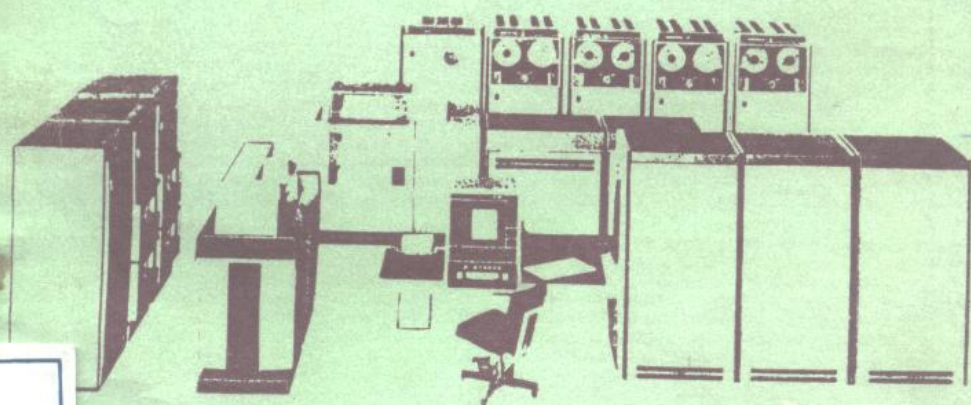


计算机基础知识丛书

数据库及其应用

姚卿达 编著



科学出版社

3

1

计算机基础知识丛书

数据库及其应用

姚卿达 编著

科学出版社

1987

内 容 简 介

数据库是六十年代末期开始发展起来的一门新技术,是管理和利用信息资源的先进工具。

本书着重介绍了数据库的基本概念,几种典型的数据库模型,举例说明了数据库管理系统的运用。此外,作者还结合近年来在数据库应用系统建设方面的实践,介绍了数据库设计方法以及数据库应用系统的分析与设计技术。在取材方面,注意了先进性与实用性相结合,书末还附有数据库管理系统的实用材料。

本书在叙述方法上,力求通俗易懂。可供具有中等文化程度以上的,初具计算机基本知识的学生、教师和管理人员,以及有关的工程技术人员阅读参考。

J5505/23
计算机基础知识丛书

数据库及其应用

姚卿达 编著

责任编辑 曾美玉

科学出版社出版

北京朝阳门内大街137号

中国科学院印刷厂印刷

新华书店北京发行所发行 各地新华书店经售

1987年4月第一版 开本: 787×1092 1/32

1987年4月第一次印刷 印张: 9 1/4

印数: 0001—9,000 字数: 209,000

统一书号: 15031·800

本社书号: 5108·15-8

定价: 2.20元

前 言

为迎接新的技术革命的到来，一个推广应用计算机的高潮正在全国兴起。广大科技人员、干部、工人和学生都迫切要求学习计算机基础知识。为了适应这一令人鼓舞的形势发展需要，中国计算机学会普及委员会决定编辑一套《计算机基础知识丛书》，将于近年内陆续出版。本书就是这套丛书中的一本。

计算机是信息处理的工具。现代计算机的迅猛发展加速了社会信息化的进程。信息的巨大价值日益被人们所认识。信息的快速生成、广泛传播和有效利用使人类更能高度发挥智力劳动的效益和施展社会组织的才能，从而极大地促使科学技术的进步和推动社会生产力的发展。普及计算机知识，特别是计算机应用方面的知识，正是整个信息生成、传输、利用过程中不可缺少的重要环节。我们就是基于这样的认识来组织编写《计算机基础知识丛书》的。它应该能为各行各业科技人员提供在自己的专业领域中应用计算机急需的计算机基础知识，同时也为计算机专业科技人员更新和扩展自己的知识领域创造条件。

