

• 专论:医学科学数据管理 •

编者按: 医学科学数据广泛涉及人类生命健康相关的各个领域, 类型复杂多样。作为国家重要的基础性战略资源和实施健康中国战略的重要前提, 在医药卫生科技创新、公共卫生事件应对、医疗卫生服务改革、居民健康服务等方面发挥着关键作用。随着未来医学模式向精准医学、智能医学、生物医学工程等以生物医学数据和计算方法为主的新型医学范式转变, 大规模、高质量和开放共享的医学科学数据建设与管理的重要性日益凸显。医学科学数据管理是一项复杂的系统工程, 需要政府、医疗机构、科研单位、企业等多方主体的协同努力。当前我国在这一领域已取得一定进展, 但仍面临着法律法规、标准规范、组织管理、伦理隐私和技术实现等方面的挑战。本期专论着眼于医学科学数据管理, 内容包括 AI 就绪数据治理、医学科学数据质量评价、医学高校科学数据管理服务等, 以期医学科学数据高质量建设和高效利用提供参考。

美国国立卫生研究院 AI 就绪数据治理实践与启示

王茜 刘湘闽 陈凌云 方安

(中国医学科学院/北京协和医学院医学信息研究所 北京 100020)

[摘要] **目的/意义** 探索医学领域人工智能 (artificial intelligence, AI) 就绪数据治理路径, 为我国 AI 就绪数据建设提供经验借鉴。**方法/过程** 运用案例分析、内容分析等方法, 基于美国国立卫生研究院战略规划、研究专项及科研项目等信息, 从实践举措、推进机制、实施路径等维度分析其 AI 就绪数据治理经验。**结果/结论** 医学领域 AI 就绪数据治理是涵盖标准、工具、伦理和能力的系统工程, 应整合生物医学研究、AI 研发、应用场景等多元要素协同推进。

[关键词] 医学; AI 就绪数据; 治理路径; 美国国立卫生研究院

[中图分类号] R-058 **[文献标识码]** A **[DOI]** 10.3969/j.issn.1673-6036.2025.07.001

Practice and Implication of AI-ready Data Governance of the U.S. National Institutes of Health

WANG Qian, LIU Xiangmin, CHEN Lingyun, FANG An

Institute of Medical Information, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100020, China

[Abstract] **Purpose/Significance** To explore the governance path of AI-ready data in the medical field, and to provide references for the construction of AI-ready data in China. **Method/Process** By using methods such as case analysis and content analysis, based on information such as the strategic planning, research projects and funded scientific projects of the U.S. National Institutes of Health (NIH), the experience of AI-

[修回日期] 2025-04-13

[作者简介] 王茜, 博士, 副研究馆员, 发表论文 20 余篇; 通信作者: 方安, 研究馆员, 硕士生导师。

[基金项目] 中央级公益性科研院所基本科研业务费项目 (项目编号: 2024-ZHCH630-01); 国家重点研发计划项目 (项目编号: 2022YFF0711902-02); 国家科技图书文献中心子任务“医学领域规范数据加工和丰富化” (项目编号: 2024XM0601)。

ready data governance is analyzed from the perspectives of practical measures, promotion mechanisms and implementation paths. **Result/Conclusion** AI-ready data governance in the medical field is a system engineering that encompasses standards, tools, ethics, and capabilities. It is necessary to integrate multiple elements such as biomedical research, AI development, and application scenarios for coordinated advancement.

[**Keywords**] medicine; AI-ready data; governance path; U. S. National Institutes of Health

1 引言

医学以照护人类健康和生命为使命，具有不容错性^[1]，这决定了在医学科研和临床实践中对数据与知识精确性、准确性的要求极高。人工智能（artificial intelligence, AI）作为科技革命与产业变革的重要驱动力，与医学的融合成为国际科技竞争新焦点，在疾病诊疗、康复管理、健康管理及医疗资源配置等方面潜力巨大^[2]。随着以大模型为代表的新一代 AI 技术在医学领域的深入应用，模型“幻觉”、灾难性遗忘等现象凸显。在增加模型参数规模、扩大模型训练数据体量的同时，提高训练数据质量、建设 AI 就绪数据成为业界共识^[3]。就 AI 就绪数据建设而言，当前在高质量数据集生产、多模态数据管理、人员能力协同、管理效能提升等领域面临诸多挑战。在此背景下，本文以美国国立卫生研究院（U. S. National Institutes of Health, NIH）为例，探究 AI 就绪数据治理逻辑和路径，为我国面向 AI 场域的数据建设提供经验借鉴。

