

关于中国科技期刊出版及其数据库建设的思考

李遵先 翁佩萱

收稿日期:2009-05-05
修回日期:2009-08-03

华南师范大学学报编辑部,510631 广州中山大道西 55 号, E-mail: xb04@scnu.edu.cn

摘要 通过中、外科技期刊出版及其数据库建设的模式的对比,剖析在中国科技学术期刊出版中其数据库建设存在的问题,提出关于中文科技期刊数据库合理发展的思考和建议。

关键词 科技期刊 数据库 收录 发展思路

2008 年是中国改革开放 30 周年,30 年前,党的十一届三中全会作出了把党和国家的工作重心转移到经济建设上来的决策,开启了中国改革开放的历史新时期。30 年来,中国取得了世人瞩目的巨大成就,中国的现代科技得到迅速的发展^[1],中国科技学术期刊事业在这 30 年中也有了长足的发展。作为中国科技成就发布的载体之一,科技期刊的数目从 1978 年的约 2800 份^[2]发展到 2007 年的 4713 份^[3],中国学术文献出版网络出版总库(中国知网)收录 1978 年科技论文 61 885 条,而 2008 年收录科技论文 2 259 027 条,是 1978 年的 36.5 倍。自 1989 年由重庆维普资讯有限公司创建以软盘形式发行的中文科技期刊篇名数据库至今,我国的文摘数据库已有十余家,二次科技文献也已达到十余家^[4]。我国的科技学术期刊出版和科技学术论文数据库经历了一个腾飞的过程,成绩喜人。但是,欣喜之余,我们也看到存在的一些与世界发达国家科技期刊出版界的差别和问题。或许,经过这么多年的发展,是该做一下问题的梳理的时候了。在本文中,笔者尝试在中国科技期刊出版和数据库收录方面与国际先进国家的模式做一些比较和反思,旨在与同行们交流,探讨今后我国科技学术期刊出版与数据库发展思路。

1 中外科技学术期刊出版及其数据库的模式对比

到目前为止,西方先进国家的科技学术期刊的出版一般有下述三种主要的模式。

第一种模式是编辑部与出版社合作的模式,首先由主办单位的期刊编辑部负责收稿、稿件审理和修改、决定稿件录用与否等工作。一旦稿件定稿,则以后的排版、版权认定、校对、出版、(网络)单行本寄送、纸本期刊发行、数据库收录发

行等工作全部由出版社负责,包括与作者的联系和互动。一般来说,一个出版社与很多期刊社或编辑部签订合作协议,其中网络投稿软件的开发通常由出版社统一负责开发,交给期刊社或编辑部使用,因此通常可以看到,由同一家出版社出版的不同期刊有着类似的投稿系统和主页,它们都在出版社的主页下有链接。

第二种模式是学术团体式的出版形式,有些大的学术团体或研究所,同时主办好几份期刊,学术团体同时兼备出版社的资格和功能,专门出版相应的科技期刊,这些期刊通常由学术团体负责审稿和出版的全过程,例如,美国电气电子工程师协会(IEEE)、美国化学会(ACS)、美国数学会(AMS)、英国皇家学会(RS)和英国皇家化学会(RSC)等都有自己的科技期刊及数据库。

第三种是期刊社把所有的流程都包揽了,这种模式通常是一些电子版期刊,他们实行网络化出版和 Open Access(OA)的模式,让读者免费下载浏览。

与上述三种模式相对应的数据库通常是:第一种模式的全文数据库就在出版社,这个数据库是出版社的商业盈利资源,如 SpringerLink 数据库、Elsevier Science Direct 数据库;第二种模式的全文数据库在学术团体如学会,虽然学会是非营利机构,但是数据库也是盈利资源;第三种模式的全文数据库就在期刊社,通常主要以交流被引用为目的,不以营利为目的。目前,也有一些出版社旗下的部分期刊实行 OA 模式,例如, Handawi Publishing Corporation 旗下就有不少实行 OA 的期刊。

