

数据库应用基础

(第三版)

刘 畅 等 编著



上海交通大学出版社
SHANGHAI JIAO TONG UNIVERSITY PRESS

数据库应用基础

(第三版)

刘 畅 等 编著

上海交通大学出版社

内 容 提 要

本书以培养基于数据库技术的数据管理和应用能力为目的,以 Access 2007 为平台,系统地介绍了数据库技术的发展及应用知识。本书的主要内容包括:数据库技术的基本概念、数据建模和数据分析的理论与方法;数据库各种对象的基本知识以及基本操作;多种数据交换技术;数据库应用的基本方法和常用技术;数据库安全、维护、配置等常用知识。本书所介绍的内容并非仅适用于 Access 2007 环境,其中主要的概念和方法适用于各类关系数据库,有关 Access 的主要知识也同样适用于 Access 的各个流行版本。

本书内容涵盖面广、结构合理、深入浅出、循序渐进。并且示例丰富、解释清晰,具有较强的启发性和实用性。本教材适合作为非计算机类专业学生的数据库课程教材使用,也可以作为数据库应用系统开发人员的数据库技术参考书籍。

图书在版编目(CIP)数据

数据库应用基础/刘畅等编著. —3 版. —上海:上海
交通大学出版社,2011(2012 重印)

ISBN 978-7-313-05136-3

I. 数... II. 刘... III. 数据库系统 - 高等学
校 - 教材 IV. TP311.13

中国版本图书馆 CIP 数据核字(2011)第 001852 号

数据库应用基础

(第三版)

刘 畅 等编著

上海交通大学出版社出版发行

(上海市番禺路 951 号 邮政编码 200030)

电话:64071208 出版人:韩建民

上海交大印务有限公司印刷 全国新华书店经销

开本:787mm×1092mm 1/16 印张:18.25 字数:450 千字

2008 年 2 月第 1 版 2011 年 1 月第 3 版 2012 年 2 月第 5 次印刷

印数:2030

ISBN 978-7-313-05136-3/TP 定价:30.00 元

版权所有 侵权必究

前 言

数据是管理的基本资源,信息产生于对数据的有效管理及合理利用。数据库技术是当今计算机应用的重要技术之一,其在信息管理领域同样占有重要地位,数据库技术为数据及信息管理提供了强有力的技术支持,数据库相关知识和技能极具实用价值。凡与各种数据和信息资源管理及应用相关的各类专业均需学习和掌握数据库及其应用方面的知识,以适应目前的专业学习需要,并为将来工作中的数据处理和数据管理打下基础。因为,没有丰富的数据管理知识,不掌握有力的数据管理工具,将无法面对蕴藏着丰富信息资源的大量业务数据,无法归纳出崭新的管理理念,无法实现现代的管理模式。与对数据库应用系统开发人员的要求不同的是,管理人员应作为数据库系统的拥有者和使用者来学习数据库知识,以便了解如何从现实业务中提取和组织数据,如何进行数据管理,如何对数据进行加工和处理,如何获取所需的信息,从而使数据和信息在管理和决策中发挥重要作用。同时,掌握由现实业务系统到数据库系统的转换和实现方法,以便在数据库系统的开发和应用过程中,积极参与、严格控制,以确保数据库系统的功能、品质和性能。

本教材不同于一般介绍数据库管理系统使用方法的教材,而是以介绍数据库技术的实际应用为主要目的,重点在于介绍数据库技术的作用及主要功能。本教材取材着重于被广泛采用的较成熟的数据库技术,并吸取了作者们长期从事教学与数据库系统开发的经验体会,从实用的角度出发,较全面、系统地介绍了数据库应用的基本知识和技术。同时,以 Access 2007 为基本环境,通过丰富的示例介绍了各有关技术的实际应用方法。本书所介绍的内容并非仅适用于 Access 2007 环境,其中主要的概念和方法适用于各类关系数据库,有关 Access 的主要知识也同样适用于 Access 的各个流行版本。

本教材共分 17 章,分别介绍了以下内容:数据库概述、Access 概述、数据模型、表、数据的有效性控制、基本数据操作、数据分析、查询基础、选择查询、操作查询和 SQL 特定查询、窗体、报表、宏、模块与 VBA 基础、VBA 程序设计基础、VBA 对 Access 常用对象及数据的访问、数据库安全与维护。各章均附有相应的概念性或操作性习题,以便于同学们练习或检查相关知识的掌握情况。本教材配有完整的示例数据库、实验用数据库、演示文稿等教学资源,以及在局域网环境下的考试系统及题库。本教材适用于各类非计算机专业数据库应用课程教学,也可为数据库应用系统的设计和开发人员提供有益的参考。

