

煤炭地质企业大数据平台建设研究

苏均生

中国煤炭地质总局, 北京

收稿日期: 2023年5月14日; 录用日期: 2023年6月12日; 发布日期: 2023年6月20日

摘要

在日常的生产运营过程中, 煤炭地质企业产生了大量的钻探地质、水文地质、物探地质、化探地质、遥感测试等多类数据, 这些海量数据以纸质方式存放于档案馆, 部分保存于电子介质中, 用于日常阅览, 结构性较差, 地质数据检索及三维可视化分析效率较低。为加强数据资源的管理、共享应用和深度开发, 需建设煤炭地质大数据平台, 创新数据管理和应用方式, 通过数据采集、汇聚、建模、共享与可视化应用, 发挥数据要素化作用, 赋能煤炭地质企业高质量发展。

关键词

煤炭地质, 大数据, 数据建模, 数据可视化

The Research of Big Data Platform Construction for Coal Geology Enterprise

Junsheng Su

China Coal Geography Corporation, Beijing

Received: May 14th, 2023; accepted: Jun. 12th, 2023; published: Jun. 20th, 2023

Abstract

During the daily operations, a large number of data about drilling, hydrogeology, geophysics, geochemical prospecting and remote sensing is produced by coal geology enterprise. These data are saved in the archives in the way of paper, and some data stored in electric media are used to read. The constitutive property of the data is sufficient, the geological data retrieval and the 3D visualization analysis are not enough. In order to strengthen management, shared application and deep development for data resource, big data platform for coal geology enterprise is needed. Through data acquisition, collection, modeling, sharing and visualization application, data element transformation is produced and coal geology enterprise can be empowered to realize high-quality de-

velopment.

Keywords

Coal Geology, Big Data, Data Modeling, Data Visualization

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

经过多年的资源勘探开发和项目建设,煤炭地质企业已形成了一大批具有丰富业务知识和地质项目资料数据,包括结构化数据和非结构化数据,这些数据大部分以纸质文件形式零散存在,有的存在于本单位档案室,有的存在于项目人员个人电脑,有的存在于工程现场信息管理系统中,这些数据蕴藏着非常宝贵的地质规律信息,通过结构化处理、运算建模和分析挖掘,可以被用于辅助资源勘探、矿业开发和生态修复,提高生产作业效率和管理决策能力。

当前我国煤炭地质行业大数据平台建设和应用一直处于探索阶段,部分单位建设了数据平台,主要偏向于技术资料的汇集与数据共享,实现信息化档案管理,支持资料借阅和查询,但很多关键信息未从地质资料中提取,数据三维建模及可视化不足。本文通过构建煤炭地质大数据平台,对各类数据进行采集、整合和管理,以矿区为单位进行地质体三维建模与可视化应用,将进一步挖掘各类煤炭地质数据的价值,提升数据资产化进程,支撑企业高质量发展。

2. 大数据管理思路

煤炭地质大数据涉及钻探地质、水文地质、物探化探、遥感测绘、野外作业等多种类型,数据分散存储于不同的介质中,为推进煤炭地质数据的统一入湖、交换共享、流通交易及分析应用,需要从源头上加强数据管理,深入开展数据治理,提升数据资产质量,构建科学规范的大数据体系。结合煤炭地质企业的特点,可从数据预处理、基础数据整理、业务数据采集、数据建模分析等方面着手研究,以大数据平台为载体,逐步实现数据资产化。

2.1. 数据预处理

收集已有地理、地质、物探、化探、遥感、钻探、采矿等资料,包含栅格、矢量、文本等多种数据形式,涵盖原始数据(野外地质填图路线、产状等)、过程数据(密度三维数据体、磁化率三维数据体等)、成果数据(地形图、地质图、钻孔、勘探线剖面图等)等一系列数据。以项目为重点,开展对现有地质资料的扫描、编排、归类,必须统一所有数据的坐标系和比例尺,建立原始资料库。

2.2. 基础数据整理

基础数据是基于地图的资料数据检索、数据分析建模的重要支撑,主要包括矿区信息、项目信息、钻孔信息、剖面信息、地形信息等数据。其中项目信息把项目名称、矿区、项目负责人、开始时间、结束时间等属性,钻孔信息包括钻孔编号、坐标系、开孔坐标、开孔方位、开孔倾角、开孔日期、中控日期、施工单位等。基础数据需要从原始资料中进行选取,确保基础数据与资料数据的一致性和关联性,为钻孔建模、层状建模、构造建模、岩性建模等提供数据支撑。

