

我国健康医疗科学数据权属管理现状、问题与对策研究*

李赞梅 刘懿 蔡妙芝 赵嘉玮 冯凤翔 李晓瑛 栾冠楠 任慧玲

(中国医学科学院/北京协和医学院医学信息研究所/图书馆 北京 100005)

杨俊涛

(中国医学科学院/北京协和医学院基础医学研究所 北京 100005)

[摘要] 介绍国内外学术界和实务界对数据权属问题的研究与实践发展现状,从数据确权、隐私泄露、跨境流通等角度分析健康医疗科学数据权属问题不明带来的几大现实困境与挑战,提出相关建议,包括加快数据权立法、平衡数据主体权益、实施分级分类管理等。

[关键词] 健康医疗大数据;科学数据;数据权;数字经济;数据流通

[中图分类号] R-058 **[文献标识码]** A **[DOI]** 10.3969/j.issn.1673-6036.2022.11.001

Status, Problems and Possible Solutions for Health and Medical Science Data Ownership Management in China LI Zanmei, LIU Yi, CAI Miaozhi, ZHAO Jiawei, FENG Fengxiang, LI Xiaoying, LUAN Guannan, REN Huiling, Institute of Medical Information & Library, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100005, China; YANG Juntao, Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100005, China

[Abstract] The paper introduces the research and practice development status of data ownership in the academic and practical circles at home and abroad, analyzes several practical dilemmas and challenges caused by the unclear ownership of health and medical science data from the perspectives of data ownership confirmation, privacy disclosure, cross-border circulation, etc., and puts forward relevant suggestions, including speeding up the legislation of data rights, balancing the rights and interests of data subjects, implementing hierarchical and classified management, etc.

[Keywords] health and medical big data; scientific data; data rights; digital economy; data flow

1 引言

大数据时代,数据逐渐演变为一种新型资产和生产要素,成为社会经济高质量发展的强大推动剂^[1]。目前,全球数字经济规模占国内生产总值总比重稳步上升,成为促进经济复苏、重塑竞争优势和提升治理能力的关键力量^[2]。在数据资

[修回日期] 2022-10-25

[作者简介] 李赞梅,助理研究员,发表论文 15 篇;通信作者:任慧玲,研究馆员,发表论文 50 余篇。

[基金项目] 科技创新 2030——“新一代人工智能”重大专项课题“中文医学术语体系构建”(项目编号:2020AAA0104901)。

产化和资源化背景下,由于数据具有所涉及主体众多、利益多元及非物属性等特点,其权属一直是大数据发展应用的核心问题,直接影响数据流通利用和数据产业健康发展,在消费、医疗等领域更为严重。国务院要素市场化配置机制、“十四五”规划^[3]等明确提出“研究根据数据性质完善产权性质”“建立数据资源产权等基础制度和标准规范”,以打破数据权属界定与产业数据交易规则制定的僵局。数据权属问题引起国内外学术界和实务界广泛讨论和争论^[4],焦点涉及数据权法律概念、数据权利性质、权利内容、权利归属、确权原则等方面。

科学数据是大数据的一种特殊来源和类型,是国家科技创新和经济社会发展的重要基础性战略资源^[5-6]。健康医疗科学数据是生物健康领域研究过程中产生的数据,以及通过观测监测、考察调查、检验检测等方式取得并用于生物健康领域研究活动的原始数据及其衍生数据,包括统计数据、实验结果、测量结果等类型。通过人工智能、5G等新兴技术,健康医疗科学数据可广泛应用于临床决策支持、药物研发、精准医疗、健康管理和智慧医疗等领域。美国、欧盟、中国等相继发布和出台精准医学计划、开放科学运动、科学数据管理办法等战略计划与政策,持续推动科学数据共享利用。

