

ICS 35.240

CCS A01

# 团体标准

T/BSIA 00X-2026

## 人工智能 智能健康术语歧义消解测试要求

Requirements for intelligent health terminology disambiguation testing for artificial intelligence

(征求意见稿)

2026-XX-XX 发布

2026-XX-XX 实施

北京软件和信息服务业协会 发布

T/BSIA 00X-2025

# 目 次

1 范围 .....	1
2 规范性引用文件 .....	1
3 术语和定义 .....	1
4 基本原则 .....	2
5 技术流程 .....	2
5.1 术语采集与预处理 .....	2
5.2 多模态特征融合 .....	2
5.3 歧义检测 .....	3
5.4 歧义消解 .....	3
5.5 标准化映射 .....	3
5.6 结果输出与反馈 .....	3
6 核心要求 .....	3
6.1 术语采集与预处理要求 .....	3
6.2 多模态语义嵌入要求 .....	3
6.3 歧义消解要求 .....	3
6.4 动态术语映射库要求 .....	4
7 质量评估 .....	4
7.1 评估指标 .....	4
7.2 评估方法 .....	4
7.3 评估周期 .....	5
8 安全保障 .....	5
8.1 数据隐私保护 .....	5
8.2 术语溯源管理 .....	5
8.3 权限控制 .....	5
8.4 数据备份与恢复 .....	5
附录 A .....	7
附录 B .....	10
参考文献 .....	13

## 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。

本文件由北京健康有益科技有限公司提出，由北京软件和信息服务业协会归口。

本文件起草单位：

本文件主要起草人：

# 人工智能 智能健康术语歧义消解测试要求

## 1 范围

本文件规定了健康术语歧义消解与标准化处理的术语和定义、基本原则、技术流程、核心要求、质量评估及安全保障。

本文件适用于医疗健康领域的智能交互系统（如在线问诊平台、健康管理 APP、智能科普工具）、电子健康档案系统、医疗数据分析平台等场景中健康术语的处理，相关系统开发者、医疗机构信息部门、健康科技企业及监管机构可参照使用。

## 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 35273-2022 信息安全技术 个人信息安全规范

GB/T 38667-2020 人工智能 术语

WS/T 447-2014 电子健康档案基本架构与数据标准

WS 363-2011 卫生信息数据元目录

WS/T 559-2017 糖尿病患者膳食指导

《国际疾病分类第十一次修订本（ICD-11）》（世界卫生组织）

《中国高血压防治指南》

《中国 2 型糖尿病防治指南》

## 3 术语和定义

下列术语和定义适用于本文件。

### 3.1

#### 健康术语 health terminology

在医疗健康领域中用于描述疾病、症状、体征、诊疗行为、药物、饮食、运动、生理指标等概念的词语或短语，如“高血压”“有氧运动”“空腹血糖”。

### 3.2

#### 术语歧义 term ambiguity

同一健康术语在不同语境下具有多种语义，或不同表述的健康术语指向同一概念的现象。前者如“运动”可指“有氧训练”或“关节活动训练”，后者如“锻炼”“活动”与“运动”表述不同但语义相近。

