

科学数据咨询服务探索与实践^{*}

张 玢 李 越 刘晓婷 唐小利

(中国医学科学院医学信息研究所/图书馆 北京 100005)

〔摘要〕 在阐明数据服务的概念和发展现状的基础上,介绍中国医学科学院图书馆开展数据咨询服务的实践,包括数据源推荐、数据检索、数据分析、数据规范、数据更新监测等服务内容,分析目前开展科学数据咨询服务存在的问题并提出建议。

〔关键词〕 医学图书馆;数据咨询服务;科学数据

〔中图分类号〕 R-058 〔文献标识码〕 A 〔DOI〕 10.3969/j.issn.1673-6036.2014.12.015

The Exploration and Practice of Scientific Data Consulting Services in Library ZHANG Bin, LI Yue, LIU Xiao-ting, TANG Xiao-li, Institute of Medical Information/Library, Chinese Academy of Medical Sciences, Beijing 100005, China

〔Abstract〕 Based on elaborating the concept and development status of data services, the paper introduces the data consulting services practice in Library of Chinese Academy of Medical Sciences, which including the data sources recommendation, data retrieval, data analysis, data standardization, data update monitoring and so on. At last it analyzes the existing problems in current scientific data consulting service and puts forward advices.

〔Keywords〕 Medical library; Data consulting services; Scientific data

1 前言

科学数据是指科研人员在研究过程中所产生的有关试验、调查等的相关数据、图片等资源^[1]。作为科研活动的产物,科学数据是信息时代最基本的科技资源,对科技的发展和创新具有支撑作用,其重要性被越来越多的科研人员所认同,致使科学研究工作向着数据密集型转变。为了最大限度地发挥科学数据的价值,需要对其进行系统的管理,以方

便科学研究的合作与资源共享,有效提高科研效率。作为社会信息、文献收藏、利用、加工、服务机构的图书馆,需要关注与研究科学数据服务,使其成为馆服务的新领域^[2]及参考咨询的重要内容^[3]。本文在概述国内外图书馆科学数据服务现状的基础上,介绍中国医学科学院医学信息研究所/图书馆(简称“医科院图书馆”)尝试开展科学数据咨询服务的实践案例,同时分析目前开展科学数据咨询服务存在的问题和注意事项。

2 科学数据服务的概念与发展现状

2.1 概念

科学数据服务是以科学数据为资源管理对象的一种信息服务形式,通过统一的技术平台,对具有研究利用价值的科学数据进行收集、整理、分析、

〔收稿日期〕 2014-09-26

〔作者简介〕 张玢,副研究馆员,发表论文 20 余篇;通讯作者:唐小利。

〔基金项目〕 国家科技图书文献中心委托专项任务(项目编号:2013XM052,2013XM056)。

挖掘，再将经过整理分析的高价值数据传递给科研信息需求者，帮助科研人员解决问题，是一项围绕科学数据而开展的数据资源管理、利用和研究活动^[4]。科学数据服务的核心是协助科研人员开展科学研究，实现科学数据的现存价值和价值增值。科学数据服务是一种跨界服务、嵌入服务、动态服务，既适应科研人员的外在需求，也符合图书情报机构自身发展的内在需要。

2.2 发展现状

随着海量数据时代的来临，学术交流环境和科研环境的变化，科学研究正在向数据密集型转变，澳大利亚、美国、英国等一些大学图书馆已把科学数据服务作为图书馆服务的重要内容，服务形式包括数据存储与发布、数据发现与获取、数据管理规划、数据分析、数据引用、数据咨询和用户社区服务等^[5]。美国康奈尔大学图书馆从 2010 年开始提供各种数据管理服务，内容包括存储备份、数据分析、元数据加工、数据发布等，并且建立了数据仓储 DataStaR，主要保存生态系统和农业学科的研究数据。学科馆员在数据监管服务中提供数据咨询服务，同时帮助研究人员依据基金申请的需求和数据的特点来规划数据的管理，了解研究人员对数据的管理需求，参与数据保存标准的制定和数字仓储的建设^[6]。国外一些高校图书馆，如罗彻斯特大学图书馆、爱丁堡大学图书馆开始设置“数据馆员”（Data Librarian）或者“研究数据馆员”（Research Data Librarian），招聘专门从事数据管理服务的工作人员^[7]。王婉通过网站访问和文献研究，发现有 34 所澳大利亚高校图书馆开展科学数据管理服务^[8]。国内图书情报机构也已经开始关注科学数据的研究，但是真正付诸实践较少。2011 年武汉大学图书馆尝试在校内搭建科学数据共享平台并开始面向全校提供数据存储和共享服务^[9]。

