引用格式：薛晓斐，肖毅，贺晨，等. 基于限制性合作博弈的跨域多源数据收益分配机制研究[J]. 技术经济，2025，44(9)：143-154.
Xue Xiaofei，Xiao Yi，He Chen，et al. Cross-domain and multi-source data revenue distribution mechanism based on restricted cooperative game theory[J]. Journal of Technology Economics，2025，44(9)：144-154.

技术经济评价

基于限制性合作博弈的跨域多源数据收益分配机制研究

薛晓斐[1,2]，肖 毅[1]，贺 晨[1]，张进澳[1]
（1. 华中师范大学信息管理学院，武汉 430070；2. 河南工程学院管理工程学院，郑州 451191）

摘 要：数据要素作为数字经济时代的重要新质生产力，其收益分配机制是促进数据要素市场健康发展的关键。跨域多源数据收益分配机制的研究仍处于起步阶段，尚缺乏成熟的解决方案。本文基于合作博弈理论，将数据要素的非排他性、数据质量和数据整合等因素作为激励机制融入到 Shapley 值法，在此基础上结合成员之间的连接关系，提出改进的 Myerson 值法作为合作博弈的解，用于解决跨域多源数据收益分配问题。研究结果表明，网络结构对数据资源的价值创造、边际贡献和收益分配具有显著影响，激励机制能够有效调节各成员之间的收益分配，最终形成更加公平合理的收益分配格局。研究成果为跨域多源数据收益分配机制提供了新的思路和方法，对构建完善的数据要素市场交易体系、实现数据要素价值释放、促进数字经济高质量发展具有重要的理论价值和实践意义。

关键词：数据要素；收益分配；跨域多源数据；限制性合作博弈；Myerson 值
中图分类号：C933 **文献标志码**：A **文章编号**：1002-980X(2025)09-0143-12
DOI：10.12404/j.issn.1002-980X.J24091606

一、引言

数据作为新质生产力的重要组成部分，在数字经济时代具有至关重要的作用，是经济社会发展的战略性基础性资源。自党的十九届四中全会首次提出将数据作为生产要素参与分配以来，国家层面相继出台一系列重要政策文件。其中，2022 年 12 月颁发的《关于构建数据基础制度更好发挥数据要素作用的意见》（简称《数据二十条》）从数据产权、流通交易、收益分配、安全治理等基础制度出发，加速推动我国数据要素的价值释放。党的二十大、2024 年的两会均对数据要素的发展进行顶层设计，从过去几年国家的重要文件和战略规划可以看出，数据要素将是未来几年中国战略布局与发力的重要领域。数据要素市场化是建设统一大市场的重要组成部分[1]，加快培育数据要素市场，有助于推动数字经济和实体经济深度融合，将数据资源“势能”转变为高质量发展的“动能”，实现“中国制造”向“中国智造”转型升级，对驱动中国传统生产力面向新质生产力转换具有重要意义。

数据资源规模庞大，但价值潜力还未充分释放是中国数字经济发展面临的问题和挑战之一。根据国际数据公司（IDC）的预测，2025 年中国的数据量将达到 48.6 字节（1 字节 = 10000 亿吉字节），数量是 2022 年

收稿日期：2024-09-16
基金项目：教育部人文社会科学研究规划基金项目“基于群智能优化策略的大数据混合预测方法及其应用研究”（21YJA630098）；中央高校基本科研业务费“融合知识驱动的大数据混合多任务预测建模研究”（CCNU24ZZ131）
作者简介：薛晓斐（1989—），华中师范大学信息管理学院博士研究生，研究方向：数据要素收益分配；（通信作者）肖毅（1975—），博士，华中师范大学信息管理学院教授，博士研究生导师，研究方向：智能决策；贺晨（1995—），华中师范大学信息管理学院博士研究生，研究方向：时间序列预测与模型可解释性；张进澳（2000—），华中师范大学信息管理学院硕士研究生，研究方向：人机协作与用户行为。

的6倍(8.1字节)。国际商业机器公司(IBM)的研究报告表明,大多数企业仅对其拥有数据的1%进行了分析应用,大量数据仍处于未开发、未利用的沉淀状态。近年来国内相继成立的数据交易所数量高达65个,但仍面临数据要素流通困难、市场生态发育不良、场内数据交易量不足等一系列挑战[2]。为充分释放数据要素价值,政府出台了一系列相关政策及措施。2023年8月,财政部印发《企业数据资源相关会计处理暂行规定》,明确数据资源的确认范围和会计处理适用准则等,有利于显化数据资源价值[3]。2023年12月,国家数据局等17部门联合发布《"数据要素×"三年行动计划(2024—2026年)》,明确提出在交通运输、医疗健康及科技创新等重点领域,通过跨领域跨主体间数据要素的流通与融合,依托大数据技术进一步提高数据的内在价值,赋能经济社会的高质量发展。实践中,单一数据在特定应用场景下无法满足现实要求,需要跨域多源数据的整合。例如,跨域多源数据在推进智能汽车创新发展的应用场景中,需要打通车企、第三方平台、运输企业等主体之间的数据壁垒,促进道路基础设施数据、交通流量数据、驾驶行为数据等多源数据融合应用,提高智能提车创新服务、主动安全防控水平。

然而,在数据要素市场化进程中,数据开放共享阻力大、顾虑多,存在"不愿开放、不敢开放、不会开放"的难题。造成大量数据"深藏闺中"的因素有很多,主要集中在产权不明确、数据质量参差不齐、收益分配机制不清晰、交易机制不完善、数据安全及隐私保护、治理制度不健全等方面。当数据产品需要跨领域跨主体的基础数据资源时,各种问题更加凸显,特别是多源数据主体之间的收益分配机制,存在"不患寡而患不均"的情况。科学合理的收益分配机制是数据要素流通交易的必要条件,能够激活市场供需主体的积极性,增强数据流通意愿。因此,构建公平合理的收益分配机制是本文的核心研究问题,如果不能得到有效解决,则无法平衡数据要素市场诸多参与者的利益诉求,势必会降低数据交易的活跃度,不利于数据要素市场的健康发展。

