



# 科学数据管理与服务 新进展

*NEW PROGRESS IN SCIENTIFIC DATA MANAGEMENT AND SERVICES*



复旦大学 张计龙

2019年10月17日



# 目录

**C**ontents

时代背景 01

科学数据管理新进展 02

新探索与新实践 03

总结与展望 04

PART  
01

# 时代背景





# 大数据时代的到来

## 机遇

在科学发展的历程中，数据作为对自然、社会现象和科学实验等的定量或定性记录，一直是科学研究的重要基础。

计算机的发明提高了人类的计算能力，也带来了新的数据存在形式和存储介质。计算机、互联网为代表的通信技术和传感技术的持续创新和广泛应用使人类的数据化能力和范围迅速发展：

- 丨 能够测量和记录的东西越来越多
- 丨 对事物、现象的测量更加频繁、细致

21世纪，从宏观到微观、从自然到社会的观察、计算、传播等仪器设备和活动正在快速产生着海量、多样的数据，几乎每个领域都面临着空前的“数据爆炸”。







# 第四范式的出现

## 机遇

图灵奖的得主Jim Gray 2007年提出：

仪器设备和模拟方法正在生产出大量的数据，科学的世界发生了变化，继实验科学、理论科学和计算科学之后，出现了第四种科研范式——**数据密集型科学研究范式**。

- 以数据为中心思考、设计和实施科学研究
- 科学发现依赖于对海量数据的收集和分析处理





## 大数据促进科学与人文融合

机遇

大数据技术将自然世界及人类思想、行为的“数据化”，将自然、社会和人类思维映射为一个由数据构成的数据世界，无论是自然科学还是社会科学，其实质都成为对数据的挖掘、计算和预测。

大数据背景下，**自然科学**通过观察或受控实验所获取的数据寻找自然规律，**社会科学**有了大数据同样可以获取客观、真实、精准的数据，并通过挖掘经验数据，发现数据间的相关关系，寻找蕴藏其中的社会规律，**社会科学与自然科学一样可以用数据的形式表征其规律，实现算法化。**





# 大数据与科学研究

## 挑战

- n 大数据意味着**新的科研方法**，以**数据为中心**来思考、设计和实施科学研究的第四范式正在成为一种普遍范式，科学发现依赖于对海量数据的收集、管理和分析。然而，随着数据规模、复杂度和产生速度的增加，数据获取、管理、分析和应用的难度必然增大。
- n **如何保证数据的流动性和可获取性？如何有效管理大数据并保证其可被高效访问？如何继承和融合多种来源的复杂海量数据？如何快速处理大数据以及如何高效地从大数据分析和挖掘出信息和知识等？只有解决了这些关键的问题，科研人员才能真正掌控大数据。**
- n 解决这些问题不仅依赖于研发和应用新的技术，而且依赖于强大的数据基础设施、配套的科技政策的建立健全，以及科研体制的转型和新型人才的培养。

PART

02

# 科学数据管理新进展







## “FAIR 原则” 成为通用原则和导向

- n 2014年，各国学术界、产业界、资助机构和学术出版机构齐聚荷兰莱顿大学，举办了Jointly Designing a Data FAIRPORT研讨会。与会代表共同设计、起草了FAIR原则，旨在发挥指引作用，促进科研数据的可重复使用，使相关利益方能够更加容易地发现和获取，适当地整合和重复利用，充分地引证当代数据密集型科学研究中产生的大量信息。
- n “FAIR 原则”：Findable（可发现）、Accessible（可访问）、Interoperable（可交互）、Reusable（可重用）
- n 2016年FAIR指导性原则正式发布在《科学数据》第三期，阐述了其理论基础和典型案例。
- n FAIR 原则将科学数据资源依开放状态分为 6 类，其中“FAIR 化的元数据”、“FAIR 化有限开放数据”、“FAIR化开放数据”和“FAIR化增强版开放数据”等4类被认为是开放数据的主要形式。
- n FAIR原则正在欧盟、美国、澳大利亚等地更加普及和流行，“FAIR原则”正成为科学数据管理、监管与出版的最新通用原则。FAIR 化数据资产的计量研究和以“Go change, Go build, Go train”为主题的 FAIR运动，进一步使该原则落地。



