

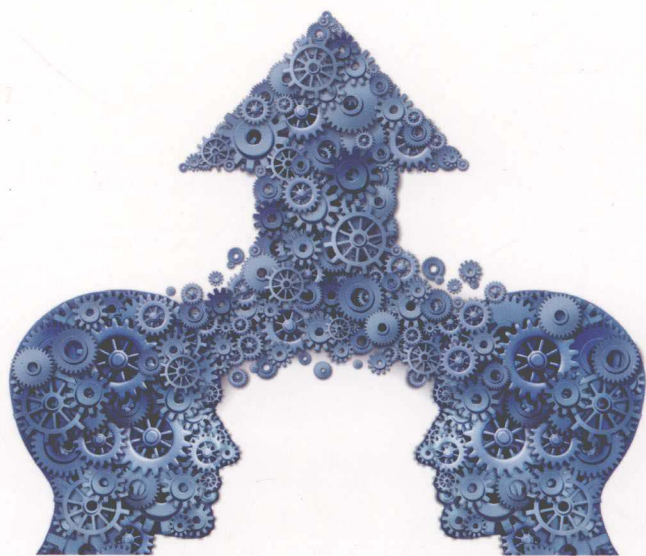
数据中国“百校工程”项目系列教材
数据科学与大数据技术专业系列规划教材

瑞翼教育

商务智能 方法与应用

张小梅 许桂秋 ● 主编

郜丹 张晴 温远远 秦朝阳 陈婷婷 ● 副主编



BIG DATA
Technology



中国工信出版集团



人民邮电出版社
POSTS & TELECOM PRESS

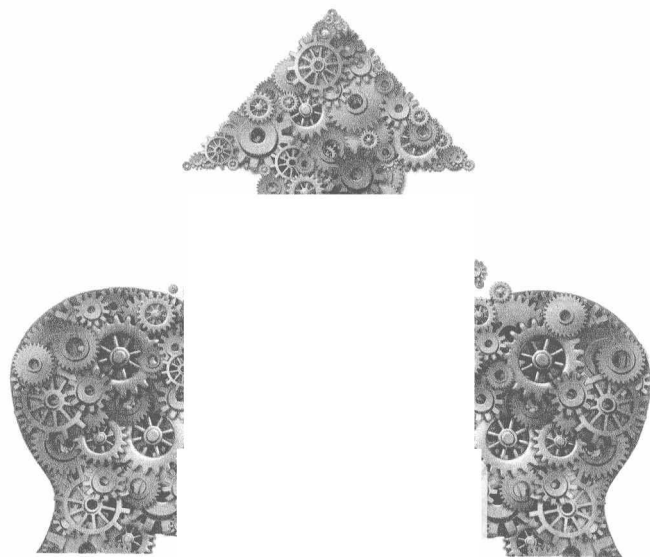
“工程”项目系列教材
技术专业系列规划教材

瑞翼教育

商务智能 方法与应用

张小梅 许桂秋 ● 主编

郜丹 张晴 温远远 秦朝阳 陈婷婷 ● 副主编



BIG DATA
Technology

人民邮电出版社

北京

图书在版编目 (C I P) 数据

商务智能方法与应用 / 张小梅, 许桂秋主编. — 北京: 人民邮电出版社, 2019.5
数据科学与大数据技术专业系列规划教材
ISBN 978-7-115-50348-0

I. ①商… II. ①张… ②许… III. ①数据处理—教材 IV. ①TP274

中国版本图书馆CIP数据核字(2019)第046743号

内 容 提 要

本书从实用的角度出发,采用理论与实践相结合的方式,介绍商务智能的基础知识、技术与方法,力求培养读者使用商务智能技术解决企业实际问题的能力。全书内容包括商务智能的基本概念,商务智能的架构,商务智能涉及的数据处理关键技术,数据仓库,维度建模,联机分析处理,以及商务智能在零售、客户关系管理、电子商务等领域的应用。

本书作为商务智能的入门教材,目的不在于覆盖商务智能技术的所有知识点,而是介绍商务智能的主要应用,使读者了解商务智能的基本构成以及如何根据行业的特点构建商务智能系统。为了增强实践效果,本书引入了4个基础实验,帮助读者了解商务智能涉及的基本技术知识和技能;在此基础上又引入了3个综合实践案例,帮助读者掌握如何在不同行业场景下构建商务智能系统。

本书可以作为高校商务智能基础与应用课程的教材,也可供对商务智能感兴趣的读者阅读。

◆ 主 编 张小梅 许桂秋
副 主 编 郜 丹 张 晴 温远远 秦朝阳 陈婷婷
责任编辑 张 斌
责任印制 陈 彝

◆ 人民邮电出版社出版发行 北京市丰台区成寿寺路11号
邮编 100164 电子邮件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
固安县铭成印刷有限公司印刷

◆ 开本: 787×1092 1/16
印张: 14.5 2019年5月第1版
字数: 335千字 2019年5月河北第1次印刷

定价: 59.80 元

读者服务热线: (010)81055256 印装质量热线: (010)81055316
反盗版热线: (010)81055315

广告经营许可证: 京东工商广登字 20170147 号

随着信息时代的发展,尤其是大数据时代的到来,数据已经成为企业重要的信息化资产。各种企业管理系统积累的数据正以惊人的速度增长,这些数据具有重要的商业价值,基于数据的决策已经得到了企业的全面关注,但是目前大多数企业对数据的利用率还很低。如何将数据转换为信息,将信息转换为知识,并将知识有效地运用到企业的决策当中,是各个企业亟待解决的问题,商务智能概念便应运而生。企业的长足发展需要灵敏的嗅觉以及快速的反应能力。提高企业的反应速度和决策的准确性,快速抢占市场,这些都离不开商务智能。