二、文献回顾

本文分别从跨域多源数据整合、数据要素特征、收益分配机制等方面探讨跨域多源数据主体之间的收益分配机制。

数据越来越被视为21世纪最具战略意义的资源,其重要性与黄金和石油相似[4]。随着数字技术不断发展、数字经济占GDP的比重持续上升,区块链、隐私计算、联邦学习等技术的应用推动隐形信息和知识显性化,为高效利用大数据提供了"发动机",使得大量数据转化为生产力成为可能[5]。以ChatGPT为代表的人工智能生成内容(AIGC)技术取得突破离不开高质量数据的发展,人工智能正在从"以模型为中心"转向"以数据为中心",数据要素的战略地位进一步凸显。在复杂应用场景下,仅依赖单一数据源往往难以反映问题的全貌,多源数据的整合成为破解这一难题的关键途径。数据整合是将多个不同的数据对象创建成一个统一的数据对象的过程,也称为数据合并或连接[6]。通过这种整合方法,不仅能够深入挖掘数据的潜在价值,还能显著提升数据产品的经济效益。常见的多源数据整合分两种,即纵向整合和横向整合。纵向整合指数据资源所具有的属性相同或者相似,多源整合增加了数据的样本量;横向整合指数据资源所具有的属性不同,多源整合增加了数据的维度,样本量保持不变[7]。跨域多源数据整合属于横向整合的范畴,不同行业的数据资源经过整合后会产生更高价值的数据产品,其价值提升来源于数据协同增值效应。Yan[8]发现第三方"数据销售商"将多个企业的数据收集后进行整合,向市场提供综合性数据产品,产生的价值远高于企业单独向市场提供独立的数据。Xin和Srivastata[9]指出当与其他数据链接或者融合时,数据的价值会爆炸式增长,大数据整合对于实现大数据的前景极具挑战性。一般情况下,数据的整合度越高,其价值越大,但Goodhue等[10]发现数据价值与整合度呈现抛物线关系,20%的整合度可以达到80%的效用价值。跨域多源数据整合在产业界应用层面上,欧阳日辉[11]认为生产和消费数据融合有助于推动元宇宙、脑机接口、量子信息、人形机器人、生成式人工智能、新型储能等技术研发和产业发展。协同平台基于海量数据的建模分析,精准预测客户需求和市场动向。数据要素需要融合才能构建"数据-信息-数据生产力"的价值链条,具体包括价值实现的融合、活动空间的融合、应用场景的融合[12]。跨域多源数据整合在数字政府建设过程中,数据共享是驱动政府决策范式转变的创新源动力[13]。其中,两个或多个部门之间的跨部门数据共享,有助

于改善公共服务、政策制定有据可循、更加有效地建设组织系统，以及更好地应付紧急情况[14-16]。政府机构的跨部门数据能够为政策制定提供参考，不同领域、行业的数据融合也为企业决策提供支持。互联网广告、资讯服务和企业供应链管理等营销、风控领域的业务运转均依赖于海量数据的流转、整合与应用[17]。企业通过内外部跨部门跨行业跨领域的多源数据整合，实现更加精准的市场分析、用户画像和风险控制等，以达到降本增效提质的运营管理目标。此外，数据整合对学科建设也是至关重要的，如生态学与环境科学、人工智能与管理学，将不同的数据整合到一起，以揭示新的见解。由此可见，在数字政府建设、企业数字化转型和跨学科融合进一步深化的背景下，跨域多源数据在数据要素市场中的地位将越来越重要。

目前围绕跨域多源数据整合的研究多数在理论层面上[18-21]，涉及多源主体之间收益分配机制的研究相对较少。《数据二十条》指出健全由市场评价贡献、按贡献决定报酬的机制，更好地发挥政府在数据要素收益分配中的引导调节作用，建立保障公平的数据要素收益分配体制机制。数据要素收益分配主要面临“向谁分、分多少”的问题，重点解决向谁分不清楚、分多少无依据的困境。王衍之等[22]认为理想的数据要素市场不仅需要通过数据资源的整合、加工服务，实现数据资源向数据产品的转化，还需要通过科学合理的定价与收益分配机制，实现数据产品在供需双方之间的有效匹配。并在此基础上设计了符合中国国情的数据交易流通模式，构建了基于 Shapley 值的三类主体之间的收益分配机制，但未考虑数据要素的特征。王子玄等[7]提出了模糊修正的 Shapley 值模型用于联邦学习中多源主体的预期收益分配，虽然模型涉及了数据产品的影响因素，同样未将数据要素特征融入到收益分配机制中。之所以要考虑数据要素的特征，是因为传统生产要素（如土地、劳动力、资本等）的研究框架和模型并不能适用于数据要素，需要根据数据要素的特征进行修正。数据要素与传统要素相比，具有以下显著差异。首先，数据要素具有空间和时间上的非竞争性，可同时重复使用且不排他[23]。传统生产要素的收益分配机制通常基于要素的稀缺性和排他性，核心在于要素所有权的转移或单独使用。例如，土地所有者通过出租土地获得收益，但同一时间只能有一个主体使用，劳动力与资本亦如此。而数据要素可以被多个主体同时使用但不会损耗其价值[24]，传统生产要素的分配逻辑在数据要素领域失效。其次，数据要素的复制成本极低，再生产边际成本几乎为零[25]，零边际成本带来显著的规模经济效应。这意味着增加一份数据的副本或者将其提供给新的使用者，几乎不需要额外的成本。这与传统生产要素形成鲜明对比，每增加一件产品的生产均需要相应的原材料、劳动力和设备等成本投入。零边际成本对收益分配机制提出了更高的要求，如果仅仅以成本为基础的分配方式难以准确反映数据资源的真实贡献。最后，数据要素具有时效性。数据的价值通常随着时间的推移而快速衰减，特别是在动态环境下，时效性是数据价值的重要体现。例如，股票的交易数据在交易时段内具有极高的参考价值，但在收盘后，其价值会迅速降低。而传统生产要素的价值可能随着时间的推移显著上升，如土地和劳动力的成本持续上升。故在数据要素的交易和收益分配中，需要充分考虑数据的时间价值，只有及时整合和共享数据资源，才能充分释放数据的潜在价值，避免因时效性降低而导致数据资源的价值浪费。鉴于数据要素所具备的上述特性，在构建跨域多源数据收益分配机制时，应考虑数据要素的特征以适应市场化的需要。

在跨域多源数据收益分配机制的已有文献中，除了未考虑数据要素特征之外，还未考虑多源主体之间的连接关系。现有的多源数据收益分配机制参考了供应链系统中的相关规则，考虑了数据整合过程中供应链上各参与成员之间的协作。合作联盟的成立在数量上要求至少两个数据持有者的参与，并且还需要参与者之间建立连接关系，其中连接关系是主体间共创价值分配时不可忽略的重要因素[26-27]。已有的多源数据收益分配机制使用了 Shapley 值法[7,22]，该方法的求解思路是按照成员对联盟的边际贡献将收益进行分配[28]，前提假设是合作联盟内所有参与者均建立连接并能完全合作。换言之，Shapley 值法是基于完全合作博弈的。然而实践中跨越多源数据主体之间由于技术差异或商业壁垒等因素直接建立连接的可能性相对较低，完全合作的情形难以实现，这限制了 Shapley 值法的直接应用。Myerson[29]在 Shapley 值法的基础上提出了基于不完全合作博弈（或限制性合作博弈）的 Myerson 值法，已广泛应用在企业创新网络共创价值分配[26]、农产品供应链网络的收益分配机制[30]、医疗联合体的收益分配等[31]。限制性合作博弈的前提假设是连通的集合是可行联盟，不连通的集合是不可行联盟，不连通集合的收益可以视为各个连通分支的收益

