

A User Guide for USMARC Format

USMARC 格式使用手册

谢琴芳 刘素清 徐 韶 合编

高等教育文献保障系统管理中心

北京大学图书馆

A User Guide for USMARC Format

USMARC 格式使用手册

谢琴芳 刘素清 徐 韶 合编

高等教育文献保障系统管理中心
北京大学图书馆
1998年7月16日

前言

图书馆历来依赖目录提供馆藏信息服务,可以说,图书馆目录是与图书馆共存的。最早的书本式目录只能提供题名和著者等少量检索点,而且出版周期长,更新困难,其局限性是很大的。卡片式目录显然比书本式目录进了一步,通过合理的目录组织,能使读者有较多的途径获取馆藏书目信息。但是由于目录柜空间、款目制作和维护的人力和财力的限制,以及读者必须到图书馆才能检索等因素的限制,图书馆资源的开发和利用仍然受到目录系统的限制。

随着计算机技术在图书馆的应用, MARC 格式应运而生。MARC 是 Machine Readable Catalogue 的缩略形式。通过 MARC,可以把人能够识别的书目信息,使计算机也能识别;计算机又能通过 MARC 格式,把 MARC 记录信息反馈成人所需要的各种书目形式或书目产品。如果一个图书馆的馆藏书目都回溯转换成 MARC 记录,图书馆就具备了基于 MARC 记录处理后可能开展的各种服务。如联机公共目录检索、新书报导、目录产品、统计报表、馆际资源共享等。MARC 记录编制质量越高,自动化管理系统越成熟,读者和图书馆工作人员从 MARC 中获益就越多。

USMARC 格式由美国国会图书馆开发,故又称 LCMARC。最初开发的格式(MARCI)较多地照顾了编制程序的方便,与传统编目方法差距较大,为此又作了较大的修改,于 1968 年 1 月正式出版,称之为 MARCII。1969 年 3 月 MARCII 磁带记录正式发行,成为图书馆发展的一个里程碑。70 年代初, MARCII 被接纳为国际标准,其后陆续出版了专著、连续出版物、档案与手稿资料、视觉资料、乐谱和地图等资料类型的 MARC 格式文件。到 80 年代,原来分开出版的 USMARC 格式合并成活页,不断更新出版。为了适应多载体资料编目工作的需要,简化格式,清除在实际工作中使用格式时会出现的歧意和偏差,1988 年美国国会图书馆网络开发和 MARC 标准办公室颁布了 USMARC 一体化的书目格式,并于 1993 年底付诸实践。所谓一体化,就是对各种资料类型均提供一种统一的书目格式,当编制某种类型的书目记录时,在格式中选择相应的数据代码、相应的字段、子字段和指示符即可。1988 年版的一体化格式为非印刷型资料 and 含有多多种资料类型的文献的编目提供了较完善的处理方案,但是对与书目文献有关的网络资源和电子资源缺少连接机制,而 1995 年的一体化格式更新版较好地解决了这一问题。

编者自 80 年代后期就投身于 USMARC 格式的研究和应用。深知基于 MARC 格式建立的数据库在图书馆自动化建设中的重要性。数据库中的 MARC 记录好比手工目录中的款目,它们都是编目工作的产物,就其性质而言,并无本质上的区别。但是,由于 MARC 记录以计算机技术为检索手段,记录中的任何信息,都有可能作为检索点,为了最大

程度地开发馆藏信息，就要求 MARC 记录所包容的信息量比传统目录更全面、更丰富、更准确；MARC 记录必须被计算机识别、阅读和处理，才能提供基于 MARC 的各种优质服务，只有遵循格式规定编制的 MARC 记录，才能达到预期的效果；MARC 记录的最高层次服务是资源共享，为了使不同语种、不同类型的信息资源能在不同国度互相传递和交流，必须使用统一的标准。因此 MARC 记录编制过程中所要求的标准化、规范化程度要远远高于传统编目的要求。编制一条高质量的 MARC 记录，实际是严格执行编目规则和书目信息交换标准的过程。这里虽然强调了 MARC 格式的重要性，但并不想把读者引入掌握了 MARC 格式，就能编制高质量 MARC 记录的误区。

