

# 中小型馆回溯书目数据库建设的 常见问题及对策

陈亦宜 薛 静

(安徽中医学院图书馆 合肥 230038)

书目数据的回溯建库工作的标准化、规范化是保障数据库正常运行的前提。我馆从上海医科大学图书馆引进北京金盘 GDLS 系统,在实际工作中遇到一些问题,并作出了相应的对策。

## 1 问题

### 1.1 数据来源的历史问题

一般大中型综合图书馆回溯编目都采用国家图书馆 CNMARC 数据套录,建库进度快,数据质量也有保障。但考虑到我馆是专业中小型图书馆,经费有限,数据命中率低,套录不是最理想的办法。经过调研,决定选择原先的公务目录卡片,作为书目数据的著录依据。在实际操作中,曾遇到以下问题:

(1) 书卡不符。以往的工作中一些图书流失,由于各种原因没有在公务目录中反映,形成了公务目录与实际藏书脱节。公务目录只能反映图书入库时的原始状况,不能反映变化了的实际馆藏。加之以往的图书改正、变更时部分卡改而书未改的历史事实,如果直接以卡片目录为依据录入,会产生有机读数据而无书,或数据的索书号不一致等现象。

(2) 款目不全。由于早期图书馆著录规则控制不严,编目人员频繁变动等原因,造成了目录卡片中款目残缺,尤其以 20 世纪 80 年代以前的藏书多见。如有的卡片仅有书名、分类号、价格,而没有责任者、出版项等一些最基本的书目信息,以及近年藏书目录中 ISBN、主题款目等著录项目也有不完整的情况。

### 1.2 数据录入的质量问题

(1) 不熟悉 MARC 格式。以往图书馆工作者流动较多,往往有人不熟悉编目工作的同志临时参与,使得部分书目数据质量难以保证。如著录格式有误,甚至遗漏必要的检索项。如直接以卡片目录上的书名录入错误数据“200@ a 现代中医药应用与研究大系:第六卷 内科”等。

(2) 录入标准前后不一。中文方式下的字母、数字、符号等录入有全角与半角两种方式,而《中国

机读目录格式使用手册》对此没有明确规定,如随意录入,就形成了前后不一致;又如录入时遇到汉字编码字符集中没有的字,录人员可能用同音字代替,或用一个或二个空格,甚至用“?”等处理,录入处理前后不一致积累问题较多。

(3) 粗心致错。有的录人员追求速度而不顾质量,导致录入错误,如把 1998 录入为 1997;或登录号给号重复,造成一个登录号同时存在于两条或两条以上不同的书目数据,这种现象占有很大的比例。

## 2 对策

### 2.1 对数据录入人员进行 CNMARC 格式培训

减少数据录的操作差错是提高书目数据质量的关键。首先要求数据录人员有很强的责任心,要掌握 CNMARC 格式和操作基本知识,还要全面了解《中国文献编目规则》、《中国图书馆分类法》等相应标准。为此必须对录人员进行系统的培训,首先要熟悉《中国机读目录通用格式》,明确 GDLS 系统的 CNMARC 格式所包含的各字段、子字段以及字段和子字段的标识符及指示符的含义,然后按机读目录格式的要求对每一种数据进行详细著录。

### 2.2 制定编目细则保证书目数据标准和规范

首先必须严格把握统一的国家标准。同时结合中小型馆实际,对某些项目进行了简化和省略。其中必须对 CNMARC 字段、子字段的选用等细节一一作出明确规定,避免在具体录入过程中因人而异,相互不一致。对于《中国机读目录格式使用手册》在文字录入标准化方面未作出明确界定的现状,我们制定了既便于操作又便于检查的条文,使数据录入有章可循,如:书名、责任者中的字母、数字和符号,汉语拼音、外文姓名的原文及缩写,占一个字节,取半角为准;遇到汉字编码字符集中没有而无法录入的字一律用同音字代替等。

### 2.3 做好藏书剔旧和书卡核对等藏书基础工作奠定建库基础

(下转第 49 页)

对应的属于该笔画汉字所在的页面链接起来;在某一确定笔画的汉字所在的页面中按第一、第二笔的笔形一丨ノ、フ 将其分成若干部分,并将各部分的名称与所包括的汉字根据实际情况使用超级链接或书签建立连接关系。再将每一部分根据字体结构按左右结构、上下结构、半包围结构等进一步划分并根据实际情况建立超级链接或是使用书签与该部分的难检字建立检索关联。

