

科学数据跨境流动治理的现状、思考与展望

王琚, 周小林, 沈云怡, 李子愚, 任孝平, 杨云*

科技部科技评估中心, 北京 100081

* 联系人, E-mail: yangyun@ncste.org

伴随着数据密集型科学研究范式的兴起, 科学数据已成为传播速度最快、影响面最宽、开发利用潜力最大的战略性、基础性科技资源, 各国对科学数据的管理、共享和综合利用的需求与日俱增。在百年变局加速演进、大国科技博弈愈加激烈、开放科学持续推进以及我国加快构建具有全球竞争力的开放创新生态背景下, 科学数据跨境流动逐步成为影响科学研究和学术交流的重要因素之一。掌握数据跨境流动规则的主导权, 也将成为推进开放科学、获取科技战略竞争优势的关键^[1-4]。

本研究聚焦科学数据的跨境流动治理, 从科学数据开放共享所形成的“被动跨境”和科学数据跨境传输所形成的“主动跨境”两个维度, 对多边机制/国际组织、世界主要经济体和中国的科学数据跨境流动的发展历程与现状进行分析。在此基础上, 凝练当前中国科学数据跨境流动存在的问题, 并提出政策建议。以期在有效把握中国科学数据跨境治理主动性的同时, 为世界科技强国和开放创新生态建设提供有力支撑。

1 科学数据跨境流动的概念与内涵

世界各主要经济体对科学数据的定义及内涵基本相同但稍有区别。在中国, 依据2018年国务院办公厅发布的《科学数据管理办法》, 科学数据被定义为“在自然科学、工程技术科学等领域, 通过基础研究、应用研究、试验开发等产生的数据, 以及通过观测监测、考察调查、检验检测等方式取得, 并用于科学研究活动的原始数据及其衍生数据”。与之对应, 欧盟提出“研究数据(Research Data)”的概念, 并将其定义为“被收集, 且被作为推理、讨论或计算基础的所有信息, 尤其是需要审查的事实或数字。包括统计数据、实验结果、测量结果、实地观察结果、调查结果、访谈记录和图像等。”欧盟的定义相较于中国更加宽泛, 覆盖了与研究相关的信息, 而不仅限于科技领域。同时, 在部分研究中, 也将科学数据称为“科研数据”。在本研究中, 上述科学数据、科研数据和研究数据所表述的内涵均称作“科学数据”。

伴随着信息技术的迅猛发展和国际合作的日益紧密, 科学数据跨境流动也逐步影响着经济社会的多个方面。经济合作与发展组织(Organization for Economic Co-operation and



杨云 研究员, 科学技术部科技评估中心副总评估师。长期从事科技评估与科技政策研究。近年来, 重点开展了包括国际科技合作规划, 开放创新政策以及国际化人才政策研究与评估等。

Development, OECD)于1985年发布的《跨境数据流动宣言》(Declaration on Transborder Data Flow, <https://legalinstruments.oecd.org/public/doc/108/108.en.pdf>)就已关注到数据跨境流动的议题, 并将其定义为“计算机化的数据或者信息在国际层面的流动”。在中国, 国家互联网信息办公室于2022年发布的《数据出境安全评估申报指南(第一版)》中, 将数据出境行为分为三类: “(1) 数据处理器将在境内运营中收集和产生的数据传输、存储至境外; (2) 数据处理器收集和产生的数据存储在境内, 境外的机构、组织或者个人可以查询、调取、下载、导出; (3) 国家网络安全和信息化委员会办公室规定的其他数据出境行为。”对上述定义的分析显示, 数据跨境行为的内涵主要分为两类: 一是数据跨越国界的传输、转移行为; 二是数据并未跨越国界, 但能被境外主体访问的行为。因目前尚无专门针对科学数据跨境流动的法律法规, 在现行情况下, 科学数据的跨境流动遵循一般性的数据跨境流动规则。基于以上分析, 并结合《科学数据管理办法》及《数据出境安全评估申报指南(第一版)》的相关规定, 科学数据跨境流动行为具体可分为“科学数据开放共享”和“科学数据跨境传输”两类。第一类是指将科学数据汇交在国内的各类数据平台上, 被境外机构、个人访问或下载的开放共享行为, 该种行为形成科学数据的“被动跨境”。开放科学背景下开放数据的相关活动, 也属于此类科学数据的“被动跨境”。2012~2021年, 中国拥有的开放科学数据集占全球总量的2.56%, 仅低于美国的3.51%, 位列全球第二^[5]。这也意味着中国开放科学数据集所形成的被动跨境的规模已位居世界前列。第二类是指将科学数据传输至境外个人或平台的行为,

形成科学数据的“主动跨境”。典型行为包含科学家基于国际科技合作的需要将数据传输至境外,以及科学家基于发表论文或其他需求将科学数据集存储至境外的数据平台等。本研究依据以上分类,从科学数据开放共享的政策与举措,以及科学数据跨境流动所应遵循的一般性数据跨境流动的法律法规两方面,对科学数据跨境流动治理的历程与现状进行分析讨论。

2 科学数据跨境流动的国际治理现状

2.1 多边框架下科学数据跨境流动的治理现状

目前,科学数据的多边跨境治理体系主要由国际组织关于科学数据开放获取的相关倡议和规定、专门的科学数据国际组织的相关举措,以及贸易协定和国际组织中关于数据跨境流动的政策与法律法规共同构成,整个治理体系呈现立体化、多层次、碎片式的规制特征^[6]。

2.1.1 国际组织及多边机构发布的科学数据开放共享相关政策与举措

国际组织及多边机构发布的科学数据开放获取的相关倡议与指南是科学数据全球治理体系的重要组成部分。目前,包括OECD、联合国教科文组织(United Nations Educational, Scientific, and Cultural Organization, UNESCO)和七国集团(Group of Seven, G7)在内的多边机构已就科学数据开放共享提出了一系列宪章、原则和指南。

OECD于2004和2006年就公共资助的研究数据分别发布了《公共资助的研究数据开放获取宣言》(Declaration on Access to Research Data from Public Funding, <https://legalinstruments.oecd.org/public/doc/157/157.en.pdf>)和《关于公共资助的研究数据获取的原则与指南》(OECD Principles and Guidelines for Access to Research Data from Public Funding, https://read.oecd-ilibrary.org/science-and-technology/oecd-principles-and-guidelines-for-access-to-research-data-from-public-funding_9789264034020-en-fr),为研究数据的开放共享确定了包括“开放、灵活、透明和遵守法律”在内的十三条原则,开启了国际组织制定科学数据开放共享标准规范的先河。UNESCO则在2013年发布的《开发和推广开放获取政策指南》(Policy guidelines for the development and promotion of open access, <https://www.unesdoc.unesco.org/ark:/48223/pf0000215863>)中,提出了五项科学数据开放共享原则。2021年,UNESCO又发布了《开放科学建议书》(Recommendation on Open Science, <https://www.unesco.org/en/legal-affairs/recommendation-open-science>),提出科学数据是开放科学中最重要的内容,并将相关议题设为优先行动的七项领域之一。八国集团(Group of Eight, G8, 2014年后成为“七国集团”)也早在2013年就签署了《开放数据宪章》(G8 Open Data Charter, <https://assets.publishing.service.gov.uk/media/>

