

图书馆声像信息资源的 Web 检索与播放

孟 珊, 张 焱, 黄 闽

(河北师范大学图书馆, 河北 石家庄 050016)

[摘 要] 文章提出了要实现声像信息资源在网络上的快速检索与播放, 所应具备的四个基本要素以及实现的具体方法, 希望借此能引出更深入的研究。

[关键词] 声像信息资源; 元数据; 对象数据; 搜索平台; Web 检索

[中图分类号] G255.73 **[文献标识码]** A **[文章编号]** 1002-1167(2002)04-0067-02

声像信息资源具有直观、生动、迅速、准确等独特的优点, 它能充分调动读者的视觉和听觉感官, 使听觉和视觉化虚为实, 把一些抽象概念形象化、实体化, 将一些综合性教育内容的复杂过程进一步简化, 使读者在生动、轻松的学习环境中获取知识, 同时声像信息资源存储量大、保存周期长等特点, 使它越来越受到广大读者的欢迎。随着计算机技术、网络技术以及多媒体技术的迅速发展, 声像信息成级数增长, 图书馆所收藏的声像信息资源品种及数量也随之大量增加。图书馆在声像信息资源方面所提供的服务也由原来的视听阅览室、多媒体阅览室的服务扩展到广域网上的信息服务。但声像信息资源及其类型如何? 在浩瀚的信息海洋中又如何快速查找和实现播放这类媒体? 这是信息用户网上搜寻时常遇到的问题。图书馆等信息服务部门应针对以上问题, 提出切实可行的实现方法。本文正是通过我馆的一些实践, 提出声像信息资源要实现 Web 检索与播放, 所应具备的基本要素以及实现方法, 并希望借此能引出更深入的研究。

1 图书馆声像信息资源要实现 Web 检索与播放, 应具备四个基本要素: 声像信息的元数据资源、对象数据资源、存储声像信息资源的数字仓库及搜索平台

1.1 元数据资源

声像信息元数据, 可采用中国机读目录格式(CN-MARC)并在此基础上定义必要的扩充字段, 对原始声像信息资源(即对象数据)进行描述, 其作用是使数据用户能够通过检索, 对相关原始声像信息资源进行选择定位, 并提供调取对象数据的手段。

1.2 对象数据资源

对象数据存储格式应分为原始备份格式和服务用格式两种:

(1) 原始备份格式采用无压缩格式, 是在对各种信息的数据化采集过程中产生的信息格式。无压缩格式信息的保存

形式可以是磁带、光盘等。这种格式的信息经压缩处理后形成服务用格式, 用于网上播放。图书馆仍应保存这部分信息备份, 其意义是随着计算机存储技术、网络通讯技术和数据压缩技术的发展, 将来可能采用新的信息压缩技术对信息进行二次加工处理, 采用新的服务用格式。图书馆就可以避免进行重复的数据采购。

(2) 服务用格式是为声像信息在现有计算机存储设备和网络通讯带宽条件下为降低存储成本并提高传输效率, 通过技术手段对无压缩格式信息加工后产生的信息。这些信息的存放格式, 应是社会上普遍使用的网络浏览器支持的格式。

1.3 搜索平台

声像信息资源的搜索平台是由检索系统、数据加工系统和相应的支持环境组成。

检索系统、数据加工系统包括:

(1) 用户界面

用户界面分为两类: 信息使用者用户界面和数据加工者用户界面。

信息使用者的用户界面为浏览器/服务器体系结构, 浏览器端用户使用的软件为在 INTERNET 网络上广为使用的浏览器软件, 如: IE4.0 及以上的版本、NETSCAPE4.0 及以上的版本。

数据加工用户界面为客户机/服务器体系结构, 客户端使用图书馆集成系统中的编目模块, 我馆使用的是北京丹诚软件公司的 DATATRANS—1000 图书馆集成系统的联机编目系统, 进行声像信息资料的元数据加工。

(2) 服务端管理系统

服务端管理系统根据声像信息资源服务功能的需要, 分别采用浏览器/服务器、客户机/服务器这两种体系结构。除提供数据库检索服务外, 还提供 WWW 页面信息服务、电子公告服务等。

