



中华人民共和国国家标准

GB/T 21336.1—2023

代替 GB/T 21336—2008, GB/T 21337—2008

地理信息 数据质量 第1部分:总体要求

Geographic information—Data quality—Part 1: General requirements

(ISO 19157-1:2023, MOD)

2023-11-27 发布

2023-11-27 实施

国家市场监督管理总局
国家标准化管理委员会 发布

目 次

| | |
|----------------------------|-----|
| 前言 | III |
| 引言 | V |
| 1 范围 | 1 |
| 2 规范性引用文件 | 1 |
| 3 术语和定义 | 1 |
| 4 缩略语和包缩写 | 4 |
| 4.1 缩略语 | 4 |
| 4.2 包缩写 | 4 |
| 5 一致性 | 5 |
| 5.1 概述 | 5 |
| 5.2 数据质量模型内容 | 5 |
| 5.3 数据质量模型 XML 编码 | 5 |
| 6 对地理信息质量的一般要求 | 6 |
| 6.1 概述 | 6 |
| 6.2 数据质量——一般要求、建议和许可 | 6 |
| 7 数据质量概述 | 7 |
| 8 数据质量组成部分 | 8 |
| 8.1 组成部分概述 | 8 |
| 8.2 数据质量单元 | 9 |
| 8.3 数据质量元素 | 10 |
| 8.4 扩展数据质量信息模型 | 12 |
| 8.5 数据质量元素描述符 | 12 |
| 9 数据质量度量 | 16 |
| 9.1 概述 | 16 |
| 9.2 标准化数据质量度量 | 16 |
| 9.3 用户定义的数据质量度量 | 19 |
| 10 数据质量评价 | 19 |
| 10.1 数据质量评价过程 | 19 |
| 10.2 数据质量评价方法 | 21 |
| 10.3 综合和派生 | 22 |
| 11 数据质量报告 | 22 |
| 11.1 概述 | 22 |
| 11.2 特殊情况 | 23 |
| 12 XML 编码要求 | 24 |
| 附录 A (规范性) 抽象测试套件 | 26 |
| A.1 数据产品规范的内容 | 26 |

| | |
|-----------------------------|----|
| A.2 XML 编码 | 26 |
| 附录 B (资料性) 数据质量概念及其应用 | 27 |
| B.1 数据质量概念框架 | 27 |
| B.2 数据集结构和质量描述的组成部分 | 27 |
| B.3 何时使用质量评价程序 | 28 |
| B.4 报告质量信息 | 29 |
| 附录 C (规范性) 数据质量的数据字典 | 31 |
| C.1 数据字典概述 | 31 |
| C.2 数据质量包数据字典 | 32 |
| C.3 代码表 | 48 |
| 附录 D (资料性) 评价和报告数据质量 | 50 |
| D.1 简介 | 50 |
| D.2 数据集描述 | 50 |
| D.3 质量评价过程 | 53 |
| D.4 报告数据质量 | 59 |
| D.5 其他示例 | 67 |
| 附录 E (资料性) 用于评价的抽样方法 | 73 |
| E.1 简介 | 73 |
| E.2 批和单位产品 | 73 |
| E.3 样本量 | 73 |
| E.4 抽样策略 | 74 |
| E.5 基于概率抽样 | 76 |
| 附录 F (资料性) 质量元素使用指南 | 80 |
| F.1 概述 | 80 |
| F.2 数据质量元素类别 | 80 |
| F.3 数据质量元素间的关系 | 82 |
| F.4 数据质量元素——应用示例 | 83 |
| F.5 疑难问题讨论 | 86 |
| 附录 G (资料性) 数据质量结果的综合 | 87 |
| G.1 概述 | 87 |
| G.2 100%通过/不通过 | 87 |
| G.3 加权通过/不通过 | 87 |
| G.4 最大/最小值 | 88 |
| 附录 H (规范性) XML 编码描述 | 89 |
| H.1 简介 | 89 |
| H.2 XML 命名空间 | 89 |
| H.3 XML 模式 | 89 |
| 参考文献 | 90 |

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 21336《地理信息 数据质量》的第 1 部分。GB/T 21336 已经发布了以下部分：

——第 1 部分：总体要求。

本文件代替 GB/T 21336—2008《地理信息 质量评价过程》和 GB/T 21337—2008《地理信息质量原则》。与 GB/T 21336—2008 和 GB/T 21337—2008 相比，除结构调整和编辑性改动外，主要技术变化如下：

- 更改了数据质量的组成部分(见 8.1, GB/T 21337—2008 的 5.1), 删除了数据质量量化元素、数据质量量化子元素、数据质量非量化元素的相关表述(见 GB/T 21337—2008 第 5 章)；
- 增加了扩展数据质量信息模型的概念(见 8.4)；
- 更改了数据质量元素描述符(见 8.5, GB/T 21337—2008 的 5.2.3、6.1.6)；
- 增加了数据质量评价结果的产生方式(见 10.3), 以及报告派生的相关规定(见 11.2.2 和附录 D)；
- 增加了数据质量定义的 XML 编码要求(见第 12 章和附录 H)。