计算机科学技术是一门综合性很强的学科，发展又十分迅速。我们要求丛书能尽力做到先进性和实用性相结合，科学性和通俗性相结合，硬件和软件相结合。为了保持题材新颖、内容先进，我们采取分批确定选题、随定随写、尽快出版的方针。为了对广大科技人员应用计算机提供实际指导，我们希望能结合一些应用计算机的实例来讲解原理，不但介

绍计算机有什么用，而且还要介绍它怎样用。我们要求丛书的大多数选题都具有中级科普读物的性质，便于自学，也能提供给相应程度的培训班用作参考教材。作为普及读物，我们自然要更加注意叙述生动形象、文笔通畅易懂，使不同程度的读者都能利用丛书获得收益。

由于我们经验不足，水平有限，疏漏不妥之处在所难免，尚希望广大读者批评指正，并提出宝贵意见。

《计算机基础知识丛书》编辑委员会

目 录

绪 论	(1)
0.1 信息与信息资源	(2)
0.2 信息资源管理与数据库技术	(5)
第一章 数据库与数据库系统	(8)
1.1 什么叫数据库?	(8)
1.2 数据库带来的好处	(12)
1.3 数据库系统	(15)
1.4 数据库应用系统及其主要问题	(24)
第二章 数据模型	(28)
2.1 基本知识	(28)
2.2 现实世界的信息结构及其表示	(41)
2.3 数据结构图	(46)
2.4 E-R图	(55)
2.5 数据模型的功能与分类	(59)
第三章 关系数据库及其应用	(62)
3.1 关系数据模型	(62)
3.2 大众数据库dBASE II	(71)
3.3 一个例子——商品管理数据库模型	(75)
3.4 数据库建立与运用的基本操作	(81)
3.5 进一步的数据操作——关系操作	(98)
3.6 打印各种格式的报表	(106)
3.7 命令文件作成与调用	(112)
3.8 SQL、QBE和SEQUITUR语言	(118)
第四章 IMS数据库及其应用	(128)
4.1 层次数据模型	(128)

4.2	如何描述IMS数据库	(131)
4.3	逻辑数据库记录型与程序说明块	(139)
4.4	逻辑关系和辅助索引	(143)
4.5	如何建立IMS数据库	(148)
4.6	一个IMS数据库实例	(155)
第五章	网络模型数据库	(169)
5.1	DBTG网络数据模型	(169)
5.2	DBTG模型的编码	(176)
5.3	DBTG子模式与数据操作	(180)
5.4	DBTG模式的实现	(182)
5.5	TOTAL二级网络数据模型	(185)
5.6	PDM系统的运用与实例	(195)
第六章	数据库应用系统	(208)
6.1	数据库应用系统的开发	(208)
6.2	数据库设计	(223)
6.3	数据库设计的支援工具	(237)
6.4	基于规范关系的数据库设计	(257)
附录 I	dBASE II 命令简表及说明	(267)
附录 II	SQL语法	(276)
参考文献		(285)

绪 论

数据库是六十年代末期开始发展起来的一门新技术，是管理信息资源的先进工具。这门技术已广泛应用于各个方面，如科技情报检索、图书管理、库存管理、企业管理与经营决策、财会统计、银行存取款管理、火车或飞机票预订、人事材料管理、旅业管理、海关进出口情报收集与管理、地图数据库、种子数据库、医院管理、交通管理、案件管理与分析、生活服务等等；有的还用于办公室自动化、科学研究、辅助设计、人工智能、资源探测、太空技术等多个方面。目前，它已受到各行各业人员的高度关注，了解数据库、运用数据库和研究数据库技术的人越来越普遍。

广大科技人员和管理人员，为了使用计算机进行辅助管理和处理信息，通过各种方式来学习数据库技术，他们需要数据库方面的读物和资料。本书正是从广大科技人员和管理人员需要出发而写的数据库及其应用的入门书，也可以作为大专院校师生的参考读物。

