

DOI: 10.16538/j.cnki.fem.20230719.401

# 数据价值化：研究评述与展望

朱秀梅<sup>1</sup>, 林晓玥<sup>2</sup>, 王天东<sup>3</sup>, 苗淑娟<sup>1</sup>

(1. 吉林大学 商学与管理学院, 吉林 长春 130012; 2. 哈尔滨工业大学 经济与管理学院, 黑龙江 哈尔滨 150001; 3. 哈尔滨工程大学 经济管理学院, 黑龙江 哈尔滨 150001)

**摘要:** 数字经济时代, 如何释放数据价值已经成为实业界和学术界关注的焦点和热点。数据价值化以数据资源化、数据资产化、数据商品化和数据资本化为核心阶段, 从动态演化的新视角揭示了数据创造价值的过程。目前相关研究比较分散, 缺乏系统梳理, 尚未形成对数据价值化的完整认知。基于此, 本文系统回顾和梳理数据价值化相关研究, 首先系统梳理2011至2022年发表的250篇英文文献和117篇中文文献, 科学呈现和对比中西方有关数据价值化的细分研究主题; 其次整合已有研究明晰数据价值化的概念内涵和特征; 接着基于价值链理论构建数据价值化的整合研究框架, 明确数据资源化、数据资产化、数据商品化和数据资本化四个阶段的关键数据行为, 以及四个阶段的前置因素和结果产出; 最后提出未来值得研究的五大议题。本研究明确了数据价值化的知识脉络和研究边界, 有助于推动数据价值化的进一步研究。

**关键词:** 数据价值化; 文献综述; 整合研究框架; 未来展望

**中图分类号:** F270 **文献标识码:** A **文章编号:** 1001-4950(2023)12-0003-15

## 一、引言

数据是数字经济时代的战略性资源和重要生产力, 作为新兴生产要素, 已和其他要素一起融入经济价值创造过程之中, 不断催生新业态、新产业、新模式, 并推动产业加速升级。2023年初国务院提请审议组建国家数据局, 意味着数据要素的转化已经上升到国家战略高度。数据价值化、数字产业化、产业数字化、数字化治理共同组成了数字经济的“四化框架”(中国信息通信研究院在2021年发布的《中国数字经济发展白皮书》), 其中, 数据价值化成为核心驱动因素。然而国际知名咨询公司埃森哲发布的研究报告显示, 三分之二以上的企业对于如何用好数据仍然一头雾水。与此同时, 我国数据要素市场化配置改革刚刚起步, 许多尝试犹如“摸着石头过

收稿日期: 2023-04-17

基金项目: 国家自然科学基金面上项目(71972086); 国家自然科学基金重大项目(72091310, 72091315); 黑龙江省自然科学基金项目(LH2022G005)

作者简介: 朱秀梅(1975—), 女, 吉林大学商学与管理学院教授, 博士生导师;

林晓玥(1996—), 女, 哈尔滨工业大学经济与管理学院助理研究员(通讯作者, 1079809837@qq.com);

王天东（1989—），男，哈尔滨工程大学经济管理学院讲师；  
苗淑娟（1969—），女，吉林大学商学与管理学院副教授。

河”，亟待以系统性、创新性的理论研究指引企业数据价值化实践，数据价值化研究也成为激发学者深入探讨的“新蓝海”。

在理论层面，学者们从不同角度出发对数据为什么创造价值及如何创造价值进行讨论，但这些研究分散于多个学科、多个领域，缺少能够“穿针引线”的整合研究。西方学者率先对数据的特征和数据管理问题展开讨论（Günther等，2017），并结合供应链管理（Brinch，2018）、营销管理（Johnson等，2019）、人力资源管理（Garcia-Arroyo和Osca，2021）等具体场景讨论数据的作用和影响。我国学者则特别关注数据形态的演化问题（李海舰和赵丽，2021），并围绕数据市场化流通问题展开研究（何玉长和王伟，2021）。总体看，数据价值化的中西方研究成果侧重点有差异，关于数据价值化的内涵和特征、数据价值实现过程等关键问题并没有形成清晰完整的回答。为此，本文拟在文献梳理的基础上融汇已有观点，尝试回答以上关键问题，以响应学界和业界关于数据理论创新的急呼。

