

科学数据管理平台“社区化”建设研究

——以 Digital Commons Data 为例

王 茜 王 晨

(上海市研发公共服务平台管理中心, 上海 200235)

摘 要:[目的/意义] 本文旨在研究开放式科学数据管理平台的先进做法, 为国内建设科学数据管理平台, 实现科研模式的数字化、开放化和社群化提供借鉴。[方法/过程] 本文对国内外科学数据管理平台的相关研究文献进行分析, 并采用探索性单案例研究法, 总结知名科学数据管理平台“Digital Commons Data”的运营机制、建设模式及服务功能。[结果/结论] 开放社区平台通过数据共享和群体协作发挥科学数据的最大利用价值。以企业为主体进行市场化运作是实现科学数据管理平台“社区化”建设的有效途径; 多领域、全球化是提高平台开放性的关键所在; 关联应用、定制推送是平台专业服务的基本保障; 数据安全是平台建设的中中之重。

关键词: 开放科学 科学数据管理 Digital Commons Data 数据共享

分类号: G203; G352

DOI: 10.31193/SSAP.J.ISSN.2096-6695.2022.02.05

0 引言

大数据时代推动了科学研究组织范式的不断演进, 全球化趋势下的科研合作带来了更大范围与更深层次的知识创造与共享, 科学领域的开放程度日益提升, 科学研究正由封闭模式转向开放模式。在此背景下, 美国、欧盟和日本等陆续实施了以推动“开放科学数据”为核心的“开放科学”战略。众多高校、科研院所等机构采取了基于建设数据共享平台的数据驱动发展模式, 掌握开放共享的主动权, 以实现更高效的科研管理^[1]。截至2022年2月, 全球80多个国家和地区在国际研究数据仓储注册系统 re3data.org 中注册的科学数据仓储数量达到3810个, 我国注册数量为49个, 约占1.29%^[2]。

[作者简介] 王茜 (ORCID:0000-0001-6463-4326), 女, 工程师, 硕士, 研究方向为科技资源管理、科技人才研究分析, E-mail: qwang@sgst.cn (通讯作者); 王晨 (ORCID:0000-0002-7099-3971), 女, 硕士, 研究方向为科技人才数据, E-mail: cwang@sgst.cn。

随着开放科学战略的发展, 科学数据在网上进行开放共享、众筹、众包。OpenAIRE 的调查显示, 一种“开放身份、开放报告、开放参与、开放互动、开放预审稿、开放最终版本评论、开放平台”的 Open Peer Review 模式成为了未来科研活动的发展趋势^[3]。这种开放参与和开放互动需要一个开放的“平台社区”来支持科研人员进行互惠式的讨论, 以及数据资源的交互分享^[4]。因此, 科学数据管理平台的发展趋势也不再仅止于数据的共享与再利用, 而是尝试打造科研群组共享的“社区化”平台, 通过科研众包和群体协作, 为科研人员提供充分便利的数据共享和群体协作。

本文针对“开放科学战略下如何进行科学数据管理的社区化建设”这一问题, 采用探索性单案例研究方法, 以“Digital Commons Data”平台为例总结国外科学数据管理平台“社区化”建设模式, 以为国内科学数据管理平台的“社区化”建设提供参考。

1 文献综述

国外启动科学数据管理平台建设时间较早且相对成熟, 已进入快速发展阶段。而国内的科学数据管理平台目前仍处于积极建设阶段, 已建成的平台主要具备数据提交、收集、组织、存储、管理、分析、共享及发布等功能。