根据商务智能课程实践性较强的特点,本书采用理论与实践相结合的方式,在介绍商务智能的基础知识、核心技术,以及商务智能在零售、客户关系管理、电子商务等领域的应用的同时,力求培养读者使用商务智能技术解决企业实际问题的能力。

通过学习本书,读者能够提升对商务智能的认识;通过练习本书相关的实践案例,读者能够初步具备设计商务智能系统的能力。

全书共 7 章,主要内容如下。

第 1 章概述商务智能的基本知识与常用技术,并设置了一个帮助读者掌握商务智能基本技术的实验。

第 2 章介绍了商务智能的重要技术手段——数据仓库,并设置了一个搭建数据仓库的实验。

第 3 章介绍了商务智能的重要技术手段——维度建模,并设置了一个构建数据立方体的实验。

第 4 章介绍了商务智能的重要技术手段——联机分析处理,并设置了一个对数据立方体进行切片、切块、旋转等数据分析操作的实验。

第 5 章从目前零售业商务智能的现状、客户关系管理、日常经营分析以及零售业案例等几个方面讲解商务智能在零售业中的应用,并提供了一个综合性较强的实践案例——购物清单关联性分析。该案例利用 Weka 智能分析软件,对销售数据进行关联性分析,研究超市如何摆放产品将有利于提高销售额度。

第 6 章从客户关系管理概述、客户细分、客户识别和客户洽谈、客户维度与属性以及复杂的客户行为等几个方面,讲解商务智能在客户关系管理中的应用,并提供了一个综合性较强的实践案例——航空客运信息挖掘。该案例从航空客运信息源数据中挖掘处理数据,建立数据模型,根据数据模型预测潜在客户。通过对本章内容的学习,读者可以初步掌握商务智能在客户关系管理中的应用方法。

第7章从智能搜索、电子商务情感分析、智能推荐技术等几个方面讲解商务智能在电子商务领域的应用，并提供了一个综合性较强的实践案例——消费者评论数据情感分析。通过对本章内容的学习，读者可以初步掌握商务智能在电子商务领域的应用方法。

本书可以作为高等院校计算机和信息管理等相关专业的商务智能课程的教材。建议安排课时为64课时，教师可根据学生的接受能力以及高校的培养方案选择相应的教学内容。

由于编者水平有限，编写时间仓促，书中难免存在疏漏和不足之处，恳请广大读者批评指正。

编者

2019年1月

目 录

第1章 商务智能概述1	第2章 数据仓库40
1.1 商务智能产生的背景.....1	2.1 数据仓库概述.....40
1.1.1 商务智能产生的原因.....1	2.1.1 数据仓库的概念.....40
1.1.2 商业决策需要商务智能.....3	2.1.2 数据仓库的特点.....41
1.1.3 企业智能化管理需要商务智能.....4	2.1.3 数据仓库的结构.....41
1.2 商务智能简介.....4	2.1.4 数据仓库与数据库.....42
1.2.1 商务智能的概念.....5	2.1.5 数据仓库和商务智能的关系.....43
1.2.2 商务智能的发展.....7	2.2 ETL 过程.....43
1.2.3 商务智能的要求.....8	2.2.1 数据抽取.....43
1.2.4 商务智能的价值.....9	2.2.2 数据转换.....44
1.3 商务智能基础.....10	2.2.3 数据清洗.....44
1.3.1 商务智能的基本架构.....10	2.2.4 数据加载.....45
1.3.2 商务智能的功能.....11	2.3 数据仓库工具 Hive.....45
1.4 商务智能的关键技术.....12	2.3.1 Hive 的数据类型与存储格式.....45
1.4.1 数据预处理.....12	2.3.2 Hive 的数据模型.....50
1.4.2 数据仓库.....13	2.3.3 查询数据.....52
1.4.3 数据挖掘.....13	2.3.4 用户定义函数.....53
1.4.4 联机分析处理.....15	实验2 数据仓库的建立.....54
1.4.5 数据可视化.....15	第3章 维度建模68
1.5 商务智能的相关应用.....16	3.1 维度建模简介.....68
1.5.1 商务智能在金融业的应用.....16	3.1.1 维度建模的概念.....69
1.5.2 商务智能在保险业的应用.....16	3.1.2 维度建模的基本原则.....69
1.5.3 商务智能在教育领域的应用.....16	3.2 维度表技术基础.....71
1.5.4 商务智能在客户关系管理的应用.....17	3.2.1 维度表的结构.....71
1.5.5 商务智能在零售业的应用.....17	3.2.2 维度代理键.....71
1.5.6 商务智能在电子商务领域的应用.....18	3.2.3 多维体系架构.....72
1.5.7 商务智能在制造业的应用.....18	3.2.4 缓慢变化维度.....75
实验1 销售数据预处理.....19	3.3 事实表技术基础.....76
	3.3.1 事实表的结构.....76
	3.3.2 可加、半可加、不可加性事实.....77

