

● 陈光祚

论“图书情报学虚拟图书馆”的建设

摘 要 目前国外网上的图书馆学信息已很丰富,相比之下,国内已落后,必须加强此项工作。图书情报学虚拟图书馆的建设,收集信息范围要广,角度要从整个图书情报事业出发,收录地域应以国内为主;著录项目要比较全面,所用软件要经过科学选择。

关键词 图书情报学虚拟图书馆 图书情报学数字图书馆 图书情报学信息

分类号 G250

ABSTRACT There is an abundance of information on networks of foreign countries, while there is not so much in China. To catch up with the latest development, we should collect and fully index a wide variety of information mainly in China.

KEY WORDS Virtual library in library and information science Digital library in library and information science Information in library and information science

CLASS NUMBER G250

目前我们正在进行一项图书情报学虚拟图书馆建设的科研项目。在近一年的时间里,我们从网上搜集了近1000个有关网站网页的网址,为其中某些条目进行了标引及撰写摘要,同时正编制虚拟图书馆的软件,进行初步的调试。我们估计,将来的网址可达3000~5000条,并且将在武汉大学传播与知识信息管理学院的网站上用这个虚拟图书馆向公众提供服务。为了使研究工作逐步深入,我们就实践中遇到的一些问题,谈点认识和初步做法,以求教于同行。

1 提倡建立虚拟图书馆

虚拟图书馆是因特网上组织信息资源的一种有效而又经济的形式。具体说来,虚拟图书馆就是根据特定的目标,选定信息资源的学科领域,对有关的网站网页进行搜索和收集,加以鉴定核实,并对核实后的网址进行合理组织,使之能够提供检索、浏览和链接的信息集合。

从广义来说,就其处理电子化信息而言,虚拟图书馆是电子图书馆的一种类型。但是它与一般的数字图书馆不同,不需要对馆藏文献与数据进行数字化处理。虚拟图书馆不需自己的藏书,而是

针对网上已经存在的信息资源,给出链接指针,使用户随时选择调用。相对于基于自己实体文献的数字图书馆来说,它具有下列明显的优点:

(1) 虚拟图书馆建设的投入少、易建易行。它以网上的已有资源为基础,类似于为其编制主题索引。它的投入仅为目标的确定,网上相关资源的查寻与核实,为有关网址进行特征描述(赋予标引词和摘要简介),编制或采用一定的软件以供信息存贮与检索,并实现对网站网页的链接。虚拟图书馆的投入还包括数据的不断维护与更新,包括对原有网址的定期核实和新网址的纳入。此外,需要计算机系统开销一定的空间。以上这些投入比起数字图书馆的建设来说是很微小的。而且这些投入主要是人的智力劳动的投入,对设备的要求不高。可以这样认为,目前大部分图书馆尚不具备建设数字图书馆的条件,但是已具备了建设虚拟图书馆的条件。

(2) 虚拟图书馆建设见效快。由于不需要具备自身馆藏,更不需对这些馆藏进行数字化处理,而是“拿来主义”,将网上资源选择组织,为我所用,见效快。

(3) 虚拟图书馆对网站网页的信息是经过再组织的,因而具有系统性和易用性。一般来说,一个专业的虚拟图书馆是按一定的逻辑次序组织

的。如分列学科总论、各分支学科的进展、学术刊物和其他出版物、有关的教育机构、学术会议等活动、学术团体、著名学者等等。这种系统的分类纲目,使众多的网址条理分明,便于使用。此外,虚拟图书馆还具备关键词等的布尔检索、顺序扫描检索、字段限定等的检索功能,提供信息集合中网址的直接检索和深度检索。这种对信息的系统化再组织,也是虚拟图书馆较之一般综合性搜索引擎的一大优势。

