
Inhaltsverzeichnis

| | | |
|----------|---|-----------|
| 1 | Datenbanken: Grundlagen und Überblick | 19 |
| 1.1 | <i>Dateien und Dateisysteme</i> | 19 |
| 1.2 | <i>Terminologie</i> | 20 |
| 1.3 | <i>Merkmale eines DBMS</i> | 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 | 25 |
| 1.4 | <i>Architektur eines Datenbanksystems</i> | 26 |
| 1.4.1 | Architekturen | 26 |
| 1.4.2 | Schemaarchitektur | 27 |
| 1.4.2.1 | Datenbankschema | 27 |
| 1.4.2.2 | Drei-Ebenen-Architektur | 28 |
| 1.4.3 | Datenunabhängigkeit | 29 |
| 1.5 | <i>Benutzerrollen bei Entwurf und Betrieb von Datenbanken</i> | 30 |
| 1.6 | <i>Datenbanken und Informationssysteme</i> | 32 |
| 1.7 | <i>Fachgebiet Datenbanken im Überblick</i> | 33 |
| 1.7.1 | Themenbereiche und Zusammenhänge | 33 |
| 1.7.2 | Einordnung des Fachs innerhalb der Informatik | 35 |
| 1.8 | <i>Historische Entwicklung</i> | 37 |
| 1.8.1 | Frühzeit | 38 |
| 1.8.2 | Prärelationale DBMS | 38 |
| 1.8.3 | Durchbruch der relationalen Datenbanken | 39 |
| 1.8.4 | Neue Anwendungsfelder für Datenbanken | 40 |
| 1.8.5 | Neuzeit | 40 |
| 1.9 | <i>Datenbanken in der Forschung</i> | 41 |
| 1.9.1 | Fachverbände | 41 |
| 1.9.2 | Aktuelle Forschungstrends | 42 |
| 2 | Informationsmodellierung | 44 |
| 2.1 | <i>Datenbankentwurf</i> | 44 |
| 2.1.1 | Phasenmodell | 44 |
| 2.1.2 | Anforderungsanalyse | 45 |
| 2.1.3 | Konzeptioneller Entwurf | 46 |
| 2.1.4 | Logischer Entwurf | 47 |
| 2.1.5 | Datendefinition | 48 |
| 2.1.6 | Physischer Entwurf | 49 |
| 2.1.7 | Implementierung und Wartung | 49 |
| 2.2 | <i>Grundlagen des Entity-Relationship-Modells (ERM)</i> | 49 |
| 2.2.1 | Semantische Datenmodelle | 50 |
| 2.2.2 | Grundkonzepte des klassischen ERM | 50 |
| 2.2.2.1 | Konzepte auf der Instanzebene | 51 |

| | | |
|----------|---|-----------|
| 2.2.2.2 | Konzepte auf der Typebene | 52 |
| 2.2.2.3 | Rollenmodellierung | 53 |
| 2.2.3 | Kardinalitäten von Beziehungstypen | 54 |
| 2.2.3.1 | Kardinalitäten im klassischen ERM | 54 |
| 2.2.3.2 | Intervalle und Komplexitätsgrade | 54 |
| 2.2.4 | Existenzabhängigkeit vs. Optionalität | 55 |
| 2.2.5 | Rekursive und n -äre Beziehungstypen | 56 |
| 2.2.5.1 | Rekursive Beziehungstypen | 56 |
| 2.2.5.2 | N -äre Beziehungstypen | 57 |
| 2.2.6 | Attribute | 58 |
| 2.2.7 | Modellierungsbeispiel | 59 |
| 2.3 | Erweiterungen des ERM | 60 |
| 2.3.1 | Erweiterungen bei Attributen | 60 |
| 2.3.2 | Generalisierung und Spezialisierung | 61 |
| 2.3.3 | Aggregation | 64 |
| 2.3.4 | Modellierung zeitlicher Aspekte | 65 |
| 3 | Relationales Datenmodell | 67 |
| 3.1 | Konzepte und Grundbegriffe des relationalen Datenmodells | 67 |
| 3.1.1 | Relationen, Tupel, Attribute und Wertebereiche | 67 |
| 3.1.2 | Eigenschaften von Relationen | 69 |
| 3.2 | Integritätsbedingungen | 70 |
| 3.2.1 | Grundbegriffe | 70 |
| 3.2.2 | Entitätsintegrität | 71 |
| 3.2.3 | Referenzielle Integrität | 71 |
| 3.2.3.1 | Begriff | 71 |
