



中华人民共和国国家标准

GB/T XXXXX—XXXX

城市全域数字化转型 城市智能中枢 数据 治理要求

Citywide digital transformation—City intelligent center—Requirements of data
governance

（征求意见稿）

（本草案完成时间：2024/12/19）

在提交反馈意见时，请将您知道的相关专利连同支持性文件一并附上。

XXXX—XX—XX 发布

XXXX—XX—XX 实施

国家市场监督管理总局
国家标准化管理委员会 发布

目 次

前言 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 2

5 数据治理范围 2

6 参考模型 3

7 数据治理实施过程 4

 7.1 概述 4

 7.2 统筹与规划 4

 7.3 组织设立 4

 7.4 资产梳理 4

 7.5 规范建立 4

 7.6 治理实施 5

 7.7 监控和评价 5

 7.8 改进和优化 5

8 归集 6

 8.1 数据归集方式 6

 8.2 数据对账 6

 8.3 数据归档 7

9 治理 8

 9.1 数据标准 8

 9.2 数据质量 9

 9.3 主数据 12

 9.4 数据融合 13

 9.5 数据互操作 13

10 管理 13

 10.1 元数据管理 13

 10.2 目录管理 15

 10.3 资源管理 17

 10.4 任务管理 17

11 安全 18

 11.1 数据加密保护 18

 11.2 敏感数据识别 18

 11.3 数据脱敏 19

11.4 数据分类分级..... 19

11.5 数据溯源..... 19

参考文献..... 20

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。

本文件由全国信息技术标准化技术委员会（SAC/TC 28）提出并归口。

本文件起草单位：

本文件主要起草人：

城市全域数字化转型 城市智能中枢 数据治理要求

1 范围

本文件给出了城市智能中枢数据治理的范围和参考模型，规定了数据治理实施过程以及数据归集、治理、管理和安全等方面的要求。

本文件适用于指导城市智能中枢数据治理体系以及相关系统平台的规划、设计与建设。

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 18391.2 信息技术 元数据注册系统（MDR） 第2部分：分类

GB/T 18894-2016 电子文件归档与电子档案管理规范

GB/T 22239-2019 信息安全技术 网络安全等级保护基本要求

GB/T 34960.5-2018 信息技术服务 治理 第5部分：数据治理规范

GB/T 39786-2021 信息安全技术 信息系统密码应用基本要求

GB/T 42093.1-2022 标准文档结构化 元模型 第1部分：全文

GB/T 43697-2024 数据安全技术 数据分类分级规则

3 术语和定义

下列术语和定义适用于本文件。

3.1

城市智能中枢 city intelligent center

综合应用新一代信息技术，融合城市多源数据资源，整合状态感知、建模分析、城市运行管理、应急指挥等功能，通过人机交互与协同，提供态势全面感知、趋势智能研判、协同高效处置、调度敏捷响应、平急快速切换等智能化服务，提升城市精准精细治理水平，支撑城市数字化转型的一种复杂系统。

3.2

数据生存周期 data life cycle

数据获取、存储、整合、分析、应用、呈现、归档和销毁等各种生存形态演变的过程。

[来源：GB/T 34960.5-2018，3.7]

3.3

数据治理 data governance

基于数据生存周期，进行数据全面质量管理、资产管理、风险管理等统筹与协调管控的过程。

注：多学科视角下，数据治理需要解决数据权属关系问题，明确数据利益相关方的角色，权利和权益及其责任关系和工作任务，避免数据风险、提高数据质量，实现数据资产能长期有序地、可持续地得到管理和利用。信息技术视角下，数据治理指对数据进行管控、处置、格式化和规范化的过程，数据治理是数据和数据系统管理的基本要素，数据治理涉及数据生存周期管理，无论数据是处于静态、动态、未完成状态还是交易状态。

[来源：GB/T 37043-2018, 2.3.1, 有修改]

3.4

数据管理 data management

数据资源获取、控制、价值提升等活动的集合。

[来源：GB/T 34960.5-2018, 3.2]

3.5

元数据 metadata

关于数据或数据元素的数据（可能包括其数据描述），以及关于数据拥有权、存取路径、访问权和数据易变性的数据。

[来源：GB/T 34960.5-2018, 3.6]

