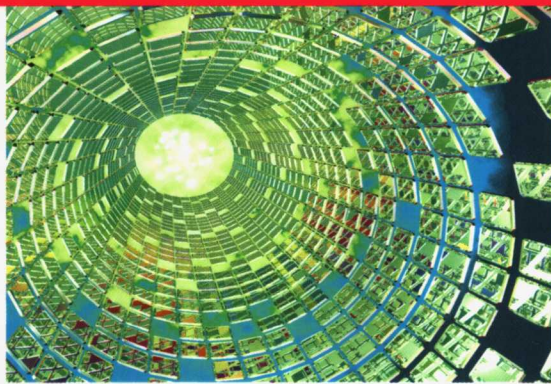


数字时代的图书馆丛书

总编 胡越

数字时代的

图书馆网络信息系统



李广建 主编



北京图书馆出版社

数字时代的图书馆网络信息系统

李广建 主编

李广建 汪语宇 张丽 编著

图书在版编目(CIP)数据

数字时代的图书馆网络信息系统/李广建主编. —北京:北京图书馆出版社,2006. 6
(数字时代的图书馆丛书)

ISBN 7-5013-3138-3

I. 数… II. 李… III. 计算机网络—信息系统—应用—图书馆工作 IV. G250.7
中国版本图书馆 CIP 数据核字(2006)第 048309 号

书名 数字时代的图书馆网络信息系统

著者 李广建 主编

出版 北京图书馆出版社 (100034 北京西城区文津街7号)

发行 010-66139745 66151313 66175620 66126153

66174391(传真) 66126156(门市部)

E-mail cbs@nlc.gov.cn(投稿) btsfxb@nlc.gov.cn(邮购)

Website www.nlcpress.com

经销 新华书店

印刷 北京华正印刷有限公司

开本 787×1092 毫米 1/16

印张 23.5

版次 2006年6月第1版 2006年6月第1次印刷

字数 540(千字)

书号 ISBN 7-5013-3138-3/G·659

定价 50.00 元

前 言

目前,我国图书馆事业正处于一个蓬勃发展的时期,为配合我国政治、经济、文化发展的需要,许多图书馆都在建设新的馆舍。另外,还有大量的图书馆虽没有新建馆舍,但也在进行旧馆改造、更新设备或者是进行扩建,以期提高服务能力,满足广大用户日益增长的需求。

随着计算机技术和网络技术在图书馆中的广泛应用,图书馆新建、扩建或改造所面临的一项重要问题就是如何建设图书馆网络信息系统,形成网络化、数字化的信息资源服务体系,图书馆网络信息系统的建设正在成为我国当前图书馆事业建设的一个热点。

本书试图较全面地概述图书馆网络信息系统建设过程中所涉及的有关知识和问题,共分十一章。第一章为绪论,论述了图书馆网络信息系统概念、内容、特征和技术体系,图书馆网络信息系统的建设原则和建设方式,概要介绍了国内外数字图书馆的发展。第二章至第四章为硬件建设,涉及图书馆网络信息系统建设中的网络技术、计算机技术和存储技术。第五章至第九章为应用系统建设,涉及图书馆自动化系统建设、信息资源建设、图书馆服务系统建设、管理信息系统建设以及图书馆网络信息系统的互操作。第十章论述了与购买及引进图书馆软硬件系统有关的内容。第十一章阐述了图书馆如何开发图书馆网络信息系统。

第一章、第七章和第九章由李广建执笔,第二章、第四章、第八章和第十章由张丽执笔,第三章、第五章、第六章和第十一章由汪语宇执笔,全书由李广建统稿。

朱强老师审阅了全书,并提出了宝贵的修改意见,使本书内容得以丰富。张树华老师、胡越老师、陈源蒸老师、人天书店邹进同志在本书的写作过程中给予了大量的支持和帮助。在此对他们深表谢意。同时我们还要感谢本书具名和未具名的参考文献的作者。

李广建

2005年夏于北京师范大学

目 录

前言	(1)
第一章 绪论	(1)
第一节 图书馆网络信息系统	(1)
第二节 图书馆网络信息系统的建设	(9)
第三节 国内外图书馆网络信息系统建设概况	(12)
第二章 图书馆网络系统	(25)
第一节 网络技术基础	(25)
第二节 图书馆网络系统设计	(37)
第三节 国内外主要网络厂商和产品	(59)
第三章 图书馆计算机系统	(63)
第一节 计算机技术基础	(63)
第二节 图书馆计算机系统建设	(75)
第三节 国内外主要计算机厂商及产品	(81)
第四章 图书馆存储系统	(86)
第一节 数据库系统	(86)
第二节 存储系统	(97)
第三节 国内外主要数据库系统和存储系统厂商及产品	(106)
第五章 图书馆自动化系统	(114)
第一节 图书馆自动化系统概述	(114)
第二节 图书馆自动化系统的功能	(119)
第三节 图书馆自动化系统的发展趋势	(124)
第四节 国内外主要自动化厂商和产品	(127)
第六章 数字资源建设	(143)
第一节 数字资源的类型	(143)
第二节 数字资源的建设	(151)
第三节 数字资源建设的标准规范与资源共享	(163)
第七章 数字资源服务系统	(173)
第一节 图书馆网站	(173)
第二节 多媒体导航系统	(180)
第三节 信息检索服务	(184)
第四节 跨库集成检索服务	(193)
第五节 数字参考咨询服务	(199)
第六节 个性化信息服务	(204)

