

### 乡村振兴大数据 交换和共享规范

Big data for rural revitalization—Data exchange and sharing standard

(征求意见稿)

— XX — XX 发布

XXXX — XX — XX 实施

## 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 数据交换与共享范围 .....	2
5 交换和共享框架 .....	2
6 元数据属性 .....	2
6.1 概述 .....	2
6.2 属性说明 .....	2
7 元数据扩展 .....	3
7.1 元数据扩展的原则 .....	4
7.2 元数据扩展的类型 .....	4
7.3 元数据扩展的实施 .....	4
7.4 元数据扩展的报备 .....	4
8 数据交换与共享元数据 .....	5
8.1 概述 .....	5
8.2 描述 .....	5
9 数据交换与共享方法 .....	8
9.1 身份认证 .....	8
9.2 数据整理 .....	8
9.3 数据安全 .....	8
9.4 数据发布 .....	8
9.5 数据获取 .....	8
10 数据接口方式及说明 .....	8
10.1 接口描述规范 .....	8
10.2 接口报文规范 .....	10
10.3 数据获取接口方式 .....	10
10.4 数据获取接口说明 .....	10
参 考 文 献 .....	13

## 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。

本文件由北京市农业农村局提出并归口。

本文件由北京市农业农村局组织实施。

本文件起草单位：

本文件主要起草人：

# 乡村振兴大数据 交换和共享规范

## 1 范围

本文件规定了乡村振兴大数据交换与共享范围、交换和共享框架、数据交换与共享的元数据属性、元数据扩展、数据交换与共享元数据、数据交换与共享方法、数据接口方式及说明。

本标准适用于乡村振兴大数据交换和共享管理。

## 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 7408 数据元和交换格式 信息交换 日期和时间表示法

GB/T 26816 信息资源核心元数据

GB/T 40094.3 电子商务数据交易 第3部分:数据接口规范

NY/T 3501 农业数据共享技术规范

## 3 术语和定义

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。

### 3.2

**元数据元素 metadata element**

元数据的基本单元。

注:元数据元素在元数据实体中是唯一的

### 3.3

**元数据实体**

一组说明信息资源相关特性的元数据元素。

注:可以包含一个或一个以上元数据实体。

### 3.4

**数据交换与共享 data exchange and sharing**

实现不同地域、不同计算机、不同软件的用户获取、交换非本机数据并进行各种操作、运算和分析的过程。

### 3.5

#### 数据接口 data interface

进行数据传输时，数据提供方和接收方需遵守的接口约束或规定。

## 4 数据交换与共享范围

乡村振兴大数据交换和共享范围如下：

- 乡村振兴数据在农业农村局内部现有相关应用系统的数据交换与共享。
- 乡村振兴数据与下属各事业单位相关应用系统之间的数据交换与共享。
- 乡村振兴数据与农业农村部相关应用系统之间的数据交换与共享。
- 乡村振兴数据与其他横向部门信息系统之间为了实现业务协同，进行的数据交换与共享。

