

科学数据共享工程技术标准

SDS/T 2132—2004

数据元标准化的基本原则与方法

The basic principles and methods for the standardization
of data elements

（征求意见稿）

（本稿完成日期：2005年5月）

××××-××-××发布

××××-××-××实施

中华人民共和国科学技术部 发布

目 次

前 言.....	IV
1 范围.....	1
2 规范性引用文件.....	1
3 术语和定义.....	1
4 数据元概念.....	4
4.1 数据元基本概念.....	4
4.2 数据元与其它数据概念的关系.....	5
5 数据元的表达格式.....	6
5.1 数据元基本属性.....	6
5.2 数据元属性描述符.....	7
6 数据元的提取.....	7
6.1 自上而下的数据元提取法.....	7
6.2 自下而上的数据元提取法.....	8
7 数据元的分类.....	8
7.1 数据元分类方案.....	8
7.2 数据元的分类属性.....	9
8 数据元的标识.....	9
8.1 数据的标识结构原则.....	9
8.2 数据注册标识的规则.....	10
9 数据元的命名规则.....	11
9.1 名称语义内容的规则.....	11
9.2 名称格式的规则.....	13
9.3 唯一性规则.....	14
10 数据元的定义.....	14
10.1 数据元定义规则.....	14
10.2 数据元定义指南.....	14
11 数据元值域及其表达模式.....	15
11.1 穷举域.....	15
11.2 可重复利用域.....	15
11.3 数据元值域的规范定义和来源.....	15
12 数据元目录标准编写格式.....	16
12.1 数据元目录标准格式.....	16
12.2 数据元目录的一般格式.....	17
附录 A (规范性附录).....	18
A.1 数据元属性的描述方法.....	18
A.2 标识类属性.....	18
A.3 定义类属性.....	19
A.4 关系类属性.....	20
A.5 表示类属性.....	21
A.6 管理类属性.....	23

A.7 附加类属性	23
附录 B (规范性附录)	24
B.1 数据元的标识	24
B.2 数据元的注册	24
B.3 注册系统	25
B.4 注册参加机构	26

前 言

科学数据共享活动既涉及科学数据的汇交、分发、服务和应用环节，也涉及科学数据的采集、加工和处理环节。在这一系列环节中，常见的工作内容之一是用数据元（如数据库中的数据字段）的形式确立某种类别的数据在其名称、含意、表示格式、标识等方面的特征，用数据的分类与编码（如科学数据分类与代码）的形式确立对某种类别的数据所作的进一步的分类并对分类结果赋予特定代码，以达到对该种数据类别或其分类结果在语义上的无歧义理解。但需要特别强调的是：一方面，数据元遴选和确立的好坏，将直接影响数据组织、生产、交换和共享的效率和质量，严重时还会导致重复的投入和建设；另一方面，数据元的遴选和确立结果要被不同的信息系统或应用系统所共享，就必须在各相关方之间建立大家一致认可的相关标准或规范，否则会导致大量的转换工作甚至无法实现共享。为此，采用何种原则和方法开展数据元标准化工作，可有效规避风险，提高数据组织和生产的效率和质量，达到共享目标，是科学数据共享活动面临的任务之一。

本标准正是基于上述现实要求而提出和制定的，其目的是确立面向科学数据共享活动的数据元标准化的基本原则与方法，指导和规范科学数据共享工程各建设单位研究和制定与各种科学数据的采集、加工、处理、汇集、分发、服务和应用相关的通用或专用的数据元目录（又称字典）标准，推动和促进科学数据的集成与共享。

按照向上积极采用适宜的国际标准或国家标准，向下突出科学数据共享活动的特点和需求的原则，本标准参考了包括 GB/T 18391《数据元的规范与标准化》在内的大量国家标准，并结合了各种类型的科学数据库建设和整合以及科学数据交换、共享、服务和应用对数据元标准化工作的特定需求。

本标准为第一次制定。

本标准由科学数据共享工程办公室提出。

本标准由科学数据共享工程办公室归口。

本标准主要起草单位：

本标准主要起草人：

数据元标准化的基本原则与方法

1 范围

本标准围绕各种类型的科学数据库建设和整合以及科学数据交换、共享、服务和应用对数据元标准化的需要，规定了数据元的提取、命名、标识、描述、分类、值域范围确立的基本原则与方法，数据元目录的格式规范，以及整个科学数据共享工程中数据元的维护与管理层次与机制。

本标准适用于科学数据共享工程各建设单位研究和制定通用和专用的数据元目录（又称数据元字典）标准。

本标准的目的在于给出科学数据共享工程、各领域、主体数据库三个层面的数据以及关联数据，在其特定的语境中，数据单元标准化的原则、方法以及管理的机制和模式。它不仅是科学数据共享中三个层面各自内部数据规范的原则与方法，同时也是相关元数据元素和元数据内容规范的原则与方法。它不仅是科学数据大平台建设中的数据规范、公共数据元目录编写必须遵循的原则与方法，更是各试点单位相关库结构设计和数据字典编写的依据和基础。

2 规范性引用文件

下列文件中的条款通过本标准的引用而成为本标准的条款。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标准，然而，鼓励根据本标准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标准。

GB/T 18391.1 -2002 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 (idt ISO/IEC 11179-1: 1999)。

GB/T 18391.2 - 2003 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(idt ISO/IEC 11179-6: 1997)。

GB/T 19488.1-2004 电子政务数据元 第1部分：设计和管理规范。

3 术语和定义

下列术语和定义适用于本标准

3.1

数据元 data element

用一组属性描述其定义、标识、表示和允许值的数据单元，在一定语境下，通常用于构建一个语义正确、独立且无歧义的特定概念语义的信息单元。

3.2

数据元概念 data element concept

能以数据元的形式表示，且与任何特定的表示法无关的概念。

3.3

数据元目录 data element dictionary

列出并定义了所有相关数据元的一种信息资源。

3.4

数据元名称 data element name

用于标识数据元的主要手段，由一个或多个词构成的命名。

3.5

标识符 identifier

注册机构内与语言无关的唯一标识符。同数据标识符。给定语境的对象的无歧义的名称。

3.6

版本 version

注册机构内，一套逐渐完善的数据元规范中的一个数据元规范发布的标识。

3.7

数据元注册系统 data element registry

由注册机构保存的用于描述数据元含义和表示形式的信息资源，包括注册标识符、定义、名称、值域、元数据和管理属性等。

3.8

实体 entity

任何具体或抽象的事物，包括事物间的联系。

3.9

对象 object

可以想象或感觉的世界的任一部分。

3.10

对象类 object class

对象集。现实世界中的想法、抽象概念或事物的集合，有清楚的边界和含义，并且特性和其行为遵循同样的规则而能够加以标识

3.11

属性 attribute

某个对象或实体的特征。

3.12

定义 definition

表述人和事物的基本特性、或者人或事物类别的词或短语；要回答“X是什么？”或“一个X属于什么？”这样的问题；一个词或词组含义的表述[韦氏世界英语大词典，第三版，1986]。一个数据元基本特性的陈述，并使之区别于其它数据元。

