

●苏新宁

音像资料的 MARC 格式研究

摘 要 根据音像资料的特点及其与 CNMARC 的分析比较,设计了音像资料 MARC 格式。对于 CNMARC 不能完全表述的音像资料字段进行修改和补充;对音像资料来说,CNMARC 的冗余字段做必要删除;对 CNMARC 不能表述的音像资料进行字段增加。其格式已在南京电视台的管理信息系统中运用。参考文献 3。

关键词 MARC 音像资料

分类号 G254.36

ABSTRACT Based on the characteristics of A/V materials and an comparative analysis with CNMARC, the author designs a MARC format for A/V materials, in which some fields are added and modified. The format has been used in the management information system of the Nanjing Television Station. 3 refs.

KEY WORDS MARC. A/V materials.

CLASS NUMBER G254.36

1 引 言

MARC 格式是一种表达文献信息的标准机读目录格式。该格式已被图书馆等信息管理与服务部门广泛用于文献信息的交换和数据共享。

为了解决国与国之间机读目录易造成的不兼容问题,国际图联于 1977 年发表了 UNIMARC 格式,使各类 MARC 格式可依赖 UNIMARC 这一中间格式进行数据交换,UNIMARC 逐渐成为世界范围内进行文献信息交换的标准格式。

我国自 80 年代末 90 年代初开始了将 USMARC 向 CNMARC 转换的工作,并于 1995 年正式出版了《中国机读目录格式使用手册》。CNMARC 的出版大大推进了中国书目数据规范与统一,加速了文献信息网络化的建设,实现了国内各单位之间以及国际间书目信息的交换和共享。

1991 年美国国会图书馆开始了 USMARC 书目格式一体化研究^[1]。这项研究的

主要目的是将 7 种不同文献类型(图书、计算机文件、地图、乐谱、连续出版物、可视资料、档案—手稿)的 MARC 格式统一为一种通用格式。1993 年底,USMARC 的一体化格式开始付诸实践^[2]。MARC 格式的一体化着重于各类信息资源的统一著录,提高编目信息的共享性。

但是,我们必须看到,由于各类文献都有着它的独特性,通常难以用一个格式将不同类型文献的特别之处都充分表达出来。尤其是音像资料,它与纸质文献的差别更大,在制作与管理过程中,音像资料的动态性数据也是现有 MARC 所不能表达的。南京电视台在构架自己的音像资料管理系统时,感到现有的 MARC 很难对音像资料的某些信息进行表述。为了实现该台与兄弟电视台信息资源的共享与交流,1998 年 2 月南京电视台委托我们,专门对音像资料的机读目录格式进行研究。

2 研究线路

在着手进行音像资料 MARC 格式设计之前,我们首先对音像资料的特点以及 CN-

MARC 进行了分析比较。我们认为,与书目资料相比,音像资料在著录款目上与其有共性的地方,但在存储介质、描述细节、著录详简程度等方面具有自己的特性。由于电视台等影视单位对音像资料的应用更强调实用性,对它的管理和使用要求紧密配合整个的采编播等业务。因此,它有别于一般图书馆面向的“静态文献资料”的管理。为此我们首先以南京电视台为基地,进行了大量充分详尽的前期调研工作,不断听取反馈意见,使其著录项目的制定更具有实用性。

同时,我们还走访了国家图书馆、中央电视台、中央教育电视台、北京电视台、中央音乐学院等单位,拜访许多专家学者。我们的工作得到专家们的肯定,同时,我们也从同专家们的交流中得到了许多启迪。

在前期充分调研的基础上,我们对 CN-MARC 进行逐项分析,紧密结合影视单位业务特点和音像资料的特别之处,力求设计出汇集音像、书刊资料的 MARC 格式。我们采取的设计准则为:保证格式具有较强的实用性和通用性和规范化,使其既满足音像资料这种特殊数据的表达需要,也兼顾书刊资料的数据表达要求。因此,我们决定采取如下研究原则:

