

国内外科学数据的组织与管理研究进展*

李慧佳 马建玲 王楠 王思丽 张秀秀

(中国科学院资源环境科学信息中心, 甘肃兰州, 730000)

摘要: 通过对现有研究文献的主题统计分析, 发现国内外科学数据组织和管理的研究热点分布在科学数据的获取、共享、利用和评价四个方面。文章以这四个方面为主线对国内外科学数据的组织和管理研究进行了梳理和分析, 提出了我国科学数据在政策保障机制、标准制定、数据服务等建设方面的发展建议。

关键词: 科学数据 开放获取 数据共享 数据引用 数据评价

Reviews on Scientific Data Organization and Management at Home and Abroad

Li Huijia, Ma jianling, Wang Nan, Wang Sili, Zhang Xiuxiu

(The Scientific Information Center for Resources and Environment, Chinese Academy of Sciences, Gan Su, Lan Zhou, 730000)

Abstract: Through the statistical analysis of scientific data articles, we found the topics and hot spots of scientific data organization and management is: scientific data access, sharing, utilization and evaluation. Article teases and analyzes scientific data organization and management, which make those four aspects as the main line. Finally, we make further study and development of scientific data, including the mechanism of data guarantee, the standards of data, the services of data etc.

Key words: scientific data; open access; data sharing; data citation; data evaluation

随着以数据密集型科学发现为主要科学研究特征的科学研究新范式的兴起与发展, 科学数据成为近年来国内外相关领域科学家的研究课题之一, 且随着研究的深入发展与大量研究文献的产生, 科学数据的理论体系已初步形成。同时, 美国国家航空航天局(National Aeronautics and Space Administration, NASA)、Digital Curation Centre (DCC) 等机构在科学数据的共享、引用、质量控制等方面的成功实践, 在为科学数据的实践应用提供借鉴的同时, 也进一步完善了科学数据的理论体系。本文在对科学数据的产生、内涵等基本概念及国内外科学数据研究的主题分布概述基础上, 重点从科学数据的组织与管理角度, 对科学数据的开放获取政策、科学数据的共享政策及应用实践、科学数据的引用规范与政策、科学数据的质量控制和评价四个方面的研究进展进行了归纳, 并对我国科学数据的政策保障机制、标准制定、数据服务等建设方面提出了发展建议。

1 科学数据的内涵

* 本文系中国科学院文献情报能力专项项目“开放知识资源登记系统”研究成果之一。

数据库技术、计算机技术、网络通讯技术等现代信息技术的快速发展,使科学研究的管理体系发生了变化,呈现出了一些新的特征,如:数据的采集、管理和使用一体化向相对分离方向转移,数据的传播、共享日益受到重视;数据的管理不只是保存与保管,增值加工、关联发布与数据服务收到更多的关注;数据的价值和使用范围进一步被挖掘和延伸,使其可以服务于社会的更多目标和更大范围;数据的传播、共享与应用进一步向标准化、数字化和语义化方向转变等等。所有这些,使得以数据开放、传播与共享为核心的科学数据管理应运而生,世界数据中心(World Data Center,WDC)、国际科学数据委员会(Committee on Data for Science and Technology,CODATA)的成立以及科学数据作为一门独立学科的诞生,也都标志着科学数据管理已成为了现代科学研究管理的主要组成部分。

科学数据的概念最早产生于“国际地球物理年”(International Geophysical Year, 1957年至1958年)期间的WDC建立初期^[1]。但目前最受业界认可的科学数据定义是世界经济合作与发展组织(Organization for Economic Co-operation and Development, OECD)在《OECD关于公共资助科学数据获取的原则和方针》中对科学数据的定义:科学数据是来源于科学研究的事实记录,如实验数值、图像等,并被科学团体或科学研究者所共同认为对研究结果有用的数据^[2]。值得注意的是,OECD认为科学数据不包括实验样本、实验室笔记、科学论文草稿、科学研究计划、同行评议以及同行交流等^[2]。我国科技部发布的《科学数据共享工程技术标准研究报告》则将科学数据定义为“科学数据是在科技活动时或通过其它方式所获得的能反映客观世界本质、特征、变化规律的原始基本数据,以及根据不同科技活动需要,进行系统加工整理的各类数据集”^[3]。综上,我们可将科学数据的含义定义为:科学数据是指来源于科学研究的非文本的事实记录,包括地图集、基因组、化学化合物、数学公式、医学数据和临床试验等非文本材料。

