



# 一流高校研究数据管理服务体系的研究与建设

——以北京大学图书馆为例

□ 崔海媛\* 罗鹏程 李国俊 朱玲

**摘要** 通过调研国内外数据驱动研究现状、趋势和用户需求,分析研究数据管理服务的国内外发展情况,研究提出高校研究数据管理服务体系,并介绍北京大学研究数据管理服务的研究成果、框架设计、构建经验与服务效果,为高校建设一流研究数据管理服务、支持用户数据驱动研究范式需求,提供研究与实践经验。以促进数据驱动研究,推动开放科学。

**关键词** 数据驱动研究 研究数据管理 数据服务体系 数据服务 数据政策

**分类号** G251

**DOI** 10.16603/j.issn1002-1027.2019.02.008

## 1 引言

随着云计算、移动互联网的应用普及,以及可穿戴智能设备的出现,一场全新的、以大规模数据生产、分享、使用为代表的技术革命正在发生,数据成为宝贵的资源、资产和生产要素。对海量数据及其隐含信息、知识的收集、分析、挖掘、有效整合,将为科学研究、社会经济发展带来巨大的红利。大数据的广泛应用开启了一个全新的智能时代<sup>[1]</sup>。2017年,以谷歌阿尔法狗(AlphaGo)和阿尔法元(AlphaZero)为代表,深度学习技术取得突破,标志着人工智能元年到来。大数据和人工智能时代,世界各国对数据的依赖快速上升,抢占大数据发展先机,已成为国家战略。美国、日本、韩国、中国、欧盟等国家和地区纷纷将大数据利用上升为国家战略,积极推进数据政策,希望利用大数据技术实现研究和产业领域的突破<sup>[2]</sup>。

数据科学已经成为全球高校发展最快的专业。数据科学社区(Data science Community)的数据显示,截至2018年12月,全球共有26个国家的361所高校开设了597个数据科学相关学位专业,其中美国高校开设的数据科学专业数量最多。全美有247所高校共设置461个数据科学学位专业,占全球总量的77%<sup>[3]</sup>。在中国,2016年2月,北京大学、

对外经济贸易大学、中南大学首次成功申请到“数据科学与大数据技术”本科新专业。2017年3月,第二批32所高校获批。2018年3月,教育部公布的高校新增专业名单中,有248所学校获批,是过去两次审批通过额度的近8倍<sup>[4]</sup>。

2017年11月24日至2018年4月3日,由北京大学图书馆和北京大学信息管理系发起与组织,多家机构合作,成功举办了“首届全国高校数据驱动创新研究大赛”(以下简称大赛)。大赛吸引了全国众多高校学生积极参与。比赛期间,有近4万用户访问大赛主页,共吸引来自北京大学等全国169所高校的1892名同学报名参赛,共有593支队伍,涉及56个一级学科。最终,来自121所高校的968人(共289支队伍)成功提交了参赛作品,涉及45个一级学科。成功提交作品最多的前20个一级学科中,图书馆、情报和档案学科队伍最多,达到59组,其次是应用经济学、社会学、管理科学与工程、计算机科学、统计学也分别达到50组、27组、23组、19组、19组。在通过形式审核的233篇论文中,有103篇论文使用平台已有数据,135篇论文使用自产的原创性数据,同时使用平台已有数据和自己原创数据的共有5篇。参赛队伍使用数据情况与笔者前期调研情况表明:数据已是绝大多数学科研究的基础,研

\* 通讯作者:崔海媛,ORCID:0000-0001-5541-7100,邮箱:cuihy@lib.pku.edu.cn。



研究者需要高质量的大数据提交、发布和研究服务平台;高质量调查数据、招聘数据与社交媒体数据仍是人文社科经济学领域研究者重点关注的研究数据。

随着数据量与数据需求的飞速增长,大数据和海量数据的管理、共享、发布和保存面临着巨大的成本与效率考验。研究数据的存储形式、存储容量与存储效率制约着研究使用数据规模与利用形式,成为研究数据管理的挑战。重视与发展研究数据的管理与服务,已成为国家层面与科研机构未来提升学术竞争力和影响力的重要发展战略之一。2018年4月2日,国务院国务院办公厅关于印发科学数据管理办法的通知,以研究数据全生命周期为中心,加强和规范研究数据的管理,推动开放共享<sup>[5]</sup>。顺应教学和研究需求,建设一流的研究数据管理服务体系,提供研究数据支撑服务,成为全球研究型图书馆的必然选择。