之和[26]。Myerson值法设定只有直接或者间接连通的参与者所形成的联盟才能实现合作,适用于跨域多源数据主体的收益分配机制,并为其提供了理论支撑。

现阶段我国数据开放仍处于起步阶段,急需加快推进数据开放共享[32]。有关数据要素收益分配机制问题,国内外至今未有成熟的解决方案。基于此,本文将对核心问题进行详细描述,即在跨域多源数据整合过程中如何进行收益分配,之后结合数据要素特征,在限制性合作博弈的基础上构建基于资源和贡献率的收益分配模型,最后进行实例检验,并结合理论模型解释结果的实践意义。通过跨域多源数据收益分配机制的不断完善以规范市场化交易,有助于促进数据要素市场高质量发展。

三、问题描述

数据供应链系统内有N个($N>2$)在运营上保持独立的企业、组织或者机构。作为数据持有者,每个成员均具备进行数据的收集、清洗、存储、治理等有关数据要素市场化前置处理的基本能力,经过一系列步骤,成员将自身的原始数据整理成数据资源。来自多个成员的数据资源经过整合后,进而形成数据产品,之后投入到数据要素市场。各成员来自不同的领域或者行业,在运营业务上不存在交叉重复,确保成员的数据资源之间是异质性的,能够进行横向整合,并且对数据产品的整合具有互补作用。跨域多源数据的整合能产生更高价值的数据产品。例如,将医疗数据、运动数据、饮食数据、日常行为数据等资源整合后形成的数据产品有助于精准反应一个人的健康状况。从原始数据到数据资源,多个数据资源到整合成数据产品,在数据要素的形态改变过程中(图1),价值增值随之发生[33-34]。进一步地,原始数据到数据资源的过程发生在成员内部,由职能部门统一指挥统一协调。而多源数据整合成数据产品的过程则发生在多个成员之间,形成基于数据提供的合作联盟。因此,如何对各成员进行收益分配是合作联盟稳定的关键因素之一。从原始数据到数据资源再到数据产品需要经过多个步骤,涉及多个参与收益分配的主体。本文的主要研究问题是构建收益分配机制,旨在将数据产品增值后所产生的收益公平合理地分配给多个数据资源的持有者。此外,从收益分配制度的阶段来看,本文主要解决的是数据要素在初次分配阶段的相关问题。

数据产品的整合与投入数据要素市场需要在诸多前置条件下进行,结合本文的研究问题,提出以下三个设定。

设定1: 所有成员的数据资源均已完成权属认定,不存在数据未确权的情况。数据确权目前在一些领域还存有争议,本文不考虑成员在获取收益后是否将收益分配给创造原始数据的个人或者其他提出者(如果按照生成场景来确权,提供数据的个人具有收益权,但不在本文考虑的环节中),仅考虑合作联盟成员之间的收益分配。

设定2: 数据产品具有价值和使用价值,整合后的数据产品价值要高于所有成员单独向数据要素市场供给时的价值总和。如果跨域多源数据整合产生价值增值低于单独向数据要素市场供给时的价值总和,便失去了组建合作联盟的意义。这样的设定是为了更好地进行收益分配机制构建。事实上,只要是在效用上互补的跨域数据资源,所形成的数据产品均能够实现价值增值。

设定3: 所有提供数据资源的成员均是在同一时间加入到合作联盟中。加入时间对于合作联盟的形成而言是一个重要的因素,较早加入应该享受更多比例的分配,但本文考虑模型的简化,设定所有成员在同一时间加入联盟。实践中,依照客户需求定制一款数据产品时,通常情况下,符合条件的成员需要在规定的时间段内提供数据资源。

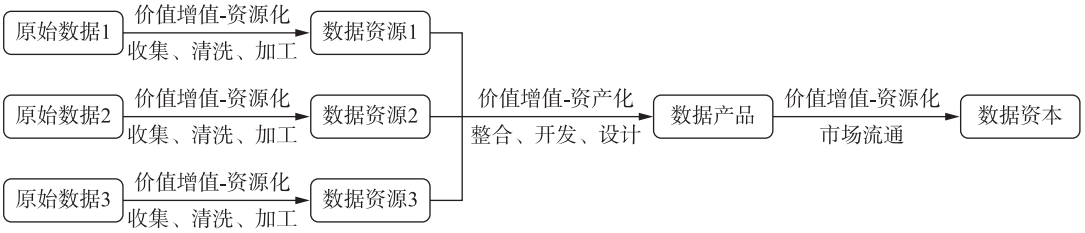


图1 数据要素价值增值过程

四、跨域多源数据收益分配机制构建

（一）合作博弈的适用性分析

供应链系统内的数据持有者可以单独出售数据资源，或者通过与其他数据持有者合作进行数据整合，实现数据资源的协同利用和价值增值。整合后的数据产品能够更好地满足市场需求，从而在市场竞争中占据优势地位。数据持有者均是以各自的利润最大化为目标，如何通过合理的收益分配机制激励数据持有者参与合作，并确保供应链的整体稳定性，成为构建跨域多源数据供应链的关键问题。数据持有者单独共享数据时不如与其他持有者进行数据整合后产生的收益高，因此数据持有者之间的收益分配问题可以使用合作博弈理论进行描述和解决。对于每个数据持有者而言，为整合后的数据产品带来价值增值是该持有者能够加入数据供应链的前提；同时，在加入数据供应链之后，持有者获取的收益至少不低于加入数据供应链之前获取的收益是维持数据供应链稳定性的保证。因此，数据供应链中持有者之间的收益分配问题可以看作是多人合作对策的收益分配问题，使用供应链成员之间的收益分配方法求解。具体而言，将这些数据持有者构成的合作博弈记为(N,v)，$N=\{1,2,\cdots,n\}$为数据持有者的集合，每个数据持有者可以单独决策是否参与合作。集合中的元素也称为局中人，v表示合作联盟的收益特征函数；$x=(x_1,x_2,\cdots,x_n)$为收益分配方案，x_i为成员$i\in N$所得的份额。合作联盟成立需要收益特征函数v满足以下几个原则：

（1）有效性原则。合作联盟存在$\sum_{i\in N}x_i=v(N)$，$x_i\geqslant v(i)$，$\forall i\in N$，$v(N)$为联盟总收益。数据产品获取的收益应该在所有局中人之间完全分配，局中人单独出售数据资源时的总收益小于他们之间合作的总收益。

（2）对称性原则。若局中人$i,k\subset N$且$S\subset N$，总有$v(S\cup\{i\})=v(S\cup\{k\})$，则$v(i)=v(k)$，表示合作博弈局中人的平等关系。这意味着只要是能为联盟提供同等价值的数据，在分配时将得到相同的收益。

