

引用格式: 邓辉, 王毅, 张辉, 等. 标识符在企业主数据结构化中的应用[J]. 标准科学, 2025(6): 80-85.  
DENG Hui, WANG Yi, ZHANG Hui, et al. Application of Identifier in the Structuring of Enterprise Master Data [J]. Standard Science, 2025(6): 80-85.

## 标识符在企业主数据结构化中的应用

邓辉<sup>1</sup> 王毅<sup>2</sup> 张辉<sup>1\*</sup> 郎大字<sup>1</sup>

(1.北京飞航捷迅科技有限公司; 2.中国物品编码中心)

**摘要:** 【目的】统一主数据结构是企业数字化建设的基础和难点。【方法】研究了标识符在数据标识中的功能作用, 分析了数据元、元数据、主数据与数据结构化的关系, 提出给企业主数据构成的每个数据元赋予唯一的标识符, 建立数据元字典, 分类分级构建元数据集和结构化的主数据模型, 并给每个主数据模型赋予唯一的标识符, 以保证不同信息系统的数据结构一致。【结果】按数据元字典和主数据模型, 开发主数据编码系统, 统一主数据编码, 使跨系统、跨业务、跨部门、跨网络拥有单一的数据来源, 实现企业主数据的无障碍协同共享。【结论】为企业主数据结构化和数据编码建设提供了理论参考。

**关键词:** 标识符; 主数据; 数据元; 元数据; 数据模型; 数据结构; 编码

DOI编码: 10.3969/j.issn.1674-5698.2025.06.010

## Application of Identifier in the Structuring of Enterprise Master Data

DENG Hui<sup>1</sup> WANG Yi<sup>2</sup> ZHANG Hui<sup>1\*</sup> LANG Dayu<sup>1</sup>

(1.Beijing Feihang Jiexun Technology Co., Ltd.; 2. Article Numbering Center of China)

**Abstract:** [Objective] The unified master data structure is the foundation and difficulty of enterprise digital construction. [Methods] This paper studies the functional role of identifiers in data identification, analyzes the relationship between data elements, metadata, master data, and data structuring, proposes assigning a unique identifier to each data element composed of enterprise master data, establishes a data element dictionary, classifies and hierarchically constructs metadata sets and structured master data models, and assigns a unique identifier to each master data model to ensure consistency in data structures across different information systems. [Results] It develops a master data encoding system based on the data element dictionary and master data model, unifies the master data encoding, and provides a single data source across systems, businesses, departments, and networks, achieving seamless collaborative sharing of enterprise master data. [Conclusion] The paper provides theoretical reference for the construction of data structuring and data coding for enterprise owners.

**Keywords:** identifier; master data; data element; meta data; data model; data structure; code

**基金项目:** 本文受2022年国家科技支撑计划项目“工业化建造自动识别与数据采集技术国际标准研制”(项目编号: 2022YFF0610500)资助。

**作者简介:** 邓辉, 本科, 高级工程师, 研究方向为供应链管理、信息化、标准化。

王毅, 硕士, 研究员, 研究方向为物品编码、自动识别技术。

张辉, 通信作者, 硕士, 研究员, 研究方向为供应链管理、标准化。

郎大字, 硕士, 工程师, 研究方向为信息化。

## 0 引言

在企业数字化转型中,跨系统、跨网络、跨业务、跨层级的数据协同是大势所趋。但是企业信息系统及基础数据库建设的年代不同、开发单位不同、主管单位不同,不同系统、不同业务所执行的数据结构及标识规则各异,形成了一个数据“孤岛”。这导致同类数据不能自动查询、采集、统计、分析和互操作,大大削弱了数据的价值,严重影响数字化转型进程。本文尝试用唯一标识符规范企业数据元及元数据和主数据,达到数据的定义统一、结构统一、格式统一、标识统一、标准统一、源头统一,实现企业数据协同共享的目标。

## 1 标识符在数据标识中的功能作用

### 1.1 标识符的定义与功能作用

在信息系统建设中,面对庞杂的数据和数据关系,为保证源头数据的唯一性、关联性,提高数据的检索、追溯和自动化处理效率,需要给重复使用的数据确定名称或标识符以避免重复和混淆。

