

中华人民共和国科学技术部基础研究司

科学数据

STUDY ON THE MANAGEMENT OF SCIENTIFIC DATASHARING

共享管理研究

黄鼎成 郭增艳 主编



中国科学技术出版社

中华人民共和国科学技术部基础研究司

73.962
542

科学数据共享管理研究

黄鼎成 郭增艳 主编

ABE26/09



2003年1月18日

中国科学技术出版社

· 北 京 ·

00984

图书在版编目 (CIP) 数据

科学数据共享管理研究/黄鼎成, 郭增艳主编. —北京: 中国科学技术出版社, 2002.11

ISBN 7-5046-3409-3

I. 科… II. ①黄…②郭… III. 自然科学-数据管理-研究 IV. N37

中国版本图书馆 CIP 数据核字 (2002) 第 090339 号

中国科学技术出版社出版

北京市海淀区中关村南大街 16 号 邮政编码: 100081

电话: 62179148 62173865

新华书店北京发行所发行 各地新华书店经售

中国科学院印刷厂印刷

*

开本: 850 毫米 × 1168 毫米 1/32 印张: 12.75 字数: 284 千字:

2002 年 11 月第 1 版 2002 年 11 月第 1 次印刷

印数: 1-3000 册 定价: 40.00 元

(凡购买本社的图书, 如有缺页、倒页、
脱页者, 本社发行部负责调换)

科 学 指 导:

徐冠华

孙 枢

马福臣

张先恩

本书各章作者:

第一章 “基础性工作”调研组

第二章 “科学数据共享”调研组

第三章 黄鼎成 马福臣

第四章 孙九林

第五章 施慧中

第六章 李集明

第七章 李晓波 祝礼强 徐 枫 李集明 贾光宇

第八章 徐 枫

第九章 苗良田 李志雄 李兴良 陈华静 路 鹏

第十章 李集明 汤 绪 王国复 熊安元

第十一章 李望平

第十二章 孙九林 施慧中 李 爽

第十三章 刘 闯 王正兴

第十四章 周 叙 陈 军

序一

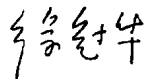
人类社会在走过充满发现、创造和发展的 20 世纪，又坚定地踏上了知识时代新的世纪征途。21 世纪将是科学技术突飞猛进的世纪，人类的未来和国家的繁荣更加依赖于科技创新及其产业化实现能力和效率。科学技术将成为人类社会变革与发展的主导力量，将为人类提供一个充满希望的未来。正如江泽民同志所指出：“一个没有创新能力的民族难以屹立于世界先进民族之林。”加快国家创新体系建设，全面提高科技创新能力是时代的要求，是关系到中华民族前途和命运的大事。

把握世界科技发展趋势和当代科技创新的特点，从国家发展需求出发，将创新能力提高到新的水平，支持与保证我国科技进步与创新、经济和社会的可持续发展，是全社会共同的历史责任。有鉴于此，科技部为了增强我国在日趋激烈的国际竞争中的主动性，在深入研究和广泛听取专家意见的基础上，决定加强“科技基础条件平台”建设。科学数据是具有明显潜在价值及其在广泛应用中得以增值的巨大社会财富，能够从根本上满足科技进步与创新、经济增长、社会发展和国家安全等多种需求，从而构成信息时代一种最基本、影响面最宽的科技创新资源，是“科技基础条件平台”的重要组成部分。面对科技创新国际化的严峻挑战，整合离散分布的海量科学数据资源，建立健全数据资源的共享机制，发

挥科学数据资源的最大效用，将贯穿“平台”建设与发展的全过程。

科技部基础研究司会同国家自然科学基金委员会综合计划局先行组织对基础性工作的调研以及基础研究司对科学数据共享问题及时组织研究，所形成的“科学数据共享工程”的基本思路，既反映了实施科学数据共享的难度和可能性，又体现了当代信息管理的主流。可以说，“科学数据共享工程”是在国家发展的迫切需求和现代信息技术迅猛发展的背景下产生的，是一项崭新的共享管理技术。一个好的管理秩序的形成，要在国家整体规划与管理下，通过政策的调控、法规体系的保障和共享技术的支撑，并在共享实践中不断增强共建共享观念，在实践中彼此交流、融合而不断完善。只有良好的共享管理秩序，才能保证科学数据资源的潜在价值在其广泛应用中得以发挥，从而不断增加科技资源的存量与战略储备，提高我国科技可持续创新能力。

我衷心希望本书的出版，能在科学数据共享工程的建设与发展中做出积极的贡献。希望社会各界有识之士，特别是科技、教育界和政府公务员以及数据服务业人士，对科学数据共享工程的建设与发展，对数据资源的共享管理献计献策，为增强原始创新能力，提高科技整体水平，为中华民族的伟大复兴而共同努力。



