

中华人民共和国农业行业标准

NY/T 4061—2021

农业大数据核心元数据

Core metadata of agricultural big data

2021-12-15 发布

2022-06-01 实施



中华人民共和国农业农村部 发布



目 次

前言 II

1 范围 1

2 规范性引用文件 1

3 术语和定义 1

4 元数据元素构成和元数据属性 1

5 核心元数据 2

6 核心元数据扩展 6

附录 A(规范性) 农业大数据更新频率分类 7

附录 B(规范性) 农业大数据安全限制等级分类 8

附录 C(规范性) 农业大数据共享规定 9

附录 D(规范性) 农业大数据格式分类 10



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。

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。

本文件由农业农村部市场与信息化司提出。

本文件由农业农村部农业信息化标准化技术委员会归口。

本文件起草单位：中国农业科学院农业信息研究所、北京佳格天地科技有限公司、中国农业科学院农业资源与农业区划研究所。

本文件主要起草人：张弓、许世卫、张艳忠、白洁、吴文斌、张文鹏、顾竹、庄家煜、吴众望、李建政、白俊松、刘佳佳、于啸晨、彭欣、简敏。

农业大数据核心元数据

1 范围

本文件规定了农业大数据的元数据元素构成、元数据属性、核心元数据和核心元数据扩展等要求。
本文件适用于农业大数据采集、存储、处理以及信息共享和发布活动中的数据管理。

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 2659 世界各国和地区名称代码
GB/T 7408 数据元和交换格式 信息交换 日期和时间表示法
GB/T 18391.1 信息技术 元数据注册系统(MDR) 第1部分:框架
GB/T 19710 地理信息 元数据
GB/T 26816—2011 信息资源核心元数据

3 术语和定义

GB/T 18391.1、GB/T 19710 界定的以及下列术语和定义适用于本文件。

3.1

元数据 metadata

定义和描述其他数据的数据。

[来源:GB/T 18391.1—2009,3.2.16]

3.2

农业大数据 agricultural big data

在农业生产经营管理中产生的类型多样、结构复杂的数据集合,具有数据量大、来源多样、生成极快、多变等基本特征,以及地域性、季节性、多样性、周期性等农业行业特征。

3.3

元数据元素 metadata element

元数据的基本单元。

[来源:GB/T 19710—2005,4.6]

3.4

实体 entity

任何现存的、曾经存在的或可能存在的具体的或抽象的事物,包括事物间的关联。

注:实体的存在不依赖于是否有关于它的可用数据。

[来源:GB/T 18391.1—2009,3.2.10]

4 元数据元素构成和元数据属性

4.1 元数据元素构成

采用摘要表示的方式定义和描述元数据,包括以下属性:中文名称、英文名称、缩写名、定义、约束/条件、最大出现次数、数据类型、值域等。

4.2 元数据属性

4.2.1 中文名称

元数据实体或元素的中文名称。

4.2.2 英文名称

元数据实体或元素的英文名称,采用元数据实体或元素的英文全称或其组合。

4.2.3 缩写名

元数据实体或元素的英文缩写。缩写规则如下:

- a) 缩写名在本文件范围内应唯一。
- b) 元数据实体或元数据元素对象存在国际或行业领域惯用英文缩写等,采取该英文缩写为其缩写名。
- c) 对于根据英文名称或其他自定义的缩写名,在保持唯一性的前提下,统一取每个单词前3个字母作为其缩写名;当如此取词不能保证唯一性时,应后推一位字母延展取词位数,直至保证唯一性为止。
- d) 元数据实体缩写名的写法是,所有组成词汇的缩写为无缝连写,并且每个词汇缩写的首字母大写。
- e) 元数据元素缩写名的写法是,所有组成词汇的缩写为无缝连写,首词汇全部采用小写字母,其余每个词汇的缩写的首字母采用大写。
- f) 缩写名不应包括任何空格、破折号、下划线或分隔符等。

