

团体标准

T/BSIA 00X-2026

衍生数据认定方法 第3部分 矿山行业

Method standard for testing the identification of derivative data

part 3: mining industry

(征求意见稿)

2026-xx-xx 发布

2026-xx-xx 实施

北京软件和信息服务业协会 发布

目 次

前言	IV
引言	IV
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 认定原则	1
4.1 客观性原则	2
4.2 可追溯性原则	2
4.3 行业全面性原则	2
4.4 审慎性原则	2
4.5 处理过程优先原则	2
5 认定条件	2
5.1 认定基础条件	2
5.2 认定专项条件	3
6 认定要点	4
6.1 数据处理过程分析法认定要点	4
6.2 数据依赖关系分析法认定要点	7
6.3 数据特征比对法认定要点	7
6.4 人工复核与确认法认定要点	8
7 判定准则	9
7.1 认定方法的使用准则	9
7.2 基于数据处理过程的判定准则	9
7.3 基于数据来源关系的判定准则	9
7.4 基于数据特征比对的判定准则	10
7.5 基于人工复核的判定准则	11
8 记录与报告	11

8.1 认定记录	11
8.2 认定报告	13
附录 A（资料性） 认定案例	15
A.1 案例一：矿山资源储量评估数据认定	15
A.2 案例二：矿山设备故障预测数据认定	17
参考文献	19

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。

《衍生数据认定方法》分为5个部分：

- 第1部分：总则；
- 第2部分：通信行业；
- 第3部分：矿山行业；
- 第4部分：医疗健康行业；
- 第5部分：情感认知行业。

本部分为《衍生数据认定方法》第3部分。

本文件由北京市大数据中心提出，由北京软件和信息服务业协会归口。

本文件起草单位：煤炭科学研究总院有限公司、北京市大数据中心、北京软件和信息服务业协会。

本文件主要起草人：程健，骆意，杨培培，黄晴，李海昌，邓院圆，李和平，王冠杰，贾琨，谢子轩、赵章界、赵莹、方方。

引 言

随着矿山行业数字化转型与人工智能技术的深度应用，地质勘探、设备运行、生产调度、安全监控等全业务链条产生了海量多源异构数据，衍生数据作为挖掘数据价值、支撑智能决策的核心载体，其合规认定成为数据安全流通、价值评估及权益保护的关键前提。

本文件严格遵循《衍生数据认定方法 第1部分 总则》（以下简称《总则》）确立的“客观性、可追溯性、全面性、审慎性、处理过程优先”核心框架，结合矿山行业“时空强关联、安全优先级高、多模态数据融合、业务场景绑定紧密”的鲜明特征，以矿山实际数据处理场景为依托，参考《煤矿安全监控系统通用技术要求》《矿山数据治理指南》等行业规范及大量实践案例，将抽象认定原则转化为具象化、可操作的行业专用认定流程与准则。

本文件核心技术内容基于以下依据制定：一是《总则》的通用认定逻辑与方法体系；二是矿山行业数据处理典型场景（如资源储量评估、设备故障预测、安全风险预警等）的处理流程特征；三是矿山安全合规、数据治理的行业特殊要求；四是信息熵量化分析在矿山结构化与非结构化数据中的适配性验证结果。通过明确矿山特有数据转换操作、细化行业专属判定阈值、规范认定记录编码规则等，旨在消除行业认定过程中的主观性与模糊性，为矿山行业衍生数据认定提供统一、权威、可落地的技术依据。

衍生数据认定方法 第3分：矿山行业

1 范围

本文件规定了矿山行业衍生数据认定的范围、规范性引用文件、术语和定义、认定原则、认定条件、认定要点、判定准则、记录与报告要求，以及相关资料性附录。

本文件适用于矿山行业中通过数据处理活动生成的数据是否属于衍生数据的判定。

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/T 15259-2008 矿山安全术语

GB 16423-2020 金属非金属矿山安全规程

GB/T 34960.5-2018 信息技术服务 治理 第5部分：数据治理规范

GB/T 35295-2017 信息技术 大数据 术语

GB/T 37973-2019 信息安全技术 大数据安全管理指南

GB/T 46010-2025 信息技术 矿山大数据 技术要求

3 术语和定义

3.1

矿山数据 Mine Data

矿山生产、安全、经营、环保、职业健康等环节产生的以及不直接由矿山产生但与矿山安全、生产相关的数据。

[来源：GB/T46010-2025,3.3，注：该标准将于2026年2月1日实施，当前暂按此版本引用，实施后执行最新版本]

