

# 团 体 标 准

T/ISC XXXX—XXXX

## 医疗健康行业智能体 语音电子病历生成智能体技术要求

Intelligent agents in healthcare industry-Technical requirements for speech-based electronic medical record generation intelligent agents

（工作组讨论稿）

在提交反馈意见时，请将您知道的相关专利连同支持性文件一并附上。

XXXX - XX - XX 发布

XXXX - XX - XX 实施

中 国 互 联 网 协 会 发 布

目次

前言 ..... III

1 范围 ..... 4

2 规范性引用文件 ..... 4

3 术语和定义 ..... 4

4 符号和缩略语 ..... 5

5 总体要求(待定) ..... 6

6 功能完备性技术要求 ..... 6

6.1 门诊病历生成 ..... 6

6.1.1 患者信息展示 ..... 6

6.1.2 实时问诊 ..... 7

6.1.3 辅助问诊 ..... 7

6.1.4 门诊病历生成 ..... 7

6.2 住院病历生成 ..... 8

6.2.1 患者信息展示 ..... 8

6.2.2 病程摘要 ..... 8

6.2.3 摘要日历 ..... 8

6.2.4 住院病历生成 ..... 9

6.3 医疗术语标准化 ..... 9

6.4 质控与合规体系 ..... 10

7 准确性要求 ..... 10

7.1 二分类任务 ..... 10

7.2 文书生成类任务 ..... 11

7.3 多分类任务 ..... 11

8 智能体能力要求 ..... 12

8.1 感知能力 ..... 12

8.1.1 回答时效性 ..... 12

8.1.2 推理能力 ..... 12

8.1.2.1 指代消解 ..... 12

8.1.2.2 术语与标准符合性检查 ..... 13

8.1.2.3 知识推理 ..... 13

8.2 规划能力 ..... 13

8.2.1 任务规划 ..... 13

8.2.1.1 目标拆解 ..... 13

8.2.1.2 规划策略 ..... 14

8.2.2 任务调度 ..... 14

8.2.2.1 调度机制 ..... 14

8.2.2.2 组织协调 ..... 15

8.3 记忆能力 ..... 15

8.3.1	短期记忆能力 .....	15
8.3.1.1	提示词管理 .....	15
8.3.1.2	记忆存储 .....	15
8.3.2	长期记忆能力 .....	16
8.3.2.1	记忆存储 .....	16
8.3.2.2	快速检索 .....	16
8.4	执行能力 .....	16
8.4.1	虚拟环境执行能力 .....	16
8.4.2	现实环境执行能力 .....	17
9	易用性要求 .....	17
9.1	可理解性 .....	17
9.1.1	语言表达清晰程度 .....	17
9.1.2	辅助理解手段 .....	17
9.2	易学性 .....	17
9.2.1	帮助文档完整性 .....	17
9.2.2	差错信息易理解性 .....	18
9.3	易操作性 .....	18
9.3.1	操作一致性 .....	18
9.3.2	消息明确性 .....	18
9.3.3	辅助输入手段 .....	18
10	安全性要求 .....	18
10.1	基础设施安全 .....	18
10.1.1	硬件设备安全性 .....	18
10.1.2	软件设备安全性 .....	18
10.2	数据安全 .....	19
10.3	应用安全 .....	19
10.3.1	内容安全 .....	19
10.3.2	服务安全 .....	19
附录 A (规范性)	XXX .....	错误！未定义书签。
附录 B (资料性)	XXX .....	错误！未定义书签。
参考文献	.....	20

## 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。

本文件是T/CIITA x《……》第i部分。T/CIITA x已经发布了以下部分：

——第1部分：……；

——……；

——第n部分：……。

本文件替代T/CIITA x.i《……》，与T/CIITA x.i相比，除结构调整和编辑性改动外，主要技术变化如下：

a)

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。

本文件由中国互联网协会提出并归口。

本文件起草单位：讯飞医疗科技股份有限公司、中国信息通信研究院、安徽医科大学第一附属医院、四川大学华西医院、中国医科大学附属盛京医院、安徽省合数智医科技有限公司

本文件主要起草人：

本文件及其所代替文件的历次版本发布情况为：

——

# 医疗健康行业智能体 语音电子病历生成智能体技术要求

## 1 范围

本文件规定了医疗健康行业智能体 语音电子病历生成智能体在应用过程中涉及的技术能力，从功能完备性技术要求、准确性要求、智能体能力要求、易用性要求和安全性要求等维度对智能体技术在病历生成场景中应用的能力提出要求。