[5a7c1dd340f0b61a825d68e7/Open_Data_Charter.pdf](https://assets.publishing.service.gov.uk/media/5a7c1dd340f0b61a825d68e7/Open_Data_Charter.pdf)),阐明了政府数据在开放透明国家建设过程中的原则、重点开放领域和共同行动计划^[7]。2022年发布的《法兰克福声明》(G7 Science Ministers' Communiqué, <https://www.enseignements-recherche.gouv.fr/sites/default/files/2022-06/g7-science-ministers-communiqu-2022-002-pdf-18818.pdf>)和2023年发布的《仙台声明》(G7 Science and Technology Ministers' Communiqué, https://www.whitehouse.gov/wp-content/uploads/2023/05/G7-ST-Final-Communique_05.13.2023.pdf)则分别提出审查数据共享基础设施,以及基于欧盟的FAIR (Findable, Accessible, Interoperable and Reusable, 可发现、可访问、可互操作和可重用)原则、运用公共资助资金推动开放科学发展等相关内容。

科学数据相关的国际组织——国际科学理事会(International Science Council, ISC)和研究数据联盟(Research Data Alliance, RDA),则通过成立数据的共享中心、数据委员会、构建信息交换模型和推广数据开放标准等形式,多举措地推动科学数据的开放共享。

作为全球科学界最权威的国际组织,ISC从20世纪50年代起就已关注到科学数据的管理与共享问题,并专门建立了世界数据中心(World Data Center, WDC)和国际数据委员会(Committee on Data for Science and Technology, CODATA)来开展相关工作。这两个组织也是实施UNESCO《开放科学建议书》的关键伙伴和利益攸关方。成立于1957年的WDC,在2008年与天文与地球物理联合服务中心合并,重新组建了世界数据系统(World Data System, WDS),为科学数据提供高质量的管理和访问^[8]。重组之前,WDC在全球设立的51个数据中心中9个分布在中国^[9]。近年来,全球逐步以可信数据仓储认证标准和机制来取代原有的数据中心会员认证机制。其中,作为全球三大可信数据仓储认证标准之一,由ISC和WDS联合推出的CoreTrustSeal认证,从16个方面对数据中心进行评估。截至2021年7月,全球共有111个数据中心通过该项认证,其中60%以上分布在欧洲、25%左右分布在美国,仅有不到8%的数据仓分布在中国(<https://www.pumc.edu.cn/yxbd/d69d3754feed48e8951e4ddc111f74a3.htm>)。

成立于1966年的CODATA则是ISC下属的、以数据为核心的跨学科综合性科学委员会。从2000年起,CODATA先后发布了包括《网络时代的科学原则》(Principles for science in the internet era, http://www.codata.info/data_access/principles.html)《大数据时代的开放数据》(Open Data in a Big Data World, https://council.science/wp-content/uploads/2017/04/open-data-in-big-data-world_short.pdf)在内的多项具有影响力的开放数据政策。其中,《网络时代的科学原则》提出的六条数据获取原则,奠定了科学数据开放标准的理念基础。1984年6月,中国以中国科学院为代表加入CODATA^[10]。2019年,CODATA在北京会议成果的基础上发布了《研究数据北

京宣言》(The Beijing Declaration on Research Data, <https://codata.org/the-beijing-declaration-on-research-data/>), 提出包含“研究数据具有全球公共产品特征”在内的新十项原则。中国代表还在本届会议上提出共建“全球开放科学云”的倡议(Global Open Science Cloud, GOSC, <https://www.cstcloud.cn/gihz>)。目前, GOSC已与全球主要信息基础设施和国际组织达成共识, 并建立了定期对话机制, 研发出首个中欧跨洲际云联邦实验床, 为科学数据的国际治理做出积极贡献。

由欧盟、美国和澳大利亚于2013年共同创立的研究数据联盟, 也是重要的科学数据国际组织, 在降低全球科学数据开放共享的壁垒、建立沟通的平台与桥梁方面发挥着重要作用。RDA主要面向农业、化学、数字人文等8个领域, 目前已吸引来自148个国家的超过13000名机构及个人成员。

上述组织除了独立开展开放数据相关的工作之外, 其共同行动也引领着全球科学数据开放共享的发展进程。2016年, RDA和WDS共同推出了致力于文献与数据关联的信息交换模型(Scholarly Link exchange, SCHOLIX), 联合推动全球科学数据的交换共享。2022年6月, 在欧盟地平线计划的支持下, CODATA和RDA联合推出了WorldFAIR项目。该项目通过在全球13个国家的11个学科或跨学科领域提出建议和制定框架, 来推进FAIR原则的实施(<https://codata.org/initiatives/decadal-programme2/worldfair/>)。2023年9月, UNESCO、ISC、CODATA和WDS四大组织又联合主办了以“迈向数据更加FAIR的世界——实施UNESCO开放科学建议以应对全球挑战”为主题的研讨会, 来促进以全球数据互操作为重点的FAIR原则的执行(<https://council.science/events/towards-fairer-world/>)。

近年来, 中国主导设立的多边合作平台也逐步开始关注科学数据跨境流动议题。中国面向全球开展科技创新交流合作的国家级平台——中关村论坛, 于2023年专门设置了“科学数据与开放科学”的平行论坛, 推动各国就促进科学数据

存储、开放和共享等议题展开讨论。国家自然科学基金委员会在与科学欧洲(Science Europe, SE)2023年的对话中, 也将科学数据跨境流动列为中欧科研合作与交流的重要议题之一。依托这些平台, 中国在全球科学数据跨境流动治理领域积极发声并贡献着力量。

2.1.2 贸易协定与国际组织关于数据跨境流动的法律法规

科学数据的跨境流动格局除了受开放共享倡议和声明等软法的影响以外, 贸易协定和国际组织关于数据跨境流动的一般性法律法规也对其产生约束。作为全球最重要的多边机制治理体系之一, 世界贸易组织(World Trade Organization, WTO)依托电子商务谈判开展关于数据跨境流动规则的制定。截至2023年12月, WTO电子商务联合声明倡议(Joint Statement Initiative, JSI)已在包括开放政府数据和保护个人数据等13项议题中取得实质性进展, 而数据流动、本地化以及源代码等议题的谈判预计将于2024年完成(Co-Convenors of the WTO Joint Statement Initiative on E-Commerce announce substantial conclusion of negotiations on a number of global digital trade rules, <https://www.meti.go.jp/press/2023/12/20231220004/202312004-1.pdf>)。另外, 日趋活跃的双多边经贸协定也成为许多国家寻求达成数据跨境流动治理共识的平台(部分协议、倡议等见表1)。统计显示, 目前全球已有超过180个区域贸易协定中, 增设了包含数据跨境流动规则的专门章节或条款^[11]。其中, 《跨太平洋伙伴关系协定》(Trans-Pacific Partnership Agreement, TPP, 美国退出后更名为Comprehensive and Progressive Agreement for Trans-Pacific Partnership, CPTPP, <http://www.mofcom.gov.cn/article/zwgk/bnjg/202101/20210103030014.shtml>) 是首个包含数据跨境流动约束性规则的贸易协定, 其从个人信息保护、通过电子方式跨境传输信息和计算设施的位置三部分规制了数据问题。数字经济伙伴关系协定(Digital Economy Partnership Agree-

表1 国际组织与多边协定中涉及数据跨境流动的相关政策与法规

Table 1 Policies and regulations related to cross-border data flow in international organizations and multilateral agreements