本手册是作为“211”工程高等教育文献保障系统西文图书联合目录培训班的讲义而编写的，旨在为全国高校建立西文书目数据库和处理西文书目数据提供较为规范的参照和依据。在编写过程中，参照了自 1988 年以来美国国会图书馆出版发行的各种 USMARC 书目格式，包括美国国会馆网页 (<http://www.locweb.loc.gov>) 上的相关文件和最新更改信息，同时，根据我国西文编目的实践和经验，从实用的角度，对 USMARC 书目格式作了较全面的介绍和分析。

由于时间紧迫，来不及征求同行专家的意见就匆匆排版印刷，难免在内容上和录入编辑方面出错，若有疏漏和不当之处，欢迎批评指正。

编者

1998 年 7 月 13 日

目 次 表

前言	v-vi
记录头标	1
001-006 控制字段	7
001 控制号	7
003 控制号标识	8
005 最近的处理日期和时间	8
006 定长数据元素—附件特征	9
007 载体形态定长字段	15
地图	15
计算机文件	16
球仪	17
放映图象	18
缩微资料	20
非放映图象	22
电影	23
录音	26
文本	29
录象	29
资料类型不明	31
008 定长数据元素	33
008/00-17, 35-39 字符位数据	33
008/18-34 字符位	35
图书	35
连续出版物	39
视听资料	43
混合资料	45
地图	45
乐谱	48
计算机文件	52
01X-09X 号码和代码字段	54
010 国会图书馆 (LC) 控制号	55
020 国际标准书号 (ISBN)	56
022 国际标准连续出版物号 (ISSN)	58
024 其它标准号	59
027 标准技术报告号 (STRN)	60
030 缩称代码 (CODEN)	60
034 制图编码数学数据	61

目 次 表

040 编目来源	63
041 语种代码	64
043 地理区域代码	65
050 美国国会图书馆索书号	66
082 杜威十进分类号	67
1XX 主要款目标目	70
100 主要款目—个人名称	70
110 主要款目—团体名称	72
111 主要款目—会议名称	74
130 主要款目—统一题名	76
20X-24X 题名字段	79
210 缩略题名	79
222 识别题名	80
240 统一题名	82
245 题名说明	84
246 变异题名	86
247 前题名	90
250-29X 版本、出版说明字段	92
250 版本说明	92
254 乐谱演奏类型说明	93
255 制图数学数据	94
256 计算机文件特征	95
260 出版发行等项	96
263 计划出版日期	98
3XX 载体形态项等字段	100
300 载体形态	100
310 当前出版频率	102
321 先前出版频率	103
362 出版日期卷期标识	104
4XX 丛编说明	106
440 丛编说明/附加款目—题名	106
490 丛编说明	107
5XX 附注字段	110
500 一般附注字段	110
501 合订附注	112
502 学位论文附注	112
504 书目等附注	113

目 次 表

505 格式化内容附注	114
6XX 主题检索字段	117
600 主题附加款目—个人名称	117
610 主题附加款目—团体名称	119
611 主题附加款目—会议名称	121
630 主题附加款目—统一题名	123
650 主题附加款目—主题标引词	124
651 主题附加款目—地理名称	126
700-75X 附加款目字段	129
700 附加款目—个人名称	129
710 附加款目—团体名称	131
711 附加款目—会议名称	133
730 附加款目—统一题名	134
740 附加款目—非控制相关/分析题名	136
76X-79X 连接款目字段	138
760 主从编款目	141
762 附属丛编款目	142
765 原著款目	143
767 译著款目	144
770 补编/特辑款目	145
772 正编记录款目	146
773 主文献款目	147
775 其它版本款目	148
776 其它载体形态款目	149
777 合订款目	149
780 先前款目	150
785 后续款目	152
787 其它相关款目	154
80X-84X 丛编附加款目	155
800 丛编附加款目—个人名称	155
810 丛编附加款目—团体名称	156
811 丛编附加款目—会议名称	157
830 丛编附加款目—统一题名	158
841-88X 馆藏、交替图形等字段	159
856 电子信息检索字段	159
附录: USMARC 字段一览表	163

记录头标(Leader)

记录头标，简称头标，位于每个记录的开头，固定长度为 24 个字符位置(00-23)，数据元素由一系列按字符位定义的数字代码和字母代码组成，提供对记录进行处理的参数。

