

知识图谱技术在全国取用水平台的应用

王晨雨,刘庆涛,沈红霞

(水利部信息中心,北京 100053)

摘要: 全国取用水平台整合是实现取用水业务“一网统管”的重要途径,但因拟整合的取用水业务系统繁多,信息资源分散,平台整合中存在一数多源、标准不统一、户证点关系不清晰等现实问题,为此尝试提出将知识图谱技术应用于平台整合过程,建设一套全国统一的数据库表结构标准,形成一套真实且唯一的户证点关系,构建一套智能且有效的监管数据产品。采用自顶向下的知识图谱构建方式,根据各系统的结构化数据,构造模式层中的本体及其相互关系,形成对应的概念模型和规则关系,再依照此模式从数据中抽取实体及关系,进行数据融合,构造数据层,并及时进行知识更新,完成全国取用水平台知识图谱的构建。最后基于统一规范的数据库表结构标准,依据对取用水户取用水行为的监管需求,开展时空尺度的数据信息统计分析,构建6类监管产品,形成取用水管控一张图。

关键词: 全国取用水平台;知识图谱;系统整合;自顶向下

中图分类号: TV213.4; TP18

文献标识码: A

文章编号: 1674-9405(2023)04-0007-07

0 引言

“十二五”以来,国家逐步加大了水资源监管信息化建设力度,先后启动了国家和地方水资源监控能力建设一期、二期项目,初步建成了覆盖国家、流域、省级的水资源信息管理系统,基本建立了取用水监控体系。“十三五”期间先后建成全国取水许可电子证照、用水统计调查直报管理、取水工程(设施)核查登记等系统,各地也结合水资源管理需要,建成了取水许可审批、取水计划、水资源税费及地方水资源管理等系统。这些系统在一定的领域和时期,发挥了特定作用,有效支撑了水资源管理、水资源税改、取水许可管理、取水工程专项整治等业务^[1]。取用水监管作为水资源管理的重要内容,通过建设这些系统,监管能力得到有效提升。

但因水资源管理数据资源涉及水利部、流域机构、地方水利部门,以及行政审批局、税务等多部门,信息资源类型多样,存储管理分散,各自独立,标准不统一,平台不联通,业务不协同,无法满足社会服务和监管决策需求;涉及取用水业务系统繁多,信息资源分散,数据共享难,存在一数多源、异常值偏多的情况,制约水资源数据算据效力的发挥,导致数据深度融合分析及业务服务支撑能力不足。对标“十六字”治水思路、国家“以水而定,量水而行”、水资源刚性约束制度^[2]、水利高质量发展、智慧水利建设^[3],以及

政务信息系统整合^[4-5]等管理要求,从全局和根本上支撑和服务国家、流域、区域水资源精细化管理,还存在一定的差距。

为解决上述问题,水利部依托全国一体化在线政务服务平台,着力完成各层级取用水管理业务信息系统整合,建成国家、省两级取用水管理平台,实现数据资源的有效汇聚、充分共享,全过程留痕,全流程监管,提升在线监管整体水平,形成水资源监管一张图,全面提升数据监管、智慧监管能力,实现取用水业务“一网统管”。为优化全国取用水平台整合过程,本研究以知识图谱技术为主要抓手,就系统整合中亟须突破的一数之源、数据融合、监管产品等难点进行探索研究。

1 知识图谱

1.1 有关概念

知识图谱的早期理念源于万维网之父 Tim Berners-Lee 关于语义网的设想,旨在采用图结构建模和记录世界万物之间的关联关系和知识,以便有效实现更加精准的对象级搜索^[6]。应现代知识图谱规模化扩展的要求,通常采用以三元组为基础的较为简单实用的知识表示方法,基本组成单位为“实体-关系-实体”“实体-属性-属性值”三元组,实体间通过关系相互联结,构成网状的知识结构^{[7]2-3[8]}。知识图谱的

相关技术已经在搜索引擎、智能问答、语言理解、推荐计算、大数据决策分析等众多领域得到广泛的实际应用。近年来,随着自然语言处理、深度学习、图数据处理等众多领域的飞速发展,知识图谱在自动化知识获取、知识表示学习与推理、大规模图挖掘与分析等领域又取得了新进展,并在行业(领域)^[9]得以有效应用,如自然灾害应急知识图谱^[10]、电力系统知识图谱^[11]、海洋知识服务体系^[12]等。

1.2 构建方法

知识图谱由模式层和数据层 2 部分构成^[13]。知识图谱在构建方法上分为以下 3 种^{[7]6-8}: 1) 自底向上。从底层数据源中提取实体、属性和关系,加入到知识图谱的数据层,然后将这些知识要素进行归纳组织,逐步往上抽象为概念,最后形成模式层。2) 自顶向下。从最顶层的概念开始构建模式层本体,然后细化概念和关系,形成结构良好的概念层次树,将知识抽取得到的实体匹配填充到构建的模式层本体中。3) 二者混合。在知识抽取的基础上归纳构建模式层,之后可对新的知识和数据进行归纳总结,从而迭代更新模式层。

