

大数据视角下的临床医学学科服务

赵志清 邹立君

(华中科技大学图书馆医学分馆 武汉 430030)

〔摘要〕 探讨大数据时代下的临床医学学科服务,包括临床研究和临床医疗两个方面,指出开发利用医疗大数据要求学科馆员和医生具有一定的信息素养,对学科馆员和医生信息素养的提升提出建议。

〔关键词〕 大数据;临床医学;学科服务

〔中图分类号〕 R-056 〔文献标识码〕 A 〔DOI〕 10.3969/j.issn.1673-6036.2015.10.017

Subject Service of Clinical Medicine from the Perspective of Big Data ZHAO Zhi-qing, ZOU LI-jun, Medical Library, Huazhong University of Science and Technology Library, Wuhan 430030, China

〔Abstract〕 The paper discusses the subject service of clinical medicine in the age of big data, including two aspects: clinical research and clinical medical treatment. It points out that the development and utilization of medical big data require for subject librarians and doctors with certain information literacy and proposes suggestion for enhancing the information literacy of subject librarians and doctors.

〔Keywords〕 Big data; Clinical medicine; Subject service

1 引言

随着计算机技术的飞速发展,人们获得、储存、分析和应用数据的能力不断增强,大数据时代已悄然来临并带来了生活、工作和思维的大变革。如今大数据应用在许多国家和地区的众多领域,如在生物医药领域,美国、欧洲和日本分别建立了世界 3 大生物数据中心,即美国国家生物技术信息中心(National Center for Biotechnology Information, NCBI),欧洲生物信息研究所(European Bioinformatics Institute, EBI)和日本 DNA 数据库(DNA Data Bank of Japan, DDBJ)。中国政府也认识到大数据的重要意义,李克强总理 2014 年 11 月 15 日提出在疾病防治、灾害预防、社会保障和电子政务等

领域开展大数据应用示范,这为医务工作者们打开了生物学大数据共享的大门。

大数据时代下信息呈爆炸式增长,据统计,每年有超过 200 万篇文献在 21 000 余种生物医学期刊上发表,且以每年 4% 的速度递增。如何在信息海洋中有效而准确地获取本专业的最前沿信息,应用于自身临床工作和科学研究,提高工作和科研效率,是急需解决的问题,也成为医学图书馆开展临床医学学科服务的关键环节。

2 大数据临床医学学科服务

2.1 大数据临床研究学科服务

2.1.1 调查用户需求 提供学科服务的前提是知道对方需要什么,因此学科馆员首先要开展需求调查,通过问卷、深入座谈及参加科室研讨会等形式,了解医生的信息行为,获取信息的途径和使用图书馆资源的频率及偏好等,掌握其主要研究方向

〔修回日期〕 2015-06-26

〔作者简介〕 赵志清,硕士,中级职称,发表论文 6 篇。

和内容,以及希望从图书馆和馆员处获得哪些帮助。各临床专科研究的疾病不同,但也存在交叉,如肺癌可进行内科治疗,也可进行外科治疗,学科馆员在深入了解各课题组的研究内容后,可将有类似研究领域的服务对象分门别类,划分成不同的小组,这样更有利于开展针对性学科服务,也有利于不同领域专家的学术争论与合作,激发创新点,促进研究更快发展。

2.1.2 搭建对口数据平台 北京市神经外科研究所张亚卓认为,大数据时代下的医学科研应该具有 3 个主要特点:一是临床医学资源发现、收集、整理和利用(标准化、科学化);二是大样本、多中心的临床研究(病因、诊断、治疗及药物开发);三是依托临床信息的生物医学资源整理、利用及研究(基因、蛋白和代谢)^[1]。可见大数据的整理是良好利用的前提,并且要开展机构间的合作,扩大大数据的范围,实现更大范围的数据共享。学科馆员可根据医生的需求,尽量丰富对口馆藏资源;收集服务对象对口的大数据资源,如生物医学大型通用数据库等;辅助医生收集整合本院数据,建立具有专科特色的小数据库,开展合作,统一标准,逐步扩大数据量;定向追踪竞争对手科研信息。国家科学图书馆机构知识库、清华大学 ArnetMiner 和爱思唯尔(Elsevier)的 Mendeley 等都是很受关注的学术网络分析挖掘系统,馆员可参考其功能设计,将上述资源整合到一个平台上,与医生共同丰富平台资源,提高馆员工作效率和医生科研效率。适用于某一临床医学专科的数据库或大数据处理平台非一朝一夕就能建成,因此,馆员要持续关注现有大数据平台发布的相关信息,辅助医生为其添砖加瓦,扩大数据覆盖面积,要积极利用图书馆现有数据资源和数据处理能力,开展优质学科服务。