2.3. 业务数据采集

业务数据采集前, 需要对主数据及业务数据进行统一的元模型定义, 建立采集校验规则和质量评估规则。数据采集时, 应针对不同类型的数据采取不同的采集方式, 确保数据入湖。对存量的以电子化资料为基础的数据, 可由人工录入每类数据所对应的项目、矿区及省市信息, 通过文件上传遥感测绘、电磁法、物探、化探、重力异常、资源储量等数据的方式进行收集[1]。对于增量的以信息系统为基础的业务数据, 可通过系统集成的方式, 获取数字矿山、智慧生态、智慧工地等平台提供的归档数据。

企业应建立统一的数据管理部门, 安排专人专岗对所有入湖的数据资源进行统一管理, 建立健全数据管理制度, 对数据采集、审核、发布、流通、交易、作废等所有环节进行管控。完善数据安全防护机制, 强化分级分类管理, 推进数据存储、调用、分析、共享等全过程的安全管理, 加强数据备份。建立数据质量校验机制, 推进数据基于校验规则、核验机制统一采集入库, 并定期核对和评估分析, 强化数据治理, 推进数据管理提升。

2.4. 数据建模分析

为有效应对煤炭地质数据类型多样、规模及复杂等特点, 可通过分布式计算技术来实现高效的数据处理[2], 采用流式计算、批处理计算、图计算、内存计算等技术进行大数据计算, 基于分布式文件系统、列式数据库、NoSQL 数据库等技术进行存储[3], 对各类数据进行清洗、转换等处理, 确保数据规范。对于业务类数据, 用表征矿床中矿、岩体、地层和构造的空间分布和相应空间位置的属性数据, 利用专业软件在勘探线位置进行剖面数据圈定, 在此基础上根据钻孔信息建立矿床地质构造模型, 如矿体模型、岩体模型、构造模型[4] [5], 从钻孔数据库中提取地层顶、底板点, 然后利用空间插值方法进行空间插值[6], 形成三维地层的顶、底板面, 最后将地层顶底板合并, 形成三维地层实体模型, 是矿岩圈定、矿量品位计算和设计、计划的基础。数据建模后, 可以按线、矩形、面进行剖切, 也可以进行特效渲染, 还可以按岩性、层状、钻孔、构造等进行分类展示, 实现数据模型的可视化。

3. 大数据平台架构设计

煤炭地质大数据平台架构主要分为硬件设施层、数据层、平台层和应用层, 在统一的安全体系和标准体系支撑下提供数据共享应用服务, 如图 1 所示。

1) 硬件设施层: 硬件设施为数据存储和软件运行提供必须的算力、存储、网络等硬件基础(见图 2)。通过云计算虚拟化技术对物理服务器实现资源池化, 数据处理系统运行在云平台虚拟机上, 大大提高了服务器利用率。采用分布式存储系统针对各类应用需求提供相应系列的硬件节点类型, 以满足不同应用对性能、IO 和容量的多样化需求。采用物理分散的分布式数据库系统来存储和管理海量的数据。服务器节点之间采用光交换机连接, 高速光纤确保网络畅通。配备防火墙、IPS、漏洞扫描、堡垒机、WAF、数据库审计、日志审计等安全产品, 保障各业务系统的安全稳定持续运行。配备两路市电和一路柴油机供电, 确保云服务持续稳定。配备电信、移动、联通等多路宽带, 确保互联网运行畅通无阻。

2) 数据层: 通过数据预算处理、业务数据采集、大数据计算、数据建模等过程, 将钻探地质、水文地质、物探化探、遥感测绘等各类数据通过一定的规则汇聚至统一的数据湖, 通过数据治理形成数据仓库, 基于数据仓库提供大数据分析展现, 形成各种基于数据模型的分析图形, 支撑企业决策分析与生产管理。数据层包括元数据、主数据、业务数据、主题数据等各类类型的数据。

3) 平台层: 为数据层与应用层的中间交互枢纽, 为各类信息系统的稳定运行、集成共享等提供支撑, 包括集成服务平台、服务资源池、应用支撑平台、业务流程平台、通用组件平台等内容, 支持将各类平台资源以服务形式对外提供, 为大数据应用提供技术及数据支撑。

