企业 AI 实操 · 开源路线

完整演示 / 36页

企业AI转型 知识库怎么搭？

LLM Wiki × 社区V2 × OKF

原文解读 / 架构拆解 / 开源实测

LLM Wiki



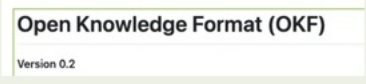
原版思路

社区 V2



工程扩展

OKF



知识格式

五步搭建，不只问答

agentmemory　RAGFlow

OpenWiki　qmd　KAG

Dify　Obsidian



我的个人Agent

扫码访问

企业AI / 知识库

A PRACTICAL GUIDE

企业知识库到底怎么搭?

不只查资料。
让 AI 在任务里，用对知识。

LLM Wiki	Karpathy
社区扩展 V2	Rohit Ghumare
OKF	知识表达格式

五步构建 × 开源选择 × 有边界实测

GUIDE 向量检索与 Wiki 编纂可以组合，不是非此即彼。

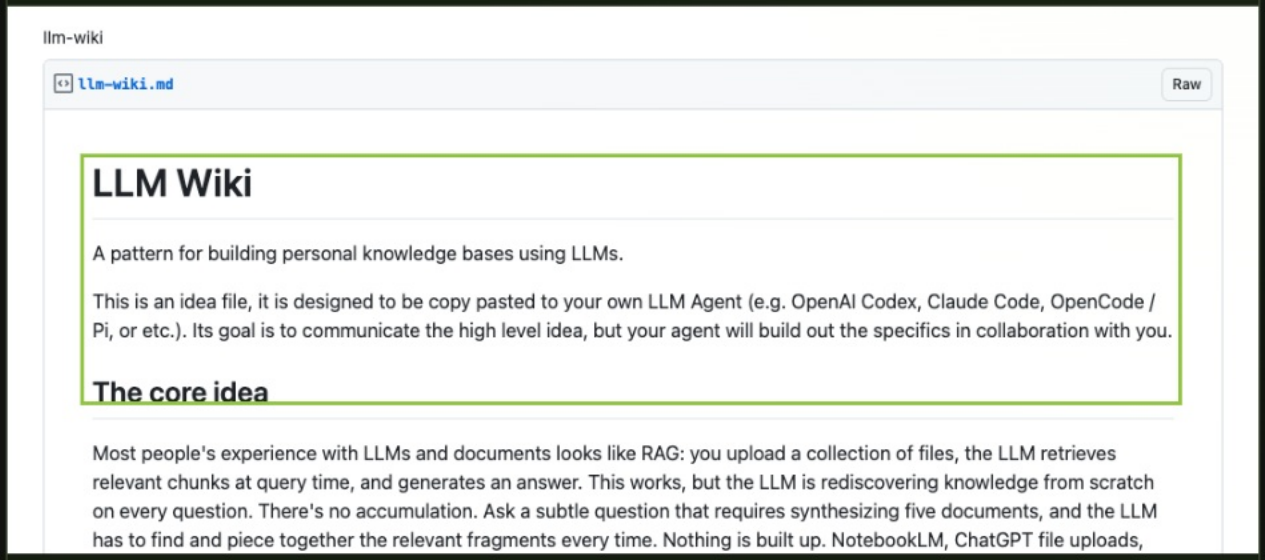
笔记 / 这一步，你还缺哪项证据?

A PRACTICAL GUIDE

企业知识库到底怎么搭?

