

# 推进科学数据与信息资源管理共享的思路与对策

苏靖<sup>1</sup> 石蕾<sup>1</sup> 王正<sup>1</sup> 闫小克<sup>2</sup>

(1. 国家科技基础条件平台中心, 北京 100862; 2. 中国计量科学研究院, 北京 100029)

**摘要:** 科学数据和信息是重要的科技资源, 是科技创新活动的重要基础和工具。科学数据和信息是国家科技基础条件平台建设体系中非常重要的一类资源, 加强科学数据和信息资源管理与共享研究, 是推进科技平台建设的重要理论基础。本文在对比分析国外科学数据和信息资源管理与共享现状的基础上, 系统总结了我国在科学数据和信息管理利用方面取得的主要工作成效, 提出完善制度体系和管理机制、推动数据中心建设、加强数据和信息全生命周期管理等意见和建议。

**关键词:** 科技基础条件平台; 科技资源管理; 科学数据; 信息管理; 信息共享; 信息利用

中图分类号: G311

文献标识码: A

DOI: 10.3772/j.issn.1674-1544.2015.05.008

## Thoughts and Countermeasure to Promote Management and Sharing of Scientific Data and Information Resources

Su Jing<sup>1</sup>, Shi Lei<sup>1</sup>, Wang Zheng<sup>1</sup>, Yan Xiaoke<sup>2</sup>

(1. National Science and Technology Infrastructure Center, Beijing 100862; 2. National Institute of Metrology, China, Beijing 100029)

**Abstract:** The scientific data and information are not only important science and technology resources(S&T), but also important basis and instruments for science S&T innovation. Scientific data and information are one of very important resources for platform construction systems in the national S&T infrastructure condition, and to strengthen scientific data and information resources management and sharing is an important theoretical basis for promoting the construction of S&T platform. This paper is based on the comparison of management and sharing with foreign scientific data and information resources, systematically summarizes the main achievements in management and use of scientific data and information in China. Based on the overseas advanced experience, further perfection of management policy and mechanism of scientific data and information management and sharing are proposed. Additionally, some opinions and suggestions are made to promote the construction of data center and strengthen management of data and information lifecycle.

**Keywords:** science and technology infrastructure, science and technology resource management, scientific data and information management, scientific data and information sharing, scientific data and information utilization

**作者简介:** 苏靖(1968-), 男, 国家科技基础条件平台中心副主任, 研究员, 研究方向: 科技管理、科技资源管理; 石蕾\*(1982-), 女, 国家科技基础条件平台中心副研究员, 研究方向: 科技管理、科技资源管理; 王正(1986-), 男, 国家科技基础条件平台中心助理研究员, 研究方向: 科技管理、科技资源管理; 闫小克(1974-), 男, 中国计量科学研究院研究员, 研究方向: 计量。

**收稿时间:** 2015年6月24日。

科学数据与信息资源主要是指各类科技活动产生的基本科学技术数据和资料、各种数据分析产品以及各类文献信息等,是最基本、最活跃的科技资源,既是科技创新活动的重要产出,也是新一轮创新活动和经济社会发展的重要基础和工具。在大数据时代,科学研究、政府决策、产业发展更加依赖于科学数据,可靠、系统、丰富的科学数据和信息资源已经成为一种新的生产要素,成为提高生产力和竞争力的强大引擎<sup>[1]</sup>。随着科技体制改革和科技计划管理改革的不断深入,进一步明确了科技资源管理与共享工作的定位和发展方向。科学数据和信息资源作为一类重要的科技资源,根据其资源属性特征,形成有针对性的管理政策和制度,并建立相应的开放方式、评价方法和支持引导措施,对于推动资源共享利用,支撑科技创新具有重要意义<sup>[2]</sup>。本文在总结我国科学数据和信息资源发展现状的基础上,综合借鉴发达国家科学数据和信息资源管理先进经验,结合当前国家科技基础条件平台工作新形势和新要求,提出进一步推进我国科学数据和信息资源管理与利用的思考与建议。