（3）可加性原则。对于任何子集$S,T\subset N,S\cap T=\varnothing$，均有$v(S\cup T)=v(S)+v(T)$。实际上，数据供应链成员构成的合作博弈还存在超可加性，即对于任何子集$S,T\subset N$且$S\cap T=\varnothing$，均有$v(S\cup T)\geqslant v(S)+v(T)$。在此条件下合作能够产生协同效应[35]。超可加性描述的是合作比分离时获取的收益高，大联盟比小联盟收益高。跨域多源数据在一定范围内，相关数据资源的维度越高，整合后的数据产品价值越大。

满足以上3个原则的收益分配方案的解被认为在合作博弈(N,v)中是稳定的。在收益分配合作博弈中，所有稳定解的集合构成了该博弈的核。当合作博弈满足凸性条件时，可使用Shapley值法对成员进行收益分配，该方法的思想是按照联盟中成员边际贡献的比例进行分配收益，得到各方均认为公平、合理的唯一分配解，能够使合作联盟保持长期稳定。在此基础上，成员i的收益，即Shapley值$\varphi_i(N,v)$如式（1）所示。

$$\varphi_i(N,v)=\frac{1}{N!}\sum_{S\subset N\backslash i}|S|!\,(|N|-|S|-1)!\,[v(S\cup\{i\})-v(S)] \tag{1}$$

其中：$|S|!$为联盟S有多少种可能；$|S|!\,(|N|-|S|-1)!/N!$组合为加入成员i后的联盟S的权重；$[v(S\cup\{i\})-v(S)]$为联盟S加入成员i后的边际贡献。当满足对称性原则、有效性原则和可加性原则时，$\varphi_i(N,v)$具有唯一值。

供应链合作是一种较为稳定且高效的合作方式，通过整合各数据持有者的资源和能力有利于数据供应链整体竞争力的提高。鉴于跨域多源数据主体之间直接使用传统Shapley值法进行收益分配存在的局限性，结合本文的研究问题做出以下改进。一方面，由于在组建合作联盟时，Shapley值法是建立在所有参与者为全连接的基础上，即任意数量的参与者均可组成合作联盟。而实践中因各种因素的限制参与者之间不是全连接的状态，导致部分成员之间并不能形成合作联盟。因此，在可行联盟与不可行联盟的基础上，考虑参与者之间的连接关系，本文使用Myerson值法进行跨域多源数据整合的收益分配。另一方面，传统的Shapley值法仅考虑参与成员的边际贡献，无法设计激励和惩罚机制。在跨域多源数据整合中，数据要素的独特特征使得合作博弈的构建和稳定性具有显著差异性，需要结合数据要素特征进行改进以适应市场化的需要。因此，在合作博弈的收益分配模型中在融入基于数据要素特征的激励与惩罚机制，使各参与者更加积极参与数据整合，提升合作联盟的稳定性。

(二)基于限制性合作博弈的收益分配模型构建

在 Shapley 值的基础上,考虑参与者之间的连接关系,结合复杂网络理论,构建限制性合作博弈模型,以实现更合理的收益分配。Myerson 值法通过在合作博弈的特征函数上引入图结构,重新定义联盟的价值,从而调整收益分配,使其更符合实际的合作关系。在该模型中,参与者之间的连接关系可以用图论中的节点和边来表示,节点代表参与者,边表示参与者之间的直接合作关系。将参与者集合 N 上二元联盟的集合称为图,记为图 L,图 L 中的元素为两个联盟成员的边。集合 N 上的图 L 记为图(N,L)。$e=\{ij\}\in L$,表示节点 i 和节点 j 存在连边,i 和 j 分别是 e 两个端点,边 $\{ij\}$ 可以简记为 ij 或 ji。Myerson 值法定义了受限特征函数 (v^L),该函数考虑了参与者之间的连接关系,对不可行联盟的价值进行调整。并在此基础上计算每个参与者的边际贡献,最终确定其应得的收益分配。

具体而言,在合作博弈(N,v)和图(N,L)的基础上构成限制性合作博弈,(N,v,L)称为交流局面,全体组成的集合记为 CS_N。结合 Shapley 值并根据连边给出的分配方案 $\varphi(N,v^L)$ 即为 Myerson 值,记 $\mu_i(N,v,L)=\varphi_i(N,v^L)$。Myerson 值又被称为图限制合作博弈 (N,v^L) 的 Shapley 值,$\mu(N,v,L)$ 是由分支有效性和公平性唯一确定的分配法则,具体定义如下:

(1)分支有效性。对于任何 $(N,v,L)\in CS_N$,$T\in S/L$,均有 $\sum_{i\in T}\mu_i(N,v,L)=v^L(T)$ 成立,则称 $\mu(N,v,L)$ 是分支有效的,分配给每个分支中成员的支付之和等于该分支的效用。

(2)公平性。对于 $(N,v,L)\in CS_N$,任何 $e=\{ij\}\in L$ 均有 $\mu_i(N,v,L)-\mu_i(N,v,L\backslash e)=\mu_j(N,v,L)-\mu_j(N,v,L\backslash e)$ 成立。$L\backslash e$ 表示由 L 删去 e 这条边所得到的图,则称 $\mu(N,v,L)$ 是公平的,即直接连通的两个参与者 i 和 j,当它们的边 ij 断开之后,i 和 j 双方获得收益和损失是相同的。

(三)融合数据要素特征的收益分配模型构建

不管是 Shapley 值法还是 Myerson 值法,在刻画参与者的边际贡献时隐含了诸多因素,在跨域多源数据整合中,这些因素的具体权重和关系难以在单一特征函数中完全被刻画。因此,本节构建的基于数据要素特征的激励体系,旨在对理论模型中隐含假设的进一步显性化具体化处理,使其更契合跨域多源数据整合场景下的实际需求,增强模型的解释力和适用性。

结合数据要素特征及数据整合的特点,本文从三个核心维度构建激励指标体系,并融入到收益分配机制中。首先,考虑数据要素的非排他性影响了数据产品的稀缺性和竞争力,进而影响其收益。因此,将非排他性融合到收益分配的激励机制中,以保障数据持有者的权益及整合后数据产品的价值,从而激励更多数据资源的共享与开放。二级指标包括数据开放性、数据共享协议和数据访问权限。其次,单个数据资源的质量对于整合后的数据产品至关重要,需要设计数据质量相关指标[36],有助于鼓励持有者提供高质量、可信赖的数据资源,进一步提高整合后数据产品的市场竞争力。二级指标包括准确性、完整性、一致性、时效性、有效性、可追溯性等。最后,数据整合是多源数据协作的核心环节,需要数据持有者在数据治理过程中积极参与,以确保整合过程的高效性和最终数据产品的质量。二级指标包括规范性、易用性和稳定性等。表 1 列出了基于上述三个维度的激励指标体系,以及 12 个具体指标。这些指标为数据要素收益分配提供了理论依据和实践参考,但并非一成不变,根据不同的应用场景和需求,可对指标体系进行调整和扩展。

