

● 华薇娜 许美琪

美国过刊数据库概况及对 我国开发过刊数据库的思考*

摘要 20世纪末,美国开始出现过刊数据库,例如JSTOR、Wilson公司期刊文献回溯数据库、MUSE项目及期刊题录索引等。国内目前尚没有1989年之前的过刊回溯数据库。在国内建设过刊数据库具有减小馆藏压力、挽救过刊、优化服务等重要意义。建设过刊数据库既有不利因素,也有可操作性。参考文献8。

关键词 期刊数据库 回溯数据库 期刊管理 期刊服务

分类号 G255.2

ABSTRACT At the end of the twentieth century, some databases of non-current periodicals appeared in the United States, such as JSTOR, H. W. Wilson's retrospective databases, Project MUSE full-text journals. However, there has not been any retrospective database for periodicals published in China before 1989. In this paper, the authors discuss the significance, difficulties and operability for the development of such a database. 8 refs.

KEY WORDS Database of periodicals. Retrospective database. Management of periodicals. Services of periodicals.

CLASS NUMBER G255.2

按收藏的时间分类,期刊可分为现刊和过刊。人们一般把当年的期刊称为现刊,往年的期刊经过整理装订成合订本,加工入藏之后便成为过刊。

20世纪90年代中、后期国外又开始出现并逐步发展起一些回溯性的期刊数据库。这类数据库主要用于收录过去一定时间范围内的某些期刊文献资料信息,大有与现期刊数据库一争高低之势。仅人文、社科方面,国外近期出现的回溯性数据库就有美国的JSTOR、MUSE、期刊题录索引等,各有200多种过刊的全文,最早的过刊回溯至1683年。这类数据库收录的原始文献,无论是类型还是范围,相对都比较集中,但时间跨度大,也充分利用了图书馆馆藏资源的优势。相比较而言,国内的全文期刊数据库(网络版)目前主要侧重于当前期刊,最早的回溯至1989年。因此,国内数据库的开发在过刊回溯方面还有很大的发展空间。本文试图就这一主题对国内外过刊数据库做一番比较与分析,以引起人们的重视。

1 美国主要过刊数据库简介

1.1 JSTOR 西文过刊全文数据库

JSTOR (<http://www.jstor.org>)全名为Journal Storage,是一个对过期刊进行数字化处理的数据库。该数据库创始于1995年8月,由美国Mellon基金会的会长William G. Bowen策划创建。该库最初的目标是要建立一个规模虽小,但包罗各种由创刊号开始的重要期刊的数据库。开始时是将10种经济类和历史类期刊共75万页数据进行数字化,并连接电子图像版,网络中心设在密西根大学(University of Michigan),并通过普林斯顿大学(University of Princeton)的镜像站向读者发送。该数据库目标定位在有计划地建立核心学术性过期刊的数字化文档,以节省空间,同时提供资料检索的功能,有效提高使用的便利性的层面

上。为了消除期刊出版单位对纸本订户可能因此而取消订购印刷本期刊的忧虑,JSTOR采用“活动墙”(moving wall)的模式,可以将过刊锁定在距离当前2至5年内。近来,由于有些期刊出版单位的积极努力,有些期刊的范围从过刊延伸到了现刊。还有些期刊自建和发行网络版,于是将原有的活动墙固定在一定的发行年代上,其后暂时停止在JSTOR的服务,故又有“固定墙”(Fixed Wall)之称。目前JSTOR的全文库是以政治学、经济学、哲学、历史等人文社会学科主题为中心,兼有一般科学性主题共20多个领域的代表性学术期刊的全文库。从创刊号到最近两三年前过刊都可用影像来浏览全文。最早的刊期回溯至17世纪。至2003年12月1日,目前该数据库共收录(包括即将收录)224个期刊出版单位的共包括6个专题的360多种期刊,用户已覆盖世界上70多个国家,其中美国以外的用户近1/3。

JSTOR的具体内容包括:

(1)Arts & Sciences I Collection 人文社科综合主题一:收录人类学、亚洲学、非裔美国人研究、经济学、生态学、数学、哲学、政治学、教育学、财政、历史、文学、人口统计学、社会学、统计学等15种人文社会学科领域117种学术性期刊全文资料。

(2)Arts & Sciences II Collection 人文社科综合主题二:JSTOR另一人文社会主题的资料库,包括历史、经济、亚洲研究、考古,以及非洲研究、拉丁美洲研究、中东研究等新增主题的期刊。已经有122种刊上网。

(3)Arts & Sciences III Collection 人文社科综合主题三:这是一个正在进行的文献集合,将包括至少120种艺术与人文期刊,如音乐、语言文学、文化研究、电影、民俗、表演艺术、宗教等。这个数据子库将于2005年底完成。

