

TUSHU GUAN ZIDONGHUA  
YU YINGYONG BIDU

图书馆自动化与应用必读

汪金蕊 吴继伟 编著

青岛海洋大学出版社

# 图书馆自动化与数据库应用必读

汪金蕊 吴继伟 编著

青岛海洋大学出版社

(鲁)新登字 15 号

图书馆自动化与数据库应用必读

汪金蕊 吴继伟 编著

\*

青岛海洋大学出版社出版发行

青岛市鱼山路 5 号

邮政编码 266003

新华书店经销

大众日报社第二印刷厂印刷

\*

1997 年 10 月第 1 版 1997 年 10 月第 1 次印刷

32 开本(787×1092 毫米) 9 印张 191 千字

印数 1—3100

ISBN 7-81026-907-0

G·100 定价:12.00 元

# 目 录

<b>第一章 图书馆自动化系统概述</b> .....	(1)
第一节 图书馆自动化发展概况.....	(1)
第二节 我国图书馆自动化发展概况.....	(6)
<b>第二章 电子计算机在图书馆管理自动化中的应用</b> .....	(9)
第一节 电子计算机及其硬件设施.....	(9)
第二节 电子计算机软件及其在图书馆自动化中 的应用 .....	(25)
<b>第三章 图书馆自动化系统</b> .....	(37)
第一节 图书馆自动化的范畴 .....	(37)
第二节 图书馆计算机自动化管理系统 .....	(39)
第三节 图书馆自动化操作系统及其应用 .....	(40)
<b>第四章 图书馆自动化的数据库建设</b> .....	(61)
第一节 图书馆数据库 .....	(61)
第二节 图书文献管理工作的内容和图书文献的 著录 .....	(63)
第三节 图书文献的分类和主题标引 .....	(69)
第四节 机读目录 MARC 的格式 .....	(76)
第五节 机读目录 MARC 的字段和应用 .....	(91)
第六节 规范化机读目录的编制.....	(213)
<b>第五章 图书文献的计算机情报检索</b> .....	(221)
第一节 计算机情报检索原理和研究对象.....	(221)

第二节	计算机情报检索技术简介·····	(223)
第三节	联机情报检索的方法·····	(226)
第四节	光盘检索技术·····	(233)
<b>第六章</b>	<b>ILAS 系统在图书馆的应用</b> ·····	<b>(237)</b>
第一节	ILAS 系统发展概况 ·····	(237)
第二节	ILAS 各子系统功能简介 ·····	(241)
第三节	读者怎样使用 ILAS 检索子系统 ·····	(255)
<b>参考文献</b>	·····	<b>(283)</b>
<b>后记</b>	·····	<b>(284)</b>

# 第一章 图书馆自动化系统概述

本世纪 40 年代第一代电子计算机的问世,标志着人类进入了技术革命的全面发展时期。新技术革命的一个突出特点,是利用现代物质手段延伸和强化人的脑力和神经活动,使人利用信息、交换信息、处理信息的能力产生飞跃。据英国科学家詹姆斯·马丁推测,在 19 世纪,人类科学知识的拓展每 50 年增长一倍。世界每年出版的图书有 70 多万种、期刊 15 万种,其他文献资料 400 万件,而且文献的增长还在继续,20 世纪世界每年处理的信息多达几十亿。而对日益膨胀的文献量,靠传统的工作方法显然已无法对付。因此,图书情报部门采用现代电子计算机来提高处理文献信息的速度;采用缩微、复制和光盘技术来提高贮存密度;采用现代通讯手段和建立网络来扩大传播范围。如今图书馆自动化已经是图书情报部门的十分重要和迫切的事情了。

## 第一节 图书馆自动化发展概况

图书馆自动化源于 1936 年,美国的拉尔夫·帕克在得克萨斯州立大学建立了霍勒瑞斯穿孔卡系统,进行图书流通管理,后来,又把这个系统用于连续出版物的记到和管理。1942 年,美国新泽西州公共图书馆安装了两台在穿孔卡片上自动记录每份出版物的登记机。1950 年,美国国会图书馆开始使

用穿孔卡片系统作为存贮和检索图书情报的手段,现在,穿孔卡片仍是将图书情报信息转换成机读数据形式的重要手段。

1946年第一台计算机ENIAC(Electronic Numerical Integrator and Computer)在美国出现,这台机器名叫“电子数值积分和计算机”,是由电子管制造的,一共用了18000个电子管、1500个继电器以及其他部件。它的总体积约 $90\text{m}^3$ ,占地 $170\text{m}^2$ ,总重量达30吨,耗电功率为140kW,真可谓庞然大物。此后计算机研究发展很快,随着第二代、第三代计算机的产生,现在小型机、大型机、多媒体计算机均已投入市场。

