

# 团体标准

T/ CSSC 001—2024

## 12345 政务服务便民热线 数据质量评价规范

Specification for 12345 Government Service Hotline Data Quality Evaluation

(征求意见稿)

2024-XX-XX 发布

2024-XX-XX 实施

中国商业统计学会 发布



### 版权保护文件

版权所有归属于该标准的发布机构。除非有其他规定，否则未经许可，此发行物及其章节不得以其他形式或任何手段进行复制、再版或使用，包括电子版，影印件，或发布在互联网及内部网络等。使用许可可于发布机构获取。

# 目 次

前 言 .....	IV
引 言 .....	V
1 范围.....	6
2 规范性引用文件.....	6
3 术语和定义.....	6
4 基本要求.....	7
5 评价对象.....	7
6 评价指标.....	9
7 评价方法.....	13
8 评价流程.....	13
附 录 A（规范性附录） 数据质量评价指标.....	15
参 考 文 献.....	21

## 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。

本文件由北京零点有数数据科技股份有限公司、国家信息中心提出。

本文件由中国商业统计学会归口。

本文件起草单位：北京零点有数数据科技股份有限公司、国家信息中心、XXX。

本文件主要起草人：XXX。

## 引 言

12345 政务服务便民热线是企业 and 群众反映问题建议、推动政务服务问题解决的重要渠道。高质量数据是热线正常运行的重要保障，对热线服务能力提升、分析决策深化、应用场景探索、数据价值挖掘与资产化等工作具有重要的意义。

本文件结合 12345 政务服务便民热线业务的数据特点，明确了热线数据质量评价的数据范围，并给出了可执行的评价指标、评价方法和评价流程，指导评估机构、热线单位开展数据质量评价，持续提升热线数据质量，并进一步实现热线数据的价值转化。

# 12345 政务服务便民热线数据质量评价规范

## 1 范围

本文件规定了12345政务服务便民热线（以下简称“热线”）数据质量评价的基本要求、评价对象、评价指标、评价方法与评价流程。

本文件适用于政府部门、热线单位、评估机构对12345等政务服务便民热线数据开展数据质量评价、数据质量管理等工作。

## 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 35295-2017 信息技术 大数据 术语

## 3 术语和定义

下列术语和定义适用于本文件。

### 3.1

**12345政务服务便民热线** 12345 government service hotline

由政府及其职能部门设立的，由电话12345和其他方式组成的，回应企业和群众等法人、自然人或其他组织诉求的非紧急公共服务平台。

[来源：GB/T 44189-2024, 3.1]

### 3.2

**12345政务服务便民热线数据** 12345 government service hotline data

依托各级政府设立的政府热线，通过电话12345、短信、信箱、网站及其他媒体等方式，为组织或个人参与社会治理、获取公共服务提供政策信息咨询、诉求受理与回访等公共服务而形成的数据。

[来源：GB/T 33358-2016, 3.2, 有修改]

### 3.3

**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。

[来源：GB/T 36344-2018, 2.3]

### 3.4

**热线数据质量评价** hotline data quality evaluation

对12345政务服务便民热线数据开展的的分析评价其数据质量水平的活动的集合。

### 3.5

**服务对象 service objects**

向12345政务服务便民热线咨询信息、反映诉求、提出意见建议的自然人、法人或其他组织。

[来源：GB/T 33358-2016, 3.4, 有修改]

3.6

**工单 workflow sheet**

记录服务对象基本信息、事项内容、办理情况、督办、回访结果等热线服务运行全过程的工作流转单。

[来源：GB/T 33358-2016, 3.6]

3.7

**知识库 knowledge base**

热线服务提供者对有关法律、法规、政策文件、部门职责、业务事项等政府政务公开与公共服务的信息,进行搜集、整理,形成可存储、可维护、可查询的信息集合。

[来源：GB/T 33358-2016, 3.5]

4 基本要求

4.1 安全合规

热线数据质量评价工作应满足数据安全合规要求，避免在评价过程中造成数据风险。

4.2 科学严谨

热线数据质量评价指标和方法应设置科学、逻辑严谨，真实全面地反映热线数据质量水平。

4.3 切实可行

数据质量评价方法应具有现实可操作性。

4.4 适配通用

数据质量评价方法、指标应适配热线业务，具有热线场景的通用性。

5 评价对象

5.1 话务类数据

5.1.1 来电数据

来电数据是指在热线来电服务过程中产生的数据，包括：

- a) 来电相关：来电时间、等待时间、是否接通、接通时间、结束时间、是否有效来电等；
- b) 来电内容：来电人的咨询、诉求、意见建议等内容等；
- c) 来电人：来电号码、来电人信息、是否匿名等；
- d) 接待人：坐席号码、坐席人员信息等。

