 DOI: https://doi.org/10.12349/iser.v6i7.6697

Analysis of Hospital Information Management Platform Construction Based on Big Data Analysis

Guanjun Chen

Zhongtongfu Consulting and Design Research Institute Co., Ltd., Nanjing, Jiangsu, 210019, China

Abstract

In the information age, hospitals have accumulated a large amount of data information during their informatization construction. These data information have diverse sources, huge quantities, and various forms. With the growth of data information and the evolution of technology, the position of big data management platforms in hospital data management has become increasingly prominent. However, there are still some problems and challenges in the construction of hospital informatization management platforms. This article elaborates on the background and development of the big data industry, and also analyzes the construction of hospital big data systems. The aim is to provide more convenience for big data processing, ensure fast processing speed and close to user needs, while ensuring the safe circulation of medical data and maximizing the value of data. It provides precise and comprehensive services and support for numerous stakeholders - from patients to medical staff, researchers, and managers.

Keywords

Big data analysis; Informationization; Hospital management

基于大数据分析的医院信息化管理平台建设分析

陈冠军

中通服咨询设计研究院有限公司，中国 · 江苏 南京 210019

摘　要

在信息时代，医院在信息化建设过程中累积了大量的数据信息，这些数据信息来源广泛、数量庞大、形式多样，随着数据信息的增长和技术的演进，大数据管理平台在医院数据管理中的地位日益凸显，然而在医院信息化管理平台建设中还存在一些问题和挑战，本文对大数据产业的背景，发展进行阐述，也针对于医院大数据系统建设进行分析，希望能更好的为大数据处理提供便利，确保处理速度快捷、贴近用户需求，同时保障医疗数据的安全流通，实现数据价值的最大化，为众多利益相关者——从患者到医务人员，再到科研人员和管理人员——提供精准、全面的服务与支持。

关键词

大数据分析；信息化；医院管理

1 引言

近年来，全球大数据产业迎来了飞速发展的新阶段，技术和应用的进步日新月异。据国际知名机构 Statista 预测，到 2035 年，全球数据量将激增至惊人的 2142ZB，其中健康医疗大数据占据重要地位。随着医疗领域数据量的迅猛增长，健康医疗大数据正成为大数据应用领域的焦点。这一趋势不仅反映了科技进步的力量，也突显了大数据在推动健康医疗领域创新发展中的巨大潜力 [1]。

2 问题与挑战

2.1 健康医院医疗大数据汇聚困难

健康医疗大数据呈现出类型复杂、规模庞大的特征，广泛分布于各个数据池中，包括医疗机构的电子病历、费用结算数据，药械厂商的多维度信息，区域卫生信息平台的居民健康档案，以及政府掌握的庞大人口与公共卫生数据等。这些数据不仅分散且异源异构，质量参差不齐，不完整性明显。随着个体健康信息的加入，数据量将呈现出爆炸性增长，这无疑给数据的存储、分析、处理和传输带来了巨大的挑战。因此，构建一个能够高效整合、处理和应用这些数据的医疗大数据平台，将成为未来医疗健康领域的重要课题。

2.2 数据标准化应用程度低

尽管医疗数据已具备一定的标准基础，但在汇集后形成的健康医疗大数据领域中，统一的国家或行业标准仍然缺失。这导致了平台厂商、医院生产系统等各自为政，数据标准难以统一，给健康医疗大数据平台的数据质量和数据治理效果带来了严重挑战，这成为了制约健康医疗大数据发展的瓶颈。因此，为了确保健康医疗大数据的优质高效治理，我

【作者简介】陈冠军（1985–），男，中国江苏南京人，硕士，高级工程师，从事智慧医疗、信息化咨询研究。

们急需建立健全的标准规范运行监管机制，统一数据标准，确保各平台、系统之间的数据互通与共享 [2]。

2.3 大数据治理难度高

医疗数据，尤其是病历信息，充斥着大量非结构化数据，如文本和图像，这些数据的处理与分析都异常复杂。在实践中，我们不得不面对数据标准多样化、数据量激增、新技术突破以及管理协调的繁琐等问题。因此，在构建医疗大数据平台时，必须根据医疗数据的特点精心选择相关技术，特别是在数据采集、数据挖掘、数据治理等核心环节，我们必须从作业、模型、物理资源等多维度进行综合评估，确保所选技术和架构能够紧密贴合医疗业务的实际需求。这样，我们才能在保证数据质量的同时，充分挖掘医疗大数据的价值，为患者、医务人员和管理者提供更精准、更全面的服务和支持 [3]。

