

# 地球大数据白皮书 (2023 年)



CCSA TC601 大数据技术标准推进委员会 地球大数据工作组

北京英视睿达科技股份有限公司

2023年9月

## 前 言

2022年12月，中共中央、国务院印发了《关于构建数据基础制度更好发挥数据要素作用的意见》，其中明确了数据作为新时代生产要素的重要价值。地球大数据包括来自地球大气圈、水圈、土壤圈、生物圈以及人类活动的多源时空数据，是数据要素产业的重要组成部分。对地球大数据的开发利用将为人类同自然环境的交互过程提供决策支持，促进可持续发展，最终实现人与自然和谐共生的现代化。

地球大数据的应用领域范围广阔，包括生态保护、自然资源管理、气象服务、城市规划、应急容灾等方面，但目前仍处于发展初期，面临着数据来源、结构多样，数据管理门槛高，数据应用场景复杂、落地方法论较少等痛点。如何充分发挥地球大数据的优势和特点，促进其社会价值、经济价值的高效释放成为产业界亟需解决的问题。

本白皮书将梳理地球大数据的内容范畴、数据资源、关键技术、行业生态、应用场景以及实践案例等，并初步提出地球大数据实践方法，旨在厘清地球大数据的行业发展现状，并为产业各方开展地球大数据开发利用的相关实践提供参考。

## 编制说明

### 指导单位：

生态环境部卫星环境应用中心、生态环境部环境规划院、中国信息通信研究院

### 编委会名单：

水利部水土保持监测中心：赵永军、罗志东、张红丽、时宇

大数据技术标准推进委员会：王超伦、马健瑞、韩晓璐、魏凯、姜春宇、马鹏玮

北京英视睿达科技股份有限公司：田启明、沈诗润、廖炳瑜、王伟、邹克旭、徐炜达、王鼎元、徐彬仁、李怀瑞、杨昱锟、常鹏慧、仝俊涛、朱珊娴、王雪瑞、郭东宸

易智瑞信息技术有限公司：刘洋、他浩飞、陈斌、马静丽、康来成、高淑鸿、曾智

阿里云计算有限公司：张辉、曾志明、陈呈举、陈辉、胡伯涛、贾雨宾、李坤、佟天泽、郭伟、郭伟

杭州数梦工场科技有限公司：念灿华、高智世、苑涵征、孔令君、魏森

中数（深圳）时代科技有限公司：吴英东、王海涛、马超、沈思泽、韩光

广东元能星泰孪生科技创新有限公司：凌晖、伍亮、岑权军、王晓枫、林子扬

特力惠信息科技股份有限公司：周红霞

**北京捷泰天域信息技术有限公司：汪维莉**

**支持单位：**

中关村睿宸卫星创新应用研究院、北京航空航天大学、北京邮电大学、北京交通大学、北京工商大学

# 目 录

一、 地球大数据总述 .....	1
(一) 地球大数据的定义 .....	1
(二) 地球大数据的价值 .....	2
(三) 地球大数据价值释放的要点 .....	4
二、 地球大数据的数据资源 .....	6
(一) 地球大数据的数据来源 .....	7
(二) 地球大数据的几何与时空属性 .....	11
三、 地球大数据的关键技术 .....	13
(一) 数据采集 .....	13
(二) 数据管理 .....	16
(三) 数据分析 .....	20
(四) 人工智能 .....	21
(五) 数据安全 .....	23
四、 地球大数据的行业生态与行业应用 .....	26
(一) 地球大数据产业图谱 .....	26
(二) 地球大数据行业应用 .....	27
五、 总结与未来展望 .....	38
(一) 总结 .....	38
(二) 地球大数据的发展趋势展望 .....	38

## 图 目 录

图 1 地球大数据实践体系视图 .....	5
图 2 地球大数据描述的五大圈层 .....	7
图 3 AIoT 技术助力空气质量监测案例 .....	15
图 4 多源数据融合示意图 .....	17
图 5 网格层次索引结构示意图 .....	19
图 6 EFDC 水动力模拟仿真示意图 .....	23
图 7 地球大数据产业图谱 1.0 .....	27
图 8 AIoT 技术助力水环境监测案例 .....	28
图 9 某地基于卫星遥感的甲烷监测结果 .....	30
图 10 移动源远程在线监测 .....	33
图 11 农作物长势监测 .....	35
图 12 土壤墒情监测 .....	35

## 表 目 录

表 1 常用的公开数据的遥感卫星清单 .....	8
表 2 常用的气象数据及其相关链接 .....	9

## 一、地球大数据总述

### （一）地球大数据的定义

地球大数据主要指来自地球大气圈、水圈、土壤圈、生物圈以及人类活动的的数据。这些数据通常具备不同的来源，但都具备时空属性，包括卫星遥感数据、气象数据、人类活动统计数据、地理信息数据等。地球大数据是数据要素产业的重要组成部分，对地球大数据进行收集、整合和分析能帮助人们对生态环境系统、人类社会和自然资源进行更深入的洞察和理解，为人类活动的开展提供决策支持，从而最终实现人与自然和谐共生的现代化。

地球大数据具备传统“大数据”的 4V 特征，即：

**数量 (Volume) 大：**地球大数据具有海量数据的特征，其来源于卫星影像、地理信息、物联网等渠道的数据在空间和时间维度上都具有较大的跨度，量级可达 PB 级。

**种类 (Variety) 多：**地球大数据的数据类型可以是包括文本、图像、音视频、地理位置信息等多种形式，具有种类繁多、结构复杂的特征。

**速度 (Velocity) 快：**地球大数据每天可产生大量的新数据，在生产生活中，为了实时应用新数据辅助决策，其传输、存取、分析对实时性具备较高的要求。

**价值 (Value) 高：**对地球大数据的开发利用可以获取具有高价值的环境信息，为人与环境的交互提供决策支持。

除此以外，相较于其它类别的“大数据”，地球大数据还具备

如下特征：

**时空性：**地球大数据与时间和空间密切相关，具有很强的时空关联性和物理关联性。

**多源性：**地球大数据的来源多样，涉及遥感影像数据、地面监测站点数据、气象数据、社会统计数据等，通常来源于不同的机构。

**普惠性：**相较于其它类型的数据要素，地球大数据的应用领域更加具备普惠性，如助力研究地球的演化过程，分析气候的变化趋势，预测潜在的气象灾害，合理地管理自然资源，从而支撑人与自然关系的可持续发展。

## （二）地球大数据的价值

地球大数据是大数据产业的重要组成部分。地球大数据以对地观测和地理空间数据为主体，是地球环境相关现象的记录。对地球大数据的分析和应用有助于揭示地球的自然环境、资源分布、气候变化、人类活动等方面的规律，并为生态环保、自然资源、应急管理、城市治理等领域提供决策支持。

我国在政策层面对地球大数据的应用实践高度关注。早在 2016 年 3 月，我国环保部即发布了《生态环境大数据建设总体方案》，为生态环境大数据建设提供了强有力的政策支持，这也意味着地球大数据在我国的正式起步。自党的十八大以来，习近平总书记非常重视生态文明建设，明确提出“绿水青山就是金山银山”的重要论

断，并强调“共谋大数据支撑可持续发展之计，加强国际合作”，体现了大数据在支撑可持续发展中所扮演的重要作用。在中共中央、国务院发布的《关于构建数据基础制度 更好发挥数据要素作用的意见》中，也将节能降碳、绿色建造、新能源等可持续发展相关场景列入数据要素赋能的重点领域。2023 年，中共中央、国务院印发《数字中国建设整体布局规划》，其中明确将“建设绿色智慧的数字生态文明”作为数字技术全面赋能经济社会发展的重点，并强调了“推动生态环境智慧治理，加快构建智慧高效的生态环境信息化体系，运用数字技术推动山水林田湖草沙一体化保护和系统治理，完善自然资源三维立体“一张图”和国土空间基础信息平台，构建以数字孪生流域为核心的智慧水利体系”等具体工作方向。

地球大数据的价值主要体现在社会和经济两个方面：

**社会价值方面：**一是地球大数据提供全球范围内的环境监测和资源管理信息，帮助监测和评估环境变化，预测和应对环境问题，支撑环境保护和可持续发展；二是通过地球大数据的分析和预测，可以提前预警自然灾害，支持应急响应和减灾；三是地球大数据可以揭示气候模式、趋势和影响因素，帮助制定应对气候变化的政策和行动；四是地球大数据可以提供城市发展和规划所需的大量地理空间信息，支持城市基础设施建设、交通规划、土地利用规划等方面的决策，促进城市规划和智慧城市建设；五是地球大数据可以提供关于土地利用、森林覆盖、水资源、能源等自然资源的信息，支持资源的科学管理和可持续利用，促进资源的合理分配和保护；六

是对地球大数据的研究应用，能帮助科学家进一步理解地球的结构和运行机制，鼓励公众参与到对地球的环保工作中，有助于科学研究和教育宣传工作的开展。

**经济价值方面：**一是地球大数据可以为农业生产提供土壤水分监测、作物生长预测、病虫害预警等方面的决策支持，提高农作物产量和质量，降低生产成本，促进农业生产和精准农业；二是地球大数据可以提供交通流量监测、交通拥堵预测、优化交通路线等，可以提高城市交通效率，降低能源消耗和交通成本，促进城市交通管理和经济发展；三是地球大数据可以为能源资源勘探开发和管理提供支持，提高能源产量和效率，减少能源浪费和环境影响，推动清洁能源的发展和可持续能源的利用，促进能源开发和管理；四是地球大数据可以为旅游业和文化产业提供包括景点选择、游客行为、文化活动等方面的数据支持，帮助制定旅游线路和推广策略，提升旅游体验和文化产业的发展。

### **(三)地球大数据价值释放的要点**

地球大数据的价值释放包括数据资源、关键技术、行业应用这三个要点。其中，数据资源是指通过卫星遥感、地理信息、物联网等系统获取的数据本身及其加工品。关键技术是指在数据获取、数据管理、数据分析、数据安全等方面相关技术。行业应用是指地球大数据在多个不同领域多种应用场景下的落地方式。下面本节将简要概括这几个要点的内容，具体的内容则会在后续的章节进行详细阐述。



图 1 地球大数据实践体系视图

**数据资源：**地球大数据的数据资源包含了地球大气圈、水圈、土壤圈、生物圈、人类圈等多个圈层的基本信息，以及人类活动与之相关的相互影响关系，可用“天、空、人、地”四个方向概括：

“天”是指来自于航天设施的观测数据，比如卫星遥感影像；“空”是指地球近地面的数据，比如气象数据、空气质量数据等；“人”是指人与自然相互作用产生的数据，包括如农活动相关数据、工业活动相关数据等；“地”是指通过物联网采集的地面数据或来自于地理信息系统的数据。对地球大数据的数据资源的分类，也可以按照其几何特征（即点、线、面、体等）和时空属性（时空静态、空间静止时间动态、时空动态等）进一步总结。在白皮书第二章，我们将进一步介绍地球大数据的数据资源情况。

**关键技术：**地球大数据的应用流程包括了数据采集、数据管理、数据分析、人工智能、数据安全等方面，每一个部分都在传统大数据领域的关键技术体系基础上，纳入了地球大数据领域独特的数据处理技术。在数据采集方面，主要涉及卫星遥感、AIoT 等方面的技术；在数据管理方面，主要包括多源数据融合、地理信息管理、时空数据库等技术；在数据分析和人工智能方面，主要包括基于时空数据的统计分析、可视化、模拟仿真、逻辑推理、机器学习等技术；此外在数据安全方面，还需涉及可信计算技术以及区块链技术等。在白皮书的第三章，我们将具体介绍地球大数据应用流程中的关键技术。

**行业应用：**地球大数据可用于多个对一定区域范围内的自然环境进行分析的行业领域，如生态环保、双碳战略、自然资源、智慧应急、城市更新、农林牧渔、气象服务、孪生流域等。在白皮书的第四章，我们将介绍地球大数据的相关行业及其应用。

## 二、 地球大数据的数据资源

地球大数据主要描述了地球大气圈、水圈、生物圈、土壤圈、人类圈这五个圈层，以及各圈层的相互作用，在数据类型和属性方面有着一定的特殊性，开展地球大数据分析应用实践之前需充分了解。地球大数据按照其数据来源，可以分成“天、空、人、地”四个部分。地球大数据可以根据自身的时空特征包括其几何属性和时空属性进一步抽象化概括。几何属性中，地球大数据可以分成“点、线、面、体”等多种几何形态。在时空属性中，地球大数据可以分

成“时空静态，空间静止时间动态，时空动态”三种形式。



图 2 地球大数据描述的五大圈层

## (一)地球大数据的数据来源

### 1.天

来自于“天”的数据指的是从太空中观测地球产生的数据，主要来源于卫星遥感。卫星遥感是指基于人造卫星的一种非接触性、大范围探测电磁波特性，并通过记录分析、揭示出探测物体的特征及其变化的空间探测技术。自 1957 年前苏联发射第一颗人造地球卫星以来，卫星事业进入了繁荣发展的阶段。利用人造地球卫星搭载的对地观测仪器，人们可以获取与地球系统息息相关的各方面的数据。按照观测对象的不同，卫星还可以分为陆地观测卫星、海洋观测卫星、大气观测卫星等。卫星遥感可以收集到光谱数据，空间数据以及随着时间变化的地球表面观测数据。在表 1 中，我们列举了当前在轨的典型卫星。

