

高校图书馆科研数据管理与服务探讨*

李秋实

(武汉大学信息管理学院, 武汉 430072)

摘要: 立足高校图书馆开展科研数据管理服务的现实需求, 主要从政策支持、技术支撑和服务推广等方面对科研数据管理服务加以分析, 并结合国外图书馆实践加以印证, 最后针对我国高校图书馆科研数据管理服务的现状, 提出建议, 以完善相关方面的研究并提供借鉴。

关键词: 科研数据; 数据管理; 高校图书馆

中图分类号: G250

DOI: 10.3772/j.issn.1673-2286.2015.05.011

1 引言

大数据环境下, 随着e-science时代的到来, 科研数据管理 (Scientific Data Curation) 正迅速成为信息科学、图书情报领域关注和研究的热点。2002年8月, 微软首席研究员、计算机图灵奖获得者Jim Gray, 首次提出了Data Curation的概念^[1]。2011年, 科研数据管理列入欧洲数字图书馆理论与实践国际会议 (TPDL) 的研究主题之一, 英国成立数据管理中心 (Digital Curation Center, DCC) 主办专业期刊和学术年会^[2]。2013年3月全球研究数据联盟 (Research Data Alliance, RDA) 得以组建^[3]; 2013年9月在意大利阿纳尼召开科研数据共享的跨学科研讨会 (Scientific Data Sharing: An Interdisciplinary Workshop) ^[4]。由此可见, 十余年来, 数据管理作为一个新兴研究领域不断发展成熟, 受到了广泛关注。

现阶段已有众多研究成果对科研数据管理的相关概念进行界定。2005年9月, 美国国家科学基金会在提供数据监护服务的数据网 (Data Net) 计划中, 将“数据”定义为“所有能以数字化形式存储并能以电子方式获取的信息, 包括数字、文本、出版物、视频、音频、

算法、软件、模型、模拟、图像等”^[5]。杨鹤林认为, 科研数据即科学研究中通过测算、计量、观察、访谈、调查、设计、建模等方法获得的, 并能以现代信息技术保存和获取的记录^[6]。英国联合信息系统委员会 (Joint Information Systems Committee, JISC) 在2003年的相关报告中提出, “数据管理”是指从数据产生开始即对其管理, 以促进数据的利用, 确保数据能够在需要时被再发现和利用^[7]。英国数据管理中心 (Digital Curation Center, DCC) 对数据监护过程进行了详细研究, 创建了数据监护生命周期模型, 将其分成描述和表示信息、建立存储计划、对科研组织活动的观察和参与、数据监护和保存四个阶段^[8], 而澳大利亚国家数据服务项目 (Australian National Data Service, ANDS) 将数据监护分成8个动作: 创造、存储、描述、标识、登记、获取、检索、开发^[9]。2006年美国研究图书馆协会 (Association of Research Libraries, ARL) 在其报告中将“科研数据管理”定义为对保存数据的组织、展示和再利用的方式^[10]。

在开展理论研究的同时, 美国、英国、澳大利亚等许多国家在科研数据管理的实践领域做出了实质性的探索成果, 特别是在高校图书馆。相比其他类型的图书馆,

* 本研究得到教育部人文社会科学研究青年项目“政府信息资源公益性增值利用对策研究” (编号: 11YJC870025) 资助。

高校图书馆更容易接触到各个学科的科研数据,与研究 人员建立密切的联系,成为高校的数据监管中心,并通过 搭建开放、共享的数据平台来实现高校科研数据的有效 管理和共享,更好地为科研人员提专业化的服务^[11]。

结合已有的研究成果,本文将高校图书馆科研数据 管理服务概括为:图书馆专业数据管理人员为研究 人员提供的基于科研数据生命周期的嵌入式全方位的 跟踪服务,通过对科研数据的收集、管理、保存、发布、 共享等一系列的活动总和,从而实现科研数据长久利用、 再利用,维持科研数据价值,为科学研究提供支持。 澳大利亚悉尼大学图书馆在科研数据管理方面取得了 丰富经验。本文通过网络调研和文献调研,分析其 科研数据管理方面的具体措施,以期为国内高校图书馆 开展数据管理服务提供借鉴。