## 1 我国科学数据和信息资源管理工作取得的积极成效

### 1.1 大力推进国家科技基础条件平台建设,夯实科学数据与信息资源管理与共享服务基础

我国对科学数据与信息资源进行系统整合与共享的工作起始于2001年科学数据共享工程。2004年纳入国家科技基础条件平台建设专项统筹推进,累计支持14个科学数据共享平台建设,整合了农业、气象、地震、人口健康、材料、能源、地质等10个技术领域32大类科技资源数据库共计5万余个,数据总量超过700TB,构建了由主体数据库、科学数据中心或数据网、门户网站构成的三级结构的数据管理与共享服务体系。同时,研究实验基地类共享平台如国家生态系统观测研究网络平台也开展了大量科学数据共享服务,累计整合各类观测、监测和研究数据超过4000GB。此外,通过平台建设也整合了大量的

科技文献信息和科普信息,如中国数字科技馆整合数字科普资源超过9TB。通过多年努力,初步建立了科学数据与科技信息资源共享管理模式和服务机制,制定了一系列数据信息标准规范,树立了一批在领域内知名度较高的数据信息共享服务品牌,形成了一批专业化的科学数据和信息资源管理与共享服务机构。

### 1.2 推进国家科技基础条件平台绩效考核,提升科学数据与信息资源共享服务能力

自从2011年国家科技基础条件平台开展绩效考核与奖励补助以来,对各共享平台数据信息资源建设、质量管理、服务能力等的规范化管理工作不断加强,各共享平台资源服务能力稳步提升,数据信息服务效果显著增强。2014年,地球系统科学数据、气象科学数据等6个科学数据共享平台数据共享服务量继续稳步增长,平台网站访问量超过5000万次,同比增长近40%;科学数据和资源信息服务数量超过160TB,同比增长超过30%;服务国家重大科技专项、国家重大工程项目(课题)以及各级各类科技计划项目(课题)近3000项,同比增长近70%;支撑发表论文3000余篇,同比增长超过10%;在科技创新和公共服务供给方面发挥了重要作用。科技文献类共享平台2014年网站访问量超过1亿人次,文献服务量超过130万篇,同比增长均超过30%,对于科技创新和经济发展的支撑保障能力进一步增强。

同时,各科学数据和信息资源共享平台聚焦重大需求和科技热点,组织开展了多项综合性、系统性、知识化的多平台联合专题服务。例如:在大气污染防治、远程医疗、水土保持、科技救灾、传染病预测预警、材料腐蚀、科学普及等方面开展了多项联合专题服务,形成了一系列科学数据 and 信息服务产品,推动了科学数据与信息资源的深度挖掘与综合集成,有效释放了科学数据与信息资源共享内生动力。

### 1.3 多渠道开展科学数据和信息资源整合,促进科技基础条件资源持续更新

各科学数据和信息资源共享平台根据科技资

源需求和领域发展推进数据持续整合，将优质科学数据和信息资源不断补充更新到国家科技基础条件平台体系。2014年科学数据和信息资源类共享平台累计新增科学数据资源约13TB，科学数据和信息资源进一步集聚。与此同时，为进一步完善科学数据和信息资源更新机制，拓展资源整合渠道，科技部于2011年启动实施了国家科技计划项目科技资源的汇交与共享工作。目前，已累计整合了国家科技计划项目所形成的科学数据库（集）1万余个，各类科技资源信息共计18万余项，审核通过的资源已向社会开放共享，并将纳入各平台开展数据共享服务。另外，通过国家重点基础研究发展计划（973计划）资源环境领域项目数据汇交和科技基础性工作专项数据汇交等工作，也已整合各类型数据资源超过2TB，数据信息整合规模继续扩大和服务能力持续提高。

