|  |  |
| --- | --- |
| ICS  |   |
| CCS  | 点击此处添加CCS号 |

|  |
| --- |
|  4104 |

平顶山地方标准

DB 4104/T XXXX—2024

智慧城市数据中台大数据元数据

2024 - XX - XX发布

2024 - XX - XX实施

平顶山市市场监督管理局  发布

目次

[前言 II](#_Toc154385689)

[1 范围 1](#_Toc154385690)

[2 规范性引用文件 1](#_Toc154385691)

[3 术语和定义 1](#_Toc154385692)

[4 描述方法 2](#_Toc154385693)

[5 核心元数据 3](#_Toc154385694)

[6 技术元数据 6](#_Toc154385695)

[7 业务元数据 10](#_Toc154385696)

[8 管理元数据 14](#_Toc154385697)

1. 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。

本文件由平顶山市政务服务和大数据管理局提出并归口。

本文件起草单位：平顶山市政务服务和大数据管理局、平顶山市卫生健康委员会、平顶山市财政局、平顶山市城市管理局、平顶山市智慧城市运行服务中心、平顶山市宝丰县大数据中心、平顶山市湛河区政务服务和大数据管理局、平顶山市大数据产业发展有限公司、平顶山数字智能产业运营有限公司、广电运通集团股份有限公司。

本文件主要起草人：邵瑞卿、孙愿强、高周、王堃、朱宝、李刚、宋小朵、林军军、张晓力、蔡高尚、张鲁非、王晓光、张鹏辉、张博翔、王广林、陈晓强。

智慧城市数据中台大数据

元数据

* 1. 范围

本文件规定了平顶山市智慧城市数据中台大数据管理中的核心元数据、技术元数据、业务元数据、管理元数据。

本文件适用于指导智慧城市数据中台大数据管理中的元数据管理工作指引。

* 1. 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 18391.3 信息技术 元数据注册系统(MDR) 第3部分：注册系统元模型与基本属性

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 35295-2017 信息技术 大数据 术语

GB/T 37043 智慧城市 术语

* 1. 术语和定义

下列术语和定义适用于本文件。

智慧城市

智慧城市是运用信息通信技术,有效整合各类城市管理系统，实现城市各系统间信息资源共享和业务协同,推动城市管理和服务智慧化,提升城市运行管理和公共服务水平,提高城市居民幸福感和满意度，实现可持续发展的一种创新型城市。

[来源:GB/T 37043—2018]

数据中台

数据中台是建立在大数据技术基础上，用于集中收集、整合、存储、管理、处理和共享政企内外所有数据资源的平台，对各业务单元业务与数据的沉淀，构建包括数据技术、数据治理、数据运营等数据建设、管理、使用体系，实现数据赋能。数据中台是智慧城市框架体系中的核心，为智慧城市提供统一、标准化、高质量的数据服务，推动数字化转型和升级。

大数据

具有体量巨大、来源多样、生成极快、且多变等特征，并且难以用传统数据体系结构有效处理的包含大量数据集的数据。

[来源:GB/T 35295-2017,2.1.1]

元数据

定义和描述其他数据的数据。智慧城市数据中台大数据领域的元数据中，应包括但不限于核心元数据、技术元数据、业务元数据、管理元数据等。

[来源:GB/T 18391.3-2009,3.2.18]

元数据元素

元数据的基本单元。

1. 元数据元素在元数据实体中是唯一的。

[来源:GB/T 21063.3-2007，3.1]

元数据实体

一组说明信息资源相关特性的元数据元素。

1. 可以包含一个或一个以上元数据实体。

[来源:GB/T 21063.3—2007，3.2]

核心元数据

描述数据资源基本属性的元数据元素和元数据实体。对于特定领域或组织而言，具有重要意义和关键价值的元数据。它包含了对于核心业务和数据资产至关重要的信息，用于支持业务流程、数据管理和决策等方面。

