21st CENTURY
实用规划教材

21世纪全国应用型本科电子商务与信息管理
系列实用规划教材

数据仓库与数据挖掘

主　编　廖开际
副主编　刘凤英　胡建军

北京大学出版社
PEKING UNIVERSITY PRESS

21世纪全国应用型本科电子商务与信息管理系列实用规划教材

数据仓库与数据挖掘

主　编　廖开际
副主编　刘凤英　胡建军
参　编　邝云英　吴　君　杨　进

北京大学出版社
PEKING UNIVERSITY PRESS

内 容 简 介

本书比较系统地介绍数据仓库与数据挖掘的理论体系和应用。本书总的指导思想是在掌握基本知识和基本理论的基础上，强调实际应用能力的培养。全书力求深入浅出，通过通俗的语言及案例分析，介绍数据仓库及数据挖掘的基本概念及相关理论与方法。从数据仓库的定义、结构、设计、构建方法及联机分析处理应用等方面对数据仓库进行较为详细的介绍；从数据挖掘的定义、数据预处理、数据挖掘中的常用算法等方面对数据挖掘的基本知识和算法等理论进行介绍。本书强调数据仓库和数据挖掘工具的应用，重点介绍 SQL Server 2005 数据仓库和数据挖掘工具的应用。附录 A 详细介绍一个简易的数据挖掘工具——Weka，该工具可作为读者学习数据挖掘时的实验工具。

本书可作为普通高等学校电子商务、信息管理、计算机应用及其他相关专业的本科教材，也可作为经贸、管理类专业的研究生教材，以及各类管理人员的培训与自学用书。

图书在版编目(CIP)数据

数据仓库与数据挖掘/廖开际主编. —北京：北京大学出版社，2008.11
(21 世纪全国应用型本科电子商务与信息管理系列实用规划教材)
ISBN 978-7-301-14313-1

Ⅰ. 数… Ⅱ. 廖… Ⅲ. ①数据库系统—高等学校—教材②数据采集—高等学校—教材
Ⅳ. TP311.13 TP274

中国版本图书馆 CIP 数据核字(2008)第 152773 号

书　　名：数据仓库与数据挖掘
著作责任者：廖开际　主编
责 任 编 辑：刘　丽
标 准 书 号：ISBN 978-7-301-14313-1/TP・0972
出　版　者：北京大学出版社
地　　址：北京市海淀区成府路 205 号　100871
网　　址：http://www.pup.cn　http://www.pup6.com
电　　话：邮购部 62752015　发行部 62750672　编辑部 62750667　出版部 62754962
电 子 邮 箱：pup_6@163.com
印　刷　者：河北滦县鑫华书刊印刷厂
发　行　者：北京大学出版社
经　销　者：新华书店
787 毫米×1092 毫米　16 开本　16.75 印张　380 千字
2008 年 11 月第 1 版　2008 年 11 月第 1 次印刷
定　　价：28.00 元

丛 书 序

随着电子商务与信息管理技术及应用在我国和全球的迅速发展，政府、行业和企业对电子商务与信息管理的重视程度不断提高，我国高校电子商务与信息管理人才培养的任务也不断加重。作为一个新兴的跨学科领域的专业，电子商务与信息管理的教育在快速发展的同时还存在着许多值得我们思考和改进的问题。特别是开办电子商务专业和信息管理专业的学校学科背景不同，有文科的、理工科的、经管类学科等，使得不同学校对核心课程的设置差异很大；另外，近年来有关电子商务与信息管理方面的教材出版的数量虽然不少，但适合于财经管理类知识背景本科生的电子商务系列与信息管理系列教材一直缺乏，而在开办电子商务和信息管理本科专业的高校中，财经管理类的高校占的比重很大。为此北京大学出版社于 2006 年 11 月在北京召开了《21 世纪全国应用型本科财经管理系列实用规划教材》研讨会暨组稿会，会上出版社的领导和编辑通过对国内经管类学科背景的多所大学电子商务与信息管理系列教材实际情况的调研，在与众多专家学者讨论的基础上，决定成立电子商务与信息管理系列丛书专家编审委员会，组织编写和出版一套面向经管类学科背景的电子商务与信息管理专业的应用型系列教材，暨《21 世纪全国应用型本科电子商务与信息管理系列实用规划教材》。