本教材为集体编写,参与编写的老师们将自己多年数据库教学经验和实际应用技术凝练于本教材之中。本教材在原版本基础上,参考了多位老师的教学建议和意见,对部分内容进行了调整,并更换了 Access 的版本。在此,对支持本教材建设及使用本教材的老师们表示感谢。

尽管编写中我们为保证内容的合理、正确做了不少努力,但错误和纰漏可能难免,欢迎您的批评指正,我们将虚心接受您的意见和建议。如教学需要,我们将乐于提供相关教学资源。

编 者

2010 年 10 月

目 录

第 1 章 数据库概述	1
1.1 数据与信息	1
1.1.1 数据	1
1.1.2 信息	3
1.1.3 数据与信息的差异	4
1.1.4 数据处理技术	5
1.2 数据库系统概述	8
1.2.1 数据库	9
1.2.2 数据库管理系统	10
1.2.3 数据库系统	11
习题 1	12
第 2 章 Access 概述	13
2.1 Access 的启动与主界面	13
2.1.1 启动 Access	13
2.1.2 Access 的主界面	14
2.2 Access 的主要对象	18
2.2.1 表	18
2.2.2 查询	18
2.2.3 窗体	19
2.2.4 报表	19
2.2.5 宏	20
2.2.6 模块	21
2.3 创建 Access 数据库	21
2.3.1 使用模板创建数据库	21
2.3.2 创建空白数据库	22
2.3.3 关闭数据库	22
习题 2	23
第 3 章 数据模型	24
3.1 数据模型	24
3.1.1 数据模型的基本概念	24
3.1.2 数据模型的主要元素	25

3.1.3 弱实体	34
3.2 数据建模过程	35
3.2.1 获取实体	35
3.2.2 上下文数据模型	36
3.2.3 基于键的数据模型	37
3.2.4 具有完整属性的数据模型	37
3.3 评价数据模型	39
习题 3	39
第 4 章 表	41
4.1 表结构及字段	41
4.1.1 表结构	41
4.1.2 字段	42
4.2 建表	42
4.2.1 基本转换规则	42
4.2.2 表的视图	43
4.2.3 创建表的基本操作	44
4.3 表的基本操作	50
4.3.1 重命名、删除	50
4.3.2 剪切、复制、粘贴	50
4.3.3 表属性	51
4.3.4 打印	51
习题 4	51
第 5 章 数据的有效性控制	53
5.1 输入掩码	53
5.2 字段有效性规则	54
5.3 记录(表)有效性规则	56
5.4 参照完整性	57
5.4.1 建立表之间的关系	58
5.4.2 参照完整性的实施情况	60
5.5 表的查阅列	61
5.5.1 创建行来源类型为值列表的查阅列	61
5.5.2 创建行来源类型为表/查询的查阅列	62
5.5.3 创建多值的查阅列	64
习题 5	66
第 6 章 基本数据操作	67
6.1 添加记录	67

6.2 浏览数据	69
6.3 修改数据	72
6.4 删除记录	72
6.5 数据透视表和数据透视图	73
6.5.1 数据透视表	73
6.5.2 数据透视图	75
6.6 导入、导出	77
6.6.1 导入	77
6.6.2 导出	79
习题 6	82
第 7 章 数据分析	83
7.1 规范化	83
7.2 依赖	84
7.2.1 函数依赖	84
7.2.2 部分依赖	86
7.2.3 传递依赖	88
7.3 冗余与更新异常	89
7.3.1 冗余	89
7.3.2 更新异常	90
7.4 关系的分类	91
7.4.1 第一范式	92
7.4.2 第二范式	93
7.4.3 第三范式	94
7.4.4 Boyce-Codd 范式	95
7.5 反规范化	96
习题 7	97
第 8 章 查询基础	99
8.1 查询语言	99
8.2 查询简介	100
8.2.1 查询分类	100
8.2.2 查询创建方式	102
8.3 简单的选择查询	104
8.3.1 SELECT 语句的基本格式	104
8.3.2 查询条件	106
8.3.3 查询条件示例	108
8.3.4 控制结果集中数据行的出现情况	112
习题 8	113

