

●肖希明 黄连庆

以需求为导向的数字信息资源开发*

摘 要 数字信息资源开发要以用户需求为导向。网络环境下数字信息用户需求,在获取方式、内容、检索和成本方面都呈现出许多新特点。针对这些特点,数字信息资源开发应在数据库建设、资源整合、信息集成服务,以及与大众搜索引擎合作等方面体现以用户需求为导向。联盟合作和知识网络应成为以需求为导向的数字信息服务平台。参考文献 29。

关键词 数字资源开发 资源建设 用户需求 需求导向

分类号 G253

ABSTRACT Digital information resource development should be oriented by user demand. In the network environment, digital information user requirements have many new characteristics in acquisition methods, contents, retrieval and costs. Based on these characteristics, digital information resource development should reflect user demand orientation in database development, resource integration, integrated information services and cooperation with public search engines. Consortium cooperation and knowledge grids should be demand-oriented digital information service platforms. 29 refs.

KEY WORDS Digital resource development. Resource development. User demand. Demand orientation.

CLASS NUMBER G253

在信息资源建设走向数字化、网络化和多媒体化的今天,蓦然回首,我们却发现,面对以指数增长的、海量的数字信息,人们急切需要的信息通常被实际不需要的信息垃圾淹没,陷入奈斯比特所说的“信息淹没,知识荒饥”之中^[1]。人们的信息需求并没有因为信息的增长而同步得到满足,因而,寻求以需求为导向的数字信息资源开发的模式,成为我们面临的重要课题。

1 数字信息资源开发要以需求为导向

1.1 “需求导向”的界定

需求导向是相对于产品导向和市场导向的一种战略定位导向^[2]。需求导向强调市场调研,分析顾客需求,并根据未来的需要做出分析和判断,为顾客创造和提供满足其需求的产品或服务。注重需求调研和以用户为中心是需求导向策略应遵循的两个重要原则。需求调研是需求导向策略确立的前提和基础,如果缺省了需求调研,需求导向就无从谈起;以用户为中心是需求导向的立足点和出发点,它是需求导向策略不会重蹈产品导向和市场导向覆辙的根本保证。

1.2 以需求为导向是数字信息资源开发的宗旨

作为一种重要的战略性资源,信息资源开发利用的水平是一个国家的综合国力和国际竞争力的重要

标志。然而,在“Internet 信息商场(Data warehouse)”上,只有得到增值利用的那部分信息资源,才算得上是真正的战略性资源和财富。以需求为导向的数字信息资源开发,就是确立一种在用户需求调研和分析的基础上,以满足用户现实信息需求和未来发展需求为指导原则和价值取向,充分调集各种信息产品、先进技术和管理等因素,为用户创造符合其需求的信息产品和服务的发展战略。

数字信息资源开发要以需求为导向,实质是要求数字信息资源的开发利用应以所服务的机构和用户群体的需求为中心,以符合用户的信息行为习惯和心理特征为前提。它是一种以人为本、以用户为中心的资源建设策略,是强调以人为中心,而不是以信息资源或信息系统为中心的开发理念。这种人性化的理念,应贯穿于数字信息资源开发和建设的整个过程。

1.3 数字信息资源开发中存在的问题

目前,我国的数字信息资源开发主要存在两个方面的问题:一是数据库建设主要采取技术驱动的模式^[3],技术推动内容建设,而对内容开发和人文环境的重要性认识不足,造成了生产存储信息的能力大大提高,但转化为内容的能力却很低的状况;二是图书情报机构的数字信息资源建设,过多地强调数字信息资

* 本文系教育部哲学社会科学重大课题攻关项目“数字信息资源的规划、管理与利用研究”(项目批准号:05JZD00024)的成果之一。

源数量种类的完备性,对资源的集成整合利用却没有给予足够的重视,结果是数据库由几种、几十种到几百种的增加,但用户在“能够找到”的基本需求满足之后,随之面对的却是大量以孤岛形式存在的数字资源在使用中带来的困惑。数字资源整合服务能力的不足,成为了制约数字信息资源经济效率和社会效益正常发挥的瓶颈。