| 3.2.3.2 | Regeln für Fremdschlüssel | 72 |
| 3.2.3.3 | Gewährleistung der referenziellen Integrität bei kritischen DML-Operationen | 73 |
| 3.3 | Abbildung des EERM auf das relationale Datenmodell | 75 |
| 3.3.1 | Problemstellung | 75 |
| 3.3.2 | Abbildungsregeln für Attribute und Entitytypen | 75 |
| 3.3.3 | Abbildungsregeln für Beziehungstypen | 76 |
| 3.3.4 | Abbildungsregeln für die Generalisierung | 79 |
| 3.4 | Optimierung von Relationen | 80 |
| 3.4.1 | Problemstellung | 80 |
| 3.4.2 | Anomalien bei DML-Operationen auf Relationen | 81 |
| 3.4.3 | Abhängigkeiten | 82 |
| 3.4.3.1 | Funktionale Abhängigkeiten | 82 |
| 3.4.3.2 | Mehrwertige Abhängigkeiten | 83 |
| 3.4.4 | Verbundtreue und Abhängigkeitstreue | 84 |
| 3.4.5 | Normalformenlehre | 85 |
| 3.4.5.1 | Erste Normalform (1NF) | 85 |
| 3.4.5.2 | Zweite Normalform (2NF) | 86 |
| 3.4.5.3 | Dritte Normalform (3NF) | 87 |
| 3.4.5.4 | Boyce-Codd-Normalform (BCNF) | 88 |
| 3.4.5.5 | Vierte Normalform (4NF) | 89 |

| | | |
|----------|---|-----------|
| 3.4.5.6 | Fünfte Normalform (5NF) | 90 |
| 3.4.5.7 | Denormalisierung | 92 |
| 3.5 | <i>Operationen der Relationenalgebra</i> | 92 |
| 3.5.1 | Einführung | 92 |
| 3.5.2 | Relationenorientierte Operationen | 93 |
| 3.5.3 | Mengenoperationen | 95 |
| 3.5.4 | Relationenalgebra und relationale Sprachen | 96 |
| 3.5.5 | Relationenkalkül | 97 |
| 4 | Die Datenbanksprache SQL | 99 |
| 4.1 | <i>Grundkonzepte</i> | 99 |
| 4.2 | <i>Historie</i> | 101 |
| 4.3 | <i>Spezifikationsdokumente</i> | 102 |
| 4.4 | <i>Beispieldatenbank</i> | 103 |
| 4.5 | <i>Datenbankanfragen</i> | 103 |
| 4.5.1 | Einführende Beispielanfragen | 104 |
| 4.5.2 | Grundgerüst von Anfragen | 105 |
| 4.5.3 | Anfragen mit Aggregatfunktionen | 107 |
| 4.5.4 | Anfragen mit Tabellenfunktionen | 108 |
| 4.5.5 | Anfragen mit Mengenoperationen | 108 |
| 4.5.6 | Anfragen mit Verbundoperationen | 110 |
| 4.5.7 | Verschachtelte Anfragen | 112 |
| 4.5.8 | Rekursive Anfragen | 114 |
| 4.5.9 | Anfragen mit sortierter Ausgabe | 115 |
| 4.6 | <i>Datenmanipulation</i> | 116 |
| 4.6.1 | Einfügen von Tabellenzeilen | 116 |
| 4.6.2 | Ändern von Tabellenzeilen | 117 |
| 4.6.3 | Löschen von Tabellenzeilen | 118 |
| 4.6.4 | Zusammenführen von Tabellenzeilen | 118 |
| 4.7 | <i>Datendefinition</i> | 119 |
| 4.7.1 | SQL-Datentypen | 119 |
| 4.7.2 | Erzeugen und Löschen von Schemata | 124 |
| 4.7.3 | Erzeugen, Ändern und Löschen von Tabellen | 124 |
| 4.7.4 | Erzeugen und Löschen von Domänen | 127 |
| 4.7.5 | Erzeugen und Löschen von Integritätsbedingungen | 128 |
| 4.7.6 | Erzeugen und Löschen von Sichten | 130 |
| 4.7.7 | Erzeugen und Löschen von Routinen | 131 |
| 4.7.7.1 | Prozedurale SQL-Konstrukte | 131 |
| 4.7.7.2 | Erzeugen von SQL-Routinen | 135 |
| 4.7.7.3 | Erzeugen von externen Routinen | 136 |
| 4.7.7.4 | Löschen von Routinen | 136 |
| 4.7.8 | Erzeugen und Löschen von Triggern | 137 |
