

我国高校数字图书馆综合门户构建

李文文, 陈 雅, 郑建明

(南京大学 信息管理系, 江苏 南京 210093)

摘要:以CALIS、CADAL和CASHL等资源为基础,在高校CCC模式的基础上,构建包括资源层、接口层、管理层、服务层和用户接入层的我国高校数字图书馆的综合型垂直门户,实现我国高校数字图书馆资源和服务共建、共享的目标。

关键词:高校数字图书馆;图书馆门户;资源整合;身份认证;P2P

中图分类号:G350 **文献标识码:**A **文章编号:**1007-7634(2012)01-55-05

The Construction of China Academic Digital library Integrated Portal

LI Wen-wen, CHEN Ya, ZHENG Jian-ming

(Department of Information Management, Nanjing University, Nanjing 210093, China)

Abstract: The integrated portal of China academic digital library including the resource layer, interface layer, management layer, service layer and user access layer based on CALIS, CADAL and CASHL has been constructed in this paper in order to achieve the target of sharing the resource and service of academic digital library.

Key words: academic digital library; library portal; resource integration; identity authentication; P2P

近几年,伴随着互联网技术的发展,门户(Portal)已成为IT及其他行业的热门词汇,信息门户、Web门户、校园门户、企业信息门户逐渐被人们认可,信息门户被引用到各个领域。信息门户技术引入到数字图书馆,即数字图书馆门户正逐渐兴起。随着越来越多的用户远程访问图书馆的资源,用户更倾向于通过一个界面,及统一的接口去无缝地访问数字图书馆所有的信息资源和服务^[1]。美欧一些国家的数字图书馆门户已经在网上运行,而中国高等教育数字图书馆(CALIS)2002年也把数字图书馆门户纳入到数字图书馆建设的总体规划当中^[2]。目前国内各个高校数字图书馆都在积极的开展门户建设,但是在目前信息膨胀和资金有限的条件下,各个高校的单一门户具有一定的限制性,因此本文主

要探讨以CALIS、CADAL和CASHL等资源为基础,在高校CCC模式基础上,构建包括资源层、接口层、管理层、服务层和用户接入层的我国高校数字图书馆的综合型垂直门户,实现我国高校数字图书馆资源和服务共建和共享的目标。

1 我国高校数字图书馆综合门户建设可行性分析

1.1 门户(Portal)

所谓Portal,指的是集成了多样化内容和多种服务模式的Internet网站。由于这种类型的Internet网站提供搜索引擎、免费资源服务、个性化服务、

收稿日期:2011-10-04

基金项目:教育部规划基金项目(10YJA870002)

作者简介:李文文(1986-),女,辽宁盘锦人,硕士研究生,主要从事数字图书馆建设与评价相关研究。

新闻发布、虚拟商店、虚拟社区等多种功能,集成了当前Internet网络的多种服务模式,因而使得它们成为了Internet网民进进出出的必经之门户,同时也构成了Internet的中坚力量。不论是以AOL为代表的ISP,还是以YAHOO为代表的搜索引擎,都是人们进入网络环境、开发利用网络资源的一个门户。从实质上讲,Portal的出现是Internet技术综合利用的结果,而它本身又代表了一整套的Internet信息服务^[3]。

1.2 构建我国高校数字图书馆综合门户可行性分析

数字图书馆门户的实质是一个读者用来进行Web内容存取的信息网关。它是一个具备用户身份识别,并根据用户特征提供简单、直觉、个性化以及用户定制的Web服务系统。通过门户,用户可以根据自己的喜好和兴趣方便地存取数字图书馆门户中的数字资源内容,使用数字图书馆的服务^[4]。构建一个由资源层、接口层、管理层、服务层和用户接入层的我国高校数字图书馆的综合型垂直门户是具有一定的实际意义和可行性的:

① 资源优势:本文构建的综合型数字图书馆门户是以CALIS、CASHL、CADAL等作为基础,而且已经基本上形成了高校CCC模式,经过一段时间的发展和推广,全国高校图书馆的资源和服务共享已经达到了一定的水平,为我国高校数字图书馆综合门户的建设提供了丰富的资源,为高校数字图书馆综合门户的建设打下了坚实的基础。