与大数据发展应用相类似,健康医疗科学数据权属问题一直悬而未决^[7],成为科学数据共享利用的一大制度阻碍。个人信息保护、生物信息安全和国家数据主权浪潮的兴起,推动健康医疗科学数据管理和保护能力的转型升级,也为健康医疗科学数据权属问题带来创新发展机遇。本文基于数据权属问题国内外现状,针对我国健康医疗科学数据权属问题面临的困境与挑战,提出对策建议,以期对建立和完善健康医疗科学数据权属制度,促进健康医疗大数据发展应用目标实现提供参考。

2 数据权属问题内涵与现状

2.1 数据利用与保护失衡激发数据权属问题

数据权是基于大数据迅猛发展实践而产生的新

型法律问题。所谓数据权是指数据作为客体或载体所承载的权利,是一种新型民事权利^[8]。大数据利用及保护与传统数据利用、保护有着极大不同,其所隐含的诸多问题已突破现有法律框架。有研究者认为“大数据公然蔑视传统的数据保护方法”^[9]。数据权属问题产生的根源,主要是数据价值凸显后数据利益分配不当以及不同主体的数据保护意识增强^[10]。数据权属要解决的不是单一所有权的归属,而是确定哪些利益需要保护,构建科学的数据权利体系,促进不同利益主体之间的激励相容^[11]。

2.2 立法尚未对数据权做出明确规范

数据权属及其分配规则不清已成为数字经济发展的一大制度障碍。欧美等发达国家司法界率先针对数据权体系开展立法规范。美国于2018年3月通过《澄清境外合法使用数据法》(Clarifying Lawful Overseas Use of Data, CLOUD)^[12],试图打破既往属地管辖的数据主权判断标准框架,改为以网络运营商管辖权为标准,从而发挥美国在数字经济和信息技术领域的领先优势。欧盟以构筑内部乃至全球单一数字市场为战略目标,于2018年5月制定《通用数据保护条例》(General Data Protection Regulation, GDPR),该条例被认为是最严格的个人数据和隐私保护条例,明确了当数据私权与公权冲突时公权优先原则^[13]。我国已开始对数据权属规范立法做出尝试,其表达均谨慎而开放,《民法典》《数据安全法》等未对数据权做正面回应而采取了开放式规定。深圳市于2021年出台国内首个基础性、综合性数据法规《深圳经济特区数据条例》^[14],率先在立法实践中探索数据相关权益范围和类型,明确自然人和数据合法处理者的权益。虽然对数据监管和保护进行政策法律规制,但国内外法律界并未对数据权属问题做出明确界定和说明。

2.3 学术界对数据权属持多种不同观点

大数据时代,数据权属争议一直是学术研究的热点^[15-17,19],形成了人格权、财产权、知识产权、新型财产权、物权、债权、复合权利说等多个不同观点。如数据财产化相关理论主张通过赋

予数据一种财产化权利^[20], 突破个人信息人格权保护的阻碍, 解决个人隐私存在差异偏好的问题, 促进数据大规模流通和交易。虽然在数据权利归属问题上尚未形成统一认识^[19], 但数据权属问题方面逐步凝聚更多共识, 明晰了数据权是多元权益, 可以分为数据主权和数据权利两种。数据主权是科技进步下对国家主权内涵的延伸, 数据权利则是自然人和法人对数据享有的综合权利, 是人格权和财产权的综合体^[10]。

2.4 科学数据知识产权学说争议尚存

科学数据责任的履行主要依赖于所有权归属^[21]。国内外出现一些开放数据许可协议来帮助明确数据权属问题^[22], 如知识共享许可协议、开放数据共用许可协议等。我国曾就政府预算资金支持开展的科学数据管理出台文件^[23], 提及科学数据的发表、使用等应遵守知识产权相关规定。中国科学院也对科学数据管理与开放共享的职责分工等进行明确^[24]。但我国科学数据权属划分更多停留在理论研究层面, 在现有法律框架下探讨科学数据共享服务、数据出版、数据保护等权属问题^[25]。如王舒等^[26]梳理了科研数据的知识产权属性、归属和保护机制, 罗娇等^[27]构建了以知识产权为基础的科学数据安全策略, 司莉等^[28]从知识产权角度分析高校科学数据产权归属和利益协调机制, 盛小平等^[7]则对科学数据开放共享面临的数据确权、侵权与维权 3 方面问题进行分析并提出数据权利治理框架与对策。对于数据权利的知识产权属性, 也有研究者不认同, 如刘新宇^[29]、李爱君^[8]、韩旭至^[19]等认为, 数据不具有独创性、期限性、法定性等知识产权必要特征, 也不必然是智力劳动成果, 更不需经过知识产权取得的相关法定程序, 难以实现知识产权上的垄断和独占。以知识产权保护数据适用效果不佳, 存在巨大的局限性。