这些卫星可以获取地球表面的图像、数据和其他信息，用于环

境监测、资源管理、灾害响应、城市规划、农林业、气象等领域。

表 1 常用的公开数据的遥感卫星清单

分类	卫星	国家/地区	分类	卫星	国家/地区
陆地观测	Worldview2	美国	海洋监测	Sentinel-3A	欧洲
	Worldview3	美国		Sentinel-3B	欧洲
	Quickbird	美国		GOCI	韩国
	Landsat-5	美国	大气观测	Aura	美国
	Landsat-7	美国		Sentinel-5P	欧洲
	Landsat-8	美国		GOME	欧洲
	Landsat-9	美国		OCO-2	美国
	NPP/SUOMI	美国		OCO-3	美国
	Terra	美国		GOSAT	日本
	Aqua	美国		风云三号 D	中国
	SPOT	法国		高分五号	中国
	Sentinel-2A	欧洲		Cloudsat	美国
	Sentinel-2B	欧洲		Calipso	美国/法国
	高分一号	中国		Polder	法国
	高分二号	中国		Parasol	法国
	高分三号	中国		Himawari-8	日本
	高分六号	中国		风云三号 B	中国
	资源一号 A	中国		风云三号 C	中国
	环境一号	中国		风云四号 A	中国
	吉林一号	中国		风云四号 B	中国

来源：整理自网络

## 2. 空

来自于“空”的数据主要是指和大气圈相关的数据，包括气象数据和空气质量数据。

气象数据是描述和记录天气和大气状态的数据，包括温度、湿

度、气压、降水量、风速、风向、云量、可见光等变量。这些气象数据通过气象观测系统、气象卫星、雷达等各种气象设备进行收集和处理，它们对于气象学的研究、天气预报、气候变化分析、农业、航空等领域具有重要的应用价值。常用的气象数据可从多个渠道进行获取，包括中国气象局公开发布的基本气象数据，美国国家环境预报中心的全球数据同化系统 GDAS 数据和全球预测系统 GFS 数据，以及欧洲中期天气预报中心的 ERA5 再分析数据。

表 2 常用的气象数据及其相关链接

数据	来源	相关链接
气象数据	中国气象局	<a href="https://data.cma.cn/">https://data.cma.cn/</a>
GDAS	美国国家环境预报中心	<a href="https://www.ncei.noaa.gov/products/weather-climate-models/global-data-assimilation">https://www.ncei.noaa.gov/products/weather-climate-models/global-data-assimilation</a>
GFS	美国国家环境预报中心	<a href="https://www.ncei.noaa.gov/products/weather-climate-models/global-forecast">https://www.ncei.noaa.gov/products/weather-climate-models/global-forecast</a>
ERA5	欧洲中期天气预报中心	<a href="https://www.ecmwf.int/en/forecasts/dataset/ecmwf-reanalysis-v5">https://www.ecmwf.int/en/forecasts/dataset/ecmwf-reanalysis-v5</a>

空气质量数据是描述和记录空气中各种有害物质浓度的数据，包括颗粒物 PM2.5 和 PM10、臭氧、二氧化氮、二氧化硫、一氧化碳、挥发性有机物 VOCs 等。这些空气污染物数据通常通过空气质量监测站点、移动监测设备和卫星等途径进行收集和记录。

### 3.人

来自于“人”的数据主要反映了人类与自然环境之间的相互作用情况，它们为了解人与自然之间的相互作用、可持续发展的实现以及环境保护与资源管理提供了重要的基础。主要包括几个方面的数据：

**工业活动数据：**如化工污染物排放数据、工业废水排放数据、工业能源消耗数据等。

**农业活动数据：**如农作物种植面积、化肥和农药的使用数据、灌溉用水数据、畜牧业排放数据、由于农业活动导致的土地退化和侵蚀程度等。

**城市化建设活动数据：**如城市扩张、人口迁移、城市人口规模和人口密度、土地利用变化、交通流量、建筑能源消耗等数据。

**能源消费和管理数据：**如能源类型及其消费量数据、可再生能源数据、电力需求和供应数据等。

**消费行为和生活方式数据：**如废弃物生成和处理、个人出行方式、食品消费等。

#### 4.地

来自于“地”的数据主要包括地面的物联网设备采集的数据和地理信息数据。

地面物联网设备（IoT）通过传感器采集各种类型的地表环境数据，包括：

**环境监测数据：**如空气质量、水质、土壤质量、噪声水平、辐射水平等。

**气候气象数据：**气象站点配备了各种物联网传感器，可收集温度、湿度、风速、风向、气压、降雨量、海洋温度、海洋盐度等。

**农业和林业数据：**物联网可以通过土壤湿度传感器、植被生长传感器、病虫害监测设备等收集农作物的生长情况、土壤状况、病

虫害情况等数据；可以监测森林火情、病虫害等情况。

**能源数据：**通过智能电网、智能家居等方式，利用电表、水表、燃气表等设备收集能源使用量的数据，通过太阳能板、风力发电机等设备收集可再生能源的产量数据。

**交通和物流数据：**通过定位系统设备、道路监控设备、交通信号灯等收集车辆位置、交通流量等数据；通过物流追踪设备收集货物位置、状态、温度、湿度等数据。

**城市管理数据：**物联网技术可以用于城市管理，如公共空间中的安全摄像头可以监测公共安全，停车场的传感器可以监测停车状况等。

地理信息数据是描述地球表面和其特征的数据，包括：

**地图数据：**包括地形图、街道地图、航海图等数据。

**地理坐标数据：**包括经度和纬度坐标、高程数据等。

**地貌地形数据：**包括山脉、河流、湖泊、沙漠、森林等自然地理特征的位置和特性。

**土地利用和土地覆盖数据：**包括农田、城市、工业区、公园、湿地等土地使用情况。

**建筑和基础设施数据：**包括建筑物的位置、高度、用途等；道路、铁路、桥梁、管道、电力线路等基础设施的位置和状况。

## **(二)地球大数据的几何与时空属性**

为对地球大数据进一步分类以简化其建模程序和资源消耗，地球大数据的分类可以按照其空间和时间属性进一步抽象为几个不同

类别。在空间属性上，按照数据要素的几何维度，我们可以将其考虑为 0-3 维的数学抽象，即点（0 维）、线（1 维）、面（2 维）、体（3 维）。在时间属性上，我们可以综合其时空属性进行划分，分为时空静止、空间静止时间动态、时空动态这三种类别。

## 1. 几何类型

地球大数据的种类繁多，对地球大数据的分类可以根据其几何属性进行。下面列举地球大数据的几何类型及其相应的数据示例：

**点数据：**即地球表面的离散的点位置数据，可以通过经纬度或其他坐标系统表示。这类数据有城市地标数据，监测站点数据等。

**线数据：**即地球表面上的一条连接两个点或多个点的路径。这类数据可以是道路数据、管道数据、轨迹数据等。

**面数据：**地球表面上的封闭区域。面数据用于表示陆地边界、行政区划、湖泊、建筑物轮廓等地理要素的形状和边界。在面数据中，一类特殊的类型为栅格数据，其由一系列像素组成的网格数据，每个像素代表地球表面的一个小区域。常见的栅格数据由遥感影像、地形模型数据等。

**体数据：**是地球表面上方的具有三维实际意义的数据，包括气象场数据、建筑物数据等。

## 2. 时空类型

对于地球大数据的时空属性，下面列举三种类型及其相应的数据示例：

**时空静止数据：**指在时间和空间上保持静止或不变的数据，通

常用于描述地理实体、属性和现象在特定时间点的状态，如地理位置数据、地形数据、地物分类数据等。

**空间静止时间动态数据：**指在空间上保持静止，但随着时间变化而变化的数据。它可以用来描述地理实体在不同时间点或时间段内的状态变化。这个类型的数据包括：时序地理位置数据、环境监测站点的时序数据、时序遥感影像数据等。

**时空动态数据：**指地理要素在时间和空间上都发生变化的数据。它用来描述地理实体随着时间推移在不同空间位置上的演变和变化。这类数据包括：移动设备轨迹数据、交通流数据、人口迁移数据等。

对于不同的几何属性和时空属性的数据，可以结合起来进一步归结为 12 种具有不同几何时空属性的数据。这些抽象的数据类型，为地球大数据的数据建模提供了范式。

### 三、 地球大数据的关键技术

地球大数据的应用流程包括：数据采集、数据管理、数据分析（人工智能与数据挖掘）、数据安全等步骤。各步骤在通用大数据处理技术基础上，纳入了本领域独特的数据处理技术。在本章节，我们将对重要技术进行梳理。

#### （一）数据采集

数据采集是指通过各种方法和技术收集、记录和获取数据的过程。其涉及到从不同来源和渠道收集数据，为建立数据集、数据库或信息库，以及后续的数据处理、分析和应用奠定基础。在地球大数据的采集过程中，卫星遥感数据采集和 AIoT 数据采集是主要的两

个关键技术。

## 1. 遥感数据采集

卫星遥感是一种利用卫星或者其他航天器搭载的遥感器获取地球表面信息的技术。它通过远距离感知和获取电磁波辐射，从而获取地球表面的图像和信息，以非接触式的方式提供对地表特征、环境、资源和人类活动的观测。遥感器是远距离感知地物环境肤色或反射电磁波的仪器，常见的遥感器有可见光摄像机、红外摄像机、紫外摄像机、红外扫描仪、多光谱扫描仪，微波辐射和散射计、成像光谱仪等。卫星遥感获取的信息进一步通过校正、变换、分解、组合等光学处理或图像数字处理过程，提供给用户分析、判读，或制成专题地图或统计图表，为资源勘察、环境监测、国土测绘等提供信息服务。

## 2. AIoT 数据采集

物联网（IoT）是通过互联网连接和交互的各种物理设备、传感器、对象和系统的网络。物联网技术使得不仅仅是计算机和人类可以通过互联网进行通信和交互，而且物体之间也可以相互连接和通信。在物联网中，物理设备和对象通过嵌入式传感器、标识符和网络连接等技术，能够实时收集和交换数据，通过云计算和网络基础设施进行存储、处理和分析。这样的互联互通使得物联网系统具备感知环境、获取信息、自动控制和实现远程监测的能力。物联网相关的信息传输技术包括 Wi-Fi、蓝牙、Zigbee、LoRaWan 等无线通信技术和协议。

在物联网基础上，人工智能物联网（AIoT）将物联网设备与人工智能技术相结合，实现智能感知、智能决策和智能交互、以更智能化、自动化的方式处理和应用物联网数据。AIoT 的核心思想是利用人工智能技术对物联网设备收集的大量数据进行处理和分析，以提取有价值的信息和洞察，并通过智能决策和控制使物联网系统更加智能和高效。在 AIoT 中，物联网设备通过传感器、节点和连接设备等采集各种类型的数据。这些数据被传输到云平台或边缘计算设备，通过人工智能技术进行分析、学习和推理。AIoT 在多个领域有广泛的应用，如智慧城市、工业自动化、智能交通、智慧农业、健康医疗等。它能够提升物联网系统的智能化水平，改善生活和工作效率，并带来更多的商业机会和创新。

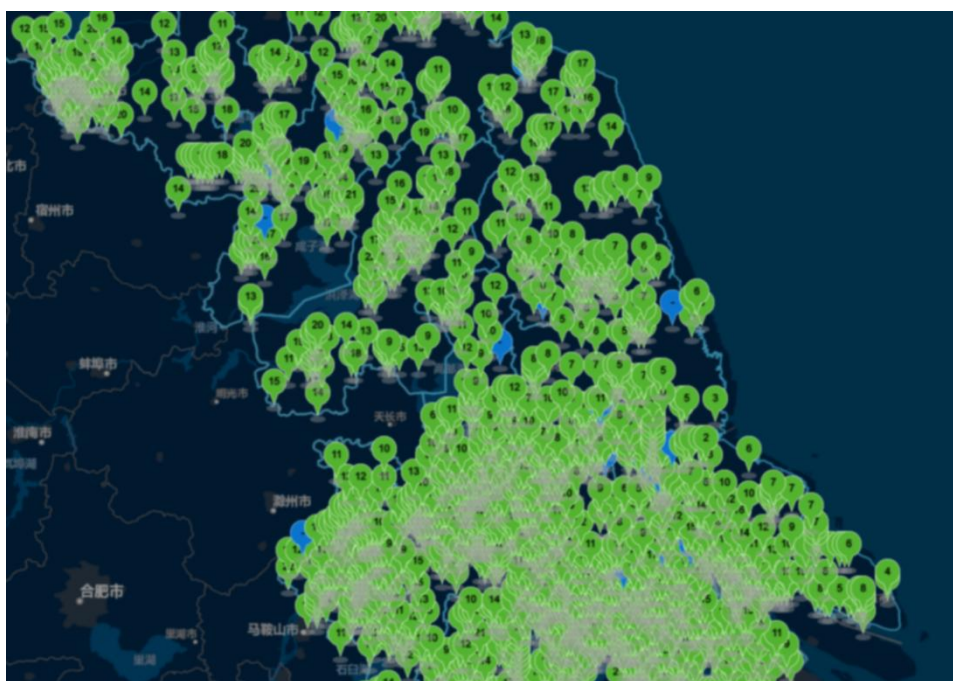


图 3 AIoT 技术助力空气质量监测案例

## （二）数据管理

### 1. 多源数据融合

多源数据融合是指将来自不同数据源和多个数据类型的地球观测数据进行整合和融合，以生成更全面、准确和有价值的信息和知识。这种融合可以在空间、时间和属性维度上进行。在地球大数据中，不同的传感器、设备和技术收集到的数据具有多样性，包括遥感影像、气象观测数据、传感器网络数据、社交媒体数据等。这些数据可能具有不同的空间分辨率、时间分辨率、观测范围和属性特征。多源数据融合的目的是通过将这些数据整合起来，消除数据间的差异性和不一致性，增加数据的覆盖范围和精度，提高数据的可信度和可用性。多源数据融合主要可以实现以下目标：

