辰极智脑 AI 平台项目立项书

项目编号: CJZN-AI-2025-001
立项日期: 2025 年 10 月 13 日
项目周期: 2025 年 10 月 - 2027 年 12 月 (27 个月)
项目负责人: 李泽宇
项目类型: 公司级 AI 战略核心项目
版本: V3.3
修订日期: 2025 年 10 月 19 日
修订说明: 结合 2025 年 10 月 17 日会议内容，重新梳理，简化内容，符合快速交付立项书的需求

目录

辰极智脑 AI 平台项目立项书
- 目录
- 第一章 项目背景与战略价值
  - 1.1 项目背景与行业痛点
  - 1.2 战略定位与核心目标
  - 1.3 商业价值与 ROI 预期
- 第二章 核心功能与产品规划
  - 2.1 近期规划（2025 年 10 月-12 月，2.5 个月）
    - 2.1.1 瑶光需求管理系统 AI 增强
      - 任务一:需求合规性校验功能测试
      - 任务二:对话系统开发
    - 2.1.2 MBSE 1.6 版本 AI 建模能力集成
    - 2.1.3 服务中台完善
  - 2.2 中期目标（2026 年 1 月-6 月,6 个月）:26 年 6 月 MVP
    - 2.2.1 为什么是 26 年 6 月?
    - 2.2.2 MVP 核心功能
      - 功能一:MBSE Copilot(AI 助手)
      - 功能二:模型增删改查
      - 功能三:基础设施与部署运维
  - 2.3 远期规划（2026 年 7 月-2027 年 12 月,18 个月）
    - 2.3.1 功能完善阶段(26 年 7 月-12 月)
      - 功能一:模型校验层
      - 功能二:CI/CD 自动化管线

功能三:仿真报错 AI 解决
功能四:针对仿真场景的专项优化
2.3.2 生态培育阶段(27 年 1 月-12 月)
战略一:打造 MBSE App Store
战略二:解决方案交付(模型即产品)
战略三:智能体社区与低代码构建
战略四:培育 MBSE 行业生态
第三章 技术架构与实现难点
3.1 总体技术架构
3.1.1 架构图
3.1.2 技术栈汇总
3.1.3 核心设计理念:Human-in-the-Loop(人在回路)
3.2 关键技术难点
3.2.1 难点一:SysML v1.6 的 XMI 格式对 AI 不友好
3.2.2 难点二:BS 架构全面云化的运维复杂性
3.2.3 难点三:100-200 人并发 AI 服务支撑
3.2.4 难点四:多智能体协作与记忆管理
3.2.5 难点五:保密行业的高准确率要求
3.3 知识库策略
3.3.1 知识库来源
3.3.2 知识库技术架构
第四章 实施计划与排期
4.1 25 年底前排期
4.1.1 第一阶段:需求合规性校验测试
4.1.2 第二阶段:对话系统开发
4.1.3 第三阶段:1.6 版本 AI 建模集成
4.1.4 第四阶段:服务中台完善(贯穿全程)
4.2 26 年上半年排期(2026 年 1 月 1 日-6 月 30 日,6 个月)
4.2.1 里程碑 4:知识库与 Agent 基础(1-3 月,3 个月)
4.2.2 里程碑 5:MBSE 2.0 与 AI 能力开发(1-6 月,6 个月)
4.2.3 里程碑:基础设施与部署运维(1-6 月,贯穿全程)
4.3 26 年下半年至 27 年排期(2026 年 7 月-2027 年 12 月,18 个月)
4.3.1 功能完善阶段(26 年 7 月-12 月,6 个月)
4.3.2 生态培育阶段(27 年 1 月-12 月,12 个月)
第五章 团队与资源配置
5.1 团队组成（3 人核心团队）
5.2 人力分配与时间安排
5.2.1 25 年(2025 年 10 月-12 月,3 个月)
5.2.2 26 年(2026 年 1 月-12 月,12 个月)

5.3 协作机制
5.3.1 内部协作
5.3.2 外部协作
第六章 财务预算与 ROI 分析
6.1 总预算（500 万元）
6.1.1 预算明细
6.1.2 GPU 服务器配置方案
6.2 收入模型（27 年底 5000 万元）
6.2.1 收入来源构成
6.2.2 软件订阅收入（1400 万，28%）
6.2.3 License 买断收入（600 万，12%）
6.2.4 项目交付营收（1750 万，35%）
6.2.5 AI 增值服务收入（650 万，13%）
6.2.6 生态平台抽成收入（600 万，12%）
6.3 ROI 分析
6.3.1 投资回报率计算
6.3.2 收入增长路径
6.3.3 成本效益分析
第七章 风险管理与应对
7.1 重大风险识别
7.1.1 风险一：MBSE 2.0 软件开发进度滞后（高优先级）
7.1.2 风险二：1.6 版本 AI 建模功能不好用（中优先级）
7.1.3 风险三：技术难度超预期（中优先级）
7.1.4 风险四：市场接受度不确定（中优先级）
7.1.5 风险五：外部 API 不稳定（低优先级）
7.2 质量保障（保密行业、军工行业准确率要求）
7.2.1 质量目标
7.2.2 质量保障措施
第八章 预期成果与价值
8.1 技术成果
8.2 产品成果
8.3 商业成果
8.4 战略价值
附录:术语表

第一章 项目背景与战略价值

1.1 项目背景与行业痛点

当前 MBSE(模型驱动系统工程)行业正处于关键转型期。传统建模工具面临三大核心痛点:**学习成本高、建模效率低、协作能力弱**。据统计,工程师掌握专业 MBSE 工具平均需要 6-12 个月,单个中等复杂度系统建模耗时 2-4 周,这严重制约了 MBSE 技术的推广与应用。

随着大语言模型技术的成熟,AI 与工程软件的深度融合已成为行业共识。Gartner 预测,到 2027 年,**70%的工程软件将内置 AI 能力**,AI 辅助建模将成为标配功能。然而,当前市场上的 AI 集成方案大多停留在浅层对话阶段,缺乏对工程领域的深度理解,无法真正解决实际工程问题。

公司现有的 MBSE 产品线(天玑/玉衡建模工具、瑶光需求管理系统)在行业内具有一定竞争力,但面临**技术架构老化**和**功能分散**两大挑战。现有软件基于 SysML v1.6 构建,采用 CS 架构,XMI 文件格式对 AI 不友好,难以实现深度 AI 集成;各产品线独立建设,缺乏统一管理,导致重复开发、资源浪费。

在此背景下,建设统一的**辰极智脑 AI 平台**,不仅是技术升级的必然选择,更是公司抢占 AI+MBSE 赛道制高点的战略举措。

1.2 战略定位与核心目标

辰极智脑 AI 平台是公司 AI 战略的**核心基础设施**,定位为"**公司级 AI 服务中台+三大智能化子系统**"。平台承载三大战略使命:

1. 技术领先:打造业内领先的 AI+MBSE 解决方案

平台将实现从"文本描述"到"可执行模型"的自动化生成,提供 MBSE Copilot 功能,降低建模门槛,提升建模效率 **300%以上**。同时支持 100-200 人协同能力(基于 Git 异步协作),满足大型项目需求。

2. 商业突破:从软件销售向解决方案交付转型

传统模式下,公司销售建模软件 License,客户需自行建模。创新模式将**直接直观展示客户需要的模型，如果客户需要直接购买**,实现从"卖工具"到"卖结果"的商业模式转型。同时构建 MBSE 生态平台,培育开发者社区,形成平台经济。

3. 生态构建:建立 MBSE 行业的"苹果 App Store"

平台开放 API 接口,允许第三方开发智能体和行业应用。建立智能体市场,开发者发布智能体,用户付费购买,平台抽成。沉淀历史项目模型,形成行业知识图谱,打造核心竞争壁垒。

1.3 商业价值与 ROI 预期

本项目总投入 **500 万元**,预期在 **27 年底实现 5000 万元营收**,投资回报率(ROI)达到 **10 倍**。商业价值体现在四个维度:

维度一:直接营收(5000 万元，至 27 年底)

收入来源	占比	金额	客户构成
软件订阅(企业版、大型企业版、教育版)	28%	1400 万	45-50 企业客户
License 买断(军工/国企内网部署)	12%	600 万	5-7 家大客户
项目交付营收(军工类甲方项目)	**35%**	**1750 万**	**军工/航空航天客户**
AI 增值服务(高级功能包)	13%	650 万	70-80 个企业客户
生态平台抽成(智能体+模型交易)	12%	600 万	生态开发者
总计	**100%**	**5000 万**	**~130-150 客户**

项目交付营收增长逻辑

项目交付营收(占比 **35%**,金额 **1750 万元**)是本项目的核心收入来源之一。其增长逻辑基于 AI 辅助建模能力对项目交付效率的显著提升。AI 模块使建模工程师效率提升 **10 倍**(从 2-4 周缩短至 2-3 天),项目交付周期从 6 个月缩短至 **2-3 个月**,客户体验显著改善,形成**"缩短交付周期→提升客户体验→促进项目复购→增强客户粘性→获取更多商业项目订单"**的正向循环。

在军工、航空航天等高价值行业,单个 MBSE 建模项目的实施收入可达 **50-200 万元**。随着客户对 AI 辅助建模能力的认可,项目复购率预计提升至 **60%以上**,形成稳定的项目交付营收来源。项目交付过程是展示平台能力的最佳场景,成功交付的项目将带动软件订阅和 License 买断销售(占比合计 **40%**,金额 **2000 万元**),形成"项目带动产品、产品反哺项目"的双轮驱动模式。

收入增长路径

25 年底试用期阶段,通过早期客户验证,预计实现 **50 万元**收入;26 年 6 月 MVP 发布后正式销售启动,预计累计实现 **500 万元**收入;26 年底完整功能交付,进入规模化销售阶段,预计累计实现 **1500 万元**收入;27 年底生态成熟,订阅续费与生态抽成双轮驱动,目标共计实现 **5000 万元**营收。

维度二:效率提升(内部降本增效)

建模效率从 2-4 周缩短至 **2-3 天**,效率提升 **10 倍**。单项目可减少**建模工程师投入**,大幅度节约人力成本。项目交付周期从 6 个月缩短至 **2-3 个月**,显著提升客户满意度。

维度三:市场竞争力

在 AI 辅助建模能力上,功能对标超越 IBM Rhapsody、Siemens Cameo 等国际巨头。订阅制定价 **20-50 万/年**,低于国际产品 **50-100 万/年**的订阅价格,形成明显价格优势。同时支持私有化部署、保密行业定制,满足军工/国企需求,提供本地化服务。

维度四:长期战略价值

历史项目模型转化为**知识资产沉淀**,形成行业知识图谱,打造不可复制的行业壁垒。开发者社区形成后,产生**生态平台效应**,平台价值指数级增长。用户使用越多,模型质量越高,吸引更多用户,形成**数据飞轮**正向循环。

第二章 核心功能与产品规划

2.1 近期规划（2025 年 10 月-12 月，2.5 个月）

当前时间节点为 **2025 年 10 月 19 日**,距离年底剩余不足 **2.5 个月**。这一阶段的核心目标是**完成现有功能的集成与测试**,为 26 年的大规模开发奠定基础。

2.1.1 瑶光需求管理系统 AI 增强

任务一:需求合规性校验功能测试

背景与现状

10 月 17 日已完成需求合规性校验功能的初步开发,但**未经过系统测试**。保密行业对准确率要求极高,一旦误报可能导致合规风险,需要大量测试和调优。

工作内容

本阶段将设计 100 个典型需求测试用例,覆盖合规、不合规、边界情况。测试准确率、召回率、误报率等核心指标,目标准确率**≥85%**。针对测试发现的问题进行优化,重点解决**军工保密行业的高准确率要求**。同时完成用户手册和操作指南,培训内部测试人员。

交付成果

经过充分测试的需求合规性校验功能,准确率达标,可用于实际项目。

任务二:对话系统开发

功能定位与技术背景

现有 MCP 接口已开发完成,知识库(RAG)已部署,联网搜索功能已集成,但**缺乏统一的对话系统**将三者整合。本功能将在需求管理系统右侧嵌入对话框(类似 Cursor、Copilot),支持用户通过自然语言与系统交互。

