

ICS 35.240.01
L 70



中华人民共和国国家标准

GB/T 36344—2018

信息技术 数据质量评价指标

Information technology—Evaluation indicators for data quality

2018-06-07 发布

2019-01-01 实施

国家市场监督管理总局
中国国家标准化管理委员会 发布

目 次

前言	I
1 范围	1
2 术语和定义	1
3 指标框架	2
4 概述	2
5 指标说明	2
5.1 评价表中表头信息说明	2
5.2 规范性	3
5.3 完整性	4
5.4 准确性	4
5.5 一致性	4
5.6 时效性	5
5.7 可访问性	5
附录 A (资料性附录) 数据质量评价过程	6
参考文献	7

前 言

本标准按照 GB/T 1.1—2009 给出的规则起草。

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这些专利的责任。

本标准由全国信息技术标准化技术委员会(SAC/TC 28)提出并归口。

本标准起草单位:中国电子技术标准化研究院、御数坊(北京)科技咨询有限公司、上海市信息投资股份有限公司、中国科学院计算机网络信息中心、深圳市华傲数据技术有限公司、贵阳信息技术研究院(中科院软件所贵阳分部)、国网浙江省电力有限公司。

本标准主要起草人:卫凤林、宾军志、甘似禹、胡良霖、于文渊、黎俊茂、陈峰、杨达、王静、董建、张群、张展新、赵菁华、李冰、李易昂、秦俊宁、陈利跃。

信息技术 数据质量评价指标

1 范围

本标准规定了数据质量评价指标的框架和说明。
本标准适用于数据生存周期各个阶段的数据质量评价。

2 术语和定义

下列术语和定义适用于本文件。

2.1

数据 data

信息的可再解释的形式化表示,以适用于通信、解释或处理。

注:可以通过人工或自动手段处理数据。

[GB/T 5271.1—2000,定义 01.01.02]

2.2

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。

[GB/T 5271.17—2010,定义 17.06.05]

2.3

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。

2.4

原始数据 raw data

终端用户所存储使用的各种未经过处理或简化的数据。

注:原始数据有多种存在形式,如文本数据,图像数据,音频数据或者几种数据混合存在。

2.5

数据生存周期 data lifecycle

将原始数据转化为可用于行动的知识的一组过程。

2.6

数据集 dataset

具有一定主题,可以标识并可以被计算机化处理的数据集合。

2.7

数据模型 data model

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。

注 1:在从高到低的不同抽象层次表示数据时,通常会区分概念模型(与某些努力相关的概念组成的模型)、逻辑模型和物理模型。

注 2:所使用数据模型的使用周境的边界的正规描述,称为上下文模式。

注 3:数据模型标识实体、域(属性)以及与其他数据的关系(关联),提供数据和数据间关系的概念视图。