

科研数据的开放： 进展、模式与新探索

张丽丽, 黎建辉

中国科学院计算机网络信息中心, 北京 100190

摘要

明确了科研数据的概念和分类,从原则、政策、技术、配套措施等方面总结了科研数据开放的当前进展,并将科研数据开放主要概括为自上而下的开放模式、纵横联合的开放模式与自下而上的开放模式3类。进一步分析指出,科研数据开放需要在强有力的行政力量支持下,配套更多的激励措施,调动数据拥有者的参与意愿。作为推动研究数据共享的一种有效形式,研究数据出版提供了可行方案,以《中国科学数据》为例,阐释了数据期刊在推动科研数据出版中的积极实践及其战略意义。

关键词

科研数据 ; 数据开放 ; 数据共享 ; 数据出版

中图分类号 : G203

文献标识码 : A

doi: 10.11959/j.issn.2096-0271.2016063

Research data openness: development, models and new exploration

ZHANG Lili, LI Jianhui

Computer Network Information Center, Chinese Academy of Sciences, Beijing 100190, China

Abstract

The concept and clarification of research data were analyzed. Judging from the aspects of principals, policies, technologies and procedures, the current state of research data openness was described. The practices have been summarized into three main kinds with top-down model, down-top model as well as horizontal-vertical model for research data openness. Further analysis shows that, besides the government forces, we still need more incentives to increase the willingness to open research data while the research data publication provides another effective solution. Take "China Scientific Data" for example, as a scholar journal publishing scientific data papers, it may help to promote research data open in a way.

Key words

research data, data openness, data sharing, data publishing

①

<http://www.nature.com/sdata/about>

②

<http://www.bu.edu/datamanagement/background/whatisdata/>

③

<https://www.elsevier.com/about/open-science/research-data>

④

<http://ec.europa.eu/digital-agenda/en/open-data-0>

⑤

<http://www.budapestopenaccessinitiative.org/background>

⑥

<http://www.oecd.org/sti/sci-tech/38500813.pdf>

⑦

<http://www.earthobservations.org/dswg.php>

⑧

http://www.iode.org/index.php?option=com_content&view=article&id=51&Itemid=95

⑨

<http://opendata.cern.ch/about>

⑩

<http://data.worldbank.org/about/open-government-data-toolkit/knowledge-repository>

1 引言

自几千年前的经验科学到几百年的理论科学,再到最近几十年兴盛的计算科学乃至今天流行的e-Science,科研范式不断变化,科研数据生长周期从未停止,数据共享话题同样经久不衰。本文立足大数据时代的科研变革,探讨新时期科研数据开放的新趋势,以期望对推进科研数据共享有所帮助。

2 科研数据概述

“数据是用于荷载情报的物理符号,是对客观事物的数学表示”^[1],是未经整理的可被判读的数字、文字、符号、图像、声音样本等^[2]。“科学数据 (scientific data)”描述的是有科研价值的数据集^①,主要来源于假设科学中证伪实验所收集的数据和发现科学中捕获并加工而成的数据。在此基础上,科研数据是通过收集、观察、生产、分析加工并为科学研究成果所用的数据^②,它是科学、技术、医学等领域知识的构建基础^③。

科研数据来源广泛,主要产生于政府行政管理活动、公益事业运行和科学研究等场景。政府部门负责社会运行的方方面面,其中包含大量的科学研究基础数据,这些公共数据具有在新产品和服务中重用的巨大潜力,能够通过分享数据收获更多^④。此外还包括大量公共部门生产与捕获的公益性数据资源。而科学研究活动本身也是研究数据使用和再生产的主要途径之一。