## 5 交换和共享框架

乡村振兴大数据交换和共享主要流程为数据提供、数据交换和数据获取三部分，具体流程见图1。

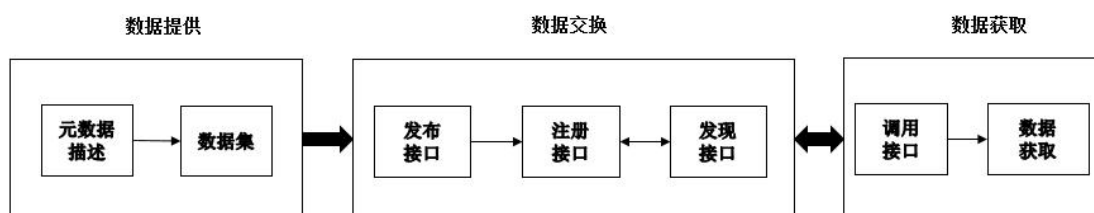


图1 总体交换和共享框架

## 6 元数据属性

### 6.1 概述

采用摘要表示的方式定义和描述元数据，摘要内容包括以下9个属性：中文名称、定义、英文名称、数据类型、值域、缩写名、约束/条件、最大出现次数、备注。

### 6.2 属性说明

#### 6.2.1 中文名称

指元数据元素或元数据实体的中文名称。

#### 6.2.2 定义

描述元数据实体或元数据元素的基本内容，给出信息资源某个特性的解释和说明。

#### 6.2.3 英文名称

元数据实体或元素的英文名称，一般用英文全称。

所有组成词汇为无缝连写。元数据元素的首词汇全部采用小写字母，其余每个词汇的首字母采用大写；元数据实体的每个词汇的首字母大写。

#### 6.2.4 数据类型

说明元数据元素或元数据实体的数据类型，对元数据元素的有效值域及允许的有效操作进行了规定。例如整型、实型、布尔型、字符串、日期型等。

#### 6.2.5 值域

规定了元数据元素的有效取值范围。

#### 6.2.6 缩写名

元数据元素或元数据实体的英文缩写名称。缩写规则如下：

- a) 缩写名在本标准范围内必须唯一；
- b) 缩写名不应包括任何空格、破折号、下划线或分隔符等；
- c) 缩写名不应使用复数形式的英文单词，除非该单词本身就是复数形式，如“Goods”；
- d) 元数据实体缩写名描述应采用 UCC 方式，即每个英文单词的首字母均大写；元数据元素缩写名描述应采用 LCC 方式，即除第一个英文单词外，每个单词的首字母大写，并把这些单词组合起来；
- e) 对存在国际或行业领域惯用英文名称缩写的，采用惯用缩写。

#### 6.2.7 约束/条件

元数据元素或元数据实体的约束/条件如下：

- a) M：必选，表明该元数据元素或元数据实体必须选择。
- b) O：可选，根据实际应用可以选择也可以不选的元数据元素或元数据实体。已经定义的可选元数据元素和可选元数据实体，可指导部门元数据标准制定人员充分说明其信息。如果一个可选元数据实体未被使用，则该实体所包含的元素（包括必选元素）也不选用。可选元数据实体可以有必选元素，但只当可选实体被选用时才成为必选。
- c) C：条件必选，当满足约束条件中所定义的条件时应选择。条件必选用于以下三种可能性之一：
  - 1) 当在多个选项中进行选择时，至少一个选项必选，且必须使用；
  - 2) 当另一个元数据元素已经使用时，选用一个元数据实体或元数据元素；
  - 3) 当另一个元数据元素已经选择了一个特定值时，选用一个元数据元素。

#### 6.2.8 最大出现次数

说明元数据元素或元数据实体可以出现的最大实例数目。只出现一次的用“1”表示，多次重复出现的用“N”表示。允许不为1的固定出现次数用相应的整型数值表示，例如“2”、“3”、“4”等。

#### 6.2.9 备注

可根据需要对元数据元素或实体进一步补充说明。

### 7 元数据扩展

## 7.1 元数据扩展的原则

- 7.1.1 选取元数据时，既要考虑数据资源单位的数据资源特点以及工作的复杂、难易程度，又要充分满足乡村振兴大数据的利用以及用户查询、提取数据的需要。
- 7.1.2 选取的元数据不但要满足乡村振兴大数据标准化需求，更应该考虑将来一定时间内可能产生的标准化需求。扩展过程中，可以积极参考国内和国外先进标准。
- 7.1.3 新建的元数据不应与第6章定义的元数据中的现有的元数据实体、元素、代码表的名称、定义相冲突。
- 7.1.4 增加的元数据元素应按照第6章所确定的层次关系进行合理的组织。如果第6章现有的元数据实体无法满足新增元数据的需要，则可以新建元数据实体。
- 7.1.5 新建的元数据实体可以定义为复合元数据实体，即可以包含现有的和新建的元数据元素作为其组成部分。
- 7.1.6 允许以代码表替代值域为自由文本的现有元数据元素的值域。
- 7.1.7 允许增加现有代码表中值的数量，扩充后的代码表应与扩充前的代码表在逻辑上保持一致。
- 7.1.8 允许对现有的元数据元素的值域进行缩小（例如，在第6章中规定的某元数据元素的值域中有7个值，在扩展后可以规定它的值域只包含其中的4个值）。
- 7.1.9 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，在第6章中定义为可选的元数据，在扩展后可以是必选的；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）。