核心能力说明

系统具备四大核心能力。**知识检索**能力结合知识库(公司软件使用手册、方法论文档、垂类行业专业知识)。**项目上下文感知**能力通过 MCP 协议实时获取当前项目的需求数据,提供针对性建议。**联网搜索**能力对于知识库未覆盖的问题,自动联网搜索最新资料。**多智能体协作**能力包含多个专门的 Agent（初期计划）:如信息收集 Agent 负责从知识库、项目数据、互联网收集信息;内容整合 Agent 负责整合多源信息,生成连贯回答;质量校验 Agent 负责检查回答的准确性、完整性,特别是涉及保密行业的内容。此外,系统维护多轮对话的上下文,支持连续提问和追问。

技术难点分析

保密行业、军工行业对信息准确性要求极高,不能出现错误或误导性回答,这是准确性要求带来的首要挑战。多源信息融合需要平衡本地知识库、外部知识库、实时联网信息的优先级和可靠性。多智能体协作增加了处理时间,需要优化流程,确保首字输出响应时间。

交付成果

功能完整的对话系统,支持自然语言需求分析、项目问答、知识检索。

2.1.2 MBSE 1.6 版本 AI 建模能力集成

背景与现状问题

研究生团队已完成基于 LangGraph 的"文本到模型"生成工作流,该工作流能够将自然语言描述转化为 SysML 模型,并输出 XMI 格式文件，预计经过开发可以导入天玑/玉衡工具。然而,该功能存在三方面问题:一是**黑盒化**,用户输入描述后系统自动生成完整模型,中间过程不可见、不可干预;二是**不可编辑**,生成的模型无法在工作流中进行修改,必须导入到天玑/玉衡后才能编辑;三是**体验差**,实际工程中用户需要在建模过程中不断调整,而当前工作流不支持增量修改。

功能定位

将已完成的工作流集成到天玑/玉衡工具中,作为**展示能力的功能**,但不进行进一步开发。重点是**展示公司在 AI 建模领域的技术储备**,而非追求实际工程可用性。

工作内容

开发接口,将 LangGraph 工作流与天玑/玉衡工具集成。设计简单的用户界面,允许用户输入文本描述、触发生成、导入 XMI 文件。准备 3-5 个典型场景的演示

案例(如自行车系统、无人机系统),用于客户展示。编写功能说明文档,明确该功能的定位和局限性。

技术难点

SysML v1.6 的 XMI 格式对 AI 不友好,结构复杂,解析和生成难度大。需要确保生成的 XMI 文件能够被天玑/玉衡正确导入,不会出现格式错误或数据丢失。生成中等复杂度模型可能需要 5-10 分钟,需要提供进度提示,避免用户误以为系统卡顿。

时间安排

11 月进行接口开发与集成,12 月完成功能测试与演示准备。

交付成果

集成到天玑/玉衡的 AI 建模功能,可用于客户演示和技术展示。

2.1.3 服务中台完善

背景与现状

辰极智脑服务中台目前已支持双模式查询(检索型+对话型)、Provider 管理、SSO 认证,但在**工程化、稳定性、性能优化**方面仍有大量工作要做。

前端优化任务

对大量页面进行重构,提升用户体验。修复已知 Bug(详见 Bug 修复报告目录)。优化响应速度,减少页面加载时间。

后端优化任务

对 ConfigService 进行重构,提升配置管理的灵活性。对编排器进行重构,支持更复杂的 AI 工作流，对注册中心进行二期开发。进行性能优化,确保并发支撑能力达标,目标是支持 **100 并发用户**。

核心技术任务详解

服务注册与发现机制方面,MVP 已完成基础的服务注册功能,支持 AI 子服务的动态注册与注销,但需要深化健康检查机制、服务版本管理、灰度发布支持,目标是支持 100+ AI 子服务的稳定注册与自动发现。

转发机制与依赖队列管理方面,需要实现智能请求转发,根据服务负载、响应时间动态选择最优服务实例。构建依赖队列管理系统,处理 AI 子服务间的复杂调用关系。支持请求优先级队列、超时重试、熔断降级等高可用机制。

子服务间依赖复杂性与保活机制是核心技术难点。AI 工作流中存在大量 A→B→C 的依赖链(如"需求分析 Agent → 模型生成 Agent → 合规校验 Agent"),单个服务

故障可能导致整个工作流失败。100+ AI 子服务需要长期保活,但部分服务(如大模型 API 调用)存在超时、限流、间歇性故障等问题,如何确保整体系统稳定性是工程挑战。解决方案是设计分层保活策略(关键服务常驻内存,次要服务按需加载)、依赖拓扑分析、故障隔离与自动恢复机制。

长期持续开发特性需要特别说明。服务中台**不是一次性交付的模块**,而是需要随 AI 子服务数量增长、业务复杂度提升而持续演进的核心基础设施。预计在整个项目周期(2025-2027 年)内持续开发,每个季度迭代优化。工程复杂度高、依赖关系深,是平台稳定性的基石。

运维部署任务

完善 Docker 部署流程,编写详细的部署文档。建立监控和日志系统,确保问题可快速定位。进行压力测试,验证系统在高并发场景下的稳定性。

技术难点

当前系统在 10-20 并发时表现稳定,但 100 并发时可能出现性能瓶颈,需要优化数据库连接池、缓存策略、负载均衡等。AI 服务依赖外部 API(如 DeepSeek、Kimi),当外部 API 不稳定时,需要有降级策略和重试机制。

交付成果

稳定、高性能的 AI 服务中台,支持 1000 并发用户,可用性≥99%（长期目标,非 25 年实现）。

2.2 中期目标（2026 年 1 月-6 月,6 个月）:26 年 6 月 MVP

26 年 6 月是**关键里程碑**,该时间节点的核心目标是交付 **MBSE Copilot MVP 版本**,实现"可用、可演示、可销售"的 AI 辅助建模能力。

2.2.1 为什么是 26 年 6 月?

背景分析

公司计划在 25 年底或 26 年初启动 **MBSE 2.0 软件的开发**(基于 SysML v2.0,BS 架构)。然而,大型软件开发周期通常为 12-18 个月,这意味着 2.0 软件可能在 26 年底甚至 27 年初才能完成。如果 AI 功能的开发完全依赖 2.0 软件的完成,将导致 AI 能力推迟到 27 年,**错失市场窗口期**。

战略选择

将 AI 能力开发与 2.0 软件开发**并行推进**。在 2.0 软件的开发过程中,AI 团队提前完成核心 AI 功能的开发,待 2.0 软件完成后,快速集成 AI 能力。26 年 6 月作为

中间检查点,要求 AI 团队交付可独立演示的 MVP 版本,即使 2.0 软件尚未完成,也能通过模拟数据展示 AI 能力。

风险说明

如果 2.0 软件开发进度严重滞后(如推迟到 27 年 Q2),将影响 AI 功能的最终集成。为此,需要在风险管理章节明确应对策略(见第七章)。

2.2.2 MVP 核心功能

功能一:MBSE Copilot(AI 助手)

功能定位

MBSE Copilot 定位类似 GitHub Copilot,为建模工程师提供实时 AI 辅助。系统具备三大核心能力。

知识问答能力

用户可以询问"如何建立状态机?""SysML 中的 Block 和 Class 有什么区别?"等问题,系统结合本体论知识、方法论文档、案例库给出详细解答。这一能力基于四层知识库支撑体系。

代码/模型补全能力

用户在编辑 SysML 文本(SysML v2.0 采用文本化建模)时,系统自动提示补全内容,大幅提升建模效率。

智能建议能力

系统分析当前模型结构,提示可能遗漏的元素、不合理的关系、潜在的优化点,帮助工程师提升模型质量。

知识库支撑体系

这些能力由四层知识库支撑。**本体论知识**层包含 SysML 规范、MBSE 理论基础。**方法论知识**层包含公司内部的建模方法论、最佳实践。**案例库**层包含历史项目的典型模型,作为参考模板。**外部知识库**层包含合作伙伴提供的 MBSE 行业知识(如 OMG 官方文档、学术论文)。

功能二:模型增删改查

功能背景

传统建模工具采用图形化界面,操作繁琐。SysML v2.0 引入文本化建模,为 AI 辅助提供了可能。

自然语言建模

用户输入"创建一个名为 MotorController 的 Block,包含两个端口:输入端口 powerIn 和输出端口 motorOut",系统自动生成对应的 SysML 代码。这一能力大幅降低建模门槛。

模型修改与删除

用户输入"将 MotorController 的 powerIn 端口改为双向端口",系统自动修改模型。用户输入"删除 MotorController 中的 motorOut 端口",系统自动删除。

模型查询

用户输入"列出所有继承自 Controller 的 Block",系统返回查询结果。这一能力帮助工程师快速理解模型结构。

功能三:基础设施与部署运维

BS 架构全面云化

前端基于 Vue 3 的 Web 应用,支持浏览器访问。后端基于 Spring Boot 的微服务架构,支持水平扩展。数据库采用 PostgreSQL(关系型数据)+ Neo4j(知识图谱)+ Qdrant(向量数据库)的混合存储方案。AI 服务通过统一的 AI 服务中台调用各子服务。

100-200 人协同能力(基于 Git 异步协作) 

协作模式说明。 系统支持**单人编辑模式**,即同一时刻一个用户编辑一个模型项目,避免实时协同冲突。**100-200 人协同能力**指的是通过 **Git 版本管理机制**实现的**异步协作**,而非传统意义上的"多人同时在线编辑"。这种设计继承了传统 MBSE 工具(如 Rhapsody、Cameo)的文件锁机制,确保数据一致性。

技术实现方案。 采用**基于 Git 的分支-合并工作流**,每个工程师在独立分支上编辑模型,完成后通过 Git 合并到主分支。借助 SysML 2.0 的 **KerML 文本化格式**,模型可像代码一样进行版本控制、差异比对、冲突解决。开发智能合并工具,辅助解决模型合并冲突,确保引用关系和语义完整性(详见 3.2.1 节)。

并发支撑能力说明。 支持 100-200 人**同时在线使用系统**(查看模型、查询知识库、使用 AI 服务),而非同时编辑同一模型。通过负载均衡、缓存、消息队列等技术,确保系统在高并发场景下的稳定性。

AI 服务如何支撑高并发。 当并发请求超过 AI 服务处理能力时,将请求放入**请求队列**,依次处理。部署**多个 AI 服务实例**,通过负载均衡分发请求。对常见问题(如"什么是 Block?")的回答进行**缓存**,减少重复调用大模型 API。当外部 API 不可用时,执行**降级策略**,返回预设的备用回答或提示用户稍后重试。

技术挑战与复杂性强调

Git 版本管理与模型合并是软件工程领域的经典难题。MBSE 模型的 Git 异步协作需要克服模型结构复杂、依赖关系多、冲突检测困难等挑战,需要开发智能合并工具确保引用关系和语义完整性。

云端部署与运维方面,传统 CS 架构软件部署简单,但 BS 架构需要考虑**高可用、容灾、备份、监控、日志、安全**等多方面问题。支持 100-200 人在线意味着需要配置负载均衡、数据库主从复制、Redis 缓存集群等复杂基础设施。

AI 服务稳定性方面,AI 服务依赖外部大模型 API,而这些 API 可能出现限流、超时、返回错误等情况。如何保证 AI 服务在外部 API 不稳定时仍能正常工作,需要设计完善的容错机制。

交付成果

可独立演示的 MBSE Copilot MVP 版本,支持基本的知识问答、模型增删改查、Git 版本管理与异步协作。

2.3 远期规划（2026 年 7 月-2027 年 12 月,18 个月）

远期规划分为两个阶段:**功能完善阶段(26 年 7 月-12 月)**和**生态培育阶段(27 年 1 月-12 月)**。

2.3.1 功能完善阶段(26 年 7 月-12 月)

本阶段的目标是在 MVP 基础上,补充高级功能,形成完整的 AI+MBSE 解决方案。

功能一:模型校验层

功能背景

模型建立后,需要校验其正确性,包括**语法正确性**(是否符合 SysML 规范)、**语义正确性**(模型是否自洽,是否存在逻辑矛盾)、**工程正确性**(是否符合行业标准和最佳实践)。

AI 能力说明

系统自动检测模型中的语法错误、语义矛盾、潜在风险。提供修复建议,甚至自动修复简单错误。学习历史项目的校验规则,形成公司内部的校验标准库。

功能二:CI/CD 自动化管线

功能背景

传统建模流程是手动建模、手动保存、手动提交。当模型规模增大、团队协作增多时,这种方式效率低、容易出错。

核心能力

自动提交功能允许用户完成一段建模工作后,系统自动将修改提交到版本库。**自动保存**功能定期自动保存,避免数据丢失。**AI 建议**功能由系统分析提交历史,提示用户何时提交、如何编写提交说明。**版本管理**功能类似 Git,支持分支、合并、回滚等操作,AI 辅助解决合并冲突。

