



中华人民共和国国家标准

GB/T 42381.8—2023/ISO 8000-8:2015

数据质量 第8部分： 信息和数据质量：概念和测量

Data quality—Part 8: Information and data quality: Concepts and measuring

(ISO 8000-8:2015, IDT)

2023-03-17 发布

2023-10-01 实施

国家市场监督管理总局 发布
国家标准化管理委员会

目 次

前言 I

引言 II

1 范围 1

2 规范性引用文件 1

3 术语和定义 1

4 基础概念 2

5 信息和数据质量 2

 5.1 语法质量 2

 5.2 语义质量 3

 5.3 语用质量 4

6 一致性 4

附录 A（规范性） 文件标识 5

附录 B（资料性） 语法质量规则 6

附录 C（资料性） 语用质量 7

附录 D（资料性） 用于测量信息和数据质量的活动模型 8

参考文献 12

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。

本文件是 GB/T 42381《数据质量》的第 8 部分。GB/T 42381《数据质量》已经发布了以下部分：

- 第 8 部分：信息和数据质量：概念和测量；
- 第 61 部分：数据质量管理：过程参考模型。

本文件等同采用 ISO 8000-8:2015《数据质量 第 8 部分：信息数据质量：概念和测量》。

本文件做了下列最小限度的编辑性改动：

- 在第 4 章中增加了对附录 A 的规范性引用。

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。

本文件由中国机械工业联合会提出。

本文件由全国自动化系统与集成标准化技术委员会(SAC/TC 159)归口。

本文件起草单位：中国标准化研究院、十堰市工业产品质量检验检测所、南京市特种设备安全监督检验研究院、安徽华普生产力促进中心有限公司、福建中信网安信息科技有限公司、福建亿榕信息技术有限公司、中国科学院计算机网络信息中心、中国人民大学、中国科学院计算技术研究所、广东华昌集团有限公司、浙江大学、北京金谷远见科技有限公司、深圳鹏锐信息技术股份有限公司。

本文件主要起草人：王志强、高峻峰、冯月贵、陶影海、金华松、庄莉、朱艳华、安小米、洪学海、郭加林、王会方、顾复、张伟群、杨青海、臧传庐、王彦军、洪岩、岳高峰、刘守华、徐凯程、尹书蕊。

引言

创建、收集、存储、维护、传输、处理、呈现信息和数据以便及时、经济地支持业务过程,既需要了解决定信息和数据质量的相关特征,也需要测量、管理、报告信息和数据质量。

GB/T 42381 定义了决定信息和数据质量的相关特征,并提供了管理、测量、提高信息和数据质量的方法。这些给定方法可用于评估信息和数据的质量。面向当前业务案例的期望和需求,对给定方法进行剪裁也很重要。

GB/T 42381 包括适用于所有数据类型的部分,以及适用于特定数据类型的部分。GB/T 42381 可独立使用,也可与质量管理体系结合使用。

GB/T 42381《数据质量》由以下系列组成。

- 第 1 到 99 部分:数据质量综述。该系列给出了数据质量系列标准的结构、数据质量标准涉及的基本概念、数据质量管理、数据质量评估等综合内容,ISO 8000-1 给出了数据质量部分的综述。
- 第 100 到 199 部分:主数据质量。主数据描述了个人、组织、地点、物品、服务、过程、规则 and 标准。该系列标准描述定义主数据质量的特性。这些特性包括:语义、语法编码、需求的一致性、数据源、准确性、完整性和数据治理。该系列规定了一些主数据信息,这些信息应在总体上确保信息发送方和接收方数据通信的可靠性。ISO 8000-100 给出了主数据质量部分的综述。
- 第 200 到 299 部分:事务数据质量。事务数据描述了时间事件,包括个人、组织、地点、物品、服务、过程、规则 and 标准。该系列标准描述定义事务数据质量的特性。这些特性包括:语义、语法编码、需求的一致性、数据源头、准确性、完整性和数据治理。该系列规定了一些业务事务数据信息,这些信息应在总体上确保信息发送方和接收方数据通信的可靠性。ISO 8000-200 给出了事务数据质量部分的综述。
- 第 300 到 399 部分:产品数据质量:产品数据质量是产品数据正确性和适用性的度量,产品数据可保证数据能及时地提供给需要这些数据的用户,产品数据是产品从概念到制造需要的数据。因此,产品不仅仅包括计算机辅助设计(CAD)数据,还包括计算机辅助制造(CAM)数据、计算数据工程(CAE)数据、产品数据管理(PDM)数据以及其他类型的数据。该系列的主要意图是提高协调产品开发的效率,减少数据接收端的重复工作。ISO 8000-300 给出了产品数据质量部分的综述。