## 7.2 元数据扩展的类型

信息资源在不同行业、不同地区的应用需求可能存在变化，要根据需求在应用中对本文件规定的元数据进行补充。在对本文件元数据内容进行扩展时，应包含第6章所定义的元数据，扩展的类型如下：

- a) 扩展元数据元素的值域；
- b) 增加新的元数据元素；
- c) 增加新的元数据实体类型；
- d) 对已有的元数据元素施加更严格的限定；
- e) 对已有元数据元素值域施加更多的限定；
- f) 创建新的元数据代码表元素（扩展代码表）。

## 7.3 元数据扩展的实施

- 7.3.1 在扩展元数据之前，应仔细地查阅第6章中现有的元数据及其属性，根据实际需求确认是否缺少适用的元数据。
- 7.3.2 对于每一个增加的元数据，应按照本标准第6章规定，采用摘要表示的方式，定义其中文名称、英文名称、数据类型、值域、注解。
- 7.3.3 对于新建的代码表和代码表元素，应说明代码表中每个值的名称、代码以及定义。

## 7.4 元数据扩展的报备

- 7.4.1 当元数据内容不能满足用户需求时，将相关的元数据标准文档报农业农村局信息管理部门进行审核。审核通过后，可以按照上述方法扩展。
- 7.4.2 数据检查、对齐后，返结果。

## 8 数据交换与共享元数据

### 8.1 概述

数据交换共享元数据包括12个元数据实体及元数据元素，如下：

- a) 交换与共享数据名称；
- b) 交换与共享数据代码；
- c) 交换与共享数据来源；
- d) 交换与共享数据摘要；
- e) 交换与共享数据关键字；
- f) 交换与共享数据格式；
- g) 数据类型；
- h) 发布日期；
- i) 涉密情况；
- j) 共享属性；
- k) 更新周期；
- l) 关联资源代码。

### 8.2 描述

#### 8.2.1 交换与共享数据名称

定义：缩略交换与共享数据内容的标题

英文名称：exchangandsharingDateName

数据类型：字符型

值域：自由文本

注解：必选项；最大出现次数为1

#### 8.2.2 交换与共享数据代码

定义：交换与共享数据唯一不变的标识代码

英文名称：exchangandsharingDateCode

数据类型：字符型

值域：自由文本

注解：可选项；最大出现次数为1

#### 8.2.3 交换与共享数据来源

定义：说明数据与共享提供方及其联系方式的信息

英文名称：exchangandsharingDateSource

数据类型：复合型

注解：必选项；最大出现次数为N

##### 8.2.3.1 交换与共享数据提供方

定义：提供信息的单位

英文名称：exchangandsharingDateProvider

数据类型：字符型



值域：自由文本

注解：必选项；最大出现次数为1

#### 8.2.3.2 交换与共享数据提供方联系方式

定义：提供信息的单位联系方式

英文名称：providerTelephone

数据类型：字符型

值域：7~18位数字字符（包括企业或人员的固定联系电话号码，完整的电话包括国际区号、国内长途区号、本地电话号和分机号，之间用“-”分隔。）

注解：必选项；最大出现次数为1

#### 8.2.4 交换与共享数据信息摘要

定义：对信息资源内容的概要描述

英文名称：exchangandsharingDateAbstract

数据类型：字符型

值域：自由文本

注解：必选项；最大出现次数为1

#### 8.2.5 交换与共享数据关键字

定义：用于概况描述数据信息内容的通用词、形式化词或短语

英文名称：exchangandsharingDateKeywords

数据类型：字符型

值域：自由文本

注解：可选项；最大出现次数为1

#### 8.2.6 交换与共享数据格式

定义：信息资源的存在方式

英文名称：exchangandsharingDateFormat

数据类型：字符型

值域：电子信息文档、数据库格式

注解：必选项；最大出现次数为1

#### 8.2.7 数据类型

定义：标明该信息项的数据类型

英文名称：dataType

数据类型：字符型

值域：电子文档数据存储类型，见GB 13000—2010

注解：必选项；最大出现次数1

#### 8.2.8 发布日期

定义：信息资源提供方发布共享信息资源的日期

英文名称：releaseDate

数据类型：日期型

值域：按GB/T 7408—2005执行，格式为CCYY-MM-DD

注解：可选项；最大出现次数1

### 8.2.9 涉密情况

定义：对资源密级状态的说明。

英文名称：secretLevel

数据类型：字符型

值域：“公开”、“秘密”、“机密”、“绝密”