设定 R_i^{E} 为成员 i 获取的非排他性激励得分,R_i^{Q} 为成员 i 获取的数据质量激励得分,R_i^{A} 为成员 i 获取的数据整合激励得分,α 为非排他性的权重,β 为数据质量的权重,γ 为数据整合的权重,且 $\alpha+\beta+\gamma=1$,R_i 为成员 i 获取的总激励得分。计算公式如式(2)和式(3)所示。

$$R=\begin{bmatrix}R_1^{\mathrm{E}} & R_1^{\mathrm{Q}} & R_1^{\mathrm{A}}\\ \vdots & \vdots & \vdots\\ R_n^{\mathrm{E}} & R_n^{\mathrm{Q}} & R_n^{\mathrm{A}}\end{bmatrix}\begin{bmatrix}\alpha\\ \beta\\ \gamma\end{bmatrix}=\begin{bmatrix}R_1\\ \vdots\\ R_n\end{bmatrix} \tag{2}$$

$$R_i=\alpha R_i^{\mathrm{E}}+\beta R_i^{\mathrm{Q}}+\gamma R_i^{\mathrm{A}} \tag{3}$$

激励指标体系的权重反映了其在收益分配中的相对重要性,需依据具体问题和应用场景进行动态设置。各指标的权重设定有多种方法,如层次分析法(AHP)、模糊评价法、熵值法、TOPSIS 法、德菲尔法等。具体

表 1　融合数据要素特征的激励指标体系

一级指标	二级指标	指标解释	指标来源
非排他性(R_i^{E})	数据开放性($R1$)	数据是否向公众或特定群体开放	文献[37-41]
	数据共享协议($R2$)	是否有明确的数据共享条款及限制性	文献[13,42-45]
	数据访问权限($R3$)	哪些用户可以访问数据,以及访问的难以程度	文献[46-48]
数据质量(R_i^{Q})	准确性($R4$)	数据内容真实、表述清晰明确、无错误、无歧义	文献[46-51]
	完整性($R5$)	数据集的完整程度	文献[46-50,52-53]
	一致性($R6$)	同一数据在不同时间保持一致的程度	文献[46-47,49,52]
	时效性($R7$)	数据更新时间和更新频率	文献[48,50,53-54]
	有效性($R8$)	数据对预期用途具有相关性	文献[38,50]
	可追溯性($R9$)	数据从采集到存储等过程种涉及到的主题、时间等数据信息能够被追踪的程度	文献[40-41,55-56]
数据整合(R_i^{A})	规范性($R10$)	数据在格式上的标准化程度	文献[49,52,57]
	易用性($R11$)	为数据使用提供便利的程度	文献[37,49,51,57]
	稳定性($R12$)	数据供给的稳定性,不易出错的程度	文献[47,50,51,53,57]

计算细节本文不再列出。

借鉴已有关于 Shapley 值法的改进思路[28,58],设定 θ 为数据供应链上所有数据持有者均认可的激励基数。在当前大数据和人工智能快速发展的背景下,算力和算法的提升使得数据处理能力显著增强,数据量的大小已不再是限制数据处理能力的关键因素。数据治理特别是数据的规范性和标准化,在数据整合过程中尤为重要。规范的数据不仅有助于提高数据整合的效率和质量,还能降低数据清洗和转换的成本,促进数据价值的充分挖掘。因此,在数据量相同维度相差不是特别大的情况下,对所有的成员使用相同的激励标准是合理的。基于此,成员 i 通过激励获取的得分占所有激励得分的份额为 $R_i \Big/ \sum R_i$,成员 i 的收益调整如式(4)所示。

$$\varphi_i(N,v)' = \varphi_i(N,v) + \theta \times \left(\frac{R_i}{\sum R_i} - \frac{1}{N} \right) \tag{4}$$

其中:$\varphi_i(N,v)$ 为 Shapley 值;$\varphi_i(N,v)'$ 为改进的 Shapley 值。当 $R_i \Big/ \sum R_i > 1/N$ 时,成员 i 因非排他性、数据质量和数据整合获得正激励收益;当 $R_i \Big/ \sum R_i < 1/N$ 时,成员 i 因非排他性、数据质量和数据整合获得负激励收益,受到一定程度的惩罚;当 $R_i \Big/ \sum R_i = 1/N$ 时,该成员的收益与改进前相同。该激励惩罚机制下合作联盟的总收益并无变化如式(5)所示。

$$\sum \varphi_i(N,v)' = \sum \varphi_i(N,v) + \theta \times \frac{\sum R_i}{\sum R_i} - \theta \times N \times \frac{1}{N} = \sum \varphi_i(N,v) \tag{5}$$

在此基础上,计算合作博弈的 Myerson 值,则参与主体 i 在合作联盟中的收益分配如式(6)所示。

$$\mu_i(N,v,L)' = \varphi_i(N,v^L)' \tag{6}$$

五、实例检验与对比

(一)考虑数据要素特征的合作博弈解

4 个不同的数据持有者($i = 1,2,3,4$),提供的数据资源在整合成数据产品过程中能够实现资源上的互补。单独出售数据资源和组成合作联盟整合后出售数据产品的收益如表 2 所示,$v(S)$ 表示联盟 S 的总收益。

表 2　$v(S)$ 的值

S	{1}	{2}	{3}	{4}
$v(S)$	8	7	12	19
S	{1,2}	{1,3}	{1,4}	{2,3}
$v(S)$	20	30	40	25
S	{2,4}	{3,4}	{1,2,3}	{1,2,4}
$v(S)$	35	45	60	70
S	{1,3,4}	{2,3,4}	{1,2,3,4}	{ø}
$v(S)$	80	75	100	0

假定该合作博弈的激励收益系数 θ 取值为 100,激励指标体系中一级指标的权重为(0.22,0.5,0.28),二级指标的权重分别为[(0.3,0.5,0.2),(0.2,0.2,0.2,

0.2,0.2,0.2),(0.4,0.3,0.3)],各指标均采用百分制,经过计算各成员的最终激励得分为(65.8,74.45,85.96,95.35),激励指标体系得分见表 3。

本实例中参与合作博弈的成员有 4 个,可形成 16(2^4)种连接结构。为简化计算,仅列出同种类型的图,各图边的集合见表 4,图的连接结构如图 2 所示。

根据联盟收益(表 2)、各指标激励得分(表 3)和成员之间的连接结构(表 4 和图 2),得出各成员的最终收益,如表 5 所示。L^8 为各成员全连接下的结果,与 Shapley 值法的分配方案相同,其余均为限制性合作博弈下的结果。

