



中国科学院大学

University of Chinese Academy of Sciences

# 数据管理与共享

第十一组

高红敏，黄梦洁，鲜槟丞，李文灏  
王传银，杨思琦，余子阳，李兆轩，张翔，刘霞，  
秦如雨，黄重庆，戴新环，王杏灵



中国科学院大学  
University of Chinese Academy of Sciences

1

数据管理与数据共享的背景

2

管理与共享平台

3

数据管理与数据共享政策

4

总结

目录  
Contents





中国科学院大学  
University of Chinese Academy of Sciences

01

## 数据管理与数据共享的背景



**什么是科学数据**



**什么是数据管理**



**数据共享好处**



**研究人员对于共享数据的态度**



**数据诚信在学术不端问题中的占比**



**数据管理平台建设必要性**



## 什么是科学数据？

随着科技的进步，尤其是信息技术的广泛应用，使得科学数据日益积累的规模和复杂度呈指数级增长。科学数据是指通过实验、观测、探测、调查等方式所获取的反应客观世界本质、特征、变化规律的原始数据，并根据不同科技活动需要，进行系统加工整理的各类数据集。科学数据已经成为信息社会的核心资源，在科学研究中，科学数据是不可缺失的支撑资源，它既是先前研究的产出，也是后续研究的输入。



## 什么是数据管理？

然而，零散、无序的数据很难被理解和使用，为了能够被有效地开发利用科学数据资源，还要进行有效的科学数据管理。

- 科学数据管理是针对不同的应用目标，通过对分散的科学数据资源进行整合、存储、加工、传播和利用等过程的管理，实现科学数据资源的合理配置，并建立健全共享机制，从而有效满足全社会对科学数据资源的需求。
- 科学数据管理的最终目的是实现科学数据资源的共享，使科学数据资源在广泛应用过程中增值。实施科学数据的管理与共享，对推动科技自主创新，减少科技领域的资源浪费具有积极作用。

## 数据共享的好处

科学进步的根源是知识和数据共享。

- 研究结果可以通过复制研究来验证，或由同一个数据集发展出一个新的假设。**数据共享**是推动开放科学的重要一步，也为科学发展开拓更多可能。数据共享是科学的核心，因此对学术界对数据共享的态度改变将促进科学的进步。促进科研数据的开放共享，以最大化利用科研数据的价值已成为全球趋势。

## 数据共享的好处

中国科学院发布的《数据管理与开放共享》办法文件中也明确表示科学数据应适时向院内外用户开放共享。

**第四条** 科学数据工作遵循统筹协调、规范管理、安全可控、持续发展的基本原则，落实主体责任，创新管理机制，加强能力建设，促进开放共享。

### 第五章 科学数据开放共享

**第十九条** 科学数据应按照分等级、可发现、可访问、可重用的原则，适时向院内外用户开放共享。



## 研究人员对于共享数据的态度

开放科研数据涉及许多利益相关者，科研人员在其中扮演着重要角色。科学发现的公开和透明化趋势对研究来说具有极大益处，虽然大多数研究人员认为数据共享是科学的理想，然而开放数据让大众取得还是有质疑的声音在。根据Wiley进行的大型调查，研究人员不愿分享原始数据的主要原因有：

- 知识产权/机密问题
- 伦理道德问题
- 资助单位/机构不要求数据分享
- 害怕没有被适当的引用
- 害怕想法被偷走
- 不知道在哪里分享数据
- 害怕曲解和误用

研究人员愿分享原始数据的主要原因有：

- 在他们的领域里“进行研究”
- 为他们的研究“增加影响和可见性”

数据诚信在学术不端问题中的占比

出版道德规范委员会（COPE）于1997年成立，有覆盖所有领域的9000多成员。COPE网站共收录了582个学术不端案例，其中涉及数据诚信问题的案例共125件，占比21.48%。美国记者Oransky和Marcus于2010年建立了“Retraction Watch”网站，致力于检查撤销文献，是监测跟踪撤销论文的实例。截至2017年，Retraction Watch共收录撤销文献18893篇，其中与数据诚信有关的撤销论文共2638篇，占比13.96%。

表1 COPE 学术不端案例统计

分类		案例数/件 占比/%	
出版后的更正	Post-publication discussions and corrections	215	36.94
道德监督	Ethical oversight	182	31.27
投诉和上诉	Complaints and appeals	134	23.02
知识产权	Intellectual property	134	23.02
作者和贡献者	Authorship and contributor ship	132	22.68
数据和可重复性	Data and reproducibility	125	21.48
期刊管理	Journal management	108	18.56
同行评审流程	Peer review processes	73	12.54
利益冲突/竞争利益	Conflicts of interest / Competing interests	70	12.03