3.2

矿山衍生数据 Mine derived Data

基于矿山源数据，遵循矿山行业特定的数据转换操作处理生成，在信息内容、结构或特征上与输入数据存在显著差异，可辅助矿山业务活动开展的衍生数据。

4 认定原则

4.1 客观性原则

认定过程必须基于矿山行业实际业务数据和事实证据，不应受主观判断（如对矿山数据价值的主观认定）或外部因素（如业务部门利益诉求）影响。

4.2 可追溯性原则

认定结论应能够追溯到所依据的矿山输入数据（如地质原始钻探数据、设备传感器数据）、完整的矿山数据处理过程（如数据处理流程图、算法代码）、采用的认定方法及其结果。

4.3 行业全面性原则

在界定的认定范围内，认定工作应结合矿山行业业务特点，覆盖矿山行业数据处理的完整链条，分析矿山数据处理关键环节，不应遗漏可能产生矿山衍生数据的处理操作。

4.4 审慎性原则

对于矿山行业复杂或难以确定的情况，宜组合使用多种认定方法，并必须进行人工复核，确保认定的准确性和可靠性。

4.5 处理过程优先原则

在判定过程中，矿山数据处理过程分析是基础性、优先性的判定依据。当处理过程分析和结果数据特征比对的初步结论冲突时，应以处理过程分析的结论为主，并加强人工复核。

5 认定条件

5.1 认定基础条件

5.1.1 数据条件

矿山行业衍生数据认定应具备的数据条件如下：

a) 应获取认定对象的完整数据样本，样本应具有矿山行业代表性，能够反映认定对象的整体特征；

b) 应获取认定对象对应的输入源数据，输入数据应完整、准确，能够支撑认定分析；

c) 应获取数据处理者享有数据使用权及衍生数据生成过程保护各方合法权益的证明文件（如合规法律意见书、授权协议、产权登记凭证等）。

5.1.2 技术条件

矿山行业衍生数据认定应具备的技术条件如下：

a) 应具备分析矿山数据处理过程的技术工具，如代码分析工具、数据流程图绘制工具、数据血缘分析工具等；

- b) 应具备数据特征比对的技术能力，如统计分析工具、数据可视化工具、数据比对脚本；
- c) 若涉及矿山行业复杂模型，应具备模型解析能力，能够获取模型的输入输出参数、核心算法逻辑、训练数据来源等信息。

5.1.3 人员条件

矿山行业衍生数据认定应具备的人员条件如下：

- a) 认定人员应具备矿山行业相关专业（如采矿工程、地质工程、安全工程、矿山机电等）本科及以上学历或中级及以上专业技术职称，熟悉矿山业务流程（如地质勘探、生产调度、安全监控等）和矿山数据特点（如时空关联性、多源异构性）；
- b) 认定人员应具备数据技术能力，能够理解矿山数据处理逻辑，掌握数据特征分析方法；
- c) 人工复核人员应包含矿山行业专家、数据技术专家，必要时可包含法律合规专家；
- d) 认定人员应具备对应矿山类型的专业适配性。煤矿：需包含采矿工程（煤矿方向）、安全工程（瓦斯防治方向）专业人员；金属非金属矿山：需包含地质工程（金属矿方向）、岩土工程（边坡治理方向）专业人员；露天矿山：需包含采矿工程（露天开采方向）、测绘工程专业人员。

5.2 认定专项条件

5.2.1 数据处理过程分析法认定条件

数据处理过程分析法认定前提条件为：能够访问或获取描述矿山数据处理过程的清晰文档、脚本、代码或配置信息。应获取的关键素材如下：

- a) 应获取矿山数据处理相关文档，包括但不限于矿山数据处理流程图、系统设计文档、处理脚本或代码、矿山算法模型定义文件、计算逻辑说明、日志文件等；
- b) 应获取矿山数据源相关文档，包括但不限于矿山数据源清单、数据字典、数据血缘图；
- c) 应获取矿山业务相关文档，包括但不限于矿山业务流程说明书、矿山行业标准、业务规则说明。

5.2.2 数据依赖关系分析法认定条件

数据依赖关系分析法认定前提条件为：能够获取认定对象与其输入数据之间的数据血缘信息、ETL 元数据、数据库依赖关系、API 调用记录或可解析的数据流配置。应获取的关键素材如下：

- a) 矿山时空数据血缘图，支持字段级血缘，包含采区 ID、时间粒度、坐标范围等时空

标签：

- b) 矿山多源 ETL 元数据，包含数据抽取规则、跨系统转换规则、时空对齐规则等；
- c) 矿山数据库依赖关系图，包含业务表关联、时空表关联等；
- d) 矿山设备 API 调用日志，包括传感器 API、第三方系统 API、调用频率等；
- e) 其他数据，包括矿山数据采集规范、数据中台数据管理规程等。