# 史上最严数据保护法GDPR

- 2016年4月14日，欧盟议会投票通过了讨论四年之久的《通用数据保护条例》（General Data Protection Regulation, GDPR），该条例将替代1995年的《个人数据保护指令》。
- 2018年5月25日，旨在保护数据时代的欧洲公民免于隐私数据泄露GDPR正式生效，开始施行。
- GDPR共11章99条，主要从数据主体、数据控制者与处理者、数据监管者三方面，重新审视欧盟的个人数据保护法规，其核心内容确立“知情权、访问权、反对权、个人数据可携权、被遗忘权”5种公民权利。
- GDPR对个人信息保护的范围之广、实施标准之高、监管力度之大均前所未有，因此，也被称为“史上最严”的数据保护条例，亦被视为近20年来最重要的数据隐私规章。

GDPR





# 中国出台《科学数据管理办法》

科学数据是国家科技创新发展和经济社会发展的重要基础战略性资源。

近年来，我国科学数据呈现出“井喷式”增长态势，为进一步加强和规范科学数据管理，保障科学数据安全，提高开放共享水平，更好地为国家科技创新、经济社会发展和国家安全提供支撑，**国务院办公厅于2018年4月2日印发《科学数据管理办法》。**

《办法》明确了我国科学数据管理的总体原则、主要职责、数据采集汇交与保存、共享利用、保密与安全等方面的内容，着重从五方面提出了具体管理措施。



中华人民共和国中央人民政府

www.gov.cn



国务院

总理

新闻

政策

互动

服务

首页 > 信息公开 > 国务院文件 > 科技、教育 > 科技

索引号: 000014349/2018-00052

发文机关: 国务院办公厅

标题: 国务院办公厅关于印发科学数据管理办法的通知

发文字号: 国办发〔2018〕17号

主题词:

主题分类: 科技、教育\科技

成文日期: 2018年03月17日

发布日期: 2018年04月02日

## 国务院办公厅关于印发 科学数据管理办法的通知

国办发〔2018〕17号

各省、自治区、直辖市人民政府，国务院各部委、各直属机构：

《科学数据管理办法》已经国务院同意，现印发给你们，请认真贯彻执行。

国务院办公厅

2018年3月17日

(此件公开发布)

(新华社北京电)

2018年3月17日

国务院办公厅

《科学数据管理办法》已经国务院同意，现印发给你们，请认真贯彻执行。

各省、自治区、直辖市人民政府，国务院各部委、各直属机构：



## 《上海市公共数据开放暂行办法》

2019年8月16日，市政府第61次常务会议审议通过《上海市公共数据开放暂行办法》（以下简称《办法》），目前已经正式公布。《办法》作为国内首部针对公共数据开放进行专门立法的政府规章，将于2019年10月1日起正式施行。

《办法》共8章48条，分为**总则、开放机制、平台建设、数据利用、多元开放、监督保障、法律责任以及附则**。重点考虑了四个关系：**开放服务和管理责任、数据开放和数据安全、国际通行原则和上海实际情况、统一平台开放和多元合作生态**，将指导上海市公共数据开放利用工作进入全新阶段。

《办法》在国内**首次提出分级分类开放模式**，对开放数据进行标准化、精细化管理，提升开放数据质量，不断满足社会公众对公共数据的需求。

## 上海市人民政府令

第 21 号

《上海市公共数据开放暂行办法》已经2019年8月16日市政府第61次常务会议通过，现予公布，自2019年10月1日起施行。

市 长

2019年8月29日





# 《中国科学院科学数据管理与开放共享办法（试行）》

为进一步加强科学数据管理，保障科学数据安全，提高科学数据开放共享水平提供制度规范，中国科学院在2019年2月21日率先制定并发布了《中国科学院科学数据管理与开放共享办法（试行）》。

《办法》分为8章，共32条，对中科院科学数据管理与开放共享的总体原则、职责分工、管理要求、保障机制及安全保密等方面的内容进行了明确。是落实国家大数据战略和《科学数据管理办法》的重要举措，亦为各大高校、科研院所和企业等法人单位在科学数据管理制度规定方面提供了参考。





## 《科学数据引用》正式发布实施

“

2017年12月29日，国家标准化管理委员会正式发布《信息技术 科学数据引用》国家标准，并于2018年7月1日起正式实施。

”