3.3.3 事实表中的空值.....	77	4.5 OLAP 操作语言.....	115
3.3.4 事实表的基本类型.....	77	4.5.1 MDX	115
3.4 维度建模的主要流程.....	78	4.5.2 MDX 查询语句.....	117
3.4.1 选择业务流程.....	79	4.5.3 SQL 和 MDX 的区别.....	118
3.4.2 声明粒度.....	79	4.5.4 MDX 表示.....	119
3.4.3 确认维度.....	80	4.5.5 成员属性和单元属性.....	120
3.4.4 确认事实.....	80	4.5.6 MDX 查询结构.....	122
3.5 对维度建模的误解.....	80	4.6 主流的 OLAP 工具.....	124
3.5.1 误解 1: 维度模型仅用于汇总 数据.....	80	4.6.1 OLAP 产品.....	124
3.5.2 误解 2: 维度模型是部门级的 而不是企业级的.....	81	4.6.2 OLAP 的实现过程.....	125
3.5.3 误解 3: 维度模型是不可 扩展的.....	81	实验 4 联机分析.....	127
3.5.4 误解 4: 维度模型仅可用于 预测.....	81	第 5 章 商务智能在零售业 的应用	134
3.5.5 误解 5: 维度模型不能集成.....	81	5.1 零售业商务智能现状.....	134
实验 3 使用 Schema Workbench 创建 Cube.....	82	5.2 客户关系管理.....	135
第 4 章 联机分析处理	96	5.3 零售管理业务优化.....	136
4.1 OLAP 简介.....	96	5.4 日常经营分析.....	136
4.1.1 维度模型的基本概念.....	97	5.4.1 商品分析.....	136
4.1.2 OLAP 的多维数据结构.....	100	5.4.2 销售分析.....	137
4.1.3 OLAP 的应用.....	103	5.4.3 会员卡分析.....	138
4.2 OLAP 多维数据分析.....	104	5.4.4 财务分析.....	138
4.2.1 切片和切块.....	105	5.5 零售业案例.....	140
4.2.2 钻取.....	106	5.5.1 数据仓库的搭建.....	141
4.2.3 旋转/转轴.....	106	5.5.2 粒度设计.....	141
4.3 OLAP 分类.....	107	5.5.3 星形模型设计.....	142
4.3.1 ROLAP、MOLAP 与 HOLAP.....	107	5.5.4 ETL 设计.....	146
4.3.2 多维数据模式.....	109	5.5.5 OLAP 的实现.....	148
4.3.3 OLAP 体系结构.....	111	5.5.6 数据挖掘.....	151
4.3.4 OLAP 与 OLTP 的区别.....	112	实验 5 购物清单关联性分析.....	156
4.4 从 OLAP 到数据挖掘.....	113	第 6 章 商务智能在客户关系 管理中的应用	162
4.4.1 数据仓库应用.....	113	6.1 客户关系管理概述.....	162
4.4.2 OLAP 和数据挖掘的关系.....	113	6.1.1 客户智能.....	162
4.4.3 多维数据挖掘.....	114	6.1.2 数据挖掘在客户关系管理中 的应用.....	164
		6.2 客户细分.....	166
		6.3 客户识别和客户流失.....	168

6.3.1	数据挖掘应用于客户识别	168
6.3.2	通过当前客户了解潜在客户	169
6.3.3	客户流失	170
6.4	客户维度与属性	171
6.4.1	姓名和地址的语法分析	171
6.4.2	国际姓名和地址的考虑	173
6.4.3	以客户为中心的日期	174
6.4.4	基于事实表汇聚的维度属性	174
6.4.5	分段属性与记分	175
6.4.6	客户维度变化的计算	177
6.4.7	低粒度属性集合的维度表	177
6.4.8	客户层次的考虑	178
6.5	复杂的客户行为	179
6.5.1	行为类型分析	179
6.5.2	连续行为分析	180
6.5.3	行为分析模型	181
6.5.4	时间范围事实表	183
6.5.5	使用满意度指标标记事实表	185
6.5.6	使用异常情景指标标记事实表	185
实验 6	航空客运信息挖掘	186

第 7 章 商务智能在电子商务领域的应用

7.1	智能搜索	199
7.1.1	网络机器人	200
7.1.2	文本分析	201
7.1.3	搜索条件的获取和分析	203
7.1.4	信息的搜索和排序	204
7.2	电子商务情感分析	206
7.2.1	评论数据收集及处理	207
7.2.2	扩展特征向量构造	207
7.2.3	情感词库构建	207
7.2.4	情感分析模型	208
7.2.5	情感倾向值计算	208
7.3	智能推荐	209
7.3.1	智能推荐产生背景及定义	209
7.3.2	智能推荐主要算法	211
7.3.3	智能推荐在电子商务中的应用	213
实验 7	消费者评论数据情感分析	215
参考文献		224

第 1 章

商务智能概述

本章采取理论和实践相结合的方式介绍商务智能的基础知识。首先介绍商务智能产生的背景，然后引出商务智能的概念、发展、基本架构、关键技术及其相关应用等。

本章介绍的是商务智能的基本原理和基础知识，对商务智能已有一定理论基础的读者可有选择地学习本章内容。

本章重点内容如下。

- (1) 商务智能的概念。
- (2) 商务智能的基本架构。
- (3) 商务智能的关键技术。

1.1 商务智能产生的背景

随着信息时代的发展，尤其是大数据时代的到来，企业的海量数据正在以惊人的速度增长。数据是企业最重要的资源。据统计，目前我国国内企业对数据的有效利用率不足 7%，很多企业还是依靠传统观念进行商业决策，并没有充分有效地利用信息。正确、有效地利用企业信息化资产，将数据转化为对企业有利的信息和知识，使商务智能 (Business Intelligence, BI) 有效地应用到企业的决策中，提高企业管理水平，已经成为智能企业与传统企业的主要区别。商务智能的兴起并非偶然，它将会为企业带来新的生命力。