2002年8月20日

序二

“创新是一个民族进步的灵魂，是一个国家兴旺发达的不竭动力。”回顾人类理性日益成熟的 20 世纪，科学、技术和工程领域的发展和成就，无一例外地依赖于对科学技术活动本质认识的升华，依靠科技工作者的创新活动和现代科学基础设施与科学数据资源的支撑。展望 21 世纪，科学技术必将获得新的飞跃性发展，并将对经济社会的发展产生巨大、直接和全面的影响。在这样一个世界背景下，中国现代科学技术也走过了它的萌芽与发展时期，并将在新的世纪步入创新时期。

科学数据资源凝聚着人类科技活动、社会活动和经济活动的价值，是当代科技创新和国家发展需求的一种战略资源。如果我们把创新体系比作知识经济的动力系统，那么科学数据与相关信息则是这个动力系统的燃料资源。科学数据资源的存量及其效用，是科技进步与创新的强有力支撑，体现一个国家的创新能力与整体科技水平。随着科学技术的飞速发展，科学数据的采集、存储、处理和传播数量与日俱增，也出现了人们始料未及的新问题：一是科学数据量猛增，但对其利用严重不足；二是数据资源共享与数据占有相矛盾；三是数据产品共享亟待规范；等等。如何使海量的科学数据资源在全社会流动起来，规范数据的管理，最大限度地发挥科学数据作为战略资源的作用，已成为科技发展所面

临的新挑战。多年来，不少科学家曾为此奔走呼吁。可以说，科学数据共享工程是解决我国科学数据共享管理问题的战略举措，是一项功于当代、利泽千秋的大事。

科技部基础研究司主持编写的《科学数据共享管理研究》一书，集中了“基础性工作”和“科学数据共享”调研组同志们三年的研究成果，他们对科学数据共享管理的若干理论、方法和技术问题展开的讨论，颇有新意。希望该书的出版有助于提高人们对科学数据共建共享的意识，有助于创建科学数据共享管理秩序，为进一步提高我国科技原始创新能力而做出积极贡献。



2002 年 8 月

前 言

随着人类社会进入信息时代，科学数据已成为一种重要的资源和巨大的社会财富，它贯穿着科技活动的全过程，渗透到人类社会的方方面面，影响着整个社会的发展。众所周知，正是由于科学数据信息资源的迅速积累和现代信息技术的迅猛发展，使科学数据应用范围大大拓宽，使用价值成倍增长。知识、信息和数据成为信息时代最活跃、影响面最宽的战略资源。科学数据资源及其管理与共享水平，已成为衡量一个国家整体科技水平和综合国力的一项重要标志。这不仅因为科学数据是科技进步与创新所不可缺少的信息资源，是政府和相关组织机构决策的信息保障，而且在促进经济增长和社会进步向依靠知识转变过程中，科学数据发挥着越来越重要的作用。

在过去的 10 多年中，全球经济的发展趋势在深刻地影响着每一个国家的生存与发展。各国为了在全球竞争中把握科技创新和经济增长的主动权，纷纷制定了本国的信息发展战略。由于科学数据具有基础性、共享性和应用增值的特点，一些发达国家还建立了保证科学数据有效管理与传播的政策、法规体系，并把整个世界作为事业发展的空间，寻求扩大国际合作与竞争的效益。经过半个多世纪的建设与发展，我国通过大量的观测、探测、监测、调查、试验等公益性业务活动和重大科学工程、重点实验室、工程中心、野外观测研究站网的建设，以及一系列科技计划的实施，积累了

海量的科学数据。同时，在科学数据的采集和积累方面，初步形成了以部门为主体与科研院所、高等院校互为补充的格局。这些宝贵的科学数据资源在科技进步、社会发展、经济增长和国家安全等方面发挥了重要的作用。伴随国家信息基础设施的建设与发展，科学数据资源的管理将朝着数字化、网络化和国际化方向发展。在以信息技术带动产业结构调整，用信息促进科学技术发展的今天，科学数据资源的高效管理与传播是最基本的、影响广泛的基础性工作。

为此，国家科技部基础研究司于 1999 年 3 月会同国家自然科学基金委员会综合计划局联合组织了基础性工作调研；针对基础性工作中最具普遍意义的科学数据的积累与共享问题，于 2000 年 11 月又组织了科技基础数据库建设与发展的调研工作。调研组在黄鼎成教授的主持下，先后完成了“基础性工作及其国家政策与管理研究”、“国家科技基础数据库建设与发展的研究报告”，并于 2001 年 10 月提出了“实施科学数据共享工程，增强国家科技创新能力”的建议。

