

文章编号: 0254-5357(2006)04-0369-04

地质矿产实验室分析测试元数据管理软件系统

吴晓军, 甘露, 吴淑琪, 邓赛文

(国家地质实验测试中心, 北京 100037)

摘要: 在已制定的《地质矿产分析测试元数据建库工作指南》的基础上, 设计了地质矿产分析测试元数据库管理系统模型和库结构内容, 研究开发了相应的数据管理软件, 介绍了软件各部分功能。该管理软件已在内部网站上运行。

关键词: 分析测试; 元数据; 元数据标准; 数据库; 软件

中图分类号: TP311.52; TP39:P59 **文献标识码:** A

Management Software System for Analytical Metadata in Geological Laboratory

WU Xiao-jun, GAN Lu, WU Shu-qi, DENG Sai-wen

(National Research Center for Geoanalysis, Beijing 100037, China)

Abstract: Based on the <Guide for Construction of Analytical Metadata Base in Geological Laboratory>, the management system model and database structure for the analytical metadata were designed. The relevant data management software was also developed and the functions of the software were introduced. The management software has been applied to the analytical metadata management in author's laboratory.

Key words: analysis; metadata; metadata standard; database; software

地质工作是我国现代化建设的先行性、基础性工作, 它对提高资源供给能力和保障程度, 对于全面建设小康社会具有重要意义。目前地质工作和新一轮地质大调查工作已经从单纯的陆域找矿向多元化方向发展, 测试领域亦从单一的元素分析扩大到多目标的无机、有机、形态分析, 其数据之繁多、类型之复杂、质量之高前所未有。现代化和智能化仪器的广泛使用, 使实验室数据以指数形式急剧上升, 对实验室的数据管理、实验信息资源共享, 提出了更高的要求, 实验室信息化势在必行。为此, 必须在分析技术和数据管理两方面进行开拓和创新, 首先要研究制订地矿实验室元数据标准, 并支持电子资源的定位、发现、评估和选择等, 为地矿

实验室数据管理和共享提供技术支撑, 为国土资源信息化作出努力。

元数据就是“关于数据的数据”, 是专门用来描述数据的特征和属性, 说明数据内容、质量、条件、查询和其他有关特征的背景数据^[1]。最基本的用途是元素的使用和表现, 要对数据包括容量大、无法进行人工管理的数据信息资源提供描述、分类和管理, 从而实现查询、阅读、交换和共享^[2]。信息网络系统的建设、开发和应用, 已成为当今衡量实验室整体水平的重要标志。网络建设是信息系统运作和共享的物理平台, 实验室的数据接口是构建信息管理系统的主体, 在统一接口标准和网络环境下, 实验数据才能真正成为数据信息交换和共享的社会资源, 为科学研究

收稿日期: 2005-12-05; 修订日期: 2006-03-05

基金项目: 国土资源地质大调查项目(03D3-56)

作者简介: 吴晓军(1963-), 女, 北京市人, 高级工程师, 从事计算机应用工作。E-mail: wuxiaojun@cags.net.cn。

和生产服务。因此元数据标准的制定和编写是共享的基础,对于不同领域一般都会根据需求来定义一个标准或几个标准。美国材料实验学会(ASTM)发布的实验室信息管理系统(LIMS)标准 E1578、分析数据交换标准 E1947、色谱数据交换标准 E1948、质谱分析数据标准 E2077、LIMS 认证标准 E2066、IC/QC 质量管理标准等已广泛用于实验室的信息技术中^[3-4]。笔者以此作为本项工作的参考依据,严格执行《标准化工作导则》^[5]的规定,引用中国《地质矿产术语分类代码》^[6]《分析化学术语》^[7]等,并参照国际分析数据交换和信息存储标准(ADISS),结合地学科技发展和新一轮地质大调查需求,针对地质矿产实验室数字资源,充分考虑数据用户、数据字典、应用项目工程和元数据成本等四方面需要,研究了实验室各种元数据特征和要素,制订编写了地质矿产样品分析测试元数据内容与格式标准、相关术语代码、标准数据文件规范;建立了实验室元数据建库标准;设计了元数据库管理系统模型;开发了全面覆盖地质调查实验室的实验测试范围的元数据库管理软件;为地质大调查实验室数据管理和数据共享提供了平台,为实验室信息化建立了基础。

1 系统运行环境

本系统在 Microsoft Windows NT 4.0 服务器操作系统、J2SE Development Kit 5.0、Tomcat 5.5 的网络平台上,数据库选用 Microsoft Sql Server 7.0,客户端选用 Windows 2000、Windows XP 操作系统。

