

浙江天宇信息技术有限公司协办

中国少数民族语言 数字信息分布式共享研究

赵生辉 (武汉大学信息管理学院 武汉 430072)

摘要 文章在介绍中国少数民族语言信息技术发展概况的基础上,探讨了少数民族语言数字信息共享的重要意义,总结了少数民族语言数字信息共享的现状,分析了少数民族语言数字信息共享的需求结构,提出了兼顾“多元性”需求和“一体化”需求的分布式信息共享模式,并设计了实现该模式的数据库系统技术架构、组织架构和制度架构。

关键词 数字信息管理 少数民族语言 多元|体架构 分布式集成 跨语种信息 共享

Distribution and Sharing of the Digital Information of Minority Languages in China

Zhao Shenghui (School of Information Management, Wuhan University, Wuhan, 430072)

Abstract Based on overview of development of information technology of the minority languages of China, importance of information sharing is discussed, the status of sharing of digital information in minority languages is summarized, and the requirements of digital information sharing is analyzed. Then, the author proposes the distributed information sharing model based on distributed database system which takes requirement of diversity and integrity into account together. The technology architecture, organizational structure and institution framework support distributed information sharing are also designed.

Keywords digital information management, the minority languages of China, diversity & integrity architecture, distributed integration, cross-language information sharing

中国是统一的多民族国家,各民族在长期的历史发展中形成了多元一体的格局。这种多样性的统一体现在语言文字上,除了全国通用的汉语普通话和规范汉字以外,大多数少数民族还使用着自己民族的语言,部分民族还有记录自己语言的文字。中国政府在各民族地区实行民族平等的双语或多语政策,少数民族文字信息处理研究得到国家相关机构的支持。20世纪80年代以来,蒙古文、满文、维吾尔文、哈萨克文、柯尔克孜文、藏文、朝鲜文、彝文、傣文等计算机系统、字处理软件相继诞生,少数民族文字信息处理有关标准也相继制定。近年来,随着少数民族地区电子政务建设的推进,藏文、蒙古文、维吾尔文等少数民族文字办公系统在部分地区的政府机构得到了应用。与此同时,众多的少数民族语言网站在互联网上

亮相。此外,民族地区广播、电视、出版等文化机构也大多采用兼容少数民族语言的编辑和排版系统,民族古籍和语言文字资料整理过程中也会用到少数民族语言文字处理系统。随着少数民族语言文字信息处理的研究进展和民族地区信息化进程的推进,如何稳妥、科学、有效地管理随之产生的少数民族语言数字信息,是民族地区信息化建设的重要课题。

少数民族语言数字信息是信息时代各少数民族和相关机构进行信息传播和沟通交流的手段,是少数民族文化的载体,也是全民族共同的精神和文化财富,少数民族语言数字信息的长期保存和资源共享具有多维价值:(1)从信息传播角度而言,少数民族语言数字信息共享可以防止语言文字异构造成的“信息孤岛”,实现跨语种信息沟通 and 理解。(2)从经济角度看,

少数民族语言数字信息共享可以使需要了解相关信息的人非常便捷地获取同一主题下各类语言文字的数字信息,避免在各个地区间跋涉所付出的资金和时间成本。(3)从科研角度看,计算语言学研究需要以大量的数字信息为基础,构建多语种语料数据库,少数民族语言数字信息共享可以为少数民族语言语料库建设奠定基础。(4)从管理角度看,少数民族语言数字信息共享必须以一体化技术和管理标准为前提,将极大地促进各民族地区信息管理工作的规范化。(5)从政治角度而言,少数民族语言多语种信息共享可以体现政府保护少数民族文化和维护国家统一管理的政治态度。(6)从社会心理角度而言,少数民族语言多语种信息共享可以增强各民族对国家整体性的认同感,增强民族凝聚力和向心力。总之,少数民族语言数字信息共享已经成为民族地区信息和档案管理工作必须面对的重要课题。