在我国科学数据和信息资源管理与共享工作取得显著进展、成效的同时，我们也面临着一些管理和技术方面的问题和不足：一是科学数据和信息资源管理与开放共享的制度政策缺失，对科学数据和信息资源共享缺乏有效的知识产权保护；二是针对数据信息从生产、保存到使用的全生命周期缺少统一规划和系统管理；三是海量异构数据库的综合集成、合理存储、快速处理、有效分析和深度挖掘能力不足；四是对科学数据库建设缺乏持续更新，导致很多数据库逐渐变为“死库”，对粮食安全、水安全、能源安全等战略性科学数据储备不足。

## 2 发达国家科学数据和信息资源管理共享的主要做法

长期以来，发达国家一直非常注重科学数据和信息资源的保存、开放和利用。特别是对国家财政投入支持产生的科学数据和信息资源进行了系统管理和规范利用，综合分析其主要做法，突出表现为政策措施保障健全、生命周期管理严密、数据中心建设规范、知识产权保护得力等特点。

### 2.1 多层次政策措施保障科学数据和信息资源的管理与开放

发达国家在推进数据与信息开放过程中普遍有多个层面的法律法规予以支持和保障，尤其是对于国有公益性数据信息的管理和共享提出了要求，对于私营企业投入产生的数据多以市场化机制进行管理。以美国为例：作为现有数据政策相对完备的国家之一，美国通过多次修订《信息自由法》，逐步构建了国家信息公开和数据资源共享的制度框架<sup>[3]</sup>，将“完全与开放”的科学数据共享政策作为信息时代的一项基本国策，由联邦政府负责统筹和规划科学数据管理工作<sup>[4]</sup>。同时，在联邦政府的统筹推进下，美国许多联邦机构和组织也制定了相关的数据管理政策，如美国国家科学基金会、美国国立卫生研究院、能源部、教育部、环保部等部门都对本部门支持产生科学数据的管理与开放提出了明确要求，例如：制定数据管理计划、提交指定数据中心保存、规定数据保存年限、对项目数据生产进行专门资助等。

### 2.2 全生命周期的管理机制保障科学数据和信息资源可用、易用和可追溯

科技资源具有形成、成长、成熟、衰亡的生命过程<sup>[5]</sup>。美国、英国等国家通常都按照科学数据生命周期进行全链条、系统化管理，也取得了突出的进展和成效。科学数据和信息资源的全生命周期通常包括：生产、处理、分析、保存、访问及重新使用6个阶段。典型生命周期模型如图1所示。具体来说，生命周期的数据生产阶段主要包括：制定数据生产和共享计划、生成元数据等；数据处理阶段主要包括：数据录入、转录及翻译、检查、验证及清理等；数据分析阶段主要包括：解释数据、产生研究成果、准备数据保存等；数据保存阶段主要包括：数据转移为最佳格式、数据转移的最佳媒介、数据备份及保存、制作数据档案等；数据访问阶段主要包括：数据分类及共享、数据访问、数据版权、数据推广等；数据重新使用阶段主要包括：在后续研究、研究评述、审查成果中重新利用数据。

