DOI 10.15302/J-SSCAE-2025.12.009

大模型驱动的水利数据流通体系与建设路径研究

王刚[1]，黄浩轩[1]，刘钰[2]，贺新武[3]，刘晓亮[4]，张威威[1*]

（1. 湖南工商大学管理科学与工程学院，长沙 410205；2. 中国电建集团中南勘测设计研究院有限公司，长沙 410014；
3. 中国电建集团华东勘测设计研究院有限公司，杭州 311100；4. 中南大学商学院，长沙 410083）

摘要：随着人工智能技术的快速发展，大模型已成为推动各行业数字化转型的重要驱动力。在水利领域，数据的高效流通对于提升水资源管理能力、优化水利工程运行、增强防灾减灾能力具有重要意义。但大模型赋能水利数据流通面临安全脆弱、协同机制不健全、时效性不足、模型可解释性欠缺等问题。本文基于大模型技术，从水利数据共享程度低、监测能力薄弱、数据质量低、分析能力有限、数据价值挖掘不足、运维成本高的业务痛点出发，基于开放性、可拓展性、安全防护能力、标准统一四项原则，构建涵盖多源采集、混合存储、智能治理、敏捷服务、场景应用及内生安全的六层技术架构。系统剖析政策法规、标准规范、技术平台、人才梯队与试点示范五大维度的耦合机制与协同需求，构建面向大模型驱动的水利数据市场化、安全化、智能化流通的建设路径，为推进水利数据市场化配置和高效流通提供了理论支撑和实践指导。建立基于大模型的水利数据流通技术体系能够显著提升数据的可用性、可解释性与流通效率，驱动水利行业实现高质量数字化转型，为智慧水利体系的构建提供坚实的理论与技术基础。

关键词：水利数据；大模型技术；多模态融合；数字孪生流域；数据流通

中图分类号：TV213.4；TD82　　**文献标识码：**A

Large Language Model-Enabled Circulation of Water Conservancy Data: Architecture and Development Path

Wang Gang[1], Huang Haoxuan[1], Liu Yu[2], He Xinwu[3], Liu Xiaoliang[4], Zhang Weiwei[1*]

(1. School of Management Science and Engineering, Hunan University of Technology and Business, Changsha 410205, China;
2. PowerChina Zhongnan Engineering Corporation Limited, Changsha 410014, China; 3. PowerChina Huadong Engineering Corporation Limited, Hangzhou 311100, China; 4. Business School, Central South University, Changsha 410083, China)

Abstract: The rapid advancement of artificial intelligence, particularly large language models (LLMs), has become a central catalyst for the digital transformation across multiple sectors. In the water conservancy field, the efficient circulation of data elements is critical for advancing water resource management, optimizing infrastructure operations, and enhancing disaster resilience. However, current LLM-enabled data circulation faces multiple challenges, including security vulnerabilities, weak cross-domain coordination, latency in data processing, and limited interpretability. Leveraging LLM technologies, this study starts by identifying critical challenges in the water conservancy domain, which include inadequate data sharing, subpar monitoring capabilities, poor data quality,

收稿日期：2025-12-10；**修回日期：**2026-02-04

通讯作者：*张威威，湖南工商大学管理科学与工程学院副教授，研究方向为水资源管理与决策；E-mail: 3022@hutb.edu.cn

资助项目：中国工程院咨询项目“健全现代化基础设施建设体制机制研究”(2024-JJZD-03)；国家自然科学基金项目(52408364)；湖南省自然科学基金项目(2024JJ6188)

本刊网址：sscae.engineering.org.cn

constrained analytical abilities, underutilized data values, and elevated operational and maintenance expenses. Guided by the principles of open access, elastic scalability, full-chain security assurance, and standards-based interoperability, this study proposes a six-layer, data-centric architecture, encompassing multi-source data acquisition, hybrid storage, intelligent data governance, agile service orchestration, scenario-driven applications, and endogenous security. Furthermore, it analyzes the coupling mechanisms and collaborative requirements across five systemic dimensions: regulatory frameworks, data and model standards, platform technologies, talent ecosystems, and pilot-scale experimentation. Building on these insights, a development pathway is proposed for constructing an LLM-enabled ecosystem for the market-oriented, secure, and intelligent circulation of water conservancy data, providing theoretical support and practical guidance for facilitating the market-oriented allocation and efficient circulation of water conservancy data. The proposed architecture can significantly enhance data usability, interpretability, and circulation efficiency, thereby propelling the water conservancy sector toward high-quality digital transformation, and providing a solid theoretical and technical foundation for the establishment of a smart water conservancy system.

Keywords: water conservancy data; large language model technology; multi-modal fusion; digital-twin watershed; data circulation

一、前言

随着数字经济的快速发展，数据已成为基础性资源和关键性生产要素，并广泛融入到经济社会发展的各个环节[1-2]。面向“国家水网–数字孪生流域”建设的双轮驱动需求[3-4]，国家推动水利数据流通，核心是将水利数据作为生产要素，在保障安全的前提下，通过市场化、智能化等手段，实现跨部门、跨层级、跨地域的高效共享和有序流动，以支撑国家水网和数字孪生水利建设。水利部高度重视水利数据建设及其在水利高质量发展中的赋能作用，并将“加快建设现代化水网”和“构建数字孪生水利体系”并列为“十五五”水利发展的核心战略任务之一。

在水利行业履行职责过程中制作或获取的各类非涉密数据资源，覆盖其归集、存储、共享、应用、安全等全生命周期的管理活动。通过打破“数据孤岛”，提升决策、调度和服务的效率与精准度，最终服务于国家水安全战略，因此水利数据的建设和流通已迫在眉睫[5]。《中共中央 国务院关于构建数据基础制度更好发挥数据要素作用的意见》明确提出建立数据可信流通体系[6]。水利部发布的《关于大力推进智慧水利建设的指导意见》明确提出，要基于全国水利“一张图”推进水利行业跨业务、跨层级数据调用以及与其他行业数据共享，实现水利信息有序共享[7]。同时，《数字孪生流域共建共享管理办法（试行）》强调，需统一技术标准，确保水利数据在统一规范框架下实现跨区域、跨层级、跨系统的规范整合与有序流通。

然而，部–省–市–县四级水利基础设施每天产生拍字节级监测、调度、遥感数据，因标准不一、权属不清、收益机制缺位而沉淀在“孤岛”里，流通率不足5%，难以有效支撑极端天气下“预报、预警、预演、预案”的防灾减灾业务、水资源优化配置与精细调度及水生态保护修复。数据流通的深层障碍源于权责界定模糊、标准规范缺失、激励补偿机制不足等制度性困境，以及由此衍生的“数据即权力”的部门本位主义。大模型在隐私计算、数据脱敏及自动化确权上的技术突破，为解决这些问题提供了全新路径[8-9]，也为化解“数据共享意愿不足、数据确权机制缺失、数据流通能力薄弱等问题”提供了关键的技术底座与安全屏障。

目前，学界在大模型与水利行业融合的相关研究也取得了一些进展[10-11]。大模型的数据分析能力、模拟预测能力、逻辑推理能力和智能决策能力，为水利工程领域的发展注入强劲动力[12]。大模型通过整合多源数据，利用其预测能力，能够为水利工程中的水资源调度优化和风险评估提供精准的管理技术支持[13]。依托大模型强大的通识知识库与自然语言理解能力，并结合领域微调技术，可以将海量的历史水文资料、实时监测数据、专业文献与行业规范作为“教材”，对基础大模型进行深度训练。催生出专业的“水文专家大模型”，使其不仅能实时解读复杂的流域水情数据，还能模拟资深专家的决策思维，为防汛抗旱、水资源精准调度与应急响应提供过去难以实现的、高度智能化的决策支持[14]。

本文通过构建“水利大模型+机理模型”双层驱动体系，可在保障数据所有权与安全的前提下，把碎片化数据转化为可计算、可推理的数字商品，实现“工程–数据–知识–价值”闭环。同时提出大模型驱动的水利数据流通体系与建设路径，立足于解决水利数据流通中的标准壁垒与价值实现难题，系统构建了大模型驱动的“机理–技术–路径”全栈框架。

二、水利数据流通现状与面临的主要挑战

（一）水利数据流通现状

水利数据是指在水利活动中产生、收集、处理、存储和使用的各类数据资源，根据数据来源和性质，主要分为基础数据、监测数据、工程数据、管理数据、服务数据等[15]（见表1）。当这些资源作为生产要素参与流通与应用时，呈现出价值性、稀缺性、强时效性和高关联性四大特征。价值性是指水利数据蕴含着丰富的信息，能够为水利管理决策提供重要支撑，是水利行业数字化转型的重要基础。稀缺性是指高质量的水利数据获取成本高，需要大量的人力、物力和财力投入，因此具有较高的经济价值。时效性是指水利数据具有很强的时效性，特别是实时监测数据，其价值会随着时间的推移而快速衰减，需要及时处理和应用[16-17]。关联性是指不同水利数据之间存在复杂的关联关系，如降雨与径流、水位与流量等，上述关联关系是进行综合分析的重要基础[18-19]。