第 9 章 选择查询	115
9.1 聚合函数	115
9.1.1 Count	115
9.1.2 Sum	116
9.1.3 Avg	116
9.1.4 Max、Min	117
9.2 分组查询	117
9.2.1 分组查询示例	117
9.2.2 筛选分组查询的结果	119
9.3 联接查询	119
9.3.1 内部联接	120
9.3.2 外部联接	121
9.3.3 自联接	123
9.4 子查询	123
9.4.1 利用子查询组织查询条件	124
9.4.2 利用子查询生成结果列	126
习题 9	127
第 10 章 操作查询和 SQL 特定查询	130
10.1 操作查询.....	130
10.1.1 生成表查询.....	130
10.1.2 追加查询.....	131
10.1.3 更新查询.....	132
10.1.4 删除查询.....	133
10.2 SQL 特定查询	134
10.2.1 联合查询.....	134
10.2.2 数据定义查询.....	135
10.2.3 交叉表查询.....	138
习题 10	140
第 11 章 窗体	142
11.1 窗体概述.....	142
11.1.1 窗体的功能.....	142
11.1.2 窗体的视图.....	143
11.1.3 打印窗体.....	144
11.2 创建窗体.....	145
11.2.1 自动创建窗体.....	145
11.2.2 利用窗体向导创建窗体.....	147

11.2.3 利用窗体设计视图创建窗体·····	148
11.3 窗体设计·····	150
11.3.1 窗体的节·····	150
11.3.2 建立数据联系·····	151
11.3.3 窗体控件·····	152
11.3.4 窗体中的计算·····	161
11.3.5 窗体的常用属性·····	163
11.4 其他窗体·····	164
11.4.1 数据透视表窗体与数据透视图窗体·····	165
11.4.2 在窗体中插入图表·····	166
11.4.3 设计主/子窗体·····	167
习题 11·····	168
第 12 章 报表 ·····	169
12.1 报表概述·····	169
12.1.1 报表的功能·····	169
12.1.2 报表的类型·····	170
12.1.3 报表的视图·····	170
12.2 创建报表·····	171
12.2.1 自动创建报表·····	171
12.2.2 利用报表向导创建报表·····	172
12.2.3 使用标签向导创建标签·····	173
12.2.4 利用报表设计视图创建报表·····	175
12.3 报表设计·····	176
12.3.1 报表的节·····	176
12.3.2 建立报表数据源·····	177
12.3.3 报表常用控件·····	178
12.3.4 报表的常用属性·····	180
12.3.5 节的常用属性·····	181
12.4 报表设计的其他问题·····	182
12.4.1 报表的排序·····	182
12.4.2 分组与汇总·····	183
12.4.3 含有子报表的报表·····	186
12.4.4 多列报表·····	188
习题 12·····	189
第 13 章 宏 ·····	190
13.1 宏的基本概念·····	190
13.1.1 宏的作用·····	190

13.1.2 宏的存在形式	194
13.2 创建宏	194
13.2.1 宏生成器	194
13.2.2 创建独立的宏	196
13.2.3 创建宏组	197
13.2.4 创建嵌入的宏	197
13.3 运行宏	198
13.3.1 直接运行宏	198
13.3.2 运行宏组中的宏	198
13.3.3 从另一个宏或者从 VBA 过程中运行宏	199
13.3.4 为响应窗体、报表或控件中的事件而运行宏	199
13.3.5 数据库系统启动时运行宏	199
13.4 宏的示例	199
13.4.1 利用宏实现查找	200
13.4.2 在报表中嵌入宏	201
13.4.3 创建启动窗体	201
习题 13	204
第 14 章 模块与 VBA 基础	205
14.1 VBA 程序设计基本要素	205
14.1.1 程序与程序设计	205
14.1.2 面向对象的程序设计	206
14.1.3 VBA 的过程	207
14.2 模块	208
14.2.1 类模块	208
14.2.2 标准模块	209
14.2.3 创建和编辑模块	209
14.2.4 Visual Basic 编辑环境	210
14.2.5 模块调试、编译	212
14.3 VBA 数据类型与表达式	212
14.3.1 数据类型	213
14.3.2 变量	215
14.3.3 符号常量	217
14.3.4 函数	217
14.3.5 表达式	221
习题 14	224
第 15 章 VBA 程序设计基础	226
15.1 顺序结构程序设计	226