功能三:仿真报错 AI 解决

功能背景

MBSE 模型建立后,通常需要进行仿真验证(如使用 Simulink、Modelica)。仿真过程中可能出现报错,工程师需要根据报错信息定位问题、修改模型、重新仿真,这个过程耗时且需要专业知识。

报错分析能力

用户将仿真报错信息输入系统,AI 分析报错原因,定位到具体的模型元素。

修复建议能力

AI 提供修复建议,如"端口类型不匹配,建议将 powerIn 改为 Integer 类型"。

自动修复能力(远期)

对于简单的报错(如端口类型错误、参数缺失),AI 自动修复并重新触发仿真。

重要限制说明

不做全自动化闭环。即不做"AI 自动运行仿真 → 发现报错 → 自动修改模型 → 重新仿真 → 再次报错 → 再次修改"这种完全无人干预的闭环。原因包括三方面:一是技术难度极高,需要 AI 理解仿真工具的报错信息、理解模型的工程语义、理解仿真工具的执行逻辑,当前技术水平难以达到;二是风险高,自动修改模型可能引入新的问题,在保密行业、军工行业,这种风险不可接受;三是用户信任问题,工程师对完全黑盒化的自动修复缺乏信任,仍然希望人工确认。

功能四:针对仿真场景的专项优化

优化背景

仿真是 MBSE 的核心应用场景,但当前 AI 在仿真场景下的表现不够理想。

仿真领域知识库建设

收集仿真工具的文档、常见报错及解决方案,构建专项知识库。

多智能体协作优化

部署三个专门 Agent:**仿真分析 Agent** 专门分析仿真报错,定位问题;**模型修复 Agent** 提供修复建议;**验证 Agent** 验证修复后的模型是否符合预期。

记忆管理与上下文工程

维护仿真任务的上下文,支持多轮对话调试。优化 Prompt 设计,提升 AI 在仿真场景下的准确率。

2.3.2 生态培育阶段(27 年 1 月-12 月)

本阶段的目标是从"软件销售"向"平台运营"转型,构建 MBSE 行业的生态平台。

战略一:打造 MBSE App Store

战略背景

不同行业、不同企业对 MBSE 工具的需求差异巨大。汽车行业需要符合 AUTOSAR 标准的建模工具,航空航天行业需要符合 DO-178C 标准的工具,传统的"通用软件"模式难以满足所有需求。

平台化实施方案

将 MBSE 2.0 软件的核心能力(模型管理、版本控制、仿真接口等)通过 **API 开放**给第三方开发者。开发者可以开发行业专属的智能体(如"汽车 AUTOSAR 建模助手"、"航空 DO-178C 合规检查助手"),发布到**智能体市场**。用户根据需要购买或订阅智能体,平台抽成。对活跃开发者提供奖励,举办开发者大赛,吸引更多开发者加入,形成**生态激励**。

战略二:解决方案交付(模型即产品)

创新模式背景

传统模式下,客户购买建模软件后,需要自己招聘工程师、学习软件、建立模型,周期长、成本高。很多中小企业因此放弃使用 MBSE。

模型交付流程

客户输入**需求**,如"我需要一个电动汽车的动力系统模型"。系统从历史项目的模型库中检索相似模型,结合 AI 生成定制化模型,完成**模型生成**。客户支付费用,获得可直接使用的模型,模型可导入到 CAD、CAE、EDA 等工具中,实现**模型交付**。

商业模式创新

客户为模型付费,而非为软件付费。类似 Canva 的设计模板、Envato 的代码素材。

技术支撑体系

将历史项目的模型转化为**知识图谱**,建立模型之间的关联关系。基于知识图谱的检索增强生成(**Graph RAG**),快速定位相似模型。将多个模型模块组装成完整模型(**模型组装**),类似乐高积木。

战略三:智能体社区与低代码构建

战略背景

企业内部可能有特定的建模流程、特定的规范,需要定制化的智能体,但不是每个企业都有能力开发智能体。

低代码平台实施

用户通过拖拽的方式,进行**可视化编排**智能体的工作流(如"先检索知识库 → 再调用外部 API → 最后生成报告")。提供常见智能体**模板库**,用户基于模板修改,快速创建定制化智能体。用户可以将自己创建的智能体发布到**社区**分享,其他用户可以使用、评价、二次开发。

战略四:培育 MBSE 行业生态

订阅制商业模式

教育版(5-10 万/年)提供基础功能,面向高校和培训机构,培养 MBSE 人才。**试用版**免费试用 30 天,功能受限(如不能导出模型),吸引潜在客户。**小微企业版**(1-5 万/年)支持 10 人以下团队,提供基础 AI 功能。**企业版**(20-50 万/年)支持 50-200 人团队,提供高级 AI 功能。**大型企业版**(80-150 万/年)支持 200+人团队,提供定制化服务。**买断版(军工/国企)**(200-500 万)支持私有化部署、内网隔离,满足保密要求。

预期成果

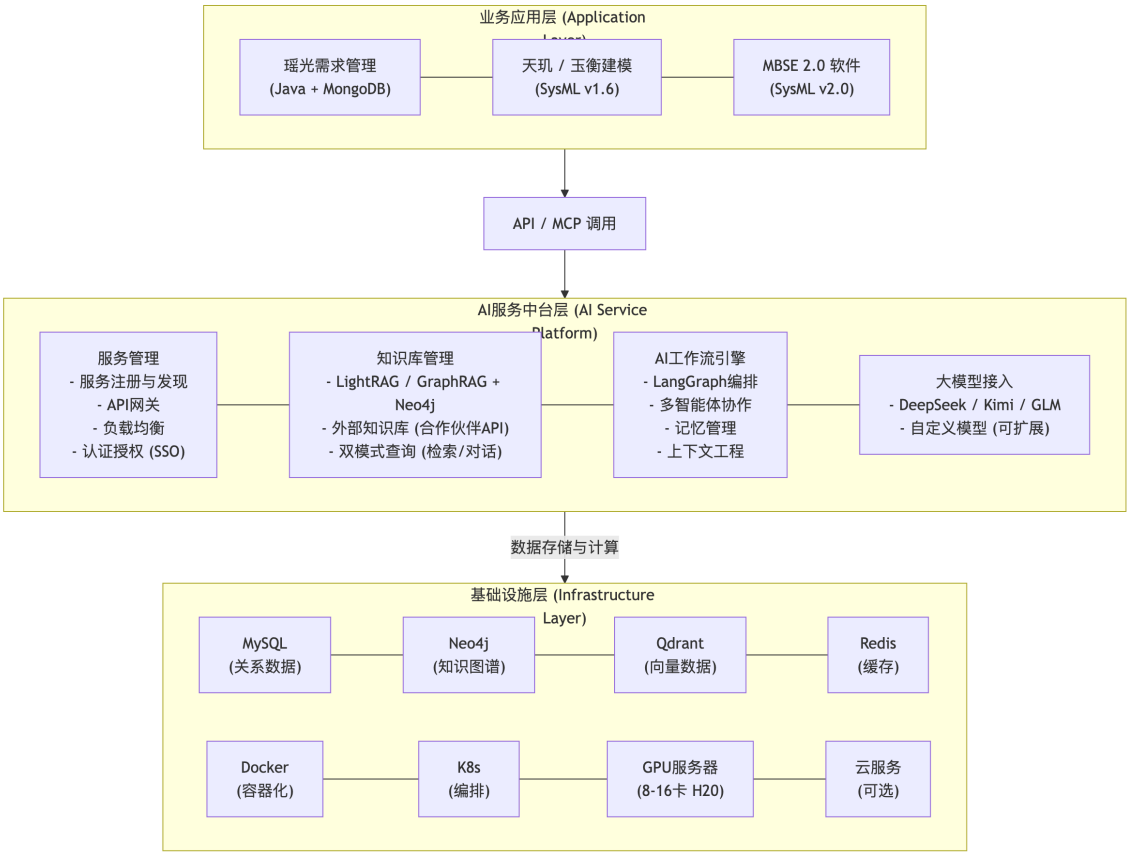
形成 MBSE 行业的平台生态,开发者、企业、用户多方共赢。

第三章 技术架构与实现难点

3.1 总体技术架构

辰极智脑 AI 平台采用**微服务架构**,分为三层:**AI 服务中台层**、**业务应用层**、**基础设施层**。

3.1.1 架构图



3.1.2 技术栈汇总

层级	组件	技术栈	说明
业务应用层	瑶光需求管理系统	Java 17 + Spring Boot 3 + MongoDB + Vue 3	现有系统,提供 MCP 接口
	天玑/玉衡建模工具	SpringBoot 2 + Vue (SysML v1.6)	现有系统,集成 AI 能力
	MBSE 2.0 软件	Spring Boot 3 + Vue 3 (SysML v2.0)	新开发系统,AI 原生集成
AI 服务中台层	服务管理	FastAPI + Nginx + Redis + Vue3 + React	API 网关、负载均衡、缓存
	知识库管理	LightRAG + Neo4j + Qdrant	混合检索(向量+图谱)
	AI 工作流引擎	LangGraph + LangChain + etc.	多智能体协作编排
	大模型接入	DeepSeek / Kimi / GLM API	国产大模型,可扩展

基础设施层	数据库	MySQL + Neo4j + Qdrant + Redis	多数据库协同
	部署	Docker + K8s + GPU 服务器	容器化部署,支持扩展

3.1.3 核心设计理念:Human-in-the-Loop(人在回路)

本系统设计的核心理念之一是**人在回路(Human-in-the-Loop)机制**,确保在 AI 自动化生成、建模、推理等关键环节保留人工确认,提高系统可控性与可靠性。

设计原则

AI 生成 + 人工审核。AI 自动生成模型元素后,需由工程师审核确认后方可正式写入模型库。系统提供"预览"功能,工程师可查看 AI 生成的模型结构、属性、关系,判断是否符合预期。同时支持"部分接受"操作,工程师可选择接受部分生成结果,拒绝或修改其他部分。

关键决策人工介入。在涉及系统架构调整、关键需求变更等决策点,AI 提供建议,最终决策权在人。例如,AI 建议"将控制器分解为三个子系统",工程师可选择接受、修改或拒绝。系统记录决策历史,便于后续追溯和审计。

可解释性增强。AI 提供的每个建议都附带推理依据,便于工程师理解和判断。例如,AI 建议添加某个接口时,说明"根据需求文档第 3.2 节和类似系统参考模型,该接口是必需的"。引用溯源功能使工程师可点击查看 AI 引用的原始文献、历史案例或知识库条目。

渐进式自动化。初期以 AI 辅助为主,随着用户信任度提升,逐步扩大自动化范围。系统提供"自动化级别"设置,用户可根据项目风险等级调整 AI 的自主权限:保守模式下 AI 仅提供建议,所有操作需人工确认;标准模式下 AI 可自动执行低风险操作(如补全属性、格式化),高风险操作需确认;激进模式下 AI 拥有更大自主权,仅在检测到异常时请求人工介入。

应用场景

MBSE Copilot 知识问答方面,AI 回答时标注信息来源和置信度,工程师可验证答案正确性。**模型自动生成**方面,AI 生成模型草稿后,工程师审核、修改、确认后才写入正式模型库。**模型校验**方面,AI 检测到模型错误时,提供修复建议,由工程师决定是否采纳。**需求追溯**方面,AI 自动建立需求与模型元素的追溯关系,工程师审核确认后生效。

行业适配性

这一理念尤其适用于**军工、航空航天等高可靠性行业**,确保 AI 技术在提升效率的同时,不牺牲工程质量和安全性。通过人机协作,既发挥 AI 的高效性,又保留人类工程师的专业判断和经验积累。

3.2 关键技术难点

3.2.1 难点一:SysML v1.6 的 XMI 格式对 AI 不友好

问题描述

SysML v1.6 采用 **XMI(XML Metadata Interchange)**格式存储模型,XMI 是基于 XML 的复杂结构化格式。XMI 文件冗余度高、嵌套层次深、可读性差,即使对人类工程师也难以直接阅读和编辑。AI 大模型在处理 XMI 时,面临以下挑战:首先是**Token 消耗大**,一个中等规模的模型,XMI 文件可能有数千行,消耗大量 Token;其次是**结构理解困难**,AI 需要理解 XMI 的 Schema、元素的嵌套关系、引用关系,而这些信息在 XMI 中是隐式的;第三是**生成错误率高**,AI 生成 XMI 时,容易出现格式错误、引用错误、Schema 不符合规范等问题。