序号	发布单位	名称	生效时间
1	亚太经合组织(APEC)	《APEC隐私框架》(https://www.apec.org/Publications/2017/08/APEC-Privacy-Framework-(2015))	2005年
2	日本、澳大利亚、加拿大等11国	《全面与进步跨太平洋伙伴关系协定》(CPTPP)	2017年
3	世界贸易组织	《关于电子商务的联合声明》	2019年
4	二十国集团(Group 20)	启动“大阪轨道”(Osaka Track, https://www.mofa.go.jp/policy/economy/g20_summit/osaka19/en/documents/final_g20_osaka_leaders_declaration.html)	2019年
5	新加坡、智利、新西兰	《数字经济伙伴关系协定》(DEPA)	2020年
6	中国、日本、韩国、澳大利亚、新西兰和东盟十国	《区域全面经济伙伴关系协定》(RCEP)	2022年
7	七国集团	《促进可信数据自由流动计划》(G7 Action Plan for Promoting Data Free Flow with Trust, http://www.g7.utoronto.ca/ict/2022-05-11-g7-ministerial-declaration-digital-ministers-meeting-en-data.pdf)	2022年

ment, DEPA, <http://gjs.mofcom.gov.cn/article/wj/ftar/202111/20211103216433.shtml>)作为全球第一个数字经济的重要规则安排,其数据跨境传输条款直接援引了CPTPP第14.11条的规则内容。2022年1月正式生效的《区域全面经济伙伴关系协定》(Regional Comprehensive Economic Partnership, RCEP, http://fta.mofcom.gov.cn/rcep/rcep_new.shtml)则将“数据自由流动”作为基础性原则,将“数据安全流动”作为例外性原则,兼顾数据的自由流动和充分保护^[12]。

2.2 欧盟科学数据跨境流动治理现状

2.2.1 欧盟科学数据开放共享的发展历程与现状

在开放科学环境下,过去10年,欧盟的科学数据开放获取经历了从倡导到试点开放,再到明确具体操作规范的演进历程。欧盟及其成员国主要通过协同制定开放共享的政策与规范,以及开展基础设施建设等举措来推动各成员国科学数据的开放。

2013年,G8国家在北爱尔兰峰会上签署《开放数据宪章》(以下简称《宪章》),拉开了欧盟国家开放政府数据的序幕。《宪章》不仅涉及政务数据的开放,对于政府所掌握的研究数据的开放也产生重要影响。基于《宪章》,英国(Open Data Strategy, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/330382/bis-14-946-open-data-strategy-2014-2016.pdf)、法国(French Prime Minister releases the G8 Open Data Charter Action Plan for France, <https://www.etalab.gouv.fr/frenchprime-ministerreleasestheg8opendatacharteractionplanforfrance/>)、德国(The Federal Government's National Action Plan to implement the G8 Open Data Charter, <https://www.bundesregierung.de/breg-en/service/information-material-issued-by-the-federal-government/the-federal-government-s-national-action-plan-to-implement-the-g8-open-data-charter-736006>)和意大利(Italy, Second Action Plan, 2014~2016, <https://www.opengovpartnership.org/documents/italy-second-action-plan-2014-16/>)分别制定了各自的数据开放方案。随后,欧盟依托科技计划逐步推出了一系列开放科学数据相关的政策与举措。如在2015年开启的第七框架计划(Framework 7, FP7)中就启动了对20%的研究项目进行开放获取资助的试点。到地平线2020计划(Horizon 2020, H2020),开放获取已成为所有资助项目的必选项(Results of the FP7 post-grant Open Access pilot, <https://digital-strategy.ec.europa.eu/en/news/results-fp7-post-grant-open-access-pilot>)。欧盟还依托H2020出台了基于FAIR原则的数据管理指南(Guidelines on FAIR Data Management in Horizon 2020, https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf),并明确了H2020计划中研究数据开放获取的四个步骤:一是建立项目的数据管理计划(Data Management Plan, DMP);二是把

研究数据集和其元数据放置在研究数据仓储中;三是让第三方用户能够免费发现、访问、挖掘、分析、重用及分发的原则;四是提供数据使用和验证的必要工具(信息)^[7]。另外,欧盟还通过实施欧洲开放科学云计划(European Open Science Cloud, EOSC)和构建FOSTER电子学习平台,来提升开放共享的基础设施和科研人员的数据管理共享技能水平。

2.2.2 欧盟数据跨境流动的制度体系

欧盟没有专门针对科学数据跨境的法律法规,科学数据跨境流动所遵从的法律法规,主要依据科学数据中涉及的是个人数据还是非个人数据,分别遵从相应的一般性数据跨境流动的规则体系。个人数据方面,欧盟委员会2016年发布的《通用数据保护条例》(General Data Protection Regulation, GDPR, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02016R0679-20160504&qid=1532348683434>)第五章规定,个人数据跨境流动主要包括3种途径:一是基于充分性认定机制(即“白名单”机制)的跨境流动,即将通过欧盟充分性认定的国家被视为具备充足数据保护能力的“白名单”国家,欧盟可直接向其传输数据;二是基于保障措施机制的跨境流动,如通过签订标准合同、约束性企业规则、认证机制、行为准则等协议,实现数据跨境流动;三是基于获得数据主体同意和签订同意书的跨境流动。非个人数据的跨境流动治理体系主要由《非个人数据自由流动条例》(Framework for the free flow of non-personal data in the European Union, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32018R1807>)《欧洲数据战略》(A European strategy for data, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0066>)《数据治理法案》(Data Governance Act, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32022R0868>)《数据法案》(Data Act, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2022%3A68%3AFIN>)等政策与法规构成。2020年发布的《欧洲数据战略》中,欧盟强调了对外加强欧盟全球竞争力和维护数据主权的重要性,提出在域外对欧盟数据进行访问时,必须保证其行为符合欧盟的基本价值观与立法框架。具体地,《数据治理法案》通过第七章“国际转入与转让”引入类似GDPR的“充分性保障”监管规则,《数据法案》的第七章则强调了向非欧盟国家传输数据时的安全评估议题。总的来说,无论是个人数据还是非个人数据,欧盟均倡导数据在欧盟内部的自由流动,以形成单一的“数字市场”。对于数据向非欧盟国家跨境的传输,欧盟则坚持着高标准与严要求。

2.3 美国科学数据跨境流动治理现状

2.3.1 美国科学数据开放共享的发展历程与现状

美国的科学数据开放共享起步较早^[13],经历了从联邦政府部门政务信息数据开放逐渐到科学数据开放的历程。早在1966年,美国政府颁布的《信息自由法》(Freedom of Infor-