3 开放科研数据进展

3.1 全面开放的公益原则

2001年11月,开放社会基金会 (Open

Society Foundations, OSF) 在匈牙利召开有关开放访问的国际研讨会,会上发表的《布达佩斯开放存取计划》^⑤提出开放存取 (open access, OA) 概念,成为“免费存取”的前奏。一系列国际组织纷纷提出或持续更新基于项目或科研领域的数据共享原则。其中,国际核苷序列联合数据库联盟 (INSDC)^[3]在《国际核酸序列数据库合作政策》中提出免费、非限制性的数据存取政策原则。《OECD公共经费支持的科研数据存取原则与指导》^⑥强调了开放、灵活、透明、依法、知识产权保护、负责、专业、互操作性、质量、安全、效率、义务、可持续性原则。地球观测组织 (GEO)^⑦提出,数据、元数据和数据产品都默认为开放数据,在最短时间内免费、无障碍地提供给注册用户使用,并在重用中予以标识。此外,联合国政府间海洋学委员会 (IOC)^⑧的《IOC海洋学数据交换政策》、欧洲核子研究组织 (CERN) 数据开放门户^⑨等也都提出了所在行业领域的数据共享原则或内容等。国际科学理事会 (ICSU)、国际科学院组织 (IAP)、发展中国家科学院 (TWAS) 以及国际社会科学理事会 (ISSC) 等联合发布的《大数据时代的数据开放国际协定》^[4]也倡导势在必行的数据开放与全球力量参与。这些不同机构组织与行业领域的科研数据开放都贯穿着“全面、开放 (full and open)”的基本原则。

3.2 政府主导的开放政策

开放数据应同时满足技术层面和法律层面的开放^⑩。政府开放数据建设,借助法律手段与技术实现,将开放科学理念推向新的发展高度。

各国及地区政府在颁布数据共享法规政策的同时,也成为开放科研数据的积极践行者,这在开放政府的专门网站中也

有体现,如美国的Data.gov、英国的Data.gov.uk、挪威的data.norge.no、澳大利亚的Data.gov.au、加拿大的Data.gc.ca、肯尼亚的opendata.go.ke、荷兰的data.overheid.nl、欧盟的publicdata.eu等^①国家和地区进行的政府数据开放实践。

来自美国的开放政府支持者提出“开放政府数据八项原则”^②,包括完整性、原始性、及时性、可获取性、机器可处理、无差别性、非专属性、免授权,为开放数据实践提供了参考指南。在我国,研究数据共享政策建设依托行业实践逐步推进,如《促进大数据发展行动纲要》《加快推进云计算与大数据标准体系建设》等。诸多公共领域实践也在日趋深入,如中华人民共和国国家统计局推行的统计数据公开共享(data.stats.gov.cn)、国家科技基础条件平台资助建立的学科数据共享平台(www.escience.gov.cn)等所推进的科研数据开放工作。

3.3 开放技术的全面深入

信息技术广泛渗透科研数据管理的全周期,从数据的生产与捕获、数据管理(加工处理保存)到数据的分享重用,涵盖诸如语义关联、元数据可扩展性与互操作、数据的结构化与关联化、数据的长期保存与智能化分析处理、数据集成平台的一揽子服务等。

整体框架设计方面,技术探索主要面向一体化、协同化数据中心共享平台建设探索,通过统一平台、网络、系统的建设来发挥数据共享的统一管理与数据中心的优势^③,实现不同数据集的访问操作。

数据的生产与捕获方面,我国近90%的重大科研基础项目实现科研数据自动化采集、70%以上实现自动录入网络传输^④,这为高通量数据的生产与服务提供了基

础,观测与加工处理仪器设备以及自动化采集、传输与处理技术皆得到长足发展。

数据分析处理方面,R语言、Rapid Miner、Mahout等主流数据分析工具的模型算法、可视化、大数据处理能力在面向大数据集分析的工作中得以发展,但仍需优秀的可视化能力与分布式的数据分析模式建设^⑤。面向服务架构(service-oriented architecture, SOA)利用Web Service技术对格式各异的数据建立统一的数据共享接口设计方案,实现对各种数据资源的集中管理与统一调度,提高数据返回的速率^⑥。回归、聚类、决策树、可视化、KNN算法、PCA、统计、随机森林、时间序列、文本挖掘等仍是数据科学的流行算法^⑦。云计算方面的主流工具包括Amazon EC2、Microsoft Azure以及OpenStack;面向大数据的分析处理方面,MapReduce、Apache Spark、Hadoop、MongoDB等提供了各具特色的框架与解决方案。

数据的发布共享与利用方面,技术研究包括利用新一代网络技术来探索数据统一发布和数据共享框架建设,同时在数据安全与数据的权限控制等方面提供切实可行的技术解决方案^⑧。依托研究数据与科技文献交叉引用^⑨,利用DOI关联模型进行数据深层价值挖掘与利用情况的追踪,为数据价值体现提供计量基础。此外,相关技术还包括大数据处理环境下关于海量数据长期保存与高效访问等内容。

