

• 医学信息组织与利用 •

大数据环境下医学信息服务平台研究与实践^{*}

虞颖映 马凌飞 胡天天

辛均益

(浙江省医学科学院 杭州 310012)

(杭州医学院 杭州 310053)

〔摘要〕 分析大数据环境对传统医学信息服务的挑战, 阐述构建医学信息服务平台的基础及建设目标、架构、功能特点, 提出基于数字化平台的医学信息服务模式, 包括医学信息资源、情报研究及卫生决策支持服务。

〔关键词〕 信息服务平台; 医学资源; 大数据; 情报研究; 卫生决策

〔中图分类号〕 R-056 〔文献标识码〕 A 〔DOI〕 10.3969/j.issn.1673-6036.2016.12.016

Research and Practice of Medical Information Service Platform under the Big Data Environment YU Ying-ying, MA Ling-fei, HU Tian-tian, Zhejiang Academy of Medical Sciences, Hangzhou 310013, China; XIN Jun-yi, Hangzhou Medical College, Hangzhou 310053, China

〔Abstract〕 The paper analyzes the challenges of big data environment to the traditional medical information service, illustrates the base of constructing medical information service platform and its goals, framework function characteristics, proposes the medical information service mode based on the digital platform, including the medical information resource, intelligence research and health decision-making supporting service.

〔Keywords〕 Information service platform; Medical resources; Big data; Intelligence research; Health decision-making

1 引言

大数据是全球数据自发式、爆发式增长以及数据处理高效率要求的综合体现, 伴随着这一浪潮的兴起, 信息技术正逐渐与社会生态系统深度融合,

“数字化”已渗透到各个领域和行业, 医学领域当然也不例外, 无论是教学、科研、临床, 还是管理模式、服务流程等都受到了前所未有的冲击, 发生了巨大变化^[1-2]。医学信息服务作为其中的分支领域, 大数据环境不仅使医学信息本身具有种类多、数量大、更新快的特点, 而且也改变了医学信息的生产、收集、存储、加工、组织和传播方式^[3]。

20 世纪 90 年代以来, 信息资源建设进入快速发展期, 实体馆藏文献数字化建设、数据库建设、网络信息资源建设和信息资源共建共享作为海量信息有效整合的载体, 受到广泛关注^[4]。在此大环境

〔收稿日期〕 2016-09-02

〔作者简介〕 虞颖映, 硕士, 副研究员, 主要从事医学情报研究, 发表论文 30 余篇。

〔基金项目〕 浙江省科技厅项目“基于数据挖掘技术的卫生行业网络舆情监测预警平台建设及应用”(项目编号: 2014F30019)。

下, 各类大型数据库如 CNKI、CBM、PubMed、ScienceDirect 等迅速发展, 医学类门户网站如丁香园、科学网、Nature、Science 等快速更新, 政府组织和机构如世界卫生组织、国家食品药品监督管理总局、疾控中心等也积极开通官方网站实现政策指南共享, 数据采集的范围、种类、数量和渠道越来越多样化, 海量数据的存储和访问技术日益更新, 数字化时代的发展趋势对各类信息资源的组织、开发和利用产生了新的要求。从信息服务的角度, 通过大数据提取知识和信息挖掘分析能力正成为信息服务的主要竞争指标^[3]。传统的医学信息服务主要基于实体馆藏资源和各类纸质文本资料, 而大数据时代下服务机构的实体化功能正在弱化^[5]。且传统服务模式存在信息响应时间长、处理过程复杂、服务流程不规范、各部门间难以共享、用户反馈滞后、信息难以重复使用、信息利用率不高等问题, 无法形成信息服务的整体良性循环, 难以满足大数据环境下医学创新对信息的要求。因此, 及时调整服务模式, 运用智能化方式强化对医学信息资源的筛选, 突破时间和空间的局限, 实现信息、用户与平台的共建共享, 从而提供更有价值、满足个性化需求^[6-7]的医学信息服务, 是医学信息工作探索和改革的重点。

2 构建医学信息服务平台的基础

2.1 数据库资源和信息化基础

传统的医学馆藏文献资源以及电子文献的发展, 形成了科学、合理的数据库资源体系^[8], 如中国知网、万方数据知识服务平台、维普数据库、国家科技图书文献中心、医药卫生科技成果数据库、Medline、Web of Science、外文全文数据库等大型数据产品, 以及特色专题、二次文献数据库等针对性更强的专题数据库, 这些数据库覆盖医药卫生的各分支领域, 为信息服务提供了基础。此外, 管理数据资源能力以及信息化建设基础也至关重要。通常, 医学信息服务细分为医学信息评价分析、医学数字资源建设、医药卫生决策支持等方面, 各项服务均具备较强的内容管理和数据集成处理能力, 这

