

雄安新区数据资源目录设计规范

目 录

一、概述.....	278
(一) 目的.....	278
(二) 适用范围.....	278
(三) 规范性引用文件.....	278
(四) 术语定义.....	279
二、数据资源目录.....	280
(一) 数据资源目录分类.....	280
(二) 数据资源目录编码.....	281
(三) 数据资源目录管理.....	283
三、元数据要求.....	287
(一) 业务元数据要求.....	287
(二) 技术元数据要求.....	289
(三) 管理元数据要求.....	292

一、概述

（一）目的

《河北雄安新区规划纲要》明确提出要“建立健全大数据资产管理体系，打造具有深度学习能力、全球领先的数字城市”。为推动新区构建多层次的数据资源管理体系，探索数据共享授权机制，实现政务数据、位置数据和城市生产、生活、运行数据的底层融合和授权共享使用，制定本导则。

（二）适用范围

本数据目录通用要求规定了数据资源目录的分类、管理模式、运行模式和总体框架，以及目录建设的元数据要求、功能要求、技术要求和安全要求，适用于雄安新区全区范围内数据资源目录建设的行为及过程。

（三）规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件，仅注日期的版本适用于本文件。凡是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 21063.1—2007 《政务信息资源目录体系 第1部分：总体框架》

GB/T 21063.2—2007 《政务信息资源目录体系 第2部分：技术要求》

GB/T 21063.3—2007 《政务信息资源目录体系 第3部分：核心元数据》

GB/T 21063.4—2007 《政务信息资源目录体系 第4部分：政务信息资源分类》

GB/T 21063.6—2007 《政务信息资源目录体系 第6部分：
技术管理要求》

《政务信息资源共享管理暂行办法》

《政务信息系统整合共享实施方案》

《政务信息资源目录编制指南（试行）》

《加快推进落实政务信息系统整合共享实施方案工作方案》

《河北省政务信息资源目录编制指南（试行）》

（四）术语定义

GB/T 21063.1—2007 界定的以及下列术语和定义适用于本文件。

数据资源：以一定形式记录、保存的文件、资料、图表和数据等各类数据的集合。

元数据：描述数据资源属性的数据。

业务元数据：从业务角度描述数据资源各种属性的数据。

技术元数据：从技术角度描述数据资源各种属性的数据。

管理元数据：从管理协调角度描述数据资源各种属性的数据。

数据资源目录：以元数据形式记录、按照一定的分类方法进行排序和编码的一组数据，用以描述各个数据资源的特征，以便于对数据资源的检索、定位与获取。

共享库：用于共享和开放数据资源的数据库。

标签：用于揭示数据资源内容和特征的标志体。

数据挂接：将数据资源与数据资源目录对接的过程。

任务调度：将数据使用方的数据获取需求转化为数据交换或

数据治理能够识别的数据内容。

二、数据资源目录

（一）数据资源目录分类

1.资源属性分类

按资源属性不同，数据资源目录分为部门目录、主题目录、基础目录。

（1）部门目录

按照资源目录标准，根据资源提供方职能职责梳理形成的资源目录。

（2）主题目录

是针对某个特定对象或某一业务领域而专门梳理形成的资源目录，如精准扶贫资源目录等。

（3）基础目录

基础目录是广泛应用的国家基础政务数据，具有基础性、基础性、标识性等特征，主要包括人口基础目录、法人基础目录、空间地理基础目录等。

2.功能分类

按流程环节中功能不同，数据资源目录可分为业务目录、数据目录和管理目录。

（1）业务目录

以资源提供方履行职责过程中的应用视角梳理形成的数据资源目录。

（2）数据目录

根据业务目录建设共享数据库,并对该共享数据库依据元数据标准进行描述而形成的数据资源目录。

(3) 管理目录

依据数据共享要求形成数据服务调度任务,提供具体数据服务的数据资源目录。

3. 层级属性分类

按目录所属的层级属性可划分为三类:

(1) 上级目录

指河北省面向全省或者雄安新区提供的数据资源目录。

(2) 本级目录

指雄安新区独立建设的数据资源目录。

(3) 下级目录

指雄安新区下辖行政区划建设的数据资源目录。

(二) 数据资源目录编码

数据资源标识符由前段码(6位数字)、中段码(10位数字和字母)和后段码(5位数字)三部分组成,中段码和后段码之间用字符“/”隔开,其形式如图1所示。

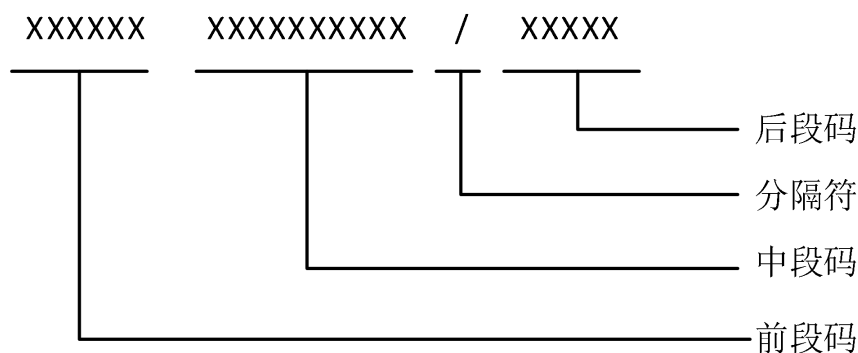


图1 数据资源标识符构成

1.前段码

按照《政务信息资源目录编制指南(试行)》(发改高技[2017]1272号)中关于政务信息资源代码结构规则的要求,隶属于河北省的政务信息资源的前段码统一为330003。

雄安新区数据资源的前段码可参照设置,以确保与国家政务信息资源目录的对接。

2.中段码

中段码由行政区划代码(6位数字)、部门类别码(1位字母)和部门编码(3位数字)组成。其中,雄安新区行政区划代码为139900,部门类别参考表1(其中A-K对照河北省政府资源目录部门类别),部门编码中参考政府部门代码和企业登记号设置。

表 1 部门类别码表

序号	部门类别码	部门类别名称
1	A	党委
2	B	人大
3	C	政府
4	D	政协
5	E	民主党派
6	F	群众团体
7	G	法院
8	H	检察院
9	J	中央驻地单位
10	K	其他政务部门
11	L	公共事业单位
12	M	企业