| 4.7.9 | Erzeugen und Löschen von Sequenzgeneratoren | 138 |
| 4.8 | <i>Transaktionssteuerung</i> | 138 |
| 4.8.1 | Transaktionsanweisungen | 139 |
| 4.8.2 | Isolationsebenen | 140 |
| 4.8.3 | Überprüfung von Integritätsbedingungen | 141 |

| | | |
|----------|--|------------|
| 4.9 | <i>Zugriffskontrolle</i> | 141 |
| 4.9.1 | Vergabe von Zugriffsrechten | 141 |
| 4.9.2 | Zurücknahme von Zugriffsrechten | 143 |
| 4.9.3 | Erzeugen und Löschen von Rollen | 143 |
| 5 | Datenbank-Anwendungsprogrammierung | 144 |
| 5.1 | <i>Grundlagen der Datenbank-Anwendungsprogrammierung</i> | 144 |
| 5.1.1 | Impedance Mismatch | 144 |
| 5.1.2 | Einbettungstechniken | 145 |
| 5.1.3 | Einbettungsarten | 145 |
| 5.1.3.1 | Statische Programmierspracheneinbettung | 145 |
| 5.1.3.2 | Dynamische Programmierspracheneinbettung | 146 |
| 5.1.4 | Architekturansätze | 147 |
| 5.1.4.1 | Fat-Client-Architektur | 148 |
| 5.1.4.2 | Thin-Client-Architektur | 148 |
| 5.1.4.3 | Thin-Client-Architektur mit Applikationsserver | 150 |
| 5.2 | <i>Embedded SQL</i> | 151 |
| 5.2.1 | Grundidee und Architektur | 151 |
| 5.2.2 | Syntax | 152 |
| 5.2.3 | Host-Variablen | 152 |
| 5.2.4 | Das Cursor-Konzept | 152 |
| 5.2.5 | Statische Einbettung | 154 |
| 5.2.6 | Dynamische Einbettung | 155 |
| 5.2.7 | Indikatorvariablen | 156 |
| 5.2.8 | SQLJ | 156 |
| 5.3 | <i>Aufrufschnittstellen</i> | 158 |
| 5.3.1 | Überblick | 158 |
| 5.3.2 | SQL/CLI | 158 |
| 5.3.3 | JDBC | 160 |
| 5.3.3.1 | Treibertypen | 161 |
| 5.3.3.2 | Verbindung aufbauen | 161 |
| 5.3.3.3 | Verbindung schließen | 163 |
| 5.3.3.4 | Transaktionen | 163 |
| 5.3.3.5 | Leseanweisungen – das ResultSet | 164 |
| 5.3.3.6 | Änderungen | 165 |
| 5.3.3.7 | Zugriff auf Metadaten | 167 |
| 5.4 | <i>Relationale Datenbankprogrammiersprachen</i> | 168 |
| 5.4.1 | Datenbankinterne Ansätze | 168 |
| 5.4.1.1 | SQL/PSM | 168 |
| 5.4.1.2 | Benutzerdefinierte Routinen | 169 |
| 5.4.1.3 | Trigger | 170 |
| 5.4.2 | 4GL-Programmiersprachen | 174 |
| 5.4.2.1 | Überblick | 174 |
| 5.4.2.2 | Beispiel: ABAP | 174 |
| 5.5 | <i>Objektrelationale Abbildung</i> | 177 |
| 5.5.1 | Java Persistence API (JPA) | 179 |
| 5.5.1.1 | Schemaabbildung | 180 |

| | | |
|----------|---|------------|
| 5.5.1.2 | Datenzugriff | 181 |
| 5.5.1.3 | Formulieren von Anfragen | 183 |
| 5.5.2 | JDO – Java Data Objects | 184 |
| 5.5.3 | Entity Beans | 188 |
| 6 | Datenbanken im Web | 189 |
| 6.1 | <i>Grundlagen des Web</i> | 189 |
| 6.2 | <i>Eigenschaften von Webanwendungen</i> | 191 |
| 6.2.1 | Anforderungen | 192 |
| 6.2.2 | Webseiten | 192 |
| 6.2.3 | Dynamische Webseiten | 195 |
| 6.2.4 | Adressierung | 197 |
| 6.2.5 | Kommunikation | 198 |
| 6.3 | <i>Datenbankanbindung im Web</i> | 200 |
| 6.3.1 | Architekturen | 201 |
| 6.3.2 | Programmierung | 202 |
| 6.3.3 | Überblick | 204 |
| 6.4 | <i>Datenbankanbindung über Datenexport</i> | 206 |
| 6.5 | <i>Clientseitige Datenbankanbindung</i> | 206 |
