

北京资产评估协会专业技术委员会 执业问题解答（数据资产专家问答第二期）

特别提示

北京资产评估协会专业技术委员会组织行业内资深专家，针对数据资产评估实践中的核心难点进行了深入的研究与实务探讨，并据此形成本期专业观点，旨在为资产评估机构及其资产评估专业人员在执行相关业务时提供技术参考。需明确指出的是，本提示内容不具备法律法规或资产评估准则的强制性效力，亦无法替代资产评估专业人员在具体执业过程中基于独立客观原则所形成的职业判断。资产评估机构及其资产评估专业人员在开展具体评估项目时，应当紧密结合项目实际背景、数据资产特性及特定经济行为目的，在充分理解并遵循相关法律法规与准则要求的前提下，对本提示内容进行审慎参考与适用，不能在执业过程中机械化地照搬照抄，以免引致不必要的执业风险。

随着数字经济的发展，数据已逐步演变为推动经济社会实现高质量发展的核心引擎。数据资产作为一种新型生产要素，正日益成为企业驱动创新发展、构建核心竞争优势的关键资源。在此宏观背景下，企业在融资担保、并购重组、资产入表及财务报告披露等多种经济行为场景中，对数据资产进行科学、合理、客观价值评估的需求呈现出增长趋势。因数据资产具有的非实体性、依托性、可复制性及时效性等特征，这使得其与传统的实物资产在物理形态及价

值形成机制上存在差异。传统资产评估中常用的实物盘点、成新率勘察等核查手段难以直接适用于数据资产，评估实践中普遍面临权利边界模糊、数据质量难以量化、价值贡献拆分缺乏标准等诸多困惑与技术难点。特别是在数据资产评估对象范围的界定、合规性盘点程序的实施以及全链路血缘分析的开展等方面，表现出较强的行业特殊性与复杂性。

本期解答聚焦于数据资产评估实务中的四个核心环节，即如何规范化描述数据资产评估对象与范围、如何实施有效的数据资产盘点（监盘）、为何开展数据资产血缘分析以及如何实施血缘分析，为资产评估专业人员提供具有操作性的实务指引。

问题 1：在进行数据资产评估时，如何描述评估对象与评估范围？

在执行数据资产评估业务时，资产评估专业人员应当遵循《数据资产评估指导意见》及相关执业准则的规范要求。针对资产评估委托合同中约定的评估对象与评估范围，从信息属性、法律属性、价值属性等三个关键维度进行全面、客观地分析和描述。这种多维度的分析和描述，有助于厘清数据资产的物理边界与权利边界，为数据资产的评定估算工作奠定基础。

1. 数据资产的信息属性

信息属性主要包括数据名称、数据结构、数据字典、数据规模、数据周期、产生频率及存储方式等。

(1) 数据名称：数据资产名称一般具有高度概括性和良好的识别性，在实际应用中能够快速、准确地被理解和使用。通常情况下，

数据资产名称通常涵盖三个描述维度，即对象维度、功能维度以及数据类型维度。对象维度主要描述数据资产所涉及的主体或内容范围，用于明确数据资产所涵盖的具体实体类别和业务场景中的关键要素，例如用户、商品、订单、设备、门店、车辆等；功能维度明确数据资产的具体用途或作用，例如分析、监控、预测、推荐、风控、运营等；数据资产类型维度旨在说明数据资产的类别，主要包括：①数据类，如数据集、盒装数据、API 数据等；②信息类，如统计报表、数据可视化图表、仪表盘趋势分析报告等；③知识类，如用户画像、推荐系统、算法模型、行为标签等；④智慧类，如智能决策系统、自动化策略引擎、知识图谱、大模型等。

(2) 数据结构指的是数据元素之间的组织关系及其存储形式。常见类型有：①结构化数据，例如数据库表结构，它通过行列规范来存储关系型数据；②半结构化数据，例如 JSON 格式，以键值对形式灵活表达层级关系；③非结构化数据，例如文本和图像，需要通过元数据来定义组织逻辑等。不同的数据结构会直接影响数据资产的可复用性和分析价值。

(3) 数据字典是数据库系统或信息系统中用于描述数据结构、属性及其含义的元数据集合。它是一个集中存储“关于数据的数据”（metadata）的存储库，旨在帮助用户、开发人员和数据库管理系统（DBMS）理解、管理和使用数据。数据字典的内容涵盖以下方面：①基础标识信息，包含字段名称、所属表或实体、唯一标识符；②技术属性，包含数据类型、数据长度精度、数据格式；③约束规则，包含主键/外键、空值允许性、默认值、唯一性与检查约束等；④