从已发表的研究文献来看, 目前国内外关于科学数据管理平台的研究主要集中在以下四个方面: (1) 平台的功能定位和特点研究。Lee J.-S 和 Jeng W 指出, 美国校际社会科学数据共享联盟存储库 ICPSR 是全球最大的社会科学数据基础设施, 存储了超过 50 万个文件的社科数据^[5]; Scot M 指出, 英国数据存档库 UKDA 解决了诸如数据收集、存放等问题, 并通过元数据制作促进了数据二次分析和定性材料的再利用^[6]; 另外, 普渡大学 PURR 平台、Dataone 等平台的目标均是实现科学数据的安全存储和共享^[7]; 项英^[8]、张慧^[9]、朱玲^[10]以及卫军朝^[7]等分别研究了国内的科学数据管理平台, 认为武汉大学社会科学数据管理平台、北京大学开放研究数据平台、中国科学院科学数据管理平台等的主要功能是实现数据的保存和管理; Steinhart G 和 Dietrich D 等人研究了美国康奈尔大学 DataStaR 存储库在研究数据整个生命周期中的作用^[11]; Wilson J 指出, 英国牛津大学嵌入式机构数据监管服务 EIDCSR 旨在确保数据生命周期各个阶段生成的研究数据可以得到安全的保存和记录, 以备未来使用^[12]; 张计龙认为, 国内部分平台, 如复旦大学社会科学数据管理平台, 也在逐步从数据的保存和共享过渡到为科研活动提供全周期服务中^[13]。(2) 平台开发模式和技术选型研究。概括而言, 国内外科学数据管理平台的构建主要有自主开发、商业软件购买和在成熟的开源数字资产系统上进行二次开发三种方式。如: 哈佛大学自主开发了 Dataverse^[14], 复旦大学^[13]、北京大学^[15]则购买了 Dataverse 软件作为科学数据管理基础平台, 康奈尔大学的 DataStaR 是以自主开发的开源软件 Fedora 作为科学数据管理基础平台^[16], 武汉大学是基于 Dspace 进行了二次开发^[17], 中国科学院的科学数据管理平台则是以自主开发为主^[7]。(3) 平台绩效评估体系研究。司莉^[18]和周宇^[19]等人基于层次分析法构建了科学数据管理平台的绩效评估体系, 主要评价指标涉及数据资源、数据管理制度、平台功能、服务效能、平台界面和软件系统等; 辛一研究了我国九个省份的科学数据管理平台建设, 并从网站页面、数据资源、数据获取、共享服务等方面构建了评价体系^[20]。(4) 开放科学背景下, 平台的

数据共享模式研究。马合等指出,从20世纪中叶起,美国、澳大利亚和一些欧洲国家如英法德等相继制定了一系列科学数据管理和开放政策来推动科学数据管理的开放与共享建设^[21];我国在“十四五”规划中也提出了构建国家科研论文和科技信息高端交流平台、建设国家数据统一共享开放平台、保障国家数据安全^[22];Wilson J^[12]和Ayris P^[23]等指出英国牛津大学嵌入式机构数据监管服务EIDCSR以及欧洲开放科学云EOSC等平台均建设了共享和开放的数据环境;Lee J.-S和Jeng W还研究了开放科学背景下ICPSR 35年的归档数据,揭示了代表数据共享趋势的数据特征^[5];Maynard M指出Data-PASS采用了开放、低壁垒的合作模式且其成员能够互补^[24]。

总体来看,国内外现有的对科学数据管理平台的研究主要集中在平台功能、开发模式、绩效评价、数据共享模式等方面。虽然已有不少学者提到了在开放科学背景下科学数据共享的重要性和必要性,但是目前国内外科学数据共享模式的研究主要集中在如政策倡导型数据共享、开放科学电子基础设施共享、开放共享数据环境建设以及低壁垒合作数据共享等。在实际建设中,由于数据开放共享政策不统一、不同利益相关者需求不同,目前国内外科学数据管理平台建设仍存在数据来源单一、数据共享政策不完善、共享机制不健全、数据管理流程与规范不成熟等诸多问题。对于科学数据管理平台,以数据管理者需求为导向,构建“社区化”合作共享机制,是推动科学数据共享的有益探索。

目前,仅有美国地球数据观测网Data One Community、Dryad Community^[7]和Digital Commons Data三个平台建设了“开放社区”功能。这三个平台可以分为两种发展模式,一种以“推动平台发展”为主,另一种以“推动数据交流”为主。两种发展模式存在较大不同,前者面向平台合作伙伴,后者面向所有科研人员;前者旨在推动平台的发展,后者旨在推动科研人员的数据交流。