至于二次文摘数据库,通常与学会、信息研究所(公司)、科技情报研究所等研究单位相联系在一起,如美国数学会(AMS)的 MR、美国化学文摘服务社(CAS)的 CA、美国工程

信息公司(Engineering Information Inc.)的EI、美国科学研究所(ISI)的SCI和ISTP、德国数学文摘社的ZBL MATH、英国机电工程师学会(IEE)的INSPEC、俄罗斯全俄科学技术信息研究所(VINITI)的《文摘杂志》AJ。这类数据库主要提供一次文献的摘要、评论、索引信息、检索,旨在为科学研究者和工程技术人员提供专业、实用的数据、知识等信息服务,支持最新的研究。

科技期刊数据库的发展是伴随科学技术的发展而发展的,数据库产品的传播由纸本、磁带、软盘、光盘等发展到今天的网络数字化传播,西方发达国家的学术期刊出版和数据库已经形成了成熟的模式,严格遵循版权保护的基本原则,除了OA期刊,全文数据库和二次文献数据库基本是商业操作的模式。

长期以来,中国的科技学术期刊的出版基本上停留在编辑部负责审稿出版的全过程(各自为政)的阶段,当然有些编辑部把排版印刷工作交给一些专业的排版公司和印刷厂,发行方面一般是邮局发行、代办发行、自办发行3种形式。目前,已经有一些英文版的中国科技学术期刊与国外的出版社合作,走上上述的第一种模式的路子,但是这只占很小的比例,因为中文期刊毕竟占绝对多数。

我国的科技期刊数据库始于1989年由重庆维普资讯有限公司(其前身为中国科技情报所重庆分所数据库研究中心)创建的以软盘形式发行的中文科技期刊篇名数据库,后发展成为网路版的中文期刊数据库(简称维普)^[5]。维普与中国知网(CNKI)、万方数据化期刊合称为3大中文电子期刊数据库。其中中国知网是中国知识基础工程(National Knowledge Infrastructure, CNKI)中的一个重要组成部分,由1995年立项的清华大学《中国学术期刊(光盘版)》发展而来,是目前国内最大的网上期刊信息资源之一。万方数据化期刊是科技部“九五”重点科技攻关项目“科技期刊网络服务系统”的一个组成部分,由中国科技信息研究所主办,1990年建成并出版了涵盖1.6万家企业的中国第一家商业化商情类数据库。3家中文科技期刊数据库的主体都是针对科技期刊的全文收录,其操作模式基本上是由期刊编辑部提供期刊的纸本或电子资源,再制作成数据库的资源。与国外文献数据库载体的发展模式相同,我国的文献数据库同样经历了从磁带—光盘—网络的发展过程,且目前网络型数据库已成为中文科技期刊数据库的主要形式。

与国外科技期刊数据库的发展不同,中国的国情决定了我国的全文数据库与西方发达国家的学术期刊全文数据库有明显的不同。

首先,中文科技期刊数据库与其所收录的科技期刊之间不存在从属关系,这是与外文科技期刊数据库的巨大差别。上面提到的国外科技期刊第一种和第二种出版模式,它们与科技期刊数据库有明显的从属关系和统一的操作模式。但我国的科技期刊数据库的服务商是由各情报所、研究院所和高等院校衍生出的公司,数据库的内容是由各科技期刊社(编辑部)所出版的科技期刊,两者间不存在从属关系。公司要以经济效益为首要目的。而数据库的建设要借助于各家独立于这些公司的科技期刊社(编辑部),其协调过程的困难性可想而知。

