

香港高校科研数据管理服务调查分析*

赵小兰 刘桂锋

(江苏大学科技信息研究所, 镇江 212013)

摘要: 采用网络调查法与内容分析法对香港10所高校图书馆的科研数据管理服务实践现状进行比较和总结, 从数据生命周期的各个环节深入剖析数据服务内容和特色, 并借助机构知识库进行数据集出版与共享服务的补充调查, 最后从数据管理政策、数据生命周期理论、机构知识库3个角度进行总结并提出相应建议。

关键词: 科研数据; 数据管理; 数据服务; 数据科学; 机构知识库; 数据出版

中图分类号: G252

DOI: 10.3772/j.issn.1673-2286.2018.06.006

1 科研数据管理服务概述

科研数据 (Research Data) 指基于参数、理论、测试或假设得到的事实、观测、图像、计算机程序、音频、测量或实验等各种形式的数据^[1]。数据类型可能是数字型、描述型、视觉型或触觉型的, 这些类型的数据也可以是原始的、干净的或经过加工的, 并可保存在任何形式的媒介上。

我国习惯把“科研数据”称作“科学数据”。但科学数据的概念相对宽泛, 在国务院办公厅印发的《科学数据管理办法》中, 科学数据主要指在自然科学、工程技术科学等领域, 通过基础研究、应用研究、试验开发等产生的数据, 通过观测监测、考察调查、检验检测等方式取得并用于科学研究活动的原始数据及其衍生数据^[2]。

近年来, 科研数据管理逐渐受到科学界及图书馆界的普遍关注。图书馆界开展科研数据服务, 既是顺应大数据背景下科学研究范式转变的时代要求, 满足科学研究基金组织对数据管理的基本要求, 也是科研人员数据管理的内在需求, 以及图书馆学科服务与机构知识库转型升级的必然需要^[3]。

科研数据管理服务是科研数据管理活动的重要组成部分, 国内外学者通常采用网络调查、实证案例、

比较分析等方法探析图书馆开展科研数据管理服务的现状, 归纳总结特点, 并提出对策建议。2008年, Choudhury^[4]开始探索美国约翰霍普金斯大学数据管理的原型系统。此后, 随着基金资助机构政策、期刊出版政策等的推动, 科研人员撰写数据管理计划的需求迫使图书馆提供科研数据管理服务。美国、澳大利亚、加拿大及欧洲等国家在开展科研数据管理服务上较有代表性^[5-10]。Searle等^[11]分析澳大利亚格里菲斯大学科研数据管理服务的典型案例, 包括科研数据政策、基础设施、技能提升和咨询培训等方面, 重点探讨图书馆员应发挥的作用; Yu^[12]分析高校图书馆在科研数据管理计划、元数据咨询与工具选择、数据归档、机构知识库运用、数据共享与访问等科研数据服务中面临的挑战与机遇; Cox等^[13]比较澳大利亚、加拿大、德国、爱尔兰、荷兰、新西兰和英国等高校图书馆的科研数据活动、服务与技能, 结果表明图书馆在科研数据活动中发挥主导作用, 尤其是在宣传和政策制定方面。除港澳台以外, 我国科研数据管理服务具有代表性的高校主要有武汉大学^[14]、复旦大学^[15]和北京大学^[16]等。目前, 香港高校的科研数据管理服务水平虽然有限, 但在实践过程中的经验能够给予内地高校科研数据管理服务参考与借鉴。

本文重点针对香港10所知名高校的图书馆网站进

*本研究得到国家社会科学基金一般项目“开放科学理念下的科研数据治理研究” (编号: 17BTQ025) 资助。

行调研,比较分析其科研数据服务的内容,重点选取香港大学、香港理工大学和香港科技大学3所高校的服务内容作为典型案例,从数据生命周期视角分析各自特色,最后对内地高校开展科研数据服务提出建议与启示。

