

PUL MARC Format

中文机读编目格式

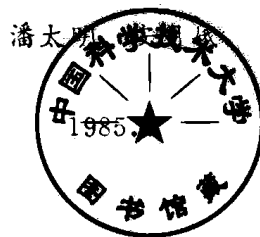
北 京 大 学
图 书 馆 自 动 化 研 究 室

1985.3

前 言

《中文机读编目格式》是在《国际机读编目格式》(UNIMARC 1980 版)的基础上,根据《文献著录总则》(GB 3792.1—83)及其他有关的国家标准和国际标准,结合我校图书馆的具体情况编写而成的。适用于图书、期刊等资料。由于我们查阅的有关标准不完全,再加能力和专业知识的限制,格式中一定会有不妥甚至错误之处,欢迎批评指正。

在整个编制过程中,得到馆领导以及陈绍业先生、韩荣宇先生、成素梅主任的热情指导,邵科同志提供了期刊部份作例用的详细材料,同时还得到馆内各部门、中编组各位同志的大力帮助,在此表示衷心的感谢。



1474/29

Note

PUL-MARC format is edited according to UNIMARC format and most of the national standards of China, and the specific conditions of Peking University Library. It is applicable to the MARC processing of monographs, serials, and parts of the non-book materials in Chinese.

As the national and international standards consulted are incomplete there must be some imperfections and defects in the format. We sincerely hope our colleagues and readers send us their suggestions and advices, so that we can correct and improve them in our practice.

We would like to express our acknowledgment here to Mr. Pan Yongxiang, deputy director of our library, Mr. Chen Shaoye, Mr. Han Ronyu, and Ms. Cheng Sumei, who gave enthusiastic supports and guidance to us in the editing of the format.

Pan Taiming. An Yue'e

目 录

定义.....	(1)
注释.....	(3)
中文机读编目格式.....	(5)
1. 交换格式结构框图.....	(5)
2. 记录头标区.....	(5)
3. 目次区.....	(8)
4. 数据区.....	(10)
5. 0——标识段.....	(13)
6. 1——编码数据段.....	(21)
7. 2——著录资料段.....	(67)
8. 3——附注段.....	(75)
9. 4——连接款目段.....	(87)
10. 5——相关题名段.....	(97)
11. 6——主题分析段.....	(107)
12. 7——责任者段.....	(115)
13. 8——国际使用段.....	(121)
14. 9——国家使用段.....	(311)
附录:	
附录一: 卡片格式举例.....	(123)

附录二：字段、子字段一览表·····	(125)
附录三：国家（或地区）代码表·····	(151)
附录四：地区代码表·····	(174)
附录五：语言代码表·····	(212)
附录六：货币代码表·····	(233)
附录七：年代范围代码表·····	(248)
附录八：读者对象代码表·····	(250)
附录九：中文罗马字母拼写法（节录）·····	(251)
附录十：中国历代纪元表·····	(254)
附录十一：文献类型与文献载体代码表·····	(268)
附录十二：国内地区代码表·····	(271)
附录十三：国内收藏单位代码表（待定）·····	(308)
附： 北大图书馆馆藏资料字段和代码表·····	(309)
A. 馆藏资料字段·····	(309)
B. 分馆（室）代码表 ·····	(313)
C. 特藏符号表·····	(314)

