K KNOWLEDGE BASE TRAINING

面向在校学生的入门实战课

企业级知识库

从概念到实战

读懂企业 IT 里的"第二大脑"，

并以 WeKnora 完成一次完整的分级配置实战



主讲 · AI 应用教研组 | 实战平台 · WeKnora

2026 / 09

目录

CONTENTS

五个模块，从企业视角理解知识库，
再到 WeKnora 上手实战。

01 知识库在企业IT中的位置
三层架构里的"能力中枢"，向下接数据、向上供应用

02 知识库的企业应用场景
智能外呼、智能客服、员工培训、战略决策……

03 知识库的种类与选型
FAQ 型 / 文档型 / 图谱型 / RAG 型，企业该怎么选

04 知识库的核心概念
RAG、Embedding、Rerank、Wiki、知识图谱、向量库

05 WeKnora 配置与实战
知识空间 → 知识库 → 文件，五步完成分级配置

个人知识库 vs 企业知识库

同样是"知识库"，个人要的是好用，企业要的是可信

对比维度	个人知识库	企业知识库 本次主题
结果严谨程度	够用就行，偶尔出错可接受	答案可溯源、口径统一、可审计
易用程度	越简单越好，单人即开即用	全员易上手，更重权限与治理
标准刚性要求	随性记录，无强制规范	命名 / 分类 / 格式 / 审批强统一
部署私有性	云端或本地，数据自己管	数据不出内网，私有化 / 专有云
协作与权限	单人维护为主	多人协作，角色权限隔离
更新与治理	自己更新即可	版本管理、审核、责任到人

