

用户输入问题

"我最近工作压力很大，经常失眠..."
对话管理层 (Dialogue Manager)

统一协调 · 上下文构建 · 流程控制

双知识库RAG模块

(创新点 ①)

专业心理知识库

- SmileChat 16K
- CBT/MBSR文献
- 治疗指南
- 50K知识片段

↑ 向量检索

权重: 60%

个人信息知识库

- 用户基本档案
- 历史问题记录
- 干预历史追踪
- 动态更新

↑ 向量检索

[时间衰减] [重排序]
权重: 40%

并行检索

检索结果融合

三层记忆模块

(创新点 ②)

长期记忆 (L3)

- 用户档案
- 情绪趋势
- 历史摘要

短期记忆 (L2)

- 会话级摘要
- 关键信息提取
- 近期20会话

工作记忆 (L1)

- 当前会话上下文
- 最近10轮对话
- 实时情绪状态

上下文构建模块

系统提示词 +
专业知识检索结果 (3-5文档) +
个人信息检索结果 (2-3文档) +
长期记忆 (档案+趋势) +
短期记忆 (近期摘要) +
工作记忆 (当前对话10轮)

增强型上下文

大语言模型层 (LLM)

□ Qwen2-7B-Instruct 或 OpenAI GPT-4

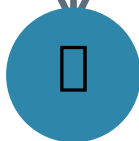
基于增强上下文生成专业且个性化回复

记忆更新模块

- 添加新轮次到工作记忆
- 会话结束 → 生成摘要 → 短期
- 更新用户档案 → 长期记忆

【系统回复】

"理解你的压力。根据你上次提到的
项目deadline问题，建议尝试：
1. 艾森豪威尔矩阵优先级管理
2. 4-7-8呼吸法改善失眠
3. 固定'担忧时间'技术..."



用户

关键创新

① 双知识库
RAG架构

② 三层记忆
系统

③ 端到端
实现

④ 多维度
评估