

● 刘苏雅

USMARC 社区信息格式

——图书馆自动化信息服务的新发展

ABSTRACT Community information is a kind of non-document information for members in a community. *USMARC Format for Community Information* was published in US in February 1993. In this paper, the compilation process, characteristics, applications and problems of this format are described. 6 refs.

KEY WORDS Community information. MARC format.

CLASS NUMBER G254.36

社区信息(community information),作为一个新出现的图书馆学专业术语,指满足某一社区内各类成员需求的非文献信息(non-bibliographic information)。它与传统图书馆提供的文献信息不同,是有关机构、团体、服务、规划、工程、事件等的信息。与文献信息相比,这些信息的最大特点是:它们是动态的、不断发展着的,因而具有更强的新颖性和及时性。开发社区信息,是美国图书馆优秀服务传统的延伸与发展。美国图书馆不满足于在自动化系统中仅仅收入有关文献的信息,文献以外的活信息也成为他们收集的重点。USMARC 社区信息格式正是美国图书馆发扬注重社区服务、注重参考工作的优秀传统的自然结果。

1 USMARC 社区信息格式的制订经过

大约在1985年时,美国公共图书馆协会社区信息专业组技术委员会发现,一些图书馆已经自发地以机读形式把社区信息输入到本馆的自动化系统中。该委员会经过摸底,集成一份名单,共收录15个州的27所提供自

动化社区信息服务的图书馆。1988年,该委员会与这些图书馆联系,请他们把本馆自动化社区信息记录中包含的数据元素开列一份清单。从提交的清单来看,各馆采用的数据元素有许多相同之处。根据这些清单,委员会拟定了一个“社区信息标准化数据元素表”。这些图书馆提供的数据元素虽然很接近,但采用的格式有较大差别:有的试图将社区信息填入原有的USMARC文献格式中;有的在文献格式基础上改良出了自用的社区信息格式;有的则采用了非MARC格式。

1989年,技术委员会与美国国会图书馆网络开发和MARC标准办公室(Net Dev/MISO)联系,并送上一份《社区信息标准化数据元素表》,希望该办公室拟定一个能够容纳这些数据元素的USMARC社区信息格式。技术委员会要求,社区信息和原有的文献信息要能够共存于同一个数据库中,以便把同一主题的文献信息和非文献信息集中在一起,方便读者检索。例如,当一个读者寻找有关如何对待吸毒少年的信息时,在这样的数据库中,他应该不仅能够找到有关该论题的文献,而且还可以了解到与该问题相关的机

构和服务的情况。

1990 年初, Net Dev/MSO 在 U SMARC 图书格式基础上, 增删了一些字段和子字段, 拟定了一个文件, 提交 1990 年 6 月 ALA 年度会议评审。评审者对这个文件不甚满意。他们认为, 社区信息与文献信息差别较大, 例如, 文献格式中的 008 字段, 以及其它一些必选字段(300 字段、245 字段)不适用或基本不适用社区信息, 而社区信息则需要大量使用附注字段。因而, 将社区信息数据元素装入文献格式中, 显得比较牵强。

为此, 1990 年 12 月初, 又召开了一个扩大范围的评审会议, 参加者包括社区信息的提供者、MARC 专家、图书馆系统供应商、高校及公共图书馆员。会议一开始, 首先宣布, 美国国家标准 Z39.2《美国国家文献信息交换标准》正在修订。修订的意向之一是, 将标准名称中的“文献的”(bibliographic)一词删除, 改为《美国国家信息交换标准》。头标区 07 字符位置也不再定义为“文献级别”(bibliographic level)。U SMARC 格式是以 Z39.2 为基础的。Z39.2 的修订意味着不仅可以有面向文献的 U SMARC, 还可以有面向非文献实体的 U SMARC, 没有必要把社区信息硬塞进文献格式中去。于是会议决定, 社区信息格式将作为独立于文献格式之外的一个特殊格式来制订。

会后, Net Dev/MSO 根据会议精神草拟了一个新文件。经过几次评审之后, 该文件以《U SMARC 社区信息格式》(U SMARC Format for Community Information)为题, 终于在 1992 年获批准^[1]。1993 年 2 月, 《U SMARC 社区信息格式》作为《U SMARC 简明格式》(U SMARC Concise Formats)的四个格式之一正式出版。