虚拟图书馆的主要作用是实现网上信息资源的共享。对于从事一定学科专业的研究工作者来说,使用本专业虚拟图书馆,无异于拥有自己的专业资料室,而且其价值大大超过这种资料室。主要表现在:它的信息资源比专业资料室要广泛得多,举凡网上已有的重要资源皆能为我所用;一些只在网上发布的信息不具有相应印刷版,唯有通过虚拟图书馆而获得;信息的更新要比专业资料室快,许多专业论文集、期刊全文在网上出现要比专业资料室通过采购获得快得多;虚拟图书馆不存在借阅证和借阅资格问题,且每天 24 小时都可存取,而无节假日闭馆及下班之虞,使用者在虚拟图书馆中所做的仅为检索、浏览和下载,不需办理借阅手续,不需归还和归架,同时也保护了自己阅读的隐私权。虚拟图书馆还提供了建立个人本机文档的可能性,即使用者根据自己的需要,从虚拟图书馆及其链接网站下载自己感兴趣的全文、事实或书目信息,并利用一定的软件(如 Qins, Lotus Notes 等数据库系统),在自己的微机中建立个人文档系统,从而成为个人化的电子文库。

虚拟图书馆也有局限性:一是网上信息资源的不足。文献信息的数字化需费时日,许多印刷文件(特别是全文型)尚未进入网络,特别是历史性文献尤其如此,从而不能在网上找到。二是网上信息较不稳定。网站的出现、改名与消失随时都有可能发生,即使是长期存在的网站也会随着时间的推移而更新内容和软件,因此使以前曾经存在的资料消失。这些网站及其内容的动态性不利于信息资源的长期保存。另外,免费存取的资料虽然不少,但许多有价值的信息是要付费检索和下载的。它们在虚拟图书馆中的出现,只起线索指引作用,或者仅起广告作用,对学者来说无实惠可言。其三,即使是提供公开浏览的网站网页,有的通过技

术措施,不能进行内容下载。第四,目前国内网络传播速度很低,要通过虚拟图书馆获得有关信息需花费很长时间,有时甚至由于传输超时而自动中断。联系到我国电信费用偏高,这对使用者是一不能不考虑的经济负担。

虽然虚拟图书馆不能取代资料室,但两者可以共存,形成互相补充的关系。

2 国外网上的图书情报学信息

在浩如烟海的网上信息中,图书情报学信息占有一席之地。其数量很难有一个精确的统计。1999 年 7 月 13 日,笔者使用全球公认的比较齐全的搜索引擎——Altavista 进行检索,检索到“Library”之下有 2100 多万个网页。在“Library Science”之下有 37233 个网页,在“Library Science journals”之下有 774915 个网页(这里显然不是指图书馆学期刊的种数,而是以期刊中文章全文出现的词为基础索引的)。在 Altavista 搜索引擎中,笔者还以中文“图书馆”进行检索,得到的网页数为 44795 个。上述不精确的数字表明,图书馆及图书馆学的网上资源是比较丰富的。更重要的是,我国各类型图书馆正在纷纷上网,建立自己的网站网页。这一前景是令人鼓舞的。

网上图书情报学有关信息大致分为以下一些类型:各国国家图书馆的网站;各国高校图书馆的网站网页;各国科研机构及专业图书馆的网站网页;各国省(州)、县、中学等图书馆的网站网页;国际和各国图书馆学协会的网站网页;图书情报学专业期刊的网站网页;图书情报学教育机构(包括招生、教学计划、课程设置、学位论文等)的网站网页;图书情报界学术活动(如会议、展览、讲座等)的网页网站;图书情报学科研成果和正在进行中的科研项目的网站网页;图书情报界著名人士介绍的网页;与图书馆学有关的出版机构及有关书目及数据库的网站网页;网上学术论坛(讨论组)有关图书情报学的网站网页;有关的电子图书馆、数字图书馆、虚拟图书馆的网页;有关的搜索引擎、网络资源指南以及图书馆名录等的网站网页;与图书情报工作自动化有关的数据库生产者、软件公司及其产品的网站网页;有关的个人网页。

以上列举的 16 种类型的网上资源还可能不全面,但总的来说,网上图书情报学有关资料是较为丰富的。要充分掌握和利用这些资源,编制专科性质的图书情报学网上资源的虚拟图书馆是必要的。

在国外,已经有人在图书情报学虚拟图书馆或网络资源指南研究方面作了工作,其成果已经上网。例如:

Internet Library for Librarians (图书馆员用的因特网图书馆), www.itcompany.com/inforetriever/

Internet public Library services for Librarians (图书馆员用的因特网公共图书馆服务)