1 现状概述

1.1 少数民族语言数字信息共享面临诸多现实困难

民族地区自然和社会环境下,要实现数字信息贡献面临着很多实际的困难,主要有:(1)民族地区地理环境的制约。中国的少数民族聚居区大多处在边疆地区,地域广阔、人口稀少,西藏、新疆等少数民族自治区还有大范围的高寒地区,气候条件恶劣,对数字信息共享必须的通讯线路的铺设等造成困难。(2)民族地区经济社会发展阶段的制约。由于受到多方面因素的制约,目前中国的少数民族地区经济发展相对滞后,支撑民族地区经济的主要产业是农牧业,人口的整体文化程度相对偏低,现代信息技术的应用存在一定的困难。(3)公共信息管理体制的制约。按照《中华人民共和国民族区域自治法》和《档案法》的规定,中国目前的公共信息管理体制是“统一领导、分级管理”,这就意味着要实现跨语种、跨地区的信息共享面临着法律上的制约,必须探索出与体制相适应的管理模式。

1.2 当前中国少数民族语言数字信息共享的特征

(1)使用人口较多的几种少数民族文字的信息处理技术基本成熟。少数民族文字的数字化问题已经引起国家和各民族地区的重视,相关科研院所和企业在少数民族文字处理技术研究开发方面做了大量的工作,并取得丰硕成果,如表1所示。

(2)多语种信息共享技术研究已经取得部分成果。1991年4月由内蒙古电子计算中心课题组完成了“蒙、藏、维、哈、朝、满、汉文操作系统”,该系统属于“七五”国家重点科技攻关项目“少数民族文字处理技术开发”的子课题,可以同时处理蒙、藏、维、哈、朝、满、汉、英等文字。2006年,国家高技术研究发展计划(863计划)信息技术领域“中文为核心的多语言处理

表1 中国主要少数民族文字处理技术的研究进展^[1]

语言名称	少数民族语言文字处理系统
藏文	青海师范大学“TCDO S2.0 版藏文系统”;中国藏学研究中心“藏文处理及激光编辑排版印刷系统”;华光书林“藏文排版和激光照排系统”;维思“基于 Windows 的藏文彩色印刷系统”;SPDOS“汉藏文操作系统和藏文字处理平台”。
蒙古文	内蒙古计算中心“蒙汉英操作系统”;华光书林“蒙文书刊、图表、报纸激光照版系统”;MPS“蒙汉混合字处理系统”;北大方正“蒙文电子出版系统”;内蒙古社会科学院“蒙古文处理系统”。
维、哈、柯文	博格达“维、哈、柯、汉、英、俄多文种排版系统”;北大方正“维、哈、柯文排版系统”;新疆“多文种系统 VTDOS6.1”。
朝鲜文	北大方正“朝鲜文书版系统”;华光书林“朝鲜文系统”。
满文	新疆“满文处理和轻印刷系统”;内蒙古大学满文系统;辽宁省档案馆满文操作系统。
彝文	北大方正“彝文书版软件系统”;西南民族大学“VCDOS 汉彝双语平台、SP-DOS 汉彝双语操作系统、基于 Win95 的彝文信息处理平台”;云南省“规范彝文排版系统”。
傣文	北大方正“傣文电子出版系统”;潍坊华光“傣文电子排版系统”。
壮文	北大方正“壮文书版系统”。
苗文	光亚电子“苗文排版系统”。

技术”重点项目启动^[2],将“多语言基础资源库研制和共享”、“多语言语音关键技术研究与应用产品开发”、“跨语言跨媒体网络搜索关键技术研究和服务产品开发”以及“多语言综合信息服务应用示范”作为研究重点。2007年,“统一平台少数民族文字识别系统”在清华大学研制成功^[3],该系统是全球首款在统一平台上支持我国主要少数民族文字文档的识别系统,在汉字和英文文档识别的基础上可识别四种类型六种类别的少数民族文字,即蒙古文、藏文、维吾尔文、哈萨克文、朝鲜文和柯尔克孜文(混排汉英),文档识别系统集成在一个统一的平台系统中,使我国最主要的几种少数民族文字文档能够自动识别输入计算机。