3 我国健康医疗科学数据权属困境与挑战

3.1 健康医疗科学数据保护和应用体系无法满足发展应用需求

健康医疗科学数据具有显著的领域特殊性, 包

含大量个人健康医疗数据和生物学数据, 涉及个人隐私、公众健康利益和国家生物安全。国际组织和各国都将个人健康医疗数据视为个人隐私和敏感数据加以保护, 而数据跨境流动引起的生物数据安全也考验和挑战国家数据法制韧性和生命力。美国《健康保险携带与责任法》(Health Insurance Portability And Accountability Act, HIPAA)、日本《关于为推动医疗领域的研究开发而匿名加工医疗信息的法律》(简称“医疗大数据法”)、我国《人类遗传资源管理条例》《人口健康信息管理办法(试行)》等, 均不同程度对患者个人信息相关权利、隐私信息保护与规则、数据流通规则与保护机制、数据跨境传输规制等做出规定。然而, 健康医疗科学数据保护和应用体系难以满足发展应用需求, 有关健康医疗科学数据的权属问题迄今没有明确的定论, 由此造成数据孤岛形成及数据流通受阻、个人隐私泄露与追责维权困难、数据流失削弱科技创新优势以及国家生物信息安全受潜在威胁等诸多影响和问题。

3.2 数据权利难衡量, 制约多元权益主体积极性, 阻碍数据共享开放

健康医疗科学数据发展应用受多个重要政策文件支持, 但权属不清导致数据确权授权难, 制约了数据资源建设、管理和利用方的主动性, 影响数据流通和要素化。如美国国家生物技术信息中心(National Center of Biotechnology Information, NCBI)对其所管核酸序列数据库 GenBank 的数据使用或发布无任何限制, 但由于部分提交者主张专利权、版权等权利, NCBI 只能对部分内容做出暂停发布处理。健康医疗科学数据采集生产、加工整理、开放共享和管理使用的全生命周期, 涉及众多类型的参与者和权益者^[30], 权利主体包括数据被采集对象(人或其他生命体)、医疗及科研人员、学术社群、科研机构、研究资助者、出版商、企业机构、社会公众乃至国家等众多类型。权益主体多元、权利义务关系复杂, 使得健康医疗科学数据所涉各方权利义务很难衡量。国外某些数据法虽然对个别数据做利归属做出界定, 如欧盟 GDPR 凸显个人信息权、

美国 CLOUD 法案强调运营商管辖权,但国家立法层面尚未对数据权做出明确规定,健康医疗科学数据的归属也因此难以明晰。

3.3 数据泄露频发,侵犯个人隐私和安全,追责维权难

数据泄露是大数据开放应用面临的重大风险和最大挑战之一^[31]。健康医疗数据是数据泄露的重灾区。IBM 报告显示^[32],超过 93% 的医疗保健组织在过去 3 年出现过数据泄露事件。健康医疗科学数据包含丰富的个人敏感和隐私信息,如姓名、身份证号、银行账户、通信信息、生物识别信息、病史、家族史、生活方式、用药情况等。其泄露或滥用不仅侵害用户数据权益,影响个人名誉、生活质量、物质财产乃至人身安全,也严重扰乱数字经济市场秩序。健康医疗科学数据权属不明往往还导致数据泄露追责难,出现问题时无法有效界定责任人,维权途径不明确。