② 人才和技术的优势:门户的建设和维护需要大量的专业人才和技术支撑,如自动搜索、分类技术、信息存储和挖掘技术等。高校具备大量的专业技术人才,特别是图书馆员,他们都是文化水平较高、信息素质较高的业务管理人才,不但具有扎实的专业知识,还具有熟练的信息收集、整理和加工的能力,有着先天的技术优势和丰富的实践经验。

③ 资金和设备优势:在门户网站建设过程中,不论是人员的培训及数据库的引进,还是新技术的应用及机器的购置与日常的维护,都离不开经费的保障。高校的资金一般比较雄厚,图书馆的投资在高校的财政预算中占有很高的比例,这为高校图书馆门户网站的建设提供了物质保证。另外,高校图书馆在计算机等设备的购置方面也具有一定的优势。同时,许多高校图书馆在场地、设备等方面仍有很大的潜力^[5]。

2 高校数字图书馆综合门户构建

所谓数字图书馆门户即充分利用网络环境,打破传统的信息服务模式,在Internet上建立具有本馆特色的网站、网页,这也是目前高校数字图书馆建设面临重要任务之一。高校数字图书馆门户建设能将信息服务提供上网,一改信息服务的被动形式,通过计算机网络主动将信息数据服务面向全校、面向地区、面向全国、面向全世界。数字化资源以及门户服务器后台所有的应用或服务,可根据相应的规则向该用户提供类似“一站式”的服务。基于此点,本文构建了一个由资源层、接口层、管理层、服务层和用户接入层组成的我国高校数字图书馆垂直综合型门户^[6](参见图1)。

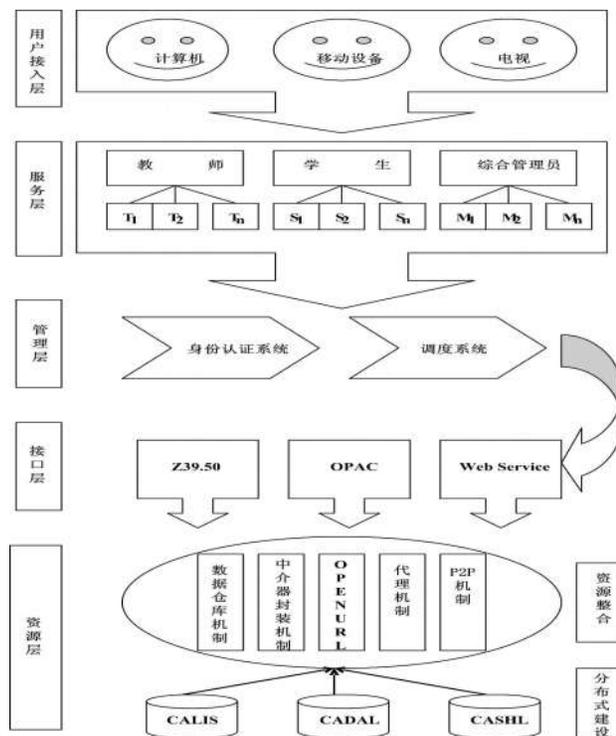


图1 我国高校数字图书馆垂直综合型门户结构图

3 高校数字图书馆门户层次分析

Web2.0技术使得网络由过去的中心化向分散化转变,用户可以在网上获得更多传播、分享和交流的自由。加之开放存取、openURL、网络技术、和数字出版的出现和普及,当前人们获取信息的方式和渠道更加的多样化,以集中快捷的方式为用户提供个性化的服务,也就是说建立综合门户是摆在图书馆人面前的紧迫任务。门户的理论和应用起源于企业领域,最初目的是为了解决“信息孤岛”和内外网

互动问题,将其引入到我国高校数字图书馆综合门户的建设当中,有利于缩小不同高校数字图书馆资源数量和服务水平之间的差距,避免我国高校数字图书馆中出现资源利用不用和信息贫瘠等问题。

3.1 资源层

3.1.1 资源分布式建设机制

数字图书馆采用分布式建设是有其必然性的。一方面,随着互联网的发展,信息量激增;另一方面,每个图书馆每年资源建设资金又是有限的,因此任何一个图书馆都无法保证购买全部的信息资源,分布式资源建设机制解决了这一问题。所谓的分布式资源建设机制就是指每个数字图书馆个体在根据自身用户需求和馆藏特色进行资源建设之后,提供各个图书馆之间的开放存取,将各自的资源进行共享。其中涉及OA、用户参与机制、原生网络数字信息、“众包”等概念。