2 AI 就绪数据治理相关研究

AI 技术具有数据驱动特征，异构、多源、不完整、有偏差的低质量数据阻碍其深入应用。目前，AI 就绪数据的定义和治理路径尚未形成通用范式，依据用户类型（组织、个人、领域）及 AI 应用场景而演化^[4]。例如，“地球科学信息合作伙伴”^[5]从数据准备、数据质量、数据文档、数据获取 4 个维度界定开放环境数据集的 AI 就绪内涵；《英国国家人工智能战略》^[6]将 AI 就绪数据界定为有助于数据整体状况的数据特征，不限于适用导向、标准化格式、FAIR 状态；麦肯锡^[7]从可感知、可理解、可获取、场景导向、安全 5 个维度定义支撑 AI 应用的

数据标准；德勤^[8]强调建设跨机构流通的高质量、可访问、可信赖、可理解、标注型数据；高德纳^[9]将 AI 就绪数据界定为支持特定用例的高质量数据，涵盖训练或运行 AI 模型所需的模式、类型、边界案例及非预期情况，认为其是以元数据驱动的动态实践，通过持续数据对齐、验证与治理，确保数据满足 AI 应用需求。在 AI 就绪数据治理方面，新加坡人工智能研发计划 AI Singapore^[10]强调采用既定数据策略、流程和实践，确保数据准确性、可靠性和完整性；Thomas D M 等^[11]从质量控制、预处理和用户智能化使用等维度，建设适用于营养和肥胖研究的 AI 就绪数据；还有学者从数据准备工具^[12]、FAIR 数据评估指南^[13]、数据偏差治理方法^[14]、多模态 AI 就绪数据库建设^[15]层面，研究面向大模型的数据质量提升技术，实现数据规模化生产、分类、清洗、关联和融合。综上，实践层面和学术界已对 AI 就绪数据治理的内涵、措施进行了诸多探索，形成可借鉴的通用经验。但医学领域应用场景多元、数据需求复杂、任务多样，AI 就绪数据治理需多重考量，既要遵循数据治理的一般原则和通用路径，又要兼顾医学数据表征、伦理、隐私、社会影响等特点。结合医学领域典型案例，分析和总结其 AI 就绪数据建设经验，能为医学领域 AI 就绪数据治理理论发展、政策制定、实践推进提供决策支持。鉴于 AI 技术体系和 AI 就绪数据特征的动态性、建设过程的复杂性，本文基于 NIH 发布的战略规划、实施的研究专项、资助的科研项目、推行的数据治理实践、出版的研究论文及报告、发布的新闻报道等文献信息，分析和总结其关于 AI 就绪数据治理的实践、建设路径及实施逻辑。选择 NIH 作为典型案例的原因：一是 NIH 是美国联邦政府的主要医学研究机构，在全球卫生健康研究领域影响力大，进入数智化时代，NIH 积极布局数据科学和 AI 战略，探索适合 AI 应用的医学数据建设方案，取

得较大成效，具有典型示范效应；二是 NIH 医学领域 AI 就绪数据建设实践具有连续性，自 2021 年持续布局和实施多个专项，从数据、工具、人才、能力等多维度同步推进，能够提供丰富的分析视角。

3 NIH 建设 AI 就绪数据的实践举措

3.1 实践进展

NIH 拥有丰富的生物医学数据，涵盖电子健康

记录数据、组学数据、医学图像数据、专病数据等，为使这些数据助力 AI 技术研发与应用，NIH 依托其数据科学战略办公室（Office of Data Science Strategy, ODSS），联合下属研究所、研究中心、美国政府机构及企业等，先后实施医学领域 AI 数据治理人才培养、研究数据 AI 就绪程度提升、多模态可信 AI 建设等举措，从标准机制、智能工具、能力建设等方面，构建可信赖、无偏差、符合伦理、模型适配的 AI 就绪数据，见表 1、图 1。

