

●张 岩

数字图书馆给编目工作带来的变化

摘 要 数字图书馆给传统编目工作带来了许多变化,包括编目对象、编目环境、编目理论、编目格式和对编目人员要求等的变化。参考文献7。

关键词 数字图书馆 编目工作 MARC 都柏林核心集 USMARC

分类号 G254

ABSTRACT Digital library has brought quite a lot of changes to traditional library cataloging, including the changes of objects, environment, theory and format of cataloging, and the requirements to catalogers. 7 refs.

KEY WORDS Digital library. Library cataloging. MARC. Dublin Core. USMARC.

CLASS NUMBER G254

数字图书馆的产生和发展,意味着传统图书馆在信息时代的重大转型,图书馆的许多方面将产生变化。本文将从编目工作角度,分析数字图书馆在编目对象、环境以及人员要求等方面产生的变化。

1 编目对象的变化

传统图书馆主要以图书、期刊等印刷型出版物为编目对象,编目对象容易确定。数字图书馆主要以数字化的信息为编目对象,是在一种动态开放的环境中对数字信息进行描述和标引,编目对象往往难以识别。比如:在网络环境下,一系列相关信息源汇集在一个主页下,或是一个数据库中的一个组成部分,这时对信息源的著录是以整体为编目对象,还是以具体的个体为编目对象?又比如:网页的超文本链接方式使用户可以选择最适合自己的次序阅读文献,用户阅读文献的方式决定了文档之间的关系,这就使得编目员难以确定符合用户需要的编目对象。

2 编目环境的变化

传统图书馆面对的是以纸介质或缩微介质等实体介质为载体的环境。数字图书馆面对的是以网络技术、多媒体技术为支撑,以数字化信息为表现形式的虚介质为载体的环境,各种联机数据库、电子出版物、网络信息资源等数字化信息激剧增长,数字图书馆在整个社会中将扮演信息中介者的角色。

3 编目理论的变化

作为图书馆组织信息和提供检索的主要手段,描

述性编目的理论已发展了一百多年,环境的巨变,正在动摇这些传统编目理论存在的基础。在网络环境下,“书目数据”这一狭隘的概念将被废弃,代之以“元数据”这一较为模糊,能为更多人使用的概念。元数据是用于提供某种资源的有关信息的结构化数据,是一种数据结构标准。简单地说,是关于数据的数据,它规定了数字化信息的组成,其作用在于规范对数据的组织,以利于检索和传递。通过元数据,能使我们了解到某个站点的资源类型,某个网页的标题、著作、主题及内容摘要、权限管理等信息。对网络化信息的管理,越来越依赖于对元数据的有效管理。

4 编目格式的变化

传统的印刷型出版物是以MARC作为机读目录标准格式,广泛用于各种书目记录。但对于数字图书馆信息资源的著录不太适用,因而用于描述数字图书馆中数字资源的元数据格式应运而生。经过几年的实践,元数据已发展出多种用来进行数字图书馆数字信息资源描述和标引的格式,由于数字图书馆收藏的内容不同,对格式的要求也就不同。因而,用于数字信息资源的元数据格式也有多种。目前,较常用的元数据格式有USMARC格式、都柏林核心集(Dublin Core)等。

4.1 USMARC 格式

USMARC随着“美国回忆”项目的启动,已被推荐为数字图书馆所使用的几种元数据格式之一,它进行了多次局部修改,以适应数字图书馆信息资源的编目。所作出的修改主要体现在以下几个方面:

(1) 使用 5 一字段记录信息资源格式和检索方式。

与印刷型出版物相比,数字图书馆信息资源的格式呈多样化。如:ASCII、SGML、HTML 等,不同格式的文献类型也不同。如网页中,既有用 HTML 语言编制的 ASCII 文件,也有与其它网页相链接的文本、声音、图像等内容。另外,数字图书馆信息资源的检索方式也是多种多样的:有 WWW、Gopher、Telnet、FTP 等方式。因而,采用 5 一字段来描述信息资源格式和检索方式。

