

团体标准

T/BSIA 001—2023

智慧城市数据中台建设规范

Code for the construction of smart city data platform

2023 - 04 - 06 发布

2023 - 04 - 07 实施

目 次

前言	II
引言	III
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 缩略语	1
5 基本原则与体系架构	2
5.1 基本原则	2
5.2 体系架构	2
6 功能要求	3
6.1 数据集成平台	3
6.2 数据开发平台	4
6.3 数据治理平台	6
6.4 数据分析平台	10
6.5 资产管理平台	10
6.6 数据开放平台	11
6.7 运维监控平台	12
7 数据安全平台	13
8 运维要求	13
8.1 人员要求	13
8.2 运维内容要求	13
9 一般建设流程	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。

本文件由北京软件和信息服务业协会提出并归口。

本文件起草单位：云粒智慧科技有限公司、北京豪尔赛智慧城域科技有限公司、北京企点时代科技有限公司、北京嘉诚瑞杰信息技术有限公司、中国航空结算有限责任公司、北京梦天门科技股份有限公司、联动优势科技有限公司、国信优易数据股份有限公司、数翊科技（北京）有限公司、北京柏睿数据技术股份有限公司、北京滴普科技有限公司、北京智网易联科技有限公司、北京中奥淘数据科技有限公司、无明智囊（北京）科技有限公司、北京赛目科技股份有限公司、亚信科技（中国）有限公司、北京国研网信息股份有限公司、北京领雁科技股份有限公司、博彦科技股份有限公司、北京猎豹移动科技有限公司、北京惠钰博科技有限公司。

本文件主要起草人：殷伟、柳宁、吴海燕、郑扬勇、戴聪棋、刘兵、李少杰、马小玲、白戈、王尧、丁学利、程晨、刘胜、余柯、田敬莲、王莲、吴学军、李强、王彬、薛晓卿、郭建超、张静燕、李琼、李超、傅盛、郭濮铭。

本文件为首次发布。

引 言

本文件的发布机构提请注意，声明符合本文件时，可能涉及到本文件[5.1条]与[一种视频监控集群的控制方法、装置、设备及存储介质]、本文件[5.2条]与[一种多关系型数据引擎的SQL合并执行系统及方法]、本文件[5.3条]与[跨集群数据处理方法、装置、电子设备以及存储介质]、本文件[5.4条]与[一种异常血缘关系的确定方法、装置、设备及介质]相关的专利的使用。

本文件的发布机构对于该专利的真实性、有效性和范围无任何立场。

该专利持有人已向本文件的发布机构承诺，他愿意同任何申请人在合理且无歧视的条款和条件下，就专利授权许可进行谈判。该专利持有人的声明已在本文件的发布机构备案。相关信息可以通过以下联系方式获得：

专利持有人姓名：云粒智慧科技有限公司。

地址：北京市西城区阜成门外大街甲28号西楼13层06室。

请注意除上述专利外，本文件的某些内容仍可能涉及专利。本文件的发布机构不承担识别专利的责任。

智慧城市数据中台建设规范

1 范围

本文件规定了智慧城市数据中台建设的体系架构、功能要求、数据安全、运维要求、一般建设流程等内容。

本文件适用于地市级、区县级智慧城市数据中台整体规划及具体项目的规划、设计、建设与运维。

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 25069 信息安全技术 术语
- GB/T 34960.5 信息技术服务 治理 第5部分：数据治理规范
- GB/T 37043 智慧城市 术语
- GB/T 37973 信息安全技术 大数据安全管理指南
- GB/Z 38649 信息安全技术 智慧城市建设信息安全保障指南
- GB/T 39477 信息安全技术 政务信息共享 数据安全技术要求

3 术语和定义

GB/T 25069、GB/T 37043界定的以及下列术语和定义适用于本文件。

3.1

智慧城市数据中台 smart city data platform

智慧城市数据中台是一套可持续的让数据用起来机制，一种战略选择和组织形式，是依据特有的业务模式和组织架构，通过有形的产品和实施方法论支撑，构建一套持续不断把数据变成资产并服务于业务的机制体系。

注：数据中台具备多源异构数据统一存储、计算的能力，同时让分散杂乱的数据通过规范化的方式管理起来。数据中台借鉴了传统数仓面向主题域的数据组织方式，基于维度建模理论，构建统一的数据公共层和应用层。数据中台依赖于大数据平台完成数据研发全流程，同时增加了数据治理和数据服务化以及数据资产内容，是数据要素服务的基础能力平台。