表1 水利数据分类体系

一级分类	二级分类	数据内容说明
基础数据	地形地貌数据	地形高程、地貌类型、坡度坡向等
	河流水系数据	河流网络、湖泊水库、水系分布等
	行政区划数据	省、市、县、乡镇等行政边界
	地质条件数据	地质构造、岩性分布、地下水等
监测数据	水文监测数据	水位、流量、流速、含沙量等
	气象监测数据	降雨、气温、湿度、风速风向等
	水质监测数据	pH值、溶解氧、浊度、污染物等
	工程监测数据	结构变形、应力应变、渗流等
	生态监测数据	生物多样性、植被覆盖、生态流量等
工程数据	设计数据	工程设计图纸、计算书、规范等
	施工数据	施工进度、质量检测、材料用量等
	运行数据	运行状态、调度记录、能耗等
	维护数据	维修记录、设备更换、保养计划等
	安全监测数据	安全评估、风险预警、应急预案等
管理数据	水资源管理数据	取水许可、用水统计、配置方案等
	防洪管理数据	洪水预报、调度方案、抢险记录等
	供水管理数据	供水计划、水质检测、管网信息等
	生态管理数据	生态流量、环境保护、修复措施等
	应急管理数据	应急预案、响应记录、救援资源等
服务数据	公共服务数据	信息公开、政策解读、办事指南等
	应急响应数据	预警发布、响应指令、处置结果等
	决策支持数据	分析报告、预测结果、优化建议等
	用户服务数据	用户需求、服务记录、反馈意见等
	政策法规数据	法律法规、标准规范、政策文件等

当前，我国水利数据流通已初步形成“政府主导、多模式并存、四渠道协同、四级平台支撑”的格局，同时也面临标准与质量、安全可信、技术与协同三大核心瓶颈。

截至2025年，我国水利数据流通仍以政府主导型为主，市场驱动型和混合型为辅，呈现出显著的“三足鼎立”格局[20]。水利数据流通渠道主要包括内部流通、跨部门流通、社会流通。水利数据流通平台已初步形成“国家–流域–地方–行业”四级水利数据流通平台体系，国家水利数据中心作为国家级水利数据平台，提供数据存储、管理和服务功能，承担着全国水利数据的统一管理和服务职责，是水利数据流通的核心枢纽。流域数据中心建立在各大流域，实现流域内数据共享，为流域综合治理和统一调度提供数据支撑，体现了流域管理的特色需求。地方水利数据中心在省、市、县各级建立，服务地方水利工作，满足不同层级水利管理的差异化需求，形成了覆盖全国的数据服务网络。行业数据平台由企业或行业协会建立，提供专业化的数据服务，在特定领域具有技术优势和专业特色，是水利数据流通的重要补充[21]。

（二）水利数据流通面临的主要挑战

1. 数据标准与质量约束

当前，水利数据在标准体系、格式规范与质量控制方面存在明显短板，这直接影响大模型技术的应用效果[22-23]。具体而言，不同部门、不同业务系统采用的编码规则、数据模型和元数据体系缺乏统一规范，导致数据难以共享与复用，无法满足大模型训练对标准化数据的需求。同时遥感监测、站点观测和工程运行等多源数据格式差异显著，难以形成统一的数据流通与处理方案，制约了大模型多模态能力的发挥。受采集条件、技术手段和管理水平制约，数据准确性、完整性、一致性等质量指标参差不齐，直接影响大模型训练质量和推理准确性。

2. 数据安全与可信风险

水利数据承载国家安全、公共安全和生态安全等敏感信息，隐私保护、访问控制与安全审计成为

流通环节的核心诉求[24]。具体而言，洪水风险图、重大工程运行状态、重要水源地信息等一旦泄露，将引发重大安全风险。目前在身份认证、权限分级、日志追溯等机制上仍存在薄弱环节，未授权访问和恶意攻击隐患突出。水利数据从采集、传输、存储到交换、应用的全生命周期中，其安全态势尤为脆弱，面临着技术漏洞、人为失误、管理短板及供应链风险等多维度的威胁。由于缺乏一套贯穿始终、统一管控的端到端安全保障体系，导致安全防护措施呈现碎片化状态，难以形成协同防御能力，从而对国家水安全构成了系统性风险[25]，亟需从制度、技术与管理三方面构建高可信度保障框架。

3. 技术与协同能力瓶颈

水利数据具有海量、实时、多源异构的典型特征，对算力资源、算法模型与平台架构提出极高要求。例如，洪水预报与调度决策、工程联合优化调度等业务对数据处理的时效性和可靠性要求苛刻，而现有分布式计算、边缘协同和存储体系尚难全面支撑。多源异构数据融合、大模型在线推理与模型可解释性等关键技术仍在探索阶段，制约了水利数据价值的深度挖掘[26]。同时，受限于部门壁垒和利益分配机制的滞后，跨部门、跨地区协同治理缺乏有效的组织保障与激励政策。

这些挑战相互交织、层层制约，共同构成阻碍大模型技术深度赋能水利行业现代化的主要障碍。若不能系统性解决标准不统一、质量参差、安全保障薄弱、技术支撑滞后、协同不畅等关键问题，水利数据潜力将难以充分释放，大模型驱动的智能预警、智能调度、智慧决策等核心场景落地也将受到根本性制约。需从标准体系完善、质量治理提升、安全可信框架构建、技术能力突破、跨部门协同机制创新等多维度协同发力，为水利数据安全高效流通与大模型深度融合应用提供坚实保障。

三、创新技术赋能大模型在水利行业的应用

大模型技术作为人工智能领域的重要突破，其强大的数据处理能力和知识理解能力为水利数据流通提供了新的技术路径。从大模型的特点出发，对大模型技术在水利数据流通中的技术机制和应用特征进行分析，为建立基于大模型的水利数据流通技术体系提供理论依据，大模型赋能水利数据流通的技术特点和驱动机制如图1所示。

（一）大模型赋能水利数据流通特征及驱动作用

大模型是指参数量达到数十亿甚至数千亿的深度学习模型，通常基于Transformer架构构建[27]，大模型具有规模效应、多模态能力、知识压缩、泛化能力四个核心特征。规模效应是指大模型通过增加参数量和训练数据规模，展现出的强大的涌现能

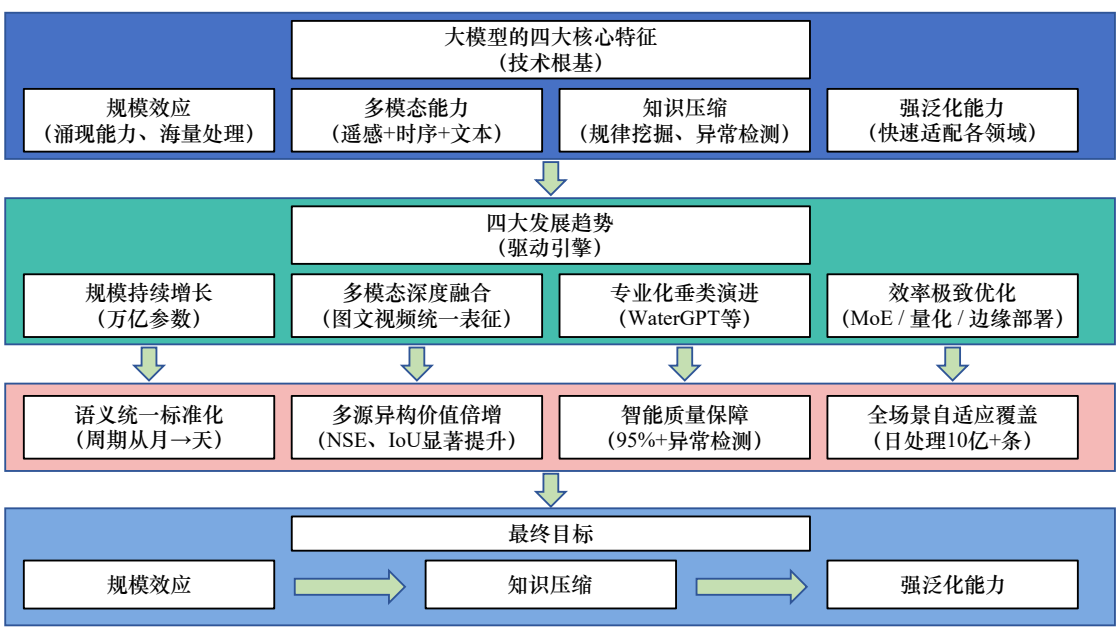


图1　大模型赋能水利数据流通的技术特征与驱动机制

注：MOE为混合专家模型，NSE为纳什效率系数，GPT为生成式预训练变换器，IoU指交并比。

力[28]。在水利数据流通中，规模效应使大模型能够处理海量的水文、气象、地质等多源异构数据，支撑流域级、国家级水利数据流通平台的建设[29]。其次，在水利领域，多模态能力使大模型能够融合卫星遥感影像、传感器时序数据、文本报告等不同模态的水利数据，提升水利数据流通的效率和价值[30]。在水利数据流通中，知识压缩能力使大模型能够从历史水利数据中学习规律和模式，为数据质量评估、异常检测、决策支持等环节提供智能支撑[31-32]。泛化能力使大模型能够适应不同流域、不同业务场景的需求，快速部署到防洪、水资源管理、工程运维等不同应用场景。

