



全国中医药行业高等教育“十三五”规划教材



全国高等中医药院校规划教材（第十版）

# 医药数据库 系统原理与应用

（供中药学、药学、管理学、计算机科学与技术等专业用）

主 编 杜建强 胡孔法

全国百佳图书出版单位  
中国中医药出版社

全国中医药行业高等教育“十三五”规划教材

全国高等中医药院校规划教材（第十版）

# 医药数据库系统原理与应用

（供中药学、药学、管理学、计算机科学与技术等专业用）

## 主 编

杜建强（江西中医药大学）

胡孔法（南京中医药大学）

## 副主编（以姓氏笔画为序）

米 鹏（山东中医药大学）

孙艳秋（辽宁中医药大学）

杜 清（北京中医药大学）

李志敏（浙江中医药大学）

罗 维（成都中医药大学）

## 编 委（以姓氏笔画为序）

曲 雷（天津中医药大学）

任 真（甘肃中医药大学）

孙扬波（湖北中医药大学）

李 艳（山西中医药大学）

杨丽琴（上海中医药大学）

余侃侃（南京中医药大学）

姜 姗（河南中医药大学）

殷云霞（安徽中医药大学）

崔茂应（云南中医学院）

程春雷（江西中医药大学）

中国中医药出版社

· 北 京 ·

图书在版编目 (CIP) 数据

医药数据库系统原理与应用/杜建强, 胡孔法主编. —北京: 中国中医药出版社, 2017. 12

全国中医药行业高等教育“十三五”规划教材

ISBN 978 - 7 - 5132 - 4548 - 7

I. ①医… II. ①杜… ②胡… III. ①数据库管理系统—应用—医药学—中医学院—教材 IV. ①R319

中国版本图书馆 CIP 数据核字 (2017) 第 255730 号

中国中医药出版社出版

北京市朝阳区北三环东路 28 号易亨大厦 16 层

邮政编码 100013

传真 010 - 64405750

保定市西城胶印有限公司印刷

各地新华书店经销

开本 850 × 1168 1/16 印张 16.75 字数 428 千字

2017 年 12 月第 1 版 2017 年 12 月第 1 次印刷

书号 ISBN 978 - 7 - 5132 - 4548 - 7

定价 49.00 元

网址 [www.cptcm.com](http://www.cptcm.com)

社长热线 010 - 64405720

购书热线 010 - 89535836

维权打假 010 - 64405753

微信服务号 zgzyycbs

微商城网址 <https://kdt.im/LIdUGr>

官方微博 <http://e.weibo.com/cptcm>

天猫旗舰店网址 <https://zgzyycbs.tmall.com>

如有印装质量问题请与本社出版部调换 (010 - 64405510)

版权专有 侵权必究

全国中医药行业高等教育“十三五”规划教材

全国高等中医药院校规划教材（第十版）

## 专家指导委员会

### 名誉主任委员

王**国强**（国家卫生计生委副主任 国家中医药管理局局长）

### 主任委员

王**志勇**（国家中医药管理局副局长）

### 副主任委员

王**永炎**（中国中医科学院名誉院长 中国工程院院士）

张**伯礼**（教育部高等学校中医学类专业教学指导委员会主任委员  
天津中医药大学校长）

卢**国慧**（国家中医药管理局人事教育司司长）

### 委 员（以姓氏笔画为序）

马**存根**（大同大学党委书记）

王**键**（安徽中医药大学教授）

王**省良**（广州中医药大学校长）

王**振宇**（国家中医药管理局中医师资格认证中心主任）

方**剑乔**（浙江中医药大学校长）

孔**祥骊**（河北中医学院院长）

石**学敏**（天津中医药大学教授 中国工程院院士）

匡**海学**（教育部高等学校中药学类专业教学指导委员会主任委员  
黑龙江中医药大学教授）

吕**文亮**（湖北中医药大学校长）

刘**力**（陕西中医药大学校长）

刘**振民**（全国中医药高等教育学会顾问 北京中医药大学教授）

安**冬青**（新疆医科大学副校长）

许**二平**（河南中医药大学校长）

孙忠人（黑龙江中医药大学校长）  
严世芸（上海中医药大学教授）  
李占永（中国中医药出版社副总编辑）  
李秀明（中国中医药出版社副社长）  
李金田（甘肃中医药大学校长）  
杨柱（贵阳中医学院院长）  
杨关林（辽宁中医药大学校长）  
余曙光（成都中医药大学校长）  
宋柏林（长春中医药大学校长）  
张欣霞（国家中医药管理局人事教育司师承继教处处长）  
陈可冀（中国中医科学院研究员 中国科学院院士 国医大师）  
陈立典（福建中医药大学党委书记）  
陈明人（江西中医药大学校长）  
武继彪（山东中医药大学校长）  
范吉平（中国中医药出版社社长）  
林超岱（中国中医药出版社副社长）  
周仲瑛（南京中医药大学教授 国医大师）  
周景玉（国家中医药管理局人事教育司综合协调处副处长）  
胡刚（南京中医药大学校长）  
洪净（全国中医药高等教育学会理事长）  
秦裕辉（湖南中医药大学校长）  
徐安龙（北京中医药大学校长）  
徐建光（上海中医药大学校长）  
唐农（广西中医药大学校长）  
彭代银（安徽中医药大学校长）  
路志正（中国中医科学院研究员 国医大师）  
熊磊（云南中医学院院长）