基于此，本文以建构式综述的形式，从四个方面推进数据价值化研究。第一，系统回顾和梳理已有研究，明确已有成果的研究脉络和研究主题。利用Cite Space5.8R3软件对与数据价值化相关的250篇英文文献和117篇中文文献进行量化分析，科学呈现和对比中西方研究主题。第二，分析数据价值化的内涵，明晰数据价值化研究的范围和边界。数据价值化作为数据转化为价值的关键过程，能够整合现有较为分散和独立的数据相关研究，并成为极具价值和挑战性的科学问题。本文结合来源、参与主体、过程、目标结果等价值创造要素提出数据价值化的完整内涵，并比较数据要素和传统要素在对象、范围、效益、模式、创新等方面的差异。第三，构建数据价值化的整合研究框架，推动数据价值化理论体系的构建。中国信息通信研究院2021年发布的《中国数字经济发展白皮书》指出“数据价值化”包括但不限于数据采集、数据确权、数据定价、数据交易等，而后学术研究引用和发展了这一观点，提出数据资源化、数据资产化、数据商品化和数据资本化四个阶段。本文整合现有研究，利用价值链理论，揭示数据资源化、数据资产化、数据商品化、数据资本化四个阶段的关键数据行为，并依据“前置因素—实现过程—结果产出”的研究逻辑建立数据价值化的整合研究框架。第四，提出未来需要深入探究的五大研究议题。

总的来看，数据价值化的相关研究数量多、涉及面广，但鲜有研究对此进行整合、总结和提升，研究脉络和研究重点不甚清晰，缺少高质量的综述性学术成果。本文旨在系统展示中西方数据价值化研究主题，并对比分析研究的差异及其原因，进而提出未来值得研究的问题，从而对现有研究进行归纳、总结和提升，以启发和引领后续研究。

## 二、数据价值化研究的量化分析

为了全面把握已有研究的进展，本文首先以“big data”“data driven”为关键词，组合“value”等关键词，以“article”“review”为文献类型在Web of Science、ProQuest、EBSCO等英文数据库进行搜索，以“数据”与“价值”“赋能”“驱动”为组合关键词在中国知网进行搜索，进行文献初筛。第二步，在时间方面，由于与“大数据”相关的期刊文章在2011年开始频繁出现，因而重点审查2011—2022年发表的文献，这一时间范围也与大数据广泛应用的时间相吻合（Wamba等，2015）。在领域方面，由于数据价值化具有明显的跨学科特质，因而广泛搜索管理学、经济学、计算机科学、信息系统等领域的文献。英文文献剔除非核心合集和开放获取的文献，中文文献保留期刊来源，共获得英文文献971篇，中文文献315篇。第三步，阅读每篇文献的摘要和结论，以题目、摘要或关键词能否反映出“数据价值化”这一研究主题为标准评估文献的主题匹配性和