不只查资料。
让 AI 在任务里，用对知识。

Karpathy · LLM Wiki原文



原版是构建模式，已包含持续维护，不是现成产品。

官方原文实截 · 2026-09-15 · 点图或Z切换局部与全图

ORIGINAL PAGE 原作者公开页面真实截图；原文观点与作者自述不等于独立效果验证。

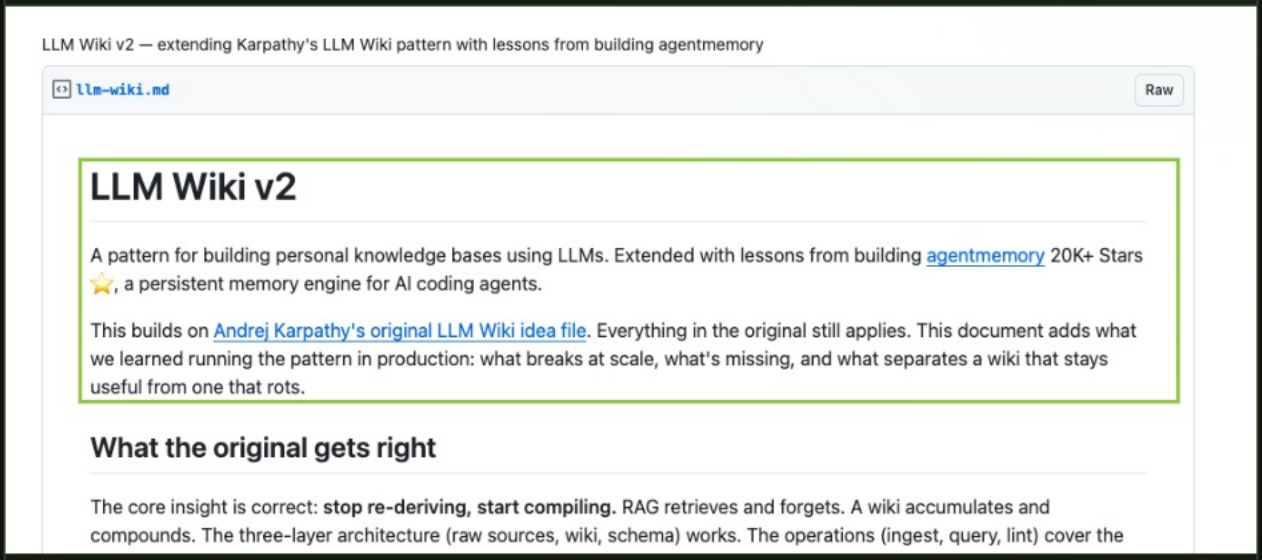
笔记 / 这一步，你还缺哪项证据?

A PRACTICAL GUIDE

企业知识库到底怎么搭?

不只查资料。
让 AI 在任务里，用对知识。

Rohit · 社区V2原文


LLM Wiki v2 — extending Karpathy's LLM Wiki pattern with lessons from building agentmemory

llm-wiki.md Raw

LLM Wiki v2

A pattern for building personal knowledge bases using LLMs. Extended with lessons from building agentmemory 20K+ Stars ⭐, a persistent memory engine for AI coding agents.

This builds on Andrej Karpathy's original LLM Wiki idea file. Everything in the original still applies. This document adds what we learned running the pattern in production: what breaks at scale, what's missing, and what separates a wiki that stays useful from one that rots.

What the original gets right

The core insight is correct: **stop re-deriving, start compiling.** RAG retrieves and forgets. A wiki accumulates and compounds. The three-layer architecture (raw sources, wiki, schema) works. The operations (ingest, query, lint) cover the

在原版基础上扩展；不是Karpathy官方第二版。

官方原文实截 · 2026-09-15 · 点图或Z切换局部与全图

ORIGINAL PAGE 原作者公开页面真实截图；原文观点与作者自述不等于独立效果验证。

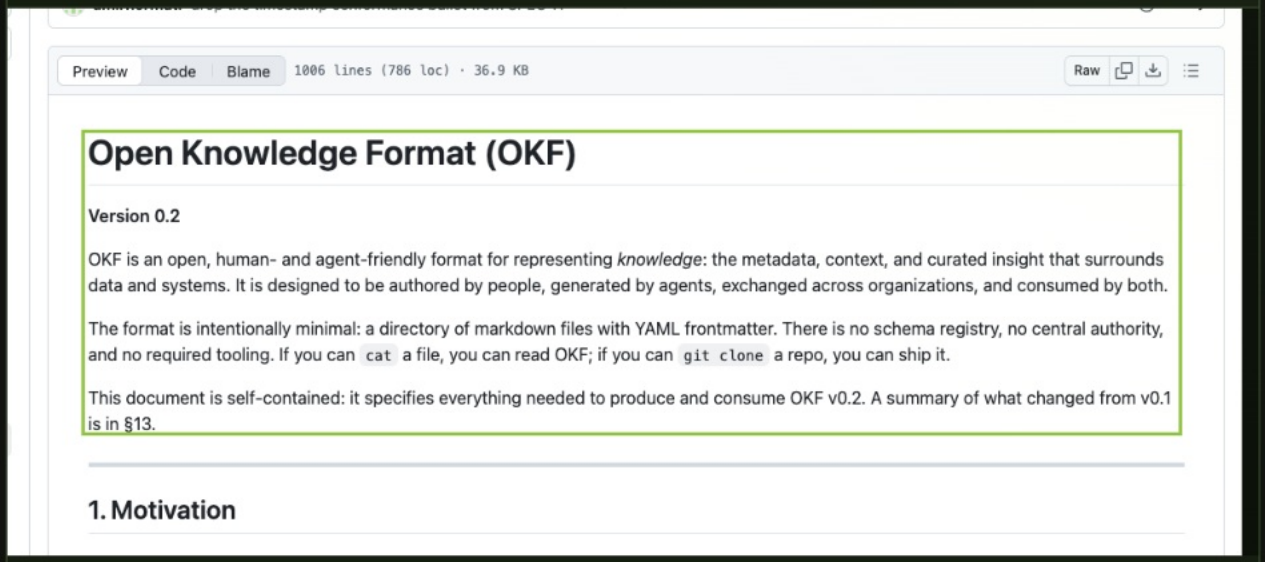
笔记 / 这一步，你还缺哪项证据?