大模型技术呈现出模型规模持续增长、多模态融合、专业化发展、效率优化等主要发展趋势[33]。从GPT-3的1750亿参数到GPT-4的数万亿参数，模型规模仍在快速增长。在水利数据流通中，模型规模的持续增长将进一步提升大模型处理海量水利数据的能力，支撑更大规模的水利数据流通平台建设。多模态融合将使大模型更好地整合遥感影像、传感器数据、文本报告等多种水利数据，实现更全面的水利数据流通和价值释放。专业化发展将推动水利领域专用大模型的开发，如洪水预报模型、水资源优化配置模型等，提升水利数据流通的专业性和准确性[34]。效率优化将降低大模型的部署和运行成本，推动水利数据流通的普及化。

目前，水利行业大部分的数据是文档、预案、巡检记录，而传统水利信息化往往聚焦于水位、雨量等结构化数据，而忽视了海量的文本文档、调度预案、专家经验记录等非结构化数据。这些数据因难以被计算机直接理解而无法流通。通过引入大模型技术，利用大模型的长文本理解能力，自动从几十万份水利历史文档中提取实体、关系和事件，构建水利知识图谱，将“死档案”变成可流通的“活知识”。针对不同流域数据标准不一的现状，也可以利用Embedding向量化技术实现跨系统的语义模糊匹配与自动映射，这些都是传统规则引擎无法实现的。

大模型已在水利数据流通过程中展现出显著的跃升效应，其实证成效显著。规模增长驱动的语义统一与标准化跃升，参数量从百亿级跃升至万亿级后，大模型的语义理解能力出现质变，可在零样本条件下实现水利专业术语的跨标准语义对齐。实证显示，基于LLaMA-3-70B的领域微调模型在七大流域32种水位–流量编码体系上的语义一致性准确率达98.7%，将传统人工标准化周期从3~6个月压缩至3~5天[35]，直接打通了“国家级–流域级–省级”三级水利数据流通的语义壁垒。

多模态融合驱动的多源异构数据价值倍增。多模态大模型（如GPT-4o、Gemini 3Pro、Qwen-VL-Max、DeepSeek-VL2等）已具备对遥感影像、站点时序数据、无人机/巡检视频、文字报告等多种异构水文环境数据的统一表征能力。这种跨模态的联合表征学习显著提升了模型在复杂水文场景下的感知与推理能力。如基于对比式视觉语言预训练模型（CLIP）或Unified-IO式框架的多模态模型在融合光学遥感影像与地面站降雨–径流序列后，可使降水及流域洪水预报的NSE较单模态基线显著提升[36]。在河道变化与土地利用等地表解译任务中，SatCLIP（L=10）模型在生态区域和国家地理位置分类任务上的准确率分别达到89.32%和94.28%，较GeoCLIP基线提升4.67%和3.56%[37]。上述进展表明，原本孤立的多源异构数据通过多模态大模型的统一嵌入空间得以高效流通与深度融合，从而在水文监测、防灾减灾和水资源管理等领域实现显著的价值倍增。

专业化发展趋势驱动的知识推理与智能质量保障。水利行业垂类大模型（如“WaterGPT”“HydroGPT”）通过在百亿条专业标注数据上的持续预训练，具备了强水利知识推理能力[31,38]。这些模型支持多模态数据处理与异常检测，如利用大模型框架自动识别时序数据中的异常值（如水位/流量突变），并结合遥感谱指数提升河道沙洲监测精度至95%以上[39]。在洪水预报场景中，集成大模型的混合模型可将NSE较传统水文模型显著提升，并通过知识图谱辅助决策[40]。这些进展表明，专业化大模型不仅保障数据质量，还驱动智能质量控制与风险预测，在流域防洪、水资源调度等领域实现高效知识推理与应用落地。

效率优化发展趋势驱动的自适应部署与全场景覆盖。通过MoE、4-bit量化、知识蒸馏等技术，水利垂类大模型的推理效率显著提升，如MoE-LLM在三维模拟内存计算（AIMC）中显著节省内存、降低能耗、加速推理[41-42]，并支持边缘设备部署[43]。这些优化使模型适用于资源受限场景，如珠江三角洲洪水响应模拟误差通过经验正交函数（EOF）分

析改善10%~20%[44]，松辽盆地径流通过神经网络+GRACE数据预测中国典型区陆地水储量变化模拟Kling-Gupta效率系数（KGE）达0.8以上（误差<15%）[45]。这些进展表明，轻量化大模型实现“一模型多场景”覆盖，在我国多个省级水文平台高效部署，支持日处理实时传感器数据超10亿条[46]，从而推动全国流域水文监测与防灾系统的实时融合与决策。

上述四大发展趋势相互叠加，使水利数据流通从“能连能通”迈向“可信可控、可感可知、高值高用”，流通效率通过模型压缩与高效推理显著提升[47-48]，数据资产价值释放率在水文人工智能应用中实现数倍增长[49]，为构建全国统一的水利数据市场提供了坚实技术底座。

（二）水利数据脱敏技术促进数据流通

在大模型驱动的水利数据流通体系中，数据脱敏是平衡数据安全与价值释放的关键环节，也是构建内生安全层的核心技术手段。面对水利数据多源异构、敏感级别复杂以及实时性要求高的特点，传统的静态规则脱敏已难以满足高效流通的需求。本研究结合水利数据结构和数据特征，建立了水利数据脱敏技术路径图。如图2所示，通过大模型的语义理解能力，识别非结构化文本和结构化文本中的敏感信息，构建一个从敏感数据智能识别到动态脱敏服务的闭环流程，从而在保障国家水安全与部门隐私的前提下，最大化数据的可用性。

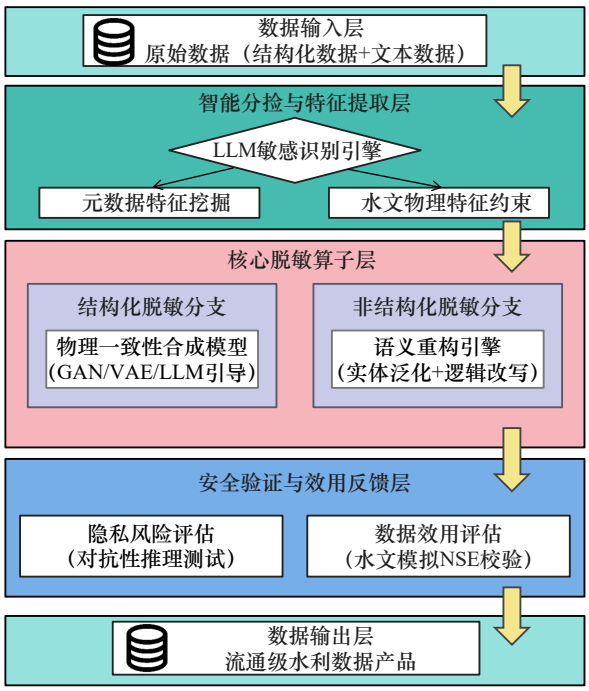


图2　数据脱敏技术框架图

注：LLM为大语言模型，GAN为生成对抗网络，VAE为变分自编码器。

该系统始于数据输入层，接收来自物联网传感器、遥感平台、业务系统、人工上报等多源原始数据，包括结构化数据（如数据库表格、时序监测数据）与非结构化数据（如巡检报告、调度指令、应急预案文本、遥感解译报告），随后进入智能分拣与特征提取层，通过长短期记忆网络时序分析引擎与无监督挖掘技术，从海量数据中识别关键事件并提取具有物理意义的水文特征。处理后的数据分流至核心模型算子层，结构化数据经由融合GNN/VAE/LIM等先进架构的物理一致性合成模型，确保其符合自然规律并完成效用转化，结构化数据使用条件变分自编码器（CVAE）或GAN，以物理方程（如圣维南方程组、水量平衡方程）作为约束条件，训练生成模型。模型输入为原始敏感数据及其上下文特征（如降雨、上游来水），输出为在统计分布与物理规律上与原始数据一致、但不泄露真实值的合成序列。非结构化文本数据针对报告、预案、记录等文本数据，采用提示工程与知识增强生成（RAG），确保大模型在脱敏过程中不引入虚假信息或逻辑矛盾语义脱敏。在保留文档逻辑与知识的前提下，泛化或替换敏感实体与细节以实现知识的凝练与组织。

