

李致忠

关于古籍联合目录数据库的构建

摘要 建设中国古籍联合目录数据库,涉及一系列准备工作:中国古籍的界定;怎么设置款目;机读格式、著录规则、使用手册的统一;分类法、主题词、书名和著者姓名规范问题。参考文献2。

关键词 中国古籍 联合目录数据库

分类号 G250.74

ABSTRACT: To create a database of union catalog of ancient Chinese books, we should do some preparatory jobs: definition of ancient books, selection of entries, MARC format, cataloging rules, classification Systems, subject headings, authority control, etc. 2 refs.

KEY WORDS: Ancient Chinese books. Database of union catalog.

CLASS NUMBER G250.74

联合目录是文献资源共建的基础,是共享的津梁。但凡欲编制联合目录,不管是哪方面文献的联合目录,也不管是书本式的还是机读式的,都必须先界定其收录范围。中国古籍网上联合目录的编制,非但有界定范围问题,还有古籍定义、目录编法,是编制品种目录还是版本目录,品种目录怎么界定品种,版本目录怎么确定款目,联合多少家图书馆,用什么机读格式,用什么分类工具标引,用什么著录手册以及书目数据中心库的建立、数据使用原则等,都需要事先有个明确的说法。本文不可能对所有问题都一一加以论列,但对其中几个重要问题想谈点看法,以利联合目录编制工作的有效开展。

1 什么是中国古籍

编制中国古籍联合目录好像不存在先界定其收录范围问题,因为说的就是古籍。但由于对什么是中国古籍历来存在着不同认识,所以首先遇到的仍是确定什么是中国古籍。

对什么是中国古籍似有3种说法:抄写或版印在1911年以前的书籍;成书并抄写或版印在1911年以前而又具有中国古典装帧形式的书籍;成书并抄写或版印在1911年以前、以研究中国古代传统文化为主要内容、用中国传统著作方式进行著述而又具有中国古典装帧形式的书籍。这3种不同的界定,就导出了

3种不同的范围。

第1种定义,认为在中国凡抄写或版印在1911年以前的书籍都是中国古籍,这范围就大了。而大就在这么一来,势必将晚清以来编译、介绍西方数理化、声光电、天地生、军科教等西学方面的书籍也囊括了进来。这就不伦不类了。

第2种定义,认为成书并抄写或版印在1911年以前而又具有中国古典装帧形式的书籍都是中国古籍,这范围还是大。清代以来,许多西方传教士来到中国。他们为传播教旨,吸引信徒,影响中国社会,写了不少宣传教义的书籍。这些书籍,论成书和传抄版印时代都在1911年以前,论形式也多是当时流行的线装书。这些书能算是中国古籍吗?

第3种定义,界定比较严谨,也比较符合实际情况。成书并抄写或版印在1911年以前,这讲的是成书和行世的时代界限。以研究中国传统文化为主要内容,讲的是书籍内容主要是谈中国传统学问、传统文化,这是内容方面的界限。用中国传统著作方式进行著述,讲的是书籍著作方式是文是诗是赋,是词是曲是剧,是传是注是笺,是辑是编是摘,是论是评是介,这不仅囊括全部中国古代作品,一些西方人士,或传技与中华各地,或供奉内延撰介西方天文历算、经纬测绘等方面的著作,也可囊括进来,这是著作方式方面的界限。具有中国古典装帧形式,讲的是书籍的外部形态,直到清朝末年,书籍的装帧形式虽然

也有了一些西方的花样,但那是书海中的一叶扁舟,主流还是包背装或线装,这是装帧形式方面的界限。有此4条界限,则什么是中国古籍,其定义明矣,其范围定矣,其联合目录之编制亦行之有据矣。

2 怎么编制古籍联合目录

怎么编制中国古籍联合目录,即怎么设置款目,这是著录、编目之前必须要明确的原则。以品种设置款目的编法,可以叫作品种目录;以版本设置款目的编法,可以叫作版本目录。这两种不同类型的目录,虽然在中国目录学史上都曾出现并流行过,但那时情况简单,又无联合共知共享的需要,所以怎么编制都无所谓。编制联合目录就不同了。编制联合目录的宗旨,是使图书馆和读者共知谁家有什么藏品,以及藏品的版本情况,因而藏品、版本、收藏情况便构成了联合目录的3个基本要素。不管你编什么类型的目录,即不管是品种目录还是版本目录,都要以完成上述这3个基本任务为宗旨。