3.2

数据血缘 data pedigree

是指数据的全生命周期中，数据从产生、处理、加工、融合、流转到最终消亡，数据之间自然形成一种关系，其记录了数据产生的链路关系，可用于数据溯源、数据质量评估、价值评估等。

4 缩略语

下列缩略语适用于本文件。

API：应用程序编程接口（Application Programming Interface）

CSV：一种通用的文件格式，可以非常容易被导入各种PC表格及数据库中（Comma-Separated Values）

DWD：细节数据层（data warehouse detail）

DWS：服务数据层（data warehouse service）

SQL：具有数据操纵和数据定义等多种功能的数据库语言（Structured Query Language）

SSH：专为远程登录会话和其他网络服务提供安全性的协议（Secure Shell）

VM：虚拟机主机（Virtual Manufacturing）

5 基本原则与体系架构

5.1 基本原则

智慧城市数据中台建设宜遵循以下基本原则：

- a) 可靠性：数据中台作为把数据变成资产并服务于业务的机制，其提供的数据应准确、可靠、不应出错；
- b) 可追溯性：对数据进行标签化管理，一旦发生数据错误能够快速定位错误来源，并能明确错误影响范围；
- c) 可扩展性：城市信息化建设的加快，使得数据种类越来越多，数据量越来越大；智慧城市数据中台应具备可扩展性，以满足数据增长对数据中台的需求，同时不应对原有业务造成影响以及避免资源浪费。

5.2 体系架构

智慧城市以连接、算法、分析驱动数据要素流转，深刻重塑城市经济生产组织方式和社会活动运行流程，推动要素更加高效配置，带动城市产业、治理服务全面升级，是促进管理和服务智慧化的新理念和新模式。

智慧城市数据中台体系架构应包括数据集成平台、数据开发平台、数据治理平台、资产管理平台、数据开放平台、数据分析平台、数据安全平台与运维监控平台等。

数据集成平台负责实现多源异构数据的汇聚；数据开发平台负责汇聚的全结构化、实时、离线数据的融合建设；数据治理平台依托相关标准规范对数据质量进行稽核与数据标准化建设；数据资产管理平台是针对治理后的数据进行资产化的分类、展现与发布；数据开放平台是对外通过数据开放共享的能力；数据分析平台结合算法、知识图谱、报表分析等能力提供行业数据分析应用；数据安全与运维平台保障整个中台建设的过程中数据全过程的安全管控，系统全链路运维监控，保障系统安全正常运行。智慧城市数据中台体系架构见图1。

