ICS 35.020
CCS L70

团 体 标 准

T/ISC XXXX—XXXX

医疗健康行业智能体 电子病历质量管理智能体技术要求

Technical requirements for Electronic medical record quality control intelligent agent in the healthcare industry

（征求意见稿）

2025-11-01

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

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：

医疗健康行业智能体 电子病历质量管理智能体技术要求

1 范围

本文件规定了医疗健康行业智能体 电子病历质量管理智能体在应用过程中涉及的技术能力，从功能要求、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电子病历质量管理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。

T/CHIA 39-2023 互联网诊疗线上病历质量管理标准

T/GDWJ 030-2025 大语言模型在患者智能交互中应用技术指引

YD/T 4899-2024 面向数字医院的医疗设备管理平台技术要求

IEEE P3394 大语言模型智能体界面标准（Standard for Large Language Model Agent Interface）

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 病历 medical record

医疗机构的医务人员对门诊、住院患者（或保健对象）临床诊疗和指导干预等医疗活动过程中形成的，记录服务对象的基本信息、病史、检查检验、处方、病情、诊断、治疗、护理等过程的文件，包括运行及归档病历资料。

4 缩略语

下列缩略语适用于本文件。

BROKE：背景，角色，目标，关键结果，改进提示词模板（Background，Role，Objectives，Key Result，Evolve）

CPU：中央处理器（Central Processing Unit）

CRISPE：角色扮演与模拟，情境模拟，个性化互动，多样化输出提示词模板（Capacity and Role，Insight，Statement，Personality，Experiment）

EHR：电子健康档案（Electronic Health Record）

EMR：电子病历（Electronic Medical Record）

GPU：图形处理器（Graph Processing Unit）

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

ICIO：任务，背景，输入数据，输出格式法提示词模板（Instruction，Context，Input Data，Output Indicator）

MCP：模型上下文协议（Model Context Protocol）
PACS：图像储存与通信系统（Picture Archiving and Communication System）

5 总体要求

医疗健康行业智能体 电子病历质量管理智能体应支持门诊病历质控、病案首页质检、住院病历质控任务并具备病历质控模型与病历质量管理平台。通过规范智能体的感知、规划、记忆、执行四大基础能力，明确医疗健康行业智能体 电子病历质量管理智能体的应用优势及业务范围，提升医疗健康服务的效率及质量。

医疗健康行业智能体 电子病历质量管理智能体应在以下方面满足要求：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、安全性要求。

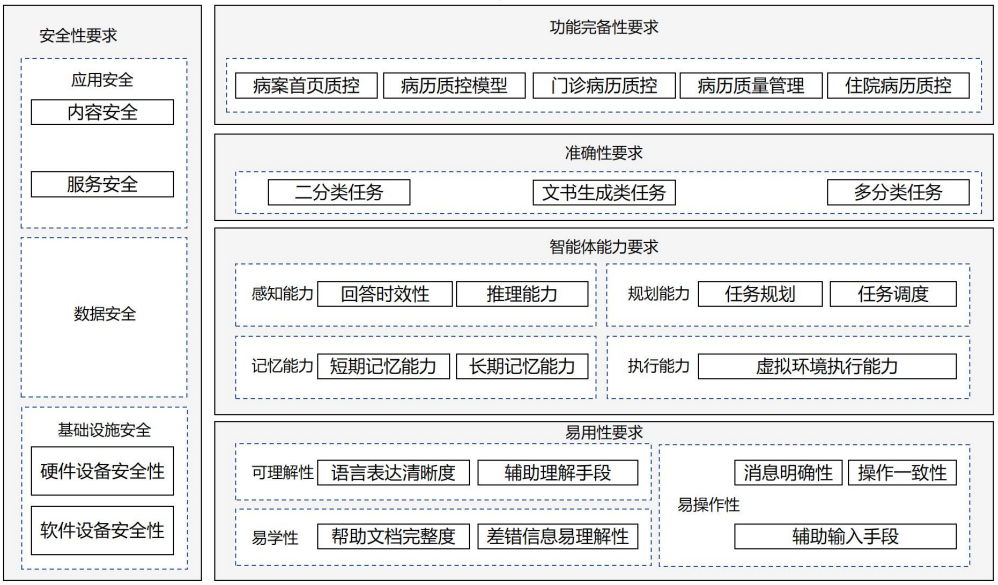

图 1 架构图

6 功能完备性技术要求

6.1 病历质控数据模型

医疗健康行业智能体 电子病历质量管理智能体按照质控要求根据不同业务场景建立文书领域模型，业务场景应包含日间病历、24h 入出院记录（死亡）、住院病历、门（急）诊病历等，文书领域模型应支持医生对质控不符合项进行询问。