表1 国外科学数据管理平台“开放社区”的建设现状

类型	平台社区名称	社区目标	社区成员	功能、形式
“推动平台发展”型	美国地球数据观测网 Data One Community	代表社区的需求和利益;为 Data One 的发展和活动提供指导	研究人员、信息管理人员、图书管理员、数据作者、用户(成员存储库)和其他合作伙伴	每年召开一次社区会议
	Dryad Community	促进数据的发布、管理和保存	向所有利益相关组织开放,如期刊、学术团体、出版商、研究机构、图书馆	在活动上发表观点;投票选举董事会章程;与其他成员合作制定解决方案
“推动数据交流”型	美国地球数据观测网 Data One 成员存储库 member repository	成为开放数据运动的领导者;促进数据共享	免费注册,对所有个人和组织开放	通过 DataOne API 公开的数据集合可以被更广泛的受众发现;积极支持现有的数据工具以及为 DataOne 定制开发的新工具;让科研工作被科学家发现,在未来创造合作的机会;支持将项目数据的访问限制到合作伙伴范围内
	Digital Commons Data	分享想法;了解领域的新事物;讨论重要的发展	免费注册,对所有个人和组织开放	为研究人员提供一个与所在领域的其他人联系的专用空间,与世界各地的同事合作、分享正在阅读的内容、随时进行讨论;私有组允许共享添加到组中的 PDF 文档的全文,实现组内同步共享;提供定制的推荐内容;支持个性化选择在个人资料中展示哪些信息;查看自身数据何时何地引

以“推动数据交流”为主的科学数据管理“开放社区”平台顺应了科研模式的数字化、开放化和社群化要求, 能够有效地增强科研共同体建设, 提升科研质量^[25]。目前我国的科学数据管理平台服务对象多局限于国内或本校科研人员, 尚未形成面向全球科研人员的“社区化”平台, 因而, 学习科学数据管理平台“社区化”建设经验对加强我国的科学数据、科研人员的交流与共享具有非常重要的现实意义。本文通过单案例研究法, 对“Digital Commons Data”进行系统调研与分析, 以期为推动国内科学数据管理平台“社区化”建设提供借鉴。

2 研究设计

2.1 研究方法和案例选择

本文采用单一案例研究方法, 选取具有代表性、典型性和启发性的案例, 以“增强对同类事物的理解”^[26]。

本文选取 Digital Commons Data 平台作为案例研究对象, 探索开放科学战略下如何进行科学数据管理“社区化”建设。其作为科学数据管理“社区化”建设的典型代表主要体现在以下两方面:(1) 具备开放科学战略所要求的典型功能。它是由荷兰 Elsevier 公司自主开发的面向政府机构、高校、科研院所或企业用户进行科学数据管理的服务平台, 既能存储科学数据, 也可以在线进行学术沟通交流, 用户还可以基于该平台建设自己的数据一体化系统, 实现数据的生产、管理、传播和评估等功能。(2) “社区化”建设的典型性。Digital Commons Data 对接了全球超过 1700 个数据存储库, 来自世界各地的 600 多万名研究人员在平台上分享创意、了解领域新事物、探讨领域发展动态、寻找创新研究途径。已有文献关于 Digital Commons Data 平台的研究主要集中在对其阅读指标应用的探索以及对其数据治理流程的分析, 从科学数据管理平台的“社区化建设”角度进行研究是一个全新的视角, 将会带来全新的认识。

2.2 数据收集与分析

在资料获取上, 本文采取了访谈和文献资料收集相结合的方法。本文选取了荷兰 Elsevier 上海分公司和北京分公司的 Digital Commons Data 产品负责人、开发工程师以及上海一流大学建设高校 A 大学的科研数据管理专家共 6 人作为访谈对象, 以“Digital Commons Data 开放社区的运营建设模式与服务功能”为切入点, 采取半结构式访谈法, 访谈提纲见表 2, 同时, 对访谈中遇到的相关主题进行拓展与追问。此外, 从 Digital Commons Data 官网、Digital Commons Data 产品介绍、新闻报道以及相关研究成果中获取文献资料。通过对资料的筛选、翻译与整合, 建立能够相对全面描绘 Digital Commons Data 平台运营与服务机制的案例资料库。

