

# 大型综合医院临床数据中心的设计与应用

杨建兴

浙江医院 浙江 杭州 310000

**摘要:** 以三甲综合性医院为例, 根据国家卫生计生健康委员会颁发执行的《电子病历基本数据集》规范, 围绕医院管理制度和临床大数据应用, 运用信息交换与系统集成技术, 提出了大型综合医院临床数据中心 (clinical data repository, CDR) 的设计与应用。建设遵循“总体规划、顶层设计、分布实施”的原则, 整合现有信息系统, 不断更新临床数据, 以数据反哺业务, 推进医院高质量发展。

**关键字:** 大数据; 临床数据中心; 电子病历

## 引言

CDR是一个实时数据库, 用以整合和统一展示单个患者分布在不同临床数据资源库中的相关信息<sup>[1]</sup>。临床数据中心是医疗信息化、智能化的重要组成部分, 其核心功能是整合来自院内分散的各个信息系统, 如电子病历系统、检验系统以及医学影像存档系统等。临床数据中心系统负责对汇集的数据进行集中管理和操控, 实现数据访问的标准化和统一化, 为医疗服务、科研教学和医疗机构管理提供可靠的数据支持和接口访问。医院各个信息系统在使用过程产生了大量临床数据, 难以实现数据共享, 无法进行有效利用<sup>[2]</sup>。因此, 需要建立一个以患者为中心的高效、安全、简洁的临床数据中心, 实现数据的统一管理利用。

## 1 系统建设目标

建设遵循“整体规划、高层设计、分步实施”的准则, 整合现有的信息系统, 以加速推进“更全面、更深度、更覆盖”的数字化医院建设。

1.1 夯实基础, 全面升级改造现有医院信息系统, 夯实打好医院信息化基础, 保证医院信息化建设可持续性发展。

1.2 全院融合, 基于全院信息系统集成平台, 建设临床数据中心, 保证系统健壮性和加强数据的可持续应用。

1.3 创新发展以“服务优化、创新支持”提升业务、管理、运行能力和水平, 提升医疗卫生服务质量、效率和能力。

1.4 以患者为核心, 重构患者就诊业务, 强化患者对医院的粘度。

1.5 积聚智慧, 依托积累的大量的临床业务数据, 构建医院知识库体系和科研体系, 提升医院临床智能化水平和质控管理能力, 提升医院科研大数据分析能力。

1.6 精细管理, 打通医院财务和业务系统之间的数据

共享通道, 实现管理的科学化、规范化和智能化, 实现医院精细化管理的目标。

## 2 需求分析

### 2.1 面向医务人员的需求

医生在诊疗过程中形成了大量的医疗数据, 有效识别的医疗数据可以帮助医生、护士和其他医疗专业人员快速准确的获取到医学知识。系统需满足常见疾病的知识库详情的解读功能, 包括诊断详情、药品、手术、检验、检查等知识库。需满足科研数据采集需求, 医院存储的病历通常为非结构化数据, 医生难以对其进行分析再利用, 希望借助临床数据中心完成病历文档后结构化处理, 要求对病历原始数据, 修改痕迹可留痕, 实现对历史病历的结构化处理。

### 2.2 面向患者的需求

临床数据中心的设计, 需要满足以患者为中心的医疗服务, 重构患者就诊业务流程, 提升医院整体运营能效。医生需要查看患者的历次治疗情况, 用药史, 过敏史, 检查记录等信息, 虽然医院早已将数据集成在医生站, 但是展示界面不够简洁, 需要多次跳转。希望通过临床数据中心的建立, 在不同的组织机构, 不同的信息系统间的资源整合, 形成以患者为主索引的电子病历新引擎, 以患者历次就诊时间节点为主线, 将患者在医院历次诊断、用药史、手术史等医疗服务活动有机地关联起来, 自动采集转换为符合国家标准临床信息数据并存储到数据中心, 直观的展示患者历次就诊所有医疗数据。