### **秘 书 长**

王键（安徽中医药大学教授）  
卢国慧（国家中医药管理局人事教育司司长）  
范吉平（中国中医药出版社社长）

### **办公室主任**

周景玉（国家中医药管理局人事教育司综合协调处副处长）  
林超岱（中国中医药出版社副社长）  
李秀明（中国中医药出版社副社长）  
李占永（中国中医药出版社副总编辑）

## 编审专家组

### 组 长

王国强（国家卫生计生委副主任 国家中医药管理局局长）

### 副组长

张伯礼（中国工程院院士 天津中医药大学教授）

王志勇（国家中医药管理局副局长）

### 组 员

卢国慧（国家中医药管理局人事教育司司长）

严世芸（上海中医药大学教授）

吴勉华（南京中医药大学教授）

王之虹（长春中医药大学教授）

匡海学（黑龙江中医药大学教授）

王 键（安徽中医药大学教授）

刘红宁（江西中医药大学教授）

翟双庆（北京中医药大学教授）

胡鸿毅（上海中医药大学教授）

余曙光（成都中医药大学教授）

周桂桐（天津中医药大学教授）

石 岩（辽宁中医药大学教授）

黄必胜（湖北中医药大学教授）

# 前言

为落实《国家中长期教育改革和发展规划纲要（2010—2020年）》《关于医教协同深化临床医学人才培养改革的意见》，适应新形势下我国中医药行业高等教育教学改革和中医药人才培养的需要，国家中医药管理局教材建设工作委员会办公室（以下简称“教材办”）、中国中医药出版社在国家中医药管理局领导下，在全国中医药行业高等教育规划教材专家指导委员会指导下，总结全国中医药行业历版教材特别是新世纪以来全国高等中医药院校规划教材建设的经验，制定了“‘十三五’中医药教材改革工作方案”和“‘十三五’中医药行业本科规划教材建设工作总体方案”，全面组织和规划了全国中医药行业高等教育“十三五”规划教材。鉴于由全国中医药行业主管部门主持编写的全国高等中医药院校规划教材目前已出版九版，为体现其系统性和传承性，本套教材在中国中医药教育史上称为第十版。

本套教材规划过程中，教材办认真听取了教育部中医学、中药学等专业教学指导委员会相关专家的意见，结合中医药教育教学一线教师的反馈意见，加强顶层设计和组织管理，在新世纪以来三版优秀教材的基础上，进一步明确了“正本清源，突出中医药特色，弘扬中医药优势，优化知识结构，做好基础课程和专业核心课程衔接”的建设目标，旨在适应新时期中医药教育事业发展和教学手段变革的需要，彰显现代中医药教育理念，在继承中创新，在发展中提高，打造符合中医药教育教学规律的经典教材。

本套教材建设过程中，教材办还聘请中医学、中药学、针灸推拿学三个专业德高望重的专家组成编审专家组，请他们参与主编确定，列席编写会议和定稿会议，对编写过程中遇到的问题提出指导性意见，参加教材间内容统筹、审读稿件等。

本套教材具有以下特点：

## 1. 加强顶层设计，强化中医经典地位

针对中医药人才成长的规律，正本清源，突出中医思维方式，体现中医药学科的人文特色和“读经典，做临床”的实践特点，突出中医理论在中医药教育教学和实践工作中的核心地位，与执业中医（药）师资格考试、中医住院医师规范化培训等工作对接，更具有针对性和实践性。

## 2. 精选编写队伍，汇集权威专家智慧

主编遴选严格按照程序进行，经过院校推荐、国家中医药管理局教材建设专家指导委员会专家评审、编审专家组认可后确定，确保公开、公平、公正。编委优先吸纳教学名师、学科带头人和一线优秀教师，集中了全国范围内各高等中医药院校的权威专家，确保了编写队伍的水平，体现了中医药行业规划教材的整体优势。

## 3. 突出精品意识，完善学科知识体系

结合教学实践环节的反馈意见，精心组织编写队伍进行编写大纲和样稿的讨论，要求每门

教材立足专业需求,在保持内容稳定性、先进性、适用性的基础上,根据其在整个中医知识体系中的地位、学生知识结构和课程开设时间,突出本学科的教学重点,努力处理好继承与创新、理论与实践、基础与临床的关系。

