

# 论采访查重的延伸与完善

李金庆, 卓晓云

(厦门大学图书馆, 福建 厦门 361005)

[摘要] 提出查重实践中不同的查重方式和查重范围, 指出查重的误区与问题, 认为“复本”范畴不限于同一种书, 应延伸查重的外延, 改进与完善查重工作。

[关键词] 文献采访; 查重; 改进

[中图分类号] G254.3

[文献标志码] A

[文章编号] 1005-6041(2010)06-0013-03

## 1 查重概述

查重对于图书馆的采访、验收和编目等工作流程来说, 是必不可少的一个重要环节。文献采访过程中的查重, 即查检、核对书商提供、读者推荐或采访人员拟采购的图书是否已入藏或预订过, 以决定是否订购或增补册数, 避免重订。采访查重是控制文献复本量、节约经费的重要措施, 也是了解馆藏情况、提高信息资源采选质量的方法之一。查重是否准确、科学, 在很大程度上影响着图书馆馆藏的数量和质量, 并直接关系到图书馆的藏书建设和为读者服务的效果, 因此搞好图书采访查重具有深远的意义。

要做好查重工作, 首先要明确什么是复本。在工作中, 不同的采编人员对复本的概念认识不一, 这种认识上的偏差, 给查重工作造成一定的混乱。一般认为“若在编图书与某一馆藏记录的题名、责任者、出版项、ISBN等著录信息完全相同, 则为重书, 也称复本”<sup>[1]</sup>。各个著录信息完全相同是复本, 这没有歧义, 但具体的理解与操作并不简单, 值得探讨。

## 2 查重的外延

对复本的认识有差异, 查重的外延也就引申出几种不同的范围:

(1) 第一种是狭义的查重, 专指同一种书, 即书名、作者和书号等基本信息都一样, 内容完全相同的才判断为复本。同一本书如果以后翻印、重印, 根据 ISBN 编号原则, 即使内容没变化, 只要装订与版式有显著变化或是重印书的出版者发生变化, 以及再版书或重印过去无国际标准书号的图书时, 都需另外申请新的 ISBN 编号。例如, 平装、精装、线装等不同装帧形式但内容相同的图书, 通常有不同的 ISBN 号。唐元基、罗庆泗合著的《乾隆传》, 平装本的 ISBN 为 7-01-001922-3 精装本的 ISBN 为: 7-01-001953-3<sup>[2]</sup>。这种情况下, 虽然图书内容没

变化, 但书号变了, 如果查重时没当做复本, 就会导致一馆重复订购很多内容完全相同的书的后果。

(2) 第二种是较为广义的查重, 指对于内容相似或相近的图书, 即将论述内容及学术层次、读者对象方面近似的图书, 也作为“准复本”考虑。例如, 莫泊桑短篇小说选有多家出版社出版, 虽然所选内容不会完全相同, 但大同小异。笔者认为图书馆选购人民文学出版社、外语教学与研究出版社等少数两三家出版社出版的版本即可。

(3) 第三种是载体扩展的查重, 查重范围不限于一种载体, 同一种文献或内容相似的文献, 不管载体是否相同, 都视为复本。例如, 厦门大学图书馆 2008 年试行查重对照中文期刊的印刷版与电子版, 考虑停订上千种已收入数据库且非核心刊的印刷刊物。

(4) 第四种是馆外延伸的查重, 即查重的范围不限于本馆, 必要时延伸到对共建馆的查重, 例如参加外刊协调共建的单位, 对于一些高价刊是否要订购, 还需查检其他共建馆的订刊情况。

## 3 查重的误区与问题

在查重工作过程中, 存在着一些认识上的误区和较普遍存在的问题, 主要有下面几点。

### 3.1 外延狭隘

很多图书馆查重时只限于第一种意义的查重。他们把内容近似甚至内容相同只是版本有别的书定义为非复本, 使这些书就不能作为复本排除。例如, 某图书馆先后 26 次订了十几个出版社出版的法国作家雨果写的《巴黎圣母院》包括同一出版社不同时期出版的版本。这里面的主要问题是把不同书号的都当作非复本购入, 结果产生了大量实际意义上的复本, 造成极大浪费。

### 3.2 没有考虑不同载体的互补

大部分图书馆采访时不考虑不同载体的互补,

查重时只查同一载体的文献,从而造成了经费上不必要的浪费。虽然目前电子文献客观上还很难大范围替代纸本文献,但是在某些情况下,例如对高价书刊的采访,还是应该考虑电子书刊的替代功能。

