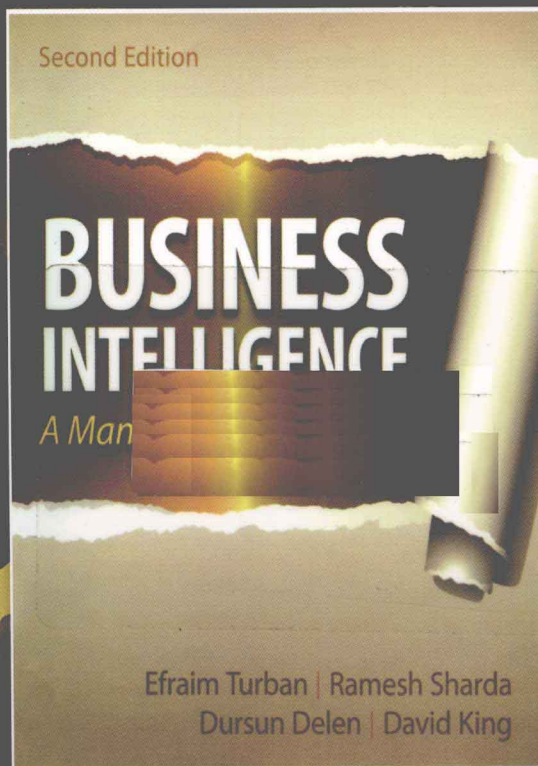


商务智能 管理视角

(美) Efraim Turban Ramesh Sharda Dursun Delen David King 著
秦秋莉 姚家奕 王英 译

Business Intelligence
A Managerial Approach Second Edition



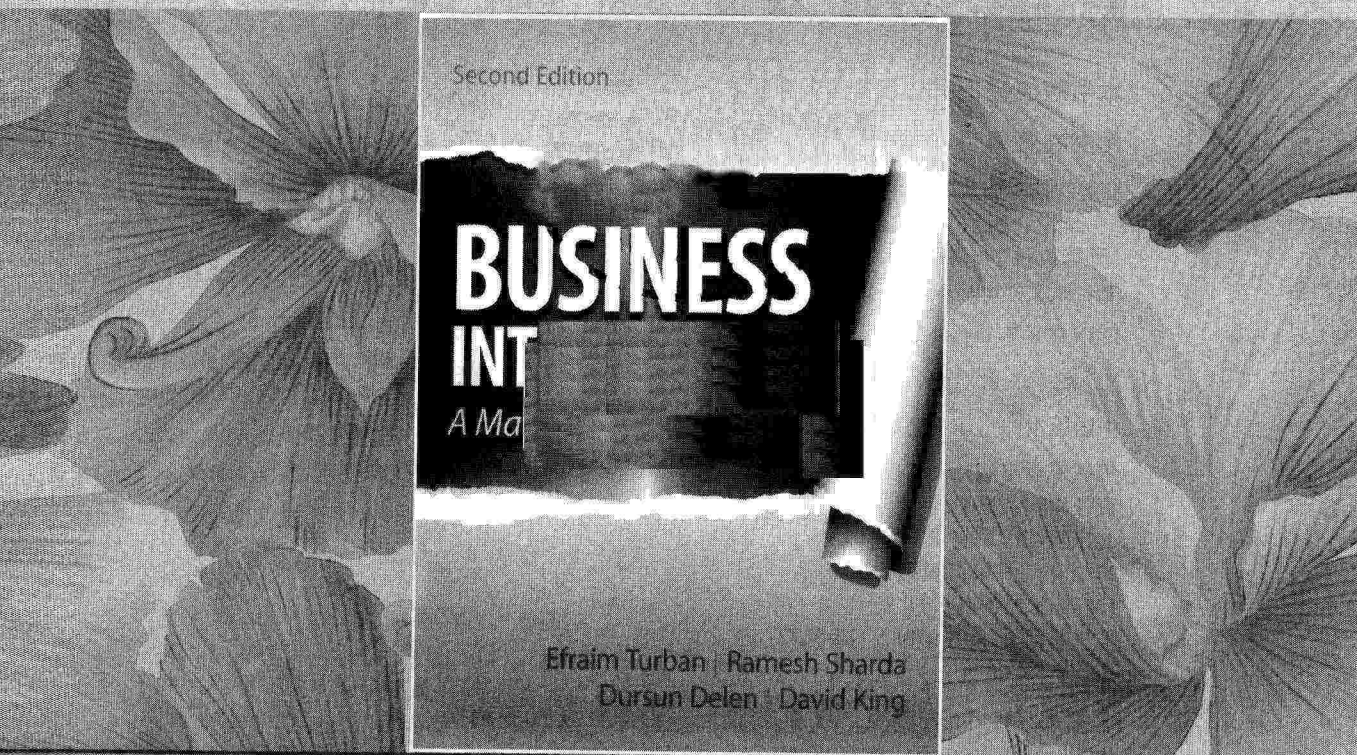
计 算 机 科 学 丛 书

原书第2版

商务智能 管理视角

(美) Efraim Turban Ramesh Sharda Dursun Delen David King 著
秦秋莉 姚家奕 王英 译

Business Intelligence
A Managerial Approach Second Edition



机械工业出版社
China Machine Press

本书主要介绍商务智能、知识管理、数据挖掘和其他智能系统的基础和应用知识,并通过例子、产品、服务和练习,以及基于网络的问题讨论扩展了读者对网络世界的理解。第2版的改进主要集中在3个领域:数据挖掘、文本挖掘和Web挖掘、实施和新技术。

本书可以作为商务智能课程的教材,也可以作为管理信息系统简介或者商务战略的教材,还可以作为MBA技术管理课程的补充读物,或者注重管理视角的管理科学和管理信息系统项目的教材。

Authorized translation from the English language edition, entitled BUSINESS INTELLIGENCE, 2E, 9780136100669 by Efraim Turban, Ramesh Sharda, Dursun Delen, David King, published by Pearson Education, Inc., publishing as Prentice Hall, Copyright © 2011.

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 © 2012.