1.1.1 商务智能产生的原因

随着信息时代的发展，特别是大数据时代的到来，企业面临前所未有的机遇和挑战。如何正确、及时地响应市场需求，快速占领市场，是每个企业都急需解决的问题。

将企业收集的海量数据正确、及时、有效地转化为信息，再转化为知识，最终支持企业决策，这是商务智能发展的驱动力。商务智能产生的原因可概括为如下几点。

1. 急切的分析型需求

20 世纪 90 年代以来，我国经济快速发展，涌现出了一大批有财力、有活力的公司，特别是改制后的大型国有企业、知名民营企业等，其整体建设都逐步向国际领先企业靠拢，陆续建设了核心业务系统。例如，电信行业的计费系统、生产制造行业的制造企业生产过程执行系统 (Manufacturing Execution System, MES)、零售分销行业的 ERP 系统

等，基本都在这一时期建成，这些系统大大提高了操作人员的工作流程规范化水平，资金流也得到了严格控制和监管。

随着这些系统的完善，特别是业务系统数据的积累，公司业务分析及决策人员发现基于业务系统基础数据的各种分析对其决策非常有帮助，于是在工作过程中越来越依赖系统数据，这一时期的分析工作普遍采取如下流程。

(1) 先从业务系统导出数据到 Excel 软件，然后通过 Excel 进行加工，最后生成报表。

(2) 生成的报表除了支持自己工作之外，还可以传递给相关部门和领导。

在这个分析过程中，普遍存在如下突出的问题。

(1) 业务系统压力大：业务系统的数据量越来越大，从业务系统查数据、导出数据越来越慢，并且频繁地导入、导出数据也会严重影响业务系统的工作效率。

(2) 制作手工报表耗时长：一般情况下，导出数据的时间加上数据处理加工的时间，再加上撰写周报、月报的时间，至少需要 1~2 天，耗时太长。

(3) 数据不统一：每个部门都有分析人员，也各自都在出报表和做分析，数据和分析结果很难共享，汇总给领导的数据容易“打架”，数据不统一。

(4) 决策难于深化：员工花费大量时间做数据处理，没有时间对数据进行细致分析，发现问题之后难以做深入的相关分析，最终出现“员工很忙，领导很急”的状况。

(5) 电子商务行业的核心数据是日志，如点击日志、搜索日志等。这些数据一开始就非常庞大，传统行业常用的 Excel 软件对系统日志的数据支撑是无能为力的。

上述问题在业务系统运营之后一直存在，而且随着时间的推移越发严重，特别是发展迅速的公司数据量激增，这些问题更加明显。

2. 企业精细化管理需要商务智能的支撑

企业精细化管理的核心思想是“快、精、准”，这些都需要商务智能的强力支撑。

(1) 快：要求自上而下的实时把控，第一时间发现问题需要商务智能支持，特别是高层领导对全公司的问题都能及时发现，需要迅速调出各种流程控制系统的数据进行分析，发现异常。

(2) 精：能对问题追根溯源，需要商务智能的向下钻取、向上钻取、交叉分析、关联分析等基本技术支撑，否则看到问题也不知道原因在哪里。

(3) 准：就是要求问题落实到人。例如，问题出在谁身上，谁该受到惩罚，谁该进行工作改进等。

3. 数据中蕴含的知识可以帮助企业进行优化升级

数据—知识—操作—数据这样一个信息闭环其实就是实践—总结—再实践的一个螺旋式上升过程，如果这个过程中缺少数据到知识的一步，那就是简单重复的操作；而加入知识总结这样一个分析之后的实践，则是有提升的实践，其结果能促使员工的工作不断得到修正和优化，企业管理不断升级。因此，BI 系统对数据进行知识化是企业优化升级的必然需求。

4. 知识产品化

在电子商务行业，不论是企业与企业间 (Business to Business, B2B)，还是企业与消费者间 (Business to Customer, B2C)，商务智能产生的知识都可以协助电子商务网站

的设计者提升网站的友好性，让网站设计者了解用户的习惯和行为，设计出符合用户操作的流程和功能；能根据买家的行为提供个性化的商品推荐，帮助用户提升工作效率和采购效果，提升他们对网站的认可度。另外，企业将产品放在电子商务网站进行推广之后，也迫切希望知道推广有没有效果，以及可对哪些方面进行改进，这也正是商务智能可以提供给企业的。

企业为了更迅速、更准确地把握自身的问题和市场的状况，都需要将数据转换为知识，合理地运用知识来获得利润，帮助企业快速崛起，这就是商务智能产生的原因。

1.1.2 商业决策需要商务智能

在信息时代，智能化已经成为企业生存之本，企业资源计划（ERP）、客户关系管理（Customer Relationship Management, CRM）和供应链管理（Supply Chain Management, SCM）等提高企业管理效率的平台积累了大量的业务数据，但到目前为止，很多企业的这些数据还没有被有效地利用起来。如何将企业的信息化资产转变为企业需要的信息和知识，为管理者的决策提供有效的数据支持，是商务智能首要关心的问题。商务智能从业务上可以划分为以下 5 个层次。

