

团 体 标 准

T/ISC XXX—XXXX

医疗健康行业智能体 智能语音随访技术 要求

Intelligent agents in the healthcare industry

Requirements for intelligent voice follow-up technology

在提交反馈意见时，请将您知道的相关专利与支持性文件一并附上。

（征求意见稿）

2025-11-01

XXXX – XX – XX 发布

XXXX – XX – XX 实施

中 国 互 联 网 协 会 发 布

目 次

前 言	3
医疗健康行业智能体 智能语音随访技术要求	4
1 范围	4
2 规范性引用文件	4
3 术语和定义	4
3.1 智能体 AI agent	4
3.2 医疗健康行业智能体 healthcare AI agent	4
3.3 医院信息系统 (his) hospital information system	4
4 总体要求	5
5 功能要求	5
5.1 完备性	5
5.1.1 专病随访	错误！未定义书签。
5.1.2 满意度随访	7
5.1.3 满意度结果分析	7
5.1.4 科研随访	9
6 准确性要求	9
6.1 二分类任务	9
6.2 文书生成类任务	10
6.3 多分类任务	11
6.4 语音合成任务	12
7 智能体能力要求	12
7.1 感知能力	12
7.1.1 回答时效性	12
7.1.2 推理能力	12
7.1.2.1 指代消解	12
7.1.2.2 术语与标准符合性检查	13
7.1.2.3 知识推理	13
7.1.3 语音识别能力	13
7.2 规划能力	13
7.2.1 任务规划	13
7.2.1.1 目标拆解	13
7.2.1.2 规划策略	14
7.2.2 任务调度	14
7.2.2.1 调度机制	14
7.2.2.2 组织协调	14
7.2.3 路径规划	14
7.2.3.1 病种路径	15

7.2.3.2 患者个性化路径	15
7.2.4 资源规划	15
7.2.4.1 人力资源	15
7.2.4.2 算力资源	15
7.2.5 风险规划	15
7.2.5.1 医疗风险	15
7.2.5.2 业务风险	15
7.3 记忆能力	15
7.3.1 短期记忆能力	15
7.3.1.1 提示词管理	15
7.3.1.2 记忆存储	16
7.3.2 长期记忆能力	16
7.3.2.1 记忆存储	16
7.3.2.2 快速检索	16
7.4 执行能力	16
7.4.1 虚拟环境执行能力	16
7.4.2 现实环境执行能力	16
7.4.3 计算规划策略准确率	17
8 易用性要求	17
8.1 可理解性	17
8.1.1 语言表达清晰程度	17
8.1.2 辅助理解手段	17
8.2 易学性	17
8.2.1 帮助文档完整性	17
8.2.2 差错信息易理解性	17
8.3 易操作性	17
8.3.1 操作一致性	17
8.3.2 消息明确性	17
8.3.3 辅助输入手段	18
9 安全性要求	18
9.1 基础设施安全	18
9.1.1 硬件设备安全性	18
9.1.2 软件设备安全性	18
9.2 数据安全	18
9.3 应用安全	18
9.3.1 内容安全	18
9.3.2 服务安全	19
参考文献	19

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。

本文件由工业和信息化部人工智能标准化技术委员会提出，中国通信标准化协会归口。

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、四川大学华西医院、安徽医科大学第一附属医院、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：

医疗健康行业智能体 智能语音随访技术要求

1 范围

本文件规定了医疗健康行业智能体 智能语音随访在应用过程中涉及的技术能力，从系统功能、使用场景、功能、智能体能力、易用性及安全性等维度对智能体技术在术后管理的语音随访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 41813.1-2022 信息技术 智能语音交互测试方法 第1部分：语音识别
GB/T 41813.2-2022 信息技术 智能语音交互测试方法 第2部分：语义理解
GB/T 36464.1-2020 信息技术 智能语音交互系统 第1部分：通用规范
GB/T 36464.3-2018 信息技术 智能语音交互系统 第3部分：智能客服
GB/T 35312-2017 中文语音识别终端服务接口规范
GB/T 34145-2017 中文语音合成互联网服务接口规范
GB/T 21023-2007 中文语音识别系统通用技术规范
GB/T 21024-2007 中文语音合成系统通用技术规范
YD/T 6481-2025 智能语音助手服务接入指南