本文将调研国内外研究数据管理服务的发展情况,提出建设研究数据管理服务体系的建议。并以北京大学研究数据管理服务体系建设经验为例,为图书馆推动研究数据管理服务提供参考。

## 2 研究综述

原中国科学院图书馆馆长张晓林指出,在全学科数据驱动研究成为趋势的时代,图书馆需要重新定义服务目标、内容、方式和能力,在大数据的基础上,基于新技术和方法,提供全新知识服务。美国麻省理工学院、美国国家医学图书馆和英国大英图书馆等都已经调整战略,融汇数据资源,重构知识服务的平台、机制<sup>[6]</sup>。原上海图书馆馆长吴建中通过对20年来全球知识开放运动的梳理和分析,提出研究数据管理对推进知识开放和共享的重要意义,并提出4点建议:①加大宣传力度,增强开放共享意识;②制定相关政策,推进开放科学发展;③整合各类资源,建设数字基础设施;④培育专业人才,提升数据管理水平<sup>[7]</sup>。

国际上,研究数据作为科学研究的重要成果受到国际学术界和出版界越来越多的重视。2009年以来,美国、英国、澳大利亚、欧盟等国家和地区纷纷出台开放数据政策。2011年,美国自然科学基金委要求项目申请书必须包含“数据管理计划”<sup>[8]</sup>,美国多个资助机构的数据管理政策陆续发布。欧盟与欧洲资助机构出台一系列数据政策。2011年,英国联合信息系统委员会(Joint Information Systems Committee)和电子科

学核心计划(E-Science Core Programme)联合组建的“国家数据监管中心(National Digital Curation Center)”发布《制定数据管理与共享计划》<sup>[9]</sup>。英国皇家学会(Royal Society)于2012年7月发布报告《科学是开放事业》(Science as an Open Enterprise),确立共享和开放科学数据的基本原则、机遇与面临的挑战<sup>[10]</sup>。欧盟2020计划规定,2017年开始全面实施科研数据开放制度,推动“开放科学”战略<sup>[11]</sup>。截至2018年6月,在资助机构开放获取政策查询数据库SHERPA/JULIET中,144个资助机构,41个资助机构政策中要求必须开放研究数据,19个机构鼓励开放<sup>[12]</sup>。

在各种政策驱动下,研究数据管理仓储发展迅速。截至2018年12月,根据国际研究数据仓储注册组织(Registry of Research Data Repositories)的统计,全世界至少有2250多个研究数据仓储<sup>[13]</sup>。哈佛大学等世界顶尖高校均建立了自己的数据仓储。科学数据出版期刊不断涌现,2011年2月,美国《科学》(Science)刊登“数据处理(Dealing with Data)”专题<sup>[14]</sup>,Wiley公司于2012年推出了期刊《地球科学数据》(Geoscience Data Journal)<sup>[15]</sup>,自然集团在2014年推出期刊《科学数据》(Scientific Data)<sup>[16]</sup>,中国科学院也于2016年推出期刊《中国科学数据》<sup>[17]</sup>,研究数据逐渐成为出版领域抢占的制高点。数据引用及引文数据库也得到越来越多的应用,原汤森路透公司于2012年在Web of Science中推出了“数据引用索引”(Data Citation Index),用于数据的检索和引用跟踪。

国家、基金会和机构数据政策推动数据管理服务需求,研究数据管理(Research Data Management, RDM)成为研究型图书馆的重要服务,相关研究与实践发展迅速。

卡罗·泰诺普(Carol Tenopir)等在2016年2月向333个欧洲大学图书馆发送邮件调查研究数据管理服务的开展情况,共有119个图书馆回复邮件。90%以上的图书馆都已经开展或计划开展研究数据的管理计划制定、培训、开发等服务,31.9%的图书馆已经参与到研究项目中,提供数据服务<sup>[18]</sup>。阿杨·尹(Ayoung Yoon)等在2015年10月—12月期间,对185个美国研究型大学图书馆的网站进行调查分析后,发现美国研究型图书馆开展的研究数据管理服务主要包括:咨询帮助(help/ask librarians,