| 6.5.1 | Skripteinbettung | 207 |
| 6.5.2 | Programmeinbettung | 209 |
| 6.5.3 | Weitere Techniken | 212 |
| 6.6 | <i>Serverseitige Datenbankanbindung</i> | 212 |
| 6.6.1 | Webseitengeneratoren | 212 |
| 6.6.2 | Skripteinbettung | 215 |
| 6.6.3 | Programmeinbettung | 219 |
| 6.6.4 | Applikationsserver | 221 |
| 6.6.5 | Webservices | 223 |
| 7 | Komponenten eines Datenbankmanagementsystems | 227 |
| 7.1 | <i>Architektur von DBMS</i> | 227 |
| 7.1.1 | Schichtenmodell | 227 |
| 7.1.2 | Prozessarchitektur | 229 |
| 7.2 | <i>Pufferverwaltung</i> | 230 |
| 7.2.1 | Notwendigkeit und Aufgabe | 231 |
| 7.2.2 | Speicherzuteilung | 232 |
| 7.2.3 | Seitenersetzung | 233 |
| 7.3 | <i>Speicher- und Zugriffssystem</i> | 234 |
| 7.3.1 | Aufgabe | 235 |
| 7.3.2 | Seiten und Sätze | 235 |
| 7.3.3 | Adressierung von Sätzen | 237 |
| 7.4 | <i>Anfrageprozessor</i> | 238 |
| 7.4.1 | Basisoperatoren | 238 |
| 7.4.1.1 | Unäre Operatoren | 239 |
| 7.4.1.2 | Binäre Operatoren | 241 |
| 7.4.2 | Anfrageplanung und -optimierung | 242 |
| 7.4.3 | Kosten und Statistiken | 247 |

| | | |
|----------|--|------------|
| 7.5 | <i>Transaktionsverwaltung</i> | 249 |
| 7.5.1 | Aufgabe | 249 |
| 7.5.2 | Serialisierbarkeit | 251 |
| 7.5.3 | Sperrverfahren | 253 |
| 7.5.4 | Nicht sperrende Verfahren | 257 |
| 7.6 | <i>Recovery</i> | 259 |
| 7.6.1 | Fehlerklassen | 259 |
| 7.6.2 | Recovery-Strategien | 260 |
| 7.6.3 | Logging | 261 |
| 7.6.4 | Wiederanlauf im Fehlerfall | 264 |
| 7.6.5 | Schattenspeicherverfahren | 266 |
| 7.7 | <i>Datenbanktechniken für moderne Hardware-Architekturen</i> | 266 |
| 8 | Dateiorganisation und Indexe | 269 |
| 8.1 | <i>Organisation von Dateien</i> | 269 |
| 8.1.1 | Formen der Dateiorganisation | 269 |
| 8.1.2 | Dateiorganisationsformen im Vergleich | 269 |
| 8.1.2.1 | Basisoperationen | 270 |
| 8.1.2.2 | Kosten | 270 |
| 8.2 | <i>Zugriffsstrukturen</i> | 271 |
| 8.2.1 | Grundlagen von Zugriffsverfahren | 271 |
| 8.2.2 | Eigenschaften von Indexen | 272 |
| 8.3 | <i>Baumbasierte Verfahren</i> | 274 |
| 8.3.1 | ISAM-Bäume | 274 |
| 8.3.2 | Balancierte Mehrwegbäume | 275 |
| 8.3.3 | Digitale Bäume | 277 |
| 8.4 | <i>Hash-Verfahren</i> | 278 |
| 8.4.1 | Prinzip des Hashing | 278 |
| 8.4.2 | Erweiterbares Hashing | 279 |
| 8.4.3 | Weitere Hash-Verfahren | 280 |
| 8.5 | <i>Mehrdimensionale Zugriffsverfahren</i> | 281 |
| 8.5.1 | Mehrdimensionale Baumverfahren | 281 |
| 8.5.2 | Grid-File | 281 |
| 8.6 | <i>Clusterung und Partitionierung</i> | 283 |
| 8.6.1 | Clusterung | 283 |
| 8.6.2 | Partitionierung | 283 |
| 8.7 | <i>Umsetzung in SQL-Systemen</i> | 284 |
| 8.7.1 | Definition von Tabellen | 285 |
| 8.7.2 | Definition von Indexen | 285 |
| 9 | Optimierung von Datenbanken und Leistungsbewertung | 287 |
| 9.1 | <i>Motivation der Datenbankoptimierung</i> | 287 |
| 9.1.1 | Kosten von Datenbankabfragen | 288 |
| 9.1.2 | Optimierungspotenzial | 289 |