在生成最终水利数据产品前，进入严格的质量与安全管控阶段，即安全验证与效果反馈层。该层采用对抗性推理测试，模拟攻击者尝试从脱敏数据中重构原始敏感信息，通过训练逆向模型，评估数据被重新识别的风险。若风险高于阈值，则调整脱敏参数或策略，同时将脱敏数据代入下游水文模型或业务系统，检验其是否仍支撑正确的决策与分析，构成可靠的双重校验机制。通过全部风险监控的数据，可封装为分级可流通的数据产品，并附带“脱敏说明标签”，明确标注为L1完全匿名合成、L2部分泛化、L3轻度扰动进入数据。同时输出数据以标准化格式Parquet或JSON-LD发布，支持应用程序编程接口（API）接口按权限动态调用，实现“数据可用不可见、场景可控可计量”。通过上述流程将水利数据塑造为标准化、可流通、可直接支撑业务决策的水利数据产品。

（三）大模型+机理模型增强预测精度

为实现水利数据预测在精度、效率与物理可信度上的统一突破，本研究针对传统机理模型计算效率低、对复杂下垫面刻画不足，以及纯数据驱动大模型的物理一致性差、极端场景易失效的双重瓶颈，提出“物理嵌入与残差校正”深度融合范式（见图3）。具体而言，该框架以简化的耦合水动力模型（基于圣维南方程组的快速求解器）作为机理分支，其核心作用并非进行高精度、高耗时的全尺度模拟，而是在少量时间内，利用地形、糙率、工程边界等基础数据，快速生成流域关键断面的水位、流量物理基准序列及其伴随的中间物理状态场。与此同时，采用融合了注意力机制的Vision-Transformer作为大模型分支的主体架构，设置一个物理投影模块，用于将机理模型输出的低维物理状态向量（如流速、水力坡度）映射为与大模型隐层同维的高维特征嵌入。嵌入特征将与经过多模态编码的实时监测数据、雷达降水预报、遥感影像解译成果乃至文本形态的调度指令进行深度融合，共同作为条件输入，引导大模型学习机理模型因参数简化而无法解释的预测残差。整个训练过程通过设计混合损失函数，将数据驱动损失与物理驱动损失（将模型最终预测值反代回圣维南方程组计算的质量、动量守恒残差）相结合，从数学优化层面硬性约束模型的输出空间，确保其严格遵循水动力学基本定律。

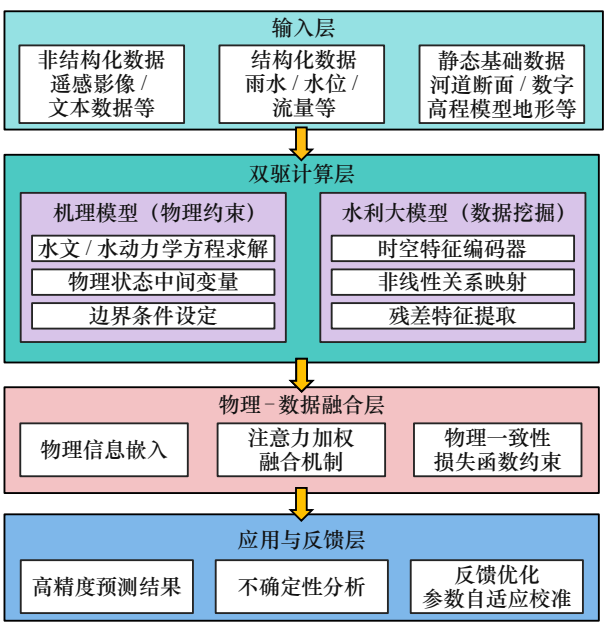


图3　大模型+机理模型融合架构

通过“物理状态先验生成－高维特征投影嵌入－条件残差学习”的闭环架构，模型不仅具备“物理基线+智能修正”的可解释性结构，更显著提升了对历史罕见极端工况的推理鲁棒性。本研究选取长江中下游的防洪控制断面汉口站数据进行仿真验证，在确保水量平衡误差小于1%的物理合理性约束下，将融合模型与传统一维水动力模型进行对比。结果表明，融合模型将洪水峰值预报的预见期平均提前了3.2 h，同时将关键断面水位预测的均方根误差降低了24.7%。该框架为构建下一代水利数字孪生平台中“可解释、强泛化、快响应”的智能模拟引擎提供了关键技术路径，有力推动水利业务从“模型驱动”和“数据驱动”向“物理机理与数据双驱动”的范式演进。

四、大模型驱动的水利数据流通技术体系

基于大模型的水利数据流通技术体系应遵循开放的技术体系、可扩展的技术体系、安全的技术体系和符合相关标准的技术体系四个原则。

（一）技术体系总体架构

基于大模型的水利数据流通技术体系采用分层架构设计（见图4），涵盖多源采集、混合存储、智能治理、敏捷服务、场景应用及内生安全的六层架构，形成完整的技术支撑体系，各层之间既相对独立又紧密协同，共同构建大模型驱动水利数据高效流通的技术基础。

多源采集层作为技术体系的源头，负责从多元化数据源实时、准确地采集水利数据。该层利用大模型的模式识别能力，在数据采集过程中实时检测异常数据，通过语义理解自动识别和分类采集数据，基于知识推理预判数据质量，提前发现潜在问题，提升数据采集的智能化水平。

混合存储层提供高性能、可扩展的分布式数据存储服务，采用混合存储策略满足水利数据的多样化需求。针对结构化数据，采用关系型数据库和时序数据库进行高效存储。针对半结构化数据，采用文档数据库进行灵活存储；针对非结构化数据，采用对象存储和分布式文件系统进行大规模存储。同时，实施数据分层存储策略，通过自动化数据生命周期管理优化存储成本和访问性能。该层利用大模

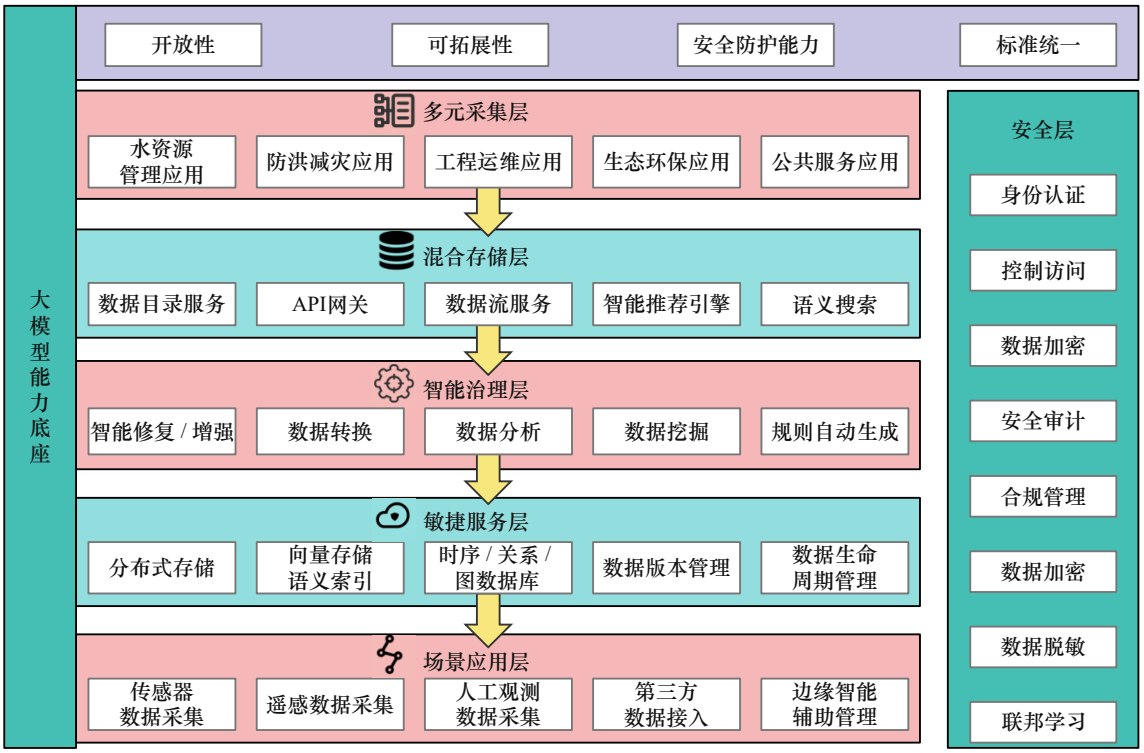


图4　基于大模型的水利数据流通技术体系架构

型的语义理解能力，自动生成数据的语义索引，实现基于语义的数据检索，提升检索准确性。基于大模型自动对数据进行分类和标签化，优化存储结构，提高数据管理的智能化水平。

智能治理层基于大模型技术，对异构水利数据进行处理和分析，利用大模型的多模态理解和生成能力，将不同格式的数据和标准进行统一处理。利用大模型的异常检测和模式识别能力，通过大模型训练后的模式识别模型自动检测数据中的错误值、缺失值、异常值，并根据数据的历史模式进行修复。同时利用大模型的语义理解能力，将不同格式的语义映射、标准转换规则生成，并将不同格式的数据和标准转换为统一格式。结合水利领域知识进行知识推理，对水文预测、洪水预警、水资源优化配置等进行分析。利用大模型的关联发现能力，通过模式识别、规律提取，发现隐藏的规律和模式。