(3)部分小规模多语种信息共享应用项目得到实施。例如,中央电视台2004年开始多语种数字电视字幕的尝试^[4];吉林省延吉市延边朝鲜族自治州在2008年建成多语种科技服务平台^[5];2009年,“人民网”的蒙文、藏文、维吾尔文、哈萨克文、朝鲜文、彝文、壮文七种少数民族语言版本全部上线,实现了七种少数民族文种的全覆盖^[6];云南省从2009年开始研发“云南少数民族语言文字资源数据库”项目^[7],旨在加强民族语言文字规范化、标准化和信息化建设,记录、抢救和保护少数民族的语言文字。

1.3 中国少数民族语言数字信息共享中的问题

我国各类机构在少数民族语言数字信息共享方面已经做了大量工作,但是整体而言,目前仍然处在起步阶段。存在的问题主要有:

(1)缺乏战略性思考。主要表现在目前的研究和实践大多关注单一语言数字信息,对于各类少数民族语言数字信息的共性问题关注较少,从国家层面进行整体性、战略性研究研究的成果还很少。现有的少数民族语言数字信息共享的实践项目大多局限在单一组织和局部区域,还没有启动一个从国家层面上整体规划的少数民族数字信息共享应用项目。

(2) 侧重文字处理技术, 对信息管理问题关注不够。主要表现为研究和实践中普遍更加关注少数民族语言文字计算机处理的技术实现, 对于信息处理系统所产生的数字信息如何管理关注不够, 特别是在少数民族语言数字信息著录方面还没有统一的标准, 不利于数字信息检索和信息交换。

(3) 没有建立强有力的组织协调机构。当前少数民族语言数字信息共享程度较低, 民族地区的大多数机构和个人的信息共享意识不强, 促进信息共享的动力不足。没有一个强有力的组织协调机构来妥善处理信息共享中的各类政治、法律、经济、文化等关系, 整合各个民族地区和机构的数字信息资源, 是造成这种局面的重要原因。

2 需求分析

少数民族语言数字信息共享是一个系统工程, 它包含了技术、政治、法律、文化等多方面的因素。要解决信息共享过程中的各类问题, 首先必须对共享的功能需求进行系统分析。这里的需求, 是指相关各方对少数民族语言数字信息共享体系在功能上的要求。一般而言, 少数民族语言数字信息共享需求分为民族地区和国家两个层面: 民族地区居民和机构一般更多的强调少数民族语言数字信息所代表的民族文化的保护, 也就是对多元性的要求; 与此相对, 国家层面上需要尽可能实现国家各类数字信息资源的共享, 维护国家的一体化管理。因此, 中国少数民族语言数字信息共享必须同时考虑多元性和一体性两个方面的需求, 实现两类需求的平衡, 保证最终的解决方案科学、合法并且可以顺利实施。

2.1 少数民族语言数字信息共享的多元性需求

从各民族地区的视角来看, 少数民族语言数字信息共享应该致力于保护少数民族文化的多元性, 主要体现在: (1) 保护少数民族语言文字权。重点是保证少数民族人口在社会生活中使用本民族语言文字的权力, 确保少数民族语言数字信息的法律效力。(2) 少数民族文化保护。将少数民族语言数字信息视为蕴含了少数民族文化的数字资源予以保护, 保证实现长期可读。(3) 体现《中华人民共和国民族区域自治法》的法律精神^[8]。中国在各少数民族聚居区实行民族区域自治制度, 在宪法规定的范围内, 给予少数民族人口独立处理日常事务的自由。少数民族语言数字信息的共享同样要体现这种自治的精神, 给予各少数民族地区一定的自由权。(4) 落实《中华人民共和国档案法》中的“分级管理”原则^[9]。档案分级管理原则是指各个地区按照国家法律法规的规定分级组织档案信息管理工作, 同一级机构对与之对应级别的机构所产生的档案资料负责。这就意味着, 各少数民族地区要对自己区域内的少数民族语言数字信息的管理负责。(5) 保

