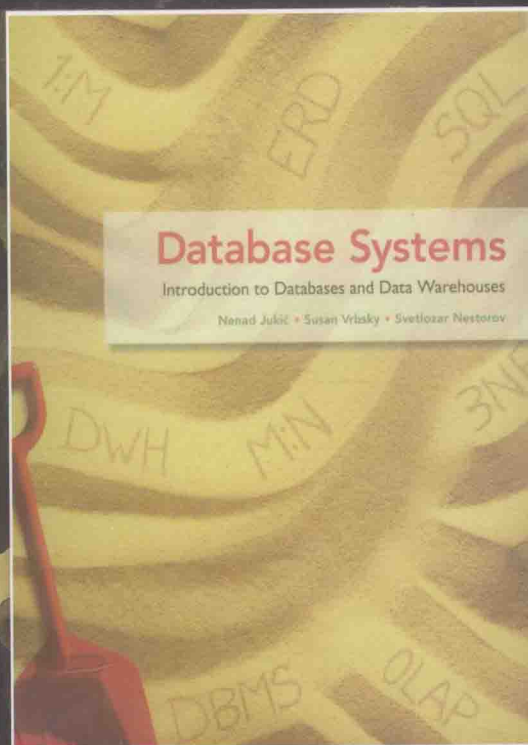


数据库系统

数据库与数据仓库导论

内纳德·尤基克 (Nenad Jukić) 著
[美] 苏珊·维布斯基 (Susan Vrbsky) 著
斯维特洛扎·奈斯特罗夫 (Svetlozar Nestorov) 著
李川 刘一静 等译

Database Systems
Introduction to Databases and Data Warehouses



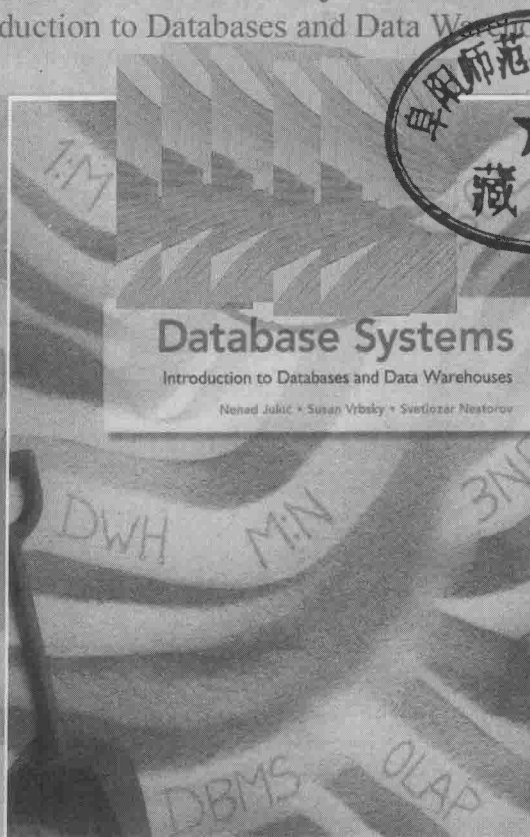
计 算 机 科 学 丛

数据库系统

数据库与数据仓库导论

内纳德·尤基克 (Nenad Jukić)
[美] 苏珊·维布斯基 (Susan Vrbsky) 著
斯维特洛扎·奈斯特罗夫 (Svetlozar Nestorov)
李川 刘一静 等译

Database Systems
Introduction to Databases and Data Warehouses



机械工业出版社

图书在版编目 (CIP) 数据

数据库系统:数据库与数据仓库导论 / (美) 尤基克 (Jukić, N.), (美) 维布斯基 (Vrbsky, S.), (美) 奈斯特罗夫 (Nestorov, S.) 著; 李川等译. —北京: 机械工业出版社, 2015.3 (计算机科学丛书)

书名原文: Database Systems: Introduction to Databases and Data Warehouses

ISBN 978-7-111-48698-5

I. 数… II. ① 尤… ② 维… ③ 奈… ④ 李… III. 数据库系统-高等学校-教材
IV. TP311.13

中国版本图书馆 CIP 数据核字 (2014) 第 281468 号

本书版权登记号: 图字: 01-2013-3388

Authorized translation from the English language edition, entitled *Database Systems: Introduction to Databases and Data Warehouses*, 9780132575676 by Nenad Jukić, Susan Vrbsky, Svetozar Nestorov, published by Pearson Education, Inc., Copyright © 2014.

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage retrieval system, without permission from Pearson Education, Inc.

Chinese simplified language edition published by Pearson Education Asia Ltd., and China Machine Press Copyright © 2015.