其次,我国科技期刊数据库所收录的期刊具有很大的重复率。无序竞争、政企不分、以权代法是造成这一现象的重要原因。另一方面,全社会知识产权意识不强,没有形成尊重知识的社会氛围,是这种现象产生的社会根源。中文科技数据库建设初期,由于各数据库的建设单位多为情报所、研究院所和高等学校等国家事业单位,缺乏市场经济的体制规范和竞争意识,各个数据库并没有在信息资源建设上争取各家科技期刊社(编辑部)的独家授权。而各科技期刊社(编辑部)对被哪家数据库的收录所产生的知识产权等方面的问题并不太在意,任由各大数据库对期刊文章进行收录、整理和分类,甚至以进入数据库的多寡作为期刊的评价指标,似乎进入越多数据库就越“高档”和光荣。在这样的背景下,形成了中文科技期刊数据库所自身特有的一个不同于外文科技期刊数据库的怪现象,这就是数据库所收录期刊的高重复率,尤其以三大中文科技期刊数据库最为明显(见表1,数据来源于文献[6])。

表1 2005年3大中文科技期刊重复收录情况表(量级:种数)

| 数据库种类 | 重复量 |
|-----------------------|------|
| CNKI VS 维普 | 5818 |
| CNKI VS 万方数字化期刊 | 2863 |
| 维普 VS 万方数字化期刊 | 2722 |
| CNKI VS 维普 VS 万方数字化期刊 | 2463 |

表1表明三大中文科技期刊数据库虽然所收录期刊不完全相同但存在相当高的重复率。这一方面造成了资源的大量浪费;一方面也为使用者带来巨大的不便。由于各数据库均为收费数据库,各图书馆在购买时往往需要购买2到3种,而其中一大部分都是相互重复的。针对这种情况,一些文章试图从各个数据库的异同进行比较来对图书馆的购买进行指导^[6,7]。

2 中文科技学术期刊数据库发展现状和挑战

由于中文科技期刊数据库必然走向商业化之性质,数据库商之间的竞争其实从各自成立的那天起就已经存在。只是数据库建设之初,由上所述各种原因,表现得不甚明显。进入 21 世纪以来,随着数据库产品在市场中所展现的巨大盈利能力和产品开发的逐渐完善,各数据库商之间的竞争日益激烈。尤其在数据库内容的争夺上表现得尤其明显。2000 年某数据库公司被多家期刊杂志社因著作权问题起诉,这当然是由各期刊社日益增加的版权意识所导致的必然结果,但更深层次的是数据库商之间的竞争结果。

这一时期,各家中文科技期刊数据库服务商为了扩大规模从而取得更大的经济效益,开始了针对科技期刊的“争夺战”。这个阶段的特点主要有几方面:

首先,以内容争夺为主,通过内容争夺实现经济效益最大化。由于数据库技术的普遍性和成熟性,数据库服务商所提供的服务业趋同化,即通过对所收录的各种科技期刊整理成数据库(网络版和光盘版)来提供文献检索和阅读的服务。因此,数据库的内容建设变得重要起来,从而形成了数据库服务商对信息资源——科技期刊的全文信息这一资源的争夺,而且各数据库商更加倾向于对科技期刊的独家收录。2008 年下半年,多家科技期刊社(编辑部)收到了某数据库商关于独家收录的协议(征求意见稿),竞争的激烈性由此可见一斑。目前,部分期刊社已实行独家授权的方式,这无疑会调整现有产业链中的关系,且迫使数据库商和刊社加强合作,从期刊内容的单纯数字化和网络发布向深层次的网络出版和信息服务拓展^[8]。科技期刊全文信息资源独家授权的争夺将是这段时期最显著的特征。

其次,过程中更重视著作权问题的处理。各个数据库在建设之初,我国关于数据库的版权问题尚无相应的法律法规,使得数据库商在信息资源建设过程中缺乏法律观念,忽视甚至侵害著作权人的著作权。为日后科技期刊数据库的发展带来了诸多隐患。随着我国法治化进程的加快,各家科技期刊著作权意识不断增强,且对科技期刊数据库重复收录的利弊有了更加深刻的认识。各数据库服务商也更加重视其所收录期刊文章的著作权问题。因此,在竞争过程不论是数据库商还是各种科技期刊都更加重视法律法规,通过运用法律法规来维护自身权益。

