

HANSER

Taschenbuch Datenbanken

Herausgegeben von Thomas Kudraß

ISBN-10: 3-446-40944-0

ISBN-13: 978-3-446-40944-6

Inhaltsverzeichnis

Weitere Informationen oder Bestellungen unter
<http://www.hanser.de/978-3-446-40944-6>
sowie im Buchhandel

Inhaltsverzeichnis

1	Datenbanken: Grundlagen und Überblick	19
1.1	Dateien und Dateisysteme	19
1.2	Terminologie	20
1.3	Merkmale eines DBMS	23
1.3.1	Aufgaben eines DBMS	23
1.3.2	Vorteile des Datenbankeinsatzes	24
1.3.3	Nachteile von Datenbanksystemen	25
1.3.4	Produkte	26
1.4	Architektur eines Datenbanksystems	27
1.4.1	Architekturen	27
1.4.2	Schemaarchitektur	27
1.4.3	Datenunabhängigkeit	30
1.5	Benutzerrollen bei Entwurf und Betrieb von Datenbanken	31
1.6	Datenbanken und Informationssysteme	32
1.7	Fachgebiet Datenbanken im Überblick	34
1.7.1	Themenbereiche und Zusammenhänge	34
1.7.2	Einordnung des Fachs innerhalb der Informatik	36
1.8	Historische Entwicklung	37
1.8.1	Frühzeit	38
1.8.2	Prärelationale DBMS	39
1.8.3	Durchbruch der relationalen Datenbanken	39
1.8.4	Neue Anwendungsfelder für Datenbanken	40
1.8.5	Neuzeit	41
1.9	Datenbanken in der Forschung	42
1.9.1	Fachverbände	42
1.9.2	Aktuelle Forschungstrends	42
2	Informationsmodellierung	45
2.1	Datenbankentwurf	45
2.1.1	Phasenmodell	45
2.1.2	Anforderungsanalyse	46
2.1.3	Konzeptioneller Entwurf	47
2.1.4	Logischer Entwurf	48
2.1.5	Datendefinition	49
2.1.6	Physischer Entwurf	50
2.1.7	Implementierung und Wartung	50

2.2	Grundlagen des Entity-Relationship-Modells (ERM) . . .	50
2.2.1	Semantische Datenmodelle	51
2.2.2	Grundkonzepte des klassischen ERM	51
2.2.3	Kardinalitäten von Beziehungstypen	55
2.2.4	Existenzabhängigkeit vs. Optionalität	57
2.2.5	Rekursive und n -äre Beziehungstypen	57
2.2.6	Attribute	59
2.2.7	Modellierungsbeispiel	60
2.3	Erweiterungen des ERM	62
2.3.1	Erweiterungen bei Attributen	62
2.3.2	Generalisierung und Spezialisierung	63
2.3.3	Aggregation	66
2.4	Datenmodellierung mit dem UML-Klassendiagramm . .	67
2.4.1	Klassen und Attribute	67
2.4.2	Assoziationsbeziehungen	69
2.4.3	Aggregations- und Kompositionsbeziehungen .	71
2.4.4	Generalisierung und Spezialisierung	71
3	Relationales Datenmodell	73
3.1	Konzepte und Grundbegriffe des relationalen Datenmodells	73
3.1.1	Relationen, Tupel, Attribute und Wertebereiche	73
3.1.2	Eigenschaften von Relationen	75
3.2	Integritätsbedingungen	76
3.2.1	Grundbegriffe	76
3.2.2	Entitätsintegrität	77
3.2.3	Referenzielle Integrität	77
3.3	Abbildung des EERM auf das relationale Datenmodell .	81
3.3.1	Problemstellung	81
3.3.2	Abbildungsregeln für Attribute und Entitytypen	82
3.3.3	Abbildungsregeln für Beziehungstypen	83
3.3.4	Abbildungsregeln für die Generalisierung	85
3.4	Optimierung von Relationen	87
3.4.1	Problemstellung	87
3.4.2	Anomalien bei DML-Operationen auf Relationen	87
3.4.3	Abhängigkeiten	89
3.4.4	Verbundtreue und Abhängigkeitstreue	91
3.4.5	Normalformenlehre	92