第七节 馆际互借与原文传递服务	(210)
第八章 图书馆管理信息系统	(217)
第一节 图书馆一卡通系统	(217)
第二节 图书馆门禁系统	(220)
第三节 电子阅览室系统	(223)
第四节 图书馆办公自动化系统	(227)
第五节 图书馆知识管理	(233)
第九章 图书馆网络信息系统的互操作	(244)
第一节 图书馆网络信息系统的互操作及其模式	(244)
第二节 XML 基础	(246)
第三节 Z39.50	(257)
第四节 OAI - PMH	(263)
第五节 LDAP	(272)
第六节 STARTS/SDLIP/SDARTS 协议	(278)
第七节 OpenURL	(283)
第八节 RSS	(293)
第九节 OAIS	(297)
第十节 Web Service	(303)
第十一节 图书馆门户	(311)
第十章 图书馆网络信息系统的选型与管理	(321)
第一节 图书馆网络信息系统的选型的步骤	(321)
第三节 图书馆网络信息系统项目招标与签订合同	(326)
第四节 图书馆网络信息系统选型中的注意事项	(336)
第十一章 图书馆网络信息系统的开发	(339)
第一节 图书馆网络信息系统分析	(339)
第二节 图书馆网络信息系统设计	(347)
第三节 图书馆网络信息系统的实现与测试	(352)
第四节 图书馆网络信息系统的维护	(356)
第五节 图书馆网络信息系统的项目管理	(358)
第六节 图书馆网络信息系统开发过程中应该注意的问题	(361)

第一章 绪 论

图书馆是收集、整理、存储和传递文献信息,为一定社会政治、经济服务的科学、文化和教育机构。随着现代社会信息化程度提高和信息技术的广泛应用,为了充分满足社会多方位、多层次、多门类的信息需求,图书馆正在从传统的手工式的信息服务模式,向开放性、自动化的图书馆服务模式转变。目前,数字技术和网络技术日趋成熟,为图书馆网络信息系统的建设奠定了技术基础,使图书馆真正实现信息资源共享、快速有效地提供各种层次的信息服务、满足各式各样的信息需求,促进了图书馆信息服务的社会化。

第一节 图书馆网络信息系统

一、图书馆网络信息系统概念与内容

要理解图书馆网络信息系统的含义,应首先了解什么是网络信息系统。网络信息系统是一种通过网络存取复杂数据并进行互操作的信息系统,这种系统能够处理各种异构、动态、分布式的数据,并能为分散在不同位置的大量用户服务。

图书馆网络信息系统实际上是网络信息系统在图书情报领域的应用,它应用先进的信息技术,收集、加工、存储数字化的图书、期刊、图像、音频、视频资料等信息资源,借助计算机网络实现用户对所有资源的访问,提供各种信息服务。在这种环境下,用户使用图书馆各种资源的地理界限已经被打破,所有人都能够通过计算机网络在任何时间、任何地点享用图书馆的各种资源和服务,并且不必关心这些资源具体物理位置。简单地说,图书馆网络信息系统是一种基于计算机网络的数字化信息资源管理和服务系统。

从功能角度看,一套完整的图书馆网络信息系统包括资源建设、资源服务、自动化管理三个部分,如图1-1所示。

(一)资源建设

资源建设是图书馆服务的基础,图书馆网络信息系统通过馆藏资源数字化、网络信息采集、第三方资源订购等方式形成数字化馆藏,对数字资源进行有序化管理和资源的整合,同时,建立完整的书目系统供用户查阅。

(二)资源服务

完备的资源服务体系是图书馆网络信息系统的重要组成部分,它将各种图书馆资源整合到一起,消除资源孤岛,屏蔽信息异构差异,为用户提供统一、方便、快捷的信息资源服务。信息资源服务包括自动化书刊借阅服务、各种信息检索服务、数字化参考咨询服务、门户发布和管理、个性化信息服务等。

(三)自动化管理

图书馆网络信息系统能够协助图书馆实现内部办公和外部服务管理的自动化,减轻工作

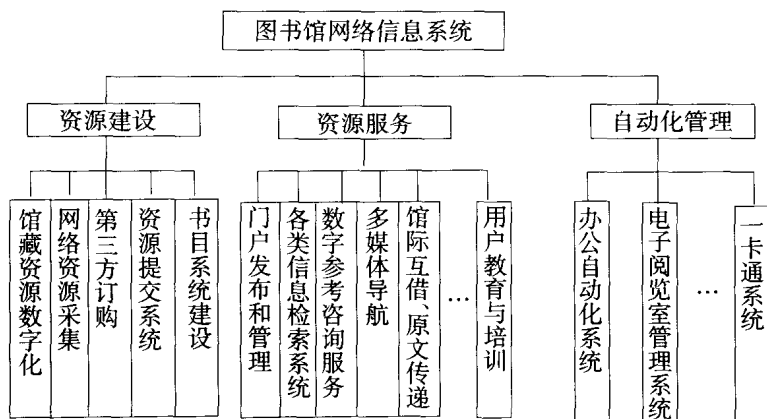


图 1-1 图书馆网络信息系统结构

人员的工作负担,为工作人员提供信息交流平台,加强图书馆内部人员的交流与学习,提高服务工作效率。图书馆自动化管理包括办公自动化系统、一卡通系统、电子阅览室管理系统、资源提交系统等。

