

论全球变化科学研究数据出版

刘 闯

(中国科学院地理科学与资源研究所, 北京 100101)

摘要: 随着大数据时代的到来和全球问题研究的深入, 全球变化科学研究数据已经成为科学研究成果中重要组成部分。其数据量之大、涉及面之广、发展速度之快、参与科学家之多已经超出人们预想。数据集约型科学研究已经成为全球变化研究的主导方式, 科学数据共享是数据集约型科学研究的必经之路。全球变化科学研究数据出版为完善我国科学数据共享机制开拓新途径。数据投稿、数据评审、数据出版、数据服务与数据引用是数据出版过程中的关键性环节, 数据论文和实体数据配套出版是数据的特殊出版方式, 经同行专家评审通过后的首期数据集通过 DOI (Digital Object Identifier) 注册系统出版 (DOI: 10.3974), 为中国全球变化科学研究数据知识产权保护和数据共享关系调整、为国家相关政策、链带技术、科学家研究方式、科技成果评价体系、商业机遇等多方位促进数据共享提供典型案例。数据出版对全球变化科学研究产生的影响将具有划时代意义。

关键词: 全球变化; 科学数据; 数据论文; 数据出版; 共享服务

DOI: 10.11821/dlxb2014S000

1 问题的提出

全球变化科学研究是一个多学科交叉融合、综合性、全局性的学科领域。在科学界, 没有几个学科类似全球变化研究这样对科学数据共享的要求如此之迫切、数据共享对学科的发展如此至关重要。可以说, 没有数据共享, 全球变化科学研究几乎成为一句空话。自1994年孙枢院士等学者呼吁中国科学数据共享应该引起重视后^[1], 中国在如何解决科学数据共享这个问题上至今花费了20年的时间。期间, 一个重要的国家战略举措是在第196次香山会议(2002)后, 在时任国家科学技术部部长徐冠华的亲自倡导和推动下, 中国正式启动了“国家科技基础条件平台”建设计划^[2], 其中包括科学数据共享平台(2003)。这项举措被列入中国中长期科技发展规划。“规划”明确了至2020年国家科技基础条件平台建设总目标, 确定了整合、共享、完善、提高的平台建设方针^[3], 制定了政府主导、强化顶层设计、优化平台布局、重点项目部署、大项目数据汇交等一系列实施战略和措施^[4]。

随着大数据时代的到来, 科学数据领域发生了超出人们预想的态势以数据量大、涉及面广、发展速度快、参与科学家多等特点迅速膨胀, 数据以横断学科的特点贯穿于每一个科学领域, 并引发科学研究方式的根本转变—数据集约型科学研究方式的出现。在全球变化研究领域, 这种改变的显著特征是每一位科研人员只要开展科学研究活动, 都将没有选择地被卷入到科学数据的主体队伍中。在这种形势下, 能否有效调动作为应用科学数据、产生科学数据的主体—科研人员—的主动性、积极性成为科学数据共享的关键。自国家科

收稿日期: 2014-04-20; 修订日期: 2014-06-20

基金项目: 国家自然科学基金项目(70240004; 7087311) [Foundation: National Natural Science Foundation of China, No.70240004; No.7087311]

作者简介: 刘闯(1948-), 博士, 研究员, 国际地球观测组织(GEO)数据共享工作组成员, 中国专家组副组长, 国际科学技术数据委员会发展中国家任务组秘书, 国际科学技术数据奖获得者(2008), 中国科学技术协会联合国事务咨询委员会委员, 长期从事全球变化科学数据研究。E-mail: lchuang@igsrr.ac.cn

技基础条件平台计划实施以来,无论是国家政府主管部门、还是参与平台建设和运行的科学家们都在尽心尽力促进科学数据的整合和共享。然而,经历了10年,中国科技界科学数据共享的局面仍不容乐观。

以笔者所在的中国科学院地理科学与资源研究所(以下简称“地理资源所”)为例,这是一个有70多年历史,以地理、资源、生态、环境和人地关系为主要研究领域的国家级研究所,是世界该领域在职科研人员最多的大研究所。2011年国家确定了23个国家科技基础条件平台认定名单中,就有二个由地理资源所主持(地球系统科学数据共享平台、国家生态系统观测研究网络)^[5],这在全国是独一无二的。不仅如此,中国科学院资源环境科学数据中心、国家发展与改革委员会电子政务空间数据分中心、国际科学技术数据委员会(CODATA)发展中国家科学数据保藏与共享任务组、国际科学理事会世界数据系统(ICSU-WDS)可再生资源与环境数据中心等与科学数据共享相关的国际级、国家级、科学院级数据共享项目都有地理资源所科学家们的主导或参与。这些项目虽各有侧重,但有一个共同的政策——科学数据共享。然而至今,这些项目成果中很大一部分不仅不能实现全社会共享,也不能在本所科研人员中全面实现共享。这种怪现象或许多少年后我们的后代不会理解为什么会是这样子,但是,它确实是目前实实在在存在的、即便是有行政管理权的几任所长也没能够消除的现象。