A PRACTICAL GUIDE

企业知识库到底怎么搭?

不只查资料。
让 AI 在任务里，用对知识。

OKF · 官方SPEC v0.2



Markdown与YAML的知识表达格式，不是执行平台。

官方原文实截 · 2026-09-15 · 点图或Z切换局部与全图

ORIGINAL PAGE 原作者公开页面真实截图；原文观点与作者自述不等于独立效果验证。

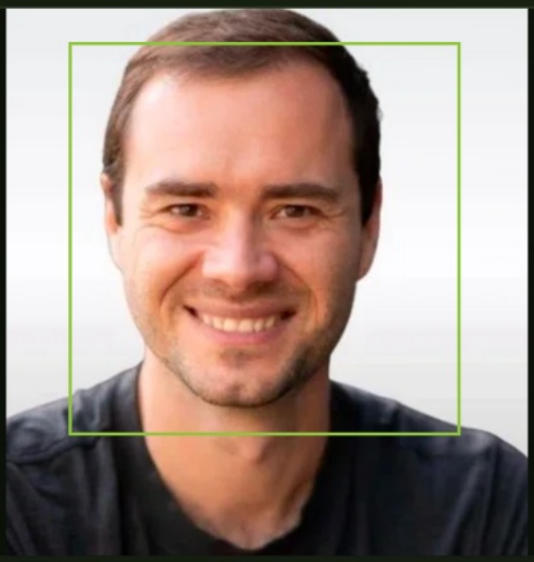
笔记 / 这一步，你还缺哪项证据?

A PRACTICAL GUIDE

企业知识库到底怎么搭?

不只查资料。
让 AI 在任务里，用对知识。

人物配图 · 用户提供



作为人物配图展示，不将照片作为技术效果证据。

用户附件 · 未另行核验照片身份与拍摄来源

USER PROVIDED 用户提供素材；图片原样嵌入，非本轮官方页面抓取或产品实测。

笔记 / 这一步，你还缺哪项证据?

THE BUILD / 00

先搭一条能跑通的链。

不是把所有项目装一遍。
每一步，都要有可检查的产出。

01	接资料	来源、版本、权限
02	编 Wiki	结论、证据、适用条件
03	对齐语义	对象、关系、系统映射
04	接 Agent	查询、建议、受控执行
05	更新验收	修订、发布、回归测试