2 数据来源分析

根据QS亚洲高校排名和香港Top Universities的排名,本文从香港排名前20的高校中选取已经开展科研数据管理服务的10所高校作为调查对象,包括由香港大学教育资助委员会资助的8所公立高校和2所私立高校。调查结果呈现3个特点。①服务内容多以Research Service/Support (RS)命名,也有以Research Data Management (RDM)、Digital Repository (DR)、Digital Initiative (DI)或Research Repository (RR)命名,其中香港公开大学和香港树仁大学以Archive和E-resource命名。②服务活动的水平和层次差异较大。香港大学、香港理工大学、香港中文大学、香港科技大学和香港浸会大学的数据管理服务内容相对丰富;香港教育大学提供收集、存储和引用服务;香港城市大学、岭南大学和香港树仁大学分别提供基础性的收集服务、出版服务和引用服务;而香港公开大学暂时尚未提供具体数据服务,仅依靠档案库提供科研数据搜索服务,未作详细过程分类服务。③服务活动的侧重点不同。调查的高校中有7所开展数据出版、数据共享、数据引用服务,6所开展数据收集服务,5所开展数据计划、存储服务,3所提供数据组织服务,2所提供数据再利用服务。

3 基于网络调查的科研数据管理服务内容概览

3.1 科研数据管理服务内容概览

本文采用网络调查法和内容分析法对香港10所高校的数据管理服务内容进行调研,发现香港高校图书馆主要对数据的管理计划、收集、组织、存储、出版或共享、引用和再利用等主要环节提供服务,且不同高校的服务内容不尽相同,各有侧重。根据调研可知,目前香港高校开展科研数据管理服务的内容见表1。

通过对调研学校的服务方式、服务资源、服务内容

表1 香港高校科研数据管理服务内容调查

学校名称	服务内容						
	数据管理计划	数据收集	数据组织	数据存储	数据出版或共享	数据引用	数据再利用
香港大学	√	√	√	√	√	√	×
香港理工大学	√	√	√	√	√	√	√
香港中文大学	√	√	×	×	√	√	×
香港科技大学	√	√	√	√	√	√	√
香港浸会大学	√	√	×	√	√	√	×
香港教育大学	×	√	×	√	×	√	×
香港城市大学	×	×	×	×	√	×	×
岭南大学	×	×	×	×	√	×	×
香港公开大学	×	×	×	×	×	×	×
香港树仁大学	×	×	×	×	×	√	×

和服务特色的综合分析,香港大学、香港理工大学和香港科技大学的服务特色鲜明,服务内容丰富。

香港大学的科研数据服务主要包括数据管理政策、知识产权与道德研究、文档/文件组织和文件格式、存储计划、研究出版、香港大学学术库。香港大学制定了数据管理计划和数据管理政策,明确知识产权与道德研究问题,详细描述文档/文件组织与文件格式、存储计划和研究出版要求等,同时还提供数据集服务。总之,香港大学科研数据服务内容不仅涵盖整个数据管理活动,还明确规定学校和研究服务人员的职责,以最高标准严格保证各个环节的有效执行。

香港理工大学的科研数据服务主要包括数据管理计划、数据收集、文档和元数据、道德与法律法规、存储备份与安全、选择与保存、数据共享。同时,明确研究项目中的责任,确保数据管理计划的有效实施。香港理工大学的科研数据服务依靠最新数字技术和专业服务团队,提供从研究项目到传播研究成果的一系列服务。

香港科技大学的科研数据服务主要包括数据管理计划、数据存储库和数据发现工具、数据引用、数据资助要求和发布政策、引用检索认证服务,以及最佳实践案例和培训服务。香港科技大学帮助科研人员计划、创建、组织、管理、访问和分享项目科研数据。香港科技大学允许每位参与研究的人员注册ORCID并保留研究成果,也可以通过DataSpace随时存储数据信息,并支持数据生命周期计划的过程管理,协助研究人员进行数据存储,提供最佳实践与培训,并利用DateCite提供数据引用服务,其引用检索认证服务也在逐步建立

