

北京市水务数据目录编制研究

尹晓楠

(北京市智慧水务发展研究院, 101117, 北京)

摘要:近年来国家大力推进数字政府建设、政务大数据体系建设。水务数据是政务数据的一部分,水务数据目录是水务数据一体化管理的基础底账,编制清晰完整的水务数据目录十分必要。在分析现有数据分类和元数据模型标准规范、梳理北京市水务数据资源现状基础上,提出基于职责的线分类法,对北京市水务数据进行了三级分类:一级分类包括基础数据、监测感知数据、业务数据、行政管理数据四类,进一步细化为38个二级分类、160个三级分类。借鉴政务数据元数据、水利相关数据元数据和数据分类标准,研究提出符合北京市水务业务人员使用习惯的水务数据目录元数据和数据目录编码规则。研究内容已在北京市水务数据目录梳理工作中得到实践应用,显现分类较稳定且边界清晰不交叉特点,可为厘清数据分类脉络提供参考,并为其他领域数据目录梳理和各地水务数据目录梳理提供思路和方法。

关键词:水务数据;数据目录;数据分类;元数据;元数据模型

Study on the compilation of Beijing water affairs data catalog//YIN Xiaonan

Abstract: In recent years, China has vigorously promoted the development of digital government and government big data systems. As a part of government data, water affairs data requires a well-structured and comprehensive data catalog to serve as the fundamental basis for integrated data management. Based on an analysis of existing data classification and metadata model standards, and a review of the current status of water data resources in Beijing, this study proposes a responsibility-based line classification method. The data is categorized into three hierarchical levels: the first level includes four categories, including basic data, monitoring and sensing data, business data, and administrative management data, which are further refined into 38 secondary and 160 tertiary categories. Drawing on government metadata standards, water-related metadata, and data classification guidelines, metadata elements and coding rules tailored to the usage habits of Beijing's water affairs professionals are proposed. The outcomes have been applied in the practice of compiling Beijing's water affairs data catalog, demonstrating a stable classification structure with clear, non-overlapping boundaries. This approach offers a valuable reference for structuring data classification and provides methodologies applicable to other sectors and regional water affairs data catalog development.

Keywords: water affairs data; data catalog; data classification; metadata; metadata model

中图分类号: TV+TP39 文献标识码: B 文章编号: 1000-1123(2025)14-0066-07

DOI: 10.3969/j.issn.1000-1123.2025.14.011

收稿日期: 2025-02-24 修回日期: 2025-05-06

作者简介: 尹晓楠, 高级工程师, 主要从事水务信息化相关工作。E-mail: xiaonanyin@126.com

2022年以来,国家《关于加强数字政府建设的指导意见》《全国一体化政务大数据体系建设指南》《关于加快公共数据资源开发利用的意见》等文件相继发布,要求建立全国标准统一、动态管理的政务数据目录。在国家政策指导下,信息资源管理方向的专家学者提出深入研究数据目录体系。曾建勋提出强化数据目录体系研究,在数字环境下拓展各类信息资源目录研究,丰富目录学体系。王平等提出目录学的重塑性发展,以“文化”和“数据”作为目录学发展场域。一些专家学者在自然资源数据、农业数据、军事数据的数据目录编制方面开展研究,提出了不同领域数据体系和目录编制方法。山东、辽宁、上海等地也对水利、海洋数据目录编制方法进行了研究和实践。国际上,“开放政府数据”(Open Government Data, OGD)是学术研究热点,学者对万维网联盟(World Wide Web Consortium, W3C)推荐的数据目录词汇表(Data Catalog Vocabulary, DCAT)新版本十分关注,并进行了研究解释,在物流业等方面对政府开放数据目录进行了应用探索。数据目录是理清数据资源内容的数据“一本账”,对有序推动政务数据共享开放、释放数据潜能、推动实现“一数之源”、不断提升政务数据质量和管理水平起到基础支撑作用。