2 国内外科学数据研究的地理分布与主题分布

在开放获取环境下,国内外对传统文献资源的研究日趋完善。因此,更多的研究机构和学者将研究目标锁定在了科学数据上。但由于对科学数据的科学研究是以不同学科领域的科学数据来支撑科学问题的解决和科学结论的论证,因而对科学数据的研究分散于不同的学科领域。通过对国内外有关“科学数据”期刊文献的地理位置、主题进行聚类和分析,可以从宏观上掌握目前科学数据研究的发展现状和前沿问题。

本文以Web of Science(WOS)数据库中的期刊文献为数据来源,以“主题='scientific data'or'research data'”、“时间=2000年至2013年”为检索式对WOS进行检索,并对检索结果进行去重、清洗和整理,最终得到2854篇来自WOS的期刊文献。统计发现,2854篇文献的第一作者分布在了全球的80个国家和地区之中,美国、英国和中国大陆占据了发文数量排名的前三名(论文发表数量排名前10的国家列表见表1)。

表 1 论文发表数量排名前 10 的国家列表

国家	论文数量
USA (美国)	875
England (英国)	243
Peoples R China (中国大陆)	153
Turkey (土耳其)	137
Australia (澳大利亚)	116
Canada (加拿大)	111
Germany (德国)	101
Netherlands (荷兰)	66
Taiwan (中国台湾)	65

为了掌握国内外科学数据研究的现状和热点，本文利用 Thomson Data Analyzer 对 2854 篇关于科学数据的文献进行了关键词清洗聚类 and 词频分析，发现目前国际上对科学数据的研究主题主要有：对科学数据进行挖掘、集成、存储的方法研究；对科学数据存储、登记系统的数据库建设和应用研究；对科学数据的知识组织与管理研究；对科学数据共享的技术、政策和实践方面的研究；对科学数据元数据描述、来源遴选等方面的质量控制研究；对科学数据的开放存取政策、方法等方面的研究（科学数据的主要研究主题分布列表见表 2）。

表 2 科学数据的主要研究主题分布列表

研究主题	研究内容
Methodology (方法论)	数据挖掘 (data mining)、数据采集 (data collection)、数据管理 (Data management)、数据分析 (Data analysis)
Database (数据库)	登记系统 (registry system)、数据仓储 (data curation)、数据图书馆 (data library)、数据仓库 (data warehouse)、数据存储 (data repository)
Knowledge management (知识管理)	关联数据 (linked data)、本体 (ontology)、语义网 (Semantic Web)、知识获取 (Knowledge acquisition)、知识共享 (knowledge sharing)、知识转化 (knowledge transfer)
Data sharing (数据共享)	科学数据共享的技术、实践应用和政策研究
Data Quality (数据质量)	包括科学数据来源遴选、评价指标、评价方法及科学数据元数据描述标准等
Open data (开放数据)	包括科学数据开放政策、开放存取方法以及开放存取的实践研究等内容

从对 WOS 获取的文献进行分析还可看出我国近年来对科学数据研究持有巨大的热情，在 2000 年至 2013 年期间发表收录在 WOS 中的以“科学数据”为主题的文献数量达到了总量的 5%，高居第三位。对我国科学数据研究文献进行主题统计分析，发现我国的科学数据研究具有以下特征：①科学数据共享是我国学者最为关注的研究主题；②科学数据组织和管理各阶段的前沿方法、技术得到了来自不同领域学者的关注；③已经产生了较多的有关科学数据存储数据库方面的实践应用；④科学数据的开放存取和科学数据的知识化将成为今后国