### 5.1.2 外呼数据

外呼数据是指在热线外呼过程中产生的数据，包括：

- a) 外呼相关：外呼发起时间、接通时间、结束时间、是否接通、是否有效外呼等；
- b) 外呼内容：外呼类型、外呼内容、关联工单等；
- c) 外呼对象：呼叫号码、外呼对象信息等；
- d) 外呼人：坐席号码、坐席人员信息等。

### 5.1.3 录音数据

录音数据是在热线来电、外呼过程中系统记录的录音数据，包括：

- a) 录音相关：录音开始时间、录音结束时间、录音时长、录音格式等；
- b) 关联信息：外呼任务、关联工单等。

### 5.1.4 挂机评价数据

服务对象在服务完成后挂机前对坐席人员进行的评价数据，包括：

- a) 评价相关：评价时间、评价问题、评价结果等；
- b) 评价人相关：评价人信息等；
- c) 评价对象相关：坐席号码、坐席人员信息等。

### 5.1.5 其他受理数据

通过短信、信箱、网站、移动端、小程序、公众号等国家和地方诉求受理渠道获取的诉求数据，包括：

- a) 受理信息：来源渠道、诉求发起时间、受理时间等；
- b) 服务对象信息：诉求人、诉求内容、诉求类型、是否有效诉求等；
- c) 接待人：坐席号码、坐席人员信息等；
- d) 关联文件信息：图片、视频、录音、文档等相关文件信息。

## 5.2 工单类数据

### 5.2.1 工单基础数据

工单数据是指在热线服务产生的，需要相关机构、部门、人员处置、反馈的任务数据，包括：

- a) 时间相关：创建时间、响应时间、完结期限、完结时间等；
- b) 类型内容：工单编号、工单分类、工单标题、工单内容等；
- c) 工单任务：工单来源、工单状态、是否重复工单等；
- d) 关联信息：问题点位、责任单位、工单评价等。

### 5.2.2 工单流转数据

工单流转数据是指工单转派、接单、处理、退回、督办等流转数据，包括：

- a) 流转环节：流转环节类型、流转时间、责任人等；
- b) 关联信息：流转进度信息、关联文件等。

### 5.2.3 工单反馈数据

工单反馈数据是指相关部门对工单诉求、问题做出的反馈，包括反馈部门、反馈时间、反馈内容等。

#### 5.2.4 工单评价数据

工单评价数据是指对企业、群众对工单事项是否解决、是否满意的评价数据，包括：

- a) 评价相关：评价时间、评价点、评价结果、是否解决、是否满意等；
- b) 评价人：评价人信息等。

#### 5.3 知识库数据

热线在服务过程中使用的知识信息，如政策、活动、联系方式等相关知识，包括：

- a) 知识相关：知识类型、知识内容、关键词等；
- b) 维护相关：创建时间、创建人、维护时间、维护人、当前版本等。

#### 5.4 统计数据

热线管理部门对热线数据进行建模、分析得到的各种统计数据，包括：

- a) 指标定义：指标名称、指标含义、计算公式、关联维度等；
- b) 指标统计：统计周期、统计范围、统计结果等；
- c) 关联信息：数据归口部门等。

#### 5.5 智能模型数据

热线智能导航、智能问答、智能外呼、坐席辅助、智能质检、智能知识库等人工智能模型应用相关的数据，包括：

- a) 多模态数据：语音数据、图像数据、文本数据、地理位置数据等智能模型学习和处理的多模态数据；
- b) 标注数据：用于智能模型训练的人工或半人工标注数据集；
- c) 合成数据：因模型训练、隐私保护等目的，通过模拟或生成技术生成的数据；
- d) 训练数据：用于人工智能模型训练的数据集；
- e) 验证和测试数据：用于热线智能模型验证和测试使用的数据集；
- f) 推理数据：智能模型对输入数据进行处理分析后推理生成的数据集。