Library and Information Science (图书情报学) www.clark.net/pub/lischank/web/library.htm

Library and Information science: General Resources (IFLA), www.nlc/hnc.ca/ifla/libdoc.htm

The Librarian's Guide to the Internet (图书馆员的因特网指南), www.concentric.net/~A.chenson

PICK: Quality Internet Resources in Library and Information science (图书情报学因特网资源) www.aber.ac.uk/~tp/www/e/

这些网上工具难以一一列举。它们可以作为编制图书情报学虚拟图书馆的借鉴参考。

3 我国网上的图书情报学资源

随着我国计算机网络建设的发展,国内大中型图书馆(其中特别是高等学校图书馆)已经纷纷上网,向社会公众提供本馆概况、书目查询、非本馆信息资源的镜像服务及其它参考检索服务等,从而实现了图书馆服务的更大范围的社会化。然而,图书情报学的教学与研究活动的上网却大大滞后,目前刚刚开始萌芽。在这个领域看来有着广阔的发展前景,它是提高图书情报学教学与研究信息交流层次、加速学科建设的必要。

最近我们利用北大的“天网”等搜索引擎,对网上的图书情报学的教学与研究信息进行了较广

泛的调查,发现其信息的分布大致如下:

(1) 高校图书情报学院系或专业的介绍。这是作为这些院系专业所在的高校信息的一部分而出现的。在这种“介绍”或“概况”中,一般包括院系专业的名称、历史、培养目标、课程设置、教师队伍、学科带头人、在校学生人数、招生计划,以及有关邮政编码、电话号码、电子邮件邮箱等信息,有的还配有一定的照片。

虽非高校,但具有硕士与博士点的图书情报单位也在网上发布有关研究生招生、培养等信息。例如中国科学院文献情报中心的网页(<http://www.las.ac.cn>)上就有博士生导师、招生计划等信息。

(2) 书目信息与期刊目次信息。关于图书情报学专科书目已经上网,这就是浙江大学图书馆叶鹰编纂的“国内图书馆学理论研究论文目录”(<http://www.lib.zju.edu.cn/yy/libj.htm>)及情报学理论研究目录(<http://lib.zju.edu.cn/yy/abslib.htm>)检索功能,看来数据是从中国科技信息所重庆分所的“中国科技期刊篇名数据库”光盘中套录而加以组织的。北京图书馆出版社(前书目文献出版社)和科技文献出版社是图书情报领域的二个专业性出版社。它们在网上均有自己的网页。其中都刊登了一些主要出版社书籍的目录,但数量极少,不能反映其全貌。有一些图书情报学专业期刊的每期目次也已上网。例如浙江大学图书馆在“电子期刊”栏下有《浙江高校图书情报学刊》1996 年各期的目次(英文),可惜仅限于 1996 年的 4 期,1997 年后中断。北京图书馆的《文献》杂志,《北京图书馆馆刊》目次等也已上网。这些目次均无检索功能,仅能逐期浏览(<http://www.nlc.nlc.gov.cn/zxxx/shamu.htm>)。《大学图书馆学报》已将 1999 年起各期目次(包括内容简介)上网(www.lib.pku.edu.cn/chnl/xuebao/index.htm)。此外,也允许直接用电子文献本投稿。

(3) 全文信息。《图书馆杂志》(上海)从 1997 年起的全文已由上海图书馆上网(<http://www.libnet.sh.cn/tsgzz/>),按期浏览,无全文检索功能。但通过天网搜索引擎可以检索篇名。《情报学报》从 1998 年起全文上网(<http://chinainfo.gov.cn/periodical/qbxb/index.htm>)。此网上刊物编得较好,有索引,有全文,提供著者、著者单位、关键

词等途径的检索手段,这是一个较理想的网页。此外,上网的全文文件还有《汕头大学图书馆通讯》(www.lib.stu.deu.cn/htdocs/tsgrx),《上海医科大学图书馆消息:上海医科大学信息快报》(www.shmu.edu.cn/newsn)。这些电子刊物主要是面向广大读者的,但也包含一些专业性文章。北京语言文化大学、广州外语外贸大学等校图书馆也将论文集全文上网。

