高校
科学数据管理与实证研究

GAOXIAO KEXUE SHUJU GUANLI YU SHIZHENG YANJIU

刘　霞　主编

武汉理工大学出版社
WUTP Wuhan University of Technology Press

高校科学数据管理与实证研究

主　编：刘　霞
副主编：饶　艳　刘　颖
谢春枝　洪正国

武汉理工大学出版社
·武汉·

内 容 简 介

本书以高校科学数据管理为研究对象，系统研究了国内外科学数据管理的最新进展及发展趋势。通过广泛调研，掌握了我国高校科学数据及科学数据管理的现状和特点，在此基础上，提出了适用于高校科学数据管理的标准规范体系、平台选择策略，并以武汉大学为例，从自然科学和社会科学各选择一个案例，详细介绍了从需求调研、元数据设计到平台构建的全部过程。最后，结合理论研究和实证经验，提出了符合我国国情的高校科学数据管理方案。

本书内容全面，可操作性强，既有理论研究，又有需求调研，还有实证案例。因此，本书一方面可用于指导高校开展科学数据管理的实践活动；另一方面则可作为本科生、研究生科学数据管理素养教育的教材。

图书在版编目(CIP)数据

高校科学数据管理与实证研究/刘霞主编. —武汉：武汉理工大学出版社，2013.9
ISBN 978-7-5629-4208-5

Ⅰ. ①高… Ⅱ. ①刘… Ⅲ. ①高等学校-自然科学-数据管理-研究 Ⅳ. ①N37

中国版本图书馆 CIP 数据核字(2013)第 232839 号

项目负责人：史卫国　　**责 任 编 辑**：史卫国
责 任 校 对：余士龙　　**封 面 设 计**：兴和设计
出 版 发 行：武汉理工大学出版社
地　　址：武汉市洪山区珞狮路 122 号
邮　　编：430070
网　　址：http://www.techbook.com.cn
经 销 者：各地新华书店
印 刷 者：武汉兴和彩色印务有限公司
开　　本：787×960　1/16
印　　张：19.75
字　　数：400 千字
版　　次：2013 年 9 月第 1 版
印　　次：2013 年 9 月第 1 次印刷
印　　数：1—1000
定　　价：58.00 元

凡购本书，如有缺页、倒页、脱页等印装质量问题，请向出版社发行部调换。
本社购书热线电话：027-87515778　87515848　87785758　87165708(传真)

序

人类在生产实践、科学实验和社会活动的过程中产生了大量的数据，这些数据中包含着丰富的信息和知识。数据经过整理和分析，形成科学的结论，指导人们的生产实践、科学实验和社会生活，在这些过程中又产生新的数据，由此周而复始。这些数据是人类宝贵的财富。随着科学技术和社会经济、政治的发展，数据的数量日趋庞大，如无科学的收集、组织和管理方法，数据将变为一堆杂乱无章的信息，从而丧失了数据的作用。在这种形势下，科学数据管理的概念应运而生。

科学数据是一个国家的战略资源，普遍受到各国的重视。发达国家较早注意到这个问题，一些国家已经形成了国家层面的科学数据管理法规，学术机构也建立了科学数据收集、管理和发布的制度。特别是学术机构之间的数据共享，更得到充分的重视，在相互交换数据、数据合作分析、建立公共数据档案等方面取得积极的进展，大大提高了科学研究工作的质量和效率。

我国对科学数据管理关注较晚，但已经有了可喜的进展。国家科技部的一些重大项目都有对研究项目所产生数据的管理规定。中国科学院对科学数据的管理走在全国的前列，已经有了全院范围内的关于研究数据的收集、管理和发布的制度。

