HZ BOOKS
华章科技

Mc Graw Hill Education

商务智能与信息化技术丛书

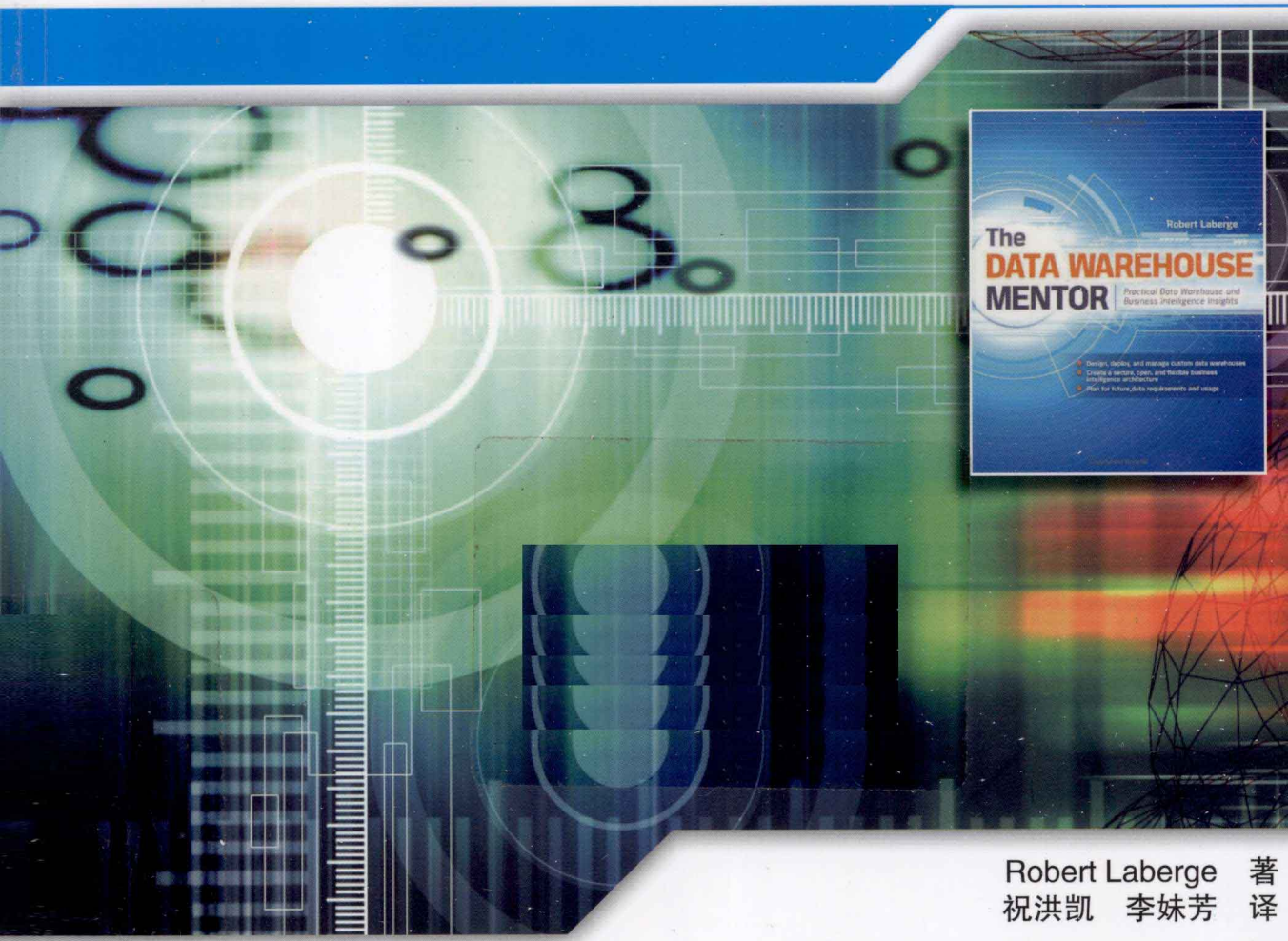
设计、部署和管理自定义数据仓库
创建安全、开放和灵活的商务智能架构
规划未来数据需求和使用