每个记录都必须具备头标，但不可重复。头标也不包含字段标识符、字段指示符和子字段代码。头标包含以下数据元素：

数据元素名称	字符数	字符位置	备注
逻辑记录长度	5	00-04	由计算机自动生成
记录状态	1	05	代码
记录类型	1	06	代码
书目级别	1	07	代码
控制类型	1	08	代码
未定义	1	09	空位
指示符长度	1	10	由计算机自动生成
子字段代码长度	1	11	由计算机自动生成
数据基地址	5	12-16	由计算机自动生成
编目等级	1	17	代码
著录格式	1	18	代码
连接记录要求	1	19	代码
款目布局	4	20-23	由计算机自动生成

除了由计算机自动生成数据元素的字符位外，其它字符位一般由编目员在建立记录时根据编目资料选择合适的代码。

逻辑记录长度(00-04)

表示整个记录的字符总数（包括它自身和记录终止符），由 5 个十进制数字组成，若记录总数不足 5 位数，则右边对齐，左边空位填零。逻辑记录长度由计算机自动生成。

记录状态(05)

1 位字母代码，反映书目记录的维护状态，05 字符位共定义 5 种代码：

a 由简编升级的记录

表示由最简级记录升级而成的记录。最简级指依据一定的编目规则编制的简编级记录。

c 修改过的记录

表示修改过或更新过的记录。例如,对原有记录追加了信息或者对其中的一处或数处数据作了修改,对简编记录的升级不用代码 c,而用代码 a。

d 删除的记录

表示该记录不再有效。这里删除的含义指的是逻辑删除,而不是物理删除。

n 新记录

表示进入书目文档后一直没有被修改过的记录。新记录包括编目文献在出版前就编制的记录,如在版编目记录,也包括几十年前就已建立,但进库后没有作过任何修改的记录。

p 由在版编目升级的记录

表示由在版编目记录(CIP)或由美国国家连续出版物协调规划(NSDP)记录升级而成的记录。对简编记录的升级不用代码 p,而用代码 a

记录类型(06)

1 位字母代码定义记录类型。本字符位定义 14 种记录类型(缩微制品不包含在内)。计算机通过记录类型代码识别书目文档中的书目类型。在联机检索中,这一代码可以作为检索中包含或排斥某类特殊介质的基础。但是记录类型代码提供的是一般信息,要提高检索资料的专指性还需结合头标其它代码或可变控制字段的相关代码来限定。

a 印刷型文字资料 (Language material)

记录内容是印刷型语言资料,包括印刷资料的缩微品,但不包括手稿型文字资料。

c 印刷型乐谱 (Printed music)

记录内容是印刷型乐谱或者是印刷型乐谱的缩微品。

d 手稿型乐谱 (Manuscript music)

记录内容是手稿型乐谱或者是手稿型乐谱的缩微品。

e 印刷型地图 (Printed map)

记录内容是印刷型地图或者是印刷型地图的缩微品。

f 手稿型地图 (Manuscript map)

记录内容是手稿型地图或者是手稿型地图的缩微品。

g 放映资料 (Projected medium)

记录内容包括电影,录像制品(video recording),胶卷(film strip),幻灯片,透明片(transparency),投影资料(包括档案投影资料)

i 非音乐性录音资料 (Nonmusical sound recording)

记录内容为讲话、演讲等非音乐性录音资料。

j 音乐性录音资料 (Musical sound recording)

记录内容为音乐录音资料。

k 二维非放映图形 (Two-dimensional nonprojectable graphic)

记录内容指图表(charts)、抽象派拼贴画(collages),计算机图形(computer graphics)、素描(drawing)、闪视卡(flash cards)、油画(paintings)、复制品的原件(copy masters)、负片(photonegatives)等平面图形及其复制品,也包括档案图形资料。

m 计算机文件 (Computer file)

记录内容指可以被计算机处理的信息体。计算机文件类型包括数值型、文本型、计算机软

件以及这些类型的组合。

o 多载体配套资料 (Kit)

由两种或两种以上资料类型组成的混合型资料,无法确定以哪一种资料为主时,可以定义为多载体配套资料。例如实验室配套资料,一套包含教科书、练习册、指南等资料的教材,一套包括试题、答案、评分标准等内容的考试资料。

p 混合型资料 (Mixed material)