3.后段码

后段码，即数据资源顺序码，是数据资源标识符中对单位所管理或拥有的数据资源进行的唯一标识。后段码共5位，由阿拉伯数字（0~9）组合而成。

（三）数据资源目录管理

1.参与方

数据资源目录全生命周期主要涉及提供方、使用方、运营方和监管方4类参与方。

（1）提供方

- 负责对资源进行目录编制、注册、更新。
- 建立并维护共享信息库。
- 审批数据服务申请。

（2）使用方

- 在协议范围内对数据进行使用，并保证数据不扩散泄露。

（3）运营方

- 与上级平台级联。
- 对提供方提交的目录进行审核、发布。
- 审批数据服务申请。
- 核查目录更新、数据更新。
- 调度服务任务。

（4）监管方

- 查看平台的目录、共享开放等各事项指标结果。

2. 本级目录

本级目录管理与服务流程如图2所示。

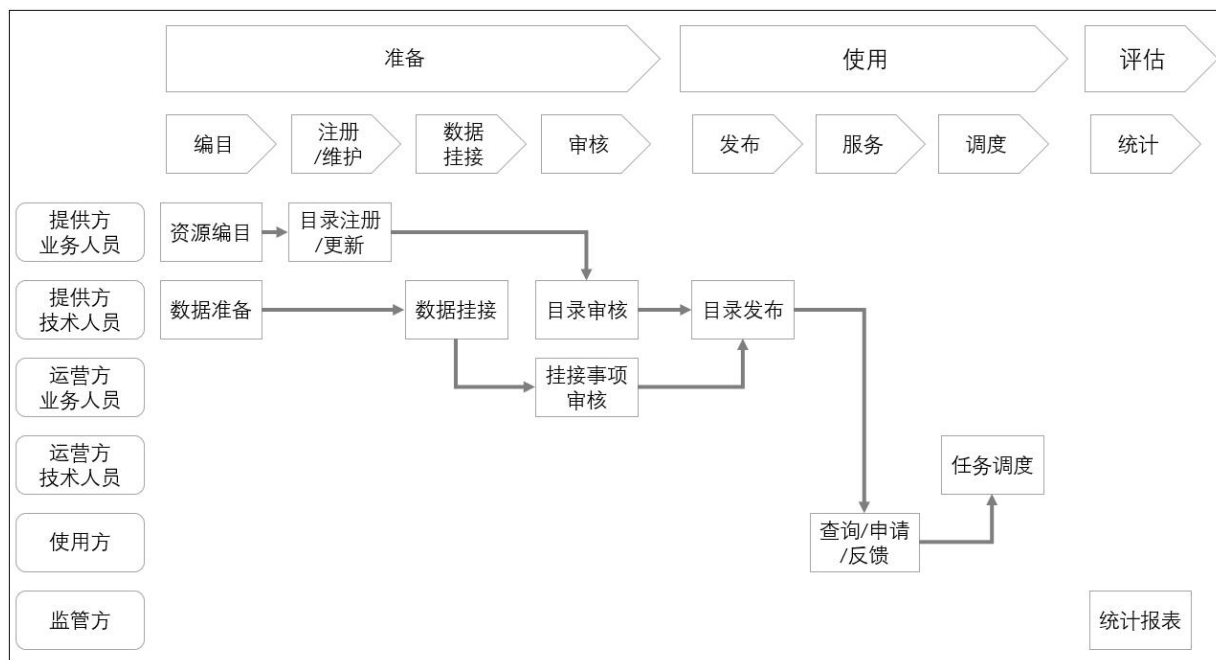


图2 本级目录管理与服务整体流程

(1) 参与角色

本级目录管理与服务整体流程涉及4类参与方的6种角色，具体如下：

1) 提供方

——提供方业务人员

——提供方技术人员

2) 运营方

——运营方业务人员

——运营方技术人员

——使用方

——监管方

(2) 主要活动

本级目录管理与服务整体流程涉及3个环节8项活动，具体描述如下：

1) 数据准备

——编目：数据资源提供方业务人员，以业务视角对资源进行梳理，通过编目工具或编目系统形成业务目录内容；资源提供方技术人员，依据业务目录建立对应的共享信息库，并提取技术特征形成技术目录。

——注册/维护：数据资源提供方业务人员将目录内容注册至目录链节点，维护共享库数据及目录内容。

——数据挂接：数据资源提供方技术人员将共享信息库与对应的资源目录进行挂接。

——审核：运营方业务人员对注册的资源目录内容进行审核，审核通过可进行发布上架；平台技术人员对数据挂接的技术信息进行审核，审核通过后与资源目录一并待发布上架；发布上架的目录内容变更，需经过审核后重新发布上架。