高等学校学者云集，学术氛围活跃，历来是产生科学数据的大户，也是科学数据管理研究和实践最为活跃的地方之一。国内高等学校同样是科学研究的重要阵地，承担了大量的国家资助项目和企事业单位的委托项目。这些项目所产生的数据，既是评价项目执行情况的基本依据之一，也是开展科研创新的基础。但是，从我国高校的科学数据管理的整体情况看，目前还处于起步阶段。在政府层面，国家教育部和各省教育主管部门还没有对高校颁布关于科学数据管理的法规；在学校层面，鲜有高校实行对校内科研数据的有效管理；在教师和研究人员个人层面，对于呈交科研数据、实现数据资源共享也存在不少疑虑，如：整理、保存和上交数据需要耗费额外的时间和精力，担心个人或团体的知识产权得不到保护等，这同样是在高校推行科学数据管理过程中存在的障碍。如何构建有效的管理机制，充分利用已有的数据收集、组织及开发利用相关成果，实现高校科学数据的汇交和共享，是一个值得认真思考和积极探索的问题。

科学数据管理并不在图书馆传统的工作领域之内。在网络化、数字化环境下，由

于图书馆在处理各类型文献信息方面的经验和实际工作成效使其在承担科学数据管理方面的能力逐渐得到社会的认可。2007 年，美国国家科学基金会(NSF)启动了 DataNet 计划，明确提出以图书馆为主体，实施科学数据管理。美国一些高校图书馆开始调整其机构和岗位设置，或设立专门的数据管理部门，或培训和招聘数据管理专员，以帮助科研人员有效地完成其数据管理计划。

2011 年，为了推动高校科学数据管理的发展，“211 工程”中国高等教育文献保障体系(CALIS)三期设立了预研性项目“中国高校科学数据管理与服务机制和平台的研究”，由武汉大学图书馆承担。项目组对国内外文献和相关网站进行了详细调研，通过问卷调查和走访了解高校科研人员数据生产和管理需求现状，在武汉大学确定了若干试点院系，搭建了实验性数据共享平台，并开始面向全校提供数据存储和共享服务。本书既是对这一研究过程和实践成果的总结，也是为推动高校科学数据管理发展所做的宣传。

实现对科学数据的科学管理，需要解决三个方面的问题，即：机制、需求、技术。

科学数据管理涉及方方面面，唯有进行顶层设计、政府行政推动，形成国家政策和制度，构建具有强制性和可操作性的运行机制，才有可能取得全面的、可持续的成效。这虽然是目前最难解决的问题，但是已经有了国外比较成型的制度法规做借鉴，以及国内某些政府部门和高校、科研单位的实践经验，已经具备了推进政府层面进行制度设计的基本条件和基础。

需求问题是解决促使研究人员参与科学数据管理的驱动力和积极性的问题。科学研究的数据掌握在研究人员个人或研究团队手里，能否及时整理和完整地汇交数据，完全取决于研究人员的自觉性和共享意识。推行科学数据管理的意义很大程度上是要实现数据共享，这将极大地提高研究效率，节省社会资源。而由于传统思维方式的惯性，我们的研究人员往往习惯于将研究数据私有化，对于社会化大生产条件下的共享理念缺乏足够的认识。因此，在有了可行的制度之后，还需要对研究人员做大量的工作，使他们认识到科学数据管理的社会化，不仅能够节省他们自己用于管理数据的时间，而且可以更多地利用他人的研究成果，有利于提升科学研究的质和量，使接受科学数据管理成为研究人员自己的客观需求，自觉地做科学数据管理和共享的践行者。

技术问题是指要解决数据汇集、管理、发布和保存的平台问题。在目前的技术环境下，已经有了很多科学数据管理平台的成功案例，我们也有了构建、管理和运行诸如 CALIS、CADAL 这样一些全国范围的信息资源共享平台的经验。相对而言，在三个问题中技术应该是最容易解决的问题了。

科学数据管理并不是一个全新的研究领域，但实施起来却非常困难，需要政府、机构和个人协调运作才有可能取得成效，而且还需要各方面的力量持续地推动。我

们应该做的一项工作就是利用各种条件，积极开展科学数据管理的研究和实践，可以是宏观层面的，也可以是某一方面的或某一学科的，积累新鲜经验，为决策层提供决策依据，为科学数据共享制造舆论环境和实验平台。水滴石穿，集腋成裘，水到渠成。本书就是在做这样的工作，是值得鼓励的。

