

浅议编目数据的适时维护

王庆雯

摘要 在编目工作中, 编目人员对数据进行适时维护, 不仅可以减少重复的劳动, 还有利于提高数据的准确性。本文从图书采购的监控、数据的整体性、著录的|致性、记录的唯|性等几个方面, 探讨编目工作中如何对数据进行适时维护。

关键词 数据维护; 编目

分类号 G254.3

Abstract Timely maintenance of catalogue data not only reduced the workload, but also improved the accuracy of data. In this paper, some problems in data maintenance were discussed from the control of procuring books, data entirety, description consistency and record uniqueness.

Keywords Data maintenance; Catalogue

Class Number G254.3

编目是图书馆文献信息资源建设中一个关键环节, 它通过对文献的整理、加工, 揭示文献的内容与形式, 以利于文献管理与使用。编目数据质量的好坏, 直接影响到图书馆整体效能的发挥。编目的首要环节是查重检索, 目的是为了“避免‘同书异号’或‘异书同号’现象, 提高图书分类、归类的准确性。查重检索还有助于发现数据中存在的其它错误, 因为这些数据错误会随着数据库的扩大、标准的更新而增加, 这将严重影响到数据库的准确性和使用效率。面对数量庞大, 内容繁多、结构复杂的馆藏数据库, 单纯依靠维护人员定期查找、修改这些错误不仅困难, 而且容易遗漏。因此, 培养编目人员在编目工作中有意识地查找并改正错误, 实现编目数据的适时维护, 不仅可以减少重复劳动, 还可以保证数据维护的质量, 有事半功倍的效果。但目前这些工作还不能有效开展, 因为还存在一些问题。本文结合我馆的实际情况, 浅议编目人员如何对数据库进行适时维护。

1 数据监控与重购预防

随着自动化、网络化的快速发展、采访查重软件的不断更新及采集器的使用, 采访工件较

好地控制了图书的重购。但是, 由于馆内文献信息服务系统未升级, ISBN 号由 9 位升至 13 位后系统不能互认; 出版业的不规范, 名义上具有唯一性的 ISBN 并不具备唯一性; 出版物的排版及题名形式的复杂、多样化; 书商提供的书目数据不完善等原因, 不可避免地会造成图书的重购。在订购与验收环节不好判重的情况下, 编目人员是把好订购质量的关键角色, 他们通过待编图书的题名、责任者、出版说明等信息的检索, 能够发现一些与馆藏内容相同的图书, 特别是题名多样化而内容完全相同的图书, 对待这样的待编图书应挑选、退回, 以保证入馆图书的质量, 减少图书经费的浪费。

例如: 在待编“文芳”著的一套《黑色大历史系列丛书》时, 题名形式多样化, 以其中的一本为例, 书名分别为“黑色大历史”、“亲历讲述私人记忆中的黑色往事”、“官场黑幕”、“官祸”。书商的订购记录为: 2001# \$a 黑色大历史 \$i 官场黑幕 \$f 文芳编。通过 ISBN 号、责任者检索到馆藏有一条记录为: 2001# \$a 官祸 \$f 文芳主编 2252# \$a 百祸民生系列丛书。与待编图书比较后发现除正题名、丛书名、价格、印刷次数不同外, ISBN、题名“官祸”、责任者、出版社、装订尺寸、页码均相同。可以判定这两套丛

2010 年 1 月 January, 2010

书内容完全相同,应视为重购图书予以退回。此套书总价 220 元,以 5 本的复本数计算可以节约图书经费 1100 元。

从上述例中看到,在订购数据与馆藏数据差别较大的情况下,采访与验收人员很难判断,只有通过编目多途径检索才能准确判断,予以控制。

在对待编图书进行检索时,还能够发现一些无效的订购数据。有些是图书馆采用预订与现采结合的图书采选方式,一些书采到后,订购记录未及时删除;另一些是预订的图书未收到而订购数据超过使用周期,没有进行删除及转订单处理。遇到这些数据,编目人员要及时地反馈给订购人员核对、处理,以免出现重购或变为无效数据,给读者查询书目记录及馆藏情况带来不便。

在对连续出版物及多卷书检索时,应核查馆藏记录中缺失的卷次、册数,并将这些信息及时反馈给采购人员进行补购,以保证馆藏文献的完整性和连续性。

2 数据的整体性

文献著录的结果不是一条条独立的书目记录在数量上的叠加,而是在形式特征或内容特征上相互联系的一个有机的书目信息集合。检索中遇到出版物名称的变更、正编与其补编、教材与其配套书以及具有派生关系的姊妹篇等有关联的图书时,在对新书进行规范著录之前,要在原先的记录上增加连接字段,使机读目录中相关记录真正形成网络联接,达到输入任何一个题名及其相关联的记录就全部能检索到。这样才能发挥计算机的一次输入多条输出的检索功能,为用户提供更多的信息。