敏捷服务层提供标准化、高性能的数据服务接口，为上层应用提供统一的数据访问能力。该层利用大模型的语义理解能力，支持自然语言查询，用户可用自然语言描述需求，系统自动理解并返回相关数据。基于大模型的知识推理能力，根据用户历史查询和业务场景，智能推荐相关数据资源。通过大模型实现跨数据源的语义搜索，即使数据格式不同，也能基于语义关联找到相关数据。

场景应用层是利用数据服务层开发适用于水利业务的应用系统，实现数据的释放。防洪减灾业务提供洪水预报、风险评估、应急调度等功能，通过大模型调用专业水文水力学模型进行计算，并利用大模型对计算结果进行语义化解读和报告生成。水资源管理业务提供水资源配置、用水效率评价、节水管理等功能，实现水资源利用效率的提升。工程运维业务提供水利工程健康监测、故障诊断、维护决策等功能，提升工程安全运营水平。生态环保业务提供水环境监测、生态流量保障、水生态修复等功能，实现水生态文明建设。公众服务业务提供水情信息查询、用水服务、政策解读等功能，提升水利公共服务水平。

内生安全层贯穿整个技术体系，构建全方位、多层次的安全防护体系。该层采用零信任安全架构，实现“永不信任，始终验证”的安全策略。对传输中的数据采用传输层安全协议 / 安全套接层协议（TLS/SSL）加密，对存储的数据采用高级加密标准（AES）等算法加密，对敏感数据支持同态加密等高级加密技术。安全审计记录完整的操作记录和访问记录，方便安全事件追溯和分析。合规管理

按照国家网络安全等级保护、《中华人民共和国数据安全法》等的有关要求来实现合规性，通过安全层防护保障水利数据的安全，同时保证水利数据的有效流通。

三峡工程调度涉及水文、气象、机组运行、电网等多系统海量异构数据。在传统模式下，多源数据协同依赖人工，决策效率受限。当前已有研究探索流域多模态大模型技术，融合水文站、气象雷达、卫星遥感、地质灾害监测等多源感知数据，借助跨模态融合与知识图谱技术，支撑防洪、发电、生态多目标调度智能辅助决策。此外，大模型的异常检测功能成功预警库区某岸段的地质微变形风险，展现了其在工程安全领域的前瞻性决策支持能力。

上述案例表明，大模型技术能够有效打通重大水利工程中的“数据孤岛”，实现多源异构数据的智能融合与价值挖掘，为防洪调度、水资源优化、工程安全等核心业务提供精准、高效的决策支持，验证了所提技术架构的可行性与实用性。

（二）多源数据采集与智能预处理技术

水利数据采集和预处理是水利数据流通的基础。多源数据采集包括多种数据采集方式，如以物联网采集水位、流量、水质等数据，建立全国水利监测网络；以卫星遥感、无人机遥感等获得大量水利数据，如水体、植被、土地利用等；以人工观测数据补充现场调查、专家评估、公众报告等；以第三方数据接入与气象、环保、农业等共享数据，实现跨部门数据共享。数据预处理基于大模型技术实现智能化处理，利用大模型的异常检测和模式识别算法，自动识别和修复数据中的错误、异常值、缺失值等问题。基于大模型的语义理解能力，实现不同格式数据的智能转换，建立统一的数据模型和编码规则。同时利用大模型的多模态能力，将传感器时序数据、卫星遥感影像、文本报告等异构数据进行跨模态融合。基于大模型的自然语言处理能力，自动为数据添加语义标签。通过大模型技术的全面应用，数据采集与预处理过程实现了从人工处理向智能化自动处理的转变，大幅提升了数据处理效率和质量，为后续的数据分析和应用奠定了坚实基础。

（三）混合存储与智能管理技术

数据存储与管理技术是确保数据质量和可用性的重要保障。根据数据特性和访问模式，采用关系型数据库、NoSQL 数据库、时序数据库、对象存储等不同的存储方式。分层存储热数据、温数据、冷数据在不同性能存储设备上，通过数据生命周期管理降低存储成本，提升访问性能。通过实时备份、增量备份、异地备份等完成数据备份恢复。通过大模型可自动生成元数据描述和语义标记，记录数据的来源、格式、质量等信息，跟踪数据的变更历史，自动管理数据的生命周期。该层利用大模型的语义理解能力，自动生成数据的语义索引，实现基于语义的数据检索，提升检索准确性。基于大模型自动对数据进行分类和标签化，优化存储结构。通过大模型技术的深度应用，可实现从被动管理向主动智能管理的转变，显著提升了数据管理的自动化水平和智能化程度，为大规模水利数据的存储和管理提供高效、可靠的技术支撑，为水利数据的高效流通提供了坚实的管理基础。

（四）敏捷共享与交换技术

数据共享与交换技术是开展水利数据流通的基础。数据共享协议采用 API 接口标准，如数据 API 接口标准、ResTful、graphQL 等，基于大模型生成语义化的接口描述，为各个系统数据交换提供标准接口，降低集成成本和复杂度。数据交换采用 JSON、XML、Parquet 等格式进行数据交换，利用大模型的语义理解能力，进行智能匹配，实现数据的互操作性和兼容性。基于 OAuth2.0、JWT 等标准的认证授权机制，确保数据访问的安全性和可控性，支持细粒度的权限管理和访问控制。数据目录服务利用大模型生成语义化数据目录，支持基于语义的智能数据发现和检索，帮助用户快速找到所需的数据资源。数据市场利用大模型的智能匹配和推荐能力，支持数据交易和共享，为数据提供方和使用方提供交易平台，促进水利数据的市场化配置，实现数据的价值变现。数据协作平台促进跨部门、跨地区的合作。通过建立完善的数据共享协议体系和数据交换平台，实现了不同系统、不同平台之间的数据互联互通，构建了完整的数据流通生态，实现了从数据发现、数据交易到数据协作的全流程服务，为水利数据的跨系统、跨平台流通提供了标准化的技术规范和安全保障，有效促进了数据的共享和协同应用，为水利数据的高效流通和价值释放提供了全方位的平台支撑。

（五）内生安全与隐私保护技术

水利数据安全与隐私保护技术架构采用多层次、全方位的技术体系。数据加密是保证水利数据安全的最基本手段，利用TLS/SSL对数据传输进行加密，也可利用AES等算法对存储数据进行加密，保证静态数据。也可利用同态加密支持加密数据计算，可在不暴露原始数据的情况下对数据进行分析、处理，为保护隐私计算提供技术支持。访问控制是保障水利数据安全访问的重要手段，基于角色的访问控制按照用户角色设置数据访问权限进行访问控制。基于属性的访问控制按照用户属性、资源属性、环境属性等进行访问控制，提供了更加灵活和可变的访问控制机制。对所有访问进行验证，并通过持续验证、监控保障了数据的安全访问。

隐私保护技术是平衡数据利用和隐私保护的重要手段，通过对敏感数据进行脱敏处理，同时使用数据替换、数据泛化、数据抑制等技术，在保持数据可用性的同时保护敏感信息。通过添加噪声和隐私预算控制，确保个体隐私不被泄露，同时保持统计分析的准确性。使用联邦学习技术，在不共享原始数据的情况下进行模型训练，实现多方数据的协同建模，在保护数据隐私的同时提高模型性能。通过构建传输加密、存储加密和同态加密的多层次加密体系，结合访问控制体系，以及数据脱敏、差分隐私和联邦学习的隐私保护体系，实现数据在传输、存储和计算全生命周期的安全保护，构建了多层次、多维度的安全防护体系，为水利数据流通提供了坚实的安全保障，既确保了数据的机密性、完整性和可用性，又支持了数据的安全共享和协同应用，有效平衡了数据开放共享与隐私保护之间的关系。

五、水利数据流通的建设路径与发展建议

（一）水利数据流通的建设路径

水利数据流通的建设路径是一个系统性工程，需要从多个维度协同推进[50-51]。如图5所示，建设路径包括政策法规、标准体系、技术平台、人才培养、试点示范五个核心维度。

