

大学图书馆数字资源组织体系探索

康桂英

(北京理工大学,北京:100081)

摘要:本文主要论述了大学图书馆的各种数字资源的组织与描述体系。

关键词:大学图书馆;数字资源;组织体系。

中图分类号:G258.6

文献标识码:A

文章编号:1009-3370(2003)04-0011-03

一、前言

随着现代信息技术的飞速发展,作为各项技术服务的图书馆无论是在服务上,还是在技术上也得到了飞速发展。随着网络资源的丰富,在超星数字图书馆、书生之家数字图书馆等各种数字图书馆如雨后春笋一样蓬勃发展起来时,为适应潮流,跟上时代的步伐,各大学图书馆也建立了高素质、高水平的数字图书馆,如北京大学的数字图书馆,清华大学的建筑数字图书馆,中国人民大学的数字图书馆研究所等等不一而足。在大学图书馆数字图书馆建设中最基本最重要的问题是什么呢?显而易见的,是资源数字化的问题。资源数字化是建立数字图书馆最基础的部分,也是最重要的部分。本文主要针对大学图书馆数字资源的组织体系进行粗浅的论断。

二、大学图书馆数字资源组织体系

数字图书馆的基础是海量的信息资源库,但海量信息资源并不等同于馆藏资源的数字化,也不等同于单纯的网络资源目录,或各种数字资源的简单拼凑。据笔者所知,现在有相当一部分大学图书馆数字图书馆的建设仅仅停留在购买一部分电子资源或者是把一部分纸型资源数字化,就认为已经建成了数字图书馆。笔者认为,仅有这些是微不足道的,大学图书馆的资源建设应该是建设一个包含不同层次、不同类型,相互联系、紧密配合的资源库群,通过先进的技术手段为读者提供随时随地的、快速的信息服务。

1. 数字化资源组织的内容。

从文献内容上:作为国家的大学图书馆,除了要收集整理具有本校特色的各种信息资源,如军工院校的军工文献、兵器文献等,师范院校的关于教育方面的信息等等;还要搜集整理其他基础学科,如数理化天地生及高新技术学科的文献;

从文献类型上:重点是本校师生的学位论文、期刊、会议录、科技报告等;

从文献载体上:包括印刷型、缩微型、光盘型、电子型、网

络型和数字型等各种载体类型的文献;

从文献来源上:包括本馆收藏的各种传统文献,国内兄弟院校图书馆收藏的文献,购买的各种光盘数据库、网络电子信息资源,通过专用搜索引擎搜集的免费网络信息资源等。

2. 数字信息资源组织模式。

书目数据库:各种图书书目信息、本校博硕士学位论文书目信息、期刊书目信息、各种回溯书目信息。书目数据库一般由图书馆自动化集成管理系统建立和提供,建成了各大学图书馆专有的书目数据库。

全文数据库:大学图书馆数字图书馆全文数据库主要包括两部分内容:第一,各大学图书馆都收集有相当一部分反映本馆馆藏的特色文献,在数字图书馆建设中,要对这些特色馆藏进行全文加工。第二,互联网是一个无限庞大的信息资源库,这个信息资源库中存在着大量有价值的科学文献,也有大量的所谓的“垃圾”信息,各大学图书馆数字图书馆全文数据库建设的一个重要方面是对网上的有价值的科学文献信息进行收集、整理,为用户提供更好地服务。第三,大学图书馆购买的各种全文数据库,如,优秀博士论文数据库 PQDD,中国期刊网,Elsevier Science, Kluwer, AIP 和 APS 等等全文数据库。第四,各大学学位论文全文数据库,据笔者所知,现在有不少大学都要求研究生或者优秀本科生提交学位论文电子版,然后对可以公开的那部分学位论文做成全文数据库,供本校读者使用。

文摘数据库:大部分大学图书馆都购买了光盘文摘数据库和网络文摘数据库,如:中文网络文摘数据库比较典型的有万方文摘数据库,重庆维普网络文摘数据库;西文方面主要有 EI, SCA, CA 等等。

网络信息资源导航数据库:各大学图书馆还根据自己的重点学科建立了一系列的网络信息资源导航数据库,如,中国人民大学的经济类导航数据库,清华大学的 CALIS 学科导航数据库,上海交通大学的管理学导航数据库,武汉大学的水利水电导航数据库,北京理工大学建立的光学工程、化学与化工、车辆工程与电子学与通讯等四个学科的重点学科导航。各大学图书馆还应该根据各馆文献特色再建立一系列的导航数