在不同的领域,标识符的概念有所不同。在软件编程中,标识符用于命名变量、函数、类、模块或其他用户自定义元素。由于早期编程语言的兼容性问题,长期以来标识符默认使用英文字符和数字。在电子数据交换中,标识符是用于标识或命名一个数据项并可能指出该数据项的确定性质的一个字符或一组字符<sup>[1]</sup>。在供应链物流领域,使用“应用标识符(AI)”来标识贸易项目、产品、资产、物流单元等数据含义与格式,AI由国际物品编码协会(GS1)负责维护管理。在工业制造领域,使用“数据标识符(DI)”来标识制造过程相关的包装、运输、度量、位置等数据的含义与格式,DI由美国物料储运学会(MHI)负责维护。GB/T 38606—2020给出了数据标识符代码表和部分应用标识符与数据标识符的映射关系表<sup>[2]</sup>。在信息与文献领域,使用“数字对象唯一标识符(DOI)”来定位检索学术资源,DOI注册服务机构需经ISO

26324批准。国际标准化组织(ISO)还建立了一种信息对象注册的分层结构,由对象标识符(OID)唯一地标识信息对象,全球OID根注册系统由法国电信公司维护,各国OID注册中心按授权负责相关节点及其分支节点的注册、管理和维护<sup>[3-4]</sup>。

我国已经广泛应用了一些标识符,如居民身份证号码、组织机构代码、统一社会信用代码、商品条码、中国标准书号等,在社会、政治、经济、生活中发挥着重要作用。

### 1.2 不同领域的标识符标准与可替代性

由于标识符在大数据应用中的重要作用,我国加速与国际标识符标准的接轨,并研制了一些国内标识符标准,例如:流通领域的GB/T 16986—2018《商品条码 应用标识符》、GB/T 28532—2012《承运人标识符 编码规则》、GB/T 26319—2010《国际货运代理单证标识符 编码规则》、GB/T 36596—2018《国际贸易商业发票标识符编制规则》等,文献领域的GB/T 12906—2008《中国标准书号条码》、GB/T 42713—2023《信息与文献 参与者名称标识符》、GB/T 36369—2018《信息与文献 数字对象唯一标识符系统》、GB/T 41132—2021《科普信息资源唯一标识符》、GB/T 40989—2021《新闻出版 知识服务 知识对象标识符(KOI)》、GB/T 23730.1—2009《中国标准视听作品号 第1部分:视听作品标识符》等,以及物联网领域的对象标识符GB/T 41810—2022《物联网标识体系 对象标识符编码与存储要求》、GB/T 38606—2020《物联网标识体系 数据内容标识符》、GB/T 42587—2023《信息技术 自动识别与数据采集技术 数据载体标识符》等。这些与标识符相关的标准为企业数据结构化、标准化提供了指引。

若国家、行业标准已经规定了数据元、元数据的标识符,企业直接采用更经济,还能与外部采用相同标准的系统接口保持一致,方便信息协同共享。但是,不同部门、不同行业制定的数据元、元数据标准大多是从本行业管理习惯出发的,导致标识符的设置各具特色,不同标准之间的数据定义、结构、格式、标识符等数据属性有

差异。企业原封不动照搬不同行业的标识符,大概率不能实现数据交互,因为不同的标识方法在不同的标识系统中是不能相互替代的。例如:10.978.12345/99990是一个DOI名称,978-12345-99990是一个中国标准书号(ISBN),这2个标识符标识了同一指示物;但是10.978.12345/99990不能用ISBN的销售订单系统进行验证,也不能转换为一个GS1条码作为ISBN条码使用,因为它不符合ISBN的语法规则;978-12345-99990同样也不能交给DOI的解析系统进行验证,因为它不符合DOI的语法规则<sup>[5]</sup>。

### 1.3 企业内部标识符的功能作用

为确保标识符的科学性、准确性,以及在企业范围内的唯一性,企业应结合自身业务、所处行业、与外部数据交互的需求,继承已有的数据命名规则和标识符,制定企业内部标识符的编码规则。企业内部标识符只限在企业范围内使用,也就是在企业数据交互范围内,每个标识符只表示一个特定含义,以避免标识符的重复和二义性。