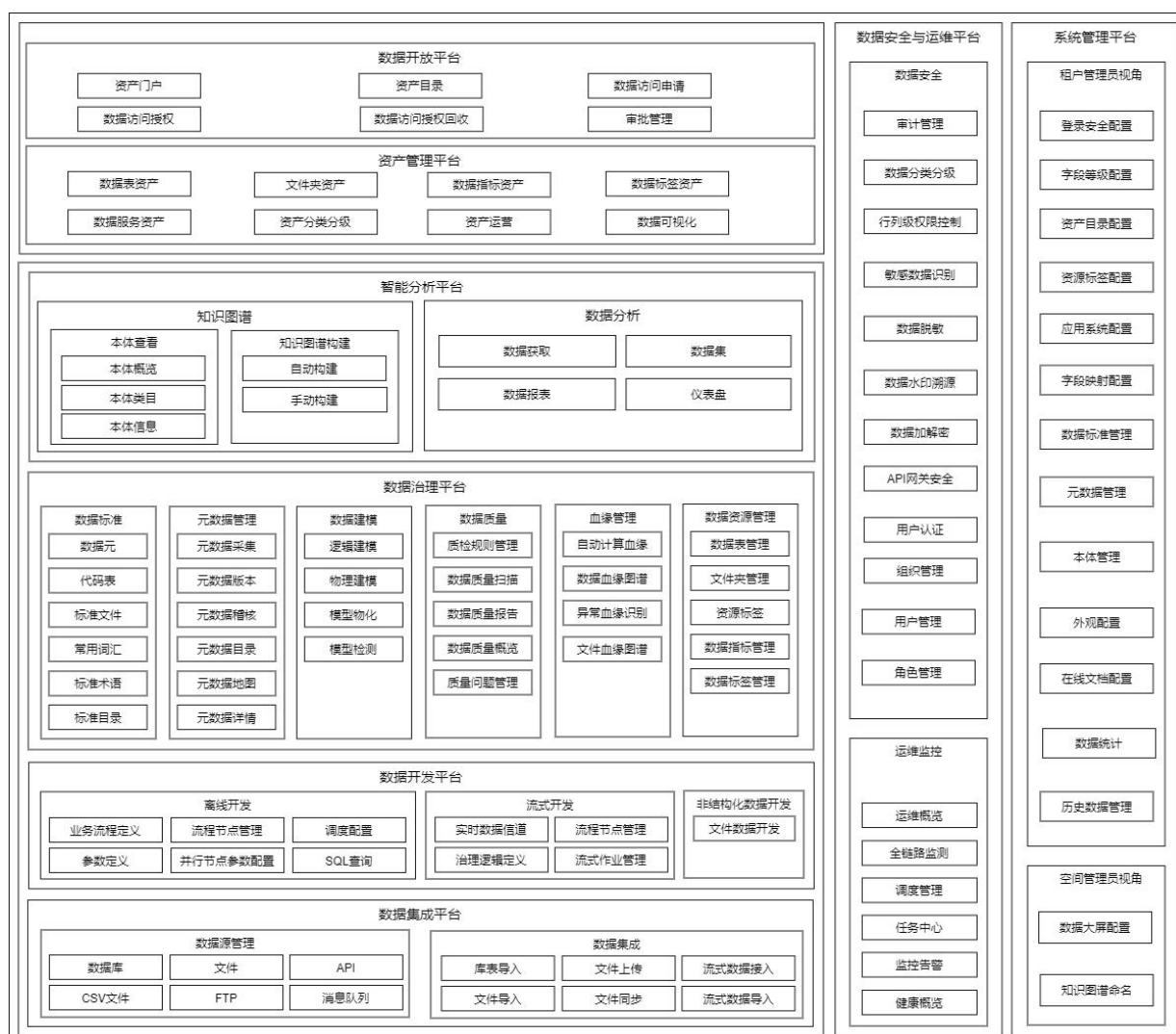


图1 智慧城市数据中台体系架构

6 功能要求

6.1 数据集成平台

6.1.1 概述

数据集成是指利用数据抽取、消息服务、数据交换、填报采集、数据接口等技术手段，把各种异构网络、异构数据源、不同类型的数据采集到数据中台进行集中存储。数据集成是数据中台数据接入的入口，数据集成的方式有结构化数据接入、非结构化数据接入、API数据接入、流式数据接入、数据源接入、实时信道数据接入等。

6.1.2 数据源管理

智慧城市数据中台数据库连接配置应支持手动配置MySQL、SQLServer、PostgreSQL、Oracle、DM（达梦）、kingbaseES（金仓）、Highgo（瀚高）等数据库的数据源驱动，当默认驱动不适用时，可手动新增驱动，保证数据源连接正常。。

6.1.2.1 数据库

数据库连接配置应支持MySQL、SQLServer、PostgreSQL、Oracle、DM（达梦）、kingbaseES（金仓）、Highgo（瀚高）等。数据源连接信息的配置管理，包括增加、删除、修改、查看、测试连通性，可查看每一数据源关联的业务流程。

6.1.2.2 CSV 文件

CSV文件可作为数据集成的数据源进行数据同步。

6.1.2.3 文件

文件即非结构化数据，包括pdf文件、word文件、图片等，均可作为非结构化数据集成的数据源进行数据同步。

6.1.2.4 FTP

支持FTP文件服务器的连接配置，对文件服务器的管理维护包括增加、删除、修改、查看、测试连通性。可查看该数据源关联的业务流程明细。

6.1.2.5 API

支持对API数据源的配置管理（协议、URL、报文格式等），对API管理维护包括增加、删除、修改、查看、测试连通性。可查看该数据源关联的作业明细。

6.1.2.6 消息队列

消息队列（实时数据信道）是管理流式数据的载体，作为数据中台内部数据源进行管理维护，包括对Topic的增加、删除、修改、查看。

6.1.3 数据集成

6.1.3.1 库表导入

数据开发者通过定义库表导入作业完成结构化数据导入中台。库表导入为周期性作业，宜支持通过选择来源数据库的表，将数据同步至目标数据表。

6.1.3.2 CSV 文件导入

数据开发者通过定义文件导入作业完成数据入中台。CSV文件导入为一次性作业，应支持通过上传CSV、Excel文件，将数据导入至目标数据表。

6.1.3.3 文件上传

文件上传应用场景为：很多政府工作报告、公文、图片、PPT等文件需集成进数据中台，以作为资产进行共享。通过定义文件上传作业完成数据入中台，文件上传为一次性作业，支持通过上传文件夹或批量上传多个文件将文件数据同步至目标文件夹。

