

# USMARC 对中文图书的编目处理

高 红

(国家图书馆 北京 100081)

[摘要] USMARC 作为世界上使用范围较广泛的机读目录格式,主要用于描述、存储、交换、控制和检索机读书目数据。本文的目的就是探讨 USMARC 针对多文种文献资料采用的著录方法及其意义和作用。

[关键词] MARC 格式 多文种文献处理 中文图书编目

[分类号] G254.364

## Cataloguing Chinese Books By USMARC

[ABSTRACT] USMARC is one of widely used machine-readable catalogs in the world. Recent years, there are more different language material of publications, to adapt this change, USMARC format has been improved.

[KEY WORDS] Bibliographic Format. Multilanguage materials Process. USMARC for Chinese Books.

[CLASS NUMBER] G254.364

近年来,由于各语种的文献资料在国际上有了越来越多的交流,尤其象中国、日本和韩国这些使用图形字符的国家,出版的文献越益受到西方国家的重视。随着计算机网络的应用和发展,不同国家、使用不同语言的人员均可通过计算机查询所需文献。特别是计算机的发展,使它能够很好处理多种字符集,并且能够在—个编目软件上利用—个 MARC 格式处理多文种文献。长期以来,西方图书馆界的计算机编目对于图形文字资料,在著录时均采用以拉丁字母拼音来表示该种文字。比如,对于中文图书,著录时将题名、著者中的汉字改用韦氏拼音或汉语拼音形式来表示,这种方法称为图形字符的罗马化表示方法。对于如何处理不同文种的机读目录格式,问题主要集中在如何使机读目录格式更加适应于不同文种资料的编目。本文的主要目的就是探讨 USMARC 近年来针对多文种文献(仅以中文图书为例)的处理而采用的著录方法。

### 1. 设置 066 字段(字符集说明)

066 字段是字符集说明,用来说明书目记录中出现的非缺省 MARC 拉丁字符集(ASCII 和 ANSEL)的信息,以便计算机据此作相应处理。其表示方法是用特殊的标识符号来表示相应的字符集。066 字段的两个指示符均未定义,填写时用空位(“##”)表示。

\$a 包含标识基本字符集的代码,此代码由标识和采用缺省的 G0 字符集(中国国家标准的基本集)的转义序列的中间和最后的字符组成。

066 ##\$a(N

(表示记录采用 G0 字符集。中间字符 hex 28(即 ASCII 的“(”),表示一个字符位由一个字节构成。

最后字符 hex 4E(即 ASCII 的“N”),表示本记录采用基本基里尔字符集)

\$b 包含基本扩展字符集代码,此代码由标识和采用缺省 G1 字符集(中国国家标准的基本集)的转义序列中的中间和最后的字符组成。

066 ##\$b\$)1

(表示记录采用基本 G1 字符集。中间字符 hex 24、hex 29(即 ASCII 的“\$”),表示一个字符位由多个字节构成;最后字符是 hex 31(即 ASCII 表示的“1”),标识本记录采用的是 MAR CJK(中文、日文

和朝文)字符集)

\$c 是交替 G0 和 G1 字符集, 转义序列中间和最后的字符, 用来说明记录中使用交替图形字符集, 为可重复必备子字段, 它与 880 字段配合使用, 用来说明 880 字段中使用交替图形字符集是什么字符集, 因此, 当记录中有 880 字段出现时, 该记录中就一定要包含 066 字段的 \$c 子字段。

066 ##\$c\$1

表示记录采用交替 G0 和 G1 字符集。标识序列的中间字符 2416(ASCII 的表示是“\$”), 一个字符位由多字节构成, 最后字符是 3116(ASCII 的表示是“1”), 标识本记录采用的是 MARC CJK(中文、日文和朝文)字符集。

## 2. 设置 880 字段(交替图形表示内容)

MARC21 书目数据格式采用字段连接的方法处理非罗马化文种的资料, 将罗马化字符的著录内容与非罗马化字符著录的内容连接起来。

880 字段用来输入非 ASCII 码的图形字符的著录内容, 其著录内容与本记录中相应的用罗马字符进行描述的字段内容成对出现, 可能与 880 字段成对出现的字段有:

020	国际标准书号
040	编目来源
041	语种代码
043	地理区域代码
1XX	主要款目
2XX	题名及相关题名
3XX	载体形态描述
4XX	专著丛编
5XX	附注
6XX	主题
7XX	附加款目
76X-79X	连接款目
800-830	丛编附加款目

从以上各个字段可以看出, 几乎所有的 USMARC 字段均可与 880 字段配对, 也就是说, 几乎所有字段的内容都可使用非罗马化字符来表示。

## 3. 880 字段的应用

880 字段(交替图形表示内容)用来输入非 ASCII 码的图形字符著录内容, 所有 880 字段以及与它对应字段的第一个子字段必定是 \$6, 因此在一般情况下, \$6 子字段是成对出现的, 这样连接后, 即可把罗马化字符所著录的内容与图形字符集著录的内容完全对应起来。\$6 的内容包括一个与之对应的标识号(Tag number), 还包括一个随机分配的序号(Occurrence number), 此序号由编目员任意分配给每个字段, 供计算机进行识别, 每一记录中每对字段的序号相同, 但每对不同字段的序号不可重复使用, 这里的序号在记录中仅表示字段的匹配关系, 并不表示它在记录中的任何顺序。当 MARC21 规范格式 880 字段没有相应字段与之相连时, \$6 控制子字段中序号为“00”。

\$6 控制子字段的结构为:

\$6<连接字段标识号>-<序号>/<交替字符集说明>/<被连接的字段号>

\$6 控制子字段中的交替图形字符集的表示符号, 即文种的识别代码有以下 6 种供选择使用:

- (3 阿拉伯文
- (B 拉丁文
- \$1 中文、日文、韩文
- (N 基里尔文
- (2 希伯来文
- (S 希腊文

例:

100 1#\$6880-01\$aShen, Fuwen.  
 245 10#\$6880-02\$aZhongguo qi yi mei shu shi /\$cShen Fuwen bian zhu.  
 880 1#\$6100-01/\$1\$a 沈福文.  
 880 10#\$6245-02/\$1\$a 中国漆艺美术史 /\$c 沈福文编著.

(这个例子中“\$1”是交替图形字符集的代表符号，表示在这个字段中遇到的第一个非罗马化的交替字符是中文)

#### 4. 各字段处理细则

##### 1) 1XX 字段（主要款目标目）

中国人名标目通常采用汉语拼音形式，外国人名标目采用其原名的罗马化形式；文献上的团体、会议名称若有英文等形式，则按英语等形式著录；如果名称仅仅是中文，著录时可用中文的汉语拼音形式。无论哪种形式，都要注意规范控制，如：

100 1#\$6880-01\$aLi, Shan'gang.  
 880 1#\$6100-01/\$1\$a 李山冈.

100 1#\$6880-01\$aMuller, Marcia.  
 880 1#\$6880-01/\$1\$a 米勒, 玛西亚.

100 1#\$6880-01\$aSuzuki, Hiramitsu, \$d1949-  
 880 1#\$6100-01/\$1\$a 铃木平光.

110 1#\$6880-01\$aChina.\$bGong ye jiao tong tong ji si.  
 880 1#\$6110-01/\$1\$a 国家统计局.\$b 工业交通统计司.

如果外国人名没有出现在编目文献上，亦可采用汉语拼音形式著录。

100 1#\$6880-01\$aLingmu, Pingguang,\$d1949-  
 880 1#\$6100-01/\$1\$a 铃木平光.

##### 2) 245、246 字段（题名说明）

- 245 题名字段采用汉语拼音形式，应严格按题名页著录。当责任说明项不是取自题名页时，应置于“[]”内；隐含的著者不必做责任说明（如，吴湖帆画集）。
- 当并列题名出现在题名页上，应与文献内容语言相同的题名转录为正题名(245\$a)，另一种语言的题名转录为并列题名(245\$b)，并为并列题名做附加款目(246)。与此对应，“880-245”字段须著录\$b 并列题名，如：

245 10#\$6880-02\$aDai er gong si chuan qi = \$bBusiness the Dell way /  
 \$cRebecca Saunders.  
 246 31\$a Business the Dell way

880 10\$6245-02/\$1\$a 戴尔公司传奇 = \$bBusiness the Dell way /  
\$c 丽贝卡·桑德斯著.

