

T/HNBX

海南省标准化协会团体标准

T/HNBX XXXX—XXXX

T/CIET XXXX—XXXX

基于人工智能的病历辅助生成技术规范

Technical specification for artificial intelligence-based assisted generation of medical records

(征求意见稿)

XXXX - XX - XX 发布

XXXX - XX - XX 实施

海南省标准化协会
中国国际经济技术合作促进会
发布

目 次

前言	错误! 未定义书签。
1 范围	错误! 未定义书签。
2 规范性引用文件	1
3 术语和定义	1
4 核心技术	1
5 应用场景	4
6 系统安全与合规	5
7 运行与维护	7

前 言

本文件按照 GB/T 1.1—2020 《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由中国国际经济技术合作促进会标准化工作委员会提出。

本文件由海南省标准化协会、中国国际经济技术合作促进会共同归口。

本文件起草单位：

本文件主要起草人：

基于人工智能的病历辅助生成技术规范

1 范围

本文件规定了基于人工智能的病历辅助生成的核心技术、应用场景、系统安全与合规、运行与维护。本文件适用于医疗机构中AI驱动的病历生成系统研发、部署与应用。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 20988—2025 网络安全技术 信息系统灾难恢复规范

GB/T 35273—2020 信息安全技术 个人信息安全规范

《电子病历应用管理规范（2025修订版）》

《病历书写基本规范》

《疾病分类与代码》

3 术语和定义

本文件没有需要界定的术语和定义。

4 硬件配置要求

4.1 终端采集设备

4.1.1 语音采集：配备 ≥ 4 麦阵列的采集设备，支持硬件降噪，适配门诊、ICU等不同场景，移动采集设备需支持蓝牙/离线录制等多种方式，续航 ≥ 8 小时。

4.1.2 移动终端：查房终端配置 ≥ 8 GB运行内存，高清麦克风，支持Wi-Fi 6/4G/5G，离线缓存及医用级消毒。

4.1.3 影像设备：对接医院现有DR/CT/MRI等设备，支持DICOM 3.0标准，手术室/检查室配置 ≥ 4 K医用显示屏，响应延迟 ≤ 200 ms。

注：硬件设备的配置可根据实际情况进行调整。

4.2 服务端与存储设备

4.2.1 服务端：核心服务器配置多核高性能CPU、专用AI推理GPU，支持集群化部署和双机热备，算力满足多任务并发推理需求。

4.2.2 存储设备：可采用分布式存储架构，总有效容量 ≥ 100 TB且支持弹性扩展，静态数据加密存储，实现冷热数据分级管理，多副本备份防丢失。

4.3 网络传输设备

宜符合以下内容：

——有线网络：核心带宽 ≥ 10 Gbps，接入层 ≥ 1 Gbps，与医院HIS/EMR等系统逻辑隔离，仅开放标准医疗接口。

——无线网络：可全院覆盖医用级Wi-Fi 6，信号覆盖率 $\geq 99\%$ ，传输延迟 ≤ 50 ms，丢包率 $\leq 0.1\%$ ，支持WPA3加密。

4.4 硬件通用要求

- 4.4.1 兼容性：支持主流操作系统和医疗行业标准接口，与医院现有信息系统无缝对接。
- 4.4.2 院感与供电：临床终端采用医用抗菌材质，耐受常规消毒剂；核心硬件配置 UPS（续航≥2 小时），关键科室终端双路供电。
- 4.4.3 环境适配：服务端部署于专用机房，临床终端适应医院现场环境，具备防摔、抗干扰特性。
注：硬件设备的配置可根据实际情况进行调整。

5 核心技术

5.1 语音识别技术（ASR）

5.1.1 性能指标

5.1.1.1 准确率：

- 安静环境（≤55 dB）：
 - 门诊诊室、单人病房等场景；
 - 准确率≥95%。
- 基础嘈杂环境（56 dB~70 dB）：
 - 普通病房、急诊分诊区等场景；
 - 准确率≥92%。
- 高噪声环境（>70 dB）：
 - 抢救室、ICU、手术准备区等场景；
 - 准确率≥88%。

