

基于流程视角的健康医疗大数据处理关键技术探讨*

高景宏¹ 任明星² 宋晓琴³ 马倩倩⁴ 刘福荣⁵ 王成增¹

(¹ 郑州大学第一附属医院河南省医院管理研究院 郑州 450000

² 郑州大学第一附属医院科研处 郑州 450052 ³ 郑州大学第一附属医院体检中心 郑州 450052

⁴ 郑州大学第一附属医院互联网医疗系统与应用国家工程实验室 郑州 450052

⁵ 北京大学肿瘤医院病案统计室 北京 100142)

〔摘要〕 目的/意义 深入探讨健康医疗大数据的处理环节及关键技术,为进一步推动健康医疗大数据的应用提供信息支撑与保障。方法/过程 参照大数据驱动的科学研究的范式,基于流程视角,在介绍健康医疗大数据来源与特点的基础上,探析大数据处理环节及关键技术,并论述目前面临的挑战和未来发展方向。结果/结论 国内外健康医疗大数据处理技术已取得一定创新发展,但在多维度数据采集、多模态数据融合、大模型数据分析与数据安全等方面仍面临挑战。未来应构建一体化大数据处理综合平台,研发或集成有针对性的功能模块,以应对相关问题。

〔关键词〕 健康医疗;大数据;流程视角;处理技术

〔中图分类号〕 R-058 〔文献标识码〕 A 〔DOI〕 10.3969/j.issn.1673-6036.2024.11.012

Discussion on Key Technologies of Health Care Big Data Processing Based on the Process Perspective

GAO Jinghong¹, REN Mingxing², SONG Xiaoqin³, MA Qianqian⁴, LIU Furong⁵, WANG Chengzeng¹

¹Institute for Hospital Management of Henan Province, the First Affiliated Hospital of Zhengzhou University, Zhengzhou 450000, China;

²Office of Academic Research, the First Affiliated Hospital of Zhengzhou University, Zhengzhou 450052, China; ³Physical Examination Center, the First Affiliated Hospital of Zhengzhou University, Zhengzhou 450052, China; ⁴National Engineering Laboratory for Internet Medical Systems and Applications, the First Affiliated Hospital of Zhengzhou University, Zhengzhou 450052, China; ⁵Medical Record Statistics Room, Beijing Cancer Hospital, Beijing 100142, China

〔Abstract〕 **Purpose/Significance** The processing links and key technologies of health care big data are deeply discussed to provide information support and guarantee for further promoting the application of health care big data. **Method/Process** According to the scientific research paradigm driven by big data, and based on the process perspective, the paper introduces the sources and characteristics of health care big data, analyzes the processing links and key technologies of big data, and discusses the current challenges and future development directions of health care big data processing. **Result/Conclusion** Both China and foreign countries have achieved innovative development in technologies for health care big data processing. However, there are still some challenges in multi-dimensional data collection, multi-mo-

〔修回日期〕 2024-07-11

〔作者简介〕 高景宏,博士,副研究员,发表论文60余篇;通信作者:王成增。

〔基金项目〕 国家自然科学基金面上项目(项目编号:72374183);河南省重点研发与推广专项(科技攻关)(项目编号:232102310062);河南省医学科技攻关计划软科学项目(项目编号:RKX202202022)。

dal data integration, large model data analysis, and data security. In the future, through building a comprehensive platform of big data processing, and developing or integrating targeted functional modules, the existing problems can be effectively addressed.