4) 应用层：面向各种最终用户，提供三维可视化分析、企业经营决策分析、地质灾害监测等信息，支持进行煤炭资源潜力评价、资源勘查、矿业开发、生产监测预警、矿井灾害治理、应急救援、远程会商决策等应用。

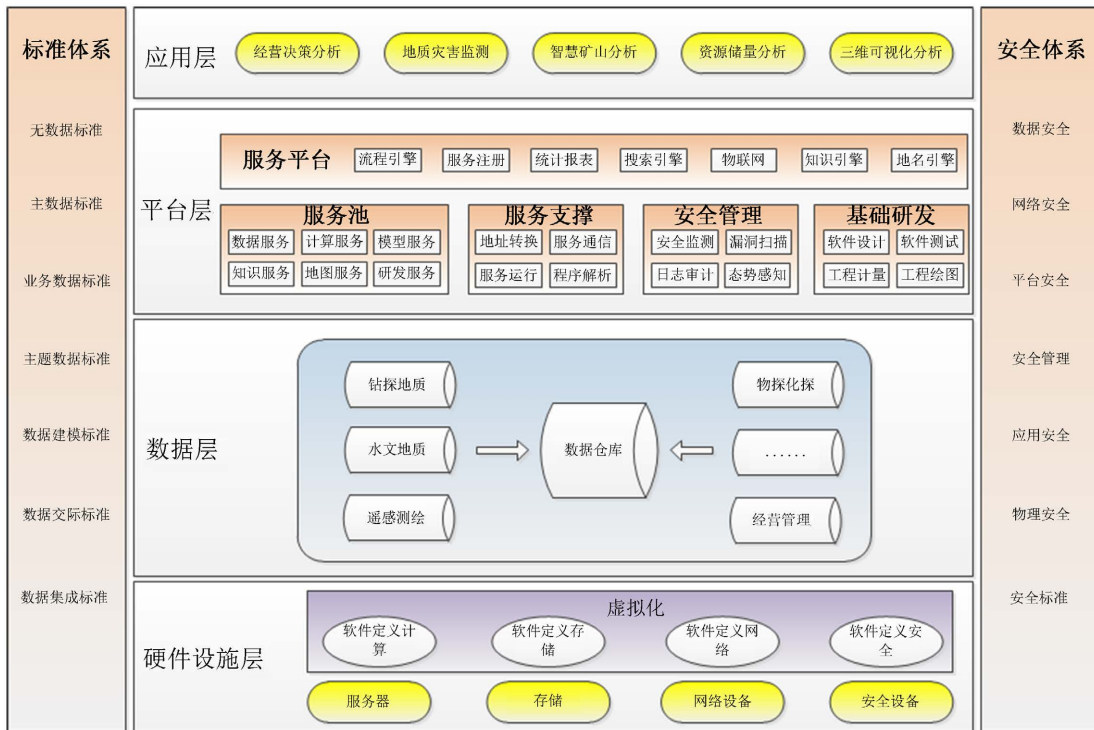


Figure 1. Diagram of the framework of the big data platform for coal geology enterprise
图 1. 煤炭地质企业大数据平台架构图