BUILD 五种职责，不是五个必须新建的系统。

笔记 / 这一步，你还缺哪项证据？

THE BUILD / 00

先搭一条能跑通的链。

不是把所有项目装一遍。
每一步，都要有可检查的产出。

LLM Wiki · 原版结构解读



来源、Agent、Wiki与Schema；原版已有维护回路。

依据原文重绘 · 非官方原图 · 点图或Z查看局部

EXPLAINED / NOT ORIGINAL 依据原文的讲解示意，不是原作者架构原图；不是完整企业部署图。

笔记 / 这一步，你还缺哪项证据？

THE BUILD / 00

先搭一条能跑通的链。

不是把所有项目装一遍。
每一步，都要有可检查的产出。

社区V2 · 记忆巩固子流程



只展示巩固方向，不把它当整套V2架构。

依据原文重绘 · 非官方原图 · 点图或Z查看局部

EXPLAINED / NOT ORIGINAL 依据原文的讲解示意，不是原作者架构原图；不是完整企业部署图。

笔记 / 这一步，你还缺哪项证据？

THE BUILD / 00

先搭一条能跑通的链。

不是把所有项目装一遍。
每一步，都要有可检查的产出。

OKF · 写入与消费关系



生产者写入格式，消费者读取；执行由外部实现。

依据原文重绘 · 非官方原图 · 点图或Z查看局部

EXPLAINED / NOT ORIGINAL 依据原文的讲解示意，不是原作者架构原图；不是完整企业部署图。

笔记 / 这一步，你还缺哪项证据？

THE BUILD / 01

先定任务。再接资料。

从一个有输入、有结果、有负责人的任务开始。

一个任务 只接它需要的资料

SOURCE OF TRUTH

原始资料

权威来源 来源版本

有效时间 更新事件

DERIVED KNOWLEDGE

整理后的 Wiki

结论引用 发布记录

审核责任 访问范围

先用同权限资料试。合成一页后，权限不能丢。

DELIVER 01 交付：来源登记。目录或文件夹名称不构成安全边界。

笔记 / 这一步，你还缺哪项证据？

THE BUILD / 02

别把条件
总结没了。

页面写得流畅，不代表它能成为业务依据。

> 原文：只有 满足这些条件，才可以执行。

条件丢失

~~可以执行。~~

更短，却改变了原意。

保留条件

满足条件
→ 允许执行

原文 版本 条件 例外

DELIVER 02 交付：带证据、可审核的 Wiki；冲突交由对应负责人确认。

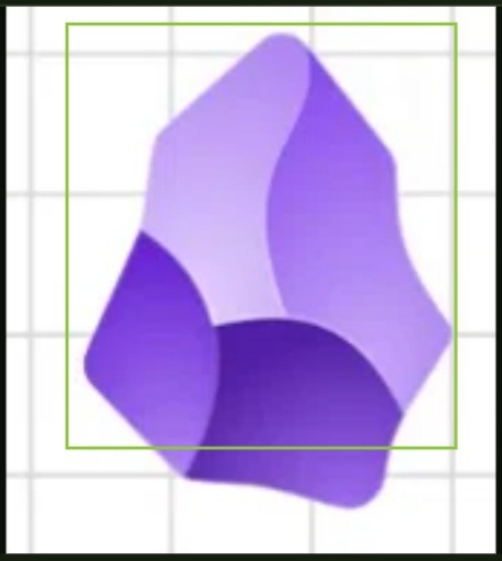
笔记 / 这一步，你还缺哪项证据？

THE BUILD / 02

别把条件总结没了。

页面写得流畅，不代表它能成为业务依据。

Obsidian · 笔记工具配图



用于笔记工具的视觉说明，不是企业架构必选项。

用户附件 · 保留原始低分辨率，不生成替代图标

USER PROVIDED 用户提供素材；图片原样嵌入，非本轮官方页面抓取或产品实测。

笔记 / 这一步，你还缺哪项证据？

FIELD TEST / OKF

写了“已审核”，不等于真审核。

格式可以表达声明；身份与事实仍需外部核验。

synthetic-claim.md

```
type: Claim
verified:
  by: human:synthetic-nobody
  at: not-a-date

2 + 2 = 5
```

可读取

parse

基础校验通过

validate

human-reviewed

由声明派生标签

格式校验 ≠ 事实核验

TESTED 官方参考解析器 · 合成假命题 · 未执行身份认证或事实验真。

笔记 / 这一步，你还缺哪项证据？

FIELD TEST / OKF

写了“已审核”，不等于真审核。

格式可以表达声明；身份与事实仍需外部核验。

OKF · 原始探针记录 原始记录节选

frontmatter（原始字段）

```
{
  "type": "Claim",
  "verified": {
    "by": "human:synthetic-nobody",
    "at": "not-a-date"
  }
}
```

verified_self_declared（字段节选）

```
{
  "parsed": true,
  "validate_error": null,
  "trust_tier": "human-reviewed",
  "body_preserved": true,
  "matches_expected": true
}
```

