

基于 LLM 的新疆多语种医院智能客服系统构建

杨金平¹, 何方劲¹, 徐海峰², 杨鹏程^{2*}

(1.中国人民解放军陆军第九四七医院信息科, 新疆 喀什 844000)

(2.新疆军区总医院信息科, 新疆 乌鲁木齐 830001)

摘要 **目的/意义**针对新疆多民族聚居区语言障碍导致的就医流程复杂、导诊效率低等问题, 构建高效、公平的本地化智能客服系统, 提升少数民族患者就医体验与服务可及性。**方法/过程**基于 Qwen3 大模型与 RAG 技术, 构建汉、维、哈三语交互系统, 集成多语种语音交互与 RAG 检索, 采用轻量化部署以适应边疆网络条件, 并在新疆某三甲医院试点三个月。**结果/结论**预约挂号成功率由 60%提升至 92%, 平均响应时间从 8min 缩短至 15s, 多语种用户满意度达 91%。该系统有效缓解语言壁垒, 显著提升医院服务效率与患者满意度, 为人工智能赋能民族地区医疗公平提供可复制的技术范式。

关键词: 大语言模型; 多语种; 医院智能客服

Construction of a Multilingual Hospital Intelligent Customer Service System for Xinjiang Based on Large Language Models

Jinping Yang¹, Fangjin He¹, Haifeng Xu², Pengcheng Yang^{2*}

(The 947th Hospital of PLA, Information Technology Department
,Kashi,Xinjiang,844000,China)

Information Department, General Hospital of Xinjiang Military Area Command, Urumqi
830001, China

【Abstract】 Objective/Significance: To address the issues of complex medical procedures and low triage efficiency caused by language barriers in multi-ethnic regions of Xinjiang, this study develops an efficient and equitable localized intelligent customer service system, aiming to improve healthcare accessibility and the medical experience for ethnic minority patients.**Methods/Process:** Based on the Qwen3 large language model and retrieval-augmented generation (RAG) technology, a trilingual interactive system (Chinese, Uyghur, and Kazakh) was constructed, integrating multilingual voice interaction and RAG-based knowledge retrieval. The system adopts lightweight deployment strategies to adapt to the limited network conditions in border areas and was piloted for three months at a tertiary hospital in Xinjiang.**Results/Conclusion:** The appointment registration success rate increased from 60% to 92%, the average response time decreased from 8 minutes to 15 seconds, and user satisfaction among multilingual patients reached 91%. The system effectively alleviates language barriers, significantly enhances hospital service efficiency and patient satisfaction, and provides a replicable technical paradigm for leveraging artificial intelligence to promote healthcare equity in ethnic regions.

Keywords: large language model; multilingual; hospital intelligent customer service

1 引言

新疆维吾尔自治区是我国多民族聚居的边疆地区，尤其南疆部分地区少数民族人口比例较高，部分中老年患者以民族语言为主要交流方式，语言障碍易影响病情描述、医患沟通及医疗信息理解，降低就医效率和服务体验。根据新疆维吾尔自治区卫生健康委员会2025年发布的相关调查报告，南疆地区仍有较高比例少数民族患者受到语言障碍影响。因此，提升医疗服务多语言保障能力，是推进民族地区智慧医院建设的重要需求。

目前，医院患者咨询、导诊等服务主要依赖人工服务，多语种服务需要配备具备民族语言能力的工作人员，不仅人力成本较高，而且服务质量易受人员数量和专业水平影响，高峰时段还容易出现沟通效率低、挂号错误、重复排队等问题。随着人工智能、大语言模型及检索增强生成（Retrieval-Augmented Generation, RAG）等技术的发展，智能客服系统在自然语言理解、多轮对话和知识问答等方面取得了显著进展，为解决医疗场景中的语言沟通问题提供了新的技术路径^[1]。然而，现有医疗智能客服系统大多面向汉语应用场景，对维吾尔语、哈萨克语等少数民族语言支持不足，难以满足新疆地区多语种医疗服务需求。

