der Top-10-Teil	22
Symbole, die in diesem Buch verwendet werden	23
Wie es weitergeht	23
 <b>TEIL I</b>	
<b>WAS IST EIN DATA WAREHOUSE?</b>	<b>25</b>
 <b>Kapitel 1</b>	
<b>Ein Beispiel zur Einführung</b>	<b>27</b>
Daten und ihre Verarbeitung	27
Daten und Datenbanken	27
Die Verarbeitung von Daten	28
Analyse von Absatzmengen und Planzahlen als Beispiel	29
Besonderheiten analytischer Aufgabenstellungen	31
Wenn personenbezogene Daten ins Spiel kommen	34
 <b>Kapitel 2</b>	
<b>Das Data Warehouse im Umfeld</b>	
<b>der betrieblichen Informationssysteme</b>	<b>35</b>
Hierarchie betrieblicher Informationssysteme	35
Zusammenfassung: Analytische Informationssysteme	38
Beispiele für analytische Informationssysteme	39
Beispiel 1: Analytische Informationssysteme im CRM	39
Beispiel 2: Kennzahlen-Analysesysteme im Rechnungswesen	41
Beispiel 3: Website-Analysesysteme	43
Fazit: Data Warehouse und analytische Informationssysteme	45

## 12 Inhaltsverzeichnis

<b>Kapitel 3</b>	
<b>Definition und Abgrenzung des Begriffs »Data Warehouse«</b>	<b>47</b>
Die 3-Schichten-Architektur analytischer Informationssysteme	47
Definitionen des Begriffs Data Warehouse	50
Definition von Inmon	50
Definition von Kimball	52
Vergleich der beiden Definitionen	53
Anwendungsfall: Das Data Warehouse und Business Intelligence	54
 <b>TEIL II</b>	
<b>ARCHITEKTUR EINES DATA-WAREHOUSE-SYSTEMS</b>	<b>57</b>
 <b>Kapitel 4</b>	
<b>Überblick über die Architektur eines Data-Warehouse-Systems</b>	<b>59</b>
Die Phasen des Data Warehousing	59
Ein allgemeines Data-Warehouse-Architekturmodell	61
Vorgehensweisen bei der Erstellung eines Data Warehouse	64
Projektdefinition und Machbarkeitsstudie	65
Analyse, Entwurf und Einführung für einen Anwendungsbereich	66
 <b>Kapitel 5</b>	
<b>Der ETL-Prozess</b>	<b>69</b>
Überblick	69
Ein einführendes Beispiel	70
Extraktion	71
Das Pull-Prinzip	71
Das Push-Prinzip	72
Beispiele	72
Transformation	77
Datenbestandsanalyse	77
Datenbereinigung	78
Datenintegration	80
Laden	82
 <b>Kapitel 6</b>	
<b>Die Basisdatenbank</b>	<b>85</b>
Merkmale der Basisdatenbank	85
Unterschied zwischen operativen Datenbanken und der Basisdatenbank	87
Die operativen Quellsysteme des Beispiels	88
Die Basisdatenbank des Beispiels	89

<b>Kapitel 7</b>	
<b>Das Analyse-Subsystem .....</b>	<b>93</b>
Dimensionen und Fakten .....	93
Dimension oder Metrik? .....	95
Metriken als Dimension .....	96
Dimensionen als Metrik .....	97
Klassifizierung von Dimensionen .....	98
Fachliche Dimensionen .....	98
Kategorische Dimensionen .....	98
Strukturelle Dimensionen .....	99
Hierarchien von Dimensionswerten .....	99
Parallele Hierarchien .....	100
Unausgeglichene Hierarchieebäume .....	101
Strukturänderungen in Hierarchien .....	102
Slowly Changing Dimensions .....	102
Typ 1: Überschreiben .....	103
Typ 2: Neue Zeile .....	104
Typ 3: Spalten mit altem und neuem Wert .....	105
Typ 4: Mini-Dimension .....	105
Zusammenfassung .....	106
Verknüpfung von Dimensionen über Metriken .....	106
Aggregationstypen von Fakten .....	107
Die Themen Datenqualität und Datenschutz .....	108
Datenqualität .....	108
Datenschutz .....	109
Architekturvarianten für ein Analyse-Subsystem .....	109
Möglichkeiten für die Architektur .....	110
Die Hub-and-Spoke-Architektur .....	111
Auswertungen und Analysen .....	112
 <b>Kapitel 8</b>	
<b>Metadaten .....</b>	<b>113</b>
Was sind Metadaten? .....	113
Metadaten im Data-Warehouse-Kontext .....	114
Das Metadaten-Management in einem Data-Warehouse-System .....	114
Standards für Data-Warehouse-Metadaten .....	118
Ein kleines Beispiel .....	119
 <b>TEIL III</b>	
<b>ANWENDUNGSBEREICHE FÜR EIN DATA WAREHOUSE .....</b>	<b>121</b>
 <b>Kapitel 9</b>	
<b>Reporting .....</b>	<b>123</b>
Das Berichtswesen eines Unternehmens .....	123
Überblick und Definition .....	123
Erzeugung und Verteilung von Reports .....	125

