

●马建玲 (中国科学院 资源环境科学信息中心, 甘肃 兰州 730000)

基于 TRS 系统构建网上专题阅览室

[关键词] 虚拟阅览室; TRS 全文检索系统

[摘要] 介绍了如何基于 TRS 全文检索系统建设网上专题虚拟阅览室。阐述了虚拟阅览室的结构和功能设计、实现方法以及虚拟阅览室的发展趋势。

[中图分类号] G250.72; G250.73

[文献标识码] A

[文章编号] 1005-8214(2005)06-0121-02

2003年,中国科学院资源环境科学信息中心在引进的 TRS 全文检索系统之上尝试构建了网上专题虚拟阅览室,通过网络为用户提供基于 Web 的专题性的、多资源类型集成的阅览服务,真正实现了阅览服务的虚拟化、网络化,并收到了比较好的服务效果。

1 TRS 全文检索系统概述

TRS 信息发布检索是基于对大量的非结构化的数据进行管理、检索、发布的系统,其功能覆盖到整个信息的生存周期,包括信息的采集、信息发布、信息分类导航和查询、信息服务以及信息传播等。TRS 全文检索服务器由 TRS Server、管理工具和开发接口三部分组成。TRS 全文检索系统包括多个功能模块,虚拟阅览室主要是在 TRS 全文数据库服务器支持下,基于 TRS 的 Web Application Server 4.0 (TRS WAS 4.0),即 TRS 推出的新一代 Web 应用服务器系统开发的,TRS WAS 4.0 利用数据库技术和动态网页生成技术,实现信息的动态发布。

TRS WAS 由两个模块组成:一个模块是系统控制模块,称为 TRS WAS 管理控制台,用于管理和设置要发布的信息;另一个模块是应用模块,用于将数据库中的信息动态发布到 Web 服务器上。在此,TRS 提供两个工具,其中 TRS Admin (TRS 数据库管理工具)负责完成对 TRS 索引和检索服务器的所有管理工作,包括建库、数据维护、安全控制、备份和恢复、优化、词典维护等各项管理任务。TRS WAS 4.0 (TRS 动态发布工具)利用数据库技术和动态网页生成技术,对 Web 服务器进行扩展,或采用第三方的 Web 应用服务器,以实现数据库信息的动态自动发布功能。在虚拟阅览室建设中,WAS 4.0 的作用有两个,一个是 TRS 数据库动态发布,另一个是生成数据库 Web 录入和发布界面。

2 虚拟阅览室结构功能设计

2.1 用户需求分析

中国科学院资源环境科学信息中心面对的用户主要是科研人员及在读的研究生或博士生,他们从事的是各类科研工作,其信息需求具有鲜明的学科特点。通过对我中心服务的各科研院所和高校的科研工作的调研获知,他们的

科研工作主要集中在化学化工、核科学和资源环境科学三个方面。针对以上分析我们确定,专题虚拟阅览室的构建也相应地分为三个专题虚拟阅览室,即化学化工虚拟阅览室、核科学虚拟阅览室和资源环境科学虚拟阅览室。即按照学科特色将不同类型的信息资源集成在同一个虚拟阅览室中,使科研人员进入某一个虚拟阅览室就可以方便地获取所需的各类资源。

确定虚拟阅览室的学科专题之后,我们又对科研人员使用的信息资源类型进行了调研。科研人员在从事科研工作时,需要及时了解国外同类研究的发展动态,而这部分信息约 70% 来自于外文期刊。其次,随着互联网的飞速发展和不断普及,越来越多的人将网络作为获取信息、学习知识的最佳工具,网络资源也成为科研工作者经常使用的一类重要资源,特别是一些免费的有价值的网络资源,使用频率非常高;另外科研人员也会经常查阅相关的外文图书以及学科新闻。根据用户的这些需求特点,每个专题阅览室分别设置了外文现刊目次、网络资源、中英文图书和学科新闻 4 个栏目。

