

Te Zhong Wen Xian
Ji Du Shu Mu Shu Ju

韦衣昶 著

特种文献 机读书目数据

TEZHONG
WENXIAN
JIDUSHUMUSHUJU

北京图书馆出版社

特种文献机读书目数据

韦昶 著

北京图书馆出版社

图书在版编目(CIP)数据

特种文献机读书目数据/韦衣昶著. —北京:北京图书馆出版社, 2002. 12

ISBN 7-5013-2064-0

I. 特… II. 韦… III. 机器可读目录—文献著录 IV. G254.364

中国版本图书馆CIP数据核字(2002)第100486号

书名 特种文献机读书目数据

著者 韦衣昶 著

出版 北京图书馆出版社 (100034 北京西城区文津街7号)

发行 (010)66126153 传真(010)66174391

E-mail Btsfxb@publicf.nlc.gov.cn

Website www.nlcpress.com

经销 新华书店

印刷 北京华正印刷厂

开本 850×1168 毫米 1/32

印张 11.375

版次 2002年12月第1版 2002年12月第1次印刷

字数 260(千字)

印数 1—3000

书号 ISBN 7-5013-2064-0/G·529

定价 22.00 元

前 言

书目记录规范化、标准化是实现信息资源共享的前提。信息资源,通常是指一切以记录形式存在的信息载体,其范围广泛、种类多样。就信息资源的基本类型而言,它可以是书籍、图书、期刊、报纸、论文、标准、档案等各种类型;就存在的形式而言,它可以是各种传统的印刷型资源,也可以是缩微型、机读型资源以及非实体资源——网络信息等。而只有对这些不同类型的信息进行规范化的组织,才能够被人们更好的利用。

普通图书、期刊在这方面经过了比较长的发展,现已进入了成熟阶段。但其它类型信息资源的标准化道路还是曲折漫长的。笔者结合自己多年来编目的实践及教学和科研成果,以《文献著录总则》、《普通图书著录规则》、《连续出版物著录规则》、《非书资料著录规则》、《地图资料著录规则》、《档案著录规则》、《古籍著录规则》、《检索期刊条目著录规则》等文献著录标准为依据,尝试用中国机读目录格式对特种文献进行著录。

本书具有实践性强,内容新颖,适用范围广的特点。本书以文献编目原理为指导,从实际工作出发,突出编目技术和方法的实践性,书中大量的文献著录实例,使读者易于接受。本书既适用于本科专业教育和专科教育以及专业技术人员培训,也适用于档案、出版和情报部门。

由于笔者才疏学浅,疏漏之处不少。恳请编目界的专家、学者、师长、学友批评指正。

作 者

2002年5月于福建师范大学

目 录

| | |
|--------------------|-----|
| 绪 论 | 1 |
| 第一节 记录头标 | 5 |
| 第二节 标识块 | 11 |
| 第三节 编码信息块 | 15 |
| 第一章 古籍机读书目数据 | 32 |
| 第一节 古籍基本知识 | 32 |
| 第二节 记录头标、标识块及编码信息块 | 48 |
| 第三节 著录信息块的著录 | 64 |
| 第四节 附注信息块的著录 | 78 |
| 第五节 连接款目块的著录 | 87 |
| 第六节 相关题名块的著录 | 97 |
| 第七节 主题分析块的著录 | 101 |
| 第八节 知识责任块的著录 | 110 |
| 第九节 国际国内使用块的著录 | 115 |
| 第十节 金石拓片的著录 | 117 |
| 第二章 非书资料机读书目数据 | 151 |
| 第一节 音像资料的著录 | 151 |
| 第二节 缩微资料的著录 | 191 |
| 第三节 电子资源的著录 | 200 |
| 第三章 其他类型文献的机读书目数据 | 213 |
| 第一节 测绘制图资料的著录 | 213 |
| 第二节 乐谱的著录 | 243 |
| 第三节 标准文献的著录 | 268 |

| | | |
|--------|----------------------------|-----|
| 第四节 | 科技报告的著录 | 275 |
| 第五节 | 学位论文的著录 | 279 |
| 第六节 | 会议文献的著录 | 285 |
| 第七节 | 技术档案的著录 | 290 |
| 第八节 | 专利文献、产品样本及零散资料的整理与编目 | 294 |
| 附 录 | | 298 |
| 一、 | 中国机读目录格式字段表 | 298 |
| 二、 | 中国机读目录格式编码信息块常用代码表 | 329 |
| 三、 | 古籍著录业务知识用表 | 343 |
| 四、 | 名称规范说明 | 355 |
| 主要参考文献 | | 357 |

