

●朱同同

图书馆编目工作原理^{*}

摘要 图书编目,尤其是实行机编,不能只满足于按规则、标准行事,必须有理论作指导。图书馆编目工作原理有四条,即: 确认原理:著录的基本要求是确认在编文献; 标引原理:核心工作是规范控制; 属性原理:书目系统本质上是专家标引系统; 简明原理:编目的最高境界。参考文献 5。

关键词 文献编目 编目工作原理 编目理论 计算机编目

分类号 G254.3

ABSTRACT In this paper, the author thinks that library cataloging, especially computer cataloging, should be done according to not only rules and standards, but also theories. The principles of library cataloging include: 1) confirmation principle: the basic requirement of the description is to confirm the item at hand; 2) indexing principle: authority control is the key; 3) attribute principle: a bibliographical system is essentially an expert indexing system; 4) simplicity principle: it is the highest level of cataloging. 5 refs.

KEY WORDS Cataloging. Cataloging principles. Cataloging theory. Computer cataloging.

CLASS NUMBER G254.3

编目工作原理是有关编目工作的最基本的道理。一般认为,编目难点在“怎么办”,只要熟知国际国内一系列编目规则、条例和标准,问题就可迎刃而解。原理比较宏观,与具体编目关系不大,懂不懂不要紧。然而,这个观念到了今天机编和书目建库的时代,应该改变了。

研制图书馆编目系统,要系统了解编目专业知识。只知其然,不知其所以然,就不可能将编目系统的功能要求搞清楚,也不可能研制出好的编目系统。

机编使编目工作的整体性大大加强。编一条机读目录,要利用机编系统,在广域、共享的联机环境下,按照MARC格式一并完成8大项著录、4种标引、馆藏注记和目录组织等工作,这不是一个手工编目人员一下就能胜任的。

编目的规则、条例和标准等,曾几次修订,目前仍处于不断完善、逐渐成熟的过程中。过去的规定为什么要修订?今后遇到一些矛盾,包括一些规则、条例和标准之间的矛盾,应如何处理?为什么?要找到这些问题的答案,有必要首先在理论上有所总揽。工作原理对于机编工作是非常需要的。

从学科建设角度看,研究编目工作原理也很必要。编目作为一项专业性很强的工作,若只谈如何操作和执行,不谈“为什么”,就不会上升为一门应用学科。

为了便于记忆,也为了更具针对性,我们将编目工作原

理归纳为4条,分别称之为确认原理、标引原理、属性原理和简明原理。

1 确认原理:著录的基本要求是确认在编文献

著录就是对在编文献的识别特征进行描述。描述在编文献识别特征的目的,是为了得到一条能确认该在编文献的记录。手工编目如此,机编也是如此。

如果对一种在编文献的识别特征,描述到足以确认的地步,那么这条记录的著录任务就完成了。书目体系中的每一条记录对其在编文献是否构成明确的一一对应关系,是否有足够的确认意义,是判断这次著录是否成功的基本标准。

从编目员个人操作的层面上看,这确实与“每种文献只做一条记录”的意思差不多。只不过“确认原理”更强调问题的效果:进入书目体系的每一条记录对特定文献的确认性应该是唯一的。

为了确保和提高这种确认性,至少应采取以下专业方法和手段。

1.1 规定描述范围和格式

ISBD(G)中所规定的8大著录项,是世人对文献识别特征的描述范围做的梳理。一般认为,在这8大项范围内,足以完成著录对确认特定在编文献的基本要求。

框定著录描述的范围,不是说每条书目记录都必须做8

注 本文系江苏省“九五”哲学社会科学项目《江苏省中文图书机读目录数据库建设研究》的成果之一。

大项的著录,一个都不能少,而是说无须超范围进行。只要能满足著录对确认特定在编文献的基本要求,少一些著录项并没有什么不好。

8 大项的著录格式,在卡片中是采取一套标识符和段落式方法解决的,在机读目录中是采取机读格式来解决的。这些格式,是框定著录范围的不同手段。

1.2 规定信息源

文献的主要识别特征,在具体文献的不同部位往往不很一致。为了确保不同编目员对同一文献的主要识别特征所做的描述是一致的,国际国内各编目规则对每个著录项的信息来源都做了严格规定。采用规定信息源(或主要信息源)的概念,最终也是为了提高著录描述对具体在编文献的确认性。