2 高校图书馆科研数据管理服务实施

美国图书馆协会在2015年4月中旬发布的《2015年 美国图书馆状况报告》中指出,高校图书馆为学生、教 师和全体人员的学习、教学、研究提供资源和服务的 支持^[12]。“支持(support)”是高校图书馆服务的重要 理念,即对教学和科研的全方位支持。为此,高校图书馆 开设e-scholarship服务,为研究者提供从信息收集分析、 存储到最后的出版发布的全生命周期的科研数据 管理服务,使科研人员可以在更大的网络空间中进行学 术交流、科研创新。数据管理服务是其研究支持服务中 的一项,研究支持服务一般由学校跨部门合作开展,在 学校科研管理部门的统一领导下,由图书馆联合学校 研究部门、财务部门、人力资源部门等,协调组织开展 工作。负责该工作的馆员一般被称为科研数据管理协 调员(Research Data Management Coordinator),面 向的服务对象包括大学教职人员、研究层次的学生和 其他附属教研人员。

2.1 科研数据管理服务政策支持

科学数据管理在国家科技发展中发挥的基础性和 引导性作用已成为共识。科学数据管理是我国参与国际 科技竞争的重要支撑条件和基础保障,只有将科学数据 管理工作上升和融合到国家战略中去,资金投入、政 策法规完善、技术标准建立、专业人才队伍建设、组织 保障等各个方面才能在国家层面获得系统安排、统筹

规划^[13]。从国家层面上来说,制定并颁布实施相关的政 策能够为研究机构和高校的政策制定、实施准则提供方 向引导。在此基础上,全国性的科研机构、中心及联盟和 学会、协会合作推动,从专业研究的角度制定细化的准 则方案。在以上政策规范的指导下,高校图书馆根据本 校拥有的资源和服务实施情况,结合自身发展需求,对上 述文件进行总体把握并灵活调整,参考学科专家的意见, 制定本校的科研数据管理政策,并使之具有可操作性。

(1) 研究机构准则:2007年澳大利亚政府联合国 家健康医疗研究委员会(National Health and Medical Research Council)、澳大利亚研究委员会(the Australian Research Council)和澳大利亚大学联盟 (Universities Australia)发布了《澳大利亚研究行为 准则》(Australian Code for the Responsible Conduct of Research),准则规定了研究机构和研究人员应遵守 的原则和实践,并提供了一个解决违反本准则的学术不 端行为框架^[14]。

(2) 学校政策:为响应《澳大利亚研究行为准 则》,悉尼大学发布了本校的《研究行为准则2013》

(Research Code of Conduct 2013),并于2013年5月27 日开始实行。该准则的制定在结构和内容上参考了《澳 大利亚研究行为准则》,针对本校的具体情况做出了调 整,具体说明了如何规范、正确地实施科研实践和对违 反原则、学术不端行为的处理。政策共分13部分,包括 政策名称,开始时间,约束范围,意向声明,适用范围, 概念定义,科研数据,元数据和记录文件,科研数据管 理计划(Research Data Management Plan, RDMP), 科研数据、原始资料和研究记录的保存,科研数据、原 始资料的所有权,获取科研数据集,角色和职责。总体 上对悉尼大学科研数据管理进行了概要说明,并提供 指导性管理思路。

根据上述准则、政策,研究人员和机构有义务和责 任维护和管理科研数据,通过制定一系列政策规范, 能够使研究人员在进行科学研究时,有明确具体的规 章制度可循,促进了科研行为的规范化、有序化、流程 化,加强了对科研行为的科学规范管理。

2.2 科研数据管理服务技术支撑

2.2.1 创建数据管理计划

科研数据管理计划是高校科研数据管理政策中提

出的要求。科研数据管理计划帮助研究人员在收集、存储、分享科研数据过程中进行规划和归档整理工作,提供访问学校的科研数据存储库的服务。科研数据管理计划为研究者进行学术研究提供帮助;向资金资助机构证明科研经费合理性;通过科研数据的再利用,提升研究成果的引用次数和学术影响力;提升学校在支持研究方面的基础设施建设和服务水平。