数据仓库应用指南

数据仓库与商务智能最佳实践

The Data Warehouse Mentor

Practical Data Warehouse and Business Intelligence Insights



Robert Laberge 著
祝洪凯 李妹芳 译

机械工业出版社
China Machine Press

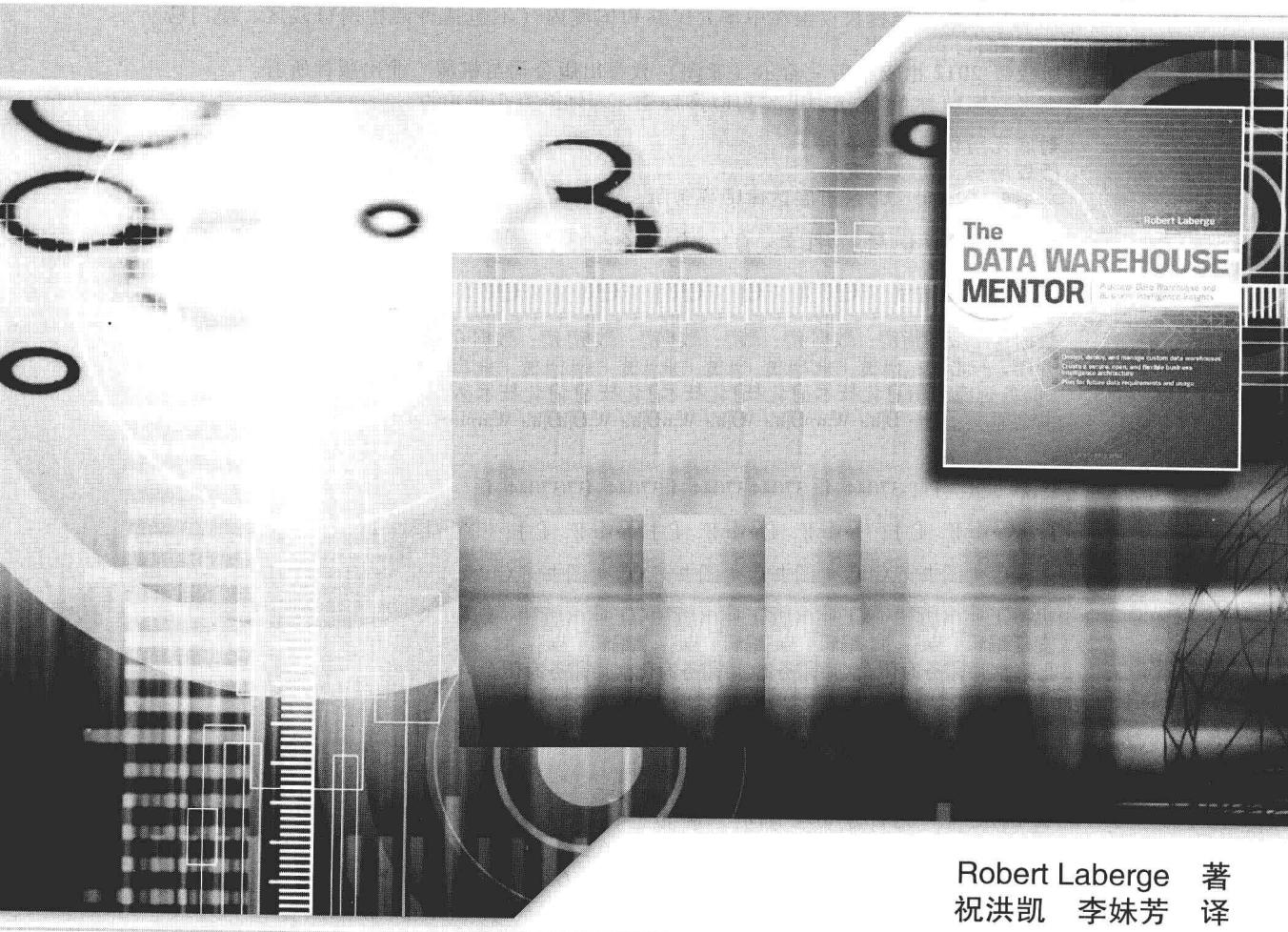
商务智能与信息化技术丛书

数据仓库应用指南

数据仓库与商务智能最佳实践

The Data Warehouse Mentor

Practical Data Warehouse and Business Intelligence Insights



Robert Laberge 著
祝洪凯 李妹芳 译

机械工业出版社
China Machine Press

Robert Laberge: The Data Warehouse Mentor: Practical Data Warehouse and Business Intelligence Insights (ISBN: 978-0-07-174532-1).

本书版权登记号：图字：01-2011-3603

图书在版编目（CIP）数据

数据仓库应用指南：数据仓库与商务智能最佳实践/拉伯格（Laberge, R.）著；祝洪凯，李妹芳译. —北京：机械工业出版社，2012.3
（商务智能与信息化技术丛书）
书名原文：The Data Warehouse Mentor: Practical Data Warehouse and Business Intelligence Insights

ISBN 978-7-111-37044-4

Ⅰ.数… Ⅱ.①拉… ②祝… ③李… Ⅲ.数据库系统 Ⅳ.TP311.13

中国版本图书馆 CIP 数据核字（2012）第 002411 号

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：吴　怡
北京京师印务有限公司印刷
