

海洋观测资料共享服务平台设计与实现

陈志勇¹ 赵明月¹ 乔佟¹ 王征²

1. 航天规划设计集团有限公司 北京 100000

2. 外交部服务中心 北京 100000

摘要: 海洋观测作为洞察海洋环境变化、助力海洋资源开发的关键手段,其数据的有效共享与利用至关重要。本文设计并实现了海洋观测资料共享服务平台,该平台具备多源数据采集接入、混合存储管理、多维度查询共享及智能分析可视化等功能。采用分层架构与模块化设计,兼顾性能、安全与可扩展性,通过严格权限控制与隐私保护技术,为海洋领域各用户提供高效、安全的数据服务。

关键词: 海洋观测资料共享服务平台;设计;实现

引言:海洋占据地球表面积的绝大部分,其环境变化深刻影响着人类社会的发展与生态平衡。海洋观测能获取水温、盐度、海流等关键数据,为海洋研究、资源开发、灾害预警等提供依据。然而,目前海洋观测数据存在分散、格式不统一、共享困难等问题。为解决这些痛点,设计并实现海洋观测资料共享服务平台十分必要,它能整合多源数据,提供便捷查询与高效分析功能,推动海洋领域的数据共享与应用发展。

1 海洋观测资料共享服务平台需求分析

1.1 功能需求

(1) 数据采集与接入:支持卫星、浮标、船舶、岸基雷达等多源海洋观测设备的数据接入,能自动识别不同设备的数据格式,实现实时采集与批量导入,确保数据覆盖水温、盐度、海浪、海流等核心观测指标,且具备异常数据初步过滤功能。(2) 数据存储与管理:采用混合存储架构,对结构化数据(如观测站点信息、实时监测数值)用关系型数据库存储,对非结构化数据(如卫星遥感图像、船舶观测日志)用分布式文件系统存储;同时提供数据版本管理、备份与恢复功能,保障数据完整性与可追溯性。(3) 数据查询与共享:构建元数据目录,支持用户按观测时间、区域、设备类型等多维度精准查询;提供标准化API接口,方便第三方平台调用共享数据;设置数据访问权限分级,平衡数据共享便利性与安全性。(4) 数据分析与可视化:集成时空分析工具,可对观测数据进行趋势分析、异常值检测、空间插值等操作;通过折线图、热力图、海洋流场图等多种可视化形式展示数据,帮助用户直观理解海洋观测规律,辅助决策制定。(5) 用户权限与安全控制:划分管理员、数据提供者、普通用户等不同角色,分配差异化操作权限;对敏感数据(如军事海域观测数据)进行脱敏

处理,限制非授权访问,同时记录用户数据操作行为,确保数据使用合规^[1]。

1.2 非功能需求

(1) 性能需求:单条数据查询响应时间不超过3秒,批量数据导出响应时间不超过30秒;支持至少500名用户同时在线操作,峰值时段并发处理能力不低于1000次/分钟,保障平台在数据高峰期稳定运行。(2) 安全性需求:采用传输层加密(SSL/TLS)与存储加密技术,防止数据在传输和存储过程中泄露;建立完善的审计日志系统,记录数据访问、修改、删除等操作,日志保留时间不少于1年,便于安全事件追溯与责任认定。(3) 可扩展性需求:采用模块化设计,支持新增观测设备接入模块、数据分析算法模块等,满足未来功能扩展需求;兼容公有云、私有云等多种部署模式,可根据数据量增长弹性扩展存储与计算资源,适应业务规模扩大。

2 海洋观测资料共享服务平台总体设计

2.1 架构设计

(1) 分层架构:采用四层垂直架构,各层职责清晰且协同联动。①数据层:负责多源海洋观测数据的集中存储与管理,包含关系型数据库(存储结构化数据)、分布式文件系统(存储非结构化数据)及缓存数据库(提升高频查询效率),同时集成数据备份与容灾模块,保障数据安全;②服务层:作为核心业务支撑层,封装数据采集、清洗、分析、共享等服务能力,通过服务注册与发现机制,实现模块间灵活调用,支持横向扩展;③应用层:面向用户实际需求,提供数据查询、可视化分析、权限管理等功能模块,支持Web端、移动端多终端适配,满足不同场景使用需求;④用户层:针对管理员、数据提供者、普通用户等不同角色,设计差异化操作界面与权限入口,确保用户高效获取所需服务。

(2) 技术选型: 采用微服务架构拆分核心业务模块, 将数据采集、存储、分析等功能拆分为独立微服务, 降低模块耦合度, 便于单独升级与维护; 部署层面采用 Docker 容器化技术封装应用与依赖环境, 结合 Kubernetes (K8s) 实现容器编排, 支持自动扩缩容、故障自愈, 提升平台部署效率与稳定性; 存储层面选用分布式存储技术, 如 Ceph 分布式文件系统, 适配非结构化数据的海量存储需求, 同时搭配 MySQL 集群存储结构化数据, 平衡存储性能与成本^[2]。