全书包括下列内容：

绪论——谈谈数据库技术在信息资源管理中的作用，

第一章——介绍数据库形象与数据库管理系统的组成，

第二章——讲解数据库结构方面的基本知识，以及数据库的重要概念：数据模型，

第三章——介绍关系数据模型，并结合微型机上广泛流行的关系数据库管理系统dBASE II的使用，讲解如何建立和使用关系数据库；

第四章——介绍层次数据模型，并以实例说明如何建立

和使用IMS数据库;

第五章——介绍DBTG网络数据模型以及TOTAL模型,举例说明网络模型数据库的建立与操作过程;

第六章——讨论数据库应用系统的分析与设计,讨论数据库设计方法。

在取材方面,注意了先进性与实用性,从数据库应用角度考虑较多。在叙述方法上,力求通俗,列举例子来说明概念与方法。当我们介绍实用的数据库软件——数据库管理系统时,着重从宏观上讲解运用步骤与方法,在不影响阅读的前提下,略去了一些类似于使用手册那样的详细叙述,旨在给读者使用该系统时的一个引导,如果辅以相应的手册和说明,那就可以很快地让读者深入到该系统的使用环境之中。由于数据库技术是计算机科学技术中的新课题之一,它是在六十年代萌芽、七十年代大发展的新技术,所以涉及到计算机科学与数学各个分支的知识,如数据结构、操作系统、程序语言(特别是COBOL语言)、离散数学、软件工程、系统分析与设计等。为了使读者不致因为某些知识不了解而读不下去,凡涉及到计算机基本知识以外的问题,将作必要的说明或给出参考线索。

本书写作过程中得到中国人民大学萨师煊教授的指教与协助,北京大学杨美清教授和她的研究生对本书初稿提出了宝贵意见,出版前,萨师煊教授和中国科技大学研究生院罗晓沛副教授作了最后审定,在此表示衷心感谢!

现在让我们转入本绪论的正题。

0.1 信息与信息资源

信息指的是各种形式的数据、资料、消息、信号…中所包

含的意义，而载荷信息的物理符号（如数字、字母、文字、图形、声音…）称为数据。信息与数据是互相联系又有一定区别的两个概念，信息直接反映着现实的概念，向接受者（人或接收信息的机器）提供关于现实世界新的事实的知识，而数据则是信息的具体表现。譬如想用电报通知南京的李明同志“数据库会四月三日上午开”，首先将这个信息写成电文（汉字形式），然后由译报员译成电报码（数字形式），即：

2422 2207 1655 2585 0934 2588 0005 2480

（数）（据）（库）（会）（四）（月）（三）（日）

0006 0552 7030

（上）（午）（开）

这样就可交发报员发报了，这个信息再由电报机转换为电信号通过线路传送到南京，最后由译报员或译码机还原为电文，送到李明手里。这里，李明是信息的接收者，这个信息先后用了汉字、数字以及电信号三种不同形式的数据来表示。通常情况下，往往不严格地区分信息与数据，而是把信息看作各种形式表示的数据之总称。在不需要严格区分的地方，“信息”与“数据”可以互换使用，例如说“信息处理”与“数据处理”是同义的，“信息资源”也可以说为“数据资源”。信息一词的英文是Information，有多种译法，日本人叫“情报”。关于信息的解释，不同的学派有不同的观点，这里就不讨论了。

信息作为一种资源，并且与能源、材料相提并论，这种认识是电子计算机问世以后才逐步确立的。信息一直在现实世界中存在，而且被人们自觉或不自觉地加工和利用着，促进着社会 and 科学技术的发展。为了收集和交换信息，古代就产生了信息的各种载体，那就是通过语言来交换信息，通过

文字来记载、积累和传播信息；到了二十世纪，随着电磁波
的发现，先后发明了电报、电话、电影、广播与电视等传播信
息的先进工具，并用录音机、录像机等设备来收集和保存声
音、图象形式的信息。另一方面，又先后发明了各种加工处
理信息的工具，如算盘、计算尺、计算机等，使人们更有效
地利用信息。为什么人们一直那么重视信息、而且不断创造
出新的工具来收集、保存、加工与传播信息呢？这是因为人
们的社会活动、生产活动和科学技术的发展离不开信息，信
息的本质与威力又逐渐被人们所认识的缘故。

