



中华人民共和国国家标准

GB/T 20533—2026

代替 GB/T 20533—2006

生态科学数据元数据

Metadata for ecological scientific data

2026-02-27 发布

2026-06-01 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

引言 V

1 范围 1

2 规范性引用文件 1

3 术语和定义 1

4 缩略语 3

5 元数据描述方法 3

6 一致性要求 4

7 元数据要求 4

 7.1 元数据应用对象范围 4

 7.2 元数据包和依赖关系 4

 7.3 元数据包的类图 5

 7.4 元数据基础包 5

 7.5 元数据依赖包 11

 7.6 生态科学数据核心元数据 14

 7.7 制定元数据专用标准的规则和方法 14

 7.8 本文件定义的元数据与科技资源核心元数据的关系 14

 7.9 元数据应用示例 14

 7.10 元数据实施方法 14

附录 A（规范性） 数据字典 15

 A.1 元数据基础包 15

 A.2 元数据依赖包 49

 A.3 代码表和枚举 60

附录 B（规范性） 元数据扩展与元数据专用标准 71

 B.1 背景 71

 B.2 元数据裁剪 71

 B.3 元数据扩展的类型 71

 B.4 元数据扩展原则与方法 71

 B.5 元数据专用标准 71

 B.6 元数据专用标准制定方法与规则 72

附录 C（规范性） 抽象测试套件 73

 C.1 简介 73

 C.2 现有元数据测试套件 73

C.3 用户自定义的扩展元数据测试套件..... 74

C.4 元数据专用标准测试套件..... 75

附录 D（规范性） 核心元数据 76

附录 E（资料性） 生态科学数据元数据与科技资源核心元数据对照表 77

附录 F（资料性） 元数据示例 79

 F.1 概述..... 79

 F.2 中国干旱半干旱区典型草地和荒漠生态系统植被地上生物量数据集元数据..... 79

附录 G（资料性） 元数据实施 93

参考文献 94

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。

本文件替代 GB/T 20533—2006《生态科学数据元数据》，与 GB/T 20533—2006 相比，除结构调整和编辑性改动外，主要技术变化如下：

- a) 增加了“生态科学数据”的术语的定义(见 3.1)，删除了术语“列表型数据”“栅格数据”“矢量数据”“元数据注册系统”“编码体系”“词汇编码体系”“句法编码体系”“数据类型”“值域”“概念模式”“概念模型”“论域”“模型”“数据模型”和“数据交换中心”(见 2006 年版的 3.1、3.2、3.3、3.12、3.14、3.15、3.16、3.18、3.19、3.20、3.21、3.22、3.24、3.25、3.26)，术语“元数据应用专规”更改为“元数据专用标准”(见 3.10，2006 年版的 3.13)；
- b) 删除了对“UML 模型符号”“UML 模型关系”和“UML 模型构造型”的说明(见 2006 年版的 4.2、4.3、4.4)；
- c) 更改了对元数据描述方法的说明，删除了 XML Schema 语言形式的元数据描述方法，将对 UML 模型图形式和数据字典形式元数据描述方法的说明更改为引用 GB/T 30522—2014 相关章节的内容(见第 5 章，2006 年版的第 6 章)；
- d) 更改了组成生态科学数据元数据的包及包之间的依赖关系(见 7.2，2006 年版的 7.2)；
- e) 更改了元数据信息包(见 7.4.1，2006 年版的 8.1.1)，更改了标识信息包(见 7.4.2，2006 年版的 8.1.2)，更改了数据实体信息包(见 7.4.3，2006 年版的 8.1.9)，更改了方法信息包(见 7.4.4，2006 年版的 8.1.4)，更改了数据质量信息包(见 7.4.5，2006 年版的 8.1.3)，增加了共享服务信息包(见 7.4.6)，增加了标注信息包(见 7.4.7)，删除了场地信息包(见 2006 年版的 8.1.5)、项目信息包(见 2006 年版的 8.1.6)、分发信息包(见 2006 年版的 8.1.7)、元数据参考信息包(见 2006 年版的 8.1.8)、空间参照系信息包(见 2006 年版的 8.1.10)和空间表示信息包(见 2006 年版的 8.1.11)；
- f) 增加了标识符信息包(见 7.5.1)，更改了责任者信息包(见 7.5.3，2006 年版的 8.2.1)，更改了覆盖范围信息包(见 7.5.4，2006 年版的 8.2.2)，增加了补充文件信息包(见 7.5.5)，删除了数据量信息包(见 2006 年版的 8.2.4)和文件格式信息包(见 2006 年版的 8.2.5)；
- g) 更改了元数据基础包 UML 模型图对应的数据字典(见 A.1，2006 年版的 A.1)；
- h) 更改了元数据依赖包 UML 模型图对应的数据字典(见 A.2，2006 年版的 A.2)；
- i) 更改了代码表中代码的编码方式(见 A.3，2006 年版的 A.3)，更改了“数据实体类型代码”“职责代码”“关联类型代码”的代码组成(见 A.3.4、A.3.6、A.3.9，2006 年版的 A.3.2、A.3.5、A.3.12)，增加了“知识产权类别代码”“关联资源类型代码”“栅格类型代码”“色彩代码”“数据采集方法分类代码”“观测方式代码”“观测位置代码”“场地类型代码”“场地形状代码”“样品分类代码”“共享范围代码”“共享途径代码”“标识符体系代码”等 13 个代码表(见 A.3.5、A.3.8、A.3.13、A.3.14、A.3.15、A.3.16、A.3.17、A.3.18、A.3.19、A.3.20、A.3.22、A.3.23、A.3.24)；删除了“密级代码”“资金来源代码”“介质名称代码”“关系基数代码”“波段组织方式代码”“字节顺序代码”“维名称类型代码”“像元定位代码”“拓扑等级代码”“成像条件代码”“学科分类代码”“研究对象代码”等 12 个代码表(见 2006 年版的 A.3.7、A.3.10、A.3.11、A.3.13、A.3.15、A.3.16、A.3.17、A.3.20、A.3.21、A.3.23、A.3.24、A.3.25)；
- j) 更改了对元数据扩展与元数据专用标准制定方法和规则的表述(见附录 B，2006 年版的附

