

新疆少数民族语言数字资源检索平台初探

吴倩

(新疆维吾尔自治区图书馆, 新疆 乌鲁木齐 830011)

摘要: 阐述了少数民族语言数字资源整合的必要性、整合背景、意义及检索平台实现技术, 以新疆维吾尔自治区图书馆为例探讨了少数民族语言数字资源检索平台的构建。

关键词: 少数民族语言; 数字资源; 资源整合; 信息检索

中图分类号: G250

文献标识码: A

文章编号: 1002-1248 (2014) 01-0037-03

Primary Exploration of Retrieval Platform for Digital Resources of Minority Languages in Xinjiang

WU Qian

(Library of Xinjiang Uygur Autonomous Region, Urumqi 830011, China)

Abstract: The necessity, background, significance of integrating digital resources of minority languages, as well as technologies for realizing the retrieval platform were clarified in this paper. The construction of the retrieval platform for digital resources of minority languages was explored by taking Library of Xinjiang Uygur Autonomous Region as an example.

Keywords: Minority Languages; Digital resource; resources integration; information retrieval

新疆是多民族聚居的地区, 许多古老民族曾在这里繁衍生息, 发展经济、文化。在漫长的社会进程中, 各民族相互融合, 逐步完成了现代民族的发展进程。目前自治区共有 55 个民族, 少数民族人口占全区总人口的 60.25%。其中有 13 个世居民族, 它们分别是维吾尔族、汉族、哈萨克族、回族、柯尔克孜族、蒙古族、锡伯族、塔吉克族、满族、乌孜别克族、俄罗斯族、达斡尔族、塔塔尔族。在这 13 个世居民族中, 除了汉族、回族、满族和斡尔族主要使用汉语外, 其它民族都有自己的语言, 这使得新疆地区的文化呈现出异常的多样性, 同时也给各民族文化间的交流和沟通带来了一定的困难。开发一款基于网络的多语种文献信息检索平台, 将优秀民族文献数字化加工、整合, 通过检索平台提供信息服务就成为资源共享、信息共通的重要渠道之一。

1 新疆少数民族语言数字资源平台建设背景

目前, 新疆地区使用较广的文种有: 维吾尔文、哈萨克文、蒙文, 使用范围和人数极少的文种有柯尔克孜文、乌孜别克文、塔塔尔文等。文字种类繁多,

给共建网络平台带来困难。20 世纪 80 年代以来, 在国家民族委员会、国家技术监督局、国家科委、电子工业部等有关部门的关心支持和帮助下, 集结起各方面的力量协同攻关, 使民族文字的信息处理工作取得了很大的进展^[1]。其中, 由新疆维吾尔自治区民族语言工作委员会、新疆大学、四川省民族语言工作委员会等多家单位研制完成的维吾尔文、哈萨克文、柯尔克孜文信息处理项目有:《维吾尔、哈萨克、柯尔克孜文字符集国际标准》、《维吾尔、哈萨克等文的键盘、字模国家标准》。在多文种操作系统方面, 新疆未来软件制作室开发了新一代多文种操作系统——《未来多文种系统 UTDOS 6.1》, 实现了民族文字、汉、俄、英一体化的多文种混合处理^[2]。但针对少数民族语言数字资源而开发的多文种检索平台尚属空白。经过专家论证, 认为在新疆地区开发一款目前使用范围较广的维吾尔文、哈萨克文与汉文兼容的共用平台是解决多语种跨平台信息资源检索难题的有效办法。在条件许可的情况下逐渐开发其它文种, 最终完成一个能够容纳更多文种的平台, 以达到网络信息资源共建共享。

