



中华人民共和国国家标准

GB/T XXXXX—XXXX

数据基础设施 标识要求

Data infrastructure—Identification requirements

(点击此处添加与国际标准一致性程度的标识)

(草案)

在提交反馈意见时，请将您知道的相关专利连同支持性文件一并附上。

XXXX – XX – XX 发布

XXXX – XX – XX 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

1 范围 1

2 规范性引用文件 1

3 术语和定义 1

4 标识编码规则 1

 4.1 主体标识码 1

 4.2 平台、接入连接器编码 1

 4.3 数据产品、数据资源编码 2

5 标识赋码要求 2

 5.1 区域/行业功能节点标识赋码要求 3

 5.2 业务节点标识赋码要求 3

 5.3 接入连接器标识赋码要求 3

 5.4 数据标识赋码要求 3

6 标识解析要求 3

7 标识安全要求 3

附录 A（资料性） 标识编码查询流程 4

附录 B（资料性） 校验码 5

参考文献 7

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。

本文件由全国数据标准化技术委员会（SAC/TC609）提出并归口。

本文件起草单位：

本文件主要起草人：

数据基础设施 标识要求

1 范围

本文件规范了数据基础设施标识的术语和定义、标识编码、赋码、解析及安全要求等。

本文件适用于数场、可信数据空间、数联网、数据元件、隐私保护计算、区块链等技术体系支撑的各类层级的数据基础设施建设，包括区域、城市、行业、企业、个人等数据基础设施。

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 32100-2015 法人和其他组织统一社会信用代码编码规则

GB/T 2260-2007 中华人民共和国行政区划代码

GB/T 4754-2017 国民经济行业分类

GB/T 17710-2008 信息技术 安全技术 校验字符系统

3 术语和定义

下列术语和定义适用于本文件。

3.1

标识对象 identification object

数据基础设施涉及的特定物理或逻辑对象，包括机构、人、物品、数据产品、数据资源、平台系统等。

3.2

标识码 identification code

标识码是对标识对象的唯一编码，本文件所述标识指标识码。

4 标识编码规则

4.1 主体标识码

主体标识码共18位，如下：

- 法人和非法人组织。对于拥有组织机构代码的法人或非法人组织，以其统一社会信用代码作为其主体标识码，按照 GB 32100，由 18 位数字或字母组成。
- 自然人。对于中华人民共和国公民，以其中华人民共和国公民身份证号码作为其主体标识码，完成加密处理后生成，由 18 位数字或字母组成。
- 其他类型主体。其他主体可向标识码登记注册平台申请，由平台分配 18 位数字或字母作为其主体标识码。

4.2 平台、接入连接器编码

平台标识码共32位，由类型码（1位）、主体标识码（18位）、区域/行业代码（4位）、随机码（8位）、校验码（1位）组成：

- 全域功能节点编码。类型码为“1”、主体标识码（18位）、区域代码（4位）、随机码（8位）、校验码（1位）。
- 区域功能节点编码。类型码为“2”、主体标识码（18位）、区域代码（4位）、随机码（8位）、校验码（1位）。

- c) 行业功能节点编码。类型码为“3”、主体标识码（18 位）、行业代码（4 位）、随机码（8 位）、校验码（1 位）
- d) 业务节点编码。类型码为“4”、主体标识码（18 位）、区域/行业代码（4 位，业务节点编码由区域功能节点分配时用区域功能节点的区域代码，业务节点编码由行业功能节点分配时用行业功能节点的行业代码）、随机码（8 位）、校验码（1 位）。
- e) 接入连接器编码。类型码为“5”、主体标识码（18 位）、区域/行业代码（4 位，接入连接器编码由区域功能节点分配时用区域功能节点的区域代码，接入连接器编码由行业功能节点分配时用行业功能节点的行业代码）、随机码（8 位）、校验码（1 位）。