《科学数据引用》针对科学数据传播机构和数据使用者，规定了科学数据引用元素描述方法、引用元素详细说明、引用格式等内容。其中，科学数据传播机构可根据该标准设计数据引用系统，并声明数据引用规则，数据使用者可根据该标准著录科学数据引用信息。《科学数据引用》国家标准的正式发布，标志着科学数据可以像学术论文一样被同行标准化引用，这将在一定程度上促进数据拥有者开放共享其数据。



## 数据出版为科学数据开放管理提供新思路

- 数据期刊积极探新与模式，数据论文类型更加丰富，出版形式更加多样。  
e.g. Nature出版集团的数据期刊Scientific Data在已有Data Descriptor基础上，再扩充文章类型，新增Analysis和Article两类文章。
- 数据期刊与数据论文出版对象延伸至各类研究产出，以发表单个观察结果、研究产出等非传统出版为目的的期刊正在兴起。  
e.g. **Pensoft** 于**2015年12月**创办新的开放获取期刊**Research Ideas and Outcomes**支持**出版整个科研生命周期中的每一步骤**，这既包括传统研究论文，也包括非传统研究产出，包括但不限于：**科研设想、资助建议书、实验设计、 workflow、数据管理计划、软件管理计划、博士后项目书、见解论文、数据论文、软件描述、会议摘要、研究介绍、研究海报、项目报告、政策与交流简报、案例研究等。**
- 更多语种的数据期刊开始萌芽。  
e.g. 《中国科学数据》、《地球大数据》、《图书馆杂志》...

PART

03

# 新探索与新实践







## 成立大数据研究院人文社科数据研究所

依托大数据学院与大数据研究院的技术力量、先进设施和成果转化服务能力，充分发挥复旦大学人文社会科学的学科优势，构建人文社科科学大数据共享服务平台和服务体系：

1. 整合、开发、共享、利用人文与社科科学领域的各类数据资源；
2. 实现人文和社会科学数据监护理论创新和研究范式的转变；
3. 培养适应大数据时代需求的人文社科复合型研究人才；
4. 推动人文社科科研成果的转化应用；
5. 为人文社会科学学科的研究模式提供智慧化和精准性的科学依据。



# 成立大数据研究院人文社科数据研究所

## I 资源基础

彭希哲：  
社会科学专题数据库

社会科学数据平台  
中国人口展望(CPP)  
长三角地区社会变迁调查 (FYRST)  
长江流域经济带数据库

.....

潘悟云：  
东亚语言数据库

1100个汉语方言字音库  
302个少数民族语言点  
数据质量高，可直接用于研究

张乐天：当代中国社会生活资料库  
(纸本)

各类资料60余万件  
书信31万多封  
日记3200册  
10多个生产大队的详细、完整的会计、统计资料  
基层政府文书资料约10万件  
企事业单位文书资料约9万件

图书馆：  
古籍、地方志数据库

古籍40万册，善本6万册  
民国图书10万册  
地方志10万册（已数字化1万册）  
少数民族文献2.3万册

.....

13家研究机构：  
特色资源

河西地方文献  
哈萨克方言数据库  
南诏大理文献  
贵州17个世居少数民族文献  
东北边疆历史文献

.....



# 成立大数据研究院人文社科数据研究所

## I 发展规划





# 复旦大学人文社科数据平台

復旦大學 | 人文社会科学  
数据平台

数据资源



可视化展示



数据管理与服务



常用工具



合作伙伴



平台简介







# 复旦大学人文社科数据平台

## 数据资源

数据资源

复旦大学 | 人文社会科学数据平台

自建品牌数据库 社会调查数据 特色文献数据库 区域经济社会发  
展数据 大数据资源 政府托管数据 合作资源 国内外数据资源

当代中国社会生活资料数据库

点击查看

| 地区 | 人口 (万人) |
|----|---------|
| 华北 | 486.844 |
| 华东 | 225.223 |
| 华南 | 6.790   |
| 华中 | 194.0   |
| 西南 | 22.0    |
| 西北 | 23.0    |
| 东北 | 9.0     |
| 东南 | 11.0    |
| 华南 | 17.0    |
| 西南 | 14.0    |
| 西北 | 18.0    |
| 东北 | 1.0     |