电子计算机的诞生,促进了图书馆自动化的研究和发展。尽管初期的计算机笨重且运算不稳定,但是,计算机从技术上为图书馆自动化提供了可能。1954年,美国海军兵器中心图书馆使用IBM型计算机实现了单元词组配检索,成为最早使用计算机的图书馆系统。1957年,美国人H. P. 卢恩(H. P. Luhn)完成了以关键词的统计处理为基础的自动文摘和引文索引的编制。1958年,美国密苏里大学图书馆采用IBM360/25型计算机建立了图书流通自动控制系统。1962年,美国加利福尼亚大学图书馆建成期刊管理自动化系统,该系统还可以编印全部馆藏目录、新书目录和未到新书目录。1964年,美国化学文摘社(CAS)建立了资料处理自动化系统,同时用书本和磁带两种形式发行化学题录等多种刊物;同年,美国国立医学图书馆建立了医学文献分析与检索系统(MEDLARS),文献标引采用医学标题表(MESH),每篇文献可以从13个角度查找,实现了资料加工与检索自动化。1965年,美国国会图书馆开始实施机器可读目录MARC(Machine Readable Cataloguing的缩写)计划,并于1969年3月正式向

世界范围发行。MARC 的研制工作共经历了两个阶段,第一阶段是 1966 年 11 月至 1968 年 11 月,这期间研制试验的是 MARC I 格式,并在哈佛大学图书馆、国家农业图书馆、耶鲁大学图书馆等 16 个各种类型图书馆内进行了试验。这一阶段称之为试验阶段。第二阶段是从 1968 年 11 月至今,这期间在原来的基础上,形成了目前所看到的 MARC II 格式机读目录磁带。1973 年国际标准化组织(ISO)将 MARC II 审定为国际标准。这是在图书馆自动化中影响最广,也是最大的一项工作,它把手工书目著录转换到磁盘上,供世界各国共享,这大大推动了世界各国的图书情报自动化工作的发展。这一期间,工业较发达的国家,如日本、苏联、西德、意大利、法国、澳大利亚等国家的图书情报自动化系统也纷纷出现。这一阶段是科学技术大发展时期,同时也为图书馆自动化提供了良好基础,促使图书馆自动化系统进一步完善并开始走向广泛的实用阶段。

随着第三代、第四代电子计算机的问世,科学技术中的一部分——信息处理变得十分活跃,该时期被称为信息处理阶段。这期间研究的内容主要有微电子学、激光、射流、仿生学、数理逻辑、通讯理论、计算机结构、程序理论、信息系统等,对这些理论的研究和实践,产生了软件技术;计算机分时技术由实验阶段也进入了实用阶段,出现了分时处理系统 TSS (Time-Shared System)。这种属于软件技术的操作系统不但能够进行大量的数据处理,而且为各种工业部门的自动生产实现了过程控制。此时图书情报检索已由批式处理向联机检索发展,也就是把人和计算机以会话的方式联结起来,计算机从终端引入提问,并立即回答。此过程中的最大特点是读者



直接同计算机会话,并可以反复进行提问检索。人类开始把人和计算机的关系以会话的方式联接起来,并向大型的网络化方向发展。这时,一些系统如美国的 NASA 系统、DIALOG 系统、ORBIT 系统、BRS 系统、MEDLINE 系统、欧航天局的 RECON 系统等,都在原有的基础上形成了跨国跨州的国际网络。其中,DIALOG 系统和 TYMNET 网络具有一定的代表性。网络的形成,使图书馆相继出现了联机图书流通系统、联机期刊管理系统、联机图书采购系统等,也使图书情报自动化进入了一个联机处理阶段。联机处理进一步使图书情报自动化进入综合系统编目,它融编目、协调采购、馆际互借、参考咨询等管理业务为一体,形成一个综合自动化网络中心。其中网络组织比较成功的例子是美国俄亥俄大学图书馆中心 OCLC(Ohio College Library Center),它已成为这种全功能的协作网络中心。网络化可以使网络内的情报资源得到协调和平衡,达到资源共享,可以满足不同地理环境下的用户对图书情报的需求。网络化是图书馆自动化发展的高级阶段,也是我国以及全世界追求的目标。