表 3　激励指标体系得分

一级指标	二级指标	成员 1	成员 2	成员 3	成员 4
$R_i^{\mathrm{E}}(0.22)$	$R_1(0.3)$	78	63	92	95
	$R_2(0.5)$	60	56	79	95
	$R_3(0.2)$	68	91	49	73
$R_i^{\mathrm{Q}}(0.5)$	$R_4(0.2)$	59	86	60	94
	$R_5(0.2)$	62	43	88	83
	$R_6(0.2)$	67	69	57	82
	$R_7(0.2)$	65	96	89	85
	$R_8(0.2)$	75	45	88	92
	$R_9(0.2)$	46	75	99	83
$R_i^{\mathrm{A}}(0.28)$	$R_{10}(0.4)$	65	54	76	90
	$R_{11}(0.3)$	46	87	64	71
	$R_{12}(0.3)$	30	64	84	89
最终得分		65.8	74.45	85.96	95.35

表 4　图 L 上的集合

L	L^1	L^2	L^3	L^4
边的集合	{1:2, 3:4}	{1:2, 1:3}	{1:2, 1:3, 2:3}	{1:2, 2:4, 1:3}
L	L^5	L^6	L^7	L^8
边的集合	{1:2, 1:3, 2:3, 2:4}	{1:2, 3:4, 2:4, 1:3}	{1:2, 1:4, 2:3, 2:4, 3:4}	{1:2, 1:3, 1:4, 2:4, 3:4, 2:3}

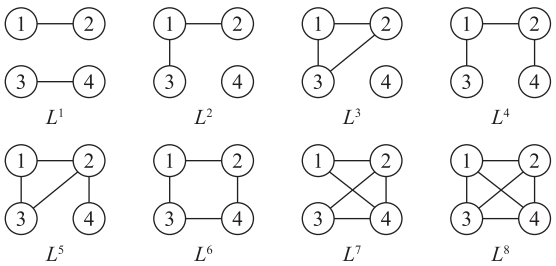


图 2　图 L 的结构示意图

表 5　图 L 基础上的 Myerson 值

L	L^1	L^2	L^3	L^4	L^5	L^6	L^7	L^8
{1}	8	8	8	8	8	8	8	8
{2}	7	7	7	7	7	7	7	7
{3}	12	12	12	12	12	12	12	12
{4}	19	19	19	19	19	19	19	19
{1, 2}	20	20	20	20	20	20	20	20
{1,3}	20	30	30	30	30	30	20	30
{1,4}	27	27	27	27	27	27	40	40
{2,3}	19	19	25	19	25	19	25	25
{2,4}	26	26	26	35	35	35	35	35
{3,4}	45	31	31	31	31	45	45	45

续表

L	L^1	L^2	L^3	L^4	L^5	L^6	L^7	L^8
{1,2,3}	32	60	60	60	60	60	60	60
{1,2,4}	39	39	39	70	70	70	70	70
{1,3,4}	53	49	49	49	49	80	80	80
{2,3,4}	52	38	44	47	75	75	75	75
{1,2,3,4}	65	79	79	100	100	100	100	100
Myerson 值	(10.5,9.5,19,26)	(21.5,15.5,23,19)	(19.5,16.5,24,19)	(26.33,24.83,20.5,28.33)	(18.33,27.66,23.33,30.17)	(20.75,18.25,26.58,34.42)	(20.5,18.5,25.17,35.83)	(21.33,17.67,26,35)
加入激励后的 Myerson 值	(5.96,7.65,20.73,30.65)	(16.96,13.65,27.73,23.65)	(14.96,14.65,25.73,23.65)	(21.79,22.98,22.23,32.98)	(13.29,25.81,25.06,34.82)	(16.21,16.4,28.31,39.07)	(15.96,16.65,26.9,40.48)	(16.79,15.82,27.73,39.65)

(二)解的比较

从表 5 可以看出图 L^1、L^2 和 L^3 均未达到联盟收益的最高值,其余的情况均实现了整个联盟收益的最高值(100)。各种合作联盟的收益分配结果存在显著差异,但均超过了单独共享数据资源所带来的收益。在无激励机制的情况下,比较 Shapley 值法和 Myerson 值法的结果。在 L^2 与 L^3 两种情况中,虽然合作联盟总收益一致,但各成员的收益分配却存在差异。主要是因为成员 1、2、3 之间存在连接关系,成员 4 与其他成员不存在连接关系所导致,成员 4 的收益仅仅来源于自身的高价值数据资源的增值。L^4 ~L^8 这五种情况中,任何两个成员之间均有路径实现互通,因此达到了合作联盟收益最大值,但在收益分配方面仍存在差异。具体而言,成员 2 的收益在 L^5 中取得最大值,主要原因是成员 2 在网络结构中占据了重要地位。成员 2 与联盟内其他三个成员均存在直接连接,特别是作为成员 4 的唯一邻居节点,起到了关键的桥梁作用。进一步地,成员 4 与成员 1、2 之间存在明显的结构洞,而成员 2 属于“结构洞占据者”,具有战略位置的优势。由于结构洞的存在,成员 2 能够更高效地获取信息和资源,与其他成员展开合作,同时在数据整合中产生更高的价值增值,从而在合作联盟中实现了收益的最大化。这一结果表明,在复杂网络中,节点的连接属性和结构位置对收益分配具有显著影响,特别是结构洞的占据者往往能够在合作中获取更大的收益,这为收益分配的优化提供了重要的理论依据。同时也表明虽然大联盟能够形成,但是在联盟内部部分成员之间没有直接的连接关系,导致这些成员所提供的数据资源价值未能充分释放及产生价值增值,从而收益分配受到影响。此外,该分配方法仅考虑了数据持有者的边际贡献,忽略了数据要素的其他关键特性。这种单一维度的考量可能导致部分成员为了追求更高的收益,而忽视数据要素的非排他性及数据治理的重要性,不利于数据要素市场的长期可持续发展。

在引入基于数据要素特征的激励机制情况下,对比 Myerson 值法前后的收益分配结果,发现收益分配的变化主要体现在激励得分的影响上。激励机制的作用并不是仅仅取决于自身的得分,还受到整个合作联盟全局的制约和影响。以 L^2 与 L^3 为例,成员 4 在无激励机制时,其收益仅来源于自身提供的高价值数据资源。然而在激励机制的作用下,成员 4 通过非排他性、数据质量和数据整合等方面所做出的积极贡献,其收益得到了显著提高。这表明,本节构建的激励机制能够有效鼓励数据持有者通过多维度的努力提升整体价值。同时,成员 1、2 的收益变化进一步说明忽视非排他性、数据质量和数据整合的重要性,其收益将会受到明显限制。激励机制引入后,按照资源和贡献率的收益分配结果将再次被修正,还能够对部分数据持有者起到了约束作用,从而使得收益分配更加客观和公平。在数据要素边际成本几乎为零的前提下,这种机制有效避免了具有数据资源优势的持有者加入合作联盟后再次将数据共享到数据要素市场或其他需求者的现象,从而增强合作联盟整合数据产品的竞争力,维护其他数据持有者的合法权益,保障联盟整体收益的最大化。实际上,当数据持有者的激励得分低于联盟的平均水平时,激励部分的数值将是负数,此时可以理解为对数据持有者的一种惩罚机制。在这种情况下,激励机制不仅能够促进收益分配的公平性,还能矫正数据持有者的行为导向,激励其更加关注非排他性、数据质量和数据整合,从而为合作联盟做出更大的贡献。这种双向调节的机制,既能奖励优质贡献,又能约束消极行为,为合作联盟的长期稳定与可持续发展提供了重要保障。