从「个人好用」到「组织可信」——企业知识库多了一层治理与安全的刚性约束，这正是本课强调"企业级"的原因。

CHAPTER 01

知识库在企业IT中的位置

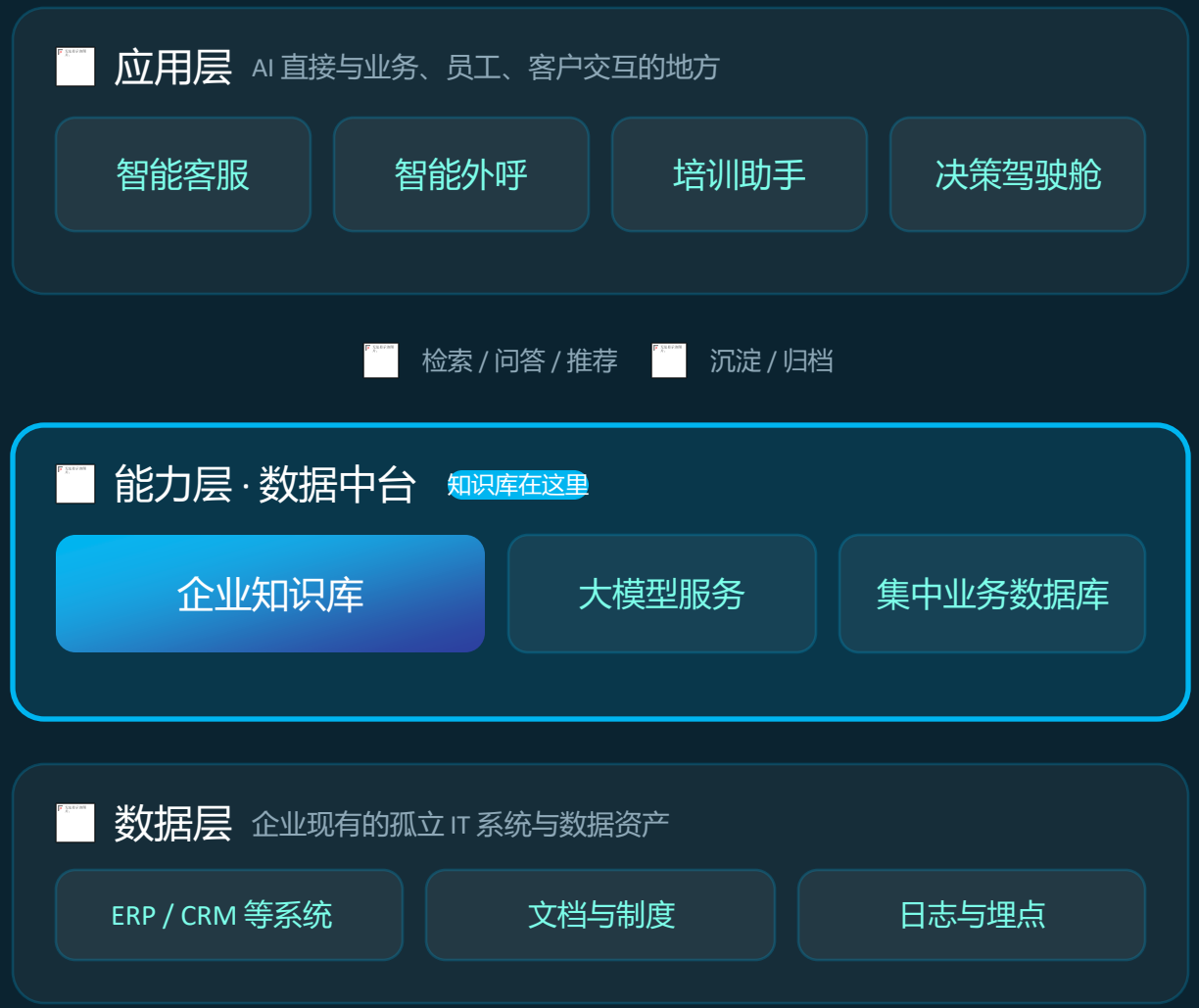
01

先看清全景：企业 IT 系统如何分层，

知识库为什么是其中不可替代的"能力中枢"。

知识库在企业IT系统中的位置

三层架构：向下承接数据资产，向上支撑智能应用



承上启下的"能力中枢"

知识库不直接面对客户，而是把散落的数据变成"可被 AI 调用的知识"，供上层应用统一取用。

一次建设，多处复用

客服、外呼、培训、驾驶舱共用同一份知识资产，避免每个系统重复维护文档。

打破孤岛，集中供数

ERP、CRM 等系统各自孤立，知识库把分散数据统一沉淀为"可被 AI 调用的知识"，应用无需逐一对接。

> 一句话定位：知识库 = 企业数据的"AI 就绪层"——没有它，大模型只能通用闲聊；有了它，AI 才能说"企业自己的话"。

为什么需要企业知识库

知识库不是"网盘备份"，而是企业运转方式的升级

多数企业的现状：知识"散落"且"沉默"

制度文档躺在网盘深处，项目经验留在离职同事的电脑里，客户问答全靠老师傅记忆——知识存在，但 AI 和新员工都找不到、用不上。

建立知识库之后：知识"入库"且"在线"

所有资料统一解析、切分、入库存档；员工用自然语言提问，AI 基于企业自己的文档作答——知识从"找文件"变成"问问题"。

> 企业最大的浪费，是让已经付过学费的经验反复重新学习。

数据私有，不外泄

私有化部署后资料不出企业内网，公网大模型碰不到核心数据

知识沉淀，可复用

项目复盘、客户话术、踩坑记录持续入库，人走了知识留下

新人上手更快

制度、规范即问即答，新员工不用再"追着老员工问"

回答有据可溯

AI 答案附引用来源，可信可查，避免"一本正经地胡说"

CHAPTER 02

知识库的企业应用场景

02

从打电话到做决策：看看领先企业

把知识库用在了哪些"真金白银"的业务里。

知识库在企业中的应用全景

业务跑到哪里，知识库就支撑到哪里

6

大核心应用场景

几乎所有企业级 AI 应用，
第一步都是让 AI 读懂
"企业自己的资料"。

判断：知识库已成为企业 AI 落地的通用底座，而非某个部门的小工具。

智能外呼

话术+客户画像自动外呼，7×24 触达客户

智能客服

FAQ 秒回、多轮对话，工单自动分流

员工培训

制度规范即问即答，新人带教自动化

战略决策

行业情报+经营数据沉淀，支撑管理层研判

研发辅助

规范、API 文档、历史方案随时检索引用

营销内容

基于产品资料批量生成文案与投放素材

场景拆解一：智能外呼与智能客服

对话类场景：知识库既"喂话术"，也"收回经验"