2 全国取用水平台知识图谱

2.1 需求分析

全国取用水平台依托水利部政务云平台建设,需对原分散在国家水资源信息管理、取水许可审批、取

水许可电子证照、用水统计直报、水资源税水量核定、取水计划、取水计量监管等多个系统的信息进行融合,但各系统内有关取用水户的信息都不全面,单个系统均不能提供满足管理部门对于取用水户、许可审批、计划执行、计量设施安装、计税水量核定等情况开展监管的全面信息,同时,涉及系统比较多,逻辑关系复杂。前期初步剔除多个系统 18 个模块 80 张表的不相关信息,整合为 8 个模块 18 张表,如表 1 所示。各模块间主要关联关系图,如图 1 所示,其中取水许可审批信息来源于取水许可审批系统,证照管理信息来源于取水许可电子证照系统,取水计划信息来源于取水计划系统,水资源税费信息来源于水资源税水量核定系统,用水统计信息来源于用水统计直报系统,计量管理信息来源于取水计量监管系统,水量监测信息来源于国家水资源信息管理系统,行政区划信息在多系统均有涉及。虽已初步整合,但模块内依然包含大量过程类数据信息,如证照管理模块中的发证机关、发证机关代码等无关取用水直接监管数据产品需求的属性数据。各模块间还存在一数多源、标准不统一、计量单位不一致等问题,如首次发证证件编号属性,在水资源税费、取水计划、计量管理等其他模块也存在,且因存在同一社会信用代码对应多个证件编号现象,造成各模块间同一属性的属性值不尽相同,导致户证点关系不准确、不唯一,户对应对象(户)数量不一致,名称不匹配等情况。

表 1 取用水平台表结构

序号	表结构类别	表标识	表名
1	取水许可审批信息	EC_APPLY_INFO	取水许可审批数据表
2	证照管理信息	EC_BASE_INFO_WR	取水许可电子证照基本信息
3	取水计划信息	WR_PLAN_MONTH_WATER	证照月度取水计划表
4		WR_PLAN_YEAR_WATER	证照年度取水计划表
5		WR_PLAN_AREA_GROUP	区域取水计划表
6	水资源税费信息	WI_WATER_SUPERVISION_MAIN	监审意见书主表
7		WI_WATER_SUPERVISION_SUBSET	监审意见书子表
8		WI_TAX_FUND	税源信息表
9		WI_TAX_CRITERION	税额标准表
10	用水统计信息	WRTJ_ML_ALL_VIEW	直报名录的基本信息表
11		WRTJ_TB_ALL_VIEW	直报调查信息表
12	计量管理信息	WI_STD_BK_MEASURE_B	计量设施信息表
13		WI_WR_DAY_W_R	在线日水量表
14		WI_MEASURE_RCD_R	远程抄表信息表
15	水量监测信息	WR_MP_B	国控监测点基本信息表
16		WR_DAY_W_R	监测点日水量信息表
17		WI_EXAMINE_RELATION	点证关系表
18	行政区划信息	GW_SYS_ADDVCD	行政区划表

