



中华人民共和国国家标准

GB/T 18391.1—2002/ISO/IEC 11179-1:1999

信息技术 数据元的规范与标准化 第 1 部分：数据元的规范与标准化框架

Information technology—Specification and standardization
of data elements—Part 1: Framework for the specification
and standardization of data elements

(ISO/IEC 11179-1:1999, IDT)

2002-07-18 发布

2002-12-01 实施

中华人民共和国
国家质量监督检验检疫总局 发布

目 次

前言	I
ISO/IEC 前言	II
引言	III
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 缩略语	7
5 方法论无关性	7
6 数据元的基本概念	7
7 数据元与其他数据概念的关系	9
8 GB/T 18391 第1部分至第6部分综述	9
附录 A(资料性附录) 数据表示和管理的基本概念	13
附录 B(资料性附录) 元数据和数据元	23

前 言

GB/T 18391 在总标题《信息技术 数据元的规范与标准化》下,包括以下几部分:

- 第 1 部分:数据元的规范与标准化框架;
- 第 2 部分:数据元的分类;
- 第 3 部分:数据元的基本属性;
- 第 4 部分:数据定义的编写规则与指南;
- 第 5 部分:数据元的命名和标识原则;
- 第 6 部分:数据元的注册。

本部分等同采用国际标准 ISO/IEC 11179-1:1999《信息技术 数据元的规范与标准化 第 1 部分:数据元的规范与标准化框架》。

本部分的附录 A 和附录 B 是资料性附录。

本部分由中国标准研究中心提出并归口。

本部分起草单位:中国标准研究中心。

本部分主要起草人:刘植婷、邢立强、史立武。

ISO/IEC 前言

国际标准化组织(ISO)和国际电工委员会(IEC)是世界性的标准化专门机构。作为 ISO 或 IEC 成员的国家团体机构,通过各类技术委员会参与国际标准的研制。技术委员会由参加各类专业领域技术活动的相关组织组建而成。ISO 和 IEC 的技术委员会在共同感兴趣的领域进行合作。同 ISO 和 IEC 有联系的官方与非官方国际组织,也可以参与该项工作。

在信息技术领域,ISO 和 IEC 已组建了联合技术委员会,即 ISO/IEC JTC1。由联合技术委员会通过的国际标准草案,提交给各国家团体进行表决。一项国际标准的颁布,至少需要 75% 参与表决的国家团体投赞成票。

需要提请注意的是:ISO/IEC 11179 本部分中某些内容可能已经获得了专利权。ISO 和 IEC 并不对全部或部分具有专利权的内容负责。

国际标准 ISO/IEC 11179-1 是由信息技术联合技术委员会(ISO/IEC JTC1)下属的数据管理与交换分委员会(SC32)负责起草的。

ISO/IEC 11179 在总标题《信息技术 数据元的规范与标准化》下,包括以下部分:

- 第 1 部分:数据元的规范与标准化框架;
- 第 2 部分:数据元的分类;
- 第 3 部分:数据元的基本属性;
- 第 4 部分:数据定义的编写规则与指南;
- 第 5 部分:数据元的命名和标识原则;
- 第 6 部分:数据元的注册。

本标准的附录 A 和附录 B 是资料性附录。

引 言

背景

人类通过自然界中事物的特性来认识事物,而这些特性则通过数据来表示。数据的基本单元——数据元的规范,涉及到为每一个数据元的相关特性进行规范化说明,以确保自然界事物表示的一致性与准确性。经过仔细、认真规范和标准化的数据将会大大加强其在系统间及环境间的实用性和共享性。数据共享包括所需数据的查询检索以及与其他用户的数据交换。当数据元按照 GB/T 18391 的要求进行编写,且其文件在一个数据元登记系统得到管理时,则从分散的数据库中查找、检索以及通过电子通信手段发送和接收它们将变得相当简便。

通过自动信息处理系统对用于交流的数据元进行识别和标准化,是一项正在进行的重要工作。该项活动的成功及其在世界范围内的应用将对提高政府、企业界及学术界间的交流起到至关重要的作用。

GB/T 18391 的制定致力于解决以下几个最基本的数据共享与标准化问题(但不仅限于此):