## 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 20113—2019 医学术语系统命名

WS 310.1—2010 电子病历基本架构与数据标准 第 1 部分：基本架构

GB/T 22239—2019 信息安全技术 网络安全等级保护基本要求

GB/T 38642—2020 信息安全技术 医疗健康数据安全指南

## 3 术语和定义

### 3.1

**语音电子病历生成智能体** speech-based electronic medical record generation intelligent agent

具备语音信号采集、识别、临床信息提取、医疗术语标准化及结构化病历自动生成等核心能力，适配门诊、住院等诊疗场景，辅助医生高效完成电子病历书写的医疗健康行业智能体。

### 3.2

**有效诊疗语音** effective diagnosis and treatment speech

医患交互过程中与患者病情、诊疗行为直接相关的语音内容，包括症状描述、病史陈述、检查建议、诊断结论、治疗方案沟通等。

### 3.3

**医疗术语标准化** medical terminology standardization

将非规范医疗表述（如口语化症状描述、药品商品名、手术简写）转换为符合国家 / 行业标准的医学术语，确保病历术语统一、数据互通的处理过程。

### 3.4

**指代消解** coreference resolution

在医患对话及病历生成过程中，识别代词或简化表述（如“该症状”“此检查”）所对应的具体医疗实体（如特定症状、检查项目、药物名称）的技术能力。

### 3.5

**病历质控与合规体系** medical record quality control and compliance system

通过自动校验机制，对病历的完整性、准确性、合规性、逻辑一致性进行核查，确保病历符合《病历书写基本规范》等法规要求的体系。

### 3.6

#### **短期记忆能力 short-term memory ability**

智能体在单次诊疗交互中，留存医患对话、临床要素、操作痕迹等数据，支撑实时病历生成的记忆能力。

### 3.7

#### **长期记忆能力 long-term memory ability**

智能体跨就诊周期留存患者诊疗数据（如既往病史、检查结果、治疗方案），支撑多轮次就诊病历关联生成的记忆能力。

### 3.8

#### **SOAP 框架 SOAP framework**

一种临床文书结构化书写框架，包含主观资料(Subjective)、客观资料(Objective)、评估(Assessment)、计划(Plan)四个核心部分，用于梳理诊疗逻辑并生成规范病历。

## 4 符号和缩略语

下列符号和缩略语适用于本文件。

Acc: 准确率 (Accuracy)

AES: 高级加密标准 (Advanced Encryption Standard)

AI: 人工智能 (Artificial Intelligence)

CRISPE: CRISPE 提示词框架 (Context, Role, Intent, Style, Example)

FN: 假阴性 (False Negative)

FP: 假阳性 (False Positive)

F1: F1 值 (F1-Score)

HIS: 医院信息系统 (Hospital Information System)

Hit\_Rate: 关键词命中率 (Hit Rate)

ICD: 国际疾病分类 (International Classification of Diseases)

ICIO: 提示词框架 (Issue, Context, Input, Output)

LIS: 检验信息系统 (Laboratory Information System)

Macro-F1: 宏 F1 值 (Macro-F1 Score)

Macro-P: 宏精准率 (Macro-Precision)

Macro-R: 宏召回率 (Macro-Recall)

MTTF: 平均无故障时间 (Mean Time To Failure)

PACS: 医学影像存档与通信系统 (Picture Archiving and Communication Systems)

Precision: 精准率 (Precision)

P\_D: 标准符合准确率 (Precision of Standard Compliance)

P\_M: 指代消歧平均准确率 (Precision of Coreference Resolution)

P\_P: 规划策略准确率 (Precision of Planning Strategy)

R: 召回率 (Recall)

ROUGE-N: ROUGE-N 指标 (Recall-Oriented Understudy for Gisting Evaluation-N)

Sen: 灵敏度 (Sensitivity)

Spe: 特异度 (Specificity)

SOAP: SOAP 框架 (Subjective, Objective, Assessment, Plan)

SSL: 安全套接层 (Secure Sockets Layer)

TP: 真阳性 (True Positive)

TN: 真阴性 (True Negative)

HTTPS: 超文本传输安全协议 (Hyper Text Transfer Protocol Secure)

5 总体要求

医疗健康行业智能体 语音电子病历生成智能体总体要求架构包括在应用过程中涉及的技术能力，从功能完备性技术要求、准确性要求、智能体能力要求、易用性要求和安全性要求等维度对智能体技术在病历生成场景中的能力提出要求。



图 1 架构图

6 功能完备性技术要求

6.1 门诊病历生成

6.1.1 患者信息展示

实现患者信息的全面整合、异常指标智能预警、精准检索与实时同步，为医生诊疗提供完整且高效的信息支持。

- a) 支持对接医院 HIS 系统，自动整合患者基础信息、历史就诊记录、预问诊数据及关键医疗数据，生成结构化全景档案，提炼核心诊疗信息（如一句话就诊摘要、既往疾病史、过敏史）并以可视化方式高亮重要事件。
- b) 实时解析门诊患者检验检查报告数据，自动识别异常指标（如血液生化异常、影像学阳性发现），通过智能弹窗提醒、颜色高亮标注等方式向医生推送预警信息。

- c) 提供完整患者列表，支持按科室、病历生成状态（如“待生成”“已生成”“需修订”）等多维度筛选，展示患者病历生成进度及关键诊疗信息（最后就诊时间、主诊断、异常指标），并支持通过患者 ID、姓名精准检索。
- d) 在 HIS 系统选取患者后，客户端应用可实时同步切换，确保医生操作的病历信息与接诊患者精准匹配，实现无缝衔接。

#### 6.1.2 实时问诊

保障医患对话采集清晰、转写准确、记录可回溯，并实现多段语音整合与关键信息标准化提取，为病历生成提供高质量数据基础。

- a) 采用医疗级降噪技术，实现医生与患者语音信号的独立通道采集，确保嘈杂诊室环境下对话内容清晰识别。
- b) 基于深度学习语音识别引擎，将医患对话实时转化为结构化文本，支持转写过程中 AI 即时编辑修正，语音转写准确率不低于 98%。
- c) 提供完整时间轴管理界面，可按就诊日期及时间多维度检索定位历史对话记录，支持音频回放与文本对照查看，同时实现医患角色分离，保存单次就诊中的所有历史对话。
- d) 具备多段语音融合生成功能，可将一次就诊中分散的医患对话片段整合，无缝衔接形成连贯、全面的对话记录，真实反映医患交流全过程。
- e) 通过智能语音分析实时识别对话中的关键临床要素（如症状、三史、医疗行为、时间描述），自动执行标准化转换（如将口语“拉肚子”转为医学术语“腹泻”、模糊时间“前几天”量化为“3 天前”），并建立症状 - 时间 - 诱因的关联结构。