### 6 评价指标

#### 6.1 数据质量评价维度

数据质量评价维度包括规范性、完整性、一致性、准确性、时效性、可访问性共6个方面：

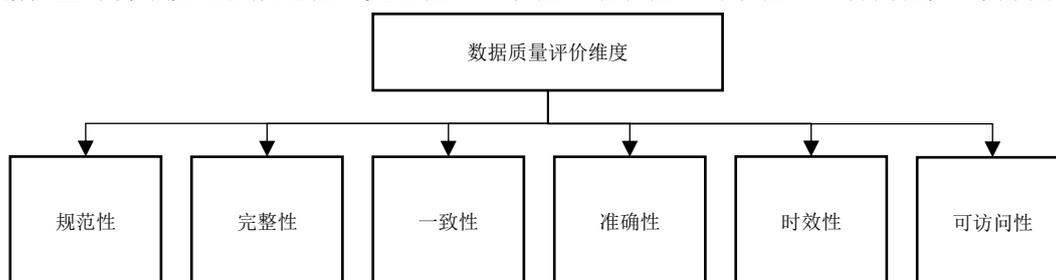


图1 数据质量评价维度

## 6.2 数据质量评价指标

### 6.2.1 总体要求

热线数据质量评价指标通过定量、定性的方式对热线数据满足规范性、完整性、一致性、准确性、时效性、可访问性的情况进行评价，既包含结构化数据的质量评价，也包含对热线文本、录音等非结构化数据的质量评价。数据质量评价指标的设置既评价数据字段自身的质量情况，也评价数据在跨行、跨列、跨表情况下的一致性、规范性等情况。数据质量评价指标见附录A。

### 6.2.2 数据规范性

热线数据规范性衡量热线数据与所定义的值域、语法、格式等规范的符合程度。

表1 数据规范性评价指标表

序号	指标名称	指标描述	计算方式
1	数值值域规范性	热线数据中数值型字段的值域应满足数据定义的要求。	全部检查数值型数据条数 A，全部满足规范性数据条数 B，数值值域规范性得分=B/A*100。
2	枚举值域规范性	热线数据中枚举型字段的值应在预定义的枚举值范围内。	全部检查枚举型数据条数 A，枚举全部满足规范性数据条数 B，枚举值域规范性得分=B/A*100。
3	数据语法规范性	应按照系统字段定义的语法格式记录数据。	全部检查数据条数 A，全部满足语法规范性数据条数 B，数据语法规范性得分=B/A*100。
4	数据长度规范性	数据字段的长度应满足数据定义要求。	全部检查数据条数 A，全部满足长度规范性数据条数 B，数据长度规范性得分=B/A*100。
5	数据格式规范性	应按照系统定义的格式存储和记录数据。	全部检查数据条数 A，全部满足格式规范性数据条数 B，数据格式规范性得分=B/A*100。

### 6.2.3 数据完整性

热线数据完整性衡量热线数据按规则要求被完整存储、赋值、描述的程度。

表2 数据完整性评价指标表

序号	指标名称	指标描述	计算方式
1	结构化数据字段完整性	结构化数据的字段设置应满足热线业务运行和数据分析的要求。	应涵盖字段数量 A，检查数据覆盖的数据字段数量 B，结构化数据字段完整性得分=B/A*100。
2	结构化数据字段赋值完整性	满足非空条件的数据字段应被完整赋值。	全部检查数据条数 A，字段全部完整赋值的数据条数 B，结构化数据字段赋值完整性得分=B/A*100。
3	非结构化文件完整性	应完整保存的热线业务相关的录音、视频、图片、文档等文件信息完整。	全部检查文件数据条数 A，文件数据全部完整的数据条数 B，非结构化文件完整性得分=B/A*100。
4	非结构化文件元数据完整性	热线业务运行过程中形成的图片、视频、音频、文档等文件的名称、描述、关联受理记录等元数据应记录完整。	全部检查文件条数 A，非结构化文件元数据全部完整的数据条数 B，完整性得分=B/A*100。

序号	指标名称	指标描述	计算方式
5	内容要素完整性	热线在业务运行、知识库维护、模型训练等场景中形成的内容要素应完整。	全部抽查数据条数 A，内容要素全部完整的数据条数 B，内容要素完整性得分=B/A*100。
6	代码含义完整性	热线数据中的代码字段应可通过数据关联可追溯到其数据描述。	全部检查代码数据条数 A，代码数据描述可追溯的数据条数 B，代码含义完整性得分=B/A*100。