表2 Retraction Watch 因数据诚信撤销的论文统计 (截至2017年)

撤销原因		撤销文献数/篇
关于数据的担忧/问题	Concerns/Issues About Data	773
重复数据	Duplication of Data	226
数据错误	Error in Data	759
伪造/制造数据	Falsification/Fabrication of Data	791
剽窃数据	Plagiarism of Data	44
不可靠的数据	Unreliable Data	45
总计		2 638

## 数据管理平台建设的必要性

- 科学数据开放共享在促进科学技术快速发展的同时，也造成了“数据获取”和“数据安全及隐私保护”之间的不平衡。
- 科学数据的管理既能保证对已有研究成果的检验，又能促进进一步的研究。可以有效减少“韩春雨事件”，同时加速我国创新能力的建设。
- 科学数据管理与共享有利于提高研究的开放性、完整性和可重复性，提高科学研究的透明度和严谨性，减少学术不端问题，有利于科研诚信建设。
- 随着《科学数据管理办法》的推出，也表明了国家重视发展科学数据管理平台的建设。





中国科学院大学  
University of Chinese Academy of Sciences

02

管理与共享平台



中国科学院大学  
University of Chinese Academy of Sciences



**科研数据管理与共享要求**



**国内平台软件建设模式**



**我国科学数据管理平台建设中的缺失之处**

Contents

## 科研数据管理与共享要求

公共资助机构对于科研数据管理与共享的要求主要集中在**数据管理计划**、**数据存放位置**、**存储期限**三方面。

- 对于数据管理计划，资助机构一般要求将其作为资助申请的一部分。
- 对于数据保存位置，如果资助机构自身拥有数据存储中心（如NERC、ESRC），一般要求数据保存在存储中心；如果没有，则要求将数据存放在受资助者所在机构的机构知识库中或第三方数据存储中心。
- 对于数据存储时间要求分为数据访问时滞期和保存期限两类，有的科研资助机构要求数据一经核实便要求即刻共享，但一般允许数据有一定的时滞期。



# 国内平台软件建设模式

- 自主研发模式。
- 采用开源软件开发模式。

表 1 国内外科学数据管理平台代表性软件开发模式比照

国内平台名称		平台软件	国外平台名称		平台软件
在 Re3data 上注册的平台 <sup>①</sup>	国家基础科学数据共享服务平台和各分学科平台 <sup>②</sup> http://www.nsdata.cn/	自主研发	美国密歇根大学校际政治和社会科学研究联盟 (ICPSR) https://www.icpsr.umich.edu/icpsrweb/		自主研发
	北京大学开放研究数据平台 http://opendata.pku.edu.cn/	开源 Dataverse	美国哈佛大学一麻省理工学院数据中心( HMDc) https://dataverse.harvard.edu/		自主研发 Dataverse
	华大基因 http://www.genomics.cn/	自主研发	美国国家冰雪数据中心( NSIDC) https://nsidc.org/data		自主研发
未在 Re3data 上注册的平台	清华大学中国经济社会数据研究中心 <sup>③</sup>	由国家统计局自主研发	美国国家生物技术信息中心( NCBI) https://www.ncbi.nlm.nih.gov/		自主研发平台系统 及生物数据搜索软件 Entrez
	中山大学学术研究数据库共享计划 <sup>④</sup>	开源 Dspace			
	中国人民大学中国国家调查数据库 http://cnsda.ruc.edu.cn/index.php	自建 SDA	美国地球观测数据网( DataOne) https://www.dataone.org/		自主研发
	华中科技大学中国高校社会科学数据中心 https://cmis.csdc.info/toIndex.action	-	美国康奈尔大学罗普公众舆论研究中心( ROPER) https://ropercenter.cornell.edu/about-the-center/data-curation/		自主研发 RoperExpress
	武汉大学高校科学数据共享平台 http://www.lib.whu.edu.cn/kxsj/aboutus.htm	开源 Dspace	美国芝加哥大学全国民意研究中心的综合社会调查项目 ( GSS) https://gssdataexplorer.norc.org/		自主研发 Data Explorer
	复旦大学社会科学数据平台 https://dyn.fudan.edu.cn/home/	开源 Dataverse	美国劳工统计局的全国跟踪调查( NLS) https://nsidc.org/data		自主研发
	湖南大学经济数据研究中心 http://edrc.hnu.edu.cn/	-	宾州州立大学图书馆科学数据服务 https://libraries.psu.edu/research/research-data-services		自主研发 Scholar Sphere
	同济大学科研数据管理与服务平台 <sup>⑤</sup>	开源 CHAIR, DSpace	英国数据档案馆( UDA) http://data-archive.ac.uk/		与挪威研究数据中心 ( NCRD) 合作研发 Nesstar
	上海外国语大学数据学术服务平台 https://datam.shisu.edu.cn/home/	开源 Dataverse	澳大利亚国家数据服务网( ANDS) https://www.ands.org.au/working-with-data		自主研发
	全国农村固定观察点调查管理和数据分析系统 <sup>⑥</sup>	自主研发	弗吉尼亚大学图书馆科学数据服务平台 https://data.library.virginia.edu/		与康奈尔大学合作研发 Fedora
	国家卫计委流动人口数据平台 http://www.chinaldrk.org.cn/wjw/#/home	自主研发	美国麻省理工学院图书馆 https://libraries.mit.edu/		与惠普实验室合作研发 Dspace
	土地调查成果共享应用服务平台 http://tdc.mlr.gov.cn/to_Login	自主研发			