本系列教材的特点在于，按照高等学校电子商务专业与信息管理专业对本科教学的基本要求，参考教育部高等学校电子商务专业与信息管理专业的课程体系和知识体系，定位于实用型人才培养。

本系列教材还体现了教育思想和教育观念的转变，依据教学内容、教学方法和教学手段的现状和趋势进行了精心策划，系统、全面地研究普通高校教学改革、教材建设的需求，优先开发其中教学急需、改革方案明确、适用范围较广的教材。此次教材建设的内容、架构重点考虑了以下几个要素。

(1) 关注电子商务与信息管理发展的大背景，拓宽经济管理理论基础、强调计算机应用与网络技术应用技能和专业知识，着眼于增强教学内容的联系实际和应用性，突出创造能力和创新意识。

(2) 尽可能符合学校、学科的课程设置要求。以高等教育的培养目标为依据，注重教材的科学性、实用性和通用性，尽量满足同类专业院校的需求。

(3) 集中了在电子商务专业与信息管理专业教学方面具有丰富经验的许多教师和研究人员的宝贵意见，准确定位教材在人才培养过程中的地位和作用。面向就业，突出应用。

(4) 进行了合理选材和编排。教材内容很好地处理了传统内容与现代内容的关系，补充了大量新知识、新技术和新成果。根据教学内容、学时、教学大纲的要求，突出了重点和难点。

(5) 创新写作方法，侧重案例教学。本套教材收集了大量新的典型案例，并且用通俗易懂的方式将这些案例中所包含的电子商务与信息管理的战略问题传授给读者。

前任联合国秘书长安南在联合国2003年电子商务报告中说："人类所表现出的创造力，几乎都没有像互联网及其他信息和通信技术在过去十年中的兴起那样，能够如此广泛和迅速地改变社会。尽管这些变革非常显著，然而消化和学习的过程却只是刚刚开始。"可以说没有一个学科像电子商务与信息管理这样如此完美地融技术与管理于一体，也没有哪一个人的知识能如此的全面丰富。参与本系列教材编写的人员涉及国内几十所高校的几十位老师，他们均是近年来从事电子商务与信息管理教学一线的高校教师，并均在此领域取得了丰富的教学和科研成果。所以本系列教材是集体智慧的结晶，它集所有参与编写的教师之长为培养电子商务与信息管理人才铺垫基础。

在本系列教材即将出版之际，我要感谢参加本系列教材编写和审稿的各位老师所付出的辛勤劳动。由于时间紧，相互协调难度大等原因，尽管本系列教材即将面世，但一定存在着很多的不足。我们希望本套系列教材能为开办电子商务和信息管理专业的学校师生提供尽可能好的教学用书，我们也希望能得到各位用书老师的宝贵意见，以便使编者们与时俱进，使教材得到不断的改进和完善。

2007年11月于大连

李洪心 李洪心博士现任东北财经大学教授，教育部高等学校电子商务专业教学指导委员会委员，劳动和社会保障部国家职业技能鉴定专家委员会电子商务专业委员会委员，中国信息经济学会电子商务专业委员会副主任委员。

前　言

随着信息化的高速发展，各行各业、各组织单位积累了大量的业务数据，这些数据存在于各单位的数据库，各种报表、文档之中，真可谓是数据的海洋。这些数据中蕴涵着组织业务活动的大量规则，包含着组织管理决策所需要的重要知识。如何从这些数据中发掘出有价值的知识，为管理决策提供支持，是政府和企事业单位共同面临的问题。这个问题的解决依赖于两项技术：一是对整个组织各部门产生的各种业务数据进行统一和综合，把业务数据转换成商业信息，形成决策支持数据管理环境，即数据仓库；二是发现隐藏在各种数据之中有用的知识，即数据挖掘。数据挖掘也称为数据库知识发现，就是将高级智能计算技术应用于大量数据中，让计算机在有人或无人指导的情况下从海量数据中发现潜在的、有用的知识。

在发达国家，企业把数据仓库和数据挖掘看成是继 Internet 之后，企业在信息经济时代获得竞争优势的一个关键因素。美国 Meta Group 市场调查机构早在 1995 年发布的资料表明，《幸福》所列的全球 2 000 家大公司中已有 99% 将 Internet 和数据仓库这两项技术列入企业计划。目前，发达国家的大型企业几乎无一例外地把实施了以数据仓库和数据挖掘技术为核心的商业智能作为其发展战略的重要组成部分。在我国，数据仓库和数据挖掘应用虽然与发达国家相比存在一定的差距，但在电信、金融、证券、税务、零售业等行业的应用已取得了可喜的成果。可以预计，数据仓库和数据挖掘在我国同样会有广阔的应用前景。