79.5%)、数据管理(Data deposit, 60%)、研究数据管理计划(Data management planning, 41.1%)、数据咨询(Data consultation, 38.4%)、数据出版和共享(data publishing and sharing, 27.6%)、数据方法(data methodology, 15.7%)<sup>[19]</sup>。

2017年3月,美国联机计算机图书馆中心(Online Computer Library Center, OCLC)发布“研究数据管理现实”系列报告,报告以12个北美、欧洲和澳大利亚的图书馆的研究为基础,重点对爱丁堡大学(英国)、伊利诺伊大学厄巴纳-香槟分校(美国)、蒙纳士大学(澳大利亚)和瓦赫宁根大学(荷兰)等4所大学的服務进行案例分析,研究这些机构的研究数据管理能力,总结研究数据管理服务为三大领域:教育类、解决方案类和长期保存类<sup>[20]</sup>。

澳大利亚国际数据服务中心(The Australian National Data Service)提出数据管理能力的5个要素:政策和流程、信息架构、服务能力、元数据管理和研究数据管理。并给出五个发展阶段的成熟度<sup>[21]</sup>。伊莲·马丁(Elaine R. Martin)提出基于图书馆以用户为中心开展数据管理服务的框架,主要涵盖:数据服务、数据管理实践指导、信息素养、存档保存和政策服务<sup>[22]</sup>。英国的数字策展中心(The Digital Curation Centre)根据研究生命周期,提出研究数据管理支持服务框架应涵盖:政策和规划、可持续性发展、研究数据管理计划、元数据管理、数据存储、数据存档策略、数据管理、培训指南等<sup>[23]</sup>。于飞(Fei YU)等(2017年)介绍了澳大利亚昆士兰大学在对大学用户进行调查后,设计与开展基于生命周期的研究数据管理服务内容包括:①研究数据管理计划和准备(数据指南、数据培训、研究数据管理计划等);②研究过程支持服务(数据工具、管理和分析、培训等);③数据保存、出版和传播(数据发布平台、长期保存和数据出版等)。见图1<sup>[24]</sup>。

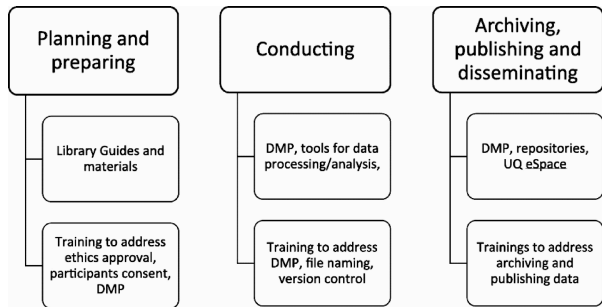


图1 澳大利亚昆士兰大学研究数据管理服务内容

### 3 研究数据管理服务体系的设计与建设

2015年,北京大学图书馆制定2018行动计划,调整组织机构,重新组织与设计服务,将原系统部更名为信息化与数据中心,设置数据管理与服务岗位,全面开展研究数据管理平台、工具和服务建设,设计数据管理服务框架。另外组建研究支持中心,设置数据馆员,开展信息素养和数据咨询等服务。数据资源成为资源建设的重要组成部分。2015年12月25日,北京大学图书馆正式发布开放研究数据服务平台(以下简称“开放数据平台”)、机构知识库、开放出版一期刊网和学者主页,构建新型学术交流生态系统,推动开放获取与研究数据管理服务,成为国内高校相关建设的引领者。并在开放数据平台的基础上,开展平台、服务、培训、政策制定与推进等一系列研究和服务工作,推动北京大学建设成为研究数据汇集、管理与服务的引领者。

#### 3.1 需求调研

2013年起,笔者团队从调研开始,定位研究者需求,掌握国际发展动态,开发数据服务平台,全面推广服务,完成北京大学研究数据管理服务框架的初步构建,为未来研究和服务奠定基础。图2列出北京大学研究数据管理服务的建设历程。

2014年,笔者团队对校内教师进行问卷调查,获取50份教师的反馈,并对26名教师进行一对一需求访谈。调研发现:高达87.5%的受访者愿意在一定条件下开放研究数据,开放的最大动机是通过开放数据,提升数据价值、提高成果引用率与增强数据曝光度。开放的最大疑虑是他人利用数据抢先发表成果。开放条件中,受访者最需要“审核使用者身份和数据使用用途后提供”功能。通过调研,准确定位研究者对研究数据管理服务的实际需求。