3.6

数据归集 data aggregation

将不同来源的数据按照一定的规则进行收集、汇聚、整合的行为和过程。

4 缩略语

下列缩略语适用于本文件。

ETL：数据抽取、转化、加载（Extract-Transform-Load）

FTP：文件传输协议（File Transfer Protocol）

HDFS：Hadoop分布式文件系统（Hadoop Distributed File System）

HTTP：超文本传输协议（Hypertext Transfer Protocol）

HTTPS：超文本传输安全协议（ypertext Transfer Protocol Secure）

MPP：大规模并行处理（Massively Parallel Processing）

SFTP：SSH文件传输协议（SSH File Transfer Protocol）

SQL：结构化查询语言（Structured Query Language）

5 数据治理范围

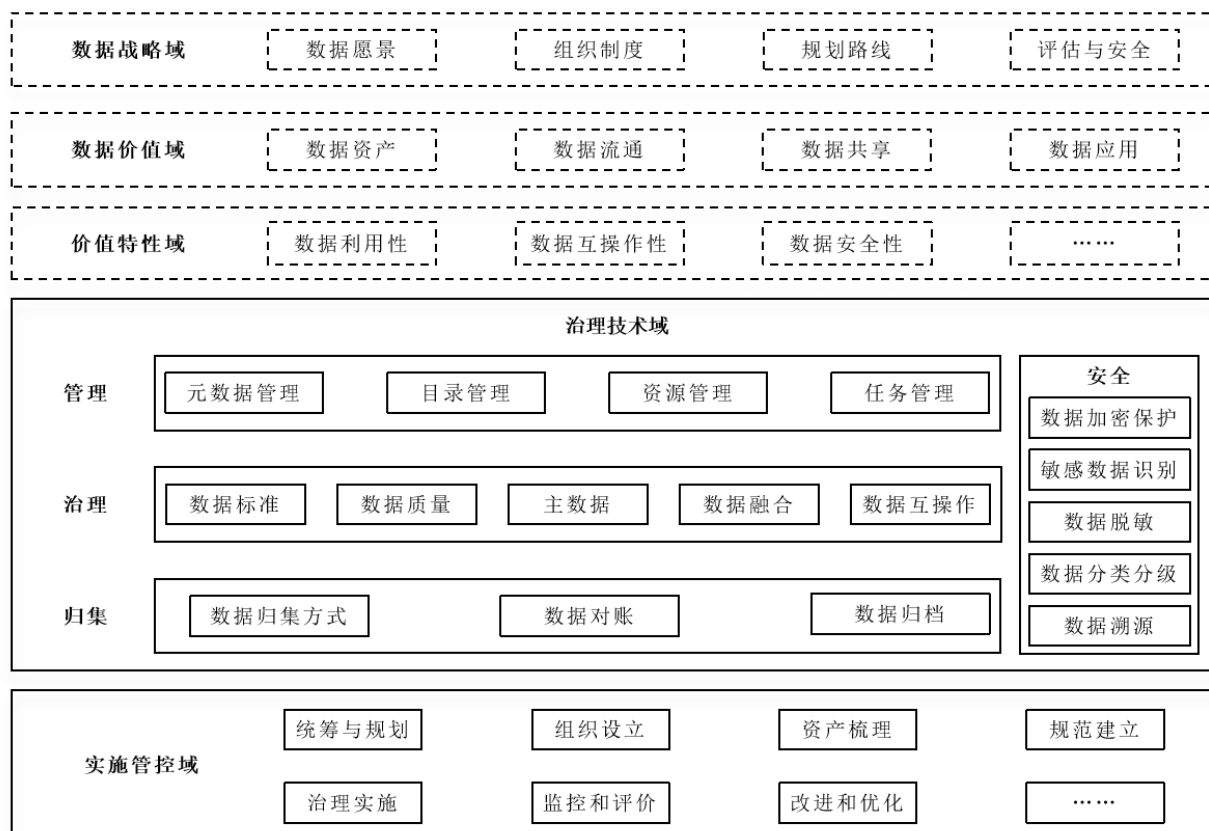
城市智能中枢数据治理的范围主要包括以下几类：

- a) 城市公共数据：城市范围内具有管理公共事务职能的组织、提供公共服务的组织在依法履行职责或者提供公共服务过程中收集、产生或获得的数据，包括但不限于交通、能源、水务、环保、城市规划以及各类城市基础设施运行的数据；
- b) 政府数据：各级政府和行政部门生成的数据，包括但不限于行政事务、政务公开、社会治理等数据；
- c) 社会组织数据：各类社会组织、行业协会及非政府机构收集和维护的数据，包括但不限于经济、文化、教育、医疗等领域的数据；
- d) 企业数据：城市中各类企业的经营活动数据，包括但不限于商业数据、生产数据、市场分析数据等；
- e) 个人数据：指能够单独或与其他信息结合识别自然人身份的各类信息，包括但不限于个人健康、出行、消费等行为数据；
- f) 物联网（IoT）数据：通过物联网设备、传感器、终端等产生的数据，包括但不限于传感器采集的数据、设备状态数据、环境数据等。

6 参考模型

城市智能中枢的数据治理参考模型兼顾数据治理各相关方的利益关注点,以数据利用、数据互操作、数据安全为驱动,从城市数据资源拥有者和管理者的视角出发,构建可持续发展的数据治理体系和可操作性框架模型,按照成体系、可迭代、可演进、有弹性的治理方式,有序组织实施治理过程,实现城市智能中枢数据资源可持续治理。

城市智能中枢数据治理参考模型由数据战略域、数据价值域、价值特性域、治理技术域、实施管控域五个部分组成,见图1。



注：虚线框部分不作为本文件约束的内容。

图1 城市智能中枢数据治理参考模型

城市智能中枢数据治理参考模型各部分描述如下：

- 数据战略域：位于模型顶层，是城市智能中枢数据治理的全局描述，包括但不限于数据愿景、组织制度、规划路线、评估与安全等方面；
- 数据价值域：位于数据战略域之下，城市智能中枢数据治理以数据价值为核心驱动力，可持续积累和沉淀城市数据资源与资产，城市数据价值体现在数据资产、数据流通、数据共享、数据应用等方面；
- 价值特性域：是数据价值域的延伸，包括但不限于数据利用、数据互操作、数据安全等特性；
- 治理技术域：位于数据价值域之下，描述实现数据价值所需的数据治理方式和技术手段。将已有分散的城市多源数据资源以物理统一或逻辑统一的形式归集至统一的数据平台并进行治理，对治理好的数据资源进行有效的管理和安全保护；

- e) 实施管控域：位于模型底层，表示城市智能中枢数据治理是在数据战略规划下，由组织和人员实施的一系列数据治理的过程和结果，描述所需的组织机制、治理流程、评估体系等内容。

7 数据治理实施过程

7.1 概述

城市智能中枢数据治理实施过程主要包括统筹与规划、组织设立、资产梳理、规范建立、治理实施、监控和评价、改进和优化等七个步骤，根据实际情况和环境因素可适当裁剪相应步骤。

7.2 统筹与规划

全面梳理和规划城市智能中枢的数据资源，应符合GB/T 34960.5-2018中9.1规定的要求，保障数据资源的系统性、可控性和可持续发展，包括数据资源识别、数据分类分级和数据需求分析等。统筹与规划要求如下：

- a) 规划数据资源应开展数据资源现状盘点，确定数据资源规划目标并编制数据资源规划文件；
- b) 规划的数据处理活动可包括但不限于数据收集、数据存储、数据处理、数据交换、数据应用、数据删除和长期保存等；
- c) 数据资源识别应识别并分类城市各类数据资源，如政府数据、企业数据、个人数据、公共数据等；
- d) 数据分类分级应基于数据的来源、敏感度、用途对数据进行分类分级，明确数据流动的规则和边界；
- e) 应开展数据需求分析，明确当前和未来的数据需求，实现数据资源能满足城市管理和服务中的业务需求。

7.3 组织设立

构建和完善数据治理的组织架构，明确数据治理任务分配、角色职责，组织设立要求如下：

- a) 应组建跨部门治理团队，由技术、业务、法律、管理等多方参与，负责数据治理工作的推进与监督；
- b) 应分配角色与责任，明确各成员的角色职责，如数据管理员、数据分析师、数据安全员等，保障每个环节有专人负责；
- c) 宜成立数据治理委员会，负责战略层面的数据决策、政策制定和跨部门协调工作。

7.4 资产梳理

资产梳理在对数据资产进行系统化的梳理与评估，形成城市数据资产目录，明确数据资产的价值和权属，包括以下内容：

- a) 数据资产分类：对元数据、主数据、参考数据、一般数据进行系统梳理，分类归档；
- b) 数据资产评估：评估数据的价值、质量、使用情况和潜在应用场景，为数据治理提供支持。

7.5 规范建立

制定数据治理框架中的相关政策、标准和流程，包括以下内容：

- a) 数据治理政策制定：明确数据管理的原则、目标和要求，覆盖数据生命周期的各个阶段；
- b) 数据标准化：制定数据格式、数据质量、数据交换、数据安全等方面的标准，保障数据的一致性、可操作性和安全性；

