



国家图书馆俄文图书回溯书目数据制作外包回顾

Outsourcing as a Solution to the Russian Retrospective Bibliographic Cataloguing Data in the National Library of China

张 芳 陈玉玉 (国家图书馆 北京 100081)

[摘 要] 因工作量大且人员不足,国家图书馆俄文图书回溯书目数据制作采用外包方式。但在俄文图书回溯书目数据制作过程中存在卡片信息著录不全、参考分类号信息著录不全、早期部分卡片上信息不全、外包人员业务水平有限及质量评估方式存疑等问题。据此,俄文图书回溯书目数据制作工作应加强回溯书目数据制作外包的组织安排;注重具体业务科组与外包公司的配合与协作;加强外包人员规章制度的管理。国家图书馆的经验可为其他图书馆相关工作的开展提供一定的借鉴。

[关键词] 俄文图书回溯书目数据制作 业务外包 国家图书馆

[中图分类号] G254.3 [文献标识码] B

[Abstract] The paper gives a brief survey of the outsourcing to the Russian retrospective bibliographic cataloguing data in the National Library of China, which includes reasons and problems for the project, and other relevant considerations in the outsourcing process. From the experience of the National Library of China, some suggestions are proposed, such as reinforcing the organization of the outsourcing to the Russian retrospective bibliographic cataloguing data, emphasizing the cooperation of concrete business teams and outsourcing factories, and paying attention to the regulation management of the outsourcing staff. The experience of the National Library of China will be helpful to other libraries.

[Key words] Russian retrospective bibliographic cataloguing data; Outsourcing; National Library of China

为适应信息社会的要求、实施国家图书馆“十一五”规划纲要,2007-2009年国家图书馆通过招投标方式对馆内俄文图书卡片回溯工作进行了外包。目前,该项工作已接近尾声,书目数据已全部制作完毕,正处于对制作数据的全面检查验收阶段。本文拟通过对国家图书馆俄文图书回溯书目数据制作外包的原因、外包工作的开展及质量控制、外包过程中存在的问题及相关注意事项的介绍,增进各界对国家图书馆俄文图书回溯工作的了解,促进资源的利用与共享,同时,也希望国家图书馆的经验能对其他图书馆有一定的借鉴和推动作用。

1 俄文图书回溯书目数据制作外包的原因

国家图书馆俄文藏书是馆内外文藏书的重要组成部分。自上世纪50年代起至2006年底,国家图书馆共收藏俄文及其他小语种图书621 097册(件)。这些图书由12个独联体国家、波罗地海三国及保加利亚、塞尔维亚、马其顿和我国共19个国家出版,文种共46种。其中,俄文图书604 431册,其他文种图书16 666册。2003年,国家图书馆俄文图书采编工作开始使用ALEPH500自动化集成系统(以色列),数据格式采用MARC21格式。数据制作方式采取自建

或通过Z39.50下载俄罗斯国立图书馆的数据(年费600美元)。2003-2006年共制作俄文新书书目数据22 853种/23 046册,完成2000-2002年已编目图书的回溯书目数据制作19 642种册,并补做流通过程中的快速编目数据2 570种册,共计45 065种/45 258册。

2006年,国家图书馆制订了本馆“十一五”规划纲要。其中主要任务为:“有计划地进行馆藏文献卡片目录的机读数据转换^[1]。因国家图书馆俄文图书采编组共有7名员工,每年承担约6 000种俄文新书的采编及约5 000种回溯书目数据制作等专项工作,无法在余下的4年时间内完成尚留有的57万多册图书回溯书目数据制作工作。为实施国家图书馆“十一五”规划,2007年国家图书馆通过招投标方式,与北京社宇科技发展有限公司签署合同,完成俄文图书的回溯书目数据制作外包及馆藏挂接外包工作。

2 俄文图书回溯书目数据制作外包工作的开展

图书书目数据回溯工作是一件细致又烦琐的工作,特别是小文种图书。国家图书馆俄文图书采编工作已进行了将近60年。在此期间,因历史原因造成加工的俄文和其他小语种图书在分类体系、著录规则、索书号的采用及排架

方式上均有较大差异,且同一文种图书在不同的历史时期也有较大的变化,因此需要我们去了解、认识,并找到数据制作最恰当的表现方式。同时,由于外包公司人员语言水平的局限和制作小语种图书书目数据经验的缺乏,在前期的调研准备阶段及数据制作过程中都有大量的问题需要解决,管理及验收工作任务艰巨。

