

基于 BERT 模型的社区慢性病患者服务体验分析

徐若昕^{1,2} 杨婷婷^{1,2} 江宽列³ 刘 静^{1,2} 万金花¹ 朱小娇¹

(¹ 汕头大学医学院附属粤北人民医院 韶关 512026 ² 汕头大学医学院 汕头 515000)

³ 海南医科大学公共卫生学院 海口 571199)

〔摘要〕 **目的/意义** 挖掘慢性病患者诊疗体验深层信息, 为优化分级诊疗政策提供参考。**方法/过程** 收集广东省和重庆市 4 家医院 539 例慢性病患者访谈文本, 基于 BERT – base – Chinese 模型进行主题识别和情感分析。

结果/结论 共分析 4 234 组问答对, 识别出 11 个主题类别。其中医疗服务与机构相关主题占比最高 (37.72%), 其次为就诊流程 (27.23%) 和健康教育 (10.75%)。情感分析显示, “医护人员”和“挂号”等主题情感积极, “转诊”与“经济负担”等主题情感消极。机器学习方法能有效识别慢性病服务中的短板。

〔关键词〕 机器学习; 文本挖掘; 慢性病管理

〔中图分类号〕 R – 058 **〔文献标识码〕** A **〔DOI〕** 10.3969/j.issn.1673 – 6036. 2025. 11. 006

Analysis of Service Experience for Community Chronic Disease Patients Based on the BERT Model

XU Ruoxin^{1,2}, YANG Tingting^{1,2}, JIANG Kuanlie³, LIU Jing^{1,2}, WAN Jinhua¹, ZHU Xiaojiao¹

¹The Affiliated Yuebei People's Hospital, Shantou University Medical College, Shaoguan 512026, China; ²Shantou University Medical College, Shantou 515000, China; ³School of Public Health, Hainan Medical University, Haikou 571199, China

〔Abstract〕 **Purpose/Significance** To mine the deep information of the diagnosis and treatment experience of patients with chronic diseases, and to provide references for optimizing the hierarchical diagnosis and treatment policy. **Method/Process** The interview texts of 539 patients with chronic diseases from 4 hospitals in Guangdong province and Chongqing are collected. Topic identification and sentiment analysis are conducted based on the BERT – base – Chinese model. **Result/Conclusion** A total of 4 234 question – answer pairs are analyzed, from which 11 topic categories are identified. Topics related to medical services and institutions account for the highest proportion (37.72%), followed by the appointment and treatment process (27.23%) and health education (10.75%). Sentiment analysis reveals that positive emotions toward topics such as “medical staff” and “appointment registration,” while “referral” and “financial burden” elicits negative sentiments. Machine learning methods can effectively identify the shortcomings in chronic disease services.

〔Keywords〕 machine learning; text mining; chronic disease management

〔修回日期〕 2025 – 10 – 29

〔作者简介〕 徐若昕, 硕士研究生, 发表论文 1 篇; 通信作者: 刘静, 副研究员, 硕士生导师。

〔基金项目〕 国家自然科学基金青年项目 (项目编号: 72204153); 广东省卫生经济学会科研课题 (项目编号: 2024 – WJMF – 68); 韶关市社会发展科技协同创新体系建设项目 (项目编号: 230330118034783)。