3.4 跨境流通失范,削弱国家科技竞争优势,危害国家数据主权和生物信息安全

健康医疗科学数据是国家基础战略资源的重要组成部分。在日趋激烈的国际科技竞争和总体国家安全观下,科学数据未能在国内得到充分利用和安全审查就向境外流失等现象,以及由此带来的国家数据主权丧失和国际创新竞争陷入被动局面等负面效应,引起各界高度重视。科学数据外流表现形式多样,包括科学数据存储在国外平台、科技论文发表在境外期刊、国际科研合作中数据出境违法违规、云存储与云计算依托域外企业等^[27,33]。华大科技、华山医院、药明康德等机构在国际科研合作中违规向境外传输我国人类遗传资源而受科技部处罚,再次敲响健康医疗科学数据跨境流通安全的警钟^[34]。以我国公民和国内生物等为研究样本收集而来的生物材料、基因序列等敏感健康医疗科学数据,是关乎国家医药科技发展与生物信息安全稳定的重要资源,违规传递出境所造成的问题与影响的严重程度不言而喻。维护健康医疗科学数据开放的合法界限,建立对外交流的安全机制是我国科学数据管理

面临的紧迫而重要的问题^[35]。

4 我国健康医疗科学数据权属问题破局对策建议

4.1 加快数据权立法以适应大数据时代需求

大数据时代的数据权属问题需要法律来调整规制。数据权属制度建立的目的是通过明晰权利调动各类主体积极性,促进数据资源开发利用和数字经济发展。因此,健康医疗科学数据权属的分配应当有利于资源流动,同时寻求各方利益,以及数字经济效益与公民隐私保护之间的平衡。加快数据权立法可从多方面入手。首先,需厘清数据权属关系,明确数据权利主体和主体权利,做好各类数据权利的科学界定,这是解决数据流通中的权利关系、规范数据应用秩序等的先决条件。其次,加强数据多元权属规范,构建健康医疗科学数据主体和数据控制者等的利益平衡机制,这既包括强化国家对具有准公共属性数据的管理和使用权,也包括赋予机构在特定条件下的数据使用权和收益权,以及赋予用户更多数据权利以提升其在数据权属中的地位 and 话语权。立法同时也需对健康医疗科学数据侵权维权提供有效通道。

4.2 构建平衡不同主体权益的数据权属制度

数据确权的本质是要解决利益平衡问题。与传统生产要素相比,健康医疗科学数据在不同的流转和使用状态中存在多种利益维度。如数据被收集者享有隐私权、个人信息权等,数据采集者和挖掘者则可能对数据深度利用,从而将智力成果转化为具有社会经济价值的财产权益。以美国国家医学图书馆建立的医学超级叙词表(Unified Medical Language System, UMLS)为例,UMLS 是集成数十种受控词表形成的庞大语义网络,但 UMLS 声明其仅是集成语料库的产权人^[36],而非来源词表的版权持有人,用户需尊重每一个来源词表持有者的权益。健康医疗科学数据利用涉及个体数据权利与机构产权利益、行业数据应用与社会公众健康利益等诸多问题。如何基于健康医疗科技和社会经济发展现状,

平衡不同主体数据利益,也是健康医疗科学数据权属制度构建的核心、关键点和最大难点。

4.3 分级分类管理制度从源头规避数据权属问题出现

数据分级分类管理是实施全生命周期保护的起点,是平衡数据安全与发展的关键。因类施策的分级分类管理制度,有助于明确重点管制和保护的健康医疗科学数据范围以及开发利用的底线和红线,从源头上尽可能规避权属问题的出现。《科学数据管理办法》虽提出数据分级分类要求,但尚未有相关配套文件对如何分级分类做出说明。《信息安全技术 健康医疗数据安全指南》^[37]根据数据重要程度和泄露后可能造成的损害和影响级别将数据分为5级,但在实际操作过程中缺乏数据分级准确判断依据。建议可结合《数据安全法》《健康医疗数据安全指南》等各政策规范要求,以健康医疗科学数据管理和利用全生命周期为主线,按照明确责任、定义标准、评估风险、建立目录、持续监控等流程制定实施分级分类制度,并按照业务发展和法规标准要求适时调整。对于涉及国家秘密、人类遗传资源数据、出境安全管理、重大科技计划项目等的重要数据管理,则按相关部门管理规定执行。