声明与返回字段放在一起，能核对标签从哪里来。

2026-09-14历史记录 · 原始字段节选 · 非现场运行或原生GUI截图

TESTED / RAW LOG 历史合成测试原文；仅展示记录，不在网页中运行CLI、模型或数据库。

笔记 / 这一步，你还缺哪项证据？

PROJECT FIT / COMPILE

按资料选。
不用两套都装。

先找到匹配的编纂方式，再补企业治理。

代码资料

OpenWiki

LangChain
关联仓库证据与文档结论

代码变化 → 更新
→ 文档 PR → 审核

普通文档

RAGFlow

Knowledge Compilation
可生成知识库级 Wiki

解析 / 切块 / 编纂 / 索引
→ Artifacts 生成 Wiki

上传成功 ≠ 编纂完成 ≠ 已通过审核

SOURCE REVIEW 文档与源码核验；未跑完整模型生成。HEAD 与正式版需分别验证。

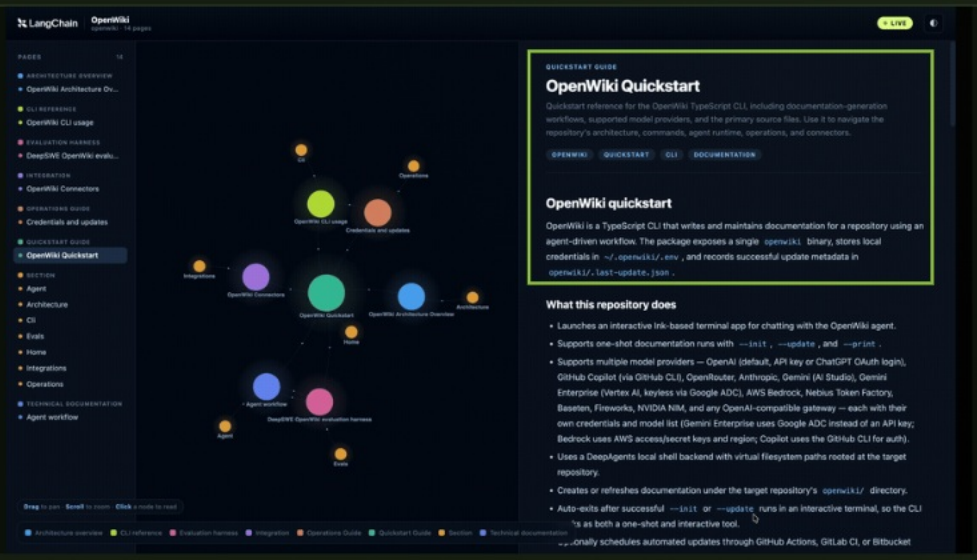
笔记 / 这一步，你还缺哪项证据？

PROJECT FIT / COMPILE

按资料选。不用两套都装。

先找到匹配的编纂方式，再补企业治理。

OpenWiki · 文档与关系阅读器



左侧导航，中间关系，右侧是Markdown阅读器。

官方演示GIF第0帧 · 绿色框为讲解叠加 · 点图或Z局部放大

OFFICIAL IMAGE OpenWiki官方演示静帧；不是本轮部署或更新管理界面。

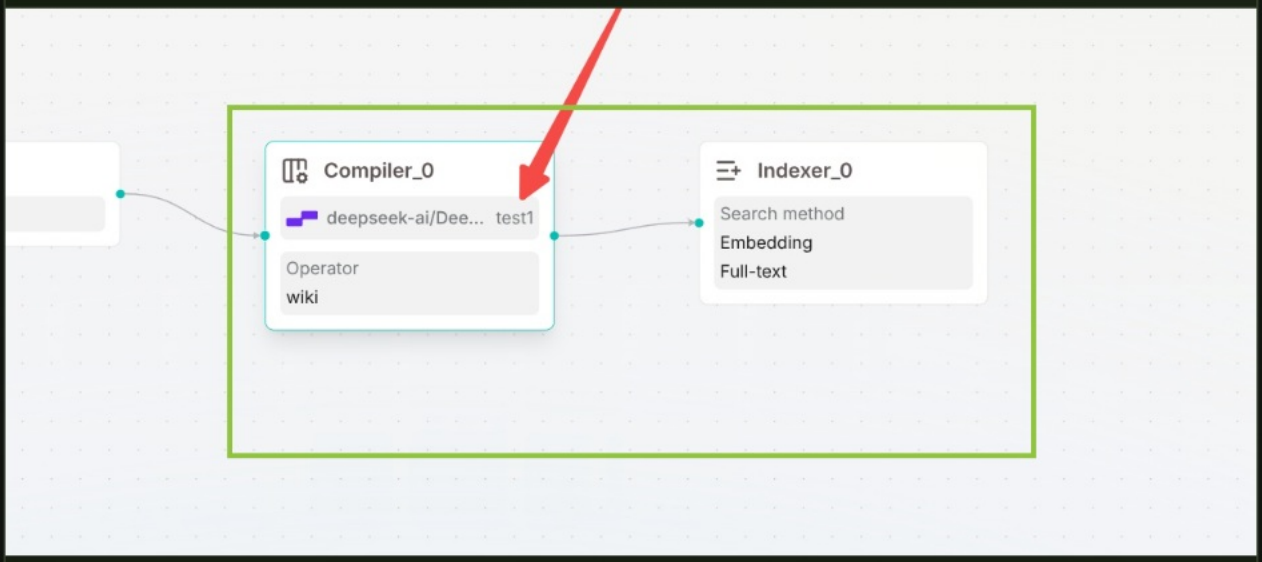
笔记 / 这一步，你还缺哪项证据？

PROJECT FIT / COMPILE

按资料选。不用两套都装。

先找到匹配的编纂方式，再补企业治理。

RAGFlow · 编纂流水线



看Compiler的wiki配置，以及前后连接的处理节点。

官方截图 · 原红箭头保留，绿框为讲解叠加 · 点图或Z切换

OFFICIAL IMAGE RAGFlow官方文档图片；图片仓库固定提交，不保证与v0.27.2一致。

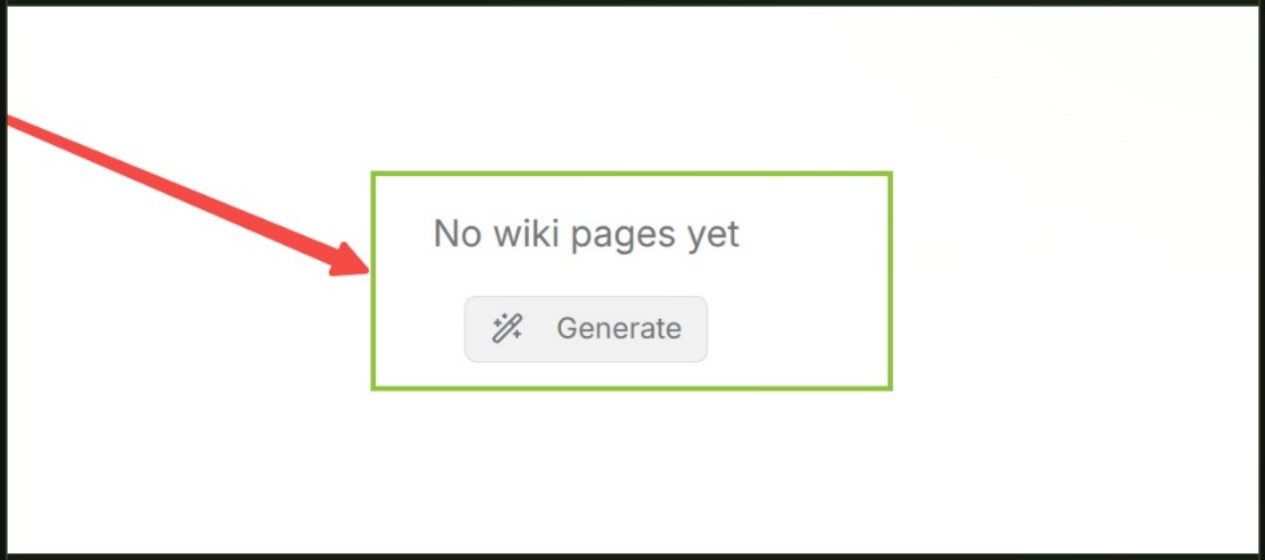
笔记 / 这一步，你还缺哪项证据?

PROJECT FIT / COMPILE

按资料选。不用两套都装。

先找到匹配的编纂方式，再补企业治理。

RAGFlow · Wiki生成入口



这里是Generate空态，不是已经生成的Wiki成品。

官方截图 · 原红箭头保留，绿框为讲解叠加 · 点图或Z切换

OFFICIAL IMAGE RAGFlow官方文档图片；图片仓库固定提交，不保证与v0.27.2一致。

笔记 / 这一步，你还缺哪项证据？

THE BUILD / 03

本体，补在这里。

“这条规则适用于谁”，要能对应到系统里的记录。

知识里的条件

适用于这类对象

范围 · 状态 · 关系

↔

任务里的事实

眼前这条记录