GB/T 42381《数据质量》拟由以下部分构成:

- 第 1 部分:综述;
- 第 2 部分:术语;
- 第 8 部分:信息和数据质量:概念和测量;
- 第 51 部分:数据治理:数据交换的策略声明;
- 第 61 部分:数据质量管理:过程参考模型;
- 第 62 部分:数据质量管理:组织过程成熟度评估:过程评估相关标准的应用;
- 第 63 部分:数据质量管理:过程测量;
- 第 65 部分:数据质量管理:过程测量量表;
- 第 66 部分:数据质量管理:生产运作管理数据处理的评估指标;
- 第 81 部分:数据质量评估:数据收集;
- 第 82 部分:数据质量评估:创建数据规则;

- 第 100 部分:主数据:特征数据交换:概述;
- 第 110 部分:主数据:特征数据交换:语法、语义编码和对数据规范的符合性;
- 第 115 部分:主数据:质量标识符交换:语法、语义和解析要求;
- 第 116 部分:主数据:质量标识符交换:ISO 8000-115 授权实体标识符的应用;
- 第 120 部分:主数据:特征数据交换:溯源性;
- 第 130 部分:主数据:特征数据交换:准确性;
- 第 140 部分:主数据:特征数据交换:完整性;
- 第 150 部分:主数据:质量管理框架;
- 第 210 部分:传感器数据:数据质量特性;
- 第 311 部分:形状产品数据质量(PDQ-S)应用指南。

本文件是 GB/T 42381《数据质量》的第 8 部分,可单独使用,也可与其他部分结合使用。

本文件供信息或数据质量利益相关方使用,重点在于关注一个或多个信息系统,包括组织间和组织内部的视图,以及贯穿于所有生命周期阶段的信息系统。当评估信息和数据的质量是否达标时,有必要为每个要测量的对象建立与业务相关的阈值,本文件没有设置这些阈值。明确测量值所采用的尺度也非常重要,本文件没有定义测量信息和数据质量的尺度,但要求对这些尺度做出说明。当传达信息和数据质量量化的结果时,接收者能够理解结果的置信度是十分有用的,尤其重要的是,要知道是否有任何规则未被应用,或者是否有任何信息或数据未经检查。

数据质量 第8部分： 信息和数据质量：概念和测量

1 范围

本文件描述了信息和数据质量的基础概念，以及这些概念应用于质量管理过程和质量管理体系的方式。

本文件还规定了在质量管理过程和质量管理体系中进行信息和数据质量测量的先决条件。

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。

ISO 8000-2 数据质量 第2部分：术语(Data quality—Part 2: Vocabulary)

3 术语和定义

ISO 8000-2 界定的以及下列术语和定义适用于本文件。

3.1

数据 data

信息的可解释形式化表示，以适用于通信、解释或处理。

[来源：ISO/IEC 2382:2015，2.12，有修改—删除注释。]

3.2

实体 entity

关注域内的具体或抽象的事物。

[来源：GB/T 16642—2008，3.29，有修改—删除的“任何”一词。]

3.3

信息 information

关于客体（如事实、事件、事物、过程或思想，包括概念）的知识，在一定的场合中具有特定的意义。

[来源：ISO/IEC 2382:2015，2121271，有修改—删除应用领域和注释。]

3.4

元数据 metadata

定义和描述其他数据的数据。

[来源：GB/T 18391.1—2009，3.41]

3.5

要求 requirement

明示的，通常隐含的或必须履行的需求或期望。

[来源：GB/T 19000—2016，3.6.4，有修改—删除注释。]