在技术平台构建方面，应打造以“水利数据要素流通与交易平台”为核心的一体化支撑体系。该平台围绕“多源供给–促进流通–安全管理–场景驱动”四大功能模块展开。其中多源供给模块负责数据的汇聚、质量评估与标准化，通过沙箱环境与追踪系统确保数据来源清晰、质量可控。促进流通模块借助大模型与智能合约实现数据智能匹配、可信交易与自动化结算，降低流通成本。安全管理模块构建覆盖全生命周期的防护体系，实施细粒度访问控制、分层加密与隐私计算，筑牢数据安全与信任

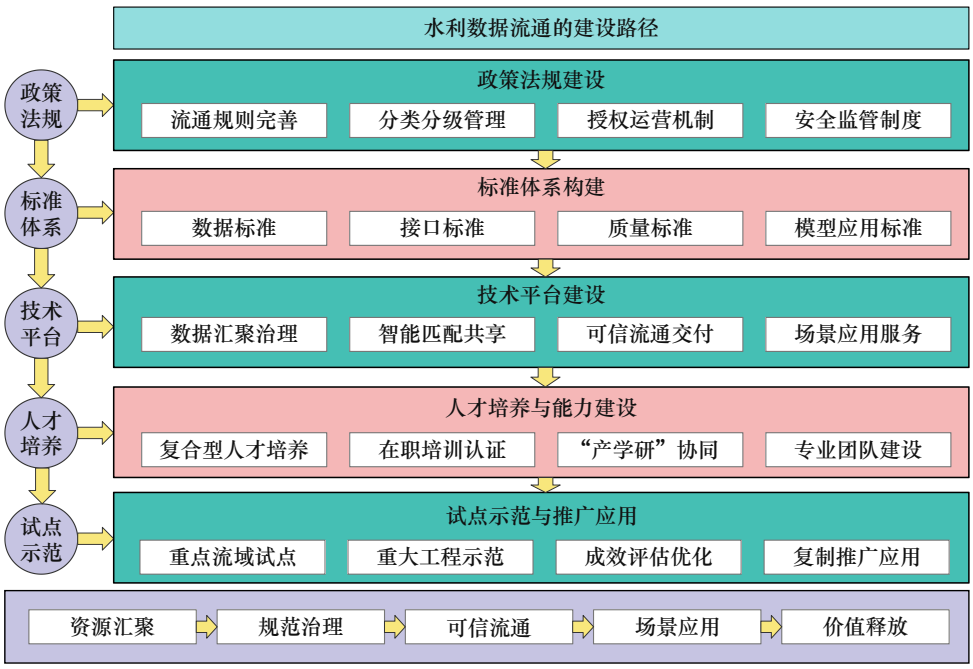


图5　水利数据流通的建设路径

基石。场景驱动模块面向防洪、水资源调配等业务，提供情境数据匹配与协作工具，推动数据在具体业务中深度应用与价值释放。

通过上述五大维度尤其是技术平台的有序建设与协同推进，可逐步打通水利数据从资源汇聚、治理加工、安全流通到场景赋值的全链条，形成“制度为纲、标准为尺、平台为基、人才为核、试点先行”的可持续发展格局，最终推动水利数据要素市场化配置改革落地，全面赋能智慧水利与数字孪生流域建设。

1. 政策法规建设

（1）完善水利数据流通政策体系

完善水利数据流通政策体系是发展水利数据流通的重要保障。制定《水利数据流通管理办法》，确定水利数据流通原则、管理职责、流通程序等，为水利数据流通提供制度保障，规范水利数据流通，建立数据分类分级制度，根据数据敏感程度和重要程度建立数据分类分级管理制度，对不同等级数据采取不同管理方法及安全措施，保障数据安全。建立数据定价机制，合理定价，促进水利数据市场化配置，通过市场手段将数据价值进行合理分配，调动数据提供者的积极性。

（2）加强数据安全法规建设

制定数据安全法实施细则，针对水利数据特点，明确水利数据安全管理的具体要求与标准，为水利数据安全提供依据。建立数据泄露应急响应机制，确保发生数据安全事件能够及时响应和处理，从而降低数据安全事件的损失。建立数据跨境流通监管制度，保护国家安全，对涉及国家秘密与敏感信息的水利数据要严密监管，防止泄露。

2. 标准体系构建

（1）数据标准制定

数据标准的制定是水利数据流通的前提。制定数据格式标准，包括数据结构、编码规则等，不同系统的数据要正确处理，提高数据的可移植性。数据质量标准制定包括数据的准确性、完整性、一致性、时效性等。数据交换标准包括数据交换格式、交换协议、交换流程等。

（2）技术标准制定

技术标准制定是确保水利数据流通技术规范的重要措施。比如制定统一的接口标准，确保系统互操作性，包括接口规范、数据格式、错误处理等标准，为不同系统间的集成提供技术规范。安全技术标准确保数据安全，包括加密算法、认证机制、访问控制等安全技术标准，为数据安全提供技术保障。制定大模型在水利领域应用的技术标准，包括模型训练、部署、评估等，为大模型在水利领域的应用提供技术指导。

3. 共享平台搭建

水利数据智能共享平台旨在构建一个覆盖数据供给、智能匹配与可信流通全链条的完整数据生态系统，以打通数据壁垒、激活数据价值。如图 6 所

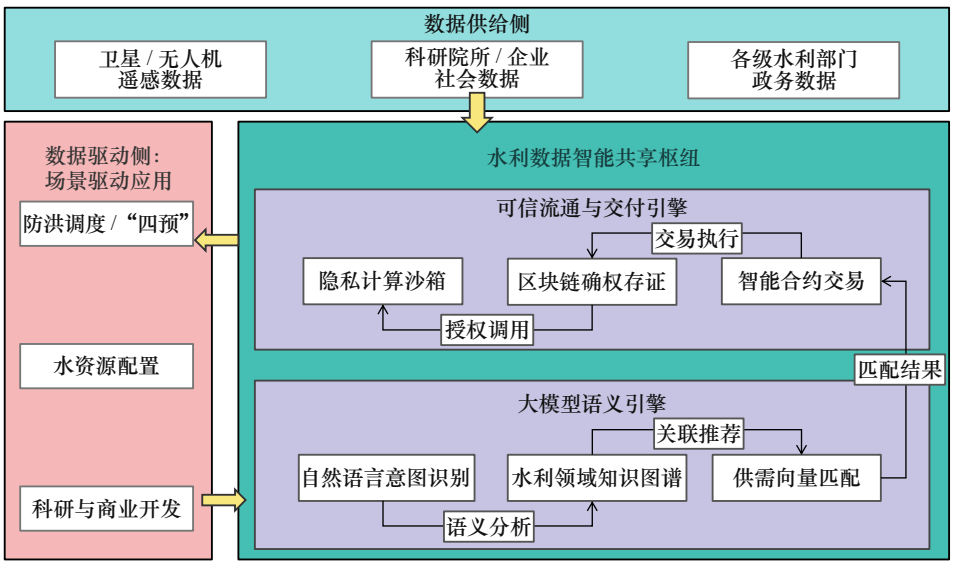


图 6　水利数据共享平台架构流程图

注：“四预”为预防、预警、预演、预案。

示水利数据共享平台的数据供给侧，平台广泛汇聚来自各级水利部门的政务数据、科研院所及企业的社会数据，以及物联网终端、无人机与卫星遥感等多源异构数据，形成统一管理、动态更新水利数据资源池。所有接入数据均需经过标准化注册与元数据描述，确保数据来源清晰、格式规范、权属明确。

在平台核心处理层，集成两大智能引擎，实现从数据到服务的智能转化。其中大模型语义引擎依托自然语言意图识别、水利领域知识图谱与向量化匹配技术，将用户以自然语言提出的数据需求自动解析为结构化查询指令。该引擎能够理解水利专业术语与业务场景，实现数据资源的精准发现与智能推荐，显著降低数据检索与获取门槛。而可信流通与交付引擎基于区块链技术实现数据上链存证和流转追踪，确保数据流通过程中权属清晰、不可篡改；结合隐私计算沙箱，支持在原始数据“不出域”的前提下进行联合建模与分析；通过智能合约自动执行数据使用协议、计费与结算，实现流通过程的自动化与可信化。最终以“可用不可见、可控可计量”的方式，向需求侧输出数据服务或数据产品。

在数据需求侧，平台面向防洪调度、水资源配置、科研创新与商业应用等多类场景，提供便捷、安全、合规的数据接入与使用渠道，形成从数据汇聚、智能处理、可信交付、场景应用、反馈优化的闭环价值链路，持续推动水利数据要素的安全高效流通与价值释放。

