

Broadview
WWW.BROADVIEW.COM.CN

数据仓库与数据挖掘
技术应用丛书

Corporate Information Factory,
Second Edition

企业信息工厂

[美] W.H.Inmon Claudia Imhoff Ryan Sousa 著
邢国庆 等译

 **WILEY**



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

<http://www.phei.com.cn>

数据仓库与数据挖掘技术应用丛书

企业信息工厂

Corporate Information Factory, Second Edition

W.H.Inmon

[美] Claudia Imhoff 著

Ryan Sousa

邢国庆 等译

电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

本书是由誉为“数据仓库之父”的 W. H. Inmon 撰写的。本书高屋建瓴，全面介绍了企业信息工厂的总体架构、组件及其相互关系，其中包括数据仓库、数据挖掘、数据集市和操作型数据存储等技术。本书简明地概述了各个组件及其相互之间的关系，说明了如何创建企业信息工厂，为企业提供决策支持，以及介绍了怎样把现有的系统改造成企业信息工厂的方法和步骤。

本书是一本数据仓库技术的入门教科书，也可作为数据仓库管理员和 IT 主管的参考书，使读者能够对数据仓库、决策支持系统以及信息系统架构等有一个全面的了解和整体把握。

All rights reserved. Authorized translation from the English language edition published by John Wiley & Sons, Inc.

本书简体中文专有翻译出版权由 John Wiley & Sons Inc. 授予电子工业出版社，未经许可，不得以任何方式复制或抄袭本书的任何部分。

版权贸易合同登记号：图字：01-2002-5366

图书在版编目 (CIP) 数据

企业信息工厂 / (美) 英蒙 (Inmon, W.H.) 等著; 邢国庆等译. —北京: 电子工业出版社, 2004.1

数据仓库与数据挖掘技术应用丛书

书名原文: Corporate Information Factory, Second Edition

ISBN 7-5053-9100-3

I. 企... II. ①英...②邢... III. 企业管理—数据库管理系统 IV. F270.7

中国版本图书馆 CIP 数据核字 (2003) 第 075193 号

责任编辑: 黄志瑜

印 刷: 北京增富印刷有限公司

出版发行: 电子工业出版社 <http://www.phei.com.cn>

北京市海淀区万寿路 173 信箱 邮编 100036

经 销: 各地新华书店

开 本: 787×980 1/16 印张: 24.75 字数: 456 千字

版 次: 2004 年 1 月第 1 版 2004 年 1 月第 1 次印刷

印 数: 4000 册 定价: 48.00 元

凡购买电子工业出版社的图书, 如有缺损问题, 请向购买书店调换。若书店售缺, 请与本社发行部联系。联系电话: (010) 68279077。质量投诉请发邮件至 zlts@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

出版说明

如果没有对海量数据进行科学分析的能力，沃尔玛的老板再精明，也绝对想不到“啤酒与尿布”这两个风马牛不相及的东西之间还有着千丝万缕的联系。而将它们放在一起，竟然增加了啤酒销量，可见数据分析的巨大威力。

信息系统数年中收集了海量数据，且数据还正以指数级增长，企业迫切地需要高效、精确、科学地分析数据，以找出其背后的寓意，进而了解企业的经营状况和外部环境，做出科学的决断，在现代激烈的竞争中胜出。所以，如何将数据点石成金，更是摆在我们面前很现实也很诱人的一个问题。

现在，很多人已经意识到数据中潜在的大量商机，并踏踏实实地进行着从数据中沙里淘金的工作。特别是在信息化的大潮中，上至政府，下到企业，从银行到电信，再到网站、超市，人们都希望用数据分析这根魔杖赢得先机。与此同时，人们也在期盼着相关书籍，以便工作中学习参考。在广泛征询专家和用户的基础上，秉着选题全面、内容经典、译者严谨的原则，我们适时地推出了这套《数据仓库与数据挖掘技术应用丛书》，以飨读者。本丛书有如下几本：

- 数据仓库基础
- OLAP 解决方案：多维信息系统的构建技术
- 数据仓库工具箱：维度建模的完全指南（第二版）
- 数据仓库生命周期工具箱：设计、开发和部署数据仓库的专家方法
- 数据仓库及其在电信领域中的应用
- 疑难数据仓库专家解决方案
- IBM 数据仓库和商业智能工具
- 可视化数据挖掘：数据可视化和挖掘的技术和工具
- 点击流数据仓库
- Web 数据挖掘：将客户数据转化为客户价值
- 企业信息工厂
- 机器学习与数据挖掘：方法和应用

