

衍生数据权利归属研究

刘欢欢 覃远春

贵州财经大学 贵州贵阳 550025

摘要：在 2022 年，国务院发布了数据 20 条的意见中，明确指出了数据作为新的生产要素，需要建立数据产权制度，随着人工智能与大数据技术的快速发展，在原始数据的基础之上，应运而生了具有更高商业价值的衍生数据，然而，由于衍生数据生成过程的复杂性和参与主体的多样性，其权属界定成为亟待解决的难题，制约着数据要素市场的健康发展，因此对于衍生数据的权利归属的研究也成为了相关热点研究对象，本文从不同的角度衍生数据的权属，明确衍生数据的归属，以更好促进经济的发展。

关键词：数据；衍生数据；衍生数据归属

对原始数据进行收集，加工、处理、整合之后得到的数据为衍生数据，衍生数据是经过不同主体的多个流程合作才得到的，具有不菲的商业价值，需要对衍生数据进行保护，理论中存在两种主流观点，一种认为需要通过赋予权利进行保护，可以得到更好的保护效果，另外一种观点认为赋予权利保护具有较大的局限，不利于技术的进步，社会的发展，通过反不正当竞争保护；其次，对于衍生数据的归属在理论与实践中都存在较大的争议，有观点认为衍生数据的所有权应该给原始数据收集者，相反，也有观点支持赋予数据处理者衍生数据所有权；除此之外，对于委托加工的衍生数据、职务中完成的衍生数据、共同合作完成的衍生数据的所有权，使用权的归属问题都是目前具有较大争议，尚未解决的问题。

1. 衍生数据的确权的合理性基础

关于衍生数据权属的界定，需以其法律属性为逻辑起点。当前理论界与实务界存在两种典型立场：其一为“赋权保护论”，主张将衍生数据纳入排他性权利体系，通过创设新型财产权或邻接权确认持有者的独占性权益。该路径强调衍生数据的创造性劳动投入与商业价值属性，认为赋权机制可明晰权责边界，激励数据加工主体持续创新，同时为司法裁判提供确定性依据。其二为“法益保护论”，反对以绝对权形式固化衍生数据权属，主张通过反不正当竞争、侵权责任等一般性法律规则实现有限保护。其核心关切在于，数据作为新型生产要素具有公共属性，若赋予过强的排他性权利，可能抑制数据要素的流通共享，反而不利于数字经济发展与社会整体效率提升。该观点特别警示，过度赋权易形成

数据垄断，阻碍技术创新与公共福祉增进，故更倾向以动态利益平衡机制替代静态权利配置。

1.1 劳动财产理论与智力成果保护

依据洛克“劳动财产理论”，当个人通过劳动将共有资源转化为具有独立价值的成果时，应当享有排他性权利。衍生数据的生成需经历数据清洗、脱敏处理、算法建模、价值挖掘等复杂技术流程，本质上是数据加工者投入创造性劳动与资本后的智力成果。法律通过确权承认其劳动价值，符合“无耕耘则无收获”的公平原则。

若衍生数据已脱离原始数据的可识别性，通过算法重组形成新的信息聚合体或预测模型，则其具备“独创性表达”或“实质性转换”特征，符合知识产权法保护“增量价值创造”的基本逻辑。

1.2 平衡私权保护与公共利益

若缺乏确权将导致“强技术主体”通过技术优势无偿攫取他人衍生数据成果，加剧市场垄断。确权规则可通过许可使用、收益分成等机制，实现数据价值链的公平分配。确权不等于绝对封闭，而是通过“权利束”设计，如设定数据共享义务、强制许可条款，在保护私权的同时满足公共科研、社会治理的数据需求。欧盟《数据治理法案》即采用此类“有限排他性”思路。

就保护模式的选择而言，赋权保护相较于利益保护具有显著的制度优势。根据科斯定理，明晰的产权界定是资源优化配置的前提条件。通过赋予数据加工处理者独立的财产权利，不仅能够有效克服传统利益衡量模式的不确定性缺

陷,更可形成稳定的制度预期,激发市场主体持续投入数据生产要素的积极性。我国《数据安全法》第7条确立的“数据权益”概念,以及《民法典》第127条对数据权益的原则性规定,均为赋权保护模式提供了规范基础。

2. 衍生数据权利界定

数据经过一系列的加工生成衍生数据,对于衍生数据的归属认定时,存在不同的情况,当收集,加工处理的是同一主体,其归属主体是明确的,没有争议,但是当收集,加工处理属于不同一主体时,对于多人合作完成的衍生数据以及职务作品,合同有约定按合同约定,但当合同没有明确其归属时,其归属问题就成了一个巨大的争议,该问题也是学术理论界的热门话题,目前主要存在三种主流的观点,单一主体所有,共有权属,社会公众所有。