地理资源所是中国科研单位中承担国家自然科学基金项目最多的研究所之一。在2010-2013年4年间该所共获得国家自然科学基金仅面上项目就达175项之多(国家自然科学基金资助项目统计资料,2013,2012,2011,2010年),仅面上基金项目经费就超过1亿元人民币(1 1419万元)^[6-9](重点、重大、国际合作、青年基金等项目还没有统计在内)。在这些研究中,几乎每一个项目的执行都需要数据的支撑,而每个项目又都产生新的数据。目前,这些数据大多还处于零散状态,既不能作为科学研究成果得到认可和保护,也不能在科研活动中广泛得到共享。国家科研经费在数据不能有效共享方面造成的浪费至今没有人能够给出一个估算结果,数据集约型科学研究从操作层面上更难以深入。

地理资源所目前的状态在中国不是个案,科研人员、教学人员渴望共享别人数据却多数拿不到或不能及时拿到,自己的数据也不能主动、积极、及时地公开和共享的现象在我国依然普遍存在^[10-11]。中国近10年来的经历表明:①中国科学数据共享平台发展速度落后于科学数据发展速度;②落后的根本原因是数据共享机制不健全,仅有由上向下科学数据共享的一种机制尚不能调动大多数科研人员在科学数据共享中的主动性和积极性;③中国科学数据共享需要有进一步完善性的新思路和新举措。

地理资源所目前的状态,在世界也带有普遍性。在2014年6月在北京召开的科学大数据高层国际研讨会上,CODATA(国际科学技术数据委员会)主席、中国科学院遥感与数字地球研究所所长郭华东院士指出:“科学数据共享是大数据科学必经之路,然而,这个问题还远没有解决”。前CODATA主席、CODATA奖获得者、美国前国家标准局官员John Rumble博士认为,除非数据作者看到了商业可以回报的价值,否则,他(她)们是不会把数据拿出来共享的。另一位CODATA奖获得者、美国国家科学研究数据和信息委员会主席Paul Uhler博士认为,除非数据科学家在数据方面的科学贡献能够得到科技界的认可(Recognized)和回报(Rewarded),否则,科学数据共享问题就难以从根本上得到解决。前CODATA主席、日本东京大学Shuichi Iwata教授,前CODATA主席、前印度科学院院长Krishan Lal院士表示日本和印度存在着同样的问题。来自南非、肯尼亚、印度尼西亚、泰国、巴西、哥伦比亚等发展中国家的科学家们异口同声地表示了赞同的意见。可见,要实现根本意义上的数据共享,科技界对科学家在科学数据方面的认可和回报是一个关键性

的世界问题。

与科学数据不能有效共享的状态相反,科学论文领域却是另外一番欣欣向荣景象:同样作为科学研究主体的科研人员争先恐后发表论文,而且还力争尽早、尽快在影响因子高的杂志发表。再审视一下科学研究成果在科技界的认可和评价工作,科研成果的评价主要基于论文基础上进行的,也包括专著、专利、决策咨询报告等。这些成果是科研人员职称晋升最重要衡量标准,是可持续获得科研经费不可或缺的指标,是得到同行认可并获得奖励的必不可少的条件,在很多单位,科研人员的工资收入和福利待遇也与发表论文密切挂钩。而科学研究人员职称能否晋升、科研项目经费能否获得、能否在本领域获得科技成果奖,甚至是工资多少、福利待遇如何是科研人员生存和发展核心利益四要素。因此,科学数据共享的关键性问题就转到了科学家科研数据成果的评价上。

要解决这个问题,即对科学家科学数据成果评价的首要问题是科学数据成果的出版。只有经过类似于科学论文发表一样,经过同行专家评审的科研数据的正式出版,对科学家科学数据成果的评价才有可能有序地进行。这样也才能够有利于应对这些数据成果的作者对科学的贡献是否和如何纳入到科研成果评价体系,在科学家最关切的利益四要素中如何体现和回报等一系列问题的解决。只有这些问题得到恰当解决,数据作者参与数据共享的主动性、积极性、迫切性才有可能得到充分发挥,国家科研经费在数据产出方面的成果滞留、遗失和流失现象才能从最大程度上给予解决。