本丛书既包括商业智能（BI）的基础——数据仓库（DW），也包括数据仓库上的两类不同目的的数据增值操作——联机分析处理（OLAP）和数据挖掘（DM）；既覆盖基础理论，如数据仓库基础，又提供不同领域的解决方案，如数据仓库在电信、银行、保险等领域的应用。

本丛书来自国外数据库领域一些著名作者的畅销书，以及国内第一线实施者的精心总结。如一直位居 AMAZON 畅销书榜的数据仓库领域的畅销书作家 Ralph Kimball 的《数据仓库工具箱：维度建模的完全指南（第二版）》、《数据

仓库生命周期工具箱：设计、开发和部署数据仓库的专家方法》，数据仓库之父 William H.Inmon 的《企业信息工厂》(Corporate Information Factory) 等。

丛书的译者均来自工作在该领域一线的人员，既有该领域的理论和实践经验，又具备中英文翻译的功底。且多位译者先前均已读过原著，所以，自感翻译的过程不再是枯燥，而是情趣盎然，乐在其中。

出版高品位、高品质的图书是博文视点的努力目标。希望您对我们的工作多提宝贵意见。您的意见是我们创造精品的动力源泉。

如果您希望将您的工作经验感悟等总结成书，我们将为您提供一流的服务，共创精品图书。

我们的联系方式如下：

地址：北京复兴路 47 号天行建商务大厦 604

邮编：100036

电话：010-51922832，68216158

传真：010-51922823

E-mail: jsj@phei.com.cn; zsh@phei.com.cn

博文视点资讯有限公司

2003 年 10 月

博文视点资讯有限公司 (BROADVIEW Information Co.,Ltd.) 于 2003 年 6 月 18 日正式成立,是信息产业部直属的中央一级科技与教育出版社——电子工业出版社 (PHEI) 与国内最大的 IT 技术网站 CSDN.NET 和最具专业水准的 IT 杂志社《程序员》合资成立的以 IT 图书出版为主业、开展相关信息和知识增值服务的资讯公司。

我们的理念是:创新专业出版体制;培养职业出版队伍;打造精品出版品牌;完善全面出版服务。

秉承博文视点的理念,博文视点的产品线为面向 IT 专业人员的出版物和相关服务。博文视点将重点做好以下工作:

- (1) 在技术领域开发专业作(译)者群体和高质量的原创图书
- (2) 在图书领域建立专业的选题策划和审读机制
- (3) 在市场领域开创有效的宣传手段和营销渠道

博文视点有效地综合了电子工业出版社、《程序员》杂志社和 CSDN.NET 的资源 and 人才,建立全新专业的立体出版机制,确立独特的出版特色和优势,将打造 IT 出版领域的著名品牌,并力争成为中国最具影响力的专业 IT 出版和服务提供商。

作为合资公司,博文视点的团队融合了各方面的精英力量:原电子工业出版社 IT 图书专业出版实力的代表部门——计算机图书事业部的团队;《程序员》杂志和 CSDN 网站的主创人员;著名 IT 专业图书策划人周筠女士及其创作群。这是一个整合专业技术人员和专业出版人员的团队;这是一个充满创新意识和创作激情的团队;这是一个不断进取,追求卓越的团队。

原 PHEI 计算机图书事业部以及武汉的创作团队,在合资之前就已经各自具备了很强的策划和出版实力,并创造出了优秀的出版业绩。而《程序员》杂志则是国内最有影响力的 IT 专业杂志,《程序员》杂志的年发行量超过 80 万份。CSDN 网站是国内最大的 IT 技术网站,到 2003 年 5 月底,CSDN 网站的注册会员已超过了 54 万,这两者已成为中国 IT 技术交流与推广的最佳平台,积累了丰富的作者和读者资源。

电子工业出版社与《程序员》杂志社和 CSDN 网站的合作以最有效率的方式形成了出版资源、媒体资源、网络资源的整合和互动,成为 2003 年 IT 出版界倍受瞩目的事件。

“技术凝聚实力,专业创新出版”,BROADVIEW 与您携手共迎信息时代的机遇与挑战!