- c) 评估与考核制度：建立数据治理过程中的评估与考核机制，保障各部门按规定执行数据管理要求。

7.6 治理实施

治理实施是将数据治理的规划、规范和架构具体落实到实际操作中，保障数据治理活动的顺利进行，包括以下内容：

- a) 数据集成与管理：整合来自不同部门和领域的的数据，建立统一的数据管理平台，实现数据资源的集中管理；
- b) 数据质量控制：通过数据清洗、数据去重、数据校验等技术手段，保障数据的一致性、完整性、准确性、时效性、规范性、唯一性和可访问性；
- c) 数据安全性与隐私保护：实施数据加密、访问控制、审计追踪、数据备份和恢复等技术，保障数据的安全性和隐私合规性；
- d) 培训和宣传：通过培训和宣传提高人员数据保护意识和能力。

7.7 监控和评价

对数据治理过程进行监控和评价，应符合GB/T 34960.5-2018中9.3规定的要求，并建立一套完整且系统的机制，保障数据的有效管理、使用和保护，包括以下内容：

- a) 监督与评估：建立数据治理的监督与评估机制，定期对数据治理活动进行检查和评估；
- b) 绩效指标设定：制定衡量数据治理成效的关键绩效指标，如数据质量提升率、数据利用率、数据安全合规率等；
- c) 定期审查与报告：通过定期的数据治理审查，生成报告，评估各项工作是否按计划推进，有无偏差；
- d) 反馈与改进建议：基于评估结果，提出改进建议，推动治理框架的持续优化。

7.8 改进和优化

7.8.1 差异分析

针对数据治理过程中出现的问题和偏差进行分析，找出治理执行与预期目标之间的差距，包括以下内容：

- a) 目标与执行偏差识别：通过对比实际执行结果和预期目标，找出差异点，如数据利用效率低下、数据共享不畅等；
- b) 根因分析：分析差异的产生原因，可能涉及技术问题、组织问题或外部环境变化等；
- c) 风险评估：评估差异对整体数据治理的影响，判断问题的严重程度和优先级。

7.8.2 改进执行

基于差异分析的结果，采取具体行动对数据治理体系进行优化和提升，包括以下内容：

- a) 改进措施制定：针对差异分析的结果，制定具体的改进措施，如优化数据治理流程、加强技术支持、调整角色分工等；
- b) 改进实施计划：制定清晰的实施计划，包括时间节点、负责人员、具体行动方案等，保障改进工作有序进行；
- c) 持续监控与优化：在改进执行过程中，持续监控各项措施的实施效果，实现问题的有效解决，并为下一轮数据治理提供反馈。

8 归集

8.1 数据归集方式

8.1.1 库表归集

库表归集要求包括。

- a) 库表归集按归集类型划分包括整库迁移、多表迁移、单表同步等，要求如下：
 - 1) 应提供源表与目标表结构一致的整库数据迁移功能；
 - 2) 应提供源表与目标表结构一致的批量表数据迁移功能；
 - 3) 应提供源表与目标表结构不一致时数据迁移功能，支持表字段自由映射。
- b) 库表归集按归集方式划分包括全量归集、增量归集，要求如下：
 - 1) 应提供源表数据全量覆盖到目标表的功能；
 - 2) 应提供源表数据增量迁移到目标表的功能，支持更新、追加、覆盖增量模式。
- c) 库表归集按归集频率划分包括离线归集、实时归集，要求如下：
 - 1) 应提供源表数据周期性自动迁移到目标表的功能，支持分钟、小时、天、周、月等频率；
 - 2) 应支持实时感知源表数据增删改等变化，并自动同步到目标表；
 - 3) 应支持实时感知源表结构变化，并自动同步修改目标表结构。

8.1.2 文件归集

文件归集包括文件数据源管理、归集任务配置、归集任务调度等，要求如下：

- a) 应提供 HDFS、FTP、SFTP 等文件数据源迁移功能；
- b) 应提供组件拖动、组件连线、任务调试、设计布局等可视化任务配置功能；
- c) 应提供指定源端文件目录或自定义文件筛选规则功能；
- d) 应提供解析源端半结构化文件为结构化数据，并迁移到目标库表功能；
- e) 应提供源端文件周期性自动迁移功能，支持分钟、小时、天、周、月等频率。

8.1.3 接口归集

接口归集包括接口数据源管理、归集任务配置、归集任务调度等，要求如下：

- a) 应支持连接多种网络数据传输协议，如 HTTP/HTTPS、WebService 等；
- b) 应提供归集接口的发布、配置、任务调试功能；
- c) 应提供源端接口传入变量请求参数功能；
- d) 应提供自定义时间频度调用源端接口归集功能；
- e) 应提供主动调用接口归集和被动接收接口推送归集功能；
- f) 应提供解析源端接口返回数据为结构化数据，并迁移到目标库表功能。

8.2 数据对账

8.2.1 概述

数据对账提供事前即时对账和事后盘点对账功能，通过对数据条数、数据内容等进行核对和检验，提供差异数据和数据对账结果报告，作为漏数问题依据，帮助快速定位数据不标准、数据质量差、操作不规范等原因，同时提供漏数自动补发机制，支持漏数任务自动补发。

8.2.2 对账要求

对账要求包括结构化数据对账、非结构数据对账等功能，要求如下：

- a) 应提供源表和目标表在某一对账时间节点比对数据量和数据内容的完整性和一致性功能；
- b) 应提供源端文件和目标端文件在某一对账时间节点比对文件个数和文件内容的完整性和一致性功能；
- c) 应提供数据归集成功后自动触发统计对账功能；
- d) 应提供配置数据对账异常告警规则功能；
- e) 应提供对账问题数据在线查看/下载功能。

8.2.3 及时对账

在数据分发的过程中，实现对提供方与分发方数据条数的实时核验，要求如下：

- a) 应支持对自动对账后形成的对账单进行查询，展示如分发资源、提供方数据条数、分发方数据条数、对账时间、对账状态、对账结果，支持查看对即时账单详情；
- b) 应支持查看数据提供方对账单和数据分发方对账单，显示数据条数、发送时间、对账时间、对账方式等，满足提供方对账单和分发方对账单的比对需求；
- c) 应支持生成提供方与分发方即时对账单，支持查看对账单后，进行手动核销。

8.2.4 盘点对账

针对数据分发过程后的场景，配置不同的对账方案实现对提供方与分发方数据条数的盘点，要求如下：

- a) 应支持配置方案生成盘点对账方案。支持对数据资源设置对账频次，可按单次、星期循环、月循环配置对账；
- b) 应支持查看由配置方案生成的对账列表，主要包括盘点对账的对账名称、最后一次对账日期、对账频次、执行状态、对账结果、补发结果；
- c) 应支持查看不同对账单的对账详情。支持查看数据提供方对账单和数据分发方对账单。显示数据条数、发送时间、对账时间、对账方式、补发结果等；
- d) 应支持生成提供方与分发方盘点对账单，支持查看对账单后，进行手动核销；
- e) 应支持导出对账差异附件，提供方与接收方的差异主键的数据，为对账人员提供纠错依据；
- f) 应支持根据对账差异附件的数据，根据对账方案区间对账结果的差异数据自动补发。