mation Act, FOIA, <https://www.govinfo.gov/content/pkg/PLAW-107publ347/pdf/PLAW-107publ347.pdf>)开创了联邦政府信息公开化的先河. 2002年出台的《电子政府法》(E-Government Act of 2002, <https://www.govinfo.gov/content/pkg/PLAW-107publ347/pdf/PLAW-107publ347.pdf>)则对包括科学数据在内的各类政府数据的开放提供了法律保障. 在此基础上, 美国联邦各部门积极制定扩大公共资助研究成果开放获取的方案. 从2002年开始, 包括美国地质调查局(U. S. Geological Survey, USGS, <https://www.usgs.gov/office-of-science-quality-and-integrity/public-access-results-federally-funded-research-us>)、美国国立卫生研究院(National Institutes of Health, NIH, <https://sharing.nih.gov/data-management-and-sharing-policy>)、美国国家海洋与大气管理局(National Oceanic and Atmospheric Administration, NOAA, <https://www.noaa.gov/information-technology/open-data-dissemination>)、美国国家科学基金会(National Science Foundation, NSF, <https://www.nsf.gov/pubs/2023/nsf23104/nsf23104.pdf>)和美国国家航天航空局(National Aeronautics and Space Administration, NASA, <https://ntrs.nasa.gov/api/citations/20150020926/downloads/20150020926.pdf>)等在内的管理部门和科研机构逐步出台本机构包含开放共享内容的科学数据管理政策, 并由首席自由法官员负责监督执行. 白宫科技政策办公室也针对受联邦资金资助的科研成果的开放共享, 分别于2013和2022年发布了《提高公众获取联邦资助科研成果的备忘录》(Increasing Access to the Results of Federally Funded Scientific Research, https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf)和《确保免费、立即和公平地获取联邦政府机构资助研究的备忘录》(Ensuring Free, Immediate, and Equitable Access to Federally Funded Research, <https://www.whitehouse.gov/wp-content/uploads/2022/08/08-2022-OSTP-Public-Access-Memo.pdf>), 要求研发支出在1亿美元以上的联邦机构制定开放共享计划. 基于此, 美国联邦政府各机构通过统一的data.gov平台向社会发布开放数据库. 目前, 该平台上已有超过28万个开放数据库. 此外, 在国会授权下, NSF、NASA、NOAA、USGS、NIH等机构均已建成各自领域的数据库基础设施和国家级科学数据中心. 2016年12月, 美国政府开放数据部际协调小组与国家科学技术委员会共同制定发布《通过国际科技合作促进联邦资助科学数据开放共享的原则》(Principles for Promoting Access to Federal Government-Supported Scientific Data and Research Findings Through International Scientific Cooperation, https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/NSTC/iwgodsp_principles_0.pdf), 明确了美国联邦科学技术部门在跨国科学数据交流合作中应遵循的原则. 提出国际科技合作活动的合作伙伴, 应在项目启动时建立DMP, 各伙伴方应确保科学数据在技术和法律上

的互操作性, 联邦政府和国际科技合作伙伴, 应实施共同的数据开放共享政策和数据标准, 以促进开放数据政策在国际科研合作中的应用.

2.3.2 美国数据跨境流动的制度体系

美国没有专门针对科学数据跨境流动的法律法规, 科学数据的跨境流动原则上遵从一般性数据跨境流动的法律法规. 美国数据跨境流动的治理经历了以《信息自由法》的颁布为标志的起步探索阶段, 以《关键基础设施法》(Critical Infrastructure Protection, <https://www.presidency.ucsb.edu/documents/executive-order-13010-critical-infrastructure-protection>)、《电子政府法》(E-Government Act of 2002, <https://www.justice.gov/opcl/e-government-act-2002>)、《网络安全研究与发展法》(Cyber Security Research and Development Act, <https://www.govinfo.gov/content/pkg/COMPS-1842/pdf/COMPS-1842.pdf>)和《联邦信息安全管理法》(Federal Information Security Modernization Act, <https://security.cms.gov/learn/federal-information-security-modernization-act-fisma>)为代表的积累完善阶段, 以及以《网络安全法》(Strengthening American Cybersecurity Act of 2022, <https://www.congress.gov/bill/117th-congress/senate-bill/3600#:~:text=Strengthening%20American%20Cybersecurity%20Act%20of%202022%20This%20bill,ongoing%20and%20continuous%20assessments%20of%20federal%20risk%20posture>)、《境外澄清数据保护法》(Clarifying Lawful Overseas Use of Data Act, CLOUD法案, https://www.justice.gov/d9/pages/attachments/2019/04/09/cloud_act.pdf)的颁布为代表的发展扩张阶段. 目前美国已形成了属人原则为主、属地原则为辅的数据跨境管辖模式, 以及准许数据自由流入和限制数据自由流出的政策体系. 前者主要体现在《韩美自由贸易协定》(U. S.-Korea Free Trade Agreement, <https://ustr.gov/trade-agreements/free-trade-agreements/korus-fta>)《APEC跨境隐私规则体系》(<https://www.apec.org/docs/default-source/Groups/ECSCG/CBPR/CBPR-PoliciesRulesGuidelines.pdf>)和《美国-墨西哥-加拿大协定》(United States-Mexico-Canada Agreement, USMCA, <https://ustr.gov/trade-agreements/free-trade-agreements/united-states-mexico-canada-agreement>)等双多边协定和政策中. 后者则通过CLOUD法案来实施长臂管辖, 确保政府可跨境调取境外数据; 通过实施《出口管理条例》(Export Administration Regulations, EAR, <https://www.bis.doc.gov/index.php/regulations/export-administration-regulations-ear>)以及其中的《商业管制清单》(Commercial Control List, CCL, <https://www.bis.doc.gov/index.php/regulations/commerce-control-list-ccl>), 和《外国投资风险审查现代化法案》(Foreign Investment Risk Review Modernization Act of 2018, FIRRMA, <https://www.congress.gov/bill/115th-congress/house-bill/5841/text>), 对部分关键技术与特定领域的数据出口进行限制. 另外, 2024年2

月,拜登政府颁发的《阻止有关国家获取美国公民大规模敏感个人数据和美国政府相关数据的行政命令》(Executive Order on Preventing Access to Americans' Bulk Sensitive Personal Data and United States Government-Related Data by Countries of Concern, <https://www.whitehouse.gov/briefing-room/presidential-actions/2024/02/28/executive-order-on-preventing-access-to-americans-bulk-sensitive-personal-data-and-united-states-government-related-data-by-countries-of-concern/>),首次明确提出限制向包括中国在内的6个“受关注国家”出售或转移包括基因组数据、生物识别数据、个人健康数据和地理位置数据在内的“敏感个人数据”。这也体现出,当前中美之间“科技竞争”与“数字冷战”的叠加,使得科学数据跨境流动议题也逐步成为大国竞争中新的博弈点。

3 中国科学数据跨境流动治理现状

3.1 中国科学数据开放共享的发展历程与现状

在全球开放科学的背景下,中国科学数据开放共享的政策体系建设可以分为3个时期——基础设施建设时期(“十五”规划到“十一五”规划)、开放共享机制建设时期(“十二五”规划和精细化管理时期(“十三五”规划之后)(如图1所示)。

2001年,“十五”规划伊始,科学技术部就启动了“科学数据开放共享工程”。2004年,在《2004~2010年国家科技基础条件平台建设纲要》的支持下,包括6个科学数据开放共享平台在内的23个科技平台开展建设,为国家层面的科学数据中心体系建设奠定基础^[14]。