①开放存取机制。OA即开放存取,是不同于传统学术传播的一种全新机制,其核心特征是在尊重作者权益的前提下,利用互联网为用户免费提供学术信息和研究成果的全文服务。近年来开放存取发展迅速,原因如下:首先,开放存取是基于互联网的学术传播机制。其次,开放存取是免费提供全文的信息服务方式。在开放存取模式下,科研人员不需要通过付费(包括个人订阅或者团体订阅)就能访问学术信息的全文。再次,开放存取充分尊重作者的权益,并不违背知识产权的精神。目前OA逐渐受到业界的认可,其主要实现途径有两种,即自我存档和开放存取期刊,即OA仓储和OA期刊,两者最大的区别在于OA期刊是经过同行评审的^[7-8]。

②原生网络数字信息、“众包”理念、用户参与机制。网络原生数字信息是指人们在网络交互工程中以数字形态输出的思维成果,在形成之初并无其他如纸质、胶片、磁带等物理形态的替代,网络是其“面世”的唯一载体,是信息聚合的主体。“众包”(crowd sourcing)则是由美国《连线》杂志的记者杰夫·豪提出的。“众包”任务通常由个人来承担,但如果涉及到需要多人协作完成的任务,也有可能以依靠开源的个体生产的形式出现。在“三网融合”背景下的信息资源建设,用户可以通过数字电视、移动设备和互联网来利用资源,同样也可以通过参与到资源的建设当中,即以“用户创造内容”为代表的创新民主化正在成为一种趋势即“众包”。数字图书馆建设中可以应用应用这一理念,充分挖掘用户的价值,发

挥用户参与机制进行原生网络数字信息必将成为数字图书馆分布式资源建设的主流方式。

3.1.2 资源整合机制

从国内外的研究现状来看,“整合”可以分为面向资源的整合、面向异构数据的整合和面向应用的整合三种。其中面向资源的整合是指各个高校馆根据自身的馆藏特色进行分布式建设,资源类型多种多样,需要一种机制将大量的异构系统及庞杂的资源重组和整合起来,形成一个能统一提供和便于使用的整合环境,主要有以下五种实现机制^[9]。

①数据仓库机制。数据仓库整合机制是一种物理集成方式,它将不同来源的数字资源按特定的方式建模并存储在单一物理位置,提供给用户一个新的、统一的目标数据模式,使得用户能够一站式地访问各种数字资源,从而达到整合的目的。数据仓库整合机制主要涉及的技术有海量存储技术、ETL技术、信息源的监控与更新。目前应用数据仓库机制的系统中比较有代表性的有法国和德国联合开发的Xyleme、美国斯坦福大学的WHIPS、美国佐治亚大学的InfoHarness等。

②中介器封装整合机制。中介器封装整合机制是一种虚拟整合方式,在基于这种机制的整合系统中,并不真正存储需要整合的数字资源,而是通过中介器和封装器来实现。它们均为软件组合,位于用户和数据源之间,中介器负责处理用户提问和查询结果的整合,封装器负责对信息源的链接和具体查询。在基于这种整合机制的系统中,用户按照全局模式进行查询,中介器接受用户查询并将之转换成中间格式,然后提交给相应的封装器,封装器进一步将中间格式的查询转化为信息源模式或本地模式的查询,并与参加整合的相应信息源进行连接,实现对信息源的查询,将查询结果返回到中介器进行处理,以统一的形式提供给用户。主要是通过GAV、LAV、GLAV、BAV4四种方式实现全局模式与信息源模式或本地模式之间的映射。

③OpenURL机制。开放链接(OpenURL)是一种解决不同的数字资源系统互操作、进行资源整合的方法,也是一项的技术标准。开放链接(OpenURL)是一种附带有元数据信息和资源地址信息的“可运行”的URL。可用来解决二次文献数据库到原文服务的动态链接问题,服务提供方(通常是图书馆)维护的链接解析器能够在相关服务网页上动态生成开放链接,是数字图书馆资源建设的主要实现方式。