为信息支撑平台在设计开发过程中进行数据采集、数据集成、数据分发并最终形成全业务信息化系统打下基础。

2.2 人力资源

从事医学信息服务的人员兼具医药卫生学专业背景和信息情报研究能力, 对现有多种类型的医学信息系统如医院管理信息系统、政府卫生信息网络平台、科研项目查新与定题服务平台、科技文献共建共享平台等开发建设和应用过程中, 均积累丰富的实践经验, 能承担项目开发、测试、维护等工作, 人员的信息服务和平台管理综合能力可满足构建医学信息服务平台的人力资源需求。

2.3 社会需求

根据工作业务范围, 医学信息服务对象通常来自医药卫生系统政府机关、医院、科研院校、制药企业等, 具有较高的专业程度和社会认可度。这些用户通常要求服务单位具有反应快速、信息准确高质、服务个性化、市场竞争力强等能力, 这就为构建医学信息服务平台提供了机遇和挑战。以此整合医学信息资源和专业技术资源, 改善信息获取、信息加工和传递滞后的现状, 由面向单个单位、独立项目的“点对点”服务模式转变为面向各级各类用户的“全方位立体”服务模式, 具有现实需要和社会意义。

3 医学信息服务平台构建过程

3.1 建设目标

任何信息源的效用都取决于它的相关性, 随着准确性提升而倍增, 且随着信息获取工作量的增加而倍增, 因此, 医学信息服务平台的建设目标可概括为高相关性、高准确度和(用户的)低工作量^[9]。平台以医学信息资源为基础, 以用户服务为中心, 以信息服务技术为手段, 结合医药卫生行业信息需求、医学情报服务模式特点进行建设, 构建具有权威性、实时更新、易于查询和阅读的数字化平台。

