



[专家简介] 郭华源,解放军总医院医学创新研究部高级工程师,清华大学计算机科学与技术专业博士,主要研究方向为医学影像人工智能和医学大数据处理。参与和承担科技部、工信部、国家自然科学基金等多项课题,在医院信息系统集成、数字几何造型、裸眼3D显示等领域研究较深入。现任北京市卫生信息职工技术协会电子病历专委会委员,中国生物医学工程学会某分会第五届委员会委员。

医院大数据中心服务规范化建设研究

车贺宾,徐洪丽,吴欢,乌日力格,胡苹,何昆仑,郭华源
解放军总医院医学创新研究部,北京 100853

摘要: **背景** 信息时代数据价值越来越受到人们的重视,近年来一些大型医院陆续建立大数据中心,系统化开展医疗数据资源开发利用工作。**目的** 梳理解放军总医院医学大数据中心建设要点,总结服务经验。**方法** 介绍解放军总医院医学大数据中心服务体系架构、数据服务流程化模式、人员管理和数据安全。 **结果** 医学大数据中心支了解放军总医院多个国家、军队科研平台申报,成立6年产出与数据服务相关论文70篇,指导研究生近百名,有效发挥了医院科技创新公共平台的支撑作用。**结论** 医院大数据中心服务规范化建设带来了医院管理及临床科研服务上的变革,有助于医院科研能力顺应时代潮流快速发展。

关键词: 医院大数据中心; 医学数据服务; 流程化模式

中图分类号:R197.1

文献标志码:A

文章编号:2095-5227(2025)02-0134-06

DOI: 10.12435/j.issn.2095-5227.23060102

引用本文: 车贺宾,徐洪丽,吴欢,等.医院大数据中心服务规范化建设研究 [J].解放军医学院学报, 2025, 46(2): 134-139.

Research on standardized construction of hospital big data center services

CHE Hebin, XU Hongli, WU Huan, WU Rilige, HU Ping, HE Kunlun, GUO Huayuan
Medical Innovation Research Department of PLA General Hospital, Beijing 100853, China
Corresponding author: GUO Huayuan. Email: guohuayuan@301hospital.com

Abstract: Background People pay more and more attention to the value of data. Some large hospitals in the medical field have successively established big data centers to systematically carry out the development and utilization of medical data resources. **Objective** To sort out the key points of medical big data center service construction of Chinese PLA General Hospital, and summarize the experience. **Methods** The service system architecture, data service process-oriented mode, personnel management and data security management of medical big data center were introduced. **Results** The medical big data center had supported the application of many national and military scientific research platforms of the general hospital, produced 70 papers related to data services in the six years since its establishment, guided nearly 100 graduate students, and effectively played the supporting role of the public platform for scientific and technological innovation. **Conclusion** The standardized construction of the hospital's big data center service has brought about changes in hospital management and clinical research service, which is conducive to the rapid development of hospital scientific research ability in line with the trend of the times.

Keywords: hospital's big data center; medical data services; process-oriented model

Cited as: Che HB, Xu HL, Wu H, et al. Research on standardized construction of hospital big data center services [J]. Acad J Chin PLA Med Sch, 2025, 46(2): 134-139.

收稿日期: 2023-07-10

基金项目: 科技创新2030-新一代人工智能重大项目(2021ZD0140406)

第一作者: 车贺宾, 硕士, 医学数据工程师。Email: chehebin@163.com

通信作者: 郭华源, 博士, 医学人工智能高级工程师。Email: guohuayuan@301hospital.com

医院所积累的大量电子病历(electronic medical record, EMR)数据具有巨大潜力, 众多医院科室与医疗大数据厂商合作, 通过引入数字化医疗管理解决方案或工具以更好地管理患者, 提高临床研究效率^[1]。然而, 科室与厂商直接合作存在诸多弊端, 如患者信息来源多个信息系统, 信息科针对每个厂商的系统开发数据接口工作量大, 实施困难; 全院临床数据治理缺乏顶层设计, 很多数据质量问题长期得不到解决, 治理工作反复做; 多个科室对接多个厂家, 医院数据管理困难, 数据安全没有保障。

