

中国科协所属期刊在国内三大期刊数据库全文上网情况*

程维红¹⁾ 任胜利^{2,3)} 王应宽⁴⁾ 严 谨³⁾

收稿日期:2010-05-03

修回日期:2010-06-21

1) 中国农业科学院作物科学研究所《作物学报》编辑部, 100081 北京市中关村南大街 12 号, E-mail: chengwh@mail.caas.net.cn

2) 国家自然科学基金委员会, 100083 北京市双清路 83 号, E-mail: rensi@mail.nsf.gov.cn

3) 《中国科学》杂志社, 100717 北京市东黄城根北街 16 号, E-mail: yanjin@scichina.org

4) 农业部规划设计研究院《农业工程学报》编辑部, 100125 北京市朝阳区麦子店街 41 号, E-mail: wangyingkuan@163.com

摘 要 2009 年 6 月的统计结果表明,中国科协所属 954 种期刊在国内三大期刊全文数据库(CNKI、万方、维普)的全文上网情况为:CNKI,872 种,占总数的 91.4%;万方,856 种,占 89.7%;维普,763 种,占 80.0%;有 729 种期刊在三个数据库同时全文上网(占 76.4%)。调查分析了中国科协期刊在三个数据库全文上网的回溯年代、全文更新的时滞。2010 年 1 月,基于 CNKI 的“中国引文数据库”在线统计了中国科协期刊的上网文献数、总下载频次和篇均下载频次,比较分析了不同学科期刊和不同类别期刊的平均上网文献数、平均总下载频次和平均篇均下载频次。结合中国科协期刊的数字化状况讨论了我国科技期刊多重数字化现象及数字化产业升级方面存在的问题。

关键词 中国科协所属期刊 CNKI 万方 维普 全文上网 数字化

科技期刊的商业数据库传播模式是伴随着计算机技术、网络技术和现代通信技术的发展应运而生的一种传播模式,以代替科技期刊传统的纸质传播模式来满足人们阅读方式改变的需要。在我国,科技期刊商业数据库已经发展了十几年,现在已相对成熟^[1]。主要数据库有 CNKI(国家知识基础设施, National Knowledge Infrastructure)系列数据库之一的“中国期刊全文数据库”(以下简称“CNKI”)^[2]、万方数据资源系统的“万方数据——中国数字化期刊群”数据库(以下简称“万方”)^[3]和重庆维普资讯的“中文科技期刊数据库”(以下简称“维普”)^[4]。它们是我国科技期刊数字化的主要实践者和推动者,在读者中具有较高的认知度和认可度,有相对成熟的营利模式和绝大部分市场占有率。三个数据库不仅将期刊内容本身数字化,而且将大量的文献、信息等资源进行分析、加工和整合,形成了海量信息库。信息库具有强大的检索功能,并实现了系统内相关文献之间的动态互链接,大大提高了期刊的利用效率和利用效果。

受中国科学技术协会学会学术部的委托,2009 年 6 月,本文作者分别以中国科协所属 954 种期刊的期刊刊名为检索词,逐刊检索并统计这些期刊被三个数据库全文收录情

况,重点调查三个数据库全文上网的回溯年代和更新时滞;2010 年 1 月,通过 CNKI 的“中国引文数据库”在线统计了它们在 CNKI 的上网文献数、总下载频次和篇均下载频次,以期分析中国科协期刊在国内三个期刊全文数据库上网情况和网上学术影响力。

1 中国科协期刊被 CNKI、万方和维普全文收录情况

1.1 全文收录概况

中国科协所属期刊被三个数据库“全文收录”和“未收录”的统计结果见表 1。中国科协 954 种期刊中有 872 种(占 91.4%) 在 CNKI、856 种(占 89.7%) 在万方、763 种(占 80.0%) 在维普全文上网。954 种期刊中有 729 种在三个数据库同时全文上网,占 76.4%。与 2007 年相比^[5],被 CNKI 和维普全文收录的期刊比例基本没有变化,被万方全文收录的期刊比例增加了 3.7 个百分点,被三个数据库同时全文收录期刊所占比例增加了 1.0 个百分点。CNKI 和万方收录的数字化期刊同时具有目次、摘要和全文,并且目次、摘要和全文同步;而维普收录的期刊中有 140 种没有全文,只有目次和/或摘要。