[Keywords] health care; big data; process perspective; processing technologies

1 引言

近年来,健康医疗大数据因其增长快速、应用广泛及价值潜力巨大,已成为国内外研究的热点^[1-2]。健康医疗大数据广泛应用于疾病预测、诊断辅助、药物研发等领域^[3-4],正在重新界定医学研究与实践的方法和流程。在大数据驱动的科研背景下,全链条、需求导向的数据处理流程尤为重要。然而,目前尚无研究从流程视角系统分析健康医疗大数据的处理环节及关键技术^[4-5]。鉴于此,本研究拟在概述健康医疗大数据来源和特点的基础上,参照大数据驱动的研究工作流程,深入探讨健康医疗大数据处理的关键环节与技术前沿,同时剖

析当前面临的挑战,旨在为进一步推动健康医疗大数据的应用提供信息支撑。

2 健康医疗大数据的来源与特点

健康医疗大数据贯穿人的全生命周期,涉及医疗服务、公共卫生、营养保健、健康管理、保险、用药、生物组学等多源信息,见表 1。除了数据容量大(volume)、数据种类和来源多样化(variety)、增长更新速度快(velocity)、数据蕴含的价值大但密度低(value)、对数据真实性要求高(veracity)等大数据“5V”特征外^[6],健康医疗大数据还具有初始性、时序性、不完整性、隐私性、冗余性等特点^[7-8]。

表 1 健康医疗大数据的主要来源

来源	内容
医疗机构诊疗数据	医院信息系统、检验信息系统、影像存储与传输系统、电子病历、心电、病理数据以及产生于医院日常诊疗、科研和运维过程的各种门/急诊信息、住院记录、护理日志、医疗费用、用药、手术、随访和医保等数据
生物信息数据	生物样本,蛋白质组学、代谢组学、基因组学等组学数据
区域卫生信息平台数据	居民健康档案个人体征监测、生活习惯与喜好、膳食营养状况、健康知识等数据
公共卫生数据	公共卫生专题调查,专病监测,膳食调查,疾病防控、妇幼保健、职业病防护等产生的数据,气象、空气污染等环境监测数据
互联网数据	有关健康、疾病或寻医的互联网搜索、话题、浏览、分享数据,购药、健康网站访问等行为数据
行业数据	医疗保险,商业健康保险,新药研发与临床试验,医药、医疗器械和耗材销售等数据
自我健康管理数据	移动或物联网连接的可穿戴设备产生的数据,健康医疗设备通过“云+端”方式收集的各种生命体征信息
监管和服务数据	政府、主管部门和行业协会的监管、服务相关数据

3 大数据驱动的科学研究的范式

大数据正在推动自然科学研究从经验研究、理论研究、计算机仿真研究向大数据科学研究的第四研究范式跃迁^[9]。在大数据驱动的研究范式中,首先按照研究需求纳入全样本数据,然后在

标准化数据处理基础上,通过机器学习算法进行集成学习、关联规则分析、特征提取及模型拟合等操作,剖析研究对象与其影响因素之间的复杂关系^[10]。通过将此复杂关系表达为具有较强因果解释能力的整体理论模型或细粒度复杂模型,可减少研究者对既有理论的依赖,更好地拟合、解释真实世界^[9-10]。

4 健康医疗大数据的处理流程

参照大数据驱动的研究范式，从流程视角出发，将健康医疗大数据的处理流程划分为数据采集、清洗、存储、分析、结果解释与可视化等环节，见图 1。为确保数据被有效分析与应用，健康医疗大数据处理需依托稳定可靠的平台，并贯穿始终地实施质量控制与安全管理^[11]。标准化数据处理可有效支撑决策支持、预警预测、医疗人工智能等应用。