国际科学联合理事会 (ICSU) 科学数据和信息发展评估报告中明确指出 (2004 年)^[12], 促进科学家平等地共享科学数据和信息是 ICSU 的重要任务,在保护科学数据和信息知识产权基础上的数据共享是当前国际科技界最重要的前沿问题^[13]。2005 年美国开始每一个科学数据逐一给予 DOI 注册的工作^[14]。经过近十年的积累,美国和欧盟在 DOI 标准和技术支撑下的科学研究数据知识产权保护与共享的机制开始形成并运转^[15-21]。2012 年 5 月,国际标准组织正式批准 DOI 作为国际标准 (ISO 26324, Digital Object Identifier System)^[22-24]; 2012 年, CODATA 与国际科技信息委员会 (ICSTI) 联合成立科学数据引用任务组 (Data Citation Task Group); 2013 年,科学研究数据全球联盟 (RDA) 与世界数据系统 (WDS) 联合成立数据出版共识组 (Data Publishing Interesting Group)。在这些国际组织世界性的大动作驱动下,在知识产权清晰的基础上,全球性科学研究数据大汇集正在涌动并开始形成可持续潮流。以美国主导的全球生物多样性基础设施 (GBIF) 为例^[25],目前已经有 3 亿多数据集 (库) 业已出版 (获得 DOI 注册)^[26-28]、300 多个数据出版中心纳入到该共享网。在德国,仅马普学会 80 个研究所目前已经有 100 多万万个数据集 (库) 注册了 DOI/Handel^[29]。

2 全球变化科学研究数据出版

根据国际 DOI 标准和中国 DOI 注册管理办法^[30],中国科学院地理资源所于 2013 年在国家科技部中国科学技术信息研究所 DOI 注册与服务中心注册 DOI 科学数据出版单位获得批准 (DOI:10.3974)。之后,地理资源所相继开发了全球变化科学数据注册与出版规则和程序、DOI 编码系统、同行评价流程与规则、元数据标准,数据论文格式、数据归档办法,数据共享政策,数据服务内容与网络实现等一系列与全球变化科学研究数据出版相关的准备工作,现已就绪。第一期全球变化科学研究数据出版由地理资源所、中国地理学会联合主办, CODATA 发展中国家任务组、肯尼亚 Jomo Kenyatta 农业与技术大学 (JKUAT)、国家地球系统科学数据共享平台、数字化林超地理博物馆联合协办,网站: <http://www.geodoi.ac.cn>。

2.1 全球变化科学研究数据出版关键环节^[31]

科学数据与科学论文都是科学研究活动的产出成果,科学论文的出版过程和管理办法有了100多年的历史经验,已经形成了完整的投稿、评审、发表、出版、检索、评估模式,虽然,随着数字化出版物的出现,论文出版体系也在不断的改进,但是,其核心精华部分仍在延续。科学数据出版与科学论文出版有相似之处,更有不同的方面。有关数据投稿、评审、存储、可视化、检索、服务、共享、引用、评估^[32]、网络系统等一系列与数据出版相关的环节均处于新生阶段,新问题很多,机遇也处处存在。这些问题和机遇,不仅涉及到技术和标准,也关系到机制、政策以及各个相关方面的利益和责任再分摊问题。全球变化科学研究数据出版以率先探索性的行动,探讨中国由下向上机制(每位科学家都有平等参与机会作为起点)促进数据共享的关键环节。诚然,科学数据出版流程中各个环节都有新问题,但是,数据投稿、数据评审、数据出版、数据服务与数据引用五大环节尤为关键,鉴于本文篇幅,本文仅就数据投稿、数据出版和数据共享服务简要论述如下,其他有关内容将另行文论述。

2.2 全球变化数据集(库)作者投稿

数据集(库)作者投稿核心内容包括作者承诺和数据投稿。

2.2.1 作者投稿自律承诺四原则

申请数据出版的作者需要承诺:① 每一位数据作者(包括合作作者)确认对该数据集(库)具有自主知识产权;作者同意将该数据著作权中的编辑权、不同介质复制权、依注册中数据公开范围内的数据散发权、网络传播权、多语种翻译权、印刷权和上述产权的转让权与数据出版者共同具有;该数据集(库)版权由数据出版者单独所有^[33-37];② 符合中华人民共和国国家安全、个人隐私和相关科学研究项目政策规定^[38];③ 保证科学数据集(库)的真实性,即没有任何作假数据(部分或全部,包括数据产生过程);④ 没有一稿(一个数据库)二投或多投(一个数据库仅能出版一次,更新数据库以不同版本出版除外)。