近年,随着移动通信、物联网、云计算、人工智能等信息技术在水务行业的应用,北京市水务行业积累了多种来源、内容丰富、类型多样的大量数据资源。水务数据是政务数据的重要组成部分,如何将汇聚的大量水务数据更好地共享、开发、利用,是当务之急。本文提出了北京市水务数据资源目录的分类、元数据模型和数据目录编码方法,并在北京市水务局数据目录编制工作中进行了应用,初步形成了水务数据底账,使数据“看得见,用得上,管得清”。

一、北京市水务数据资源现状

北京市水务局在2021年开展了智慧水务1.0建设,构建了以水务物联感知、数据汇聚共享为核心的智慧水务基础底座。目前,已汇聚至底座的数据主要有以下方面:

①基础数据。水务对象相对稳定的属性信息、主要特征信息、地理空间数据等,如河流、湖泊、水利工程等水务对象的基本属性数据、空间位置等。

②监测感知数据。通过各种监测感知设备采集感知的数据及设施计量数据,包括雨量、河道水情、水质

源利用量、机井取水量、供水量、排水量、用水户用水量等。

③业务数据。水旱灾害防御、水资源管理、水利工程、水生态、水环境等各类核心业务工作中产生的计划、运行、整编、统计数据。

④行政管理数据。水务管理机构日常运行管理过程中收集和产生的数据,包括综合办公、财务与审计、人事与教育、国际合作、科技管理等数据。

二、水务数据目录

《政务信息资源目录编制指南(试行)》对政务信息资源目录作了定义:政务信息资源目录是通过对信息资源依据规范的元数据描述,按照一定分类方法进行排序和编码的一组信息,用以描述各个政务信息资源特征,便于对政务信息资源检索、定位与获取。水务数据是从事水务活动各单位在履行职责和开展业务活动过程中采集、产生、获取的以电子或者其他方式记录的数据资源。水务数据目录是按照元数据标准,就水务业务数据特点进行分类,描述水务数据特征及管理相关属性的数据资源清单,便于水务数据服务与管理。水务数据目录概念模型由元数据、数据分类、数据编码等关键要素构成(见图1)。

