

●索传军 张怀涛

网络环境下虚拟馆藏的建设*

摘要 虚拟馆藏与图书馆传统的现实文献资源一起构成图书馆信息资料的总和。网上信息资源极为丰富且数据类型繁多,对图书馆来说实际能够开发利用的却十分有限。目前基本的反映网上资源的方式是主页。网上虚拟馆藏的开发就是有目的地组织一些信息资源,帮助用户搜寻、发现一些专业学科的有用资源。参考文献 9。

关键词 网络环境 虚拟馆藏 组织模式

分类号 G253

ABSTRACT Library information resources consist of virtual collections and traditional real document resources. Although there is quite a lot of information on the Internet, there is little easy to be used for library services. To develop virtual collections is to organize information resources and provide them for users. 9 refs.

KEY WORDS Network environment Virtual collection Organizational pattern

CLASS NUMBER G253

1 虚拟馆藏概念与特点

虚拟馆藏(virtual holdings)源于虚拟图书馆(virtual library)。所谓虚拟图书馆,其核心就是利用虚拟馆藏开展信息服务。从以下一些虚拟图书馆的观点就可以分析出虚拟馆藏的含义与特点。

(1) 虚拟图书馆就是图书馆的服务不局限于物理意义上的馆藏,而是通过通信网络联结各地区、全国甚至全球信息资源的逻辑意义上的馆藏,用户可以在其中检索比本馆馆藏多得多的信息^[1]。

(2) 虚拟图书馆不一定基于一个实际的图书馆,它可以是存储在由网络连在一起的许多图书馆中的资源和服务的混合物^[2]。

(3) 虚拟图书馆能够使用远程通信网络将大量的全文和图像等数据从图书馆传输到远程用户^[3]。

(4) 虚拟图书馆是一种知识管理实体,它将传统图书馆范畴与远程通信和计算机技术的应用有效地结合起来,通过将图书馆自身拥有的资源、图书馆母体机构专有信息、外部的世界范围内的信息

资源无缝地整合,促进每个用户快速地存取和有效地使用信息^[4]。

(5) 虚拟图书馆就是没有围墙的图书馆。它以电子方式将世界范围内的图书馆、个人、机构及商业公司连接起来,并提供检索其存储的学术信息资源^[5]。

从以上五种观点可以看出,对虚拟图书馆概念的描述,主要是对馆藏信息资源和信息服务的描述。其共同特点就是借助计算机系统、通信网络可以利用本馆以外的电子信息资源,也就是我们要论述的虚拟馆藏。因而可以将虚拟馆藏定义为:虚拟馆藏就是本馆读者借助于计算机系统、通信网络等可以广泛利用(检索、传输等)的本馆以外网上的电子信息资源的总和。虚拟馆藏具备下面4个特点:网络环境下图书馆的虚拟馆藏是电子信息资源。虚拟馆藏具有广泛的共享性,简单地说,一个图书馆的网上(Web 服务器)信息资源可以是多个图书馆的虚拟馆藏。虚拟馆藏具有动态随机性。就是说,网上虚拟馆藏随时都在更新,在发生变化。另外,虚拟馆藏大小与你每次的利用能力和程度有密切的关系。虚拟馆藏对计算机、通信网络等有较强的依赖性。

* 本文为基于网络的图书馆信息资源组织与服务课题组的国家社会科学规划基金资助项目。

2 虚拟馆藏与现实馆藏的关系

现实馆藏是本馆自己拥有的可以为读者开展服务的实有馆藏。它可以是多种载体形式的信息资源,如纸质文献(图书、期刊等),电子文献(电子图书、期刊、数据库)等。简单地说,它是本馆拥有的各类信息资料的总和。但在网络环境下,只有网上数字化的电子文献可以被共享,就是说现实馆藏只有在网上能够为读者开展服务的那部分电子信息资源可以作为其他馆的虚拟馆藏。目前,我国一些图书馆已经建了计算机局域网,上了图书馆计算机集成管理系统,建立了书目信息数据库,实现了书目信息在因特网上检索。但书目信息所反映的未数字化的馆藏文献资料不能够通过网络共享。这样,在网络环境下书目信息数据库既是本馆的现实馆藏,又是其他馆的虚拟馆藏。而书目信息数据库所反映的其他一些非数字化的文献信息资料,由于不能在网上共享而只是本馆的现实馆藏,不是其他馆的虚拟馆藏。