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。

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

封底无防伪标均为盗版

版权所有,侵权必究

本书法律顾问 北京市展达律师事务所

本书版权登记号:图字:01-2010-6650

图书在版编目(CIP)数据

商务智能:管理视角(第2版)/(美)特班(Turban, E.)等著;秦秋莉,姚家奕,王英译. —北京:机械工业出版社,2011.12

(计算机科学丛书)

书名原文:Business Intelligence: A Managerial Approach

ISBN 978-7-111-36727-7

I. 商… II. ①特… ②秦… ③姚… ④王… III. 电子商务-教材 IV. F713.56

中国版本图书馆CIP数据核字(2011)第254566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:盛思源

北京瑞德印刷有限公司印刷

2012年2月第1版第1次印刷

185mm×260mm·14.75印张

标准书号:ISBN 978-7-111-36727-7

定价:55.00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

文艺复兴以降,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中取得了垄断性的优势;也正是这样的传统,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襄助,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,有的还专程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,我们的联系方法如下:

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



华章教育

华章科技图书出版中心

译者序

Business Intelligence: A Managerial Approach, 2E

商务智能是一门新兴的边缘学科，近年来引起了学术界和产业界的广泛关注。商务智能是企业利用现代信息技术收集、管理和分析结构化和非结构化的商务数据和信息，创造和累计商务知识和见解，改善商务决策水平，采取有效的商务行动，完善各种商务流程，提升各方面商务绩效，增强综合竞争力的智慧和能力。

商务智能的出现是一个渐进、复杂的演进过程，而且仍处在发展中。20 世纪 90 年代后期，商务智能出现了突飞猛进的发展，越来越多的企业提出了他们对商务智能的需求，把商务智能作为帮助企业达到经营目标的一种有效手段。计算机界很多著名公司已经认识到商务智能巨大的发展潜力，纷纷加入到商务智能研究和软件开发的行列。商务智能技术已从决策支持系统开始，经历了数据仓库、联机分析、数据挖掘的发展，最后到达了可视化信息技术阶段。2010 年，从全球范围来看，商务智能领域并购不断，商务智能市场已经超过 ERP 和 CRM 市场成为最具增长潜力的领域。从中国市场来看，商务智能已经度过了一个从知到行，从概念到实践的阶段。当前金融业、电信业、零售业、服务业都在广泛地应用各种类型的商业智能工具，体验到了数据分析、报告、挖掘的力量，在经营分析、客户选择、绩效管理、运行预警方面得到很大的帮助。

本书作者是具有多年在 IT 相关领域的学术界和产业界工作的博士和专家，不仅承担了很多相关领域研究项目，在国际学术一流期刊发表了大量的学术论文，而且具有在跨国公司从事咨询工作的经验，因此可以从全方位角度向读者介绍商务智能的技术，展现如何从管理的视角去理解商务智能。

正如本书作者所言，本书可作为商务智能课程的教材，也可作为管理信息系统简介或商务战略的教材，同时还可以作为 MBA 技术管理课程的补充读物，或者是注重管理视角的管理科学和管理信息系统项目的教材。另一个目标是向实践管理者提供商务智能、知识管理、数据挖掘和其他智能系统的基础和应用。

本书第 2 版较第 1 版有了很大的改动，内容更加丰富、形式更加新颖、资料更加全面，同时配有生动及时的网站，相信可以满足广大读者的要求。

本书第 2 版包括 6 章。第 1 章和第 6 章，目录、前言和术语，由北京交通大学秦秋莉博士、副教授负责翻译；第 2 章、第 3 章和第 5 章由北京交通大学姚家奕副教授负责翻译；第 4 章由北京交通大学王英老师负责翻译。全书的译文由秦秋莉负责统一定稿。

感谢机械工业出版社的编辑们，是他们的远见使得本书能够尽快与读者见面。

需要特别指出的是，2010 年我接手了本书的翻译工作。当时正值我将以访问学者身份前往美国亚利桑那大学 ELLER 商学院管理信息系统系人工智能实验室（AI LAB）访学交流之际。通过这一年的访学过程，我更好地理解了数据挖掘和商务智能，对翻译本教材具有很大的帮助。特别是 Chen Hsinchun 教授的 AI LAB 实验室研究课题和他教授的“数据挖掘”课程资料，Sudan Ram 教授的“商务智能”课程，Paulo Goes 教授的“商务智能中的数据挖掘”课程，Zhu Zhang 助理教授的“数据挖掘与网络控制”课程，对我深入了解数据挖掘和商务智能提供了很大的帮助，在此表示诚挚的感谢。

由于译者水平有限，译文中的错误和不当之处，敬请读者朋友指正。意见请发往 qlqin@sohu.com，我们将不胜感激。

希望读者喜欢这本译著，希望这本译著有助于进一步推动我国的商务智能研究与应用的深入开展。

秦秋莉

2011 年 11 月于美国亚利桑那大学人工智能实验室 AI LAB

基于计算机的决策支持的应用越来越广泛。许多公司，比如 IBM、Oracle 和 Microsoft 正在建立基于分析的组织单元，以便商业运作更加灵活有效。越来越多的懂计算机和网络的决策者使用更多的计算机工具来支持他们的工作。同时，消费者和组织在交互的过程中产生着不计其数的数据。这些数据存储能够用来开发和提升适当的产品、服务，促进消费者消费，优化组织内的运作。

本书的目的就是向读者介绍商务智能技术。在一些圈子里，商务智能（Business Intelligence, BI）也称为商务分析（business analytic）。我们不加区别地使用这两个术语。本书介绍了这些系统构建和使用中的基本技术和方式。

第2版的改进主要集中在3个领域：数据挖掘、文本和 Web 挖掘、实施和新技术。这一版尽管有这么多的变化，但仍然保留了综合性和用户友好性，这使得本书在市场上占据领先地位，最终呈现给读者最准确及时的知识，而这在别的教科书上是没有的。