| 9.1.3 | Zielbestimmung der Datenbankoptimierung | 290 |
| 9.2 | <i>Phasen der Datenbankoptimierung</i> | 292 |

| | | |
|-----------|---|------------|
| 9.3 | <i>Phase 1.1 – Optimierung des Datenbankschemas</i> | 293 |
| 9.3.1 | Konzeptuelles Schema | 294 |
| 9.3.1.1 | Attribute | 294 |
| 9.3.1.2 | Tabellen | 294 |
| 9.3.1.3 | Redundanz | 296 |
| 9.3.2 | Externes Schema | 297 |
| 9.3.2.1 | Sichten | 298 |
| 9.3.2.2 | Prozedurale SQL-Erweiterungen | 298 |
| 9.3.3 | Internes Schema | 298 |
| 9.3.3.1 | Materialisierte Sichten | 298 |
| 9.3.3.2 | Zugriffspfadstrukturen | 299 |
| 9.4 | <i>Phase 1.2 – Anwendungsoptimierung</i> | 301 |
| 9.4.1 | Optimierung von Unternehmensfunktionen | 301 |
| 9.4.2 | Optimierung der Anwendung | 301 |
| 9.4.3 | Optimierungen im Mehrbenutzerbetrieb | 302 |
| 9.4.4 | Formulierung von SQL-Anweisungen | 304 |
| 9.5 | <i>Phase 2 – Hauptspeicheroptimierung</i> | 306 |
| 9.5.1 | Gestaltung des Datenbankpuffers | 306 |
| 9.5.1.1 | Komponenten des Datenbankpuffers | 307 |
| 9.5.1.2 | Größe des Datenbankpuffers | 308 |
| 9.5.1.3 | Blockfüllgrad | 308 |
| 9.5.2 | Schreiben des Datenbankpuffers | 309 |
| 9.5.2.1 | Sicherungspunkt-Intervalle | 309 |
| 9.5.2.2 | Protokolldatei | 310 |
| 9.5.3 | Optimierer | 310 |
| 9.5.3.1 | Statistiken | 311 |
| 9.5.3.2 | Planhinweise (Hints) | 311 |
| 9.6 | <i>Phase 3 – Optimierung der Sekundärspeicherzugriffe</i> | 312 |
| 9.6.1 | Zusammenspiel mit dem Betriebssystem | 312 |
| 9.6.2 | Verteilung der Eingabe-/Ausgabelast | 313 |
| 9.6.2.1 | Verteilung der Daten- und Indexdateien | 313 |
| 9.6.2.2 | Verteilung der Daten | 313 |
| 9.6.2.3 | Verteilung der Protokolldatei | 314 |
| 9.6.2.4 | RAID-Level | 314 |
| 9.6.3 | Optimierung physischer Speicherstrukturen | 315 |
| 9.6.3.1 | Blockgröße | 315 |
| 9.6.3.2 | Cluster-Techniken | 316 |
| 9.6.3.3 | Kompressions-Techniken | 317 |
| 9.6.3.4 | Reorganisation | 317 |
| 9.7 | <i>Leistungsbewertung</i> | 318 |
| 9.7.1 | Transaction Processing Performance Council | 319 |
| 9.7.2 | Vergleichbarkeit der Benchmark-Ergebnisse | 320 |
| 10 | Objektrelationale Datenbanken | 322 |
| 10.1 | <i>Objektorientierte Konzepte</i> | 322 |
| 10.1.1 | Objekte | 322 |
| 10.1.2 | Methoden | 323 |

| | | |
|-----------|--|------------|
| 10.1.3 | Kapselung | 323 |
| 10.1.4 | Objektidentität | 323 |
| 10.1.5 | Klassen | 324 |
| 10.1.6 | Spezialisierung | 325 |
| 10.2 | Objektorientierung in Datenbanken | 326 |
| 10.3 | Objektrelationale Standard-SQL-Konzepte | 327 |
| 10.3.1 | Typkonstruktoren | 328 |
| 10.3.2 | Distinct-Typen | 330 |
| 10.3.3 | Strukturierte Typen | 330 |
| 10.3.4 | Methoden | 333 |
| 10.3.5 | Benutzerdefinierte Konstruktoren | 335 |
| 10.3.6 | Benutzerdefinierte Casts | 336 |
| 10.3.7 | Benutzerdefinierte Ordnungen | 337 |
| 10.3.8 | Typisierte Tabellen | 338 |
| 10.3.9 | Typisierte Sichten | 339 |
| 10.4 | Objektrelationale Anfragen | 341 |
| 10.4.1 | Anfragen auf Kollektionen | 341 |