### 3.3 几个极端

一是认为由单一途径查重,可减少漏检率。但是因为降低查准率,检出结果有不少信息垃圾,影响采访人员快速、准确地判断。偏好单一检索途径的另一种情况是采访人员图省事,检索时只选用一个检索途径进行查重,结果是把许多同名异书、同号异书的文献当作复本去重,遗漏了一些本该购置的重要文献。

二是认为同时由多个途径混合查检,可提高查准率。殊不知,由于书目信息不一定很准确,有时对于是否有复本还需要人工辅助判断。如果同时采用过多的限制条件来检索,就会造成不应有的漏检,从而重复购置复本。因此,采访要根据具体情况选择、变换检索方式和检索词。

三是认为机检可完全取代人工识别,无需核对。事实上,虽然不少查重软件功能强大,但由于检索过程的复杂性,查重结果往往离不开人工判断。特别是碰到检索词错误或不规范的情况,容易出现同书异名、同名异书的情况,都需要人脑识别。

### 3.4 粗心大意

有的采访人员在查重时粗心大意,造成重订、漏订。对于国外作者由于音译、意译差别或输入字母、符号时大小写切换及全角半角状态不同等原因,录入的责任者也可能不相同;国内作者也会因本名或笔名等原因,同一责任者有不同的表示形式;同书异名、同名异书的情况也不少见,特别是翻译著作会因不同译者而采用不同的题名<sup>[3]</sup>;有的采购系统没解决新旧 ISBN 号转换的问题,也需要查重人员细心辨识。

多卷集查重时如不细心核对,容易看错卷期,从而漏掉所缺卷期,重订已经入藏的卷期。同时,查重时不看复本数,不管是否需要补购复本,一概去重,也是普遍存在的问题。

### 3.5 时间差导致重复订购问题

外出采访是图书采访的重要形式之一。由于在馆采访人员和外出采访人员同时办订,如未合理协调,外出采访的图书和订购数据滞后一段时间到馆后,可能会出现不少重订情况,这就是采访的“时间差”问题<sup>[4]</sup>。解决的办法是外出采访数据在确认订购之前要连线馆藏采访库和中央书目库进行查重。没有条件在外连线的图书馆,相应时间内的在馆采访数据,要等外出采访数据查重确认后再发订。我馆使用汇文的无线外采软件,可随时上网连线本馆

服务器进行查重,但有的书商的采集器只采集 ISBN 号,图书信息不全,所以仍会存在重订现象。

## 4 查重的改进与完善

在查重过程中,通过拓展范围、灵活检索、综合考虑、细心核对可提高查重质量。

### 4.1 内容的拓展

“复本”应不限于图书内容和书目信息完全相同的书,同一种书的不同版本,要分析判断其是否新增有重要参考价值的内容,如果没有就不必购进新版本,除非原来的复本量不足。

对于内容近似的书,既要考虑各馆对这方面文献的需求量以及现藏量,还要分析待选文献的学科性质和学术价值,与原有馆藏中同种类文献的相关程度,待选文献是否比已有馆藏同类文献的学术价值更大,或者是否简单重复,确实需要时才购置。

### 4.2 载体的延伸

不同载体形式的可替代文献十分丰富,查重时应兼查不同载体的复本情况。电子文献在一定条件下可替代纸本文献。笔者认为,在一般情况下,图书馆可考虑下面一些文献的替代:1)介于买与不买之间的文献。这类文献在采访选择时经常碰到,主要如:文献内容属于采访范围,但水平一般,较为重要的书刊,但价格偏贵;已有馆藏,只是复本稍嫌不足;新版本频出的图书。2)价格昂贵,但利用率不高的外文书刊。文献的查重范围应包括各种载体,例如,厦门大学图书馆对于推荐订购的外文期刊,查重后发现馆藏没有纸版但有电子版,或属 OA 期刊,依照采访原则概不新订。

### 4.3 灵活查重

查重工作有不少技巧,灵活运用可提高效率和准确性。在具体工作中,采访人员常常根据自己的习惯和偏好来选择不同的检索点,但为了避免漏查,查重的方式应灵活多样,可以选用多种方式混合查重,并根据需要选择不同的匹配方式(如前方匹配、中间匹配或者是完全匹配)来提高查全和查准率。有时会碰到某些前冠小字或数字是否为题名一部分的疑惑或者前置词不确定的问题,这时可通过截词查重回避难题。

在查重过程中碰到某些记录有疑问时,采访人员应根据常识进行判断,看看是不是检索词不准确、不规范,或是同书异名、同作者异名(如“卡内基”与“卡耐基”)等,要分析判断可能正确或变异的检索词,加以试查,或改变、结合多种检索途径试查。查重时要正确处理标点符号、空格、数字、汉语拼音及外文字母,避免漏检。