**数据补充和增强：**通过融合多源数据，可以填补数据的空缺、提高数据的空间分辨率和时间分辨率，增强数据的信息内容和质量。

**信息提取和知识发现：**多源数据融合可以融合不同数据源中的信息，从中提取有用的特征、模式和关联。通过数据融合，可以发现隐藏在数据中的关联和趋势，提供更全面的地球观测信息和洞察。

**数据一致性和准确性：**不同数据源可能存在数据间的不一致性，如不同的坐标系统、观测误差等。多源数据融合通过校正、配准和一致化处理，可以提高数据的准确性和一致性。

**业务应用和决策支持：**多源数据融合可以为不同领域的业务应用提供更全面和综合的数据支持，如环境监测、资源管理、灾害响应、城市规划等。融合后的数据可以为决策制定者提供更准确和全

面的信息，支持决策的制定和优化。

在实际应用中，多源数据融合技术是地球大数据在各个应用场景中的关键一环。这体现在多个方面：其一，多源数据融合技术（如变分同化、卡尔曼滤波等方法）提升了获取数据的准确性，为后续利用数据进行分析以及人工智能应用提供了更有价值的信息；其二，多源数据融合技术可以全局展示全方位的信息，为实际应用中更科学合理的决策提供必要的保障。下图展示的是一个多源数据融合的例子。

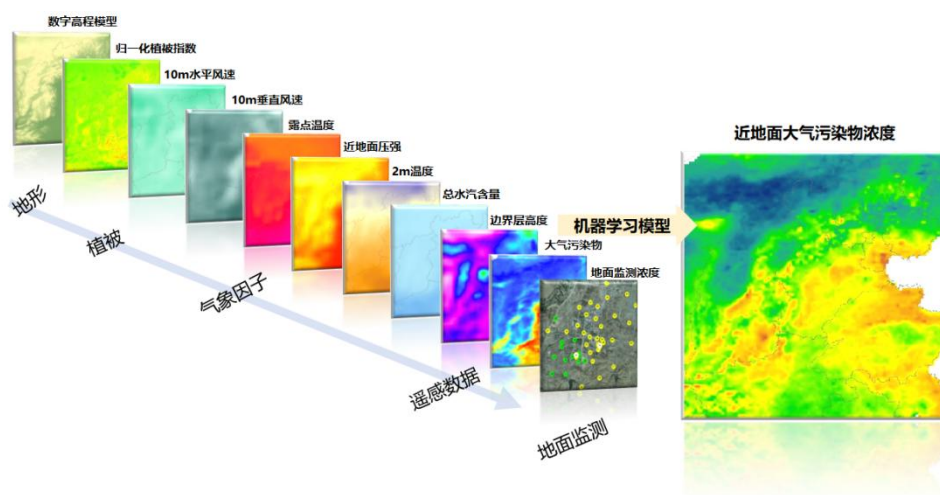


图 4 多源数据融合示意图

## 2.地理信息管理

地理信息系统（GIS）是一种用于收集、存储、管理、分析和展示地理空间数据和与之相关的属性数据的技术系统。它包含了地理学、地图学、数据库技术和计算机科学等技术，用于处理和分析与地理位置有关的信息。地理信息系统广泛应用于地球大数据的多个领域，如城市规划、环境管理、农业、林业、水土保持、地质勘探、

交通规划、电力网络管理、地理教育等。它提供了对地理空间数据进行存储、分析和可视化的能力，为理解和解决与地理位置相关的问题和挑战提供了重要的工具和方法。

在地球大数据的实际应用中，人们广泛地使用 GIS 作为数据可视化的基座，为生态环境一张图，智能决策驾驶舱等相关场景应用提供了技术支持。

### 3. 数据存储：时空数据库

时空数据库是一种用来有效存储和处理与时间和空间位置相关数据的数据库。时空数据库与传统数据库的本质区别在于其采用特殊的索引和查询技术来管理大规模的时空数据。时空数据库关键要点是建立数据的快速索引引擎，这可以利用层次存储结构结合时空索引来实现。

层次存储结构是指将空间划分为多个网格，每个网格可以再进一步细分成多个子网格，以此类推将研究区域用多层次网格形式表示。常用的网格有正四边形网格、六边形网格、三角形网格等。正四边形网格是将地球的墨卡托投影平面划分成多个正四边形，每个四边形里再进一步划分成四个正四边形，采用这样的层级划分方式得到对不同区域的空间存储方式，常用的时空数据库 GeoHash 即采用这种正四边形的层次存储方式。类似的，六边形、三角形网格也是将地球表面划分成多个层级的六边形、三角形。常用的层次存储方式有 GeoHash, Google S2, Uber H3 等。

构建时空数据的快速索引引擎的另一个关键是创建时空索引。

常用的时空索引方式有 R 树、四叉树、空间填充曲线（space filling curves）等。其中空间填充曲线是一种用来建立有效索引的重要工具，其关键在于能够建立多维空间和一维数据的映射，同时保留数据的邻近性，即，我们可以通过空间填充曲线将地理坐标（经纬度）映射到一维的索引值，并用这个一维的索引值唯一确定该地理坐标，从而对该地理位置的经度和纬度查询转换为对该一维索引值的查询。常用的空间填充曲线有希尔伯特曲线（Hilbert curves）和 Z-order 曲线。

在地球大数据应用领域，时空数据库是必不可少的一个关键技术。它作为地球大数据多场景应用的基础设施，为实际应用中快速查询定位业务所需的时空数据提供了技术保障。

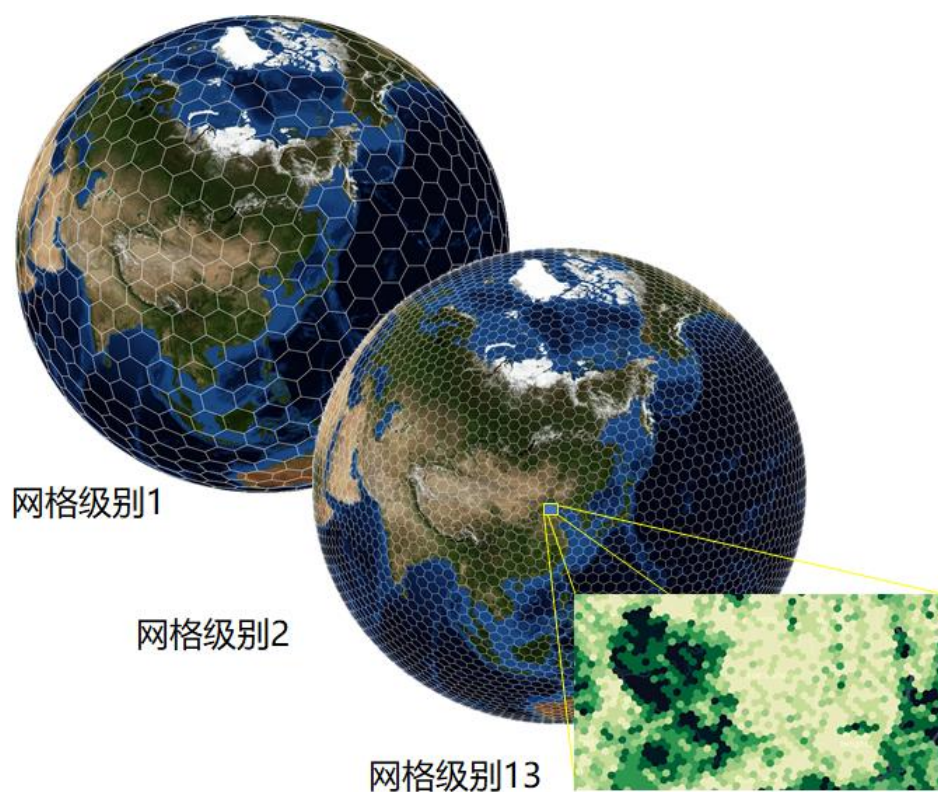


图 5 网格层次索引结构示意图

### （三）数据分析

数据分析相关技术可以帮助人们理解和解释复杂的地球大数据，是地球大数据应用过程中的重要一环。围绕地球大数据的分析技术可以进一步细分为通用统计分析、地理信息分析、可视化交互、模拟仿真四个部分。

#### 1. 通用统计分析

统计方法是分析地球大数据的重要工具。它可以用来总结和描述地球大数据的主要特征，检验地球数据相关的假设，研究数据的时空分布等。地球大数据中常用的统计方法有：描述性统计分析、时间序列分析、空间统计模型等。描述性统计分析可以用来刻画地球大数据的统计量，便于直观了解数据本身的基础信息；时间序列分析则侧重于利用数据时间维度上的相关性来进行分析，比如自回归模型  $AR(p)$ 、移动平均模型  $MA(q)$  等；空间统计模型则可以用来刻画空间维度上数据的相关性，常用的有克里金（Kriging）方法、空间自回归（Spatial Auto-Regressive）模型、地理权重回归（Geographical Weighted Regression）等。

#### 2. 可视化交互

可视化交互技术能对地球大数据和其分析结果进行直观的展示。通常涉及 3D 可视化技术，如 Google Earth Engine 的三维展示，Cesium 三维可视化开发工具等，以及虚拟现实（VR）和元宇宙（Metaverse）等技术。

可视化交互是地球大数据应用场景中的一项关键技术。其可以将对地球大数据的分析结果以丰富多彩的图像/影像等方式展示，并提供数据交互渠道，使得人们可以更为直观且便捷地了解地球大数据的分析结果。

#### **(四)人工智能**

人工智能（AI）在地球大数据领域扮演了重要角色。人工智能技术能够从复杂多源异构的具有时空属性的地球大数据中发现有价值的模式和信息，为各种应用和决策提供支持。传统的人工智能在地球大数据背景下，继续发展成为具有时空智能的人工智能技术。

下面几个小节，我们将介绍地球大数据中的人工智能技术。

##### **1.基于知识的逻辑推理**

基于知识的逻辑推理（Knowledge-based Logical Reasoning）可以帮助我们处理地球大数据中的海量数据，并将其转换成知识体系。其主要有几个步骤：1）知识表示：需要把现有的知识转换成计算机可以理解的形式，可以通过逻辑公式、图形模型、语义网络等形式；2）推理算法：利用逻辑推理算法来处理这些现有的知识，并将其转换成结果；3）解释与验证：对推理的结果进行解释和验证，以确保其正确性和可靠性。通过上述三个步骤的持续迭代，从而从海量数据中推演出有价值的信息，并扩充我们的知识体系。常用的基于知识的逻辑推理方法有知识图谱、演绎推理、神经符号学习、因果推理、逻辑神经网络、知识蒸馏等。

## 2. 机器学习

在地球大数据领域，机器学习方法可以用来处理时空数据、并利用时空相关性进行预测或决策。具有时空智能的机器学习方法有时空点过程模型（ST-point process）、随机森林（random forest）、支持向量机（SVM）、K-means 聚类、深度学习中的卷积神经网络（CNN）、时空卷积神经网络（ST-CNN）等。

机器学习（包括深度学习）等技术为人们更好地利用地球大数据提供了新的发展机会。比如，在生态环保领域，人们可以对卫星遥感数据应用机器学习算法（如随机森林、支持向量机等），以实现地物识别，准确掌握潜在污染源的地理分布。近年来，基于深度学习框架实现的基础大模型也在地球大数据领域具有广阔的应用前景，比如盘古气象大模型，为台风路径预测、海浪预测等方面提供了更准确的预测结果。

## 3. 模拟仿真

对地球大数据的开发利用过程中，常用机理模型来模拟地球系统的运行机制，从而进行预测性分析。机理模型是基于科学原理和物理规律构建的数学模型，它通过建立各种物理、化学、生态等方面的方程和参数，模拟地球系统的行为，从而揭示地球系统中的关联性、动态变化和未来趋势。常用的机理模型有：1）气候模型，用于模拟气候系统，如 WRF，WRF-Chem，CMAQ，CAMx 等；2）水文模型，用于描述水在陆地表面的运动和分布的模型，包括降雨、蒸发、径流和地下水流等过程，如 MIKE 21，MIKE SHE，Delft3D 等；3）生

态系统模型,用于描述生态系统的结构和功能的模型,如 Biome-BGC, BioMod2 等。这些机理模型帮助我们对地球系统有更深入的了解。

模拟仿真是地球大数据应用的重要技术。对现有数据的建模分析,可以帮助人们更深刻地了解地球各个部分的演变,包括对过去历史发展规律的追溯和总结,以及对未来可能的发展趋势和结果的预测。这在各个应用场景中都是关键的一环。

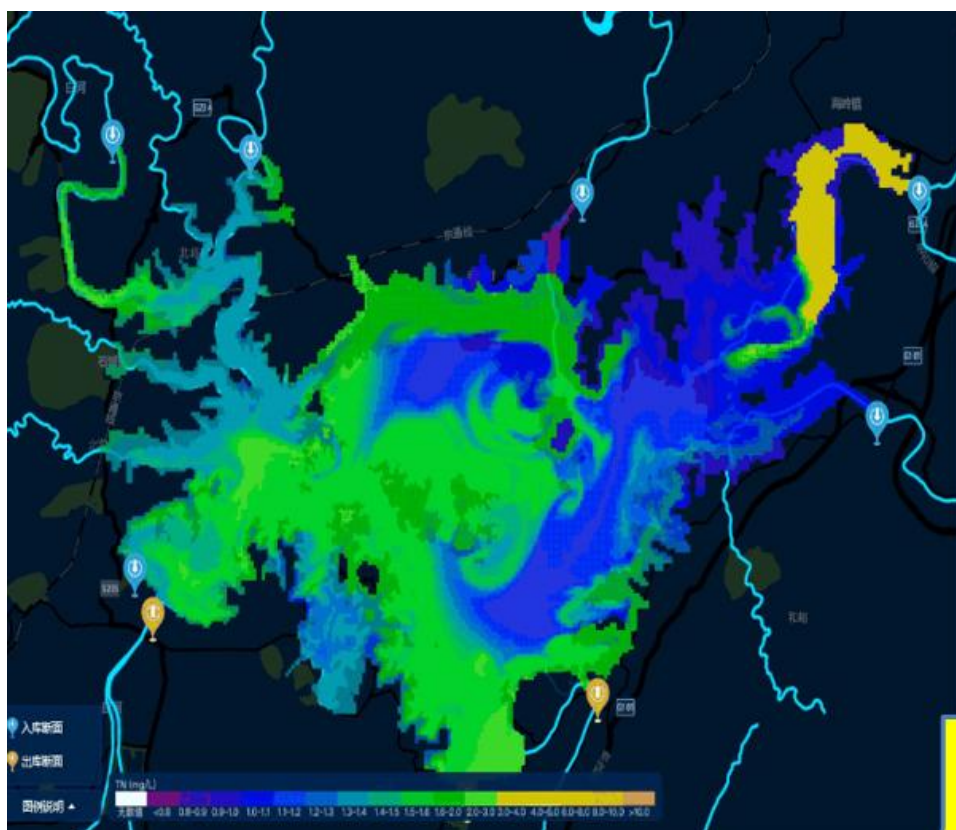


图 6 EFDC 水动力模拟仿真示意图

## (五)数据安全

数据安全是地球大数据关键技术的重要组件。在收集、存储、处理和传输地球大数据时需采取各种措施和方法,以确保数据的机密性、完整性和可用性,防止未经授权的访问、篡改、泄露和破坏。各个国家和地区对数据安全都有相应的立法,比如 2021 年 6 月 10