2 U SMARC 社区信息格式的特点

U SMARC 社区信息格式与其它三个

U SMARC 格式的结构一样, 由头标、目次和可变字段三部分组成。

U SMARC 社区信息格式的头标区删除了个别字符的定义和代码值。如, 17~19 字符位改为了“未定义”, 05 字符位删除了“a”和“p”两个代码值。其头标区的主要特点是, 在 06 字符位置(记录类型)设代码值“q”, 表示“社区信息”, 把 07 字符定义为“数据种类”, 分别用 5 种代码标识 5 类社区信息:

n, 标识个人信息。表示该记录提供有关某个具有特定专长的个人(如故事员、市领导等)的信息;

o, 标识组织信息。表示该记录提供有关组织或团体(如公司、俱乐部)等的信息;

p, 标识规划或服务信息。表示该记录提供有关一个活动(如驾驶员培训、血库等)的信息。这种活动贯彻了一个组织或团体的目标;

q, 标识事件信息。表示该记录提供有关预期发生事件(如演讲、定期会议等)的信息;

z, 标识其它信息。表示该记录提供有关上述代码不能反映的实体(如位于大学校园的天文馆这类设施)的信息。

在 U SMARC 社区信息格式的可变控制字段(0XX 字段)中, 增加了一个 004“编码日期定长字段”, 占 32 个字符位置, 分别记录 4 类日期: 活日期(需要复审的日期)、删除日期(系统中被删除信息的日期)、事件或规划开始的日期、事件或规划结束的日期。每一类日期占 8 个字符, 4 位记年, 两位记月, 两位记日。不适用的日期类型用代码填充, 4 类代码均不适用, 则省略 004 字段。007 字段仍名为“物理描述固定字段”, 改为 10 个字符位置, 专门记录为残疾人提供的设施, 如是否有专用电梯、休息室, 是否使用手语等。008 字段(定长数据元素字段)设 15 个字符位置, 提供记录的整体信息或特殊社区信息。如记录的入档日期、是否录用或提供志愿者, 是否照管儿童, 是否提供会议室或其它设施, 以及实

体所用的语言等。

可变数据字段 010~ 9XX 分为 10 个块。它们是:

- | | |
|----------------|---------------|
| 0XX 控制信息、号码和代码 | 5XX 附注 |
| 1XX 基本名称 | 6XX 主题检索字段 |
| 2XX 题名、地址 | 7XX 主题以外的附加款目 |
| 3XX 物理信息等 | 8XX 替代图形 |
| 4XX 系列信息 | 9XX 为当地实施保留 |

其中, 0XX、5XX、6XX、9XX 的英文名称与文献信息格式完全相同。社区信息格式共有 68 个可变数据字段, 字段设置尽可能与文献格式相对应。如 052 地理分类代码字段、511 参加者或执行者附注字段和 546 语言附注字段都直接移植了文献信息格式的相应字段。1XX~ 7XX 主要字段设置如下: 1XX 设 3 个字段: 100“基本名称——个人”、110“基本名称——机关团体”和 111“基本名称——会议”, 分别记录 3 类实体的基本名称。

2XX 设 3 个题名字段和 3 个地址字段。题名字段有: 245“题名”、246“题名变异形式”和 247“前题名”字段。与文献格式不同的是, 245 字段不再是必选字段。因为大部分实体的名称都记录在 1XX 块中, 245 只记已经结束的事件的名称。对于正在发展中的事件, 其题名也应记录在 1XX 块中。地址字段是社区信息特有的字段。其中, 270 基本地址字段和 271 其它地址字段与 1XX 字段一起使用, 当一个实体有不止一个地址时, 例如某机构有一个所在地址和一个通讯地址, 就要同时使用 270 和 271 字段。而 275 字段是与 245 字段一起使用的。