#### 4. 尝试形式创新,注重实践技能培养

为提升对学生实践技能的培养,配合高等中医药院校数字化教学的发展,更好地服务于中医药教学改革,本套教材在传承历版教材基本知识、基本理论、基本技能主体框架的基础上,将数字化作为重点建设目标,在中医药行业教育云平台的总体构架下,借助网络信息技术,为广大师生提供了丰富的教学资源 and 广阔的互动空间。

本套教材的建设,得到国家中医药管理局领导的指导与大力支持,凝聚了全国中医药行业高等教育工作者的集体智慧,体现了全国中医药行业齐心协力、求真务实的工作作风,代表了全国中医药行业为“十三五”期间中医药事业发展和人才培养所做的共同努力,谨向有关单位和个人致以衷心的感谢!希望本套教材的出版,能够对全国中医药行业高等教育教学的发展和中医药人才的培养产生积极的推动作用。

需要说明的是,尽管所有组织者与编写者竭尽心智,精益求精,本套教材仍有一定的提升空间,敬请各高等中医药院校广大师生提出宝贵意见和建议,以便今后修订和提高。

国家中医药管理局教材建设工作委员会办公室

中国中医药出版社

2016年6月



## 编写说明

数据库技术自从 20 世纪 60 年代末期诞生以来,表现出强劲的发展势头和旺盛的生命力,成为计算机科学最为活跃最为实用的分支之一。基于数据库技术设计开发的各种信息管理系统已经成为现代生活的基础性设施,遍布工作生活的各个领域。在医学领域,数据库技术的应用也极为普遍,国家卫生信息“十二五”规划提出,要建立居民健康数据库和电子病历数据库,医院信息系统(HIS)、电子病历系统(EMR)、实验室检验系统(LIS)和医学图像档案系统(PACS)等也都是以数据库技术为核心来构建的。

数据库课程已经成为世界范围内计算机学科的核心课程之一,是计算机科学与技术、软件工程、信息管理与信息系统、电子商务等专业的必修课程。因为其重要性和普及程度,也成为许多非计算机专业的选修课程,已经由一种计算机应用的专门技术发展为现代计算机环境下理论研究与应用技术的核心组成部分。

编者致力于为高等院校本科生提供一本实用管用的数据库教材,本教材具有如下特点:

1. 以一个统一的门诊数据库案例贯穿全书各个章节:从第一章绪论开始,到数据模型、SQL 语言、数据安全性、数据库的设计和应用开发,乃至数据库的查询优化和并发控制,都是围绕统一的门诊数据库案例展开。通过统一的案例,读者可以深刻而系统地体会数据库的概念和原理,掌握数据库设计开发技术。

2. 编写内容突出了实用性。本书围绕门诊数据库案例,详细讲解了数据库系统的设计,介绍了当前流行的数据库系统体系结构和数据库访问技术,并给出了 ADO.NET 访问实现示例和 JDBC 访问实现示例,帮助读者掌握数据库系统的开发技术。同时,本书介绍了当前若干种典型的医药数据库系统,对于一些过时的技术不再展开讨论。

3. 设计了丰富的实验内容。本书围绕数据库的安装、SQL 查询、数据库安全性和完整性、数据库的管理等方面设计了多个实验,并提出具体的实验要求。同时,也提出了课程设计的要求。

本书第一章由杜建强编写,第二章由崔茂应编写,第三章由李志敏、孙扬波、任真编写,第四章由李艳编写,第五章由胡孔法、余侃侃编写,第六章由米鹏、曲雷编写,第七章由杜清、程春雷编写,第八章由杨丽琴编写,第九章由孙艳秋编写,第十张由姜姗编写,第十一章由殷云霞编写,第十二章由罗维编写。全书由杜建强和胡孔法统稿。

本书定位于高等中医药院校计算机科学与技术、软件工程、信息管理与信息系统、医药信息工程及其相关专业数据库课程教材,也可作为其他非计算机专业的数据库课程教材。本书也可用作软件开发人员及医药卫生行业信息化工作者的参考书。