8.3 数据归档

数据归档是指将数据转移到受控的存储媒体中进行长期保存，要求如下：

- a) 存储媒体应符合GB/T 18894-2016中8.2.2规定的要求；
- b) 归档文件格式应符合GB/T 18894-2016中8.4规定的要求，实现归档数据长期可读；
- c) 应设置适当的保护程序来验证存储媒体的可访问性，实现存储媒体在数据保存期限内可访问；
- d) 应明确归档数据的访问权限和访问方式；
- e) 应制定数据恢复机制，在需要时能快速、准确地恢复归档数据；
- f) 应对涉及个人隐私的数据进行脱敏或匿名化处理，采取加密、访问控制等措施，保障归档数据安全；
- g) 应提供归档策略配置、归档数据恢复、归档数据监控等功能；
- h) 应提供管理组织内部和外部监管的数据归档需求功能，预测归档数据对系统性能、维护成本、数据安全产生的影响；
- i) 应提供自定义 workflow 设计数据归档流程任务功能；
- j) 应提供设置数据归档库表或物理存储属性功能；

- k) 应提供对数据资源进行按时间或按记录进行归档功能；
- l) 应支持在归档数据时同步迁移归档要素，如结构定义、数据字典等信息；
- m) 应提供恢复已归档数据功能；
- n) 应提供分析并记录归档数据增长趋势、归档操作的实施成本功能。

9 治理

9.1 数据标准

9.1.1 数据标准文件管理

数据标准文件管理包括标准文件维护、标准文件版本管理、标准文件引用情况等，要求如下：

- a) 应对标准文件进行分类、编号和归档管理；
- b) 应提供数据标准文件新增、引用、查询、修改、删除等功能；
- c) 应提供数据标准文件浏览查询、关键字模糊查询功能；
- d) 应提供对标准文件版本管理功能，可查看历史版本；
- e) 应提供以列表或可视化方式查看数据标准中对标准文件的引用情况功能。

9.1.2 数据字典标准管理

数据字典标准管理指对数据字典标准的数据元值域进行管理，要求如下：

- a) 应根据组织业务管理需求，制定组织内数据字典分类管理框架，保障数据字典分类框架的全面性和各分类之间的独立性；
- b) 应支持创建数据元标准的代码集，配置数据字典的内部标识符，并自定义维护代码集属性，包括但不限于：代码集名称、代码集编码、代码集字段信息，填写数据字典状态及生效日期；
- c) 应支持引用数据元标准中审批通过的数据项，支持设置约束信息；
- d) 应提供代码集全生存周期的管理功能，包括但不限于代码集数据的创建、发布、审核、下线、版本管理和删除；
- e) 应支持对数据字典中类别、特性等进行更换；
- f) 应支持对数据字典的正确性、安全性、完整性进行测试，并生成测试结果；
- g) 应支持通过手动录入、库表导入、文件导入等方式填充代码集数据；
- h) 应提供建立代码集与数据元的链接功能，规范数据元的值域范围。

9.1.3 数据元标准

数据元标准是指对数据的基本单元进行规范化的定义和描述的规则集合，要求如下：

- a) 应建立数据元的分类和命名规则，规范数据元标准的命名、类型、长度、值域及其他相关说明信息，制定数据元的命名、描述与表示规范；
- b) 应建立数据元的统一目录，根据数据元的分类及业务管理需求，建立数据元管理的目录，对组织内部的数据元分类存储；
- c) 应提供自定义数据元标准属性功能，标准属性包括但不限于：业务属性、技术属性、管理属性、质量属性、生存周期属性；
- d) 应支持对数据元类目进行分类管理和维护，支持自定义数据元类目，对数据元进行管理；
- e) 应支持数据元的查找和引用，提供数据元查找和引用的在线工具，支持按照数据元名称、内部标识符、英文名称等不同维度对数据元进行查找；
- f) 应支持根据元数据生成数据元标准；

- g) 应支持对数据元标准进行导入和导出；
- h) 应支持数据元标准的版本管理；
- i) 应支持数据元标准全文检索；
- j) 应支持数据元标准生存周期管理，包括但不限于新增、发布、停用、变更、删除；
- k) 宜支持根据数据元标准定期进行引用情况分析，了解各应用系统中对数据元的引用情况。

9.1.4 编码规则管理

编码规则用于约束和引用该数据标准的编码生成规则，可提升标准编码的规范性，降低手动配置成本和管理成本，要求如下：

- a) 应支持自定义编码和基于编码规则自动生成二种方式，具有唯一性；
- b) 应支持根据自增序列（按标准创建顺序）、固定字符串、所属标准集编码创建编码规则；
- c) 应支持根据指定序列的位数、起始值和步长，按照标准 ID 的生成时间自增；
- d) 应支持选择公共标准属性类型，如业务属性、技术属性、管理属性、质量属性、主数据属性、生命周期属性、安全属性等；
- e) 应定义数据编码的标准化格式，梳理组织内部的数据，形成统一的编码规则；
- f) 应支持引用相关标准新建编码规则；
- g) 应支持根据引用的标准模板配置编码生成规则自动生成编码值，生成后可手动修改；
- h) 应支持自定义组合编码规则，包括但不限于流水号、字母、数字、常量、时间变量；
- i) 应支持导入和导出编码规则；
- j) 应支持对编码规则进行分类管理和维护；
- k) 应提供编码规则生存周期管理功能，包括但不限于新增、发布、停用、变更、删除；
- l) 应定期对编码规则的使用情况进行检查和评估，发现问题及时纠正；
- m) 应提供编码规则生成接口功能，供第三方平台调用。

9.1.5 标准配置

标准配置提供数据标准内部标识符规则生成配置功能，依据9.1.4编码规则管理，在创建数据元、数据项集、限定词时自动生成编码规则。标准配置要求如下：

- a) 应支持对数据元的内部标识符定义自动生成规则，在手动创建本地数据元标准时内部标识符可通过定义的规则自动生成；
- b) 应支持设定智能对标优先级配置，在数据标准智能对标时优先遍历优先级高的标准分类中的数据标准项；
- c) 应支持自定义标准信息项，满足业务标准的差异化维度及使用需求。

9.2 数据质量

9.2.1 质量维度

宜参照GB/T 36344-2018设置数据质量维度，描述如下：

- a) 一致性：数据与其他特定上下文中使用的数据无矛盾的程度；
- b) 完整性：按照数据规则要求，数据元素被赋予数值的程度；
- c) 准确性：数据准确表示其所描述的真实实体（实际对象）真实值的程度；
- d) 时效性：数据在时间变化中的正确程度；
- e) 规范性：数据符合数据标准、数据模型、业务规则、元数据或权威参考数据的程度；
- f) 唯一性：同一数据只能有唯一的标识符；