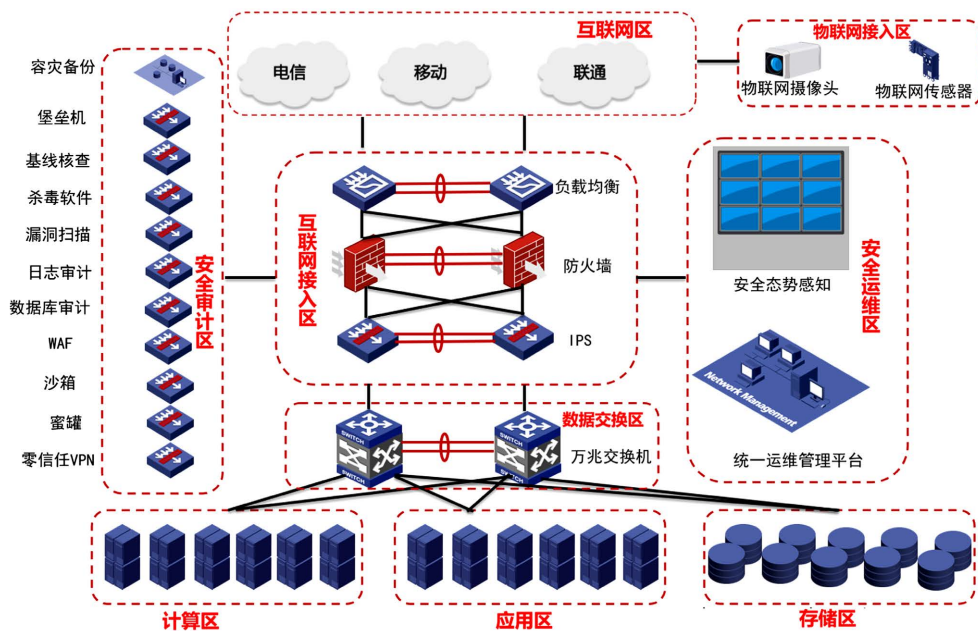


Figure 2. Diagram of the network infrastructure for hardware
图 2. 硬件设施网络架构图

4. 煤炭地质大数据平台关键技术

煤炭地质大数据平台建设过程中涉及的关键技术包括数据存储技术、多模态的数据处理技术、数据可视化技术等。

4.1. 地质数据存储技术

对各种不同来源的数据采集并处理后,需要进行灵活可扩展的存储,存储框架具有可扩展性和容错能力,需要高效的存储管理方法。Hadoop 分布式文件系统、基于分布式文件系统的 NoSQL 数据库、分布式数据库 HBase、关系型数据库、对象存储系统和时序数据库等都适合捕捉和存储海量数据[7]。针对不同的应用场景产生的结构化数据、时序设备监测数据、非结构化的勘探资料、遥感测绘数据、水文地质数据等,一方面通过分布式文件系统、对象存储系统和时序数据库等多种存储框架进行综合运营,另一方面可通过元数据定义主数据、业务数据的语义结构,实现对数据的跨界连接、组织融合和协同查询[7]。

4.2. 地质数据处理技术

基于 GPU 集群的分布式计算、自然语言处理、预测分析、基于神经网络的深度学习等先进技术均可应用于地质大数据处理中。地质大数据应根据不同的应用场景运用不同的分析处理技术,如批处理、迭代处理、流处理、交互式分析等[8],可通过分布式文件系统,如 NoSQL 数据库或 Hive、Pig、Spark 等实现离线批量处理或迭代处理,对各个系统产生的海量生产或过程数据进行回归、分类、排序、协同过滤等。另一方面,可通过 Storm、Spark Streaming 等技术实现实时在线处理。

4.3. 地质数据可视化技术

可通过建模软件进行数据层状建模、钻孔建模、构造建模及岩性建模,通过模型叠加方式实现三维地质数据可视化,利用数据库数据建立覆盖层、夹层、透镜体、岩层、岩脉、断层、溶洞等不规则地质体。可使用测绘的数据生成地质体三维地面模型。利用勘探的钻孔、平硐、出露地质点、历史勘探剖面、CAD 地质图形、地层分界面点、地层分界面迹线和产状等数据建立地质年代层、岩性层、风化层、断层等地质分界面,创建地下水、卸荷界面[9]。

5. 煤炭地质大数据平台实现及应用

本文以某区域地质数据为研究对象,结合某煤炭地质企业地质资料管理、建模及可视化应用需求对大数据平台进行实现,该平台采用模块化的微服务架构,基于 java 技术进行实现,服务器操作系统为 Centos7.2,数据库采用 MySQL、MangDB 等进行存储,使用 3Dmine 软件进行三维地质数据建模,三维地质模型发布到大数据平台上实现可视化应用。

5.1. 地质数据管理

对于结构化的钻孔数据、储量数据等,通过数据录入的方式实现数据存储(见图 3)。钻孔地质数据库一般包括:定位表和测斜表、地层岩性表、地层表、化验表(样品化验)、构造表等各种表格。对于遥感测绘、水文地质、物探化探、剖面地下等专业数据,可通过文件上传的方式,通过大数据平台的程序接口实现在线浏览。

5.2. 三维数据模型

根据地质地形图、三维建模软件通过对等值线的提取、三维赋值,可以形成具有实际高程的三维地面模型,如图 4 所示。依据钻孔、地质剖面等数据进行层状建模,实现矿区各地层结构三维建模,如图 5 所

示。依据钻孔、地质剖面、地球物理数据等创建矿区岩性三维模型，实现矿区各地层的岩性分布三维建模。依据矿区地质调查构造信息及地球物理探测数据，通过人工干预的方法绘制出矿区构造三维模型，实现矿区相关构造的三维建模。