内学者的研究重点；⑤科学数据遴选和质量控制方面的研究较为薄弱，亟待加强。

3 科学数据的开放获取政策

3.1 科学数据的开放获取发展历程

2003 年德国马普学会在柏林会议上通过的《柏林宣言》指出开放获取的内容既包括原始的科学研究成果，也包括原始的科学数据。在此之后的十年时间里，科学数据开放获取运动在不断的推进和发展。OECD 的 34 个成员国分别在 2004 年和 2006 年签署发布了《开放获取公共资助研究数据的宣言》和《开放获取公共资助研究数据的原则和指南》两份科学数据开放获取政策报告，并在第二份报告中对开放数据的范围和定义首次进行了明确界定^[3]。2010 年，潘顿协议启动了“开放数据协议”（ODate），其中要求科学数据供应商必须遵守协议规定开放其所拥有的科学数据^[4]。2011 年，LinkedScience.org 推出了用以实现开放共享和互联数据集、科研方法、工具以及科学词汇表的关联开放方法^[5]。

3.2 科学数据的开放获取政策发展

目前，各国并不是对所有类型的科学数据都实现开放获取，如，美国政府对科学数据的管理分为三个层次：①对于有可能危害到国家安全、影响到政府政务、涉及到个人隐私的数据纳入到有十分严格、明确规定的保密性运行机制中进行管理；②对于公有的数据，如政府资助产生的数据纳入到“完全与开放”的共享管理机制之中；③对于私有的数据，如私营公司投资产生的数据纳入到“平等竞争”的市场化共享管理机制之中^[6]。

为了形成统一的科学数据开放获取政策体系，《欧洲科学数据开放获取政策研究》(RECODE)项目致力于联合对科学数据开放获取研究感兴趣的欧洲利益相关者，共同努力解决科学数据开放获取的技术、基础设施、伦理和法律、国家体制等问题。同时，提出支持科学数据开放获取的政策框架，该框架积极吸收和采纳涵盖广泛的各种建议。通过 RECOD 项目建立科学数据开放获取参与机制，进行科学数据开放获取相关内容的研究和交流，重新编制针对不同利益相关者和决策者的科学数据开放获取政策^[7]。

2013 年 7 月，欧盟委员会邀请研究人员、相关行业、基金委以及专业数据中心、出版商和图书馆共同参与有关科学数据开放获取研究的相关讨论。这些利益相关者共同修订了欧盟委员会制定的有关科学数据开放获取的政策，从而为欧盟研究计划 2020 展望奠定基础，讨论会主要围绕五个方面的问题展开：①如何定义科学数据的类型，以及何种类型的科学数据应该开放获取；②如何进行科学数据的开放性限制；③如何解决科学数据的再利用问题；④科学数据的存储和获取问题；⑤如何提高“数据意识”和“文化共享”^[8]。

4 科学数据共享政策及应用实践

4.1 科学数据的共享政策

科学数据共享的研究最早起源于欧美等发达国家和地区，这些国家和地区目前也已经制定、发布了许多有关科学数据共享利用的规范和政策。如，英国的 PubMed Central 倡议学

者在发表论文时，还需向期刊社提供支撑论文正文的数据及附加材料^{[9][10]}；美国的 Inter-university Consortium for Political and Social Research (ICPSR)要求只有交纳了年费的机构会员用户，如政府部门、高校、研究机构、数据中心、非营利性组织等，才能免费获取数据，而非机构会员或没有缴纳年费的会员用户需交纳一定的费用才能获取其数据^[11]；美国国立卫生研究院(National Institutes of Health, NIH)规定：①只共享那些被科学界普遍认可、对研究结果有重要意义的最终数据；②为了避免重复，需保证共享数据的唯一性；③为保证数据的时效性，所公开、共享的数据要早于基于最终数据的研究成果刊出时间；④作者可以通过网站（包括个人网站和机构网站）的数据发布实现数据共享，也可以通过将数据存储到提供获取、控制、存档以及传播的数据库(Data Archive) 中进行共享。对那些可能危及到国家安全、涉及个人隐私的敏感数据，还可以通过有资深研究者利用和分析的受控安全环境 Data Enclave 中经对受限数据(Restricted Data)的分析后实现共享^{[12][13]}。

