



mitp

Gunter  
Saake

Kai-Uwe  
Sattler

Andreas  
Heuer

3. Auflage

# Datenbanken

## Implementierungstechniken

# Vorwort

Das Gebiet der *Datenbanksysteme* gehört zu den klassischen Ausbildungsbereichen der Informatikstudiengänge. Datenbanksysteme kommen immer dann zum Einsatz, wenn an die Datenhaltung besondere Anforderungen hinsichtlich der Zuverlässigkeit, des zu speichernden Volumens, der Ausfallsicherheit, des Mehrbenutzerzugriffs, der Komplexität der Datenbeschreibung oder der Datenqualität gestellt werden. Zu Beginn des Informationszeitalters ist es daher nicht verwunderlich, dass der Umgang mit Datenbanksystemen für viele Absolventinnen und Absolventen der Informatikstudiengänge zum Berufsalltag gehört.

Viele Grundkonzepte der Datenbanktechnologie wurden in den 70er-Jahren entwickelt und seitdem beständig weiterentwickelt. Nach einer Stabilisierung in den 80er-Jahren – die damalige relationale Basistechnologie wurde fälschlicherweise als das Universalrüstzeug für alle Arten von Anwendungen angesehen – wird die aktuelle Entwicklung von speziellen Anwendungsanforderungen (Geoinformationssystemen, Datenanalyse, Multimediaanwendungen) und neuen Architekturen (Internetdatenbanken, erweiterbare Datenbanken, Datenhaltung auf mobilen Kleinrechnern, Datenbankföderationen) geprägt, die die etablierten Techniken in vielfältiger Weise weiterentwickeln. Als Resultat ist das Gebiet der Implementierungstechniken derart vielgestaltig geworden, dass es in einem üblichen Grundlagenlehrbuch und einer Datenbankgrundvorlesung nur noch knapp angerissen werden kann.

Bereits in der ersten Auflage des Buches *Datenbanken – Konzepte und Sprachen* [HS95a], das sich auf Modelle und Benutzersprachen konzentriert, wurde daher ein spezieller Folgeband zum Thema Implementierungstechniken für Datenbanken angekündigt, der hiermit vorliegt. Diese Aufteilung auf zwei Bände erlaubt es, Implementierungskonzepte ausführlicher vorzustellen und auch Spezialentwicklungen für neue Anwendungen nicht unter den Tisch fallen zu lassen.

Nur wenige in Deutschland ausgebildete Informatiker oder Wirtschaftsinformatiker werden je an der Implementierung eines „echten“ Datenbankmanagementsystems beteiligt sein – wie üblich bei derartiger Grundsoftware wird

der Markt von wenigen (in der Regel amerikanischen) Firmen dominiert. Trotzdem gibt es eine Reihe guter Gründe, sich intensiv mit den in diesem Buch diskutierten Implementierungstechniken zu beschäftigen:

- Auch wenn man selten ein komplettes DBMS implementieren dürfte, können ursprünglich für Datenbanksysteme entwickelte Einzelalgorithmen und Datenstrukturen auch in anderer Software erfolgreich eingesetzt werden.
- Komplexe Anwendungen von Datenbanksoftware, die einen Schwerpunkt in kommerziellen Arbeitsgebieten für Absolventen bilden, erfordern ein tiefes Verständnis der Abläufe in Datenbanksystemen, um diese auch bei sehr großen Datenbeständen und Transaktionslasten effizient zu tunen.

Nicht zuletzt ist die Implementierung von Datenbanksystemen ein Gebiet, das hohe Anforderungen an Datenstrukturen und Algorithmen stellt und daher per se interessant für alle ist, die sich intensiv mit Informatik und Software Engineering beschäftigen.

### *Die dritte Auflage*

Das Gebiet der Implementierung von Datenbanksystemen entwickelt sich stetig weiter. Schon die zweite Auflage dieses Buches [SHS05] erreichte eine Dicke von über 850 Seiten. Auch bei Weglassen von Techniken, die sich in den letzten sechs Jahren als nicht so zentral wie 2005 gedacht herausgestellt haben, hätte ein einfaches Hinzufügen neuerer Entwicklungen zu einem Lehrbuch von über 1000 Seiten geführt – zu viel für das Lehrbuchbudget eines typischen Studenten, und zu schwer, um es zum Stöbern in die Tasche zu stecken. Schweren Herzens haben sich die Autoren daher dazu entschlossen, drastische Einschnitte bei der Themenauswahl vorzunehmen.

Das Buch behandelt jetzt die Basistechnologien zentraler, insbesondere relationaler Datenbankmanagementsysteme: Architektur, Datenorganisation, Anfragebearbeitung, Synchronisation im Mehrbenutzerbetrieb und Recovery. Diese Themen werden detailliert und ergänzt um neuere Entwicklungen dargestellt und behandelt. Dieses Basiswissen soll den Leser insbesondere in die Lage versetzen, die Eigenschaften eines relationalen DBMS zu verstehen und bei auftretenden Problemen kompetent kausale Zusammenhänge zwischen Problemsituationen und Implementierungs- bzw. Parametrisierungsentscheidungen herzustellen.