绪 论

机读目录格式(Machine Readable Cataloging Format)是根据文献特点和文献机构之间信息交换的需要建立的标准化的计算机可读形式。在机读目录中,一条书目记录(record)相当于手工目录中的一条款目,它是对于一种文献相关信息的完整记录。记录按一定顺序排列而成的集合称为文件(file),内容与时间上具有完整意义的一个文件,相当于一个功能很强的手工目录体系。机读目录文件经过计算机加工处理,可以按照需要输出题名、责任者、主题、分类等多种目录。每一条机读书目记录都由四个部分组成,即:

| | | | |
|-------|-------|-------|-------|
| 记录头标区 | 地址目次区 | 数据字段区 | 记录分隔符 |
|-------|-------|-------|-------|

1. 记录头标区

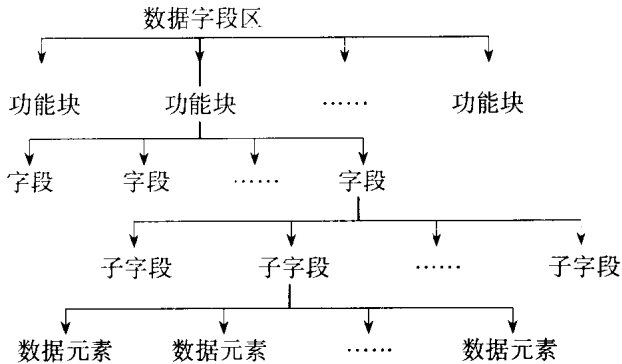
记录头标区包含处理记录时可能需要的有关记录的一般信息。整个头标区为24个字符,由固定长数据元素组成,并通过字符位置标识。

2. 地址目次区

地址目次区的各段(12个字符为一段)以各字段的字段标识符(3位字符)为首。依次记录每一字段的长度(4位字符)和起始位置(5位字符),因此它的作用类似于文献的目次。地址目次区的数值全部由计算机自动生成。

3. 数据字段区

在地址目次区之后为数据字段区,其中的数据按五个层次组织起来。结构图示如下:



(1) 功能块

CNMARC 参照 UNIMARC, 并根据我国文献著录的实际, 将数据字段区划分为 10 大功能块, 具体如下:

0-- 标识块: 共分 12 个字段, 包括记录和文献的标识号码(如控制号、ISBN、ISSN 等)

1-- 编码信息块: 共分 7 个字段, 包括描述文献的各个方面的定长数据元素(通常是编码数据)

2-- 著录信息块: 共分 6 个字段, 包括 ISBD 和我国《著录总则》规定的除附注项和文献标准编号以外的全部著录项目。

3-- 附注块: 共分 24 个字段, 包括对文献各方面的文字说明, 除少数附注(如内容附注)外, 没有预先规定的格式。

4-- 连接款目块: 共分 31 个字段, 包括以数字和文字形式(如 ISBN、ISSN、识别题名、丛编附加款目、责任者/题名附加款目)对其他记录的标准连接。此外, 还可包括相关记录的控制号对其他记录的连接。

5-- 相关题名块: 共分 16 个字段, 包括可作为检索点的本文档的其他题名(通常出现在文献上)。

6-- 主题分析块: 共分 17 个字段, 包括分类及主题标识(如《中图法》分类号、DDC 分类号、汉语主题词, 以及作为主题的个人

名称和团体名称等)。

7-- 责任者块:共分9个字段,包括对文献负有责任的个人及团体的名称。这些名称分为个人名称、家族名称及团体名称,并进一步区分为主要责任者、等同责任者及次要责任者。

8-- 国际使用块:共分2个字段,包括对记录负有责任的机构标识等字段。

9-- 国内使用块:只有馆藏信息1个字段,保留给国内使用。

字段可包含1个子字段,也可包含多个子字段。如200字段就含有正题名、一般资料标识、另一责任者的正题名、并列正题名、副题名及其说明题名文字、第一责任说明、其他责任说明等15个子字段。有些字段不能重复,有些字段可以重复,子字段也是如此。

在CNMARC中,以下字段为必备字段,即:

001—记录标识号

100—一般处理数据(仅为某些数据元素)

101—作品语种(当作品有语种时)