2 数字信息用户需求特点

数字信息资源开发所面向的用户,可以划分为不同的类型。从用户组织形式的角度,可将数字信息用户划分为机构用户(又称“行业用户”community)和个人用户(又称“最终用户”)。个人用户是数字信息资源的最大消费者,他们要求数字资源系统可以为其个人提供资料查询和获取、终身学习、科学研究、信息交流等信息支持和服务环境。IFLA在“实体——关系”模型分析基础上把用户的需求归纳为找到、辨别、选择和获取四个方面^[4]。毛军等在此基础上,将用户的需求行为归纳为发现(discovery)、定位(locate)、查询(request)、存储(storage)、交互(interactive)和发布(publishing)六种类型^[5]。

我们认为,分析信息用户的需求,还应从用户所处的社会环境对其信息行为和需求心理所产生的影响进行更加深入的分析,惟此才能挖掘用户潜在的和长远的发展需求,确保所制订的资源开发战略具有前瞻性和持久的优势。当今信息用户所处的社会环境是一个竞争激烈、创新制胜的环境,也是一个数字化生存的网络环境。这一环境必然深刻影响着信息用户需求的心理和行为。数字信息资源开发必须对这一环境有深刻的理解和准确的把握。

2003年年初,由美国密西根大学的Daniel Atkins教授牵头的一个“蓝带委员会”提交了一份题为《赛百基础结构带来科学与工程革命》的报告^[6]。OCLC也于2004年和2005年先后发布了两篇重要调查报告——“The 2003 OCLC Environmental Scan: Pattern Recognition”^[7](简称“Scan”)和“OCLC Perceptions of Libraries and Information Resources”^[8](简称“Perceptions”)。参考这几份调查报告,以及中国互联网络信息中心(CNNIC)最新发布的“中国互联网络发展状况统计报告”有关互联网用户“上网目的”和“经常使用的网络服务”等项目的调查统计数据^[9],我们可以看出,从整体来看,数字信息用户的需求具有以下特点:

2.1 方式需求:无缝链接和自助服务

“Scan”和“Perceptions”两个调查报告显示,当前新兴信息用户获取信息的三个主要趋势是:“自助服务”、“用户满意”和“无缝性”。在网络环境下,信息用户越来越多地“自行操作”,并且获得满意的结果;用户期望利用信息就像“用电或用水”一样方便,不受地域和载体的限制,并且不再像传统上那样把休闲与学习、工作截然分开,工作地点和家的界限变得模糊。

2.2 内容需求:知识化、精品化

随着网络信息资源的指数扩张,而大众搜索引擎关键词检索方法相关度太低,令人们越来越感觉到信息挑选和甄别的痛苦,信息过量已经成为一种社会公害。人们已经意识到信息精品才是互联网上最稀缺的资源,信息的深度远比广度更重要。内容需求知识化是指用户信息需求的重点逐步从简单文献获取转移到知识发现,用户的核心需要不仅是支持简单的和已知状态的文献获取,而且是支持在未知状态下的探索,支持对“不知道的东西”的探索与发现,支持知识创造^[10]。所谓内容需求精品化是指用户希望信息服务机构能从海量异构的信息存储中,按其真实需求加以集成、浓缩、重组、提炼成“高浓度”的信息“精品”,以加快其研究开发与创新的进程^[11]。

2.3 检索需求:一站式、全文化

国外的多项调查研究显示^[12-15],图书馆员和科研人员等图书馆用户在检索资源的选择上已经出现了分流,用户越来越倾向于使用搜索引擎作为自己的检索起点;习惯了通过搜索引擎进行一站式检索的潜在用户,已经没有耐心一家家地走访图书馆网站。用户希望以单一入口访问所有信息,并可实现与网络资源间的互动服务(service on demand)和个性用户对所需资源检索的一步到位(one click is enough)。调查还表明^[16],信息用户的典型期望是:什么都是全文的,并且可下载、可打印。

