



中华人民共和国国家标准

GB/T XXXX—XXXX

信息技术 生态环境大数据 系统框架

Information technology—Big data for ecological environment—System framework

(征求意见稿)

在提交反馈意见时，请将您知道的相关专利连同支持性文件一并附上。

XXXX - XX - XX 发布

XXXX - XX - XX 实施

国家市场监督管理总局
国家标准化管理委员会 发布

目 次

前 言.....	III
1 范围.....	1
2 规范性引用文件.....	1
3 术语和定义.....	1
4 生态环境大数据参与主体.....	2
4.1 概述.....	2
4.2 生态环境大数据系统协调者.....	2
4.3 生态环境数据提供者.....	2
4.4 生态环境大数据应用提供者.....	2
4.5 生态环境大数据计算框架提供者.....	3
4.6 生态环境大数据使用者.....	3
5 生态环境大数据系统框架.....	3
5.1 概述.....	3
5.2 计算框架层.....	3
5.3 数据资源层.....	4
5.4 应用支撑层.....	5
5.5 数据应用层.....	5
5.6 配套保障体系.....	6
参 考 文 献.....	7

前 言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

本文件由全国信息技术标准化技术委员会（SAC/TC28）提出并归口。

本文件起草单位：清华大学、生态环境部信息中心、中国电子技术标准化研究院、重庆市生态环境监测中心、生态环境部环境规划院、绿巢智慧数据科技（北京）有限公司、中移系统集成有限公司、福建省生态环境信息中心、广东柯内特环境科技有限公司、北京思路创新科技有限公司、安吉两山转化数字研究院、北京嘉诚瑞杰信息技术有限公司、北京东方金信科技股份有限公司、北京市大兴区生态环境局、福建大数据一级开发有限公司、杭州数梦工场科技有限公司、北京工业大学、山东省生态环境监测中心、深圳市生态环境智能管控中心、九江市生态环境局、中电万维信息技术有限责任公司、中科三清科技有限公司、江苏擎天工业互联网有限公司、北京百分点信息科技有限公司、智慧神州（北京）科技有限公司。

本文件主要起草人：

信息技术 生态环境大数据 系统框架

1 范围

本文件确立了生态环境大数据的系统框架，明确了生态环境大数据参与主体，规定了生态环境大数据系统框架的功能组成等。

本文件适用于为各级政府部门、企事业单位、社会团体等开展生态环境大数据系统的规划设计、建设及其升级、运维运营、安全保障等工作提供指导，并且为生态环境大数据系列标准制定提供基础。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 35589-2017 信息技术 大数据 技术参考模型

3 术语和定义

下列术语和定义适用于本文件。

3.1

生态环境大数据 big data for ecological environment

在生态环境治理和生态修复等过程中产生和应用的数据集合。

[来源：GB/T AAAAA—202X，3.1]

3.2

生态环境大数据系统 system of big data for ecological environment

对生态环境大数据进行采集、处理、存储和应用的基础设施、技术和应用的综合系统。

3.3

生态环境大数据系统协调者 system coordinator of ecological environment big data

是指用于综合类的政府部门或生态环境部门内的综合类的组织机构定义并明确所需的数据应用活动并将其整合到可运行的系统中。

3.4

生态环境数据提供者 data provider of ecological environment

指生态环境部门、其他相关政府部门、企业、组织机构和公众等，通过各种采集方式和渠道，将生态环境数据和信息引入到生态环境大数据系统中。

3.5

生态环境大数据应用提供者 application provider of big data for ecological environment

指从事生态环境大数据系统建设和服务的企业、组织机构等，通过执行生态环境数据生命周期操作，以满足数据使用者的应用需求以及系统协调者定义的安全和隐私保护需求。

3.6

生态环境大数据框架提供者 framework provider of big data for ecological environment

指综合类的政府部门或生态环境部门内的综合类的组织机构，通过执行应用转换，为生态环境大数据应用创建提供资源和服务，同时保护数据完整性和隐私。

3.7

生态环境大数据使用者 data user of big data for ecological environment

指生态环境部门、其他相关政府部门、企业、组织机构和公众等，通过调用生态环境大数据应用提供的接口按需访问信息，并与其产生可视的、事后可查的交互。

4 生态环境大数据参与主体

4.1 概述

生态环境大数据参与主体包括生态环境大数据系统协调者、生态环境数据提供者、生态环境大数据应用提供者、生态环境大数据计算框架提供者、生态环境大数据使用者等。

4.2 生态环境大数据系统协调者

生态环境大数据系统协调者主要功能是规范和集成各类所需的生态环境数据应用活动，以构建一个可运行的系统。具体工作包括以下几个方面：

- a) 收集并分析生态环境大数据使用者提出的需求；
- b) 组织并协调生态环境数据提供者按需提供数据；
- c) 组织、协调并监督管理生态环境大数据应用提供者按照应用需求、技术要求、安全和隐私要求、建设运营相关规范、管理体制机制等开展应用系统平台建设和运维；
- d) 配置和管理生态环境大数据计算框架提供者中各组件资源，为其他组件分配对应的物理或虚拟节点，对各组件的运行情况进行监控，以保障数据提供、应用提供和数据使用工作的运行；
- e) 建立和完善管理机制体制、制定标准保障其他各实体稳定高效运行。