针对存在的以上问题, 一些大型医院陆续建立大数据中心, 系统化开展医疗数据资源开发利用工作^[2-3]。解放军总医院于2016年在国内率先成立了医学大数据中心, 为全院临床研究提供医疗数据服务, 在服务体系建设和管理模式方面率先进行了探索。

1 医院大数据中心服务体系

医院大数据中心服务体系以数据安全管理体系、数据存储和计算资源为基础, 配套数据治理和应用工具, 遵循数据服务流程, 实现临床数据价值转化应用。该服务体系既适用于临床科研的

一般需求, 如病例检索、数据整理和统计分析等, 也适用于医学影像深度学习、组学数据处理分析等医学大数据新需求。

1.1 大数据中心服务体系

医院大数据中心服务体系见图1, 基于数据安全管理体系防护, 自底向上主要包括数据资源、数据存储和计算资源、数据服务流程规范、数据服务工具体系和数据价值转化应用等^[4]。中心服务体系的构建在保障临床数据安全的同时, 实现了数据服务有序开展^[5]。

数据资源是医院大数据中心的基础, 在信息科的帮助下, 中心完成了解放军总医院8个医学中心医疗数据的融合汇聚, 集成住院病例600万例, 门诊病例1.2亿例, 初步形成了多中心的医疗数据资源。

数据安全管理体系贯穿整个数据服务流程, 是临床数据安全访问利用的保障, 主要包括数据库安全设置、运维安全审计和终端防护。数据库采用权限最小化原则设置用户权限, 开发测试环境与正式环境部署网络分离、密码分开设置^[6]。终端防护实现对所有网络终端的安全管理和准入控制, 保证合法授权用户进入网络、拒绝非法无授权用户进入网络或拷贝数据^[7]。运维安全审计对外

