



OmniTian

企业级 AI 知识与智能体工作平台

产品白皮书

版本 V1.0

2026-08-21

目录

目录.....2

一、前言3

二、行业背景与挑战.....3

三、产品定位与设计理念4

 3.1 产品定位.....4

 3.2 设计理念.....4

四、系统架构5

 4.1 接入层5

 4.2 应用层5

 4.3 能力引擎层.....6

 4.4 数据与治理层.....6

五、核心功能模块.....7

 5.1 AI 知识问答（检索增强生成·RAG）7

 5.2 Cowork 智能体：自主任务执行8

 5.3 知识文库与知识检索9

 5.4 文案办公：一键生成正式文档10

 5.5 会议纪要：录音转写与结构化整理.....10

 5.6 提示词市场.....10

六、引擎与治理中心.....11

 6.1 多模型引擎中心11

 6.2 MCP 外部工具协议.....11

 6.3 数据看板与费用看板12

 6.4 部门与权限治理12

 6.5 身份认证与登录集成.....12

 6.6 敏感信息保护.....12

七、安全与合规.....13

八、部署与集成.....13

 8.1 部署方式.....13

 8.2 与企业现有生态的集成.....13

九、典型应用场景.....14

十、总结与展望.....14

一、前言

在企业数字化转型进入深水区的今天，知识分散、协作低效、重复性文案与事务性工作占用大量人力，已成为制约组织效率提升的共性问题。与此同时，通用大语言模型虽然能力强大，却普遍存在“不了解企业内部知识”“无法直接完成业务动作”“缺乏权限与合规约束”等落地障碍。

问小天 (OmniTian) 是一套面向企业内部场景设计的 AI 知识与智能体工作平台，将知识库检索增强问答 (RAG)、自主任务执行智能体 (Agent)、办公文档自动生成、会议纪要转写、多模型引擎管理与企业级权限治理整合于统一入口，帮助企业以较低的实施成本，把大模型能力转化为员工可以直接使用的生产力工具。

让每一位员工都拥有一个懂企业知识、能动手做事的 AI 同事。

本白皮书旨在从产品定位、系统架构、核心能力、安全治理与应用场景等维度，系统介绍问小天平台的设计理念与技术实现，为企业客户、合作伙伴与内部决策者提供全面参考。

二、行业背景与挑战

大模型技术的成熟为企业智能化带来了新的可能性，但在实际落地过程中，企业普遍面临以下挑战：

- 知识分散：制度文件、产品文档、历史邮件与会议记录散落在不同系统中，员工检索成本高，新人上手慢。
- 模型选型不确定：国内外模型供应商众多，接口协议不一，企业难以在成本、效果、合规之间灵活权衡，且容易被单一供应商锁定。
- 数据安全与权限合规：企业知识往往涉及部门机密与个人信息，直接接入公有云通用问答工具存在数据外泄与越权访问风险。
- 任务停留在“问答”层面：员工需要的不仅是一个答案，而是能直接产出文档、表格、纪要等工作成果的能力。

- 协同与集成成本高：企业已有企业微信、飞书、钉钉等办公生态，新工具若不能融入现有身份体系与 workflow，很难被真正采用。

问小天正是围绕上述痛点展开设计：以知识库为地基、以多模型引擎为动力、以智能体为执行手臂、以部门权限与合规治理为护栏，构建可在企业内部私有化部署、安全可控的一体化 AI 工作平台。

三、产品定位与设计理念

3.1 产品定位

问小天定位为“企业专属的统一 AI 工作台”，覆盖从日常知识问答、文案办公，到复杂多步骤任务自动执行的完整需求光谱，是知识管理系统、办公自动化工具与智能体平台三者能力的融合体。

3.2 设计理念

- 有据可查：所有知识类回答均基于企业自有知识库检索结果生成，而非模型凭空生成，答案可追溯至原始文档。
- 模型中立：采用统一的 OpenAI 兼容协议对接各类大模型服务商，企业可自由选择或混合使用不同模型，避免被单一厂商锁定。
- 权限先行：知识库、对话、任务在部门与权限模板体系下运行，“看不到的知识就检索不到”，从源头保障数据安全。
- 知行合一：AI 不止负责回答问题，更能通过 Cowork 智能体调用工具、生成文档，直接产出工作成果。
- 开放集成：通过 MCP (Model Context Protocol) 协议对接企业已有的第三方系统与工具，通过 OAuth2 对接企业微信 / 飞书 / 钉钉等现有身份体系，降低推广门槛。