# 复旦大学人文社科数据平台

## 数据可视化

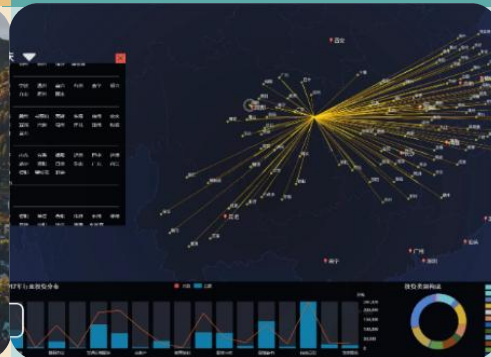
复旦大学数据版图



当代中国社会生活资料分布



银联刷卡消费可视化系统



国家珍贵古籍名录





# 复旦大学人文社科数据平台

## 数据工具

### 数据分析与统计工具



### CGE建模工具



### 系统动力学建模工具



### 质性数据分析工具



### 地理信息系统工具



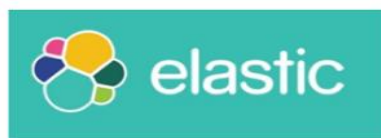
### 微观仿真建模工具



### 可视化工具



### 大数据工具







## 合作探索数据出版

- 复旦大学大数据研究院人文社科数据研究所联合《图书馆杂志》，聚焦大数据时代的数据出版领域，共同研发推出了图情领域国内第一家期刊数据管理平台，携手打造“数据论文”栏目，以发表图情领域高质量、高价值、可机读的科学数据集为方向，以期为读者提供一手、新型的研究数据资料。通过数据出版，让数据充分共享，发挥数据自身的巨大潜在价值，促进图情研究领域的理论创新与发展。





# 合作探索数据出版

- 截至2019年9月，已成功在《图书馆杂志》数据出版了30篇论文，其中包含5篇数据论文、5篇含数据支撑的论文和20篇含富媒体格式的论文，平台已发布课题40个，已发布文件183个，平台首页访问量已达5998次，下载量共计34726次，其中数据论文的数据下载量高达34672次，另待审核的数据论文已有近10篇。

84 · 数据论文 ·

库刊联合 (2019年第4期) 总第40期  
Library Journal (Vol.7) No.4

## 2015年复旦大学师生中文电子期刊资源访问行为数据集

殷沈琴 张计龙 汪东伟 胡杰 (复旦大学文科数据研究所)

摘要 通过ERU系统采集2015年全年复旦大学师生访问中文电子期刊数据库的行为数据，经过完整性、准确性和代表性三个维度的筛选，筛选出两个中文期刊数据库平台的访问行为数据，并对个人信息进行加密不可恢复性脱敏处理，结果共有记录10,119,795条。本数据集可以用于电子资源采购需求分析和效益分析、过量下载监控，也可以用于替代计量分析、学科热点分析和学科交叉趋势分析，具有实用价值和科研价值。

关键词 期刊数据库 用户访问行为 ERU 复旦大学 2015年 数据爬取

DOI: 10.13663/j.cnki.112019.08.012

Dataset of Chinese Electronic Journals Resources Access Behavior at Fudan University in 2015

Yin Shenqin, Zhang Jilong, Wang Dongwei, Hu Jie

### 2015年复旦大学师生中文电子期刊资源访问行为数据集

作者: 殷沈琴, 张计龙, 汪东伟, 胡杰

描述: 通过ERU系统采集2015年全年复旦大学师生访问中文电子期刊数据库的行为数据，经过完整性、准确性和代表性三个维度的筛选，筛选出两个中文期刊数据库平台的访问行为数据，并对个人信息进行加密不可恢复性脱敏处理，结果共有记录10,119,795条。本数据集可以用于电子资源采购需求分析和效益分析、过量下载监控，也可以用于替代计量分析和学科交叉趋势分析，具有实用价值和科研价值。

hdl:20.500.12304/ERU2015

22539 下载

最新发布: Sep 10, 2018

### 2016年复旦大学人文社会科学领域中文电子期刊资源访问行为数据集

作者: 胡杰, 张计龙, 殷沈琴, 汪东伟

描述: 2016年复旦大学人文社会科学领域中文电子期刊资源访问行为数据集通过ERU系统采集而来，共有27个字段，共2182755条记录，具备完整性、准确性和代表性。本数据集对于人文社会科学研究热点分析、用户行为分析、高校图书馆电子资源需求分析等众多研究方向都具有丰富的应用和研究价值。