（1）第一个层次是告诉企业发生了什么。商务智能可以提供事先预制好的报告、企业平衡记分卡或综合管理“仪表盘”，利用集中管理的关键绩效指标（Key Performance Indicator, KPI）解决企业运营绩效问题，监控企业的发展，将复杂的报告用简单的方式表达出来。

（2）第二个层次是让用户探索为何会出现问题，也叫例外分析。业务部门可以从固定的报表、报告和一些关键的 KPI 中，得到很多相关的信息，但是当发现问题时，他们需要了解为何发生了这些问题。这时，就需要用到联机分析处理（Online Analytical Processing, OLAP）。业务分析员经常需要根据发现的问题完成自己的分析和报告。在很多情况下，业务分析员和决策制定者需要一套商务智能的工具，通过访问集成好的数据仓库（Data Warehouse），获得需要的信息。

（3）第三个层次是让用户实时看到现在发生了什么，这个层次是实时的信息分析。为了实现该层次的分析，企业决策层要制定实时情况下的业务战略和决策，获得实时的数据，查询并解决当时发生的问题。因此，运营模式和业务流程会发生较大的变化。例如，当客户因某种原因对服务不满，要退出服务或者退货时，相关客户服务人员发现这个客户是企业的大客户，他应该迅速将情况反馈给大客户经理，大客户经理可以很快查找该客户的消费记录，并马上和客户联系，争取挽留客户，而不是等到客户流失后才急迫地与客户联系。

（4）第四个层次是帮企业预见即将发生什么。企业的决策者仅仅了解现在还远远不够，了解将来会发生什么以及对风险的预测和评估也都是非常重要的。企业还需要统计分析的功能，帮助企业对客户进行细分、预测客户的行为、预测客户业务的趋势、辨认欺诈行为等。若要满足以上的需求，要有复杂的算法、统计模型和大量的数据，因此需要具有支持大数据量的处理能力。

（5）第五个层次是希望发生什么。决策依据是由系统提供的，而系统的数据是由运营系统得到的，例如，由 Web 页获得，或者以基于市场条件和用户需求进行的特价、促

销活动得到。商务智能可以建立清楚的决策和业务政策，让事物沿着正确的轨迹、朝着预定的方向行进，达到预期的目标。

商务智能的技术基础是数据仓库、联机分析处理、数据挖掘等。其中，数据仓库用来存储和管理数据，其数据从运营层获取；联机分析处理把这些数据转化成信息，支持各级决策人员进行复杂查询和联机分析处理，并用直观易懂的图表把结果展现出来；数据挖掘（Data Mining）指从海量的数据中提取出隐含在数据中的有用知识，以便各级决策人员做出更有效的决策，提高企业决策能力，如图 1-1 所示。

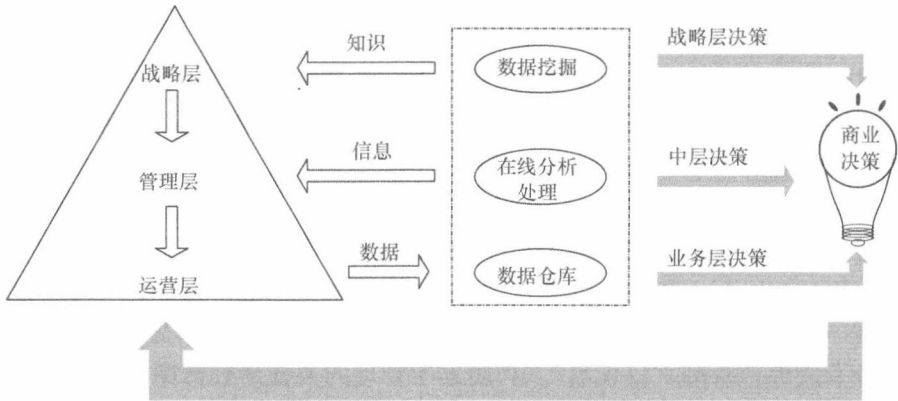


图 1-1 商务智能运行示意图

一个好的商务智能解决方案，可以帮助企业从大量的数据中获取有价值的信息和知识，并提供分析和统计预测的工具。例如，民生银行采用商务智能后，使用 Hyperion Intelligence（Oracle 公司的产品）帮助民生银行的各级人员实现数据查询、报表展示和决策分析等；业务人员可在基于 Web 的客户端进行数据查询、分析，直接生成灵活多样的图表报告；管理人员可通过使用 Hyperion Intelligence 对业务现状和发展趋势进行交互式分析，跟踪业务发展动态，及时解决经营过程中出现的问题。

1.1.3 企业智能化管理需要商务智能

企业智能化指企业如果想在瞬息万变的市场中生存，必须根据企业内部数据和相关市场数据，快速有效地处理企业内部问题、调整企业发展方向，以便适应客户需求的变化，采取正确的客户解决方案。领导者制定的决策决定企业的发展方向。因此，如何使企业管理策略得到科学的数据评价分析是企业决策者们最关心的问题。一个企业是否具有有效地利用在各种业务系统、数据集或数据仓库中的信息的能力决定了企业的发展。企业可以通过一个联系信息生产者和信息使用者的完整的信息供应链，实现企业商务智能化所带来的价值。要实现企业智能化管理必须借助商务智能技术。