技术元数据

描述技术系统、应用程序和数据存储结构的元数据。它包含了对连接名称、连接唯一标识、连接方式、IP地址、元数据类型等的描述，用于支持系统开发、维护和管理等方面。

业务元数据

描述组织业务过程、数据资产和业务规则的元数据。它包含了对所属部门、所属分层、业务描述、表数量、本月新增等方面的描述，用于支持业务理解、数据管理和决策等方面。

管理元数据

描述和记录组织内各种信息资产的元数据。它包括对版本号、采集方式、登录时间、版本更新时间、上一次同步时间等方面的描述和定义。管理元数据提供了对这些信息资产的全面了解，帮助组织有效地管理和利用这些资产。

* 1. 描述方法

采用摘要表示的方式定义和描述元数据，摘要内容包括：

1. 中文名称：元数据元素或元数据实体的中文名称；
2. 定义：描述元数据实体或元数据元素的基本内容，给出公共数据资源某个特性的概念和说明；
3. 英文名称：元数据实体或元数据元素的英文名称，是公共数据资源某个特性的英文名称，应用英文全称，不应包括任何空格、破折号、下划线或分隔符等；
4. 数据类型：说明元数据元素或元数据实体的数据类型，如复合型、字符型、日期型、数值型、布尔型等；
5. 值域：说明元数据元素可以取值的范围；
6. 短名：元数据元素或元数据实体的的英文缩写名称，应遵循以下规则:
	1. 短名应唯一，
	2. 短名不应包括任何空格、破折号、下划线或分隔符等，
	3. 在保持唯一性的前提下，取英文名称中每个英文单词的前三个字母作为其短名缩写标识，
	4. 当取词不能保证唯一性时应延展取词位数，直至保证唯一性为止；
7. 注解：对元数据元素或元数据实体含义的进一步解释,包括元数据元素或元数据实体的约束(必选、可选)和最大出现次数。
	1. 核心元数据
		1. 组成

智慧城市数据中台大数据的核心元数据有6个独立元数据元素和3个元数据实体组成，见表1。

1. 核心元数据元素及元数据实体

| 序号 | 元数据名称 | 信息类别 | 备注 |
| --- | --- | --- | --- |
| 1 | 标识符 | 基本信息 | 必选 |
| 2 | 名称 | 基本信息 | 必选 |
| 3 | 关键字 | 基本信息 | 必选 |
| 4 | 访问限制 | 基本信息 | 必选 |
| 5 | 摘要 | 基本信息 | 必选 |
| 6 | 信息资源链接地址 | 基本信息 | 必选 |
| 7 | 责任者 | 创建者 | 基本信息 | 必选 |
| 发布者 | 基本信息 | 必选 |
| 维护者 | 基本信息 | 必选 |
| 8 | 信息资源分类 | 主题分类 | 基本信息 | 必选 |
| 格式分类 | 基本信息 | 可选 |
| 类型分类 | 基本信息 | 可选 |
| 9 | 日期 | 创建日期 | 基本信息 | 必选 |
| 修改日期 | 基本信息 | 必选 |
| 发布日期 | 基本信息 | 必选 |