2013 年 8 月

目　录

1 科学数据管理概述

近年来，科学数据管理与共享得到学术研究机构和政府部门的广泛重视。美国、英国、澳大利亚等发达国家制定了相关法规和政策，启动了各种科学数据共享项目。高校作为科学研究的重要阵地之一，其科学活动所产生的研究数据，虽然在数量级别上远低于某些专门研究机构，但在学科覆盖面和零散程度方面则更为复杂。与发达国家相比，目前国内高校在科研项目数据管理规范化方面基础薄弱，而科学数据的管理和共享将更为复杂和困难。因此，只有认真研究高校科学数据的特点和规律，探索构建适合我国国情的高校科学数据管理机制，实现科研项目数据的汇交和共享，才能有效地提升科学数据的价值，加速科研进程并最终促进科学的交流和创新。

1.1 科学数据与科学数据管理

1.1.1 科学数据

1. 科学数据的概念

何为科学数据？国外的名称主要为 Research Data，即研究数据；也有 Scientific Data，科学数据。国内的名称则经由科技数据[①]演化为科学数据。关于科学数据的定义，国内外学者主要有以下一些观点：

我国科学数据共享工程中对科学数据的定义如下：科学数据，是指人类在认识世界、改造世界的科技活动中所产生的原始性、基础性数据，以及按照不同需求系统加工的数据产品和相关信息。它即包括了社会公益性事业部门所开展的大规模观测、探测、调查、实验和综合分析所获得的长期积累与整编的海量数据，也包括国家科技计划项目实施与科技工作者长年累月科学实践所产生的大量数据。科学数据具有分离性、驾驭性、共享性、客观性、长效性、积累性、公益性、非排他性、不对称性、增值性、可传递性和资源性等特点[②]。

---

① 孙九林，等. 我国科技数据管理和共享服务的新进展[J]. 世界科技研究与发展，2002，24(5)：15-19.

② 科学数据共享概念与术语[EB/OL]. [2012-03-11]. http://www.sciencedata.cn/pdf/2.pdf.

陈传夫认为科学数据是指各类科技活动产生的原始性、基础性数据及其分析研究信息，是国家创新体系中最活跃的要素之一①。

王学勤认为科学数据包括科研论文、专利、研究报告、实验观测数据和元数据、参考资料、照片和图表、学术类多媒体资源等，不仅包括公开出版和可公开获取的数据，还包括很多的灰色科学数据②。

美国自然科学基金（National Science Foundation，NSF）在2005发布的“长期保存的数字数据集”中指出，数据指的是任何可以数字形式存储的信息，包括文本、数字、图片、视频或电影、音频、软件、算法、方程式、动画制作、模型、模拟等。这些数据通过不同方法产生，如观察、实验、计算③。

澳大利亚国立大学（Australia National University，ANU）将科研数据（Research Data）定义为数字形式的研究数据，指产生于研究过程中并能存贮在计算机上的任何数据，也包括能转换成数字形式的非数字形式数据。如传感器读取的数据、遥感勘测数据、神经图像、实验数据、调研结果及来自测试模型的仿真数据等④。

经济合作与发展组织（Organization for Economic Co-operation and Development，OECD）在其存取公共基金资助研究数据的原则与建议中，将“研究数据(Research Data)”定义为：作为科学研究基本来源的事实记录（数值、文本记录、图像和声音），被科学团体所共同接受的对研究结果有用的数据。但不包括这些内容：实验室笔记、初步分析、科学论文的草稿、未来的研究计划、同行评论以及个人和同行的交流、实物（如实验样本、细菌和测试的动物）等。另外还强调数据为数字化的计算机可读的科学数据⑤。

美国国立卫生研究院（National Institutesof Health，NIH）对“研究数据”的定义为：记录事实材料，受到科学界普遍认同的能对研究成果进行验证的必要材料。它不包括：初步分析、科技论文的草稿、未来的研究计划、同行评审、与同行

---

① 陈传夫. 中国科学数据公共获取机制：特点、障碍与优化的建议[J]. 中国软科学，2004(2)：8-13.

