

# 电子病历智能检索系统设计与实现

刘 谦

牟雪菲 李俊伟

羊海锋

(海南省人民医院  
海口 570311)(海南易建科技股份有限公司  
海口 570125)(海南省人民医院  
海口 570311)

〔摘要〕 以海南省人民医院为例, 阐述电子病历智能检索系统设计思路、技术路线、架构以及具体功能, 指出该系统应用有助于提升对临床科研的支撑保障能力, 为实现精准医疗和智慧医疗服务体系奠定坚实基础。

〔关键词〕 电子病历; 智能检索; 大数据; 分布式存储

〔中图分类号〕 R-056 〔文献标识码〕 A 〔DOI〕 10.3969/j.issn.1673-6036.2021.01.011

**Design and Realization of Electronic Medical Records Intelligent Retrieval System** LIU Qian, Hainan General Hospital, Haikou 570311, China; MU Xuefei, LI Junwei, Hainan Yijian Science and Technology Co., Ltd., Haikou 570125, China; YANG Haifeng, Hainan General Hospital, Haikou 570311, China

〔Abstract〕 Taking Hainan General Hospital as an example, the paper expounds the design idea, technical route, architecture and specific functions of the electronic medical record intelligent retrieval system, points out that the application of the system is helpful to improve the support ability of informatization to clinical scientific research, and lays a solid foundation for the realization of precision medicine and smart medical service system.

〔Keywords〕 Electronic Medical Records (EMR); intelligent retrieval; big data; distributed storage

## 1 引言

目前国内大部分医院都上线了电子病历系统, 但这些系统往往只解决病历数据电子化问题, 没有很好地实现电子病历的大数据分析, 也没有有效利用和服务于临床。国家卫健委电子病历应用水平评价指南明确指出电子化之后病历数据交换、共享以及后续数据挖掘、利用才是决定电子病历应用水平的重要指标<sup>[1]</sup>。由于历史原因, 医院同时运行着上百种医疗业务信息系统, 这些信息系统的数据构成

电子病历数据的重要组成部分。但这些信息系统往往多源、异构, 彼此割裂, 形成信息烟囱, 导致各种电子病历医疗数据处于孤立状态, 无法得到有效、合理的利用<sup>[2]</sup>。海南省人民医院是海南省最大的综合性三级甲等医院和全省疑难危重症诊治中心, 希望引入大数据与人工智能等新兴技术, 提高医院信息化建设整体水平, 开展医疗健康大数据技术探索和产品研发, 促进大数据与人工智能技术在医院的落地应用。其中电子病历智能检索系统是医院研发的一个大数据应用系统。

## 2 系统结构

电子病历智能检索系统在已有电子病历数据和

〔修回日期〕 2020-11-19

〔作者简介〕 刘谦, 研究员, 发表论文 70 余篇。

各种临床业务信息系统数据的基础上,利用自然语言处理技术,通过设置电子病历相关参数与关键信息节点,通过大数据和人工智能处理算法,从各种临床业务信息系统中自动抓取电子病历关键数据,将患者分散在医院业务信息系统中的临床诊疗数据按照患者主索引进行深度融合,以患者为中心汇集到一起重新进行梳理,实现以时间轴集成视图方式展现每位患者门诊、住院病历、医嘱、医技检查、检验信息,从而形成患者全方位的临床诊疗信息检

索系统<sup>[3]</sup>。实现集临床数据采集、存储、分析、挖掘于一体,符合临床需求的电子病历智能检索系统。以此为基础探索建设以患者为核心的临床信息系统,通过对临床诊疗大数据的分析、挖掘和利用,逐步将医院信息化建设重心从对医疗数据管理转变为对临床知识的管理,从而提升医院信息化对临床科研的支撑保障能力,为最终实现精准医疗和智慧医疗服务体系奠定坚实基础。系统结构,见图1。



### 3 系统设计

#### 3.1 设计思路

采用以患者为核心的设计思路,通过对医疗大数据的分析、挖掘和利用,实现对于患者临床诊疗数据的智能化管理与检索。电子病历智能检索系统涉及的数据量巨大,因此采用主流的分布式存储架构进行数据存储与备份。数据采集功能设计上,充分考虑目前临床诊疗多元化数据来源和多形态数据特征的复杂情况,主要对医院已经运行的业务信息系统数据进行抽取、整理和清洗,同时结合区域人口数据库主索引,利用互联网技术采集患者院外临床诊疗数据。采集的数据分为结构化、半结构化、非结构化3大类。为配合数据采集工作需要,采用多模式的采集方案,同时对数据采集过程及质量进

