

博昇慧源AI研究院

Enlighten AI Hub

技术能力概览 · V1.0 · 2026年5月

研究院简介

博昇慧源AI研究院 (Enlighten AI Hub)

专注于工业AI大模型落地与垂直领域AI系统研发的技术机构，依托20年产业咨询积累与自研**Enlitecore 架构体系**，为工业制造、医疗、生物医药、文旅、跨境电商等行业提供从需求分析、系统架构到私有化交付的全链路AI解决方案。

研究院当前已在多个垂直领域完成生产级系统研发，覆盖**计算机视觉、知识图谱推理、多源数据预测、LLM行为架构**等核心AI方向，同步推进学术研究与产业落地的双轨发展路径，致力于成为行业信赖的AI技术合作伙伴。



工业制造

智能制造升级



智慧医疗

诊疗辅助决策



Enlitecore
核心架构



生物医药

药物研发加速



跨境电商

多语言智能服务



数字文旅

沉浸式体验



AI创作

AIGC内容生成

核心技术底座 Enlitecore — 定义



Enlitecore是博昇慧源AI研究院自研的模块化全栈架构体系，历经多个生产级项目严苛验证，具备极强的跨领域适配与快速部署能力，是连接底层算力与上层业务的坚实桥梁。

核心设计理念：共享底座 + 场景架构

所有项目共用一套高可用基础设施层，在此之上灵活适配垂直领域的业务场景架构，兼顾企业级工程一致性与行业专业深度。

场景覆盖广

私有化友好

基于开源技术栈，支持客户本地完全离线部署，无云厂商依赖

工程一致性

极速启动

新项目无需从零搭建，选定匹配架构即可快速启动开发与验证

基础设施层（上）

基础设施层（所有架构共用）

关系型数据库

PostgreSQL + SQLite +
SQLAlchemy + Alembic（版本
迁移）

缓存与任务队列后端

Redis：提供内存级高速缓存，
并作为任务队列的可靠存储

图数据库

Neo4j：专注于复杂关系的知
识图谱存储与多跳查询能力

异步任务调度

Celery：分布式任务处理框架，
轻松支撑系统高并发场景

向量数据库

Qdrant：为AI应用提供高性能、
准确的高维语义向量检索

实时通信 & AI推理

WebSocket 双向推送 | ONNX
Runtime 统一推理运行时



PostgreSQL
稳定可靠 · 功能丰富



Neo4j
知识图谱 · 关系查询



Qdrant
向量检索 · AI原生



Redis
极速缓存 · 队列基石

基础设施层（下）



前端交互层

Vue 3 + TypeScript +
Pinia + Vue Router，构建
高性能响应式应用与状态
管理。



容器化部署交付

Docker + Docker
Compose 标准化容器交
付，确保开发、测试与生
产环境的高度一致性。



移动端跨平台

React Native + Flutter 双
栈支持，实现一套代码，
iOS / Android 双端原生
体验。



自动化流水线

GitHub Actions 构建