| 10.4.2 | Anfragen mit Pfadausdrücken | 342 |
| 10.4.3 | Anfragen mit Methodenaufrufen | 342 |
| 10.4.4 | Anfragen auf flachen Tabellenextensionen | 343 |
| 10.4.5 | Typspezifische Anfragen | 343 |
| 10.4.6 | Anfragen mit temporärer Typanpassung | 343 |
| 11 | XML und Datenbanken | 345 |
| 11.1 | Überblick über XML | 345 |
| 11.1.1 | Der XML-Standard und verwandte Standards | 345 |
| 11.1.2 | XML Schema | 347 |
| 11.2 | Anfragesprachen für XML | 350 |
| 11.2.1 | Pfadausdrücke (XPath und XQuery) | 351 |
| 11.2.2 | XQuery | 352 |
| 11.3 | XML und relationale Datenbanksysteme | 356 |
| 11.3.1 | Relationale Speicherung von XML | 356 |
| 11.3.2 | SQL/XML | 358 |
| 11.3.3 | Realisierung in kommerziellen Systemen | 361 |
| 11.4 | Reine XML-Datenbanksysteme | 366 |
| 12 | NoSQL-Datenbanksysteme | 368 |
| 12.1 | Motivation und Grundbegriffe | 368 |
| 12.2 | Klassifikation | 369 |
| 12.2.1 | Key-Value-Datenbanksysteme | 370 |
| 12.2.1.1 | Datenmodell und Schema | 370 |
| 12.2.1.2 | Anfragen und Datenmanipulation | 371 |
| 12.2.1.3 | Einsatzbereiche und Systeme | 371 |
| 12.2.2 | Dokumentenorientierte Datenbanksysteme | 372 |
| 12.2.2.1 | Datenmodell und Schema | 372 |
| 12.2.2.2 | Anfragen und Datenmanipulation | 373 |
| 12.2.2.3 | Einsatzbereiche und Systeme | 374 |

| | | |
|-----------|--|------------|
| 12.2.3 | Column-Family-Datenbanksysteme | 374 |
| 12.2.3.1 | Datenmodell und Schema | 374 |
| 12.2.3.2 | Anfragen und Datenmanipulation | 377 |
| 12.2.3.3 | Einsatzbereiche und Systeme | 377 |
| 12.2.4 | Weitere NoSQL-Datenbanksysteme | 378 |
| 12.3 | Datenmodellierung | 378 |
| 12.3.1 | Generelle Aspekte der Datenmodellierung in NoSQL-Datenbanksystemen | 378 |
| 12.3.2 | Datenmodellierung für dokumentenorientierte Datenbanken | 380 |
| 12.3.3 | Datenmodellierung für Column-Family-Datenbanken | 382 |
| 12.3.3.1 | Eingebettete Speicherung in Column-Family-Datenbanksystemen | 382 |
| 12.3.3.2 | Spaltenfamilien | 384 |
| 12.3.4 | Datenmodellierung für Key-Value-Datenbanken | 385 |
| 12.4 | Anwendungsentwicklung mit NoSQL-Datenbanksystemen | 385 |
| 12.4.1 | MapReduce | 386 |
| 12.4.1.1 | MapReduce-Prinzip | 387 |
| 12.4.1.2 | MapReduce-Beispiel | 388 |
| 12.4.1.3 | MapReduce-Frameworks | 389 |
| 12.4.1.4 | MapReduce-Trends | 389 |
| 12.4.2 | Schema-Management | 389 |
| 12.5 | Skalierbarkeit, Verfügbarkeit und Konsistenz | 390 |
| 12.6 | Auswahl eines geeigneten Datenbanksystems | 392 |
| 12.6.1 | Kriterienkatalog | 392 |
| 12.6.2 | Performance | 392 |
| 12.6.3 | Polyglotte Persistenz | 393 |
| 13 | Verteilte und föderierte Datenbanksysteme | 394 |
| 13.1 | Überblick, Grundbegriffe, Abgrenzung | 394 |
| 13.1.1 | Verteilte vs. parallele Datenbanksysteme | 395 |
| 13.1.2 | Verteilte vs. föderierte Datenbanksysteme | 397 |
| 13.2 | Schemaarchitektur und Entwurf verteilter Datenbanksysteme | 400 |
| 13.3 | Fragmentierung | 401 |
| 13.3.1 | Horizontale Fragmentierung | 401 |
