

我国科学数据开放共享的实现路径研究

陈晓玲^{1*}, 全志薇², 马卓³, 刘臻²

¹吉林省科学技术信息研究所大数据中心, 吉林 长春

²吉林省科学技术信息研究所信息资源中心, 吉林 长春

³吉林省科学技术信息研究所情报理论与方法研究中心, 吉林 长春

收稿日期: 2022年8月1日; 录用日期: 2022年9月5日; 发布日期: 2022年9月14日

摘要

为了更好地促进我国科学数据开放共享的践行。本文运用文献分析法、网络调查法和内容分析法, 查阅国内外相关文献和科学数据平台, 梳理出国内外科学数据开放共享开展现状, 并对其开放共享的基础条件和相关因素进行剖析, 提出了我国科学数据开放共享的实现路径。基于现有的科学数据政策、理论、技术、人才等基础条件, 详细分析科学数据的拥有者、利用者、受用者和保障者的关系、职责和作用, 提出建立政产学研协同合作模式形成科学数据的汇聚, 开展我国科学数据平台的融合服务, 形成我国科学数据的互联互通, 最后从科学数据的政策、资源、服务、流通、应用、安全、人才与交流合作角度提出相关对策建议。

关键词

科学数据, 开放共享, 实现路径, 对策建议

Research on the Realization Path of Scientific Data Open Sharing in China

Xiaoling Chen^{1*}, Zhiwei Quan², Zhuo Ma³, Zhen Liu²

¹Big Data Center, Jilin Provincial Information Institute of Science and Technology, Changchun Jilin

²Information Resource Center, Jilin Provincial Information Institute of Science and Technology, Changchun Jilin

³Information Theory and Method Research Center, Jilin Provincial Information Institute of Science and Technology, Changchun Jilin

Received: Aug. 1st, 2022; accepted: Sep. 5th, 2022; published: Sep. 14th, 2022

*通讯作者。

文章引用: 陈晓玲, 全志薇, 马卓, 刘臻. 我国科学数据开放共享的实现路径研究[J]. 社会科学前沿, 2022, 11(9): 3791-3802. DOI: 10.12677/ass.2022.119519

Abstract

In order to better promote the practice of open sharing of scientific data in China. By using the methods of literature analysis, network investigation and content analysis, this paper reviews the relevant literature and scientific data platforms at home and abroad, combs the current situation of the opening and sharing of scientific data at home and abroad, analyzes the basic conditions and related factors of its opening and sharing, and puts forward the implementation path of the opening and sharing of scientific data in China. Based on the existing basic conditions such as scientific data policies, theories, technologies and talents, this paper analyzes in detail the relationships, responsibilities and roles of the owners, users, recipients and guarantors of scientific data, and proposes to establish a collaborative cooperation model between government, industry, university and research to form a convergence of scientific data, develop the integration services of China's scientific data platforms, and form the interconnection of China's scientific data. Finally, from the perspective of scientific data policies, resources, services, circulation put forward relevant countermeasures and suggestions from the perspective of application, safety, talents, exchange and cooperation.

Keywords

Scientific Data, Open Sharing, Realization Path, Countermeasures and Suggestions

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

“科学数据”概念最早提出于1957~1958年“国际地球物理年”的世界数据中心WDC (World Data Center)建立初期[1]。科学数据是人类社会科技活动所产生的基本数据,或者是按不同要求加工的数据产品和相关信息。它是信息时代最基础、最活跃、最有影响力的大数据,具有重大的社会、经济和科学价值,在应用中能获得高质量的知识服务和数据产品,不仅使科学数据增值,还提高科技创新能力的重要支撑。科学数据开放共享是国家信息化建设的重要组成部分,是国家创新体系和科技创新能力建设中急需发展的现代科学基础设施[2]。在国家统一规划、政策调控和相应法规的保障下实施科学数据开放共享工程,应用先进技术,整合离散的科学数据资源,构建面向全社会的共享服务体系,实现对科学数据资源的规范管理和高效利用,从而为政府决策、经济增长、社会发展、科技创新和国家安全提供科学数据资源的保障。

本文试图解决以下3个研究问题:1)梳理国内外科学数据开放共享的现状,并对其开放共享的基础条件进行分析;2)构建科学数据开放共享的实现路径,解决数据壁垒和全面开放共享等问题;3)基于实现路径模型,提出科学数据开放共享的合理化建议,为我国科学数据资源的开放共享的具体实施提供了决策参考。

2. 相关综述

国外科学数据共享起步于20世纪50年代,成长于80年代和90年代,2010年后在开放科学和开放数据的指导下,数据联盟的发展得到迅速推进,国际间科学数据的广泛参与有助于其资源在更广泛的领

域和地域内获得更高效的开放共享与利用[3]。科学数据政策法规体系是数据开放共享的保障,国外众多的发达国家建立了相对完善的政策。例如:美国《开放政府数据法案》[4]、英国《英国数据能力发展战略规划》[5]、国际科学联合会(International Council for Science, ICSU)《大数据时代开放数据公约》[6]等促进科学数据开放共享的政策法规。科学数据管理平台是提高科学数据管理的重要手段,具有制定科学数据管理方案、数据组织、数据描述、数据存储、数据分析、数据共享和数据采集等功能[7]。例如:美国校际社会科学数据共享联盟(Inter-University Consortium for Political and Social Research, ICPSR)[8]、地球科学数据观测网(Data ONE)、英国数据存档库(The National Archives(United Kingdom), UKDA)等科学数据管理服务平台[9]。