软件开发人员习惯用自然语言、缩写词自主命名和定义标识符,但由于自然语言和缩写词具有多义性,不同人的理解可能会有差异,如CAD在不同的语境中有计算机辅助设计、加拿大元、冠状动脉疾病等完全不同的含义。另外,同一事物可用不同的词语表达,如Manufacturer、Factory、OEM,以及汉语的生产厂、生产厂家、生产单位、制造单位、生产制造单位等词汇实质上表达的是同一个事物。为避免人们对同一单词或短语的不同理解,除了对概念进行定义之外,“数字”“数字+字符”作为标识符应用越来越广泛。

## 2 企业数据结构化与数据元、元数据、主数据的关联关系

数据元、元数据、主数据是信息领域的专用术语,由英文翻译而来。企业产生的大量结构化数据是靠主数据规范的。主数据是用元数据描述的,元数据是由数据元组成的。数据元、元数据、

主数据对于数据结构化有不同的功能作用。

### 2.1 数据元

数据元(Data Element)是用一组属性描述定义、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,也有翻译成“数据元素”的。数据元一般由对象类、特性和表示3部分组成。对象类是人们希望采集和存储相关数据的事物,如人、汽车等。特性是用来区别和描述对象,是对象类的所有成员共有的特征,如性别、颜色等。表示是描述数据被表达的形式,由值域、数据类型、计量单位、表示类组成。值域是数据元允许值的集合<sup>[6]</sup>。以“姓名”为例,数据元的表示如图1所示。

中文名称: 姓名
英文名称: name
标识符: R00101
定义: 人的姓氏名称
数据类型: 字符型
数据格式: a...50 (不定长, 最长50个字节)
值域:
约束/条件: 必备
注释: 中国公民填写在公安户籍管理部门正式登记注册的、在人事档案中记载的、正在使用的中文姓名, 汉字之间不空格; 外国人按护照等法定有效证件中的姓名填写。

图1 “姓名”的数据元表示

数据元一般包括“中文名称”“标识符”“定义”“数据类型”“数据格式”“值域”“约束/条件”等。图1还给出了姓名的“英文名称”,便于中英文准确对照;用“注释”规范了姓名的填写要求。使用信息系统管理数据元的,还可包含数据生成、注册、更改、时间等管理属性,保持数据元状态的可追溯性。企业统一数据元表示时,应设置相同的数据类型和格式,这样才能实现信息系统集成和数据协同共享。

### 2.2 元数据

元数据(Metadata)是定义和描述其他数据的数据,特别是描述基础数据的结构和特性,提供数据准确性和一致性的要求。我国发布了一系列元数据标准,企业宜直接引用,也可制定企业元数据

标准,统一主数据需要的各项数据元名称、结构和格式。以照明产品为例,其元数据包括产品基础信息、产品发布信息、产品价格信息、产品分类信息、产品生产信息、技术参考信息等。每一类信息相关的特性都很繁杂,一般选定并规范一些必要的、在较大范围内共享的、变动频率不高的特性作为该类信息的核心元数据集,以保证该类信息在一定范围内拥有准确的、单一的理解。表1摘录了几项照明产品的基础信息元数据<sup>[7]</sup>。

### 2.3 主数据

主数据 (Master Data) 是企业需要跨系统、跨

部门进行共享的核心业务实体数据<sup>[8]</sup>。主数据跨越企业各部门,存在于人、财、物、产、供、销等多个应用系统中,具有共享性、唯一性、原始性、稳定性等基本特征,能够被重复利用,需要各系统保持一致。不同领域的企业主数据范围不尽相同,以大型生产制造企业为例,典型的主数据包括组织机构(客户、供应商、部门)、人员、财务、质量、经营、物品、合同、设计、工艺、生产制造等方面的能被共享、重复利用的数据,如图2所示。

企业主数据包罗万象,依据数据范围,对数据进行分类,制定每一类数据的结构化规则、编

表1 照明产品核心元数据字典描述 (摘自GB/T 37600.3—2018表1)

标识符	中文名称	英文名称	标记	定义	数据类型	值域	约束/条件	备注
003.1.1	产品名称	name	name	明确产品的中文名称	文本型	自由文本	M	
003.1.3	产品品牌	product brand	productBrand	产品的品牌	文本型	自由文本	O	
003.1.4	产品型号	product model number	productModel Number	产品的型号	文本型	自由文本	M	
003.1.5	净重	net weight	netWeight	明确产品的重量	数值型	正实数	M	单位: kg
.....								