本书中文简体字版由 Pearson Education (培生教育出版集团) 授权机械工业出版社在中华人民共和国境内 (不包括中国台湾地区和香港、澳门特别行政区) 独家出版发行。未经出版者书面许可, 不得以任何方式抄袭、复制或节录本书中的任何部分。

本书封底贴有 Pearson Education (培生教育出版集团) 激光防伪标签, 无标签者不得销售。

本书基于互联网和商务智能应用环境, 介绍操作型数据库和分析型数据库的基本概念、设计方法和使用技巧, 兼具时效性、理论性和实用性。主要内容包括: 数据库基础知识, 数据库需求与 ER 建模, 关系数据库建模, SQL, 数据库的实现与使用, 数据仓库概念, 数据仓库与数据集市建模, 数据仓库的实现与使用, DBMS 功能与数据库管理。书中包含丰富的实例、图示、代码和练习, 配有教学网站和课程资源, 帮助读者举一反三、学以致用。

本书适合作为高等院校计算机相关专业数据库原理与设计课程的教材, 也可作为数据库技术人员的参考书。

出版发行: 机械工业出版社 (北京市西城区百万庄大街 22 号 邮政编码 100037)

责任编辑: 曲 熠

责任校对: 殷 虹

印 刷: 三河市宏图印务有限公司

版 次: 2015 年 4 月第 1 版第 1 次印刷

开 本: 185mm × 260mm 1/16

印 张: 23.75

书 号: ISBN 978-7-111-48698-5

定 价: 79.00 元

凡购本书, 如有缺页、倒页、脱页, 由本社发行部调换

客服热线: (010) 88378991 88361066

投稿热线: (010) 88379604

购书热线: (010) 68326294 88379649 68995259

读者信箱: hzsj@hzbook.com

版权所有·侵权必究

封底无防伪标均为盗版

本书法律顾问: 北京大成律师事务所 韩光 / 邹晓东

文艺复兴以来，源远流长的科学精神和逐步形成的学术规范，使西方国家在自然科学的各个领域取得了垄断性的优势；也正是这样的优势，使美国在信息技术发展的六十多年间名家辈出、独领风骚。在商业化的进程中，美国的产业界与教育界越来越紧密地结合，计算机学科中的许多泰山北斗同时身处科研和教学的最前线，由此而产生的经典科学著作，不仅擘划了研究的范畴，还揭示了学术的源变，既遵循学术规范，又自有学者个性，其价值并不会因年月的流逝而减退。

近年，在全球信息化大潮的推动下，我国的计算机产业发展迅猛，对专业人才的需求日益迫切。这对计算机教育界和出版界都既是机遇，也是挑战；而专业教材的建设在教育战略上显得举足轻重。在我国信息技术发展时间较短的现状下，美国等发达国家在其计算机科学发展的几十年间积淀和发展的经典教材仍有许多值得借鉴之处。因此，引进一批国外优秀计算机教材将对我国计算机教育事业的发展起到积极的推动作用，也是与世界接轨、建设真正的世界一流大学的必由之路。

机械工业出版社华章公司较早意识到“出版要为教育服务”。自1998年开始，我们就将工作重点放在了遴选、移译国外优秀教材上。经过多年的不懈努力，我们与 Pearson, McGraw-Hill, Elsevier, MIT, John Wiley & Sons, Cengage 等世界著名出版公司建立了良好的合作关系，从他们现有的数百种教材中甄选出 Andrew S. Tanenbaum, Bjarne Stroustrup, Brian W. Kernighan, Dennis Ritchie, Jim Gray, Alfred V. Aho, John E. Hopcroft, Jeffrey D. Ullman, Abraham Silberschatz, William Stallings, Donald E. Knuth, John L. Hennessy, Larry L. Peterson 等大师名家的一批经典作品，以“计算机科学丛书”为总称出版，供读者学习、研究及珍藏。大理石纹理的封面，也正体现了这套丛书的品位和格调。

“计算机科学丛书”的出版工作得到了国内外学者的鼎力相助，国内的专家不仅提供了中肯的选题指导，还不辞劳苦地担任了翻译和审校的工作；而原书的作者也相当关注其作品在中国的传播，有的还专门为其书的中译本作序。迄今，“计算机科学丛书”已经出版了近两百个品种，这些书籍在读者中树立了良好的口碑，并被许多高校采用为正式教材和参考书籍。其影印版“经典原版书库”作为姊妹篇也被越来越多实施双语教学的学校所采用。

权威的作者、经典的教材、一流的译者、严格的审校、精细的编辑，这些因素使我们的图书有了质量的保证。随着计算机科学与技术专业学科建设的不断完善和教材改革的逐渐深化，教育界对国外计算机教材的需求和应用都将步入一个新的阶段，我们的目标是尽善尽美，而反馈的意见正是我们达到这一终极目标的重要帮助。华章公司欢迎老师和读者对我们的工作提出建议或给予指正，我们的联系方式如下：

华章网站：www.hzbook.com

电子邮件：hzjsj@hzbook.com

联系电话：(010) 88379604

联系地址：北京市西城区百万庄南街1号

邮政编码：100037



译者序

Database Systems: Introduction to Databases and Data Warehouses

数据库技术诞生于 20 世纪 60 年代, 现已发展成为一门内容丰富的学科, 形成总产值达数百亿美元的产业。随着大数据时代的来临, 数据库技术正以惊人的速度将现实世界中的信息转化为数据存储到各类计算机系统中, 而且这一过程的发展态势可能超出人类的有限预想。其中蕴含的不仅是自然和生命, 还有人类的行为、情感和历史。同我们生存其中的真实自然界一样, 新兴的“数据自然界”中也潜藏着无尽的奥秘和巨大的财富, 因此吸引着大批科学界、人文界及商界的学者和技术人员投身其中。正确解读和有效利用这些数据是新技术革命时代指引人类前行的灯塔。

