

长三角跨区域科学数据中心建设研究^{*}

储节旺 杨婷婷

(安徽大学管理学院, 合肥 230601)

摘要: 在科学研究的新范式中, 科学数据已经成为重要的基础科研资源, 促进区域间科学数据的整合, 建立跨区域科学数据中心显得尤为必要。本文在研究长三角跨区域科学数据中心建设必要性的基础上, 分析建设基础, 设计建设框架, 并提出建设重点, 具体包括协调落实科学数据汇交政策、确立规范的数据标准体系、建设分布式数据处理平台、完善科学数据出版机制、健全数据中心的配套制度。

关键词: 跨区域科学数据中心; 长三角一体化; 科学数据开放; 科学数据标准体系

中图分类号: G203 **DOI:** 10.3772/j.issn.1673-2286.2022.06.008

引文格式: 储节旺, 杨婷婷. 长三角跨区域科学数据中心建设研究[J]. 数字图书馆论坛, 2022 (6) : 58-64.

在科学研究过程中积累的科学数据作为一种基础性数据, 伴随着知识经济时代的到来, 已经成为支撑国家和地区经济发展的重要战略性资源。广义上的科学数据包括科研工作者在进行科学研究全过程中所产生的过程数据以及研究成果^[1]。国务院办公厅2018年印发的《科学数据管理办法》对科学数据的外延进行了确定, 科学数据主要包括在自然科学、工程技术等领域, 通过基础研究、应用研究、实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^[2]。

互联网时代, 科学数据呈现出5V: Volume (海量)、Velocity (高速)、Variety (多样)、Value (低价值密度)、Veracity (准确)的典型大数据特征^[3]。与此同时, 科学研究进入了“第四范式”——数据密集型科学, 数据密集型科学以海量科学数据的分析应用为基础^[4]。科学数据中心作为科学数据存储与管理的重要载体, 随着全球科学数据的开放共享进程不断推进, 其建设进程也不断加快。1957年国际科学联合理事会(ICSU)成立世界数据中心(World Data Center)^[5]。此后, 欧美大多数国家开始探索建立自己的科学数据中心, 建成了如美国国家空间科学数据中心(National Space Science Data Center)、英国数字保存中心(Digital

Curation on Centre)、欧洲数据联盟中心(EUD-CA)等数据中心^[6]。国内科学数据中心的建设相较于国外起步较晚, 2003年科技部开始试点建设气象科学数据中心、测绘科学数据中心、水文水资料科学数据中心等6个数据中心。2011年, 科技部部署了23个国家科技基础条件平台, 其中包括6家科学数据共享领域和5家自然科技资源共享领域的的数据资源平台^[7]。到2019年, 我国的科学数据中心建设取得了较大的成果, 建成了包括国家地震科学数据中心、国家人口健康科学数据中心等20个国家级科学数据中心。目前国内已经建成的科学数据中心以自然科学领域为主, 其中国家高能物理科学数据中心、国家青藏高原科学数据中心、国家地震科学数据中心、国家空间科学数据中心、国家极地科学数据中心是国内5个典型的科学数据中心^[8]。

在新的科学研究范式中, 促进科技资源的开放与共享是当前优化与整合科技资源的重要途径^[9]。2018年国务院发布的《关于建立更加有效的区域协调发展机制意见》指出, 要推动京津冀、长三角等几大城市群发展, 强化区域板块间的科技创新合作, 促成区域间的优势互补与协同发展^[10]。当前科学数据受物理空间位置、行政管辖等因素的约束, 区域内部科学数据的同质性较高, 而异质性科学数据较为缺乏。各省市已建成

^{*} 本研究得到国家社会科学基金项目“开放创新环境下大数据知识生态系统协同演化与智慧涌现研究”(编号: 20BTQ069)资助。

的科学数据服务平台的科学数据标准体系尚未统一、数据服务能力较为欠缺,致使科学数据的流动性大幅降低。为了健全国家科技创新体系,迫切需要探索建设一批跨区域科学数据中心,对区域间的科学数据进行有效整合,建立起共建共享、开放高效的科技研发支撑体系^[11]。