5.2.3 数据特征比对法认定条件

数据特征比对法认定前提条件为：能够获取认定对象及其潜在输入数据的样本，及其数据字典或模式定义。应获取的关键素材如下：

- a) 认定对象与输入数据样本；
- b) 矿山数据字典，包括字段业务含义、安全指标范围、时空粒度定义等；
- c) 数据统计工具与脚本，包括时空聚合脚本、空间分析工具、时序特征提取工具等；
- d) 数据比对脚本或工具。

5.2.4 人工复核与确认法认定条件

前提：组建符合要求的复核小组，且可提供完整的前期认定资料（分析报告、数据样本、处理文档、比对结果等），支撑专家全面审查。

6 认定要点

6.1 数据处理过程分析法认定要点

6.1.1 信息获取和解析

应获取与认定对象生成相关的完整矿山行业数据处理脚本、代码或详细配置文档。应使用人工审查、代码分析工具或矿山业务数据处理平台界面，解析并理解数据的输入来源、中间处理环节、逻辑和规则。

6.1.2 识别关键数据转换操作

矿山行业关键数据转换操作按矿山“地质勘探→生产开采→安全监控→设备运维→综合管理”业务流程，各类型矿山关键转换操作如下：

6.1.2.1 地质勘探阶段

地质勘探阶段关键数据转换操作主要包括：

多源矿山数据复杂的关联（JOIN）或合并（UNION）：整合钻探、物探、化探等不同系统、不同来源的异构数据，通过作业区域 ID、坐标等唯一标识进行多字段关联，形成标准化

地质数据集，支撑综合地质模型构建。

表 1 各类型矿山地质勘探阶段关键数据转换操作

序号	主要矿山类型	关键数据转换操作	应用场景
1	煤矿	煤层厚度三维插值计算、多源地质数据（钻探、物探、化探）关联整合等。	通过作业区域坐标、煤层编号等唯一标识实现异构数据标准化，支撑煤层储量估算模型构建。
2	金属非金属矿	矿体品位空间分布建模、多源勘探数据合并等。	基于矿体赋存规律实现数据整合，为矿体资源量计算提供基础。

6.1.2.2 生产开采阶段

生产开采阶段关键数据转换操作主要包括：

a) 生产数据分组聚合

按采场、班次、开采周期等维度，对产量、品位、瓦斯浓度、顶板压力等实时单点数据执行聚合运算，实现数据粒度跃迁，生成采场日产量、阶段性生产汇总等报表。

b) 矿山专业指标计算

基于矿山行业公式及业务逻辑，通过出矿数据与地质数据等多字段组合计算，生成开采贫化率等源数据中不存在的核心指标，反映生产、资源等业务状态。

表 2 各类型矿山生产开采阶段关键数据转换操作

序号	主要矿山类型	关键数据转换操作	应用场景
1	煤矿	瓦斯浓度时空聚合（按采掘工作面、小时粒度进行聚合）、通风网络阻力计算等。	基于煤矿通风系统拓扑结构及瓦斯监测数据实现动态分析，服务于瓦斯超限预警与通风系统优化。
2	金属非金属矿山(地下)	矿石品位加权平均计算、采矿贫化率核算等	结合出矿数据与地质数据按开采块段进行组合运算。
3	金属非金属矿山(露天)	剥离量—开采量动态平衡计算、采剥进度时序统计等。	按开采周期对剥离量、开采量数据进行整合分析，支撑露天开采规划制定。

6.1.2.3 安全监控阶段

安全监控阶段关键数据转换操作主要包括：

矿山文本数据结构化转换：对矿山隐患排查报告、故障报告等非结构化文本进行解析，

通过实体提取、关系梳理，转化为结构化隐患清单、风险标签等数据，支撑安全隐患管理。

表 3 各类型矿山安全监控阶段关键数据转换操作

序号	主要矿山类型	关键数据转换操作	应用场景
1	煤矿	瓦斯突出风险指数建模、煤尘浓度聚合分析等。	融合瓦斯浓度、通风量、煤层透气性等多维度数据构建评估模型，支撑瓦斯突出预警与煤尘防治。
2	金属非金属矿山(地下)	边坡位移时序特征提取、岩体稳定性评分等。	对边坡位移监测数据进行趋势分析与特征挖掘，实现边坡失稳预测。
3	金属非金属矿山(露天)	排土场沉降量聚合分析、边坡角合规性校验等。	按排土场分区对沉降数据进行整合，结合行业边坡角标准进行合规性判定，保障排土场安全管控。