hdl:20.500.12304/ERU2016HSS

5414 下载

最新发布: Nov 2, 2018

### 2016年复旦大学自然科学领域中文电子期刊资源访问数据集

作者: 汪东伟, 伏安娜, 张计龙, 殷沈琴, 胡杰

描述: 通过ERU系统采集2016年复旦大学全校用户访问自然科学领域中文电子期刊资源的行为数据。本数据集共有记录225644条，分为12个数据文件。本数据集有较强的应用价值和科研价值，包括为数据资源采购提供决策支持，分析用户信息行为和学术行为，了解科学研究热点和学科交叉趋势等。

hdl:20.500.12304/ERU2016NATURALSCIENCES

1118 下载

最新发布: Sep 11, 2018

### 2016年全球知名智库影响力指标数据集

作者: 李洁, 黄晨, 余敏杰, 陈国钢, 刘涛

描述: 通过Web of Science、CNKI、ProQuest、中华数字书苑数据库和搜索引擎等第三方网站，采集截止2016年底的全球各大知名智库的学术影响力、知名度影响力和客观性数据。经过广泛调研，参照宾夕法尼亚大学《2015 Global Go To Think Tank Index Report》和中国社会科学院《全球智库评价报告(2015)》，秉承公开性、完整性、代表性的原则筛选出169家全球著名智库，从基本情况、学术影响力、知名度影响力和同行评议四个角度来采集数据。本数据集可以用于全球著名智库的发展分析和研究，也可以用于对智库的建设现状以及其对公共政策影响力的评估。

hdl:20.500.12304/TTIMPACT2016

33 下载

最新发布: Aug 27, 2018

| 篇名              | 作者                      | 刊名    | 发表时间       | 被引  | 下载    | 阅读   | 收藏 |
|-----------------|-------------------------|-------|------------|-----|-------|------|----|
| 1 图书馆需要怎样的“大数据” | 樊伟红; 李晨晖; 张兴旺; 秦晓珠; 郭自宽 | 图书馆杂志 | 2012-11-15 | 237 | 10298 | HTML | ☆  |





# 中国高校研究数据管理推进工作组

- 为促进我国研究数据管理的发展，2014年10月23日，复旦大学图书馆、北京大学图书馆、清华大学图书馆、上海交通大学图书馆、浙江大学图书馆、武汉大学图书馆、北京理工大学图书馆、上海外国语大学图书馆、同济大学图书馆9家单位共同成立了中国高校图书馆研究数据管理推进工作组，工作组秘书处设立在复旦大学。2016年工作组年会上，决定正式更名为中国高校研究数据管理推进工作组。





# 中国高校研究数据管理推进工作组

工作组目标：

- ü 宣传推动研究数据的科学管理；
- ü 推进研究数据管理措施和政策的制订和施行；
- ü 提升图书馆员和科研人员的数据素养及研究数据管理理念和技能；
- ü 推进研究数据的创建、收集、存储、发布、应用、重用、出版及共享；
- ü 在中国乃至全球范围内，促进跨学科、跨部门、跨区域研究数据共享与利用；
- ü 推进研究数据相关标准规范的制订和实施；
- ü 推荐或提供可信任的研究数据管理与共享的平台与工具；
- ü 为高校图书馆的研究数据管理建设提供最佳实践案例；





## 当代中国社会生活资料共建共享联盟

- 2016年11月，东亚高校图书馆国际论坛上复旦大学图书馆和张乐天教授等共同提出，倡议在校际图书馆间建立联盟机制，实现当代中国社会生活资料共建共享，充分利用，服务于全球学术研究。

- 2017年，图书馆与大数据研究院成立人文社会科学数据研究所，筹建当代中国社会生活资料馆，联合启动资料数字化和数据库建设。



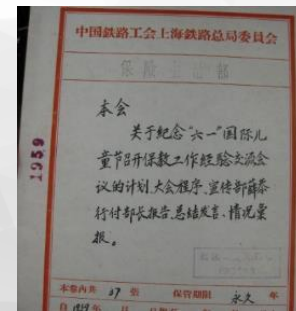
基层政府文书  
约20万件



民间书信  
超过50万封



日记和笔记  
5000多册



企事业单位文书  
约80万件



## 当代中国社会生活资料共建共享联盟

- 2017年，复旦大学人文社会科学数据研究所成立同时，筹建当代中国社会生活资料馆，联合启动资料数字化和数据库建设。
- 2018年初，正式启动联盟筹备工作。
- 2018年10月16日上午，召开联盟第一次工作研讨会
- 2018年10月17日，联盟正式成立揭牌。