收稿日期: 2013-06-09

作者简介: 吴倩 (1977-), 女, 馆员, 新疆维吾尔自治区图书馆技术部, 研究方向: 文献资源数字化加工、资源译制、网站维护。

2 少数民族数字资源整合的意义及其对服务平台的要求

网络环境下,文献信息能否得到充分利用的前提是看其能否转化为数字化信息。少数民族文献是各民族长期发展中形成的文化典籍,是祖国灿烂文化的重要组成部分,其在少数民族经济建设、科学研究及文化建设等方面都有重要的作用。从20世纪90年代开始,许多少数民族地区图书馆尝试建设特色数据库,具有代表性的有内蒙古自治区图书馆研制开发的《蒙文书目机读目录数据库》,四川省高校图工委组织研制的《四川省高校图书馆馆藏西南少数民族文献数据库》,兰州大学敦煌学研究所与图书馆联合开发的《敦煌学数据库》,四川省图书馆《中国西南民族文化多媒体资料库·羌族历史文化》,宁夏大学推出的《西夏学文献题录数据库》等,都具有非常高的学术价值和社会价值^[3]。数字化后的民族文献信息不仅可供研究人员从不同的途径进行检索,作各种目的的统计分析,而且可供网上流通,实现资源共享。因此,将优秀民族文献数字化,并通过资源整合平台发布使用,对充分开发利用民族文献,促进民族文化事业向现代化和国际化方向发展具有重要作用。

随着数字化资源日渐丰富,如何将不同文种数据库、不同数据商及自建专题特色数据库、进行合理的组织并提供有效的利用,对多语种、多平台、多类型、多领域的数字资源进行整合,是当前一项重要的任务和值得研究的课题,其整合程度直接关系到数字资源能否被高效吸收与利用。

无论是民族文献的数字化过程还是多语种数字资源整合平台,民族文字的信息处理技术既是关键也是难点,要求其检索平台能够综合运用各种技术、方法和手段对图书馆所拥有的众多数字资源进行系统化和优化,实现以统一友好的界面向用户提供集成检索、浏览和使用所有数字化资源,以更好的实现少数民族文化现代化和国际化发展。

3 新疆少数民族语言数字资源检索平台建设构想

新疆作为多民族聚居地区,资源建设要突出新疆地方特色、民族特色和文化特色,以少数民族语言资源建设为重点,同时侧重新疆地方文献数字化建设。

3.1 新疆维吾尔自治区图书馆数字资源现状

新疆维吾尔自治区图书馆由于所处

区域的民族特点、风俗文化的特殊性,以及所承担的公共文化服务任务,其数字资源建设应结合本馆及本地区的实际需要,建设具有地方特色的数字资源库,为地方经济建设和各项事业的发展提供信息服务。

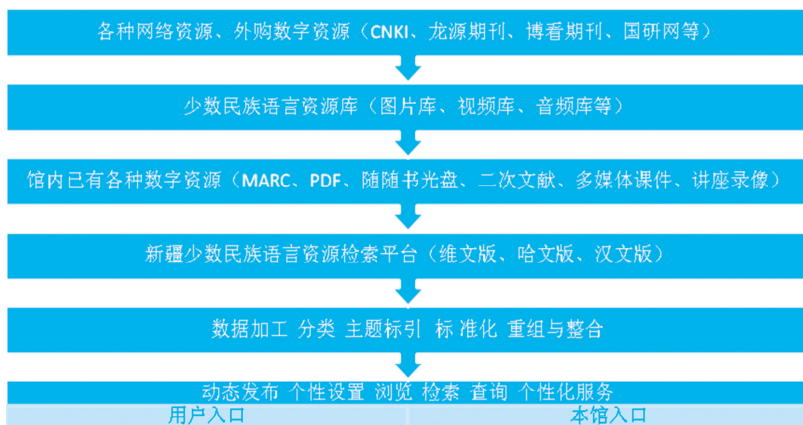
截至目前,新疆维吾尔自治区图书馆已建成的少数民族语言数据库有农牧区实用技术库、少数民族爱国主义题材影视库、非物质文化遗产库、少数民族少儿“双语”及基础教育库等,正在建设的特色数据库有新疆地方文献数据库、新疆印象之民风民俗数据库、新疆印象之红色记忆图片数据库等,外购数据库有CNKI中国知网、国研网、龙源期刊、博看期刊、万方视频、万方政策法规、中国数字图书馆电子图书等。