6.2 门诊病历质控

医疗健康行业智能体 电子病历质量管理智能体应支持门诊病历的形式质控与内涵质控：

a) 完整性检查：应支持对门诊病历中主诉缺乏症状、时间，体格检查中缺乏一般生命体征，过敏史缺失等情况进行缺失验证与质量检测。
b) 合理性检查：应支持对主诉冗余、一般生命体征不在正常范围内、年龄性别异常等情况进行规范性质量检测。
c) 一致性检查：应支持对门诊病例中主诉与现病史症状、时间、部位不一致，诊断与性别年龄矛盾等情况进行质量检测。
d) 专科诊疗合规检查：应支持针对缺失专科阳性体征描述、诊断依据不充分、专科处置措施和诊断、检验检查报告不符或矛盾等情况进行质量检测。

6.3 病案首页质控

医疗健康行业智能体 电子病历质量管理智能体应支持病案首页的形式质控与内涵质控：

a) 完整性检查：应支持对病案首页的基本信息、病情信息（诊断信息、手术信息、其他信息）、费用信息等模块的字段完整度进行质量检测。
b) 一致性检查：应支持对病案首页各字段与参考信息、首页与入院记录字段不一致情况的质量检测。应支持对病案首页基本信息、病情信息的数据规范程度进行质量检测，应支持对病案首页与其他文书的前后矛盾情况进行质量检测。
c) 时效性检查：应支持验证病案首页在患者出院后24小时内完成的病历数占同期出院患者病历总数的比例。
d) 诊断漏写检查：应支持基于检验检查报告结果，提示疾病未在病案首页填写的质量检测。
e) 手术漏写检查：应支持对于诊疗操作类手术未在病案首页填写的质量检测。
f) 编码检查：应支持针对合并编码、编码互斥、非主诊断编码等情况进行质检并智能推荐合理编码。

6.4 住院病历质控

医疗健康行业智能体 电子病历质量管理智能体应实现住院病历全流程形式质控与内涵质控：
a) 完整性检查：应支持对全病历文书完整性与字段完整性进行质量检测，包括日常病程记录、上级医师查房记录、手术记录、术前小结、输血记录、抗生素使用记录文书。
b) 规范性检查：应支持对入院记录中不符合格式要求、内容不规范的内容进行质量检测。
c) 一致性检查：应支持对入院记录、手术记录、术后病程记录、出院记录中各字段与内容不一致情况的质量检测。
d) 时效性检查：应支持对入院记录、手术记录、出院记录等病历基于病案管理质量控制指标要求进行时效性质检。
e) 专科病历检查：应支持针对专科病历中专科检查未填专科患者检查数据的情况进行质控，包括：产科病历未记录宫高、腹围；胆囊炎病历未记录墨菲氏征。
f) 不合理复制检查：应支持针对全住院病历文书进行不合理复制质量控制。

6.5 病历质量管理

6.5.1 病历质量统计分析

医疗健康行业智能体 电子病历质量管理智能体应提供多维度的病历质量统计分析功能：
a) 质量指标监控：应支持自动统计甲级病历率、缺陷率、整改率等27项病案管理质量控制指标。
b) 趋势分析：应支持按时间维度分析病历质量变化趋势，自动识别异常波动。
c) 科室对比：应支持不同科室、病区的质量横向对比分析。
d) 医师个人质量档案：应支持医师端建立医师个人病历质量画像，支持个性化改进建议。

