

石油企业电子信息技术数据治理的关键技术研究

李子玉¹ 张全¹ 车越²

1. 天津瑞洋海洋工程有限公司 天津 300000

2. 天津北海油人力资源咨询服务有限公司 天津 300000

摘要: 石油企业电子信息技术数据治理关键技术研究聚焦数据孤岛、质量参差、安全合规等痛点,通过构建统一数据湖与主数据管理体系打破数据壁垒;采用AI清洗标注、多模态处理技术提升数据质量;应用隐私计算、区块链技术强化数据安全;依托边缘计算与工业互联网平台实现实时数据处理。研究还涵盖数据资产化、可视化及跨部门协同机制,助力企业向智能化、绿色化转型。

关键词: 石油企业; 电子信息技术; 数据治理; 关键技术

引言

在数字化转型浪潮下,石油企业电子信息技术数据治理已成为提升竞争力的核心要素。面对勘探开发、生产运输等环节产生的海量异构数据,传统治理模式面临数据孤岛、质量参差、安全合规风险等挑战。本研究聚焦数据全生命周期管理,通过整合物联网、AI、区块链等前沿技术,构建覆盖数据采集、存储、分析、共享的智能化治理体系,助力石油企业实现数据资产化运营与价值最大化,为行业高质量发展提供技术支撑。

1 石油企业电子信息技术数据治理现状与挑战

1.1 行业现状

(1) 石油企业已实现勘探、生产、运输等核心业务环节的信息化覆盖,各类专业系统(如勘探数据管理系统、生产执行系统)已成为业务运转的关键支撑。但由于早期系统建设缺乏统一规划,各系统多为独立部署,形成“数据孤岛”,导致跨环节、跨部门数据流通性差,难以形成全产业链数据协同价值。(2) 数据类型呈现多元化特征,既包含勘探参数、生产指标等结构化数据,也涵盖地质报告、设备运维文档、视频监控等非结构化数据。但数据质量参差不齐,一方面中文专业语料稀缺,制约AI技术在数据处理中的应用;另一方面数据标注缺乏统一标准,不同系统数据格式不兼容,增加了数据整合难度。(3) 数据合规压力持续增大。随着《数据安全法》《个人信息保护法》等法规落地实施,叠加石油行业部分业务涉及跨境合作,数据跨境流动审核要求趋严;同时,企业内部员工信息、客户数据等隐私信息保护需求提升,对数据全生命周期合规管理提出更高要求^[1]。

1.2 核心挑战

(1) 数据质量管控难度大。石油勘探开发等环节数据采集环境复杂,易产生噪声数据、缺失数据;多源异

构数据来源分散、标准不一,数据融合过程中易出现语义冲突,难以形成高质量数据资产。(2) 技术整合协同不足。当前石油企业虽引入物联网、云计算、AI等新技术,但各技术应用多处于单点突破阶段,缺乏系统性整合。例如,物联网数据采集与云计算平台数据存储、AI数据分析衔接不畅,技术优势难以充分发挥,未能形成全流程数据治理技术体系。(3) 安全合规技术应用滞后。数据分类分级标准尚未完全落地,对核心数据、敏感数据的识别与管控不够精准;匿名化处理、隐私计算等前沿合规技术应用不足,难以在保障数据安全的前提下实现数据价值挖掘。(4) 组织架构支撑薄弱。多数石油企业缺乏专职数据治理角色,如首席数据官(CDO)、数据保护官(DPO),数据治理责任分散;跨部门数据协作机制不完善,业务、IT、合规等部门沟通壁垒明显,导致数据治理效率低下。

2 石油企业电子信息技术数据治理的关键技术框架与解决方案

2.1 高质量数据集构建技术

(1) 数据清洗与标注:依托AI技术实现数据治理提质增效,采用DataAgent等智能代理工具,通过机器学习算法自动识别勘探数据中的异常波动、生产数据中的缺失字段等问题,完成噪声数据过滤与精准修正,同步实现数据的自动分类分级。针对石油行业多模态数据并存的特点,构建融合文本(地质勘察报告、钻井日志)、图像(岩心薄片影像、设备巡检照片)、音视频(现场作业监控、设备运行声响)的一体化处理模型,结合行业术语库完成实体抽取与关系映射,搭建全链条知识图谱,实现数据语义关联与高效检索,为后续分析应用奠定基础。(2) 数据标准化:以业务场景为导向,采用ER图建模方法构建统一数据模型,明确勘探、生产、运输等各

