

论我国图书馆规范数据库的合作建立与共享

胡云霞 (浙江传媒学院图书馆 杭州 310018)

摘要 文章介绍了当前我国图书馆规范控制的发展现状,通过对欧美图书馆界联机共享规范数据库先进经验的分析,进而为更好地开展我国图书馆规范工作提出建议。

关键词 规范控制 规范数据库 书目数据 共建共享

On the Cooperative Establishment and Sharing of Library Authority Database in China

Hu Yunxia (Library of Zhejiang Institute of Media and Communications, Hangzhou, 310018)

Abstract The paper makes an account of the status quo of the authority control of library in China. By making a detailed analysis of the advanced experiences of online authority control database possessed by European and American libraries, the paper aims to make some useful suggestions regarding the improvement of the authority control of library in our country.

Keywords authority control, authority control database, online bibliographic data

规范控制是指为确保文献信息目录中著录标目的唯一性和一致性,而建立、维护与应用规范系统的过程。在网络协作和资源共有的时代要求下,如何提高书目数据库质量和书目检索的查准率、查全率,如何科学地提示文献信息资源的内容和形式特征,如何最大限度地实现书目信息资源共建共享已经成为图书馆界编目人员关注的热点问题。近年来,我国图书馆在规范控制工作中取得了可喜的成绩,制作了大量的规范数据,但总的来说,我国图书馆规范控制工作尚处于初期的探索阶段,相比之下,欧美图书馆界在规范控制理论和实践领域均取得了丰硕的成果,在这种情况下,分析和探讨国外成熟的发展经验,为更好地开展我国的规范控制工作提供借鉴。

1 国内图书馆规范控制的发展现状

1.1 大陆地区规范控制的发展

1.1.1 中国国家图书馆规范数据库

国家图书馆于1995年开始制作名称规范数据。2003年3月,该馆的“名称规范组”与“中文书目数据组”合并,在新引进的以色列编目软件ALEPH500系统下,成功地实现规范数据与书目数据的挂接,结束规范数据与书目数据分别建档、长期割裂的局面。这使规范数据和书目数据可以同时建立、同步维护,从而大幅度提高书目的质量,进一步拓展国家图书馆书目发展与完善的空间^[1]。该馆规范数据库包括个人名称、

团体名称与题名名称三个部分,已收录五十万余条记录,每月新增记录六千条,其规范记录基于UNIMARC,统一标目为中文简体形式,汉语拼音以单字为单元,记录格式采用《中国机读规范格式》。

1.1.2 CALIS联合目录规范数据库

CALIS联合目录项目组于2003年9月启动联机规范控制系统项目,开始建设规范数据库,对中、西文文献信息标目实施规范控制。其建设方案是:要求成员馆分工回溯各自的规范记录,走一条把成员馆各自编制的规范档加入到统一的规范档中的道路。2004年,该项目组已完成中文个人名称标目自动生成规范记录的软件需求,并启动了规范系统软件开发项目。目前,已在实验规范数据库中建立46万余条中文个人名称规范记录。其规范记录的统一标目分别采用符合书目特点的中文简体、繁体和外文形式,同时建有方便海外和香港地区使用的汉语拼音形式;记录格式基于UNIMARC规范格式,但提供MARC21规范格式的转出。该项目组的下一步工作是:在联机建立书目记录的同时实现联机规范控制^[2]。

1.2 港台地区规范控制的发展现状

1.2.1 香港中文名称规范数据库(HKCAN)

香港中文名称规范数据库(HKCAN)是香港的岭南大学图书馆、香港中文大学图书馆等7所图书馆采用类似合作编目的方式,旨在建立一个能反映中文著者和机构名称特点的规范数据库。其库将包括人名、

团体名称、会议名称、划一题名及丛编题名等,目前HKCAN资料库已经有超过16万条规范记录,并且有望每年稳定增加2万条新规范记录^[3]。其统一标目采用与《中国人名汉语拼音拼写法》一致的汉语拼音拼写形式,编目规则采用AACR 2R,格式采用MARC 21规范格式^[4]。