(2) 增加 856 字段。

电子信息资源代表了一种新的信息交流方式,用户与电子文献之间的交互性使用户可以直接浏览文献全文,或在由电子文献组成的全文数据库中实现检索。对用户而言,重要的是提供有效的获得途径——网址,以及一些相关信息(如文件的字节数等)。为了实现从书目信息到文献本身信息的检索,USMARC 增加了 856 字段,即“电子信息资源地址与存取”字段,主要记录被著录的数字对象或与之相关的其它电子信息资源的主机地址、电子资源名、检索方式等具有检索意义的信息,建立了从书目记录到电子信息资源的链接,并将 856 字段列为必备字段。856 字段的设置,开创了编目的新时代,其超文本链接,有利于读者获得电子信息资源,从而实现了从目录到资源本身的链接,为读者检索提供了极大方便。

(3) 制定“机读目录文献类型定义”。

基于 MARC 格式必须在专门的系统中使用,为了更好地促进 MARC 格式在网络环境下的应用,美国国会图书馆正在研究制订“机读目录文献类型定义”(Machine Readable Cataloging Document Type Definition,简称 MARC DTD),使基于 ISO2709 的 MARC 格式数据自动转换到基于 ISO8879 的 SGML 格式,以适应各种网络软件和浏览器,而 MARC 格式本身也会产生相应变化。

4.2 都柏林核心集

都柏林核心集是针对网络资源的描述而制定的。随着网络资源的迅速增加,编目人员已不可能对其全部编目,只能选择一些著名数据库及较稳定和权威性的网址进行编目,即使这样,编目量依然非常大。要从根本上解决对网络资源的编目问题,就需要拓宽编目渠道,即由著者及出版者对其提供的电子资源进行编目。都柏林核心集结构简单,只包含 15 个数据元素,且每个元素含义都通俗易懂,

非专业人员也能很快掌握,而且还可以将所包含的传统字段通过映射转换为 MARC 格式,与图书馆原有的目录联为一个整体。但是非专业人员著录的题名、责任者、主题词有可能是非规范的,这就需要有经验的编目人员进行规范工作。

5 对编目人员提出了更高的要求

在联机环境下,编目工作已实现对数据的共享,原始编目量大大减少,从原始编目中解脱出来的编目人员应把重点放在数字信息资源的整合。数字图书馆需要大量使用高技术设备和计算机网络,它对编目人员提出了更高的要求。编目人员不仅要具有计算机编目方面的知识,良好的外语水平,还应具备较高的计算机和网络技术方面的知识,这对编目人员来说是一个不小的挑战。数字图书馆信息资源的格式多种多样,已有不同制作形式的电子资源出现在网上,编目人员必须熟悉各种电子资源的格式,才能准确地加以描述。因此编目人员应加强计算机和网络技术方面的知识,更新自身的知识结构,尽快掌握数字图书馆编目方法,同时利用网络资源建立本馆有特色的专业网络资源数据库以及各种特藏的数字化。

数字图书馆的发展意味着传统图书馆在信息时代的重大转型。它的产生,给编目工作带来了许多有待解决的难题,这将促使我们不断地探索新理论、研究新方法,推动编目事业的更大发展。

参考文献

- 1 胡燕蕊. 谈用编目的手段整理网络资源. 现代图书情报技术, 2000(4)
- 2 朱蓓, 司莉. 远程电子文献编目浅议. 大学图书馆学报, 1999(2)
- 3 真濤, 王新. 未来图书馆编目技术. 情报理论与实践, 2000(3)
- 4 王新. 都柏林核心集综述. 情报理论与实践, 2000(5)
- 5 郑巧英, 杨宗英. 数字图书馆原型体系结构初探. 情报学报, 2000(3)
- 6 林海青. 数字图书馆的信息组织. 中国图书馆学报, 2000(1)
- 7 肖珑. 元数据格式在数字图书馆中的应用. 大学图书馆学报, 1999(4)

张 岩 北京大学图书馆编目部馆员, 北京大学信息管理系在职研究生. 通讯地址: 北京大学图书馆. 邮编 100871.
(来稿时间: 2002-06-17)