“十二五”时期发布的《国家重大科技基础设施建设中长期规划(2012~2030年)》,提出将开放共享程度作为设施运行

考核的重要指标。在这一时期,国家自然科学基金委员会和中国科学院等机构,逐步出台受财政资金资助论文开放获取的相关政策,促进了开放共享制度的落地。

进入“十三五”时期,中国的科学数据开放共享工作进入了精细化管理阶段。2018年国务院发布的《科学数据管理办法》(以下简称《管理办法》),是首部国家层面对于多领域科学数据开放共享工作的制度安排。《管理办法》要求“政府预算资金资助形成的科学数据应当面向社会和相关部门开放共享”。同年,科学技术部、财政部发布的《国家科技资源共享服务平台管理办法》也做出类似规定。2019年,科学技术部在重点研发计划项目指南中进一步要求“在科研活动过程中的模型、算法、数据、数据产品、评估方法、研究结果、参数必须公开发表或在线免费共享”(https://service. most.gov.cn/2015Stztg_all/20190125/2844.html)。这一时期中国无论是科学数据中心的建设还是与论文关联的通用型数据存储库的建设,都取得了一定的进展,并在全球相关领域研究中发挥积极作用。2019年,科学技术部、财政部对于国家科技资源共享服务平台进行优化,形成20个国家科学数据中心^[15]。包括国家高能物理科学数据中心和国家人口健康科学数据中心仓储等在内的多个数据中心已通过了CoreTrustSeal认证,并逐步成为各领域全球主要的科学数据中心。中国科学院从2015年开始自主研发科学数据银行,目前已成为全球万余种学术期刊推荐的论文关联数据存储与发布平台,在支撑广大学者工作的同时,有效汇聚了全球优质科学数据资源^[16]。另外,相关部委、地方和机构也积极开展本行业领域科学数据管理办法实施细则的制定。据不完全统计,截至2020年6月,包括农业农村部、交通运输部、陕西省和中国科学院在内的20余个部委、地方和机构已制定了科学数据

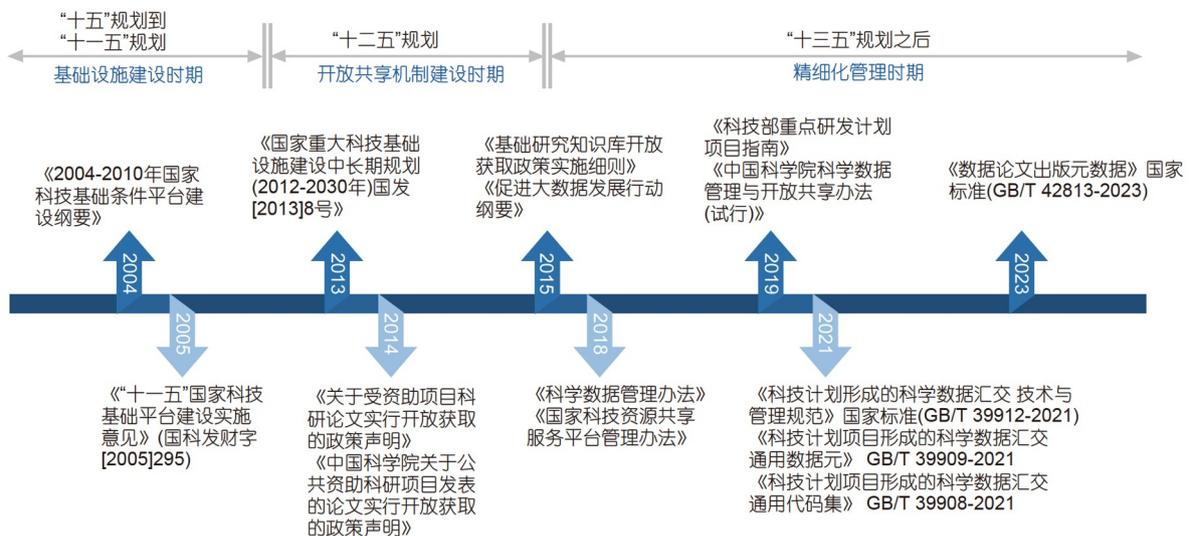


图1 中国科学数据开放共享的相关制度建设

Figure 1 Institutional construction of scientific data's opening and sharing in China

相关的管理办法或实施细则,科学数据开放共享机制逐级落实。总体来讲,基于“开放为常态,不开放为例外”原则,目前,中国已形成政府、行业机构和领域数据中心为主体的科学数据管理和建设体系。

3.2 中国科学数据跨境流动现状

3.2.1 中国数据跨境流动的制度体系

与国际惯例一致,中国科学数据的跨境流动也需遵从一般性数据跨境的法律法规。2017年以来,中国逐步形成了以《网络安全法》《数据安全法》《个人信息保护法》(以下简称“数据三法”)为核心,《数据安全评估指南》《个人信息出境标准合同办法》和《个人信息保护认证实施规则》等标准、指南为支撑的、数据跨境流动的法律法规体系(如图2所示)。

中国最早对个人信息和重要数据的跨境流动做出统一规定是在2017年6月1日正式施行的《网络安全法》。《网络安全法》规定了关键信息基础设施运营者对个人数据和重要数据的本地化存储要求,及数据的出境评估义务。2021年9月出台的《数据安全法》则从数据的分类分级、重要数据目录制定,以及数据跨境流动的具体方式等方面做出规定。个人信息跨境流动的管理则主要依据2021年11月1日实施的《个人信息保护法》。《个人信息保护法》主要从数据的本地化存储要求、个人信息处理者向境外提供个人信息所具备的条件,以及个人信息跨境流动的路径等方面来规范个人信息跨境有序流动。

为保障“数据三法”的实施,国家互联网信息办公室等部门又陆续发布了《数据出境安全评估办法》(以下简称《评估办法》)、《个人信息出境标准合同办法》《信息安全技术—个人信息安全规范》《个人信息跨境处理活动安全认证规范》等管理办法及标准,全面支撑数据跨境流动体系的

构建。

3.2.2 中国科学数据跨境流动体系的现状与实践

根据一般性原则,当前中国科学数据跨境流动遵循“数据三法”及相关的制度与标准所构建的一般数据跨境流动的体系。具体而言,除根据《数据安全法》和《保守国家秘密法》规定涉及国家秘密的数据禁止出境外,当前中国科学数据跨境流动的方式主要分为3种(如图3所示):一是自由跨境流动;二是有条件地跨境流动,包括进行出境安全评估、订立标准合同和开展个人信息认证保护的方式跨境流动;三是限制性跨境流动。

第一,科学数据的自由跨境流动。依据国家网络安全和信息化委员会办公室2024年3月发布的《促进和规范数据跨境流动规定》,学术合作等活动中产生的数据出境,不包含个人信息或者重要数据的,不需要申报数据出境安全评估、订立个人信息出境标准合同、通过个人信息保护认证。

第二,科学数据有条件地跨境流动。该种类型主要包含以下三类跨境机制:(1)科学数据采取数据出境安全评估的方式跨境流动。结合《网络安全法》《个人信息保护法》《数据出境安全评估办法》《数据出境安全评估申报指南》等法律法规,采取此种方式出境的科学数据应符合下列情形之一:①数据处理者向境外提供重要数据;②科技领域的关键信息基础设施运营者和处理包含100万人以上个人信息的处理者向境外提供个人信息;③自上年1月1日起累计向境外提供10万人个人信息或者1万人敏感个人信息的处理者向境外提供个人信息。(2)科学数据采取订立个人信息标准合同的方式跨境流动。依据《个人信息保护法》《个人信息出境标准合同办法》等法律法规,采用该种方式开展数据跨境流动的科学数据处理者,应当同时符合下列情形:①非关键信息基础设施运营者;②处理的个人信息不满100万人的;③自上年1月1日起累计向境外提供个人信息不满10万人的;④自