注解：必选项；最大出现次数1

### 8.2.10 共享属性

定义：说明信息共享情况的信息

英文名称：SharingRule

数据类型：复合型

注解：必选项；最大出现次数1

#### 8.2.10.1 共享类型

定义：信息资源的共享类型，包括无条件共享、有条件共享、不予共享

英文名称：sharingType

数据类型：字符型

值域：无条件共享为1、有条件共享为2、不予共享为3

注解：必选项；最大出现次数1

#### 8.2.10.2 共享条件

定义：不同共享类型的信息资源的共享条件

英文名称：sharingCondition

数据类型：字符型

值域：自由文本

注解：必选项；最大出现次数1

#### 8.2.10.3 共享方式

定义：获取信息资源的方式

英文名称：sharingMode

数据类型：字符型

值域：自由文本

注解：必选项；最大出现次数1

### 8.2.11 更新周期

定义：对信息资源更新的频率

英文名称：accessPeriodicity

数据类型：字符型

值域：自由文本，如：实时、每日、每周、每月、每季、每年等

注解：必选项；最大出现次数1

### 8.2.12 关联资源代码

定义：相关联资源的符合规范标识体系的代码

英文名称：relationCode

数据类型：字符型

值域：自由文本，如该信息资源同属于其他资源分类体系，需要标注该资源在其他分类体系中的代码

注解：可选项；最大出现次数N

## 9 数据交换与共享方法

### 9.1 身份认证

接口调用方(客户端)向接口提供方(服务器)申请接口调用账号，申请成功后，接口提供方会给接口调用方一个app\_id和app\_key。

### 9.2 数据整理

对需要共享的数据利用元数据进行定义、标识、表示及允许值等一系列属性描述，方便数据对象在不同部门、不同系统之间进行流通和流转，并确保流过程中数据标准的一致性。

### 9.3 数据安全

在乡村振兴大数据共享的过程中应符合相关法律法规的规定，应采取措施保证数据的有序安全开放，如根据数据等级对传输数据进行加密，防止所开放的数据被泄密或篡改，并可对加密数据进行跟踪和审计。

### 9.4 数据发布

数据发布方法如下：

- 数据提供方使用元数据定义数据元；
- 定义数据获取接口；
- 与共享中心连接；
- 身份验证获取唯一标识；
- 携带身份标识，注册并发布服务接口,完成数据接口注册共享。

### 9.5 数据获取

数据获取方法如下：

- 与共享中心连接，身份验证，获取身份标识；
- 携带身份标识，能过查找获得所需服务接口；
- 按接口定义，调用服务，获取数据；
- 数据检查、对齐后，返结果。

## 10 数据接口方式及说明

### 10.1 接口描述规范

### 10.1.1 接口描述规则

接口服务描述是业务服务的集中描述，包含其提供的功能以及如何进行调用；是在系统发布的服务提供方和服务调用方之间的规约。接口服务描述必须包含对各个方面的人员有意义并可理解的信息，接口服务描述要包含概要和详细服务信息，使用平台/技术实现无关的语言进行描述以业务为导向。