### 3.3

### 歧义消解 disambiguation

通过技术手段分析健康术语的上下文信息，确定其在特定场景下唯一语义的过程。

## 3.4

### 术语标准化 term standardization

将非标准健康术语映射到权威标准术语体系中的过程，确保健康术语表述的统一性和语义的准确性。

## 3.5

### 多模态语义嵌入 multimodal semantic embedding

将文本、生理信号、行为日志等多源异构数据映射到统一语义空间，形成具有上下文感知能力的术语向量表示方法。

## 3.6

### 动态术语映射库 dynamic terminology mapping library

支持实时更新和扩展的健康术语知识库，包含标准术语、歧义术语及其上下文映射关系，用于支撑术语的自动消解和标准化。

## 4 基本原则

### 4.1 权威性原则

标准化结果应基于国家卫生健康委发布的诊疗规范、世界卫生组织 ICD 标准、权威医学指南等权威来源，确保术语语义的准确性。

### 4.2 语境适配原则

歧义消解需结合用户健康状况、交互场景、任务目标等上下文信息，确保术语语义与实际应用场景匹配。

### 4.3 动态演化原则

支持新兴健康术语的增量接入与语义更新，适应健康领域知识快速发展的需求。

### 4.4 可解释性原则

术语歧义消解与标准化过程应保留关键决策依据，确保结果可追溯、可解释。

## 5 技术流程

### 5.1 术语采集与预处理

从用户输入文本、电子健康档案、健康科普内容等数据源中采集健康术语，进行清洗（过滤特殊符号、错别字修正）、分词及词性标注处理。

### 5.2 多模态特征融合

将预处理后的文本术语与关联的生理信号（如心率、步态数据）、行为日志（如运动时长、饮食记录）通过多模态对比学习框架映射到统一语义空间，生成术语语义向量。

### 5.3 歧义检测

通过查询动态术语映射库，判断术语是否存在多义性或表述多样性；对存在歧义的术语，标记其候选标准术语集合。

### 5.4 歧义消解

计算歧义术语语义向量与候选标准术语语义向量的相似度，结合上下文语境（如用户慢病类型、当前任务场景）确定最优标准术语。

### 5.5 标准化映射

将消解歧义后的术语映射到标准术语体系，并记录映射关系及置信度；对未匹配到标准术语的新兴术语，启动增量嵌入流程。

### 5.6 结果输出与反馈

输出标准化术语及映射依据，同时收集用户或专业人员的反馈意见，用于优化语义嵌入模型和动态术语映射库。

## 6 核心要求

### 6.1 术语采集与预处理要求

#### 6.1.1 输入方式支持

系统应支持文本、语音转文字等多种输入方式的术语采集，语音转文字的术语识别准确率应不低于 97%。

#### 6.1.2 错别字修正

预处理阶段需完成常见错别字修正（如“高雪压”修正为“高血压”），修正准确率不低于 95%。

#### 6.1.3 数据清洗

应过滤特殊符号、无意义字符，保留有效术语占比不低于 80%的样本。

### 6.2 多模态语义嵌入要求

#### 6.2.1 数据融合

应融合至少 3 种异构数据（文本术语、生理信号、行为日志）进行语义嵌入，确保术语语义表征的完整性。

#### 6.2.2 语义向量质量

语义向量在标准术语测试集上的余弦相似度均值应不低于 0.85。

#### 6.2.3 模型性能

术语聚类纯度应不低于 0.92，支持增量训练时新术语与既有语义空间的一致性保持。

### 6.3 歧义消解要求

#### 6.3.1 消解准确率

对高频歧义术语（如“运动”“锻炼”“活动”“调理”）的消解准确率应不低于 96%。

#### 6.3.2 响应性能

歧义消解响应时间应不超过 1 秒，支持每秒不少于 100 条术语的并发处理。

### 6.3.3 上下文感知

应结合用户健康状况、历史交互记录、当前任务场景等上下文信息进行歧义消解。

## 6.4 动态术语映射库要求

### 6.4.1 术语覆盖

应包含至少 5000 条核心标准健康术语，覆盖疾病、症状、药物、饮食、运动等主要健康领域。

### 6.4.2 增量更新

支持新兴术语的增量嵌入，从术语提交到完成语义对齐的时间应不超过 24 小时。

### 6.4.3 更新频率

术语库更新频率不低于每季度 1 次，同步权威医学指南及标准的更新内容。

## 7 质量评估

### 7.1 评估指标

术语歧义消解系统的质量评估应包括以下核心指标：

表 1 质量评估指标要求

评估指标	指标要求	计算方法
标准化准确率	≥95%	正确标准化的术语数量/总评估术语数量×100%
歧义消解 F1 值	≥94%	$2 \times (\text{精确率} \times \text{召回率}) / (\text{精确率} + \text{召回率}) \times 100\%$
术语库覆盖率	≥90%	可匹配到标准术语的输入术语数量/总输入术语数量×100%
用户满意度	≥4.3 分（5 分制）	用户满意度评分平均值
响应时间	≤1 秒	单次术语消解平均耗时
并发处理能力	≥100 条/秒	单位时间内处理的术语数量

