

灵枢 AI 智能体开发平台

产品白皮书



启创动力
QI CHUANG DONG LI

文档编号：

版本号：

V1.1

生效日期：

2026-04-07

编纂单位：

武汉启创动力数字科技有限公司

目录

灵枢 AI 智能体开发平台	1
1. 产品概述	3
2. 建设背景与需求	3
3. 平台架构	4
4. 适用对象	4
5. 产品能力介绍	5
5.1 知识与对话服务能力	5
5.2 流程编排与行业应用能力	6
5.3 协同智能体能力	8
5.4 部署与交付能力	9
5.5 账号权限管理能力	11
5.6 工作空间管理能力	12
6. 关键技术机制	12
7. 典型应用场景	13
8. 简易使用流程	13
9. 集成能力说明	14
10. 演进方向	14

1. 产品概述

AI 智能体平台是一套面向企业场景的智能应用构建与运行平台，提供模型接入、知识库管理、工作流编排、工具调用、对话服务、多智能体协同和部署交付等全链路能力。平台既支持快速构建单一智能助手，也支持面向复杂业务场景打造具备自主决策与协作能力的智能体系统。

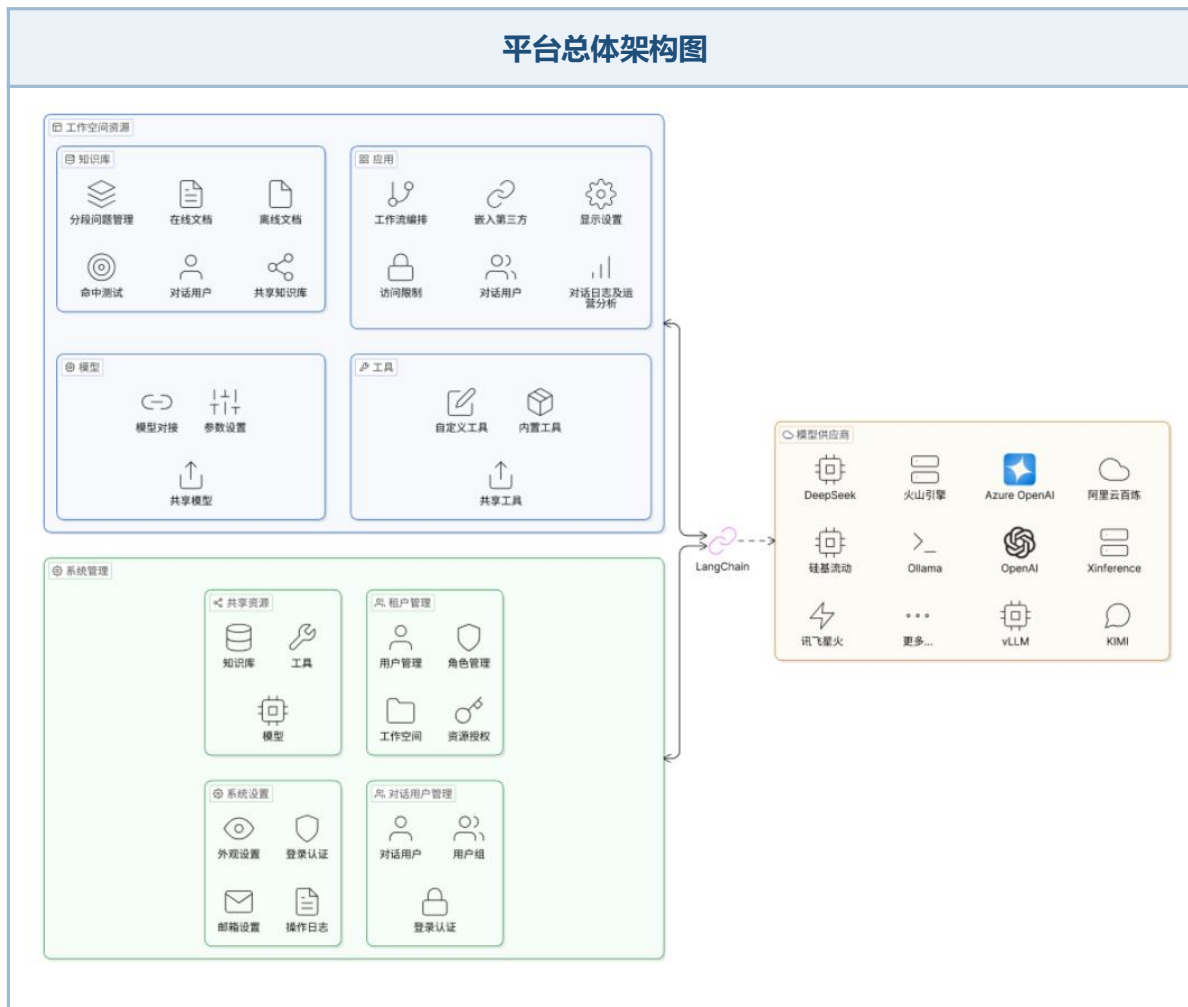
产品定位并非单纯的“问答系统”或“模型调用中台”，而是面向业务落地的智能体生产平台。其目标是在保证安全、可靠、可交付的前提下，将大模型能力沉淀为可复用、可编排、可集成、可溯源的企业级应用能力。