本书的基本内容包括 10 个章节及 10 个附录。第 1~6 章主要介绍操作型数据库, 第 7~9 章介绍分析型数据库、数据仓库与数据集市的相关内容, 第 10 章从一个更高的视角给出数据库管理相关概述。本书由浅入深, 由数据库的基本概念论及数据库的现状, 最后对数据库技术的发展趋势进行展望, 讲解过程中注重理论联系实际, 令读者的学习和理解更为深刻和透彻。本书既可作为初学者的入门教程, 也可供较有经验的用户借鉴, 特别适合作为大学本科数据库课程的参考读物。

本书的翻译工作是在极其紧张的条件下, 经过所有团队成员的艰辛拼搏最终完成的, 其中凝聚着每位参与者的真诚与责任。翻译工作由四川大学李川副教授统一负责, 刘一静老师协助进行最终统稿, 参与人员还有吴诗极、聂章艳、张永辉、李艳梅、谢世娜等。大家在节假日加班工作, 对译文字雕句琢, 最终有了本书的诞生。机械工业出版社的编辑在本书的出版过程中给予我们大力支持和理解, 没有他们的奉献, 也不可能有本书的面世。

尽管译者心正意诚, 然则受限于自身水平, 本书一定存在不少问题, 还期望各位读者给予批评指正, 各位的反馈将使本书更趋完善。最后, 真诚期望本书对大家有益, 这是对我们翻译工作的最大认可!

译者

2014 年 12 月 9 日夜

于四川大学家中

本书是一本入门级的综合性数据库教材，适用于研究生或本科生的信息系统数据库课程。虽然选择本书作为教材的学生大多已经掌握了一定的信息系统基础知识，然而从本书的内容安排来看，它依然适用于那些没有预修相关基础课程的学生。每章中都包含必要的引言，之后再详细介绍数据库各方面的知识，旨在站在数据库专家的高度，为读者和用户讲解数据库知识。

操作型数据库和分析型数据库

当今的实用领域，如信息系统、商业数据分析及决策支持查询等，都同时需要操作型和分析型两种数据库系统功能。本书在介绍这两种类型的数据库时兼具时效性、理论性和实用性的特点。采用本书作为教材的学生，将学会如何设计和使用操作型和分析型数据库，并将所学内容运用到现今的商业环境。

操作型数据库和分析型数据库都是目前信息系统研究的主流，因此应采取的教学方法是：即使应用环境仅涉及其中一种类型，我们也将以一种有意义的方式同时讲授两种类型的数据库知识。根据我们的想法和经验，采用本书作为教材可以实现简单明了的教学。我们已基于这本教材进行了多年的数据库课程教学（包括本科生教学和研究生教学，涵盖学期授课和季度授课），每次都能实现预定的教学目标，并且好评如潮。此外，我们还收到大批已毕业或在读的本科生、研究生以及用人单位（包括实习和全职）的书面反馈，证明学习过这门课程的学生在面临有数据库相关技能需求的公司招聘时胸有成竹。

本书特点

本书将所有关于设计和使用操作型、分析型数据库的基本内容分为 10 章及 10 个附录。

第 1～6 章主要介绍操作型数据库以及一些数据库基本问题，包括：数据库需求、ER 建模、关系建模、数据库约束、更新异常、规范化、SQL 语言、数据库前端、数据质量问题。

第 7～9 章主要介绍分析型数据库、数据仓库与数据集市的相关内容，包括：数据仓库概念、维度建模（星形模式）、数据仓库/数据集市建模方法、提取/转换/加载（ETL）处理、在线分析处理（OLAP）/商务智能（BI）功能、数据仓库/数据集市前端。

第 10 章从一个更高的视角（忽略细节）给出数据库管理相关概述。

附录（A、B、C、D、E、F、G、H、I、J）中是一些其他数据库相关主题的概述，包括：扩展 ER 建模（EER）、更高范式（高于第三范式）、企业资源计划（ERP）、数据管理和主数据管理、面向对象数据库、分布式数据库、并行数据库、云计算、数据挖掘、XML、NoSQL 数据库、大数据。

在本书的网站中，可以得到一个免费的基于 Web 的数据建模套件 ERDPlus（erdplus.com），该套件是结合本书内容设计和开发的。学生和教师可以利用该套件，特别是套件里针对学术研究的设置来创建 ER 图、关系模式及维度模型（星形模式）。我们鼓励教师和学生尝试该数据建模套件，体验其简洁性、易用性、轻便性及学术研究适用性。当然，也欢迎教师和学生使用其他工具来创建 ER 图、关系模型及维度模型。本书中所包含的练习可以用

ERDPlus 完成,也可以通过其他建模工具和方法(如 Visio、ERwin、ER/Studio、MS Word、MS Excel、MS PowerPoint、free drawing 等)完成。

配套资源

本书还有一些专为教师和学生准备的配套资源,内容如下。

配套网站(dbtextbook.com),包括:

- ERDPlus 的链接,它是一个数据建模套件(结合本书内容开发)。
- SQL 脚本及数据集。
- 关于免费得到 DBMS 和 OLAP/BI 软件的教程。
- 本书作者的邮箱。

教师资源网站(www.pearsonhighered.com/jukic)^①,包括:

- PPT(快速版、经典版及完整版)。
- 教师手册,包括:
 - 各章最后的复习题、练习及小案例的答案。
 - 额外的练习题(配有答案)。
 - 教学大纲及教学计划。
 - 教学方案。
- 测试题文件,内容为涵盖各章节内容的测试问题集,问题形式包括多项选择题、判断题、简答题,每一道题目都有相应章节的索引及难度等级。