3.13

语境 context

对使用名称或产生名称的应用环境或应用规程的指明或描述。

3.14

特性 property

对象类的所有个体所共有的某种性质。

3.15

特性词 property term

数据元名称的一个部分，用于表述对象类的特性。（数据元名称的一个成分，表述数据元所属类别）

3.16

限定词 qualifier term

帮助定义和区分数据库中某个名称的词或词组。

3.17

表示词 representation term

数据元名称的成分，用于描述数据元的表示形式。

3.18

表示 representation

值域、数据类型的组合，必要时也包括度量单位或字符集。

3.19

注册机构标识符 (RAI) registration authority identifier

赋予注册机构的标识符。

3.20

版本标识符 version identifier

赋予版本的一个标识符。用于提交或修改数据元注册。

3.21

分类方案 classification scheme

依据对象所具有的共同特征，如：起源、组成、结构、应用和功能，而对其进行的排列或划分。

3.22

关键字 keyword

用于数据元检索的一个或多个有意义的字词。

3.23

分类法 taxonomy

依据类及子类间的固有关系进行分类的方法。

3.24

主题词表 thesaurus

按给定顺序排列参照词汇，其中显示和标识了词汇间的关系。

3.25

分类的成分 classified component

可以依据一个以上分类方案进行分类的数据元的成分。这些成分包括对象类、特性、表示类、数据元概念、值域和数据元。

4 数据元概念

4.1 数据元基本概念

数据元，又称数据类型，通过一组属性描述其定义、标识、表示和允许值的数据单元。在特定的语义环境中被认为是不可再分的最小数据单元。

数据元一般来说由三部分组成：

a) 对象类：思想、概念或真实世界中的事物的集合，它们具有清晰的边界和含义，其特征和行为遵循同样的规则。

b) 特性：对象类中的所有成员共同具有的一个有别于其他的、显著的特征。

c) 表示：它描述了数据被表达的方式。

对象类是人们希望研究、搜集和存储它们的相关数据的事物，比如汽车、人、房屋、订单等。

特性是人们用来区分和描述对象的一种手段。特性的例子包括颜色、性别、年龄、收入、地址等等。

表示与数据元的值域关系密切。一个数据元的值域是数据元的所有允许值的集合。

一个数据元概念由一个对象类与一个特性的组合。因此，一个数据元是由一个数据元概念和一个表示组成。

图 4-1 给出了数据元的结构模型。

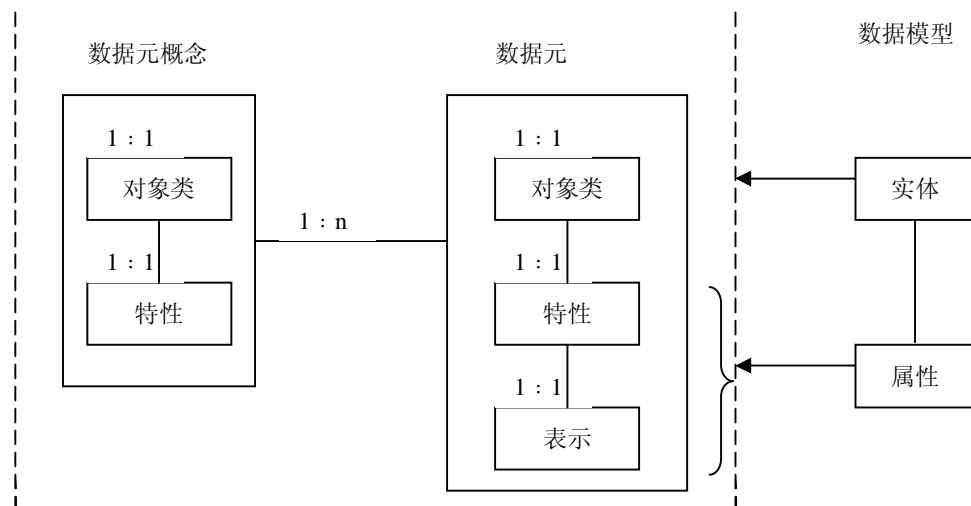


图 4-1 数据元的结构模型

数据元由数据元概念和表示两部分组成。当一个表示被联合到一个数据元概念时就能够产生一个数据元。根据图中所示，数据元和数据元概念之间存在多对一的关系，也就是一个数据元必须要一个数据元概念，而一个数据元概念可以对应多个数据元，换句话说，多个数据元可以共享一个数据元概念。

数据元需要并且只能有一个表示。当数据元的概念模型相同而表示不同时就是两个不同的数据元，数据元中的表示是描述数据元中的数据元概念中的特性的。也就是说在数据元中，特性都具有且只有一个表示。

在数据元概念中对象类和特性之间是一一对应的关系，一个对象类需要且只需要一个特性（或者特性类），一个特性只描述一个对象类，当一个特性和一个对象类建立关联时就产生了一个数据元概念。

同实体关系类的数据模型相比，模型中的实体相当于数据元中的对象类，而实体的属性相当于数据元中的特性和表示。在面向对象的词中，数据元概念还可以表示为对象类和特性的组合。另一方面，在实体/关系的术语中，数据元概念是实体的属性。

4.2 数据元与其它数据概念的关系

图 4-2 是数据层级的简化表示，表明数据元所在的那些层。数据元出现在数据库、文件和事务集中。数据元是一个组织管理数据的基本单元，因而它必然是组织内部数据库和文件设计，并用于建立与其它组织交流的事务集的组成部分。

在组织内部，数据库或文件由记录、段和元组等组成，而记录、段和元组则由数据元组成。数据元本身包含有字符、图像、声音等多类数据。

一个组织需要将数据传输给其它组织时，数据元构成了事务集的基本单元。事务主要发生于数据库间或文件间，但组织间的文件和数据库结构（如记录或元组）并不一定相同。信息（数据加上理解）传输的公共单元是数据元。