持各类信息资源之间的逻辑联系。按照档案管理的原则, 档案信息资源的保存应该尽可能保持各类信息资源之间的逻辑联系, 尽量不要人为割裂。由于少数民族语言数字信息在实践应用中的非官方性, 往往是与国家通用语言文字版本的文件配合使用, 在保存的时候, 要尽量保持这些文件之间的原始联系。

2.2 少数民族语言数字信息共享的一体性需求

从国家整体视角来看, 少数民族语言数字信息共享应该尽最大的可能保证统一的多民族国家管理的一致性, 主要体现在: (1) 政治上的统一象征。少数民族语言的使用与政治之间具有密切的联系, 常常成为极端民族主义分子煽动民族矛盾的借口。作为统一的多民族国家, 中国需要各个地区在语言的使用上体现出统一性特征。(2) 维护国家通用语言文字的法律地位。《中华人民共和国通用语言文字法》规定, 国家的通用语言文字是普通话和规范汉字, 国家必须依法维护通用语言文字的法律地位^[10]。(3) 满足跨语种信息检索的需求。中国是统一的多民族国家, 所有少数民族语言数字信息是人民的共同财富, 在有跨语种检索要求的时候, 应该可以检索出相同主题下各类语言文字版本的数字信息。(4) 数字文件的价值保证。按照档案学理论, 数字文件价值必须以必要的元数据著录为保证, 如果关键元数据缺失, 其真实性就不能被承认。为此, 国家应该对少数民族语言数字信息的著录做出统一的规定。(5) 数字信息长期保存。由于数字信息对生成环境具有依赖性, 在脱离原生技术环境后, 如果无法得到原生环境的技术参数, 这些数字信息可能会面临无法读取的风险。因此, 出于对语言数字信息长期可读性的目的, 必须对技术系统参数的元数据著录进行统一规定。(6) 民族语言语料库建设的需要。计算语言学需要以大量少数民族语言语料为基础, 需要积累大量的少数民族语言语料资源, 集中统一管理是实现这一目标的有效方式。(7) 少数民族语言数字信息的知识管理。少数民族语言数字信息当中可能包含文学、医药、手工艺等方面的知识, 而识别这些知识也需要将大量数字信息集中管理后发掘其中隐藏的规律性。

2.3 少数民族语言数字信息共享的需求平衡

中国少数民族语言数字信息共享面临模式的选择问题。信息资源的管理模式一般分为“集中式”和“分散式”两类。“集中式”要求将所有的信息资源在物理上集中到同一场所, 进行统一管理; “分散式”则是各个地区独立开展信息资源管理工作。“集中式”可以最大限度的提高信息资源共享的程度, 但是少数民族语言信息资源的利用也具有地域性特征, 往往产生信息的机构所在地区的利用需求最为强烈, 集中式管理在满足局部信息利用需求方面效率相对较低。

“分散式”管理可以最大限度地保证各个地区的

信息利用需求,但是难以实现跨地区的信息共享,无法进行有效的知识管理,信息利用的范围和利用的类型有限。如果强调少数民族语言数字信息共享的“多元性”需求,则管理模式应该选择“分散式”,给予各个民族地区较大的自主权;如果强调少数民族语言数字信息共享的“一体性”需求,则管理模式应该选择“集中式”,在国家层面上建立一个大型数字信息管理中心,收集和保存来自全国的所有少数民族语言数字信息。“分散式”管理模式将会造成多个信息孤岛,无法满足信息共享的要求。“集中式”管理模式则可能违背分级管理原则,破坏产生地区多种类型信息之间的逻辑关联,给本地用户的利用造成不便,因而也不能有效满足系统的功能需求。