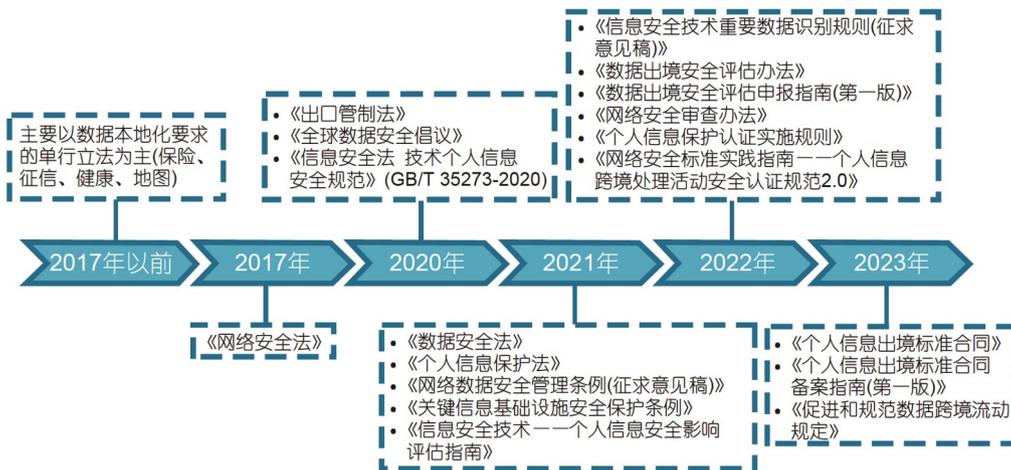


图2 中国数据跨境流动制度建设历程
Figure 2 Construction process of data's cross-border flow system in China

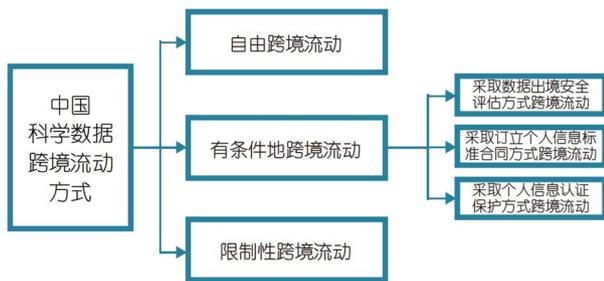


图3 中国科学数据跨境流动方式

Figure 3 China's pathway for scientific data's cross-border flow

上年1月1日起累计向境外提供敏感个人信息不满1万人的。

(3)科学数据采取个人信息认证保护的方式跨境流动。依据《个人信息保护法》《个人信息保护认证实施规则》等法律法规，采用该种方式开展包含个人信息的科学数据跨境流动的数据处理者包含以下主体：①申请认证的个人信息处理者应取得合法的法人资格，正常经营且具有良好的信誉、商誉；②跨国公司或者同一经济、事业实体下子公司或关联公司；③在中国境外处理中国境内自然人个人信息以分析、评估境内自然人的行为的活动。

第三，科学数据的限制性跨境流动。《出口管制法》第四条规定，对于管制物项相关的技术资料等数据，国家实行统一的出口管制制度，通过制定管制清单、名录或者目录、实施出口许可等方式进行管理。因此，《中国禁止出口限制出口技术目录》中所涉及的技术和军民两用技术的资料及相关数据，应当遵循《出口管制法》的相关规定，限制或禁止其跨境流动。

在中国科学数据跨境流动的制度体系不断完善的过程中，一些地区也积极开展科学数据跨境流动的实践探索。目前，包括医疗健康、航空、汽车等领域的相关企业已依据数据出境安全评估办法成功实现数据有序跨境流动。北京还在积极探索跨国药企数据出境的“绿色通道”，并上线了全市首个数据安全与治理公共服务平台。另外，依据个人信息出境标准合同备案和个人信息保护认证两种方式开展的数据跨境流动在北京和澳门等地也已有成功案例。这标志着3种有条件的数据跨境流动均已成功实践。

4 中国科学数据跨境流动面临的问题与挑战

4.1 科学数据跨境流动治理的顶层设计不足

《数据安全法》第六条规定，工业、电信、交通、金融、自然资源、卫生健康、教育、科技等主管部门承担本行业、本领域数据安全监管职责。目前，工业和信息化部、自然资源部、财政部和人民银行等部门均已发布了各自领域的科学数据管理办法，而科技领域数据治理的思路、原则与指

导性文件均未出台。同时，一般性的数据跨境治理体系与科学数据开放共享的“被动跨境”规制目前仍未有效衔接；对于国际科技合作中的科学数据跨境流动机制，目前也尚无具体的政策安排，一定程度上挤压了当前国际科技的合作空间。另外，中央网络安全和信息化委员会办公室在2024年3月公布的《促进和规范数据跨境流动规定》中所提出的“因学术合作等活动产生的数据出境，不包含个人信息或者重要数据的，不需要通过数据出境安全评估等方式出境”，虽然一定程度上减轻了国内科研人员和国际合作伙伴的担忧，但是顶层设计缺位，以及“重要数据”制定的主办单位尚未明晰等问题，依然使科学数据跨境流动可能产生的风险无法得到有效防范，治理体系整体呈现碎片化的特征。

4.2 科学数据跨境流动治理的具体规制尚未完全实施

一是用于指导科学数据出境的具体规范性文件尚未出台，尤其是处于核心地位的科学数据分类分级管理办法、重要数据目录和关键信息基础设施认定规则等文件均处于“缺位”状态。二是“数据管理计划”制度有待实施。欧盟和美国均将DMP作为数据管理的关键手段，并将其贯穿于科研项目及合作的整个数据管理周期。在中国，数据管理计划机制尚未大范围实施，使得大量科研人员在合作中面临必要的数据交换时，“无规则可依”，为开放合作带来风险，也不利于中国在国际科技合作谈判时占据主动。三是新旧规则在原则与具体管理细则等方面有较大差异，已开放共享的科学数据面临违规风险。如《科学数据管理办法》中对于政府预算资金资助形成的科学数据提出“开放为常态、不开放为例外”的管理原则，而《数据安全法》中则强调“国家保障数据依法有序自由流动”，新旧治理规则的指导原则已有较大差异。又如《数据安全法》规定“国家建立数据分类分级保护制度”，而《科学数据管理办法》则规定“法人单位要对科学数据进行分级分类”，两种数据分类分级制度建立的主体也有所不同。因此，有必要依据上位法，修订现有科学数据管理政策，同时对已开放共享的科学数据，在新的规制全面实施后重新合规。

4.3 科学数据中心的国际化、标准化和人才队伍建设有待提升

一是科学数据中心和仓储库的国际化水平不足。中国通过CoreTrustSeal认证的数据库数量仅占全球的不到8%，与欧洲(60%)和美国(25%)有较大差距。Springer Nature等认可的6所科学数据存储库中，中国只占一所，另外三所分布在美国、两所分布在欧洲。二是科学数据国际化建设严重不足。相较于欧洲积极制定并在全球推广FAIR原则，中国在科学数据开放共享原则和标准体系建设方面还需大幅提升。另外，目前全球主要的可信数据仓储认证标准和机制基本以美

国与欧盟国家制定的为主,中国在该领域的话语权严重缺失。三是科学数据跨境流动治理的人才队伍建设有待加强。中国相关规则的建设处于起步阶段,专业的管理人才数量远远不足^[17]。同时,科学数据跨境流动治理覆盖管理学、法学和国际合作等多个领域,复合型人才有待培养。