200—题名与责任说明

801—记录来源

其他字段的取舍,取决于各类型文献转换为机读形式的具体记录。记录数据内容由编目条例和负责建立记录的书目机构实际执行的规范所决定的,即数据元素的有无不仅由格式的要求确定,而且由国家编目条例和实际执行的规范确定。

(2)子字段与数据元素

所谓子字段,泛指字段内所定义的数据单元(数据元)。而数据元素则是被明确标识的数据最小单元。在可变长字段内,数据元素构成子字段,并用子字段标识符标识。在头标区、目次区、和固定长子字段内,由代码构成的数据元素由其字符所在位置标识。

(3)标识系统

数据字段区采用的标识符号有三种:标识符,又叫内容标识符(Content designator);指示符(Indicator);分隔符(Separator)。

标识符位于数据的前位,标志数据开始。标识符包括字段标识符(Tag)、字段指示符(Indicator)和子字段标识符(Subfield identifier)。字段标识符由三位数字组成,用于识别不同的字段。其中第一位数字用以区分不同的功能块。如第一位数字为“2”,代表著录信息块;为“3”,代表附注块。也就是说,在同一个功能块中,不同字段的第一位数字均相同。例如,“200”标识题名与责任者项,“225”标识丛编项。子字段标识符由两位字符组成。第一个字符为国际标准 ISO646 中规定的专用符号 IS1,格式文本中用“\$”或“@”表示;第二个字符可以是数字,也可以是字母。子字段标识符用以表示子字段,例如,在 200 字段中,“\$a”表示正题名,“\$e”代表副题名,“\$f”代表责任者;在 701 字段中,“\$4”代表责任方式。在目录产品输出时,计算机自动把字段标识符和子字段标识符转换成相应的著录用标识符号。如“200”字段在中“第一责任者”的子字段标识符“\$f”自动转换成“/”,“副题名”的子字段标识符“\$e”自动转换为“:”等。

指示符由两位数字或字母,位于字段标识符之后。指示符只出现在可变长字段中,向计算机提供某些数据的操作信息。暂时没有定义的指示符用空格填充,格式文本中记作“#”。例如,101 文献语种字段中,第一个指示符为翻译指示符,可以取值“0”、“1”或“2”;第二个指示符没有定义,填空“#”。

分隔符位于数据的结尾,用于分隔不同的数据集合。分隔符包括字段分隔符和记录分隔符两种。字段分隔符(Field separator)用于分隔字段,它采用 ISO646 中规定的专用字符 IS2,在格式文本中出现时,用“*”。记录分隔符(Record separator)为 ISO646 中规定的专用字符 IS3,在格式文本中用“#”表示。记录分隔符位于每个记录的结尾处,用以分隔记录。两种分隔符均由计

算机自动生成。

本书主要尝试讨论特种文献机读书目数据的编制技术。包括古籍(含拓片)、非书资料(含音像资料、缩微制品、机读文件)和其他类型文献(含地图、乐谱、标准、科技报告、学位论文、会议文献、技术档案、专利资料、产品样本等)。但无论这些文献如何特殊,它们都与普通图书有着共同之处。因此,本书是在中文普通图书机读格式的基础上重点讨论适用于特种文献的机读格式字段和子字段,未做详细说明的有关字段请参看《中国机读目录格式使用手册》(修订版)。

在机读目录格式中,无论那种文献数据记录都必须具有以下三个部分:

记录头标(Record Label)

标识块(Identification Block)

编码信息块(Coded Information Block)

在这三部分中有适用于各种文献的字段,也有只适用某一类文献的专用字段。绪论讨论的是记录头标、标识块和编码信息块中的适用于各种文献的通用字段部分。

第一节 记录头标

记录头标是按照 ISO - 2709 的规定设立的,它包含处理记录时可能需要的有关记录的一般信息。记录头标位于每个记录之首,它是必备的和不可重复的。记录头标没有字段标识符、子字段标识符和指示符。整个头标区长度为 24 个字符,字符位置规定为从 0 至 23,由固定长数据元素组成,这些数据元素通过字符位置来标识。

记录头标定长数据元素表:

| 数据元素名称 | 字符数 | 字符位置 |
|-------------|-----|-------|
| (1)记录长度 | 5 | 0-4 |
| (2)记录状态 | 1 | 5 |
| (3)执行代码 | 4 | 6-9 |
| (4)指示符长度 | 1 | 10 |
| (5)子字段标识符长度 | 1 | 11 |
| (6)数据基地址 | 5 | 12-16 |
| (7)记录附加定义 | 3 | 17-19 |
| (8)地址目次项结构 | 4 | 20-23 |