研究人员登录图书馆账号,即可创建自己的科研数据管理计划(RDMP)。创建一个RDMP既可以获得2TB以内的免费数据存储空间,同时也响应了学校的政策要求。悉尼大学要求研究人员提供详细的数据管理计划清单,并严格遵循该准则,鼓励研究者(包括高年级研究层次的学生)在每个研究项目开始时进行科研数据管理计划。

2.2.2 升级机构库,建设研究数据存储库

近年来,图书馆机构库的建设在国内外已普遍开展,但在发展过程中出现了一些问题,主要表现在政策、版权和管理方面。因此高校图书馆可以在已建成的存储系统机构库的基础上,开展科学数据管理工作。如康奈尔大学图书馆以数据仓储DataStaR为数据组织管理平台进行数据管理。悉尼大学的科研数据存储库(Research Data Store, RDS)是一个企业级的网络附加存储(Network Attached Storage, NAS)设备,其存储容量大,可扩展性强。

数据存储库的功能为研究人员需通过完成科研数据管理计划来使用科研数据存储库。

(1) 云存储:通过授予访问权限,用户可以简单地使用对数据存储库网络硬盘的存储空间进行分区,保证用户可以在校园网内或者通过校外VPN网络的任何一台计算机访问存储空间。

(2) 存储空间规划:数据存储库提供一定的免费存储空间,超出将另外收费;为确保研究人员的需求得到满足,图书馆信息技术部门应对存储空间容量作出规划,并对存储空间的使用情况实行配额限制并监控实际使用情况。

数据存储库安全措施,主要表现在:①硬件,通过使用不同层次的数据存储实现磁盘级保护;②定期进行备份并保存,可通过图书馆信息技术部门新要求数据进行恢复;③数据灾难恢复,当大学数据中心不可用时,可通过激活数据中心访问已经同步的数据。

总而言之,数据存储库帮助研究人员在数据保存、与元数据链接、可获得性、安全性、备份、可恢复性、容量(可扩展性)、共享性、文件管理、可组织性等方面满足了数据管理行为准则和创建科研数据管理计划的各项政策要求。

2.2.3 发布数据注册平台

图书馆数据注册系统是为科研数据管理服务的数据搜集工具,该系统的建设目标为:对研究数据集合进行管理、规划、提升研究数据可见度;利用适当的方式发现和重用数据;建立对有价值的数据进行搜索的数据库;加强研究人员间的沟通联系和协作。通过联系(发送邮件)图书馆科研数据管理服务部门的工作人员或者负责本院系的学科馆员来将自己的研究成果在系统上进行注册。

以悉尼研究数据注册系统(Sydney Research Data Registry)为例,该系统是悉尼大学和澳大利亚国家数据服务(Australian National Data Service, ANDS)开发的研究数据管理工具,用来跟踪、捕获、收集悉尼大学研究人员产生的研究数据集合的在线系统,包含的信息是研究数据的集合即元数据,而非研究数据本身。系统实现与外部平台的对接,可以把关于研究数据集合的描述性信息发布到ANDS旗下的澳大利亚研究数据(Research Data Australia)^[15]平台上。

2.3 科研数据管理服务推广

2.3.1 建设指导性网站

图书馆网站作为图书馆提供服务的重要门户,是图书馆对外展示形象的窗口和拓展延伸读者服务的重要渠道^[16]。高校图书馆通过建设科研数据管理专题网站,有助于图书馆加强与研究人员的联系,通过推介提升服务知晓度,进而鼓励研究人员获取利用。

网站通过设立研究数据管理计划清单指导研究人员完成研究数据管理的一系列流程。内容一般包括管理数据、保存数据、数据共享和存档、发现数据、道德规范与研究数据、知识产权和政策等方面的内容,涵盖数据管理服务项目中一系列的流程,并在每一部分详细列出具体的建议、操作规范、疑难解答等。

悉尼大学在官方网站主页的研究支持栏目中设

立了数据管理服务项目的指导性网站,由图书馆和研究组合部共同进行建设维护。作为澳大利亚国家数据服务(Australian National Data Service, ANDS)的 Seeding the Commons 项目^[17]的一部分,网站借鉴了澳大利亚其他高校如昆士兰大学、纽卡斯尔大学的建设资源和经验,向用户普及数据管理相关信息,并提供数据管理方面的服务支持。

