

# 关于网络信息资源编目的探讨

高 红

**【摘要】** 文章探讨了网络资源编目的特点、网络资源编目的可行性,介绍了几种常见的元数据格式的编目应用,最后寄望我国图书馆界关注网络资源编目待解决的课题。

**【关键词】** 网络信息资源 电子资源 编目实现

随着 Internet 的发展以及网络信息资源的不断增多,正如传统图书馆中图书、期刊的收藏需要编目一样,网络信息资源编目应运而生。目前,这一新生事物越来越受到人们的重视。

## 1 网络资源编目的特点

网络信息资源是指“通过国际互联网可以利用的各种信息资源”。计算机、光学存贮和网络技术使得信息的制作、存贮、传递和使用发生了根本变化,从编目的角度讲,明确如何识别各种形式的文献、如何处理网络信息的易变性、如何解决网络资源易于复制的问题,有助于理解网络资源编目的含义。

### 1.1 格式多样化

与传统文献形式相比,网络资源格式的多样化更加突出。由于因特网上存在大量正式和非正式标准,如 ASC II、SGML、HTML、WORD、TEX、RTF 等等,所以文献的类型也不相同,那么如何在编目中处理这些新型格式? 1993 年随着 USMARC 书目记录格式的一体化实施,采用在 5xx 字段记录资源格式的内容,如:

- 500 Map on lining paper
- 516 Computer programs
- 516 Text(Law reports and digests)
- 516 hypertext(electronic journal)

以上各例对信息资源格式进行了描述,对提高书目的捕获能力及满足用户检索需求起到了重要作用。

### 1.2 资源的易变性

网络信息资源容易修改,不像传统文献的书目记录各个项目真实可靠,用户可以很有把握地利用。因此,如何保证网络信息的真实可靠将成为重要课题。

### 1.3 资源的易复制性

网络资源极易复制,由此带来许多新问题。出于安全和缩短用户与特定资源距离的考虑,同一文本被保存在众多的网址上,造成重复,书目记录和这些网址的关系以及对网址的著录选择,都成为网上编目须解决的问题。

### 1.4 特有的检索方式

在网络环境下,资源的实际存贮位置与其检索地点可能十分遥远,检索用户须有特定资源的通讯地址(通常指主机的 IP 地址),精确地址说明特定资源的位置,会给检索带来极大方便。此外,编目还必须提供充分信息,便于用户以后检索。同样,USMARC 采用 538—系统细节附注字段,对有关检索方式字段进行描述,如:

- 538 Data in extended ASC II character set
- 538 System requirements: IBM 360 and 370; 9k bytes of internal memory; OS SVS and OSMVS
- 538 System requirements: PC; World Wide Web access

上面例子分别说明数据采用的是 ASC II 扩充字符集和对系统要求的说明。

## 2 网上资源编目的可行性

目前,我国图书馆还没有对网络资源进

行编目的经验,但国外的成功实验,证明了网上资源编目的可行性。

### 2.1 USMARC 一体化的实现

由 OCLC 和美国国会图书馆发起并实施的修改 USMARC 格式之建议,提出了 MARC 的“电子定位与存取”字段—856 字段,开创了编目的新时代。当文献资源或其子集可以通过电子方式获取时,可在书目记录中启用 856 字段,该字段的超文本连接有助于用户获得和连接网上资源。此项提议于 1993 年 1 月,由“书目信息机读格式委员会”通过,成为记录网络信息资源的重要字段。

### 2.2 OCLC 的 InterCat 计划

OCLC 的 InterCat 计划有 500 多个成员馆参与,项目实验期间对 18000 条网络信息资源进行了编目,已经建立了拥有十万条记录的数字化资源书目记录数据库,可提供题名、作者、关键词和高级检索、浏览等检索途径,并证实 MARC/AACR2 可适于网络资源编目。

### 2.3 网上资源编目步骤

首先,由负责馆藏发展的馆员选择符合馆藏计划与用户要求、具有研究价值和相对稳定的网络资源,并填写“网络资源著录请求表”;其次,由编目人员按相应格式对网络资源进行编目;再次,由自动化系统人员将网址和检索方式添加在 856 字段中。

### 2.4 元数据格式

USMARC 格式对网络资源进行编目,速度慢且成本高,赶不上网络信息资源的发展速度,因此,在信息技术界,许多研究者设计了多种元数据格式来描述电子信息资源。元数据就是数据的数据,元数据与网络信息资源之间的关系如同书目数据与图书之间的关系一样,元数据能否帮助图书馆实现网络信息的高效编目,是目前图书馆界关注的热点,上海图书馆元数据课题组就编制了 DC 元素与 USMARC 格式的元素对照表,做了元素的一一映射。