20世纪90年代我国首次提出地球科学领域的数据共享,较晚于发达国家。为了加快科学数据管理和共享应用,制定了国家和地方的数据管理、数据中心建设、数据标准规范等政策。例如:《科学数据管理办法》为科学数据管理提供了重要的政策保障,指导国家重点研发科技计划项目数据汇交管理。目前,有17个省市也出台了科学数据的管理办法或实施细则,其整体框架与核心内容保持一致,部分政策则根据实际情况略有扩展,地方也在大力推进科学数据的汇交、开发与发布,构建统一的科学数据资源服务体系。促进数据中心建设,相继颁布了《关于加快构建全国一体化大数据中心协同创新体系的指导意见》、《全国一体化大数据中心协同创新体系算力枢纽实施方案》、《新型数据中心发展三年行动计划(2021~2023年)》等,从数据中心建设体系、算力枢纽和发展规划出发,提升算力基础设施,不仅在软实力方面强化数据要素的汇聚和培育,还在基础设施方面加强数据中心建设和提升算力服务。国家重点基础研究等资助项目要求承担单位建立规范、健全的项目和成果数据库,也对相关科研成果的开放共享提出了明确要求。相关数据标准以政府资助的科技计划形成的科学数据汇交标准为主,例如:《科技计划形成的科学数据汇交通用代码集》(GB/T 39908-2021)、《科技计划形成的科学数据汇交通用数据元》(GB/T 39909-2021)、《科技计划形成的科学数据汇交技术与管理规范》(GB/T 39912-2021)等,包括标准的通用代码集、通用数据元、技术与管理规范等。科学数据中心是国家信息基础设施的重要组成部分,目前有“国家高能物理科学数据中心”等20个国家科学数据中心、“国家重要野生植物种质资源库”等31个国家资源库,极大地推动了我国科学数据治理进程[10]。国家级科学数据中心开放共享平台的建设无论从组织体系、管理模式,还是资源体系和功能设计都形成了丰富的数据储备和汇聚。

许多学者也开展相关学术研究,例如:盛小平等阐述了国内外科学数据开放共享现状,从开放学、开放共享、开放数据三者关系剖析科学数据开放共享现状,开放共享障碍因素及实证研究[2]、影响因素[3]、数据安全[11]和权利治理研究、合作行为分析[12]等相关研究,为我国科学数据开放共享提供了丰富的理论研究和实证分析。王卷乐等开展科学数据标准体系[13]、数据治理研究[14]、数据中心建设[15]、数据平台建设研究,开展了国家科学数据战略规划层面和实践应用研究。温亮明等探究数据政策落实[16]、共享伦理研究[17]、数据实践进展与发展路径研究[18],从科学数据开放共享面向的问题、表现和现状出发进行理论研究。研究发现部分文献综述了国外科学数据开放共享研究,但是鲜少对我国地方科学数据资源建设与管理工作的深入揭示。为此,本文聚焦我国科学数据的开放共享,重点对促进地方科学数据的开放共享与建设进行阐述和分析,为了推动地方科学数据开放共享建设,立足我科学数据总体建设目标,深入剖析我国科学数据全面开放共享的建设模式和实现路径,以期为我国科学数据资源建设、管理与利用工作提供参考与借鉴。

3. 科学数据开放共享的基础条件

无论国家社会经济发展还是科技创新发展都离不开科学数据,因此要发挥科学数据的支撑作用,提升科学数据资源的应用价值。数字经济加剧发展亟须提升科学数据开放共享服务,要求相关科学数据机

构面向社会公众多渠道提供有针对性、精细化的科研推送式服务[19]。科学数据的开放共享将极大地促进基础资源利用、服务工作的开展以及对其隐性信息的挖掘分析, 利于改进相关机构的工作内容及方式、助力科研服务的提升。优化科技创新资源和创新生态的新需求。将通过信息化手段搭建虚拟科学数据资源整合环境, 克服区域协同发展的时空差异和跨区域行政障碍, 集聚各类创新主体和创新要素, 将创新生态系统的网络性、关联性和系统性引入科学数据资源体系建设, 基于数据资源分类分析、关联模型构建和属性定义, 形成以创新主体为核心的科学数据资源知识图谱, 通过不同服务功能模块集成实现虚拟科学数据资源创新生态系统在线运行, 提供信息共享、决策支撑、智慧服务等集成化服务。

我国实施科学数据开放共享已经具备了充分的基础条件(见图 1), 1) 从政策法规看, 开放共享政策是先行之举, 开放科学数据是重点, 也是开放获取的补充, 开放共享是桥梁和纽带, 开放科学数据是价值所在[20]。国家开放共享的治理理念为科学数据的开放共享提供了稳定的外部环境。2) 从理论研究看, 科研人员对科学数据开放共享等研究和探讨, 为科学数据开放共享提供了良好的理论支撑, 我国科学数据平台主要包括政府部门的公共数据、高等院校的机构知识库、网络科学数据库、科学数据中心、领域数据、数据出版等。3) 从基础设施看, 为了适应海量的科学数据资源存储, 国家正在不断引入和加强科学数据基础设施的建设, 也建立了实验室、数据中心、集成平台、计算中心等基础设施, 突破时间、空间和地域的限制, 加快建立全方位的科学数据基础设施。各种技术基础设施的投入和技术的研究应用, 为科学数据的开放共享提供了技术基础支撑。4) 从技术支撑看, 为了充分发挥科学数据的价值, 促进科学研究和科技创新, 科学数据的管理主要依托海量数据的存储、采集、快速索引、关联与融合、实时分析等处理技术[21]。新兴技术和工具[22]、区块链技术[23]等促进了科学数据的融合、开放共享、技术创新能力。5) 从成功案例看, 数据平台建设等行业案例为科学数据开放共享提供了成功案例, 如, 国际开放科学数据资源平台包括联合国数据、非洲开放数据、OECD 图书馆、data 等领域性或综合性平台。6) 从科学数据管理和专业人才看, 培养数据人才为科学数据开放共享提供了人力资源。

