

国家馆实物地质资料信息一体化管理设计与实现

米胜信¹, 姚聿涛¹, 高志新², 王琳¹

(1. 国土资源实物地质资料中心, 河北 三河 065201; 2. 中国地质调查局发展研究中心, 北京 100037)

摘 要: 国家实物地质资料馆通过近些年工作的开展, 已经采集了全国范围内大量珍贵的实物地质资料, 并按照相关标准规范对每份实物资料进行了整理加工。如何对国家馆藏实物资料数字化资源进行有效管理, 并向社会提供服务是我们亟需解决的问题。本文主要阐述如何将国家馆整理的实物资料数据进行有效管理和整理入库, 实现实物资料信息的一体化管理并在互联网进行发布。

关键词: 国家实物地质资料馆; 实物地质资料; 一体化管理

中图分类号: G271 **文献标识码:** A **文章编号:** 1004-4051(2016)SI-0114-05

Integrated design and implementation of geological data in Cores and Samples Center of Land and Resources

MI Sheng-xin¹, YAO Yu-tao¹, GAO Zhi-xin², WANG Lin¹

(1. Cores and Samples Centre of Land and Resources, Sanhe 065201, China;

2. Development Research Center, China Geological Survey, Beijing 100037, China)

Abstract: Through the work in recent years, Cores and Samples Center of Land and Resources has collected a large amount of precious physical geological data. And all of the data has been processed according to relevant standards. The key problem is how to manage the data effectively and provide services to the community. This article mainly introduces how to organize and manage the kind of physical geological data storage in cores and samples center of land and resources, and achieve the integration of management of physical information to provide services through the Internet.

Key words: Cores and Samples Center of Land and Resources; physical geological data; integration of management

国家实物地质资料馆(国土资源实物地质资料中心, 以下简称国家馆)承担国家重要实物地质资料采集、管理、开发研究和利用, 为政府主管部门提供决策与业务技术支撑, 向社会提供公益性服务。国家馆采集实物地质资料, 均按照“典型性、代表性、特殊性和系统性”的原则, 由全国范围内进行筛选, 通过汇交和专项采集, 构建反映我国地质矿产条件和重大工作成果的国家级实物资料体系^[1]。

国家馆通过近几年的工作, 理清了实物地质资料信息管理和流程, 在此基础上利用信息化技术, 将整理数字化后形成的多种实物资料类型文件和数据进行了规范化设计, 实现了实物资料一体化管理, 并将实物资料信息在网上提供服务利用。

1 国家馆资源建设情况

1.1 国家馆资源情况

截至 2015 年 10 月, 国家馆共收藏实物地质资料 394 档, 主要包括 557 个钻孔、总计孔深 298274.24m 的 229134.00m 岩芯、岩屑 9109 件; 标本 15736 块、大标本 236 块; 薄片 40814 片、光片 64 片; 副样 389 件、煤样 106 件。

1.2 数据的获取

国家馆对采集的每档实物地质资料, 主要包括实物资料(岩芯、标本、薄片、岩屑、样品)和相关资料信息(项目成果报告、采集报告、勘探线图、剖面图、钻孔柱状图)组成。按照《实物地质资料建档工作指南》、《实物地质资料整理工作指南》和《实物地质资料数字化工作指南》中的要求, 将采集的实物地质资料进行整理。

1) 岩芯扫描: 对岩芯进行扫描、将岩芯分层描

收稿日期: 2016-03-21

述信息进行整理录入(通过岩芯扫描仪系统)生成岩芯扫描图像数据库和扫描文件。

2)标本照相:对标本的切面照相 2~6 张,然后对标本描述信息进行整理,生成 EXCEL 表格。

3)薄片显微照相:对薄片信息进行电子显微镜照相,对每个薄片的不同区域,不同角度和不同倍数,每个薄片照相 10 张左右,然后将显微照相结果进行描述,将显微图像和描述信息按照一定的存储结构进行标准化文件信息存储。

4)相关资料整理:将收集的所有纸质、电子资料进行整理,生成 PDF 或 DOC 格式的文档。

5)资料信息著录:通过实物资料业务管理系统,将整理的实物资料信息进行汇总,对每档实物地质资料进行著录,生成每档实物地质资料信息的基本信息表。

整理后生成的实物资料信息主要包括实物资料基本信息表、岩芯扫描图像信息、岩芯描述信息、标本图像信息、标本描述信息、薄片显微图像信息、薄片鉴定信息和单张图片信息、相关资料目录信息和文档扫描信息等。