6.1.2.4 设备运维阶段

设备运维阶段关键数据转换操作主要包括：

a) 矿山智能算法计算

应用机器学习等智能算法，对采煤机等矿山设备运行数据进行特征提取、异常检测，生成设备故障预测结果、运行异常标识等结构化分析结果，辅助设备故障预判。

b) 设备图像、点云特征提取

对输送带激光点云、设备图像等非结构化空间数据进行处理，提取几何特征、缺陷特征等语义特征，生成可用于设备磨损程度检测、故障定位的结构化数据。

表 4 各类型矿山设备运维阶段关键数据转换操作

序号	主要矿山类型	关键数据转换操作	应用场景
1	煤矿	采煤机振动频率特征提取、液压系统压力异常检测等。	对采煤机运行参数进行时序分析与特征提取，实现采煤机故障预测。
2	金属非金属矿	提升设备运行参数趋势分析、制动系统可靠性评分等。	整合提升设备运行数据进行趋势研判，支撑提升系统运维决策。

6.1.2.5 综合管理阶段

通用关键数据转换操作包括多模态数据合并，基于位置坐标、时间戳等业务主键，整合设备运行数据、生产数据、安全监控数据等不同类型数据，支撑综合生产报表编制。

表 5 各类型矿山综合管理阶段关键数据转换操作

序号	主要矿山类型	关键数据转换操作	应用场景
1	煤矿	安全风险分级分类模型推理等。	融合瓦斯、通风、采掘等多维度数据构建分级模型,实现安全管控优先级判定。
2	金属非金属矿	资源回收率动态核算模型推理等。	结合开采量、资源储量等数据进行实时核算,支撑资源利用效率评估。

6.1.3 评估操作对数据的影响

分析矿山数据处理过程的复杂度和对输入数据的改变程度,判断是否通过计算、聚合、关联、分析等方式生成了新数据,以及新数据是否是矿山行业特有的、在信息内容或结构上不同于输入数据,而非简单的数据格式转换或筛选。

6.2 数据依赖关系分析法认定要点

6.2.1 获取和构建数据血缘关系

利用现有数据治理工具、日志分析、代码解析或人工分析等方式,获取或构建认定对象的数据血缘信息。矿山行业血缘关系构建主要包括:

a) 构建矿山行业时空维度血缘:利用矿山数据系统的时空血缘工具,追溯认定对象的时空依赖;

b) 构建矿山业务维度血缘:按矿山业务域(安全、生产、环保等)拆解数据依赖关系,并关联矿山行业专属业务规则。

6.2.2 识别直接输入数据

沿数据血缘链条向上追溯,识别直接生成认定对象的上游矿山行业数据集或数据项。

6.2.3 追溯源数据来源

从直接输入数据开始,持续向上追溯,宜识别数据处理链条中的关键中间数据,并最终追溯到符合矿山行业标准的源数据或其他已知数据来源。

6.2.4 分析依赖关系性质

《总则》7.2.3.4 列举的三类依赖关系及以下依赖关系适用于矿山行业:

a) 时空依赖:认定对象的时空范围由输入数据决定,且需符合矿山时空管理规则;

b) 行业规则依赖:认定对象的生成必须依赖符合矿山行业规范的特定数据源,否则无业务意义。

6.3 数据特征比对法认定要点

6.3.1 确定比对维度

根据矿山大数据特性和预期的数据处理效果，选择合适的比对维度。其中，数据类型维度详见《总则》7.3.3.1，时空维度、业务维度、安全维度如下：

a) 时空维度：应对比时间粒度、空间粒度、时空连续性等；

b) 业务维度：应对比生产指标、设备指标、行业规则匹配度（认定对象是否符合矿山业务逻辑）等；

c) 安全维度：应对比安全指标范围（认定对象是否在输入数据的合理推导范围内）、敏感数据变化、灾害关联度（认定对象与输入数据的关联强度）。

6.3.2 执行数据特征比对

应使用合适的工具、编程脚本或统计分析方法，在选定的维度上对认定对象和潜在输入数据集样本进行定量或定性比对。

6.3.3 分析并量化差异

详细分析比对结果，识别输入数据集和认定数据集之间的差异，并尽可能对差异进行量化描述。

6.3.4 结合处理过程或依赖关系分析判断

将数据特征比对结果与已知的数据处理过程分析结果或数据依赖关系分析结果相结合。判断观察到的显著差异是否与已知的加工、分析、整合等处理活动一致。

6.4 人工复核与确认认定要点

6.4.1 组织复核小组

组建由熟悉数据治理、业务流程、相关技术、法律合规等领域的专家组成的复核小组。矿山行业一般场景（如普通生产数据统计、非安全类经营数据认定等）组建3人专家组复核；高敏感场景（如矿山安全监控数据衍生认定、重大资源储量评估数据认定、涉及个人信息或国家重要数据的衍生数据认定、影响重大生产决策的衍生数据认定等）组建5人专家组复核。