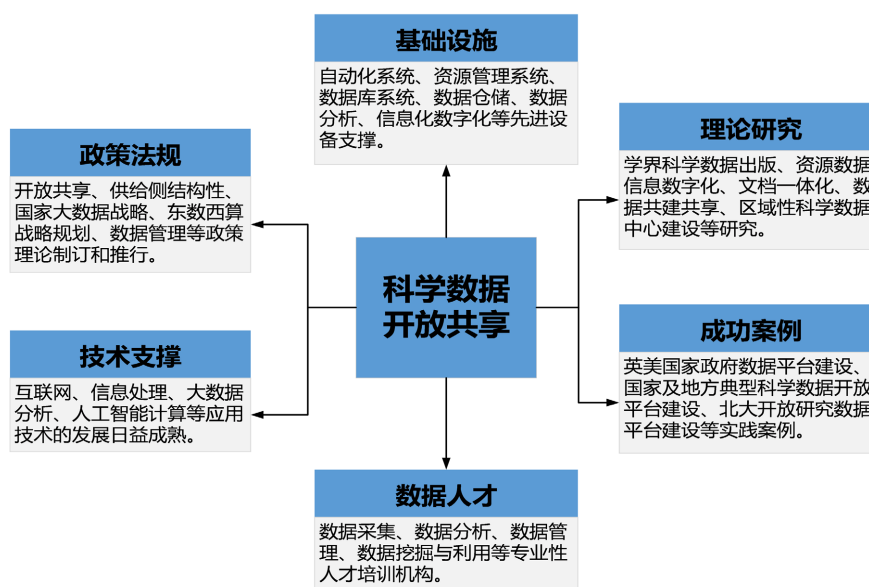


Figure 1. Basic conditions for open sharing of scientific data

图 1. 科学数据开放共享的基础条件

4. 科学数据开放共享的相关因素

为了充分发挥多元主体科学数据开放共享的作用, 科学数据只有不断被使用, 才能产生更多的价值。

科学数据可复用性、可共享性、可重现性，为了最大限度地发挥科学数据的增值作用[24]。科学数据开放共享主要涵盖数据的拥有者、利用者、受用者及保障者四类(见图 2)，是连接四方主体的桥梁，也是新时期提升科学数据开放共享的重要路径[25]。

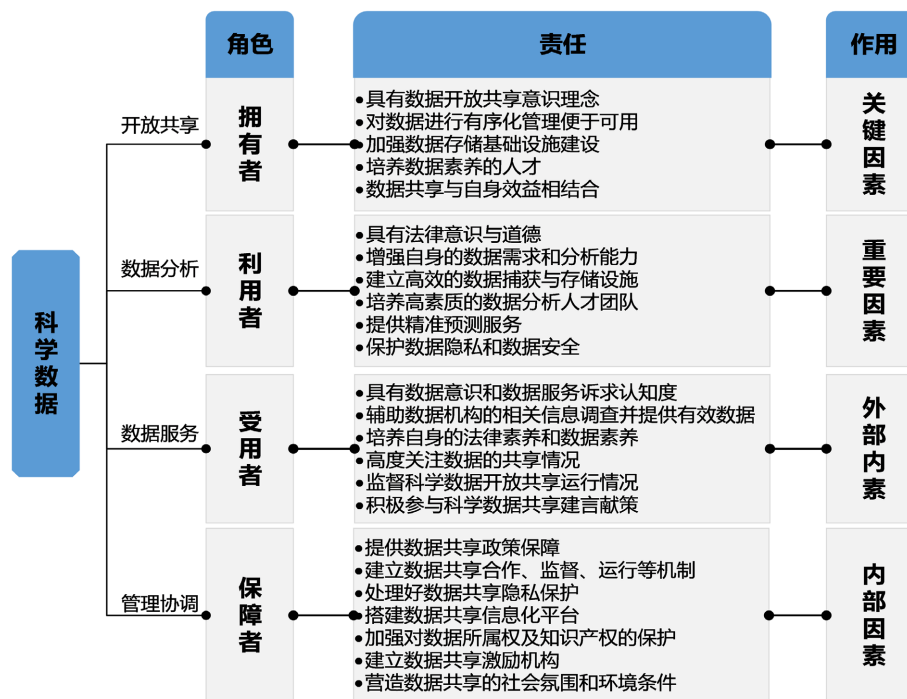


Figure 2. The perspective, responsibility and role of multiple subjects in the open sharing of scientific data
图 2. 科学数据开放共享多元主体的角度、责任和作用

4.1. 科学数据拥有者

科学数据拥有者的开放共享是检验自身科学数据工作整体水平的重要指标之一，有利于及时调整工作方向，提高社会竞争力。目前，科学数据的来源主要有两种：一是科技计划项目实施和科研人员在科学研究与实践过程中产生的实验、观察、调查等科研数据；二是政府部门长期采集和管理的公共数据。科学数据管理即是对这两类数据进行有效的管理，并促进其广泛共享，使之价值最大化[26]。拥有者对科学数据开放共享影响主要表现为科学数据管理的基础建设和数字化建设，以及相关机构或部门对其开放共享的意识理念，是实现其开放共享的关键；各科学数据服务机构及部门须对自身科学数据实施有序化管理，使其处于可用状态，加强自身办公自动化、电子文件管理系统及扫描、缩微等基础设施建设；培养员工科学数据意识，以促进其对有保存利用价值的科学数据的归档保存与系统化梳理，辅以健全的工作制度与激励奖惩制度提高工作效率。此外，科学数据机构作为盈利性组织，科学数据的开放共享与其自身发展效益相契合时，才能从根本上推动科学数据开放共享，鉴于此，可实行科学数据开放共享有偿服务等。

4.2. 科学数据利用者

科学数据利用者对科学数据信息的挖掘与分析，对自身发展方向有重要的指导作用，可促进自身的转型升级。利用者是否具备道德法律意识关系，与相关组织机构或部门的沟通协作及科学数据获取能力，是否具备有效分析科学数据的能力，是否提供有效的数据服务需求，这些都是关系到科学数据开放共享

的重要因素；科学数据利用机构须开展基础设施建设，以实现开放共享的科学数据的获取、存储与分析；培养具有科学数据管理和分析能力的专业团队；加强对科学数据服务诉求信息的整合，对科研人员进行精细化分类，创建对科研人员科学数据服务信息需求评估预测工作体系，提供精准预测服务，实现开放共享科学数据的高质运用[27]；同时，注意科学数据利用行为的合法合规性，尊重和保护科学数据隐私，支付利用科学数据的费用，促进多方再合作等。