#### 6.1.3 辅助问诊

基于患者数据动态提供问诊、诊断及检查建议，辅助医生完善病史采集、优化诊断思路，提升问诊全面性与诊疗精准度。

- a) 基于患者电子健康档案和实时对话内容，自动生成个性化问诊建议，推荐需重点关注的问诊方向（如遗漏的症状细节、既往相关病史），辅助医生全面采集病史。
- b) 在医患对话过程中，结合患者历史病历、检验检查数据及临床知识库，动态分析潜在诊断方向，按症状关联性、疾病优先级和临床指南对诊断建议排序，并标注置信度等级，辅助医生调整问诊重点。
- c) 依据当前诊疗进展和患者临床特征（如症状、病史、初步检查结果），AI 自动预判后续可能需要的检查检验项目，为医生提供前瞻性决策支持，优化诊疗流程安排。

#### 6.1.4 门诊病历生成

实现病历自动生成、可溯源、可编辑，并通过临床思维可视化与质量核查，确保病历规范、准确且符合诊疗逻辑。

- a) 基于 AI 智能分析医患对话、患者历史数据及检验检查结果，自动生成符合《门诊病历书写规范》的电子病历，智能填充主诉、现病史、既往史、体格检查、诊断及处理意见等核心模块，支持一键引用至医院 HIS 系统。
- b) 提供全链路数据溯源功能，清晰展示病历生成依据（如关联的医患对话片段、检验报告数值、历史就诊记录），支持逐项核对与修订，确保诊断过程可追溯、可审计。



- c) 支持医生对 AI 生成病历灵活编辑，包括自由文本修改、模块调整、关键信息增删，且保留所有修改记录，兼顾自动化效率与人工干预精准性；当病历不符合临床需求时，支持一键触发病历重新生成，系统基于最新数据优化内容。
- d) 基于患者数据动态生成诊断逻辑路径，运用 SOAP 框架（主观资料、客观资料、评估、计划）可视化展示 AI 分析过程，梳理问诊思路、多维度分析症状、补全 SOAP 结构，并对病历进行逻辑一致性、诊疗合规性、用药安全风险三重核查。

## 6.2 住院病历生成

### 6.2.1 患者信息展示

实现住院患者信息的分类筛选、生成记录跟踪、任务智能排序与快速床位切换，提升住院病历管理效率与便捷性。

- a) 系统提供全面的患者列表视图，方便医生快速浏览和管理所有患者信息。列表支持按姓名、床号等信息展开查询，也支持通过我的患者/全科患者筛选，便于医生快速定位到特定患者。
- b) 展示住院患者今日已生成的文书列表，医生可以查看文书生成记录的状态、生成时间以及简要内容，便于跟踪和管理。
- c) 列出尚未生成或即将触发生成的文书任务，系统根据患者住院情况（如治疗进度、出院时间）和预设规则，智能推荐任务生成顺序，帮助医生高效安排工作。
- d) 支持按床位顺序快速切换患者（上一床/下一床），并提供全院或当前科室床位一览表，可按姓名、床号查询并跳转至目标患者，适用于查房、集中文书处理场景。

### 6.2.2 病程摘要

自动整合住院全周期数据，生成多类型结构化摘要并实时更新，同时标记关键事件，为医生快速掌握患者病情变化提供支持。

- a) 利用自然语言处理技术，自动分析患者住院期间的医嘱、检查报告、护理记录、手术记录等所有病历资料，识别提取关键信息（如主要诊断、治疗方案、重要检查结果、并发症）。
- b) 从时间维度生成住院期间每日 24 小时患者全维度数据快照，从事件维度在诊疗关键决策点（入院、会诊、手术、危急值、转科）生成定向文本摘要，包括入院摘要、检查检验摘要、手术摘要、会诊摘要、查房摘要、危急值摘要、出院摘要。
- c) 实时同步患者最新临床数据（如生命体征、检验结果、用药记录），确保医护人员查看的摘要始终为最新状态，避免因信息滞后导致误判。
- d) 智能识别并标注诊疗过程中的重要节点（如手术、会诊、病情变化），形成时间轴标记，帮助医生快速定位关键信息，梳理患者诊疗历程。

### 6.2.3 摘要日历

以日历形式可视化展示住院关键事件，同时自动识别风险、发出预警并跟踪处理，辅助医生直观掌握诊疗进度与患者风险状况。

- a) 将提取的住院关键信息按时间顺序整理，以日历形式展示患者住院周期，按日期排列关键事件（如入院、检查、手术、会诊），支持周视图，方便快速定位特定日期或事件。
- b) 自动识别并标记重要医疗事件，用不同颜色或图标区分事件类型（如红色标记“危急值报警”、蓝色标记“手术”、绿色标记“检查报告生成”）。

- c) 分析患者病历资料中的数值型数据（如生命体征、实验室检查结果），结合医学知识和临床经验，自动识别可能存在的风险或异常（如药物过敏、感染风险、电解质紊乱）。
- d) 对识别出的风险或异常自动生成预警提示信息（含风险类型、具体数值、建议关注事项），以醒目方式（如红色标记、弹窗）展示，并支持对预警处理情况记录和跟踪，根据处理结果更新预警状态。