智能课堂电子教材

智能课堂电子教材(CourseSmart eTextbook)可以节省学生购买教材或商家推荐教材的开销。学生可以使用主流信用卡获得所需课程的全部讲授内容,课程教材可以通过书名、作者来查找。使用智能课堂电子教材,学生能够搜索特定的关键词或页码,在线作标记,打印带有课堂记录的读书作业,并能标记出重要章节以方便复习。想获取更多信息或购买智能课堂电子教材,请访问 www.coursesmart.com。

教学方法

本书集合了描述清晰的理论概念、简单易懂的实例、广泛且具有普适性的实用构件。对于学生需要掌握的每一项技能,每一章最后都提供了大量相应的练习和小案例。

在大多数章末都有标题为“问题说明”的部分,这部分内容可作为额外主题进行选择阅读或是作为相同主题的内容延伸(取决于课程难度及课程时间)。

下面给出各章教学安排的概述。

第1章 引言

主要内容:本章给出各章内容的快速概览:数据库相关的基本术语、概念及组件,如数据和信息、数据库管理系统(DBMS)、数据库系统开发步骤,以及操作型数据库与分析型数据库的对比。

^① 关于本书教辅资源,用书教师可向培生教育出版集团北京代表处申请,电话:010-57355169/57355171,电子邮件:service.cn@pearson.com。——编辑注

教学方法：这一章采用了简短的描述和简明的实例，目的是简要地给出本书的总框架，为后面的章节做好铺垫。

第 2 章 数据库需求与 ER 建模

主要内容：实体-联系 (ER) 建模是将用户数据库需求形式化的一种概念化方法，本章将对其进行全面介绍。本章应用了陈氏 ER 标记的一种变形版本，当然，其他标记方法和概念数据建模方法也有提及。本章强调 ER 建模的目的是收集并可视化用户需求。本章将介绍 ER 模型的各种组件：实体 (包括弱实体)、属性 (常规属性、单一属性、复合属性、多值属性、派生属性)，以及一对一、一对多、多对多联系 (二元或一元)。

附加内容：章末探讨了 ER 建模方面的一些更深入的问题 (在相同实体之间的多实例多对多联系、联合实体，以及三元或更高元的联系)。

教学方法：本章内容基于实例，针对需求收集和 ER 模型创建提供了相应的综合案例。通过本章的练习、小案例以及免费软件 (ERDPlus-ER 图功能) 强化所介绍的概念。这一章的目的是让学生对 ER 建模相关的描述性概念以及需求可视化有深入了解，并通过大量实用练习来强化这些内容。

第 3 章 关系数据库建模

主要内容：本章全面介绍了关系数据库模型，包括关系概念、关系模式、完整性约束及用户自定义约束。此外，还讲解了将 ER 图 (实体、属性以及一元和二元的一对一、一对多、多对多联系) 映射为关系模式的过程。

附加内容：章末探讨了关于关系数据库建模的一些更深入的问题 (映射关联实体、映射三元联系、设计者创建的主码和自动编号选项、ER 建模和关系建模的必要性)。

教学方法：这一章的教学基于实例说明 ER 图概念的映射过程及关系模式的创建过程。本章的练习、小案例及免费软件 (ERDPlus-关系模式功能) 强化了本章所介绍的内容。这一章的目的是让学生对关系数据库建模概念有深入了解，并且通过大量练习来强化这些内容。

第 4 章 更新操作、更新异常与规范化

主要内容：本章介绍了更新操作 (插入、删除和修改)、规范化和更新异常 (以此说明规范化的必要性)，介绍和探讨了函数依赖的概念，同时还介绍了第一范式 (1NF)、第二范式 (2NF) 及第三范式 (3NF) (其他范式的介绍可以在附录 B 中找到)。

附加内容：章末探讨了有关规范化的更深入的问题 (规范化例外、逆规范化、规范化与 ER 建模、为流数据库内容增添新表)。

教学方法：本章内容基于实例，说明更新操作、更新异常及规范化过程，并通过练习强化所介绍的概念。这一章的目的是让学生对更新和规范化概念有深入了解，并且通过大量练习来强化这些内容。

第 5 章 SQL

主要内容：本章全面介绍了 SQL (结构化查询语言)，包括 SQL 创建、更新语句及关系型数据库查询，以及检索数据命令，如 SELECT 语句 (伴随多条件，采用 AND、OR 和 NOT 操作符)、聚集函数 (SUM、COUNT、AVG、MIN、MAX)、GROUP BY、ORDER

