

●陈传夫

信息资源公共获取的社会价值 与国际研究动向*

摘 要 信息资源公共获取是指信息被便捷地、免费或通过合理付费方式被公众无障碍获取。目前国际研究偏重于:信息资源公共获取的社会与法律问题,信息资源公共获取差异研究,信息资源开放存取研究。参考文献 15。

关键词 信息资源 公共获取 社会价值 研究动向

分类号 G253

ABSTRACT Public access to information resources is the access by the public easily, freely or through reasonable charged services without barriers. The researches on this topic are focused on the social and legal issues, the differences in the public access to information resources and the open access to information resources. 15 refs.

KEY WORDS Information resources. Public access. Social values. Research trend.

CLASS NUMBER G253

1 信息资源公共获取的社会价值

1.1 推进信息公共获取,是发挥信息资源效益的需要

信息资源的开发利用是国家信息化建设取得实效的关键,也是衡量国家信息化水平的一个重要标志。随着经济的发展和社会的进步,信息资源的这种重要性将更加突出。资源短缺是全球经济发展必须面对的一个重大问题。要保持我国经济持续快速健康发展,必须把开发利用信息资源摆在重要战略位置。

随着信息化速度的加快,全球信息量急剧增长。仅 Alta Vista 数据库就达 3 亿网页,并以每天 300 万网页的速度增长,全球网络信息资源达 80 亿网页以上。加利福尼亚大学伯克利分校信息管理及系统学院莱曼教授领导的研究人员发现,全球每年新增由纸张、胶片以及磁、光存储介质所记录的信息生产总量达到 5 万亿兆字节^[1]。

我国数字化资源增长很快。截止到 2004 年 2 月 1 日,全国在线数据库的总量达 169867 个。这些数据库分别为企业(占全部在线数据库的 54.3%)、商业网站、其他非营利机构网站、个人网站、教育科研机构网站、政府网站等所有。中国数字图书馆工程被列为国家“十五”计划重点建设项目,投资达到 11.42 亿元。中国知识基础工程已经建设成具有 1000 万篇全文文献的超大型文献数据库,每天追加更新 3000 篇原文以

上。中国期刊全文数据库积累全文文献 800 万篇,题录 1500 余万条,收集国内公开出版的 6100 种核心期刊与专业特色期刊的全文,数据完整性达到 98%。

2002 年,我国公共图书馆累计总藏书量达 42628.3 万册^[2],普通高校图书馆有一般图书 72551.898 万册,电子图书 1809.7765 万片。中国科学院图书馆系统拥有图书 771 万册,期刊 2095 万册,其他各类文献共 855 万余册(件),院内开通或试用的电子全文期刊约 11770 余种^[3]。中国社会科学院图书馆系统共有藏书 494.2258 万册^[4]。2003 年,公共图书馆总流通 21440 万人次,书刊文献外借 18775 万册次,为读者举办各种活动 92701 次,参加 2432 万人次^[5]。但与欧美发达国家图书馆相比,我国图书馆资源还存在着利用率低,公众获取较困难,大量资源被闲置、浪费的情况。2001 年,美国公共图书馆图书流通册次为 178998 万册次,人均 6.5 册次,英国 2000-2001 年的流通为 40561 万册次,人均 6.79 册次,我国 2000 年公共图书馆图书流通册次为 16913 万册次,人均只有 0.13 册次。人均册次英国是我国的 50 倍。图书馆资源公共获取具有很大的潜力。因此专家认为,应“拓展社会的公共信息空间”^[6]。