1 跨区域科学数据中心相关研究现状

通过文献梳理可以发现目前国内尚无直接针对跨区域科学数据中心的研究,但部分科学数据中心相关研究可以为跨区域科学数据中心研究提供借鉴。①对科学数据中心的数据共享关键技术进行探讨。石京燕等^[12]在介绍国家高能物理科学数据中心分布式数据处理平台的基础上为科学数据中心的跨区域数据资源共享与数据访问提供了参考方案。卢逸航等^[13]对科学数据中心间互操作模式现状进行分析,总结了两大类科学数据中心间互操作的模式,即全局互操作模式和局部互操作模式。②对科学数据中心涉及的数据政策以及协议的评估分析。如崔雁^[14]利用re3data.org注册机制,对科学数据中心的数据类型、使用许可、元数据标准等进行了多角度的分析,为科学数据的进一步开放提供相关建议。文禹衡等^[15]以扎根理论为基础构建科学数据中心用户注册协议合规性评价体系,经评价当前多数科学数据中心的用户注册协议合规程度不高,需要在用户信息的保护、风险预警机制等方面做出进一步的完善。

长三角区域作为国内科技发展先行区,具有丰富的科技资源,但各地受制于行政区域的划分,整体上对科技资源的配置和利用效率不高。因此本文以长三角区域为例,探索建设跨区域科学数据中心相关问题,对其建设的必要性、建设基础进行探讨,分析科学数据中心建设过程中的重点问题,对长三角跨区域科学数据中心的建设和发展提出对策建议。

2 长三角跨区域科学数据中心建设的必要性

目前,长三角区域一体化发展已经上升为国家战略,在国务院印发的《“十四五”数字经济发展规划》中提到要加快构建全国一体化大数据中心体系,并在长三角等区域建设数据枢纽节点,优化数据中心的总

体统筹规划,而建设长三角跨区域科学数据中心正是该发展规划在科学数据领域的落脚点,该数据中心起着作为国家级平台的对接与补充的作用^[16]。长三角地区的各省市都分别上线运行了各自的科学数据共享平台,包括江苏省科技资源统筹服务中心^[17]、安徽省科技创新云服务平台^[18]、上海科技创新资源数据中心^[19]等。目前这些平台都由各政府部门运营,各省市之间经济与科技发展的不均衡导致了科学数据共享平台建设进度的差异,上海地区作为全球科技创新中心,其科学数据资源共享的发展走在其他省市的前列。2020年2月21日上海科技创新资源数据中心正式成为欧洲开放科学云EOSC (the European Open Science Cloud) 的首家非欧洲会员机构,这也标志着上海市的科技资源共享迈出了与国际接轨的重要一步^[20]。其他省份虽然也有相应科技资源共享平台的建设基础,但是进展相对较慢,呈现出各地区数据中心建设进度不协调的“分裂”现状,也极大阻碍了科学数据中心跨区域数据资源的传递与共享。此外,目前各个平台制定的科学数据标准与规范存在较大差异,容易造成后续科学数据之间对接的困难。因此打破行政壁垒,探索建设长三角地区分布式科学数据库,推动区域内科学基础设施、大型科研仪器、科技成果库等科技资源的开放共享与合理流动,才能最大程度地发挥这些科技资源的价值。长三角跨区域科学数据中心的建立能够加快形成长三角科技创新共同体,加速科研成果转化效率,推动长三角区域建设成为科技创新高地以及高质量发展先行区。

3 长三角跨区域科学数据中心的建设基础

长三角区域一直是我国经济活力较强,科技发展较快的地区。2021年长三角地区GDP总量占全国比重为24.13%,能够为区域的科技创新提供坚实的经济基础^[21]。长三角地区聚集了大量的科研机构 and 科技创新人才,历经20余年的科技合作发展,各省市之间具有深厚的合作基础,在多个层面具备建立跨区域科学数据中心的基础条件。