2012 年 3 月第 1 版第 1 次印刷
186mm×240mm · 21 印张
标准书号：ISBN 978-7-111-37044-4
定价：69.00 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
购书热线：（010）68326294；88379649；68995259
投稿热线：（010）88379604
读者信箱：hzjsj@hzbook.com

译 者 序

数据仓库和商务智能在各个领域的应用已经如火如荼。本书作者给我们分享了他30多年的工作经验。本书涉及面广泛，内容全面，从基础概念的介绍、各个组件的剖析，到实践中的问题，作者都给出了细致的描述，深入浅出、高屋建瓴地阐述了数据仓库和商务智能的方方面面。本书理论和实践相结合，是一本不错的数据仓库和商务智能方面的整体指南。

本书分为三个部分来讲解数据仓库这个复杂系统，以及实现商务智能的有效方案。第一部分介绍了关于商务智能和数据仓库的基础概念，旨在介绍基础知识，为管理者思考为何、如何建立数据仓库提供了思考的方向。第二部分介绍数据仓库系统的基础组件，这部分涉及数据仓库和商务智能系统的技术方面，探讨了如何建立数据仓库系统，来维护企业资产并提供商务智能支持工作。第三部分从实践角度说明了如何构建数据仓库系统，包括经典的构建场景以及后期工作。

说实话，业余翻译了几本书，总是很担心译得不好误导了读者。在翻译的过程中我们查阅了大量文献和网络资源，为了译好一些不常见的词汇也反复琢磨了作者原意。但是还是觉得时间紧迫，翻译仓促，惶恐不安。在这里要感谢机械工业出版社华章分社编辑吴怡老师的很多辛苦付出，也感谢所有其他为本书付出努力的编辑们。

由于时间、精力、能力有限，本书的疏漏、错误之处在所难免，还望各位读者不吝指正。

前　言

本书对数据仓库世界中很多主题进行了探讨。本书旨在从业务和技术角度说明数据仓库系统的构建，侧重于简单朴实地描述如何构建切实的解决方案。这些见解来源于我30多年在20多个国家中50多家企业的亲身经历，在这些经历中，我曾作为独立顾问、员工以及IBM产业模式和资产实验室的合伙人，见证了很多数据仓库的实施过程。