在一个企业的各项活动中，形成了一股包括人力、实物
（原料、设备、半成品、成品、运输工具、能源等）以及货
币在内的物流，这些人力、实物和货币在各个企业之间以
及一个企业内各个部门之间相互传递。在物流进行的同
时，必然会伴随着一股信息流。例如，一个工厂的生产活
动，除了购买原料与设备、组织调动人力、产品出厂、资金
周转的物流外，还有相伴的生产规划、新产品投资计划、
购销合同、生产进度表、设备清单与管理册、产品清单、人
事材料册、各种传票、各种报表、财务帐本等组成的信息
流；财务部门发放工资时，除了发给货币之外，必须有职工
的工资信息——即包含编号、姓名、单位、职务、应发工
资、扣除金额、实发金额、签章等项目的工资清单。信息流
总是与物流同时产生、同时存在、紧紧相依的，这也说明
了信息的重要性。一个工厂的决策人员，必须全面地掌握厂
内关于人力、设备、技术水平、资金、库存等方面的信息，
还要掌握国内外市场动态信息、上级制订的发展规划、其他
同行厂家的生产情况，才能有正确的决策，制订出有远见又
切合实际的生产规划和新产品投资计划。任何一个单位或部
门的管理人员，都必须善于开发和掌握信息、善于利用信

息，才能使管理水平不断提高。

在电子计算机出现以前，人们只知道信息的重要性，但还未认识到信息是一种资源。电子计算机的出现与应用，使信息处理（包括信息的收集、存贮、检索、加工与计算、传播）技术面貌一新，从而推动着科学技术的飞跃发展。据统计，七十年代，各先进国家大量使用电子计算机的结果，科技成果比人类有史以来到六十年代的总和还多。计算机的应用已深入到几乎一切部门，在大量的信息处理实践中，逐渐使人们对信息有了本质的认识，开始把它作为一种重要资源来对待，同时也进一步认识到开发信息资源、利用信息资源对于推动生产和社会发展的巨大威力。据1981年统计，美国拥有电子计算机约两百万台，每年处理信息的工作量相当于四百亿人的工作量，这一工作量靠美国两亿人来，做，根本是不可想象的事。由于美国将计算机、能源和材料并列为国家三大重点，目前每年用于计算机科学研究的投资是十分巨大的，因而为有效地开发和利用信息资源，促使国家技术进步、富有，提供了条件。日本在战后的高速发展，与他们重视信息的开发与利用是分不开的，他们把以计算机为代表的“信息工业”或叫“知识工业”作为带头工业，日本政府颁布“信息处理振兴法”，以立法方式定为国策，还提出了建立“信息化社会”的口号。我国近年来也重视了信息科学的发展，计算机科学技术已列为国家重点发展项目，并取得了可喜的成绩，这将大大地推动四化建设的发展。

0.2 信息资源管理与数据库技术

随着现代科学技术的发展，社会上产生的信息量急剧增加。每一个人都通过各种渠道得到许多信息，例如：听报

告，参加会议，听广播，看电视，阅读文件、报刊、书籍和交谈，以及处理各项事务…都会接收到形形色色的信息。一个办公室的秘书，每天要处理的公函文件可能是数以百计，全世界每年发表的科技论文达几百万篇，一种报纸所刊登的新闻每天可达十几个甚至几十个版面，一台快速打印机每分钟可以印出几万行数据，在太空技术、地下资源探测、海洋开发等方面所获得的信息更是惊人，而且形式多样、相互联系、错综复杂。有人用“信息海洋”、“信息爆炸”、“信息时代”等词汇来形容这种情景。