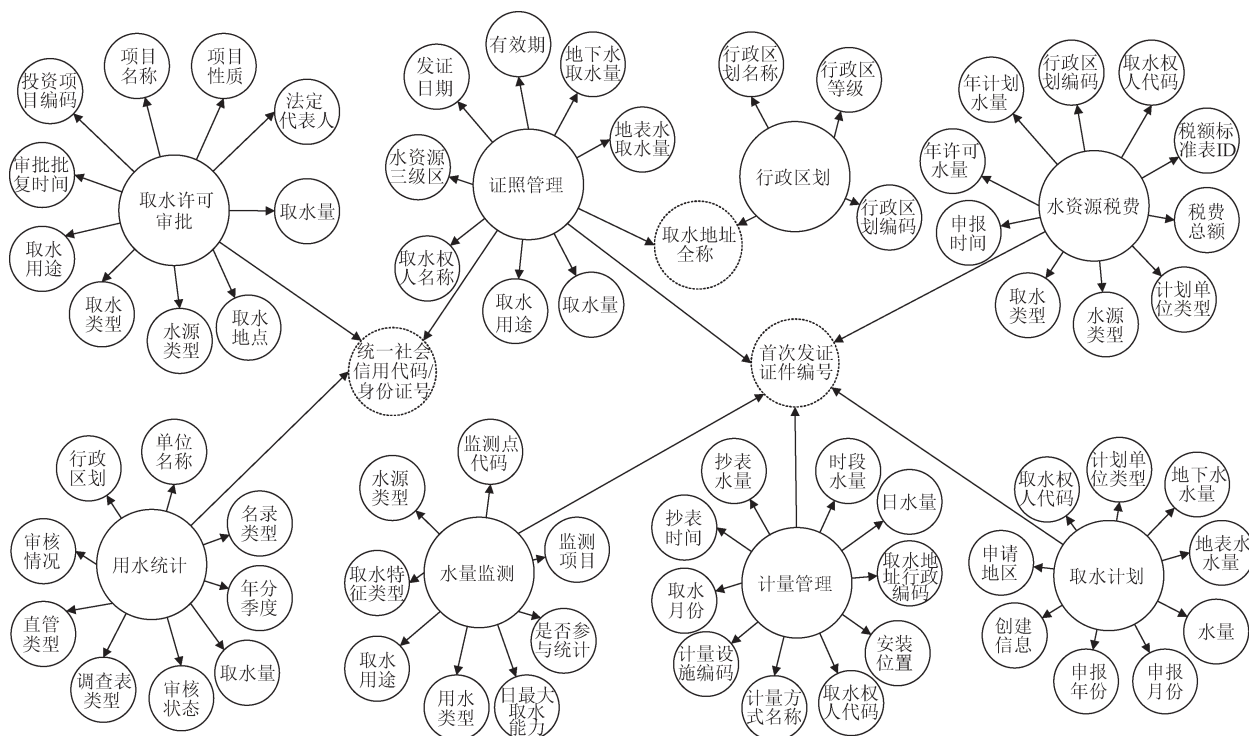


图 1 各模块间主要关联关系图

因此,利用知识图谱技术,以取用水户为核心监管对象进行系统性梳理,整理数据源,建设一套全国统一的数据库表结构标准,形成一套真实且唯一的户证点关系,构建一套智能且有效的监管数据产品集,是实现“一网统管”“互联网+监管”系统功能需求的重中之重,其中统一规范的数据库表结构标准建设是关系到全国取用水平台整合成败的关键环节。

2.2 技术架构

全国取用水平台知识图谱是将知识图谱技术应用于取用水管理的技术形式,旨在梳理多模块对象间的关系,挖掘各模块关键信息项,剔除数据源重复项、差异项、不确定项等,更好地支撑全国取用水平台整合及各模块间的知识推理,实现取用水管理辅助分析及决策支持,属于在取用水领域开展知识图谱技术应用的尝试。基于全国取用水平台初步整合后各模块的数据库表结构标准、领域类知识图谱构建特性^[14-15],采用自顶向下方式构建全国取用水平台知识图谱,技术架构如图2所示。

3 全国取用水平台知识图谱的构建

3.1 模式层构建

模式层是知识图谱的概念模型和逻辑基础^{[7]3},全国取用水平台知识图谱模式层构建,主要根据现有的

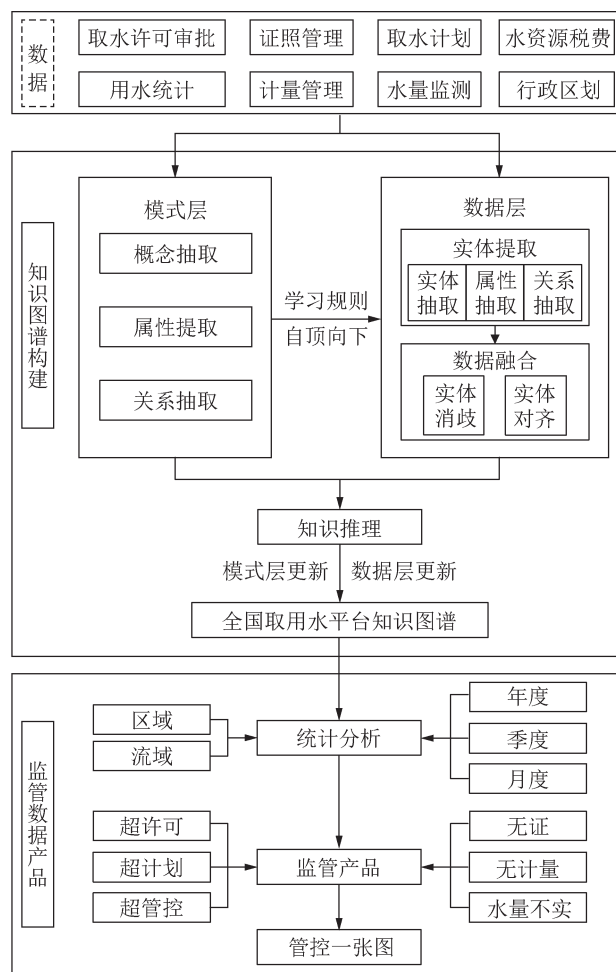


图2 全国取用水知识图谱技术架构图

取水许可审批、证照管理、统计直报等模块信息整合后的结构化数据库,针对全国取用水平台整合要求,构造模式层的本体库^{[16]584-585},进行本体概念定义、层次关系划分,属性关系提取,以及概念间语义关系定义。全国取用水平台知识图谱的模式层构建主要是

