

ICS 35.240.50

CCS L70

**DB21**

辽宁省地方标准

DB21/TXXXX—XXXX

# 工业数据流通 数据质量管理规范

(征求意见稿)

XXXX-XX-XX 发布

XXXX-XX-XX 实施

辽宁省市场监督管理局 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 总体要求.....	1
5 职责要求.....	2
6 数据质量管理指标.....	2
7 质量管理流程.....	3
7.1 概述.....	3
7.2 数据采集.....	3
7.3 数据审核.....	3
7.4 数据校验.....	3
7.5 数据更正.....	4
8 质量管理工具.....	4
8.1 数据质量管理规则.....	4
8.2 数据质量检测任务.....	4
8.3 数据质量监控.....	4
8.4 数据质量报警.....	5
8.5 数据质量检测报告.....	5
8.6 数据质量检测报告分析.....	5
8.7 数据质量整改.....	5
8.8 数据质量跟踪.....	5
参考文献.....	6



## 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。

本文件由辽宁省工业和信息化厅提出并归口。

本文件起草单位：沈阳华睿博信息技术有限公司等。

本文件主要起草人：邵华等。

本文件发布实施后，任何单位和个人如有问题和意见建议，均可以通过来电和来函等方式进行反馈，我们将及时答复并认真处理，根据实际情况依法进行评估及复审。

归口管理部门通信地址：辽宁省沈阳市皇姑区北陵大街45-2号。

归口管理部门联系电话：024-86913384。

标准起草单位通讯地址：辽宁省沈阳市和平区青年大街386号华阳国际大厦2396。

标准起草单位联系电话：18698849086。



# 工业数据流通 数据质量管理规范

## 1 范围

本文件给出了工业数据流通过程中数据质量的总体要求和职责要求，规定了数据质量管理指标、质量管理流程、质量管理工具等方面的要求。

本文件适用于工业数据流通过程中对工业数据进行质量管理的工作。

## 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 36344-2018 信息技术 数据质量评价指标

## 3 术语和定义

下列术语和定义适用于本文件。

### 3.1

**工业数据** industrial data

在工业领域中，涉及企业的所有生产活动和服务所产生的数据。

[来源：GB/T 42128-2022, 3.1.1]

### 3.2

**数据质量** data quality

数据的一组固有特性满足要求的程度。

注：固有特性一般指永久性的特性。

[来源：GB/T 39400-2020, 3.1]

### 3.3

**数据质量管理** data quality management

指导和控制某机构数据质量的协调活动。

[来源：GB/T 39400-2020, 3.2]

### 3.4

**数据质量管理工具** data quality management tool

运用信息化技术和管理方法来规划、实施与控制数据质量的识别、评估、提升等一系列活动。

## 4 总体要求

工业数据流通过程中，相关单位应按本文件的要求建立数据质量管理体系，编写出台相应的文件，在执行过程中应对其有效性进行持续改进，实现数据质量的持续提高。具体包括：

- a) 识别数据质量管理工作覆盖的范围，确定数据质量管理涉及的环节和流程；
- b) 确定这些环节和流程的顺序和相互关系；
- c) 确定所需采取的准则和方法，以确保这些环节和流程的运行及控制有效；
- d) 确保管理职责明晰、保障机制健全，以支持这些环节和流程的运行和监控；
- e) 对这些环节和流程进行分析、监督和考核；
- f) 采取必要的措施，以实现数据质量管理目标和对这些环节和流程的持续改进。

### 5 职责要求

数据质量管理者应对本单位与数据质量管理相关的工作作出规定，具有以下方面的职责：

- a) 指导推进、监督考核本单位工业数据流通、利用相关工作；
- b) 根据需要确保按照本文件的要求，建立、实施、保持并持续改进数据质量管理体系；
- c) 提出工业数据的数据目录规范、数据质量指标要求，制定数据质量管理规则，确保数据质量管理工作有明确的目标和评估标准；
- d) 向最高管理者报告数据质量管理情况和改进需求，提供数据质量监测和评估报告；
- e) 负责与数据质量管理工作有关的外部联系工作；
- f) 对其他部门提出的数据质量要求及时进行处理和反馈，定期执行数据质量评估。