编制品种联合目录,在款目上,即书目数据记录的条数上,可能会大大简化,但在品种的界定上却会大费周折和煞费苦心;在版本表白、考定、合并与排列上,也会遇到种种难题;在藏家标注与登录方式上也会前后重复,形成藏家简称成排或代码成串,使人眼花缭乱,目不暇接的现象。

仅以品种的界定而论,那就非常不容易。特别是正经、正史,历来加工者众。带有不同加工内容的就应单算一品种。可是古籍情况复杂,有白文,有传注,有单行,有合刻。有原书与一种加工内容合刻,有与两种三种加工内容合刻。有与一人加工内容合刻,有与多人加工内容合刻。有的带释音,有的带释义,有的纂图互注,有的重言重意,有的通考,有的丛说,有的要义,有的辑录,有的长笺,有的纲领,凡此种种,不一而足。要编品种目录,这些问题都要逐一辨析清楚,才能确定款目的废置,谈何容易!并且即使是品种款目确立了,接下来便是在各该款目下著录若干不同的版本。中国古籍,特别是古籍的要籍,如四书五经、前四史、先秦诸子、楚辞文集等,一种书或一种加工,常常是历朝历代各种刻书机构或私宅坊肆都版印行世,形成多种不同的版本。以《文选》为例,将李善注、五臣注、六臣注的各时各种版本列出,光中国国家图书馆善本所藏就有50部之多。要把几十家几百家

图书馆所藏版本都集中起来,再加以条分缕析,合并相同版本,你说得有多难!因为编制联合目录,各家都只能提供款目记录,而对于版本最多提供一些原始描述。机编可能会好一点,可能会更多地提供一些能够反映版本异同的原书特征材料,多给一些判别异同的依据,从而多解决一些问题,但仍会有若干举棋难定、棘手难决的问题。况且,各书各种版本在流传过程中,藏书家在其上进行批校题跋者亦属不少。相同版本有后人批校题跋的亦不宜与无批校题跋者合并。所有这些问题,要想在一个品种款目下逐一加以解决,也要费很多的心血。即便是这些问题都解决了,某一品种款目下某一版本的藏家标准也很难处理。均以藏书单位的简称著录之,会出现著不胜著的黑字现象;均以藏书单位代码著录之,会出现烦人的长串数码现象;款目编序号,版本编序号,藏家编代码,目后编索引,以号索号,又无异于玩数码捉迷藏,也够烦人的。当然,机读目录可能会减少这些烦人的景象。

半个多世纪以来,中国人编制古籍综合目录、专科目录,乃至联合目录,多是版本目录。即在部类之下,以某书的不同版本为设置款目的原则,这样就使上述以品种立目带来的棘手问题迎刃而解。一种书因版本不同可以重立款目,重复是有些重复,但同一品种因版本不同而书名稍异或特异的问题,也迎刃而解。况且一个版本一个款目也好著录藏家。无论在版本款目下直接著录藏家简称或代码,还是款目编序号,藏书单位编代码,以号索码,都比较好安排。当然,机读目录可能会由不同检索点上,解决书本式联合目录难以解决的问题,这就更需要我们精心策划、合理安排了。

3 机读格式、著录规则、使用手册的统一

编制网上可以查询检索的联合目录,首先需要确定的就是使用什么样的机读格式。机读格式确定了,大家才能有所遵循。但哪一种机读格式都是西方人设计出来的适于处理西方文献的格式,现代中文图书使用起来可以勉强对号入座,大体可用。中国古籍的著录处理也使用这种格式就不完全适应。可是现行的通用格式又成为国际标准,各国在处理自己的文献并上网提供检索时,又不能不采用这种格式,这就有了在不违背格式原则的前提下,调整制定适合中国古籍使用的字段表的问题。这个问题各地各馆,乃至海

峡两岸都有了一些尝试,也都取得了一些经验。但由于各地各馆过去手工著录编目的程规并不完全相同,问题技术处理的习惯不同,因而各自使用的字段表也不完全相同。这个问题在实施机读编目之前,必须予以统一。这个问题的统一,看似简单,实则在思想上和理解上却存在不少难点。思想问题主要是指长期手工著录编目形成的定式思想和近几年机编时各自形成的格式定见。这在统一机读目录编目格式时,会引起热烈的讨论和意见纷争。表面上看是公说公有理,婆说婆有理,背景则是这么做我改得多,那么做你改得多,本质是个工作量大小的承担责任问题。在技术处理上的理解认识问题,主要是指一些问题在著录时怎么处理,放在什么字段,在哪做附注等,常常因理解认识不同而引起不同的意见纷争。这些问题需要经过充分讨论,博采众长,最后确定统一的古籍机读目录格式。

