

● 吴开华

文献资源共享和机读目录数据库建设

ABSTRACT The only way to meet better the informational demands of the society is to give full play to the integral superiority of social literature resources. To establish the MARC data base is none other than the base of literature resource sharing. But the crux of setting up the data base is the standardization and normalization of it which should be improved constantly. Suggestions are proposed after analyzing the present situation of the Chinese MARC data base.

SUBJECT TERMS Literature resource sharing — MARC data bases — China

CLASS NUMBER G252

1 文献资源共享的重要意义

近年来,随着科学技术的高速发展,图书情报资料迅猛激增。任何一个图书馆都无法也无能力收齐世界上所有的出版物。过去那种各自为政的图书馆已经远远不能满足广大读者越来越复杂的情报检索需求。只有把整个社会的所有图书馆当作一个整体,发挥全社会文献资源的整体优势,才能对整个社会的情报需求达到一个比较满意的保障程度。

我国是一个发展中国家,实行经济体制改革和对外开放政策,正在进行规模宏大的社会主义现代化建设,需要更多的信息资源。我国现时文献资源却很贫乏。据统计,我国所订期刊总数从未超过世界现有期刊总数的20%,且在有限的经费中,又存在大量的重复订购,若不通过地区乃至全国性的资源共享很难满足需求。我们不仅要对我们所拥有的文献信息资源进行开发和利用,以求最充分的共享,而且还要与国际文献信息网络接轨。

2 机读目录数据库的建设是文献资源共享的基础

机读目录数据库,简单地说,就是一种经过人们编辑组织、以机读形式出现的书目记录的集合。它将传统的卡片目录的内容以标准数据形式记录在计算机的存贮载体上,用计算机识别和阅读。

较之卡片目录,机读目录具有许多优越性。根据需要,人们可以通过编目用软件,重组、修改、更换数据元,编辑输出书本式目录、卡片目录以及各种专题目录和联合目录等。机读目录可以提供包括题名、责任者、分类号、主题词、ISBN号等检索途径,满足读者多角度、多方位的文献检索的需求。机读目录可以一次输入,多次输出,还可以转录和通过通信线路进行远距离传送,以达到一馆编制,多馆使用,一方生产,八方共享的目的。人们可以利用现代复制技术,或是通过远距离终端查寻、检索异地、它馆的馆藏目录,以利用各

馆的馆藏文献。知识信息进入数据库,犹如产品摆在商店里,可以供各个地方、各个阶层的人们自由选用。

机读目录有助于推进联合编目的发展,加强馆际互借工作,实现情报检索网络化。机读目录是知识信息最有效的开发和利用,将最大范围地满足读者不断增长的检索需求,达到最大程度的文献资源共享。

机读目录数据库是控制和处理知识的现代化手段,也是信息时代文献书目信息存贮和利用的必然要求。目前我国各大中型图书馆、情报单位的自动化系统设备已初具规模,建立文献资源共享网络技术的硬件也渐趋具备。但是,先进的设备只是一种工具,真正起作用的还是机读数据库中的书目数据。由于过去我国在图书馆现代化方面重系统,轻数据,导致今天书目数据库的建设远远落在后面,以至于花费昂贵经费购买的设备,或是闲置一旁,或是利用不充分。书目数据库是图书馆自动化的基础,没有足够容量的标准数据库的支持,实现图书馆自动化只能是一句空话。我们应该加速书目数据库的建设,尽快形成符合标准的 MARC 数据记录,为我国文献资源共享奠定基础。

3 书目数据的标准化、规模化是数据库建设的关键,也是资源共享的先决条件

书目数据的质量对于数据库建设至关重要。书目数据是图书馆计算机数据中唯一具有永久性保存价值的数据。其质量的优劣高低,直接关系到数据库的存亡与兴衰,关系到图书馆自动化事业的百年大计,是数据库的生命线。书目数据库的质量主要取决于数据库中全部书目数据的标准化、规模化程度。如果没有统一格式、统一编码和标准著录、规范标引的数据,机器无法处理也就失其意义。没有标准化、规范化,就没有自动化。