2.4 成本需求:最小化

“齐夫(Zipf)最小努力原则”和“穆尔斯(Mooers)定律”告诉我们:人们力图把他们可能付出的平均工作消耗最小化;如果用户取得信息比不取得信息更麻烦和伤脑筋的话,他将放弃对这一信息的需求。这两条定律说明“省时省力”是用户选择情报系统的重要标准,用户希望用最短的时间获取最有用的信息。随着社会竞争日趋激烈,求新求快成为了人们的普遍心理。当前互联网上预印本服务(arXiv)和“BT”(网上点对点传播)盛行,究其原因这些服务迎合了人们求新求快的心理需求。

综上所述,在数字化、网络化的环境下,信息用户的需求、获取知识的方式与习惯正在发生深刻的变化。数字信息资源建设与开发必须适应信息环境的转变,顺应用户的需求,以满足用户需求为导向探索全新的发展路径,在数字信息资源开发方式以及服务平台建设等方面进行创新性的研究和实践。

3 以需求为导向的数字信息资源开发策略

1975年,美国学者 Richard Saul Wurman 首次提出了信息构建理论(Information Architecture, IA)。该理论自问世以来,对电子信息的资源建设和开发利用产生了深远的影响。其核心观点是:信息构建应“化复杂为明晰”(making the complicated clear)和“使信息可理解”(making information understandable)^[17]。IA的实质是使信息可视和可理解,帮助人们更加成功地找到和管理信息,通过构建信息路径,使资源的利用变得简单与方便,最大限度减少用户查询和获取资源的时间与难度,减少和消除用户的“信息焦虑”(Information Anxiety)。信息构建理论反映了当前社会大众对信息资源开发利用的呼声和要求,因此,图书情报机构应以“化复杂为明晰”和“使信息可理解”为目标,使数字信息资源开发方式做到贴近用户的需求。具体可从以下方面入手:

3.1 数据库建设

特色数据库建设是图书馆特色馆藏数字化的主要形式。由于缺乏宏观规划与指导,目前各馆的特色数据库建设基本上是采取单库模式,强调内容特色,不考虑数字资源的可广泛使用性、可互操作性和可持久使用性,形成了一个又一个无法互通互联的“信息孤岛”。这种单库模式无法满足网络用户一站式、无缝检索所有相关资源的愿望。以需求为导向的数据库建设,首先要强化信息产品的标准化建设,形成一整套的关于数字资源建设、资源描述和资源检索利用的核心标准规范,保证数字信息资源建设能够按照统一的标准规范进行,从而支持在整个数字信息环境中的有效使用、广泛获取和长期保存信息,为特色化数字信息资源在地区乃至全国的共知共享奠定坚实的技术基础。数字信息资源系统的标准规范建设作为互通互联的数字图书馆建设的重要方面,在西方发达国家已经过了多年的研究并取得了瞩目的成就,近年来在我国也在此领域进行了积极的探索^[18-19]。

3.2 信息资源整合

数字资源整合(Digital Resource Integration),将以

数据库资源为中心内容的各种数字化文献资源通过中间技术处理,整合一体,统一在一个平台上,提供一站式服务,以提高资源利用效率。其实质是为了解决多平台、多结构信息资源的集成检索问题。从整合的范围、规模看,数字资源整合既有本单位范围内的数据库的资源整合,也有跨单位、跨区域、全国性乃至跨国界的数字资源整合,如 Dialog、OCLC 的联机数据库,参与机构达几百家之多,我们可以通过统一界面获取题录信息,又迅速通过各种传递方式获取到全文。数字资源整合是实现资源整合的关键技术,通常采用两种方式^[20]:依赖元数据收集的元数据仓储(Metadata Repository)/聚合型(又称“元数据方式”)、依赖计算机能力和标准协议的元搜索(Metasearch)/技术型(又称“元搜索方式”),两种方式各有优势和劣势。开放链接技术是当前数字资源整合系统性能改进的技术热点,它通过 OpenURL(开放环境信息传递的规范化语法)解决二次文献与一次文献之间的动态链接,应用前景不可估量。此外,基于个性化空间与信息定制的数字资源整合^[21]也是未来技术探索的热点之一。