注：噪声等级依据GB/T 3222.1—2023在设备麦克风位置实测值。

——方言准确率≥75%，支持实时纠错与方言自适应。

- 5.1.1.2 鲁棒性：配备多麦克风阵列与主动降噪算法，鲁棒性：配备多麦克风阵列与主动降噪算法，在诊室噪声>60 dB 时仍保持识别稳定性，具体为语音识别准确率≥90%、单句转写延迟≤800 ms、连续识别无中断时长≥30 分钟。

5.1.2 功能要求

支持医疗专业术语库动态加载（如药品名、解剖学术语）。

5.2 医疗大语言模型（LLM）

5.2.1 模型架构

5.2.1.1 基于大规模医疗垂直模型构建，需满足以下性能基准：

- 场景覆盖能力：支持门诊、住院、急诊等全场景病历生成：
 - 实时语音生成：医患对话录音需清晰（信噪比≥40dB），AI同步解析并生成结构化病历，关键字段（主诉、现病史等12项核心要素）完整率≥98%；
 - 历史数据生成：基于患者既往医疗数据自动整合，需校验信息一致性，完整率≥98%。

注：系统需保留录音原始文件及病历修改痕迹，生成内容需经医生审核签名方可生效。

——知识承载能力：内置医疗知识库覆盖《疾病分类与代码》最新版≥99%的疾病实体；

——效率约束：生成标准门诊病历端到端延迟≤10 秒（复杂病历≤15 秒）”。

5.2.1.2 可动态融合最新临床指南（更新延迟≤72 小时）。

5.2.2 核心能力

5.2.2.1 结构化抽取：自动提取主诉、现病史、诊断依据等 12 类临床实体。

5.2.2.2 逻辑校验：内置矛盾检测引擎（如症状与诊断冲突预警）。

5.2.2.3 诊断依据的文献溯源需按临床重要性分级实施：

- 主要诊断依据（强制溯源）：
 - 范围：确定疾病诊断的核心依据（如病理结果、关键影像学特征、确诊性检验指标）；
 - 溯源深度：必须引用最新临床指南并标注文献名称、版本及具体条款
 - 覆盖率要求：100%主要诊断需满足此要求。

——次要诊断依据（条件溯源）：

- 范围：并发症、既往史、伴随症状等辅助性依据；
- 溯源规则：若存在争议性结论（如罕见病关联性判断），需提供文献支持；常规描述性内容（如高血压分级）允许仅标注标准名称；
- 错误率控制：文献引用错误率 $\leq 0.5\%$ 。

5.3 医学影像与文本处理技术

5.3.1 OCR 识别

5.3.1.1 外院报告关键字段（如肿瘤尺寸、实验室指标）提取准确率 $\geq 92\%$ 。

5.3.1.2 支持非结构化 PDF/手写体识别字符级错误率 $\leq 2\%$ ，语义级错误率 $\leq 1\%$ ，字段级错误率 $\leq 8\%$ 。

注1：若业务中存在大量极端潦草样本，可提前通过样本增强技术扩充训练数据，进一步优化模型适配性。

注2：建立“机器预警+人工复核”双重机制：模型识别后会置信度低于95%的字符或字段进行标红预警，由人工复核确认，确保最终输出结果符合业务要求。同时，针对高频出现的潦草手写体样本，定期纳入模型迭代数据集，持续提升模型对复杂手写体的适配能力。

5.3.2 多模态融合

按医疗场景差异化实现影像调取与报告生成：

——手术场景（实时记录）：

- 调取对象：患者术前关键影像（如 CT，MRI，DSA 等）。
- 触发条件：医生在手术过程中明确口述解剖结构或区域（如“胆囊三角区域”、“肝门部”）。
- 术中辅助：实时检索并显示关联影像切片，辅助术者决策。
- 术后记录：自动将关联影像切片嵌入手术记录文档；基于医生口述及影像内容，辅助生成该步骤的描述文本（如“解剖胆囊三角区域，分离出胆囊动脉与胆囊管，见附图”）。

