

KSSJ/BM41-2023

智能化矿山数据融合共享 指标数据管理规范

Intelligent mine data fusion and sharing

Specifications for index data management

国家矿山安全监察局
2023年6月

目次

前言	II
1. 范围	1
2. 规范性引用文件	1
3. 术语和定义	1
4. 目标与总体要求	2
4.1. 完整性要求	3
4.2. 唯一性要求	3
4.3. 准确性要求	3
4.4. 规范性要求	3
5. 指标体系框架	3
6. 指标数据定义规范	3
6.1. 指标数据定义框架	3
6.2. 指标数据分类	4
6.3. 指标数据标准属性定义	5
7. 指标数据质量管理相关规范	11
附录 A	13
附录 B	14
附录 C	15
附录 D	16
附录 E	17
参考文献	19

前 言

本文件参照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。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别这些专利的责任。

本规范起草单位：国家能源投资集团有限责任公司、国能信息技术有限公司、中国华电集团有限公司、陕西煤业化工集团有限责任公司、应急管理部信息研究院、华电煤业集团有限公司、华电煤业集团数智技术有限公司、中安智讯（北京）信息科技有限公司、河南能源集团有限公司、晋能控股集团有限公司、煤炭科学研究总院有限公司矿山大数据研究院、航天智控（北京）监测技术有限公司、华能煤炭技术研究有限公司。

本文件技术指导：杨荣明、徐会军、田臣、王海春、王致兵、王鹏、胡而已、孙长春、赵宇波、张玉波。

本规范主要起草人：辛华、王晓燕、张冬阳、李文慧、闫计栋、韩培强、邱爽、朱拥军、牟岩、王宇静、王宁、徐金陵、黄金、陈帅领、侯尚武、宋文兵、李俊、郭军、胡勇、胡兵。

智能化矿山数据融合共享 指标数据管理规范

1. 范围

本文件规定了智能化矿山指标数据的体系框架、指标数据定义框架、指标数据分类和指标数据标准属性定义。

本文件适用于智能化矿山指标数据的统一管理和约束。

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；凡是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/T 34679-2017 智慧矿山信息系统通用技术规范

GB/T 36073-2018 数据管理能力成熟度评估模型

DAMA国际.DAMA-DMBOK2数据管理知识体系指南

DZ/T 0376-2021 智慧矿山建设规范

煤矿智能化建设指南（2021年版）

3. 术语和定义

3.1. 指标数据 index data

指标数据是智能化矿山在战略发展、业务运营和管理支持等各领域业务分析过程中衡量某一个目标或事物的数据。指标数据反映数量特征的概念和数值，包括指标数据标准和指标数据值两部分。

3.2. 基础项指标 basic index

基础项指标是表达业务实体原子量化属性的概念集合，具有明确的业务定义且业务定义不可进一步拆解的指标。通常是可以直接对单一变量的明细数据进行简单计算得到的不可进一步拆解的指标。如“调运量”、“销售量”等。具有如下特征：

- 1、指标计算规则中仅包含一个变量；

2、其稳定性高，业务定义、计算公式和统计口径不易随业务管理视角频繁变化。

3.3. 复合项指标 composite index

复合项指标是建立在基础项指标之上，由若干个基础项指标通过一定运算规则计算形成，在业务角度无法拆解，如“签约率”、“利润率”。具有如下三种类型：

- 1、由基础项指标计算得到；
- 2、由基础项指标和复合项指标计算得到；
- 3、由复合项指标计算得到。

3.4. 派生类指标 derived index

派生类指标是基本类指标（即基础项指标或复合项指标）与一个或多个维度值相结合产生的指标。如“月计划调运量”、“月日均销售量”等。具有如下两种类型：

- 1、由基础项指标和维度组合得到；
- 2、由复合项指标和维度组合得到。

3.5. 维度 dimension

维度是对智能化矿山在业务经营过程中所涉及对象的属性进行划分的方式。维度作为观察事物的视角，并不孤立存在，而是通过与指标数据结合使用，可以对指标数据的不同方面进行对比与分析。

4. 目标与总体要求

指标数据设计以智能化矿山业务定位为基础，遵循数据编码专题组数据分类整体框架，形成智能化矿山统一、完善、适用的指标分析体系框架，通过指标在能化矿山范围内的统一定义、存储、管理和使用，有效提升智能化矿山指标统计规范和指标数据质量，推动智能化矿山各层级单位指标数据资源的整合与共享，全面提高智能化矿山数据资产的利用价值。

