

深入未来计算机网络

数字图书馆

建设理论与应用技术

Theory on Building Digital Library and Its Application Technology

蒋太岩 等主编

Theory on Building
Digital library
and Its Application Technology

高等教育出版社

PLA 2000-2001

数字图书馆

建设理论与应用技术

Theory and Technology of Digital Library Construction and Application Technology

李永强 等主编

Digital library

清华大学出版社

数字图书馆 建设理论与应用技术

蒋太岩 等主编

高等教育出版社

责任编辑：禹 婷

责任校对：隋 敏

装帧设计：王 野

图书在版编目 (CIP) 数据

数字图书馆建设理论与应用技术/蒋太岩 等主编

—北京：高等教育出版社，2003.1

ISBN 7-04-009498-5

I. 数… II. 蒋… III. ①图书馆学—工具书 ②高等教育—图书馆
教学参考书 IV. G31.11

中国版本图书馆 CIP 数据核字(2003)第 64331 号

数字图书馆建设理论与应用技术/蒋太岩 等主编

出版发行	高等教育出版社
社 址	北京市东城区沙滩后街 55 号
经 销	新华书店北京发行所
印 刷	北京利华印刷厂印装
开 本	850×1168 1/32
印 张	18.5
字 数	540 千字
版 次	2004 年 2 月第 1 版
印 次	2004 年 2 月第 1 次
印 数	1-1000 册
定 价	68.00 元

ISBN 7-04-009498-5/G·50

(版权所有 侵权必究)

编委会名单

主 编：蒋太岩 王 强 刘万国
 高效伟 姜曼丽
副主编：王 宇 隋 敏
编 委：王 强 王 宇 刘万国
 徐大叶 李大庆 高效伟
 孙 浩 陈 爽 徐殿军
 姜曼丽 蒋太岩 隋 敏

序 言

本书是渤海大学有关人员在实施辽宁省教育厅批准的重点科研项目“中美数字图书馆建设比较研究”课题的过程中,与承担国家数字图书馆建设 863 计划的北京超星数字信息技术有限公司的专家共同探讨数字图书馆建设最新成果的基础上著成此书。20 世纪 90 年代以来,随着计算机多媒体技术和网络通信技术的飞速发展,人们传统的获取、利用、传递、交流知识的方式开始发生革命性的变化。数字图书馆正是凭借一系列现代电子通信技术的成长和发展应运而生,并以其规模超大、覆盖面广、信息快捷、可以远程查询和检索等一系列独具特色的优势,向传统图书馆发起了挑战。十几年来,数字图书馆应用技术迅猛发展,日新月异,风靡全球。

建设数字图书馆,其实质就是通过计算机网络通信技术,利用和开发各种信息资源,建立服务于人类、服务于社会的强大的信息资料库。

以美国为首的西方发达国家在数字图书馆建设上起步较早,发展较快,取得了举世瞩目的骄人业绩。我国数字图书馆建设虽然起步稍晚,但发展势头强劲,无论是网络通信技术,还是信息资源开发,均在健康发展,努力赶超国际先进水平。

近几年来,随着数字图书馆建设的不断发展,关于数字图书馆建设的论著相继问世,给人们,尤其是业内人员以教育和启迪。看了这本《数字图书馆建设理论与应用技术》以后,觉得它与众不同,独具特色。它除了系统地介绍了中外数字图书馆发展建设的历史、基本理论、基本技术及应用的相关政策法规外,还有以下两个特点吸引了读者的视线,引起了人们的关注:一是它比较全面地展示了中外著名图书馆,尤其是数字图书馆的不同的建筑风格和使用功能。二是它将触角不仅仅停留在图书馆,而且投向了为数字图书馆蓬勃发展的国内外许多知名的网络信息技术开发公司,比较充分地展示了为数字图书馆建设开发研制的一系列新产品和新技术,使人们看了以后耳目一新,受益颇深。

这本书已经正式发行了,相信它不仅会成为各级图书馆工作人员一本不可或缺的工具书,成为高等院校广大师生教学中的一本好教材,而且对于希望了解和掌握数字图书馆发展、建设与应用的人们将是一本有价值的藏书。