目前国际联网已成为新一代人们获取文献信息的方式。国际联网(Internet 以后简称英特网)起源于 20 世纪 60 年代后期,它是美国国防部高级科研计划局(APPA)首先创立的。早期的 APPA 网络只是一个松散的计算机通讯网络,到了 80 年代中期美国全国科学基金会投资,在 APPA 网络技术的基础上,建立了全国科学基金会网络(NSFNET),使网络成为可以提供像巨型计算机或专业软件那样的资源,并联结着许多小型网络的横跨全国的超大型网络。英特网是国际性的网络,它是用以连接比其层次低的网络的网络。现在它已经把全

球成千上万的网络连接在一起,连接于其上的计算机可互通电子邮件。如今使用英特网的用户以每月 12% 的速度在增长。1983 年美国国防部将英特网分为军用与民用两部分。民用部分划规 NSFNET, 主要供科研教学使用, 1989 年正式改名为英特网(Internet)。现在美国的大学生成为最积极使用英特网的用户, 他们可从英特网这一巨大的信息库中获取无穷无尽的知识。1995 年初, 全世界七个最发达的工业大国聚会布鲁塞尔, 共同制定开发全球信息高速公路的大计, 会议制定了多个合作项目。在图书馆的项目规划中所提出的设想是: 旨在创建一座全球性的电子图书馆, 可通过多媒体技术进行检索, 并鼓励各国共同利用已转成软件的藏书。

另外, 美国七大有线电视服务商之一的琼斯交互有线电视公司, 最近也推出了作为信息高速公路上的全球电子图书馆模式。全球电子图书馆的基本设想是把全球各地大型图书馆内已转变为数字信号的收藏内容汇总起来, 借助于全球最大的计算机网络——交互网络, 逐步使其成为地方小型图书馆以及家庭都能享用的信息资源。

为了加速信息传播的速度和扩大传播的范围, 信息高速公路的设想开始产生。自从美国首先提出建设信息高速公路以来, 世界各国纷纷响应, 日本派出科技厅、通产省和邮政省等 4 个省厅联合考察团赴美国考察, 得出的结论是, 克林顿政府提出的信息高速公路计划将给美国经济带来新的信息文明与繁荣。日本政府立即采取行动, 提出了日本的“信息高速公路”计划——曼陀罗计划(Mandara 计划), 决心用 22 年的时间(即至 2015 年)使日本成为第一流的信息大国。1992 年 10 月, 法国、德国、英国、意大利、西班牙 5 国提出了泛欧网络计

划(Global European Networks),简称 GEN,并于1995年开始过渡到经营型的欧洲传输网。1993年,加拿大政府已投资7.5亿加元,建设“信息高速公路”,并于1994年4月19日宣布了由30人组成的“公众咨询顾问委员会”,同时,还准备成立一个名为“加拿大信息传播公司”的全国性机构,负责组织管理全国“信息高速公路”的实施。

信息高速公路的建设将使信息流通量呈几何级数增加。信息高速公路是集计算机技术、通讯技术、网络化技术为一体的信息技术集成,它将使图书馆的开放面向世界公众。图书馆的馆藏文献载体将由现在的以纸为主变为以电子载体为主。大部分馆藏将由数字数据库和模拟数据库组成。

## 第二节 我国图书馆自动化发展概况

随着发达国家计算机情报信息管理科学的发展,我国从70年代中期以后,也开始了研究和试验计算机在图书情报工作中的应用。1974年8月,中央批准了“汉字处理信息工程”的研制工作,简称“7482工程”。主要包括汉字通讯、汉字排版、汉字情报检索的计算机应用软件、主题词表及机器翻译等。我国最早研制计算机软件的单位是中国科学院计算技术研究所和第一机械工业部技术情报所。中国科学院计算技术研究所于1973年开始用国产111型计算机的汇编语言编写情报检索QJ111程序,1975年投入运行,1976年与中国科学院图书馆合作试编了机读目录数据库,输入文献5000余篇,图书300余册,除进行成批处理SDI服务外,还进行了编制藏书目录和新书通报等试验。1979年,中国科学院图书馆与