建立取水证照、取水计划、取用水监测、取用水计量、用水统计、取用水税等6类核心要素的取用水综合本体,利用本体的思想构建全国取用水平台知识图谱的模式层如图3所示,可为数据层中具体实例的要素属性分解奠定理论框架基础。

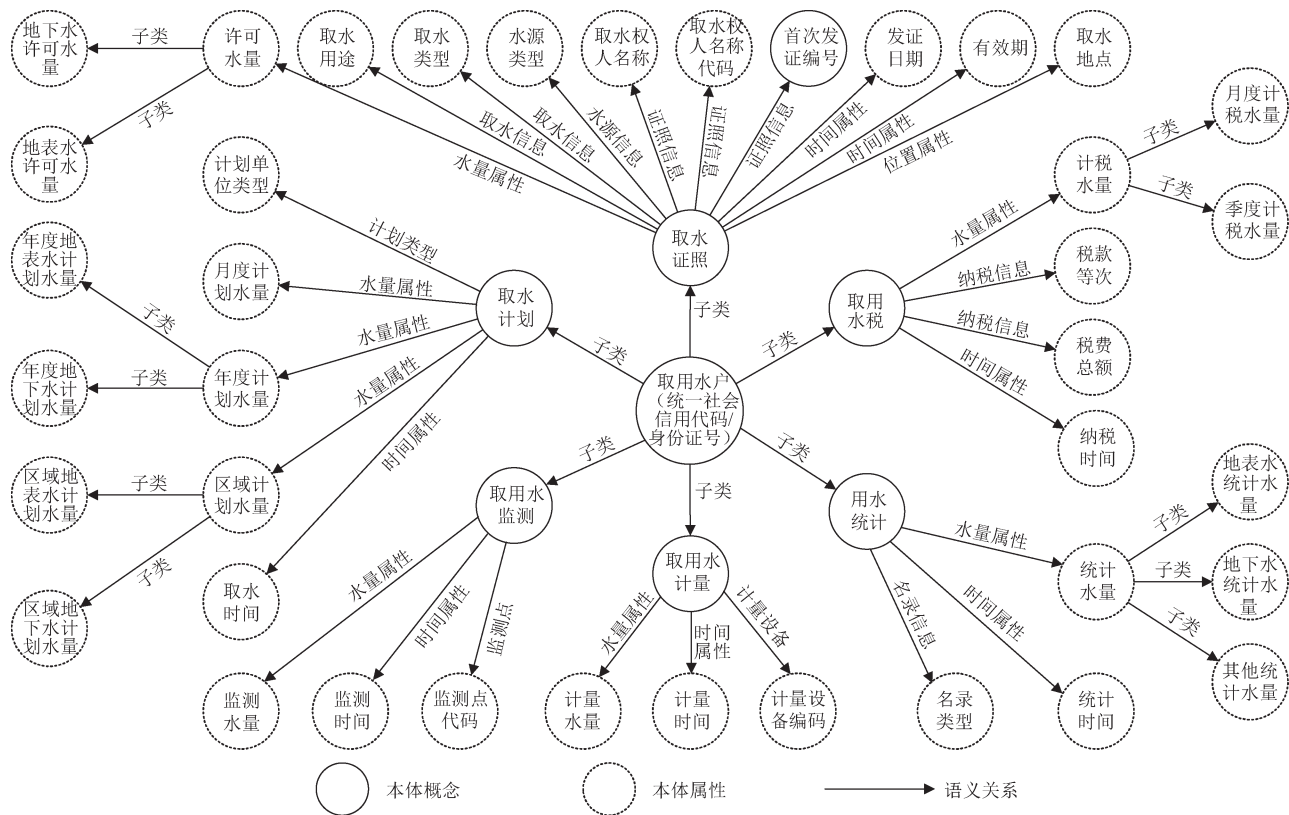


图3 全国取用水平台知识图谱模式层图

3.1.1 概念抽取

概念抽取用于获取原数据库中的概念在取用水平台的概念定义及概念层次分类的集合。概念定义包括延续和重新定义2种形式:对各表中名称、语义关系相同,整合至全国取用水平台不会发生冲突歧义的概念,可采用延续定义,如统一社会信用代码、取水权人代码、发证证件编号等;在各表中名称相同、语义关系不相同,整合至全国取用水平台产生歧义的概念,则采用重新定义,如多个系统中均含有“取水量”名称,在平台整合中需重新定义证照管理模块中的取水量为许可水量、用水统计模块中的取水量为统计水量、水量监测模块中的取水量为监测水量等。概念层级分类可表示为取水计划中月度、年度,区域、流域等时空尺度层级,以及统计水量中包含地表水、地下水、其他等统计水量的从属关系层级。

