

DOI:10.7522/j.issn.1000-0240.2021.0035

WU Adan, SHI Yanmei, GUO Jianwen, et al. Design and implementation of new digital Heihe River information system[J]. Journal of Glaciology and Geocryology, 2021, 43(5):1606-1616. [吴阿丹, 史艳梅, 郭建文, 等. 新版数字黑河信息系统的设计与实现[J]. 冰川冻土, 2021, 43(5):1606-1616.]

新版数字黑河信息系统的设计与实现

吴阿丹^{1,2}, 史艳梅³, 郭建文¹, 李建轩¹, 李恒¹

(1. 中国科学院西北生态环境资源研究院 黑河遥感试验研究站 甘肃省遥感重点实验室, 甘肃 兰州 730000;

2. 中国科学院大学, 北京 100049; 3. 中国人民解放军32016部队, 甘肃 兰州 730000)

摘要: 随着黑河流域时空数据的快速增长, 已有的黑河信息系统在共享效率、数据分析及平台架构等方面存在缺陷, 已无法满足现实的需要。本研究针对目前平台的缺陷重新设计新版数字黑河信息系统(NDHRIS)的整体框架, 从观测、数据及模型三个角度实现新版数字黑河信息系统的新功能。首先从基础设施、服务及应用三个层面介绍系统整体设计, 其次详细阐述系统实现的关键技术及功能改进, 最后, 讨论了该系统仍然存在的问题及未来发展方向。目前该系统已业务化运行, 数据量超过5 T, 共1 058条元数据, 向大约100个研究所和50个项目提供了大约8 TB数据集和5 000个数据服务, 有效提升了黑河流域科学数据的共享效率, 支撑全球寒区监测计划的实施。

关键词: 黑河流域; 观测实验; 信息系统; 数据共享; 数据可视化; 模型代码

中图分类号: P208; TP311 **文献标志码:** A **文章编号:** 1000-0240(2021)05-1606-11

0 引言

虽然国内外对数字流域的理解存在差异^[1-2], 但都一致认为数字流域是流域科学的重要研究内容之一。黑河流域是中国第二大内陆河流域, 很早就开展了数字流域研究, 并且提出了“数字黑河”的概念^[3]: 黑河数字由基础观测网络设施及信息系统组成, 前者由星-空-地观测系统组网构建, 后者是为黑河流域科学研究和流域集成管理而搭建的集数据、模型和观测系统于一体的信息化平台, 是黑河流域研究的数据中心, 又是它的虚拟实验室, 模拟和观测平台^[4]。

自1990年以来, 黑河流域共开展过五次大规模的综合观测试验:

(1)1990年6月—1992年10月, 在黑河流域中游开展了“黑河地区地气相互作用野外观测实验研究”(HEIFE)^[5], 开创了黑河流域综合观测研究的先河, 为之后的观测奠定了基础, 其共享数据集至今仍为研究者所广泛使用。

(2)2004年6—8月期间, 在黑河流域西部子水系中游金塔绿洲开展了为期两个多月的“绿洲系统能量与水分循环过程”观测试验, 之后的多个夏季中, 又在这一地区开展后续试验, 这些试验统称为“金塔试验”。

(3)2007—2009年, 中国科学院西部行动计划项目“黑河流域遥感—地面观测同步试验与综合模拟平台建设”和国家重点基础研究发展计划项目“陆表生态环境要素主被动遥感协同反演理论与方法”共同发起了“黑河综合遥感联合试验(以下简称WATER实验)^[6]”。初步实现了建立一个开放的试验平台和发展多尺度、多分辨率、高质量并最终完全共享的综合数据集的目标, 全部数据通过黑河综合遥感联合试验信息系统进行发布。

(4)2012—2015年, “黑河流域生态—水文过程综合遥感观测联合试验(以下简称HiWATER实验)^[7]”也在黑河流域进行, 其目的是目标是显著提升对流域生态和水文过程的观测能力, 建立国际领

收稿日期: 2020-03-19; 修订日期: 2021-01-03

基金项目: 国家重点研发计划项目“生态监测物联网关键技术研发”(2016YFC0500105); 国家自然科学基金重大研究计划项目“黑河流域水-生态-经济系统的集成模拟与预测”(91425303)资助

作者简介: 吴阿丹, 博士研究生, 主要从事信息服务、寒旱区数据科学研究. E-mail: wuadan@lzb.ac.cn

通信作者: 郭建文, 副研究员, 主要从事寒旱区遥感与地理信息系统应用研究. E-mail: guojw@lzb.ac.cn

先的流域观测系统,提高遥感在流域生态-水文集成研究和水资源管理中的应用能力。

(5)2016—2020年,生态系统关键参量监测设备研制与生态物联网示范项目^[8]也在黑河流域部署生态监测物联网开展监测示范,构建了天地一体化的群落-生态系统-景观多尺度生态物联网监测系统,对推进黑河流域生态文明建设具有重要支撑作用。

针对这些不同时段、不同规模的黑河流域观测实验,中科院西北生态环境资源院(原中科院寒区旱区环境与工程研究所,简称西北院)经过20年的时间建立了一系列不同功能的数据服务系统,初步实现了实验介绍、数据共享、模型展示等信息服务,如旧版数字黑河^[9]、西部数据中心^[10]、黑河流域生态水文过程综合遥感观测联合实验、黑河流域遥感-地面观测同步实验与综合模拟平台、中国陆面数据同化网、黑河流域模型集成、黑河数据可视化在线平台^[11]、中科院黑河遥感实验研究站等信息系统。这八个系统都涉及到了黑河流域相关的数据和模型,从不同角度提升了黑河实验的数据共享效率,实现了数字黑河信息系统向数据中心的过渡,但這些系统尚不能满足“数字黑河”的现实要求,存在一些急需问题:①信息平台门户较多,不利于平台的维护和数据的高效共享;②一些平台样式陈旧,在不同的终端中会有样式变形问题,用户体验较差;③功能不够全面,缺少数据分析及数据实时监测等功能;④这些系统由不同的团队开发,使用不同的技术框架,很难进行系统间的互操作,随着数据量的不断增多,系统性能也受到挑战。