(1) 科学数据资源丰富。长三角地区拥有全国近四分之一的“双一流”高校,“211”高校有25所,上海、合肥两地分别建有国家综合科学中心,三省一市共配备大科学装置15个、国家重点实验室74个、创新型机构17 686家,高端科研人才和高科技企业众多,省级以上

企业技术中心在500家以上,科技创新链的各环节分布均衡^[22]。如江苏省科技资源统筹服务中心收录生物样本1 730 123份,农业种质就达到38 934份^[17]。

(2) 政策支持力度大。2016年12月长三角地区的三省一市共同签订《关于共同推进长三角地区协同创新网络建设合作框架》;2018年11月举办的“长江经济带科技资源共享论坛”上,四地签署了《长三角科技资源共享服务平台共建协议书》^[23];2019年5月通过的《长江三角洲区域一体化发展规划纲要》中也明确描述了长三角地区的科技资源共享的发展规划,要推动长三角地区科技资源的合理流动与开放共享^[24]。

(3) 协同创新基础深厚。在2022年发布的《长三角区域协同创新指数2021》中,长三角区域协同创新指数的年均增速达到了9.54%,2020年的创新指数较2011的指数翻了一番;就优秀科技人才的流动性而言,长三角三省一市的科技创新人才的跨区域流动达到了165万人次;就科学成果开发与共享而言,长三角区域内专利转移数量达到17 741件,41个地级以上城市全部参与到科技论文合作网络中^[25];就科技成果转化效果而言,在上海张江建立的长三角国家技术创新中心即将成为长三角产学研深度融合的示范中心^[26]。

4 长三角跨区域科学数据中心的建设框架

长三角跨区域科学数据中心的建设框架(见图1)包括科学数据整合体系、科学数据分析体系以及科学数据服务体系3个模块的内容。

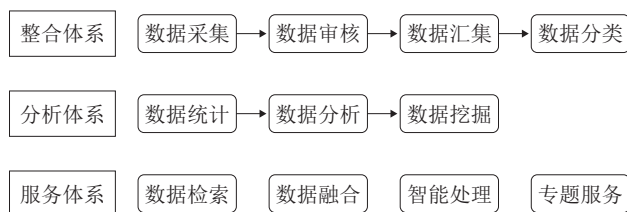


图1 长三角跨区域科学数据中心的建设框架

(1) 科学数据整合体系重点面向长三角区域的四大优势特色产业(生物医药、物联网、集成电路、智能网联汽车),整合上述领域的科学数据库,推进长三角区域科学数据的集成化和一体化进程。科学数据的采集是整个科学数据整合工作开展的起点,跨区域科学数据中心可以采用目前大多科学数据中心所使用的“自

建+提交+外采”的数据采集方式,数据生产者在对科学数据进行登记时不仅要上传科学数据实体,同时还应补充对科学数据进行描述的信息,对处理数据所使用的软件也应进行说明或者一并提交^[22]。随后科学数据中心对采集到的科学数据开展详细的数据审核,审核通过后科学数据中心将与用户签订数据汇交协议,对于审核没有通过的数据也应进行说明。跨区域科学数据中心应带动省级科学数据的汇入,积极探索与各种科研单位的合作与共建,不断提高数据中心数据的数量以及质量,形成丰富的科学数据资源体系,让科学数据资源更好地服务长三角区域的科技创新^[27]。跨区域科学数据中心必须对不同来源和不同类型的数据进行统一的整合与规划。科学数据的分类体系可以按照“学科+区域”相结合的方式进行,按照数据资源所属的不同学科和区域建立完整的科学数据分类体系。

(2) 科学数据分析体系是数据中心通过整合长三角区域内的分布式计算机以及各类大型科学分析仪器等基础设施,以“一平台多中心”的管理模式对跨区域数据进行一站式处理^[12]。跨区域科学数据中心在对收集的海量科学数据进行基础性统计分析的基础上,利用人工智能等大数据分析技术对科学数据进行系统性挖掘与分析,形成可访问、可复用的数据产品,更深度地揭示科学数据之间的内在关联,将科学数据中心打造成为一个直观、立体的可视化数据访问系统^[28]。