(1)记录长度(字符位置0-4):5个十进制数字,右边对齐,不足部分用零补齐。记录长度表示整个记录的字符长,包括记录头标、地址目次区及全部数据字段。本项数据由计算机自动产生。

(2)记录状态(字符位置5):用1位字母代码表示记录处理状态。

c = 修改过的记录:对原已发行且记录状态为n,o,p的记录,经修改更新后的记录,其记录状态位置应置c,以替换原记录。

d = 被删除的记录:任何一条记录不再有效时,其代码应改为d,并可在300附注字段说明其原因。

n = 新记录:凡新建档的记录,包括出版前的在版编目记录,其代码均用n,只有一种情况除外,即记录如适用代码o时,则优先用o。

o = 较高层次的记录:当新记录层次较低时,即另有一比其层次高的记录存在时,该新记录用代码o,此时层次等级填2。

p = 曾为不完整的预编记录:当一以前发行的预编记录现为出版后的记录所取代时,其代码改为p。

(3)执行代码(字符位置6-9)

①记录类型(字符位置6):常用的记录类型代码如下:

- a = 文字资料(印刷品)
- b = 文字资料(手稿)
- c = 乐谱(印刷品)
- d = 乐谱(手稿)
- e = 地图和天体图资料(印刷品)
- f = 地图和天体图资料(手稿)
- g = 投影和录像资料(电影、胶卷、幻灯片、录像)
- i = 录音(非音乐节目)
- j = 录音(音乐节目)
- k = 二维图像(图片、设计图等)
- l = 计算机载体
- m = 多媒体资料
- r = 三维手工制品和实观教具
- u = 拓片
- v = 善本书

根据文献制作情况,选择记录类型代码。如刻本、活字本、铅印本、石印本、影印本、油印本、拓片等选印刷型“a”,稿本、写本、抄本、墨迹等选手写型“b”,有特殊要求必须区别善本书时选“v”;地球仪选“r”,光盘选“l”。若一文献包含两种记录类型,如手写和拓片两部分,可根据主要部分或第一部分的制作方法选择“b”或“u”。

②书目级别(字符位置7):机读目录格式定义了四种可用值:

a = 分析型文献:该记录所描述的实体在物理形态上包含在另一实体之中,如单行本古籍的某一部分、专著中的一个章节、会议录中的一篇论文、丛帖子目等。

c = 汇集型文献:人为配套的著作集,如汇集在一个函套内的古籍资料,某作家的全部手稿集、集拓本、合裱本等。

m = 专著:以一册或若干分卷册出全的出版物。如单部专著、单幅拓片、多卷集或丛书的单册等。

s = 连续出版物:以连续的分卷册并计划无限期连续出版的出版物。如期刊、报纸、丛刊、年鉴、专著丛编等。

根据所著录文献个体的分合状况,在以上的级别中选择一个。

③层次等级(字符位置8):本代码标识本记录以层次性的关系与其它记录连接,并指明它在层次中的相对位置。

= 未定义层次级别:对多层次出版物不使用多层次著录方法,或不拟拦截各层次记录。如丛编中的单册与丛编,若不设丛编检索点,亦未建丛编记录。

0 = 无层次的记录,既本记录与其他记录无任何关联。

1 = 最高层的记录,当一条记录高于与它有关联的其他任何记录时,该记录即为最高层记录,如丛编记录。记录中有向下连接的462、463、464字段,必须选“1”。

2 = 低于高层次的记录(在最高层以下的记录)。当另有一条与本记录有关联、且高于本记录的记录存在时,本记录即为低层次记录。如丛编中的单册,当建立丛编记录时,该单册记录层次等级代码值为“2”,此时,字符位置5(记录状态)填为0。记录中有向上连接461、462、463字段,必须选“2”。

④未定义(字符位置9):填空格“#”。

(4)指示符长度(字符位置10):表示指示符长度的1位十进制数字。本格式指示符占2位字符长,故取值为2。

(5)子字段标识符长度(字符位置11):表示子字段标识符长度的1位十进制数字,本格式子字段标识符占2位字符长,故取值为2。

(6)数据基地址(字符位置12-16):本数据基地址表示第一个数据字段相对于记录的起始字符位置,占5个字符位,用十进制数表示。右边对齐,不足部分用零补全。它等于头标区和目次区的字符总数。由计算机自动生成。