内容适宜性，并进一步筛查和补充上述文献的参考文献，最终共得到英文文献250篇，中文文献117篇。

### (一) 数据价值化英文文献分析

本文利用Cite Space5.8R3软件分析了数据价值化英文文献关键词聚类图谱和关键词突现情况(见图1)。

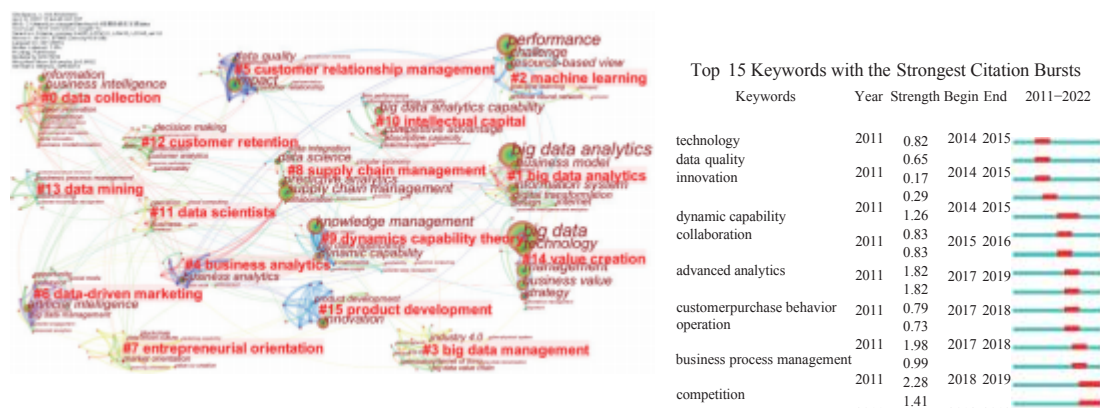


图1 英文文献关键词聚类图谱和关键词突现分析

文献分析后生成15个聚类，这些聚类是确定研究主题的基础，包括数据收集（data collection）、机器学习（machine learning）、大数据管理（big data management）、商业分析（business analytics）、用户关系管理（customer relationship management）、数据驱动营销（data-driven marketing）、创业导向（entrepreneurial orientation）、供应链管理（supply chain management）、动态能力理论（dynamics capability theory）、智力资本（intellectual capital）、数据科学家（data scientists）、用户留存（customer retention）、数据挖掘（data mining）、价值创造（value creation）和产品开发（product development）。聚类模块值（modularity） $Q=0.7618>0.3$ ，聚类结构显著，聚类平均轮廓值（weighted mean silhouette） $S=0.8273>0.7$ ，聚类非常合理。

由关键词突现分析结果可知，数据价值化英文研究的演变路径为：数据自身的管理问题→数据在具体场景中的应用。据此，本文将数据价值化英文文献的具体研究内容归纳为以下两大研究主题。

#### 1. 数据自身的管理问题

LaValle等（2011）较早地注意到数据经过分析后能够产生洞察力指导行动进而创造价值，可视作为早期数据价值化研究的雏形。数据自身的管理问题涉及数据行为、大数据管理、价值创造、大数据能力等细分问题。（1）数据行为。现有研究常基于价值链理论讨论从数据获取到数据决策的整个数据管理过程中的数据行为（Miller和Mork，2013），涉及数据生成、数据采集、数据预处理、数据存储、数据分析、数据可视化和数据利用等（Janssen等，2017；Faroukhi等，2020a；b）。（2）大数据管理。学者们主要围绕数据特征以及隐私与安全等问题展开讨论。数据特征被认为是影响价值实现的关键技术特征，包括数据数量、数据速度、数据多样性、数据准确性、便携性和互联性等（Günther等，2017；Ghasemaghaci和Calic，2019）。同时，大数据平台的安全和隐私执法不力被认为是企业激发数据价值的主要障碍（Colombo和Ferrari，2015），可能产生包括主体歧视等13个数据伦理挑战和社会风险（Clarke，2016；Breibach和Maglio，2020）。（3）数据价值创造。部分学者从不同视角构建了大数据价值创造的模型，例如Comuzzi和Patel（2016）建立

支持企业利用大数据创造价值的五层次成熟度模型；Günther等（2017）构建大数据价值实现的集成模型，指出数据价值实现是工作实践、组织模式和利益相关者之间持续互动的结果；

Ylijoki和Porras (2019) 改进了DIKW (data-information-knowledge-wisdom, 数据—信息—知识—智慧) 和VVC (virtual value chain, 虚拟价值链) 模型, 提出大数据价值创造过程模型, 包含资产创造、能力创造、转化、竞争等过程; Wang和Wang (2020) 开发了大数据和知识管理协同关系的概念模型; Santarsiero等 (2023) 提出了大数据与智力资本交互并增强组织价值创造的概念框架。(4) 大数据能力。大数据能力是数据发挥生产要素作用, 改善微观运行效率和提高企业创新绩效的必要条件 (谢康等, 2020), 其细分能力包括大数据分析能力、大数据预测能力、商业智能能力等 (Yasmin等, 2020)。