(3) 科学数据服务体系是直接面向用户的模块,科学数据检索是跨区域科学数据中心所提供的最基础的数据服务——面向所有平台注册和认证通过的用户提供开放科学数据的检索服务。跨区域科学数据中心应根据长三角地区科技发展的重点需求为导向,提供具有区域特色的数据服务,促进跨学科、跨领域的数据融合,长三角跨区域科学数据中心所存储的资源类型涉及物理、生物、化学等多个基础学科,这些资源是众多科学研究领域的关键基础资源,科学数据中心在积极推动跨领域数据深度融合的基础上,激发科学数据的应用价值,让科学数据更好地服务于科研创新活动。跨区域科学数据中心可以向用户提供机器学习、数据同化、时间序列分析等多种数据分析的方法库与模型库,帮助用户多角度、全方位地对科学数据进行智能处理与分析^[29]。跨区域科学数据中心的建设要避免成为传统的单一数据仓储,不能单纯地着眼于数据内容的建设,目前已建成的科学数据共享平台所提供的数据服务形式较为单一,更好地为科研从业人员提供服务才是建

立科学数据中心的根本目的,应该探索多样化的服务形式,更好地服务于用户,表1列举了一些跨区域科学数据中心可以提供的数据服务^[30]。

表1 跨区域数据中心可提供的数据服务

数据服务类型	具体内容
特定需求数据爬取和清洗	数据中心面向用户的特定需求,根据具体数据使用场景,提供数据爬取以及数据清洗的一站式服务,帮助用户高效、精准地获取所需数据
特色数据库	面向长三角地区的优势产业(如生物医药、物联网、集成电路等)建立起专项科学数据集
主题数据库	面向长三角区域的重点基础科学研究课题,建立主题科学数据库

5 长三角跨区域科学数据中心的建设重点

5.1 协调落实科学数据汇交政策

稳定的数据资源体系是科学数据中心建设的关键。跨区域科学数据中心在建设过程中要重视科学数据汇交制度的建设,目前长三角区域各省市的科学数据汇交政策(见表2)中除浙江省外,都在《科学数据管理办法》基础上根据当地实际情况制定了本地区的科学数据管理实施细则。其中安徽省和上海市的管理实施细则中明确提出科学数据中心建设的细则以及相应职责。上海市还就“市科学数据管理中心”的主要职责进行细致的规定,对科学数据汇交政策做出了要求,为后续行动的规范提供很好的规章。在这几项管理办法中,只有安徽省强制要求由省级政府资金资助的项目所产生的科学数据必须汇交到相对应的科学数据中心^[31]。以上实施细则虽然为科学数据的汇交与管理提供了依据,但是其落实仍然存在很多阻碍,例如它们更多是针对政府部门和数据中心等机构的主要职责进行规范,但是并未针对科学数据的利益相关者制定明确的行为规范,可能会导致实施细则的落实受到科学数据利益相关个人的自觉性的影响。因此,一方面应该细化科学数据利益相关人员的行为规范,另一方面还应加强对科研工作者行为的引导。上述实施细则都是各地根据自身情况制定,所以难免会出现一些差异与分歧,各地政府主管部门应该加强政策沟通与交流,保持长三角跨区域科学数据中心建设过程中政策

制定的一致性^[32]。

表2 长三角区域各省市的科学数据汇交政策

发布省份	政策	发布时间
上海市	《上海市科学数据管理实施细则(试行)》 (草案)	2021年1月6日
江苏省	《江苏省科学数据管理实施细则》	2019年2月19日
安徽省	《安徽省科学数据管理实施办法》	2018年11月18日
浙江省	《浙江省公共数据条例》	2022年3月1日