智能外呼：一条越用越聪明的闭环

1. **客户名单与画像进入系统**
   来源、等级、历史沟通一键带出
2. **知识库匹配话术与应答策略**
   按产品线、地区、客户类型自动选择
3. **AI 外呼机器人自动对话**
   7×24 小时批量触达，旺季不占人力
4. **实时识别意图，动态调整话术**
   感兴趣→深入介绍；拒绝→礼貌收尾
5. **通话记录回流知识库**
   优质应答沉淀为新话术，下次直接复用

第 5 步的经验重新进入第 2 步——闭环让系统随使用持续进化

智能外呼

主动触达

- 营销触达：新产品、活动邀约批量外呼，话术统一不出错
- 回访催办：贷款提醒、售后回访、账单催缴全自动化
- 线索清洗：先由 AI 粗筛意向，人工只跟进高意向客户

智能客服

被动应答

- FAQ 秒回：退换货、资费、网点查询等高频问题直接命中
- 多轮对话：结合上下文连续追问，不要求用户重复描述
- 工单分流：复杂问题自动带摘要转人工，坐席效率倍增

场景拆解二：员工培训与战略决策

对内：让每个人快速上手；向上：让决策有据可依

员工培训：把"老师傅"装进系统

制度、产品手册、操作规范统一入库后，培训从"集中授课"变为"随用随学"——新人遇到的 90% 的问题，都能在知识库里得到与官方口径一致的答案。

即问即答

报销流程、休假制度、审批链路随时查

新人带教

按岗位推送学习路径与必读清单

自动考核

基于文档生成题库，学完即测

客户投诉质量问题的处理时限是多少？

根据《客户服务管理规范》第 4.2 条：24 小时内首次响应，48 小时内给出处理方案。 来源：《客户服务管理规范》v3.1

▲ 培训助手实测：答案可溯源到具体制度条款，口径与官方一致

战略决策

把经营数据、行业情报、历史复盘统一沉淀，管理层问"值不值、进不进、怎么打"，知识库用事实回答。

- **汇聚** 经营数据 + 行业报告 + 竞品动态
- **整合** 知识库去重归档，建立情报专题库
- **研判** 生成决策简报，附数据与来源引用

典型用法：市场进入评估 · 竞品动态周报 · 经营异常预警

CHAPTER 03

知识库的种类与选型



FAQ、文档、图谱、RAG 各有分工——

企业该把钱花在哪一种上?

知识库的四大种类

按"知识形态"与"智能程度"排列：从人工维护到自动理解

◀ 人工维护为主 · 建设成本低

▲ 企业部署主流选择

AI 自动理解 · 智能程度高 ▶

01 FAQ 型

02 文档型

03 图谱型

04 RAG 型

FAQ 型

一问一答的知识对（Q&A、SOP）。命中精准、上线快，但覆盖面有限，适合高频客服问题。

文档型

制度、手册等长文档的目录化管理和关键词全文检索。建设快，但"搜得到"不等于"答得准"。

图谱型

把人、产品、事件建成实体关系网络。擅长关联推理与风控溯源，但构建成本最高。

RAG 型

全量非结构化文档向量化后语义检索，自然语言直接提问。企业部署的主流形态。

✦ 本次培训的实战对象

适合企业部署的知识库类型

类型定了，还要选对"部署形态"——数据放在哪里是底线问题

企业级正解

RAG 型知识库 + 私有化部署

语义问答体验 + 数据不出内网，是中大型企业能同时兼顾"好用"与"安全"的组合。

补充：已有 Wiki/图谱资产的企业，可在 RAG 底座上叠加图谱增强，不必推倒重来。

公网 SaaS 知识库

轻量起步

开通即用、按需付费，适合小微企业与非敏感资料；但核心数据上传第三方，存在合规与泄露风险。

适用：初创团队 · 公开资料 · 内部轻量协作

私有化部署（本地/专有云）

企业主流

知识库与大模型都跑在企业自己的服务器上，数据不出域、模型可定制、权限可审计；需要一定算力与运维投入。

适用：金融/制造/政务 · 涉密文档 · 合规要求高

混合部署

折中方案

敏感文档留在私有环境，公开资料放 SaaS；通过权限分层兼顾成本与安全，架构与治理更复杂。

适用：多分支机构 · 数据密级差异大

1. 数据敏不敏感?
2. 文档量级多大?
3. 有没有人管运维?