录 B);

k) 更改了生态科学数据核心元数据的组成(见附录 D,2006 年版的附录 E);

l) 删除了数据集全集元数据应用专规(见 2006 年版的附录 D)。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由中华人民共和国科学技术部提出。

本文件由全国科技平台标准化技术委员会(SAC/TC 486)归口。

本文件起草单位:中国科学院地理科学与资源研究所、国家科技基础条件平台中心、南京大学、中国科学院计算机网络信息中心、中国标准化研究院、中国科学院植物研究所、中国科学院南京土壤研究所、中国科学院大气物理研究所、天津师范大学。

本文件主要起草人:何洪林、苏文、徐波、侯艳飞、郭学兵、黎建辉、张黎、任小丽、胡良霖、王志强、吴冬秀、潘贤章、胡波、徐文婷、郭欣妍、孟伟庆。

本文件及其所代替文件的历次版本发布情况为:

——2006 年首次发布为 GB/T 20533—2006;

——本次为第一次修订。

引 言

当前全球变化和可持续发展等一系列全球化的资源环境问题日益严峻,成为国际社会关注及科学研究热点,生态学研究重点由小尺度的个体、群落和生态系统向区域、大陆乃至全球尺度转变。生态系统观测也从过去的短时间、定位及小规模合作观测逐渐向长期、区域甚至全球大规模联合的网络化观测、天-空-地立体综合观测方向转变。生态系统观测研究所获取的数据逐渐由传统的数值型数据为主向多模态数据并存转变。同时,随着大数据和人工智能技术的发展,数据驱动、人工智能与生态学交叉融合的新型科研范式发展方兴未艾。这些变化迫切需提升生态科学数据管理、共享和应用水平。

近年来,特别是 2018 年我国《科学数据管理办法》颁布以来,相关领域一系列新的国家标准相继发布,如 GB/T 32843—2016《科技资源标识》、GB/T 35294—2017《信息技术 科学数据引用》、GB/T 30523—2023《科技资源核心元数据》、GB/T 43705—2025《科学数据安全分类分级指南》等。国际上生态学领域相关元数据规范《生态元数据语言》(Ecological Metadata Language, EML)也多次完善并发布新的版本。这对生态学领域科学数据在管理与共享、数据安全、知识产权保护及与国际接轨等方面也提出了新的要求。

生态科学数据作为生态科学研究创新的基础,数据规模、质量、管理和共享服务水平是影响其价值发挥的重要因素。而元数据标准是规范生态科学数据管理和实现生态科学数据共享应用的关键。GB/T 20533—2006《生态科学数据元数据》在我国过去近 20 年的生态科学数据管理与共享进程中发挥了重要作用,但目前其内容已不能适应新形势和新要求,亟需对其进行修订。

本文件面向数据集层次生态科学数据,用 UML 模型图和数据字典方式描述了通用元数据的构成,并定义了生态科学数据核心元数据的组成及制定元数据专用标准的规则与方法。本文件可为生态学领域科学数据中心、生态系统野外科学观测研究站以及其他与生态科学数据生产、管理与服务相关的组织和个人使用。本文件对我国相关部门或机构建立面向全社会的网络化、智能化的生态科学数据管理和共享服务体系具有重要支撑作用。本文件的实施有助于我国生态科学数据资源的建设积累,促进生态科学数据的规范化管理,提高生态科学数据的共享利用水平。

生态科学数据元数据

1 范围

本文件规定了一个在数据集层次上描述生态科学数据的概念模型,确立了构成该模型的一组必选、条件必选和可选的元数据子集、元数据实体和元数据元素,各种需求层次的元数据应用所需要的最小元数据元素集(简称核心元数据)以及为满足特殊需求对元数据进行扩展和制定元数据专用标准的规则,并描述了制定元数据专用标准的方法。

本文件适用于生态科学数据集和数据集系列的编目、描述与规范化管理,也适用于科学数据中心的数据服务。

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 4880(所有部分) 语种名称代码
- GB/T 7408(所有部分) 日期和时间 信息交换表示法
- GB/T 7714 信息与文献 参考文献著录规则
- GB/T 30522—2014 科技平台 元数据标准化基本原则与方法
- GB/T 30523—2023 科技资源核心元数据
- GB/T 32843—2016 科技资源标识
- GB/T 35294—2017 信息技术 科学数据引用
- GB/T 36369—2018 信息与文献 数字对象唯一标识符系统
- GB/T 39908—2021 科技计划形成的科学数据汇交 通用代码集

3 术语和定义

下列术语和定义适用于本文件。

3.1

生态科学数据 ecological science data

生态学领域科学研究活动中,通过观测、试验、调查等方式获取的原始基本数据,以及根据科研需求对原始基本数据和收集的其他数据进行加工处理、分析应用等产生的各种衍生数据的统称。

3.2

数据集 dataset

可识别的数据集合。

注:通过对覆盖范围或属性类型的限制,数据集在物理上能成为更大数据集中的一个比较小的数据组。

[来源:GB/T 19710.1—2023,3.3,有修改]

3.3

数据集系列 dataset series

具有共同特征的数据集的集合。