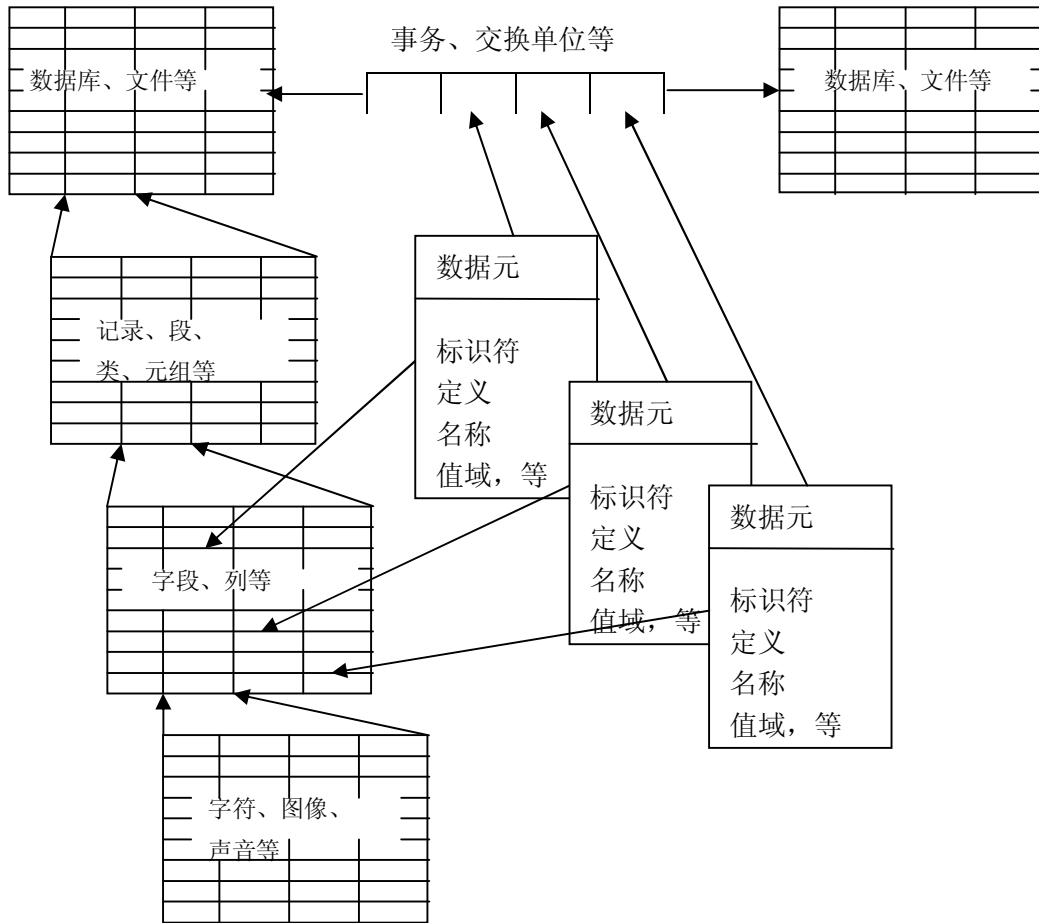


图 4-2: 各数据层级的数据元

5 数据元的表达格式

5.1 数据元基本属性

一个数据元规范由一组属性组成。本标准规定了一组基本属性（参见表5-1）。纵列“约束”表示：在数据元字典中，一个属性是“必选（M）”，还是“条件选（C）”，或者是“可选（O）”。

表5-1 数据元属性

属性种类	数据元属性名称	约束
标识类	—— 名称	M
	—— 标识符	C
	—— 版本	C
	—— 注册机构	C
	—— 同义名称	O
	—— 语境	C
定义类	—— 定义	M

关系类	—— 分类方案 —— 关键字（词） —— 相关数据参照 —— 关系类型	O O O C
表示类	—— 表示类别 —— 表示形式 —— 数据元值的数据类型 —— 数据元值的最大长度 —— 数据元值的最小长度 —— 表示格式 —— 数据元允许值	M M M M M C M
管理类	—— 责任机构 —— 注册状态 —— 提交机构 —— 备注	O C O O

5.2 数据元属性描述符

数据元属性应依照一种标准方式来注册和控制，以便数据元字典中的数据元在信息交换中保持一致性，并且能够在不同的数据管理环境中进行数据元比较。表 5-2 给出了关于描述数据元属性的常规描述符集。纵列“约束”表示描述数据元属性时，一个描述符是“必选”（M），还是“条件选”（C），或者是“可选”（O）。

表 5-2 常规描述符

属性描述符	约束
—— 名称（中、英文）	M
—— 定义	M
—— 约束	M
—— 条件	C
—— 最多实例数	O
—— 数据类型	M
—— 最大长度	O
—— 字符集	C
—— 语言	C
—— 备注	O

6 数据元的提取

为科学数据共享数据元提取提供一个方法论指南是确保提取数据元具有科学性和互操作性的关键。科学数据共享数据元的提取方法有两种：自上而下（top-down）提取法和自下而上（down-top）提取法。6.1 条节和 6.2 条节将分别简要介绍这两种数据元提取方法。

6.1 自上而下的数据元提取法

对于新型领域的数据元提取，一般适用这种“自顶而下”的提取法。基本步骤是，在业务流程和功

能分析的基础上，通过建模分析、确立关心“对象”，从数据方案的角度看，数据元主要解决数据的概念和组成。在数据模式三个层次中概念信息模式和逻辑数据模式的基础上，分析提取数据元及其属性。一般步骤为：业务功能建模、业务流程建模、信息建模、数据元的提取、数据元的提交。

6.2 自下而上的数据元提取法

对于已有的科学数据共享系统，可根据其自身数据库系统的ERD图进行数据元的提取。对于科学数据可能存在的信息模型、数据模型、数据流程图、数据库设计（模式、文件、格式）以及接口规范和计算机程序中数据元都可以进行共享。对这些系统中的已存在的数据元要依据本标准，结合科学数据业务和管理要求，逐部门、系统收集、筛选、梳理、排重、找出共性，在协调的基础上定义、分类整理，以实现标准化。

7 数据元的分类

对数据元进行分类有几个目的：分类可帮助用户从众多的数据元中找出某个单一的数据元；方便对数据元进行数据管理分析；通过继承使原本借助其他属性（如名称和定义）不能完整表述的语义内容得以表达。

国际规范、通行的数据元分类一般通过数据元的分类方案和数据元的分类属性来进行。

7.1 数据元分类方案

规范的数据元的分类方案一般包括：关键字、主题词表、分类法和本体论，其主要作用在于：

- 派生和形成抽象数据元和应用数据元；
- 确保适当属性和属性值的继承；
- 从参照词汇表中派生名称；
- 消除歧义；
- 辨识上位类、同位类和下位类的数据元概念；
- 辨识数据元概念和数据元之间的关系；
- 辅助模块化设计的名称和定义的开发。

7.1.1 关键字

关键字作为基本属性可应用于对象类、特性、表示、数据元和数据元概念。关键字的描述如下：

名称：关键字

定义：用于数据元检索的一个或多个有意义的字词。

约束：条件选

数据类型：字符串

7.1.2 主题词表术语

主题词表术语能够与数据元和数据元概念关联起来。本标准并没有规定主题词表的结构。7.2 中的属性可以用来捕捉与数据元或数据元概念关联的主题词表信息。

7.1.3 分类法和本体论分类单元

分类法是基于概化或特化以及集、子集和集隶属关系这样的数学概念的概念或分类单元的层次结构。本体论是关于分类单元的网状结构，目的在于为自然界某些部分提供模型，由关于对象的类别、对象特性以及自然界中该部分对象间可能联系的理论组成。一个本体论可以包括对分类单元的解释以及对符合语法规则的使用做出限定的正式通则。分类法和本体论中的分类单元可能与下列已分类的数据注册成分相关联：对象类、特性、表示类和数据元概念。本部分并不阐明分类法或本体论的结构，但是，列于7.2中的属性将被用来把握此类结构信息。

7.2 数据元的分类属性

数据元的若干成分导致了分类的需求。主要的成分包括对象类、特性、表示、值域、数据元概念以及数据元本身。属性的使用应使各分类方案与数据元选定的成分联系起来。记录一个数据元任何成分的分类信息可能用到下列属性：

- 被分类成分标识符
- 被分类成分名称
- 分类方案类型
- 分类方案名称
- 分类方案版本
- 分类方案项类型
- 分类方案项值

每个成分与分类方案的关联使用可以使数据元应用的相关方能够更好地运用分类方案：

- 分析对象类、数据元概念和数据元
- 在对象类、特性、表示、数据元概念以及数据元间进行比较
- 减少数据元概念和数据元的多样性
- 明确无误地识别、描述和定义数据元概念和数据元
- 辅助用于指定注册状态的数据元分析
- 解决同义词和多义词的有关问题
- 于数据注册处检索数据元概念和数据元
- 辨识数据元概念和数据元之间的关系