2. 建设背景与需求

在企业数字化阶段，大模型已从“能力展示”进入“业务落地”阶段。客户普遍不再满足于独立的聊天页面，而是更加关注知识接入效率、场景流程闭环、复杂任务协同、部署合规以及投入产出比。

传统知识问答方案难以承担流程执行任务，单模型单助手模式在复杂业务中缺乏角色分工，仅支持单一部署模式又难以满足不同客户的网络、安全和预算要求。因此，企业需要一套统一的平台，以承载从知识问答到流程执行、从单助手到多智能体协作、从在线使用到离线部署的完整路径。

3. 平台架构



4. 适用对象

- 适用行业：政务、制造、能源、医疗、金融、教育、零售等
- 适用组织类型：集团企业、中大型企业、政府单位、园区、机构客户等
- 适用角色：管理层、业务部门、信息化部门、运营部门、客服团队、知识管理团队等

5. 产品能力介绍

能力域	能力说明	典型输出	对应价值
对话与知识服务	基于 RAG 知识库快速构建聊天机器人，支持网页、企业门户、客服窗口等渠道嵌入。	智能问答助手、制度问答、产品咨询机器人	缩短上线周期，提升知识复用效率，降低人工答疑成本
流程编排与行业应用	通过 workflow 配置将模型、知识、工具和审批节点串联，形成面向垂直场景的标准化流程。	行业助手、表单流转、报告生成、业务协同流程	从单点问答升级为可执行业务能力
多智能体协同	多个具备自主决策能力的智能体在统一任务框架下分工、通信与协作。	研究分析、方案生成、任务拆解与闭环执行	处理复杂任务，提高结果完整性与执行深度
部署与交付	同时支持 SaaS 与私有化部署，并可根据客户网络、数据与安全要求进行弹性交付。	公有云租户、专属实例、内网私有化版本	适配不同规模客户，兼顾扩展性与安全合规

5.1 知识与对话服务能力

平台内置基于 RAG 的知识库能力，支持文档导入、切片处理、向量化、检索召回、重排与答案生成。业务团队可在较短时间内构建面向制度、产品、客服、运维和售前等场景的聊天机器人，并通过 Web 页面、嵌入式组件、API/SDK 等方式快速集成到现有系统。

该能力强调“快集成、可控答、可溯源、可持续迭代”。在建设过程中，可按知识源类型建立多知识库、多权限隔离和多应用复用机制，使知识资产从静态文档转化为可服务的智能能力。



5.2 流程编排与行业应用能力

平台提供可视化工作流配置能力，支持将知识检索、模型推理、规则判断、工具调用、人工审批和结果输出串联成标准流程。通过流程节点、条件分支、变量管理和异常处理机制，平台可将大模型能力沉淀为垂直行业中的可执行业务能力。