6.4.2 审查判定依据

人工复核小组应全面审查前期方法（基于数据来源关系、数据处理过程、数据特征比对的判定）得出的初步判定结论及其所依据的所有证据、分析报告和原始信息。

6.4.3 讨论与综合判断

复核小组应重点关注矿山行业复杂场景和判定依据存在争议的情况，结合矿山行业业务

语境、技术实现细节、法律合规要求和对矿山衍生数据定义的深刻理解，进行综合判断；

必要时，应要求数据处理人员、矿山业务人员提供进一步解释、数据样本或系统演示。

6.4.4 形成最终结论

若复核小组通过协商或表决机制形成一致结论，该结论为最终认定结论；若存在分歧且无法协商一致，应邀请更多或更权威矿山行业专家和数据技术专家参与复核，直至形成最终结论。

6.4.5 详细记录复核过程

应详细记录人工复核的过程，包括参与人员审查的依据、讨论的主要观点、存在的分歧意见及其理由、最终结论及其确定的理由。

7 判定准则

7.1 认定方法的使用准则

矿山行业衍生数据使用准则如下：

a) 矿山行业衍生数据认定应采用初步判定和人工复核相结合的模式，初步判定结果应经过人工复核小组复核与确认后，形成最终结论；

b) 初步判定应遵循本文件 4.5 处理过程优先原则，当因认定条件受限导致数据处理过程分析法无法形成有效初步判定结论时，可选用数据依赖关系分析法、数据特征比对法单独或组合开展初步判定。

7.2 基于数据处理过程的判定准则

基于数据处理过程的判定准则如下：

a) 若数据处理过程分析发现认定对象的生成依赖于矿山输入数据，且过程中执行了本文件 6.1.2 中所列的任何一类矿山关键数据转换操作，则将认定对象判定为矿山衍生数据；

b) 若矿山数据处理过程仅涉及数据格式转换、编码转换、简单的字段更名、数据物理传输、数据复制或基于简单条件的数据筛选，未执行 6.1.2 中所列关键数据转换操作，则不应仅凭此认定为矿山衍生数据，但应结合其他方法（如数据依赖关系分析、数据特征比对）进行综合判断；

c) 若矿山数据处理过程中存在多个连续的处理步骤，且至少包含一个 6.1.2 中所列关键数据转换操作，则整体生成的认定对象判定为矿山衍生数据，即使部分步骤为简单处理操作。

7.3 基于数据来源关系的判定准则

基于数据来源关系的判定准则如下：

a) 若数据来源关系分析证明认定对象在内容或结构上明确来源于一个或多个矿山输入数据，且该生成过程涉及本文件 6.1.2 中所列的关键数据转换操作，则初步判定认定对象为矿山衍生数据；

b) 若认定对象仅是对矿山输入数据的物理复制、传输或简单的基于输入数据的子集筛选，则不应仅凭此判定为矿山衍生数据，可结合其他方法（如数据处理过程分析、数据特征比对）进行综合判断或由人工复核小组进行综合研判；

c) 若无法追溯到认定对象明确的输入数据来源，或者其唯一来源是自身（例如矿山生产日志数据，无其他输入数据），则不应判定为矿山衍生数据（应视为矿山源数据），除非有其他强有力证据（如处理过程分析显示其是通过复杂矿山计算或模型生成）。例外条款：若认定对象为多源间接衍生数据，且部分输入数据可追溯，同时处理过程包含本文件 6.1.2 所列关键数据转换操作，即使部分输入数据来源无法完全追溯，仍可判定为矿山衍生数据，但需在认定报告中详细说明未追溯数据的情况及对认定结论的影响程度。

7.4 基于数据特征比对的判定准则

基于数据特征比对的判定准则如下：

a) 若数据特征比对显示认定对象与主要矿山输入数据在以下任一维度存在矿山显著差异，且该差异能够通过 6.1.2 识别的关键处理操作或 7.2 追溯到的复杂依赖关系得到合理解释，则应支持认定为矿山衍生数据：