5.2 确立规范的数据标准体系

长三角区域的科学数据类型多样,必须建立一个统一的数据标准规范体系对科学数据进行有效的整合与集成,应根据科学数据生命周期的全过程建立一个完善的标准体系(见图2)。科学数据标准体系的专用标准应该包括数据采集、数据审核、数据分类到数据发布的数据生命周期全流程。同时,跨区域科学数据中心作为综合型的数据仓储,元数据也是其进行高质量的数据管理所要考虑的问题之一,因此数据中心应该基于自身的主要侧重学科与数据中心的定位,在保持现有元数据组织方式的传统优势的基础上进行创新,为科学数据中心所存储的数据提供更加多样化的关联方式,并在数据的易用性和简洁性之间寻求平衡^[33]。此外,除了确定整个数据中心的专用数据标准,对于一些学科数据库、主题数据库以及专题数据库,数据中心可以在专用标准的基础上提出针对这些特殊数据库的指导标准。如针对主题数据库,可以在指导标准中对数据库的总体架构进行统一定义,并明确该类型数据在运行以及服务等方面的要求。

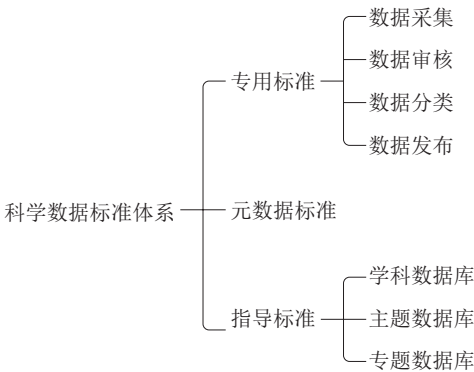


图2 科学数据标准体系

5.3 建设分布式数据处理平台

长三角跨区域科学数据具有数据规模大、结构复杂、分布广等特点, 如何对海量分布式异构数据进行有效管理对数据共享技术提出了更多的挑战。目前已经建成的科学数据中心主要是通过各区域分中心以及数据资源点之间以“物理上分布, 逻辑上统一”的形式构成数据共享服务网络系统^[13]。长三角区域的科学数据分散在各个地区的科研机构中, 长三角跨区域科学数据中心同样可以采取上述形式来建设, 建成1个总中心、4个区域分中心的分布式数据处理平台(见图3)。赵瑜等^[34]针对各分布式节点数据资源之间协同与共享, 提出了一种基于元数据的分布式数据统一访问技术。该方法利用元数据检索服务和数据库统一访问消除异构资源之间的差异。由于分布式云服务平台的用户认证将更加复杂, 因此必须采用安全的跨域联盟认证方式, 目前比较通用的是基于tokens的统一身份认证方式, 用户通过认证后会得到经过加密后的token, 随后用户在异地节点资源的合法作业身份也将由token来提供^[12]。通过上述技术构建起的分布式处理平台将有效帮助科学数据的整合。

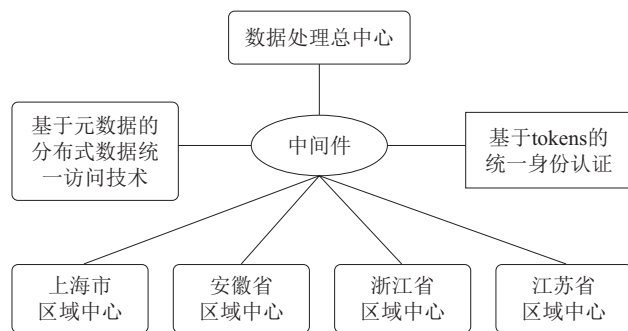


图3 分布式数据处理平台

5.4 完善科学数据出版机制

跨区域科学数据中心是由政府统筹协调、多元主体参与建设的数据平台, 数据中心的高效运行需要充分激发数据价值链上各主体的积极性和主动性。根据普丽娜等^[35]于2019年对上海市571名科研从业人员的调查显示, 94.8%的科研从业人员有意愿进行科学数据共享, 但共享的前提是在后续的科研成果中对他们的科学数据进行数据权益保护。数据出版是在科学数