国际上制定的各类文献著录标准,行世已经有一段历史了。中国引进这些标准加以翻译理解并制定相应的文献著录规则,经国家技术监督局批准推荐实施,也有一段历史了。其中关于中国古籍的著录标准,早在80年代后期就已被批准为国家标准而荐用。90年代前期又经修改而被批准为国标继续荐用。这个著录规划,既符合国际著录标准原则,又从中国古籍自身的特点出发,做了必要的调整和修改,是一个可用的古籍著录标准。可是迄今为止,由于文献著录标准始终是推荐性的使用标准,而不是强制推行性的使用标准,因此各古籍收藏单位在为自藏古籍进行著录时,用与不用这个标准一由自任。这个问题在中国内地各类型图书馆普遍存在。不编联合目录,这个问题无庸提起,也无碍大局。可是要编联合目录,实现文献信息资源共知共享,这个标准著录规则的使用,至少是联合目录的参编成员馆,便要必须执行。否则就会著录项目不一,位置不一,符号不一,格式不一。这个联合目录也就无法编辑起来了。

机读格式、著录规划都统一了,并不意味着编目人员在著录过程中就没有任何问题了。事实上再好的机读格式和著录规则,也不可能对古籍文献自身所有问题都规定出统一处理的办法,让人在著录时都能一目了然,对号入座。在实践上还会遇到很多问题,让编目著录人员认识不一,理解不一,因而处理的原则、方法和结果也不一。同一种书同一个款目中的相同问题却有各自不同的处理方式,还是个五花八门,这个目录同样是无法编成。这个问题国家图书馆古

籍部门久已注意到,并责成有关人员起草了这方面的使用手册,待在一定范围内征求意见后即可印行试用。试用一段时间之后再搜集意见,进行修订,令其进一步完善,再普遍荐用。

只有在上述机读格式、著录规则和使用手册都基本统一并共识的前提下,才能着手书目数据的制作,并不断输送到书目数据中心库加以总校和统一处理,使之成为规范的书目数据记录,再正式批准入库。这种统一、制作、总校、批准的程序,说起来简单,做起来却不是很容易。这一点,我们务必要有清醒的认识和思想准备。

4 分类法、主题词、名称规范的确定

机读目录重检索点,因而分类就显得不那么重要。但至少在目录情况下,要编制机读联合目录,尚不能取缔分类。中国古籍的编目,从《隋书·经籍志》以降,绝大多数情况下都采用四部分类法。这说明四部分类法有强大的生命力。之所以如此,盖因这样的分类法反映了中国学术发展的基本情况。“《七略》之流为四部,如篆隶之流为行楷,皆势之所不容已者也。史部目繁,不能悉隶以《春秋》家学,四部之不能返《七略》者一;名、墨诸家,后世不复有其支别,四部之不能返《七略》者二;文集炽盛,不能定百家九流之名目,四部之不能返《七略》者三;钞辑之体,既非丛书,又非类出,四部之不能返《七略》者四;评点诗文,亦有似别集而实非别集,似总集而非总集者,四部之不能返《七略》者五。凡一切古无今有、古有今无之书,其势判如霄壤,又安得执《七略》之成法以部次近日之文章乎?然家法不明,著作之所以日下也;部次不精,学术之所以日散也。就四部之成法而能讨论流别,以使之恍然于古人官师合一之故,则文章之病可以稍救,而《七略》之要旨其亦可以有补于古人矣。”^[1]这段议论十分精到,将四部分类法之不能回返《七略》六分法的道理论列得十分明白。然而这仅是说四分法不能返回六分法的一面,却未说四分法尚有不能包容未来的一面。章学诚是乾隆时代的人,无以见到乾隆以后学术的新发展。要编中国古籍联合目录,乾隆以后这100多年的古籍亦不能不予收录。而要收录就不能不部次甲乙,别类部居,以便仍能使人即类求书,因书究学。例如类书,章学诚时便说:“类书自不可称为一子,隋、唐以来编次皆非也。然类书之体亦有二:其有源委者如《文献通考》之类,当附史部故事之后;其无