中。此外，香港科技大学知识库可以为研究人员提供索引服务，公布研究人员的研究成果、研究项目和研究兴趣，提高其影响力并促成合作。

3.2 科研数据管理政策分析

科研数据管理政策对高校相关科研和管理人员的科研数据管理起规范和制约作用，有益于提高科研数据管理活动的效率和效益。在调研的10所高校中，香港大学明确提出数据管理政策，研究资助者遵守并履行数据相关政策声明和义务，确保建立完善的体系，并且捍卫研究过程和结果。香港大学的科研数据和记录管理政策于2015年5月5日获得批准，同时成立科研数据和记录管理专门小组，负责监督政策执行。相关研究人员、部门、中央行政单位、服务提供者、研究申办者和外部合作者在适当的情况下，需要满足并践行相关立法和监管要求。科研数据和记录必须准确、完整、真实、可靠、可识别检索和具有安全性。相关科研和管理人员须遵守法定义务，在适用情况下遵守机构审查委员会（IRB）和人类研究伦理委员会（HREC）批准的资助机构和项目特定协议要求。同时规定研究人员及学校相关负责人的具体职责，而高校研究委员会负责指导政策的制定和更新。

香港理工大学在科研数据计划中提供研究室、研究指引、一对一研究咨询、专家协助和文献检索服务；香港浸会大学在科研数据计划中帮助确认资金赞助，并计划收集、管理数据和源材料；其他高校未见具体描述。

4 基于网络调查的科研数据管理服务内容分析

数据管理计划是实现科研数据管理活动的宏伟蓝图。依据科研过程和科研数据管理生命周期理论，将服务内容划分为数据收集、数据组织、数据存储、数据出版或共享、数据引用、数据再利用，据此结合表1，对香港10所高校的科研数据管理服务进行更细致的分析。

4.1 数据管理计划

数据管理计划描述数据集的产生、采集、处理等过程的数据生命周期，通常包括收集不同类型的数据，

元数据格式和内容，访问、共享和再利用策略，存档和管理计划，保密和知识产权等。越来越多的资助机构要求研究人员在申请阶段准备数据管理计划，最大限度地提高受资助项目的投资回报。许多期刊出版商也要求研究人员将其论文的支撑数据分享或存放在开放存储库中，以提高论文的可见性和影响力。在调研的10所高校中有5所高校提供数据管理计划服务。

由表2可知，香港大学按照项目开展前、实施中和完成后三个阶段制定数据管理计划。香港中文大学的计划内容更加宽泛，涉及专家咨询、软硬件支撑及培训等方面；香港科技大学的计划包括数据收集、数据组织、数据保存与共享、数据再利用等环节，主要涉及生成或收集不同类型、元数据格式和内容、访问、共享和再利用策略，以及存档、管理计划、保密、知识产权等；香港浸会大学的数据管理计划突出相关科研和管理人员的角色与责任，以及数据安全等；香港理工大学的计划更加全面，除上述内容外，还强调道德与法律法规及资金来源。研究项目中各自的角色和责任需明确划分，以确保数据管理计划顺利实施，确保各阶段有效执行。合理预算数据管理任务所需的资源有专业知识、培训、硬件或软件、数据存储库、归档人员成本等。

综上，目前大部分香港高校的数据管理计划是按照数据生命周期理论进行制定，且不拘泥于科研过程，涉及角色与责任、道德与法律，以及配套条件保障。

表2 香港高校数据管理计划服务内容

学校名称	服务内容
香港大学	规划项目各阶段的数据收集、存储和维护数据，项目生命周期内处理日常基础信息，项目完成后规划处理已获取的数据
香港中文大学	免费的专家建议，包括数字项目的合适硬件、软件和程序，提供后勤支持
香港科技大学	生成或收集不同类型的数据，规定元数据格式和内容，涉及访问、共享和再利用策略，制定存档和管理计划，完善保密和知识产权协议
香港浸会大学	角色与责任，数据生成与收集，数据管理、存储和保存，数据安全
香港理工大学	管理信息，数据收集，文档和元数据，道德与法律法规，存储、备份和安全，选择和保存，数据共享，职责和资金来源

4.2 数据收集

数据产生或收集是科研过程的第一步，也是数据

管理计划的重要一环,在调研的10所高校中,共6所高校提及数据收集服务(见表3)。香港大学提出期刊论文的数据集副本和科研项目产生的独立数据集都可能成为收集对象,并且着重强调在数据收集时遵守相关协议及伦理规范。香港大学知识产权政策规定了工作人员和学生知识产权的所有权,并表明了对版权和专利的立场,收集和储存个人资料必须符合《个人资料(私隐)条例》;香港浸会大学提供数据收集的各种类型与格式;香港中文大学规定收藏的类型及应遵守的规范;香港科技大学指明通过DataSpace上传与提交数据,并提出具体操作要求;香港教育大学通过EdUHK为师生提供各种出版物的收集与保存服务;香港理工大学提出数据收集需要考虑的问题,即研究人员需要选择数据格式,列出计划用的文件类型、可能用到的软件及简要原因,同时需要考虑存储、备份和访问方面的影响。