- g) 可访问性：数据能被访问的程度。

9.2.2 数据质量管理

数据质量管理旨在保障数据在其生存周期内的质量，要求如下：

- a) 应支持识别影响数据质量的数据流程，具体包括但不限于数据采集、数据稽核、数据标记、数据更正等环节，并梳理数据流向；应针对个流程制定具体的管理措施，实施流程管控；
- b) 应明确数据质量稽核流程的程序、内容和岗位职责，编制程序文件和相应的作业指导文件；
- c) 应明确数据质量标记流程和内容，建立数据质量业务规则，监督、指导相关部门及业务人员对业务规则进行执行；
- d) 应明确数据更正流程的工序和内容，及时督查整改质量不合格的数据；
- e) 应建立数据可追溯机制，如通过审计日志等方式在采集、审核校验、更正等流程中对数据进行唯一性标识，并保留文件化信息；
- f) 应收集数据使用部门和人员的数据质量反馈信息，并采取纠正措施。

9.2.3 质量规则

应基于数据质量维度，设置数据质量规则。质量规则要求如下：

- a) 应支持库级、表级、字段级、跨字段级规则和自定义规则，自定义规则支持通过 SQL 语句定义质量规则，支持按照字段配置监控规则；
- b) 应支持 SQL 命令、谓词逻辑、正则规则、字典规则和值域规则等多种规则形式；
- c) 应支持用户按需批量导入和导出质量规则；
- d) 应支持对规则的语法正确性进行检查并反馈检查结果，对于语法不正确的规则提供编辑和修正功能；
- e) 应支持对表级增量数据行及数据大小波动的检测，以及字段级检测；
- f) 应支持配置数据过滤条件，控制检测数据范围；
- g) 应支持质量规则分级控制，如强规则异常时阻断下游任务，防止脏数据传播；弱规则异常时仅触发报警；
- h) 应支持波动率检验，提供绝对值和相对值（上升或下降）两种比较方式；
- i) 应支持数值型检验，提供绝对值比较方式，涵盖空值、唯一值、重复值等关键指标的统计和分析；
- j) 应支持自定义配置质量规则告警条件，包括但不限于问题数据比例、问题严重程度和问题通知对象等；
- k) 应支持通过并行计算引擎和分布式计算能力，实现跨库、跨表、跨字段的数据质量检测；
- l) 应支持对质量规则进行测试并生成测试结果。

9.2.4 质量检测任务

质量检测任务包括质量检测对象的质量规则编排、任务调度配置、任务监控等，要求如下：

- a) 应支持根据质量检测方案自动生成质量检测任务，并提供任务的运维管理功能；
- b) 应支持指定检测数据范围，包括全量数据与增量数据，并支持按日、周、月、年及具体事件周期进行调度配置；
- c) 应支持一个检测对象配置多个分支质量校验流程；
- d) 应提供图形化界面，支持用户以并行或串行的方式编排质量规则；
- e) 应支持通过系统接口对接和人工录入登记的方式进行纠错登记，配置错误数据基本信息、申请人基本信息、反馈内容、附件及补充说明等关键要素，并提供错误数据的处理机制；

- f) 应支持手动或自动执行质量检测任务，应支持对检测任务进行优先级管理；
- g) 应支持查看质量检测任务的详细信息，并支持将任务分派给指定的处理对象；
- h) 应提供新增、送审、变更、停用、启用、删除等质量检测任务的维护功能；
- i) 应支持根据数据质量规则进行定量检测和分析；
- j) 应支持在整个数据集上进行数据质量规则的验证，并提供频率分布和度量指标；
- k) 应支持测量和监控数据项对数据质量规则的符合度；
- l) 应支持流动式和批量式两种控制方式；
- m) 应支持对检测任务的监控，包括对任务的执行情况、历史执行情况、任务告警记录的查看等；
- n) 应支持用户配置质量告警内容模板，并支持验证告警模板是否正确运行；
- o) 应支持对检测任务进行测试，并生成测试结果。

9.2.5 数据质量分析

数据质量分析要求如下：

- a) 应明确数据质量问题的分析要求，对关键数据质量问题的根本原因、影响范围进行分析；
- b) 应建立数据质量分析案例库；
- c) 应提供向下钻取功能，支持用户根据需要深入到数据的各个层次进行细致分析；
- d) 应支持统计分析、趋势分析、关联分析等多种数据分析形式，支持不同数据质量问题的分析需求；
- e) 应支持将质量分析结果以图形化的方式展现。

9.2.6 质量问题管理

质量问题管理要求如下：

- a) 应支持自动生成问题工单，工单信息包括但不限于工单产生时间、工单状态、工单处理进度、检测总数、问题总数和问题明细；
- b) 应支持用户对问题工单进行有效性确认；
- c) 应提供问题工单派发功能，支持自动派发至预设的问题责任人、手动派发至指定的处理人；
- d) 应支持用户根据需求自定义问题工单的处理闭环流程；
- e) 应支持在线浏览问题数据明细；
- f) 应支持对问题数据明细进行打标，标识出问题数据违反的质量规则；
- g) 应提供问题数据明细的导出功能，支持将数据导出至本地存储；
- h) 应支持查看问题工单检测对象的字段级血缘关系；
- i) 应支持建立质量问题知识库，将质量问题解决方案转化为知识点进行沉淀。

9.2.7 质量报告

数据质量报告要求如下：

- a) 应支持创建数据质量报告模板，涵盖质量评分、问题字段明细、问题分布、告警分布、问题趋势、评分趋势等关键指标；
- b) 应支持对数据质量报告模板进行编辑，包括模板的新增、变更、停用、启用、删除等；
- c) 应支持通过图形化界面，结合图形、列表与文字相结合的方式，展现数质量状况，同时提出数据质量改进建议；
- d) 应支持基于名称、分类等条件模糊或精准搜索数据质量报告；
- e) 应支持针对全局视角、业务系统视角、数据责任部门视角、数据库技术视角等多维度的数据质量分析；

- f) 应支持按年、季度、月等时间维度查看数据质量得分变化趋势；
- g) 应支持展示问题数据分布情况，支持对数据质量检测过程中的各类信息进行汇总分析；
- h) 应支持根据当前问题数据分布，自动生成数据质量提升建议；
- i) 应支持查询指定时间段内容的质量评分明细；
- j) 应支持深入分析数据质量事件，追踪错误数据血缘，识别问题类型及起源，并提出解决方案选项；
- k) 应支持自动化生成定期数据质量管理运行报告，包括过程监控和结果总结，并建立报告发布机制；
- l) 应支持输出整表质量报告，包括抽样条数、表大小、字段个数、整体空置率、注释完整率、整列为空字段数、整行重复行数等关键指标；
- m) 应支持输出字段级质量报告，包括字段名称、字段注释、字段类型、是否主键、抽样值、空值个数、空置率、唯一值个数、零值数、字段探查得分等关键指标。