② 王学勤，等. 建立数据驱动的 e-Science 图书馆服务：机遇和挑战[J]. 图书情报工作，2011(7)：80-83.

③ NSF. NSB-05-40 Long-lived digital data collections：enabling research and educating in the 21st century[EB/OL]. [2012-3-26]. http://www.nsf.gov/pubs/2005/nsb0540/nsb0540.pdf.

④ ANU data management manual：Managing digital research data at the Australian National University [EB/OL]. [2013-02-20]. http://regnet.anu.edu.au/sites/default/files/files/ANU_Data_Management_Manual.pdf.

⑤ OECD. OECD principles and guidelines for access to research data from public funding [EB/OL]. [2012-02-20]. http://www.oecd.org/dataoecd/9/61/38500813.pdf.

的通讯记录；物理对象（例如，实验室样品、音频或视频磁带）、商业秘密、商业信息、保密材料；在法律保护之下的信息（例如，知识产权）、人事与医疗文件和类似的文件；可能造成侵犯个人隐私的信息或者能够被用来识别身份的信息①。

剑桥大学（University of Cambridge）在 Incremental Project 的术语解释中，对研究数据的定义非常广泛，包括任何数字化的输出及输入和研究者在研究过程中使用或创造的陈述语言。如 Excel 电子表格和 SPSS 统计或测量的数据集、用于转换数据的计算机代码、研究员野外记录、用于研究或演示的电子图片、包含学术信息的电子邮件通信记录以及学术论文的草稿等。尽管多数研究数据会以数字化形式存在，但一些实体或论文记录，如实验室笔记，因其是了解数据的部分，也应当作为研究数据②。

还有许多机构或组织对专业领域的科学数据进行了定义。澳大利亚南极数据中心（Australian Antarctic Data Center）对数据管理的对象——南极科学数据的范畴进行了说明：南极科学数据几乎包括在南极地区采集的各种科学数据，从原始数据到处理过的数据，从纸质存贮的数据到电子媒质存贮的数据。具体包括：①原始数据或者未经处理的观测数据；②对原始数据进行校正或者分析处理后形成的数据；③对数据集进行精炼或分析而形成的数据（产品）；④说明科学数据的采集环境与相关背景的数据，或者有助于提高科学数据的价值与作用的说明性数据③。

我国国家农业科学数据共享中心将农业科学数据定义为：从事农业科技活动所产生的基本数据，以及按照不同需求而系统加工整理的数据产品和相关信息④。

我国测绘科学数据共享服务网所收集的测绘科学数据是指在现有测绘数据基础上，主要面向科学家和科学研究的需要，经过对原有测绘数据的整合、加工、集成、保密等技术处理得到的可以直接服务于科学研究，重大科技工程及国民经济、社会发展其他领域的空间定位基础数据⑤。

综合以上的研究成果来看，对科学数据的定义可以分为三种：一种是仅仅将验证过的科学数据作为科学数据管理的对象，而科学研究过程中产生的笔记、初步的实验分析结果等未经过验证的数据、资料不纳入管理对象的范畴，如美国国立卫生研究院的定义；第二种是将与科学研究相关的实验数据、实验笔记、图像、音视

---

① NIH[EB/OL]. [2013-03-17]. http://grants.nih.gov/grants/policy/nihgps_2011/nihgps_ch2.htm.

② University of Cambridge. Explanation of Terms[EB/OL]. [2013-03-27]. http://www.lib.cam.ac.uk/preservation/incremental/glossary.html.

③ 凌晓良，等. 澳大利亚南极科学数据管理综述[J]. 地球科学进展，2007，22(5)：532-539.