1.3 数据的整理与组织

1.3.1 岩芯信息

岩芯图像信息是利用引进的 YXCJ-VZZ 型图像高分辨率岩芯扫描仪,对岩芯进行扫描,扫描后的单张岩芯图像扫描长度为 1 m(行业通用的岩芯盒长度为 1 m),宽度为 5~10 cm(岩芯直径),生成的单张岩芯图像的大小为 5~10 MB。岩芯描述信息由采集的电子版或者文本资料进行分层描述,通过岩芯扫描仪系统进行录入。最终将扫描后的岩芯描述信息和岩芯图像信息存储到 CoreDB_YXFB 数据库中,岩芯图像按照项目(档号)-钻孔编号-扫描图像的树形结构进行存储。

1.3.2 标本信息

标本信息主要由标本描述信息和标本图像信息组成,标本描述信息存储在 EXCEL 表格中,标本图像信息按照项目(档号)-标本编号-标本图像的树形结构进行存储。标本图像的存储格式主要为 TIF 类型,单张图像在 300 MB 左右。

1.3.3 薄片信息

薄片信息主要有薄片描述信息和薄片显微图像。薄片描述信息存储在 EXCEL 表格中,主要字段包括薄片的基本属性信息,如薄片编号、采集层位、结构构造、采集人等。单张薄片显微图像的存储结构为图幅-剖面-薄片-薄片图像+TXT 文件。TXT 文件除存储薄片的基本属性信息外,还存有光

性、放大倍数、显微镜下结构描述信息。薄片图像的存储格式主要为 JPG 类型,单张图像大小在 10 MB 左右。

1.3.4 相关资料

相关资料信息主要有目录信息和文件信息,目录信息包含相关资料的名称、类型和文件存储位置。文件信息多为 PDF 和 DOC 格式或 EXCEL 表格文件。

2 数据表设计

根据国家馆实物资料整理存储和实物资料信息关联情况,对国家馆实物资料信息进行了数据标准规范设计,实现实物资料的一体化管理,为实物资料信息的社会化网络服务提供基础。部分主要实物资料数据表字段设置如下所述。

2.1 实物地质资料基本信息数据构成

基本信息表作为项目的索引,反映该档实物资料的属性信息,包括项目信息、馆藏信息、汇交信息和实物信息等字段。

2.2 岩芯信息表

岩芯信息主要包括岩芯属性信息、岩芯图像信息和岩芯分层信息。岩芯属性信息包括岩芯所属钻孔的编号、位置、取样长度、见矿范围等;岩芯的图像信息主要包括岩芯深度和其所对应的图像;岩芯分层信息主要包括岩芯分层深度、岩石名称和岩性描述。

2.3 标本信息表

标本信息表主要存储标本的属性信息、标本采集信息、标本位置信息、地质描述信息和标本图像信息。

属性信息包括标本编号、所在项目(图幅)、剖面等;采集信息包括标本的采集单位、采集位置、采集时间等;地质描述信息包括标本的鉴定名称、结构构造、矿物成份等;图像信息包括图像的存储格式、存储位置。

2.4 薄片信息表

薄片信息表主要存储薄片的属性信息、薄片采集信息、地质描述信息,薄片显微图像和显微图像的属性信息。

属性信息包括薄片的编号、所在项目(图幅)、剖面等;采集信息包括薄片的采集单位、采集位置、采集时间等;地质描述信息包括薄片的鉴定名称、结构构造、肉眼描述信息;图像信息包括显微图像的存储格式、存储位置、镜像描述和显微镜操作信息。

2.5 相关资料信息表

相关资料信息表主要存储资料的属性信息、数量信息和数字化信息。属性字段包括资料名称、资

料类别;数量信息包括资料单位、数量;数字化信息包括资料的存储格式和存储位置。

3 系统的实现

3.1 系统设计

国家馆实物资料一体化的目的是在不改变国家馆实物资料信息化工作模式的前提下,对实物资料信息的管理,将分散在不同工作区域、不同类型的数据按照数据库的规范要求和信息发布要求进行处理并最终入库,从而轻松实现数据的安全管理和发布。

3.1.1 数据库设计

3.1.1.1 岩芯数据表数据库映射设计

岩芯信息表、岩芯影像信息表和岩芯描述信息表中某些字段信息由 CoreDB_YXFB 数据库中的 YTK_JH、YTK_YXPM 和 YTK_YXTX 的数据表字段中映射而来。