3 医科院图书馆科学数据咨询服务实践

3.1 融入科研全过程的图书馆数据咨询服务

参考咨询是图书馆的核心专业工作内容，图书

馆拥有大量的信息，有做好科学数据咨询的资源、人才与技术条件，应当嵌入科研团队中开展科学数据咨询、管理与协助服务。国外许多大学图书馆都把数据服务作为参考咨询工作的重要内容^[10]。图书馆科学数据咨询人员应当围绕科学数据的生命周期，为科研人员提供个性化、有针对性的咨询服务，包括数据的需求调查、科学数据的规划、收集、加工、分析、发布、归档、保存和维护管理等一系列工作^[11]。如表 1 所示，科学数据咨询人员要从科研项目立项之初就主动融入项目组，从科研项目的调研、开展具体研究到结题，提供全程的科学数据跟踪式服务，这样不仅能节约科研人员管理数据的时间，而且还能收集到高质量的科学数据，为以后的科学数据再利用奠定基础^[12]。医科院图书馆从 2010 年开始面向国家的“重大新药创制”和“艾滋病和病毒性肝炎等重大传染病防治”重大科技专项开展专业化信息服务，服务团队积极嵌入到课题组和科研工作过程中，及时了解和掌握科研工作人员的数据需求，帮助他们检索、获取并协助管理科学数据。

表 1 融入科研全过程的图书馆数据咨询服务

科研阶段	咨询问题	服务内容
科研立项阶段	同研究领域的基金项目有哪些？	数据检索与分析
	有哪些研究者或机构可以合作？	数据检索与分析
	去何处注册临床试验？	数据源选择
科研实施阶段	与同类研究试验的异同点如何？	数据检索与分析
	有无新的基金项目、临床试验？	数据更新监测
	数研究数据如何进行存储与管理？	数据规范与管理
成果产出阶段	如何共享研究数据？	数据整理、格式化、提交
	如何选择合适的期刊投稿？	数据检索与分析
科研结题阶段	如保长期保存研究数据？	数据规范与管理
	如何评价研究质量？	数据检索与分析

3.2 数据源选择与获取

数据服务的开展依托于其信息源，数据源的质量决定数据服务的质量。在开展咨询服务的初期，科研人员多希望图书馆推荐可靠权威的数据来源，

或由图书馆进行数据源的导航服务。例如重大传染病临床研究领域的科研人员非常关注在其研究领域类似的临床试验情况,以及临床试验的设计方案。解决此类的临床研究问题,传统的图书文献已经不是最佳的信息来源,服务团队需要了解目前主要的国内外临床试验研究数据库及其获取途径。可以通过世界卫生组织(WHO)、中国临床试验注册中心(CHICTR)、美国国立医学图书馆(NLM)官方网站上免费注册获取部分数据,还可以通过采购商业数据库如 Clinical Trial Intelligence, Integrity, Cortellis 等直接检索获取。

3.3 数据检索

数据检索是图书馆的专业优势与特长。数字科研环境下图书馆员更应做好科学数据检索与查找服务,使科研人员能够及时、准确地了解其所需的科学数据。在为重大专项课题组开展服务的过程中,科研人员提出希望了解其关注的疾病领域发布了多少临床实践指南(包括诊疗标准、规范或者专家共识),有哪些领域专家参与指南撰写,指南中规定的疗效指标是什么,因此图书馆服务人员需要充分检索各类文献信息资源、学协会网站以及政府组织官方网站,全面获取该疾病研究领域的临床指南文献,对临床指南文献及其作者进行初步的统计分析,得出该疾病领域发布了临床实践指南的文献数量和专家数量。此外,服务人员还需要通过阅读指南文献,提取相关的疗效指标数据。

3.4 数据分析

数据分析是实现数据服务增值的重要手段。服务人员必须对数据内容和价值进行深入分析,充分挖掘科学数据与其他科学信息的关联。目前开展的数据咨询服务多数采用的是一些较为简单的分析方法和工具。例如有科研人员为了解某研究主题的发展脉络,希望服务人员对该主题研究文献情况、作者情况、主要国家、机构以及研究方向进行分析。服务人员检索 PubMed, Embase, Web of Science 及 Scopus 4 个文献数据库,借助 Thomson Data Analyzer

软件,采用文献计量学方法对检索结果进行详细分析,揭示当前国际研究发展状况、主要学者及机构,采用内容分析及分类汇总的方法探讨该主题领域的主要研究方向。服务人员也可利用 ESI, Innography, Incites, SciVal 等科研评价工具开展学科评价、引进人才或合作伙伴的学术影响力分析等服务,用以支持科研管理与决策。