美国计算机协会专业与教育委员会的创始人之一，Google 公司的高级技术经理 Kevin Scott 认为，员工的数据挖掘、统计建模等数据分析方面的技能是现代企业最看重的技能。人才的短缺是限制我国数据仓库和数据挖掘应用发展的一个重要“瓶颈”。如果在 Google 或百度搜索引擎中输入关键词“数据仓库　数据挖掘　人才招聘”进行搜索，就可以发现数据仓库与数据挖掘人员已成为当前企业的热门人才。数据仓库与数据挖掘人才分为两类：一类是开发型人才，其主要工作是从事数据仓库与数据挖掘工具的开发，这类人才主要受雇于大型软件开发公司，相对而言，需求量不是太大；另一类是应用型人才，对他们的要求是既懂业务，又懂工具的应用，他们的主要工作就是利用先进的数据管理工具，进行数据资源的管理与价值开发，为企业管理和决策提供支持。随着企业对数据资源管理的日益重视，高层次的数据资源管理人才将成为每一个现代化企业的必备人才。为适应这种形势的发展，数据仓库与数据挖掘人才的培养引起了各高校的高度重视，许多学校的经贸、管理类专业的本科生、研究生都开设数据仓库与数据挖掘课程。编者认为，数据仓库与数据挖掘应用人才应在具备基本知识和基本理论的基础上，注重实际应用能力的培养。本书正是在这样的指导思想下编写的。

本书共分为 8 章和 1 个附录。第 1 章介绍企业数据资源管理，旨在说明数据仓库与数据挖掘是企业数据资源管理的高级阶段，也是必然趋势。同时也介绍了数据仓库与数据挖掘能为企业做什么。第 2 章介绍数据仓库的概念与结构，读者在学完本章后可以明白什么是数据仓库，数据仓库是怎样构成的。第 3 章介绍数据仓库的设计与开发，在学完本章后，读者可以根据需求设计自己的数据仓库，并逐步地建立自己的数据仓库(或数据集市)。第 4

章介绍数据仓库的一个重要应用——联机分析处理，读者可以理解多维数据分析的方法。第 5 章介绍数据挖掘概述，在学完本章后，读者应能明白数据挖掘能做什么，有些什么方法。第 6 章介绍数据挖掘中的一个重要的也是工作量最大的一个环节——数据预处理，在学完本章后，读者应明白为什么要进行预处理，预处理的各种方法，并学会怎样进行数据的预处理。第 7 章介绍数据挖掘中的常用算法，读者可以理解这些算法的基本思想，掌握典型算法的应用。第 8 章结合一个实际案例介绍 SQL Server 2005 数据仓库与数据挖掘工具及其应用。建议读者在本章学习的基础上，自己深入钻研 SQL Server 2005 的数据仓库与数据挖掘工具。最后，以附录 A 的形式给出了一个简易的数据挖掘工具——Weka，旨在为本课程的教学提供一个方便的实验环境。

本书适合作为非计算机专业的教材及自学读物。为了让读者更容易理解数据仓库与数据挖掘的基本原理及其应用，本书每章都提供了引例、案例和适量的例题。同时，为了开阔读者的视野，每章提供一定量的阅读材料。建议在本课程的教学过程中，要理论联系实际，安排一定的实验。建议采用开放式实验方法，即学生自己准备实验数据和实验环境，解决一个实际问题，最终达到理论联系实际的目的。在教学过程中，也可以将附录 A 与数据挖掘的各章节结合起来进行讲解和学习。

参与编写本书的教师都具有丰富的教学经验。第 2 章由杨进负责编写；第 5 章由胡建军负责编写；第 6 章由邝云英负责编写；第 7 章由刘凤英负责编写；第 8 章由吴君负责编写；其他章节由廖开际负责编写。全书最后由廖开际、刘凤英和胡建军修改和统稿。在本书的编写过程中，得到了作者家人们的大力支持，在此对他们表示感谢！研究生赵兴庐为附录 A 的编写做了大量工作，在此表示感谢！

