

《衍生数据认定方法 第1部分 总则》团体标准 编制说明

一、任务来源

2023年2月，中共中央、国务院印发《数字中国建设整体布局规划》，明确将数字中国建设作为推进中国式现代化的重要引擎，并指出要夯实数字中国建设基础，必须畅通数据资源大循环。同年12月，财政部发布《关于加强数据资产管理的指导意见》，要求有序推进数据资产化，并鼓励行业根据发展需要，自行或联合制定数据资产相关标准。这些战略部署对数据要素治理的规范化、标准化提出了紧迫要求。

北京市作为全国数字经济发展的“排头兵”，率先出台了一系列纲领性政策法规。例如，《北京市数字经济促进条例》规定，“市经济和信息化部门会同市场监管等有关部门推进数字经济地方标准体系建设，建立健全关键技术、数据治理和安全合规、公共数据管理等领域的地方标准”；又如，《关于更好发挥数据要素作用进一步加快发展数字经济的实施意见》明确提出，要“探索建立结构性分置的数据产权制度”，推动“界定数据来源、持有、加工、流通、使用过程中各参与方的合法权利”。

要实现数据产权的“结构性分置”，首先必须能科学、客观地区分不同形态的数据。特别是经过“加工”环节产生的“衍生数据”，是“数据加工使用权”和“数据产品经营权”等新型权利的核心载体。若缺乏一套公认的认识标准，产权分置将沦为空谈。

基于以上需求，北京市政数局在国家数据要素综合试验区（北京）2025年试验任务书中，设立了《衍生数据认定标准试验任务》，由北京联通、智慧足迹作为任务承担方，同时与北京市大数据中心、百度、电子四院、海天瑞声、认知洞察、煤科总院、北京健康有益科技有限公司、滴滴出行等衍生数据研发专业团体组织共同形成了专项工作组，并形成了一套1+N的《衍生数据认定方法标准》整体标准制定与实践体系。

其中，北京联通、智慧足迹承担总则与《通信行业衍生数据认定标准》的制

定；认知洞察承担《情感认知行业衍生数据认定标准》；煤科总院承担《矿山行业衍生数据认定标准》、北京健康有益公司承担《健康医疗行业衍生数据认定标准》等。

二、制定规范的必要性和意义

（一）填补国内外标准体系空白，抢占数据要素基础规则制定高地

当前，国内外现有数据标准体系虽然已在数据治理、数据质量管理及数据资产管理等方面取得阶段性成果，但仍存在一个普遍且关键的盲区。即，现有标准大多建立在“数据对象已明确”的前提之上，主要侧重于对已知数据对象的“过程管理”与“状态评价”。然而，在数据要素从原始形态向具备新价值的产品转化的关键阶段，最基础的问题——“它是什么？”，即如何对其进行“身份认定”，却始终缺乏一个统一、科学的标准。

为填补上述空白领域，本标准首次系统性地构建了“衍生数据”的认定原则、流程与方法。本标准具有以下三方面独特价值：

基础性：它位于数据价值链的前端，是后续所有数据资产管理、质量评价、价值评估和合规流通活动的基础。

前瞻性：欧盟《数据法案》（Data Act）等国际前沿法规已开始关注数据的“共同生成”和再利用，本标准的研究制定使北京市在这一关键基础规则领域走在了国际前列。

协同性：本标准并非孤立存在，而是对现有国家、地方数据相关系列标准的有力补充，构成了数据标准体系中承上启下的关键一环。

（二）为数据产权界定、资产入表、交易流通等核心环节提供操作性支撑

本标准将为北京市数据要素市场化配置改革中的三大核心难题提供基础性、可操作的解决方案。

一是为数据产权的“三权分置”提供技术标尺：本标准提出的“数据处理过程分析法”，为判断“加工”行为是否达到“实质性改变”、从而创设新的数据产品经营权提供了客观、可量化的技术标尺。

二是为数据资产“入表”提供可信依据：本标准通过规范化的认定流程与统一的结果表述要求，确保了衍生数据资产的生成过程可追溯、可验证、可审计，为会计师事务所与资产评估机构采信数据资产提供了可靠依据。