业务属性，包含字段中文名、数据来源；⑤取值范围，包含枚举值、最值等；⑥治理属性，包含安全合规信息，如敏感数据标记、分级分类、GDPR 等合规标签；⑦元数据管理信息，包含创建/修改记录、版本号、关联业务流程。这些信息共同构成了一个完整的描述体系，确保了数据的一致性和可理解性。

(4) 数据规模是数据资产量化维度的核心指标之一，是对数据资产物理体量与增长态势的综合度量，直接影响数据存储成本、处理效率及价值的评估。它不仅包含基础的数据量度量，比如记录条数、文件大小（GB/TB/PB），还需涵盖数据维度，例如字段数量、时间跨度，以及增长速率，像日增/月增数据量、复合增长率等关键要素。

(5) 数据周期，即时间跨度，是数据资产时效性维度的核心指标，是对数据资产时间属性的综合度量，直接影响数据资产的价值衰减速率与应用场景适配性。它不仅包含数据所覆盖的时间范围，比如“2023 年度”，还需涵盖更新频率，例如“实时更新”“每日更新”，以及数据生成周期，也就是采集或生产的时间间隔，留存期限，即合规要求或业务需求的存储时长。

(6) 产生频率是指单位时间内数据被创建或更新的速率，反映了数据的活跃程度和时效性需求。根据业务场景不同，数据产生频率可分为：①实时：毫秒至秒级，如 IoT 传感器、交易系统；②高频：分钟至小时级，如 APP 埋点；③中频：日级，如用户行为日志汇总；④低频：周/月/年，如财务报表；⑤事件驱动型：由特定操作触发，如用户注册等。

(7) 存储方式作为数据资产可访问性维度的核心指标，是对数据资产存储形态和访问路径的综合衡量，直接影响数据资产的调用效率、安全风险和维护成本。它包括：①数据的存储形态和位置，如数据库、数据湖、API 接口；②存储架构，例如集中式或分布式存储；③存储介质，即 SSD、HDD 或云存储；④访问权限控制，如角色隔离或加密访问；⑤容灾备份策略，如多副本或异地容灾；⑥数据检索效率，即查询响应时间等要素。

2. 数据资产的法律属性

法律属性决定了数据资产财产性权利的边界，是数据资产评估过程中确定评估对象与范围权属边界的核心依据。法律属性主要包括授权主体信息、产权持有人信息，以及权利类型、权利路径、权利范围、权利期限、权利限制等权利信息。

(1) 权利类型：依据《中华人民共和国民法典》《中华人民共和国数据安全法》《中华人民共和国个人信息保护法》等相关法律法规，以及中共中央、国务院发布的《关于构建数据基础制度更好发挥数据要素作用的意见》（“数据二十条”）的政策精神，数据产权主要包括数据资源持有权、数据加工使用权以及数据产品经营权。国家数据局于 2025 年 1 月发布的《数据领域常用名词解释（第二批）》，对数据产权、数据持有权、数据使用权和数据经营权作出了名词解释。

①**数据产权**是指权利人对特定数据所享有的财产性权利，包括数据持有权、数据使用权、数据经营权等。②**数据持有权**是指权利人自行持有或委托他人代为持有合法获取的数据的权利，其目的在

于防范他人非法违规窃取、篡改、泄露或者破坏持有人持有的数据。③**数据使用权**是指权利人通过加工、聚合、分析等方式，将数据用于优化生产经营、提供社会服务、形成衍生数据等方面的权利。通常情况下，使用权是权利人在不对外提供数据的前提下，将数据用于内部使用的权利。④**数据经营权**是指权利人通过转让、许可、出资或者设立担保等有偿或无偿的方式对外提供数据的权利。不同的权利类型有着各自独特的价值实现路径以及相应的收益分配机制。评估人员在开展价值评估工作时，需对这些不同的权利类型进行清晰识别与划分，并在评估报告中详细阐述其具体特征和差异，以此确保评估结果的准确性与专业性。

（2）权利路径：数据资产的权利路径通常是指数据从初始产生、加工处理到流转使用全链路的权属归属与授权关系链条，涵盖三个环节：初始权属确认，即明确数据采集阶段的原始授权主体与合法范围；流转授权验证，也就是核查数据在内部部门或外部合作方间共享的授权协议、期限及用途限制；终端使用合规性审查，指验证数据最终应用场景是否符合初始授权约定。

（3）权利范围：数据资产的权利范围通常是指权利人对数据资产依法享有的合法使用边界，通常通过四个方面来表述，即使用权限，比如独占使用、共享使用、转授权使用；地域范围，例如境内使用、跨境使用；时间期限，比如授权有效期、数据生命周期；用途限制，例如内部分析、商业变现、研发支撑等。