为实现这三部分的功能,除硬件以外,图书馆网络信息系统一般包括数字资源采集与加工系统、图书馆业务集成管理系统、信息检索服务系统、馆际文献服务系统、Web 信息服务系统、信息导航系统、管理系统等几个子系统,这些子系统共同构成了图书馆网络信息系统,如图 1-2 所示。

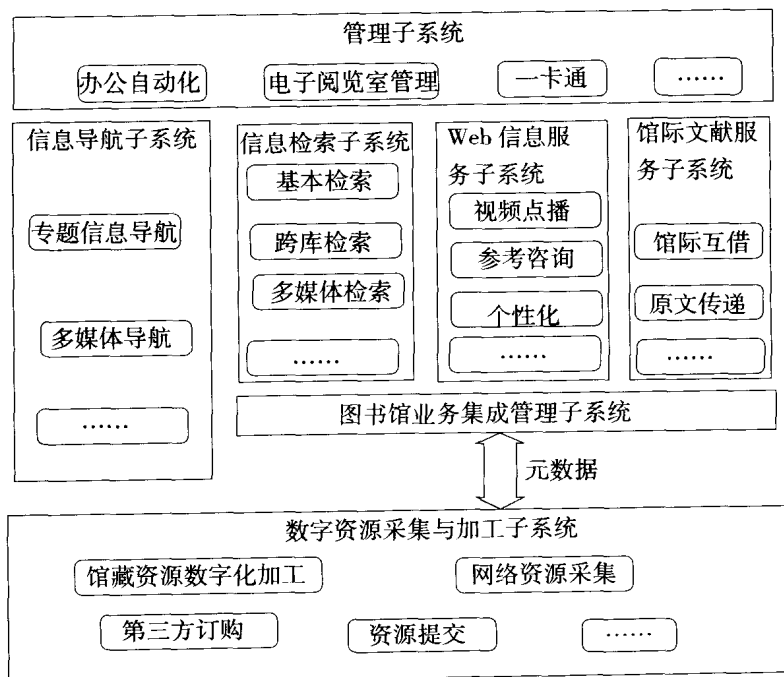


图 1-2 图书馆网络信息系统功能体系结构

(一)数字资源采集与加工子系统

数字资源采集与加工子系统是实现图书馆数字资源建设的重要子系统。图书馆网络信息系统的数字资源包括电子图书、电子期刊、网络数据库、光盘、图像资源、音频资源和视频资源以及其他电子资源。数字资源采集与加工子系统主要通过馆藏资源数字化加工和管理、自建数据库、购买第三方数字资源、网络资源采集、接收所属机构内外资源提交等方式来构建自己的数字资源体系。

(二) 图书馆业务集成管理子系统

图书馆业务集成管理子系统即图书馆自动化系统,它以数字资源采集与加工子系统为基础,通过元数据与数字资源进行交互,同时,支持上层服务(信息导航、信息检索服务、Web 信息服务、馆际文献服务)的实现。图书馆业务集成管理子系统一般由采访模块、编目模块、流通模块、期刊管理模块、公共查询模块等构成,以书目数据为中心、涵盖现代图书馆各个业务管理环节,包括采访、编目、典藏、流通、阅览、期刊管理等,能实现多文种、多类型的实体文献著录处理,包括图书、连续出版物、视听资料、音乐文献、地图及其他混合型资料的著录处理,并对广域网环境下的联合采访、联合编目、馆际互借、原文传递等扩展功能提供支持。同时,图书馆业务集成管理子系统还可将网络资源信息作为虚拟实体进行著录管理。同时,图书馆业务集成管理子系统的各业务环节之间有着广泛的信息交换和共享,能形成完整的业务统计信息。

(三) 信息检索服务子系统

信息检索服务子系统是图书馆网络信息系统所提供的信息服务的一个重要组成部分,信息检索服务也是用户服务中的主要内容。信息检索服务系统包括文献信息检索、连续出版物信息检索、视听资料信息检索、网络信息检索、全文检索、跨库检索、个性化检索以及与检索相关的服务,包括读者借阅查询、资源导读等。信息检索服务有 Web 方式和非 Web 方式的联机服务方式。

(四) Web 信息服务子系统

Web 信息服务系统是指图书馆网络信息系统通过网络,以数字资源为基础,对用户提供的各种服务,包括数字化参考咨询服务、视频点播(VOD)服务、Email 服务、FTP 服务、网上读书服务、电子教室、远程教育、视频会议、个性化信息服务等等。虽然不同的图书馆网络信息系统由于其自身的网络速度、带宽以及服务重点等方面的不同,其 Web 信息服务子系统所包含的信息服务的类型有所不同,但用户不用到馆即可使用这些 Web 信息服务,这是传统图书馆所不具备的。

(五) 馆际文献服务子系统

馆际文献服务子系统支持向用户提供馆际互借、原文传递等服务,目的在于提高信息资源的共享性。用户利用馆际文献服务系统,从馆际互借协作单位借入文献时,系统按本馆流通政策将文献出借给最先提出请求的读者,同时馆际文献服务系统也可按照馆内用户借阅文献的操作作为馆际互借单位或个人向外单位或个人借出文献。用户还可通过馆际文献服务系统以电子邮件或邮寄等方式获取检索到的文献原文,包括期刊论文、专利文献、学位论文、会议文献等。