CHAPTER 04

知识库的核心概念

RAG、Embedding、Rerank、Wiki、知识图谱、向量库——六个词，一次讲透。

核心概念（上）： RAG · Embedding · Rerank

三个词构成一条链：知识如何被 AI"找到并讲对"



RAG

检索增强生成 · Retrieval-Augmented Generation

让大模型回答前"先查资料再开口"：从企业文档中检索相关片段，连同问题一起交给模型，生成有依据的答案。

解决：模型不知道企业内部知识、爱编造（幻觉）



Embedding

向量化 · Embedding

把文字压缩成一串数字（向量），语义越接近、向量距离越近——"退货政策"和"退款规则"会被放在一起。

解决：机器如何"理解语义"而不是只匹配关键词



Rerank

重排序 · Rerank

先粗筛出 50 条候选，再由精排模型按"与问题的相关度"重新打分排序，只把最相关的几条交给大模型。

解决：检索结果太多太杂，答案质量不稳定



一个类比：企业知识库 = 图书馆

文档入库	图书采购登记
Embedding 向量化	给每本书编索书号
向量数据库	按索引排列的书架
语义检索	按主题找书，而不是只按书名
Rerank 精排	从一堆候选里挑最相关的三五本
大模型生成	读完这几本，替你写一份摘要

记忆口诀：先入库、再索引、后检索、精排序、生成答案带引用。

核心概念（下）：Wiki · 知识图谱 · 向量数据库

知识库不是孤立产品，而是"知识从哪来、如何关联、怎么被找到"的组合

三个概念，回答三个问题

- 知识从哪里来？ Wiki
- 知识如何关联？ 知识图谱
- 知识怎么被找到？ 向量数据库

实际项目中三者常常同时存在：Wiki 供员工沉淀，向量库供 AI 检索，图谱补足复杂关联。

Wiki · 协作知识库

定位：知识的"上游产地"

Confluence、语雀等工具里由员工共同编辑维护的结构化文档体系。它解决"知识持续产出"的问题，是 RAG 知识库最重要的原料来源——Wiki 更新，AI 的答案随之更新。

关键词：共同编辑 · 版本管理 · 组织沉淀

知识图谱 · 实体关系网络

定位：知识的"关系网"

把人、产品、订单、流程抽象为实体，再用关系互相连接。擅长多跳关联推理，例如"A 供应商的哪些原料来自 B 地区"。风控溯源、供应链穿透等场景价值极高。

关键词：实体 · 关系 · 多跳推理

向量数据库 · 语义检索基础设施

定位：RAG 的"发动机"

专门存储 Embedding 向量并提供相似度检索的数据库，常见有 Qdrant、Milvus、Faiss 等。文档向量化入库后，用户的自然语言问题也转成向量去"最近邻"匹配。

关键词：Embedding 存储 · 相似度检索 · 低延迟

一张图看懂 RAG 全流程

本章核心

离线阶段

建库

- **文档解析**：PDF/Word/网页统一抽取
- **切分 Chunking**：长文按语义拆成小段落
- **Embedding**：每个片段转成向量
- **入库**：向量存入向量数据库

同一份向量库存，被在线问答反复调用

在线阶段

问答

- **用户提问**：自然语言直接问
- **问题向量化**：用同一模型转向量
- **相似检索**：召回 Top-K 相关片段
- **Rerank 精排**：按相关度重新排序
- **生成答案**：大模型综合片段作答
- **答案+引用**：附来源，可溯源

关键理解：大模型本身没有被"喂"企业文档，而是每次回答时现场查资料——所以文档一更新，AI 立刻"学会"。

两处最容易踩坑：切分粒度太粗/太细、检索阶段缺少 Rerank——都直接决定答案质量。

知识库的调优：方向由知识类型决定

上线只是开始，真正拉开差距的是持续调优——而调优没有"一套参数打天下"