三、北京市水务数据分类方法

1. 水务数据分类标准现状

依据国家政务信息资源目录体系的数据分类,政务数据分为21个一级类、133个二级类。国家政务信息资源分类标准面向全部政务数据,将水利数据和农

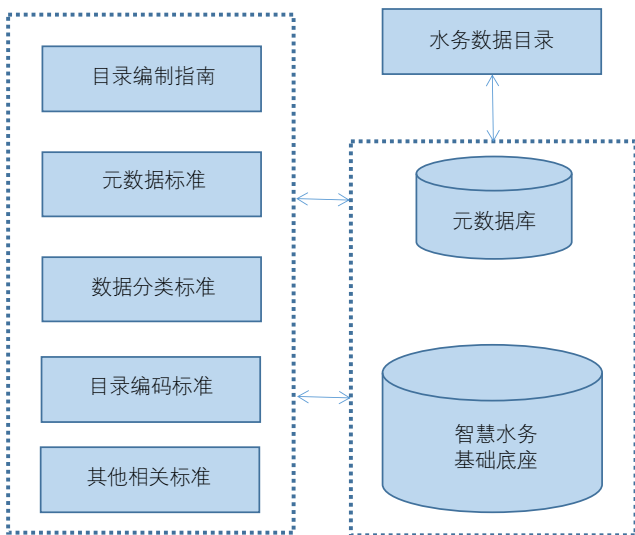


图1 水务数据目录概念模型

业数据合为一个分类,水文与气象、测绘、地震合为一个分类,分类过于宏观,还需各行业进一步细化。《水利信息分类与编码总则》(SL/T 701—2021)设计了水利业务、水利行业、业务阶段三个维度的水利数据分类结构,此分类具有扩展性,各水利部门还需按框架进一步编制分类。2022年,水利部印发《水利数据分级分类指南(试行)》,提出水利数据一级类6个、二级类37个、三级类178个,分类基本涵盖水利数据,但不涵盖全部水务管理数据(水务数据不仅包括水利数据,还包括供水、排水等城市管理数据)。住房和城乡建设部在2015年发布行业标准《城镇供水管理信息系统供水水质指标分类与编码》(CJ/T 474—2015),对供水监测数据进行了分类与编码。山东省、山西省、上海市和内蒙古巴彦淖尔市也发布了本地水利数据分类标准,山东省对水利数据进行主题分类,山西省提出水利数据三维结构,上海市将水利管理对象分类,巴彦淖尔市按照用途和业务属性将水利数据分类。现有水务数据分类相关标准及规范见表1。北京市尚未发布涵盖全部水务数据的分类标准规范,只对用水基础数据进行了分类。为了便于数据检索、共享和开发利用,需要制定北京市水务数据分类方法。

2. 基于职责的线分类法

北京市水务数据由北京市水务局各内设机构在履职过程和业务活动中采集、产生、获取。北京市水务局现有27个内设机构、21个所属单位。27个内设机构主要有三类:行政管理部门,包括办公室、研究室、法制处、政务服务处、规划管理处、投资计划处等;业务管理部门,包括水资源管理处、北京市节约用水办公室、水利工程建设处、供水管理处、污水与再生水管理处、水旱灾害防御处等;内部管理部门,包括财务处、审计

处、人事教育处等。各内设机构职责分工明确,职责不交叉,且编制用语规范。水务数据与职责关联度很高,通过线分类法可以参考内部机构职责,根据业务专题进行分类。

内设机构在履职过程中,会产生基础数据、监测感知数据、业务数据和行政管理数据等四类数据。结合北京市水务局内设机构职责和现有相关标准规范,按照业务管理人员的数据使用习惯,将四类数据按业务细化分类,便于数据的检索、定位及应用。

(1) 分类步骤

①提取职责关键内容及关键词。查找各内设机构职责,分析可能产生的数据资源,提取转化为数据分类关键词,关键词采用描述职责的标准用语。

②查找现有标准规范。查找与业务相关的国家标准、行业标准、地方标准等,对比职责内容,在基础数据、监测感知数据、业务数据、行业管理数据一级分类基础上,细化数据的二级分类和三级分类。

③进行数据分类。在一级分类下,细化水务业务专题为二级分类,细化专题内容为三级分类。

(2) 分类实践案例:供水管理相关数据分类

①提取供水管理处职责关键内容及关键词。供水管理处主要承担北京市供水行业管理工作,“三定”职责为“编制城乡供水发展规划和年度计划并组织实施。拟定供水行业的政策、标准、规程规范并监督实施。指导、监督供水行业的服务质量。监督管理公共供水单位及自建设施供水单位的供水水质和安全运行。组织供水行业特许经营的具体实施工作”。提取供水管理工作的关键内容及关键字为:供水管理、供水规划、年度计划、供水行业政策、供水行业标准、供水行业规范规程、供水监督、公共服务。

表1 现有水务数据分类相关标准及规范

分类	标准规范
国家标准	政务信息资源目录体系 第4部分:政务信息资源分类(GB/T 21063.4—2007)
行业标准及规范	水利信息分类与编码总则(SL/T 701—2021)
	水利数据分级分类指南(试行)(办信息〔2022〕56号)
	城镇供水管理信息系统 供水水质指标分类与编码(CJ/T 474—2015)
地方标准	水利数据服务技术规范(DB37/T 4318—2021)(山东省地方标准)
	水利技术标准体系 总体框架(DB14/T 2727—2023)(山西省地方标准)
	水利信息分类规范(DB1508/T 133—2021)(巴彦淖尔市地方标准)
	用水管理信息系统 基础信息分类和编码规范(DB 11/T 1954—2022)(北京市地方标准)

②查找供水管理相关标准规范。在全国标准信息公共服务平台查找供水相关的数据分类标准。如选取住房和城乡建设部在2015年发布的行业标准《城镇供水管理信息系统 供水水质指标分类与编码》(CJ/T 474—2015),将供水监测数据分为供水厂监测数据及供水管网监测数据,监测内容包括水量、水质、水压等。