注：主体标识码。编码规则参见4.1。

区域代码。参照GB/T 2260 第 5 章中的前4位数字作为本文件区域代码。

行业代码。1~2位固定为数字“00”，3~4位对应GB/T 4754第5章中规定2~3位作为本文件行业代码。

随机码。当用户申请多个标识码出现代码重复时, 随机码进行区分, 确保标识码唯一性。

校验码。校验码采用GB/T 17710的有关规定。

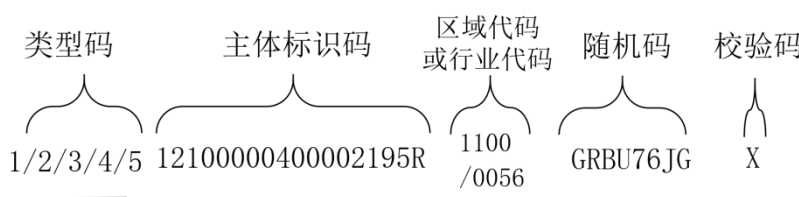


图1 平台、接入连接器编码

4.3 数据产品、数据资源编码

数据产品、数据资源编码共32~65位，由类型码（1位）、主体标识码（18位）、平台区域/行业代码段（4位）、随机码（8位）、校验码（1位）、扩展码（最多32位）组成，如选择扩展码，扩展码之前加“.”（1位）分隔：

- 类型编码。数据产品类型编码为“6”，数据资源类型编码为“7”。
- 主体标识码。编码规则参见 4.1。
- 平台区域/行业代码段。与受理平台的区域代码或行业代码一致。
- 随机码。当用户申请多个标识码出现代码重复时，随机码进行区分，确保标识码唯一性。
- 校验码。校验码采用 GB/T 17710 的有关规定，见附录 B。
- 扩展码。为可选项，编码长度为 1~32 位，可兼容其他标识编码，不选时省略。

如某数据产品分配的标识码为:

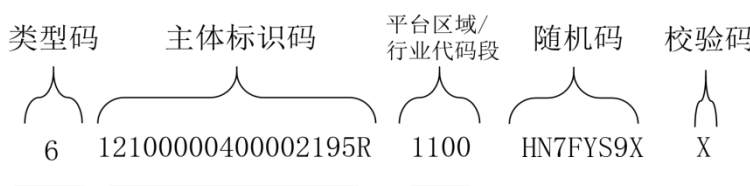


图2 数据产品编码

如某平台登记的数据资源码为：12（登记类型）15（行业分类）20250228（日期码）1BH95JZY（随机码），为其分配的标识码为：

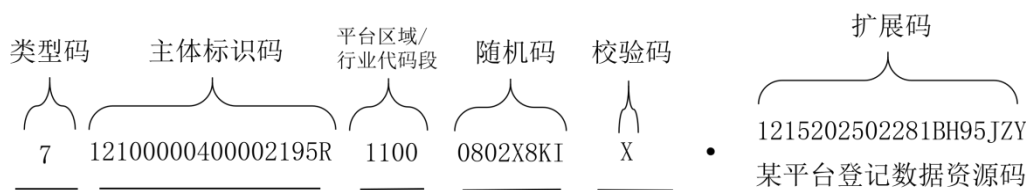


图3 数据资源编码（选扩展码）

5 标识赋码要求

5.1 区域/行业功能节点标识赋码要求

区域/行业功能节点标识编码由全域功能节点负责分配，标识码一旦分配不可变更。

5.2 业务节点标识赋码要求

业务节点向区域/行业功能节点注册时，由区域/行业功能节点向业务节点分配标识码，标识码一旦分配不可变更。

5.3 接入连接器标识赋码要求

数据主体向区域/行业功能节点、业务节点注册接入连接器时，由区域/行业功能节点分配接入连接器标识码。

5.4 数据标识赋码要求

数据标识含数据产品标识和数据资源标识，赋码要求如下：

- a) 用户通过业务节点或区域/行业功能节点登记数据产品/数据资源时，由区域/行业功能节点分配标识码。
- b) 用户通过接入连接器首次登记某一数据产品/数据资源时，由区域/行业功能节点为该数据产品/数据资源分配标识码。
- c) 存量登记平台可通过区域/行业功能节点提供的登记接口进行资源或产品登记，区域/行业功能节点为其分配相应标识码，原有标识可作为扩展码保留，完成原有标识与新分配标识码转换。