例《历史文献研究》原名为《中国历史文献研究集刊》在长沙出版,1986 年起改名《中国历史文献研究》在武汉出版,1990 年由北京出版改为现名,现在又由华中师范大学出版社出版。在编

《历史文献研究》时,通过书中序言和版本说明了解到题名变更信息后,分别以《中国历史文献研究集刊》《中国历史文献研究》题名为检索点进行查重。并对这两条记录增加连接字段。

(1) 2001# \$ a 中国历史文献研究集刊 \$ f 中国历史文献研究会编

311# # \$ a 本书第一、二辑由湖南人民出版社出版,第三—五辑由岳麓书社出版。1986 年起改名《中国历史文献研究》,由华中师范大学出版社出版三辑。嗣后移北京,改名《历史文献研究》。现由华中师范大学出版社出版。

440# 0 \$ 12001 \$ a 中国历史文献研究 \$ 1210 \$ a 武汉

440 # 0 \$ 12001 \$ a 历史文献研究 \$ 1210 \$ a 武汉

(2) 2001# \$ a 中国历史文献研究 \$ f 张舜徽主编 \$ g 中国历史文献研究会,华中师范大学历史文献研究所编

311# # \$ a 本书原名《中国历史文献研究集刊》,第一、二辑由湖南人民出版社出版,第三—五辑由岳麓书社出版。1986 年起改名《中国历史文献研究》,由华中师范大学出版社出版三辑。嗣后移北京,改名《历史文献研究》。现由华中师范大学出版社出版。

430# 0 \$ 12001 \$ a 中国历史文献研究集刊 \$ 1210 \$ a 长沙

440# 0 \$ 12001 \$ a 历史文献研究 \$ 1210 \$ a 武汉

(3) 2001# \$ a 历史文献研究 \$ f 中国历史文献研究会

311# # \$ a 本书原名《中国历史文献研究集刊》,第一、二辑由湖南人民出版社出版,第三—五辑由岳麓书社出版。1986 年起改名《中国历史文献研究》,由华中师范大学出版社出版三辑。嗣后移北京,改名《历史文献研究》。现由华中师范大学出版社出版。

430 # 0 \$ 12001 \$ a 中国历史文献研究集刊 \$ 1210 \$ a 长沙

430 # 0 \$ 12001 \$ a 中国历史文献研究 \$ 1210 \$ a 武汉

在原馆藏记录(1)、(2)中增加了 311、430 或 440 字段后与新编文献(3)记录形成了有效连接。不管使用哪一个题名查寻,都可以检索这三条记录。

3 著录的一致性

CNMARC 的发展经历了几个阶段,而且日趋规范。但回溯过去的数据库,其字段、子字段、标识符和指识符甚至其内容和含义均有一定的变化;各馆在建库的初期大都采取突击式的回溯建库,数据记录简单、规范性差;同一种丛书、多卷书集中和分散著录形式多样化等因素都造成了数据著录的差别,影响了数据库数据的准确性和检索功能。在编目过程中,编目人员一定要注意数据著录的规范性和一致性。

3.1 题名著录的一致性

题名是图书编目的重要著录项,也是读者确认、选择文献信息资源必不可少的依据。题名著录的完整准确直接影响到书目数据的质量和读者的检索效率。检索中对含标点符号、空格等特殊字符的题名一定要检查标点符号、空格的著录是否规范,对 200 字段繁体字正题名要在 518 字段提供简体汉字形式的正题名检索点,对题名前冠有“钦定”、“袖珍”、“实用”、“新编”、“简明”等字样的正题名,在 517 字段提供去掉这些冠词的题名检索点。含有年代的题名,如果是集中著录的连续出版物,应将去掉年代后的正题名著录于 200 字段 \$ a 子字段,将年代著录于 010 字段 \$ b 子字段,如果是分散著录的单本图书,200 字段应客观著录题名页提供的文献题名,在 517 字段提供经规范处理后的题名检索点。

3.2 责任者著录的一致性

在 CNMARC 中责任者著录主要涉及 200 字段和 7XX 字段。在 200 字段责任者按照题名页上的形式客观著录。对同一责任者要使用

统一姓名在 7XX 字段提供检索点。在检索中,最常见的是外国责任者的汉译姓、宗教人物及帝王名称、作者笔名及艺名等责任者姓名著录的不规范。书目数据中只有责任者检索点著录的统一,才能保证数据检索的准确率。因此,对原记录中的不规范责任者一定要修正,根据需要增加 314 字段对责任姓名进行规范说明,并在 7XX 字段著录规范的姓名。