Weggefallen sind einige Techniken, die für spezielle Anwendungsszenarien und andere Datenmodelle relevant sind, so Indexstrukturen für multimediale Daten oder die Realisierung von Objektidentifikatoren in objektorientierten Datenbanken.

Der größte Einschnitt allerdings betrifft den Verzicht auf die Behandlung der Aspekte verteilter Datenhaltung sowie der Parallelisierungspotentiale. Ei-

nige wenige Punkte nur erschienen den Autoren so zentral, dass sie weiterhin behandelt werden, so die Fragen der Partitionierung oder der Commit-Protokolle. Dieser Verzicht ist den Autoren besonders schmerzhaft, da die aktuellen Entwicklungen im Hardwarebereich und der Softwareinfrastruktur (Stichworte Cloud-Computing, P2P-Datenhaltung, Multi-Core-Prozessoren) gerade hier viel Entwicklungspotential zeigen. Aber ohne diesen Einschnitt wäre ein Seitenlimit unter 700 Seiten nicht zu halten gewesen. Hier besteht in den nächsten Jahren sicher der Bedarf an einem dedizierten Lehrbuch zu diesen Themenkreisen.

Das Themenspektrum dieser dritten Auflage deckt aber – zusammen mit dem Lehrbuch über Konzepte und Sprachen von Datenbanken [SSH10] – genau die Aspekte von Datenbanksystemen ab, die nach den Empfehlungen des deutschen Fakultätentags für Informatik in einem Informatik-Bachelor-Studium zwingend erforderlich sind. Natürlich werden innerhalb dieses Themenspektrums die Inhalte in diesem Buch so breit gefächert und vertieft, dass auch spezielle Vorlesungen für Master-Studiengänge aus dem hier dargestellten Stoff abgeleitet werden können. Umgekehrt sind aber keine Themengebiete in dieser dritten Auflage entfallen, die zum Pflichtkanon eines Informatik-Bachelors gehören.

## *Danksagungen*

Buchprojekte wie dieses haben – zumindest, wenn man die Autorentätigkeit nicht als Vollzeitjob betreibt – die Eigenschaft, dass der zeitliche Aufwand meist völlig unterschätzt wird. So müssen dann immer die Familien gestresster Autoren ertragen, die „nur noch“ ein Kapitel, einen Abschnitt etc. fertigstellen müssen. Natürlich wissen wir, dass dies durch Widmungen oder Danksagungen nicht ausgeglichen werden kann. Dennoch möchten wir uns an dieser Stelle bei unseren Familien für ihr Verständnis, ihre Geduld und ihren Rückhalt bedanken – ohne sie wäre dieses Buch sicher nie fertig geworden.

Andreas Heuer möchte sich bei seinen beiden Koautoren bedanken, dass sie ihn auf diese – für ihn meist passive – Reise zur dritten Auflage mitgenommen haben, obwohl er sich zwei Jahre lang vorrangig eher in der Wunderwelt der Universitätskliniken aufgehalten hat. Ein weiterer Dank gilt vielen Kolleginnen und Kollegen eben aus diesen Kliniken für den Entwurf und die sehr erfolgreiche Umsetzung einer komplexen und langen Transaktion. Die Mitarbeiter des Rostocker Lehrstuhls Datenbank- und Informationssysteme, insbesondere Holger Meyer und Meike Klettke, haben Forschung und Lehre am Lehrstuhl über zwei Jahre auch ohne ihren Chef aufrechterhalten, auch dafür ein dickes Danke. Und schließlich gilt ein Dank noch Renate Heuer, ohne die die Motivation und die Kraft gefehlt hätte, auf das `commit` der langen Transaktion hinzuarbeiten.

Ein besonderer Dank gilt David Broneske für die Unterstützung bei Abbildungen und Beispielen. Für das Korrekturlesen der dritten Auflage bedanken

wir uns bei Stephan Baumann, Felix Beier, Ingolf Geist, Alexander Grebhahn, Francis Gropengießer, Andreas Lübcke, Veit Köppen, Martin Kuhlemann, Eike Schallehn und Reimar Schröter. Einige der Übungsaufgaben entstammen dem in den Jahren entstandenen Fundus der beteiligten Lehrstühle, ohne dass sich die einzelnen Urheber der Aufgaben noch feststellen lassen.

Ergänzende Informationen zum Buch wie Verweise zu begleitenden Vorlesungsmaterialien und gegebenenfalls erforderliche Fehlerkorrekturen sind im Web unter folgender Adresse zu finden:

<http://www.biberbuch.de>

Magdeburg, Ilmenau und Rostock, im Oktober 2011

Gunter Saake, Kai-Uwe Sattler und Andreas Heuer