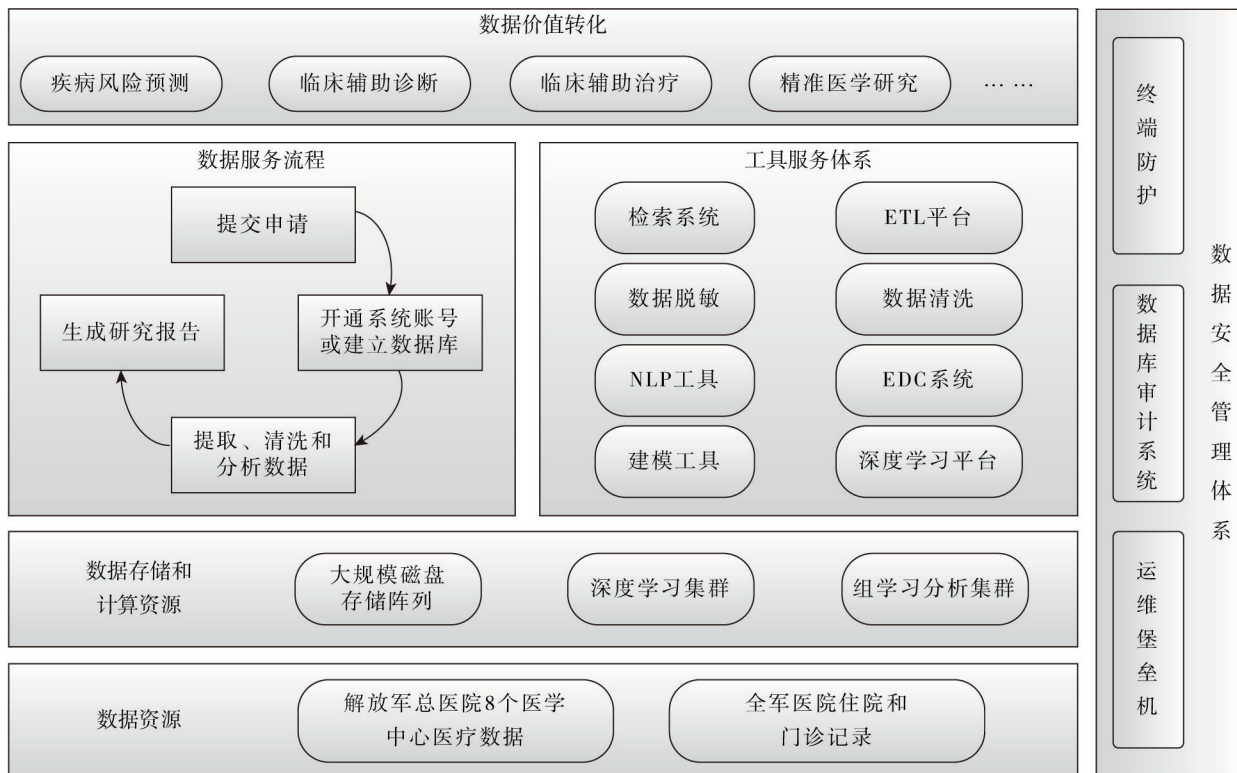


图1 医院大数据中心服务体系

Fig. 1 Architecture of medical big data center

提供统一登录入口, 监控人员通过数据库SQL操作和日志监控, 对数据库进行精准管理, 确保数据安全。

数据存储和计算资源是大数据中心的硬件基础^[8]。目前, 中心数据存储资源达到PB级, 存储内容主要包括医院临床数据, 如患者基本信息、就诊记录、检查报告、检验结果、门诊病历等结构化数据, 住院病历等文本数据和X线检查、CT、磁共振、超声等影像数据, 另外包括部分科研项目搜集的基因组学数据。多中心患者的临床数据融合汇聚, 主索引是患者唯一的身份识别途径, 可将医院不同数据表有效关联, 确保患者信息的统一完整^[9]。计算资源主要包括深度学习平台集群和组学分析平台集群, 配备高性能GPU, 满足医学影像深度学习和多组学数据整合分析^[10]。

数据服务内容主要包括特征提取、数据清洗、描述统计、数据预处理、建模分析、模型解释等, 建立相关的工具体系是大数据中心将数据处理分析能力赋能临床的重要手段, 主要包括病历检索系统、自然语言处理工具、专病数据库系统平台等, 主要解决临床科研人员研究中常见的数据不好查、提不到、难处理等问题, 实现医疗数据可检索、可处理、可统计分析。见图2。

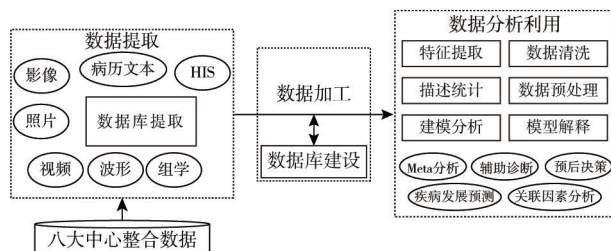


图2 医学大数据中心数据服务内容

Fig. 2 Data services of medical big data center

病历检索系统包括数据多维度筛选和查询结果可视化展示等功能模块, 目前多维度可检索变量包含人口学基本信息、就诊记录、诊断、医嘱、手术、生命体征、实验室检验结果、检查报告、电子病历等, 查询结果可以启发研究者新的思路或完善现有研究方案^[11]。自然语言处理工具包括原文预处理、变量管理、函数复用、结果审核等功能模块, 通过分析临床医学文本, 标识其核心概念术语及标志符的语义结构, 支持临床科研人员按不同需求从检查报告或病历文本中后结构化数据元^[12]。专病数据库系统平台根据研究方案设计个性化表单, 方便临床整理患者院内数据, 收

集随访数据, 长期存储管理并提供在线分析服务^[13]。以上数据服务科研工具与人工服务相辅相成, 有效满足临床大量常规化、部分复杂化的科研需求^[14]。

1.2 数据服务流程化模式

医院大数据中心建立数据服务流程化模式, 通过制度约束、系统监管相结合的方式实现服务申请、审核批准、数据服务(检索提取、清洗加工、分析建模)、成果复核、资料归档等过程全面监管, 确保数据服务工作运行有序, 保障数据处理分析质量和数据利用安全可控。

1.2.1 数据服务申请制度 数据服务申请制度旨在加强和规范医院大数据中心公共科研平台服务流程, 明确人员职责, 确保平台健康有序发展, 具体操作流程见图3。

(1)在线申请。通过解放军总医院公共科研平台, 临床科研人员在线提交数据服务申请, 服务内容包括技术资源及指导、数据检索与提取、文本特征结构化、文本数据清洗与描述性分析、数据建模与模型评价、影像数据提取、影像/波形数据分析、组学数据分析、专病库建设与数据集成、软件工具包租用、硬件资源租用等23项计时有偿条目。中心指定专人为公共科研平台管理员, 负责对提交的材料进行初步审核, 资料完备, 初步审核通过; 若资料不完整, 则退回修改, 并备注修改要求。