——矿山数据结构发生重大变化；

——矿山数据粒度发生重大变化；

——矿山核心内容类型发生重大变化；

——矿山关键统计特性（如均值、方差、分布形态）或数据分布与输入数据存在显著差异，且非简单筛选或抽样所致；

b) 若比对发现仅存在矿山数据格式、编码、字段名称等表层变化，或认定对象仅为输入数据的简单子集，且无 7.1 所述的矿山关键处理操作支持，则不应仅凭此判定为矿山衍生数据，但应结合其他方法（如数据处理过程分析、数据依赖关系分析）进行综合判断或由人工复核小组进行综合研判；

c) 矿山显著差异的量化判定应优先采用信息熵量化分析法，当相对熵超过预设阈值（推

荐阈值 1.2) 或信息熵 (推荐阈值 40%) 变化率超过预设阈值时, 应支持认定为矿山衍生数据。上述阈值的确定依据: 基于矿山行业 5 类典型数据 (地质勘探数据、设备运行数据、安全监控数据、生产统计数据、环境监测数据) 的样本统计分析, 结合 20 位矿山行业专家与数据技术专家的论证结果; 阈值调整机制: 针对特殊矿山类型或特定数据场景 (如稀疏地质数据、高频传感器数据), 可根据实际数据特征在 $\pm 20\%$ 范围内动态调整, 调整需经 3 名及以上行业专家签字确认并记录调整理由。

7.5 基于人工复核的判定准则

基于人工复核的判定准则如下:

a) 人工复核小组应全面审查前期方法 (基于数据来源关系、数据处理过程、数据特征对比的判定) 得出的初步判定结论及其所依据的所有证据、分析报告和原始信息;

b) 复核小组应重点关注矿山行业复杂场景和判定依据存在争议的情况, 结合矿山行业业务语境、技术实现细节、法律合规要求和对矿山衍生数据定义的深刻理解, 进行综合判断;

c) 若复核小组通过协商或表决机制形成一致结论, 该结论为最终认定结论; 若存在分歧且无法协商一致, 应邀请更多或更权威矿山行业专家和数据技术专家参与复核, 直至形成最终结论。

8 记录与报告

8.1 认定记录

结合矿山行业安全优先、时空强关联、多源异构、业务场景绑定的核心特性, 构建“全链路可追溯、行业属性突出、安全合规贯穿”的记录与报告体系, 重点覆盖矿山安全敏感数据、生产运营数据、跨矿协同数据等认定场景。

8.1.1 认定记录内容

认定记录应包含矿山衍生数据认定的全过程信息, 包括但不限于:

a) 认定任务基本信息: 认定任务标识 (编码规则: MS-YYYY-YW-XZ-LSH, 其中“MS”为矿山行业统一代码, “YYYY”为认定年份, “YW”为业务领域代码 (地质勘探 DZ、设备管理 SB、安全监控 AQ、生产运营 SC、经营管理 JY), “XZ”为认定类型代码 (常规认定 CG、高敏感认定 MG), “LSH”为 3 位流水号, 示例: MS-2025-DZ-CG-001)、任务名称 (明确指向矿山具体应用)、认定对象标识 (详见《总则》10.2)、认定启动日期、认定完成日期、认定执行者等;

b) 源数据信息：按照《总则》10.4 要求明确源数据的标识符或描述，各类型矿山源数据如下：

表 6 各类型矿山核心源数据

序号	主要矿山类型	核心源数据
1	煤矿	瓦斯浓度、通风量、煤层厚度、采掘进度、煤尘浓度等。
2	金属非金属矿 (地下)	矿石品位、边坡位移、提升设备负载、爆破振动速度、采矿贫化率等。
3	金属非金属矿 (露天)	剥离量、采剥比、排土场边坡角、设备调度效率、边坡沉降量等。

c) 处理过程信息：按照《总则》10.5 要求描述生成认定对象的关键数据处理过程和逻辑；

表 7 各类初步判定方法的描述信息

序号	初步判定方法	描述信息
1	数据处理过程分析法	矿山数据处理流程图、处理脚本及代码、关键数据转换操作识别结果、处理过程分析报告等。
2	数据依赖关系分析法	数据血缘关系、时空依赖完整性、数据依赖关系分析报告等。
3	数据特征比对法	比对维度、比对方法、比对结果、特征比对报告等。