解决方案

方案一:XMI 简化与预处理。对 XMI 进行简化,去除冗余信息,保留核心结构。将 XMI 转换为更友好的中间格式(如 JSON、YAML),供 AI 处理。处理完成后,再将中间格式转换回 XMI。

方案二:文本化建模(面向 SysML v2.0)。SysML v2.0 引入 **KerML(Kernel Modeling Language)**文本化建模语言,模型以纯文本形式表示,类似编程语言。文本化建模对 AI 更友好,AI 可以像处理代码一样处理模型。在 2.0 时代,逐步淘汰 XMI,全面转向文本化建模。

协作模式与版本管理机制

单人编辑模式说明。本系统采用 B/S 架构下的**单人编辑模式**,即同一时刻仅允许一个用户编辑某个模型项目,避免实时协同编辑带来的冲突问题。**不支持多人同时在线编辑**:传统 MBSE 工具(如 IBM Rhapsody、Cameo)大多基于文件锁机制,本系统继承此设计理念,**不支持多人同时在线编辑同一个模型项目**。文档第 2.2.2 节提到的"100-200 人在线协同编辑"指的是**异步协作能力**(通过 Git 版本管理),而非实时协同编辑。

异步协作通过 Git 管线实现。虽然不支持实时协同编辑,但通过引入**模型合并机制(基于 Git 版本管理)**,实现多人异步协作:每个工程师在独立分支上编辑模型;完成后通过 Git 提交、合并(Merge)操作整合到主分支;借助 SysML 2.0 的 KerML 文本化表示,模型文件可像代码一样进行版本管理和冲突解决。

基于 SysML 2.0 的文本化版本管理

SysML 2.0 的优势。传统 SysML v1.6 采用 XMI(XML 格式)存储模型,对版本管理不友好。**SysML 2.0 引入 KerML(Kernel Modeling Language)文本化建模语言**,模型以纯文本形式表示,类似编程语言源代码。文本化格式天然适合 Git 版本控制,支持 Diff、Merge、Blame 等标准 Git 操作。

数据结构重新设计。为支持 Git 管线的模型合并机制,我们需要**重新设计数据存储结构**:从二进制 XMI 格式迁移到 KerML 文本格式;设计模型元素的差异比对(Diff)算法,识别模型变更(新增、删除、修改);开发智能合并工具,辅助解决模型冲突(类似 Git 的 Merge Conflict Resolution);建立模型元素 ID 稳定性机制,确保合并过程中引用关系不被破坏。

系统级适配。SysML 2.0 标准已针对文本化和版本化做了优化(如命名空间、元素 ID 稳定性),我们基于此进行系统级适配,将模型纳入标准 Git 工作流。开发 VS Code/Eclipse 插件,提供模型冲突的可视化解决界面。集成 CI/CD 流程,自动检测模型合并后的一致性和完整性。

时间规划。25 年底支持 1.6 版本的 AI 集成;26 年开发基于 KerML 的文本化建模能力,为 2.0 版本做准备;26 年 Q3-Q4 设计并实现基于 Git 的模型版本管理系统,支持多人异步协作。

3.2.2 难点二:BS 架构全面云化的运维复杂性

问题描述

传统 CS 架构软件(如天玑/玉衡 1.6)部署简单,用户安装到本地电脑即可使用,无需复杂的服务器配置。BS 架构软件需要部署到云端服务器,涉及**负载均衡、数据库主从复制、缓存集群、日志监控、容灾备份**等复杂运维工作。支持 100-200 人在线协同编辑,意味着需要应对**高并发、高可用、高性能**的挑战。

技术挑战

实时同步问题:用户 A 修改模型,用户 B 的界面需要实时更新,如何实现?(WebSocket? Server-Sent Events?) **数据库性能**问题:Neo4j 图数据库在高并发写入时性能下降,如何优化?**AI 服务瓶颈**问题:100 个用户同时请求 AI 服务,如何保证响应时间?如何避免外部 API 限流?

解决方案

架构设计方面。前端采用 Vue 3 + WebSocket,实现实时编辑。后端采用 Spring Boot 微服务,支持水平扩展。数据库层面,关系型数据(用户、权限)采用 MySQL,主从复制,读写分离;知识图谱采用 Neo4j,采用分片策略,将大图拆分为多个子图;向量数据(知识库)采用 Qdrant,支持分布式部署;缓存采用 Redis 集群,缓存热点数据。负载均衡采用 Nginx + K8s,自动扩缩容。AI 服务部署多个 AI 服务实例,通过

消息队列(RabbitMQ / Kafka)分发请求;引入**请求优先级**,VIP 用户的请求优先处理,普通用户排队;引入**降级策略**,当外部 API 不可用时,返回缓存结果或提示用户稍后重试。

运维保障方面。监控采用 Prometheus + Grafana,实时监控服务器 CPU、内存、磁盘、网络。日志采用 ELK(Elasticsearch + Logstash + Kibana),集中管理日志。告警机制:当服务异常(如响应时间>5 秒、错误率>5%)时,自动发送告警。容灾机制:数据库定期备份,部署在多个可用区,当一个可用区故障时,自动切换到备用可用区。

投入估算

人力方面需要 1 名专职运维工程师(或外包),负责部署、监控、故障排查。硬件方面除了 GPU 服务器(200-300 万),还需要配置负载均衡器、Redis 集群、数据库服务器,预计额外投入 50-100 万。云服务(可选)如果采用阿里云、腾讯云等公有云,年度费用 10-20 万。

3.2.3 难点三:100-200 人并发 AI 服务支撑

问题描述

AI 服务依赖外部大模型 API(如 DeepSeek、Kimi),每个 API 调用耗时 1-3 秒。当 100 个用户同时发起 AI 请求时,需要调用 100 次 API,如果串行处理,总耗时 100-300 秒,显然不可接受。外部 API 通常有**并发限制**(如每秒最多 10 次调用)和**流量限制**(如每天最多 10 万次调用),超过限制会被限流或封禁。

解决方案

请求队列。所有 AI 请求先进入消息队列(RabbitMQ / Kafka),再由多个 Worker 并行消费。根据外部 API 的并发限制,动态调整 Worker 数量(如 API 限制每秒 10 次调用,则部署 10 个 Worker)。

请求合并。如果多个用户在短时间内提出相同或相似的问题,合并为一次 API 调用,返回相同结果。例如,用户 A 和用户 B 都问"什么是 Block?",系统只调用一次 API,结果同时返回给 A 和 B。

缓存机制。对常见问题的回答进行缓存(Redis),下次再问时直接返回缓存结果。缓存有效期设置为 1 小时,避免缓存过期后仍然返回旧答案。

降级策略。优先级队列:VIP 用户的请求优先处理,普通用户排队。限流:当并发请求超过系统处理能力时,返回"系统繁忙,请稍后重试"。备用模型:当主模型(如 DeepSeek)不可用时,自动切换到备用模型(如 Kimi)。

成本控制

外部 API 按调用次数收费,假设平均每次调用成本 0.01 元,每天 10 万次调用,年度成本约 365 万元。通过缓存、请求合并等优化,预计可将 API 调用次数降低 50%,年度成本降至 180 万元。

3.2.4 难点四:多智能体协作与记忆管理

问题描述

单一大模型难以处理复杂任务,需要引入**多智能体协作**。例如,对话系统包含:信息收集 Agent、内容整合 Agent、质量校验 Agent,三者需要协作完成任务。多智能体协作涉及:**任务分解、任务分配、结果聚合、冲突解决**,技术难度高。同时,需要维护**对话记忆**(多轮对话的上下文)和**任务记忆**(历史任务的经验),避免重复工作。

解决方案

LangGraph 工作流编排。使用 LangGraph 定义智能体之间的工作流(如"信息收集 → 内容整合 → 质量校验")。每个智能体是一个节点,节点之间通过边连接,边上可以定义条件(如"如果校验失败,返回内容整合节点重新生成")。

记忆管理。短期记忆:当前对话的上下文,存储在 Redis 中,对话结束后清除。长期记忆:历史对话的关键信息(如用户偏好、常见问题),存储在 MySQL 中,用于个性化推荐。任务记忆:历史任务的执行记录(如"用户 A 在项目 X 中问过问题 Y,回答是 Z"),存储在 Neo4j 知识图谱中,用于知识积累。

上下文工程。优化 Prompt 设计,减少不必要的上下文,降低 Token 消耗。使用**向量检索**,从历史对话中检索相关内容,而非将所有历史对话都放入 Prompt。设计**结构化 Prompt 模板**,引导 AI 按照固定格式输出,便于后续解析。

3.2.5 难点五:保密行业的高准确率要求

问题描述

公司客户以军工、航空航天等保密行业为主,这些行业对信息准确性要求极高。AI 的回答如果出现错误或误导性信息,可能导致严重后果(如合规风险、安全隐患)。目前大模型存在"幻觉"问题,即 AI 可能编造不存在的信息,这在保密行业是不可接受的。

解决方案

知识库优先策略。AI 回答时,优先从知识库检索信息,只有知识库中没有时才依赖大模型生成。对于关键信息(如法规、标准、规范),只允许从知识库检索,不允许 AI 生成。

多轮校验机制。引入**质量校验 Agent**,对 AI 生成的回答进行二次校验。校验内容包括:是否存在明显错误?是否引用了不存在的文献?是否与知识库内容矛盾?如果校验不通过,重新生成或标注"该回答未经验证,仅供参考"。

引用溯源。AI 回答时,标注信息来源(如"该信息来自《SysML 规范 1.6》第 3.2 节")。用户可以点击查看原始文献,验证 AI 回答的正确性。

人工审核。对于关键任务(如合规校验、仿真报错分析),AI 只提供建议,最终决策由人工审核。建立**审核日志**,记录谁在何时审核了哪些内容,保证可追溯。

测试策略

设计 100 个典型测试用例,涵盖常见问题、边界情况、异常情况。目标准确率**≥85%**,误报率**<5%**。定期进行回归测试,确保系统升级后准确率不下降。

3.3 知识库策略

3.3.1 知识库来源

辰极智脑 AI 平台的知识库分为**三类**:

1. 外部合作知识库(MBSE+各垂类行业知识)

来源为合作伙伴提供的 MBSE+各垂类行业知识库(如 OMG 官方文档、学术论文、行业标准、卫星领域文档、航海领域文档)。内容涵盖 SysML 规范、MBSE 理论、建模方法论、垂类行业 MBSE 最佳实践、垂类行业 MBSE 相关文档。接入方式通过 API 接入,调用合作伙伴的对话服务,获取专业领域知识。优势在于专业性强、覆盖面广,节省自建成本。

2. 内部自建知识库(软件知识+AI 知识)

软件使用知识包括天玑/玉衡用户手册、操作指南、常见问题解答;MBSE 2.0 软件的功能说明、API 文档、开发指南;公司内部的建模方法论、项目流程、质量规范。

AI 智能体知识包括提供给智能体的工具文档(如"如何调用 MCP 接口?""如何查询 Neo4j 数据库?");智能体工作流的设计文档、调试经验、常见问题;Prompt 模板库、优化策略、最佳实践。

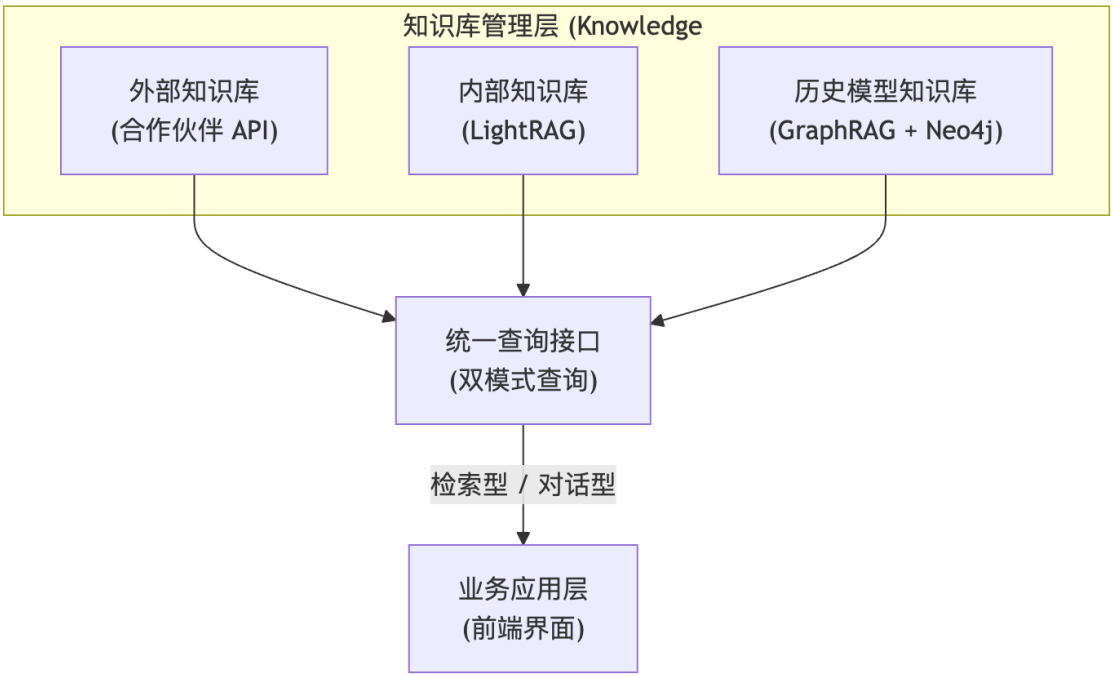
3. 历史模型知识库(GraphRAG)

