Product Requirements Document · V1.1

医药多智能体AI助手

BTDT LightRAG · 产品需求文档

把可信度做成产品体验的医药垂直问答平台

三智能体协同 + RAG 知识库 + 区块链溯源 + 内地海外双标并列

项目档案

产品名称	医药多智能体AI助手（BTDT LightRAG）
文档版本	V1.1（基于用户调研迭代）
项目周期	2025.11 — 至今
项目角色	产品负责人（端到端独立产出）
文档负责人	徐玉欣
最后更新	2026.05

横评成绩
24 人盲评综合得分 9.2 / 10 · 双域命中率 86% · 领先 GPT-3.5、通义千问、药智数据 APP

徐玉欣　　意向岗位：大模型产品经理（实习）
联系方式：1191091728@qq.com

项目	信息
产品名称	医药多智能体AI助手（BTDT LightRAG）
文档版本	V1.1
文档负责人	徐玉欣
项目角色	产品负责人（端到端）
项目周期	2025.11 — 至今
最后更新	2026.05
文档状态	进行中（V1.1 基于用户调研产出，正在推进 P0 迭代）

版本变更记录

版本	日期	变更内容	负责人
V1.0	2025.11	完成基础需求分析、三智能体架构方案、原型设计与首版落地	徐玉欣
V1.1	2026.04	引入用户调研与竞品横评结果，重新校准需求优先级，输出下一阶段 P0/P1 迭代计划	徐玉欣

本 PRD 既是产品落地文档，也是个人作品集投递版本。V1.0 部分对应已交付内容，V1.1 部分体现基于真实用户反馈的迭代思考。

一、需求分析

1.1 项目背景与机会

AI 医药问答正处于一个尴尬阶段：模型能力快速进步，但「能不能用」与「敢不敢用」之间的差距越来越大。

核心问题：

- 通用大模型在医药垂直场景下幻觉严重，会自信编造药品名称、用法用量、适应症
- 错误信息以严谨的语气输出，用户极难辨别真伪
- 医药信息是高敏感、高代价的内容——一次错误回答可能直接影响用药安全
- 当前主流 AI 医药产品普遍缺乏来源透明化与内容核验机制，用户即便看到答案，也不敢真正信任
- 内地/海外用药标准存在差异（剂量单位、规格、疗程长度等），现有通用模型几乎全部按单一域回答，缺乏跨域核对意识

机会洞察：

在「可信」这个维度上，行业整体的供给是不足的。谁能率先把「答案可被验证 + 跨域可核对」做成产品级体验，谁就能在医药垂直 AI 上建立壁垒。

1.2 目标用户与使用场景

用户群体	占比预估	核心诉求	典型场景
高校学生（医学/药学）	35%	精准的药物知识与考点查证	课程学习、考试备考、文献阅读辅助
普通用药人群	50%	用药安全、副作用、配伍禁忌	家庭用药咨询、症状自检后的用药选择
医药从业者	15%	快速可靠的专业核验工具	药剂师查证、基层医生辅助、医药内容创作

1.3 用户调研（V1.1 新增）

调研方式： 竞品横评问卷 + 半结构化用户访谈 + 医药专业同学盲评 + 自体验拷打

调研样本： 有效问卷 24 份，深度访谈 8 人

样本构成：

- 医疗行业从业者 9 人（含临床药师 3、基层医生 4、医药内容审核 2）
- 高校学生 11 人（医学/药学相关 6，非医学背景 5）
- 科研人员 4 人

调研时间： 2025.05 — 2025.07

调研方法： 设计标准化 7
题医药问答评测集（覆盖剂量、规格、疗程、儿童用药、最大日剂量等高频场景），对本产品（BTDT LightRAG）与 GPT-3.5、通义千问、药智数据 APP 进行同题盲评，按 10/5/0 三档评分（10 分=内地+海外双域信息完整，5 分=单域完整，0 分=信息缺失）。