3.2 平台架构模型

根据服务对象、使用角色和功能划分等原则，

将平台划分为数据集成、数据加工、信息分析、信息应用及管理 4 个基本系统。具体的平台架构，见图 1。

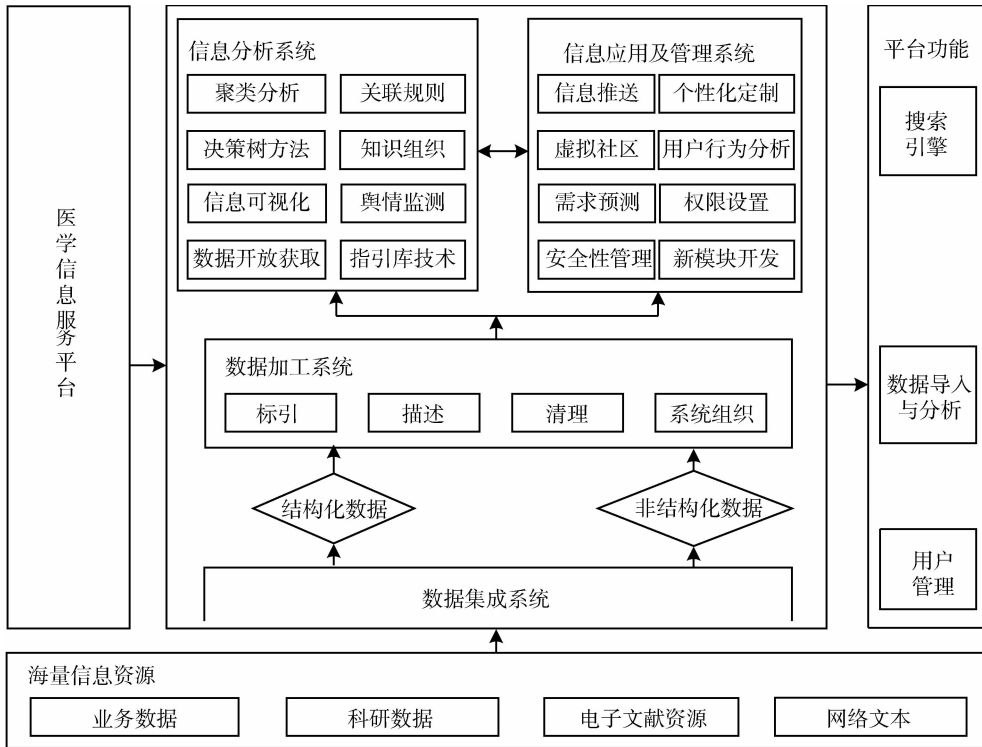


图 1 医学信息服务平台架构

数据集成系统对各类医学信息进行自动化采集，将业务数据、科研数据记录、电子文献资源、网络文本信息等相关数据进行整合，将不同来源、格式、含义特点的数据进行有序集中，为数据加工做好准备^[10]。数据加工系统在数据集成的基础上，将数据类型分为结构化（存储在各类型专业数据库里的数据）和非结构化（所有格式的办公文档、文本、图片、XML、HTML 等），采用标引、描述、清理、系统组织等方法进行规范化处理。大数据环境下数据加工更关注相关度的深度揭示，语义网和医学领域本体将发挥重要作用。信息分析系统主要基于数据挖掘技术及信息分析技术对规范化数据进行加工处理，如聚类分析技术识别不同数据集进行分类整合；关联规则进行数据信息之间相关性分析，了解用户需求、提供个性化服务；决策树方法对数据进行规律挖掘，从而通过分类规则对相应对象的类别进行预测；知识组织研究方法开展医学文献和

健康数据研究；此外还有信息可视化研究、热点追踪及舆情监测、科研数据开放获取、指引库技术等^[11-12]。信息应用及管理系统旨在与用户发生交互共享，借助人工智能和挖掘技术分析用户行为特征并预测用户需求，实现信息推送、个性化定制、虚拟社区等功能；同时对用户信息进行不同级别权限设置和安全性管理，保证系统内的数据不会在未经授权的情况下被访问或恶意篡改，具有相关新兴模块的可拓展性潜力。

3.3 平台功能特点

基于 4 个基本系统创建的平台不是对已有信息资源废弃重建，而是在思维上摆脱物流的观念，在技术上寻求人工智能、虚拟技术的支持，在理论上以系统工程等为依据，找到一种医学信息服务相关人员都满意的信息共建共享模式^[13]。平台的功能特点如下：（1）强大的搜索引擎，包括信息检索、浏

览与导航, 满足用户对不同类型数据、不同检索申请、不同知识的需求。(2) 数据导入与分析, 平台支持内部和外部用户进行数据的直接导入、录入等功能, 能进行数据的关联分析和知识挖掘, 提升知识资源的增值理念。(3) 用户管理, 能即时统计资源上载下载情况, 对用户类别、需求特点、使用反馈等统计分析, 构建动态的数字化平台^[14]。

4 基于数字化平台的医学信息服务模式 (图 2)

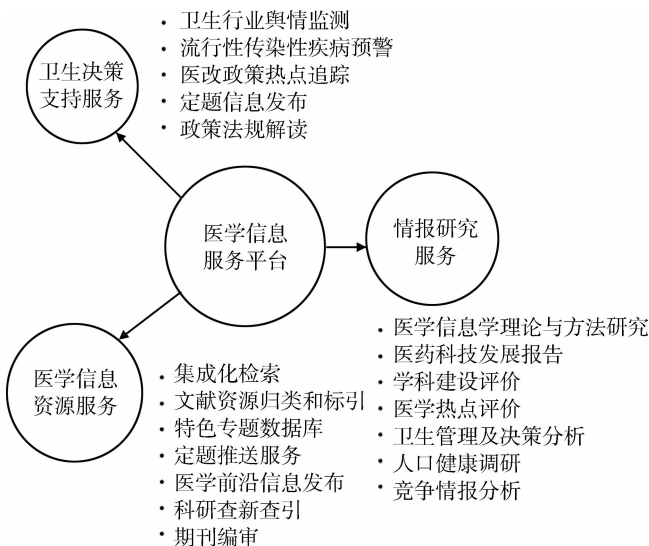


图 2 基于数字化平台的医学信息服务模式

4.1 医学信息资源服务

平台集成化的检索界面, 能有效帮助用户缩短检索时间, 提高检索效率, 完成文献请求和传递; 同时情报服务人员通过整合纸质和数字文献资源, 按照研究方向进行归类和标引, 可创建多样化的特色专题数据库, 进行定题服务, 满足用户个性化需求。此服务是平台最基本的功能, 即实现“一站式服务”。用户注册并通过审核, 登录可获取不同开放程度的信息资源服务。医学科研前沿信息发布、科研项目查新和期刊编审模块面向全体用户开放, 其中文献检索、特色资源模块面向科研人员开放, 主要提供各类文献、信息数据库系统链接以及分类信息推送和专题检索服务。用户还可根据需要, 浏览搜索最新医学研究进展, 提交科研项目查新查引