(4) 学术会议信息。北京大学图书馆为纪念北大100周年校庆和新馆落成,1998年在网上发布了“廿一世纪大学图书馆的使命”的学术会议消息,包括会议日程、讨论主题、征稿事宜等内容。随后将100多篇论文全文上网,成为一个学术质量很高的会议文集,可浏览下载,可惜无全文检索功能。中山大学图书馆则将1997年在该校举行的第四届海峡两岸图书资讯学会会议的AB两集会议论文集的目次以及会议讨论分组情况在网上公诸于世。此外,大陆与香港高校联合举办的学术讨论会论文也已全文上网。

(5) 科研成果信息。吉林工业大学信息工程系在网上公布了历年来获奖科研成果的目录,包括项目名称、主持人姓名、颁奖单位及奖励级别等。

(6) 个人网页。在图书情报学领域中已出现个人网页,较突出的是浙江大学图书馆叶鹰的网页(heep://lib.zju.edu.cn/yy/yy.hen1)。在这个网页上,不仅提供本文前面提到的图书情报学理论研究书目,而且也发表他的著作与论文目录以及一些文章的全文。北京大学信息管理系的网站也包含了10多个学生的个人主页。这种个人主页目前较少,但将来会有发展。个人主页是作者倾注全力办好的园地,一般都会显得有所特色。这是对图书情报学资源的一种补充。

我国网上图书情报学信息正在迅速发展,但也存在如下若干问题:

(1) 上网的信息量太少,与我国图书情报学实际的研究水平很不相称。我国有50多种专业期刊,每年发表数千篇文章。加之非专业期刊和报纸上发表的图书情报学文章,笔者估计至少在六七千篇以上。然而只有极少数的几种刊物在网上发布每期目次。全文上网者更是十分罕见。其实,不少图书情报学刊物都有电子文本,并纳入清华大学的“中国学术期刊全文库”系统。清华的这个库

无论是光盘版或联机版,均要收费。一般用户难于存取。为什么不能象《情报学报》那样,既纳入清华的系统,又单独全文上网?至于上网的图书情报学研究人员介绍,仅限于博士生导师及个别学校(中山大学信息管理系)的教师,总人数只有一二十人。图书情报界的学术会议的公告和会议论文目录的报导,仅限于北京大学、中山大学等极少数网页。

(2) 上网信息门类不全,还存在严重的缺口。目前除教学信息较普遍外,研究信息、会议信息、专科学历信息等太少。特别是学位论文(硕士及博士论文)仍是缺口(有极少数在清华大学的“高校学位论文数据库”中有反映)。学位论文质量较高,但一般只印一二十本,属“难得资料”,如能通过互联网通报其题录、摘要或乃至全文,对于学科的研究有重要意义。北大的数字图书馆包含了图书情报学学位论文题目及导师姓名。这是一个良好的开端。其次,“在研科研项目”一项也是缺口。如能公布正在进行中的科研项目,可以避免选题的重复,也有助于人们了解学科发展的前沿。高校图书情报学专业毕业生信息(包括姓名、年龄、性别、籍贯、成绩、志愿等)的信息如能上网,也有助于毕业生寻找工作。河北大学信息管理系作了这方面的工作。此外,网上的学术论坛(新闻讨论组)有关图书情报学的,也无人进行整理汇编。

(3) 网上信息更新不及时。许多院系网页一年难得更新一次。由于网上信息更新不及时,致使其内容失去现实性,降低了可信度。

(4) 网上信息缺乏检索功能。上海的《图书馆杂志》提供1997年以来的每期文章的全文,只能按期浏览,不能从作者、分类、主题等途径进行检索。虽然“天网”搜索引擎对其每篇文章的题名进行自动抽取反映,但这是搜索引擎的功能,而不是杂志信息本身提供的功能。浙江大学图书馆叶鹰提供的书目也是如此。作为书目,没有检索功能,就缺乏本身作为检索工具的价值。《情报学报》的网页具备了检索功能。这是一个榜样。

4 图书情报学虚拟图书馆的设计

作为中文的专业性的图书情报学虚拟图书馆

至今很少有人做过。笔者仅见到的是台湾大学图书资讯学系陈雪华教授编纂的“图书资讯学网路资源指南”(www.lis.ntu.edu.tw/news/contents.htm)。