本书介绍了构建数据仓库的组件和不同选择，以及选择某种方式的利弊。每家企业的数据仓库构建都是具有其特色的，但可以借鉴全球范围内很多企业的各种数据仓库和商务智能环境中获取的知识。本书首先从高层角度介绍了数据仓库主题，以确保对术语和上下文理解一致，然后详细说明了各个主题。这些主题都和数据仓库、商务智能和性能管理相关。

对于数据仓库的构建不存在规则，但是有很多指南。本书的主要根本点是根据具体的和对业务需求的理解，构建适应特定企业需求的解决方案，同时为今后的工作创建一个开放、灵活的架构基础。很多企业在初始包含商务智能报表的集中式数据仓库的构建上花费了大量的预算，结果却发现其创建的解决方案过于具体，只适合一两个用途，而无法满足后期的需求。当然，我们无法对未来进行预测，但是可以在一定程度上预期今后的数据需求和使用方式，确保设计和构建环境灵活、开放，对于变化可扩展而不需要每次重新设计和构建。

很多企业的领导人意识到企业数据是企业的基础资产，必须对它进行组织、结构化和维护，以保证其业务信息有较好的质量和管理，从而在整个企业范围内共享。如果没有信息系统，企业就无法运作，而如果没有商业目的，信息系统就不复存在。它们相互依存，应该充分意识到信息架构和使用方式，以使得企业变得更加智慧。

本书结构

第一部分：准备

第一部分介绍商务智能和数据仓库的基础概念，旨在介绍基础知识，为管理工

作奠定基础。

第 1 章：数据仓库和商务智能概述　该章概要介绍了商务智能和数据仓库，最后提出了和数据仓库实现相关的高层次问题。

第 2 章：企业中的数据　该章探讨了数据如何作为企业资产，并提出关于如何组织数据的见解。

第 3 章：为什么创建数据仓库　该章探讨了支持和反对构建数据仓库的各种理由。“支持”的理由在于已经有一些构建数据仓库的经典场景，而“反对”的理由在于企业的文化和局限性能否推动项目向前发展。

第 4 章：数据仓库和商务智能战略　该章给出了构建数据仓库和商务智能行动的一些规划，探讨从何处以及如何启动项目，这取决于这项工作是面向商业报表解决方案，还是努力将数据进行组织和结构化。

第 5 章：项目资源：角色和洞察力　该章讨论了数据仓库项目的关键角色，以及最佳实践的团队结构。

第 6 章：项目总结概论　该章简要介绍了项目章程、项目范畴和工作说明书的内容。

第二部分：组件

第二部分介绍了数据仓库系统的基础组件，深入分析了数据仓库和商务智能系统的技术方面。这部分具体探讨了数据仓库系统中用以维护企业资产和提供商务智能支持工作的各个组件。

第 7 章：商务智能：数据集市及其使用方式　该章从数据模型到性能问题，详细探讨了数据集市及其使用方式。

第 8 章：企业数据模型　该章讨论了企业数据模型、如何构建企业数据模型的一些实例以及一般问题。

第 9 章：数据仓库架构：组件　该章从建模和数据流角度探讨了数据仓库架构的不同类型。

第 10 章：ETL 和数据质量　该章探讨了数据仓库中的数据采集层和分发层的一些普遍特征，并提出关于数据质量问题的一些见解。

第 11 章：项目规划和方法论　该章讨论了数据仓库和商务智能项目规划的一些方法。

第三部分：构建

第三部分从实践角度说明了如何构建数据仓库系统。这部分旨在介绍经典的构建场景和工作，以及数据监理和对后期工作的审查。

第 12 章：工作场景　该章介绍了如何使用自上而下、自下而上和混合式方法来构建数据仓库和商务智能系统，并讨论了一些其他主题，包括简要介绍企业信息架构。

第 13 章：数据监理　该章探讨了企业数据监理，包括企业结构、数据质量、所有权和变更管理。

第 14 章：项目后评审　该章探讨了数据仓库和商务智能项目在开发完成后的一些方面。

本书力争做到成为构建数据仓库系统的完整指南，目标是理解当今数据仓库系统中的很多问题，并从多个角度提出自己的观点。作者希望本书能够帮助你构建好自己的数据仓库。

希望你喜欢本书！

致谢

特别感谢本书的技术编辑 David Marcotte 和 Ken Yu，感谢他们帮我审查书稿，并提出他们的观点和建议。他们的宝贵意见为本书最终的成功出版提供了方向指引。真心感激！