收稿日期:2003-05-04

作者简介:康桂英(1974-),女,北京理工大学图书馆助理研究员,研究方向为信息资源的组织与揭示。

据库。

学位论文数据库:各大学图书馆已经建立了各校的博硕士学位论文数据库。

特色专题数据库:这是一种深度加工形成的信息组织。由专业人员结合学科和学科建设、结合科研课题和研究领域等,搜集相关网络信息资源并加以专业化的整理和组织,就不同的专题组建各种专题数据库,并跟踪网络信息的变化,定期对数据库的内容进行更新的补充。数据的来源可以是教学、科研机构等上载的学术资源,也可以是信息机构分析、筛选后下载的相关网络资源。其内容更具专业性针对性,检索简便,对学术研究具有很高的价值。

电子图书:有很多高校图书馆购买了大部分的电子图书,国内目前为止在电子图书制作和资源提供比较不错的数字图书馆有超星电子图书、书生之家数字图书和北大方正的 APABI 电子图书等几家数字图书馆。如,北京理工大学图书馆就购买了北大方正集团的 APABI 电子图书制作软件,这个软件不仅可以把以往图书扫描后制作成电子图书,还可以购买各出版社出版的最新图书的电子版本,本校读者可以在校园网内自由借还这些电子图书。

3. 数字信息资源组织技术。

XML 技术(元数据技术)XML 技术主要用来编制各种元数据,现在有很多元数据都是通过 XML 技术来实现的,例如, APABI 电子图书以及学位论文著录工具 DESI 就是 XML 技术来实现的。

扫描技术:扫描技术在资源数字化中占有举足轻重的作用,现在各大学图书馆采访部提倡购买电子图书,要求研究生提交电子版学位论文,以便能很方便地把这些资源数字化,提供给读者使用。那么对于以前的纸质文献究竟怎样数字化呢?这个最好的解决办法就是借助于扫描技术, 将其扫描成 PDF 或者是 TIF 等格式类型,进行组织与揭示后,以供读者在网络在线阅读。

三、资源描述体系

资源描述体系是大学图书馆数字资源组织中最重要的一环。就目前我们身处的这个信息资源大爆炸的社会来说,我们缺少的不是资源,而是能更好地满足用户需求的资源。这就要求如何把数字资源更好组织与描述出来,能让读者最方便地查找到自己需求的信息资源。在笔者看来,各大学图书馆应具备以下几种最基本的资源描述体系。

1. 以 MARC 格式为基础的编目体系。

标准化、规范化是图书馆编目系统中最重要的因素,是信息资源共享的基础。MARC 是把复杂的著录格式计算机化的一个成功发明,它以字段和子字段的组合来表达灵活多变的书目信息,它以自己的字段、子字段的可重复可变量的特性,兼容了传统目录著录项目的可重复与整条款目变量的特性。因此,它让不同的计算机系统共享数据,极大地推动了图书馆自动化和网络化地发展。

就目前来说,各高校针对各种文献信息资源主要有两种最基本的 MARC 编目格式:西文文献资源主要使用 USMARC 格式,中文文献资源则使用 CNMARC 格式。这两种 MARC 格式对各种文献的组织与揭示在一定程度上起到了很重要的作用,一直占据编目体系的主导地位。而且我国最大的联合目录——CALIS 联合目录也主要运用这两种 MARC 格式。

2. 以 XML 和 Dublin Core 为基础的元数据应用体系。

各大学图书馆纷纷建立了一系列网络信息资源导航数据库,对于这些资源,如何进行组织与揭示,也是一件很重要的事情,我们这里所说的“组织与揭示”绝不仅仅意味着将网上学术资源简单地归类、排序,而是序化进而优化,对信息内容进行深加工、再创造以达到信息增值的目的。早在 1994 年, Michigan 大学图书馆信息学院 L.B. Rosenfield 教授就指出:信息管理人员对信息资源进行“重新包装”(Repackaging)的知识和技术,会在使 Internet 变成一个有价值的信息空间中发挥更为重要的作用。那么,究竟如何对这些网络资源进行组织与揭示呢?这就要依靠我们的元数据技术了。