2.2 虚拟阅览室系统结构和模块设计

2.2.1 系统结构

虚拟阅览室系统采用三层体系结构进行设计,后台通过 TRS 全文数据库平台对数据进行统一的存贮和管理,前端通过 Web 应用服务器 (Tomcat 4.0) 及 JSP 动态网页技术实现对信息的发布和维护,用户则通过 Web 浏览器对信息进行浏览和检索,所有应用都封装在 Web 应用服务器层,具有良好的扩展性和灵活性 (见图 1)。

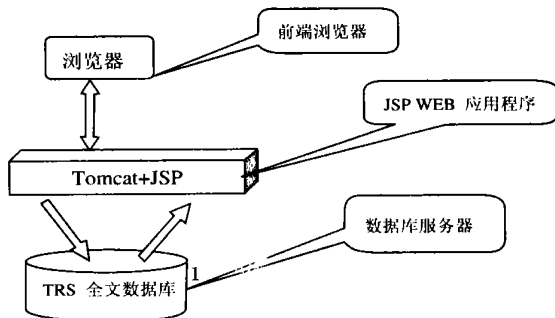


图 1

2.2.2 功能模块设计

虚拟阅览室系统主要由三大功能模块组成:数据加工与维护、信息检索与发布及附加功能 (见图 2)。数据加工维护及信息检索发布实现的主要步骤包括:①建立与数据库的连接;②获取用户的检索请求;③从数据库中检索

* 中国科学院知识创新工程重大项目 (SCDL2002-11) 资助

出符合要求的记录; ④显示记录概览列表及每一条记录的详细信息。

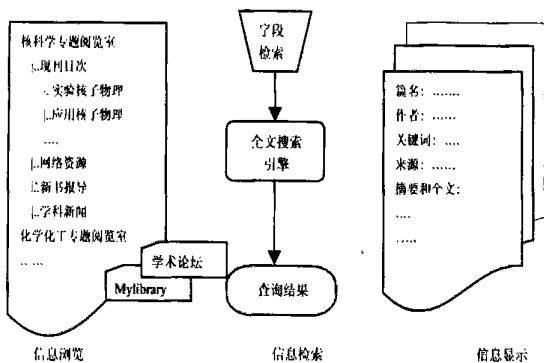


图 2

(1) 数据加工与维护。首先采用 TRS admin (TRS 数据库管理工具) 完成对虚拟阅览室包括建库、数据维护、安全控制、备份和恢复、系统设置等管理任务。按照我中心各专题虚拟阅览室的服务内容, 我们为每个阅览室都设计了 4 个数据库: 包括现刊目次库、网络资源库、新书目录库和学科新闻库。其中在现刊目次库的结构设计中除了常用的篇名、作者、关键词、文摘等字段外, 还增加了电子版全文字段和期刊网址字段, 使现刊目次数据与我中心已订购的电子期刊全文数据库有效链接, 方便用户浏览或下载电子版全文 (包括 PDF 格式和 HTML 格式)。现在国外的期刊绝大部分都有自己的期刊网站, 网站内容包括出版社或杂志社介绍、出版刊物、热点文章、投稿信息等内容, 期刊网址字段的设立, 可以使用户随时了解自己感兴趣的出版社或杂志社。网络资源库的字段设置基本按照 Dublic Core 元数据标准, 对所有精选的网络资源都进行了详细的描述, 方便用户选择使用。

数据库结构设计完成之后, 利用 TRS WAS 4.0 设计了基于 B/S 的数据加工与维护功能, 包括数据的新增、修改、删除和外部数据转换 (TRS 提供多种外部格式的转换, 包括 ISO2709、文本、PDF) 等。通过外部数据转换功能可以方便快速地将新书的 MARC 数据等转入虚拟阅览室数据库。