3.2 新疆少数民族语言数字资源检索平台建设构想

新疆少数民族资源检索平台的构建以资源服务为核心,将多文种文献信息及多语种数字资源管理与服务的全过程看成一个统一系统,融用户服务、资源管理和检索结果控制为一体,达到用户服务满意、管理过程经济高效、检索结果准确的最大价值。检索平台的总体建设思路体现了用户统一检索、元数据统一收割、接口标准相对统一的规划思想,实行统一系统框架下二级系统之间的信息共享和信息交换机制。平台构建采取基于资源导航并与OPAC系统无缝对接的整合模式,将现有数字资源的检索入口整合在一起,建立数字资源导航库,提供按资源名、关键词、资源标识等获取数字资源的途径,用户可通过不同语种界面登陆,由检索口到数据库查找的方法浏览资源。

新疆少数民族语言资源检索平台基本框架如附图。

数字资源整合研究是一项系统工程,在建设过程中,必须总体规划,长远布局,协调发展,大力合作,避免重复建设。在整合数字资源的组织上,要实现各种数字资源要素的结构性优化;在数字资源的类型上,要实现从目录、文摘深入到全文,包括引文、全文的



附图 新疆少数民族语言资源检索平台基本框架

无缝链接;在数字资源载体上,要实现从文本资料到图片、音频资料、视频资料及网络信息资源的并轨^[4]。

3.3 新疆少数民族语言资源检索平台页面设计

后台结构设计完成后,要通过页面设计显示出来。根据资源结构及用户需求,检索平台首页以Web网页形式发布。其特性是将图形、音频、视频信息集合于一体,同时易于导航,可以建立各信息之间的链接,用户可以从一个链接跳到另一个链接,在各页各站点之间进行浏览。设计者要了解熟悉数据库内容,针对用户设计出特色鲜明、主题突出、编排合理、便于大众理解、使用方便的页面。页面设计的根本目的是揭示资源内容,设计人员只有熟悉、了解平台结构及数据库内容,做到形式和内容统一协调,将丰富多样的内涵用简洁明了的设计展示出来,才能从方便用户的角度使数据库的内容得到形式上的最佳体现。

4 结语

平台开发人员提交平台建设工作总结后,在平台验收会上要进行平台演示、测试说明,由上级主管部门聘请专家对少数民族语言资源检索平台的任务完成情况、系统功能、数据规范等页面效果进行综合验收

评审,验收合格将在新疆维吾尔自治区试运行。检索平台功能的认识、维护、管理和完善是一个循序渐进的过程。作为系统管理人员,应对其应用过程中常出现的问题,及时进行跟踪、分析和研究,并把自己的工作实践与解决方法提出来,反馈给软件研发人员,通过双方的共同努力使问题得到圆满的解决,进而提高检索平台的性能。新疆少数民族语言资源检索平台开发、测试完成后,可以使现有少数民族数字资源环境得到全面优化,帮助少数民族地区图书馆提升信息服务水平,实现从信息服务到知识服务的转变,并进一步构建出满足少数民族地区需要的知识库和知识网络。

参考文献:

- [1] 包和平,王学艳.中国少数民族文献数字化建设研究[J].情报方法,2002,(2):34
- [2] 包和平,王学艳.中国少数民族文献数字化建设研究[J].图书馆理论与实践,2002,(2):34
- [3] 陆凤红.以特色数据库的建设与共享展示西部少数民族文献的价值[J].图书馆理论与实践,2009,(6):101,
- [4] 王淑娟.数字资源整合平台优化研究[J].情报探索,2011,(3):64.