（4）权利期限：数据资产的权利期限通常是指权利人对数据资产依法享有合法权利的有效存续周期，通常可以通过授权有效期，

即明确数据使用权的起止时间，比如 5 年授权期；数据生命周期，也就是数据从产生到销毁的合规留存时长；权利终止条件，比如授权到期未续期、数据主体撤回授权或数据销毁三个方面表述。

（5）权利限制：数据资产的权利限制通常是指对权利人使用数据资产的合法约束条件，通常可以从用途限制，比如禁止商业变现、仅限内部分析；地域限制，例如禁止跨境传输、仅限境内使用；转授权限制，比如禁止对外授权、仅限特定主体共享；合规限制，例如符合《中华人民共和国个人信息保护法》《中华人民共和国数据安全法》等法规要求等四个方面表述。

（特别提示：上述权利范围、权利期限和限制需要结合全链路授权内容的衔接性与具体应用场景相结合综合判断。）

3. 数据资产的价值属性

价值属性反映了数据资产创造经济利益的能力、模式及质量特征，是选择评估方法及设定参数的关键。价值属性主要包括数据覆盖地域、数据所属行业、数据成本信息、数据应用场景、数据质量、数据稀缺性及可替代性等。

（1）数据覆盖地域：数据资产的覆盖地域一般指的是数据采集、使用及流转所涉及的地理范围，通常从以下三个维度进行描述，①采集地域，也就是数据的初始来源地理区域；②使用地域，即数据合法应用的地理边界，需与权利范围中的地域限制保持一致；③合规地域，即符合当地数据法规要求的区域，若涉及跨境情况，需满足《数据出境安全评估办法》（国家互联网信息办公室令第11号；2022年7月7日公布，自2022年9月1日起施行，现行有效）。

在数据资产评估过程中，覆盖地域的合规性与广泛性会直接影响数据资产的价值。

（2）数据所属行业：数据资产的所属行业一般指的是与数据资产相关联的主营业务领域以及行业分类属性，通常可从以下三个维度进行描述：一是行业分类依据，例如国民经济行业分类标准、数字经济细分领域；二是行业合规要求，如金融行业需符合《金融数据安全治理实施指南》（由中国互联网金融协会组织编制，于2024年10月25日正式发布的标准），医疗行业需遵循医疗行业相关规范；三是行业特性影响，如高价值行业的数据资产具有更强的收益潜力。在数据资产评估中，所属行业的合规性和行业价值会直接影响价值测算；在评估过程中，需要关注所属行业是否属于高增长合规领域，如人工智能、数字金融，以及所属行业是否存在监管限制，如某些敏感行业等行业因素。

（3）数据成本信息：数据资产的成本信息一般是指数据从采集、加工、存储到传输整个链路的投入成本与合规成本之和，通常可从以下四个维度加以描述：一是采集成本，包含人力、工具及授权费用；二是加工成本，涵盖算力、软件及清洗标注费用；三是存储成本，包含硬件、维护及灾备费用；四是合规成本，涉及数据安全认证、隐私保护及审计费用。在数据资产评估中，成本信息的完整性与合理性将直接影响评估结果。评估过程中可重点关注成本缺失或隐性合规成本情况对价值的影响。

（4）数据应用场景：数据资产的应用场景一般是指数据资产能够产生价值的具体业务场景及使用方式，通常可以从以下三个维

度进行描述：一是业务场景，比如营销、风控决策、产品优化；二是用户场景，例如 C 端个性化推荐、B 端行业解决方案；三是合规场景，比如符合行业监管要求的应用模式。在数据资产评估中，应用场景的多样性与商业价值直接影响评估结果，需重点关注应用场景是否广泛且变现能力强，是否场景单一或存在合规限制等因素对价值的影响。

（5）数据质量：数据资产的数据质量一般是指数据满足业务需求以及合规要求的程度，通常可从以下五个维度来描述：①准确性，即数据与真实情况的相符程度；②一致性，即跨系统数据的逻辑统一性；③完整性，即数据字段的缺失比例及覆盖范围；④规范性，即符合隐私保护和行业监管要求；⑤时效性，即数据的更新频率和新鲜程度；⑥可访问性，即数据资产在需要时能被合法授权主体顺畅获取使用的特性。在数据资产评估中，数据质量的好坏直接影响评估结果；若各质量维度均达标且合规性较高，评估价值相对就高；若存在质量缺陷，例如关键信息缺失、时效性不足，则评估值相对较低。

