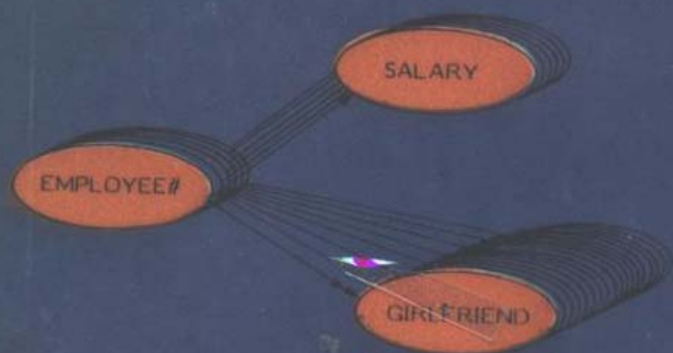




数据库入门

[美] 詹姆斯·马丁
机械工业出版社

• A *James Martin* BOOK

数据库入门

[美] 詹姆士·马丁

焦雄华 译

林钧永 校



机械工业出版社

本书原文名为《数据库用户指南》。

数据库技术几乎是计算机中一切数据处理工作的支柱。作者深入浅出地介绍了什么是数据库，数据库的生产性及灵活性，数据库有关人员的工作，数据模型的建立，数据库的所有权及保密等方面的内容。

本书对于开发数据库技术和加速计算机应用的发展非常有益，它适合于计算机专业人员、管理人员及使用人员学习参考。

AN END-USER'S GUIDE TO DATA BASE

JAMES MARTIN

PRENTICE-HALL, INC. 1981

数据库入门

〔美〕詹姆斯·马丁

熊雄华·译 林钧永·校

机械工业出版社出版（北京阜成门外百万庄南里一号）

（北京市书刊出版业营业许可证出字第117号）

中国农业机械出版社印刷厂印刷

新华书店北京发行所发行·新华书店经销

开本 787×1092 1/32·印张 5 5/8·插页 1·字数 121 千字

1986年7月北京第一版·1986年7月北京第一次印刷

印数 0,001—9,330·定价 1.70 元

统一书号：15033·6286

I

数据库是能够满足多个用户的多种需要而互相联系的、可以被共享的数据集合。

采用这种方法存储起来的数据可以不受使用程序的限制，并且可以使用一个通用的、可控的方法来添加新的数据，或修改和检索现有的数据。

采用数据库技术的目的是为了加速计算机应用的发展，降低应用维护费用，并以尽可能高的效率给用户提供工作中所需要的数据。

数据库技术将几乎成为一切数据处理工作的支柱。只有用户在数据库设计过程中，在各个方面密切配合，数据库的设计才可能圆满成功。

译者序

本书作者鲁姆士·马丁是计算机科学方面有成就的专家。曾在美国国际商业机器公司(IBM)工作,还担任过其他公司的职务。他已有许多著作,如《数据库管理原理》、《数据传输系统分析》、《通讯卫星系统》等。由于他有实践及教学经验,所以他的著作内容实用具体、条理清楚、文字流畅、举例生动,在美国成为畅销的书籍。

本书是一本普及读物,于1981年出版。出版本书的目的是使广大的计算机最终用户懂得数据库的原理及方法,能够与数据处理专业人员密切配合,以便有效地开展计算机应用。书中由基本概念讲起,没有使用艰涩的术语和公式,生动地讲述了数据库技术的由来、发展及趋向,数据模型的建立方法,数据库语言的使用,数据库的结构及各类人员的职责、保密性及安全性等等。本书从开始读起,步步深入,犹如曲径通幽,最后升堂入室,自然而然地得出了结论:“原来数据库是这么回事,照这样作法是可以办得到的。”

我国计算机应用正处于蓬勃发展的前期,数据库技术应用的迫切性已为人们所认识,但尚未进入普遍广泛应用的阶段。所以,正如作者在书中一再强调的,应当在开始时就有一个正确的理解,打下良好的基础,才能避免重蹈一些技术先进国家所曾遇到的灾难及苦恼。因此,普及数据库技术是一件十分重要的事。这也就是为什么要把这本书介绍给我国读者的原意。本书不仅对业务部门的用户有用,对从事计算机数据处理的专业技术人员以及计算机科研教学人员也将有所裨益。

译文如有不当之处,希望读者批评指正。

前 言

由于数据库技术逐渐普及，所以希望所有受到数据库影响的用户如会计师、工程师、业务人员、管理人员、车间主任、部门领导、审计师、各类专业人员和公司总裁等对数据库系统的原理都要有所了解。