——检查场景（检查过程中辅助与报告实时生成）：

- 调取对象：本次检查产生的实时影像流（如超声图像序列、内镜视频流）。
- 触发条件：医生在检查过程中描述病灶特征（如“发现肝脏占位”、“十二指肠球部溃疡”）。
- 报告中插入：自动捕获当前影像关键帧/截图并插入对应检查报告中。
- 结构化标注：在插入的影像上自动标记医生描述的病灶位置。
- 结构化描述生成：同步自动生成针对该病灶的结构化描述文本（包括但不限于：尺寸测量、具体位置、回声/色泽/形态特征等关键要素），

5.4 安全性

5.4.1 数据安全

5.4.1.1 采用联邦学习框架，训练数据不出医疗机构。

5.4.1.2 隐私保护：符合 GB/T 35273—2020，敏感信息加密存储。

5.4.2 内容标识与追踪

5.4.2.1 生成病历需嵌入隐式数字水印，符合《人工智能生成合成内容标识办法》。

5.4.2.2 操作日志全链路审计，修改记录可追溯至具体医师。

5.5 伦理治理与合规

5.5.1 公平性与偏见控制

5.5.1.1 医疗大语言模型训练数据应覆盖不同性别、年龄、民族、地域、疾病谱人群，避免因数据偏差导致诊疗建议歧视。

5.5.1.2 系统需内置偏见检测模块，定期评估输出内容是否存在以下倾向：

- 对特定人群诊断概率异常偏高/偏低；
- 治疗建议资源分配不均（如更倾向于推荐昂贵药物给某类患者）；

5.5.1.3 发现显著偏见时，触发模型再训练或加权调整机制。

5.5.1.4 可解释性与决策溯源所有由 AI 生成的关键临床判断（如初步诊断、用药建议）应提供依据来源说明，包括：

——引用的临床指南名称及版本；

——相似病例匹配度分析；

——推理链可视化（如“发热+咳嗽+肺部湿啰音 → 考虑社区获得性肺炎”）。

——支持医生点击查看“AI 思考过程”，提升信任度与临床采纳率。

注1：核心场景（如危重症用药、手术指征）：必须提供完整依据，包括指南引用+病例匹配+推理链可视化。

注2：常规场景（如慢性病随访建议）：可简化为“指南摘要+关键特征匹配”，无需完整推理链。

5.5.2 患者知情同意与参与机制

5.5.2.1 若系统使用真实患者数据用于模型优化（如联邦学习更新），须事先取得患者明示授权，并符合《个人信息保护法》第十三条关于敏感信息处理的规定。

5.5.2.2 在数字人问诊、语音采集等环节，应通过弹窗、语音播报等方式告知患者：

——“当前对话将被 AI 记录并用于生成病历”；

——“您的数据不会用于商业用途”；

——“您有权拒绝 AI 辅助记录”。

5.5.3 责任边界与人类监督

5.5.3.1 AI 仅作为辅助工具，不得替代执业医师做出最终诊断或签署法律文书。

5.5.3.2 所有 AI 生成内容必须经过医生审核、修改、确认签字后方可归档生效。

5.5.3.3 系统应在界面显著位置标注：“本内容由 AI 生成，仅供参考，最终责任由主治医师承担”。

6 应用场景

6.1 门诊病历生成

6.1.1 实时交互增强：医患对话通过 ASR 实时转写（延迟 ≤ 500 ms），结合 LLM 动态提取主诉、现病史等，生成初稿后自动关联历史病历对比逻辑矛盾。