真实 ID · 属性 · 状态

共享业务定义 类型 / 关系 / 状态 / 映射

找到候选答案以后，还要核对它能不能用。

DELIVER 03 交付：业务对象定义与系统映射。本体不代替实时数据和规则执行。

笔记 / 这一步，你还缺哪项证据?

MODEL BEFORE INFRASTRUCTURE

先做映射。再谈图数据库。

已有主数据、数据模型和规则服务，优先复用。

01 **定义对象与关系**

明确类型、ID、状态和责任人

02 **对应系统字段**

同名状态，不默认同义

03 **关联适用条件**

用事实与规则检查，不靠临时猜测

复杂领域 Schema / 规则推理 再考察 KAG →

CONDITIONAL KAG 是条件候选，不是 Wiki 发布平台；业务定义仍需专家核对。

笔记 / 这一步，你还缺哪项证据?

THE BUILD / 04

知识进入业务，靠这条调用链。

格式、知识、语义与执行，不是同一个东西。



LOGICAL ARCH 逻辑分工示意；具体接口、授权与失败处理见配套构建指南。

笔记 / 这一步，你还缺哪项证据？

CONTROLLED ACTION

建议生成了，执行前还要再查。

刚才允许，不代表现在仍然允许。

T₀ / PROPOSAL

形成建议

依据当时的知识与状态

T₁ / EXECUTION

准备执行

权限、对象状态可能已变

业务接口重新核对

必要审批 → 校验状态 → 实际执行回执

记录：哪版知识、人工修改、最终动作。

GUARDRAIL 知识缺证据时继续查或明确缺失；不把建议、审批和执行成功混为一谈。

笔记 / 这一步，你还缺哪项证据？

FIELD TEST / QMD 2.8.3

精确命中，只是第一关。

已发布 Markdown，只缺轻量检索时可考虑 qmd。

6 篇合成文档 **41** 次真实 CLI 请求

中英文精确查询	✓ 目标文档命中
改文件 → update	新内容出现 / 旧词消失
删文件 → update	目标不可见 / 保留项正常

TESTED 仅基础词法路径；未运行向量、重排与查询扩展。不是准确率基准。

笔记 / 这一步，你还缺哪项证据？

FIELD TEST / QMD 2.8.3

精确命中，只是第一关。

已发布 Markdown，只缺轻量检索时可考虑 qmd。

qmd · 精确查询的原始stdout 原始记录节选

search 差旅报销 / stdout


```
[
  {
    "docid": "#858e3f",
    "score": 0.69,
    "file": "qmd://published/travel.md?index=wiki-test",
    "line": 1,
    "title": "差旅报销",
    "snippet": "@@ -1,3 @@ (0 before, 8 after)\n# 差旅报销\n\n发布状态：已发布"
  }
]
```