(4)Business Collection 商业主题:有关商业方面的期刊46种,包括经济与金融类的期刊,另将增加20种会计、市

* 本文为江苏省社会科学规划基金资助课题的研究成果之一。

场、管理、运营方面的刊物。

(5) Ecology & Botany Collection 生态学与植物学主题: 收录生态学与植物学主题的 29 种刊。其中有 6 种与“人文社科综合主题一”重合。

(6) General Science Collection 综合科学主题: 收录 7 种物理、数学、工程学和生物学方面经典的学术性期刊全文资料, 如《科学》(Science), 多数回溯至 17 世纪。

(7) Language and Literature Collection 语言与文学主题: 目前收录多个国家语言文学方面的期刊 46 种, 涉及多种语种, 如阿拉伯语、意大利语、德语等, 其中不少期刊与人文社科类所收刊物重合, 另有 33 种期刊将在近期推出的“人文社科综合主题三”中出现。

(8) Mathematics and Statistics 数学与统计主题: 收录数学与统计方面的核心期刊 30 种, 不过, 这些期刊几乎都可在前文所介绍的主题中的期刊相重合。

(9) Music 音乐主题: 收录多个国家的音乐研究和音乐理论期刊 32 种, 这是 JSTOR 数据库中国际化范围最宽的主题, 包括荷兰、克罗地亚、匈牙利、德国、法国等国。所有这 32 种期刊都包括在即将面世的“人文社科综合主题三”中。

1.2 期刊题录索引

期刊题录索引 (Periodicals Contents Index, <http://pci.chadwyck.com/>), 简称 PCI。这原是英国的一个期刊数据库, 后由美国 ProQuest 的信息与学习公司 (Information and Learning Co) 并购。该数据库主要包括大量的期刊题录, 尤以欧洲的期刊为最。目前收录 1770~1995 年 200 多年间人文、社会科学领域 37 个学科, 包括英语、法语、德语、意大利语、西班牙语及其他西方语种的 40 多种语言的学术性期刊 4100 多种, 并每年增加 300~500 种期刊, 预计近期内将发展到 5000 种。但目前该数据库中有 250 多种期刊的 80 多万篇论文提供全文数据, 预计 2003 年底全文期刊将扩大到 300 种。同时, 在 PCI 中还有 130 多种期刊的题录著录可直接链接到 JSTOR 的人文社科期刊主题一和主题二的相关全文。

1.3 美国 Wilson 公司期刊文献回溯数据库

美国 Wilson 公司是以出版人文社会科学期刊文献索引、文摘著称的出版公司。90 年代中后期, 在其原印刷版检索刊物的基础上, 发行了多个主题的基于网络的索引文摘数据库。近年来, 该公司又相继发行了多个期刊文献回溯数据库, 例如: Readers' Guide Retrospective: 1890 - 1982 (期刊文献读者指南: 1890 - 1982), Humanities Index Retrospective: 1907 - 1984 (人文科学索引: 1907 - 1984), Social Sciences Index Retrospective: 1907 - 1983 (社会科学索引: 1907 - 1983), Art Index Retrospective: 1929 - 1984 (艺术索引 1929 - 1984) 等, 将其相应的印刷版的内容自创刊期起全面收录进来。

1.4 MUSE 项目

MUSE 项目 (Project MUSE, <http://muse.jhu.edu/>), 最早于 1995 年由美国约翰·霍普金斯大学出版社与该校米尔顿·艾森豪威尔图书馆合作, 通过万维网提供该校全文期刊的服务项目。1999 年发展成为独特的非赢利性的出版机构。现该数据库已扩大到提供 30 多个出版单位的 220 多种全文期刊的检索与浏览服务。该数据库的学科范围以人文、艺术和社会科学为主, 具体包括文学评论、历史、视觉与表演艺术、文化研究、教育、政治学、性研究、经济学等。目前仅对机构团体提供服务, 其用户已遍及美国全国并扩大到加拿大、英国、法国、德国、意大利、奥地利、荷兰、芬兰、瑞

士、瑞典、澳大利亚、墨西哥、埃及、希腊、以色列、日本、韩国等国家。这是个以近期期刊为主的人文社科专题期刊数据库。近来, 该项目也启动了部分期刊与 JSTOR 数据库中相关过刊全文的连接。

2 国内过刊的现状与发展的需求

目前, 我国期刊数量庞大, 内容丰富。随着期刊在传递信息方面作用的日益显著, 学术性研究性以及一些历史悠久的公共图书馆中“重书轻刊”的现象得到了扭转, 即使是在经费紧缺的情况下, “舍书保刊”也成为许多单位的文献采购政策。因此, 经过时间的累积, 过期刊已普遍成为这些图书馆馆藏的重要构成部分, 数量很是庞大。于是, 科学的期刊管理与开发逐渐成为人们需要面对的重要问题。