3XX 记实体的营业时间 (301)、从属实体 (303)、现有会议室和设施 (311)、现有设备 (312)。

4XX 只有一个 440“系列题名”字段, 记录事件所属系列的题名。

5XX 设 21 个附注字段。500 是“一般附注字段”, 501~ 587 是各种专门附注字段。这些专门附注字段可能带有显示常用词 (dis-

play constant)。其中, 520 和 531 是使用率较高的两个字段。520“描述附注”字段描述实体的目标或服务; 531“符合条件者、费用、手续附注”字段记载实体的特有信息, 如使用、出席人数、会员费、加入条件等。

6XX 记录主题检索点。与文献格式不同的字段有 630“主题附加款目——出版物题名”字段, 记主标目为一个出版物题名的主题词; 653“索引词——非受控的”字段, 记与标准题词/叙词不一致的索引词; 654“主题附加款目——分面论题术语”字段, 记录根据分面词表拟定的主题词。

7XX 记录其它检索点。除设有与文献格式相应的 700、710、711 字段外, 还设有 730“附加款目——出版物题名”字段, 记录出版物题名做入口词的附加款目标目; 740“附加款目——具体规划题名”字段, 记录实体所设具体规划的题名。

3 U SMARC 社区信息格式的应用

这里介绍美国一些大学、公共图书馆和图书馆系统供应商对 U SMARC 社区信息格式的应用情况。

(1) 高校图书馆。缅因大学是一所州立大学, 有七个校区。主校区位有一个自动化图书馆系统, 称作 URSUS。它与其它各校区图书馆、缅因州图书馆、Bangor 公共图书馆, 以及校外的 83 个缅因州社区学院图书馆相连, 并与因特网连通。缅因大学的七个校区都分别开展了收费、免费或义务的公众服务 (publicservice), 利用大学和社区信息资源, 向缅因州公民提供不计学分的课程、服务和活动。最初, 大学对此未作整体规划。1990 年, 缅因大学试验在 URSUS 系统上建立一个全校的联机公共资源数据库, 使之成为既可以检索图书馆联机目录, 又可以作为公共服务网络中心的信息系统。这个联机公共资源数据库由个人、规划和服务、设施及事件的

记录组成。记录格式采用当时还未正式出版的 U SMARC 社区信息格式。数据库记录分为三类: 教员和行政管理人员个人记录、规划和服务的记录、设施记录。他们采用发放调查表的方法收集数据。一张调查表发给教员和行政管理人员。填写本着自愿的原则, 凡愿意将个人信息输入数据库的个人, 在调查表中可填写自己的专长、咨询经验和希望为之服务的客户类型。另一张调查表调查了大学现有的各种中心、研究所的信息^[2]。印第安那大学也是一所州立大学。与其它大学一样, 学校常常爆发种族事件。在 90 年末的一次学校少数民族事务会议上, 副校长助理 Coldwell-Colbert 博士提出, 学校应当有一个集中提供多样化信息的中心。所谓多样化信息, 这里指不同文化背景、不同民族、不同能力的人(包括残疾人)所需要的各种信息。在座的图书馆代表 Pat Ensor 曾参加过学校有色人种老师的招聘工作, 对建立这样一个中心的必要性深有体会。会后, 他组织人力在印第安那大学图书馆 LU SHI 系统上开发了一个“多样化信息联机(D DN)”数据库, 集中、及时地提供大学所属各学院、学校和行政部门各项有关活动的进展情况。数据库收录有关个人、事件、课程、奖学金、教学计划、展览等实体的记录。记录采用未正式出版的 U SMARC 社区信息格式。该数据库 1992 年秋向公众开放, 用户可以通过拨号方式, 从印第安那自动化网络进入州立大学图书馆, 也可以通过因特网访问这个数据库^[3]。

(2) 公共图书馆。美国公共图书馆在提供社区信息方面一直走在前列。至 1989 年, 马里兰东海岸的 8 个农村县公共图书馆都已经编制了社区信息的卡片文档, 其中一个馆还建立了社区信息数据库。但各馆所用的记录格式不同。且大多数馆都没有及时更新这些文档的信息。1992 年, 他们启动 Shorebirds 工程, 开始建立一个 8 馆联合的 I&R (Information & Referral) 数据库, 收录了东