4.3. 科学数据受用者

科学数据受用者可享有针对性的科学数据精准服务。受用者除去其经济状况等外在因素，科学数据受用者的数据意识、对数据开放共享致使自身受益的认知程度、数据信息共享意识、对自身权益的了解程度、对数据服务诉求的表达能力等均影响着科学数据开放共享的顺利开展[28]。科学数据受用者应积极参与科学数据服务机构相关信息调查，提供有效数据；提高自身法律素养和数据素养，增强个人隐私信息保护与维权意识；提高自身科学数据意识及表达科学数据服务诉求的能力；提升对科学数据开放共享情况的关注度；还可组建受用者团体，对科学数据开放共享进行监督，积极为科学数据开放共享建言献策；通过各种途径提高自身互联网参与度，增强线上提供与查取利用科学数据信息的能力等。

4.4. 科学数据保障者

科学数据保障者可提高科学数据管理的行政能力，解决和协调科学数据的开放共享过程中涉及的开放与保密的权衡、数据隐私保护、信息安全、各主体之间的利益、知识产权等问题[12]。保障者(政府)是否为科学数据开放共享提供外部政策保障，以及相关的开放程序与范围、合作共享规范机制、对知识产权的保护机制、市场运行监督机制、数据挖掘机制技术及人才扶持政策的颁布执行程度等影响科学数据开放共享的运行效率[29]。科学数据保障者须起到模范带头作用，处理好自身科学数据开放共享与隐私保护的关系，进一步搭建科学数据信息共享平台；加强对档案信息系统的安全保护，防止档案信息的损坏与丢失，促进科学数据信息数字化建设等；加强对科学数据所属权及知识产权的保护，严厉打击盗用滥用科学数据信息等行为[30]，建立科学数据开放共享激励机制，鼓励各组织机构间进行科学数据信息共享[11]。

5. 科学数据开放共享的实现路径

因此，在科学数据开放共享中，应从不同组织和国家、不同领域、不同应用场景、不同服务形式的角度，建立不同的机制或模式推动科学数据开放共享，明确科学数据的供给方和需求方的利益与诉求，实现有效的流通机制、激励机制、利益分配机制，有序推动科学数据开放共享。统筹我国科学数据资源是科技创新发展的重要任务，关键在于尽快建立一套以政府为主导、以需求为导向、以企业为创新主体、政产学研协同合作的科学数据开放共享服务体系，积极发挥各主体的积极性和主动性，大力促进我国科学数据共享服务体系中各机构、各环节、各流程的全面发展[31]。目前，我国科学数据开放共享以国家科学数据中心为主体，建立地方分中心或分站，加快科学数据的统筹规划与开放共享。为了实现我国科学数据开放共享的路径，第一步是采用政产学研协同合作模式加快科学数据资源开放共享，第二步是加强我国科学数据平台融合服务。

5.1. 建立政产学研协同合作模式

由于政府机构、高等院校、科研机构和企业是推动我国科学数据发展的主体，从而导致在数据资源开放共享的地位和作用各不相同[32]。政府机构的作用主要表现在引导和支持数据开放共享的政策，以及数据管理平台的主导建设。同时，通过加大科学研究 R & D 投入，提供科学供给和资金支持。社会数据包

括政府及公共机构在依法开展活动中收集的各类数据及其衍生数据，如自然资源数据、经济社会数据等。企业是最直接面对市场需求，负责生产和开发产品及技术需求，重点是科技成果的产业化和应用。企业数据是指企业在生产经营管理活动中产生或合法获取的各类数据及其衍生数据。高等院校主要从事基础研究和科学研究，与科研机构一起为科学数据资源开放提供技术积累、人才培养、知识创新和科研设备等。科研数据是指能够识别从事科研活动的机构或人员在科研活动中产生的数据。在政产学研各主体间相互协作和配合下，我国科学数据资源开放共享通过建设科学数据融合公共服务平台，包括上下联动、机构联通、学科交叉等，有效促进了科学数据资源的统筹互动与共享发展。

依托政产学研建立科学数据战略联盟(见图3)，集成共享各主体间的优势科学数据资源，与科研院所、高等院校、企业及社会组织团体等建立协同服务，以科学数据开放共享为基础，共同开展机构间科学数据的分析挖掘方法、科学数据的升值与应用、科学数据作为生产要素参与生产等方面研究，实现新旧动能转化。对策建议：创新科学数据管理体制，破除机制约束，打破科学数据在地区、机构、领域的分散，整合多源、异地、异构的科学数据资源；加强科学数据合理流动协同和协调机制，增强共享意愿。建设我国科学数据开放共享平台融合服务，形成与机构、项目、成果、论文、知识产权等相关的数据关联，促进科技中介服务；有利于完善科学数据开放共享的数据安全制度、知识产权制度、监督与考核制度以及激励机制；构建我国科学数据开放共享应用生态系统，促进大数据科学发展，为数据时代我国科技创新发展提供支撑[33]。建立政产学研科学数据合作和跨区域合作，国家科学数据中心与省内外高等院校、科研机构开展数据合作，形成我国科学数据中心合力，在战略规划、技术研发、人才培养形成优势互补；项目驱动合作，通过承担国家资源资助的科研项目，联合开展技术攻关和数据储备建设[34]；重点领域合作，通过与地方的优势产业、重点机构进行合作，建立相关领域、特色的科学数据资源存储；宣传培训合作，通过不定期的联合举办专题讨论会、研讨会、培训会等开展交流合作。综上所述，我国通过不断加强政产学研在更广泛范围内的大量合作与协同创新，积极推动了科学数据资源的互动与开放共享。