优势与不足，为服务优化提供更可靠的循证依据。

1 引言

我国慢性病管理政策以分级诊疗制度为核心，通过强化基层医疗机构的服务功能，构建多层次协同体系^[1]。在此框架下，患者的健康管理需求驱动其在不同层级医疗机构间流转，形成包含基层首诊、双向转诊、上下联动、急慢分治的复杂服务接触网络。从患者视角出发，评价其对各服务接触点的满意度，有助于遵循“识别接触点-量化评价-优化改善”路径，分阶段指导慢性病管理改进，最终形成良性闭环^[2]。目前，关于慢性病患者评价的研究多聚焦理论框架构建^[3]和模型应用^[4]，多采用问卷、半结构化访谈等方法。这些方法主观性较强，难以捕捉患者未经引导的真实反馈，且分析大规模定性数据时效率较低。近年来，自然语言处理技术，特别是基于 Transformer 架构的预训练模型（如 BERT），为深度理解非结构化医疗文本提供了强大工具^[5]。Devlin J 等^[6]提出的 BERT 模型通过双向编码器捕捉文本上下文语义，在多项自然语言理解任务中取得了突破性进展。随后，Alsentrer E 等^[7]开发的 ClinicalBERT 验证了此类模型对临床专业领域文本（如电子健康记录）的有效性和适应性。自然语言技术的进步推动了患者生成内容相关研究，例如，Greaves F 等^[8]利用情感分析挖掘在线患者评论，证明从自由文本中量化患者体验的可行性。然而，既往研究多聚焦在线医疗平台上相对简短自发的评论，针对深度访谈获取的、具有上下文信息的慢性病患者体验文本研究仍不足。更重要的是，此类研究通常直接将通用模型应用于特定领域^[9]，忽略了医疗文本的专业性和场景复杂性，导致分析精度和深度受限。

为此，本研究结合专家知识与深度学习，先由医学专家团队基于访谈文本标注患者关注的核心维度，构建高质量监督学习数据集；再选择 BERT-base-Chinese 模型进行有监督微调，实现新语料的精准主题分类与情感分析。该方法旨在融合深度访谈的丰富性、专家知识的准确性与机器学习的高效性，从而以数据驱动的方式客观识别慢性病管理过程中的

2 资料与方法

2.1 资料来源

选取 3 所三级医院（广东省一所、重庆市两所）接收的社区卫生服务中心转诊慢性病患者，以及一所社区卫生服务中心（广东省）就诊的慢性病患者，进行深度访谈。纳入标准：符合《国际疾病分类第十一次修订本（ICD-11）》中高血压、糖尿病等慢性非传染性疾病的诊断标准，且病程达 1 年及以上；具备完整的语言认知功能，能够独立完成半结构化访谈；具有双向转诊医疗经历（社区首诊后经绿色通道转至三级医院，完成系统治疗后回归社区进行健康管理）。排除标准：合并心脑血管急性事件或糖尿病酮症酸中毒等严重并发症；近 3 个月参与其他干预性研究；电子健康档案数据完整性低于 85%。

为系统评估慢性病患者的全周期诊疗体验，采用半结构化访谈，核心维度包括：对医护人员技术与服务态度的感知（如您对本次接诊的医生是否满意？原因是什）；对就诊流程（挂号、检查、取药）效率的评价；对转诊经历顺畅度的反馈；对医疗经济负担与医保政策的看法；对健康教育与长期管理支持的评估以及对医院硬件设施与环境的评价。所有访谈均由统一培训的调查员执行，采用开放式提问，最大限度获取患者自发、真实的反馈，避免引导性偏差。

2.2 数据清洗与预处理

采用基于自然语言处理的自动化文本切片技术对半结构化访谈语料进行系统性处理。实验环境为：CPU（AMD Ryzen 9 8445HS）、GPU（NVIDIA GeForce GTX 4060 Ti, 16GB VRAM）。通过 Python 构建正则表达式匹配脚本，依据语音转写文本中的话轮转换标记（如“访谈人”和“患者”）实施初始语句分割，具体切分过程包括以下步骤。一是使用正则表达式模式：(r‘访谈人：|患者：’) 检测话轮边界，将连续文本分割为独立的发言片段。二是

基于内容分析排除基线资料提问语句,如涉及人口统计学或背景信息的提问,以避免噪声干扰。三是提取有效语料片段,确保每个片段代表一个完整的问答交互。四是通过自动配对相邻发言,形成结构化问答对,并以随机抽样的人工审核实施质量检查,以验证切分准确性和一致性。最终生成的结构化语料库包含 4 234 组高质量问答对。该处理方法能够提升非结构化文本数据的机器学习适配性,为后续分类模型的监督式训练与情感得分计算提供高质量数据集。

