

Frank Geisler

Datenbanken Grundlagen und Design

5., aktualisierte und erweiterte Auflage



Inhaltsverzeichnis

	Vorwort zur 5. Auflage	15
	Über den Autor	16
Teil I	Grundlagen	17
1	Einführung in das Thema Datenbanken	19
1.1	Warum ist Datenbankdesign wichtig?	24
1.2	Dateisystem und Datenbanken	26
1.2.1	Historische Wurzeln	27
1.2.2	Probleme bei der Datenhaltung im Dateisystem	28
1.2.3	Datenredundanzen und Anomalien	35
1.3	Das Fallbeispiel	37
1.4	Zusammenfassung	38
1.5	Aufgaben	41
1.5.1	Wiederholung	41
1.5.2	Zum Weiterdenken	42
2	Datenbanksysteme, Datenbankanwendungen und Middleware	43
2.1	Datenbanksysteme	43
2.2	Verschiedene Arten von Datenbanksystemen	47
2.3	DBMS-Funktionen	48
2.4	Datenbankmodelle	51
2.4.1	Hierarchische Datenbanken	52
2.4.2	Netzwerk-Datenbanken	57
2.4.3	Relationale Datenbanken	59
2.4.4	ER-Datenbankmodelle	65
2.4.5	Objektorientierte Datenbanken	68
2.5	Datenbankanwendungen	73
2.5.1	Einschichtige Datenbankanwendungen	74
2.5.2	Zweischichtige Datenbankanwendungen	74
2.5.3	N-schichtige Datenbankanwendungen	76
2.6	Middleware	77
2.6.1	ODBC	78

	2.6.2	ADO	79
	2.6.3	ADO.NET	81
2.7		Zusammenfassung	82
2.8		Aufgaben	90
	2.8.1	Wiederholung	90
	2.8.2	Zum Weiterdenken	90
3		Das relationale Datenbankmodell	93
3.1		Entitäten und Attribute	94
3.2		Tabellen	95
3.3		Schlüssel	98
3.4		Relationale Operatoren	104
	3.4.1	DIFFERENCE	104
	3.4.2	DIVIDE	105
	3.4.3	INTERSECT	105
	3.4.4	JOIN	106
	3.4.5	PRODUCT	108
	3.4.6	PROJECT	109
	3.4.7	SELECT	110
	3.4.8	UNION	110
3.5		Beziehungen innerhalb der Datenbank	111
	3.5.1	1:1-Beziehung	111
	3.5.2	1:N-Beziehung	113
	3.5.3	M:N-Beziehung	114
	3.5.4	Optionale und nicht-optionale Beziehungen	115
	3.5.5	Primär-/Fremdschlüssel und Datenredundanzen	115
3.6		Metadaten	116
3.7		Indizes	119
3.8		Zusammenfassung	120
3.9		Aufgaben	125
	3.9.1	Wiederholung	125
	3.9.2	Zum Weiterdenken	125
Teil II Datenbankdesign und Implementierung			129
4		ER-Datenbankmodellierung	131
4.1		Datenmodelle und Abstraktion	132
	4.1.1	Das konzeptionelle Modell	133
	4.1.2	Das interne Modell	135

4.1.3	Das externe Modell	136
4.1.4	Das physikalische Modell	137
4.2	Das Entity-Relationship-Modell	138
4.2.1	Entitäten	138
4.2.2	Attribute	139
4.2.3	Primärschlüssel	145
4.2.4	Beziehungen	147
4.3	Erstellen eines ER-Diagramms	162
4.4	Zusammenfassung	170
4.5	Aufgaben	173
4.5.1	Wiederholung	174
4.5.2	Zum Weiterdenken	174
5	Normalisierung	177
5.1	Warum Normalisierung?	177
5.1.1	Das Normalisierungsbeispiel	179
5.1.2	Erste Normalform	180
5.1.3	Zweite Normalform	184
5.1.4	Dritte Normalform	186
5.1.5	Boyce-Codd-Normalform (BCNF)	191
5.1.6	Höhere Normalformen	192
5.2	Normalisierung und Datenbankdesign	196
5.3	Denormalisierung	199
5.4	Zusammenfassung	200
5.5	Aufgaben	202
5.5.1	Wiederholung	202
5.5.2	Zum Weiterdenken	203
6	SQL-Grundlagen	205
6.1	Einführung	206
6.1.1	Historischer Überblick	207
6.1.2	Datentypen	208
6.1.3	Die SQL-Komponenten	210
6.1.4	Logische Verknüpfungen	216
6.2	Daten mit SQL abfragen	221
6.2.1	Einfache Abfragen	222
6.2.2	Tabellen verknüpfen mit Joins	243
6.2.3	Verschachtelte Abfragen	253
6.2.4	Sichten	254