在对访谈记录进行整理时, 本文采取了三级编码的方式(见表 3)。一级编码是开放式编码, 基于前期广泛收集的原始数据资料, 经过编码与整合, 提炼概括了 28 个概念化类属(A1-A28)。二级编码是主轴式编码, 在一级编码的基础上, 对形成的 28 个概念化类属进行对比和归纳, 总结出 8 个范畴化类属(B1-B8)。三级编码是选择式编码, 在二级编码的基础上, 建立各范畴之间的联系, 形成 3 个核心类属(C1-C3)。

表2 访谈提纲

访谈目的	了解 Digital Commons Data 开放社区的运营、建设模式及服务功能
访谈方式	电话访谈
访谈对象	荷兰 Elsevier 上海分公司和北京分公司的 Digital Commons Data 产品负责人、开发工程师、A 大学专家等。
访谈问题	1. Digital Commons Data 平台的建设主体是什么?
	2. Digital Commons Data 平台的建设理念是什么?
	3. Digital Commons Data 平台有哪些主要功能?
	4. Digital Commons Data 平台的资助来源有哪些?
	5. Digital Commons Data 平台的后台架构是什么?
	6. Digital Commons Data 平台有哪些独特的技术实现?
	7. Digital Commons Data 作为开放数据体系体现在哪些方面?
	8. Digital Commons Data 平台如何进行开放数据社群建设?
	9. 为实现科学数据管理平台“开放社区”建设,我国应在哪些方面做出努力?

表3 三级编码过程

三级编码 (选择式编码)	二级编码 (主轴式编码)	一级编码(开放式编码)
C1 运营机制	B1 建设主体	A1 Elsevier 自主开发; A2 基于 Digital Commons Data, 建设自己的数据一体化平台(高校、研究机构、企业或政府机构)
	B2 建设理念	A3 数据的充分利用 A4 科研数据全生命周期管理 A5 数据留存备份 A6 便捷访问 A7 安全共享
	B3 资助来源	A8 Elsevier 自身研发经费 A9 高校、科研院所、政府单位等使用者的购买费用
C2 建设模式	B4 后台架构	A10 基于云端分布式部署 A11 基于云端模块化应用 A12 各模块独立或组合使用 A13 与其他科研数据管理(RDM)工具结合 A14 使用开放式应用程序接口(API)
C3 服务功能	B5 全球科学数据库索引功能	A15 对接约 2000 个数据存储库 A16 深度索引超过 35 个存储库的 1000 多万个数据集
	B6 电子实验室记录本功能	A17 使用 DANS 进行永久的 DOI 与研究数据保存 A18 电子实验室记录本 A19 数据采集和数据重现 A20 提交一键生成的报告 A21 单个数据集存储空间超过 100GB
	B7 科研形象管理和社交网络功能	A22 数据发布同步到 OpenAIRE(欧洲最大发布中心) A23 研究数据协作和工作流工具 A24 Feed 跟踪与推荐功能 A25 数据安全保障
	B8 招聘就业与基金资助信息服务功能	A26 提供 2000 多个机构的科学、技术和卫生工作的岗位 A27 职业搜索与建议 A28 2000 多个机构的基金资助信息

3 案例描述与分析

3.1 Digital Commons Data 开放社区的运营机制

Digital Commons Data 平台的建设理念在于实现科学数据的充分利用 (A3), 帮助用户对科学数据进行全生命周期管理 (A4), 使科研人员能够随时随地、安全地访问和共享信息 (A6、A7)。另外, Digital Commons Data 旨在实现科学数据在科研群组内的共享, 通过打造公共群组, 提供相同领域研究人员相互联系的专用空间, 契合科研合作网络生态。Digital Commons Data 产品研究和开发的资金主要来自 Elsevier 公司这一市场化主体的自主投入 (A8), 高校、科研院所、政府单位等若要使用该平台, 需要进行服务购买 (A9)。