2.3 BERT 语言模型训练与聚类

在模型训练前,须构建监督学习所需的标注数据集。初步分析清洗后的文本,依据患者访谈中频繁提及的核心体验议题,确定文本分类的类别体系。例如,将患者表达的“医生解释病情很耐心”“护士一直在引导我”等积极反馈,以及“医生头都不抬”“多问几句就很不耐烦”等负面感受,归类为“医护人员服务态度”维度;将“转诊手续复杂不知道找谁”“社区医院和上级医院互相推诿”等典型抱怨,归类为“转诊流程”维度;将“看病花光了积蓄”“医保报销比例太低”等涉及费用的描述,归类为“经济负担”维度。为了确保患者体验分类的科学性与临床相关性,由 3 名研究人员(具有临床医学和公共卫生背景)组成专家小组,基于上述框架,并参考既往研究^[10-11]对随机抽取的 1 000 条文本样本进行人工标注。在标注过程中,通过多次讨论,修正歧义类别定义,达成一致,确立 11 个核心维度作为标注标准。该过程保证了输出结果的多样性和独特性,排除了与患者体验无关的内容。

采用 BERT - base - Chinese 预训练模型^[10],基于切片后的问答对微调 BERT 模型。BERT - base - Chinese 模型以 12 层 Transformer 架构平衡模型表达能力和计算效率。BERT 模型是基于 Transformer 架构的预训练语言模型,在自然语言处理任务,如文本分类(患者需求)和实体识别(药物种类)中具有优秀性能。在 BERT - base - Chinese 模型上添加分类层,并使用标注的中文文本数据集进行训练,

以适配特定分类任务^[6]。为减少过拟合并优化适应性,冻结嵌入层的前 8 层,保留预训练知识,仅重训练最后 4 层,专注于医学特征。为确保评估客观性,将标注数据($N=1\,000$)按 8:1:1 随机分为训练集、验证集和测试集。训练集用于参数迭代,验证集用于超参数调优和早停判断,测试集用于最终性能评估,所有结果在测试集进行报告。微调使用 AdamW 优化器,输出各类别概率,详细参数,见表 1。

表 1 微调模型参数

名称	值或说明
预训练模型	BERT - base - Chinese
标签类别数	11
隐藏层 dropout 概率	0.1
注意力层 dropout 概率	0.1
冻结层数	BERT 前 8 层
优化器	AdamW
BERT 学习率	1e - 5
分类器学习率	3e - 5
权重衰减	0.01
训练批大小	8
评估批大小	8
训练轮数	6
模型保存策略	每个 epoch 保存
训练结束加载最佳模型	是
最佳模型评价指标	加权 F1 分数
混合精度训练	是
数据加载进程数	0
是否移除未用列	否
标签列名	labels
warmup 比例	0.1
日志记录步数	25
梯度累积步数	2
学习率调度类型	线性
梯度检查点	是
早停策略	容忍 2 个 epoch 无提升

2.4 情感分析

聚类分析后,使用 BERT 架构的情感分析功能

对问答对进行细粒度情感分析，通过 Softmax 层输出情感极性（积极或消极）和连续情感强度值^[11]。模型校准遵循通用语言理解评估基准测试框架^[12]，并使用 5 折交叉验证确保鲁棒性，最终获得标准化情感强度，计算其中位数和平均值。而后采用 Mann-Whitney U 检验分析组间情感得分差异的统计学意义。

3 结果

3.1 描述性统计分析

本研究共收集 539 例慢性病患者的访谈资料，年龄范围为 23~88 岁，以 60~75 岁中老年人为主，男性占 54.6%，见表 2。

表 2 研究对象基线特征

类别	特征	数量 (例)	占比 (%)
性别	男性	294	54.6
	女性	245	45.4
学历	初中及以下	393	72.8
	中专、高中	115	21.4
	本科及以上	31	5.8
慢性病	高血压	369	68.4
	糖尿病	230	42.6
	冠心病	96	17.8
	其他	107	19.8