有关单位协作编写了定题服务应用软件 SDI-111 系统,同时与南京大学等单位协作编制了载有我国全部国家标准的国际交换用磁带,在联合国科教文组织验收后为 ISO 组织所接受。1979 年上海交通大学建立了检索系统。1979 年 4 月,一机部情报所、四川省机械局研究设计院和情报站、武汉大学图书馆学系、陕西机械学院经济系协作在国产 TQ-16 型电子计算机上设计了 QJWS-2 检索程序。南京大学数学系情报检索教研室 1978 年为本校图书馆设计了 NDTs-78 西文图书检索系统。1979 年,北京图书馆、北京大学图书馆和图书馆学系、清华大学、中国图书进出口公司、中国科学院图书馆共同协作引进美国 MARC 磁带,开始对国外机读目录进行研究。1982 年武汉大学图书馆学系与湖北省情报所在 EC-1022B 计算机上研制了 SDI 系统,1984 年该系现代技术教研室在 MC68000 微机上研制了图书情报管理系统和流通系统,在图书编目、通报、检索、流通、统计、打印等方面,具有较强的处理功能。1985 年 4 月 18 日中国科学技术情报所重庆分所开始实现联机国际检索。同年 4 月 27 日辽宁省情报所国际公用数据终端投入使用,能查找世界 400 多个数据库的信息资源。目前我国图书馆正在普及自动化,1995 年以后,我国图书馆界在图书馆自动化方面开始大踏步前进,全国图书馆大力普及计算机在图书馆的应用。在网络建设上,有些地区正在致力于校园网、地区网的建设,有些图书馆已和国际上联网。在软件上我国为了加速文献信息网络建设引用了 MARC 格式(Machine Readable Catalogue 机读目录),根据我国国情建立了 CNMARC 格式(China MARC),现在正在进一步研究完善 CNMARC。CNMARC 在图书馆的应用,使我

国图书馆计算机的机读目录趋于规范化、标准化,为进一步实现全国图书馆情报信息的联网,并争取同国际上联网奠定了基础。如今美国经营英特网的斯普林特公司同我国国家电信总局签约,将把英特网延伸入我国,以支持我国的金桥工程。这样我国将有 1000 多个大学和科研单位可以联结到网上,将有 100 万人能使用到英特网。由于全国网络化的发展,我国图书馆网络也随之发展,从一个局域网、一个系统发展为多个子网、子系统。

近年来随着我国通信事业和计算机联网的迅速发展,我国正在为 CERNET 的建设做准备。CERNET 是中国教育和科研计算机网示范工程(China Education and Research Network)的英文简称。该网络的建设工作在国家教委主持下,由清华大学、北京大学等十几所高校承担。其总体建设目标是利用先进实用的计算机技术和网络通讯技术,把全国大部分高等学校连接起来,并与 INTERNET 网互联。CERNET 的网络中心设在清华大学,地区网点分别设在北京、上海、南京、西安、广州、武汉、成都、沈阳。目前利用国家公用数据网联接 CERNET 网络中心和清华大学、北京大学等地区网点的 CERNET 业已开始工作。

总之,图书馆自动化从 50 年代穿孔卡片系统开始,到 60 年代出现了一个脱机处理阶段,而后相继出现了联机图书馆系统、联机期刊管理系统、联机图书采购系统等联机处理阶段,它融图书馆自动化各子系统管理为一体,形成了一个综合自动化网络中心。如今网络系统正迅速发展,信息高速公路正在建立,国际联网正在实施,图书馆自动化也正向电子图书馆阶段发展。

## 第二章 电子计算机在图书馆管理 自动化中的应用

### 第一节 电子计算机及其硬件设施

#### 一、电子计算机的基本概念与术语

##### (一)数据和信息

从键盘上输入到计算机中需要处理的数字、字母、符号和文字等称为数据；经计算机处理后输出的数字、字母、符号和文字则称为信息。数据是表示信息的，但并非任何数据都能表示信息，信息是经过处理后的数据，它是客观现实世界的反映，而数据则具有任意性，用不同的数据形式可以表示同样的信息。

##### (二)数据处理

数据处理是指将数据转换成信息的过程。它包括对数据的收集、存贮、传送、检索、分类、加工及输出等一系列的过程。数据处理的作用是从大量的、杂乱无章的数据中整理出有价值、有意义的信息。因此，我们可以用以下公式表示数据处理的作用：