2.3.2 培训与互动

高校图书馆在整合人力资源的基础上,结合学科数据管理需求,先在重点学科进行科研数据管理服务试点,进而逐步推广至全校各学科。可采用举办系列讲座、开设培训课程、邀请专家进行案例展示等方式,通过网站发布培训、讲座等信息,重视互动宣传,设置最新报道、新闻存档、投稿、RSS订阅等栏目,更新服务内容,提升服务项目的时效性、动态性,以增进用户对数据管理相关内容的了解。还可利用图书馆各类型社交媒体平台(微博、微信等),对重要信息重点宣传推广。

悉尼大学图书馆开展的有关数据管理服务的培训包括学术讲座、会议等,如发表高质量论文的讲座、核心研究工具年会,以吸引用户参与到数据管理服务项目中来。

3 对我国高校图书馆数据管理服务的启示

3.1 发挥政策先导性作用

纵观国外图书馆的数据管理服务,可明显看出,在政策制定方面,积极响应国家政策,制定切合本校的长期发展准则、政策,同时通过图书馆各部门的协调合作,争取多方面的力量支持。对高校图书馆而言,可积极联系所在学校的管理部门,通过实际案例和实施效果说服政策制定部门,将图书馆的主张上升为学校的政策,同时争取获得学校经费、人员、设备等方面的支持。

3.2 构建技术支撑环境

在科研数据管理的技术应用方面,国外高校图书馆投入建设开发了一系列的基础设施和软件工具,包括大型存储设备、高性能服务器、软件系统等,搭建服务

器终端。因此,高校图书馆应致力于建设便于用户科研数据管理与利用的环境,把校园内外的技术设施与研究人员的需求整合到一起,通过自建、共建、合作共享、购买、租用等方式,构建用户科研数据管理平台,并把其嵌入到用户的科研环境、科研过程,满足科研人员信息交流、数据共享的需求,同时满足数据存储、管理、长期保存、复用的需要。

3.3 加强研究人员的科研数据管理素养培训

虽然国内对数据管理理论研究已达到一定深度,一些高校图书馆数据管理平台也已经建成投入使用,但由于各方面的原因,数据管理服务也陷入了信息服务“最后一公里”的困境,并未被科研人员获取利用,造成了研发资源的浪费。对高校图书馆而言,应针对科研人员的数据管理需求,提供普遍的和学科化的信息素养教育,宣传推广数据管理服务为他们带来的便利。如提供集成发现获取数据的途径,包括政府网站、学术机构库、大学网站、搜索引擎等;从文件命名规则、组织文件夹和文件版本控制入手,指导研究人员如何组织数据,如何对数据长期保存与处置;明确进行研究数据管理时应遵循的道德要求,涉及隐私、保密和同意、文化敏感性、商业敏感性等问题。

3.4 跟踪定位,开发深层次服务

由调查结果可以发现,目前高校图书馆现有的数据管理服务虽然全面,但不够深入。国内高校图书馆在进行该服务时,应当在此基础上继续拓展延伸,在数据获取服务、数据分析服务、信息素养教育等方面继续加强。在开发数据管理工具和硬件平台的基础上,将图书馆的学科馆员进行参考咨询的智能分析工作成果嵌入到数据发现、数据管理、数据共享、数据交流出版的指导和实践中,支持个性化的发现、集成、定制,促进数据服务层次不断深化。

4 结语

“数字化浪潮”正深刻影响着知识流向,在此背景下高校图书馆如何提升与拓展服务能力成为图书馆业界值得思考和关注的问题。2015年3月23日在中国图书

馆学会专业图书馆分会举办的主题为“新技术环境下的数字图书馆建设与服务”学术报告会上,张智雄研究员将“信息管理的模式由信息本身的生命周期管理转向科研生命周期的信息管理”作为数字图书馆未来发展的八大趋势之一,因此以服务教学科研为己任的高校图书馆更应当与不同领域、不同类型的研究机构合作,进行深层次、细粒度的知识揭示,提供专业化的、深入内容的知识服务,肩负起高校图书馆科学研究使命,实现多元化发展,在知识开放和流动中焕发活力。