生态环境大数据系统协调者通常是综合类的政府部门或生态环境部门内的综合类的组织机构。

4.3 生态环境数据提供者

生态环境数据提供者的主要功能是为生态环境大数据应用提供者提供数据源，通过各种采集方式和渠道，将数据和信息引入到生态环境大数据系统中，供生态环境大数据系统发现、访问和转换。

生态环境数据提供者包括生态环境部门、其他相关政府部门、企业、组织机构和公众。其中，其他相关政府部门包括除生态环境部门外承担生态环境保护职责的其他政府部门、提供生态环境相关数据的相关政府部门，以及各部门的派出机构和直属单位；企业包括生态环境监管对象和提供生态环境咨询、治理服务的企业；组织机构包括科研机构和其他组织机构。

生态环境数据提供者通过以下渠道开展数据采集活动：

- a) 生态环境业务管理活动：是指生态环境部门和承担生态环境保护职责的政府部门及各部门的派出机构和直属单位在履行业务职能过程中，通过生态环境业务系统或本地文件等形式采集生态环境相关业务数据的过程；
- b) 相关政府部门共享：是指相关政府部门履行部门职责过程中产生与生态环境管理需求密切相关的数据，并通过共享交换提供数据。例如气象数据、水文数据、社会经济数据等；
- c) 物联网监测：是指承担生态环境保护职责的政府部门和企业利用各种传感物联网设备，采集的以时间序列为主的生态环境实时监测数据，包括大气、水、声、辐射、土壤等环境数据，以及污染排放在线监控、视频实时监控数据等；
- d) 互联网采集：是利用互联网数据采集技术，对生态环境领域相关的数据或信息进行动态采集；
- e) 其他：包括各种组织机构开展科研项目、第三方服务等产生的数据等。

4.4 生态环境大数据应用提供者

生态环境大数据应用提供者主要围绕生态环境大数据使用者需求，在系统协调者的组织下，将来自数据提供者的数据进行处理与分析，提供给数据使用者，同时需满足系统协调者规定的技术要求、安全性和隐私要求等。

生态环境大数据应用提供者包含从事生态环境大数据系统建设和服务的企业、组织机构等。

4.5 生态环境大数据计算框架提供者

生态环境大数据计算框架提供者主要为生态环境大数据应用提供者在创建具体应用时提供使用的资源和服务。

生态环境大数据计算框架提供者通常是综合类的政府部门或生态环境部门内的综合类的组织机构。

4.6 生态环境大数据使用者

生态环境大数据使用者通过调用生态环境大数据应用提供的接口按需访问信息，与其产生可视的、事后可查的交互。

生态环境大数据使用者包括生态环境部门、其他相关政府部门、企业、组织机构和公众。其中，其他相关政府部门包括承担生态环境保护职责的其他政府部门、需要生态环境相关数据的其他政府部门，以及各部门的派出机构和直属单位；企业包括生态环境监管对象以及生态环境数据的社会化利用主体；组织机构主要包括科研机构和其他组织机构等。