海岸地区俱乐部、组织和规划的 3000 个记录。他们参照未完全脱稿的 U SMARC 社区信息格式, 统一了 8 个馆的记录格式。并计划在 1993 年建立马里兰州州级 I&R 数据库, 届时正式采用新推出的 U SMARC 社区信息格式^[4]。此外, 在 U SMARC 格式出台之后的 1994 年 1 月, Queens Brough 公共图书馆的专门服务部(该馆的对外服务机构)和编目部也曾合作开发了一个联机社区信息数据库, 称做“infolinq 社区服务”^[5]。

(3) 图书馆系统供应商。在国会图书馆动手设计 U SMARC 社区信息格式的前几年, 一些自动化图书馆系统的供应商就已经应图书馆的请求, 尝试在联机数据库中以机读形式表示社区信息。由于没有标准格式作指导, 各家对数据元素的取舍、格式的设计都不相同。U SMARC 社区信息格式草案下发之后, ALA 曾对图书馆系统供应商做过一次调查, 了解他们提供社区信息的现状和对新标准的反映。共发放调查表 17 份, 收回 16 份。调查结果表明, 在 16 家供应商销售的系统中, 仅一家不含社区信息功能。在余下的 15 家中, 4 家已采用新标准, 9 家用类似 MARC 的格式, 2 家用非 MARC 格式。在接受调查时, 15 家系统商已经知道发布了新标准。在 16 家系统供应商中, 15 家认为需要统一的标准格式, 并表示愿意采用新格式^[6]。

4 U SMARC 社区信息格式的问题

以往每一种文献格式的设计, 都是编目规则问世在先, 机读格式推出在后。机读格式主要说明记录的结构、字段、子字段的内容和标识符。数据元素的界定、取舍、标目的选取、标目形式的确定, 则由编目规则负责。唯有美国社区信息格式是在未制订编目规则的情况下颁布的。美国公共图书馆协会社区信息专业组技术委员会的主席解释这样做的原因时说, 各机构在自行建立社区信息数据库的路

上已经走得太远了。也就是说,当时的形势已经不容许等待编目规则制订之后再设计社区信息格式。正因为如此,尽管这个格式对社区信息数据库的设计发挥了重要的指导作用,但在应用这一格式创建记录的过程中,也遇到了一些问题。

4.1 信息来源

个人、组织机构、活动、服务和规划不像文献那样,有统一的题名页或类似题名页的信息源。其信息源的确定是一个更为复杂的问题。机读格式是不负责说明信息源的。编目机构不得不自行规定信息源。有的机构规定,个人记录以其本人介绍为依据,机构、服务和规划等以这些实体本身发表的宣传材料为依据。文献若以作者和出版者的评介为编目依据,可能会丧失客观性,而非文献实体这样选取信息源,似乎有一定可行性。因为实体能够提供什么信息,能够为哪些客户服务,确实要他们自己来说明。但是,实体本身介绍中的促销言词是否也可能误导系统的检索者?对于夸张的信息,编目员应如何剔除?这些也是社区信息编目需要解决的问题。

4.2 编目对象

非文献实体种类之繁多更甚于文献,如何选择非文献记录的编目对象,向我们提出了一些新问题。如一个大团体下设若干小团体,是以大团体为编目对象,还是以小团体为编目对象?一个机构开展了多项服务,设立了多项规划,那么,应该以机构为编制记录的对象,还是以服务或规划为对象,亦或它们各自都作为编目对象?个人大多属于一个组织,个人又是服务与规划的提供者,个人与组织、服务、规划的记录如何取舍,如何联系?这些问题与多层次文献和多载体文献的著录有相似之处也有不同之处,需要认真加以研究。

4.3 数据界定

美国社区信息格式引进了一些新的术语,但并没有做严格界定。如社区信息格式规定,规划的题名记录在 110 字段,事件的题名

记录在 245 字段。但是,规划与事件如何区分?据说,国会图书馆后来做过一项补充说明:已知起始时间和结束时间的为事件,正在进行的事件为规划或活动。但是,由于社区信息格式毕竟不是编目规则,其中涉及的许多其它术语还有待在编目实践中,逐渐划定概念的内涵与外延。这个问题不解决,即使有了统一的格式,仍然会出现同一实体的记录不一致的情况。