参考文献

- [1] Jim Gray,Alexander S.Szalay,Ani R Thakar,et al.Online scientific data curation,publication,and archiving [R/OL].[2015-01-14].<http://research.microsoft.com/pubs/64568/tr-2002-74.pdf>.
- [2] 崔宇红. E-Science环境中研究图书馆的新角色:科学数据管理[J].图书馆杂志,2012(10):20-23.
- [3] Research Data Alliance Launch and First Plenary[EB/OL].[2015-01-07].<https://rd-alliance.org/plenary-meetings/first-plenary.html>.
- [4] Scientific Data Sharing: An Interdisciplinary Workshop[EB/OL]. [2015-01-07]. <http://www.globaleventslist.elsevier.com/events/2013/09/scientific-data-sharing-an-interdisciplinary-workshop.html>.
- [5] National Science Board.Long-Lived Digital Data Collections Enabling Research and Education in the 21st Century[EB/OL].[2015-1-14].http://www.nsf.gov/nsb/meetings/2005/LLDDC_draftreport.pdf.
- [6] 杨鹤林. 数据监护:美国高校图书馆的新探索[J].大学图书馆学报,2011(2):18-21,41.
- [7] Lord P,Macdonald A.Data curation for e-Science in the UK: an audit to establish requirements for futurecuration and provision[EB/OL].[2015-01-14]. http://www.jisc.ac.uk/uploaded_documents/e-ScienceReportFinal.pdf.
- [8] DCC Curation Lifecycle Model[EB/OL].[2015-01-14]. <http://www.dcc.ac.uk/resources/curation-lifecycle-model>.
- [9] Adrian Burton,Andrew Treloar,Designing for Discovery and Reuse: the ANDS Data Sharing Verbs' Approach to Service Decomposition[J].INTERNATIONAL JOURNAL OF DIGITAL CURATION,2009,4(3):44-56.
- [10] Friedlander A,Adler P.To Stand the Test of Time:Long -Term Stewardship of Digital Data Sets in Science and Engineering.A Report to the National Science Foundation from the ARL Workshop on New Collaborative Relationships--The Role of Academic Libraries in the Digital Data Universe[J]. ASSOCIATION OF RESEARCH LIBRARIES,2006(9):160.
- [11] 邢文明,吴方枝,司莉. 高校图书馆开展科研数据管理与共享服务调查分析[J].图书馆论坛,2013(06):19-25,61.
- [12] The State of America's Libraries 2015: A Report from the American Library Association [EB/OL].[2015-05-21].<http://www.ala.org/news/state-americas-libraries-report-2015>.
- [13] 黄如花,王斌,周志峰.促进我国科学数据共享的对策[J].图书馆,2014(3):7-13.
- [14] Australian Code for the Responsible Conduct of Research[EB/OL].[2015-04-12]. <http://www.nhmrc.gov.au/guidelines-publications/r39>.
- [15] Research Data Australia[EB/OL].[2015-04-12].<http://researchdata.ands.org.au/>.
- [16] 郭继棠.浅谈省级公共图书馆的网站建设[J].图书与情报,2010(6):131-133.
- [17] Seeding the Commons project[EB/OL].[2015-04-12]. http://sydney.edu.au/research_support/data/seeding-the-commons-project.shtml.

作者简介

李秋实,女,1992年生,硕士研究生,研究方向:信息资源管理,E-mail:apple645@163.com。

Research Data Management and Service in Academic Libraries

LI QiuShi

(Information Management School of Wuhan University, Wuhan 430072,China)

Abstract: Based on the practical needs of academic library research data management services, this article gives an analysis of the research data management services primarily from policy support, technical support and services promotion, and combines with foreign library practice to be confirmed. According to current situation of the academic libraries in China, it puts forward proposals for the research data management services, aiming to improve the related research and provide reference.

Keywords: Research Data; Data Management; Academic Library

(投稿日期:2015-04-21;编辑:雷雪)