表3 香港高校数据收集服务内容

学校名称	服务内容
香港大学	收集和储存个人资料必须符合《个人资料(私隐)条例》,从人类或活的脊椎动物收集的资料需要伦理委员会的批准
香港浸会大学	数据收集的类型与格式应考虑地理空间数据库、仪器、模型、多媒体、数字格式、软件等
香港中文大学	期刊、论文、手稿、专题特藏(甲骨文、香港文学)等分类明确的收藏,提交收藏应遵守规定的文件组织格式。如违反《版权条例》的作品会被下线,但是印刷版可以存取
香港科技大学	提交前需遵守主题域、文件格式、共享策略、元数据标准等要求,通过DataSpace服务师生,可以随时上传数字作品
香港教育大学	以数字化格式长期保存;所有文件均经过版权许可处理;出版物包括项目说明书和手册、高校出版物、工作人员出版物、学生出版物、课程材料和内部文件
香港理工大学	数据收集需要考虑数据收集方式、数据组织方式、文件命名规则、版本控制、质量保证流程

4.3 数据组织

数据组织在数据生命周期中起承前启后的作用,主要包括元数据、数据文档、文件格式、命名、版本等。对科研数据进行组织,有利于数据的理解和再利用,可以节省时间和资源,从而提高研究效率。

在调研的10所高校中,香港大学、香港科技大学

和香港理工大学都提供数据组织服务。其中香港大学的服务最详细,主要采用DC元数据,具体提供保留元数据库与文件链接,将readme.txt文件与提供基本解释性细节的数据一起存储,在实验记录本或相关论文和报告中记录相关内容并链接相关网站,以及利用文件夹和标签等对科研数据文件进行重叠,按主题、作者等对文件进行分类的服务。此外,文件组织和文件格式通常采用简洁、信息丰富的文件名称进行文件命名,无须打开文件就能分辨内容。为减少无法读取的风险,应尽量选择开放格式,或者通用格式。开放格式包括PNG、FLAC、WebM、HTML、gzip、CSS等,非开放格式包括CDR、DWG、PSD、RAR、WMA等。香港理工大学指出,文档和元数据伴随数据产生,应尽可能使用现有的描述标准;没有额外的解释性文件,科研数据集通常无法理解。因此,需考虑如何获得或创建文档和元数据、记录文件的工具、数据字典等支持文档和元数据标准,以便研究人员更好地理解数据。香港科技大学提出不同学科有不同的元数据标准,研究者可以选择适合的需求标准,相关组织和机构为科研数据管理提供有价值的指导;此外,香港科技大学还提供科研数据管理最佳实践和在线培训材料的相关链接。

4.4 数据存储

科研数据的存储在整个数据管理计划中的意义显著。在调研的10所香港高校中有5所高校提供该服务。

香港大学的存储方式包括学术库、部门及个人的存储设备,以及云存储等。香港大学提供保存实验室笔记的政策保护和实践指南(ipHandbook of Best Practices),比较受欢迎的参考管理软件包括EndNote、RefWorks、Mendeley、Zotero和Colwiz。香港大学拥有EndNote网站许可证,且由香港大学图书馆提供培训。此外,香港大学制定电子通信和个人数据存储的便携式存储设备、个人电脑和公共云服务指导方针,其云服务可以在任何地点提供24小时访问,另有ITS(信息技术服务部)提供便携式存储设备、个人电脑、公共云服务等数据存储指导。香港大学的学术库可永久存储有价值的科研数据(含数据集)并附加“黑暗文档”存储服务。