4.4 数据取舍

在众多的有关编目对象的信息中,选择哪些,摒弃哪些,历来是由编目规则定夺的。机读格式只提供了一个信息记录的格式。而且,美国社区信息格式是面向个人、组织、规划和服务、事件及其它 5 种信息制订的。至于每一种信息选用哪些字段,则有待编目规则做出详细规定。

4.5 标目形式

如团体类实体的标目是用多级标目,以其上级机构的名称为入口词,还是以具体团体的名称为标目?现在有些编目机构是参照 AACR2 团体标目的选取办法来执行的。但是,社区信息格式囊括许多最基层的实体,数量更多,情况更复杂,应当有一个专门的规则来解决相应的问题。

4.6 系统维护

社区信息记录的最大特点是,除已经结束的事件外,绝大多数实体是动态的、不断变化的。例如,组织机构宗旨、从事领域和名称的改变;个人兴趣的转移和专长的扩展;教学计划、服务和活动的发展等,都使相应记录随时可能失去真实性。因此系统必须建立一种维护机制,保证记录的永久有效性。

5 建议

绝大多数非文献实体的信息比文献信息更及时、更新颖,因而具有更大的使用价值。图书馆作为信息单位之一,占据这一领域,才

能更充分地展示自己在信息社会举足轻重的作用。目前我国共享编目刚刚起步,各馆忙于分散建立自己的文献数据库,编目部门还没有从文献编目的重负下解脱出来,自然无暇顾及非文献信息的编目问题。共享编目发展将证明当前一些图书馆所做的低水平建库工作是得不偿失,甚至是完全无效的。要推动我国的图书馆事业尽快赶上发达国家图书馆界前进的步伐,图书馆就应放弃小我的利益,携起手来,努力推进共享编目工作。一旦共享编目在我国具有了一定规模,各馆的原始编目量大大降低,编目部门就可以腾出手来,不失时机地向社区信息领域迈进。为了在这一天到来之时,我们能有的充分的精神准备,从容地跨入这一领域,我建议:

(1) 关注国际、国外社区信息格式的发展。编目理论界、机读目录设计者应仔细研究 USMARC 社区信息格式,切实搞清楚其字段、子字段、数据元素的含义和设置的意义,并注意国际编目界在有关方面的进展动向。

(2) 广泛征集非文献实体的数据元素。标准化文献情报技术委员会从现在起,就应开始调查有哪些机构,特别是信息业有关机构已经建立了非文献数据库,并请他们提供数据库中包含的数据元素。在此基础上汇总、拟定出符合我国国情的社区信息基本数据元素表。

(3) 及早着手制订非文献实体的编目规则。我们应接受美国的教训,对非文献实体的信息源、格式、标识符号、数据定义、数据取舍、标目选择、标目形式、参照尽早进行调查

研究,争取在社区信息MARC格式出台之前,制订出比较完整、实用的编目规则来。

参考文献

- 1 Bruns, Phyllis The USMARC Community Information Format: a history and brief description, In: Information Technology and Libraries, 1992, Vol 11 No. 4 387~ 396
- 2 Lutz, Marilyn Using the community information format to create a public service resource network Sharon Quinn Fitzgerald, and Thomas Zantow. In: Information Technology and Libraries 1992 Vol 11 No. 4 373~ 383
- 3 Ensor, Pat Diversity information online: the development of a CIF-based database In: Information Technology and Libraries 1992 Vol 11 No. 4 384~ 387
- 4 Gray, Douglass MARC and I & R. In: Information Technology and Libraries 1992 Vol 11 No. 4 397~ 398
- 5 Ranta, Judith A. Queens Borough Public Library's guidelines for cataloging community information In: Cataloging & Classification Quarterly. 1996 Vol 22(2). 51~ 69
- 6 McClintock, Patric J. Community information and automated library systems In: Information Technology and Libraries 1992 Vol 11 No. 4 399~ 403

刘苏雅 北京大学信息管理系副教授。通讯地址:北京市海淀区。邮编 100871。

(来稿时间: 1997. 8. 20, 编发者: 赵薇)