* + 1. 描述
			1. 标识符

定义：用于对特定智慧城市数据中台大数据进行唯一标识的一组字符

英文名称：identifier

数据类型：字符型

短名：ide

注解：必选项，最大出现次数为1

* + - 1. 名称

定义：核心元数据的名称

英文名称：name

数据类型：字符型

短名：nam

注解：必选项，最大出现次数为1

* + - 1. 关键字

定义：能够最大程度概括用户所要查找的信息内容的词或短语

英文名称：keyword

数据类型：字符型

短名：key

注解：必选项，最大出现次数为1

* + - 1. 访问限制

定义：对用户访问对网站、数据库或其他资源权限的限定

英文名称：accessrestrictions

数据类型：字符型

短名：accres

注解：必选项，最大出现次数为1

* + - 1. 摘要

定义：对智慧城市数据中台大数据资源内容（或关键字段）的概要描述

英文名称：abstract

数据类型：字符型

短名：abs

注解：必选项，最大出现次数为1

* + - 1. 信息资源链接地址

定义：智慧城市数据中台大数据存在指向特定信息资源的网络链接或URL（统一资源定位符）

英文名称：linkaddress

数据类型：字符型

短名：linadd

注解：必选项，最大出现次数为1

* + - 1. 责任者

定义：对智慧城市数据中台大数据负有直接责任的个人、组织或机构

英文名称：responsible

数据类型：复合型

短名：res

注解：必选项，最大出现次数为1

* + - * 1. 创建者

定义： 智慧城市数据中台大数据元数据的创建者的描述

英文名称：creator

数据类型：字符型

短名：cre

注解：必选项，最大出现次数为1

* + - * 1. 发布者

定义： 智慧城市数据中台大数据元数据的发布者的描述

英文名称：publisher

数据类型：字符型

短名：pub

注解：必选项，最大出现次数为1

* + - * 1. 维护者

定义： 智慧城市数据中台大数据元数据的维护者的描述

英文名称：vindicator

数据类型：字符型

短名：vin

注解：必选项，最大出现次数为1

* + - 1. 信息资源分类

定义： 说明智慧城市数据中台大数据信息资源分类的类目信息

英文名称：resourcecategory

数据类型：复合型

短名：rescat

注解：必选项，最大出现次数为1

* + - * 1. 主题分类

定义： 说明智慧城市数据中台大数据信息资源分类的主题类目信息

英文名称：theme

数据类型：字符型

短名：the

注解：必选项，最大出现次数为1

* + - * 1. 格式分类

定义： 说明智慧城市数据中台大数据信息资源分类的格式类目信息

英文名称：format

数据类型：字符型

短名：for

注解：可选项，最大出现次数为1

* + - * 1. 类型分类

定义： 说明智慧城市数据中台大数据信息资源分类的类型类目信息

英文名称：type

数据类型：字符型

短名：typ

注解：可选项，最大出现次数为1

* + - 1. 日期

定义： 智慧城市数据中台大数据提供时间

英文名称：date

数据类型：复合型

短名：dat

注解：必选项，最大出现次数为1

* + - * 1. 创建日期

定义： 智慧城市数据中台大数据提供创建时间

英文名称：found

数据类型：复合型

短名：fou

注解：必选项，最大出现次数为1

* + - * 1. 修改日期

定义： 智慧城市数据中台大数据提供修改时间

英文名称：edit

数据类型：复合型

短名：edi

注解：必选项，最大出现次数为1

* + - * 1. 发布日期

定义： 智慧城市数据中台大数据提供发布时间

英文名称：publish

数据类型：复合型

短名：pub

注解：必选项，最大出现次数为1

* 1. 技术元数据
		1. 数据源技术元数据
			1. 组成

智慧城市数据中台大数据的数据源技术元数据有8个独立元数据元素和1个元数据实体组成，见表2。

1. 数据源技术元数据元素及元数据实体

| 序号 | 元数据名称 | 信息类别 | 备注 |
| --- | --- | --- | --- |
| 1 | 数据源名称 | 基本信息 | 必选 |
| 2 | 唯一标识 | 基本信息 | 必选 |
| 3 | 数据库类型 | 基本信息 | 必选 |
| 4 | 连接方式 | JDBC连接 | 基本信息 | 必选 |
| ODBC连接 | 基本信息 | 必选 |
| ADO.NET连接 | 基本信息 | 必选 |
| 其它连接 | 基本信息 | 必选 |
| 5 | IP地址 | 基本信息 | 必选 |
| 6 | 数据库名称 | 基本信息 | 必选 |
| 7 | 数据库版本 | 基本信息 | 必选 |
| 8 | 端口号 | 基本信息 | 必选 |
| 9 | 用户名 | 基本信息 | 必选 |