由于编者水平有限,书中存在错误和不足,敬请读者提出宝贵意见,以便再版时修订提高。

# 目 录

<b>1 绪论</b>	<b>1</b>	<b>3 SQL 语言</b>	<b>33</b>
1.1 数据库系统概述	1	3.1 SQL 概述	33
1.1.1 数据库系统基本概念	1	3.1.1 SQL 的产生与发展	33
1.1.2 数据管理技术的发展	3	3.1.2 SQL 的特点	34
1.1.3 数据库系统的特点	4	3.2 医生门诊系统数据库	34
1.2 数据模型	6	3.3 数据定义	36
1.2.1 数据模型概述	6	3.3.1 数据库的定义与撤销	36
1.2.2 E-R 模型	8	3.3.2 基本表的定义和维护	37
1.2.3 层次模型	10	3.3.3 索引的建立和删除	40
1.2.4 网状模型	11	3.4 单表查询	41
1.2.5 关系模型	12	3.4.1 查询列	42
1.2.6 面向对象模型	14	3.4.2 查询元组	43
1.2.7 XML 模型	14	3.4.3 查询结果排序	47
1.3 数据库系统结构	16	3.4.4 简单统计查询	47
1.3.1 三级模式结构	16	3.4.5 分组查询	48
1.3.2 二级映像及数据独立性	17	3.5 复合查询	49
<b>2 关系模型与关系代数</b>	<b>19</b>	3.5.1 连接查询	49
2.1 关系数据结构	19	3.5.2 嵌套查询	52
2.1.1 关系	19	3.5.3 集合查询	55
2.1.2 关系模式	21	3.5.4 SELECT 语句一般格式	56
2.1.3 关系实例	22	3.6 数据更新	57
2.2 关系操作	23	3.6.1 插入数据	57
2.2.1 基本的关系操作	23	3.6.2 更新数据	58
2.2.2 关系数据语言的分类	23	3.6.3 删除数据	60
2.3 关系约束	24	3.7 视图	61
2.3.1 关系的三类完整性约束	24	3.7.1 定义视图	61
2.3.2 实体完整性	24	3.7.2 查询视图	62
2.3.3 参照完整性	25	3.7.3 更新视图	62
2.3.4 用户定义的完整性	26	3.7.4 视图的作用	63
2.4 关系代数	26	<b>4 数据库完整性与安全性</b>	<b>67</b>
2.4.1 传统的集合运算	26	4.1 数据库完整性	67
2.4.2 专门的关系运算	27	4.1.1 实体完整性	67
		4.1.2 参照完整性	70

4.1.3	自定义完整性	72	6.2.3	需求分析的表示	115
4.1.4	触发器	73	6.2.4	需求分析示例	116
4.2	数据库安全性概述	75	6.3	概念结构设计	118
4.2.1	三类安全性问题	75	6.3.1	概念结构概述	118
4.2.2	安全性标准	76	6.3.2	概念结构设计策略	118
4.3	数据库安全性控制	78	6.3.3	以 E-R 图表示的概念结构设计	118
4.3.1	用户标识与鉴别	79	6.3.4	概念结构设计实例	125
4.3.2	存取控制	79	6.4	逻辑结构设计	127
4.3.3	视图机制	85	6.4.1	概念模型向关系模型的转换	127
4.3.4	审计	86	6.4.2	关系数据模型的优化	130
4.3.5	数据加密	86	6.4.3	用户子模式设计	131
4.3.6	统计数据库安全性	87	6.4.4	逻辑结构设计示例	132
<b>5</b>	<b>关系数据理论</b>	<b>89</b>	6.5	数据库的物理设计	133
5.1	关系模式设计中的问题	89	6.5.1	数据库物理设计的内容与方法	133
5.2	关系模式规范化	90	6.5.2	存储结构与存取方式的选择	134
5.2.1	函数依赖	90	6.6	数据库设计评价	137
5.2.2	码	91	6.6.1	数据库评价	137
5.2.3	1NF	92	6.6.2	数据库逻辑设计评价和物理设计评价	137
5.2.4	2NF	92	6.7	数据库的实施和维护	138
5.2.5	3NF	93	6.7.1	数据库的初始化	139
5.2.6	BCNF (Boyce - Codd 范式)	93	6.7.2	数据库的试运行	140
5.3	多值依赖与 4NF	94	6.7.3	数据库的运行和维护	141
5.3.1	BCNF 关系模式存在的问题	94	<b>7</b>	<b>数据库应用开发</b>	<b>143</b>
5.3.2	多值依赖	95	7.1	数据库系统体系结构	143
5.3.3	4NF	96	7.1.1	常用开发体系结构	143
5.4	数据依赖的公理系统	97	7.1.2	C/S 结构	144
5.4.1	函数依赖的逻辑蕴涵	97	7.1.3	B/S 结构	147
5.4.2	Armstrong 公理系统	97	7.1.4	体系结构小结	148
5.4.3	函数依赖集的等价和覆盖	100	7.2	数据库访问技术	148
5.5	模式的分解	101	7.2.1	ODBC	149
5.5.1	分解的无损连接	101	7.2.2	MFC DAO	150
5.5.2	保持函数依赖	105	7.2.3	RDO	151
<b>6</b>	<b>数据库设计</b>	<b>110</b>	7.2.4	OLE DB	152
6.1	数据库设计概述	110	7.2.5	ADO	152
6.1.1	数据库设计的特点	110	7.2.6	ADO.NET	153
6.1.2	数据库设计的方法	110	7.2.7	JDBC	154
6.1.3	数据库设计的步骤	112	7.3	数据库应用开发实践	155
6.2	需求分析	114	7.3.1	存储过程	156
6.2.1	需求分析的任务	114	7.3.2	ADO.NET 访问实现示例	160
6.2.2	需求分析的方法	114	7.3.3	JDBC 访问实现示例	167