6.5.2 人工质控管理

医疗健康行业智能体 电子病历质量管理智能体应支持与人工质控的协同工作：
a) 质控任务分配：应支持按科室、病种、医师级别等维度分配人工质控任务。
b) 质控标准管理：应支持对于质控标准库的维护和更新，支持不同等级医院的差异化要求。
c) 质控结果反馈：应支持实时反馈人工审核意见。
d) 数据分析统计：应支持对病历质量指标数据进行趋势对比

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完整性检查、合理性检查、一致性检查、时效性检查等形式质控与内涵质控判断，聚焦缺失项与不一致项的判断，比如区分病案首页内容是否存在与查房记录、手术记录等病历内容不一致的情况。

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）

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

式中：

Sen——灵敏度

TP——真阳性

FN——假阴性

b) 特异度（Spe）

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

式中：

Spe——特异度

TN——真阴性

FP——假阳性

c) 准确率（Acc）

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

式中：

Acc——准确度

TP——真阳性

TN——真阴性

FP——假阳性

FN——假阴性

7.2 文书生成类任务

文书生成类任务准确性要求针对不符合项自动生成解答与针对医生疑问生成解答两项任务，核心是确保生成文本与参考答案的语义契合度、关键信息的命中率，贴合《门诊病历书写规范》与医院特殊病历要求等标准。

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

$$\mathrm{ROUGE-N} = \frac{\sum S \in \{\mathrm{ReferenceSummaries}\} \sum \mathrm{gram_n} \in \mathrm{sCount_{match}}(\mathrm{gram_n})}{\sum S \in \{\mathrm{ReferenceSummaries}\} \sum \mathrm{gram_n} \in \mathrm{sCount}(\mathrm{gram_n})}$$

式中：

N——即 n-gram，文本内容滑动窗口字节数，参考值为 2；

$\mathrm{Count_{match}}(\mathrm{gram_n})$——参考摘要和机器生成摘要中共有的 n-gram 的数量；

$\mathrm{Count}(\mathrm{gram_n})$——参考摘要中 n-gram 的数量；

b) 关键词命中率

$$\mathrm{Hit\_Rate} = \frac{\mathrm{Count_{Hit}}}{\mathrm{len(keyword)}}$$

式中：

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

$\mathrm{len(keyword)}$——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：

$$sim(x_i, y_i) = \frac{Emb(x_i) \cdot Emb(y_i)}{\| Emb(x_i) \| \| Emb(y_i) \|}$$

$$Precision = \frac{1}{|x|} \sum_{x_i \in x} \max_{y_i \in y} sim(x_i, y_i)$$

$$Recall = \frac{1}{|y|} \sum_{y_i \in y} \max_{x_i \in x} sim(x_i, y_i)$$

$$BERTScore = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall}$$

式中：

$Emb(x_i)$——句子 x 中词语在经过编码器后的嵌入向量；

$Emb(y_i)$——句子 y 中词语在经过编码器后的嵌入向量；

$sim(x_i, y_i)$ ——两词语嵌入向量的余弦相似度；

$Precision$——精确率；

$Recall$——召回率；

7.3 多分类任务

多分类任务准确性要求针对多维度医疗类别划分类任务，例如疾病编码质控、临床路径选择质控、紧急程度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

$$\mathrm{Macro-P} = \frac{1}{\mathrm{k}} \sum_{\mathrm{i=1}}^{\mathrm{K}} \mathrm{Precision_i}$$

式中：

K——分类任务的总类别数（如三分类时 K=3）

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

b) Macro-Recall：宏召回率

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

式中：

K——分类任务的总类别数（如三分类时 K=3）

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

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

8 智能体能力要求

8.1 感知能力

8.1.1 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智能体返回结果的时间计算响应实时性，计算方式参见式：