因此，本研究构建一种面向新疆多语种环境的医院智能客服系统。以 Qwen3 为核心，融合 RAG 技术构建汉语、维吾尔语和哈萨克语多语言知识库，结合多语言语音识别与语音合成技术，实现多语种语音交互和智能问答，并与医院业务系统集成，为患者提供智能导诊、挂号预约、政策咨询和健康宣教等服务。

2 系统需求分析

提取分析医院近 10 年的门诊、住院患者数据发现，维吾尔族患者占就诊总人数的 62.9%。受医保差异化报销政策影响，基层医疗机构报销比例高于上级医疗机构，未按规定逐级转诊至高等级医疗机构就诊的患者医保报销比例相应降低，因此大量维吾尔族患者首先在基层医疗机构就诊^[2]。二级以下医院维吾尔族患者占比较高，2000 年以前出生患者日常交流多使用维吾尔语，医院线上咨询中维吾尔语使用比例也较高。为进一步了解患者实际需求，对 200 名少数民族患者进行了结构化访谈，结果显示，70.5% 的患者认为就医过程中存在挂号、导诊及医患沟通困难，主要原因是语言障碍导致症状表达不准确、科室选择错误及就诊流程不熟悉；另有部分患者反映，对医保政策、费用构成及检查流程了解不足，影响了就医效率和满意度。因此系统需满足多语种医疗咨询、智能导诊、医保政策查询及信息服务等需求，同时兼顾系统性能、安全性和可维护性。

2.1 功能需求

（1）多语种智能导诊与健康宣教。系统应支持汉语、维吾尔语和哈萨克语三种语言交互，根据患者描述的症状推荐相应就诊科室，并提供疾病预防、合理用药、术后护理等健康宣教内容，帮助患者提高就医效率和健康管理能力。

（2）医疗信息与医保政策咨询。系统应支持医保报销政策、费用构成、检查报告及就诊流程等信息查询，为患者提供统一的信息服务入口，减少重复咨询，提高医疗信息获取效率。

（3）知识库管理与人工协同。系统应支持医疗知识、药品信息及医保政策的动态更新，保证知识内容的准确性和时效性；对于复杂咨询或高风险问题，应能够及时转接人工客服，保障医疗服务质量。

（4）服务管理与数据分析。系统应具备患者意见反馈、工单管理及咨询数据统计分析能力，为医院优化导诊流程、完善医疗服务和辅助管理决策提供数据支撑。

2.2 非功能需求

系统应满足医院日常业务高并发访问需求，保证文本问答和语音交互的实时响应，提高维吾尔语、哈萨克语等少数民族语言的识别与理解准确率，确保导诊及咨询结果的可靠性^[3]。同时，系统应采用适老化交互设计，支持语音输入、大字体显示等功能，方便老年患者使用；严格落实患者隐私保护要求，对敏感医疗数据进行安全管理，并具备良好的可扩展性和可维护性，以适应医院业务流程及医保政策的持续更新^[4]。