5 生态环境大数据系统框架

5.1 概述

生态环境大数据的系统框架包括计算框架层、数据资源层、应用支撑层、数据应用层，以及支撑系统运行的安全管理、运维管理等配套保障体系（见图1）。

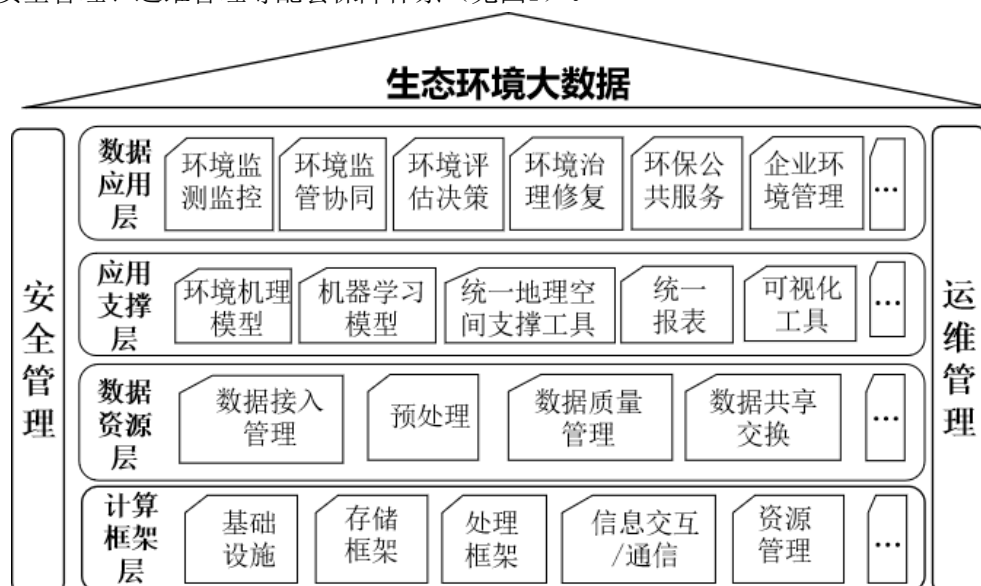


图1 生态环境大数据系统框架

5.2 计算框架层

5.2.1 概述

计算框架层包括基础设施、存储框架、处理框架、信息交互/通信和资源管理等。

5.2.2 基础设施

基础设施为生态环境大数据系统中的所有其他要素提供必要的资源，这些资源是由一些物理资源的组合构成，可控制/支持相似的虚拟资源。基础设施分为以下四个方面的内容：

- a) 网络：从一个资源向另一个资源传输数据的资源；
- b) 计算：用于支持系统运行和处理数据的计算能力和执行环境，包括图形处理器、中央处理器、内存等；
- c) 存储：用于为系统提供保存数据的资源；
- d) 环境：承载系统需要用的外部物理资源，包括房屋、门禁、机柜、温湿度控制、电力、消防等。

5.2.3 存储框架

存储框架是在基础设施之上构建的软件层，提供数据存储、计算和管理的功能。这些平台通常包括分布式文件系统、数据库管理系统、数据集成工具等，用于管理和操作大规模的数据。

5.2.4 处理框架

处理框架是指用于进行大数据处理和分析的软件工具集合，能够并行处理大量数据，并提供分布式计算、数据挖掘、机器学习和实时数据处理等功能。

5.2.5 信息交互/通信

信息交互/通信是指用于支撑生态环境数据在各个组件之间进行的交互和通信，包括数据的传输、共享和交换等操作，通过网络或其他通信手段实现，通信方式可以是点对点传输、存储转发或多播等。

5.2.6 资源管理

资源管理用于有效地管理和分配基础设施的资源，优化资源的利用和性能，以保障任务能够按时运行，包括调度器、容器化技术、负载平衡等。

5.3 数据资源层

5.3.1 概述

数据资源层主要是指构建生态环境数据资源中心，通过对需要面向多主体、多应用的生态环境数据进行统一收集、预处理、分析和存储，形成数据集，并提供数据产品和服务共享的基础数据平台。

包括数据接入管理、预处理、数据质量管理、数据共享交换平台等内容。

5.3.2 数据接入管理

数据接入管理负责处理与生态环境数据提供者的接口和生态环境数据的引入。

生态环境数据来源、数据类型、数据结构多样且复杂，可采用数据接口、数据前置库、数据包等形式进行数据导入，并根据技术的发展，迭代升级数据采集技术和方式。

5.3.3 预处理

数据治理是根据生态环境大数据应用需求，对获取的生态环境原始数据进行处理操作，以提供正确、规范及有价值的信息，包括数据验证、清洗、标准化、格式化和存储等活动。

5.3.4 数据质量管理

数据质量管理负责对从数据接入到分析应用各个阶段的数据质量的评价与控制。包括制定数据管理目标、制定数据质量规则、开展数据监控、数据校验、数据质量评价、数据问题报送和反馈闭环管理等。

5.3.6 数据共享交换

数据共享交换是根据数据使用者需求提供数据共享服务，包括目录服务、数据访问服务、数据交换服务和离线数据包服务等。

5.4 应用支撑层

应用支撑层主要是指生态环境专业支撑工具，从生态环境应用系统中抽象出来的具有通用性支撑作用的专业工具和组件。主要包括环境机理模型、机器学习模型、统一地理空间支撑工具、统一报表和可视化工具等。

5.5 数据应用层

应用层主要包括基于数据层汇聚的数据经过加工处理后，按照生态环境大数据使用者开展水环境管理、大气环境管理、应对气候变化和温室气体排放管理、土壤环境管理、固体废物及化学品管理、自然生态保护、核与辐射安全监管、污染排放管控等各项业务的管理和决策需求以及面向社会的公共服务需要，形成的各类应用系统。用于支撑环境监测监控、环境监管协同、环境评估决策、环境治理修复管理、环保公共服务、企业环境管理以及其它相关活动等应用场景。具体包括：