以查找浙江清代学者俞樾的“樾”字为例:先数该字的笔画数是十六画,“樾”的第一笔和第二笔分别是“一丨”,左右结构。在手册首页的笔画数列表中找到笔画数十六(已设置为超级链接),双击进入笔画数是十六的下一级相关联页面,在网页列表中找到“一丨”(已设置为超级链接),双击打开其下一级页面,根据该网页上提供书签,将目标定位到网页的包含左右结构难检字部分:这样就可以很轻易的找到“樾(yue):两树交聚而成的树荫。也指道旁成荫的树。其陈桥五笔键值是 sht”。

## 2.2 用户信息反馈模块的设计与实现

设计一个基于网页的留言板来实现。首先,用 Frontpage 2000 表单功能设计一个用户界面。在同一个表单中应包括三个文本框,两个发送留言的按钮。这三个文本框分别命名为 name、E-mail、homepage,供用户使用完手册后填写反馈信息使用。第二步是使用 Access 建立一个数据库。其数据表字段将其与表单中文本框的命名相一致: name、E-mail、homepage。保存并命名为 liuyan.dbf。该数据库用来保存用户在留言板传输到服务器上的反馈信息,用户使

用时运行在服务器端。表单界面做好之后,再利用 Frontpage 2000 的站点设置功能建立与数据库的连接。在页面上表单属性中,点选“发送到数据库”按钮,单击“选项”按钮,在“数据库结果”选项卡中,分别选中文本框的表单项,单击“修改”按钮,在修改域中分别修改表单项和数据库中的字段一一对应;然后一路确定,以更改表单中文本框的值,以便单击表单中的命令按钮能修改数据库中对应字段;最后对表单中的各个文本框进行相关方面的设置。这样就完成了用户与数据库相关联的设置。用户使用过该手册后根据网页上关于留言板的提示,在留言板的三个文本框中分别输入自己的姓名,E-mail 地址,使用意见及使用过程中出现的问题等相关内容,按“提交”按钮将留言提交系统,按“修改”按钮重新填写反馈意见然后再提交给系统。收集反馈意见,管理员可以对电子手册进行不断的完善。

## 参考文献

- 1 方其桂等编著. Frontpage 多媒体 CAI 课件制作实例教程. 清华大学出版社, 2002
- 2 杨宗义等编. 难字大字典. 西南师范大学出版社, 1998
- 3 吕新平, 张强华编. 五笔字型速查手册. 西安电子科技大学出版社, 2000
- 4 张世伟编. 数据高手—Access 数据库轻松做. 北京电力出版社, 2003

纪少平 1965 年生,女,陕西理工学院图书馆馆员,文献检索教研室主任。

程荣芳 1977 年生,女,陕西理工学院图书馆信息部助理工程师。

(收稿日期:2004-08-19 编审:方子丽)

\*\*\*  
(上接第 47 页)

我馆利用此次回溯建库工作对馆藏图书全部整理和清点。首先按照先中医药,后西医的顺序做好图书剔旧。第二步进行书、卡核对,核对时注意逐项校对,及时改正卡片目录著录中的错误或模糊记录,及时核对解决补充书卡不一的问题。书卡核对结束后,依分类顺序排列,方便数据录入。同时,启动图书条形码粘贴工序,完毕后及时抽出书根卡,依分类号排序,供下一步录入条形码数据。实践表明,虽然在前期准备工作花费了较多的人力、时间,但可有效提高建库效率。

## 2.4 严格校对保证建库顺利进行

校对工作主要有两方面内容:一是建库前的书卡核对,二是对录入的各种数据进行校对,这是书目数据质量的重要保证。要求校对人员认真细致,一

丝不苟。采编部编目员长期从事图书著录,具备丰富的编制机读目录的经验,被列为校对人员首选。一旦出现录入差错,他们可凭借丰富的经验及时作出“诊断”并予以纠正。在具体操作上,建库初期要重视录入数据的跟查校对,将录入错误控制在最小范围内,建库后期可利用软件程序进行登录和索书号查重以及抽查审校核对。

## 2.5 关注反馈以加强书目数据库的维护

关注读者检索和流通服务效果,通过各种反馈信息,及时掌握机读目录数据库存在的错误并进行修正,也是提高回溯书目数据库质量的一种有效途径。编目人员不能仅仅埋头建库,还要与一线服务人员加强沟通,获取反馈信息以提高工作质量。

(收稿日期:2004-07-06 编审:王宗义)