CI/CD 流水线，实现从代
码提交到自动测试、构建
与部署的全流程自动化。



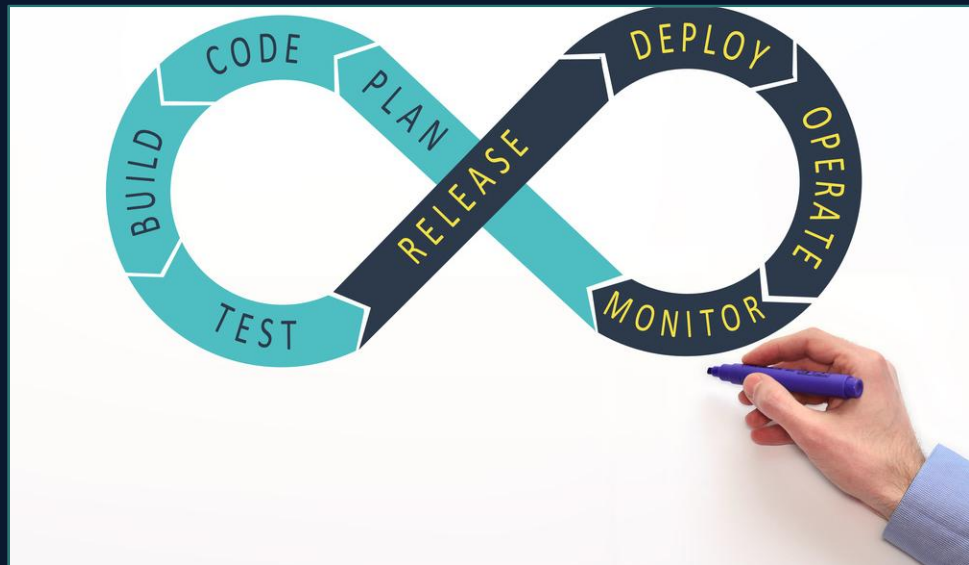
身份与权限认证

基于 JWT 的无状态鉴权
机制，支持多租户架构下
的细粒度权限隔离与安全
管控。



生产级服务治理

Nginx 流量分发 + PM2 进
程守护，基于 Ubuntu
22.04 稳定系统，保障服
务高可用。



代码架构：标准三层分离模式

- Endpoints (接口层)：负责处理 HTTP 请求路由与参数校验，解耦业务逻辑。
- Services (业务层)：封装核心业务规则与事务处理，实现逻辑复用与统一管理。
- Repositories (数据层)：抽象数据访问接口，屏蔽底层存储差异，确保数据持久化一致性。

表1.1 六大场景架构（上）



星熔炉Forge

Enlitecore · Forge

适用场景：

企业管理系统、工业数据看板、设备档案、标准业务交付

核心技术组合：

FastAPI + SQLAlchemy + PostgreSQL + Vue 3 + TypeScript

代表项目：

pipediag、高升创新设备管理、智练猿



星慧眼Vision

Enlitecore · Vision

适用场景：

AI视觉分析、目标检测与跟踪、医疗图像处理

核心技术组合：

YOLOv8 + SAM2 + Real-ESRGAN + ONNX Runtime + OpenCV + FFmpeg

代表项目：

轶德医疗AI手术显微镜、小人国人流分析



星河图Nexus

Enlitecore · Nexus

适用场景：

知识图谱推理、复杂多跳查询、科研数据挖掘、生物学AI

核心技术组合：

GraphRAG + Neo4j + Qdrant + Celery + Redis + FastAPI + WeasyPrint

代表项目：

RekonBio 药物靶点发现系统

表1.1 六大场景架构（下）

 星脉动Signal Enlitecore · Signal ■ 适用场景 多源数据融合、时序预测、市场情报、消费行为分析 ■ 核心技术组合 时序预测模型 + NLP热点分析 + 多平台API + 爬虫 + FastAPI + Vue 3 ■ 代表项目 · 海外电商海外选品 · 小人国游客倾向预测	 星法则Protocol Enlitecore · Protocol ■ 适用场景 LLM行为架构、多层协议系统、AI行为契约与审计 ■ 核心技术组合 HEP（H1-H8）+ DSIP + ABE + SDE + Flask + React Native ■ 代表项目 · PathCoach 职业发展平台	 星创世Studio Enlitecore · Studio ■ 适用场景 AI内容创作、视频生成、多模型路由调度 ■ 核心技术组合 Seedance 1.5 Pro + 首尾帧中继 + 多模型路由 + Flask + 五层导演架构 ■ 代表项目 · AIW-mbai AI视频创作平台
--	---	--