其中,YTX_JH 表中,JH 字段存储的类型为档号(6 位阿拉伯数字)+矿区名称(汉字)+钻孔编号(字母开头的英文字符)。

YTK_YXTX 表中,存储岩芯扫描的图像信息,QSSD 为起始深度,ZZSD 为终止深度,ZQM 为原

图名称和存储位置,ZQM2 为缩略图名称和存储位置。

YTK_YXPM 表中,存储岩芯钻孔的描述信息,QSSD 为起始深度,ZZSD 为终止深度,YX 为岩性,YXMS 为岩性描述。

在系统新建数据库的 Drill 数据表中,PKIIN 字段提取 JH 字段中的档号生成,DRILL_NO 提取 JH 字段中的钻孔编号生成。

岩芯描述表和岩芯图像表中的字段由 YTK_YXPM 和 YTK_YXTX 中的相应字段映射生成。

3.1.1.2 薄片数据关系设计

薄片数据表,主要有薄片属性和薄片显微图像信息数据构成。薄片属性信息的可以通过 EXCEL 表格中的字段提取。薄片显微照相数据表中,薄片的光性、放大倍数、镜下结构描述等字段需要从与单张显微图像对应的 TXT 文件中读取。

在 TXT 文件中,1~7 行的字段属性在薄片数据表中已经存在,8 行、9 行和 11 行字段内容分别对应薄片的光性、放大倍数、镜下结构描述信息,需要提取到数据库中(图 1)。

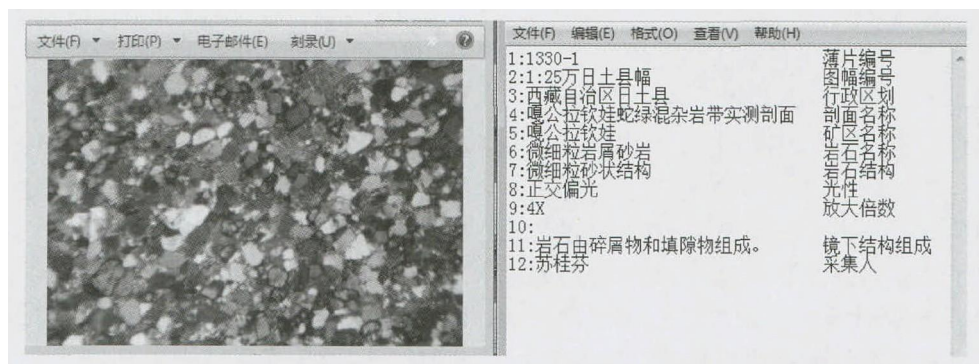


图 1 薄片显微图像 TXT 文件内容

在数据库设计时,首先将 TXT 文件名与薄片编号对应,与薄片信息表中的薄片编号匹配。然后读取该 TXT 文件的 8 行、9 行和 11 行字段中的内容,分别赋予薄片显微图像信息表中的光性、放大倍数和镜下结构描述信息。

3.1.1.3 标本数据关系设计

标本数据表主要有标本属性和标本图像信息构成。标本属性信息通过 EXCEL 表格的标本信息整理表中的对应字段提取,并将标本图像存储位置信息进行存储。

前文提到标本图像的格式多为 TIF 格式,TIF 格式下的图像数据量较大,单张标本图像大小为 300 MB 左右,需要将 TIF 格式的图像转换为 JPG

格式,以便于网络发布。在数据设计过程中,设置图像存储位置信息关联过程表,用于存储转换后的 JPG 格式的标本图像存储位置信息。

3.1.1.4 数据关系设计

根据国家馆实物资料数据规范要求,在数据库建立时设计表格之间的关系。其中,案卷基本信息表作为目录级数据的顶层,统领各文件级数据表。岩芯信息、标本信息、薄片信息和相关资料信息以档号为关键字,在案卷基本信息表下,按照档号为管理字段并列存储(图 2)。