译者序

《企业信息工厂 (Corporate Information Factory)》一书是由被誉为“数据仓库之父”的 W. H. Inmon 先生继《建立数据仓库》、《使用数据仓库》以及《面向对象的数据仓库设计》之后撰写的又一部力作。至今,他已经写了40多本有关数据库、数据库管理以及数据仓库方面的书,并经常在高层次的数据库会议上做主题报告。W. H. Inmon 也是 www.billinmon.com 网站的合作伙伴,该网站专门讨论信息工厂、数据仓库及其有关技术,值得一览。

本书高屋建瓴,全面介绍了企业信息工厂的总体架构、各个构件及其相互关系,其中包括数据仓库、数据挖掘、数据集市和操作型数据存储等技术,精辟说明了如何创建企业信息工厂,以及怎样把现有的系统改造和集成为企业信息工厂的方法和步骤。

企业信息工厂给出了一种信息系统的逻辑架构,其目的是利用运营业务数据,提供智能化的业务管理和决策支持能力。本书提供了一种组织和管理信息系统的逻辑架构和方法,用于把数据仓库、数据挖掘、数据集市、操作型数据存储以及 Internet 等技术集成为一体,以满足企业的实际业务决策需要。本书提出了现今和将来创建整体信息系统解决方案的发展方向。现有的实践证明,企业信息工厂可以适应任何规模的企业创建数据仓库和决策支持系统。

当今,企业的 IT 主管和数据仓库管理人员需要组建一个小型的专门技术队伍,确保能够快速地为企业高层管理人员的业务决策提供准确的信息。W.H.Inmon 的《企业信息工厂》满足了这些人的需要。

本书对数据仓库及有关的周边技术做了全面的介绍,有益于对此领域有一个全面的了解和整体把握,可以为进一步地深入研究、开发和应用打下良好的基础。

本书由邢国庆翻译。在翻译过程中,得到了黄志瑜、Joe Simpson、邸静、邢梦可、成效禹、伊晓强、魏承明、贾磊、仇鹏涛以及张秋慧等同志的大力协助。尤其是黄志瑜先生,在本书的整体把握等方面提出了极为宝贵的指导意见,本书的名字也是根据黄志瑜先生的建议确定的。Joe 在细节的推敲方面给予了极大的帮助。邸静女士承担了本书的全部文字录入工作。在此一并表示感谢。

限于译者的水平和能力,错误和不当之处在所难免,恳请广大读者给予批评指正。

译者

2003 年 3 月

前言

有的书主要介绍“怎样”做事情，这一类书籍可以通称为操作指南或手册；有些书主要解释“为什么”要这样做，这通常是讲述基本原理一类的书籍；另外一些书描述的是所需信息在“什么地方”，这是如地图册之类的图书；还有一些书只描述应当做“什么”，本书就是这样一本书。

1. 数据仓库的起源

计算机和信息技术行业正以惊人的速度发展着。在打孔卡片和纸带输入系统的时代，我们只是把计算机作为一种不堪重负的笨重计算设备来使用，这种系统通常只适用于工资和费用的统计、收发等业务。

随着大容量磁盘存储器、廉价的内存、性能优异的操作系统以及直接的最终用户接口设备的出现，一种称做联机处理的全新计算模式成为现实。采用联机处理方式，计算机的应用日新月异，订票系统、自动柜员机以及其他新的应用系统开始日益普及。

接踵而来的是最终用户端的革命性变化。个人计算机、电子表应用，以及第四代语言（4GL）技术开启了先前曾被关闭的计算领域的大门。当计算技术的费用直线下降时，潘多拉魔盒打开了。计算机应用现在已延伸到传统的信息系统领域之外。任何人，只要有预算经费就能够开始掌握自己的信息处理的命运。

最终用户非常愿意完全由自己控制一切，但这种自主性只能是一种海市蜃楼。而且很快就会发现，即使采用最先进的计算技术，如果把整个控制都交给最终用户，自然还会带来其他新的问题，如缺乏集成和有效性，很难形成一定的处理规模等。

同时，人们还发现很难访问交易处理过程中产生的数据，因而无法有效地做出决策，因为不管决策处理用到的是汇总数据还是详细数据，都需要涉及历史的、集成的数据。

数据仓库于是应运而生。但此后不久，随着数据量和最终用户需求的不断增长，超出了数据仓库能够提供的数据定制与调整能力。由于它的全部设计重心在于集成与管理通用的企业数据，因此，数据仓库无法满足一些特定的业务信息需求。为了应对这种情况，不同的部门发现，定制的数据仓库子集，也就是数据集市能够提供它们需要的自主性以及对企业信息的解释和使用。