技术演进方向

Enlitecore 持续扩展技术边界，以下方向已纳入下一阶段架构规划：



TimescaleDB

时序数据库

工业传感器数据、游客流量时序分析，替代通用数据库处理高频时序写入



Elasticsearch

搜索引擎

企业知识库全文检索，与Qdrant向量检索形成关键词+语义双引擎



MongoDB

文档数据库

非结构化配置数据、导演架构Profile存储、动态Schema场景



Apache Kafka

消息中间件

高吞吐实时数据流，对应视觉感知系统的摄像头数据流与多源数据管道

CURRENT STACK

- PostgreSQL
- Neo4j
- Qdrant
- Redis

NEXT STAGE

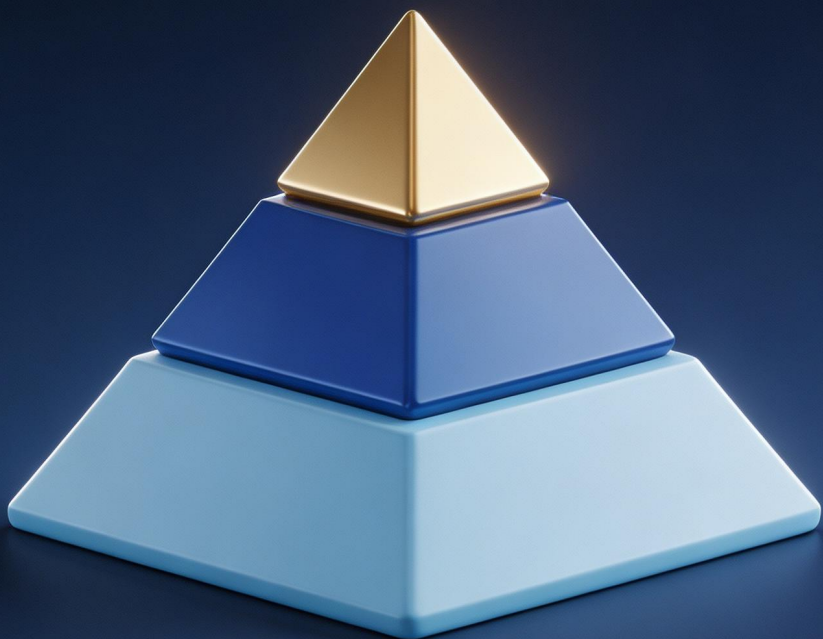
- TimescaleDB
- Elasticsearch
- MongoDB
- Apache Kafka



构建高可用、高吞吐、语义丰富的
下一代混合数据架构

第二部分：大模型工程能力 — 概述

博昇慧源AI研究院的大模型能力覆盖从日常应用到前沿研究的完整层次，既能与业界主流工具无缝接轨，又具备独立研发AI架构协议的深度能力，构建了从“应用落地”到“前沿探索”的全栈式能力体系。