对于期刊来说, 建立数据库已成为目前最为有效的典藏方式, 而且便于用户检索使用, 不受时间和地域的限制。但是, 就目前的期刊数据库来看, 一般收录的都是近期期刊。目前收录期刊时间最早的是中国科技期刊全文数据库 (维普), 它收录了 1989 年至今的主要科技期刊。另一个比较著名的中国学术期刊全文数据库则是收录了 1994 年至今的期刊, 这对于 80 年代以及 80 年代之前的期刊文献的检索需求很难满足。而且从数据库整体来说, 比较侧重于科技期刊的收藏, 对于人文学科的覆盖面相对较小。

造成上述情况, 除了客观因素之外, 思想认识上的偏颇是重要的根源所在。对于过刊来说, 面临的主要问题是“老化”。随着科学技术的发展, 知识“老化”的速度加快, 科学文献的时效性缩短。随着时间的推移, 期刊文献就会逐渐失去其时效性, 被更新的知识所代替。但是, 文献的老化并不意味着图书馆的过刊就失去了作用。实际上, 过刊因为能反映科学的系统性和完整性, 展现某一学科的发展历史, 真正体现期刊的学术价值。其资料价值与历史价值会随着时间的流逝而更显得珍贵。

但是, 过刊在保存方面具有一定的难度。首先因其面广量大, 需准备大量的放置空间和投入大量的人力财力用于日常管理。但倘若建立过刊数据库, 将纸质文献转换为电子文献, 使过刊通过数据库得到了开发, 就可节省常规期刊管理方面所需要的人力、财力和空间。

其次, 近年来, 随着人们对过刊需求的增加, 过刊也逐渐实现开架式管理、外借式服务。过刊开架借阅, 能开发期刊资源, 充分利用其价值, 是提高过刊利用率、更好地为教学科研服务的重要方式。开架不仅有利于读者根据已知文献线索直接、方便、自由、迅速地在书架上挑选所需期刊, 而且通过浏览过刊可以获得某些意想不到的收获, 有利于读者查找收集到与自己正从事课题研究有关的文献资料, 也有利于读者获得与该课题间接相关的其他学科资料。开架借阅还方便了一些需要反复查阅大量资料的读者, 大大提高了过刊的利用率。但是, 这种开架式管理、外借式服务也存在着一定的弊端。目前相当部分的期刊, 每种几乎只订一份, 这一份杂志可谓名副其实的“孤本”了。有的杂志在现刊阅览室就受到了不同程度的损害, 如果开架管理, 恐怕会面目全非、污损、割页、丢失将会更严重。而且这些缺憾有时很难弥补。因此, 过刊的完整性、系统性势必遭到破坏, 利用价值势必受到影响。

再者, 开架式管理要有足够大的空间容纳读者。但是, 事实上, 过刊数量庞大, 即使进行开架式管理, 还是只能满足一小部分读者的需求。还有很多读者需要在过刊库中查找资料, 面对浩如烟海的文献资料, 很多人只能望“海”兴

叹了。

3 关于建立国内过刊数据库的一些思考

3.1 建立过刊数据库的意义

(1)是减小馆藏压力的需要。

随着时间的推移,馆藏的过刊必然发生数量与质量的变化。累积种类和累积册数不断增加,所占馆舍空间逐年膨胀;期刊日渐老化,保存到一定年份,复印需求减少,某些期刊甚至已无读者需求;在借阅和保管过程中,部分过刊严重损坏,难以修复,不能提供急需借阅;在馆藏年代上连续性极差或缺期较多的期刊,其使用价值也受到很大影响;某些学科领域的读者群消失或者骤减,原本适应这些读者群需求的期刊也无人问津。虽然图书馆每年都进行“剔除”工作,但是面对数量庞大且都较为珍贵的过刊,“剔除”工作也很困难,需要耗费大量的人力和物力。建立相应的数据库是解决这个问题的最好方法。数据库无需实际空间,能够解决馆藏空间问题;只需将相应内容保存到数据库中,就可以不必再花过多的精力保存原始过刊。而且无需进行“剔除”,所有的文献资料都可以收录到数据库中。

(2)是挽救过刊的需要。

除了近期过刊增长快,收藏压力大的矛盾外,还存在着远期过刊本身散失、破损、严重残缺不齐的问题。这是我国期刊发展过程中的一个特殊问题。即使是在藏的过刊,也会因为收藏条件的限制而遭损毁。这些过刊与古籍还不同,它们没有相当于古籍收藏的待遇,在常温、灰尘、日光的诸多影响下,许多期刊已不堪翻阅,失去了被利用的机会。即使有馆藏,有些也已不堪多次复印了。倘若多家馆藏单位合作建立过刊数据库,甚至联合海内外、国内外的馆藏,对较为完整地挽救、保存与开发国内的过刊将会是十分有意义的。因此,建立过刊数据库,实际上也是挽救过刊的一种好方法。