<b>8 数据库系统管理</b>	<b>173</b>	<b>10.4 物理优化</b>	<b>209</b>
8.1 数据库管理员职责	173	10.4.1 存取路径优化	209
8.2 数据库恢复技术	174	10.4.2 基于代价的优化	209
8.2.1 事务的基本概念	174	<b>11 并发控制</b>	<b>212</b>
8.2.2 故障的种类	175	11.1 问题的提出	212
8.2.3 恢复的实现技术	176	11.2 并发控制	214
8.3 恢复的策略	178	11.2.1 封锁机制	214
8.3.1 事务故障的恢复	178	11.2.2 封锁的粒度	217
8.3.2 系统故障的恢复	179	11.3 并发实现	219
8.3.3 介质故障的恢复	179	11.3.1 活锁和死锁	219
8.4 具有检查点的恢复	180	11.3.2 并发调度的可串行化	221
8.5 数据库镜像	181	<b>12 数据库技术新发展</b>	<b>225</b>
<b>9 典型的医学数据库系统</b>	<b>184</b>	12.1 数据库系统发展特点	225
9.1 医院管理信息系统 HIS	184	12.1.1 数据模型的发展	225
9.1.1 医院信息系统的发展	184	12.1.2 数据库技术与新技术结合	226
9.1.2 典型医院信息系统的功能	185	12.1.3 数据库技术在新领域的拓展	227
9.1.3 典型医院信息系统的设计与开发	186	12.2 数据库技术发展趋势	227
9.2 电子病历系统 (EMRS)	188	12.2.1 发展趋势	228
9.2.1 电子病历系统的发展	188	12.2.2 几项主流技术发展趋势	229
9.2.2 典型电子病历系统的功能	189	12.3 几种新型数据库系统	230
9.2.3 典型电子病历系统的设计与开发	191	12.3.1 分布式数据库系统	231
9.3 检查检验系统 (LIS)	192	12.3.2 对象关系数据库系统	232
9.3.1 检查检验系统的发展	192	12.3.3 XML 数据库系统	235
9.3.2 典型检查检验系统的功能	193	12.3.4 NoSQL 数据库系统	240
9.3.3 典型检查检验系统的设计与开发	194	12.3.5 数据仓库与数据挖掘	241
9.4 医学图像系统 PACS	195	<b>实验</b>	<b>244</b>
9.4.1 医学图像系统的发展	195	实验1 DBMS 的安装	244
9.4.2 典型医学图像系统的功能	196	实验2 数据库、数据表的创建与维护	244
9.4.3 典型医学图像系统的设计与开发	196	实验3 数据表的单表查询	246
9.5 医药数据库系统的发展趋势	197	实验4 数据表的复合查询	247
<b>10 关系查询原理及优化</b>	<b>200</b>	实验5 数据表的数据更新	247
10.1 关系查询处理	200	实验6 视图的创建与更新	248
10.1.1 查询处理步骤	200	实验7 数据库的完整性与安全性控制	249
10.1.2 查询处理算法	201	实验8 SQL Server 数据库管理	250
10.2 关系查询优化	203	课程 (大作业) 设计	250
10.2.1 查询优化概述	203	<b>参考文献</b>	<b>251</b>
10.2.2 查询优化示例	204		
10.3 代数优化	206		
10.3.1 关系代数转换	206		
10.3.2 启发式规则	207		

# 1 绪论

数据库技术自从 20 世纪 60 年代末期诞生以来,表现出强劲的发展势头和旺盛的生命力,成为计算机科学最为活跃最为实用的分支之一。基于数据库技术设计开发的各种信息管理系统已经成为现代生活的基础性设施,遍布工作生活的各个领域。从银行储蓄系统、ATM 自动取款机、大型超市购物系统、火车票飞机票购票系统、校园一卡通系统、图书馆信息系统、微博、QQ 即时通讯、淘宝购物到企业资源计划系统(ERP)、电子政务系统(e-Government)、地理信息系统(GIS)等无不构建于数据库之上。在医学领域,数据库技术的应用也极为普遍,国家卫生信息“十二五”规划提出,要建立居民健康数据库和电子病历数据库,医院信息系统(HIS)、电子病历系统(EMR)、实验室检验系统(LIS)、医学图像档案系统(PACS)、医疗保险系统、区域卫生信息平台、药品零售企业 GAP、药品生产企业 GMP 等也都是以数据库技术为核心来构建的。可以说,数据库系统建设的规模、数据量的大小、应用范围等已经成为衡量一个国家、地区、单位和部门信息化程度和管理水平的重要标志。