2.1 单一主体所有

单一主体所有主张衍生数据的排他性权利应归属于某一特定主体(如原始数据收集者或加工处理者)。其逻辑基于“权益延续性”或“劳动赋权”——原始数据方认为衍生数据是初始数据的自然延伸,而加工方则强调自身通过技术投入创造新价值,类比知识产权保护。这种模式的优势在于权责清晰,可激励数据开发投入(如企业核心数据产品的商业保护),但可能导致数据垄断或流通壁垒,例如平台企业独占用户画像数据阻碍市场竞争。

2.1.1 原始数据控制权优先论

原始数据控制权优先论主张衍生数据的所有权应归属于原始数据收集者,其核心逻辑在于承认数据价值的“源头依赖性”,认为原始数据是衍生数据的必要前提,其生成依赖原始数据的“母体”,因此所有权应追溯至原始数据来源,若允许加工处理者独立拥有衍生数据,可能导致数据垄断风险,如原始数据方失去对下游价值链的控制。

同时该理论认为,衍生成果本质上是原始数据的延伸与增值形态,因为原始数据控制者基于合法采集(如用户授权、合同约定)与数据安全责任(如GDPR对数据控制者的义务规定),应享有对衍生数据的持续控制权。例如,中国《数据安全法》虽未直接规定数据权属,但强调数据控制者对数据安全的义务,可能间接支持原始数据方的权益。法律实践中,这一观点常通过“权益延续性”原则体现(如德国判例要求数据加工需原始方授权),旨在防止数据控制链断裂导致的隐私泄露或垄断风险,同时强化数据全生命周期

的合规性。然而,批评者指出过度强调原始方权利可能抑制数据二次创新,需通过合理使用规则平衡各方利益。

2.1.2 数据加工劳动成果论

数据加工劳动成果论的核心主张是,数据加工处理者通过技术投入与创造性劳动对原始数据进行增值处理,应享有衍生数据的所有权或排他性权益。其法理基础源于洛克“劳动赋权”理论,强调加工者在算法设计、算力投入、数据分析等环节的实质性贡献,使衍生数据脱离原始数据的原始形态并形成独立价值(如商业洞察、预测模型)。实践中,这一观点常通过知识产权路径实现保护,例如美国《数字千年版权法》(DMCA)对独创性数据库的保护,或中国司法判例(如“淘宝诉美景案”)承认加工方对衍生数据的竞争性财产权益,认为其劳动成果具有排他性价值。

当数据经深度加工形成具有独立价值的衍生数据时,其财产权归属应当遵循“劳动赋权”原则。数据加工处理者通过算法创新、清洗脱敏、建模分析等技术投入,使原始数据发生质变并产生增量价值,依据洛克劳动财产理论,理应获得相应的数据财产权。

再者,从产业实践维度考察,将衍生数据权属赋予加工处理者符合数字经济运行规律。以金融科技领域的信用评分模型为例,各机构通过海量数据处理形成的风险评估模型,其商业价值主要来源于持续的算法优化与数据训练。若简单适用共同共有或社会公有规则,将导致“搭便车”现象,抑制企业的创新投入。欧盟《数据法案》第35条关于“数据生成者权利”的规定,以及美国《加州消费者隐私法》对衍生数据的除外条款,均体现了对数据处理者权益的保护趋势。

然而,该理论面临伦理与法律争议:一方面,过度强调加工者权利可能削弱原始数据方的控制权,甚至引发隐私风险(如未经充分脱敏的数据再利用);另一方面,单纯依赖劳动赋权原则可能忽视数据来源的合法性基础。

而笔者认为,首先强调加工者权利并非必然削弱原始数据方控制权或加剧隐私风险,关键在于建立权责对等的制度框架与技术约束。首先,加工者的权利主张以合法获取原始数据为前提,其权益范围通常限于经脱敏或聚合处理后的衍生成果,而非原始数据本身。例如,GDPR等法规明确要求数据处理者承担合规义务,包括匿名化标准、数据最小化原则,确保加工环节不回溯至原始可识别信息。其次,现代技术手段(如差分隐私、联邦学习)已能实现“数据可用不

可见”，在保障加工者开发权益的同时，通过算法机制阻断原始数据泄露风险。此外，原始数据方可通过合同保留核心控制权（如限制衍生数据用途、设置数据回传审计条款），形成“权利分层”结构。因此，加工者权利与隐私保护并非零和博弈，而是需通过法律细化（如明确衍生数据再处理规则）与技术创新实现动态平衡。