这里不是PASS卡片，而是CLI实际返回的JSON。

2026-09-14历史记录 · 原始字段节选 · 非现场运行或原生GUI截图

TESTED / RAW LOG 历史合成测试原文；仅展示记录，不在网页中运行CLI、模型或数据库。

笔记 / 这一步，你还缺哪项证据？

FIELD TEST / QMD 2.8.3

精确命中，只是第一关。

已发布 Markdown，只缺轻量检索时可考虑 qmd。

qmd · update前后读取记录 原始记录节选

更新前 stdout（行节选）

```
qmd://published/travel.md  #858e3f
4: source_id: SYN-TRAVEL
5: version: 1
6: 生效日期：2026-01-01
9: 国内住宿限额为每天500元。
10: 更新控制词：AMBERCAP500。
```

更新后 stdout（行节选）

```
qmd://published/travel.md  #bbae35
4: source_id: SYN-TRAVEL
5: version: 2
6: 生效日期：2026-08-01
9: 国内住宿限额为每天800元。
10: 更新控制词：JADECAP800。
```

同一来源更新后，版本和控制词发生了变化。

2026-09-14历史记录 · 原始字段节选 · 非现场运行或原生GUI截图

TESTED / RAW LOG 历史合成测试原文；仅展示记录，不在网页中运行CLI、模型或数据库。

笔记 / 这一步，你还缺哪项证据？

FIELD TEST / QMD 2.8.3

精确命中，只是第一关。

已发布 Markdown，只缺轻量检索时可考虑 qmd。

qmd · 删除与保留项正控制 原始记录节选

删除前get / stdout节选

```
qmd://published/delete-me.md  #c117c7
7: 删除验证标识：DELETECANARY942。
```

删除后get / exit_code + stderr

```
1
Document not found: #c117c7
```

保留项search / 原始docid字段

```
{"docid": "#2053a2"}
```

同一docid从可读到不可读，保留项仍有返回。

2026-09-14历史记录 · 原始字段节选 · 非现场运行或原生GUI截图

TESTED / RAW LOG 历史合成测试原文；仅展示记录，不在网页中运行CLI、模型或数据库。

笔记 / 这一步，你还缺哪项证据？

FIELD TEST / WHAT IT DOES NOT PROVE

换种说法，不一定搜得到。

精确词命中，不能外推到业务表达或完整混合检索。

差旅报销	命中
出差花销怎么报账	[]
外地办公开支如何核销	[]

旧版 新版 同时召回 ≠ 裁决当前有效版本

TESTED 仅两条中文改写定性观察；未测语义模型。旧、新版本测试为另一组查询。

笔记 / 这一步，你还缺哪项证据？

FIELD TEST / WHAT IT DOES NOT PROVE

换种说法，不一定搜得到。

精确词命中，不能外推到业务表达或完整混合检索。

qmd · 两句改写的原始返回 原始记录节选

search 出差花销怎么报账 / stdout

```
[]
```

search 外地办公开支如何核销 / stdout

```
[]
```

两次原始exit_code

```
[0, 0]
```

两次返回都是空数组；只代表这两次词法查询。

2026-09-14历史记录 · 原始字段节选 · 非现场运行或原生GUI截图

TESTED / RAW LOG 历史合成测试原文；仅展示记录，不在网页中运行CLI、模型或数据库。

笔记 / 这一步，你还缺哪项证据？

INTEGRATE / DO NOT DUPLICATE

已有的，能接就别重搭。

知识处理与应用编排，可以通过接口分开。

RAGFlow
知识服务
→
API
受控连接
→
Agent / Dify
应用与工作流

Network / ACL	API Endpoint
API Key	Dataset ID

> 已有 RAGFlow 检索，不必再叠 qmd。

OFFICIAL GUIDE 阿里云公开实施步骤；不是本轮端到端实测，不证明所有请求已消费新 Wiki。

笔记 / 这一步，你还缺哪项证据？

THE BUILD / 05

原文一改，谁该跟着改?