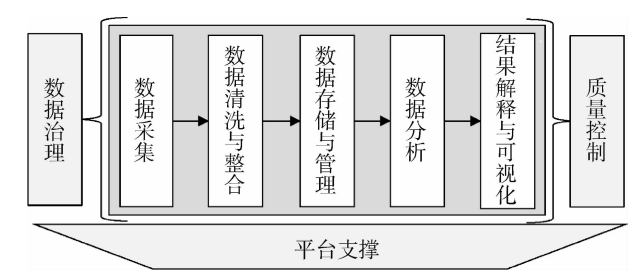


图 1 健康医疗大数据处理流程

5 健康医疗大数据处理环节及关键技术

5.1 数据采集

数据采集通过多种方式从数据生产环境获取数量庞大、类型众多的原始数据并进行预处理，为后续数据处理提供满足需求的数据集^[11]。传统数据采集方法包括射频识别、传感器、日志文件、数据库同步与抓取、条形码、Web 爬虫、社交网络交互、移动互联网以及信息检索分类工具（如百度、谷歌）等^[12]。近年来，Flume、Kafka、Sqoop 等大数据采集技术在健康医疗领域广泛应用。其中，Flume 作为实时数据采集的开源框架，能够有效汇集各种网站服务器上的日志、时间等信息，并存储到本地集中存储器中，为后续分析提供基础数据。Kafka 是基于分布式架构的开源消息发布-订阅系统，以其高吞吐率、速度快、低延迟等特点，成为实时平台中解决数据采集速度与处理速度不同步的关键中间件^[13]。Sqoop 作为开源的离线数据传输工具，在 Hadoop 与传统数据库之间的数据传递中发

挥重要作用，通过从结构化存储器抽取数据并加载到 Hive、Hbase 和 HDFS 等大数据系统中，可显著提高数据采集的并发性和容错性^[12-13]。

鉴于健康医疗大数据的复杂性和多样性，能够自动识别、汇集和整合多源异构数据的采集工具和技术日益受到关注。例如，抽取、转换和加载（extraction - transformation - loading，ETL）工具通过抽取、转换和清洗，再将数据加载到目标数据库，为后续数据分析和应用提供支撑^[13]。Robotic Process Automation 通过模拟人类在计算机上执行的操作，可从多个数据源自动化收集、整合和提取数据。自然语言处理技术通过解析多源文本数据，帮助计算机识别、理解 and 处理人类语言的结构和含义，实现文本信息的高效采集。深度学习模型可自动学习和提取复杂海量数据中的特征，实现多源数据格式和结构的自动识别与解析^[10,12-13]。

5.2 数据清洗与整合

数据清洗是面向原始数据中“脏数据”的源头，通过针对性技术手段，将“脏数据”修复为满足要求的高质量数据。数据清洗一般采用数理统计、数据挖掘及预定义清洗规则的方法，通过分析产生原因、评估影响、考察分布情况、提取规则，对异常、缺失、重复、不一致等“脏数据”进行修正、转换或删除，达到提高数据质量的目的^[14]。常用的数据清洗技术包括 K - 均值聚类算法、Canopy 算法、K 近邻算法、邻近排序算法、贝叶斯分类算法和神经网络算法等^[11,14]。

数据整合是大数据分析与应用的基础保障。美国健康信息交换和欧洲泛欧健康数据空间等项目均致力于实现跨机构、跨地域的数据清洗与整合，旨在促进数据的共享和分析^[15-16]。数据整合需要对存储在文本文件、数据库中的数据进行清洗和转换，以消除相似、重复数据，并抽取元数据信息，实现数据的统一融合。建立综合服务云平台，在数据实时性和可靠性要求下，通过基于分布式架构的高速 ETL 技术，对健康医疗大数据进行整合，已成为行业趋势^[15]。例如，数据湖和数据仓库技术能够将多源异构数据抽取到临时中间层，进行清洗、转

换、标准化和整合等，然后按需调取至数据集市供后续处理，为大数据的联机分析提供支撑^[16]。另外，可利用云计算的弹性和可扩展性实现大数据的高效整合，数据联邦技术可通过在不同数据源之间建立虚拟连接实现关系数据库、NoSQL 数据库、文件系统、云服务等不同格式、不同结构和不同位置数据的无缝整合^[17-18]。

5.3 数据存储和管理

大数据存储和管理指通过专用设备对收集的数据进行存储，建立数据库并进行管理和调用。传统的数据存储主要通过 MySQL、Oracle 等关系型数据库实现，具有快速存储结构化数据、支持随机访问等优点^[19-20]，但已逐渐无法满足健康医疗大数据处理的需求。为了解决海量健康医疗数据的存储问题，弹性可扩展、高容错、高吞吐量且成本低的分布式存储系统成为首选^[2,19]。目前代表性的分布式架构大数据存储技术有 HDFS、Kosmos Distributed File System、Google File System（GFS）等，支持各类型数据存储，并可通过增加硬件设备、冗余存储等保证存储系统的扩展性和可靠性^[20-21]。

