

59

G250.76  
W31

# 数字图书馆

王大可 编著



A0999529

海天出版社

## 图书在版编目(CIP)数据

数字图书馆/王大可编著. — 深圳:海天出版社,  
2002.2

ISBN 7 - 80654 - 641 - 3

I. 数... II. 王... III. 数字图书馆 IV. G250.76

中国版本图书馆 CIP 数据核字(2001)第 098217 号

海天出版社出版发行

(深圳市彩田南路海天大厦 518026)

<http://www.htph.com>

责任编辑:陈丹 封面设计:王晓珊 张浩

责任技编:王颖

海天电子书开发公司排版制作 电话:2720730

深圳彩帝印刷实业有限公司印刷 海天出版社经销

2002年2月第1版 2002年2月第1次印刷

开本:850mm×1168mm 1/32 印张:7.375

字数:240千 印数:1-5000册

定价:22.00元

海天版图书版权所有,侵权必究。

海天版图书凡有印装质量问题,请随时向承印厂调换。

## 前 言

当今世界正在迅速地进入信息社会。自1992年美国克林顿政府提出建设“信息高速公路”计划以来,信息浪潮波及全球,以发展信息产业和信息技术为目标而形成新的高潮,其影响之大、范围之广、投资之巨、发展之快史无前例。

图书馆作为信息战线上的服务机构,也在由传统图书馆向现代化图书馆过渡,其面貌发生了极大的变化。出现了“电子图书馆”、“虚拟图书馆”、“网络图书馆”、“数字图书馆”等新的概念。虽然其提法各异,描述的内容略有区别,但就其本质来讲都是一样的,只是从不同的侧面反映而已。一般地,人们喜欢用“数字图书馆”作为其统称。

90年代以来,数字图书馆理论与技术已成为图书情报学界的一个国际性热点论题,并成为美国、英国和日本等发达国家图书馆高新技术应用的一个新领域。西方国家陆续开始了一些试验性或示范性数字图书馆项目的研究,我国也有几个数字图书馆研究计划或试验项目开始启动。

数字图书馆是未来图书馆的发展方向,这已经为图书馆界普遍接受,基本形成共识。但数字图书馆到底应该是一种什么样的形态,在实际环境中如何建设与实现,却见仁见智、众说纷纭,许多学者提出了富有创见的新思想、新观点,在许多方面探

索着、实践着,这充分反映了数字图书馆在图书馆学研究和工作中的生机与活力:

本书是作者根据自己参加“中国试验型数字图书馆”项目以来,分别对美国、英国、法国、德国、澳大利亚等国家进行技术考察,走访 30 余所图书馆及有关数字图书馆项目研究部门后,对国外数字图书馆的发展现状进行了较全面的调研,在收集大量资料的基础上,结合国内图书情报界同行们提出的许多观点,分析目前国内一些单位或个人所开展的数字图书馆研究工作的进展和存在的问题,并根据自己的研究工作实践和体会编写而成。

本书试图简要地叙述数字图书馆的基本理论和实现模式,并通过一些应用实例,以帮助正在研究和实践数字图书馆建设的领导和实际工作者思考和借鉴,故着力于知识性与实践性相结合:

众所周知,数字图书馆是一个极泛的大概念,它并不是指某一个图书馆、一个资源库或一个应用软件系统,它所涉及的面十分广,要研究的内容极其丰富,甚至有些概念还没有严格的定论。本书不着力于对这些理论性问题进行深入的探讨,而只是把一些概念提出来。全书重点围绕如何构筑数字图书馆系统这个主线而展开,因而书中侧重于对数字图书馆系统的信息体系结构、信息组织方法、系统构成模式及实现技术进行介绍。

本书共分为 10 章。第一章简述图书馆信息技术的变革,第二章给出数字图书馆的产生背景,数字图书馆的定义、功能与特征。第三章介绍数字图书馆的信息体系结构及信息处理的一些关键技术,并给出一个典型的数字图书馆信息体系结构的实际例子。第四章至第五章,主要介绍数字图书馆系统的设计与开发技术,数字图书馆资源建设的方法和建库工作中应注意的一些问题。第六章给出数字图书馆资源建设中将用到的一些标

准,同时通过实例介绍元数据格式的定义和不同格式数据的转换方法等。第七章通过引入知识产权保护的一些基本概念来说明数字图书馆建设中应注意到有关版权方面的约束和合理应用问题。第八章介绍国外数字图书馆的发展概况。第九章简述我国数字图书馆的研究现状,结合我国的国情,提出我国数字图书馆建设的发展战略。第十章简述数字图书馆的发展与未来。