④ 农业科学数据共享中心简介[EB/OL]. [2013-01-15]. http://www.agridata.cn/homepage/ch_intro.asp.

⑤ 测绘科学数据共享服务网[EB/OL]. [2013-03-22]. http://sms.webmap.cn/.

频、模拟系统等半成品作为科学数据管理的对象，如剑桥大学的定义；第三种则是含义最为广泛的理解，即将科学研究完成前的所有过程数据、半成品以及科学研究完成后的成果都作为科学数据管理的对象，如王学勤所给的定义。在本书中，引用的是最宽泛的定义，即第三种定义。

2. 数字化科研背景下科学数据的新含义

随着信息技术的推动，科研信息化日趋成熟，科学计算促使人们描述人类社会复杂对象的能力不断提升，各种网络传输系统和数据保存与分析设施不仅帮助科学家获得了“史上最强”的观察能力、分析能力甚至实验能力，更促进科研方式从理论分析和观察向科研对象的模拟和仿真发展，并推动“以数据为基础的科学研究第四范式”的形成。该范式是“数据密集型”科学研究新范式，其概念由 Jim Gray 于 2007 年提出。Jim Gray 认为，人类科学研究的第一范式是最早期的利用经验描述自然现象；第二范式是数百年以前所进行的利用理论模型开展研究；第三范式是数十年以前开始的模拟复杂现象的科学计算研究；所谓第四范式则区别于这些研究方法，指在数字化科研背景下，由假设驱动向直接基于科学数据进行探索的科学方法转变①。科学研究与发现范式的转变赋予科学数据以新的含义，并向科学数据的管理提出新的挑战。数字化科研背景下的科学数据含义具体包括以下各方面：

由探测器等各类高端仪器设备获取、高性能计算机模拟等方式所产生的海量、原生科学数据。这些科学数据不仅仅包含传统条件下由理论预测及实验观测得到的实验观测结果，还包括将科研活动中各方面的因素，如科研人员、科研仪器，甚至科研过程及管理机制进行聚合，以及将科学研究对象（包括社会科学领域的研究对象）通过计算机仿真和模拟分析等方式产生的数字表达等。这种数字表达相对于传统表达的优势在于它能描述大尺度或微尺度的实体，以及按照科研人员的需要，根据科学研究所需进行各种形式的组合、变化和数字表达。

初始数据包括科学实验未经处理的数据及科学研究对象的数字表达。从初始数据到中间数据乃至最终到研究结果的科学工作流（Scientific Workflow）技术，使得科学数据的可视化方法、技术及软件资源也以一种资源的形式存在，人们能以一种可重复、可验证、分布式的方式来描述科学研究或科学实验的过程，了解科学研究过程中所采用的数据处理方法、模型和工具，从而对科学研究进行验证②。

3. 科学数据与科学文献的区别与联系

同为科学研究的结果记录及成果体现，科学数据与科学文献既有着显著的区别，又存在着密不可分的联系，表现为：

---

① Tony Hey, et al. The Fourth Paradigm Data-Intensive Scientific Discovery[EB/OL]. [2013-7-19].

② 钱鹏. 高校科学数据管理研究[D]. 南京：南京大学，2012.

在类型与格式方面，科学文献较为统一而规范，科学数据则较为复杂，既包括实验过程中通过观察获取的记录数据，以及直接通过数字化仪器设备获取的数据，还包括各类文档中包含的数据，以及社会科学研究中的统计数据和问卷调查结果所产生的数据等。

科学文献一般来说是科研活动的总结，同一项研究可以出版侧重点不同的多篇文献。科学数据则不是静态的研究结果，而是需要记录和反映科研活动这一动态过程，同一项研究可能存在初始数据、中间数据以及最终数据等一系列研究数据。