指标数据需要遵循的以下总体要求：

4.1. 完整性要求

指标的信息应避免缺项，保证内容完整。

4.2. 唯一性要求

保证指标选取的全面，避免指标之间重复。

4.3. 准确性要求

每项指标都必须准确体现业务需求，能够科学地反映评价对象的某一方面信息，避免出现同名不同义或者同义不同名等有歧义的指标。

4.4. 规范性要求

智能化矿山对指标数据标准的定义和分类提出了明确的要求，各专业、各层级应严格按照本规范的要求开展指标数据标准工作。

5. 指标体系框架

智能化矿山指标数据体系遵循数据编码专题组整体框架，围绕智能化矿山核心生产运营业务，以全面性、整体性和稳定性为基本原则，构建指标数据体系框架，指导后续相关指标的梳理与规范编制工作，综合、全面反映智能化矿山运营情况，为智能化矿山的各级决策提供统计分析依据，也为未来的大数据深度应用提供基础数据保障。具体指标体系框架参见数据编码规范。

6. 指标数据定义规范

6.1. 指标数据定义框架

指标数据定义框架规范了指标数据分类和指标数据标准属性。指标分为基本类指标和派生类指标，其中基本类指标又分为基础项指标和复合项指标。指标数据标准属性规定了指标数据标准的业务属性、技术属性和管理属性，可应用在智能化矿山数据资产管理的相关领域。其中，业务属性、技术属性和管理属性中的必填属性称为基础属性，其他可根据实际情况进行梳理的属性称为扩展属性。