* 矿区 :	<input type="text"/>	项目编号 :	<input type="text"/>	项目名称 :	<input type="text"/>
* 钻孔编号 :	<input type="text"/>	* 坐标系统 :	—请选择—	* 开孔坐标E :	<input type="text"/>
* 开孔坐标N :	<input type="text"/>	* 开孔坐标W :	<input type="text"/>	* 最大孔深 :	<input type="text"/>
* 轨迹类型 :	—请选择—	* 是否见矿 :	—请选择—	* 勘探线号 :	<input type="text"/>
* 开孔方位 :	<input type="text"/>	* 开孔倾角 :	<input type="text"/>	* 开孔日期 :	<input type="text"/>
* 终孔日期 :	<input type="text"/>	钻进方法 :	<input type="text"/>	工作阶段 :	<input type="text"/>
* 施工单位 :	<input type="text"/>				

深度 (m) 方位角 (°) 倾角 (°) 测量方法

+ 新增 保存 删除

Figure 3. Diagram of the geological data input interface

图 3. 地质数据录入界面示意图

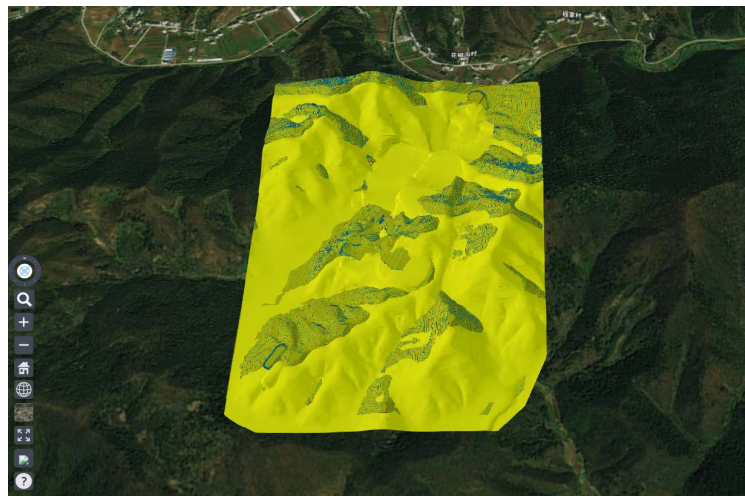


Figure 4. Diagram of the 3D geological ground model

图 4. 三维地质地面模型图



Figure 5. Diagram of the 3D geological layered model

图 5. 三维地质层状模型图

5.3. 地质资料检索

基于 GIS 展示全国各地的煤矿资源分布情况,支持通过模糊搜索(煤田、矿区、市区名等)到相关煤矿的所有历史资料情况,通过类别(预查、普查、详查、勘探、拟建、新建、在建、生产等)、企业性质(国有企业、股份合作、股份有限、有限责任、私营合伙、私营独资、集体企业、私营企业等)、产能、时间区间高级检索和定位,定位的地区在地图上圈出。

5.4. 数据可视化应用

在传统方式下,煤炭地质勘查研究主要采用平面图、剖面图及表格等二维地质资料来表达地下三维空间,信息检索和地质分析效率低,直观性不足。通过对试点区域相关数据的采集、处理、建模和可视化应用,对地质体进行三维可视化应用,能够完整准确地表达了复杂地质现象的边界条件及地质体内含的各种地质构造,直观地再现地质单元的空间展布及其相互关系,最大限度地提高区域地质分析的直观性和准确性。

6. 煤炭地质大数据平台的发展方向与意义

6.1. 煤炭地质大数据平台发展方向

随着人工智能、数字孪生、区块链等数字技术的不断发展,数据处理能力不断提升,地质勘查技术与数字技术的结合度更高,数据在推动煤炭地质行业转型发展方面的作用将更加突出,未来煤炭地质企业大数据平台将朝着智能化方面不断发展。

- 地质勘查技术与大数据平台的融合应用:通过大数据平台提供的数据服务,为特定区域地质勘探开发提供地质数据、水文数据、遥感测绘数据等支撑,便于精准发现区域资源分布情况、资源储量情况、地质条件情况等信息,提升找矿的精准性。
- 大数据平台与人工智能技术的融合应用:通过人工智能在大数据平台的应用,便于将水文、物探化探、遥感测绘等信息与现有三维地质模型进行融合,完善层状、构造、岩性等三维模型,便于解决岩性识别等煤炭地质行业的难点。
- 大数据平台与数字孪生技术的融合应用:通过数字孪生技术,将物理勘探开发与虚拟的三维数字模型进行关联,便于支撑管理人员通过平台监控勘探开发工作推进情况,同时将勘探开发的信息传递给三维数字模型,动态更新三维模型信息。