2000 年 6 月,由科技部等单位组建的国家科技图书文献中心是一个基于网络环境的虚拟的科技文献信息服务机构,每年订购外文期刊 15500 种以上,占全国

* 本文系教育部优秀青年资助计划项目的研究成果之一。

采购品种数的60%多;外文会议录2500多种;中文期刊、会议录、学位论文基本齐全。它覆盖数学、物理、化学、地球科学、生命科学、工程技术各学科、林学、农学、医学等国家科技创新急需的主要领域。近20年来,我国投入巨资开展了大规模的地质、气象、海洋、水文、环境、地震调查和观测,积累了海量科学数据,先后建设了5000多个科学数据库^[7]。“九五”期间,国家在资源环境领域直接投入探测、调查的费用就不低于200亿元^[8]。同期,为实施国家级科技计划的经费投入也接近2000亿元,其中约有30%~50%用于科学数据的采集与整理^[9]。有关方面估计我国科学数据资源总量占世界的10%,每年发表的科技文献量约占全球的1/8。但是,数据的共享所产生的效益则只占全世界的0.1%,大量资源处于闲置状态,科学数据获取方面的障碍造成国家在文献资源领域投资的效益得不到发挥,已经妨碍了科学创新活动。

根据英国 Southampton 大学的 Stevan Harnad 观测,全球每年200万篇学术性文章包含在2万种期刊中(eprintblog, Sept. 25, 2002)。根据中国科技信息研究所的分析,2003年全国发表在中文期刊上的学术论文达到24万篇,被国际三大检索刊物SCI、EI和ISTP收录的共93352篇,占世界论文总数1834994篇的5.1%。然而许多发表的学术资源因为期刊发行量的限制,不能被公共获取。

1.2 推进信息公共获取,是提高政府信息化效益的需要

一国政府往往是最大的信息生产者也是最大的信息拥有者。政府信息具有权威性、行政性、系统性等特点。政府信息与公众的政治、文化、经济生活关系十分密切。政府信息源是人们全面考察、评价社会,从事社会政治、经济、文化等活动不可或缺的国家资源。我国80%以上的信息资源掌握在政府手中。但目前我国的政府信息资源只有20%是公开的。这种封闭、闲置或半封闭、半闲置状态,不仅会造成信息资源的极大浪费,还会致使资源配置失调、经济活动成本增加等社会问题^[10]。

信息公开与政府透明是满足社会公众和公民知情权的需要。信息公开首先是政务信息公开。不仅政务信息,其他信息,如市场信息、环境信息、食品卫生安全信息、传染病流行信息、自然灾害和其他突发事件信息等,都应该通过一定途径和形式公开,让全体社会公众或相应领域的社会公众知晓^[11]。政府信息资源的公共获取对于公众掌握国家政策、了解政府行为具有重要意义,这是政府信息化取得实效的标志。通过信息资

源公共获取和及时的信息反馈,政府也可以在重大决策事项、经济社会发展目标、财政财务收支、组织人事等多方面听取公众意见。将一切行为均置于群众监督之下,这无疑对构建诚信政府起着促进和推动作用。

国家在信息化领域的投资如何才能产生效益,关键是要信息能便捷地、以合理价格或免费为公众所获取。新华社的调查表明,公众从政府信息化中获得的服务很少。而据美国的一份调查报告,93%的被调查政府机构拥有网站,向用户(公民、企业、非营利机构)提供政府文件、选举信息、意见投诉、医疗保健、社会福利、执照申请等公共服务和网上办公信息。3/4被调查个人登录过政府站点,大多数对政府网站的界面和内容感到满意。而我国,各地、各行业、各部门各搞一套,采用的标准也各不相同,形成一个个“信息孤岛”,信息资源的价值体现不出来。推进信息公共获取要求政府依照法定义务提供其拥有的信息资源,减少和避免信息资源的闲置和浪费,促进政府与公众之间联系与交流,促使不同的政府部门、机构之间以及政府部门与社会之间实现资源共享,保障社会公众依据法定权利要求政府提供信息。政府信息有着潜在的巨大的经济价值。及时获取政府信息,可以抓住时机创造物质财富。也有助于社会公众通过公开、公平、合法的渠道及时获取政府信息,最大限度地减少信息阻塞、信息浪费现象和信息寻租现象,消除腐败,减少信息获取的社会成本。

