



[上]

高凡◎主编

大学图书馆

Daxue Tushuguan Fazhan Yu Chuangxin

发展与创新



西南交通大学出版社
[Http://press.swjtu.edu.cn](http://press.swjtu.edu.cn)

Library

大学图书馆发展与创新（上）

主 编 高 凡

西南交通大学出版社

· 成 都 ·

内容简介

本辑丛书主要讨论了信息时代大学图书馆面临的发展与创新问题。共分为图书馆事业、图书馆管理、信息组织与信息技术、文献信息资源建设、信息服务、用户教育、队伍建设等 7 个栏目, 共刊登论文 55 篇。深入广泛地探讨了信息组织、数字图书馆技术、大学图书馆管理与改革、学科资源建设、学位论文管理、图书馆个性化信息服务、信息素质教育、图书馆馆员岗位培训、学科馆员制度等有关理论和实践问题。

图书在版编目 (C I P) 数据

大学图书馆发展与创新 / 高凡主编. —成都: 西南交通大学出版社, 2004.5
ISBN 7-81057-871-5

I. 大... II. 高... III. 院校图书馆 - 图书馆管理
- 文集 IV. G258.6-53

中国版本图书馆 CIP 数据核字 (2004) 第 047269 号

大学图书馆发展与创新 (上)

主编 高 凡

*

责任编辑 王 婷

封面设计 何东琳设计工作室

西南交通大学出版社出版发行

(成都二环路北一段 111 号 邮政编码: 610031 发行部电话: 87600564)

<http://press.swjtu.edu.cn>

E-mail: cbsxx@swjtu.edu.cn

成都市蜀丰印刷厂印刷

*

开本: 787mm×1092mm 1/16 印张: 14.75

字数: 346 千字

2004 年 5 月第 1 版 2004 年 5 月第 1 次印刷

ISBN 7-81057-871-5/G · 066

售价: 56.00 元

图书如有印装质量问题, 本社负责退换

版权所有, 盗版必究, 举报电话: (028) 87600562

序 言

近两年来,发展与创新成为图书馆界的一大研究热点。图书馆界有识之士对这一问题的深入研究,已不仅仅简单地满足于图书馆自身厚重的历史积累和目前图书馆发展的表面繁荣,也不仅仅满足于业务更新、技术更新,而是充满着忧患意识、危机意识和强烈的历史使命感,正在深入地从多个侧面探求促进图书馆发展的内在动力及其发展前景,探求图书馆变革的途径和方式。大学图书馆是学校的文献信息中心,在当前急剧变化的环境中,已从“重形式”发展为“重内容”,即馆舍从重硬件建设到重内涵发展,资源从重外形(概念)揭示到重内容开发,管理从以“书”为本到以“人”为本。具体而言,大学图书馆在未来的发展中,还面临着以下几方面的关键问题:

1、新的角色和任务:需要修订图书馆的发展目标以满足网络环境下不同读者的现实和潜在需求;积极参与高等教育的教学改革,扩大图书馆在学校中的影响,使图书馆的发展方向与学校的方向保持一致;就经费、知识产权和电子资源的使用等问题,加强图书馆和学校各院、系、处、所的沟通;加强用户,尤其是学生的信息素质教育,发挥图书馆强有力的教育功能;

2、馆藏资源:重视大学图书馆在学术出版中的作用,加强学术信息交流体系的动态合作;采集用户所需的非文本信息(如图像、数据、声音、GIS等);重视收集网络信息资源,尤其是学术信息资源;资源采集强调使用权,而不是实际资源的拥有权;在不同的e-Resource之间建立无缝连接;

3、服务:加强合作,增加国际合作机会;图书馆需要做更深入和广泛的用户调查和评估;需要构建便于用户使用的服务系统,为用户提供发现和使用图书馆资源的途径,提高馆藏利用率和管理效率;为用户提供可广泛获取的E-Service及其工具;

4、人员队伍:发挥图书馆员的潜力,保持一线工作人员能够适应未来的挑战;重视人才,提高图书馆员的待遇,应对激烈的人才市场竞争;加强馆员培训,

提高员工变革和创新的能力;强调学科馆员的重要性,加强学科馆员队伍的建设;

5、图书馆领导:保持图书馆和校园内其他部门的良好关系,加强沟通,加强合作;加强管理,强调以人为本的人性化管理;研究图书馆发展趋势,重视图书馆战略发展规划;