15.1.1 注释语句	226
15.1.2 赋值语句	227
15.1.3 输入对话框(InputBox)	228
15.1.4 消息对话框(MsgBox)	228
15.2 分支结构程序设计	230
15.2.1 IF 语句	230
15.2.2 Select Case 语句	233
15.3 循环结构程序设计	234
15.3.1 Do...Loop 循环	235
15.3.2 For...Next 循环	236
15.4 数组	237
15.4.1 数组的定义	238
15.4.2 数组的访问	239
15.5 过程	240
15.5.1 过程定义	240
15.5.2 过程的调用	241
15.5.3 参数传递	242
15.6 VBA 代码设计示例	243
15.6.1 统计不同类别的字符个数	244
15.6.2 计算个人所得税	245
习题 15	247
第 16 章 VBA 对 Access 常用对象及数据的访问	249
16.1 Access 内置对象及引用	249
16.1.1 Application 对象	249
16.1.2 Forms、Reports、Moudules 集合	251
16.1.3 AccessObject、CurrentProject、CurrentData 对象	252
16.1.4 DoCmd 对象	253
16.1.5 示例:浏览数据库对象	256
16.2 ADO 数据对象	258
16.2.1 ADO 的作用	258
16.2.2 设置对 ADO 对象库的引用	259
16.2.3 ADO 对象模型	259
16.2.4 利用 ADO 对象访问数据库	260
16.2.5 Recordset 对象的常用操作	263
16.2.6 示例:利用 ADO 对象查找和编辑数据记录	266
习题 16	270
第 17 章 数据库安全与维护	271
17.1 数据库的加密和解密	271

17.1.1	加密数据库.....	271
17.1.2	解密数据库.....	272
17.2	数据库管理.....	272
17.2.1	压缩和修复数据库.....	272
17.2.2	备份和恢复数据库.....	273
17.2.3	为数据库建立文档.....	273
17.3	系统选项.....	275
17.3.1	常用选项.....	275
17.3.2	与信任中心相关的选项.....	277
习题 17		280

第 1 章 数据库概述

伴随着人类的生产、生活以及各种活动,数据源源不断地产生,数据是对客观实体属性的记录,是获取信息的基本材料。面对大量的数据,如何才能有效地加以利用,一直是人们尤其是管理人员所关注的问题。经过了长期不断的探索,新的数据处理技术陆续被人们创建和采用,并且取得了显著的成效。20 世纪 60 年代诞生的数据库技术日臻成熟,其主要目的是有效地组织、管理和存取大量的数据资源。数据库技术使数据有了统一的结构,可对数据实行统一、集中、独立的管理,从而实现了数据的共享,保证了数据的完整性和安全性,提高了数据管理效率。各种数据库管理系统软件被人们熟悉和使用,成为数据和信息处理工作不可或缺的重要工具。

本章中,将从基本知识出发,介绍数据和信息的概念及主要特性;介绍数据处理(或信息加工)的基本内容以及数据处理技术发展的主要过程;介绍数据库与数据库管理系统方面的基本概念和基础知识。以便使同学们在学习数据库应用技术之前,对数据、数据库及其主要应用有一个较系统、较全面的了解。

1.1 数据与信息

在当今这个信息时代,“数据”和“信息”这两个词大概已经进入使用频率最高词汇的排行榜了,它们经常被人们使用,随处可以听到。那么到底什么叫数据? 什么叫信息? 数据就是通常所说的数字吗? 信息就是消息吗? 数据和信息之间有什么联系? 又有什么区别? 本节将讨论这些问题。

1.1.1 数据

数据,人们从不同的角度看待和利用它时可以有不同的理解和定义,比如:

- 数据是能够由计算机处理的数字、字母和符号;
- 数据是对事实、概念或指令的一种特殊表达形式,它可以用人工或自动化装置进行通讯、翻译或处理;
- 数据是对客观实体属性的记录;

.....