香港理工大学强调数据存储与备份需考虑空间、费用、备份方式、安全性、恢复方式等问题。系统备份数据应该包括存储条款,涉及存储空间、可能的额外服

务费用、数据备份、数字和非数字数据及副本如何保存到安全位置、数据恢复等。研究人员需选择保留的数据和保存时间，厘清契约、法律或监管规定须保留或销毁的数据及保留条件和保留时间。对于需要长期保存的数据，保存位置也需要在考虑范围内，如存储库、档案库、第三方存储库及其存储费用等。香港科技大学对教师和学生的保存服务内容稍有差别，但都可以将二者的科研数据存入DataSpace@HKUST，支持在校研究人员在整个研究周期内的数据管理。

香港浸会大学规定数据存储的要求，如数据结构、存储媒介、相关政策、版权问题（如数据拥有者）和归档等。香港教育大学鼓励科研人员申请注册ORCID（开放研究者与贡献者ID），便于科研数据的长期保存。

4.5 数据出版或共享

数据出版或共享是实现数据价值的重要途径和条件。在调研的香港高校中，7所高校提供数据出版或共享服务（见表4），大部分是通过机构知识库提供服务，如香港大学的学术库、香港科技大学的DataSpace等。香港大学指出，研究成果应及时发布，出版物应以专业客观的分析和准确的科研数据为基础。

表4 香港高校数据出版或共享服务内容

学校名称	服务内容
香港大学	香港大学学术库中心数据在用户授权情况下可以开放存取，方便研究人员分享、合作
香港理工大学	数据共享需考虑数据受众、潜在用户如何查找数据，数据分享对象及分享条件，共享数据处理机制，持久标识符，以及政治、商业、专利等限制因素的应对措施
香港中文大学	数字库服务、年度科研海报展及版权保护信息服务
香港科技大学	通过机构知识库DataSpace进行出版与共享
香港浸会大学	学术交流服务支持科研数据开放获取，并提供E-Book和E-Journal发表平台（如中外医学哲学和Asian Energy）
香港城市大学	通过学术库提供数据集共享服务
岭南大学	岭南大学数字库1-Search提供数据共享服务

研究人员一般可通过数字服务平台进行数据出版或共享。如香港浸会大学的研究者合作加强科研数据集研究，将主要资料整理索引并数字化；开发在线数字学术项目，将数据集嵌入图书馆的机构资源库；支持“数字奖学金计划”，提供电子期刊和电子书出版平台。

4.6 数据引用

数据引用对于使用他人科研数据或成果是必要的。调研中有7所高校提供该服务，具体服务内容见表5。香港科技大学图书馆为正在申请中国科学院、英国皇家学会等学术机构奖学金的教师提供引文检索认证服务，该服务包括根据用户提供的出版物清单，搜索一个或多个工具（如Web of Science），服务费可能会根据搜索出版物和引用工具的使用数量而发生变化；同时指出引用数据与引用文献不同，一些期刊、数据库和组织对数据引用有更具体的指导。DataCite旨在改善数据引用的国际组织，DataCite建议数据引用的基本格式至少应包括创作者（按发表年份）、题名、出版商和ID，其中版本或资源类型为可选项。

表5 香港高校数据引用服务内容

学校名称	服务内容
香港大学	遵守机构审查委员会（IRB）和人类研究伦理委员会（HREC）协议要求，合规利用他人科研数据
香港理工大学	图书馆提供相关线上或线下服务
香港中文大学	引文搜索、管理工具和引文指导指南
香港科技大学	引文检索认证服务
香港教育大学	Edlink平台提供资料印证链接（引文）服务
香港浸会大学	数字学术服务处培训或告知用户如何引用科研数据
香港树仁大学	图书馆在线资源提供在线引用工具

4.7 数据再利用

数据的再利用能够使数据发挥更大的价值和意义。调研的高校中，香港科技大学推荐不同学科数据发现工具，如re3data、Genbank、Harvard Dataverse、DataOne等；香港理工大学的数据发现平台OneSearch为发现和传递本地和远程资源提供一站式解决方案。其他高校多数可利用数据机构库或学术库等平台供研究人员进行数据开放共享并挖掘潜在合作对象。

