

蔡韶莹. 伦敦卫生与热带医学院医学科研数据管理经验及其启示[J]. 中华医学图书情报杂志, 2018, 27(8): 8-12.

DOI: 10.3969/j.issn.1671-3982.2018.08.002

· 专题 ·

伦敦卫生与热带医学院医学科研数据管理经验及其启示

蔡韶莹

[摘要]通过对伦敦卫生与热带医学院的医学科研数据管理实践的案例研究,分别从实践背景、政策制定、基础设施建设等方面剖析了伦敦卫生与热带医学院的成功经验,以期对我国医学科研数据管理的研究与实践提供有益的启示。

[关键词]医学科研数据;数据管理;伦敦卫生与热带医学院

[中图分类号]G252;R-058

[文献标志码]A

[文章编号]1671-3982(2018)08-0008-05

Experiences with medical research data management in London School of Hygiene and Tropical Medicine and their enlightenments

CAI Shao-ying

(Library of Xiamen Medical College, Xiamen 361023, Fujian Province, China)

[Abstract] The successful experiences with medical research data management in London School of Hygiene and Tropical Medicine were analyzed in aspects of its practical background, policy making and infrastructure construction according to the case study on its medical research data management in order to provide their enlightenments on the research and practice of medical research data management in our country.

[Key words] Medical research data; Data management; London School of Hygiene and Tropical Medicine

随着 e-Science 的发展,很多大学越来越意识到整理集成后的科研数据是非常有价值的资源。对科研数据进行管理不仅可以使科研人员利用现有数据加快研究进程,还可以让资助的公共研究成果为公众所用,提升全社会的创新和研究水平。高校作为科学研究的重要阵地,其科研活动产生的数据具有学科范围广、数据零星分散、类型多样等特点,管理起来复杂^[1]。依靠科研人员自主管理科研数据的传统方式管理效率和水平较低,给科研数据的共享和再利用造成了很大困难。因此,如何对这些科研数据进行有效管理已成为亟需解决的问题。

国外高校已经在科研数据服务实践方面进行了大量的尝试,积累了丰富的经验,如伦敦大学科研数据管理的实践经验给我们提供了有用的借鉴,值得深入研究。本文详细介绍了伦敦卫生与热带医学院的科研数据管理服务经验,希望能够启发研究机构特别是医学院校利用有限的资源提供有效的医学科研数据管理服务。

1 伦敦卫生与热带医学院的科研数据管理历程

伦敦卫生与热带医学院^[2]是伦敦大学享有盛名的 3 所学院之一,在研究生培养、前沿领域研究中为公众健康和热带医学的发展做出了重大贡献。英国三分之一的研究生医疗教育和研究都是在伦敦卫生与热带医学院进行的。与其他综合性大学相比,伦敦卫生与热带医学院的规模比较小,但在公共卫生、医疗服务、初级卫生保健等领域有显著的影响力。学院现有员工 1 300 多名,这些研究人员每年都产生大量的数据,而这些数据亟待以安全的方式

[基金项目]厦门市社会科学院基金项目“厦门市医学科学数据管理与共享保障体系研究”(厦社科研[2018]D09号)成果之一

[作者单位]厦门医学院图书馆,福建 厦门 361023

[作者简介]蔡韶莹(1986-),女,福建漳州人,硕士,馆员,发表学术论文

存储和管理。

学校的管理层早就意识到数据在科研活动中的重要作用。2002 年,英国联合信息系统委员会(Joint Information Systems Committee, JISC)资助的一项研究项目引起了学校的重视。该项研究认为,伦敦卫生与热带医学院应该建立一个中央数据服务机构帮助科研人员管理数据,确保这些数据被合理地保存,并在未来可以被访问^[3]。由于当时学校的档案管理部门没有足够的资源或专业知识承担任务,也没有制定相应的保障政策,因此学院的科研数据管理进展缓慢。

在现实需求的驱动下,学校高层经过研究讨论决定建立科研数据管理(Research Data Management, RDM)服务中心。2011 年该服务中心提交了一份报告,提出了促进科研数据管理实施的一系列建议。这些建议包括建立相关的制度政策、建设一个 RDM 网站以及用于数据发现的门户网站、增强机构档案部门的服务功能等^[4]。