1.3 照实著录

照实著录,指对主要著录项的描述,要求严格按在编文献规定信息源上的文字如实进行。即不论原题是繁体字、异体字,还是外文字母、符号、有意义空格,甚至是错字、漏字等,都照实著录。这些主要著录项指题名与责任说明项、版本项、文献特殊细节项、出版发行项和丛编项等5项。

强调5个著录项要按规定信息源照实著录,是为了保证不同编目员著录同一种文献时,在主要识别特征的描述上,保持高度一致。这种高度一致,说到底,还是为了满足著录的基本要求:确认特定在编文献。

注意,照实著录是就文字本身而言的,不是就文字的形式或形态而言的。照实著录的不是文字的形式或形态,如颜色、大小、字体等方面情况,而是文字本身。同时,照实著录不妨碍执行一些通行的书写规则,如省略著录、外文大小写和汉语拼音等。

1.4 确认性的检测等问题

一条书目X对文献A是否构成确认关系,可以用实践来检测。一般认为,根据一条给文献所编的书目X,多次组织进行检索,如果都确切地找到文献A,而不至于误以为所指是另一文献B,那么可以认为,书目X对文献A构成了确认关系。这一检测方法,理论上相当可靠,但很费事,实际上没人会这样做。

编目专家有自己习惯的检测思路:只要严格按编目规则办,不同编目员背靠背地对同一文献进行编目所形成的两条书目记录 X_1 和 X_2 ,其主要著录项的照实著录部分应完全相同,次要著录项可能有详简之别,但应基本一致。否则,不是违反编目规则,就是有笔误。在联机编目、远程相互套录的环境中,这种方法往往更加实用,更有可操作性。

从严格意义上讲,提要作为《文献著录总则》在ISBD(G)基础上增设的著录项,其功能已超出书目范围,涉及了文摘应予解决的任务。为了兼顾我国传统,不妨认为这一延伸书目功能的做法,操作上并无不可,但不能因此模糊

对著录的基本要求的认识。

总之,人们是通过规定范围内文献识别特征的描述,并采用规定信息源和照实著录等方法,以保证对在编文献构成确认关系的。每条书目记录对在编文献的这种确认作用,是构建书目体系,发挥整体书目功效、完成书目任务的基础。

2 标引原理:标引的核心工作是规范控制

一般认为,编目工作主要由两部分组成:一是著录,二是标引。标引可以大致分为四种:名称标引、题名标引、主题标引和分类标引。一条书目记录一旦完成了这四种标引工作,就可以在书目体系与它自己本身之间建立起全面的排检联系。

作为检索词的名称(包括个人名称、团体名称和地名等)、题名和主题,相当一部分有规范形式与非规范形式之分;作为检索点的分类号,一部分会在本馆不用或用其它类号取代。为了建立这些检索词、分类号之间的单纯参照、相关参照和一般参照,需要建立规范档,以便对排检系统实行规范控制,最终确保对书目体系进行检索的有效性和高质量。

2.1 规范控制的方向是自动连接的批处理

在卡片目录中,规范档与书目文档可以有三种关系:一是供工作人员参与的关系,即两者独立,自成系统,书目文档的标目不一定采用规范档的标目,规范档仅作参考,不承担规范控制作用。二是供用户参考的关系,即将规范档作为书目文档的组成部分进行混排。三是规范控制关系,即两者虽自成系统,但要求书目文档的标目一定要采用规范档的标目,规范档承担规范控制的作用。