第2版更新了什么

第2版增加了许多最新的内容，删掉了过时的内容。以下是几个重要的变化：

- **全部修订或新增的章节。**第5章（彻底进行了修订）以全面但易懂的方式研究了2个最流行的商业分析工具。这章提供了很多应用案例，使得主题有趣并且吸引读者。第6章是新增加的，这章调查了几个新的现象，这些现象正在或有可能改变商务技术和实践，它们分别是：无线射频识别（Radio Frequency Identification, RFID）、云计算、社交网络、Web 2.0、虚拟世界等。商务智能实施的重要主题和面向需求的计算战略也增加到了这一章中，同时第6章还更新了计算机化的决策支持对个人、组织和社会的影响。
- **流线型覆盖。**我们通过删除在线的格式文本来缩短教材篇幅，但是我们仍然保留了大量通用的内容。我们通过网站及时提供最新的内容和一些链接。另外，减少了每一章参考文献的数量。而且，我们对第1章中商务智能的介绍性的报道和数据挖掘实现了流水线式，这种综述形式能够让学生在学期开始提前准备如何以一个学术项目的方式进行学习和思考（指导老师可能也需要）。我们还删除了本章网上已有的内容并且合并了一些内容。
- **新的作者团队。**与上一版相比，本书增加了一名作者，还有一名作者的角色扩充了。前一版作者（Turban、Sharda、Aronson 和 King）构建了很好的内容，Ramesh Sharda 和 Dursun Delen 主要修订这一版。Ramesh 和 Dursun 主要工作在决策支持系统和数据挖掘领域，他们具有相关的行业和研究经验。
- **新的幻灯片图形。**尽管印刷版本中图形保留了以前版本中的内容，但是新内容中增加了新的图，所有的图都重新画成了彩色，并且可以从在线图库中获得，用于幻灯片演示。
- **更新及时的网站。**本书的使用者能够进入网站，网站包括与本书主题相关的新故事、软件、学习指南，甚至还包括 YouTube 的视频。
- **重新修订和更新的内容。**所有的章节都有新的开篇场景和结束案例，这些都是基于最近的真实应用故事和事件。除此之外，更新了贯穿本书的应用案例，包括最近的一个特殊技术/模型应用的例子。新的网站链接也增加到本书中。本书删掉了许多旧产品的链接和参考文献。最后，许多章节都有新的练习、网络练习、讨论题等。

第2版其他特殊改动归纳如下：

- 第1章 包括新的开篇故事和结尾案例，以及整章内流线型的材料。
- 第2章 包括数据仓库的新材料，包括大量的在线分析处理（Online Analytical Processing, OLAP）和多维数据模型。一款改编的可亲自动手的 MicroStrategy 软件演示也加入了本书。除此之外，我们还将架构和实施问题部分实现了流水线型结构。最后，还加入了数据仓库的未来部分。
- 第3章 是以上一版多个章节为基础合并而成的。除了更新的开篇场景、结束案例以及整篇的讨论这种流水线型结构，还包括了新的关键绩效指标（Key Performance Indicator, KPI）和运营指标、精益六西格玛、数据可视化、业务流程管理架构等部分。
- 第4章 采用体现标准的数据挖掘项目流程的方法，全面深度地展现数据挖掘的内容。与上一版相应的章节相比，本章重新撰写，使得读者更加容易使用和获得数据挖掘的信息。特别是延伸了文本和 Web 挖掘（有专门的章节），更重要的是扩展了数据挖掘的方法和方法论。本章还详细地描述了人工神经网络和它们在管理决策方面的使用，重点详细地描述了最流行的人工神经网络架构，阐述了它们在不同决策类型问题中使用的差别。这一章还加入了人工神经网络模型灵敏度分析的解释。
- 第5章和第6章 是新增章节。

我们保留了许多上一版不错的内容并更新了相关的内容，这些内容概括如下：

- **Teradata 大学网络（TUN）的连接。**许多章节都包括 TUN 的连接（teradatauniversitynetwork.com）。Teradata 网站（Teradata 学生网站 TSN，teradatastudent.network.com）主要包括了学生的作业。登录 TSN 网站允许学生阅读案例，观看网站研讨会，回答问题，查询资料等。
- **减少穿插的小板块，组织结构更合理。**小板块减少了 50% 以上，重要素材合并到正文中。现在只有两个小板块：应用案例和技术前沿。
- **软件支持。**TUN 网站提供了免费的软件支持。除此之外，网站还提供了软件的练习，我们的网站还包括了附加的一些链接。

目标和覆盖范围

目前组织能够很容易地使用内部网和因特网发布高价值的性能分析应用给全球的决策者。公司定期地开发分布式系统、内部网和外部网，用来方便地访问存储在许多地点的协作和通信数据。各种信息系统应用彼此集成起来，或者与其他基于网络的系统集成起来，有些集成甚至超越了组织的边界。由于管理者手头有很多精确的信息，所以他们能够更好地决策。

今天的决策支持工具使用网络来进行分析，它们使用图形用户界面，使得决策者可以利用熟悉的网络浏览器更加灵活、有效和容易地观看和处理数据和模型。企业信息、知识和其他高级系统的这种容易使用和阅读能力移植进了个人电脑和个人数字助理。管理者们通过使用一系列的无线掌上设备，包括移动电话和个人数字助理，通过电脑和网络进行沟通。这些设备使得管理者能够访问重要的信息和有用的工具，进行沟通和协作。数据仓库及其分析工具（例如在线分析处理、数据挖掘）极大地提高了穿过组织边界的信息接入和分析。

本书可作为商务智能课程的教材，也可作为管理信息系统简介或商务战略的教材，同时也可以作为 MBA 技术管理课程的补充读物，或者是注重管理视角的管理科学和管理信息系统项目的教材。本书的另一个目标是给实践管理者提供商务智能、知识管理、数据挖掘和其他智能系统的基础和应用。

这次修订版本的主题是商务智能和为企业决策提供支持的商务分析。除了传统的商务智能

应用外,这一版全篇通过例子、产品、服务和练习,以及基于网络的问题讨论扩展了读者对网络世界的理解。本书突出网络智能和网络分析,这些和电子商务及其他应用的商务智能和商务分析是平行的。本书有网站支持 (pearsonhighered.com/turban),网站上提供了许多在线的文件。通过特殊的网站部分还可以链接到许多软件指导书。