我们将在网站上(www.pup6.com)发布本书相关的教学资料，包括课件、实验环境和教学体会等。

由于编者水平有限，加之编写时间仓促，疏漏之处在所难免，欢迎广大读者批评指正。

编　者

2008 年 8 月

目　录

第1章 企业数据资源管理

教学目标

通过本章的学习，应认识到数据是企业的一种基本资源，了解企业数据按照不同的用途可分为操作型(业务)数据和分析型(决策)数据，相应地用数据库和数据仓库进行管理。

教学要求

知识要点	能力要求	相关知识点
数据资源管理	(1) 理解为什么数据是一种资源 (2) 了解数据资源管理的技术历程 (3) 理解数据资源管理的意义	(1) 数据文件 (2) 管理信息系统 (3) 决策支持系统
数据资源管理的技术体系	(1) 了解数据仓库产生的背景 (2) 理解数据库与数据仓库的区别与联系 (3) 了解数据仓库的应用	(1) 数据库 (2) 数据仓库 (3) OLAP 和数据挖掘

引例

美国的美洲银行2005年的数据仓库已拥有800GB存储信息。银行副总裁走入工作室可以毫不费力地查询“硅谷地区有多少居民拥有高尔夫球会员资格，多少人拥有家庭游泳池”，由此为美洲银行带来了竞争优势。因为银行通过了解自己客户的生活方式，来制定银行的服务规范，以满足客户的个性化需求，扩大客户群。

但是，要对如此大量的数据进行有效管理，让用户实现联机任意查询，并获得有用的查询结果，对企业来说是一种挑战。思考以下问题。

(1) 数据仓库中大量的数据从何而来？

(2) 数据要如何组织才有利于数据的分析查询？

(3) 如何保证查询分析结果的有效性？

随着企业信息化的不断深入，用户数据的积累已达到一定的规模。对数据进行简单的、局部的和浅层次的查询、统计是数据资源最基本的应用，而对数据资源的增值应用则是对企业财务、业务进行全面的、历史的和多角度的分析。传统定制的报表应用由于技术上的缺陷，无论从功能还是效率方面都难以满足企业对信息快速增长的需求。企业业务系统沉淀下来的数据就像是一座尚未开发的金矿，不但没有得到很好的利用，相反还可能成为一种负担。企业在信息化建设方面的巨大投入却只获得了部分回报，造成了资源极大的浪费。

本章介绍数据资源管理的基本概念和主要内容。

1.1 数据资源的概念

人、财、物是企业的有形资源，描述这些资源和企业过程的数据是企业的另一种重要资源，称为数据资源。掌握数据资源概念的关键在于理解其内涵：数据资源不仅限于数据本身，还包括用以产生数据的资源和人的因素。

1.1.1 企业资源

企业通过资源的交换与外部环境相互作用。企业经营管理者的任务是优化企业资源，最大限度地利用企业资源。

企业资源包括如下要素。

(1) 人：人力资源。

(2) 财：资金资源。

(3) 物：包括材料、设备和能源在内的资源。

(4) 数据：对物理资源的特征、状态及其相互作用关系的描述。

人、财、物资源是物理存在的，是有形的，称为物理资源。

数据资源的价值不在于它的物理存在，而在于它所表现的内容。数据资源是对物理资源的描述和表达，也称为元资源。利用数据资源可以有效地对物理资源进行调度、控制和

计划。因此，数据资源是一种管理资源。

在企业中，管理人员利用数据资源来管理物理资源。企业管理人员的任务是管理这些资源，以便最有效地利用它们。

1.1.2 数据资源

1. 定义

狭义的数据资源是指数据本身，即企业运作中积累下来的各种各样的数据记录，如客户记录、销售记录、人事记录、采购记录、财务数据和库存数据等。

广义的数据资源涉及数据的产生、处理、传播、交换的整个过程，包括数据本身、数据的管理工具(计算机与通信技术)和数据管理专业人员等。广义的数据资源概念更能反映现代数据资源开发利用的要求。