在机读目录数据库中,规范档可通过电子链针与书目文档自动连接,因此一般不会考虑参考关系,而总是考虑利用机编软件发挥其规范控制作用。但这种自动连接,实际上有个别与批量两种处理方式。

自动连接的个别处理,是由编目员在编制MARC书目记录时,利用机编软件对需要进行规范检查的检索词一个个地用鼠标“划黑”后,再到MARC规范记录里去进行检查处理。这种处理方式比较容易为人接受和理解,但处理效率低一些,会出现规范的遗漏。目前国内研制的“汇文”图书馆编目软件可以完成这种个别性的规范检查。

自动连接的批量处理,是由编目员利用MARC规范记录,定期对书目文档进行批量的连接处理。这是规范控制最理想的连接方式。目前国内尚无此类编目软件。

2.2 规范档的一致性需要确保

规范控制是编目的核心工作。规范档的一致性,在共享性和整体性日益加强的联机编目环境下必须得到确保。目前全国图书馆联合编目中心的规范档格式,与CALIS的不完全一致,给各家图书馆具体应用带来不便。国家标准

部门在这些问题上应该有所作为,做好统一协调工作。

3 属性原理:书目系统本质上是专家标引系统

MARC 数据是一种典型的必须由专家编制的元数据。图书馆书目系统的属性,从本质上看,是专家标引系统而不是自动标引系统。

从构成看,自动标引系统,包括:一是计算机和网络系统,二是应用软件和数据库管理系统,三是大量组织起来的应用数据,四是系统员和自动标引操作工。专家标引系统则不但包括自动标引系统的四方面,而且还包括编制应用数据的专家。

说图书馆书目系统“本质上是专家标引系统而不是自动标引系统”,就是说,不管在图书馆书目系统局部功能自动化方面取得什么进展,从根本上看,它的书目等应用数据的编制,包括数据的描述和标引,最终要依赖编制应用数据的专家来解决,而不能仅靠几个擅长数据自动处理技术的计算机专家解决,更不能仅靠几个数据录入操作工解决。

3.1 书目系统恰恰不同于全文数字化系统

很多人受简单类比思维定势的影响,认为复杂的图书馆全文数字化系统,除了用几个系统员之外,也仅仅用几个工人就操作起来了,为什么比它似乎简单的图书馆书目系统,却非要“额外”增加几个编制 MARC 数据的专家,而且在可以共享、可以购得、可以套录下载 MARC 数据的条件下还一直依赖他们呢?

答曰:图书馆全文数字化系统,是自动标引系统,不是专家标引系统。全文数字化系统的应用数据,可以用扫描加校对的办法获取。在编定了自动标引处理程序的前提下,这些工作只要操作工即可解决,无须依赖专家。将在编文献的识别特征按规定描述下来,再进行标引,形成记录,并组织到书目体系中,这些工作涉及对著录项的理解,涉及对规定信息源的理解,涉及对描述内容的提炼取舍,涉及名称、题名和主题的规范,涉及主题分析,涉及分类标引,涉及对本条记录在整个书目体系中的位置和相互联系的理解。这些分析、理解、描述、规范、标引等工作,虽有编目规则可循,但这些规则有点象写作课上的“小说写法”,必须由人具体掌握执行,并不能完全程序化、机械化,也不是目前自动和智能技术所能解决的。换言之,编制 MARC 数据的规则,包括种种著录、标引和目录组织规则等,尚不能全部转化为数据计算关系,因而 MARC 数据的编制、标引和维护等,最终不能完全自动标引,而必须有赖于机编专家来完成。

3.2 即使可以套录还是需要编目专家

即使利用联机编目条件,可方便地套录下载原始编目机构提供 MARC 数据,图书馆还是需要本馆自己的编目专家。因为 MARC 原始编目单位的立场和工作任务,与各馆

书目建库的立场和任务不完全相同,与一个馆群的联合书目建库的立场和任务也不完全相同:

第一,MARC 原始编目数据上的描述性著录,与图书馆的具体文献不完全一一对应,不完全适用,往往要略做修改后才能适用。如原始记录是指平装两册本,而本馆在编文献却是精装全一册,且两者价格也不同。又如原始记录著录的是一种仍在刊行的完整的连续出版物,而本馆原已收藏其中一辑,现在又新获四辑在编,等等。

第二,原始 MARC 记录可以提供检索字段、规范字段、连接字段,但不可能提供各馆书目体系内部自己的种种有机联系,如层次记录、前后记录、相关记录之间的联系等。

第三,原始编目单位一般不提供规范记录档,也不可能越俎代庖,为各馆建立单纯参照、相关参照和一般参照,进行规范控制,或进行数据质量检查工作。

第四,像“下挂”馆藏注记这样看似简单的操作,也不是缺乏机编专业知识的人员做得好的。编目部“张冠李戴”,流通部不断为“露馅”的“下挂”进行改正,这种情况并不少见。

可见,MARC 书目数据的下载和套用,与其它数据的下载使用不同。此事大中型图书馆必须由 MARC 数据专家进行,小型图书馆至少要有 MARC 数据专家参与把关,否则时间一长,数据差错积重难返,无法收拾。

总之,图书馆书目系统是世界上最复杂的人工记录系统,也是世界上最复杂的专家标引系统。书目数据专家是书目建库的关键。

4 简明原理:简明是编目的最高境界

所谓简明,指简洁和明了。编目有如写文章再编个索引,只要达意、顺畅、明了,则越短越好。

简明原理具体意思是:描述性著录若能达到确认在编文献的要求,则应越简越好,而非越详越好;全面标引太繁太难,不标引又不能发挥书目功能,因此,标引要按书目体系的实际需要来进行。具体以“简明”两字要求即可。

可惜“红楼梦”三字在书目体系中,对题名叫《红楼梦》的某一特定文献还缺乏完整的确认意义,否则不妨承认“红楼梦”三字也构成了一条合格的书目记录。因为“红楼梦”三字不能被确认是指“甲戌本”还是指“脂斋本”,是指人民文学版还是指译林版,是指横排本还是指竖排本,是指繁体字本还是指简体字本,等等。

说简明是编目的最高境界,实际上是给编目提出了一个很高的要求:每条记录不但要达意、顺畅、明了,而且要尽量简洁。

4.1 MARC 格式对著录和标引做了可详可简的设计安排
机读目录中的著录,主要靠标识块里的各文献标识号字段、著录块和附注块来承担。其中,各文献标识号字段和著录块对文献识别特征承担主要描述任务,而附注块则

对文献识别特征承担次要描述任务。

若以绘画艺术来打比方,则著录更像速写,而不是像高清晰度的拍摄。“抓住特征,两三笔勾勒出人物形象特征”的著录,最为理想。但如果简单而起不了确认作用,简单而不可得,就只好详一点,繁一点。比方在附注字段做进一步说明。

机读目录中的标引,目前主要围绕名称、题名、主题和分类等四种标引设计。机读目录的款目连接块、相关题名块、主题分析块、知识责任块等,以及用机读规范格式建档的规范数据,都是用来解决这四种标引及其有关参照的。

对其它一些可能需要标引的文献特征,诸如读者对象、出版周期、文献类型、索引附录、资料规格、文种等,则以代码方式,集中存于编码信息块中。一旦需要,可利用信息技术自动标引处理。

由此可见,机读目录中关于著录和标引两部分的 MARC 格式设计,实际上构成了一种可详可简的安排:著录上分主要描述项和补充描述项,标引上分显性类标引和隐性类标引。这种安排,一方面为发挥机读目录的应有功能和潜在功能提供了充分余地,为编制简洁明了的机读目录提供了充分的执行余地,另一方面,也使得对目录编制质量不宜以“对错”论之,而只能以“完善不完善”、“简明不简明”来判定。