3.1.2 属性提取

属性提取用于获取原数据库中可服务于“一网

统管”的平台整合目的本体的相关属性,如时间、地点、水量等属性。提取过程中以具有一定法律效力的证照管理模块信息所含属性为基准,提取其他模块中不同属性进行补充,同一属性及不直接服务于监管数据产品的信息不提取。如提取某用水户同一本体在证照管理模块基本信息中包含的时间、地点等属性信息,则不再提取其他模块相关信息。

3.1.3 关系抽取

关系抽取用于构建概念之间的关系。对于全国取用水平台中的概念,通过识别概念间的语义关联关系,将关系抽取出来。如平台整合以取用水户为核心监管对象,以电子证照模块的首次发证证件编号、统一社会信用代码/身份证号为“根”,基于证照管理、取水计划、水量监测、计量管理、用水统计、水资源税等模块信息,梳理户、证、点、水量(许可、计划、监测、计量、统计、计税等水量)对应关系,抽取取用水户真实且唯一的关联关系。

3.2 数据层构建

数据层是模式层的实例化和事实应用,全国取用水平台数据层的构建,是基于已有数据库、模式层良好的概念层次知识体系和规则关系,实现数据层对模式层的映射,主要包括实体抽取、数据融合。抽取的实体可以根据对应的概念,按照模式层的关系层次组成实体间的结构关系。

3.2.1 实体提取

实体提取是从数据中抽取实体、属性与实体间的相互关系。针对结构化数据字段定义明确的特点及监管数据产品的需求,基于构建的全国取用水平台知识图谱模式层,设计相应字段的映射规则,从关系数据库中直接抽取实体名称及属性信息,对象之间的语义关系可通过数据库的字段链接进行映射。全国取用水平台整合中实体、属性及语义关系的提取,一般以具有一定法律效力的证照管理模块所含的信息为基准,并合并实体在其他模块的属性、与其他实体的关联关系。根据管控需要,实体需着重以下几个方面的提取:1)户、证的对应关系梳理提取,以证照管理模块中的首次发证编号、用水统计模块中户的目录为主,其他模块户证信息为辅,完成取用水户目录的全量提取;2)户、证、点、水量、计量设施等全链条的对应关系梳理提取,基于户、证对应关系的梳理,完成各模块、监测或计量等水量对应的取水点及水量计量设施等关系的梳理提取。

3.2.2 数据融合

由于数据来源和质量不同,信息抽取得到的数据层中可能存在大量的冗余和错误,因此需要通过数据融合对抽取得到的数据进行梳理和规范化整合。数据融合的关键步骤包括实体对齐^{[7]7}和实体消歧^{[16]588-589}:

1) 实体对齐。实体对齐主要解决同一实体采用不同表达方式的问题,将同一实体的不同表达方式归一化表示,解决一数多源问题。例如,相同身份标识的用水户,在证照管理模块中用户名和水量监测模块中用户名可能不一致,实体对齐过程就是将不同模块抽取到的不同用户名明确为一个用户名。

2) 实体消歧。实体消歧主要解决不同实体采用相同名称的问题,从而建立准确的实体链接。在取用水平台的实际语言环境中,存在某个名称对应多个具体实体的问题。例如,用水统计模块的“统计水量”、水资源税模块的“计税水量”在各自原系统中均使用“水量”代表,平台整合后易混淆,实体消歧过程就是根据水量所在不同模块及关联关系,区分不同模块的水量,并准确链接。

3.3 知识推理

知识推理^{[16]592-593}是针对知识图谱中已有事实或关系的不完备性,挖掘或推断出未知或隐含的语义关系。全国取用水平台知识图谱中的知识推理,通过对已有知识或关系的逻辑分析,制定推理规则,并从各模块中提取相关知识,从而挖掘或推断出新的知识或语义关系。如超许可、超计划等监管业务规则,一般提取取用水户的计量水量信息,与对应的许可、计划水量等信息指标值比对,大于指标值即判断为疑似违规行为;无证取水监管规则为存在于用水统计模块的取用水户,未在证照管理模块里溯源的,即判断为疑似无证取水行为。取用水户关系梳理及监管产品逻辑关系如图4所示。

3.4 知识更新

全国取用水平台中的取用水户监测计量、基础及业务管理等信息内容在不断增加和更新,平台整合完成后还需要动态构建和迭代更新,不断增加新的知识、删除旧的知识并相应调整知识图谱的结构,保障知识的时效性。知识更新有以下2种层次^[17]:

1) 模式层更新。当新增的知识中包含了概念、关系、属性及其类型变化时,需要在模式层中更新知识图谱的数据结构,包括对概念、关系、属性及其类型的增、删、改操作。例如,随着监管水平的提升,监管需求的增加,增添新的本体概念、属性等。

2) 数据层更新。主要指新增实体或更新现有实体的关系、属性值等信息,更新对象为具体的知识(如三元组),更新操作一般通过知识图谱构建技术自动化完成。在进行更新前,需要经过知识融合等步骤,保证数据的可靠性和有效性。例如,证照过期更新,形成监管问题清单核验后重新上报的数据更新。