指标数据定义框架如图1所示：

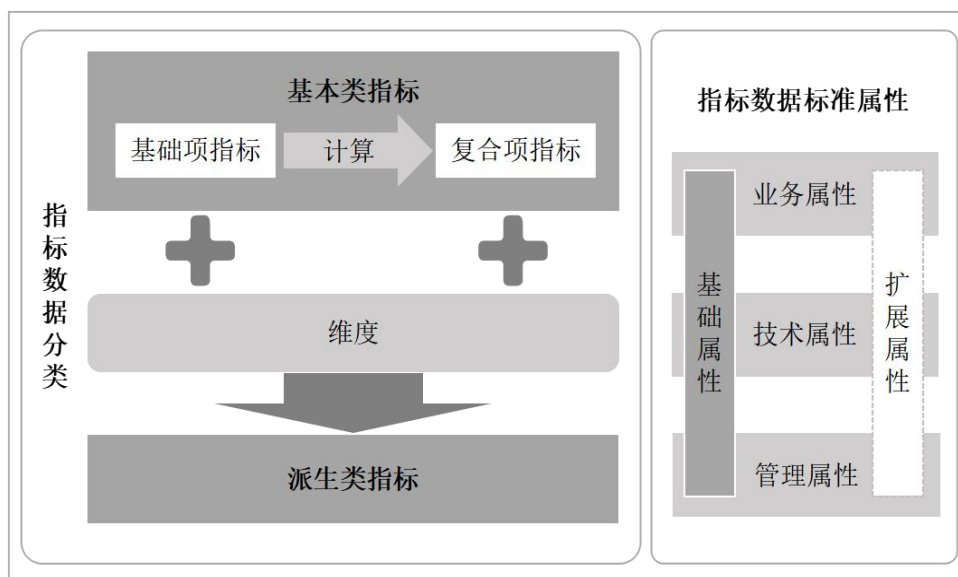


图 1 指标数据定义框架

6.2. 指标数据分类

6.2.1. 基本类指标

6.2.1.1. 基础项指标

基础项指标是表达业务实体原子量化属性的概念集合，是可以直接对单一变量的明细数据进行简单计算得到的不可进一步拆解的指标。如“调运量”、“销售量”等。具有如下特征：

- 1、指标计算规则中仅包含一个变量；
- 2、其稳定性高，业务定义、计算公式和统计口径不易随业务管理视角频繁变化。

6.2.1.2. 复合项指标

复合项指标是建立在基础项指标之上，由若干个基础项指标通过一定运算规则计算形成，在业务角度无法拆解，如“签约率”、“利润率”。具有如下三种类型：

- 1、由基础项指标计算得到；
- 2、由基础项指标和复合项指标计算得到；
- 3、由复合项指标计算得到。

6.2.2. 派生类指标

派生类指标是基本类指标（即基础项指标或复合项指标）与一个或多个维度值相结合产生的指标。如“月计划调运量”、“月日均销售量”等。具有如下两种类型：

- 1、由基础项指标和维度组合得到；
- 2、由复合项指标和维度组合得到。

6.3. 指标数据标准属性定义

1、业务属性

业务属性是数据在业务层面的定义，描述数据和智能化矿山业务相关联的特性，是数据业务含义的统一解释及要求。业务属性包括：一级主题、二级主题、三级主题、四级主题、指标名称、业务定义、处理逻辑、维度、基础计量单位、参考标准、上报频度、统计时间、映射类型。

2、技术属性

技术属性是数据在技术层面的定义，描述了数据与信息技术实现相关联的特性，是数据在信息系统项目实施时统一的技术方面定义。技术属性包括数据类型、数据格式、数据源系统、数据源表名、数据源字段名。

3、管理属性

管理属性是数据在管理层面的定义，描述了数据与数据管理相关联的特性，是数据管理在数据标准管理领域的统一要求。管理属性包括：指标编码、数据主责部门、标准管理部门、颁布日期、废止日期。