4.4 中国参与科学数据跨境流动的治理仍显不足

一是在科学数据跨境流动的国际治理中发声不足。目前,以美欧为代表的两大数字治理规则体系相对成型,RDA和WDS等专门性的科学数据国际组织仍然由欧美国家主导。这一现状使得中国在现有的科学数据国际组织中发声不足,国际组织相关规则易受欧美规则影响。二是中国主导建设的科学数据国际治理平台的数量和影响力有限,使中国科学数据跨境流动在国际治理中处于被动地位。中国虽逐步依托中关村论坛、香山科学会议等平台开展相关治理对话,但平台数量及国际影响力非常有限,制约了中国参与这一领域国际治理的深度与广度。三是中国对于科学数据跨境流动全球治理的相关研究还不够全面深入,尤其是从具有中国视角、全球视野开展的科学数据跨境流动研究非常有限,对中国参与全球创新治理的支撑不足。

5 思考与展望

5.1 加强科学数据跨境流动的顶层设计,创新治理方式

一是平衡发展与安全,充分认识在国际化科研环境和开放创新生态建设背景下,科学数据开放共享的必要性。从开放科学和风险管理^[18,19]两方面对科学数据跨境流动治理进行整体布局。二是依据科学数据的特点,创新治理方式。如探索在国家数据跨境治理体系中对科学数据跨境实施单列管理,积极推动以科学数据开放共享为依托的科学数据自由跨境流动;设置科学数据跨境流动负面清单;对通过研发产生的和用于科学研究的非敏感领域数据建立跨境流动“绿色通道”试点,允许未进入负面清单的科学数据自由跨境流动等创新性治理举措^[20]。三是将科学数据跨境流动治理与人工智能、数据密集型科学的发展相结合,对科学数据跨境治理进行系统性、前瞻性布局。

5.2 自上而下与自下而上相结合,迭代完善科学数据的跨境治理规制

一是依据《数据安全法》《关键信息基础设施安全保护条例》《网络安全标准实践指南——网络数据分类分级指引》等要求,参考《信息安全技术重要数据识别指南》(征求意见稿)、《工业和信息化领域数据安全管理办法(试行)》等文件,推动和加快科学数据分类分级规则、重要数据

目录和科技领域关键信息基础设施认定规则等标准和政策的制定。二是引导包括国家自然科学基金委、各国家科技计划管理部门等科研资助机构,出台科学数据管理和流动的制度和政策操作指南,尽快在中央、地方财政科技计划中完善以DMP为基础的数据管理体系,并在全国范围内进行推广。三是引导科研机构、高校等科学数据生产单位制定和实施本单位科学数据管理和跨境流动的方案。四是在科学数据治理规制全面实施之后,对已开放共享的科学数据开展全面的合规性审查。同时,管理部门应在指导保证制度政策落实的同时,及时跟踪评估实施情况,总结相关经验、逐步优化完善制度机制。

5.3 大力推进科学数据中心的国际化建设,提升跨境治理的标准化和人才队伍水平

一是优化科学数据中心的建设布局,打造一批兼具规模、质量和数据共享能力的,具有世界影响力的科学数据中心,提高科学数据中心的国际化水平。在保证安全的前提下,为全球科学界的研究提供高质量的数据供给。另外,还可依托大科学装置与国际大科学计划,加快建设相应领域的世界科学数据中心。二是提高现有科学数据中心吸纳、汇交、管理和分析科学数据的能力^[21],在中国全球领先的学科领域,大力提升数据中心建设的质量。三是借鉴欧盟建立和推广FAIR原则的举措,加快中国科学数据管理规范的制定与国际推广。四是加强科学数据跨境治理人才队伍建设,开展常态化的人才能力建设。

5.4 加强具有全球视野的系统研究,深度参与科学数据跨境流动的治理

一是积极开展具有全球视野的科学数据跨境流动治理的系统研究,有效支撑政策决策。二是积极参与中国优势领域科学数据的国际治理。目前,中国在高能物理、大气研究以及人类基因等领域的科学数据开放共享水平居世界前列,可依托这些领域的优势,积极参与以UNESCO,WDS和ICODATA为代表的国际科学数据组织的议事机制,从领域数据中心(存储仓)建设、国际标准认证、管理制度建立与推广,以及数据挖掘工具的开发等方面,深度参与并适时引领领域数据中心的国际治理。三是依托制度型开放地区的优势及已有的国际科技合作基础,试点探索科学数据跨境规则的衔接和先行先试。如探索香港面向英美、澳门面向欧盟的科学数据跨境规则衔接,探索广西面向RCEP,上海、海南等面向CPTPP、DEPA等相关多边经贸协定中科学数据跨境规则的衔接。同时,结合已有的国际科技合作布局,如广东面向以色列、成渝面向共建“一带一路”国家、广西面向东盟国家、宁夏面向阿拉伯国家、湖北湖南面向非洲国家等,开展面向更大区域的科学数据跨境探索,扩大中国科学数据跨境规则的影响力。