4 全国取用水平台知识图谱的监管数据产品

基于知识图谱技术,全面构建全国取用水平台基础信息数据库,形成标准统一、联通共享、一数一源的取用水平台数据库表结构标准,探索全国取用水平台中开展统计分析、构建监管数据产品集、形成水资源管控一张图等的典型应用。

4.1 统计分析

基于全国取用水平台数据库表结构标准,分别开展时间和空间尺度的计划水量、许可水量、用水总量、地下水管控指标、电子证照数量、取用水户户数等数据统计,时间尺度分为月度、季度、年度3种类型,空间尺度分为流域/区域类型。基于基础数据的统计,

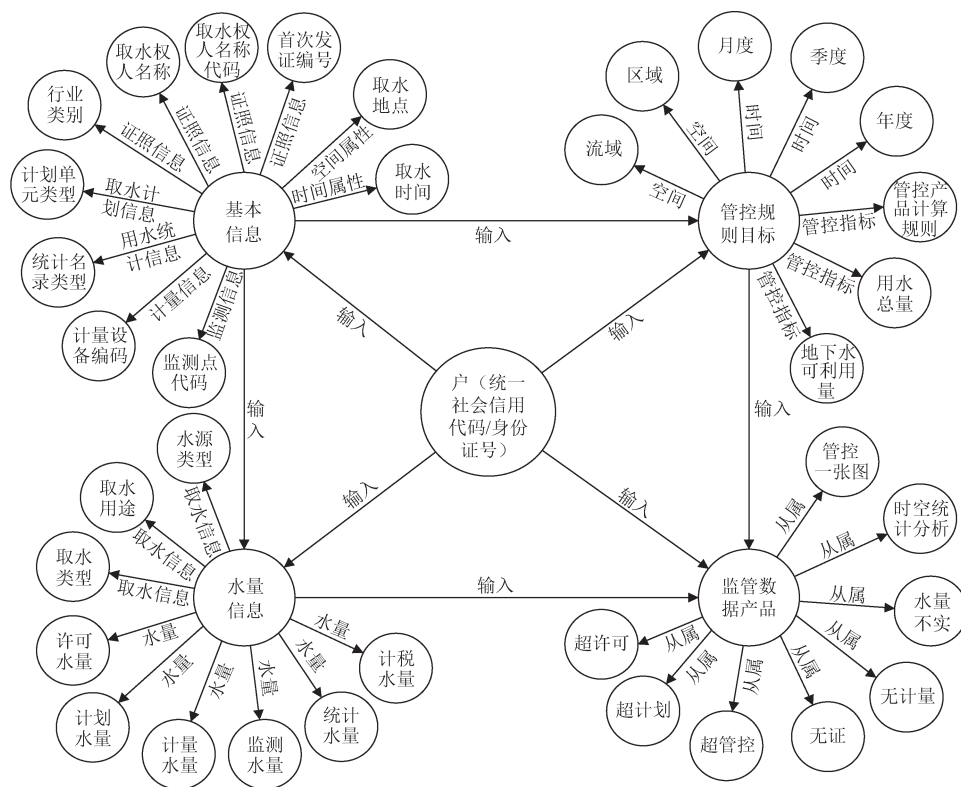


图4 取用水户关系梳理及监管产品逻辑关系图

可加强流域/区域取水水源供水、重点行业季度或年度用水变化趋势监测分析,强化流域/区域水资源开发、经济社会用水、水资源承载能力发展趋势研判,助力取用水总量管控、“以水四定”和经济社会发展涉水决策,提升管理的科学性、预见性和有效性。

4.2 监管产品

基于知识图谱技术对户、证、点信息进行融合比对分析,确定户、证、点真实且唯一性的“血缘关系”。根据取用水户取用水行为监管需求,以及知识推理规则,提取相关监测、统计信息,对比相应监管指标,评价取用水户行为是否存在疑似超许可、超管控、水量不实等违规行为。根据监管需求分为6类监管评价产品。

4.2.1 超许可监管产品

分为户、区域2种监管数据评价产品。对户的监管,技术路线为从数据库中提取被评价取用水户计量水量或监控水量年度总量信息,比对取水证照模块的许可水量信息(一户多证情况为各证许可量之和),如前者大于后者即判断为户疑似存在超许可取水行为。对区域的监管,按照“行政区域内批准取水的总水量,不得超过流域管理机构或者上一级水行政主管部门下达的可供本行政区域取用的水量”的规定^[18],技术路线为从数据库中提取被评价区域从河道内和地下直接取水的取用水户(数据库中被标识为无多级取水