数据库课程已经成为世界范围内计算机学科的核心课程之一,是计算机科学与技术、软件工程、信息管理与信息系统、电子商务等专业的必修课程。因为其重要性和普及程度,也成为了许多非计算机专业的选修课程,已经由一种计算机应用的专门技术发展为现代计算机环境下理论研究与应用技术的核心组成部分。

本章主要介绍数据库系统这门学科中的一些重要概念、数据管理技术的发展、数据库系统的特点和数据模型的知识等,使读者对数据库课程有一个大致的了解和整体的把握,为后面的学习奠定一个良好基础。需要指出的是,本章对于全面正确认识数据库系统的特征与功能,把握数据库原理与技术都相当重要,具有基础性的意义。

## 1.1 数据库系统概述

在系统地学习数据库课程之前,先介绍一些数据库最常用的基本概念:数据(Data)、数据库(Database, DB)、数据库管理系统(Database Management System, DBMS)、数据库系统(Database System, DBS)。

### 1.1.1 数据库系统基本概念

#### 1. 数据

数据是描述事物的符号记录。数字是描述事物最常用的符号,如一个病人年龄 35 岁,身高 175cm,体温 38.5℃等。当然,还可用字符串“张三”“江西中医药大学”来描述病人的姓名和工作单位,用文本记录病历。这些字符串、文本,以及病人的心电图、X 光片图像等也是数据。因此,数据的种类是很多的,包括数字、文字、图形、图像、音频、视频等。

光从数据本身有时很难完全理解其含义。例如,下面一段话是用来描述一位病人的:

(张三, 男, 45, 60, 39)

除了能猜测出病人姓名是张三, 性别是男外, 后面三个数字很令人费解。如果将上面这段话加上一些内容, 变成表 1 的样子, 就一目了然了。

表 1-1 病人基本信息表

姓名	性别	年龄	体重 (kg)	入院体温 (°C)
张三	男	45	60	39
...	...	...	...	...

因此, 数据的表现形式必须同数据的解释结合起来才能被准确理解, 数据和关于数据的解释是不可分的。具有特定意义的数据也称为信息, 信息是对现实世界中存在的客观实体、现象和关系进行描述的具有特定意义的数据, 是经过加工处理的数据。数据和信息是两个既紧密联系、又相互区别的概念, 数据是信息的具体表现形式, 信息是数据有意义的表现。很多情况下, 数据库系统也称为信息管理系统。

## 2. 数据库

数据库是长期储存在计算机内, 有组织的、可共享的大量数据的集合。为了有效管理病人, 可以建立病人数据库。首先要收集大量病人的数据, 包括病人基本信息、病人临床表现、病人检查信息、病人诊断信息和病人治疗信息等等, 然后把收集到的大量数据按照某种数据模型有序组织起来, 再存储在计算机内, 既可以供医生、又可以供病人查询使用, 实现医生和病人的数据共享。同其他形式的数据存储相比, 数据库中的数据具有永久存储、有组织和可共享三个基本特点。

## 3. 数据库管理系统

在建立病人数据库的时候, 面临的挑战是如何科学地组织和存储大量的数据, 如何高效地获取和操作数据, 如何有效地维护数据。数据库管理系统承担了这些艰巨的任务。数据库管理系统是一种操纵和管理数据库的大型软件, 用于建立、使用和维护数据库, 简称 DBMS。它具有强大的数据定义、数据组织、数据存储、数据操纵和运行维护等功能, 对数据库进行统一的管理和控制, 以保证数据库的安全性和完整性。用户通过 DBMS 访问数据库中的数据, 数据库管理员也通过 DBMS 进行数据库的维护工作。

数据库管理系统是位于操作系统和用户之间的数据管理软件。同操作系统一样, 数据库管理系统也是计算机系统中的基础软件, 是一个大型复杂的软件系统, 目前常用的数据库管理系统有甲骨文公司的 Oracle, 微软公司的 SQL Server、Access, IBM 公司的 DB2, 开源数据库 MySQL 以及国产数据库达梦等。

## 4. 数据库系统

数据库系统是为适应实际数据处理的需要而发展起来的一种较为理想的数据处理系统, 通常由软件、数据库和数据库管理员组成。其软件主要包括操作系统、数据库管理系统和各种实用程序。数据由数据库管理系统统一管理, 数据的插入、删除、修改和检索均要通过数据库管理系统进行。数据库管理员负责创建、监控和维护整个数据库, 使数据能被任何授权用户安全有效地使用。医院信息系统、电子病历系统、实验室检验系统、医学图像档案系统和医疗保险系统都是为满足实际数据处理需要建立的数据库系统。数据库系统组成可以用图 1-1 表示。