说明: 约束/条件中M表示必选, O表示可选。

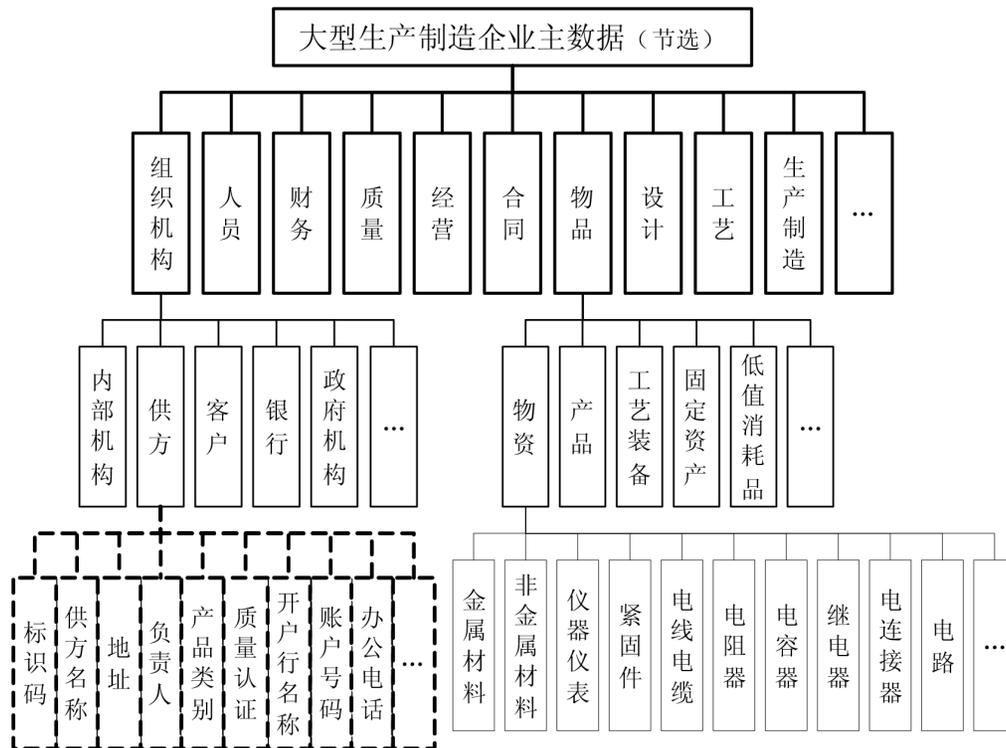


图2 大型生产制造企业的主数据结构 (节选)

码标准,再根据编码标准构建数据模型,开发主数据平台,借助主数据平台形成标准化的、源头的数据库。各个信息系统使用这个源头数据库就能实现企业内部数据的协同共享。

### 3 唯一标识符在企业主数据结构化中的应用

#### 3.1 统一企业数据元标识符

我国发布了一系列公共性公益性数据元、元数据和数据分类与代码标准,如性别(GB/T 2261.1)、民族(GB/T 3304)、学历(GB/T 4658)、学位(GB/T 6864)、专业技术职务(GB/T 8561)、货币(GB/T 12406)、计量单位(GB/T 17295)、国家和地区(GB/T 2659.1)、行政区划(GB/T 2260)等。这些标准规定了一些重要术语、概念的名称、含义、数据类型、数据长度、值域等特性,企业应直接引用,减少理解差异,提高数据的规范性和权威性。

企业根据本单位数据管理的范围,引用或结合相关标准,梳理出所需的数据元,规范每个数据元的名称、定义、数据类型、数据格式或字节长度、值域等特性,结合主数据分类,为每个数据元赋予一个唯一的标识符,建立企业数据元字典,保证企业信息系统的数据库的源头准确、唯一,实现企业范围内同一数据元的无障碍交互。

不同企业对数据颗粒度的要求不尽相同,必要时应对数据的内涵及外延进行解释,还可根据需要细分数据元,如“面积”使用“顷”“亩”“平方尺”“平方千米”“平方米”“平方英寸”“平方英尺”“平方码”等不同的计量单位会产生不同的数值。例如GB/T 16986详细定义了国际通用的物流计量单位的应用标识符,其中“面积”的应用标识符见表2。