中国少数民族语言数字信息的共享模式必然要在“多元性需求”和“一体性需求”之间做出平衡,最终的解决方案必须兼顾这两个方面的需求,这样的模式就是“分布式”共享模式。

分布式共享就是通过建立分布式数据库,实现多区域、多类型的数字信息一体化共享。所谓“分布式数据库”就是物理上分散而逻辑上集中的数据库系统。分布式数据库系统使用计算机网络将地理位置分散而管理和控制又需要不同程度集中的多个逻辑单位连接起来,共同组成一个统一的数据库系统^[1]。分布式数据库系统具有三大特征:(1)物理上的分散性。分布式数据库当中的数据不是存储在一个站点上,而是分散存储在由计算机网络连接起来的多个站点上,而这种分散存储对于用户来说是感觉不到的。(2)逻辑上的整体性,分布式数据库系统中的数据在物理上分散在各个站点中,但是这些分散的数据逻辑上却构成一个整体,它们被分布式数据库的所有用户共享,并由一个分布式数据库管理系统统一管理,使得“分布”对用户而言是透明的。(3)站点的自治性。也叫场地自治性,各个站点上的数据由本地的DBMS管理,具有自治处理能力,完成本站点的局部应用。

综上所述,基于分布式数据库的中国少数民族语言数字信息共享模式可以有效地平衡“多元性需求”和“一体性需求”的矛盾,在保障各民族地区数字信息管理的独立自主和分级管理的同时,又能够实现少数民族数字信息的一体化共享,是中国少数民族语言数字信息共享的理想模式。

3 架构设计

中国少数民族语言数字信息的分布式共享的核心是构建分布式数据库,除此之外,要保证整个架构的正常运转,还必须建立相应的组织架构和制度架构。

3.1 技术架构

基于分布式数据库的中国少数民族语言数字信息共享系统的技术原理如图1所示。

图1中,分布式数据库中的数据主要由两部分组成,一部分是应用所需要的数据集(Data Base, DB),它是数据库的主体,一部分是对数据库当中数据结构的定义,以及全局数据的分片、分布的描述,称为数据目录(Data Directory, DD)。在分布式系统当中,数据和描述它的数据目录有全局和局部之分,局部数据称为LDB,局部数据目录称为LDD,全局数据称为GDB,全局数据目录称为GDD。与此相对应,数据库管理系统也被分为局部数据库管理系统LDBMS和全局数据库管理系统GDBMS,用户也根据检索所涉及的数据范围的不同被分为Local Client和Global Client。

上述架构中,所有的数据在物理上是分散的,但在逻辑上是集中的。各个局部数据库可以独立运作,在产生全局应用需求的时候,在网络通信管理(Communication Management, CM)接口的协助下通过网络调取全局数据。其中,局部数据库的主要作用是满足本地数字信息的保存、管理、检索和利用,而全局数据库是一个虚拟数据库,全局用户可以非常方便地使用全局数据库中的资源,尽管这些数据库在物理上可能是分散在各个地区的,实现数字信息的跨地区共享。

中国少数民族语言数字信息分布式共享还需要处理好以下问题:

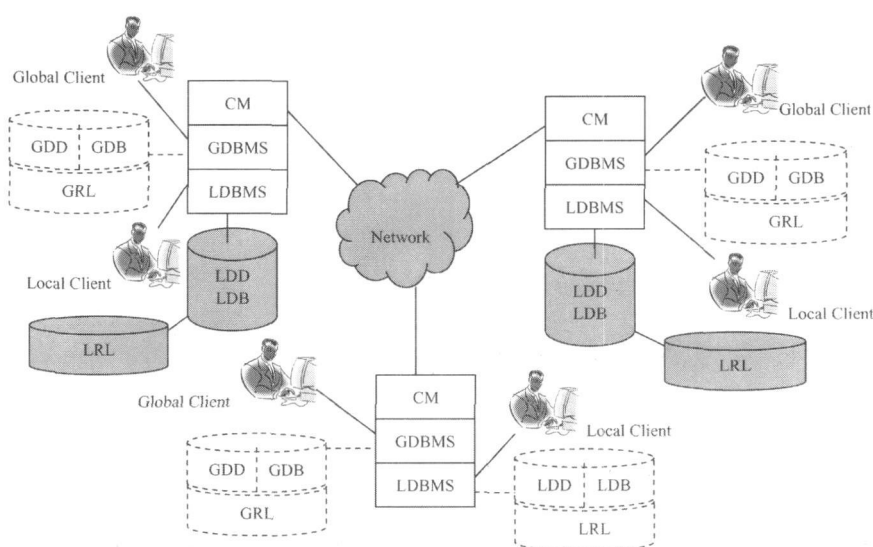


图1 少数民族语言数字信息分布式共享示意图^[11]