* 基金项目:中国科协科技期刊数字化、网络化建设及其变化发展趋势——《中国科协科技期刊发展报告(2010)》专题之一,项目编号: QK2009002 - B

表1 中国科协期刊被 CNKI、万方和维普全文收录情况(2009年6月)

数据库	收录情况	2009年		2007年	
		期刊数量	占中国科协954种期刊的比例(%)	期刊数量	占中国科协898种期刊的比例(%)
CNKI	全文收录	872	91.4	823	91.7
	未收录	82	8.6	75	8.3
	合计	954	100.0	898	100.0
万方	全文收录	856	89.7	772	86.0
	未收录	98	10.3	126	14.0
	合计	954	100.0	898	100.0
维普	全文收录	763	80.0	727	81.0
	未收录或收录无全文	191	20.0	171	19.0
	合计	954	100.0	898	100.0
3个数据库同时全文收录		729	76.4	677	75.4
2个数据库同时全文收录		131	13.7	127	14.1
1个数据库全文收录		42	4.4	41	4.6
3个数据库均未全文收录		52	5.5	53	5.9
合计		954	100.0	898	100.0

1.2 收录全文的回溯年代

中国科协954种期刊在CNKI、万方和维普全文上网的回溯年代分布差异较大(表2)。CNKI收录期刊70.3%回溯至创刊号,万方和维普回溯至创刊号的期刊分别只占4.7%和11.7%。由于CNKI致力于从创刊号起收录期刊,因此回溯年代无明显的个别年度分布高峰。万方回溯年代的分布高峰在2000年(253种,占29.6%)、1999年(227种,占26.5%)、1998年(171种,占20.0%)和2001年(145种,占16.9%)。维普回溯年代分布高峰在1989年(470种,占61.6%)。

表2 CNKI、万方和维普全文收录中国科协期刊的回溯年代(2009年6月)

全文收录 回溯年代	CNKI		万方		维普	
	刊数	所占比例(%)	刊数	所占比例(%)	刊数	所占比例(%)
1950年以前	12	1.4	0	0	0	0
1951-1960年	55	6.3	0	0	1	0.1
1961-1970年	33	3.8	0	0	1	0.1
1971-1980年	165	18.9	0	0	4	0.5
1981-1990年	287	32.9	0	0	503	65.9
1991-2000年	270	31.0	651	76.0	186	24.5
2001年以后	50	5.7	205	24.0	68	8.9
合计	872	100	856	100	763	100
从创刊号起收录	613	70.3	40	4.7	89	11.7

1.3 全文上网的更新时滞

中国科协954种期刊在CNKI、万方和维普全文上网的更新情况显示(表3),在CNKI,“滞后1期”所占比例最大(占50.3%);在万方,“滞后2期”和“滞后3期”所占比例最

大(分别占34.2%和33.0%);在维普,“滞后4期及以上”所占比例最大(占41.3%)。CNKI中与印刷版同步“现刊”所占比例明显高于万方和维普。

值得注意的是,虽然CNKI全文上网的更新情况最好,但“滞后4期及以上”的比例却高于万方。原因在于,111种中华医学会系列期刊与万方签署了独家期刊数据库合作协议,CNKI数据库中没有这些期刊2006年或2007年以后的全文;7种英文版期刊分别与IoPP、Nature Publishing Group、Oxford University Press、Springer和Wiley-Blackwell等国外出版商合作签订了排他性的出版合同,CNKI数据库中只有期刊与国外出版社合作前的全文。

表3 中国科协期刊在CNKI、万方、维普全文更新时滞(2009年6月)