4.2 科学数据在国内外的共享应用实践

4.2.1 美国

美国的科学数据共享最佳实践代表机构是 NASA。NASA 下设的空间科学数据运行办公室是负责航天局科学数据的机构，且由于数据主要来源于航天局的空间飞行计划，因而数据集集中于天文和空间科学领域，主要负责数据永久存档以及提供天体物理学、空间物理学数据等工作。为了实现数据的多种服务与应用，科学数据运行办公室还设置了空间物理学数据运行中心(SPDF)，负责多学科、多任务的科学数据服务的开发、设计和实现^[14]。此外，美国的专业图书馆凭借着其在专业领域的服务与资源建设优势，进而成为了国家科学数据共享体系中的主要组成部分，如美国国家医学图书馆(The United States National Library of Medicine, NLM)凭借着其在创立、维护生物信息学数据库方面的丰富经验，承担了一个关于计算分子生物学的研究计划。在该计划中，作为 NIH 在 NLM 的一个分支的国家生物技术信息中心(National Center of Biotechnology Information, NCBI)主要通过自己开发的一系列工具和软件，如基因序列注册软件 BankIt、数据搜索软件 Entrez、基因序列比对分析软件 BLAST 等来实现数据资源的网络共享^[15]。

4.2.2 日本

日本科学数据共享的实践代表是科研数据公开数据库。该数据库拥有 70 个主题数据库，数据来源于日本规模最大的科研院所——日本产业技术综合研究所的各机构的科研项目，日本产业技术综合研究所的工作和研究领域主要分布在能源与测量技术、电子学与纳米技术、生命科学与 IT 技术、材料科学、制造业与环境科学、地球科学等领域，数据由各机构自己整理、共享到该数据库后通过网络向科研机构、工业企业提供免费服务^[16]。

4.2.3 中国

我国各类图书馆、科技信息研究机构也在尝试搭建科学数据共享服务平台。如, 2001 年底,为促进气象数据的社会共享, 科技部启动了第一个科学数据共享试点项目“气象科学数据共享试点”^[17]; 中国西部环境与生态科学数据中心和中科院国家科学图书馆兰州分馆合作建立的文献、数据资源知识服务平台“中国(西部)环境与生态科学知识积累平台”^[18]; 中国科学院地理科学与资源研究所的研究团队为了有效的整合地学领域科学数据资料而构建了“地学科学数据共享平台”^[19], 等等。

4.2.4 国际合作

有关科学数据共享的应用实践除了有各国独立研究外, 还有部分跨地区、跨学科的科研机构和国际合作项目对这一问题进行了探索。如, CODATA 作为国际科学联合会 (ICSU) 下的一个跨学科的科学委员会, 致力于提高对整个科技领域中有重要变化的数据质量、可靠性、管理与可访问性。CODATA 通过关注各学科领域之间的科学数据的质量、获取和传播, 其目的在于为科学家提供科学数据共享的途径, 从而推动科学活动的国际合作以及加强通过科学数据创造新知识的能力^[20]; 美国的《全球科学信息共有先导》(Global Information Commons for Science Initiative, GICSI) 计划主要是激励人们尝试新模式的创造、传播以及合作利用科学数据信息, 并为计划的主要参与成员提供科学数据的全球化共享平台^[21]。英国的“促进发展中国家科学数据共享与应用全球联盟”(e-SDDC) 计划侧重于建立不同领域科学数据共享的沟通机制, 进而促进发展中国家科学数据的共享和利用。该计划的目的在于帮助发展中国家的科学数据共享平台建设, 从而缩小与发达国家之间的数字差距^[22]。

5 科学数据的引用规范与政策

虽然科学数据在过去经常被共享, 却很少像期刊文献引用一样进行标准化的引用, 如果科学数据集被引用, 它们将会在与之相关联的学术交流的有效周期内创造出学术价值和重要意义。随着有关科学数据的数据库、仓储中心以及知识库的建设越来越多, 科研人员、科学数据存储服务中心不得不思考科学数据的规范化引用问题。一些科学数据数据库、知识库的有关科学数据引用方面的标准规范和政策也陆续公布。2009 年, 德国国家科学技术图书馆 (German National Library of Science and Technology, TIB)、大英图书馆 (British Library)、法国科学技术信息研究所 (French Institute for Scientific and Technical Information, INIST)、丹麦技术信息中心 (Technical Information Center of Denmark)、加拿大科学技术信息研究所 (Canada Institute for Scientific and Technical Information, CISTI)、澳大利亚国家数据服务中心 (Australian National Data Service, ANDS) 等机构联合签署“提升网络科学数据开放获取”的学术合作备忘录, 以建立一个用于登记科学数据, 并为其分配永久标识符, 从而使科学数据集可以作为独立的、可引用的、唯一的科学对象被科学家使用的非盈利代理服务——DataCite^[23]; 澳大利亚国家数据服务中心 (ANDS) 作为 DataCite 的参与者, 认为科学