$$\text{数据} + \text{处理} = \text{信息}$$

数据是“原材料”，信息是“成品”，而计算机则是一个“处理”数据、“生产”信息的工厂。有时，当两个以上的数据处理过

程前后相继时,信息和数据的概念就会产生交叉,表现出数据和信息的相对性。如图 2-1 所示。

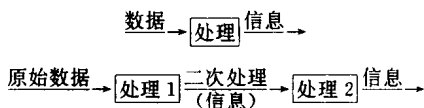


图 2-1 数据与信息的关系

### (三)位、字节和字

计算机内只使用由 0 和 1 两个字符表示的二进制数,二进制数中的每个 0 或 1 就是信息的最小单位,称为二进制的“位”(bit)。

在计算机中,若干个位组成一个字节(byte)。字节应该由多少位组成,这取决于计算机的结构。通常,微型计算机中一般多由 8 个位组成一个字节,即由 8 个位组成的一个字节可以代表一个字符(例如一个英文字母)。此外,十进制数 0~9 用二进制表示时仅占用 4 位,而一个字节是 8 位,有两个 4 位,所以一个字节最大可以表示一个 2 位的十进制数 0~99。

在计算机存储器中占据一个单独的地址并作为一个单元处理的一组二进制数位称为“字”(Word)。字指的是数据字,它由若干个位或几个字节组成。比如一个字含 8 位、16 位、32 位等等,或者说一个字可以包含 1 个、2 个、4 个字节等等。

对于计算机的控制器来说,一个字或数个字是一条指令;对于计算机的运算器和存储器来说,一个字或数个字是一个数据;在处理数据时,字是被当作一个数值整体来处理的。换句话说,计算机在同一时间内所处理的一组二进制数称为“一个计算机字”。每个计算机字所包含的二进制的位数叫“字长”。字长是计算机的重要指标之一,字长越长,说明机器所能

处理数的有效位越多,该计算机的精度越高。一个字由多少位组成,这取决于计算机的体系结构,由计算机内部的累加器、存贮器和主要寄存器的存储能力来决定。

#### (四)指令和程序

我们将数据输入到计算机中,或要求计算机对数据进行某种处理,都必须按一定的规范进行操作。每一个基本的操作步骤称为指令,指令又可称为命令。有时,为完成一次处理需要许多指令共同组成一个序列,这个序列就称为程序。

当我们把要处理的数据(信息)输入到计算机时,作为实现这一处理的程序也同时要输入到计算机中,计算机在执行程序时不需要人来干预,它能自动进行信息处理,但处理的方法是由人选择的,程序是由人来编制的。因此,计算机的处理仍是按照人的意愿来进行的。

#### (五)硬件与软件

计算机的各个部件和设备都是由元器件构成的有形物体,这些有形物体被称之为计算机的硬件;而计算机的软件是指用来指挥计算机运行的各种程序的总和,以及开发、使用和维护这些程序所需要的技术资料。

### 二、计算机硬件的组成

从微型计算机的硬件构成上看,它主要由中央处理器、存贮器、输入设备、输出设备及接口总线等组成。

#### (一)中央处理器(CPU)

运算器和控制器一起组成了计算机的核心,称为中央处理器,即CPU(Centrl Processing Unit)。

运算器是对信息进行加工处理的部件,它在控制器的控制下与内存交换数据,负责进行各类基本的算术和逻辑运算。



在运算器中,还含有能暂时存放数据或结果的寄存器。

控制器是整个计算机的指挥中心,它负责协调计算机各部件的工作,保证数据、信息的运算能按预先规定的目的和步骤有条不紊地操作及处理,确保系统自动进行运算。

CPU 与内存贮器合在一起,再加上输入输出接口就叫做主机,这是微机的主要组成部分。

近几年来,CPU 型号不断出新,各项指标越来越高。目前市场上微机主要的 CPU 型号有 8088,80286,80386,80486,80586 等。

## (二)存贮器

微型计算机的存贮器专门用来存放处理程序和待处理的数据,也可以存放运算完的结果。

存贮器是计算机的记忆装置,用于存放原始数据、中间数据、最终结果和处理问题的程序,存贮器内的信息是按地址存取,向存贮器存入信息也称为“写入”。写入新的内容则覆盖了原来旧的内容。从存贮器里取出信息,也称为“读出”。信息读出后并不破坏原来存贮的内容。因此,信息可以重复取出,多次利用。

计算机的存贮器分为内存贮器和外存贮器:

1. 内存贮器。内存贮器又称为主存贮器,因装在主机机箱内,故称为内存贮器,简称内存。内存存取信息的速度快,但价格比较贵。早期的内存主要是磁芯存贮器,现已被体积小、速度快的半导体集成电路存贮器所代替。

内存贮器根据其功能又分为只读存贮器 ROM (Read Only Memory) 和随机存取存贮器 RAM (Random Access Memory) 两种。