这些用户中的许多用户可以说对计算机和软件知识是一无所知。这本书就是为这些用户写的。本书不去叙述数据库管理系统如何工作，也不叙述各种技术细节，主要是叙述用户应当懂得的有关知识，还要叙述一下用户会从数据库中获得什么好处。

虽然这不是一本纯技术性的书籍，但也可供数据处理人员阅读。因为在本书中讨论了影响数据库成败的各种原因。数据库成功与否在很大程度上由用户所决定——取决于他们是否参与数据定义的工作以及数据逻辑模型建立的工作，取决于如何使用数据库语言，是否承认数据库共享原则等。很多数据库管理人员变成了软件专家和计算机方面的专家，但却不重视数据模型的建立和用户的需要。

用户一定要读懂本书所讲的内容。这本书的内容可以在一天之内讲完，最多不超过两天的课程。在这两天内教会用户初步画出泡状图，并用自己的数据建立简单的模型。

如果希望较多地了解数据库技术，可以在读完本书以后，再读本书作者写的另一些书，如《数据库管理原理》(Principle of Data-Base Management) 如果还希望更多地了解数据库管理方法，则可以读作者所著的《数据库环境的管

理》(Managing the Data-Base Environment)。同时,建议用户及分析员看一下这两本书的录象课程,录象是作者与戴尔塔克公司(Deltak Inc.)共同摄制的。在《数据库环境的管理》一书的录象课程中详细地探讨了数据库系统成功或失败的原因。

还有一本补充读物为《数据库用户指南的问题解答》,由Prentice-Hall公司发行。

詹姆士·马丁

目 录

译者序

前 言

第一章 什么是数据库.....	1
一、引言	1
二、什么是数据库	2
三、数据的价值	5
第二章 生产性及灵活性.....	8
一、引言	8
二、数据的多次使用	10
三、文件系统的问题	10
四、维护	13
五、数据库管理系统	15
六、数据的独立性	16
七、特定请求	17
八、不断改变	20
九、失败	21
十、误解	21
十一、基础结构	23
十二、逐步扩充	23
第三章 数据库有关人员的工作.....	25
一、引言	25
二、数据共享的等级	25
三、数据项定义	26
四、数据模型化	30

五、数据管理员	30
六、数据库设计员	31
七、逻辑数据和物理数据	31
八、设计过程	32
九、主题数据库	34
十、领导层的考虑	38
十一、策略	39
十二、漫长的路程	41
第四章 什么是数据	43
一、泡状图	43
二、数据项	43
三、数据项间的联接	44
四、单箭头与双箭头	45
五、数据项的多个具体值	46
六、反向联接	47
七、数据项组	47
八、关键字及属性	48
九、数据项组	50
十、实体	50
十一、组合关键字	53
十二、数据项间的多种联接	55
第五章 数据模型的建立	56
一、逻辑设计	56
二、数据管理	57
三、用户的反复设计	59
四、表示方法的澄清	60
五、综合	61
六、综合过程说明	62
第六章 用户的设计方法	66

一、引言	66
二、数据字典	66
三、数据模型化方法	67
四、应用的压力	73
五、泡状图的编码	74
第七章 怎样成功地建立数据模型	78
一、稳定性	78
二、对未来的考虑	78
三、添加属性	79
四、隐含关键字	79
五、数据字典校核	81
六、组合关键字的注意事项	82
七、关键字间的逆映象	83
八、从属于关键字整体	84
九、隐含主关键字的避免	85
十、组合关键字	86
十一、用户数据视图的大小	86
十二、转换及兼容	88
十三、公布式数据	89
十四、逐步实现	89
第八章 用户的数据库语言	91
一、引言	91
二、交互方式还是脱机方式	92
三、自发性	93
四、数据库语言的能力	94
五、分布式文件	95
六、多种语言	96
七、表格式语言	96
八、通用信息系统(GIS)	100

九、易用数据库语言	104
十、QUERY-BY-EXAMPLE语言(举例查询语言)	104
十一、不具备计算机语言知识的初学者	107
十二、更新	108
十三、算术运算	108
十四、举例查询	109
十五、更复杂的查询	110
十六、AND(与)及OR(或)的隐含	112
十七、表间的链接	113
十八、函数	115
十九、条件框的使用	116
二十、插入及删除	117
二十一、用户自用表	118
二十二、与COBOL语言的比较	120
二十三、分布式数据库	124
第九章 数据的所有权及保密	126
一、数据的所有权	126
二、安全性	128
三、十二条要点	130
四、保护层次	131
五、破坏安全的种类	133
六、三级处理方法	135
七、数据损坏的两种类型	136
八、专门的研究	136
第十章 影响计算机性能的几种考虑	138
一、引言	138
二、数据库查询	138
三、专用计算机	139
四、费用问题	140