#### 6.2.4 数据一致性

热线数据一致性衡量热线数据与上下文中使用的数据和数据定义无矛盾的程度。

表3 数据一致性评价指标表

序号	指标名称	指标描述	计算方式
1	数据标识一致性	热线数据应使用统一编码、逻辑一致的标识编码。	全部检查数据条数 A，全部满足一致性数据条数 B，数据标识一致性得分=B/A*100。
2	数据格式一致性	系统记录的数据格式应与数据定义的格式一致。	全部检查数据条数 A，全部满足一致性数据条数 B，数据格式一致性得分=B/A*100。
3	数据等值一致性	数据值应与实际数据条件下的相关数据计算的值一致。	全部检查数据条数 A，全部满足一致性数据条数 B，数据等值一致性得分=B/A*100。
4	字段定义一致性	话务、工单、统计数据同一字段定义、字段口径应一致。	全部检查相同字段数据条数 A，全部满足一致性数据条数 B，字段定义一致性得分=B/A*100。
5	字段逻辑一致性	字段之间的关系应与其业务逻辑保持一致。	全部检查数据条数 A，全部满足一致性数据条数 B，字段逻辑一致性得分=B/A*100。

#### 6.2.5 数据准确性

热线数据准确性衡量热线数据表示其所描述的事实相符合的程度。

表4 数据准确性评价指标表

序号	指标名称	指标描述	计算方式
1	事实符合性	话务、工单、知识库、统计数据应与其所代表的事实相匹配。	全部抽查数据条数 A，全部满足真实性数据条数 B，数据准确性得分=B/A*100。

#### 6.2.6 数据时效性

热线数据时效性衡量热线数据满足及时性要求的程度，包括数据入库及时性、数据计算及时性。

表5 数据时效性评价指标表

序号	指标名称	指标描述	计算方式
----	------	------	------

1	入库及时性	热线数据从发生到数据入库时间应满足业务场景需求。	全部数据条数 A，全部满足时效性数据条数 B，入库及时性得分=B/A*100。
2	计算及时性	热线明细数据产生到统计数据计算完成的时间应满足业务场景需求	全部统计数据条数 A，全部满足时效性数据条数 B，计算及时性得分=B/A*100。

### 6.2.7 数据可访问性

热线数据可访问性衡量热线数据能被访问和使用的程度。

表6 数据可访问性评价指标表

序号	指标名称	指标描述	计算方式
1	数据可读性	热线数据应具有完备的详实的数据说明文档，并及时更新。	具备完善数据说明文档，数据实际情况与数据文档一致，100分；具有说明文档，但未随数据变化及时更新导致，50分，无相关说明文档，0分。
2	访问技术条件	热线数据应具备标准的 API 接口、消息队列、数据文件等技术访问条件。	已开发接口、消息队列等访问方式，并配备说明信息，满足多种数据访问需求，100分；开发有数据接口、消息队列，配备说明信息，但不能满足全部数据需求场景，50分；无现成访问方式的，0分。
3	访问合规条件	热线数据应具备访问数据的合规条件，包括数据权属明确，无隐私数据、涉密数据等。	数据满足全部合规条件，100分；存在不满足合规条件的数据，0分。

## 6.3 数据质量评价指标计算

### 6.3.1 计算方式

热线数据质量评价指标根据评价对象类型，分别对话务数据、工单数据、知识库数据、统计数据 and 智能模型数据进行数据质量评分，再单个类别数据质量评分计算得出热线数据质量总体评分。

### 6.3.2 单个类别数据质量评分

单个类别的数据质量评分根据其数据评价指标的得分加权计算得出，计算方式如式（1）：

$$Q = (\sum q \times w / \sum w) \times 100 \dots \dots \dots (1)$$

式中：

Q——数据质量总体评分；

q——单个数据质量评价指标得分；

w——数据质量评价指标权重。

### 6.3.3 热线数据质量总体评分

热线数据质量总体评价得分根据各个类别数据的数据质量评价得分求和得出，计算方式如式（2）：

$$SQ = \sum Q / N \dots \dots \dots (2)$$

式中：

SQ——热线数据质量评分；

Q——单个类别的数据质量得分。

N——数据类别数量。

## 7 评价方法

### 7.1 分层抽样法

- a) 分层抽样法通过对热线数据进行分层并抽取样本进行数据质量评价，实现样本的均匀覆盖，数据可基于时间、一级分类进行分层；
- b) 分层抽样法各层抽取的样本量应相对均匀，数据规模具有统计学意义，各层抽样比例可遵循 GB/T 10111-2008 的规定；
- c) 分层抽样法适用于工单内容、反馈内容、知识内容、录音等非结构化数据的质量评价场景。