环节的数据实体、属性及关联规则，消除跨系统数据语义歧义。严格对标DCMM国家标准，建立覆盖数据全生命周期的企业级数据标准体系，规范数据编码规则、格式要求、元数据描述等核心要素，形成“采集-存储-处理-应用”全流程标准化规范。同时建立标准动态更新机制，适配业务升级与技术迭代需求，保障数据“同源、同标、同态”，破解跨部门数据流通壁垒^[2]。（3）场景适配性优化：针对油田勘探野外复杂环境、生产环节连续作业、运输管网广域覆盖等差异化场景，定制数据采集与处理方案。例如在勘探场景，部署抗干扰、低功耗采集设备，优化地震波数据、地质样本数据的采集精度与传输稳定性；在生产场景，基于设备运行周期设计数据过滤规则，优先保留关键工况参数，提升实时处理效率；在运输场景，适配管网分布式布局，构建边缘计算节点，实现数据就近处理与本地预警，降低传输延迟。

2.2 AI驱动的智能治理技术

（1）AI原生治理架构：打破传统治理模式的技术割裂问题，构建“AI×数据×流程”一体化智能治理平台，实现业务流、数据流、智能流深度协同。平台集成数据采集、清洗、分析、应用全环节功能，通过AI算法自动优化治理流程，动态适配业务需求变化。典型案例为胜利油田，其依托该架构整合油井生产、设备运维等全维度数据，通过深度学习模型实现生产动态实时模拟，提前72小时识别设备故障、产量波动等异常情况，预警准确率达92%以上，显著提升生产管控效能。（2）自动化治理工具：部署智能元数据管理系统，通过自然语言处理技术自动识别数据来源、格式、业务含义等元数据信息，构建动态更新的元数据知识库；应用基于机器学习的自动分类分级工具，结合行业数据特征库精准识别核心数据、敏感数据，实现分级标记与权限自动关联；引入联邦学习、安全多方计算等隐私计算技术，在不泄露原始数据的前提下，实现多部门、多企业数据协同分析，平衡数据价值挖掘与安全保护需求。（3）多模态数据处理：聚焦石油行业非结构化数据占比高的痛点，利用AI技术突破传统处理瓶颈。通过计算机视觉技术解析地震勘探三维图像，快速识别储层分布、断层构造等关键信息；采用自然语言处理技术提取钻井报告、设备运维文档中的核心参数与故障描述，转化为结构化数据；结合语音识别与语义理解技术，处理现场作业语音记录，实现数据实时录入与分析，将数据处理周期缩短60%以上，大幅提升决策效率^[3]。

2.3 合规与安全保障技术

（1）数据分类分级：建立多维度数据分类分级体系，

结合石油行业特点，从业务敏感度、数据价值、泄露影响等维度，将数据划分为核心数据（如勘探开发核心技术参数、战略规划数据）、敏感数据（如员工隐私信息、客户商业数据）、一般数据三个等级。针对不同等级实施差异化管控策略，核心数据采用“多重认证+加密存储+全程审计”模式，敏感数据实施访问权限最小化管理，一般数据简化流程提升使用效率，实现安全与效率的平衡。（2）隐私计算与匿名化：规模化应用差分隐私、同态加密等前沿技术，在数据共享与分析过程中保障隐私安全。通过差分隐私技术在原始数据中加入可控噪声，确保数据统计特性不变的同时隐藏个体信息；采用同态加密技术实现加密状态下的数据运算，避免原始数据泄露；对涉及个人信息的数据实施假名化、去标识化处理，去除身份识别字段，确保数据使用符合《个人信息保护法》等法规要求。（3）可信数据空间：构建企业级可信数据共享平台，集成区块链、加密存储等技术，实现数据确权、访问控制、审计追踪全流程闭环管理。通过区块链技术记录数据产生、传输、使用等全环节操作，确保流转过程不可篡改、可追溯；建立数据确权机制，明确数据权属主体与使用范围，为跨部门、跨企业数据共享提供信任基础；搭载智能审计模块，实时监控数据操作行为，自动识别违规操作并触发预警^[4]。（4）合规流程嵌入：将合规要求贯穿数据全生命周期，在数据产品设计阶段开展合规风险评估，明确数据使用边界与安全要求；在数据采集环节嵌入合规校验规则，自动过滤违规数据；在模型训练阶段建立数据合规审核机制，杜绝使用不合规数据训练模型；在数据应用环节实施动态合规监控，确保数据使用符合法律法规与企业制度，实现“事前预防、事中控制、事后追溯”的全流程合规管理。