——支持唯一的且明确无误的标识和引用对象类、数据元概念和数据元，其形式应做到语言上的中立和应用到信息技术上的可能。

8 数据元的标识

8.1 数据的标识结构原则

8.1.1 属性标识

为了区分不同的数据元，用一组相关的5个属性，对每个数据元进行命名和标识。这些属性是：

- 名称；
- 语境；
- 注册机构标识符；
- 数据标识符；
- 版本标识符。

8.1.2 名称和语境

在注册机构中注册的一个数据元至少应当有一个名称，根据该数据元使用的语境，可以分配多个名称。每个名称在某个特定的语境中有着特殊的作用。为了便于数据管理，可建立严格结构化的名称；使用者可以规定首选名称；在特定的软件环境（如：一个特定程序语言或数据库管理系统）中可以产生简称。

在每个语境中可以为许多数据元赋予名称，每个语境建立一个命名约定（一般是一组规则），用以详细说明在该语境中如何规范名称。一个命名约定应包括语境的全部有关方面，这包括：

- 命名约定的范围，如“行业首选名称”；
- 确定名称的机构；
- 指导名称的用词来源和内容的语义规则，如：由数据模型生成的词汇，规范中常用词汇等等；
- 关于用词顺序的语法规则；
- 关于受控词表、名称长度、字符集、语言词法规则；
- 制定在同一语境中名称是否必须具备唯一性的规则。

8.1.3 注册数据标识符

属性注册机构标识符（RAI）、数据标识符（DI）和版本标识符（VI）构成了注册数据标识符。一个数据元至少需要一个注册数据标识符。数据标识符由注册机构分配；数据标识符在一个注册机构的范围内必须是唯一的。

由于每个注册机构可以决定其各自的分配方案，因此不能保证用某注册机构的数据标识符能对一个数据元进行唯一标识。例如，如果两个注册机构都用连续的6位号码，就会有一组数据元有相同的DI，而且完全有可能，不同的数据元具有相同的DI。相反，如果同一个数据元在两处机构注册，它就有两个DI。因此，一个数据元的标识不仅需要DI还需要RAI。

如果数据元的特定属性改变了，就应产生并注册数据元的新版本，在这种情况下就需要VI来完成对数据元的唯一标识。

一个注册数据标识符是信息系统、组织或其它希望共享一个特定数据元（但不能利用相同的名称和语境）的参与者之间交换数据的关键。当注册数据标识符与建立多于一种自然语言的语境有关联时，注册数据标识符也有助于语言的翻译，并且对由不同的注册机构管理的数据元集合之间是一个参考。

8.2 数据注册标识的规则

8.2.2 数据注册标识的规则

- a) 在一个注册机构的注册系统中，每个数据元应有一个唯一的数据标识符。

b) 注册机构标识符、数据标识符和版本标识符的组合，便构成数据元的唯一标识。

9 数据元的命名规则

每个数据元是由一组成分构成，它们选自各自语境的结构设置。

数据元可以（一般）由其成分的名称构成，在一个名称中设置每个成分的含义（语义）和相对或绝对的位置（语法）。它们可以用分隔符来分界。每个成分值的集合或范围应当由一个机构严格管理。

- 语法规则使含义能够传达；
- 句法规则按一致性和指定顺序关联成分；
- 词法（词形和词汇）规则减少冗余，增加精确度。

9.1 名称语义内容的规则

语义涉及到名称成分的含义，以及对其界定的分隔符。

9.1.1 语义规则

a) 对象类表示在一个领域内有关的事物，例如：在某领域的一个模型中发现的事物。

示例：费用

b) 需要有一个且仅有一个对象类词。

c) 特性词应当从特性体系结构设置中产生，并表示出数据的类别。

d) 需要有一个且仅有一个特性词。

e) 限定应由专业领域机构决定产生，当需要描述一个数据元并使其在特定的语境中唯一时，即可增加上限定词。限定词的顺序并不重要。限定词是可选的。

示例：预算周期

f) 表示数据元的有效值集合由表示词来描述。

g) 需要有一个且仅有一个表示词。

9.1.2 名称成分语义

成分由离散词组成。本标准名称成分由以下词描述：

- 对象类词；
- 特性词；
- 表示词；
- 限定词。

9.1.2.1 对象类词

对象类词是构成数据元名称的一个成分，它表示某一语境中的一项行为或一个对象。

例如，在下面的数据元中：

——雇员的姓氏

——费用预算周期合计总额

——树的高度测量

——成员的姓氏

成分雇员、费用、树和成员是对象类词。

9.1.2.2 特性词

一组特性词是由一个特性分类中的一组名称成分构成的。这个组必须是由离散的（每个定义不能与其它定义相重叠）和完整的（全部的特性，这个组表示了所有说明数据元所需的信息概念）词构成。

例如，在数据元中：

——雇员的姓氏

——费用预算周期合计总额

——树的高度测量

——成员的姓氏

成分姓氏、总额和高度是特性词。

特性词当然要出现在数据元的定义中。使用两个结构设置中的成分，提供了一种分类的补充方法。用数据元对象类和特性两个成分形成的一个名称，包含了有关数据元极为重要的信息，而且排除了当没有约定使用而被采用的不重要的或不合理的元素。

9.1.2.3 表示词

表示词是一个数据元名称中描述数据元表示形成的一个成分。每个表示词由一个受控单词表或一个分类体系形成。表示词的类别表现形式如下：

- 名称
- 测量
- 数量
- 总额
- 数目
- 文本

这类词描述了数据元有效值集合的形式。通常这类表示词可能与特性词有部分重复，此时，可以从结构化名称中将一个词或词的一部分删除。

例如，在数据元中：

——树的高度测量

——雇员的姓氏

成分测量和姓氏是表示词。注意姓氏是一个特性词，为了表达的清楚，冗余的字可以删去。

9.1.2.4 限定词

如果必须对一个数据元进行唯一标识，可以将限定词加到对象类词、特性词和表示词上，这些限定词也许是从一个语境规定的结构设置中产生的。在（确立）命名约定的规则中（时），建议对限定词的数量予以限定。

例如，在数据元“费用预算周期合计总额”中，成分“预算周期”是限定词。

9.1.3 分隔符语义

词的成分由分隔符来界定。它们可以是：

—— 没有语义含义。可以用一项命名规则说明分隔符由一个空格或一个确切的特定字符（如一个连字符或下划线）组成，而不管各成分间的语义关系如何。这样的规则简化了名称的生成过程。

—— 有语义含义。语义含义能由分隔符表达，如将限定词之间的分隔符和其它成分之间的分隔符设定不同的分隔符。用这种方法，分隔符就把限定词从名称的其它部分清楚地标识出来了。