1.2.2 台湾地区中文名称权威数据库

1990年,中国台湾汉学研究中心和台湾大学图书馆联合成立了中文名称权威数据库工作组,专门开展台湾地区的中文名称规范工作,该计划后因故中止,于1998年又重新启动。台湾汉学研究中心现已收录中文名称规范记录52万余条,其记录格式采用本馆自有的格式,但可转出符合《中文机读权威格式》的规范记录^[2]。

1.3 国内规范数据库的合作与共享发展现状

近几年来,我国的一些大型图书馆及其联合编目系统,如国家图书馆、中国高等教育文献保障系统管理中心(CALIS)等不仅在大陆开展了合作活动,还将合作范围扩展到港台,两岸三地图书馆界在网络环境下开展中文名称规范的区域合作初见成效。

2003年10月,由中国国家图书馆、中国高等教育文献保障系统管理中心(CALIS)、香港地区大学图书馆协作咨询委员会(JULAC)和台湾汉学研究中心联合成立图书馆合作组织中文名称规范联合协调委员会(CCCNA),并制定“中文名称规范联合协调委员会章程”。该协调委员会的宗旨为“协调规划、标准兼容、联合建设、资源共享”,其工作任务是本着“求大同,存小异”的原则,协调两岸三地有关中文名称规范工作的规范与研究,促进和推动相关地区中文名称规范工作;制定适合本地区使用的、能够和各地区中文名称规范及国际名称规范标准接轨的名称规范标准,并且本着联合建设的原则,首先建立本地区规范文档的合作建设与共享机制,待时机成熟,再考虑多地区规范文档的合作建设与共享。目前,该协调委员会正在不定期的以会议、论坛、E-mail等形式交流、研讨各地区中文名称规范标准的实施进展和国际发展趋势,为推动中文名称规范控制工作的发展发挥了很大的作用^[5]。

从目前进展情况来看,虽然国内一些图书馆已进行了区域化的合作,也在积极探索联合模式,但由于我国图书馆行政管理体制等方面的原因,国内的规范控制与共享受到限制,在全国范围内还未真正实现书目规范控制,还未建立较大规模的且具有权威性的规范数据库,所以国内的规范控制工作尚处于起步的阶段,其合作的范围也不够深入,与国外规范工作的国际化相比还存在着很大的差距。在这种情况下,如何建立共享规范数据库,如何加快我国规范数据文档共建共享的步伐,需要图书馆界同仁的积极探索。

2 欧美图书馆界联机共享规范数据库的先进经验

在国外图书馆界,已有很多合作与共享的经验,最有影响的是NACO与VIAF项目。

2.1 名称规范合作计划(NACO)

美国国会图书馆(LC)在20世纪70年代初期,就相继建立了主题规范和名称规范的计算机处理系统,在80年代初实现了联机规范控制系统,并向其他国家发行规范数据。目前,LC的规范数据库已经实现商品化,特别是名称规范数据,为世界各国文献机构建立自己的名称规范文档提供有益的参考。最有影响力的就是LC倡导建立的NACO,参加NACO的图书馆需依据计划的标准及指引,新增或修改规范记录,从而建立一个国际性规范数据库。目前,NACO的成员馆已遍及世界各地,已有400多所图书馆参与此计划,该数据库已有超过625万条规范记录^[6]。参加NACO成员馆可以得到许多利益但也必须遵守一定的要求:包括遵守统一的标准,必须参加一个联机合作编目系统,按规定每年提交规范记录。为确保在该数据库内的规范记录合乎统一的标准,LC会按新会员的要求提供为期五天的训练课程,培训之后,经过授权才能开始工作。其他非NACO的图书馆,亦可透过商业中介机构,使用该数据库的规范记录,借此达到资源分享的目的,提高标目的统一性及节省各馆得利建档的成本。例如美国的OCLC/WLN的MARS(MARC Record Service),服务范围包括提供最新的及定期更新规范记录,并修改及更新相关的书目记录等。

2.2 虚拟国际规范档(VIAF)

