

DBJT45

广西壮族自治区交通运输行业指南

DBJT45/T 052—2023

交通运输政务数据治理技术指南

Technical guide for transport government data governance

2023-04-09 发布

2023-04-30 实施

广西壮族自治区交通运输厅 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 总则 2

5 数据框架与治理范围 2

5.1 数据框架 2

5.2 治理范围 2

6 数据标准维护管理 3

6.1 维护管理内容 3

6.2 技术要求 3

6.2.1 数据标准录入 3

6.2.2 审核与发布 3

6.2.3 数据标准映射 3

6.2.4 数据标准评估 4

6.2.5 数据标准维护 4

6.2.6 数据标准查询 4

7 元数据管理 4

7.1 基本内容 4

7.2 技术要求 4

7.2.1 元数据采集 4

7.2.1.1 适配器管理 4

7.2.1.2 数据源管理 4

7.2.1.3 采集任务配置 4

7.2.1.4 采集日志管理 4

7.2.2 元数据采集监控 4

7.2.3 元数据存储 5

7.2.4 元模型管理 5

7.2.5 元数据查询 5

7.2.6 元数据分析 5

8 数据质量审计 5

8.1 审计规则管理 5

8.1.1 审计规则 5

8.1.2 审计规则管理 6

8.2 审计任务管理 6

8.2.1 审计任务要求 6

8.2.2 审计任务执行 6

8.2.3 审计任务监控 6

8.3 审计结果管理..... 6

 8.3.1 结果展示..... 6

 8.3.2 例外管理..... 6

 8.3.3 数据质量统计管理..... 7

 8.3.4 审计报告管理..... 7

9 数据质量提升..... 7

 9.1 事前预防..... 7

 9.2 事中监控..... 7

 9.3 事后改善..... 7

10 安全要求..... 7

参考文献..... 8

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。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由广西壮族自治区交通运输厅提出并宣贯。

本文件由广西交通运输标准化技术委员会归口。

本文件起草单位：广西交通设计集团有限公司、广西壮族自治区交通运输信息管理中心。

本文件主要起草人：罗海宇、李道飞、杨凯、韦扬、王长海、陈成伟、李世锋、陈猷辉、邓定丰、刘鹏博、周敏璐、刘海南、周开制、黄中章、岑汝喜、韦宇飞、刘亮、覃薇、张昕、陈科先、陆海鹏、吴宇昊、胡文君、黄梦凡、姚西桐。

本文件主要审查人：李文勇、张小健、李琳、姚育章、吴川、杨晓东、王勇。

交通运输政务数据治理技术指南

1 范围

本文件界定了交通运输政务数据治理技术的术语和定义,规定了交通运输政务数据治理技术的总则、数据框架与治理范围、数据标准维护管理、元数据管理、数据质量审计、数据质量提升、安全要求。

本文件适用于广西壮族自治区行政区域内公路、水路、道路运输政务数据的治理。

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 34960.1 信息技术服务 治理 第1部分:通用要求
GB/T 34960.5 信息技术服务 治理 第5部分:数据治理规范
GB/T 36344 信息技术 数据质量评价指标
GB/T 37378-2019 交通运输 信息安全规范
JT/T 697 (所有部分) 交通信息基础数据元

3 术语和定义

下列术语和定义适用于本文件。

3.1

交通运输政务数据 transport government data

由交通运输政务部门或者为交通运输政务部门采集、加工、使用、处理的数据。包括交通运输政务部门依法采集的数据、交通运输政务部门在履行职能过程中产生和生成的数据、交通运输政务部门投资建设和外购服务获取的数据、交通运输政务部门依法授权管理的数据等。

3.2

数据元 data element

通过定义、标识、表示以及允许值等一系列属性描述的数据单元。在特定的语意环境中被认定为不可再分的最小数据单元。

3.3

数据标准 data standard

数据的命名、定义、结构和取值范围方面的规则和基准。

3.4

数据治理 data governance

数据资源及其应用过程中相关管控活动、绩效和风险管理的集合。

3.5

数据质量 data quality

在指定条件下使用时,数据的特性满足明确的和隐含的要求的程度。

3.6

数据管理 data management

数据资源获取、控制、价值提升等活动的集合。

3.7

数据安全 data security

采用技术和管理措施来保护数据的保密性、完整性和可用性等。

4 总则

- 4.1 数据质量的评估衡量应符合 GB/T 36344 的要求，主要有规范性、完整性、准确性、一致性、时效性、可访问性 6 类指标。
- 4.2 本文件涉及的技术架构，宜符合 GB/T 34960.1、GB/T 34960.5、《交通运输政务信息资源目录编制指南（试行）》、《交通运输信息资源交换共享与开放应用平台省级工程建设指南》的规定。
- 4.3 本文件涉及交通运输政务数据治理的要求，参照《广西政务数据资源管理与应用改革实施方案》、《广西政务数据“聚通用”实施细则（试行）》、《广西政务数据治理实施方案》的要求开展。