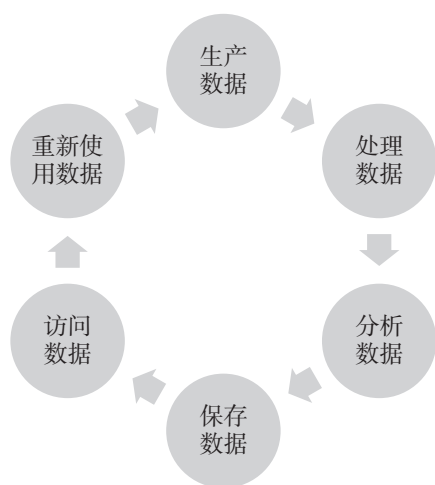


图1 科学数据典型生命周期模型

### 2.3 数据中心是科学数据和信息资源管理的一类重要载体

国外数据中心是实现科学数据和信息资源汇集、管理、存储和共享的专业化机构，美国、英国、澳大利亚等国家都建立了人员齐备、运行机制健全、运行服务规范的国家数据中心。总体来看，国外数据中心在数据管理方面重点开展以下工作<sup>[6]</sup>：(1)开展数据评价；(2)转换数据格式；(3)安全保存数据；(4)定期备份数据；(5)提供数据在线查询；(6)控制数据访问；(7)进行数据标识和认证；(8)建立标准化的数据引用机制；(9)促进数据使用。此外，值得注意的是，为满足科研基金组织要求项目申请者递交数据管理计划的要求，国外一些数据中心还提供帮助科研人员制定专业数据管理计划的专业软件工具或类似的服务。

目前，世界知名的科学数据中心如美国海洋大气局（NOAA）、美国国家空间科学数据中心（NSSDC）、英国数据保存中心（DCC）、英国数据档案中心（UK Data Archive）以及澳大利亚国家数据服务中心（ANDS）等，基本上都针对数据的全生命周期进行了规范化的运行和管理。在2015年由数据创新中心发布的评述八国集团（G8）数据开放宪章进程的报告中<sup>[7]</sup>，英国在国家数据开放程度方面得分最高，除了其在科学数据管理政策、制度保证非常完善外，分领域建设和

布局专业化的科学数据中心也是推动其数据和信息公开共享的重要因素。

### 2.4 处理好数据公开与保密的关系，保障科学数据和信息资源的有效管理与开放共享

处理好科学数据和信息资源公开与保密的关系，是实现数据信息资源整合、开放、共享和利用的根本性问题。各发达国家和国际数据组织也一直在研究和探讨数据安全与保密等相关问题。目前，各国在推进科学数据和信息资源管理及开放共享过程中主要依靠明确数据所有者、确定数据使用许可协议、细化数据安全等级等方法解决数据信息资源公开与保密中的矛盾问题。例如：美国数据管理机制主要包括保密性管理机制、“完全与开放”管理机制和市场管理机制三类<sup>[8]</sup>，除危及国家安全、影响政府政务和涉及个人隐私采用保密机制外，原则上由政府投资产生的科学数据和信息资源都应采取完全开放与共享的管理机制向全社会开放，由企业自己投入而产生的科学数据和信息资源可以采取市场机制保护投资者权益。

与此同时，随着对于科学数据和信息资源公开与保密问题研究的不断深入，世界各国逐渐形成共识：科学数据和信息资源标识是保证数据和信息产生者的利益的有效手段。因此，多个发达国家政府和知名数据中心都在积极推广数据标识和引用相关工作。如澳大利亚在全社会大力倡导科学数据和信息资源标识，并建立了相应的科学数据和信息资源引用评价及奖励机制，确保科学数据和信息资源完全开放与共享，支撑国家的创新战略实施。

## 3 深入推进我国科学数据和信息资源管理的思考与建议

面对新形势和新要求，深入推进科学数据与科技信息管理与共享是提升科技创新能力，落实创新驱动发展战略，支撑科技创新引领经济发展新常态的重要举措。要进一步加强科学数据和信息管理，需要在继续健全政策制度和管理机制的同时，统筹考虑科学数据和信息资源的全链条管

理，做好科学数据与信息资源的采集、加工、挖掘、利用和分级分类，支持科学数据与信息资源的持续更新与积累。围绕以下几个方面推进相关工作。

### 3.1 完善科学数据和信息资源管理与开放共享制度体系，推进数据和信息公开与共享

要深入推进科学数据和信息资源的公开与共享，根本性问题是破除体制机制障碍。要研究制定国家科学数据共享管理条例，特别是要对国家财政支持产生的科学数据和信息资源的整合集成与开放共享提出明确要求，建立科学数据和信息资源开放获取政策，解决科学数据共享过程中存在的安全保密、知识产权保护、资源信息化等问题。推动行业部门建立科学数据共享行业规章制度，以破除行业数据资源信息孤岛为突破口，彻底打破科学数据共享壁垒。加强机构合作，逐步推动公共领域的基础性、公益性科学数据对外开放。推动科学数据共享与市场主体有机结合，促进技术创新数据公开，推动技术集成创新和产业模式创新。