作为管理资源的数据资源不仅限于数据本身，还包括用以产生、加工、存储和使用数据的资源。

2. 从数据中获取信息和知识

企业的信息系统由输入(数据)部分、数据处理部分和输出(信息和知识)部分组成，如图 1.1 所示。

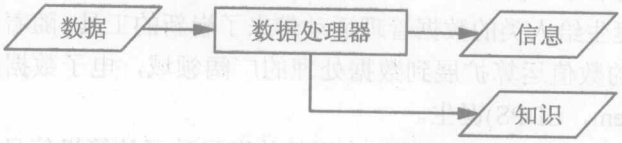

图 1.1 企业的信息系统

(1) 数据：信息系统的输入。

(2) 信息和知识：信息系统的输出，体现数据资源的经济价值。信息反映的是数据本身之间的关系，知识反映的是数据描述的物理资源之间的关系。

(3) 数据处理器：将数据转换成信息或者从数据中挖掘出知识，是对数据资源价值的提升，也是数据资源的基本要素。数据处理器既包括计算机元素，也包括非计算机元素。从广义上讲，数据处理器是企业数字神经系统的基础，其中包括系统的硬件和软件、开发和使用系统的人员、计算机网络以及安置这些资源的设施。

3. 数据资源的组成

数据资源由以下 3 个要素组成。

(1) 有用的数据。

(2) 作为企业神经系统的信息基础设施(Information Infrastructure)，如计算机硬件、软件以及网络系统。

(3) 人的因素，如系统人员和用户。

1.1.3 数据资源管理及其发展历程

数据资源管理涉及数据资源生命周期的各个过程。

(1) 数据获取：确保能够收集到必要的原始数据。

(2) 数据加工：将原始数据加工成有用的信息。

(3) 数据利用：确保能够在适当的时间、以适当的形式得到必要的数据和信息，从数据中发现决策所需要的知识。

(4) 数据报：在适当的时候废弃过时的数据，并代之以及时、准确的数据。

狭义的数据资源管理是指对数据本身的管理，即获取数据、有效地利用数据、在适当的时候废弃过时数据的整个活动过程。

广义的数据资源管理是指在对数据本身进行管理的基础上，增加信息系统管理的内容，即对涉及数据活动的各个要素(数据、技术、人员、机构等)进行合理的计划、集成和控制，以实现对数据资源的充分开发和有效利用，从而满足社会数据的需求。

本章讨论的数据资源管理是指狭义的概念。

随着信息技术的飞速发展和社会竞争的日趋激烈，组织的数据管理活动日渐活跃，各种各样的信息系统应运而生。为适应社会信息环境的变化和组织经营管理的需要，组织的数据管理模式也在不断地进行调整和变革。这种演变大致遵循“电子数据处理系统→管理信息系统→决策支持系统”的路线。

传统上，企业等社会组织的信息系统和图书馆一样，管理对象主要是纸质文献资料，管理作业基本靠人力和手工劳动，主要解决的是文献资料的收集、整理和保存问题。

电子计算机的诞生给人类的数据管理活动带来了崭新的工具。随着计算机技术的发展，计算机应用从单纯的数值运算扩展到数据处理的广阔领域，电子数据处理系统(Electronic Data Processing System，EDPS)诞生。

20世纪50年代至60年代企业计算机应用的热潮导致了计算机信息系统的形成和发展，并带来了组织信息系统的首次繁荣。但是，随着时间的推移，它也逐渐暴露出许多局限性和不足。EDPS 只能完成单纯的数据处理工作，缺乏分析预测能力，不能满足组织经营管理的需要。

管理信息系统(Management Information System，MIS)是在 EDPS 的基础之上于 20 世纪 60 年代中期逐步发展而来的。MIS 避免了 EDPS 的一些弊端，在信息处理的方法、手段和技术方面都有明显的进步。较之 EDPS，MIS 具有如下特点。

(1) 更加强调科学管理方法和定量化管理模型的运用，强调系统优化的作用。

(2) 更加强调对数据的深层次开发利用，强调信息系统对生产经营过程的预测和控制作用。

(3) 更加强调科学的、系统化的开发方法，强调高效率、低成本的系统结构和数据处理模式。

MIS 采用标准的工具和技术手段对组织的数据进行加工处理，在组织的数据管理中得到了十分广泛的应用。但是，随着信息技术突飞猛进的发展和组织信息环境日新月异的变化，MIS 的管理决策功能薄弱、只管理内部数据而不管理外部数据、只有业务信息而没有办公信息的局限性日益突出。针对这些不足，20 世纪 70 年代以后又先后兴起了决策支持系统(Decision Support System，DSS)和办公自动化系统(Office Automation System，OAS)。