3 系统架构设计

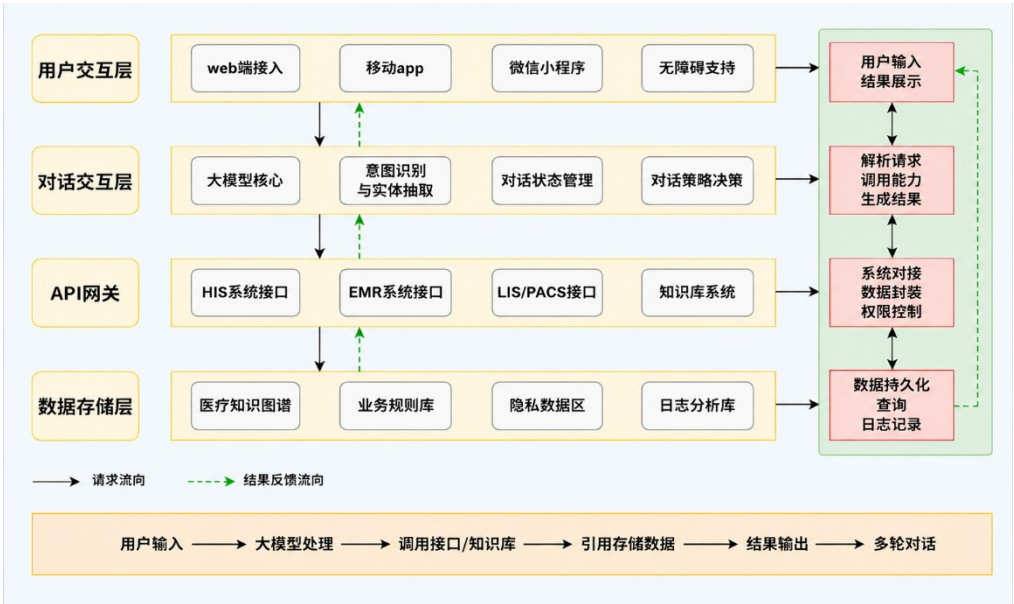


图 1 系统架构设计图

如图 1 所示，系统采用分层微服务架构，兼顾高可用性、多语种适配性与医疗场景安全性，整体分为四层：用户交互层、对话交互层、API 网关层与数据存储层。各模块松耦合部署，支持横向扩展，可在边疆地区有限算力条件下通过模型量化与缓存策略实现轻量化运行。各层具体设计如下：

3.1 用户交互层

系统支持 Web 端、移动 App 及微信小程序等多种接入方式，为患者提供统一的智能客服入口。界面采用适老化设计，支持大字体、语音输入与语音播报等功能，提高老年患者的使用便利性。针对新疆多民族地区应用需求，系统支持汉语、维吾尔语和哈萨克语三种语言自由切换，并兼容维吾尔语从右向左的书写方式，为不同民族患者提供一致的交互体验。

3.2 对话交互引擎

对话交互引擎是系统的核心模块，以大语言模型为基础，结合多语言自然语言理解技术，实现患者意图识别、实体抽取、多轮对话管理及智能回复。针对维吾尔语、哈萨克语等低资源语言，结合本地医疗语料，提高医疗术语识别和跨语言语义理解能力。在回答生成过程中，引入上下文管理机制，实现连续对话状态保持，并结合 RAG 检索结果生成准确、可追溯的医疗咨询回复，有效降低大模型幻觉，提高导诊、医保咨询和健康宣教等场景的回答准确率^[5]。

3.3 API 网关集成

系统采用统一 API 网关实现与医院 HIS、EMR、LIS、PACS 等业务系统的数据交互，统一完成身份认证、接口调用和权限控制。通过标准接口获取挂号信息、检查报告、费用

10.12201/bmr.202609.00005V1

信息及电子病历摘要，并结合多语言知识库，为不同语言用户提供一致、准确的医疗咨询服务，实现业务数据与智能客服的深度融合。

3.4 数据存储

3.4.1 多语言医疗知识库

针对汉语、维吾尔语和哈萨克语语义差异较大的特点，系统采用“独立知识库+跨语言语义映射”策略构建多语言医疗知识库。以统一医学标准术语作为语义锚点，实现三种语言医学实体的语义关联，在保持各民族语言表达习惯的同时，保证不同语言查询结果的一致性，为 RAG 检索提供统一知识支撑。

3.4.2 向量数据库

系统构建多语言向量数据库存储医疗指南、药品说明书及医保政策等非结构化知识，并采用基于语义的动态文本切分方法生成向量索引。在检索过程中，根据用户问题进行语义匹配，召回最相关知识片段，为大语言模型提供事实依据，提高医疗问答的准确率和可解释性。

3.4.3 业务规则与安全管理

依据《个人信息保护法》将数据划分为 L1（公开）至 L4（高度敏感）四个等级：L1、L2 级数据采用库表级透明加密及基础访问控制；L3、L4 级敏感字段实施增强型字段级保护。加密采用 AES-256-GCM 认证加密算法，每条记录独立生成 12 字节随机 IV，将密文、IV 及 Tag 一并入库。架构上采用信封加密，由独立 KMS 集中管理数据密钥（DEK）并执行月度轮换，解密时严格依据数据绑定的密钥版本获取对应密钥，所有加解密操作均记录独立审计日志^[6]。查询结果在执行脱敏后再落盘至日志，全链路贯彻“数据可用不可见”原则^[7,8]。