定 义

文献目录记录——简称记录。由记录头标区、目次区和一个或多个数据字段组成。

结构——构成一个记录的各个组成部分的安排。

记录头标区——位于记录开始处的部份。在该区内给出了处理本记录的有关参数。

目次区——记录中所有数据字段起始位置的索引。

目次区说明——说明目次区内各登录项结构的一组参数。

数据字段——含有一个具体范畴的数据，在记录中属于可变长部份。位于目次区之后，并与目次区内的一个登录项相对应。一个数据字段可包含一个或多个子字段。

分隔符——为一控制符号，用来从逻辑上或等级上分隔和限定数据单位。

内容标志符——标识数据元素的符号，或者用来提供有关数据元素的附加信息。内容标志符包括字段标识符、指示符、和子字段标识符。

字段——一种特定的字符串。用字段标识符识别，包含一个或多个子字段。

字段标识符——用来标识有关字段的三位字符的字符串。

定长字段——长度不变的字段。本格式对全部数位内容都做了规定。一个定长字段可以包含一个或几个子字段。每一子字段又可包含一个或多个数据元。

变长字段——长度可变的一种字段，其长度随数据元素的具体内容（包括指示符，子字段标识符和字段分隔符）而定。变长字段可以含有一个或多个数据元素或子字段。

字段分隔符——一位控制符，置于每一变长字段的末尾，用来分隔本字段与下一字段。

指示符——与变长字段相关的符号（数字或字母）。它可以为字段内容、记录中的一字段同另一字段的关系、以及某些数据处理过程中所需要的操作提供附加信息。

本书目实体——200字段所描述的文献。即著录资料源。亦称本著录实体。

相关书目实体——与本书目实体有关的作品。例如：特定卷向上联系的丛书或向下连接的册分析。

子字段——字段中的一种特定的信息单位（见数据元素）

2 中文机读编目格式

定长子字段——长度不变的子字段。根据本格式的规定可具体确定有关的全部内容。一个定长子字段可以包含一个或多个数据元。它可以用在定长字段中，如字段100的子字段\$a,也可以用在变长字段中，如字段200的\$z。

子字段标识符——在变长字段中，为识别各个子字段，用两位字符所组成的代码。

数据元（素）——资料的最小单位。在可变长字段中，由数据元素构成子字段，并用子字段标识符加以识别；在记录头标区、目次区和定长子字段中，数据元由它们的字符位置来识别。

分层级别——简称层级。按逻辑记录分配给本书目实体的值。此值表明该实体在分层结构中所处的位置。例如：多卷集个别卷的书目叙述同多卷集本身的书目叙述，可以看做是处在两个彼此相关的不同层级上。当一个为另一个的组成部份时，只有整个书目实体及其各别部份都制有书目记录，才可使用层级。例如，期刊中的论文，专著中的章节，丛编的附属丛编，以及丛书的单册。

注 释

- (1) 对于不同媒体的同一数据元素,尽可能分配给相同的子字段标识符。
- (2) 子字段标识符首先赋予的是字母,需要时可以扩展成数字。对于子字段标识符的顺序,未作规定。
- (3) 说明性资料包含在附注中,不准备用做检索点。
- (4) 编码资料(多种媒体类型的共同数据元素,以及媒体特性的数据元素)要按位标示在编码数据字段中。
- (5) 必备字段——机读记录中必需提供的字段:
- 001——记录标识
 - 100——一般处理数据
 - 101——作品语言(如果作品有文字)
 - 123——编码数据字段:地图和天体图资料—比例尺和坐标
 - 200——题名和责任者叙述
 - 206——资料特殊细节项:地图和天体图资料—数学数据
 - 801——编制机构字段
- 其他字段的有无,不只是决定于格式的要求,而更主要的是由编目机构的规范所决定。
- (6) 字段和子字段“可重复”。
- 如果“可重复”指的是字段,则在记录中该字段可以出现一次以上;如果“可重复”指的是子字段标识符,则在字段中该子字段可以出现一次以上。
- (7) 补缺符号(Fill character)采用竖线“|”表示,用来代替编目机构不能判明的内容标识符或编码信息。
- (8) 美元符号(\$)代替ISO646的IS₂符号,用作子字段标识符的第一个符号,如果在资料中出现了实际的美元符,则在其下加一横线,即(\$),以示两者的区别。
- (9) 有关指示符的“未做规定”,指的是对该指示符位未赋值。
- (10) 各例中的符号b用来表示空格,字段分隔符没有明显表示出来。
- (11) 单位(Single)记录的使用:

4 中文机读编目格式

有关同一书目题名的几个不同期次的编目数据，可以包含在一个单位记录中。

- (12) 如果字段或子字段只适用于某些载体，则在字段名称之后（用括号），或在子字段注释栏内注明载体类型。
- (13) IFLA——国际图书馆协会联合会的简称。全名为：International Federation of Library Associations and Institutions.
- (14) ISO——国际标准化组织，全名为：International Organization for Standardization.
- (15) ISBD——国际标准书目著录（规则）。全名为：International Standard bibliographical description.
- (16) ISBN——国际标准书号。全名为：International Standard Book Number.
- (17) ISSN——国际标准连续出版物号。全名为：International Standard Serial Number.
- (18) ISDS——国际连续出版物资料系统。全名为：International Serials Data System.
- (19) GB——中华人民共和国国家标准。

中文机读编目格式

1. 交换格式结构框图

交换格式包括记录头标区、目次区、数据字段区和记录分隔符等四个部份。各部份的位置如下图所示：

记 录 头 标 区
目 次 区
数 据 字 段 区
记 录 分 隔 符

头标区：固定长，长度规定为24个字符。

目次区：可变长。

数据区：可变长。

记录分隔符：标志本记录结束。

2. 记录头标区—记录标示

头标区为24个字符长的定长区，其结构为：

字符位置：0 ~ 4	记 录 长 度	↑ 记 录 头 标 区 ↓
5	记 录 状 态	
6 ~ 9	执 行 代 码	
10	指 示 符 长 度	
11	子字段标识符长度	
12~16	数 据 基 地 址	↑ 目 次 区 说 明 ↓
17~19	记 录 附 加 注 释	
20	每个登录项中“数据字段长度”的长度	
21	每个登录项中“起始字符位置”的长度	
22	每个登录项中“限定执行部份”的长度	
23	未 定	

定长数据元表

数据元名称	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) 记录附加注释	3	17~19
(8) 目次区说明	4	20~23

字符位置

(1) 记录长度	0 ~ 4
----------	-------

用五位十进制数，表示一个文献记录的字符总数。即头标区、目次区、数据区和记录分隔符所占字符位的总和。右边对齐，不足五位时（左边）填零。

(2) 记录状态	5
----------	---

代码：

- c = 修改过的记录
- d = 已删除的记录
- n = 新记录
- o = 以前为较高层级的记录
- p = 以前为不完整的在版记录

(3) 执行代码	6 ~ 9
----------	-------

①记录类型 代码：

6

- a = 印刷式文字资料
- b = 手稿式文字资料
- c = 印刷式乐谱
- d = 手稿式乐谱
- e = 印刷式地图和天体图资料
- f = 手稿式地图和天体图资料
- g = 放映和录象资料
- h = 缩微制品

i = 非音乐性录音
 j = 音乐性录音
 k = 平面(二维)图(图片, 图纸等)
 l = 计算机存贮介质
 m = 多种媒体

②书目级别 代码:

7

m = 专题出版物——单行本或成套图书
 s = 连续出版物——定期或不定期
 a = 单册分析——本书目源为另一书目源的一部份, 其物理形态并不独立。
 c = 收集品, 如某一作者的全部原稿。

③分层级别

8

该代码提供具有分层关系的各记录之间的相对位置(见连接款目段的等级分析部份)。

代码: b = 分层关系未定

0 = 无分层关系

1 = 最高层级记录

2 = 低于最高层级记录(最高层次以下的全部等级)

④未定。填空格符。

9

(4)指示符长度

10

用一位十进制数表示指示符所占字符的位数。本格式取值为2。

(5)子字段标识符长度

11

用一位十进制数, 表示子字段标识符的字符位数。本格式取值为2。

(6)数据基地址

12~16

占五个字符位, 用十进制数表示。所记内容等于头标区和目次区的字符总数。右边对齐, 不足五位时左边填零。

(7)记录附加注释

17~19

①著录级次

17

用一位字符代码说明记录的详简程度。规定:

b = 空格。标志在版编目记录。

1 = 简要级次(只著录主要项目)。即第一著录级次。

8 中文机读编目格式

2 = 基本级次(著录主要项及部份选择项),即第二著录级次。

3 = 详细级次(包含主要项及全部选择项),即第三著录级次。

②著录格式

18

用一位字符代码,表示记录中的各描述项是否采用ISBD格式。

代码: b = 完全采用ISBD格式

n = 不采用ISBD格式

i = 部分或者不完全采用ISBD格式。

③外文书目语言

19

代码: b 无

L 拉丁语

E 英语

R 俄语

A 阿拉伯语

S 西班牙语

F 法语

J 日语

G 德语

M 多种语言

(8)目次区说明

20~23

①每一登录项中“数据字段长度”的长度

20

用一位十进制数,指明目次区内每一登录项的“数据字段长度”所采用的字符位数。本格式取值为4。

②每个登录项中“起始字符位置”的长度

21

用一位十进制数,指明目次区内每一登录项的数据字段“起始字符位置”所采用的字符位数。本格式取值为5。

③每个登录项中“限定执行部份”的长度

22

用一位十进制数,指明目次区内每一登录项中“限定执行起始位置”所采用的字符位数。本格式允许取值为0或5。当取值为“0”时,说明无限定执行部份。

④未定。填空格符。

23

3.目次区

目次区为可变长。结构如下图示:

字符位数: 3	字段标识符	↑	↑
4	数据字段的长度	↑	↑
5	起始字符位置	↑	↑
(5)	限定执行起始位置(任选)	↓	↓
12 或 17	登 录 项	↓	↓
12 或 17	登 录 项	↓	↓
	⋮		
1	字 段 分 隔 符	↓	↓

目次区包括若干个登录项和区末的字段分隔符(标志本字段结束)。每一登录项对应于数据区内的一个数据字段。登录项由字段标识符(3位)、数据字段的长度(4位)、起始字符位置(5位)、限定执行起始位置(5位。如果需要的话)等几部份组成。每一登录项的字符数可取12位(无限定执行部份)、也可取17位,但在一个目次区内的所有登录项必需采用同一值。登录项的排列顺序是依照字段标识符的取值由小到大。

3.1 字段标识符

用三位十进制数字,标识相对应的数据字段。其取值为001~999。

- 0 -- 标识段
- 1 -- 编码数据段
- 2 -- 著录资料段
- 3 -- 注释段
- 4 -- 连接款目段
- 5 -- 相关题名段
- 6 -- 主题分析段
- 7 -- 责任者段
- 8 -- 留作国际使用段
- 9 -- 国家使用段(未定)

3.2 数据字段的长度

用四位十进制数,表示本登录项中“字段标识符”所示数据字段的字符总数(包

括指示符和字段分隔符)。右边对齐,不足四位时左边填零。

3.3 起始字符位置

用五位十进制数,表示本登录项中“字段标识符”所示数据字段的第一个字符,相对于数据基地址的位置。目次区的第一个“起始字符位置”为“0”,(即数据区的第一个字段的起始位置为“0”)。其后各项中的“起始字符位置”,都从第一个数据字段的起始位置算起。数字要右边对齐,不足五位时左边填零。

3.4 限定执行起始位置

用五位十进制数,表示本登录项内“字段标识符”所示数据字段的限定执行部份的第一个字符,相对于数据基地址的位置。数据字段中的限定执行部份,记有外文书目资料,置于整个字段的尾部。如果本记录无外文书目资料,则目次区就不存在“限定执行(部份的)起始位置”。如果本记录的某些字段含有外文书目资料,而本登录项所指数据字段没有,则本登录项中的“限定执行起始位置”填“99999”。

4. 数据区

本区由若干个数据字段组成,每一数据字段都以一个字段分隔符结束。在数据区之后置一记录分隔符,表示该记录结束。

本格式有三种不同类形的数据字段:

记录识别字段,字段标识符为001;

保留字段,字段标识符为002—009;

书目数据字段,字段标识符为010~999。

4.1 记录识别字段(001)