2.2.1 VIAF项目的概况

2003年8月,德国国家图书馆、美国国会图书馆和OCLC联机计算机图书馆中心联合开发了一个个人名称的虚拟国际规范文档VIAF,此项目由OCLC提供软件和技术支持,将德国国家图书馆和美国国会图书馆两家的个人名称规范档合并并提供服务。该项目的关键点是利用规范记录和其对应书目记录的信息开发自动名称匹配算法。VIAF项目的长远目标是连接来自许多国家图书馆和其他权威数据源的规范名称,形成一个共享的个人、团体、会议和地名的全球规范文档服务,使并全世界用户通过Web免费使用^[7]。

VIAF系统特点是:(1)VIAF是一个元数据文档,允许用户从一个国家书目机构个人名称规范档的记录连接到另一个国家规范档的匹配记录;(2)通过专门设计的用户界面提供Web检索;(3)支持多语言和多脚本;(4)使用OAI(Open Archive Initiative),从各机构获取规范档的元数据,继而将元数据添加到共享的服务器以保证文档更新;(5)系统的设计能够链接任意数量的规范档。

2.2.2 VIAF的模式

对于分布式模式检索者将使用一种标准协议,类似于下一代的Z39.50或者SRU/SRW(查询/获取网络服务协议)去检索有多个国家书目机构或地区性权威机构共同参与的相互独立的规范文档。另一种是采用一个中心规范文档同时连接所有其他的规范文档的模式,这样不需要每个国家书目机构与其他所有参加者进行连接。编目员通过检索中心文档就可以检索到同一个实体在世界范围内的全部规范记录,如果在中心文档没有匹配,通过Z39.50就会进行对其他文档的检索。第三种是集中式模式。这种模式最适合记录维护。OAI(Open Archives Initiative)协议通过服务器从每个国家规范文档接收元数据。只要国家规范文档的信息有变化,服务器信息随时就会更新^[8]。

2.2.3 VIAF的理念

VIAF建立在原有AACR名称规范(个人名称、团体名称、地理名称、会议名称)、题名规范(统一题名、丛编名等)的指导思想之上,以AACR标目、参照等内容为依据,以国家级书目机构建立的规范档为基础,面向数字环境提出的规范控制解决方案,这样建立的规范档将不限于某个系统、某局部地区、某个国家,将成为国际性、跨行业、跨专业的规范档,可广泛用于文献、情报、资源提供与服务部门^[7]。

3 国外成熟的发展经验对我国图书馆规范工作的启示

3.1 走合作化道路,促进书目信息资源共建共享

从欧美图书馆发展的经验可以看出,规范控制工作是一项复杂、艰巨的编目基础工作,需要依靠每一个图书馆的共同努力来实现,仅凭一个或少数几个图书馆的力量是有限的,各馆自建规范档不仅会造成浪费,且不利于规范工作标准化的实现。美国国会图书馆NACO的成功实施让我们看到合作建库的许多优点,成员馆的小投入可以获得大收益;VIAF所取得的初步成果向我们证明了共享国际规范文档的好处,不仅可以减少工作量,降低编目成本而且可以在国际范围内简化规范数据的创建和维护,最关键的是使用户以自己偏好的语言、文字和形式来检索信息。因此,我们确信只有联合大多数图书馆的力量,采取合作建立规范数据库的方式,走合作化的道路,才是明智之举,才符合我国规范工作的发展方向。

3.2 加强规范控制规则的标准性和一致性

国内的两大编目机构国家图书馆和CALIS制订了一系列的规则与标准,为我国规范工作的开展奠定了基础。尤其是《中国文献编目规则》中“标目法”一章的推出,《规则》第二部分题名为标目法,实际包括了规范控制的部分内容,明确提到规范控制的作用:“……通过规范控制,实现书目的检索功能与汇集功

能,保证书目记录的查全率和查准率”,这对推动我国书目规范工作的标准化具有重大意义,但是两大系统之间还存在着比较大的差异,例如在标目的选取、标目的形式等方面遵循了不同的编目规则,使两者之间的规范数据不能共享。目前在中文规范编目格式方面只存在一个文化行业标准《中国机读规范格式》,在编目规则方面只有2005年4月出版的《中国文献编目规则》(第二版),却没有权威的国家标准。面对目前的情况,改进这些标准研制与修订,呼吁有关领导部门成立专门机构,组织一些专家持续进行研究,尽快出台与时俱进的国家标准是当前图书馆界迫在眉睫的任务。只有这样才能保证规范数据款目编制的一致性,实现书目数据的共建共享。