（6）数据稀缺性及可替代性：数据资产的稀缺性及可替代性一般是指数据资产在市场中的稀缺程度及被其他数据替代的可能性；①稀缺性，即数据的独特性、获取难度及市场供给量；②可替代性，即同类数据的替代成本及替代效果。在数据资产评估中，稀缺性与可替代性的组合直接影响评估结果。需要注意的是，《数据资产评估指导意见》对不同评估方法中稀缺性与可替代性的处理存在差异：在市场法中，稀缺性的影响被纳入“供求调整系数”的考量范围；

而在成本法或收益法中，“价值调整系数”的确定方式主要为结合质量因素综合确定或结合剩余经济寿命确定，稀缺性与可替代性尚未被明确列为独立调整因素。

问题 2：如何对数据资产进行盘点及清查核实？

在数据资产评估业务中，现场调查是不可或缺的核心环节，数据资产的盘点工作由数据资产产权持有人完成，资产评估专业人员是在产权持有人完成资产清查盘点的基础上实施监盘。

1. 资产评估专业人员如何对数据资产进行清查核实

在数据资产清查核实过程中，需关注数据资源、数据资产持有权、数据加工使用权和数据产品经营权等权利状况。鉴于数据资产具有的虚拟性、无形性等特点，形成需经过数据源确认、加工、清洗等一系列过程，形成过程需要律师协助确认资产权利，并经有权部门（通常指各地数据局、数据交易所等）登记确权，取得确权证书，同时结合一定的应用场景，才可以带来收益，从而符合数据资产的特性。即已经历过从数据资源到数据资产的转化过程，因此在数据资源转化为资产进入到评估环节时，一般情况下拟评估的数据资产应具有有效的数据资产权利证书，以明晰其权利归属、证明其合规性，有明确的评估目的、评估对象及评估范围，具体的应用场景或产品使用范围，资产评估专业人员以此为基本前提基础展开评估。

（1）收集数据资产的产权资料 and 基本信息

资产评估专业人员应根据委估数据资产情况，收集数据资产的法律意见书、数据资产质量报告、数据资产权属登记证书等文件，

了解相关专业机构的资质、经验和能力，关注该部分文件的披露信息，引用数据资产产权人提供的专业报告（通常由有资格的第三方机构出具），核实评估对象和评估范围，确定委估数据资产是否具备评估的前提和基础。

（2）了解数据资产的存储状态和保管状态

在现场勘查过程中，资产评估专业人员应通过向管理人员访谈，了解数据资产的存储状态和存储地点。并收集与之对应的数据存储证明文件，包括但不限于云服务合同、服务器租赁合同，服务器购买合同等证明数据存储的文件。

（3）关注数据资产的权利内涵及其与评估对象及评估范围的一致性

资产评估专业人员执行数据资产清查核实程序时，需要明确评估目的所对应的数据资产权利内涵，分析数据资产权利类型，并关注评估对象相应权利是否与评估目的相一致；核实律师出具的法律意见书中对数据资产权利来源及权力行使是否符合相应的法律法规规定的表述，是否符合经济行为的用途。

（4）了解数据资产的应用场景和质量情况

资产评估专业人员通过对产权持有人专业管理人员的访谈，了解数据资产的应用场景及使用和运营情况。

资产评估专业人员需关注数据资产质量，取得第三方专业机构出具的数据质量评价专业报告或者其他形式的数据质量评价专业意见等加以引用。

（5）对数据资产进行抽查盘点和有效性确认

资产评估专业人员可取得数据汇总目录或数据统计情况资料，制定数据抽查盘点表，对抽盘的数据由产权持有人的专业管理人员带领，由管理人员逐条查验所抽查的数据，由资产评估专业人员进行记录和监盘，确认数据存储状态和存在性，以及与应用场景下的有效性。

在完成上述抽查盘点的基础上，资产评估专业人员和产权持有人一方就盘点表进行签字确认，形成数据资产抽查盘点底稿。

2. 数据资产产权持有人的盘点工作及实施程序

数据资产的形态和价值往往难以通过简单物理计量或直观查看来确认，这就要求其盘点过程必须依托更为严谨的技术手段和系统化方法。数据资产的动态性和多样性决定了盘点工作需要与数据处理流程各阶段紧密结合，以保障数据的完整性、准确性和一致性。因此在实务中对数据资产的盘点不仅要关注其最终存在状态，还应追溯其生命周期中的每个关键节点，从而全面评估其真实价值与潜在风险。评估专业人员应了解数据资产的盘点流程，才能有效地完成现场调查中的监盘工作，故将数据资产产权持有人通常履行的盘点技术手段和程序步骤介绍如下：