特点

- **面向管理视角** 商务智能能够从两个视角来看:技术和管理。本书主要从第二个视角来看。本书包括许多关于商务智能的应用和实施的展示。然而,我们承认技术的重要性,因此每1章恰当地表述了技术概要。在学习指南中以博客链接到本书网站的形式提供一些技术细节。
- **面向真实世界** 大量有效的有关全世界大公司、小企业、政府、非营利组织的案例使得概念更加生动。这些案例给学生展示了商务智能的能力、成本和使用它的理由,以及现实中公司在内部使用商务智能的创新方法。
- **Teradata 大学网络 (TUN) 连接** TUN 是由 Teradata 赞助的一个免费学习平台,是 NCR 的一个分支,其目标是帮助职员与其他商务智能领域的同仁之间进行学习、教学、沟通和合作。几百所大学和学院都参与和使用 TUN。Teradata 也支持学生平台 (teradatastudentnetwork.com),上面包含了大量学习资源,例如案例、网络研讨会、学习指南、练习和资源的链接。本书与 TUN 主要通过给学生提供使用平台来完成各章不同类型的作业相链接。
- **大多是当前的主题** 有证据显示,本书提供的内容大多数是在 2008 年和 2009 年被引证的有关商务智能的主题。
- **集成系统** 相比其他教材重视孤立的基于互联网的商务智能系统,本书强调那些支持企业和企业众多用户的系统。
- **全球视角** 全球竞争、合作伙伴、贸易的重要性正在快速地增长,因此,全书提供的都是国际案例。
- **在线内容** 可以在线获取本书的附加材料。包括家庭作业的数据文件、许多报告、视频和软件的链接。
- **用户友好性** 本书不仅覆盖所有主要的商务智能主题,而且清晰、简单、结构组织合理。它提供了所有基本定义和逻辑概念支持。进一步说,本书易懂,并且充满有趣的真实世界案例,能激发读者的阅读兴趣。在每节结束还提供相关复习题,以便读者复习和消化新的材料。

补充包: Pearsonhighered.com/Turban

一个全面灵活的基于技术支持的补充包可以用来提高教学和学习经验。下面的教师指导补充材料会放在本书网站 pearsonhighered.com/turban 上,采用本书作为教材的教师需要联系所在地的培生代表处,申请下载权限:

- **指导手册** 指导手册包括整门课程和各章节的学习目标,各章结尾的问题和练习答案,教学建议(包括项目指导)。
- **幻灯片** 描述和构建了本书关键概念的幻灯片(可以下载获得)。
- **测试题和 TestGen 测试软件** 测试题包括判断正误、多项选择、填空、问答等类型。问题都是按照难度水平分类的。答案都引用了书的页码。测试题可以采用微软的 Word 格式和 Pearson TestGen 的计算机格式。TestGen 是一个全面的测试和评价套件工具,它允许指导老师创立和发布课程测试题,可以采用传统打印后分发的方法,也可以通过本地局

域网在线分发。TestGen 特征向导帮助用户在整个项目中移动，软件获得了全技术支持。

- **在线课程材料** Pearson Prentice Hall 支持本书使用者使用在线课程，通过黑板课程管理系统提供文件上传，用来测试、考问和其他辅助。请联系你所在地的培生代表处，获取特殊课程的进一步资料。

致谢

许多人帮助我们完成了本书和第 1 版出版后的修订。首先，我们感谢那些为本书和另一本决策支持系统教材——《Decision Support and Business Intelligence Systems》（第 9 版，Prentice Hall，2011）正式审稿的人们，他们给我们提供了很大帮助，他们是：

Ann Aksu	中部皮德蒙特社区学院
Bay Arinze	德雷克塞尔大学
Banjit Bose	新墨西哥大学
Kurt Engemann	约那学院
Badie Farah	东密歇根大学
Gary Farrar	哥伦比亚大学
Jerry Fjermestad	新泽西技术学院
Martin Grossman	布里奇沃特州立大学
Jahangir Karimi	科罗拉多大学丹佛分校
Huer Lee	东密歇根大学
Natalie Nazarenko	纽约州立大学弗雷多尼分校
Kala Chand Seal	洛约拉玛莉曼特大学
Roger Wilson	费尔门特州立大学
Vincent Yu	密苏里大学理工学院
Fan Zhao	佛罗里达高尔夫海岸大学

其次，感谢那些为我们的文本和支撑材料提供资料的人。Barbara Wixom 为第 1 章撰写开场白，介绍了 Teradata 大学网络与本书的关系。Dan Power (Dssresources.com 和北爱荷华州大学) 允许我们使用他的“虚拟世界”栏目上的信息。Deborrah C. Turban (菲律宾圣托马斯大学) 为本书第 6 章做出了贡献。还要感谢亚利桑那州立大学的 Haluk Demirkan。

第三，感谢为第 1 版做出贡献的 Janine Aronson (乔治亚大学)，他是我们的合作者，为数据仓库一章做出了贡献，Mike Goul (亚利桑那州立大学) 为第 1 章做了贡献。T. P. Liang (台湾中山大学)，为神经网络贡献了资料。

第四，感谢提供开发和演示软件的零售商，Acxiom (小岩城，阿肯色州)、California Scientific Software (内华达城，加利福尼亚州)、Catalyst Development 的 Cary Harwin (尤卡谷，加利福尼亚州)、Demandtec (圣卡罗，加利福尼亚州)、DS Group 公司 (格林威治，康涅狄格州)，KD-Nuggtes.com 的 Gregory Piatetsky-Shapiro (盖恩斯维尔，佛罗里达州)、NeuroDimension 公司 (盖恩斯维尔市，佛罗里达州) 的 Gary Lynn、Promised Land Technologies (新罕布什尔，康涅狄格州)、Salford Systems (拉荷亚，加利福尼亚州)、Sense Network (纽约，纽约州)、Statsoft 公司的 Gary Miner (塔尔萨，俄克拉何马州)、Ward System Group 公司 (弗雷德里克，马里兰州)、Wordtech Systems (奥林达，加利福尼亚州)。