6.3	Daten mit SQL verändern	258
6.3.1	INSERT	258
6.3.2	UPDATE	259
6.3.3	DELETE	260
6.4	Weitere wichtige SQL-Befehle	261
6.4.1	Mengenfunktionen	261
6.4.2	Stringfunktionen	266
6.4.3	Numerische Funktionen	268
6.4.4	Datetime-Funktionen	269
6.5	Zusammenfassung	271
6.6	Aufgaben	275
6.6.1	Wiederholung	275
6.6.2	Zum Weiterdenken	275
Teil III Weiterführende Themen		277
7	Projektablauf bei der Erstellung einer Datenbank	279
7.1	Der System Development Life Cycle	281
7.1.1	Planung	282
7.1.2	Analyse	284
7.1.3	System-Design	286
7.1.4	Implementierung	287
7.1.5	Wartung	289
7.2	Der Datenbank-Lebenszyklus	289
7.2.1	Grundlegende Analyse	290
7.2.2	Datenbankdesign	295
7.2.3	Implementierung und Datenimport	307
7.2.4	Test und Evaluierung	309
7.2.5	Betrieb	309
7.2.6	Wartung und Evolution	310
7.3	Zusammenfassung	310
7.4	Aufgaben	312
7.4.1	Wiederholung	312
8	Transaktionen und konkurrierende Zugriffe	313
8.1	Was ist eine Transaktion?	313
8.1.1	Eigenschaften einer Transaktion	317
8.1.2	Transaktionsverwaltung mit SQL	318

8.1.3	Das Transaktionsprotokoll	320
8.2	Konkurrierende Zugriffe	322
8.2.1	Lost Updates	322
8.2.2	Dirty Read	323
8.2.3	Nonrepeatable Read	324
8.2.4	Phantome	325
8.3	Sperrmechanismen (Locks)	326
8.3.1	Granularität	327
8.3.2	Sperrtypen	332
8.3.3	Zwei-Phasen-Locking	333
8.3.4	Deadlocks	334
8.4	Zusammenfassung	336
8.5	Aufgaben	339
8.5.1	Wiederholung	339
8.5.2	Zum Weiterdenken	340
9	Die Client-Server-Architektur	341
9.1	Was ist Client-Server?	341
9.1.1	Geschichte von Client-Server	342
9.1.2	Vorteile von Client-Server	343
9.2	Client-Server-Architektur	346
9.2.1	Client-Komponenten	349
9.2.2	Server-Komponenten	350
9.2.3	Middleware	351
9.2.4	Netzwerk-Protokolle	355
9.3	Zusammenfassung	361
9.4	Aufgaben	364
9.4.1	Wiederholung	364
10	Verteilte Datenbanksysteme	365
10.1	Vor- und Nachteile verteilter Datenbanksysteme	366
10.2	Verteilte Datenverarbeitung vs. verteilte Datenbanken	371
10.3	Komponenten eines verteilten Datenbanksystems	373
10.4	Transparenz beim Datenzugriff	376
10.4.1	Transparente Datenverteilung	377
10.4.2	Transparentes Transaktionsmanagement	382
10.5	Datenfragmentierung	385
10.6	Replikation	388
10.7	Zusammenfassung	389