调研问卷： https://v.wjx.cn/vm/e3XDs0D.aspx

关键发现：

1. 「准确」的优先级压倒一切

用户对答案准确性的在意程度，远大于对响应速度的在意程度。受访者多次提到「宁可慢一点，也要对」。

1. 内地/海外双域信息差是被忽视的真痛点

横评结果显示，通用大模型（GPT-3.5、通义千问）几乎只输出单域信息（多为内地或海外其一）；本产品 BTDT LightRAG 在 7
道题中均明确区分了内地与海外标准。受访的医药从业者反馈：「这种区分才是临床真实需要的」。

1. 用户存在「下意识复核」行为

约 80% 的受访者在拿到 AI 回答后，会下意识「再问一句」或者去搜索引擎/官方说明书二次确认。这说明现有产品没有真正解决信任问题。

1. 不同用户群体的关注点差异显著

- 医药从业者最关心「来源是否权威」「是否区分内地/海外标准」
- 高校学生最关心「能不能引用、能不能写进作业/论文」
- 普通用户最关心「我能不能这样用、有没有风险」

1. 溯源的「展示」比「可点」更重要

用户大多不会主动点开来源链接，但「有溯源」和「没溯源」对信任感的影响完全不同——溯源不只是功能，更是一种「专业感」的可见性。

1. 多轮对话需求被反复提及

单轮 QA 无法满足真实咨询场景。用户希望能「这个药能和那个药一起吃吗」这样追问下去。

洞察转化为需求：

调研发现	产品改动方向	优先级调整
准确性压倒一切	强化核验智能体的拦截阈值，宁拒答不乱答	P0 维持
内地/海外双域差异是核心差异化点	默认双域并列展示，并在 UI 上标识来源地域	新增 P0
用户会下意识复核	把溯源信息前置展示，而非折叠在「查看来源」按钮里	P2 → P0
用户群体诉求差异	回答结构按场景差异化展示（学生看引用 / 医生看权威 / 普通用户看「能不能用」）	P1
多轮对话需求强	将多轮对话从 P2 上调至 V1.1 P0	P2 → P0

1.4 竞品分析

1.4.1 定性对比

竞品	优势	短板	我们的差异化
通用大模型（ChatGPT、通义、文心等）	通用对话能力强、生态完整	医药垂直能力弱、无溯源、敏感问题无兜底、几乎不区分内地/海外标准	医药垂直 RAG + 三智能体核验 + 区块链溯源 + 双域信息并列
丁香医生 AI 助手	品牌信任度高、内容权威	偏 C 端医疗咨询，缺少结构化核验机制	三智能体分工，回答可被独立验证
药智数据 APP 等垂直工具	数据库准确、垂直度高	交互形态老旧、缺少 AI 自然语言能力	AI 对话 + 知识库的融合体验
各类医药知识库网站	信息权威	检索体验差、需要专业知识才能用	自然语言一问即答 + 自动定位证据片段

1.4.2 横评定量结果（V1.1 新增）

评测集：7 道高频医药问答题（剂量/规格/疗程/儿童用药/最大日剂量/起始剂量）
评分标准：10 分=内地+海外双域信息完整；5 分=单域完整；0 分=信息缺失或错误
评测人：24 名医疗从业者、医药专业学生、科研人员盲评，取平均分

产品	平均得分（满分 10）	双域命中率	主要短板
BTDT LightRAG（本产品）	9.2	86%	部分罕用药覆盖不足
药智数据 APP	6.4	14%	仅内地数据、无 AI 对话能力
通义千问	5.1	0%	单域输出、无溯源

