

TUSHUGUAN SHIYE

HUIGU YU ZHANWANG



世纪之交

图书馆事业回顾与展望

中国图书馆学会编

北京图书馆出版社

庆祝中国图书馆学会成立 20 周年

世纪之交：图书馆事业回顾与展望

中国图书馆学会 编

北京图书馆出版社

图书在版编目(CIP)数据

世纪之交:图书馆事业回顾与展望/中国图书馆学会编.北京:北京图书馆出版社,1999.7
ISBN 7-5013-1639-2

I.世… II.中… III.①图书馆学会-中国-纪念文集②图书馆事业-概况-中国 IV.
G259.22-53

中国版本图书馆 CIP 数据核字(1999)第 25421 号

书名 世纪之交:图书馆事业回顾与展望

著者 中国图书馆学会 编

出版 北京图书馆出版社(原书目文献出版社)

发行 (100034 北京西城区文津街7号)

经销 新华书店

印刷 北京市双桥威宁侯印刷厂

开本 787×1092 毫米 1/16

印张 23.25

字数 530(千字)

版次 1999 年 7 月第 1 版 1999 年 7 月第 1 次印刷

印数 1—3000

书号 ISBN 7-5013-1639-2/G·438

定价 45.00 元

序

在人类即将迈入 21 世纪之际,我们迎来了新中国建国 50 周年,同时也迎来了中国图书馆学会成立 20 周年。20 年来,伴随着共和国改革开放的步伐,中国图书馆事业得到了空前的繁荣和发展。中国图书馆学会在中国科学技术协会、文化部和各届理事会的领导下,作为桥梁和纽带,团结、组织广大会员和图书馆工作者,开展多种形式的学术活动,促进图书馆学研究和图书馆事业的发展,为社会主义物质文明和精神文明也做出了积极的贡献。

为庆祝中国图书馆学会成立 20 周年暨 1999 年年会的隆重召开,总结我国图书馆理论研究与事业发展的成就和经验,探索新世纪图书馆的发展战略,学会在全国图书馆范围内开展了征文活动。本次征文的总主题是:“世纪之交:图书馆事业回顾与展望”;分主题有:(1)数字图书馆与网上图书馆,(2)网络环境下的文献资源建设与共享,(3)新世纪图书馆管理与改革,(4)走向 21 世纪的文献信息服务,(5)文献信息标引与编目工作的发展态势,(6)知识经济与图书馆。

征文活动得到了全国各地、各级、各类型图书馆工作者的积极响应和大力支持。短短几个月,年会组委会收到了文章 750 余篇,数量之多,仅次于 1996 年北京国际图联大会的应征规模。由全国图书馆界知名专家、学者组成的论文评选委员会对这些文章进行了初选、终审两次评选,最后评出优秀论文 70 篇,大会交流论文 341 篇。由于篇幅所限,本论文集只将优秀论文 66 篇结集出版,其余论文也有一定参考价值。

本书是一部凝聚着广大图书馆工作者心血和汗水的论文集。虽然不能算鸿篇巨制,但是不同层面、不同角度对我国图书馆事业的历史和发展进行了回顾与展望,比较全面地反映了我国图书馆工作者的学术水平。

随着知识经济时代的到来,作为人类知识宝库的图书馆,对推动知识经济的发展正起着不可低估的作用。1998 年 12 月 22 日中共中央总书记江泽民同志对国家图书馆的视察,充分体现了党和国家对图书馆事业的高度重视,表明了图书馆在社会政治、经济、文化中的重要地位。这次视察,是以江泽民同志为核心的党中央实施“科教兴国”战略的一次具体行动,寓意深远,影响巨大,激励了国家图书馆和全国图书馆工作者为社会进步和发展做出更大的贡献。

回顾过去,展望未来,我们相信:在邓小平理论伟大旗帜的指引下,在以江泽民同志为核心的党中央的领导下,在全国广大图书馆工作者和图书馆学会会员的共同努力下,我国图书馆事业必将以崭新的面貌跨入 21 世纪,并在新世纪得到更新、更大的发展!

谨以此书庆贺中国图书馆学会成立 20 周年;同时,为国庆 50 周年献上我国广大图书馆工作者的一份节日祝愿。在此,我代表中国图书馆学会,向为本书出版给予大力支持和帮助的北京图书馆出版社表示衷心的感谢。

周和平
1999年6月8日

目 录

数字图书馆与网上图书馆

关于建设我国数字图书馆的若干理论与实践问题·····	莫少强(1)
网上图书馆的构建·····	黄群庆(10)
数字图书馆与数字化标准文献资料的保护·····	李 旒(15)
探讨网络环境中农业核心图书馆系统的建立·····	张芃芃 吴限忠(19)
数字图书馆和网上图书馆·····	杨宗英 郑巧英(27)
中国数字图书馆:希望与挑战并存 ·····	张伟云(34)
试论区域性中心城市公共图书馆在数字化图书馆中的定位 ·····	陈秀英 彭俏东 明安生(38)
从传统图书馆到数字图书馆:趋势与变革 ·····	初景利(42)