企业也发现，他们还需要集成的操作型数据，而数据仓库却不能为需要

操作型数据集成的人提供任何帮助。经过争论之后，出现了操作型数据存储。最后，随着 Internet 和低廉的硬件技术的出现，利用数据仓库、数据集市以及操作型数据存储，可以使用更多的数据，支持广泛的用户群体，因而可以逐步解决这些越来越复杂的问题。就像处理一个简单的问题一样，让用户（包括业务伙伴甚至客户）访问来自若干数据源的数亿字节的集成数据已经逐渐成为很现实的事情了。

2. 由此及彼，展望未来

我们当中的某些人可能已经经历或见证过一些开发过程中遇到的问题。第一个问题就是我们对技术的重要性，太过于专注技术的本身和技术的成熟性。我们只关注其细节而忽略了大的方面。因此，就像盲人描述大象一样，只能理解自己最熟悉的那一小部分。

第二个问题是只注重一天天的日常开发事项和开发进度。一叶障目，不见泰山，最终只能猜测明天将会出现什么。

本书的目的就是为了克服上述障碍，从头开始，仔细地考察整个业务信息系统的发展过程，重点介绍一种已经过实践证明的架构及其有关的技术。另外，我们还将介绍这个架构中的关键构件，说明怎样发挥它们各自的作用，并使之相互协调、有序地运行。这个架构就是企业信息工厂（Corporate Information Factory, CIF）。

我们特别感兴趣的是，当企业试图以企业信息工厂架构之外的其他方法建立它们的信息系统时会发生什么情况。企业信息工厂的简单发展历史已经表明：

- 肯定会建立一个不同于本书所描述的信息系统结构；
- 当建立的信息系统结构不同于本书描述的企业信息工厂架构时，可能会在以下几个方面付出代价：
 - ✓ 基础设施费用
 - ✓ 有效的性能
 - ✓ 集成的缺乏
 - ✓ 无缝的技术连接
 - ✓ 最终用户的满意度
 - ✓ 应对变化

企业信息工厂并不是建立信息系统的惟一方法，但它是最好的方法，能够满足企业信息处理的长期目标。

在许多情况下，某些设计人员不一定愿意遵守这个架构推荐的建设原则。但设计人员必须记住，这是要付出代价的。

写作本书的目的就是希望为读者指出一个已经过实践证明的组织信息系统的方法。如果不采用这个方法，将会产生许多意料之中或者根本无法预料的不良后果。这些不良后果的影响，从浪费大量的开发资源，到无法提供有效的信息资源，其范围是很广的。必须对此有足够的思想准备和应对措施。

大量的因素促进了企业信息工厂的发展，其中包括：

- 不断增长的业务需求；
- 硬件价格不断下降；
- 技术不断地翻新和成熟；
- 用户需求的不断增长；
- 硬件、软件和网络功能的提高。

此外，许多企业已经逐渐开始转向完全不同风格的处理方式。例如，从操作型的传统处理逐步转变为数据仓库，之后又引出了数据集市；电子报表等新的技术开启了许多分析功能。所有这一切导致了新的企业处理模式的出现，而这也是我们将要在本书中介绍的内容。

3. 为什么要写第二版

本书的第一版介绍了企业信息系统以及有关信息技术的发展。在写作第一版时，书中的描述是准确的，而且也最好地反映了这一领域当时的技术成果。但自该书出版之后，我们开始注意到，还有一些新的现代信息系统架构的重要构件没有及时补充进来。

探索数据仓库、备用存储和决策支持系统（DSS）应用显然是信息系统架构的重要构件。此外，许多大型的国际企业已开始建立多个数据仓库，这自然又引申出一个大型企业可以拥有多个企业信息工厂的新课题。而且随着IT技术的不断发展，Internet、客户关系管理（CRM）以及企业资源规划（ERP）也成为重要的一部分。所有这一切改变了信息系统的远景规划，因此，必须写第二版，把它们纳入到企业信息工厂架构中。