所谓信息处理或数据处理，就是对信息（各种形式的数据）进行加工处理，它包括对于数据进行操作的一系列活动，如数据的收集、存贮、传播、检索、分类、加工或计算，打印各种报表或输出各种所需要的图形。信息处理的基本目的是从大量的、混乱的、不易理解的数据（事件和数字）中，提取出对某一特定的对人们有价值的、或满足特定要求的小量的数据——有用的信息。或者说从大量事实中提取对用户有价值的要素，并以某种形式提供出来。

数据处理并不是有了电子计算机以后才出现的。自1890年发明了第一台卡片制表机并用其编制人口普查表以来，就已经开始了数据处理的历史。通常把这种用机械设备来进行数据处理的系统称为机械数据处理系统（其英文缩写名是MDPS）。由于机械数据处理系统里存贮数据的介质（如穿孔卡片）、计算速度以及处理能力的限制，其效率是很低的。因而，其应用范围也是有限的，仅仅局限于金融会计及政府或企业部门的统计制表等方面。电子计算机的出现，加上六十年代以来存贮数据的介质迅速发展——出现了磁带、磁鼓、磁盘等大容量的快速存贮介质，使得数据处理的应用范围有了飞速的发展。为了与机械数据处理系统相区别，

人们把用电子计算机进行数据处理的系统称之为电子数据处理系统（英文缩写是EDPS）。由于电子计算机硬件和软件的发展，使得电子数据处理系统已经远远超出了早期机械数据处理系统所局限的数值数据的范围，而扩展到处各种形式的信息。

自七十年代初以来，在数据处理领域中广泛采用数据库技术，使数据处理进入了一个新的阶段，这就是以数据库系统为核心机构的数据处理阶段。这一阶段实现了计算机化的全面信息系统——管理信息系统MIS(Management Information System)。其特点是把一个组织（企业或部门）有关的各种数据集中存贮与管理，建立起综合化的数据库(Data Base),并提供了管理数据库的软件,使各个业务部门方便地使用数据库存贮的数据;同时采用计算机网络技术,进行分布式处理,通过计算机网络把不同部门、不同地区甚至不同国家的计算机或终端连在一起,共享丰富的信息资源,使信息的交换、传递和处理变得更为方便,从而更有效地利用信息资源。由于数据库技术在信息(数据)处理中显示了巨大的威力,所以许多计算机科学家把这个阶段称为“数据库时代”。

在信息处理的一系列活动中,数据的收集、存贮、传播、检索、分类等操作是基本环节,这些基本环节统称为信息管理或数据管理,而数据的处理(加工、计算、打印报表等)则是千变万化的,不同的业务就有不同的处理。在计算机信息处理系统里,数据管理技术所解决的是上述基本环节,而数据的处理则是靠事务处理程序或应用程序来实现的。数据库技术是数据管理的最先进技术,它所解决的正是信息处理的基本问题——信息资源管理问题。

值得指出的是,关于信息资源的开发问题,不是目前的数据库技术所能完全解决的,这方面的概念和方法都有待探讨。

第一章 数据库与数据库系统

1.1 什么叫数据库？

数据库一词是英文database的中译名，而database则是由data和base两个词合并而成的，有时也写成data-base，可缩写为DB，直译为“数据基地”。就是说，当需要某种数据的时候，数据库都能随时提供。目前对这个源语的解释还没有定论^[1]，自六十年代末以来，有许多从事数据库研究的专家都给出了数据库的定义，从不同的角度去描绘数据库，在这里不打算去讨论它们，而是引用比较普遍的一种定义（C.J. Date^[2]）：所谓数据库是被存贮起来的运用数据的汇集，其数据被特定的组织（如公司、银行、医院、大学、政府机关等大型企业或事业单位）的应用系统所利用。从这个定义出发，可将数据库的形象描述如下：

第一，有大量的数据被集中存贮在计算机的存贮器中，汇集成“数据基地”。这大量的数据汇集是包含多种多类的，而且不同种类数据之间相互不是孤立的。例如，一个企业或工厂的生产活动中，往往有下列种类的数据：

采购计划；

向其他厂商定购原料或设备的合同数据；

供销计划；

向外供货合同数据；

物资分配计划与领料单位明细表；

库存物资明细帐本；

器材进库单、出库单与调拨单记录，
生产计划方面的数据，
生产情况管理方面的流动性数据，
财务数据（托收、付款、收款、转帐等），
工资数据，
人事材料，
市场信息，
...