4.2.4 定义

对元数据实体或元素概念与内涵进行的说明。

4.2.5 约束/条件

说明元数据实体或元素是必选的、可选的或条件必选的。该描述可以有如下的值:

- “M”表示必选,元数据或元数据元素应当选用;
- “O”表示可选,元数据实体或元数据元素可以选用,也可以不选用;
- “C”表示条件必选,说明元数据实体或元素是否选用的条件,当该条件满足时,至少有一个元数据实体或元数据元素必选。

4.2.6 最大出现次数

元数据实体或元素可以具有的最大实例数目。只出现一次的用“1”表示,重复出现的用“N”表示。允许不为1的固定出现次数用相应的数字表示,如“2”“3”“4”等。

4.2.7 数据类型

描述该元数据所适用的数据类型,数据类型包括文本型、字符型、枚举型、日期型等基本类型。

4.2.8 值域

对于元数据元素,域表示该元素的允许取值范围或与之对应的类或数据类型的名称;对于元数据类,域表示在数据字典中描述该类的行的范围;角色名称的域表示与之关联的类的名称。

5 核心元数据

农业大数据核心元数据应符合表1的规定。

表 1 农业大数据核心元数据

序号	中文名称	英文名称	缩写名	定义	约束/条件	最大出现次数	数据类型	值域
1	标识	Identification	ID	唯一标识资源所需的基本信息	使用参照对象的约束条件	使用参照对象的最大出现次数		
1.1	中文名称	Chinese title	chiTitle	农业大数据的中文名称	M	1	文本型	自由文本
1.2	英文名称	English title	engTitle	农业大数据的英文名称	M	1	字符型	自由文本
1.3	版本	Edition	edition	农业大数据版本	O	1	字符型	自由文本
1.4	摘要	Abstract	abstract	农业大数据内容的简要说明	M	1	文本型	自由文本
1.5	关键词	Keyword	keyword	描述农业大数据主题的通用词或短语	O	N	文本型	自由文本
1.6	数据量	Memory Size	memSize	以农业大数据资源所占的物理存储空间表示的数据量	O	1	字符型	自由文本,用户可根据需要选择 KB、MB、GB 等单位
1.7	语种	Language	language	描述农业大数据所使用的语种	M	N	字符型	自由文本,按 GB/T 2659 执行
1.8	数据采集方式/来源	Data acquisition	dataAcquisition	描述农业大数据的数据来源或者采集方式	M	1	文本型	自由文本,包括自动采集、人工统计等
1.9	更新频率	Update frequency	upFrequency	描述农业大数据在多长的时间内更新一次	O	1	文本型	按照附录 A 的规定执行
1.10	信息资源链接地址	Physical address	phyAdd	可以访问农业大数据资源的有效网络地址	O	N	字符型	自由文本
1.11	安全限制等级	Right	right	农业大数据使用时的限制信息	M	1	字符型	按照附录 B 的规定执行
1.12	共享属性	Share property	shareProperty	农业大数据在共享时的可读写性	O	1	枚举型	按照附录 C 的规定执行
1.13	共享类型	Sharetype	shareType	农业大数据共享条件	M	1	枚举型	按照附录 C 的规定执行
1.14	共享方式	Sharemode	shareMode	农业大数据共享方式	O	1	枚举型	按照附录 C 的规定执行
2	数据格式	DataFormat	DataFormat	农业大数据的存储文件格式,包括类型、版本等	使用参照对象的约束条件	使用参照对象的最大出现次数		