数据引用应该和研究人员发表文章时提供参考文献一样有规范化标准,因而对其所有的科学数据引用进行了规范,并强调“国家发展战略的一个重要目的是为了让更多的研究者能够重复利用科学数据,为了达到这个目的,ANDS 提供科学数据的重要性和数据引用的标准程序,以便科学数据更好的被共享利用”^[24]。

现阶段,尽管大量的实验数据和观测数据可以进行同行评议、结论检验,同时,无论该科学数据集是独立的,还是与其他数据集关联的,都可以被其他研究人员进行分析和利用,但这样的承诺仅仅是以对科学数据版权、完整性和出处的可识别与可获取为基础的。科学数据的引用还存在诸多问题,比如科学数据不像期刊文献一样有规范化的版本和出处记录,既不能用期刊文献的引用格式对其进行引用,也不能用数据库截屏图对科学数据进行标示。科学数据引用标准及其最佳实践是提升科学数据应用效果的基础,然而在目前所有的科研领域中却是一个严重的缺失。在科学数据管理计划逐渐成为科研基金项目遴选和审批的必备条件时,CODATA 和其他相关机构组成任务团队,专门对科学数据引用的相关关键问题进行调研,最终制定了《科学数据引用标准》,以此来协调各学科领域的科研活动,促进科学活动的常规化和标准化运作^[25]。此外,英国经济与社会委员会出版了《数据引用小册子》^[26]; IASSIST 成立了特别兴趣小组对科学数据的引用进行研究^[27]; 哈佛大学的 IQSS 研究所于 2011 年出台了“数据引用原则”,以促进出版商、数据存档系统和数据引用研究之间的合作,同时通过该原则清晰阐述三者之间的核心问题^[28]。

在数据引用的实践方面,首当其冲的是汤森路透(Thomson Reuters)集团,它根据开放获取环境下科学活动的要求,发布了将数字化研究与强大的发现工具相连接的“科学数据索引”(Scientific Data Index, SDI)^[29],使得研究人员能够快速、轻松地识别并获取最相关的数字研究资源。SDI 还结合了 Web of Knowledge 强大的引文检索功能和导航功能,使得科学文献中所包含的科学数据的效能实现最大化利用,借助 Web of Knowledge 的影响力,SDI 中的科学数据还可被数以百万计的研究人员共享利用,以让科研人员更好的展示自己的研究成果并寻找潜在的合作者。

6 科学数据的质量控制和评价

NIST 与 TRC 的“热物理特性数据库”认为影响科学数据的因素有数据的不确定性和数据的完整性,可通过“试验样本纯度”、“试验方法的误差”、“元数据完整性”和“数据记录的正确性”等指标来评价和判断这两个因素^[30]; Matthew Gamble 认为科学数据的质量可通过“数据本身是否符合某些标准”、“数据来源的可信度”和“数据是否符合科学需求”三个指标来判断^[31]; 我国学者郭秋梅、宋扬则从软件方法的角度,认为科学数据的质量问题主要是由数据采集、整合和录入失误以及显示不当造成的,需要在数据的采集、存储和传输过程中满足相关的质量要求^[32]。

科学数据中心、机构知识库等对科学数据的质量控制与评价是通过制定相应的要求标

准、引用规范及进行同行评议来实现的。如，DCC 对科学数据的出版商、期刊、作者进行了相关约束^[33]；ANDS 的 RDA 对科学数据的遴选原则、数据来源制定了详细的政策说明，认为良好的科学数据应该具有良好的元数据描述、高质量的网络链接、数据的可获取和可再利用等属性，且数据的内容经得起再调查和再试验^[34]；《Nature》对科学数据的来源、上传格式等做出了明确要求，对需要提交的科学数据进行严格的专家审核^[35]。