3.3 信息集成服务

信息集成服务(Information Integration Service)目前尚无统一定义。有学者认为:“信息集成服务是针对某一特定领域或特定组织机构的需求,以信息为对象,信息资源为本体,服务为动力,网络技术为手段,协同作业为方法,把信息资源诸要素有机地链接成一个整体的动态过程。”^[22]可以看出,信息集成服务是建立在网络化、数字化的虚拟环境下,以满足用户需求为目的,通过对原有资源、系统、流程和人员等各方面的集成,为用户构建一个全方位、多层次、无缝链接和个性化信息利用环境的一种新型信息资源开发方式。学科信息门户(Subject Gateway)是学术信息集成服务的常用模式。但目前的学科信息门户只是在统一访问界面、资源导航、统一检索、统一认证等方面做得较好,而内容聚合、个性化信息定制、支持 Portlets(通过 Portlets 将不同资源与应用中的各类信息集中在同一页面上呈现给出用户)以及协作能力等核心功能则未得到充分展示^[23]。近几年在国外大学图书馆兴起的 IC(Information Commons,即信息共享中心)为深化学科信息门户的核心服务功能提供了有益的启示。IC 通过整合图书馆技术部、参考咨询部、多媒体部和数据服务部功能,将用户的需求和期望的变化纳入到一个整体中,为学生和教职员创造一种全力支持其整个学习和科研过程的环境^[24]。IC 服务模式的突出特点是,信息服

务台处于IC的功能中心,它是用户接触图书馆数字资源的第一道门槛,IC各个分部门的很多工作都要围绕它接收的信息资源问题而开展^[25]。

3.4 与大众搜索引擎合作

为了适应用户自助服务和对搜索引擎依赖程度日益提高等信息行为的变化,2004年初,OCLC率先推出了OPEN WorldCat计划,将其拳头产品WorldCat的馆藏数据陆续加入Google及Yahoo两大搜索引擎中,创建图书馆频道(library channel),使用户可以方便快捷地检索全球图书馆的资源。2005年底,中国国家图书馆也宣布了与Google Scholar的合作计划。这些通过信息推送等方式将图书馆信息资源主动纳入到用户情景中的有益尝试,提高了图书馆信息资源在用户心目中的可见度和易用性,可视为信息资源开发利用的方式方法之一。由于版权等限制,目前用户通过搜索引擎查找到的图书馆电子资源只能是二次文献及其馆藏信息,或者是无版权意义的电子版全文。当前图书馆与搜索引擎联合开展的资源开发,还属于浅层次的开发,我们期望今后通过妥善解决版权问题、采用策略联盟等途径^[26],拓展资源合作开发的深度与广度,提高经济效率和社会效益。

4 以需求为导向的数字信息服务平台建设

数字信息服务平台作为图书情报机构网络化服务的信息基础设施,它提供了一个通用和高效的开发环境和运行环境,允许信息机构根据需要开发和部署各种数字资源的发布、门户、检索、计费、点播、参考、个性化信息定制等服务,是数字图书馆功能框架的重要组成部分^[27]。为满足数字用户多元化、一站式、全方位的信息需求,以及随着新技术、新标准的出现,以需求为导向的数字信息服务平台建设思路发生了巨大的转变,由以单个数字图书馆为中心转向区域型、行业型、乃至全国型数字图书馆的标准化、互通互联的发展模式。