四、系统架构

问小天采用清晰的分层架构设计，自下而上分别为数据与治理层、能力引擎层、应用层与接入层，各层之间通过标准化接口解耦，便于独立扩展与替换。



图 4-1 问小天系统总体架构图

4.1 接入层

支持标准 Web 浏览器访问，并提供 PWA (Progressive Web App) 离线能力，保证弱网环境下的基础可用性；同时支持通过企业微信、飞书、钉钉等企业身份体系单点登录 (OAuth2)，无需员工重新记忆账号密码。

4.2 应用层

面向终端用户的六大核心功能模块：AI 知识问答、Cowork 智能体、文案办公、会议纪要、知识库与检索、提示词市场，均通过统一的顶部导航与首页入口呈现，用户可按需切换。

4.3 能力引擎层

包括多模型引擎中心（统一管理各类大模型的接入配置）、工具调用（Function Calling）框架、MCP 外部工具协议适配层、向量检索（Embedding）服务，以及支撑大文件异步解析的任务队列，是平台对外能力扩展的核心枢纽。

4.4 数据与治理层

以 MySQL 关系型数据库为核心存储，承载知识库文档与分片、部门与权限模板、费用与用量统计、敏感信息识别规则等治理类数据，为上层能力提供合规、可审计的数据基础。

五、核心功能模块

问小天以“统一 AI 工作台”为核心理念，围绕企业员工日常最高频的知识获取、任务执行与文案办公场景，构建了六大核心功能模块。

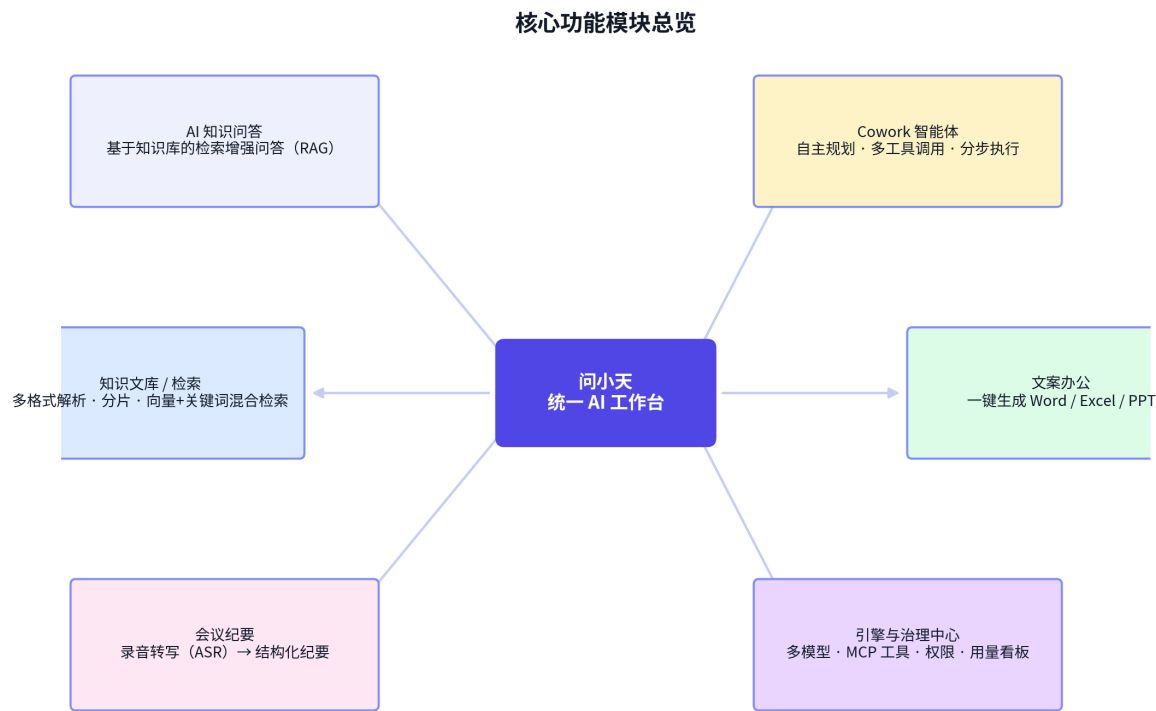
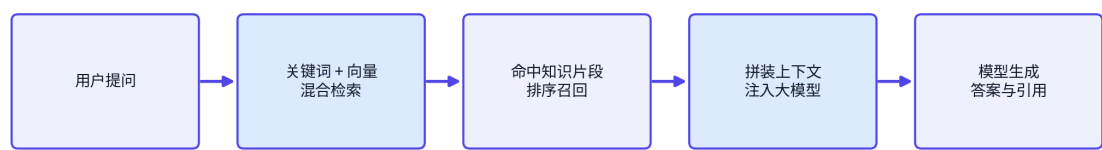