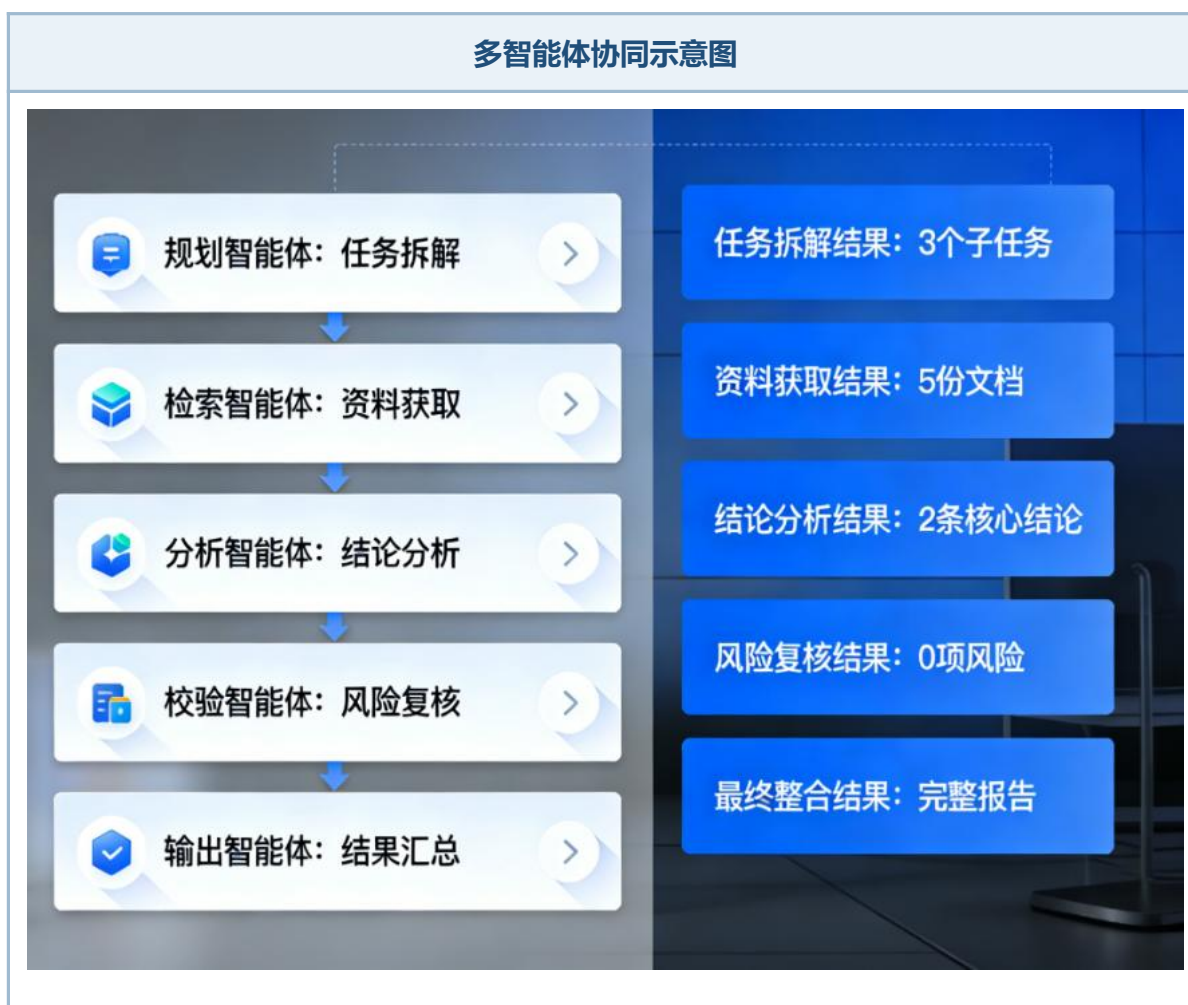
该能力适用于报告生成、表单预审、工单分派、客户意图识别、业务问答转办理等场景。对于行业项目，流程编排是把“会回答”升级成“会办事”的关键能力。



5.3 协同智能体能力

对于需要多步骤分析、角色分工或跨域知识协同的任务，平台支持构建多个具备自主决策能力的智能体，并通过通信、交互与协作机制共同完成复杂任务。例如，可设置规划智能体、检索智能体、分析智能体、校验智能体和输出智能体，分别承担任务拆解、资料获取、结论分析、风险复核和结果汇总等职责。