4.1 联盟合作型

联盟合作是图书馆文献资源共建共享的重要形式。以往,由于缺乏统一的标准规范等多种原因,联盟内的各个数字信息服务平台之间实质上是无法达到真正的互联和共享的。随着数字图书馆标准化工作的推进,以及开放应用技术的发展,数字信息服务平台的可广泛使用性、可互操作性和可持续性逐步得到了实现。例如CALIS“十五”项目“中国高等教育数字图书馆(CADLIS)”按照“标准化、开放性、协作性、互操作”

等原则,陆续推出新型门户系统、统一用户认证系统、统一检索系统、OpenURL资源调度系统、馆际互借与文献传递系统、分布式联合虚拟参考咨询系统、重点学科网络资源采集与导航系统、联机编目系统、元数据联合著录系统等十几种数字图书馆软件,使得高校图书馆不仅可以采用新技术和新标准来完善或重新构建自己的数字资源服务体系,同时,还能够将本馆的数字信息服务体系与省级、地区级乃至全国级的高校数字服务体系真正进行集成,实现全省、全地区、全国范围内的资源与服务的共建共享和自动化协同服务。

4.2 知识网格型

知识网格是一个智能互联的环境,能使用户有效地获取、发布、共享和管理知识资源,并为用户和其他服务提供所需的知识服务,辅助实现知识创新、协同工作、问题解决和决策支持^[28]。长期以来,知识的控制单位基本上停留在文献级别上,制约了信息资源的有效利用。知识网格则致力于利用网格、数据挖掘、推理等技术从大量在线数据集中抽取和合成知识,使搜索引擎能够智能地进行推理和回答问题,并从大量数据中得出结论。知识网格将在知识获取与知识表示、知识可视化、知识关联和集成、在动态虚拟组织间进行有效的知识传播和知识管理等多个方面取得突破,从而构建一个基于“知识元”水平上全文关联的、数据结构标准化的、可以实时以多媒体信息交互的知识服务平台,并通过标准化将各种动态更新的专业知识仓库关联起来,无论数据库的物理位置何在,无论数据库专业性质和具体结构如何,都可以使其在知识信息的内容上跨越数据库界限而浑然成为一体。欧盟于2004年启动的DILIGENT原型项目向世人表明,网格型数字信息服务平台的目标是结合e-Science、e-Learning等环境,让用户根据自身研究课题、学习的需要,允许从用户自己的计算机出发,使用户可以使用自我定义的标准创建个人虚拟数字图书馆,把所有的信息动态地集成到一起,并最终成为以用户信息活动为基础的数字信息环境^[29]。

数字信息服务平台开发从联盟合作型到知识网格型,是一个具有里程碑意义的跨越,我国图书情报界在这方面的工作还处于研究与试验的初级阶段。现阶段的首要任务,应该是加强和完善数字信息资源开发的技术标准规范建设,夯实信息资源共享的技术基础,使各种应用系统和异构的资源与服务具备统一的接口,增强对网格服务的支持。

(下转第87页)

论文,涉及其个人研究领域、研究经历及项目进展等,通常还友情链接相关站点,利用价值也较高。如 Open Access News、American Scientist Open Access Forum、Open Access Now Forum、SPARC Open Access Forum 等网站为用户提供了大量有关开放存取的信息。

参考文献

- 1 Directory of open access journals. [2006-11-08]. <http://www.doaj.org/articles/about>.
- 2 Module 31 of the CONSER Cataloging Manual (Library of Congress). [2006-11-15]. <http://www.loc.gov/acq/cons->

[er/Module31.pdf](#).