网络环境下的文献资源建设与共享

资源共享是知识经济时代图书馆发展的根本出路·····	邵小鸥(48)
网络环境下我国图书馆文献信息资源建设与共享·····	刘省泉 曹 臻(53)
我国民族文献资源建设的概况及其开发利用的未来展望·····	包和平 包智慧(58)
面向 21 世纪的图书馆信息资源建设 ·····	许 玲(64)
与 ISP 机构合作 实现图书馆资源共享 ·····	刘洪辉(68)
新的信息环境中的图书馆藏书发展政策·····	肖希明(73)
网络环境下的文献资源共享·····	张晓林(77)
网络环境下华侨华人文献资源的建设与共享·····	郭呈祥 陈晓燕(85)
试论电子环境下图书馆馆藏发展趋向·····	王 珊(89)

新世纪图书馆管理与改革

重组我国大学图书馆业务机构的新思路·····	朱常红(95)
也谈图书馆馆长领导艺术·····	郑太平 刘慧慧 李泽红 敬 卿 赵云亭(100)
20 世纪与 21 世纪:处在十字路口的中国图书馆学教育 ·····	刘延章 谢琳惠(105)
浅谈高校图书馆改革发展中的各种关系·····	胡 勤(109)
试论新形势下的企业图书馆工作·····	李法成(113)
关于在我国建立图书馆协会的构想·····	傅安平(119)
21 世纪阅读社会与公共图书馆读书活动 ·····	范正文 王学熙(122)

21 世纪图书馆人力资源管理战略	孟雪梅(127)
论图书馆岗位考核——关于图书馆考核工作的对比研究	胡 越 黄柏楼(132)
江苏公共图书馆管理改革二十年	吴 林(138)
21 世纪中小学图书馆管理工作亟待解决的几个问题	陈守海(144)
试论跨世纪民族图书馆员的素质	金贞爱(148)
《普通高等学校图书馆规程》的地位作用及修正的若干问题	夏 勇 马景梯(153)
建立可持续发展的图书馆管理体系——知识经济带给图书馆管理的思考	王世伟(158)
论面向 21 世纪以需求为基础的图书馆管理原则	肖 燕(164)
图书馆面临新世纪挑战的对策	马远良(170)

走向 21 世纪的文献信息服务

面向 21 世纪的图书馆信息服务	孙世菁 刘颖奇(175)
世纪之交图书馆信息服务中的误区及对策	沈媛芬(179)
加强高校图书馆的信息用户教育	孙 平(184)
走向 21 世纪的图书馆文献信息服务	杨文祥(189)
走向 21 世纪的农村信息服务	谢坤生(194)
现代文献信息服务的发展与前景	张春红(199)
对我国图书馆公益性原则的反思	王陆军(206)
走向新世纪的信息加工	倪晓建(211)
网络化与大学图书馆参考咨询工作	冯白云(215)
知识经济时代的信息管理和服 务	党跃武(221)
加强中小学图书馆建设 为素质教育发展服务	杞友存(230)
面向知识经济的图书馆电子信息服务体系	郑朝晖(234)
走向未来 走向社会——论我军情报服务现状与优化改革	韩冬梅 敬 卿(240)

文献信息标引与编目工作的发展态势

从情报检索语言到知识组织语言——试论检索语言的术语演变	周绍萍 田书格(247)
国家图书馆中文图书主题标引格式演变过程的思考	朱 芊(254)
数字化时代的电子资源编目	胡从玉(263)
我国发展联机编目的障碍及对策	张云瑾(268)
《文献编目》课程教学改革实践的回顾与前瞻	杨玉麟(272)
我国文献信息工作标准化与书目数据库建设	段明莲(278)

知识经济与图书馆

知识经济与图书馆	张廷川(283)
论知识经济与馆员素质	崔 波(289)

图书馆如何迎接知识经济时代的到来.....	竺海康	赵美娣(294)
向知识经济过渡中我国图书馆职能变革与定位.....	曲静涛	(300)
知识经济与文献情报学研究.....	刘 君	(307)
公共图书馆迎接知识经济挑战的策略.....	刘勤玖	梁 华(316)
知识经济与图书馆、图书馆学教育纵横谈	詹德优	(321)
知识经济与图书馆发展方略.....	钱佳平	(326)
知识经济:图书馆面临的新机遇和新挑战	邱均平	(331)
论知识经济时代的图书馆.....	柯 平	白 华(341)

其 他

蓬勃发展的军校图书馆事业.....	郝沫平	李 颖(346)
Internet 上免费全文科技期刊调研	马景娣	刘 琼 胡晓珍(352)
知识创新与图书馆.....	孙 坦	李广建 冷伏海(357)
发展民族地区图书馆事业的几点思考.....	阿 图	(362)

关于建设我国数字图书馆的若干理论与实践问题

莫少强(广东省中山图书馆)

一、引言

数字图书馆是世界各国图书馆和信息产业界正在竞相研究和大力发展的重大课题。中国政府十分重视数字图书馆的建设,1997年将“中国试验型数字化图书馆”列入国家重点科研项目,由北京图书馆、上海图书馆、广东省中山图书馆、深圳图书馆、辽宁省图书馆和南京图书馆联合开发研制。两年多来,研制组对世界各国发展数字图书馆的趋势、技术路线、相关技术标准等进行了广泛调研和深入分析,上海图书馆提出了5个技术研究报告,为发展我国数字图书馆进行了有益的探索 and 提供了重要的参考依据。最近,国家图书馆、广东省中山图书馆和广东省湛江市图书馆同北京市超星电子技术公司合作,开展大规模数字化馆藏资源建设,利用具有国际先进水平的中文全文检索软件开发图书搜索引擎,利用163、169网络建立资源共享平台,将300多万页(约5000种)文献上网供读者检索和阅读,访问人次急剧上升,每月访问量超过200多万人次。与其他现有的图书馆网络相比,该系统已具备数字图书馆的雏形。笔者作为上述项目的参加者和组织策划者之一,拟结合实际对建设我国数字图书馆的若干理论与实践问题发表如下意见。谬误之处,敬请指正。