(1) 参照 UNIMARC 和 GB3792.1 - 83《文献著录总则》、GB3792.4 - 85《非书资料著录规则》、《中国机读目录格式使用手册》、《音像资料著录规则》等国际标准和国家标准,并结合音像资料的特点、最新发展动态和未来趋势。

(2) 以 CNMARC 格式为主要依据,在著录项的设置、著录项的排列以及标识符的运用上都力求与国家标准保持一致,在强调共性的同时,又注重广播影视系统的特点和采编播导人员检索使用的要求,保证有较强的实用性和针对性。

(3) 以 CNMARC 为基础,但其侧重点放在音像资料方面,注重对其细化描述,以满足实际应用的需要。

根据以上原则,我们勾画了如下研究思路:

对于 CNMARC 不能完全表述的音像资料字段进行修改和补充;对音像资料来说,CNMARC 的冗余字段进行必要的删除;对 CNMARC 不能表述的音像资料信息进行字段增加。

3 研究成果

CNMARC 格式主要著录对象是图书资料,而音像资料所具有的个性化特征和“动态性”数据,它很难充分反映。这就要求我们在制定音像资料 MARC 时,既要谋求两者之间的兼容,又要针对音像资料的特征加以突出反映。

音像资料具有和图书资料相通的方面,但两者的侧重点也各有所不同。音像资料,尤其是电视台等业务性较强的单位更强调动态性和实时性,如实际播放时间是广播电影电视音像资料的重要数据,每天安排的节目时间表中都必须用到,在采用计算机控制的自动切换系统后,每个节目的播放时间就更显得重要,每个节目的起始时间都是非常重要的检索点。此外本格式对于音像资料,尤其是电视台的大量素材信息、版权信息、获奖信息、节目信息、附件信息、播出信息、审查信息、细节项等,都在 9XX 内部使用块中作了相应地规范化处理,从一台综合性文艺节目到一首歌、一个小品甚至是一个水滴音效都可能进行著录。这类信息既可以作为台内信息使用,也可以作为交换信息使用。现就我们在研究中,对 CNMARC 字段的主要修改与补充内容阐述如下。

3.1 对 001 字段的修改

001 字段为记录惟一性标识字段。CNMARC 规定该字段的长度为 10 个字符,共包含 3 部分内容:资料类型(2)、编目年代(2)、编目流水号(6)。原格式采用用两个字符来表示编目年代,不言而喻会遇到 2000 年问题,因此本格式将其由 2 位增加到 4 位。下为该字段修改后的表达。

(1) 资料库类型代码(代码类型)2 0 - 1

(2) 编目年代 4 2 - 5

(3) 编目流水号 6 4 - 9

例如:001 21199700001

表示中文的音像资料,于 1997 年编目,流水号为 00001。

3.2 对 115 字段的补充

115 字段为定长的编码数据字段,主要用于表达非书资料(投影片、录像制品和电影片)。该字段的 \$a 子字段中的第 12 项“视频规格-录像制品(字符位置 16)”的“Z”标识(原为其他),为满足现代录制技术发展的需要,用它记录各大公司不断推出新的录像带或 CD 激光唱片产品。当该位置被标记为“Z”时,可将新的录像制品数据著录于 307 字段中。目前,这种处理方法已被影视部门广泛采用。CNMARC 目前尚不能表达的录像制品的录制规格如下:

Betacam(盒式录像带):日本 SONY 公司研制的专业视频规格;

Betacam sp(盒式录像带):日本 SONY 公司研制的专业视频规格;

Digital Betacam(盒式录像带):日本 SONY 公司研制的广播级数字视频规格;

Betacom sx(盒式录像带):日本 SONY 公司研制的数字压缩视频规格;

M 型(盒式录像带):日本松下公司研制的专业视频规格;

M2(盒式录像带):日本松下公司研制的专业视频规格;

DVCPRO(盒式录像带):日本松下公司研制的数字压缩视频规格;

DVCPRO50(盒式录像带):日本松下公司研制的数字压缩视频规格;

Digital s(盒式录像带):日本胜利(JVC)公司研制的数字压缩视频规格;

DV(盒式录像带):日本胜利公司研制的家用数字压缩视频规格;