3 术语和定义

下列术语和定义适用于本文件。

3.1

智能体 AI agent

又称人工智能代理，是指驻留在某一环境下，能持续自主地发挥作用，具备驻留性、反应性、社会性、主动性等特征的计算实体。

3.2

医疗健康行业智能体 healthcare AI agent

在通用智能体的基础上，结合医疗健康行业特点设计的智能体，与医疗健康相关任务的适配度较高。

3.3

医院信息系统 (his) hospital information system

医院信息化建设的核心组成部分，为管理和运营医院而设计和开发的一套综合性的信息系统。

4 总体要求

智能语音随访系统的设计和实现应以用户需求为导向、以目标达成成为宗旨、以任务实现为要点，符合“高内聚、低耦合”的设计理念，满足应用及服务的扩展与集成，在保持各功能模块独立性的同时，也能促使模块间的相互统一、数据共享与互操作。

系统架构分为基础层、能力层和应用层。基础层通过虚拟化技术，实现对基础运行环境的监控和资源调度，为平台提供计算和存储服务支持，是平台正常运行的基础。能力层集成了大模型能力、语音合成、语音识别、自然语言理解等核心模块，为平台输出智能化能力。应用层实现与用户交互界面，支持数据输入、方案配置、模板定制、报表展示等操作功能，同时提供任务调度、外呼过程管理、业务日志管理、交互质量控制、呼叫流程管理等服务。



图1 智能随访系统架构图

5 功能完备性要求

5.1 专病随访

5.1.1 患者信息管理

系统可通过对接his拉取患者信息，也可通过新增、导入两种方式创建患者信息。信息包含患者姓名、电话、性别、年龄等基本信息，以及主要诊断、就诊时间、主治医生等就诊信息。

5.1.2 患者档案查询

系统应支持查询患者历次的门诊、住院记录，以及针对患者全周期的随访记录。

5.1.3 患者随访计划管理

可根据患者的主要诊断、就诊时间自动创建对应的随访计划，也可由医生、护士动态创建患者的随访计划。系统应具有以下基础功能。

——支持在院患者、门诊/出院患者列表中选中需要随访的患者，人工创建随访计划；支持医生/护士人工创建患者的随访计划，可设置即时发送任务、定时发送任务等，实现更为灵活的患者随访管理。

——支持设置随访任务的任务节点、执行路径、执行时间、重拨设置、是否发送短信、短信模板选择等。

5.1.4 随访管理

支持管理多种形式的随访及其过程管理，应包括但不限于以下单体或组合随访形式：

——智能语音外呼随访：智能外呼系统可模拟医生护士与患者进行互动。在过程中系统会理解对方说话的意图并对应回复，并通过自然语言理解和大数据分析，将对话中的关键信息进行结构化处理。针对各个科室的出院随访、门诊随访、满意度调查等业务需求，可定制电话外呼话术，形成后期的话术模板。

——短信随访：针对患者宣教、复诊提醒等通知场景，可使用短信快速群发短信，让患者及时了解信息。支持用户查看短信的发送状态、发送记录。

——问卷随访：针对信息采集数量较多、较复杂的随访场景，如科研随访，可通过电话、短信、问卷组合的随访模式，来解决复杂类信息的采集，弥补电话短信的短板，丰富应用场景。此外，系统可结合电话提醒患者去填写问卷，以提高问卷的填写率。

——人工随访：针对一些特殊患者，如精神内科、老干部病房等，需要使用人工随访。系统可安排此类科室的需求，配置人工随访的模板，设计人工随访的问题、选项，人工在随访时可快速记录随访内容。

5.1.5 病种随访路径管理

针对各个科室中存在患者在就诊后的一段时间需要多次随访，系统可一次性设置各个病种多次随访的时间点和随访内容，届时可快速给患者调用该随访路径。随访路径的各个节点支持电话、短信、问卷以及人工随访的灵活组合，应具备以下功能。