④代理机制。代理机制使用了三种基本的 Agent,即用户 Agent、资源 Agent、和代理方 Agent。用户 Agent 负责维护用户信息,并提供系统接口,以方便用户与整合系统之间的交互。资源 Agent 负责对分布式资源进行处理,将数字资源按照整合系统的表示形式进行描述和转换,代理方 Agent 负责将从用户 Agent 发出的查询请求与所要查询的资源 Agent 进行匹配。主要涉及的技术有通信技术、协调技术等,目前使用代理机制的比较有代表性的整合系统有美国德克萨斯州奥斯汀微电子和计算机技术公司开发的 InfoSleuth、美国德克萨斯技术大学开发的 AgentRAIDER、意大利摩德纳大学开发的 MIKS 系统等。

⑤P2P 机制。P2P 即(Peer-to-Peer),是近几年兴起的一种新的计算机模式,它能够使 PC 和其他非服务器计算是实现以对等的方式联网,彼此共享对方的资源,支持互联主机之间的动态变化。主要实现原理是,存在多个分布式的对等 Peer,每一个 Peer 都拥有一套自己的数据模式,在整合的过程中,通过 Peer 与本地资源模式的映射,实现对本地资源的访问,同时依靠 P2P 映射来完成 Peer 之间的模式转换,实现 Peer 之间的通信。通过 P2P 方式,任何一个 Peer 的查询均可以在其他 Peer 之间执行,达到充分访问分布式信息资源的目的。P2P 整合机制不仅能够实现大规模数字资源的整合,而且可以实现 Web 资源的动态整合,是整合系统具有强大的扩展性,是一种比较有生命力的整合机制。

3.2 接口层

所谓接口层主要用于链接客户机和服务器。即用户通过哪一种操作协议获得数字图书馆资源服务,目前常用的主要有 Z39.50、OPAC 和 Web Service 三种。

Z39.50 是严格基于 ISO 的 OSI(开放系统互联)参考模型的应用层协议,是一个美国国家标准。Z39.50 的目的是为了信息系统的开放互联,由于各信息系统分别采用各自的数据库软件,数据的描述格式、访问方式等都各不相同,必须为各自数据库系统建立一个抽象、通用的用户视图,将各个系统的具体实现映射到抽象模型上,才能使不同的系统在一个相互理解的、标准的通信平台上进行交互,满足互操作的需要。

OPAC(全称 Online Public Access Catalogue),在图书馆学上被称作“联机公共目录查询系

统”。读者可以万维网实现图书的查找和借阅。原来的电子书目用 TELNET 作为技术支持,但是已不能满足现今图书馆联盟间的书目进行汇总。现在的网上书目多以地区性图书馆的书目加以汇总,能使读者的查询结果覆盖更大范围的图书馆。根据图书的特性,在网上书目的查找也有着不同的方式。

Web Service 是一种新的 web 应用程序分支,是自包含、自描述、模块化的应用,可以发布、定位、通过 Web 调用。Web Service 可以执行从简单的请求到复杂商务处理的任何功能。一旦部署以后,其他 Web Service 应用程序可以发现并调用它部署的服务。Web Service 是一种应用程序,它可以使用标准的互联网协议,像超文本传输协议(HTTP)和 XML,将功能纲领性地体现在互联网和企业内部网上。可将 Web 服务视作 Web 上的组件编程。

3.3 管理层

身份认证是用户获得服务的关键,针对本文提出的综合门户,笔者提出将 Liferay 技术引入到综合门户建设中。Liferay 是一个开源门户,代表了完整的 J2EE 应用,使用了 Web、EJB 以及 JMS 等技术,特别是其前台界面部分使用 Struts 框架技术,基于 XML 的 portlet 配置文件可以自由地动态扩展,使用了 Web Services 来支持一些远程信息的获取,使用 Apache Lucene 实现全文检索功能。其主要特点包括:提供单一登陆接口,多认证模式(LDAP 或 SQL)。管理员能通过用户界面轻松管理用户,组,角色,用户能可以根据需要定制个性化的 Portlet。