日全国人大常委会通过了《中华人民共和国数据安全法》，对数据安全的重视提升到了国家战略层面；欧盟也在 2016 年发布了《General Data Protection Regulation》，即通用数据保护条例，对涉及到的数据隐私等相关内容做了阐述。下面我们主要介绍区块链和可信计算这两种用于保障数据安全的技术。

## 1.可信计算

可信计算（Trusted Computation）是一种为计算过程和数据提供安全保护的计算机技术，旨在确保计算过程和计算结果的可信性和安全性，其在云计算与物联网等方面具有重要应用。

可信计算的一个关键技术是可信执行环境（Trusted Execution Environment, TEE）。它提供了一种安全的执行环境，保护计算机中的代码和数据免受恶意攻击和未经授权的访问。TEE 由硬件（例如 Intel SGX, ARM TrustZone）和软件（例如 Intel Software Guard Extensions）组成，采用隔离和加密技术确保在环境内运行的代码和数据不受外部威胁的影响。TEE 通过硬件隔离或虚拟化技术将其自身与操作系统和其他应用程序隔离开来。这种隔离防止了来自其他软件或系统的恶意攻击和干扰，确保 TEE 中的代码和数据安全可靠。TEE 使用加密技术来保护存储在其中的敏感数据。数据在进入 TEE 之前进行加密，只有在 TEE 内进行解密后才能被使用。TEE 的启动过程也是一个重要环节，确保 TEE 在正常运行之前没有被篡改或受到恶意软件的影响，这通常涉及硬件或固件级别的校验和数字签名验证，确保 TEE 的完整性和真实性。此外，TEE 在运行时会监控自身的

状态，以监测是否受到任何未经授权的修改或篡改。如果发现任何可以更改，TEE 可能会拒绝继续执行，以保护其完整性。

可信计算作为物联网技术的关键部分，在地球大数据应用场景中也有重要应用。它确保了所获取的地球大数据的真实性和有效性，为地球大数据的数据安全提供了技术保障。

## 2. 区块链

区块链技术是一种去中心化的分布式账本(distributed ledger)技术，通过密码学、共识算法和网络协议等技术手段，实现了多方之间的信任和数据共享。区块链是由一系列按时间顺序链接的数据块组成的链式结构，每个数据块包含了一批交易记录。每个数据块都包含一个唯一的标识符（哈希值），该标识符由数据块的内容和前一个数据块的标识符计算得出。这种链接关系使得区块链中的数据具有连续性和不可篡改性。区块链使用了多种密码学算法来确保数据的安全性和隐私性。其中，哈希函数是常用的密码学工具，用于将数据转化为固定长度的哈希值。常见的哈希函数有 SHA-256 和 SHA-3 等。非对称加密算法（如 RSA 和椭圆曲线加密）用于实现身份验证、数字签名和加密通信等功能。在区块链中，共识算法用于解决分布式网络中多个节点之间的数据一致性和信任问题。常见的共识算法有 PoW (Proof of Work) 和 PoS (Proof of Stake)。区块链是建立在分布式网络上的，它由许多节点组成，每个节点都维护了完整的账本副本。节点之间通过点对点的通信协议进行数据传输和共识达成。分布式网络的特点是去中心化、鲁棒性强，没有单点

故障。区块链根据权限和访问控制的不同，可以分为公有链、私有链和联盟链。公有链是开放的，任何人都可以参与其中，如比特币和以太坊；私有链是限制访问的，只有特定的节点可以参与，如企业内部的区块链；联盟链是由多个组织共同管理的区块链，参与者需经过授权。

区块链具备数据可追溯、多方信认等诸多特性，适用于需要各方协同的地球大数据相关项目中，如在双碳领域记录各方碳排放量、碳交易情况等，从而构建一个多方背书的信任体系。

#### **四、 地球大数据的行业生态与行业应用**

##### **(一)地球大数据产业图谱**

为促进地球大数据产业发展，中国通信标准化协会大数据技术标准推进委员会地球大数据工作组于 2023 年 8 月开展了《地球大数据产业图谱 1.0》的编制工作，图谱覆盖了地球大数据产业的各个环节，共收录了 30 余家领域内头部机构，旨在梳理地球大数据行业生态，为相关领域的研究和发展提供参考。



图 7 地球大数据产业图谱 1.0

## (二)地球大数据行业应用

地球大数据在各个行业中有着广泛的应用。通过整合、分析和挖掘海量的地球观测数据和其他相关数据，地球大数据可以提供深入的洞察和智能化的决策支持，帮助实现更高效、更可持续的发展。本节将讨论地球大数据在生态环保、双碳战略、自然资源、气象服务、城市更新、农林牧渔、智慧应急、孪生流域等方面的应用。

### 1.生态环保

生态环保领域，地球大数据作为直接的业务支撑数据具有重要作用。下面具体介绍几个相关的应用。

**生态系统监测与保护：**地球大数据可以用于监测和评估各种生态系统的健康状况和变化趋势。通过卫星影像、遥感数据和智能物联网，可以获取大规模的生态数据，如空气质量、水体水质、森林覆

盖、湿地变化、土地利用和植被指数等。利用人工智能算法对这些数据进一步分析，可以了解评估空气质量的变化状况、水体范围的演化、森林覆盖的变化情况、土地利用类型的演变、以及植被变化等，帮助科学家和环境保护机构了解生态系统的动态变化，及时发现和应对生态破坏和环境污染等问题。

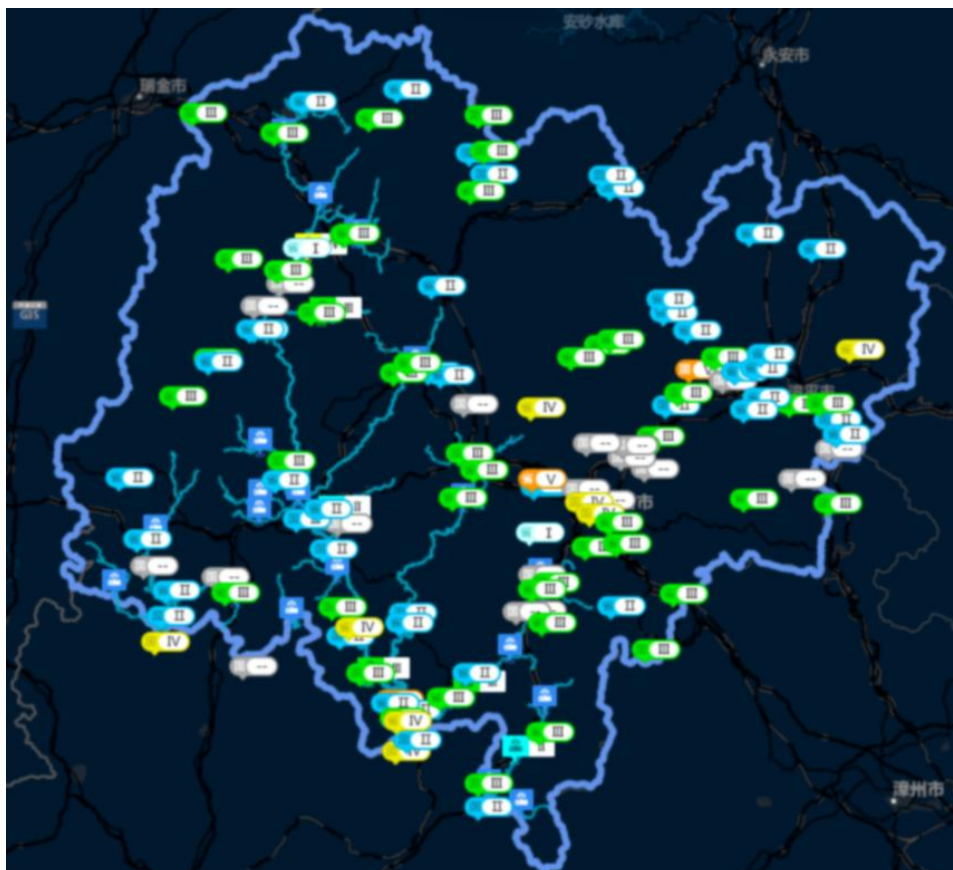


图 8 AIoT 技术助力水环境监测案例

**污染监测与环境治理：**地球大数据可以通过遥感数据、传感器数据和地理信息系统，监测大气污染、水污染和土壤污染等问题。对地球大数据的分析，可以有效识别污染高值区，帮助环境执法部门按图索骥进行环境执法，可以探索污染发生的模式，判别潜在的污染源及其扩散趋势，从而达到“未污先治”的效果。

**生物多样性与物种保护：**地球大数据可用于监测生物多样性和保

护濒危物种。卫星影像和物联网传感器数据可以获取物种的栖息地变化数、迁徙路径和数量变化等数据。对这些数据的分析处理，对于制定物种保护计划、评估保护区的有效性以及研究生物多样性和生态系统功能具有重要意义。

**水土流失状况监测评价：**地球大数据可用于典型生态系统水土保持功能评价。以遥感影像为基础，融合土地利用类型、植被覆盖度和地形等数据，开展农田、森林、草原、荒漠等生态系统水土流失状况评价，定量评价水土保持工程实施效益，从水土保持功能角度反映生态系统质量、稳定性状况及提升方向。

**应对气候变化：**地球大数据对于研究气候变化、评估气候风险以及制定适应措施具有重要意义。通过收集和分析大量的气象数据、大气数据和海洋数据，可以揭示气候变化的模式和趋势，帮助科学家和政策制定者了解气候变化对生态系统的影响，制定相应的应对策略。

## 2. 双碳战略

在 2020 年 9 月，习近平主席在第七十五届联合国大会一般性辩论上宣布中国“碳达峰”和“碳中和”目标。双碳战略正获得越来越多的人认可并积极参与其中。地球大数据有效地支撑中国双碳战略的实施，具有重要的应用价值。其主要有如下应用场景：

**温室气体排放监测：**地球大数据可以用于监测和评估温室气体的排放情况。通过卫星遥感和地面监测站点，可以获得空气质量数据，获取各个地区和不同行业的温室气体排放数据。通过对这些数据的

分析，可以精确了解掌握不同行业和企业的具体排放情况，帮助政府部门制定减排策略和监测减排效果。



图 9 某地基于卫星遥感的甲烷监测结果

**能源管理与优化：**地球大数据通过监测能源消耗数据、能源生产设施的运行状况和能源转换效率等，可以提供节能减排的建议和优化方案。对在能源领域获得的地球大数据的分析和建模可以帮助优化能源供应链和能源系统，减少不必要的能源消耗和碳排放。

**可再生能源规划：**地球大数据通过分析太阳辐射、风速、地形等数据，可以帮助确定适合建设太阳能和风能发电设施的地点。这样可以有效利用可再生能源资源，减少对化石燃料的依赖，降低碳排放。

**碳排放交易和碳市场：**地球大数据可以在碳排放交易和碳市场中发挥重要作用。通过监测和验证碳排放减少的效果，可以生成可信

的碳排放证书。对碳排放数据的分析，也可以识别和检测异常数据，有效监管碳排放交易。

**生态系统碳储量评估：**地球大数据可以用于评估生态系统的碳储量。通过卫星遥感数据、激光雷达扫描和地面调查等手段，可以估算出森林、湿地和土地等生态系统的碳储量。这些数据可以帮助监测碳汇、评估生态系统服务和制定相应的保护和恢复措施。

### 3. 自然资源

地球大数据在自然资源领域也具有广泛应用，可以帮助管理和保护自然资源，促进可持续利用。下面列举一些具体的应用场景：

**土地利用和覆盖监测：**地球大数据可以用于土地利用和覆盖的监测和评估。通过卫星遥感数据、空间图像和地面调查等，可以实时获取土地利用变化、森林覆盖、农田分布和城市扩张等信息。这些数据可以帮助决策者进行土地规划、资源管理和生态保护，制定合理的土地利用政策和保护措施。

**水资源管理：**通过监测水体质量、水位、流量和降雨量等数据，可以实时了解水资源的状况，并进行合理的水资源分配和管理。地球大数据还可以用于湖泊和河流的水生态保护，帮助评估水生生态系统的健康状况、监测水生态变化和保護水生生态多样性。

**人为水土流失监管：**地球大数据可以用于人为水土流失监管，通过卫星遥感数据、空间图像和现场核查等，及时发现、制止和查处生产建设活动水土保持违法违规行爲，根据管理需求实现对人为水土流失情况监控实时化、智能化。

**矿产资源勘探和管理：** 通过卫星影像、地质调查数据和地球物理探测等技术，可以帮助发现和评估矿产资源的分布和储量。这些数据可以用于矿产资源的合理开发、矿区环境保护和矿产资源管理规划。

**生物多样性资源：** 地球大数据还可以用于评估和监测生物多样性的状况和变化，帮助相关部门掌握生物资源、物种分布等信息。

#### 4. 气象服务

地球大数据可以用于提高气象预报的准确性、提供实时的气象信息和支持气候变化研究。以下是一些具体的应用场景：

**天气预报和预警：** 地球大数据可以用于天气预报和气象预警。通过收集和分析大量的气象数据，包括气象观测、卫星图像、雷达数据等，可以建立气象模型和预测算法，提供准确的天气预报。地球大数据还可以用于气象灾害的预警，如暴雨、风暴和台风等，及时通知人们并采取相应的应急措施。

