

基于CiteSpace分析的数据确权研究热点与趋势

黄苏宁

北京工业大学, 中国·北京 100124

【摘要】随着数字经济的不断发展,国家对数据资源的重视度不断提高,数据的产权保护问题也成为学界、实务界研究关注的热点。有关数据确权的观点或成果不断涌现,研究内容也逐步深入。为全面了解数据产权研究状况,本文基于citespace软件对我国数据确权相关研究进行了可视化梳理,重点围绕研究主题、热点等展开图谱分析,系统归纳并提炼了相关研究的演进脉络、研究侧重与发展趋势。结果显示,数据确权领域的发文量变化大概分为萌芽、缓慢发展、快速发展的阶段。从研究热点变动来看,数据确权领域研究经历了从关注数据作为新型生产要素权利的确定,到聚焦数据作为经济资产在数据市场中流动和交易的相关问题,再到现阶段促进数据的管理与治理以及数据跨境交流的发展历程。

【关键词】数据确权;数据赋权;citespace;区块链;可视化分析

【基金项目】本文系2024年北京工业大学第二十五届星火基金项目:数字经济时代数据确权的困境及法律对策研究(编号XH-2024-09-11)。

在数字经济时代,数据已成为数字化、网络化和智能化的基础性元素,逐渐演变为中国经济发展的关键生产要素。在国民经济建设和日常生活中,数据的作用愈加凸显。2020年4月9日,中共中央和国务院发布了《关于构建更加完善的要素市场化配置体制机制的意见》,首次将数据与土地、劳动、资本和技术并列,确立其作为新型生产要素的重要地位。《意见》强调需根据数据特性完善其产权属性,以支撑数据要素市场的建设。数据确权涉及法律层面上对数据持有者、加工者等主体的权利内容的界定,确定数据在不同形态下的权属。本研究使用CiteSpace软件对中国数据确权相关文献进行可视化梳理,重点分析研究主题、热点和前沿,系统总结研究演进、侧重及发展趋势,旨在为我国数据确权研究提供参考依据。

1 数据来源与研究方法

1.1 数据来源

本研究基于中国知网数据库(China National Knowledge Infrastructure, CNKI),检索字段选定关键词=“数据确权”“数据赋权”“数据权属”“个人数据权”“企业数据权”“数据权利”“数据持有权”“数据携带权”等归属于数据确权领域的词语,检索时间为2024年4月18日,时间跨度为2006—2024年得到有效文献1140篇。为保证分析结果的有效性,本研究利用CiteSpace软件对数据进行除重等操作,最终数据集由1117篇文献组成。

1.2 研究方法

本研究采用文献计量学分析的方法,运用可视化分析软件CiteSpace对搜索到的文献进行分析。首先,从宏观层面分析数据确权研究的历史进程及发展趋势;然后,通过知识图谱的关键词共现和聚类,得出数据确权研究存在的一

些主要研究面向;最后,结合知识图谱结论总结得出数据确权研究的学术热点及研究维度。

2 结果与分析

2.1 研究现状与趋势分析

在2006至2015年的萌芽阶段,国家初步关注数据确权相关问题,科技和市场需求对数据确权的重视逐渐觉醒,但整体活动较少。2016至2020年的缓慢发展阶段,随着《网络安全法》等法规的实施,公众对数据保护的需求增加,学术界的研究也开始增多。自2021年以来的快速发展阶段,国家密集出台《数据安全法》和《个人信息保护法》,“数据二十条”的发布进一步明确和完善了数据权属及数据基础制度。同时,大数据、人工智能和区块链等技术的发展为数据确权提供了更成熟的技术支持。市场需求和学术研究呈现爆发式增长,公众对数据隐私的关注也达到

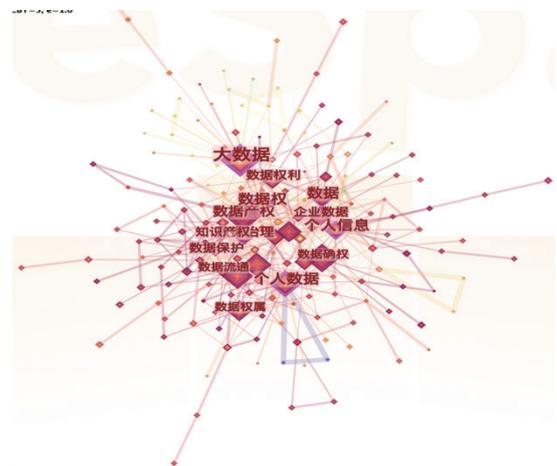


图1 数据确权关键词共现图谱

注:圆圈表示关键词,圆圈的面积越大则关键词出现的频次越高。

序号	词频	中心度	首次出现年份	关键词
1	169	0.08	2014	数据产权
2	137	0.08	2016	数据确权
3	117	0.45	2012	大数据
4	98	0.04	2019	数据合规
5	86	0.06	2017	数据权属
6	82	0.02	2020	数字经济
7	75	0.26	2006	个人数据
8	73	0.09	2015	数据权利
9	66	0.02	2020	数据要素
10	64	0.26	2012	数据权
11	61	0.09	2019	企业数据
12	53	0.02	2013	数据
13	50	0.18	2015	数据交易
14	50	0.08	2019	数据安全
15	44	0.16	2016	个人信息
16	42	0.07	2012	数据治理
17	31	0.07	2017	数据保护
18	24	0.21	2021	数据流通
19	23	0.11	2014	知识产权
20	21	0.08	2018	区块链