马自卫

2003 年 12 月于北京

目 录

第 1 章 数字图书馆概述	1
1.1 数字图书馆的基本概念及特征.....	1
1.2 数字图书馆的基本功能.....	2
1.3 美国数字图书馆产生和发展状况.....	4
1.4 西方其他国家数字图书馆发展建设现状.....	6
1.5 中国数字图书馆产生和发展现状.....	13
1.6 中美数字图书馆建设比较研究.....	20
第 2 章 数字图书馆的基本系统构成	27
2.1 资源加工制作系统.....	28
2.2 信息存储管理系统.....	37
2.3 资源调度系统(handle system)	42
2.4 信息检索服务系统.....	44
2.5 网络运营支撑系统.....	51
2.6 数字图书馆基本系统的逻辑关系.....	56
第 3 章 数字存储技术	58
3.1 数字存储技术.....	58
3.2 数据压缩技术.....	78
3.3 信息检索技术.....	88
3.4 多媒体检索技术.....	91

第 4 章 数字图书馆的主要标准规范	101
4.1 数字图书馆元数据概论	101
4.2 机读目录	106
4.3 都柏林核心元数据标准	119
4.4 置标语言概述	127
4.5 标准通用置标语言简介	128
4.6 超文体置标语言	131
4.7 可扩展置标语言	143
4.8 资源描述框架	160
4.9 中美国家数字图书馆标准规范比较及我国的对策	167
第 5 章 数字图书馆信息安全技术	174
5.1 网络安全	174
5.2 网络安全技术	183
5.3 PDF MERCHANT 技术	203
5.4 数字水印技术	214
第 6 章 电子出版物的基本类型	229
6.1 电子出版物概论	229
6.2 书目数据库和全文数据库	246
6.3 电子图书	257
6.4 电子期刊和电子报纸	264
第 7 章 数字图书馆知识产权保护	274
7.1 知识产权的基本概念	274
7.2 知识产权的特征	275
7.3 网上作品的著作权保护	277
7.4 数据库著作权保护	278
7.5 著作权集体管理问题	280

7.6 中美数字图书馆知识产权保护比较研究及对象	281
第8章 基于内容管理平台的数字图书馆	288
8.1 什么是内容管理	288
8.2 内容管理与数字图书馆建设	289
8.3 主流内容管理厂商及其数字图书馆方案比较	290
8.4 数字图书馆应用软件平台	292
8.5 视频点播系统解决方案	297
8.6 大学图书馆内容存储解决方案	305
第9章 图书馆计算机综合管理系统	314
9.1 国外管理系统	314
9.2 国内图书馆计算机综合管理系统	323
第10章 存储设备和数据库	358
10.1 存储设备介绍	358
10.2 国内外数据库及其公司介绍	374
10.3 国外数据库	382
10.4 其他专业数据库	390
第11章 数字图书馆软件平台及图书馆建设	405
11.1 国内数字图书馆软件平台	405
第12章 数字图书馆建筑	435
12.1 文献数字存储对图书馆建筑的影响	435
12.2 数字化图书馆建筑的设计模式、空间结构与布局	446
12.3 数字化图书馆的建筑文化与形式	467
12.4 数字图书馆建筑将从现代建设体系走向生态建筑体系	478
12.5 智能化数字图书馆建筑	491

12.6 实例:渤海大学新校区图书馆	503
12.7 结语	506
第 13 章 附 录	508
13.1 与数字图书馆知识产权保护相关的国际公约简介	508
13.2 与数字图书馆相关的国际组织	512
13.3 国外与数字图书馆相关的法律法规内容简介	514
13.4 国内与数字图书馆相关法律法规条款	518
13.5 国内外重要图书馆网站网址	571
参考文献	576
后 记	581

第 1 章 数字图书馆概述

1.1 数字图书馆的基本概念及特征

数字图书馆是一个覆盖全球的庞大的数字化系统。它将分散于不同载体、不同地理位置的各种信息资源以数字化的形式存储其中,并以网络化的方式互相连接,以满足人们方便快捷地查询、阅览和下载各种所需的信息资料。

数字图书馆主要具备以下特征:

1. 拥有海量的数字化资源

海量的数字化资源是数字图书馆的“物质”基础。所谓数字化就是将信息用电磁介质,按二进制编码的方法加以存储和处理,将传统的纸介质存储信息的形式转变为用计算机存储和处理信息的形式。数字化资源不仅包含多种文字,而且还可将声音、图像、影像等资料数字化;数字化资源不仅可包含大量的各类图书馆、博物馆的资料,而且还可包含大量的电子出版物和全世界难以数计的各类网站中集聚的无数的数字化、多媒体的信息资源。可见,拥有海量的数字化资源是数字图书馆的一个显著特征。

2. 信息传递存取网络化

高速的数字通信网络即信息高速公路是数字图书馆存在的基础。数字图书馆依附于网络而存在,其对内的业务组织和对外的业务服务都是以网络为工具和载体,使得它得益于网络也受制于网络。只有建立四通八达、畅通无阻的网络,数字图书馆才能得以