5 数据框架与治理范围

5.1 数据框架

交通运输政务数据框架包括数据资源池、数据治理、资源目录、数据共享、数据应用、数据安全等内容。本文件仅对数据治理做规范化要求，如图 1 所示。

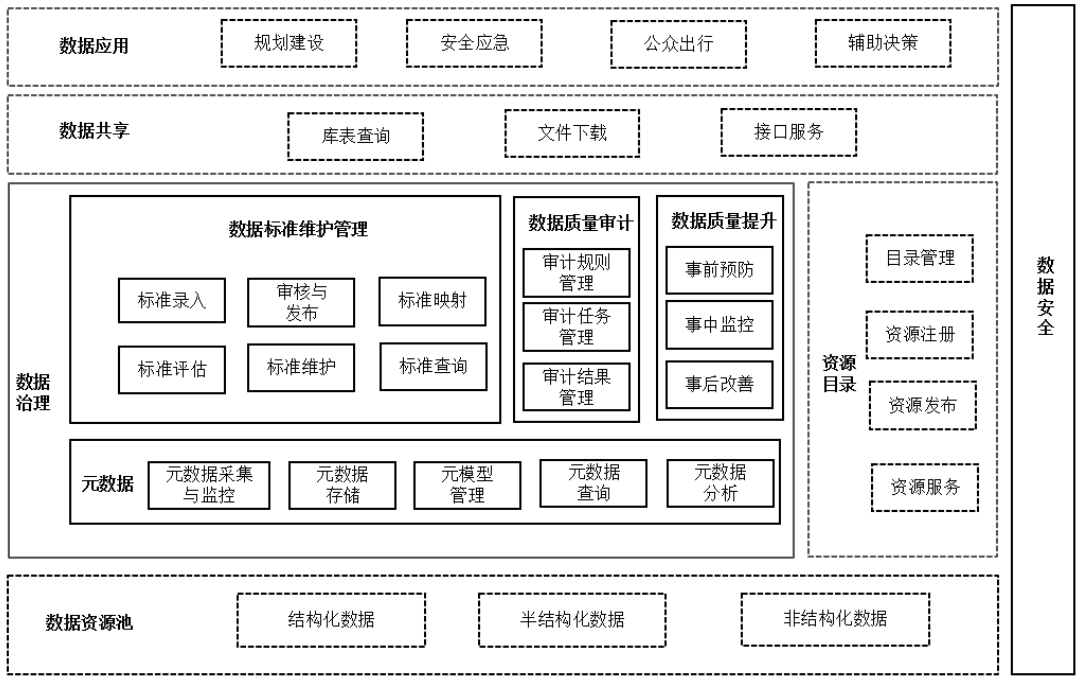


图1 交通运输政务数据治理框架

5.2 治理范围

交通运输政务数据治理范围包括数据治理平台域、数据治理实施域，具体要求如下：

- 数据治理平台域应包括数据治理平台框架、平台应用、平台维护，并选择合理的平台工具；
- 数据治理实施域宜包括数据标准维护管理、元数据管理、数据质量管理等专题实施工作。