接口服务描述主要由2部分组成：接口服务属性和接口服务规约。

#### 10.1.1.1 接口服务属性

接口服务属性是业务接口服务的描述，它提供组织和性能相关要求，是接口服务提供方和接口服务消费方之间服务协议的依据。定义一个接口服务应该包含的属性如下所示：

- 服务编码（必选）：能体现出接口服务的业务含义的唯一识别编码，应遵循统一的编码规则，定义请参考交易服务代码规范；
- 接口服务中文名称（必选）：能体现接口服务的意义，简单概括，应少于 30 个汉字；
- 接口服务描述（必选）：清晰描述接口服务的业务功能；
- 服务提供方（必选）：提供此接口服务的业务系统；
- 服务消费方（可选）：可能会使用到此接口服务的业务系统；
- 服务等级（必选）：分为紧急、普通和低级；
- 报文文件名：指请求报文文件名和响应报文文件名；
- 数据量（必选）：运行期间单次访问的数据量均值和范围，比如均值为 200K，范围：100K-300K；
- 运行效率（可选）：单次访问的响应时间均值和范围；
- 支持并发（可选）：此接口可支持的并发访问量；
- 是否可重（可选）：接口是否能被多次重试调用；
- 超时说明（必选）：提供超时时间及其说明。
- 返回码说明（必选）：业务上需提供返回码和返回信息给服务消费方。
- 技术方式（必选）：明确此接口的技术方式；
- 调用说明（必选）：提供调用此接口服务的必要条件。
- 安全要求（必选）：描述此接口服务的安全要求。
- 接口服务调用权限（必选）：描述此接口服务的调用权限。
- 描述该接口可以被哪些系统调用，即服务的调用权限。

#### 10.1.1.2 接口服务规约

接口服务规约是接口及参数的技术描述，定义所有绑定和传输信息，以及所有支持的操作及相关输入、输出的格式等以XSD格式描述；而Webservice协议通常以WSDL形式存在，被称为服务WSDL。接口服务规约包含：详细接口服务操作描述、数据类型、消息格式和结构、绑定的传输协议和服务的位置。

#### 10.1.2 接口服务实现规范

接口服务实现需要遵循的规范如下：

- 接口服务命名要遵循一致的服务命名规范；
- 接口服务应该遵守统一的报文规范；
- 接口服务为了重用可以适当提高接口的颗粒度；
- 接口服务的设计和定义应该与接口的实现分阶段进行；
- 接口服务中传输的报文要求都是经过校验的，符合业务规则的，否则不符合报文会被返回；
- 接口服务应充分考虑到扩展性，例如在定义数据结构时多个字段封装为一个可扩展的对象；

——接口服务应尽可能通用，对于同一业务对象，应避免为不同系统开发不同接口服务。

### 10.1.3 接口服务检查规范

接口服务抽取后要对服务的交互进行检查，通过此检查来发现该接口服务是否合理。以下列出的是对接口服务进行检查的指导性原则，这些原则在运用的时候可以根据具体的项目情况进行裁减和选择。

- 业务服务能否可以被一个以上的流程使用；
- 业务服务是否是灵活和完整的；
- 业务服务是否是业务相关、有业务含义并被业务人员所能理解的；
- 该服务在多大程度上依赖于其它业务服务；
- 该服务是否是无状态的服务；
- 该服务的描述是否和其它服务的描述相冲突；
- 是否没有类似或重复的服务存在。

## 10.2 接口报文规范

### 10.2.1 报文组织格式

根据应用集成标准的要求，为使不同应用系统能够直接有效的进行集成，系统主要使用json或xml格式报文。无论在何种传输协议上，不同系统之间面对的报文格式需要统一。无论是json格式还是xml格式，两者语义必须等价。

### 10.2.2 XML 报文规范

采用XML方式来传输接口数据，XML报文满足通用性和扩展性原则，能满足业务的各种需要，而且方便开发人员应对不断变化的业务需求，方便新交易的开发。

### 10.2.3 JSON 报文规范

JSON是一种轻量级的数据交换格式，它基于 ECMAScript的一个子集，采用完全独立于编程语言的文本格式来存储和表示数据。JSON数据交换语言应易于人阅读和编写，同时也易于机器解析和生成，并有效地提升网络传输效率。

## 10.3 数据获取接口方式

数据获取接口方式包括但不限于：

- 文件接口：以电子文件形式传输数据、交换信息的接口方式；
- API 接口：https（超文本传输安全协议，Hypertext Transfer Protocol Secure）网络协议的网络服务，用于交易主体之间传输获取数据的接口方式。