第五，特别感谢 Teradata 大学网络及其执行总裁 Michael Goul；TUN 的副总裁 Barb Wixom；TUN 的创始人 Hugh Watson；项目经理 Susan Baxley；Teradata 的 Mary Gros，她是 Teradata 与学术团体的联络人。正是由于他们的鼓励促成这本书和 TUN 的联系，为本书提供了许多有用的材料。

第六，感谢许多帮助我们处理行政事务、编辑、校对、制图和做准备工作的人员。包括：Subramanian Rama Lyer（俄克拉荷马州立大学）、Mike Henry（俄克拉荷马州立大学）、Angie Jungermann（俄克拉荷马州立大学）、Brittany Solomon（俄克拉荷马州立大学）和 Ivan C. Soballos II（菲律宾的 DE LA SALLE LIPA）。Judy Lang 与我们所有的人进行合作，包括编辑，在第 1 版的制作过程指导我们。

最后，Pearson Prentice Hall 团队非常令人称许：Bob Horan 执行主编精心安排了这个项目，编辑项目经理 Kelly Loftus 给我们控制时间表，Shanthi Lakshmipathy 为我们的手稿做文字编辑工作，正是由于 Pearson Prentice Hall 的创作团队和 PreMedialGlobal 的工作人员，才使我们的手稿变成了一本书。

我们在此感谢以上所有人员的合作，没有他们的帮助，本书就不可能创作完成。

E. T.

R. S.

D. D.

D. K.

作者简介

Business Intelligence: A Managerial Approach, 2E

Efraim Turban (工商管理硕士、博士, 加利福尼亚大学伯克利分校) 是夏威夷大学太平洋管理信息系统研究院的访问学者, 他曾经任职于多所大学, 包括: 香港城市大学、里海大学、佛罗里达国际大学、加利福尼亚州立大学长滩分校、东伊利诺伊大学, 以及南加利福尼亚大学。Turban 博士在一流学术期刊上发了 100 多篇学术论文, 例如《Management Science》、《MIS Quarterly》、《Decision Support Systems》。他还出版了 20 本书, 包括:《Electronic Commerce: A Management Perspective》、《Information Technology for Management》。他还是世界各地许多公司的咨询师。Turban 博士目前的研究领域是基于 Web 的决策支持系统、智能代理在电子商务系统中的使用, 以及全球电子商务的协同问题。

Ramesh Sharda (工商管理硕士、博士, 美国威斯康星大学麦迪逊分校) 是信息系统研究所 (the Institute for Research in Information Systems, IRIS) 的所长, 康菲石油公司技术管理主席, 俄克拉荷马州立大学 (Oklahoman State University, OSU) 西尔斯工商管理学院管理科学与信息系统的杰出贡献教授。Ramesh Sharada 博士在许多学术期刊上发表了 100 多篇与他的研究相关的论文, 这些期刊包括:《Management Science》、《Information Systems Research》、《Decision Support Systems》、《INFORMS Journal on Computing》、《Production Operations Management》、《Journal of Management Information Systems》、《Interface》等。他是信息系统协会决策支持系统和知识管理特殊团体 (SIGDSS) 的共同创办人。Sharada 博士服务于多个编辑委员会, 包括《INFORMS Journal on Computing》、《Decision Support Systems》、《ACM Transactions on Management Information Systems》。他还是《Annals of Operations Research》的编辑, 施普林格 (Springer) 集成信息系统系列和操作研究计算机科学界面丛书的编辑。他目前的研究兴趣是决策支持系统、协同应用、信息过载管理技术。他为许多组织进行咨询, 并且在国际执行教育项目进行讲学。Sharada 博士也是 iTradeFair.com 的共同创办人, 这是一家运营虚拟交易的公司。

Dursun Delen (博士, 俄克拉荷马州立大学) 是俄克拉荷马州立大学西尔斯工商管理学院管理科学与信息系统的副教授。他在 2001 任命为俄克拉荷马州立大学助理教授之前, 他在得克萨斯州大学城的基于知识的系统公司工作, 作为研究员工作了 5 年。期间, 他领导了一系列与决策支持和其他信息系统相关的研究项目, 资金是由联邦机构资助的, 如: 美国国防部 (DoD)、美国国家航空和宇宙航行局 (NASA)、美国国家标准技术研究所 (NIST)、美国能源部 (DOE)。他的研究成果发表在许多一流学术期刊上, 包括:《Decision Support Systems》、《Communications of the ACM》、《Computers and Operations Research》、《Computers in Industry》、《Journal of Production Operations Management》、《Artificial Intelligence in Medicine》、《Expert Systems with Applications》等。最近他和 David Olson 教授出版了一本有关高级数据挖掘技术的书。他是《International Journal of RF Technologies: Research and Applications》期刊的副主编, 他还是《Journal of Information Technologies》、《International Journal of Intelligent Information Technologies》、《Journal of Emerging Technologies in Web Intelligence》、《International Journal of Service Sciences》的编委会成员。他的研究和教学兴趣是决策支持系统、数据和文本挖掘、知识管理、商务智能, 以及企业建模。

Dave King (博士) 具有 25 年主持决策支持开发、性能管理、企业系统软件的经验。目前, 他是位于亚利桑那州斯科茨代尔市的 JDA 软件公司新产品开发部的副部长。他在 Comshare 公司作为首席信息官和产品开发部高级副部长工作了几年后, 于 2004 年加入 JDA。King 博士发表了许多论文和书籍, 他是《Electronic Commerce: A Managerial Perspective》(Prentice Hall) 一书的合作者。他还服务于许多工业咨询委员会和大学董事会, 包括乔治亚大学的管理信息系统咨询委员会和亚利桑那州立大学专家政治咨询委员会。