为了进一步提高数字黑河信息服务,本文在黑河流域空间观测和地面测量等数据基础上,立足中

科院西北院信息基础设施,借鉴已有系统资源,重构新版数字黑河信息系统(New Digital Heihe River Information System,以下简称NDHRIS),期望进一步提升黑河流域科学数据的共享效率。本文主要介绍NDHRIS的设计和实现,重点在新功能及功能优化的考虑和实现上,最后讨论了系统仍存在的问题和未来需要改进的三个方向。

NDHRIS在信息服务方面的主要贡献如下:

(1)将多语言编写的系统统一集成为基于Python的Web系统,一方面能够简化系统的维护,方便用户一站式浏览。另一方面,支持与大数据平台的对接,为后期方法库、模型库的在线运算提供支持。

(2)替换开源元数据编辑器Geonetwork,开发了一套自有知识产权的元数据管理工具,具有较高的可扩展性和稳定性;

(3)基于ISO19115元数据标准,统一数据共享和交换的规则,实现多平台数据的统一管理与共享。

(4)进一步优化数据汇交流程、数据申请流程。在保证数据质量的前提下,进一步提高了数据管理和使用的效率。

1 整体设计

1.1 系统框架

NDHRIS以满足数据“共享为本,可视化为根”的原则进行建设。针对已有的平台存在的缺陷,从观测、数据、模型三个角度出发,构建一个风格现代化、内容多样化,标准统一化,流程简单化的信息服务平台。系统整体框架包括基础设施层、服务层和应用层三个层面,如图1所示。

基础设施层:提供整个系统运行的硬件环境及

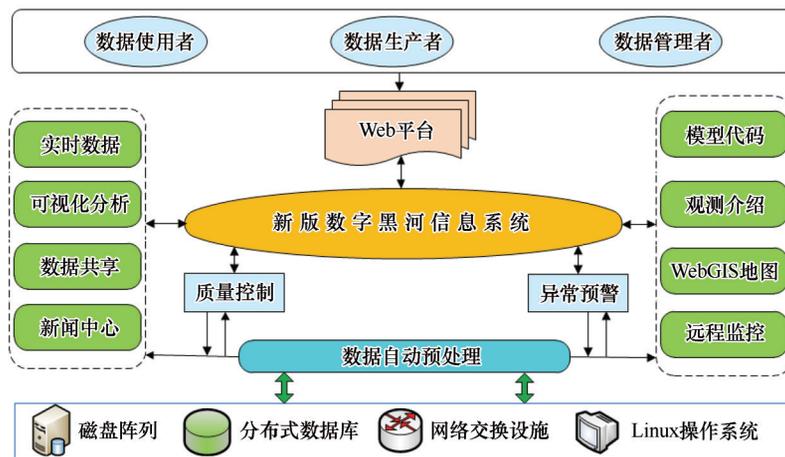


图1 系统整体框架

Fig. 1 Overall framework of the system

软件环境。本系统基础设施使用中科院青藏高原研究所提供的云环境：一个 Web 服务器，一个 500 T 的磁盘阵列，四个处理服务结点，这四个节点后期可用于基于 Spark 的分布式计算。

服务层：由一系列 Web 服务组成，包括数据共享服务、可视化服务、地图服务、数据实时监测服务等。

应用层：即 NDHRIS 网站界面，通过调用服务

层 Web 服务提供的各种方法，实现观测实验介绍、数据汇交、数据浏览及下载、数据可视化及模型代码共享等功能。

1.2 功能设计

NDHRIS 是在已有八个旧系统基础上进行功能的升级，重点从观测、数据、模型三个方面进行功能改进(图2)。

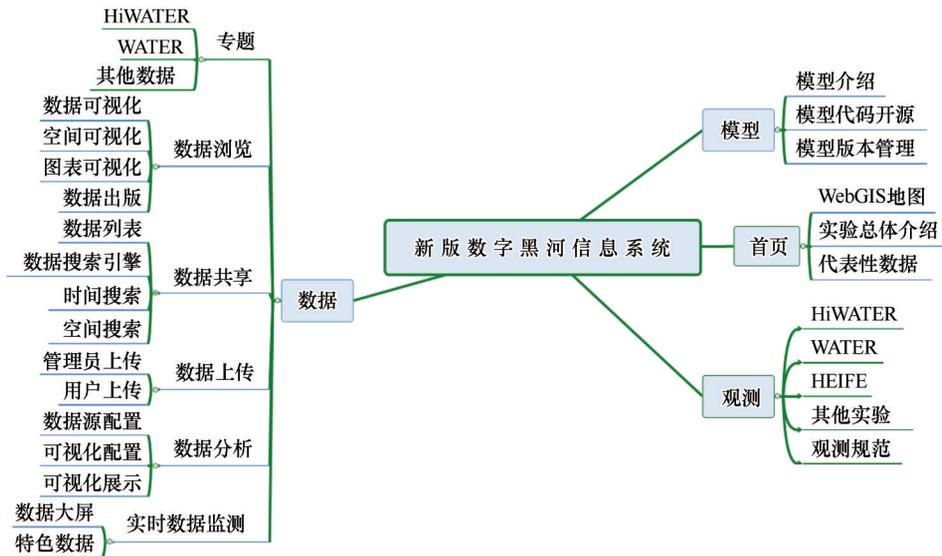


图2 系统主要功能

Fig. 2 Main functions of the system

2 系统实现

2.1 开发技术和运行环境

NDHRIS 采用浏览器/服务器模式(B/S)架构。浏览器端采用Nodejs作为开发语言，服务器端采用Python技术，利用瓦片服务、Echart插件实现不同类型、不同体量数据的可视化。为了提高用户体验同时用到了Jquery、Ajax等技术。同时采用分布式数据库GreenPlum^[12-13]和分布式文件系统Hadoop^[14]对不同结构的数据实现差异化的管理策略。Web系统和数据库均运行在Ubuntu16操作系统环境中。

2.2 功能实现

2.2.1 观测数据输入

随着信息技术的发展，中科院西北院从2012年至2015年对黑河流域观测系统网络进行了数据自动采集、远程传输^[15]、数据质量控制^[16]等方面的信息化改造，部署了多尺度嵌套无线传感器网络^[17]，形成“一条样带+多个核心观测站+遥感综合观测试验+各研究项目自主观测”组成的综合、立体、多尺度的流域观测系统。从2012年至今，在黑河上