例如，在数据元中：

—— 费用-预算_周期-合计-总额

在限定词之间的分隔符是下划线；其它名称成分之间的分隔符是连字符。

9.2 名称格式的规则

9.2.1 句法规则

句法原则详细说明了一个名称中各成分的排列。这一排列可能是相对的或者绝对的，或者二者兼有。

a) 相对排列是依据其它成分确定该成分。如在一个约定中的一项规则可能要求一个限定词必须总是出现在被限定词的前面。

b) 绝对排列是确定该成分的固定位置。如，一项规则可能要求特性词总是一个名称的最后成分。

9.2.2 针对数据元英文名称的词法规则

a) 名词仅用单数形式，动词（若有的话）为现在时。

b) 名称的各个成分间和多个词之间用空格分隔。不允许用特殊字符。

c) 名称中所有单词是组合在一起的。

d) 允许使用缩写词、首字母缩略词和大写首字母。

9.2.3 针对数据元中文全拼的格式

数据元中文全拼由中文名称中的每一个汉字的拼音组成，拼音中间用连字符“-”连接，并全部使用小写。

9.2.4 句法规则

a) 对象类词应处于名称的第一（最左）位置。

b) 限定词应位于被限定成分的前面，限定名称的顺序不应用于区别数据元名称。

c) 特性词应处于第二位置。

d) 表示词应处于最后位置。假如表示词中有任何字与特性词中的字重复，则删去冗余词。

示例：费用 预算 周期 合计 总额

9.3 唯一性规则

同一语境的所有名称应是唯一的。

10 数据元的定义

10.1 数据元定义规则

数据元的定义应：

a) 具有唯一性（在出现此定义的任何数据字典或数据元目录中）。

说明：每个定义必须区别于（字典或目录中的）任何其它定义，以保证专一性。定义中所表述的一个或多个特性必须使被定义的概念与其它概念相区别。

b) 要阐述其概念是什么，而不是阐述其概念不是什么。

说明：编写定义时，仅阐述其概念不是什么并不能对概念作出唯一的定义。

c) 用描述性的短语或句子阐述。

说明：必须使用短语来形成包含概念的基本特性的准确定义。不能简单地陈述一个或几个同义词，也不能以不同的顺序简单地重复这些名称。如果一个描述性短语不够，则应使用完整的、语法正确的句子。

d) 仅可使用人们普遍理解的缩略语。

说明：对缩略语，包括简称和英文首字母含义的理解，通常受到特定环境的限制。环境不同，同一缩略语也许会引起误解或混淆。因此，为了避免词义不清，在该定义中就应使用全称，而不用缩略语。

所有简称在第一次出现时，必须予以说明。

e) 表述中不要加入不同的数据元定义或引用下层概念。

说明：

10.2 数据元定义指南

数据元的定义宜：

a) 阐述概念的基本含义；

b) 准确而不含糊；

c) 简练；

d) 能单独成立；

e) 表述中不应加入理论说明、功能说明、范围信息或程序信息；

f) 避免相互依存；

g) 相关定义使用相同的词和一致的逻辑结构。

11 数据元值域及其表达模式

数据建模者通常称属性的表示形式为它的“值域”，或简称为“域”。在此，我们称数据元的表示部分为**表示**。

数据元在商务运作中呈现为值，并在信息交换等功能中得到共享。数据元通常有一个允许值的集合。这个允许值的集合被称之为**值域**。

数据元从不表示为一个单独的数值，因为它是一个类（如，数据值完整的集合）而不是一个单个事例。比如，雇员标识符是一个数据元，它的值域由一个特定企业中允许值的一个完整列表来描述。这里的数据值仅是雇员标识符所有实例的一个列表。数据元的一个实例只有一个单个数据值并被称之为一个“数据元实例”。

一个特性有定义并隶属一个对象类。相反，一个表示没有定义，但有一个格式类、允许值、最大字符数，如果可以度量，还应有一个度量单位。比如，数据元“月名称”就有以下表示：

- 格式类 = alpha
- 允许值 = January, February, March, April, May, June, July, August, September, October, November, December
- 最大字符数 = 9
- 度量单位 = （不适用，因为不存在度量的问题）

11.1 穷举域

尽管数据元的元数据的结构可以从对象类或特性的角度考察，也可从表示的角度考察，这最容易用一个穷举域的例子来表明。穷举域是一个可以由所有允许值列表指定的值域。比如，人们给国家标上名称，就很容易引证它们。包含现今世界上所有国家名称的集合就是穷举域的一个例子。

尽管特性和（表示的）有关的值域可以被考虑用来标识数据元一个潜在有效数据值的集合，对象类通常将宽泛的值域限定为数据值的一个具体的子集，以用于某个具体的数据元。

11.2 可重复利用域

一些域有被许多数据元使用的潜在可能。比如，所有人类眼睛颜色的名称可被用于雇员、承包商、夫妻和数据库中其它类的人员。

一个严格定义并被指定的域可以被再利用。这些可再利用的域在用于数据元时，具有相同的值域、表示和特性。比如，雇员眼睛颜色或承包商眼睛颜色。我们称这些可再利用域为通配数据元。

11.3 数据元值域的规范定义和来源

名称：值域

定义：根据相应属性中所规定的数据类型、数据格式而决定的数据元的允许值的集合。

约束：必选

出现次数：1：1

数据类型：字符串

备注：该集合可以通过以下四种方式给出：

- 通过名称给出，即直接指出值域的名称，比如数据元“两字母国家代码”的值域是 GB/T 2659 《世界各国和地区名称代码》中的全部两字母代码；
- 通过参考资料给出，比如数据元“产品条码”的值域是已经在物品编码中心注册的所有产品的条形码；
- 通过一一列举的方式给出所有可能的取值以及每一个值对应的实例或含义；
- 通过规则间接给出，比如数据元“无线电频率”的值域是从 3kHz 到 300GHz，并且遵循 IEC-50 的规范。
- 无要求

本标准建议科学数据中的数据元值表达格式基本采用 GB/T 19253 规定的的数据元值的数据类型与字符长度表示方法，可根据需要做适当的修改和补充。

12 数据元目录标准编写格式

在数据的应用、管理和交换中，数据元目录或字典的编写是很重要的手段和途径之一。

12.1 数据元目录标准格式

数据元目录标准一般由以下部分构成：

- 1 范围
- 2 规范性引用文件
- 3 术语和定义
- 4 数据元的描述格式
- 5 数据元的类目分组
- 6 数据元的维护与管理
- 7 数据元目录
- 8 数据元索引
 - 8.1 按数据元标记的顺序索引
 - 8.2 按数据元名称的顺序索引（汉语拼音字母顺序）

12.1.1 数据元描述格式

作为一个与国际接轨的、管理完善和使用方便的数据元标准，为能在更大范围内使用，在更高层次上共享，应尽量按上述要求对数据元的属性进行全面地描述。但由于目前数据元标准的制定还处在初级阶段，数据元的提交、注册和管理还不够完善，甚至还没有建立起相应的机构。鉴于这种客观现实，目前在制定数据元标准时，往往只选择了数据元几个最主要的属性进行规定。

12.1.2 数据元索引

为了便于数据元的检索，数据元标准宜建立数据元索引，一般分别以数据元标记（即内部标识）顺序和数据元中文名称汉语拼音字母顺序建立两个索引。两个索引中都只需列出“标记”和“中文名称”两个属性，只是在两个索引中这两个属性前后排列顺序不同而已。