记录内容主要指档案与手稿

r 三维制品或自然生成物 (Three-dimensional artifact or naturally occurring object)

记录内容指人工制品,如模型、智力玩具、雕塑品、三维展品、机器、服装等。自然生成物,如显微样品、标本等。

t 手稿型文字资料 (Manuscript language material)

书目级别(07)

1 位字母代码,说明书目的级别,共定义 6 种书目级别。

a 专著分析级(专著的组成部分)

编目实体是某专著的析出部分。如编目实体是专著中的某一章节,某篇文章。

b 连续出版物分析级(连续出版物的组成部分)

编目实体是某连续出版物的析出部分。如某刊物的一篇文章,一个栏目或专辑。

c 合集

指汇集性书目实体。这些资料不一定是一起出版发行或生产的,有时为了管理之便而把一些资料汇集在一起。如个人手稿汇总在一起形成一个编目实体、将不同形式的资料按某种目的汇集在一起。合集的书目记录应以其主要书目实体为描述对象。

d 子集

编映实体是某个总集的析出部分。子集的书目记录中应包含描述子集的数据元素以及标识主体资料(总集)款目的数据(773 字段)

m 专著

一册或多册按一定计划出版的资料,通常指单部书或多卷书。如单行本专著、多卷本专著、单张地图、地图集、专著丛编、单份手稿等。

s 连续出版物

以连续的分册或卷期无限期连续出版的出版物。包括现刊、绝版刊、报纸等。

控制类型(8)

1 位字符代码表示控制类型,共定义 2 种代码。

6 无控制类型

a 档案控制

指示符长度(10)

1 位数字代码,说明字段指示符的长度。本格式的指示符长度为 2。

子字段代码长度(11)

1 位数字代码,说明子字段代码(如 \$a)的长度。本格式的子字段代码长度为 2。

数据基地址(12-16)

由 5 位十进制数组成,说明 MARC 通讯格式中第 1 个数据字段的起始字符位置。如某条 MARC 记录,其头标固定为 24 个字符位,其日次区为 24 个字符位,那么 MARC 记录中第一个数据字段的起始符位置为 261 字符位。数据基地址由系统自动生成。

编目等级(17)

1 个字符位代码,说明书目记录内的书目信息或内容标识的完整程度,共设 8 种代码。编目级别主要用于合作编目和套录数据的用户对源数据进行识别。

6 完全级

指最完整的 MARC 记录。编制记录时与编目实体核对过。若是连续出版物,必须至少核对过其中的一期。

1 完全级(未核实文献实体)

虽然是最完整的 MARC 记录,但未与编目实体核对过,主要用于回溯转换的记录。

2 次完全级(未核实文献实体)**4 核心级(Core level)****5 部分级**

未经编目机构审定的正式记录,一般指在制作过程中的中间记录,如因采购新书需临时建立的一条记录,为实现计算机借书而临时建立的一条记录等,该记录有可能被编目机构升级为一条符合要求的记录,也有可能不被升级而被其它级别的记录取代。

7 最简级

按照简化级编目要求编制的记录,标目符合规范化要求。

8 预编级

指在版编目记录(CIP)或由美国国家连续出版物协调规划项目(NSDP)编制的记录。

u 级别不详**z 不适用****著录标准(18)**

1 个字符位代码说明编制记录依据的标准和规则,共设 4 种代码。

6 非 ISBD 标准

主要指 ISBD 标准颁布(1974 年)前使用的 AACRI、ALA 编目规则等。有时也称 ISBD 前的标准。

a AACR2

指依据英美编目条例第二版编制的记录。

i ISBD

依据国标准书目著录(ISBD)编制的记录,其标目的选择原则可能与 AACR2 不一致。

u 标准不详

若标准不详,使用代码 u,若使用的标准要作进一步说明,可选择 040 字段 \$e(著录条例)

连接记录要求(19)

1 个字符位的代码说明是否能根据连接款目字段(76X-78X)产生相关记录附注。共设 2 个

代码:

x 不要求相关记录

用于以下两种情况: ①当书目记录中没有连接款目字段, ②记录中有 1 个或多个连接款目字段, 都著录有相关记录的基本识别信息, 或者有连接款目综合附注(580 字段)

r 要求相关记录

当记录中有连接款目字段, 但未著录相关记录的基本识别信息, 也没有连接款目综合附注时, 选用 r。

款目布局(20-23)