背景:公司历史上交付了大量 MBSE 项目,积累了丰富的模型资产。

价值:这些模型是公司的核心竞争力,可以作为新项目的参考模板,显著提升建模效率。

技术方案:将历史模型转化为**知识图谱**(Neo4j),建立模型元素之间的关联关系。采用**GraphRAG(Graph Retrieval-Augmented Generation)**技术,基于知识图谱进行检索。用户输入需求后,系统从知识图谱中检索相似模型,推荐给用户或自动组装。

3.3.2 知识库技术架构



双模式查询

检索型(Retrieval):用户提出明确问题(如"SysML 中的 Block 定义是什么?"),系统从知识库中检索最相关的文档片段,直接返回。

对话型(Conversational):用户进行多轮对话(如"Block 和 Class 有什么区别?""能举个例子吗?"),系统维护对话上下文,结合知识库生成连贯回答。

第四章 实施计划与排期

4.1 25 年底前排期(2025 年 10 月 20 日-12 月 31 日,2.5 个月)

当前时间节点为 2025 年 10 月 20 日,距离年底剩余 72 天。在这一关键时期,项目将并行推进四个核心阶段,为明年的全面展开奠定坚实基础，由于后端团队人手紧缺，后端开发工作由 AI 工程师兼任。

4.1.1 第一:需求合规性校验测试

本阶段聚焦于验证 AI 系统在军工保密行业需求分析中的准确性和可靠性。测试团队将设计覆盖合规、不合规及边界情况的 100 个测试用例,通过系统化测试和迭代优化,确保系统能够准确识别需求文档中的合规性问题。

周次	时间	工作内容	负责人	交付物
W1	10.20-10.26	设计测试用例(100 个),覆盖合规、不合规、边界情况	测试工程师	测试用例库
W2	10.27-11.02	执行测试,记录问题,统计准确率、召回率、误报率	测试工程师	测试报告 V1
W3	11.03-11.09	针对问题进行优化,重点解决军工保密行业的准确率要求	朱思雨	优化版本
W4	11.10-11.15	回归测试,验证优化效果,编写用户手册	测试工程师+技术文档	最终测试报告

验收标准: 准确率≥85%,误报率<5%,测试用例通过率≥90%。这些指标的达成将确保系统在实际业务场景中的可靠性和可用性。

4.1.2 第二:对话系统开发

对话系统是平台的核心交互界面,采用多智能体协作架构设计。系统将整合信息收集、内容整合和质量校验三大 Agent 模块,实现从知识库检索、MCP 接口调用到联网搜索的全方位信息获取能力。通过多源信息融合和上下文管理,系统将为用户提供准确、完整、可追溯的专业解答。

周次	时间	工作内容	负责人	交付物
W1-2	10.20-11.02	设计多智能体协作架构(信息收集、内容整合、质量校验)	李泽宇	架构设计文档
W3-4	11.03-11.16	开发信息收集 Agent(知识库检索+MCP 接口+联网搜索)	朱思雨	Agent 代码
W5-6	11.17-11.30	开发内容整合 Agent(多源信息融合+上下文管理)	朱思雨	Agent 代码
W7-8	12.01-12.14	开发质量校验 Agent(准确性校验+引用溯源)	朱思雨	Agent 代码
W3-9	11.03-12.20	集成测试、性能优化、用户界面开发	全员	对话系统 V1.0

验收标准: 系统需支持多轮对话并正确维护上下文,响应时间≤3 秒,准确率≥80%。这将确保用户获得流畅的交互体验和可靠的信息服务。

4.1.3 第三:1.6 版本 AI 建模集成

本阶段将 AI 建模能力集成到天玑/玉衡建模工具中,实现从自然语言描述到 SysML 模型的自动化生成。系统将通过 LangGraph 工作流调度,接收文本输入,自动生成符合 SysML v1.6 规范的模型文件。前后端协同开发确保功能的完整性,演示案例的准备将验证系统在真实场景中的可用性。

周次	时间	工作内容	负责人	交付物
W1-2	11.01-11.14	设计接口,连接 LangGraph 工作流与天玑/玉衡工具	李泽宇	接口设计文档
W3-6	11.15-12.12	开发前端界面,支持文本输入、生成触发、XMI 导入	前端工程师	前端界面
W3-6	11. 15-12.12	开发后端接口,调用 LangGraph 工作流,处理 XMI 文件	李泽宇、胡俊超	后端接口
W7-8	12.13-12.26	集成测试,准备演示案例(自行车、卫星、无人机等)	测试工程师	演示案例库
W9	12.27-12.31	编写功能说明文档,明确定位和局限性	胡俊超	功能说明文档

验收标准: 成功集成到天玑/玉衡工具,演示案例运行正常,XMI 文件格式正确可导入工具。完整的功能说明文档将帮助用户理解系统能力边界。

4.1.4 第四阶段:服务中台完善(贯穿全程,10 月 20 日-12 月 31 日)

服务中台作为整个 AI 平台的核心基础设施,需要在前后端、运维部署和测试等多个维度进行系统化完善。这是一个长期演进的工程,2025 年底前将完成 MVP 级别的核心功能。

类别	工作内容	负责人	完成时间
前端优化	既有页面重构、新页面开发、Bug 修复、性能优化	李泽宇	12 月 31 日
后端优化、新功能开发	配置服务重构、编排器重构、性能优化、注册中心二期开发	李泽宇、胡俊超	12 月 31 日
运维部署	Docker 部署流程完善、监控日志系统搭建	李泽宇(临时负责)	12 月 31 日
测试	压力测试(目标 100 并发),稳定性测试	测试工程师(内部协作)	12 月 31 日

验收标准: 支持 100 并发用户,可用性≥99%,Docker 部署流程完整且文档齐全。

服务中台持续性开发特性: 服务中台是整个 AI 平台的核心基础设施,不会在 2025 年底前完全完成,而是需要在整个 AI 组工作周期中持续开发和演进。2025 年第四季度(当前阶段)将完成 MVP 级别的服务管理、知识库管理和注册中心;2026 年全年将深化服务注册发现机制、转发机制、依赖队列管理和保活机制以及一系列开发过程中产生的新需求;2027 年将优化 100+AI 子服务的依赖监控体系、性能调优和高可用保障。服务中台的工程复杂度高、依赖关系深,是长期维护和演进的核心模块,预计在整个项目周期(2025-2027 年)内持续投入研发资源。

4.2 26 年上半年排期(2026 年 1 月 1 日-6 月 30 日,6 个月)

2026 年上半年的核心目标是在 6 月交付 MBSE Copilot MVP 版本。这一目标的实现需要知识库建设、Agent 框架开发和 MBSE 2.0 软件集成的协同推进。

4.2.1 里程碑 4:知识库与 Agent 基础(1-3 月,3 个月)

知识库是 AI 系统的知识源泉,本阶段将系统化收集和整理 MBSE 领域的本体论知识、公司方法论文档和历史项目案例。同时开发支持多智能体协作的 Agent 框架,为后续功能开发提供技术底座。

月份	工作内容	负责人	交付物
1 月	收集 MBSE 本体论知识(SysML 规范、MBSE 理论),构建本体论知识库	公司团队	本体论知识库 V1
2 月	收集公司方法论文档、历史项目案例,构建方法论知识库和案例库	数据标注团队	方法论知识库+案例库
3 月	设计开发 Agent 框架(任务分解、任务分配、结果聚合、冲突解决)	AI 团队	Agent 框架 V1.0

验收标准: 知识库覆盖率≥80%(核心知识点),知识检索召回率≥85%、精确率≥90%,Agent 框架支持多智能体协作。高质量的知识库和稳定的框架是后续功能开发的基石。

4.2.2 里程碑 5:MBSE 2.0 与 AI 能力开发(1-6 月,6 个月)

MBSE 2.0 软件本体由独立团队负责(另行立项),AI 团队聚焦 AI 能力开发。两个团队并行工作,通过 API 接口实现集成,确保即使 2.0 软件进度滞后,AI 功能也能独立演示。

AI 团队将重点开发知识问答、智能建议和模型增删改查三大核心功能。知识问答功能结合本体论、方法论和案例库,为用户提供专业解答;基于 KerML 文本化建模,提升建模效率;智能建议通过分析模型结构提供优化建议;模型增删改查实现自然语言到 SysML 代码的转换,大幅降低建模门槛。

月份	AI 团队工作内容	MBSE 2.0 团队工作内容	集成点
1 月	设计 MBSE Copilot 架构(知识问答、智能建议)	设计 MBSE 2.0 总体架构(前端+后端+建模引擎)	-
2 月	开发知识问答功能(结合本体论、方法论、案例库)	开发前端界面(Vue 3)、后端基础框架(Spring Boot)	定义 API 接口
3 月	开发智能建议功能(分析模型结构,提供优化建议)	开发建模引擎(SysML v2.0 解析器、验证器)	API
4 月	开发模型增删改查功能(自然语言 → SysML 代码)	开发多人协同编辑功能(OT/CRDT 算法)	模型分析 API
5 月	开发模型增删改查功能(自然语言 → SysML 代码)	完成核心建模功能(创建、编辑、删除模型元素)	模型操作 API
6 月	集成测试、性能优化、准备 MVP 演示	集成测试、用户界面优化	联调与测试

验收标准(6 月 30 日): MBSE Copilot 可独立演示,支持知识问答、智能建议、模型增删改查,响应时间≤20 秒,准确率≥80%。

风险说明: 若 MBSE 2.0 软件开发进度滞后,6 月无法完成核心功能,AI 团队将使用模拟数据开发和测试,确保 AI 功能能够独立演示。待 2.0 软件完成后,替换为真实数据,快速完成集成。

26 年 AI 团队上半年人员配置: 核心团队包括李泽宇、朱思雨、胡俊超(3 人全职),根据开发量需要按需及时扩充,协作团队包括前端工程师(内部协作,50%工作量)和测试工程师(内部协作,按需投入)。26 年 6 月 MVP 节点,全员投入集成测试与演示准备,确保 MVP 质量达标。

4.2.3 里程碑:基础设施与部署运维(1-6 月,贯穿全程)

为支撑 MBSE 2.0 的 BS 架构和多人在线建模需求,本阶段将完成云化部署方案设计、协同功能开发和生产环境搭建。重点解决负载均衡、数据库集成、缓存策略等关键技术问题,确保系统能够支持 100-200 人的同时建模工作。

验收标准: 支持 100-200 人协同工作(Git 异步协作模式),响应时间≤500ms,可用性≥99%。

4.3 26 年下半年至 27 年排期(2026 年 7 月-2027 年 12 月,18 个月)

4.3.1 功能完善阶段(26 年 7 月-12 月,6 个月)

在 MVP 版本基础上,本阶段将补齐模型校验、CI/CD 自动化和仿真优化三大关键功能。模型校验层将提供语法、语义和工程正确性的三层校验,确保生成模型

的质量;CI/CD 自动化管线将实现自动提交、保存和版本管理,提升开发效率;仿真报错 AI 解决功能将分析仿真错误并提供修复建议,降低问题排查成本。

时间	工作内容	交付物
7-9 月	开发模型校验层(语法、语义、工程正确性校验)	校验功能 V1.0
9-11 月	开发 CI/CD 自动化管线(自动提交、自动保存、版本管理)	CI/CD 管线 V1.0
11-12 月	开发仿真报错 AI 解决功能(报错分析、修复建议)	仿真优化功能 V1.0