首先,要摸清家底。2006年国家图书馆便开始对馆内需回溯图书的数量、排架方式、卡片状况进行调研。截止到2006年底,国家图书馆尚有1999年以前编目、需制作回溯书目数据及馆藏挂接(即粘贴条形码,并通过其将图书与数据连接)的俄文图书57万多册;图书的排序按编目年代分固定排架和分类排架两种:1957年以前的图书按托罗帕夫斯基编制的《十进分类法》进行排序,1958-1974年的按固定排架排序,1975-1999年的按《中国图书馆分类法》排序。为加快速度、保证质量,同时又能在制作回溯数据的过程中及时完成馆藏挂接,我们选择了与图书排架顺序一致的卡片目录为依据进行俄文图书回溯工作。因当时俄罗斯国立图书馆1998年以前的回溯书目数据数量较少,故国家图书馆俄文图书回溯书目数据制作一律采取原始编目的方式。

其次,明确编目原则。国家图书馆俄文图书回溯书目数据的制作遵循《MARC21书目数据格式使用手册》、《西、俄文图书编目著录条例》、《俄罗斯编目条例》(俄文版),参考《西文文献著录条例》及其他相关标准。卡片回溯原则上要求严格依据卡片信息进行著录,对目录卡片中在格式、标识符、著录及多卷书的处理等方面与现行标准不一致的地方进行修改、删除、补充或找到其他相对合理的处理方式。如大写字母仅运用于著录项目第一个单词的首字母、专有名词和题名,单元内容按语法规则著录;为便于查重,将从编题名保留在丛编项中,而将卡片上著录在丛编题名前的团体机构名称改著在500字段,并加说明词“В надзаг.”(题名页题名上方);将卡片上为传记所做的根查著录在600字段个人主题附加中,并删除多余的文字说明“о нем”(论其);将其他责任者在700字段作个人名称附加等。

但针对某些不适宜按现行标准进行著录的卡片,依旧采取据实如片著录的方式处理。如1954-1958年和1975-2002年间编目的卡片,其索取号由分类号加著者号构成,且著者号与标目一致。当文献有多个责任者时,通常按其第一责任者取著者号。但根据目前国家图书馆俄文图书编目规则,文献有多个(4个以上,含4个)责任者时,只需将第一责任者在700字段作个人名称附加即可。在此情况下,在回溯过程中如果一概而论将多个责任者著录在700字段,势必造成数据中著者号与标目不一致。因此,尽管与现行标准不完全一致,但为了力求数据的统一,我们仍选择按卡片取多个责任者中的第一责任者作标目。对于俄文多卷书不同于西文图书的或分散著录,或集中著录,而是根据图书到馆情况择其连续卷册进行集中著录,采取了尊重卡片的集中著录方式,既在一定程度上避免了系统内

大量书目数据重复,又相对集中地对多卷书进行了揭示。

俄文图书卡片回溯过程中还应注意语言方面的某些特性问题。在将245字段中的其他责任者、会议等在7--字段做附加时,应将这些名称的变化形式(语法第二格,属格)还原为第一格(主格)。如245字段著录为:под ред. А. И. Иванова, 700字段应著录为:Иванов, А. И., 否则责任者的性别就由男人变为女人了。

另外,俄文字母A、B、E、K、M、H、O、P、C、T、X外形与英文字母相同,但在计算机系统内的编码完全不同,若编目时在检索字段著录混淆,会直接影响系统的查重和检索。如在索书号中出现错误,则会导致数据排序与图书实际排序不一致,给读者服务和数据使用带来不必要的麻烦。国家图书馆俄文图书编目所采用的索书号分别为:1954-1958年采用托罗帕夫斯基编的《十进分类法》和哈芙金娜编的《两位著者号码表》,1958-1974年实行图书固定排架,分类采用《中小型图书馆图书分类表(草案)》,1975-2002年采用《中国图书馆分类法》和哈芙金娜编的《三位著者号码表》。在上述分类法和索书号中,除《中小型图书馆图书分类表(草案)》和《中国图书馆分类法》类号中的字母用英文著录外,其余分类号、著者号、固定排架号中的字母均应用俄语表示。书目数据中ISBN号中的字母、罗马数字一律用英文著录,卷册号、丛编编号用俄文著录。