为了增强我国在日趋激烈的国际竞争中的主动性，我们必须加快国家创新体系建设，建立健全科技创新资源共享机制，加强现代科技基础设施的建设，以全面提高科技创新能力。科技部在 1999 年推出的基础性工作与公益性研究专项计划的实践基础上，又于 2002 年 6 月做出了加强科技基础条件平台建设的战略部署。科学数据共享工程正是实施这一战略部署的主要任务之一。它是在国家整体规划与管理下，应用现代信息技术整合各部门、各单位的科学数据资源，通过制定共享政策法规、完善管理体制与服务体系、建立健全共享机制，使各部门、各单位所积累的科学数据资源纳入国家科学数据共享管理的统一框架；通过科学技术不同领域的

国家科学数据中心和科学数据共享服务网的建设以及共享技术的研究开发，形成跨部门、跨地区、跨学科、多层次、分布式的国家科学数据共享服务体系。从而使国家长期投入而获得和积累的基础性、公益性科学数据实现分类分级共享，使海量的科学数据资源的潜在价值得以充分发挥与增值，在提升我国科技创新能力方面发挥重大作用。

然而，在当前我们不能不看到实施这项工程的艰巨性，公众意识是否达到了一定的水平？长期存在的部门分离屏障能否有效消除？有关法规体系的建设能否顺利推进？技术和资金保障能否及时落实？等等。因此，我们组织出版这本书的目的，一方面是向公众介绍即将实施的我国“科学数据共享工程”，另一方面也希望进一步听取各方面的真知灼见，使这项事业能够得到健康发展。在这里，我要特别提到的是调研组的专家们，他们来自不同单位，日常工作都十分繁忙。然而，他们将调研任务视为一项光荣的使命，做了大量深入细致的研究工作，提出了适合我国国情的合理建议，为构建“科学数据共享工程”做出了重要贡献。若干年后，当我们看到“科学数据共享工程”的实施，对我国的科技发展、社会进步和经济增长产生促进作用时，所有为此做出贡献的人，无论是科学家还是管理者，都会感到无限欣慰。

科技部基础研究司司长 张先恩

2002年8月12日

目 录

序 1	徐冠华 (1)
序 2	孙 枢 (3)
前言	张先恩 (1)

上篇 调研报告

第一章 基础性工作及其未来发展与政策研究	(3)
第二章 实施科学数据共享工程, 增强国家科 技创新能力	(34)

下篇 科学数据共享管理研究

第三章 科学数据管理的走向——资源共享	(59)
第四章 科学数据资源及其开发利用	(85)
第五章 科学数据分类分级	(117)
第六章 科学数据发布策略	(136)
第七章 科学数据共享技术框架与共享技术平台	(162)
第八章 元数据技术及其在科学数据共享中的 应用	(178)

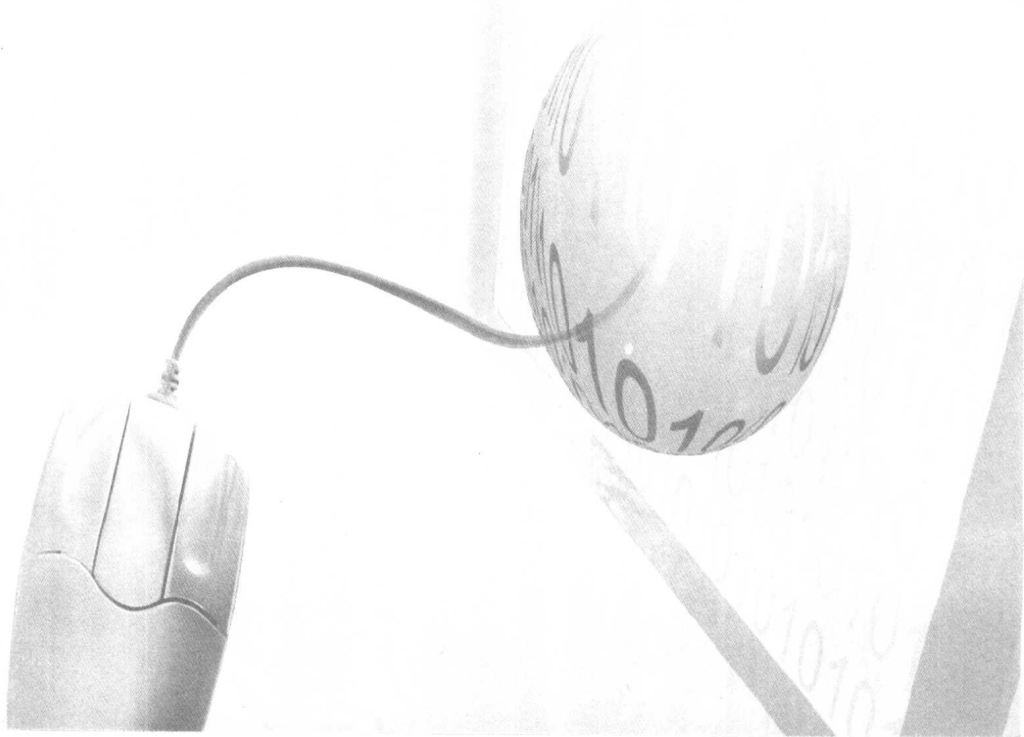
第九章	国家科技计划项目科学数据汇交管理制度的建设	(197)
第十章	气象科学数据共享试点实践	(213)
第十一章	科学数据库网络共享平台建设	(236)
第十二章	WDC—D 数据工作历程与展望	(245)
第十三章	美国国有科学数据“完全与开放”共享国策剖析	(267)
第十四章	国内外地理空间信息网络化服务情况分析	(293)