### 6 数据质量管理指标

工业数据质量管理指标框架见图 1，应基于此框架对工业数据进行数据质量的符合性评价。

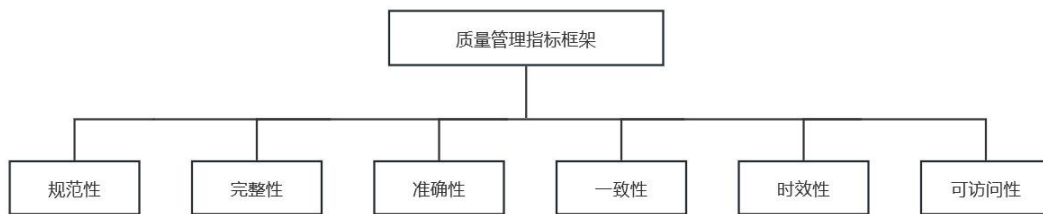


图 1 质量管理指标框架

各质量特性的说明如下：

- a) 规范性：数据符合数据标准、数据模型、业务规则、元数据、权威参考数据及安全规范的程度。
  - 1) 数据标准是数据的命名、定义、结构和取值规范方面的规则和基准。
  - 2) 数据模型是对分析的图像和文本表述，该分析识别了组织为完成其使命、功能、目标、目的和战略，以及管理和评价组织所需要的数据。
  - 3) 业务规则是一种权威性原则或指导方针，用来描述业务交互，并建立行动和数据行为结果及完整性的规则。
  - 4) 元数据是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。
  - 5) 权威参考数据是系统、应用软件、数据库、流程、报告或平台日志记录用来参考的特定字段的有效数据集合。
  - 6) 安全规范是安全和隐私方面的规则，包括数据权限管理、数据脱敏处理。
- b) 完整性：按照数据规则要求，数据元素被赋予数值的程度。即数据信息是否存在缺失的状况，包括数据元素完整性和数据记录完整性；



- c) 准确性：数据准确表示其所描述的真实实体(实体对象)真实值的程度，即数据记录的信息是否存在异常或错误，包括数据内容正确性、数据格式合规性、数据重复率、数据唯一性、脏数据出现率；
- d) 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，即数据是否遵循了统一的规范数据集是否保持了统一的格式，主要体现在数据记录的规范和数据是否符合逻辑，包括相同数据一致性和关联数据一致性；
- e) 时效性：数据在时间变化中的正确程度，包括基于时间段的正确性、基于时间点及时性、时序性；
- f) 可访问性：数据能被访问的程度，包括可访问和可用性。

数据质量管理过程中数据质量评价应按GB/T 36344-2018的规定进行。

注：由于工业产品生命周期同一阶段的数据具有强关联性，如产品零部件组成、工况、设备状态、维修情况、零部件补充采购等，工业产品生命周期的研发设计、生产、服务等不同环节的数据之间也需要进行关联；工业企业生产、经营流程具有较强时序性。因此，工业数据质量特性评价宜侧重于一致性和时效性。

## 7 质量管理流程

### 7.1 概述

工业数据质量管理应贯穿于工业数据流通的各个阶段，数据质量管理者应按照数据质量管理的要求，对影响数据质量的数据流程加以识别，制定管理措施，并实施流程控制，从数据采集源头开始梳理数据流向，识别影响数据质量的数据流程，关键流程至少应包括数据采集、数据审核、数据校验、数据更正等。

### 7.2 数据采集

数据质量管理者应明确数据采集流程的工序、内容和岗位职责，编制程序文件和相应的作业指导文件。在数据采集工作中应满足以下要求：

- a) 数据事权部门应制定基础数据采集的业务规则，并负责监督在各级各类相关系统中实现；
- b) 承担数据采集任务的操作人员应由数据事权部门培训考核后上岗，数据事权部门对于保证数据的规范性、完整性、准确性、一致性、时效性等质量特性具有指导的义务和责任；
- c) 基础数据应与相关标准所定义的数据项集合相吻合，以保证原始数据的完整性；
- d) 基础数据中的数据项应和数据项名称对应的标准数据代码相吻合，且符合相关标准规定的逻辑校验关系，并且同时符合数据项名称表征的客观实体，以保证原始数据的准确性；
- e) 基础数据中的数据项应和工业数据标准规定的的数据属性相吻合，以保证原始数据的规范性；
- f) 基础数据应和数据事权部门操作日志所记载的最末一次数据属性相吻合，以保证原始数据的时效性和一致性。

### 7.3 数据审核

数据质量管理者应明确数据审核流程的工序、内容和岗位职责，编制程序文件和相应的作业指导文件。在数据审核工作中满足以下要求：

- a) 数据事权部门应建立基础数据采集的审核程序，对录入系统的数据质量进行审核；
- b) 在基础数据的采集过程中，数据事权部门应采用抽验方式对所录入数据的质量进行人工或计算机检验，抽验样本数量应按一定比率进行抽取，抽验样本周期和抽验样本数据跨度应分别控制在规定时间之内；
- c) 对录入数据采用了计算机自动审核方式的，应建立相应的审核制度和规程；
- d) 应保持上述质量审核的全部证实性记录。

### 7.4 数据校验

数据质量管理者应明确数据校验流程的工序、内容和岗位职责，编制程序文件和相应的作业指导文件。在数据校验工作中满足以下要求：

- a) 数据事权部门应建立在线校验程序和相关校验规程，对在线数据规范性、完整性、准确性、一致性、时效性等质量特性进行校验；
- b) 相关部门应设立相应的数据检验管理岗位，承担在线数据校验管理职责，对校验结果进行统一汇总及处置；
- c) 有针对性地制定相应的校验方法，对不同数据库中的数据进行交叉校验；
- d) 应保持上述数据校验的证实性记录。