4 系统实现

4.1 多语种支持实现

针对新疆多民族医疗服务场景，系统采用 Qwen3 大语言模型构建多语种交互能力，支持汉语、维吾尔语和哈萨克语三种语言的原生理解与生成，避免传统“翻译——处理——回译”模式造成的语义损失^[9]。并针对维吾尔语、哈萨克语词形变化特点进行优化，提高少数民族语言的医疗术语识别和语义理解能力。

系统支持文本、语音等多种交互方式，能够自动识别用户语言并完成相应的语音识别、语义理解和语音合成，实现跨语言一致的交互体验。同时，构建统一的多语言医疗知识库对疾病、药品、检查项目及医保政策等知识进行统一管理，保证不同语言下咨询结果的一致性。系统测试结果表明，汉语、维吾尔语和哈萨克语三种语言的问答准确率均达到 90% 以上，满足医院实际应用需求。

4.2 知识库构建与 RAG 技术

为系统性抑制大模型“幻觉”，构建 RAG（Retrieval Augmented Generation）架构，实现“事实驱动生成”^[10]。

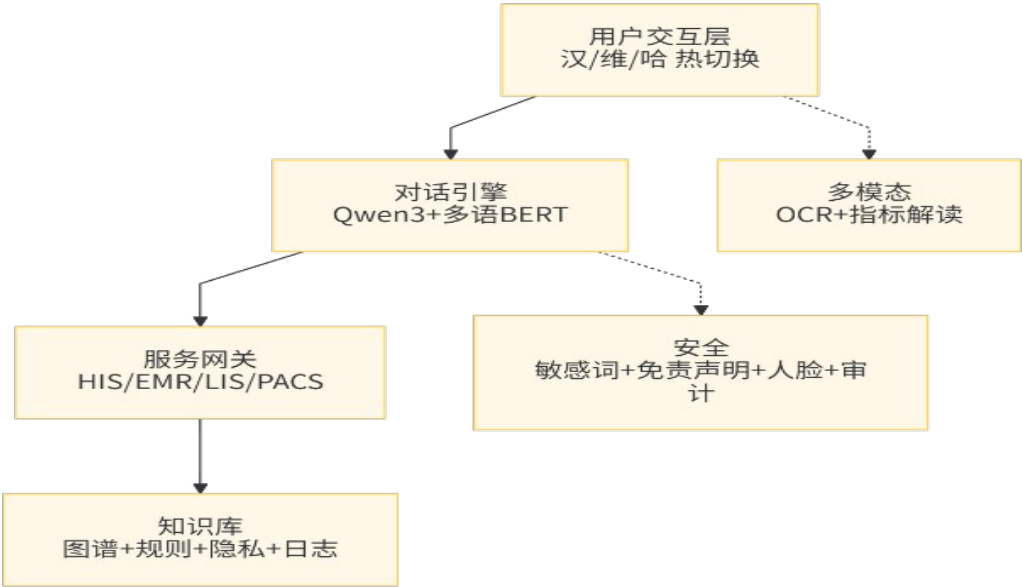


图 2 客服系统编排设计图

如图 2 所示，整个流程以用户交互层为入口，支持中文、维吾尔语、哈萨克语的实时语言切换，满足边疆地区多元语言需求。用户输入（文本或图像）首先由对话交互（Qwen3+多语 BERT）进行意图理解和查询解析；若包含化验单、病历图片等，则通过 OCR 模块提取文本，再由医学指标解读模块转化为结构化临床数据^[11]。

系统通过 RAG 机制有效抑制大模型的“幻觉”问题，避免了对海量医学语料进行昂贵且耗时的全量训练^[12]。在用户提问时，系统动态检索底层知识库——该知识库由医学知识图谱、临床诊疗规则、药品说明书、权威指南等结构化与非结构化数据生成字典组合，并经过隐私脱敏和版本管理。