③供水数据分类。“供水管理”关键词涉及基础数据、监测感知数据,结合各类标准和水务局数据现状,基础数据主要包括供水单位、供水厂(站)、供水厂(站)附属设施、供水设施项目、供水计量设施等内容,监测感知数据包括监测的水量、水质、水压。“供水监督”“公共服务”关键词,涉及业务数据,包括供水监管、供水统计、公共服务。“供水规划”“年度计划”“供水行业政策”“供水行业标准”“供水行业规范规程”关键词,涉及行政管理数据,包括水务规划、政策法规(见表2)。

(3)北京市水务数据分类

按照上述基于职责的线分类法,分析北京市水务

局内设机构职责,梳理水务数据,设计数据分类。共设三级分类,一级分类为基础数据、监测感知数据、业务数据、行政管理数据四类,进一步细化为38个二级分类,160个三级分类(见表3)。

四、北京市水务数据元数据模型

1. 现有水务数据元数据标准

政务数据元数据标准需要符合政府信息特点,满足实际管理需求,并便于资源利用。水务数据管理在行业上遵照水利数据管理的要求。针对水利数据元数据标准,国内外已有一些研究成果。国际上有ISO 19115系列标准“地理信息—元数据(Geographic information—Metadata)”,主要侧重于地理信息元数据规范。在我国,水利部2010年发布了行业标准《水利信息核心元数据》(SL 473—2010),在政务元数据基础上增加了数据质量评价信息作为可选指标,主要包括数据可信度、数据状况、数据质量信息,可以据此体现数据可信度、可用度;标准规定水利信息

表2 供水管理职责与数据分类对应

提取职责关键词	一级分类	二级分类	三级分类	数据内容
供水管理	基础数据	监测站(厂、点、户)	供水单位	供水单位基本情况、联系方式、资产结构、设施基础情况、运营情况等
			供水厂(站)	城镇供水厂、乡镇供水厂、农村供水站、城镇自建供水设施和水厂等的基本信息、联系方式、水源地、规模等
			供水厂(站)附属设施	供水管网、清洁池、水质监测站点、水压监测站点等基础信息
			供水设施项目	项目基本情况、投资情况、项目设计情况、项目审批情况、项目计划情况、地表水水厂净水工艺、地下水水厂净水工艺
			供水计量设施	取水、出厂水、独立计量区(DMA)水表名称、规格型号、准确度登记、常用流量、生产厂家、安全使用地点、状态、首次检定时间及周期检定时间等基础信息
	监测感知数据	供水	水量	供水厂水量监测数据
			水质	水质监测点水质监测数据
			水压	供水管网水压监测数据
	业务数据	供水	供水监管	供水行业监督管理信息
			供水统计	供水量等统计数据
			公共服务	供水审批等公共服务信息
供水监督	行政管理数据	政策法规与规划	水务规划	城乡供水发展规划和年度计划
公共服务				
供水规划			政策法规	供水行业的政策、标准、规程规范
年度计划				
供水行业政策				
供水行业标准				
供水行业规范规程				