12.2 数据元目录的一般格式

数据元目录是数据元标准的核心内容。一般来说：

- a) 为节省版面，将数据元的内部标识、中文名称和英文名称放在一行中。
- b) 内部标识码之间留有空位，以便扩充。但如针对同一数据元概念的名称和代码型数据元，应分配相连的内部标识符。
- c) 数据元和数据元之间用一条实线分割开来，使读者看起来方便、清晰。

编写数据元目录时，建议基本采用示例中的书写格式，但要增加值域的内容，如下所示：

内部标识符 中文名称 英文名称

说明：

语境：

表示：

值域：

同义名称：

关系：

备注：

示例：

1104 币别 Name of currency in Chinese

说明：货币的中文名称。

语境：通用

表示：C..20

值域：采用 GB/T 12406-1996《表示货币和资金的代码》中的货币名称。

同义名称：无

关系：无

备注：详见 Q/PICC 003《保险业务代码集》

1105 币别代码 Currency code

说明：货币名称的代码。

语境：通用

表示：a3

值域：采用 GB/T 12406-1996《表示货币和资金的代码》中的 3 字母货币代码。

同义名称：无

关系：无

备注：详见 Q/PICC 003《保险业务代码集》

附录 A
(规范性附录)
数据元的表示规范

A.1 数据元属性的描述方法

数据元属性应依照一种标准方式来描述。下面的描述符只对数据元属性的描述有效(对数据元的描述无效)。

a) 名称

赋予数据元属性的标记。名称是唯一的。名称以字符串形式表示。

b) 定义

属性的描述,可使一种属性与其他属性清晰地区别开来。定义以字符串形式表示。

c) 约束

显示一个属性是始终还是有时出现的描述符。该描述符可以有两个取值:必选或可选,前者表示该属性必须出现,后者表示该属性可以出现,也可以不出现。

d) 出现次数

显示一个属性出现多少次的描述符。该描述符有以下四种情况:0:1(表示不出现或出现1次),0:n(表示不出现或出现n次),1:1(表示出现且仅出现1次),1:n(表示出现1次或多次)。

e) 数据类型

描述属性的所有取值的类型。属性值的数据类型示例有:“字符”、“字符串”、“数字”、“日期”。

f) 备注

与属性应用有关的注释。

A.2 标识类属性

A.2.1 名称:中文名称

定义:赋予数据元的单个或多个中文字词的指称。

约束:必选

出现次数:1:1

数据类型:字符串

备注:参见中文名称的命名规则。

A.2.2 名称:英文名称

定义:赋予数据元的单个或多个英文字词的指称。

约束:必选

出现次数:1:1

数据类型:字符串

备注:参见英文名称的命名规则。

A.2.3 名称:中文全拼

定义:数据元中文名称的汉语拼音。

约束:必选

出现次数:1:1

数据类型:字符串

备注:参见中文全拼的命名规则。

A.2.4 名称:标识符

定义:在一个注册机构内,由注册机构分配的、与语言无关的数据元的唯一标识符。

约束：必选

出现次数：1：1

数据类型：数字

备注：参见标识符的分配规则。

A.2.5 名称：版本

定义：在一个注册机构内的一系列逐渐完善的数据元规范中，某个数据元规范发布的标识。

约束：必选

出现次数：1：1

数据类型：字符串

备注：参见版本的格式规则。

A.2.6 名称：注册机构

定义：经授权对数据元实施注册、维护和管理功能的组织。

约束：必选

出现次数：1：1

数据类型：字符串

备注：属性“注册机构”、“标识符”和“版本”的组合可用来唯一地确定一个数据元。

A.2.7 名称：同义名称

定义：一个数据元在应用环境下的不同称谓。

约束：可选

出现次数：0：n

数据类型：字符串

备注：一个数据元可以有多个同义名称，第一个同义名称应该是统一的约定名称。

A.2.8 名称：语境

定义：产生或使用数据元的应用环境或应用规程的说明。一个语境可以是一个业务领域、信息系统、数据库、文件或数据模型等，也可以是他们的组合。

约束：必选

出现次数：1：n

数据类型：字符串

备注：必须为数据元指定属性“语境”。

A.3 定义类属性

A.3.1 名称：定义

定义：表达一个数据元的本质特性并使其区别于所有其他数据元的陈述。

约束：必选

出现次数：1：1

数据类型：字符串

备注：参见有关正文中的数据元定义规则。

A.3.2 名称：对象词

定义：数据元名称成分之一，表达了该数据元所属事物或概念的集合。

约束：必选

出现次数：1：1

数据类型：字符串

备注：标识出数据元的对象词（词）有助于实现对数据元的规范化命名、分析、类比和查询等功

能。

A.3.3 名称：特性词

定义：数据元名称成分之一，表达了该数据元所属的对象类的某个显著的、有区别的特征。

约束：必选

出现次数：1：1

数据类型：字符串

备注：标识出数据元的特性词有助于实现对数据元的规范化命名、分析、类比和查询等功能。

A.3.4 名称：应用约束

定义：数据元在实际应用中的相关约束。

约束：可选

出现次数：0：1

数据类型：字符串

备注：该属性侧重于描述从应用中提取出的约束需求，比如数据元的保密问题、用法、规则等等。

A.4 关系类属性

A.4.1 名称：分类方案

定义：根据对象的共性如：来源、构成、结构、应用、功能等将其排列或分组。

约束：可选

出现次数：0：n

数据类型：字符串

备注：采用分类方案的标识符进行标识。

A.4.2 名称：分类方案值

定义：一个分类方案中的某一分类的代码。

约束：可选

出现次数：0：n

数据类型：字符串

备注：分类方案值与分类方案有一一对应关系。

A.4.3 名称：关系

定义：当前数据元与其他相关的数据元之间关系的一种描述。

约束：可选

出现次数：0：n

数据类型：字符串

备注：表 A-1 给出了数据元的几种基本关系的表示格式。

表 A-1 数据元基本关系的表示格式表

关系中文名称	关系表示符	关系描述
派生关系	derive-from	描述了数据元之间的继承关系，一个较为专用的数据元是由一个较为通用的数据元加上某些限定词派生而来，例如“derive-from B”（B 是数据元的标识符，下同），表明当前数据元由数据元 B 派生而来。
组成关系	compose-of	描述了整体和部分的的关系，一个数据元是另外若干个数据元组成，例如：“compose-of B、C、D”，表示当前数据元是由数据元 B、C、D 共同组成。

替代关系	replace-of	描述了数据元之间的替代关系，例如：“replace-of B”表明当前数据元替代了数据元 B。
连用关系	link-with	描述了一个数据元与另外若干数据元一起使用的情况，例如：“link-with B、C、D”，表明当前数据元需要和数据元 B、C、D 一起使用。