申请, 向期刊编辑部投稿等。

4.2 情报研究服务

情报研究服务是平台的重点功能, 根据情报调研科学性、预测性、服务性的特点, 旨在展示最新情报研究成果、发布情报调研与评价信息, 有利于各级用户了解服务范围和研究方向, 开展深度合作交流。展示的研究成果包含科研课题、学术论文、专利、科研奖励等类型, 涉及医学信息学理论与方法、医药科技发展、学科建设、医学热点评价、公共卫生研究、人口健康决策等领域, 根据政府、企业、研究院所、个人用户对信息需求的不同, 设置不同的权限和信息匹配资源, 确保研究成果的准确、高效传递。情报调研与评价服务借助信息分析技术和数据挖掘技术, 按照用户要求进行内容加工整合, 提供相关研究领域前沿性和竞争性的专题信息, 为有关部门提供解决方案和决策支持。用户也可借助平台进行服务咨询、情报调研申请等。

4.3 卫生决策支持服务

“十二五”期间政府加大了公共卫生和应急反应信息系统建设力度^[15], 信息化与公共卫生的关系日趋紧密。卫生决策支持服务可充分借助平台整合相关卫生信息化数据和政策资源, 发挥决策支持作用。卫生行业舆情监测预警服务基于互联网信息采集技术和数据挖掘技术, 实时动态监测新闻门户、社交网络等相关互联网站点, 及时发现网络突发性事件、敏感舆情、流行性传染性疾病疫情并发出预警, 为卫生管理部门提供及时、可信的决策支持参考。医疗卫生改革及基本药物制度热点分析服务旨在对海量的医改信息进行系统全面的提炼处理, 通过定题信息发布、动态追踪、政策法规解读等为相关决策部门提供参考依据。

5 结语

大数据时代下的医学信息服务具有知识性、增值性、创新性、集成性、专业性和个性化的特点^[16], 在此基础上, 借助智能检索、信息推送、数

据挖掘等搭建医学信息服务平台,可以有效整合国内外医学信息资源,规范服务工作流程,提升医药卫生行业情报服务和决策支持的效率,形成多种功能、互为支撑的完整服务体系。通过平台的建设和推广应用,可实现情报服务人员、政府部门、科研院所、企业和社会用户之间的良性互动,充分发挥平台的信息库和智库作用,实现共赢的同时,也可为医学科技创新发展带来动力。