6.1.2 模板自适应：基于疾病类型（如呼吸系统/心血管）自动匹配专科模板，支持医生语音指令修改。

6.2 住院病历管理

6.2.1 住院文书全周期生成

6.2.1.1 移动查房与语音病历采集

应符合以下内容：

——支持在 Pad 或手持终端上通过语音方式记录查房内容，ASR 实时转写准确率 $\geq 90\%$ ；

——查房语音自动同步至 HIS/EMR 系统，并结合患者、既往病史、实时生命体征（心率、血压、SpO₂等）、检验检查结果，生成或更新：

- 入院记录；
- 首次病程记录；
- 日常病程记录；
- 结合语音及 HIS 系统相关数据，生成或更新记录后，记录同步至 HIS/EMR。

——技术要求：

- 支持离线缓存，在网络中断时暂存本地，恢复后自动上传；
- 与医院时钟系统对齐，时间戳精确至秒级。

6.2.1.2 数字人辅助问诊生成入院记录

应符合以下内容：

——部署虚拟数字人（Digital Human）于门诊分诊区或病房前导环节，开展结构化问诊：

- 主诉采集、既往史、过敏史、家族史、个人史等；

- 支持自然语言对话与语音唤醒交互。
- 问诊过程由医疗大模型动态理解语义，自动提取临床实体并生成初步《入院记录》草稿。
- 医生可查看数字人原始问答日志、AI 生成内容及修改痕迹，确认后正式归档。
- 合规要求：
 - 数字人不得独立做出诊断或开具医嘱；
 - 所有生成内容需经执业医师审核签字方可生效；
 - 保留完整操作审计链，符合《电子病历应用管理规范》。

注：企业可自行选择。

6.2.1.3 智能出院摘要与出院小结一体化生成

宜包括但不限于以下内容：

- 出院前系统自动整合本次住院全周期诊疗数据，覆盖医疗文书、诊疗操作、检验检查、护理记录等核心维度，具体数据源如下：
 - 核心医疗决策数据：入院诊断、修正诊断、最终诊断及诊断依据；手术/操作名称、术时、术中关键发现、并发症记录；
 - 全程诊疗过程数据：日常病程记录、首次病程记录、上级医师查房记录、疑难 / 危重病例讨论记录；长期/临时医嘱单（含用药医嘱、治疗医嘱执行情况）；
 - 检验检查全量数据：住院期间所有实验室检验报告（血常规、生化、病原学等）、影像学检查报告（CT/MRI/ 超声等）及关键影像截图、病理报告；
 - 治疗与康复相关数据：抗菌药物使用天数、用药种类、调整记录；输血记录；术后恢复曲线（引流液量、疼痛评分、生命体征变化趋势）；护理记录（病情观察、护理措施、患者依从性）；康复评估记录（功能恢复情况、康复训练效果）；
 - 随访与预后相关数据：出院时病情转归、出院后用药指导、饮食/运动建议、复诊时间及复查项目、需警惕的异常情况及应急处理方式。
- LLM 自动生成《出院小结》《出院指导》《康复建议书》三类文档：
 - 出院小结：面向医疗机构流转，严格遵循《病历书写基本规范》格式要求，完整整合上述全量数据，重点突出诊疗核心节点、关键决策依据及转归情况，确保数据与住院病历一致性；
 - 出院指导：面向患者及家属，采用通俗化语言转化专业医疗信息，明确用药剂量/频次、禁忌事项、生活方式调整建议，同步提示复诊必要性及延误风险；
 - 康复建议书：基于患者年龄、基础疾病、手术类型及恢复情况个性化生成，包含分阶段康复训练方案（强度、时长、动作要点）、康复期营养支持建议、并发症预防措施，可关联权威康复指南相关内容。

6.2.2 多端协同

版本一致性：医生工作站与移动端采用存证技术，确保病历修改记录可追溯。

7 系统安全与合规

7.1 数据全生命周期安全

7.1.1 存储与传输

7.1.1.1 本地化部署：医疗数据强制存储于院内私有云或政务云专区，禁止跨境传输。

7.1.1.2 加密标准：传输层采用国密 SM4/SM9 算法，静态数据使用 AES-256 加密，密钥由医院硬件加密机管理。

7.1.2 隐私保护

7.1.2.1 动态脱敏：病历生成时自动屏蔽身份证号、联系方式等字段，支持按角色授权查看完整信息。

7.1.2.2 匿名化处理：训练数据需通过 k-匿名化 ($k \geq 5$) 或差分隐私 ($\epsilon \leq 0.8$) 技术处理，满足 GB/T 35273—2020。