二、关于数字图书馆的概念和基本框架

随着我国广域网技术的迅速推广,我国已建立了覆盖全国的计算机通信网络,建立网站5000多个,发展网络用户200多万个,网上中文信息也快速增长。网上出现了一批名为“网络书库”、“网上书廊”、“电子书屋”、“电子图书馆”的网站,有的则称为“XX数字化图书馆”。这些网站大部分仅有少量的数字化文献资源(一般不超过1000部作品),采用文本格式输入,用HTML设计网页,使用通用浏览器阅读;有的图书馆将书目数据同部分文本记录进行链接,实现了通过书目检索而获得文本原文;有的图书馆将数以百计的国内外图书馆的网址进行分类列表和链接,也称之为“网上图书馆”、“虚拟图书馆”等;有的文献更是将因特网称之为超级数字图书馆。出现这些情况并不奇怪,因为对“什么是数字图书馆?”国内外尚无统一和严格的定义。

根据课题组对国内外发展数字图书馆趋势的调研分析和两年多来的试验研究,我们认为建立一个数字图书馆必须具备以下要素:(1)拥有独立和规模宏大的馆藏数字化文献资源;(2)拥有高速、可靠和开放的网络资源,可每天24小时向社会提供电子信息服务;(3)拥有高效率和操作方便的搜索引擎和浏览器,可自动进行分布式文献检索,并通过网络实现远距离获得原文;(4)拥有一套数字化资源制作、存储、发布、维护的软件系统。我国建立数字图书馆还需要考虑中文化等诸多因素,系统还必须满足以下要求:(1)印刷载体文献的数字化基本保持书刊

原貌,即读者在屏幕看到的必须与原件一致;(2)数字图书馆必须能够同时存储文本、图像、音频、视频等文件,其中书刊资料要以方便阅读的方式提供使用;(3)提供具有中文全文检索功能的搜索引擎,能够对书名、著者、分类、主题、目次、文摘甚至内容进行检索,并通过网络提供原文。因此,根据上述要求分析,目前许多称之为“网上图书馆”、“数字化图书馆”的网站并不是真正意义上的数字图书馆,或只能叫作数字图书馆的“社会主义初级阶段”。

数字图书馆的数字化馆藏,应由三部分组成:(1)印刷载体的数字化文献;(2)多媒体文献(音频、视频节目);(3)二次开发文献(由图书馆加工的文摘、剪报资料等)。因此必须创建能够存储和处理上述数据的数字图书馆基本框架。根据我们的研究,数字图书馆资源的基本框架可定义为三层:即对象存储层、元数据存储层和网页数据存储层。对象存储层永久存储文本、图像、音频、视频等信息,其物理存储地址一经确定不再更改;元数据存储层存储书名、著者、主题词、分类号、文摘、提要以及其他标引信息,可动态更新和建立跨服务器的对象链接;网页数据存储层主要存储数字图书馆网站的网页信息和相关的管理信息。图1为我们构建的数字图书馆基本框架示意图。