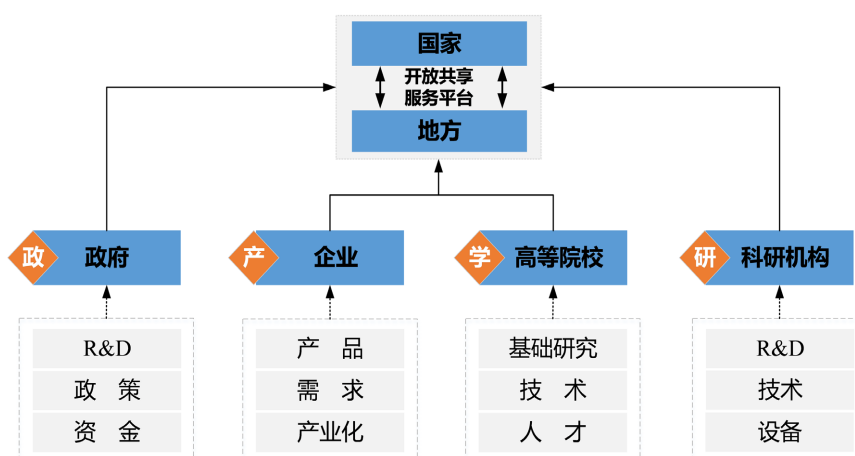


Figure 3. Realization path of open and shared scientific data in politics, industry, university and research
图3. 科学数据开放共享的政产学研实现路径

5.2. 开展我国科学数据平台融合服务

在开放科学理论指导下，产生了涵盖开放的政策、资源、数据、开源软件、开放实验室等元素的科研全生命周期。因此，在建立科学数据开放共享的融合服务平台时，所有与科学数据相关的元素都应该是互联互通的(见图4)，如依托机构、拥有者、科研项目、学术论文、科技成果、标准、知识产权等元素，利于科学数据的系统化与关联化。

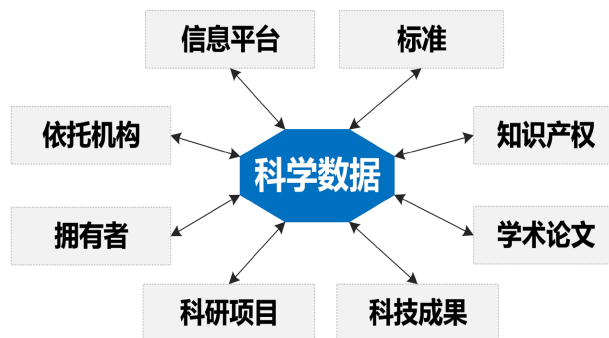


Figure 4. Scientific data interconnection
图 4. 科学数据互联互通

科学数据资源是一种专业性较强的稀缺资源，从产生到利用易出现信息不对称性和独立性，这将严重影响科学数据资源的开放共享。国内外实践证明，构建科学数据管理平台是实现科学数据资源的高效利用。我国在科学数据开放平台已经取得了一定成效，如国家科技基础条件平台中心致力于推动科技资源优化配置，承担科学数据、生物种质、实验材料等科技基础条件资源的建设与管理工 作，实现开放共享。浙江省在全国率先启动数字化改革，围绕高水平创新型省份和科技强省建设，引领全省科技创新数字化改革持续迭代升级，建立“浙里关键核心技术攻关应用平台”、升级“中国浙江网上技术市场 3.0 服务平台”和“职务成果转化在线”，以数字思维、数字技术实现职务科技成果转化流程再造和制度重塑。江苏省以推进资源开放共享与服务平台基地建设为抓手，建立“江苏省科技资源统筹服务平台”，围绕创新要素共享体系建设“一条主线”，推进线上服务云平台、线下服务共同体全域全链融通。因此，国家和地方省市开展了不同层次的科学数据、公共数据、科技资源的建设，并取得了一定的成效，统科学数据资源优势已转化为创新优势、产业技术优势和经济优势。

开展我国科学数据平台融合建设(见图 5)，1) 在服务范围上，科学数据平台应具备数据共享、仪器设备、产业监测、科技金融、创新主体、公共研发和交流合作等服务，同时集成相关平台；2) 在数据导航上，

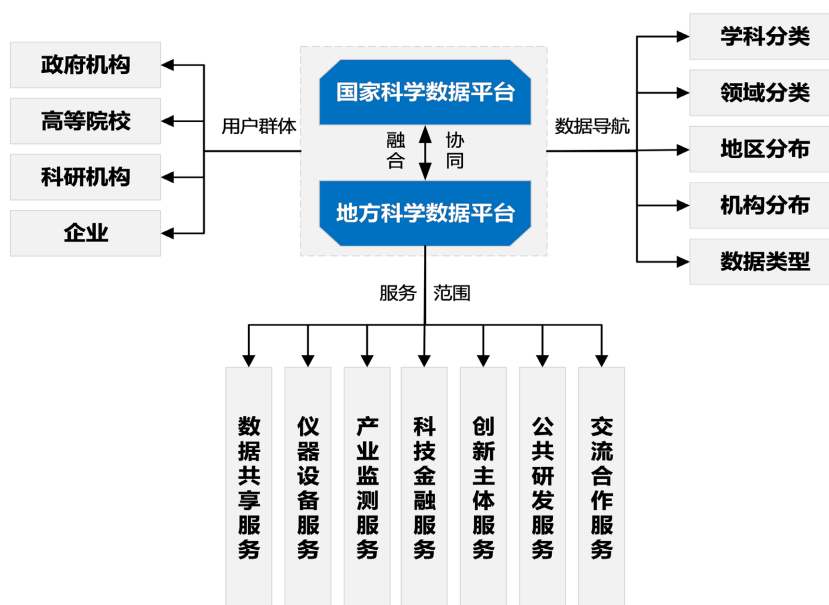


Figure 5. China's scientific data platform integration service
图 5. 我国科学数据平台融合服务

提供按学科分类、领域分类、地区分布、机构分布、数据类型等导航功能；3) 在用户群体上，平台应面向政府机构、高等院校、科研机构以及企业等创新主体。对策建议：为用户提供基本的科技文献服务、科学数据、仪器设备配套服务、国际交流合作等应用场景，进一步推动科学数据资源向社会开放；利用市场要素和产业需求，不断加强平台建设，通过开放共享的方式，最大限度地开发、挖掘和配置资源，为科技型企业、孵化器、新型研发机构提供科研服务；促进创新主体的科技成果转化，总体在盘活众多数据资源、优化整体创新环境；形成我国科学数据平台的协同合作、融合发展和互补互联；在完善平台构建的同时，应提供相应的政策支撑、项目资助、资金扶持、税收优惠补贴、激励措施。