此外,基于本体^⑩的逻辑构建与语义检索、关联数据技术(如RDF语义链接功能实现^⑪等智能化技术实践)也取得了长足发展。面向大数据时代,用于解决大数据异构化加工处理、不确定性分析、复杂系统建模、大数据价值挖掘等方面的技术探索也为数据密集型科研带来了新的解决方案。

① <http://zh.wikipedia.org/zh/%E9%96%8B%E6%94%BE%E8%B3%87%E6%96%99>

② <http://opengovdata.org/>

③ <http://www.kdnuggets.com/2016/09/poll-algorithms-used-data-scientists.html>

3.4 开放数据的配套举措

为补偿数据共享供给者的劳动付出、维护数据共享链条的持续生长,适当的数据发布时机与发布内容规范、数据获取授权机制的有效补充以及数据引用溯源等方面的激励与监督措施,成为推动开放共享理念与实践互联的积极策略。

(1) 设置保护期限

为了更好地鼓励科研数据的共享,一些数据共享计划涵盖了延迟科研数据发布时限的要求。以期刊为例,提交《自然》^⑭的论文不得早于成果发表一周向媒体发布。中国科学院、国家自然科学基金委员会声明^⑮,其资助项目产生的论文需提交指定机构或知识库,并在发表后12个月内开放获取。施普林格(Springer)将先前12个月数据滞后发布的使用范围由针对有相应管理要求的机构资助出版扩展为全部。Emerald将先前允许机构存储库存取数据并及时发布的政策修改为根据机构要求执行24个月的发布时限,但作者自愿存储行为则不存在发布时间限制,即“自愿分享,义务除外”模式^⑯。

(2) 内容分级控制

为解决信息发布的控制性问题,开放存取运动提出了“金色道路(golden route)”和“绿色道路(green route)”^⑰,以促进有限制的科研数据共享。

涉密数据的开放则需要相关法律与行业准则规范。例如,我国《气象资料共享系统建设气象数据分级与分类(草案)》规定,气象数据对外服务分为绝密、机密、保密、内部和公开5级,以规范数据开放工作。

科研数据的共享方式还受到科研数据组织形式与价值的影响。仍以气象数据为例,《气象资料共享管理办法》规定,气象科研数据的提供利用,按用户对象和级别分为无偿、有偿两类。数据的精细化加工

意味着更多的成本投入。为此,面向客户需求所进行的数据深度分析与加工仍采取一定的有偿机制来推动数据交流。

当科研数据仍处于政策法规合理约束的未解禁期或者知识产权、保密与隐私权保护下的未公开状态,仍然可以建立资源授权的补充机制来满足“灰色地带”的数据开放需求。授权机制包括两方面:一是对数据拥有者占有权力的认可授权;二是对数据使用者申请使用行为的认可授权。例如免费开源软件授权(free and open source software license)^⑱、知识共享(creative common)^⑲和开放数据共用(open data common)^⑳等。英国的“皇家版权(Crown Copyright)”^㉑、美国国家生物信息技术中心(NCBI)的dbGaP存储库^㉒等都进行了相关探索实践。

(3) 引用、关联与溯源

从数据引用的表现形式来分,科研数据的引用可分为传统文献引用、独立数据引用和混合引用3种模式。和传统科学学术认可方式类似,科研数据的使用也需要通过标识、引用的方式来跟踪数据集的利用情况,并通过计量分析与大众评议等渠道实现对发布数据的真实性、准确性和著作权等方面的重复校验。数据引用探讨包括微观层面的技术标准规范、工具方法以及宏观方面的数据引用政策、指南、数据引用声明等。我国的冰川(雪冰)冻土学科中心、寒区旱区科学数据中心、中国科学院计算机网络信息中心等也提出了数据引用规范指南。

与此同时,关联数据通过网络标识符认定事实、概念、人、地方、现象以及拥有共同特性的文件,从而使人发现数据的关联,并借此增加数据集的应用价值^㉓。建立在关联数据基础上的溯源模型,帮助数据使用者更好地校验数据质量,推动数据共享。

^⑭ <http://www.nature.com/nature/authors/policy/embargo.html>

^⑮ http://www.ercim.eu/publication/Ercim_News/enw64/jeffery.html

^⑯ <http://creativecommons.org/licenses/>

^⑰ <http://opendatacommons.org/>

^⑱ <http://www.nationalarchives.gov.uk/information-management/re-using-public-sector-information/copyright/crown-copyright/>