表 1 NIH 建设 AI 就绪数据的实践进展

时间	实践举措	建设内容
2021 年	ODSS 联合 NIH 主任办公室等 15 家 NIH 下属研究所/中心/办公室，实施“信息科学、人工智能和机器学习、生物医学交叉领域劳动力发展”资助专项 ^[16]	面向不同职业发展阶段的生物医学研究人员，提供 AI 就绪数据技能课程或培训计划
2022 年	ODSS 联合 NIH 主任办公室等 9 家 NIH 下属研究所/中心/办公室，实施“生物医学和行为科学中推进 AI 伦理化研发和应用”资助专项 ^[17]	开展 AI 伦理相关的理论、实践、工具、技术、模型及资源等方面的研究
2021—2023 年	ODSS 联合美国国立眼科研究所等 22 家 NIH 下属研究所/中心/办公室，连续实施“提升 NIH 资助研究数据 AI 就绪度”专项 ^[18]	资助 AI 就绪数据治理方法、技术、软件、工具等的研发，覆盖基因、组学、细胞、分子等多种生物医学数据类型
2022—2026 年	NIH 共同基金（NIH common fund）资助“人工智能之桥” ^[19-20] 专项：“The Artificial Intelligence Ready and Equitable Atlas for Diabetes Insights, AI-READI” “Clinical Care AI through the CHoRUS Network, CHoRUS” “Cell Maps for AI, CM4AI” “Bridge2AI-Voice”	围绕数据汇聚、数据共享、社区参与、能力建设、团队促进等内容，建设和共享 AI 就绪专病数据集、临床诊疗数据集、大规模图谱数据集、声学数据集
2021 年	NIH 启动“促进健康公平和研究多样化的人工智能/机器学习联盟计划” ^[21-22]	提升 AI 模型开发中代表性不足群体的参与度和代表性，降低 AI 应用中的偏差及偏见
2024 年	ODSS 发布“多模态 AI 伦理化赋能健康研究”专项 ^[23]	解决多模态 AI 就绪数据统一表示、不同模态数据量相对不足、数据对齐、伦理考量等建设挑战

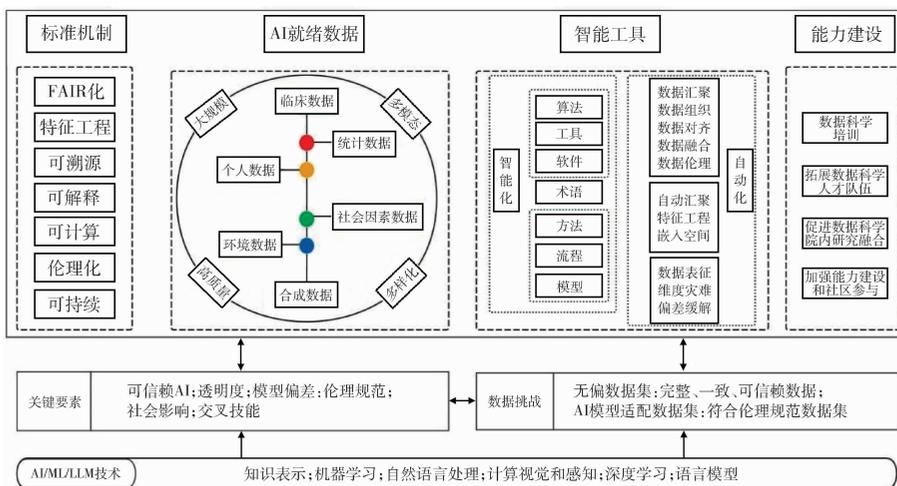


图 1 NIH 面向人工智能的数据治理机制

3.2 标准引领

Bridge2AI 将 AI 就绪数据定义为经合规化汇聚和规范化处理的数据, 可用于训练、分类、预测、文本/图像生成或模拟等场景, 能支撑 AI 模型生成可解释结果, 具备可信赖、高质量、伦理化、可解释、可共享、可计算、特征化等特征^[24]。Bridge2AI 从 FAIR 化、可溯源、特征工程、可解释、伦理化、可持续、可计算 7 个维度提出医学领域 AI 就绪数据建设标准, 确保特定医学场景下 AI 分析的伦理规范性、可解释性、正确性和可信度。其中, FAIR 化标准确保数据可发现、可访问、可互操作和可重用; 可溯源标准确保数据来源、处理步骤、生产主体透明化; 特征工程标准确保数据描述语义化、标准化、丰富化; 可解释标准要求提供数据背景信息、适用性及完整性信息; 伦理化标准确保数据采集、管理、分发符合伦理安全要求; 可持续标准提供长期保存与管理, 保障数据持久应用; 可计算标准要求以机器可读方式提供关于数据描述、访问及背景的信息。