日常应用层 | 与主流AI工具无缝接轨
能力覆盖：多模型调用、本地化私有部署、智能 Agent 系统搭建、多模态视频生成

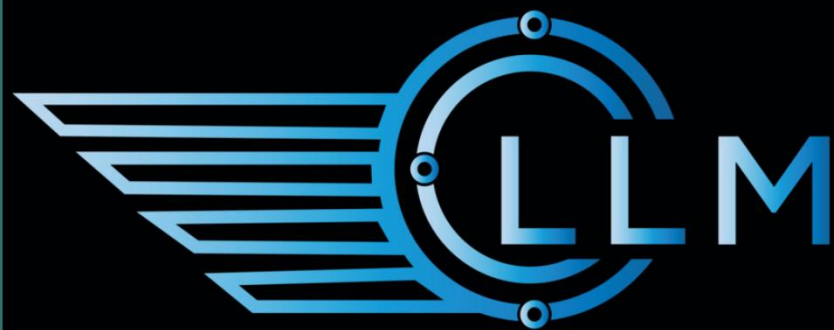


技术实施层 | 工程化落地生产级系统
核心能力：多模型动态路由调度、企业级 RAG 知识库构建、GraphRAG 推理引擎落地



前沿研究层 | 自研架构协议与学术探索
探索方向：HEP 行为契约理论、五步推理 GraphRAG 核心算法、追踪 arXiv 前沿论文复现

大模型能力：日常应用层



多平台调用与精准选型

熟悉国内外主流模型，可根据任务类型、成本与效果进行组合调用与选型。



本地化私有化部署

在本地GPU环境独立部署开源模型，确保企业核心数据安全，满足无网络场景需求。



AI智能体与 workflow编排

基于主流框架构建自动化Agent workflow，实现复杂业务逻辑的多步骤智能执行与闭环。



Prompt工程与指令设计

建立系统化的提示词设计方法论，适配各垂直行业，显著提升AI工具落地效果。



长视频生成与叙事控制

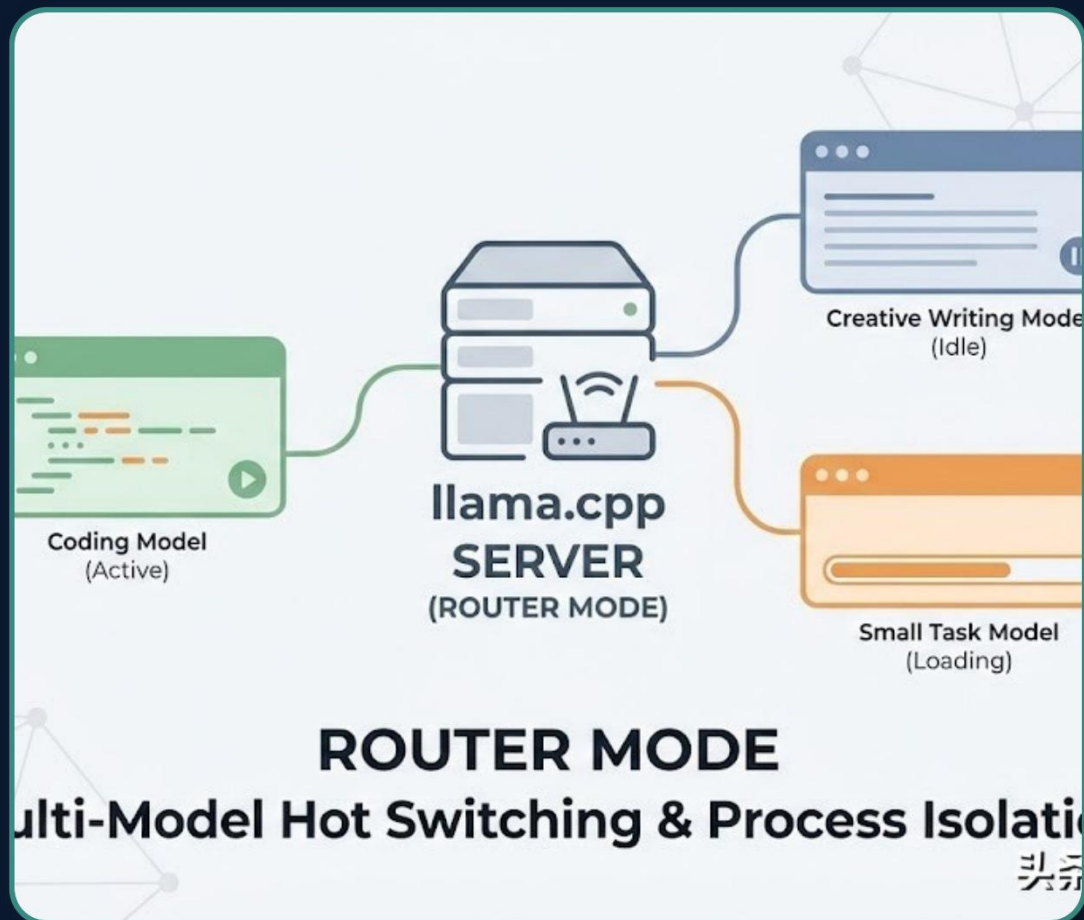
突破单次生成时长限制，实现多镜头、连贯叙事的长视频内容自动化生产。



全模态内容辅助创作

构建图文、音乐、视频全类型内容的AI生成 workflow，大幅提升内容生产效率与多样性。

大模型能力：技术实施层



多模型路由与私有化

根据任务类型动态调度多个大模型，兼顾性能与成本，支持完全离线私有化部署。

AI原生开发环境与协作

构建多AI协同工作的开发环境，以AI驱动方式完成全栈系统工程交付，提升开发效率与系统适应性。

RAG企业私有知识库

基于检索增强生成技术构建专属知识库，支持多租户物理隔离，确保数据安全边界。

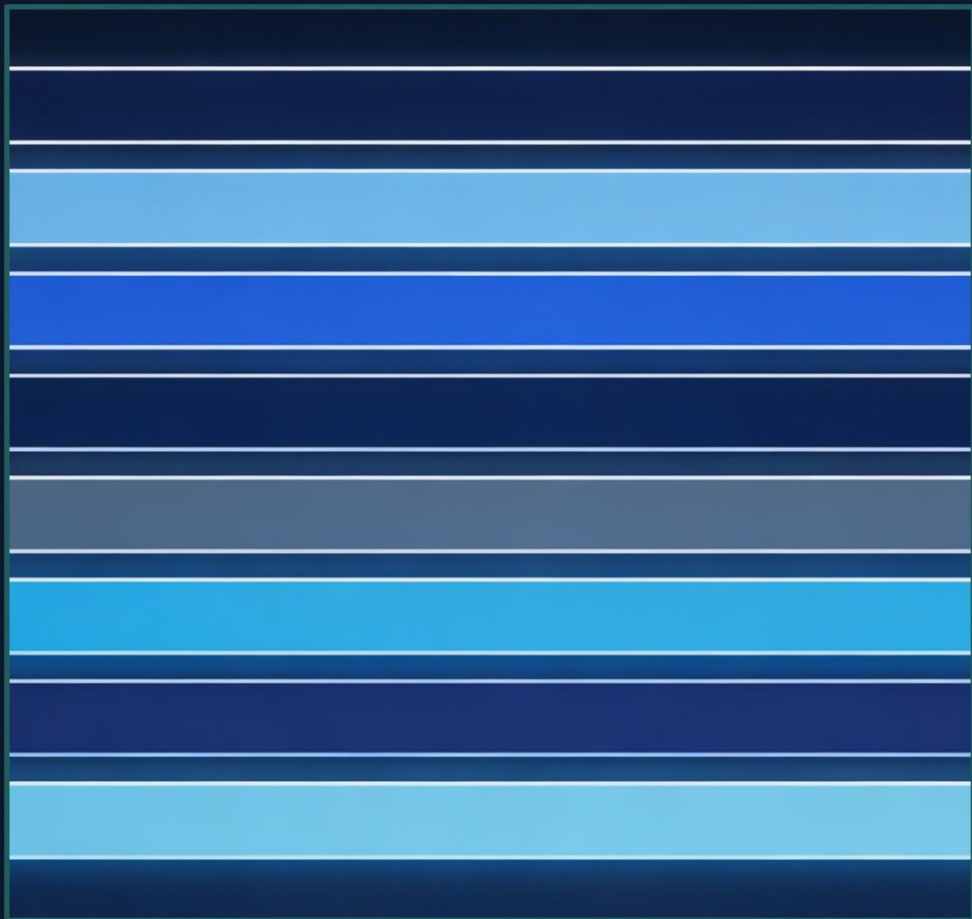
LLM结构化报告生成

将大模型的非结构化自然语言输出转化为标准化的业务PDF报告，直接对接业务端交付需求，减少人工干预成本。

GraphRAG推理引擎

将大模型与图数据库融合，实现复杂多跳关系推理，突破传统向量检索的语义局限。