### 3.2 实施数据标识与引用标注，健全科学数据和信息资源管理与开放共享评估监督体系

建立数据标识与引用标注制度是保护知识产权，加强科学数据和信息资源管理，推动数据开放共享的有效手段。我国也应该建立统一的科学数据标识和引用标准制度，规范科学数据的注册与引用标准，探索建立我国科学数据出版机制。在数据标识和引用标注制度的基础上，针对科学数据和信息资源的多学科性和多类型性，制定分级分类考核标准，建立健全科学数据和信息资源管理与开放共享评估监督体系，构架基于数据全生命周期的科学数据评估监督体系。同时，以科学数据评价为基础，推动科学数据建设者和共享服务人员评价制度改革，建立激励机制，保障科技资源生产者、服务者利益，保证人才队伍稳定及人才自身发展，激发创新活力。

### 3.3 加快科学数据和信息资源集聚，打造数据中心

以科学数据和信息类科技基础条件平台为基础，突出重大问题和需求导向，紧密衔接重大

科技创新活动，加快资源集聚，集中力量打造一批具有领域、行业优势的权威性科学数据中心。遴选我国学科发展和科学数据资源优势学科，借鉴国外先进经验，探索推进具有国际水平的科学数据中心建设。围绕我国经济社会发展的重大战略需要，以及粮食安全、水安全、能源安全等问题，部署储备战略性科学数据中心建设，分级分类分时限开放共享。围绕一带一路、京津冀一体化、长江经济带等国家、区域发展战略部署，设立区域科学数据和信息资源服务中心，集聚优势资源开展数据支撑服务。

### 3.4 加强科学数据和信息资源全生命周期管理，实现数据持续积累与更新

建立重点科学数据库建设长效机制，支持重点领域大型科学数据库、信息库的数据采集、整理与保藏，确保重要数据和信息长期保存与持续更新。结合中央财政科技计划（专项、基金等）管理改革工作部署，结合各类科技计划（专项、基金）布局，聚焦国家重大战略任务，支持科学数据和信息资源采集与保存，推动数据共享和利用。继续完善科学数据汇交标准和 workflows，推动国家科技计划项目、国家科技重大专项、行业专项等项目形成的科学数据统一汇交，将科学数据汇交纳入项目管理流程。探索多元化经费支持方式，调动社会力量建立数据库建设基金，鼓励更多科研单位、科学家、科研工作者开展数据库建设并向权威数据库集成。

### 3.5 完善科学数据与信息资源管理技术与标准，探索基于数据的科研新模式

制定并出台科学数据和信息资源管理与共享技术标准，对科学数据和科技信息资源的生产、发布、存储、使用进行规范化管理，破除数据共享技术障碍。研制开发一批数据采集、加工、管理、挖掘、共享的共性技术软件，支撑科技数据和信息资源管理与共享。针对需求组织开展多领域跨库数据整合与集成，实现数据互联互通，形成标准化的数据产品生产流程和服务管理模式，构造基于数据的、开放协同的研究与创新模式。

（下转第 55 页）

所以这些维度和指标还没有得到明确的构建。从这种意义上来说, 后续的研究还需要对评价要素中的维度和指标开展进一步的分析, 以期最终构建出一套准确、科学、易于操作且符合资源共享实际的利用效果评价指标体系。