BY、HAVING、嵌套查询、UNION 和 INTERSECT 操作符、IN、EXISTS、各种连接、其他 SQL 语句和函数。

附加内容：章末探讨了有关 SQL 的更深入的问题（SQL 中观测值的不合理使用、SQL 标准，以及常见 RDBMS 工具包中的不同 SQL 语法）。

教学方法：本章内容基于实例，讲解如何用 SQL 语句实现关系数据库的建立、插入及查询，包含用 SQL 命令实现上述操作的具体实例。本书的 Web 主页（dbtextbook.com）提供了六种常见 DBMS 包（Oracle、MySQL、Microsoft SQL Server、PostgreSQL、Teradata、IBM DB2），其中包含了本章所有相关 SQL 语句的脚本。教师可以在本章授课中直接利用这些脚本在 RDBMS（根据教师选择）中复制、粘贴、执行 SQL 语句。通过这样的做法，教师可以向学生介绍 SQL 命令，同时还可以展示数据库的建立、插入和查询。数据集、练习、小案例可以强化本章内容。此外，本书 Web 主页介绍了怎样获得免费的、无访问限制的最新关系型 DBMS 软件。这一章的目的是让学生对 SQL 概念有深入了解，并且通过大量练习来强化这些内容。

第 6 章 数据库的实现与使用

主要内容：这一章讲述关于数据质量的问题——数据的准确性、完整性、一致性、唯一性、实时性及统一性。这些问题所涉及的数据都存储在数据库系统中。这一章还包括数据库前端接口（数据库形式、报表、应用）、参照完整性选项（删除和更新选项：级联、限制、设置为空、设置为默认值）、索引，以及用户自定义约束的实现。

附加内容：章末探讨了断言和触发器。

教学方法：本章在教学安排上将介绍一些精简但有意义的关于数据库实现和使用方面的最基本问题，这些问题在第 5 章中并未提及。本章将通过大量的例子来讲解所提到的概念，并且通过大量实用练习来强化这些概念。

第 7 章 数据仓库概念

主要内容：本章给出了数据仓库和数据集市这两个术语的定义，并且介绍了数据仓库的基本构件和基本概念（资源系统、ETL（提取、转换、加载）、集成的分析数据仓库、面向主题的数据库、OLAP/BI 前端）。同时，本章还给出了关于数据仓库开发步骤的概述。

教学方法：本章采用一些简短的描述和简要的实例对数据仓库进行入门级讲解，并为后面两章的内容做好简明扼要的铺垫。

第 8 章 数据仓库与数据集市建模

主要内容：本章介绍了维度建模——一种用于分析型数据库的概念和逻辑数据设计技术（如数据仓库和数据集市）。本章介绍的概念包括：维度表和事实表、星形模式、雪花模型、星座模型、缓慢变化维度。本章介绍了用于分析型数据库建模的 ER 建模技术（相对于第 2 章中提到的用于操作型数据库的 ER 建模技术）。本章还概述了不同数据仓库项目的开发方法：数据仓库总线结构（Kimball 方法）、一致维度、规范化数据仓库（Inmon 方法），以及独立数据集市。

附加内容：章末将维度建模和 ER 建模这两种数据仓库 / 数据集市设计方法进行了比较。

教学方法：这一章的教学通过实例说明基于单数据源或多数据源的维度模型（星形模

式)、精细的和聚集的事实表、缓慢变化维度,以及其他维度建模内容。本章还列举了 ER 建模或规范化的数据仓库实例。练习、小案例、免费软件(ERDPlus-星形模式功能)强化了本章所介绍的概念。这一章的目的是让学生对数据仓库和数据集市建模概念有深入了解,并且通过大量实用练习来强化这些内容。

第 9 章 数据仓库的实现与使用

主要内容: 本章给出 ETL 过程的概述,包括用于从操作型数据库中提取有用数据(以供分析使用)的相关基础设施和过程的建立,将这些数据进行格式变换以适应目标数据仓库模型的结构,通过数据清理和清洗来确保已变换数据的质量,以及将已变换和具有质量保证的数据加载到目标数据仓库的方法。本章定义了“在线分析处理”(OLAP)和“商务智能”(BI)这两个术语,它们通常与分析型数据库的前端应用有关。本章还介绍了 OLAP/BI 工具中的常用功能。

附加内容: 章末讨论了用于 OLAP/BI 工具的不同数据库及不同的 OLAP/BI 架构。

教学方法: 本章的教学基于实例说明所介绍的概念。配套的网站(dbtextbook.com)提供了如何免费且无限制地获取最新 OLAP/BI 软件的教程、数据集、练习题。本章简明扼要地讲解了有关数据仓库实现与使用方面的最基本问题。

第 10 章 DBMS 功能与数据库管理概述

主要内容: 本章给出了 DBMS 功能和组件的全面概述,以及数据库管理的相关问题,如数据安全、备份、恢复、性能及优化。

教学方法: 本章对所介绍的内容进行了快速概述,让学生粗略地了解 DBMS 的功能和数据库管理的相关内容。

附录

主要内容: 附录部分给出附加的数据库相关内容概述,包括扩展的 ER 建模(EER)、更高范式(高于第三范式)、企业资源计划(ERP)、数据管理与主数据管理、面向对象数据库、分布式数据库、并行数据库、云计算、数据挖掘、XML、NoSQL 数据库及大数据。