对读者来说,由于书目记录标识符号统一,书目结构规范,便于理解,容易掌握,通用性强,可以最大限度地方便读者使用,提高检索效率。

书目数据的社会化、公用化,要求书目记录必须具备兼容性和可交换性。不同编目部门所编的书目记录应该能够互换。一个图书馆所编的书目记录能够被其他图书馆或文献部门顺利地组织到自己目录中去。一个国家所编的目录记录也能够被其他国家比较容易地组织到本国的书目中去。只有这样,才能有利于馆际、国际书目资源的交换,进而实现馆际、国际文献资源的共享。

书目数据的标准化、规范化主要包括:文献著录、文献标引(分类标引和主题标引)、机读格式的标准化以及编目工作规范控制等。

标准是标准化工作的核心。我国文献工作标准化起步较晚,一定程度上影响了书目数据的标准化建设进程。目前关于西文图书我们可以利用 LCMARC 光盘作为标准编目数据源,还可以引进 OCLC 等覆盖面大、命中率高的其他数据源。对于中文图书,其标准书目数据只有靠我们自己去编制。

我国文献工作标准化委员会自 1979 年 12 月宣布成立后,在短短的十几年中,标准化工作还是很有成效的。我们的文献工作国家标准经历了从无到有、从少到多、从低到高的发展过程,已初成体系,但要继续努力,不断完善。

目前,我国书目记录编制工作可依据的标准有:(1)国际标准:《国际标准书目著录》(ISBD)、UNIMARC 格式和手册等。(2)国家标准:GB3792. 1-83《文献著录总则》、GB3792. 2-85《普通图书著录规则》、GB3792. 3-85《连续出版物著录规则》、GB3792. 4-85《非书资料著录规则》、GB3860-83《文献主题标引规则》等。(3)国家试用标准:《中国图书馆图书分类法》(第三版)、《汉语主题词表》(1980 年和 1991 年自然科学增

订本)。(4)在实际工作中正在起着行业性标准作用的《中国机读目录通讯格式》(书目文献出版社,1991 年)。

标准的执行更为重要。有了再多再好的标准,如果不在实际中应用、贯彻,那就什么效果也达不到。而且一个标准的质量,也只有实践检验中,才能得到进一步的改进和提高。我们的原则应该是,有国家标准的,就执行国家标准,暂时没有国家标准的,就参照国际标准或是国外先进标准执行。国际标准是按照严格的程序制定出来的,代表着科学技术的发展水平,包含着科技成果。向国际标准靠拢,可以提高我国的标准化水平。在国际统一的大环境中,标准的国际化是发展的必然趋势。

已选定使用的标准,要正确理解,全面执行,始终如一,贯彻到底。不要随意改动和省略规定项目,因为有时一个空格或一个标点符号,都会影响整个数据库的标准化水平。

4 我国机读目录数据库建设的现状与对策

80 年代中后期,我国书目数据库建设如雨后春笋,呈现出一派欣欣向荣的喜人景象。但是,要建成一个标准化、规范化并达到实用规模的机读目录数据库,并不是一件轻而易举的事,这项艰苦、复杂、细致而又浩大的系统工程,需要大量的人力、物力和财力投入。当前我国书目数据库建设主要存在以下几方面问题。

4.1 低水平重复建库。

图书情报系统基本上是延续五六十年代模仿原苏联的模式按行政系统设计的,各机构之间条块分割。这些单位功能相仿,藏书相似,固守着“大而全”、“小而全”的观念。系统之间各自为政,同一情报源多家重复建库。许多自建库标准化水平低、数据质量差,影响了使用与交换,有的甚至成了废库或死库。

4.2 大型的数据库少。

国产数据库的数据量一般为几万条,有的只有几千条,超过 10 万条记录的为数甚少。这当然与记录量高达 2300 多万条,且以每年约 200 万条速度继续递增的 OCLC 无法相比,就是与国外那些记录量为几十万条的一般数据库相比也差之甚远。数据量少、覆盖面窄、检索命中率低的结果必然是使用者少、效益差。

4.3 资金投入不足。

书目数据的前期处理、数据录入及系统建设均需要资金投入,尤其是一些大型的数据库更需要一笔可观的资金。在书刊大幅度涨价,图书情报单位购书经费严重不足的今天,仅仅依靠各馆自筹显然是不够的。