## 2. 数据在具体场景中的应用

有些学者将数据价值化置于战略导向、供应链管理、用户管理、营销管理、业务流程管理、人力资源管理、数字技术创新等具体场景中进行讨论。(1) 战略导向。学者们已经发现数据能够驱动企业战略导向发生改变, 从而更有效地产生和捕获市场价值, 例如Aker等 (2021) 系统综述了国际企业大数据战略导向研究, 确定了数据驱动战略导向的三个中心主题: 数字平台导向、市场导向、创新和创业导向, 并建立大数据战略导向的整体框架。(2) 供应链管理。主要涉及供应链管理中的大数据采集来源、大数据存储、大数据分析、大数据在网络中的应用、大数据增值等细分主题 (Addo-Tenkorang和Helo, 2016)。Brinch (2018) 基于价值发现、价值创造和价值捕获三个价值创造的维度, 构建了大数据供应链管理框架。(3) 用户管理。Line等 (2020) 研究发现互惠的大数据价值创造是用户在使用数据方面获得的利益和控制水平的函数, 并提出了四种大数据互惠价值创造的类型。此外, 在线社交大数据研究中已经形成了数据产生和收集的来源、表征数据的属性、数据处理技术, 以及数据分析的价值挖掘四个细分主题 (Solazzo等, 2021)。(4) 营销管理。研究主要关注营销部门如何通过大数据分析改进营销策略, 例如Johnson等 (2019) 提出营销部门需要经历大数据分析实施的五个阶段: 萌芽、认可、承诺、文化转变和数据驱动营销; Martens等 (2016) 构建针对用户行为的大规模精细数据营销预测模型。(5) 业务流程管理。数据分析和流程改进实践能够互为补充 (Wu等, 2020), 在业务流程中使用大数据可以培养企业的灵活性、敏捷性和活力, 从而提高组织的探索和利用能力 (Dezi等, 2018)。Zhan等 (2017) 和Sultana等 (2022) 分别构建了数据驱动的创新能力和产品创新流程的大数据框架。(6) 人力资源管理。数据对人力资源实践的影响包括信息、学习、知识、效率等方面 (Garcia-Arroyo和Osca, 2021), 更重要的是, 数据能够改变企业的决策文化, 从而产生数据驱动的决策 (Frisk和Bannister, 2017; Tian和Liu, 2017), 并出现了数据分析师、数据科学家、首席数据官等新职位 (Carillo, 2017)。(7) 数字技术创新。工业4.0 (Klingenberg等, 2021)、云计算 (Shou等, 2020)、区块链 (Bhatti等, 2021)、人工智能 (Lee等, 2021) 等技术的发展能够帮助企业更好地利用数据创造价值。

## (二) 数据价值化中文文献分析

本文基于Cite Space 5.8R3软件分析了数据价值化中文文献关键词聚类图谱和关键词突现情况 (见图2)。

文献分析呈现11个聚类, 这些聚类是确定研究主题的基础, 包括大数据、数据要素、数据赋能、数据价值、反垄断、价值、价值运动、企业信息化、大数据交易平台、价值孵化和场景。聚类模块值 $Q=0.7101 > 0.3$ , 聚类结构显著, 聚类平均轮廓值 $S=0.8844 > 0.7$ , 聚类非常合理。

由关键词突现分析结果可知, 数据价值化中文研究的演变路径为: 数据价值挖掘→数据市场化流通。据此, 本文将数据价值化中文文献的具体研究内容归纳为以下两大研究主题:

### 1. 数据价值挖掘

国内对数据价值化的研究起步较晚，目前研究成果集中讨论了数据赋能、数据价值、数据

6

外国经济与管理（第45卷第12期）

---



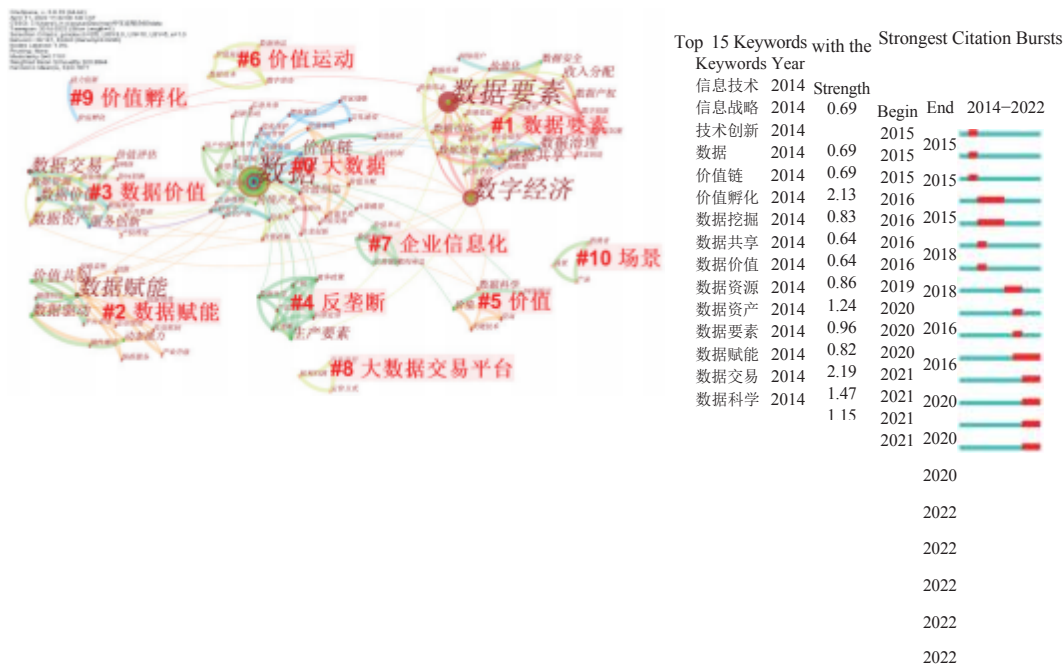


图 2 中文文献关键词聚类图谱和关键词突现分析

要素、场景、反垄断等关键数据问题，并关注了数据形态的演化问题。（1）数据赋能。主要包括数据赋能与价值共创、与制造升级、与社群问题、与弱势群体等四个方面（孙新波等，2020）。（2）数据价值。主要以数据价值链为视角展开数据价值化路径研究，例如叶春森等（2014）研究了云计算——大数据的价值链驱动产品与服务创新平台化的机制问题；王雷等（2016）基于价值链构建了大数据驱动的共享价值创造模型；李晓华和王怡帆（2020）分析了数据价值链在生产环节中的价值创造机制。值得注意的是，已有学者研究提出数据形态沿着“数据资源—数据资产—数据商品—数据资本”动态进化，各阶段价值形态分别对应潜在价值、价值创造、价值实现和价值增值（李海舰和赵丽，2021）。（3）数据要素。主要聚焦数据要素价值创造的实现方式展开研究，例如王建冬和童楠楠（2020）讨论了数据要素与其他生产要素协同联动的“五链协同”机制并构建理论模型；张昕蔚和蒋长流（2021）探究了数据要素的形成过程及其融合发展机理。此外，安小米等（2021）、梁正和吴培熠（2021）还关注释放数据要素价值过程中的数据治理问题。（4）场景。数据应用场景的研究较少，仅关注到在产品突破性创新（李树文等，2021）和营销（杨扬等，2020）等场景中的作用。（5）反垄断。数据的爆炸式增长引起学界围绕大数据具体行为的认定以及如何规制等问题展开讨论，现有关于数据垄断问题的研究主要分为理论探讨型和规制对策研究型两种类型，例如反对大数据限制竞争的学者从竞争与创新两个维度探讨大数据对市场与消费者的影响，而支持对数据垄断进行限制的学者则认为数据垄断将数据优势转化为竞争优势，并获取消费者剩余，可能造成的危害构成了一种社会成本（费方域和闫自信，2018）。