(六) 信息导航子系统

信息导航子系统的功能是引导用户了解并使用图书馆及其网络信息系统或者获得专题信息,信息导航子系统包括专题门户网站、多媒体导航系统等。图书馆专题门户网站将与某一专题有关的所有信息资源与服务功能通过主页的形式展现出来并为用户服务。多媒体导航系统

一般以多媒体为表现形式,采用触摸屏方式,内容由声音、图像、视频构成,有些多媒体导航系统还使用了三维动画或虚拟现实技术。读者通过点击屏幕即可了解图书馆的建筑、资源、服务,为读者更好地利用图书馆及其网络信息系统起到了直观的引导作用。

(七)管理子系统

管理子系统实现对整个图书馆的自动化管理功能,对馆内工作业务的管理和对图书馆网络信息系统进行管理,主要包括办公自动化系统(电子馆务)、电子阅览室管理系统、一卡通等。

二、图书馆网络信息系统的特征

(一)信息资源数字化

图书馆网络信息系统是基于计算机的信息系统,图书馆中的各种信息资源,包括文字信息、图形图像信息以及声音、视频、动画等信息,都必须转化为数字形式,才能为图书馆网络系统所处理。信息资源数字化不仅是图书馆网络信息系统运行的基础,也是图书馆网络信息系统的特征之一。信息资源的数字化包括两个方面,一是馆藏文献内容的数字化,一是已数字化资源的加工处理。

馆藏文献内容的数字化是指采用数字技术,将馆藏文献,包括印刷型文献、缩微制品和视听资料的内容进行数字转换和处理,使之能为图书馆网络信息系统所管理并提供给用户使用。

已经数字化的全文信息包括各种电子出版物(电子期刊、电子书、电子报纸等)和网上下载的数字资源。对于电子出版物来说,需要进行格式转换,对数据元对象、内容、章节、目次进行标识,对内容进行分类和主题标引,并利用超链接技术将它们和正文连接起来,使用户能够利用图书馆网络信息系统对电子出版物进行检索和利用。对于网上下载的数字资源,应保证其格式与图书馆网络信息系统资源的格式一致,必要时也需要进行格式转换和作相应的处理。

(二)信息资源共享化

图书馆网络信息系统是一个基于网络信息系统,任何用户都可以通过网络获取系统所拥有的大量信息资源,因此,信息资源的共享化是图书馆网络信息系统的另一个基本特征。当然,共享化是相对的,由于条件和环境等诸多因素的制约,例如出于知识产权、信息安全的考虑,信息资源共享化是有一定条件的,是在一定范围内的共享化。

(三)业务管理自动化

图书馆网络信息系统可以利用计算机技术、通信技术、网络技术对各项图书馆业务工作实行自动化的管理,包括对各种信息资源的数字化存储、加工、处理,对数字化信息的标识、描述和组织,对图书馆工作人员的任务分配和考核、对用户进行管理等等。图书馆网络信息系统的这一特征决定了它贯穿于图书馆业务工作的始终,极大地提高了图书馆工作的效率,增强了图书馆的服务能力。

(四)信息服务虚拟化

网络和通讯技术使得用户可以在任何时间、任何地点接受图书馆网络信息系统提供的服务,而不需要亲自到图书馆,这是图书馆网络信息系统服务与传统图书馆服务的根本差别,这种虚拟化的信息服务,大大缩短了信息传递的时间以及信息提供者和信息使用者之间的距离,加快了信息交流和反馈的速度,使得图书馆的价值得到了充分发挥。

三、图书馆网络信息系统的技术体系

(一) 图书馆网络信息系统技术体系的基本特征

概括地讲,图书馆网络信息系统涉及信息采集(扫描、缩微、OCR 技术、语音识别、人工智能等)、信息表现(多文种平台、图像压缩、多媒体技术等)、信息组织(元数据、MARC 数据、XML 等)、信息检索(全文检索技术、跨库检索等)、信息传递(WEB 技术、发布、PDF 格式等)、信息安全(加密、签名、水印等)等技术领域,是一个综合性的领域,Sun 公司 2002 年 8 月 5 日发布的技术白皮书《Digital Library Technology Trends》认为,用于构建图书馆网络信息系统的技术具有以下几个基本特征:

1. 服务驱动(Service Driven)

图书馆网络信息系统的技术是由图书馆所提供的信息服务以及开展这些信息服务所需要的工具所决定的,换句话说,只要是图书馆服务需要的技术,都属于图书馆网络信息系统的技术。

2. 开放结构(Open Architecture)

图书馆网络信息系统的技术必须是开放的、可扩展的,并且支持异构的分布式系统之间的互操作。

3. 可伸缩性(Scalability)

图书馆网络信息系统的技术必须是健壮的、可伸缩的和可靠的,能够为各种环境下的各种背景和有不同信息需求的大量用户方便地使用。

4. 持久性(Persistence)

图书馆网络信息系统的技术必须能够保证持久地访问图书馆的资源。一方面,这些技术必须有一定的稳定性,能够保证一年 365 天不间断的服务。另一方面,技术持久性的另一个含义是应能够保证系统的更新换代不会影响对原有信息资源的访问。

5. 私密性(Privacy)

图书馆网络信息系统的技术必须对私密性敏感,不仅能控制对资源的访问和获取权限,并且还能支持对资源的匿名和定制访问,以确保资源的知识产权不受侵犯,用户的隐私不被泄露。