但是,只要认真总结经验,合理规划,统筹安排,加强引导,这些问题是可以在发展中得以解决的。

5 关于机读目录数据库建设的几点建议

5.1 以标准化为纲,改造非标准库为标准数据库

中文自建库标准化水平普遍低。究其原因,主要是初期建库时大多数图书情报单位只着眼于利用计算机对本馆本部门文献资源进行开发利用与实现自身的业务管理。封闭型独立开发,没有认真考虑如何与其他馆进行书目交换、实现资源共享等问题。无论是记录格式还是数据内容和形式,都没有严格的标准要求,造成数据质量的低劣。对这些数据库进行改造的关键是:组织这些单位与数据库建设的有关人员进行系统的学习和培训,提高对书目数据标准化意义的认识,掌握编制标准化数据的技能。对新入藏图书,利用北京图书馆图新公司发行的机读目录(1991 年开始发行),或一些地区性的联合编目中心,如北京高校图书馆联采统编部发行的机读目录(1993 年下半年开始发行,约 50 所北京地区高校图书馆中文图书联合目录)为标准数据

源进行编目。这样既保证书目质量、加快编目速度、提高标准化水平,又比较省时省力;对原有书目库中的记录逐一重新审查,进行标准化处理。这样能将成千上万个非标准库改造成成为标准数据库。

5.2 加强宏观调控,重点扶持高水平数据库

机读目录的共享性是其最突出的优点。将书目记录转换成机读目录是一项费时、费力、高投入的工作,一旦转换为机读目录,却可提供给任何一个图书馆或情报部门使用,共享编目成果。如对同一数据源,多馆重复加工纯属不必要。上级主管部门应该加强宏观调控,选择那些基础好,数据要求严格,标准化水平高,有发展前途的数据库,实行政策倾斜,重点投资扶持。使其功能社会化,结构网络化,参加国际竞争,进入国际市场,为加入世界大循环作出贡献。

5.3 发展横向联合,创建大型书目数据库

要使一个数据库在决策中发挥参谋作用,在教学、科研、生产中发挥助手作用,在动

态跟踪和监视国外高技术发展中起耳目作用,它必须拥有大量的、且覆盖面相当广泛的书目记录,这是任何一个图书馆或情报单位无法胜任的。只有将各地区、各系统的高水平数据库联合起来,建成国家级、大型的、综合性的书目数据库,形成全国性联机服务网,才是理想的选择。这是缓和文献增长和经费短缺,情报需求与文献匮乏两大矛盾,构建我国文献资源保障体系的重要途径。

参考文献

- 1 赵保佑,李正平,李允豹.情报社会学.东方出版社,1992
- 2 郑耀东,牧心.论世界数据库发展趋势和我国数据库建设问题.情报科学,1993,(2)
- 3 任世勇,王在平.从 OCLC 看我国图书馆自动化的发展.四川图书馆学报,1992,(6)

吴开华 女,1969年毕业于北京大学图书馆学系。现为清华大学图书馆编目部主任,副研究馆员。
通讯地址:北京清华大学,邮编 100084。

(来稿时间:1994-10-17。编发者:翟凤岐)

(上接第 61 页)

了图书信息的传播速度,扩大了图书馆影响,增强了图书馆吸引力,同时对教学和科研也能起促进作用。

3. 加快横向联合,实现分工协调。加强兄弟馆,特别是加强相同专业图书馆的联系,以对那些价格比较昂贵的进口原版书刊进行订购协调,共同使用。这样,既可缓和经费紧张形势,又可减少遗漏。如果以此为基础,再发展到跨地区乃至全国图书馆,将会收到更大的效益。

4. 适当减少复本量,大力搞好阅览工作。从长远看,扩大藏书种数比增加图书复本量更为重要,特别是对专业性和技术性较强的图书馆更是如此。1985 年我国高校图书馆平均复本量为 8 册,现在是 3~5 册。从发展趋势看将来很可能维持在 1~2 册。这样,新书大都

进阅览室,很少进入流通部门,读者外借将日益减少。要想解决这一难题,唯一的途径是把图书馆办成阅览中心,为读者提供所需的复印资料。尽管我国现在基础较差,大量复印的条件还不充分具备,但这是今后发展的必然方向。

5. 积极向有关部门申请,适当增加购书经费。为使图书馆能充分地为社会建设提供“能源”,各级图书情报部门应联合起来,开展协作,并积极向主管部门呼吁,要求增加购书经费,至少应补上书刊涨价差额,以维持馆藏的完整性和连续性。

陈 钢 男,现在大连医科大学图书馆工作,通讯地址:大连市沙河口区,邮码 116023。

(来稿时间:1994-11-24。编发者:丘峰。)