1.3 推进信息公共获取,是缩小数字鸿沟的需要

信息化的推进一方面丰富了社会信息资源,另一方面也带来了信息差距。拥有计算机、网络等资源的人群,正快速拉大与未拥有者在知识取得、财富以及社会地位的距离,形成新的社会不公平现象。1999年美国发表《定义数字鸿沟:在网络中落伍》研究报告指出:在信息社会,个人计算机以及网络等信息工具对于个人的经济成就以及生涯发展具有关键性的影响力,有无计算机以及运用计算机能力的高低,将成为主宰贫富差距的力量。全球每年新增加的信息量中,美国制造的存储信息总量就达到40%,其中印刷信息占30%,磁记录信息占40%^[12]。

在我国,因特网、数据库等信息化基础设施集中在大城市。东部和西部地区、城市与农村之间差距极大。在加速国民经济与社会发展网络化、信息化的同时,必须高度重视地区间以及城乡间的信息差距问题,应当把缩小信息差距、消除知识贫困作为21世纪的重大发展战略,而且应当把缩小信息差距作为网络

化与信息化建设的重要内容^[13]。

在我国有相当一部分人属于弱势群体,例如下岗职工、进城的务工人员。信息已经成为影响我国农村劳动力转移的四大障碍之一^[14]。在信息化时代,公正和权益应当充分体现在不同行业、社会阶层对信息资源、发展机会的公平占有上。这些弱势群体,在获取信息方面也处于弱势地位。应当改善他们的弱势地位。另外,我国在推进信息化过程中,各行业、各部门、各系统往往从自身利益出发,纷纷建起各自的信息资源系统,例如教育系统、公共系统与科学院系统。有些系统(网络)之间不能融合,无法做到资源共享,实际上形成一种“信息割据”的局面。

1.4 信息资源公共获取,是提高研究效率和研究公平的体现

科学数据获取可以提高科研的效率。我国科学家很早就注意到科学数据不能共享给科学发展造成的损失,并呼吁在我国实施科学数据共享。中科院地理科学与资源研究所全球变化信息研究中心主任刘闯博士说:“由于缺乏共享数据库,我们的科研一直在重复着数据采集和整理工作,既浪费了财力、人力,也限制了研究水平的提高。”例如,一套基本气象数据或一套青藏高原基础地理数据的使用费都在20万元以上,“已使我国中等规模科研项目不堪重负”^[15]。

我国持续的科研投入为科技论文的产出创造了条件。2003年财政科技拨款达1000亿元人民币。我国科技论文产出总量很大,目前年发表量达274438篇居世界前列。这些科技论文蕴涵巨大的科技能量需要通过开放渠道为公众获取。但是,印刷型期刊科学交流有发行量小、渠道窄、周期长等天然缺陷,不能满足公共获取的需要。比如教育部系统每年获得国家2/3的科研基金,产出了全国70%的科技论文,需要公共获取科技论文的群体包括教师、研究人员、研究生、本科生、开发人员等。但该系统每年至少有10万篇科技论文找不到发表渠道(仅仅每年毕业需要发表论文的硕士、博士研究生就达40万人)。

由于公共获取渠道不畅,我国科学家的优秀科研成果可能不能为我国同行首先获悉,影响创新效果。世界主要出版商包括Kluwer, Springer, 垄断期刊出版价格,导致“学术交流危机”。据美国研究图书馆协会统计,2003年美国大学图书馆用于购买期刊的费用相对于1986年而言,涨幅高达260%,平均涨幅为215%。中国需要花大量外汇购买自己的成果。有些期刊根本没有订阅或只有少数图书馆订阅。

科技论文的公共获取是研究平等的体现,也是教育平等的体现。通过开放获取,那些没有进入211等平台的大学可以获得更多的交流机会,地方院校可以获得重要大学的成果,偏远地区的师生可以获得更多的交流机会。拓展科技论文发表与获取渠道,使国家科技投入产生实效是实现研究平等的需要。