三是为数据交易流通提供“标准化商品说明书”：本标准实际上为“衍生数据”这一核心数据产品提供了一份详尽的“商品说明书”，有助于降低交易成本，减少交易纠纷，建立市场信任，促进数据要素安全、合规、高效地流通。

（三）为数据要素市场的健康有序发展和数据价值的充分释放提供根本保障

制定《衍生数据认定方法标准》，是一项具有全局性、基础性和紧迫性的工作。一方面，能够明晰规则，防范风险。通过为衍生数据提供清晰的“身份证”，有助于明确数据处理活动中各方的权利、义务和责任，从源头上规范数据处理行为，防范数据滥用和安全风险。另一方面，能够激发创新，释放价值。明确的规则是市场活力的催化剂。通过本标准，企业能够清晰地预见其数据加工投入可以形成受保护、可交易的合法资产，这将极大激发企业投资于数据技术、开发数据产品的积极性，为千行百业的数字化转型注入强大动能。

三、主要编制过程

1. 第一阶段，成立标准研制起草工作组（以下简称工作组），编制项目立项材料和标准草案。2025年11月，北京市大数据中心牵头成立标准编制工作组，开始标准起草工作。工作组通过查阅文献，形成理论框架和工作方法，并对已有相关内容基础进行了梳理，多轮研讨沟通，形成立项申请材料和标准草案。工作组内部多次召开立项准备沟通会，为项目立项制定完整方案。

2. 第二阶段，项目立项评审。工作组于2025年12月在中关村知识产权大厦A座，北京软件和信息服务业协会会议室召开了项目立项专家评审会。评审专家包括：首经贸数据要素市场化专题团队带头人、财税学院副院长陈蕾、清华大学软件学院研发主任敖然、京东集团公共业务副总裁孔鑫、北京大成律师事务所高级合伙人陈福、百度资深标准化专家杨庄媛。标准立项通过专家评审后，信息在全国团体标准信息平台和协会官方网站和北京软协官网发布。

3. 第三阶段，编写规范征求意见稿。2026年1月-7月，围绕标准草案及各方意见，工作组开展了进一步调查研究，通过多轮内部研讨和标准内容文件修订，形成征求意见稿。

四、制定规范的原则和依据，与现行法律、法规、标准的关系

本标准依据《中华人民共和国标准化法》及其《实施细则》，并参照《标准

化工作导则第 1 部分：标准的结构和编写》GB/T 1.1-2020 编制，严格遵守以下五个原则。

科学性原则：标准的编制必须建立在科学的基础之上，确保标准的制定符合基于数据科学和信息技术等学科实际，在衍生数据的判定方面，应遵循科学的方法和规范。

实用性原则：确保标准能够满足实际工作需要，易于理解和操作，具有实用性。同时，标准应易于理解和实施，方便各相关方使用和推广。

动态性原则：充分考虑数据技术、信息技术的发展趋势，以及相关技术的最新进展，标准应具有一定的灵活性，以适应技术和市场的变化。

统一性原则：应追求统一性和协调性，避免不同标准之间的冲突和矛盾。在编制过程中，将充分参考国内外相关的评估标准和最佳实践，确保之间的衔接和互操作性。

开放性原则：编制工作坚持开放、透明的原则，广泛征求各方意见和建议，确保标准的制定过程公开、公正、公平。

五、主要条款的说明，主要技术指标、参数、试验验证的论述

本标准的核心技术内容构建了一个完整的“原则-流程-方法”体系。

（一）四大认定原则：

1. 客观性原则：认定过程必须基于事实和证据，不应受主观判断或外部因素影响；

2. 可追溯性原则：认定结论应能够追溯到所依据的输入数据、完整的处理过程、采用的认定方法及其结果；

3. 全面性原则：在界定的认定范围内，认定工作应尽可能覆盖数据处理的完整链条，分析关键环节，不应遗漏可能产生衍生数据的处理操作；

4. 审慎性原则：对于复杂或难以确定的情况，宜组合使用多种认定工作方法，并必须进行人工复核，确保认定的准确性和可靠性；

（二）七步标准化认定流程

1. 认定启动与计划：根据数据资产管理、安全合规、业务需求等触发条件，明确界定认定对象标识、认定范围和任务计划。

2. 信息获取与梳理：应系统地收集与认定对象生成过程相关的所有必要信息。应包括认定对象的元数据、描述其生成逻辑的数据处理脚本或代码、系统配置信息、ETL 作业定义、数据流图、数据血缘信息(如有)、相关技术文档、业务规则说明以及必要时对数据处理人员和业务专家的访谈记录。