说明 MARC 记录中每个字段目次区的结构。由 4 个字符位的数字组成。20 字符位定义字段长度的长度, 即定义目次区中字段长度部分所含的字符数, 其值总为 4。21 字符位定义起始字符位置的长度, 即定义目次区中起始字符位置部分所含的字符数, 其值总为 5。22 字符位定义为执行定义部分的长度, 由于目次区不包括执行定义部分, 故其值总为 0, 23 字符位未定义, 其值总为 0。

相关字段:

040 编目源

580 连接款目综合附注

760-787 连接款目字段

相关文献:

USMARC Specifications for Record Structure, Character Sets, and Exchange Media.

内容标识符史:

1. 06 记录类型

06 字符位的代码 b (档案与手稿控制) 于 1995 年废止, 用头标 08 字符位(控制类型)来反映。

h 缩微形式出版物, 1968-1972 年使用, 后改用 007

n 特殊资料说明

1975-1983 年使用, 之后分别用代码 k (二维非放映图形 (Two-dimensional nonprojectable graphic)) 和代码 r (三维制品或自然生成物) (Three-dimensional artifact or naturally occurring object)。

2. 18 著录标准

p 部分采用 ISBD 标准

1974-1987 年使用, 之后改用 i 表示

r 根据 AACR1 著录

1976-1981 年使用, 现在用 i 表示

使用说明:

1. 头标里的代码表示的是记录特征, 而不是直接表示编目实体本身的特征

2. 多数数据元素由软件自动生成, 仅记录状态(05)、记录类型(06)、书目级别(08)、编目等级(17)和著录标准(18)等数据元素要根据编目实体和记录编制要求选择相应代码。

3. 通过头标 05、17、18 字符位的代码，大体可以判定一条书目记录的质量，可为编目员套录数据、批处理替代数据提供参考信息。

001-006 控制字段

00X 字段包含记录控制号和其它控制信息, 用于处理 USMARC 书目记录。控制字段可以由单个数据元素组成, 也可以由一系列通过字符位识别的定长数据元素组成, 目前定义了 5 个字段:

001 控制号	必备, 不重复
003 控制号标识	选用, 不重复
005 最近一次作业的时间和日期	必备, 不重复
006 定长数据元素—附件特征	选用, 可重复
007 载体形态定长字段	选用, 可重复
008 定长数据元素	必备, 不重复

说明:

1. 00X 字段均为可变控制字段, 包含数据元素和字段终止符, 不包含指示符和子字段代码。
2. 00X 可变控制字段的数据元素可以是单个数据元素, 也可以是定长的字符串, 每个数据元素通过相关的字符位识别。

001 控制号 (Control number)

控制号是 MARC 记录的识别号, 具有唯一性, 由编制、使用或发行记录的机构设置。

001 控制号是必备字段, 不能重复使用。

实例:

例 1 001 bbb86104385b

注: 美国国会图书馆控制号: ~~bbb~~(0-2 位)—控制号前缀; 86(03-04 位)—记录入档年; 104385(05-10 位)—记录流水号; b(11 位)—补编号

例 2 001 pu197000736

注: 北京大学图书馆在光盘环境下编目时设置的控制号: pu1—机构代码; 97—记录入档年; 000736—记录流水号。

例 3 001 0297000736

注：北京大学图书馆中央书目数据库记录控制号：02—编目语种(西文)；97—记录入档年；000736—记录控制号。

应用：

用于识别特定的 MARC 记录，主要用于业务部门对特定记录的检索和批处理。

相关字段：

010 美国国会图书馆控制号

035 系统控制号

使用说明：

1. 控制号是一个联机系统的唯一识别号，通常由系统自动生成。
2. 由于控制号具有唯一性，为方便交换，提供交换数据方一定要提供控制号结构和有关输入方面的文件，而接受记录方应重新分配本系统的控制号，而把原控制号移至 035 字段或 010 字段。例如：当套录的数据源取自 BiblioFile 光盘数据库，可以把源记录 001 字段的控制号移至 010 字段(美国国会图书馆控制号)；当套录的数据源取自其它数据库，通常把源记录 001 字段的控制号移至 035 字段。
3. 根据控制号具有唯一性这一特点，若利用数据库中相近的记录建立另一条新记录(如利用某文献第一版的记录建立该文献第二版的新记录，必须产生新的控制号。