（1）数据资产盘点技术手段与工具

数据资产盘点的技术手段与工具是实现数据资产盘点的核心支撑。这些工具不仅能够借助自动化元数据采集、跨系统数据映射等技术，识别并整合分散在不同业务系统的数据资源，使数据资源通过元数据实现透明化，进而达到可管理、可控制、可使用的状态；还能运用脱敏处理、权限隔离等技术手段，在不暴露原始数据内容

的情况下完成盘点分析，切实满足“数据可用不可见”的基本要求，保障数据安全与隐私合规。数据资产产权主体在资产清查的实务中，可以根据数据资产规模等因素选择适合规模的数据盘点工具。

数据盘点工具通常有以下类别：一是基础型盘点工具，如数据库管理系统自带的查询功能、Excel 等电子表格软件，适用于数据规模较小、结构简单的场景，可快速统计数据条目、字段类型等基础信息；二是专业型盘点工具，包括元数据管理工具（如 Apache Atlas、Collibra）、数据目录工具（如 Alation、Informatica Data Catalog）等，这类工具能自动化采集数据资产的元数据信息，构建数据资产目录，支持数据资产的分类、检索与关联分析；三是智能型盘点工具，融合了人工智能与大数据技术，具备数据质量检测、敏感数据识别、数据血缘分析等高级功能，可深度挖掘数据资产的价值属性与潜在风险，典型代表有阿里云 DataWorks、华为云数据资产目录等。部分专注于数据科技的企业以及一些大型资产评估机构，也着手开发并推出了专门用于数据盘点的专业化工具。

（2）数据资产盘点的核心步骤

①元数据的识别与采集

元数据作为“关于数据的数据”，是数据资产盘点的核心基础，其采集需结合多种方法以实现全面覆盖：一是自动化采集工具，通过 Apache Atlas、Informatica Metadata Manager、Collibra 等专业元数据管理工具自动扫描数据库、数据仓库及大数据平台，提取表结构、字段定义、数据血缘等信息，提升效率并减少人工错误；二是利用数据库系统自带功能，通过编写 SQL（Structured Query

Language，结构化查询语言）查询从系统表或视图中提取元数据；三是自定义脚本开发，针对特殊需求采用 Python（如 SQLAlchemy 库）、Java 等编写脚本灵活提取存储元数据；四是借助 ETL（Extract Transform Load 抽取、转换、加载）数据资产盘点与血缘分析工具如 Informatica PowerCenter、Talend 等工具在数据流转过程中自动捕获元数据；五是手动补充录入，针对无法自动获取的业务描述、数据所有者等信息，建立 Web 表单等手动录入流程。

②数据分级分类

分级分类是确保数据安全管理和合规使用的前提，依据《数据安全法 数据分类分级规则》（GB/T 43697-2024，2024 年 3 月 1 日国家市场监督管理总局、国家标准化管理委员会发布，自 2024 年 10 月 1 日起施行）通常遵循“先分类，后分级”的原则。数据分类需从多维度展开：业务维度按所属领域划分，例如财务、人力资源、客户数据等，可采用“行业领域—业务职能—内容主题”的多级架构；来源维度分为内部生成与外部获取数据；管理维度关注结构化特征，包括结构化、半结构化、非结构化，以及存储方式和更新频率等。分类需遵循同一维度下，一条数据仅归一类、类目划分符合普遍认知且用语一致的原则。

数据分级则根据数据遭篡改、破坏、泄露或非法使用后对国家安全、公共利益、个人或组织合法权益的危害程度定级。常见级别划分模型包括三级模型，即核心数据/L6/L5、重要数据/L4-L3、一般数据/L2-L1，其中核心数据常涉及国家安全重点领域或国民经济命脉，以及六级模型，如武汉市指南细分为 L1 至 L6，实现精细化

管理。分级规则明确：同一类别数据含不同级别时按最高级别定级，分级结果直接决定共享开放策略与安全保护措施，例如核心数据严格限制共享。

③数据血缘分析

数据血缘分析旨在追踪数据从初始产生、加工处理到流转使用的全链路过程，核心目标是明确数据的来源、流动轨迹及处理环节，确保数据真实可靠、权属清晰，是数据资产评估中验证权属完整性、数据质量合理性及加工合规性的核心环节，可有效支撑评估中的问题根因分析与结论可信度提升。在实践中，数据血缘分析需通过系统化方法实施：一是利用自动化工具解析结构化数据，例如数据库表、ETL 脚本的血缘关系，标记非结构化数据的流转路径，追踪实时流数据的生产与消费端信息，并生成可视化血缘图谱；二是关联全生命周期操作日志，形成“血缘链路+操作记录”的证据闭环，当发现数据异常时可快速定位源头与责任人；三是基于血缘分析链路提取关键信息支撑评估测算，包括权属信息，即初始授权文件、流转范围、变更记录；质量评价信息，即源头错误率、加工改进幅度、更新频率；成本分摊信息，即采集、加工、存储等环节投入；收益贡献信息，即直接/间接/长期收益及关联度。