科技论文在线交流是学术资源公共获取的重要方式,有发表快捷高效、共享面广、发表成本低、开放参与、科学信用、消除学术腐败等优点。值得注意的是,教育部科技发展中心主办的“中国科技论文在线”网站,利用现代信息技术手段,打破传统出版物的概念,免去传统的评审、修改、编辑、印刷等程序,给科研人员提供了一个方便、快捷的交流平台,成为及时发表成果和新观点的有效渠道,这是开放获取的一个很好的尝试。“中国科技论文在线”方向正确,社会效益很好,受到全国学术界的欢迎。该网站集成学术论文资源,促进信息增值,提高了国家科技竞争力,实现了教育部对科技论文的有效管理,使科技评奖、学术监督、成果管理规范化。根据笔者观察,“中国科技论文在线”平均每天发表文章的数量相当一本科技期刊的容量,目前累计访问量超过100万人次,正在发挥很好的社会效益,值得提倡。

2 国际研究动向

2.1 信息资源公共获取的社会与法律问题

美国、加拿大、英国和德国的多个独立研究机构,结合法律、技术与公共利益开展跨学科综合研究,如哈佛大学政府管理学院1990年创办“国家信息化基础设施工程”跨学科论坛,致力于全球范围信息政策的研究与讨论。每年集中于一个国家和主题,特别强调知识产权等政策对信息资源的公共获取的影响,哈佛大学法学院的伯克曼研究所则致力于信息社会问题研究。斯坦福大学因特网与社会中心则主要探讨信息化对公众带来的利益与损害。出版物有日本信息处理开发协会《信息化白皮书》、哈佛大学出版的《法律与公共政策杂志》、《法律与技术杂志》,斯坦福大学出版的《法律与政策评论》,加州大学伯克利分校出版的《技术与法律杂志》等等。斯坦福大学在这些杂志中遴选的前沿文章如《公平信息实践与隐私建设》(Marc Rotenberg, 2001)、《源代码公开案例》(Mathias Strasser, 2001),表明他们的研究工作已经深入到信息公开的具体问题。美国计算机学会、大学协会、国际图书馆协会联合会、美国研究图书馆协会等组织的研究则集中于

数字化信息的合理使用方面。国际图联认为,信息是为所有人的,无论何种形式的信息均应被公共获取,版权不应当成为信息与思想获取的障碍,也不能仅仅为付得起费用的人所获取,国家或国际版权法中关于合理使用的规定应做适当修改以使数字化的信息如同印刷本那样可以被平等地获取,作为电子文献传递服务的一部分,图书馆员对数字化作品的复制与暂时存储不应视为侵犯版权,国际协定应保障跨国信息流的技术控制和合法用户对信息获取的平衡。电子前沿基金会则致力于推动信息的公共获取,其创始人约翰·巴洛在著名的《思想经济》中指出,信息是要流动的,信息不能被占有。

2.2 信息资源公共获取差异研究

早在1995年,美国商务部、经济与统计局、国家电信与信息局就发表了《网络失败:偏远与都市中的信息贫困调查报告》。报告显示在美国信息化的早期,信息资源的公共获取差距很大。1998年的《定义信息鸿沟》研究报告,认为美国公众通过电话、计算机、因特网获取信息的能力急剧增加。2000年发表《朝着数字化参与》报告。2002年2月美国电信与信息管理局和统计局出版《连线国家:美国人正在推进因特网的使用》报告,认为信息公共获取的差距正在缩小。华盛顿大学设立数字鸿沟工程,亚利桑那大学设立数字鸿沟解决方案论坛。1999年9月,美国商务部长主持召开了数字鸿沟高峰会。2000年7月22日,在日本冲绳举行的八国首脑会议发表《全球信息社会冲绳宪章》,强调承诺采取措施,让世界各地的人们都能分享到信息技术进步带来的利益,克服数字鸿沟。2002年8月,因特网与多媒体国际峰会在瑞士召开,讨论全球网络信息鸿沟问题。

2.3 知识产权保护与信息资源公共获取的平衡研究

知识产权代表的是知识创造人的个体,公共获取代表的是社会的公共利益。完善的知识产权制度会鼓励信息传播,反之将阻碍信息的社会利用。如何掌握这个“度”?二者的边界在哪里?这就是政府公共政策应当安排的领域。公共政策就是政府基于公共利益,动用大量资源引导个人和团体行动的规则与政府行为。知识产权的垄断性与公共利益的公共性是一对矛盾。全球知识产权实在法都有向权利人倾斜的趋势,例如TRIPs协议和《WIPO版权条约》、《协调欧盟信息社会版权与相关权特定方面的指令》。1999年,美国公共利益促进科技数据库可获取性研究委员会、国家研究理事会出版《平衡问题:科学与