产品	平均得分（满分 10）	双域命中率	主要短板
GPT-3.5	4.7	0%	多为海外标准、与内地实际处方差异大

差异化定位：

> 我们不是「另一个医药 AI 问答」，而是「第一个把可信度与跨域核对做成产品体验的医药 AI 问答」。

二、产品规划

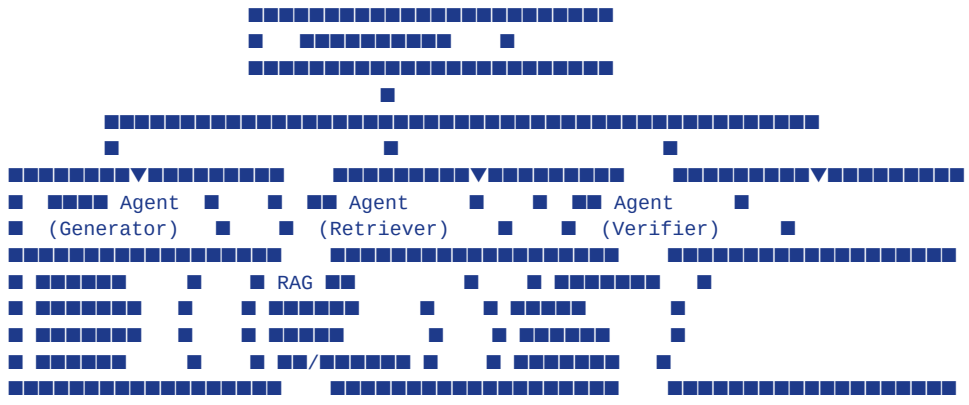
2.1 产品定位

面向学生、普通用户、医药从业者的垂直可信医药问答与信息核验平台。

2.2 价值主张

关键词	承诺
可信	每条回答都看得见来源、经得起核验
严谨	经过领域调参，控制术语严谨度与回答边界
全面	内地 + 海外双域用药标准并列，临床决策更安全
安全	敏感问题自动拦截，永远不在底线上冒险
稳定	阿里云 API 部署，保证响应速度与并发可用

2.3 整体架构 — 三智能体协同

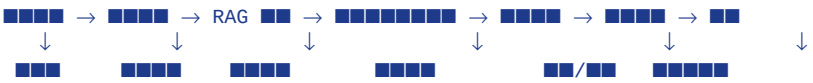


底层基础设施（自下而上）：

1. 医药知识库（三域：内地 CCKS / 海外 PubMed / 自建联合数据集）
2. RAG 检索层（向量化 + 关键词混合检索，按地域分域）
3. 多智能体协同层（任务路由、Agent 调度）
4. 前端交互层（Web / 移动端）

分工原则： 一个负责讲清楚，一个负责找证据，一个负责把关。

2.4 核心业务流程



关键节点说明：

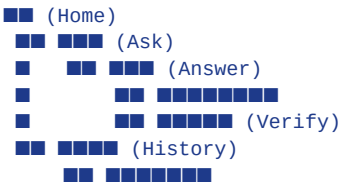
节点	输入	处理	输出
意图识别	用户原始问题	判断类型与敏感等级	路由策略 + 敏感标记
RAG 检索	标准化问题	知识库向量检索 + 元数据 + 分域召回	候选证据片段（Top-K）
协同生成	证据片段 + 用户问题	基于证据生成回答	结构化回答草稿（双域并列）
内容核验	回答草稿 + 证据	一致性 + 安全性 + 双域校验	通过/拦截/改写
溯源展示	通过的回答	附加来源 + 上链信息	最终输出

2.5 功能模块清单

模块	子功能	优先级	版本
医药垂直问答	单轮问答、术语理解、用药咨询	P0	V1.0
RAG 知识库	检索召回、证据返回、知识库扩充	P0	V1.0
双域信息并列	内地/海外标准分别召回与展示	P0	V1.0
溯源核验	来源标注、可信链路、区块链存证	P0	V1.0
异常处理	敏感拦截、无结果兜底、网络异常重试	P0	V1.0
多轮对话	上下文理解、连续追问、对话历史	P0（V1.1 上调）	V1.1
溯源前置展示	关键证据片段直接展示在回答中	P0（V1.1 新增）	V1.1
差异化回答结构	按用户场景调整回答深度与结构	P1	V1.1
个性化健康建议	结合用户画像与历史问答	P1	V2.0
专业工具集成	药物相互作用、症状自检	P2	V2.0