五、数据系统的分类	140
六、物理数据库设计	142
七、调度问题及响应时间问题	143
第十一章 分立的用户系统	144
一、集中化与分散化	144
二、保存自己数据的用户	145
三、用户的积极性	147
四、独立的数据输入	148
五、小型计算机的数据库	149
六、快速增长造成的问题	149
七、分立的信息系统	151
附录一 典型的数据库设计步骤	156
附录二 课堂练习问题	160
参考书目	166

第一章 什么是数据库

一、引言

在数据处理领域中，数据库技术变得越来越重要了。将来大量的数据处理要建立在数据库的基础之上。因此，方便适用的数据库对于公司和机关团体将是必不可少的。

并不是所有的数据库系统都是成功的，很多数据库的效果令人失望。总结许多成功和失败的经验以后，可以清楚地看出，能不能把用户有效地结合进来是数据库系统能否成功的关键。为了达到这个目的，就应当使用户了解某些数据库技术。对用户来说，数据库是一个新的领域，复杂难懂，常会使他们感到迷惑不解。通常，用户既不懂得数据库的原理，也不知道他们自己在搞好数据库方面会起到重要的作用。

本书向用户讲解应当掌握的有关数据库的知识，并向用户讲解他们在设计和改进数据库时可能起到的作用。在讲解时尽量避免使用专业术语，以免把数据库技术说得过于复杂。主要目的是让所有最终用户充分理解书中所叙述的基本概念。

很多用户对采用数据库系统十分担心，担心采用数据库以后会改变他们的工作，好象进了黑屋子找不到东南西北一样。要解除这种顾虑，最好的方法是给他们“照个亮”。这就是，一方面要靠弄懂本书所叙述的原理；另一方面要靠参加数据定义和设计的实践。

这里，“最终用户”的含意是数据库的最终使用者，而不

是为最终用户编写程序的程序员之类的中间使用者。数据库最终用户可以是从事各种工作的人员，如会计师、工程师、公务员、管理人员、车间主任、部门领导、审计师、专业人员、保险统计员和公司总裁及其助手等等。

不论是管理人员、计划人员或其他专业人员，如果在他们未来的工作中将采用数据库系统的话，就应当搞懂本书所讲述的内容，并成为数据库工作的积极参加者。

二、什么是数据库

数据库是一个为多种应用提供共享数据的数据集合。

任何一个用户，都不可能看到数据库中所有的各种数据，而只能看到他（她）工作中所需要的那一种数据。可能一个用户只看到一种数据文件，这种文件常常是具有共同的结构，并且很简单，但它却是从相当复杂的数据结构中导出来的。而其他用户看到的是从同一数据库中导出的不同的数据文件。

因此，一个数据库不仅可以为多个用户所共享，而且不同的用户对它会有不同的理解。这种情况可以比作“盲人摸象”。一个人摸到象腿，觉得它象一棵树干；另一个人摸到象尾，觉得它象一根绳子；还有人则摸到了象牙，如此等等。与此类似，数据库的不同用户所看到的是不同的数据视图。因此，必须有人来设计这个“大象”。在设计时，必须保证满足许多不同用户的不同要求。换句话说，所有有关的用户的数据视图全都可以从这个数据库中导出。

设计数据库是一件复杂的工作，但设计技术已经能作到高度的自动化。所以，目前设计中最大的困难在于透彻地了解用户所需要的数据，并在数据库中表示它们。要作到这一点很不容易，因为，不同的用户可能把同一个数据项赋予不

同的名字，而对有些实际上截然不同的数据项又赋予同一个名字。

为了使计算机能正确地进行工作，在描述用户的数据时，需要高度的精确性。

用户部门的责任是要保证他们所需的数据都存放在数据库中，并给予正确的定义，适当的组织，还要设有防止破坏和不正确使用的保护措施。同时，数据库设计员在了解用户对数据的需要时，也必须取得用户部门的协助。