游、中游及下游建立9个观测区域，362个观测点，包括数采仪385个，2145个传感器。目前仍在运行的观测点有11个(图3)，详细信息如表1所示。

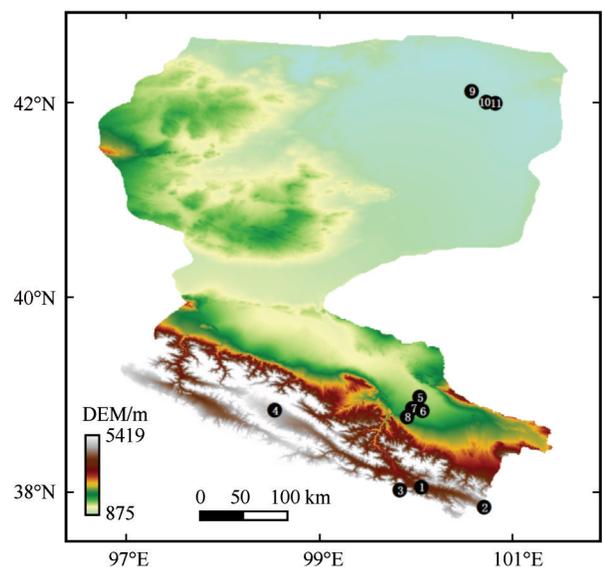


图3 黑河流域11个运行站点分布图

Fig. 3 Distribution map of 11 operating stations in heihe river basin

表 1 观测站点详细信息(阿柔、大满和四道桥是超级站,其他是普通气象观测站)

Table 1 Details of observatory stations (Arou, Daman, and Sidaoqiao are superstations; the other stations are ordinary stations)

编号	站点名称	下垫面类型	位置	经度/E	纬度/N	高程/m	观测时期
1	阿柔站	亚高山草甸	上游	100.46°	38.05°	3 033	2012年6月至今
2	景阳岭站	高寒草甸	上游	101.12°	37.84°	3 750	2013年8月至今
3	埡口站	高寒草甸	上游	100.46°	38.05°	3 033	2014年1月至今
4	大沙龙站	高寒草甸	上游	98.94°	38.84°	3 739	2013年8月至今
5	张掖湿地	芦苇	中游	100.45°	38.98°	1 460	2012年6月至今
6	黑河遥感站	沼泽草甸	中游	100.48°	38.83°	1 560	2014年8月至今
7	大满站	玉米	中游	100.37°	38.86°	1 556	2012年5月至今
8	花寨子荒漠站	盐爪爪荒漠	中游	100.32°	38.77°	1 731	2012年6月至今
9	沙漠站	红砂荒漠	下游	100.99°	42.11°	1 054	2015年4月至今
10	四道桥站	怪柳	下游	101.14°	42.00°	873	2013年7月至今
11	混合林站	胡杨、怪柳	下游	101.13°	41.99°	874	2013年7月至今

NDHRIS 重点对 WATER、HiWATER 等几个观测实验网站进行归纳(图4),在集成原有网站内容的基础上,重新设计“观测”栏目,从实验计划、实验组织、实验参与者、合作交流、实验进展、出版物、支持单位等角度详细介绍黑河观测实验,并将观测实

验与其观测数据关联,即用户在浏览实验内容的同时可快速获取相关数据(图5)。

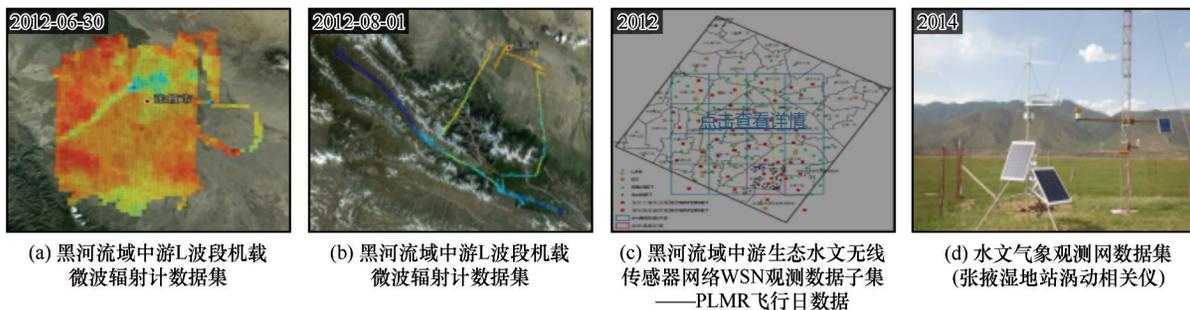
2.2.2 数据管理、分析和共享

“数字黑河”已阶段性地完成了数据集成,主要数据来源为黑河计划有关项目的数据成果、黑河生



图 4 集成已有网站

Fig. 4 Integration of existing web sites



(a) 黑河流域中游L波段机载微波辐射计数据集
 (b) 黑河流域中游L波段机载微波辐射计数据集
 (c) 黑河流域中游生态水文无线传感器网络WSN观测数据子集——PLMR飞行日数据
 (d) 水文气象观测网数据集(张掖湿地站涡动相关仪)

图 5 观测实验与对应数据相互关联

Fig. 5 Correlation between observation experiment and corresponding data

态水文遥感试验、以及黑河流域的各种基础地理数据。数据量超过5 T,共1 058条元数据。根据这些流域数据特点,我们建立了具有高稳定性、高扩展性的分布式数据库管理系统,在此基础上是实现了数据共享功能:用户可以按照专题浏览、所有数据、空间检索、关键词浏览、缩略图浏览、文献浏览及作者浏览等方式查询及下载数据,并提供了数据可视化、数据实时监测等功能。本节重点阐述数据共享功能中的几个新的进展:

(1) 细化数据共享流程中用户角色

NDHRIS的数据共享流程如图6所示。各类数

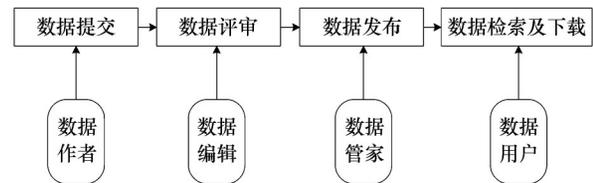


图6 数据共享流程

Fig. 6 Data sharing process

据贡献者将数据提交到数据平台,数据经过专业评审之后正式发布,并注册DOI。各类数据用户可通过数据共享门户以多种方式检索、查看和按需下载数据。在该业务流程中,主要包括四种角色,具体如下如表1所示。