3.3 全域数据

NIH 通过多元参与机制, 聚焦重大卫生健康问题, 加强对科学数据、临床数据、社会因素数据、环境因素数据、个人健康数据等的汇聚、组织与关联, 建设覆盖生物医学群体、个体及研究社区的 AI 就绪数据空间, 提升 AI 技术理解健康、环境、群体及个体关系的能力和效率。AI-READI 项目汇聚 2 型糖尿病相关的人群统计数据、视觉功能数据、社会因素数据、基因数据、诊疗数据、认知功能数据、检查检验数据等, 并从性别、民族、健康状态 3 个维度平衡数据的多样性。CHoRUS 项目通过电子健康档案、医学影像、健康监测等系统, 从美国 14 个临床数据分布点汇集 100 000 例患者的结构化文本数据、影像数据、临床病历数据、体征数据等。CM4AI 提供来自各种数据流的分层细胞图谱, 侧重与癌症、神经疾病、心脏疾病相关的染色质修饰剂、代谢酶及蛋白质等数据类型。Bridge2AI-Voice 旨在建立符合伦理规范的生物声学数据库, 将声音作为健康的生物标志物, 整合到临床诊疗之中, 推动声学 AI 研

究及预测模型的构建, 覆盖神经退行性疾病 (阿尔茨海默病、帕金森病、脑卒中等)、情绪和精神障碍 (抑郁症、精神分裂症、双相情感障碍等)、呼吸系统疾病 (肺炎、慢性阻塞性肺疾病等)、儿科疾病 (孤独症、言语发育迟缓等) 5 种疾病类型, 实现声音数据与电子健康记录数据、放射组学数据、基因组学数据等的汇聚、整合、关联。

3.4 工具赋能

建设覆盖 AI 应用全过程的医学领域 AI 就绪数据处理工具和流程, 研究数据清洗、组织、丰富、增值等关键方法技术, 将大型、多源、多模态数据建成 AI 技术友好型数据集。研发适用于多模态数据汇聚、清洗、融合、对齐、存储、共享的工具和流程, 利用低代码或零代码技术研发面向生物医学研究社区用户的软件平台, 适用于大规模和多样化生物医学数据的可视化工具, 用于器官、系统、个体和群体建模的数字孪生方法, 以及新型隐私保护计算技术, 以增强数据的利用效能。

3.5 能力建设

面向生物医学研究人员不同职业生涯阶段, 特别是面向下一代科学研究人员及处于职业生涯早期的科学研究人员, 组建政策学、生物医学、数据科学、语义科学、信息科学、计算机科学、医学信息学、智能计算学、公共卫生学等跨学科师资队伍, 提供面授、远程授课、自主学习、专题讲座、实践研讨、能力训练营、社交媒体等混合式多元化培养渠道, 设计短期、中期、长期培训课程体系, 提升基础统计和信息素养, 培养 AI 技术和大型数据集思维, 拓展数据和 AI 模型相关知识, 如数据表示、处理、融合、AI 概念、应用程序及可解释性等, 提升生物医学研究人员的 AI 就绪数据生成能力, 缩小与数据科学家的技术及沟通差距。

3.6 可信保障

一方面, 通过分析 AI 伦理规范、用户需求、数据共享、知情同意、多样性要求等, 从政策和实践层面提取确保数据包容性和无偏性的关键要素;

另一方面,提升代表性不足群体在 AI 就绪数据治理中的作用,增加临床、监测、护理中代表性不足群体的数据体量,平衡 AI 模型设计、使用和部署领域研究人员的比例,将贫困、教育、压力、健康食品和医疗保健可获胜、危险要素接触情况等健康社会决定因素数据纳入数据汇聚类型范畴。同时,建立基于 AI 社区反馈的数据质量提升机制,通过 AI 黑客马拉松、小型 AI 应用程序开发、公民科学挑战赛等活动,识别和确定医学领域 AI 就绪数据建设的差距、不足及改进措施;推进“标准规范-组织监管-分类分级”三位一体的 AI 就绪数据伦理化应用实践,加强 AI 应用伦理标准内涵和范畴研究,引入机构审查委员会评估人类研究的合规性,制定数据分类分级管理制度,根据数据分级制定使用共享协议。