3.2 Digital Commons Data 开放社区的建设模式

Digital Commons Data 平台在技术框架上基于云端进行分布式部署 (A10) 和模块化应用 (A11), 既可以灵活组合也可以与其他科学数据管理 (RDM) 工具结合使用 (A12、A13)。平台对外提供开放式应用程序接口 (API) (A14), 可以与全球科学数据管理生态系统和其他 Elsevier 科研情报管理系统集成使用。另外, Digital Commons Data 整合了 Elsevier 投稿系统, 科研人员在使用 Digital Commons Data 投稿的同时可以提交研究过程中的数据集。

3.3 Digital Commons Data 开放社区的服务功能

3.3.1 全球科学数据库索引

Digital Commons Data 收录了全球 1700 多个公共数据存储库中约 2050 万个数据集 (A15), 每个数据集均有超过 100GB 的存储空间, 科研人员可以发布、共享、展示和存储科研数据。平台对其中 35 个存储库的 1000 多万个数据集进行了深度索引 (A16), 支持多维度检索, 并提供数据文件内容的在线预览, 不仅使得科研人员能够轻松准确地找到相关数据, 同时也显著提高了科研项目的可见度和影响力。

3.3.2 电子实验室记录本

作为一个安全的云存储库, 可以在平台长期存储科学数据, 无论用户在哪里, 数据都很容易被共享、访问和使用, 独立科研人员或团体可以在同一位置组织、批注和共享数据, 有效提升了科研人员的工作效率。同时, 先进的电子实验室记录本功能 (A18), 可以存储、采集、重现不同版本的科学数据 (A19), 以支持纵向研究, 还可向各级组织提交一键生成的数据报告 (A20)。独特的 DOI 使得用户的研究成果与引用数据集实现关联 (A17)。

3.3.3 科研形象管理和社交网络

用户科研形象的个性化管理是 Digital Commons Data 平台的主要特色之一, 通过状态查询, 用户可以查看个人研究成果的影响力、所在国家、研究领域等^[27], 还可以与平台上志同道合的科研人员进行联系, 进一步拓展自己的学术研究网络。另外, Digital Commons Data 可以设置不同用户群组, 并给予不同权限, 用户可阅读、修改组内所有文献。此外, “Feed” 功能可以跟踪用户的研究主题与社交网络, 并进行相关研究及新研究人员推荐 (A24)。

3.3.4 招聘就业与基金资助信息服务

作为一个学术网络的社交平台, 借助庞大的用户体量和科研领域的专业服务, Digital Commons

Data 面向机构和用户提供双向招聘就业服务。截至 2022 年 2 月,平台已发布 262361 个科学、技术和卫生工作的就业岗位 (A26),用户可根据地区或领域进行搜索 (A27)。同时, Digital Commons Data 能够从 2000 多个机构及时收集到基金资助信息,帮助科研人员快速匹配适合申请的基金项目 (A28)。

4 结论与启示

4.1 市场化运作模式是实现平台建设的有效途径

Digital Commons Data 平台是由 Elsevier 公司自主开发并管理的,政府、高校、科研机构或其他企业可通过购买服务的方式使用该平台。我国目前的科学数据管理平台多集中于科研院所和高等院校,往往基于自身研究需求而建,数据孤岛的现象普遍存在,开放共享的意识也还没有形成。“开放社区”的本质是推进科学数据共享,推动科研人员群体协作,借鉴 Digital Commons Data 平台的运营模式,推进以科研人员需求为导向、以企业为建设主体的科学数据管理平台建设,是实现“开放社区”理念和打造优质服务功能的重要途径。

4.2 多领域、全球化是提高平台开放性的关键所在

Digital Commons Data 平台拥有全球 600 多万名科研用户,用户数量多且覆盖领域广,全球化的数据共享基本能满足各个领域研究人员的科研需求。有学者研究指出, Digital Commons Data 的访问量与传统被引量呈正向强相关关系,说明越来越多的研究人员选择在 Digital Commons Data 上阅读和整理文献,进而进行引用^[28]。目前我国自建的科学数据管理平台多集中于某一单一领域,与全球科学数据平台的链接也比较少,因此打造多领域、全球化的科学数据管理平台,对提升用户覆盖范围、融入全球科学研究生态体系、提高平台影响力具有深远的意义。