6 数据标准维护管理

6.1 维护管理内容

数据标准维护管理的主要内容应包含数据标准录入、审核与发布、标准维护、标准查询、标准映射与评估等，见图2。

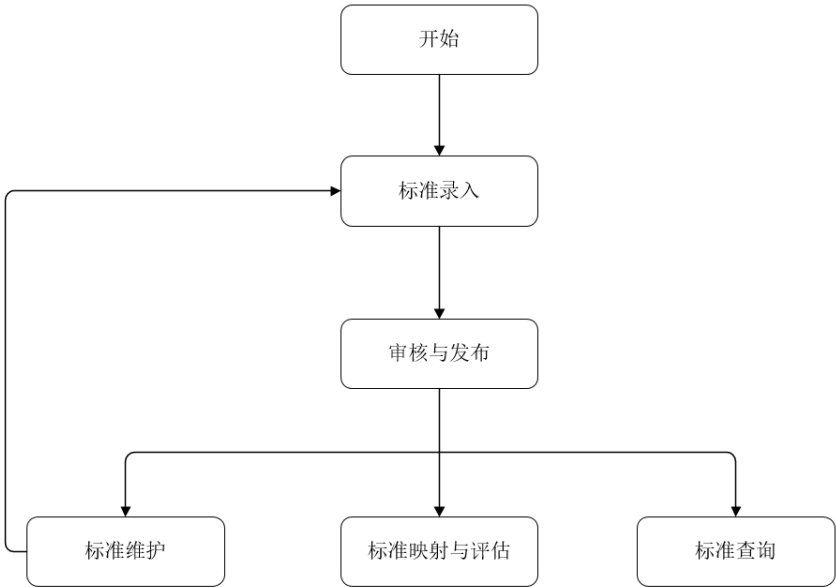


图2 数据标准维护管理内容

6.2 技术要求

6.2.1 数据标准录入

数据标准录入的要求如下：

- 遵照国家标准、行业标准、地方标准的内容和要求，以及数据标准业务性质，提供信息实体、基础数据元、基础代码集、术语定义及其相关属性的录入；
- 基础数据元应符合 JT/T 697（所有部分）的要求，同时预留出兼容其他相关国家标准、行业标准、地方标准等要求的字段；
- 数据标准录入应支持列表自动导入和人工录入功能。

6.2.2 审核与发布

数据标准录入后，应根据实际管理要求，设置审核流程，审核人员通过后，经过主管部门的审批后可予以发布。

6.2.3 数据标准映射

数据标准映射的要求如下：

- 实现数据标准与元数据的关联与映射，包括元数据关联标准和标准关联元数据两种方式；

——数据标准与元数据映射前，应充分搜集元数据相关的国家标准、行业标准、地方标准，并全部录入到标准管理系统中，实现与元数据的映射。

6.2.4 数据标准评估

应包括强制落标和非强制落标。评估内容应包括评估属性、标准编号、标准值、元数据路径、元数据实际值、评估结果等。

6.2.5 数据标准维护

6.2.5.1 交通运输数据标准出现新增、变更、废止等情况时，应及时完成标准的录入、变更、删除等操作，应注意与数据采集端的标准执行同步，确保标准的准确性。

6.2.5.2 数据标准管理信息系统应支持对标准进行备份与恢复，保证数据标准与数据的版本对应。

6.2.6 数据标准查询

系统支持通过标准编号、标准中文名称、标准英文名称、标准所属类别对已发布的标准进行查询。

7 元数据管理

7.1 基本内容

元数据管理基本内容包括元数据采集、存储、查询、分析、监控、元模型管理等内容。

7.2 技术要求

7.2.1 元数据采集

7.2.1.1 适配器管理

元数据采集应支持常见的适配器，包括：

- 关系型和非关系型数据库；
- ETL 工具；
- 对象和文件存储。

7.2.1.2 数据源管理

数据源类型应覆盖适配器中支持的所有类型的数据源。

7.2.1.3 采集任务配置

在配置采集任务之前，根据数据量和数据特点有序、分步骤开展元数据采集。采集任务配置应包括数据源名称、采集方式、入库策略、执行周期等内容。

7.2.1.4 采集日志管理

元数据采集应同步记录采集任务的执行日志。采集日志应包括任务名称、任务描述、数据源名称、任务状态、开始时间、结束时间、耗时、日志详情等内容。

7.2.2 元数据采集监控

元数据采集监控的指标应包含但不限于表1中的内容。

表1 元数据采集任务监控项说明表

监控项	说明
采集任务编号	每个采集任务均赋予唯一标识
任务名称	采集任务的名称，要求唯一性
任务描述	简要描述采集任务的相关信息
数据源名称	采集任务的数据源名称
任务开始时间	采集任务的开始时间，格式为“yyyy-mm-dd hh:mm:ss”
任务结束时间	采集任务的结束时间，格式为“yyyy-mm-dd hh:mm:ss”
任务调度周期	采集任务调度的周期
最近一次执行结果	采集任务最近一次的执行结果。成功，取值Y；失败，取值N
任务运行日志	记录采集任务每次调度的详细日志信息

7.2.3 元数据存储

元数据存储在业务逻辑上应满足多版本管理的要求。存储中应设置标签区分元数据版本。技术逻辑上应按照结构化数据的存储方式，宜选择关系数据库存储。

7.2.4 元模型管理

交通运输政务数据元模型应包括关系型数据库和非关系型数据库、ETL工具、对象和文件系统等类型。关系型数据库应包括库、表空间、表、视图、字段、主键、外键、函数、存储过程、触发器等模型，具体模型属性以满足实际治理需求为准。其他类型的元模型应根据实际技术特点支持技术人员定义。

7.2.5 元数据查询

对元数据名称、代码、类型、版本、创建时间等元数据的查询。

7.2.6 元数据分析

- 元数据分析包括血缘分析和影响分析：
- 血缘分析。针对数据采集（产生）的元数据、数据存储的元数据、数据加工处理的元数据、数据流转交换的元数据依次记录，全面反映当前数据对象的血缘关系；
 - 影响分析。依据当前数据对象与下游数据对象的关系，模拟元数据变更后对下游数据的影响。