概括地说,网络环境下图书馆的现实馆藏和虚拟馆藏具有以下6种关系:

(1) 计算机网络环境下现实馆藏和虚拟馆藏是相互的。网上自有馆藏信息资源既是本馆的现实馆藏,同时又是其他馆的虚拟馆藏,这部分馆藏是指在网上可供其他馆共享的那部分数字化了的文献。

(2) 在本馆的现实馆藏中,存在未数字化的纸质等文献信息资料的一部分单机版电子文献,不能在网上共享。这样,现实馆藏>现实电子馆藏(指馆藏中已经数字化的文献信息资源)>可作为其他馆虚拟馆藏的信息资源。

(3) 本馆对现实馆藏具有更新、修改、利用和支配等权利(仅对本馆开发建设的那部分资源);对虚拟馆藏只有利用(共享)权,一般没有更新、修改和支配权。

(4) 现实馆藏是本馆馆藏;虚拟馆藏是外馆(或外单位、个人等)馆藏。

(5) 本馆的读者或用户对现实馆藏的使用具有直接性;虚拟馆藏,除网上免费信息资源外,一般需要通过某种协议或申请(设立账号、密码等)后才能使用——检索、浏览、下载等,也就是说虚拟馆藏的使用具有间接性。

(6) 虚拟馆藏是多个图书馆、机构、个人可共享的信息资源的总和。从理论上讲,在因特网上的图书馆,其虚拟馆藏大到可以是除本馆以外的整个因特网信息资源,因而,虚拟馆藏>现实馆藏。

因而,网络环境下图书馆的现实馆藏和虚拟馆藏既相互联系,又相互区别。

3 虚拟馆藏的规模与影响因素分析

3.1 虚拟馆藏影响因素分析

因特网上信息资源丰富,但地理上分散,组织上松散;数据类型多,随机变化大。目前任何人对因特网上信息资源的开发利用都是十分有限的。所以,对一个图书馆来说,理论上其虚拟馆藏十分巨大,除本馆现实馆藏外,整个因特网上的信息资源都可以作为其虚拟馆藏,但网上实际能够开发利用的信息资源却是十分有限的。这主要取决于三个方面:

一是因特网上的信息资源虽然十分丰富,但有许多信息对我们图书馆是无用的,如:与本馆收藏无关的信息;因特网上的一些无用信息,如黄色东西;还有一些难以开发利用的信息,如电子布告栏等;另有一些非常有价值但缺乏有效的组织管理,对图书馆来说就很难开发利用等。这就决定我们有效虚拟馆藏是有限的。

二是取决于图书馆的开发利用能力。从某种程度上说,这是决定我们实际虚拟馆藏大小的主要因素。如:网络环境与信息技术设备条件,如Web服务器的档次,线路的通信带宽和选用的浏览器等;人员的素质与技术水平,如计算机应用水平,外语阅读水平等。

三是本馆对虚拟馆藏的开发模式与政策。图书馆领导对虚拟馆藏的理解、重视程度,决定了对虚拟馆藏开发的投入力度的大小。虚拟馆藏的开发有多种模式,每种模式虚拟馆藏的表现方式不尽相同,开发难度也不完全相同,因而造成虚拟馆藏的规模、容量也不相同。

3.2 虚拟馆藏的规模

根据图书馆网络信息资源的开发利用情况,网络环境下图书馆的虚拟馆藏存在三种形式:理论虚拟馆藏、有效虚拟馆藏(或者潜在虚拟馆藏)、实际虚拟馆藏。这三种形式的虚拟馆藏既相互联系,密