图 5-1 核心功能模块总览

5.1 AI 知识问答（检索增强生成 • RAG）

AI 知识问答是员工获取企业知识最直接的入口。用户提出问题后，系统在其有权限访问的知识库范围内，采用“关键词检索 + 向量语义检索”混合策略召回最相关的知识片段，再将其拼装为上下文注入大模型，生成有据可查、附带引用来源的回答，避免模型“凭空编造”。

知识问答 (RAG) 工作流程



权限过滤：按部门 / 知识库可见范围自动过滤检索结果，确保回答“有据可查、权限可控”

图 5-2 知识问答 (RAG) 工作流程

在检索能力上，系统内置支持 OpenAI Embeddings 协议的向量服务对接（如 BAAI/bge 系列中文向量模型），并将向量相似度与关键词命中按可配置权重融合排序，兼顾语义理解与精确匹配，适合中小规模知识库的高性价比部署；同时预留了后续平滑升级至专用向量数据库的扩展空间。

5.2 Cowork 智能体：自主任务执行

对于超越“一问一答”的复杂目标，用户可以在 Cowork 中用一句话描述任务目标（并可上传附件），智能体会自主规划步骤、按需调用工具、持续推进，直至任务完成或明确无法完成时向用户说明原因。

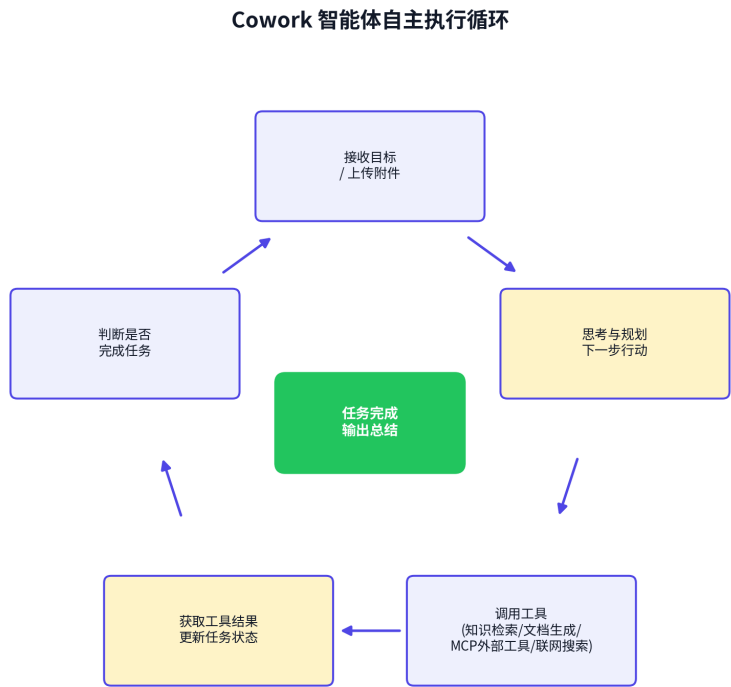


图 5-3 Cowork 智能体自主执行循环

Cowork 智能体当前可调用的典型工具包括：

- 知识检索类：列出知识库、检索知识片段、按标题获取文档完整内容；
- 文档生产类：一键生成 Word / Excel / PPT，或根据会议记录直接生成结构化会议纪要文档；
- 通用能力类：精确计算器、当前日期时间、联网搜索最新信息、图片生成；
- 安全合规类：敏感信息（手机号、身份证号、银行卡号、邮箱、IP 等）检测与脱敏；
- 个性记忆类：记录用户长期有效的偏好与背景信息，供后续任务自动复用；
- 外部集成类：通过 MCP 协议调用企业接入的第三方系统工具（详见 6.2 节）。