构建了多语言医学知识库（支持汉语、维吾尔语、哈萨克语三种语言）为 RAG 提供检索支撑。在回答生成过程中，系统通过 RAG 关联并注入相关语料片段作为上下文，引导 Qwen3 生成回答。模型生成的回答中，约 60%的内容直接来源于 RAG 检索到的权威语料引用，而非模型参数自主生成，确保每一条临床建议均可追溯至原始医学文献或诊疗规范^[13]，为验证 RAG 机制对系统性能的实际提升效果，研究设计了专门的消融实验，相关量化评估结果详见 5.2 节。

为保障安全合规，所有请求经服务网关与医院 HIS/EMR/LIS/PACS 系统对接，在权限控制下获取患者个性化信息（如历史病历），同时通过敏感词过滤、人脸识别、操作审计和免责声明四重机制确保数据安全与法律合规。整个架构实现了“多模态输入→语义理解→知识检索→安全增强→多语言生成”的闭环^[14]，是 RAG 技术在垂直医疗领域落地的典型范式。

4.3 安全防护机制

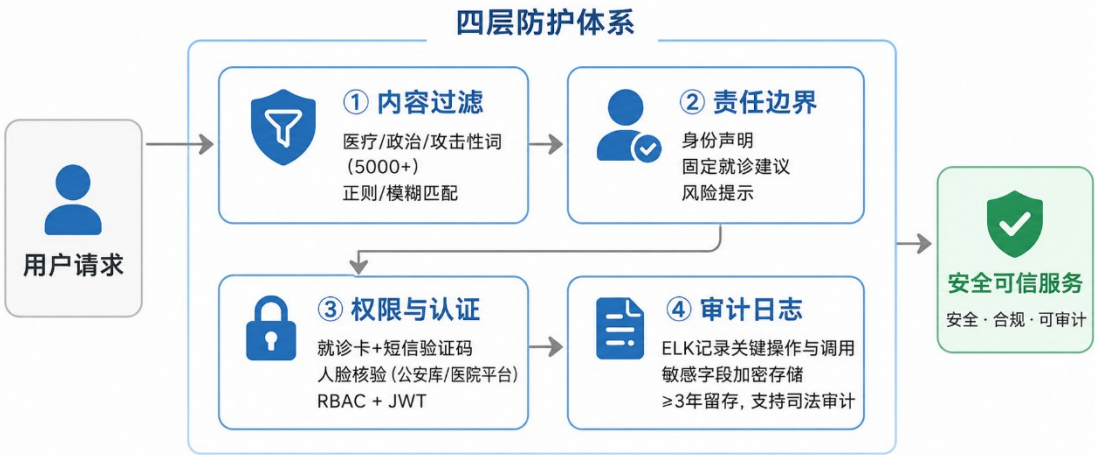


图 3 安全防护机制流程图

为确保系统安全性与合规性，构建“四层防护”体系。内容过滤层集成医疗、政治及攻击性敏感词库，过滤不当内容。责任边界层在对话关键节点插入非执业医师身份声明，对疾病诊断、用药咨询等高风险问题固定回复格式并附加风险提示，避免越界建议。权限认证层采用就诊卡绑定与手机号验证，敏感操作触发活体人脸识别；基于 RBAC 模型严格限定患者仅可访问本人数据，API 调用全程鉴权^[15,16]。操作审计层记录用户 ID、时间戳、问题、检索知识、生成回答及操作结果等全链路日志，患者信息字段加密存储，留存期不少于 3 年，满足《数据安全法》《个人信息保护法》及司法审计要求^[17]。

4.4 多模态交互扩展

为提升服务的包容性与智能化水平，设计了多模态交互接口，支持图像输入。技术架构包括图像输入接口，支持 JPG、PNG、PDF 格式，最大 10MB，前端自动进行压缩与旋转校正。OCR 识别方面，采用 PaddleOCR，识别内容包括检查项目名称（如“血常规”）、指标项（如“白细胞计数”）、数值、单位及参考范围^[18]。关键信息抽取通过命名实体识别（NER）模型识别“指标数值单位”三元组，例如从“白细胞：12.5×10⁹/L”中提取。生成解读报告时，将结构化数据输入大模型，结合 RAG 中的“检验指标参考范围知识库”生成解读，示例输出为：“检查项目：血常规，白细胞计数：12.5×10⁹/L（参考范围：4.0-10.0），提示：数值偏高，常见于细菌感染、炎症等。建议结合临床症状，咨询医生进一步评估。”输出形式支持 Markdown 和 HTML，便于在 App 中展示。在安全控制方面，图像上传后立即脱敏^[19]，仅保留核心数据，OCR 结果与原始图像分离存储，降低泄露风险^[20]。