三、原型设计

3.1 核心页面与信息架构



3.2 关键页面说明

> 原型仅用于呈现核心信息结构，详细视觉稿见 Figma 链接（面试时可现场展示）。

3.2.1 首页 Home

- 快捷提问入口（输入框 + 发送按钮）
- 热门场景引导（药物查询 / 用药咨询 / 配伍禁忌等）
- 常见问题列表（按类别归集）
- 入口标识：明显的「可信医药 AI · 内地海外双标」品牌承诺

3.2.2 提问页 Ask

- 输入框（支持多行）
- 快捷指令（@药品名 / #症状 / !紧急咨询）
- 敏感词前置提示（如涉及紧急症状直接引导就医）
- 历史对话入口

3.2.3 回答页 Answer

- 结构化回答展示（要点 + 详情分层）
- 双域并列卡片：内地标准 / 海外标准 左右分栏展示
- 溯源前置：关键证据片段直接显示在回答下方（V1.1 新增）
- 来源外链一键跳转
- 「重新生成」与「不满意反馈」按钮
- 进入核验详情页入口

3.2.4 核验详情页 Verify

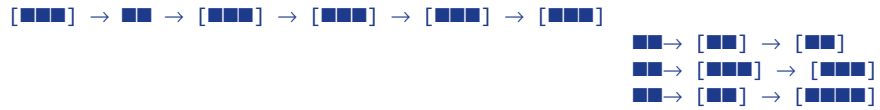
- 可信标识（已上链 / 已核验徽章）
- 完整来源列表（文档名、章节、地域标签、上链时间戳）
- 区块链存证查询入口（可选展开）

3.3 关键字段定义

字段	类型	必填	说明
question	string	是	用户提问，长度上限 500 字
intent_type	enum	是	知识查询 / 用药咨询 / 敏感问题 / 闲聊
sensitivity_level	enum	是	low / medium / high（high 触发拦截）
region_scope	enum	是	mainland / overseas / both（默认 both）

字段	类型	必填	说明
evidence_list	array	是	RAG 召回的证据片段（含来源元数据 + 地域标签）
answer_status	enum	是	success / blocked / no_result / error
trace_id	string	是	区块链存证 ID
sources	array	是	引用来源列表（文档名 + 章节 + 地域 + URL）

3.4 状态流转



3.5 异常场景处理

场景	处理逻辑	用户感知
无检索结果	友好提示 + 给出问题改写建议	"未找到可信信息，您可以试试这样问……"
敏感问题（如自伤、未成年用药）	拒绝回答 + 触发安全提示 + 引导就医	"这个问题建议咨询医生，您可以拨打……"
网络异常	本地缓存最近回答 + 自动重试 3 次	自动重试，失败时提示「网络异常，请稍后重试」
核验未通过	触发改写或直接拦截	用户不感知失败，仅感知到回答更谨慎
知识库覆盖不足	明确告知「当前知识库不覆盖该领域」	不编造，承认能力边界
单域信息缺失	标注「海外标准暂未收录」或反之	不强行编造另一域信息

四、落地思考

4.1 技术方案理解（PM 视角）

技术方案	解决什么问题	PM 视角的关键判断
RAG 外挂知识库	用证据约束模型自由发挥，降低幻觉	知识库质量决定产品上限；数据治理比技术选型更重要
三域知识库架构	内地/海外/自建分域，支撑双域并列输出	是本产品最核心差异化，必须前置投入
模型调参 + API 部署	让模型在医药场景稳定可用	关注响应速度、并发稳定性；调参 ROI 高于换模型

技术方案	解决什么问题	PM 视角的关键判断
区块链溯源	把可信度做成可被验证的事实	用户感知到的是「专业感」而非区块链本身；作为差异化背书
多智能体协同	分工解耦，让每个 Agent 专注一件事	比单一大模型更可控、可调试、可解释