- 如果其他语言的并列题名未出现在题名页,而是出现在其他书页上(如封面、书脊、版权页等),245 字段不必著录其他语言题名,将其著录在 246 字段;与此对应,“880-245”字段不著录其他语言的并列题名,但需在“880-500”字段予以说明,如:

245 10\$6880-02\$aMei guo yu ya zhou : \$bmei guo xin zhan lue he bing li tai  
shi /\$cZhalemi, Halilezhade deng zhu.

246 33\$aUnited States and Asia : \$btoward a new U.S. strategi y and force posture

880 10\$6245-02/\$1\$a 美国与亚洲 : \$b 美国新战略和兵力态势 /  
\$c 扎勒米,哈利勒扎德等著.

880 ##\$6500-00/\$1\$a 原文题名: The United States and Asia: toward a new U.S.  
strategi y and force posture.

### 3) 250 字段(版本说明)

版本说明项通常采用汉语拼音和阿拉伯数字表示,如,第二版著录为 Di 2 ban 和第 2 版,如:

250 ##\$6880-02\$aDi 1 ban.

880 ##\$6250-02/\$1\$a 第 2 版.

### 4) 260 字段(出版发行说明)

出版地、出版家一般采用汉语拼音,如:

260 ##\$6880-02\$aBeijing : \$bZhongguo tong ji chu ban she, \$c2001.

880 ##\$6260-02/\$1\$a 北京 : \$b 中国统计出版社, \$c2001.

### 5) 300 字段(载体形态说明)

300 ##\$6880-04\$a14, 148, [3] p., [36]leaves of plates : \$bill.(some col.) ; \$c28 cm.

800 ##\$6300-04/\$1\$a14, 148, [3]页, [36]叶图版 : \$b 插图(部分彩色); \$c28 公分.

### 6) 440 字段(丛编说明)

丛编字段采用汉语拼音。当丛编题名既有英语题名,又有中文题名时,440 字段须著录中英文题名(汉语拼音)，“880-440”字段对应著录中英文题名,如:

440 #0\$6880-05\$aQin zi le yuan ; \$v21

880 #0\$6440-05/\$1\$a 亲子乐园 ; \$v21

440 #0\$6880-06\$aShi jie yi shu quan ji = A comprehensive collection for appreciation  
of world art

880 #0\$6440-06/\$1\$a 世界艺术全鉴 = A comprehensive collection for appreciation  
of world art.

### 7) 5XX 字段(附注说明)

如果在 880 字段中用其他字符集中的字符所表示的著录项目在本记录中无相应的罗马化字段,则应分配其序号为“00”,因此,序号为 00 时,则专用于 880 无匹配字段的情况,如:

245 10\$6880-02\$aZhongguo qi yi mei shu shi /\$cShen Fuwen bian zhu.

246 33\$aHistory of the traditional Chinese lacquer art

880 10\$6245-02/\$1\$a 中国漆艺美术史 / \$c 沈福文编著.

880 ##\$6500-00/\$1\$a 封面题名: The history of the traditional Chinese lacquer art.

此外, 处理中文图书时, 常著录的附注说明有:

Partial text in English	部分英文篇幅
Parallel text in Chinese and English	中英对照
Text written in traditional Chinese scripts	繁体字本
Parallel text in Chinese and Hanyu Pinyin	拼音注本
Cover title	封面题名
Original title	原文书名
Title on spine and colophon	书脊及版权页题名
Author's original name	著者原名
Contents	内容
Includes index	附索引

## 5. 实例

例 1:

(书名页示意图)

(版权页示意图)

<p>开发孩子潜能的 300 种方法</p> <p>[美]埃丽斯·格里菲斯 著 王文忠 徐金灿 译</p> <p>新华出版社</p>	<p>Every Child Is a Genius 365 Fun and Easy Ways to Develop Your Child's Gifts Copyright© 1998 by Elise M. Griffith 中文简体字版专有权属于新华出版社 开发孩子潜能的 300 种方法 [美]埃丽斯·格里菲斯 著 王文忠 徐金灿 译 新华出版社出版发行 (北京宣武门西大街 47 号) 2001 年 9 月第一版 ISBN 7-5011-5274</p>
--	---

机读目录格式:

066 ##\$c\$1

100 1#\$6880-01\$aGriffith, Elise M.