5 香港高校机构知识库及数据集服务内容研究

在调查的10所香港高校中，只有岭南大学、香港树仁大学和香港公开大学暂未独立设置机构知识库服务，其他7所高校均已设置独立的数字机构库，为用户提供服务（见表6）。

表6 香港高校机构知识库

学校名称	机构知识库名称	服务内容			机构知识库网址
		论文	专利	数据集	
香港大学	学术库	√	√	√	http://hub.hku.hk/advanced-search?field1=title&thesis=1
香港理工大学	机构研究档案库	√	√	√	http://ira.lib.polyu.edu.hk/
香港中文大学	数字库	√	√	×	http://repository.lib.cuhk.edu.hk/#
香港城市大学	学术库	√	×	√	https://scholars.cityu.edu.hk/en/
香港科技大学	机构库	√	√	√	http://repository.ust.hk/ir/
香港教育大学	研究库	√	√	×	https://repository.edu.hk/
香港浸会大学	数字学术服务	√	×	×	https://repository.hkbu.edu.hk/res_units/

已开设机构知识库的高校中,服务水平和层次亦不尽相同。服务内容比较全面的有香港大学、香港科技大学和香港理工大学。香港大学提供研究人员、研究成果、机构组织、资助基金、数据集、论文和专利服务,香港大学学术库中心长期储存和保存香港大学的研究资料,研究人员和研究生可将数据集存入该中心,包括期刊论文的数据集副本和科研项目产生的独立数据集。香港科技大学机构库收集、保存和传播由本校成员创建的学术成果,所有存放的文件都可以在互联网上公开访问,为本校研究人员的学术出版物提供索引,并为现有教职人员提供学术资料导航,以展示研究人员的出版物、科研兴趣和项目。其他高校亦提供部分服务内容。

综上,已经开设机构知识库服务的7所高校均提供机构组织和论文服务,4所高校提供数据集服务。香港大学的数据集数量在360个左右,通过DOI进行下载,元数据类型有题名、院系、作者、教师、关键词、学科、RGC Code、ANZSRC等;数据集主要集中在临床肿瘤学系和计算机科学系,主题最多的是拟南芥。香港科技大学提供7个数据集,提供基础检索和高级检索,如Air Quality和Urban Cultural GIS Data of Republican Beijing;元数据类型有题名、期刊标题、作者、主题、ISBN及ISSN,用户可通过DataSpace@HKUST获取。香港城市大学提供5个数据集,只提供基础检索,如Martial Arts、Dance和Sports (MADS) Dataset;元数据类型有创建者、创建日期、创建者单位、出版商、DOI、数据描述及相关论文链接,注册账号可全文下载(部分有光盘或其他媒介资料)。香港理工大学在机构研究档案库的科研成果形式中虽提到期刊论文、学位论文、专利及数据集,但在检索界面却没有发现数据集的检索入口及具体的存储数量。

6 结论与启示

本文通过对香港10所高校图书馆的调研发现,每所图书馆均根据自身条件开展服务,主要呈现以下特点。

6.1 科研数据管理政策是开展服务的外在推动力

2018年3月17日,国务院办公厅颁布《科学数据管理办法》,以进一步加强和规范数据管理,保障科学数据安全,提高开放共享水平,更好地支撑国家科技创新、经济社会发展和国家安全^[2]。《科学数据管理办法》的颁布,表明科研数据的共享与管理已上升到国家战略层面,作为国家级数据管理政策,为高校、科研机构等各部门开展科研数据管理进行顶层设计并提供宏伟蓝图。科研数据管理政策是开展科研数据管理服务的重要外在推动力,对于相关科研人员的科研数据管理活动起到规范和制约作用,提升科研数据管理的效率和效益。正所谓“没有规矩,不成方圆”,制定并遵守一定的数据管理政策,了解并遵循行业道德规范,有利于形成良好的科研数据管理体系,更好地保护个人或集体的科研数据成果。香港高校开展科研数据管理服务的一个重要推动力是受发达国家科研数据管理氛围,以及国际基金资助机构与知名期刊对数据管理计划要求的影响,香港高校中虽然只有香港大学明确制定了数据管理政策,但其他高校均借鉴相关科研基金资助政策、政府开放数据政策、期刊开放数据政策等工作。数据管理政策的制定和有效执行,对香港高校开展数据管理服务具有十分重要的意义,更为未来相关数据服务的发展提供重要保障。目前,内地高校单独制定数据管理政策的还相对较少,我国科研人员对科