3.5 数据规范

与传统图书馆管理的图书文献相比,科学数据的形式多样,包括文本、数字、图像、视频、算法、动画、模型等。学科类别繁多、专业交叉频繁。数据间的逻辑关系非常复杂,增加了科研人员管理数据的难度。为了科学合理地存储和共享海量的科研数据,图书馆服务人员可以利用自己的优势,与研究人员合作建立所关注领域的数据规范。图书馆可以根据用户的习惯与要求设计元数据,以便后期数据管理与共享。数据规范可保障同一事物、同一概念具有唯一的数据表达方式,为后期的统计、分析带来方便。例如课题组在整理科研成果过程中,同一期刊有刊名大写全称、小写全称、缩写、电子 ISSN、印本 ISSN 等不同形式,服务人员可以根据图书馆期刊数据库的相关信息,对成果论文期刊进行名称或 ISSN 的规范。

3.6 数据更新监测

数据咨询服务另外一项重要的内容是对数据源更新数据的监测。目前各类数据源都保持日更新、周更新、月更新或年更新。科研人员需要及时了解相关领域有无新的基金项目(临床试验)、研究论文和研究动态,数据服务人员需要掌握数据源的更新频率,熟练使用 Email Alert、RSS 等获取更新数据的工具或软件,及时对科学数据进行更新维护。随着为重大专项课题组服务与合作的深入,科研人员会在以往提出的检索分析需求基础上,提出数据更新的需求,要求对变化情况进行描述。服务人员通过对数据进行时间序列分析,找出发展趋势与规律,分析变化的原因,为用户提供有价值的参考建议。

4 问题与思考

4.1 亟待推动完善国家科学数据服务政策体系

国际组织制定了一系列科学数据共享指南与政策声明,为各国制定科学数据共享政策提供了指导和借鉴。我国目前科学数据共享政策十分薄弱,缺乏相对完善的科学数据共享政策体系,导致在科学数据服务工作中很难寻找到有关规范的支持。因此应当加快推动国家科学数据共享服务政策的完善,借鉴国际组织关于科学数据共享的指南与声明,根据中国国情建立相应的科学数据共享政策体系,一方面使广大科研用户认识到科学数据服务的重要性,另一方面也使国内的科学数据服务有章可循。目前医科院图书馆开展的数据咨询服务主要是帮助用户检索与管理科学数据,无法提供类似国外图书馆的数据保存与提交服务,这离实现国家科学数据共享体系还有很大的差距。

4.2 加快构建专业研究图书馆科学数据服务机制

科学数据服务是一项非常复杂的服务内容,它包含各个领域的问题,如数据描述、存储、分析、检索与利用、导航服务等。构建图书馆科学数据服务机制需要充分考虑各方因素,包括政策体系、制度体系、技术平台、管理机制等。只有深刻认识到科研创新与科学数据服务的重要关系,制定合理完善的服务管理制度机制,才能有效解决数据服务中的诸多问题,有效开展科学数据服务。目前我国还未出台有关科学数据服务的条例规定,缺乏相应的制度规范,导致科学数据服务与管理主体不明,数据资源分散在科研工作者和项目课题组中,很难实现资源共享以及有效的利用、增值。就医科院图书馆来说,目前尝试开展的数据咨询服务零散、不成系统,以后应当推动科学数据服务机制建设,用完善的制度来规范科学数据服务行为,才能真正实现科学数据资源的共建与共享。

4.3 用户需求尚需进一步培养与开发

就科研人员而言,在科研任务繁重的前提下,

主动存储和共享数据主要还是源于研究工作和发表成果的需要。在目前科学数据共享政策体系还不健全的情况下,许多科研人员倾向于自己对科学数据进行管理,没有认识到通过图书馆获到数据服务的便利性。一些科研人员欠缺元数据、数据收集、整理、分析、管理等知识与技巧。数据馆员需要加强与科研人员的合作,深入了解其对科学数据的需求,尽可能提供符合用户需求的个性化服务,帮助研究者在需要的时候获得相关数据支持。图书馆还可以开展科学数据管理培训服务,从数据管理、学术交流、知识产权等层面为科研人员开展培训,使用户数据管理理念快速提升。

4.4 数据服务水平与馆员知识技能有待提高

图书馆的数据服务水平受多种因素的影响,如数据服务人员数量、数据服务人员的知识与水平、数据资源采购资金、计算机软硬件设备、用户需求以及网络提供的相关服务水平等。数据咨询服务包括上述多种服务内容,所有这些服务都需要知识与技巧,不是所有的图书馆员都可以进行科学数据服务。与普通的参考服务工作者相比,数据馆员必须拥有更多的知识和技能,不但要熟练使用各类操作系统,还要能对数据以各种形式进行导入导出、清洗和可视化分析等。国外许多地区都提供了数据馆员学习与培训的机会,国内图书馆需要加强数据馆员的教育与培训,也可直接引入所需人才,使馆员具备科研数据管理与服务所需的专业知识与技能。