我们武汉大学大众传播与知识信息管理学院目前正组织力量从事图书情报学虚拟图书馆的研究与编制工作。现就设计问题与同行进行商讨。

一是收集材料的学科范围确定问题。由于国内关于图书情报学的定名及其范围界定的讨论没有得出一个明确的肯定意见。加之近年来不少学校的图书馆学系和情报学系改名信息管理系之后,又与编辑出版、大众传播、管理信息系统等学科融合,使这个问题显得更为模糊。笔者认为,学科的研究范围应当与学校行政建制有所区别。左右行政建制有许多其他因素,并且将来还可能变更,因此不能作为确定虚拟图书馆收录材料范围的主要依据。

二是收录材料的角度问题。即图书情报学与图书情报工作是否要区分的问题。顾名思义,作为“学”,主要是报导有关图书情报学教学研究的网址,其中包括教育机构及其教学计划,学科各分支的理论与方法研究,有关学科的书刊,有关的国际国内组织及其活动,有关的研究人员等等。至于国内外的各个图书馆、各具体书目索引及参考工具书与数据库等是否收录?笔者认为,对各图书馆的网址,仅收主要图书馆,以及包含一批图书馆名录的网址。后者是经过信息加工的产品,如清华大学图书馆的“国内上网图书馆”、“国外上网图书馆”等。对于各参考工具书、出版社、网上书店等也宜采取类似的做法。至于一些学科边缘性的材料,如计算机程序设计、JAVA等语言等,笔者认为只应收录直接与图书馆自动化有关的。但对于搜索引擎应充分收录。因为对用户来说,搜索引擎的使用可以在一定程度上弥补专业虚拟图书馆的不足。

三是收录材料的地域问题。笔者认为应以我国为主,特别是大陆地区的网址为主,同时收录台湾及国外的网址。台湾大学陈雪华先生编“图书资讯学网路资源指南”对大陆的有关网址收录极少。虚拟图书馆的编制体现“以我为主”的原则,不仅有利于我国同行学者的充分交流,而且由于具有自身特色,在国际上更能得到重视。

四是款目的著录事项问题。整个虚拟图书馆相等于一部主题索引,而每一网站网页在虚拟图书馆中相当于主题索引中的一个款目。款目的著录事项是决定其信息含量、引得深度、检索途径的基础。笔者认为应有下列一些基本的著录事项:网站或网页中文名称、英文名称;网址(由于镜像等原因产生多于一个网址的,都应予以记录,各网址间用一定的分隔符隔开);网站地址可用文字型地址,也可用数字型地址;关键词,可以给出一个以上的关键词;分类号:需要编制一个专用的分类表;网站网页维护者;内容简介,可数十字至数百字,除指明其复盖面外,还应指明其收录材料的起始时间、更新周期、有无检索手段等。

五是软件的选择。虚拟图书馆应有相应的软件以存贮数据、显示数据、检索数据、链接网站网页。在存贮数据方面,应具备数据录入的增删改、外部数据的成批转换入库,以及浏览网页时网页地址自动入库功能。在网页地址自动入库方面,还应解决Big5码的自动转换成GB码问题,以便在摘取网页地址的同时摘取网页名称和简介。在显示数据方面,应能对检索结果数据进行多链排序,如按与检索式贴近程序排,按网页名称排,按更新日期排等。检索数据方面应具有布尔逻辑检索、字段限定检索、全库顺序扫描检索等功能。根据检索结果链接有关网站网页,当然是作为虚拟图书馆的基本要求。台湾大学陈雪华教授编的《网路资源指南》仅提供分类浏览检索功能,而无其它检索功能,显得美中不足。随着虚拟图书馆库容的增加,应该重视检索功能的实现。款目中的各个著录字项(字段),除内容简介外,都应设为索引点,使之成为可检索的途径。

图书情报学虚拟图书馆的建设是图书情报学学科建设的一项基础工作。它的成功需要同行们的共同努力。

陈光祚 武汉大学大众传播与知识信息管理学院教授,博士生导师。通讯地址:湖北武汉。邮编 430072。

(来稿时间 1999-07-20。编发者:李万健)