

# 《中国古代著者规范数据库》编制概况

唱春莲

(国家图书馆善本特藏部善本组)

进入信息化时代,图书馆作为传统意义上的主要信息提供源,正在将自己丰富的资源通过网络及便于交流的方式向社会提供,以满足社会需求,并在信息化社会中继续充当重要信息源的角色。数字化图书馆正是在这一形势下以一种新的模式迅速形成。为顺应形势的发展需要,我们在制作馆藏善本古籍书目数据库基础上,开始了《中国古代著者规范数据库》的制作工作。国图这个工作项目,旨在将散落在各处的中国古代著者资料加以集中整合,建成一个完整的古代著者规范数据库,为数字化图书馆工作提供可链接的古代著者规范信息,为古文献书目数字化及社会需求服务。

早在八十年代中期,善本部领导及有关人员已经意识到编制中国古代著者规范文档是一项方便学界,造福未来的工作,即开始筹划、拟定相关条例,并安排有关业务组人员负责具体编制工作。至九十年代末,共整理出约四万条古籍著者名称及有关著者规范资料约二百条。

由于业务工作调整,自2000年始,善本组与馆业务处正式签订了编制《中国古代著者规范数据库》项目,是善本组的主要业务工作项目之一。当年完成六千余条。2001年完成五千余条。预计今年完成八千至一万条。

## 一、收录范围、原则及内容

《中国古代著者规范数据库》收录范围包括1911年辛亥革命以前出生的,有过著作或诗文,至今有文字记载的著者简要生平资料。

著者规范名称的确立是本数据库的重要内容,一般为人们所熟悉的通行名称。如人们熟知的中国古代四大名著之一《红楼梦》的作者曹雪芹,名曹霑,号雪芹,以号行,字梦阮,别号芹圃、芹溪。此条数据的规范名称就是“曹雪芹”,而其名、字、别号则入附注项说明,并分别在根查项反映。古代僧人的规范名称要在法号前加“释”字,其他名号资料同样入附注项及根查项反映。有些生活在朝代交替之际的人,他们的事迹资料往往被跨时代的典籍记载,在这种情况下,一般以此人登科、入仕或主要活动的时代为准,认真做好认定和查重工作。对古代著者的称谓要准确,一般以为后世普遍认可的为准。如:政治家、文学家、科学家、艺术家、诗人等等,不可随意冠名。对于生卒年不详的著者,其生平事迹的内容依具体情况尽量详细。

每条规范数据内容主要包括:著者名称、性别(女性要注明)、生卒年(用阿拉伯数字记公元年,生年或卒年不详者用问号代替,均不详者注明“生卒年不详”字样)、字号(包括本名、字、号、别名、别称等)、籍贯、主要事迹、著述(如作者无著作而仅存诗、文作

品的也收录并注明)等。其中在姓名、性别、字号、生卒年、籍贯、著述等方面,经查考、勾沉获得的资料,能详则详,全部反映。而对于著者的生平事迹,则重点反映登科及第情况、学术方向、仕途的主要经历(初官、末官、官至最高级)等,其余则从略。

每条数据字数根据资料内容的多寡决定,一般在一、二百字左右。本数据库主要是古代著者名称规范文档,因此有的作者所用字、号、别称虽然很多,为方便读者检索,也要全部录入,不得省略。有的著者著述很多,不论其存世与否,只要有确切记载都要反映,以全面揭示著者的著述情况。

我们在查找资料、确立著者规范名称时,遇到著作人姓名、字号均无考者,一般不予收录。

## 二、使用软件及格式

本数据库使用丹诚图书馆系统软件,依据 UNIMARC 规范数据格式录入。

具体录入格式内容如下:

记录头标

001 记录标识号

100##\$a 通用处理数据

152##\$a 编目条例代码

200#0\$a 个人名称(主要名称。如姓名或字、号、别称等)\$c 名称的附加成分(女性要注明)\$f 生卒年(阿拉伯数字表示,生或卒年不详者以“?”代替,均不详者注“生卒年不详”)

200#0\$7 另外一种文字的代码\$a 另外一种文字的个人名称(汉语拼音)

3000#\$a 附注内容(著者生平事迹)

400#0\$a 单纯参照根查个人名称(除主要名称外的其他名称。注:一个名称做一个此字段)

400#0\$6 连接说明码、连接号\$7 另外一种文字的代码\$a 另外一种文字的个人名称(汉语拼音)

801#0\$a 国家代码\$b 机构名称和代码\$c 处理日期

810##\$a 参考数据来源(数据内容的来源、出处。注明书名、卷次)

规范数据样例一:

01072nx 2200193 45

-01/著者规范库/ctlno/0001402|94673c3c7a0500001f

001????????????