指标数据标准属性定义如图2所示：

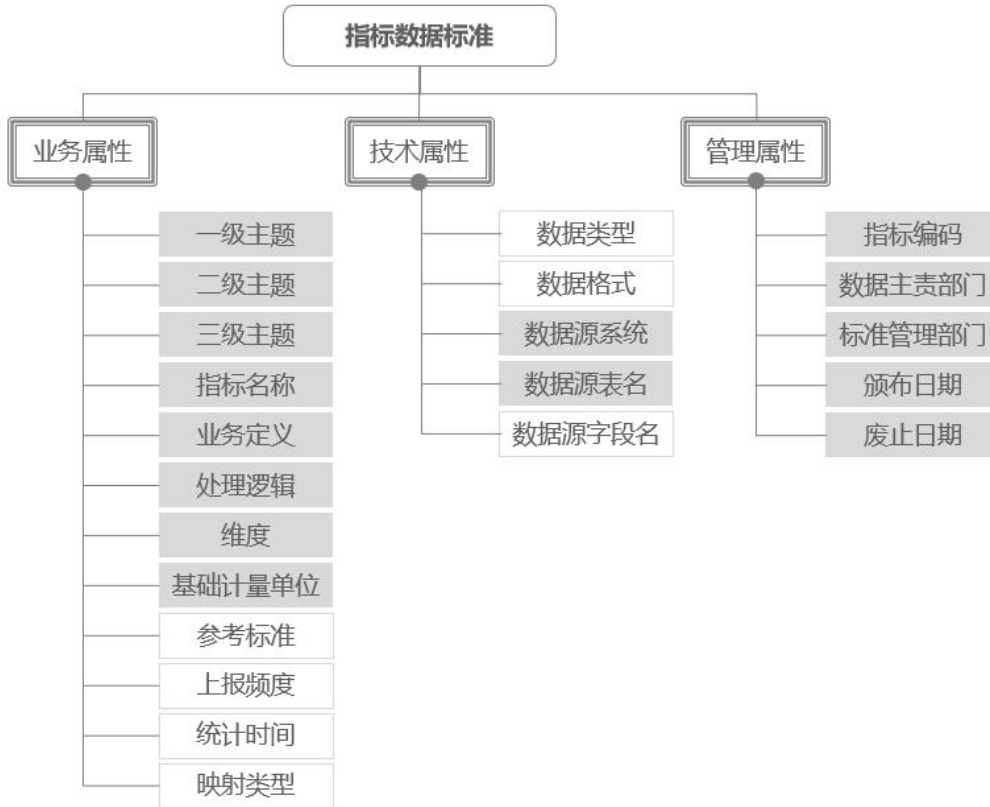


图 2 指标数据标准属性定义

说明：

1) 基本类指标数据标准（即基础项指标和复合项指标）属性包含标记为灰色背景部分的属性；

2) 派生类指标数据标准属性包含全部属性。

6.3.1. 业务属性**6.3.1.1. 一级主题**

一级子主题是指标数据按业务分析应用的实际情况对主题进一步细化，包含基础类、安全类、生产类、管理类四个一级子主题。详见附录 E

6.3.1.1. 二级主题

二级子主题是指按业务细分的实际情况对一级子主题进一步细化，包含主要采煤、掘进、供配电等 57 个二级子主题。详见附录 E

6.3.1.2.三级主题

三级子主题是指按业务细分的实际情况对二级子主题进一步细化。

6.3.1.3.指标名称

指标名称是指标项的统一中文命名，每个指标具有唯一的规范化名称。通过指标名称可以从业务上区分不同的数据，指标名称应易于被数据使用人员理解和识别。

6.3.1.4.业务定义

业务定义是指标项在业务上的解释、描述或者说明，是基于创建数据的业务流程对数据业务口径和相关业务场景的详细描述，是数据业务含义的自然语言表述。

业务定义应该精准、细致，以利于数据使用人员理解，不应有循环引用或直接用指标名称进行定义。业务定义可参考相关国家标准和行业标准、外部监管机构的定义、公司内部业务制度、信息系统业务需求定义、以及行业经验的总结性归纳。

1) 业务定义主要从指标的原理或概念解释、指标用途、指标统计所覆盖业务范围、智能化矿山的业务特色、与通用概念和监管要求的差异性等方面进行描述；

2) 考虑到一些指标的业务定义过于复杂,需要较长篇幅说明,在描述指标的业务定义时,如果该指标的相关定义在智能化矿山其他相关文件或者权威机构发布文献中已有详细描述,可以描述为参见该文献说明。

6.3.1.5.处理逻辑

处理逻辑是指标数据对应的计算处理方式。即指标数据核算、平衡、拆分等业务规则(计算公式\函数等)，描述指标的本质和内涵，是满足“口径统一”的要求。

6.3.1.6. 维度

维度是报表分割显示统计数值的角度，主要用来描述在业务活动中会从哪些角度对指标项进行使用和分析，体现为报表的一行或者表头中的一列。一般来说，维度具有离散化取值的特性，即取值可以枚举。常用维度包括时间、组织、业务板块等。

1) 维度设计的参照系通常为基础数据标准中的标准代码。

2) 对于具体的指标，其业务场景可能只与部分维度结合，也可能只在某一维度中取其中的某几个维度值，因此在指标数据标准设计过程中，将具体描述每个指标适用的维度。

3) 维度定义模板详见附录B-表2 维度定义（模板）。

6.3.1.7. 基础计量单位

基础计量单位是指标在经营管理或者报表统计中通常使用的最小度量衡，包括千伏安/千瓦、元、百分比、万吨、万方、美元等。如采剥量描述露天煤矿开采过程中土岩剥离和采煤的工作总量，开采过程中使用的最小计量单位为立方米，则其基础计量单位为立方米。

6.3.1.8. 参考标准

参考标准是指标数据标准的业务依据来源类型和详细依据，包括国家标准、行业标准、制度发文等。参考标准应遵循如下规范：

1) 对于依据文件应填写完整的文件名称，例如：“国家标准_《GB/T 2261.1-2003 个人基本信息分类与代码 第1部分 人的性别代码》”；

2) 如一个数据标准有多个制定依据，则用顿号“、”分隔。

6.3.1.9. 上报频度

上报频度是数据上报（产生）的频度定义，包含日、月、季度、年度等。

6.3.1.10. 统计时间

统计时间是指数据汇总时间，如月初、每月十号等。

6.3.1.11.映射类型

映射类型描述指标项是否为基础指标项或基础指标项计算得来，取值包括直接映射和计算映射。

6.3.2. 技术属性

6.3.2.1.数据类型

数据类型是根据数据的业务定义、业务规则和常见表现形式定义其所采用的数据类。数据类型包括：编码类、代码类、标志类、文本类、金额类、比例类、数值类、日期类、时间类、日期时间类。