附 录

附录 1	国家科技基础数据库建设与发展的研究报告	(319)
附录 2	调研报告参与专家名单	(378)
主要参考文献		(382)
后记		(388)

上篇

调研报告



第一章 基础性工作及其未来发展与政策研究*

1987年全国基础性研究状况调研组所定义的基础性研究包括三类工作：①以探索未知、认识自然为主要目的，无明显应用背景的纯基础研究；②有广泛应用背景或应用目的，但以获取新知识、揭示新规律、发现新原理和新方法为目标的定向性研究；③对基本科学数据、资料和信息进行系统的收集、鉴定和评价、积累和综合分析，以探索基本规律的研究。其中，第三类所涉及的正是多年科技管理实践取得共识的基础性工作的重要内容。实践表明，人类在认识世界、改造世界的科技活动中所产生的，关于自然、社会乃至思维的数据、资料和相关信息，对科技进步而言，是最基本、影响面最宽的科技创新资源。加强基础性工作，意味着增强对国家科技创新能力的建设。

* 科技部基础研究司、国家自然科学基金委综合计划局 1999 年 3 月委托开展“基础性工作及其国家政策与管理研究”。课题组在基础研究司和综合计划局直接领导下，先后在北京、武汉、南京召开了由地学、生物学、农学、医学、环保、天文、空间、物理、力学、化学、技术科学以及 WDC-D、CODATA 国际科学数据组织等方面近 230 位专家参加的 21 个座谈会。课题组在听取到会专家意见的基础上，分析了国内外有关资料，于 1999 年 6 月提交了调研阶段报告《我国基础性工作状况和发展计划的总体构想与近期实施方案》；同年 12 月完成了“全国基础研究‘十五’计划和 2015 年远景规划”《基础性工作》专题报告。本章是在上述报告的基础上于 2000 年 3 月所做的课题总结，主要针对基础性工作的内涵、作用、发展趋势与现状、发展目标与任务、以及政策与管理的基本思路进行讨论。

一、基础性工作的内涵、特点和作用

(一) 基础性工作具有独特的内涵及其自身特点

分析认为,服务于科技进步与创新、推动基础科学发展、促进经济增长与社会发展的基础性工作是指对基本科学数据、资料和相关信息(统称科技信息)进行长期系统的采集、整理与保存,以探索基本规律,并推动这些科技信息的流动与使用。

从工作性质来看,基础性工作的内涵可以归纳为三个环节和一个关键技术,即:采集是应用多种技术手段,通过观测、探测、监测、调查、试验或实验以及发掘、汇集等工作,获取相关的科技信息,包括数据、文字资料、图片、实物标本和样品等。整理与保存是对所获取的科技信息进行分类、整理、鉴定与评价、积累与保存(包括活体的培养与延续),并通过综合分析研究制定各类科技标准,探索基本规律等方面的研究工作。流动与使用是指利用现代技术对上述科技信息,按照统一的技术标准和相关的政策法规,进行广泛的传播与共享。显然,在上述三个环节中,贯穿始终的则是技术标准与质量控制,它是保证科技信息流动与便捷使用的关键。技术标准与质量控制及相关法规体系,既是科技信息共享的前提,又是经济和社会发展的综合性技术基础。

从上述科技信息的产生过程来看,依据学科领域及其研究对象的不同,又可以将基础性工作分为两大类。一类是对自然现象的客观认识,需要通过各种手段进行不同规模的观测、探测、调查、发掘、采集和试验以获取海量的数据、资料(文字的或图形的)、标本(或样品)及其相关信息,主要集中在地学、生物学、天文学、空间科学、基础农学以及