六、结论与展望

在跨域多源数据整合过程中,存在“价值共创、贡献难辨”的特点,导致收益分配结果难以体现公平性、合理性。在此背景下,本文将数据要素的非排他性特征、数据质量和数据整合等因素融合到Shapley值法,并在此基础上使用Myerson值法作为数据合作联盟的收益分配方法。该方法不仅考虑了成员提供的数据资源的边际贡献,还将非排他性、数据质量和数据整合等因素作为激励机制融合到收益分配机制中,引导数据持有者重视非排他性、提升数据质量、加强数据治理等,促进数据要素价值的充分释放。基于改进的Myerson值法表明,网络结构不同导致成员对最终数据产品价值创造的贡献度存在差异,激励机制显著影响收益分配结果,从而形成更加公平合理的收益分配格局。在限制性合作博弈中,占据“结构洞”位置的成员能够与更多成员形成小合作联盟,具有更高的边际贡献,从而在各种情境下获取最大收益。在数据资源既定前提下,持有者可以通过激励机制优化其贡献方式以获取更多的收益。特别地,数据持有者应该认识到非排他性的重要性,因为数据要素的边际生产成本几乎为零,共享给更多的需求者并不会带来额外的成本,但是会造成整合后的数据产品竞争力下降,损害合作联盟的整体利益。因此,数据持有者应该与其他类似组织或机构建立连接,加强沟通,同时重视非排他性、数据质量和数据整合等激励因素,以便在合作联盟中获取更多的收益。本文所构建的合作博弈解,为跨域多源数据收益分配提供了一般性的分配方法,在解决收益分配问题方面具有合理性和有效性,并具有广泛的应用价值。在实践应用中,该方法能够平衡参与者的利益诉求,增强合作联盟的稳定性,为数据要素市场的规范化发展提供理论依据和实践指导。

在未来研究中,将此合作博弈解运用到具体的跨域多源数据收益分配机制时,可根据实际需要,进一步探讨数据要素特征对收益分配的影响。同时,可设计更加精细化的激励指标体系,并探索更加科学合理的权重设置方法。此外,还可将成员加入联盟的先后顺序、不同成员的风险偏好等因素纳入模型改进中,构建更合理的分配机制以保障各方利益,促进数据要素市场的健康和可持续发展。

参考文献

[1] 林春, 王子绮, 张雯婧, 等. 数据要素市场化赋能银行数字化转型了吗?——来自中国城市层面的经验证据[J]. 技术经济, 2024, 43(9): 56-71.

[2] 陈舟, 郑强, 吴智崧. 我国数据交易平台建设的现实困境与破解之道[J]. 改革, 2022(2): 76-87.

[3] 王艳, 杨达. 中国式管理会计体系变革: 从数据要素到数据资产[J]. 管理世界, 2024, 40(10): 171-189.

[4] ALHARTHI A, KROTOV V, BOWMAN M. Addressing barriers to big data[J]. Business Horizons, 2017, 60(3): 285-292.

[5] 刘洋, 董久钰, 魏江. 数字创新管理: 理论框架与未来研究[J]. 管理世界, 2020, 36(7): 198-217, 219.

[6] SCHILDHAUER M. Data integration: principles and practice[M]//RECKNAGEL F, MICHENER W K. Ecological Informatics: Data Management and Knowledge Discovery. Cham: Springer International Publishing, 2018: 129-157.

[7] 王子玄, 黄倩倩, 任明. 联邦计算中的数据交易预期收益分配模型[J]. 现代情报, 2024, 44(11): 108-117.

[8] YAN R. Profit sharing and firm performance in the manufacturer-retailer dual-channel supply chain[J]. Electronic Commerce Research, 2008, 8(3): 155-172.

[9] XIN L D, SRIVASTAVA D. Big data integration[M]. Cham: Springer International Publishing, 2015.

[10] GOODHUE D L, WYBO M D, KIRSCH L J. The impact of data integration on the costs and benefits of information systems[J]. MIS Quarterly, 1992, 16(3): 293-311.

[11] 欧阳日辉. 数字基础设施促进电子商务创新发展的机理与路径[J]. 广西社会科学, 2024(1): 1-11.

[12] 孔艳芳, 刘建旭, 赵忠秀. 数据要素市场化配置研究: 内涵解构、运行机理与实践路径[J]. 经济学家, 2021(11): 24-32.

[13] 黄甄铭, 魏娜, 梁正. 跨部门数据共享源于机构变革还是动机转变?——基于浙江省M区的案例研究[J]. 管理世界, 2024, 40(5): 87-106.

[14] VAN PANHUIS W G, PAUL P, EMERSON C, et al. A systematic review of barriers to data sharing in public health[J]. BMC Public Health, 2014, 14(1): 1144.

[15] FOLEY P, ALFONSO X, SAKKA M A. Information sharing for social inclusion in England: A review of activities, barriers and future directions[J]. Journal of Information, Communication and Ethics in Society, 2006, 4(4): 191-203.

[16] MCGUIRK P M, O'NEILL P M, MEE K J. Effective practices for interagency data sharing: Insights from collaborative research in a regional intervention[J]. Australian Journal of Public Administration, 2015, 74(2): 199-211.

[17] 中国信息通信研究院. 数据要素白皮书(2023年)[M]. 北京: 中国信息通信研究院, 2023: 32-33.