表1 四种用户角色

Table 1 Four user roles

用户角色	描述
三类数据贡献者	数据期刊、数据作者自主发布数据、大数据平台分析产生的数据
两类数据用户	注册用户和非注册用户具有不同的浏览权限
两类数据评审专家	数据编辑和数据评审人
数据管家	系统管理员,具有数据发布和管理权限

针对数据共享流程中涉及的四类用户角色,NDHRIS提供了四个对应的功能:

1)数据提交功能:为三类数据贡献者提交数据提供支持。

2)数据检索及下载功能:为两类数据用户检索、查看(可视化)和获取数据提供支持。

3)数据评审功能:为数据编辑和数据评审人组织评审或评审数据提供支持。

4)数据发布与管理功能:为数据管家发布经过评审的科学数据集和注册DOI提供支持。

(2) 数据发布流程的优化

高质量数据是数据共享的基础,因此在汇交数据时必须进行数据的质量检查。旧的平台已经具备数据汇交的功能^[18]:数据提供者需要首先登陆数据中心平台,然后撰写元数据并通过FTP将数据实体进行上传,经过数据中心检查后发布数据。然而该汇交流程还存在一些不足:汇交数据编辑功能不够灵活、数据提供者必须安装FTP客户端工具、数据评审无法按照不同层级进行评审(如项目、课题、子课题等)、汇交环节出现错误后,数据中心管理人员只能通过邮件或者电话沟通解决。

新版数字黑河信息系统,在总结了大量数据作者及用户的意见后,重新凝练了数据评审流程,从以下三个方面进行了功能改进(图7):

1)从数据提交者的角度,整个流程都是向导性操作,显著提高数据汇交效率,体现在两个方面:开发了基于ISO19115标准的元数据编辑工具,替换开源元数据编辑器Geonetwork,使得元数据编辑更加能活,并且支持中文和英文两种语言的元数据编写;增加了网盘方式的数据上传,提高用户体验;

2)从评审专家角度,细化了角色,可以从项目、课题、子课题等不同角色评审数据,实现逐级审核的过程,保证了提交数据的质量;

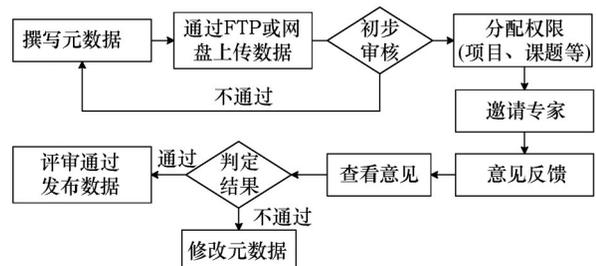


图7 NDHRIS数据发布流程

Fig. 7 New data publishing process of NDHRIS

3)从数据中心管理员的角度,减少了其操作步骤,只需根据系统提示邀请专家及根据意见判断是否发布即可。

(3) 基于元数据的数据检索

元数据互操作技术的发展,以及元数据互操作

实践成果提供的支撑使得元数据互操作具备可行性。因此,新版数字黑河系统基于 ISO19115 标准^[10],对所有元数据进行重构,实现元数据之间的互操作:当用户在检索资源时,系统可以在已经建立关系的元数据元素字段中搜索(图 8),可以将不同的元数据标准描述的相关信息资源检索出来,实现黑河数据的一站式获取(图 9、图 10)。目前系统包含的数据集主要是基于中国寒区和干旱区关键变量的数据,如冰川、雪、冻土、沙漠、地表大气强迫、土壤质地、植物功能类型图、生态水文遥感产品等模型数据集。

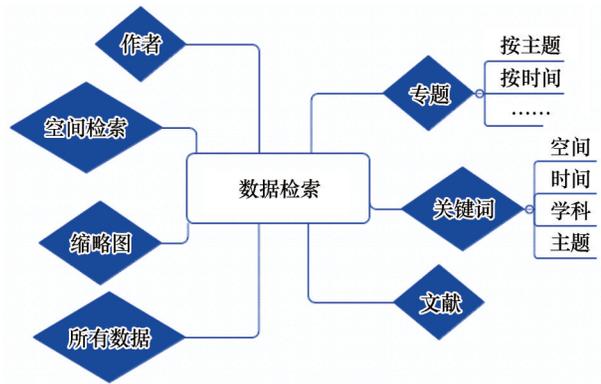


图 8 基于元数据的数据检索

Fig. 8 Data retrieval based on metadata

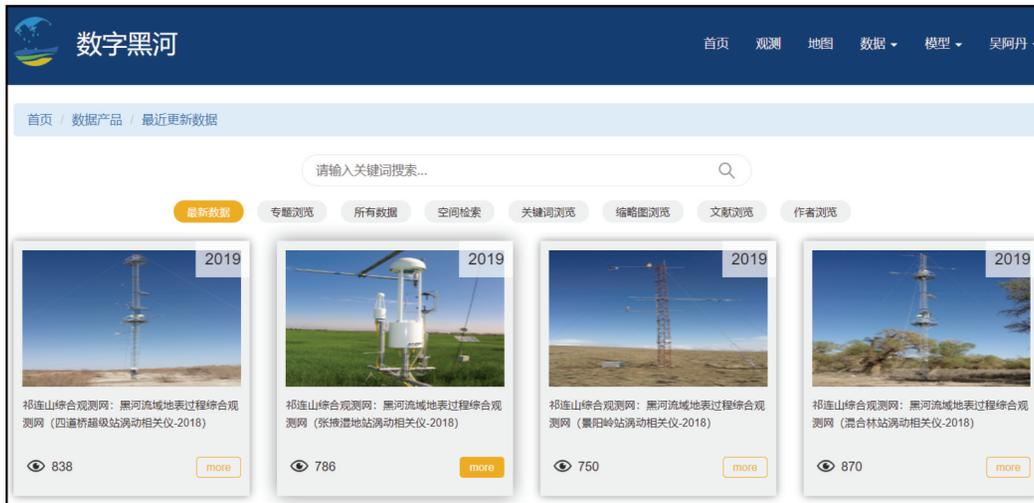


图 9 数据检索界面

Fig. 9 Data retrieval interface



图 10 元数据浏览界面

Fig. 10 Browsing interface of metadata

此外,NDHRIS 也实现了多平台数据的统一管理与共享:目前新系统与时空三极大数据平台、国家青藏高原数据中心、三江源国家公园数据监测平台以不同域名对外发布,但这些平台均采用同一套数据库管理系统,这样的设计策列有利于平台数据之间的互操作,如,对于元数据的录入及编辑,这些系统均采用统一的元数据管理系统,管理员可根据数据的特点选择不同的数据发布平台。对于用户来讲,可在国家青藏高原数据中心搜索到NDHRIS 中的数据。