5.3 知识库与知识检索

知识库支持 PDF、Word、TXT、Markdown、Excel、PPT 等多种格式文档的批量上传（单文件最大 100MB，单次最多 20 个文件），系统自动完成文本解析、分片（默认每片约 800 字，片间重叠 100 字以避免语义割裂）与向量化处理。对于大文件，解析任务会自动转入后台

异步队列处理，避免上传请求超时，提升大规模知识库建设的稳定性。知识检索页面则为用户提供独立于对话场景的“搜索引擎式”检索体验，便于快速定位原文出处。

5.4 文案办公：一键生成正式文档

文案办公模块让用户通过自然语言描述即可生成正式的 Word 文档、Excel 表格或 PPT 演示文稿，可直接下载使用，大幅减少格式排版等重复性劳动，尤其适合周报、方案、数据整理表、汇报材料等高频办公场景。该能力同时作为工具被 Cowork 智能体调用，实现“调研 + 生成文档”的端到端自动化。

5.5 会议纪要：录音转写与结构化整理

会议纪要模块采用“浏览器本地录音 + 服务端转写”的架构：浏览器仅使用标准 MediaRecorder API 在本地录制音频，不依赖任何特定云端语音服务，录制完成后上传至企业自有服务器，再转发给可自由配置、兼容 OpenAI Whisper 协议的语音转写接口。管理员可根据部署所在地区选择连通性最佳的服务商，从而保证境内、境外均可稳定使用。转写文本经大模型整理后，自动生成包含概述、要点、决议与待办事项的结构化会议纪要，并可导出为 Word 文档。

5.6 提示词市场

提示词市场为团队提供高质量提示词（Prompt）模板的沉淀与复用能力，帮助不同岗位的员工无需具备提示词工程经验，也能通过“套用模板”的方式获得稳定、专业的 AI 输出效果，逐步沉淀为企业自身的 Know-how 资产。

六、引擎与治理中心

6.1 多模型引擎中心

问小天不绑定单一大模型厂商，而是统一通过 OpenAI 兼容协议对接各类模型服务，管理员可在“引擎中心”后台自由添加、切换、并行管理多个模型实例，并为每个模型单独配置输入/输出单价，便于成本核算与合理调度。目前系统预置以下主流服务商的快速接入模板：

服务商	接入协议	默认模型示例	说明
DeepSeek	OpenAI 兼容协议	deepseek-chat	系统默认预设
OpenAI	OpenAI 官方协议	gpt-4o-mini	支持视觉理解可选开启
Moonshot (Kimi)	OpenAI 兼容协议	moonshot-v1-8k	
智谱 GLM	OpenAI 兼容协议	glm-4-flash	
通义千问	OpenAI 兼容协议	qwen-plus	
豆包 (火山方舟)	OpenAI 兼容协议	doubao-pro-32k	
自定义服务商	任意 OpenAI 兼容接口	按需填写	满足私有化 / 信创模型接入需求

这种“协议统一、厂商可插拔”的设计，既降低了企业更换或新增模型供应商的迁移成本，也为未来接入信创、私有化部署大模型预留了空间。

6.2 MCP 外部工具协议

MCP (Model Context Protocol) 是问小天对接企业已有系统与第三方能力的标准化通道。管理员在“MCP 服务”中配置外部工具服务器地址后，系统会自动完成握手、拉取可用工具列表，并将其转化为 Cowork 智能体可调用的工具定义。这意味着企业可以将内部 OA、CRM、工单系统等能力以“工具”的形式暴露给智能体，使其在自主执行任务时具备操作外部系统的能力，而无需对问小天本身进行代码改造。

6.3 数据看板与费用看板

管理员可通过“数据看板”实时掌握知识库规模、文档解析状态、用户活跃度等运营指标；

“费用看板”则基于各模型配置的输入/输出单价与实际调用的 Token 用量，自动计算并汇总 AI 使用成本，支撑企业对 AI 投入产出比进行量化管理。

6.4 部门与权限治理

知识库支持按可见范围（全员 / 指定部门 / 仅个人）进行精细化授权，并可套用预设的权限模板批量配置，管理员与知识库所有者可随时调整授权部门。系统在检索、问答、Cowork 任务等所有涉及知识调用的环节，均会依据用户所属部门与知识库可见范围自动过滤结果，确保“员工只能看到、检索到自己有权限的内容”。