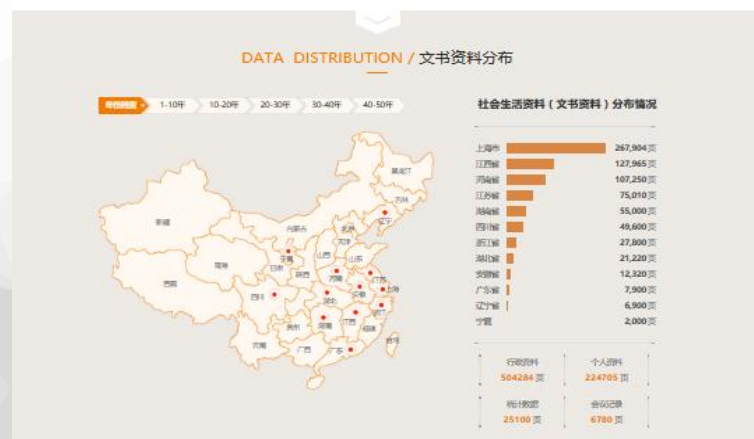






# 当代中国社会生活资料数据库建设

- 2019年CADAL研究项目《当代中国社会生活资料数据库共建共享》正式启动。
- 引入哈佛大学Datatags数据分级思路，为当代中国社会生活资料数据设计**数据分级模型和分级方案**。
- 修订当代中国社会生活资料数据申请与审核模式，完成数据申请流程、数据申请表的设计。
- 计划后续逐步对已经完成数字化的50万页生活资料完成数据分级，实现当代中国社会生活资料联盟成员数据共享。
- 截至2019年8月，数据库网站可以访问约2.8万条资料元数据和13万页资料图片，平台访问量超22万。







# 当代中国社会生活资料共建共享联盟

## 联盟成员

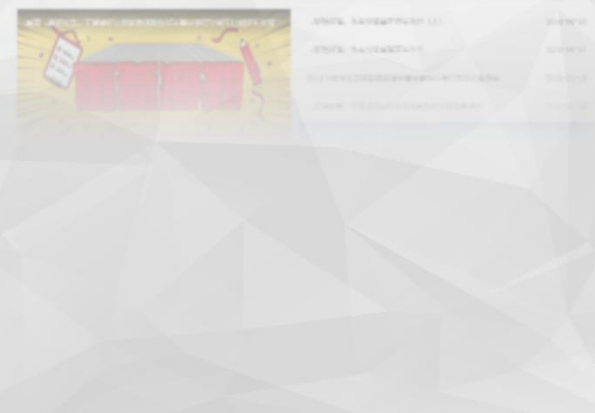
- 国内图书馆**14**家，包括北京大学图书馆、北京师范大学图书馆、复旦大学图书馆、南开大学图书馆、浙江大学图书馆、人民大学图书馆、香港中文大学图书馆等。
- 海外高校图书馆和研究机构**13**家，包括加州伯克利大学东亚图书馆、加州斯坦福大学东亚图书馆、哥伦比亚大学图书馆、普林斯顿大学东亚图书馆、日本东洋文库等。





## 慧源上海地区高校优质资源共建共享平台

- 市教委牵头，上海地区多家高校图书馆共同参与的资源共享平台。
- 2014年正式上线，2017年改版升级，现整合了沪上10所高校20个特色库，惠及沪上150万师生。
- 在大数据背景下，慧源共享平台技术架构进一步升级扩展，支持去中心化的科学数据资源的共享。
- 截至2019年6月，慧源平台访问量超过96万。
- 2016年获“2016上海市智慧城市建设优秀实践成果奖”。
- 2017年获“上海科学技术情报成果奖二等奖”。
- 截至2019年8月，平台访问量超97万。





# 上海高校开放数据创新研究大赛



# 慧源共享

## 首届上海高校开放数据 创新研究大赛 2019年4月—12月·上海

- Ø 推动和促进上海市高校间数据资源的开放共享
- Ø 鼓励上海市高校师生基于开放数据资源进行创新应用与研究
- Ø 培养和提升沪上大学生的数据素养与数据能力

**慧源共享 2019年“慧源共享”首届上海高校开放数据创新研究大赛**

更多信息，敬请关注：[www.huiyuan.sh.edu.cn](http://www.huiyuan.sh.edu.cn)