8 数据质量审计

8.1 审计规则管理

8.1.1 审计规则

- 审计规则类型包括：
- 技术规则：包括数据库是否可访问、数据类型与数据标准规定是否匹配、数据长度和精度与数据标准规定是否匹配、取值是否存在空值、非结构化数据的索引是否空值等；
 - 业务规则：包括车辆号牌位数约束、船检编号规范约束、车辆类型规范约束、路线基础数据编码规范约束、高速公路车辆通行费金额值域约束等。若存在国家标准、交通运输行业标准、地方标准或交通运输部、自治区有相应的规范与要求，应满足其要求；
 - 其他规则：通过政务服务、其他部门多源比对等方式设定规则，对数据质量进行校核。

审计规则内容应包含规则名称、规则代码、主题表、检查字段、规则描述、启用状态等信息。

8.1.2 审计规则管理

审计规则管理包括：

- 新建规则：包括新建单条规则、批量新建规则，通过标准新建规则。新建规则时应对规则进行分组；
- 修改规则：包括修改规则名称、规则代码、主题表、检查字段、规则描述、启用状态等信息；
- 查询规则：按照规则名称、规则代码、描述查询规则；
- 删除规则：包括单条删除和批量删除。

8.2 审计任务管理

8.2.1 审计任务要求

将规则按照实际业务逻辑配置成相应的审计任务，进行相关数据质量的审计。

8.2.2 审计任务执行

应按照一致性、完整性、准确性、非冗余性、可访问性、及时性、可用性等原则审计数据质量问题。

8.2.3 审计任务监控

审计任务监控应至少包括表2中的监控内容。

表2 审计任务监控项说明表

监控项	说明
审计任务编号	每个审计任务均赋予唯一标识
任务名称	审计任务的名称，要求唯一性
数据集名称	审计任务对应的数据集
任务开始时间	任务的开始时间，格式为“yyyy-mm-dd hh:mm:ss”
任务结束时间	任务的结束时间，格式为“yyyy-mm-dd hh:mm:ss”
任务调度周期	任务调度的周期
数据量	审计任务扫描的数据量
错误数	不满足规则的错误数量
任务执行人	审计任务的执行人
最近一次执行结果	审计任务最近一次的执行结果。成功，取值Y；失败，取值N
任务运行日志	记录审计任务每次调度的详细日志信息

8.3 审计结果管理

8.3.1 结果展示

质量审计结果应以图表、数据的形式展示。

8.3.2 例外管理

当数据审计规则出现特例情况时应予以标注，亦可取消对例外情况的标注。

8.3.3 数据质量统计管理

数据质量审计结果宜多维度进行分析统计。

8.3.4 审计报告管理

审计结果应生成详细的数据质量审计报告。审计任务管理应采用信息化管理技术，管理系统审计结果导出包括分类（如按业务系统、数据所属机构等）导出和整体导出。导出形式应支持 Excel、Html 和 XML 格式。

9 数据质量提升

9.1 事前预防

在数据生产器（即信息系统、传感器等）建设过程中，应基于第6章数据标准维护管理技术，约束各类数据生产器的设计与研发。

9.2 事中监控

事中监控包括以下环节：

- a) 质量监控：应利用第 8 章中数据质量审计技术，对数据源、采集工具（ETL 等）进行审计监测；
- b) 问题定位：应利用第 7 章中元数据血缘分析技术，定位数据质量问题发生的环节和具体位置；
- c) 质量提升：应引入闭环管理流程，引入工单流程引擎技术，提升数据质量。

9.3 事后改善

应在数据质量事中提升后，结合数据血缘和影响分析，形成体系化的数据质量提升措施，实现数据流转上、下游协同，提升数据质量。

10 安全要求

交通运输政务数据治理技术中用户终端、载运装备、基础设施、计算中心、网络与通信等，应符合 GB/T 37378-2019 中第5章交通运输信息系统安全技术体系架构和第6章通用技术的要求。

参 考 文 献

[1]中华人民共和国交通运输部办公厅 交办科技〔2017〕123号 关于印发《交通运输政务信息资源目录编制指南（试行）》的通知

[2]中华人民共和国交通运输部办公厅 交办科技〔2017〕124号 关于印发《交通运输信息资源交换共享与开放应用平台省级工程建设指南》

[3]广西壮族自治区党委办公厅、广西壮族自治区人民政府办公厅 厅发〔2019〕141号 关于印发《广西政务数据资源管理与应用改革实施方案》的通知

[4]广西壮族自治区数字广西建设领导小组办公室 桂数广办发〔2019〕13号 关于印发《广西政务数据“聚通用”实施细则（试行）》的通知

[5]广西壮族自治区数字广西建设领导小组办公室 桂数广办发〔2020〕13号 关于印发《广西政务数据治理实施方案》的通知