2012 年,伦敦卫生与热带医学院的 RDM 项目正式开始运营。初期项目运营的经费是由威康信托基金会提供的,该基金会主要用来支持建立或重新开发具有重要战略意义的基础性建设^[5]。2015 年 7 月,威康(Wellcome)信托基金会提供的经费到期,转而由伦敦卫生与热带医学院为 RDM 服务中心提供经费支持。

2 伦敦卫生与热带医学院医学科研数据管理的经验

为了更好地提供服务,RDM 服务中心将 RDM 细分为 3 个目标来实现:一是加强科研数据管理政策的制定,二是加强基础设施建设,三是加强机构内部研究人员的数据管理实践。

2.1 加强科研数据管理政策的制定

这个目标主要是确保科研数据管理政策被纳入机构的政策框架,并保证制定的政策既要符合英国研究理事会(Research Councils UK, RCUK)的数据管理准则,也要匹配其他机构正在开发的相似政策文件^[6]。制定科研数据管理政策是顺利开展科研数据管理服务的关键性因素,不仅可以促进科研人员积极参与科研数据管理,还可借此嵌入新的服务。2012 年,RDM 服务中心将数据管理政策的第一版提交给学校的管理层,并进行了为期 12 个月的公

示,其中政策执行的灵活性成为最受争议的问题。

经过多次协商探讨,2018 年 2 月,RDM 服务中心提交了伦敦卫生与热带医学院科研数据管理政策的第二版。第二版的政策主要包括 10 方面的内容,并且允许某些科研人员根据实际情况决定是否执行。科研数据管理政策 2.0 版本^[6]主要包括以下内容:由学校主导的研究项目必须包含数据管理计划,在基金资助者同意的情况下,数据管理所产生的费用应该列入项目经费预算。在整个项目研究中数据权利应该尽早得到承认和支持,科研数据必须存储在规范化的环境中,无论科研人员获取的数据是否存储在学校或者其他地方都必须在学校登记在册,必须尽早通过合适的方法使支持基础发现的科研数据可以获取,应有足够的文件辅助理解和分享科研数据,通过标准和规范化的管理支持数据的重复利用,可获得的科研数据必须包含机构学会代码、基金信息、开放研究者与贡献者身份(Open Researcher and Contributor ID, ORCID),研究过程中所使用的科研数据必须在研究报告中以引用的方式标注出来。

2.2 加强基础设施建设

RDM 服务的第二个目标是结合 IT 服务,审视和加强机构内的数据管理基础设施。在前期数据收集阶段,开发人员对科研人员在数据管理过程中可能遇到的问题进行了调查,结果显示排名前 5 的问题是:数据归档操作的不确定性、制定数据共享协议过程中的问题、数据标准问题、数据安全问题以及数据存储空间不足的问题。

根据调查结果,围绕以下两个方面展开工作。一是进行数据管理系统支持。RDM 服务中心对学院工作人员进行了培训和指导,以帮助他们找到适合自己数据的内部或外部系统;系统选择侧重主题领域、灵敏度、内容类型和集合的大小 4 个标准。二是建设机构知识库。机构知识库可以对数据集进行编目、分类、存储,进而共享数据集;机构知识库的建设可以保证资源有限的 RDM 服务中心能够充分履行其收集并保存科研数据的职责,并填补目前学院在数据管理支持方面的空白。

为了保证机构知识库的可持续发展,RDM 服务中心规定:如果科研人员不能把科研数据合理地存储在某个地方(如英国数据服务中心),则必须把它

存储在学院的机构知识库。技术人员经过对几种常见机构知识库建库软件(如 Alfresco, CKAN, DSpace, EPrints and Fedora)的评估,最终决定采用南安普顿大学的 Eprints 平台^[7],并通过若干个第三方插件满足特定的建库需求。建成后的机构知识库服务器托管在伦敦大学的计算机中心,以便学校能够利用 EPrints 的优势。机构知识库承担的责任主要包括支持对科研数据集的格式处理、内容描述、长期保存等,提供查询、调用、显示等基本服务,提供对数据进行转换、交汇、可视化等软件工具,支持数据发布等服务^[8]。