出版者的话

译者序

前言

作者简介

第1章 商务智能简介 1

开篇场景: Norfolk Southern 利用商务智能

进行决策支持获取成功 1

1.1 变化的商务环境和计算机化的决策
支持 3

1.2 商务智能框架 4

1.2.1 BI 的定义 5

1.2.2 BI 的历史 5

1.2.3 BI 的架构 6

1.2.4 BI 的形式 8

1.2.5 BI 的好处 8

1.2.6 事件驱动预警 10

1.3 智能创造和使用与商务智能治理 11

1.3.1 智能创造和使用的循环过程 11

1.3.2 智能与窃取 12

1.4 交易处理和分析处理 12

1.5 成功的 BI 实施 13

1.5.1 典型的 BI 用户群体 13

1.5.2 合适的计划及其与商业战略的
一致性 131.5.3 实时的、基于需求的 BI 是可
达到的 14

1.5.4 开发或获得 BI 系统 14

1.5.5 理由和成本-利润分析 14

1.5.6 隐私安全和保护 15

1.5.7 系统集成和应用 15

1.6 商务智能的主要工具和技术 15

1.6.1 技术和工具 15

1.6.2 选择 BI 供应商 16

1.7 本书计划 16

1.8 相关资源、链接和 Teradata 大学网络
的连接 17

1.8.1 资源和链接 17

1.8.2 案例 17

1.8.3 供应商、产品和演示 17

1.8.4 期刊 17

1.8.5 Teradata 大学网络的连接 17

1.8.6 本书的网站 18

本章重点 18

关键术语 18

讨论题 18

练习 19

本章结尾应用案例 19

参考文献 20

第2章 数据仓库 21

开篇场景: DirecTV 的蓬勃发展与实时

数据仓库 21

2.1 数据仓库的定义和概念 23

2.1.1 什么是数据仓库 23

2.1.2 数据仓库的特点 23

2.1.3 数据集市 24

2.1.4 业务数据存储 24

2.1.5 企业数据仓库 24

2.1.6 元数据 26

2.2 数据仓库流程概述 27

2.3 数据仓库架构 28

2.3.1 可选的数据仓库架构 30

2.3.2 哪种架构是最好的 33

2.4 数据集成功以及提取、转换和加载的
过程 34

2.4.1 数据集成 34

2.4.2 提取、转换和加载 36

2.5 数据仓库的开发 38

2.5.1 数据仓库供应商 39

2.5.2 数据仓库开发方法 40

2.5.3 数据仓库开发的其他思考 41

2.5.4 数据仓库中的数据表示 42

2.5.5 数据仓库中的数据分析 43

2.5.6 OLAP 与 OLTP 43

2.5.7 OLAP 操作 43

2.6 数据仓库的实施问题 46

2.7 实时数据仓库	50	3.8.2 商业 BPM 套件	89
2.8 数据仓库管理系统、安全问题和 未来发展趋势	53	3.8.3 BPM 市场与 BI 平台市场对比	90
2.9 相关资源、链接和 Teradata 大学 网络的连接	55	3.9 绩效仪表盘和记分卡	90
2.9.1 资源和链接	55	3.9.1 仪表盘和记分卡	91
2.9.2 案例	55	3.9.2 仪表盘设计	91
2.9.3 供应商、产品和演示	56	3.9.3 仪表盘展示的内容	92
2.9.4 期刊	56	3.9.4 数据可视化	92
2.9.5 其他参考文献	56	本章重点	94
2.9.6 Teradata 大学网络的连接	56	关键术语	96
本章重点	57	讨论题	96
关键术语	57	练习	96
讨论题	58	本章结尾应用案例	98
练习	58	参考文献	100
本章结尾应用案例	60	第 4 章 商务智能中的数据挖掘	102
参考文献	61	开篇场景：数据挖掘来到好莱坞	102
第 3 章 业务绩效管理	64	4.1 数据挖掘概念和定义	104
开篇场景：Harrah 公司加倍下注	64	4.1.1 定义、特征和好处	106
3.1 业务绩效管理概述	66	4.1.2 数据挖掘的工作原理	109
3.1.1 BPM 定义	67	4.2 数据挖掘应用	112
3.1.2 比较 BPM 和 BI	67	4.3 数据挖掘流程	115
3.2 制定战略：我们想到哪里去	68	4.3.1 步骤 1：理解业务	115
3.2.1 战略规划	69	4.3.2 步骤 2：理解数据	115
3.2.2 战略差距	70	4.3.3 步骤 3：数据准备	116
3.3 计划：我们如何达到那里	70	4.3.4 步骤 4：建模	118
3.3.1 运营计划	70	4.3.5 步骤 5：测试和评估	118
3.3.2 财务计划和预算	71	4.3.6 步骤 6：部署	119
3.4 监控：我们做得怎么样	71	4.3.7 其他标准化数据挖掘过程和 方法	120
3.4.1 诊断控制系统	72	4.4 数据挖掘方法	121
3.4.2 差异分析的困难	72	4.4.1 分类	121
3.5 行动和调整：我们需要做什么不同 的吗	74	4.4.2 分类模型正确性估算	122
3.6 绩效评价	76	4.4.3 数据挖掘聚类分析	127
3.6.1 KPI 和业务指标	76	4.4.4 关联规则挖掘	128
3.6.2 现有绩效评价系统存在的问题	77	4.5 数据挖掘中的人工神经网络	130
3.6.3 有效的绩效指标	78	4.5.1 人工神经网络的要素	131
3.7 BPM 方法	80	4.5.2 人工神经网络应用	132
3.7.1 平衡记分卡	81	4.6 数据挖掘软件工具	134
3.7.2 六西格玛	83	4.7 关于数据挖掘的一些谎言和谬误	139
3.8 BPM 技术和应用	87	本章重点	139
3.8.1 BPM 架构	87	关键术语	140
		讨论题	141
		练习	141