3.1.2 功能设计

系统具备的主要功能有数据导入、数据编辑、数据查询、数据输出、图像处理和质量检查等。

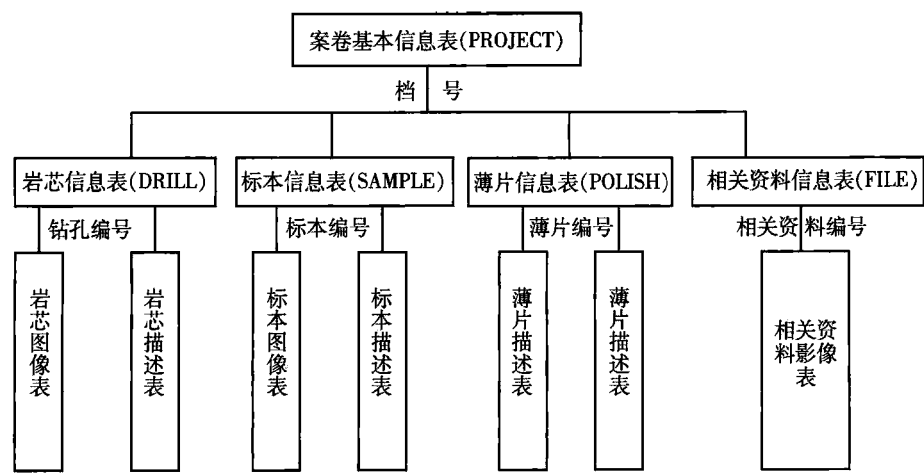


图 2 系统数据库总体结构图

- 1)数据导入:在资料加工过程中,基本信息表、标本信息表、薄片鉴定信息表和相关资料信息表信息均为 EXCEL 表格,因此系统具备数据导入的功能。
- 2)数据编辑:入库后的实物资料信息具备对某些信息进行处理和加工的功能。
- 3)数据查询:设定一定的查询条件,便于管理检索与利用。
- 4)数据输出:根据不同用户的需求,可以自定义导出不同格式、不同字段的数据表格和数据库,用于产品开发和数据服务。
- 5)图像处理:国家馆加工处理的实物资料图像,有些图像的格式和大小不符合实物资料网络发布的要求,因此要求系统在图像入库时能够对图像进行缩放、切片、格式转换等处理。

6)质量检查:在所有的实物资料入库后,对实物资料进行质量检查,包括必填项、数据格式、数据逻辑关系等方面的检查,并生成质量检查报告,对存在问题的数据进行错误提示和数据导出。

3.1.3 技术流程

系统在开发过程中,主要按照数据层、开发层、应用层三个层次进行设计,按照需求开发系统^[2],见图 3。

3.2 系统实现

系统具有高度可视化的编辑环境,所见即所得的用户界面,其主体功能界面实现对国家馆实物资料数据的管理、一体化编辑和质量检查等。

- 1)案卷信息管理。在案卷信息管理中,将实物地质资料业务管理系统导出的案卷信息表进行数据导入,保存到系统中。并可对导入的数据进行编辑、删除、按需导出等操作。
- 2)岩芯信息管理。岩芯信息管理,首先将

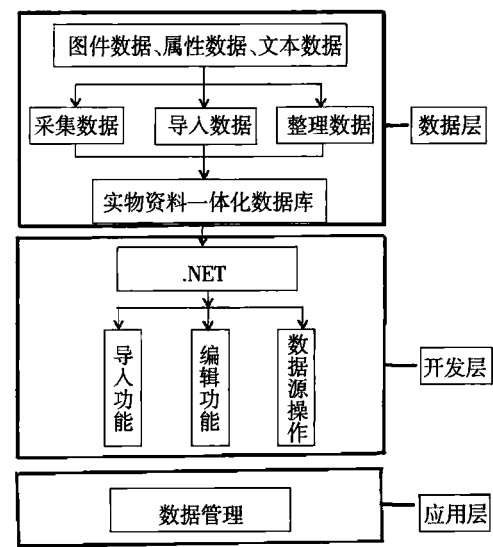


图 3 系统开发流程图

CoreDB_YXFB 数据库安装在同一服务器下,实现岩芯信息与 CoreDB_YXFB 中数据字段的自动映射,抽取相应的字段信息并插入到岩芯信息表、岩芯描述表和岩芯图像表中,然后将扫描整理后的图像信息存储在指定位置,通过系统实现对岩芯信息的导入、导出和数据质量检查。

3)薄片信息管理。首先将显微镜采集的图像信息按照标准的存储结构存放在系统指定位置,然后将薄片鉴定信息表导入到系统中,通过添加项目、添加剖面等不同方法,可以将薄片图像导入到系统中,在图像导入的同时,系统会对薄片图像进行压缩处理,达到网络查询和发布的要求(图 4)。通过系统实现薄片信息的导入、导出和数据质量检查。