- 缺乏促使全球数据获得与交换的机制,特别是在应用领域;
- 全球唯一的标准数据元的标识符尚不存在;
- 有关数据元特性的文档尚不足以支持数据全面自动化的共享,包括数据的查找、检索与交换;
- 目前尚未形成关于数据元标识、开发和描述的统一指南;
- 在成千上万的数据元中查找和检索某一具体的标准数据元是非常困难或者是几乎不可能的事情;
- 目前尚不存在通用的方法对标准数据元进行有效的组织;
- 尽管有些数据在一个机构内已经标准化,但机构间公用的数据标准却很少;
- 机构间数据的交换导致了专用数据交换表示方法的激增;
- 数据定义和描述的精确程度尚不能支持数据的再利用以及用户多样化的需求;
- 现有用于缩减冗余逻辑数据的目录结构尚不能满足需要;
- 由于标准数据元的缺乏而使全球实施电子数据交换(EDI)受到阻碍;EDI 报文的内容需要标准数据元。

为了推动电子通信的全球化,国际标准界通过细心而持续不断的努力确定了开放系统互连环境(OSIE)。在该环境内,不同的计算机硬件及应用可实现信息共享。开放性信息处理系统所需的四个基本要素(硬件、软件、通信和数据)中的三个(硬件、软件和通信)要素,已经或正在制定为相关标准,而 GB/T 18391 作为数据规范,为开放性信息系统的第四个基本要素,提供了在 OSIE 内实现数据共享的机制。

系统要做到真正的开放,数据必须是易于交换的且可以在各种不同的应用环境内及其相互间进行共享,它可以跨越区域化的网络和分布式的网络。为了数据的共享,数据的使用者与拥有者必须对其含义、表示和标识的理解是一致的。例如,为了理解数据的含义,数据使用者必须在数据元登记系统获得关于数据的描述。数据必须得到充分的描述,并应使使用者可以通过很方便的途径得到这些描述。数据元登记系统提供了对数据元内容及表示进行组织的方法,以确保数据描述的一致性,并使数据设计者和使用者较为方便地对其进行定位。数据的统一规范化将为整个软件开发生命周期中的数据检索、交换和一致性的使用提供便利。具有规范意义和形式的信息单元则被称为“已标准化数据元”。

GB/T 18391 的目的

GB/T 18391 描述数据元标准化和注册,使理解和共享数据成为可能。相对于传统的数据管理办法,以 GB/T 18391 来描述数据元标准化及其注册将使一个共享数据环境的产生花费更少的时间和精力。

GB/T 18391 的目的在于如何就离散性数据元的描述及其语义内容(元数据,它以一种一致、标准的方式来表达数据元)的表达与维护提供具体的指南。它还就数据元注册系统的建立提供指南。

GB/T 18391 的初衷是通过电子信息交换以满足国际范围数据开放化交换的需要,它还能:

- 便利数据的获取和注册;
- 加速数据的访问及应用;
- 依据元数据描述的特性,运用智能软件简化数据操作;
- 促成用于 CASE 工具和数据存储的数据表示元模型的开发;
- 方便电子数据的交换与数据的共享。

GB/T 18391 将使以下信息系统和人们之间的数据传递获益:

- 一个机构内;
- 不同的机构间;
- 跨越所有软件、硬件层面和地理、机构及政治的界限。

关于数据元的元数据被存储于数据元注册系统中。数据元注册系统支持数据及其描述的共享。注册是将支持数据共享的元数据文件化的过程。注册应在数据元层面进行,使其含义得到最大程度的理解。GB/T 18391 将使最终用户能自信、准确、无歧义地解释其所想要表达的意思。

GB/T 18391 的用户

GB/T 18391 为数据用户和管理者阐明了共享数据所必需的数据元特性中的一个基本集。并对重要的数据元特性如标识符、定义及分类类型做了特意的强调。GB/T 18391 对数据元注册系统的描述,旨在帮助共享数据的用户对数据元的含义、表示和标识有一个共同的理解。如果已获得数据值,用户就可以找到该数据的确切含义。用户还可以通过辨识所需数据的类型从数据库中检索所期望的数据值。