6、长期保存:积极保存现有的特色资源,尤其是本校教师的数字原生资源;为读者长期利用古籍、特藏等易损资源提供不同的服务策略;与商家合作,考虑电子资源的长期保存。

要解决以上问题,大学图书馆必须顺应时代变革,以变应变,以变革求创新,以创新迎接挑战,推动图书馆的可持续发展。我们需要不断探索新世纪图书馆应有的功能、价值观、发展目标及使命。西南交通大学图书馆针对这一热点问题组织开展专题讨论,并集结成文,是积极的、有意义的探索。我相信通过我们图书馆界同仁的努力,从理论层面和实践层面来总结、归纳大学图书馆的发展规律,探索、发现大学图书馆发展和创新的途径,将极大地推动我国图书馆事业的发展。

徐引篪

2004年5月于北京中关村

前 言

在传统的大学学术信息交流体系中,图书馆是主要的甚至是对信息进行可靠收集、存储、组织和直接提供信息服务的机构,图书馆的核心能力主要体现在文献组织、检索与传递等方面。但是,数字信息环境正急剧地改变着这一切。现代科学技术的迅速发展,使文献信息从产生、传递到使用的全过程都发生了巨大的变化。大量数字化、网络化的资源成为除传统印刷型资源外的重要资源。日新月异的网络环境不仅改变了大学图书馆多年来传统的收藏、使用和服务模式,而且打破了图书馆之间彼此独立、封闭的壁垒。数字资源的产生和使用、高校网络化进程的不断加速,使信息获取更加直观和便捷,信息资源、信息服务、信息交流和用户工作过程聚合在同一数字空间,使用户的信息获取与利用空间整体地处于数字化状态,整个校园的信息交流平台和模式发生了剧烈变化。同时,新兴的信息服务机构纷纷出现,它们凭借自身的计算机、通信技术优势,灵活的经营机制,使得图书馆不再是用户寻求信息服务的唯一渠道。E-university, E-science 环境下的大学图书馆何去何从?大学图书馆如何应对网络化、数字化的挑战?大学图书馆如何在对手如林的网络环境下寻找自己的生存与发展空间?

历史的经验告诉我们,世上没有不朽的存在,有变革才能有创新,才能有发展。创新是图书馆事业发展的灵魂。我们应深入研究大学图书馆在 E-university、E-science 中的地位和作用,变革根植于传统图书馆框架下的业务流程、服务模式以及组织机构设置,加强信息资源建设,调整工作重心和方向,提高管理水平,改善人员结构,以创新作为大学图书馆改革和发展的根本原则,使大学图书馆真正成为校园学术交流体系中的信息枢纽和信息资源集散地。

正是基于这样的想法,并本着原铁路高等院校图书情报书刊一贯追求的“求实、探索、改革、创新”的精神,我们确定了 2004 年的书名为《大学图书馆发展与创新》,其主题是:信息时代大学图书馆发展与创新的理论与实践,副主题

有：(1) 大学图书馆发展模式与功能；(2) 大学图书馆管理与改革；(3) 网络环境下信息资源管理、开发与利用；(4) 大学图书馆服务模式与用户需求；(5) 大学图书馆的社会职能与作用；(6) 数字图书馆、复合图书馆与图书馆未来；(7) 知识管理、知识服务与图书馆发展、创新；(8) 核心竞争力与图书馆发展、创新；(10) 其他。

发出征文通知后，共收到来自北京交通大学、石家庄铁道学院、中南大学、大连铁道学院、兰州交通大学、华东交通大学、中国科学院文献情报中心、中国科技信息研究所、中山大学、武汉理工大学、吉林大学、湖南吉首大学、绵阳师范学院、广西中医学院、新泰市职工中专、重庆电力高等专科学校和西南交通大学等单位的来稿百余篇。经认真审阅，本书共刊登论文 55 篇，主要讨论了信息时代大学图书馆面临的发展与创新问题，分为图书馆事业、图书馆管理、信息组织与信息技术、文献信息资源建设、信息服务、用户教育、队伍建设等 7 个栏目，深入广泛地探讨了信息组织、数字图书馆技术、大学图书馆管理与改革、学科资源建设、学位论文管理、图书馆个性化信息服务、信息素质教育、图书馆馆员岗位培训、学科馆员制度等有关理论和实践问题。我们希望，通过本次专题探讨，能为网络环境下大学图书馆的发展与创新做一些有益的探索与思考。