(3)是深度开发过刊资源,发挥过刊价值的需要。

期刊的价值并不完全在于收藏,而是需要人们将其内容和形式特征揭示出来,将其全文显示出来,面向更广大的用户,充分发挥其价值。过刊文献的开发是提高过刊利用率的基础,其开发的程度和深度直接影响过刊文献的利用程度与利用效益。开发工作做得越细致深入,越符合读者需要,则越有读者利用。完善的过刊目录索引体系是建立过刊数据库的重要基础,而开发和建立过刊数据库又必将推动和加强过刊的二次文献和三次文献信息的整理和加工,诸如主题目录、索引、篇名目录、索引的建立,以及过刊文献内容的摘要和综述的编制等。而这正是数据库所能达到的目标。利用现代信息技术对图书馆丰富的过刊资源进行开发利用,让读者通过计算机快速检索到过刊资源中的各种信息,这是提高过刊资源利用率的方便有效的途径,也是期刊管理工作的一个必然趋势。

(4)是优化读者服务的需要。

由于过刊的数量庞大,分散零乱,管理也不够有序化,人们普遍感到使用十分不便。建立过刊数据库,可以解决过刊管理落后的局面,大大提高服务质量和工作效率。工作人员要了解和掌握过刊库的全部情况,只需通过调用数据库就可以知道本馆过刊的入藏情况,检索、查找、清点都十分方便。

(5)是促进开发人文社会科学研究的需要。

相对于科技期刊,人文社科期刊在某种程度上具有一定的社会效益的滞后性。这种期刊所刊载的观点、思想,要发挥作用,需要一个吸收、转化、实验、检测、推广的较长过程。其价值效益也具有间接潜在性。它所承载的内容是一种潜在的生产力,一般是为科学技术、经济建设、社会进步提供科学的指导思想、思维方式和价值观念,起着推动经济发展和社会进步的作用。从这个角度来说,开发过刊数据库对国内人文社科的开发具有十分重要的意义。

3.2 建立过刊数据库的不利因素

建立过刊数据库也存在一些不利因素。首先,过刊数据库的市场相对较小,其检索需求不及现刊数据库强烈。而且过刊多集中在人文社科类期刊,在检索的迫切性上不及自然科学类期刊。其次,由于过刊原始文献不如现刊的可获得性强,在生产环节会产生一些附加的开支,有些还需要得到海外国外资源的补充,故过刊数据库的生产成本相对较高,并相应带来其使用成本的提高。再有,过刊数据库的利用率相对较低,其主要用户群多集中在学者专家,其学科主要在人文社会科学领域。另外,随着数据库产业的国际化趋势以及世界范围的大规模使用,知识产权保护问题也会逐渐显现出来。当然,这些不利因素并不应成为不建库的理由。

3.3 建立过刊数据库的可操作性

随着人类迅速进入信息进代和知识经济时代,知识创新和技术创新给图书情报机构带来了巨大压力和巨大动力。一个单位的图书数据已远远不能满足需要,资源共享成了人们的共识,基于网络的数字图书馆在国内已经有了较好的基础,而这正是开发国内过刊数据库的基础。人们可以根据各馆的馆藏,在建立统一标准的前提下,逐步采取诸如馆际联网、地区联网、全国联网,甚至国际联网的形式,最终实现我国过刊在互联网上的大团圆。

总之,在众多的信息资源中,期刊是最主要的信息源之一。而过刊在文献信息的交流环节中又发挥着很大的作用,是人们进行科学研究不可缺少的参考数据。在期刊现刊全文数据库十分红火的今天,建立过刊全文数据库的动议是对国内信息资源整合的一种思考,希望引起人们的重视。

参考文献

- 1 赵燕群主编.连续出版物工作.北京:北京图书馆出版社,2001
- 2 于凤霞.过刊资源的优势及其开发利用.图书馆论坛,2001(6)
- 3 陈水英.高校图书馆过刊管理刍议.现代情报,2002(4)
- 4 彭仁贤,高玉华.电子全文期刊数据库:东方的中国期刊网与西方的JSTOR.中国图书馆学报,2002(2)
- 5 苗洁.简述中国期刊发展的历史.河南社会科学,2001(3)
- 6 柳丹枫.新世纪:中国期刊发展趋向及其启示.图书馆建设,2002(2)
- 7 吴绮云.高校图书馆过刊计算机管理述论.广东行政学院学报,2001(3)
- 8 华薇娜.网络学术信息资源检索与利用.北京:国防工业出版社,2002

华薇娜 南京大学信息管理系教授.通讯地址:江苏南京.邮编210093。

许美琪 南京大学信息管理系99级本科生.通讯地址同上。

(来稿时间:2003-12-10)