9.2.8 数据质量提升

数据质量提升要求如下：

- a) 应支持数据质量校正，采用数据标准化、数据清洗、数据转换和数据整合等技术手段，对不符合质量要求的数据进行处理，解决现有数据质量问题；
- b) 应支持对数据质量提升全流程的监控，包括数据质量提升的申请、审核、批复流程等各个流程节点的状态监控；
- c) 应建立数据质量知识库，收集各类数据质量案例、经验和知识，形成组织的数据质量知识库。

9.3 主数据

9.3.1 主数据管理

主数据管理要求如下：

- a) 应支持对主数据进行增、删、改、查等操作；
- b) 应支持对主数据的状态进行管理，包括主数据的申请、审批、发布；
- c) 应支持对主数据的变更进行版本管理；
- d) 应支持对主数据管理流程的审批，包括主数据模型、主数据维护、主数据编码、模型加工关系审批，能够根据实际管理需求自定义审批流程；
- e) 应提供主数据质量提升功能，包括定义质检规则、生成质检方案；
- f) 应支持通过提供主数据查询接口，为各信息系统提供主数据服务；
- g) 应支持通过可视化界面查看系统内主数据整体情况，包括但不限于主数据的分布、来源及变化情况；
- h) 应提供主数据标准定义能力，能够定义与管理主数据属性的约束条件和校验规则。

9.3.2 主数据编码规则

主数据编码规则要求如下：

- a) 应提供主数据编码规则功能，能够对主数据进行编码规则定义；
- b) 应支持提供多种编码格式类型，包括但不限于流水码、固定字符串、日期时间码等；
- c) 应支持多个码段组合的编码，也可设置一个码段的顺序编码；每个码段可设置编码规则，指定编码格式及编码长度。

9.3.3 主数据模型

主数据模型要求如下：

- a) 应提供主数据模型及类目管理功能，能够根据业务域及自定义分类方式对模型进行分类管理；
- b) 应支持对模型基础信息和模型属性信息进行管理，包括但不限于创建、编辑、删除、启停、查询等；
- c) 应提供主数据模型批量导入导出功能；
- d) 应提供主数据模型在线加工功能，能够建立数据源表和目的表直接的关系，进行数据关系映射；
- e) 应支持多种映射规则，包括但不限于垂直拆分、水平拆分、多表联合、多表连接、自定义映射、主数据比对等。

9.4 数据融合

数据融合是指集成多个数据源以产生比任何单独的数据源更有价值信息的过程。数据融合方式包括感知融合、特征融合和决策融合，要求如下：

- a) 感知融合：应支持对不同来源的感知数据进行关联性统计分析，进行原始数据融合；
- b) 特征融合：应支持从不同来源的感知数据中提取有代表性的特征，这些特征融合成单一的特征矢量；
- c) 决策融合：应支持在特征层融合之后，对提取出的特征矢量进行联合判断和处理，从而得出对观测目标的一致性结论。

9.5 数据互操作

数据互操作是指在不同系统或平台之间进行信息交换和共享的能力，要求如下：

- a) 应按照统一通用的数据格式，将不同来源的数据进行解析和处理；
- b) 应按照统一的目录和规范，将不同来源、不同格式的数据进行统一索引；
- c) 应支持不同数据空间的数据相互集成，支持同步数据资源目录，实现空间内和跨空间的数据资源快速发现与访问；
- d) 应构建通信协议和协议更新机制，包括但不限于输入参数、输出结果、调用方式、传输方式和错误处理机制等关键要素。

10 管理

10.1 元数据管理

10.1.1 元数据类型

元数据类型是指描述数据元素的属性的种类，应符合GB/T 18391.2规定的要求。元数据类型帮助标准化和规范化不同系统间的数据交换和共享，保障元数据的一致性和可比性。元数据类型描述如下：

- a) 基本元数据：描述数据元素的基本属性，如数据的类型、结构和格式，保障数据表现形式的一致性；
- b) 复合元数据：将多个基本元数据组合成一个结构化的单元，适用于描述如文档、表单、记录、对象等复杂信息；
- c) 分类元数据：通过将数据元按照功能、用途、领域或其他特征进行分组，帮助用户理解不同数据元之间的关系，并方便在不同应用系统之间共享。

10.1.2 元数据归集

元数据归集是指获取数据元数据并对元数据进行组织、存储和管理的过程，涵盖城市智能中枢建设运营的数据全生存周期，包括配置归集任务、监控归集任务等。元数据归集要求如下：

- a) 应支持多种元数据归集方式，包括但不限于实时归集、定时归集、手工导入；
- b) 应支持自动归集，通过元数据管理工具的相关适配器进行归集，适配器可适配各类数据库、ETL、数据仓库和报表产品，以及各类结构化或半结构化数据源，包括但不限于：关系型数据库、MPP数据库、大数据、文件服务器、接口、搜索引擎、NoSQL数据库；
- c) 应支持根据元模型动态生成元数据导入模板，自动适配不同类型的数据源，简化数据的采集和标准化过程；
- d) 应支持内置归集数据源的元模型组合；
- e) 应支持根据元数据模型或元数据自动识别和自动关联关系；
- f) 应支持云服务方式；
- g) 应支持监控元数据归集进度；
- h) 应支持归集完成后将元数据组织成符合公共仓库元模型（CWM）的结构，将元数据存储于关系型数据库中；
- i) 应支持数据源更新时同步更新元数据；
- j) 宜支持归集数据的基本信息、空间参考信息、质量评价信息、维护信息、限制信息、生产信息、分发信息等元数据；
- k) 各元数据项应逐项填写内容，不应有空项。

注：有值时应如实记录；无值时应记为“无”（数值型记录“0”）；值未知时应记为“未知”（数值型记录“0”）。

10.1.3 元数据查询

元数据查询是指对已归集的元数据进行搜索和查看的过程，要求如下：

- a) 基于关键字的查询：用户可通过输入与元数据字段相关的关键字（如数据集名称、主题分类、地理范围等）进行搜索；
- b) 基于时间和空间的查询：用户可根据数据的时间维度（例如某个时间段的数据）或空间维度（例如某个特定地区的数据）进行查询；
- c) 基于分类体系的查询：用户可通过导航层次结构进行检索；
- d) 查询结果应以用户易于理解的方式呈现，包含相关元数据字段信息时应支持按照关键字进行全文搜索；
- e) 应支持查看元数据详细信息，如元数据基本信息、空间参考信息、质量评价信息等。