(2)资料审核。数据服务申请通过后, 科研人员在系统导出《医学大数据服务申请审批表》及附件, 经所在科室领导签字同意后交平台管理员, 中心负责查验相关材料, 并形成审核记录, 审核通过后提交评估技术组论证。

(3)项目论证。评估技术组对项目研究方案的可行性、服务项目的合理性和各项服务预估耗时产生的费用进行讨论, 形成会议纪要、论证意见并确定具体实施人员, 报科室领导审核。

(4)协议签订。服务内容、费用等与申请方确认无误后, 签订协议, 主要内容包括项目名称、签约双方信息、合作内容、合作目标及分工、保密要求、合作经费和成果分配等, 其中合作经费由科研课题外协服务费以内部转账的形式支付。

(5)组织实施。合作协议签订完毕后, 双方协同组织实施, 在约定时间内, 保质保量地完成协议相关内容。具体实施人员负责对数据分析结果进行技术总结, 数据及代码归档并形成翔实研究

报告。实施期间，由评估技术组定期对项目实施进度、质量进行督导。

(6)结果反馈。项目实施完成后，评估技术组对研究报告进行评估，评估通过后，上报科室领导审核，办理项目交付。

1.2.2 数据服务协议收费制度 数据服务协议收费制度，明确了参与方的责任和权益，有利于保障服务工作持续有质量的开展。收费核算标准按成本测算，主要考虑服务过程中投入的人力成本、设备折旧/维修成本、房屋折旧成本和其他成本等。评估技术组对具体服务内容进行拆解，并按收费项目进行预计费用的合计，双方无异议后签订服务协议。对方在缴纳总预计费用的20%作为押金后即可进行服务项目开展，并按项目进展进度缴纳完成50%、80%、100%费用，项目完成后由大数据中心开具可在院内进行课题报销的服务清单。

医学大数据中心开设专门收入账户，以内部转账形式收取服务课题中外部协作费，与原有支出账户区分，收入所得作为中心运行办公经费。

1.2.3 数据处理分析流程 数据处理分析流程见图4，包含数据提取、数据清洗、描述统计分析、数据预处理、建模分析、效果评估和模型解释等一系列工作流程，涉及技术包括结构化数据的规则提取、非结构化数据的正则匹配技术、自然语言处理技术、机器学习建立预测模型等^[15]。数据处理分析工作比较复杂，需要多学科人员协同合作，工作中经常会根据过程结果修正临床研究方案，重新调整数据提取和分析策略。该项工作占整个研究的很大比重，规范流程可以显著提高临床研究的执行效率，是获得具有科学性和标准性