同为科研产出，科学数据与科学文献相互关联，其表现为：科学文献以科学数据为研究对象进行研究，通过对科学实验数据产生过程及产生结果进行描述并得出科研结论从而产出科学文献。在科学数据管理中则需建立两者的关联关系，一方面通过外部相似性建立交叉引用关系，另一方面在内容层面直至知识层面实现两者的聚合和知识关联，使科学数据和科学文献紧密整合共同服务于科学研究①。一些传统科学文献出版社非常重视科学数据与科学文献的关联服务。例如 Elsevier 出版社就在其期刊论文和图书等文献访问平台上提供与文献相关的科学数据关联服务，用户可以通过该平台获得该文献研究过程中所涉及的科学数据等资料②。

1.1.2 科学数据管理

1. 科学数据管理的内涵

科学数据管理是指对科研工作者在科学研究活动中产生的科学数据进行统筹协调、科学配置、整合管理，涉及对各类型科学数据进行采集、分类、标准化、发布及共享，以形成管理科学数据的理念、政策、规范、环境、措施与体系，发挥科学研究数据资源的最大效益。

毋庸置疑，科学数据管理的对象即为科学数据。对应于科学数据三个不同层次的定义，科学数据管理的内涵可包括宏观与微观两方面。

从宏观上说，科学数据管理是针对不同的应用目标，通过对原始性或者分散性的科学数据资源进行整合、存储、整理、加工、传播和利用等过程的管理，使其实现系统化、标准化和规范化；对科学数据活动的各要素（科学数据、技术与设备、人员与机构等）进行合理的计划、组织、指挥、调控，实现科学数据资源和相关资源的合理配置，建立健全共享机制，从而有效满足全社会对科学数据资源的

---

① 韩涛，等. 科学数据与科学文献相关性研究——以生物信息学为例[J]. 图书情报知识，2008 (12)：42-46.

② Jsbrand Jan Aalbersberg. Supporting Science through the Interoperability of Data and Articles[J]. D-LibMagazine，2011 (Number 1/2)[EB/OL]. [2012-11-18]. http://www.dlib.org/dlib/january11/aalbersberg/01aalbersberg.html.

需求①。

从微观上说，涉及科学数据的创建、获取、转换、共享、保护、记录和保存的过程均可称之为科学数据管理②。数据管理并非生成结果而是处理数据所得到的必然结果③。一般而言，对于具体机构来说，科学数据管理主要是微观上的，针对数据、资源本身的管理。

2. 科学数据管理的意义

近年来，随着各国的科技投入增大，科学观测和分析能力已得到快速的提升，导致科学数据的产生和积累呈指数级增长。科学数据的数量大约每年翻一番④。仅以科学实验数据增长为例，2009 年麻省理工学院（Massachusetts Institute of Technology，MIT）图书馆员调查了该校一些重要领域科学家的数据量，发现 6 个案例中的科学家每年产生数据总量大约为41 000TB，如物理系教授的数据量为 20 600 TB，神经影像学教授的数据量为 5.4TB，气候变化研究的科学家的数据量为 200TB⑤。由于仪器的精确性，数据质量也呈迅速上升趋势。当数据海量化、网络化、开放化和计算化时，它的作用也在发生革命性的变化。在大数据时代，科学家不仅可以通过对广泛的数据进行实时、动态地监测与分析来解决科学问题，更可把数据作为科学研究的对象和工具，基于数据来思考、设计和实施科学研究。数据不再仅仅是科学研究的结果，而是变成科学研究的活的基础和工具；人们不仅关心数据建模、描述、组织、保存、访问、分析、复用和建立科学数据基础设施，更关心如何利用泛在网络及其内在的交互性、开放性，利用海量数据的知识对象化、可计算化，构造基于数据的知识发现和协同研究⑥。正是科学研究态势的变化，使科学数据管理更具有潜力与价值。

根据“大科学”与“小科学”的差异、自然科学与人文社会科学的差异，科学数据管理的作用和内容有所不同，但总体而言，科学数据管理的目标是使数据更好地为研究者所利用，具体包括：

---

① 董诚，黄鼎成. 科学数据资源的管理[J]. 中国基础科学，2006(6)：20-24.