数据管理方面，现有单表存储结构已很难适应健康医疗大数据对海量存储、高并发读写、数据组织及去冗余等方面的要求^[22]。目前常见的大数据管

理技术是 Hadoop 基于 HDFS 的 HBase 和谷歌公司基于 GFS 的 Big Table。二者均既可为健康医疗大数据应用提供类似传统数据库的简单数据组织、查询、融合及结构化存储等功能，也可对各种大数据并行处理提供数据源及数据分析结果的存储^[11,19,21]。近年来，云数据管理日益流行，其允许用户根据需要调整数据存储和处理能力，并可随时随地使用任何设备访问，兼具可扩展性、灵活性和成本效益^[19,21,23]。

5.4 数据分析

数据分析用于发现数据的价值，是健康医疗大数据处理的核心，常用技术包括数理统计、数据挖掘、机器学习、自然语言处理、文本分析与图像、语音识别等，见表 2。机器学习在大数据分析中的应用日益广泛。国外 TensorFlow、PyTorch 等开源框架有效推动了基于大数据分析的决策和预测模型研究^[14,19]。国内在积极采用国际开源大数据分析技术的同时，也贡献了自己的开源项目，如百度的 PaddlePaddle 深度学习平台。近年来，还出现了联邦学习、大模型系统、边缘计算和增强分析等新兴技术，通过增强算力、节约成本、保护隐私、增加精度和效率、设计通用大模型等，为健康医疗大数据分析与应用提供了新的路径与手段，典型应用如医渡云、扁鹊健康大模型、全诊通等^[24]。

表 2 常见大数据分析方法

数据分析方法	方法描述
数理统计	基于统计学原理，对数据进行收集、整合、分析和结果解释，主要用于分析变量间可能出现的定性或定量关系。常见方法包括统计检验、相关性分析、判别分析、时间序列分析、聚类分析、回归分析等
数据挖掘	从大量、随机、有噪声、不完全数据中提取过去未知、有潜在价值信息的过程，是统计学、数据库管理和人工智能的综合运用。通过归纳演绎或聚类分析，构建依赖、分类或预测模型，发现相关关系、异常或潜在趋势等。常用工具有 IBM SPSS、Python、Oracle Darwin、SGI MineSet 和 Weka 等
机器学习	分为有监督学习、无监督学习、半监督学习、强化学习和深度学习等。在大数据分析中主要用于搜索、迭代优化和图计算。常用技术包括贝叶斯网络、支持向量机、决策树、神经网络、聚类、序列分析、回归拟合、迁移学习、隐马尔可夫模型和概率图模型等
自然语言处理	基于语言学和计算机科学，利用计算机算法对人类自然语言进行识别和分析。主要用于语义、情感和舆情分析以及机器翻译、信息检索和过滤等。关键技术涉及语音识别、文本生成、词法和句法分析、语义分析等
文本分析	从原始文本中挖掘隐藏有用信息或知识的过程，涉及信息检索、数理统计和计算机语言等
图像、语音识别	利用计算机和现代通信技术，实现对图片、音频等的编解码和匹配，从而识别或分类多源异构信息。图像识别主要涉及机器视觉和人工智能技术，语音识别则主要用到数字信号处理、人工智能和数理统计等

健康医疗大数据分析对运算架构、时效性、运算性能和计算域存储单元的数据吞吐率等要求较高,传统的数据分析工具已难以胜任^[10]。因此,能提高数据处理速度,使系统具备高扩展性、时效性和可用性等优势的分布式并行数据处理技术趋于流行。目前,分布式高效并行处理健康医疗大数据的模式有 3 种:离线批处理计算、流式实时处理计算和内存计算。谷歌公司提出的 MapReduce 是离线批处理模式的典型,其对数据的一致性无特殊要求,兼具高扩展性和高可用性,适用于大规模多源异构数据的混合处理^[6,11],但其时延较大,不适合实时在线流式处理、机器迭代学习。对此,有公司分别研发针对实时数据处理的架构 s4 系统和 Storm 系统^[11,20],能提供高性能的实时分析、全天候计算、在线机器学习等。Spark 分布式内存计算系统启用了内存分布数据集,具有快速反应、灵活迭代运算能力,能够满足在小数据集上处理复杂数据的计算需求^[8]。