4 启示

以大数据、AI 为代表的战略性数智技术叠浪式发展,特别是以 ChatGPT 为代表的大模型技术广泛应用,凸显高价值语料在 AI 时代的重要地位^[25-26]。建设面向 AI 场域的高质量数据成为我国政策和实践重点布局方向。如国务院印发的《新一代人工智能发展规划》^[27]要求构建智能化数据基础设施等;国家互联网信息办公室等 7 部门印发的《生成式人工智能服务管理暂行办法》^[28]提出推动公共训练数据资源平台建设等;国家数据局等 17 部门联合印发的《“数据要素×”三年行动计划(2024—2026 年)》^[29]布局建设高质量语料库等。卫生健康数据是医学领域数字技术赋能的基础,在医学研究、临床诊疗、政策制定等方面发挥重要作用,但也面临采集存储标准不一、质量参差不齐等挑战^[30],亟待从源头加强数据治理,提升质量和共享程度。NIH 关于 AI 就绪数据治理的实践经验,为我国构建面向 AI 就绪的医学数据治理体系提供参考与启示。

一是动态辩证认识医学领域 AI 就绪数据治理范畴。在强 AI 技术驱动下,医学数据治理对象由结构化、静态、截面型文本数据向非结构化、动态、时序型多模态数据拓展,治理方式由劳动密集

型转向自动化、敏捷化、工程化,治理内容由单一维度拓展为多元维度,治理取向由面向人类用户拓展为面向机器及模型群体。如面向医学和人工智能快速融合发展趋势,上海市经济和信息化委员会、上海市卫生健康委员会、上海市疾病预防控制局等机构联合打造智慧健康 AI 底座,按照专业化、适配性原则,立足最全知识体系、最丰富场景适配、最广泛数据来源,构建卫生健康领域具有前瞻性的高质量“语料魔方”,发布上海市首个卫生健康行业语料库,而且同步构建语料专项治理规程和工具链体系,推动卫生健康领域数据治理从基于规则到模型驱动的模式转变^[31]。

二是建设面向 AI 场域的敏捷化生物医学数据生产体系。构建并向生物医学研究社区提供 AI 就绪数据生产软件、硬件和工具,实现个人健康数据、公共数据、临床诊疗数据、科学研究数据的动态化、伦理化、透明化、自动化汇聚,减少环境和人为操作带来的偏差,确保数据质量控制和分析过程透明化,提升数据内容和元数据的完整性。

三是实施建制化、标准化的数据治理机制。首先实施跨学科团队协同科研模式,整合不同科研主体,建立共享工作机制。其次加强交叉复合型人才培养,弥合专业知识差距。最后面向 AI 应用场景和需求,制定和实施数据标准,实现数据标准与自动化数据处理算法等的融合。

5 结语

在 AI 技术深度赋能社会发展的背景下,数据作为关键战略性资源,提高其利用效率、实现广泛互联、挖掘数据价值,成为建设 AI 适配型基础设施的关键。本文聚焦医学领域 AI 就绪数据建设,以 NIH 实践为研究对象,运用案例分析、内容分析等方法,分析其建设举措、推进机制、实施路径,总结其在标准、工具、伦理、能力等方面的实践经验。医学领域 AI 就绪数据治理是系统工程,应整合生物医学研究、AI 研发、应用场景等多元要素协同推进。后续将围绕医学领域高水平重大创新场景,进一步探索数据治理与 AI 融合