表3 北京市水务数据分类

一级分类	二级分类	三级分类	一级分类	二级分类	三级分类
基础数据	河流湖泊	河流、湖泊、流域	业务数据	水文	整编、预报、预警
	水利工程	水库、水电站、水闸、橡胶坝、泵站、引调水工程、堤防、塘坝、窖池、蓄滞洪区、防汛仓库、其他水利工程		水资源	配置调度、水资源保护、水资源调查评价、取水统计、取水口管理、监督检查、水资源税
	水土保持	水土保持区划、水土保持治理项目、生产建设项目、水土保持单位		水生态环境	水环境治理、水生态保护修复
	监测站(厂、点、户)	水文站、雨量站、水位流量监测站、地下水水位监测站、水质监测站、水生态监测站、蒸发量监测站、墒情站、水源地、地表水取水口、取水户、取水计量设施、供水单位、供水厂(站)、供水厂(站)附属设施、供水设施项目、供水计量设施、居民用水户、非居民用水户、村级用水户、用水计量设施、积水监测站、水土保持监测站(点)、水务视频监控点、其他监测站(点)、污水处理厂(再生水厂)、农村污水处理厂站、污泥处置设施、污水处理厂附属设施、再生水利用工程、雨水滞蓄工程、排水户、排水计量设施		水利工程	水利工程建设、水利工程运行、安全监督
	地理空间数据	基础地图、影像数据、数字高程模型/数字表面模型、倾斜摄影影像/激光点云、建筑信息模型、水下地形、水务专题地理空间数据		水旱灾害防御	水旱灾害防御调度、洪水影响评价
	个人信息	个人基本信息、个人身份信息、个人健康生理信息、个人财产信息、个人位置信息、身份鉴别信息、其他个人信息		水土保持	水土流失治理、水土保持监管
监测感知数据	其他管理对象	其他管理对象		用水	用水计划、用水统计、节水型社会建设
	水文	降水、水位、流量、泥沙、地下水、蒸发、墒情、其他水文监测数据	行政管理数据	供水	供水统计、供水监管、公共服务
	水资源	地表水资源量、地下水资源量		排水	设施运行监管、污水处理费、海绵城市建设
	水生态环境	水质、水生生物		河湖长制	河湖长、河湖事件管理
	水利工程	运行监测、安全监测		专项调查与统计	水务普查、水务调查、水务统计
	水旱灾害防御	积水		综合办公	公文处理、档案管理、会务管理、信访工作、年鉴公报、其他
	水土保持	坡面径流、水土流失		人事教育	编制管理、机构改革、干部人事管理、教育培训
	用水	居民用水户用水、非居民用水户用水、村级用水户用水		财务与审计	财务管理、资产监管、领导干部经济责任审计、财务审计、专项审计和稽查
	供水	水量、水质、水压		纪检与监察	纪律检查、党风廉政建设
	排水	污水、再生水、污泥		政策法规与规划	政策法规、水务规划
	水务视频监控	视频影像		科技与信息化	科技项目及成果、水务标准化、对外合作与交流
	其他	其他		政务服务	行政许可、行政确认、其他职权、行政复议、行政检查、其他政务服务
				公共服务	接诉即办、其他公共服务
				移民管理	移民安置、移民后期扶持

核心元数据由8个元数据子集构成,包括标识信息、内容信息、覆盖范围信息、限制信息、维护信息、数据质量信息、参照系信息、分发信息,其中标识信息为必选子集,其余为可选子集。有相关研究提出在《水利信息核心元数据》基础上增加属性信息和结构信息等元数据子集,用于描述水利对象关联属性的概况信

息和数据库连接信息等。地方层面,内蒙古巴彦淖尔市在2021年发布《水利信息资源目录与元数据规范》(DB1508/T 137—2021),对水利数据目录编制和元数据标准进行了规范。

2. 北京市水务数据元数据模型设计

结合北京市水务数据管理现状,在相关标准基础上,细化数据共享开放范围,建立数据目录与数据库信息的链接,增加数据质量评价信息,提出北京市水务数据核心元数据模型,主要包括基本信息、共享开放属性、结构化数据存储信息、数据质量信息、其他五部分(见图2)。

①基本信息。包括数据资源名称、数据资源编码、数据摘要(包括采集数据项、采集范围等)、数据分类、数据分级(参照《政务数据分级与安全保护规范》《水利数据分类分级指南(试行)》等相关标准规范)、关联数据资源编码、更新周期、站点(涉及采集站点的数据资源需标明站点名称)、数据获取途径(包括人工填报、数据共享、测站接入等)、数据开始日期、数据量、数据格式、数据提供方、监督处室。

②共享开放属性。包括共享属性(按共享条件分为有条件共享、无条件共享、不予共享,按共享范围分为水务系统内部共享、政务系统内部共享)、共享条件(有条件共享的数据资源需标明)、不予共享的依据(不予共享的数据资源需标明)、开放属性(分为有条件开放和无条件开放)、开放条件(有条件开放数据资源需标明)。