^⑲ <http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/about.html>

^⑳ <http://www.nsf.gov/eng/general/dmp.jsp>

4 开放科研数据的典型模式

参考科研数据的来源与开放实践所选取的路径,大致可分为自上而下的开放模式、横纵联合的开放模式与自下而上的开放模式3种情况。

4.1 典型的开放模式归纳

在自上而下的开放模式中,科研数据开放的政策与经费发起者(多为政府)成为推动科研数据开放的原动力。这一类数据的开放共享以不断积累的数据为基础,相应提升数据开放体量,并通过数据生产与获取能力的不断提高扩展数据共享服务的能力和水平。对这些数据而言,开放共享并非道德层面的高尚行为,而是所有权性质的必然要求。例如,美国国家科学基金会(NSF)积极推进新的数据政策,要求所有项目申请需同时提交一份“数据管理计划”,对项目产生数据的存取归档与共享等方面做出明确规定。英国自然环境研究理事会(NERC)①数据政策规定项目期间产生的有长期利用价值的数据需提交NERC数据中心,以保障其价值的持续性发挥等。

横纵联合的共享模式多以存在隶属或管辖权利的部门内部以及跨部门之间的多元化合作为主,尤以数据共享的国际、区域与学科交叉方向的联盟组织的科研数据开放较为常见,如地学领域的地球系统科学联盟(ESSP)、国际地球观测组织、灾害风险综合研究计划(IRDR),生物学领域的物种2000(Species 2000)、全球生物多样性信息网络(GBIF)以及著名的人类基因组计划(HGP)、国际虚拟天文台联盟(IVOA)等展开的科研数据

开放活动。参与联盟组织的会员按照统一的章程,共同维护数据资源的积累和服务能力的提升。这种看似松散的组织恰恰成为当前国际范围内科研数据积累与推动开放的中坚力量,甚至在某些领域也形成垄断局面。

自下而上的开放模式多立足于一定的激励措施基础上,推动数据的拥有者根据一定的规则主动地参与数据开放。这一开放模式的组织形式更为松散、开放,同时也成为推动科研数据开放共享不可忽视的重要组成部分。例如,在期刊方面,越来越多的出版机构要求研究论文的发表包括数据方面的配套开放。其中,《科学》等杂志对来稿涉及的数据提出管理要求。许多机构存储库也纷纷提出数据开放资源计划与实践,如PANGAEA(德国)、arXiv(美国)、RePEc(英国)、PubMed Central(美国)、Figshare(英国)等在数据资源的持续积累与开放共享方面提供了多样化的支持^[16]。

4.2 开放中的焦点问题

利益机制研究是解决共享难题的焦点。首先引入相关利益者的视角,无论哪种数据开放模式,政府部门、公共服务机构、科研人员与社会公众都是重要的参与者(见表1),他们在科研数据开放活动中肩负不同责任,利益焦点各有侧重。

其中,政府部门作为数据监管者,在负责立法与实施的同时,还需配套合理的激励机制,以促进数据的开放重用。而激励措施的建立,既要满足数据生产者的利益需要,又应该维护数据使用者的合法权益,并通过公共服务机构助推科研数据开放的深度和广度。

从利益相关者的视角分析中不难看出

① <http://www.nerc.ac.uk/research/sites/data/>

表1 科研数据开放中的相关利益者

相关利益者	科研数据开放职责	利益焦点
政府	数据监督管理者	合法合规
公共服务机构	图书馆	满足服务需要
出版社	数据服务	学术出版
数据中心	数据服务	满足服务需要
大学	数据教育、数据服务	提升数据教育水平
企业	数据服务	数据价值增值
联盟	制定原则、推广实践	数据服务能力提升
科研人员	数据生产者 and 使用者	数据工作被认可 数据易于获取使用
社会公众	数据使用者	数据利用与受益

出，科研数据的共享不意味着“杀鸡取卵”，数据共享工作的展开应以认可数据生产者与发布者的科研价值为第一要务。将相关利益者置于开放共享模式之中，可以继续识别出3种主要开放模式的关键领导力。其中，自上而下的科研数据开放模式中，政府或组织的行政命令成为推进数据共享的核心力量；横纵联合的共享模式则是组织章程与自愿分享两股力量共同推进的杰作；自下而上的共享模式中，数据开放