#### 参考文献

- [1] 高峰, 王剑. 知识管理在项目管理中的应用研究——以农业科学数据共享中心项目为例[J]. 科技管理研究, 2012(16):184-187.
- [2] 朱艳华, 黎建辉, 胡良霖. 科技数据应用效果评估模型初探[J]. 中国科技资源导刊, 2014, 46(2): 7-14.
- [3] 宋立荣, 孟宪学, 周国民. 我国农业科学数据共享中信息质量管理的措施与建议[J]. 中国农业科技导报, 2009, 11(6):37-42.
- [4] 贺丹. 智力密集型城市科技创新资源利用效率分析[D]. 武汉: 武汉理工大学, 2008.
- [5] 蒋仁爱, 高昌林, 冯根福. 中国研究机构的科技资源利用效率评价研究[J]. 统计研究, 2013, 30(6):19-29.
- [6] 罗序斌, 陈运平. 我国地方普通高校资源利用效率的DEA测度及其管理对策[J]. 南昌大学学报, 2012, 43(4):155-160.

- [7] 王剑, 高峰. 数字开放环境下科技资源你共享服务评价模式研究[J]. 现代情报, 2014, 34(10):31-34.
- [8] 肖鹏, 胡许萍, 孙玉红. 科技型小微企业科技资源利用效率评价[J]. 科技进步与对策, 2013, 30(22):133-136.
- [9] 孙建军. 网络公共信息资源利用效率影响因素实证分析[J]. 图书情报工作, 2012, 56(10):35-40.
- [10] 周琼琼, 冯楚建. 我国科技信息资源影响力评价体系设计与研究[J]. 中国科技资源导刊, 2013, 45(5):1-5.
- [11] 杨行, 彭洁, 赵伟. 科学数据共享网站资源可见性评价框架构建[J]. 中国科技资源导刊, 2014, 46(6):101-107.
- [12] 孙九林. 科学数据资源与共享[J]. 中国基础科学, 2003(1):30-33.
- [13] 黄鼎成, 郭增艳. 科学数据共享管理研究[M]. 北京: 中国科学技术出版社, 2002.
- [14] 杨从科. 中国农业数据资源建设研究[D]. 北京: 中国农业科学院, 2007.
- [15] 王剑, 王健, 赵华, 等. 农业科学数据共享中心服务效果的评价方法研究[J]. 中国科技资源导刊, 2014, 46(5):84-89.
- [16] 师荣华, 刘细文. 基于数据生命周期的图书馆科学数据服务研究[J]. 图书情报工作, 2011, 55(1):39-42.

(上接第49页)

#### 4 结语

加强科学数据和信息资源管理, 推进资源开放共享和综合利用, 既是深化科技体制改革的重要任务, 也是大数据时代科技管理工作面临的重要机遇和挑战。长期以来, 我国在推进科学数据和信息资源管理和共享方面已经开展了大量工作并取得了积极成效, 综合借鉴发达国家在科学数据和信息资源共享方面的典型成功案例和主要做法, 我国要进一步深化科学数据和信息资源的管理与共享, 必须继续深化科技平台工作, 从完善制度体系、加强监督评估、打造数据中心、保障数据更新以及优化技术标准等方面多管齐下。

#### 参考文献

- [1] 曹凌. 大数据创新: 欧盟开放数据战略研究[J]. 情报理

论与实践, 2013, 36(4):118-122.

- [2] 叶玉江. 加强科技平台工作推进科技资源管理[J]. 中国科技资源导刊, 2015, 47(2):1-6.
- [3] 罗辉. 美国关于科技资源共享的法律和法规[J]. 全球科技经济瞭望, 2011, 26(4):31-36.
- [4] 陈传夫, 曾明. 科学数据完全与公开获取政策及其借鉴意义[J]. 图书馆论坛, 2006, 26(2):1-5
- [5] Daniel Castro, Travis Korte. Open Data in the G8: A Review of Progress on the Open Data Charter[EB/OL]. [2015-04-25]. <http://www.datainnovation.org/2015/03/open-data-in-the-g8/>.
- [6] 陈大庆. 英国科研资助机构的数据管理与共享政策调查及启示[J]. 图书情报工作, 2013, 57(8):5-11.
- [7] 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):34-39.
- [8] 贺威, 刘伟榕. 大数据时代的科研革新[J]. 未来与发展, 2014(2):1-4.