一对一访谈发现:①数据开放程度与学科特点密切。对于数据驱动、数据密集特征明显的学科,数据开放比较标准和规范,可以推进到原始数据的程度,例如社会调查学科和生物信息学科。在生物信息学科,发表期刊时发表数据已经是趋势,大多向期刊指定数据库提交,或者自己建立数据库。对于其他学科,数据开放也存在,但向大多数同行开放的仅是整理后的成果数据,原始数据只与直接合作者共享。②数据开放具有明显的时滞期。几乎所有的学者都强调,数据必须在自己的成果发表之后,才考虑开放。这与学者对“数据被他人抢先发表成果”的担





心是一致的。③研究数据开放行为多为学者自发,缺乏政策激励机制和维护更新平台,缺乏基于数据引用、认可和学术声誉的良性反馈,亟需规范的数据服务平台。

2017年,笔者对校内本科生与研究生的课程调研发现:全校约26个院系开设数据驱动研究的研究课程约有56门,本科生课程有85门,不仅分布于理工学部,也分布于人文与社会科学学部。对9名课程教师进行访谈,调研发现:教师课程教学急需高质量开放研究数据平台。教师获取数据的难度远远

超过获取文献。

2018年,笔者对国外27所大学(美国12所、英国11所和澳大利亚4所)的研究数据管理服务主页进行调研,梳理国内外研究机构和大学的研究数据管理服务内容。调研结果见表1,可以发现研究数据管理服务内容主要包括:数据管理计划、数据服务、数据素养、数据管理、长期保存和数据政策几部分。围绕研究和研究数据生命周期,提供全流程的数据管理服务。