例: 2001# \$ a 网络星河 \$ e 对互联网、商业和社会的反思 \$ f(美)曼·纽尔·卡斯特著 \$ g 郑波, 武炜译

314 # # \$ a 责任者 Castells 规范汉译姓: 卡斯特利斯

701# 0 \$ a 卡斯特利斯 \$ g(Castells, Manuel), \$ f1942- \$ 4 著

例: 2001# \$ a 搏斗 \$ f(法)罗曼·罗兰(Romain Rolland)著 \$ g 陈实, 黄秋耘译

701# 0 \$ a 罗曼·罗兰 \$ g(Romain Rolland) \$ 4 著

注意,罗曼·罗兰是广为人知的中文名称,可以不按照规定的形式记录,而将普遍接受的形式直接著录到责任者检索点。数据库中同一责任者的不同书,要保持责任者著录形式的一致。

3.3 分类标引的一致性

编目人员查重检索不仅仅是为了确定复本或者判断文献版本的不同,更多的是针对文献的类型和内容特点,通过相应的检索点对馆藏中待编文献具有相同、相近学科内容的文献分类标引的查看,为新书分类标引提供参考依据。在检索中,对数据的对比和分析,馆藏数据分类标引不一致的问题就会凸现出来。编目人员在遵循中图法分类原则的基础上,依据图书馆的性质、藏书结构、管理方式、本馆的分编细则及时修改,以保持同类、同种文献分类标的一致。

例如,对英汉对照读物的分类,《中图法》第四版“H319.4 读物”类目下注释规定:“以提高阅读能力为目的的各种简易读物、对照读物、注

释读物入此, 如愿细分, 可用组配编号法。依据本馆分编细则, 对英汉对照读物的分类标引不作组配编号, 直接归入 H319.4 类目下。但要组配编号数据著录在第二个 690 字段, 以便为读者查寻提供参考。

2001# \$ a 妙语真言\$ e 英汉对照\$ f 游径海编译

690# # \$ aH319.4 \$ v4

690# # \$ aH319.4: H03 \$ v4

905# # \$ aCCNU \$ dH319.4/ YJH1

4 记录的唯一性

在 CNMARC 中, 每一种图书在馆藏数据库中都对应该一条书目记录。但在检索中发现多条记录并存的现象较多。除了编目人员打字输入错误、检索查重工作不到位造成的外, 更多的是因编目人员对著录的规则及图书的认识不统一造成的。如印刷年的改变误作为版本的变更; 丛书与多卷书集中与分散的选择不一致, 共同题名与分辑题名在正题名的选取上存在差别等等, 都可能造成重复记录, 既占用了数据库的有效空间, 影响运行速度和管理效能, 又容易使读者困惑。遇到这些问题编目人员要及时清理、合并, 注意合并后删除多余的记录。例《萌芽》2006 年度佳作/ 中国名刊 2006 年度佳作系列与 2007 年度采用集中著录的形式著录, 而《萌芽》2005 年度佳作/ 中国名刊 2005 年度佳作系列采用分散著录的形式著录。

集中著录形式:

010# # \$ a978-7-5407-3837-2 \$ b2006 \$ dCNY24.80

010# # \$ a978-7-5407-4020-7 \$ b2007 \$ dCNY29.80

2001# \$ a《萌芽》年度佳作\$ f 萌芽杂志社选编

2252# \$ a 中国名刊年度佳作系列

分散著录形式:

010# # \$ a7-5407-3407-8 \$ dCNY24.80

2001# \$ a《萌芽》2005 年度佳作\$ f 萌芽杂志社选编

2252# \$ a 中国名刊 2005 年度佳作系列

查重时针对上述情况, 应将分散著录形式合并到集中著录形式中, 在集中著录形式中增加 010 字段即: 010# # \$ a7-5407-3407-8 \$ b2005 \$ dCNY24.80, 并在馆藏业务窗口中增加年代说明。合并后删除分散著录记录。

数据的管理和维护是一项极其重要的日常工作, 它直接关系到图书馆实现自动化管理和数字化服务的效能发挥。编目人员必须充分认识到这项工作的重要性, 认真对待检索时发现的每一个数据问题, 并做到适时维护和修改, 以保证数据资源的完整性、准确性、实用性, 为读者提供更好的服务。

参考文献:

- [1] 北京图书馆《中国机读目录格式使用手册》编委会. 中国机读目录格式使用手册[M]. 北京: 华艺出版社, 1995. 12.
- [2] 谢琴芳. CALIS 联机合作编目手册: 上[M]. 北京: 北京大学出版社, 2000.
- [3] 周明华, 刘荣珍. 书目数据清理与合并[J]. 大学图书馆学报, 2005(6): 51-56.
- [4] 牛书东. 图书馆馆藏书目数据库的维护探讨[J]. 图书馆工作与研究, 2007(3): 41-42.
- [5] 修晓. 论 MARC 著录的整体现[J]. 图书馆建设, 2007(5): 65-68.

作者简介:

王庆雯(1966-)女, 馆员, 华中师范大学图书馆, 武汉, 430079.