## 2. 数据市场化流通

随着数据市场化步伐加速，数据交易和流通的问题同时引发学者的关注并成为未来研究的核心问题之一，主要涉及数据市场和数据权利两个方面。（1）数据市场。数据要素市场化是市场经济条件下资源优化配置的需要，涉及市场配置、市场定价、市场交易、市场竞争、市场制度等具体问题（何玉长和王伟，2021），例如张敏（2017）基于交易安全视域讨论了大数据交易法律

监管的原则和制度构建问题；黄倩倩等（2022）提出“报价—估价—议价”相结合的数据价格生成路径和“冷启动期—成熟期”分步走的数据产品动态定价策略。也有学者提出数据市场化流通的新业态和新模式，例如尹西明等（2022）提出以数据银行为业务模式载体的“要素—机制—绩效”动态整合价值化理论模型。（2）数据权利。数据权利形成于一组多个权利集合的“权利束”，包含数据所有权、数据运营权、数据使用权的“三权分置”，涉及个人、集体、组织、国家等多元主体（赵瑞琴和孙鹏，2021）。

### (三) 数据价值化中英文文献对比分析

本文对关键词出现频次、第一次出现的年份和中心性进行计算与整理，对数据价值化中英文文献排名前10位的关键词进行共现统计，如表1所示。

表 1 中英文关键词聚类及分布对比

排名	关键词	频次	中心性	年份	关键词	频次	中心性	年份
1	big data	141	0.01	2011	数据	31	0.56	2014
2	big data analytics	137	0.23	2012	数据要素	22	0.08	2020
3	performance	96	0.02	2014	数字经济	15	0.18	2018
4	technology	59	0.01	2014	数据赋能	10	0.19	2018
5	impact	51	0.21	2014	数据交易	6	0.05	2017
6	management	50	0.06	2016	价值链	5	0.01	2014
7	supply chain management	46	0.05	2012	数据驱动	4	0.07	2018
8	information system	41	0.13	2012	价值共创	4	0.01	2018
9	Challenge	39	0.23	2015	生产要素	4	0.02	2018
10	big data analytics capability	37	0.01	2017	数据治理	4	0.09	2016

文献分析可以发现，数据价值化中英文文献的焦点并不完全一致，但却形成互补之势。数据价值化的英文文献主要侧重于大数据分析在具体场景中的应用，数据的价值主要从数据资产的角度进行解释，对数据形态及其动态演化问题关注有限。数据价值化的中文文献则围绕数据要素的形态变化和市场流通问题展开研究，但对数据在具体场景中如何发挥作用的讨论不如西方充分。值得注意的是，中西方学者都注意到数据价值链在揭示数据价值流动方面的重要作用，本文借此进一步揭示数据价值的实现过程问题。

中西方研究焦点和演变路径的差异很大程度上源于所处背景的不同，中国数字经济起步晚，但势头猛、动力足，形成“立足产业和市场优势，有效市场和有为政府相互促进”的中国模式<sup>④</sup>。中国数字经济实现了跨越式发展，数字技术和数据在生产和生活领域充分应用，政策加速推动数据要素市场化改革。西方特别是美国数字经济布局早，核心技术的领先实力和地位为其数字经济发展提供强大动力，形成“依托持续领先的技术创新，巩固数字经济全球竞争力”的美国模式，数据随着新技术、新产业、新业态、新模式的不断迭代而不断创造新价值。

### 三、数据价值化的内涵与特征

在梳理数据价值化已有研究的基础上，本文进一步引入数据价值创造的要素视角提出数据价值化的完整内涵，并对比数据要素价值化与传统要素价值化的核心特征，从而整合提升已有数据价值化内涵与特征的研究，明晰数据价值化的理论研究基础和边界。

#### (一) 数据价值化的概念内涵

数据价值化的定义存在两种视角：一是关系视角，该视角认为大数据是价值创造的重要因素，现有研究主要构建了数据相关变量与价值相关变量之间的关系模型并剖析影响路径，例如 Pinochet等（2021）的研究发现大数据分析能力积极影响基于流程导向的动态能力，进而影响产

品创新绩效。二是形态视角，该视角从数据资源、数据资产、数据商品、数据资本等数据形态的演进出发，破解数据价值识别与配置问题，例如中国信息通信研究院提出数据价值化是指以数据资源化为起点，经历数据资产化、数据资本化阶段，实现数据价值化的经济过程<sup>②</sup>。何伟（2020）和孙克（2021）认为全面激发数据要素价值需要经历三个阶段，即数据资源化阶段提升

①资料来源：中国信息通信研究院. 全球数字经济白皮书[R].北京：中国信息通信研究院，2021.