大模型能力：前沿研究层



HEP 行为契约架构

自研八层行为协议(H1-H8)，通过确定性与随机性分离的设计原则，实现LLM驱动系统的行为边界可控与全链路可审计。



五步 GraphRAG 推理引擎

融合疾病-基因-蛋白-通路-文献知识图谱，自研五步推理与4维评分体系，实现复杂科研场景的自动化靶点发现。



工业AI大模型落地方法论

面向制造、医疗、生命科学等垂直领域，建立涵盖知识结构化、私有化部署与多模型协作的完整工程化落地体系。



学术发表方向

基于GraphRAG的工业级药物靶点发现系统研究成果，拟以研究院名义发表arXiv高水平学术论文。

第三部分：项目全景总览（上）



RekonBio

架构：星河图 | 领域：生物医药AI

从海量文献自动发现、评分潜在药物靶点，生成结构化PDF报告，赋能药企研发。

★ 核心：GraphRAG知识图谱 / 多跳推理 / 4维评分

轶德医疗显微镜 · 星慧眼

领域：医疗AI视觉 | 场景：眼科手术辅助

为国产手术显微镜叠加AI能力，实时识别标注手术结构，超分增强画质并自动生成术后报告。

★ 亮点：<50ms低延迟识别 / YOLOv8+SAM2双模型 / 4倍图像超分

小人国人流分析 · 星慧眼

领域：文旅视觉AI | 场景：景区精细化运营

融合摄像头与门禁数据，实时追踪游客动线，生成热力图，辅助景区客流管控与资源调度。

★ 优势：100+目标并发追踪 / 统计误差<5% / 热力图秒级刷新

小人国游客预测 · 星脉动

领域：文旅数据智能 | 场景：营销与决策辅助

整合电商、投流、新闻热点等5类异构数据，构建时序预测模型，精准预测游客消费倾向与未来30天客流趋势。

★ 价值：趋势捕捉提前2-3周 / 提升营销投放ROI

PathCoach · 星法则

领域：AI职业发展 | 场景：个性化成长规划

基于自研的HEP分层协议与行为契约机制，通过LLM驱动，为用户提供可解释、可审计的职业发展路径规划。

★ 技术壁垒：HEP八层行为契约引擎 / 决策链路100%可追溯

第三部分：项目全景总览（下）