2 程序设计原理

2.1 结构设计及关键技术

客户机和服务器结构(C/S)与浏览器和服务器结构(B/S)是当今世界开发模式技术架构的两大主流技术。地质矿产实验室分析测试元数据库管理软件采用了当今最先进的浏览器和服务器结构。传统的 C/S 体系结构虽然采用的是开放模式,但在特定的应用中客户机和服务器端都还需要特定的软件支持。由于未能提供用户真正期望的开放环境,C/S 结构的软件需要针对不同的操作系统开发不同版本的软件,加之产品的更新换代十分快,已经很难适应百台电脑以上局域网用户同时使用,而且造价高、效率低。而 B/S 结构是对 C/S 结构的一种变化或者改进的结构。在这种结构下,用户工作界面是通过 WWW 浏览器来实现,极少部分事务逻辑在前端实现,主要事务逻辑在服务器端实现,形成了所谓的三层结构。其优点是

简化了客户端电脑载荷,减轻了系统维护与升级成本和工作量,降低了用户的总体成本。能实现一次性到位的开发,实现不同的人员,从不同的地点,以不同的接入方式访问和操作共同的数据库;有效地保护数据平台和管理访问权限,服务器数据库也很安全。特别是在 Java^[8]这样的跨平台语言出现之后,B/S架构管理软件更加方便、快捷、高效。

自 1999 年,Sun 公司推出了 JSP 动态网页技术以来,它被迅速和广泛地用于 Web 上的应用软件开发,并具有跨平台、安全、高效、面向对象的特点。可以允许 Web 开发人员和设计人员非常容易地创建和维护动态网页,作为 JavaTM 技术的一部分,JSP 能够快速开发出独立于平台的 Web 应用程序。JSP 是基于 Java 和 Servlet 技术,它是直接和服务器相关联的,因此速度快,开发难度也要小得多。本系统采用三层结构(图 1),前台采用 Servlet/JSP 技术设计,中间层使用 Tomcat 作为中间服务器,后台使用 Sql Server 作为数据库层^[9]。



图 1 三层结构示意图

Fig. 1 Schematic diagram of three-layer structure

Tomcat 因其技术先进、性能稳定,而且免费,是目前比较流行的 Web 应用服务器。因此 Tomcat 作为本系统 JSP 和 Servlet 的运行平台。

2.2 数据库设计

地质矿产实验室分析测试元数据库管理软件的主要关系如图 2。

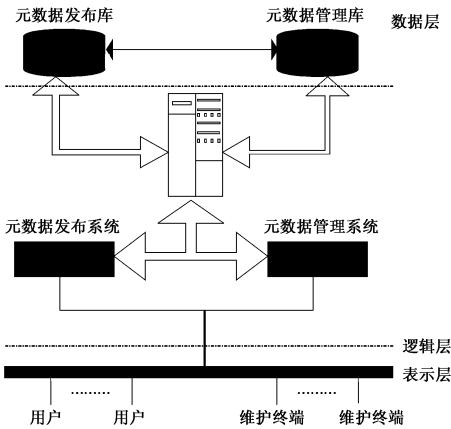


图 2 系统结构关系

Fig. 2 Schematic diagram of the system connection frame

2.3 软件功能模块设计

地质矿产实验室分析测试元数据库管理系统是以分析测试类型和分析项目为核心的静态数据库,实现自动导入的编辑功能;易于数据处理与转换,提高数据的检索能力;XML 和 OLE 技术的使用解决了N 层系统开发的打印问题。其软件功能模块见图3。

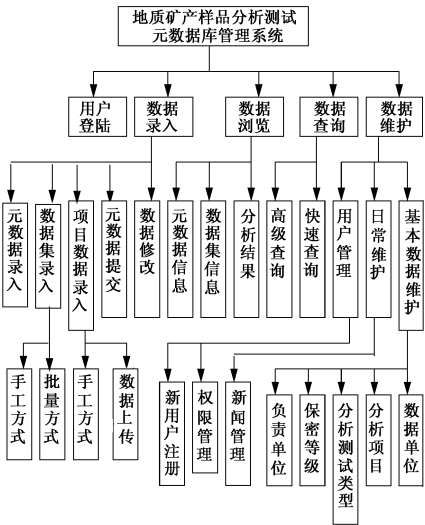


图3 软件功能模块

Fig. 3 The function modules of th software

3 软件的功能介绍

地质矿产实验室分析测试元数据库管理系统包括用户登录、数据录入、数据浏览和查询、数据维护等功能,并以网页形式出现,有最新新闻、文件下载、相关连接、快速查询等。系统的主界面见图4。

3.1 用户登录

本系统分维护级、工作级、一般用户级三个权限。权限不同,对系统访问的内容不同。维护级是最高级别,管理下属的工作级和一般用户级;工作级可以对元数据进行录入、修改;一般用户只可对系统进行浏览查询。