7.2 内容安全与质控合规

7.2.1 生成内容监管

7.2.1.1 双重质控引擎

7.2.1.1.1 AI 内涵质控：基于医疗大模型分析诊疗合理性（如用药剂量超指南预警）。

注：保持内置，仅在医生端显示预警提示。

7.2.1.1.2 可溯源标识：生成病历嵌入隐式数字水印，包含模型版本、生成时间及操作者 ID，符合《生成式 AI 服务管理暂行办法》。

7.2.2 法规适配性

7.2.2.1 电子病历格式强制符合《电子病历应用管理规范（2025 修订版）》；

7.2.2.2 AI 生成病历中的诊断依据、治疗建议等内容，应基于权威医学知识库或最新临床指南，并满足以下要求：

——来源可溯：

- 系统内置知识库应来源于国家卫生健康委发布或认可的临床诊疗指南、路径及标准；
- 支持对接中华医学会系列指南、国家临床医学中心推荐方案、NICE/NCCN 等国际权威指南（经本地化审核后启用）；
- 每项关键诊断建议应能回溯至具体指南条目或证据等级（如 I A 类推荐）。

——动态更新机制：

- 知识库更新周期不得超过 72 小时，自官方发布新版指南起计算；
- 更新日志应记录版本变更、生效时间及影响范围。

7.3 基础设施与访问控制

7.3.1 网络安全架构

7.3.1.1 零信任防护：部署零信任网关，实现设备、用户、应用三重认证。

7.3.1.2 边界防御：配置下一代防火墙（支持入侵防御/防病毒）及堡垒机审计运维操作。

7.3.2 权限管理

7.3.2.1 分层授权：按角色划分权限（如住院医师仅可修改病程记录，主任医师拥有终审权）。

7.3.2.2 双因子认证：敏感操作（如批量导出病历）需短信+生物特征验证。

7.4 审计追溯与司法存证

应符合表 3 内容。

表 1 审计追溯与司法存证

要求	技术实现
全链路审计	病历生成系统：负责日志实时采集，存储≤6 个月后自动触发转交流程。 医院档案平台：接收日志后进行加密归档，按《病历管理规定》留存≥10 年，配置独立访问权限（如仅审计人员可调取）
区块链存证	诊断结论、患者知情同意书等关键数据哈希上链
司法调阅接口	支持法院通过 API 调取区块链存证及操作日志