DSS 是 20 世纪 70 年代初期在 MIS 的基础上发展起来的支持决策者对半结构化管理问

题进行决策的信息系统，它的进步在于将信息系统的注意力转向高层管理决策者，并相应地引入外部数据，以及强调人机交互和用户友好。决策模型和用户共同驱动系统的运行最终为决策者提供了切实可行的决策方案。但是，DSS 对 MIS 的发展仍然还是沿着技术的路线，试图通过技术手段和模型化的方法提高决策的效益，这对于在当今急剧变化的社会信息环境下日益复杂化的战略决策问题是很难奏效的。

1.2 数据资源管理的意义

数据资源管理的意义可从以下 3 个方面来理解。

(1) 从信息系统的发展过程来看，对数据资源进行有效的管理是信息系统进入成熟阶段的重要标志。

(2) 数据资源管理是解决企业内部由于数据重复而导致的各种问题的根本途径。

(3) 数据资源是企业取得竞争优势的关键。

1.2.1 信息系统进入成熟阶段的重要标志

企业的信息系统的成长过程分为 5 个阶段：初级、普及、整理、集成、成熟。

在初级阶段，计算机刚进入企业，只作为办公设备使用，应用非常少，通常用来完成一些报表统计工作，甚至大多数时候被当做打字机使用。随着企业对计算机应用认识的深入，人们体会到计算机应用的价值，开始学习、使用和维护计算机。

在普及阶段，计算机应用在一些部门见到成效，从最初的一些应用部门向其他部门蔓延，大量的人工数据处理转向计算机处理，人们对计算机的热情增加，需求增长。

在整理阶段，由于人们对计算机信息处理需求的增长，造成财务支出的大幅度上涨，企业领导不得不对之进行控制，注重采用成本/效益去分析应用开发。并针对各项已开发的应用项目之间的不协调和数据冗余等问题进行统一规划。这一阶段的效益可能比第二阶段还要低。

在集成阶段，企业高层领导意识到信息战略的重要，信息成为企业的重要资源，企业的信息化建设也真正进入到数据处理阶段。在这一阶段中，工作的重点是对数据资源进行管理和控制，包括数据的质量控制、数据的有效利用和数据价值的提升。企业开始选定统一的数据库平台、数据管理体系和信息管理平台，以统一数据的管理和使用，各部门、各系统基本实现资源整合、信息共享。

在成熟阶段，信息系统已经可以满足企业各个层次的需求，从简单的事务处理到支持高效管理的决策。工作的重点是建立以数据资源为基础的系统计划和战略计划，将信息系统作为取得竞争优势的有力手段。

1.2.2 解决企业内部数据不一致问题的根本途径

在计算机化事务处理的初级阶段，事务相关的数据以文件的形式组织和保存，并作为编制处理程序的基础。同一组数据在不同的应用系统中不能共享，需要重复存储，如图 1.2

所示。因而导致在数据文件之间数据的重复和不一致问题，这将对企业的经营活动产生严重的影响。

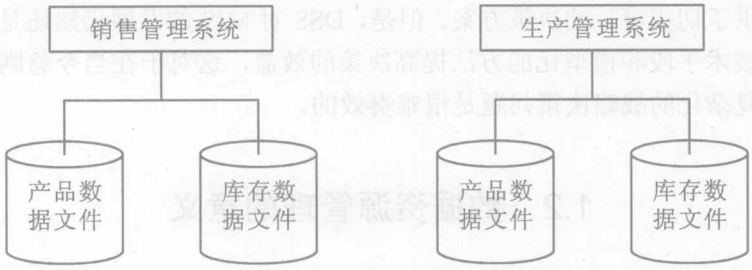

图 1.2　以文件系统方式管理数据造成的数据重复问题

数据重复从根本上影响着信息系统的管理和应用。为了解决数据重复的问题，必须通过数据库系统对数据资源进行全面管理，确保相互重复的数据文件在任何时候都可以得到同步更新，如图 1.3 所示。