6. 推进科学数据开放共享的对策建议

6.1. 立足社会发展需求完善地方的科学数据政策法规和标准规范

为了制定或遵循适合科学数据管理的政策和标准。1) 立足经济社会发展战略需求，面向经济社会发展需要，按照《科学数据管理办法》等要求，依托国家级科学数据中心，建立辐射范围广的地市级科学数据中心。坚持立足科技创新，充分发挥科学数据在科研和产业发展中的技术引领和支撑作用，协调各行业共同发展。以科学数据中心为载体和纽带，从科学数据的角度全面融入数字经济建设，使我国科学数据能够服务大局、参与全局决策。同时，在新的国际国内形势下，面对新的任务和要求，进一步建设我国的科学数据战略资源储备。2) 建立健全组织保障体系，根据科学数据资源多源异构的复杂特点，制定完善的科学数据标准体系[35]，包括数据采集与加工、元数据、数据管理、数据质量、数据服务等相关的标准和规范，内容需要覆盖科学数据资源全生命周期的关键业务节点，并根据需要制定相应的指导规范。3) 建立我国数据平台的对接合作机制，确保接口顺畅和数据格式互通。同时，各地科学数据中心之间要建立交流机制，不断学习先进技术和典型服务案例。4) 优化科学数据的生态环境，切实发挥科学数据的作用，体现科学数据的价值，把科学数据“搞活”，让科学数据更好的“说话”，使科学数据能够“繁衍”。5) 在全国范围内持续推进科研项目科学数据汇交管理，按照《科学数据管理办法》定期开展科学数据组织、管理、呈交、共享等方面的奖惩工作，将科学数据呈交情况纳入科研项目考核评价，将科学数据汇交情况纳入科技发展计划项目结项环节。6) 鼓励发表科学数据相关学术论文，根据国际形势发展的要求，对于拟在国外期刊发表的论文中的科学数据进行认证，使其效力等同于将数据发给国外期刊认证。

6.2. 形成我国的纵向流通横向合作的科学数据资源储备

为了有效建立科学数据的资源体系。1) 在科学数据采集方面，构建数据生命周期的全流程数据链，推进数据采集、标注、存储、传输、管理、审核等全流程管理，打破政府、高校、科研机构、企业等数据壁垒，实现数据传输、控制、管理、应用等多源数据一体化集成和汇聚。构建不同机构间的数据采集和共享机制，推动不同领域数据标注关联与管理应用的实施。推动不同设备的数据采集和互通互操作，推广各级统一的数据交换架构，制定关键设备数据接口标准。2) 整合省级科学数据资源向国家科学数据中心的汇聚，针对成长期已形成的省级科学数据资源的收集和保存，在科学数据资源采集的前期，应统筹规划和设计与国家科学数据中心平台的数据接口服务，建立一体的数据标准规范，从而为后期省级科学数据资源深度融入国家科学数据中心平台提供技术和理论框架支撑。3) 逐步建立地市级科学数据资源向省级科学数据中心的集成，加强省级科学数据中心的汇聚整合能力，服务于地方科技创新和社会经济发展。促进科学数据汇交制度的实施，坚持“开放为常态、不开放为例外”的理念，充分分析、挖掘、共享科学数据，打好“科学数据”产品。对存储的科学数据，建立数据的公开权限、开放条件、开放对象和审核流程。建立数据应急机制和管理系统，进行有效的科学数据进行备份或异地备份。

6.3. 构建我国的科学数据服务与流通机制

国家科学数据管理机构与科学数据平台、地方科学数据管理机构与科学数据平台形成互联互通，建立跨地区、跨部门、跨领域的科学数据管理协同和融合发展，形成行业和领域的分工协作管理模式，避免重复建设，加强我国科学数据中心之间的优势互补。加强国家科学数据中心的国际化建设，推动其全球影响力和服务能力，带动国际间的科学数据合作交流，进而全面提升国家的科学数据治理能力。大力推进地方科学数据的基础建设，推动地方科学数据的建设和发展，带动本地区的社会、经济和科技的发展，从而加快地方的数据治理水平。建立科学数据相关主体的相关权利和责任，构建科学数据权利的基本框架，明确数据权利类型、权利主体、控制边界和使用权限，建立科学数据统一登记制度，对原始数据、脱敏化数据、模型化数据和标准化数据等权属界定和流转进行动态管理，分层分类分级推进科学数据的价值和流通，以科学数据的应用需求为导向，完善科学的数据流通环境，精准对接市场需求，坚持多元协同共治原则，充分发挥政府、高校、科研机构、企业数据资源优势，营造可持续发展环境。

6.4. 开展科学数据的情报分析和智慧应用服务

为了进一步探索和提高科学数据的情报分析和智慧应用服务。1) 深化数据驱动的全流程实践应用，提高数据情报分析的智能化，实现不同领域、行业科学数据的全流程和综合应用，提升科学数据的应用水平。2) 扩充科学数据应用模式，建立重点行业、领域等特色数据库，建立重点行业数据应用标杆，分行业或领域建立科学数据应用路径、模式方法和发展重点。3) 建立科学数据指导规范和应用指南，编制特定的科学数据应用指南，为科学数据应用提供指导方向，建立科学数据资源目录，指导法人单位对自身科学数据进行分级分类。4) 采用先进技术搭建科学数据应用平台，应用适当的技术开发科学数据平台。该平台负责承担全省科学数据的汇交、审核、宣传、培训、咨询与技术支持的保障服务，通过线上下载、离线共享、定制服务等方式向社会开放共享。