## 14 Inhaltsverzeichnis

Arten von Berichtssystemen .....	125
Was sich Anwender vom Reporting wünschen und wie die Wirklichkeit oft aussieht .....	126
Einige Tipps für die Report-Gestaltung .....	127
Graphische Darstellungen im Report .....	128
Die Hichert-Success-Regeln .....	131
Grundformen für Reports .....	132
Ist-Ist-Vergleiche .....	132
Plan-Ist-Vergleiche .....	133
Plan-Wird-Vergleiche .....	134
Berücksichtigung dynamischer Dimensionsstrukturen .....	135
Report as-is .....	136
Report as-of .....	136
Report as-posted .....	137
Ein praktisches Beispiel .....	137

## **Kapitel 10 Online Analytical Processing .....**

**139**

Motivation und Definition .....	139
Charakteristika von OLAP .....	141
Abgrenzung OLAP und OLTP .....	141
Die Codd'schen Regeln .....	142
FASMI .....	143
Spezielle OLAP-Operatoren .....	144
Pivotierung bzw. Rotation .....	144
Roll-up und Drill-down .....	145
Slice und Dice .....	146
Beispiel .....	148

## **Kapitel 11 Data Mining .....**

**151**

Einführung .....	151
CRISP-DM .....	153
Methoden und Verfahren beim Data Mining .....	154
Assoziationsanalyse .....	155
Clusteranalyse .....	160
Klassifikation mit der Diskriminanzanalyse .....	164
Entscheidungsbaumverfahren .....	166
Spezielle Data-Mining-Fragestellungen im Kontext von Data-Warehouse-Daten ..	171
Welche Artikel werden gemeinsam gekauft? .....	172
Unterscheiden sich gute, normale und schlechte Kunden? .....	172
Welche Kunden besitzen eine bestimmte Produktaffinität? .....	173
Praxisbeispiel »Predictive Analytics« .....	174
Kollaboratives Filtern .....	175

**TEIL IV**  
**MODELLIERUNG EINES DATA-WAREHOUSE-SYSTEMS ..... 177**

**Kapitel 12**  
**Data Vault ..... 179**

Einführung .....	179
Hubs, Satelliten und Links .....	180
Hubs .....	180
Links .....	182
Satelliten .....	183
Beispiel .....	185

**Kapitel 13**  
**Semantischer Entwurf eines Data Warehouse ..... 191**

Zur Wiederholung: das Entity-Relationship-Modell .....	191
Drei Schritte bei der Modellierung einer Datenbank .....	192
Das ER-Modell: Entitätstypen, Attribute und Beziehungen .....	192
Das multidimensionale ER-Modell .....	194
ADAPT .....	196

**Kapitel 14**  
**Relationale Modellierung der Datenwürfel ..... 199**

Einführung .....	199
Das Star-Schema .....	200
Beispiel .....	201
Besondere Merkmale des Star-Schemas .....	204
Das Snowflake-Schema .....	207
Vergleich von Star- und Snowflake-Schema .....	209
Das Galaxy-Schema .....	211

**TEIL V**  
**ZUGRIFF AUF EIN DATA WAREHOUSE ..... 213**

**Kapitel 15**  
**Multidimensionale Abfragen mit SQL ..... 215**

Zugriff auf ein Data Warehouse mit SQL .....	215
Erzeugen der Tabellen .....	216
Typische analytische Fragestellungen .....	218
OLAP-Erweiterungen von SQL .....	220
Die WINDOW-Klausel .....	220
Erweiterungen der GROUP-BY-Option .....	225
Statistische Funktionen .....	228