6. 实用性(Practicality)

图书馆网络信息系统的技术应该是成熟稳定的技术,符合图书馆工作的需要,能促进图书馆工作的发展。

7. 模块化(Modularity)

各种技术都应以模块化的方式加入图书馆网络信息系统中来,新技术与原有技术能方便地结合到一起,任何一种技术的升级不仅不应对其他技术利用产生影响,而且能与原有技术进行互操作,同时不降低整个系统的性能。

8. 时限性(Timeframe)

图书馆网络信息系统的技术应支持系统迁移,换代时间不低于 3 至 5 年。

9. 客户支持(Client Support)

图书馆网络信息系统的技术应支持通用的计算机配置和软件,只有在特殊情况下才需要专用的客户端。

10. 基础设施(Infrastructure)

图书馆网络信息系统的技术应符合通用、流行的技术平台和相关的技术标准(如网络、网络协议、身份验证等技术)。

(二)图书馆网络信息系统的技术体系

图书馆网络信息系统利用先进的计算机设备和技术,通过计算机网络,向用户提供准确、全面的数字化信息资源,并提供相关的信息服务,其技术体系如图 1-3 所示。

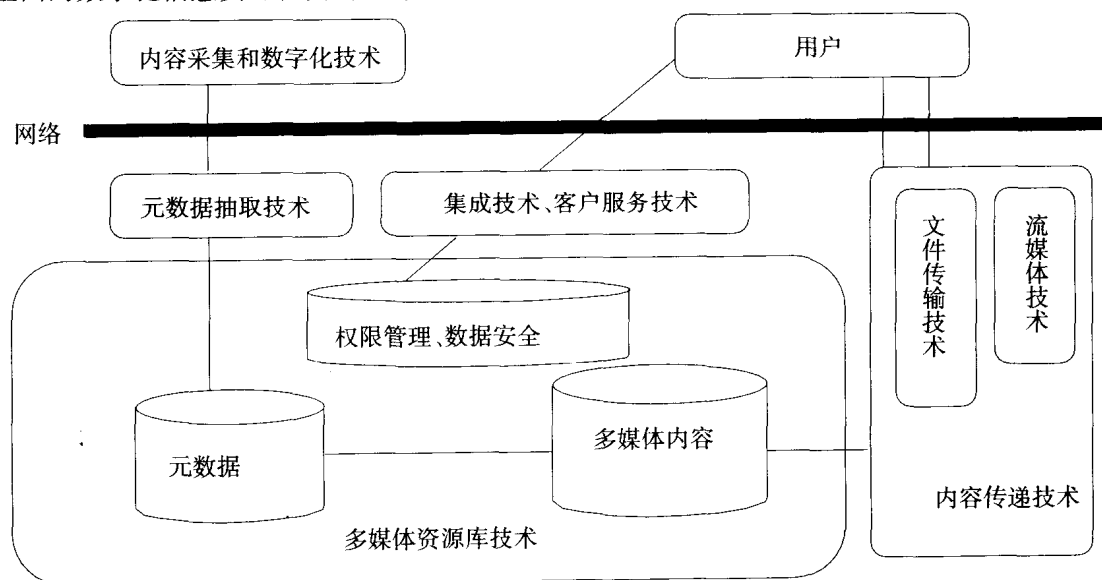


图 1-3 图书馆网络信息系统的技术体系

从图 1-3 中可以看出,图书馆网络信息系统的技术体系包含了以下几个方面:

1. 内容采集和数字化技术

内容采集和数字化技术,是指获取信息资源,并将内容从物理形式转换成数字形式技术。

目前图书馆采用购买和获取因特网资源两种方式进行内容采集。因特网信息资源是对传统图书馆资源的一种有力支持和扩充,利用搜索引擎和信息采集系统,实现图书馆的因特网资源的自动获取,将大大提高图书馆资源的服务水平。

信息资源数字化只需创建一次内容,即可在不同环境下进行管理和利用,大大提高了信息的可用性。数字化技术包括对文本信息、图形图像信息、音频信息和视频信息的数字化。文本数字化技术有人工键盘录入技术、手写板技术、扫描技术、纠斜技术、清理技术、版面分析技术、OCR 技术、自动标引技术、压缩技术等。图形处理技术包括 CAD 技术等。图像采集生成技术包括 PIC 技术、GIF 技术、JPEG 技术、数码相机摄影技术、扫描技术、图像软件获取技术、图像软件捕捉技术等。音频信息数字化技术主要包括 CD 技术、MP3 技术、RM 技术、WAV 技术以及声卡捕捉技术、提取光盘上的声音文件等技术。视频信息数字化技术包括 MPEG 技术、AVI 技术、VCD 技术和 DVD 技术等。

2. 元数据抽取技术

元数据抽取从数字化的信息资源中抽取元数据,或根据数字化资源的内容创建其元数据或索引信息。目的是便于管理、保存、搜索和发现这些信息资源。元数据抽取技术主要有基于

元数据复用的元数据抽取技术、基于 XML DTD 的元数据抽取技术、基于词语分析的元数据抽取技术和基于自动数据赋值的元数据抽取技术。

3. 多媒体资源库技术

多媒体资源库用于存储数字化内容及其元数据。多媒体资源库技术涉及到数据压缩、数据库、数据仓库等相关技术。

数据压缩是通过数学运算将原来较大的文件变为较小文件,通常可分为无损压缩和有损压缩两种类型。无损压缩是利用数据的统计冗余进行压缩,压缩后的数据经过重构还原后与原始数据完全相同,而没有任何失真。有损压缩是利用了人类视觉对图像中的某些频率成分不敏感的特性,允许压缩过程中损失一定的信息。虽然压缩后的数据经过重构还原后与原始数据有所不同,但是所损失的部分对理解原始图像的影响较小或基本没有影响。