由于数字图书馆是近年来才开始的新的研究课题,其研究的基本理论和相关内容尚不够成熟,加上自己对该课题的研究也不够深入,收集的资料不够全面,故在编写此书时感到力不从心。书中肯定会有不少缺点和错误,恳请从事该项研究工作的同志们和广大读者予以指正。本书在编写过程中参阅了许多相关文章、材料和专著,引用了一些同志的研究成果,特别是得到了深圳图书馆领导和有关同志的关心和大力支持,在此一并表示谢意。

王大可

2000年8月于深圳

# 目 录

第一章 图书馆信息技术的变革·····	( 1 )
第一节 信息技术与图书馆·····	( 1 )
第二节 图书馆信息技术的应用与发展·····	( 7 )
第二章 数字图书馆的基本概念·····	( 17 )
第一节 数字图书馆的产生·····	( 17 )
第二节 数字图书馆的有关定义·····	( 22 )
第三节 数字图书馆的特征·····	( 34 )
第四节 数字图书馆的功能·····	( 39 )
第五节 数字图书馆引起的变化·····	( 42 )
第三章 数字图书馆的体系结构·····	( 52 )
第一节 数字图书馆的信息结构·····	( 52 )
第二节 数字图书馆系统的结构模型·····	( 56 )
第三节 信息处理的关键技术·····	( 65 )
第四节 一个典型实例·····	( 70 )
第四章 数字图书馆系统设计与开发工具·····	( 93 )
第一节 系统设计的基本原则与模式·····	( 93 )

---

第二节	系统的构成	(97)
第三节	系统的信息组织方法	(105)
第四节	系统设计的相关技术	(108)
第五节	IBM 数字技术开发平台	(114)
第六节	Tamino 数据库技术开发平台	(121)
<b>第五章</b>	<b>数字图书馆的资源建设</b>	<b>(136)</b>
第一节	馆藏文献数字化的意义	(136)
第二节	我国文献数字化建设工作的进展	(138)
第三节	文献数字化中的关键性问题	(141)
第四节	我国文献数字化工作面临的问题及解决方案	(146)
<b>第六章</b>	<b>标准及应用</b>	<b>(150)</b>
第一节	机读目录格式——MARC	(150)
第二节	都柏林核心——Dublin Core	(155)
第三节	超文本置标语言——HTML	(163)
第四节	标准通用置标语言——SGML	(171)
第五节	可扩展的置标语言——XML	(172)
第六节	应用举例	(174)
<b>第七章</b>	<b>知识产权</b>	<b>(182)</b>
第一节	知识产权的基本概念	(182)
第二节	知识产权保护的主要原则	(186)
第三节	著作权制度	(189)
第四节	数字图书馆建设中的著作权保护问题	(198)

第八章 国外数字图书馆的发展概况·····	(212)
第一节 美国数字图书馆的发展状况·····	(212)
第二节 日本数字图书馆项目研究状况·····	(220)
第三节 英国的数字图书馆试验计划·····	(226)
第四节 法国数字图书馆项目实施计划·····	(229)
第五节 G8 数字图书馆联盟计划·····	(231)
第六节 其他国家的数字图书馆项目·····	(235)
第九章 我国数字图书馆建设·····	(240)
第一节 我国数字图书馆研究现状·····	(240)
第二节 我国数字图书馆建设的必要性·····	(252)
第三节 建设数字图书馆的条件·····	(257)
第四节 数字图书馆建设的具体做法·····	(260)
第五节 对我国数字图书馆建设的思考·····	(264)
第十章 数字图书馆的发展与未来·····	(275)
第一节 数字图书馆建设中需要解决的问题·····	(275)
第二节 数字图书馆未来的发展趋势·····	(279)
主要参考文献·····	(283)



## 第一章

# 图书馆信息技术的变革

### 第一节 信息技术与图书馆

回顾近代科学发展的历史，人类经历了两次技术革命。第一次是从18世纪60年代到19世纪中期的技术革命，其主要标志是蒸汽机的发明和应用；第二次是19世纪70年代第一次世界大战期间的技术革命，其主要标志是电力和电机应用。这两次技术革命大大促进了人类文明的发展，同时也开创了图书馆工作与技术相结合的历史，使图书馆的局部工作实现半机械化半自动化，以减轻图书馆员的劳动强度，提高工作效率，加快图书文献的流通。