## 我国科学数据管理平台建设中的缺失之处

- 资金来源单一，造成数据管理平台不能均衡发展
- 软件开发缺少开源理念，提高了我国整体平台系统开发成本
- 部分平台存在单打独斗现象，缺乏合作建设理念
- 平台服务功能不全面，会影响数据使用的用户黏度
- 高校图书馆科学数据管理平台数量少，尚未形成规模



中国科学院大学  
University of Chinese Academy of Sciences

03

## 数据管理与数据共享政策





中国科学院大学  
University of Chinese Academy of Sciences



**国内外数据管理与开放共享政策**



**我国管理与共享政策体系的不足**



**我国管理与共享政策体系未来的发展**

Contents



国内数据管理与开放共享政策

表 2 我国有关科学数据管理与共享的政策

政策类型	政策名称	发布机构	发布日期
法律	中华人民共和国科技进步法(2007 修订)	全国人大	2007.12
	中华人民共和国促进科技成果转化法(2015 年修正)	全国人大	2015.10
行政法规	国家中长期科学和技术发展规划纲要(2006-2020 年)	国务院	2005.12
	国务院关于改进加强中央财政科研项目和资金管理的若干意见	国务院	2014.03
	关于深化中央财政科技计划(专项、基金等)管理改革的方案	国务院	2014.03
	国务院关于国家重大科研基础设施和大型科研仪器向社会开放的意见	国务院	2014.12
	“十三五”国家科技创新规划	国务院	2016.07
	科学数据管理办法	国务院	2018.04
部门规章	国家科技计划项目科学数据汇交暂行办法(草案)	科技部	2003.10
	“十一五”国家科技基础条件平台建设实施意见	科技部、财政部、发改委、教育部	2005.07
	国家重点基础研究发展计划资源环境领域项目数据汇交暂行办法	科技部、财政部	2008.03
	国家高技术研究发展计划(863 计划)管理办法	科技部、财政部、中国人民解放军原总装备部	2011.08
	国家科技支撑计划管理办法	科技部、财政部	2011.09
	国家重点基础研究发展计划管理办法	科技部、财政部	2011.11
	国家重大科研基础设施和大型科研仪器开放共享管理办法	科技部、发展改革委、财政部	2017.09
	气象资料共享管理办法	中国气象局	2001.12
	地震科学数据共享管理办法	中国地震局	2006.06
	中国极地科学考察数据管理办法实施细则	国家海洋局极地考察办公室	2006.08
	中国极地科学考察样品和数据管理办法(试行)	国家海洋局	2010.11
	交通运输科学数据共享管理办法	国务院交通运输主管机构	2007.07
	水文水资源科学数据共享管理办法实施细则(试行)	中国水科院	2011.07
	水文水资源科学数据共享管理办法(试行)	水利部水文局	2011.07
地方规章	太原科技资源开放共享条例	太原市人大常委会	2010.11
	重庆市科技资源开放共享管理办法	重庆市科学技术委员会、财政局、教育委员会、人力资源和社会保障局、物价局	2016.11
数据共享中心政策	农业科学数据汇交管理办法	农业科学数据共享中心	-
	农业科学数据共享管理办法	农业科学数据共享中心	-
	人口健康科学数据共享平台科学数据汇交管理办法	人口健康科学数据共享平台管理中心	2012.01
大学政策	武汉大学数据管理办法	武汉大学	2016.10