致谢 感谢国家自然科学基金(L2124029)和亚太经济合作组织(APEC)科技创新政策伙伴关系基金(PPSTI 052023A)的资助。

推荐阅读文献

- 1 Tolle K M, Tansley D S W, Hey A J G. The Fourth Paradigm: Data-intensive scientific discovery. *Proc IEEE*, 2011, 99: 1334–1337
- 2 Gewin V. Data sharing: An open mind on open data. *Nature*, 2016, 529: 117–119
- 3 Qu J, Huang K. The development logic and future mission of open science (in Chinese). *Chin Sci Bull*, 2022, 67: 4312–4325 [曲建升, 黄河敏. 开放科学的发展逻辑与未来使命. *科学通报*, 2022, 67: 4312–4325]
- 4 Wen K, Song D C, You D Y, et al. Features and current situation of the evolution of China's scientific data open sharing policy from the perspective of policy bibliometric (in Chinese). *Chin Sci Bull*, 2024, 69: 1175–1182 [温珂, 宋大成, 游玓怡, 等. 政策计量视角下中国科学数据开放共享政策演进与体系构建. *科学通报*, 2024, 69: 1175–1182]
- 5 Guo H D, Chen H S, Yan D M, et al. Strengthening the construction of open data infrastructure and promoting the development of open science (in Chinese). *Bull Chin Acad Sci*, 2023, 38: 806–817 [郭华东, 陈和生, 闫冬梅, 等. 加强开放数据基础设施建设, 推动开放科学发展. *中国科学院院刊*, 2023, 38: 806–817]
- 6 Liu Y. Regulation and improvement of cross-border flow of data in the World Trade Organization reform process (in Chinese). *Intellect Prop*, 2023, (4): 108–126 [刘影. 世界贸易组织改革进程中数据跨境流动的规制与完善. *知识产权*, 2023, (4): 108–126]
- 7 Jiang E B, Li N. Research on research data open sharing of eu in open science environment (in Chinese). *World Sci-Tech R&D*, 2020, 42: 655–666 [姜恩波, 李娜. 开放科学环境下的欧盟研究数据开放共享研究. *世界科技研究与发展*, 2020, 42: 655–666]
- 8 Wang J L, Sun J L. Review, reform and prospect analysis of world data center (in Chinese). *Adv Earth Sci*, 2009, 24: 612–620 [王卷乐, 孙九林. 世界数据中心(WDC)回顾、变革与展望. *地球科学进展*, 2009, 24: 612–620]
- 9 Wang W T. The WDC-D Coordinating Committee was formed in January of this year (in Chinese). *J Glaciol Geocryol*, 1989, 11: 260 [王文伟. WDC-D协调委员会于今年元月成立. *冰川冻土*, 1989, 11: 260]
- 10 A R H, Wu C, Li X X. International practice of open research data and its enlightenment to China (in Chinese). *Bull Chin Acad Sci*, 2020, 35: 11–18 [阿儒涵, 吴丛, 李晓轩. 科研数据开放的国际实践及对中国的启示. *中国科学院院刊*, 2020, 35: 11–18]
- 11 Liu Y. International governance on data cross-border flow (in Chinese). *Lilun Daobao*, 2023, (5): 61–63 [刘影. 数据跨境流动的国际治理. *理论导报*, 2023, (5): 61–63]
- 12 Hong Z G, Huo J X. RCEP's regulation on cross-border data flow and its important impact (in Chinese). *Southwest Fin*, 2022, (4): 83–94 [洪治纲, 霍俊先. RCEP对数据跨境流动的规制及其重要影响. *西南金融*, 2022, (4): 83–94]
- 13 Wang L. Policies and practices on scientific data management by US Federal Government (in Chinese). *Glob Sci Technolo Econ Outlook*, 2018, 33: 47–51 [王炼. 美国联邦政府科学数据管理政策及实践. *全球科技经济瞭望*, 2018, 33: 47–51]
- 14 Huang X Z, Mi J N, Zhang C P, et al. The evolution, knowledge system and method tools of scientific data reuse—A concurrent discussion of the influence of the fourth research paradigm (in Chinese). *Sci Res Manag*, 2022, 43: 100–108 [黄欣卓, 米加宁, 章昌平, 等. 科学数据复用研究的演化、知识体系与方法工具——兼论第四科研范式的影响. *科研管理*, 2022, 43: 100–108]
- 15 Gao Li. Make scientific data “open and shared” in science and technology management (in Chinese). *Chin J Sci*, 2022-06-08 [高雅丽. 在科技管理中, 让科学数据“开放共享”. *中国科学报*, 2022-06-08]
- 16 Wang J L, Li Y, Wang S Q, et al. Global impact analysis and implementation strategy recommendations of the FAIR principles (in Chinese). *Chin Sci Bull*, 2024, 69: 1183–1191 [王卷乐, 李扬, 王淑强, 等. FAIR原则全球影响分析及其实施策略建议. *科学通报*, 2024, 69: 1183–1191]
- 17 Leng S H Y, Su T Y, Zhang L. Practice and reflections on open sharing of marine scientific data (in Chinese). *Chin Sci Bull*, 2023, 68: 2239–2249 [冷疏影, 苏天贽, 张亮. 海洋科学数据开放共享实践与思考. *科学通报*, 2023, 68: 2239–2249]
- 18 Guo H D, Zou Z M, Chen G, et al. Thoughts on the support of scientific data for major national strategic requirements (in Chinese). *Chin Sci Bull*, 2024, 69: 1116–1122 [郭华东, 邹自明, 陈刚, 等. 科学数据支撑国家重大战略需求的若干思考. *科学通报*, 2024, 69: 1116–1122]
- 19 Liao F Y, Hu L L, Wang J, et al. Research and suggestions on scientific data security standards (in Chinese). *Chin Sci Bull*, 2024, 69: 1142–1148 [廖方宇, 胡良霖, 王健, 等. 科学数据安全标准研究与工作建议. *科学通报*, 2024, 69: 1142–1148]
- 20 Su J. Consideration and suggestion on strengthening scientific data management in the era of big data (in Chinese). *Chin Soft Sci*, 2022, (9): 50–54 [苏靖. 大数据时代加强科学数据管理的思考与对策. *中国软科学*, 2022, (9): 50–54]
- 21 Li Y, Su J B. Towards good governance of data: A case study in geoscience data governance (in Chinese). *Chin Sci Bull*, 2024, 69: 1149–1155 [李新, 苏建宾. 走向数据善治: 以地球科学数据治理为例. *科学通报*, 2024, 69: 1149–1155]

Summary for “科学数据跨境流动治理的现状、思考与展望”

Current status and prospects for scientific data's cross-border governance

Jun Wang, Xiaolin Zhou, Yunyi Shen, Ziyu Li, Xiaoping Ren & Yun Yang*

National Center for Science and Technology Evaluation, Beijing 100081, China

* Corresponding author, E-mail: yangyun@ncste.org

With the rise of the fourth scientific research paradigm featuring Big Data and the need of international Science and Technology (S&T) cooperation, scientific data has become as a significant component of fundamental S&T resources. Scientific data's cross-border flow rules have also gradually become a key issue for S&T governance. As China accelerates the construction of an open innovation ecology and participates actively in open science, the demand for sharing, management, and comprehensive utilization of scientific data is also increasing. Therefore, mastering the rules of scientific data's cross-border flow is necessary and has a critical meaning in this context.

This study first clarifies the scientific data's definition and its cross-border flow's main features, then analyzes the institutional construction of multilateral /international organizations, European Union and United States' scientific data cross-border flow from two dimensions—passive cross-border flow formed by open and sharing, and active cross-border flow formed by cross-border transmission. The multilateral governance system presents a multilevel and fragmented feature, while the governance system of the E.U. reveals a feature of “loose inside, strict outside”. Concurrently, the governance system of the U.S. restricts the free outflow but allows the free inflow of scientific data.

Secondly, the research sorts out the history of China's scientific opening and sharing data based on China's Five-Year Plans. Then, it focuses on the institutional construction of scientific data's cross-border transmission after the enactment of China's “Three Key Laws on Data”, specifically Cyber Security Law, Data Security Law, and Personal Information Protection Law. The research finds three primary ways with five routes available for scientific data to move cross-border under “Three Key Laws on Data”. They are: (1) scientific data's free flow; (2) flow with conditions in which data exit security assessment, personal information standard contract, and personal information authentication protection are its three routes; (3) restricted cross-border flow.

After the above analysis, based on the problems and challenges faced, the research proposes four recommendations for China's cross-border governance of scientific data: (1) Strengthening the top-level design and innovative governing measures of scientific data's cross-border flow. Specifically, balancing the relationship between development and security, and fully understanding the necessity of open sharing of scientific data in the context of international scientific research environment and open innovation ecology construction. In addition, trying out innovative governance methods, such as setting up a negative list of the scientific data's cross-border flow, establishing a “green channel”, and other appropriate steps. (2) Improving the whole governance system from both top-down and bottom-up. Specifically, accelerating the formulation of standards and policies on scientific data's classification, important data categories, and critical information infrastructure; guiding national research funding agencies to improve their data management system with data management plan (DMP); and conducting comprehensive compliance review of scientific data that has been open and shared after the full implementation of new regulations. (3) Enhancing the extent of scientific data center's internationalization, standardization, and capacity building by optimizing the construction layout of centers, and improving existing data centers' ability to absorb, transfer, manage, and analyze scientific data. (4) Accelerating the pace of deep participation in the global governance of scientific data's cross-border flow from aspects of actively participating in the international governance activities; relying on the institutional open regions and existing S&T cooperation base to make the trial and convergence of different rules; and actively conducting related research from a global perspective.

scientific data, cross-border flow, open and sharing, global governance, data cross-border

doi: [10.1360/TB-2023-1057](https://doi.org/10.1360/TB-2023-1057)