3.5	Operationen der Relationenalgebra	99
3.5.1	Einführung	99
3.5.2	Relationenorientierte Operationen	100
3.5.3	Mengenoperationen	102
3.5.4	Relationenalgebra und relationale Sprachen ...	104
3.5.5	Relationenkalkül	105
4	Die Datenbanksprache SQL	106
4.1	Grundkonzepte	106
4.2	Historie	108
4.3	Spezifikationsdokumente	109
4.4	Beispieldatenbank	110
4.5	Datenbankanfragen	111
4.5.1	Einführende Beispielanfragen	111
4.5.2	Grundgerüst von Anfragen	112
4.5.3	Anfragen mit Aggregatfunktionen	114
4.5.4	Anfragen mit Tabellenfunktionen	115
4.5.5	Anfragen mit Mengenoperationen	116
4.5.6	Anfragen mit Verbundoperationen	117
4.5.7	Verschachtelte Anfragen	120
4.5.8	Rekursive Anfragen	122
4.5.9	Anfragen mit sortierter Ausgabe	123
4.6	Datenmanipulation	124
4.6.1	Einfügen von Tabellenzeilen	124
4.6.2	Ändern von Tabellenzeilen	125
4.6.3	Löschen von Tabellenzeilen	126
4.6.4	Zusammenführen von Tabellenzeilen	126
4.7	Datendefinition	127
4.7.1	SQL-Datentypen	127
4.7.2	Erzeugen und Löschen von Schemata	132
4.7.3	Erzeugen, Ändern und Löschen von Tabellen ..	132
4.7.4	Erzeugen und Löschen von Domänen	135
4.7.5	Erzeugen und Löschen von Integritätsbedingungen	136
4.7.6	Erzeugen und Löschen von Sichten	138
4.7.7	Erzeugen und Löschen von Routinen	139
4.7.8	Erzeugen und Löschen von Triggern	145
4.7.9	Erzeugen und Löschen von Sequenzgeneratoren	147
4.8	Transaktionssteuerung	147
4.8.1	Transaktionsanweisungen	147
4.8.2	Isolationsebenen	149

	4.8.3	Überprüfung von Integritätsbedingungen	150
4.9		Zugriffskontrolle	150
	4.9.1	Vergabe von Zugriffsrechten	150
	4.9.2	Zurücknahme von Zugriffsrechten	151
	4.9.3	Erzeugen und Löschen von Rollen	152
5		Datenbank-Anwendungsprogrammierung	153
5.1		Grundlagen der Datenbank- Anwendungsprogrammierung	153
	5.1.1	Impedance Mismatch	153
	5.1.2	Einbettungstechniken	154
	5.1.3	Einbettungsarten	154
	5.1.4	Architekturansätze	155
5.2		Embedded SQL	159
	5.2.1	Grundidee und Architektur	159
	5.2.2	Syntax für die Einbettung	159
	5.2.3	Host-Variablen	159
	5.2.4	Das Cursor-Konzept	161
	5.2.5	Statische Einbettung	162
	5.2.6	Dynamische Einbettung	163
	5.2.7	Indikatorvariablen	164
	5.2.8	SQLJ	164
5.3		Aufrufschnittstellen	166
	5.3.1	Überblick	166
	5.3.2	SQL/CLI	166
	5.3.3	JDBC	168
5.4		Relationale Datenbankprogrammiersprachen	174
	5.4.1	Datenbankinterne Ansätze	174
	5.4.2	4GL-Programmiersprachen	181
5.5		OO-DB-Frameworks	183
	5.5.1	Java EE	185
	5.5.2	JDO	194
6		Datenbanken im Web	198
6.1		Grundlagen des Web	198
6.2		Eigenschaften von Webanwendungen	201
	6.2.1	Anforderungen	201
	6.2.2	Webseiten	202
	6.2.3	Dynamische Webseiten	205
	6.2.4	Adressierung	207
	6.2.5	Kommunikation	208