6.1.3.4 文件同步

文件同步应用场景为：本地FTP服务器的对应文件目录下有较多政府工作报告、公文、图片、PPT等文件，并会定期做更新，需要周期性同步文件进中台。

6.1.3.5 流式数据接入

数据开发者通过定义流式数据接入作业完成流式数据离线导入中台。流式数据接入为周期性作业，支持通过选择来源实时数据信道将数据同步至目标表。

6.1.3.6 流式数据导入

流式数据导入应用场景为：IOT设备流式数据经解析转换后批量入仓，入仓后可基于实际业务需求，开展对数仓数据的开发治理。

6.2 数据开发平台

6.2.1 概述

数据开发是一套数据加工并对加工过程管控的工具，数据开发人员可以利用数据开发模块提供的功能，把数据加工成对业务有价值的形式并供给业务使用。

智慧城市数据中台数据开发模块应面向不同的数据处理场景完成数据开发，实现离线数据、实时数据的快速融合处理，结合原生的SQL探索能力，提高数据开发效率并降低数据开发成本。

6.2.2 离线开发

6.2.2.1 业务流程定义

业务流程指针对业务实体，抽象出业务流程的概念，从业务视角组织数据的开发治理，提高任务管理效率。

6.2.2.2 流程节点管理

流程节点管理包括库表导入、库表导出、流式数据导入、SQL加工、数据汇聚、参数提取、流程依赖。

6.2.2.3 调度配置

定义整个业务流程的执行频率，如分钟、小时、日、周、月，生成周期性调度作业。支持对多个业务流程批量设置调度时间，以完成批量流程调度。

6.2.2.4 参数定义

定义流程参数，供业务流程中各节点针对数据表分区数据导入来使用。新增一组参数，包含参数名称、参数值或表达式。

6.2.2.5 并行节点配置

支持参数并行节点数（可并行运行的最大节点数量）。

6.2.2.6 SQL 查询

基于SQL语言开发模型是数仓数据治理的核心，模型开发的过程中，数据建模人员需要先探查数仓中有哪些数据再进行建模。同时，当有些数据导入数仓后，发现有误，数据开发者无需配置业务流程，即可快速完成数据修正。

SQL查询应可支持查询数据、插入数据、清空表、添加/删除/查询分区。

6.2.3 流式开发

流式开发基于实时流计算引擎，对无边界数据集进行连续不断的处理、聚合和分析，获取结果具备低延迟性，可有效控制在毫秒或秒级别。

流式数据开发应支持基于模型、规则，计算生成事件，并将处理结果输出到指定目标源。

6.2.3.1 实时数据信道

实时数据信道只是管理流式数据的载体，作为数据中台内部数据源进行管理维护，包括对Topic的增加、删除、修改、查看。

6.2.3.2 流式作业管理

支持按作业名称、状态查询流式数据开发作业。状态为“未上线”的流式数据开发作业可查看、修改、删除、上线、查看任务；状态为“运行中”的流式数据开发作业可查看、下线、查看任务。

6.2.3.3 治理逻辑定义

流式数据处理逻辑的开发需要使用JAVA语言，依赖中台提供的流式数据治理SDK，可结合实际的业务场景处理数据流转中的逻辑，最后打包成一个JAR文件。

6.2.3.4 流程节点管理

通过不同类型的节点，如数据来源节点、SQL加工节点、数据去向节点，定义节点之间依赖关系。

6.2.4 非结构化数据开发

非结构化数据开发宜支持以可视化方式呈现非结构化数据治理的全链路，便于文件数据的问题定位与溯源管理。

6.3 数据治理平台

6.3.1 概述

GB/T 34960.5将数据治理框架分为顶层设计、数据治理环境、数据治理域和数据治理过程四大部分。本章主要针对数据治理域、数据治理过程进行要求。

数据治理是从根本上建立数据标准和规范，统一数据资产管理，统一数据调度，保证数据可控、可用、可信的工程；同时，以数据运营的方式，从政务数据和社会数据中提升社会效益和经济效益，为数字政府和智慧城市的建设提供可信、可用的数据支撑。数据治理指数据中台依托标准规范体系，对归集的数据进行清洗整合、主题建模、指标定义、标签生产、质量检核、血缘计算等操作，以提升数据质量、挖掘数据价值。

数据治理的主要内容包括数据标准、元数据管理、数据质量、血缘管理、数据建模、数据资源管理。

6.3.2 数据标准

6.3.2.1 数据元

创建数据表时可已将已维护的元数据作为可选字段。通过定义数据表的元数据，规范数据表结构，进一步规范数据资产。可查看数据元历史版本并支持相邻版本之间的数据元对比，便于用户了解数据变化差异。