### 7.2 评估方法

#### 7.2.1 自动化测试

使用包含不少于 1000 条标注歧义术语的测试集，自动计算标准化准确率、歧义消解 F1 值及术语库覆盖率。测试集应覆盖疾病、症状、药物、饮食、运动等 6 大核心场景。

#### 7.2.2 人工评估

邀请 5-10 名医疗健康领域专家对随机抽取的 200 条标准化结果进行审核，评估语

义匹配的准确性。人工评估一致性系数（Kappa 值）应不低于 0.8。

### 7.2.3 用户调研

选取不少于 500 名系统实际用户，收集其对术语标准化结果的满意度评分。调研应覆盖不同年龄、教育背景、健康状况的用户群体。

### 7.3 评估周期

系统上线前应完成全面评估，上线后每半年进行一次复评，重大版本更新后应重新评估。

## 8 安全保障

### 8.1 数据隐私保护

#### 8.1.1 加密存储

用户生理信号、健康档案等敏感数据需按照 GB/T 35273-2022 要求进行 AES-256 加密存储。

#### 8.1.2 安全传输

数据传输过程应采用 SSL/TLS 1.3 协议，确保传输过程的完整性和保密性。

#### 8.1.3 访问控制

实施基于角色的访问控制（RBAC），敏感数据访问需经过身份认证和授权。

### 8.2 术语溯源管理

#### 8.2.1 日志记录

记录每条术语的标准化过程日志，包括输入术语、候选标准术语、相似度得分、映射结果及操作人员。

#### 8.2.2 日志保存

日志保存期限不少于 1 年，关键操作日志保存期限不少于 3 年。

#### 8.2.3 审计追踪

支持对术语处理过程的全程审计追踪，确保操作可追溯。

### 8.3 权限控制

#### 8.3.1 分级管理

对动态术语映射库的修改权限进行分级管理，仅授权人员可执行新增、修改、删除等操作。

#### 8.3.2 操作痕迹

保留所有数据修改操作痕迹，包括操作时间、操作人员、修改内容等。

#### 8.3.3 安全审计

定期进行安全审计（每季度至少一次），检查权限配置合理性和操作合规性。

### 8.4 数据备份与恢复

#### 8.4.1 备份策略

术语库数据应每日进行增量备份，每周进行全量备份，备份数据异地存储。

#### 8.4.2 恢复演练

每半年进行一次数据恢复演练，确保在数据丢失或损坏时能够及时恢复。

## 附录 A

(规范性附录)

## 常见健康歧义术语及标准映射表

本附录列出医疗健康领域 100 条高频歧义术语及其对应的标准术语，覆盖疾病、症状、诊疗、饮食、运动、药物等 6 大核心场景。

表 A.1 常见健康歧义术语及标准映射表

歧义术语	上下文场景	标准术语	权威依据
运动	减肥塑形	高强度间歇训练	《中国成人超重和肥胖症预防控制指南》
运动	术后康复	被动关节活动训练	《临床康复医学》
运动	慢病管理	中等强度有氧运动	《中国慢性病防治工作规划》
调理	中医体质改善	中医辨证施调	《中医体质分类与判定》
调理	慢病指标控制	慢病综合干预	《中国慢性病防治工作规划》
调理	亚健康改善	生活方式干预	《健康管理师国家职业技能标准》
感冒	儿童发热咳嗽	急性上呼吸道感染（儿童）	ICD-11 J06.9
感冒	成人鼻塞流涕	急性鼻咽炎（成人）	ICD-11 J00
感冒	流感季节	流行性感冒	ICD-11 J10-J11
降糖	药物治疗	抗糖尿病药物治疗	《中国 2 型糖尿病防治指南》
降糖	饮食干预	糖尿病饮食血糖管理	WS/T 559-2017 糖尿病