A.5 表示类属性

A.5.1 名称：表示词

定义：数据元名称成分之一，表达了该数据元值域的表达形式。

约束：必选

出现次数：1：1

数据类型：字符串

备注：标识出数据元的表示词有助于实现对数据元的规范化命名、分析、类比和查询等功能。表 A-2 给出了国际范围内认可的表示词。

表 A-2 国际范围内认可的表示词的列表

表示词	含义
金额	货币单位的数量，通常与货币类型有关。
日期	特定的年月日，格式参照 GB/T 7408。
日期时间	特定的年月日中的特定时间点，格式参照 GB/T 7408。
代码	表示一组值中的一个值的字符串（字母、数字、符号）。
描述	表示一个人、客体、地点、事件或概念的一系列句子，既可用于定义（通常用一两个句子），也可用于较长文本。在数据元的中文名称中通常使用“说明”、“备注”、“意见”等词。
名称	为一个人、客体、地点、事件或概念指定的一个词或短语。该词或短语是该人、客体、地点、事件或概念的称谓。
号码	一个特定的值的数字表示。它通常暗示了顺序或一系列中的一个。
百分比	具有相同计量单位的两个值之间的百分数形式的比率。
量	非货币单位数量，通常与计量单位有关。
比率	一个计量的量或金额与另一个计量的量或金额的比。
指示符	两个且只有两个表明条件的值，如 On/Off、True/False，又称标志。

A.5.2 名称：数据类型

定义：用于表示数据元的符号、字符或其他表示的类型。

约束：必选

出现次数：1：1

数据类型：字符串

备注：表 A-3 包括了数据类型可能的取值列表，但不限于下表中所列。

表 A-3 数据类型可能的取值列表

数据类型	说明
字符型 (string)	通过字符形式表达的值的类型。
数字型 (number)	通过从“0”到“9”数字形式表达的值的类型。
日期型 (date)	通过 YYYYMMDD 的形式表达的值的类型，符合 GB/T 7408
日期时间型 (datetime)	通过 YYYYMMDDhhmmss 的形式表达的值的类型，符合 GB/T 7408

布尔型 (boolean)	两个且只有两个表明条件的值, 如 On/Off、True/False。
二进制 (binary)	上述无法表示的其他数据类型, 比如图像、音频等

A.5.3 名称: 数据格式

定义: 从业务的角度规定的数据元值的格式需求, 包括所允许的最大和/或最小字符长度, 数据元值的表示格式等。

约束: 必选

出现次数: 1: 1

数据类型: 字符串

备注: 1、数据格式中使用的字符含义如下:

a =字母字符

n =数字字符

an =字母数字字符

m(m=为自然数) =定长 m 个字符 (字符集默认为 GB 2312)

..ul =长度不确定的文本

..p,q (p,q 均为自然数) =最长 p 个数字字符, 小数点后 q 位

.. =从最小长度到最大长度, 前面附加最小长度, 后面附加最大长度

YYYYMMDDhhmmss=“YYYY”表示年份,“MM”表示月份,“DD”表示日期,“hh”表示小时,“mm”表示分钟,“ss”表示秒可以视实际情况组合使用。

例 1: an5(aannn) 表示定长 5 个字母数字字符, 前 2 个为字母字符, 后三个为数字字符;

例 2: n..17, 2 表示最长 17 个数字字符, 小数点后两位。

例 3: an3..8 表示最大长度为 8, 最小长度为 3 的不定长的字母数字字符。

2、如果“数据类型”是“二进制”, 在本属性中应标识出二进制的具体格式, 如“JPEG”。格式参照 RFC 2046[1]。

A.5.4 名称: 值域

定义: 根据相应属性中所规定的数据类型、数据格式而决定的数据元的允许值的集合。

约束: 必选

出现次数: 1: 1

数据类型: 字符串

备注: 该集合可以通过以下四种方式给出:

——通过名称给出, 即直接指出值域的名称, 比如数据元“两字母国家代码”的值域是 GB/T 2659《世界各国和地区名称代码》中的全部两字母代码;

——通过参考资料给出, 比如数据元“产品条码”的值域是已经在物品编码中心注册的所有产品的条形码;

——通过一一列举的方式给出所有可能的取值以及每一个值对应的实例或含义;

——通过规则间接给出, 比如数据元“无线电频率”的值域是从 3kHz 到 300GHz, 并且遵循 IEC-50 的规范。

——无要求

A.5.5 名称: 计量单位

定义: 属于数值型的数据元值的计量单位。

约束: 可选

出现次数: 0: 1

数据类型：字符串

备注：详见 GB/T 17295-1998《国际贸易用计量单位代码》中的计量单位的名称。

A.6 管理类属性

A.6.1 名称：状态

定义：数据元在其注册的全生存期内所处状态的标示。

约束：必选

出现次数：1：1

数据类型：字符

备注：数据元在其注册的全生存期内包括四种阶段：

——草案：该数据元的内容处在草案阶段，相关单位和部门可以广泛提出意见和建议；

——试用：该数据元的内容经过一定的技术程序，可以在一定范围内进行试用，并反馈实验意见；

——标准：该数据元的所有内容已经成为各方应当遵守的标准；

——废止：该数据元的内容即将从标准中删去。

A.6.2 名称：提交机构

定义：提出对数据元进行增加、变更（导致版本变更）或废止的组织或组织内的部门。

约束：可选

出现次数：0：1

数据类型：字符串

A.6.3 名称：批准日期

定义：数据元进入“标准”阶段的日期。

约束：可选

出现次数：0：1

数据类型：字符串

备注：当数据元进入“标准”阶段时，该属性必选。采用 YYYYMMDD 的格式。

A.7 附加类属性

A.7.1 名称：备注

定义：数据元的附加注释。

约束：可选

出现次数：0：1

数据类型：字符串

备注：在上述属性未能描述的其他注释。

附录 B
(规范性附录)
数据元的维护与管理

数据元的维护包括数据元生命周期管理、数据元的添加、删除和更新，这些工作应由数据元注册机构负责。数据元的生命周期包括四个阶段，即草案阶段、试用阶段、标准阶段和废止阶段。

B.1 数据元的标识

数据元的维护与管理首先是数据元的标识。依据前面的规定，注册数据元的唯一性由三个标识属性值的组合来确定：注册机构标识符（RAI）、数据标识符（DI）和版本标识符（VI）。

B.1.1 数据标识符（DI）的赋值

每一个列入注册系统的新数据元，应赋予一个新的数据标识符（DI）。如果现有数据元被修改，其内容被改变，或其值变换了表示形式，这时，应该赋予现有数据元一个新的 DI。例如，改写了必需属性的定义，或修改了表示形式就要赋予新的 DI。但是，对某个定义做编辑性的修改，不构成新数据元的理由，只要某个定义维持着实质性内容的描述就行。基于包括注册系统在内的诸多因素的需要，每个注册机构（RA）应适当制定和发布一批详细指南，对要求赋予新 DI 做出规定，即对新构成数据元的附加条件加以导向，因为在注册系统里已确定的属性值改变了。

B.1.2 版本标识符（VI）的赋值

通常，当任何属性值发生变化（相对要求有一个新 DI 而言），就要出现一个新的版本标识符（VI）。但是，每个数据元可以要求不同版本的处理，例如，对于雇员姓名，当变更雇员 DE 允许值时，不需要新版本，如改变帐户类型 DE 允许值时，那就需要一个修改的版本了。每个注册机构都应制定一批专项指南，对诸多情况进行专门研究、细化，并对其负全责。

B.2 数据元的注册

B.2.1 新数据元的注册程序

向注册机构（RA）申请数据元注册，应由主办机构（SO）提出。之前，应尽量同有关主管机构（RO）磋商。

注册机构（RA）应决定是否接受申请。如果收下了报告，注册机构（RA）就在注册状态注上“已注册”的字样。

如果收下了申请，注册机构（RA）应赋予注册数据标识符值。RA 应将该注册数据标识符值连同申请书提供的信息，登入数据元注册系统。这时，主办机构（SO）与其主管机构（RO）将收到一份项目的复印件，要求对该项目进行检查，如果查出有错，应即向注册机构（RA）提出建议。经 RA、RO 和 SO 三方核实无误，收到报告的注册状态应当标上“注册注册”或“标准化”的字样。