7.5 应急与容灾管理

7.5.1 业务连续性

7.5.1.1 容灾等级：符合 GB/T 20988—2007 5 级标准。

7.5.1.2 应急切换：主系统故障时，自动降级至基线版病历生成模式。

7.5.2 漏洞响应

7.5.2.1 高危漏洞修复时限≤24 小时，中危漏洞≤72 小时。

7.5.2.2 每季度开展第三方红队攻防演练，渗透测试覆盖率达 100%。

8 运行与维护

8.1 运维监控体系

8.1.1 性能指标

8.1.1.1 响应时效

病历辅助生成系统的响应时效应根据病历类型、输入规模和临床场景需求进行差异化定义，并满足以下分级性能要求：

表 2 响应时效

病历类型	输入特征描述	平均耗时	适用场景
简易病历	主诉+简要随访信息≤150字；无影像联动、无复杂逻辑校验	8~12 秒	门诊慢病随访、复诊小结
标准门诊病历	门诊初诊语音输入（主诉+现病史≤300字）+基础历史病历对比；含常规结构化抽取	15~20 秒	门诊初诊记录
常规住院病程病历	查房语音+实时生命体征数据+简要病程补充；含基础逻辑校验与字段补全	均值≤8 秒, P95 ≤12 秒	住院日常病程记录
复杂住院首程病历	查房语音+完整现病史+初步诊断依据 + 全面历史病历对比；含深度结构化抽取与矛盾校验	均值≤12 秒, P95 ≤16 秒	住院首次病程记录
综合入院记录	主诉+现病史+既往史+过敏史+家族史+检验/影像数据+多模态融合；需整合全维度临床数据	≤25 秒	入院记录生成
跨学科会诊摘要	多学科诊疗意见+多源病历数据+文献溯源 + 推理链可视化；需跨学科知识融合	≤30 秒	MDT 会诊摘要、疑难病例讨论记录

8.1.1.2 可用性

系统全年可用率≥99.9%，故障自动切换时间≤30秒。

8.1.2 智能运维

8.1.2.1 异常预测：基于日志分析预判硬件故障（准确率≥85%），当预测模型检测到高风险故障（如硬盘 SMART 指标异常），自动触发双备份同步，确保数据一致性。

8.1.2.2 资源弹性：CPU 利用率>80%时自动扩容，<30%时缩容以减少成本。

8.2 持续优化机制

8.2.1 模型迭代

8.2.1.1 月度评估：按疾病类型抽样测试诊断准确率，未达 90%的模型触发重训练。

8.2.1.2 知识库更新：临床指南发布后 72 小时内同步至系统，模板库季度更新。

8.2.2 用户反馈闭环

8.2.2.1 医生标注系统

支持在线标注生成错误（如实体提取偏差），自动加入训练集。

8.2.2.2 版本灰度

8.2.2.2.1 灰度范围

8.2.2.2.1.1 初始部署覆盖不超过两个临床科室或≤10%活跃用户数，优先选择非急诊、低风险科室（如门诊慢性病管理、康复科）进行试点。

8.2.2.2.1.2 灰度期间收集真实场景下的使用数据与医生反馈。

8.2.2.2.2 错误分类与容忍阈值

系统运行期间应对AI生成内容进行抽样质控，错误按严重程度分为三级，并设定不同容忍限值：

表 3 容忍限值

错误等级	定义	接受标准（日均）	处置机制
严重错误 (Severe)	事实性医学错误（如错误诊断、解剖结构混淆） 用药剂量/禁忌症错误导致潜在高风险 AI 替代医师签署法律文书或做出独立诊疗决策	0例	立即触发警报，暂停相关模块服务； 启动医疗安全事件报告流程； 72小时内完成根因分析与整改
主要错误 (Major)	关键临床字段缺失（如主诉、过敏史、既往史） 模板错配导致信息结构混乱； 存在明显逻辑矛盾（如无发热记录却开具退热药）； 文献溯源失败（主要诊断无指南支持）	≤0.5% 的生成任务	连续3天超标即启动模型再训练流程； 向科室负责人发送预警通知
次要错误 (Minor)	语句不通顺、术语使用不精准 标点符号或格式错误 同义词替换不当但不影响临床理解	不设硬性上限，但需满足： 年度趋势下降 ≥10%/年； 单机构月度投诉率<1‰	自动归集至反馈闭环系统； 每季度汇总用于模型微调与模板优化
<p>注1：抽样方法：每日随机抽取不少于总生成量5%的病历，由第三方质控团队或医院医务处进行盲评。</p> <p>注2：患者因隐私隐瞒（如性病史）、语音录入失败（方言/噪音导致未识别）、数字人对话中明确拒绝回答，不属于病历生成错误。</p>			

8.2.2.2.3 应急回滚机制

应符合以下内容：

——若灰度期间出现以下情况之一，立即暂停并回滚：

- 发生一起及以上疑似因AI误导导致的医疗纠纷苗头；
- 单日主要错误率超过1.5%；
- 用户投诉集中爆发（>5起/日，经核实属实）。

——回滚应在 1 小时内完成，并启动根因分析流程。