（二）大模型驱动的水利数据流通发展建议

1. 构建国家水利数据流通治理体系

构建国家水利数据流通治理体系需要从治理框架、法律法规和标准化三个维度协同推进。在国家层面建立水利数据流通治理委员会，建立“国家–流域–地方”三级治理体系，制定《水利数据流通管理办法》，明确数据所有权、使用权、收益权的界定原则，建立准入、监管、退出机制。修订《中华人民共和国水法》《中华人民共和国防洪法》等相关法律，制定《水利数据安全保护条例》，明确水利数据的分级分类标准和安全保护要求，设立数据纠纷仲裁机构。制定水利数据流通的国家标准、行业标准、团体标准，建立数据质量认证体系，制定数据格式、接口规范、安全标准等技术标准，确保不同系统间的互操作性，为大模型应用提供标准化数据基础。

2. 创新水利数据市场化配置机制

创新水利数据市场化配置机制需要建立交易市场、完善利益分配和创新商业模式。在国家级数据交易所设立水利数据交易专区，建立基于数据质量、稀缺性、时效性的动态定价体系，建立数据泄露保险、交易纠纷处理等风险防控机制。建立“谁投入、谁受益”的利益分配原则，明确数据提供方、使用方、平台方的收益分配比例，对水利数据流通相关企业给予税收减免，设立专项资金支持基础设施建设。鼓励发展数据银行、数据信托、数据合作社等新业态、新模式，建立投融资机制，引导社会资本参与，形成政府引导、市场主导、社会参与的多元化流通格局。

3. 强化水利数据流通的技术支撑

强化水利数据流通的技术支撑需要建设技术平台、加强技术创新和建立安全保障体系。建设覆盖全国的水利数据流通技术平台，整合各级水利部门的数据资源，采用云计算、大数据、人工智能等先进技术，建设高性能的数据处理和分析平台，建立API网关，提供统一的数据访问接口。设立水利数据流通技术创新基金，支持大模型在水利数据流通中的应用技术研发，建立“产学研用”协同创新机制，重点突破多源异构数据融合、大模型在线推理、模型可解释性等关键技术。建立多层次的数据安全防护体系，采用区块链、同态加密等先进技术，建立数据泄露应急响应机制，确保数据在采集、传输、存储、交换过程中的安全。

4. 推进水利数据流通的试点示范

推进水利数据流通的试点示范需要选择重点流域开展试点、支持重大水利工程应用和建立人才培养体系。选择长江、黄河、珠江等重点流域开展水利数据流通试点，探索不同流域的水利数据流通模式，建立示范区，形成可复制、可推广的经验模式，定期评估试点效果。以重大水利工程为依托，开展水利数据流通的深度应用，建立重大水利工程的水利数据流通标准，探索大模型技术在工程调度、风险预警、生态保护等方面的应用。在高校设立水利数据流通相关专业，培养既懂水利业务又懂数据科学和大模型技术的复合型人才，建立培训体系和人才激励机制，为水利数据流通提供人才保障。

六、结语

本研究立足于水利数据流通的现实痛点，系统创新性地构建了大模型驱动的“机理–技术–路径”全栈框架，填补了通用大模型在水利垂直领域应用架构设计的空白。在机理层面，本研究首次系统深入剖析了大模型规模效应、多模态感知、知识压缩及泛化适应四大特征，揭示了其在异构语义统一、多源数据价值倍增、知识智能推理及全场景自适应部署中的核心驱动机制。在技术层面，突破了传统单一的数据治理模式，创新设计了涵盖多源采集、混合存储、智能治理、敏捷服务、场景应用及内生安全的六层技术架构，特别是提出了基于大模型的智能预处理、语义化存储检索及多层次隐私防护等关键技术解决方案。在路径层面，提出了“技术底座+制度创新”的双轮驱动策略，构建了“国家–流域–地方”三级协同治理体系，并从政策法规完善、标准体系确立、基础设施夯实及市场化配置等方面提出了系统化实施路径。未来研究应进一步聚焦大模型与水利机理的深度融合，重点突破多源异构数据智能融合、大模型在线推理优化、模型可解释性增强等关键技术，推动水利科学与数据科学、经济学、法学等学科的交叉创新，为水利数据价值的高效释放与行业数字化转型提供持续的理论支撑和技术引擎。

利益冲突声明

本文作者在此声明不存在任何利益冲突或财务冲突。

Received date: December 10, 2025; **Revised date:** February 4, 2026

Corresponding author: Zhang Weiwei is an associate professor from the School of Advanced Interdisciplinary Studies, Hunan University of Technology and Business. His major research field is water resources management and decision-making. E-mail: 3022@hutb.edu.cn

Funding project: Chinese Academy of Engineering project “Research on Improving the Institutional Mechanisms for Modernizing Infrastructure Development” (2024-JJZD-03); The National Natural Science Foundation of China Project (52408364); Hunan Provincial Natural Science Foundation (2024JJ6188)

参考文献

[1] 陈晓红，肖粲然，曹文治，等. 我国统一数据要素大市场框架体系与建设路径研究[J]. 中国工程科学，2025，27(1)：40–50.
Chen X H，Xiao C R，Cao W Z，et al. Framework and pathway for the construction of a unified data-element market in China[J]. Strategic Study of CAE，2025，27(1)：40-50.

[2] 王浩，李晓晴. 国家水网主骨架和大动脉不断完善[N]. 人民日报，2024-10-09(14).
Wang H，Li X Q. Continuous improvement of the main framework and major arteries of the national water network[N]. People’s Daily，2024-10-09(14).

[3] 中共中央 国务院印发《国家水网建设规划纲要》[J]. 大坝与安全，2023(3)：21.
Central Committee of the Communist Party of China，State Council issuance of the outline for *the construction planning of the national water network*[J]. Dam & Safety，2023(3)：21.

[4] 数字中国建设整体布局规划[EB/OL].(2023-02-27)[2026-01-21]. https://www.gov.cn/zhengce/2023-02/27/content_5743484.html.
Overall layout plan for building a digital China[EB/OL].(2023-02-27)[2026-01-21]. https://www. gov. cn/zhengce/2023-02/27/content_5743484.html.

[5] 李永安. 大型水电项目融资与资本运作——三峡工程10年来的融资策略及实践[J]. 中国工程科学，2003，5(10)：1-7.
Li Y A. Financing & capital operations of large hydropower projects：A summary review of financing strategies & practices in Three Gorges Project[J]. Strategic Study of CAE，2003，5(10)：1-7.

[6] 中共中央 国务院关于构建数据基础制度更好发挥数据要素作用的意见[J]. 自然资源通讯，2022(24)：14-18.
Central Committee of the Communist Party of China，State Council. Opinions on establishing a basic data system to better tap the role of data as a production factor[J]. National Land & Resources Information，2022(24)：14-18.

[7] 关于大力推进智慧水利建设的指导意见[EB/OL].(2022-12-30)[2026-01-21]. http://www. mwr. gov. cn/zw/tzgg/tzgs/202201/t20220106_1564240.html.
Guiding opinions on vigorously promoting the construction of smart water conservancy[EB/OL].(2022-12-30)[2026-01-21]. http://www.mwr.gov.cn/zw/tzgg/tzgs/202201/t20220106_1564240.html.

[8] Wang L，Ma C，Feng X Y，et al. A survey on large language model based autonomous agents[J]. Frontiers of Computer Science，2024，18(6)：186345.

[9] Xi Z H，Chen W X，Guo X，et al. The rise and potential of large language model based agents：A survey[J]. Science China Information Sciences，2025，68(2)：121101.

[10] Loukatos D，Kondoyanni M，Alexopoulos G，et al. On-device intelligence for malfunction detection of water pump equipment in agricultural premises：Feasibility and experimentation[J]. Sensors，2023，23(2)：839.

[11] Luviano Soto I，Sánchez Y C，Raya A. Water quality polluted by total suspended solids classified within an artificial neural network approach[J]. Water Quality Research Journal，2025，60(1)：214-228.

[12] 白东东，丁扬威，田嘉皓，等. 国内数字孪生水利研究进展分析[J]. 人民长江，2025，56(S1)：313-321.
Bai D D，Ding Y W，Tian J H，et al. Analysis on research progress of digital twin water conservancy in China[J]. Yangtze River，2025，56(S1)：313-321.

[13] Fereshtehpour M，Esmaeilzadeh M，Alipour R S，et al. Impacts of DEM type and resolution on deep learning-based flood inundation mapping[J]. Earth Science Informatics，2024，17(2)：1125-1145.

[14] 柴华，侯爱中，成建国，等. 人工智能水文预报模型技术研究进展[J]. 水文，2026，46(1)：1-11.
Chai H，Hou A Z，Cheng J G，et al. Research progress on artificial intelligence-based hydrological forecasting models[J]. Journal of China Hydrology，2026，46(1)：1-11.