(4) 针对不同类型数据的可视化分析策略

数据可视化使抽象的数据转化为易于辨识的图像,为用户提供更加直观的信息和知识,是应对数据共享平台“数据丰富,信息贫乏”的有效手段^[11]。NDHRIS 针对不同体量、不同类型的数据使用不同的可视化策略,本节重点介绍地图和结构化数据的可视化方法(图 11)。

地图可视化包括基础图层、专题图层的可视化。

① 基础图层可视化

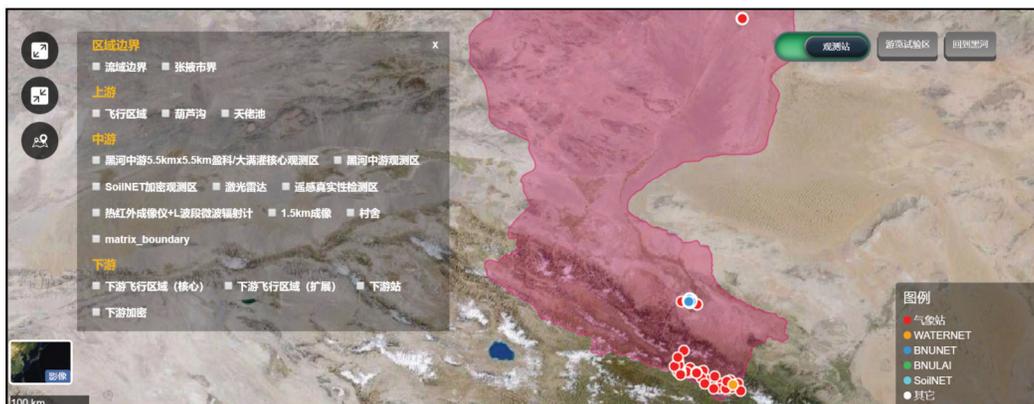


图 12 以瓦片服务发布的研究区域

Fig. 12 Research areas published as tile services

② 专题图层可视化

根据专题图数据量的大小,采取不同的渲染策略。如果是较小的矢量文件,服务器直接将该数据转换为 Geojson 格式的数据,并且通过 Gzip 压缩后再发送到客户端,客户端浏览器中可以自动解压并利用 Openlayer 直接显示该 Geojson 数据(图 13)。

如果是较大的矢量数据或者是栅格数据,需要调用瓦片服务(WMTS),该过程和基础图层的可视化过程一致。

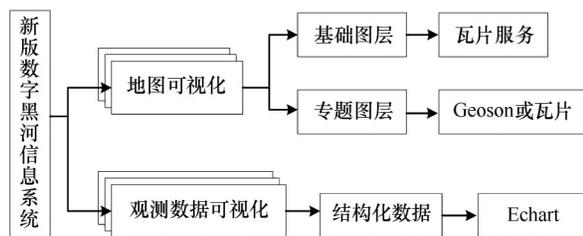


图 11 系统主要数据的可视化方案

Fig. 11 Visualization scheme of main data in the system

使用地形、研究区域、河流、湖泊等基础地理信息数据绘制的地图被作为基础底图使用,由于这些数据普遍数据量大,并且具有类型多、属性多的特点,直接浏览如此大量的数据往往效率较低,所以需要引入瓦片地图服务:预先将这些基础数据在服务器端生成切片并存储,利用 Geoserver 将切片数据发布为 WMTS 服务(Web Map Tile Service),最后在 Web 前端使用 Openlayer 加载该服务即可高效显示基础图层。如图 12,该地图中使用了 Mapbox 发布的瓦片地图以及经过瓦片处理后的黑河流域(粉色区域),这种方式可以极大的提高了 Web 访问速度和用户体验。

对于结构化的观测数据,其格式一般为 XLS 或 CSV,如常规气象水文观测数据(降雨量、水位、流量等),该类数据我们利用 Python 程序对其进行自动解析并保存到分布式数据库中,所有的观测数据都保存在 Tb_value 表中,但我们对表进行了分表及索引的优化,在这些工作的基础上利用前后端交互的方式进行可视化绘图;客户端将目前浏览的数据时间范围和空间范围发送到服务器,服务器端则根据对应的数据范围在存储数据库中查询,然后将查询结果转换为 JSON 格式^[19],并返回给客户端,客户