10.8	Aufgaben	392
10.8.1	Wiederholung	392
11	Data Warehouses	393
11.1	Die Notwendigkeit der Datenanalyse	395
11.2	Decision-Support-Systeme	396
11.2.1	Der Unterschied zwischen operationalen Daten und DSS-Daten	399
11.2.2	Anforderungen an eine DSS-Datenbank	407
11.3	Das Data Warehouse	410
11.3.1	Data-Marts	414
11.3.2	Zwölf Eigenschaften, an denen man ein Data Warehouse erkennen kann	415
11.4	OLAP (Online Analytical Processing)	416
11.4.1	OLAP-Architekturen	421
11.4.2	Relationales OLAP (ROLAP)	425
11.4.3	Multidimensionales OLAP (MOLAP)	430
11.5	Das Sternschema	433
11.5.1	Fakten	433
11.5.2	Dimensionen	434
11.5.3	Attribute	435
11.5.4	Attribut-Hierarchien	437
11.5.5	Sternschemata in der Praxis	438
11.5.6	Techniken zur Erhöhung der Performance	439
11.6	Das Snowflake-Schema	442
11.7	Slowly changing Dimensions	444
11.7.1	Typ 1	445
11.7.2	Typ 2	446
11.7.3	Typ 3	449
11.7.4	Typ 4	450
11.7.5	Typ 6/Hybrid	451
11.8	Zusammenfassung	452
11.9	Aufgaben	453
11.9.1	Wiederholung	453
12	Data-Mining	455
12.1	Der Data-Mining-Prozess	458
12.1.1	Das Problem definieren	459

12.1.2	Daten vorbereiten	459
12.1.3	Die Daten sichten	460
12.1.4	Ein Data-Mining-Modell definieren	461
12.1.5	Data-Mining betreiben	463
12.1.6	Die Ergebnisse zur Verfügung stellen	464
12.2	Zusammenfassung	464
12.3	Aufgaben	465
12.3.1	Wiederholung	465
13	LINQ	467
13.1	Unverträglichkeit zwischen Relationen und Objekten	467
13.1.1	Das Problem, Objekte auf Tabellen abzubilden	468
13.1.2	Wem gehört das Schema?	472
13.1.3	Das Doppel-Schema-Problem	473
13.1.4	Identitätsprobleme bei Entitäten	473
13.1.5	Rückgewinnung der Daten	475
13.2	Die Architektur von LINQ	477
13.3	Spracherweiterungen, die LINQ ermöglichen, am Beispiel von C#	481
13.3.1	Anonyme Typen	481
13.3.2	Objekt-Initialisierer	483
13.3.3	Collection-Initialisierer	483
13.3.4	Partielle Methoden	484
13.3.5	Implizit deklarierte lokale Variablen	485
13.3.6	Erweiterungsmethoden	487
13.3.7	Lambda-Ausdrücke	487
13.3.8	Abfrage-Ausdrücke	490
13.4	Aufgaben	492
13.4.1	Wiederholung	492
14	Big Data	495
14.1	Strukturierte, semistrukturierte und unstrukturierte Daten	495
14.2	Die Evolution der Datenverarbeitung	498
14.2.1	Datenstrukturen erstellen	501
14.2.2	Data Warehouses, Datamarts und BLOBs	501
14.2.3	Content-Management-Systeme	502
14.2.4	Die dritte Stufe der Evolution	502
14.3	Was genau ist eigentlich Big Data?	503

14.4	Der Big-Data-Projektzyklus	504
14.5	Die Architektur eines Big-Data-Projekts	506
14.6	Map Reduce	511
14.7	Big Table	511
14.8	Hadoop	511
14.9	Aufgaben	511
14.9.1	Wiederholung	512
A	Lösungen zu den Wiederholungsaufgaben	513
A.1	Kapitel 1	513
A.2	Kapitel 2	516
A.3	Kapitel 3	518
A.4	Kapitel 4	521
A.5	Kapitel 5	523
A.6	Kapitel 6	525
A.7	Kapitel 7	532
A.8	Kapitel 8	534
A.9	Kapitel 9	536
A.10	Kapitel 10	538
A.11	Kapitel 11	540
A.12	Kapitel 12	542
A.13	Kapitel 13	543
A.14	Kapitel 14	545
	Stichwortverzeichnis	547