更新时滞	CNKI		万方		维普	
	刊数	所占比例(%)	刊数	所占比例(%)	刊数	所占比例(%)
现刊	46	5.3	25	2.9	11	1.4
滞后1期	439	50.3	293	34.2	164	21.5
滞后2期	163	18.7	282	33.0	183	24.0
滞后3期	56	6.4	122	14.3	90	11.8
滞后4期及以上	168	19.3	134	15.6	315	41.3
合计	872	100	856	100	763	100

1.4 未被全文收录的中国科协期刊

调查表明(表4),中国科协954种期刊中有225种未被三个数据库中的一个、两个或三个收录,未被收录的期刊分别占中国科协学术类、技术类、科普类、检索类和综合类期刊^[6]的17.2%、18.0%、90.0%、100.0%和43.5%,其中科普

类和检索类期刊所占比例明显较高。有 34 种英文版期刊未被一个、两个或三个数据库收录,占中国科协全部 69 种英文版期刊的 49.3%,这与许多英文版期刊与国外出版商合作签订了排他性的出版合同有关。

有 52 种期刊的全文未被三个数据库中任何一个收录,这其中有科普类期刊 32 种,占 61.5%,占中国科协全部 60 种科普类期刊的 53.3%。52 种期刊中有 30 种另有自建的网站,22 种无自建的网站。

表 4 中国科协期刊未被 CNKI、万方、维普全文收录情况
(2009 年 6 月)

类别	未收录期刊数				954 种期刊 类别分布(B)	A/B (%)
	一个 数据库	两个 数据库	三个 数据库	合计 (A)		
学术类	91	12	11	114	661	17.2
技术类	26	7	0	33	183	18.0
科普类	7	15	32	54	60	90.0
检索类	1	0	3	4	4	100.0
综合类	6	8	6	20	46	43.5
合计	131	42	52	225	954	-

2 中国科协期刊在 CNKI 的上网文献数、总下载频次和篇均下载频次

为了解中国科协期刊在 CNKI 的上网文献数和被使用情况,利用 CNKI 的“中国引文数据库”^[7]逐一检索被 CNKI 全文收录的 872 种期刊(表 1)的上网文献数、总下载频次,并根据每种期刊的上网文献数和总下载频次计算该期刊的篇均下载频次。

2.1 期刊上网文献数分布

中国科协期刊在 CNKI 上网的全文文献数量分布表明:有 71.0% 期刊的上网文献数在“5000 篇以下”,分布高峰是“2001~3000 篇”(占 19.6%);上网文献数在“10000 篇以上”的期刊有 54 种,其中“医药卫生”类期刊 33 种,占 61.1%,远高于全部 872 种期刊中“医药卫生”类期刊所占比例 26.2 个百分点;上网文献数在“1000 篇以下”的 51 种期刊中有英文版期刊 23 种,占全部 57 种英文版期刊的 40.4%,远高于全部 872 种期刊中英文版期刊所占比例(6.5%)。

上网文章数最多的 10 种期刊是《现代中西医结合杂志》(39380 篇)、《中国组织工程研究与临床康复》(37953 篇)、《中国误诊学杂志》(37410 篇)、《中国科技信息》(32106 篇)、《实用医技杂志》(28996 篇)、《计算机工程与应用》(25067 篇)、《职业与健康》(22963 篇)、《福建医药杂志》(21834 篇)、《中医杂志》(21015 篇)和《陕西中医》(19143 篇)。

2.2 总下载频次分布

中国科协期刊在 CNKI 的总下载频次分布情况显示:有

77.8% 期刊的总下载频次在 200000 次以下(50000 次以下占 18.7%,50001~100000 次占 28.7%,100001~150000 次占 19.1%);总下载频次介于 200001~300000 次的期刊有 90 种,占总数的 10.7%;总下载频次大于 300000 次的期刊有 97 种,占总数的 11.5%。