数据类型的详细定义见附录C。

6.3.2.2.数据格式

数据格式描述数据在精度、长度、形态上的定义，包括所允许的最大和/或最小字符长度，数据的表示格式等。

数据格式的详细定义见附录D。

6.3.2.3.数据源系统

数据源系统是指标对应数据源所在的系统名称。

6.3.2.4.数据源表名

数据源表名是指标对应的数据源系统表名称。

6.3.2.5.数据源字段名

数据源字段名是指标对应的数据源系统字段名称。

6.3.3. 管理属性

6.3.3.1.指标编码

指标编码是对指标数据标准的统一编号，按照规定的编码规则由系统自动生成指标数据编码。

基本类指标编码由标识字符、主题域编码和流水码三部分共11位字符组成

（如下图基本类指标编码规则），具体说明如下：

1.一位标识字符：标识该标准类型为指标数据标准，英文缩写为Q（Quota的首字母）。

2.五位主题域编码：1位一级主题域编码+2位二级主题域编码+2位三级主题域编码，各级主题域编码参见附录E智能化矿山指标数据标准主题域编码对照表。

3.五位流水码：采用十进制数字顺序编码法，用5位阿拉伯数字表示，在不同主题域

编码下从00001开始按升序编码。

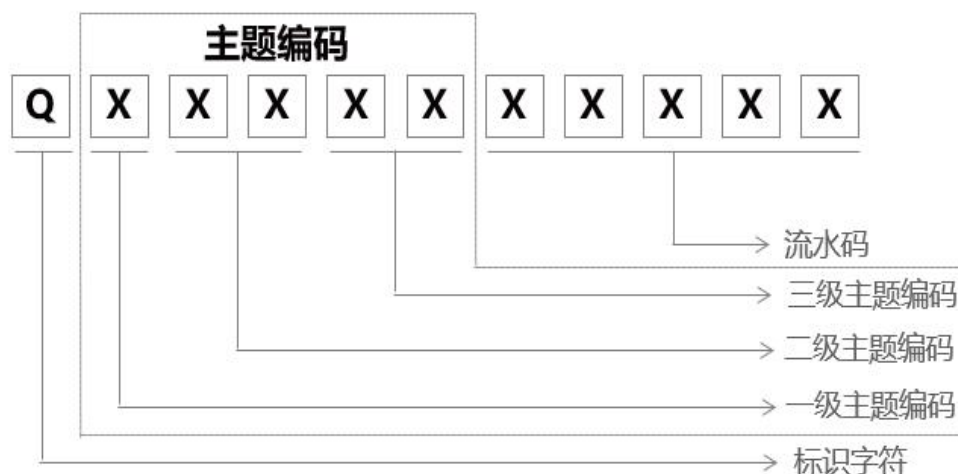


图 3 基本类指标编码规则

表 1 基本类指标编码示例

指标名称	指标编码	一级主题域	二级主题域	三级主题域
万吨掘进率	Q2040200001	2 生产数据	04 主要系统	02 掘进工作面

派生类指标编码由标识字符、日期和流水码两部分共14位字符组成（如下图派生类指标编码规则），具体说明如下：

1.一位标识字符：标识该标准类型为指标数据标准，英文缩写为Q（Quota的首字母）。

2.八位日期编码：系统生成编码的日期，格式：YYYYMMDD。

3.五位流水码：采用十进制数字顺序编码法，用5位阿拉伯数字表示，在不同日期下从00001开始按升序编码。

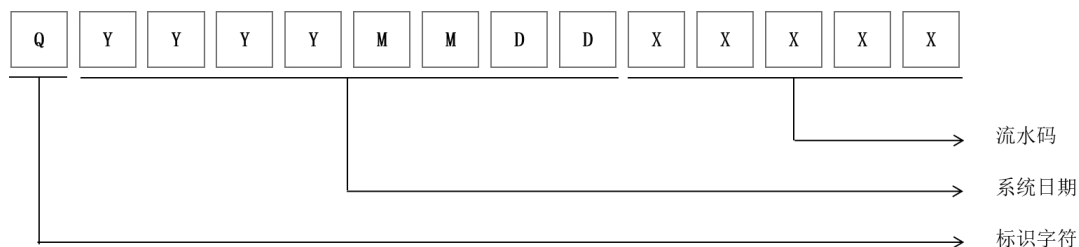


图4 派生类指标编码规则

6.3.3.2.数据主责部门

数据主责部门是指对指标数据标准的业务属性拥有最终业务解释权的业务单位。数据主责部门通常为该数据所涉及智能化矿山业务的主管部门。

6.3.3.3.标准管理部门

标准管理部门是指对指标数据标准管理负责的业务单位。原则上指标数据标准由信息部统一归口管理，数据管理者为信息部。

6.3.3.4.颁布日期

颁布日期是指指标数据标准在公司正式颁布启用的日期，格式为“YYYY-MM-DD”，例如“2023-01-05”。

6.3.3.5.废止日期

废止日期是指指标数据标准在公司正式停止使用的日期，格式为“YYYY-MM-DD”，例如“2022-10-01”。