表1 国外27所大学研究数据管理服务内容的调研结果

|      |           | 数据管理计划   | 数据服务   |      |      |      | 数据素养     |        |            |        |        | 数据管理  |      |     | 长期保存 |         |         |        | 数据政策    |
|------|-----------|----------|--------|------|------|------|----------|--------|------------|--------|--------|-------|------|-----|------|---------|---------|--------|---------|
|      |           | 数据管理计划指南 | 数据需求咨询 | 数据共享 | 数据引用 | 数据版权 | 数据生命周期介绍 | 数据政策指南 | 数据管理计划工具培训 | 工具软件培训 | 数据仓储指南 | 元数据描述 | 文件格式 | 文件名 | 数据存储 | 机构的数据仓储 | 数据安全与备份 | 软件代码保存 | 机构的数据政策 |
| 美国   | 哈佛大学      | ✓        | ✓      |      | ✓    |      |          | ✓      | ✓          |        | ✓      |       | ✓    | ✓   | ✓    | ✓       |         |        |         |
|      | 麻省理工学院    | ✓        | ✓      |      |      |      |          | ✓      | ✓          | ✓      |        | ✓     | ✓    | ✓   | ✓    |         |         |        |         |
|      | 约翰霍普金斯大学  | ✓        | ✓      | ✓    | ✓    | ✓    |          | ✓      |            | ✓      | ✓      | ✓     |      |     | ✓    | ✓       | ✓       | ✓      | ✓       |
|      | 斯坦福大学     | ✓        | ✓      | ✓    | ✓    | ✓    |          |        | ✓          | ✓      | ✓      | ✓     | ✓    | ✓   | ✓    | ✓       | ✓       |        |         |
|      | 西北大学      | ✓        | ✓      |      | ✓    |      |          | ✓      |            |        | ✓      | ✓     |      |     |      | ✓       |         |        |         |
|      | 哥伦比亚大学    | ✓        | ✓      | ✓    | ✓    | ✓    | ✓        | ✓      |            |        | ✓      | ✓     | ✓    | ✓   | ✓    | ✓       | ✓       |        | ✓       |
|      | 康奈尔大学     | ✓        | ✓      | ✓    | ✓    | ✓    |          | ✓      |            |        |        | ✓     | ✓    | ✓   | ✓    | ✓       | ✓       |        |         |
|      | 耶鲁大学      | ✓        | ✓      |      | ✓    |      |          |        |            | ✓      | ✓      | ✓     |      |     |      |         |         |        |         |
|      | 新罕布什尔大学   | ✓        |        |      | ✓    |      |          | ✓      |            |        |        | ✓     |      | ✓   |      |         | ✓       |        | ✓       |
|      | 加州大学伯克利分校 | ✓        | ✓      | ✓    | ✓    | ✓    |          |        |            |        |        | ✓     |      |     | ✓    | ✓       | ✓       |        |         |
|      | 明尼苏达大学    | ✓        | ✓      |      | ✓    | ✓    |          | ✓      | ✓          | ✓      | ✓      | ✓     |      | ✓   |      | ✓       |         |        |         |
|      | 弗吉尼亚大学    | ✓        | ✓      | ✓    | ✓    | ✓    | ✓        | ✓      | ✓          | ✓      | ✓      | ✓     | ✓    | ✓   | ✓    | ✓       | ✓       |        |         |
| 英国   | 牛津大学      | ✓        | ✓      | ✓    | ✓    | ✓    |          | ✓      |            | ✓      |        | ✓     |      | ✓   | ✓    | ✓       | ✓       |        | ✓       |
|      | 剑桥大学      | ✓        | ✓      | ✓    | ✓    |      |          | ✓      |            | ✓      |        | ✓     | ✓    | ✓   |      | ✓       |         |        | ✓       |
|      | 伯明翰大学     | ✓        | ✓      |      |      |      |          |        |            |        |        | ✓     | ✓    | ✓   |      | ✓       |         |        | ✓       |
|      | 爱丁堡大学     | ✓        | ✓      |      |      |      |          | ✓      |            | ✓      |        |       |      |     |      | ✓       |         |        | ✓       |
|      | 伦敦国王学院    | ✓        | ✓      | ✓    | ✓    | ✓    | ✓        | ✓      |            | ✓      |        | ✓     | ✓    | ✓   | ✓    | ✓       | ✓       |        |         |
|      | 利物浦大学     | ✓        | ✓      |      | ✓    | ✓    |          | ✓      |            |        | ✓      |       | ✓    | ✓   |      |         |         |        |         |
|      | 曼彻斯特大学    | ✓        |        | ✓    | ✓    |      |          | ✓      |            |        |        | ✓     | ✓    | ✓   | ✓    | ✓       | ✓       |        |         |
|      | 诺丁汉大学     | ✓        |        |      | ✓    |      | ✓        |        |            |        |        | ✓     |      | ✓   |      |         | ✓       |        |         |
|      | 伦敦大学学院    | ✓        |        | ✓    | ✓    | ✓    |          | ✓      |            |        |        |       | ✓    |     | ✓    | ✓       |         | ✓      | ✓       |
|      | 约克大学      | ✓        |        | ✓    | ✓    | ✓    |          | ✓      |            |        |        | ✓     | ✓    | ✓   | ✓    |         |         |        | ✓       |
|      | 谢菲尔德大学    | ✓        |        | ✓    | ✓    | ✓    |          | ✓      |            |        | ✓      | ✓     | ✓    | ✓   |      | ✓       | ✓       |        | ✓       |
| 澳大利亚 | 昆士兰科技大学   | ✓        |        |      |      | ✓    |          | ✓      |            |        |        | ✓     | ✓    | ✓   | ✓    | ✓       |         |        | ✓       |
|      | 莫纳什大学     | ✓        |        |      |      | ✓    |          | ✓      |            |        | ✓      | ✓     | ✓    | ✓   | ✓    | ✓       | ✓       |        | ✓       |
|      | 墨尔本大学     | ✓        | ✓      |      |      |      |          | ✓      | ✓          | ✓      |        |       |      |     |      |         |         |        | ✓       |
|      | 纽卡斯尔大学    | ✓        |        |      | ✓    |      |          | ✓      |            | ✓      | ✓      | ✓     | ✓    | ✓   | ✓    |         | ✓       |        |         |



### 3.2 服务体系设计

通过全面了解国际进展,调研用户需求,并与院系合作,调研数据管理需求活跃的若干典型院系、课题组或个人,发掘和把握研究者对研究数据管理和再利用的切实需求,探索最佳协作机制。建立与北京大学中国社会科学调查中心、科学研究部、社会科学部和部分院系教师的合作机制,共同建设研究数据管理服务。笔者团队设计提出了研究数据管理体系(表2)。并在服务体系下,建设和推动研究数据管理服务(建设历程见图2)。