10.1.4 元模型管理

元模型管理是指元模型创建、维护、更新和应用的过程，包括元模型的创建与定义、发布与授权、维护与更新、应用与集成以及文档与培训等，要求如下：

- a) 应符合 GB/T 42093.1-2022 规定的要求；
- b) 应支持概念模型、逻辑模型、物理模型的归集和管理；
- c) 应完整地描述标准文档的结构和内容，包括层次结构、要素表述形式和内容要素等方面；
- d) 应与不同标准文档中的应用保持一致性，保障元模型的通用性和可复用性；
- e) 应支持在审核中更改缺乏语义描述、缺少字段、类型错误、编码缺失或不可识别的字符编码等问题，支持模型的创建、更新、查询、维护等操作；
- f) 应支持元数据版本管理，确定元数据版本并发布基线版本，可实现对存异的或错误的元数据进行追溯、检查和回复；
- g) 应进行动态调整和维护，实现元模型与标准文档的实际情况保持一致。

10.1.5 元数据存储

元数据存储要求如下：

- a) 一致性和完整性：应保障数据的一致性和完整性，当元数据进行更新时，相关的引用、依赖关系和索引不应被破坏。元数据库应支持数据库管理系统、文件系统等不同物理存储方式；
- b) 版本控制：应支持对元数据的存储进行版本控制，以便跟踪不同版本的元数据的变化；
- c) 权限控制 and 安全性：应保障安全性和访问控制，对元数据的存取设定不同的权限，保障敏感数据的安全。

10.1.6 数据血缘

元数据血缘包括血缘管理、血缘检索、血缘统计等，要求如下：

- a) 应支持手工管理元数据血缘关系，包括新增、修改、删除等操作；
- b) 应支持根据数据归集和数据处理任务自动解析形成元数据血缘关系功能；
- c) 应提供库、表、字段级血缘关系功能；
- d) 应支持基于字段级血缘关系的数据质量追踪；
- e) 应提供以可视化数据流向关系呈现元数据血缘功能；
- f) 应提供查看单个元数据血缘和影响分析链路图功能；
- g) 应支持字段级元数据关系构建，实现图形化的血缘分析；
- h) 应支持数据血缘关系识别，包括数据来源、传输过程、处理过程等；
- i) 应提供统计无血缘关系的孤立元数据信息功能；
- j) 应支持以历史事实的方式记录每项数据的来源、处理过程、应用对接情况；
- k) 应支持从数据表之间的血缘、数据库之间的数据流向血缘、数据处理任务血缘、工作流血缘等多个方面，记录数据表在数据治理过程中的全链血缘关系；
- l) 应支持对数据开发的全过程追踪，实现以某个数据对象为起点的所有与该对象相关的元数据和关联关系追踪；
- m) 应支持采用自动解析方法归集数据血缘，可采用系统跟踪、机器学习、手工收集等多种方式相配合；
- n) 描述血缘关系时应采用可直观表示数据使用的链路形态，描述每个节点应包括名称、类型、状态等信息，以及包含关系、上游名称、类型、状态等上下游节点信息；
- o) 宜支持统计元数据重要程度使用热度。

10.2 目录管理

10.2.1 目录编制

目录编制要求如下：

- a) 应提供定义层级目录功能；
- b) 应提供导入导出目录功能；
- c) 应提供调整目录展示顺序功能；
- d) 目录服务应支持多用户模式，支持不同服务使用机构申请逻辑独立的目录服务；
- e) 应提供目录在线编制、导入编制等编制功能；
- f) 目录编制后，应支持将目录和对应的数据库表关联，并赋予唯一的目录标识符和编码；
- g) 应支持自动目录生成，包括积累式生成和增量式生成等；
- h) 应支持通过元数据的管理，自动实现目录的生成与使用；

- i) 应支持自动建立不同类型的目录，包括但不限于数据归集对象目录、数据存储目录、数据服务目录、数据产品目录等；
- j) 应提供目录缓存功能，快速响应服务使用机构的使用需求；
- k) 应支持展示目录之间的关联关系；
- l) 应支持通过图形化的拖拉拽方式配置目录模板；
- m) 应提供目录和资源的版本管理功能。

10.2.2 目录资源挂接

目录资源挂接要求如下：

- a) 应支持对不同类型数据进行资产编目，包括但不限于文件、库表、接口、消息管道等；
- b) 应支持基于归集元数据实现资产盘点，自动识别资产、补充资产信息并进行编目；
- c) 应提供目录注册接口，支持数据服务方将数据产品、服务、应用等上架到目录；
- d) 应支持对数据和服务资源元数据自动审核并进入元数据库，形成目录；
- e) 应提供已编目资源清单的导入导出功能；
- f) 应支持通过目录的上报、下发、申请、审批、更新、撤销等，将数据资源目录在纵向及横向级联；
- g) 应提供资源生存周期管理功能，包括但不限于新增、送审、变更、停用、启用、删除。

10.2.3 目录属性

资源目录属性主要包括基本属性、业务要素、管理要素和价值要素，描述如下：

- a) 基本属性：包括数据的来源、类型、结构、规模、更新周期、标准和质量等；
- b) 业务要素：包括业务描述、业务指标、业务规则和关联关系等；
- c) 管理要素：包括数据权属、分类分级、安全信息、数据溯源、职责权限和应用情况等；
- d) 价值要素：包括市场信息、领域信息、地域信息和应用价值等。

10.2.4 目录发布

目录发布要求如下：

- a) 应支持将已编目好的资源发布至资产门户或共享系统，对外提供目录信息服务；
- b) 应支持能够自定义设置特定部分的元数据是否可对外提供服务；
- c) 应支持目录的查询、核查、更新、同步等目录服务，各级用户按照权限查看数据资源目录；
- d) 应支持根据数据资源目录相关属性和数据项进行数据资源查询；
- e) 应支持自动化审核，对发布的数据资源进行质量检查和合规性验证；
- f) 应支持目录智能搜索，搜索条件包括但不限于关键词、分类、标签等；
- g) 应支持数据提供方设置资源分级分类权限，限制资源发布对象范围，权限可授权对象包括但不限于部门、角色、用户等。

10.2.5 目录更新

数据目录更新要求如下：

- a) 应支持通过多级审批方式，对数据资源名称、数据来源、数据资源摘要、数据资源更新周期、数据规模、数据存储周期、共享范围、服务种类服务提供方式、服务系统名称、服务系统地址等相关属性及详细数据项列表等信息进行更新；
- b) 应支持将已发布的资源下线，下线后的资源在资产门户不可见；
- c) 应支持对审批通过后的目录进行删除、修改，并对修改后的目录进行审批；

- d) 应提供流程管理功能，支持自定义审批流程、管理流程设计，包括流程的创建、复制、编辑、导入、导出、删除等。

10.3 资源管理

10.3.1 资源注册

资源注册包括系统注册、资源登记、资源导入等，要求如下：

- a) 应提供注册资源所属业务系统功能，系统属性包括但不限于：系统名称、事权部门、管理部门、系统在用标识、系统访问地址、系统管理员、管理员联系方式、承载网络、承建单位、承建单位联系人、上线日期、下线日期、系统描述；
- b) 应提供注册不同数据类型资源功能，包括但不限于：文件、库表、接口、消息管道；
- c) 应支持多种资源注册方式，包括但不限于：元数据注册、导入注册、手工新增注册、复制注册；
- d) 应提供统计业务系统所属资源数量功能；
- e) 应提供按业务系统视角检索资源功能。