进入20世纪40年代以来，一个以原子能工业、信息技术、空间技术和生物工程等为代表的第三次技术革命出现了。许多新技术、新理论、新科学出现了，这标志着一个新的技术时代来临，世界正在出现一个以信息技术为主导的新技术革命浪潮。新的技术革命迅速地在改变着世界的面貌，使工业、

农业、交通、能源、金融、教育、科技以及其他领域发生着巨大的变化。在新技术革命的挑战面前，世界上许多国家的政府、专家和学者都在紧张地酝酿和采取对策。现代信息技术正在使人类从工业化社会向信息化社会发展，信息已和自然资源一样成为人类不可缺少的宝贵的资源。

90年代以后，随着计算机科学技术的迅猛发展，国际信息高速公路的开通和网上信息资源的开发利用，促使图书馆现代信息技术迈向新的领域。时代在呼唤着图书馆的伟大变革，新技术革命为图书馆实现现代化带来了新的技术手段。图书馆正朝着网络化、电子化和数字化的方向发展，图书馆信息技术的发展出现了“质”的飞跃，以至提出了“无围墙的图书馆”、“电子图书馆”、“网络图书馆”、“虚拟图书馆”和“数字图书馆”的概念，并正在逐步成为现实。

## 一、图书馆传统信息技术

### 1. 印刷与排版技术

迄今为止，人们所接触的各种书籍、报刊等纸质印刷物，是图书馆保存和传递信息的主要载体。其信息大都是以排版印刷形成的。排版印刷术的发明是中国人民对世界文明的一大贡献。印刷术从发明至今，在各方面都取得了巨大的发展，印刷的设备越来越精密，印刷的速度越来越快，印刷的质量越来越好，印刷的范围越来越广。印刷是人类文化发展和信息交流的有力工具，是促进社会文明发展的重要手段。

过去的印刷大都采用铅字排版印刷技术，这种活字排版技术在我国流传了上千年。

近年来，随着电子计算机排版和输出技术的发展，使印刷

技术出现了一次飞跃性的变革，由我国发明的计算机汉字排版技术开始大规模投入使用。近年来出现的轻印刷系统逐渐代替了传统的铅版印刷，计算机系统中的喷墨打印机、激光打印机也大量投入应用，给图书馆现代文献信息的传播和利用提供了极为方便的手段。

## 2. 复印技术

复印技术即通常人们所指的静电复印技术，它作为传递信息的重要手段之一，已在图书馆信息技术中扮演一个重要角色。

静电复印是利用某些材料的光电导静电特性对被复印稿件进行照相，并以复印品准确、迅速输出的一种图文复制技术。

静电复印曾被看作是印刷技术中的一个分支，因而也被称作“特殊印刷”或“无压印刷”。它在图书馆文献复印中的应用，改变了过去人们对所需文献内容的抄抄写写的繁琐工作，使得文献交换和传递速度大大加快。因此，复印技术的应用是图书馆信息技术中不可缺少的部分。

## 3. 缩微技术

缩微技术是以高密度的信息贮存、记录模拟影像为特点，以胶片为介质，运用光学摄影方法，将文献中的文字、图像按比例缩小，加工成为缩微品。要利用文献时，一般通过放大阅读器设备，按一定倍率放大，就可以进行阅读，也可以还原后进行复印输出。

缩微技术为图书馆大量珍贵的历史文献提供了保存和利用的技术手段，提高了文献收藏和利用的水平。缩微技术具有以下优点：

- (1) 信息存储密度高、体积小、节省藏书空间

缩微技术是一种信息密集贮存技术。图书文献经缩微后,可以减少 80% ~ 98% 的存储空间。因此缩微是目前缓解馆舍紧张行之有效的办法之一。

#### (2) 保存时间长

缩微品的使用寿命要比纸质印刷品长得多,用同样正确的方法保存,纸质印刷品几十年后就会老化损坏,而缩微品的寿命在保证了冲洗条件和规定的保存条件下,可以长期或永久地保存。