6.3.2.2 代码表

支持批量创建，配置后空间用户可见，代码表作为数据元的属性扩展，可被数据元引用。

6.3.2.3 标准文件

标准文件统一纳管相关国家标准、行业标准，以供资源目录做引用，实现数据治理有据可依。

6.3.2.4 常用词汇

主要是指数据中心通用的一些业务概念定义，如名词或名词短语，是对一种共识的定义。支持常用词汇的增加、删除、修改、查看。

6.3.2.5 标准术语

标准术语为元数据提供业务元数据进行支撑，帮助用户对数据进行更好地理解。可将标准术语直接引用为数据元，间接作为数据标准规范被数据表引用。

6.3.2.6 标准目录

宜提供标准目录管理功能，用于对数据元进行目录编排，实现数据元的分类管理。

6.3.3 元数据管理

6.3.3.1 元数据采集

元数据采集宜支持自动扫描数据中台各数据源中的数据表结构信息，并识别表结构变更信息，完成数据源元数据信息的采集汇聚。

6.3.3.2 元数据版本

元数据版本宜支持查询、以列表形式展示元数据版本信息，每个来源系统对应列表中的一条版本信息、查看该采集作业对应的元数据历史版本。

6.3.3.3 元数据稽核

在元数据采集过程中，支持元数据稽核。宜提供静态说明‘元数据稽核开启后，若采集的元数据相比上一版本发生变化，则自动对元数据字段是否为中文名、表名是否为中文名进行稽核，同时用数据元标准匹配做稽核（仅数据元标准中的数据元英文名与数据类型），否则不做稽核’。

6.3.3.4 元数据类目

元数据采集完成后，宜自动生成元数据类目。

6.3.3.5 元数据详情

可获取元数据详细信息，包括：数据表基础信息、数据预览、业务流程、文件导入、元数据质量、数据血缘、共享次数、参考依据、数据总量、存储容量、更新时间。

6.3.3.6 元数据地图

元数据地图宜有序视图呈现系统所有已采集元数据与数据中台数据仓的所有数据表，进行信息整合，支持展示元数据的全域信息。包括但不限于元数据列表、数据表分布、数据量分布。

元数据地图宜根据元数据采集作业自动生成元数据目录，目录名称宜为被采集元数据的数据源名称。每个数据表宜可显示以下内容：

- 数据表中英文名；
- 共享次数；
- 参考依据；
- 字段总量；
- 数据总量；
- 存储容量；
- 更新时间。

6.3.4 数据质量

6.3.4.1 质量概览

智慧城市数据中台应支持数据提供质量概览，便于从全局了解组织数据质量全貌，具体内容如下：

——重点关注：

- 数据表总数；
- 已质检数据表数；
- 质检覆盖率；
- 质检不合格表数；
- 质检不合格率。

——规则应用趋势：

统计数据表总数、质检数据表数、质检不合格表数的变化趋势，直观呈现质检工作的覆盖情况以及数据质量的变化趋势。

——部门质检覆盖率：

对各部门数据的质检覆盖率进行统计，以督促质检覆盖率低的部门开展数据质检工作。

——部门质检不合格率：

对各部门数据的质检不合格率进行统计，以督促质检不合格率高的部门提升数据质量。

6.3.4.2 规则管理

6.3.4.2.1 概述

6.3.4.2.1.1 质量规则是数据质量管理的基础。规则管理包括：告警设置、分区表达式配置、质量规则配置、调度配置。

6.3.4.2.1.2 质量管理列表应能展示中台所有的数据表，以对数据表配置质量规则。质量管理列表包括但不限于以下内容：

- 数据表英文名；
- 数据表中文名；

- 责任人；
- 规则更新时间；
- 有效质量规则数；
- 校验结果；
- 最新校验时间；
- 配置规则；
- 质量报告；
- 其它。

6.3.4.2.2 告警设置

如果质量规则检验的结果出现异常，应有质量告警信息通知相关人员。

6.3.4.2.3 质量规则配置

6.3.4.2.3.1 新建质量规则

新建质量规则应支持内置规则与自定义规则两种方式，具体如下：

- 内置规则：
 - 规则名称；
 - 规则类型；
 - 对象类型；
 - 分区表达式；
 - 比较方式；
 - 期望值；
 - 告警等级；
 - 内置规则类型。
- 自定义规则：
 - 规则名称；
 - 规则类型；
 - 对象类型；
 - 分区表达式；
 - 采样方式；
 - 过滤条件；
 - 比较方式；
 - 期望值；
 - 告警等级。