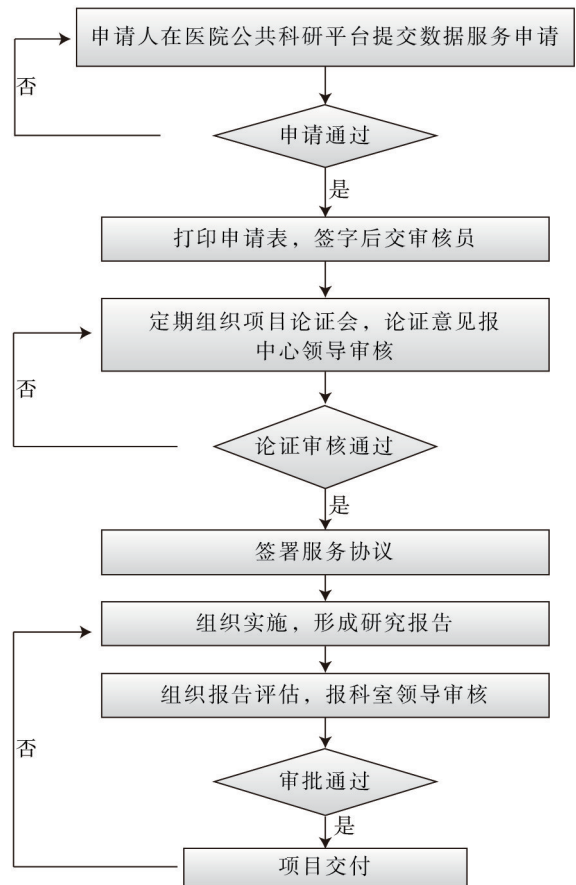


图3 数据服务申请制度

Fig. 3 Application process of data services

研究结论的保证^[11]。见图4。

为方便临床人员参与数据处理分析过程，中心通过云桌面系统(预装常用数据处理分析软件)将所需数据提供到临床科室内网环境，并分配独立访问账号。云桌面系统通过限制用户导入导出操作，实现了原始数据不出科的跨科使用，既保证了数据安全，又减轻了中心工位负担。

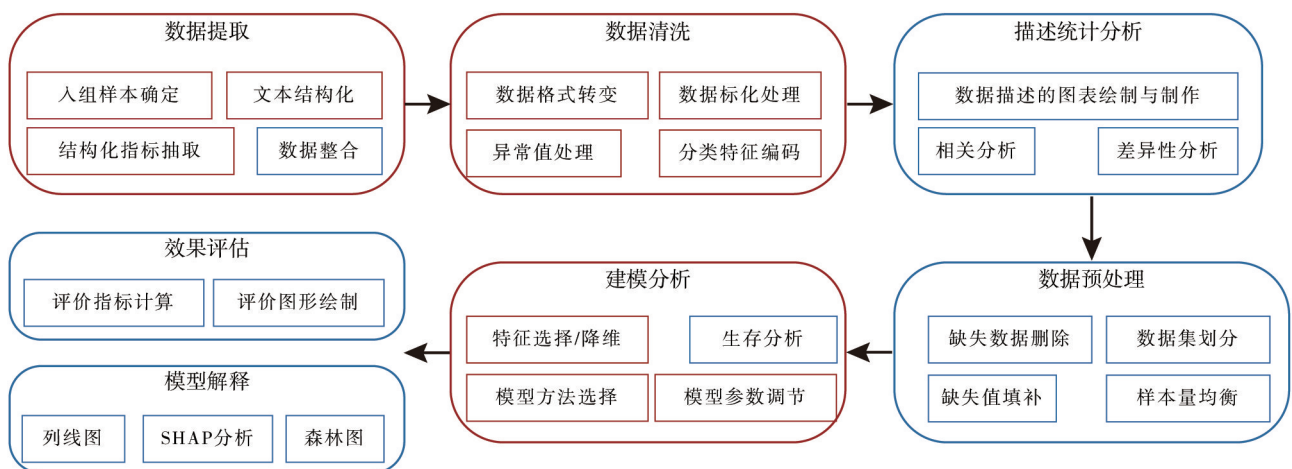


图4 数据处理分析流程(蓝色：可实现自动化操作；红色：需人工与代码结合操作)

Fig. 4 Data processing and analysis process (Blue: automated operation; red: manual operation combined with code)

2 人员管理与数据安全

医院大数据中心作为临床大数据科研的支撑，通过优化体系化管理和注重技术的联合运用，提升自身价值，有效支持全院临床科研的快速发展，实现合作共赢，但在建设过程中需要着重把握好以下几点。

2.1 人员服务水平

人员服务水平的提升是实现高效医疗服务和科研支持的关键，涉及多个维度，包括服务规范、专业能力和服务意识。

服务规范：中心人员应严格遵守既定的科室规章制度和服务流程规范。要求在日常工作中，确保服务过程的可追踪性、分析结果的可复现性以及数据资产的安全性。通过制定和执行严格的服务标准，可以提高服务的质量和可靠性。

服务专业：中心人员需具备扎实的大数据技术知识，包括数据处理、统计分析等能力。此外，还应涉猎流行病学等相关知识，掌握临床方案设计的方法。只有不断学习和实践，才能深化对临床科研的理解，从而提供更加专业和高效的服务^[16]。

服务意识：服务意识是指在日常工作中所体现的主动性、热情度和周到性。中心人员应认识到服务的重要性，以及优质服务对中心和个人职业发展所带来的积极影响。通过提高服务意识，可以激发中心人员的积极性和创造性，不断提升服务水平。