教学方法: 附录部分以简短的说明和实例来陈述内容,以此帮助学生粗略地了解一些数据库相关的附加内容。

致 谢

Database Systems: Introduction to Databases and Data Warehouses

我们要感谢各位审稿人在本书形成过程中提出的真知灼见，他们是：

Gary Baram, 天普大学

Jeff Hassett, 犹他大学

Emily Kelly, 芝加哥洛约拉大学

Barry King, 巴特勒大学

Mark Llewellyn, 中佛罗里达大学

Brian Mennecke, 艾奥瓦州立大学

Sree Nilakanta, 艾奥瓦州立大学

Janet Renwick, 阿肯色大学

John Russo, 温特沃斯理工学院

Julian Scher, 新泽西理工学院

Linda Zhou, 马里兰大学

真诚地感谢他们在时间、精力、建议、专业知识方面给予的慷慨帮助。感谢行业专家 Benjamin Korallus (普华永道)、Robin Edison 和 Douglas Marcis (OptumInsight)、Stan Ozenbaugh 和 Kevin Little (天睿公司)、Mark Watson 和 Abhishek Sharma (堪萨斯城联邦储备银行)、Ryane Bohm (通用电气公司)、Alexander Rukhostkiy (波音公司)、Creighton Lang (NTT 数据公司)、Zachary Relli (博思艾伦咨询公司), 以及 Gregory Roszczybiuk (李奥贝纳广告公司)。还要感谢我们的同事 Stefan Lessmann (汉堡大学)、Mary Malliaris 与 Faruk Guder (芝加哥洛约拉大学)、Michael Goul (亚利桑那州立大学)、Hugh Waston (佐治亚大学)、Boris Jukić (克拉克森大学), 以及 Miguel Velasco (明尼苏达大学)。同样感谢芝加哥洛约拉大学的学生: Melanie Ruiz、Jordan Braun、Subash Pant、Sarah Shea、Adam Chorazy、Anton Dokov 和 Karan Desai。

特别感谢世界级的数据库管理员 / 数据库设计者 / 程序设计者 Zoran Avramović, 他对本书的仔细审阅和直率见解促使我们坚持不懈地工作。同时特别感谢 Ivan Bare, 他在本书的准备阶段提供了大量的后勤保障及技术支持。

在此, Nenad 还想对他有幸遇到的信息技术领域的优秀教师表达谢意, 他们是: Zoran Vlačić 教授、Damir Kalpić 教授、Allen Parrish 教授及 Susan Vrbsky 教授。

在撰写本书的过程中, 我们得到了 Paul Gray 博士友好、中肯的建议及鼓励, 他已于 2012 年去世。Paul 是信息系统教育与研究领域的创始先驱之一, 在此对他表示深切的怀念。

Nenad Jukić
Susan Vrbsky
Svetlozar Nestorov

Nenad Jukić: 信息系统领域教授，芝加哥洛约拉大学昆兰商学院商务智能与数据仓库研究生课程中心负责人。

Jukić 博士从 1999 年起在昆兰商学院信息系统与运营管理系从事本科生、研究生和行政教育教学工作。2005 ~ 2007 年，Jukić 博士还是北京国际 MBA 项目位于北京大学的中国经济研究中心信息系统领域的访问教授。1997 ~ 1999 年，他在位于密歇根州艾伦代尔的伟大州立大学计算与信息系统学院任教。

Jukić 博士在克罗地亚萨格勒布大学计算与电气工程学院获得计算科学与电气工程学士学位，在位于美国阿拉巴马州塔斯卡卢萨的阿拉巴马大学获得硕士和博士学位。

Jukić 博士积极开展各种信息技术研究工作，涉及的领域包括：数据库建模与管理、数据仓库、商务智能、数据挖掘、电子商务和 IT 战略。他的研究成果公开发表于众多管理信息系统及计算机科学领域的学术期刊、会议出版刊物和书籍上。除学术工作之外，他还为多家公司和机构的数据库、数据仓库、商务智能项目担任技术顾问，涵盖创业公司、财富 500 强公司、美国政府部门及军事机构。

Suan V. Vrbsky: 副教授，阿拉巴马大学计算机科学研究生课程中心负责人。

Vrbsky 博士从 1992 年起在阿拉巴马大学计算机科学系从事本科生及研究生课程教学工作，现任阿拉巴马大学云计算与集群计算实验室主任。1983 ~ 1986 年，她同时在南伊利诺伊大学计算机科学系任教。

Vrbsky 博士在美国西北大学获得学士学位，在南伊利诺伊大学卡本代尔分校获得计算机科学硕士学位，在伊利诺伊大学香槟分校获得计算机科学博士学位。

Vrbsky 博士的研究领域是数据库及云计算，涵盖数据密集型计算、实时数据库、数据库安全、移动数据库及绿色计算。她在计算机科学学术期刊、会议出版刊物和书籍上合作发表了 100 余篇文章，并且是自然科学基金获得者。

Svetlozar Evtimov Nestorov: 芝加哥大学计算研究所高级研究助理。

Nestorov 博士曾任芝加哥大学计算机科学系助理教授，在此期间，他从事本科生和研究生数据库与计算机系统课程的教学工作。之后，他参与创立了 Mobissimo，这是一个得到风险投资的旅游搜索引擎，在 2004 年被《时代》杂志评选为 50 个最酷网站之一。

Nestorov 博士在斯坦福大学获得计算机科学与数学学士学位及计算机科学硕士学位，凭借“结构化和半结构化数据的数据挖掘技术”这篇论文获得该校的计算机科学博士学位。他的导师是 Jeffrey Ullman 教授。

Svetlozar 负责尼尔森公司数据中心数据仓库项目的设计和开发，这也是芝加哥大学布斯商学院的一部分。他感兴趣的研究领域还包括数据挖掘、高性能计算和 Web 技术。