(2) 信息检索与发布。采用 JSP 开发动态 Web 页面, 主要分为信息浏览和信息检索两部分。信息浏览可以实现按专题阅览室、按信息资源类型及按学科主题浏览信息, 信息检索则基于 TRS 全文搜索引擎, 可以实现对数据库中所有字段包括文摘进行全文检索, 另外可以实现各字段间组合检索以及在结果集中的二次检索; 除此之外, 我们还设计实现了通过作者、关键词、分类号、刊名的超链接检索, 在检索出的某条记录中, 如果点击某个关键词超链接, 系统会自动对此关键词进行检索, 从而检索出包含此关键词的所有记录。这种超链接检索将所有记录组织成了一个网状结构, 极大地方便和拓展了用户的检索。

(3) 附加功能。虚拟专题阅览室应建设成为一个专题学科小门户, 向用户提供具有学科特色的各项服务。我们除了提供信息检索和浏览服务外, 还集成了用户常用的数据库系统, 开发了针对用户个性化需求的 Mylibrary 系统,

使各学科专家可以方便地建立自己的 Mylibrary, 集成自己常用的数据库、网络资源、电子期刊等。另外还设计了针对核科学、化学化工和资源环境的学科论坛, 给科研人员建立了一个可以讨论学术问题的平台。

3 虚拟阅览室发展趋势展望

3.1 逐步构建集成服务平台

构建网上虚拟阅览室实际是建立一个网上专题学科小门户, 它打破了传统阅览室只集中一类文献的局限, 可以集成与专题有关的多钟类型、多种格式的信息资源, 以及多类型、多层次的信息服务。如除了印本文献的揭示外, 还可以集成全文电子期刊、电子书、网络数据库、多媒体资料等信息资源, 以及文献提供、定题服务、学科情报研究、学术论坛、BBS 等信息服务项目。另外专题虚拟阅览室因具有鲜明的学科特色, 还可以相应地建立学科专家工作室, 提供学科专家的个人资料和研究成果等。虚拟阅览室将逐步建成综合的集成服务平台, 目的是为用户提供一站式的服务。

3.2 充分利用先进技术支撑

虚拟阅览室的建设需要网络技术、数据库技术、个性化服务技术等多种技术的支撑, 而计算机技术及网络的发展是日新月异的, 这就需要在构建虚拟阅览室时充分使用成熟的具有可扩展性的先进技术。如支持非结构化数据的全文数据库技术, 支持信息服务扩展的开放链接技术, Mylibrary 个性化服务技术等。先进的技术支撑一方面可以保障虚拟阅览室服务的稳定性, 另一方面将进一步拓展服务的内容并且将极大地方便用户的使用。

3.3 从提供文献单元服务向提供知识单元服务转化

传统的阅览室主要是提供图书、期刊、工具书及各类目录等文献单元的服务, 即用户只能从文献单元层次获取信息; 而虚拟阅览室是对各类数字化信息的集成, 这种集成可以通过对信息的知识化加工, 如主题标引、数值数据、事实数据的深度加工等形成基于知识的信息组织, 从而使虚拟阅览室成为专题的知识组织系统。如对于化学化工专题阅览室可以从化合物结构、化学反应、化学配方、化工产品等多个角度、多个层次揭示和组织信息资源, 使科研人员可以方便地从专业角度检索信息、使用信息, 对信息的知识化加工组织与服务将成为虚拟阅览室的重要发展方向之一。

[参考文献]

- [1] Wu Bao hua, Miao Xiao Yuan, Gao Fei. On the Characteristics of the Digital Library and the Influence to the Work of Reader Service [J/OL]. <http://www.jsai.or.jp/afita/afita-conf/2002/part7/p529.pdf>, 2004-09-12.
- [2] 北京拓尔思公司. <http://www.tris.com.cn/products/>, 2004-09-10.
- [3] 王志刚, 唐文忠, 杨宗煦. 用 JSP 和 TRS 开发文献管理系统 [J]. 计算机应用, 2001, (8): 62-63.

[作者简介] 马建玲 (1969-), 女, 毕业于兰州大学图书情报学系, 中国科学院资源环境科学信息中心副研究员, 发文多篇。

[收稿日期] 2004-12-21 [责任编辑] 张京生