——支持按照科室维度展示专病随访路径。

——支持展示路径名称、版本号、病种名称等信息。

——支持列表按照病种、路径名称、版本号筛选路径；

——支持编辑、删除专病路径。

5.1.6 随访异常结果自动提醒

根据患者是否接通电话、患者在人机交互中是否回答有出现病情未控制的情况，系统可将此异常情况自动筛查出来，生成待办任务，由护士进一步人工跟进处理。

系统支持将随访过程中的一些异常情况以及需要人工进行随访的任务单独展示出来，数据概览的首页也会展示快捷入口，提醒医护人员及时处理，以提高随访的效果，系统应具备以下基础功能。

——支持以列表形式展示有随访异常任务的患者，根据异常是否处理分组为：未处理、已处理；

——支持根据患者姓名电话、就诊类型、诊断、主治医生、科室名称等维度的列表筛选；

——支持在详情页面查看患者已随访内容并再次进行人工随访，可标注随访状态。

5.1.7 随访结果统计分析

针对医生/护士用户，系统会智能分析每次外呼结果，支持用户查看历次外呼信息，包括外呼时间，接通人数，应用话术等，以及每次外呼的具体信息，包括外呼接通情况，患者回复的分类结果，患者回答的明细内容等。还支持查看每个患者的详细信息，听取互动录音，查验调查结果。

针对监管侧用户，系统支持查阅各个科室、随访人员的智能语音随访、短信随访、问卷随访、人工随访的情况，了解随访覆盖率、随访效果等指标。

5.2 满意度随访

5.2.1 患者信息管理

系统应支持患者信息管理，具备以下基础功能。

——支持对门诊的患者信息进行增删查等管理，主要包含患者姓名、电话、就诊科室、就诊时间等信息。

——支持对出院的患者信息进行增删查等管理，主要包含患者姓名、电话、就诊科室、入院时间、出院时间等信息。

——支持通过Excel表格的方式导入患者名单至满意度系统。

——支持对接医院his系统，可每天自动拉取当天就诊的患者信息。

5.2.2 满意度任务管理

系统应支持满意度任务管理，具备以下基础功能。

——支持设置“全部随访”“按人数随机抽查”“按比例随机抽查”设置随访人数。

——支持设置“XX天内不可重复外呼”，天数支持自定义。

——支持选择患者，手动发起满意度调查任务。

——支持据患者的就诊信息，结合满意度调查方案，自动发起满意度调查任务。

——支持AI外呼进行满意度调查，AI电话任务包括电话模板选择、执行时间、电话重拨次数、电话重拨间隔等。

——支持问卷方式进行满意度调查，包括问卷模板选择、执行时间、有效期内的短信/电话提醒设置等。

5.2.3 满意度方案管理

满意度方案管理应具备以下基础功能。

——支持自定义编辑随访内容，选择执行方式、模板、执行时间、重拨次数、短信提醒模板等。

——支持对新建方案保存，并可以一键复用。

——支持定制智能外呼话术，如患者整体满意度情况、考核医院有无收受红包、有无制定院外就诊、有无院外购药、考核患者不满意原因等，并针对话术进行单独识别引擎与知识图谱优化，形成后期的话术模板。

——支持满意度问卷发送，可自定义创建问卷内容，包含单选题、多选题、填空题、评分题，可设置题目之间的关联。

——支持短信发送问卷链接，支持为每个患者生成单独的问卷地址，支持电话提醒用户填写。

——支持问卷预览、暂存、复制、编辑、删除、搜索等。

5.3 满意度结果分析

5.3.1 自动生成满意度报表

自动生成满意度应具备以下功能。

——支持根据统计结果自动生成满意度报表，按照科室展示呼叫人次、接通人次、接通率等。

——支持按照外呼时间以及考核周期（月、季度、年、自定义周期）查看满意度报表。

——支持各科室满意度调查明细查看，针对满意度问题中各问题回答明细下钻查看，并支持明细导出。

根据满意度结果自动生成的满意度报表，可以直观的展示各科室的满意度人次、接通率等。并支持各科室的详情查看，支持满意度问题下钻等。利于监管人员全局把控院内各科室的满意度情况。