2.4 医院医疗数据安全隐患

随着医疗数据电子化的进程不断加速，医疗数据和应用呈现出惊人的增长态势。这一转变不仅为动态数据的安全监控和隐私保护带来了前所未有的挑战，还使得我们面临着一系列严峻的数据安全风险，如非法登录、越权访问、异常调阅等。医疗数据中含有的个人隐私信息无比丰富，一旦隐私泄露，不仅患者的人格和尊严会受到侵害，还可能给他们的健康和生活带来不良的影响。因此，构建健康医疗大数据平台时，数据安全管理体系的建立显得尤为重要。这一体系应全面涵盖数据存储、网络设备和基础设施的安全保障，并采取数据流转全程留痕、数据安全监测和预警、以及数据泄露事故的可查询和可追溯等具体措施，确保医疗数据的安全性和患者的隐私得到充分的保护。

2.5 维护医院医疗数据安全的挑战

2021 年 11 月 1 日，《个人信息保护法》正式生效，其标志着中国个人信息保护的法律体系更加成熟、稳固。对于与公众个人信息紧密相关的医疗行业而言，确保数据合规不仅是一项职责，更是对数据管理与使用者的重大挑战。当前，行业内的 HIPPA、JCI 等评审多依赖于外国标准，这无疑可能引发国家安全方面的隐忧。此外，当中国的医疗机构面临个人数据侵权投诉或纠纷时，由于缺乏系统性的处理与应对措施，一旦遭遇大规模数据泄露，后果将不堪设想。因此，为了满足《个人信息保护法》的要求，维护医疗数据的安全，中国亟需完善相关的操作指引，并构建一套系统的应对策略，以确保在面对潜在风险时，能够迅速、有效地进行应对。

2.6 健康信息平台运营体系缺乏

自“十四五”规划实施以来，中国区域全民健康信息平台建设已取得了显著的阶段性成果，不仅整合了各卫生健康机构的信息数据，实现了互联互通与资源共享，还在公共卫生、医疗服务、医疗保障等多个领域发挥着重要作用。然而，随着大数据、云计算、物联网等先进技术的广泛应用，区域平台建设中的一些问题开始浮出水面，如“重建设轻应用，重业务轻治理，重数量轻质量”的倾向，导致平台在服务创新、数据质量、开放共享、安全合规等方面出现短板。为了解决这些问题，我们需要加强对平台的应用创新，提升数据质量，促进开放共享，并确保安全合规 [4]。

3 医院大数据系统建设分析

本文关注医院大数据平台的建设，采用了“一湖三台两中心”的大数据核心架构。所谓“一湖”，即指利用数据湖作为存储、调用、分析和挖掘海量异构健康医疗数据的技术支撑，实现高效稳定的数据库架构。在“三台”方面，包括数据中台、业务技术中台和开放应用平台，面向多级医疗数据汇聚、数据治理和数据共享提供了技术支撑。而“两中心”则分别为运营维护和安全管控提供了支持，以探索健康医疗大数据的运营模式、加密数据脱敏和确保数据应用流程的安全。通过以上架构的设计和实施，医疗机构可以更好地管理和应用医疗数据，推动医疗信息化进程，为医疗服务和研究提供更好的支持和保障。



图 1 医院大数据系统管理架构

医院大数据系统主要模块包括

（1）数据湖

基于前沿大数据处理技术，平台将 Hadoop 与 MPP 数据库相结合，构建了湖仓一体架构，为数据处理提供了强大的支持。这一架构不仅能够实现数据的汇聚存储，还能进行开发治理和分析挖掘等操作。通过批处理引擎、流处理引擎和大规模并行处理引擎等计算引擎，平台为用户提供了底层存储和计算能力。整体系统采用分布式架构，可以通过水平扩展满足不断增长的业务需求。这一综合的解决方案，为企业提供了一体化的数据处理平台，有效提升了数据的处理效率和质量。

（2）数据中台

数据中台是一种面向数据治理开发人员的智能、敏捷的数据管理平台，通过全方位的监控和治理，实现数据的统一管理，保证数据在整个流程中的完整性、准确性、一致性和时效性，帮助用户建立符合自身特征的数据架构和数据治理体系。数据中台包含结构化与非结构化数据治理工具集、A1 标注和 AI 训练等子模块。AI 标注可以实现对医疗文本和影像数据的人工或自动标注，而 AI 训练则支持基于标注

数据集的 AI 建模，包括训练、管理和评估等功能，集成了 TensorFlow、Caffe 等开源 AI 框架。通过数据中台，企业能够更好地掌握数据的全貌，实现数据的有效管理和应用。通过数据中台的应用，企业能够更快地适应市场的需求，更好地应对竞争压力，实现数据驱动的智能化发展。