基于平台化技术支撑的大数据综合处理与分析已成为趋势。在国外,Amazon Web Services、Microsoft Azure、Google Cloud Platform 等云计算平台均提供强大的大数据分析服务和工具^[5,19,25]。我国也在积极推进自主研发和构建健康医疗大数据平台,如华为的 FusionInsight、阿里的 MaxCompute、腾讯的 TencentDB for TBase 等平台,旨在整合各类健康医疗数据资源,推动数据的标准化、互联互通以及深度分析应用。在中医药领域尤为特色,例如,通过构建中医药一体化大数据平台,不仅能够深度挖掘中医药大数据的潜在价值,还能建立中医药大数据知识图谱,梳理患者整体画像与诊断结果之间的关系,可实现简化患者就医流程、基于现代科学诠释中医药理论、提升中医诊断体系标准化和诊断水平、形成个性化中医健康管理方案等目标^[26-27]。

5.5 可视化

可视化指将抽象信息以贴近人类自然感知的图形形式予以呈现,属于人机交互、统计学、计算机图形学、信息科学等有机结合的交叉学科^[11]。数据

处理可视化主要通过视觉和图形化的方式,溯源、演绎数据分析过程,基于数据交互的可视化表达,发现数据中隐含的信息,以形象直接的形式将结果展示给用户,提高分析结果的可解释和接受程度^[4]。常见的可视化技术有历史流、标签云、空间信息流等,涉及基于图像、集合、图标、分布式、面向像素等的可视化技术^[28]。近年来,可视化技术快速发展,Tableau、Power BI、D3.js 等工具使复杂的数据分析结果以直观的方式呈现^[14,25]。另外,人工智能驱动的可视化技术(如 ChartGPT),以及交互式与沉浸式相结合的可视化技术(如增强现实与虚拟现实)等,因其能使用户更直观地探索和理解数据的优点,得到快速发展与应用^[28-29]。

6 面临挑战

目前,健康医疗大数据的处理依然面临诸多挑战^[4,6,8,11,13,14,20]。在数据采集环节,各信息系统和数据采集工具的接口标准、传输协议存在不同程度的差异,无缝对接和采集多模态数据的接口或工具需要针对性定制开发或集成,导致数据采集在人力、物力、时间方面的投入过高,且进一步处理效率与质量并不理想。数据清洗方面,面临数据缺失、异常、重复、不一致等问题,每种问题均需一种或多种不同方法予以处理,这既增加了工作量,也对技术创新提出了更高的要求。数据融合中,需平衡保留源数据信息完整性与有效变量特征识别提取之间的关系,避免信息无法满足分析需求或算力浪费等情况。数据存储和管理方面,多源异构健康医疗数据的汇总存储和管理使不同数据指标间的关联更加紧密,但也可能增加敏感信息泄漏的风险。因此,在确保高效数据存取和软硬件资源有效利用的同时,要重视数据安全和隐私保护。数据分析方面,鉴于健康医疗大数据的“5V”特性,目前基于单一或有限维度与类型数据的分析,以及基于未经多模型交叉验证的单一算法分析结果,难以较好地反映和解释真实世界的研究命题,限制了数据分析的深度和广度。