#### 6.2.4 住院病历生成

实现住院病历自动预生成、多模块结构化重构、个性化调整与重新生成，减轻医生文书负担，确保病历完整规范。

- a) 通过预设规则或关键节点检测（如治疗完成、出院前特定时间），自动整合患者最新住院数据（入院情况、诊疗经过、检查结果、手术记录），结合医学知识库和标准模板，实时生成结构化初步住院病历（如出院记录草稿），支持动态更新与智能预测（如药物清单、随访计划）。
- b) 自动提取患者姓名、性别、年龄、住院号、入院日期、出院日期等基本信息，按标准格式生成患者基本信息段落；同时实现入院情况重构（提取主诉、现病史、体征等）、诊疗经过重构（记录检查检验、会诊、手术情况）、出院情况重构（汇总出院时症状、体征及带药信息）。
- c) 支持根据医院或科室的个性化病历规范，自动生成符合要求的标准化文本段落，医生可通过预置模板或自定义规则（如术语风格、段落结构）调整输出内容；根据医嘱自动生成出院医嘱分析，包括药物治疗、康复建议、随访计划、特殊指示。
- d) 当患者住院信息（如检查结果、治疗方案）发生变化时，系统可自动检测并重新生成病历；医生也可手动触发重新生成操作，系统基于最新数据生成病历并允许预览、编辑和引用。

#### 6.3 医疗术语标准化

对检验检查结果、药品名、手术记录等医疗术语进行语义化转换与标准化处理，确保病历术语规范统一，支持精准医疗与数据互操作。

- a) 对检验检查结果进行语义化转换：将数值结果转化为定性描述（如“血糖 7.8mmol/L”转为“血糖轻度升高”）、标注异常指标临床意义（如“ALT 120U/L”转为“谷丙转氨酶显著升高，提示肝功能异常”）、规范形态学与微生物检测结果描述（如“痰培养出金黄色葡萄球菌”转为“痰培养阳性，检出金黄色葡萄球菌”）。
- b) 实现药品名语义化转换：将药品商品名转为通用名（如“奥格门汀”转为“阿莫西林克拉维酸钾”）、统一剂量单位（如“0.5g qd”转为“每日一次，每次 0.5 克”）、规范给药途径与药物作用描述（如“ivgtt”转为“静脉滴注”、“头孢呋辛酯”转为“头孢类抗生素，用于抗感染治疗”）。
- c) 自动从手术记录中提取关键信息并标准化：将术式简写转为全称（如“LC”转为“腹腔镜胆囊切除术”）、补充手术目的与入路描述（如“阑尾切除术”转为“为切除炎症阑尾而行阑尾切除术”）。
- d) 对治疗方案、疾病诊断、症状体征、医学影像描述进行语义化转换与规范（如“头疼”转为“头痛”、“CT 示脑部低密度影”转为“CT 扫描显示脑部存在低密度病灶，考虑缺血性改变”），同时综合多检查结果进行分析并语义化总结。

6.4 质控与合规体系

通过自动校验，确保病历质量达标、操作可追溯且数据互通，降低医疗风险。平台支持对病历进行自动审核校验，校验内容包括：开展完整性校验（确保包含患者基本信息、诊断、治疗方案等所有必要内容）、准确性校验（核对诊断、治疗与病历资料一致性，验证时间线与数据准确性）、合规性校验（确保符合医疗文书书写规范与术语标准）、逻辑一致性校验（检查病历内容前后无矛盾，诊断与治疗相符）。

7 准确性要求

7.1 二分类任务

二分类任务准确性要求主要针对语音类语料处理，聚焦医患语音信号的分类识别，比如区分有效诊疗语音与无干扰语音、特定症状表述的有无判断等场景。

二分类混淆矩阵

分类		人工智能分类	
		阳性	阴性
参考标准分类	阳性	真阳性（TP）	假阴性（FN）
	阴性	假阳性（FP）	真阴性（TN）

- a) 灵敏度（Sen）：针对有效诊疗语音的识别，灵敏度不低于97%，确保不遗漏关键医疗对话内容，避免因漏检导致病历核心信息缺失。

$$Sen = \frac{TP}{TP + FN} \times 100\%$$

式中：  
Sen——灵敏度；  
TP——真阳性，正确识别的有效诊疗语音片段数；  
FN——假阴性，未识别出的有效诊疗语音片段数。

- b) 特异度（Spe）：针对无干扰语音的过滤，特异度不低于98%，减少无效信息对病历生成的干扰，确保病历内容聚焦医疗核心。

$$Spe = \frac{TN}{TN + FP} \times 100\%$$

式中：  
Spe——特异度；  
TN——真阴性，正确识别的无干扰语音片段数；  
FP——假阳性，误判为有效诊疗语音的干扰片段数。

- c) 准确率（Acc）：为保障语音数据筛选的精准性，整体二分类任务准确率不低于97.5%，为后续病历生成提供高质量数据基础。

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \times 100\%$$

式中：  
Acc——准确率；

TP——真阳性，正确识别的有效诊疗语音片段数；  
 TN——真阴性，正确识别的无关干扰语音片段数；  
 FP——假阳性，误判为有效诊疗语音的干扰片段数；  
 FN——假阴性，未识别出的有效诊疗语音片段数。

## 7.2 文书生成类任务

文书生成类任务准确性要求面向门诊/住院病历等医疗文书的自动生成，核心是确保生成文本与参考摘要的语义契合度、关键诊疗信息的命中率，贴合《门诊病历书写规范》等标准。

- a) 客观指标（ROUGE-N）：对生成任务，计算客观指标ROUGE-N，聚焦病历文本的语义连贯性与核心信息完整性，N参考值为2，要求门诊病历ROUGE-2值不低于92%；住院病历含病程记录、出院小结，ROUGE-2值不低于90%。参考摘要需以医生手写规范病历或临床认可的标准病历为基准，确保评估贴合临床实际。其计算公式如下：

$$\text{ROUGE} - N = \frac{\sum S \in \{\text{ReferenceSummaries}\} \sum \text{gram}_n \in \text{sCount}_{\text{match}}(\text{gram}_n)}{\sum S \in \{\text{ReferenceSummaries}\} \sum \text{gram}_n \in \text{sCount}(\text{gram}_n)}$$