 **智练猿** | 星熔炉 · 企业知识培训

用途：企业私有知识库培训平台，数据物理隔离，支持RAG问答与资产管理。
技术：RAG知识库 · FastAPI · SQLite · Vue 3
AI能力：企业专属RAG问答 · 零跨租户风险 · 私有化部署

 **电商选品** | 星脉动 · 跨境电商选品

用途：聚合Amazon/Shopee多平台数据，结合热点趋势为卖家推荐高潜力选品。
技术：LLM趋势分析 · 选品评分模型 · 多平台API · Vue 3
AI能力：4+平台同时接入 · 分析时间 天 → 分钟级

 **高升创新设备管理**

星熔炉 · 半导体工业设备

建立全生命周期档案，AI预测零件更换窗口，主动发现商机。

★ 商机主动前置发现

 **pipediag** | 星熔炉 · 工业管道检测全流程数字化管理

工业管道检测流程的全面数字化转型，实现AI辅助异常识别与自动化检测报告生成。

核心技术：FastAPI · PostgreSQL · Vue 3 · TypeScript · Nginx | **AI赋能：**报告生成 小时 → 分钟级 · 数据结构化率 100%

 **AIW-mbai** | 星创世 · 基于 Seedance 1.5 Pro 的AI视频创作平台

采用“首尾帧中继技术”突破生成限制，构建五层导演架构，实现连贯长视频的高效创作。

核心技术：Seedance 1.5 Pro · 多模型路由 · Flask · 五层导演架构 | **AI赋能：**单段5秒扩展至30-45秒 · 精细风格控制

第四部分：技术能力图谱 — 概述与全栈工程

博昇慧源AI研究院的技术能力覆盖从基础设施到AI推理的完整纵深，能够独立交付端到端的生产级AI系统，而非依赖外部集成商拼凑。三大能力支柱互为支撑：全栈工程层保障系统稳定可交付，AI/ML工程层决定智能化深度，数据工程层决定系统能“看到”多少真实世界。

■ 4.1 全栈工程能力：私有化部署与高可用架构

具备从后端API、数据库到前端交互界面的完整研发与交付能力。所有系统均可完全脱离公有云依赖，在客户本地服务器独立运行。采用成熟的异步任务架构支撑高并发场景，确保生产环境下的系统稳定性与任务调度的高效性。