再次,数据库商以合作协商的姿态与各科技期刊社(编辑部)进行沟通。在尊重科技期刊社(编辑部)的意愿上,本

着互利互惠、合作共赢的精神争取各科技期刊的全文收录。例如,中国知网将全年的财务收支情况公开发给各科技期刊社(编辑部),并通过提供各种有益信息,提供各种服务(如召开编辑培训会议)等方式来与各科技期刊社(编辑部)维护相互之间的关系,使各社(编辑部)对其增进了解、增进感情,赢取各社(编辑部)的信任和支持。

这一时期,各数据库商把主要精力放在数据库的内容建设上,通过各种方式争取更多科技期刊的收录直至独家收录。由于科技期刊社(编辑部)与数据库商的非从属关系,这个过程将是漫长、曲折的。实际上,早期一些英文版的科技学术期刊曾经被国内的这几家全文数据库所收录,随着它们投于 Springer 或 Elsevier 出版社麾下,它们所属的论文已经全部撤出中文数据库。那么,可以想象得到,如果要做到像国外一样,全部科技期刊独家收录,中国的全文数据库将经历一场痛苦的重组。

中文科技期刊数据库的发展也面临着诸多挑战。

首先,是来自国外科技期刊数据库的挑战。虽然随着我国科技实力的不断增强和科研人员数量的不断增多,各科技期刊的稿件资源比较充沛。但由于国外科技期刊影响力大,大部分国内科研人员的优秀稿件都被国外期刊吸引过去,这一现实直接影响了国内科技期刊的质量,进而间接影响了中文科技期刊数据库的质量。

其次,一些先进的出版理念也给科技期刊数据库甚至整个出版行业带来挑战。如 OA 出版。OA 出版理念是一种新的出版理念,并在实践中被积极探索。OA 是 Open Access 的缩写,国内翻译为开放存取或开放获取。其倡导为促进学术交流而进行文献的免费获取。OA 文献的出版是通过一些官方机构及非营利组织建立的开放存取网站或其他形式的出版物,出版成本一般由作者、政府及出版机构等提供。目前,已形成一些有较大规模的网络 OA 出版项目,如美国学术出版与资源联盟项目(SPARC)、科学公共图书馆项目(PLoS)等。国内也有一部分高校学报实行免费全文下载,教育部科技发展中心所建立的科技论文在线网站也已颇具规模。这种新的出版模式无疑将挑战现有的科技期刊数据库的利益链条。

3 关于中文科技期刊全文数据库发展的思考

数据库资源的争夺自始至终无论在国内国外都在进行着,但国外的竞争多是在稿件(上游)资源的争夺上,而国内则是在期刊(下游)资源收录的争夺上。因此,中文科技期刊数据库的各服务商必须在合理处理与各科技期刊社(编辑

部)的关系上多下功夫。尤其重要的是,各数据库服务商和各科技期刊社(编辑部)要有明确定位。

首先,数据库服务商定位。

针对中文科技期刊数据库的特点,数据库服务商要合理定位,结合自身优势,发掘数据库自身特色。一定要认识到数据库向精专细化发展的大趋势,例如做好某些学科期刊的数据库,专门收录某一学科的期刊。另外,要密切联系各中文科技期刊社(编辑部),争取收录更多的优秀刊物;同时,要把握市场规律,积极探索合理的市场化道路。积极开拓国内外市场,扩大数据库在国内外市场的影响力,这不仅能为数据库商带来更多的利润,而且也为中文科技期刊扩大影响力作出贡献,这将是中文科技期刊社(编辑部)在是否接受被该数据库独家收录时所考虑的一个重要因素。另外,必须明确把独家收录作为数据库收录的原则,不能一味迁就期刊社,必要时必须摒弃违反原则的期刊入库。实际上,各全文数据库必须明确意识到,放弃部分期刊,是势在必行的事情,除非有哪家数据库全面倒闭,不然你总得有所舍弃,让所有的数据库都能生存下去。