其次劳动赋权原则与数据来源合法性并非互斥，二者本质上是数据权益链条的“前置条件”与“后续分配”关系。劳动赋权的适用前提已隐含对数据来源合法性的要求——加工者主张权益的基础必须建立在原始数据的合法获取（如用户授权、合同约定或法定许可）之上，例如 GDPR 明确规定数据处理者需证明其处理活动的合法性基础（如同意、公共利益等），否则劳动成果因“污染数据”而丧失正当性。因此，劳动赋权原则并非忽视合法性，而是将合法性作为权益主张的必要门槛。实践中，司法判例（如“淘宝诉美景案”）在认可加工者权益时，均严格审查原始数据来源的合规性，仅对合法获取的数据衍生成果予以保护。劳动赋权与合法性审查共同构成数据权益的完整逻辑闭环，前者解决价值创造后的利益分配问题，后者确保价值创造起点的正当性，二者缺一不可。

2.2 共有权属

该观点认为应该根据数据要素在生成、加工、流通各环节的贡献，分配权属。例如，《数据二十条》提出“数据产权分置”模式，将数据资源持有权、加工使用权、产品经营权分离，由不同主体分别享有，同时单一主体所有可能导致数据壁垒，阻碍流通。共有模式通过强制许可、数据开放共享等制度平衡私权与公共利益。这种观点综合考虑了原始数据所有者和数据处理者的贡献。原始数据是衍生数据产生的基础，而数据处理者的工作则赋予了原始数据新的价值，因此双方都对衍生数据享有一定的权利，应共同拥有衍生数据。

但是笔者认为，若归共同所有，原始数据所有者和数据处理者的贡献难以精确量化，导致在确定双方对衍生数据的具体权益比例时，容易产生争议，其次责任界定不清：在数据管理过程中，如数据安全保护、数据质量维护等方面，原始数据所有者和数据处理者可能会因为责任界定不明确而出现相互推诿的情况。

2.3 社会公众所有

当数据具有公共属性或者是在公共领域中产生的，那么

基于这些数据产生的衍生数据也应归社会公众所有。这种观点强调数据的公共利益属性，认为数据及其衍生数据应该为社会大众所共享，以促进社会的整体发展和公共利益的提升。

如政府部门收集的人口普查数据、环境监测数据等公共数据，经过分析处理后得到的衍生数据。

若将衍生数据归属于社会公众所有，可能引发一系列弊端：一方面，创新激励受阻，数据创造者因难以从自身劳动成果中获取合理回报，投入创新与资源收集的积极性大幅降低，例如科技企业花费巨资研发用户行为分析模型，若数据归属公众，企业可能因收益无法覆盖成本而放弃研发，阻碍数据技术的进步；另一方面，数据治理混乱，公众对数据的归属认知存在差异，难以形成统一、有效的管理规则，导致数据滥用、泄露风险增加，且在数据侵权等纠纷中，责任界定困难，维权成本高昂。

3. 结语

衍生数据作为数字经济时代的核心生产要素，其法律保护机制的建构已成为现代经济治理体系的重要命题。当前理论界围绕数据权益保护存在“赋权保护”与“利益保护”两种范式之争，本文基于制度效能与激励机制的比较分析，主张构建以赋权保护为主导的规范体系。在此框架下，针对衍生数据的权利归属这一基础性法律问题，本文通过类型化分析方法，将权利主体区分为个人单独所有、共同共有及社会公众所有三种模式，进而从法律经济学视角论证数据加工处理者作为独立权利主体的正当性依据。

参考文献：

- [1] 朱瑞. 图书馆衍生数据的权利确证、保护困境及路径选择 [J]. 图书馆建设, 2024, (04): 47-57.
- [2] 田大治. 数字社会衍生数据确权的法理证成 [J]. 学习与实践, 2022, (08): 44-50.
- [3] 王利明. 论数据权益：以“权利束”为视角 [J]. 政治与法律, 2022, (07): 99-113.
- [4] 严宇, 孟天广. 数据要素的类型学、产权归属及其治理逻辑 [J]. 西安交通大学学报 (社会科学版), 2022, 42 (02): 103-111.
- [5] 杨翱宇. 数据财产权益的归属判定 [J]. 重庆大学学报 (社会科学版), 2023, 29 (01): 241-254.
- [6] 王渊, 黄道丽, 杨松儒. 数据权的权利性质及其归属研究 [J]. 科学管理研究, 2017, 35 (05): 37-40+55.