- 3 黄水文.基于数字图书馆门户的资源建设.图书情报工作,2004(4)
- 4 中国互联网络发展状况统计报告. [2006-11-14]. <http://www.cnnic.net.cn/uploadfiles/pdf/2006/7/19/103651.pdf>

陈传夫 武汉大学信息资源研究中心教授,信息管理学院院长。通讯地址:武汉。邮编 430072。

王云娣 浙江师范大学文献信息研究中心研究馆员。通讯地址:浙江金华。邮编 321004。

(来稿时间:2006-12-26)

(上接第 68 页)

参考文献

- 1 Knowledge Lost on Information. Report of the NSF Workshop on Research Directions for Digital Libraries, June 15-17, Chatham, MA. [2007-3-15]. <http://www.sis.pitt.edu/~dlwshop/JISC/NSFreport.pdf>
- 2 项保华.基于顾客价值创新的企业战略定位研究.浙江大学博士学位论文,2001
- 3 曾民族.沟通无限——中文信息资源深度开发与利用研讨会专家观点集锦.数字图书馆论坛,2005(9)
- 4 刘素清. IFLA 书目记录功能需求 (FRBR) 初探. 大学图书馆学报,2004(6)
- 5 CSDL 中国科学院国家科学数字图书馆项目管理中心. 中国国家数字图书馆工程需求调查研究报告. [2006-12-20]. <http://www.maojun.com/doc/cdl-user-req.pdf>
- 6,12 Revolutionizing Science and Engineering Through Cyberinfrastructure; Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure. http://www.communitytechnology.org/nsf_ci_report/report.pdf [2006-4-10]
- 7,13 The 2003 OCLC Environmental Scan: Pattern Recognition. [2006-7-14]. http://www.thealbertainlibrary.ab.ca/netspeed/netspeed2004/Keynote_Descriptions2004.pdf
- 8,14 College Students' Perceptions of Libraries and Information Resources——A Report to the OCLC Membership. [2006-7-14]. <http://www.oclc.org/asiapacific/zhen/reports/pd/studentperceptions.Pdf>
- 9 中国互联网络信息中心. 第 19 次中国互联网络发展状况统计报告. [2007-3-20]. <http://www.cnnic.net.cn/html/Dir/2007/01/22/4395.htm>
- 10 张晓林.构建数字化知识化的信息服务模式.津图学刊,2003(6)
- 11 舒明全.基于用户个性化需求的精品化信息服务模式.情报科学,2004,22(6)
- 15 Steve Carroll. Googled Science. Library Connect Newsletter,2004,2(2):pp5
- 16 Lesley M. Moyo. Electronic Libraries and the emergence of

new service paradigms. The Electronic Libraries,2004(3)

- 17 刘青华.信息构建(IA)与图书馆数字化服务浅谈.数字图书馆论坛,2006(4)
- 18 张晓林等.数字图书馆建设的标准与规范.中国图书馆学报,2002(6)
- 19 CALIS 管理中心编制.中国高等教育数字图书馆技术标准与规范.2004,10
- 20 姜爱蓉.数字资源整合系统的技术发展趋势.图书馆杂志,2006,25(12)
- 21 胡昌平,汪会玲.信息资源的异构整合及其对平台建设的要求.中国图书馆学报,2006(3)
- 22 赵英.图书馆信息集成服务初探.晋图学刊,2002(3)
- 23 王文清.构建高校数字图书馆综合服务门户.数字图书馆论坛,2006(2)
- 24 Russell B, Barbara T. Information commons redux: concept, evolution, and transcending the tragedy of the commons. Journal of Academic Librarianship, 2002, 28(5): 227-286
- 25 李平. Information Commons: 一种新的研究型图书馆服务模式.图书情报工作,2005(4)
- 26 韩红.数字资源建设与服务的信息网格——“数字资源策略联盟”的构建.数字图书馆论坛,2005(10)
- 27 黄如花.数字图书馆原理与技术.武汉:武汉大学出版社,2005
- 28 诸葛海.利用知识网格实现知识创新.中国信息化,2004(3)
- 29 黄国彬.从 DILIGENT 建设看欧洲数字图书馆发展.数字图书馆论坛,2006(6)

肖希明 武汉大学信息管理学院图书馆学系主任,教授,博士生导师。通讯地址:武汉珞珈山。邮编 430072。

黄连庆 佛山科技学院图书馆副研究馆员,武汉大学信息管理学院硕士研究生。通讯地址:广东省佛山市。邮编 528000。

(来稿时间:2007-04-16)