2.3 加强科研人员的数据管理实践

该目标主要通过 3 个方面实现:一是确定机构内科研数据管理需求,二是确保科研人员从项目伊始就制定数据管理计划,三是支持科研人员不断变化的数据管理需求。

2.3.1 确定机构内科研数据管理需求

不同科研人员的科研数据管理需求各有不同,如进行田野调查研究的科研人员的需求就不同于在实验室工作的研究人员。因此,在项目的初期阶段,调查分析科研人员的需求至关重要。服务中心是通过以下方法确定机构内的科研数据管理需求的:一是通过网络调查的方式了解科研人员的研究方式、数据类型和潜在的需求;二是通过调研,了解 IT 服务部门和学术部门有关研究业务的政策和程序;三是列出经常支持学院研究项目的资助者名单,确定这些机构和项目对科研数据管理的具体要求;四是对国家或国际相关法律法规(如数据保护法^[3])进行研究,确保科研数据管理符合法律规定,并激发科研人员的数据共享意愿。

2.3.2 确保科研人员从项目伊始就制定数据管理计划^[9]

数据管理计划(Data Management Plan, DMP)是一份集中概括所有科研数据管理活动有关事项的文档,可以提高数据管理的透明度,并保证整个研究团队遵循统一的数据管理流程和规范,促进工作的一致性、规范化和高效率。但是在很多时候,科研人员在开始的时间并没有考虑到要进行科研数据管理,直到在研究过程中需要的时候才想起来。因此, RDM 服务中心将数据管理计划列入数据管理政策,

即必须创建所有研究项目数据管理计划,并提交给 RDM 服务中心审查。目前,制定数据管理计划日益成为学术界的标准做法,被越来越多研究型大学采用。

在指导制定数据管理计划的过程中,有 3 个需要注意的地方。一是优先重点项目:服务中心应确定机构内的重点研究项目,并确保这些项目制定合适的数据管理计划;二是避免不必要的重复:有些项目已经有出资者文件管理计划或者研究协议文件的,则不应再要求其完成本机构的数据管理计划;三是只提供指导:虽随时对科研人员在制定数据管理计划过程中遇到的问题进行指导或提供建议,但最终的决定权还是在科研人员手上。

2.3.3 支持不断变化的数据管理需求

许多学校 RDM 的核心活动之一是为科研人员的需求提供量身定制的服务,但刚开始运行的 RDM,既没有精确估算需求的数据,也没有可对比的关于处理支持请求的资料。因此在服务规划的初期阶段,RDM 服务中心要判断不同情况下的潜在需求,通过提交给 RDM 的请求支持详细记录信息来发现经常和不经常使用 RDM 服务的系(部),以及需要建议指导的地方。这些记录显示,排名前 5 的经常被用户提及的问题包括数据共享、资助者数据管理计划、伦敦卫生与热带医学院数据管理计划、数据许可、数据存储和安全系统^[10]。对于这些常见问题,RDM 制作了相应的问题指南以减少处理这些问题所需的时间,如制作了公共科学图书馆(The Public Library of Science, PLoS)数据政策概要指南^[11]和威康信托基金会的数据管理计划^[12]。在研究过程中遇到有关 PLoS 的问题时,科研人员可以根据现有的问题指南进行处理,从而减少等待时间。另外项目组还面向重点用户开展深入访谈,和科研人员共同讨论科研数据共享面临的挑战,详细了解用户具体需求,争取合作机会。除此之外,RDM 服务中心通过各种方式随时了解用户意见和建议,进而不断完善服务功能。

3 对医学院校医学科研数据管理的启示

数据密集型科研范式时代,对机构科研数据进行管理已成为共识。根据本文的分析结果,医学院校在进行科研数据管理活动时,要注意以下几方面的问题。

3.1 积极开展科研数据管理需求调研

对科研人员的数据管理需求进行调研可以为开展科研数据管理打下坚实的基础,伦敦卫生与热带医学院取得的成功与其科研数据管理需求调查密不可分。通过调查可以深入了解研究者所处的环境和遇到的挑战。调研内容主要围绕科研数据现状、数据的开发程度、研究者数据共享意愿、遇到的困难、希望得到的服务和培训等方面展开。

从英国高校的实践来看,参照数据资产框架(Data Asset Framework, DAF)进行需求调研是大多数高校的做法。DAF 是由英国 JISC 资助的一项关于数据资产审计的研究,主要目标是构建一种通用的高校数据资产审计框架。高校在进行科研数据管理时,可以参照这个框架进行数据资产的调研^[13]。据统计,越来越多的高校都以 DAF 框架进行机构数据需求的调研,并以此开展科研数据管理实践,如牛津大学、爱丁堡大学、谢菲尔德大学、伦敦大学的卫生与热带医学院等。我国高校在开展调研时,可以认真学习并参考 DAF,结合自身的特点,制定适合自己的科研数据管理需求的调研方案。