| 13.3.2 | Vertikale Fragmentierung | 403 |
| 13.3.3 | Kombinierte Fragmentierung | 403 |
| 13.4 | Verteilungstransparenz | 404 |
| 13.4.1 | Vorteile verteilter Datenbanksysteme | 404 |
| 13.4.2 | Transparenzeigenschaften verteilter Datenbanksysteme | 404 |
| 13.5 | Verteilte Anfrageverarbeitung | 405 |
| 13.5.1 | Datenlokalisierung | 407 |
| 13.5.2 | Globale Optimierung der Join-Auswertung | 409 |
| 13.6 | Transaktionsverwaltung in verteilten Datenbanksystemen | 412 |
| 13.6.1 | Koordination | 412 |
| 13.6.2 | Synchronisation | 414 |
| 13.6.3 | Deadlock-Behandlung | 415 |
| 13.6.4 | Synchronisation bei Replikation | 416 |

| | | |
|-----------|---|------------|
| 13.7 | <i>Föderierte Datenbanksysteme</i> | 417 |
| 13.7.1 | Autonomie und Heterogenität | 418 |
| 13.7.2 | Architektur föderierter DBS | 419 |
| 13.7.3 | Integrationsprozess | 420 |
| 13.7.4 | Anfrageverarbeitung in föderierten DBS | 421 |
| 13.7.5 | Synchronisation in föderierten DBS | 422 |
| 13.8 | <i>Cloud-Datenbanken</i> | 423 |
| 13.8.1 | Cloud Data Management und Big Data | 423 |
| 13.8.2 | Das CAP-Theorem | 424 |
| 13.8.3 | Database as a Service (DBaaS) | 425 |
| 13.8.4 | Skalierbarkeit und Verfügbarkeit | 426 |
| 13.9 | <i>Trends</i> | 428 |
| 14 | Data Warehouse | 430 |
| 14.1 | <i>Architektur</i> | 430 |
| 14.1.1 | Datenquellen | 430 |
| 14.1.2 | Back-End-Bereich | 430 |
| 14.1.2.1 | Monitore | 431 |
| 14.1.2.2 | Extraktionskomponenten | 432 |
| 14.1.2.3 | Transformationskomponente | 433 |
| 14.1.2.4 | Ladekomponente | 434 |
| 14.1.3 | Datenbank | 434 |
| 14.1.3.1 | Data Warehouse | 434 |
| 14.1.3.2 | Data Marts | 436 |
| 14.1.3.3 | Archiv-Datenbank | 437 |
| 14.1.4 | Front-End-Werkzeuge | 437 |
| 14.1.4.1 | Berichts- und Anfragewerkzeuge | 437 |
| 14.1.4.2 | OLAP-Werkzeuge | 438 |
| 14.1.4.3 | Data-Mining-Werkzeuge | 438 |
| 14.1.4.4 | Sonstige Front-End-Werkzeuge | 438 |
| 14.1.5 | Sonstige Werkzeuge | 439 |
| 14.1.5.1 | DWS-Manager | 439 |
| 14.1.5.2 | Metadaten-Repository | 439 |
| 14.2 | <i>Multidimensionale Datenmodelle</i> | 440 |
| 14.2.1 | Statische Aspekte | 440 |
| 14.2.2 | Dynamische Aspekte | 444 |
| 14.3 | <i>Speicherung und Schemagestaltung</i> | 446 |
| 14.3.1 | Relationale Speicherung | 446 |
| 14.3.2 | Multidimensionale Speicherung | 447 |
| 14.3.3 | Spaltenorientierte Speicherung | 448 |
| 14.4 | <i>Erweiterung relationaler Datenbanken</i> | 451 |
| 14.4.1 | Materialisierte Sichten | 452 |
| 14.4.2 | Partitionierung | 453 |
| 14.4.3 | Bitmap-Index | 455 |
| 14.4.4 | SQL-Erweiterungen zum Einfügen | 456 |
| 14.4.5 | Komplexes Gruppieren | 457 |

| | | |
|-----------|---|------------|
| 14.4.6 | Star Query | 458 |
| 14.4.7 | Bulk Loader | 460 |
| 15 | Data Mining | 461 |
| 15.1 | <i>KDD-Prozess</i> | 461 |
| 15.2 | <i>Clustering</i> | 462 |
| 15.2.1 | Definition und Beispiele | 462 |
| 15.2.2 | Anforderungen und Probleme | 463 |
| 15.2.3 | Verfahren | 463 |