(3) 检索系统

数据库检索系统以 ISO2709 格式存放的各种形态的声音、录像资料、其它电子出版物和已经转换为数字化文件并存储在计算机系统内的多媒体信息的描述信息, 其检索途径根据

声像信息的特征设置为题名(包括丛书名、副题名、其他题名)、责任者、分类号、主题词、载体类型、登录号等。

支持环境包括:计算机设备、计算机操作系统、其它辅助设备以及应用软件。

1.4 数字仓库

数字仓库包括存储数字对象资源的设备和数字对象资源加工系统。数字仓库的硬件可以是普通计算机系统设备,也可以是网络硬盘服务器类的专用存储设备。对象数据加工系统的实际内容包括各种页面制作工具、图像制作工具、多媒体信息的采集加工程序、各种文字处理软件和必要的数据备份工具等。

2 图书馆声像信息资源 Web 检索和播放的实现方法

(1) 利用图书馆专业化加工处理数据的优势,建立声像资料的元数据库,是实现声像资料目录检索的第一步。当对声像信息资源进行分类和特征描述后,形成共享的元数据库,信息用户在这元数据库中检索所使用的时间和检索出信息的准确率与使用简单的全文检索相比将获得极大的改善。

我馆声像信息元数据的形式是利用 CNMARC 格式。

(2) 利用模/数转换技术,通过对音频、视频、图像、文字等媒体进行数字转换、编辑、压缩等技术处理,把馆藏磁带、光盘等声像资料以及对应的一些文字资料转换成计算机可识别的数字化对象数据格式。存储在网络服务器的硬盘上。

以下表格列举了所采用的对象数据格式:

数据类型	原始备份格式	服务用格式
录像及其它连续播放的影像	AVI	MPEG、RM
图像	BMP、TIF	JPG、GIF
声音	WAV	MP3、MID
文字	TXT、HTML、PDF	TXT、HTML、PDF

(3) 利用超文本链接技术,实现声像资料的目录元数据与音频、视频、图像、文字等媒体数字化资料的链接。网络用户通过浏览器,浏览检索输出的结果,通过对声像信息的描述信息,确认自己感兴趣的信息,点击我们在描述信息上的链接,实现声像的播放,播放链接在必要时可以形成下一级详细目录页面,以供读者选择具体播放哪一段或哪一课。

声像信息资源 Web 检索与播放是图书馆信息服务中,技术含量较高的新形式的信息服务。由于计算机技术、网络技术与多媒体技术的飞速发展,图书馆元数据使用标准与格式、声像资料的数字化处理、网络传输以及声像的播放等技术方案也将随之重新调整。

参考文献:

- [1] 邗纛枢,等. 声像资料网上阅览的实现方法[J]. 图书情报知识. 2000, (4).
- [2] 朱岩. Dublin Cord与XML与MARC. (<http://www.unihan.com.cn/digital/js-xml.htm>)
- [3] 北京丹诚软件有限责任公司. 数字化图书馆项目建设建议书(<http://www.datatrans.com.cn>).

Sound & Image Information Resources' Web Search and Play in Library

MENG Shan, ZHANG Yan, HUANG Min

(Library of Hebei Teachers' University, Shijiazhuang 050016, China)

Abstract: The paper puts forward four chief elements and concrete ways to achieve rapid searching and transmitting of sound and image information resources, and hopes that this can lead to further studying of the topic.

Key words: sound & image information resource; metadata; object data; search flat roof; Web searching

[作者简介] 孟珊,女,馆员,现在河北师范大学图书馆期刊部工作,已发文多篇。

[收稿日期] 2002-03-05

(上接第99页)

How Library Better Serve for Scientific Research & Production

ZHENG Tian-ling

(Sun Yet-san Library of Guangdong Province, Guangzhou 510110, China)

Abstract: The paper discusses the importance of library in scientific research, and how library plays a better role in providing information resources for scientific research.

Key words: library; scientific research; information service

[作者简介] 郑田菱,女,1986年毕业于华南理工大学化学工程系,现在广东省中山图书馆外文部工作,馆员。

[收稿日期] 2002-04-23