据开放过程中通过引入数据引用等机制, 对科学数据生产者的权益进行保护的科学数据共享方式。当前国内外的数据出版模式大体上可以分为学术论文关联出版、数据存储平台出版和数据论文出版三类, 目前多数科学数据中心的数据出版模式主要是第二类^[36]。中国科学院则是第三类科学数据出版机制的主要推行者, 分别于2016年和2017年创办了《中国科学数据》和《全球变化数据学报》这两本探索数据论文出版模式的期刊^[37]。在科学数据出版中, 最重要的问题就是建立科技资源标识体系, 当前中国科学院科学数据中心体系主要采用科技资源标识CSTR(China Science and Technology Resource)对科技资源进行唯一标识, CSTR具有以下特点: 唯一性、持久性、兼容性、互操作性、动态更新^[38]。长三角跨区域科学数据中心与中国科学院数据中心体系一样具有多学科、跨区域的特点, 因此同样也可以采用科技资源标识CSTR来对科技资源进行标识, 这也便于建立起统一的数据接口, 实现与国家级科学数据中心的资源对接。在对科学数据进行统一标识之后, 在后续的科学数据引用以及出版制度中, 科学数据也将成为数据生产者的重要研究成果, 因而能够激发其进行数据共享的积极性。

5.5 健全数据中心的配套制度

跨区域科学数据中心的可持续发展离不开良好的配套制度, 包括人才队伍的建设、资金来源的管理, 以及绩效考核制度。建设一支具有专业技能的高素质人才队伍是科学数据中心运行的重要基础。因此, 必须完善人才培养体系, 采用科学合理的激励机制, 拓宽员工的晋升渠道, 留住数据中心建设所需的专业人才, 同时吸引更多的青年人才加入到队伍中来^[39]。数据中心应该定期对员工进行专业技能培训, 提高员工的专业素养, 确保员工能够与最新的前沿技术接轨。在资金保障制度方面, 跨区域科学数据中心在运行的多个方面, 如科学数据的管理、科技服务的提供、人才队伍的建设都需要长期稳定的资金支持, 因此应吸引更多的政府资金与科研基金参与到科学数据中心的运营中来, 为跨区域科学数据中心建设一个稳定的资金平台。在绩效考核制度方面, 应该建立一套有效的绩效考核指标体系, 不断完善制度体系的建设, 推行标准化管理, 有效提高跨区域科学数据中心的服务水平。

6 结语

建设跨区域科学数据中心是当前探索区域科技协同发展、整合优化科技资源的重要途径。本文以长三角地区为典型案例分析跨区域科学数据中心的建设路径, 长三角地区虽然在建设科技资源共享平台上已经有了一定的基础与经验, 但在具体实现的过程中还是必须要坚持一体化的建设原则, 在建设过程中探索出一套能够高效运营的管理机制, 在发挥各省市基础优势的基础上开展跨区域、跨学科的开放科学合作, 将长三角跨区域科学数据中心打造成科技资源的集聚地。