式中：

N——即 n-gram，文本内容滑动窗口字节数，参考值为 2；

$\text{Count}_{\text{match}}(\text{gram}_n)$ ——参考摘要和机器生成摘要中共有的 n-gram 的数量；

$\text{Count}(\text{gram}_n)$ ——参考摘要中 n-gram 的数量。

- b) Hit\_Rate（关键词命中率）：关键词需覆盖病历核心要素，包括患者基本信息、主诉、现病史关键症状、既往史、过敏史、检查结果、诊断结论、治疗方案等。要求核心关键词命中率不低于95%。其计算公式如下：

$$\text{Hit\_Rate} = \frac{\text{Count}_{\text{Hit}}}{\text{len}(\text{keyword})}$$

式中：

Hit\_Rate——关键词命中率；

$\text{Count}_{\text{Hit}}$ ——生成病历中准确命中的核心临床关键词数量；

$\text{len}(\text{keyword})$ ——关键词的数量。

## 7.3 多分类任务

多分类任务准确性要求则针对多维度医疗类别划分，覆盖疾病诊断分型、症状严重程度分级、检查结果异常等级判定、诊疗阶段划分等场景，保障分类结果的精准可靠。

- a) Macro-Precision（宏精准率）：要求不低于93%，确保各类别划分的精准性，避免因分类错误导致病历诊断描述、病情评估等内容失真。

$$\text{Macro} - P = \frac{1}{K} \sum_{i=1}^K \text{Precision}_i$$

式中：

Macro - P——宏精准率；

K——分类任务的总类别数（如三分类时 K=3）；

$Precision_i$ ——第  $i$  类精准率，计算公式为  $Precision_i = \frac{TP_i}{TP_i + \sum_{j \neq i} FP_{j \leftarrow i}}$  ( $TP_i$  为  $i$  类真阳性数， $FP_{j \leftarrow i}$  为实际非  $i$  类但预测为  $i$  类的假阳性数， $j \neq i$  时属于  $i$  类的误判)。

- b) **Macro-Recall (宏召回率++)**：要求不低于92%，确保各类别临床信息无遗漏，全面覆盖患者诊疗相关的多维度特征。

$$Macro - R = \frac{1}{K} \sum_{i=1}^K Recall_i$$

式中：

$Macro - R$ ——宏召回率++；

$K$ ——分类任务的总类别数（如三分类时  $K=3$ ）；

$Recall_i$ ——第  $i$  类精准率，计算公式为  $Recall_i = \frac{TP_i}{TP_i + \sum_{j \neq i} FN_{i \rightarrow j}}$  ( $TP_i$  为  $i$  类真阳性数， $FN_{i \rightarrow j}$  为实际  $i$  类但预测为  $j$  类的假阴性数， $j \neq i$  时属于  $i$  类的漏判)。

- c) **Macro-F1 (宏精准率与宏召回率的调和平均)**：要求不低于92.5%，综合平衡精准率与召回率，保障多分类结果既准确又全面，贴合病历生成的临床实用性要求。

$$Macro - F1 = \frac{2 \times Macro - P \times Macro - R}{Macro - P + Macro - R}$$

式中：

$Macro - F1$ ——宏精准率与宏召回率的调和平均；

$Macro - P$ ——宏精准率；

$Macro - R$ ——宏召回率++。

## 8 智能体能力要求

### 8.1 感知能力

#### 8.1.1 回答时效性

语音电子病历生成智能体应具备一定的回答时效性，聚焦病历生成全流程的实时响应需求，从医患对话启动、临床数据查询到病历模块生成，确保响应速度不影响诊疗效率。要求单次响应时间不超过1.5秒，且连续多轮诊疗交互中响应延迟波动不超过0.3秒。

通过用户从发起请求到系统或应用程序返回结果的时间计算响应实时性，计算方式参见式：

$$ES\_T = R\_T_{finish} - R\_T_{start}$$

式中：

$ES\_T$ ——响应时间；

$R\_T_{finish}$ ——语音电子病历生成智能体返回结果的时间，对应语音转写完成、核心临床要素提取完毕、病历模块初步生成或数据同步至病历系统的时间；

$R\_T_{start}$ ——用户发起请求的开始时间，特指医生发起语音转写、临床要素提取、病历生成或数据同步请求的时间。

#### 8.1.2 推理能力

##### 8.1.2.1 指代消解

语音电子病历生成智能体在指代消解能力上应具备一定的准确率。针对病历生成场景中医疗实体的精准匹配，核心覆盖症状、体征、检查项目、药物名称、疾病诊断、手术方式等关键临床要素的指代识

别，要求指代消解准确率不低于 96%。

计算对话系统的指代消解准确率，即多轮对话中某个轮次代词或名词可能指代多种不同事物情况下识别正确。计算公式如下：

$$P_M = \frac{m}{M} \times 100\%$$

式中：

$P_M$ ——本轮指代消歧平均准确率，特指正确识别“该症状”“此检查”“该药物”等代词或简化表述对应的具体医疗实体的次数；

$m$ ——本轮中每个代词或名词短语被正确识别的次数；

$M$ ——本轮中所有代词或名词短语数量，即对话及病历生成过程中所有医疗相关指代表述的总数量。

### 8.1.2.2 术语与标准符合性检查

语音电子病历生成智能体在术语与标准符合性检查能力上应具备一定的准确率，这其中包括识别术语错误、格式不符、缺失项检测与矛盾项识别，聚焦病历文本的术语规范与格式合规，确保生成内容符合医疗术语标准及病历书写规范，要求检查准确率不低 95%。