关系的取用水户)的许可水量总量信息,比对区域用水总量指标,如前者大于后者即判断为区域疑似存在超限审批违规行为。

4.2.2 超计划监管产品

对于具有一定规模量以上的取用水户,按有关要求填报取水计划,监管产品同样分为户、区域2种监管数据评价产品。对户的监管,技术路线为从数据库中提取被评价取用水户计量水量或监控水量月、年度总量信息,比对取水计划模块的月、年度计划水量信息,如前者大于后者即判断为户疑似存在“超计划”取水行为。对区域的监管,技术路线为从数据库中提取被评价区域从河道内和地下直接取水的取用水户(数据库中被标识为无多级取水关系的取用水户)月、年度计量水量总量信息,比对取水计划模块上报的月、年度区域总量指标,如前者大于后者即判断为区域疑似存在超计划取水行为。

4.2.3 超管控监管产品

对于已明确地下水取用水量管控指标的13个省,评价区域监管产品,技术路线为从数据库中提取该区域各取水水源为地下水的取用水户计量水量或监控地下水量年度总量信息,比对该区域地下水取用水量管控指标,如前者大于后者即判断为区域疑似存在超管控取水行为。

4.2.4 无证监管产品

对于取用水户无证取水监管产品,技术路线为从数据库中提取用水统计模块取用水户相关信息,以区域为单元比对取用水证照模块总数、基础信息等,对于在取水证照模块中未能溯源的取用水户和取水口信息,判断为疑似无证取水行为。

4.2.5 无计量监管产品

对于取用水户无计量取水监管产品,技术路线为从数据库中提取用水统计模块取用水户水量信息,以及取用水计量模块水量信息,比对两者水量信息,对于在用水统计模块中存在统计水量信息,而在取用水计量模块中未能溯源到对应取用水户的计量水量信息,判断为疑似无计量取水行为。

4.2.6 水量不实监管产品

对于取用水户计税水量监管产品,技术路线为从数据库中提取取水税模块季度、年度水量信息,比对用水统计模块统计水量信息,以及水量监测模块监测水量信息,对于水量信息数据不一致的,形成问题清单发至各省重新核对数据;如核对后计税水量明显低于统计或监测水量的,判断为疑似水量不实行为。

4.3 管控一张图

全国取用水平台与水利一张图对接,提高监管数据可视化水平,提升智慧监管能力。基于全国取用水平台基础信息数据库、监管产品的业务规则,汇集取用水户、取水口(地表、地下)取用水监测站点、取水许可电子证照、用水名录、计量设施、水量等各类基础信息,以及6类监管产品成果信息,分区域、时间进行展示,形成直观明晰的取水管控一张图。

5 结语

全国取用水平台整合是加强取水管理、严格水资源源头管控、强化水资源刚性约束的一项重要考核任务。目前正在积极推进水利部部本级平台建设和与省级平台的对接工作。本研究以知识图谱技术手段,探索涉及水资源管理的多系统集成、数据融合,挖掘各系统业务关联关系,将分散和独立的信息系统整合为一个互联互通、业务协同、信息共享的大系统,全面提升取水监管、智慧监管能力,将为实现取水管理“四预”,实施水资源刚性约束制度,促进生态文明建设和高质量发展提供有力支撑。

参考文献:

[1] 沈红霞,王圆圆,王旖.全国取用水管理信息系统建设[J].

水利信息化,2022(5):74-78.

- [2] 中华人民共和国水利部.加快建立水资源刚性约束制度 奋力推进“十四五”水资源管理工作[EB/OL].[2023-02-20].http://szy.mwr.gov.cn/zyzt_25278/1/02/202106/t20210601_1520892.html.
- [3] 水利部网络安全与信息化领导小组办公室.智慧水利建设顶层设计(水信息(2021)323号)[A].北京:中华人民共和国水利部,2021:9-10.
- [4] 中华人民共和国国务院办公厅.国务院办公厅印发《政务信息系统整合共享实施方案》[EB/OL].[2023-02-20].http://www.gov.cn/xinwen/2017-05/18/content_5195038.html.
- [5] 国务院.国务院关于加快推进政务服务标准化规范化便利化的指导意见[J].中华人民共和国国务院公报,2022(8):29-35.
- [6] 漆桂林,高桓,吴天星.知识图谱研究进展[J].情报工程,2017,3(1):4-25.
- [7] 黄恒琪,于娟,廖晓,等.知识图谱研究综述[J].计算机系统应用,2019,28(6).
- [8] 朱木易洁,鲍秉坤,徐常胜.知识图谱发展与构建的研究进展[J].南京信息工程大学学报(自然科学版),2017,9(6):575-582.
- [9] 刘烨宸,李华昱.领域知识图谱研究综述[J].计算机系统应用,2020,29(6):1-12.
- [10] 杜志强,李钰,张叶廷,等.自然灾害应急知识图谱构建方法研究[J].武汉大学学报(信息科学版),2020,45(9):1344-1355.
- [11] 高海翔,苗璐,刘嘉宁,等.知识图谱及其在电力系统中的应用研究综述[J].广东电力,2020,33(9):66-76.
- [12] 宋丽丽,姜晓轶,赵龙飞,等.领域知识服务体系构建与实践——以海洋知识为例[J].自然资源信息化,2023:1-9.
- [13] 徐增林,盛泳潘,贺丽荣,等.知识图谱技术综述[J].电子科技大学学报,2016,45(4):589-606.
- [14] 倪奕.基于知识图谱的市政设计院知识管理系统研究与实现[D].上海:上海交通大学软件学院,2019:1-65.
- [15] 牟天昊,李少远.流程工业控制系统的知识图谱构建[J].智能科学与技术学报,2022,4(1):129-141.
- [16] 刘岍,李杨,段宏,等.知识图谱构建技术综述[J].计算机研究与发展,2016,53(3).
- [17] 田玲,张谨川,张晋豪,等.知识图谱综述——表示、构建、推理与知识超图理论[J].计算机应用,2021,41(8):2161-2186.
- [18] 中华人民共和国水利部.取水许可管理办法[EB/OL].[2023-02-20].http://szy.mwr.gov.cn/zcfg/gz/201812/t20181231_1069501.html.