## 10.4 数据获取接口说明

### 10.4.1 文件接口

#### 10.4.1.1 概述

文件接口是数据需求方获得数据授权后由数据服务平台为其提供的下载端口，通过端口可直接获取数据文件。

#### 10.4.1.2 接口说明

文件接口说明包括但不限于：

a) 文件下载地址：可获取到数据文件的 URL 地址；URL 一般包括模式或协议、服务器名称或 IP（互联网协议地址，Internet Protocol Address）地址、路径和文件名等；

示例：协议：//子域名.域名.顶级域名:端口号/目录/文件名.文件后缀?参数=值#标志

b) 文件类型：数据文件的存储格式说明；

示例：rar 文件

c) 文件大小：数据文件的数据量描述；

示例：10M

d) 附加说明：数据文件的其他说明。

## 10.4.2 API 接口

### 10.4.2.1 概述

API接口是乡村振兴大数据平台中数据提供方、数据获取方分别共享和获取共享数据集的服务方式。

### 10.4.2.2 接口参数

接口参数要求如下：

a) 公共参数

公共参数是每个接口都要携带的参数,描述每个接口的基本信息,用于统计或其他用途,放在Header或url参数中。主要参数及说明见表1。

表 1 主要参数及说明

参数	说明	备注
app_id	唯一标识用户ID	app_id是全局唯一的，每个app_id将对应一个客户
app_key	加密key	app_key用于参数签名使用。
timestamp	时间戳	时间戳，是客户端调用接口时对应的当前时间戳，时间戳用于防止DoS攻击。
request_id	请求ID	用户请求ID，是客户端随机生成的值,要保证全局唯一。
sign	签名	一般用于参数签名，防止参数被非法篡改。
token	系统调用的唯一凭证	访问令牌access token，用于接口中，用于标识接口调用者的身份、凭证，减少用户名和密码的传输次数。

一般token、timestamp、request\_id和sign 四个参数会在接口中会同时作为参数传递，每个参数都有各自的用途，其中首次获取token需要app\_id、timestamp、request\_id、sign，客户端获取token后不再需要传递app\_id。

b) API 参数

访问接口时必要的或可选的参数,应说明参数的名称、类型、是否必填、示例值和描述说明等属性,参数格式统一采用json格式在header中传递。具体实现由数据提供方定义。

### 10.4.2.3 接口说明

API接口说明包括但不限于：

a) 接口标识：赋予 API 接口的唯一标识；由平台或数据提供方自定义；

- b) 接口名称：接口传输的数据核心内容的简要、概况性描述，应突出数据内容指向的对象以及对象具有的显著特征，如时间特征、地域特征、行业特征等；由数据提供方自定义；
- c) 接口描述：接口的功能描述；由数据提供方自定义；
- d) 接口地址：接口的 URL 访问地址；由数据提供方自定义；
- e) 请求方式：对数据进行操作采用的类型，可采用 GET 或 POST 等方式；
  - 1) GET 请求数据的接口地址中参数禁止直接使用中文，如果在 URL 中使用中文必须统一转码。
  - 2) Post 请求数据的接口地址中除调用身份参数 token 直接在 Header 或 RUL 中。
  - 3) 请求参数：访问接口时必要的或可选的参数，应说明参数的名称、类型、是否必填等属性；具体实现由数据提供方自定义。
- f) 返回数据：请求数据后，系统返回的相关结果数据。数据包含三个属性：状态码（code）、信息描述（message）、响应数据（data）。
- g) 错误码参照：请求数据后，系统返回的错误码及其返回说明。客户端根据状态码及信息描述可快速知道接口，如果状态码返回成功，再开始处理数据。由系统开发方自定义。

示例：

错误码	错误描述
3004	请求参数不存在
3005	请求过于频繁，请稍后再试
3006	access_token 参数不能为空
3007	该 IP 已被禁用
3008	接口需要授权
3009	接口调用异常，请稍后再试
3010	access_token 令牌无效
3011	该数据只允许企业用户调用
3012	调用 api 网关失败，请与客服联系
300	缺少名称参数
301	未查询到数据

### 参 考 文 献

- [1] GB/T 39477—2020 信息安全技术 政务信息共享数据安全技术要求
  - [2] DB43/T 1754—2020 政务大数据中心数据交换规范
-