在各种元数据格式中,DC(Dublin Core)都柏林核心元数据格式主要是用于网络信息资源的组织与揭示的。最初由美国 OCLC 公司发起,国际性合作项目 Dublin Core Metadata Initiative 设计,由参与合作项目的机构共同维护修改,每年都召开 workshop,并在会上提出新的修改意见。其最初应用目的是为了网络资源的著录与挖掘,由于 DC 元素简单易用,加之 OCLC 的大力推广和网络资源著录的巨大需求,DC 发展成为可用于任何媒体的组织与揭示。

对于网络信息资源导航数据库的描述,主要用 DC 来实现了。对于那些收集的网络信息资源,大学图书馆可以根据自己的馆藏特色和学科特点来确定应该用 DC 的哪些元素对这些资源描述和揭示,以提供给用户使用。

3. 以其他形式的元数据为辅的元数据应用体系。

元数据的广泛应用是因为现代信息资源处理上的挑战而发展起来的,一是数字资源逐渐成为信息资源的主流,而这些资源从产生、存档、管理到使用都远远不同于传统的纸的介质文献;二是网络和数字化技术使信息的发表既快又便捷,由此而来的海量信息要求有能与现代计算机技术和网络环境相适应的方便、快捷、有效的数据发现和获取方法。

除了 DC 元数据以外,国际上运用得比较多的还有这些元数据:CDWA (Categories for the Description of Works of Art),艺术作品描述类目;EAD (Encoded Archival Description),编码档案描述;FGDC 元数据格式 Federal Geographic Data Committee(联邦地理数据委员会);GILS(Government Information Locator Service),政府信息定位服务;TEI(Electronic Text Encoding and Interchange),电子文本编码与交换;VRA (Visual Resources Association Data Standards Committee),VRA 视觉资料核心类目。

这 6 种主要元数据格式主要针对某一具体的信息资源进行描述,在各种信息资源的组织与揭示方面起着举足轻重的作用。由于大学图书馆的数字资源由各种载体形式和各种内

容组成,为了对某些特殊的信息资源进行描述,各大学图书馆有必要发展这些除 DC 以外的其他元数据的发展,形成以 DC 为主,其他元数据为辅的元数据应用体系。同时,各大学图书馆在组织不同类型的资源时,应根据各自的情况,结合这些元数据的特点制定出特色馆藏元数据体系标准。

4.各大学图书馆应根据各自特色馆藏和技术力量,研究开发相关软件,促进新型元数据的应用。

如上面所述,各大学图书馆都有各自的特色馆藏,如各校的电子版博硕士学位论文全文数据库,为了能有效地组织与揭示这些资源,都八仙过海,各显神通。也有一部分软件公

司针对此开发了一系列的相应软件,有的大学购买了这些软件以后,根据自己的实际情况做出相应的二次开发,制定出一套适合本馆发展的元数据标准在这一方面,据笔者所知,北京大学和清华大学图书馆就做得相当不错,他们根据自己的实际情况开发了不少新型元数据标准。在电子版学位论文的制作和著录方面,北京理工大学图书馆也迈出了关键性的步伐,我们购买了北大方正公司的 APABI 软件来制作与著录电子版学位论文,在实际的应用中,我们摒弃了不适合我们图书馆的一些部分,根据我们的实际情况做了相应的开发,到目前为止,我们的电子版学位论文数据库已经正式面向读者使用了。

参考文献:

- [1] 马喻.网络中文社科学术资源的组织与开发利用[J].现代图书情报技术,2001年第3期.
[2] <http://www.lib.pku.edu.cn>.

On the Organization System of the Digital Resources in a College Library

KANG Gui-ying

(Beijing Institute of Technology, Beijing:100081)

Abstract: This paper mainly discusses the organizing system of the digital resources of a College Library.

Key words: College Library; Digital resource; Organizing system.

(上接第 10 页)

参考文献:

- [1] 苏汉林.教育创新:重在培育创新型教师队伍[J].云南教育,2003年第6期.
[2] 刘瑜.关于高校实施创新教育的思考[J].中国高教研究,2001年第4期.
[3] 常顺英.课堂教学要重视学生创新能力培养[J].高等理科教育,2002年第2期.

The Innovative Education of 21th-century Colleges in China

CHANG Shun-ying XI Qiao-juan

(Beijing Institute of Technology, Beijing: 100081)

Abstract: The Innovative Education of 21th-century Colleges in China plays an important role in the compliment of overall qualities education among the students. This paper discusses some problems of how to adjust the educational guidance to a better model,how to stimulate the students' desire for innovation and cultivate their spirit of innovation,on the basis of which we tend to develop the innovative education up to a new surging movement in the new century.

Key words: College; Innovative education.