### 2.3 面向管理者的需求

临床数据中心需要构建一个面向医院管理者的质控平台, 为医院的管理、运营和决策提供可靠的数据支持, 从而实现医院的高质量发展。满足管理者对多种关键数据进行分析, 包括医院的门诊量、手术量、出入院患者数、床位使用率等。同时, 数据平台应具备对对比分

析功能，能够将当前数据与既往数据比较分析，以便于医院管理者做出更为精准和科学的决策。

### 3 系统部署及总体框架

#### 3.1 概述

临床数据中心涉及的系统规模大、业务复杂，为保障系统顺利进展，采用了Scrum敏捷软件工程管理，并结合相应的版本管理工具、过程管理工具、bug跟踪、任务跟踪自动化管理工具，以技术手段来保障项目管理的科学性、连续性、闭环可监控性。

针对医院的需求分析，临床数据中心基于《电子病历基本数据集》规范，进行科学分类和抽象描述，采用基于面向服务的体系架构（SOA）、高内聚低耦合系统架构，设计了符合临床要求，医院管理的临床数据中心。

#### 3.2 系统部署

临床数据中心数据平台分为大数据处理集群和大数据应用集群两部分。大数据处理集群基于linux操作系统，提供大规模高性能分布式数据存储和计算能力。大数据应用集群采用微服务架构提供灵活、高效的应用开发、部署、运维能力。

##### 3.2.1 搜索引擎部署

为了提高搜索效率，在基础环境中搭建分布式搜索

引擎，这样可以实现千万份病历中秒级返回搜索结果，可以支持复杂的查询条件并支持全文检索。

(1) 搭建高可用高性能可扩展的分布式搜索引擎系统，支持数十亿级数据的存储和索引，支持全文检索、数值范围、正则匹配、层级关联等多种查询方式。

(2) 提供查询语言智能构建引擎，对各业务场景查询条件进行智能分析，生成搜索引擎能够识别的搜索语句，并从搜索引擎系统中召回精确相关的患者病历数据。

(3) 提供医疗语言语义引擎系统，可对用户查询语句进行智能纠错、主干词识别、相似查询、归一匹配等分析，识别医护真正的搜索需求。

##### 3.2.2 微服务架构部署

数据应用层的整体框架设计基于面向服务（SOA）的思路，以微服务的设计理念，将业务功能按医疗业务域抽象、建模，并在此基础上形成一个个的服务单元，实现服务的注册与发布。整个微服务平台还提供了可视化的服务治理平台，能动态配置服务的路由规则、权重调节、负载均衡。提供多种数据集成方式，除数据平台集成方式外还提供webservice、数据库视图等多种常规接口方式。