4.3.2 生态培育阶段(27 年 1 月-12 月,12 个月)

2027 年将聚焦于平台生态建设,通过开放 API 接口、建设智能体市场、开发低代码平台和完善解决方案交付能力,构建开放共享的 MBSE 生态系统。开发者可以基于平台开发定制化智能体并在市场上架,企业可以通过解决方案交付平台快速获取所需的模型和服务,教育机构和企业可以根据需求选择不同版本的订阅服务。

时间	工作内容	交付物
Q1 (1-3 月)	开放 API 接口,发布开发者文档,启动智能体市场	MBSE App Store V1.0
Q2 (4-6 月)	开发低代码平台(可视化编排、模板库、社区分享)	低代码平台 V1.0
Q3 (7-9 月)	开发解决方案交付功能(模型检索、模型组装、模型交易)	解决方案交付平台 V1.0
Q4 (10-12 月)	完善订阅制商业模式,推广教育版、试用版、企业版	商业化运营

关键指标(27 年底): 平台注册开发者≥100 人,智能体市场上架智能体≥50 个,解决方案交付订单≥20 单,订阅客户≥100 家。这些指标的达成将标志着 MBSE 生态的初步形成和平台商业价值的验证。

第五章 团队与资源配置

5.1 团队组成（3 人核心团队）

角色	姓名	2025 职责	技能背景
项目负责人兼架构师	李泽宇	整体架构设计、服务中台开发、前后端技术决策、运维管理、团队协调	MBSE 领域开发工程师,AI 技术背景,全栈开发能力

AI 工程师(合规与对话)	朱思雨	合规性校验工作流开发、对话系统集成、知识库管理	Python/LangGraph、RAG 系统、多智能体协作
AI 工程师(建模与服务)	胡俊超	前期学习为主，一个月后配合开发 MBSE 1.6 集成与优化、服务注册发现模块深化开发	Python/LangGraph、RAG 系统、多智能体协作 (**2025 年 10 月 13 日入职**)

当前团队构成(3 人核心团队)

李泽宇(项目负责人兼架构师)统筹全局,负责架构设计、服务中台全栈核心开发、前后端技术决策、运维部署管理,协调 AI 团队与 MBSE 2.0 团队的工作。

朱思雨(AI 工程师)负责合规性校验工作流、对话系统开发、知识库管理,调用 AI 服务中台的能力实现业务层 AI 功能。

胡俊超(AI 工程师,2025 年 10 月 13 日入职) 先学习，后期配合 MBSE 建模功能 1.6 集成与优化、参与服务注册发现模块的深化开发,协助服务中台的工程实现。

外部协作与资源支持

算法工程师：26 年新增 1 人,负责后训练、优化、知识库构建、RAG 性能优化(正式入编)。

前端工程师：公司内部其他项目团队协作,负责工业软件前端界面开发。

后端工程师：公司内部其他项目人手紧缺，由 AI 工程师兼任。

测试工程师：公司内部测试团队协作,负责功能测试、性能测试、安全测试。

运维工程师：当前暂缺,由李泽宇临时负责运维工作,26 年计划招聘专职运维工程师或通过公司内部调配。

数据标注团队：外部协作,负责知识库构建、语料标注。

5.2 人力分配与时间安排

5.2.1 25 年(2025 年 10 月-12 月,3 个月)

角色	人员	工作内容	工作量
项目负责人兼架构师	李泽宇	总体规划、架构设计、服务中台全栈开发、运维管理	全职
AI 工程师(合规与对话)	朱思雨	对话系统开发、合规性校验测试	全职

AI 工程师(建模与服务)	胡俊超	前期学习，后期 MBSE 1.6 集成、服务注册发现模块开发	全职
前端工程师	内部协作	前端优化、对话界面开发	兼职(50%)
后端工程师	AI 兼职	天玑玉衡摇光的后端开发由 AI 组兼任	全职
测试工程师	内部协作	功能测试、性能测试	按需投入

任务分配细化

李泽宇的主要任务包括：服务中台架构设计与核心模块开发(服务注册、转发机制、编排器重构);MBSE1.6AI 建模集成的核心开发（玉衡后端，AI 建模服务后端）;临时负责运维部署(Docker、监控日志、压力测试);团队技术评审与风险控制。

朱思雨的主要任务包括：对话系统开发(整合 MCP 接口、知识库 RAG、联网搜索);需求合规性校验功能的系统测试与优化;知识库内容管理与质量保障。

胡俊超(10 月 13 日入职,新员工培养期)的主要任务包括：熟悉现有系统架构与代码库(第 1-4 周);参与 MBSE 1.6 集成的接口开发与测试(11 月);深化服务注册发现模块开发,学习微服务架构实践(12 月)。

5.2.2 26 年(2026 年 1 月-12 月,12 个月)

角色	人员	工作内容	工作量
项目负责人兼架构师	李泽宇	整体规划、核心功能开发管控、团队管理、对外协调、架构演进	全职
AI 工程师	朱思雨、胡俊超	知识库构建、Agent 框架、MBSE Copilot 开发、模型校验优化	全职(2 人)
算法工程师	**新增 1 人(下半年加入)**	后训练、优化、RAG 性能优化	全职(**正式入编**)
新增 AI 工程师	**新增 2 人(按需加入)**	扩展 AI 能力、支撑规模化开发	全职(**正式入编**)
前端工程师	内部协作	前端界面开发	兼职(50%)
测试工程师	内部协作	功能测试、性能测试、安全测试	按需投入
运维工程师	**计划招聘或内部调配**	Docker 部署、监控日志、故障排查	全职(**26 年 Q2 加入**)

2026 年团队扩编计划

上半年(1-6 月)：首先维持 3 人核心团队(李泽宇、朱思雨、胡俊超),按需在必要时完成团队扩充以确保完成 26 年 6 月 MVP 交付，数据标注团队外包(外部协作)。

下半年(7-12 月)：团队规模预计扩展至 **6 人正式研发团队**。随上半年按时间共新增 2 名 AI 工程师(正式入编),负责扩展 AI 能力、支撑规模化开发;新增 1 名算法工程师(正式入编),负责后训练、优化、RAG 性能优化;新增 1 名运维工程师(正式入编或内部调配),专职负责运维部署;前端与测试继续由公司内部其他项目团队协作，数据标注团队外包(外部协作)。

团队结构演进逻辑：从 MVP 阶段的 3 人核心组(2025 Q4 - 2026 Q2)逐步扩展为 6 人正式研发团队(2026 H2),形成长期稳定的工程体系,支撑 27 年的生态建设与规模化推广。

5.3 协作机制

5.3.1 内部协作

周会：每周一上午全员参加,同步项目进度、讨论技术问题、分配工作任务,确保团队成员对项目整体状况有清晰的认识。

每日：每天团队负责人会在必要时与团队成员讨论,记录当天完成的工作内容、遇到的问题及解决方案、明天的工作计划,保持团队透明度和可追溯性。

技术评审：对于关键技术决策(如架构设计、技术选型等)召开技术评审会议,邀请外部专家参与,充分论证技术方案的可行性和风险,确保技术路线正确。

代码审查：所有代码提交前必须经过同事审查,通过代码审查机制保证代码质量,降低缺陷率,促进团队成员之间的技术交流和知识共享。

5.3.2 外部协作

与 MBSE 2.0 团队协作：定义清晰的 API 接口,AI 团队与 2.0 团队通过接口集成,避免系统耦合。每周召开联调会议,同步进度、解决集成问题。AI 团队提前使用模拟数据开发,待 2.0 团队完成后快速替换为真实数据,确保开发并行推进。

与合作伙伴协作(外部知识库)：签订合作协议,明确知识库的使用范围、API 调用次数、费用结算等关键条款。定期评估合作伙伴的服务质量,如不满足项目要求,及时切换到备用方案,降低供应商依赖风险。

与测试/运维团队协作：外包团队（如数据标注）按需投入,避免长期养团队的固定成本。公司共用团队确保建立清晰的交付标准和验收流程,保证共用团队的工作质量符合项目要求,通过明确的合同条款保障项目进度和质量。

第六章 财务预算与 ROI 分析

6.1 总预算（500 万元）

6.1.1 预算明细

类别	25 年 Q4	26 年全年	27 年全年	合计	说明
人力成本	30 万	200 万	-	**230 万**	3-6 人核心团队+内部协作+外包团队
GPU 服务器（一次性）	-	250 万	-	**250 万**	8-16 卡 H20，含 CPU、内存、存储
存储资源	1 万	5 万	-	**6 万**	向量数据库、图数据库存储（云化）
第三方 API	2 万	8 万	-	**10 万**	DeepSeek、Kimi 等大模型 API（测试，非运维支出）
云服务	-	10 万	-	**10 万**	阿里云/腾讯云（可选）
其他支出	2 万	6 万	-	**8 万**	差旅、培训、会议等
总计	**35 万**	**479 万**	**0**	**514 万**	**实际控制在 500 万以内**

预算说明：27 年不再投入新的预算,依靠 26 年底交付的产品开始商业化运营并产生营收。GPU 服务器是一次性投入,采购 8-16 卡 H20,单卡价格约 15-20 万,总计 200-300 万,加上 CPU、内存、存储等配套设备,预算 250 万。人力成本控制在 230 万,占总预算的 45%,符合软件项目的通常比例(人力成本占 50%-60%)。第三方 API 费用按每次调用 0.01 元、每天 10 万次调用、年度 365 天计算约 365 万,但通过缓存、请求合并等优化预计可降低 50%,实际年度费用约 180 万;为保守起见,预算按 8 万/年估算以覆盖开发测试阶段的低频调用。

6.1.2 GPU 服务器配置方案

方案一：8 卡 H20 配置（预算 200 万）

- GPU：8 卡 H20，单卡 15 万 × 8 = 120 万
- CPU：双路 AMD EPYC，192–384 核心、384–768 线程= 10 万
- 内存：512GB DDR4 ECC = 8 万
- 存储：20TB NVMe SSD（向量数据库+图数据库） = 2 万

- 网络：万兆网卡 + 交换机 = 5 万
- 机架/电源/散热等 = 10 万
- **总计**：**155 万 ≈ 170 万**（含安装调试）

方案二：16 卡 H20 配置（预算 300 万）

- GPU：16 卡 H20，单卡 15 万 × 16 = 240 万
- CPU：双路 AMD EPYC，192–384 核心、384–768 线程 = 10 万
- 内存：1TB DDR4 ECC = 15 万
- 存储：30TB NVMe SSD =3 万
- 网络：万兆网卡 + 交换机 = 8 万
- 机架/电源/散热等 = 15 万
- **总计**：**291 万 ≈ 300 万**（含安装调试）

推荐方案：本项目推荐采用方案一(8 卡 H20),主要基于以下理由:8 卡 H20 已能满足当前需求(知识库构建、模型推理、Agent 开发);如果后续需求增长,可再采购第二台 8 卡服务器实现水平扩展;同时节省预算,将省下的 100 万用于人力成本或 API 费用。

6.2 收入模型（27 年底 5000 万元）

6.2.1 收入来源构成（总览）

收入来源	占比	金额	客户构成	单价范围
软件订阅（企业版、大型企业版、教育版）	28%	1400 万	45-50 企业客户	20-150 万/年
License 买断（军工/国企内网部署）	12%	600 万	5-7 家大客户	100-500 万/套
项目交付营收（军工类甲方项目）	35%	1750 万	军工/航空航天客户	50-200 万/项目
AI 增值服务（高级功能包）	13%	650 万	70-80 个企业客户	5-15 万/年
生态平台抽成（智能体+模型交易）	12%	600 万	生态开发者	30%抽成
总计	**100%**	**5000 万**	**~130-150 客户**	-

6.2.2 软件订阅收入（1400 万，28%）

客户类型	数量	单价范围	收入贡献	说明
大型企业版	5 家	80-150 万/年	550 万	200+人团队，定制化服务
企业版	30 家	20-50 万/年	750 万	50-200 人团队，标准功能
小微企业版	8 家	1-5 万/年	25 万	10 人以下团队，基础功能
教育版	10 所	5-10 万/年	75 万	高校、培训机构
小计	**53 家**	-	**1400 万**	-

订阅制定价策略：采用按人数定价与按功能定价相结合的模式。企业版 20 万/年支持 50 人,平均每人 4000 元/年;大型企业版 80 万/年支持 200 人,平均每人 4000 元/年。基础功能(建模、保存)免费,高级功能(Git 协作，AI 辅助、校验、仿真优化)收费。教育版享受 50%折扣,通过培养 MBSE 人才形成长期用户粘性。