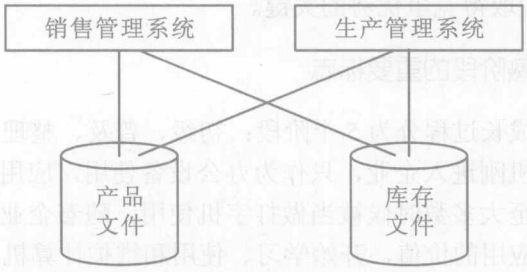

图 1.3　以数据库方式管理数据解决数据重复问题

1.2.3　数据资源的管理和应用是取得竞争优势的关键

在高度信息化时代，数据资源的管理和应用是取得竞争优势的关键。关于这一观点，可以通过对企业价值链和价值系统的分析来得到答案。

1. 企业价值链

企业通过一系列的价值活动来实现价值增值，企业内部的所有价值活动联结在一起形成了价值链。

价值链由一系列的价值活动构成，每个价值活动又都包括资源输入、人力资源和技术等基本要素，并且每个价值活动都使用和产生数据资源。

在这一系列的活动中，涉及信息基础设施(如数据库系统和计算机设备)、人的因素(如系统开发人员、使用系统的领导)和数据本身这 3 个信息资源的基本要素。

2. 价值系统

随着互联网和电子商务的普及，经营管理者在建立和完善企业内部价值链的同时，更加致力于将企业内部的价值链与其他企业的价值链相连，以取得进一步的附加价值。这种企业间的价值链的连接被称做价值系统。例如，一个制造厂家可以通过采购管理系统将内部价值链与供应商的价值链相连，以便在必要的时候获得必要的输入资源，达到降低采购

和材料库存成本的目的。同时，该厂家也可以通过销售系统与销售商的价值链相连，以达到降低销售成本和产品库存成本的目的。与企业内部价值链一样，价值系统的每个环节也都使用和产生数据资源。

无论是在企业价值链还是在连接企业价值链的价值系统中，每个价值活动都包括数据元素，数据资源的利用成为企业取得竞争优势的关键。在高度信息化时代，竞争优势意味着利用数据获得更大的市场份额和利益。企业要想在竞争中获胜，不仅需要优越的物理资源，还需要优越的数据资源。企业的经营管理者应该综合利用物理资源和数据资源，来实现企业的战略目标。

1.3 信息资源管理的相关技术

本节主要从技术构成的角度来讨论数据资源管理过程中的关键要素和它们之间的相互关系。

1.3.1 数据资源管理的技术框架

图 1.4 给出了一个数据资源管理的技术框架。从图中可以看出，数据资源管理的技术框架主要由 3 部分构成。

(1) 面向业务操作的数据资源管理：包括数据库、事务处理系统以及管理信息系统。

(2) 面向决策分析的数据资源管理：包括数据仓库以及与之紧密相关的决策支持系统。

(3) 知识资源的管理和利用：包括知识库以及基于知识的系统。

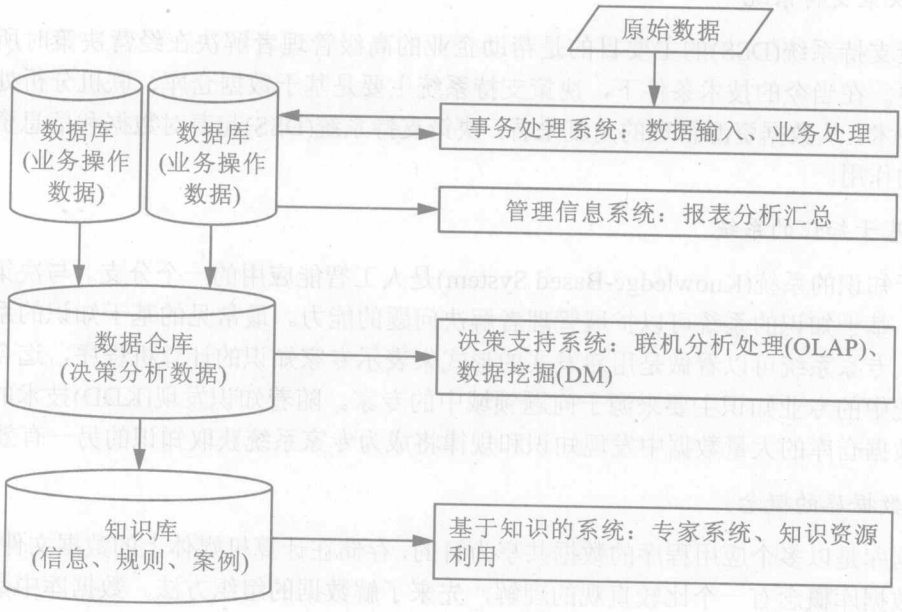

图 1.4 数据资源管理的技术框架