- a) 环境监测监控：政府部门、环境监管在对生态环境基本情况和污染物排放情况进行数据采集的基础上，进一步对数据进行管理和分析，为生态环境管理、决策和污染排放控制提供基础数据的应用系统。按照监测监控对象不同，分为生态环境质量监测和污染源监控场景，生态环境质量监测场景包括环境质量监测和生态系统状况监测场景，污染源监控场景包括监督性监测、在线监测、执法监测、应急监测等场景；
- b) 环境监管协同：生态环境部门和其他相关政府部门使用数据开展生态环境监管、环境违法行为监管、跨部门跨区域联合执法、突发环境事件应急等。按照监督管理范围的不同，将监管协同分为内部协同和外部协同。其中，内部协同包括部门和区域内的协同管理，主要为常规场景下的监管业务（信息跟踪、目标管理、任务调度等）和特定场景下的监管业务（监察执法、应急管理）；外部协同包括跨部门和跨区域的行政审批信息互通、监测数据共享、执法协同、应急调度协同等场景；
- c) 环境评估决策：生态环境部门和其他相关政府部门根据履行业务职能的需要，使用数据开展生态环境质量评价、环境污染溯源和预警预报、科学决策等活动，以及为相关活动而进行的数据共享。
- d) 环境治理修复：政府部门或参与环境治理的企业对环境污染问题进行诊断、评估评价，制定治理修复方案并采取相关措施，对治理修复结果进行评估反馈；
- e) 环保公共服务：政府对其他生态环境保护的参与主体（如企业、公众、科研机构等）提供专题数据以及模型算法、资讯、宣传教育、咨询、生态环境权益市场管理等服务；
- f) 企业环境管理：生态环境监管对象企业进行的数据的上传上报、利用相关数据进行生态环境保护和完善相关生产、商业活动；
- g) 环保科技研发：包括组织机构开展生态环境数据再生产再加工、科学研究、公益活动等，以及生态环境数据社会化利用主体的企业，对开放数据进行分析挖掘，提供相关的社会化服务等；
- h) 其他相关活动。

5.6 配套保障体系

5.6.1 安全管理

安全管理是指通过不同的技术手段和安全管理措施，构建生态环境大数据安全防护体系，实现覆盖硬件、软件和上层应用的安全保护，从网络安全、主机安全、应用安全、数据安全等方面来保障生态环境大数据的安全性。

按GB/T 35589-2017中7.6，安全管理主要包括以下四个方面的内容：

- a) 网络安全：通过网络安全技术，保障数据处理、存储安全和维护正常运行；
- b) 主机安全：通过对集群内节点的操作系统安全加固等手段保障节点正常运行；
- c) 应用安全：具有身份鉴别和认证、用户和权限管理、数据库加固、用户口令管理、审计控制等安全措施，实施合法用户合理访问资源的安全策略；
- d) 数据安全：从集群容灾、备份、数据分级、数据完整性、数据分角色存储、数据访问控制等方面保障用户数据的安全。

同时提供一个合理的灾备框架，提升灾备恢复能力，实现数据的实时异地容灾功能，跨数据中心数据备份。

提供隐私保护，在不暴露用户敏感信息的前提下进行有效的数据挖掘，根据需要保护的内容不同，可分为位置隐私保护、标识符匿名保护和连接关系匿名保护等。

5.6.2 运维管理

运维管理是指通过运维管理制度、运维管理措施、运维服务团队等，构建统一的运维管理体系，保障生态环境大数据系统高效稳定运行。

按GB/T 35589-2017中7.7，主要包括以下内容：

为生态环境大数据提供统一的运维管理系统，实现对包括数据中心、基础硬件、平台软件和应用软件等的集中运维、统一管理，实现安装部署、参数配置、监控、告警、用户管理、权限管理、审计、服务管理、问题定位、升级和修复等功能。

具有自动化运维的能力，通过对基础层的基础硬件、平台软件等资源进行统一管理，合理分配和调度业务所需要的资源，做到自动化按需分配；同时提供对数据层、支撑层和应用层的多个应用软件进行集中运维的能力，自动化监控生态环境大数据系统各功能构件的运行状态、性能指标等，对异常情况进行告警和事件处置，实现从业务维度来进行运维的能力。

对主管理系统节点及所有业务组件中心管理节点实现高可靠性的双机机制，采用主备或负荷分担配置，避免单点故障场景对系统可靠性的影响。

参 考 文 献

- [1] GB/T 35589-2017 信息技术 大数据 技术参考模型
 - [2] GB/T AAAAA—202X 信息技术 生态环境大数据 数据分类与代码
-