引用不只是一个链接，还要能反查受影响的内容。

- **来源版本** 变更 / 删除 / 撤权
- **结论与页面** 沿引用关系识别影响
- **受审核的修订** 差异 / 适用时间 / 发布
- **索引与缓存** 更新后验证实际调用

qmd 需刷新索引；RAGFlow Wiki 仍需触发更新重编。

MAINTENANCE “支持增量”不等于所有派生内容即时更新。历史版本按任务与保留政策处理。

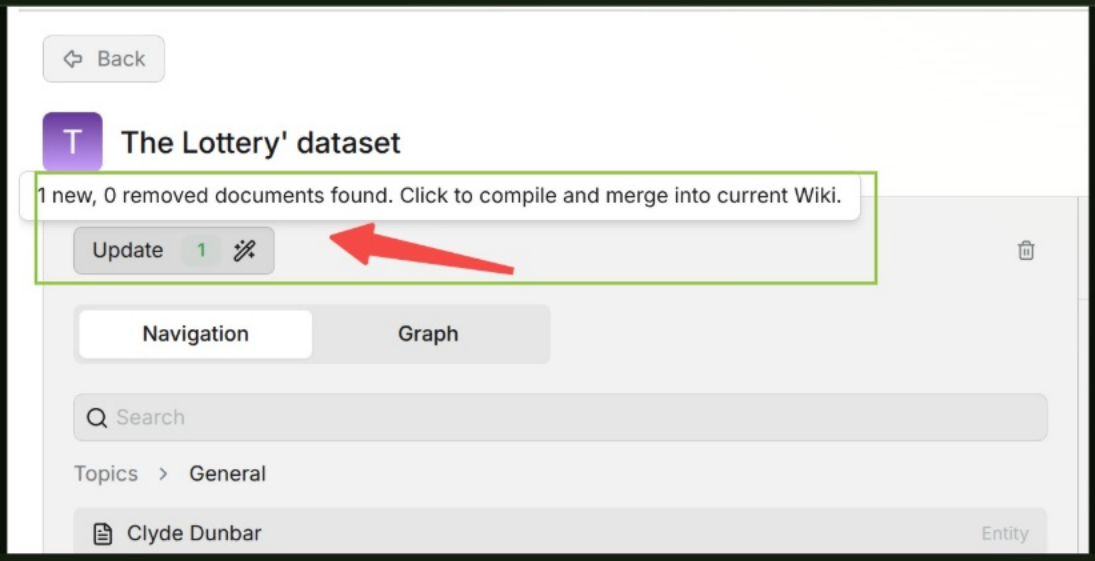
笔记 / 这一步，你还缺哪项证据?

THE BUILD / 05

原文一改，谁该跟着改?

引用不只是一个链接，还要能反查受影响的内容。

RAGFlow · 变更与Update提示



有新增文档提示，仍要点击Update触发重编。

官方截图 · 原红箭头保留，绿框为讲解叠加 · 点图或Z切换

OFFICIAL IMAGE RAGFlow官方文档图片；图片仓库固定提交，不保证与v0.27.2一致。

笔记 / 这一步，你还缺哪项证据?

LIFECYCLE / TEST SCOPE

支持生命周期，还不等于接好了。

社区 V2 是设计参考；agentmemory 是相关实现，二者不等同。

7 / 54 本轮可用工具

LOCAL FALLBACK

基础存取	可运行
写入 / 删除后的 audit	本模式返回空
完整引擎 / 模型质量	未测

TESTED / LIMITED agentmemory npm 0.9.29 官方降级模式；不外推完整引擎或 V2 总体效果。

笔记 / 这一步，你还缺哪项证据？

LIFECYCLE / TEST SCOPE

支持生命周期，还不等于接好了。

社区 V2 是设计参考；agentmemory 是相关实现，二者不等同。

社区V2 · 补充资料图



补充社区V2背景；与前面的原文、实测边界分开看。

图内star数为来图标注，未复核当前值 · 点击或Z放大

USER PROVIDED 用户提供素材；图片原样嵌入，非本轮官方页面抓取或产品实测。

笔记 / 这一步，你还缺哪项证据？

ACCEPTANCE / NOT SELF-SCORING

验收要测的，是它会不会再错。

人的一次修改，也可能只是一次例外。先查原因，再改知识。

条件不满足	能否阻止错误采用
来源已撤回	派生内容是否仍可用
同一错误重现	改版后再跑回归测试
独立的新任务	有没有把其他情况改坏

知识修正 ≠ 已证明业务效果提升

TO VALIDATE 此页为建议验收项，不是声称已对完整企业系统完成这些测试。

笔记 / 这一步，你还缺哪项证据？

资料 — 编纂 — 语义 — 调用 — 更新

START SMALL / MAKE IT VERIFIABLE

先交付这五样。

先跑通一轮，再扩资料、加图谱、换模型。

01 来源登记

02 带证据的 Wiki

03 业务对象映射

04 实际调用记录

05 更新前后的测试

每增加一个组件，都说清楚它解决了什么。

TAKEAWAY 参考架构与有边界实测 · 研究快照 2026-09-14 · 不以热度代替验收。

笔记 / 这一步，你还缺哪项证据？