2.1.3 利用平台辅助研究和成果发表 医疗大数据可应用于疾病诊断和治疗研究,如采用目的序列捕获测序技术对已知的特定基因组区域设计探针,捕获后测序,分析人类基因组中相关区域与疾病的相关性;可应用于疾病流行病学研究,如利用大数据挖掘某些地区大气污染物与恶性肿瘤发病和死亡

的相关性;还可应用于药物副作用分析等方面,如从大量病人数据中挖掘到与某种药物相关的不良反应或从社交网络中搜索服用某种药物的不良反应记录。学科馆员的数据挖掘应围绕这些方面展开,针对课题组的研究方向有的放矢。学科馆员可借助汤森路透数据分析(Thomson Data Analyzer, TDA)等情报分析工具和科学引文索引(Science Citation Index, SCIE)、期刊引用报告(Journal Citation Reports, JCR)或基本科学指标(Essential Science Indicators, ESI)等数据库中的分析功能,采用文献计量分析等方法,挖掘某一领域的研究热点、核心期刊、高被引文献、核心人员、核心机构和专利申请情况等,帮助医生们掌握自身研究领域的最新进展和发展动态,遴选投稿杂志,选择合作人员和机构,跟踪竞争对手,发现学科交叉创新点,过滤已有研究观点和专利,提高研究创新性和专利申请成功率。学科馆员需要对海量信息进行剔重、过滤、归纳、推理及联想,发现数据间的内在关联,发掘潜在的有价值的数据,以图表等直观形式展示给医生。

2.1.4 利用平台汇总科研成果 2011 年美国国家科学基金会(National Science Foundation, NSF)要求所有基金申请者必须提交包括数据的长期保存、访问方式与共享等内容的研究数据管理计划^[2],提醒研究者们要有意识地汇总保存自己的研究数据。学科馆员可协助建立课题组和医生个人研究档案,使其研究均有迹可循,辅助课题申请和成果发表,使其研究成果(论文、专利和获奖情况等)一目了然,随时可用于成果鉴定或报奖。

2.2 大数据临床医疗学科服务

2.2.1 挖掘医疗大数据辅助临床诊疗 专科医生不可能也没必要通晓所有专科疾病知识,当遇到疑难杂症或需跨专科诊疗时,在拥有大型临床数据系统的理想情况下,可将病人的各项指标输入临床数据系统,通过挖掘提供相似病例或可能诊断,再结合病人的具体情况和医生的经验进一步进行鉴别,制定较佳临床诊疗路径,降低医疗成本和病人经济负担。若临床数据系统中没有此类病例,医生可收

集其整套记录并录入系统,丰富大数据系统的内容,供他人借鉴。但建立标准、大型的临床数据系统存在一定困难,美国特拉华大学生物信息学和计算机生物学中心吴慧华主任表示,不同系统和科研机构之间难以实现标准化的数据共享和分析是生物医学与信息科学结合过程中遇到的关键难题^[3]。如一个病人会在不同医院就诊,但各家医院拥有各自的电子病历系统,疾病资料不能被便捷共享,医生要调阅某个病人某一疾病自始至终的发生发展过程及治疗经历几乎不可能,从一定程度上妨碍了疾病的快速诊断和治疗。建立临床数据系统还涉及病人的隐私安全问题,由于国内缺乏相关法律法规,使信息安全和保密等无法可依。另外,各国语言不同和非结构化数据需要被结构化处理等,也阻碍了大数据的形成和利用。

2.2.2 匹配小数据 医学的服务对象是人,由于个体情况千差万别,即使患有同一种疾病,表征也可有不同,因此在医疗大数据使用中,要兼顾病人的个体差异,利用大数据发现规律并用小数据匹配个人,寻找最佳诊疗方式。如肺癌术后,可将手术方式、转移情况、KRAS 状态、病理分型和标志物类型等匹配大数据,寻找适合的化疗药物和方案,再根据病人对药物的敏感性和经济条件等个人特点,选择最佳的化疗药物和方案。

2.2.3 提供健康信息服务 无论是健康人群还是病人及其家属,都希望获取健康信息,以促进健康或减轻痛苦,尤其是患有儿科疾病、慢性病和肿瘤等的病人和家属,更迫切地想了解疾病相关知识,但统计显示我国居民健康素养仍处于较低水平。国家卫生和计划生育委员会 2014 年 12 月 17 日发布的报告显示,2008 年我国居民健康素养水平为 6.48%,2012 年为 8.80%,2013 年为 9.48%^[4],虽然逐年有所增加,但仍有待提高。学科馆员可协助医生设立健康信息服务网页、微信或博客等,普及本专科疾病或常见病的诊疗相关知识。彭琰等^[5]认为健康信息还可包括各种治疗方案及其效果的描述、选择医院和医生的方法、健康保险和替代疗法等相关信息。健康信息服务就像在医生和病人之间架起一座沟通的桥梁,可以提高病人及家属的健康