表2 北京大学研究数据管理体系

| 数据服务         | 数据管理            | 数据素养   | 长期保存       | 数据政策                |
|--------------|-----------------|--------|------------|---------------------|
| 数据服务主页(建设中)  | 数据管理工具软件服务(建设中) | 数据素养培训 | 框架体系(建设中)  | 研究数据管理政策制定与发布(建设中)  |
| 开放数据平台(发布)   | 数据管理私有云服务(建设中)  | 工具软件使用 | 元数据标准      | 知识产权培训              |
| 数据需求咨询(数据馆员) | 数据管理平台(建设中)     | 研究培训   | 存储保存       | 政策培训(国家、资助机构、机构、期刊) |
| 数据驱动研究大赛     |                 | 写作培训   | 安全备份       | 政策咨询                |
|              |                 | 数据资源导航 | 源代码保存(建设中) |                     |

### 3.3 建设过程

#### (1) 数据建设

在数据建设方面,截至2018年12月,开放数据平台发布数据空间37个、数据集236个、数据文件1096个,收录中国家庭追踪调查、中国健康与养老追踪调查、综合型语言知识库、中国历代人物传记资料库等多个高质量、具有国际影响力的数据集。

(2) 比较与分析不同研究数据发布平台,推出符合北京大学需求的开放数据平台

在调研与测试多个软件平台与工具的基础上,确定应用开源软件Dataverse建设北京大学研究数据开放服务平台,开发实现中英双语界面、校统一认证、用户审核管理、DOI注册、数据统计等新功能,满足数据提交、管理与服务需求。实现数据成果交互与揭示,数据平台与机构知识库实现互操作,将基于平台数据集的研究成果通过机构知识库进行保存与服务。推动国际平台交互与揭示,数据平台与国际数据仓储Re3data、国际数据引用索引平台(Data

Citation Index)实现数据交互、收录与揭示,提高数据平台国际可见度。

建设云存储、云笔记、唯一标识符管理系统、长期保存框架等相关平台和服务。在提供开放数据服务的基础上,研究与应用项目管理工具、平台服务。

#### (3) 全面开展北京大学研究数据管理服务

以合作促发展,通过建立校内多部门、多院系合作机制,为课程、项目与教师提供持续服务,并与院系、教师课题组、政府、企业建立数据合作关系,立足高校实际需求,与国际同步发展,建立良性、可持续发展机制。

(4) 推动建设北京大学研究数据长期保存与共享服务机制

调研研究数据管理服务政策、长期保存与共享服务机制,将数据政策研究纳入到北京大学开放获取政策推动项目中。推出北京大学数据唯一标识符服务,提供数据的长期保存与国际数据中心互操作服务。

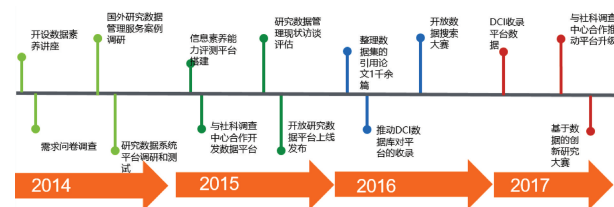


图2 北京大学研究数据管理服务的建设历程

#### (5) 服务营销

在中国研究数据相关政策滞后于国际的情况下,如何让更多研究者了解、参与和受益于研究数据管理服务,成为一项挑战。笔者团队通过策划多种方案,推动研究数据管理服务建设。

首先,通过营销提升影响力。

对开放数据平台进行搜索引擎优化(Search Engine Optimization),改进与提高搜索引擎排名,优化后,平台搜索进入百度和谷歌搜索结果第一。在北京大学主页科学研究栏目下发布开放数据平台链接,增加平台入口。通过北京大学37个院系图书馆分馆发放研究数据管理宣传册,并在校图书馆管理委员会年度会议、科研秘书年度会议和在国内学术会议上广泛宣传北京大学研究数据管理服务。与国际数据仓储Re3data和国际数据引用索引平台Data Citation Index合作,将平台数据在国际平台上收录与揭示,增强国际可见度,扩大数据的国际影