6.3	Datenbankanbindung im Web	211
	6.3.1 Architekturen	211
	6.3.2 Programmierung	213
	6.3.3 Überblick	215
6.4	Datenbankanbindung über Datenexport	216
6.5	Clientseitige Datenbankanbindung	217
	6.5.1 Skripteinbettung	218
	6.5.2 Programmeinbettung	220
	6.5.3 Weitere Techniken	223
6.6	Serverseitige Datenbankanbindung	224
	6.6.1 Webseitengeneratoren	224
	6.6.2 Skripteinbettung	227
	6.6.3 Programmeinbettung	231
	6.6.4 Applikationsserver	233
	6.6.5 Webservices	236
7	Komponenten eines Datenbankmanagementsystems	239
7.1	Architektur von DBMS	239
	7.1.1 Schichtenmodell	239
	7.1.2 Prozessarchitektur	241
7.2	Pufferverwaltung	242
	7.2.1 Notwendigkeit und Aufgabe	243
	7.2.2 Speicherzuteilung	244
	7.2.3 Seitenersetzung	245
7.3	Speicher- und Zugriffssystem	247
	7.3.1 Aufgabe	247
	7.3.2 Seiten und Sätze	247
	7.3.3 Adressierung von Sätzen	249
7.4	Anfrageprozessor	251
	7.4.1 Basisoperatoren	251
	7.4.2 Anfrageplanung und -optimierung	255
	7.4.3 Kosten und Statistiken	260
7.5	Transaktionsverwaltung	262
	7.5.1 Aufgabe	262
	7.5.2 Serialisierbarkeit	264
	7.5.3 Sperrverfahren	266
	7.5.4 Nicht sperrende Verfahren	270
7.6	Recovery	272
	7.6.1 Fehlerklassen	273
	7.6.2 Recovery-Strategien	274
	7.6.3 Logging	274

	7.6.4	Wiederanlauf im Fehlerfall	278
	7.6.5	Schattenspeicherverfahren	279
8		Dateiorganisation und Indexe	281
	8.1	Organisation von Dateien	281
		8.1.1 Formen der Dateiorganisation	281
		8.1.2 Dateiorganisationsformen im Vergleich	281
	8.2	Zugriffsstrukturen	284
		8.2.1 Grundlagen von Zugriffsverfahren	284
		8.2.2 Eigenschaften von Indexen	285
	8.3	Baumbasierte Verfahren	286
		8.3.1 ISAM-Bäume	287
		8.3.2 Balancierte Mehrwegbäume	288
		8.3.3 Digitale Bäume	290
	8.4	Hash-Verfahren	291
		8.4.1 Prinzip des Hashing	291
		8.4.2 Erweiterbares Hashing	291
		8.4.3 Weitere Hash-Verfahren	292
	8.5	Mehrdimensionale Zugriffsverfahren	293
		8.5.1 Mehrdimensionale Baumverfahren	294
		8.5.2 Grid-File	294
	8.6	Clustering und Partitionierung	296
		8.6.1 Clustering	296
		8.6.2 Partitionierung	296
	8.7	Umsetzung in SQL-Systemen	297
		8.7.1 Definition von Tabellen	298
		8.7.2 Definition von Indexen	298
9		Optimierung von Datenbanken und Leistungsbewertung	300
	9.1	Motivation der Datenbankoptimierung	300
		9.1.1 Kosten von Datenbankanfragen	300
		9.1.2 Optimierungspotenzial	301
		9.1.3 Zielbestimmung der Datenbankoptimierung ...	303
	9.2	Phasen der Datenbankoptimierung	304
	9.3	Phase 1.1 – Optimierung des Datenbankschemas	306
		9.3.1 Konzeptuelles Schema	307
		9.3.2 Externes Schema	310
		9.3.3 Internes Schema	311
	9.4	Phase 1.2 – Anwendungsoptimierung	314
		9.4.1 Optimierung von Unternehmensfunktionen ...	314
		9.4.2 Optimierung der Anwendung	314