参考文献

- [1] 秦小燕, 初景利. 科学数据素养内涵结构研究[J]. 图书情报工作, 2019, 63(18): 30-39.
- [2] 国务院办公厅关于印发科学数据管理办法的通知[EB/OL]. [2022-06-02]. http://www.gov.cn/zhengce/content/2018-04/02/content_5279272.htm.
- [3] 王国成. 从3V到5V: 大数据助推经济行为的深化研究[J]. 天津社会科学, 2017(2): 94-99.
- [4] 刘敬仪, 江洪, 廖宇. 德国地球科学领域科学数据中心调查与启示[J]. 数字图书馆论坛, 2019(12): 52-58.
- [5] 王卷乐, 孙九林. 世界数据中心(WDC)中国学科中心数据共享进展[J]. 中国基础科学, 2007(2): 36-40.
- [6] 曹乔卓然, 陈祖刚, 李国庆, 等. 科学数据中心资源和用户访问控制体系[J]. 大数据, 2022, 8(1): 98-112.
- [7] 黄铭瑞, 李国庆, 李静, 等. 国家科学数据中心管理模式国际对比研究[J]. 农业大数据学报, 2019, 1(4): 14-29.
- [8] 李正超. 国内科学数据共享平台建设现状及发展策略研究[J]. 图书馆理论与实践, 2018(8): 108-112.
- [9] 种国双, 赵衡, 裴小兵. 科技资源共享: 需求、服务机制和服务模式[J]. 科技管理研究, 2019, 39(22): 28-34.
- [10] 中共中央 国务院关于建立更加有效的区域协调发展新机制的意见[EB/OL]. [2022-06-02]. http://www.gov.cn/zhengce/2018-11/29/content_5344537.htm.
- [11] 李剑锋, 刘杨, 付强. 吉林省科学数据共享服务平台建设方案探讨[J]. 科技资讯, 2019, 17(12): 21-22.
- [12] 石京燕, 黄秋兰, 汪璐, 等. 国家高能物理科学数据中心分布式数据处理平台[J]. 数据与计算发展前沿, 2022, 4(1): 97-112.
- [13] 卢逸航, 李国庆, 陈祖刚. 科学数据中心间互操作模式研究[J]. 数据与计算发展前沿, 2022, 4(1): 69-83.
- [14] 崔雁. 科学数据开放中数据中心政策分析与建议[J]. 图书情报工作, 2016, 60(8): 73-78.
- [15] 文禹衡, 金铭. 科学数据中心用户注册协议的合规性评估研究[J]. 图书情报知识, 2022, 39(1): 119-129.
- [16] 国务院印发《“十四五”数字经济发展规划》[EB/OL]. [2022-06-02]. http://jsz.mof.gov.cn/zt2019/ysjg/zcfg/202204/t20220427_3806519.htm.
- [17] 江苏省科技资源统筹服务中心[EB/OL]. [2022-02-09]. <https://www.jssic.cn/#/index>.
- [18] 安徽省科技创新云服务平台[EB/OL]. [2022-06-02]. <http://kjt.ah.gov.cn/bsfw/abmflbsml/ahskjcxwfwpt/index.html>.
- [19] 上海科技创新资源数据中心[EB/OL]. [2022-06-02]. <http://www.sstir.cn/>.
- [20] 上海科技创新资源数据中心成欧洲开放科学云亚洲首家成员机构[EB/OL]. [2022-06-02]. <https://baijiahao.baidu.com/s?id=1659158720370259654&wfr=spider&for=pc>.
- [21] 张泉, 白冬梅, 邹成东. 长三角副中心城市促进区域经济协同发展研究——以合肥市为例[J]. 现代管理科学, 2022(3): 19-29.
- [22] 长三角科技资源共享服务平台[EB/OL]. [2022-06-02]. <https://www.csjpt.cn/researchBase/researchBaseIndex>.
- [23] 长三角率先构建我国区域协同创新共同体[EB/OL]. [2022-06-02]. http://m.xinhuanet.com/ah/2018-03/18/c_1122552784.htm.
- [24] 中共中央 国务院印发《长江三角洲区域一体化发展规划纲要》[EB/OL]. [2022-06-02]. http://www.jiaxing.gov.cn/art/2021/12/16/art_1228968107_59498168.html.
- [25] 储晓露. 2021长三角区域协同创新指数发布[J]. 今日科技, 2022(2): 14-15.
- [26] 长三角国家技术创新中心在上海揭牌成立[J]. 华东科技, 2021(7): 10.
- [27] 刘杨, 陈晓玲, 全志薇, 等. 国家级科学数据中心建设对吉林省科学数据中心建设的示范效应[J]. 中国管理信息化, 2022, 25(3): 173-178.
- [28] 米琳莹, 崔辰州, 樊东卫, 等. 国家天文科学数据中心发展思路浅析[J]. 农业大数据学报, 2019, 1(4): 37-45.
- [29] 国家青藏高原科学技术中心[EB/OL]. [2022-06-02]. <http://data.tpdc.ac.cn/zh-hans/data/analyze/list>.
- [30] 司莉, 李月婷, 邢文明, 等. 我国科学数据共享平台绩效评估实证研究[J]. 图书馆理论与实践, 2014(9): 30-35.
- [31] 支凤稳, 赵宇晨. 我国科技计划项目科学数据汇交政策分析[J]. 图书馆学研究, 2021(23): 45-52.
- [32] 丁亚男, 单艺, 刘东来, 等. 我国各地区科学数据管理办法比较