本文件修改采用 ISO 19157-1:2023《地理信息 数据质量 第 1 部分：总体要求》。

本文件做了以下编辑性改动：

- 改正了印刷错误：将 6.2 中“第 4 章”更改为“第 5 章”；将 8.5.5 中“元质量评价示例见 E.5.3”更改为“元质量评价示例见 D.5.3”；将 C.1.6 注中“6.5.2”更改为“7.2”；将 D.1 中“E.5 补充若干示例”更改为“D.5 补充若干示例”；将 D.3.5.2 中“逻辑一致性”更改为“逻辑一致性”；将表 D.3 注释 a 中“实际包含的单位产品数目”更改为“多余的单位产品数目”；将表 D.14 第 7 行中“见 E.4.2”更改为“见 D.4.2”；将 D.5.3 中“拓扑测量绝对位置准确度”更改为“地形测量绝对位置准确度”；
- 修改了表 C.2：为质量元素增加“度量”(measure)关联(见表 C.2 第 6 行), 后续类、类属性和关联的行号顺次后延；
- 将部分示例更改为符合我国国情的案例：图 D.1~D.4、D.3.5.1 中，将道路名称和住户名称更改为中国名称；E.4.3.2 示例中，将“UTM”更改为“高斯-克吕格投影”；F.3.3 中，将特定地块的价格单位“\$”更改为人民币“¥”；F.4.1.2 示例 1 中，将数据集范围“England”“Scotland”更改为“河北省”“德州市”；F.4.1.3 中，将“England and Wales”更改为“德州市”；
- 删除了附录 I(资料性)“与 ISO 19157:2013 比较的向后兼容性”。

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。

本文件由中华人民共和国自然资源部提出。

本文件由全国地理信息标准化技术委员会(SAC/TC 230)归口。

本文件起草单位：国家基础地理信息中心、国家测绘产品质量检验检测中心、中国农业科学院农业资源与农业区划研究所、北京中科数遥信息技术有限公司。

本文件主要起草人：刘若梅、周旭、陶舒、贾云鹏、郭建坤、程滔、李雪梅、李广泳、杜娟、张莉、张鹤、姚艳敏、毕建涛、杨刚。

本文件所代替文件的历次版本发布情况为：

GB/T 21336.1—2023

- GB/T 21336, 2008 年首次发布；
- GB/T 21337, 2008 年首次发布；
- 本次为第一次修订。

引 言

地理数据正越来越多地被共享、交换和使用,远超出数据生产者预期的目的。由于数据的价值与其质量直接相关,在选择地理数据的过程中,地理数据质量信息显得至关重要。对地理数据用户而言,可能会有多个数据集供其选择,用户通过比较数据集的质量来确定最能满足其需求的数据集。

描述地理数据质量的目的是便于比较和选择最适合应用需求的数据集。数据集质量的完整描述能促进数据集的共享、交换和应用。地理数据质量信息使数据生产者能够评价其数据集对产品规范规定的符合程度,并帮助数据用户评价产品满足其特定应用要求的能力。为评价数据集质量,需要按一致的方式采用明确定义的评价程序。

为了便于对比,应用可比较的方式表达质量报告结果,并对所使用的数据质量度量达成共识。这些数据质量度量通过与论域的比较,提供了地理数据质量描述符。若使用不兼容的度量,质量对比将无法开展。GB/T 21336 旨在提供定义地理数据质量的框架、实现方法与维护注册要求等。拟由 3 个部分构成。

- 第 1 部分:总体要求。确立描述地理数据质量的原则,规范数据质量度量的组成和结构,定义通用的数据质量度量,给出评价数据质量的通用程序。
- 第 2 部分:XML 模式实现。定义数据质量 XML 编码。
- 第 3 部分:数据质量度量注册簿。规定根据第 1 部分建立、维护和发布数据质量度量注册簿的程序。

本文件认为数据生产者与数据用户可从不同的角度看待数据质量。因此,可按照数据生产者的产品规范或数据用户的数据质量要求确定一致的质量水平。如果数据用户要求的数据质量信息多于数据生产者提供的信息,则数据用户可按照数据生产者的数据质量评价流程来获取附加信息。在这种情况下,为了使用数据生产者的数据处理流程,数据用户的要求被视为产品规范。

本文件旨在提供定义地理数据质量的框架,包括地理数据质量的原则、处理地理数据质量信息的概念、数据质量度量的结构和内容,以及报告质量信息的指导方案。该框架包含如何增加数据质量度量的规则,可通过这种方式进行扩展,也可用来定义具有复杂维度的数据质量。

地理信息 数据质量 第 1 部分：总体要求

1 范围

本文件确立了描述地理数据质量的原则。本文件：

- 定义描述数据质量组成部分的完备体系；
- 明确其他特定领域定义描述数据质量组成部分的过程；
- 规定数据质量度量的组成部分和内容结构；
- 给出评价地理数据质量的通用程序；
- 确立报告数据质量的原则。

本文件还规定了一组用于评价和报告数据质量的数据质量度量。

本文件未涉及地理数据的可接收最低质量水平。此类信息通常作为数据产品规范中的一项要求出现，可依据 ISO 19131 定义。

本文件适用于数据生产者提供质量信息，描述和评价其数据集符合产品规范的程度，以及数据用户探究特定的地理数据质量是否满足其特定应用要求。

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 19710.1—2023 地理信息 元数据 第 1 部分：基础(ISO 19115-1:2014,MOD)

注：GB/T 19710.1—2023 被引用的内容与 ISO 19115-1:2014 被引用的内容没有技术上的差异。

GB/T 35647—2017 地理信息 概念模式语言(ISO 19103:2015,IDT)

ISO 19109:2015 地理信息 应用模式规则(Geographic information—Rules for application schema)

3 术语和定义

下列术语和定义适用于本文件。

3.1

准确度 accuracy

测试结果或测量结果与真值间的一致程度。

注：本文件中，真值可以是可接受为真的参照值。

[来源：GB/T 3358.2—2009,3.3.1,有修改]

3.2

一致性 conformance

conformity

满足要求。

注：无歧义时，修饰语“一致性”可省略。例如，“测试报告”等同“一致性测试报告”。

[来源：ISO 19105:2022,3.4]