### 7.2 全量统计法

- a) 全量统计法通过对全量热线数据进行规则匹配和数据质量分析，实现数据质量评价；
- b) 全量统计法适用于话务数据、工单数据、知识库数据、统计数据中结构化字段的质量评价。

### 7.3 调查分析法

- a) 调查分析法通过调研系统、文献资料、访谈等方式，分析、评价数据质量；
- b) 调查分析法可支持对数据质量评价指标进行评价和修正；
- c) 调查分析法适用于热线时效性、可访问性等数据质量评价场景。

## 8 评价流程

数据质量评价流程包括导入数据质量评价环境，分析数据并确定评价指标，确定权重和建立评价模型，提取规则库并进行检测，核验检查结果得到指标评分，形成数据质量评价结论。评价流程见图2。

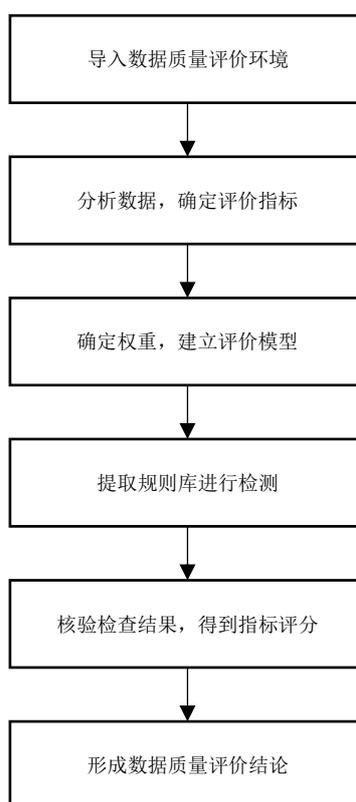


图 2 数据质量评价流程

流程环节如下：

- a) 导入数据质量评价环境，确定评价对象和评价范围。开展数据质量评价准备工作，建立数据质量评价团队，梳理待评价单位的业务现状、系统建设现状、数据现状，整理热线业务要求、数据设计文档及适用的标准规范文档，明确待评价数据范围；
- b) 分析数据特征并确定评价指标。对热线数据进行盘点，识别数据特征，并确定适用的评价指标；
- c) 建立评价模型，明确数据质量评价指标权重。根据热线数据特征，建立评价模型，并对数据质量评价指标权重进行适当调整；
- d) 根据数据质量评价指标要求，进行数据质量检查，形成检查评分结果；
- e) 核验检查结果，计算评价指标得分和总体数据质量评分；
- f) 形成数据质量评价结论，出具质量评价报告。

**附录 A**  
**(规范性)**  
**数据质量评价指标**

话务数据、工单数据、知识库数据、统计数据、智能模型数据等类别数据质量评价指标分别见表A. 1、表A. 2、表A. 3、表A. 4、表A. 5，时效性评价指标和可访问性评价指标是各类别数据质量评价的通用评价指标。

**表 A. 1 话务数据质量评价指标表**

序号	一级指标	二级指标	三级指标	指标描述
1	完整性	结构化数据字段完整性	话务字段数据完整性	来电时间、接通时间、是否接通、挂机时间、来电人信息、坐席人员信息等数据字段完整。
2		结构化数据字段完整性	录音数据完整性	录音开始时间、录音结束时间、录音时长、录音关联信息等数据完整。
3		非结构化文件完整性	关联文件完整性	受理过程中形成的图片、视频、音频、文档等文件数据完整。
4		非结构化文件元数据完整性	文件元数据完整性	受理过程中形成的图片、视频、音频、文档等文件的名称、说明、关联受理记录等信息完整。
5		代码含义完整性	话务数据代码含义完整性	话务类型代码、部门代码、地理位置代码及其他代码字段，通过数据关联可追溯到其数据描述。
6	规范性	数值值域规范性	录音数据值域规范性	录音时长等数值型字段的值域满足数据定义的要求，如录音时长的数值 > 0。
7		数值值域规范性	评价数据值域规范性	评价结果等数值型字段的值域满足数据定义的要求，如五分制评价的得分应[1, 5]。
8		枚举值域规范性	话务数据枚举值域规范性	话务数据中枚举类型的数据满足枚举值域的定义，如外呼类型、是否匿名等字段的值应在枚举值范围内。
9		数据语法规范性	话务数据字段语法规范性	按照系统字段定义的语法格式记录数据，如手机号码格式满足 GB/T 31186. 5-2014 的要求等。
10		数据格式规范性	话务数据字段格式规范性	按照系统定义的格式存储和记录数据，如使用时间格式记录来电时间、呼出时间等。
11	准确性	事实符合性	话务数据真实性	呼入、外呼记录的话务信息是企业、群众的真实行为，无弄虚作假，话务记录信息与实际情况一致。
12	一致性	数据标识一致性	话务数据标识一致性	热线数据中使用统一编码、逻辑一致的通话编码等标识。
13		数据格式一致性	话务数据格式一致性	系统记录的数据格式与数据定义的格式一致。
14		数据等值一致性	话务数据等值一致性	话务数据相关字段的值与其相关字段计算后的值一致，如呼叫时长=呼叫结束时间-呼叫接通时间。
15		字段定义一致性	话务字段定义一致性	代表相同含义的话务字段数据定义上应保持一致。
16		字段逻辑一致性	话务字段逻辑一致性	呼出时间、呼入时间应早于呼出结束时间、呼入结束时间。