技术数据库中的私人权利与公共利益》,深入讨论了数字时代数据库中私人权利与公共获取的平衡问题。2000年,美国国家科学院、美国工程院以及由计算机科学、物理、数学家和应用科学家代表组成的小组在国家科学基金会的支持下,发表《数字困境:信息时代的知识产权》的著名报告。欧洲学者F. Willem Grosheide则发表《从用户方看版权:用户获取权》专论(EIPR, 2001)、《为什么全球化知识产权不平衡》(Peter M. Gerhart, 2000)、《施舍馅饼给南半球》(EIPR, 2000. 22(10))。

2.4 信息资源开放存取研究

为解决国际上学术资源交流与共享的障碍,开放存取(open access,简称OA)运动产生了。开放存取也可译为公共获取、开放共享、开放取阅、开放近取、免费阅览等。根据开放社会研究所(Open Society Institute, OSI)布达佩斯开放存取倡议(BOAI)的定义,开放存取是“论文的全文可以免费供用户阅读、下载、拷贝、散发、印刷、搜索或链接”。2003年10月,Max Planck Society起草的“柏林宣言”(Berlin Declaration),OA定义为“全文在线免费获取,完全版本的论著储存在至少一个在线服务器中,以确保免费阅读,不受约束地传播……和长期的数据库式(Archive)储存”。开放存取的一个重要特征是学术资源的开放出版、发布与保存。中国科学院和国家自然科学基金委员会已于2004年签署了“柏林宣言”,表示中国科研机构和科研人员积极支持并推动开放存取活动。OA是一种全新的学术资源交流理论,是在现有的法律框架下建立的学术资源的新型信息传播系统。它打破期刊等出版预定的既定模式,倡导学术信息资源免费使用。它消除学术资源获取的价格障碍(订阅费,许可证费,数据库的浏览费)以及许可障碍(包括版权与许可证限制)。国际上学术资源由传统传播向数字化机构仓储等开放存取方式转移是不可忽视的重要趋势,例如国际上逐步对开放存取的内涵、组织形式达成共识,提出了两种开放存取策略:“自存档”的电子文献档案和开放式网络期刊。这标志着开放存取理念已在科学交流领域取得了较为全面的理解和支持。1998年3月和2001年8月,拉丁美洲和加勒比海地区的健康科学信息代表分别发布《San Jose虚拟健康图书馆宣言》和《Havana宣言》。

很多机构致力于开放存取技术的开发。如:高级学术出版国际联合会(ICAPP)提供期刊管理和出版软件、以SGML标记的内容传递系统,并为那些独自

出版学术期刊的学者提供服务。《开放档案倡议》(OAI)是一个致力于提高开放存取系统间互操作性的标准,它使数据提供者 and 数据服务者分别在 OAI 门户上注册,然后由数据提供者按都柏林元数据元素集标准向数据服务者提供元数据记录,后者将这些分散的元数据信息统一整理并提供相关信息服务。2002 年 2 月,来自 8 个国家的 8 个主要研究图书馆组织联合建立“国际科学交流联盟”(ISCA),旨在将现有的科学交流过程转变为开放存取的方式。

著名的《科学》杂志十分关注该领域。《科学》自 1999 年以来连续发表数十篇研究报告、论文或社论等探讨这一问题。国际顶尖杂志连续发表文章论述这一问题,可见该问题之重要。《科学》2004 年 305 期发表《寻求开放存取》等文指出,开放存取公众资助的研究资料对于科学进展与公共财来说是非常必要的,但目前始终缺乏一致性的政府政策及科学团体来处理有关开放存取资料的事宜。发表的报告又指出,开放存取研究资料可以促进长期的经济获益、利于政府政策的决定以及加快科学发展的脚步。这是该报告的国际研究团队针对经济合作发展组织(OECD)当中具有代表性的资料存取议题持续进行研究的发现。