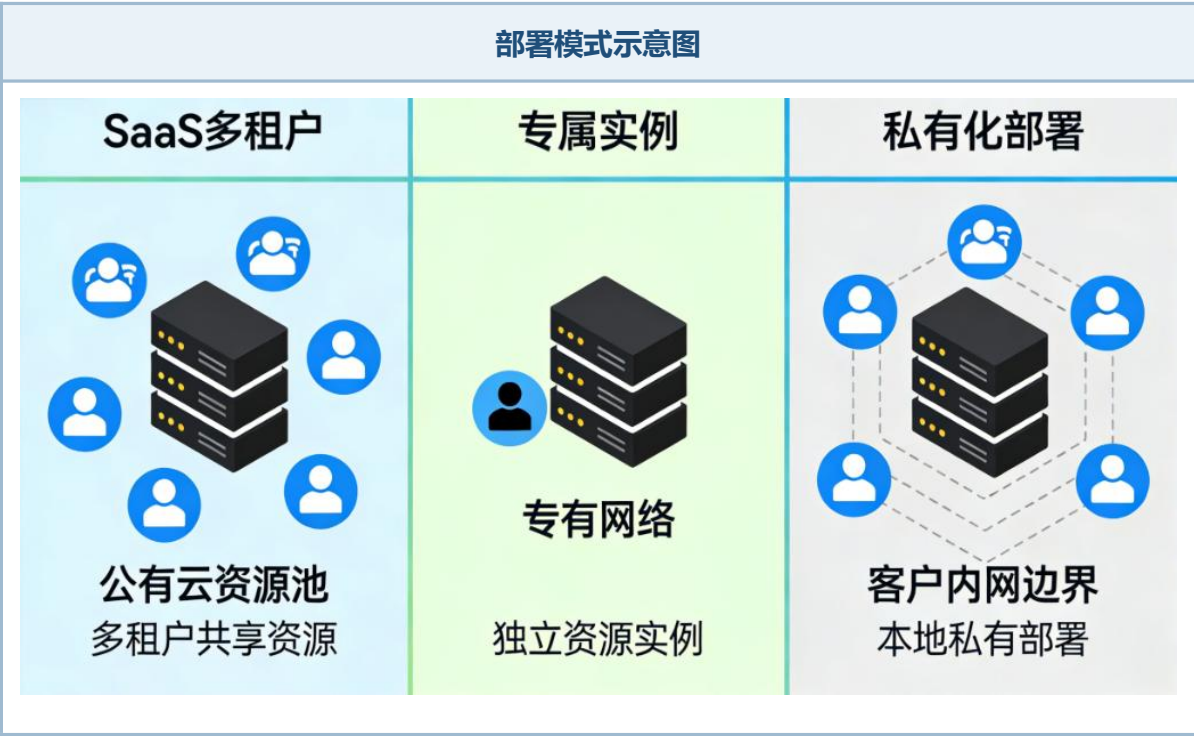
多智能体协同并非简单并发调用，而是在统一任务框架下形成有角色、有状态、有反馈的协同网络。平台支持任务分配、消息传递、上下文共享、阶段结果回传和最终整合，从而增强复杂任务处理深度。



5.4 部署与交付能力

平台支持 SaaS 订阅与本地私有化部署两种主要模式。对于希望快速试点、低门槛上线的客户，可采用 SaaS 方式进行租户级开通；对于存在内网隔离、数据不出域、模型专属接入或安全审计要求的客户，可采用私有化部署方式，将平台落于客户专属云环境或本地数据中心。

部署能力不仅体现在安装方式，更体现在交付适配性。平台可结合客户组织规模、预算水平、网络架构和安全要求，提供标准 SaaS、专属实例、混合部署、全私有化等不同策略。



私有化部署配置要求:

模型: Qwen3.6-27B

基础配置

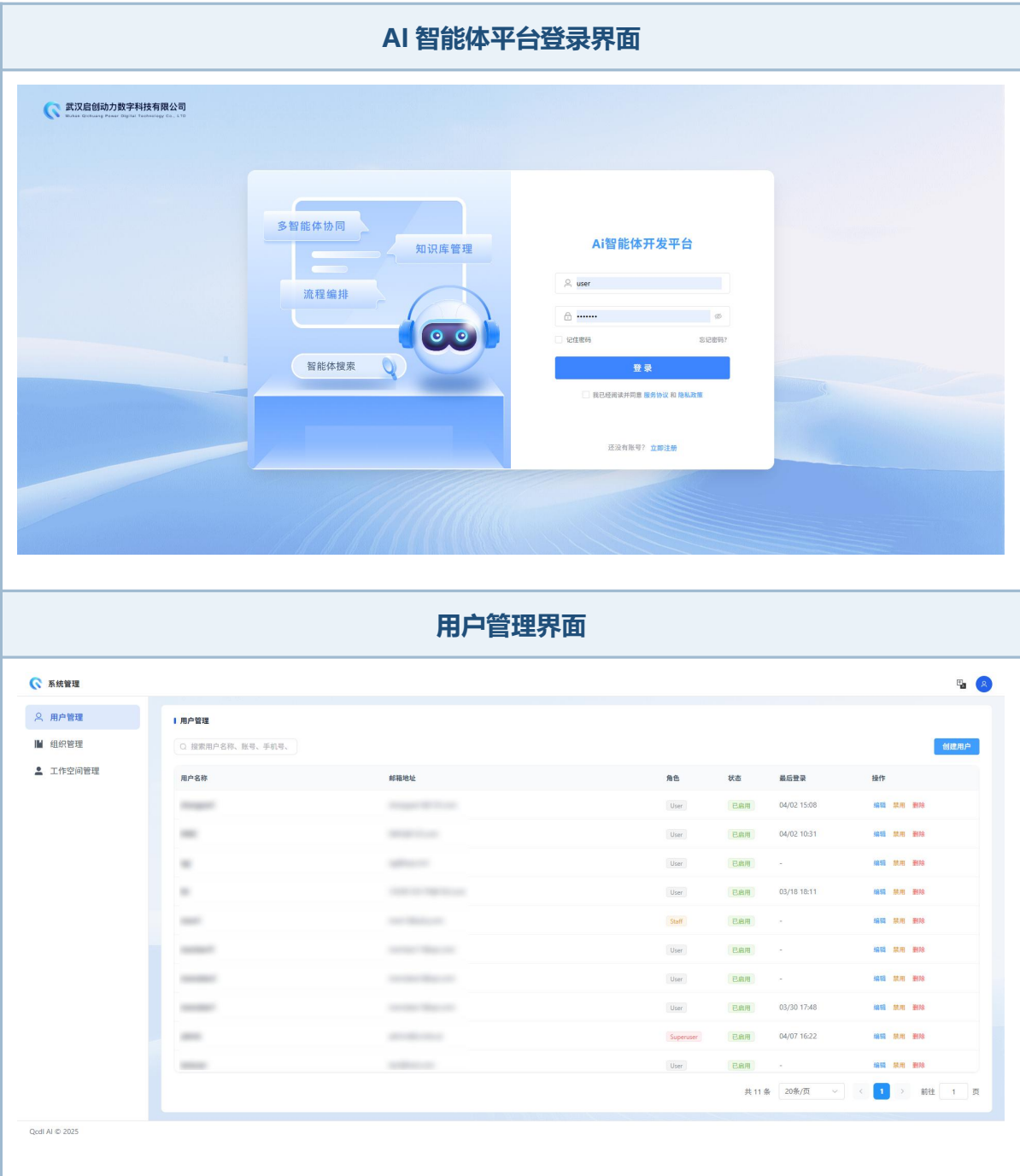
配置项	建议要求
GPU	NVIDIA 显卡两张, 显存 32GB
内存	DDR5 64GB
CPU	8 核处理器
存储	100GB NVMe SSD

推荐配置

配置项	建议要求
GPU	NVIDIA 显卡两张, 显存推荐 48GB 及以上
内存	DDR5 128GB 及以上
CPU	16 核及以上处理器
存储	1TB 及以上 NVMe SSD

5.5 账号权限管理能力

平台支持完善的账号与权限管理，可结合企业组织架构、岗位职责和业务分工，实现用户与角色多维度统一管理。系统支持功能权限、数据权限和工作空间的灵活配置，满足不同层级、不同岗位、不同场景下的精细化授权需求，满足企业级应用对安全性、规范性和可管可控的建设要求。



5.6 工作空间管理能力

平台提供统一的智能工作空间，支持围绕不同业务场景对智能体、知识库、工作流、工具能力和成员权限进行集中管理，帮助企业快速构建面向具体任务的专属应用环境。通过工作空间机制，用户可按部门、项目、业务条线或应用场景进行灵活划分，实现资源隔离、权限分级和独立配置，满足多团队并行建设与统一管理需求。