调度系统是数字图书馆门户的重要组成部分。数字图书馆由各种成分构成,它们包括人员、电脑、网络、对象库、查询系统、web 服务器、数字对象、对象的元数据等。为了识别这些成分,系统使用一种标识方法,即调度码和调度系统。调度系统主要解决图书馆文献服务中上下文敏感帮助和适当复本问题,通过为馆藏电子文献加上 OpenURL 链接并根据调度知识库自动向读者提供上下文相关的资源获取服务。系统分为中心子站点式和中心开户式,是基于 OpenURL 标准的多级调度系统,能够自动更新调度知识库,允许自定义本地调度规则并定制个性化调度服务,向第三方提供资源注册标准和接口,方便数据商批量更新资源注册信息。

3.4 服务层

高校数字图书馆门户(Portal)是由一个个的

Portlet组成的,这也个性化服务的实现提供了可能。由于在管理层对用户的管理采用了Liferay技术,即采用用户—用户组—角色—Portlet的方式,即将高校数字图书馆的用户分为教师、学生和综合门户管理员三个一级用户组,教师组和学生组再根据职称、专业院系的不同分为二级用户组,二级用户组下面的个体用户,即角色,提供不同的Portlet,角色通过不同的Portlet来获得个性化的服务。具体的实施方法是对不同的角色赋予不同的ID号,根据ID号的不同取得各自的访问权限,获得跟自己专业和兴趣相关的个性化的资源和服务^[10]。

3.5 用户接入层

接入层通常指网络中直接面向用户连接或访问的部分。接入层目的是允许终端用户连接到网络,因此接入层交换机具有低成本和高端口密度特性。接入交换机是最常见的交换机,它直接与外网联系,使用最广泛,尤其是在一般办公室、小型机房和业务受理较为集中的业务部门、多媒体制作中心、网站管理中心等部门。接入层是最终用户(教师、学生)与网络的接口,它应该提供即插即用的特性,同时应该非常易于使用和维护。接入层由无线网卡、AP和L2Switch组成,按照宽带网络的定义,接入层的主要功能是完成用户流量的接入和隔离。随着信息技术和电子产品的不断发展,人们获得服务的方式越来越多样化、便捷化,电脑、自主服务机等成为了人们钟爱的服务获得工具^[11]。特别是在“三网融合”的背景之下,移动设备、IP电视等也成为了可供人们选择的工具。

4 结 语

提倡建设世界一流大学首先要建设一流的高校

图书馆,高校数字图书馆综合门户作为我国高校数字图书馆建设的重要组成部分,对于服务高校教学和科研起到决定性作用^[12]。特别是CALIS、CASHL和CADAL经过近些年的发展,使我国高校数字图书馆资源共享化已达到一定的水平,然而由于经费、人才和技术等因素的限制,高校数字图书馆需要一个综合的门户,以高校CCC资源共享为基础,以综合门户为平台,为用户提供一站式的服务^[13]。

参考文献

- 肖小勃,乔亚铭.数字图书馆门户建设[J].大学图书馆学报,2004,(1):23-28.
- 张宏伟,张振海.CNKI网络资源共享平台[J].现代图书情报技术,2005,(4):6-10.
- 张 岚.信息门户与数字图书馆门户建设[J].江西图书馆学刊,2006,36(1):87-89.
- 章 静,郭吉安.2.0时代高校数字图书馆知识门户构建研究[J].图书馆学刊,2010,(2):105-107.
- 李 武.开放存取出版的两种主要实现途径[J].大学图书馆学报,2005,(4):58-63.
- 黄 颖.中国开放存取研究现状分析[J].图书情报工作,2008,(6):1-5.
- 李广建.图书馆信息系统:技术、实现与应用[M].北京:国家图书馆出版社,2010:39-48.
- 中文Liferay[EB/OL].<http://liferay.cn/web/guest/home>,2011-04-28.
- 李 静.基于Liferay的个性化数字图书馆门户的研究与实践[J].图书馆理论与实践,2009,(8):48-50.
- 秦嘉杭,许 鑫,苏新宁.数字图书馆门户的设计和构建[J].情报学报,2006,25(10):546-554.
- 王文清.构建高校数字图书馆综合服务门户[EB/OL].[Http://www.chinalibs.net](http://www.chinalibs.net),2011-4-30.
- 付婷波.高效数字图书馆门户建设研究[J].图书馆建设,2006,(3):49-51.
- 张晓青.高校数字图书馆门户网站构建探析[J].情报理论与实践,2005,28(5):617-619.

(责任编辑:孙晓明)