——支持根据统计结果自动生成满意度报表，按照科室展示呼叫人次、接通人次、接通率等。

——支持按照外呼时间以及考核周期（月、季度、年、自定义周期）查看满意度报表。

——支持各科室满意度调查明细查看，针对满意度问题中各问题回答明细下钻查看，并支持明细导出。

——支持按门诊和出院分别统计调查的患者人数、接通人数、接通率。

5.3.2 整体调查情况分析

整体调查情况应支持按门诊和出院分别统计调查的患者人数、接通人数、接通率。

5.3.3 科室调查情况排名

系统应支持科室调查情况排名功能。

——支持按照科室维度统计满意度调查结果，对科室满意度进行排名。

——支持对不满意原因进行统计分析，如护士长满意度，医生服务态度、医生技术等

——支持对不满意医护人员、满意医护人员统计出现次数。

——支持科室满意度报表的导出。包括科室名称、调查人数、接通率等。

——支持对导出报表的表头进行配置。

5.3.4 调查内容分类

系统应支持调查内容分类功能。

——支持按照话术或者问卷问题对结果进行分类和统计，例如对医院整体服务、医生护士态度、技术水平、排队时间、医院环境、有无收受红包等的评价等。

——支持收集患者提出的不满意建议归类。例如：环境不好、饭菜不好等。

——支持对不满意患者人工电话随访以完成满意度调查闭环。

5.3.5 电话接通情况分析

系统应支持电话接通情况分析，具备以下基础功能。

——支持按照任务时间段进行筛选，展示呼叫人数、接通人次、整体接通率等。

——支持按照接通结果分类进行统计，包括正常回答、自动留言、接通不便、号码错误、中断、不愿配合等。

5.3.6 未接通原因分析

系统应支持针对未接通电话系统自动分类，包括：停机、号码错误、无应答等。

5.3.7 满意度问卷任务统计

系统应支持将患者填写的满意度问卷信息进行汇总展示，使医生护士能够查看满意度问卷的填写率、满意度评价情况。

5.3.8 全员满意度情况

对全员满意度情况分析，系统应具备以下功能。

- 支持分别对门诊和出院满意度数据情况进行统计分析。展示门诊、出院的累计服务人数、满意度环比。
- 支持展示门诊、出院患者满意度前5科室排名。
- 支持展示门诊、出院患者不满意原因分析。例如护士态度、医生服务态度等
- 支持通过柱状图展示就诊人数、调查人数、有效随访人数的趋势。
- 支持通过曲线图展示满意度得分的趋势。
- 支持自定义时间，月度、季度、年度的数据统计。

5.3.9 生成满意度分析报告

- 系统应支持生成满意度分析报告，具备以下基础功能。
- 支持根据使用的满意度话术生成医院或院区分析报告，并支持报告导出。
 - 支持根据使用的满意度话术生成科室整改报告，并支持报告导出。

5.3.10 科研随访

- 科研患者队列分组：通过科研随访管理，将科研患者加入队列并分组，维护科研管理路径和随访指标，配置指标对应的采集方式（AI 电话，问卷量表，指标监测，人工随访），创建科研随访待办任务，人工与 AI 耦合协同的方式便捷采集患者院外数据，提升医护人员科研随访的效率。应具备以下功能：
- 支持新增患者进入科研随访管理，区分对照组、试验组队列。
 - 科研指标配置方面，应支持按照病种配置对应的科研随访指标采集方案；支持维护指标名称，采集来源（问卷、AI电话、人工、患者上传指标方式采集）等。
 - 随访路径创建和执行方面，应支持制作科研随访路径，支持配置对照组和试验组随访任务，生成科研随访人员待办处理任务。
 - 人机耦合任务处理和指标查看方面，应支持人工处理科研随访任务，支持人工拨打电话确认和指标确认和修改；支持查看历次随访任务采集记录；支持查看科研患者采集数据，支持患者指标数据导出。

6 准确性要求

6.1 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关干扰语音、特定症状表述的有无判断等场景。

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）

表 1 二分类混淆矩阵

a) 灵敏度 (Sen)

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

式中:

Sen——灵敏度

TP——真阳性

FN——假阴性

b) 特异度 (Spe)

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

式中:

Spe——特异度

TN——真阴性

FP——假阳性

c) 准确率 (Acc)

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

式中:

Acc——准确度

TP——真阳性

TN——真阴性

FP——假阳性

FN——假阴性

6.2 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) ROUGE-N: 对生成任务，计算客观指标ROUGE-N，其计算公式如下:

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

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) 关键词命中率

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

式中:

$\text{Count}_{\text{Hit}}$ ——机器生成文本中命中关键词的数量

$\text{len}(\text{keyword})$ ——关键词的数量

c) BERTScore: 对生成任务, 计算客观指标BERTScore, 计算公式如下:

$$\begin{aligned} \text{sim}(x_i, y_i) &= \frac{\text{Emb}(x_i) \cdot \text{Emb}(y_i)}{\|\text{Emb}(x_i)\| \|\text{Emb}(y_i)\|} \\ \text{Precision} &= \frac{1}{|x|} \sum_{x_i \in x} \max_{y_i \in y} \text{sim}(x_i, y_i) \\ \text{Recall} &= \frac{1}{|y|} \sum_{y_i \in y} \max_{x_i \in x} \text{sim}(x_i, y_i) \\ \text{BERTScore} &= \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \end{aligned}$$

式中:

$\text{Emb}(x_i)$ ——句子 x 中词语在经过编码器后的嵌入向量;

$\text{Emb}(y_i)$ ——句子 y 中词语在经过编码器后的嵌入向量;

$\text{sim}(x_i, y_i)$ ——两词语嵌入向量的余弦相似度;

Precision——精确率;

Recall——召回率。

6.3 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, 例如疾病诊断分型、症状严重程度分级、检查结果异常等级判定等多类别区分场景, 保障分类结果的精准可靠。

a) Macro-Precision: 宏精准率

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

式中:

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 $\text{Precision}_i = \frac{\text{TP}_i}{\text{TP}_i + \sum_{j \neq i} \text{FP}_{j \leftarrow i}}$ (TP_i 为 i 类真阳性数, $\text{FP}_{j \leftarrow i}$ 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++

$$\text{Macro} - R = \frac{1}{K} \sum_{i=1}^K \text{Recall}_i$$

式中:

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 $\text{Recall}_i = \frac{\text{TP}_i}{\text{TP}_i + \sum_{j \neq i} \text{FN}_{i \rightarrow j}}$ (TP_i 为 i 类真阳性数, $\text{FN}_{i \rightarrow j}$ 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

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

6.4 语音合成任务

系统采用科大讯飞语音合成系统提供语音合成能力。语音合成的准确性要求是通过一套由客观量化指标和主观评估标准共同构成的体系来定义和衡量的。指标从不同维度（内容、发音、韵律）评估准确性，主要体现在三个层面：文本内容准确性、发音准确性和韵律准确性。

a) Word Error Rate: 词错误率

$$WER = \frac{(Ins + Del + Sub)}{Ref_Words} \times 100\%$$

式中：

Ins——插入错误数，指生成语音经识别后，出现了原文中没有的词；

Del——删除错误数，指生成语音经识别后，遗漏了原文中有的词；

Sub——替换错误数，指生成语音经识别后，原文中的词被错误替换为其他词；

Ref_Words——参考文本的总词数，即原始输入文本的总词数。

b) Phoneme Error Rate: 音素错误率

$$PER = \frac{Ins(Ph) + Del(Ph) + Sub(Ph)}{Ref_Phonemes} \times 100\%$$

式中：

Ins (Ph)——插入错误音素数，指生成语音中出现了标准音素序列里没有的音素）

Del (Ph)——删除错误音素数，指生成语音中遗漏了标准音素序列里有的音素）

Sub (Ph)——替换错误音素数，指生成语音中某个音素被错误替换为其他音素）

Ref_Phonemes——参考音素总数，即输入文本经G2P转换后得到的标准音素序列总数量）

c) Mean Opinion Score: 主观自然度评分

使用平均意见得分评估语音自然度和整体质量。评估者聆听语音样本，并给出1-5分的综合评分（1分最差，5分最好），最终取所有评分的平均值。商用级TTS系统的MOS评分通常要求在4.5分以上。