运行和发挥其神奇的作用。目前,数字图书馆正通过由 Internet 和 Web 组成的高速、大容量的计算机网络系统将世界各国的图书馆和无数台计算机联为一体。如美国正在实施建立的国家信息基础设施计划,这一计划要求在 2015 年以前,在全美建立起一个连接所有社会机构和家庭的光纤通讯网络。而中国国家信息化基础设施(CNII)也正在加紧实施,其近期网络建设目标是建立金桥网信息高速国道,将以光纤、卫星、微波、程控、无线移动等多种通信方式,形成天地一体的网络结构,这是我国数字图书馆发展赖以生存的信息基础设施。可以预见,此规划一旦完成,我国数字图书馆高速公路建设将会有一个新的飞跃。

3. 分布式管理是数字图书馆发展的高级阶段

分布式管理意味着全球数字图书馆遵循统一的访问协议之后,数字图书馆可以实现“联邦检索”,全球的数字图书馆将像现在的因特网联接网站一样,把全球的数字化资源联为一体,组成一个巨大的图书馆,实现人类所有知识的完全共享。分布式管理之所以是数字图书馆的基本要素,在于它强调标准协议的重要性,只有全球共同遵循 TCP/IP 协议,才有因特网今天的现状,数字图书馆技术还没有这样一个公认的标准协议,因此,数字图书馆技术标准的选择和参与制定,对未来数字图书馆的发展和建设都是至关重要的。

1.2 数字图书馆的基本功能

与传统的图书馆相比,数字图书馆具有其独特的功能与作用,它不仅仅只是对人类科学文化知识的各种文本、图片、音频、视频作品的收藏、管理、咨询和阅览,而且已经成为各类信息资源搜集、整理、存储、传播、开发和利用的中心。

数字图书馆将图书、期刊、照片、声像资料、数据库、网页、多媒

体资料等各类信息载体与信息来源在统一数字化创建的基础上,将他们有机地组织并链接起来,以动态分布的方式为广大用户提供全方位的服务。

1.2.1 书刊目录检索服务功能

联机书刊检索目录(简称 OPAC)是数字图书馆的主要服务项目之一。OPAC 不仅完全取代了传统的卡片目录的功能,而且还能准确地显示该图书或资料库存和流动状态,同时,它还支持通过书名、作者、主题词、关键词及分类号等各种检索途径,实现检索目的。

1.2.2 联机全文自动检索服务功能

全文检索的基本原理就是以文字、声音、图像等各类数据为处理对象,提供按照数据资料的内容来实现信息的检索。它通过快捷的数据管理工具,能快速帮助人们进行文档资料的整理和管理工作,通过 Z39.50 协议或 WWW 技术,运用强大的数据查询手段,帮助人们迅速查找到所需的任何信息,包括获得所需的原文。自动检索系统可以在 10 分钟内用 30 种语言浏览 2300 多种杂志的 10 万篇文章,查阅完一个指定的课题。这样的工作如果改用人工来做,即使不考虑语言障碍,也需要五年的时间才能完成。可见联机全文自动检索具有神奇的功能。

1.2.3 各种电子出版物利用功能

电子出版物,是指以数字形式存储在光、磁、电等存储介质上,并通过计算机或远程通讯进行阅读的出版物。电子图书、电子期刊、联机数据库、软件出版物、多媒体出版物、数字化音频、视频出版物均属电子出版物范畴。在数字图书馆中,电子出版物是数字图书馆信息资料的重要来源。因此,通过建立各类光盘数据库与数字图书馆的连接,使数字图书馆方便快捷地为读者提供各类电子系列产品信息资源。

1.2.4 因特网资源连接功能

因特网是数字图书馆的重要环境条件。数字图书馆通过与因特网连接,迅速进入了全球性计算机网络,不仅可以获取国际、国内大量的信息,而且可以与各个数字图书馆相连接,通过各馆书目服务机构,获取各馆的馆藏信息。

1.3 美国数字图书馆产生和发展状况

数字图书馆起源于美国。20世纪60年代末,随着世界上第一台电子计算机的诞生,人类开始第一次尝试应用计算机存储和处理图书馆信息技术。美国俄亥俄州法律自动研究系统的出台,首次实现了法律问题的联机检索,当时美国空军建立的法律信息电子系统(LITE),索引了美国所有的法律法规和司法解释,为军队系统人员提供自动检索服务,开创了数字技术应用于图书馆的先河。20世纪70年代,IBM公司基于文本存储与检索技术,开发了STAIRS系统,将计算机应用于当时一些图书馆信息流通管理方面。进入80年代以来,随着计算机与通信技术相结合,促进了图书馆联机系统和网络化的发展,一批联机编目网络和一些商业性联机检索系统相继出现,并进一步发展在各馆之间建立起互联网络。到了80年代末期和90年代初,由于计算机技术,通信技术、网络技术、高密度存储技术和多媒体技术的飞速发展,使图书馆自动化进入了高级发展阶段。进入90年代以来,随着因特网的建立和普及,美国率先开展了真正意义的数字图书馆的研究和建设。