1.2 商务智能简介

提到“商务智能(Business Intelligence)”这个词，人们普遍认为是加特纳集团(Gartner

Group) 在 1996 年首次提出来的,但事实上 IBM 公司的研究员汉斯·彼得·卢恩(Hans Peter Luhn)早在 1958 年就用到这一概念。他将“智能”定义为“对事物相互关系的一种理解能力,并依靠这种能力去指导决策,以达到预期的目标”。“商务智能”通常被理解为将企业中现有的数据转化为知识,帮助企业做出明智的业务经营决策的工具。这里所说的“数据”包括:来自企业业务系统的订单、库存、交易账目等的数据;来自客户和供应商等的的数据;来自企业所处行业和竞争对手的数据;来自企业所处的其他外部环境中的各种数据等。而商务智能能够辅助的业务经营决策,既可以是操作层的决策,也可以是战术层和战略层的决策。为了将数据转化为知识,需要利用数据仓库、联机分析处理(OLAP)工具和数据挖掘等技术。因此,商务智能从技术层面上来讲不是新技术,它是将数据仓库、数据挖掘、联机分析处理等技术进行整合,最终形成一个完整的解决方案,帮助企业进行决策。它可以根据大量的企业内、外数据发现隐匿在其中的商机或威胁,了解企业和市场的现状、把握趋势、识别风险、理解企业业务状况,从而提高企业核心竞争力,为企业带来更多价值。

1.2.1 商务智能的概念

现代意义上的商务智能的概念是 1996 年由加特纳集团(Gartner Group)提出的,加特纳集团将其定义为:商务智能描述了一系列的概念和方法,通过应用基于事实的支持系统来辅助商业决策的制定。商务智能技术提供使企业迅速分析数据的技术和方法,包括收集、管理和分析数据,将这些数据转化为有用的信息,然后分发到企业各处。

商务智能自产生以来发展非常迅速,但是目前尚不成熟,不同企业基于自己的出发点对商务智能的概念或多或少都有着不同理解,如表 1-1 所示。

表 1-1 不同企业对商务智能的定义

企业	商务智能的定义
Business Objects (SAP)	商务智能是一种基于大量数据的信息提炼的过程。这个过程与知识共享和知识创造密切结合,完成了从信息到知识的转变,最终为商家创造更多的利润
IBM	商务智能是一系列技术支持的简化信息收集、分析的策略集合
Microsoft	商务智能是任何尝试获取、分析企业数据,以更清楚地了解市场和顾客,改进企业流程,更有效地参与竞争的过程
IDC	商务智能是下列软件工具的集合:终端用户查询和报告工具、联机分析处理工具、数据挖掘软件、数据集市、数据仓库产品和主管信息系统
Oracle	商务智能是一种商务战略,能够持续不断地对企业经营理念、组织结构和业务流程进行重组,实现以顾客为中心的自动化管理
SAP	商务智能是通过收集、存储、分析和访问数据帮助企业更好决策的技术
Data Warehouse Institute	商务智能是把数据转换成知识,并把知识应用到商业运营的一个过程

通过表 1-1 可以看出,企业对商务智能的定义倾向于从技术、应用的角度,更多的是从商务智能的过程去描述并理解商务智能。在商务智能发展的早期,加特纳集团认为商务智能是数据仓库、数据集市、查询报表、数据分析、数据挖掘以及数据备份与恢复等辅助企业解决问题的技术及其应用。在 2007 年的商务智能峰会上,商务智能有了新的

定义，有学者认为商务智能可以是一个伞状的概念，其内容包括分析应用、技术架构、平台以及实践。这意味着业界对商务智能的认识跳出了技术的范畴，商务智能不再仅仅指技术工具的集合。

综上所述，商务智能是融合了先进信息技术与创新管理理念的结合体。商务智能集成了企业内外的数据，对数据进行加工处理并从中挖掘出知识，为企业创造更多的商业价值；商务智能面向企业战略并服务于战略层、管理层、运营层，指导企业经营决策，提升企业核心竞争力，达到从数据到知识再到利润的转变，从而为企业创造更多的效益。

商务智能集成大数据的关键技术（数据采集、数据预处理、数据分析与挖掘、数据可视化）于一体，构建了辅助企业决策的商务智能系统。本书中提到的商务智能与商务智能系统含义完全一致。

1. 数据

数据（Data）是用来记录、描述和识别事物的符号，是对客观事物的性质、状态以及相互关系等进行记载的物理符号或这些物理符号的组合。它是可识别的、抽象的符号。

数据不仅指狭义上的数字，也可以是具有一定意义的文字、字母、数字符号的组合，还可以是图形、图像、视频、音频等，同时也是客观事物的属性、数量、位置及其相互关系的抽象表示。例如，“0、1、2……”“阴、雨、下降、气温”“学生的档案记录、货物的运输情况”等都是数据。数据经过加工后就成为信息。