3.2 文本聚类结果

为评估方法有效性，对比微调后的 BERT 模型与两种经典无监督聚类算法（K-means、DBSCAN）。在相同数据集上，K-Means 算法类别语义一致性差，存在大量跨主题混合簇，如将“医护人员态度”与“医院环境”评价混为一谈；DBSCAN 算法因文本向量密度问题产生过多被判定为无法归类的噪声点，导致聚类结果可用性低。相比之下，微调后的 BERT 模型凭借强大语义理解能力，归类的 11 个类别与专家标注高度吻合，且内部主题鲜明、语义一致性强。具体性能对比，见表 3。

表 3 模型精度对比

模型	准确率	F1	召回率	轮廓系数	主题一致性
BERT-base-Chinese	0.775	0.704	0.793	0.32	0.62
K-means	0.613	0.541	0.640	0.15	0.38
DBSCAN	0.553	0.532	0.519	-0.27	0.22

共收集有效访谈文本 539 份，切片后得到问答对 4 836 对，数据清洗后有效问答对 4 234 对，有效率 87.55%。使用微调后的模型对所有有效问答对进行分类，结果显示：医院（包括设备、环境和医护人员）相关评价占 24.3%，患者医院选择倾向性相关占 37.72%（略高于其他类别），就诊流程相关占 27.23%，健康教育相关占 10.75%，见表 4。各类占比均衡，无极端结果。

表 4 文本聚类结果

主类别	子类别	数量 (对)
医护人员	服务态度、患者个性化定制	510
信任程度	医疗技术、其他	495
健康教育	管理合并症教育、信息获取、是否配合治疗	455
开药取药及挂号	开药取药、挂号、开具的检查	422
随访复查	电话随访、上门随访、复查	406
医院口碑	他人推荐、医院评价	395
经济负担	无	390
转诊	其他转诊、县级医院转诊、社区医院转诊	325
便利程度	病情严重程度、医保地点、就医距离	317
医院设备	精密仪器、医疗设备	272
医院环境	病房等基础设施、卫生环境	247

3.3 情感分析结果

情感分析获得每条患者反馈的情感极性得分（范围 [-1, 1]，值越大表示情感越积极）。各主题得分平均值均高于 0.54，显示整体评价积极正向。为评估患者体验一致性，计算标准差（standard deviation, SD）以衡量评分波动，同时考虑平均值

和中位数。标准差越小，共识度越高；标准差越大，意见分歧越显著。共识度判断标准： $SD < 0.20$ 为高共识度； $0.20 \leq SD < 0.30$ 为中等共识度； $0.30 \leq SD < 0.35$ 为低共识度； $SD \geq 0.35$ 为极低共识度。详细数据，见表 5。

表 5 各主题情感得分的描述性统计与共识度评估

主题类别	平均值	中位数	标准差	共识度评估
随访复查	0.59	0.59	0.14	高共识度
医护人员	0.61	0.61	0.15	高共识度
医院口碑	0.62	0.60	0.18	高共识度
便利程度	0.58	0.53	0.22	中等共识度
健康教育	0.61	0.54	0.25	中等共识度
开药取药及挂号	0.57	0.47	0.28	低共识度
医院设备	0.60	0.45	0.30	低共识度
医院环境	0.56	0.42	0.31	低共识度
转诊	0.57	0.41	0.32	低共识度
信任程度	0.54	0.34	0.35	极低共识度
经济负担	0.57	0.33	0.37	极低共识度

患者体验呈现高度共识与显著分歧并存。“医护人员”“随访复查”“医院口碑”主题得分最高且标准差最小 ($SD < 0.20$)，表明其服务体验获得高度一致认可，是当前服务体系中最可靠的支柱。相比之下，“经济负担” ($SD = 0.37$) 和“信任程度” ($SD = 0.35$) 标准差较高，达到极低共识度水平，表明反馈出现极化：积极评价拉高平均值，但大量患者体验集中于低分，突显服务或政策未均衡惠及所有人群，存在公平性和普适性挑战，应优先关注和精细管理。其余主题分布于中等及低共识度区间，表明服务体验的稳定性和一致性存在梯度差异。