4.2 知识库构建方法论

知识库不是「把数据塞进去」，而是一套从原始数据到可用知识的工程化治理流程：

■■■■■■■■■ → ■■■■■■■■ → ■■■■■ → ■■■■/■■■■■ → ■■■■■

三域架构：

域	数据来源	覆盖范围	数据处理要点
a 域（内地）	CCKS 医疗数据集	中国药典、药品说明书、临床指南	单位统一为内地规范（mg→g 换算）、适应症中文标准术语对齐
b 域（海外）	PubMed	FDA 橙皮书、临床文献摘要、海外药品说明书	中英双语对齐、剂量单位 mg/g 等效标注、海外标准中文化展示
c 域（自建）	AI 自整理 + 联合数据集	a/b 域交叉校验结果、专家标注 QA 对、罕用药补充	人工校验+AI 辅助标注、定期治理更新

关键处理步骤：

1. 元数据标签：为每条数据打上地域标签（内地/海外）、药品分类、来源权威等级、最后更新时间
2. 筛选清洗：去掉重复条目、格式不统一、信息残缺的记录
3. 无效数据删除：过时药品、已退市品种、与当前药典冲突的旧版信息
4. 交叉校验：a/b 域同药同题对比，不一致处人工核验后录入 c 域

4.3 风险与应对

风险类别	具体风险	应对方案
内容风险	知识库存在过时信息	建立定期更新机制 + 标注「最后更新时间」
合规风险	涉及处方药咨询的法律边界	在敏感场景强制引导就医，不替代医生
技术风险	RAG 召回不准导致回答偏差	多路召回 + 核验智能体二次校验
体验风险	过于严谨导致拒答率高、用户流失	设置可调拦截阈值，平衡严谨与可用
信任风险	用户对区块链概念无感	用「已核验」「已上链」徽章简化呈现，不强调技术
地域风险	用户误用海外标准在内地实际用药	双域并列展示时明确加注「请遵医嘱，以处方为准」

4.4 资源协同

角色	协同内容
算法/研发	RAG 链路实现、Agent 调度、模型调参、API 部署
知识库运营	三域数据治理、元数据标注规范、清洗规则维护、更新机制
内容审核	敏感词库维护、安全拦截策略
法务合规	医药内容合规审查、用户协议、免责声明
设计	交互稿、视觉规范、品牌一致性

五、迭代规划

5.1 V1.0 已交付（基础闭环）

- 医药垂直问答（基础场景）
- RAG 知识库（三域数据接入 + 元数据标签 + 筛选清洗）
- 三智能体协同架构落地
- 双域信息并列输出
- 溯源核验机制（区块链存证）
- 异常处理（敏感拦截 / 无结果 / 网络异常）
- 阿里云 API 部署上线

V1.0 成果： 完成可演示、可部署、可迭代的产品闭环；在 24 人盲评中横向对比 GPT-3.5、通义千问、药智数据 APP，综合得分 9.2 / 10，双域命中率 86%，显著领先竞品。

5.2 V1.1 进行中（基于用户调研的优先级校准）

P0 — 必做：

需求	调研依据	预期价值
多轮对话支持	受访者高频提及单轮 QA 不够用	复用率提升 + 真实场景覆盖
溯源信息前置展示	用户不会主动点开溯源，但「可见」对信任感影响显著	信任感可见化、降低用户复核成本
拦截阈值调优	用户在意准确性远超速度	减少错误回答、强化「敢用」感受
双域展示 UI 优化	横评结果证明双域是核心差异化，需在前端更显性	强化产品独特性、提升专业人群满意度

P1 — 重要：

需求	调研依据	预期价值
差异化回答结构	不同用户群体关注点显著差异	提升各群体满意度与匹配度
知识库扩充（OTC、中成药、儿科、罕用药）	横评中暴露的覆盖盲区	提升问题命中率