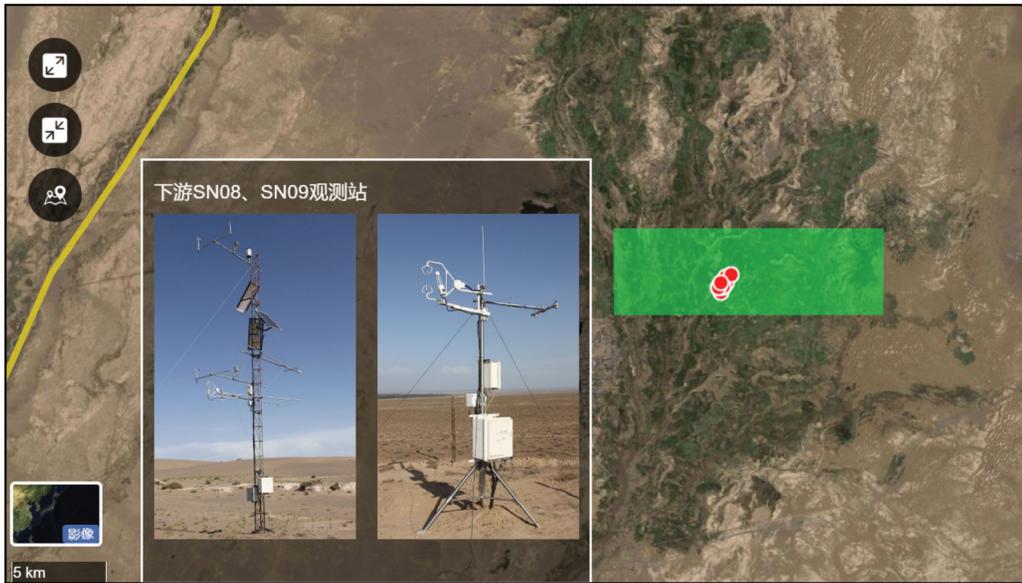


图 13 黑河下游飞行区域在基础图层上叠加(绿色部分)

Fig. 13 Heihe downstream flight area is overlaid on the base layer (green)

端接收 JSON 数据并利用开源脚本 Ecahrt 控件绘制 图像(图 14)。

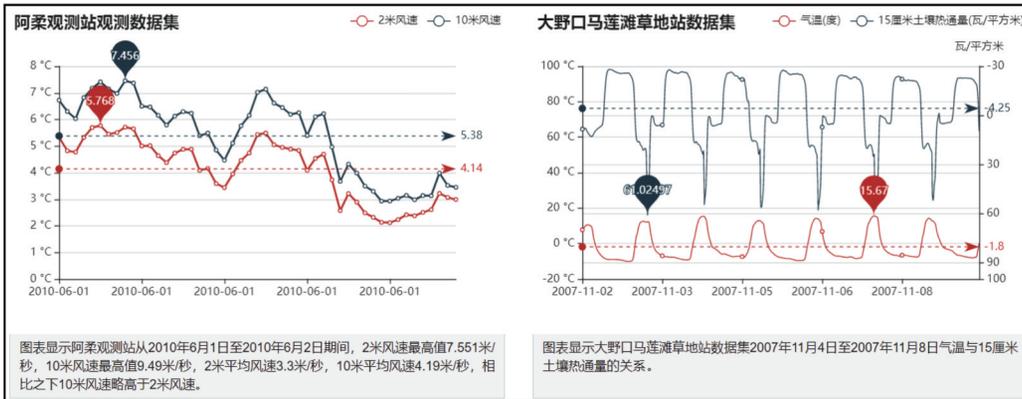


图 14 气象数据可视化效果

Fig. 14 Visualization of meteorological data

(5) 面向生态监测物联网流数据的实时监测

物联网^[20-22]的兴起为黑河流域生态观测监测数据的实时获取提供了重要保障^[23],这些数据具有高速、无限、不确定性等特征,我们称之为流数据。面对这些流数据,传统数据可视化方法存在挑战:流数据在短时间内持续产生,如果将这些流数据全部加载到内存中进行显示对于普通服务器无法满足这样得性能要求。如果采用客户端频繁请求服务器(如 Ajax)的方式更新用户界面上的可视化图表,当用户量较多时服务器负载会急剧增加。

NDHRIS 基于 Websocket 实现了客户端浏览器与服务器持续的长连接,降低了服务器的负载,并且可以持续为客户端提供最新数据及其可视化曲

线,即只要服务器端接收到野外新数据,客户端会立刻更新数据,如图 15,因黑河流域加密观测期间数据采集周期为 1 分钟,所以整个数据监测曲线每分钟自动更新一次。

2. 2. 3 模型共享

对黑河流域以及周边地区建立科学模型是开展流域研究必不可少的一种手段^[23]。数字黑河集成了流域生态-水文-经济模型,具有对冰冻圈水文过程、地表水与地下水交互、水资源利用、荒漠和绿洲的生态水文过程、人口经济社会发展等的综合模拟能力。

NDHRIS 基于开源项目 GitLab 建立模型代码共享,提供了模型代码浏览、共享及版本控制等功能,便于黑河模型集成研究的开源及共享(图 16)。

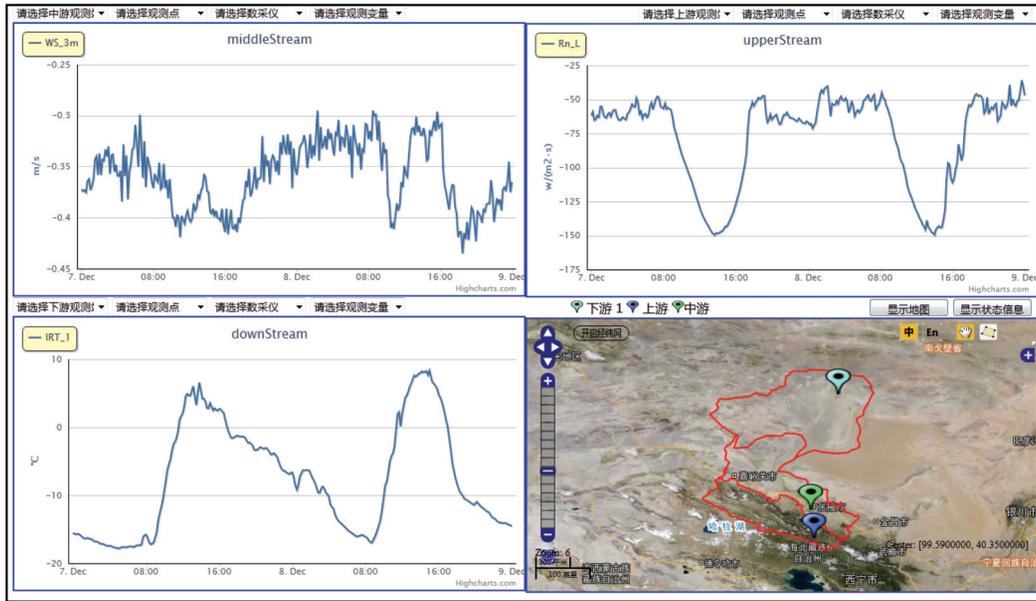


图 15 黑河上游、中游及下游数据实时更新(数据每分钟更新一次)