推荐阅读



数据挖掘：概念与技术（第3版）

作者：Jiawei Han 等 译者：范明 等 ISBN: 978-7-111-39140-1 定价：79.00元



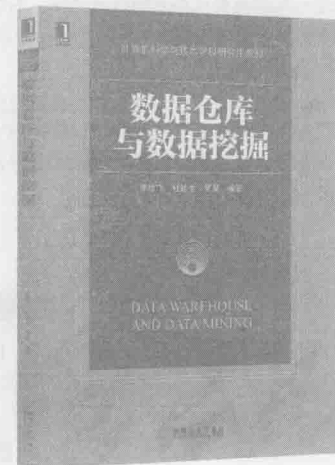
数据挖掘与R语言

作者：Luis Torgo 译者：李洪成 等 ISBN: 978-7-111-40700-3 定价：49.00元



R语言与数据挖掘最佳实践和经典案例

作者：Yanchang Zhao 译者：陈健 等 ISBN: 978-7-111-47541-5 定价：49.00元



数据仓库与数据挖掘

作者：李雄飞 等 ISBN: 978-7-111-43675-1 定价：39.00元

推荐阅读



数据集成原理

作者: AnHai Doan 等 译者: 孟小峰等 中文版: 978-7-111-47166-0 定价: 85.00元

**数据集成的第一部综合指南, 从理论原则到实现细节,
再到语义网和云计算目前所面临的新挑战。**

这是一本数据集成技术的权威之作, 书中的大部分技术都是作者提出来的。本书内容全面, 很多技术细节都介绍得非常清楚, 是数据集成相关工作人员的必读书籍。

—— Philip A. Bernstein, 微软杰出科学家

本书的三位作者对数据集成领域都有重要贡献, 既有学术背景, 又有工业界的经历。书中包含很多例子和相关信息, 以便于读者理解理论知识。本书包含了现代数据集成技术的很多方面, 包括不同的集成方式、数据和模式匹配、查询处理和包装器, 还包括Web以及多种数据类型和数据格式带来的挑战。本书非常适合作为研究生数据集成课程教材。

—— Michael Carey, 加州大学欧文分校信息与计算机科学Bren教授

数据库系统概念 (第6版)

作者: Abraham Silberschatz 等 译者: 杨冬青等 中文版: 978-7-111-37529-6 定价: 99.00元

中文精编版: 978-7-111-40085-1 定价: 59.00元

英文精编版: 978-7-111-40086-8, 定价: 69.00元

数据库领域的殿堂级作品

夯实数据库理论基础, 增强数据库技术内功的必备之选

对深入理解数据库, 深入研究数据库, 深入操作数据库都具有极强的指导作用!

本书是数据库系统方面的经典教材之一, 其内容由浅入深, 既包含数据库系统基本概念, 又反映数据库技术新进展。它被国际上许多著名大学所采用, 包括斯坦福大学、耶鲁大学、得克萨斯大学、康奈尔大学、伊利诺伊大学等。我国也有多所大学采用本书作为本科生和研究生数据库课程的教材和主要教学参考书, 收到了良好的效果。

出版者的话

译者序

前言

致谢

作者简介

第1章 引言	1
1.1 基本术语	1
1.2 数据库系统开发步骤	4
1.2.1 数据库需求收集、定义和 可视化	4
1.2.2 数据库建模	5
1.2.3 数据库实现	6
1.2.4 开发前端应用程序	6
1.2.5 数据库部署	6
1.2.6 数据库使用	6
1.2.7 数据库管理和维护	7
1.3 数据库新版本	7
1.4 数据库使用范围	7
1.5 参与数据库系统的人员	7
1.5.1 数据库分析者、设计者和 开发者	8
1.5.2 前端应用程序分析者和 开发者	8
1.5.3 数据库管理员	8
1.5.4 数据库终端用户	8
1.6 操作型数据库与分析型数据库	9
1.7 关系数据库管理系统	9
1.8 本书内容安排	9
关键术语	10
复习题	10

第一部分 操作型数据库

第2章 数据库需求与ER建模	14
2.1 引言	14
2.2 ER建模基本构件	14
2.3 实体	14
2.4 属性(唯一和非唯一)	15
2.5 联系	15
2.6 联系类型(最大基数侧)	17
2.7 联系和联系实例	18
2.8 联系属性	19
2.9 实例:数据库需求集及其ER图	20
2.10 复合属性	22
2.11 复合的唯一属性	23
2.12 多个唯一属性(候选码)	23
2.13 多值属性	24
2.14 派生属性	24
2.15 可选属性	25
2.16 实例:包含多种类型属性的 实体	26
2.17 联系中最小基数和最大基数 确切值	26
2.18 一元联系和联系的角色	27
2.19 相同实体间的多种联系	28
2.20 弱实体	29
2.21 实体、属性和联系的命名约定	31
2.22 多个ER图	32
2.23 实例:另一组数据库需求 及其ER图	33
2.24 数据库需求和ER模型的使用	34
2.25 各种ER符号体系	36
2.26 扩展的ER模型	37
2.27 问题说明:相同实体之间具有 多个实例的M:N联系	38