10.3.2 数据预览

数据预览是指对结构化数据资源样例数据进行查看，要求如下：

- a) 应提供配置数据脱敏规则功能，预览数据为脱敏后数据；
- b) 应提供零代码筛选样例数据功能。

10.3.3 版本管理

版本管理是指对资源历史变更记录的维护管理，要求如下：

- a) 应提供查询资源历史变更版本列表功能，并支持查看单个历史版本详情；
- b) 应提供选择任意两个版本比对变更差异点功能；
- c) 应提供选择任意历史版本属性恢复到当前版本功能。

10.4 任务管理

10.4.1 基本要求

数据生存周期管理按任务方式进行管理，基本要求如下：

- a) 应支持对输入的多源异构数据进行全量或增量的整合加工，处理过程应包括数据提取、清洗、转换、分发、归档、处置等；
- b) 应支持按任务方式设置数据处理任务，包括任务管理、处理监控、配置库管理等；
- c) 应支持通过数据表的列与标准数据元关联、数据表的列与数据字典关联、物理模型与逻辑模型关联等；
- d) 应支持结构化比对、非结构化比对、结构化和非结构化融合比对、关键词比对、二进制比对等方式；
- e) 宜支持通过图形化界面对标识规则进行加载、解析、编译和执行；
- f) 宜支持标识任务自动生成、智能化调度对标识规则进行管理。

10.4.2 作业管理

作业管理包括任务开发、任务调试、作业编排、作业调度等，要求如下：

- a) 应提供以组件化方式开发基本的数据加工和计算任务功能，包括但不限于数据提取、数据清洗、数据关联、数据比对、数据标识、数据分发；

- b) 应提供自定义数据处理脚本任务功能;
- c) 应支持自定义处理函数管理;
- d) 应支持离线处理和实时处理;
- e) 应提供在线调试任务功能;
- f) 应提供以可视化方式将多个数据处理任务编排成 workflow 功能, 编排方式包括但不限于串行、并行;
- g) 应支持手动执行、定时执行、事件依赖等任务调度方式;
- h) 应支持设置任务运行参数, 包括但不限于任务参数、环境变量、系统参数;
- i) 应支持管理任务生存周期, 包括但不限于启动、停止、删除、重新运行;
- j) 应支持失败任务的重跑;
- k) 应支持任务导入和导出。

10.4.3 处理监控

处理监控是指对已经提交调度的任务及历史执行情况进行监控, 包括任务运行状态监控、任务执行日志分析、异常任务告警、补数据等, 要求如下:

- a) 应提供查看任务列表功能;
- b) 应提供任务实例的基本运维操作功能, 包括但不限于启动、停止、暂停、恢复、重跑;
- c) 应提供查看任务的实时日志、历史日志与任务配置功能;
- d) 应支持监控任务运行状态;
- e) 应提供设置任务的异常告警条件、告警方式、告警对象功能;
- f) 应提供运维监控概览功能, 概览内容包括但不限于查看实例执行总体情况、运行时长最长的任务列表、失败次数最多的任务列表;
- g) 应提供任务运行失败原因诊断功能, 并支持给出调整建议。

10.4.4 配置库管理

配置库是指存储处理任务、脚本元数据的数据库, 应支持配置库构建、连接测试等。

11 安全

11.1 数据加密保护

依据国家信息系统密码应用要求, 从物理和环境安全、网络和通信安全、设备和计算安全、应用和数据安全四个层面明确等级保护要求中不同级别的密钥管理和安全管理要求, 等级保护要求应不低于三级, 具体应符合 GB/T 39786-2021 对应密码应用级别和 GB/T 22239-2019 对应等级级别的要求。

11.2 敏感数据识别

敏感数据识别要求如下:

- a) 对发现的数据资产应支持采用人工方式进行敏感数据的梳理和识别;
- b) 应支持基于用户指定或预定义的敏感数据及特征, 通过自动化程序主动识别敏感数据, 如通过关键字匹配、正则表达式等方式;
- c) 应支持敏感数据识别策略和敏感数据脱敏策略, 并进行管理和维护;
- d) 应支持个人隐私数据识别策略, 个人隐私数据包括但不限于姓名、身份证、家庭住址、电话、银行账号、电子邮箱等;

- e) 应支持用户自定义添加新的识别策略。

11.3 数据脱敏

数据脱敏包括但不限于以下要求：

- a) 应支持敏感数据的脱敏规则，包括但不限于数据混淆、替换、映射、遮蔽、加密和解密、随机、可逆脱敏算法等；
- b) 应提供对脱敏策略和脱敏规则的管理功能；
- c) 应提供自定义敏感数据类型发现规则功能，包括但不限于正则发现和关键字发现两种规则，并提供在线校验功能，对字符串进行分段设置；
- d) 应提供静态脱敏功能，能够通过全量及增量方式、手动及定时方式执行静态脱敏任务；
- e) 应提供动态脱敏功能；
- f) 应支持用户自定义配置脱敏策略；
- g) 个人信息保护及去标志化可参照 GB/T 37964-2019。

11.4 数据分类分级

应按照GB/T 43697-2024对数据进行分类分级并根据不同级别实施相应的数据安全防护措施。

11.5 数据溯源

数据溯源是指追踪记录数据全生存周期，记录各环节的过程信息，要求如下：

- a) 应建立数据溯源机制，如通过审计日志等方式在采集、审核校验、更正等流程中进行唯一性标识；
- b) 应构建统一的审计日志格式，字段包括但不限于时间戳、事件类型、操作来源等，保障日志信息的完整性；
- c) 应支持日志自动采集，从系统和应用程序中收集日志信息；
- d) 应支持采用区块链等技术，实现日志数据的不可篡改；
- e) 应建立数据标识和映射关系，赋予数据唯一的溯源标识；
- f) 应详细记录数据在各环节的操作信息，构建完整的数据溯源链；
- g) 应集成合规性检查机制，保障数据流转过程符合法律法规和内部政策，实现异常监测和告警；
- h) 应构建数据溯源可视化工具，提供直观的数据流转视图，便于用户追踪和理解。

参 考 文 献

- [1] GB/T 5271.17-2010 信息技术 词汇 第17部分：数据库
 - [2] GB/T 36344-2018 信息技术 数据质量评价指标
 - [3] GB/T 37043-2018 智慧城市 术语
 - [4] GB/T 37964-2019 信息安全技术 个人信息去标识化指南
 - [5] GB/T 42450-2023 信息技术 大数据 数据资源规划
 - [6] GB/T 44109-2024 信息技术 大数据 数据治理实施指南
 - [7] GA/T 2202-2024 公安大数据处理 数据接入 数据对账技术要求
-