- 分析研究 [C] // 财金观察 (2021年第2辑), 2022: 379-385.
- [33] 黄国彬, 王涛. 综合型科学数据仓储元数据创建服务研究 [J]. 图书情报工作, 2021, 65 (21): 131-140.
- [34] 赵瑜, 李晓东, 张新建. 基于元数据的分布式数据统一访问技术 [J]. 指挥信息系统与技术, 2019, 10 (4): 33-37, 60.
- [35] 普丽娜, 殷晓, 谢文娴. 上海科学数据管理和共享需求分析及对策 [J]. 情报工程, 2021, 7 (6): 88-100.
- [36] 袁雅琴, 胡晓彦, 佟继周, 等. 大数据开放背景下的我国空间科学数据出版实践 [J]. 中国科技资源导刊, 2022, 54 (1): 89-96.
- [37] 许玮, 王鹏涛. 作者服务视角下数据论文出版流程的再优化 [J]. 科技与出版, 2021 (10): 71-78.
- [38] 科学数据中心 [EB/OL]. [2022-02-09]. <https://www.casdc.cn/support/identifying>.
- [39] 文雯, 刘斌, 赵金国. 科技文献平台资源共享架构与运行机制研究——以贵州省产业创新科技文献服务平台为例 [J]. 农业图书情报学报, 2021, 33 (3): 68-77.

作者简介

储节旺, 男, 1971年生, 博士, 博士生导师, 研究方向: 知识管理, E-mail: chujwmail@163.com。

杨婷婷, 女, 1998年生, 硕士研究生, 研究方向: 知识管理。

Study on the Construction of Cross-Regional Scientific Data Center in Yangtze River Delta

CHU JieWang YANG TingTing

(School of Management, Anhui University Hefei, 230601, P. R. China)

Abstract: In the latest paradigm of scientific research, scientific data has become an important basic scientific research resource, and it is especially necessary to promote the integration of scientific data between regions and to establish cross-regional scientific data centers. Based on studying the construction necessity of the cross-regional data center in the Yangtze River Delta region, this article analyzes the construction basis, designs the construction framework, and proposes the construction priorities, which include coordinating the implementation of scientific data remittance policies, establishing a standardized data standard system, building a distributed data processing platform, improving the scientific data publication mechanism, and improving the supporting system of the data center.

Keywords: Cross-regional Scientific Data Center; Yangtze River Delta Integration; Scientific Data Openness; Scientific Data Standard System

(收稿日期: 2022-06-05)

资源推介

国家科技期刊开放平台提供开放服务

为实现学术资源的开放共享,为科学研究提供强有力的基础条件,我国陆续建成学术期刊开放服务平台。2013年由全国哲学社会科学规划领导小组批准,中国社会科学院图书馆负责建设的国家哲学社会科学学术期刊数据库(National Social Science Database, NSSD)上线开通,收录社会科学类学术期刊2000多种。2018年,中国科学技术信息研究所负责建设的国家科技期刊开放平台(<http://doaj.istic.ac.cn>)上线运营,成为我国科技期刊集中开放获取平台,收录期刊1300余种,其中核心期刊占比73%。普通用户可免费获得论文在线阅读和全文下载的一站式服务,还可在线浏览期刊的关键词分析、年度发文量趋势、年度被引趋势、年度影响因子趋势、年度发文机构统计、期刊近十年学科分布等情况;加入平台的期刊社不仅能增加期刊论文的传播途径,提高期刊影响力,还可以免费获得中国科学引文索引数据库(csci.istic.ac.cn)的使用权,及时掌握期刊的引证指标。