在计算机领域已变得如此复杂的现在，回忆一下狄更斯时代是很有趣的。在那个时代，数据处理是由一个戴着大礼帽，高高地坐在长腿椅上，手里拿着鹅毛笔的办事员来完成的。他面前放着一大叠装订得很好的厚厚的账簿。当接到一批货物的订单以后，他要对此笔交易进行全面的处理。他要查阅库存账，看看有没有库存现货可以满足订单上的要求，或者有哪些是必须加工的。他要把订单登记上账。而当发出一批货物时，要核减库存账，给用户开发票，在分户账目相应的户头上进行登记。这些都是很容易理解的简单过程。

假如有人要了解营业状况，查询某项库存，或者查询一笔客户欠款时，这个办事员可以将他的台账翻到有关的账页上，并立刻作出回答。在今天，人们可以设想这样的一个职员，他是用电话来承办订货并且回答查询。他在每天工作结束时，也要结算账目，如果需要成本核算，他可能还要修改它们使之符合最新的情况。

不过，令人赞叹的狄更斯时代的办事员的工作方式，只能适用于相当小的公司。当公司发展扩大以后，账簿数量不断增加，就需要有若干个办事员才能照管过来。业务分工使工作变得轻松些，一个办事员管库存账，而另一个办事员专管

开发票等等。本世纪初，采用了许多机械化的方法。为了有效地使用这些机械化方法，工作可以分批进行。例如，可以把数百项交易组成一批，这些交易中的某一些计算任务可以由一个办事员或一台计算机器来进行，而下一项任务则由另一个办事员或另一台计算机器来进行。

当采用穿孔卡片计算时，特大批量的处理就变得更加经济了。把大量卡片送到一台卡片机中予以处理，处理完以后才能改换机器进行下一项任务。同样，使用计算机磁带时，必须先用一个程序对大量的磁带文件进行处理，完了以后才能为下一种文件进行分类准备。按这种方式进行工作，就会失去往昔办事员式的工作方式的灵活性，再不能对一笔笔交易作单独的处理，也无法对一项账目情况、一个客户的信用卡结存数或一项货物的库存数等查询并迅速作出回答。如果进行这类查询。至少要经过一周或更长的时间以后，才能回答。当要把一些项目过账时，计算机必须依次扫描文件中所有的各项，才能找到要修改的那些项，而且，所有各项不论修改与否，都要在一个新的磁带上重写一遍。

把数据硬性地按应用分成许多单独的文件，再进行成批处理，并不是理想的操作方式。但对管理部门来说，这种方式可以控制本单位所有最新信息，并能随时提供给他们，这已经是非常方便了。由于数据处理技术的习惯性，管理部门也不得不对这种顺序的方式妥协，久而久之，就成了认可的操作方式，而很少再去考虑它是否适用了。

采用数据库就象有了一个持有多多种应用数据的极为聪明敏捷的狄更斯式的办事员。“他”可以把他的账簿编排得很好，只需要最少的书写量，并能够迅速地查阅账簿，以回答任何可能出现的询问。“他”同他的狄更斯时代的前辈大不相

同。那个缓慢的办事员只能在一本账簿上读写数据，而数据库系统却可以迅速查找一组组的数据，可以迅速将分散的数据收集在一起以对各式各样的信息查询作出回答。“他”确实是经营管理的“天使”。

公司经理常常需要跨部门或跨行业（如工程、财务、人事、生产和市场方面）的信息。他们需要市场决策对人事安排的影响、新的销售策略对生产的影响、或大的销售额与劳务费用的联系等等的信息。如果各部门只有各自的处理操作，那么计算机在回答上述问题时便无能为力。而类似狄更斯式的超级办事员的数据库则可以迅速地从一个部门的账目到另一个部门的账目中去检索有关的数据。存储的数据是集中的，而结构是一致的，因此，可供各部门共同应用。

公司或其他机关都越来越把数据当做公司经营中必不可少的基本资源。因此，对数据来讲，也应该象其他基本资源的管理一样，需要有专门的管理部门和机构。在已计算机化的公司中，在生产管理、市场销售、财务管理、计划编制及其他职能方面有效地使用数据变得如此重要，以致使其对公司在激烈竞争的市场中的发展和生存，有着巨大的影响。在政府各部门中，或者在拥有大量数据的现代化工厂中，狄更斯式的超级办事员将有效地使用数据、方便而准确地控制机器、或防止懒散的官僚主义造成的劳动力浪费等方面起着重要作用。

三、数据的价值

数据的价值根据数据能否发挥作用而定。目前，经常有这样的事情发生：在一项重要决策制定的过程中，所需要的数据以计算机易读的形式存放在某个单位的某处计算机中，但当决策制定人员需要时，却无法得到它们。此外，由于数