表 A.2 工单数据质量评价指标表

序号	一级指标	二级指标	二级指标	指标描述
1	完整性	结构化数据字段完整性	工单数据字段完整性	工单编码、工单类型、工单创建时间、工单标题、工单内容、工单状态、工单完结时间、工单来源、诉求人信息等数据字段完整。
2		结构化数据字段赋值完整性	工单数据字段赋值完整性	工单编码、工单类型、工单创建时间、工单标题、工单内容、工单状态、工单完结时间、工单来源、诉求人信息等数据字段在数据定义条件下赋值完整。
3		内容要素完整性	工单数据内容要素完整性	不同类型的工单数据的内容要素的完整性。其中： a) 咨询类：诉求人咨询内容完整； b) 建议类：诉求人建议内容及理由完整； c) 求助、投诉、举报类：诉求人姓名、时间、地点、事件、涉及对象、前期处理情况、核心诉求完整； d) 不稳定因素类：诉求人反映事件的时间、地点、实施行为要素完整； e) 突发事件类：事件发生的时间、地点、状态、影响范围、危害程度、目前状态内容要素完整。
4		结构化数据字段完整性	办理数据字段完整性	派单、承办、反馈、督办等环节的办理时间、办理部门、办理人员、文本内容字段完整。
5		结构化数据字段赋值完整性	办理数据字段赋值完整性	派单、承办、反馈、督办等环节的办理时间、办理部门、办理人员、文本内容字段在数据定义条件下赋值完整。
6		内容要素完整性	反馈内容完整性	反馈内容按照工单类型包含完整的内容要素。其中： a) 咨询类：承办单位、答复内容、答复时间； b) 建议类：移交转办信息； c) 已解决的求助、投诉、举报及其他类型：主责单位、办理时间、主要措施； d) 未解决的求助、投诉、举报及其他类型：主责单位、未解决原因、下一步工作措施、落实时间、反馈情况； e) 不能解决的求助、投诉、举报及其他类型：主责单位、不能解决原因或困难、有无研究工作措施、反馈情况。
7		内容要素完整性	评价内容完整性	工单评价内容应覆盖服务态度、工作效率、办理过程、办理结果的满意度。
8		代码含义完整性	工单数据代码含义完整性	工单类型代码、部门代码、地理位置代码及其他代码字段，通过数据关联可追溯到其数据描述。
9		数值值域规范性	工单数据值域规范性	办结期限等数值型字段的值域满足数据定义的要求，如办结期限的数值 > 0。
10	数值值域规范性	评价数据值域规范性	评价结果等数值型字段的值域满足数据定义的要求，如五分制评价的得分应[1, 5]。	
11	枚举值域规范性	工单枚举值域规范性	工单类型、工单状态、工单来源、责任单位等字段的值应在枚举值范围内。	
12	数据语法规范性	工单字段语法规范性	按照系统字段定义的语法规则记录数据，如手机号码格式满足 GB/T 31186.5-2014 的要求等。	
11	数据长度规范性	工单字段长度规范性	工单内容、反馈内容的长度满足规范要求。	
13	数据格式规范性	工单字段格式规范性	按照系统定义的格式存储和记录数据，如使用时间格式记录工单创建时间等。	
14	数据格式规范性	工单文本格式规范性	工单内容文本格式满足规范性要求，使用书面语完整记录诉求人的诉求信息、时间、地点、事件、现状、诉求等信息。	