Fig. 15 Real-time data in upper, middle and down stream of heihe river is updated every minute



图 16 模型代码共享界面

Fig. 16 Sharing interface of model codes

3 结论及展望

本文在集成已有数字黑河相关系统的基础上,采用最新的技术和设计理念构建了一个具有良好稳定性和支持大数据分析接口的数字化信息平台—新版数字黑河信息系统(<http://heihe.westgis.ac.cn/>)。该系统将观测、数据、模型集成为一个整体,实现黑河流域科学数据信息服务,可以为黑河流域

研究提供有力的支持。目前新系统已经业务化运行,有效弥补了八个旧版黑河数据服务系统的不足,将进一步提高黑河流域相关科研数据的共享效率和服务能力。

目前 NDHRIS 的注册用户总数约为 4 314 人,向大约 100 个研究所和 50 个项目提供了大约 8 TB 数据集和 5 000 个数据服务。虽然已阶段性地完成

了数据集成和共享,但还存在一些问题。下一步工作将重点解决以下三个问题:

(1)利用 Python 对 NDHRIS 框架进行了重新改造,具有了与大数据平台(如 Spark, Hadoop)对接的能力,但目前还未实现大数据分析功能。今后须完善 Mapreduce、Spark 等大数据计算接口,并支持关键遥感数据典型反演算法接口的 Web 调用(如植被覆盖度、植被生产力、积雪覆盖比例、雪深、冻融状态、碳吸收能力等),为实现黑河流域的决策支持提供分析平台。

(2)本系统只能与元数据结构一致的平台进行数据互操作,然而目前已有的国内外数据共享平台没有采用统一的元数据标准,开发语言各不相同,存储策略各不相同,这为不同平台之间的数据集成共享带来挑战。为了能够进一步增强 NDHRIS 系统的使用范围及影响力,下一步应该考虑实现异构平台之间的互操作:对各种元数据标准的元数据元素进行分类,建立元数据元素之间多样的关系,实现层级关系丰富的科学数据元数据之间的互操作,为异构互操作的数据共享网络平台提供支撑。

(3)基于 ISO19115 标准,我们对 NDHRIS 中所有元数据进行了重构,实现基于多要素的数据查询,提高了数据查询的效率和准确度,但是,面对时空数据的快速增长,NDHRIS 仍需面对新的挑战:由于应用任务和时空数据之间缺乏深度语义关联导致信息服务从海量数据中为特定任务提取精准信息的能力受到限制,因此笔者期望今后利用本体理论建立任务、时空数据及 Web 服务的统一描述模型,并发展多级语义约束的“任务-数据”定量关联方法,实现渐进求精的数据智能聚合,提供更加智能的数据共享服务。

参考文献(References):

- [1] Liu Jiahong, Wang Guangqian, Wang Kai. Review on advancement of study on digital riverbasin in China[J]. Journal of Hydraulic Engineering, 2006, 37(2): 240-246. [刘家宏,王光谦,王开. 数字流域研究综述[J]. 水利学报, 2006, 37(2): 240-246.]
- [2] Wang Guangqian, Lie Jiahong. The digital watershed model for the yellow river basin[J]. Water Resources and Hydropower Engineering, 2006, 37(2): 492-499. [王光谦, 刘家宏. 黄河数字流域模型[J]. 水利水电技术, 2006, 37(2): 492-499.]
- [3] Li Xin, Chen Guodong, Ding Yongjian, et al. Design of water resource information system of the Heihe River basin, Northwest China[J]. Journal of Desert Research, 2000, 20(4): 378-382. [李新, 程国栋, 丁永建, 等. 黑河流域水资源信息系统设计[J]. 中国沙漠, 2000, 20(4): 378-382.]
- [4] Li Xin, Chen Guodong, Wu Lizong. Digital Heihe River basin 1: an information infrastructure for the watershed science[J]. Advance in Earth Sciences, 2010, 25(3): 297-305. [李新, 程国栋, 吴立宗. 数字黑河的思考与实践 1: 为流域科学服务的数字流域[J]. 地球科学进展, 2010, 25(3): 297-305.]
- [5] Hu Yinqiao, Gao Youxi. Some new understanding of processes at the land surface in arid area from the HEIFE[J]. Acta Meteorologica Sinica, 1994, 52(3): 285-296. [胡隐樵, 高由禧. 黑河实验(HEIFE)——对于干旱地区陆面过程的一些新认识[J]. 气象学报, 1994, 52(3): 285-296.]
- [6] Li Xin, Li Xiaowen, Li Zengyuan, et al. Progresses on the Watershed Allied Telemetry Experimental Research (WATER)[J]. Remote Sensing Technology and Application, 2012, 27(5): 637-649. [李新, 李小文, 李增元, 等. 黑河综合遥感联合试验研究进展: 概述[J]. 遥感技术与应用, 2012, 27(5): 637-649.]
- [7] Li Xin, Liu Shaoming, Ma Mingguo, et al. HiWATER: an integrated remote sensing experiment on hydrological and ecological processes in the Heihe River basin[J]. Advance in Earth Sciences, 2012, 27(5): 481-498. [李新, 刘绍民, 马明国, 等. 黑河流域生态-水文过程综合遥观测联合试验总体设计[J]. 地球科学进展, 2012, 27(5): 481-498.]
- [8] Li Xin, Liu Shaoming, Sun Xiaoming, et al. Innovative development of equipments and internet-of-things techniques for ecosystem monitoring and its demonstration[J]. Acta Ecologica Sinica, 2016, 36(22): 7023-7027. [李新, 刘绍民, 孙晓敏, 等. 生态系统关键参量监测设备研制与生态物联网示范[J]. 生态学报, 2016, 36(22): 7023-7027.]
- [9] Li Xin, Wu Lizong, Ma Mingguo, et al. Digital Heihe River basin 2: data integration[J]. Advance in Earth Sciences, 2010, 25(3): 306-316. [李新, 吴立宗, 马明国, 等. 数字黑河的思考与实践 2: 数据集成[J]. 地球科学进展, 2010, 25(3): 306-316.]
- [10] Nan Zhuotong, Li Xin, Wang Liangxu, et al. Design and implementation of the online data sharing portal of Environmental and Ecological Science Data Center for West China[J]. Journal of Glaciology and Geocryology, 2010, 32(5): 970-975. [南卓铜, 李新, 王亮绪, 等. 中国西部环境与生态科学数据中心在线共享平台的设计与实现[J]. 冰川冻土, 2010, 32(5): 970-975.]
- [11] Wu Adan, Guo Jianwen, Li Jianxun, et al. The application of web-based visualization system for a wireless sensor network in Heihe Watershed Allied Telemetry Experimental Research[J]. Remote Sensing Technology and Application, 2013, 28(3): 416-422. [吴阿丹, 郭建文, 李建轩, 等. 基于 Web 的黑河流域生态水文 WSN 自动观测数据可视化系统应用研究[J]. 遥感技术与应用, 2013, 28(3): 416-422.]
- [12] Waas F M. Beyond conventional data warehousing-massively parallel data processing with Greenplum Database[C]//Informal Second International Workshop on Business Intelligence for the Real-time Enterprise, 2009(27): 89-96.
- [13] Antova L, El-Helw A, Soliman M A, et al. Optimizing queries over partitioned tables in MPP systems[C]//Proceedings of the 2014 ACM SIGMOD International Conference on Management of Data. 2014: 373-384.
- [14] Shvachko K, Kuang H, Radia S, et al. The hadoop distributed file system[C]//IEEE Symposium on Mass Storage Systems & Technologies. 2010: 121-127.
- [15] Guo Jianwen, Shang Qingsheng, Chang Hailong, et al. Design of field observation data automatic assembling system[J]. Remote Sensing Technology and Application, 2013, 28(3): 399-404. [郭建文, 尚庆生, 常海龙, 等. 野外观测数据自动综汇系统方案设计[J]. 遥感技术与应用, 2013, 28(3): 399-404.]
- [16] Liu Feng, Guo Jianwen. Study on quality control approach for