100 @a20000701achiy0108 ea

152 @aBDAB

200 0@a 王庭珪@f (1080~1172)

200 0@7ba@a

3000 @a 一作廷珪，字民瞻，自号卢溪真逸，吉州安福（今属江西）人。徽宗政和八年（1118）进士，调衡州茶陵县丞。宣和末年退居乡里。后除直敷文阁，传世有《卢溪集》，此外著述颇多，均已佚。

400 0@6a04@a 廷珪

400 0@6a04@7ba@a

400 0@6a05@a 民瞻

400 0@6a05@7ba@a

400 0@6a06@a 卢溪真逸

400 0@6a06@7ba@a

801 0@aCN@bNLC@c20001109

810 @a 《宋史翼》卷七十

规范数据样例二：

01072nx 2200193 45

-01/著者规范库/ctlno/0001631|b4673c3c5f0600001f

001????????????

100 @a20000701achiy0108 ea

152 @aBDAB

200 0@a 王闾@f (? ~1146)

200 0@7ba@a

3000 @a 字无功，明州慈溪（今浙江慈溪东南）人。早年曾举进士不第，弃而学佛，修净土宗。高宗绍兴十六年卒。有《净土自信录》。

400 0@6a04@a 无功

400 0@6a04@7ba@a

801 0@aCN@bNLC@c20001109

810 @a 《宋诗纪事》卷十六

### 三、工作步骤及今后的设想

纵观我国几千年的文明发展史，其间著书立说者很多，他们分布在社会的各个领域，全方位记录了人类社会文明的发展进程。根据我国历史文化悠久，历代文人层出不穷的特点，我们首先以此为切入点，先期搜集各朝代的诗文作家资料。到目前为止，我们已经完成秦、汉、三国、唐、宋、清代的部分著者规范数据的制作，计有二万余条数据。下一步将计划完成制作元、明期间的诗文著者规范数据。以后再分别将古代在政治、军事、经济、科技、医

药、艺术等方面的有过著述的人物资料陆续补充进去。这样，采用纵横交替的方法，即纵向：收集各个历史时期的著者资料，横向：收集每一时代在各个不同领域里有过著述的著者资料，从而尽可能地将我国所有古代著者收录进去，为学界服务，为后世留下宝贵资料。

古代著者的资料散落在浩如烟海的古经典籍中，查找这些古代著者的有关资料十分不易。为确保这一项目高质量早日完成，我们在搜集整理古代著者资料的工作中，充分利用前人的研究成果是十分必要的。我们参考了有关古代人物传记资料的各种工具书，如二十五史传记人名索引、各种史书中的艺文志、各类古籍书目、有关人名词典等，也审慎选用近些年出版的有关古籍整理方面的书籍，如《全唐文》、《全唐诗》、《全宋诗》等，从中确定著者规范名称，编制规范记录。

在搜集资料过程中，对比和查重工作必不可少。在使用前人有关研究成果时，我们尽量选取那些经过有关专家学者认真考证撰写的，具有一定权威性的古代著者传记资料，按照规范数据的标准要求，从中加以遴选、补充，力求使每一条规范数据详实可靠。

同时，每条数据的810字段注明资料来源。在此字段中，我们尽可能反映本数据资料的原始出处，如果原始记载不可考，则将后人研究成果的作品名称标明，以备读者查阅使用。在提供给读者使用本数据之前，还要认真核实，确保资料来源准确。

随着工作的进展，古代著者收录范围不断在扩大，可利用的前人研究成果将越来越少，更多的著者资料需要我们在古代典籍中挖掘、研究。因此，大量艰巨的查考整理和规范工作还需要我们去完成。

总之，《中国古代著者规范数据库》的建立将会给古籍书目数字化和专家学者的研究工作带来很大方便。对于我们来说，编撰制作这一数据库本身，也是一个学习和研究过程。要把在中国数千年的文明发展进程中留下记载的著者全部收编，确实不是一件容易的事情。在实际工作中往往会出现各类预想不到的问题，及时发现，及时解决，用科学严谨的态度对待这项工作，是我们所遵循的原则。因此，由于数据量大，时间有限，我们已完成的数据中难免存在一些不完善之处，我们会随时加以修正。实际上，随着有关专家学者对各个领域研究的深入，不断会有新的发现。因此，《中国古代著者规范数据库》的内容不断得到修订、补充，是必然的，也是正常的。

此项工作开展不久，即被列入海峡两岸五地图书馆文献数字化合作项目。目前香港等地图书馆也在尝试制作中国著者规范数据库，我们希望尽早在古代著者规范数据库方面进行交流、合作，以便达成中国著者规范数据内容和形式的真正统一规范，最终实现全球中文文献信息资源共享，也希望业内人士对我们的工作予以帮助，多提宝贵意见及建议。