表1 中心度较高的关键词

新高。这些因素共同推动了数据确权领域发文数量的显著增加。

2.2 研究热点与前沿分析

2.2.1 关键词共现分析

借助CiteSpace软件进行共现分析有助于科学提炼出数据确权领域的研究热点。数据确权领域的关键词共线图谱，包括203个节点（关键词）和413条线条。其中，圆圈最大所对应的关键词（数据产权）表明出现频率最高，出现169次。此外，数据确权、大数据、数据合规、数据权属、数字经济等关键词出现的频率同样较高。图谱中各节点均有线条连接，表明数据确权领域的关键词间具有较强联系，研究热点较集中。

2.2.2 关键词突现分析

Top 19 Keywords with the Strongest Citation Bursts



表2 关键词突现图谱

注：蓝色线条表示时间间隔；红色线条表示出现年份和结束年份（突发的持续时间）Strength（突发强度）表明关键词在此研究领域的重要性。

结合表2以及对现有文献资料的梳理总结，可以将我国数据确权领域研究热点的演变分为四个阶段：

个人数据保护与大数据的早期研究（2006-2012年）：在2006至2012年间，个人数据保护问题日益成为学术界关注的焦点。欧盟在1995年通过的《数据保护指令》为个人数据保护提供了法律基础，这一框架对中国在个人数据

保护方面的关注和立法进程产生了深远影响。在这一时期，“个人数据”（2006-2020年）和“大数据”（2012-2020年）这两个关键词在学术研究中的出现频率显著增加，突现强度分别为6.89和25.59，这反映了学术界对个人数据与大数据技术日益增长的关注。随着互联网的普及和信息技术的快速发展，个人数据的广泛应用也带来了隐私泄露的风险。技术进步与个人数据安全之间的矛盾，促使中国的学术界和法律界加强了对个人数据保护和立法行动。特别是在应对新兴技术方面，中国逐步加快了立法步伐。学术研究在这一过程中发挥了重要作用，为法律制定提供了理论支持和实践建议，从而进一步加强对个人数据的保护。

数据主权与数据经济的兴起（2013-2015）：在2013-2015年期间，随着大数据技术的迅猛发展，数据逐渐成为国家和企业的重要资产，数据主权与数据经济的概念开始在中国崛起。这一阶段，关键词“数据主权”（2013-2021年）和“数据价值”（2013-2015年）的突现强度分别为7.88和1.81，反映出学术界和政策制定者对这些议题的重视。中国在这一时期开始深刻认识到数据主权的重要性，特别是在国家安全和经济利益相关的领域。2017年，中国颁布了《网络安全法》，明确了数据本地化要求和跨境数据传输的规定，显示了国家对数据主权和数据安全的高度重视。2014年，中国国家互联网信息办公室发布了《个人信息保护指南》，进一步规范了个人信息的采集和使用。此外，随着数据经济的快速发展，中国学术界和政策制定者开始系统探讨数据的商业价值、产权归属和交易机制，推动了“数据价值”的研究。有学者在数据权—数据权利—数据产权的范式嬗变理论的基础上提出“数据持有者+数据访问权”的双阶构造，认为我国数据产权配置方案应以数据持有者权利理论为基础，以数据持有者与数据访问权的法律实现为路径，一方面明晰数据持有者的权利构造、构成要件、权利边界等内容，以此构建数据流通与利用规则；另一方面在数据合同流通存在市场失灵的情形下，法律有必要介入创新数据访问权或访问规则，以确保数据资源的流动性与可获取性。学者在该阶段还意识到数据交易市场的快速发展需要建立明确的数据定价机制，并详细讨论了数据的定价问题、相关交易规则的制定，以及如何在数据交易中保证透明度和公平性来避免数据滥用和市场不正当竞争。这不仅在经济学和管理学领域产生了深远的影响，也为政策制定提供了重要的理论支撑。

隐私保护与法律规制的加强（2016—2019）：从2016年到2019年，隐私保护与法律规制得到了显著加强。随着数据应用的广泛深入，隐私保护迅速成为学术界和政策制定者关注的核心议题。2018年，欧盟实施的《通用数据保护