行严格把控,实现采集数据的高可用性、高速度、高质量,为数据分析应用提供可靠的基础数据。

#### 3.2 技术路线

从医院业务数据库中抽取的原始数据首先进入分布式文件系统(Hadoop Distributed File System, HDFS)进行贴源层数据备份,经过清洗和处理后的数据保存在数据仓库Hive中供后续数据服务调用,对于传统的关系型业务数据库保存在MySQL中,需要建立全文索引的数据保存在MongoDB和ElasticSearch中。电子病历智能检索系统为医生提供基于ElasticSearch搜索引擎的数据查询工具,用于病例回顾或科研分析,实现智能化检索,同时提供查询授权功能,限制医生查询时限、范围,留存查询日志,保护数据安全<sup>[4]</sup>。基于电子病历智能检索系统实现的患者360度全景视图是在整合电子病

历数据基础上，利用机器学习和自然语言处理技术，将患者分散在院内业务系统中的电子病历数据以时间轴集成视图方式展现出来，实现每位患者门诊、住院病历数据集中调阅和展示。患者360度全景视图基于浏览器/服务器（Browser/Server, B/S）架构的开发技术，提供应用程序接口（Application Programming Interface, API），可以方便地嵌入各种业务信息系统中，实现在医生工作站、医技工作

站、手术麻醉系统等业务场景的应用。由于系统基础数据集合采用各个临床业务系统中全量数据进行整理、清洗与融合，通过患者360度全景视图可查看患者历次门诊和住院就诊记录，特别是在历次住院就诊记录中几乎完整地呈现患者电子病历全量信息，极大提升临床医生诊断治疗精准度，提高工作效率。系统技术路线，见图2。

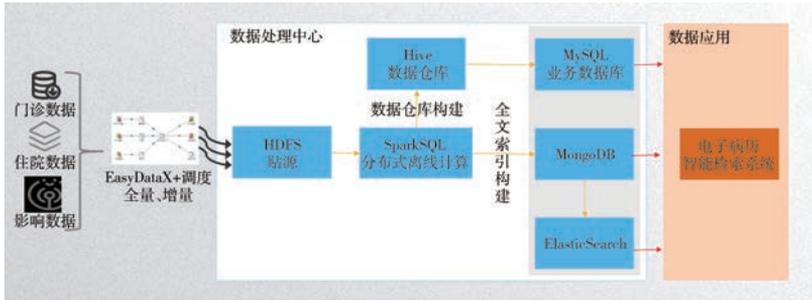


图2 电子病历智能检索系统技术路线

## 4 功能实现

### 4.1 临床诊疗数据快速检索

提供基于医疗大数据平台的电子病历智能检索服务，通过类似于Google的搜索引擎，临床医生可以便捷地检索海量电子病历数据记录，在PC、平板电脑、手机等任何终端调阅患者就医全流程资料，获得最真实、连续、完整的数据，为医院临床科研提供决策支持。

### 4.2 相似病历分析

以就诊事件为单位，对比患者间的就诊事件相似度，对同一科室就诊事件进行相似度分析，当前相似度分析只针对诊断和医嘱，根据权重值进行评分，最后加权平均求得病例间相似分数，方便临床医生检索相似病历，辅助其进行正确的临床判断和决策。