2.4 云原生与物联网支撑技术

（1）混合云架构：采用“私有云+公共云”混合云部署模式，优化资源配置与安全保障。将勘探开发核心数据、敏感数据部署在私有云环境，通过专属网络与严格的安全防护措施保障数据安全；将办公协同、行业资讯等非核心业务数据部署在公共云，利用公共云的弹性扩展能力降低IT运维成本。搭建云间数据安全流转通道，通过加密传输、访问控制等技术实现跨云数据安全共享，提升资源利用效率。（2）物联网实时采集：构建全场景物联网数据采集网络，在井场、管网、炼化厂区、运输车队等关键场景，大规模部署传感器、RFID标签、视频监控、无人机等设备，实现油井产量、设备运行参数、管网压力、环境指标等数据的实时采集。通过边缘计算节点对采集数据进行预处理，过滤无效数据、提取关键信

息,降低传输压力,确保数据传输的实时性与准确性,为生产监控与智能调度提供数据支撑^[5]。(3)实时数据库(RTDB):部署高性能实时数据库,具备海量实时数据的高速存储、检索与分析能力。针对石油生产场景中数据吞吐量高、延迟要求低的特点,优化数据库存储结构与索引算法,支持每秒百万级数据写入与毫秒级检索响应。通过实时数据库整合各环节实时数据,为生产过程监控、设备故障预警、产量动态调度等业务场景提供高效数据服务,保障生产连续稳定运行。

3 石油企业电子信息技术数据治理未来展望与建议

3.1 技术趋势

(1)图数据库与知识图谱技术普及,支撑复杂关联数据分析。未来石油企业将广泛应用图数据库技术,高效关联勘探、生产、运输全链条多维度数据,结合行业知识图谱挖掘地质构造、设备运维与供应链间的复杂关系。通过可视化分析精准定位数据关联逻辑,为勘探方案优化、设备故障根因排查等提供精准支撑,突破传统结构化数据分析的局限。(2)AI大模型与数据治理深度融合,实现自动化治理与价值挖掘。AI大模型将全面渗透数据治理全流程,替代人工完成数据清洗、标注、分类分级等工作,大幅提升治理效率与精度。同时依托大模型语义理解能力,深度挖掘多模态数据价值,例如分析海量钻井日志生成勘探优化建议,推动数据治理从“合规保障”向“价值驱动”转型。(3)区块链技术应用于数据确权与审计,增强信任机制。区块链的去中心化、不可篡改特性将有效解决石油行业跨企业、跨部门数据确权难题,明确数据权属关系。同时构建全生命周期审计追踪体系,实时记录数据流转全环节操作,确保数据可追溯、可监管,强化数据安全合规保障能力,为数据共享奠定信任基础。

3.2 实施建议

(1)企业层面:设立CDO/DPO角色,完善数据治理组织架构。石油企业需建立专职数据治理岗位,明确CDO统筹数据战略与价值挖掘,DPO聚焦安全合规管理,

厘清各部门职责边界。建立跨业务、IT、合规的协同机制,定期推进数据治理工作,破解部门壁垒,推动治理体系化、常态化开展。(2)行业层面:推动数据标准统一,建立行业级数据共享平台。由行业协会牵头,联合头部企业制定统一数据标准,规范采集、编码等关键环节,消除企业间数据“语言壁垒”。搭建行业共享平台,依托隐私计算实现数据“可用不可见”,整合优质数据资源,提升全行业治理与创新能力。(3)政策层面:加强数据安全法规宣贯,鼓励技术创新与试点应用。相关部门需针对性开展法规培训,提升企业合规意识。出台专项扶持政策,设立创新基金,支持企业与科研机构联合攻关核心技术;搭建试点平台推广优秀案例,以点带面推动全行业数据治理高质量发展。

结束语

石油企业电子信息技术数据治理是驱动行业数字化转型的关键引擎。本研究通过构建AI驱动的智能治理框架、融合隐私计算与区块链的安全防护体系,有效破解了数据孤岛、质量管控与合规风险等核心难题。未来需持续深化技术融合创新,强化跨部门协同与行业标准建设,推动数据治理向场景化、价值化方向演进,为石油企业提升运营效率、实现绿色低碳发展提供坚实的数据支撑与决策保障。

参考文献

- [1]李晓燕.大数据时代下计算机电子信息处理技术研究[J].数字通信世界,2021,(04):100-101.
- [2]秦伟光.电子信息技术在网络安全中的应用[J].数字技术与应用,2021,39(05):199-201.
- [3]王瑞珩.电子信息技术在网络安全中的应用分析[J].网络安全技术与应用,2020,No.238(10):154-156.
- [4]孔莉,孔凡华,杨各.计算机网络技术在电子信息工程中的应用探讨[J].中国设备工程,2022(10):213-215.
- [5]李永浪.自动化技术在电子信息工程中的应用分析[J].新型工业化,2022,12(04):22-26.