总下载频次在“50000 次以下”的 157 种期刊中有英文版期刊 49 种,占全部 57 种英文版期刊的 86.0%,其中有 29 种英文版期刊的总下载频次在 10000 次以下。

总下载频次最多的前 10 位的期刊为《心理科学》(1088601 次)、《系统仿真学报》(958693 次)、《计算机应用研究》(892784 次)、《中草药》(827084 次)、《生态学报》(813621 次)、《中国电机工程学报》(796242 次)、《岩石力学与工程学报》(760826 次)、《计算机工程与设计》(744353 次)、《电子学报》(738920 次)和《建筑学报》(664029 次)。

2.3 篇均下载频次分布

由每种期刊的上网文献数和总下载频次可计算出该刊篇均下载频次,结果表明,篇均下载频次分布在 50 次以下的占 74.5%(分布高峰为 21~30 次,占 19.4%;11~20 次占 18.1%;31~40 次占 17.3%;41~50 次占 12.8%),篇均下载频次大于 100 次的期刊只有 30 种(占总数 3.6%)。

篇均下载频次“10 次以下”的 58 种期刊中有 30 种是英文版期刊,占全部 57 种英文版期刊的 52.6%,远高于全部 872 种期刊中英文版期刊所占比例(6.5%)。综合英文版期刊在 CNKI 中上网文献数和总下载频次的分布情况,说明英文版期刊在 CNKI 中被使用的情况很不理想。

篇均下载频次最高的 10 种期刊为《心理科学》(192 次)、《心理学报》(186 次)、《中国土地科学》(177 次)、《人文地理》(171 次)、《中国天然药物》(152 次)、《自然资源学报》(150 次)、《中国管理科学》(149 次)、《经济地理》(146 次)、《高分子通报》(141 次)和《生态学报》(136 次),其中有 6 种为社会科学类期刊。说明社会科学类期刊在 CNKI 中被使用的情况较好。

2.4 不同学科期刊的平均上网文献数、平均总下载频次和平均篇均下载频次

为了解不同学科期刊上网文献的被使用情况,统计了中国科协不同学科期刊的平均上网文献数、平均总下载频次和平均篇均下载频次并做了排序。由表 5 可见,“医药卫生”类期刊的平均上网文献数位列第 1 位,但其平均总下载频次和平均篇均下载频次仅列第 10 位和第 12 位,说明“医药卫生”类期刊的上网文献量虽大,但文献上网后的被使用的情况却很不理想;与之相反,“文化/科学/教育/体育”类期刊的平均上网文献数虽位列第 5 位,但其平均总下载频次和平均篇均下载频次却均列第 1 位,说明“文化/科学/教育/体育”类期刊有较突出的网上影响力。

表5 中国科协不同学科期刊上网文献的被使用情况(2010年1月)

学科	刊数	平均上网 文献数	平均上网 文献数排序	平均总下载 频次	平均总下载 频次排序	平均篇均 下载频次	平均篇均下载 频次排序
工业技术	262	4047	6	166399	7	41	10
航空/航天	9	3410	10	220217	2	50	3
环境科学/安全科学	8	3638	8	186936	3	57	2
交通运输	21	2506	11	105303	11	46	8
经济	9	4488	3	171338	4	49	5
农业科学	51	4156	4	127956	9	41	11
生物科学	46	3622	9	171311	5	47	6
数理科学和化学	71	3765	7	168123	6	46	7
天文学/地球科学	47	2307	12	95970	12	42	9
医药卫生	304	5647	1	119432	10	29	12
自然科学总论	19	5560	2	160132	8	50	4
其他(文化/科学/教育/体育)	25	4108	5	268901	1	72	1
	872	-	-	-	-	-	-

2.5 不同类别期刊的平均上网文献数、平均总下载频次和平均篇均下载频次

为了解不同类别期刊上网文献的被使用情况,统计了中国科协不同类别期刊的平均上网文献数、平均总下载频次和

平均篇均下载频次并做了排序。由表6可见,学术类期刊“平均上网文献数”虽列第5位,但其平均总下载频次和平均篇均下载频次却列第1位,科普类期刊则与之相反,说明学术类期刊上网文献的使用情况要高于其他类别的期刊。