3.2 建立政府、医学院校等多层次的医学科研数据管理政策

制定科研数据管理政策是大学等学术机构开展科研数据管理的重要制度保障。国外高校在国家法律政策和科研资助机构的合力推动下,纷纷出台了科研数据管理的相关政策,从战略角度明确了科研数据管理的相关政策。2018 年 1 月 23 日,中央全面深化改革领导小组第二次会议审议通过了《科学数据管理办法》,《科学数据管理办法》的实施说明我国政府开始从国家层面重视科学数据管理。但其范围仅停留在行业层面,缺乏具体的关于医学科学数据管理的相关规定,同时国内重要的医学科研资助机构均未明确要求科研人员进行数据管理和数据提交,缺乏强制性,难以引起科研人员对医学科研数据管理的重视^[14]。从微观层面上来说,国内高校的管理层缺乏科研数据管理服务的意识,对科研数据的价值认知不足,将科研数据管理纳入高校科研发展战略议程并制定相关政策的医学院校屈指可数。

因此,尽快建立我国国家层面的医学科研数据管理政策是深入开展医学学科领域数据交流的迫切

需要,需要由卫生部主导,医学院校、医学研究机构以及各类科研资助机构等辅助建立一个统一的医学科研数据管理政策,从宏观角度设定科研数据管理服务的内容、标准、目标,为未来铺平道路^[15]。医学院校也要提升对科研数据管理重要性的认识,借鉴国内外成功的实践案例,将科研数据管理服务纳入学校的发展议程。

3.3 明确数据管理计划的重要性

完整的数据管理计划(Data Management Plan, DMP)是科研人员获取科研资助机构支持的关键因素,也是有效管理科研项目所产生的数据和应对可能产生的各类问题的重要保障。目前,科研数据管理计划已经成为国外很多科研资助机构评审项目的一个重要条件,如英国医学研究理事会(Medical Research Council, MRC)要求所有申请者提交的项目申请书都要包含一份科研数据管理计划^[16]。因此,科研数据管理服务中心要积极为医学科研人员提供相应的数据管理计划指导,不仅要就不同的基金和课题要求,为用户提供医学数据管理规划模板或指南,还要提供关于医学科研数据的类型、数据格式、数据标准、数据归档等咨询服务。除此之外,通过介绍常用的 DMP 工具可以帮助科研人员更轻松高效地创建数据管理计划,目前常用的数据管理计划创建工具有英国数字监管中心(Digital Curation Centre, DDC)开发的 DMPonline 以及美国加利福尼亚数字图书馆开发的 DMPTool。以 DMPonline 内置了英国基金会和部分大学的 DMP 内容需求模板文档,用户可按照其提供的模板及流程在线创建、编辑、分享、输出或者删除 DMP 文档^[17]。

3.4 注重服务分流,提升服务效率^[18]

由于人力资源有限,伦敦卫生与热带医学院 RDM 服务中心只有两名全职人员,其中一名是项目经理,另一名是软件开发工程师。另外还有一个由学校副校长主持的指导小组,该小组成员由各系(部)选出的学术专业人员组成,主要任务是提供建议以及发展方向。由于项目经理独自一人不能支持这么多医学院员工和学生的数据服务工作,因此项目组采取了一个三级分流的办法,确定优先顺序。首先优先支持有基金资助的科研人员;其次,对于没有基金支持的研究人员以及学生,如果他们希望获

得数据管理方面的帮助,可以通过 RDM 系统的培训和请求模块得到支持和建议;第三,详细记录每个提交到 RDM 的服务请求,并在此基础上规划 RDM 服务系统的开发工作,如通过提供书面指导和举办研讨会的方式减少同样的问题的重复率以及尽量减少处理同样请求所需的时间。

3.5 着力提高科研人员的数据素养

除上述基础的数据管理服务内容外,科研数据管理服务能否成功开展还需要提升科研人员本身的数据素养。伦敦卫生与热带医学院的 RDM 服务中心注重围绕服务内容,有针对性地开展研究人员科研数据管理技能培训、信息咨询等服务。培训的形式从网络教程、在线指导到各类面对面的研讨会、会议,培训的内容从科研数据管理的入门指导到深层次的问题辅导,一般根据研究人员、学生所处的水平和需求制定^[19]。