#### 3.3 总体架构

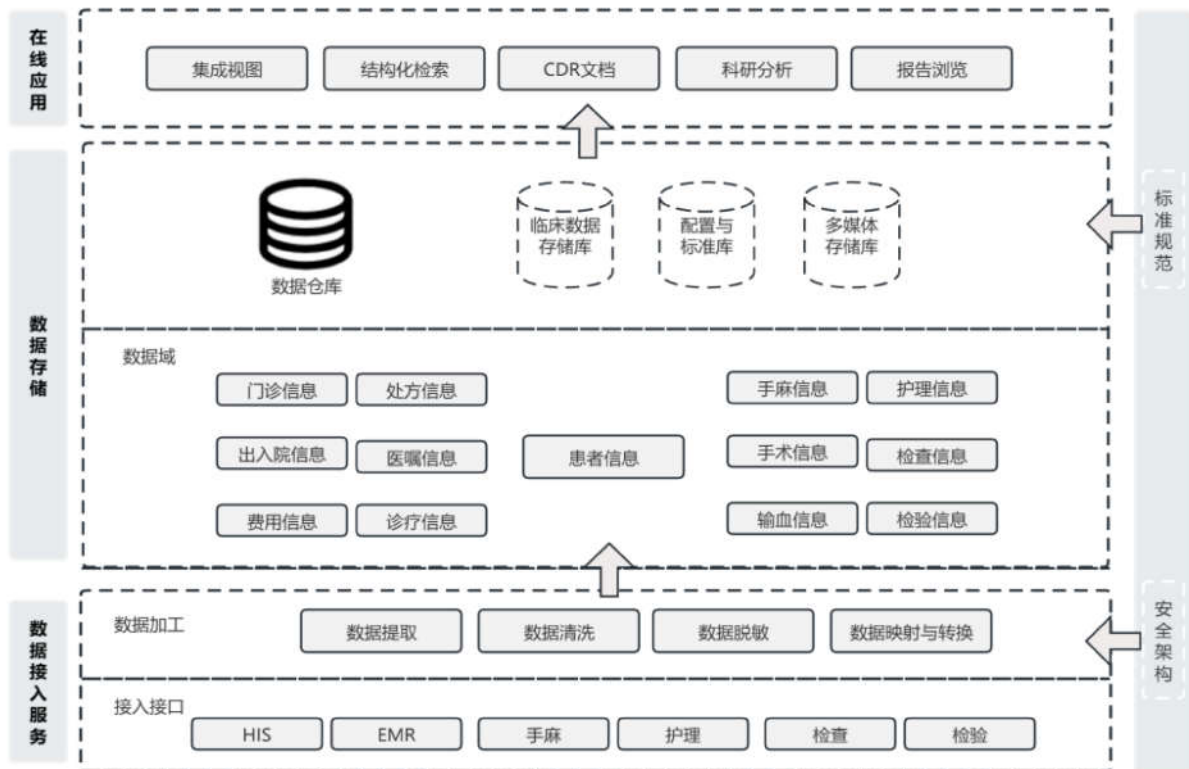


图1 系统整体架构

临床数据中心服务器的物理位置在院内机房,采用B/S技术架构,基于J2EE的技术标准,开发实现选用主流的开源体系技术, JAVA语言开发。后台基于Linux操作系统运行, J2EE应用服务器, 以保障系统7\*24小时稳定运行。linux建设在虚拟化机器上, 便于减容扩容以及其他管理工作。系统后台使用基于内存亦可持久化的缓存技术, 用做数据缓存, 采用组件化开发, 提升应性能, 具有良好的安全性, 可扩展性和容错性。

系统的总体架构包含数据接入服务、数据存储与在线应用。其中, 数据接入服务负责从各个源系统中抽取数据。数据存储负责数据建模、存储。在线应用负责充分有效地利用数据。系统整体框架, 见图1。

### 3.4 数据安全

医疗机构在日常运营中无法避免的需要收集、存储和处理大量的数据, 其中包括不仅限于医院自身的运营数据, 更主要的是患者的个人信息。这些数据属于《个人信息保护法》定义的敏感信息<sup>[3]</sup>。医院信息中心有义务有责任做好患者的个人隐私保护工作。因此, 在系统前期设计阶段, 为了数据安全存储, 设计了访问控制、权限管理以及日志记录。

访问控制和权限管理, 实施严格的访问控制和权限管理制度, 系统做了权限访问控制, 不同的角色有不同的数据访问权限, 针对医生、护士、医技等不同的角色制定相应的数据访问规则, 只有经过授权的用户才能访问他们职责范围内所需的数据, 防止未经授权的访问和数据泄露。日志记录, 建立全面的日志记录机制, 跟踪用户对数据的访问和操作, 用户在系统中所有的操作留有痕迹, 以便在发生安全事件时进行追溯和调查。

#### 3.4.1 数据脱敏和加密

临床数据中心采用SSL(安全套接层)加密, 分离密钥和加密数据, 使用过滤器和数据备份等方式, 构建数据的存储安全策略<sup>[4]</sup>。数据脱敏和加密, 包括患者信息中的敏感数据: 患者姓名、电话、身份证号、地址等, 利用脱敏或加密规则对数据进行变形处理, 以达到加密的效果。