英国下议院科技委员会在 2004 年 7 月,针对高等教育之机构典藏提出相关报告。报告认为出版形态正由印刷转变为 OA,研究了著作权法、专利等对 OA 的影响。该报告批评英国政府未针对关于科学出版相关的议题做出响应。建议所有机构针对该单位出版的出版品,包含纸本资源、免费提供或线上资源,皆应设置一典藏机制管理维护。所有英国的大学和研究机构都应该提供及时且免费的线上全文论文,由各研究委员会和政府资助的研究也都应当比照办理。为了鼓励免费公开论文的行动,该报告建议各个研究委员会资助研究者发表论文。

美国的众议院议员也建议美国国家卫生研究院所赞助的研究,在发表研究论文的同时将论文副本送到公共医学中心,且公共医学中心会在论文公开发表的 6 个月后将提供论文的原始稿;或是由政府负担研究者发表论文的费用,使公共医学中心可以马上在网络上公开论文原始稿,让大家能在网上免费且实时地取得。

网上信息资源的公共获取引起了国际上关注。2003 年 5 月 10-11 日,ICSU, UNESCO 和 ICST 共同主办了国际数字化科学数据开放存取与公共领域国际学术研讨会(巴黎)。IFLA 执行委员会发表了《学

术文献和研究文献开放存取声明》。2004 年,美国政府发布了《科学自由获取华盛顿原则》,美国国家科学基金会资助了多项研究。美国图书馆协会、研究图书馆协会等机构合作成立了“信息获取联盟”(Information Access Alliance)。为促进学术资源的开放出版与开放存取,ARL 组建了 The Scholarly Publishing and Academic Resources Coalition (SPARC) (www.arl.org/sparc)。澳大利亚发表公共获取项目报告 *OPEN ACCESS: a progress report*, 荷兰建立了国家科学数字仓储 DARE (Digital Academic Repositories)。瑞士 Lund 大学图书馆、丹麦电子研究图书馆、丹麦研究图书馆协会等机构邀请来自美国、英国和加拿大等国的学者、出版商、编辑、图书馆员等,于 2002 年 10 月 22-24 日召开了第一次北欧科学交流会议,探讨科学研究成果的自由流动,并决定以后每两年召开一次。

参考文献

- 1,12 Peter Lyman and Hal R. Varian. How much information? 2003 UC Berkeley, 2004
- 2 文化部计划财务司. 中国文物统计年鉴 2003. 北京: 北京图书馆出版社, 2003
- 3 中国图书馆学会. 中国图书馆年鉴 2003. 北京: 北京图书馆出版社, 2004
- 4 王洛林, 朱锦昌. 中国社会科学院年鉴 2003. 北京: 中国社会科学出版社, 2004
- 5 国家统计局. 中国统计年鉴 2004. 北京: 中国统计出版社, 2004
- 6 于良芝. 拓展社会的公共信息空间——21 世纪中国公共图书馆可持续发展模式. 北京: 科学出版社, 2004
- 7 张景勇. 我国将实施科学数据共享工程. 大众科技报, 2002-12-05
- 8 任建民. 科研杜绝盲人摸象. 人民日报, 2002-12-10
- 9 孙海东. 中国高科技资源“闲”得可怕. 北京晚报, 2003-07-24
- 10 黄梓良, 石彩丽. 加快推进政府信息化的措施. 经济论坛, 2004-09-02
- 11 姜明安. 程序正义与社会治理创新. 在中欧政府高层论坛——“社会治理创新”国际学术研讨会上的发言提纲. 2005
- 13 胡鞍钢. 信息化挑战中国. 北京: 北京大学出版社, 2002
- 14 王娇萍. 农村劳动力转移存在四大障碍. 人民日报, 2000-08-21
- 15 http://news.xinhuanet.com/newscenter/2003-01/06/content_679731.htm

陈传夫 武汉大学信息管理学院教授。通信地址: 武汉。邮编 430072。 (来稿时间: 2006-02-22)