④数据的安全评价

安全评价以数据分级结果为基础，核心是评估数据全生命周期的安全防护能力，这是数据资产合规性验证的关键环节，直接影响其评估价值的合法性与稳定性。评价需围绕分级结果制定差异化策略：通用安全要求方面，需建立覆盖数据安全策略、规程及资产清

单的管理体系，确保符合《中华人民共和国数据安全法》《中华人民共和国个人信息保护法》等法规；生命周期各阶段需针对性防护，如核心数据传输与存储应采用高强度加密；云环境下还需关注服务提供商是否具备透明可信的安全能力体系，包括基于生命周期的安全管理与合规保障机制，以支撑数据资产在云场景下的权属安全与合规使用。

⑤数据的质量评价

数据资产质量评价是价值量化的关键基础，直接决定评估结果的可靠性与合理性，其核心在于判断数据是否满足准确性、一致性、完整性、规范性、时效性及可访问性等核心要求。准确性体现数据与真实情况的相符程度；一致性反映跨系统数据的逻辑统一性；完整性关注数据字段的缺失比例及覆盖范围；规范性强调符合隐私保护和行业监管要求；时效性衡量数据的更新频率和新鲜程度；可访问性则指数据在需要时能被合法授权主体顺畅获取使用的特性。评价过程需技术与业务深度融合，一方面通过数据剖析、统计分析、规则引擎等技术手段量化各维度指标，另一方面结合业务部门实际使用场景验证数据适用性，确保评价结果与数据资产的业务价值紧密关联。

问题 3：数据资产评估过程中为何要进行数据血缘分析？

数据血缘分析（Data Lineage Analysis），亦被称为溯源分析，其核心目的在于利用技术手段追踪并描述数据从产生源头、经过各级采集、清洗、加工、存储直至最终应用及销毁的全生命周期流转路径。数据资产因具备非实体性、可复制性及价值动态变化等特征，

与传统实物资产在权属界定、质量控制及价值归因等方面存在差异。开展数据血缘分析是数据盘点的重要环节。通过构建清晰的数据流转脉络，能够有效执行对数据资产的权属合规性、质量可靠性相关资料的核查验证，以及识别价值形成与贡献的合理性。

1. 数据资产权属资料的核查验证

数据血缘分析能够系统性地回溯数据从源头至最终应用的整个流转环节，清晰界定各节点的权属归属和合规状况。通过追踪数据的完整流动路径，可以提取原始授权文件，如初始授权协议、行政许可证明、采购合同以及上游权属链等文件，还能提取流转授权范围，包括内部流转权限、对外共享期限和用途，以及权属变更记录，如转让、过期、争议处理凭证等权属信息，为数据资产产权核查提供核心支撑，确保数据资产权利边界的清晰性与合规性。

数据资产全链路合规性核查需要覆盖来源、清洗、加工、存储、流转传输这五大核心环节：在来源环节，需按照生成渠道分类验证合法性。对于内部业务系统数据，通过操作日志和权限记录确认企业合法持有；对于互联网采集数据，核查《隐私政策》等授权文件是否符合“告知同意”原则；对于传感器数据，追溯设备部署授权和行政许可；对于交易获取数据，调取采购合同和权属证明，并通过血缘分析验证转让链条的合规性。在清洗环节，清洗规则制定主体应在授权范围内操作，合并第三方数据时保留原始权属标识；在加工环节，核查授权文件和合作协议，标注衍生数据的权属归属以界定权利边界。在存储环节，通过血缘分析定位物理存储位置，对于自有服务器，匹配固定资产证明；对于第三方云存储，核查服务

协议中的所有权条款，规避权属旁落的风险。在流转传输环节，追踪加密协议和权限记录，确保符合《中华人民共和国数据安全法》的要求，同时核查接收方的授权范围，防止违规流转。

2. 数据质量与可靠性评价资料的核查验证

《数据资产评估指导意见》明确规定，数据质量评价应当涵盖准确性、完整性、一致性、时效性、规范性和可访问性等维度。数据血缘分析依托全链路追踪技术，能够为质量评价提供系统性的证据支撑。