一些科学数据存档发布平台则通过制定科学数据引用规范、政策来实现对科学数据质量的控制。如，哈佛大学的 IQSS 研究所于 2011 年出台的“数据引用原则”^[36]、Thomson Reuters 的 SDI^[37]、CODATA 的“科学数据引用标准”^[38]、英国经济与社会委员会的《数据引用小册子》^[39]。IASSIST 还成立了特别兴趣小组对科学数据引用的进行研究^[40]；Jan Brase 对“提升网络科学数据开放获取”的代理服务——DataCite 的科学数据质量规范进行了研究^[41]。

科学数据质量的另一种评价方法是进行同行评审。NERC 认为科学数据的专家评审应该作为数据共享利用的首要步骤，因此 NERC 环境数据中心在科学数据创建时，还同时发布该科学数据的专家评审意见，由于这种做法在传统学术出版物中的良好应用而得到广大科研人员的认可，因此 NERC 环境数据中心正在尝试建立有关科学数据完整的同行评审机制^[42]；专门出版原始科学数据（集）的国际性跨学科期刊《地球系统数据》（ESSD）也利用科学数据评论平台 ESSDD，将科学数据（集）首先在 ESSDD 中发布，经过一定时间的公众讨论、专家评议等社会评论过程后，ESSD 最终挑选高质量的科学数据（集）进行出版^[43]。

7 结语

从对国内外科学数据在获取、共享、利用和评价方面的研究进展可以看出，我国科学数据的发展起步较晚，因此，我国在科学数据的开放获取、组织利用、共建共享等方面，可借鉴先进国家和地区的成功实践经验，为我国科学数据的管理政策制定、服务内容创新、开放机制构建等提供借鉴与参考。

（1）建立科学数据开放获取的政策保障与管理机制。开放获取已成为包括科学数据的发展趋势之一，美国、英国、澳大利亚等国际上一一些发达国家政府也已制定了贯穿于科研成果产生、发布、管理与利用四大环节的科学数据获取政策体系，美国海洋与大气局、NIH 等机构对科学数据的开放获取实践有着成功的经验。反观我国在科学数据开放获取的政策制定与实践，到目前为止，政府还没出台完整的、系统的科学数据开放获取政策，科研资助机构如国家自然科学基金委员会、全国哲学社会科学规划办公室以及科技部、教育部等相关组织机构也都未在资助项目申请指南中有明确的科研成果开放规定。使得我国尽管已有一些科研机构如中国科学院通过机构知识库来开放科研成果，但科学数据的开放获取还没有形成规模。我国应借鉴发达国家在科学数据开放资助政策、科学数据开放管理政策、科学数据共享利用政策方面的成功实践，尽快建立科学数据开放获取政策保障与管理机制，并在政策的指导、规范下推动科学数据的管理、存储、开发、发现与利用。

(2) 统一科学数据的注册与引用标准。寻求和建立数据之间的关联关系、挖掘隐藏在数据关联关系之中的隐性知识以及规律是近年来科学数据的研究热点之一，也是关联数据、大数据分析等学科领域所研究的问题之一。但由于科学数据有别于一般资源数据的特点，即科学数据不像期刊论文、学位论文等资源一样有明确的数据出处、发布机构等数据信息与数据属性，因而在实现科学数据间的关联关系时会存在诸多问题。为了有效的实施科学数据的科学管理与实践应用，并与国际接轨，我国应参考国际上主流的科学数据引用标准来制定我国科学数据的注册与引用标准，提高我国科学数据的可检索性、可发现性、可解释性与可重新利用性。从而推动我国科学数据的注册与引用向规范化、标准化与国际化方向发展，让国内外更多的科研工作者充分利用我国的科学数据，从而扩大和提高我国科学数据的影响力与科研价值。

(3) 延伸科学数据的服务及范围。目前，我国虽然已有科学数据共享和利用的服务平台，如中国科学院国家科学图书馆兰州分馆建设的“中国(西部)环境与生态科学知识积累平台”，向用户提供什么科学数据管理和组织的相关服务，但从我国科学数据服务的整体发展来看，科学数据服务的发展还属于起步阶段，对作为社会信息服务与资源存储机构的图书馆来说在科学数据的服务方面有着极大的发展空间。如，可根据一些商业期刊（如Nature、Science）、开放期刊（如PLOS One、Biodiversity Journal）、组织机构（如NIH、NASA）等制定的科学数据存档政策，提供科学数据的发布等培训服务；借鉴哈佛大学的Dataverse Network和康奈尔大学图书馆的DataStaR等建立数据仓储、为用户提供数据存档、逐步实现数据共享与开放的成功实践，提供科学数据的管理与存储服务。