（3）业务技术中台

业务技术中台是指为各上层应用提供通用的技术和业务能力的平台。它可以提供基于 K8s 的容器化运行环境、容器镜像仓库、数据库服务、消息队列服务、大数据服务、开发者支持服务、统一用户管理服务、统一数据访问服务、AI 数据集管理、AI 训练环境、AI 推理服务、联盟区块链服务以及隐私计算服务等一系列功能。通过统一的服务平台，各部门可以便捷地获取所需的技术资源，实现快速的业务开发和部署。同时，业务技术中台的普及还可以促进企业的数据管理和利用效率。通过数据沙箱服务，企业可以实现数据的安全共享和隐私保护，保证数据的合规性和安全性，而可视化数据沙箱服务则可以帮助企业更加直观地理解和分析数据，为业务决策提供更有力的支持。总的来说，业务技术中台的建设不仅可以提升企业的业务创新能力和竞争力，同时也有助于提升数据管理和利用效率，实现数据驱动的智能决策。

（4）开放应用平台

开放应用平台致力于提供开放的数据服务，包括数据资源、能力和服务的共享。通过提供数据能力开放、应用能力开放等功能，为客户构建统一的数据平台和应用机制，实现健康医疗大数据平台的安全开放、应用开发和算法调用支持。这样可以促进健康医疗领域数据资源的共享和健康医疗数据应用的深化。通过开放平台，第三方可以获取健康医疗数据资源，并基于平台提供的数据能力和服务进行应用开发和部署，推动健康医疗领域的发展。开放应用平台的建设不仅提高了数据资源的利用效率，也拓展了健康医疗数据的应用领域，为行业发展带来新的机遇和挑战。

（5）运维运营中心

运营运维中心是一个涵盖多个关键模块的综合体，包括运营支撑中心、监管支撑中心、认证鉴权中心、用户管理中心、产品服务中心和运维中心。其功能涵盖数据资源申请审批、计量计费、用户管理、产品分类管理以及系统运维等方面。在系统运维中，具备资产的多维度、全方位管理，业务系统全生命周期监控、任务集中调度和统一管理、异常告警等功能。同时，运营运维中心还需要与周边系统进行集成对接，确保信息互通，实现平台的全面运营功能。通过整合各个功能模块，运营运维中心能够更好地满足用户需求，提升服务质量，促进业务发展。总的来说，运营运维中心是数字化平台运营的核心，为企业提供了全面的运营支持和系统维护功能。只有不断完善和强化各个模块之间的协作与整合，运营运维中心才能发挥最大的效益，推动企业向数字化转型迈出更坚实的步伐。

（6）数据安全管控中心

数据安全管控中心是健康医疗数据安全的坚实后盾，全程保障数据汇聚、存储、治理、开发、共享和回收的安全与合规。我们提供的安全防护措施包括数据库加密、数据脱敏和数据库权限管控等，确保数据在任何环节都得到有效保护。此外，我们支持资产梳理、分类分级、数据库漏洞扫描和数据资产风险评估等功能，全方面提升数据安全管理的效率和水平。通过数据安全运营模块，我们能够统筹日常运营和监管工作，确保风险可感知、事件可处置、策略可优化的数据安全管控系统的建立。在当前数字化时代，数据安全至关重要。我们致力于为健康医疗行业提供最可靠和最完善的数据安全保障，帮助机构和个人建立起健全的数据安全意识和应对能力。让我们携手合作，共同维护健康医疗数据的安全，推动行业的健康发展 [5]。

4 结语

大数据与人工智能的融合在当前大数据领域备受关注，在医疗领域，由于医疗大数据具有多源、异构、非统一等特点，大数据与人工智能的结合更显优势。这种融合为医疗行业带来了更多可能性，可以帮助医生更准确地诊断疾病、预测患者风险，同时也为医疗机构提供了更好的数据管理和治理方式。可以预见，大数据与人工智能的融合将进一步推动医疗行业向着智能化、精准化的方向迈进。

参考文献

[1] 王梅.面向医院的大数据治理模型设计及实施[J].大数据时代,2024,(08):35-39.

[2] 吴佳妮.大数据背景下医院创新财务管理模式分析[J].投资与合作,2024,(08):110-112.

[3] 冯雨萱,何恬睿,沈力,等.医疗大数据质量评价及管理研究综述[J].医学信息学杂志,2024,45(08):1-7.

[4] 杨冰倩.基于医疗大数据环境的疾病预测模型设计[J].科技资讯,2024,22(16):41-44.

[5] 周晓瑜,相琼,常杰,等.大数据背景下多级可拓模型在医院儿科医疗设备质量管理评价中的价值研究[J].中国医学装备,2024,21(08):159-164.