《企业信息工厂》第二版是根据《Data Management Review》杂志 2000 年 1 月这一期中的大型插图描述的架构编写的（如图 2-1 所示）。图中的介绍是当时对企业信息组织架构的完整理解和总结。如果读者想得到那个插图的拷贝，可以联系 DMR 杂志，或从 www.billinmon.com 网站中下载（免费）。在此，我们非常感谢 Ron Powell 和 Jean Schauer 对创建该插图的赞助，以及 IBI 公司对创建该插图提供的资金赞助。

4. 本书的读者对象

《企业信息工厂》一书可以适用于各种读者对象，其中包括：

IT 管理人员。IT 管理人员可以使用企业信息工厂去预测系统开发和系

统架构设计时下一步应当怎样进行。避免在项目建设过程中违背企业信息工厂推荐的设计准则，浪费不必要的人力和物力资源。在了解了企业信息工厂的基本概念和有关技术之后，管理人员可以据此把握未来的发展方向。

开发人员。在项目开始之后，开发人员可以确定项目的组织与企业信息工厂的设计准则是否协调。如果整个设计违背了企业信息工厂的设计准则，可以在成为事实之前，建议设计人员进行矫正。

投资商。企业信息工厂给出了确定技术投资能够取得什么回报的一种简单的方法。如果正在投资的系统架构不符合企业信息工厂的设计准则，投资商可以依据市场的认可情况提出修改意见，并采取相应的对策。

最终用户。企业信息工厂成功与否的关键是最终用户是否能够使它达到自己的愿望。最终用户可以评估系统是否偏离了自己的期望值。如果成功地实现了企业信息工厂，将会使最终用户的业务平稳运作并提高生产效率。

5. 本书的编排

本书是为适应从初学者到有经验的企业信息工厂专家等各种类型的读者的需求而进行编排的。对于企业信息工厂的初学者，建议读者从头到尾逐章阅读本书，因为后续每一章的内容都建立在前一章的基础之上，逐次解释了什么是企业信息工厂，以及怎样建立和管理企业信息工厂。对于有经验的读者而言，也许需要阅读头两章，然后深入研究适合自己需要的章节。他们也可能特别感兴趣第9章、第10章、第13章和第17章，这几章介绍了某些新的构件和概念。

本书共分为4部分。第一部分包括作为入门性的第1章和第2章，其中概述了企业信息工厂的基本概念和由来，使读者能够全面了解企业信息工厂的发展过程。第二部分包括第3章至第14章，其中分别介绍了企业信息工厂架构的各个构件，以及怎样把它们组织在一起，以提供决策支持能力。同时也讨论了这个架构的某些变体。第三部分包括第15章至第17章，其中讨论了怎样建立和管理企业信息工厂。第四部分即附录A，其中给出了评估和考察企业信息工厂的准则。

6. 本书的写作背景

我不知道如何表达我的家人对我许多天顾不上回家或长时间在家写作本书的第一版时所给予的理解和支持。当我询问是否能够再改写第二版时，他们又给予了肯定的回答。他们不仅同意，而且还鼓励我深入研究，写出更好的企业信息工厂。

我希望借此机会向多年来帮助充实和丰富了企业信息工厂理念的许多同行、客户和朋友们表示衷心的感谢。正是由于他们集体的努力和智慧，使得

本书得以问世。对此，我要特别感谢以下人士和机构：

John Zachman, Zachman International
John Bair, Independent Consultant
Lowell Fryman, C/Net
Roger Geiwitz, Independent Consultant
Sue Osterfelt, Bank of America
JD Welch, IBM
Dennis McCann, Ambeo
Ken Richardson, Ambeo
Dale Brocklehurst, Ambeo
Joyce Norris-Montanari, Braun Consulting
Jon Geiger, Braun Consulting
Steve Miller, Braun Consulting
Jim Kalustian, Braun Consulting
Mike Evanisko, Braun Consulting
Dave Imhoff, Intelligent Solutions
Rob Geller, Quest
Robert Frim, Independent Consultant
Pete Simcox, Genesis
Mark Mays, Arrowhead Consulting
WarehouseMCI team
John Ladley, Knowledge Interspace
Doug Laney, Meta Group
Bob Lokken, Knosys
Brian Burnett, AppsCo
Steve Murchie, Microsoft
Allen Perry, Coglin Mill
Ron Powell, DM Review
Bill Prentice, SAS
Mike Wipperfeld, Informix
Lisa Loftis, Braun Consulting
Steve Hill, Informix
Keven Gould, Sybase
Stephen Gardner, NCR