4.2 UNIMARC 对著录详简级次不作规定

一种文献可能著录的识别特征和可能进行的标引,从理论上说是相当多的。为了帮助掌握文献著录的详简分寸,ISBD 对著录详简级次规定分三级,而 UNIMARC 则没有规定(记录头标中的编目等级,是对记录完整程度的描述,不是对著录详简级次规定的要求)。

为什么不规定呢?这与书目记录构成本身有关,也与 MARC 数据的共享本质有关。

书目记录是详是简,这本是一个不宜一概而论,不宜硬性规定,必须靠专家具体掌握的问题。如对“一本专著若有卷端题名要不要标引,若含不一致的丛书名要不要另做标引”等问题,回答一般是肯定的。但谁也不能保证没例外。若卷端题名毫无检索价值,为什么不能免则免,少做一检索点?比方不一致的丛书名形式 B,若因系统首选“前方一致”匹配,而与已成检索点的丛书名形式 A 检索效果相同,为什么要多此一举,再做一检索点呢?书目记录作为文献特征的确认性描述,其详简程度要视在编文献具体情况而定。简而不能满足确认文献之要求,则必详之;详而无补确认性,未增检索之径,属多此一举,则应简之。总之,对著录的详简程度做硬性规定是不科学的。

机读目录是一种社会产品,需投入一定的制作成本。

MARC 本身是一种为了供交换而规定的著录格式。如果没有套录利用或向他人提供编目成果之必要,就不必非按 MARC 格式编目不可。一个供职于某馆的编目员,在直接或间接联机编目的环境下,实际上是为全部联机的图书馆实时编制 MARC 数据;在以软盘、光盘或 E-mail 等非实时方式交换 MARC 数据的环境下,则是非实时地为合作馆编制 MARC 数据。在这种情况下,单独考虑一个图书馆 MARC 数据著录的详简级次,实际上没有什么意义。而且随着编目自动化水平的不断提高,随着利用 MARC 数据的便利程度不断提高,这种意义将日益弱化。何况联机环境和不断交换的环境,本身有一种对每条 MARC 数据不断修订、补充和完善的机制。

从宏观操作角度看,在实时联机或定时共享的编目环境中,与其为每种文献编制三种详简级次的 MARC 数据,不如只编一条来得节约和方便。三种详简级次的 MARC 数据,在共享环境中必然归于统一:既然详的记录是套录一次,简的记录也是套录一次,为什么不套详的呢?

确认原理说了著录的基本要求和目的,标引原理强调了标引业务的核心工作。两个原理又一起概括了编目工作的基本业务内容。

属性原理从数字化高度揭示了书目系统的本质属性。书目的著录和标引,实际上是文献识别特征的描述,数据检索价值的判断,事物指称或类分的规范。编目尽管在许多局部问题、局部工作上,正在从人工处理向自动处理过渡,但从根本上说,它是不可计算、不可编程的。书目数据库有赖于一支机编专家队伍的不断维护。而简明原理则对整个机编专家工作成果提出了一个高度概括的衡量标准。

这四个编目工作原理相辅相成,互为补充,相互渗透,阐明了编目的主要目的、基本要求、工作内容、业务核心、本质属性和衡量标准等,构成了对全部编目工作基本道理的概括说明。

参考文献

- 1 熊光莹. 计算机编目技术手册. 北京:北京图书馆出版社,1999
- 2 刘苏雅. 中文文献编目. 北京:书目文献出版社,1994
- 3 刁维汉. 现代文献编目概论. 上海:华东师范大学出版社,1994
- 4 谢宗昭. 文献编目概论. 南京:南京大学出版社,1990
- 5 李纪有,余惠芳. 图书馆目录. 北京:书目文献出版社,1986

朱同同 南京师范大学图书馆副研究馆员。已发表图书馆学论文 30 余篇,出版著作 4 部。通讯地址:南京师范大学。邮编 210097。

(来稿时间:2001-06-27)