6.3.4.2.3.2 编辑质量规则

针对质量规则不同的状态，应可支持如下操作：

- 查看任务；
- 修改规则；
- 删除规则；
- 试运行；
- 上线；
- 下线；
- 调度配置。

6.3.4.2.4 调度配置

显示生成该数据表的业务流程的调度信息，作为定义质检调度周期的参考。

6.3.4.3 数据扫描

数据扫描是按既定质检规则对数据进行扫描质检。支持周期调度触发质检。调度周期指数据表质检执行的频率，如分钟、小时、日、周、月。支持对同一数据表的不同质检对象定义按不同的调度周期进行数据扫描。

6.3.4.4 质量报告

智慧城市数据中台应支持生成数据质量报告，数据质量报告应由以下三部分构成：

- 数据表基础信息；
- 质量分析结果；
- 报告详情。

6.3.4.5 问题管理

宜提供质量问题的统一管理，包含结果和问题统计、结果查询、结果和问题明细查询、问题溯源。

6.3.5 数据建模

可新建主题域（逻辑模型分类），主题域支持分级，其下可关联不同逻辑模型，每个逻辑模型中可包含多个逻辑实体（以及逻辑实体与逻辑实体之间的关系）；

支持新建、修改、发布、下架物理模型类目、关联物理模型；

支持发布后直接在数仓中创建并同步关系建模的业务表以及维度建模的事实表、维度表、汇总表；

6.3.6 血缘管理

6.3.6.1 自动计算血缘

智慧城市数据中台应能自动计算形成全局的数据血缘图谱，血缘图谱旨在形象、直观地将数据血缘关系进行展示，方便数据开发者探索数据问题本质，及时定位问题，辅助数据治理。

6.3.6.2 数据血缘图谱

血缘图谱提供分层展示数据的能力，以带箭头的连线体现数据关联关系，箭头指向下游数据。血缘图谱可覆盖数据表的源头，记录数据表的所属数据源、源头与目标表映射关系以及数据量。

6.3.6.3 异常血缘识别

一个数据源头的核心业务的改动，下游各分析应用应保持同步。当系统进行升级改造时，如果修改了数据结构、ETL程序等元数据信息，依赖数据血缘的影响分析，可以快速定位出元数据修改会影响到哪些下游数据，从而减少系统升级改造带来的风险。

6.3.6.4 文件血缘图谱

以桑基图形式展示文件血缘上下游链路关系，各个节点代表不同文件，在相对位置上，左侧为上游，右侧为下游。从左至右从来源系统开始，经过各层级文件治理，最后到达应用系统，形成完整链路，体现文件治理过程与价值。

6.3.7 数据资源管理

6.3.7.1 数据表管理

智慧城市数据中台资产管理应支持对数据表进行管理，包括数据表增加、删除、修改、发布、变更、下架、关联类目、文件导入、查询、查看等操作。

6.3.7.2 文件夹管理

智慧城市数据中台资产管理应支持对数据文件夹进行管理，包括文件夹增加、删除、查找、修改、变更、发布、下架、关联类目的管理。

6.3.7.3 资源标签

智慧城市数据中台资产管理应支持资源标签的配置，资源标签通过对标签类型和标签进行维护管理，以满足不同项目在创建资源所需属性字段不同的需求。

6.3.7.4 数据指标管理

数据指标是业务运营过程中，按照既定业务逻辑对数据进行加工处理，使之成为描述、衡量、分析、预测业务结果的基础原材料。由数据定时加工和数据获取接口两部分构成，具体如下：

- 数据定时加工负责周期性的将启用的派生指标数据从 DWD 表加工到 DWS 表；
- 数据获取接口负责通过接口方式直接输出结果集提供给应用或其它产品使用。

6.3.7.5 数据标签管理

数据标签指基于标签加工规则对数据打标签。支持数据开发者定义标签类目，标签类目是对标签作分类组织，定义标签之间的逻辑关系。通过标签加工在作业可生成数据标签，加工作业定义的内容包含标签名称、标签值、加工对象（数据表、字段）、标签加工类型。

6.4 数据分析平台

智慧城市数据中台应具备数据分析功能，应支持但不限于以下操作：

- 数据获取：数据集作为数据源和可视化展示的中间环节，承接数据源的输入，并为可视化展示输出数据表；
- 数据报表：支持在线定义报表，含明细表、分组表、交叉表等；
- 可视化分析：通过拖拉拽的形式，将图表（折线图、柱状图等）拖到仪表板布局里，通过自定义布局构建数据分析仪表盘，支持开展可视化数据分析。
- 提供定义生成行业知识图谱的工具，助力数据开发者将构建知识图谱的过程做得更高效，从而有效降低交付成本。