如果申请未被接收，就退回主办机构（SO），并附上一份拒绝受理的详细说明。如有可能，附上一份建设性的建议书，提出采用何种方式，申请就能被接受。至于拒绝受理的原因。

B.2.2 修改数据元的注册程序

要求修改注册系统的数据元，通常由原先的主办机构（SO）告之注册机构（RA）。如果要求修改的申请来自其他渠道，那么，注册机构（RA）应将其申请转给原先的 SO 和有关的 RO 去确认，以便

采取进一步的措施。

注册机构（RA）要修改注册系统的数据元，最好是给 SO 和 RO 发一份项目的复印件，让他们对登录的内容进行检查，如果查出有错，立即给 RA 提出改正建议。

如果修改的要求不被接收，最好将修改申请退回 SO 和 RO，并附上拒绝受理的理由之详细说明。如有可能，提出一份建设性的建议书，提出怎样才能使申请被接受。至于拒绝受理的原因。

除上述外，供修改用的注册程序，同修改新数据元注册程序一样。

B.2.3 注册状态的说明

这种属性（注册状态）提出“数据元在进行注册期间所处各种状态的总称”。

下列状态由注册机构（RA）认定并注册，对每个数据元处于注册过程的什么阶段，按照 B.2.1 和 B.2.2 规定的程序执行。

已登录：已注册的数据元，包括所有必需的属性值，但其内容可以同本标准其他各个部分规定的的质量要求不一样。

已审核：已登录的数据元，应同本标准其他各个部分规定的的质量要求一致。

已标准化：已审核的数据元，由注册机构（RA）确认，优先用于数据交换和新或现时的应用领域。“标准化”的数据元在注册系统中可能是唯一的，或在相似的数据元中是首选的。

已失效：已登录、已审核或已标准化的数据元，经注册机构（RA）判定为“已停止使用”后，不再使用。

注册机构（RA）应制定更详细的要求准则，在自己管理的领域内，规定每个阶段应达到的目标。

B.2.4 管理状态的说明

这是注册机构（RA）在处理注册请求期间，所处各个状态的总称。由注册机构（RA）用允许值赋予这个属性的值，证明正处在审定、草案、拒绝、提交等状态，以此推进对注册的管理。注册机构的职责，在于对这种管理提出办法，精心安排，及时公布。

B.3 注册系统

每个注册机构（RA）都应保有数据元注册系统，列入其主管领域的数据元。

B.3.1 内容

对每个数据元注册而言，数据元注册系统至少应包括标准的规定的的数据元的基本属性和管理属性。

B.3.2 语言

注册系统使用的语言，应由有关注册机构（RA）规定的程序予以确认。

B.3.3 数据元注册簿的有效性

注册系统内容的有效性，须依照有关注册机构（RA）规定的程序进行管理。

B.3.4 数据元注册系统的版权

鉴于注册机构（RA）是自行管理的机构，在本标准规定的程序内自行操作，因此，其版权应属于注册机构（RA）自己。

B.4 注册参加机构

本标准的框架里，列有注册执行机构（RAB）的三种类型：主办机构（SO）、主管机构（RO）和注册机构（RA）。每种类型的 RAB 都有各自的准则、作用和职责范围。本标准在下列条款做了规定。

B.4.1 提交机构（SO）的资格、作用和职责

B.4.1.1 提交机构（SO）的资格

根据本标准规定，凡是希望进行数据元注册的提交机构，必需执行本标准规定的程序。每个注册机构（RA），应制定自己的标准准则，来审查成员单位的资格。

B.4.1.2 提交机构（SO）的作用

希望进行数据元注册的主办机构（SO）应遵照本标准的规程和要求，归属相关的注册机构（RA）。

B.4.1.3 提交机构（SO）的职责

提交机构（SO）负责下列业务：

- 提供数据元基本属性、管理属性和附加属性等信息，其格式按注册机构（RA）的要求。
- 提供注册机构（RA）要求的补充信息，以满足其履行职责的要求。
- 保证已注册的数据元，在未提交给注册机构（RA）时，数据元属性值的规范不会改变。

B.4.2 主管机构（RO）的设定、作用和职责

B.4.2.1 主管机构（RO）的设定

通常，主管机构（RO）由某业务实体指定，该实体与提交机构协调其管理的数据元的必选属性内容。如果没有主管机构（RO）的话，提交机构（SO）即为主管机构（RO）。

B.4.2.2 主管机构（RO）的作用

主管机构（RO）或下属部门对数据元属性值的完整性和准确性负责。例如：由注册机构（RA）维护和管理的数据元的语义和允许值。

主管机构（RO）按注册机构（RA）的要求，受理提交机构（SO）对有关属性的修改提案，例如：名称、定义以及数据元允许值。主管机构（RO）需将对任何已注册数据元的说明所进行的实质性修改，通知注册机构（RA）。

B.4.2.3 主管机构（RO）的职责

主管机构（RO）必须：

- 按注册机构（RA）的要求，对提交注册的数据元，就其语义、名称和允许值提出建议。
- 对赋予主管机构（RO）的数据元所作的任何修改，要通知注册机构。
- 当赋予数据元的属性值发生混淆、争议时，作出裁决。

B.4.3 注册机构（RA）的设立、作用和职责

B.4.3.1 注册机构（RA）的设立

一个组织为了将自身建成注册机构（RA），需要完成如下工作：

- 申请获得一个注册机构标识符（RAI）、

——按照本标准对数据元注册需遵循的程序，提出规定、建议和解释。

——确定注册范围内专门说明的补充条款。

——规定数据元基本、管理属性中的每个属性和注册机构（RA）认为需要的附加属性的格式，以及提交注册数据元所使用的媒体。所规定的注册形式和相应的程序，应符合对提交机构（SO）的要求。

——确定提交申请应使用形式和媒体，注册机构（RA）还需向提交机构（SO）提供申办指南，

——制定并颁布一批规章，使注册生效。注册机构（RA）规定许可用户，访问内容，可用次数，以及提供信息的语言、媒体和格式。

B.4.3.2 注册机构（RA）的作用

注册机构（RA）应受理提交机构（SO）对其所在领域有关数据元注册方面的提案。注册机构（RA）负责维护数据元注册系统，并分配注册数据标识符。

B.4.3.3 注册机构（RA）的职责

有关进行数据元注册的申请，注册机构（RA）应该履行下列职责：

——受理有关数据元的注册申请，分配注册数据标识符值，并依据以下条款，维护数据元注册系统。

——如有些要求涉及到正在注册的数据元必选属性值时，同有关主管机构磋商。

——按照商业惯例，全面进行注册，尤其要采取一些有效措施，保护好注册系统。特别要注意注册机构（RA）的以下职责：

——受理提交机构（SO）关于数据元注册的申请。

——审查并推动申请在注册期间的进展。

——设定适当的注册状态。

——根据规定的程序，将本部门的决定通知提交机构（SO）。