4.4 借助先进技术手段实施数据安全和隐私保护

区块链具有去中心化、防篡改、可追溯等特点,有助于便捷地在参与健康医疗科学数据共享利用的多方之间建立互信关系,为可信的数据流通和规范利用提供技术平台。“工业数据空间”等可信安全的数据共享环境也可为推动健康医疗科学数据有效流通和数据价值挖掘提供关键基础架构。多方安全计算、联邦学习等数据安全技术则有助于避免健康医疗科学数据流通中的泄露等问题。

5 结语

健康医疗科学数据已成为健康医疗大数据发展应用和数据要素市场的重要内容。本文从健康医疗科学数据权属角度分析了数据权属不清制约健康医疗科学

数据健康及其可持续发展应用的几种表现,并提出相应对策建议。但受限于数据权属问题的复杂性以及我国健康医疗科学数据开放应用和管理尚处在初期发展阶段的现实情况,相关理论、制度及实践尚属于薄弱环节。健康医疗科学数据权属问题的解决需要政府、企业、学界等多方力量共同推进。

参考文献

- 1 中共中央,国务院. 关于构建更加完善的要素市场化配置体制机制的意见 [EB/OL]. [2022-07-01]. http://www.gov.cn/zhengce/2020-04/09/content_5500622.htm.
- 2 中国信息通信研究院. 全球数字经济新图景(2020年)——大变局下的可持续发展新动能 [EB/OL]. [2022-07-05]. <http://www.caict.ac.cn/kxyj/qwfb/bps/202010/P020201014373499777701.pdf>.
- 3 中华人民共和国国民经济和社会发展第十四个五年规划和2035年远景目标纲要 [N]. 人民日报, 2021-03-13 (001).
- 4 付伟, 于长钺. 数据权属国内外研究述评与发展动态分析 [J]. 现代情报, 2017, 37 (7): 159-165.
- 5 European Commission. Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020 [EB/OL]. [2022-07-08]. https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf#:~:text=Article%2029.2%20of%20the%20Model%20Grant%20Agreement%20sets,all%20peer-reviewed%20scientific%20publications%20relating%20to%20its%20results.
- 6 国务院办公厅. 关于印发科学数据管理办法的通知 [EB/OL]. [2022-07-23]. http://www.gov.cn/zhengce/content/2018-04/02/content_5279272.htm.
- 7 盛小平, 袁圆. 科学数据开放共享中的数据权利治理研究 [J]. 中国图书馆学报, 2021, 47 (5): 80-96.
- 8 李爱君. 数据权利属性与法律特征 [J]. 东方法学, 2018 (3): 64-74.
- 9 Tene O, Polonetsky J. Big Data for All: Privacy and User Control in the Age of Analytics [J]. Northwestern Journal of Technology and Intellectual Property, 2013 (11): 239-273.
- 10 楼何超. 数据权属争议及其分类界定 [EB/OL]. [2022-08-09]. http://www.cssn.cn/gd/gd_rwhz/gd_ktsb_