* + - 1. 描述
				1. 数据源名称

定义： 数据源技术元数据来源的名称

英文名称：datasourcename

数据类型：字符型

短名：datsounam

注解：必选项，最大出现次数为1

* + - * 1. 唯一标识

定义： 数据源技术元数据的唯一标识符

英文名称：uniqueidentifier

数据类型：字符型

短名：uniide

注解：必选项，最大出现次数为1

* + - * 1. 数据库类型

定义： 存储数据源技术元数据的数据库类型

英文名称：databasetype

数据类型：字符型

短名：dattyp

注解：必选项，最大出现次数为1

* + - * 1. 连接方式

定义： 数据源技术元数据的连接方式

英文名称：linktype

数据类型：复合型

短名：lintyp

注解：必选项，最大出现次数为1

JDBC连接

定义： 数据源技术元数据的JDBC连接方式

英文名称：jdbcconnection

数据类型： 字符型

短名：jdbccon

注解：必选项，最大出现次数为1

ODBC连接

定义： 数据源技术元数据的ODBC连接方式

英文名称：odbcconnection

数据类型： 字符型

短名：odbccon

注解：必选项，最大出现次数为1

ADO.NET连接

定义： 数据源技术元数据的ADO.NET连接方式

英文名称：adonetconnection

数据类型： 字符型

短名：adocon

注解：必选项，最大出现次数为1

其它连接

定义： 数据源技术元数据的其它连接方式

英文名称：otherconnnection

数据类型： 字符型

短名：othcon

注解：必选项，最大出现次数为1

* + - * 1. IP地址

定义： 可以访问数据源技术元数据信息的信息

英文名称：IPaddress

数据类型：数字型

短名：IPadd

注解：必选项，最大出现次数为1

* + - * 1. 数据库名称

定义： 数据源技术元数据的数据库名称

英文名称：databasename

数据类型：字符型

短名：datnam

注解：必选项，最大出现次数为1

* + - * 1. 数据库版本

定义： 数据源技术元数据的数据库版本

英文名称：databaseversion

数据类型：数字型

短名：datver

注解：必选项，最大出现次数为1

* + - * 1. 端口号

定义： 数据源技术元数据传输数据的通道

英文名称：port

数据类型：数字型

短名：por

注解：必选项，最大出现次数为1

* + - * 1. 用户名

定义：使用数据源技术元数据的用户名称

英文名称：username

数据类型：字符型

短名：usenam

注解：必选项，最大出现次数为1

* + 1. 数据表技术元数据
			1. 组成

智慧城市数据中台大数据的数据表技术元数据有4个独立元数据元素组成，见表3。

1. 数据表技术元数据元素及元数据实体

| 序号 | 元数据名称 | 信息类别 | 备注 |
| --- | --- | --- | --- |
| 1 | 数据源连接名称 | 基本信息 | 必选 |
| 2 | 数据源连接唯一标识 | 基本信息 | 必选 |
| 3 | 数据库名称 | 基本信息 | 必选 |
| 4 | 数据源IP地址 | 基本信息 | 必选 |

* + - 1. 描述
				1. 数据源连接名称

定义： 数据表技术元数据来源连接的名称

英文名称：datasourcelinkname

数据类型：字符型

短名：datsoulinnam

注解：必选项，最大出现次数为1

* + - * 1. 数据源连接唯一标识

定义： 数据表技术元数据连接的唯一标识符

英文名称：datasourcelinkuniqueidentifier

数据类型：字符型

短名：datsoulinuniide

注解：必选项，最大出现次数为1

* + - * 1. 数据库名称

定义： 数据表技术元数据的数据库名称

英文名称：databasename

数据类型：字符型

短名：datnam

注解：必选项，最大出现次数为1

* + - * 1. 数据源IP地址

定义： 数据表技术元数据的连接方式

英文名称：datasourceIPaddress

数据类型：数字型

短名：datsouIPadd

注解：必选项，最大出现次数为1

* 1. 业务元数据
		1. 数据源业务元数据
			1. 组成

智慧城市数据中台大数据的数据源业务元数据有4个独立元数据元素组成，见表4。

1. 数据源业务元数据元素

| 序号 | 元数据名称 | 信息类别 | 备注 |
| --- | --- | --- | --- |
| 1 | 所属部门 | 基本信息 | 必选 |
| 2 | 所属分层 | 基本信息 | 必选 |
| 3 | 业务描述 | 基本信息 | 必选 |
| 4 | 表数量 | 基本信息 | 必选 |