6 标识解析要求

全域功能节点、区域/行业功能节点、业务节点应配置相关解析模块，共同构成标识解析体系。标识解析要求如下：

- a) 全域功能节点、区域/行业功能节点、业务节点应支持使用标识码跨域查询，提供数据查询服务目录。
- b) 全域功能节点应提供标识编码、赋码、解析、寻址、维护等能力，维护标识信息库。
- c) 应支持使用标识码作为解析标识符，使用标识码可解析出全域功能节点、区域/行业功能节点、业务节点、接入连接器、数据产品等标识对象的位置信息、状态信息或者其他相关信息。
- d) 全域功能节点、区域/行业功能节点、业务节点需配置标识解析模块，完善标识与 IPv6 等网络标识的映射关系，可通过 DNS 解析等机制实现标识解析。

7 标识安全要求

标识码是身份确认、节点平台互联的关键，应满足标识安全性要求：

- a) 标识使用环境安全：涉及到标识的传输、存储时应符合国家信息安全等级保护、商用密码应用安全性评估等要求。
- b) 标识管理设施安全：涉及标识赋码、查询、解析的平台节点，宜建立标识码冗余检测与纠错机制、唯一性校验、高并发处理、全链路日志审计及容灾恢复等处理方案。

附 录 A
(资料性)
标识编码查询流程

使用标识跨域查询位置的流程描述如下：

- a) 用户在业务节点或区域/行业功能节点对数据标识发起查询请求。
- b) 业务节点查询本地标识数据库并返回结果，查询完成。
- c) 如未查询到该数据标识，业务节点将查询请求发至其对接的区域/行业功能节点，区域/行业功能节点查询本地标识数据库并返回结果，查询完成。
- d) 如区域/行业功能节点未查询到该数据标识则将查询请求发至全域功能节点，由全域功能节点查询标识数据库并返回结果（该标识所在的某节点地址），查询完成。

附录 B
(资料性)
校验码

1 校验码

校验码采用 GB/T 17710,MOD 37-2校验码系统。

1.1 校验公式

标识码中各个位置上的号码字符值应满足下列公式的校验：

$$\sum_{i=1}^{32}(a_i \times W_i) \equiv 1(\text{mod } 37) \cdots \cdots \cdots (1)$$

式中：i——表示号码字符从右至左包括校验码字符在内的置序号；

a_i ——表示第i位置上的号码字符值；

W_i ——表示第i位置上的加权因子，其数值依据公式 $W_i=2^{i-1}(\text{mod } 37)$ 计算得出。表1列出标识码中各个位置上的加权因子 W_i 数值。

表1 标识码中各个位置上的加权因子 W_i 数值

i	32	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16	15
W_i	22	11	24	12	6	3	20	10	5	21	29	33	35	36	18	9	23	30
i	14	13	12	11	10	9	8	7	6	5	4	3	2	1				
W_i	15	26	13	25	31	34	17	27	32	16	8	4	2	1				

1.2 校验码字符值的计算

当i=1时， $W_i=2^0=1$,公式(1)可表示成：

$$a_1 + \sum_{i=2}^{32}(a_i \times W_i) \equiv 1(\text{mod } 37) \cdots \cdots \cdots (2)$$

公式(2)中， a_1 即为校验码字符值，其取值范围是 $0 \leq a_1 \leq 36$ ；当 a_1 值等于36时，用符号“*”表示。

满足于公式(2)及取值范围要求的校验码字符值 a_1 可根据 a_1 与 $\sum_{i=2}^{32}(a_i \times W_i) \equiv 1(\text{mod } 37)$ 的换算关系算出，见表2。

表 2 校验码字符值 a_1 与 $\sum_{i=2}^{32}(a_i \times W_i) \equiv 1(\text{mod } 37)$ 的换算关系表

$\sum_{i=2}^{32}(a_i \times W_i) \equiv 1(\text{mod } 37)$	0	1	2	3	4	5	6	7	8	9	10	11	12	13
校验码字符值 a_1	1	0	*	Z	Y	X	W	V	U	T	S	R	Q	P
$\sum_{i=2}^{32}(a_i \times W_i) \equiv 1(\text{mod } 37)$	14	15	16	17	18	19	20	21	22	23	24	25	26	27
校验码字符值 a_1	O	N	M	L	K	J	I	H	G	F	E	D	C	B
$\sum_{i=2}^{32}(a_i \times W_i) \equiv 1(\text{mod } 37)$	28	29	30	31	32									

校验码字符值a ₁	A	9	8	7	6									
----------------------	---	---	---	---	---	--	--	--	--	--	--	--	--	--

参 考 文 献

- [1] 《国家数据基础设施建设指引》
-