## 7.5 数据更正

数据质量管理者应明确数据更正流程的工序和内容，编制程序文件和相应的作业指导文件。在数据更正工作中满足以下要求：

- a) 数据质量管理部门和数据事权部门应建立数据更正的业务规则，并负责监督、指导各级部门对业务规则的执行，对拒不执行的部门，数据质量管理部门和数据事权部门有权终止违规者的更正操作；
- b) 在基础数据审核或校验过程中，若发现质量不合格的数据，应要求采集人员重新核对，经确认后，对原录入数据进行更正；
- c) 对外部受众投诉的数据质量问题，应启动更正程序，及时更正来自外部发现的质量不合格数据；
- d) 数据事权部门和数据质量管理部门应保持上述数据更正的证实性记录。

## 8 质量管理工具

### 8.1 数据质量管理规则

数据质量管理规则要求包括但不限于：

- a) 应支持根据数据质量特性自定义质量维度；
- b) 应支持关联数据标准规范，进行基础规则定义。包括但不限于空值校验、值域校验、格式校验、长度校验、精度校验、唯一性约束校验、主外键校验等；
- c) 应支持根据基础规则和平台应用场景自定义业务规则；
- d) 应通过自定义 SQL 语句等方式完成对数据质量的检核。

### 8.2 数据质量检测任务

数据质量检测任务要求包括但不限于：

- a) 应支持质量检测任务的创建；
- b) 应支持检测任务调度规则的设定；
- c) 应支持手动执行与自动执行检测任务，支持手动、定时、轮询、每周、自定义等多种调度规则；
- d) 应支持质量检测任务运维，查看任务执行日志、任务调度次数等；
- e) 应支持问题数据的查询与导出；
- f) 应支持未达标数据的查询与导出。

注1：问题数据是指平台设置数据库表中合格率范围之外的数据。

注2：未达标数据是未达到平台设置的数据质量特性合格率的表中所有数据。

### 8.3 数据质量监控

数据质量监控要求包括但不限于：

- a) 应支持对数据进行监控，明确数据的来源和去向；

b) 应支持对规则校验的结果进行监控和分析，校验结果应定位到原始数据项。

#### 8.4 数据质量报警

应支持问题数据的报警。

#### 8.5 数据质量检测报告

数据质量检测报告要求包括但不限于：

- a) 应支持对校验结果进行记录并生成检测报告，检测报告的内容包括但不限于：
  - 1) 历史检测记录的数据波动，以及本次质量检测执行时间、检测结果等基本信息；
  - 2) 对应数据的动态质量变化情况；
  - 3) 每项业务规则检测数据量、问题数据量、合格率等详细信息。
- b) 应支持在线查看每项业务规则问题数据的详情；
- c) 应支持打包下载对应的问题数据；
- d) 应支持数据质量检测报告的查看和下载。

#### 8.6 数据质量检测报告分析

数据质量检测报告分析要求包括但不限于：

- a) 应具备对检测报告进行综合性分析的功能，包括但不限于以下内容：
  - 1) 应支持通过多维度的分析报告来展示质量检测总体结果，包含检测数据来源、检测时间、综合得分、检测数据量、合格数据量、规则数目、检测资源数目等基本信息，和各个质量维度下每项的得分，以及此次检测数据的短板问题所属维度；
  - 2) 应支持展示每项业务规则的详细检测内容，如每项的检测数据总数、问题数据数目，合格率、问题数据分布情况等；
  - 3) 应支持通过评分报告详情展示每项规则对应的得分情况及对应的比分析权重，体现不同规则的重要程度，并对高权重质量规则进行着重分析处理；
  - 4) 应支持每个业务规则横向对比，提供宏观的检测报告分析，展示数据质量问题的分布。
- b) 应具备对检测报告进行统计分析的功能，包括但不限于以下内容：
  - 1) 应支持通过图表的形式展示数据质量概览信息；
  - 2) 应支持按照数据源目录层级，即分层分域原则，按照指定的业务层级如区域、部门、指定表等多种维度进行统计分析，并生成对应的质量趋势图。

#### 8.7 数据质量整改

数据质量整改要求包括但不限于：

- a) 应支持根据质量分析结果生成对问题数据的质量提升建议；
- b) 应支持数据标准化、数据清洗、数据转换和数据整合等功能，对不符合质量要求的数据进行处理，并纠正数据质量问题。

#### 8.8 数据质量跟踪

数据质量跟踪要求包括但不限于：

- a) 应支持数据质量跟踪，记录数据质量事件的评估、初步诊断和后续行动等信息，验证数据质量提升的有效性；
- b) 应支持对系统问题进行修正、对数据标准进行完善。

## 参 考 文 献

- [1] GA/T 1000-2011 公安信息化数据质量管理规范
  - [2] DB14/T 2526-2022 工业互联网综合平台数据质量管理要求
  - [3] DB3713/T 261-2022 公共数据 开放数据质量管理规范
-