参考文献

- 1 于施洋,王建冬,童楠楠. 大数据环境下的政府信息服务创新:研究现状与发展对策 [J]. 电子政务, 2016, (1): 26-32.
- 2 王正国. 数字化时代的医学革命 [J]. 中国数字医学, 2009, 4 (1): 8-11.
- 3 陈锐,冯占英,李焱,等. 大数据对生物医学信息服务各环节的影响研究 [J]. 图书情报工作, 2015, 59 (9): 68-72.
- 4 贺德方. 数字时代情报学理论与实践—从信息服务走向知识服务 [M]. 北京:科学技术文献出版社, 2006.
- 5 李国栋. 大数据下的医学文献情报创新 [J]. 中国科技信息, 2015, (13): 72-73.
- 6 刘冬云. 云计算环境下高校图书馆数字信息资源共建共享系统的构建 [J]. 图书馆学刊, 2014, (11): 99-101.
- 7 刘海龙. 关于网络环境下提高医学信息服务质量和效益的思考 [J]. 健康导报·医学版, 2015, 20 (12): 275.
- 8 彭晓东,王茂林. 基于高校图书馆的企业情报服务平台研究 [J]. 情报理论与实践, 2011, 34 (2): 58-61, 71.
- 9 于彤,张竹绿,贾李蓉. 面向循证医学的知识服务平台概述 [J]. 中国中医药图书情报杂志, 2014, 38 (8): 55-57.
- 10 栾旭伦. 大数据环境下高校图书馆个性化信息服务系统研究 [J]. 图书馆学刊, 2014, (8): 118-121.
- 11 乔幸娟. 数据挖掘技术在数字图书馆中的应用研究 [J]. 农业图书情报学刊, 2014, 26 (12): 118-120.
- 12 张文惠. 数据挖掘技术提升高校图书馆水平 [J]. 电脑开发与应用, 2014, 27 (12): 49-51, 54.
- 13 黄东流,张旭,刘娅. 基于共建共享模式的知识服务系统建设研究 [J]. 情报杂志, 2011, 30 (3): 170-175.
- 14 王建文. 基于图书情报系统的知识服务能力优化策略 [J]. 科技创新导报, 2015, (15): 190.
- 15 万美. 卫生信息化视角下的医学信息资源建设 [J]. 医学信息学杂志, 2014, 35 (4): 77-79.
- 16 邓红巧. 略论知识服务在高校图书馆中的应用 [J]. 高校图书馆工作, 2010, 30 (6): 70-72.
- 1 于施洋,王建冬,童楠楠. 大数据环境下的政府信息服务创新:研究现状与发展对策 [J]. 电子政务, 2016, (1): 26-32.
- 2 王正国. 数字化时代的医学革命 [J]. 中国数字医学, 2009, 4 (1): 8-11.
- 3 陈锐,冯占英,李焱,等. 大数据对生物医学信息服务各环节的影响研究 [J]. 图书情报工作, 2015, 59 (9): 68-72.
- 4 贺德方. 数字时代情报学理论与实践—从信息服务走向知识服务 [M]. 北京:科学技术文献出版社, 2006.
- 5 李国栋. 大数据下的医学文献情报创新 [J]. 中国科技信息, 2015, (13): 72-73.
- 6 刘冬云. 云计算环境下高校图书馆数字信息资源共建共享系统的构建 [J]. 图书馆学刊, 2014, (11): 99-101.
- 7 刘海龙. 关于网络环境下提高医学信息服务质量和效益的思考 [J]. 健康导报·医学版, 2015, 20 (12): 275.
- 8 彭晓东,王茂林. 基于高校图书馆的企业情报服务平台研究 [J]. 情报理论与实践, 2011, 34 (2): 58-61, 71.
- 9 于彤,张竹绿,贾李蓉. 面向循证医学的知识服务平台概述 [J]. 中国中医药图书情报杂志, 2014, 38 (8): 55-57.
- 10 栾旭伦. 大数据环境下高校图书馆个性化信息服务系统研究 [J]. 图书馆学刊, 2014, (8): 118-121.
- 11 乔幸娟. 数据挖掘技术在数字图书馆中的应用研究 [J]. 农业图书情报学刊, 2014, 26 (12): 118-120.
- 12 张文惠. 数据挖掘技术提升高校图书馆水平 [J]. 电脑开发与应用, 2014, 27 (12): 49-51, 54.
- 13 黄东流,张旭,刘娅. 基于共建共享模式的知识服务系统建设研究 [J]. 情报杂志, 2011, 30 (3): 170-175.
- 14 王建文. 基于图书情报系统的知识服务能力优化策略 [J]. 科技创新导报, 2015, (15): 190.
- 15 万美. 卫生信息化视角下的医学信息资源建设 [J]. 医学信息学杂志, 2014, 35 (4): 77-79.
- 16 邓红巧. 略论知识服务在高校图书馆中的应用 [J]. 高校图书馆工作, 2010, 30 (6): 70-72.

(上接第 65 页)

- 2 孟刚,潘蕾,高菁,等. 医疗器械不良事件报表质量评估方法研究 [J]. 中国医疗器械信息, 2008, 14 (2): 43-47.
- 3 刘斌,翟伟,马宁,等. 医疗器械不良事件报告表质量评价方法探索 [J]. 中国药物警戒, 2011, 8 (3): 165-168.
- 4 施雯慧,姚捷,赵燕,等. 美国医疗器械不良事件报告数据库研究 [J]. 中国医药导报, 2014, (28): 148-152.
- 5 施雯慧,陈颖,姚捷,等. FDA 公共数据开放项目中屈螺酮炔雌醇片的分析研究 [J]. 中国药物警戒, 2015, 12 (9): 552-555.
- 6 化柏林. 知识抽取中的停用词处理技术 [J]. 现代图书情报技术, 2007, (8): 48-51.
- 7 Blei D, Carin L, Dunson D. Probabilistic Topic Models [J]. IEEE Signal Processing Magazine, 2010, 27 (6): 55-65.
- 8 Wu H, Luk R, Wong K, et al. Interpreting TF-IDF Term Weights as Making Relevance Decisions [J]. ACM Transactions on Information Systems, 2008, 26 (3): 1-37.
- 9 Grün B, Hornik K. Topicmodels: an R package for fitting topic models [J]. Journal of Statistical Software, 2011, 40 (13): 1-30.
- 10 Daniels K, Mosher WD, Jones J. Contraceptive Methods Women Have Ever Used: United States, 1982 - 2010, National Health Statistics Reports, 2013, No. 62 [EB/OL]. [2016-07-15]. <http://www.cdc.gov/nchs/data/nhsr/nhsr062.pdf>.