#### (3) 价格便宜

缩微品价格很便宜,同一出版物的缩微品要比纸型印刷品便宜得多,而且缩微版价格的上涨速度也低于印刷版。

#### (4) 便于交流

缩微复制品比印刷型复制品便宜,且体积小,图书文献缩微后,便于交流,有利于资源共享,并且既经济又迅速。

#### (5) 有利于图书文献检索

缩微摄影技术可以和计算机、光盘、复印等技术相结合,组成新的自动化检索系统,便于读者利用。

#### (6) 拷贝、还原简便

对缩微品的复制、放大、还原技术简单,操作极为简便。

正因为缩微技术在现代图书馆信息技术中具有如此重要的意义,所以,文献缩微技术的应用愈来愈受到人们的重视,开展这项工作 in 图书馆事业的发展中具有极其重要的意义。另外,在这里也要特别说明,缩微品并没有因为光盘的出现而失去其价值。国外实践证明,缩微品在保存的长久性和可靠性方面极大地优越于光盘。

## 二、图书馆现代信息技术

### 1. 计算机技术

计算机技术的应用是人类科学技术的最伟大发明之一。目前已扩展到科学技术和国民经济的各个领域，并且日益被人们广泛地使用。它的效能和作用，确实使人惊叹不已。

计算机技术的应用，已远远超出了科学计算的范围，而渗透到各门学科乃至日常生活之中。计算机技术在图书馆信息系统中的广泛应用，大大地促进了文献信息的收集、加工、处理、传播和利用，使图书情报现代化进入了一个新的发展时期。

### 2. 高密度存储技术

随着电子技术的高速发展，计算机高密度存储技术也发展得很快。在早期的计算机上，人们使用磁芯作为内存储器的主要元件，一般使用磁鼓、磁泡等设备做外存。但这些早期的存储设备，其存储密度低，读写速度也较慢。后来又出现了半导体存储器。当大规模集成电路计算机出现以后，硬磁盘、软磁盘广泛地投入应用，使计算机的存储技术出现一次新的飞跃。

80年代以来，随着计算机高密度存储技术的飞速发展，特别是大容量磁盘存储技术和光盘存储技术的应用，使计算机高密度存储技术在图书馆的应用迅速地出现了一个新的局面。人们预计，随着计算机高密度存储技术的进一步发展和在图书馆的应用，在不久的将来，所谓“袖珍式图书馆”、“手提式图书馆”将会成为现实。

### 3. 通信技术

从19世纪40年代到20世纪30年代,电磁技术广泛应用于通信领域。电报的发明,电话的出现,开始了近代电信事业,为迅速传递信息提供了方便。20世纪末,电子技术广泛用于通信网络、微波传输、电子多路通信网络、海底电缆以及卫星通信。从而形成了现代的通信网络。

通信技术与计算机技术相结合,通信网络的迅速发展,使得计算机的功能和作用发挥得淋漓尽致。今天,世界各地的计算机都可以借助通信网络相互进行信息传递,E-mail通信正在代替传统的书信、电报和电话。网上书目查询、全文检索,以至于多媒体信息的传递愈来愈普遍。现代通信技术和网络技术的利用,使得信息资源共享成为现实。

### 4. 多媒体技术

多媒体技术是指能综合处理多种媒体信息如文字、图片、图像、音频、视频和动画等,使多种媒体建立联系,并具有交互功能的信息处理技术。

#### (1) 多媒体技术的特点

①集成性,即将多种媒体有机地组织在一起,共同表达事物,做到图、文、声、像一体化。

②交互性,即人机交互,在播放多媒体节目时,能够进行人工干预,通过人机交换信息来完成任务。

③数字化,即多媒体中各类媒体都是以数字化的方式存储信息,便于计算机对各类信息进行统一加工、处理和传播。

#### (2) 支持多媒体的关键技术

多媒体技术综合处理多种媒体信息,全文文本,动态图像以及语音信息都要求有很大的存储空间,视频图像与音频信息

都具有连续性，这些都对多媒体技术提出了较高要求，其中最关键的技术包括：

① 数字压缩与还原技术；② 海量存储器；③ 多媒体操作系统；④ 数据库理论与方法；⑤ 支持多媒体的硬件与网络技术等。

### (3) 多媒体技术对图书馆的影响

多媒体技术将以信息表达直观、生动、逼真等优点而被广大读者所接受，它的出现和投入实用，将对图书馆传统的信息保存与提供方式、服务方式、工作方式等带来巨大的冲击。多媒体技术在图书馆的广泛应用，必将使图书馆信息的存储、开发、利用迈向一个新的阶段。