2.2.2 作者投稿内容

为科学数据集(库)出版的投稿,作者需要准备好数据集(库)的元数据、数据论文、实体数据三部分内容。元数据、数据论文和实体数据集(库)均需要同行专家评审。

(1) 元数据

元数据是数据集(库)出版获得DOI注册的关键信息,全球变化科学研究数据出版系统将元数据标准各项栏目逐一编撰在数据集(库)出版申请系统中。作者只要按照网站需要填写的内容逐一填写就可以完成元数据的内容上报。

(2) 数据论文

数据论文是一篇集数据集(库)的说明文和论证该数据集(库)创新性和可靠性的论文二者兼有的文章。因此,它既有说明文的特点,也有论文的特征,它是一种特殊文体的、针对一个数据集(库)的科学数据说明论文。

数据论文需要对下列信息给予清晰的交代:数据集(库)名称,作者、作者所在单位及其联系信息、通讯作者及其联系信息、数据出版者及其联系信息、数据引用格式、数据来源引用格式、关键词、数据集(库)组成及各组成部分数据格式、数据量、数据共享范围、数据获取方式、数据共享服务中心及其联系信息、数据产生资助项目信息、数据集(库)内容涵盖的时间、空间、主要内容信息。此外,数据论文中需要对该数据的创新性和可靠性给予翔实的论证。下列内容应尽量满足:数据集(库)产生的基础数据和信息的可靠性评价、数据集(库)产生和获取的方法和途径、数据集(库)内容的阐述、数据集(库)可视化展示图、对该数据集(库)可靠性检验方法和检验结果、数据集(库)产生、主要参

考文献(包括主要参考数据)等。

(3) 实体数据

实体数据是数据出版的核心。对于全球变化科学研究数据来说,数据呈多样性,表现在数据内容、覆盖区域、空间分辨率、时间频率、数据时代、数据格式、数据量、数据版本等。它正是丰富多彩的全球变化科学数据精髓所在,是科学家科学研究数字化成果最精彩的表达,是当代以至长远国家科技资源财富的重要组成部分。

2.3 数据出版—数据论文与实体数据成对配套出版并分获 DOI

全球变化科学数据出版采取数据及其相关的数据论文关联方式同时出版。第一期,数据论文采取纸质和电子版二种方式以地理学报增刊的方式出版;实体数据在 DOI:10.3974/以数字化方式出版。全部信息均可以在全球变化科学研究数据出版系统网站 (<http://www.geodoi.ac.cn>) 获得。科学数据与科学数据论文配套出版是全球变化科学数据出版的一种特殊方式。无论科学数据,还是数据论文都经同行专家评审后出版。

实体数据出版犹如科学论文发表一样是科学家科学思想和科学结论的公开发表。对于同一地理学现象,不同的科学家可能有不同的学术观点;在不同的学术观点下分析得到的数据有可能不一致。科学数据的出版将本着“百花齐放、百家争鸣”、尊重作者、保护作者知识产权的基本原则,予以无歧视地及时出版。由于有些数据研发部分成果已经以科学论文的方式出版,作为数据论文将不再重复,欠缺的内容在数据论文中进行补充。对于科学数据研发采用的基本数据相同、研发方法论相同或相近、地理空间和表达时间相同的数据集(库),将以申请为先的原则出版;对于其后申请的,如果没有足够的创新内容,则原则上不考虑准予出版。这样处理,也是遵循科学论文发表的基本原则—尊重科学家的首创劳动成果和对该成果知识产权保护的基本原则。本期出版的数据报批国家测绘局审核批准,审图号 GS (2014) 1266 号。全部数据正式归档至地理资源所数字化图书馆、中国科学院资源环境科学数据中心和国家地球系统科学数据共享平台。