2. 信息

信息（Information）是指对数据进行收集、管理以及分析的结果，是经过一系列的提炼、加工和集成后的数据。信息与数据既有联系，又有区别。数据是信息的表现形式和载体，可以是符号、文字、数字、语音、图像、视频等。而信息是数据的内涵，信息加载于数据之上，解释数据的含义。数据和信息是不可分离的，信息依赖数据来表达，数据则生动具体地表达出信息。数据是符号，是物理性的；信息是对数据进行加工处理之后所得到的并对决策产生影响的数据，是逻辑性和观念性的。数据是信息的表现形式，信息是数据有意义的表示。数据是信息的表达、载体，信息是数据的内涵，它们是形与质的关系。数据本身没有意义，数据只有对实体行为产生影响时才成为信息。

3. 知识

知识（Knowledge）是人类对物质世界以及精神世界探索的结果总和。知识也是人类在实践中认识客观世界（包括人类自身）的成果，它包括事实、信息的描述或在教育和实践中获得的技能。知识是人类从各个途径中获得的经过提升总结与凝练的系统的认识。

4. 三者之间的关系

数据、信息、知识这三者是依次递进的关系，代表着人们认知的转化过程。数据指的是未经加工的原始素材，表示的是客观的事物。而通过对大量的数据进行分析，人们可以从中提取出信息。人们有了大量信息后，会对信息再进行总结归纳，将其体系化，就形成了知识。数据、信息、知识的层次关系如图 1-2 所示。

数据是宝贵的财富，只有充分有效地利用这种财富，识别信息，获取知识，辅助商务决策，才能从中获取价值。数据、信息、知识和决策之间的关系如图 1-3 所示。

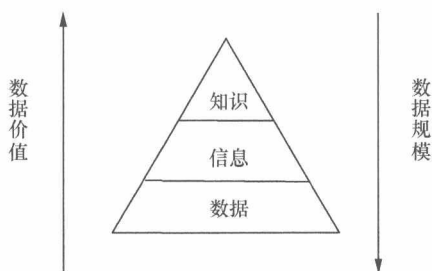


图 1-2 数据、信息、知识的层级关系

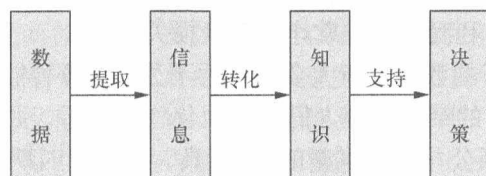


图 1-3 数据、信息、知识和决策之间的关系

1.2.2 商务智能的发展

商务智能作为企业信息化的高端产品，其发展既依赖于相关技术的进步，又依赖于竞争日益激烈的市场环境下企业对商务智能的认知。

1865年，理查德·米勒·德文斯（Richard Millar Devens）在《商业趣闻百科全书》（*Cyclopædia of Commercial and Business Anecdotes*）中提出了“商务智能”（Business Intelligence）一词。他用这个词来描述银行家亨利·福尼斯（Henry Furnese）通过收集信息并根据这些信息，先于竞争对手采取行动，从而获利。可以看出，当时的“商务智能”概念与现在的概念大相径庭。

1958年，汉斯·彼得·卢恩（Hans Peter Luhn）撰文讨论了利用技术来收集商务智能的潜力。按照今天的理解，商务智能就是利用技术来收集和分析数据，将之转换成有用的信息，并根据这些信息，“先于竞争对手”采取行动。从本质上说，现代版的商务智能是利用技术，在正确的时间，依据正确的信息，迅速且有效地做出决策。

直到1968年，具备专业技能的人员才可以把数据转换成可用的信息。当时多个来源的数据通常存储在筒仓中，研究报告呈碎片化，彼此脱节，可以做出多种不同的解读，只有那些具备专业技能的人，才能把数据转换成可用的信息。埃德加·科德（Edgar Codd）认识到这是个严重的问题。1970年，他发表文章，改变了人们思考数据库的方式。他的“关联式数据库模型”引起了广泛关注，被全世界所采纳。

之后出现的决策支持系统（Decision Support System, DSS）是第一个数据库管理系统。很多学者都认为，现代版的商务智能就是从DSS数据库演化而来。

20世纪70年代末，企业管理者开始使用互联网来收集商业信息。高层管理信息系统（Executive Information Systems, EIS）由此诞生，为企业高管提供决策方面的支持。EIS旨在提供“简化”决策过程所需的、最新的、准确的信息，强调以图表和易用界面的方式呈现这些信息。EIS的目标是把企业高管变成“亲自动手”的用户，让他们自己处理邮件、进行研究、做出任命和阅读报告，而不是通过中间人接收这些信息。但由于

其作用有限，EIS 渐渐退出市场。

20 世纪 80 年代，企业开始使用内部数据分析解决方案（由于受到当时计算机系统的限制，这通常是在下班后和周末进行），数据仓库开始流行。在数据仓库出现之前，企业需要大量的数据冗余，以向参与决策的所有人提供有用的信息。数据仓库把以往通常存储在多个地方的数据，存储在了同一个地方，从而大幅缩短了使用者访问数据所需的时间。

数据仓库还有助于推动大数据的应用。数量庞大、形式多样的数据（电子邮件、互联网、Facebook、Twitter 等）可以从同一个地方访问，这不但节约了时间和资金，而且还能访问以前访问不了的许多商业信息。在提供由数据驱动的知识方面，数据仓库潜力巨大。这些知识可以提高利润、发现欺诈、减少损失。