David Marcotte 是全球性零售业和分析业的大师，他拥有该行业的渊博知识。

Ken Yu 是数据设计和数据流方面的技术专家，他在很多部门工作过，积累了对数据建模和数据架构的很多实践经验和常识性方法。

特别感谢我的妻子 Rakhee Laberge（工商管理硕士），感谢她对本书的审查以及提出的很多意见和建议。

Robert（Bob）Laberge

联系方式：datawarehousementor@gmail.com

作者简介

Robert (Bob) Laberge 是多家互联网企业的创始人、IBM 产业模式和资产实验室的首席顾问，他的研究重点是数据仓库和商务智能解决方案。

Bob 早在 20 世纪 70 年代末就开始其职业生涯，当时比尔·盖茨还只是一个百万富翁，Bob 曾经是开发人员、数据库管理员、数据建模师、项目经理、数据架构师、企业信息架构师、数据仓库/商务智能审计员、战略师，而且还是富于创新的企业家。从那时，Bob 就跑遍全球，通过设计、优化、最佳实践和在概念层、逻辑层和物理层的常识说明，提供指导、培训和证明数据仓库和商务智能实践经验和解决方案。Bob 成功地帮助了 50 多家大型企业扩展业务，这些企业涉及零售、保险、医疗、铁路、电信、电子商务和银行等行业。

Bob 拥有英国 Durham 大学的工商管理硕士学位。你可以通过 datawarehousementor@ gmail. com 联系他。

技术编辑简介

David Marcotte 是 Kantar Retail 的美洲零售分析的高级副总裁，其专业领域涉及国际销售、采购、零售业、供应链和商务智能。他在很多技术主题上发表过演讲，并且是高深层会议的培训师，和全球范围的众多财富 500 强公司合作。最近，他是可口可乐零售研究会的《2010 年零售业新兴市场研究》的研究人员和作者，该研究总结了在中国、秘鲁、土耳其、波兰、南非和巴西的调查结果。在加入 Kantar Retail 之前，他是 IBM 的全球商务智能组的负责人。在 IBM 工作期间，他参与开发团队，提供零售业的一系列解决方案，包括损失预防、定价、分类管理过程、商品规划、供应链、商品销售和销售团队优化。他还是 IBM 零售数据仓库解决方案的业务和项目负责人，在这个项目中，他首次和本书的作者 Bob Laberge 合作。他为很多

的零售商和制造商提供解决方案，这些公司包括索尼、Sobeys、D&S、Falabella、Shoppers Drug Mart 和 Kroger 等。David 在很多行业委员会和专题讨论组工作过，包括 IRI 产品咨询委员会和高效消费者响应组（ECR）。你可以通过 David. Marcotte@KantarRetail. com 和他联系。

Ken Yu 在信息管理和数据架构领域拥有超过 25 年的工作经验。他曾经在不同行业和应用中担任过企业信息架构师、项目架构师、数据建模师、数据仓库管理员和开发人员。1995 年，Ken 加入 IBM 以支持大型数据库客户端工作，担任业务负责人。作为信息管理顾问，他在整个企业范围内提供关于数据监理、主数据和元数据管理、数据建模和参考模型的价值等主题的指导和培训。他一直是众多关键任务应用以及商务智能和数据仓库解决方案的主要设计、开发、技术架构和业务过程参与人员。当前，Ken 是 Cypselurus 公司的首席顾问，该公司在加拿大安大略省的多伦多市。

目　　录

第二部分 组 件