

mitp Professional

Data Warehouse Technologien

von

Veit Köppen, Gunter Saake, Kai-Uwe Sattler

2. Auflage 2014

Data Warehouse Technologien – Köppen / Saake / Sattler

schnell und portofrei erhältlich bei beck-shop.de DIE FACHBUCHHANDLUNG

Thematische Gliederung:

Datenbankmodelle, Datenbanktheorien

mitp/bhv 2014

Verlag C.H. Beck im Internet:

www.beck.de

ISBN 978 3 8266 9485 1

Architekturprinzipien von Data-Warehouse-Systemen Datenstrukturen und Algorithmen Anwendungsfeld Business Intelligence

Aus dem Inhalt:

Data Warehousing

- Architekturkonzepte
- Extraktion, Transformation und Laden
- Datenqualität
- Business Intelligence

Modellierung

- Multidimensionales Modell
- Relationale Umsetzung
- Star- und Snowflake-Schema
- Slowly Changing Dimensions

Speicher- und Indexstrukturen

- ROLAP und MOLAP
- Partitionierung
- Row Stores, Column Stores und In-Memory
- Bitmap-Indexe
- Mehrdimensionale Indexstrukturen

Data Warehouse: Anfragen und Verarbeitung

- OLAP-Anfrageoperatoren
- SQL-Operatoren im Data Warehouse
- Anfrageplanung
- Materialisierte Sichten

Dieses Lehrbuch behandelt Konzepte und Techniken von Data-Warehouse-Systemen, die eine wesentliche Komponente in betrieblichen Entscheidungsprozessen darstellen. Im Mittelpunkt stehen dabei Architekturprinzipien sowie die Umsetzung des multidimensionalen Datenwürfels als zentrale Komponente des Data Warehouse. Die Zusammenführung der Daten aus verschiedenen betrieblichen und externen Quellen spielt eine ebenso wichtige Rolle wie Datenstrukturen und Algorithmen für die Realisierung von Speicher- und Indexstrukturen. Die Navigation im Datenwürfel und die Anfrageverarbeitung sowie Anwendungen aus dem Themenfeld Business Intelligence geben einen Einblick in den Umgang mit dem Data Warehouse.

Detailliert werden sowohl der Aufbau als auch die Nutzung von Data-Warehouse-Systemen beleuchtet. Dabei stehen Modellierungskonzepte und die Thematik der multidimensionalen Anfragen im Vordergrund. Zudem werden Interna wichtiger Systemlösungen von Oracle, IBM und Microsoft anhand zahlreicher Beispiele erläutert.

Das Buch fokussiert auf relationale Umsetzungsstrategien des Data Warehouse. Es ist daher empfehlenswert, sich ebenfalls mit den Grundlagenwerken *Datenbanken - Konzepte und Sprachen* sowie *Datenbanken - Implementierungstechniken* auseinanderzusetzen; sie erlauben es dem Leser, die Konzepte aus Datenbanken für das Data Warehouse leichter zu transferieren. Das Buch ist geeignet für Studierende der Informatik oder verwandter Fächer im Masterbereich und bietet gleichzeitig auch dem Anwender bzw. Entwickler vertiefende Hintergrundinformationen zu aktuellen Data-Warehouse-Technologien.

Die Autoren lehren und forschen im Bereich Datenbanken und Informationssysteme sowie Business Intelligence – Veit Köppen und Gunter Saake an der Universität Magdeburg und Kai-Uwe Sattler an der TU Ilmenau.



www.mitp.de

Außerdem bei mitp:



ISBN 978-3-8266-9453-0



ISBN 978-3-8266-9467-7



ISBN 978-3-8266-9690-9

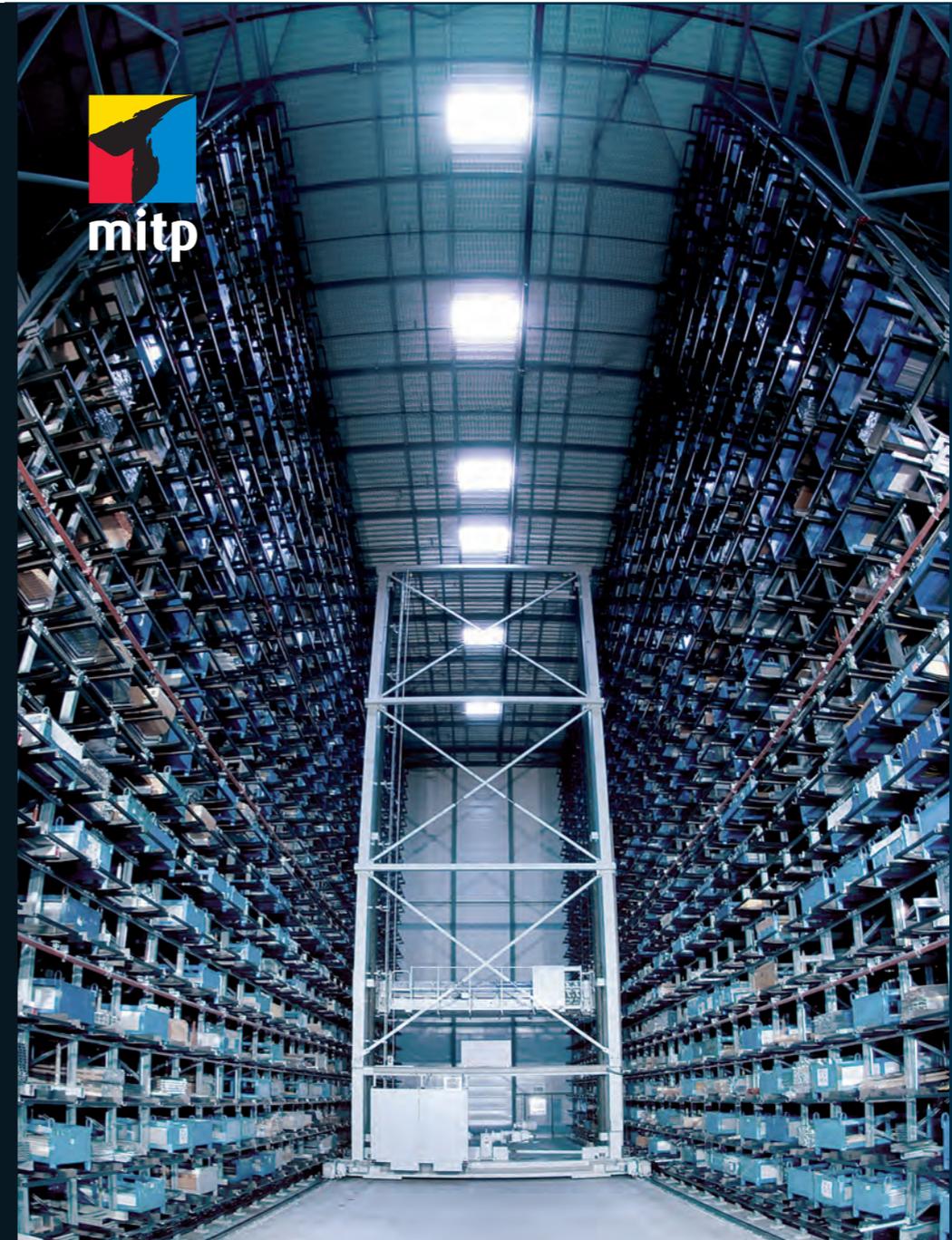


(D) € 29,99



Köppen
Saake
Sattler

Data Warehouse Technologien



Data Warehouse Technologien

Veit
Köppen

Gunter
Saake

Kai-Uwe
Sattler

2. Auflage

Vorwort

Vorwort zur 2. Auflage

Auch wenn ein Lehrbuch eher die stabilen, langfristigen Grundlagen eines Gebietes zum Inhalt hat, hat es doch die schnelle Entwicklung gerade im Bereich der Datenanalyse erforderlich gemacht, bereits nach zwei Jahren eine zweite Auflage des vorliegenden Buches zu erstellen. Hinzugenommen haben wir Erläuterungen zur Abgrenzung zum aktuellen Hype-Thema „Big Data“ sowie technische Ergänzungen zu den Themen Schichtenarchitekturen, Hauptspeicherdatenbanken, Slowly Changing Dimensions und die Umsetzung im Reporting.

Natürlich bleiben bei einem Buchprojekt stets auch Punkte offen, so dass wir im vorliegenden Lehrbuch keine Vollständigkeit gewährleisten können. Daher nutzen wir in bewährter Weise am Ende jedes Kapitels weiterführende Literaturempfehlungen. Neben der Aktualisierung der Literaturliste haben wir auch Fehler der ersten Auflage korrigiert. Hier möchten wir uns bei der Vielzahl von Leserinnen und Lesern bedanken, die uns auf Fehler hingewiesen und unverständliche Passagen aufgedeckt haben.

Danksagungen

Ein besonderer Dank gilt Herrn Dr.-Ing. Thomas Leich, der uns auf eine Vielzahl von fehlerhaften Darstellungen und Formulierungen hingewiesen hat, die wir in dieser Auflage beheben konnten. Unser weiterer Dank gilt Sebastian Dorok, Alexander Grebhahn, Martin Schäler, Reimar Schröter und Thorsten Winsemann für konstruktive Beiträge zur Verbesserung des Lehrbuches.