**参赛时间：**2019年4月~12月

**参赛形式：**基于/结合大赛数据，自选选题/根据选题指南，以研究论文为最终作品参与竞赛。

**参赛对象：**上海地区高校师生

|       |             |
|-------|-------------|
| 学术训练营 | 4月22日~5月26日 |
| 参赛报名  | 4月22日~6月3日  |
| 大赛培训  | 6月上旬        |
| 作品提交  | 6月3日~9月23日  |
| 作品评审  | 9月下旬~10月下旬  |
| 获奖公布  | 11月上旬       |
| 成果转化  | 11月上旬开始     |

大赛主图及报名二维码：

慧源共享官微：

大赛活动直播平台：

指导单位：上海市教育委员会、上海市经济和信息化委员会  
 主办单位：复旦大学、上海市教育信息中心、上海市电化教育馆  
 承办单位：慧源共享教育资源服务中心（华东大学图书馆、复旦大学图书馆、上海财经大学图书馆、上海电力大学图书馆、上海海洋大学图书馆、上海师范大学图书馆、上海外国语大学图书馆、同济大学图书馆）、中国索引学会  
 协办单位：教育部CADAL项目管理中心、上海市高校图书馆  
 数据提供：复旦大学社会科学数据研究中心、复旦大学当代中国社生活资料中心、慧源共享教育资源服务中心、上海市电化教育馆  
 媒体合作：《图书馆杂志》  
 支持合作：上海阿法通智能标签系统技术有限公司、北京汇云博图科技有限公司、北京万方数据股份有限公司、上海图盟信息科技有限公司





## 上海高校开放数据创新研究大赛（活动组织）

- Ø 指导单位：上海市教育委员会、上海市经济和信息化委员会
- Ø 主办单位：复旦大学、上海市教委信息中心、上海市电化教育馆
- Ø 承办单位（排名先后按字母顺序）：慧源共享教育资源服务中心（东华大学图书馆、复旦大学图书馆、上海财经大学图书馆、上海电力大学图书馆、上海海洋大学图书馆、上海师范大学图书馆、上海外国语大学图书馆、同济大学图书馆）、中国索引学会
- Ø 协办单位：教育部CADAL项目管理中心、上海市高校图工委
- Ø 数据提供单位：复旦大学社会科学数据研究中心、复旦大学当代中国社会生活资料中心、慧源共享教育资源服务中心、上海市电化教育馆
- Ø 媒体合作伙伴：《图书馆杂志》
- Ø 支持合作伙伴：上海阿法迪智能标签系统技术有限公司、北京汇云博图科技有限公司、北京万方数据股份有限公司、上海图趣信息科技有限公司



# 上海高校开放数据创新研究大赛（系列活动）

1

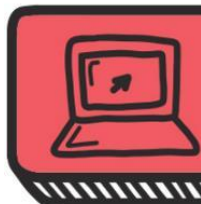
学术训练营



Ø “数据悦读”学术训练营（4.22~5.26）：邀请多领域知名数据专家，在各大高校巡回开展以数据为主题的专题讲座，同时介绍大赛情况。活动进行现场直播。

Ø “慧源共享”上海高校开放数据创新研究大赛

（4.22~6.3~10.30）：大赛提供高校图书馆业务数据集、长三角地区社会变迁调查数据、当代中国社会生活资料书信数据集（百姓家书）以及复旦大学ERU数据集、上海市中小学生学习图书馆借阅数据集、第二届学生“好问题”征集评选活动数据集（首次提问）、中小学生学习数字图书馆单页阅读记录数据集，参赛团队可自定选题或根据选题指南开展研究，参与竞赛。



数据大赛

2

Ø 成果孵化（10.31~12.31）：通过出版大赛优秀论文集、推荐发表优秀获奖论文、推荐出版优秀数据、支持优秀研究成果落地、推荐实习等途径，为参赛团队提供更多的机会和支持。