本章结尾应用案例	144	6.2.3 BI 整合的水平	185
参考文献	145	6.2.4 嵌入式智能系统	185
第5章 文本挖掘与 Web 挖掘	147	6.3 BI 系统与数据库和其他企业系统的 连接	185
开篇场景: 文本挖掘与安全和反恐	147	6.3.1 与数据库连接	186
5.1 文本挖掘的概念和定义	149	6.3.2 BI 应用和后端系统的整合	186
5.2 自然语言处理	151	6.3.3 中间件	187
5.3 文本挖掘应用	155	6.4 面向需求的 BI	188
5.3.1 市场营销应用	155	6.4.1 传统 BI 的限制	188
5.3.2 安全应用	155	6.4.2 面向需求的选择	188
5.3.3 生物医学应用	157	6.4.3 关键特性和好处	188
5.3.4 学术应用	158	6.5 法律、隐私和道德问题	190
5.4 文本挖掘过程	159	6.5.1 法律问题	190
5.4.1 任务 1: 确定素材	160	6.5.2 隐私	190
5.4.2 任务 2: 创建文献术语矩阵	161	6.5.3 决策和支持中的道德问题	191
5.4.3 任务 3: 提取知识	162	6.6 BI 中的新兴话题: 概述	192
5.5 文本挖掘工具	166	6.7 Web 2.0 创新	193
5.5.1 商业软件工具	166	6.7.1 Web 2.0 的典型特征	193
5.5.2 免费软件工具	166	6.7.2 Web 2.0 公司和新的商业模型	194
5.6 Web 挖掘概述	167	6.8 在线社交网络: 基础和示例	194
5.7 Web 内容挖掘和 Web 结构挖掘	168	6.8.1 定义和基本信息	194
5.8 Web 使用挖掘	170	6.8.2 移动社交网络	195
5.9 Web 挖掘的成功实例	171	6.8.3 主要的社交网络服务: Facebook 和 Orkut	195
本章重点	174	6.8.4 商业和企业社交网络的意义	196
关键术语	175	6.9 虚拟世界	198
讨论题	175	6.10 社交网络和 BI: 协同决策	201
练习	175	6.10.1 协同决策的崛起	202
本章结尾应用案例	176	6.10.2 虚拟团队决策中的协同	202
参考文献	177	6.11 RFID 和新的 BI 应用机会	203
第6章 商务智能实施: 整合和新兴趋势 ...	179	6.12 现实挖掘	206
开篇场景: BI Eastern Mountain Sports 增加 合作和生产力	179	关键术语	208
6.1 BI 实施: 概述	181	讨论题	209
6.1.1 BI 实施因素	181	练习	209
6.1.2 BI 实施中的管理问题	182	本章结尾应用案例	211
6.2 BI 和整合实施	184	参考文献	211
6.2.1 整合的类型	184	术语	213
6.2.2 为什么进行整合	184		

商务智能简介

学习目标

- 理解今天动荡的商务环境，描述组织机构如何在这样的环境（解决问题和探索商机）中生存并且取得成就。
- 理解管理决策中对计算机化支持的需要。
- 描述商务智能的方法论和相关概念，并将它们与决策支持系统相联系。
- 理解商务智能实施中存在的主要问题。

当今的商务环境在不断地改变并且变得越来越复杂。组织、个人、公众都面临着巨大的压力，这些压力迫使他们要对变化的环境做出快速的反应，同时还要求他们在运作方法上有创新精神。这就需要组织机构灵活并且频繁快速地在战略层、战术层、操作层做出决策。有些决策是非常复杂的，做出这样的决策需要大量相关的数据、信息和知识。在需要决策的框架中，处理这些数据就需要企业能够做出非常迅速、实时的行动，这通常需要某些计算机化的支持。

本书讲述了如何将商务智能作为一种计算机化的支持应用到管理决策中。在重点讲述针对决策支持的商务智能的理论和概念基础的同时，也涉及有效的商务工具和技术。本章一方面详细介绍了这些内容，另一方面也对本书的内容进行了概述。

开篇场景：Norfolk Southern 利用商务智能进行决策支持获取成功

在美国有4个大型的铁路货运公司，Norfolk Southern（以下简称为NS）是其中之一。每天，公司在东部的22个州、哥伦比亚区、安大略、加拿大有大约500辆货运火车在运行，运行的总里程有21 000英里。公司有超过260亿的固定资产和超过30 000名员工。

在一个多世纪的时间里，铁路行业一直是一个受到严格管制的行业。NS及其前身主要是依靠管理自己的成本来盈利的。管理者将主要的精力放在了对现有轨道车辆的优化利用上，依靠公司的固定资产来获得更多的成果。在1980年，行业开始部分放松管制，这就为公司之间合并提供了机会。与此同时，公司可以基于自己的服务来收费并和顾客订立合同。准时送货成了影响这个行业的重要因素。

在一段时间里，NS公司适应业界变化的对策是变成了一个“预定铁路”。这就意味着公司必须要制定一套固定的火车运行时刻表，为行驶在火车与码头之间的汽车制定一套固定节点。在这种情况下，管理者能够预测什么时候他们可以将货物送达客户。

NS一直用多种复杂的系统来经营自己的业务。然而，变成一个“预定铁路”就需要一个新的系统首先可以应用统计模型来决定最好的路线和连接点以使火车运输的表现最优。然后这个系统还要应用模型来制定可以指导铁路运行的计划。这些新系统叫做TOP（Thoroughbred Operating Plan），TOP是在2002年开始部署的。

NS意识到仅用TOP系统来管理铁路的运行是不够的，公司还要监测和衡量TOP计划的表现。NS的众多系统产生了成千上万的关于货物的记录、轨道车的信息、火车全球定位系统（Global Positioning System, GPS）的信息、火车燃料的信息、收入的信息、机组人员管理和历史

跟踪记录信息。不幸的是，公司在开发利用这些信息的同时还要冒着对系统的运行产生重大影响的风险。

早在1995年，公司投资引进了一个1TB的Teradata数据仓库，这个数据仓库是关于历史数据的仓库^①。数据仓库是按照以下方式组织的：一方面数据是很容易得到的（使用网络浏览器），另一方面数据可以用来做决策支持。数据仓库的数据来自公司的运行系统（也就是原始系统），并且一旦这些数据从原始系统移到数据库中，用户就可以得到数据并且使用这些数据而不必冒着影响系统运行的风险。