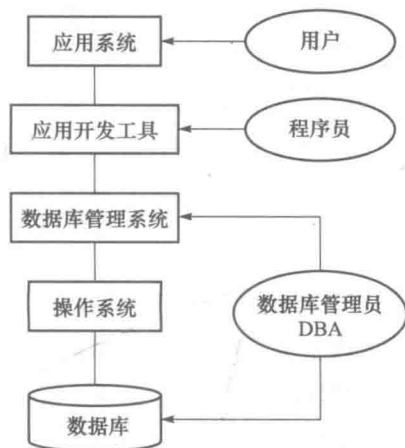


图 1-1 数据库系统组成

### 1.1.2 数据管理技术的发展

数据管理是指对数据的组织、分类、编码、存储、检索和维护，是数据处理过程的主要内容与核心部分。随着计算机硬件和软件的发展，在应用需求的驱动下，数据管理经历了人工管理、文件系统和数据库系统三个发展阶段。

#### 1. 人工管理阶段

数据的人工管理阶段出现在 20 世纪 50 年代中期以前，当时计算机主要用于科学计算，软硬件设施还处于较为原始的时期。硬件方面没有可供直接访问的磁盘等存储设备，外存只有卡片机、磁带机；软件方面没有通用的操作系统，没有数据管理方面的软件，只有汇编语言，数据采用批处理方式。人工管理数据有如下特点和不足：

(1) 数据不保存：当时计算机主要用于科学计算，一般不需要将数据长期保存。只是当计算某一问题时将数据输入，计算任务完成后就撤走。不仅对于用户数据是如此处理，对于系统软件也是同样处理。

(2) 数据由程序管理：由于没有专门的软件系统完成数据的管理工作，因此应用程序不仅要设计数据的逻辑结构，还要设计数据的物理结构，包括存储结构、存取路径和输入方式等。

(3) 数据不共享：数据依附于应用程序，一组数据只能对应一个程序。在出现多个不同程序涉及相同数据时，必须各自定义，难以相互参照利用，造成程序之间存在大量的冗余数据。

(4) 数据没有独立性：由于数据紧密依附应用程序，当数据的逻辑结构或物理结构变动时，必须对应用程序进行相应修改，程序员的负担相当沉重。

人工管理阶段数据与应用程序之间的对应关系如图 1-2 所示。

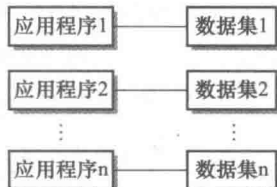


图 1-2 人工管理阶段数据与应用程序之间的对应关系

#### 2. 文件系统管理阶段

文件系统管理数据阶段出现在 20 世纪 50 年代中期到 60 年代中期。这一时期，由于数据量逐渐增加，数据存储、数据检索和数据维护的需求越来越迫切，计算机软硬件技术也有很大

的改进。硬件方面出现了磁盘、磁鼓等直接存储设备，软件方面出现高级语言和操作系统，而且在操作系统中有专门的数据管理软件，一般称之为文件系统。数据不仅能够批处理，而且能够联机实时处理。文件系统管理数据阶段有如下特点：

(1) 数据长期保存：由于大量使用计算机进行数据处理，需要对数据反复进行查询、更新（插入、删除和修改）等基本操作，因此数据以文件形式长期保存在计算机外部存储设备中。

(2) 数据由文件系统管理：文件系统把数据组织成相互独立的数据文件，实现了“按文件名访问，按记录存取”的数据管理技术。应用程序与数据文件间存在多对多的关系：一个应用程序可以使用多个数据文件，一个数据文件也可以被多个应用程序所使用。

相对于人工管理阶段，文件系统管理阶段在数据的共享性和独立性方面有一定的进步，但仍然存在以下缺点：

(1) 数据共享性差，冗余度大，一致性差：由于文件之间没有结构，缺乏联系，同样的数据有可能在多个文件中重复存储，导致文件系统数据冗余度大，而且容易导致数据不一致。

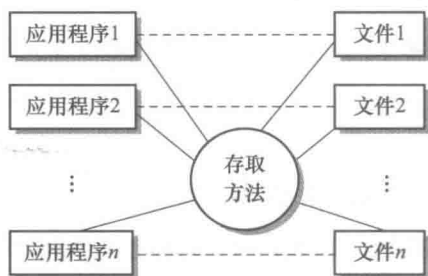


图 1-3 文件系统阶段数据与应用程序之间的对应关系

(2) 数据的独立性差：由于文件只能存储数据，不能存储文件的结构表述；数据文件的操作都要依靠应用程序实现。当文件的逻辑结构发生变化，应用程序就需要修改，文件和程序之间缺乏独立性。文件系统一般是针对特定应用程序的，在现有数据基础上增加新的应用会很困难，系统不容易扩充。