目前,我国科研人员收集传统文献资源的信息素养已经达到一定水平,但由于科研数据管理是一个新兴领域,科研人员对其的认识参差不齐。缺乏相关政策法规的鼓励以及对自己创建的科研数据的保护,影响了科研人员数据素养的提升,因此需对他们进行相关培训。具体数据素养教育服务的内容可包括国内外医学科研数据管理政策介绍、国内外数据资源介绍、数据分析工具的使用、数据安全教育等。

4 结语

伦敦大学以及英国高校科研数据管理实践的丰富经验对我国开展相关的研究与实践具有较高的借鉴意义,我国医学院校及研究机构可以借鉴英国的实践经验促进医学科研数据的管理,进而推动我国医学事业的蓬勃发展。

【参考文献】

- [1] 丁培. 国外大学科研数据管理政策研究[J]. 图书馆论坛, 2014, 34(5): 99-106.
- [2] Introducing our School[EB/OL]. [2018-05-10]. <https://www.lshtm.ac.uk/aboutus/introducing>.
- [3] Knight G. Building a research data management service for the London School of Hygiene & Tropical Medicine[J]. Electronic Library and Information Systems, 2015, 49(4): 424-439.
- [4] LSHTM Research Data Working Group. Raising standards - lowering barriers: documentation of, access to and preservation of re-

search data at the London School of Hygiene & Tropical Medicine [EB/OL]. [2018-05-10]. http://researchonline.lshtm.ac.uk/3409897/1/Raising%20standards%20lowering%20barriers%20report_draft.pdf.

- [5] Wellcome exists to improve health for everyone by helping great ideas to thrive[EB/OL]. [2018-05-10]. <https://wellcome.ac.uk/about-us>.
- [6] LSHTM Research Data Management Policy[EB/OL]. [2018-05-10]. http://researchonline.lshtm.ac.uk/612422/1/LSHTM_RDM_Policy_v2.pdf.
- [7] Southampton data survey: our experience and lessons learned[EB/OL]. [2018-05-10]. www.disc-uk.org/docs/SouthamptonDAF.pdf.
- [8] A year in the life of an RDM Service: How we helped LSHTM researchers in 2015 [EB/OL]. [2018-05-10]. <http://blogs.lshtm.ac.uk/rdmss/rdmservice2015/>.
- [9] Prepare a LSHTM Data Management Plan[EB/OL]. [2018-05-10]. http://blogs.lshtm.ac.uk/rdmss/lshtm_dmp/.
- [10] Knight G. EPSRC Research Data Management Compliance Report. [EB/OL]. [2018-05-10]. http://researchonline.lshtm.ac.uk/3716463/1/epsrc_compliance_report.pdf.
- [11] PLOS data policy summary[EB/OL]. [2018-05-10]. <http://blogs.lshtm.ac.uk/rdmss/files/2014/03/PLOS-Data-Policy-Summary.pdf>.
- [12] Writing a wellcome Trust Data Management & Sharing Plan[EB/OL]. [2018-05-10]. http://researchonline.lshtm.ac.uk/3716764/1/WellcomeTrust_DMP_Guide.pdf.
- [13] 卫军朝, 蔚海燕. 英国高校研究数据管理需求调研实践分析[J]. 图书情报知识, 2016(4): 83-92.
- [14] 利好消息! 中央深改组审议通过《科学数据管理办法》[EB/OL]. [2018-05-10]. http://www.sohu.com/a/218719508_398084.
- [15] 唐源, 吴丹. 国外医学科学数据共享政策调查及对我国的启示[J]. 图书情报工作, 2015, 59(18): 6-13.
- [16] MRC Data Sharing Policy [EB/OL]. [2018-05-10]. <https://mrc.ukri.org/documents/pdf/mrc-data-sharing-policy/>.
- [17] 王璞. 英美两国制定数据管理计划的政策、内容与工具[J]. 图书与情报, 2015(3): 103-109.
- [18] Research Data Management Support Service Project Initiation Document (PID) [EB/OL]. [2018-05-10]. <http://blogs.lshtm.ac.uk/rdmss/files/2012/12/RDMSS-PID-v12.pdf>.
- [19] Research Data Management Support Service Project Initiation Document (PID) [EB/OL]. [2018-05-10]. <http://blogs.lshtm.ac.uk/rdmss/files/2012/12/RDMSS-PID-v12.pdf>.

[收稿日期: 2018-05-29]

[本文编辑: 黄思敏]