DVCOM(盒式录像带):日本 SONY 公司研制的数字压缩视频规格;

D5(盒式录像带):日本松下公司研制的广播级数字分量视频规格;

D1(盒式录像带):日本 SONY 公司研制的广播级数字分量视频规格;

D2(盒式录像带):日本 SONY 公司研制的广播级数字复合视频规格;

D3(盒式录像带):日本松下公司研制的广播级数字复合视频规格。

3.3 对 205 字段的修改

在 CNMARC 中,205 字段用于表述版本信息。根据影视部门制作节目的工作流程,我们将 CNMARC 中版本说明(\$a)子字段改为音像资料的节目版本,将附加版本说明(\$b)用制作版本信息来替代,其他子字段内容不变。具体说明如下:

\$a 节目版本说明,亦称版本形式。指音像资料的内容未变,其载体有变的版本,翻录版即为此种情况。例如,从影片到录像带的翻制版:FK 到 VH。若著录来源中有关版来源说明,著录于附注字段,此外根据自身需要还可著录特殊版信息。该子字段包含下列版次说明:购买、自制、交换、翻录、微波复制、其他。

\$b 制作版本说明。该子字段包含下列版次:素材版、剪辑版、工作版、原版(分为自制原版和翻录原版)、播出版、副版 1、副版 2、副版 3、副版 4。

例如:205 \$a 中文版 \$b 翻录原版

3.4 对 210 字段增加子字段

CNMARC 的 210 字段用于表达音像资料的出版发行项,针对影视作品的制作发行方式的多样化特征,除制作地、出版地等还添加了拍摄录制地(\$i)和录制形式(\$j)等信息项,保证满足 MTV、外景节目标注的需要。

例如:210 \$i 黄土高原(延安) \$j 外景

3.5 补充 015 字段

在 CNMARC 中,01x 字段主要用作标识出版物国际标准号的字段(如,010 为 ISBN;011 为 ISSN)。同样,ISRC 为各种音像记录以及记录中每一条独立使用的节目提供了一个国际性的惟一标准编码。随着计算机网络、光电通讯技术的发展和广播电视事业的高科

技化,音像制品贸易流通的国际化,ISRC 将越来越显示出它的重要性。因此我们添加了 015 字段用来标识音像资料的国际标准号。该字段可由以下子字段组成。

- \$a ISRC(由国家指定的机构分配)
- \$b 限定
- \$d 获得方式和/或定价
- \$z 错误的 ISRC

例如:015 \$ACN - E01 - 95 - 0070 - 0

\$d 28.00

3.6 增加 195 编码数据字段

CNMARC 没有提供专门的字段描述光盘制品,而对目前大量涌现的光盘音像制品的记录则是一个现实问题。我们在遵循 MARC 制定原则的基础上,增加了 195 定长编码数据字段,专门用于表述各类光盘制品。

表 1 为该字段的 \$a 子字段的表达格式,具体编码内容参见文献[3]。

表 1

数据元素序号	数据元素名称	字符数	字符位置
1	激光制品类型	1	0
2	播放长度	3	1 - 3
3	颜色标志符	1	4
4	宽度、尺寸	1	5
5	附带资料	4	6 - 9
6	影像规格 - 激光制品	1	10
7	制作形式	1	11
8	卷及文件结构	1	12
9	盘片结构	1	13
10	盘片使用类型	1	14
11	数据压缩标准	1	15

3.7 增加业务管理工作等字段

音像资料的管理与服务部门以及资料的利用者出于工作上的特殊需要,希望能从资料库的 MARC 记录中获得节目信息、广播电视播出信息、审查、获奖和版权和附件信息等,而这些信息在 CNMARC 中很难有准确的定位。因此,在我们所研究的音像资料 MARC 格式中,为这些信息设立了 96X 字段。

3.7.1 960 节目信息

该字段针对广播、电视对节目信息的特殊处理要求而制定的,反映了记录与节目之间的指示关系。该字段适用于广播电视节目,作为选择性字段,字段可重复。该字段的子字段构造如表 2。