表 1（续）

序号	中文名称	英文名称	缩写名	定义	约束/条件	最大出现次数	数据类型	值域
2.1	格式类型	Format type	formatType	农业大数据格式类型	M	1	文本型	按照附录 D 的规定执行
2.2	格式版本	Format version	formatVersion	农业大数据格式版本	O	1	字符型	自由文本
3	业务范围	Subject Range	SubRange	农业大数据内容所涉及的业务分类	使用参照对象的约束条件	使用参照对象的最大出现次数		
3.1	业务类别	Subject name	subName	农业大数据内容所属业务类型	M	N	文本型	自由文本,按农业农村部数据资源分类的规定执行
3.2	分类级别	Subject level	subLevel	农业大数据内容所属业务类型级别	M	1	字符型	自由文本
3.3	专题名称	Category name	catName	农业大数据所属的数据资源分类名称	O	1	文本型	自由文本,按农业农村部数据资源分类的规定执行
3.4	专题级别	Category level	catLevel	农业大数据所属的数据资源分类级别	O	1	字符型	自由文本,按农业农村部数据资源分类的规定执行
3.5	专题代码	Category code	catCode	农业大数据所属的数据资源分类编码	O	1	字符型	自由文本,按农业农村部数据资源分类的规定执行
4	地理范围	Geographic Range	GeographicRange	可以使用地理名称表示的空间范围	使用参照对象的约束条件	使用参照对象的最大出现次数		
4.1	行政区划名称	Administrative Name	admName	按行政权力覆盖面划分的区域的名称	M	N	文本型	自由文本,如全国、省、市、县等
4.2	地理区域名称	Geographic Region	geoRange	按地理环境表面的地理特征划分的区域的名称	O	N	文本型	自由文本,如青藏高原、长江三角洲等
4.3	农业区划名称	Agricultural Region	agriRange	按农业生产条件划分的有生产结构特点的区域名称	O	N	文本型	自由文本,如东北区、黄淮海区等
5	时间范围	Temporal Range	TemporalRange	农业大数据内容所涵盖的时间范围	使用参照对象的约束条件	使用参照对象的最大出现次数		
5.1	时间点	Single Date and Time	singleDateTime	农业大数据内容所涵盖的单个时间	O	N	日期型	按 GB/T 7408 的规定执行
5.2	时间段	Range Date and Time	rangeDateTime	农业大数据内容所涵盖的时间段	O	N	文本型	

表 1 (续)

序号	中文名称	英文名称	缩写名	定义	约束/条件	最大出现次数	数据类型	值域
5.2.1	开始时间	Begin Date and Time	beginDateTime	一个时间段的起始时间	M	1	日期型	按 GB/T 7408 的规定执行
5.2.2	结束时间	End Date and Time	endDatetime	一个时间段的结束时间,与开始时间相呼应	M	1	日期型	按 GB/T 7408 的规定执行
6	日期	Date	Date	农业大数据资源的相关日期	使用参照对象的约束条件	使用参照对象的最大出现次数		
6.1	更新日期	Date of update	upDate	农业大数据资源最近一次更新日期	O	1	日期型	按 GB/T 7408 的规定执行
6.2	发布日期	Date of publication	pubDate	农业大数据发布日期	O	1	日期型	按 GB/T 7408 的规定执行
7	责任者	ResponsibleParty	RespParty	负责农业大数据的创建、保存、管理、发布或提供该服务的单位	使用参照对象的约束条件	使用参照对象的最大出现次数		
7.1	责任者名称	Responsible party name	respName	农业大数据责任者名称	M	N	文本型	自由文本
7.2	责任者类型	Responsible party type	respType	农业大数据责任者类型	M	1	文本型	自由文本
7.3	责任者通讯地址	Responsible party address	respAdd	农业大数据责任者通信地址	O	1	文本型	自由文本
7.4	邮政编码	Responsible partypostal code	respPostCode	农业大数据责任者邮政编码	O	1	字符型	自由文本
7.5	联系电话	Responsible party phone	respPhone	农业大数据责任者联系电话	O	N	字符型	自由文本
7.6	邮件地址	Responsible party electronic mail address	respEmailAdd	农业大数据责任者邮件地址	O	N	字符型	自由文本