2.4 数据共享与服务

(1) 全球变化科学数据共享服务的根本核心是工作在一线的科研人员能够免费、无偿、任何时候都可以通过计算机网络下载获取数据的共享。要做得这一点,有二种运行机制,一种是国家直接购买服务(国家公共财政资助共享服务)^[39-42],全民共享;另一种是单位购买服务(科研单位和大学等个体单位),由商业运行机制的共享服务商提供共享服务,单位内科研人员可以共享数据。目前,大多数科学论文共享采取的是后一种方式。科学数据采用哪一种共享机制还需要更多的实践。本期出版的科学数据是作为试验研究,采取全社会开放和无偿共享服务方式,可持续性全社会无偿开放共享机制的形成还需要由上向下共享机制的支持和进一步的实践研究。

(2) 数据共享服务包括对全社会用户的共享服务,对数据作者的回报服务和对特殊用户的专题服务三方面内容。目前,全球变化科学数据注册与出版系统提供对全社会用户的共享服务包括网络免费浏览服务和网络免费下载服务二项内容。由于有些数据量比较大,为了便于用户下载,对于一些数据集(库)数据量比较大,很难以一个数据文件 ftp 下载,在提供服务过程中,将数据按照一定规则(经纬度)划分成若干数据文件,根据目前网络带宽允许的情况下,确保每一个数据文件在几分钟内用户畅通下载。在对数据作者的回报服务中,全球变化科学数据注册与出版系统为每一个出版的数据单独设立一个独立的版面,并通过计算机自动计算和查询系统,在数据浏览次数、数据下载次数、数据下载量方面予以即时网上统计和显示。此外,该系统还将与中国 DOI 注册中心合作,提供数据引用次数和引用者统计和相关信息,便于作者了解自己的数据被世界范围内的用户引用情况。

对特殊需求用户的专题服务是数据增值服务的重要内容,出版方将与有关方面,如国家科技基础条件平台—地球系统科学数据共享平台等合作,在尽可能的情况下,起到“桥梁”和“催化剂”作用,为特殊需求的国家重大科研项目、研究机构和个人等提供专题服务。

3 第一期数据的出版

全球变化科学研究数据第一期出版 20 个数据集(库)。伴随每一个数据集(库)的出版,各有一个与其相对应的数据论文(中英文版)关联出版。在第一期出版的科学数据集(库)中,作者有来自中国科学院地理资源所的科研人员,有非中国科学院的科研人员、大学教授,也有非地理学专业人员,更有在籍的研究生。在数据方面,有数据量不足 1MB 的数据集,也有数据量在 GB 级别的数据集;有矢量数据、栅格数据、表格数据、图像数据、影视数据;数据空间分辨率有 30 m 分辨率、250 m 分辨率、500 m 分辨率和 1 km 分辨率,也有点、站的数据集。全球变化科学研究数据出版体现了数据及数据作者多样性的特点,这正是全球变化学科特点的体现。

4 结束语

近 20 年来,中国科学家一直呼吁和努力推动的全球变化科学数据共享问题是现代科学研究成果管理的重大问题,也是现代地理学科学研究方式变革的关键性问题。长期以来,在全球变化科学数据共享领域,由公共财政资助产生的、大量的、分布于科学研究各个角落的科学数据共享成为“老大难”。在 DOI 国际标准、技术支撑下,在科学数据知识产权保护基本准则下,科学数据出版是大数据时代的必然产物,是数据集约型科学研究不可逾越的必经之路。在全世界争先恐后抢占科学数据出版“制高点”的大形势下,在崭新的科学数据出版事业正在悄然兴起并迅速发展过程中,全球变化科学研究数据的出版,将对中国在国家层面、科研机构层面、科学家个人层面、学术组织层面、知识出版服务层面等多方位研究和探索数据共享之路提供典型案例。笔者相信,科学数据的出版,将在国家政策、相关技术标准、科学家科研方式、科技成果评价体系、商业机遇等方面提出一系列具有挑战性的新问题,它对该领域中国相关的国家战略和政策调整的影响将是前所未有的。可以预料,未来 20 年,全球变化科学研究数据出版将会出现如雨后春笋般的发展态势;它对全球变化科学研究产生的影响将是全方位的^[43]、深刻的和具有划时代意义的。