GB/T 18391 为系统分析者和数据管理者提供了能满足需求的数据元的再利用途径,或如果一个数据元不存在,设计新的数据元的途径。在用户从数据库中获得数据元以前,数据服务管理者和系统分析者必须对数据进行逻辑标识和描述,以确保用户不会无意中引进不一致的数据值。如果数据分析者想生成一个共享数据的产品,他们首先应该知道所需特性的数据元是否存在。若存在,他们可以使用它。如果系统分析者复制该数据元,就必须以包含相同信息的相同方式表示数据元。若不存在具有恰好相同特性的数据元,数据管理者就需要设计数据元使其描述适合于软件开发者的需求。GB/T 18391 的目的在于开发对数据元的精确描述。依照这个由多个部分组成的国家标准的原则产生的数据元,可以进行数据交换和检索,而与其所使用的信息处理系统和通信协议无关。

GB/T 18391 为软件开发者提供了确保数据一致的手段。注册系统可以使软件开发者在整个软件开发生命周期(SDLC)中保证数据使用的一致性。注册系统将提供数据元管理的机制并使其在软件开发生命周期(SDLC)的不同阶段具有可追溯性。

GB/T 18391 为数据字典开发者、数据元注册系统、计算机辅助软件工程(CASE)工具和其他数据管理软件提供了用于设计元模型的基础,而元模型又是获取、存储、管理和交换数据元的元数据所必需的。

信息技术 数据元的规范与标准化

第 1 部分:数据元的规范与标准化框架

1 范围

GB/T 18391 规定了数据元组成的基本内容,包括元数据。GB/T 18391 的本部分适用于人、机共享数据元的规范化表示和含义,不适用于机器级上数据以位和字节为单位的物理表示。

本部分提供了关联部分的相关环境和从概念上理解数据元的基础。

2 规范性引用文件

下列文件中的条款通过 GB/T 18391 本部分的引用而成为本部分的条款。凡是注日期的引用文件,其随后所有的修改单(不包括勘误的内容)或修订版均不适用于本部分,然而,鼓励根据本部分达成协议的各方研究是否可使用这些文件的最新版本。凡是不注日期的引用文件,其最新版本适用于本部分。

GB/T 5271.4—2000 信息技术 词汇 第 4 部分:数据的组织(idt ISO/IEC 2382-4:1987)

GB/T 10112—1999 术语工作 原则与方法(neq ISO/DIS 704:1997)

GB/T 15237.1—2000 术语工作 词汇 第 1 部分:理论与应用(eqv ISO 1087:1-2000)

GB/T 18391.2——¹⁾ 信息技术 数据元的规范与标准化 第 2 部分:数据元的分类
(idt ISO/IEC 11179-2:2000)

GB/T 18391.3—2001 信息技术 数据元的规范与标准化 第 3 部分:数据元的基本属性
(idt ISO/IEC 11179-3:1994)

GB/T 18391.4—2001 信息技术 数据元的规范与标准化 第 4 部分:数据元定义的编写规则与指南(idt ISO/IEC 11179-4:1995)

GB/T 18391.5—2001 信息技术 数据元的规范与标准化 第 5 部分:数据元的命名和标识原则
(idt ISO/IEC 11179-5:1995)

GB/T 18391.6—2001 信息技术 数据元的规范与标准化 第 6 部分:数据元的注册(ISO/IEC 11179-6:1997)

GB/T 20001.1—2001 标准编写规则 第 1 部分:术语(ISO 10241:1992,International terminology standards—Preparation and layout,NEQ)

ISO 标准手册 10 数据处理 词汇 1982

3 术语和定义

下表中的术语和定义适用于 GB/T 18391 的各部分。表头中“部分”编号列中的 X 表明该术语在该部分被定义并会在其他章中被使用。以黑体出现于一个术语定义中的词语本身也是一个术语,它将会在该章的其他地方被定义。以普通字体出现的词语采用其被普遍理解的含义。某些词这两种情况都存在。也有的情况是两个或两个以上的术语在一个定义中同时出现,且出现的新术语并未给出定义。在上述情况下,确定术语时是无歧义的。

1) 正在制定中。