2) 数据使用

——发布：运营方审核通过的资源目录进行发布上架。

——服务：使用方通过浏览器等客户端查询目录内容，提交数据获取申请。

——任务调度：数据获取申请审批通过后，由平台技术人员

调度交换或治理任务，使用方通过任务获取实时或定时的数据服务。

3) 数据评估

——统计：对流程中各环节从不同维度定制统计报表，监管方用于数据共享监督考核。

3. 上级目录

本级目录需与上级目录级联，以获取上级平台数据资源。

(1) 参与角色

上级目录管理与服务流程涉及3类参与方的4种角色，具体如下：

1) 运营方

——运营方业务人员

——运营方技术人员

2) 使用方

3) 监管方

(2) 主要活动

本级目录管理与服务整体流程3个环节7项活动，具体描述如下：

1) 数据准备

——编目：平台技术人员通过级联系统获取资源目录，并通过编目系统提取数据资源的技术特征，形成技术目录，并依据上级平台已经发布的目录编制形成本级业务目录。

——注册/维护：平台技术人员将目录内容注册至目录链节

点，依据上级平台目录内容的变化维护本级目录内容。

——审核：运营方业务人员对注册的资源目录内容进行审核，审核通过可进行发布上架。

2) 数据使用

——发布：运营方审核通过的资源目录进行发布上架。

——服务：使用方通过浏览器等客户端查询目录内容，提交数据获取申请。

——任务调度：数据获取申请审批通过后，由平台技术人员调度交换或治理任务，使用方通过任务获取实时或定时的数据服务。

3) 数据评估

——统计：对流程中各环节从不同维度定制统计报表，监管方用于数据共享监督考核。

4. 下级目录

雄安新区下级目录主要涉及下辖区县，考虑到区县目录可直接放在雄安新区目录下，无需专门建设目录平台，因此下级目录管理与服务流程与本级目录相同。

三、元数据要求

(一) 业务元数据要求

业务元数据应包括以下元数据：

1. 数据资源名称（数据资源内容的标题名称）
2. 内容摘要（对资源内容进行概要说明的文字）

3.注册日期（提供方注册数据资源的日期）

4.提供方

——提供方部门/单位：资源提供方部门/单位全称，不可使用简称。

——提供方部门/单位联系人：资源提供方部门/单位联系人姓名。

——提供方部门/单位联系人电话：资源提供方部门/单位联系人电话。

5.提供方代码（资源提供方部门/单位代码）

6.资源标识符（数据资源唯一不变的标识编码）

7.资源技术支持

——资源技术支持部门/单位：资源技术支持部门/单位名称。

——资源技术支持联系人：资源技术支持联系人姓名。

——资源技术支持联系人电话：资源技术支持联系人电话。

8.资源分类

——分类方式：资源所采用的分类方式。

——分类类目：分类方式中的类目名称。

9.共享类型（数据资源共享类型包括无条件共享、有条件共享和不予共享三类）。

10.开放类型（数据资源开放类型包括无条件开放、有条件开放和不予开放三类）。

11.数据项

——数据项名称（数据资源中具体数据项的标题）。

——数据项共享类型（数据资源中数据项共享类型包括无条件共享、有条件共享和不予共享三类）。

——数据项开放类型（数据资源中数据项共享类型包括无条件开放、有条件开放和不予开放三类）。

12.共享方式：数据采用的共享方式，包括数据交换、接口服务、文件下载。

13.数据更新周期（实时、每天、每周、每月、每季度、每半年、每年）。

14.标签（用于标识资源内容的通用词、形式化词或短语）。

（二）技术元数据要求

技术元数据应包括以下元数据：

1.技术资源名称（技术资源内容的名称，可以是系统名称、表名称、文件名称、文件夹名称，也可以根据资源内容提取）。

2.数据格式类型（资源数据的格式分类，包括数据库、电子文件、电子表格、数据接口、图形图像、视频媒体、自描述格式）。

——数据库

- 数据库类型：可用于共享的数据库类型。
- 数据库地址：可用于共享的数据库地址。
- 数据库用户名：可用于共享的数据库用户名。
- 数据库密码：可用于共享的数据库密码。
- 表中文名称：可用于共享的数据库表中文名称。

- 表英文名称：可用于共享的数据库表英文名称。
- 字段中文名称：可用于共享的数据库字段中文名称。
- 字段英文名称：可用于共享的数据库字段英文名称。
- 字段类型：可用于共享的数据库字段类型。
- 字段长度：可用于共享的数据库字段长度。

—— 电子文件

- 文件名称：电子文件名称（资源为一个文件使用此项元数据）。

- 文件夹名称：文件夹名称（资源包括多个文件使用此项元数据，统一存放在指定文件夹中）。

- 文件格式：电子文件存储格式，如 doc、docx、pdf、html、wps、txt、ppt 等格式。

- 存储地址：电子文件存储地址。

—— 电子表格

- 文件名称：电子文件名称。

- 文件夹名称：文件夹名称（资源包括多个文件使用此项元数据，统一存放在指定文件夹中）。

- 文件格式：电子表格存储格式，如 xls、xlsx、csv、et 等格式。

- 存储地址：电子文件存储地址。

—— 数据接口

- 接口名称：数据接口中文描述。

- 接口类型：数据接口类型，包括 rest、soap。
- 接口地址：数据接口服务地址。
- 服务编码：接口服务编码，如 UTF-8、GBK 等。
- 接口参数：数据接口服务参数信息。