3.4 组间情感得分差异的显著性检验

在 Mann - Whitney U 检验分析后得出：积极组包括“医护人员” (510 例, 55.67%) 和“随访复查” (406 例, 44.32%)，消极组包括“经济负担” (390 例, 44.07%) 和“信任程度” (495 例, 55.93%)。结果显示，积极组得分中位数 (0.60) 显著高于消极组 (0.35)，差异具有统计学意义 ($U = 205,500.5$,

$P < 0.001$)。

4 讨论

4.1 BERT 聚类分析揭示患者就医 3 大关切

聚类分析揭示患者对医疗服务质量的 3 大核心诉求：硬件设施先进性（如诊断精度）、治疗方案循证依据充分性、医护团队专业素养与服务态度。诊疗服务流程的时空配置效率问题突出，涉及预约挂号易用性、检查项目周转时效、分级诊疗衔接顺畅度和药品供应保障。患者评价存在矛盾：部分认可数字化流程便捷，部分批评冗余检查和等待时间。此外，在医疗信息素养维度，患者评价包括疾病知识获取渠道满意度、医嘱依从性教育有效性和数字化健康平台可及性。结果表明应构建覆盖诊前、诊中、诊后的连续性支持体系，验证了《中国防治慢性病中长期规划（2017—2025 年）》^[13] 中“健全治疗 - 康复 - 长期护理服务链”的必要性。

4.2 慢性病患者关注点：从医疗服务到全程管理

患者对核心环节（如医护人员、随访复查）的积极情感强度显著高于经济负担和信任程度。情感分析结果显示，“医护人员”与“随访复查”主题的情感得分中位数分别为 0.61 和 0.59，且标准差小于 0.20，表明患者对这些服务体验的认可具有高度共识。健康教育主题也显示积极情感（中位数 0.54），患者认同其价值并主动参与健康管理。这与陈菊红等^[14]研究结论相呼应，该研究表明，对妊娠期糖尿病孕妇实施结构化健康教育，能有效提升患者对疾病知识的掌握程度、治疗依从性，并最终改善母婴临床结局。印证了系统化、规范化的健康教育在慢性病管理中的关键价值。

相比之下，“转诊”与“经济负担”主题呈现显著消极情感，其中位数情感得分分别低至 0.41 和 0.33，且标准差高达 0.32 和 0.37，反映出患者在此类问题上的体验存在极大分歧与不满。李晓轩等^[15]指出这种差异可能与经济状况和医保满意度相关。其研究表明，仅 47.6% 的康复患者认可双向转诊，患者对基层医院环境设施的不满意度达

22.3%，对医疗费用及报销水平的不满意度为5.6%。多因素分析进一步揭示，患者年龄、学历、家庭年收入以及对下转流程的了解程度均是影响其转诊态度的关键因素。未来应着力提升基层服务水平，增强患者信任，优化绩效考核，减少趋高就医现象，以系统性降低患者经济负担。

4.3 机器学习与医疗管理的跨界融合

机器学习应用于慢性病患者评价分析涉及多学科领域，医务人员在其中起关键作用，但当前研究多由计算机领域主导，医疗领域参与不足。未来应发挥多学科优势，计算机研究者开发更准确高效的模型；医务人员应用专业知识解决慢性病管理效率低和健康负担重等问题。在人工智能与医疗融合，以及“健康中国2030”规划推动下，共同推进机器学习在慢性病管理中的应用，实现服务全覆盖。建议后续研究拓展多平台患者数据采集，为诊疗优化提供全面决策支持。