## 3 几种元数据格式的编目应用

### 3.1 USMARC 对网络资源进行编目

USMARC 因广泛用于书目记录数据而为人们所熟知,增加的 856 字段—电子定位与存取字段,主要记录被著录的数字对象或与之相关的其他电子资源的存储地址和存取方式,为可重复字段(repeatable)。指示符 1 表示检索方式,如 0—Email、1—FTP、2—Telnet、3—Dial-up、4—Http、7—检索方式由子字段 \$2 指定;指示符 2 表示 856 字段标识的电子资源与编目文献间的关系。该字段包括 27 个子字段,主要包括:主机地址(域名)、检索号码、文件压缩信息、目录/文件路径、电子文件名、统一资源名称、用户名(email@之前的名称)、口令、操作系统、文件大小、电子格式类型等,例如:

```
856 41 $ uhttp:// www.jstor.org/
journals/0277903x.html
```

第二指示符的值 1 表示编目文献本身不是电子资源,但可以通过记录在 856 字段的 URL 地址获取与编目文献内容相同的电子资源。

```
856 2 # $ aanthrax.micro.umn.edu
$b128.101.95.23
```

该字段包含一个符合远程登录协议的电子资源检索信息,\$a 和 \$b 分别记录提供电子资源的主机域名和 IP 地址。

### 3.2 Dublin Core 网络资源编目新思路

都柏林核心元数据集是 1995 年由美国 OCLC 和 NCSA(国家超级计算中心)共同研究产生的一套描述网络电子文献的方法,这一方法的核心是如何用一个简单的标准化的元数据记录来描述种类繁多的网络信息,使非图书馆专业人员也能够了解和使用这种内容描述方法,达到更有效地描述和检索网上资源。DC 由 15 个基本元素组成,分成三大部分:

a. 内容描述部分:题名 (Title)、主题 (Subject)、说明 (Description)、来源 (Source)、

语种 (Language)、关联 (Relation) 和覆盖范围 (Coverage);

b. 知识产权部分: 创建者 (Creator)、出版者 (Publisher)、其他责任者 (Other contributor) 和权限 (Rights);

c. 外形描述部分: 日期 (Data)、类型 (Type)、形式 (Format) 和标识符 (Identifier)。

都柏林核心元数据结构简单, 共 15 个数据元素, 其含义易学易记, 并且国际通用, 已有德语、日语、西语等不同语种版本, 还可以与其他元数据元素连接使用。随着都柏林核心元数据集在世界各国及各学科领域的广泛应用, 15 个元素已不能满足人们对资源作具体描述的需要, 为了提高著录的精确性, DC 数据集引进了修饰词的概念, 如语言修饰词 (Lang)、构架修饰词 (Scheme) 及子元素修饰词 (Subelement), 尤其是通过 Scheme 把 MARC/ AACR 的优点和各种已有的分类法、主题词表吸收进来, 增强了 DC 的描述性和权威性。修饰词分元素修饰词和编码体系修饰词两种, 前者缩小了元素的含义范围, 使其具有专指性, 主要在题名、说明、日期、格式、关联和覆盖范围元素中使用; 后者有助于对元素修饰词值的理解, 主要在主题、日期、格式、标识符、来源、覆盖范围中使用。可参考下面的例子:

\* 题名 (Title) 修饰词

元素修饰词:

Alternative (交替题名)

名称: 交替题名

标签: 交替题名

定义: 任何一种用资源正式题名以外的其他题名。

注释: 该修饰词可以包括缩写题名和翻译题名。

\* 主题修饰词

编码体系修饰词

LCSH (美国国会图书馆主题词表)

名称: LCSH

标签: LCSH

定义: 美国国会图书馆主题词表

参见: <http://ids.udmercy.edu/library/lcsh.htm>

### 3.3 REACH 著录单元集合

REACH 著录单元集合是美国研究图书馆组织 (RLG) 的一个研究项目成果, 用于文化艺术类作品及珍善本的著录, 由 20 个数据单元构成, 包括作品类型 (Type)、载体材料和技术 (Materials and techniques)、题名 (Title)、责任者 (Creator)、责任方式 (Role)、日期 (Date)、主题 (Subject)、版本 (Edition)、相关作品 (Related work)、收藏地点 (Repository place)、收藏号 (Repository number)、附注 (Descriptive note) 等。

网络信息资源的编目是一项复杂而艰巨的任务, 解决复杂问题的简单而有效的办法便是传统方法的延伸和扩充。尽管国内一些图书馆在网络资源编目研究方面开始尝试, 但与国外比较, 无论在理论研究、技术实施还是软件开发方面都存在不小的差距, 对此, 我们应有充分的认识, 我们要关注网络信息资源编目待解决的课题, 如对网络资源易变性的处理、856 字段的完善、元数据格式的应用等。展望未来, 网络资源的编目必将在世界编目发展史上写下新的篇章。

#### 参考文献

- 1 吴建中. DC 元数据. 上海科学技术文献出版社, 2000
  - 2 CALIS 联机合作编目手册. 下册. 北京大学出版社, 2000, 12
  - 3 肖琬. 美国国家数字图书馆项目进展. 情报学报, 1998, 17(3)
  - 4 王绍平. 合作编目—网络资源书目控制的必由之路. 图书馆杂志, 2000(4)
- (作者信息: 国家图书馆采编部, 邮编: 100081. 收稿日期: 2001 年 7 月 18 日。)