**气候监测和分析：** 地球大数据通过长期的气象观测数据、历史气候记录和遥感数据，可以分析气候的长期趋势和变化规律。这些数据可以用于气候变化研究、制定气候适应策略和评估气候风险。

**气象灾害风险评估：** 通过整合气象数据、地理信息和人口分布等，可以评估不同地区和人群面临的气象灾害风险。这些数据可以用于制定应急管理计划、提前预警和减轻灾害影响。

**农业气象服务：** 通过气象数据和农业模型，可以提供农业气象预报和农作物生长模拟。农民可以根据气象预报和农业建议，合理安排

排农事活动，决策灌溉和施肥，提高农作物产量和质量。

**航空和航海气象服务：**地球大数据在航空和航海领域提供了重要的气象服务。通过收集和分析大气数据、风场数据和海洋气象数据等，可以提供准确的航空和航海气象预报。这些数据可以帮助航空和航海运输规划航线、优化飞行和航行安全。

## 5.城市更新

在城市更新领域，地球大数据可以发挥巨大作用。具体表现在几个方面：

**智慧交通管理：**地球大数据可以用于交通管理，包括交通流量监测、交通拥堵预测和智能交通信号控制。通过使用传感器、摄像头和车载设备等收集交通数据，可以实时监测道路状况、交通流量和拥堵情况。这些数据可以用于优化交通信号灯控制、提供实时交通信息和推荐最佳出行路线，减少交通拥堵和减少尾气排放。



图 10 移动源远程在线监测

**环境监测和治理：**通过监测空气质量、噪声水平、水质和污染源

等数据，可以实时了解城市环境状况，用于及时预警和治理环境污染，改善居民的生活质量和健康状况。

**城市规划和土地利用：**地球大数据可以用于城市规划和土地利用的优化。通过分析人口密度、建筑分布、土地利用数据和交通网络等，可以帮助规划者制定合理的城市发展规划和土地利用策略。这些数据可以支持可持续城市发展，提高城市的空间利用率和环境友好性。

**公共安全和应急响应：**地球大数据通过整合各类数据，包括监控视频、社交媒体数据和传感器数据，可以实时监测城市的安全状况、异常事件和灾害风险。这些数据可以用于提供实时警报、指导应急响应和优化公共安全管理。

## 6. 农林牧渔

地球大数据可以帮助提高农业、林业、畜牧业和渔业的生产效率、资源管理和可持续发展。以下是一些具体的应用场景。

**农作物监测和管理：**地球大数据通过卫星遥感数据、气象数据和土壤传感器等，可以实时监测和评估农田的土壤湿度、植被指数、气候条件和病虫害情况等。这些数据可以帮助农民优化灌溉、施肥和农药使用，提高农作物的生长质量和产量。

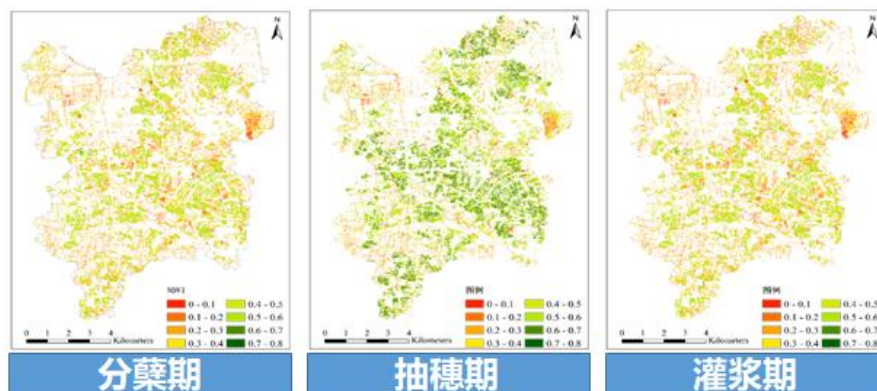


图 11 农作物长势监测

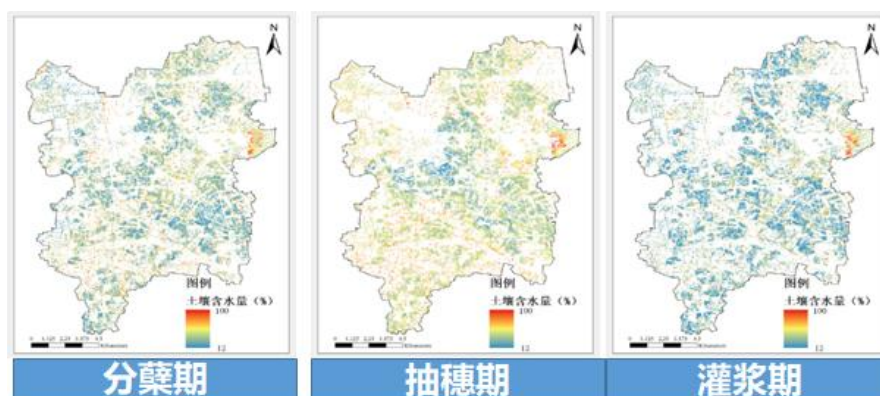


图 12 土壤墒情监测

**精准农业和智能农业：**通过使用传感器、物联网和数据分析，可以实现对农田、养殖场和渔场的精确监测和控制。农民可以根据数据分析的结果，精确施用农药、肥料和水资源，提高资源利用效率，减少环境污染。

**林业资源管理：**通过卫星遥感数据和激光雷达扫描等技术，可以实时监测森林覆盖、森林类型、树种分布和森林生态系统的健康状况。这些数据可以用于森林资源规划、病虫害监测、森林火灾预警和森林保护。

**畜牧业管理和养殖监测：**通过使用传感器和监测设备，可以实时监测牲畜的健康状况、饲料摄入量、运动情况和环境条件等。这些

可以帮助养殖场主和畜牧业者及时发现异常情况，预防疾病爆发和提高养殖效率。

**渔业资源管理和渔情监测：**通过卫星遥感数据、海洋气象数据和渔业数据等，可以实时监测海洋渔业资源的分布、数量和健康状况，用于渔业资源保护、渔船调度和渔情预警，促进渔业可持续发展。

## 7. 智慧应急

地球大数据可以帮助提高应急响应能力和减轻灾害影响。

**灾害监测和预警：**地球大数据可以用于监测和预警各类自然灾害，如洪水、地震、风暴和山火等。通过卫星遥感数据、气象数据、地震监测和火灾监测等，可以实时获取灾害相关的信息。用于提前预警、监测灾害演变趋势，并及时采取措施减少损失。

**灾害影响评估：**地球大数据可以用于评估灾害对人口、基础设施和环境的影响。通过卫星图像、激光雷达扫描和无人机图像等，可以快速获取受灾区域的情况，包括人口密度、道路状况、建筑损毁和环境比那花等。帮助决策者了解灾害影响范围和成都，有针对性地采取救援和恢复措施。

**应急资源调度和管理：**通过实时监测和分析人口密度、交通状况和物资供应数据，可以优化应急资源地分配和调度。数据分析和预测模型可以帮助决策者做出准确的资源调度决策，提高应急响应效率。

**人员定位和救援：**通过移动通信数据、社交媒体数据和卫星定位数据等，可以追踪受灾人员的位置和行动。这些数据可以用于指导

救援队伍和优化救援路线，提高人员救援的准确性和时效性。

**灾后评估和恢复规划：**通过获取灾后影像和地理数据，可以快速评估灾害造成的损失和需要恢复的区域。可以帮助制定恢复规划、优先分配资源和监测恢复进度。

## 8. 孪生流域

地球大数据在孪生流域的应用可以帮助实现流域管理的科学化、精细化和可持续发展。

**水资源管理：**地球大数据可以用于水资源管理。通过收集和分析流域内的水文数据、气象数据和水资源利用数据，可以了解流域的水资源分布、变化和利用情况。这些数据可以用来制定流域水资源管理策略、优化水资源分配和决策支持。

**洪水和干旱预警：**通过监测流域的降雨量、河流水位、土壤湿度等数据，可以实时了解流域的水文情况，并进行洪水和干旱的预测和预警。

**流域生物多样性保护：**地球大数据可以支持流域的生态环境保护。通过卫星遥感数据、生态监测数据和物种分布数据等，可以评估流域的生态状况、物种多样性和生态系统服务价值。

**水质监测和污染防控：**通过收集和分析流域内水体的水质数据、污染源分布等，可以了解水体的污染程度和来源。用于制定水质监测计划、污染防治措施和水环境管理策略。

## 五、 总结与未来展望

### （一）总结

本白皮书主要探讨了地球大数据的概念和其应用。在本白皮书中，我们先介绍了地球大数据的基本定义和范畴，阐述了地球大数据的社会价值和经济价值。随后，我们介绍了地球大数据的数据资源和关键技术，并进一步阐述了地球大数据在各个领域的广泛应用。本白皮书旨在为读者提供一个初步的了解地球大数据这一重要概念的渠道，对地球大数据的研究在持续进展中，后续将继续更新地球大数据的研究进展。

### （二）地球大数据的发展趋势展望

地球大数据是大数据领域的一个重要分支，其发展受到诸多因素的影响。可以预见的是，在地球大数据产业的不断发展过程中，其产业融合、技术进步的趋势始终保持着极高的发展势头。

#### 1. 产业趋势

**垂直领域的深度应用：**随着地球大数据的发展，其在生态环保、城市更新、气象服务、自然资源等领域的应用将得到深化。例如，通过分析和理解大规模气候数据，我们可以更好地预测和应对极端天气和气候变化。通过分析城市数据，我们可以更科学地进行城市规划和管理。

**跨行业的整合与合作：**随着数据的开放和分享，地球大数据将促进不同行业、不同领域的整合与合作。比如，环保领域的数据可以与健康领域的数据结合，帮助我们更好地理解环境污染对人类健

康地影像。

**数据服务市场的发展：**随着地球大数据地价值越来越被认识，数据服务市场将得到发展。企业将提供数据收集、处理、分析、咨询等服务，帮助政府、企业和公众更好地利用地球大数据。

## 2. 技术趋势

**人工智能与机器学习：**人工智能技术将在地球大数据地分析和应用中发挥关键作用。例如，机器学习算法可以帮助我们从小规模数据中提取模式和知识，预测未来的趋势。

**面向地球大数据的 AIoT：**结合地球大数据的应用场景，AIoT 具有巨大的发展潜力。未来面向地球大数据的 AIoT 将更注重多模态数据融合、更全面的智能决策、更安全的隐私保护、更高效的自主学习与自适应系统。

**地球大数据的管理：**在地球大数据的管理方面，如何构建一套多层次全覆盖的离散全球网格化系统受到研究者的广泛关注。面对地球大数据的高效管理的实际需求，研究者们正在加速开发更简单易用等优势离散全球网格化系统。此外，由于地球大数据本身具备的四维时空属性，如何构建具备实现数据融合和快速索引数据的时空数据库也是当前受到广泛关注的热点问题。

**数据安全和隐私保护：**随着数据安全和隐私保护问题的日益受到关注，新的技术和方法将被开发出来，以保护数据的安全和隐私。例如，区块链技术可以帮助我们确保数据的完整性和安全性。

**数据可视化：**数据可视化技术使我们更容易理解和使用地球大数据。例如，虚拟现实和增强现实技术可以帮助我们以更直观的方式展示和探索复杂的地球数据。

## 附录：地球大数据应用案例选集

### 案例一：阿里云 – 三峡数字孪生流域案例

#### 1. 案例背景：

2021 年国家水利部召开了推进数字孪生流域建设工作会议，要求大力推进数字孪生流域建设；在国家政策推动下，各流域管理机构、地方水行政主管部门和有关水利工程管理单位陆续开展数字孪生流域先行先试项目，水利行业掀起了建设数字孪生流域的高潮。

当前，水利行业的数字流域建设还存在如下几个难点：

（1）地理系统与数据分散。地理信息数据来源众多、来源系统割裂，数据未进行有效整合，未形成统一的数字孪生底座。

（2）推演仿真系统使用门槛较高。原有水情推演系统更偏重静态、离线分析，对于使用人员有非常高的专业操作要求，无法简单快速得到推演结果。

（3）计算结果不直观，无法高效的进行对比分析。在完成区间交互、水位/流量控制、调洪演算和调度调优等算法推演之后，缺乏有效的视觉辅助，对计算结果无法进行快速评估。

#### 2. 地球大数据应用：构建长江数字孪生流域

长江是亚洲和中国的第一大河，世界第三大河。发源于青海省唐古拉山，最终在上海市崇明岛附近汇入东海。长江全长约 6300 千米，长江干流宜昌以上为上游，长 4504 千米，流域面积 100 万