②资料来源：中国信息通信研究院政策与经济研究所. 数据价值化与数据要素市场发展报告[R].北京：中国信息通信研究院，2021.

数据质量、数据资产化阶段释放数据价值、数据资本化阶段扩展数据价值。谢康等（2020）、赵瑞琴和孙鹏（2021）、曲佳宝（2020）、杜庆昊（2020）分别讨论了各个数据形态的价值化机制。

本文基于现有研究，结合来源、参与主体、过程、目标结果等价值创造要素（孙新波等，2021），将数据价值化定义为企业在“数据资源化—数据资产化—数据商品化—数据资本化”的各个阶段中实施数据价值链行为，从而将数据要素纳入经济系统并与传统要素深入结合，进而激发数据自身价值、经济价值和社会价值。

## （二）数据要素价值化与传统要素价值化核心特征的对比

随着数字技术的不断发展，数据成为数字经济的核心关键要素和血液。满足原始性、机读性和具有一定质量的数据被称为数据要素（高富平和冉高苒，2022）。与劳动、资本、土地、知识、技术、管理等传统生产要素相同，数据要素也具有直接或间接被生产所需、明确的产权归属、各自功能互相组合、范围不断变化以及使用增殖性等特点（王柏玲等，2020）。与传统要素不同的是，数据要素在价值化对象、价值化范围、价值化效益、价值化模式、价值化创新等方面表现出新特征，也使数据价值化的底层逻辑发生了明显转变（见表2）。

表 2 传统要素价值化与数据要素价值化核心特征比较

维度	传统要素价值化	数据要素价值化
价值化对象	直接投入和参与到经济活动	与传统要素相结合
价值化范围	价值链、产业链、供应链	更广泛的生态
价值化效益	边际收益递减、边际成本先减后增	边际收益递增、边际成本递减
价值化模式	固定、单一、确定	价值倍增、投入替代、资源优化、挂牌交易
价值化创新	需要外部驱动或企业主动才能实现创新	催生新业态、新模式，实现自主创新

价值化对象方面，传统要素不依赖载体，能够直接投入和参与经济活动当中创造价值（张昕蔚和蒋长流，2021）。数据要素在企业内部无法单独创造价值（Ylijoki和Porras，2019），必须与其他传统要素构建新的要素组合和结构，体现出强渗透性（戴双兴，2020），经过数据整合、数据融通、数据洞察、数据赋能以及数据复用实现“两要素互补、多要素协同、全要素耦合”（白永秀等，2022；尹西明等，2022），全面激发对经济社会价值创造的乘数效应。

价值化范围方面，传统生产要素主要通过传统的价值链、产业链、供应链流动。数据要素的外部性和低复制成本、非竞争性等特性，促使数据在多主体之间广泛传播和共享，在互动过程中产生更多数据，并促进传统要素流动，推动大规模协作的广度、深度、频率进入一个新阶段（张昕蔚和蒋长流，2021），甚至出现了数字平台、虚拟团队、众包等新的主体集合形式，实现网络协同效应和要素跨界融合（Nambisan等，2017）。同时，数据要素的外溢又可以吸引更多的主体参与数据价值化生态，不断放大数据的价值产出。

价值化效益方面，传统要素具有稀缺性和不可再生性，在一定时间与空间范围内的要素总是有限的，一般遵循边际效益递减、边际成本先减后增的规律，制约着社会生产和增长能力的提高（王柏玲等，2020）。数据要素由于自身的可复用性、共享性、无消耗性等特性，能够源源不断地产生和循环使用（陈书晴等，2022）。随着数据要素规模的不断扩大，数据要素能够反映并

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/405133134001011141>