| 15.2.3.1 | Partitionierende Verfahren | 464 |
| 15.2.3.2 | Hierarchische Verfahren | 465 |
| 15.2.3.3 | Dichtebasierte Methoden | 468 |
| 15.3 | <i>Assoziationsanalyse</i> | 468 |
| 15.3.1 | Definition und Beispiel | 468 |
| 15.3.2 | Anforderungen und Probleme | 468 |
| 15.3.3 | Verfahren | 469 |
| 15.4 | <i>Klassifikation</i> | 472 |
| 15.4.1 | Definition und Beispiele | 472 |
| 15.4.2 | Anforderungen | 472 |
| 15.4.3 | Verfahren | 473 |
| 15.4.3.1 | Entscheidungsbaum-Klassifikatoren | 473 |
| 15.4.3.2 | Regelbasierte Klassifikatoren | 473 |
| 15.4.3.3 | Weitere Verfahren | 474 |
| 15.5 | <i>Anomalieentdeckung</i> | 474 |
| 15.5.1 | Definition und Beispiele | 474 |
| 15.5.2 | Anforderungen und Probleme | 475 |
| 15.5.3 | Verfahren | 475 |
| 15.5.3.1 | Grafische und statistikbasierte Verfahren | 476 |
| 15.5.3.2 | Distanzbasierte Ansätze | 476 |
| 16 | Multimedia-Datenbanken | 478 |
| 16.1 | <i>Einführung</i> | 478 |
| 16.2 | <i>Mediendaten</i> | 482 |
| 16.3 | <i>Suche nach Mediendaten</i> | 485 |
| 16.3.1 | Textsuche | 485 |
| 16.3.2 | Bildsuche | 487 |
| 16.3.3 | Audiosuche | 488 |
| 16.3.4 | Videosuche | 489 |
| 16.4 | <i>Mediendatentypen</i> | 490 |
| 16.5 | <i>Einbettung in Datenbanksysteme</i> | 494 |
| 16.5.1 | Schemastrukturen | 494 |
| 16.5.2 | Anfrageformulierung | 496 |
| 16.6 | <i>Einsatz</i> | 497 |
| 17 | Geodatenbanken | 499 |
| 17.1 | <i>Geodaten</i> | 499 |
| 17.1.1 | Eigenschaften von Geodaten | 499 |
| 17.1.2 | Metadaten | 501 |

| | | |
|----------|---------------------------------|------------|
| 17.2 | <i>Datenschemata</i> | 502 |
| 17.2.1 | Standardisierung | 502 |
| 17.2.2 | ISO 19107 Spatial Schema | 502 |
| 17.2.3 | ISO 19125 Simple Feature Access | 504 |
| 17.2.3.1 | Datenschema | 504 |
| 17.2.3.2 | Datenrepräsentationen | 506 |
| 17.2.4 | ISO/IEC 13249-3 SQL/MM Spatial | 507 |
| 17.2.5 | Räumliche Bezugssysteme | 508 |
| 17.2.5.1 | EPSG-Bezugssysteme | 510 |
| 17.2.5.2 | Lineare Bezugssysteme | 510 |
| 17.3 | <i>Funktionen</i> | 510 |
| 17.3.1 | Geometrische Funktionen | 511 |
| 17.3.2 | Topologische Prädikate | 512 |
| 17.3.2.1 | Boolesches Modell | 512 |
| 17.3.2.2 | Dimensionsmodell | 513 |
| 17.4 | <i>Räumliche Anfragen</i> | 515 |
| 17.4.1 | Räumliche Basisanfragen | 515 |
| 17.4.1.1 | Räumliche Selektion | 515 |
| 17.4.1.2 | Räumlicher Verbund | 516 |
| 17.4.1.3 | Nächste-Nachbarn-Anfrage | 517 |
| 17.4.2 | Mehrstufige Anfragebearbeitung | 517 |
| 17.4.3 | Approximationen | 518 |
| 17.5 | <i>Räumliche Indexe</i> | 520 |
| 17.5.1 | Grundtechniken | 520 |
| 17.5.1.1 | Clipping | 520 |
| 17.5.1.2 | Punkttransformationen | 521 |
| 17.5.1.3 | Raumfüllende Kurven | 521 |
| 17.5.1.4 | Überlappende Blockregionen | 522 |
| 17.5.2 | Quadrees | 523 |
| 17.5.3 | R-Bäume | 525 |
| 17.6 | <i>Geodatenbanksysteme</i> | 527 |
| | Abkürzungsverzeichnis | 529 |
| | Literaturverzeichnis | 534 |
| | Sachwortverzeichnis | 554 |