准确性验证：通过全链路核查，追踪数据在各处理节点的错误率记录，验证数据是否符合真实客观的要求，是否存在因算法缺陷导致的系统性偏差；完整性验证：追踪核查数据字段在流转过程中的增删改记录，重点核查是否存在关键业务信息如时间戳、用户 ID、授权字段的非预期丢失或遗漏；一致性验证：利用血缘关系比对多来源数据在逻辑上的冲突如同一用户在不同业务系统中的基础信息差异，识别数据孤岛导致的一致性问题；时效性验证：追踪核查数据的更新频率、ETL 处理时长及传输延迟等技术指标，判断数据到达应用端的时间是否满足业务场景的实时性要求；规范性验证：校验数据在存储、传输、使用过程中是否遵循统一的格式标准如日期格式“YYYY-MM-DD”、编码规则及字段命名规范，核查元数据描述是否完整，确保无违规的非标数据混入；可访问性验证是一个确保产品或服务能够被所有用户无障碍使用的过程，它通过核查存储介质的读取性能以及传输路径的稳定性，来保证数据在需要时能够被正常访问。

这些通过血缘分析获取的质量指标，将直接支撑成本法模型中质量调整系数的确定。数据质量越优良，评估值越高；反之，则需扣减系数，对重置成本进行贬值调整，从而确保质量核查结果与评估参数取值的一致性。

3. 合理识别数据价值形成与贡献

数据血缘分析通过全链路追溯，能够系统性地支撑成本法中的数据成本投入分析与收益法中的数据贡献价值评估，解决价值量化过程中的“模糊”问题。

（1）量化成本投入

血缘分析贯穿数据全生命周期中每一个产生成本的环节，为历史成本的归集与分摊提供了量化依据，是估算重置成本或开展价值分析的重要基础。

成本分摊需覆盖数据资产全生命周期的核心环节：①采集环节通过追踪数据爬虫工具运行时间占比等实际资源占用情况，按比例分摊人力成本与技术采购成本；②清洗加工环节提取各数据资产在 ETL 处理中的算力消耗占比（如 GPU 使用时长、云计算资源占用字节数），据此分摊算力费用与第三方服务支出；③存储与维护环节依据数据资产在生命周期中的存储空间占用比例及备份频率，合理分摊服务器硬件采购费、云存储租赁费及安全软件维护费；④传输环节则针对企业级专线带宽、CDN 加速节点等共用传输软硬件，通过血缘分析实现多维度分摊——基于数据传输流量占比分摊带宽成本，基于资源占用时长分摊节点成本，并可根据业务优先级权重调整分摊比例。这些的分摊指标，确保了成本归集与数据资产的实际

资源消耗相匹配，为估算重置成本或开展价值分析提供基本依据。

（注：不同行业不同企业网络集群架构可能不同，成本分摊涉及的基础指标可能存在不同）

（2）追溯数据贡献

血缘分析能够建立数据资产与业务价值之间的因果链条，提取各环节的降本增效指标，为收益法中的预期收益预测及价值贡献分摊提供核心支撑。

①直接收益贡献可通过追踪数据在具体业务场景中的调用记录及效果，量化其对核心 KPI 的提升作用，比如用户行为数据提升营销 ROI、风控数据降低坏账率，以此明确不同数据资产对直接收益的贡献占比；②对于间接收益贡献，则需提取业务流程优化的降本指标，像供应链数据缩短库存周转天数、降低仓储运营成本，再通过区分不同数据资产在优化过程中的参与度，衡量间接收益的分摊比例；③长期收益潜力的评估，要基于数据在历史链路中的复用率及衍生能力，预测其对未来降本增效的影响，例如用户特征数据缩短新产品研发周期、提升市场占有率，从而为收益法中未来收益期的参数预测提供动态依据。

问题 4：数据资产评估过程中如何进行数据血缘分析？

数据血缘分析的实施是一项复杂性工作，需遵循标准化的操作步骤，有效结合技术工具与管理流程。数据血缘分析可以利用数据领域专家工作成果及相关专业报告等。资产评估专业人员自行履行数据血缘分析程序时，应当确保具备相应专业知识、技术手段和经验。数据血缘分析具体操作可分为以下五个关键环节。