### 5. 数字技术

所谓数字技术，这是随当代科学技术的发展而迅速发展起来的一门新的技术，即指将各种采集到的信息，包括文字、图片、声音、图像、动画等原始信息，以数字的方式（在计算机系统中采用二进制数值 0 和 1 来表示）进行存储、加工、处理、传输，再经转换后输出。这种数字技术，目前已广泛地应用于社会的各个领域，如数字扫描仪、数字照相机、数字电视机、数字通信等。数字图书馆的研究工作，也在紧张的进行之中。人们预计，在不久的将来，数字图书馆将成为现实。

## 第二节 图书馆信息技术的应用与发展

本节主要介绍现代信息技术在图书馆的应用，包括国外图书馆现代信息技术发展概况及我国图书馆现代信息技术的应用

现状，同时提出了图书馆现代信息技术发展的三个阶段。

## 一、国外图书馆现代信息技术发展概况

现代信息技术在图书馆中应用，首先是从美国开始的。1950年，美国海军军械中心图书馆的H·E·泰利特提交了世界上第一篇建议图书馆应用计算机的研究报告。1954年，该馆首先在IBM701型计算机上建立NOIS检索系统，采用单元词组配检索，输出结果为文献号码。1958年，经过改进的NOIS系统使用IBM704型计算机，可以检索文摘、题目和作者等项目。NOIS检索系统的建立，开始了图书馆现代信息技术的新时期。进入60年代后，一些工业发达的国家都先后在图书馆应用计算机。1960年，美国人L·R·尼诺提出了编印目录卡片的机械化装置，被认为是最初的机读目录。1962年，美国加利福尼亚大学和南伊利诺斯大学图书馆分别应用计算机进行期刊和流通管理。1963年，原西德柏林大学和鸿波大学图书馆用计算机管理流通、目录编制和期刊登记。1964年，美国国家医学图书馆用计算机编制《医学文摘》，并生产MEDLARS数据库磁带。1966年，美国国会图书馆开始了MARC试验计划。同年，原西德国家图书馆开始用计算机编制全国书目；英国电气工程师学会开始用计算机检索；日本科学技术文献中心完成文献速报自动编制系统。1969年，美国国会图书馆的MARC II格式机读目录磁带公开发售，经过几年的修改，使之成为美国国家标准。

随着计算机技术的高速发展和通信技术的发展，到了70年代，图书馆自动化系统迅速由单机批处理系统发展到联机系统，并逐步发展成为计算机网络系统，各种不同规模、不同功



能的图书情报网络陆续出现。其中，资格最老，规模最大的要算美国俄亥俄学院图书馆中心的 OCLC (Ohio College Library Center) 系统。该系统创建于 1967 年，到 1972 年便发展成为美国图书馆界最有影响的服务系统。1977 年易名为联机计算机图书馆中心 (Online Computer Library Center)，并成立了 OCLC 公司，是当今世界上最大的图书馆书目数据库中心。直至 1999 年 4 月止，全世界已有 67 个国家的 35000 个图书馆使用该中心提供的书目记录或网上服务。在该联机编目中心，平均每 15 秒钟就增加一条新的记录，目前该中心的书目库中已有 4600 万条书目记录。除书目记录外，近年来，他们也开始建立文摘库、索引库和全文数据库，现已有 5400 万篇全文文献可提供给用户使用。同时，OCLC 通过联机网络，还收集了 1100 多所大学和专业刊物的电子版文献，包括全文和影像资料，利用网络向外发布、传递和提供使用。

类似这样的联机编目网络，全世界共有 10 多个，其他如英国的 BLAISE，BLCMLimited，SWALCAP，苏格兰的 SCOLCAP，荷兰的 PICA，瑞典的 LIBRIS，加拿大的 UTLAS，香港的香港图书馆网络，澳大利亚的 ABN 和日本的学术情报中心。

在美国西部还有另外两个较大的图书馆网络，一个是设在加州斯坦福大学的研究图书馆情报网络 RLIN (Research Libraries Information Network)，直接担负着向研究图书馆协会所属各高等院校图书馆和科研图书馆提供网络服务；另一个则是华盛顿州图书馆网络 WLN (Washington Library Network)，直接从华盛顿州立图书馆给北太平洋西北地区的各图书馆提供服务。现在国会图书馆已把该馆的计算机网络与上述三个大型网络系统联通起来，形成美国全国性的网络。

另外，世界上有名的几个大型情报检索系统也是在这一阶