机制等问题。

作者贡献: 王茜负责研究设计、文献调研、论文撰写; 刘湘阁负责案例调研与材料整理; 陈凌云负责论文修订; 方安负责研究设计、论文修订、提供指导。

利益声明: 所有作者均声明不存在利益冲突。

参考文献

- 1 王辰. 做“大医学、大卫生、大健康”的倡行者——王辰院士在协和医学院 2022 届毕业典礼上的讲话 [J]. 中国医学人文, 2022, 8 (8): 5-8.
- 2 王墨, 璞嘉, 严曦梦. 上海合力推动“AI+医学”创新发展 [EB/OL]. [2025-03-06]. https://paper.cnstock.com/html/2024-12/16/content_2006748.htm.
- 3 刘合. 大模型既要“吃得饱”更要“吃得好” [N]. 科技日报, 2025-03-03 (1).
- 4 HOLMSTRÖM J. From AI to digital transformation: the AI readiness framework [J]. Business horizons, 2022, 65 (3): 329-339.
- 5 Earth Science Information Partners. Are your data ready? Take stock with ESIP's new AI-ready checklist [EB/OL]. [2025-02-15]. <https://www.esipfed.org/checklist-ai-ready-data/>.
- 6 UK Office for Artificial Intelligence. National AI strategy [EB/OL]. [2025-03-04]. https://assets.publishing.service.gov.uk/media/614db4d1e90e077a2cbdf3c4/National_AI_Strategy_-_PDF_version.pdf.
- 7 McKinsey & Company. Accelerating AI impact by taming the data beast [EB/OL]. [2025-03-04]. <https://www.mckinsey.com/industries/public-sector/our-insights/accelerating-ai-impact-by-taming-the-data-beast>.
- 8 Deloitte. AI readiness & management framework (aiRMF10) [EB/OL]. [2024-12-28]. <https://www2.deloitte.com/content/dam/Deloitte/us/Documents/public-sector/ai-readiness-and-managementframework.pdf>.
- 9 Gartner. Gartner 路线图: AI 数据就绪的关键要素 [EB/OL]. [2025-01-16]. <https://www.gartner.com/ngw/globalassets/cn/information-technology/documents/2025/q1/ai-ready-data-roadmap-cn.pdf>.
- 10 AI Singapore. AI readiness index (AIRI) [EB/OL]. [2024-12-28]. <https://aisingapore.org/innovation/airi/>.
- 11 THOMAS D M, KNIGHT R, GILBERT J A, et al. Transforming big data into AI-ready data for nutrition and obesity research [J]. Obesity, 2024, 32 (5): 857-870.
- 12 WOOD D, LUBLINSKY B, ROYTMAN A et al. Data-prep-kit: getting your data ready for LLM application development [EB/OL]. [2024-12-28]. <https://arxiv.org/abs/2409.18164>.
- 13 CHEN Y, HUERTA E A, DUARTE J, et al. A FAIR and AI-ready higgs boson decay dataset [J]. Scientific data, 2022, 9 (31): 1-8.
- 14 KIDWAI-KHAN F, WANG R, SKANDERSON M, et al. A roadmap to artificial intelligence (AI): methods for designing and building AI ready data to promote fairness [J]. Journal of biomedical informatics, 2024, 6 (154): 1-8.
- 15 FENG S, CAI A, WANG Y, et al. A robotic AI-chemist system for multi-modal AI-ready database [J]. National science review, 2023, 10 (12): 1-3.
- 16 NIH Office of Data Science Strategy. Administrative supplements for workforce development at the interface of information sciences, artificial intelligence and machine learning (AI/ML), and biomedical sciences [EB/OL]. [2024-12-12]. <https://datascience.nih.gov/artificial-intelligence/initiatives/Workforce-Gap-Data-Governance-AI>.
- 17 NIH Office of Data Science Strategy. Administrative supplements for advancing the ethical development and use of AI/ML in biomedical and behavioral sciences [EB/OL]. [2024-12-12]. <https://datascience.nih.gov/artificial-intelligence/initiatives/ethics-bias-and-transparency-for-people-and-machines>.
- 18 NIH Office of Data Science Strategy. Administrative supplements to support collaborations to improve the AI/ML-readiness of NIH-supported data [EB/OL]. [2024-12-12]. <https://datascience.nih.gov/artificial-intelligence/initiatives/Improving-AI-readiness-of-Existing-Data>.
- 19 NIH Office of Strategic Coordination. Notice of intent to publish a funding opportunity announcement for research opportunity announcement for the data generation projects of the NIH bridge to artificial intelligence (Bridge2AI) program (OT2) [EB/OL]. [2024-12-27]. <https://grants.nih.gov/grants/guide/notice-files/NOT-RM-21-022.html>.

(下转第 14 页)