6.2.3 License 买断收入（600 万，12%）

客户类型	数量	单价范围	收入贡献	说明
军工/保密企业	2-3 家	150-300 万/套	450 万	内网部署，买断 License
大型国企/央企	2 家	50-100 万/套	150 万	内网部署或混合云
小计	**4-5 家**	-	**600 万**	-

买断定价策略：客户一次性付费购买永久 License,包含软件本体及 AI 功能。公司提供现场私有化部署、培训及技术支持服务。买断后每年收取维护费(License 价格的 20%),用于软件升级、Bug 修复和技术支持。

目标客户包括军工企业(航天科技、航天科工、中国兵器、中国船舶等)和大型国企(中国商飞、中国中车、国家电网等)。

6.2.4 项目交付营收（1750 万，35%）

项目类型	数量	单价范围	收入贡献	说明
保密 MBSE 建模项目	6-8 个	150-300 万/项目	1200 万	航空航天、武器装备系统建模
商业航空航天项目	2-3 个	100-200 万/项目	350 万	飞机、卫星系统建模
其他行业项目	3-5 个	50-80 万/项目	200 万	汽车、能源等行业
小计	**11-16 个**	-	**1750 万**	-

项目交付特点：AI 辅助建模使效率提升 10 倍,项目周期从 6 个月缩短至 3-4 个月,具有高附加值特征。成功交付项目后客户复购率预计达 60%以上,同时项目交付过程能够展示平台能力,有效促进软件订阅和 License 买断销售。

6.2.5 AI 增值服务收入（650 万，13%）

服务类型	数量	单价范围	收入贡献	说明
高级 AI 功能包	70-80 个	8-10 万/年	650 万	智能建模、自动校验、仿真优化

增值服务定价：基础订阅客户如需使用高级 AI 功能(如自动校验、仿真优化),需额外支付增值服务费。定价按功能模块收费:智能建模 5 万/年、自动校验 5 万/年、仿真优化 5 万/年,购买全套优惠至 10 万/年。

6.2.6 生态平台抽成收入（600 万，12%）

收入来源	交易额	抽成比例	收入贡献	说明
智能体市场	1700 万	30%	510 万	开发者发布智能体，用户付费购买
模型交易	450 万	20%	90 万	历史模型、行业方案交易
小计	**2150 万**	-	**600 万**	-

生态平台运营策略：智能体市场方面,开发者发布智能体,用户按月订阅(如 10-50 元/月)或一次性购买(如 100-500 元),平台抽成 30%、70%归开发者,热门智能体可获得推荐位提升曝光度。模型交易方面,公司整理历史项目模型并标注行业、场景、复杂度后发布到平台,客户按模型复杂度定价购买(简单模型 1000-5000 元,复杂模型 1-5 万元),平台抽成 20%、80%归模型提供方。

生态培育目标：27 年底实现平台注册开发者≥500 人、智能体市场上架智能体≥1000 个、智能体年交易额≥2000 万、模型交易年交易额≥500 万。

6.3 ROI 分析

6.3.1 投资回报率计算

总投入：500 万元
总收入（27 年底）：5000 万元
净利润（假设毛利率 70%）：5000 万 × 70% = 3500 万
投资回报率（ROI）：（净利润 - 总投入）/ 总投入 = (3500 万 - 500 万) / 500 万 = **600%**
投资回报周期：从 25 年 10 月开始投入，27 年 12 月实现 5000 万营收，共计 **27 个月**。

6.3.2 收入增长路径

时间节点	功能交付	收入启动	累计收入	客户数	说明
25 年	1.6 AI 集成 + 对话	试用期，少量早	**50 万**	~5 家	早期种子客户

底	系统	期客户			验证
26 年 6 月	MVP（MBSE Copilot）	正式销售启动	**500 万**	~20 家	企业版订阅为主
26 年底	完整功能 + 生态启动	规模化销售	**2000 万**	~80 家	订阅+买断双轮驱动
27 年底	生态成熟 + 订阅续费	生态+订阅双轮驱动	**5000 万**	~160 家	生态平台抽成占 14%

增长路径说明：25 年底阶段聚焦早期种子客户(5 家),主要为战略合作伙伴,以试用为主并收取少量付费(如试用费 10 万/家),总计 50 万。26 年 6 月 MVP 发布后正式启动销售,签约 20 家企业客户(企业版 20-50 万/年),累计 500 万。26 年底功能完善,签约 80 家客户(包括 30 家企业版、8 家大型企业版、5 家买断客户、10 所教育版、27 家小微企业版),累计 2000 万。27 年底生态成熟,签约 160 家客户,生态平台抽成 700 万,总计 5000 万。

6.3.3 成本效益分析

内部效益（效率提升）：建模效率提升 10 倍,传统建模需要 2-4 周,AI 辅助后仅需 2-3 天,节省 90%时间。单项目减少建模工程师投入量,年节约成本 60-90 万。项目交付周期从 6 个月缩短至 3-4 个月,显著提升客户满意度。

外部效益（市场竞争力）：功能对标国际巨头,AI 辅助建模能力超越 IBM Rhapsody、Siemens Cameo。订阅制定价 20-50 万/年,低于国际产品 50-100 万/年的买断价格,具有明显价格优势。支持私有化部署、保密行业定制,满足军工/国企的本地化服务需求。

长期战略价值：历史项目模型转化为知识图谱,形成不可复制的行业壁垒作为知识资产沉淀。开发者社区形成后,生态平台效应将带来平台价值的指数级增长。用户使用越多模型质量越高,从而吸引更多用户,形成数据飞轮的正向循环。

第七章 风险管理与应对

7.1 重大风险识别

7.1.1 风险一：MBSE 2.0 软件开发进度滞后（高优先级）

风险描述

MBSE 2.0 软件计划在 25 年底或 26 年初启动开发,预计开发周期 12-18 个月。如果 2.0 软件开发进度滞后(如推迟到 27 年 Q2 才完成),将严重影响 AI 功能的最终

集成。AI 团队的工作成果(如 MBSE Copilot、模型增删改查)无法在实际软件中落地,导致投资浪费。

风险等级：高(发生概率：中，影响程度：高)

应对策略

本项目针对该风险制定了三层应对策略,确保即使 2.0 软件开发滞后,AI 功能仍能按计划交付。

第一层策略是 AI 团队提前使用模拟数据开发。AI 团队在 2.0 软件尚未完成时,使用模拟数据(Mock Data)进行开发和测试。26 年 6 月交付的 MVP 版本可以独立演示,无需依赖 2.0 软件。待 2.0 软件完成后,快速替换为真实数据,完成集成。这种策略允许 AI 团队独立推进开发进度,不受 2.0 软件开发周期限制。

第二层策略是定义清晰的 API 接口。AI 团队与 2.0 团队在项目启动时,定义清晰的 API 接口(如模型创建、查询、修改、删除接口)。AI 团队基于接口开发,2.0 团队基于接口实现,避免耦合。即使 2.0 软件内部实现变化,只要接口不变,AI 功能不受影响。这种"契约式开发"模式确保了两个团队可以并行工作,降低相互依赖。

第三层策略是保留 1.6 版本作为备选方案(Plan B)。如果 2.0 软件开发严重滞后,AI 团队继续在 1.6 版本上深化功能,提升成功率、支持增量修改、优化校验能力。1.6 版本虽然体验不如 2.0,但至少能交付可用的 AI 功能,避免项目完全失败。

监控指标

为及时发现风险并启动应对措施,项目将建立定期监控机制。每月与 2.0 团队召开进度会议,跟踪开发进度。如果 2.0 团队进度滞后超过 2 个月,立即启动 Plan B,确保项目不会因外部依赖而陷入被动。

7.1.2 风险二：1.6 版本 AI 建模功能不好用（中优先级）

风险描述

1.6 版本的 AI 建模功能(基于 LangGraph 工作流)存在明显缺陷：黑盒化、不可编辑、体验差。该功能主要用于展示能力,但如果客户试用后发现不好用,可能对公司的 AI 能力产生负面评价。

风险等级：中(发生概率：高，影响程度：中)

应对策略

针对 1.6 版本功能缺陷,项目采取"明确定位、优化展示、收集反馈"三步走策略。

首先,明确功能定位至关重要。在功能说明文档和客户演示中,明确该功能是技术验证版本,主要展示公司在 AI 建模领域的技术储备。强调真正可用的 AI 建模功能将在 26 年 6 月的 MVP 版本中交付(基于 SysML v2.0)。通过设定合理预期,避免客户误以为 1.6 版本是最终产品,从而降低期望值。

其次,准备精美演示案例以展示优势。精心设计 3-5 个演示案例(如自行车系统、无人机系统),确保演示过程流畅、效果惊艳。演示案例聚焦于"快速生成模型"的能力,而非"人机协同编辑"能力,扬长避短。通过精心准备的演示,展现 AI 技术的潜力,而非暴露当前版本的缺陷。

第三,收集客户反馈指导 2.0 开发。虽然不对 1.6 功能进行进一步开发,但收集客户反馈(如"希望支持增量修改"、"希望看到中间过程"),作为 2.0 功能设计的参考。这种方法将 1.6 版本的局限性转化为 2.0 版本的改进方向,使客户的负面体验变成有价值的需求输入。

7.1.3 风险三：技术难度超预期（中优先级）

风险描述

多项技术难点(如服务中台、Git 模型合并机制、100-200 并发支撑、多智能体协作)技术难度大,可能出现攻关失败或开发周期延长。如果核心技术难题无法解决,将影响功能交付和质量。

风险等级：中(发生概率：中，影响程度：高)

应对策略

面对技术挑战,项目采取"提前预研、外部支援、分阶段交付"的综合应对策略。

技术预研是降低风险的第一步。在项目启动初期开始(25 年 10 月),对关键技术难点进行预研。通过预研评估技术可行性和实现难度。如果预研发现技术难度过大,及时调整方案或降低目标,避免在项目中期遭遇无法逾越的技术障碍。

对于公司内部无法解决的技术难题,引入外部专家提供支持。通过购买咨询服务或技术合作,快速获得关键技术 know-how,缩短攻关周期。

分阶段交付策略确保项目持续推进。将复杂功能拆分为多个阶段,先交付基础版本,再逐步完善。例如,26 年 6 月 MVP 版本只支持 50 人并发,26 年底再提升至 100-200 人并发。这种"小步快跑、迭代完善"的方法,既降低了技术风险,又能及时向客户展示阶段性成果。

7.1.4 风险四：市场接受度不确定（中优先级）

风险描述

AI 辅助建模是新兴技术,市场对其价值和可行性存在质疑。客户可能担心 AI 生成的模型质量不可靠,或担心 AI 取代工程师导致失业,从而抵触 AI 功能。如果市场接受度低,销售困难,将影响收入目标的实现。

风险等级：中(发生概率：中，影响程度：中)

应对策略

提升市场接受度需要通过实际案例、优质服务和正确定位来消除客户疑虑。

种子客户验证是建立市场信心的关键。在 25 年底,与 5 家战略合作伙伴签订试用协议,免费提供 AI 功能,收集使用反馈。根据反馈优化功能,积累成功案例,形成示范效应。真实客户的成功案例比任何营销宣传都更有说服力,能够有效降低后续客户的采购顾虑。

强化培训与技术支持确保客户能够充分利用 AI 功能。为客户提供详细的培训材料(视频教程、操作手册、常见问题解答),提供 7×24 小时技术支持,快速响应客户问题。组织线上/线下研讨会,邀请行业专家分享 AI+MBSE 的价值。通过完善的服务体系,帮助客户克服学习曲线,真正体验到 AI 带来的效率提升。

在宣传策略上,强调"人机协同"而非"AI 替代"。在宣传中强调 AI 是工程师的助手,而非替代品,帮助工程师提升效率,而非取代工程师。展示 AI 如何辅助工程师完成重复性工作(如模型校验、文档生成),让工程师有更多时间专注于创造性工作。通过正确的价值定位,消除工程师对 AI 的抵触情绪,将 AI 从"威胁"转变为"助手"。

7.1.5 风险五：外部 API 不稳定（低优先级）

风险描述

AI 服务依赖外部大模型 API(如 DeepSeek、Kimi),这些 API 可能出现限流、超时、返回错误、服务中断等情况。如果外部 API 不稳定,将影响 AI 服务的可用性和用户体验。

风险等级：低(发生概率：低，影响程度：中)