在此对所有关心和支持本书的作者和单位表示衷心感谢，特别是我的导师、中国科学院文献情报中心博士生导师、国家科学数字图书馆项目中心主任徐引麓研究馆员在百忙之中为本书作序，特表示最诚挚的谢意。西南交通大学图书馆杨元馆长长期以来对《丛书》的建设与发展给予了密切地关注和切实、有力地支持；北京交通大学图书馆萨殊利馆长、赵亚虹副研究馆员，西南交通大学图书馆赵颖梅和路强洲等同志在本书的组稿、审校、宣传、发行等方面做了大量的工作；西南交通大学出版社对《丛书》的出版给予了大力的支持，在此一并致谢。

高 凡

目 录

信息组织与信息技术

数字图书馆中知识组织系统的功能与实现	岳 红	1
美国数字图书馆中元数据的应用研究	胡利勇	7
数据仓库技术及其在图书馆的应用	熊拥军	13
基于Web的管理信息系统设计方法	陈春颖 王云祥	17
浅谈改进基于Web的OPAC功能	陈大建	21
高校图书馆网页建设应实现的功能	刘文杰	26
搜索引擎的使用体会	邓发云	29
信息资源存储与图书馆	王李霞	33

图书馆管理

数字化图书馆管理中的读者隐私保护问题	周从军	37
图书馆人性化管理探讨	肖晓燕	42
高校图书馆管理体制的比较分析	蒋建林	45
学习型社会中的图书馆	赵颖梅	49
将图书馆建成“学习型组织”	陈苏静 任 颖 王翠菊	53
数字图书馆带给图书馆管理发展的新思路	骆小平	56
大学图书馆管理与改革	方桂英	60
基于层次分析的高校图书馆经费支出预算	焦丽辉 陈艳春	64

文献信息资源建设

大学图书馆的Internet学科资源建设现状	陈定权 吴振新	67
高校学位论文的版权归属与使用尺度	许 勇	73
对高校图书馆学位论文管理工作的思考	丁 娜 郑 兰	80
关于非书资料回溯建库的思考	黎 明 朱亦丹 胡 琳	84
谈书刊附盘的管理	姚道先 杨玲艳 孙静茹	87
试论中文电子资源的著录	胡 琳	90
信息网络时代大学图书馆期刊工作探讨	邓 嵩	96
高校图书馆对网络电子期刊的开发利用研究	孙武秀 黄玉莲	99
谈影响高校图书馆文献采访工作的四个因素	郑 兰 韩志航	102
谈出版物国际交换	张爱霞	106
在外文原版采购中利用网络书评资源	高 林	111
浅谈网络环境下高校图书馆的西文采访工作	宋 雪	115
论文献资源建设中的质量控制	秦建宁	119
谈我馆特藏文献的保护	李红梅 孟宪梅	124

信息服务

论网络环境下个性化信息服务新模式·····	王 进	127
建设以用户为中心的大学图书馆服务·····	王玉梅	130
刍议高校图书馆的优质服务·····	刘润平	134
民族地区高校图书馆为社会服务的思考·····	殷 黎	138
信息技术推动图书馆从被动服务转向主动服务·····	吴振新 陈定权	142
对未来图书馆读者服务的思考·····	朱玉兰	146
浅论图书馆读者服务工作·····	陈烈英 黄萍建 何武俊	149
研究生的借阅倾向分析·····	曾国秀 方太强 郭志梅	152
北京交通大学图书馆科技图书阅览室读者利用情况统计分析·····	刘素范 陈 梅	157

用户教育

高等教育中信息素质教育的标准及主要内容·····	姜 琳	162
高校开展信息素质教育的探讨·····	邓要武 赵亚虹	167
高校图书馆信息用户教育研究·····	鄢少青 史东风	171
浅谈高校图书馆与大学生素质教育·····	孙静茹 刘丽杰 唐淑娟	175
高校图书馆教育信息化工作·····	陶昌华	178
充分发挥图书馆在素质教育中的作用·····	吴绪芹	181

队伍建设

图书馆馆员岗位培训体系研究初探·····	萨殊利 王星华 邓要武	184
康涅狄格大学学科馆员制度及其启示·····	李春旺	189
高校图书馆一线馆员结构的现状及思考·····	杨 媛 黄 琳 赵 钊	195
提高素质,做一名合格的图书馆员·····	杨晓文 杨晓丹	199