研数据管理政策的主要研究也还停留在理论方面,随着《科学数据管理办法》的稳步推进,以及科研基金组织与期刊对数据管理计划的强烈拉动,大规模高校制定数据管理政策的外部与内部条件均已成熟。

6.2 数据生命周期理论是开展服务的内部凝聚力

数据生命周期理论是香港高校科研数据服务体系构建的理论基础,其实质是依据科研过程来管理数据,是科研数据管理的关键阶段。国际上,英国数字保存中心、美国校际政治与社会科学数据共享联盟、英国数据档案馆等专业机构,以及牛津大学、弗吉尼亚大学、加州大学圣地亚哥分校等高校均制定特色鲜明的数据生命周期模型。大量理论与实践研究表明,数据生命周期模型一般分为数据管理计划、数据组织、数据存储、数据分析、数据共享和数据引用等。香港大学、香港科技大学、香港理工大学等高校按照数据生命周期理论制定数据管理计划,收集、处理、分析、分享、利用数据,将各环节的数据进行科学管理,逐步建立并完善已有的数据服务平台。《科学数据管理办法》明确规定,科学数据生命周期包括采集、汇交与保存,共享与利用,保密与安全三大部分,与国外模型相比,非常重视和强调数据的保密与安全。在我国已开展科研数据管理服务的高校中,虽然各阶段的名称有所差异,服务内容层次有所侧重,但是基本遵循了数据生命周期模型。因此,在进行科研数据管理时,按照数据生命周期理论对科研数据服务体系进行预期评估、内容框架搭建具有十分重要意义,为今后的科研数据管理奠定坚实基础。香港高校依照数据生命周期构建高校科研数据管理框架、设定科研数据服务的具体内容,为内地高校制定科研数据管理服务的具体实践提供参考和借鉴作用。

6.3 机构知识库是进行数据集共享服务的重要支撑力

《科学数据管理办法》提出,要积极推动科学数据出版和传播工作,支持科研人员整理发表知识产权清晰、准确完整、共享价值高的科学数据,同时,科学数据生产者应按照相关标准规范组织开展科学数据采集生产和加工整理,形成便于使用的数据库或数据集。目前科学数据出版与共享的模式一般分为3种:①

在已有机构设施的基础上改进或扩建,如机构知识库等;②新建开放数据共享平台,如哈佛大学的Harvard Dataverse和北京大学的开放研究数据平台;③租用校外机构(第三方)的基础设施,如数据存储库等。调研发现,香港高校(如香港大学、香港理工大学、香港科技大学等)提供科研数据管理服务的重要平台为机构知识库。通过机构知识库给读者提供论文、专利和数据集等服务,方便读者查找、下载文献,也能给作者或贡献者提供出版、保存科研成果的平台,扩大成果的影响力和利用率,并最大限度地挖掘潜在合作者,最终实现真正的数据共享和有效的数据再利用。高校机构知识库通常涵盖的科研成果类型是论文、专利、专著等科研成果数据,未来将拓展至数据集、科研数据等科研过程数据。各高校应结合自身实际情况,合理选择科研数据出版的模式。鉴于目前各高校的机构知识库均已开始运行,因此,依托机构知识库开展科研数据管理服务,实现双赢,既能方便科研人员进行科研数据与科研成果的链接集成,又能为图书馆拓展机构知识库的服务功能。香港高校的机构知识库服务虽然处于积极探索与发展阶段,但是诸如数据集等服务也初见成效,内地高校在建立和发展机构知识库时,可以对比、参照香港高校机构知识库,结合自身特点,构建和发展机构知识库,为本校科研人员提供数据集出版与共享服务。

综上所述,香港高校开展科研数据管理服务时间虽然不长,但在实践探索过程中已积累一定经验,制定数据管理政策是服务开展的规范保障。香港高校为更好地面向科研人员提供优质高效的科研数据服务,正在竞相构建各自的服务体系,依靠先进技术和专业团队最大限度地提供科研数据服务。尽管目前服务水平和层次不尽相同,但是已经对科研项目和相关人员产生积极影响。高校图书馆面向科研人员提供科研数据服务是一项复杂的系统性工程,包括科研人员服务需求识别、服务政策制定、服务理论运用、服务内容设计、服务体系构建、服务机制优化、服务效果评价等方面。内地高校应在《科学数据管理办法》的指引下,借鉴国内外丰富的理论成果与成功实践,循序渐进地推动科研数据管理服务科学、有序、高效地开展。

参考文献

- [1] D/2.8 Management of research data [EB/OL]. (2015-11-13) [2018-03-21]. http://www.mopp.qut.edu.au/D/D_02_08.jsp.