于是一个企业或工厂的生产管理数据库就包括上述全部或部分种类的数据，而且是相互有联系的，象其中的库存物资帐目、进库与出库帐和财务数据之间，就有物资数量和金额方面的直接联系；生产计划、生产管理与人力物力资金（涉及到人事材料、库存物资明细帐以及财务数据等）也有直接联系，即：某个项目需要哪些技术人员和哪些技术工人，需要哪些物资，需要多少资金，一个职工参加哪些项目，某种器材用在哪些项目上；还有其他种种联系。

第二，汇集起来的数据是为特定组织内各业务部门（叫做数据库用户）所利用的，他们可以使用类似英语（English-like）或汉语的会话语言从数据库得到所要的信息或报表，简单直观。例如，在人事数据库的人员表中，可以在终端上打入命令

```
SELECT  姓名  年龄  性别  工资  奖金  
FROM    人员表  
WHERE   职称 = '程序员' AND 工资 < 60
```

而找出所有工资60元以下的程序员名单，并列出现这些人的年龄、性别、工资和奖金表。他们也可以通过专门的业务处理程序来运用数据库中存贮的数据，也就是采用批处理方式或联机方式对存贮的数据进行检索、更新、追加、删除以及加

工处理,从而及时地得到他们所需要的信息,并以报表的形式或其他形式提供出来。例如,某个油田物资供应指挥部,为了实现物资管理现代化,将物资信息存贮于计算机,建立起物资管理数据库(图1-1),对这个数据库操作的业务处理

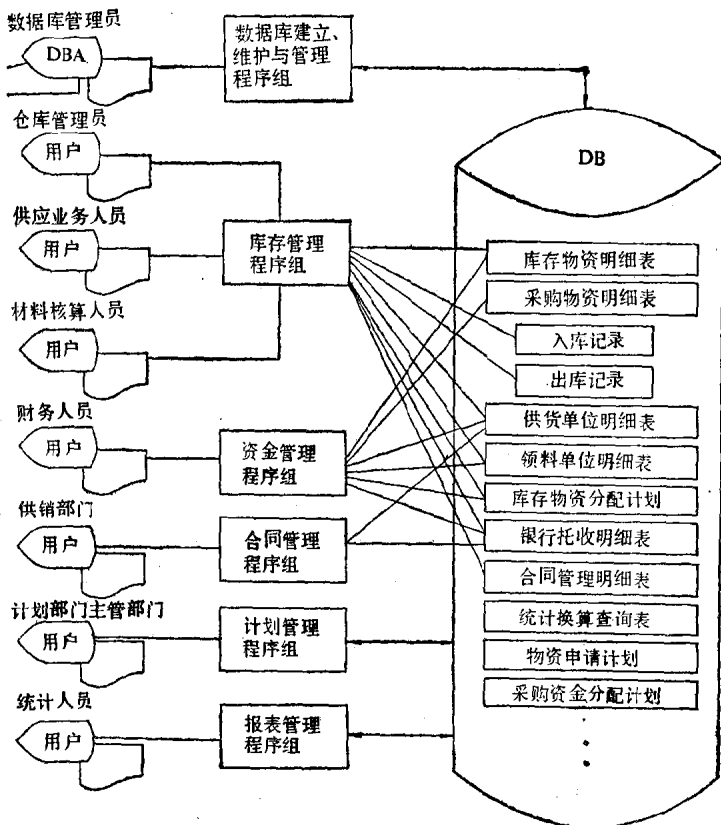


图 1-1 物资管理数据库

程序有:

①建立库存物资明细帐程序——明细帐包括每种器材的编号、名称、规格、数量、价格、库存量、保养期…等项