序号	一级指标	二级指标	二级指标	指标描述
15		数据格式规范性	反馈文本格式规范性	反馈文本格式满足规范性要求，使用书面语反馈诉求人信息的处理结果、解决方案、相关依据等信息。
16	准确性	事实符合性	工单数据真实性	工单记录数据、工单流转数据、工单反馈数据与实际情况符合。
17	一致性	数据标识一致性	工单数据标识一致性	热线数据中使用统一编码、逻辑一致的通话编码、问卷编码、工单编码等标识。
18		数据格式一致性	工单数据格式一致性	工单数据记录的数据格式与数据定义的格式一致。
19		数据等值一致性	工单数据等值一致性	工单数值字段的值与其依赖的字段的价值计算结果一致，如工单处理时长=工单完结时间-工单创建时间。
20		字段定义一致性	工单字段定义一致性	工单类型、工单来源、工单状态等字段的定义在工单数据中保持一致。
21		字段逻辑一致性	工单字段逻辑一致性	工单字段之间的逻辑关系满足数据定义的逻辑，如工单完结时间晚于工单创建时间、工单状态未完结，完结时间为空等。
22		字段逻辑一致性	评价字段逻辑一致性	工单评价数据的评价结果（满意、不满意）与其评价得分的映射规则一致。

表 A.3 知识库数据质量评价指标表

序号	一级指标	二级指标	二级指标	指标描述
1	完整性	内容要素完整性	知识内容完整性	知识库知识涵盖覆盖其应包含的所有内容要素。
2		结构化数据字段完整性	知识库数据字段完整性	知识名称、知识版本、知识维护人、更新日期、是否有效等元数据信息完整。
3	规范性	数值值域规范性	知识库数据值域规范性	知识库数据的数值型字段的值域满足数据定义的值域规范，如浏览量、搜索量等字段应为大于 0 的整数等。
4		枚举值域规范性	知识库数据枚举值域规范性	知识类型、知识区域等枚举字段的值应在预定义的枚举值范围内。
5		数据语法规范性	知识库数据字段语法规范性	按照系统字段定义的语法规则记录数据，如手机号码格式满足 GB/T 31186.5-2014 的要求等。
6		数据长度规范性	知识库数据字段长度规范性	知识标题、知识内容等字段满足预定义的数据格式要求。
7		数据格式规范性	知识库数据字段格式规范性	知识库字段按照系统定义的格式存储和记录数据，如使用时间格式记录工单创建时间等。
8		数据格式规范性	知识库内容文本格式规范性	知识库内容的文本格式满足规范性要求，使用书面语完整记录知识的名称、时间、地点、部门、事件、现状等信息。
9	准确性	事实符合性	知识库数据真实性	知识库记录的知识信息与实际参照信息相符合。
10	一致性	数据标识一致性	知识库数据标识一致性	知识库数据中同一字段应使用统一格式、逻辑一致的标识编码。
11		数据格式一致性	知识库数据格式一致性	知识库数据中同一字段的数据格式一致。
12		字段定义一致性	知识库数据字段定义一致性	知识库中知识类型、关联部门等字段的定义一致。
13		字段逻辑一致性	知识库数据字段逻辑一致性	知识库中数据字段之间的逻辑关系一致，如只是创建时间不应晚于维护时间等。

表 A.4 统计数据质量评价指标表

序号	一级指标	二级指标	二级指标	指标描述
1	完整性	内容要素完整性	指标定义完整性	指标的名称、含义、计算公式、更新周期、数据范围、维度等定义数据完整。
2		内容要素完整性	指标结果维度完整性	指标计算结果覆盖其定义和应用所需的所有维度。
3		代码含义完整性	统计数据代码含义完整性	指标统计结果中的所有代码信息通过数据关联可追溯到其数据描述。
4	规范性	数值值域规范性	统计数据值域规范性	统计数据中数值型字段的值域满足数据定义的要求，如接通率的值域范围为[0, 1]。
5		数据格式规范性	统计字段格式规范性	按照系统定义的格式存储和记录数据，如使用时间格式记录工单创建时间等。
7	准确性	事实符合性	统计数据真实性	指标的业务含义和计算逻辑与其所代表的现实业务逻辑一致。
7	一致性	数据标识一致性	统计数据标识一致性	统计数据使用与原始数据使用统一一致的标识编码。
8		数据格式一致性	统计数据格式一致性	统计数据的格式与指标定义的格式一致。
9		数据等值一致性	统计字段等值一致性	接通率、满意度等指标根据计算公式和计算条件，与实际数据条件下的明细数据的值保持一致。
10		字段定义一致性	统计字段定义一致性	统计字段的统计口径一致，如总体满意度、区域满意度、月满意度等指标的满意度计算使用的数据和计算方法保持一致。
11		字段逻辑一致性	统计字段逻辑一致性	统计字段的计算结果与明细数据按照计算公式得到的结果一致；分维度计算的数据与全部数据的计算结果一致。