2.28	问题说明: 关联实体	41	3.18	相同实体间的多个联系的映射	74
2.29	问题说明: 三元 (及更高阶) 联系	42	3.19	弱实体的映射	75
	总结	46	3.20	实例: 将另一个ER图映射为 关系模式	77
	关键术语	48	3.21	关系数据库约束	79
	复习题	49	3.21.1	隐含约束	80
	练习	49	3.21.2	用户自定义约束	80
	小案例	50	3.22	问题说明: 关联实体映射	83
第3章	关系数据库建模	54	3.23	问题说明: 三元联系映射	83
3.1	引言	54	3.24	问题说明: 设计者创建的主码 和自动编号选项	85
3.2	关系数据库模型基本概念	54	3.25	问题说明: ER建模和关系建模	86
3.3	主码	56		总结	87
3.4	将实体映射为关系	56		关键术语	87
3.5	将具有复合属性的实体映射 为关系	57		复习题	88
3.6	将具有唯一复合属性的实体 映射为关系	58		练习	88
3.7	将具有可选属性的实体映射为 关系	59		小案例	89
3.8	实体完整性约束	59	第4章	更新操作、更新异常与规范化	91
3.9	外码	60	4.1	引言	91
3.10	将联系映射为关系数据库 组件	60	4.2	更新操作	91
3.10.1	1:M联系的映射	61	4.2.1	插入操作实例	91
3.10.2	M:N联系的映射	63	4.2.2	删除操作实例	91
3.10.3	1:1联系的映射	65	4.2.3	修改操作实例	92
3.11	参照完整性约束	66	4.2.4	关于更新操作术语的说明	92
3.12	实例: 将ER图映射为关系模式	67	4.3	更新异常	92
3.13	将拥有若干候选码 (多个唯一 属性) 的实体映射为关系	69	4.3.1	实例场景	92
3.14	将具有多值属性的实体映射为 关系数据库组件	70	4.3.2	实例关系 (包含冗余数据)	93
3.15	将具有派生属性的实体映射为 关系	71	4.3.3	插入异常	95
3.16	实例: 将具有多种类型属性的 实体映射为关系模式	71	4.3.4	删除异常	95
3.17	一元联系的映射	72	4.3.5	修改异常	95
3.17.1	1:M一元联系的映射	72	4.4	函数依赖	95
3.17.2	M:N一元联系的映射	73	4.5	函数依赖实例	96
3.17.3	1:1一元联系的映射	74	4.6	简化函数依赖	98
			4.6.1	增广函数依赖	98
			4.6.2	等价函数依赖	99
			4.7	函数依赖类型	100
			4.7.1	部分函数依赖	100
			4.7.2	完全函数依赖	100
			4.7.3	传递函数依赖	101
			4.8	另一个函数依赖实例	101

4.9 规范化	103	5.19 别名	145
4.9.1 第一范式	103	5.20 多关系连接	146
4.9.2 第二范式	105	5.21 ALTER TABLE	147
4.9.3 第三范式	106	5.22 UPDATE	148
4.9.4 其他范式	108	5.23 DELETE	148
4.9.5 消除冗余和更新异常	108	5.24 CREATE VIEW 和 DROP VIEW	149
4.10 另一个规范化实例	110	5.25 集合运算: UNION、INTER- SECT、EXCEPT(MINUS)	150
4.11 问题说明: 规范化例外情况	112	5.26 使用其他 SQL 命令的实例	151
4.12 问题说明: 逆规范化的规范化 与性能	113	5.27 CREATE TABLE (附加实例)	152
4.13 问题说明: ER 建模和规范化	114	5.28 INSERT INTO (附加实例)	154
4.14 问题说明: 用于简化数据库 内容的设计者添加的 实体(表)和码	115	5.29 约束管理	157
关键技术语	117	5.30 SELECT (附加实例)	158
复习题	118	5.31 关系与自身的连接(自连接)	158
练习	118	5.32 OUTER JOIN	158
第 5 章 SQL	123	5.33 无主码/外码组合的连接	160
5.1 引言	123	5.34 IS NULL	160
5.2 SQL 命令综述	123	5.35 EXISTS	160
5.2.1 数据定义语言	123	5.36 NOT	161
5.2.2 数据操纵语言	124	5.37 从查询中插入关系	161
5.2.3 数据控制语言和事务控制 语言	124	5.38 其他 SQL 功能	162
5.3 SQL 数据类型	124	5.39 问题说明: SQL 中观测值使用 不当	162
5.4 SQL 语法简要说明	124	5.40 问题说明: SQL 标准和 SQL 语法差异	163
5.5 CREATE TABLE	125	5.40.1 SQL 语法差异 1: DATE 和 TIME 数据类型	163
5.6 DROP TABLE	127	5.40.2 SQL 语法差异 2: FOREIGN KEY	163
5.7 INSERT INTO	128	5.40.3 SQL 语法差异 3: 别名 关键词 AS 的使用	164
5.8 SELECT	130	5.40.4 SQL 语法差异 4: ALTER TABLE	164
5.9 WHERE	132	5.40.5 SQL 语法差异 5: 集合运算	165
5.10 DISTINCT	133	5.40.6 SQL 语法差异 6: FULL OUTER JOIN	166
5.11 ORDER BY	134	5.40.7 SQL 语法差异 7: 约束管理	166
5.12 LIKE	135		
5.13 聚集函数	135		
5.14 GROUP BY	136		
5.15 HAVING	139		
5.16 嵌套查询	141		
5.17 IN	142		
5.18 JOIN	143		