图书馆事业

网络文化对我国高等教育的影响及其对策·····	杜新中	201
论图书馆的人文精神和人文教育·····	马 倩	208
知识管理与图书馆创新·····	张金科	213
数字图书馆的未来·····	张 宇	218
数字图书馆的著作权问题与著作权保护·····	张晓颖	222
论转型期大学图书馆的庸俗化·····	龚 菲	225

数字图书馆中知识组织系统的功能与实现

岳 红*

(武汉理工大学 武汉)

【摘 要】 文章阐明了知识组织系统在数字图书馆中除了具有检索功能外,还必须具有强大的信息链接与动态重组功能和使数字图书馆资源面向多样化用户的功能。知识组织系统在数字图书馆的功能的实现必须进行仔细地策划。

【关键词】 数字图书馆 知识组织系统

1 引言

数字图书馆的发展历程表明:数字图书馆实行由信息提供向知识提供的转变,是其产生和发展的必然趋势,也是其迎接知识经济挑战与时俱进的必然选择。数字图书馆信息提供的知识化,必然要求数字图书馆对其数字化信息资源进行知识组织。

知识组织系统(knowledge organization systems,简称KOS)是信息(知识)组织的工具,是对内容概念及其相互关系进行描述和组织的机制。可对各信息对象按照知识内容和知识结构进行描述、连接和组织。根据对概念关系的揭示程度,知识组织体系可分为3个层次:

(1) 词汇表:强调概念的定义,一般不涉及复杂语义关系和分类结构,例如规范档、词汇表、字典、地名表等;

(2) 分类聚类体系:强调概念间的层级聚合和类别体系,例如主题词表和分类表;

(3) 关系列表:强调表现概念之间关系,例如叙词表、语义网络和概念集。

数字图书馆一般使用一个或一个以上的知识组织系统。数字图书馆中知识组织系统的目的就是管理数字图书馆的收藏、进行内容组织并实施检索功能,同时引导读者实现知识发现的过程。因此,除了检索功能外,数字图书馆中的KOS还必须具有强大的信息链接与动态重组功能和使数字图书馆资源面向多样化用户的功能。

所谓信息链接与动态重组,指在已建立的信息资源体系基础上,将有关信息内容链接起来形成新的信息链或信息集合,从而实现信息的重组。这种链接重组主要不依靠作者事先建立的链接点,而是专门建立相应的链接机制。有意识地根据用户特定需要和信息内容内在关系来进行面向用户的信息组织,可以看成是用户界面层的信息组织机制。

2 将数字图书馆资源链接到其他相关资源

利用知识组织体系,数字图书馆的信息内容(例如词汇、符号代号、段落或文献)与相

*岳红 硕士 副研究馆员 武汉理工大学图书馆 邮编:430070

关的知识组织体系链接,包括与词汇表、叙词表、分类表等的链接,可解释有关信息内容、按知识组织体系显示及组合信息内容,或进一步支持利用知识组织体系的映射,如将代码链接到全文,提供更多描述性记录,或将实体名称链接到物理样本资源。这种链接的基础是将那些能从数字资源内的信息中抽取出来,并在 KOS 中进行检索的信息识别出来。

2.1 将代码扩展到全文

一些行业采用代码体系来加快行业内部的信息交流,将这些代码体系和它所代表的全称联系起来是非常有益的。如美国国家医学图书馆(NLM)PubMed 检索服务系统,就可以通过检索档、文本分析程序和关键词索引,将 PubMed 检索结果包含的 GenBank 代码识别出来,并在文本中嵌入一个活链。GenBank 代码是美国国家生物技术信息中心网站拥有的 GenBank 和结构生物信息蛋白数据库中的生物序列号,每个生物序列号开头部分都是固定不变的数据库识别代码。点击活链,即可链接到 GenBank 中找到并显示该代码代表的生物序列号的记录。

2.2 链接到描述性记录

将实体的名称,如个人姓名、组织机构、地点,链接到关于该实体的附加信息是超链接的最早用途之一。字典、词汇表、分类表等知识组织系统能够用于将一个资源中的实体链接到另一个资源中对该实体的更为详细的描述。

2.2.1 将有机体的名称链接到分类记录

使用分类规范档作为中介规范文档,就可把含有有机体名称或图像的文本或图像文件链接到相关的附加信息。如集成分类信息系统(ITIS)就是一个分类中介文档。ITIS 是分类专家合作开发一个在线北美植物和动物名称列表。包含有关北美及世界的植物、动物、真菌和微生物的权威性分类信息。如果有机体的名称能够在数字图书馆文本资料中手工或自动识别,并作为提问递交到 IT IS 数据库,IT IS 找到匹配的记录后,就能提供包括该有机体的基本信息的记录,包括它的同义词、常用名及在更大的分类体系中的位置。