的动力多源自于数据拥有者本身的自愿行为。Boulton G等^[17]描述的“以科研数据开放为核心的开放科学框架”（如图1所示）中，也可以看到类似的情况。在知识的联合生产过程中，行政管理数据开放需要政府部门强有力的制度框架约束。这类科研数据的开放采取自上而下的道路，开放的程度和范围密切地与政府法规政策框架相适应。而对于后两类科研数据共享而言，数据开放更多地采取和自下而上相结合的手

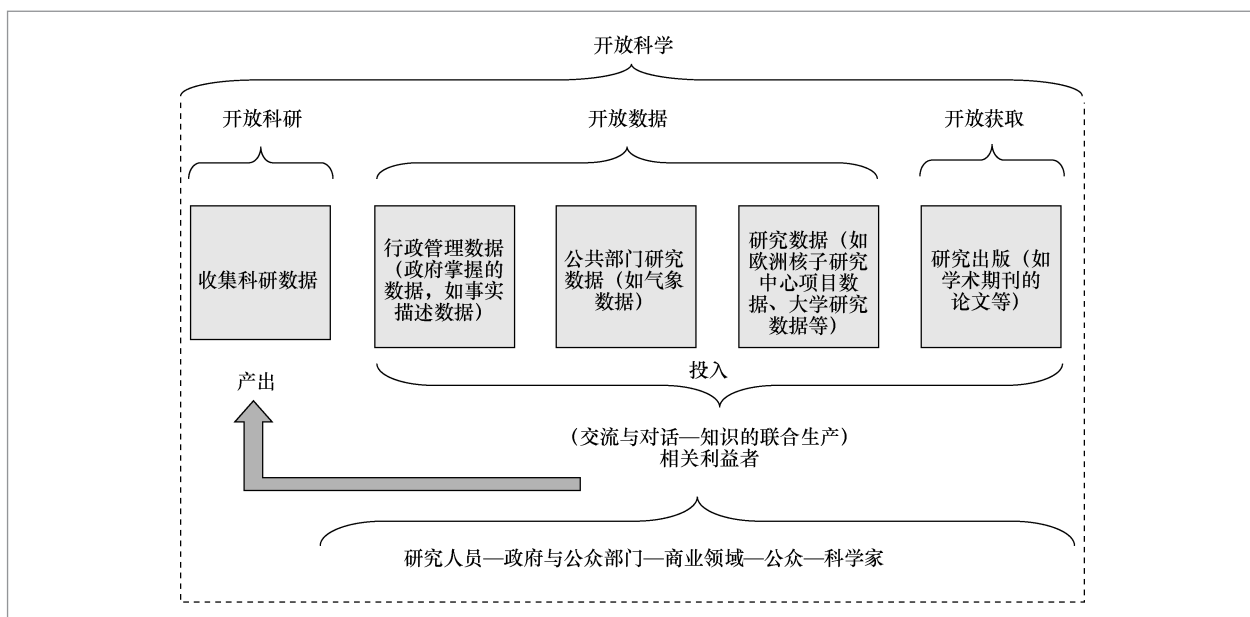


图1 以科研数据开放为核心的开放科学框架

段来推进。由此可见,针对公共部门科研数据以及科学科研数据的开放共享,为了更好地适应目前主流的3种科研数据共享模式运行,还需要合理配套激励机制,为此需要进行更多探索。

5 开放科研数据的新探索

在推动科研数据开放的探索中,学术出版机构进行的以数据为核心的出版活动提供了一种值得关注的解决方案。

5.1 关于科研数据出版

作为开放科学运动(open science movement)的核心部分,数据出版的主要目标是提升数据在科研成果中的首要性地位,包括独立数据出版(standalone data publication)、学术论文辅助数据出版(data publication by proxy)、附录数据出版(appendix data)、期刊数据档案出版(journal driven data archival)以及数据论文出版(overlay publication)等类别^[18]。其中,尤以易于数据获取与使用的独立数据出版为近年来发展热点。