4.5 服务过程中注意版权风险规避

与其他图书馆文献信息服务一样,科学数据服务也会涉及知识产权,数据引用的规范与统一问题仍然存在。在数据服务过程中,还要考虑哪些资源可以不受限制地访问,哪些需要在校园网内访问,特别是通过赠与或者代理方式以及从因特网上收集的资源,一定要注明来源并尽量限定在校园网访问。随着版权问题的日益突出,数据产品的版权问题成为一个热点问题,因此在数据服务时要考虑版权风险的规避。必要时参与数据服务的图书馆员需签署保密协议,保证数据不被泄露;同时还需帮助

科研人员提高数据管理意识和能力,使科研人员能很好地利用系统,安全有效地管理数据。

5 结语

科学数据是科学研究的基础,可以直接或间接地为科学数据创建者、使用者带来经济效益,同时,在提高全民素质、开展自我教育、监督违规行为、维护社会稳定、帮助政府决策等方面发挥着重要作用。图书馆应当把科学数据作为服务与管理的新内容,及时调整定位,进一步重视科学数据服务。国外图书馆科学数据服务已有成功经验值得借鉴,国内开展科学数据服务也要因地制宜,循序渐进。大力推动国家科学数据服务政策的制订与完善,建立健全科学数据服务制度体系,充分发挥图书馆自身的专业优势,设立专门的科学数据岗位,为数据密集型科学研究活动提供有力的数据资源支撑。

参考文献

- 徐坤,曹锦丹. 高校图书馆参与科学数据管理研究[J]. 图书馆论坛, 2014, (5): 92-98.
- 陈建新. 科学数据服务: 图书馆服务的新领域[J]. 图书与情报, 2013, (4): 93-95.
- Medforth J. Data Protection for Library and Information Services [J]. Program - electronic library and information systems, 2002, 36 (2): 131-132.
- 李晓辉. 图书馆科研数据管理与服务模式探讨[J]. 中国图书馆学报, 2011, (5): 46-52.
- 洪程. 国外科学数据服务现状研究[J]. 图书馆杂志, 2012, (10): 31-34.
- 李力. 国外研究型图书馆学科服务的发展态势及启示——以康奈尔大学为例[J]. 图书馆学研究, 2013, (14): 82-85.
- 孟祥保, 钱鹏. 国外高校图书馆数据馆员岗位设置与管理机制[J]. 图书与情报, 2013, (4): 12-17.
- 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, (3): 130-136.
- 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, (6): 33-38.
- 安艳杰, 黄良军. 图书馆参考咨询新领域——数据服务[J]. 情报资料工作, 2008, (3): 94-96.
- 师荣华, 刘细文. 基于数据生命周期的图书馆科学数据服务研究[J]. 图书情报工作, 2011, 55 (1): 39-42.
- 沈婷婷, 卢志国. 科研项目不同阶段的科学数据监管方法[J]. 图书馆建设, 2013, (3): 49-51.
- 郭向飞, 李树民, 黄晓鹏. 我国文献检索课教学方法相关研究文献分析[J]. 医学信息学杂志, 2012, 33 (12): 85-88.
- 钟宇, 胡芳. 北京地区 211 重点高校文献检索课教学内容的调查与分析[J]. 图书馆学研究, 2010, (23): 8-11.
- 柯临慧, 袁荣华. 成教班医学文献检索课的教学改革思考[J]. 农业图书情报学刊, 2008, 20 (6): 119-120.
- 黄晴珊. 国内医学文献检索网络课程开发的现状及建议[J]. 情报科学, 2005, 23 (5): 796-800.
- 朱伟丽. 基于翻转课堂的文检课教学设计[J]. 图书馆杂志, 2013, (4): 87-90, 112.
- 曹洪欣, 汤琳, 单甜甜. 非医学信息学专业医学信息学教育课程的设置[J]. 中华医学教育探索杂志, 2012, 11 (11): 1093-1096.
- 曲保丽. 医学信息检索课教学的改革实践[J]. 医学信息学杂志, 2011, 32 (11): 87-89.
- 黄利辉, 张燕舞, 张玢, 等. 8 年制医学生文献检索课程改革实践[J]. 医学信息学杂志, 2012, 33 (8): 86-88.
- 袁永旭, 贺培风, 于琦, 等. MOOC 背景下的医学信息检索课教学改革[J]. 中华医学图书情报杂志, 2014, 23 (7): 13-14.
- 张稚鲲, 李文林, 郝桂荣. 文献检索网络课程教学模式探索与实践[J]. 医学信息学杂志, 2011, 33 (12): 80-83.
- 马淑坤, 周瑛, 邸铭, 等. 文献检索课一体化教学与开放实习室教学研究[J]. 医学信息学杂志, 2013, 34 (11): 88-90.
- 陆华娟. 从传统文献检索课向信息素养教育转型的思考[J]. 医学信息学杂志, 2009, 30 (7): 89-91.