6.5. 建立科学数据安全和风险防范机制

为了确保科学数据的安全和隐私保护，应建立一个具有知识产权机制的科学数据安全体系。1) 数据安全监管体系，首要任务是形成数据安全主体的责任，数据的具体防护要求，建立全流程的数据安全监管与管理体系。2) 数据保障能力，加强数据动态感知、统计评估、监测与预算等支撑和保障能力。3) 数据安全技术，采用数据监测、防护、加密、脱敏等安全技术来加强安全保护要求，构建数据安全备案机制来提高数据安全事件的处理能力。4) 数据流通机制，建立数据流动风险防控机制，加强数据流动监测和业务协同监管。5) 数据知识产权保护政策，巩固重点领域数据基础设施的安全保障，增加关于保护专利、数字版权、商业秘密、隐私数据等知识产权保护指导政策。

6.6. 培养科学数据人才和合作交流活动

为了我国科学数据的发展，培养具有数据素质的科学数据中心管理人才和研究团队，1) 在数据管理方面，加快发展壮大从事科学数据的专业人才和人才团队，探索数据人才的激励与保障措施，通过职位、奖励、绩效考核等方式提高专业人才的收益，从而提升科学数据服务人才的专业水平。2) 在数据应用方面，要培养或引进高层次科学数据分析挖掘人才，高层次科学数据专业人才将提升科学数据分析与挖掘能力，支撑科学数据资源的管理和应用。积极引导社会组织和企业合法合规地开发科学数据资源，向社会提供增值服务。3) 科学数据纳入人才激励机制，结合“破五唯”工作的全面开展，将科学数据呈交情况纳入人才评价工作，将评奖、职称晋级、科技进步奖评选等有机结合，从而打开科技人才工作的新局面。

广泛开展国内外科学数据的交流与合作，在构建科学数据中心时，要充分考虑和借鉴国际科学数据中心的战略规划和导向作用，积极探索多种方式的科学数据联合共建与共享模式，开拓全新的各层级的科学数据中心的合作交流模式，科学数据中心的建设必须紧跟国际科学数据发展趋势，充分借鉴国际科学数据建设与发展的成熟经验。1) 开展科学数据机构“1+N”共建，保持与国内外重点机构、数据组织的紧密的合作交流，依托科学数据资源，开展科学数据的分析挖掘方法、科学数据增值应用、科学数据作为生产要素参与生产等方面的研究。2) 建立促进科技成果转化的国内外科学数据项目合作，开展科学数据中的学术研究和科学关键技术攻关，以及科技成果转化下科学数据实用模式，与科学数据的提交者、制造者共同开发科学数据，形成数据产品，部分科学数据已经广泛支持了行业和企业的发展，或者基础性与实用性较强的科学数据的研究成果。科学数据的转移转化是科技成果转化的重要组成部分。3) 全面实施科学数据宣传与培训工作，不断向全社会宣传科学数据工作的重要性，营造适合科学数据工作的好环境，在科研工作者中树立科学数据意识，使我国科学数据相关机构和人员对科学数据工作有更深入地了解，掌握一定的科学数据收集和报送能力，不断提高科学数据生产能力和质量。

综上所述，为了提升科学数据高质量发展，在重点行业或领域科学数据、重大科研攻关数据等方面形成权威的品牌科学数据资源体系；建立不同学科、不同领域、不同类型的科学数据分析工具和算法平台，从而提高科学数据的再分析、再利用、再开发和再挖掘能力；把握国家战略规划、地方社会发展需求和科技前沿，提升科学数据资源的精准推荐和定制化服务水平，形成线上、线下和离线等多种形式并存的数据共享服务体系；加强科学数据知识产权保护，形成数据安全保障体系；注重科学数据管理和技术人才的培养，提高科学数据人才的专业化水平；整体优化科学数据资源布局，深化多方国内外合作与交流，全面提升我国科学数据的国际影响力。

7. 结语

加快促进国家科学数据的开放共享，对于加快国家基础条件整合、促进政产学研用、数据融合一体化发展具有重要的理论和实践意义。首先，探索国内外科学数据的开放共享应用情况，挖掘未来的科学数据资源的开放共享发展趋势，提出我国的科学数据开放共享的可行性分析，进而释放我国科学数据资源。同时，我国科学数据的开放共享具有扎实的研究基础，无论是政策法规、理论研究、基础设施和先进技术，还是人才支撑和应用实践。其次，分析了科学数据开放共享中人的相关因素，从数据拥有者、利用者、受用者和保障者的角度阐述了科学数据开放共享的地位、责任和作用，使科学数据开放共享得到充分释放和发挥。再次，分析了科学数据开放共享中机构的相关因素，从政府机构、高等院校、科研机构、企业的数据及其数据诉求出发，探索了机构间的科学数据开放共享的实现路径，一是加强政产学研用的综合协调和融合发展，增加了机构间的共享意愿，实现了科学数据资源的互联互通和共享耦合；二是开展我国科学数据平台的互联互通，建立多用户、多功能、多服务的科学数据平台，有助于消除数据资源的信息孤岛和不对称，形成数据要素的自由流通、绩效与激励机制、知识产权保护制度等一套完备的运行机制。最终，充分利用创新主体的市场需求极大，极大盘活科学数据资源的存量，实现科学数据资源的应用价值最大化。

基金项目

2021年 ISTIC-SPRINGER NATURE 开放科学联合实验室开放基金“开放共享视角下中国科学数据的实现路径研究”(项目编号: IT2135)研究成果之一。

参考文献

- [1] Nicolet, M. (1984) The International Geophysical Year (1957-1958): Great Achievements and Minor Obstacles. *Geological Journal*, 8, 303-320. <https://doi.org/10.1007/BF00185934>