(1) 数据结构设计。绝大多数数字信息是非结构化的,而分布式共享要求数据是结构化的,一般是关系型数据库。分布式架构中,本地数据库当中所存储的应用数据集合只能按照统一的元数据模型进行著录后得到数字信息的元数据。为了对数字信息本身的共享,可以建立用来存储数字信息的资源库(Resource Library, RL),并在元数据和内容数据之间建立关联。当需要读取数字对象的时候,再访问资源库。因此,本地数据库需要保存数字信息、著录元数据、数据目录三方面内容,其中著录元数据是对数字信息的描述,数据目录是对著录元数据的描述。其中,元数据模型的设计是系统成功的关键,必须按照科学化的程序,综合多方面的需求进行设计。

(2) 语言文字的信息编码。为了保证整个数据库系统不出现乱码,建议采用国际标准 Unicode 作为统一的编码标准。为此,在数据采集时,对没有采用该标准的信息要进行转换。

(3) 跨语种信息共享问题。为了解决多语种之间信息共享问题,需要选择一种中介语言,中国国家通用语言汉语则可以充当这种角色。按照这一要求,每个局部数据库的数据著录的时候需要同时采用两种语言进行著录,一种是数字信息本身所采用的语言,另外一种汉语。按照这种模式,在全局应用当中,每个局部数据库可以采用少数民族语言关键词进行信息检索,而全局数据库执行的是按照汉语进行的检索,全局数据库当中所有同一主题而语言文字各不相同的数字信息均可以被检索到。同时,也可以通过阅读这些数字信息的汉语著录信息来了解跨语种数字信息的内容。

(4) 通讯网络的选择问题。考虑到民族地区经济发展相对滞后,可以将上述分布式架构建立在国家电子政务基础网络上,节约成本的同时也可以提高信息安全的水平。

(5) 少数民族语言数字信息的长期保存问题。相对于英语和汉语,少数民族语言文字信息处理技术发展不成熟,经过大规模应用的软件数量不多。为了保证长期可读,必须在保存数字信息的同时,也保存产生这些信息的应用软件。在分布式共享系统中,为读取信息,数字信息和应用软件可以同时下载。