7 未来方向

鉴于健康医疗大数据处理面临的问题,未来可在以下几方面谋篇布局^[14,23,25,30-32]。一是基于大数据驱动的研究工作流程,将数据处理各环节视作单独研究命题,分别形成系统的实施方法论和技术体系,针对性解决现有挑战。二是根据具体数据分析与应用需求,定制化开发接口并统一信息传输协议,保障多源异构数据采集的有效性和可行性。三是通过区块链、边缘云、联邦学习和隐私计算等技术,规范数据处理流程,明确数据的使用条件、适用范围和权责关系,进一步夯实健康医疗数据的安全与隐私保护。四是采用多种计算机技术和机器学习算法,对健康医疗大数据进行多维度、多层次分析并交叉验证模型拟合结果,以充分挖掘其蕴含价值,更好地辅助决策。五是打造一体化健康医疗大数据处理综合平台,通过现代数据栈技术,基于流程视角,构建涵盖数据处理各环节的功能模块或集成工具,突破“数据烟囱”问题,关联和打通数据处理路径,节省算力消耗和数据分析时间,进一步提高数据处理的质量和效率。

8 结语

健康医疗大数据处理可辅助临床诊疗、慢病管理、药物研发、卫生政策制定等。同时,基于大数据处理分析,能够有效反映医疗卫生机构的整体运营管理情况,有利于决策者提高管理效率,控制不合理因素,作出正确决策,优化医疗卫生资源配置,最终实现大数据驱动的疾病“防诊治”、辅助决策支持、科研助力和管理支撑等业务应用,具有可观的社会效益和经济效益。本研究在概述健康医疗大数据来源与特点的基础上,参照大数据驱动的科学范式,基于流程视角,对健康医疗大数据的处理环节及其关键技术进行解析,重点探讨大数据处理前沿技术,并对现有问题及未来发展方向进行了讨论,可为健康医疗大数据的进一步规范处理与应用提供重要的信息支撑。

作者贡献: 高景宏、王成增负责研究设计、论文撰写与修订、监督管理;任明星、宋晓琴负责资料整理;马倩倩、刘福荣负责文献检索与整理。

利益声明: 所有作者均声明不存在利益冲突。

参考文献

- 1 LYNCH C. Big data: how do your data grow [J]. *Nature*, 2008, 455 (7209): 28-29.
- 2 JIANG P, SINHA S, ALDAPE K, et al. Big data in basic and translational cancer research [J]. *Nature reviews cancer*, 2022, 22 (11): 625-639.
- 3 CHOWKWANYUN M. Big data, large-scale text analysis, and public health research [J]. *American journal of public health*, 2019, 109 (S2): S126-S127.
- 4 石乐明,郑媛婷,苏振强,等. 大数据与精准医学 [M]. 上海:上海交通大学出版社,2017.
- 5 KAR A, SUBASH A, RAO V U S. Reactive artificial intelligence using big data in the era of precision medicine [J]. *JAMA surgery*, 2020, 155 (7): 671.
- 6 高景宏,李明原,王琳,等. 健康医疗大数据在精准医疗领域的应用与挑战 [J]. *医学信息学杂志*, 2022, 43 (5): 15-20.
- 7 傅琪,毛琛. 健康医疗大数据驱动下的流行病学研究: 机遇与挑战 [J]. *中华疾病控制杂志*, 2023, 27 (2): 125-126, 237.
- 8 WESSON P, HSWEN Y, VALDES G, et al. Risks and opportunities to ensure equity in the application of big data research in public health [J]. *Annual review of public health*, 2022, 43: 59-78.
- 9 米加宁,章昌平,李大宇,等. 第四研究范式: 大数据驱动的社会科学研究转型 [J]. *学海*, 2018 (2): 11-27.
- 10 郁建兴,高翔,王诗宗,等. 数字时代的公共管理研究范式革命 [J]. *管理世界*, 2023, 39 (1): 104-116.
- 11 翟运开,高景宏,何贤英,等. 面向精准医疗的大数据处理技术研究 [M]. 北京:科学出版社,2021.
- 12 MAZHAR S A, ANJUM R, ANWAR A I, et al. Methods of data collection: a fundamental tool of research [J]. *Journal of integrated community health*, 2021, 10 (1): 6-10.
- 13 TAHERDOOST H. Data collection methods and tools for research; a step-by-step guide to choose data collection technique for academic and business research projects [J]. *International journal of academic research in management*,