4.3 关联应用、定制推送是平台专业服务的基本保障

Digital Commons Data 平台提供各类数据应用、分析软件工具及不同版本的应用下载和插件说明,为用户提供一站式的数据存储、分析和应用服务。同时,平台还能定制推送关注领域的发展动态,体现了数据共享的拓展形式。我国的科学数据管理平台多以存储功能为主,对于平台的可持续发展、功能模块、用户体验等方面考虑较少,在数据密集型科学及科学研究第四范式的推动下,以关联应用和智能推送为主要服务模式的平台建设,是科学数据管理平台提高用户科研效率、提升科研合作的关键。

4.4 数据安全是平台建设的中中之重

科学数据是科技创新与发展的基础性战略资源,科学数据安全的影响大至国家安全、产业安全,小至个人的合法权益。Digital Commons Data 可自建群组,将共享数据的权限限制在组内成员之间,大大保障了数据共享过程中的安全性问题 (A25)。我国科学数据管理平台也要将数据存储、传输、共享过程中的安全问题作为重中之重,在数据脱敏、权限控制、系统监控等各个方面健全制度和流程,做好软硬件保障,保护科学数据在其全生命周期中免受破坏性外力和非授权操作的侵害,保持科学数据的机密性、完整性和可用性。

本文虽然对 Digital Commons Data 这一平台进行了较为全面的调查分析,但限于篇幅,未能对其他具有代表性的科学数据管理平台进行调查和比较研究,因而所得出的经验启示不够全面。随着科学数据管理平台“开放社区”建设的蓬勃发展,未来将有更多的案例值得进一步挖掘与探索,也可以从技术和结构等更多维度上进行拓展。

【参考文献】

- [1] 顾立平. 科研模式变革中的数据管理服务: 实现开放获取、开放数据、开放科学的途径[J]. 中国图书馆学报, 2018, 44(6): 43-58.
- [2] re3data.org [EB/OL]. [2022-02-14]. <https://www.re3data.org/search?query=>.
- [3] Ross-Hellauer T, Deppe A, Schmidt B. Survey on open peer review: Attitudes and experience amongst editors, authors and reviewers[J]. Plos One, 2017, 12(12): e0189311.
- [4] 严炜炜, 张敏. 科研协同中的数据共享与利用行为模式分析[J]. 情报理论与实践, 2018, 41(1): 55-60.
- [5] Lee J.-S, Jeng W. The landscape of archived studies in a social science data infrastructure: Investigating the ICPSR metadata records[J]. Proceedings of the Association for Information Science and Technology, 2019, 56(1): 147-156.
- [6] Scot M. Les archives britanniques des sciences sociales. Deux études de cas: UK Data Archive (UKDA) et Qualidata[J]. Genèses, 2006, 63(2): 46-65.
- [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] 项英, 赖剑菲, 丁宁. 高校图书馆科学数据管理服务实践探索——以武汉大学社会科学数据管理为例[J]. 情报理论与实践, 2013, 36(12): 89-93.
- [9] 张慧, 程雨琦, 王传清. 科学数据开放共享合作关系实证研究——以北京大学开放研究数据平台为例[J]. 图书情报工作, 2021, 65(23): 22-30.
- [10] 朱玲, 聂华, 崔海媛, 等. 北京大学开放研究数据平台建设: 探索与实践[J]. 图书情报工作, 2016, 60(4): 44-51.
- [11] Steinhart G, Dietrich D, Green A.b. Establishing trust in a chain of preservation: The TRAC checklist applied to a data staging repository (DataStaR)[J]. D-Lib Magazine, 2009, 56(1): 147-156.
- [12] Wilson J. Embedding Institutional Data Curation Services in Research (EIDCSR) Project Final Report[R]. University of Oxford, Oxford OX2 6NN, February 2011.
- [13] 张计龙, 殷沈琴, 张用, 等. 社会科学数据的共享与服务——以复旦大学社会科学数据共享平台为例[J]. 大学图书馆学报, 2015, 33(1): 74-79.
- [14] Harvard dataverse [EB/OL]. [2022-02-21]. <https://dataverse.harvard.edu/>.
- [15] 罗鹏程, 朱玲, 崔海媛, 等. 基于 Dataverse 的北京大学开放研究数据平台建设[J]. 图书情报工作, 2016, 60(3): 52-58.
- [16] 张闪闪, 王铮. 海外基于开源软件的典型科研数据共享服务平台案例研究[J]. 图书情报工作, 2017, 61(6): 79-86.
- [17] 刘霞, 饶艳. 高校图书馆科学数据管理与服务初探——武汉大学图书馆案例分析[J]. 情报工作, 2013, 57(6): 33-38.
- [18] 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.
- [19] 周宇, 廖思琴, 阮莉萍, 等. 数据监护平台评价指标体系构建与测定研究[J]. 图书馆学研究, 2017(1): 35-42.
- [20] 辛一. 九省份科学数据共享平台网站建设比较研究[J]. 中国科技资源导刊, 2019, 51(3): 18-23, 84.
- [21] 马合, 黄小平. 欧美科学数据政策概览及启示[J]. 图书与情报, 2021(4): 84-91.
- [22] 中共中央关于制定国民经济和社会发展第十四个五年规划和二〇三五年远景目标的建议[EB/OL]. [2022-02-18]. http://www.gov.cn/zhengce/2020-11/03/content_5556991.htm.
- [23] Ayris P, Berthou J Y, Rapporteur R B, et al. First report and recommendations of the Commission High Level Expert Group on the European Open Science Cloud[R/OL]. [2022-02-25]. <https://op.europa.eu/en/publication-detail/-/>