本字段记有识别本记录所需要的编号。编号均由编制文献记录的机构确定。

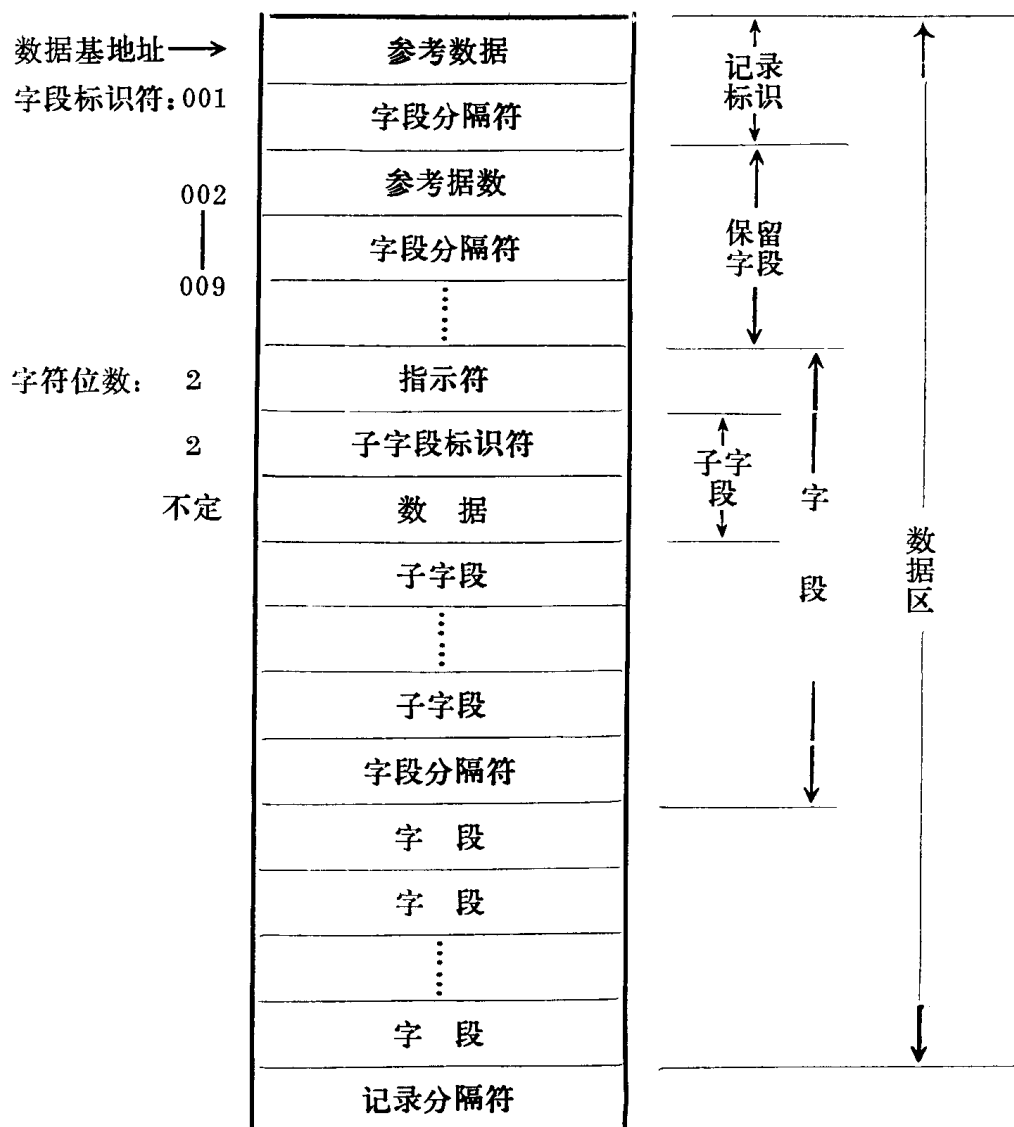
该字段不设指示符和子字段。

4.2 保留字段(002~009)

留做提供未来资料处理所需要的信息。这类字段不设指示符和子字段。

4.3 数据字段(010~999)——文献目录字段

此类字段的结构如下图示:



从上图可以看出，一个数据字段包括指示符（取 2 位长）、若干子字段和末尾的字段分隔符。

子字段由子字段标识符和数据两部份组成。子字段标识符的第一位为字符 \$，第二位为内容特征字符。