[18] 宋新平，王笑，李明星，等. 大数据驱动企业竞争情报感知能力的形成机理——基于华为公司的案例研究[J]. 情报科学，2024，42(8)：154-163.
[19] 李明堂. 数字政府治理集约化的逻辑、挑战与进路[J]. 东岳论丛，2024(2)：124-133.
[20] 林镇阳，陈荣源，郭明军，等. 多元主体协同治理的数据要素价值生态体系研究[J]. 技术经济，2024，43(11)：1-13.
[21] 周颖玉，柯平，刘海鸥. 融合·赋能·共创：区块链嵌入算法推荐的未来发展图景[J]. 图书情报知识，2024，41(3)：144-153.
[22] 王衍之，黄静思，王剑晓，等. 数据要素流通与收益分配机制研究：以风电场景融合气象数据为例[J]. 管理评论，2024，36(6)：30-41.
[23] 熊巧琴，汤珂. 数据要素的界权、交易和定价研究进展[J]. 经济学动态，2021(2)：143-158.
[24] GLAZER R. Measuring the value of information：The information-intensive organization[J]. IBM Systems Journal，1993，32(1)：99-110.
[25] LERNER J，PATHAK P A，TIROLE J. The dynamics of open-source contributors[J]. American Economic Review，2006，96(2)：114-118.
[26] 彭正银，孙莹，李妍，等. 基于 Myerson 值优化的企业创新网络共创价值的分配研究[J/OL]. 中国管理科学：1-12.[2022-09-14]. https://doi.org/10.16381/j.cnki.issn1003-207x.2022.0456.
[27] DYER J H，SINGH H，HESTERLY W S. The relational view revisited：A dynamic perspective on value creation and value capture[J]. Strategic Management Journal，2018，39(12)：3140-3162.
[28] 马士华，王鹏. 基于 Shapley 值法的供应链合作伙伴间收益分配机制[J]. 工业工程与管理，2006(4)：43-45，49.
[29] MYERSON R B. Graphs and cooperation in games[J]. Mathematics of Operations Research，1977，2(3)：225-229.
[30] 李泉林，张玉，鄂成国. 多个农民专业合作社和多个超市的网络博弈研究[J]. 系统工程学报，2019，34(1)：29-45.
[31] 蔡蕾，史纪磊，李理，等. 基于图限制合作博弈有效分配原则的医疗联合体利益分配机制[J]. 系统管理学报，2021，30(2)：393-400.
[32] 樊自甫，程姣姣. 基于微分博弈的数据开放策略及合作收益分配机制研究[J]. 运筹与管理，2021，30(12)：100-107.
[33] 于施洋，黄倩倩，虞洋，等. 数据要素市场的价值增值研究：理论构建与实施路径[J]. 电子政务，2024(2)：33-40.
[34] 李海舰，赵丽. 数据成为生产要素：特征、机制与价值形态演进[J]. 上海经济研究，2021(8)：48-59.
[35] 刁丽琳，朱桂龙，许治. 基于多权重 Shapley 值的联盟利益分配机制[J]. 工业工程与管理，2011，16(4)：79-84，91.
[36] 刘冰，庞琳. 国内外大数据质量研究述评[J]. 情报学报，2019，38(2)：217-226.
[37] 谭必勇，陈艳. 我国开放政府数据平台数据质量研究——以十省、市为研究对象[J]. 情报杂志，2017，36(11)：99-105.
[38] 张晓娟，谭婧. 我国省级政府数据开放平台元数据质量评估研究[J]. 电子政务，2019(3)：58-71.
[39] 李晓彤，翟军，郑贵福. 我国地方政府开放数据的数据质量评价研究——以北京、广州和哈尔滨为例[J]. 情报杂志，2018，37(6)：141-145.
[40] 刘于蓝，张艳丰，马家伟，等. 我国省级政府数据开放平台信息服务质量测度实证研究[J]. 情报科学，2024，42(9)：165-177.
[41] 翟军，陶晨阳，李晓彤. 开放政府数据质量评估研究进展及启示[J]. 图书馆，2018(12)：74-79.
[42] 涂淼，陈晶. 运营商数据治理体系和数据安全探究[J]. 通信与信息技术，2024(3)：120-123，132.
[43] 王新建，胡广伟. 数智环境下价值网络对社区养老服务精准化的影响机理研究——基于扎根理论的探索[J]. 现代情报，2025，45(1)：150-163.
[44] 黄楠，李竹伊，李雪岩，等. 管理学研究中的大数据：科学问题与未来发展[J]. 南开管理评论，2024，27(7)：74-85.
[45] 孟雪，郝文强，吴钟灿. 价值、组织与过程：数据开放共享隐私风险管理体系建构——基于英国的经验考察与借鉴[J]. 情报理论与实践，2024，47(6)：184-194.
[46] CAI L，ZHU Y. The challenges of data quality and data quality assessment in the big data era[J]. Data Science Journal，2015，14：2.
[47] WANG R Y，REDDY M P，KON H B. Toward quality data：An attribute-based approach[J]. Decision Support Systems，1995，13(3/4)：349-372.
[48] MILLER H. The multiple dimensions of information quality[J]. Information Systems Management，1996，13(2)：79-82.
[49] WANG F，ZHAO A，ZHAO H，et al. Building a holistic taxonomy model for OGD-related risks：Based on a lifecycle analysis[J]. Data Intelligence，2019，1(4)：309-332.
[50] 莫祖英. 大数据质量测度模型构建[J]. 情报理论与实践，2018，41(3)：11-15.
[51] 查先进，陈明红. 信息资源质量评估研究[J]. 中国图书馆学报，2010，36(2)：46-55.
[52] 郭路生，刘春年. 大数据时代应急数据质量治理研究[J]. 情报理论与实践，2016，39(11)：101-105.
[53] WANG F. Understanding the dynamic mechanism of interagency government data sharing[J]. Government Information Quarterly，2018，35(4)：536-546.
[54] VETRÒ A，CANOVA L，TORCHIANO M，er al. Open data quality measurement framework：Definition and application to open government data[J]. Government Information Quarterly，2016，33(2)：325-337.
[55] 王芳，储君，张琪敏，等. 跨部门政府数据共享：问题、原因与对策[J]. 图书与情报，2017(5)：54-62.
[56] 曹海军，李明. 中国政府数据开放平台服务质量评价——基于熵权 TOPSIS 的实证分析[J]. 上海行政学院学报，2020，21(4)：55-64.
[57] FANG F，YU S，LIU M. An improved shapley value-based profit allocation method for CHP-VPP[J]. Energy，2020，213：118805.

Cross-domain and Multi-source Data Revenue Distribution Mechanism Based on Restricted Cooperative Game Theory

Xue Xiaofei[1,2], Xiao Yi[1], He Chen[1], Zhang Jin'ao[1]

(1. School of Information Management, Central China Normal University, Wuhan 430070, China; 2. Management Engineering School, Henan University of Engineering, Zhengzhou 451191, China)

Abstract: Data elements, as an important new quality productive forces in the digital economy era, the revenue distribution mechanism is the key to promoting the healthy development of the data element market. Current work on cross-domain and multi-source data revenue distribution mechanism is still in its initial stage and lacks mature solutions. Based on cooperative game theory, the non-exclusivity of data elements, data quality, data integration, and other factors were incorporated as incentive mechanisms into the Shapley value method. On this basis, the connection relationship between members was combined to propose an improved Myerson value method as the solution of cooperative game to solve the cross-domain and multi-source data revenue distribution problem. The results show that the network structure has a significant impact on the value creation, marginal contribution and revenue distribution of data resources. The incentive mechanism can effectively adjust the revenue distribution between members and finally form a fairer and more reasonable revenue distribution pattern. New ideas and methods are provided for the cross-domain and multi-source data revenue distribution mechanism, which have important theoretical value and practical significance for constructing a perfect data element market trading system, realizing the value release of data elements and promoting the high-quality development of the digital economy.

Keywords: data element; revenue distribution; cross-domain and multi-source data; restricted cooperative game; Myerson value