7 智能体能力要求

7.1 感知能力

7.1.1 回答时效性

医疗健康行业智能体应具备一定的回答时效性。

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}——医疗健康行业智能体返回结果的时间；

R_T_{start}——用户发起请求的开始时间。

7.1.2 推理能力

7.1.2.1 指代消解

医疗健康行业智能体在指代消解能力上应具备一定的准确率。

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；

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。

7.1.2.2 术语与标准符合性检查

医疗健康行业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。

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：

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

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。

7.1.2.3 知识推理

医疗健康行业智能体在知识推理能力上应具备一定的准确率。

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。

7.1.3 语音识别能力

医疗健康行业智能体在语音识别能力上应具备一定的准确率。

应用识别的基础错误率计算语音识别能力，公式如下：

$$CER = \frac{S + D + I}{N} \times 100\%$$

式中：

S ——替换错误字符数；

D ——删除错误字符数；

I ——插入错误字符数；

N ——参考文本的总字符数。

7.2 规划能力

7.2.1 任务规划

7.2.1.1 目标拆解

医疗健康行业智能体在目标拆解能力上具备一定的性能优越度。

- a) 目标识别认知：医疗健康行业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；
- b) 目标分析预测：医疗健康行业智能体应支持对目标进行分析预测，包括但不限于目标概念、结构、复杂性、层次、目标间关系等；
- c) 拆解关联度：医疗健康行业智能体拆解目标时可满足拆解子目标间相关联；
- d) 拆解合理性：医疗健康行业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；
- e) 拆解可解释性与可视化：医疗健康行业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。

7.2.1.2 规划策略

医疗健康行业智能体应支持任务内或任务间的组织规划。

- a) 规划结构性：支持按照一定的结构的任务结构进行任务规划，如线性、分层、并行、树状、网状、条件、迭代等；
- b) 规划逻辑性：医疗健康行业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；
- c) 规划一致性：医疗健康行业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；
- 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。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；

R ——完成指定任务提供的规划策略中总的操作数量；

7.2.2 任务调度

7.2.2.1 调度机制

医疗健康行业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：医疗健康行业智能体调度机制的可选度，如先来先服务、短作业优先、轮转调度机制、优先级调度机制、最早截止时间优先等；
- b) 鲁棒性：医疗健康行业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，如自动干预及手动干预；
- 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。

7.2.2.2 组织协调

医疗健康行业智能体执行任务时应具有各项组织协调能力。

- a) 资源协调：医疗健康行业智能体对时间分配、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：医疗健康行业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；
- c) 进度监控：医疗健康行业智能体应支持监控流程执行进度，并对异常情况进行报警；
- d) 应急处置：当紧急事件发生，医疗健康行业智能体应支持灵活调整任务分配策略，具备应急能力。

7.2.3 路径规划

7.2.3.1 病种路径

医疗健康行业智能体应具备病种专属路径规划能力。

- a) 预设标准化路径：覆盖多种专病，每个路径点包含“随访时间点、提问内容、异常触发条件”；
- b) 路径可视化配置：支持医院自定义路径，并支持“节点增删、话术修改”。

7.2.3.2 患者个性化路径

医疗健康行业智能体应具备患者个性化路径规划能力。

- a) 病史关联调整：基于患者多维度病史整合分析，打破单一病种随访的局限性，实现“以患者为中心”的个性化、全维度随访路径动态适配，通过病史关联确保随访服务覆盖患者所有关键健康风险点，符合临床诊疗中“合并症综合管理”的专业要求。

7.2.4 资源规划

7.2.4.1 人力资源

医疗健康行业智能体应具备人力资源优化规划能力，减少无效人力投入。

- a) 人力替代：智能体承担主要常规随访，释放医护人员专注于异常病例处理，由护士重点跟进；
- b) 人力调度：自动将“未接通、病情异常”的随访任务汇总为待办列表，按科室归属及紧急程度分配给对应医护，避免人力浪费。

7.2.4.2 算力资源

医疗健康行业智能体应具备算力资源规划能力，平衡效率与成本。

- a) 分层部署：能力层采用云端部署，按需调用算力，降低本地硬件成本；应用层采用本地部署，保障数据安全；
- b) 并发调度：支持动态并发路数调整，避免算力闲置。

7.2.5 风险规划

7.2.5.1 医疗风险

医疗健康行业智能体应具备医疗风险规划能力。

- a) 危急值实时预警：若患者产生紧急反馈或系统对接发现患者异常情况，系统及时触发预警并生成紧急处理待办；
- b) 异常趋势预警：对慢病患者，若连续随访发现病情隐患因素，自动生成异常标记，提醒医生调整治疗方案。