- Heihe wireless sensor network observation data [J]. *Remote Sensing Technology and Application*, 2013, 28(2): 252-257. [刘丰, 郭建文. 面向黑河无线传感网络观测数据的质量控制方法研究[J]. *遥感技术与应用*, 2013, 28(2): 252-257.]
- [17] Jin Rui, Li Xin, Yan Baoping, et al. Introduction of eco-hydrological wireless sensor network in the Heihe River basin[J]. *Advance in Earth Sciences*, 2012, 27(9): 993-1005. [晋锐, 李新, 阎保平, 等. 黑河流域生态水文传感器网络设计[J]. *地球科学进展*, 2012, 27(9): 993-1005.]
- [18] Wang Liangxu, Wu Lizong, Li Hongxing, et al. Design and implementation of data collection system for integrated research on the eco-hydrological process of the Heihe River basin [J]. *Remote Sensing Technology and Application*, 2013, 28(3): 362-369. [王亮绪, 吴立宗, 李红星, 等. 面向黑河流域生态水文过程集成研究的科学数据汇交与管理[J]. *遥感技术与应用*, 2013, 28(3): 362-369.]
- [19] Han Min, Feng Hao. Research of GIS data exchange method based on JSON[J]. *Science of Surveying and Mapping*, 2010, 35(1): 159-161. [韩敏, 冯浩. 基于JSON的地理信息数据交换方法研究[J]. *测绘科学*, 2010, 35(1): 159-161.]
- [20] Xia F, Yang L T, Wang L, et al. Internet of things[J]. *International Journal of Communication Systems*, 2012, 25(9): 1101-1102.
- [21] Gubbi J, Buyya R, Marusic S, et al. Internet of Things (IoT): a vision, architectural elements, and future directions [J]. *Future Generation Computer Systems*, 2013, 29(7): 1645-1660.
- [22] Wollschlaeger M, Sauter T, Jasperneite J. The future of industrial communication: Automation networks in the era of the internet of things and industry 4.0 [J]. *IEEE Industrial Electronics Magazine*, 2017, 11(1): 17-27.
- [23] Li Xin, Xie Zhenghui. Introduction to the Water-Soil-Air-Plant-Human Nexus: modeling and observing complex land surface systems at river basin scale [J]. *Journal of Geophysical Research: Atmospheres*, 2019, 124(23): 12375-12379.

Design and implementation of new digital Heihe River information system

WU Adan^{1,2}, SHI Yanmei³, GUO Jianwen¹, LI Jianxuan¹, LI Heng¹

(1. Key Laboratory of Remote Sensing of Gansu Province, Heihe Remote Sensing Experimental Research Station, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou 730000, China; 2. University of Chinese Academy of Sciences, Beijing 100049, China; 3. Chinese People's Liberation Army Unit 32016, Lanzhou 730000, China)

Abstract: The “Digital Heihe River Basin” can be considered as an implementation of the Digital Earth concept in the Heihe River basin, the second largest inland river basin of China. The Digital Heihe River Basin is composed of an information infrastructure and its applications. The former is further built up by a data integration platform, a modeling system and an automatic observing system, and the latter are various applications using integrated models and decision support systems. With the rapid growth of the spatio-temporal data of Heihe River basin, some existing Heihe information service systems have defects in the aspects of sharing efficiency, data analysis and platform architecture, etc., which can no longer meet the practical needs. In this study, the overall framework of the new digital Heihe information system (NDHRIS) is redesigned to address the defects of the current platform, and the new functions of the new digital Heihe information system are realized from three aspects: observation, data and model. In this paper, the overall design of the system is introduced from three levels: infrastructure, service and application. Secondly, the key technologies and functional improvements of the system are described in detail. Finally, the existing problems and future development directions of the system are discussed. At present, the system has been operational and includes a lot of data with key variables of cold and arid regions in China, such as glacier, snow, frozen soil, desert, model datasets such as surface atmospheric forcing, soil texture, plant functional type map, eco-hydrology remote sensed products, and comprehensive dataset in Heihe River basin. It has a data volume of more than 5 T and a total of 1 058 metadata, providing about 8 TB data sets and 5 000 data services to about 100 research institutes and 50 projects, effectively improving the sharing efficiency of scientific data in Heihe River basin and supporting the implementation of the global cold monitoring plan.

Key words: Heihe River basin; observational experiments; information system; data sharing; data visualization; model code

(责任编辑: 车涛; 编辑: 周成林)