(下转第 58 页)

## 5 结 语

报纸是一种及时、快速的信息载体,它和期刊、图书一样,也是图书馆馆藏资源的组成部分。传统的报纸管理工作只是一种被动式服务。而今,随着互联网的高速发展,全球信息已成为一体,仍没有哪个行业的信息会像报纸信息那样庞杂、廉价和透明。世界上大多数报纸网站显示的都是全部报纸的信息内容,而且使用其报纸数据库也比使用其他类别的数据库限制要少得多。这种差距正是报纸管理工作者的优势。综合利用网络报纸及报纸数据库是网络时代开发报纸信息服务效果的最佳切入点,图书馆员应该借此东风,善加利用,开展多种形式的报纸服务和宣传工作,使国内乃至全球的报纸资源为我所用,为读者提供更优质的服务。

### [参考文献]

[1]何涵慧. 报纸索引和数据库的结合——浅谈报纸索引的现状与未来[J]. 新闻采编, 2007(2): 42-44

(上接第 14页)

#### 4.4 统筹考虑

在采访实践中碰到某一待选近似文献是否去重,或作为准复本购置,要综合考虑各方面因素:分析其价值及本馆读者对该书的潜在需求;近似书内容的相似程度及馆藏数量;是否有不同载体的版本或共建馆已订购,等等。

复本量也需统筹考虑。馆藏复本量少的,要根据该书的潜在利用率和价格等因素来考虑是否增订复本;属于“重点采选”范围且价格适当的,可考虑电子图书和纸质图书同时入藏,但应减少纸质图书入藏复本;对属于“一般采选”范围的,则可考虑只收藏纸质图书或电子图书。若属于本馆“重点采选”范围或分工协调采选范围,但其他图书馆也已订购或收藏,则采选时可减少入藏复本<sup>[5]</sup>。

此外,参与馆际联合采购或地区协调采购的图书馆,对某些文献的采访查重范围还需扩及其他图书馆。对于上文提到的粗心大意容易出错的几种情况,要特别细心辨识、核对,正式发订前要进行“二次查重”,避免误订。必要时,丛书要分别查找丛书名和单册题名,对于有分卷题名的多卷书,既要用总

[2]严三九. 当前我国上网报纸问题探析[J]. 编辑之友, 2000(6): 2-4

[3]王效锋. 网上十大报纸数据库检索系统的比较分析[J]. 图书馆建设, 2002(5): 80-82

[4]宋明亮. 我国报纸信息数据库开发的现状与对策[J]. 中国图书馆学报, 1995(1): 60-65

[5]王效锋. CNKI报纸数据库的收录范围质疑[J]. 四川图书馆学报, 2002(4): 63-66

[6]王效锋. 利用网上报纸资源为读者服务的几点尝试[J]. 情报杂志, 2002(11): 82-83

[7]韩兴雨. 我国报纸网站现状及发展策略研究[D]. 南京: 河海大学, 2007 10

[收稿时间] 2010-07-23

[作者简介]王红(1960—),女,副研究馆员,研究方向:报纸信息开发与利用,发表论文数篇;陈陶(1963—),女,研究馆员,信息咨询部主任,研究方向:信息资源开发与利用,发表论文数篇。

## 题名查重,又要用分卷题名查重。

### [参考文献]

[1]王春秀,盆民辉. 计算机编目查重问题研究[J]. 图书馆学研究, 2006(5): 58-59

[2]吴春. 试谈网络环境下如何提高图书查重的准确率[J]. 图书馆工作与研究, 2006(3): 37-38

[3]罗红丽. 查重小议[J]. 山东图书馆季刊, 2006(4): 58-59

[4]柯霞芳,贡维才. 再谈图书采购查重[J]. 图书馆建设, 2006(3): 43-44

[5]郑兰. 信息时代实体馆藏信息资源的采选[J]. 河北科技图苑, 2007, 20(4): 56-57, 21

[收稿时间] 2010-07-27

[作者简介]李金庆(1962—),男,毕业于武汉大学图书馆学专业,研究馆员,中国索引学会理事、福建省社科信息学会常务理事、厦门市图书馆学会《图书馆探索》副主编,曾在十余省的二十多种省级以上书刊发文六十余篇,出版专著1部;卓晓云(1976—),女,毕业于武汉大学图书馆学专业,馆员,已发表论文数篇。

[收稿日期] 2010-07-13

[作者简介]张文勇(1980—),男,馆员,本科,现任湖南图书馆合作协调部副主任。

(上接第 42页)

[5]莫其强. 大共享理念下高校图书馆信息资源建设的策略[J]. 图书馆学研究, 2009(7): 52-55