致谢:感谢中国科学院地理科学与资源研究所葛全胜所长,国际科学技术数据委员会主席郭华东院士,中国地理学会刘燕华理事长、张国友秘书长,《地理学报》(中英文版)编辑部何书金主任对本项工作的支持。感谢天津中科遥感信息技术有限公司开发部的合作者们为全球变化科学数据注册与出版系统做出的实用性的合作开发工作。感谢科技部中国科学技术信息研究所资源共享促进中心彭杰主任、董诚副主任多年来在推动数据出版理念所作出的扎扎实实的促进工作。感谢编委会全体学者多年来在科学数据共享领域作出的不懈努力、团队合作和辛勤劳动。由衷地感谢科技部中国科学技术信息研究所 DOI 注册与服务中心刘耀副主任,郭小峰总工程师和他们的团队在推动和建设与国际接轨的中国 DOI 注册与管理国家基础设施作出的卓有成效的工作,不无夸张地说,没有他们为大家建设的中国 DOI 注册基础设施,由中国主导的、必将在世界产生重要影响的全球变化科学研究数据出版就不可能适时启动和发展。

参考文献 (References)

[1] Department of Earth Science, Chinese Academy of Sciences. Recommend on data sharing of earth science, resources,

- environment. *Advance in Earth Sciences*, 1996, 11(1): 122-123. [中国科学院地学部. 关于进一步做好我国地球科学、资源与环境科学研究基础资料与数据共享的建议. *地球科学进展*, 1996, 11(1): 122-123.]
- [2] National Science and Technology Infrastructure Program. [国家科技基础条件平台建设纲要 (2004-2010) www.most.gov.cn/tjcw/tckytj/200708/t20070813_52439.htm.]
- [3] National Mid-Long Term Scientific and Technical Planning (2006-2020). [国家中长期科学和技术发展规划纲要 (2006-2020年)], http://www.most.gov.cn/mostinfo/xinxifenlei/gjkjgh/200811/t20081129_65774.htm.
- [4] Ministry of science and technology, the Ministry of Finance, Joint Announcement on Promoting the national science and technology infrastructure, 2008. [科技部、财政部关于进一步推动国家科技基础条件平台开放共享工作的通知, 国科发计[2008]722号, http://www.most.gov.cn/fggw/zfwj/zfwj2008/200901/t20090104_66481.htm.]
- [5] The Ministry of Science and Technology, Announcement on Identification and Performance Assessment of National Scientific and Technical Information Infrastructure, 2011. [科技部, 关于开展国家科技基础条件平台认定和绩效考核工作的通知, 国科发计[2011]318号, 2011年08月05日, http://www.most.gov.cn/fggw/zfwj/zfwj2011/201108/t20110805_88853.htm.]
- [6] Statistics of National Natural Science Foundation Projects, NSFC, 2010. [国家自然科学基金资助项目统计资料, 2010年度, 国家自然科学基金委员会编.]
- [7] Statistics of National Natural Science Foundation Projects, NSFC, 2011. [国家自然科学基金资助项目统计资料, 2011年度, 国家自然科学基金委员会编.]
- [8] Statistics of National Natural Science Foundation Projects, NSFC, 2012. [国家自然科学基金资助项目统计资料, 2012年度, 国家自然科学基金委员会编.]
- [9] Statistics of National Natural Science Foundation Projects, NSFC, 2013. [国家自然科学基金资助项目统计资料, 2013年度, 国家自然科学基金委员会编.]
- [10] LIU Chuang. Study on scientific data sharing mechanism of China. *Land and Resources Informatization*, 2004, (1): 5-7. [刘闯. 我国科学数据共享机制建设研究. *国土资源信息化*, 2004, (1): 5-7.]
- [11] LIU Chuang. Recent Developments in Environmental Data Access Policies in the P.R. China, Open Access and the Public Domain in Digital Data and Information for Sciences. *National Research Council of the National Academies, Washington D.C., USA*. 2004, 74-78.
- [12] Roberta Balstad, Jean Bonnin, Marc H Brodsky, Liu Chuang, Carlos A Correa, Norihisa Doi, Ray Harris, Andrew M Kaniki, Vitaly A Nechitailenko, Pierre L-J Ritchie, Rajashekar T B. Priority Area Assessment on Scientific Data and Information, ICSU Report of the CSPR Assessment Panel on Scientific Data and Information, International Council for Science, 2004. http://www.icsu.org/publications/reports-and-reviews/priority-area-assessment-on-scientific-data-and-information-2004/PAA_Data_and_Information_report.pdf.
- [13] LIU Chuang, WANG Zhengxing. *International Scientific Data Policies*. Beijing: Scientific and Technical Documentation Press, 2005. [刘闯, 王正兴. 国际科学数据政策与法规. 北京: 中国科技文献出版社, 2005.]
- [14] Norman Paskin, International DOI Foundation, Oxford, UK. Digital Object Identifiers for Scientific Data, *Data Science Journal*, Volume 4, 2005, <http://dx.doi.org/10.2481/dsj.4.12>, https://www.jstage.jst.go.jp/article/dsj/4/0/4_0_12/_article.
- [15] LIU Chuang. Study on USA scientific data management mechanism. *China Basic Science*, 2003, (1): 34-39. [刘闯. 美国国有科学数据共享管理机制及对我国的启示. *中国基础科学*, 2003, (1): 34-39.]
- [16] LUO Zongjun, LIU Chuang, WANG Zhengxing, The legal issues of US global change studies. *ADVANCE IN EARTH SCIENCES*, 2003,18(3): 464-470. [罗宗俊, 刘闯, 王正兴. 美国全球变化研究的法律基础. *地球科学进展*, 2003, 18(3): 464-470.]
- [17] LIU Chuang et al. Analysis of National Strategy of Global Change Studies: USA. Beijing: SinoMaps Press, 2000. [刘闯等. 全球变化研究国家战略分析: 以美国为例. 北京: 地图出版社, 2000.]
- [18] WANG Zhengxing, LIU Chuang. Freedom of information act and publicly funded information access in the UK, *Studies in Science of Science*. 2006, 24(5): 688-695. [王正兴, 刘闯. 英国的信息自由法与政府信息共享. *科学学研究*, 2006, 24(5): 688-695.]
- [19] WANG Zhengxing, LIU Chuang. Comparison study on public funded information sharing mechanism models. *China Basic Science*, 2005, (5): 36-42. [王正兴, 刘闯. 政府信息资源共享两种模式及其效益比较. *中国基础研究*, 2005, (5): 36-42.]
- [20] WANG Zhengxing, LIU Chuang, LIN Xin, MA Lin. Sustainable data sharing: The key is benefit equilibrium. *Forum*