(下转第27页)

[7] 张阿哲, 李家欢, 朱子建. 数字孪生流域数据安全问题探究及对策[J]. 水利信息化, 2022 (6): 15-19.

[8] 钱峰, 周逸琛. 数字孪生流域共建共享相关政策解读[J]. 中国水利, 2022 (20): 14-17.

Pilot construction and practice of digital twin Tanjiang river basin

DONG Wanjun^{1,2}, WANG Liujie³, DONG Bo^{2,4}, CHEN Liming⁵, HU Tengfei⁵, NIU Ruiping^{1,2}

(1. *Nanjing Research Institute of Hydrology and Water Conservancy Automation, Ministry of Water Resources, Nanjing 210012, China*; 2. *Research Center on Hydrology and Water Resources Monitoring, Ministry of Water Resources, Nanjing 210012, China*; 3. *Pearl River Water Conservancy Comprehensive Technology and Network Information Center, Guangzhou 510611, China*; 4. *Jiangsu NIHWA Technology Co., Ltd., Nanjing 210012, China*; 5. *Nanjing Hydraulic Research Institute, Nanjing 210029, China*)

Abstract: The construction of digital twin Tanjiang River Basin is carried out to improve the flood control capacity. The digital twin Tanjiang River aims to meet the needs of flood control scheduling in the basin, with the goal of finely scheduling 22 large and medium-sized reservoirs and 7 cascade gates. Focusing on the core functions of “four pre”, data base plates, model platforms, and business applications are developed. For data base plates, a L1 level of basin and a L2 level of key areas and engineering projects are established. For model platforms, 6 major types of water conservancy professional models are developed. For business applications, the prediction and regulation system platform with “four pre” functions is constructed, in order to achieve efficient management of multi-source monitoring perception and data, to support centralized management, joint scheduling, and scientific decision-making of reservoirs, gates, and dams. It also can maximize the safety of flood (tide) in the basin, address the existing shortcomings of prediction and monitoring, improve the timeliness of engineering scheduling feedback, and provide a reference for standardized management of rivers and lakes. Thus, a basin joint scheduling platform is preliminarily formed. The project achievements were put into practice during the 2022 “Dragon Boat Water”, typhoon and rainstorms, which could help the joint regulation of reservoirs and sluices in Tanjiang River basin, and achieve the best benefits of flood control without inundation, water resources utilization.

Key words: digital twin; pilot construction; joint scheduling; basin flood control; Tanjiang River basin

(责任编辑: 陆 燕)

(上接第 13 页)

Application of knowledge graph technology in national water consumption platform

WANG Chenyu, LIU Qingtao, SHEN Hongxia

(*Information Center, Ministry of Water Resources, Beijing 100053, China*)

Abstract: The integration of national water consumption platform is an important way to realize “integrated management” of water consumption business. However, there are many water consumption systems to be integrated, and the information sources are dispersed. In addition, there are some practical problems in platform integration, such as multiple sources, inconsistent standards, unclear relationship among accounts, licenses and points, etc. Therefore, this paper attempts to apply knowledge graph technology to the platform integration process, building a set of national unified database table structure standard, forming a set of real and unique relationship among accounts, licenses and points, and building a set of intelligent and effective regulatory data products. The top-down knowledge graph construction method is adopted, in which an ontology in the pattern layer and its mutual relationship can be built according to the structured data of each system, and corresponding conceptual model and rule relationship can be formed. The entity and relationship from the data are extracted, data fusion and data layer construction are carried out, and knowledge is timely updated. Thus, the construction of the knowledge graph of the national water consumption platform is completed. Finally, based on the unified and standardized database table structure standard, statistical analysis of data information on a spatial-temporal scale is carried out according to the water consumption behavior supervision demands of water users, and six types of supervision products are constructed to form a “chart” of water consumption control.

Key words: the national water consumption platform; knowledge graph; system integration; top-down

(责任编辑: 来冰华)