6.2. 煤炭地质大数据平台建设意义

- 加强地质资料信息化管理:通过对各类钻孔数据、地质档案数据、钻探数据、物探化探等数据的统一管理,将纸质档案转化为电子档案来进行永久性的存储,使原始资料得到很好的保存,极大降低了物理破坏的可能性。
- 助力企业决策分析:通过数据可视化,助力企业借助大数据技术从煤炭储量、矿区潜力等不同维度对企业地质数据资产进行综合计算、分析,为能源决策、煤炭资源规划和开发、生态修复提供了丰富详实的科学依据。
- 提升数据服务创新能力:通过数据共享、流通、交易和挖掘服务,有利于发挥数据资产化作用,提升数据互联互通能力与创新发展,为企业生产管理、技术创新、资产化应用等提供支撑,以充分释放地质大数据红利。
- 提升业务工作效率:使作业人员可以通过对三维地质结构模型的灵活运用,更加快速、形象、直观的完成对区域内地质结构各类数据的系统性认识;节省大量传统图纸、图件内相关信息的整合、分析时间,更加方便、快捷地获得预期数据样本与结论。

7. 煤炭地质大数据平台实施建议

煤炭地质大数据平台建设是一个动态的、不断丰富完善的过程，涉及多类数据和多个模型。企业要结合自身实际，遵循“循序渐进、试点推进、安全可靠、逐步深入、创新应用、赋能产业”的原则进行建设实施，同时注意以下几个方面：

- 做好总体规划与架构设计。企业应从长远角度规划设计大数据平台，着力从业务、应用、数据、基础设施、安全等多方面进行详细设计，梳理应用场景和需求，制定发展蓝图与目标，明确建设思路与内容，制定建设路线图与计划。
- 重视大数据体系建设与治理。企业应建立平台管理制度和标准体系，明确管理部门与职责，深入开展数据分级分类管理，注重数据采集、编码、维护、校验、审核、检查、评估等过程管理，加强数据安全防护，推进构建科学完善的大数据体系。
- 不断推进数据服务化。企业应以大数据与经营生产业务结合为出发点，建立数据共享、分析、评价、交易、发布等机制，推动数据在内外共享应用，深入数据资源开发和分析应用，发掘数据价值，加快数据互联互通与资产化步伐。

8. 结束语

煤炭地质大数据平台通过对各类数据的收集、处理、管理、建模与可视化应用，将大数据、三维建模等技术与煤炭地质业务进行结合，有利于提升煤炭地质企业数据管理、数据服务经营生产业务、促进技术创新等能力，加快煤炭地质企业经营方式转变和生产管理智能化步伐，支撑企业高质量发展。

参考文献

- [1] 柳波. 地质大数据采集与分析系统的配置技术研究[J]. 自然资源信息化, 2022(1): 12-18.
- [2] 王议, 孙伟, 马冬梅, 张志鹏, 何培雍. 基于3D管理系统的煤炭矿山三维建模及实践: 以陕北某煤矿为例[J]. 中国矿业, 2016, 25(7): 148-152.
- [3] 王华. 基于大数据框架的数据存储与管理[J]. 信息记录材料, 2022, 23(8): 189-191.
- [4] 翁正平. 复杂地质体三维模型快速构建及更新技术研究[D]: [博士学位论文]. 武汉: 中国地质大学, 2013.
- [5] 曹婧. 面向地质建模的数据管理技术研究及应用[D]: [硕士学位论文]. 青岛: 中国石油大学(华东), 2017.
- [6] 陈方吾. 边坡三维地质体快速建模及可视化系统研发[D]: [硕士学位论文]. 成都: 成都理工大学, 2020.
- [7] 谭章禄, 马营营, 袁彗. 煤炭大数据平台建设的关键技术及管理协同框架[J]. 工况自动化, 2018, 44(6): 16-20.
- [8] 丁富东. 大数据平台的关键技术及组网方案[J]. 电信科学, 2015(4): 158-163.
- [9] 北京华创汇翔科技有限公司. 三维地质建模浅谈[EB/OL]. https://www.hchxtech.com/page278?article_id=62, 2022-07-11.