3.3 创建全国规范控制中心,统筹全国书目数据的规范控制工作

NACO是在美国国会图书馆的倡导下进行的,而VIAF是由OCLC提供软件和技术支持,连接德国国家图书馆和美国国会图书馆的规范文档。借鉴欧美图书馆的先进经验,国内在合作建立规范数据库的时候,首先应该建立一个规范控制中心,指导各成员馆的规范控制工作。笔者认为应以国家图书馆为龙头,成立全国规范控制中心,统筹全国书目数据的规范控制工作,加大规范控制的研究力度,建设权威的规范数据库。在此基础上,国家图书馆要大力鼓励国内其他文献机构参与规范数据库的建设,打破行业界限,采取联合编制和维护规范数据的管理模式。该中心主要的职责是:组织专家学者开展规范工作的理论和实践研究,编制规范文档的规则,并在成员馆之间推行和实施,提供相关的技术和软件支持,严格校验各成员馆编目数据的质量,对各成员馆的规范工作进行指导,定期或不定期地对成员馆开展业务培训等。

3.4 在项目成员单位之间建立功能完善的书目规范联合检索平台

从NACO、VIAF所取得的成效来看,在各成员馆单位之间构建统一的书目规范联合检索平台是建设大规模合作规范数据库,实现全球书目资源共享的关键途径。构建的联合检索平台实质上是基于Web方式的分布异构统一检索平台,通过对各成员馆单位各自的规范数据库的高度、有效整合,实现跨平台的无缝链接,使编目人员用一个检索指令在一个统一的界面上一次查得所需其他成员馆单位的书目规范信息。可见,建立统一的检索平台对于我国开展书目规范工作是非常有意义的,不仅大大节省编目人员在信息查询时间和数据处理时间,提高对某一书目规范信息的检全率和检准率,而且能够把分布在不同成员馆单位的规范数据库联合起来,是使信息资源从分散独立走向互联共享,进一步为书目规范信息数据源的合作共建创造有利条件。

在这个方面,HKCAN在2003年做了初步的尝试,

所开发的一站式名称规范搜索引擎能够实现大陆、台湾、香港多家机构名称规范数据库的彼此互检。但这一检索工具操作尚不便利、功能有待进一步完备(比如当不同成员单位对同一实体采用的规范标目形式不一致时,如何通过反馈单参照实现二次检索的信息汇集问题)^[9]。今后各成馆单位应加强合作,在联合构建检索平台中应该不断加强和完善检索平台的整体功能,还要充分考虑系统采用的标准是否与国际标准接轨、与其他系统的兼容情况、资源配置是否方便快捷、系统响应速度、系统运行平台情况等,促进规范信息资源共享的渠道更加畅通和便利。

3.5 经常性的开展规范数据库的维护工作

维护规范数据库指的是对规范数据库的更新和管理,它是规范工作中的核心内容。随着文献编目工作的开展,编目标引规则不断更新,新的规范记录日益增加,已有的规范记录及关系也常发生各种变化。由于规范记录中的信息是动态的,因而对于规范数据库的更新和管理,是一项繁重的工作。只有坚持对规范数据库的经常、及时、持续的维护,才能体现其生命力,才能提高其整体功能。因此,对于合作建立共享规范数据库的维护是一项长期而又艰巨的工作,需要联合所有成员馆的力量,共同参与数据库的维护中来。