2.2 数据安全

数据安全是一个长期被讨论的重点议题，尤其在处理高度敏感的医疗数据时，其重要性更是不容忽视。数据安全与数据治理常被视为医疗数据资源有效利用的两个基本要素。随着大数据产业的兴起，医疗数据的价值日益凸显，其潜在的经济利益和科研价值引起了广泛关注。临床工作者和医疗技术人员必须认识到，数据资产是医疗机构资产的重要组成部分，任何未经授权的泄露或交易行为都是严格禁止的。

3 医院大数据中心数据服务成效

总医院大数据中心有力支撑了国家发改委、国家科技部、国家药监局、军委科技委等管辖的国家工程研究中心、医学研究中心和重点实验室的申报和建设，有效发挥了总医院科技创新公共平台的支撑作用。2018年1月—2023年12月，共

承接临床科研数据服务290项，项目类型统计结果见图5。结果显示，排名前3的项目类型分别为数据支撑、疾病风险预测、专病库建设，其比例依次为30%(86)、16%(46)和13%(37)。见图5。

科研服务产出相关论文70篇，其中中文核心期刊42篇，SCI 28篇，支持临床培养研究生近百名。论文发表初始以中文核心期刊为主，后SCI论文逐渐增多，并超过中文核心论文数量，见图6。

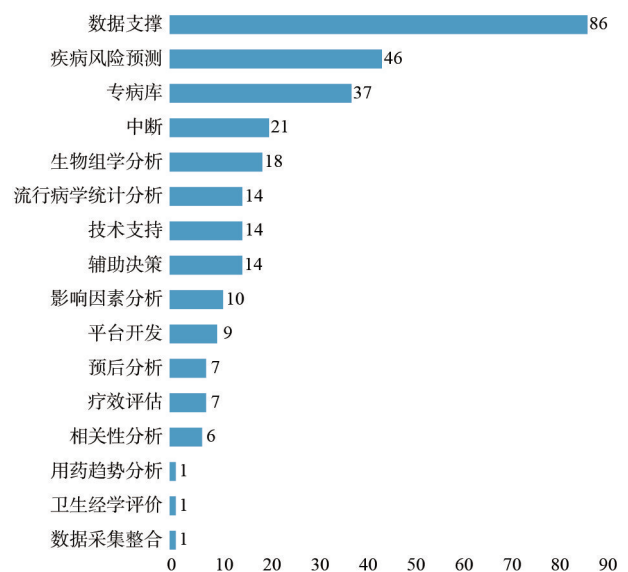


图5 数据服务项目类型统计

Fig. 5 Statistical chart of types of data service items

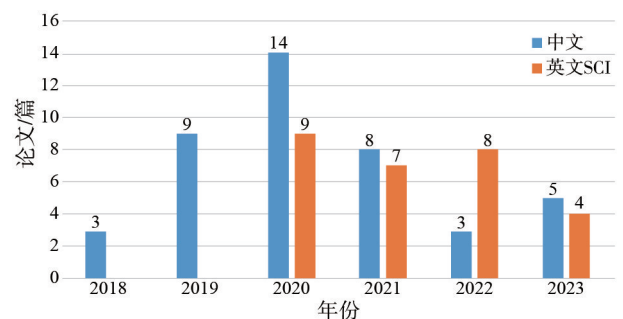


图6 科研数据服务相关论文产出情况

Fig. 6 The output of papers in scientific research data service work

通过医疗数据的整理和分析，发现新现象，总结新知识点，仅仅是医疗数据利用的起点。医疗大数据的真正价值需要将碎片化知识融合，并最终应用于实践，实现从数据到知识，知识到实践，实践到数据的良性循环发展^[17]。医院大数据中心以信息化建设形成的医院“数据湖”为基础，通过大数据技术实现AI辅助诊断、疾病预测、治疗评估、健康管理等价值转化，达到医疗活动成本降低、效果增强的目的，为医疗相关产业链带来新变化的产品及服务。

4 讨论

医院大数据中心运行一段时间后,易出现服务工作疲于应付导致项目进展拖延、报告结论不可靠以及数据安全得不到保障等问题。所以,中心要注意服务水平的持续提升和数据安全防护建设。