法律 · 医学 · 合规

求「准」 术语密集、边界严格，错一个字都可能出事故。

调优方向

- Rerank 精排拉满，只留最相关片段
- 检索阈值调高，宁缺毋滥
- 强制引用原文，答案可溯源
- 拿不准就答"不确定"，绝不编造

准确性 引用溯源 拒答兜底

管理 · 咨询

求「全」 方法论、框架、流程与案例，要结构化、逻辑自洽。

调优方向

- 引入知识图谱，串起方法间关联
- 混合检索：关键词 + 语义兼顾
- 分块按方法论模块切，保逻辑完整
- 拉大上下文窗口，容纳完整框架

结构化 逻辑链 方法论复用

文学 · 创意

求「活」 风格、语感与灵感，答案本就没有标准。

调优方向

- 放宽检索阈值，召回更多候选
- 增大 Top-K，给模型更多素材
- 调高生成随机性，鼓励发散
- 保留大块上下文，护住语感语境

风格保留 灵感发散 适度放开

先判断知识类型，再定调优方向：准、全、活三种取向，对应三种完全不同的调法——这也是企业知识库"私有化 + 按领域分库"的价值所在。

CHAPTER 05

WeKnora 配置与实战

以开源框架 WeKnora 为例，

动手完成一次"知识空间 → 知识库 → 文件"的分级配置。

WeKnora 的分级结构：知识空间 → 知识库 → 文件

三级结构是 WeKnora 一切配置的骨架：先分区、再建库、后传文件

知识空间 Knowledge Space

一级：命名空间，按团队/事业部隔离，成员与权限在这里管理

知识库 A「产品资料」

二级：独立配置单元

Embedding 模型　分块策略　检索配置

- 产品手册.pdf
- FAQ.xlsx
- 更新日志.md

三级：文件上传后自动解析并切分为 Chunk

知识库 B「制度规范」

- 员工手册.docx
- 财务制度.pdf
- 差旅标准.xlsx

同一个空间下可建多个知识库，各自独立配置

知识空间

一套部署可划出多个空间：按子公司、事业部或项目隔离；用户、角色、配额在空间层统一管理，互不可见。

知识库

按业务主题建库（产品、制度、客服话术……）；每个库绑定自己的模型、分块与检索参数，是"配置的基本单位"。

文件 / Chunk

最小导入单元。上传即触发解析→切分→向量化；更新文件只需重新导入，不影响同库其他内容。

分级的好处：空间隔离权限 → 库级精细配置 → 文件级灵活更新，大小团队都能各管各的

WeKnora 配置与使用五步法

从空系统到能回答问题，标准路径就这五步

课前准备清单

- **部署 WeKnora**
  Docker Compose 一键启动全套服务
- **准备模型**
  配好大模型与 Embedding 模型的 API
- **整理文档**
  挑 10–20 份代表性资料作为首批语料

新手建议：先只用一种文档类型跑通全流程，再逐步扩展格式与知识库数量。

1 创建知识空间

✦ 一步到位定边界，后期迁移成本高

规划隔离维度（部门/事业部/项目），添加成员并分配角色权限

2 创建知识库并配置

✦ 库级参数决定后续问答质量的地基

选 Embedding 模型、设置分块大小与重叠、选向量/混合检索方式

3 上传文件并解析

✦ 抽查切分结果：段落是否完整、有无截断

批量导入 PDF / Word / Excel / Markdown，等待解析与向量化完成

4 配置检索与 Rerank

✦ 用 10 个真实问题测试命中率再调参

开启 Rerank 精排，调整 Top-K 数量与相似度阈值

5 接入应用并持续评测

✦ 知识库是"养"出来的，不是一次配好的

接入问答界面或 API，定期用新问题评测，迭代分块与参数

THANKS FOR WATCHING

让知识成为企业的第二大脑

课后动手：用 WeKnora 建好你的第一个「知识空间 → 知识库 → 文件」，

并用 10 个真实问题检验它的回答质量。

1. 建一个空间

以你的班级或社团为隔离单位

2. 传一批文档

社团章程 / 课程笔记 / 活动复盘

3. 问十个问题

检验命中率，试着调优分块参数