- 1666/kjgmbjxdlfzdh/202010/t20201028_5202381.sht-ml.
- 11 王融, 易泓清. 数据权属大讨论中的共识凝聚 [EB/OL]. [2022-08-05]. <https://tisi.org/18958>.
 - 12 Congress Gov. S. 512 - Law Enforcement Access to Data Stored Abroad Act [EB/OL]. [2022-07-05]. <https://www.congress.gov/bill/114th-congress/senate-bill/512/text>.
 - 13 European Commission. General Data Protection Regulation [EB/OL]. [2022-08-05]. <https://gdpr.eu/>.
 - 14 深圳市人民代表大会常务委员会. 深圳经济特区数据条例 [EB/OL]. [2022-09-07]. http://www.szrd.gov.cn/szrd_zlda/szrd_zlda_fflg/flfg_szfg/content/post_706636.html.
 - 15 姬蕾蕾. 大数据时代数据权属研究进展与评析 [J]. 图书馆, 2019 (2): 27-32.
 - 16 Carroll M W. Sharing Research Data and Intellectual Property Law: a Primer [J]. Plos Biology, 2015, 13 (8): 1-11.
 - 17 崔淑洁. 数据权属界定及“卡-梅框架”下数据保护利用规则体系构建 [J]. 广东财经大学学报, 2020, 35 (6): 78-87.
 - 18 吕廷君. 数据权体系及其法治意义 [J]. 中共中央党校学报, 2017, 21 (5): 81-88.
 - 19 韩旭至. 数据确权的困境及破解之道 [J]. 东方法学, 2020 (1): 97-107.
 - 20 劳伦斯·莱斯格. 代码 2.0: 网络空间中的法律 [M]. 北京: 清华大学出版社, 2009.
 - 21 Wallis J C, Borgman C L. Who is Responsible for Data? An Exploratory Study of Data Authorship, Ownership, and Responsibility [J]. Proceedings of the American Society for Information Science and Technology, 2012, 48 (1): 1-10.
 - 22 黄如花, 李楠. 开放数据的许可协议类型研究 [J]. 图书馆, 2016 (8): 16-21.
 - 23 国务院办公厅. 科学数据管理办法 [EB/OL]. [2022-08-18]. http://www.gov.cn/zhengce/content/2018-04/02/content_5279272.htm.
 - 24 中国科学院. 科学数据管理与开放共享办法 [EB/OL]. [2022-07-25]. http://www.cas.cn/tz/201902/t20190220_4679797.shtml.
 - 25 许燕, 麻思蓓, 郑彦宁, 等. 科学数据的法律属性与知识产权管理 [J]. 科技管理研究, 2020, 40 (22): 177-182.
 - 26 王舒, 王红, 宋晓丹. 科研数据的知识产权保护与许可机制研究 [J]. 图书馆论坛, 2016, 36 (4): 65-71.
 - 27 罗娇, 刘细文. 知识产权视角下科学数据安全管理的策略选择 [J]. 图书情报工作, 2021, 65 (12): 38-46.
 - 28 司莉, 李璐. 我国高校科研数据共享中的知识产权与利益协调机制 [J]. 图书馆, 2018 (7): 18-23.
 - 29 刘新宇. 大数据时代数据权属分析及其体系构建 [J]. 上海大学学报 (社会科学版), 2019, 36 (6): 13-25.
 - 30 顾立平. 科学数据权益分析的基本框架 [J]. 图书情报知识, 2014 (1): 34-51.
 - 31 The White House. Big Data: Seizing Opportunities, Preserving Values [EB/OL]. [2022-07-27]. https://obamawhitehouse.archives.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf
 - 32 IBM. 2020 年数据泄露成本报告 [EB/OL]. [2022-08-04]. <https://www.ibm.com/cn-zh/security/data-breach>.
 - 33 李洋, 温亮明. 我国科学数据外流: 表现、问题与对策 [J]. 图书馆杂志, 2019, 38 (12): 72-81, 115.
 - 34 科技部. 人类遗传资源管理行政处罚 [EB/OL]. [2022-07-15]. <https://fuwu.most.gov.cn/html/jxxt-ml/20181201/2837.html?tab=szcf>.
 - 35 王瑞丹, 杨静, 高孟绪, 等. 加强和规范我国科学数据管理的思考 [J]. 中国科技资源导刊, 2018, 50 (2): 1-5.
 - 36 刘静羽, 顾立平, 王昉, 等. 国际大型词表开放共享协议分析与启示——以 STKOS 词表开放共享协议设计为例 [J]. 图书馆杂志, 2020, 39 (2): 41-50.
 - 37 全国信息安全标准化技术委员会. 信息安全技术 健康医疗数据安全指南 [EB/OL]. [2022-07-15]. <http://c.gb688.cn/bzgk/gb/showGb?type=online&hcno=239351905E7B62A7DF537856738247CE>.