1. 界定血缘分析范围：明确分析边界

开展血缘分析的首要任务是依据评估基准日及评估目的，从时间范围、数据范围及链路深度三个维度划定清晰的分析边界。

时间范围需严格匹配评估基准日，确保血缘分析数据的时间属性与评估对象一致，通常涵盖数据的生成、持有期；数据范围应全面覆盖评估对象关联的全链路数据，包含基础类原始数据、过程类加工数据及应用类衍生数据；链路深度则需依据血缘分析目标动态确定追溯的层级，一般通过四个方面进行确认：基于权属确认时，必须追溯至初始数据源，如用户签署授权协议的原始记录、系统生成数据的源头日志，以验证权属来源的完整性；基于质量评估时，通常追溯至直接上游数据源及关键加工节点，如 ETL 过程的输入数据，确保覆盖质量评价的五个维度，若直接上游质量存疑，则需继续向上穿透；基于成本分摊时，血缘分析范围需覆盖数据全生命周期的所有成本发生环节，包括采集、清洗、加工、存储、传输，链路深度需追溯至资源消耗的直接关联环节，比如算力消耗的具体步骤、人力投入的具体团队，确保成本分摊依据充分；基于收益贡献时，需关联数据应用场景的收益产生环节，直接收益追溯至用户行为数据的上游源头，间接收益追溯至加工整合步骤，长期收益追溯至衍生逻辑，链路深度需追溯至对收益有直接影响的关键环节，为贡献占比量化提供依据。

2. 搭建元数据基座：构建血缘分析的基础架构

元数据（Metadata）是描述数据的数据，是构建血缘关系、还原数据流转路径的基础载体。搭建坚实的元数据基座需通过需求分

析、元数据的采集、验证与优化三个步骤进行：

需求分析需明确血缘分析目标对应的元数据类型，包括技术元数据、业务元数据及操作元数据，其中技术元数据涵盖存储位置、表结构、接口协议，业务元数据包含业务含义、指标口径、应用场景，操作元数据涉及生成时间、操作人员、加工步骤、授权记录；元数据采集则利用自动化工具从各类数据源中自动抽取三类元数据，在存储与管理环节要将采集的元数据结构化存储于统一的元数据仓库中，建立技术元数据与业务元数据之间的映射关联，形成数据资产的全貌视图；验证与优化需定期校验元数据的准确性与现势性，并根据血缘分析需求的变化动态调整采集范围。

3. 链路构建：还原数据完整流转路径

链路构建是血缘分析的核心技术环节，需结合“数据血缘分析技术”与“操作日志追踪技术”，从“数据流向”与“操作行为”两个维度，还原数据从源头到评估节点的完整链路。

结构化数据解析需针对关系型数据库如MySQL、Hive，通过SQL解析工具分析存储过程及 ETL 脚本，自动提取“源表-目标表一字段”级的映射关系；非结构化数据标记则面向半结构化或非结构化数据如JSON、文件，借助标签映射工具打上血缘分析标签以记录流转过程；实时流数据追踪要针对Kafka等消息队列中的实时数据，利用流处理框架记录生产端与消费端信息；可视化图谱生成需基于上述解析结果，生成节点代表数据实体、边代表加工关系的血缘关系可视化图谱；日志关联闭环需采集全生命周期的操作日志，包括源端生成日志、加工端ETL日志、存储端访问日志，并将其与血缘

链路进行关联，例如当发现数据异常时，可通过血缘图谱定位存储位置，再结合操作日志追踪具体的访问人员与操作时间，形成完整的证据闭环。

4. 信息提取：获取评估所需量化指标

基于构建的血缘链路，提取各类关键信息以支撑评估测算：权属信息：提取初始授权文件，包括原始授权协议、行政许可证明、采购合同及上游权属链文件；流转授权范围，包括内部流转权限、对外共享期限与用途；权属变更记录，包括转让、过期、争议处理凭证。质量评价信息：提取源头错误率，即采集阶段的缺失、格式错误占比；加工后质量改进指标，即清洗环节对准确性、一致性的提升幅度；更新频率，即实时/日/周更；缺失率，即关键字段空值占比。成本分摊信息：提取各环节投入成本，包括采集环节的人力与工具成本、加工环节的算力与软件成本、存储环节的硬件与维护成本、传输环节的带宽与加密成本。收益贡献信息：提取直接收益，包括销售额增长、ROI 提升；间接收益，包括运营效率优化、成本节省；长期收益，包括研发周期缩短、潜在价值及各收益对应的上游数据环节关联度。

5. 形成血缘分析报告

根据上述血缘分析，形成血缘分析报告，直接服务于评估结论的形成。血缘分析报告通常包含数据链路图谱，即可视化展示完整流转路径；关键信息汇总表，即结构化展示权属、质量、价值贡献信息；血缘分析结论，即明确链路完整性、合规性及质量等级等；依据血缘分析结果优化和调整评估相关参数。（备注：目前数据资

产的血缘分析报告通常需借助第三方数据科技公司提供，评估专业人员引用第三方报告时，应关注血缘分析的过程及内容。）

主要执笔人：陈蕾、张月梅、殷守梅、殷霞、温云涛
胡金华、王占峰

外部协同专家：黄芳（北京融理律师事务所律师）
邵长钰（上海北京大学临港中心、数字经济
中心主任）

审核人：阮咏华

校对人：胡金华