应对策略

虽然该风险优先级较低,但仍需建立多层保障机制确保服务稳定性。

多 Provider 备份是应对 API 不稳定的核心策略。接入多个大模型 API(DeepSeek、Kimi、GLM 等),当主 Provider 不可用时,自动切换到备用 Provider。在 AI 服务中台层实现 Provider 管理和自动切换逻辑,确保单一 API 故障不会导致整体服务中断。

缓存机制可以在 API 不可用时提供基础服务。对常见问题的回答进行缓存(Redis),当外部 API 不可用时,返回缓存结果。缓存有效期设置为 1 小时,避免缓存过期后仍然返回旧答案。这种机制特别适用于文档查询、规范解释等变化较少的场景。

当所有备份方案都失效时,降级策略确保用户获得清晰的反馈。当所有外部 API 都不可用时,返回预设的备用回答(如"系统维护中,请稍后重试")或提示用户联系技术支持。避免因技术故障导致用户困惑或数据丢失,维护系统的可靠性印象。

7.2 质量保障（保密行业、军工行业准确率要求）

7.2.1 质量目标

项目针对保密行业和军工行业的高标准要求,制定了严格的质量目标体系,涵盖准确性、可用性和性能三大维度。

指标类别	指标名称	目标值	说明
准确性	合规校验准确率	≥85%	100 个测试用例,至少 85 个正确
	模型生成成功率	≥60%	10 个场景测试,至少 6 个成功
	知识检索召回率	≥85%	检索相关文档的比例
	知识检索精确率	≥90%	检索结果的正确性
可用性	服务可用率	≥95%	全年停机时间<438 小时
	故障恢复时间	<15 分钟	从故障发生到恢复的时间
性能	API 响应时间	<800ms	P95 指标
	知识库查询延迟	<15 秒	100 次查询均值
	并发用户数	≥100	压力测试验证

7.2.2 质量保障措施

为确保达成上述质量目标,项目建立了全方位的质量保障体系,涵盖测试、审查、溯源和人工校核四个层面。

测试策略

项目采用多层次测试体系确保功能质量。单元测试覆盖每个模块(如信息收集 Agent、质量校验 Agent),测试覆盖率要求≥70%,确保基础模块的正确性。集成测试验证多个模块集成后的端到端流程,确保模块间协作正常。性能测试使用 JMeter 进行压测,验证并发支撑能力是否达标。安全测试通过漏洞扫描确保无高危漏洞,保护客户数据安全。每次功能更新后执行回归测试,重新运行所有测试用例,确保新功能不影响旧功能的稳定性。

代码审查

所有代码提交前必须经过同事审查(Code Review),审查内容包括代码质量、安全漏洞、性能瓶颈、文档完整性。使用工具辅助审查(如 SonarQube、ESLint),自动检测潜在问题。通过严格的代码审查流程,在开发阶段就消除潜在缺陷,降低后期维护成本。

引用溯源

为提升 AI 回答的可信度,系统在 AI 回答时标注信息来源,例如"该信息来自《SysML 规范 1.6》第 3.2 节"。用户可以点击查看原始文献,验证 AI 回答的正确性。这种引用溯源机制不仅提升了答案的可信度,也为用户提供了深入学习的途径。

人工审核

对于关键任务(如合规校验、仿真报错分析),AI 只提供建议,最终决策由人工审核。建立审核日志,记录谁在何时审核了哪些内容,保证可追溯。这种"AI 辅助+人工决策"的模式,既发挥了 AI 的效率优势,又保证了关键决策的可靠性,特别适合保密行业和军工行业的严格要求。

第八章 预期成果与价值

8.1 技术成果

1. AI 与 MBSE 结合的工程化方案

本项目将形成一套完整的 AI+MBSE 技术方案,涵盖架构设计、技术选型、工作流编排、知识库构建、多智能体协作等核心内容。该方案的设计具有高度的通用性和可扩展性,不仅适用于 MBSE 领域,还可以复用到其他工程软件领域,如 CAD(计算机辅助设计)、CAE(计算机辅助工程)、EDA(电子设计自动化)等,为公司在工程软件智能化方向的**业务拓展合作**奠定坚实的技术基础。

2. 行业知识库与知识图谱

项目将构建一个全面的 MBSE 行业知识库,涵盖本体论、方法论和案例库三大模块,知识覆盖率预计达到 80%以上。同时,将历史项目中积累的模型转化为基于 Neo4j 的知识图谱,建立模型元素之间的复杂关联关系。这些知识资产将成为公司的核心竞争力,不仅可以直接应用于当前项目,还将持续为未来的业务发展提供价值支撑。

3. 智能体开发框架

基于 LangGraph 等技术,项目将开发一套完整的智能体开发框架,支持任务分解、任务分配、结果聚合和冲突解决等关键功能。该框架具有良好的扩展性和复用性,可用于快速开发各种行业专用的智能体应用,例如"汽车 AUTOSAR 建模助手"用

于汽车电子架构设计,"航空 DO-178C 合规检查助手"用于航空软件安全认证等,显著降低新智能体的开发成本和周期。

4. 知识产权

项目计划申请 3 项软件著作权,包括辰极智脑 AI 服务中台、瑶光 AI 助手和 MBSE Copilot 三个核心产品。同时,计划申请 2-3 项发明专利,重点聚焦在核心技术创新上,如"基于知识图谱的 MBSE 模型检索方法"可以显著提升模型检索的准确性和效率,"多智能体协作的建模优化方法"可以实现复杂建模任务的智能化分解与协作。这些知识产权将构筑起公司的技术护城河,有效保护核心技术资产。

8.2 产品成果

1. 辰极智脑 AI 服务中台

这是一个公司级的 AI 服务管理平台,实现对所有 AI 相关能力的统一管理和智能路由。平台支持双模式查询(检索型和对话型),满足不同场景下的用户需求;提供完善的 Provider 管理功能,灵活接入各类公司内外部开发的 AI 服务;集成 SSO 单点登录认证,确保系统安全性;具备强大的知识库管理能力,支持知识的高效组织和检索。该中台采用模块化设计,不仅服务于 MBSE 产品线,还可以复用到其他业务线,实现 AI 能力在公司内的全面赋能。

2. 瑶光 AI 助手

这是需求管理系统的 AI 增强版本,通过集成人工智能技术,显著提升需求管理的智能化水平。系统支持需求合规性的自动校验,及时发现需求中的错误和不一致;提供对话式的需求分析功能,帮助用户快速理解和分析复杂需求;具备强大的知识检索能力,快速定位相关的历史需求和案例。这些功能的引入将大幅提升需求分析的效率,降低因需求问题导致的项目风险。

3. MBSE 2.0 AI 建模工具

基于 SysML v2.0 标准构建的新一代建模工具,原生集成了 AI 能力,为用户提供全新的建模体验。工具内置 MBSE Copilot 智能助手,支持知识问答、和智能建议等功能,大幅降低建模门槛和提升建模效率。系统支持模型的增删改查等完整生命周期管理,集成 Git 版本管理功能,实现模型的版本控制和历史追溯。基于 Git 的异步协作机制,系统可支持 100-200 人规模的团队协同工作,完全满足大型项目的协作需求。

4. MBSE 生态平台

项目致力于打造 MBSE 行业的开放生态平台,类似于苹果 App Store 的商业模式。平台将开放完善的 API 接口,允许第三方开发者在平台上开发和发布智能体应用以及行业专用工具。建立智能体市场,开发者可以将其开发的智能体上架销售,

用户可以按需购买和使用。同时建立模型交易平台,促进行业内模型资产的共享和复用。通过生态平台的建设,吸引更多开发者和用户加入,最终形成繁荣的平台经济。

8.3 商业成果

1. 直接营收

项目预计在 2027 年底实现 5000 万元的营收目标,投资回报率(ROI)达到 600%。收入来源呈现多元化结构:软件订阅服务预计贡献 2000 万元,主要来自中小企业客户的年度订阅费用;License 买断销售预计贡献 1500 万元,主要来自大型企业和军工单位的一次性购买;AI 增值服务预计贡献 800 万元,包括定制化开发、技术咨询和培训服务;生态平台交易抽成预计贡献 700 万元,来自智能体市场和模型交易平台的分成收入。

2. 市场竞争力

在 AI 辅助建模能力方面,项目产品将超越 IBM Rhapsody、Siemens Cameo 等国际知名产品,成为行业领先者。在价格策略上具有显著优势,采用订阅制定价模式,年费仅为 20-50 万元,远低于国际产品 50-100 万元的买断价格,为中小企业降低了准入门槛。同时,提供完善的本地化服务,支持私有化部署以满足数据安全要求,提供保密行业定制化服务,充分满足军工和国有企业的特殊需求,在中国市场形成独特的竞争优势。

3. 客户积累

项目预计在 2027 年底累计签约客户约 160 家,客户结构涵盖各个细分市场。其中,大型企业版客户 8 家,主要为年营收超过 10 亿元的大型企业集团;企业版客户 30 家,面向中型企业提供标准化服务;小微企业版客户 15 家,为创业公司和小微企业提供轻量级解决方案;教育版客户 10 所,与高校和科研院所建立合作关系;买断客户 6-8 家,主要为对数据安全有严格要求的军工单位和国有企业;AI 增值服务客户 80 个,为各类客户提供定制化的 AI 技术服务。

4. 生态培育

通过开放平台战略,吸引并培养开发者生态。预计平台注册开发者数量将达到 500 人以上,形成活跃的开发者社区。智能体市场预计上架智能体 1000 个以上,涵盖各个细分领域的专业应用,为用户提供丰富的选择。智能体市场的年交易额预计达到 2000 万元以上,形成活跃的交易市场。模型交易平台的年交易额预计达到 500 万元以上,促进行业内模型资产的流通和复用。生态的繁荣将为平台带来长期的竞争优势。

8.4 战略价值

1. 知识资产沉淀

通过将历史项目中积累的模型系统性地转化为知识图谱,项目将形成不可复制的行业壁垒。这些知识资产不仅记录了项目的设计决策和经验教训,还蕴含着行业的最佳实践和隐性知识。随着知识库的持续积累和完善,其价值将不断增长,成为公司最核心的竞争力来源。竞争对手即使获得相同的工具和技术,也无法在短时间内复制这些长期积累的知识资产。

2. 平台效应

一旦开发者社区形成并达到临界规模,平台的价值将呈现指数级增长趋势。开发者为平台贡献各类智能体和应用,丰富了平台的功能和服务,从而吸引更多的用户使用平台。用户规模的增长又会吸引更多的开发者加入,为平台开发更多优质的应用。这种正向循环一旦启动,将形成强大的网络效应,大幅提升平台的市场价值和用户黏性,构建起难以撼动的市场地位。

3. 数据飞轮

随着用户使用平台的频率和规模不断增加,系统将积累越来越多的使用数据和反馈信息。这些数据可以用于持续优化模型的质量和 AI 算法的性能,从而提升用户体验,吸引更多用户使用平台。更多的用户又会产生更多的数据,进一步推动 AI 能力的提升。这种数据驱动的良性循环,即"数据飞轮效应",将使平台的 AI 能力随着时间推移不断增强,形成长期的技术竞争优势。

4. 行业引领

通过持续的技术创新和产品迭代,项目将把公司塑造成 AI+MBSE 领域的标杆企业,引领行业的技术发展方向。公司将积极参与行业标准的制定工作,将自身的技术成果和实践经验转化为行业规范,提升公司在行业内的话语权和影响力。通过技术领先地位和标准制定权,公司可以在市场竞争中占据主动,引导行业向有利于自身的方向发展,巩固市场领导地位。

附录:术语表

术语	全称	说明
MBSE	Model-Based Systems Engineering	模型驱动系统工程
SysML	Systems Modeling Language	系统建模语言
LLM	Large Language Model	大语言模型
RAG	Retrieval-Augmented Generation	检索增强生成
GraphRAG	Graph Retrieval-Augmented Generation	基于知识图谱的检索增强生成
MCP	Model Context Protocol	模型上下文协议

SSO	Single Sign-On	单点登录
API	Application Programming Interface	应用程序接口
MVP	Minimum Viable Product	最小可行产品
ROI	Return on Investment	投资回报率
QPS	Queries Per Second	每秒查询数
XMI	XML Metadata Interchange	XML 元数据交换格式
KerML	Kernel Modeling Language	SysML v2.0 的核心建模语言
CI/CD	Continuous Integration / Continuous Deployment	持续集成/持续部署

撰写人:李泽宇
撰写日期:2025.10.19