2.2.2 链接化学名称到分子结构

BIOSIS 是世界上最大的非盈利性的生物和生物化学数据库生产商。BIOSIS 指定用 CAS 注册号码(RNs)识别化学名称。通过 STN 国际在线系统,一个 BIOSIS 的用户可以从一个或更多的检索结果中提取 RN 代码。被提取的 RN 能在 CAS 注册文档检索(该文档包括有机物、无机物、生物序列、金属和合金等 21 000 000 化学物质),结果提供化学名称同义词的链接及其分子结构。利用 CAS 开发的特殊工具,还可以观看和操纵这些分子结构。

2.2.3 链接个人姓名到书目信息

人名规范文档控制人名的各种变化形式,是常见的规范文档。如 GALE 自传资源(包括 142 000 自传和 1 000 种期刊中的相关引文)就利用人名规范文档从文本中发现人名,并自动链接到自传文章。卡内基集团开发的 NameFinder 以及 IBM 公司开发的智能代理机(Intelligent Agent)是这种链接中需要用到两种文中人名定位工具。

2.3 链接实体名称到物理样本

艺术、考古、文化历史研究者或团体、展览会和博物馆正将其收藏的实物和样本目录逐步数字化,作为知识组织系统,这些目录不仅提供描述性记录,而且指明了实物或样本在博物馆、档案馆和其他收藏单位的具体位置。一般每个收藏品都会被赋予一个识别码,若数字图书馆的资源中的文献包含这些识别码,这些识别码就能被提取出来并在网络目录或数据库

中检索,检索结果中包含实物的收藏地点和联系信息,使用户可以获得数字图书馆资源中提到的实物。目前许多期刊文献都要求作者提供识别码,并可以链接到研究者或组织机构维护的样本网站。

3 使资源向多种类型的用户开放

KOS通过提供交替主题途径、用多种方式显示和组合信息内容、支持多语种检索、支持自由文本检索等方法将网络上分散的用户的联结在一起,提高对用户的作用层次和作用力度,使数字图书馆真正成为用户的核心合作伙伴。

3.1 提供交替主题途径

要使数字图书馆资源能面向不同的用户,需提供多个主题检索入口。通过以下途径可以为用户提供多个检索入口:

(1) 用多种检索语言进行标引:此方法成本较高。

(2) 保留加入数字图书馆的成员的原有知识组织体系。如美国国防部、美国能源部等机构的书目记录加入美国 NASA 航空航天数据库时,各成员的受控词条均被保留。

(3) 建立多种系统间的映射:在新的 BIOSIS 关系标引结构中, BIOSIS 建立了并维护着将自然语言的疾病名称与 MeSH 受控疾病术语联系起来的规范档。BIOSIS 将记录提供给 NLM'S TOXLINE 数据库时,根据毒物学文献中出现了自然语言词汇和 BIOSIS 中使用 Mesh 术语标引的正式自然语言关键词的映射关系,这些记录将被自动处理,适当的 Mesh 术语也被作为交替检索点被记录下来。

交替主题检索仅用于资源级的检索,如数据库、电子书、电子期刊或图像库,以便于其他团体识别感兴趣的资源然后再进行单独检索;如果数字图书馆有详细的书目记录或元数据,就有可能开发出将原有数字图书馆或资源的概念转换为交替项目的转换程序。

3.2 用多种方式显示和组合信息内容

KOS 可以为数字图书馆馆藏资源提供新的信息显示和组合方式。在数字图书馆环境,这些内容可以作为“层”添加在一个或多个标的上,如文本检索的结果可以根据名词术语出现的频次或文献的国别、刊名、著者以图像或视频形式显示。

以地理信息数字图书馆“geolibrary”为例,它主要用地名检索,如果一个 geolibrary 书目记录或全文中有经纬度坐标,坐标值就可以被提取出来递交到地名表系统,地名表系统可将地理空间区域的位置和地名联系起来,返回这个区域的名称;然后通过在地名表中按地名的再度搜索,找到一个地名对应的多个不同区域,并将地名和对应于地图上的区域显示出来;用户对地名的地理区域加以识别并确定后,用户就可以通过地理制图工具检索文本资源。