3 俄文图书回溯书目数据制作外包的质量管理

充分发挥管理的控制职能,通过前期控制、同期控制和反馈控制的综合运用,建立健全一整套行之有效的管理机制,加强与各方人员的沟通和协调,使得国家图书馆的俄文图书回溯书目数据制作外包工作在艰难中不断快速推进。截止到2009年10月底,已顺利完成并通过馆内验收的书目数据共计447 994余种,国家图书馆馆藏俄文图书卡片目录已全部实现可机读,为馆内俄文图书读者服务工作奠定了坚实的基础。

俄文文献资源本身具有的特殊性,使其在外包的过程中遇到了诸多的问题和挑战:(1)工作任务重、时间紧,要在短短几年的时间内完成几十万条数据的制作,压力之大可想而知;(2)没有现成的管理模式和经验可以借鉴,只能摸着石头过河;(3)外包公司工作人员需要跨越图书馆业务技能和语言两大壁垒,培训任务繁重;(4)卡片目录和馆藏分布情况复杂,文献年代久远,信息源不全。

面对上述挑战,国家图书馆外文采编部成立了专门的回溯工作领导小组,具体负责对外包公司的指导、监督、管理,与馆内相关部门的沟通、协调以及与各科组间的配合。在前期控制中,领导小组根据每年的任务不断调整和完善外包数据的制作标准,并与馆业务管理处就验收标准达成共识;由回溯工作领导小组组织定期和不定期的业务培训,加大对外包人员的培训力度(在前期控制中俄文图书采编组根据俄文文种的特点在部门数据制作标准和业务培训要

点中作相应说明)。在同期控制中,第一,与外包公司将工作的完成进度具体约定到月,要求其每月严格按时保质保量完成当月回溯任务,对逾期未完成或完成质量不合格部分给予惩罚;第二,实行阶段性验收制度,每月下旬按全年回溯书目数据量的一定比例进行验收,月底向部门和馆业务管理处提交验收报告,并对报告所附列表中的错误进行抽查,检查是否已及时全部修改完毕;第三,长期由专人对外包工作进度进行跟踪,防止出现前紧后松的情况,影响数据质量。在反馈控制中,不断积累经验,定期与外包公司进行沟通和反馈,总结经验、提出问题,并对问题的解决提出有针对性的解决方案。与此同时,随着俄文外包工作的不断推进,国家图书馆更加注重与上级业务监管部门的交流,积极达成各种共识,增加了外包工作的透明度。

4 俄文图书回溯书目数据制作过程中存在的问题

尽管在俄文图书回溯书目数据制作外包工作开始之前已进行了大量调研和精心准备,但今天回头看俄文回溯工作,仍发现许多问题。

4.1 卡片信息著录不全

卡片上的登录号(即财产号)作为文献的身份识别代码未被著录进书目数据,造成卡片信息回溯不全。2003年以前,国家图书馆俄文图书从采访到流通都是手工操作,文献从入藏起的全部相关信息原则上都被集中在唯一对应的登录号下并记录在馆内俄文图书采编组保存的俄文图书登记簿上。如果将登录号在书目数据的适当字段进行著录,则通过一条书目数据即可获悉一本书的全部信息,无需再去查找卡片,从而真正实现可机读的目标。但目前由于种种原因未将其著录进书目数据,若在后续工作中需要查找上机前采访的某文献的来源,数据里无登录号指示,不利于利用登录号进行查账。

4.2 参考分类号信息著录不全

上世纪50-60年代,国家图书馆曾使用过俄文全国统一编目卡片。集中编目工作由北京图书馆(1998年更名为国家图书馆)、中国科学院图书馆(2006年更名为中国科学院国家科学图书馆)、中国人民大学图书馆等单位合作,由全国第一中心图书馆委员会统一组织领导下的俄文统一编目组完成^[2]。卡片的编制以来自苏联的卡片为基础,卡片底端居中或靠右位置加打《中小型图书馆图书分类表(草案)》分类号、《中国科学院图书馆图书分类法简表》分类号和《中国人民大学图书馆图书分类法》分类号等本地信息。由于工作经验和业务水平有限,加之回溯文献涉及的年代久远、情况复杂,以致回溯之初我们误将卡片上的《中国科学院图书馆图书分类法简表》分类号与《中国人民大学图书馆图书分类法》分类号认作是俄罗斯某时期的特殊分类号,因此并未要求全部著录进书目数据,造成文献分类信息回溯不完整,甚至可能会给其他机构参考馆内数据造成一定程度的不便。