9.4.3	Optimierungen im Mehrbenutzerbetrieb	316
9.4.4	Formulierung von SQL-Anweisungen	317
9.5	Phase 2 – Hauptspeicheroptimierung	320
9.5.1	Gestaltung des Datenbankpuffers	320
9.5.2	Schreiben des Datenbankpuffers	323
9.5.3	Optimierer	324
9.6	Phase 3 – Optimierung der Sekundärspeicherzugriffe . . .	326
9.6.1	Zusammenspiel mit dem Betriebssystem	326
9.6.2	Verteilung der Eingabe-/Ausgabelast	327
9.6.3	Optimierung physischer Speicherstrukturen . . .	329
9.7	Leistungsbewertung	332
9.7.1	Transaction Processing Performance Council . .	333
9.7.2	TPC-C	334
9.7.3	TPC-H/TPC-R	335
9.7.4	TPC-W	336
9.7.5	TPC-App	337
10	Objektrelationale Datenbanken	338
10.1	Objektorientierte Konzepte	338
10.1.1	Objekte	338
10.1.2	Methoden	339
10.1.3	Kapselung	339
10.1.4	Objektidentität	339
10.1.5	Klassen	340
10.1.6	Spezialisierung	341
10.2	Objektorientierung in Datenbanken	342
10.3	Objektrelationale Standard-SQL-Konzepte	343
10.3.1	Typkonstruktoren	344
10.3.2	Distinct-Typen	346
10.3.3	Strukturierte Typen	346
10.3.4	Methoden	349
10.3.5	Benutzerdefinierte Konstruktoren	352
10.3.6	Benutzerdefinierte Casts	353
10.3.7	Benutzerdefinierte Ordnungen	353
10.3.8	Typisierte Tabellen	354
10.3.9	Typisierte Sichten	356
10.4	Objektrelationale Anfragen	357
10.4.1	Anfragen auf Kollektionen	358
10.4.2	Anfragen mit Pfadausdrücken	359
10.4.3	Anfragen mit Methodenaufrufen	359
10.4.4	Anfragen auf flachen Tabellenextensionen	359

	10.4.5	Typspezifische Anfragen	359
	10.4.6	Anfragen mit temporärer Typanpassung	360
11		XML und Datenbanken	361
	11.1	Überblick über XML	361
	11.1.1	Der XML-Standard und verwandte Standards	361
	11.1.2	XML Schema	363
	11.2	Anfragesprachen für XML	366
	11.2.1	Pfadausdrücke (XPath und XQuery)	367
	11.2.2	XQuery	368
	11.3	XML und relationale Datenbanksysteme	372
	11.3.1	Relationale Speicherung von XML	373
	11.3.2	SQL/XML	375
	11.3.3	Realisierung in kommerziellen Systemen	377
	11.4	Reine XML-Datenbanksysteme	383
	11.4.1	Überblick	383
	11.4.2	Tamino – Überblick	384
	11.4.3	Erweiterbarkeit durch Server Extensions	389
	11.4.4	Textsuche in Tamino	390
	11.4.5	Internationalisierung in Tamino	392
	11.4.6	Tamino und WebDAV	392
12		Verteilte und föderierte Datenbanksysteme	394
	12.1	Überblick, Grundbegriffe, Abgrenzung	394
	12.1.1	Verteilte vs. parallele Datenbanksysteme	394
	12.1.2	Verteilte vs. föderierte Datenbanksysteme	397
	12.2	Schemaarchitektur und Entwurf verteilter Datenbanksysteme	399
	12.3	Fragmentierung	401
	12.3.1	Horizontale Fragmentierung	401
	12.3.2	Vertikale Fragmentierung	403
	12.3.3	Kombinierte Fragmentierung	403
	12.4	Verteilungstransparenz	403
	12.4.1	Vorteile verteilter Datenbanksysteme	403
	12.4.2	Transparenzeigenschaften verteilter Datenbanksysteme	404
	12.5	Verteilte Anfrageverarbeitung	405
	12.5.1	Datenlokalisierung	406
	12.5.2	Globale Optimierung der Join-Auswertung	409
	12.6	Transaktionsverwaltung in verteilten Datenbanksystemen	412
	12.6.1	Koordination	412