- on Science and Technology in China, 2005, (6): 92-96. [王正兴, 刘闯, 林昕, 马琳. 科学数据可持续共享: 关键是利益均衡. 中国科技论坛, 2005, (6): 92-96.]
- [21] SUN Honglie, LIU Chuang. Priority areas of international scientific data development. China Basic Science, 2003, (1): 13-18. [孙鸿烈, 刘闯. 国际科学技术数据前沿领域发展研究. 中国基础科学, 2003, (1): 13-18.]
- [22] International DOI Foundation, <http://www.doi.org>.
- [23] ISO 26324: 2012, Digital Object Identifier System, http://www.iso.org/iso/catalogue_detail.htm?csnumber=43506.
- [24] DOI Handbook, Version 5, April 2012, <http://www.doi.org/hb.html>.
- [25] Penev L, Mietchen D, Chavan V, Hagedorn G, Remsen D, Smith V, Shotton D (2011). Pensoft Data Publishing Policies and Guidelines for Biodiversity Data. Pensoft Publishers, http://www.pensoft.net/J_FILES/Pensoft_Data_Publishing_Policies_and_Guidelines.pdf
- [26] Handel: <http://www.handle.net/>
- [27] J. Brase, U. Schindler, The Publication of Scientific Data by World Data Centers and the National Library of Science and Technology in Germany, Data Science Journal, Volume 5, 2006, <http://dx.doi.org/10.2481/dsj.5.205>, https://www.jstage.jst.go.jp/article/dsj/5/0/5_0_205/_article.
- [28] Florian Quadt, André Düsterhus, Heinke Höck, Michael Lautenschlager, Andreas V. Hense, Andreas N. Hense, Martin Dames, Atarabi - A Workflow System for the Publication of Environmental Data, Volume 11, 2012. <http://dx.doi.org/10.2481/dsj.012-027>, https://www.jstage.jst.go.jp/article/dsj/11/0/11_012-027/_article.
- [29] Lawrence, B., Jones C., Matthews, B., Pepler, S., Callaghan S., Citation and Peer Review of Data: Moving Towards Formal Data Publication. The International Journal of Digital Curation 6, pp. 4-37, 2011.
- [30] China DOI, Institute of Science and Technology Documentation of China: <http://www.chinadoi.cn>, <http://www.doi.org.cn>.
- [31] International Association for Social Science Information Services & Technology, Quick Guide to Data Citation, <http://www.icpsr.umich.edu/files/ICPSR/enewsletters/iassist2.html>.
- [32] Interoperability of data in e-commerce systems, <http://cordis.europa.eu/econtent/mmrcs/indexs.htm>.
- [33] Copyright Law of People's Republic of China. [中华人民共和国著作权法, 1990年9月7日通过, 2001年10月27日修正, <http://www.lawxp.com/statute/s899627.html>.]
- [34] Regulations for the implementation of the copyright law of People's Republic of China [中华人民共和国著作权法实施条例, 中华人民共和国国务院令第633号, 2013年1月30日, 2013年3月1日起施行, http://www.gov.cn/zwgg/2013-02/08/content_2330132.htm.]
- [35] Berne Convention for the Protection of Literary and Artistic Works, Signed on September 9, 1886, Amended on September 28, 1979. http://www.wipo.int/treaties/en/ShowResults.jsp?lang=en&search_what=N&treaty_id=15
- [36] WIPO Copyright Treaty, 1996, <http://www.wipo.int/treaties/zh/ip/wct/index.html>.
- [37] Universal Copyright Convention, http://portal.unesco.org/en/ev.php-URL_ID=15381&URL_DO=DO_TOPIC&URL_SECTION=201.html.
- [38] The Law on Guarding State Secrets of the people's Republic of China. The third meeting of the Standing Committee of the Seventh National People's Congress on September 5, 1988, amended on April 29, 2010, the eleventh session of the National People's Congress Standing Committee of the fourteenth meeting, http://www.gov.cn/flfg/2010-04/30/content_1596420.htm.
- [39] Declaration of the Budapest Open Access Initiative, 2002. http://en.wikipedia.org/wiki/Budapest_Open_Access_Initiative.
- [40] Bethesda Statement on Open Access Publishing, 2003. <http://legacy.earlham.edu/~peters/fos/bethesda.htm>.
- [41] Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities, 2003. http://wenku.baidu.com/link?url8YRr6L145ixIkIjwvCmJE8bJrGfwOF0pQqPGc4dpORaTKs2dW80f0YzZwXHyvqIDyI0_xeDPG94awSfRwwdajHvHzIIIOAus200SnxQ0Ni.
- [42] Kronberg Declaration on the Future of Knowledge Acquisition and Sharing, UNESCO, June 2007. <http://blog.ci123.com/weiyu/entry/184940>.
- [43] GEOSS Data Sharing Principles, 2009. http://www.earthobservations.org/geoss_dsp.shtml.