6 核心元数据扩展

6.1 核心元数据扩展原则

6.1.1 新建的元数据不应与本文件定义的元数据中现有的元数据实体、元素、代码表的名称、定义相冲突。

6.1.2 增加的元数据元素应按照本文件所确定的层次关系进行合理的组织,如果本文件现有元数据实体无法满足新增元数据需要,则可扩展元数据实体。

6.1.3 许以枚举型替代值域为自由文本的现有元数据元素的值域。

6.1.4 允许对现有的元数据元素的值域进行缩小。例如,在本文件中规定的元数据元素的值域为自由文本的,在扩展后可以规定它的值域为枚举型,采用某个代码表中的值。

6.1.5 允许对现有的元数据的可选性和最大出现次数施以更严格的限制。例如,在本文件中定义为可选的元数据,在扩展后可以是必选的;在本文件中定义为可无限次重复出现的元数据,在扩展后可以是只能出现1次。

6.2 核心元数据扩展技术要求

6.2.1 不应改变现有元数据元素的名称和定义。

6.2.2 不得将必选项变更为条件必选或可选项。

6.2.3 不得将条件必选项变更为可选项。

6.2.4 不得将代码表变更为自由文本。

6.2.5 不得改变代码表中的已有代码。

6.2.6 每一个扩展的元数据子集、实体和元素,应按照与本文件相同的格式定义其名称、缩写名、约束条件、最大出现次数、数据类型和值域等。

6.3 核心元数据扩展实施

农业大数据核心元数据进行扩展的步骤应按照 GB/T 26816—2011 中 6.3 的规定执行。

附 录 A
(规范性)
农业大数据更新频率分类

农业大数据更新频率的分类见表 A. 1。

表 A. 1 农业大数据更新频率分类

名称	说明
不定期	不定期更新数据
每分钟	每分钟更新数据
每小时	每小时更新数据
每日	每日更新数据
每候	每候更新数据
每周	每周更新数据
每十天	每 10 d(每旬)更新数据
半月	每半月更新数据
每月	每月更新数据
两月	每两月更新数据
每季度	每季度更新数据
每半年	每半年更新数据
每年	每年更新数据
按需要	定期但无法用上述名称描述的更新频率

附 录 B
(规范性)
农业大数据安全限制等级分类

农业大数据安全限制等级分类及编码见表 B. 1。

表 B. 1 农业大数据安全限制等级分类

等级分类	编码
公开	001
非公开	002



附 录 C
(规范性)
农业大数据共享规定

农业大数据共享属性、共享类型和共享方式的分类见表 C.1。

表 C.1 农业大数据共享规定

名称	说明
共享属性	可见、可读、可写 3 种属性
共享类型	共享、不共享 2 种类型
共享方式	数据库直连、FTP 下载、服务接口等方式

附 录 D
(规范性)
农业大数据格式分类

农业大数据格式的分类见表 D. 1。

表 D. 1 农业大数据格式分类

名称	说明
关系型数据库	以关系表格形式组织农业大数据,若干相关表格可共同构成一个数据库
非关系型数据库	分布式的、非关系型的、不保证遵循 ACID 原则的数据存储系统
文件系统	以数据库以外的文件形式组织科学数据,如大气文件、天文文件等
空间地理信息	利用计算机存储、处理和分析的地理环境要素数据
图像	文本之外的一种主要可视化表示方式。例如,物理对象的照片、航空影像、卫星影像、地图、图表及其他图像和图形
声音	以音频表现的农业大数据资源
视频	以视频表现的农业大数据资源
软件	以原格式或压缩格式存储的计算机程序
其他	以上类型所不能涵盖的数据集



中华人民共和国
农业行业标准
农业大数据核心元数据

NY/T 4061—2021

* * *

食典通

中国农业出版社出版

(北京市朝阳区麦子店街18号楼)

(邮政编码:100125 网址:www.ccap.com.cn)

北京印刷一厂印刷

新华书店北京发行所发行 各地新华书店经销

* * *

开本 880mm×1230mm 1/16 印张 1 字数 20 千字

2022 年 4 月第 1 版 2022 年 4 月北京第 1 次印刷

书号: 16109 · 8910

定价: 32.00 元



NY/T 4061—2021

版权专有 侵权必究

举报电话: (010) 59194261