6.5 身份认证与登录集成

除账号密码登录外，系统支持 GitHub、Google 等标准 OAuth2 登录，以及企业微信自建应用、飞书自建应用、钉钉扫码登录等企业级身份源集成，管理员可在“登录设置”中按需启用，帮助企业将问小天无缝纳入现有身份认证体系，降低员工使用门槛。

6.6 敏感信息保护

平台内置敏感信息检测工具，可自动识别文本中的手机号、身份证号、银行卡号、邮箱地址、IP 地址等敏感信息并进行脱敏（打码）处理，该能力既可在对话中按需调用，也可作为 Cowork 智能体的工具，在生成对外文档前自动执行合规检查。

七、安全与合规

- 数据自主可控：系统可私有化部署于企业自有服务器或云环境，知识库文档、对话记录、向量数据均存储于企业自有数据库，不经由第三方通用问答平台留存。
- 权限最小化原则：知识检索、文档访问、任务执行均以“部门 + 知识库可见范围”为边界，从源头避免越权访问。
- 传输与鉴权安全：管理端接口、队列触发接口均设置访问令牌（Token）校验；OAuth2 登录采用标准授权码模式，降低账号密码泄露风险。
- 敏感信息识别：内置手机号、身份证号、银行卡号、邮箱、IP 等常见敏感信息的识别与脱敏能力，降低数据外泄风险。
- 输出可追溯：知识问答类回答保留引用来源，费用看板保留每一次模型调用的用量与成本记录，便于审计与复盘。

八、部署与集成

8.1 部署方式

问小天基于 PHP + MySQL 的成熟技术栈构建，对服务器环境要求低，可灵活部署于企业自有机房、私有云或公有云环境，满足不同规模企业对信创、等保合规等场景的部署诉求。前端支持 PWA 离线基础访问能力，在弱网或离线场景下仍可保障核心页面可用。

8.2 与企业现有生态的集成

- 身份集成：企业微信 / 飞书 / 钉钉扫码登录，账号体系与企业通讯录天然打通；
- 系统集成：通过 MCP 协议接入 OA、CRM、工单、DevOps 等企业内部系统工具；
- 模型集成：兼容并支持在国内外主流大模型服务商之间自由切换，亦支持私有化部署模型的接入；
- 数据集成：知识文库支持常见办公文档格式的批量导入，平滑承接企业存量文档资产。

九、典型应用场景

应用场景	核心诉求	对应能力
企业制度与流程答疑	员工快速获取报销、考勤、入职等制度性问题的准确答案	AI 知识问答 + 知识文库
新员工快速上手	降低新人查找资料、请教同事的时间成本	知识检索 + 提示词市场
周报 / 方案 / 汇报材料撰写	将零散信息快速整理为规范文档	文案办公 + Cowork 智能体
会议记录整理	将冗长的会议录音转化为结构化纪要与待办	会议纪要 (ASR)
跨部门知识共享	在保障数据安全前提下促进部门间知识流通	部门权限治理 + 知识文库
复杂多步骤任务代办	“调研 + 检索 + 生成文档”一站式自动完成	Cowork 智能体 + MCP 外部工具
AI 使用成本管控	量化各部门 / 各模型的 Token 消耗与费用	引擎中心 + 费用看板

十、总结与展望

问小天 (OmniTian) 通过 “知识库 RAG 问答 + 自主任务智能体 + 多模型引擎 + 企业级权限治理” 的一体化设计，帮助企业以较低门槛将大模型能力转化为员工可直接使用的生产力工具，同时通过部门权限、敏感信息保护与私有化部署能力，兼顾效率与安全合规。

未来，问小天将持续在以下方向演进：

- 检索能力升级：随着知识库规模增长，平滑引入专用向量数据库，提升大规模场景下的检索性能；
- 智能体能力扩展：丰富 Cowork 可调用的工具生态，支持更复杂的多智能体协作与长周期任务；

- 多模态能力增强：拓展图像理解、语音交互等多模态场景的原生支持；
- 生态集成深化：持续扩充 MCP 工具生态与企业 IM/OA 系统的深度集成能力。

问小天希望成为每一家企业都能负担得起、信任得过、真正用起来的 AI 基础设施。