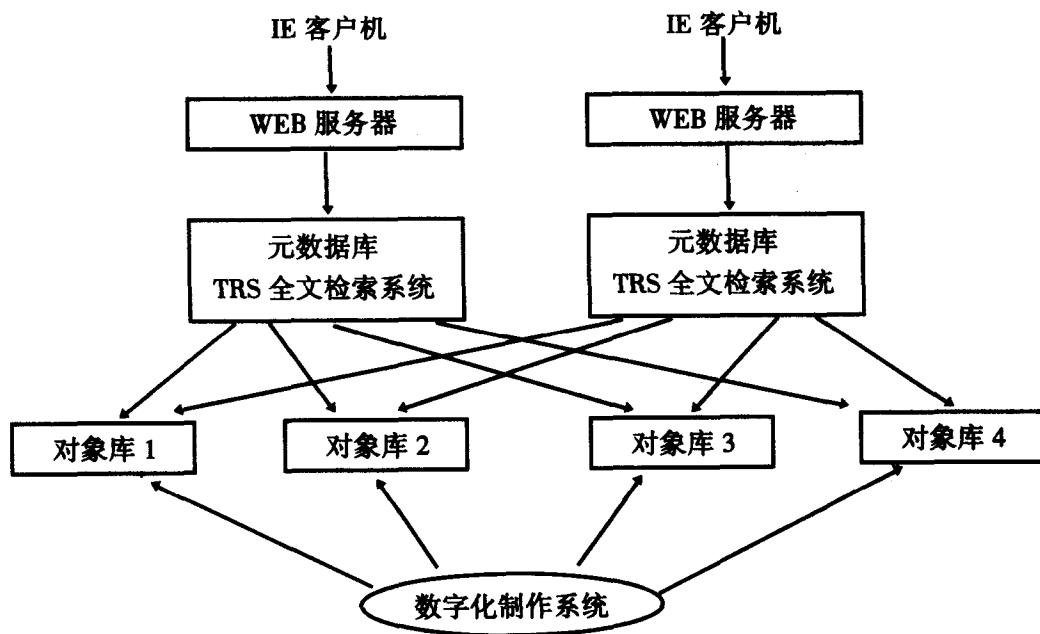


图1 数字图书馆基本框架

1998年下半年以来,国家图书馆、广东中山图书馆和湛江图书馆先后与北京超星公司合作,采用PDG文档资料扫描存储制作系统进行大规模的馆藏文献数字化,利用TRS中文全文检索系统开发图书搜索引擎,通过同邮电部门的合作在163、169网建立网站,按图1所示的数字图书馆基本框架建立起“网上图书馆”,向全社会提供免费的网上检索和提供原文,实现了文献资源的共建共享。与其他网站相比,该系统具有以下特点:(1)建立了大规模的数字化文献

资源,目前上网 300 多万页(5000 部图书)书刊,今年将上网 3000 万页(15 万部图书),这样的数量和规模是其他网站所望尘莫及的;(2)系统提供具有全文检索功能的图书搜索引擎,可从书名、作者、主题、分类等途径进行检索,也可以从图书的目次、摘要和其他标引词进行查询,响应速度快,操作方便;(3)可在网上直接获得图书原文,读者所看到的内容(包括字体、版式、插图等)与原书完全相同;(4)系统的数字化资源不仅保持了图书馆的目录组织形态,而且保证每部图书以方便阅读习惯的方式提供给每位读者;(5)使用超星浏览器或通用的 IE 浏览器,既可阅读网上图书,也可以同时访问其他网站;(6)具有分布式查询功能,通过网络可直接阅读分布在北京、广州、湛江等不同城市图书馆的数字化资源;(7)对文献的存储数量几乎没有限制(TRS 目前每库最多可存储 42 亿条记录);(8)在实现数字化资源共享和保护作者版权方面进行了有益的探索。我们认为该项目的成功实施,为建设有中国特色的数字图书馆闯出了一条新路,为实现国家、省和市三级图书馆网络的资源共建共享提供了经验,具有十分重大的现实意义。

三、关于馆藏文献的数字化处理技术

图书馆需要进行数字化处理的文献有多种类型,但占绝大多数的是印刷型文献。如何正确选择技术先进、符合我国国情并能与国际互联网接轨的数字化技术,对大量的书刊进行数字化处理,是能否建成数字图书馆的关键。对印刷载体文献的数字化有以下技术要求:(1)尽可能保持图书、期刊的版式原貌(数字化肯定会丢失一些信息);(2)符合人们的阅读习惯;(3)在满足各种要求的前提下尽可能小的存储尺寸;(4)能够通过网络快速传送(如果看一页书要等上几分钟就是不实用的);(5)能够在图书馆内组织大规模的工业化生产;(6)图书馆能够承受的生产成本。我们首先分析采用 OCR 技术的方案,现在许多网站上使用的就是这种方法。将书刊扫描并 OCR 为文本后,用 HTML 语言编辑上网,虽然有节省空间和便于网络传送的优点,但制作速度慢、成本高,需要人工校对还错漏百出,不能保持书刊版式原貌,此种方式只适合发布新闻和信息,不适合用于数字图书馆的大规模制作。美国、澳大利亚、加拿大、英国等图书馆广泛采用光学扫描、压缩存储的图像处理方法,存储格式有 TIFF、JPEG、GIF 等多种,各有其优缺点,最大的问题是存储尺寸太大,每页需占数百 KB 空间。近年来又出现了全数字化的 PDF 格式,因其可以实现全文检索又能显示原有版式而大受欢迎,国内已有《参考消息》等报社用于制作报纸全文光盘。但我们认为 PDF 格式目前尚不适用于数字图书馆,因为其昂贵的制作成本(相当于重新排版)是任何一个馆都负担不起的。

我国图书馆和档案部门近年来采用了北京超星公司开发研制的 PDG 文档资料扫描存储系统进行数字化。PDG 系统是一个非常适宜图书、档案部门进行大规模数字化资源制作的系统,目前已为国家图书馆、广东中山图书馆、湛江图书馆和全国近千个档案部门所采用,制作了数千万页的图书、期刊和档案文献,具有扫描速度快、制作成本低、高度压缩存贮和还原,以及可提供上网服务等优点。PDG 主要的技术参数为:扫描 A4 页(300DPI)存储尺寸为 30 - 50KB,转换为网上浏览和网络传输格式(150DPI)为 15 - 25KB;普通扫描仪制作速度为每分钟 2—4 页,快速扫描仪制作速度为每分钟 22 页;网上浏览需要下载和运行专用或 IE 版的超星浏览器(免费提供);可与通用浏览器软件和 TRS 中文全文检索系统实现无缝联接;PDG 格式可转换为 BMP、TIFF 等格式。我们对制作单位和网上用户的意见进行了调查和分析,普遍

认为 PDG 系统具有以下优点:(1)高度的压缩存储和还原技术先进,比其他格式节省空间 50% 以上;(2)快速扫描和转换技术领先,比其他扫描系统提高速度 2—10 倍;(3)制作成本低廉,图书、档案部门可以在本单位开展生产制作;(4)网上传输速度快,使用普通 MODEM 显示一页图书只需 2—3 秒;(5)使用方便,网络用户可进行打印、下载和存位图等操作;(6)具有向其他格式转换和进一步实现全数字化的基础。其缺点是因为采用图像方式不支持全文检索,网络用户需要先下载浏览器软件,目前不支持彩色图像等。综合比较和权衡得失,我们认为 PDG 不失为适合我国国情,经过实践检验并得到广大用户认可的数字化制作系统,如能在我国图书馆统一推广应用,必将大大推进我国数字图书馆的建设和发展。