在2002年，数据仓库变成了TOP系统的关键组成部分。NS建立了一个TOP仪表盘应用，可将数据从数据仓库中抽出，用图描绘出与运输计划不符的性能，包括火车性能和连接点性能。这个应用使用可视化技术使区域经理能够更轻松地了解如此大量的数据（如每周这里有160 000个连接点遍布整个网络）。自从此应用实施以来，消失的连接点的数量已经减少了近60%。并且在过去的5年中，轨道车的运转周期已经减少了一整天，这意味着节省了数百万的资金。

NS拥有一套企业数据仓库（Enterprise Data Warehouse, EDW），这就意味着一旦数据被放到数据仓库中，那么整个公司都可以得到数据，而不仅仅是对某个应用。虽然火车和连接点的性能数据是供TOP使用的，但公司可以将这些数据用于其他类型的应用。例如，市场部门开发了一个叫做AccessNS的应用程序，这个应用程序是为NS的那些想要进入NS广泛的运输网络的客户建立的。这些客户想要知道运送他们货物的船只现在在哪儿，有时客户还想要了解一些历史信息，如：我的货物是从哪里来的？需要多长时间到达？在运送的过程中遇到过什么问题？

AccessNS允许来自8 000多个客户组织的14 500用户随时访问系统，获得预先确定的关于他们账户的客户报告。用户可以得到时时更新的信息，也可以查看过去3年的数据。AccessNS拥有预警功能和真正简单聚合（Really Simple Syndication, RSS）的跟踪能力。事实上，每天有4 500个报告发布给用户。AccessNS提供自主服务的特性，使得NS能够给客户他们想要的信息，并且减少了从事客户服务的员工数量。事实上，如果没有AccessNS系统，要维持现在的客户报告水平公司至少需要47人。

公司的各个部门，从工程与战略规划部到成本与人力资源部都在使用EDW系统。公司内部的一个很有意思的应用程序是由人力资源部开发的。最近，为了很好地满足NS公司超过30 000名员工的需要，该部门需要确定区域办公室所在地。通过将员工的人口统计信息（如邮政编码）与原本用在工程部的地理数据整合后，人力资源部就能够清楚地勾画出员工的人口密度可视化地图，这样就使区域服务办公地点选取的优化工作变得很容易了。

现在，NS公司的数据仓库系统已经发展成一个6TB的系统。该系统可以管理公司巨大的铁路和海运服务网的海量信息。NS利用这套数据仓库系统分析趋势、制定预测时间表、存档记录并为顾客的自主服务提供便利。这套数据库系统为超过3 000名的员工和140 000个外部客户和利益相关者提供信息。

NS是第一家提供自助服务商务智能的铁路企业，它的创新使得其他的铁路企业纷纷效仿。公司还是第一家可以为外部客户提供大量历史数据的公司。

① Dashboard是一个苹果公司Mac OS X V10.4 Tiger操作系统中的应用程序，用做称为“widget”的小型应用程序的执行基础。——译者注

开篇场景的问题

1. 在 NS 公司, 信息系统是如何用来支持决策的?
2. 可视化应用可以获得哪种类型的信息?
3. AccessNS 可以提供哪种类型的信息支持?
4. NS 公司是如何将数据仓库应用到人力资源管理中的?
5. 同样的数据库是否可以应用到商务智能和优化的应用中?

从开篇场景中能够学到什么

这个开篇场景表明: 即使在一个很成熟的行业, 数据仓库技术仍然可以通过在公司的经营中获取更高的效率的方式使企业获得竞争的优势。确实, 在许多情况下, 这就是需要挖掘的前沿。从资产中获得更多的利润需要公司对其业务有及时详细的理解, 同时具有使用信息做出更好决策的能力。在本书中可以看到许多这样的应用案例。

可以在 Teradata 大学网络 (简称 TUN) 上获取更多关于案例的辅助资源, 将在后续的章节中进行详细的叙述。这些包括其他的论文和一篇题为《Norfolk Southern Uses Teradata Warehouse to Support a Scheduled Railroad》的播客。

来源: Contributed by Professors Barbara Wixom (University of Virginia), Hugh Watson (University of Georgia; 2005), and Jeff Hoffer (University of Dayton).

1.1 变化的商务环境和计算机化的决策支持

开篇场景说明了一个全球化的公司是如何在一个成熟而竞争激烈的市场中取得成就的。公司正在飞速地发展它们业务的计算机化支持。为了理解公司为什么如此地青睐计算机化支持, 包括商务智能, 我们建立一个商业压力-反应-支持模型来说明这个问题, 如图 1-1 所示。

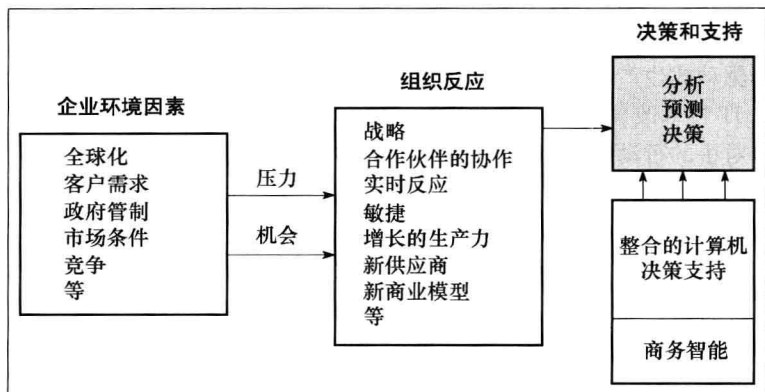


图 1-1 商业压力-反应-支持模型

商业压力-反应-支持模型

商业压力-反应-支持模型就如其名字一样包括 3 个部分: 源于今天商业环境的商业压力, 为克服压力公司采取的反应 (或者是对环境中有利机会的利用), 对环境监测提供方便并且能够提高组织反应能力的计算机化支持。

商业环境 组织今天面对的环境变得愈加复杂。这种复杂性一方面创造了机会, 另一方面也产生了问题。以全球化为例, 今天, 你可以轻松地世界上的许多国家找到供应商和顾客, 这就意味着你可以买到更加便宜的原料, 并将产品和服务卖给更多的顾客, 存在很多机会。然而全球化意味着更多和更激烈的竞争。商业环境因素可以分成主要的 4 种类型: 市场因素、顾客需求