$$\text{ES\_T}=\text{R\_T}_{\text{finish}}-\text{R\_T}_{\text{start}}$$

式中：

ES_T ——响应时间；

$\text{R\_T}_{\text{finish}}$ ——医疗健康行业智能体返回结果的时间；

$\text{R\_T}_{\text{start}}$ ——用户发起请求的开始时间。

8.1.2 推理能力

8.1.2.1 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。

计算智能体的指代消解准确率，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：

$$\text{P}_{\text{M}}=\frac{\text{m}}{\text{M}}\times 100\%$$

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。

8.1.2.2 医疗命名实体识别与术语理解能力

医疗健康行业智能体应具备强大的医疗命名实体识别能力，能够准确识别并理解病历文本中的关键医学术语、标准化编码及机构规范。计算公式如下：

$$\text{P}_{\text{M}}=\frac{\text{D}_1}{\text{D}}\times 100\%$$

式中：

P_{D}——准确率；

D_1——正确识别术语的轮次量；

D——测试总轮次量。

8.1.2.3 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。

8.1.2.4 时序关系推理能力

医疗健康行业智能体在时序感知能力上应具备一定的准确率

a) 时间表达式识别：医疗健康行业智能体应能准确识别输入中绝对时间（如“2023-10-27”）和相对时间（如“入院后第3天”、“术后2小时”）的表达。

b) 时序逻辑验证：医疗健康行业智能体应能基于识别出的时间点，推理并验证事件的时序逻辑是否合理（如检查手术记录时间是否在医嘱下达时间之后、术后病程记录是否在手术时间之后）。

8.2 规划能力

8.2.1 任务规划

8.2.1.1 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，包括关键信息（如病历类型、质控项等）和潜在障碍（如数据缺失、矛盾等）；

b) 结构化分析：医疗健康行业智能体应支持对任务目标进行结构化分析，包括理解其概念、复杂性、层次以及子任务间的依赖关系（例如，住院病历质控需在完成入院记录、病程记录、手术记录等多个文书的质控后，才能进行整体一致性检查）；

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) 人机协同拆解与确认：医疗健康行业智能体应支持基于人工质控历史数据提出人工质控初步拆解方案，并允许用户进行干预、修改和确认。智能体与系统应支持交互式调整子目标的优先级、依赖关系和可行性；

f) 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。

8.2.1.2 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：医疗健康行业智能体应支持按照任务结构进行规划，如线性（顺序执行门诊病历的完整性、合理性检查）、分层（住院病历质控按文书层级分解）、并行（同时检查多个病历与病案首页的一致性）、条件（根据诊断结果触发编码检查）或迭代（对段落在多病历中进行不合理复制检查）；

b) 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；

c) 冲突解决预案：医疗健康行业智能体应支持对内部策略冲突进行预案的能力。在规划阶段应对可能出现的策略冲突节点进行预判并准备合理冲突解决预案；

d) 规划策略有效率：医疗健康行业智能体生成的规划方案应避免冗余。该指标用于衡量规划方案中，必要的核心步骤所占的比例。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

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) 调度机制多样性：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、基于医疗规则（如病历风险等级、质控项严重度）的优先级调度机制、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至重试或医师人工处理；

8.2.2.2 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应支持对时间资源、计算资源、数据资源的预估与规划；
b) 任务分配：医疗健康行业智能体在并发请求场景下，智能体应能依据实时资源状况，将任务动态分配到不同的处理单元，实现负载均衡；
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智能体行为，如不同策略出现冲突时启动冲突解决预案或提醒医师介入。

8.3 记忆能力

8.3.1 短期记忆能力

8.3.1.1 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；
d) 模板测试与验证：智能体应提供对提示词模板进行测试和验证的功能。例如，通过一组标准测试用例，评估模板是否能稳定地引导大模型输出符合质控要求的答案。

8.3.1.2 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) 上下文窗口管理：医疗健康行业智能体应具备实时监控和管理上下文长度的能力，能在预估后续生成内容触及窗口上限时进行预警并触发优化，确保在长对话下的核心性能指标不衰减；
c) 上下文优化：医疗健康行业智能体在长对话中，应支持对上下文窗口内容进行关键词提取、精简、或通过加权方式提取关键上下文信息以适配大模型输入窗口；
d) 记忆一致性：医疗健康行业智能体在多轮对话中，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；
e) 矛盾校验：医疗健康行业智能体在长对话中，在查询到矛盾项时应主动向医生发起澄清请求；
f) 短期存储容量：医疗健康行业智能体能够记住和近期对话内容的容量。