5 应用场景与实验验证

5.1 实验设计

为验证系统在真实医疗场景中的性能与可用性，构建导诊咨询、检验报告解读及医保政策咨询 3 类典型应用场景开展实验。测试数据来源于医院实际业务，经脱敏处理后形成标准测试集，共包含 600 组问答样本，其中汉语 300 组、维吾尔语 200 组、哈萨克语 100 组，样本比例约为 3:2:1。

实验邀请 3 名临床主治医师、2 名医学信息学专家、2 名少数民族语言专家及 1 名医保政策专员组成专家组，对系统回答进行独立评分。评价指标包括回答准确率

（Accuracy）、幻觉率（Hallucination Rate）、平均响应时间（Latency）及专家综合评分。其中回答质量采用四级加权评价方法，将回答划分为致命错误、严重错误、轻微错误和正确回答四个等级，对应权重分别设为 0、0.5、0.8 和 1.0，综合准确率按下式计算：

$$Accuracy_w = \frac{\sum_{i=1}^n Score_i \times Weight_i}{\sum_{i=1}^n Weight_i}$$

式中，Score 表示单条回答评分，Weight 表示对应权重。

为验证 RAG 技术对医疗问答性能的影响，分别构建 Qwen3（无 RAG）和 Qwen3+RAG 两种实验配置，在相同测试集上开展对比实验。

5.2 RAG 消融实验

为评估检索增强生成（Retrieval-Augmented Generation，RAG）机制对医疗智能客服性能的提升效果，在保持模型参数一致的前提下，对比 Qwen3 原始模型与 Qwen3+RAG 模型的表现。RAG 知识库包含医院诊疗规范、医保政策、药品说明书、临床指南及多语言医学知识，共计约 12 万条知识片段，并采用 Milvus 构建向量索引。实验结果如表 1 所示。

表 1 RAG 消融实验结果

指标	Qwen3（无 RAG）	Qwen3+RAG	提升幅度
综合回答准确率	82.4%	94.7%	+12.3%
导诊准确率	86.2%	94.5%	+8.3%
报告解读准确率	89.6%	97.8%	+8.2%
医保政策准确率	84.8%	97.0%	+12.2%
幻觉率	18.5%	7.1%	-11.4%
平均响应时间	13.8 s	15.2 s	+1.4 s
专家综合评分	4.18	4.72	+12.9%

实验结果表明，引入 RAG 机制后，系统综合回答准确率由 82.4%提高至 94.7%，提升 12.3 个百分点；幻觉率由 18.5%下降至 7.1%，降低约 61.6%。其中，医保政策咨询场景提升最为明显，这是由于 RAG 能够实时检索最新政策文件作为生成依据，有效降低模型依赖参数记忆产生的错误回答。平均响应时间仍控制在约 15 s，能够满足医院在线咨询实时响应需求。

5.3 多语言性能对比实验

针对新疆地区多民族医疗服务需求，对汉语、维吾尔语及哈萨克语三种语言进行了性能测试。评价指标包括意图识别准确率、回答准确率、平均响应时间及用户满意度。

表 2 多语言性能比较

指标	汉语	维吾尔语	哈萨克语
意图识别准确率	96.8%	92.4%	90.8%
回答准确率	96.3%	91.2%	90.5%
平均响应时间(s)	13.8	15.6	15.9
用户满意度	94.5%	91.3%	90.8%

实验结果表明，汉语场景整体性能最佳，回答准确率达到 96.3%。维吾尔语与哈萨克

10.12201/bmr.202609.00005V1

语场景准确率分别达到 **91.2%**和 **90.5%**，能够满足医院日常导诊及咨询需求。少数民族语言准确率略低于汉语，主要原因在于医疗专业语料相对有限，部分医学术语存在同义表达及方言差异。