3. 核心判定分析：基于收集到的信息，识别关键数据转换操作和数据依赖关系，并可根据需要比对输入输出数据特征，应用规则或模型进行辅助分析，以形成对认定对象是否为衍生数据的初步判断。

4. 人工复核与最终确认：判定结果不确定、存在争议或涉及重要/敏感数据时，必须组织由具备相关专业知识的专家组成复核小组。复核小组必须审阅分析报告、初步结论及其依据，进行讨论和综合判断。必要时，应要求数据处理人员提供进一步解释或演示。

5. 认定结果记录与归档：应以结构化、规范化的方式详细记录整个认定过程的关键信息和最终认定结论。已完成的认定结果应妥善存储和管理。

6. 认定结果应用与管理：认定结果应作为数据治理活动的基础。宜进行分类分级管理，并基于分级结果，将其应用于后续的数据资产目录更新、差异化安全策略制定、数据流通权限管理等活动中。应建立认定结果的版本管理机制，定期或在认定对象的数据处理逻辑、输入数据发生重大变化时对认定结果进行复核和更新。

7. 争议处理：当认定结论存在争议时，任何异议方可在规定时限内提出申请，由指定争议处理机构组织独立专家小组，通过材料复核、调解及必要时的专家仲裁等多级程序予以处理，最终形成具备约束力的裁决结果，并完整归档执行。

（三）四种核心认定方法

1. 数据处理过程分析法：通过分析代码、脚本，识别是否发生了关键数据转换操作（如复杂关联、分组聚合、新指标计算等）。

2. 数据依赖关系分析法：通过追溯数据血缘，确定认定对象是否来源于其

他数据。

3. 数据特征比对法：通过比对输入与输出数据在结构、粒度、统计特性等方面的显著差异进行判定。

4. 人工复核与确认法：针对复杂、重要或存在争议的场景，引入多领域专家进行综合研判，确保最终结论的权威性和公正性。

（四）新兴技术场景指引

针对 AI 生成内容（AIGC）、联邦学习等前沿技术场景，提出了专门的认定原则和分析要点，确保标准的适用性与前瞻性。

六、重大意见分歧的处理依据和结果的说明

本规范研究制定过程中未出现重大分歧。

七、采用国际标准或国外先进标准程度的说明，以及国内、外同类产品或标准的对比情况

（一）参考和引用标准的标准号和标准名称

GB/T 35295—2017 信息技术大数据术语

GB/T 36344—2018 信息技术数据质量评价指标

GB/T 38667—2020 信息技术大数据数据分类指南

（二）与国内外相关标准的内容异同

本标准在数据要素标准体系中具有独特的定位，它不重复现有标准，而是填补了关键空白，为其他标准的落地应用提供了必要支撑。

对比维度	本标准《衍生数据认定方法》	现有数据治理/资产/质量标准
核心目标	身份认定与界定：提供科学的原则、流程和方法，用于识别和确认衍生数据。	过程优化与状态评估：建立管理框架、评估能力、评价质量，以提升数据管理水平。
关注焦点	数据的“生成与演变”过程：聚焦从原始数据到新数据的转换过程、加工方式和血缘关系。	数据的“管理与应用”状态：聚焦数据在组织内的存储、流动、使用、安全等状态。
关系定位	前置基础与底层支撑：衍生数据的认定是开展数据资产管理、质量评价等活动的前提。	后置应用与管理框架：在数据形态被明确后，应用这些标准进行有效的管理和使用。

八、实施标准的措施建议

按照标准要求,推动企业构建、优化、评估及验证《衍生数据认定方法标准》。

九、其他说明

无。

《衍生数据认定方法 第1部分 总则》起草工作组

2026年7月8日