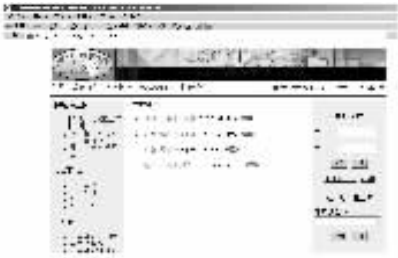


图4 地质矿产实验室分析测试元数据库管理系统主窗口

3.2 数据录入

数据录入:包括元数据、数据集、项目数据以及元数据的提交和数据修改。

元数据内容:包括数据集标识信息(基本信息、地理坐标范围、样品采集信息等,见图5)、数据集数据质量信息(项目编号、名称、类型质量控制等)、参照系统信息、数据集描述信息(数据集摘要、关键词、参考文献等)、元数据发行信息、参考信息等。

数据集内容:包括数据集名称、统一编号、项目编号、项目名称、岩石名称、分析测试类型、地质年代、样品处理、校正方法、分析单位、数据形式等信息(图6)。在实际工作中经常有大批量样品,为了减轻使用人员工作量,录入方式设计为单批和批量两种方式。

项目数据:包括项目名称、分析结果、数据单位。项目数据往往数据量很大,许多大型分析仪器所产生的分析结果均可以生成文本,进行整批的数据导入,因此开发了两种数据录入方式:手工录入(图7)、数据上传(图8)。数据上传后,还可根据需要对数据单位进行大批量、个别修改。避免了手工数据录入所产生的错误,方便了使用,提高了工作效率。

数据修改可对已录入的各数据集进行编辑、删除等操作。元数据提交后,在进行数据修改时,状态栏中即会显示“提交”(提交前为录入状态)。



图5 元数据的录入窗口

Fig. 5 The metadata input dialog window



图6 数据集的录入窗口

Fig. 6 The data input dialog window



图7 项目数据手工方式录入窗口

Fig.7 The dialog window for data input manually



图8 项目数据批量上传方式窗口

Fig.8 The file submit dialog in batch

3.3 数据浏览和数据查询

一般用户即可进行元数据详细信息和数据集信息的浏览和打印,但不能对数据进行任何修改。数据查询包括对最新的数据集进行单个查询;可以指定某个元数据的某个分析项目、统一编号或分析结果等进行快速查询;也可以按录入序号、项目名称或结果升降序进行组合的高级查询。总之可以按使用者的不同需求对元数据进行多方位的查询检索。数据浏览和查询窗口见图9。



图9 数据浏览和查询窗口

Fig.9 The dialog window for data browse and data query

3.4 数据维护

数据维护的实质是初始化工作,数据维护包括用户管理(设置用户信息和权限)、新闻管理(新闻发布及新闻管理)、基本数据维护(设置负责单位信息、保密等级、分析测试类型、分析项目、数据单位等)。在使用该系统前,必须完整地录入这些基本数据,才能有效地运行本系统。

4 结语

本系统目前已在国家地质实验测试中心的内部网站上运行,不久将在中国地质科学院网站上发布。使用者可以通过访问该网站,将实验室的元数据集成到元数据库管理系统进行统一管理,并通过元数据发布系统为使用者提供查询服务;使用者可以借助元数据深入了解数据集的各项特征,最终获取元数据的数据集;它为建立地质矿产分析测试元数据库共享平台,为各地质矿产实验室建立实验室信息管理系统,提供技术参考标准。同时,将地矿分析测试元数据库共享平台纳入地学科学数据共享平台,具有重要的意义。

5 参考文献

[1] 赵慧勤. 网络信息资源组织——元数据[J]. 情报理论与实践,2000,23(6):465-467.

[2] 数据挖掘研究所. <http://www.dmresearch.net/html/abstract/meta-data/1000000858.jsp>.

[3] ASTM. ASTM LIMS Guide. 1994 Annual Book of ASTM Standard, American Society for Testing and Materials (ASTM) [S]. 1994, Vol. 14. 01.

[4] Mary D Hinton. Laboratory Information Management System[M]. New York: Marcel Dekker Inc, 1995: 1-336.

[5] GB/T 1.1—2000, 标准化工作导则[S].

[6] GB 9649—88, 地质矿产术语分类代码[S].

[7] GB/T 14666—93, 分析化学术语[S].

[8] 於东军, 杨静宇, 李千目, 等. Java 程序设计与应用开发[M]. 北京: 清华大学出版社, 2005: 10-314.

[9] Stephen Wynkoop. SQL Server 6.5 开发使用手册[M]. 康博创作室译. 北京: 机械工业出版社, 1998: 1-219.