数据库技术是数据存储的重要技术之一。数据库管理系统的种类很多,图书馆网络信息系统中常用的有关系型数据库管理系统和面向对象型数据库管理系统。关系型数据库管理系统允许多个用户同时访问同一信息,当一个用户对此信息改动时另一个用户不会受到影响,同时可以对数据进行及时的备份,保障了数据的安全性。而面向对象型数据库管理信息系统与分布式数据库相结合,更适合应用于海量多媒体资源的存储。

数据仓库是面向主题的、集成的、在一定周期内保持稳定的、用于支持决策分析的数据集合。它具有面向主题、集成和不可更新等特点。数据仓库技术包括联机分析处理技术和数据挖掘技术等。

4. 集成技术

集成是指将异构的、分布的和自治的资源或系统有效地组织成一个有机的整体,为用户提供一个统一的信息利用环境,以方便信息利用,提高信息资源与信息系统的利用效率和服务效率。集成技术包括数字资源集成技术和系统集成技术。

数字资源集成的方式分为物理集成和虚拟集成。物理集成是通过建立一个数据仓库,将参加集成的各类信息资源按主题或多维方式建模,集中存储在数据仓库中,然后为这个数据仓库建立检索、查询等操作机制,提供给用户一种新的数据模式。虚拟集成是建立用户与信息资源之间的一个集成接口,数据仍然保存在各自的信息源上,集成系统通过一个虚拟的集成视图及其查询处理机制,自动将集成模式的查询请求转换成各个分布式信息源的查询。数字资源集成涉及的技术有虚拟数据库技术、XML 技术、数据仓库技术、智能代理技术、信息网格技术等。

系统集成不仅包括不同图书馆网络信息系统之间的集成,还包括图书馆网络信息系统与现有图书馆自动化系统之间的集成。系统集成涉及的技术有 Web Services 技术、中间件技术和基于协议的整合技术。

Web Services 是一种新型的自适应、自我描述、模块化的 Web 应用程序。这些应用程序可以跨越 Web 进行发布、定位和调用。Web Services 通过服务提供者(Service Provider)提供服务及维护注册表以使服务可用,服务代理充当服务提供者与服务请求者之间的中介,为服务提供交换场所,服务请求者与服务代理一起发现 Web 服务,然后调用这些服务以创建应用程序。

中间件技术是这样一种技术,它为系统平台与应用程序之间或应用程序与应用程序之间提供了标准接口,从而使系统平台与应用程序之间、应用程序与应用程序之间进行互操作。中间件分为数据库中间件、远程过程调用中间件、面向消息中间件、基于对象请求代理和事务处

理中间件。

基于协议的整合技术主要指各种文档描述和信息检索协议标准,包括 SGML/XML、Z39.50、HTTP/CGI、RDF、OAI、SOAP 等。

5. 客户服务技术

客户服务技术是指针对用户提供各种信息服务的技术,包括信息检索技术、个性化服务技术、信息交互技术、电子商务技术等。

信息检索是图书馆信息服务的主要内容,信息检索技术是图书馆网络信息系统的核心技术,图书馆通过有效的信息组织提供了更高效率、更为准确和更多途径的信息检索功能。信息检索技术主要有:(1)文本信息检索技术,主要包括文本检索技术、搜索引擎技术、信息过滤技术、跨语言检索技术等;(2)基于内容的多媒体检索技术,它根据媒体和媒体对象的内容语义、上下文联系,从数据库中找到具有指定特征或者含有特定内容的相似媒体对象,是一种近似匹配的技术,主要包括图像检索技术、动态视频检索技术以及音频检索技术等;(3)智能检索技术,包括自然语言检索技术、概念检索技术等。自然语言检索是指用户以自然语言作为提问式进行检索;概念检索是为用户提供与用户检索相关的词汇,从而可以起到明确用户检索需求的目的。

个性化服务是图书馆“以用户为中心”服务宗旨的体现,从根本上说,就是尊重用户,研究用户的行为和习惯,为用户选择最符合其需求的资源,提供特色服务。服务内容包括 Alerting 服务(最新资料快报服务)、Active Learning(主动学习)、数据移动服务及个性化定制等,涉及的技术有个性化定制技术、用户建模技术、信息检索与过滤技术、信息推送技术、机器学习与相关反馈技术等。

信息交互服务是指利用各种用户界面技术和可视化检索技术来方便、直观、快捷地为用户提供各种信息服务,以提高图书馆网络信息系统的易用性。信息交互服务涉及的技术有用户界面技术和检索可视化技术。