平方千米。如何快速精准的构建全流域数字孪生底座是数字流域项目的重点与难点之一。

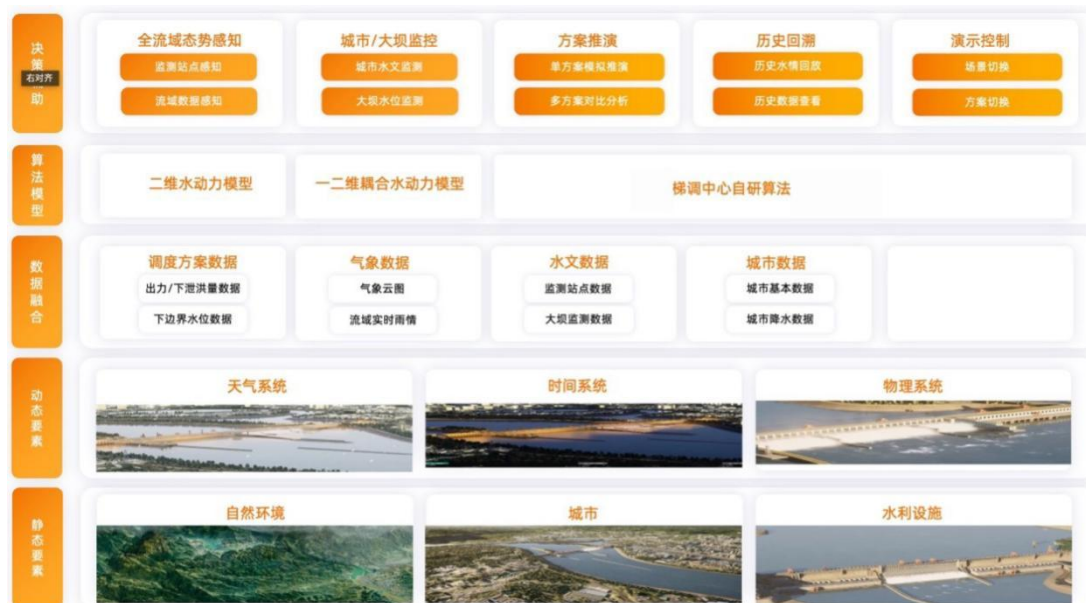
构建全流域数字孪生底座的第一步，是要融合不同类型的空间数据。GIS 行业在过去几十年的发展过程中，不同的地理信息系统大多是采用不同的数据格式，基于相互独立的、具体的和封闭的平台开发的，他们之间没有数据的交互性，因此地理数据的组织存在很大的差异和不同。在数字孪生领域，GIS 数据无疑是必备要素之一，从卫星影像、高程地形、二三维矢量地图到倾斜摄影、高精路网等数据类型，每一种数据都具有不同的数据格式和使用方法。

在基于阿里云 DataV 数字孪生平台完成数据融合之后，利用 DataV 的三维自动化生成能力，实现了全流域 100 万平方千米地形、流域沿线 7 座重点城市、6 座大坝的快速程序化建模，将场景建模时间从数月降低到数周。同时，基于数据驱动的数字孪生场景与传统的纯手工建模相比，极大提升了场景模型的地理位置精准度，同时也让场景模型变的更易于维护，当基础数据更新后三维场景自动同步更新，而不是耗费大量人力物力重新制作。在数据层面上，所有的业务数据都可在空间实体上进行挂载与关联，使数字孪生场景成为业务分析管理的重要入口和载体，这与传统按图层进行分类和管理的 GIS 系统有本质上的不同。在大型水利枢纽管理业务场景中，实现了流域、城市、大坝从宏观到微观的空间实体构建，并与水文监测数据、城市洪涝数据、大坝梯调方案数据进行挂载关联，为水利孪生场景注入了数据的血液，让业务的脉动真实可触摸。



### 3. 技术架构设计：全域时空流域孪生底座

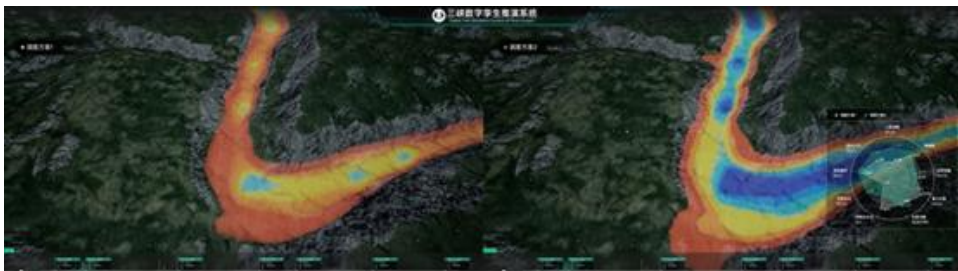
基于阿里云 DataV 数字孪生平台，完成了全域时空数据接入与海量时空数据融合，同时接入了行业仿真推演算法，构建出一套可计算的、具有生命力的数字孪生时空底座，支撑了上层全流域态势感知、水文监测预警、洪水模拟推演等丰富的业务场景。同时将地理信息数据、行业数据、经济社会数据在统一的时空底座上进行融合，使业务视角更全面、分析更精准、决策更科学。



### 4. 建设成果：实时水情推演仿真系统

每年的 5-7 月是长江流域的汛期，长江有 80% 的水量集中在汛期。由于汛期时水流量大，而长江下游地势平坦，很容易导致江水漫出河道，发生洪涝灾害。在过去几乎每隔 10 年左右，长江流域都会遭受一次较大的洪水侵袭。由于洪水规模大，而且长江中下游地区人口密度大，使得当地的经济因洪水遭受巨大损失。在一系列大型水利枢纽修建完成之后，可以拦截长江上游的洪水并减少出库水量，尽可能地为下游争取时间减少损失和影响。

长江流域涉及的河段长、模拟范围广、计算量大，为满足水利工程综合调度需求，洪水演进模型采用一维、二维水动力耦合模型，模型具有计算速度快、计算稳定性高、计算成果合理可信的特点。在完成水动力算法开发与接入之后，通过对水动力模型仿真的每一帧仿真数据进行（流向、流速、高程等维度）地理校准、重采样，并进行几何细节增强与优化，在科学仿真实现了“物理真实”的基础上，同步实现了“视觉真实”，使水情推演效果动态准确，实现了历史水情回放与实时水情推演仿真；将抽象的算法结果用视觉上可理解的方式进行呈现，实现了不同仿真结果地快速对比，未来将对长江流域的防洪调度起到重要作用。



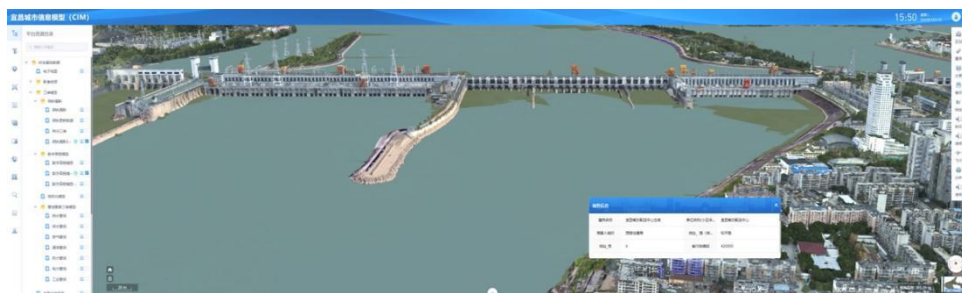
## 案例二：阿里云 - 宜昌城市大脑

## 1. 案例背景：

2021 年 5 月宜昌市市委召开城市大脑建设专题会议，提出要推动实施城市大脑建设，助力宜昌未来整体管理架构重心转移，宜昌治理手段、治理方式战略转换。按照“1234”的总体架构规划 1 个底座、2 套支撑和应用体系、3 项保障、4 类场景应用来分步骤有计划的进行城市大脑建设，其中数字底座是城市大脑的基础和核心，担负着数据资源承载、数据整合、业务协同、安全防护、运维管理等任务。

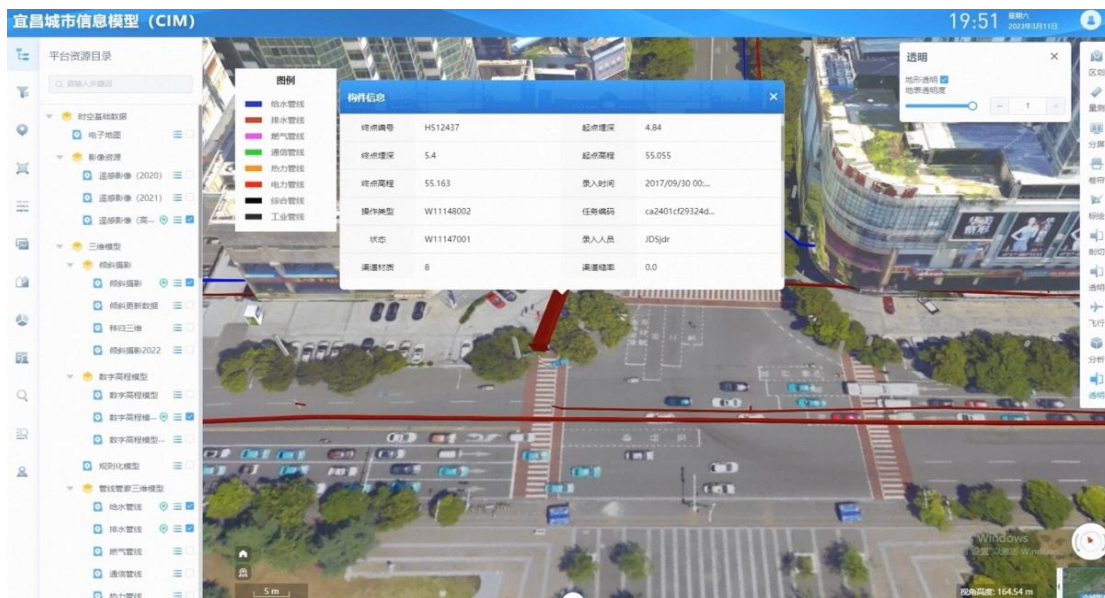
## 2. 地球大数据应用：构建城市数字基座，成为先进数字公共设施

2021 年 10 月，阿里云承接了宜昌城市大脑数字底座的建设，引入先进技术架构和服务能力，构建完善的运维、安全和质量控制体系，充分利旧现有设施，帮助宜昌打造一个具有宜昌特色的数字底座。



(1) 构建标准化、可度量的城市数据模型，支撑城市一人一企一档建设，打造城市数据中台：引入大数据计算引擎、空间计算引擎、城市数据模型、数据可视化分析等工具引擎，利用多种类型、多种技术手段增强数据归集能力；在“城市数据模型”的参考体系下，利用产品化的数据计算引擎，对全量数据进行治理与建模，建

成高起点、可度量、标准化的城市数据底盘；并通过多维智能数据标签，让更多一线的业务人员可以参与到数据探索中，全面增强数据治理能力，挖掘数据资产更大价值。



（2）打造城市信息模型 CIM 平台，充分发挥城市空间数据、物联感知数据、政务管理数据的互通共享能力：宜昌拥有丰富的空间地理数据，但存在数据来源冗杂、数据格式多样、数据类型复杂、数据空间坐标不统一等种种问题。通过建设城市信息模型 CIM 平台，以三维数据模型为基础，融合“GIS+BIM+IOT”技术，构建全市统一的城市实体空间数据支撑底座和智慧城市的空间操作系统，为宜昌城市大脑搭建服务于项目的“规、建、管、养、用、维”全生命周期的空间数据智能管理平台，实现的全时空数据融合、全周期生长记录、全要素规则贯通和全过程治理开放，逐步实现数字城市与现实世界的“数字孪生”。

### 3. 技术架构设计：



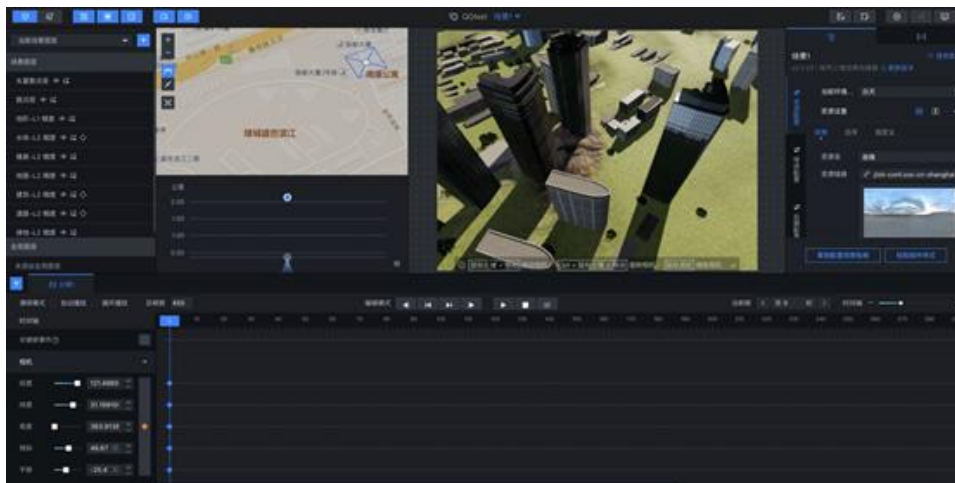
#### 4. 建设成果:

打造城市多业务、多维度的数字管理驾驶舱，实现城市管理精细化：宜昌数字管理驾驶舱基于城市信息模型平台的空间 GIS 数据，结合阿里云数字孪生仿真算法，实现 L3 级别的三维城市孪生模型的快速构建，为宜昌市打造了涵盖安全、交通、民生、文旅、政务五大模块的驾驶舱场景，为城市管理者提供精细化管理的数字驾驶舱，为城市管理提供有效的决策支撑。



帮助其他政府业务部门快速构建不同业务场景下的数据应用：宜昌数字底座的建设要实现一个底座建设，N 个应用场景的建设目标，希望通过一套统一的低代码开发平台，让不同的业务部门参与到各自部门业务小脑的建设中，所以宜昌数字底座建设通过 DataV 数据

可视化开发工具，打通数据中台与 CIM 平台的数据接入共享能力，实现让业务人员基于图形化的开发界面进行不同场景的业务应用开发，充分发挥数字底座的数据共享服务能力。



### 案例三：水利部水土保持监测中心 – 自然资源案例

人为水土流失监管既是水利部门的一项重要法定职责和社会管理职能，更是水土保持贯彻落实习近平生态文明思想、推动生态优先绿色发展的重要政治任务。通过常规的手段很难做到生产建设项目水土保持监管的全覆盖，大量违法违规行为得不到及时的制止和惩处，人为水土流失仍不能得到全面有效遏制，成为加快生态文明建设的重要制约因素。