在研究数据出版中,科研人员与数据工作者按照规范的质量管理和控制流程,以数据论文的方式,通过互联网公开发布其通过观察、实验、计算分析等科研过程所产生的原始数据,或通过已有数据进行系统化地收集、整理和再加工后形成数据产品,供其他使用者便捷地发现、获取、理解和再分析利用,且可在科研论文及相关科研成果中引用。

此外,研究数据出版的迅速发展,为多元化的科研数据存储提供了可能,进而为打破现有科研领域的数据资源垄断提供了新的契机,同时对我国研究数据的出版事

业具有积极的战略性意义。

5.2 《中国科学数据》的出版实践

《中国科学数据》于2015年创刊,由中国科学院主管,中国科学院计算机网络信息中心主办,是国家网络连续型出版物的首批试点刊物。作为目前中国唯一的专门面向多学科领域研究数据出版的学术期刊,该刊致力于研究数据的开放、共享和引用,推进研究数据的长期保存与数据资产管理,探索研究数据工作的有效评价机制,推动数据科学的发展。

《中国科学数据》收录的研究数据论文,结合传统期刊论文内容和结构化描述模式,是对具有科学价值的某类或某个数据集进行规范化描述所形成的科学研究论文,能够使数据更具发现性、引用性、解释性和重用性。该刊收录的数据论文需提供所发表数据集的描述细节,包括数据收集和加工处理方法、数据质量评估和验证的方法、便于理解和使用数据的相关信息等,但不包含新的科学假设。完整的数据论文出版应包括数据论文和对应数据集两部分,二者通过唯一标识符实现一致性关联,经同行专家评议保障数据的高质量与可读性。《中国科学数据》现已正式出版两期内容(2016年6月和9月),所发表的数据论文内容涵盖地球科学、生物科学、化学科学、空间天文与信息科学等方面的数据内容。

为更好地了解数据出版进展情况,该刊相关人员同时组织调查^②,反馈显示,数据出版作为推动科研数据共享的有效途径之一,获得越来越多的支持,配套工具发展初具规模,但在学术地位和行业认可度等方面有待提升,出版运营模式尚不成熟,引用评价机制等方面亟待发展。

② <http://www.csdata.org/paperView?id=9>

6 结束语

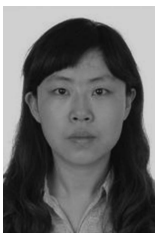
综上,笔者针对科研数据开放实践的现状与实际内容进行了梳理,从“完全开放”的公益原则、政府为主所推进的开放政策演进、开放技术的支撑与配套数据开放的诸多举措中,不难看到,典型的自上而下的开放模式、横纵联合的共享与自下而上的三大类共享模式中,数据开放的原动力需要更多地调动数据拥有者的积极参与意愿。为此,有必要在科研数据开放现有成果的基础上,深入探讨有关激励措施方面的内容。保护期限、分级控制、引用关联与溯源为此提供了一些参考,研究数据出版也提供了另一种思路。希望上述分析,尤其是研究数据出版领域的相关推介,能够为更好地推动科研数据开放共享提供一定的指导。

参考文献:

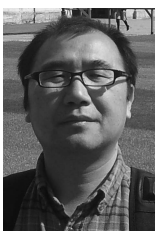
- [1] 严怡民. 情报学概论(修订版)[M]. 武汉: 武汉大学出版社, 1994.
YAN Y M. Introduction to information science (revised edition) [M]. Wuhan: Wuhan University Press, 1994.
- [2] 于良芝. 图书馆学导论[M]. 北京: 科学出版社, 2003.
YU L Z. An introduction to library science[M]. Beijing: Science Press, 2003.
- [3] BRUNAK S, DANCHIN A, HATTORI M, et al. Nucleotide sequence database policies[J]. Science, 2002, 298 (5597): 1333.
- [4] ICSU, ISSC, TWAS, et al. Open data in a big data world[R]. [S.l.:s.n.], 2015.
- [5] 罗朝宇, 衣丰超, 袁兴刚, 等. 基于一体化平台数据中心的数据共享平台在电力企业的应用[C]//2013电力行业信息化年会论文集, 11月30日—12月1日, 2013, 北京, 中国. 北京: 人民邮电出版社, 2013: 5.
LUO Z Y, YI F C, YUAN X G, et al. Application in electronic enterprise for data sharing platform based on integrated data centers[C]//The Annual Conference on Electric Power Industry Informationization, November 30–December 1, 2013, Beijing, China. Beijing: Posts and Telecom Press, 2013: 5.
- [6] 国家科技基础条件平台中心. 国家科学数据资源发展报告[R]. [S.l.:s.n.], 2015.
National Science & Technology Infrastructure Center. Annual report for national scientific data resources development[R]. [S.l.:s.n.], 2015.
- [7] 杨京, 王效岳, 白如江, 等. 大数据背景下数据科学分析工具现状及发展趋势[J]. 情报理论与实践, 2015, 3(38): 134–137, 144.
YANG J, WANG X Y, BAI R J, et al. Big data analysis tools in data science: current current status quo and development[J]. Information Studies: Theory & Application, 2015, 3(38): 134–137, 144.
- [8] 孙周军, 何惠仪, 徐美红, 等. 基于SOA架构建立气象科学数据共享系统方法研究[J]. 电脑知识与技术, 2010(8): 1888–1891.
SUN Z J, HE H Y, XU M H, et al. The research of the method of constructing meteorological scientific data sharing system based on SOA[J]. Computer Knowledge and Technology, 2010(8): 1888–1891.
- [9] 陆冬云, 张和珍, 何险峰, 等. 科学数据库建设框架——统一发布及数据共享方案[J]. 计算机与应用化学, 2004(1): 103–106.
LU D Y, ZHANG H Z, HE X F, et al. Framework for scientific data construction—uniform solution for distribution and data sharing [J]. Computers and Applied Chemistry, 2004(1): 103–106.
- [10] 郭学武. 基于引文的科学数据与科技文献关联研究[J]. 情报科学, 2014, 32(4): 59–62, 125.
GUO X W. Link study of scientific data and scientific literature based on citation analysis[J]. Information Science, 2014, 32(4): 59–62, 125.

- [11] 张乃静, 鞠洪波, 纪平. 本体构建理论在林业科学数据共享中的应用研究[J]. 西北林学院学报, 2013(6): 151-156.
ZHANG N J, JU H B, JI P. Application of ontology construction theory in forestry data sharing[J]. Journal of Northwest Forestry University, 2013(6): 151-156.
- [12] 司莉, 李鑫. 基于关联数据的科学数据集成与共享研究——以Bio2RDF项目为例[J]. 图书馆学研究, 2014(21): 51-55.
SI L, LI X. Linked data oriented scientific data integration and data sharing—case study of Bio2RDF project [J]. Researches in Library Science, 2014(21): 51-55.
- [13] 中国科学院、国家自然科学基金委员会. 中国科学院、国家自然科学基金委员会发布关于实施开放获取政策的声明[J]. 图书情报工作, 2014, 58(11): 96.
CAS, NSFC. Announcement for open access policy by CAS & NSFC[J]. Library and Information Service, 2014, 58(11): 96.
- [14] SUTTON S. Open access, publisher embargoes, and the voluntary nature of scholarship [J]. C&RL News, 2013(10): 468-472.
- [15] MORANDO F. Legal interoperability: making open (government) data business and communities [J]. JLIIS.it, 2013, 4(1): 441-452.
- [16] 黄永文, 张建勇, 谢靖, 等. 开放资源的再利用模式研究[J]. 图书情报工作, 2013, 57(21): 32-37.
HUANG Y W, ZHANG J Y, XIE J, et al. Research on recycling models of open resources[J]. Library and Information Service, 2013, 57(21): 32-37.
- [17] BOULTON G. Open data in a big data world—the science international accord[R]. [S.l.:s.n.], 2016.
- [18] LAWRENCE B, JONES C, MATTHEWS B, et al. Citation and peer review of data: moving towards formal data publication[J]. The International Journal of Digital Curation, 2011, 2(6): 4-37.

作者简介



张丽丽 (1984-), 女, 中国科学院计算机网络中心助理研究员, 《中国科学数据》期刊编辑部副主任, 主要研究方向为研究数据治理、信息经济学。



黎建辉 (1973-), 男, 博士, 中国科学院计算机网络信息中心研究员、博士生导师, 大数据技术与应用发展部主任, CODATA中国委员会秘书长, 主要研究方向为大数据管理、大数据分析与管理。

收稿日期: 2016-09-23

基金项目: 国家自然科学基金资助项目 (No.91546125)

Foundation Item: The National Natural Science Foundation of China(No.91546125)