我们采用 PDG 格式进行印刷载体馆藏文献的数字化,但并不排斥采用其他格式进行资源数字化,因为数字图书馆对信息资源存储和利用的需求是全方位、多层次的。如图书馆开发和加工的二次文献(题录、文摘、剪报资料)、重要的检索工具书等适宜采用文本格式,以便实现全文检索;网页文字说明、图像等适宜采用国际流行的 HTML 和 JPG 格式;各种彩色画册、珍贵的古籍善本可考虑采用高质量(当然成本也高)的扫描方法和存储格式;对绝大多数的印刷型书刊则采用 PDG 格式,这样就能大幅度降低制作成本,大大加快图书馆数字化的步伐。

四、关于元数据和搜索引擎技术

为在数字图书馆浩如烟海的文献信息中快速、准确地查找到所需的文献,必须进行文献的著录、标引和建立搜索引擎。这里涉及到两个关键问题,一是数字图书馆采用怎样的标准和方式建立元数据库,二是选择什么数据库做搜索引擎。我们首先讨论元数据问题。所谓元数据就是经过著录、标引的检索信息(书名、著者、主题等)加上文献地址(网址/存储路径)。国内有的数字图书馆技术研究小组依照外国的技术方案,提出数字图书馆的最小元数据集合采用 Dublin Core;最小元数据集合 Dublin Core 的实现采用 RDF 方式;元数据分成两层,第一层为 DC,第二层为 MARC 或 TEI Header,在资源建设中要求按这两层实现。笔者认为这个方案值得商榷。第一,Dublin Core 所包含的元数据仅有题名、作者或创建者、主题及关键词、描述、出版者、日期、文献类型、来源、语种等 15 个项目,可检信息比传统书目数据还少,更不可能揭示文献内容;第二,RDF 实现方式过于复杂,不利于中文全文检索系统的应用;第三,采用 Dublin Core 的出发点是便于与国际接轨,这与推广 CN-MARC 的目的相同,但实际上由于受文字和本国文化等条件所限并不能如愿,如 OCLC 以及大部分西方国家图书馆的中文编目就没有采用 CN-MARC,我国反而浪费了大量的人力、物力。有的方案提出对文献进行中文全文自动标引,这当然可以大大提高检全率,但必须先将图像 OCR 为文本,其成本图书馆承受不起,实际上大部分文献也没有必要做到每个汉字都有快速检索功能。通过调研和分析,我们提出了一个能够满足现实技术条件需要,符合图书馆编目标准(ISBD),可包容普通图书、期刊、古籍、非书资料等类型文献,具有良好扩展性能,并能满足数字图书馆检索需要的元数据方案,这就是本文第五部分介绍的《数字式中文全文文献通用格式》。该方案对元数据的处理极为灵活,对文献引得深度没有限制,在同一记录中可标识 n 篇文献的 URL(即可深入揭示文献目次、篇名和内容)。

这里讨论第二个问题,即数字图书馆应选用什么数据库作搜索引擎?数字图书馆大量的信息和知识是非结构化的,如图书、期刊、科技报告、专利文献、录音、录像、Web 页面等,并且

这些信息是由多种不同的软件工具产生的,如光学扫描、字处理软件、激光照排系统和 HTML 制作工具等,传统的关系数据库不能担此重任。我们终于找到了能够把这些格式繁杂、内容丰富的信息变成可快速检索和管理,并非常适用于中国数字图书馆的系统——TRS 中文全文检索系统。TRS(Text Retrieval System)是易宝北信公司系列产品的注册商标,源于 80 年代中期开始的对中文信息检索的开创性研究,今天已经大大超出了全文检索的领域,发展到网上信息发布、电子出版、搜索引擎、文档管理和电子商务应用等领域,是我国用户最多的全文检索系统。目前已有国家图书馆、国家版本图书馆、广东省中山图书馆、广州图书馆、深圳图书馆、福建省图书馆等数十家大型图书馆采用。TRS 内嵌汉语自动分词系统,支持按词、按字或词字混合的策略,大大提高检索的准确性和响应时间;采用先进的压缩技术,使得系统的空间膨胀率低;提供了多达 48 种检索运算符的检索方式,包括外部特征与正文内容的各种逻辑组合检索、位置检索、二次检索、渐进检索、历史检索、词根检索、大小写敏感检索、概念检索、对检索结果按与检索表达式的相关性和重要性程度排序等;支持多媒体信息的存贮和管理,支持 RTF、Microsoft Word、Excel、Powerpoint、PDF、S2/PS2/PS、HTML、ISO2709 等文档的自动装库;除了各种检索手段,TRS 还配备了强大的数据库管理功能,如数据库维护、数据库的安全性控制、数据的完整性和一致性、数据库备份和恢复、用户管理等功能,提供可定制的信息服务计费功能。TRS 版本 3.53 的性能指标如下:

- (1)可创建 60000 个以上数据库;
- (2)每库可存储、管理 42 亿条记录;
- (3)1G 字节原始信息的查询时间为 2—3 秒;
- (4)100M 原始文本信息的索引时间为 2 小时左右;
- (5)空间膨胀率在 0.9—1.5 之间;
- (6)可跨 50 个数据库检索;
- (7)支持大量并发用户的同时访问;
- (8)运行平台:Microsoft Windows NT、SUN Solaris、HP HP-UX、IBM AIX、Digital Unix、SGI IRIX、SCO Unix、Linux。

广东省中山图书馆按照上述元数据方案,利用 TRS 平台建立了网上图书馆的元数据库和图书搜索引擎,经实践证明是可行的,达到了数字图书馆的各项基本功能。这充分说明,采用先进成熟的软件作为数字图书馆的开发平台,比图书馆自行开发软件提高效率不知多少倍!图书馆应该把重点放在资源建设上,而不应重复开发类似的软件系统。

五、关于数字式中文全文文献的技术标准

建设我国数字图书馆应该直接采用先进的国际标准,如 SGML、HTML 以及其他的图像、通信标准,但也应该制定符合我国国情的相关标准。为加强我国公共图书馆文献资源建设,推动图书馆自动化、网络化和数字图书馆的发展,文化部科技司将制定《数字式中文全文文献通用格式》这一文化行业标准任务下达给广东省中山图书馆。这是建立数字图书馆元数据库的一个重要标准。该馆成立了标准研究和起草小组,经过两年多的研究和试验,最近已完成格式的基本设计和标准初稿。该标准格式是基于 SGML 国际标准和中文全文检索技术而设计的新一代机读格式,规定了文本、图像以及多媒体等类型文献元数据的规范和定义,可用于普

通图书、古籍、连续出版物、标准文献、科技报告、学位论文、地图资料、缩微资料、计算机文档等类型文献的元数据处理,是联接数字图书馆对象和网络用户的桥梁与纽带。该标准采用可变长字段的文本格式,每字段以换行符作为结束符,并严格遵循《中国文献编目规则》以及有关的国际和国家文献工作标准,坚持准确完整地进行文献著录。该标准具有通用性,既可作为中文全文文献信息的交换格式,也可作为计算机系统的内部存贮格式。目前,广东中山图书馆按照该格式建立了“网上图书馆”和数十个全文数据库(数据量超过1亿汉字),接入广东邮电169网向社会服务,接受用户的评价和实践的检验。据初步分析,该格式具有以下优点:(1)基于通用置标语言SGML和中文全文检索技术而设计,可与先进的国际标准接轨,适用于Internet网络环境;(2)实现了文献编码、目录和文献内容的一体化处理,极大地简化了中文文献的著录和标引,提高了工作效率和建库质量;(3)全方位支持中文全文检索系统,不仅传统的文献目录而且文献的目次、摘要、内容也可以实现文本全文检索;(4)采用面向对象技术,可以对文本、图像、录音、录像、计算机程序以及网址资源进行编目和数字化处理,覆盖了几乎所有的文献类型;(5)可以利用CN-MARC记录自动转换为该格式,并在依照ISBD标准编目、文献标引深度、支持数字式图书馆开发等方面有所创新。以下是用上述标准制作的样本:

<REC>

<记录控制号>=GD000005

<记录版次标识>=199811200155301

<固定长编码信息>=1998112001998mychichicnzzysy

<目录信息>=岭峤春秋-岭南文化论集(一)/广东炎黄文化研究会编.-北京:

中国社会科学出版社,1994.-630页;20cm.-(岭南文化研究论丛).-7-5000-5470-XRMB24

目 录

《岭南文化论集》序/欧初

在岭南文化研讨会上的讲话/李兰芳

岭南文化研讨会开幕词/丁希凌

岭南文化研讨会闭幕词/赖竹岩

论岭南文化的历史地位/李权时

试论岭南文化的原生型态/李勤德

古今联系探索广州的特点/关履权

当代岭南文化的考察/周大鸣

岭南传统的商品经济与商品意识/黄明同

试论近代岭南商品文化的特点/刘圣宜

珠江三角洲居民的商品意识/刘琨琰

宋代移民与岭南文化的发展/李明华

石湾大汕与方以智/姜伯勤

粤东高氏的源流与文化传统/高路加

岭南文化与海/蔡鸿生

汉代徐闻港的对外贸易/高美玲

近代西学东渐与岭南文化/袁钟

深圳特区文化初探/陈乃刚

惠能禅学与岭南文化/冯达文

 南汉高祖刘岩/邓炳权
 论屈大均的文化趣尚/覃石文
 屈大均著作版本述略/李文约、林子雄
 反侵略的爱国诗人黄培芳/管林
 康有为与岭南文化/宋德华
 粤西桂圩地区汉族宗教的研究/陈运飘
 粤剧艺术与岭南文化/李门
 浅论岭南的音乐与戏曲/邓汝邦
 岭南画概述/徐家凤
 论岭南文化与岭南文献整理/钟贤培
 胡根天老师的业绩/汪铭珊
 岭南远古文化遗存/阮晓波
 余靖风采与岭南文化/易行广
 岭南文化发展及其开拓者述略/何国华
 惠州姓对与中原民族南迁/刘灿
 中西文化交流与近代岭南文化/刘卓红
 洪秀全在广州活动对其思想影响/方忠英
 广东葡语和广东英语初探/章文钦
 近代岭南政治思想的演变与特点/张岩
 当今岭南饮食文化的特点及其影响/许少彭
 试论广东境内的多言现象/张振江
 试论清代文字狱对岭南方志的影响/张振江
 岭南文化与中原文化比较/应向东
 编后记/
 <国际标准信息> = 750005470X
 <定价> = RMB25
 <标引词> = 文化研究 广东 文集
 <分类号> = G109
 <文献内容> = 阅读图书
 <制作单位> = 中山图书馆