003 控制号标识 (Control number identifier)

记录生成 001 字段控制号的机构代码，该代码依据 *USMARC Code List for Organizations* (由美国国会图书馆维护)来著录。

实例：

003 DLC

003 OCoLC

005 最近的处理日期和时间 (Date and Time of latest transaction)

005 字段包含最近一次对本记录进行业务处理的日期和时间。本字段的数据元素长度为 16 位。日期按照 *Representation for Calendar Date and Ordinal Date Information Interchange* (ANSI X3.30) 自动生成，由 8 位字符组成，其模式为 *yyyymmdd* (其中 y 为年，m 为月，d 为日)，时间按照 *Representations of Local Time of the Day for Information Interchange* (ANSI X3.43)，也由 8 位字符组成，模式为 *hhmmss.f* (其中 h 为小时，m 为分，s 为秒，f 为秒的第一位小数)

本字段是必备字段，不能重复使用。

实例:

例 1 005 19900124103555.6

注: 本记录最近一次进行业务处理的日期是 1990 年 1 月 24 日(19900124), 时间是 10 时 35 分 55.6 秒。

应用:

记载书目记录最近进行业务处理的日期和时间。以此作为记录修改更新的标识。

相关字段(代码)

008 字段/00-05 字符位 (入档日期)

相关文件/文献

ISO 2014 信息交换用国际标准时间记载形式。

使用说明:

1. USMARC 在 1982 年设置 005 字段, 在此以前, 记录中不一定包含本记录被更新的日期和时间, 或被记载在其它字段。
2. 本字段的日期、时间信息由系统产生, 字段中包含的时间一般为地方时。
3. 本字段的时间和日期是动态的, 记录每更新一次, 本字段的时间也随之改变, 而 008 字段中的入档日期是永久不变的。

006 定长数据元素—附件特征(Fixed-length data elements—additional material characteristics)

本字段以代码形式, 描述具有多种资料类型特征的编目文献。例如: 具有连续更新页的活页出版物, 具有专著和连续出版物特征, 带盒式讲演录音带附件的教科书, 具有专著和录音特征。本字段主要记载附件的特征信息, 主文献的资料类型特征记录在 008 字段(定长数据元素)。006 字段是选用字段, 可重复使用。

数据元素:

006 字段包含 18 个字符位(00-17), 用来提供 008 字段不能反映的编目实体的特殊编码信息。字段采取自定义结构, 即由 00 字符位资料类型的代码值确定其它字符位(01-17)的代码。

006 字段的 00 字符符合包含 7 种书目资料类型, (13 种记录类型), 它们的数据元素内容分别列表说明如下:

1. 图书

符位置	数据元素内容	必备性	备注
00	006 资料类型代码	必备	
	a 图书	有则必备	
	t 手稿型文字资料	有则必备	
01-04	插图	选用	含 16 种代码
05	读者对象	必备	含 2 种代码
06	资料形式	必备	含 7 种代码
07-10	内容特征	选用	含 21 种代码
11	政府出版物	必备	含 11 种代码
12	会议出版物	选用	含 2 种代码
13	纪念文集	选用	含 2 种代码
14	索引	选用	含 2 种代码
15	未定义	选用	用空位表示
16	小说	必备	必备
17	传记	必备	含 5 种代码

各位字符的代码详见 008 字段图书记录类型的 008/18-34 字符位。

2. 连续出版物

字符位置	数据元素内容	必备性	备注
00	006 资料类型代码	必备	
	s 连续出版物	有则必备	
01	出版频率	必备	含 17 种代码
02	发行规律	必备	含 4 种代码
03	ISDS 中心	选用	含 13 种代码
04	连续出版物类型	必备	含 4 种代码
05	原版资料形式	选用	含 7 种代码
06	资料形式	必备	含 7 种代码
07	著作整体特征	选用	含 22 种代码
08-10	内容特征	选用	含 22 种代码
11	政府出版物	必备	含 2 种代码
12	会议出版物	选用	含 2 种代码
13-15	未定义	选用	用空位表示
16	题名原文字母/文字	选用	含 15 种代码