## 16 Inhaltsverzeichnis

### **Kapitel 16** **Die Abfragesprache MDX ..... 229**

Einführung .....	229
Spezielle OLAP-Operatoren und Funktionen .....	233
Tupel und Sets .....	233
Member und Children .....	234
Kreuzprodukt mittels Crossjoin .....	234
Der WITH-Operator .....	235
Häufige Fragestellungen .....	236

### **Kapitel 17** **Zusammenspiel von MDX und SQL ..... 239**

OLAP-Server .....	239
Der OLAP-Server Mondrian .....	241
MDX-Schema von Mondrian .....	241
Mondrian-Frontend-Tools .....	245

## **TEIL VI** **SPEICHERUNG UND OPTIMIERUNG AUF DATENBANKEBENE... 247**

### **Kapitel 18** **ROLAP, MOLAP und anderes ..... 249**

ROLAP und MOLAP .....	249
Spaltenorientierte und In-Memory-Speicherung .....	252
NoSQL-Datenbanksysteme .....	255
Typen von NoSQL-Systemen .....	255
NoSQL-Datenbanken bei einem Data Warehouse .....	258
Beurteilung .....	263

### **Kapitel 19** **Optimierungsmöglichkeiten bei relationalen Datenbanken .. 265**

Einführung .....	265
Partitionierung .....	266
Partition by List .....	267
Partition by Range .....	268
Partition by Hash .....	268
Partition by Reference .....	269
Materialized Views .....	270
Klassische Views vs. Materialized Views .....	270
Materialized Views bei einem Data Warehouse .....	273
Indizierung .....	274
Klassischer Index .....	274
Bitmap-Index .....	275
Mehrdimensionale Indizes .....	276



<b>TEIL VII</b>	
<b>DER TOP-10-TEIL</b>	<b>279</b>
<b>Kapitel 20</b>	
<b>10 Schritte auf dem Weg zu Ihrem ersten Dashboard</b>	<b>281</b>
Und so wird es gemacht ...	282
Festlegung der Datenquellen	282
Vorbereitung der Daten	283
Erstellung eines Dashboards	285
Daten aus mehreren Quellen	287
Integration von Landkarten	288
<b>Kapitel 21</b>	
<b>10 Schritte, die helfen, die richtige Data-Warehouse-Software zu finden</b>	<b>291</b>
Marktanalyse für BI-Software	291
Definition der eigenen Anforderungen	292
Einbindung des Managements, Projektplan	293
Marktanalyse der infrage kommenden BI-Anbieter	293
Einholung von Angeboten	293
Durchführung von Testinstallationen	294
Bewertung der Systeme	294
Ermittlung der Kosten	295
Einholung von Referenzen, Anbieterqualifikation	296
Überprüfung der Lizenzvereinbarung	296
<b>Kapitel 22</b>	
<b>10 Übungsaufgaben zur Wiederholung</b>	<b>297</b>
Aufgaben	297
Aufgabe 1: Assoziationsanalyse	297
Aufgabe 2: Diskriminanzanalyse	297
Aufgabe 3: Data Vault	298
Aufgabe 4: ADAPT	298
Aufgabe 5: MDX	299
Aufgabe 6: Star-Schema	299
Aufgabe 7: OLAP mit SQL	299
Aufgabe 8: Snowflake-Schema	300
Aufgabe 9: Optimierung	300
Aufgabe 10: Multidimensionale Datenbank	301
Lösungen	301
Lösung von Aufgabe 1	301
Lösung von Aufgabe 2	303
Lösung von Aufgabe 3	303
Lösung von Aufgabe 4	304
Lösung von Aufgabe 5	304

## 18 Inhaltsverzeichnis

Lösung von Aufgabe 6 .....	306
Lösung von Aufgabe 7 .....	307
Lösung von Aufgabe 8 .....	309
Lösung von Aufgabe 9 .....	309
Lösung von Aufgabe 10 .....	311
<b>Literaturverzeichnis .....</b>	<b>313</b>
<b>Stichwortverzeichnis .....</b>	<b>317</b>