7. 指标数据质量管理相关规范

指标数据质量管理是对指标数据规范性、准确性、一致性等方面进行质量评价，为指标数据质量改进提供依据。基本要求如下：

智能化矿山范围内各层级单位使用同一套信息化指标体系，保证指标库的唯一性、权威性；指标设计客观实用、指标计算方式简便易操作、具有同业横向可比性，根据管理需要对指标体系进行持续改进；指标设计要能够客观反映实际工作情况。指标库由指定部门统一管理，确保指标管理的统一规范；智能化矿山可按组织层级，对指标实行分层管理，一级对一级负责。对于指标设置和目标值的确定，确保下层组织目标在职能范围内对上级组织目标形成支撑；

指标值统计要认真严谨、实事求是，反映信息化工作真实成果和业绩。

附录 A

表 A1 基本类指标（即基础项指标、复合项指标）指标数据（模板）

指标编码	一级主题	二级主题	三级主题	指标名称	业务定义	处理逻辑	维度	基础计量单位	参考标准	统计频度	提报时间	指标类别	数据类型	数据格式	数据源系统	数据源表名	数据源字段名	数据主责部门	标准管理部门	颁布日期	废止日期
Q2040200001	生产数据	04 主要生产系统	02 掘进工作面	万吨掘进率	是指采掘工作中反映掘进进尺与回采产量的比例关系，即一定时期内开掘巷道的总长度（m）与原煤产量（万吨）的比率。	万吨掘进率=生产掘进进尺/矿井原煤产量	时间，组织	米/万吨	\	日	8:00:00	复合项指标	数值类	n. . 12	线下报表	YGL0092_煤炭技术经济指标表 CHN、YGL0056_煤炭生产企业技术经济指标表 CHN	\	煤炭运输部	信息化部		

附录 B

表 B1 维度定义（模板示例参考）

维度编号	维度名称	维度定义及说明	维值	维值名称	维值含义
1	时间	以时间作为描述、表达变量的度量尺度。是在分析经济变量时加上时间要素的角度而形成的	1	日	\
			2	旬	\
			3	月	\
			4	季	\
			5	半年	\
			6	年	\
2	计算规则	指标计算过程中常用的计算方法	1	比上日增量	\
			2	比上月增量	\
			3	比同期增量	\
			4	去年同期	\
			5	本年累计	\
3	组织	集团公司组织机构	1	XX 公司	\
			2	XX 公司	\
			3	XX 公司	\
			4	XX 公司	\
			5	--	引用组织架构主数据
4	产业	集团公司产业分类	1	煤炭	\
			2	火电	\
			3	新能源	\
			4	水电	\
			5	运输	\
			6	化工	\
			7	科技环保	\
			8	----	\
5	煤炭资源储量分类	描述煤炭资源储量的分类方法	1	储量	\
			2	基础储量	\
			3	资源量	\