Ron Swift, NCR

Marc Demarest, Independent Consultant

Jeanne Friedman, Independent Consultant

Greg Battas, Tandem Computers

Ralph Kimball, Kimball and Associates

博文视点公司向您郑重推荐——

数据仓库与数据挖掘技术应用丛书



数据仓库工具箱：维度建模的完全指南（第二版）

本书从维度建模的基本知识入手，通过给出零售、库存、采购、订单管理、客户关系管理、账目、人力资源管理、财经服务、电信与公用事业、交通、教育、卫生护理、电子商务以及保险等方面的实例研究，全面、深入、透彻与具体地介绍了利用维度模型设计与开发数据仓库的综合技术知识。

本书主要是为数据仓库设计人员、技术人员以及管理人员而编写的。此外，对于积极参与数据仓库初期规划工作的业务分析人员来说，本书也有很高的学习价值。



数据仓库生命周期工具箱：

设计、开发和部署数据仓库的专家方法

本书介绍了使数据仓库项目获得成功所必备的各种知识和经验教训，它们都是作者自 1982 年以来在从事数以百计的数据仓库安装和咨询任务过程中不断积累总结出来的。书末的两个附录中提供了大量的框架、任务、模板以及生动详实的样例（具体内容见本书配套光盘），所有这些都使本书别具一格。全书主题广泛，思想深刻，内容详尽，图文并茂。

本书不仅是现代信息系统开发人员的重要指南，而且是所有面向数据仓库项目的设计、开发、管理、咨询人员的高级参谋，并且适合信息管理与信息系统、计算机应用、电子商务等专业的高校师生作为教学参考用书，还可供从事传统数据库系统工作的技术人员参考阅读。



数据仓库及其在电信领域中的应用

本书首先介绍了数据仓库及其一些基本概念。从第 4 章开始，概要地分析了数据仓库在电信领域的应用情况。在第 5 章中列举了一些具体的应用案例，介绍了数据仓库技术在电信领域中的应用。第 6 章和第 7 章概要地介绍了一些数据仓库产品和进行数据仓库产品测试需要考虑的内容。

本书适合在电信领域从事数据仓库的读者学习参考，也可作为数据仓库爱好者的参考用书。

疑难数据仓库专家解决方案



本书介绍了各种类型、规模和结构的公司都会面临的 91 种常见疑难及其解决方案。9 位数据仓库建设方面的权威专家针对这些看似难以应付的危机各抒己见，并给企业管理人员、IT 经理与职员以及终端用户提供了防止和拯救这些问题的具体措施。

本书既可以作为解决具体数据仓库问题的简明手册，也可以作为了解数据仓库现实问题的入门性读物。

可视化数据挖掘：数据可视化和挖掘的技术和工具



本书描述了可视化数据挖掘技术，以及可视化数据挖掘技术能够解决的商业问题的种类。在介绍完业务问题和基本原理后，以一个完整的实例逐步讲解如何利用可视化数据挖掘技术实施商业智能项目的方法。利用可视化数据挖掘工具和技术，分析人员能够从全新的角度快速、轻松地检索信息，解决常见的商业问题。可视化数据挖掘使数据挖掘变得简单，非技术出身的业务经理们利用它能够更好地理解市场，并做出明智的决策。

另外，本书还介绍了可视化工具方面的知识，拓宽了读者的范围。

本书适合于数据可视化和可视化数据挖掘商业智能解决方案实施单位的各层次人员，包括：数据分析员、业务分析员、领域专家和决策人员。

点击流数据仓库



本书系统地介绍了点击流数据仓库的所有问题。本书的特点是内容丰富，概念清楚、准确，实例丰富，讲解详细。首先，本书内容结构的逻辑性很强，它的每一章都建立在前一章的基础上，引导读者在头脑中逐步形成一个点击流数据仓库的框架。然后，结合实例详细说明创建点击流数据仓库的过程及其所遇到的众多技术难题。

本书论述了点击流数据仓库的基础知识以及构建点击流数据仓库的全过程。本书适用于学习或者在工作中运用点击流数据仓库技术的教师、学生或工程技术人员，特别适用于对数据仓库技术有所了解但希望进一步提高构建点击流数据仓库能力的应用开发人员。