3.3 支持多语种检索

欧洲环境局 (EEA) 的综合多语种环境词表 (GEMET) 有 12 种语言。德国的欧洲书目数据来源主题中心正在开发一个将数据源和虚拟图书馆的元数据信息链接起来的系统,而 GEMET 将被这个系统用来进行语种转换。这样,用户不仅可以本国语言检索,也可以用其他语言检索。即 EEA 的所有数据库系统都可作为虚拟图书馆馆藏,用多种语言从受控词表和自由文本两种途径检索。

3.4 支持自由文本检索

自由文本检索是 WEB 网上主要的检索方法。网络资源中仅有一部分有元数据,有控制词表的更少。利用 KOS 可以帮助用户选择自由文本关键词。例如,GETTY 词表项目就越来越多地被搜索引擎用来寻找同一概念的同义词。KOS 还可用作网络挖掘工具,Northern Light 就开发了一个命中结果相关度很高的网络挖掘工具。支持 Northern Light 的 KOS 是一个容纳大量现有词表和主题表的系统。检索结果被组织在 Northern Light 开发的分类表的“文件夹”下,这些高级分类目录帮助用户鉴别同一术语的意义。例如“pitcher”可能放入两个文件夹。一个是“sports”(作为垒球投球手的意思),一个是“decorative arts”(作为水罐的意思)。选择“sports”,系统仅提供有关垒球投球手有关的网络资源,选“decorative arts”,则提供与水罐有关的资源。

KOS 一般是为一个专门学科、任务、功能,为专门馆藏和数据库做索引而开发。因此,在特定的使用领域和系统,KOS 可以非常有力地支持自由文本检索和将网络资源整合的到现有的数字图书馆。

4 数字图书馆的知识组织系统的策划与实现

4.1 策划

4.1.1 分析用户需求

数字图书馆选用 KOS 前,必须全面了解用户环境和数字图书馆内外内容链接的可能性,同时考虑可通过 KOS 进入数字图书馆的用户意见。

4.1.2 选择 KOS

由于建立一个 KOS 费时费力,所以最好找一个现成的 KOS。通过查询用户、目录、参考资料、图书馆员或 INTERNET 搜索可以发现有用的 KOS。

4.1.3 构造

对于那些可以上 WEB 网的系统,将 KOS 作为外部系统进行链接是可行的。这种体系机制要求通过脚本或检索提问来确定资源的位置,如全球资源定位符(URL)或其他提问。提问脚本传递的是从数字图书馆系统到外部 KOS 的登录信息(包括用户 ID 和口令),以进入该网络数据库。

链接到 KOS 的 URL 地址可能发生改变,而数字目标识别器、永久地址计划、使用较新浏览器的统一资源识别(URI)和统一资源名称(URN)有助于解决这一问题。

KOS 也可从其开发者那里获取并在本地安装。这时不需要登录信息,KOS 可本地控制,访问速度加快,但要考虑维护和本地系统管理以及磁盘空间、防火墙、接口设计的问题。

如果数字图书馆计划囊括许多 KOS,必须考虑系统结构的规模。美国国家图书馆医学 UMLS 包括 40 多种不同的词表。它的主要目的是开发一个在这些词表间移动的元词表,这时系统的管理是主要被考虑的问题,而映射问题则不需考虑。

许多数字图书馆系统是现有集成图书馆系统(ILS)的扩展和应用,考虑 KOS 如何与图书馆系统集成是很重要的。由于许多 ILS 开发商在设计系统时没有考虑对外部文件或数据库的链接,在一些情形下,开发商可能要求文件存贮在 ILS 的目录或服务器上才能进入 ILS。

外部的通信则可能要求必须是 Z39.50 C/S 连接。可链接或可检索的字段也受到限制。这种相对封闭的系统,使得 ILS 可能成为应用 KOS 的不利环境。所以应鼓励开发商开发出支持相对开放的结构和使传统图书馆系统支持更多的数字图书馆功能的系统。

4.1.4 维护 KOS

在策划 KOS 时系统及内容的维护都应加以考虑。KOS 版本控制非常重要,根据需要可以选择重装新版和仅传递变化的数据两种方法。

4.1.5 将 KOS 呈现给用户

KOS 可以以不同方式展现在用户面前。馆藏可以归属到数字图书馆网站 KOS 相关主题或类目下;KOS 可作为不同的用户识别检索点的高级工具;如果数字图书馆的内容包括元数据记录,KOS 可以显示为索引条目或检索导航用的词条;KOS 也可以是透明的,即词表可在幕后使用,进行同义词扩检、链接数字图书馆资源到其他信息资源,及对获得的信息过滤或排序。