publication/2ec2eced-9ac5-11e6-868c-01aa75ed71a1/language-en.

[24] Maynard M. Data-PASS partners: Plans for moving forward [EB/OL]. [2022-02-14]. <http://www.data-pass.org/presentations.jsp>.

[25] 陈秀娟, 张志强. 开放科学的驱动因素、发展优势与障碍 [J]. 图书情报工作, 2018, 62(6): 77-84.

[26] 罗伯特. 案例研究: 设计与方法 [M]. 周海涛, 译. 重庆: 重庆大学出版社, 2010.

[27] Mendeley 文献管理软件及学术社交平台简介 [J]. 国际口腔医学杂志, 2018, 45(3): 285.

[28] 于琦, 张昕瑞, 吴胜男, 等. Mendeley 与传统引用指标相关性的元分析 [J]. 情报杂志, 2020, 39(2): 191-198.

Research on “Community” Construction of Scientific Data Management Platform: Taking Digital Commons Data as an Example

Wang Qian Wang Chen

(Shanghai R&D Public Service Platform, Shanghai 200235, China)

Abstract: [**Purpose/significance**] This article aims to study the advanced practices of open scientific data platforms, and to provide reference for the construction of scientific data management platforms in China, realizing the digitization, openness and communityization of scientific research models. [**Method/process**] Based on the literature, this article analyzes the relevant research on scientific data management platforms at home and abroad, adopts the method of single case study, and summarizes the operating mechanism, construction mode and service functions of the well-known scientific data management platform “Digital Commons Data”. [**Result/conclusion**] The open community platform effectively utilizes the maximum value of scientific data through data sharing and group collaboration. Market-oriented operation with enterprises as the main body is an effective way to realize the establishment of an “community” of a scientific data management platform. Multi-domain and globalization are the keys to improving the openness of the scientific data management platform. Associated applications and customized pushes are the basic guarantee for the professional services of the scientific data management platform. Data security is the top priority in the construction of the “community” of the scientific data management platform.

Keywords: Open science; Scientific data management; Digital Commons Data; Data sharing

(本文责编: 王秀玲)