源委者如《艺文类聚》之类,当附集部总集之后,总不得与子部相混淆。或择其近似者,附其说于杂家之后可矣。^[12]类书不可称子,“总不得与子部相混淆”,这意见是对的。但主张将其附入史部和集部,仍是不伦不类。不如从实际出发,将其从子部剥离,单列为类书部。丛书比类书晚出,但发展却赓续未断。清代之《古今图书集成》似为类书做了一结,此后考据学大兴,那种摘抄、节取各书各类片段文字为类书者,为朴学家们弗信而不取。而搜求各类专书为一书再冠以总名者,为丛书。丛书保持着所收各书的独立性与完整性,为学者所宗信,故乾嘉以后继续发展,乃成为清代学术发展的一大成就,故丛书之行世蔚为大观。其内容性质则跨学科跨部类,非经非史非子非集,放在子部亦属不伦不类。也可以并且已经从子部剥离出来,单列为丛书部。这样就形成了经、史、子、集、类、丛六部。至于每部当中的属、类之设置,亦应做到科学的调整和必要的增删,形成新的类表,再给以科学的配号。这样古籍联合书目数据库的构建,在分类上才算有了张本。

在古籍联合书目数据库的构建过程中,对所收每种古籍在著录编目时要否能否给予主题标引,这是将来检索时的重要检索点,它比分类更具体更准确,更敷实用。但现行的现代文献的主题词表,不能直接拿来就用,这就又牵扯出一个古籍编目标引的词表问题。古籍分类法的类目标名,有些就可以作为主题词使用,但多数显得宽泛,机检一调出来一大堆书,仍然让人无从检览。能否在上述六分法的基础上,将类表列出,在各个相关的类目下,加详加细主题词。如子部—天文历算—历法—具注历;子部—天文历算—历法—麟德历;集部—曲类—诸宫调—刘知远;集部—曲类—传奇—刘知远白兔记;集部—曲类—诸宫调—董西厢;集部—曲类—杂剧—北西厢;集部—曲类—传奇—南西厢等等,便可标引出最后一层的主题。免得一调天文,则天文历算类的书目数据全部显出;一调历法,则各种历书的书目数据全部显出;一调西厢记,则诸宫调、杂剧、传奇等各剧种的西厢记书目数据全部显出,令人目不暇接,无以检询。但作出这么一个主题词表,亦非易事,必须事先有人起草试作,再经过讨论修订,方可使用。这是基本建设,我们应该知

难而进,接受新事物,创造新局面。

最后还有个名称规范问题。名称规范是指书名和著者姓名规范问题。同名异书或同书异名,这类现象为数不少。以哪个为规范标目,这牵涉到以书名为检索点时的检索问题。最麻烦和讨厌的是古书在传刻过程中,书名常常被穿靴戴帽,弄得人检索时不知从哪下手。如《古本董解元西厢记》、《董解元西厢记》、《新刊合并董解元西厢记》、《元本出相北西厢记》、《新刻徐文长公参订西厢记》、《新订徐文长先生批点音释北西厢》、《李卓吾先生批评北西厢记》、《李卓吾先生批点西厢记真本》、《重刻元本题评音释西厢记》、《鼎镌陈眉公先生批评西厢记》、《新刊改正全像评释北西厢记》、《张深之先生正北西厢秘本》、《重刊订正元本批点画意北西厢》、《田水月山房西厢藏本》等等,名目繁多,花样翻新。这些书著录时会照卷端题名客观描述,但在作为检索点时却不可不予规范。否则就无从下手,无法检索了。

古书著者姓名的情况也很复杂:有题名的,有题字的,有题号的,也有题地名、山名、水名,乃至室名斋号的。著录时可以客观照录,附注时必得说明其真实姓名或通用题名,以便规范,供再著录该人其他著者时作为划一的依据。这项基础工作国家图书馆已操作有年,积累著者4万余人,基本资料已经具备,数据格式也已敲定,只待形成数据与编目挂接,即可提供使用。

总之,构建中国古籍联合书目数据库,说起来话简,办起来事繁。凡事预则立,不预则废。与其想得简单,配套的准确工作未备好就仓猝上马,招致返工;不如全面筹划,分头齐做,分工协作,步步为营,有序推进更好。

参考文献

- 1.2 〔清〕章学诚.校雠通义·宗刘第二.见:王重民.校雠通义通释.上海:上海古籍出版社,1987

李致忠 全国政协委员,国家图书馆研究馆员。通讯地址:北京白石桥路39号国家图书馆。邮编100081。

(来稿时间:2000-03-24)