用户界面技术是指用户与计算机通过互相传递信息进行交互的操作方式。用户界面技术经历了字符用户界面(CUI)、图形用户界面(GUI)、社会用户界面(SUI)几个发展阶段。CUI 要求用户逐个输入字符来执行指令,操作繁杂,指令多而不易记忆,使用困难,执行结果也通过字符显示,单调且难以读懂。GUI 通过“桌面”显示计算机所有资源,用“图标”表示其中的每一个具体对象,清晰易懂,加上鼠标的配合,极大地简化了一般操作。SUI 成为一种新型的用户界面技术,它采用多媒体用户界面,如文字、图标、图形、图像、动画、声音和视频等,并能以 3D 方式显示,使用户有身临其境的感觉,充分调动了用户注意力和实现全方位互动,提高了用户使用图书馆网络信息的效率。

检索可视化是指把文献信息、用户提问、各类检索模型以及利用检索模型进行信息检索的过程中不可见的内部语义关系转换成图形,在一个二维或三维的可视化空间中显示出来。检索可视化技术能够以二维或三维的图形高效地输出检索结果,用点或其他符号来表示数据库中的文献,并且整个检索过程是连续的、可逆的,用户可以在任意节点位置处选择前进或后退。通过检索可视化技术,用户可以直接观察到检索过程,并与系统进行有效的交互操作。可视化技术涉及一般可视化技术、面向文献内容的可视化技术、面向超文本链接分析的可视化技术、面向网站拓扑结构的可视化技术、面向日志信息的可视化技术和面向信息检索结果集合分析的可视化技术等。

电子商务是图书馆网络信息系统的重要功能,同时也是客户信息服务的一项重要内容。图书馆网络信息系统的电子商务平台连接图书馆、出版社、作者、商家、读者和银行机构,通过安全认证进行版权管理,控制对资源的访问,同时处理计费收费等相关业务。

6. 内容传递技术

内容传递负责传递各种信息内容,主要通过文件传输或流媒体两种方式来完成。文件传输通过 FTP、HTTP、电子邮件、光盘发行、频道推送等形式来传输信息内容,是最常用的传递方式。流媒体是一种使用流技术在网络上传送的多媒体文件。流技术可以把连续的影像和声音信息经过压缩处理后放在网站服务器上,让用户边下载边观看、收听,而无需等整个压缩文件下载到用户机器上以后才能观看的网络传输技术。流媒体在图书馆网络信息系统的内容传递中发挥着越来越重要的作用。一个完整的流媒体平台包括流服务应用软件、集中分布式视频系统、视频业务管理媒体发布系统、视频采集制作系统、媒体内容自动索引检索系统、媒体数字版权加密系统、媒体存储系统和客户端系统。流媒体技术中的传输协议有实时传输协议 RTP 与 RTCP、实时流协议 RTSP、资源预订协议 RSVP 等。

7. 数字资源安全技术

数字资源的安全包括三个方面:认证技术、权限控制和权益管理。

认证技术包括身份认证和信息认证。身份认证是网络服务双方确认对方身份真实性的过程,是提供安全信息服务的前提。信息认证是指对信息实施加密传输与验证,以保证信息的真实性、完整性。主要的认证技术包括加密技术、消息摘要、数字签名、数字信封和认证机制。

权限控制是指允许被授权的用户(单个用户或团体组织)对某些资源的访问,同时拒绝向非授权的用户提供服务。权限控制涉及很多模型,如自主性访问控制(Discretionary Access Control, DAC)、强制性访问控制(Mandatory Access Control, MAC),这些模型适用于不同的应用程序和应用环境,他们通常都是基于采用用户—用户组的安全模型。在多数图书馆网络信息系统中,通常采用基于角色的访问控制技术(Role - Based Access Control, RBAC),这种技术具有简化管理的优点,能够有效地增强对用户访问的控制和管理。RBAC 模型利用用户—角色、角色—权限和角色—角色之间的关系来达到权限控制的目的。

权益管理即版权保护技术,主要解决信息资源的合理使用问题。权益管理涉及的技术有安全容器技术、数字水印技术、移动 Agent 技术、加密技术、CA 认证技术等。

第二节 图书馆网络信息系统的建设

图书馆网络信息系统的建设是一项非常复杂的工作,必须有明确的目标、合理的原则和科学的方法。

一、图书馆网络信息系统建设目标及原则

(一)图书馆网络信息系统的建设目标

随着社会信息化的发展,图书馆在社会信息服务业中扮演着越来越重要的角色。因此,图书馆必须完善信息服务基础设施,建设集网络环境、文献资源、系统管理、信息服务为一体的综合网络信息系统。具体地说,图书馆网络信息系统的建设目标是:

1. 建立完善的图书馆信息网络

应充分采用先进、成熟的网络支撑和信息处理技术,构建整体有序、可靠实用、信息高度共享和结构开放的信息网络。图书馆网络信息系统应具有足够的扩展性,以满足图书馆资料采访、编目、检索、流通等工作自动化和全面数字资源服务功能的要求。

2. 建设完备的数字资源体系

完备的数字资源体系是图书馆网络信息系统的重要组成部分,是发挥图书馆网络信息系统功能的前提,有了完备的数字资源,才能借助网络环境和高性能计算机实现信息资源的有效利用和共享。

3. 提供全方位的信息服务

图书馆网络信息系统应充分依托丰富的文献、信息资源、完善的功能设施和现代化的管理,通过网络及各种应用系统,最大限度满足社会的多样性信息需求。

(二)图书馆网络信息系统的建设原则

高起点、高性能、重应用是对图书馆网络信息系统的基本定位,在图书馆网络信息系统建设中,普遍遵循以下原则:

1. 先进性原则

近几年图书馆网络信息系统的建设经验充分表明,先进的技术意味着目前更高的性能和将来更低的维护费用。在综合考虑投资、风险、需求等条件情况下,应当尽可能采用先进的技术来开发系统,设计思想、软硬件设备、系统结构均应保证有先进性,以便在较长的时间内不再重新投资更换主要设备及软件。

2. 适用性原则

图书馆网络信息系统的功能和技术性能应与图书馆的职能、规模及目标相符合,能够满足图书馆的业务管理、信息加工和信息服务的需求。

3. 成熟稳定性原则

图书馆网络信息系统在处理图书馆的各项业务中,应当具有较强的稳定性和容错性。要求系统的硬件和软件成熟稳定,确保图书馆各项工作顺利进行。

4. 标准化原则

图书馆网络信息系统所采用的技术应符合通用的国际和国家标准,支持成熟的信息处理标准与规范,为实现图书馆间的信息资源共享奠定基础。

5. 开放性原则

图书馆网络信息系统建设不但需要考虑当前的需求,还需要能够适应将来的发展。系统的硬件和软件平台应具有高度的开放性,支持业界标准与协议,兼容相关的技术,并提供足够的软件硬件接口。

6. 安全性原则

由于图书馆的各项业务是对外开放的,一旦系统出现差错将给用户带来极大的不便,给图书馆造成极坏的影响。因此要求图书馆网络信息系统必须有安全可靠性的保证。

7. 高性价比原则

图书馆网络信息系统软硬件建设在各方面条件相近时,应选择较高系统性能价格比的设备与软件。根据图书馆的规模和功能,选择最适当的资源,以求达到较高的系统性能价格比。

8. 柔性化原则

功能完备的系统,不仅表现在已有功能的完备,更重要的是对各种潜在需求和变化的需求

的支持。图书馆网络信息系统的功能应具有良好的可扩充性,能够根据图书馆发展需要灵活配置和扩展,以支持图书馆的可持续发展战略。

二、图书馆网络信息系统建设方式

建立图书馆网络信息系统,其目的是促进和加强图书馆间的协调与合作,实现信息资源共享,快速有效地提供各种层次的信息服务,促进图书馆信息服务的社会化。现阶段,信息技术特别是计算机技术发展十分迅速,光纤网络技术、局域网技术、以太网技术、卫星通信技术等网络技术都已相当成熟,宽带综合业务数字网(B-ISDN)技术、ATM 异步传输技术已从实验走向实用。数据库技术、客户机/服务器技术等网络应用开发技术已被普遍应用,网络应用系统开发工具也逐步完善,这些均为图书馆网络信息系统的建设提供了技术基础。就目前来看,图书馆网络信息系统建设主要有两种方式,一种是定制方式,另一种是成套方式。

(一)定制方式

定制方式是当前构建图书馆网络信息系统的主要手段,它是指图书馆根据自身需求,提出一整套系统方案,根据软件工程、项目开发的方法,构建图书馆自己的网络信息系统。由于一个完整的图书馆网络信息系统涉及的范围很广,包括门户网站、数字资源采集加工系统、数字资源集成系统、各类信息检索系统、数字化参考咨询系统、个性化服务系统、电子馆务系统等,而且每个系统的需求、功能、服务对象、复杂程度都有很大差别,因此,必须根据图书馆的实际情况,对系统每个部分进行实事求是的分析和定位,做到有针对性地开发和建设。

采用定制方式需要有较强的技术实力和资金实力作支持,由于大部分图书馆通常不具备开发实力,因此,在图书馆网络信息系统建设过程中可以实行开发工作外包,与比较有信誉的软件开发商合作进行系统的研发。

定制开发方式的优点表现在:由于定制方式是在经过充分的调查研究和详细的系统分析设计基础上开展的,因此,开发出的图书馆网络信息系统有较强的针对性和较高的实用性,系统符合用户及图书馆自身业务需求;由于通过定制方式建立的图书馆网络信息系统属于一次性投入,不必随着时间的推移支付许可费和维护费,从长远看,是一种比较经济的实施方式;在系统建立的过程中,系统的开发和建设是可控制和可预测的,因而系统具有较好的可用性和可持续发展性,可随时根据环境的变化和需求的拓展进行系统的不断调整和完善;更为重要的是,定制方式还使得系统能很容易地进行二次开发,从而能满足不断增长的信息需求并保持系统的技术先进性。

定制开发方式的缺点在于:一次性投资较大,对技术水平要求较高,因而较适于资金实力较强、人力资源丰富、软硬件条件较优越的大中型图书馆;另外,标准的变化会导致系统缺乏与其他数字环境的互操作性,影响信息资源的共享,因此标准化问题是以定制方式建立图书馆网络信息系统的一个瓶颈。

(二)成套方式

成套方式是指图书馆通过招投标等方式,直接购买一些计算机公司和图书馆自动化厂商的产品,构建自己的网络信息系统。这些商业性的厂商大多有较雄厚的资金和技术支持,有一套自己的解决方案,有较高的利用价值,如著名的图书馆自动化厂商 Endeavor Information Systems 公司、Innovative 公司、Sirsi 公司、Ex libris 公司、VTLS 公司等都提出了自己的数字图书馆解决方案,每个公司都有自己擅长的图书情报领域,如 Endeavor Information Systems 公司的