8.3.2 长期记忆能力

8.3.2.1 知识库管理

医疗健康行业智能体应支持知识库管理。

a) 质控规则知识库创建：医疗健康行业智能体应持久化存储各类文档（如病历质控规则、医学标准、临床指南）；
b) 术语知识库创建：医疗健康行业智能体应支持持久化存储医学术语与编码规则等知识；
c) 患者病历库创建：医疗健康行业智能体应支持对患者历史病历的储存；
d) 医师质控历史库创建：医疗健康行业智能体应支持对医师质控结果记录的持久化存储以支持个性化数据统计与医师个性化质量画像生成；
e) 知识库管理：医疗健康行业智能体应支持对于知识库内容的删除、更新、逻辑验证与矛盾识别以符合最新医学指南与机构规范；
f) 知识库管理接口：医疗健康行业智能体应支持图形化或自然语言接口简化知识库更新流程。

8.3.2.2 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；
d) 元数据检索：医疗健康行业智能体应支持应支持基于元数据的筛选后检索以支持对于相同病人历史病历与相似症状病历文献的快速检索；
e) 检索结果可解释性与溯源：医疗健康行业智能体返回检索结果时，应能提供结果的来源。例如，当质控提示“主诉与现病史不一致”时，应能关联并高亮显示相关质控规则或病例库中的相似案例。

8.4 执行能力

8.4.1 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应能基于医疗工作流和当前操作上下文，在适当时机主动触发相关的质控任务或提供智能提示。
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，其他实体包括其他智能体、MCP服务器、工具等；
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；
e) 系统对接能力：医疗健康行业智能体应支持通过标准接口与外部系统（如EHR/EMR、HIS、PACS）交互的能力，能够安全的读取历史和写入新病历；
f) 操作显式化：医疗健康行业智能体应支持在执行任何对医院信息系统的写入或修改操作前以清晰、无歧义的方式向负责医生显式提示即将执行的操作内容、依据与潜在影响，并设计明确的确认机制，在获得医生授权后方可执行。智能体的角色是提供决策支持与操作建议，最终的决策权与控制权应始终由医生掌握；
g) 执行容错与回退机制：医疗健康行业智能体在执行过程中应具备处理操作失败等异常情况的能力，能进行错误提示、启动备选方案或安全回退，并记录故障信息。
h) 任务执行准确率：医疗健康行业智能体应能准确执行其制定的规划方案。本指标用于衡量智能体对规划中的单个步骤的执行可靠性。计算公式如下：

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

式中：
P_p——任务执行准确率；
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。
i) 任务执行端到端准确率：医疗健康行业智能体应能可靠地完成用户指定的完整任务。本指标用

于衡量智能体在真实工作流中，从接收任务到交付最终结果的综合成功比例。计算公式如下：

$$P_{e2e} = \frac{N_{success}}{N} \times 100\%$$

式中：

P_p——任务执行端到端准确率；

$N_{success}$——被成功完成的任务数量；

N——总任务数量。

9 易用性要求

9.1 可理解性

9.1.1 语言表达清晰程度

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并以标准医学术语表示，避免口语化内容导致的歧义。

9.1.2 辅助理解手段

医疗健康行业智能体涉及院内规则、新质检规则时应显式引用相关说明辅助医师理解学习新规。

9.2 易学性

9.2.1 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。

9.2.2 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。

9.3 易操作性

9.3.1 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。

9.3.2 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。

9.3.3 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。

10 安全性要求

10.1 基础设施安全

10.1.1 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。

10.1.2 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。

10.2 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。

10.3 应用安全

10.3.1 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。

10.3.2 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。

附 录 A
（规范性）
XXX

附 录 B
（资料性）
XXX

参 考 文 献

[1]《病案管理质量控制指标》（国卫办医函〔2021〕28号）
[2]《生成式人工智能服务管理暂行办法》（国家互联网信息办公室等令 第15号）
[3]《互联网信息服务深度合成管理规定》（国家互联网信息办公室等令 第12号）
[4]《医疗机构处方审核规范》（国卫办医发〔2018〕14号）