4.3 早期部分卡片上信息不全,很多书目数据的必备

项目完全空缺。

我们在操作时要求外包公司回溯人员按卡片实际情况进行回溯。如果卡片上无出版信息,则260字段如是著录:\$a[Б. м.: \$b б. и., \$c б. г.];如果题名不全,则用省略号代替。虽然此种做法就回溯卡片本身而言并无不妥,但从后来的实际应用中逐渐发现,书目数据,尤其是多卷集等文献的书目数据,其中大量描述性信息缺失,加之馆藏俄文文献数量巨大,因此给使用者通过馆内系统进行检索造成了一定困难。

4.4 外包人员业务水平有限

外包公司工作人员本身并未接受过系统的编目业务培训,且多数工作人员并未掌握一定的外语技能,加之“外包公司系以营利为目的,造成了其不可能在编目人员配置、业务能力提高等方面进行长期持续性的投入”^[3],因此在验收过程中经常会发现一些常识性的错误,如将题名中所含人名误当作责任者而著录进100或700字段;含多部分说明项的题名信息无法分辨如何著录;因不清楚卡片上某些说明信息的涵义造成著录错误。

4.5 质量评估方式存疑

虽然随着经验的不断积累我们一直在寻找最佳的质量评估手段,但每年的回溯数据数量巨大、任务繁重,所以只能选择按一定比例进行抽查验收。国家图书馆业务管理处外包回溯数据验收考核标准规定外包回溯数据抽检比例为回溯数据总量的10%,但在实际操作中,俄文回溯数据外包验收比例已经达到15%-20%。尽管如此,通过这样的验收方法得出的结论是否能真正体现回溯数据的质量仍待确认。

5 其他相关问题及注意事项

首先,加强回溯书目数据制作外包的组织安排。图书回溯工作是一项烦琐、细致的工作,特别是与外包公司磨合初期。全年安排的工作数量应适当,工作节奏紧凑而不忙乱,避免因赶时间、抢工期造成质量难以保障,而后又花大量时间进行校对、修改,造成人力、物力的浪费。另外,馆内各部门应全力协调、统一管理、标准一致,共同做好外包工作。

其次,在整个回溯外包工作的过程中,具体业务科组与外包公司的配合与协作也同样重要。如国家图书馆俄文图书采编组在俄文图书回溯外包的前期,为提高外包公司人员数据制作速度与质量,向外包编目人员提供了馆内工作人员长期以来积累的参考资料。如国家图书馆俄文采编组常用小语种代码、国家代码、地区代码等。为便于外包公司人员著录相关字段,将俄语单词 словарь(词典)、энциклопедия(百科)、справочник(手册)、конгресс(会议)、библиография(书目)、указатель(索引)等,以及俄文图书回溯书目数据制作中细节的处理方式均体现在馆内《外文采编部外包回溯数据MARC21格式要点说明》中,对外包公司更好地开展工作起到了一定的帮助、促进作用。

最后,加强外包人员规章制度的管理。严格执行卡片目录的交接手续,保证卡片不丢失及正确排序;挂接过程中图书不能直接码放在地面上,挂接完成后应将图书正确归架并摆放整齐;严格执行书库管理规定,严禁携带各种易燃、易爆等危险物品及食品入库,保证用电安全,禁止大声喧哗,保持环境卫生,不能将图书条形码随处胡贴乱扔。

总之,国家图书馆俄文图书回溯书目数据制作外包并非一帆风顺,期间经历了许多曲折,各种编目规则、标准的运用及规章制度的制订也是一个不断摸索完善的过程。总体来讲,虽然该项工作目前还存在着一些问题,但成绩是显著的,效果是明显的。首先,读者通过因特网可以查询本馆编的全部俄文图书,增强了图书检索的便利性和图书本身的利用率;其次,本馆工作人员在新书的采选过程中也能在系统中查重,避免或减少再版书、重印书的采选,以增加采选品种,提高采选质量;最后,目前国内许

多图书馆俄文图书新书少、旧书多。五六十年代国内图书馆普遍使用全国性集中编目的俄文统一编目卡片,国家图书馆俄文卡片回溯外包的经验或许对其他图书馆有一定的借鉴和参考意义。

参考文献:

- [1] 国家图书馆“十一五”规划纲要[EB/OL].[2009-08-09].http://www.nlc.gov.cn/service/gygt_ghgy.htm.
- [2] 宋继忠.集中编目[EB/OL].[2009-08-09].<http://www.czks.com/lib/name/nn10.htm>.
- [3] 赵伯兴,戴行德.论编目业务外包环境下的书目质量控制[J].国家图书馆学报,2008(4):79.