参考文献:

- [1]姜晓虹.国内科学数据相关研究进展分析[J].图书情报工作, 2009, 53(13):50-53.
- [2]Distributed Active Archive Centers[EB/OL].[2012-12-15].<http://nasadaacs.eos.nasa.gov/about.html>.
- [3]刘细文,熊瑞.国外科学数据开放获取政策特点分析[J].情报理论与实践, 2009, 32(9):5-9,18.
- [4]Panton Principles for Open Data in Science[EB/OL].[2013-01-21].
<http://www.pantonprinciples.org/>.
- [5]Kauppinen,T.,Espindola, G. M. D.. Linked Open Science-Communicating, Sharing and Evaluating Data, Methods and Results for Executable Papers[J].Procedia Computer Science, 2011, (04) :726.
- [6]刘闯.美国国有科学数据共享管理机制及对我国的启示[J].中国基础科学, 2003, (1): 36-41.
- [7]Policy RE Commendations for Open Access to Research Data in Europe[EB/OL]. [2013-01-21].
<http://recodeproject.eu/>.
- [8]Open Access to Research Data: The European Commission's consultation in progress[EB/OL].

- [2013-01-21]. <http://openeconomics.net/2013/07/09/open-access-to-research-data/>.
- [9]BBSRC launches policy to share research data[EB/OL].
[2013-01-05].<http://www.bbsrc.ac.uk/news/archive/2007/070501-pr-data-sharing.aspx>.
- [10]MRC.Data sharing[EB/OL].[2013-01-05].<http://www.mrc.ac.uk/Ourresearch/Ethicsresearchguidance/datas haring/index.htm>.
- [11]Nonmember access to data[EB/OL].[2012-12-15].
<http://www.icpsr.com/ICPSR/membership/nonmember.html>.
- [12]Frequently asked questions on data sharing[EB/OL].[2012-12-15].
http://grants.nih.gov/grants/policy/data_sharing/data_sharing_faqs.htm.
- [13]NIH data sharing policy and implementation guidance[EB/OL].[2012-12-15].
http://grants.nih.gov/grants/policy/data_sharing/data_sharing_guidance.htm.
- [14]Distributed Active Archive Centers[EB/OL].[2012-12-15].<http://nasadaacs.eos.nasa.gov/about.html>.
- [15]National Center for Biotechnology Information[EB/OL].[2013-02-16].
<http://www.ncbi.nlm.nih.gov/>.
- [16]National Institute of Advanced Industrial Science and Technology[EB/OL].[2013-02-23].
<http://www.aist.go.jp/index-en.html>.
- [17] 国家科技基础条件平台 中国科技资源共享网 [EB/OL].[2013-02-03].<http://www.escience.gov.cn/>.
- [18]西部数据中心知识积累平台[EB/OL].[2013-01-10].<http://westdc.westgis.ac.cn/knowledge>.
- [19]中国地质科学院数据网[EB/OL].[2013-01-10].<http://www.geoscience.cn/>.
- [20]CODATA[EB/OL].[2013-03-10].<http://www.codata.org/taskgroups/index.html>.
- [21]The Global Information Commons for Science Initiative[EB/OL].[2013-01-10].<http://www.codata.org/wsis/GlobalInfoCommonsInitiative.html>.
- [22]e-SDDC. Global Alliance for Enhancing Access to and Application of Scientific Data in Developing Countries[EB/OL].[2013-07-25].<http://www.e-sddc.org/cn/index.html>.
- [23]Jan Brase.DataCite-A global registration agency for research data[C].Fourth International Conference on Cooperation and Promotion of Information Resources in Science and Technology, 2009:257-261.
- [24]Data Citation Awareness[EB/OL].[2012-10-23].
<http://ands.org.au/guides/data-citation-awareness.html>.
- [25]Data Citation Standards and Practices Task Group[EB/OL].[2013-01-16].
<http://www.codata.org/taskgroups/TGdatacitation/index.html>.
- [26]Data Citation Brochure published by the UK's Economic and Social Research Council[EB/OL].
[2013-01-20].<http://www.iassistdata.org/blog/data-citation-brochure-published-uks-economic-and-social-research-council>.
- [27]IASSIST Publishes a Quick Guide to Data Citation[EB/OL].[2013-01-20].
<http://www.iassistdata.org/blog/iassist-publishes-quick-guide-data-citation>.
- [28]Data Citations Principles[EB/OL].[2013-01-20].
<http://events.iq.harvard.edu/events/node/2462>.
- [29]Data Citation Index on Thomson Reuters Web of Knowledge[EB/OL].[2012-11-01].