条例（GDPR）》为全球数据保护设定了新标准，对中国的隐私保护立法产生了重要影响。中国在此背景下，于2018年通过了《网络安全法》，2021通过《个人信息保护法》，进一步强化了对个人数据的保护，并明确了数据收集、使用和存储的规范。这一法律标志着中国数据保护法规的逐步完善，响应了公众对隐私保护的迫切需求。大数据和人工智能的发展显著提升了数据的收集和分析能力，这与公众对隐私权的保护需求形成了冲突，促使法律和政策对隐私保护的强化。另外，国际法律影响也对中国的隐私保护立法产生了深远影响。GDPR的实施为中国提供了宝贵的经验，中国在借鉴国际经验的同时，结合本土实际，逐步完善了数据保护的法律法规，从而推动了隐私保护法律的加强。除了法律方面的保护之外，区块链技术不断成熟，针对现有网络系统数据共享所面临的数据权属难确认、数据安全难保障等挑战，通过引入区块链技术，将设备状态和数据哈希值存储至区块链，保证数据的可靠性；使用智能合约构建去第三方数据交易平台，保证数据共享的安全性，便捷地完成数据变现和数据价值转移。区块链技术可以为数据提供防篡改，消除数据交易过程中的信任问题，进而有效确认数据权属，提升数据系统的安全性与可靠性，保证数据共享效率。

行为规制与利益平衡的探索（2020—2024）：2020至2024年间，随着人工智能和自动化技术的广泛应用，数据的使用方式日益复杂，尤其是在涉及个人行为数据的收集和分析时，行为规制逐渐成为学术界和政策制定者的新焦点。在这一背景下，中国积极探索通过法律和政策手段对数据使用行为进行规制，以确保技术进步与社会道德规范相一致。中国加快了对《个人信息保护法》的实施，进一步规范了数据追踪和使用行为，确保个人隐私得到更严格的保护。这一系列举措背后的推动力，部分源于技术伦理与社会规范之间的挑战。随着人工智能和大数据技术的普及，数据使用的伦理问题日益突出。中国的学术界和政策制定者正在积极探索如何在保障社会整体利益的同时，确保数据技术的合理和道德应用。此外，利益平衡的复杂性也成为研究的热点。在技术飞速发展的背景下，如何在商业利益、技术创新与个人权利保护之间找到最佳平衡点，成为一大挑战。中国通过不断调整立法和政策，寻求在技术进步和社会伦理之间达成最佳平衡，推动社会和谐发展。

2.3 研究趋势分析

关于数据确权的研究近年来始终是学术界研究的热点，通过对现有文献的梳理和总结，对该领域的研究趋势加以

分析。

其一，进一步加强区块链技术对数据确权领域的促进作用。未来，区块链、物联网（IoT）和人工智能（AI）的融合将显著提升数据确权的效率和安全性。这些技术的整合将推动智能城市、供应链管理等领域的发展，并为全球数据治理注入新的创新动力。

其二，进一步分析各国如何在数据主权与跨境数据流动之间实现平衡。探索通过多边合作机制建立全球认可的国际数据治理标准，促进数据在不同国家的流通和利用。

其三，进一步关注实现数据权利与隐私保护之间的平衡。未来的研究可以重点关注隐私保护技术在数据确权中的实际应用效果，特别是如何在不同场景下实现数据权利与隐私保护之间的平衡。

3 结论

本研究基于CiteSpace软件，通过发文情况分析、关键词共现分析、关键词突现分析、关键词聚类分析等，对我国数据确权相关研究进行可视化分析，系统归纳并提炼了相关研究的演进脉络、研究侧重与发展趋势。主要得到以下结论：（1）我国数据确权领域的发文数量整体呈现不断增长的态势，数据确权领域的发文量变化大概分为萌芽、缓慢发展、快速发展的阶段。（2）从研究热点变动来看，数据确权领域研究经历了从个人数据保护与大数据的早期研究到数据主权与数据经济的兴起，到隐私保护与法律规制的加强，再到行为规制与利益平衡的探索。（3）我国数据确权领域后续研究趋势主要包括：进一步加强区块链技术对数据确权领域的促进作用，进一步分析各国如何在数据主权与跨境数据流动之间实现平衡，进一步关注实现数据权利与隐私保护之间的平衡。

参考文献：

- [1] 文禹衡. 数据确权的范式嬗变、概念选择与归属主体[J]. 东北师大学报(哲学社会科学版), 2019(5): 69-78.
- [2] 周林彬, 马恩斯. 大数据确权的法律经济学分析[J]. 东北师大学报(哲学社会科学版). 2018(2): 30-3
- [3] 涂燕辉. 大数据的法律确权研究[J]. 佛山科学技术学院学报(社会科学版). 2016. 34(5): 83-87
- [4] 刘建刚. 数据权的证成[J]. 北京政法职业学院学报, 2016(4): 19-25.
- [5] 李爱君. 中国大数据法治发展报告(2017)[EB/OL]. [2019-12-03]. http://www.sohu.com/x/207852682_498735.