1992年7月,美国政府提出了《信息基础建设与科技法案》,明确界定了数字图书馆的含义、功能与技术应用范围,正式将数字图书馆建设纳入到美国国家信息基础建设(NII)之中。

1994年以来,美国先后有四项大型的数字图书馆建设规划工程先后实施,取得了明显的成效。

1.3.1 美国数字图书馆启动计划

1994年9月,美国国家科学基金会(NSF)、美国国家宇航局(NASA)、美国国防部高级研究项目中心计划署(DARPA)共同出资,由美国六所大学分别承担,为期四年的“美国数字图书馆启动计划”一期工程(DLI-1)建设正式开始。该计划的目标是:“使收集、存储和组织数字化信息的技术手段得到极大的进步,使数字化信息能通过网络查询、检索和处理,并以用户方式予以实现。”该计划的六个研究项目已于1998年8月如期完成。

“美国数字图书馆启动计划”二期工程(DLI-2)由国家科学基金会(NSF)、美国国防部高级研究项目中心计划署(DARPA)、国家人文学资助会(NEH)、国家医学图书馆(NLM)、国会图书馆(LOC)和国家宇航局(NASA)联合资助。目前,美国正通过DLI-2计划从扩展媒体、形态等方面,大力促进数字图书馆的研究和开发。

1.3.2 美国国家数字图书馆“美国的记忆”项目

1995年10月,在美国第104届国会决议中,决定投资6000万美元,在美国国会图书馆建立“国家数字图书馆”项目(NDLP),在因特网上正式名称为“美国的记忆”。“美国的记忆”数字图书馆项目的目标是,到2000年实现500万件文献的数字化,它们集中反映了美国建国200年来的历史遗产及文化。几十个资源库中,包括美国建国以来的重要历史里程碑式文献,如经杰克逊总统亲笔修改的《独立宣言》手稿,林肯总统的演说,南北战争中的一些照片,各族移民的民谣演唱等生动活泼的多媒体资源,现已不仅成为美国对青少年进行爱国主义教育的首选电子教材,而且他们在互联网上免费向全球服务,取得了良好的社会效益。

1.3.3 加州数字图书馆(CDL)项目

美国加州数字图书馆项目完成于1999年1月。其馆藏内容十分丰富,含有5000余种电子刊物,167个大型书目数据库,在主

要的内容资源数字化中,有将近 4000 种内容资源采用了 EDA/SGML 标引,CDL 采用 Interlib 集成服务,系统具有对学者电子印刷智能检索、多媒体查询服务功能。目前,加州已有 41 个单位使用其在线档案。

1.3.4 美国国家科学、数学、工程与技术教育数字图书馆(NSDL)计划

1998 年,美国国家科学基金会正式启动了美国国家科学、数学、工程与技术教育数字图书馆计划。按此计划,NSDL 将发展为一个联机环境,向数百万各种档次的学生和教师提供高质量的科学、数学、工程与技术教育资料,使 NSDL 成为美国公民终身学习的一种有效的无穷无尽的资源。

1.4 西方其他国家数字图书馆发展建设现状

1.4.1 西方八国集团(G8)项目

1995 年 2 月 25—26 日,法、日、美、加、德、意西方七国集团(后来由于俄罗斯加入,遂称之为八国集团 Group8,简称 G8)在布鲁塞尔公开了一次讨论全球信息基础设施建设工作会议。提出了十一项“全球信息社会示范计划”(GISPP)。它们是:

由欧洲委员会和日本负责的“全球总目录计划”(GIP);

由加拿大和日本负责的“全球宽带网互操作计划”(GIBN);

由法国和德国负责的“跨国文化教育培训计划”(C-CTE);

由法国和日本负责的“电子图书馆计划”(EL);

由意大利和法国负责的“世界文化遗产多媒体存取/电子博物馆计划”(MAWCH/EM);

由美国负责的“环境与自然资源管理”(ENRM);

由加拿大负责的“全球紧急管理”(ENRM);

由欧洲委员会负责的“全球保健应用”(GHAP);