表6 中国科协不同类别期刊上网文献的被使用情况(2010年1月)

类别	刊数	平均上网 文献数	平均上网 文献数排序	平均总下载 频次	平均总下载 频次排序	平均篇均 下载频次	平均篇均下载 频次排序
学术类	640	4365	5	150487	1	42	1
技术类	181	4408	4	145879	2	35	2
综合指导类	37	5813	3	113555	3	29	3
科普类	13	6572	1	93580	4	18	4
检索类	1	6005	2	68715	5	11	5
合计	872	-	-	-	-	-	-

3 讨论

3.1 我国科技期刊的多重数字化现象

中国科协所属期刊,乃至我国的科技期刊近年来呈现多重数字化现象。除本文展示的中国科协所属期刊中有80%以上在CNKI、万方、维普全文上网外,其中的许多期刊还同时在“台湾华艺中文电子期刊服务数据库”、“龙源期刊网”、“中国科技论文在线”、“首席医学网”、“光学期刊网”等期刊网全文上网,许多英文版期刊通过与国外出版商签定合作出版协议在其网络出版平台全文上网,还有许多期刊通过自建的网站将期刊数字化后发布。

有版权的我国大多数期刊出版单位只是数字资源的提供者而不是控制者有关,即真正处于我国网络期刊出版平台主导地位的不是内容版权的拥有者,而是CNKI、万方、维普等商业数据库。

近两年,这种多重数字化现象呈快速发展态势,据笔者2007年和2009年的调查^[5,6],加入“中国科技论文在线”的中国科协期刊由2007年的2种增加到2009年的13种;加入“首席医学网”的由7种增加到23种;加入“光学期刊网”的由10种增加到12种;在国外出版平台全文上网中国科协期刊由42种增加到60种。自建网站的中国科协期刊由2007年的448种增加到2009年的581种^[6]。

3.2 我国科技期刊的数字化产业升级任重道远

20世纪末以来,数字技术和网络技术的进步迅速改变着

期刊的多重数字化造成资源的浪费,其产生的原因与拥

包括科技出版在内的整个出版业的生存方式。在这些新技术的影响下,发达国家的科技出版业纷纷以积极姿态迎接数字化的挑战。目前,美国、英国、德国以及荷兰的一些大型科技出版企业和主流科技期刊基本实现了数字化转型,实现了科技出版的产业升级。不仅出版活动的全流程基本实现了数字化,而且一些大型科技出版企业(科技文献首次出版商)也嬗变为科技内容资源服务商^[8],相继于20世纪90年代中后期创建了集期刊稿件采编、文章在线发布(大部分为现刊和网络预出版)、全文数据库、并集成各类资源的一站式网络出版平台^[9]。同时探索出了数字出版的营利模式,数字出版业务逐渐成为它们的主要收入来源,如在短短十年内,Elsevier初步完成了从一个“传统出版商”过渡到“信息解决方案提供商”的转型,目前公司80%的收入来自于数字化产品[Elsevier(科技部)战略总监 Alexander van Boetzelaer(包亚峰)在“第五届中国科技期刊发展论坛”上的发言。2009-10-27,上海]。