			患者膳食指导
降糖	运动干预	糖尿病运动治疗	《中国 2 型糖尿病防治指南》
消炎	细菌感染	抗菌药物抗炎治疗	《抗菌药物临床应用指导原则》
消炎	无菌性炎症	非甾体抗炎药治疗	《中国骨关节炎诊疗指南》
消炎	中药调理	清热解毒治疗	《中药学》
血压高	确诊疾病	高血压病	ICD-11 I10-I15
血压高	暂时升高	血压一过性升高	《中国高血压防治指南》
血压高	测量误差	测量技术因素	医疗器械使用规范
头晕	脑血管疾病	脑供血不足	ICD-11 G45
头晕	耳部疾病	眩晕综合征	ICD-11 H81
头晕	心理因素	焦虑相关头晕	ICD-11 F41
锻炼	力量训练	抗阻运动	《运动处方中国专家共识》
锻炼	心肺功能	有氧运动	《运动处方中国专家共识》
锻炼	柔韧性	柔韧性训练	《运动处方中国专家共识》
活动	日常活动	身体活动	WHO 身体活动指南
活动	康复训练	功能性活动训练	《康复医学》

活动	社交参与	社会参与活动	老年健康管理指南
饮食	营养摄入	膳食营养管理	《中国居民膳食指南》
饮食	疾病治疗	医学营养治疗	《中国糖尿病医学营养治疗指南》
饮食	体重控制	能量限制饮食	《中国超重/肥胖医学营养治疗指南》

歧义术语	上下文场景	标准术语
运动	减肥塑形	高强度间歇训练
运动	术后康复	被动关节活动训练
调理	中医体质改善	中医辨证施调
调理	慢病指标控制	慢病综合干预

## 附录 B

## (资料性附录)

## 多模态语义嵌入模型训练指南

本附录详细规定多模态语义嵌入模型的训练流程、数据准备、架构设计、超参数优化及评估验证方法，为模型开发提供可操作指导。

## B.1 训练数据准备

## B.1.1 数据来源与类型

文本数据：包含权威医学文献（如 PubMed 摘要）、诊疗指南（如《中国高血压防治指南》）、用户健康咨询文本（脱敏处理），总量不少于 50 万条，单条文本长度 10-50 字符。

生理信号数据：包含心率、血压、血糖、步态加速度等时序数据，采样频率 $\geq 1\text{Hz}$ ，每条信号数据关联对应的健康术语标签（如“高血压”关联血压信号），总量不少于 10 万条。

行为日志数据：包含用户运动时长、饮食种类、睡眠时长等结构化数据，每条日志需标注核心健康术语（如“有氧运动”关联跑步运动日志），总量不少于 20 万条。

## B.1.2 数据预处理

a) 文本数据预处理：

- 1) 采用 jieba 分词工具进行分词，使用 BERT-Chinese 预训练模型进行词嵌入；
- 2) 过滤包含特殊符号、无意义字符的文本，保留有效术语占比 $\geq 80\%$ 的样本；
- 3) 对文本进行去重、去噪处理，确保数据质量。

b) 生理信号数据预处理：

- 1) 采用 Z-score 标准化处理，去除异常值（超出均值 $\pm 3$ 倍标准差的数据）；
- 2) 通过滑动窗口（窗口大小 5 秒，步长 2 秒）提取特征；
- 3) 对缺失值采用插值法或删除处理。