不可分,又相互区别。它们是动态的,随因特网的发展而发展。理论虚拟馆藏如前所说可以无穷大,其极限是整个因特网信息资源。假设理论虚拟馆藏为 X ,则 $\lim X \rightarrow \infty$ 。对一个图书馆来说,有效虚拟馆藏就是因特网上其可能开发利用的信息资源的总和。由于因特网信息资源的分布、组织特点,以及本图书馆的网络设备环境、条件等的限制,对因特网上信息资源的开发利用是十分有限的。如世界著名的网络搜索引擎 Alta Vista 文档索引库有 2100 万页, Lycos 文档索引库有 1900 万页, Opentext 文档索引库有 150 万页, Infoseek Guide 文档索引库有 100 万页。然而,这些著名的网络搜索引擎所反映的也仅是整个因特网上信息资源的很小一部分。因而,有效虚拟馆藏远小于理论虚拟馆藏。假设有效虚拟馆藏为 Y ,则 $Y < X$ 。实际虚拟馆藏是一个图书馆实际开发利用的因特网上信息资源的总和。假设实际虚拟馆藏为 Z ,则 $Z < Y$ 。实际虚拟馆藏是有效虚拟馆藏中已开发利用的部分。

4 虚拟馆藏的组织模式

因特网上信息资源不仅数量大,而且种类繁多。但最基本的反映方式是主页。所谓主页,就是 Web 服务器向网络用户提供超文本或超媒体信息时最先展示的信息页^[6]。它类似一本书的封面和目录,介绍的是该网络服务器的概貌,主要提供该主页的信息类别、内容和形式等。通常我们利用 WWW 浏览器查询信息时,首先获得的是一些网络服务器的主页地址(URL)。通常主页上有一些信息,有些就是我们所需的信息,有些则不是。因而要获得具体信息,必须选择有关栏目进行查询。一般情况,我们获得一些网络服务器主页的绝对地址后,对其主页进一步查询,可以获得一些具体信息的相对地址(相对地址指明网络信息资源在服务器的相对位置)。也就是说,因特网上信息的开发实际上分为两步。第一步获得主页的绝对地址;第二步再对有关主页上的信息进行深入开发。目前对因特网上信息的开发,还主要是第一步。对第二步的开发,还需要更深入的研究。

因特网上虚拟馆藏信息资源开发,就是有目的地组织一些信息资源,为上网查询信息的读者(或用户)提供指引和帮助。具体就是图书馆员按照某种模

式和方法通过上网搜寻、挖掘,发现一些专业学科的信息资源,通过进行分类、标引、建立索引动态连接等为读者提供查询服务。现在,因特网上信息资源的开发模式主要有三种(举例):

(1) 热门站点或相关站点推荐。它们是因特网上利用最多、最简单、直接的信息开发利用模式。这种模式被大家广泛采用,如在华中理工大学、北京大学、清华大学图书馆的主页上都可以看到这类信息服务项目。在清华大学图书馆主页上有“国内网上图书馆, Internet 导航”;在华中理工大学图书馆主页上有“网上图书馆和热点网址”;在华南理工大学图书馆主页上有“其它图书馆主页地址”;在上海交通大学图书馆主页上有“热点地址”等。

(2) 因特网专业信息指南系统建设。因特网专业信息指南系统类似传统图书馆的读者导读系统,目的是指导读者有效地利用馆藏信息资源。通常导读系统介绍的是本馆的现实馆藏,而因特网专业信息指南系统介绍的是虚拟馆藏。“指南系统”较网上热门站点的推荐来说,它经过专业人员的加工、组织,信息更加系统准确,导航作用更加突出。而较专业信息资源指引库来说,其信息量要少些。但重点突出,更加准确。比如:在西安交通大学主页(<http://www.xanet.edu.cn/xjtu/newxjtu/xjtu/chn/lib.html>)上有“网上资源”;在军事医学科学院主页(<http://202.38.153.17>)上有“Internet 生物医学信息指南系统”;主要按照资源类型分为数据库、电子期刊、专利信息、学术会议、电子论坛等,按照医学主题分为生物医学、临床医学、军事医学等;在上海交通大学图书馆主页(<http://202.120.193>)上有“重点学科网上资源收集”和“电子图书馆”;其重点学科网上电子服务有自动控制、计算机与通讯、材料科学、机械工程、船舶与海洋工程和动力机械等专业学科类目。

“专业信息指南系统”的开发建设方法,是通过多种搜索引擎对某一或某些主题信息上网查询、浏览,并参考有关文献,选择参考价值较高的信息资源(URL),总结、组织、归类,设置类目。因而,专业信息指南系统的结构可以这样设计:首先按照学科设立一级类目,在一级学科类目下设立二级学科栏目,二级学科类目下可以按照信息资源类型设三级类目(如果本学科的信息资源较少,可以不再按信息资源类型设三级类目,使二级类目直接对应于具体的信