* + - 1. 描述
				1. 所属部门

定义： 数据源业务元数据所属的部门

英文名称：affiliation

数据类型： 字符型

短名：aff

注解：必选项，最大出现次数为1

* + - * 1. 所属分层

定义： 数据源业务元数据所属的分层

英文名称：hierarchy

数据类型：字符型

短名：hie

注解：必选项，最大出现次数为1

* + - * 1. 业务描述

定义： 数据源业务的描述

英文名称：businessdescription

数据类型： 字符型

短名：busdes

注解：必选项，最大出现次数为1

* + - * 1. 表数量

定义： 数据源业务元数据所用表的数量

英文名称：tablenumbers

数据类型： 数字型

短名：tabnum

注解：必选项，最大出现次数为1

* + 1. 数据表业务元数据
			1. 组成

智慧城市数据中台大数据的数据表业务元数据有13个独立元数据元素组成，见表5。

1. 数据表业务元数据元素

| 序号 | 元数据名称 | 信息类别 | 备注 |
| --- | --- | --- | --- |
| 1 | 数据表名称 | 基本信息 | 必选 |
| 2 | 数据表中文名称 | 基本信息 | 必选 |
| 3 | 字段数量 | 基本信息 | 必选 |
| 4 | 数据总量 | 基本信息 | 必选 |
| 5 | 本年增量 | 基本信息 | 必选 |
| 6 | 本月增量 | 基本信息 | 必选 |
| 7 | 昨日增量 | 基本信息 | 必选 |
| 8 | 表数据更新时间 | 基本信息 | 必选 |
| 9 | 所属数据分层 | 基本信息 | 必选 |
| 10 | 所属部门 | 基本信息 | 必选 |
| 11 | 描述 | 基本信息 | 必选 |
| 12 | 增量字段 | 基本信息 | 必选 |
| 13 | 业务意义 | 基本信息 | 可选 |