- 2021, 10 (1): 10 – 38.
- 14 CÔTÉ P O, NIKANJAM A, AHMED N, et al. Data cleaning and machine learning: a systematic literature review [J]. Automated software engineering, 2024, 31 (2): 54.
- 15 FILLINGER S, DE LA GARZA L, PELTZER A, et al. Challenges of big data integration in the life sciences [J]. Analytical and bioanalytical chemistry, 2019, 411 (26): 6791 – 6800.
- 16 ABRAHAM R, SCHNEIDER J, VOM BROCKE J. Data governance: a conceptual framework, structured review, and research agenda [J]. International journal of information management, 2019, 49 (12): 424 – 438.
- 17 GU Z, CORCOGLIONITI F, LANTI D, et al. A systematic overview of data federation systems [J]. Semantic web, 2022 (12): 107 – 165.
- 18 MARINESCU D C. Cloud computing: theory and practice [M]. San Francisco: Morgan Kaufmann, 2022.
- 19 PRAMANIK P K D, PAL S, MUKHOPADHYAY M. Research anthology on big data analytics, architectures, and applications, edited by information resources management association [M]. Hershey, Pennsylvania: IGI Global, 2022.
- 20 DASH S, SHAKYAWAR S K, SHARMA M, et al. Big data in healthcare: management, analysis and future prospects [J]. Journal of big data, 2019, 6 (1): 1 – 25.
- 21 张习梅, 杨露, 南原. 疾病预警在健康大数据管理平台中的应用 [J]. 医学信息学杂志, 2021, 42 (2): 49 – 52, 65.
- 22 WANG X, WILLIAMS C, LIU Z H, et al. Big data management challenges in health research: a literature review [J]. Briefings in bioinformatics, 2019, 20 (1): 156 – 167.
- 23 BABAR M, AHMAD JAN M, HE X, et al. An optimized IoT – enabled big data analytics architecture for edge – cloud computing [J]. IEEE internet of things journal, 2023, 10 (5): 3995 – 4005.
- 24 CHEN J, RAN X. Deep learning with edge computing: a review [J]. Proceedings of the IEEE, 2019, 107 (8): 1655 – 1674.
- 25 IQBAL R, DOCTOR F, MORE B, et al. Big data analytics: computational intelligence techniques and application areas [J]. Technological forecasting and social change, 2020, 153 (4): 119253.
- 26 姜义琳, 张红, 马兆辉, 等. 智慧中医药服务新模式的实践与思考 [J]. 中国卫生信息管理杂志, 2023, 20 (3): 358 – 363.
- 27 周洁, 叶洪柳, 汤娟娟. 大数据视域下中医药技术化相关管理的困境与优化策略 [J]. 中医药管理杂志, 2024, 32 (2): 184 – 186.
- 28 WASKOM M L. Seaborn: statistical data visualization [J]. Journal of open source software, 2021, 6 (60): 3021.
- 29 QIN X, LUO Y, TANG N, et al. Making data visualization more efficient and effective: a survey [J]. The VLDB journal, 2020, 29 (1): 93 – 117.
- 30 WU X, LI W, TU H. Big data and artificial intelligence in cancer research [J]. Trends cancer, 2024, 10 (2): 147 – 160.
- 31 PEREZ L G, PEET E D, VEGETABLE B, et al. Big data needs and challenges to advance research on racial and ethnic inequities in maternal and child health [J]. Womens health issues, 2022, 32 (2): 90 – 94.
- 32 XIA J, WANG J, NIU S. Research challenges and opportunities for using big data in global change biology [J]. Global change biology, 2020, 26 (11): 6040 – 6061.

《医学信息学杂志》开通微信公众号

《医学信息学杂志》微信公众号现已开通, 作者可通过该平台查阅稿件状态; 读者可浏览当期最新内容、过刊等; 同时提供国内外最新医学信息研究动态、发展前沿等, 搭建编者、作者、读者之间沟通、交流的平台。可在微信中搜索“医学信息学杂志”添加, 也可扫描右侧二维码添加, 敬请关注!



《医学信息学杂志》编辑部