4.2 链接的建立

建立数字图书馆与 KOS 的链接有两部分。一是在数字图书馆资源中发现关键锚点,二是在目标文件中查询。根据不同的情况,允许“在线实时”链接和“嵌入式”链接。

第一步,查阅与数字图书馆资源有关的元数据。元数据包含需要做链接的术语(如 SIC 代码、艺术家名称、地点名称、地理坐标),且是用户将会感兴趣的特定资源的主题,这个元数据就能用来做链接。如果数字图书馆中的文本资源没有合适的元数据,就必须进行文本分析,识别关键信息。如果文本字符串有特别的特征,如数据库首字母、专门结构(如经度、纬度),可使用执行简单的字符串搜索的程序或能保留命中位置的搜索引擎。如果文本字符串不包含此类线索,则须采用文本挖掘或更复杂的文本分析工具,这些工具使用了许多语义和句法算法来发现关键信息,如 IBM 智能代理机,它包括专门的识别人名和地名的算法。

第二步,链接到 KOS。根据 KOS 是本地安装还是远程使用的区别,所采用的链接方法也不同。如果是本地安装,通过执行大量的运算处理使这两个文件匹配是有可能的。如果是远程使用,这个交互作用需要运用脚本知识和各种基于网络的检索技术。

如果 KOS 作为数字图书馆资源和其他资源的中介文档使用,必须了解执行检索和返回数据的数据和过程。如果 KOS 必须给数字图书馆返回一个值,那么这些数据 and 过程必须进行双向评估。

选择链接工具也十分重要。链接可以是固定的或实时的。采用固定形式,特定的 URL 要求被嵌入数字图书馆资源链接点。也可使用 URN,URN 要求在目标文件点创立一个名称空间,这样,检索就会针对这个名称空间而不是 URL。固定定位器(PURLs)和数字目标识别器(DOIs)也可解决这个问题。这些方案适合 HTML 形式的文献。

关于数据库的内容检索问题,美国国家医学图书馆目前通过 IGM URL 功能支持对其多个数据库的检索。IGM 用户可创立执行数据库检索的 URL,例如,下面的脚本可以执行 HealthSTAR 文档对“pneumonia”的检索:

```
http://igm-02.nlm.nih.gov/cgi-bin/IGM_robot.pl?datafile = HealthSTAR&search =  
Subject-pneumonia.
```

NLM 网站提供创立这样的 URL 的句法信息。意图是将检索的 URL 被个人用户作为书签收

藏, 同样的概念可以用来在锚点创立一个活链, 和其他的脚本一起, 名词 pneumonia 的创立能自动被链接点的活链取代。

参考文献

- 1 Gail Hodge. Systems of Knowledge Organization for Digital Libraries Council on Library and Information Resources, 2000
- 2 张晓林, 李宇. 描述知识组织体系的元数据. 图书情报工作, 2002 (2): 64~69

美国数字图书馆中元数据的应用研究

胡利勇*

(中山大学信息管理系 广州)

【摘 要】 数字图书馆中,元数据充当了重要的角色。作者考察了美国几个实际运行的数字图书馆,着重探讨它们在信息资源建设中元数据方案的应用情况和实施方法,并对这几个图书馆中元数据应用的特征进行归纳总结。

【关键词】 数字图书馆 元数据 美国

数字图书馆将分散于不同国家和地区的不同载体类型的信息资源经过有序整理后,采用数字方式存储,通过网络途径传递,使用计算机管理和利用,形成一个高度组织的、规模庞大的网络信息资源库。数字图书馆中的信息资源具有数量庞大、媒体形式复杂多样、分散式分布等特点,使得数字图书馆中信息资源的有效组织与利用成为难题,元数据则为解决这一难题提供了有效的途径。

元数据(Metadata)是关于数据的数据,是描述某种类型资源的属性、并对这种资源进行定位和管理、同时有助于资源检索的数据。元数据被用来描述和标引网络信息资源,为网络信息资源提供了有效的组织和管理机制。在数字图书馆中,元数据更是得到了前所未有的广泛应用。