中国科协所属期刊处在管理体制复杂、出版资源分散和市场主体缺失的状态中^[10],自建网站数量众多、上网形式多样从一个侧面充分印证了这一点^[5,6]。与西方发达国家相比,中国科协所属期刊的数字化、网络化总体程度较低。从出版流程看,中国科协所属期刊中仅有27%的出版业务流程实现了网络化改造,许多期刊仍然因循传统的运作方式^[6]。从出版产品形式看,除80%以上的中国科协期刊依托CNKI、万方、维普等几家商业数据库(非科技文献首次出版者)将科技期刊以二次文献集成的方式作数字化发布并提供以数据库为基础的多样性的科技信息服务外,大部分期刊出版者(科技文献首次出版者)能够为用户提供的仍然是以传统科技出版产品为主,或通过自建网站提供零散分散、大部分质量不高的科技信息服务。期刊数字出版业务的主体仍然是IT业,广大期刊出版单位更多的是不自觉地、被动地被卷入到数字化的浪潮之中,他们在目前的数字出版业务中所发挥的作用非常有限^[11]。由于多为分散办刊,处于被动的局面中,没有与几大数据库出版商讨价还价的实力,期刊数字化收入在他们目前的收入中所占比例仍然很小,完全无法弥补因数字化而造成的发行收入减少的缺口。

目前,我国科技期刊数字化发展在CNKI、万方、维普等商业数据库的推动下已经具备了相当的规模,但三个网上的全文普遍滞后于印刷版,未体现科技期刊传播科技信息的高效性,即从某种意义上说这种全文上网方式称不上是期刊网

络版的在线发布,而只是期刊论文后印本的网络仓储,读者对文献的访问大多采用图书馆集团购买或按篇付费形式,限制了科技文献的广泛传播和高效使用。三个数据库为期刊提供的主页模板千篇一面,无法推荐特色栏目,三个数据库至今尚未开发出实用的、被编辑部普遍接受的期刊稿件采编系统,以实现编辑流程的网络化。

综上,在20世纪末到21世纪初信息技术突飞猛进发展的十几年来,中国科协所属期刊,乃至我国的科技期刊,作为科技文献的首次出版者在信息化带来的产业升级中没有集中的、大规模的投入,未与发达国家同步完成科技出版业的产业升级,使我国科技期刊与发达国家相比,在数字化、网络化方面的发展规模和技术水平均明显滞后,缺乏国际竞争力。广大期刊主办单位,通过何种姿态和方式,顺应时代发展,完成由传统的期刊运作向互联网信息提供商和服务商的转变,是一个值得广泛关注和认真思考的问题。

致谢 在资料收集和论文撰写过程中,得到中国科协学会学术部沈爱民、杨文志、王晓彬、张利军、张建国等的大力帮助,在此深表谢忱。

参考文献

- 1 赵文义,王磊,宋庆国等. 学术期刊的网络传播模式分析. 中国科技期刊研究, 2008, 19(1): 80-82
- 2 CNKI中国知网, [2009-06-15]. <http://www.cnki.net/>
- 3 万方数据资源系统-中国数字化期刊群, [2009-06-15]. <http://www.wanfangdata.com.cn/>
- 4 重庆维普资讯有限公司, [2009-06-15]. <http://www.cqvip.com/>
- 5 程维红,任胜利,王应宽等. 中国科协科技期刊网络化出版现状. 中国科技期刊研究, 2008, 19(3): 359-366
- 6 程维红,任胜利,王应宽等. 第三章 中国科协科技期刊数字化和网络化建设. 见:中国科学技术协会主编. 中国科协科技期刊发展报告(2010). 北京:中国科学技术出版社, 2010
- 7 中国引文数据库, [2010-01-05]. <http://ref.cnki.net/knsref/>
- 8 方卿主编. 科技出版国际竞争力研究. 武汉:武汉大学出版社, 2008:37-38
- 9 程维红,任胜利,王应宽等. 国外科技期刊的在线出版——基于国际性出版商和知名科技社团网络平台的分析. 中国科技期刊研究, 2008, 19(6): 948-953
- 10 刘泽林,张品纯,孙际元等. 第二章 中国科协科技期刊出版体制现状与发展. 见:中国科协科技期刊发展报告(2009). 北京:中国科学技术出版社, 2009
- 11 聂震宁. 出版业与数字出版的距离有多远. 见:黄先蓉,罗紫初主编. 数字出版与网络教育. 北京:高等教育出版社, 2009:5