4)标本信息管理。标本信息管理与薄片信息管理类似。区别在于标本的原始图片为 TIF 格式。在导入标本数据同时,系统将 TIF 格式转换为 JPG

格式,然后存储 JPG 格式的原始图后生成不同大小的缩略图。对导入的数据进行质量检查,并导出标本图像缺失和错误记录。

5)相关资料信息管理。相关资料信息的管理,首先将库藏管理室加工整理后的相关资料文件存

储在指定位置,然后将实物资料案卷级目录数据导入到系统中,并将导入后的相关资料,进行处理,生成 PDF、DOC 等不同格式的文件进行数据的输出等,实现相关资料的有序管理。

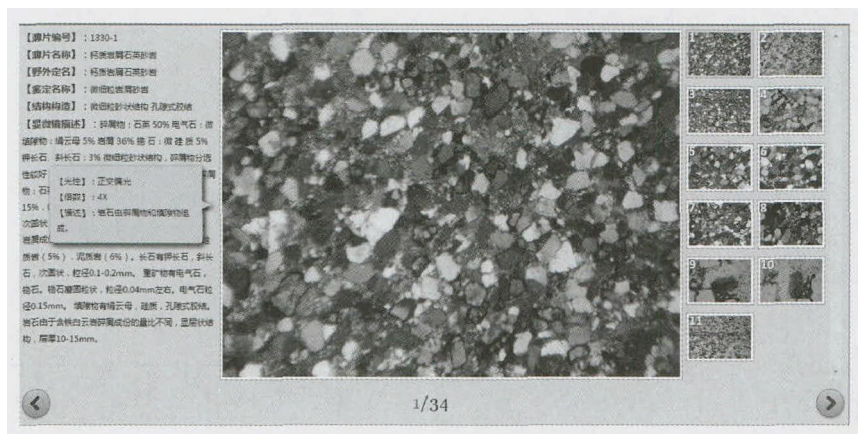


图 4 薄片信息浏览

3.3 前台信息发布

将一体化管理后的国家馆实物资料信息在中国实物地质资料信息网(www.cgsi.cn)发布,为用户提供国家馆资料信息的浏览、关键字查询、空间查询等功能。

1)空间查询。空间检索提供案卷(项目)信息的关键字、行政区划和实物类型等三种检索方式,并将检索后的案卷(项目)以列表和地图显示的方式展示,点击列表记录或者地图气泡,均可以浏览案卷详细信息页。

2)案卷信息浏览。单个案卷信息页面中,通过点击相应信息可以浏览岩芯信息、标本信息和薄片信息。

3)岩芯信息浏览。通过岩芯信息浏览页面,可以查看岩芯扫描图像、岩芯描述和其他相关信息。在页面显示过程中,利用图像延时加载技术即 Lazy-load 加载岩芯图像信息,使图像加载十分顺畅,未出现卡顿现象;利用 Ruler 控件设置岩芯深度标尺,作为岩芯图像和岩性描述信息深度对比的中间组件,使岩芯图像信息和描述信息同步显示^[3]。

4)标本信息浏览。在标本信息浏览页面,可以查询标本的属性信息和标本图像信息。点击标本信息页面中的箭头按钮,可以查询项目内所有的标本信息。

5)薄片信息浏览。在薄片信息浏览页面,可以查看薄片的属性信息和薄片显微照相信息。图 4 左侧为薄片属性信息,黄色气泡为单张薄片显微照相的属性,右侧 11 张图片分别为该薄片的不同区域和不同倍数的镜下显微图像,点击箭头按钮,查看该项目下所有薄片信息。

4 结 论

通过一体化的设计,实现了国家馆实物资料信息便捷、高效、长期的管理。系统将国家馆采集整理的 300 多档实物资料中的 500 多个钻孔,1.3 万多片薄片、8000 多块标本和相关资料信息进行了有效管理和网络发布。不仅提高了工作效率,规范了实物资料管理流程,并且为实物资料信息共享服务和专题服务产品开发打下坚实的坚实的基础。

参考文献

- [1] 刘晓文. 实物地质资料属性特点、重要性及研究意义[J]. 中国国土资源经济, 2006(8): 23-25.
- [2] 龚建辉. 基于面向对象的 GIS 制图数据与建库数据一体化系统的设计与实现[J]. 地理信息世界, 2012(5): 68-74.
- [3] 米胜信. 基于 WebService 技术的岩芯图像显示研究[J]. 中国矿业, 2015, 24(S1): 500-504.