响力。

其次,服务营销案例一首届全国高校数据驱动研究大赛。

为广泛宣传研究数据管理服务,2016年11月7日至12月16日,北京大学图书馆举办第三届搜索达人比赛——“开放数据”篇。大赛共吸引300多名师生参与,帮助师生拓展开放数据视野,提高开放数据搜索和分析的实践技能,激发利用开放数据进行教学科研的热情,以更好地开展科学数据管理与利用相关服务、助力教学科研、推动科学数据开放共享。

2017年12月至2018年4月,由北京大学图书馆发起,联合校内外机构,成功举办“全国高校数据驱动创新研究大赛”。大赛吸引全国高校万余人的关注,在报名阶段,共吸引来自全国160多所高校近600支队伍(近2000名选手)参赛报名。最终由来自全国120余所高校的289支队伍(近1000人)成功提交参赛作品。评选出13支现场答辩队伍和120支优秀队伍。2018年4月3日下午,答辩活动在北京大学举行,现场评选出特等奖、一等奖、二等奖和三等奖共计十三项大奖。大赛得到广泛关注,全国数十所高校发布相关信息,主流媒体进行报道,参赛者对大赛极为重视,提交的高质量、有创新研究成果的参赛作品,得到评审专家们的充分肯定。

大赛达到了营销开放数据平台的目标。在大赛举办之前,平台的日均访客数为52人,大赛宣传报名期间开放数据平台日均访客数为753人,大赛结束后日均访客数为477人。通过大赛,平台日均访问量增加了近10倍,大赛期间,数据下载量达到过去两年总和的5倍,新增用户数量达到过去两年总和的4倍。

#### 3.4 未来规划

抓住大数据发展机遇,将北京大学建设成为研究数据汇集、管理与服务的世界一流引领者,已成为北京大学图书馆的重要发展目标之一。北京大学图书馆未来将在以下方面推进研究数据管理:加强在数据工具、软件、管理与服务等方面的研究与应用,以数据服务为凝聚力,建设持续汇集国内外高质量研究数据的研究数据平台,提供基于数据的教学与研究的支撑平台;开展数据创新研究、教学与产业孵化活动与比赛;与校外数据资源丰富的政府、企事业单位合作,寻求互补共赢的合作模式;推动数据教学

与科研创新。

在数据服务和数据管理平台建设方面,将通过以下方式加快发展:

- 为教学提供支撑服务,促进与课程相关的数据和论文的保存、积累。

- 融入教师的科研过程,通过数据馆员与教师展开深入合作,推进研究数据的全生命周期管理。

- 推动北京大学制定研究数据管理与服务政策,推动研究数据的收集、保存、管理和共享。

- 优化平台性能与功能,发展为国际一流数据平台。

- 提升平台的数据量承载能力,构建基于大数据、高并发的数据管理系统,促进平台对生物、医学、社会科学等学科大数据的支撑。

- 提高数据平台与其他系统的互操作性,促进数据与论文成果、学者主页、项目成果的关联,加速数据的流动和共享。

#### 4 结语

在大数据和人工智能环境下,研究数据呈指数级增长,数据驱动研究成为全学科研究态势,数据管理和需求日益迫切。2018年12月2日—4日,在第14届开放获取柏林会议上,中国国家自然科学基金委、国家科技图书文献中心、中科院文献情报中心代表在会议上发布立场声明,明确表示中国支持欧盟倡议的OA2020和开放获取S计划,支持公共资助项目研究论文立即开放获取<sup>[25]</sup>。S计划包括研究数据的开放获取,此声明意味着中国的开放数据政策可能也即将出台。对研究型图书馆来说,这既是挑战,也是机遇。图书馆应重视研究数据管理服务,助力中国高等教育的“双一流”建设,把握研究数据管理服务发展契机,加快研究与实践,成为研究数据汇集、管理与服务的引领者。

#### 参考文献

- 黄晓艳,马珉.大数据开启智能时代——访中国科学院院士鄂维南[J].高科技与产业化,2017,(6):36—41.
- 张影强,张大璐,梁鹏.发达国家如何布局大数据战略[J].中国经济报告,2018,(1):87—89.
- College & university data science degrees[EB/OL]. [2018-12-20]. <http://datascience.community/colleges>.
- 283所高校获批数据科学与大数据专业[EB/OL]. [2018-05-10]. <http://36kr.com/p/5125134.html>.