6.5 资产管理平台

6.5.1 概述

数据资产的最大特点是具有重复利用性和增值性，有效的数据能为组织创造巨大价值。数据中台提供资产目录，以资产目录的形式提供了数据资产展示的窗口，同时支持不同部门间数据的共享交换，让数据在不同的部门中流通，从而创造出更大的价值。能够支持构建不同类目的数据资产，包括基础数据、主题数据等。

6.5.2 数据表资产

依据国标、行标进行数据建模，形成不同的主题域（类目）及对应类目的数据表。宜有数据资产搜索、资产详情功能。

6.5.3 文件夹资产

依据不同业务场景定义文件夹，并按场景做文件夹资产分类，形成不同的类目及对应类目的文件夹资产。宜有数据资产搜索、资产详情功能。

6.5.4 数据服务资产

为用户提供白屏化的API消费申请能力，实现能力闭环。宜有数据资产搜索、资产详情功能。

6.5.5 数据指标资产

依据不同业务场景定义数据指标，并按场景做指标分类，形成不同的类目及对应类目的数据指标资产。宜有数据资产搜索、资产详情功能。

6.5.6 数据标签资产

依据不同业务场景定义数据标签，并按场景做数据标签分类，形成不同的类目及对应类目的数据标签。宜有数据资产搜索、资产详情功能。

6.5.7 资产分类分级

支持对数据资产划分等级，由租户管理员统一定义资源分级信息，各空间用户在维护数据表元数据时，即可引用字段等级信息。如：安全等级-绝密、机密、秘密。

对数据资产做权限分级。支持跨工作空间的表级、字段级数据权限控制，以有效保障数据安全。

6.5.8 数据可视化

智慧城市数据中台应具备数据可视化功能，为用户提供可视化的资产视图，具体功能如下：

——数据看板：

- 数据总量；
- 数据接入总量；
- 治理产生的数据量；
- 质检合格率；
- 服务部门数；
- 数据被消费次数。

——数据接入：

- 接入对象统计；
- 数据接入量变化趋势。

——数据治理：

- 以血缘链路的形式呈现。

——数据消费：

- 访问热度排名；
- 资产目录。

6.5.9 资产运营

面向业务运营人员，提炼数据资产、数据质量、数据使用的核心指标，直接展示数据源连接状态；通过质检敦促不同来源优化数据质量；体现数据模型价值即模型反应的是跨源头的的数据整合，才能产生有价值的的数据洞察。

6.6 数据开放平台

6.6.1 概述

智慧城市数据中台数据开放是把数据变成一种服务能力，通过数据服务让数据资产参与到业务，激活整合数据中台，是数据中台价值重要体现之一。通过安全、可控、有序的措施保障向数据消费者开放、共享数据，充分挖掘数据资源“宝矿”，发挥最大效益，提升政府治理能力和公众服务水平。

6.6.2 数据资产门户

资产门户实现信息资源共享、业务协同和数据开放的基础，是各部门之间信息共享的依据。提供统一的资产目录，数据消费者可通过搜索、标签、最新、推荐等快速定位所需数据资产。

a) 资产概览：

统计数据表总量、数据服务总量、派生指标总量、数据标签总量、文件总量，以直观的形式展示资产概况；

b) 主题分类：

展示数据表、文件夹、数据指标、数据标签、数据服务共 5 类数据资源的一级类目，便于快速检索；

c) 数据动态：

1) 最新数据：按照资源类型分类，根据数据更新时间，倒序排列，展示各类型数据最近更新的数据信息；

2) 热门数据：按照资源类型分类，根据数据共享次数，倒序排列，展示各类型的数据信息。

6.6.3 数据资产目录

数据资产的最大特点是具有重复利用性和增值性，有效的数据能为组织创造巨大价值。数据中台提供资产目录，以资产目录的形式提供了数据资产展示的窗口，同时支持不同部门间数据的共享交换，让数据在不同的部门中流通，从而创造出更大的价值。

资产目录是实现信息资源共享、业务协同和数据开放的基础，是各部门之间信息共享及数据向社会开放的依据。资产目录便于业务人员快速定位所需数据资产，多元化的资源共享方式，助力组织实现不同类型数据资产的共享。

6.6.4 资产访问申请

数据资源消费者通过中台资源目录功能，定位想要访问的数据表，在线提交访问申请。提交访问申请需填写访问说明。支持表级、列级、行级的数据表访问，覆盖不同粒度的数据消费需求。