(7)记录附加定义(字符位置17-19):占3个字符位,为处理

记录提供必要的详细信息。

①编目等级(字符位置 17):用 1 位字符代码表示机读记录的完整程度,以及建立该记录时是否依据原文献。

= 完全级:表示建立该记录时依据原文献。

1 = 次级 1:表示建立该记录时未依据原文献,如根据目录卡片编制的机读目录。

2 = 次级 2:表示该记录是预编(在版编目)记录,这种记录一般不完整,如载体形态字段不完整或没有。

3 = 次级 3:表示记录为非完整编目的记录。

②著录格式(字符位置 18):用 1 位字符代码表示记录采用的著录规则。它表明在著录 200 ~ 225 字段时是否遵循 ISBD 规则。

= 完全采用 ISBD 格式

i = 部分采用 ISBD 格式

n = 未采用 ISBD 格式

③未定义(字符位置 19):填充格“#”。

(8)地址目次区结构(字符位置 20 - 23):此区为每一个字段的目次款目和结构提供详细的说明。4 个字符位的内容如下:

①“数据字段长度”的长度(字符位置 20):用 1 位十进制数字标识每个地址目次款目里的“数据字段长度”部分的字符数。本格式的数据字段长度占 4 位字符,故取值为 4。

②“起始字符位置”的长度(字符位置 21):用 1 个十进制数字标识每个地址目次款目里的“起始字符位置”部分的字符数。本格式的起始字符位置占 5 位字符,故取值为 5。

③“执行定义部分”长度(字符位置 22):用 1 位十进制数,标识每个地址目次款目里“执行定义部分”所含字符数。由于本格式地址目次区中不含这部分,故本格式取值为 0。

④未定义(字符位置 23):填充格“#”。

以上数据元素中,字符位置 0 ~ 4、12 ~ 16 系数字资料,10、11、

20~23 系特定数值,可由计算机自动产生或预先设定,其余数据元素则需依据具体著录实体而定。在著录时,一律使用单字节,字母代码用小写,同时注意屏幕上缺省值的修改。

例 1:00735nvm0#2200079###450#(表示一个无层次关系的单行本善本古籍新记录的头标)

| 字符位 | 值 | 注释 |
|-------|-------|-------------------|
| 0-4 | 00735 | 记录长度 |
| 5 | n | 新记录 |
| 6 | v | 记录类型—善本古籍 |
| 7 | m | 书目级别—专著 |
| 8 | 0 | 层次等级—无层次关系 |
| 9 | # | 未定义 |
| 10 | 2 | 指示符长度 |
| 11 | 2 | 子字段标识符长度 |
| 12-16 | 00079 | 数据基地址 |
| 17 | # | 编目等级—完全级 |
| 18 | # | 著录格式—完全采用 ISBD 格式 |
| 19 | # | 未定义 |
| 20 | 4 | “数据字段长度”的长度 |
| 21 | 5 | “起始字符位置”的长度 |
| 22 | 0 | “执行定义部分”的长度 |
| 23 | # | 未定义 |

例 2:00681oaa2#2200265###450#(表示一个低层次的丛帖子目新记录的头标)

| 字符位 | 值 | 注释 |
|-----|-------|--------|
| 0-4 | 00681 | 记录长度 |
| 5 | o | 较高层新记录 |

| | | |
|---------|-------|-------------------|
| 6 | a | 记录类型—拓片 |
| 7 | a | 记录级别—分析级 |
| 8 | 2 | 层次等级—低级 |
| 9 | # | 未定义 |
| 10 | 2 | 指示符长度 |
| 11 | 2 | 子字段标识符长度 |
| 12 - 16 | 00265 | 数据基地址 |
| 17 | 1 | 编目等级—次级 1 |
| 18 | # | 著录格式—完全采用 ISBD 格式 |
| 19 | # | 未定义 |
| 20 | 4 | “数据字段长度”的长度 |
| 21 | 5 | “起始字符位置”的长度 |
| 22 | 0 | “执行定义部分”的长度 |
| 23 | # | 未定义 |

地址目次区

地址目次区的各段(12 个字符为一段)以各字段的字段标识符(3 位字符)为首。依次记录每一字段的长度(4 位字符)和起始位置(5 位字符),因此它的作用类似于文献的目次。由于每条记录所录的字段数不等,所以地址目次区所占的字符数长度不定,因此,需要在其终结处加一字段分隔符表示该区终结。例:

001001100000010002500011020001700036... *

地址目次区的数值全部由计算机自动生成。

第二节 标识块

本块包含标识记录或标识文献实体并记载在实体上的号码。