#### 3.2 企业核心元数据建设

把某类主数据必需的、企业重点关注的、在企业内部能够实现统一的数据元作为该类主数据的核心元数据。采用统一建模语言(UML)描述数

表2 GS1物流计量单位的应用标识符(摘自GB/T 16986—2018/ISO/IEC 15418:2016)

应用标识符	含义(格式)	计量单位	单位符号	数据名称
334n	面积	平方米	m <sup>2</sup>	AREA
353n	面积	平方英寸	in <sup>2</sup>	AREA
354n	面积	平方英尺	ft <sup>2</sup>	AREA
355n	面积	平方码	yd <sup>2</sup>	AREA

据实体与数据元的关系,规范数据结构,分类分级构建数据模型,对主数据构成元素的名称、数据类型、数据长度、值域、权限,以及编码规则、校验规则、编码状态等特性进行定义和约束。核心元数据的数据元,应从企业数据元字典中选用,并保证数据元标识符的一致性。为便于管理,企业可对每一个数据模型进行编号,也就是建立数据模型的标识符。

#### 3.3 主数据标识码

主数据是由元数据、数据元集合组成的,如图2虚线部分的供方主数据记录由标识码、供方名称、地址、负责人、产品类别、质量认证、开户行名称、账户号码、办公电话等数据元集合构成。这些数据元能够满足企业不同数据库要求。统一规范的主数据结构是企业数据交换和资源共享的重要前提,统一主数据标识是减少数据二义性的有效方法。主数据标识是通过编码系统(或平台)实现的,编码是给事物或概念赋予代码的过程。编码的结果是代码,代码是一种标识码。近几年“编码”一词使用频率越来越高,人们经常有意无意地把“编码”当作名词使用,表示给事务赋予的“代码”本身<sup>[9]</sup>。由此看出编码、代码与标识符何其相似,编码、代码也是一种标识符。

依据数据模型,建立统一的主数据编码系统,对机构、人员、物品、供方、质量、合同等主数据实体编制具有唯一性的主数据标识码,坚持一数一源,实现一次录入全域共享,避免由于信息描述不一致而引起一码多物、一物多码及数据质量问题,使跨系统、跨部门拥有统一的数据结构、格式和

标识符,优化企业数据的流通与共享机制,推动数据管理和应用水平的提升,实现数据资源的最大化利用,进而提升工作效率和用户满意度<sup>[10]</sup>。

## 4 结语

通过梳理企业对数据资源需求,进行数据结构化分析,汇总提炼必要的数据元,结合国家、行业标准规范每个数据元的名称、定义、数据类型、

数据格式或长度、值域等特性,结合主数据分类,给每个数据元赋予唯一的标识符,建立企业数据元字典;把企业主数据必需的、重点关注的、在企业内部需要统一的数据元作为核心元数据,规范主数据结构和编码规则,制定企业主数据标识标准;分类分级构建主数据模型,开发主数据编码系统,统一主数据编码,使企业跨系统、跨业务、跨部门、跨网络拥有统一的、准确的、权威的数据来源,实现数据的无障碍协同共享。

### 参考文献

- [1] 行政、商业和运输业电子数据交换(EDIFACT)应用级语法规则 第1部分:公用的语法规则:GB/T 14805.1—2007/ISO 9735-1:2002[S].
- [2] 物联网标识体系 数据内容标识符:GB/T 38606—2020[S].
- [3] 信息技术 开放系统互连 对象标识符(OID)的国家编号体系和操作规程:GB/T 26231—2017[S].
- [4] 韩雪,周钢,马文静,等.OID标识与统一社会信用代码数据关联技术分析与研究[J].标准科学,2016(10):116-120.
- [5] 信息技术 元数据注册系统(MDR) 第1部分:框架:GB/T 18391.1—2009/ISO/IEC 11179-1:2004[S].
- [6] 信息与文献 数字对象唯一标识符系统:GB/T 36369—2018/ISO 26324:2012[S].
- [7] 邓辉,王毅,张辉.企业物品编码元数据标准化研究[J].中国标准化,2025(3):30-34.
- [8] 数据管理能力成熟度评估模型:GB/T 36073—2018[S].
- [9] 张辉,王毅,邓辉.物品编码与编目术语的辨析[J].中国标准化,2025(3):21-25.
- [10] 陈坚.基于主数据管理标准化的政府数据治理研究[J].标准科学,2024(9):35-40.