6. 关键技术机制

为保障平台在真实业务场景中的可用性与可扩展性，平台采用多层技术机制支撑智能体运行。

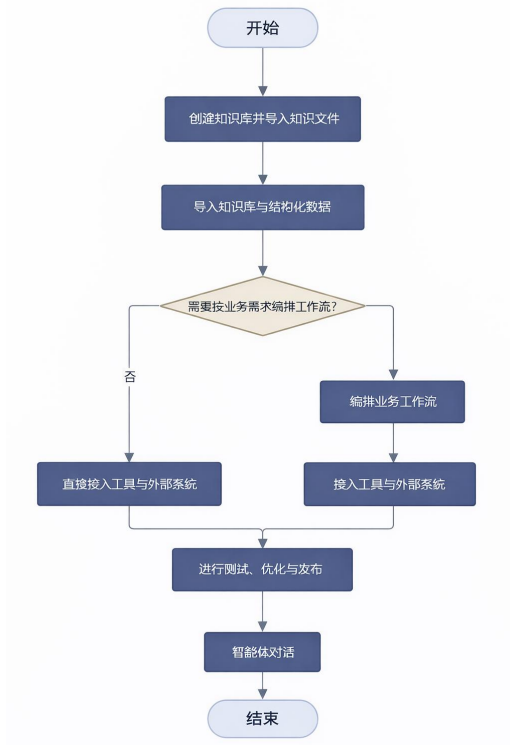
- 模型接入机制：支持多模型统一接入、版本切换、参数配置与调用监控。
- RAG 检索机制：支持多知识源接入、向量检索、关键词召回、重排融合和答案溯源。
- 工具调用机制：支持内部 API、第三方服务、数据库查询、搜索引擎和业务系统能力接入。
- 任务编排机制：支持条件分支、循环重试、节点状态记录、超时处理与人工介入。
- 多智能体通信机制：支持共享上下文、消息总线、任务回传、信息溯源和统一调度。

7. 典型应用场景

智能体	建设说明	适用场景
企业知识助手	围绕制度、项目、产品和 FAQ 构建统一问答入口，并嵌入门户或办公入口。	适用于企业制度查询、产品资料问答、项目经验复用、内部培训支持等场景。
行业办理助手	利用 workflow 将咨询、预审、填报、分派和回执串成闭环。	适用于合同审核辅助、报销问答、流程引导、信息采集、表单填报等场景。
客服与售前机器人	提供产品咨询、方案推荐、线索收集和标准答复。	适用于售前咨询、售后支持、工单分流、常见问题处理等场景。
复杂分析协同助手	通过多智能体分工完成资料收集、分析、复核和报告输出。	适用于能源运维、政务问答、医疗导诊、制造巡检、金融知识支持等行业化需求。

8. 简易使用流程

- 第一步：创建知识库，导入知识文件
- 第二步：导入知识库与结构化数据
- 第三步：配置智能体角色，接入模型
- 第四步：按业务需求编排 workflow
- 第五步：接入工具与外部系统
- 第六步：进行测试、优化与发布
- 第七步：智能体对话



9. 集成能力说明

平台支持多终端、多渠道接入，可根据客户现有数字化应用形态，将智能体能力快速嵌入门户网站、业务系统、移动应用、协同办公平台及对外服务窗口，实现统一能力输出与多场景触达。

平台支持多来源知识接入与统一治理，可对企业内部文档、结构化数据、FAQ 资料、业务数据等进行统一纳管，并通过切片、向量化、检索增强及权限控制机制，构建面向业务场景的高质量知识底座，为智能问答、流程协同及多智能体决策提供可靠支撑。

10. 演进方向

平台能够帮助客户以统一底座承载多类智能应用，减少重复建设；通过知识与流程融合，让智能体从展示能力走向业务能力；通过多智能体协同，提升复杂任务处理上限；通过 SaaS 与私有化双模式交付，满足不同客户的部署诉求。

未来，平台会继续沿着行业模板沉淀、智能体市场化复用、任务评估闭环、跨系统自治协同等方向演进，逐步形成企业级 AI 智能体生态。