[作者简介]

张 芳 女,1962年生,国家图书馆副研究馆员。

陈玉玉 女,1981年生,国家图书馆馆员。

[收稿日期:2009-08-18]

(上接第39页)期内难以改变,数据库的投资应当向利用率高的中文数据库倾斜。

4.3 调整数据库资源的学科分布,优化结构。

按2007年的统计数据,深圳大学的数据库投资额按文科、理科、综合性学科3类划分,大致的百分比为7.41:52,同年,文科、理科、综合性学科3类数据库的使用量百分比为2.4:94。理科数据库占投资总额的41%,使用量只占4%,使用量百分比为投资总额百分比的1/10;综合性学科数据库占投资总额的52%,使用量却占94%。今后应根据学校学科的发展及数据库的实际使用情况调整数据库资源的学科分布,使其结构逐步优化,以适应学校学科优势的发展。在投资额的分配上,要突出学校的学科重点,同时还要兼顾各学科的平衡。

4.4 考量其他影响利用率的因素

除了经费投入、学科种类、数据库质量外,订购数据库还应多考虑影响利用率的其他因素,如数据库的语种、所覆盖的学科面、数据库收录的文献量、文献类型、是否提供全文文献等。这些因素将直接影响到数据库的使用效果。

4.5 加大数据库的宣传和培训力度

除了保持原有的宣传和培训措施(如新生入学教育、开设多门研究生与本科生文检课、每周举办数据库培训讲座、对图书馆员工进行数据库培训)外,图书馆还可以采取其他一些措施提高数据库的利用率。例如,增加师资,改变目前信息检索课程供不应求的状况;信息检索课程按学科进一步细分,实行专业对口的教学,深化教学内容;

针对几个投资大、利用率低但学科发展又要求必须订购的外文数据库,重点举办培训讲座;根据各学院、系、所的需要,上门开展有针对性的讲座;落实学科馆员制度,学科馆员对相关学院提供对口的资源服务;将数据库培训

列入校人事处管理,要求全校教师必须参加继续教育课程。通过加强对用户的宣传培训,切实提高数据库的利用率。

此外,定期开展对数据库的评估也是非常重要的。通过评估,图书馆能够了解用户的需求,淘汰不太需要的数据库,增订学科发展最需要的数据库,优化数字化资源的结构,提高投资效益,实现印刷型资源与数字化资源的合理布局,使资源的配置更为科学、合理,促进图书馆整体文献资源保障水平和服务水平的提高。

参考文献:

- [1] 杨 毅. CALIS 组团引进数据库现状和读者调查[R]. 南京:教育部 CALIS 高校组团引进数据库年会, 2006.
- [2] 杨 毅. CALIS 组团引进数据库的发展和评估[R]. 成都:教育部 CALIS 高校组团引进数据库年会, 2007.
- [3] 李 洪. 高校用户对国外网络数据库的使用统计与分析[J]. 图书馆论坛, 2005(3):210-212, 166.
- [4] 杨浴琼, 张凯勇. 高校图书馆数据库资源评价研究[J]. 情报科学, 2008 (11):1717-1719.
- [5] 张李义. 高校图书馆商业数据库评价指标体系研究[J]. 中国图书馆学报, 2004 (1):66-69.
- [6] 汪 媛, 赖茂生. 我国高校图书馆引进网络版全文数据库的综合评价模型[J]. 情报科学, 2004 (9):1061-1065.

[作者简介]

李 洪 女,1968年生,现工作于深圳大学图书馆,副研究馆员。

伍 宪 女,1954年生,现工作于深圳大学图书馆,研究馆员。

陈大庆 男,1964年生,现任深圳大学图书馆副馆长,副研究馆员。

赖丽萍 女,1978年生,现工作于深圳大学图书馆,馆员。

[收稿日期:2009-08-18]