采用脱敏和加密算法又两种方式, MD5加密算法与DES加密算法。MD5加密算法, 是一种不可逆的算法, 展示界面可用“\*”显示。DES加密算法, 是一种可逆的算法, 方便按权限进行解密, 满足特定业务需要显示明文数据的场景。通过加密脱敏, 确保敏感数据在存储时以密文形式存在, 保护敏感隐私数据。

#### 3.4.2 数据存储与传输安全

数据存储加密, 即大数据平台处理后的重要数据先

加密再存储, 防止重要数据被非法窃取或窥探。可通过系统指定重要字段进行数据加密, 通过为系统生成专用的token来实现对系统平台的密钥的分发与管理。数据传输安全, 在网络传输中采用SSL/TLS安全协议为数据传输提供安全保障, HTTPS协议保障在使用浏览器搜查看病历全文时, 保障用户身份认证安全和数据传输安全。

通过这些综合性的安全措施和技术, 大数据中心能够有效地保护敏感数据的完整性、保密性和可用性, 降低数据泄露和违规访问的风险, 确保数据在存储层的安全加固。同时, 持续的监控、审计和改进也是保障大数据中心安全不可或缺的部分。

## 4 系统应用

### 4.1 科研分析应用

临床数据中心提供了丰富的临床数据资源, 支持临床科室医学研究工作。通过对临床数据的分析和挖掘, 可以发现疾病的发生发展规律、治疗效果和不良反应等方面的信息, 从而进行疾病研究、药物试验和临床试验等科研活动。同时, 临床数据中心为医院的跨学科研究和合作提供支持, 在临床医学、基础医学、公共卫生等领域开展跨学科研究, 提高医院科研水平和医疗质量。

### 4.2 结构化数据应用

机构化数据应用包含结构化检索、360集成视图。结构化检索, 医务工作人员通过关键字进行病历的全文检索, 可以按门诊或住院类型、性别、年龄范围、病历类型、门诊科室等进行跨不同类型病历的二次筛选。360集成视图, 作为数据中心数据展示窗口, 医生可以在一个操作界面上将历次的门诊、住院就诊信息按时间轴进行集中展现, 支持将单次就诊活动的诊断、医嘱、检查、检验、病历、手术、护理的概况全景展示。

### 4.3 报告管理应用

临床数据中心对患者历次报告进行集中管理, 包含管理报告上传、报告作废、报告更新。管理报告上传, 报告由业务系统生成后, 统一上传到报告中心, 进行集中式存储。报告作废, 原有报告存在缺陷时, 业务系统可通过消息机制通知报告中心立即作废报告, 等待新报告生成。报告变更, 将新报告上传到报告中心, 并且更新报告信息, 建立索引, 而变更前的报告需要撤销索引, 但作为历史报告痕迹备查。同时, 报告变更时, 根据报告浏览、下载或打印的记录, 自动通知医护人员和患者, 保证变更后的报告能及时送达给各相关干系人, 避免误诊等情况发生。

## 结束语

临床数据中心从业务系统中抽取数据, 将原先分布

在各业务系统中的信息整合并加工处理形成全院的数据中心,并在此基础上实现对医院信息资源的二次利用,最大限度地方便病人就医、方便医院一线医护人员工作、方便医院管理人员分析决策。同时,为医务人员提供基于大数据技术的医疗服务,深入洞察病症诊疗手段与成果,为相关病症研究提供数据支撑,提高医生医疗水平,有效推动医院高质量发展。

#### 参考文献

- [1]秦贞华. 医院临床数据中心CDR构建[J]. 电子元件与信息技术,2022,11:136-139.
- [2]姜林. 基于集成平台的临床数据中心分析[J]. 网络安全技术与应用,2020,11:140-141.
- [3]李立果. 数据安全法下的医疗数据安全建设[J]. 网络安全和信息化,2022,04:128-133.
- [4]闫慈,王鹏,杨越,任劲,吴睿豪,管音,张茜. 省级肿瘤大数据中心规划与建设[J]. 医学信息学杂志,2020,11:54-59.