2.2 数据模型设计

(1) 元数据标准: 基于 ISO19115 地理信息元数据标准进行扩展, 补充海洋观测领域专属元数据项, 包括观测设备类型 (卫星/浮标/船舶)、观测指标 (水温/盐度/海浪)、数据采集频率、数据质量等级、数据归属机构等, 形成统一的海洋观测元数据规范, 确保元数据的完整性与互操作性, 支撑跨平台数据检索与共享。(2) 数据分类与编码规则: 按数据来源与类型进行二级分类, 一级分类分为卫星观测数据、浮标观测数据、船舶观测数据、岸基雷达数据四大类; 二级分类细化至具体数据形式, 如卫星观测数据下分遥感图像数据、遥感反演数据等。编码采用“分类码-时间码-设备码-序号”16位编码规则, 例如“SB-20240501-B01-001”代表2024年5月1日由编号B01的浮标采集的第001条数据, 实现数据唯一标识与快速溯源。

2.3 接口设计

(1) 内部接口: 聚焦数据处理全流程协同, 设计数据清洗接口与数据转换接口。数据清洗接口对接数据采集模块, 通过预设规则 (如阈值过滤、异常值替换) 剔除无效数据, 输出标准化洁净数据; 数据转换接口支持不同格式数据 (如 CSV、NetCDF、JSON) 的相互转换, 将非结构化数据 (如卫星图像) 转换为可分析的结构化数据, 保障各模块间数据格式兼容。(2) 外部接口: 采用标准化接口设计, 提升平台开放性与兼容性。
① RESTful API 接口: 遵循 REST 设计规范, 提供数据查询、导出、共享等接口服务, 支持 JSON 数据格式传输, 方便第三方应用 (如海洋科研平台、海事管理系统) 调用; ② OGC 标准服务接口: 集成 WMS (Web 地图服务)、WFS (Web 要素服务), 支持将海洋观测数据与地理空间信息叠加展示, 满足用户对数据空间分布的可视化需求, 同时兼容主流 GIS 软件 (如 ArcGIS、QGIS), 实现跨系统数据交互。

3 海洋观测资料共享服务平台核心模块实现

3.1 数据采集与预处理模块

(1) 多源异构数据接入: 针对数据库 (MySQL、PostgreSQL)、文件 (CSV、NetCDF、TIFF)、实时流 (Kafka 消息队列) 等不同来源数据, 设计适配性接入适配器。数据库接入通过 JDBC/ODBC 驱动建立连接, 支持定时增量同步与全量同步; 文件接入采用 FTP/SFTP 协议接收浮标、船舶上传文件, 结合文件解析器自动识别格式并提取数据; 实时流接入通过 Kafka 消费者订阅卫星、雷达实时观测数据流, 实现毫秒级数据接收, 保障多源数据高效汇聚。(2) 数据清洗与质量控制的算法实现: 采用“规则过滤+智能检测”双重算法。规则过滤通过预设阈值 (如水温合理范围 $-2^{\circ}\text{C}\sim 30^{\circ}\text{C}$) 剔除明显异常值, 结合时间连续性校验 (如相邻数据差值超阈值标记为可疑); 智能检测引入孤立森林算法识别非线性异常数据, 采用卡尔曼滤波补全缺失数据, 同时通过 Shapiro-Wilk 正态性检验评估数据分布合理性, 输出质量等级 (优/良/差), 确保预处理后数据准确率达 95% 以上。

3.2 数据存储与管理模块

(1) 分布式文件系统与关系型数据库结合: 采用 HDFS 存储卫星遥感图像、船舶观测日志等非结构化/半结构化数据, 利用其高容错性与可扩展性支撑 PB 级数据存储; MySQL 集群存储结构化数据 (如观测站点信息、元数据、用户信息), 通过主从复制实现读写分离, 提升查询效率。同时构建数据映射机制, 将 HDFS 文件路径与 MySQL 结构化数据关联, 实现跨存储介质数据联动查询。(2) 时空数据索引优化: 针对海洋观测数据的时空特性, 采用 R-Tree 索引优化空间查询, 快速定位特定海域 (如经纬度范围) 的数据; 对时间维度数据采用 B+ 树索引, 提升按时间区间检索效率; 对高频更新的实时数据, 结合四叉树索引实现多尺度空间分割, 平衡检索速度与索引维护成本, 使时空联合查询响应时间缩短至 2 秒内^[3]。