- 药学杂志, 2024, 59 (4): 369–374.
- 18 HAIYAN L, MENG C. Quality assessment and indicator system frame of national medical scientific data sharing project [J]. World science and technology, 2009, 11 (4): 578–581.
- 19 LEE S, ROH G H, KIM J Y, et al. Effective data quality management for electronic medical record data using SMART DATA [J]. International journal of medical informatics, 2023, 180 (12): 105262.
- 20 KAHN M G, CALLAHAN T J, BARNARD J, et al. A harmonized data quality assessment terminology and framework for the secondary use of electronic health record data [J]. eGEMs, 2016, 4 (1): 1244.
- 21 SCHMIDT C O, STRUCKMANN S, ENZENBACH C, et al. Facilitating harmonized data quality assessments. A data quality framework for observational health research data collections with software implementations in R [J]. BMC medical research methodology, 2021, 21 (1): 63.
- 22 王卷乐, 陈沈斌. 地学栅格格网数据质量评价指标与方法 [J]. 测绘科学, 2006 (5): 83–85, 82.
- 23 吴贤玲, 王秀珍, 黄敬峰. 面向水稻生长模拟的网络共享数据综合评价 [J]. 杭州师范大学学报 (自然科学版), 2012, 11 (5): 464–469.
- 24 FUCHS S, NORDEN B, NEUMANN F, et al. Quality – assurance of heat – flow data: the new structure and evaluation scheme of the IHFC global heat flow database [J]. Tectonophysics, 2023, 863 (9): 229976.
- 25 KONG Q, CHEN Y, FANG W, et al. Analysis of space – borne GPS data quality and evaluation of precise orbit determination for COSMIC – 2 mission based on reduced dynamic method [J]. Remote sensing, 2022, 14 (15): 3544.
- 26 OLSEN A R, VOLDNER E C, BIGELOW D S, et al. Unified wet deposition data summaries for North America; data summary procedures and results for 1980–1986 [J]. Atmospheric environment, 1990, 24 (3): 661–672.
- 27 WANG R Y, STRONG D M. Beyond accuracy: what data quality means to data consumers [J]. Journal of management information systems, 1996, 12 (4): 5–33.
- 28 张洋, 肖燕珠. 生命周期视角下《科学数据管理办法》解读及其启示 [J]. 图书馆学研究, 2019 (15): 37–43, 13.
- 29 殷沈琴, 徐萍. 国际数据管理计划工具比较研究 [J]. 图书馆杂志, 2024, 43 (4): 108–117, 142.
- 30 陈美, 梁乙凯. 开放政府数据隐私风险控制中个人数据匿名化研究 [J]. 图书馆学研究, 2021 (11): 66–71.

(上接第 6 页)

- 20 NIH Common Fund. Bridge2AI: propelling biomedical research with artificial intelligence [EB/OL]. [2024–12–27]. <https://bridge2ai.org/>.
- 21 NIH. The NIH’s artificial intelligence/machine learning consortium to advance health equity and researcher diversity program – data and research core [EB/OL]. [2024–12–27]. <https://www.aim-ahead.net/data-and-research-core/>.
- 22 NIH. The NIH’s artificial intelligence/machine learning consortium to advance health equity and researcher diversity program – aim [EB/OL]. [2024–12–27]. <https://www.aim-ahead.net/programs/aim-ahead-dicb-program/>.
- 23 NIH Office of Data Science Strategy. Advancing health research through ethical, multimodal AI [EB/OL]. [2025–01–02]. <https://datascience.nih.gov/sites/default/files/MAI-Solicitation-outline.pdf>.
- 24 CLARK T, CAUFIELD H, PARKER J A, et al. AI – readiness for biomedical data: Bridge2AI recommendations [EB/OL]. [2025–01–02]. <https://pubmed.ncbi.nlm.nih.gov/39484409/>.
- 25 赵国屏. 生物医学研究范式的转变——大数据 + 人工智能的机遇和挑战 [EB/OL]. [2025–01–02]. <https://www.shlab.org/news/5443089>.
- 26 张智雄. 在人工智能时代贡献文献情报领域的智慧和方案 [J]. 农业图书情报学报, 2023, 35 (1): 5–8.
- 27 新一代人工智能发展规划 [EB/OL]. [2024–12–26]. https://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm.
- 28 生成式人工智能服务管理暂行办法 [EB/OL]. [2024–12–26]. https://www.gov.cn/zhengce/zhengceku/202307/content_6891752.htm.
- 29 “数据要素 ×” 三年行动计划 (2024–2026 年) [EB/OL]. [2024–12–26]. https://www.cac.gov.cn/2024-01/05/c_1706119078060945.htm.
- 30 刘辉. 卫生健康数据应用的机遇和挑战 [J]. 中国卫生, 2024 (7): 18–19.
- 31 上海发布卫生健康语料库暨首批医疗应用场景 MaaS 平台 [EB/OL]. [2024–12–26]. <https://www.shanghai.gov.cn/nw31406/20241128/ecf8db565a4944d39e2a72defbf a05e7.html>.