c) 行为日志数据预处理：

- 1) 将分类变量（如饮食种类）进行独热编码；
- 2) 连续变量（如运动时长）进行 min-max 归一化，形成固定维度的特征向量。

## B.2 模型架构设计

推荐采用 FGCrossNet 低秩张量融合架构，整体分为特征提取层、融合层、语义映射层三部分。

## B.2.1 特征提取层

文本特征：采用 BERT-base 模型（层数 12，隐藏层维度 768）提取，输出维度 256。

生理信号特征：采用 1D-CNN（卷积核大小 3，数量 64）提取，输出维度 256。

行为日志特征：采用全连接层（神经元数量 256）提取，输出维度 256。

## B.2.2 融合层

通过低秩张量融合将三模态特征融合为统一特征向量，融合公式为：

$$F = W \times T_1 \times T_2 \times T_3 + b$$

式中：

W——低秩张量权重（秩为 32）；

b——偏置项；

$T_1$ 、 $T_2$ 、 $T_3$ ——各模态特征张量。

### B. 2.3 语义映射层

采用 2 层全连接网络：

第一层神经元数量 128，激活函数 ReLU；

第二层神经元数量 64，激活函数 ReLU；

最终输出 64 维的术语语义向量。

### B. 3 超参数设置与训练策略

表 B.1 超参数设置

超参数名称	取值	说明
学习率	5e-5	采用余弦退火学习率调度，每 10 轮衰减 10%
批量大小（Batch Size）	64	GPU 显存 $\geq$ 16GB 时可设置为 128
迭代次数（Epoch）	100	验证集损失连续 10 轮无下降则早停
正则化系数	1e-4	采用 L2 正则化防止过拟合
dropout 概率	0.2	语义映射层设置 dropout 层
优化器	AdamW	权重衰减系数 0.01
损失函数	对比损失 (Contrastive Loss)	温度系数 0.07

### B. 4 模型评估与优化

#### B. 4.1 评估指标

语义向量余弦相似度：均值 $\geq$ 0.85；

术语聚类纯度： $\geq$ 0.92；

歧义消解准确率： $\geq$ 96%；

模型推理时间：单条术语 $\leq$ 50ms。

#### B. 4.2 优化策略

对低相似度样本（余弦相似度 $<$ 0.7）进行数据增强（如同义词替换、上下文扩展）；  
采用知识蒸馏技术，将大模型（BERT-large）的语义知识蒸馏到基础模型，提升推理效率；

使用混合精度训练（FP16）加速训练过程，减少显存占用。

#### B. 4.3 增量训练

新增术语数据时，采用冻结特征提取层、微调融合层和语义映射层的方式进行增量训练：

训练轮次不少于 20 轮；  
学习率设置为初始学习率的 1/10；  
确保新术语与既有语义空间的一致性，一致性检验余弦相似度 $\geq 0.8$ 。

## 参考文献

- [1] GB/T 22239-2019 信息安全技术 网络安全等级保护基本要求
  - [2] GB/T 35273-2022 信息安全技术 个人信息安全规范
  - [3] GB/T 38667-2020 人工智能 术语
  - [4] GB/T 44588-2024 数据安全技术 互联网平台及产品服务个人信息处理规则
  - [5] WS/T 447-2014 电子健康档案基本架构与数据标准
  - [6] WS 363-2011 卫生信息数据元目录
  - [7] WS/T 559-2017 糖尿病患者膳食指导
  - [8] 《国际疾病分类第十一次修订本（ICD-11）》（世界卫生组织）
  - [9] 《中国高血压防治指南》（国家卫生健康委）
  - [10] 《中国2型糖尿病防治指南》（中华医学会糖尿病学分会）
  - [11] 《中国成人超重和肥胖症预防控制指南》（卫生部疾病控制司）
  - [12] 《抗菌药物临床应用指导原则》（国家卫生健康委）
  - [13] 《中国骨关节炎诊疗指南》（中华医学会骨科学分会）
  - [14] 《中医体质分类与判定》（中华中医药学会）
  - [15] 《运动处方中国专家共识》（中华医学会运动医疗分会）
-