5 结语

本研究将医学专家知识与BERT深度学习模型相结合，对来自多中心的慢性病患者深度访谈文本进行主题识别与细粒度情感分析。该方法有效克服了传统调研的主观性与通用模型在专业领域精度不足的局限，不仅系统揭示了患者在硬件设施、治疗方案循证性及医护团队素养3大维度的核心诉求，还精准量化了其在服务体验中的共识与分歧，如“转诊”与“经济负担”主题所呈现的高分歧与消极情感。研究成果从患者视角，为构建“诊前-诊中-诊后”连续性支持体系提供了经验证据，验证了国家相关政策方向的必要性。

然而，本研究仍存在一定的局限性。样本地域覆盖相对集中，可能影响结论的普适性；受研究方法所限，重症监护患者及存在语言、认知障碍的群体未被纳入，可能导致样本对慢性病患者总体的代表性存在偏倚。基于以上局限与本次研究的经验，未来研究可着力于以下方向：一是开展跨区域、多中心合作，以验证并拓展当前发现；二是在伦理与

可行性前提下，开发适应性方法以纳入更广泛的患者群体；三是进一步探索领域自适应等前沿技术，训练医疗专用模型，持续提升对复杂医疗文本的理解与分析精度，从而为分级诊疗政策的精细化调整提供更强大的决策支持。

作者贡献：徐若昕负责实验实施、论文撰写；杨婷婷负责数据收集与整理；江宽列负责实验设计、参与论文撰写；刘静负责提供指导、论文审核；万金花、朱小娇负责问卷调研与数据收集。

利益声明：所有作者均声明不存在利益冲突。

参考文献

- 北京大学医学部血管健康研究中心，国家卫生健康委卫生发展研究中心，甘肃省第三人民医院，等. 慢病医防融合分级诊疗指南：T/CRHA 150—2025 [EB/OL]. [2025-07-16]. <https://ndls.org.cn/standard/detail/015a064220bfd1ca70ac6282f9951ac3>.
- 徐华，张海斌，高松，等. 患者评价对互联网医疗平台运营策略的影响 [J]. 中国卫生信息管理杂志，2022, 19 (6): 849-854.
- 张彭，谢洋，任嘉铭，等. 中医理论指导下患者评价量表的研制现状与思考 [J]. 中华中医药杂志，2024, 39 (8): 4419-4424.
- 王馨艺，郭静，李萍. 住院适宜性评价模型在神经内科住院患者中的应用 [J]. 中国社会医学杂志，2025, 42 (2): 226-231.
- GARDAZI N M, DAUD A, MALIK M K, et al. BERT applications in natural language processing: a review [J]. Artificial intelligence review, 2025, 58 (6): 166.
- DEVLIN J, CHANG M W, LEE K, et al. BERT: pre-training of deep bidirectional transformers for language understanding [C]. Minneapolis: The 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2019.
- ALSENTZER E, MURPHY J R, BOAG W, et al. Publicly available clinical BERT embeddings [EB/OL]. [2025-09-21]. <http://arxiv.org/abs/1904.03323>.
- GREAVES F, RAMIREZ-CANO D, MILLETT C, et al. Use of sentiment analysis for capturing patient experience from free-text comments posted online [J]. Journal of

- medical internet research, 2013, 15 (11): e239.
- 9 HE K, MAO R, LIN Q, et al. A survey of large language models for healthcare: from data, technology, and applications to accountability and ethics [EB/OL]. [2025-09-24]. <http://arxiv.org/abs/2310.05694>.
- 10 VANDENBERG S, HAIR H, HARVEY G, et al. The patient journey map: improving the emergency department communication experience for patients and their family and friends [EB/OL]. [2025-07-19]. <https://openurl.ebsco.com/contentitem/doi:10.29173%2Fcjen131?sid=ebsco:plink:crawler&id=ebsco:doi:10.29173%2Fcjen131>.
- 11 GAO Z, FENG A, SONG X, et al. Target-dependent sentiment classification with BERT [J]. IEEE access,
- 2019, 7 (10): 154290-154299.
- 12 XU L, HU H, ZHANG X, et al. CLUE: a Chinese language understanding evaluation benchmark [EB/OL]. [2025-07-19]. <http://arxiv.org/abs/2004.05986>.
- 13 中国防治慢性病中长期规划 (2017—2025 年) [EB/OL]. [2025-07-19]. https://www.gov.cn/zhengce/content/2017-02/14/content_5167886.htm.
- 14 陈菊红, 解红文, 阮永兰, 等. 妊娠期糖尿病孕妇线上线下结构化健康教育研究 [J]. 护理学杂志, 2020, 35 (11): 5-8.
- 15 李晓轩, 叶睿雪, 王昱棋, 等. 分级诊疗背景下康复患者双向转诊现状及对策研究 [J]. 卫生软科学, 2024, 38 (10): 84-88, 94.