由于 TRS 全文检索系统的支持,这个元数据记录中的任何字符都可以快速检索(包括目次和篇名),通过 URL 直接获得原文。这样的标引深度既基本满足了检索需要,同时又避免了全文本记录自动标引所产生的大量冗余信息。

建设数字图书馆需要制定一系列新的文献工作标准。在制定新标准的过程中,笔者认为有两个倾向值得引起重视:一是片面强调图书馆的特殊性,不积极向国际信息产业界标准靠拢,情愿对旧格式修修补补而不愿采用新技术对其进行脱胎换骨的改造,如 MARC、Z39.50 都是图书馆标准(或协议)而非国际信息产业界标准,新一代标准的制定不能再走老路。二是片面强调与国际接轨而不顾中国国情。数字图书馆的建设,强调要有中国特色非常重要。我国有着五千年的灿烂文明和浩如烟海的文献典籍,我国的汉字有全球四分之一以上的人使用,如果建设面向 21 世纪的图书馆不首先考虑中国人民的需求,置中国国情而不顾去寻求所谓与国

际接轨的技术标准,就是本末倒置,是不可取的。因为数字图书馆与本国文化有着密切的联系,不能等同于计算机网络的互联。

六、关于数字图书馆的共建共享与知识产权保护

数字图书馆的基本模式之一,就是不同城市甚至世界各地的图书馆进行分工协作建库,通过网络实现资源共建共享。要实现这一目标,由政府统一规划和协调,组织图书、情报部门共同推进数字化资源建设极为重要。今年1月,由国家图书馆牵头,在北京召开了规模空前的全国图书馆资源共建共享会议,文化部、教育部、国务院办公厅等国家部委的有关领导出席会议,124个图书馆签署了资源共建共享协议。为贯彻落实这次会议精神,广东省文化厅制定了《广东省图书馆资源共建共享实施方案》,目前正在组织实施。广东的具体做法是:

1. 确定和发展网络成员馆。具备条件并遵守合作协议的图书馆可申请加入广东省图书馆信息协作网:(1)已建立一定规模的特藏文献或可获得相关的文献资源;(2)有独立的WEB服务器和128KB以上的DDN专线,在广东视聆通(169)网建立网站,可提供每天24小时的信息服务;(3)建立专门的数字化制作部门,购置PDG文档资料扫描存储系统,可进行数字化文献资源制作。

2. 积极推进数字化资源合作建设。(1)中文新书数字化制作和上网服务。省馆负责本省新书出版物的筛选、著录、标引和扫描,装载到本馆服务器上网服务,并参加全国图书馆资源共享网络,建立TRS图书搜索引擎数据库,为本省网络成员馆提供全国范围的新书资源共享。(2)外文新书刊的数字化制作和网上文件传输服务。网络成员馆在就外文新书刊的采购进行充分协商和协调的基础上,负责本馆进口原版书刊的数字化制作,并通过网络进行数字化文件的远距离传输,实行成员馆合作制和免费服务,即成员馆免费提供本馆外文新书刊数字化资源,同时可免费获得其他成员馆的外文新书刊数字化资源。(3)特藏文献和地方文献的回溯建库。参加合作建库的成员馆负责本省(市)地方文献和特藏全文数据库的制作和上网服务。

3. 采用统一的资源数字化和网络共享技术。参加合作建库的成员馆统一采用PDG系统进行馆藏文献的数字化制作;在中国公众多媒体通信网(169)平台建立资源共享平台,成员馆通过DDN专线与当地的169网络联网;以广东省中山图书馆的TRS中文全文检索系统为基础建立数字图书馆元数据库和搜索引擎;按照《数字式中文全文文献通用格式》进行中文文献标引和元数据资源制作。

4. 制定有力的保证措施。(1)广东省公共图书馆自动化网络办公室负责制定我省图书馆文献资源数字化的总体规划和有关管理办法。成员馆开发和建设有关全文数据库必须向该办公室提出书面计划,由该办公室进行审核和协调,避免重复建设。(2)参加合作建库和网络共享的成员馆根据所确定的建库任务、收录范围和本馆馆藏特色进行文献资源收集,建立有特色和较为全面的图书资料馆藏,并承担有关的购书费用。(3)成员馆承担本馆网站设置的有关费用,包括服务器、网络设备、数字化制作设备、DDN专线以及PDG、TRS软件等购置费用。(4)各馆建立的中文新书和特藏文献、地方文献全文数据库免费向成员馆开放,即除向省馆TRS图书搜索引擎提供书目和链接地址等数据外,同时向其它成员馆提供同样的数据,使全体成员馆可通过169网络方便地直接访问其数据库。(5)通过169网络的虚拟网功能建立面向全体成员馆的专用网,用于建立外文图书期刊全文数据库以及共享网络,向成员馆免费提供全文文