6.6.5 资产访问授权

应支持对数据表、文件夹、数据服务、数据指标、数据标签的访问申请进行审批授权。

6.6.6 资产访问授权回收

6.6.6.1 数据表访问授权回收

应支持对已授权的数据表进行授权回收。

6.6.6.2 文件夹访问授权回收

应支持对已授权的文件夹进行授权回收。

6.6.7 审批管理

智慧城市数据中台数据开放平台应具备数据审批管理功能。

注：数据访问申请、访问授权、访问授权回收都涉及在线 workflow 审批，通过 workflow 完成数据资产访问申请、访问授权以及访问授权回收的审批管理，提供个人工作台以统一查看待审批的、已审批的、已申请的等相关信息。

6.7 运维监控平台

6.7.1 概述

运维监控是数据中台得以健康、持续运转的基础。智慧城市数据中台运维监控平台应能对数据全生命周期的监控、告警，便于迅速定位和排除故障，保障业务正常稳定运行。

6.7.2 运维概览

智慧城市数据中台运维监控平台应具备运维概览功能，运维概览宜支持展示包括但不限于以下内容：

- 需要重点关注的运维指标；
- 任务运行状态分布；
- 任务完成情况；
- 运行时长排行；
- 任务出错排行；
- 调度任务数量趋势；
- 监控日志管理；
- 任务类型分布。

6.7.3 调度管理

智慧城市数据中台运维监控平台应具备调度管理功能，对数据处理任务进行监控、告警，便于迅速定位和排除故障，保障业务正常稳定运行。

调度管理信息列表应包括但不限于以下内容：

- 作业名称；
- 作业类型；
- 状态；
- 调度周期；

——近 1 周期任务数。

6.7.4 任务中心

任务中心集中纳管中台的所有任务，包括数据集成、SQL开发以及数据质量，对任务做统一呈现，便于查看执行失败的任务并定位任务失败原因。其中任务状态宜分为以下阶段：

- 等资源；
- 运行中；
- 运行成功；
- 运行失败；
- 运行超时；
- 运行过期。

7 数据安全平台

智慧城市数据中台完善的数据安全能力，提供数据分类分级、敏感数据识别、数据加解密、API安全网关、危险操作认证、数据水印溯源等能力，保证数据的隐私合规、可审计、可回溯，多重安全保护机制，全方位保障整数据安全。

平台应具备隐私计算统一接入能力，支持以主流隐私计算技术完成高价值敏感数据服务，保障数据要素安全有序流通。

智慧城市数据中台数据安全平台应符合GB/T 37973、GB/T 39477、GB/Z 38649的要求。

8 运维要求

8.1 人员要求

智慧城市数据中台运维人员应符合以下要求：

- 熟悉 Linux 系统相关的管理工具（如 grep、sed、awk 等）及 shell、Python 等计算机语言；
- 应具备 MySQL、postgresql、国产信创数据库等关系型数据库操作技能；
- 计算机网络、操作系统等基础；
- 熟悉开源数据平台及相关技术，包括但不限于 Hadoop、MapReduce、redis、Hive、Hbase。

8.2 运维内容要求

智慧城市数据中台上线试运行完成后，应进行业务应用情况监控检查；当发现业务或数据库读写不正常时应及时处理。主要检查内容如下：

- 系统检查：
 - 检查系统日志是否有告警；
 - 检查CPU、内存、网络等系统性能是否异常。
- 业务检查：
 - 定期巡检，检查数据应用业务是否正常。

9 一般建设流程

构建一套智慧城市数据中台，需要有成熟的数据中台建设方法论作为指导，数据中台建设方法论可分为资源盘点、架构规划设计、数据资源汇集、数据治理和数据服务运营五个阶段，实现数据“汇、通、管、用、评”端到端全生命周期的数据建设体系，具体内容如下：

- a) 资源梳理盘点：对项目客户环境进行数据资源梳理盘点、开展数据与业务调研工作；
- b) 架构规划设计：根据调研梳理的情况，进行项目业务、技术、数据架构规划方案设计；
- c) 数据资源汇聚：进行多源异构数据的采集汇聚，纳入数据中台进行统一管理；
- d) 数据治理：依托标准规范体系，对归集的数据进行清洗整合、主题建模、指标定义、标签生产、质量检核、血缘计算等操作，以提升数据质量，挖掘数据价值；

- e) 数据服务运营：应把数据变成一种服务能力，通过数据服务让数据资产参与到业务中，并持续地进行数据资源的汇集、治理与数据服务开发的支撑，构建数据要素安全可持续流通的能力。
-