附录 C

C.1 范围

本附录给出了指标数据指标的数据类型。

C.2 数据类型

数据指标的数据类型包括编码类、代码类、标志类、文本类、金额类、比例类、数值类、日期类、时间类、日期时间类。

C.2.1 编码类

编码是用少量、简单的基本符号，选用一定的组合规则，表示大量复杂多样的信息。

C.2.2 代码类

代码是一套预先定义的，用来描述一个有限集合的事物或事物的属性，代码数据能够相对稳定的在一段时期内。

C.2.3 标志类

表示“是/否”意义的标志。

C.2.4 文本类

需要以文本的形式对与业务活动密切相关的对象和业务的进行说明的数据。

C.2.5 金额类

金额类指以货币金额的形式体现的数据项，适用于各类财务信息。金额类数据标准，需扩展定义度量单位（元、万元等）。

C.2.6 比例类

比例类指以比值的形式体现的数据项，适用于各类比率信息。

C.2.7 数值类

数值类指除金额类及比例类外的以整数或小数的形式体现的数据项，适用于各类以数量反映的信息。数值类数据标准，需扩展定义度量单位。

C.2.8 日期类

日期类指以需要日期的形式体现的数据项，以描述业务发生的日期。

C.2.9 时间类

时间类指以需要时间的形式体现的数据项，以描述业务发生的时间。

C.2.10 日期时间类

日期时间类指以需要以日期和当日时间的组合形式体现的数据项，以描述业务发生的日期和时间。

附录 D

D.1范围

本附录给出了指标数据标准的数据格式定义。

D.2数据格式

- a 字母字符。
 n 数字字符。
 an 字母数字字符。
 anc 字母数字汉字字符。
 M、N 表示自然数。

表 D1 数据格式

表示方式	含义
M!a	M 位字母字符，定长。
M!n	M 位数字字符，定长。
M!an	M 位字母数字字符，定长。
M!anc	M 位字母数字汉字字符，定长。
a..M	最多为 M 位字母字符。
n..M	最多为 M 位数字字符。
an..M	最多为 M 位字母数字字符。
anc..M	最多为 M 位字母数字汉字字符。
aM..	最少为 M 位字母字符。
nM..	最少为 M 位数字字符。
anM..	最少为 M 位字母数字字符。
ancM..	最少为 M 位字母数字汉字字符。
aM..N	最少为 M 位最多为 N 位字母字符。
nM..N	最少为 M 位最多为 N 位数字字符。
anM..N	最少为 M 位最多为 N 位字母数字字符。
ancM..N	最少为 M 位最多为 N 位字母数字汉字字符。
M(N)	M 位数字字符，其中包括小数点和 N 个小数位 (M>N+1)。
YYYY-MM-DD	日期格式，表示年月日。Y 表示时间元素“年”所用的数字，M 表示时间元素“月”所使用的数字，D 表示时间元素“日”所使用的数字。
HH:MM:SS	时间格式（24 小时制），表示时分秒。H 表示时间元素“小时”所使用的数字，M 表示时间元素“分钟”所使用的数字，S 表示时间元素“秒”所使用的数字。
YYYY-MM-DDTHH:MM:SS	日期时间格式，表示某年某月某日某时某分某秒。T 为时间标识符，在日期和日的组合表达式中，用以标识该日的时间表示法的开始。

附录 E

表 E1 主题域编码

编码		主题域名称
一级	二级	
1		基础类
	01	证照信息
	02	机构
	03	地质条件
	04	开采条件
	05	灾害条件
	06	IT 基础设施
2		生产类
	01	采煤
	02	掘进
	03	供配电
	04	提升
	05	主运输
	06	辅助运输
	07	通风
	08	压风
	09	供水
	10	排水
	11	降温制冷
	12	洗选
	13	调度管理
	14	生产技术管理
	15	生产计划管理
	16	机电管理
	17	爆破管理
3		安全类
	01	顶板防治
	02	冲击地压防治
	03	水害防治
	04	火灾防治
	05	瓦斯防治
	06	粉尘防治
	07	热害防治
	08	安全监控
	09	井下作业人员管理
	10	视频监控
	11	通讯联络
	12	风险分级管控
	13	事故隐患排查治理
	14	上级安全检查
	15	不安全行为管理

KSSJ/BM41-2023

	16	事故管理
	17	安全培训
	18	职业健康
	19	应急管理
	20	环境保护
4		管理类
	01	人力资源管理
	02	财务管理
	03	审计管理
	04	物资管理
	05	设备管理
	06	运销管理
	07	节能减排管理
	08	科技管理
	09	项目管理
	10	法务管理
	11	综合管理
	12	信息化管理

参考文献

- [1] GB/T 5271.18-2008 信息技术 词汇 第 18 部分：分布式数据处理
 - [2] GB/T 10113-2003 分类与编码通用术语
 - [3] GB/T 15259-2008 煤矿安全术语
 - [4] GB/T 18725-2008 制造业信息化 技术术语
 - [5] GB/T 32400-2015 信息技术 云计算 概览与词汇
 - [6] GB/T 34679-2017 智慧矿山信息系统通用技术规范
 - [7] GB/T 35295-2017 信息技术 词汇 第 2 部分：大数据及其应用领域术语
 - [8] GB/T 37700-2019 信息技术 工业云 参考模型
 - [9] 煤矿智能化建设指南（2021 年版） 国家能源局、国家矿山安全监察局
 - [10] 煤矿安全规程（2022 版） 应急管理部
-