息资源, 这样更简洁明了), 三级类目对应的就是具体的信息资源的网址。

例如: 郑州工业大学是一所综合性工科大学。现有 10 个院系, 20 多个专业。根据这一情况, 郑州工业大学图书馆(<http://lib.zzut.edu.cn>)可以在自己的主页上建立这样适合本校读者的指南系统:

(一级类目) 化学工程信息资源

(二级类目) 有机化工、无机化工、化工机械等

(三级类目) 对应于因特网上的一些化学信息资源网

· Chemical Information Sources(化学信息资源)

· Listserv 邮件列表:

订阅主题: chm inf-1

订阅地址: listserv@iubym.ucs.indiana.edu

· Listserv 邮件列表:

订阅主题: cicouwe

订阅地址: listserv@iubym.ucs.indiana.edu

· Chemistry information (化学信息)

Web:

<http://acad.mtstate.edu/~chemnet/chem.html>

· Chemistry Journals(化学刊物)

Web:

<http://www.hetacyte.pair.com/journals.html>

· Chemistry Talk and General Discussion(化学论坛)

Usenet:

Sci chem

电气与信息工程信息资源

计算机科学、工业自动化

.....

土木建筑工程信息资源

土木建筑、建筑学等

.....

水利及环境工程信息资源

水工建筑、农田水利、城市规划和环境保护

机械工程信息资源

机械学、材料科学

.....

橡塑磨具信息资源等

.....

(3) 专业信息资源指引库的开发与建设。在因特网上开发建设一个图书馆的虚拟馆藏, 或者说建设具有本馆自己特色的信息资源网, 这是对以往以信息资源相对集中和规范的数据库概念的重大突破。从它的概念、功能以及系统的体系结构等理论问题, 到网上信息的搜集、加工、组织等均有一

系列的技术问题。但从现有的技术和条件看, 要较为系统、全面地开发因特网上某一或某些专业的信息资源, 关键是信息资源指引库的建设。指引库与专业信息资源指南系统相比, 专业信息资源指南系统是静态的, 而指引库是动态的。指引库开发建设的技术难度更大, 反映信息更加全面及时。

指引库是指所建立的数据库中, 从物理上讲并不存储各种实际信息资源, 但对其进行访问却可以检索到有关的信息实资源, 即指引用户到特定的地址获取所需信息。这样就可以把因特网上与某一或某些主题相关的节点进行集中, 按照方便用户(读者)检索的原则, 用用户(读者)熟悉的语言组织起来, 向用户(读者)提供这些资源的分布情况, 指引用户(读者)查找^[7]。

专业信息资源指引库是这样—一个系统: 它将国内外该专业的信息资源通过因特网等途径进行搜集, 将这些物理上分散的大量原始信息资源进行合理整理和组织, 从逻辑上联系起来, 并且通过各种导航手段, 为读者方便地定位, 迅速获取所需信息资源提供引导。信息资源指引库主要有三部分组成: 一是反映信息资源的 URL (Uniform Resource Locator 统一资源定位器) 即信息资源的网址; 二是原始信息, 包括访问频度较高的原始信息资源的镜像等; 三是一套方便信息组织和读者查询的支持技术(检索软件)。为了搜集专业信息资源的 URL, 所需的关键技术之一是因特网自动漫游技术。从国外的经验看, 人们搜集因特网信息资源的手段是利用一种“网络机器人”的软件(automated robot, Wanderer, Spider, Harvest 等)在因特网上漫游, 自动搜集网上 WWW, Gopher 和 ftp 站点上的各种相关信息(连接、文本等), 并且把这些存于本站点上, 然后进行加工、组织和提供服务^[8]。

专业信息资源指引库的建设, 我国一些单位也做了开发研究。如南京大学软件新技术国家重点实验室研制开发的“基于 WWW 的信息搜集系统 DGS (Information Discovering and Gathering System)”。DGS 是为了在 WWW 上自动进行中英文技术资料的搜集而设计开发的。DGS 能够根据用户所提供的训练文本和系统设定, 提取目标资料的特征, 并根据目标特征自动在 WWW 上进行资料搜集, 然后将所搜集到的资料经分类整理导入文档索引库, 供用户查询。DGS 系统能够保持资料库的