3.3 数据共享与权限控制模块

(1) 基于角色的访问控制 (RBAC) 模型: 设计“用户-角色-权限”三层架构, 预设管理员 (全权限)、数据审核员 (数据审核/发布)、普通用户 (数据查询/下载) 等角色, 支持自定义角色与权限组合。通过权限矩阵定义各角色对数据的操作权限 (如查看/编辑/导出), 结合会话管理实现权限动态生效, 确保不同角色仅能访问授权范围内数据。(2) 动态水印与数据脱敏技术: 对导出数据嵌入动态水印, 水印内容包含用户 ID、操作时间等信息, 支持肉眼不可见的隐形水印 (防止截图盗用) 与可见半透明水印 (明确数据归属); 对敏感数据 (如军事海域坐标、船舶编号) 采用脱敏处理, 通过字符替换 (如“辽渔 3XXX”)、范围模糊化

(如经纬度精确到 0.1°)等方式,在保留数据可用性的同时保障信息安全。

3.4 可视化与分析模块

(1)前端框架实现交互式地图:采用Leaflet构建轻量化交互式地图,支持海洋基础底图加载、观测站点标记、数据图层叠加(如海浪高度热力图),实现地图缩放、平移、测距等操作;结合ECharts实现多维度数据可视化,如用折线图展示水温时序变化、用散点图呈现浮标分布,支持图表联动(如点击地图站点显示对应数据图表),提升数据直观性。(2)后端分析引擎:基于Spark构建分布式计算引擎,实现海量数据并行分析(如区域海洋温度均值计算、异常数据聚类);集成Python科学计算库(NumPy、Pandas、Scikit-learn),提供趋势预测(如ARIMA模型预测海浪变化)、相关性分析(如水温与盐度关联分析)等算法服务,支持用户自定义分析参数,输出分析报告与可视化结果。

4 海洋观测资料共享服务平台安全与隐私保护

4.1 数据传输安全

采用HTTPS/TLS1.3协议构建全链路加密通信通道,覆盖数据采集端(卫星、浮标、船舶)与平台、平台内部模块间、平台与用户终端的所有数据传输场景。通过证书验证机制确保通信双方身份合法性,防止中间人攻击;对传输数据进行对称加密(AES-256算法)与消息摘要校验(SHA-256算法),既避免数据被窃取篡改,又能快速识别传输过程中的数据完整性问题,保障多源观测数据在传输环节的安全

4.2 存储安全

基于透明数据加密(TDE)技术,对数据库(MySQL集群)与分布式文件系统(HDFS)中的数据进行实时加密。数据库层面,TDE加密数据文件与备份文件,密钥独立存储于专用密钥管理系统(KMS);分布式存储层面,采用块级加密方式处理非结构化数据,加密密钥按数据类型分级管理。同时,KMS支持密钥生命周期管理(生成、轮换、销毁)与访问权限控制,定期自动轮换密

钥,防止密钥泄露导致的存储数据安全风险。

4.3 访问控制

集成单点登录(SSO)系统,用户通过统一账号密码完成一次认证后,即可访问平台内授权的所有模块,减少账号管理复杂度,同时避免多系统重复登录带来的安全隐患。对外提供数据共享服务时,采用OAuth2.0协议实现第三方应用授权,支持授权码、密码等多种授权模式,严格控制第三方应用对平台数据的访问范围与操作权限,授权过程全程记录日志,确保访问行为可追溯。

4.4 隐私保护技术

在海洋数据处理中应用k-匿名化技术,对包含用户/机构标识的观测数据(如船舶编号、岸基站点归属)进行处理,确保每个数据分组中至少包含k个不区分个体的记录,避免单一主体被识别。同时,引入差分隐私技术,在数据查询、统计分析结果中添加微小噪声(基于拉普拉斯机制),既保证统计结果的可用性,又防止攻击者通过多次查询反推原始敏感数据,平衡海洋数据共享价值与隐私保护需求。

结束语

海洋观测资料共享服务平台的设计与实现,为海洋数据的整合与高效利用开辟了新路径。通过多源数据接入、混合存储及精准查询等功能,满足了海洋科研与管理多元需求。在平台运行中,严格的安全与隐私保护机制保障了数据安全。未来,平台将紧跟技术发展,持续优化功能,提升数据服务能力,为海洋事业发展提供更有力的数据支撑,助力海洋强国建设。

参考文献

- [1]刘晓丹,赵迎春,丁磊.基于SSH框架的海洋调查资料信息管理系统的设计和实现[J].海洋科学,2022,(03):31-32.
- [2]孙朝辉,李兆钦,刘增宏.智慧海洋国际共享应用平台发展趋势及其对策建议[J].海洋开发与管理,2020,(07):72-73.
- [3]张皓.面向智慧海洋的海洋信息管理服务平台设计与实现[J].科技创新导报,2020,(05):60-61.