标准符合准确率，即检测术语使用每次是否正确，计算公式如下：

$$P_D = \frac{D_1}{D} \times 100\%$$

式中：

$P_D$ ——准确率；

$D_1$ ——正确识别术语错误、格式不符、缺失项、矛盾项等问题的轮次量；

$D$ ——测试总轮次量。

### 8.1.2.3 知识推理

语音电子病历生成智能体在知识推理能力上应具备一定的准确率，基于临床诊疗逻辑与医学知识库，围绕“症状-诊断-检查-治疗”关联链条开展推理，直接服务于病历核心内容生成，要求 F1 值不低于 93%。

根据推理总数和推理正确数，计算 F1 值：

$$F1 = \frac{2 \times P \times R}{P + R}$$

式中：

$P$ ——预测正确的数量/预测出的总数量；

$R$ ——预测正确的数量/实际总数量。

## 8.2 规划能力

### 8.2.1 任务规划

#### 8.2.1.1 目标拆解

语音电子病历生成智能体在目标拆解能力上具备一定的性能优越度，针对病历生成全流程目标，拆解为贴合临床场景的专业化子任务，确保各环节衔接适配病历书写逻辑。

- a) 目标识别认知：语音电子病历生成智能体应支持对目标进行深入认知，包括但不限于关键信息和潜在障碍，聚焦主诉、现病史、既往史、体格检查、诊断、处理意见等病历核心模块的信息采集需求，明确各模块必填要素及完整性标准；
- b) 目标分析预测：语音电子病历生成智能体应支持对目标进行分析预测，结合诊疗场景、疾病类型预判病历结构复杂、所需核心信息维度及生成优先级，包括但不限于目标概念、结构、复杂性、层次、目标间关系；
- c) 拆解关联度：语音电子病历生成智能体拆解目标时可满足拆解子目标间相关联，子任务围绕“语音采集-转写-要素提取-术语标准化-模块填充-逻辑校验-病历整合”形成闭环关联，确保前序任务结果直接服务于后序任务；
- d) 拆解合理性：语音电子病历生成智能体拆解目标时宜参考拆解子目标可行性、依赖关系和优先级，保障拆解目标及可操作性。语音电子病历生成宜参考临床诊疗流程优先级、信息依赖关系，保障子任务可操作；
- e) 拆解可解释性与可视化：语音电子病历生成智能体应支持提供拆解规划方案的详细解释和可视化展示，帮助用户或开发者理解方案的生成过程和结果。以病历生成流程为脉络，可视化展示各子任务对应的病历模块填充进度、数据来源（如语音转写、HIS系统数据）及未完成项，方便医生追踪。

### 8.2.1.2 规划策略

语音电子病历生成智能体应支持任务内或任务间的组织规划，贴合病历生成的时序性与规范性要求，支持规划方案适配临床实际操作。

- a) 规划结构性：支持采用“线性+分支”结合模式，线性流程保障病历模块生成顺序，分支流程适配特殊情况；
- b) 规划逻辑性：语音电子病历生成智能体拆解任务的合理性，确保子任务之间有明确的逻辑关系。应支持病历时间线（症状出现顺序、诊疗操作时序）、因果关系（症状-诊断-治疗）的一致性，避免规划导致的病历逻辑混乱。
- c) 规划一致性：语音电子病历生成智能体在任务间或任务内部组织规划时，应具备规划一致性和协调性，避免重复任务、死循环任务、冲突任务及无效任务等不一致问题，如病历中既往史与现病史中疾病描述一致、用药记录与诊断适配，规划过程中自动校验信息一致性；
- d) 计算规划策略有效率，即智能体为了完成指定任务给出的规划步骤中，有多少步骤是必要的有效步骤，要求有效率不低于94%。计算公式如下：

$$E_p = \frac{R_1}{R} \times 100\%$$

式中：

$E_p$ ——规划策略有效率；

$R_1$ ——完成指定任务所提供的规划策略中有效的操作数量；

$R$ ——完成指定任务提供的规划策略中总的操作数量。

## 8.2.2 任务调度

### 8.2.2.1 调度机制

语音电子病历生成智能体应支持多种任务调度机制，具备一定鲁棒性。

- a) 调度机制：语音电子病历生成智能体调度机制的可选度，优先支持“关键信息优先”调度机制，对主诉、诊断等核心病历模块的生成任务优先分配资源，同时兼容优先级调度、最早截止时间优先等；
- b) 鲁棒性：语音电子病历生成智能体在面对异常情况时应能够迅速适应并重新规划任务调度，针对语音转写中断、检查数据未同步、系统临时故障等场景，应支持自动触发数据补采、任务暂停或切换至手动录入模式，恢复后可接续生成，不重复已有内容；
- c) 自主性：系统应支持自动调度医疗健康行业智能体工作和协同，支持根据病历生成进度，自动调度HIS系统数据查询、医疗术语标准化引擎、临床逻辑校验工具等协同工作，无需人工干预。

#### 8.2.2.2 组织协调

语音电子病历生成智能体执行任务时应具有各项组织协调能力。

- a) 资源协调：语音电子病历生成智能体对时间分配、计算资源管理、数据访问与存储、多源信息整合的支持度；
- b) 任务分配：语音电子病历生成智能体应支持依据实际资源，在并发等情况下进行灵活的任务分配。如在多患者同时就诊场景下，应支持按接诊顺序及病历紧急程度，动态分配计算资源，支持并行生成多份病历且互不干扰；
- c) 进度监控：语音电子病历生成智能体应支持监控流程执行进度，并对异常情况进行报警，支持实时展示各病历模块的生成状态，对超时时限的任务触发弹窗提醒；
- d) 应急处置：当紧急事件发生，语音电子病历生成智能体应支持灵活调整任务分配策略，具备应急能力，针对数据异常或紧急诊疗场景，应支持快速切换至极简模式，优先生成核心病历信息，后续补充完整内容。