献远程传真服务。这一网络的建立和使用仅限于成员馆,不直接向读者提供远程检索服务。(6)参加合作建库和网络共享的图书馆必须同广东省公共图书馆自动化网络管理办公室签订合同,确定建库的目标、内容、数量和质量,保证在规定的时间内向全体成员馆提供数字化资源,并向读者提供 24 小时的联机服务。

处在全球网络化和知识经济时代,数字图书馆能否实行有效的知识产权保护备受关注。现在有些网站用 HTML 语言编程和发布文本型的电子图书,任何人可免费浏览和下载,因为是整部作品的网上传播,对作者版权造成侵犯是毫无疑问的。我们保护知识产权的做法是:(1)图书文献统一用超星 PDG 格式制作和上网,每页作为一个显示单元,读者不容易将整部作品下载,或整部传送的通信费远高于书价;(2)用户阅读和下载要使用超星浏览器,限制其用于其他商业用途;(3)控制上网书刊图像的 DPI,使其不能直接用于盗版印刷;(4)在网页声明“本馆书刊免费上网服务无任何商业目的,如作者认为有版权争议电告即将该书从本馆网站取下”。这种模式在国家图书馆、广东中山图书馆等网站试行了半年,访问人次超过数百万,不仅至今未接到一例关于作品版权的投诉,反而有的作者还来电要求将其作品上网。事实证明,这种数字图书馆的运作模式已初步被社会接受和认可。

七、结束语

数字图书馆代表着 21 世纪图书馆的发展方向,其前途不可限量。我国在探索建设数字图书馆的方面迈出了第一步,虽然前进道路上还存在不少困难,还有许多有待探索的新课题,但我们愿同国内外的同行携手合作,共同努力,让数字图书馆在我们这一代手中建成,为人类文明的保存和传播作出我们的贡献。

参考文献:

1. 刘炜,赵亮.技术研究报告:数字图书馆中图像信息的处理和存储
2. 刘炜,赵亮.技术研究报告:元数据方案实施意见
3. Association of Research Libraries, Definition and Purposes of a Digital Library
4. Kuny, Terry and Cleveland, Gary, Digital Libraries: Myths and Challenges
<http://www.ifla.org/IV/ifla62/62-kuny.pdf>
5. Balas, Janet, Copyright in the Digital Era
<http://www.infoday.com/cilmag/jun/story2.htm>

网上图书馆的构建

黄群庆(广东省科技图书馆)

国际互联网的发展与普及,形成了一个没有国界的网络社会,网民们安逸地坐在家中轻按鼠标,即可进行通信、娱乐、购物、学习、科研、商务、贸易和政府等活动。现实社会中的行为越来越多地搬到网络社会中去实现,当然了,网上公民也离不开图书馆。人们越来越多地提及数字图书馆(Digital library)、电子图书馆(Electronic library)、无墙图书馆(No wall library)、虚拟图书馆(Virtual library)和全球图书馆(Global library)等名词,这些名词从不同的角度强调了未来图书馆的特征,这些特征归纳起来有两大要素,一是图书馆的文献信息必须经数字化处理,二是数字化的文献信息资源打破传统上的空间概念,在网络上传输,为远程的读者服务,实现高度的资源共享。可以说,前者是前提,而后者则是最终目的。当前数字图书馆和网上图书馆的建设,正倍受图书馆界的关注,已成为业内引领时尚的新事物。“数字图书馆和网上图书馆”这个会议分主题提得很符合形势发展。本文暂且以“网上图书馆”这个名词,对它的“馆舍”、“馆藏”、“读者”和“馆员”做一些思考,有了一个大致认识或构想,才谈得上去建设它,实现它。

1 关于网上图书馆的“馆舍”

数字化文献资源的贮存设备和供读者访问的服务器,可以说就是网上图书馆的“馆舍”。相对于传统图书馆的印刷型文献载体和庞大的实体建筑物馆舍,网上图书馆的“馆舍”以其独特的方式存在。

1.1 “馆舍”的大小要规划。笔者认为,要建设一个网上图书馆,至少要有自己独立的服务器。如果没有自己的服务器,数字化资源就要挂靠在别人的服务器上,工作上受很多约束,就好像“馆舍”设在别人的办公大楼里,大招牌是别人的,不能很好地树立自己的形象。根据数字化文献资源的已有数量与发展规划,贮存设备可以选择光盘塔、光盘库或硬盘阵列等,相对来讲,硬盘阵列具有可扩充和速度快等优点。据报道,1998年罗马尼亚原子物理研究所激光部的研究员帕维尔发明了一种超级光盘驱动器,可使光盘的存储容量增加上万倍。换言之,罗马尼亚科学院图书馆现存的160万册图书如用普通技术存储需要8万张光盘,而利用帕维尔的技术仅需5张光盘。可见,高密度存储技术的发展将使得网上图书馆的“馆舍”可摆脱随着资源的累积而使得光盘塔或光盘库也面临容量不够的境地,保持体积小特点。

1.2 “馆舍”的网址要直观易记。这关系着用户对网上图书馆站点的访问,也关系着网上图书馆在用户中树立的形象。目前最为常用的网址是WWW地址,例如http://www.loc.gov是美国国会图书馆的网址,由网络传输协议(http)、服务器类型(www)和域名loc.gov组成,域名与它的机构名称(Library of Congress)和性质(government)相对应。这一点很重要的。前两年企业界出现一阵域名被抢注的风波,说明了与机构名称相对应的域名也是一种无形资产,