7.2.5.2 业务风险

医疗健康行业智能体应具备业务风险管控规划能力。

- a) 质量抽检：自动抽取随访录音，检测话术合规性、语义理解准确性，生成质量报告；
- b) 合规追溯：所有随访记录（语音、文本、数据）统一进行保存，支持溯源。

7.3 记忆能力

7.3.1 短期记忆能力

7.3.1.1 提示词管理

医疗健康行业智能体应具备提示词管理相关功能。

- a) 模板丰富度：医疗健康行业智能体应具备多种预制的提示词模板，如文本生成类、知识问答类、逻辑推理类等；

- b) 框架丰富度: 医疗健康行业智能体应支持的提示词框架丰富度, 即在不同框架提问下效果稳定, 如ICIO 框架、CRISPE 框架、BROKE 框架等;
- c) 模板管理: 医疗健康行业智能体应具备提示词模板管理功能, 如创建、修改、删除等;

7.3.1.2 记忆存储

医疗健康行业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, 医疗健康行业智能体最长可以支持的历史对话轮次;
- b) 记忆一致性: 智能体在长对话中, 对同一实体的描述应保持一致。例如, 对患者年龄、既往病史的提及应全程统一;
- c) 矛盾校验: 智能体在长对话中, 在查询到矛盾项时应主动向医生发起澄清请求;
- 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;
- e) 短期存储容量: 能够记住和近期对话内容的容量。

7.3.2 长期记忆能力

7.3.2.1 记忆存储

医疗健康行业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, 医疗健康行业智能体最长可以支持的历史对话轮次;
- b) 知识更新: 医疗健康行业智能体知识更新频次与质量;
- c) 存储容量: 能够记住和存储的长期痕迹字符数量。

7.3.2.2 快速检索

医疗健康行业智能体应支持快速检索功能。

- a) 检索速度: 医疗健康行业智能体从接收到查询请求到返回检索结果所需的时间;
- b) 检索准确性: 医疗健康行业智能体返回的检索结果与用户查询意图的匹配程度;
- c) 检索覆盖范围: 医疗健康行业智能体能够检索到的信息来源和类型。

7.4 执行能力

7.4.1 虚拟环境执行能力

医疗健康行业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: 医疗健康行业智能体的交互积极性, 应可以从被动服务向主动服务转变;
- b) 交互对象多样性: 医疗健康行业智能体与软件环境中的其他实体进行交互的支持度, 其他实体包括其他agent、系统、环境本身等;
- c) 数据格式多样性: 医疗健康行业智能体需要对接收到的软件环境信息进行理解和解码的能力, 环境数据包括文本及多模态数据;
- d) 工具丰富度: 医疗健康行业智能体可以调用外部工具的数量, 如文档解析、语音识别、数据库访问、图像识别等。

7.4.2 现实环境执行能力

医疗健康行业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: 医疗健康行业智能体应支持多类传感器, 如物理传感器、温度传感器、光学传感器等, 并支持多类参数感知;
- b) 动态协作能力: 多个医疗健康行业智能体间应可以支持智能体间团队协作能力。

7.4.3 计算规划策略准确率

- a) 计算规划策略准确率，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。

8 易用性要求

8.1 可理解性

8.1.1 语言表达清晰程度

医疗健康行业智能体界面文字、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。

8.1.2 辅助理解手段

医疗健康行业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。

8.2 易学性

8.2.1 帮助文档完整性

医疗健康行业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。

8.2.2 差错信息易理解性

医疗健康行业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。

8.3 易操作性

8.3.1 操作一致性

医疗健康行业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。

8.3.2 消息明确性

- a) 医疗健康行业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) 医疗健康行业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) 医疗健康行业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。

8.3.3 辅助输入手段

医疗健康行业智能体应支持智能联想、语音、手写等多种输入方式。

9 安全性要求

9.1 基础设施安全

9.1.1 硬件设备安全性

医疗健康行业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。

9.1.2 软件设备安全性

医疗健康行业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。

9.2 数据安全

医疗健康行业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。

9.3 应用安全

9.3.1 内容安全

医疗健康行业智能体输出内容（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。

9.3.2 服务安全

医疗健康行业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 34083-2017 中文语音识别互联网服务接口规范
-