美国是数字图书馆项目实施的先导,早在20世纪80年代后期,美国就率先开始数字图书馆的研究和建设。1994年开始的“数字图书馆先导研究计划”便推动了美国数字图书馆发展,很多组织、社团都参与了数字图书馆的研究活动。随着众多数字图书馆项目的开展与实施,现在已有一批数字图书馆初具规模。笔者从中选择了几个有代表性的数字图书馆,通过实际考察,对这些数字图书馆中元数据的实际应用情况做了论述。

1 加利福尼亚在线档案馆(Online Archive of California, OAC)

OAC是加利福尼亚数字图书馆的核心组成部分之一,它专门建立了一个查找帮助(Finding Aids)数据库,利用查找帮助(如目录、索引,以及图书馆、博物馆等为了方便使用它们的资源而建立的文档等),为加利福尼亚各图书馆、博物馆、档案馆以及其他机构中的手稿、照片和艺术作品等数字化资料的联合收藏提供链接和检索功能。

查找帮助是一种用来描述和识别信息资源的数据,相当于一个资源的向导,它们提供从多种角度对资源进行分析和描述的信息。与通常的目录相比,查找帮助包含了更加完整的资

*胡利勇 硕士研究生 中山大学信息管理系 邮编:510275

源描述信息。

OAC 中查找帮助采用的是编码文档描述 (Encoded Archival Description, EAD) 元数据模式。EAD DTD (Document Type Definition, 文档类型定义) 是一种采用 SGML 或 XML 文档编码的元数据标准, 它最早是在 1993 年由加利福尼亚大学伯克利分校图书馆开发的, 最初的目标是为可供机器识别的查找帮助提供一种共享的编码标准。

EAD DTD 描述的对象是文本文档, 主要用于描述档案、手稿和查找帮助等资源, 并利用网络检索和获取这些文本信息资源。EAD 由头标, 前页, 档案描述三部分组成, 每一部分又包含多个小项及若干细项。它能适应任何长度的目录和记录, 并能描述在各种媒介上的所有类型的档案。

在美国, 采用 EAD 标准的数字图书馆项目还有很多, 如美国遗产虚拟档案项目创建一个 EAD 查找帮助的共享数据库, 以提供对美国历史和文化的检索服务。

2 美国记忆 (American Memory)

美国记忆由美国国会图书馆创建和维护, 它向用户提供取自 100 多个历史收藏的 700 多万份数字化资料, 这些资料包括数字化的历史文本、照片、地图、视频、音频等各种媒体形式。国会图书馆对不同类型的历史资料, 采取了符合国际和行业标准格式的数字化手段, 如: 文本采用 SGML 或 PDF 格式, 图像和照片采用 TIFF 和 JPEG 两种格式分别存储, 声音文件采用 WAVE 和 RealAudio 两种格式, 等等。

美国记忆采用 SGML 格式的主要有两类资源: 一是查找帮助, 采用 EAD 元数据方案; 二是全文书籍、小册子、手稿及其他历史文献, 则采用文本编码倡议 (Text Encoding Initiative, TEI) 元数据方案。

TEI 也是采用 SGML 文档编码的元数据标准, TEI 最早于 1987 年由美国多个组织机构共同创建, 起初是为人文学科方面的文本而编制的, 它试图克服古老的历史文献编码模式多样化带来的困难, 以便不同机构组织间交换编码文本。随着 TEI 模式的不断扩充和发展, 它的适用范围已不再局限于古文学方面了。

为了使元数据成为文档的一部分, 同时为编目提供准确和完整的信息来源, TEI 的成员还专门编制了一套独立的关于 TEI 标头 (Header) 的细则。TEI 标头位于 TEI 文件的起始部位, 以 <teiHeader> 作为标识符, 由文档描述、编码描述、文本特征、修订记录四个部分构成, 分别以 <fileDesc>、<encodingDesc>、<profileDesc>、<revisionDesc> 作为标识, 其中只有第一部分文档描述是必需的。

美国记忆为每个 SGML 文本文件都增加一个 TEI 标头, 由于文本文档已经用目录或者查找帮助进行标引, 因此, 美国记忆实际采用的 TEI 标题是一种经过改良的、简化了的 SGML 单元集。例如, 美国记忆中 TEI 标头如表 1 所示:

3 亚历山大数字图书馆 (Alexandria Digital Library, ADL)

ADL 是加利福尼亚大学圣巴巴拉分校的数字图书馆计划, 参与者还包括纽约州立大学布法罗