(上接第 19 页)

- 10 CHOWDHERRY A, NARANG S, DEVLIN J, et al. Palm: scaling language modeling with pathways [J]. Journal of machine learning research, 2023, 24 (240): 1-113.
- 11 王文奇, 郭梦帆, 杨杜祥, 等. 大语言模型发展与应用综述 [J]. 中原工学院学报, 2025, 36 (2): 1-8.
- 12 SOLARES J R A, RAIMONDI F E D, ZHU Y, et al. Deep learning for electronic health records: a comparative review of multiple deep neural architectures [J]. Journal of biomedical informatics, 2020, 101 (1): 103337.
- 13 孙磊, 汪安安, 宋一敏, 等. 大语言模型在临床医学领域的应用、挑战和展望 [J]. 解放军医学院学报, 2025, 46 (1): 50-60.
- 14 ZHANG H, SONG H, LI S, et al. A survey of controllable text generation using transformer-based pre-trained language models [J]. ACM computing surveys, 2023, 56 (3): 1-37.
- 15 JIA M, DUAN J, SONG Y, et al. mediKAL: integrating knowledge graphs as assistants of LLMs for enhanced clinical diagnosis on EMRs [EB/OL]. [2025-02-22]. <https://arxiv.org/abs/2406.14326>.
- 16 王博, 于志昊, 张军雁, 等. 基于电子病历数据和知识增强的医疗大语言模型构建方法研究 [J]. 解放军医学院学报, 2025, 46 (1): 97-103, 119.
- 17 李佳林, 郜斌宇, 陈卉. 基于大语言模型的电子病历文本结构化 [J]. 中国医疗设备, 2025, 40 (5): 42-46, 52.

(上接第 34 页)

- 7 安泽心, 杜静. 融合轻量化区块链的医疗信息安全网络架构设计 [J]. 信息技术, 2025 (2): 28-33.
- 8 陈英杰, 沈济南, 梁芳, 等. 医疗云环境下访问控制增强模型 [J]. 郑州大学学报 (理学版), 2022, 54 (5): 49-56.
- 9 盛朝阳. 基于智能合约的医疗数据访问控制方法研究 [D]. 郑州: 郑州大学, 2022.
- 10 上海市卫生健康委员会. 健康上海行动 (2019—2030 年) [EB/OL]. [2024-09-15]. <https://wsjkw.sh.gov.cn/agwsjkjcj2/20190916/0012-65223.html>.
- 11 江苏省人民政府. 落实健康中国行动 推进健康江苏建设实施方案 [EB/OL]. [2024-09-15]. https://www.jiangsu.gov.cn/art/2020/2/17/art_64797_8974642.html.
- 12 浙江省人民政府. 关于推进健康浙江行动的实施意见 [EB/OL]. [2024-09-15]. https://www.zj.gov.cn/art/2022/2/16/art_1229019364_2392893.html.
- 13 盛姝, 黄奇, 杨洋, 等. HL7 FHIR 框架下中国医疗领域信息交换研究与解决方案 [J]. 数据分析与知识发现, 2021, 5 (11): 13-28.
- 14 陈嘉莉, 马自强, 苗莉, 等. 基于 Hyperledger Fabric 的电子病历共享方案 [J]. 通信学报, 2024, 45 (S1): 60-74.