



中华人民共和国国家标准

GB/T 18391.4—2009/ISO/IEC 11179-4:2004
代替 GB/T 18391.4—2001

信息技术 元数据注册系统(MDR) 第4部分:数据定义的形成

Information technology—Metadata registries (MDR)—
Part 4: Formulation of data definitions

(ISO/IEC 11179-4:2004, IDT)

2009-09-30 发布

2009-12-01 实施

中华人民共和国国家质量监督检验检疫总局
中国国家标准化管理委员会 发布

目 次

前言	I
引言	II
1 范围	1
2 一致性	1
3 术语和定义	1
4 数据定义要求与建议的概要	2
4.1 要求	2
4.2 建议	3
5 规定	3
5.1 前提	3
5.2 要求	3
5.3 建议	5
参考文献	7

前 言

GB/T 18391《信息技术 元数据注册系统(MDR)》分为六个部分：

- 第 1 部分：框架；
- 第 2 部分：分类；
- 第 3 部分：注册系统元模型与基本属性；
- 第 4 部分：数据定义的形成；
- 第 5 部分：命名和标识原则；
- 第 6 部分：注册。

本部分为 GB/T 18391 的第 4 部分，等同采用 ISO/IEC 11179-4:2004《信息技术 元数据注册系统(MDR) 第 4 部分：数据定义的形成》(英文版)，仅有编辑性修改。

本部分代替 GB/T 18391.4—2001《信息技术 数据元的规范与标准化 第 4 部分：数据定义的编写规则与指南》。

本部分与 GB/T 18391.4—2001 相比主要变化如下：

- 名称变更为：《信息技术 元数据注册系统(MDR) 第 4 部分：数据定义的形成》；
- 4.1 和 5.2 中的“规则”改为“要求”；
- 4.2 和 5.3 中的“指南”改为“建议”；
- 取消了 4.1a) 具有唯一性(在出现此定义的任何数据字典中)；
- 增加了 4.2h) 适合被定义的元数据项的类型。

为了保持与英文版本的一致性，英文版中斜体表示的内容本部分也使用斜体表示，第 3 章中术语的英文名称保留英文版本中的大小写格式。

本部分由中国标准化研究院提出。

本部分由全国信息技术标准化技术委员会归口。

本部分由中国标准化研究院负责起草。

本部分主要起草人：刘植婷、邢立强、李小林、史立武、孙广芝。

本部分于 2001 年首次发布，本次为第一次修订。

引 言

数据处理和电子数据交换很大程度上依赖于数据库中数据记录的准确性、可靠性、可控性和可校验性。数据的正确而恰当的使用与解释的前提是数据的提供者和使用者都对数据的含义和表示有共同的理解。为了促进这种共同的理解,应定义数据的若干特性或属性。这些数据的特性通称为“元数据”,即“描述数据的数据”。GB/T 18391 的本部分给出了基于元数据注册系统的数据定义形成的要求与建议。这些定义的目的在于确定、描述、解释和澄清数据的含义,以促进数据元的标准化和可重复使用,并促进信息系统的数据共享与集成。

元数据注册系统的结构用一个概念数据模型的形式来规定。元数据注册系统被用来保存数据元和有关概念的信息,例如,“数据元概念”、“概念域”和“值域”。通常,它们都被看作是“元数据项”。这样的元数据对于清晰地描述、记录、分析、分类和管理数据是必需的。

GB/T 18391 的本部分规定的定义要求与建议并不总是适用于在词汇和语言字典中出现的术语定义。应用在语言字典和元数据注册系统中的要求之间存在着差异。GB/T 18391 的要求比自然语言字典的要求有更多的限制。例如,语言字典可以有包含术语或字词的多种意义的多种定义,然而为特定语境所研制的元数据定义不宜在该语境内具有多种意义。元数据定义是用来说明一个概念或多个概念的,这些概念通过数据集、数据值、数据元或其他元数据项来表示。可以确立一个单一的参考定义,其他定义可以声明与其等价(如,可以在一种语言中建立一个参考定义,其他语言的定义可以声明与其等价)。元数据项可以在特定语境内有一个首选定义,不赞成使用其他的定义。

许多元数据定义包含着自身需要被定义的术语(例如,“费用”,“津贴”,“递送”)。某些术语在不同的行业部门中可能具有不同的定义。因此,大多数元数据注册系统有必要对定义所用术语建立一个关联词汇表。

信息技术 元数据注册系统(MDR)

第4部分:数据定义的形成

1 范围

GB/T 18391 的本部分规定了构建数据和元数据定义的要求与建议。本部分仅描述了定义的语义方面,而不考虑定义的格式。尤其当用于 GB/T 18391 第 3 部分所规定的元数据注册系统的内容时,可广泛用于确定数据和元数据的定义。

这些要求与建议适合于形成数据元和其他数据构件类型的定义,这里的数据构件类型可以是实体类型、实体、关系、属性、对象类型(或类)、对象、复合物、代码条目、元数据项,以及由 XML 标记引用的数据。

2 一致性

GB/T 18391 的本部分可以独立地使用,例如,用于定义一个元数据注册系统语境外的数据元。如果在研制定义中已经遵循了这些要求与建议,就可以声明与本部分一致。

在 GB/T 18391 元数据注册系统的语境中,当确立 GB/T 18391.6 中所规定的注册状态时,本部分应构成关于定义的规则。

3 术语和定义

下列术语和定义适用于 GB/T 18391 的本部分。

3.1

概念域 **Conceptual Domain**

有效的值含义的一个集合。

注:概念域中的值含义既可枚举也可通过描述表达。

[GB/T 18391.3—2009,定义 3.3.21]

3.2

概念 **Concept**

通过对特征的独特组合而形成的知识单元。

[GB/T 15237.1—2000,定义 3.2.1]

3.3

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:数据可以由人工或自动的方式加工、处理。

[GB/T 5271.1—2000,定义 01.01.02]

3.4

数据元 **data element**

由一个属性集合规定其定义、标识、表示和允许值的一个数据单元。

[GB/T 18391.3—2009,定义 3.3.36]

3.5

数据元概念 **Data Element Concept**

以一个数据元的形式来表示的一个概念,其描述与任何特定表示法无关。

[GB/T 18391.3—2009,定义 3.3.38]