* + - 1. 描述
				1. 数据表名称

定义： 数据表业务元数据的名称

英文名称：datatablename

数据类型：字符型

短名：dattabnam

注解：必选项，最大出现次数为1

* + - * 1. 数据表中文名称

定义： 数据表业务元数据的中文名称

英文名称：datatableChinesename

数据类型：字符型

短名：datatabChinam

注解：必选项，最大出现次数为1

* + - * 1. 字段数量

定义： 与对象或类关联的变量数量

英文名称： fieldcount

数据类型：数字型

短名：fiecou

注解：必选项，最大出现次数为1

* + - * 1. 数据总量

定义： 数据表业务元数据的总量

英文名称：totaldata

数据类型：数字型

短名：totdat

注解：必选项，最大出现次数为1

* + - * 1. 本年增量

定义： 数据表业务元数据的本年增量

英文名称：annualincrement

数据类型：数字型

短名：anninc

注解：必选项，最大出现次数为1

* + - * 1. 本月增量

定义： 数据表业务元数据的本月增量

英文名称：monthlyincrement

数据类型：数字型

短名：moninc

注解：必选项，最大出现次数为1

* + - * 1. 昨日增量

定义： 数据表业务元数据的昨日增量

英文名称：yesterday'sincrement

数据类型：数字型

短名：yesinc

注解：必选项，最大出现次数为1

* + - * 1. 表数据更新时间

定义： 表内数据的更新时间

英文名称：updatetime

数据类型：日期型

值域：YYYY-MM-DD hh:mm:ss

短名：updtim

注解：必选项，最大出现次数为1

* + - * 1. 所属数据分层

定义： 数据表业务元数据所属数据分层

英文名称：layerofbelongdata

数据类型：字符型

短名：layofbeldat

注解：必选项，最大出现次数为1

* + - * 1. 所属部门

定义： 数据表业务元数据所属部门

英文名称：department

数据类型：字符型

短名：dep

注解：必选项，最大出现次数为1

* + - * 1. 描述

定义： 对数据表业务信息的描述

英文名称：describe

数据类型：字符型

短名：des

注解：必选项，最大出现次数为1

* + - * 1. 增量字段

定义： 数据表业务元数据用于记录数据的变化字段

英文名称：incrementalfield

数据类型：字符型

短名：incfie

注解：必选项，最大出现次数为1

* + - * 1. 业务意义

定义： 业务的意义

英文名称：businesssignificance

数据类型：字符型

短名：bussig

注解：可选项；最大出现次数为1

* 1. 管理元数据

8.2 数据源管理元数据

* + - 1. 组成

智慧城市数据中台大数据的数据源管理元数据有2个独立元数据元素及1个元数据实体组成，见表6。

1. 数据源管理元数据元素

| 序号 | 元数据名称 | 信息类别 | 备注 |
| --- | --- | --- | --- |
| 1 | 登记时间 | 基本信息 | 必选 |
| 2 | 更新时间 | 基本信息 | 必选 |
| 3 | 采集方式 | 自动采集 | 基本信息 | 必选 |
| 手动采集 | 基本信息 | 必选 |

* + - 1. 描述
				1. 登记时间

定义： 数据源管理元数据的登记时间

英文名称：registrationtime

数据类型：时间型

值域：YYYY-MM-DD hh:mm:ss

短名：regtim

注解：必选项，最大出现次数为1

* + - * 1. 更新时间

定义： 数据源管理元数据的更新时间

英文名称：updatetime

数据类型：时间型

值域：YYYY-MM-DD hh:mm:ss

短名：updatim

注解：必选项，最大出现次数为1

* + - * 1. 采集方式

定义：数据源管理元数据的采集方式

英文名称：acquisitionmethod

数据类型：复合型

短名：acqmet

注解：必选项，最大出现次数为1

自动采集

定义： 数据源管理元数据的自动采集方式

英文名称：automaticacquisition

数据类型：字符型

短名：autcacq

注解：必选项，最大出现次数为1

手动采集

定义： 数据源管理元数据的手动采集方式

英文名称：manualacquisition

数据类型：字符型

短名：manacq

注解：必选项，最大出现次数为1

8.3 数据表管理元数据

* + - 1. 组成

智慧城市数据中台大数据的数据表管理元数据有3个独立元数据元素及1个元数据实体组成，见表7。

1. 数据表管理元数据元素

| 序号 | 元数据名称 | 信息类别 | 备注 |
| --- | --- | --- | --- |
| 1 | 最新版本号 | 基本信息 | 必选 |
| 2 | 版本更新时间 | 基本信息 | 必选 |
| 3 | 上一次同步时间 | 基本信息 | 必选 |
| 4 | 采集方式 | 自动采集 | 基本信息 | 必选 |
| 手动采集 | 基本信息 | 必选 |

* + - 1. 描述
				1. 最新版本号

定义： 数据表管理元数据的最新版本号

英文名称：latestversion

数据类型：数字型

短名：latver

注解：必选项，最大出现次数为1

* + - * 1. 版本更新时间

定义： 数据表管理元数据的最新版本更新时间

英文名称：versionupdatetime

数据类型：时间型

值域：YYYY-MM-DD hh:mm:ss

短名：verupdatim

注解：必选项，最大出现次数为1

* + - * 1. 上一次同步时间

定义： 数据表管理元数据上一次进行数据库同步的时间

英文名称：lastsynchronizationtime

数据类型：时间型

值域：YYYY-MM-DD hh:mm:ss

短名：lassyntim

注解：必选项，最大出现次数为1

* + - * 1. 采集方式

定义： 数据表管理元数据的采集方式

英文名称：acquisitionmethod

数据类型：复合型

短名：acqumet

注解：必选项，最大出现次数为1

自动采集

定义： 数据源管理元数据的自动采集方式

英文名称：automaticacquisition

数据类型：字符型

短名：autcacq

注解：必选项，最大出现次数为1

手动采集

定义： 数据源管理元数据的手动采集方式

英文名称：manualacquisition

数据类型：字符型

短名：manacq

注解：必选项，最大出现次数为1