从数据处理和信息管理这个角度看,最后一个定义最能体现数据的本质,因此,我们采用的数据定义为:数据是对客观实体属性的记录。

这里所说的客观实体就是现实世界中的事物。可以是具体的事物,比如,学生,商人,汽车等;也可以是人为的概念,比如,学习,贷款,募捐等。属性是实体的描述性性质或特征,是对实体具有相对重要意义的某一方面情况的描述结果。例如,从学生学籍管理的角度来看,学生是实体,所关注的有关学生的特征是:学号、姓名、性别、专业、班级,等;课程是实体,所关注的特征有:课程名称、开课学院、课程性质、课时,等。

根据上述定义,数据包括了两个方面的含义:

第一,数据是客观实体属性的反映。当数据作为客观实体属性的反映时,它是属性名和属性值的统一体。例如,有一张学生登记表(如表 1-1)。

表 1-1 学生登记表

班 级	学 号	姓 名	性 别	出生日期	入学成绩
071 财政 1	2007031011	张宝文	男	1989-10-26	577
071 国贸 4	2007061246	赵思敏	女	1989-06-03	582
071 经管 2	2007151048	杨 锋	男	1989-08-17	573
...	...	...	...	...	...

表中每一个学生都是一个客观实体,而班级、学号、姓名、性别、出生日期和入学成绩等都是属性名。071 财政 1、2000031011、张宝文、男、1989-10-26、577 等都是属性值。

如果数据离开了属性名,也就失去了其所反映的客观实体属性,因而使属性值失去了意义;如果数据没有属性值,则不能表明客观实体的具体特性,也就失去了作为数据的内容。例如,属性名“入学成绩”和属性值“577”构成了一项数据。如果只有“577”而没有“入学成绩”或者相反,都是没有意义的。

第二,数据具有载体和表现形式。数据作为客观实体属性的记录,必须要有一定的载体。比如,使用纸张来记录学生成绩,利用磁带来记录音乐,利用胶卷来记录风景,利用光盘来刻录班级联欢实况,等等。这里提到的纸张、磁带、胶卷、光盘就是用以记录数据的载体。载体是在数据传输中携带数据的媒介,是数据赖以附载的物质基础。包括以能源和介质为特征,运用声波、光波、电波传递数据的无形载体,以及以实物形态记录为特征,运用纸张、胶卷、胶片、磁带、磁盘传递和贮存数据的有形载体。

数据在表现形式上也是多种多样的。属性名一般用文字符号表示,属性值既可以用数字表示,也可以用文字、符号、图形、图像和声音等非数字的形式表示。数据的表现形式虽然多种多样,但从计算机处理的角度,可以把数据划分为数值数据和非数值数据两个大类,每个大类又可以根据需要继续划分出许多具体的数据类型。

由上述讨论可知,数据是描述客观实体特征的特定的符号,是人们传达思想,进行信息交流的媒介。然而,由于对客观实体的认识存在差异,以及记录客观实体的目的不同,人们对客观实体属性的记录也会有所不同。例如,姓名、性别、年龄、民族、身高、体重、血型、血压、政治面貌、文化程度、大学语文成绩、数学成绩、数据库系统应用成绩……等等,都是“人”这个客观实体的属性。但是,医院中对“病人”(客观实体)的记录,可能只会记录姓名、性别、年龄、身高、体重、血型、血压……等属性,不会关心文化程度、大学语文成绩、数学成绩等属性;而大学对于“学生”(客观实体)的记录,只会记录姓名、性别、年龄、民族、政治面貌、大学语文成绩、数学成绩、数据库系统应用成绩等属性,不会关心身高、体重、血型、血压等属性。从现实世界的客观事物的状况到信息管理系统中的数据,需要经历一系列重要的转变过程。这个过程涉及到人对数据的理解和认识,也涉及到获取和加工数据的手段和技术。一般来说,应该根据具体应用的需要,从客观事物出发,经过分析、归纳,提取出相关事物的重要属性,然后应用一些规则使之规范化,最终形成符合应用需要的数据。

1.1.2 信息

信息作为一种客观存在,在史前的猎、牧部族之中,在农业社会和工业社会中,尤其是在当今的文明社会中,它一直都在积极地发挥着人类意识或没有意识到的重大作用。科学技术的进步,使人们终于认识到,信息是与物质和能量可以相提并论的用以维系人类社会存在及发展的三大要素之一。因此,只有科学地了解和认识信息的基本作用,才能更好地把握信息,进而使信息更好地为人类社会服务。

人们都说当今社会是信息社会,当今时代是信息时代,信息量大,信息爆炸等等,可谓言必称信息。可是,什么是信息呢?没有人能够完全说清楚,因为信息定义本身也深深陷入了信息爆炸之中,信息界各家探索者都要给信息下个自己的定义。据统计,目前有关信息的定义有几十种之多,例如:

- 信息就是信息,既不是物质也不是能量;
- 信息是收信者事先所不知道的报道;
- 信息就是消息;
- 信息就是数据;
- 信息就是知识;
- 信息是人与外界相互作用的过程中所交换的内容的名称;
- 信息是经过加工后的数据,它对接收者的行为能产生影响,对接收者的决策具有价值;
- 信息是事物之间相互联系、相互作用的状态的描述;

.....