Global change research data publishing and repository

LIU Chuang

(Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China)

Abstract: Data sharing is the critical issue of the global change studies. Since 1994, after suffering the data for charge for more than 15 years, China explored the national strategy to enhance the data sharing. One of the most important initiatives is the National Scientific Data Sharing Program (SDSP), later called the National Scientific Data Sharing Infrastructure of China. Although both the decision makers and data management teams are all try hard to make the datasets available for the global change studies, most of the datasets conducted from the public funded research programs are still not available to be shared, which are out of the circle of the National Scientific Data Sharing Infrastructure of China. The most critical reason is the data authors' contribution to the science are not recognized and rewarded in the scientific communities as the data contributor expected.

The research data publishing is a solution for encouraging the public funded data openly available, because it not only prevent the intellectual property of the data contributors, but also make the data benefit the society all. The Digital Object Identifier (DOI) is a useful mechanism, standard and technology to make the data publishing happen. The Global Change Research Data Publishing and Repository was established in the joint effort from the Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, Geographical Society of China, the National Remote Sensing Center of China, Ministry of Science and Technology of China, cooperated with the Committee on Data for Scientific and Technology Data (CODATA) Task Group on Preservation of and Open Access to Scientific and Technology Data in Developing Countries (PASTD), China Secretariat for the Group of Earth Observation (GEO), National Scientific Data Sharing Infrastructure of Earth System Sciences of China and the Digital LIN Chao Geomuseum, as well as the Jomo Kenyatta University of Agriculture and Technology (JKUAT), Kenya. The data publishing procedures, standards and policies have been identified. 20 datasets related to the global change studies have been published in the first issue. Not only the metadata, but the data papers and the data products are also published via peer review. The data papers are published via the *Acta Geographica Sinica*. All of the data including the metadata, data papers and data products all can be free accessed at www.geodoi.ac.cn. Two different services are provided through the system. Of which, one is for the end users. The principles for serving the end users are (1) openly available, (2) free download for no commercial uses; (3) timely to put the data online as soon as the data peer review completed; (4) making the data informative via data paper methodology; (5) the datasets are easy to be accessed. The second service is for the data authors or contributors. The statistics of visitors, download numbers, sizes of the data downloaded, and citations are all calculated though the system in time.

Keywords: Global change; Research data; Data paper; Data publication; Data sharing services.