### 8.3 记忆能力

#### 8.3.1 短期记忆能力

##### 8.3.1.1 提示词管理

语音电子病历生成智能体应具备提示词管理相关功能。

- a) 模板丰富度：语音电子病历生成智能体应具备多种预制的提示词模板，包含“主诉生成”“现病史结构化”“检查结果描述”“出院医嘱撰写”等病历专属提示词模板，覆盖门诊、住院不同场景及内科、外科、妇产科等专科需求；
- b) 框架丰富度：语音电子病历生成智能体应支持的提示词框架丰富度，优先支持适配医疗文书生成的专用框架，确保提示后生成内容符合病历书写规范，同时兼容ICIO、CRISPE等通用框架，保障不同提示方式下的生成效果稳定；
- c) 模板管理：语音电子病历生成智能体应具备提示词模板管理功能，支持医生根据科室特点自定义病历模块提示词模板，可创建、修改、删除并保存至个人模板库，方便复用。

##### 8.3.1.2 记忆存储

语音电子病历生成智能体应支持记忆尽量多轮次的历史对话。

- a) 历史对话轮次：在模型性能无明显下降的情况下，支持至少50轮医患诊疗对话记忆，确保完整留存单次就诊的所有语音交互内容；
- b) 记忆一致性：在长对话中，对同一医疗实体的描述全程统一，如患者年龄、既往病史、检查结



果数值，避免病历中出现信息矛盾；

- c) 矛盾校验：当记忆中检测到矛盾项，主动向医生发起澄清请求，确保病历信息准确；
- d) 上下文优化：智能体在长对话中，应支持对上下文窗口内容进行关键词提取与精简以适配大模型输入窗口，支持对长对话上下文窗口内容进行核心临床要素提取与精简，优先保留症状、病史、检查结果等关键信息，适配大模型输入窗口，不丢失病历生成必要数据；
- e) 短期存储容量：能够记住和近期对话内容的容量，支持至少10000字符的短期诊疗数据存储，满足单次复杂就诊的信息留存需求。

### 8.3.2 长期记忆能力

#### 8.3.2.1 记忆存储

语音电子病历生成智能体应支持记忆尽量多轮次的历史对话。

- a) 历史对话轮次：计算在模型性能没有明显下降的情况下，应支持患者近3年就诊的核心医疗对话记忆，在模型性能稳定的前提下，可追溯至少10次就诊的关键诊疗内容，如既往症状、治疗方案、检查结果；
- b) 知识更新：同步最新医疗术语标准、病历书写规范及临床指南，更新频率不低于每季度1次，确保长期生成的病历符合最新要求；
- c) 存储容量：支持单患者至少50000字符的长期诊疗数据存储，包含历次就诊病历核心信息、关键检查结果、诊断结论、治疗方案等，满足跨周期病历关联生成需求。

#### 8.3.2.2 快速检索

语音电子病历生成智能体应支持快速检索功能。

- a) 检索速度：从发起检索请求到返回与当前病历生成相关的历史数据，如既往病史、同类病例诊疗记录，响应时间不超过2秒；
- b) 检索准确性：应支持优先匹配与当前病历生成直接相关的核心信息，如患者既往同类症状的诊疗记录、相同诊断的病历模板，匹配度不低于90%，避免无关信息干扰；
- c) 检索覆盖范围：应支持涵盖患者历史病历、医疗术语词典、临床指南、科室专属病历模板、检查检验标准值等与病历生成直接相关的信息来源，支持多维度组合检索。

## 8.4 执行能力

### 8.4.1 虚拟环境执行能力

语音电子病历生成智能体在虚拟环境下应具备虚拟环境执行能力。

- a) 交互积极性：语音电子病历生成智能体的交互积极性，应支持可以从被动服务向主动服务转变，如主动提示医生补充病历关键缺失信息、提醒术语标准化异常，并推送个性化问诊建议；
- b) 交互对象多样性：应支持与医院HIS、LIS、PACS等核心业务系统深度交互，同时兼容医疗术语标准化工具、临床逻辑校验系统、病历模板库等专用工具，实现多系统数据无缝流转；
- c) 数据格式多样性：应支持兼容语音数据、结构化检查报告、非结构化医嘱文本、医学影像描述、PDF格式病历等多类型数据的解析与整合，直接用于病历生成；
- d) 工具丰富度：应支持调用外部工具的数量，包含语音识别引擎、医疗术语映射工具、病历模板库、临床逻辑校验工具、数据溯源工具等专用工具，且支持工具调用结果直接复用至病历对应模块。