### 4.3 构建病历及专科知识图谱（图3）

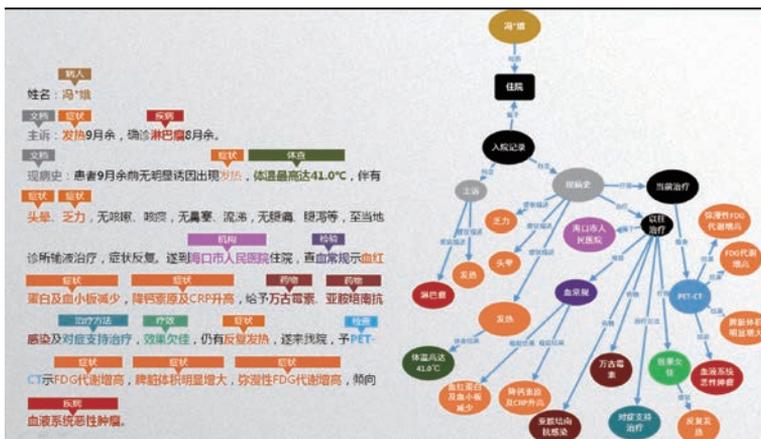


图3 电子病历智能检索系统中知识图谱应用

系统引进知识图谱技术,从电子病历中抽取临床知识,通过本体建模方法构建病历图谱,可将病历数据归纳为患者、疾病、症状、医生、科室、医院、治疗、医嘱、检验、检查、手术、药品、住院事件、门诊事件、体检事件等概念。通过与医疗知识库相关联为实现医学知识检索、疾病预测、用药推荐和相似病历等奠定良好基础,为临床科研工作者提供新的结论以供参考,达到更好地辅助临床医学研究的目的。临床电子病历知识图谱是专科知识图谱的基础,从临床病历中抽取专科电子病历知识图谱,结合专科知识库构建专科知识图谱。

## 5 结语

电子病历智能检索系统从2018年下半年开始在海南省人民医院上线运行,已经完成近50万份电子病历的后结构化处理,整合临床医嘱和检验检查结果约650万条数据。系统集成业务信息系统的

院内数据以及电子健康档案、全员人口库等院外数据,重构数据模型,以患者为中心重新整合,实现患者360全景视图,为患者和医生提供全方位的数据查看。电子病历智能检索系统的应用提升信息化对临床科研的支撑保障能力,为医生深入挖掘医疗数据价值提供可能,为最终实现精准医疗和智慧医疗服务体系奠定坚实基础。

## 参考文献

- 1 甘霖. 基于云计算的电子病历全文检索系统 [J]. 中国数字医学, 2016, 11 (12): 41-43.
- 2 赵景越, 徐松青, 蔡连忠. 全文检索引擎技术在电子病历中的应用 [J]. 数字技术与应用, 2014 (12): 89-90.
- 3 王晓, 罗二平, 张健. 基于语义的电子病历智能全文检索 [J]. 医疗卫生装备, 2008, 29 (4): 45-46.
- 4 阮彤, 高炬, 冯东雷, 等. 基于电子病历的临床医疗大数据挖掘流程与方法 [J]. 大数据, 2017, 3 (5): 83-98.

(上接第28页)

- 4 Smith B, Grenon P. Basic Formal Ontology (BFO) Home [EB/OL]. [2020-03-24]. <http://basic-formal-ontology.org/>.
- 5 Smith B. OBO Library - Basic Formal Ontology [EB/OL]. [2019-08-26]. <http://purl.obolibrary.org/obo/bfo.owl>.
- 6 ISO/IEC. Information Technology - Top - level Ontologies (TLO) - Part 2: Basic Formal Ontology (BFO) [EB/OL]. [2020-07-04]. <https://www.iso.org/standard/74572.html?browse=tc>.
- 7 Smith B. BFO Users [EB/OL]. [2020-07-06]. <http://basic-formal-ontology.org/users.html>.
- 8 Zhu Y, Liu L, Gao B, et al. TCDO: a community-based traditional Chinese drug ontology [C]. Shanghai: Proceedings of the 11th International Biocuration Conference (Biocuration-2018), 2018.
- 9 段宇锋, 黄思思. 基于BFO构建中文植物物种多样性领域本体的研究 [J]. 现代图书情报技术, 2015, 265 (12): 80-87.
- 10 Arp R, Smith B, Spear A D. Building Ontologies with Basic

Formal Ontology [M]. Cambridge: MIT Press, 2015.

- 11 Bandrowski A, Brinkman R, Brochhausen M, et al. The Ontology for Biomedical Investigations [J]. Plos One, 2016, 11 (4): e0154556.
- 12 Smith B, Ceusters W, Klagges B, et al. Relations in Biomedical Ontologies [J]. Genome Biology, 2005, 6 (5): R46.
- 13 Groza T, Kohler S, Moldenhauer D, et al. The Human Phenotype Ontology: semantic unification of common and rare disease [J]. American Journal of Human Genetics, 2015, 97 (1): 111-124.
- 14 Samtivistijai S, Lin Y, Xiang Z, et al. CLO: the cell line ontology [J]. Journal of Biomedical Semantics, 2014 (5): 37.
- 15 郭靖文, 杨晟, 史涪仁, 等. MedPortal: 面向精准医学的生物学本体资源存储和应用平台 [J]. 中国生物医学工程学报, 2017 (5): 557-564.
- 16 朱彦, 何勇群, 郑捷, 等. 基于基本形式化本体的本体构建 [M]. 北京: 人民卫生出版社, 2020.
- 17 Smith, Barry. Classifying Processes: an essay in applied ontology [J]. Ratio, 2012, 25 (4): 463-488.