参考文献

- [1] 王洋. 规范控制与国家图书馆书目数据的可持续发展[J]. 国家图书馆学刊, 2004(2): 18-23.
- [2] 孙更新. 我图书馆规范工作的现状及发展对策[EB/OL]. [2009-09-03]. http://www.nlc.gov.cn/yjfw/gtcbw_ywytlbm06ungx.htm.
- [3] 第三次中文文献资源共建共享合作会议[EB/OL]. [2009-09-03]. <http://www.library.gov.mo/cddcc/paper09>.
- [4] 谢琴芳. CALIS中文名称规范数据库建设方案及实施进展[J]. 新世纪图书馆, 2005(1): 3-6.
- [5] 王洋. FRAR与中文名称规范控制[EB/OL]. [2009-09-03]. http://www.nlc.gov.cn/yjfw/gtcbw_ywytlbm06wangy.htm.
- [6] 付蔚, 王海兰. 从FRBR到FRAR——兼述中文文献资源的规范控制工作[J]. 大学图书馆学报, 2007(2): 97-100.
- [7] Ed. O'Neill. Virtual International Authority File[EB/OL]. [2009-09-03]. <http://www.nla.gov.au/lis/stdnrcs/grps/acoc/viaf2005.ppt>.
- [8] Rick Bennett, Christina Hengel-Dittrich, Barbara B. Tillert 著, 顾犇, 译. 虚拟国际规范文档[J]. 国家图书馆学刊, 2006(4): 87-90.
- [9] 曹宁. 2005年中文名称规范数据库进展情况报[EB/OL]. [2009-09-03]. <http://www.lib.hku.hk/ccrcd/2-1.ppt>.

[作者简介] 胡云霞, 女, 1980年生, 浙江传媒学院图书馆馆员。
收稿日期: 2009-09-25

出版动态

《中国数字图书馆宏观管理研究》出版

由田国良主编的国家社会科学基金重点课题结项成果《中国数字图书馆宏观管理研究》一书于2009年9月由湘潭大学出版社出版。全书27万余字,除导言、结语,另分7章:数字图书馆宏观管理概论;数字图书馆宏观管理机构;数字图书馆信息资源的宏观管理;数字图书馆技术和人力资源的宏观管理;数字图书馆知识产权的宏观管理;数字图书馆投资的宏观管理;数字图书馆工程项目的宏观管理。

该书详细比较、梳理了国内外学者对“数字图书馆”概念的解释,提出从狭义和广义两个层次理解“数字图书馆”。提出了我国数字图书馆宏观管理的三大目标:促进全社会数字图书馆稳定、持续、健康发展;构建数字图书馆合理的结构体系;保持数字图书馆总供给和总需求的基本平衡。

该书通过考察各国数字图书馆宏观管理机构及其职能,结合我国实际,提出在现有机构框架内,传统图书馆与数字图书馆统一归口管理的思路。结合数字信息资源申报登记制度,应该对现有而且不断增加的数字信息资源进行全面调查摸底。要尽快建立健全全国文献和数字信息资源联合目录;尽快实现全国所有新出版物的数字化;根据轻重缓急,分期分批开展各类文献数字化;对因特网各种信息资源进行系统收集整理保存提供,同时建立网络数字信息资源指引库。

该书通过论证指出,我国数字图书馆的信息资源可以采取集中式和分布式结构相结合的办法。可以考虑集中建立几个类似“美国记忆”的大型数据库。但是,从长远看,还是要采用分布式结构体系。数字图书馆信息资源建设的重点是数字信息资源整合,主要涉及数字图书馆异构数字信息资源系统和因特网数字信息资源的整合,使跨时空分布的数字信息资源成为一个用户可以无障碍利用的统一的数字信息资源整体。

该书认为,应该优化数字图书馆人才群体结构,包括人才的复合化和层次化。

该书认为,没有必要在法律条文上单独列出“数字图书馆知识产权”,而只需在现有著作权法的框架内,按照国际惯用的三步检验原则,针对数字图书馆所涉及的数字复制、网络传播、制作链接等问题,设置明确的限制条款,完善著作权法规。

该书认为,数字图书馆投资的宏观管理主要是调控全社会数字图书馆投资总规模。中央财政重点加强国家数字图书馆的投入,扶持西部和边远地区、农村地区数字图书馆投入。要通过宏观政策引导企业投资方向,对以公益性服务为目标的数字图书馆企业投资项目给予减免税和优惠税收、贷款政策。

该书介绍了目前我国数字图书馆工程项目的管理架构,提出了改进和完善数字图书馆工程项目专家咨询论证制度的措施。

(田信)