素养水平,减轻对疾病的恐惧,帮助他们正确面对疾病,从而积极配合医生、护士的治疗和护理,缓解医患关系。

3 学科馆员和医生信息素养的提升

3.1 概述

开发利用医疗大数据需要馆员和医生们具有一定的信息素养,包括信息意识、信息知识、信息能力和信息道德。国外高校图书馆科学数据素养教育是从数据的生命周期角度出发,围绕科研流程所需的技能展开^[6],值得借鉴。

3.2 学科馆员信息素养的提升

由通晓生物医学和信息科学的复合型人才担任临床医学学科馆员是最佳选择,他们能更好地解决生物医学大数据利用中的融合问题。虽然国内许多高校顺势而为,开设了医学信息学专业,培养了许多跨学科人才,但远远不能满足需求,现实情况是医院或医学院图书馆员往往只具备信息科学背景。因此学科馆员要在学科服务的过程中不断增长生物医学知识,但更需培养对网络资源、数据库的数据挖掘能力,辅助医生做出决策。特别要注意的是信息道德的培养,在学科服务过程中,学科馆员会接触到服务对象的隐私和研究信息等,对这些信息要保密,不能随意泄露。

3.3 医生信息素养的提升

医生往往不具备信息科学背景,但在工作和科研的过程中,会逐渐掌握一些信息技能,一般都不系统。一些机构面向医务工作者开展了针对性培训,内容涉及信息检索、组织、评价与利用、数据挖掘等方面,如中国医科大学医学信息学院开发的卫生人员信息素质培训平台向医生提供系统化信息素养培训,取得了良好效果。医院和图书馆应加强相互合作和资源共享,学科服务的介入使得医生可以借助学科馆员的信息能力,辅助其科研和临床工作,同时也使学科馆员能更快、更准地获得所需的医学知识,互相促进,共同进步。

4 结语

大数据时代既是临床医学学科服务的机遇也是挑战,学科馆员要充分认识大数据的重要性,掌握大数据挖掘的方法,丰富自身医学知识,利用现有资源的同时,尽力开发新资源,逐步开展优质化的学科服务。但信息的爆炸式增长不可避免地会使劣质数据越来越多,数据质量问题及其所导致的知识和决策错误已经在全球范围内显现,如美国由于数据错误而引发的医疗事故每年可导致约 98 000 人死亡,约占全部医疗事故致死人数的 50%^[7]。因此学科馆员应和医生一同去辨伪存真,避免劣质数据的不良影响,促进临床和研究工作的良性发展。

参考文献

1 张亚卓. 医疗大数据 临床怎么用 [J]. 吉林医学信息,

2014, 30 (8): 32-34.

- 2 肖健, 李焱, 冯占英, 等. 大数据及其对医学图书馆的启示 [J]. 中华医学图书情报杂志, 2014, 23 (4): 29-34.
- 3 王庆. 大数据: 生物医学的待解之题 [N]. 工人日报, 2013-03-22 (6).
- 4 中华人民共和国国家卫生和计划生育委员会. 2013 年我国居民健康素养水平提高至 9.48% [EB/OL]. [2014-12-17]. <http://www.nhfp.gov.cn/xcs/s3582/201412/971753f8b9504caba6e081cb88cf6a58.shtml>.
- 5 彭琰, 严莉, 朱红. 大数据时代用户健康信息学的价值 [J]. 医学信息学杂志, 2014, 35 (1): 2-6.
- 6 孟祥保, 李爱国. 国外高校图书馆科学数据素养教育研究 [J]. 大学图书馆学报, 2014, (3): 11-16.
- 7 Kohn L T, Corrigan J M, Donaldson M S. To Err is Human: building a safer health system [M]. Washington: National Academies Press, 2000.

2016 年《医学信息学杂志》征订启事

《医学信息学杂志》是国内医学信息领域创刊最早的医学信息学方面的国家级期刊。主管: 国家卫生和计划生育委员会; 主办: 中国医学科学院; 承办: 中国医学科学院医学信息研究所。中国科技核心期刊 (中国科技论文统计源期刊), RCCSE 中国核心学术期刊 (武汉大学中国科学评价研究中心, Research Center for Chinese Science Evaluation), 美国《化学文摘》、《乌利希期刊指南》及 WHO 西太区医学索引 (WPRIM) 收录, 并收录于国内 3 大数据库。主要栏目: 专论, 医学信息技术, 医学信息研究, 医学信息组织与利用, 医学信息教育, 动态等。读者对象: 医学信息领域专家学者、管理者、实践者, 高等院校相关专业的师生及广大医教研人员。

2016 年《医学信息学杂志》国内外公开发行, 每册定价: 15 元 (月刊), 全年 180 元。邮发代号: 2-664, 全国各地邮局均可订阅。也可到编辑部订购: 北京市朝阳区雅宝路 3 号 (100020) 医科院信息所《医学信息学杂志》编辑部; 电话: 010-52328673, 52328674, 52328671。

《医学信息学杂志》编辑部