② University of Cambridge. Explanation of Terms [EB/OL]. [2013-03-27] . http://www.lib.cam.ac.uk/preservation/incremental/glossary.html.

③ ANU data management manual：Managing digital research data at the Australian National University[EB/OL]. [2013-02-20]. http://regnet.anu.edu.au/sites/default/files/files/ANU_Data_Management_Manual.pdf.

④ Jim Gray，et al. Scientific Data Management in the Coming Decade[J]. SIGMOD Record. 2005，34(4)：34-41.

⑤ Madnick S，et al. Case study summary the scientific data flood：How much information? [EB/OL]. [2013-02-16]. http://hmi.ucsd.edu/pdf/HMI_Case_Summary.pdf.

⑥ 梁娜，曾燕. 推进数据密集型科学发现，提升科技创新能力：新模式、新方法、新挑战[J]. 中国科学院院刊，2013(1)：115-121.

（1）促进科研人员的研究工作，提高科研成果生产速度

科研人员在科学研究过程的每一个阶段都会产生大量的数据，从科学研究的选题开始一直持续至科学成果的产生。可以说，科研工作的核心就是围绕数据进行的，并且工作流程越复杂的科学研究，过程数据的种类和数量会越多，数据对研究的持续性开展就更有价值，但处理、查找、利用数据的难度也相应增大。事实上，查找、整理现有数据是许多科研人员的难点。尤其对于单独从事科学研究的研究人员或小规模研究团队而言，其计算机运用能力和科学数据的管理、组织能力将直接影响其科研工作的进展。通过科学数据管理可实现数据的有序化、可网络获取、可计算、可开放关联，从而使科学研究工作的各个流程更为顺畅地运行，也能大大缩短科研人员在生成科研成果时用于查找、整理已有数据的时间，从而促进科研工作的进行。

实际上，一些大型的科研机构已经逐步形成了较为成熟的基于工作流程的数据管理规范，也具有了运用计算机技术对大量数据进行管理的能力。如澳大利亚南极中心，到2006年底为止，开发了基于Web的科学项目申请系统、元数据管理与发布系统；开发了基于内容管理的在线数据库实时管理系统；编写了1810多条元数据记录；提供了800多个在线数据文件和30多个Web数据库（如出版物、南极地名字典、南极地图目录、SCAR南极地形特征目录、南极生物多样性等）供用户查询与下载；已提供了11 000余次的数据下载服务；将多个在线数据库（科学项目、元数据、数据与出版物等）进行了关联等①。

（2）实现科学数据的共享，提高科学数据价值

科学数据管理的目标之一是实现科学数据在全球范围内的有效共享。管理是共享的前提，而共享则是管理的目标之一。科学数据的共享要求建立起统一的资源组织与交换标准，如我国科学数据共享平台要求采用统一的核心元数据标准，以便能够对核心元数据进行自动汇总，建立起跨学科、跨部门的共享体系。

科学数据的共享有助于科学研究的进行，是加强学术交流的重要途径之一。科学数据是一笔重要的资产，它的价值往往超越了其产生时对于所在课题组和科研项目的价值。科学数据的价值表现在三个方面：科学价值、经济价值和社会价值。当代科学数据的科学价值表现在它是科学研究的基础，同时它也是科学研究的“牵引力”；科学数据的经济价值表现在科学数据可以直接或间接为数据创建者、数据使用者带来经济效益；科学数据的社会价值主要体现在它可以提高信息时代全民素质、全民的自我教育、违规行为的监督、社会稳定、政府决策的监督和政府意志的潜移默化的执行能力等方面。美国作为科学技术领先的国家，其“完全与开放”的

① Mitsuo Fukuchi, et al. Report of“Workshop on science data management at the national institute of polar research”[J]. Antarctic Record, 2005, 49(1): 133-144.