d) 人工复核信息：复核小组成员、复核日期、复核讨论记录、复核结论；

e) 认定结论信息：最终认定结论、认定依据、采用的认定方法。

8.1.2 认定记录管理

a) 格式要求：认定记录应采用矿山行业结构化模板，按“任务信息—输入数据—处理过程—依赖关系—特征比对—人工复核—结论”分层归档，关键信息需单独标注，便于快速检索；

b) 介质与备份：优先采用矿山数据中台专用存储系统或符合行业安全标准的电子文档，支持与矿山现有数据治理平台对接；电子文档需进行版本控制，标注修改人、修改时间及修改原因，同时按矿山数据安全要求进行异地备份、离线存储；

c) 保存期限：遵循矿山行业数据管理规定，矿山数据生命周期分为采集、存储、处理、使用、归档、销毁 6 个阶段，其中安全敏感数据（如安全监控衍生数据、重大资源储量评估数据）生命周期为永久保存，普通生产运营数据生命周期为 10 年，经营管理数据生命周期为 5 年；认定记录保存至对应数据生命周期结束后不少于 5 年；

d) 查阅权限：建立矿山行业分级查阅机制，设置管理部门一级查阅权限，单位内部二级

查阅权限，外部单位查阅需经矿山安全监察部门及数据主管部门双重审批；

e) 审批流程规范：数据敏感等级为普通级（如经营数据）的，所有矿山类型均执行单位内部审批流程，参与审批主体为数据主管部门及业务部门，审批时限为 3 个工作日。数据敏感等级为敏感级（如生产数据）的，煤矿执行二级审批流程，参与审批主体为单位数据主管部门及矿山安全管理部门，审批时限为 5 个工作日；金属非金属矿山执行二级审批流程，参与审批主体为单位数据主管部门及行业主管部门，审批时限为 5 个工作日。数据敏感等级为高敏感级（如安全监控数据、重大资源储量评估数据）的，煤矿执行三级审批流程，参与审批主体为单位数据主管部门、矿山安全监察局及省级行业协会，审批时限为 7 个工作日；金属非金属矿山执行三级审批流程，参与审批主体为单位数据主管部门、省级应急管理厅及行业专家委员会，审批时限为 7 个工作日。

8.2 认定报告

8.2.1 认定报告内容

认定报告应基于认定记录，以规范、简洁、易于理解的方式呈现矿山衍生数据认定结果，报告应包含以下内容：

- a) 报告封面：报告名称、认定任务标识、认定对象标识、认定单位、报告出具日期；
- b) 目录：列出报告各章节的标题和页码；
- c) 引言：简要介绍认定背景、认定目的、认定范围；
- d) 认定过程概述：简要描述认定所采用的方法、认定人员组成、认定时间安排；
- e) 认定对象详情：详细描述认定对象的基本信息、数据结构、数据量；
- f) 输入数据详情：详细描述输入数据的基本信息与认定对象的关联关系，确认数据处理者享有数据使用权；
- g) 认定分析过程：详细阐述基于数据来源关系、数据处理过程、数据特征比对的分析过程和结果，引用相关证据，数据处理过程中对各方权益的保护机制；
- h) 人工复核过程：详细描述复核小组组成、复核讨论内容、复核结论及理由；
- i) 认定结论：明确给出最终认定结论，并说明认定依据；
- j) 附件：包含认定记录和相关支撑材料。

8.2.2 认定报告管理

认定报告应加盖认定单位公章，报告出具后应及时分发至相关使用方。若认定结论发生

变更，应重新出具认定报告，并在报告中说明变更原因、变更内容和变更依据，同时收回原报告或在原报告上标注变更说明。

附录A
(资料性)
认定案例

A.1 案例一：矿山资源储量评估数据认定

A.1.1 场景概述

某矿山企业基于地质勘探的钻孔坐标、矿石品位、矿体厚度等源数据，通过资源储量地质块段法计算模型，生成不同开采块段的资源量、储量、平均品位等指标矿山资源储量评估数据，需认定该资源储量评估数据是否为矿山衍生数据。

A.1.2 认定过程

A.1.2.1 数据来源关系分析

追溯认定对象（资源储量评估数据）的输入数据，确认其来源于地质勘探源数据，无其他外部数据源；

输入数据属于矿山源数据，与矿山资源储量计算业务直接相关，不存在无关数据混入。

A.1.2.2 数据处理过程分析

数据处理过程分析如下：

- a) 获取资源储量计算模型的处理脚本和计算逻辑文档，识别处理过程中的关键数据转换操作；
- b) 对地质勘探源数据进行分组聚合（按开采块段分组），计算每个块段的矿石平均品位、矿体平均厚度等统计量；
- c) 应用地质块段法数学公式，结合矿山行业标准参数，计算每个块段的资源量和储量；
- d) 处理过程包含本文件 5.2.1 中所列的“多源数据整合”和“矿山专业指标计算”两类关键数据转换操作。