其次,中文科技期刊自身定位。

中文科技期刊的自身定位是中文科技期刊数据库发展的另一个关键环节。各期刊必须充分了解全文数据库收录的原则和版权的重要性,要形成关于数据库独家收录的统一认识。要结合本刊的办刊宗旨,从提高刊物质量和影响力的角度进行思考。不要盲目服从于被动的收录,应在认识到数据库收录重要性的基础上与各个数据库进行积极沟通,从而解决刊物是否要被收录、需要被具有何种特色的数据库收录以及收录方式等问题。

再次,政府在加强针对数据库发展的法制建设,规范数据库的发展,为数据库发展提供系列、持续的指导意见也是必不可少的。

政府有关部门应该积极探讨,协调各方利益关系,使我国的中文科技期刊数据库形成科学良性的发展和有序合理的竞争,避免资源的浪费。其中引入市场机制和制定相应法律法规是重要的手段。可喜的是,数据库发展这种无序竞争、政企不分、以权代法所带来的不良社会影响,已引起国家有关部门的高度重视,并相继出台了一些政策法规。强调“激励创造、有效运用、依法保护、科学管理”。引导各数据库商遵纪守法,有序竞争,自由重组。实际上,据我们上面的分析,数据库的调整、合并、重组,应该是不可避免的。

值得提到的是;中国科技论文的各大数据库已经意识到

上述的问题,一些措施正在逐步酝酿或实施中,例如合并调整,收录期刊精专细化等等。

由于中文科技期刊的主要读者是中文读者,国际稿件和国际读者的匮乏决定了国际化之路的狭窄,因此中文科技期刊数据库的订户主要是国内的图书馆和科研单位。目前看来,我国的科技期刊还不太可能像国外一样,由出版社出版,这种各自为政的局面还会维持较长的时间。这里笔者认为有两个原因:一是期刊出版经费的各自为政所决定;二是国内原有的以出版书籍为主的科技类出版社目前是不太可能开展这种业务的,因为我国科技期刊全文数据库发展至今已经颇具规模,不可能再由出版社另起炉灶,加入原本已经激烈竞争的行列。那么除非由目前的全文数据库承担起出版社的责任,不然很难改变原来的格局和态势。笔者这里建议,全文数据库商承担起出版社全部或部分责任,也就是说,把数据库利润用来发展这一块业务,做到多功能化,或许这是中国科技期刊全文数据库做到独家收录的发展之路。这也或许可以解决部分期刊(例如高校学报)对主办高校的依赖性和综合性问题,顺利向专业性学术期刊转型的一个途径。

我们认为,过去的20年,我国科技走向世界,科技期刊全文数据库从无到有、从小到大,已经取得了世人瞩目的成绩;展望未来,各个中文科技期刊数据库必将借助国家进入世界科技强国之机和网络技术之优势,在调整创新中发展壮大。我们相信,当我们庆祝改革开放50周年的时候,我们应该可以看到我国的科技期刊全文数据库已经成为各具特色、在国内外有重要影响的科技文化产业。

参考文献

- 1 白春礼. 中科院三十年:中国现代科技发展的集中写照. 人民网, (2008-12-30) [2009-04-23]. <http://30:people.com.cn/GB/134895/8600302.html>
- 2 李桂丽,周桂莲,南红梅. 中国科技“核心期刊”评定问题研究. 西北农林科技大学学报(社会科学版),2008,(6):160-162
- 3 新闻出版总署. 新闻出版总署出版公报. 新闻出版总署网站(2008-08-01) [2009-4-23]. <http://www.gapp.gov.cn/cms/html/21/490/200808/459129.html>
- 4 罗金增. 从 Ei Village2 看我国二次文献数据库的发展趋向. 数字图书馆论坛,2006,(5):53-57
- 5 唐曙南. 科技期刊数字化网络化问题研究. 上海:华东师范大学,2001
- 6 姜琳. 合理选择中文期刊数据库. 科技情报开发与经济,2006,(6):7-9
- 7 顾立茵. 网上三大中文期刊数据库特点功能比较. 现代情报,2005,(1):135-137
- 8 曾建勋,宋培元. 我国科技期刊网络化发展动因与趋向. 编辑学报,2008,(4):283-285