Web 数据挖掘：将客户数据转化为客户价值

Web 正在改变着整个商业世界!商业世界的不断改变影响着数据挖掘技术,数据挖掘也在不断地改变整个商业世界。本书全面地展示了 Web 对于数据挖掘在商业方面的影响,包含了大量 Web 数据挖掘的实例,并将数据挖掘放入了一个学习环境中,描述了一些对于任何关注客户的企业都非常重要的概念,以及面向 Web 的与商业相关的分析类型。

本书适合于 Web 数据挖掘, Web 分析以及其他相关领域的技术专业人员,经营管理人员阅读,也可以作为大专院校相关课程的重要辅导教材。



企业信息工厂

本书是由誉为“数据仓库之父”的W.H.Inmon撰写的。本书高屋建瓴地全面介绍了企业信息工厂的总体架构、组件及其相互关系,其中包括数据仓库、数据挖掘、数据集市和操作型数据存储等技术。本书简明地概述了各个组件及其相互之间的关系,说明了如何创建企业信息工厂,为企业提供决策支持,以及介绍了怎样把现有的系统改造成企业信息工厂的方法和步骤。

本书是一种数据仓库技术的入门教科书,也可作为数据仓库管理员和 IT 主管的参考书,使读者能够对数据仓库、决策支持系统以及信息系统架构等有一个全面的了解和整体的把握。



机器学习与数据挖掘：方法和应用

本书分为 5 个部分,共 18 章,较为全面地介绍了机器学习的基本概念,并讨论了数据挖掘和知识发现中有关问题以及多策略学习方法,具体地阐述了机器学习与数据挖掘在工程设计,文字、图像和音乐,网页分析、计算机病毒和计算机控制,医疗诊断、生物医疗信号分析和水质分析中生物信号处理等方面的应用情况。

本书收集众多不同领域中数据挖掘的实际案例,以此来说明数据挖掘的具体解决方法,以期为广大读者提供一个更为广阔的数据挖掘应用视角。

本书的读者可以是任何对机器学习与数据挖掘感兴趣的工程技术、业务管理人员,或其他从事具体技术工作的人员。本书也可作为大专院校相关课程的重要辅导教材。

后期即将推出:

数据仓库基础

OLAP 解决方案:多维信息系统的构建技术

IBM 数据仓库和商业智能工具

ERP 原理·设计·实施（第二版）



企业要提高竞争力，企业的信息化建设是必由之路，企业资源计划（Enterprise Resource Planning，简称 ERP）为企业的信息化建设提供了全面集成的解决方案。本书力求从务实的角度，揭开 ERP 那层神秘的面纱，让企业对 ERP 形成一个正确的认识，推动我国 ERP 实施应用的健康发展。

为了让读者从各种角度对 ERP 进行全面的、系统的了解，并满足各种层次人员的需求，本书对 ERP 相关知识的讨论涵盖了原理、设计与应用的全部过程，同时收集了一些 ERP 实施的案例为读者提供参考。本书适用于企业的各级领导、项目经理、系统分析员及高级程序员，也可作为 ERP、MBA 的培训教材或相关专业研究生、大学生的教材。

第 1 版出版后引起了很大的社会反响，本书是根据读者来信修订的第 2 版，不仅案例和内容适时做了很大调整，还针对读者来信选出有代表性的问题在书后做了详细解答，并且更新了封面。本书第 2 版将更加贴近读者需要，为 ERP 领域的研究和实施者提供更切实的指导。

CRM 原理·设计·实践



这是一部全面系统地论述客户关系管理（CRM）的技术管理书籍。作者用提出问题、解决问题的写作手法，在准确定义 CRM 概念的基础上，对 CRM 应用系统的设计思想、设计平台、CRM 实践的最佳方法、CRM 软件项目的确立和实施等进行了多方位、详细的论述。

本书的 CRM 三角概念，客户关系计算机模型，CRM 业务流程应用设计以及以“下一个最佳活动 NBA”为核心的 CRM 实践方法论都是本书作者多年的研究与实践心得，可使读者全面、透彻地理解 CRM 的核心理念，并有效提高企业的 CRM 实践能力。

本书的读者对象是 CRM 项目设计、实施人员，广大 CRM 研究与爱好者，以及企业领导、管理人员、营销经理等。也可作为高校计算机专业、信息管理专业研究生和本科生教材。