Für die grafische Unterstützung wollen wir erneut Herrn David Broneske unseren Dank aussprechen. Letztlich muss auch das Projektmanagement für die Erstellung eines Lehrbuches unterstützt werden. Daher gilt unser Dank Herrn Ernst-Heinrich Profener vom mitp-Verlag für sein Engagement, uns jederzeit behilflich zu sein, und für sein Warten auf die finale Fassung dieser zweiten Auflage.

Schließlich wollen wir uns für die Unterstützung bei unseren Familien bedanken, die uns zahlreiche Stunden entbehren mussten, damit wir nur noch schnell die „kleineren“ Änderungen umsetzen können.

Wir hoffen, dass die vorliegende zweite Auflage auch weiterhin für die Lehre im Masterstudium hilfreiche Dienste erweist. Zusatzmaterialien wie vorlesungsbegleitende Folien finden sie auf der entsprechenden Webseite des Buches. Alle verbleibenden Mängel gehen selbstverständlich zu Lasten der Autoren. Wir freuen uns natürlich über ihr Feedback auf der Biberbuch-Webseite.

<http://www.biberbuch.de>

Magdeburg und Ilmenau, im April 2014

Veit Köppen, Gunter Saake und Kai-Uwe Sattler

Vorwort zur 1. Auflage

Die bekannten Datenbanksysteme haben ihre Wurzeln in der Unterstützung des operativen Betriebs – kurze Buchungen, Abrechnungen oder Bestellaufnahmen sind typische Operationen auf Datenbanken. Für diesen Typ gibt es mit den relationalen Datenbanken ein weitverbreitet akzeptiertes Modell. Die Ausbildung in relationaler Datenbanktechnologie und die Nutzung der Sprache SQL hat Einzug in die Ausbildung von Fachschulen bis hin zu Universitäten gefunden – dementsprechend vielfältig und ausgereift ist die Lehrbuchauswahl zu diesem Thema.

Reporting-Anwendungen sowie analytische Anwendungen, wie die Vorbereitung und Auswertung einer Werbekampagne oder eine Warenkorbanalyse, benötigen prinzipiell die Daten der operativen Systeme, stellen aber ganz andere Anforderungen an das System. So sind historische Daten für analytische Zwecke unverzichtbar, in operativen Systemen aber eher ein Ballast. Daher hat sich das Gebiet der Data-Warehouse-Systeme als eigenständiges Entwicklungsgebiet etabliert, das eine analytische Systemunterstützung für derartige Anwendungen speziell zuschneidert. Obgleich hier oft die relationale Datenbanktechnologie genutzt werden kann, ergibt sich eine Vielzahl andersgearteter technischer Probleme und Lösungen, sodass sich ein eigenes Buch zu diesem Themenkomplex anbietet.

Lehrbücher zum Data Warehousing betrachten oft den potenziellen Nutzen und den Aufwand beim Betrieb eines Data Warehouse, sind also eher aus betriebswirtschaftlicher oder Wirtschaftsinformatik-sicht geschrieben. Für technische Fragestellungen ist der Leser auf die Standardliteratur zu Datenbanken angewiesen. Diese Literatur fokussiert allerdings eher auf die Lösung der Datenhaltung in operativen Systemen. Diese Lücke in den Ausbildungsunterlagen soll das vorliegende Buch schließen. Es liegt damit aber auch in der Natur der Sache, dass ein Data-Warehouse-Buch nicht in sich abgeschlossen sein kann, sondern auf Grundlagen aus dem Datenbankbereich zurückgreift, die durch entsprechende Lehrbücher wie etwa [SSH13, SSH11] vermittelt werden. Die Autoren haben sich daher entschlossen, die Data-Warehouse-Teile aus den Biber-Büchern herauszunehmen und die dort in diesem Zusammenhang behandelten Themen in diesem Buch eigenständig und detailliert zu betrachten. Vorhandene Redundanzen in den Darstellungen dieser Bücher werden bei den nächsten Auflagen beseitigt.

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 Danksagun-

gen 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.

Das Buch basiert auf einer Vorlesung, die von den Autoren an der Universität Magdeburg und der TU Ilmenau angeboten wird. Eingeflossen sind dabei viele Ideen, Feedback und Materialien von Kollegen anderer Universitäten, die ähnliche Vorlesungen halten. Ihnen allen sei an dieser Stelle ausdrücklich gedankt, wobei wir Wolfgang Lehner, Ulf Leser und Felix Naumann besonders hervorheben möchten.

Ein außerordentlicher Dank gilt David Broneske für die Unterstützung bei Abbildungen und Beispielen. Für das Korrekturlesen bedanken wir uns bei Sebastian Bress, Ingolf Geist, Alexander Grebhahn, Julia Hawlitschek, Andreas Lübcke 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 und Ilmenau, im Juli 2012

Veit Köppen, Gunter Saake und Kai-Uwe Sattler