水利部高度重视人为水土流失监管工作，把生产建设项目监管信息化应用作为重中之重，近几年在大力推进生产建设项目水土保持“天地一体化”监管示范和全国推广，总体目标是实现对在建生产建设项目信息化动态监管，以提高监管时效性和覆盖面，使监管更精准和到位。近年来，水利部出台了大量生产建设项目遥感监管政

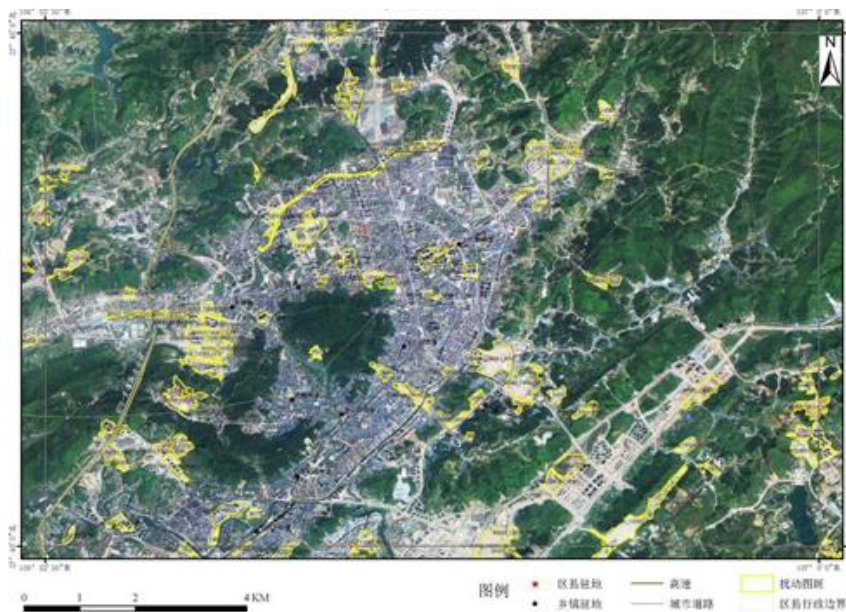
策文件推进水土保持信息化监管工作，流域机构和各级水土保持监管部门根据文件要求，以卫星遥感、无人机等先进技术为手段，先后开展了生产建设项目水土保持监管示范、部管项目“天地一体化”监管以及生产建设项目集中区动态监管等监管实践工作，信息化监管工作取得重要进展和显著成效。

在数据采集方面，通过收集当年资源 1、3，高分 1、2、6、7 号和备选资源卫星数据，以及历年录入全国水土保持信息管理系统的生产建设项目水土流失防治责任范围 shp 文件，水土保持方案特性、方案批复文件、省级遥感监管成果等数据，开展人为水土流失图斑新增、更新和属性完善，经过合规性判读和筛查，明确疑似违法违规扰动图斑。

在数据分析处理创新方法方面，利用协同解译方法对人为水土流失图斑进行解译。该方法的创新之处在于支持多角色用户远程协同在线解译，支持千人同时在线协同工作，支持遥感影像入库、历史扰动图斑入库、扰动图斑解译提取，图斑审核，扰动图斑解译成果导出，工作量统计等。

在获得创新性结果方面，通过人为水土流失遥感监管项目，发现人为水土流失违法违规行为，提升水土保持管理效能和水平。2015 年-2018 年全国选取试点进行示范工作，2019 年起，水利部正式启动实施水利部生产建设项目水土保持遥感解译与判别项目，四年共认定并查处“未批先建”“未批先弃”“超出防治责任范围”等违法违规项目 12.99 万个。

在场景应用价值方面，通过多年的工作实践，发挥了遥感监管“及时、全面、快速”的特点和优势，基本构建了一套“天上看，地面查、全覆盖”的人为水土流失监管模式，发现并查处了大量以前难以发现的违法违规行为，有效防治了生产建设项目可能造成的水土流失。



#### 案例四：易智瑞 – 城市更新案例

2019 年 12 月中央经济工作会议首次提出了“城市更新”的理念，并于 2021 年 3 月首次写入 2021 年政府工作报告和“十四五”规划文件《中华人民共和国国民经济和社会发展第十四个五年规划和 2035 年远景目标纲要》中，将其上升至国家战略层面，正式全面推开。“实施城市更新行动”的启动，代表我国城市更新已进入适应新时代要求、承载新内容、重视新传承、满足新需求的新阶段，标志着城市建设从“粗放式发展”进入“精细化运营”时代。

为了推动城市结构调整优化和品质提升，转变城市开发建设方式，某市开展了“实施城市更新行动”，要求建立“规划片区、用地地块、建筑物”等关键对象的开发区全要素时空变化城市更新信息时空数据库，并基于该数据库，围绕更新资源管理、更新谋划分析、拆征监管、安置管理等各类城市更新业务，建立二三维一体化的城市更新综合信息管理平台 and 城市更新掌上通，实现规划、拆征、招商、出让、建设、竣工、安置、维护等城市更新业务全过程数字化，助推城市更新管理更加精准高效，助力谋划城市有机更新的重点区域及其更新方向、目标、时序、总体规模和更新策略，强化推进城市有机更新空间与时间上的衔接和融合，为城市更新谋划、规划选址、重大项目落地、城市拆征安置项目建设和资金流向监管工作、城市安置房分配方案遴选提供长效助力，实现城市高质量发展，助推城市更新管理更加精准高效，城市综合承载能力不断提升，推动城市有机更新可持续发展。

## 1. 业务模型

通过大数据分析、人工智能、地理信息等先进技术，建立研判规则分析模型即城市更新智谋模型；包括按区域、按建筑年限、按用地类型、按用地亩均效益范围、按规划要求、按开发区空间布局、按用地规划、5分钟生活配套设施、15分钟生活圈配套设施、低效用地等不同条件，不断更新学习健全研判规则分析模型，形成一套可以配置的算法模型。根据“研判规则分析模型”，分析城市更新单元的优先级别，预测城市有机更新区域范围。并可结合空间地图，

二三维联动自定义区域、自定义空间位置范围，自主分析研判城市更新单元，为政府提供有力的科学依据。

## 2. 应用场景

（1）更新概况：某市提出加快实施城市更新“4321”行动，并制定了城市更新任务目标，在大屏端展示全市范围内城市更新的5年总体目标、年度任务目标、年度投资进度，直观呈现出城市的更新进度。开发区的城市更新计划、行动目标和具体项目的执行情况，实现从城市、开发区、项目层级的宏观到局部的更新概况简介。



（2）更新智管：构建基于时空影响因子的生活圈公共服务设施评价分析模型，用户可对开发区公共服务设施配置10分钟生活圈、15分钟生活圈服务情况开展分析，辅助开发区开展城市更新谋划，优化公共服务配置，提升社区空间品质。利用“研判规则分析模型”分析整个开发区内建议优先拆征区域单元，包括每个建议更新单元的地块范围、地块面积、区域价值、企业绩效、土地利用规划、房屋建筑等信息，辅助领导决策分析，发挥谋划价值，形成城市更新“看图作战”，让城市城市更新智能化、精细化、科学化。

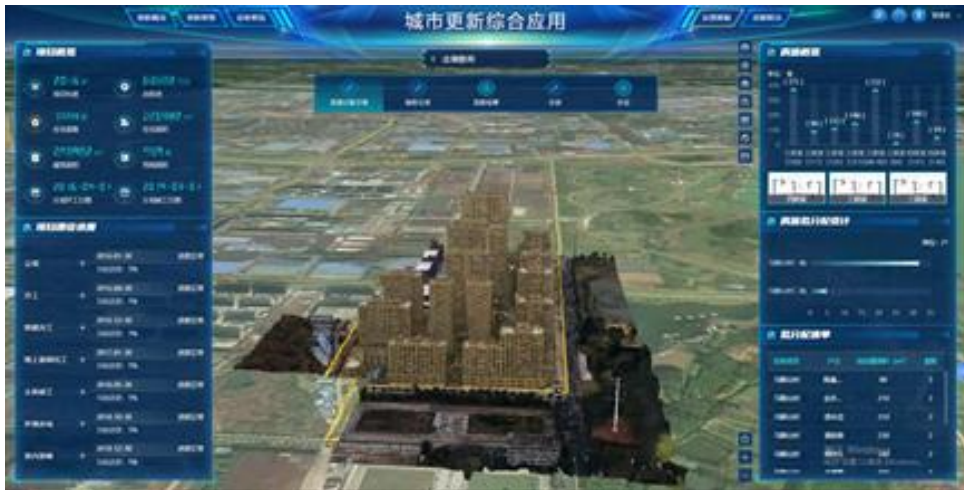


(3) 征收智监：通过大屏展示全域范围内所有拆征项目的分布、进度，统计分析每个区域、每个乡镇街道的项目开展情况、进度情况、总体拆征体量、已完成率、资金使用情况等各类信息，并分析每个项目中每个拆迁户安置滚屏展示每个项目具体概况，实现看图作战。



（4）安置智配：通过三维 GIS 平台在大屏端通过一张图展示每个项目的分布和实时进度，融合关联拆征户的安置情况等进行三维可视化分析，并将每户的房屋从拆征到安置，对各阶段的建设进度，形象进度、交付进展情况、每套安置房入住情况、剩余房源等，进行全面监管，实现对拆迁预计安置面积与实际安置面积进行比对，通过一张图实时让领导全面掌控开发区拆征安置整体过程和进展，辅助整个开发区的拆征安置资源调度和协调指挥。





（5）资源智治：利用 CIM、GIS（地理信息科学）、云计算、大数据、数据挖掘等先进技术，整合城市各类资源，并提供二三维地图联动展示，专题展示、行政区划展示招商、规划、用地、建设、拆征、安置等各类城市资源图层叠加，信息查询定位，详情展示等功能。



### 3. 应用价值

城市更新大数据三维可视化分析系统（大屏端）利用建筑物三维模型数据，以空间关系充分融合政务服务数据、项目数据、城市地理信息数据、城市谋划数据等，形成“城市更新大厅”，充分利用

强大的地理信息与三维场景还原能力、IoT 设备承载能力，以及超强级别的云运算能力，真实模拟一个和“现实世界”完全一致的孪生的“虚拟世界”。实现项目大数据实时化、可视化、空间化、智能化、便捷化管理，成为开发区对外宣传的“展示门户”，为相关部门工作提高效率，领导科学决策提供依据，形成城市更新“看图作战”，让城市城市更新智能化、精细化、科学化，以城市更新释放城市未来生命力。

## 案例五：英视睿达 - 大气环境热点网格精细化监管案例

英视睿达以地球大数据平台为基础，为生态环境部“千里眼”计划进行技术支持，项目通过利用卫星遥感、智能监测设备、大数据、人工智能等综合手段，以地球大数据平台为依托，将重点区域 39 个城市 42.8 万平方公里区域划分为 6 万个 3km × 3km 网格，从中筛选出 5000 个 PM2.5 浓度较高的热点网格进行“千米级”监控，进一步精准识别出 1000 个 500m × 500m 的小网格实施“百米级”监管。

项目自 2018 年开展以来，累计向京津冀及周边、汾渭平原 39 城市生态环境部门推送环境异常信息 8.5 万条，指导地方生态环境部门排查企业 8.4 万家，发现各类涉气环境问题 5.5 万个。项目提高了重点区域环境监管效能，颠覆了传统的“人海战术”，有效节约了国家在生态环境领域人力、物力、财力的投入，实现第一时间发现问题、解决问题，开启了环境精准监督和执法的新局面，取得了良好成效。

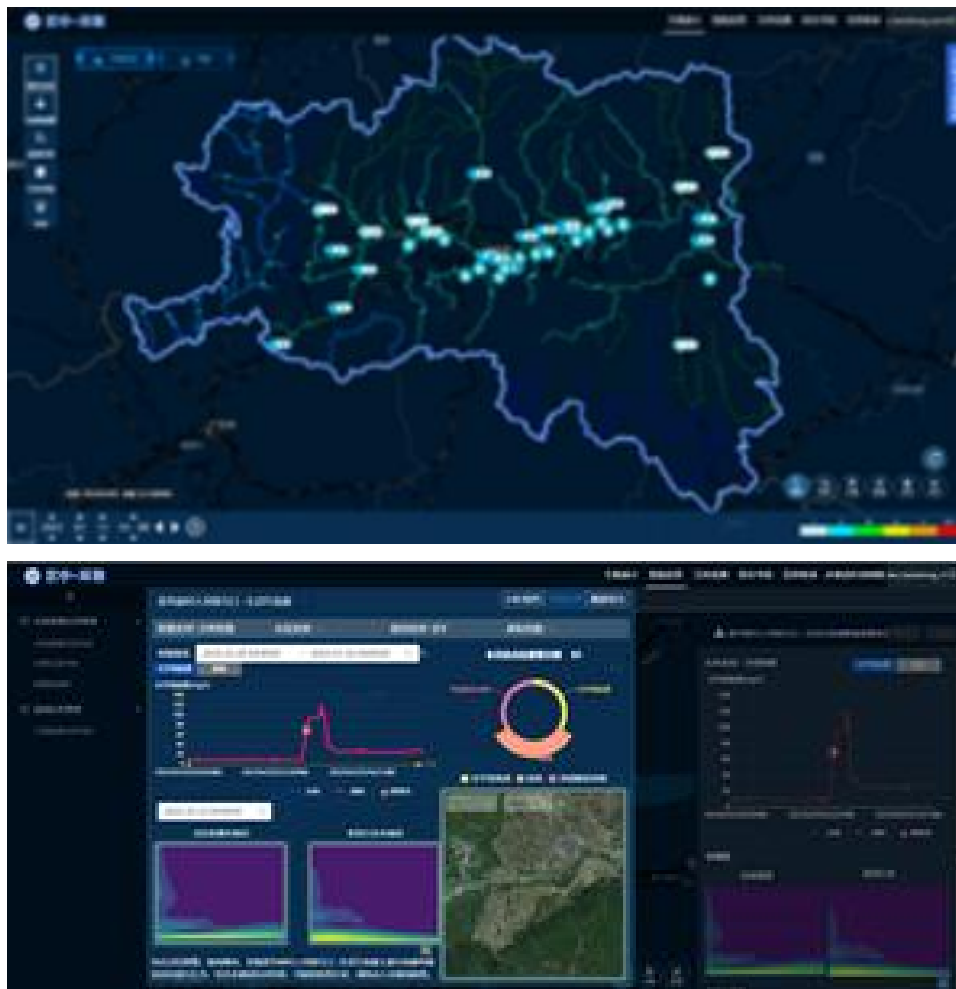


## 案例六：英视睿达 - 水基因精细化监管案例

2020 年期，为深入推进汉中市水污染防治，进一步改善汉江流域水环境质量，确保水质稳定达标，英视睿达为汉中市生态环境局搭建了水基因精细化监管系统，利用卫星遥感、高密度水基因水质监测微站等进行深度融合，对沿河污染物来源全面监控，自动识别跟踪、定位问题河段，锁定疑似排污口，使监管排查有方向，达到环境应急处置“一张图”管理，实时发现，及时处理。通过水质污染—设备报警—预警报告—现场巡查—结果反馈的闭环管理机制，根据水基因水质监测微站监测数据异常变化情况，充分发挥远程监控预警功能，通过非现场监管方式，迅速、精准锁定污染源，提高执法效能。为相关部门精准执法和汉中市水质持续向好提供技术保障。