- 5 国务院办公厅关于印发科学数据管理办法的通知[EB/OL]. [2018-05-01]. [http://www.gov.cn/zhengce/content/2018-04/02/content\\_5279272.htm](http://www.gov.cn/zhengce/content/2018-04/02/content_5279272.htm).
- 6 张晓林. 颠覆性变革与后图书馆时代——推动知识服务的供给侧结构性改革[J]. 中国图书馆学报, 2018, 44(1): 4-16.
- 7 吴建中. 推进开放数据 助力开放科学[J]. 图书馆杂志, 2018, 37(2): 4-10.
- 8 Nsf data sharing policy[EB/OL]. [2018-10-01]. <https://www.nsf.gov/bfa/dias/policy/dmp.jsp>.
- 9 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(8): 5-11.
- 10 Science as an open enterprise[EB/OL]. [2018-12-02]. <https://royalsociety.org/topics-policy/projects/science-public-enterprise/report/>.
- 11 COMMISSION E. Open access & data management[EB/OL]. [2018-07-10]. [http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-dissemination\\_en.htm](http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-dissemination_en.htm).
- 12 Research funders' open access policies[EB/OL]. [2018-05-31]. <http://v2.sherpa.ac.uk/juliet/>.
- 13 Re3data. Org [EB/OL]. [2018-12-20]. <https://www.re3data.org/>.
- 14 Special online collection: Dealing with data[EB/OL]. [2018-06-19]. <http://www.sciencemag.org/site/special/data/#opennewwindow>.
- 15 Geoscience data journal [EB/OL]. [2018-06-19]. <https://rmets.onlinelibrary.wiley.com/journal/20496060>.
- 16 Welcome, scientific data[EB/OL]. [2018-06-19]. <https://www.nature.com/news/welcome-scientific-data-1.15293>.
- 17 中国科学数据[EB/OL]. [2018-06-19]. <http://www.csdata.org/p/static/33/>.
- 18 Tenopir C, Talja S, Horstmann W, et al. Research data services in european academic research libraries [J]. *Liber Quarterly*, 2017, 27(1): 23-44.
- 19 Yoon A, Schultz T. Research data management services in academic libraries in the us: A content analysis of libraries' websites [J]. *College & Research Libraries*, 2017, 78(7): 920-933.
- 20 Bryant R, Lavoie B, Malpas C, et al. A tour of the research data management (rdm) service space. The realities of research data management, part 1 [J]. 2017.
- 21 Creating a data management framework[EB/OL]. [2018-06-07]. <https://www.ands.org.au/guides/creating-a-data-management-framework>.
- 22 Martin E R. The role of librarians in data science: a call to action [J]. *Journal of eScience Librarianship*, 2016, 4(2): 7-9.
- 23 How to develop rdm services—a guide for heis[EB/OL]. [2018-05-01]. <http://www.dcc.ac.uk/resources/how-guides/how-develop-rdm-services>.
- 24 Yu F, Deuble R, Morgan H. Designing research data management services based on the research lifecycle - a consultative leadership approach [J]. *Journal of the Australian Library and Information Association*, 2017, 66(3): 287-298.
- 25 张晓林. 让所有科研论文免费阅读,中国机构明确力挺开放获取[EB/OL]. [2018-12-20]. <http://zhishifenzi.com/depth/depth/4778.html>.

作者单位:北京大学图书馆,北京,100871

收稿日期:2018年7月7日

## Research and Implementation on Research Data Management Services Supporting the Data Driven Research Demand

—Taking Peking University as an Instance

Cui Haiyuan Luo Pengcheng Li Guojun Zhu Ling

**Abstract:** Based on comprehensive literature review, this article analyzes the situation and trends of Data Driven Research and Research Data Management and then discusses issues of implementation of Research Data Management Services to support Data Driven Research. By generalizing Research Data Management Services, it provides the framework of Research Data Management Service to Academic Universities of China, introduces the case of Peking University Research Data Management Services. It is hoped to help academic libraries to promote the development of Research Data Management Service and benefits for the improvement of open science and research innovation in China.

**Keywords:** Data Driven Research; Research Data Management; Research Data Management Service Framework; Data Service; Data Policy