5.4 专家评测

基于上述测试集，进一步开展真实医疗场景专家评测，对导诊咨询、报告查询及医保政策咨询三类业务进行综合评价。

表 3 三类典型场景专家评测结果

应用场景	综合准确率	专家评价
导诊咨询	91.0%	可用
报告查询	96.2%	高可信
医保政策咨询	95.0%	高可信
综合平均	93.9%	高可信

测结果显示，报告查询场景表现最佳，加权准确率达到 **96.2%**，OCR 识别结合 RAG 知识检索能够有效提高检验指标解读的准确性；医保政策咨询场景准确率达到 **95.0%**，能够准确完成异地医保备案、报销流程及政策查询；导诊咨询综合准确率为 **91.0%**，主要误差来源于复杂症状涉及多科室交叉判断及少数民族医学术语标准化不足。总体来看，系统已达到医院智能客服实际部署要求。

6 系统应用效果与讨论

6.1 医院运行效果分析

系统于 **2025 年 8 月至 11 月**在新疆某三甲医院开展为期 **3 个月**的试点运行，共服务患者 **3 万余**人次，其中少数民族患者约 **2 万人**，累计完成医疗咨询 **3 万余**次，问题解决率达到 **93.3%**。

试点结果表明，系统显著提升了医院咨询服务效率。患者咨询平均等待时间由人工服务模式的 **8 min** 缩短至 **15 s**；系统日均处理咨询量由 **200 余**次提高至 **500 余**次，相当于约 **5 名**专职导医的工作量。与此同时，就诊预约成功率由 **60%**提升至 **92%**，导医台及护士站约 **70%**的基础咨询得到有效分流，门诊医生日均接诊效率提高 **18%**。

在医院运营方面，系统承担了大量重复性咨询工作，基础咨询岗位需求减少约 **60%**，预计每年可节约人工运营成本约 **30 万元**，体现出较好的经济效益和管理价值。

在用户体验方面，随机发放满意度调查问卷 **2147 份**，从响应速度、回答准确性、多语言交互体验及整体满意度 **4 个**维度开展评价。结果表明，多语言交互功能获得最高评价，患者整体满意度达 **91%**，说明系统有效缓解了少数民族患者的语言障碍。

6.2 多民族医疗服务改善效果

针对新疆多民族、多语言医疗服务需求，系统支持汉语、维吾尔语和哈萨克语三语智能交互，有效缓解了少数民族患者就医过程中的语言障碍。

试点期间，少数民族患者因等待人工翻译导致的平均等待时间由 **47 min** 缩短至 **8 min**，患者平均就诊总耗时由 **156 min** 缩短至 **102 min**，重复咨询率下降 **72%**。

同时，系统能够提供多语言导诊、医保政策咨询及健康宣教等服务，提高了患者获取医疗信息的便利性，改善了少数民族患者的整体就医体验，为边疆地区智慧医疗建设提供了新的服务模式。

6.3 系统应用价值分析

实验结果及医院试点运行情况表明，系统在服务效率、回答准确率及用户体验等方面

10.12201/bmr.202609.00005V1

均取得了较好的应用效果。依托 Qwen3 大语言模型、RAG 检索增强技术及多语言医疗知识库，系统实现了医疗知识的精准检索与多语言智能生成，有效降低了大语言模型幻觉问题，提高了医疗咨询结果的可信度。

此外，系统采用统一 API 网关与医院 HIS、EMR、LIS 等业务系统集成，实现了医疗数据共享和业务流程协同，为智慧医院智能客服建设提供了一种可复制、可推广的技术方案。对于边疆民族地区而言，该系统不仅能够提升医疗服务公平性，还能够降低人工服务压力，提高医院整体运行效率。

7 结论

本研究构建并验证了面向边疆多民族地区的医院智能客服系统，提出“多语种原生支持+本地化医疗知识融合+服务流程深度嵌入”的技术范式，有效缓解语言壁垒与医患信息不对称问题，显著提升了服务效率与患者体验，为人工智能促进民族地区医疗公平提供了可行路径。未来，系统将进一步升级为覆盖全疆、贯通三级医疗机构的数字健康基础设施，为铸牢中华民族共同体意识下的边疆健康治理提供技术支撑。

作者贡献： 杨金平负责研究设计、实验结果分析、论文撰写；何方劲负责研究方案完善、数据集构建、实验实施；徐海峰，杨鹏程负责研究方案完善、论文修订。

利益声明： 所有作者均声明不存在利益冲突。

参考文献：

[1] 陈紫林, 祝帆帆, 罗宇昕, et al. 大语言模型在医疗健康领域的应用现状与前景展望[J]. 医学与哲学, 2025, 46(12): 32-37.