A.1.2.3 数据特征比对分析

数据特征比对分析如下：

- a) 结构比对：原始地质勘探数据为单条钻孔记录（字段包括钻孔 ID、坐标、品位、厚度），资源储量评估数据为开采块段汇总记录（字段包括块段 ID、资源量、储量、平均品位），结构发生根本性改变；
- b) 粒度比对：源数据粒度为单钻孔，评估数据粒度为开采块段（一个块段包含多个钻孔），粒度发生根本性改变；
- c) 量化比对：计算源数据品位分布与评估数据平均品位分布的相对熵，结果为 1.8（超过推荐阈值 1.2），信息熵变化率为 55%（超过推荐阈值 40%），存在矿山显著差异。

A.1.2.4 人工复核

复核小组由地质工程师、数据分析师组成，具体复核工作如下：

- a) 地质工程师确认资源储量计算模型符合矿山行业标准，关键数据转换操作合理；
- b) 数据分析师确认数据处理过程分析和特征比对结果准确，无逻辑漏洞；
- c) 复核小组一致认定资源储量评估数据为矿山衍生数据。

A.1.3 认定结论

基于数据来源关系分析、数据处理过程分析、数据特征比对分析和人工复核，认定该矿山资源储量评估数据为矿山衍生数据。

A.2 案例二：矿山设备故障预测数据认定

A.2.1 场景概述

某矿山企业基于采煤机原始运行数据（包括转速、电流、温度、振动等实时参数），通过设备故障预测模型（基于机器学习算法），生成了采煤机故障预测数据（包括故障类型、故障概率、预测时间窗口等），需认定该故障预测数据是否为矿山衍生数据。

A.2.2 认定过程

A.2.2.1 数据来源关系分析

追溯认定对象（故障预测数据）的输入数据，确认其来源于采煤机原始运行数据，同时参考了设备历史故障数据（已认定为矿山衍生数据）：

输入数据属于矿山源数据和已认定的矿山衍生数据，与采煤机故障预测业务直接相关。

A.2.2.2 数据处理过程分析

数据处理过程如下：

- a) 获取设备故障预测模型的算法代码和模型定义文件，识别处理过程中的关键数据转换操作：
- b) 对采煤机原始运行数据进行时间序列重采样（将每秒数据重采样为分钟平均值），提取振动频率、电流波动等复杂特征；
- c) 将提取的特征与设备历史故障数据输入机器学习模型（随机森林算法），训练模型并生成故障预测结果（故障类型、概率）；
- d) 处理过程包含本文件 5.2.1 中所列的“矿山时间序列数据的重采样、复杂特征提取”和“应用矿山机器学习模型、统计模型生成预测值、评分或分类标签”两类关键数据转换操作。

A.2.2.3 数据特征比对分析

数据特征比对分析结果如下：

- a) 结构比对：原始采煤机运行数据为实时参数记录（字段包括时间戳、转速、电流、温度、振动），故障预测数据为故障预测结果（字段包括设备 ID、故障类型、故障概率、预测时间窗口），结构发生根本性改变；
- b) 内容比对：源数据为设备运行参数数值，故障预测数据为故障相关的预测信息（源数据中无故障类型、概率等信息），内容发生根本性改变；
- c) 量化比对：计算源数据振动参数分布与故障预测数据故障概率分布的相对熵，结果为 2.3（超过推荐阈值 1.2），信息熵变化率为 68%（超过推荐阈值 40%），存在矿山显著差异。

A.2.2.4 人工复核

复核小组由矿山设备工程师、机器学习专家、合规专家组成，复核过程如下：

- a) 矿山设备工程师确认故障预测模型的输入特征（振动频率、电流波动）符合采煤机故障诊断的行业知识，预测结果具有实际业务意义；
- b) 机器学习专家确认模型算法逻辑合理，训练数据来源可靠，预测过程符合机器学习规范；
- c) 合规专家确认数据处理过程符合矿山数据隐私保护要求，无数据泄露风险；
- d) 复核小组一致认定采煤机故障预测数据为矿山衍生数据。

A.2.3 认定结论

基于数据来源关系分析、数据处理过程分析、数据特征比对分析和人工复核，认定该矿山采煤机故障预测数据为矿山衍生数据。

参 考 文 献

- [1] GB/T 43697—2024 数据安全技术 数据分类分级规则
-