12.6.2	Synchronisation	414
12.6.3	Deadlock-Behandlung	415
12.6.4	Synchronisation bei Replikation	417
12.7	Föderierte Datenbanksysteme	418
12.7.1	Autonomie und Heterogenität	418
12.7.2	Architektur föderierter DBS	420
12.7.3	Integrationskonflikte	421
12.7.4	Integrationsstechniken	422
12.7.5	Anfrageverarbeitung in föderierten DBS	424
12.7.6	Synchronisation in föderierten DBS	425
12.8	Trends	426
13	Data Warehouse	427
13.1	Architektur	427
13.1.1	Datenquellen	427
13.1.2	Back-End-Bereich	427
13.1.3	Datenbank	432
13.1.4	Front-End-Werkzeuge	435
13.1.5	Sonstige Werkzeuge	436
13.2	Multidimensionale Datenmodelle	438
13.2.1	Statische Aspekte	438
13.2.2	Dynamische Aspekte	442
13.3	Speicherung und Schemagestaltung	444
13.3.1	Relationale Speicherung	445
13.3.2	Multidimensionale Speicherung	446
13.4	Erweiterung relationaler Datenbanken	446
13.4.1	Materialisierte Sichten	447
13.4.2	Partitionierung	448
13.4.3	Bitmap-Index	450
13.4.4	SQL-Erweiterungen zum Einfügen	451
13.4.5	Komplexes Gruppieren	451
13.4.6	Star Query	453
13.4.7	Bulk Loader	455
14	Data Mining	456
14.1	KDD-Prozess	456
14.2	Clustering	457
14.2.1	Definition und Beispiele	457
14.2.2	Anforderungen und Probleme	458
14.2.3	Verfahren	459
14.3	Assoziationsanalyse	463

	14.3.1	Definition und Beispiel	463
	14.3.2	Anforderungen und Probleme	464
	14.3.3	Verfahren	464
14.4		Klassifikation	467
	14.4.1	Definition und Beispiele	467
	14.4.2	Anforderungen	468
	14.4.3	Verfahren	468
14.5		Anomalieentdeckung	470
	14.5.1	Definition und Beispiele	470
	14.5.2	Anforderungen und Probleme	471
	14.5.3	Verfahren	471
15		Multimedia-Datenbanken	474
	15.1	Einführung	474
	15.2	Mediendaten	478
	15.3	Suche nach Mediendaten	481
	15.3.1	Textsuche	481
	15.3.2	Bildsuche	483
	15.3.3	Audiosuche	484
	15.3.4	Videosuche	486
	15.4	Mediendatentypen	487
	15.5	Einbettung in Datenbanksysteme	490
	15.5.1	Schemastrukturen	490
	15.5.2	Anfrageformulierung	493
	15.6	Einsatz	494
16		Geodatenbanken	496
	16.1	Geodaten	496
	16.1.1	Eigenschaften von Geodaten	496
	16.1.2	Metadaten	498
	16.2	Datenschemata	499
	16.2.1	Standardisierung	499
	16.2.2	ISO 19107 Spatial Schema	500
	16.2.3	ISO 19125 Simple Feature Access	501
	16.2.4	ISO/IEC 13249-3 SQL/MM Spatial	505
	16.2.5	Räumliche Bezugssysteme	506
	16.3	Funktionen	508
	16.3.1	Geometrische Funktionen	508
	16.3.2	Topologische Prädikate	510
	16.4	Räumliche Anfragen	512
	16.4.1	Räumliche Basisanfragen	513

16.4.2	Mehrstufige Anfragebearbeitung	515
16.4.3	Approximationen	516
16.5	Räumliche Indexe	517
16.5.1	Grundtechniken	518
16.5.2	Quadrees	521
16.5.3	R-Bäume	523
16.6	Geodatenbanksysteme	525
16.6.1	Objektrelationale Geodatenbanksysteme	525
16.6.2	Konkrete Geodatenbanksysteme	526
17	Datenbanken in OLTP-Systemen	528
17.1	Betriebliche Informationssysteme	528
17.1.1	Operative Systeme	528
17.1.2	Analytische Systeme	529
17.2	Prozessarchitektur	529
17.3	Sperrkonzepte	532
17.4	Das Verbucherkonzept	533
17.5	Datenbankschnittstelle	533
17.5.1	Datenbankunabhängige Programmierung	533
17.5.2	Datenpufferung	536
17.5.3	Anfrage-Caching	537
	Abkürzungsverzeichnis	538
	Literaturverzeichnis	543
	Sachwortverzeichnis	560