1988 年，在罗马举行的数据分析大会结束后不久，商务智能开始作为一个技术概念出现。在这场大会上得出的结论促使人们开始简化 BI 分析，并使之对用户更加友好。BI 企业大量涌现，多家新公司都提供新的 BI 工具。在那个时期，BI 有两项基本功能：产生数据和提供报告，并以适当的方式组织和呈现数据。

从 20 世纪末到 21 世纪初，BI 服务开始提供简化的工具，降低决策者对工具的依赖度。这些工具更易于使用，而且可以为决策者提供他们所需且有效的功能。商业人士可以通过直接与数据打交道的方式，收集数据，获取知识。

1.2.3 商务智能的要求

商务智能的主要目的是帮助企业决策者进行更有效的决策。据调查，有些决策者担忧的问题已经持续多年，到目前为止仍没有被解决，主要内容如下。

- “我们收集了大量数据，但我们无法访问它。”
- “我们需要把数据按每一种方式分片和切割。”
- “商界人士需要轻松获取数据。”
- “告诉我什么是重要的。”
- “我们整个会议都在争论谁拥有正确的数字，而不是做决定。”
- “我们希望人们利用信息来支持基于事实的决策。”

这些问题反映了商务智能的基本要求。

(1) 商务智能的信息必须易于访问，且内容也必须是可理解的。数据对业务用户来说应该是直观和明显的，而不仅仅只考虑开发人员的需求。数据的结构和标签以及词汇应该模仿业务用户的思维过程。业务用户通常会希望将分析数据分离并进行无限制的组合，访问数据的业务智能工具和应用程序必须简单易用，且必须以尽可能少的等待时间将查询结果返回给用户。

(2) 商务智能必须能提供可信信息。商务智能要求数据是可信的，各种来源的数据必须仔细地收集、清洗，以保证数据质量，而且当用户需要时才能发布。

(3) 商务智能必须适应变化。用户需求、业务条件、数据和技术都可能发生变化，因此商务智能系统设计必须将这一因素考虑在内，确保现有的数据或应用程序不会失效。当业务区向数据仓库提交新问题和新数据时，不应该更改或中断现有的数据和程序。如果系统必须要进行修改，则要求设计者应进行适当的解释，使这些更改对用户透明。

(4) 商务智能必须及时提供信息。当商务智能用于操作决策时，原始数据可能需要

在几分钟甚至几秒的时间内转化为有用的信息。

(5) 商务智能必须保证信息资产的安全。例如，一个珠宝销售企业的销售订单存放在数据仓库中，数据仓库中保存了企业及其客户的信息，如果这些信息落在某些别有用心的人的手中，可能会对企业和客户产生不利的影响。因此商务智能必须有效地控制企业机密信息的访问路径。

(6) 商务智能必须帮助企业提高决策的权威性和可信赖程度。数据仓库必须有正确的数据来支持决策。商务智能最重要的输出是基于所呈现的分析数据做出的决策，这些决策提供了只有商务智能才可以带来的巨大的业务影响和价值。

(7) 企业要积极接受商务智能，并认为其可以有效地协助企业决策。使用好的产品 and 平台构建一个解决方案并不是最重要的，如果业务用户不接受、不使用商务智能，那么即使再好的解决方案也没有用武之地。商务智能和业务软件不同，用户必须使用业务软件，别无选择，而用户可以选择是否使用商务智能。

1.2.4 商务智能的价值

商务智能之所以越来越重要，是因为现代企业对各种信息的无知是可怕的，不知不觉的风险是巨大的，而一知半解可能比一无所知危害更大，因为一知半解可能会使企业带着错误的思想做出决定和采取行动。商务智能所要做的就是充分利用企业在日常经营中积累的大量数据，并将它们转化为信息和知识来避免企业的错误决策。

商务智能在商业决策中的作用和价值主要体现在以下几个方面。

1. 改善客户关系管理

企业正在逐渐由以产品为中心转换为“以客户为中心”，商务智能可利用客户关系管理(CRM)结合联机分析处理和数据挖掘等技术，处理大量的交易记录和相关客户资料，对顾客进行分类，然后针对不同的客户制定相应的服务策略。例如，金融行业利用商务智能把客户数据整合后进行分析，帮助企业减少客户流失。

商务智能为客户、员工、供应商、股东和大众提供关于企业及其业务状况的有用信息，从而提高企业的知名度，增强整个信息链的一致性。利用商务智能，企业可以在问题变成危机前很快地对它们加以识别并解决。商务智能也有助于加强客户的忠诚度，一个参与其中并掌握充分信息的客户更有可能购买你的产品和服务。

2. 提供可赢利性分析

商务智能可以得到精准、及时的信息，帮助企业获得竞争优势。商务智能解决方案可以帮助企业分析利润的来源、各类产品对利润总额的贡献程度、广告费用是否与销售成正比等。商务智能可通过详细的目标数据，协助企业制定降低成本的决策。美国有许多保险、租赁和金融服务公司都已经体会到了商务智能带来的好处。

3. 改善业务洞察力

商务智能可减少管理者收集数据、获取信息所花费的时间，加速决策过程，保证在正确的时间让决策者可以接收到正确的信息。商务智能可以用来帮助企业理解业务的推动力量，认识哪些是趋势、哪些是非正常情况和哪些行为正在对业务产生影响，及时调整策略。例如，兴业证券利用 SCAMA 把数据整合后进行分析，辅助企业高层进行企业关键业绩指标分析、竞争对手分析以及投资收益分析等，增强了企业竞争力，取得了很好的效果。