需求	调研依据	预期价值
历史对话管理	用户希望「再看一眼之前问过的」	提升复用与留存

P2 — 可选：

- 用户反馈闭环（顶踩、评论、纠错通道）
- 回答分享卡片（医药从业者诉求）
- 处方/说明书 OCR 识别后直接问答

5.3 V2.0 长期规划

- 个性化健康建议（结合画像与历史问答）
- 专业工具集成（药物相互作用查询、症状自检引导）
- 多模态输入（拍照识别药品说明书）
- 个人健康档案打通（与可穿戴设备数据联动）

六、成功指标

指标	含义	V1.0 现状	V1.1 目标	衡量方式
综合问答得分（10 分制）	24 人盲评横评分数	9.2	≥ 9.5	标准化 7 题评测集 + 盲评
双域命中率	同时给出内地+海外信息的回答占比	86%	≥ 92%	系统埋点 + 抽样核验
溯源覆盖率	可一键查证的回答占比	100%	100%（前置展示）	系统埋点
拦截准确率	敏感问题正确拦截比例	92%	≥ 95%	人工评估
平均响应时长	用户感知的回答速度	4.2s	< 3s	系统埋点（P95）
7日复用率	用户 7 日内再次使用比例	38%	提升 30%	用户行为分析
多轮对话占比	包含 2 轮以上对话的会话比例	0%	≥ 40%	系统埋点

注：V1.0 现状数据来源于内测期（2026.03-04）抽样统计与 24 人盲评，样本量较小，仅作迭代基线参考。

七、附录

7.1 名词解释

名词	解释
RAG	Retrieval-Augmented Generation，检索增强生成

名词	解释
Agent	智能体，具备特定任务能力的 AI 单元
区块链溯源	将关键信息上链存证，实现来源不可篡改
幻觉	大模型自信编造不存在的事实
召回	从知识库中检索出与问题相关的内容
双域	内地（中国）与海外（主要为 FDA / EMA）的用药标准
CCKS	China Conference on Knowledge Graph and Semantic Computing，中国知识图谱与语义计算大会，其医疗数据集涵盖中文医学实体识别、关系抽取等任务数据
PubMed	美国国家医学图书馆（NLM）维护的生物医学文献数据库

7.2 知识库数据源详述

a 域 — 内地（CCKS 医疗数据集）：

- CCKS 历年医疗 NLP 评测数据（实体识别、关系抽取、事件抽取）
- 中国药典 2025 版（结构化药品说明书）
- NMPA 国家药监局公开数据
- 国家卫健委临床用药指南

b 域 — 海外（PubMed）：

- PubMed 近 5 年中英文摘要（API 接入）
- FDA Orange Book（橙皮书，经清洗的结构化版本）
- EMA 欧洲药品管理局公开数据
- 默沙东诊疗手册（专业版，公开版权内容）

c 域 — 自建联合数据集：

- AI 辅助自整理：基于 a/b 域交叉对比生成候选 QA 对，人工校验后入库
- 医药顾问标注：约 1500 条专业校验 QA 对
- 罕用药与特殊人群用药补充数据
- 月度增量更新 + 季度全量治理

数据治理流程：

1. 原始公开医疗数据集获取
2. 附加元数据标签（地域 / 药品分类 / 权威等级 / 更新时间）
3. 筛选清洗（去重、格式统一、残缺补全）
4. 删除无效数据（过时药品、已退市品种、与当前药典冲突记录）
5. 分域入库 + 向量化索引构建
6. a/b 域同题交叉校验，不一致处人工核验录入 c 域

7.3 调研材料

- 竞品横评问卷： https://v.wjx.cn/vm/e3XDs0D.aspx
- 访谈纪要（脱敏版）： 可面试时现场展示

- 原始评分数据 Excel： 可面试时现场展示

文档结束 · 持续更新中

> 本 PRD 由徐玉欣独立产出，承担需求调研、产品方案、原型设计、迭代规划全流程。
联系方式：1191091728@qq.com