245 10\$6880-02\$aKai fa hai zi qian neng de 300 zhong fang fa /  
\$cAilisi Gelifei si zhu ; Wang Wenzhong, Xu Jincan yi.

246 33\$aEvery child is a genius : \$b365 fun and easy ways to develop your  
child's gifts.

260 ##\$6880-03\$aBeijing : \$bXin hua chu ban she,\$c2001.

300 ##\$6880-04\$a307 p.

440 #0\$6880-05\$aQing ping guo cong shu

546 ##\$aIn Chinese.

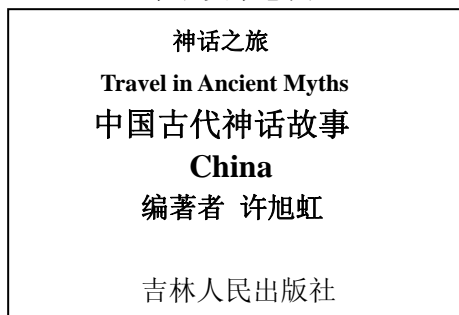
700 1#\$6880-06\$aWang, Wenzhong.

880 1# \$6100-01/\$1\$a 格里菲斯, 埃丽斯.

880 10\$6245-02/\$1\$a 开发孩子潜能的 300 种方法 /\$c 埃丽斯·格利菲斯著 ;  
 王文忠, 徐金灿译.  
 880 ##\$6260-03/\$1\$a 北京 :\$b 新华出版社,\$c2001.  
 880 ##\$6300-04/\$1\$a307 页.  
 880 #0\$6440-05/\$1\$a 青苹果丛书  
 880 ##\$6500-00/\$1\$a 原书名: Every child is a genius : 365 funand easy ways  
 to develop your child's gifts.  
 880 1#\$6700-06/\$1\$a 王文忠.

例 2:

(书名页示意图)



机读目录格式:

066 ##\$c\$1  
 100 1#\$6880-01\$aXu, Xuhong.  
 245 10\$6880-02\$aZhongguo gu dai shen hua gu shi = \$bChina /\$cbian zhu zhe,  
 Xu Xuhong.  
 246 31\$aChina  
 260 ##\$6880-03\$aChangchun : \$bJilin ren min chu ban she,\$c2001.  
 300 ##\$6880-04\$a186 p. : \$bill. ; \$c25 cm.  
 440 #0\$6880-05\$aTravel in ancient myths  
 546 ##\$aIn Chinese.  
 880 1#\$6100-01/\$1\$a 许旭虹.  
 880 10\$6245-02/\$1\$a 中国古代神话故事 = \$bChina /\$c 编著者, 许旭虹.  
 880 ##\$6260-03/\$1\$a 长春 :\$b 吉林人民出版社, \$c2001.  
 880 ##\$6300-04/\$1\$a186 页 :\$b 插图 :\$c25 公分.  
 880 #0\$6440-05/\$1\$a 神话之旅

通过以上著录, 我们可清楚看到, 记录中用罗马化字符描述的全部内容搬入 880 字段, 形成一个完整的非罗马化字符所描述的书目记录, 再通过计算机进行一些自动处理, 即可将同一文献为读者生成两种文字的书目记录, 并可保持在数据库中一本图书只产生一条书目记录, 其著录格式是一个。这样不仅大大方便了读者且满足了不同读者的需求, 同时也对多语种文献的处理方法, 提供了一个可借鉴的行之有效的办法。

参考文献:

1. MARC 21 Format for Bibliographic Data. 1999.
2. 吴开华主编. 高校图书馆编目工作研究. 北京图书馆出版社, 2000 年.
3. 《西文文献著录条例》(扩大修订版) 中国图书馆学会, 2003.