表 2

子字段标识	子字段内容	注释
\$a	节目类别说明	可重复
\$b	节目子类别说明	可重复
\$c	节目拍摄时间	不可重复
\$d	节目长度时间说明	不可重复
\$e	栏目名称(一级)	不可重复
\$f	栏目名称(二级)	不可重复
\$g	栏目名称(三级)	不可重复
\$h	节目总集(盘)说明	不可重复
\$i	节目分集(盘)说明	不可重复

3.7.2 961 广播电视播出信息

该字段用于记录广播电视节目的播出及相关信息。该字段为可选字段,字段可重复,部分子字段可重复。该字段的子字段内容如表 3。

表 3

子字段标识	子字段内容	注释
\$a	首播时间	不可重复
\$b	首播频道	不可重复
\$c	首播栏目	不可重复
\$d	再播时间	可重复
\$e	再播频道	可重复
\$f	再播栏目	可重复
\$g	时间长度	不可重复

3.7.3 962 审查信息

该字段用于著录电视、电影、文艺节目的审查信息。该字段为可选字段,可重复。它的子字段内容如表 4。

表 4

子字段标识	子字段内容	注释
\$a	审查责任者	可重复
\$b	审查日期	可重复
\$c	审查意见	可重复

3.7.4 963 获奖信息

该字段用于著录电视节目、影片、电视剧等获奖信息。这些信息可作为鉴定节目、影片、电视剧等质量的重要参考。可选择著录字段,字段可重复。表 5 为该字段的子字段内容。

表 5

子字段标识	子字段内容	注释
\$a	获奖名称	可重复
\$b	获奖时间	可重复
\$c	获奖级别	可重复

3.7.5 965 版权项

该字段著录从其他单位或国外购买、引进的节目、影片等版权信息。对引进、购买的节目数据著录,该字段为必选字段。若对非引进购买的节目的数据记录,该字段为选择字段,但该字段不可重复。表 6 为子字段描述。

表 6

子字段标识	子字段内容	注释
\$a	版权责任者	不可重复
\$b	版权责任说明	不可重复
\$c	版权购买金额	不可重复
\$z	供片单位	不可重复

3.7.6 967 附件信息

该字段主要用于记录音像资料所附带的其他资料信息。本字段可选择使用,字段也可以重复。表 7 为子字段描述。

表 7

子字段标识	子字段内容	注释
\$a	附件名称	可重复
\$b	附件类型	可重复
\$c	附件索取号	不可重复
\$d	附注信息	可重复
\$e	附件数量	不可重复

3.7.7 968 细节项

该字段只用于著录录音等音像资料。可选择字段,也可重复使用。子字段描述如表 8。

表 8

子字段标识	子字段内容	注释
\$a	细节款目	不可重复
\$b	素材说明	可重复
\$c	责任者	不可重复
\$d	起止时间与位置	不可重复
\$e	时间长度	不可重复
\$f	镜头描述	可重复
\$g	附注信息	可重复

4 结束语

经过近一年的努力,我们完成了音像资料机读目录格式的设计。该格式已在南京电视台的管理信息系统中得到运用,这在全国范围内的影视部门中尚属首次,具有一定的开创意义。本格式的制定和系统的开发,对于建立和健全广播电视系统音像资料管理的科学化,实现音像资料著录工作的标准化以及推动各部门各单位之间音像资料的交流与共享等方面都具有重大的现实意义,迎合了网络信息传递、交流和共享的大趋势。

在本课题的研究中,我的研究生崔雅萍、许晋军做了大量工作,在此对他们表示感谢。

参考文献

- 1 熊光莹. USMARC 书目格式一体化. 现代图书情报技术, 1992(3): 40 ~ 50
- 2 林容等. 编目格式的重大突破 - USMARC 的进展介绍. 现代图书情报技术, 1997(3): 41 ~ 46
- 3 苏新宁等. 音像资料机读目录 (MARC) 格式, 课题研制报告. 1998. 12

苏新宁 南京大学信息管理系教授。通信地址: 南京市。邮编 210093。

(来稿时间: 1999 - 08 - 10。编发者: 刘喜申)