信息的各种定义都有一定的道理,其实事情往往就是这样,简单的事物往往是最伟大的,人们好像都知道它是什么,对其熟视无睹,但细细考究起来却又无法给出一个完整的定义,信息就是如此。

这里不去评价各种信息定义的特点与实质,暂且以“信息是事物之间相互联系、相互作用的状态的描述”作为对信息的定义,并以此为基础来讨论信息的有关特性。

由此定义可知,只有当事物之间相互联系、相互作用时,才会有信息。一个事物由于另一个事物的影响而使其某种属性发生了变化,从信息的观点看,是因为前者得到了后者的某种信息。从信息的观点出发,可以把相互联系、相互作用的事物有目的的发展变化看作信息的采集(获取)、传输、存储、加工、变换的过程。任何事物的发展变化,都既受其他事物的影响,又影响其他事物。也就是说,既接受来自其他事物的信息,又向其他事物发送信息。

通常把信息发生者称为信源,信息的接收者称为信宿,传播信息的媒介称为载体,信源和信宿之间信息交换的途径与设备称为信道(通道)。一般情况下,信息接收者(信宿)接收到来自信息发送者(信源)的信息后,会进行处理(分析、评价、决策等),并根据处理后的信息付诸行动,而这种行动反过来又影响信息发送者,这种影响称为信息反馈。信息从发送者传到接收者,经过接收、处理、行动等各个环节又反馈到信息发送者,形成一个信息运动的循环,称之为信息循环(如图 1-1 所示)。可以看出,信源和信宿是相对的。

概括起来,信息的基本作用主要体现在以下几个方面。

(1) 信息是重要的资源。

信息对经济发展、社会进步起着巨大作用。信息的占有水平与利用程度,成为衡量一个国

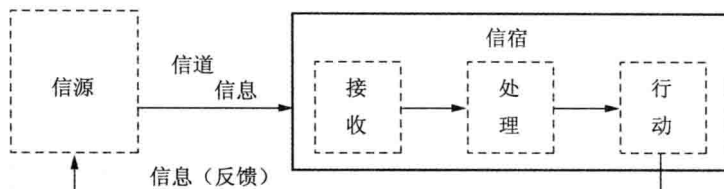


图 1-1 信息循环

家现代化水平的重要标志,是国家综合实力的重要组成部分。管理者管理的主要资源有:人力资源、原材料资源、机器资源(包括设备和能源)、资金资源和信息资源五类。其中,前四种资源都是有形的,一般称之为物质资源;而信息资源是无形的,是由于它所描述的对象而具有价值,为此称之为概念资源。管理者是利用概念资源来管理物质资源的。

(2) 信息是决策的基础。

现代管理的核心是决策。正确的决策取决于多种因素,如决策体制、决策方法、领导者的能力、技巧等。但决定性的因素,是对客观实际,及世界与未来行动及其后果的正确判断。信息系统可以帮助企业领导迅速取得和分析有关决策的信息,有助于企业降低决策中的不确定性和风险。如果信息不充分,决策就失去了根本的依据,信息不灵就可能导致决策的失败。要提高企业的竞争能力,要实现正确的决策,就必须拥有大量的信息。正所谓“知己知彼,百战不殆”。

(3) 信息是实施管理控制的依据。

管理过程就是信息的收集、传递、加工、判断、决策的过程。企业系统的全部活动可概括为两大类:生产活动和管理活动,围绕着生产活动,执行着决策、计划和指挥职能。生产活动中流动的是物,而管理活动中流动的是信息。物流是生产经营活动的主体流动,信息流伴随着物流产生,反映物流的状况。管理人员通过信息流了解物流的运动,同时通过信息流计划和调节物流。因此,在经营管理中信息流起着主导作用。