③结构化数据存储信息。包括前置库表信息(数据资源提供单位提供的数据库表名称、字段描述、取值范围等)、数据服务库表信息(水务基础底座提供数据服务的数据库表名称、字段描述、取值范围等)。

④数据质量信息。包括数据及时率(按数据更新频率送达前置库的情况)、数据完整率(按约定数据量送达前置库的情况)、数据应用情况(使用了本数据的应用系统名称)。

⑤其他。包括备注等信息。

五、数据目录编码

基于上述数据分类体系,对北京市水务数据目录进行唯一编码。数据目录编码由四部分数字组成,第一部分(第一位)为一级分类码,第二部分(第二、三位)为二级分类码,第三部分(第四、五、六位)为三级分类码,第四部分(第七、八、九、十位)为顺序码(见图3)。

六、应用及展望

在北京市水务数据目录编制工作中,按照数据资源现状进行填报,对数据分类、元数据模型、数据资源编码设计进行了应用,编制基础数据目录225条,监测感知类数据目录159条,业务类数据193条,行政管理类数据目录66条,形成了《北京市水务数据目录》。应用结果表明,本文构建的基于职责的线性数据分类方法,能够很好地进行水务数据分类,分类较稳定且边界清晰不交叉,适用于不同层级、不同门类的政务数据资源分类,为厘清数据分类脉络提供参考。本文定义的元数据模型能够较好描述水务数据资源,为数据检索、共享、开发利用提供基础。通过数据目录,数据使用人员可以快速清晰了解数据内容,数据管理人员可以分析数据关联性,为数据挖掘、数据治理工作提供便利,为数据资产入表、数据资源确权工作提供基础台账。将来,可采用人工智能方法进行元数据提取,提高目录编制效率和质量,同时需建立相关规章制度、

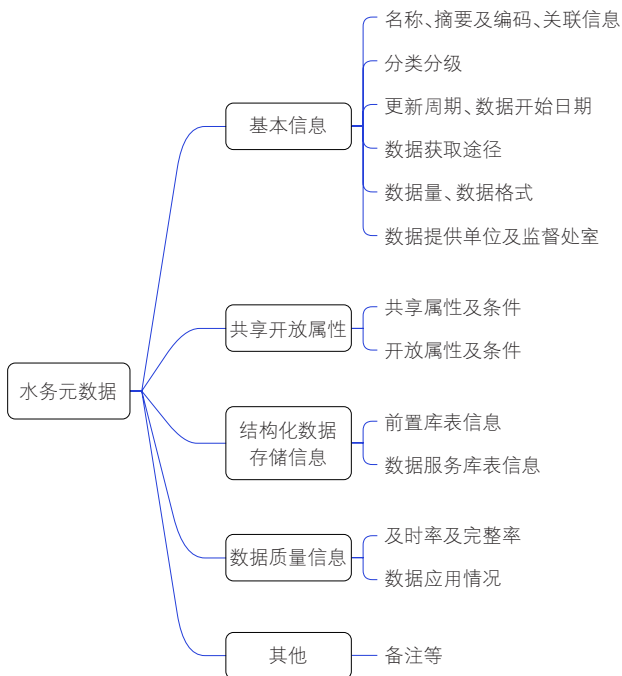


图2 北京市水务数据核心元数据模型结构

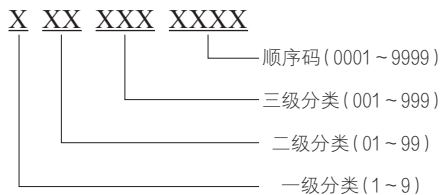


图3 数据目录编码结构

标准体系,保障数据目录常态更新,提高可用性。

参考文献:

- [1] 国务院.关于加强数字政府建设的指导意见[A].2022.
- [2] 国务院办公厅.全国一体化政务大数据体系建设指南[A].2022.
- [3] 中共中央办公厅,国务院办公厅.关于加快公共数据资源开发利用的意见[A].2024.
- [4] 曾建勋.强化数据目录体系研究[J].农业图书情报学报, 2024, 36(9):102-103.
- [5] 王平, 苏杭.面向国家重大战略的中国目录学重塑性发展[J].图书馆建设, 2024(6):158-165.
- [6] 刘皓宇, 吴洪涛, 周俊杰, 等.自然资源数据资源目录编制研究[J].自然资源信息化, 2022(5):52-57.
- [7] 杨颜如, 冯丽, 赵连静.农业数据资产的分类目录研究[J].农业展望, 2023, 19(11):67-73.
- [8] 陈辉, 叶培春, 颜祥成, 等.基于元数据的作战数据资源目录体系构建[J].火力与指挥控制, 2023, 48(10):191-197.
- [9] 王茜, 葛召华.山东省水利数据中心资源目录体系建设[J].山东水利, 2011(9):5-7.
- [10] 李锴锋.辽宁省水利数据资源目录编制方式及具体流程探讨[J].水利技术监督, 2023(6):24-26+32.
- [11] 潘崇伦.水务海洋数据的智慧治理[J].上海信息化, 2023(10):34-37.
- [12] WIRTZ B W, WEYERER J C, BECKER M, et al.Open government data: A systematic literature review of empirical research[J].Electronic Markets, 2022, 32:2381-2404.
- [13] ALBERTONI R, BROWNING D, COX S, et al.The W3C Data Catalog Vocabulary, Version 2: Rationale, Design Principles, and Uptake[J].Data Intelligence, 2024, 6(2):457-487.
- [14] AL RAHAMNEH A, SERRANO HERNANDEZ A, FAULIN J.The Impact of Integrating Open Data in Smart Last-Mile Logistics: The Example of Pamplona Open Data Catalog[J].Sustainability, 2025, 17(2):408.
- [15] 郝仲勇.打造数字水务 助力智慧城市 全面推进北京市智慧水务1.0建设[J].北京水务, 2021(S2):1-4.
- [16] 国家发展改革委, 中央网信办.政务信息资源目录编制指南(试行)[A].2017.
- [17] 国家质量监督检验检疫总局, 国家标准化管理委员会.政务信息资源目录体系 第4部分:政务信息资源分类:GB/T 21063.4—2007[S].2007.
- [18] 水利部.水利信息分类与编码总则:SL/T 701—2021[S].2021.
- [19] 住房和城乡建设部.城镇供水管理信息系统 供水水质指标分类与编码:CJ/T 474—2015[S].2015.
- [20] 山东省市场监督管理局.水利数据服务技术规范:DB37/T 4318—2021[S].2021.
- [21] 山西省市场监督管理局.水利技术标准体系 总体框架:DB14/T 2727—2023[S].2023.
- [22] 巴彦淖尔市市场监督管理局.水利信息分类规范:DB1508/T 133—2021[S].2021.
- [23] 巴彦淖尔市市场监督管理局.水利信息资源分类与编码总体规范:DB1508/T 136—2021[S].2021.
- [24] 上海市市场监督管理局.水务信息管理 第1部分:分类与编码:DB31/T 362.1—2024[S].2024.
- [25] 北京市市场监督管理局.用水管理信息系统基础信息分类和编码规范:DB11/T 1954—2022[S].2022.
- [26] 张晓娟, 唐长乐.我国政府信息元数据标准体系框架构建及其应用流程[J].信息资源管理学报, 2018, 8(3):25-36.
- [27] Geographic Information—Metadata: ISO 19115: 2023[S].2023.
- [28] 水利部.水利信息核心元数据:SL 473—2010[S].2010.
- [29] 成建国, 冯钧, 杨鹏, 等.水利数据资源目录服务关键技术研究[J].水利信息化, 2014(6):18-21+35.
- [30] 巴彦淖尔市市场监督管理局.水利信息资源目录与元数据规范:DB1508/T 137—2021[S].2021.

责任编辑 王 慧