- [2] 盛小平, 杨智勇. 开放科学、开放共享、开放数据三者关系解析[J]. 图书情报工作, 2019, 63(17): 15-22.
- [3] 张晓青, 盛小平. 国外科学数据开放共享政策述评[J]. 图书馆论坛, 2018, 38(8): 147-154.
- [4] 中国国际经济交流中心大数据战略课题组. 发达国家如何布局大数据战略[J]. 中国经济报告, 2018(1): 87-89.
- [5] 崔宇红. E-Science 环境中研究图书馆的新角色: 科学数据管理[J]. 图书馆杂志, 2012, 31(10): 20-23.
- [6] 虞路清, 肖云, 黎建辉. 科学数据库数据共享政策研究与实施[C]//中科院科学数据库办公室. 科学数据库与信息技术学术研讨会论文集第八集. 2006: 322-327.
- [7] 卫军朝, 张春芳. 国内外科学数据管理平台比较研究[J]. 图书情报知识, 2017(5): 97-107.
- [8] Swanberg, S.M. (2017) Inter-University Consortium for Political and Social Research (ICPSR). *Journal of the Medical Library Association*, **105**, 106-107. <https://doi.org/10.5195/jmla.2017.120>
- [9] 王卷乐, 王明明, 石蕾, 高孟绪, 陈明奇, 郑晓欢, 等. 科学数据管理态势及其对我国地球科学领域的启示[J]. 地球科学进展, 2019, 34(3): 306-315.
- [10] 黄铭瑞, 李国庆, 李静, 范湘涛. 国家科学数据中心管理模式的国际对比研究[J]. 农业大数据学报, 2019, 1(4): 14-29.
- [11] 盛小平, 王毅. 利益相关者在科学数据开放共享中的责任与作用——基于国际组织科学数据开放共享政策的分析[J]. 图书情报工作, 2019, 63(17): 31-39.
- [12] 王瑞丹, 高孟绪, 石蕾, 王超, 徐波. 对大数据背景下科学数据开放共享的研究与思考[J]. 中国科技资源导刊, 2020, 52(1): 1-5+26.
- [13] 王卷乐, 石蕾, 徐波, 王玉洁, 高孟绪, 王超. 我国科学数据标准体系研究[J]. 中国科技资源导刊, 2020, 52(5): 45-51+77.
- [14] 王卷乐. 后疫情时代的全球科学数据治理关切[J]. 国外社会科学, 2021(5): 58-64.
- [15] 王淑强, 王卷乐, 李扬, 王晶, 王玉洁, 李海英. 基于文献计量学的国际地球科学数据管理研究进展[J]. 全球变化数据学报(中英文), 2020, 4(3): 299-313+后插 93-后插 107.
- [16] 李洋, 温亮明. 《科学数据管理办法》落实现状、影响因素及推进策略研究[J]. 图书情报工作, 2021, 65(2): 65-74.
- [17] 温亮明, 李洋. 我国科学数据开放共享模式、标准与影响因素研究[J]. 图书情报研究, 2021, 14(1): 33-41.
- [18] 温亮明, 李洋, 余波. 基于区块链技术的《科学数据管理办法》落实路径探析[J]. 现代情报, 2021, 41(8): 136-146.
- [19] 党洪莉. 社会科学数据的开放与共享: 发展现状、障碍与出路[J]. 图书馆理论与实践, 2018(5): 70-74.
- [20] 温亮明, 李洋, 郭蕾. 我国开放科学研究: 基础理论、实践进展与发展路径[J]. 图书馆论坛, 2022, 42(2): 22-35.
- [21] 黎建辉, 沈志宏, 孟小峰. 科学大数据管理: 概念、技术与系统[J]. 计算机研究与发展, 2017, 54(2): 235-247.
- [22] 景贵飞. 开放科学对全球综合地球观测系统建设影响分析[J]. 国土资源遥感, 2020, 32(4): 1-7.
- [23] 寇蕾蕾, 祝忠明, 张伶, 王思丽. 开放科学视域下科研众包平台的功能与服务研究——以 Daemo 平台为例[J]. 图书馆学研究, 2020(5): 59-66.
- [24] 盛小平, 吴红, 胡冰洁. 科学数据开放共享障碍的实证研究[J]. 图书情报工作, 2019, 63(17): 23-30.
- [25] 盛小平, 孙倩倩. 国内科学数据开放共享主题的作者合作关系与合作研究内容分析[J]. 图书情报工作, 2021, 65(23): 13-21.
- [26] 司莉, 邢文明. 国外科学数据管理与共享政策调查及对我国的启示[J]. 情报资料工作, 2013(1): 61-66.
- [27] 盛小平, 田婧, 向桂林. 科学数据开放共享中的数据质量治理研究[J]. 图书情报工作, 2020, 64(22): 11-24.
- [28] 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.
- [29] 盛小平, 吴红. 科学数据开放共享活动中不同利益相关者动力分析[J]. 图书情报工作, 2019, 63(17): 40-50.
- [30] 盛小平, 郭道胜. 科学数据开放共享中的数据安全治理研究[J]. 图书情报工作, 2020, 64(22): 25-36.
- [31] 盛小平, 袁圆. 国内外科学数据开放共享影响因素研究综述[J]. 情报理论与实践, 2021, 44(8): 173-179+102.
- [32] 张旺, 程慧平. 科学数据开放共享策略机制及优化路径研究[J]. 情报杂志, 2020, 39(5): 154-161.
- [33] 李成赞, 张丽丽, 侯艳飞, 周园春, 黎建辉. 科学大数据开放共享: 模式与机制[J]. 情报理论与实践, 2017, 40(11): 45-51.
- [34] 盛小平, 秦颢洋. 公众科学视角下的科学数据开放共享合作行为分析[J]. 图书情报工作, 2021, 65(23): 4-12.
- [35] 齐法制, 陈刚, 程耀东. 建立权责明晰且能力健全的科学数据开放共享机制——以高能物理领域为例[J]. 中国科学基金, 2019, 33(3): 229-236.