(6) 容灾备份问题。多点异地备份是分布式系统的优势。适当地增加相关冗余信息的存储,既可实现相互备份,同时也可以减少调用全局检索的次数,提高系统整体效率。

3.2 组织架构

少数民族语言数字信息管理涉及多个部门之间的关系协调,其中最主要的是民族事务管理、档案管理部门之间的关系,按照分布式共享的思想对其组织架构设计如图2。

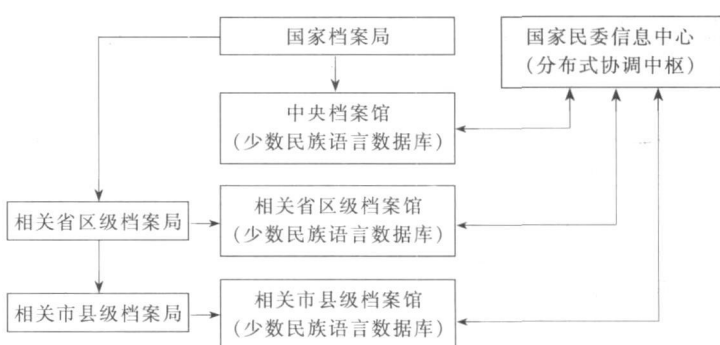


图2 少数民族语言数字信息共享的组织架构

图2将少数民族语言数字信息共享的任务做了组织分工,档案管理部门负责数字信息的采集、整理、著录、保管,国家民族事务管理委员会信息中心负责少数民族语言数字信息的共享,地方档案馆只要完成分布式数据节点少数民族语言数字信息的归档管理,就可以按照法律规定共享系统内全国范围允许公开的少数民族语言数字信息。另外,国家民族事务管理委员会信息中心也可以了解少数民族地区的信息,增加决策的准确性。这种组织架构的最大优点是符合《国家档案法》规定的“统一领导、分级管理”原则,不需要增加新的工作机构,只要在现有工作体制的基础上增加新的工作职能就可以正常运转。

3.3 制度架构

为了确保少数民族语言数字信息分布式共享的实现,还需要一系列制度的支撑,如表2所示。

表2 中国少数民族语言数字信息共享的制度保障

类别	制度名称	相关内容
国际标准	ISO 10646 Information Technology Universal Multiple Octet Coded Character Set	多语种信息编码
国家法律	中华人民共和国宪法 中华人民共和国民族区域自治法 中华人民共和国通用语言文字法 中华人民共和国档案法 中华人民共和国保密法 中华人民共和国信息公开法	民族政策 民族语言文字政策 民族语言文字政策 档案管理体制 涉密信息管理 信息公开范围
国家标准	GB/T 15418-1994 档案分类引规则 GB/T 20530-2006 文献档案资料数字化工作导则 GB/T 8894-2002 电子文件归档与管理规范	电子文件分类引 民族数字化 电子文件管理规程
部门标准	DA/T 18-1999 档案著录规则 DA/T 19-1999 档案主题标引规则 DA/T 30-2002 满文档案著录名词与术语汉译规则	电子文件元数据 电子文件主题标引 双语翻译规则
工作制度	少数民族语言数字信息管理责任归属方案 少数民族语言数字信息的认定规范 少数民族语言数字信息的采集规范 少数民族语言数字信息元数据设计方案 少数民族语言数字信息双语著录规范 少数民族语言数字信息前端控制规范 少数民族语言数字信息的安全保护规范	职能责任划分 管理对象认定 信息采集规程 元数据方案 双语著录规则 前端质量控制 信息安全保护

4 结语

基于分布式数据库的分布式共享模式可以有效地平衡“多元性需求”和“一体性需求”的矛盾,在保障各民族地区数字信息管理的独立自主和分级管理的同时,又能够实现少数民族数字信息的一体化共享,是中国少数民族语言数字信息共享的(下转第66页)

在 Web2.0 环境下,利用 Blog、Wiki、RSS、Tag 等技术,采取用户参与的图书馆信息组织模式,重视用户价值,吸引用户建设和组织图书馆资源,实现用户对资源的共建共享,能增强资源揭示的深度与广度。但要实现用户参与图书馆信息组织,还需争取来自图书馆、图书馆员以及外部信息环境等诸多方面的支持。

参考文献

- [1] Ellyssa K. The social tools of Web2.0: Opportunities for academic libraries[J]. Choice, 2007(8): 2011-2021.
- [2] 欧阳剑. 新网络环境下用户信息获取方式对图书馆信息组织的影响[J]. 中国图书馆学报, 2009(6): 97-102.
- [3] Macgregor G, McCulloch E. Collaborative tagging as a knowledge organization and resource discovery tool[J]. Library Review, 2006(5): 291-300.
- [4] Shirky C. Ontology is overrated: Categories, links and tags[EB/OL]. [2010-05-20]. <http://www.shirky.com/writings/ontology-overrated.html>.
- [5] 汪会玲. 基于 Web2.0 的信息服务实践及其改进策略[J]. 情报理论与实践, 2009(5): 56-60.
- [6] 张树人. 从社会性软件、Web2.0 到复杂适应信息系统研究[D]. 北京: 中国人民大学, 2006.
- [7] 邓胜利. 交互式信息服务的构成要素与定位分析[J]. 情报理论与实践, 2009(1): 18-21, 13.