文件系统阶段数据与应用程序之间的关系如图 1-3 所示。

### 3. 数据库系统管理阶段

进入 20 世纪 60 年代，随着计算机应用领域的日益拓展，计算机管理的数据规模越来越大，数据量急剧增长，多种应用共享数据集合的要求越来越强烈，基于文件系统的数据库管理技术无法满足实际应用的需要。在这一时期，计算机硬件技术得到了飞速发展，大容量磁盘、磁盘阵列等数据存储技术日益成熟，新型的存储硬件陆续进入市场，而价格却在不断下降；同时许多公司竞相投入到数据库管理技术的开发与研制当中，软件环境不断完善。在迫切的实际需求和良好的软硬件环境中，统一管理数据的专门软件系统——数据库管理系统应运而生。数据库管理系统的诞生，标志着数据库管理技术的飞跃。

#### 1.1.3 数据库系统的特点

同人工管理和文件系统相比，数据库系统管理数据具有显著的优点。

##### 1. 数据整体结构化

数据整体结构化是数据库系统与文件管理系统的根本区别。文件系统中，文件内部的数据是结构化的，但是文件和文件之间的数据没有联系。而数据库系统不仅要考虑数据表内的结构化，还要考虑数据表之间的联系。

例如，以文件系统管理医院的数据，存在管理病人、医生和诊疗总费用的三个文件。三个文件记录的属性组成见图 1-4。



病人文件Patient的记录结构

病人ID	姓名	性别	职业	出生日期	电话
------	----	----	----	------	----

医生文件Doctor的记录结构

医生ID	姓名	职称	科室
------	----	----	----

诊疗总费用文件CureFee的记录结构

病人ID	医生ID	总费用
------	------	-----

图 1-4 病人、医生、诊疗总费用文件结构

病人、医生、诊疗总费用三个文件记录内部是结构化的，但是文件和文件之间却没有联系。而实际上三个文件是存在联系的，诊疗总费用 CureFee 文件中的病人 ID 必须是病人文件 Patient 中的某个病人 ID，医生 ID 必须是医生文件 Doctor 中的某个医生 ID。显然文件系统不能反映这种联系，只能通过应用程序才能体现出来。在数据库系统中，记录之间的这种联系可以通过完整性约束来实现，实现了整体数据的结构化。

数据库系统数据整体结构化的另一个表现是数据不仅仅针对某个部门级的应用，而是面向整个组织的多个应用。例如，一个医院的信息管理系统不仅要考虑门诊部的门诊管理，还要考虑住院部的住院管理、人事部的人事管理、财务部的财务管理、后勤部门的物资管理等。因此，医院信息管理系统中医生数据就要面向各个部门而不仅仅是门诊科室，如图 1-5 所示。

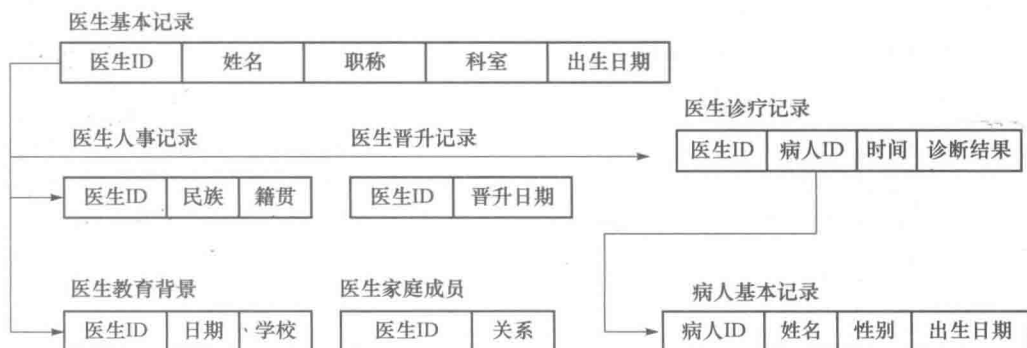


图 1-5 医药信息管理系统中的医生数据

## 2. 数据共享性高，冗余度低

数据库中的数据不是面向特定应用而是面向整个组织的，因此多个用户、多个应用可以访问同一数据，数据是高度共享的。数据共享可以大大减少数据的冗余程度，节约存储空间。数据共享还可以有效防止数据的不相容性和不一致性。

由于数据实现了整体结构化，面向整个组织，不仅可以被已有的多个应用共享，还可以增加新的应用，满足不同用户的多种需要，数据库系统具有良好的扩充性。

## 3. 数据独立性高

数据独立性是指数据和应用程序之间的独立性，分为物理独立性和逻辑独立性。物理独立性是指应用程序与存储在磁盘上的数据是互相独立的。用户程序只需关注数据库名称、数据文件名称和文件中的属性名称等逻辑概念，而不用考虑数据的实际物理存储，不需关心实际数据

NOTE