不断更新,能够向用户提供及时的网上技术资料服务^[9]。

5 虚拟馆藏的管理

在未来网络环境下的图书馆中,虚拟馆藏在文献保障中的作用将变得越来越重要。由于虚拟馆藏对图书馆的馆藏资源建设和信息资源开发还是新生事物,其建设和开发过程中还存在较多的未知因素和问题,如有关的虚拟馆藏建设政策、管理办法,虚拟馆藏的利用以及有关的知识产权和法律问题等等。

(1) 政策。我国图书情报界对文献资源共享的理论和实践问题进行了较为全面和深入的研究。但由于缺乏统一的政策指导,一直未能达到所期望的效果。现在由于技术的发展,使得资源共享成为可能。如果没有科学政策的指导,也难以发挥其作用,因而应尽快对虚拟馆藏的建设、管理和使用等有关问题进行研究,以指导其发展。

(2) 权利和义务。“图书馆个体”对虚拟馆藏的建设有建议权,没有支配权;但该“图书馆集体”对虚拟馆藏的建设具有共同的责任和义务。

(3) 虚拟馆藏的发现工具。对于未来网络图书馆来说,虚拟馆藏的使用工具就是网络资源发现工具。现在因特网上有许多网络信息检索和浏览工具,如 Gopher、WAIS 和 WWW 等信息检索工具。现因特网上的中文信息资源增加较快,但我们缺少网络中文信息检索和浏览工具,为了更好地利用虚拟馆藏,就必须开发我们自己的使用工具。

(4) 虚拟馆藏共享系统。虚拟馆藏是有若干个图书馆或信息服务机构组成的集体。对虚拟馆藏中有关文献、信息的分布加以科学地说明,有利于资源共享和对虚拟馆藏的使用。

参考文献

- 1 孙承鉴 电子图书馆雏形 北京: 中国电子出版社, 1996: 30~ 33
- 2 Alan Poulter Towards a virtual reality library. A slib Proceedings, 1993, 45(1): 11~ 17
- 3 Charles Oppenheim. Virtual reality and the virtual library. Information services & Use, 1993, 13: 215~ 217
- 4 Pamela Jajko. Planning the virtual library. Medical Reference Service Quarterly, 1993, 12(4): 51~ 67
- 5 Ken Beiser The virtual library. Computers in Libraries, 1992, 12(6): 26
- 6 朱鸽昀等 军事医学科学院生物医学 Web 服务器的建立 情报学报, 1997(2): 124~ 127
- 7 陈梅华 探索网络信息资源建设的关键技术 情报学报, 1997(2): 105
- 8 尹俊涛, 徐如镜 全球信息网络环境下信息资源开发利用技术初探 情报学报, 1997(2): 100~ 103
- 9 邹 涛 基于 WWW 的信息搜集系统 DGS 的实现 计算机世界产品与技术专辑, 1999-04-19(c10)

索传军 郑州大学图书馆副研究馆员。通讯地址: 郑州市。邮编 450002。

张怀涛 郑州纺织工学院图书馆副研究馆员。通讯地址: 郑州市中原西路 41 号。邮编 450007。

(来稿时间: 1999-06-17。编发者: 翟凤岐)

(上接第 47 页) 情报学博士论文与其他学科领域相互渗透是十分明显的。我国可试验设立图书馆学与其他学科的双学位、合作学位的硕士、硕士后等学位项目。

参考文献

- 1 Danny P. Wallace: ALISE Statistical Report/students, ed by Timothy W. Sineath, 1996, p. 86
- 2 钟守真 图书馆学专业核心课程的构建 海峡两岸第三届图书资讯学学术研讨会论文集(A 辑), 武汉大学图书情报学院, 1997

- 3 彭斐章 图书馆学定有灿烂的未来 图书情报工作, 1996(3)
- 4 孟广均 关于学科建设和名称设置之我见 图书馆工作, 1996(3)
- 5 胡昌平 面向 21 世纪的中国信息管理类专业教育 情报学报, 1999(1)

陈传夫 武汉大学大众传播与知识信息管理学院教授、副院长。通讯地址: 湖北武汉。邮编 430072。

(来稿时间: 1999-07-20。编发者: 李万健)