- http://thomsonreuters.com/content/press_room/science/730914.
- [30]Qian Dong, Xinjiang Yan, Randolph C. Wilhoit, et al. Data Quality Assurance for Thermophysical Property Databases—Applications to the TRC SOURCE Data System[J]. Journal of Information Computer Science, 2002, (42) : 473-480.
- [31]Matthew Gamble, Carole Goble. Quality, Trust, and Utility of Scientific Data on the Web: Towards a Joint Model[C]. Koblenz: Proceedings of the International Conference on Web Science 2011, 2011:1-8.
- [32]郭秋梅,宋扬.基于软件方法的科学数据质量控制[C].先进制造与数据共享国际研讨会, 2007:539-543.
- [33]Data Citation and Linking[EB/OL].[2012-12-21].
<http://www.dcc.ac.uk/resources/briefing-papers/introduction-curation/data-citation-and-linking>.
- [34]Who contributes to Research Data Australia[EB/OL].[2012-10-13].
<http://researchdata.ands.org.au/>.
- [35]Guide to Publication Policies of the Nature Journals[EB/OL].[2012-12-21].
<http://www.nature.com/authors/gta.pdf>.
- [36]Data Citations Principles[EB/OL].[2013-01-20].<http://events.iq.harvard.edu/events/node/2462>.
- [37]Data Citation Index on Thomson Reuters Web of Knowledge[EB/OL].[2012-11-01].
http://thomsonreuters.com/content/press_room/science/730914.
- [38]Data Citation Standards and Practices Task Group[EB/OL].[2013-01-16].
<http://www.codata.org/taskgroups/TGdatacitation/>.
- [39]Data Citation Brochure published by the UK's Economic and Social Research Council[EB/OL].[2013-01-20].
<http://www.iassistdata.org/blog/data-citation-brochure-published-uks-economic-and-social-research-council>.
- [40]IASSIST Publishes a Quick Guide to Data Citation[EB/OL].[2013-01-20].
<http://www.iassistdata.org/blog/iassist-publishes-quick-guide-data-citation>.
- [41]Jan Brase. DataCite - A global registration agency for research data[C]. Fourth International Conference on Cooperation and Promotion of Information Resources in Science and Technology, 2009:257-261.
- [42]Sarah Callaghan, Steve Donegan, Sam Pepler. Making Data a First Class Scientific Output: Data Citation and Publication by NERC's Environment Data Centres[J]. The International Journal of Digital Curation, 2012, 7(1): 107-113.
- [43]Earth System Science Data (ESSD)[EB/OL].[2013-01-31].<http://www.earth-system-science-data.net/home.html>.

作者简介:李慧佳(1984-),女,中国科学院资源环境科学信息中心助理馆员;马建玲(1969-),女,中国科学院资源环境科学信息中心研究馆员;
王楠(1978-),女,中国科学院资源环境科学信息中心副研究馆员;王思丽(1985-),女,中国科学院资源环境科学信息中心馆员;张秀秀(1982-),女,中国科学院资源环境科学信息中心馆员。

联系电话: 0931-8270026

联系邮箱: lihj@llas.ac.cn

尊敬的编辑老师，您好！

根据贵刊修改要求，本文在以下方面做了修改，主要有：

- 1.对全文的语句、错别字进行了梳理与纠正；
- 2.对导论部分进行了重新梳理、对结语部分进行了扩展梳理和条理化；
- 3.正如修改意见所言，本文说采用文献计量方法欠妥，在文中已取掉“文献计量”字样；
- 4.正文引用的地方全部改在了内容之后，对报告采用了书名号。

特此说明。