3



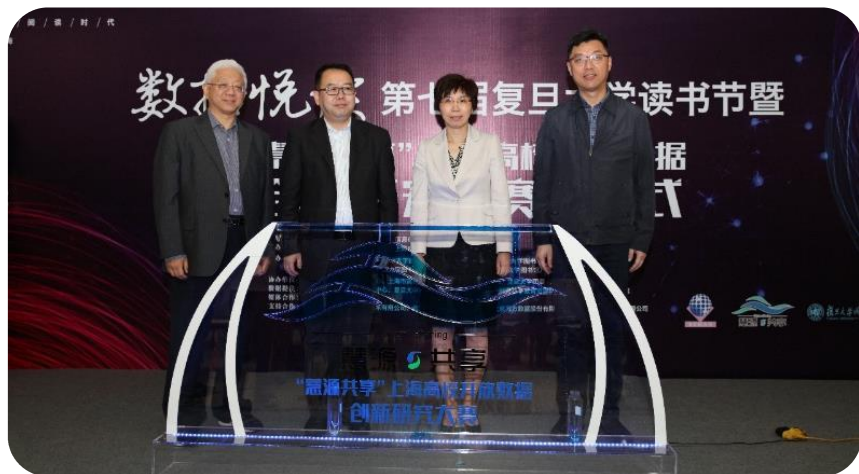
成果孵化





## 上海高校开放数据创新研究大赛（活动开幕）

在第24个“世界读书日”来临之际，4月22日下午，“数据悦读”第七届复旦大学读书节暨“慧源共享”上海高校开放数据创新研究大赛启动仪式在复旦大学举行。





## 业界发声

- 张计龙副所长带队参加哈佛大学Dataverse Community 2019年年会，年会主题为“支持、可持续和增长的数据共享”。
- 会上分享了Dataverse在不同国家/地区的应用案例，介绍了开发建设新进展，探讨了在持续增长的数据共享需求的环境下Dataverse的发展趋势。
- 殷沈琴副研究馆员作题为《Data Sharing Makes Teaching and Research Better》的报告。





## 业界发声

刘丽君研究馆员、薛崧副研究馆员和王顺箐馆员三人参加了国际图书馆协会联合会（International Federation of Library Associations and Institutions -- IFLA）在雅典召开的第85届年会，本次会议的主题是“一场关于变革带来图书馆变化的对话”。会上刘丽君研究馆员和王顺箐馆员分别做了报告《Helping Legal Aid Providers and Vulnerable Communities Access Law in China through a Robust Legal Information Services System》和《Inheritance and Innovation: Local History Collections in Academic Library》





PART

04

# 总结与展望





## 不忘初心，砥砺前行

- 科学数据的价值及数据管理与共享开放的意义更加明晰，政策导向更加清晰。
- 数据管理与数据开放的理念逐步落地实践，新兴技术提供持续助力，推动更好发展。
- 数据权益和数据安全依然是多方关注的焦点，合作与共赢是不变的发展之道。
- 不论科研领域，不论利益群体，不论地区机构，未来更大的发展需要多方为了共同的目标，不忘初心，砥砺前行。





## 主要参考文献

- CODATA中国全国委员会. 大数据时代的科研活动[M]. 科学出版社, 2014.
- 黄欣荣. 大数据促进科学与人文融合[N]. 中国社会科学报, 2018-09-11(005).
- 张玉娥, 王永珍. 欧盟科研数据管理与开放获取政策及其启示——以“欧盟地平线2020”计划为例[J]. 图书情报工作, 2017, 61(13):70-76.
- 郝世博, 徐文哲, 唐正韵. 科学数据共享区块链模型及实现机理研究[J]. 情报理论与实践.
- 张丽丽, 温亮明, 石蕾, 等. 国内外科学数据管理与开放共享的最新进展[J]. 中国科学院院刊, 2018(8).
- 张恬, 刘凤红. 数据出版新进展[J]. 中国科技期刊研究, 2018, 29(5).
- 濮静蓉, 刘桂锋, 钱锦琳. 国内外开放数据的研究进展与述评[J]. 图书与情报, 2017(6).
- 桂畅旒. 欧盟《通用数据保护法案》的影响与对策[J]. 中国信息安全, 2017(7).
- 邢文明, 刘婷. 增强出版驱动的科学数据出版: 动因、模式及路径[J]. 中国科技期刊研究, 2019, 30(08):853-861.
- <https://www.force11.org/group/fairgroup>
- [http://www.gov.cn/xinwen/2018-04/02/content\\_5279353.htm](http://www.gov.cn/xinwen/2018-04/02/content_5279353.htm)
- <http://news.ynet.com/2019/08/28/2042998t70.html>
- [https://mp.weixin.qq.com/s/q-i\\_szKupSK5SdHY2r4VHw](https://mp.weixin.qq.com/s/q-i_szKupSK5SdHY2r4VHw)
- 其他文献资源、内部教材、内部资料、团队文章
- 部分图片来自网络



Thank You