[15] 邹晓涛，金思凡，黄天称，等. 水利数据要素构建与乘数效应激活路径[J]. 中国水利，2025 (5)：15-28.
Zou X T，Jin S F，Huang T C，et al. Construction of water data elements and activation pathways for multiplier effects[J]. China Water Resources，2025(5)：15-28.

[16] 李政，胡成盛，周振. 积极推进数据要素市场化配置改革[J]. 宏观经济管理，2025(3)：70-77.
Li Z，Hu C S，Zhou Z. Actively reform the market-oriented allocation of data elements[J]. Macroeconomic Management，2025(3)：70-77.

[17] 赵放，宋健，王洁霜. 数据要素市场化对营商环境的优化效应——基于政府和市场的双重视角[J]. 浙江工商大学学报，2025(3)：104-115.
Zhao F，Song J，Wang J S. Optimization effect of data factor marketization on business environment：Based on the dual perspective of government and market[J]. Journal of Zhejiang Gongshang University，2025(3)：104-115.

[18] 金和平，杨磊，涂晶，等. 基于混合数据库的长江流域数据要素流通平台构建[J]. 人民长江，2025，56(3)：222-229.
Jin H P，Yang L，Tu J，et al. Construction of data circulation platform for Changjiang River Basin based on hybrid database technology[J]. Yangtze River，2025，56(3)：222-229.

[19] 华美芳，张勇. 政府主导型数据交易平台的发展模式、困境与对策——基于经典扎根理论的研究[J]. 信息资源管理学报，2025，15(2)：73-90.
Hua M F，Zhang Y. The development pattern，dilemmas，and countermeasures of government-led data trading platforms：An analysis based on classical grounded theory[J]. Journal of Information Resources Management，2025，15(2)：73-90.

[20] Muñoz-Carpena R，Carmona-Cabrero A，Yu Z W，et al. Convergence of mechanistic modeling and artificial intelligence in hydrologic science and engineering[J]. PLoS Water，2023，2(8)：1-24.

[21] Feng D P，Liu J T，Lawson K，et al. Differentiable，learnable，regionalized process-based models with multiphysical outputs can approach state-of-the-art hydrologic prediction accuracy[J]. Water Resources Research，2022，58(10)：e2022WR032404.

[22] Willard J D，Varadharajan C，Jia X W，et al. Time series predictions in unmonitored sites：A survey of machine learning techniques in water resources[J]. Environmental Data Science，2025，4：e7.

[23] 李强，赵铜铁钢. 基于大语言模型的洪涝灾害统计调查[J]. 水利水电技术(中英文)，2025，56(9)：60-75.
Li Q，Zhao T T G. Statistical survey on flood disasters based on large language model[J]. Water Resources and Hydropower Engineering，2025，56(9)：60-75.

[24] Fu R，Jiao Y，Chen D，et al. Research on the application of large language models in shore power management decision[C]//2024 4th International Conference on Intelligent Power and Systems (ICIPS)，2024：1348-1352.

[25] Lippolis A S，Lodi G，Nuzzolese A G. The water health open knowledge graph[J]. Scientific Data，2025，12：274.

[26] 刘海彪，赵帅，陈伟昌，等. 水利感知网中数据传输规约研究[J]. 人民长江，2025，56(S1)：364-367.
Liu H B，Zhao S，Chen W C，et al. Research on data transmission protocol in water resources perception network[J]. Yangtze River，2025，56(S1)：364-367.

[27] Gupta N S，Valarmathi B，Hiwanj I. A comparison of deep learning and machine learning methods for water quality prediction[C]//2025 3rd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT)，2025：1361-1368.

[28] 吴华赟，胡国华，王利丽. 基于大模型的数据安全技术探索与研究[J]. 中国信息安全，2025(9)：33-37.
Wu H Y，Hu G H，Wang L L.Exploration and research on data security technology based on big model[J]. China Information Security，2025(9)：33–37.

[29] Wang X H，Li Y，Qiao Q，et al. Enhancing surface water quality parameter prediction using deep learning and feature augmentation methods[J]. Water，2025，17(24)：3523.

[30] 蒋云钟，冶运涛，赵红莉，等. 水利大数据研究现状与展望[J]. 水力发电学报，2020，39(10)：1-32.
Jiang Y Z，Ye Y T，Zhao H L，et al. Research status and prospects on water conservancy big data[J]. Journal of Hydroelectric Engineering，2020，39(10)：1-32.

[31] 董雅雯，李恒，刘佳，等. 大模型技术的发展与应用：现状、机遇与挑战[J]. 人工智能，2025 (1)：110-118.
Dong Y W，Li H，Liu J，et al. Development and application of large-model technology：Status quo，opportunities and challenges [J]. AI-View，2025(1)：110-118.

[32] Aderemi I A，Kehinde T O，Daniel Okwor U，et al. Explainable AI for water quality monitoring：A systematic review of transparency，interpretability，and trust[J]. IEEE Sensors Reviews，2025，2：419-443.

[33] Shi J H，Ni L X，Liu J M，et al. Water quality prediction at large-Scale：Systematic assessment of decomposition-deep learning models[J]. Journal of Cleaner Production，2025，519：145995.

[34] Ding Y K，Zhu Y L，Feng J，et al. Interpretable spatio-temporal attention LSTM model for flood forecasting[J]. Neurocomputing，2020，403：348-359.

[35] Li X，Sun Q L，Zhang Y F，et al. Enhancing hydrological extremes prediction accuracy：Integrating diverse loss functions in Transformer models[J]. Environmental Modelling & Software，2024，177：106042.

[36] Sit M，Demiray B Z，Xiang Z R，et al. A comprehensive review of deep learning applications in hydrology and water resources[J]. Water Science and Technology，2020，82(12)：2635-2670.

[37] 雷畅，罗斌，黄艳. 面向业务场景的水利数据模型构建方法[J/OL]. 人民长江，2025-12-01. http://www.rmcjzz.cjw.cn/cn/article/pdf/preview/57b1e289-70ad-4a73-a90e-9a271be8d849.pdf.
Lei C，Luo B，Huang Y. Method for constructing water-resources data model oriented to business scenarios[J/OL]. Yangtze River，2025-12-01. http://www. rmcjzz. cjw. cn/cn/article/pdf/preview/57b1e289-70ad-4a73-a90e-9a271be8d849.pdf.

[38] Brown T，Mann B，Ryder N，et al. Advances in neural information

processing systems[C]//Advances in Neural Information Processing Systems，2020：1877-1901.

[39] Hess P，Drüke M，Petri S，et al. Physically constrained generative adversarial networks for improving precipitation fields from Earth system models[J]. Nature Machine Intelligence，2022，4（10）：828-839.

[40] Kim J，Kwon Y，Jo Y，et al. KG-GPT：A general framework for reasoning on knowledge graphs using large language models[C]// Findings of the Association for Computational Linguistics：EMNLP 2023，2023：9410-9421.

[41] Dong S J，Wang L B，Du B，et al. ChangeCLIP：Remote sensing change detection with multimodal vision-language representation learning[J]. ISPRS Journal of Photogrammetry and Remote Sensing，2024，208：53-69.

[42] Micek S，Katzenmeyer B，Nichols A，et al. HydroGPT - on demand training for hydraulic modeling workflows[EB/OL]. 2025 [2025-09-10].https：//awschool.com.au/training/hydrogpt/.

[43] Mukherjee J. Discrimination of river sandbanks for sand mining in high-mineral regions using multispectral images[J]. Discover Geoscience，2025，3（1）：100.

[44] Lin R H，Chen S，He Z Y，et al. Electricity behavior modeling and anomaly detection services based on a deep variational autoencoder network[J]. Energies，2024，17（16）：3904.

[45] Guan X X，Shan B Y，Nguyen V D，et al. Beyond observed extremes：Can hybrid deep learning models improve flood prediction?[J]. EGUsphere Preprints，2025：EGUsphere–2025–1509.

[46] Büchel J，Vasilopoulos A，Simon W A，et al. Efficient scaling of large language models with mixture of experts and 3D analog in-memory computing[J]. Nature Computational Science，2025，5（1）：13-26.

[47] Hess P. When AI's large language models shrink[J]. IEEE Spectrum，2023，60（11）：28-33.

[48] Yao Y，Yu T Y，Zhang A，et al. Efficient GPT-4V level multimodal large language model for deployment on edge devices[J]. Nature Communications，2025，16：5509.

[49] Wu Y P，Yin X W，Zhou G Y，et al. Rising rainfall intensity induces spatially divergent hydrological changes within a large river basin[J]. Nature Communications，2024，15：823.

[50] Lu S B，Li W Q，Yao G B，et al. The changes prediction on terrestrial water storage in typical regions of China based on neural networks and satellite gravity data[J]. Scientific Reports，2024，14：16855.

[51] Zhang Z Y，Shan Y L，Zhao D D，et al. City level water withdrawal and scarcity accounts of China[J]. Scientific Data，2024，11：449.