[2] 刘伟, 张远华. 分级诊疗导向下基层首诊模式比较研究[J]. 卫生经济研究, 2026, 43(2): 68-71.

[3] 余杰, 飞龙, 郭陆祥, et al. 基于通用大模型的民族语言大模型构建技术[J]. 中文信息学报, 2025, 39(8): 75-81.

[4] 莫琳芳, 李喆, 甘辉亮, et al. 全球视野下医疗人工智能中患者隐私和数据安全: 焦点与策略[J]. 海军军医大学学报, 2025, 46(8): 989-999.

[5] 徐月梅, 胡玲, 赵佳艺, et al. 大语言模型与多语言智能的研究进展与启示[J]. 计算机应用, 2023, 43(S2): 1-8.

[6] 齐义文, 乔鑫鑫, 邓飞其, et al. 基于混沌数据加密的切换系统最优控制与隐私保护[J]. 自动化学报 2026, 52(7): 1-12.

[7] 李安全, 祝顺东. 云计算时代下数字医院的信息安全[J]. 网络安全技术与应用, 2025(8): 136-138.

[8] Venkataradhakrishnamurthy P, Malathi K. A holistic framework for strengthening security of healthcare data through encryption utilizing blockchain technology[J]. Scientific reports, 2025, 16(1): 2008-2008.

[9] 杨滨瑕, 罗旭东, 孙凯丽. 基于预训练语言模型的机器翻译最新进展[J]. 计算机科学, 2024, 51(S1): 50-57.

[10] 徐月梅, 叶宇齐, 何雪怡. 大语言模型的偏见挑战: 识别、评估与去除[J]. 计算机应用, 2025, 45(3): 697-708.

[11] 范俊军, 刘贤嫻. 我国民族语言文献文本数字化识别问题——基于 OCR 及其工具[J]. 暨南学报(哲学社会科学版), 2024, 46(6): 31-45.

[12] 程平, 王骥菱, 黄靖川, et al. RAG 与 DeepSeek 融合的企业信息智能核查研究[J]. 中国注册会计师, 2025(9): 72-78.

[13] Aljohani B, Alsanoosy T. Enhancing Medical Question Answering with LLMs via a Hybrid Retrieval-Augmented Generation Framework[J]. Information, 2026, 17(2): 133-133.

[14] 胡健, 刘军, 刘华栋. 大数据环境下隐私保护数据脱敏对策研究[J]. 信息与电脑, 2026, 38(9): 39-41.

[15] Atlam H F, Yang Y. Enhancing Healthcare Security: A Unified RBAC and ABAC Risk-Aware Access Control Approach[J]. Future Internet, 2025, 17(6): 262-262.

[16] 魏郅华. 基于 RBAC 安全模型的系统用户权限设计[J]. 软件, 2025, 46(1): 157-159.

[17] 兰浩维, 吴泽宇. 医院信息化建设中网络安全防护措施研究[J]. 科技创新与应用, 2025, 15(21): 141-144.

- [18] 郑宏坤. 基于 PaddleOCR 的图像识别技术在疫情防控中的应用与实践[J]. 数字技术与应用, 2022, 40(12): 117-119.
- [19] Katta B. AI-Powered Data Governance: Advancing Privacy and Compliance Through Data Masking and Sentiment Analysis[J]. Journal of Computer Science and Technology Studies, 2025, 7(3): 837-844.
- [20] Li Z, Guo B, Mu D. Research on a Method for Recognizing Text on Book Spines in Libraries Based on Improved YOLOv11 and Optimized PaddleOCR[J]. Electronics, 2025, 14(23): 4689-4689.