- [8] 李铭薰, 卜小蝶. 电脑与数位 3C 专业虚拟社群知识分享之影响因素研究[J]. 教育资料与图书馆学, 2009(3): 377-401.
- [9] Brookes, B. C. Foundation of information science(I): Philosophical aspects[J]. Journal of Information Science, 1980(2): 125-133.
- [10] 张红兵, 和金生. 基于 Web2.0 的知识共享系统[J]. 图书情报工作, 2009(6): 118-121.
- [11] 国际农业研究协商小组 2020 未来发展宣言书[EB/OL]. [2010-05-12]. <http://www.redlib.cn/html/14414/2009/60655390.htm>.
- [12] Factors affecting user participation in video UCC(User-Created Contents) service[EB/OL]. [2010-02-20]. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.88.6789&rep=rep1&type=pdf>.
- [13] 庄玉梅, 张体勤. 个体知识分享行为的理论透视[J]. 图书情报工作, 2009(20): 118-121.
- [14] 邓有盈, 柯浩仁. 于数位典藏建立社会性标记之研究: 以杨英风数位美术馆为例[J]. 图书与资讯学刊, 2009(1): 80-106.
- [15] 刘炜, 张春景. 试论网络资源的规范控制[J]. 现代图书情报技术, 2008(12): 27-31.

[作者简介] 陈红艳, 女, 1976 年生, 湖北大学图书馆馆员。

司莉, 女, 1965 年生, 武汉大学信息管理学院教授, 博士生导师。

收稿日期: 2010-12-17

(上接第 61 页)理想模式。为了实现这一模式,需要设计并实施以分布式数据库为基础的技术架构,同时建立相应的组织架构和制度架构以支撑分布式共享体系的运作。

参考文献

- [1] 华林. 论少数民族文字历史档案的数字化技术保护[J]. 档案学研究, 2006, 132(2): 36-39.
- [2] 中华人民共和国科技部. 国家高技术研究发展计划(863 计划)“中文为核心的多语言处理技术”重点项目课题申请指南[OL]. [2010-12-04]. <http://www.most.gov.cn/tztg/t2006100136442.htm>.
- [3] 我国研制成功少数民族文字识别系统[OL]. [2010-12-04]. http://www.gov.cn/fwxx/kp/2007-01/31/content_513378.htm.
- [4] 陈力. 中央电视台对多语种数字电视字幕的尝试[J]. 现代电视技术, 2004, 7(10): 17-19.
- [5] 刘志明. 方学苑. 延边州多语种科技服务平台的建设[J]. 技术与市场, 2008, 11(7): 23-26.
- [6] 人民网. “人民网实现七个主要少数民族语种全覆盖”

[OL]. [2010-12-04]. <http://politics.people.com.cn/GB/8198/158422/158569/9520998.html>.

- [7] 新华网云南频道. 云南实施“少数民族语言文字资源数据库”项目保护民族语言文字[OL]. [2010-12-04]. http://www.yn.xinhuanet.com/newscenter/2010-09/23/content_20976381.htm.
- [8] 中华人民共和国民族区域自治法[OL]. [2010-12-04]. http://www.gov.cn/test/2005-07/29/content_18338.htm.
- [9] 中华人民共和国档案法[OL]. [2010-12-04]. <http://www.saac.gov.cn/articleaction.do?method=view&id=ff808081172649a801172660f3010001>.
- [10] 中华人民共和国通用语言文字法[OL]. [2010-10-05]. http://www.gov.cn/ziliao/flfg/2005-08/31/content_27920.htm.
- [11] 邵佩英. 分布式数据库系统及其应用[M]. 北京: 科学出版社, 2006: 22-26.

[作者简介] 赵生辉, 男, 1977 年生, 西藏民族学院管理学院讲师, 武汉大学信息管理学院博士研究生。

收稿日期: 2010-12-17