## 中科院数据管理与开放共享办法

明确科研项目数据汇交要求，提出建立论文关联数据汇交机制，明确科学数据开放共享主体责任，并对工作机制、业务流程、适用范围等方面进行了进一步阐述。

中科院数据管理与开放共享办法政策要点：

- 明确科研项目数据交汇要求 科技项目数据管理计划作为立项必要条件；项目负责人按数据管理计划汇交数据；科技项目数据管理计划完成情况作为验收必要条件
- 建立论文关联数据汇交机制
- 明确科学数据开放共享主体责任 院属法人单位、野外台站、重大科技基础设施、院科学数据中心
- 规划中科院科学数据中心体系



## 国外数据管理与数据共享政策举例

- 美国教育科学研究院(Institute of Education Science, IES)在《IES 资助拨款产生的数据实施公共获取的政策声明》中指出, 科研人员可将数据共享的成本包括在他们的资助申请和预算请求中, 包括准备和归档要共享的数据及数据文档的成本。
- 美国白宫科学技术政策办公室 (Office of Science and Technology Policy, OSTP)在《增加联邦政府资助的研究成果的获取》备忘录中强调了支持科研数据管理、分析、存储、保存和监管相关的培训、教育和员工发展的重要性。
- 德国科学基金会(Deutsche Forschungsgemeinschaft, DFG)与瑞士科学基金会(Swiss National Science Foundation, SNSF) 在科研数据管理政策中明确提及可资助科研人员的数据准备和非商业的数据上传等工作。
- 2011年12月英国商业、创新与技能部发布的《促进增长的创新研究战略》指出政府将建立开放数据研究所, 帮助商业和学术机构管理和发布数据。

## 我国管理与共享政策体系的不足

我国管理与共享政策体系虽然已经初步形成，但是还不够系统完善，数据管理仍处于条块分割和各自为政的状态，数据共享仍存在不少困难和阻碍。

例如国外的科研数据管理与共享政策大都要求科研人员在科研项目开始前制定科学合理的数据管理计划,并以数据管理计划为基础推动科研数据的管理与共享,同时这些政策有着较强的可执行与可实施性。

我国已有的科研数据管理与共享政策主要是由科学数据共享工程的各数据共享中心制定的。与国外的数据管理与共享政策相比,我国的科研数据政策在指导思想上存在着“重管理轻共享”的特点。

因此，需要借鉴国外的经验，在当前科研数据政策体系的基础上，进一步完善我国科研数据管理与共享政策体系，实现各领域、各层面科研数据的有效管理。

## 我国管理与共享政策体系未来的发展

- 国家层面的相关法律法规，包括“科研数据共享法”、“政府信息公开法”、“保密法”、“隐私保护法”、“知识产权法”等法律法规；
- 科研资助管理机构的数据政策及指南，包括“科研数据管理与共享政策”、“数据汇交政策”等；
- 科研单位的数据管理，包括“本单位的科研数据管理与共享政策”、“知识产权政策”、“学术规范政策”、“数据保密政策”等。





中国科学院大学  
University of Chinese Academy of Sciences

04

总结

- 小组想法：将数据管理和共享规则细化到学校，细化到个人，每个学校可以设立一个数据管理审查组，对学校的科研人员的相关论文数据进行不同权限管理，同时可以避免学术不端现象。
- 数据的共享和管理在信息技术高度发展的今天显得越来越重要，我国现在虽然已经制定了一些相关的政策，但还有很多不完善的地方。我们相信在未来这些政策和制度也能越来越完善，让我们科研工作者的工作变得越来越方便。

- [1]王凯,彭洁,屈宝强.科学数据管理与共享领域文献计量研究[J].中国科技资源导刊,2015,47(4):31-39.
- [2]周玉琴,邢文明.我国科学数据管理与共享政策体系研究[J].中华医学图书情报杂志,2018,27(8):1-7.
- [3]张萍. 英国高校科研数据管理及启示[J]. 情报杂志(4):159-163.
- [4]顾立平,包秦雯.国外开放科研数据的激励政策[J].图书馆论坛,2019(8):138-145.
- [5]崔旭, 赵希梅, 王铮, et al. 我国科学数据管理平台建设成就、缺失、对策及趋势分析——基于国内外比较视角[J]. 图书情报工作, 2019(9):21-30.
- [6]袁红卫, 黄松, 刘嫣. 麻省理工学院科学数据管理与共享平台调研及启示[J]. 图书馆学研究, 2019.
- [7]华小琴,司莉,李亭.我国科学数据共享中障碍因素分析及其启示[J].图书馆工作与研究,2019(11):18-26.
- [8]姚长青, 田瑞强. 开放科学中的数据诚信问题研究[J]. 科技与出版, 289(01):132-137.





中国科学院大学

University of Chinese Academy of Sciences

谢谢大家的聆听！