## 8.4.2 现实环境执行能力

语音电子病历生成智能体在现实环境下应具备现实环境执行能力。

- a) 感知参数多样性：应支持医疗场景专用设备的数据采集，如诊室语音采集设备、便携式检查仪器数据接入，同时支持温度、湿度等环境参数感知，避免环境因素影响语音采集质量；
- b) 动态协作能力：应支持多智能体协同完成复杂病历生成，如语音转写智能体、信息提取智能体、术语标准化智能体、病历整合智能体分工协作，确保各环节高效衔接；
- c) 调用接口能力：具备病历生成专属接口体系，应支持与医院HIS、LIS、PACS等核心业务系统接口，支持与医疗术语标准库、病历存储系统接口；
- d) 计算执行规划策略准确率，即智能体为了完成指定任务给出的规划步骤中，有多少步骤执行后得到了正确的结果。计算公式如下：

$$P_p = \frac{C_1}{C} \times 100\%$$

式中：

$P_p$ ——规划策略准确率；

$C_1$ ——完成指定任务所提供的规划策略中得到正确结果的操作数量；

$C$ ——完成指定任务提供的规划策略中总的操作数量。

## 9 易用性要求

### 9.1 可理解性

#### 9.1.1 语言表达清晰程度

语音电子病历生成智能体界面文字、提示及交互内容应简洁准确，避免生僻医学术语。必要术语需附通俗释义，患者端信息表述需无歧义。

- a) 生成病历文本：应支持严格遵循《病历书写基本规范》，使用标准医学术语，术语应用准确，避免模糊表述。病历结构清晰，模块划分明确，段落层次分明，关键信息用加粗或颜色标注，便于医生快速审阅与修改；
- b) 交互提示语言：应支持精准指向病历生成操作，表述简洁明确，避免模糊表述；
- c) 专业术语注释：应支持对非通用专科术语，在医生端提供悬浮窗注释。
- d) 数据来源标注：应支持生成病历中自动标注关键信息的数据来源，辅助医生理解病历内容依据，增强可信度。

#### 9.1.2 辅助理解手段

语音电子病历生成智能体涉及复杂医学知识、操作流程等操作宜辅以图文、动画或视频展示；关键环节宜设引导提示，提升信息理解效率。

### 9.2 易学性

#### 9.2.1 帮助文档完整性

语音电子病历生成智能体应配备结构化帮助文档，含功能说明、操作指南及常见问题解答，支持关键词检索，内容随平台更新同步修订。

### 9.2.2 差错信息易理解性

语音电子病历生成智能体操作错误或系统异常时，差错信息应明确原因并提供解决方案，不应以技术代码表述。

## 9.3 易操作性

### 9.3.1 操作一致性

语音电子病历生成智能体各功能模块操作逻辑、交互样式应保持统一，降低用户学习成本。

### 9.3.2 消息明确性

- a) 语音电子病历生成智能体向用户推送的各类消息，如检查提醒、复诊通知、用药提示等，内容应明确具体，包含关键信息，如时间、地点、注意事项等。
- b) 语音电子病历生成智能体向用户推送的各类消息的标题和正文应简洁明了，不应使用冗长复杂的表述。
- c) 语音电子病历生成智能体消息推送应具备合理的频率和时机，不应过度打扰用户。

### 9.3.3 辅助输入手段

语音电子病历生成智能体应支持智能联想、语音、手写等多种输入方式。

## 10 安全性要求

### 10.1 基础设施安全

#### 10.1.1 硬件设备安全性

语音电子病历生成智能体涉及的硬件设备（如网络设备、存储设备、计算设备等）的安全防护能力应包含：

##### a) 通用安全要求：

- （1）应满足物理安全保障要求，包含防火、防雷、防水、灾备、授权等；
- （2）应满足功能安全保障要求，包含设备标签、硬件接口安全、固件安全、驱动程序安全等；
- （3）应满足管理安全保障要求，包含管理机制、管理人员等；

##### b) 网络设备安全专用要求：分布式训练、推理时应满足组网安全保障要求，包含网络带宽、网络时延、网络丢包率、网络抖动等；

##### c) 计算设备安全专用要求：

（1）应具备保障人工智能加速芯片应具备通用安全保障能力，包含 AI 加速芯片信息窃取防护、架构安全漏洞防护等；

（2）应具备保障人工智能加速芯片在异构场景下应具备稳定运行的能力，包含 CPU 与 GPU 相结合的场景；

（3）应具备保障人工智能加速芯片运行环境安全的能力。

#### 10.1.2 软件设备安全性

语音电子病历生成智能体应支持多种设施如依赖库、AI 框架、向量数据库、中间件、接口等具备

安全防护能力，包含：

- a) 漏洞管理：软件设施应定期进行漏洞扫描和修复，具备完善的漏洞响应机制；
- b) 安全更新：软件设施应及时更新安全补丁，以防止新出现的安全威胁。

## 10.2 数据安全

语音电子病历生成智能体应支持数据采集、数据预处理、数据使用等数据相关内容具备安全防护能力，包含：存储安全、隐私保护、过程安全、销毁安全等。

## 10.3 应用安全

### 10.3.1 内容安全

语音电子病历生成智能体输出内容（含生成内容、决策内容）应符合全人类普适的道德伦理及医学伦理要求。

- a) 应支持尊重人权，包括医疗健康行业智能体输出内容（含生成内容、决策内容）应遵循人权的普遍性和不可侵犯性的原则，尊重人类平等、尊严和自由的权利；
- b) 应支持无偏见歧视性，包括医疗健康行业智能体输出内容（含生成内容、决策内容）避免产生偏见及歧视性结果的程度；
- c) 应符合科技伦理原则，包括增进人类福祉、坚持公平公正、推动透明可释、确保可控可信等；
- d) 应遵循科技伦理指标，包括公平性、透明可释性、数据隐私、可控可靠性、内容向善、责任可追溯、可持续性等。

### 10.3.2 服务安全

语音电子病历生成智能体应支持服务安全可信、内容安全可信等应用相关内容具备安全防护能力，包含：

- a) 服务安全：医疗健康行业智能体涉及的模型安全性应满足模型安全保障要求，包含 MTTF、服务安全性、服务合规性、反馈处置机制等。

## 参 考 文 献

- [1] 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》
-