表 A.5 智能模型数据评价指标表

序号	一级指标	二级指标	二级指标	指标描述
1	完整性	内容要素完整性	模型数据覆盖完整性	智能模型的特征/标签、训练数据、验证测试数据、推理数据等数据覆盖所有应用场景的需求。
2		代码含义完整性	代码含义完整性	智能模型的所有代码字段，通过数据关联可追溯到其数据描述。
3	规范性	数值值域规范性	模型数据数值值域规范性	标注数据、合成数据、推理数据中数值类型的数据满足值域定义的要求。
4		枚举值域规范性	模型数据枚举值域规范性	标注数据、合成数据、推理数据中数值类型的数据满足枚举值域定义的要求。
5		数据语法规范性	模型数据数据语法规范性	标注数据、合成数据、推理数据中数值类型的数据满足语法定义的规范。
6		数据长度规范性	模型数据数据长度规范性	标注数据、合成数据、推理数据中数值类型的数据满足长度定义的规范。
7		数据格式规范性	模型数据数据格式规范性	标注数据、合成数据、推理数据中数值类型的数据满足数据格式定义的规范。
8		数据格式规范性	合规规范性	推理过程具备有效的敏感词等合规检测。
9	准确性	事实符合性	训练数据准确性	训练数据与实际数据情况分类的定义保持一致。
10	一致性	数据标识一致性	模型数据数据标识一致性	标注数据、合成数据、验证测试数据、推理数据的使用统一的数据标识。
11		数据格式一致性	模型数据数据格式一致性	标注数据、合成数据、验证测试数据、推理数据的数据格式与数据定义一致。
12		数据等值一致性	模型数据数据等值一致性	标注数据、合成数据、验证测试数据、推理数据的数据值与相关值计算的结果一致。
13		字段定义一致性	模型数据字段定义一致性	标注数据、合成数据、验证测试数据、推理数据中同一数据的定义口径一致。
14		字段逻辑一致性	模型数据字段逻辑一致性	标注数据、合成数据、验证测试数据、推理数据的数据字段之间逻辑关系保持一致。

## 参 考 文 献

- [1] 国务院办公厅关于进一步优化地方政务服务便民热线的指导意见(国办发(2020) 53 号)
  - [2] 国务院关于进一步优化政务服务提升行政效能推动“高效办成一件事”的指导意见(国发[2024]3 号)
  - [3] GB/T 10111-2008 随机数的产生及其在产品质量抽样检验中的应用程序
  - [4] GB/T 36344-2018 信息技术 数据质量评价指标
  - [5] GB/T 43709-2024 资产管理信息化数据质量管理要求
  - [6] GB/T 44189-2024 政务服务便民热线运行指南
  - [7] GB/T 44190-2024 政务服务便民热线集成规范
  - [8] GB/T 44191-2024 政务服务便民热线知识库建设指南
  - [9] GB/T 44192-2024 政务服务便民热线数据应用指南
  - [10] DB11/T 2021-2022 12345 市民服务热线服务与管理规范
  - [11] DB21/T 3709-2023 12345 政务服务便民热线管理与服务规范
  - [12] T/CBA 221-2024 银行业数据资产估值指南
  - [13] 《面向人工智能的数据治理实践指南（1.0）》
  - [14] ISO/TS 8000-8:2015 Data quality Part 8:Information and data quality:Concepts and measuring
  - [15] ISO/TS 8000-81:2021 Data quality Part 81:Data quality assessment: Profiling
  - [16] ISO/IEC 5259-3 Artificial intelligence - Dataquality for analytics and machinelearning(ML) Part 3:Data quality managementrequirements and guidelines
  - [17] ISO/IEC 5259-4 Artificial intelligence - Dataquality for analytics and machinelearning(ML) Part 4:Data quality process framework
-