参数名称：接口调用时所需参数名称。

参数注释：接口调用时所需参数注释或描述。

请求位置：接口调用时参数的请求位置，default、header。

参数类型：接口调用时参数的类型，string、number、boolean。

是否必须：接口调用时参数是否必须项。

- 接口返回值：数据接口服务返回值信息。

返回数据名称：接口调用返回数据中单项名称。

返回数据注释：接口调用返回数据中单项注释或描述。

返回数据类型：接口调用返回数据中单项数据类型。

- 接口使用说明：接口使用手册地址或描述说明。

—— 图形图像

- 文件名称：图形图像文件名称（资源为一个文件使用此项元数据）。

- 文件夹名称：文件夹名称（资源包括多个文件使用此项元数据，统一存放在指定文件夹中）。

- 文件格式：图形图像文件的格式。

- 存储地址：电子文件存储地址。

—— 视频媒体

- 文件名称：视频文件名称（资源为一个文件使用此项元数据）。

- 文件夹名称：文件夹名称（资源包括多个文件使用此项元数据，统一存放在指定文件夹中）。

- 文件格式：视频文件的格式。

- 存储地址：电子文件存储地址。

（三）管理元数据要求

管理元数据应包括以下元数据：

1. 审核日期（资源审核通过的日期）

2. 发布日期（资源发布上线的日期）

3. 更新日期（资源发生变更的日期）

4. 目录版本（目录的版本号）

5. 使用方

——资源使用方部门/单位：资源使用方部门/单位全称，不可使用简称。

——资源使用方部门/单位联系人：资源使用方部门/单位联系人姓名。

——资源使用方部门/单位联系人电话：资源使用方部门/单位联系人电话。

附录（资料性附录）

核心元数据描述表

元数据	说明
1.数据资源名称	<p>定义：缩略描述数据资源内容的标题。</p> <p>数据类型：字符型。</p> <p>注解：必选项；最大出现次数为1。</p> <p>说明：缩略描述对应数据资源具体内容的标题。</p>
2.内容摘要	<p>定义：对数据资源内容（或关键字段）的概要描述。</p> <p>数据类型：字符型。</p> <p>注解：必选项；最大出现次数为1。</p> <p>说明：对数据资源内容进行概要说明（或关键字段）的描述。</p>
3.注册日期	<p>定义：提供方注册数据资源的日期。</p> <p>数据类型：日期型，格式为CCYY-MM-DD。</p> <p>注解：必选项；最大出现次数为1。</p> <p>说明：按定义要求进行描述。</p>
4.提供方	<p>定义：提供数据资源的政务部门。</p> <p>数据类型：字符型。</p> <p>注解：必选项；最大出现次数为1。</p> <p>说明：具体提供信息资源的部门和单位，原则上中央政务部门细化到司局或所属行政事业单位，地方政务部门细化到内设机构和所辖政务部门，企业细化的最小法人实体。</p>
5.资源标识符	<p>定义：数据资源的唯一不变的标识代码。</p> <p>数据类型：字符型。</p> <p>注解：必选项；最大出现次数为1。</p>
6.共享类型	<p>定义：数据资源的共享类型包括：无条件共享、有条件共享、不予共享三类。值域范围对应共享类型排序分别为1、2、3。</p> <p>数据类型：数值型。</p> <p>注解：必选项；最大出现次数为1。</p> <p>说明：对每个数据资源按不同共享类型进行描述。</p>
7.开放类型	<p>定义：数据资源面向社会开放的属性，包括“是”和“否”，对应取值分别为1和0。</p> <p>数据类型：数值型。</p> <p>注解：必选项；最大出现次数为1。</p> <p>说明：按定义要求进行描述。</p>

元数据	说明
8.提供方代码	<p>定义：提供数据资源的部门和单位代码。</p> <p>数据类型：字符型。</p> <p>注解：必选项；最大出现次数为1。</p> <p>说明：政务部门代码采用《国务院关于批转发展改革委等部门法人和其他组织统一社会信用代码制度建设总体方案的通知》中规定的法人和其他组织统一社会信用代码。提供方代码采用资源分类“项”中的政务部门代码，而非部门内设机构和直属单位代码。企业单位代码采用全国组织机构代码统一社会信用代码。</p>
9.数据格式类型	<p>定义：数据资源的存在方式（可多选），包括数据库、电子文件、电子表格、数据接口、图形图像、视频媒体、自描述格式等多种类型。</p> <p>数据类型：字符型。</p> <p>注解：必选项；最大出现次数为1。</p> <p>说明：按定义的多种格式选择确定。</p>
10.数据项名称	<p>定义：描述结构化信息资源中具体数据项的标题。适用于格式为数据库、电子表格类等的信息资源。</p> <p>数据类型：字符型。</p> <p>注解：可选项；最大出现次数为1。</p> <p>说明：描述电子表格、数据库，以及具有结构化数据内容资源中的各数据项（字段）标题，采用中文表示。</p>
11.共享方式	<p>定义：不同共享类型的政务信息资源的共享条件。</p> <p>数据类型：字符型。</p> <p>注解：必选项；最大出现次数为1。</p> <p>说明：无条件共享类和有条件共享类的政务信息资源，应标明使用要求，包括作为行政依据、工作参考，用于数据校核、业务协同等；有条件共享类的政务信息资源，还应注明共享条件和共享范围；对于不予共享类的政务信息资源，应注明相关的法律、行政法规或党中央、国务院政策依据。</p>