- [2] 国务院办公厅. 国务院办公厅关于印发科学数据管理办法的通知 [EB/OL]. (2018-03-17) [2018-04-02]. http://www.gov.cn/jzhengce/content/2018-04/02/content_5279272.htm.
- [3] 刘桂锋. 高校科研数据管理理论与实践 [M]. 镇江: 江苏大学出版社, 2017.
- [4] CHOUDHURY G S. Case study in data curation at Johns Hopkins University [J]. *Library Trends*, 2008, 57 (2): 211-220.
- [5] 张萍. 英国高校科研数据管理及启示 [J]. *情报杂志*, 2015, 34 (4): 155-159.
- [6] 韩金凤. 加拿大高校图书馆科研数据管理服务调研及启示 [J]. *国家图书馆学刊*, 2017, 26 (1): 38-46.
- [7] 陈媛媛, 柯平. 《欧洲学术研究图书馆科研数据服务》调查对我国高校图书馆科研数据服务的启示 [J]. *图书情报工作*, 2017, 61 (6): 73-78.
- [8] 王婉. 澳大利亚高校图书馆参与科研数据管理服务研究 [J]. *图书馆论坛*, 2014, 34 (3): 130-136, 149.
- [9] 张莎莎, 黄国彬, 邸弘阳. 美国高校图书馆科研数据管理服务研究 [J]. *图书馆杂志*, 2016, 35 (7): 59-66.
- [10] 刘桂锋, 卢章平. 美国高校科研数据管理实践个案研究——以伊利诺伊大学香槟分校为例 [J]. *图书情报研究*, 2016, 9 (4): 24-34.
- [11] SEARLE S, WOLSKI M P, SIMONS N, et al. Librarians as partners in research data service development at Griffith University [J]. *Program Electronic Library and Information Systems*, 2015, 49 (4): 440-460.
- [12] YU H H. The role of academic libraries in research data service (RDS) provision: opportunities and challenges [J]. *Electronic Library*, 2017, 35 (4): 783-797.
- [13] COX A M, KENNAN M A, LYON L, et al. Developments in research data management in academic libraries: towards an understanding of research data service maturity [J]. *Journal of the Association for Information Science and Technology*, 2017, 68 (9): 2182-2200.
- [14] 刘霞, 饶艳. 高校图书馆科学数据管理与服务初探——武汉大学图书馆案例分析 [J]. *图书情报工作*, 2013, 57 (6): 33-38.
- [15] 张计龙, 殷沈琴, 张用, 等. 社会科学数据的共享与服务——以复旦大学社会科学数据共享平台为例 [J]. *大学图书馆学报*, 2015, 33 (1): 74-79.
- [16] 朱玲, 聂华, 崔海媛, 等. 北京大学开放研究数据平台建设: 探索与实践 [J]. *图书情报工作*, 2016, 60 (4): 44-51.

作者简介

赵小兰, 女, 1988年生, 硕士研究生, 研究方向: 科研数据管理。

刘桂锋, 男, 1980年生, 博士, 副研究馆员, 硕士生导师, 通信作者, 研究方向: 科研数据管理, E-mail: liugui Feng29@163.com。

Investigation and Analysis of Research Data Management Service in Hong Kong Universities

ZHAO XiaoLan LIU GuiFeng

(Institute of Science and Technology Information, Jiangsu University, Zhenjiang 212013, China)

Abstract: Online survey and content analysis methods are adopted to compare and summarize the practice of research data management services in libraries of ten Hong Kong universities. Upon various stages of the data life cycle, we take in-depth analysis of data service contents and features. By taking advantages of the institutional repositories, we supply the investigation of dataset publishing and sharing services. Finally, conclusions and suggestions come from three perspectives: data management policy, data life cycle theory and institutional repository.

Keywords: Research Data; Data Management; Data Service; Data Science; Institutional Repository; Data Publishing

(收稿日期: 2018-05-05)