4.1 提升服务水平

(1)持续教育与培训:定期为人员提供专业培训,以更新知识和技能,适应医疗大数据领域的快速发展。(2)服务流程优化:鼓励中心人员总结经验,提炼服务方法,不断优化服务流程,提高服务效率。(3)跨学科学习:鼓励人员学习跨学科知识,如流行病学、临床研究方法等,以拓宽服务视野,增强服务能力。(4)质量控制与评估:建立服务质量控制和评估体系,定期对服务过程和结果进行评估,确保服务质量的持续改进。(5)激励机制:建立有效的激励机制,鼓励中心人员提高服务意识,为提供优质服务创造良好的工作环境。

4.2 提高数据安全防护

(1)制定规章制度:包括但不限于工作人员行为准则、外部合作人员管理规定、数据安全管理条例以及数据服务管理条例等,确保全体工作人员进行系统的数据安全教育。(2)提升安全意识:避免因追求便利而将医疗数据下载至个人设备或非法传输给他人使用,同时积极思考管理过程中存在的问题并提出改进意见。(3)加强数据库的管理和网络信息的安全隔离:通过权限配置、角色管理、运维监控和网络隔离等技术措施,防止数据的未授权访问和泄露,并确保数据在存储、处理和传输过程中的完整性和可用性。

数据安全防护是一个多维度、多层次的系统工程,需要管理员、安全员和其他人员共同努力,才能实现医疗数据资源的高效、安全利用。

5 结语

医院大数据中心的建设是顺应时代发展的必然,它带来了医院管理及临床科研模式的变革。

大数据中心的建设有助于医学科研数据的开放获取和共享利用,减少医院科研的重复投资和建设时间,在保证数据资产安全的前提下,更好地服务患者诊疗、临床科研和医院管理工作,促进医院高质量发展。

作者贡献 车贺宾:论文撰写;徐洪丽、吴欢:审读和修订;乌日力格、胡葶:数据收集,绘制表格;何昆仑、郭华源:监督指导。

利益冲突 所有作者声明无利益冲突。

数据共享声明 本论文相关数据可依据合理理由从作者处获取,Email: chehebin@163.com。

参考文献

- 1 王映佳.大数据时代背景下医院数据中心建设的相关思考[J].电脑知识与技术,2020,16(3):9-10.
- 2 国家医疗数据中心在北大医学部成立[J].中国数字医学,2015,10(6):113.
- 3 赣南医学院第一附属医院.赣南医学院第一附属医院成立医药大数据中心[J].肿瘤防治研究,2019,46(4):388.
- 4 方安,胡佳慧,钱庆,等.大数据环境下的医学科研服务架构[J].中华医学图书情报杂志,2019,28(1):8-12.
- 5 李尚.基于大数据的医院数据中心建设的思考[J].现代信息科技,2019,3(23):104-105.
- 6 隽军利,王小龙.解析数据中心数据库系统安全架构研究与应用[J].信息系统工程,2019(6):79.
- 7 彭晓辉.终端安全防护技术在网络管理中的应用研究[J].电子技术与软件工程,2014(5):220.
- 8 王少伟,孙咸江.医院数据中心建设研究[J].电子世界,2020(21):27-28.
- 9 韩煜.医院临床数据中心构建的思路分析[J].医学信息,2020,33(17):18-19.
- 10 李晓华.大数据视角下的医学影像技术的发展与应用探究[J].科学技术创新,2020(3):150-151.
- 11 席韩旭,张晨,张欣,等.基于临床大数据的科研平台建设与应用探讨[J].医院管理论坛,2020,37(9):67-68.
- 12 郑西川,胡燕峰,冯贞贞.结构化描述语言电子病历模型及其关键技术实现[J].中国数字医学,2011,6(5):40-43.
- 13 赵前前.基于大数据科研平台的专病数据库建设及应用[J].中国数字医学,2020,15(12):89-92.
- 14 薛万国,应俊.大数据时代的医学创新与现实挑战[J].解放军医学院学报,2019,40(8):705-708.
- 15 莫祖英.大数据处理流程中的数据质量影响分析[J].现代情报,2017,37(3):69-72.
- 16 胡德斌.提高数据中心服务能力,创造数据中心服务价值[J].中国金融电脑,2011(4):32-34.
- 17 柴扬帆,孔桂兰,张路霞.医疗大数据在学习型健康医疗系统中的应用[J].大数据,2020,6(5):29-44.

(责任编辑:潘越,施晓亚)