(4) 信息是系统内外联系的纽带。

所谓系统,是指由处于一定环境中的若干个具有相对独立功能的部分组成的集合,系统的各个组成部分之间相互联系、相互影响,并为共同完成系统的整体目标而运行。为将系统各组成部分联结为一个整体,使各组成部分能够相互协调,就要将它们之间的信息流很好地组织起来,形成系统内各组成部分间联系的纽带。此外,系统总是要与所在环境中的其他系统进行一定的沟通、交流。例如发布新产品、回答客户询问、报送财务报表,等等。此类沟通和交流都是利用系统对外部的信息流来完成的。因此,信息既是系统内部也是系统之间进行联系的纽带。

1.1.3 数据与信息差异

一般认为,可以从如下几个方面来区分数据和信息这两个概念。

(1) 数据是信息的表现形式,信息是数据的含义。

信息从信源传输到信宿,需要借助于载体,载体以某种特殊形式的变化和运动来反映信息的内容,并使接收者可以感知。在信息处理中,信息载体上表现的接收者(人或机器)可以识别的符号(或信号)就是数据,而这些符号(或信号)反映的意义则是信息。比如,中央电视台播报的天气预报,在卫星、微波、光缆、电缆等媒介上传播时,各级接收者接收、处理、转发的都是数

据,而观众听到的“丽江明天午后有雷阵雨”则是信息。

注意,数据并不限于数字,文字、声音、图像、光信号、电信号、磁信号等都是数据的表现形式。信息与载体的性质无关,而数据的具体形式却取决于载体的性质。

(2) 数据是信息的原材料,信息是数据处理(加工)后的结果。

比如,超市 POS 机采集到的鲜牛奶销售量 1500 箱,经过比较发现比去年同期增长了 10%,这里“销售量 1500 箱”是数据,“比去年同期增长了 10%”则是信息。

必须指出,信息处理有一个不断深入的过程。数据经过一级处理变成了信息,而这个信息可能又作为数据被继续处理形成新的信息。这一过程可见图 1-2。

所以,按此标准区分,数据和信息是个相对的概念。同一个内容,在此处是信息,而在彼处可能又是待处理的数据了。比如,一个城市的财政总收入对于本市来说是信息,而对于该市所在的省份来说,这只是待处理的数据。

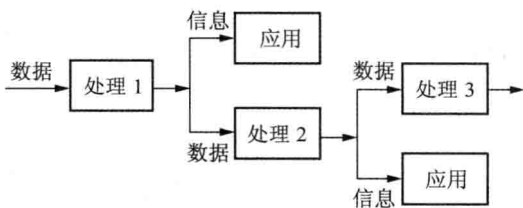


图 1-2 信息处理的层次

(3) 信息对决策有价值,而数据则没有。

信息必定有人的参与,必定包含在人的决策活动中。决策活动是信息存在的必要条件,这个属性可以很好地区分数据和信息。例如,“丽江明天午后有雷阵雨”,这条消息对于丽江地区的人和准备明天到丽江地区的人来说是信息,因为对他(她)们自身的决策有价值。但是对于处于外地或与丽江这一地区无关的人来说,没有决策价值,因此不是信息。

以上给出了区分数据与信息几条基本依据,实际上,数据和信息的辨别还取决于语义环境。例如,一个职工的工资实发数对其个人来说是信息,但是对代办工资的银行来说就是数据了。

由于数据和信息“形影不离”,人们在许多场合并不刻意区分数据和信息这两个概念,常常把信息处理也叫做数据处理。在不影响对问题理解的情况下,可以把“信息”和“数据”这两个术语不加区别地使用。

1.1.4 数据处理技术

由于在人们的劳动、生产、生活等日常活动过程中产生了大量的数据和信息,而这些数据和信息如果不能很好地收集和处理就不会产生任何价值的。因此,随着人们对数据重要意义的认识不断提高,数据的处理技术也逐步地得到发展、丰富和完善。本节将介绍一些与数据处理技术相关的知识。

1. 数据处理

数据处理的基本目的是从大量的、可能是杂乱无章的、难以理解的数据中抽取并推导出对于某些特定的人们来说是有价值、有意义的信息。所以,数据处理是将数据加工处理为信息的工作,该工作一般包括数据的采集、校验、传输、加工、存储和传输等多个环节。

(1) 数据采集。

数据采集是数据处理过程的起点。数据采集有两个方面的含义,一是指收集有关数据,二是指将数据输入到数据处理系统。

数据的收集,应该充分了解系统的需求,制定出合理的方案。由于事物的属性是多种多样