本项目共布设 42 套固定式水基因水质监测微站，23 套用于区、县界水质考核断面监测，19 套用于入河排污口监测，实现了“科学化、精细化、常态化”环境管理模式的创新与转变，为汉中市水质持续向好提供决策和技术保。2022 年，汉中市 7 个国控站水质月均

为 I 类和 II 类水，黄金峡在 10 月为 III 类水质标准，汉江和嘉陵江干流汉中段水质总体较好，水环境状况均为优；64 个手工监测断面均优于或达到 III 类以上水质，优良水体占比达到 100%，汉江、嘉陵江出境水质稳定保持 II 类标准。汉中市水污染精准执法入选 2022 年生态环境部第十批生态环境执法典型案例



## 案例七：中数科技 - 数据安全应用案例

### 1. 项目背景与目标

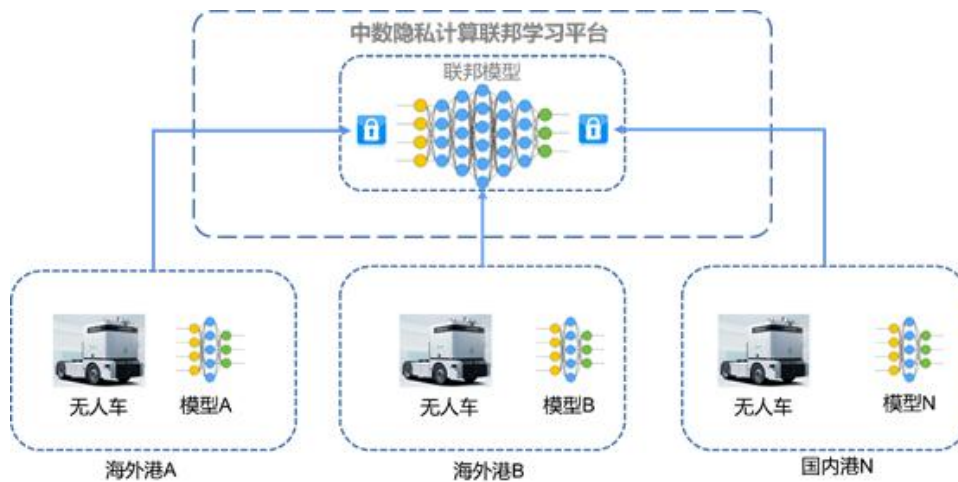
2020 年，在 5G 商用的关键节点上，我国车联网产业进入快速发展阶段：车联网上升到国家战略高度；产业政策持续利好；车联网

技术标准体系已经从国标层面完成顶层设计，C-V2X 技术标准可指导产业开发；我国已经建成较为完善的车联网产业链，在测试验证、应用示范方面形成一定规模，为后续大规模产业化和商业化奠定坚实基础。可以说，经过了具有里程碑意义的 2020 年，我国车联网产业正加速驶向商用的光明未来。

本项目目标是实现基于大数据及隐私计算技术的无人车充电与电池预维护，利用大数据、联邦学习、人工智能技术赋能智能无人车等场景，进一步研究探索车联网模式创新应用，实现无人车使用前、中、后对电池检测的线上化、自动化、智能化。主要实现的功能和解决的问题：一是搭建全在线的无人车互联体系，助推西井科技涉港无人车业务实现数字化全流程；二是构建基于车联网大数据的智能预测性维护模型，包括风险预测、风险告警；三是精准的车联网模型信息输出能力，包括风险数据采集、多维数据加工；四是车联网与电池维护管理体系设计，包括项目地块管理、硬件信息可视化、硬件损耗程度可视化、车辆工作效率可视化、区域业务分析、数据匹配分析、数据监测报告；五是优化数据中台，提供采集、计算、建模、服务的全链条数据服务，向后整合多源数据，实现全域数据的统一管理、融合共享，向前提供统一的数据服务以及可视化交互、数据分析等功能，推动业务数据化向数据业务化发展。

## 2. 创新点

采用多方数据训练模型，预测速度更快，精度更准。联邦共同建模需要不同的机构共同参与，因为联邦建模的模型效果会在实际应用中表现出来，并记录在永久数据记录机制（如区块链）上。提供数据多的机构所获得的模型效果会越好，具体取决于数据提供方对自己和他人的贡献。



辅助决策：训练各大港口收集的模型数据，将训练模型上传到联邦智行平台，管理者可以在每个港口之间数据不互通的情况下享受模型训练带来的效益。同时联邦智行平台会通过训练好的模型对管理者进行辅助决策，帮助其解决较为困难的现实问题。

### 3. 应用价值

为智慧港口的进一步智能化赋能：以隐私计算技术为基础服务行业创新。成熟经验逐步向同行业输出：实现隐私计算技术与新能源、无人车和现代智慧港口等平台的有机衔接。

### 4. 项目总结

本项目结合隐私计算技术的特点，坚持数据隐私保护的方针，运用前沿数字化手段搭建了成熟的隐私计算—车联网产业体系，在智慧港口领域首创以隐私计算技术、AI 技术、无人驾驶技术等主体技术为核心的技术闭环，并且针对智慧港口、无人车与自动驾驶客户群体的特点，形成了一套完整的业务流程，填补了行业空白，理念先进、设计合理、功能丰富。

## 案例八：元能星泰 – 孪生流域应用案例

流域物理空间实体既包括自然地貌、植被和水系，还包括水库、堤防、闸坝、泵站等水利工程体系，数据对象种类繁多、关系复杂、系统耦联，且数据精度、更新频次明显不足。

面对流域“四预”管理所面临的数据挑战，需要开拓更有效的方式。地球大数据具有宏观、动态、客观监测能力，可对包括陆地、海洋、大气及与人类活动相关的数据进行整合和分析，具有高瞬时性、任意空间性、物理相关性等特点，可以把大范围流域作为整体进行认知，为流域“四预”管理完善数据类型、数据范围和数据质量，对包括流域、水利工程等水利对象的主要属性数据和地理空间数据进行大尺度、周期变化的监测和评估，从而助力解决水利信息资源缺失、整合不足、数据鸿沟以及共享困难等问题，为“四预”管理提供算据支撑。本文聚焦流域防洪“四预”应用，以福建闽江流域为例打造基于地球大数据的数字孪生流域“四预”管理平台。

数据是数字孪生流域的基础，可以看作是一个集数据采集、通信、融合、存储与组织于一体的信息平台，需满足数字孪生流域全覆盖、高频次、高精度的数据需求，其可划分为数据感知层与数据交互层。其中数据感知层主要是与水利业务信息监测采集相关的全域感知设施，从业务方面可以划分为江河水系数据、水利工程数据、地下水数据、水质监测数据等，除了传统流域信息采集基础设施，还包括卫星遥感、无人机、无人船等新型信息获取方式，旨在实现流域数据信息实时全覆盖，获取信息的精度与频次能满足虚拟孪生世界与真实流域实时同步运行的需求。数据交互层涵盖网络连接设施与通信计算设施，主要是对采集得到的数据进行通信传输、计算存储、数据汇集，最后建立多维多时空数据模型对海量数据进行有机组织。

在该试点应用中，MetaStar 元能星泰以闽江流域为单元、时空数据为底座、数学模型为核心，对物理流域全要素的数字化映射，使用云渲染引擎加载上述高精度的数字底板数据构建的三维场景，通过数字孪生体与水动力模型、调度模型的深度融合，运用参数化驱动水粒子物理仿真系统拟合具体防洪业务，通过引擎 API 和场景模型完成物理驱动、实时渲染、动态视觉特效等进阶功能，实现洪水预报预演的流域防洪“四预”应用：

## 1. 洪涝灾害要素预报

紧扣“降雨-产流-汇流-演进”预报环节，基于降水、水位、潮位、流量等实现洪峰水位、洪峰流量、洪量、径流量、潮位、淹没面积、淹没水深、检查井溢流情景预报等，建立参数在线率定机制，

实时滚动分析已发生洪水的预报规律，提高洪水预报精度，为洪涝灾害风险预警提供基础。

## 2. 洪涝灾害风险预警

包括明确预警任务、建立预警发布机制、加强预警信息发布行业管理、构建预警汇集发布平台、打通预警发布最后一公里”等。针对河道洪水、山洪灾害、内涝灾害等洪水灾害风险事件，结合监测条件、防护对象、影响范围等，确定以水位、流量、淹没范围及淹没深度等为警示指标，实现闽江洪水预警、中小河流洪水早期预警等。

## 3. 洪涝灾害仿真预演

基于闽江流域数字化场景、构建预演模型、开发预报调度一体化系统、建设预演平台等。运用数字化、网格化、智能化手段、实现水工程预报信息与调度运行信息及边界条件的集成耦合，根据水雨情预报情况，对水库、河道、蓄滞洪区及洪水演进情况进行模拟预演，设定全面的防洪情景目标，及时滚动更新预演方案，快速分析展示风险形势，为洪涝防御提供科学决策支持。

## 4. 洪涝灾害调度预案

针对构建的洪水预演模型，进行不同场景下流域预报、水库联合调度、河道洪水演进、蓄滞洪区淹没方案计算，进行多方案的综合比选，建立方案评估指标体系，进行方案优劣的评选，最终形成用于指导流域实际防洪调度的典型预案，并形成预案知识库。

地球大数据支持下的闽江流域数字孪生防洪“四预”平台，实现数字流域和物理流域的虚实交互，通过智能干预实现流域防洪调度的智能化决策，全面科学地提升了闽江流域的智慧防洪调度水平，为智慧水利建设探索了一条可行的技术和方法，后续还可结合实际需求进行持续扩展和升级完善。



## 案例九：数梦工场 – 智慧应急案例

数梦工场服务应急管理部的数据治理系统，数据资源池约 600TB 的数据资产，涵盖部内有关司局、部属单位、地方省厅、外部委、社会单位共 1700 余类数据资源，6 万余个数据项，共计约 9095 亿条数据，391.2 万个非结构化文件。经过治理融合加工，形成管理对象、应急力量、灾害事故、应急资源等 4 大应急主题库，监测预警、监督管理、决策支持、指挥救援、社会动员、政务管理等 6 大专题库。

### 1. 方法-结果

数梦工场在应急管理部数据治理工作中成效显著。数据智能探查能力实现单表处理时长由过去的 4 小时减少到 1.2 小时，数据处理

效率提升了 3 倍。全栈多租户的服务能力，为森林消防局、地震局等 7 机关司局和部属单位开通租户，租户完成超过 400 个数据专题开发。多云跨云的云原生能力支持北京中心（华为云）和贵州分中心（阿里云）统一管理和数据备份，整体数据服务的性能利用云原生技术线性提升。

数梦工场解决了困扰客户已久的目录级联、标准级联和服务级联等关键技术难题。率先在纵向上实现应急管理部、省、市、县四级资源目录级联，实现跨部门、跨层级、跨系统、跨区域的数据资源共享。目前持续为 310 个业务应用提供数据服务累计超 30 亿次，构建 PB 级数据资源池，建成覆盖全国 31 个省级行政区和新疆生产建设兵团、47 个地市、上百个区县一体化信息资源体系。通过多级级联，市、县应急管理部门的数据资源获取周期从 1 天缩减到 1.5 小时、资源共享效率提升 92%、资源归集时间缩短 80%，数据质量提升了 3 倍。此外该平台支撑全国危化品安全生产监管系统，每天超 10 亿条安全生产感知数据实时汇聚至应急部指挥中心，同时回流预警信息到省市监管部门，使全国危化品高危事故率降低 30%以上。



## 2. 应用价值

数梦工场服务应急管理部的数据治理系统，初步形成了行业级一体化、智能化数据治理机制和数据管理能力，在四川泸定 95 地震、昆明 COP15 大会等灾害处置和活动保障工作中发挥了核心作用。在 2022 年全国数字防汛演练工作中，实现三峡库区和长江中游河段各城市之间气象、水情、灾情、应急资源信息的实时共享，全面实现跨部门、跨层级、跨区域的监测预警和应急指挥的业务协同。



**大数据技术标准推进委员会**

**地址：** 北京市海淀区花园北路 52 号

**邮编：** 100191

**邮箱：** TC601@CCSA.org.cn

**网址：** [www.tc601.com](http://www.tc601.com)