核心技术栈概览：

后端框架：FastAPI, Flask • 数据库：PostgreSQL, SQLite, SQLAlchemy

前端交互：Vue 3 + TypeScript + Pinia + Vue Router • 任务队列：Celery + Redis

部署环境：Nginx + PM2 + Ubuntu 22.04 LTS



前端交互层 (Presentation Layer)

Vue 3 生态 / TypeScript / Pinia / Vue Router



API 网关与应用层 (Application Layer)

FastAPI / Flask 高性能后端框架



服务与逻辑层 (Service Layer)

SQLAlchemy (ORM) / Celery (异步任务) + Redis



数据持久层 (Data Layer)

PostgreSQL (生产环境) / SQLite (轻量级)



基础设施与部署 (Infrastructure)：

Nginx + PM2 + Ubuntu 22.04 LTS

技术能力图谱 — AI/ML工程 + 数据工程

4.2 AI / ML 工程能力 (核心壁垒)

具备从CV实时推理(50ms级延迟)、知识图谱多跳推理到LLM垂直领域落地的完整能力，构建高壁垒的智能化系统。

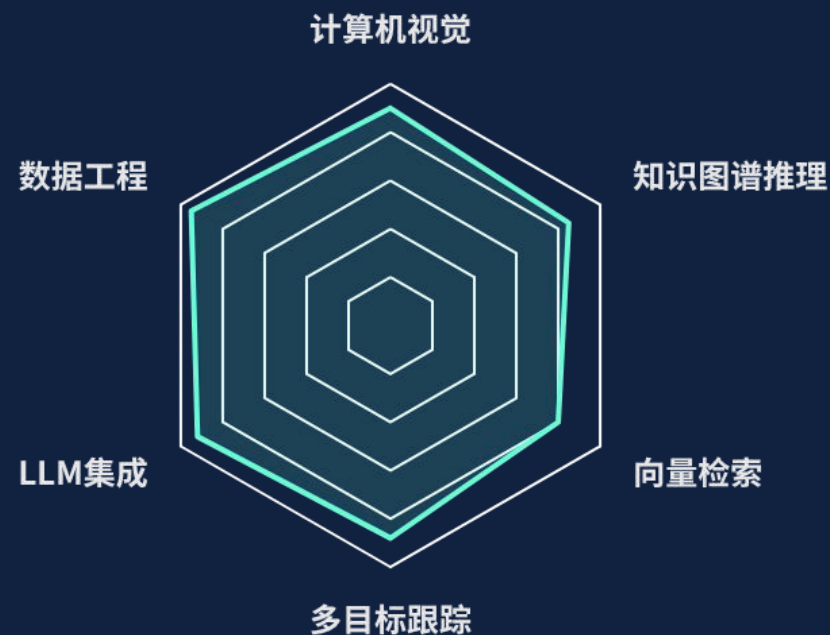
- 视觉: YOLOv8 / SAM2 / Real-ESRGAN / ONNX Runtime | 跟踪: DeepSORT
- 图谱: Neo4j / GraphRAG | 检索: Qdrant | LLM: 多模型路由 / 私有知识库RAG

4.3 数据工程能力 (信息视野)

整合多源异构数据，构建时序预测与NLP热点分析管道，将原始数据转化为可落地的业务洞察与结构化报告。

- 采集: 多平台电商API / 实时爬虫 / 生物医学库 (DisGeNET/PubMed等)
- 输出: 时序预测模型 / WeasyPrint 结构化PDF交付

技术能力维度分布



均衡发展的技术栈，实现从底层数据到上层应用的全链路覆盖

附件 A.1：企业设备数据库类



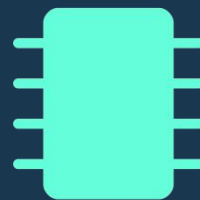
pipediag — 鼎业智能管道检测工业系统

客户：鼎业智能 | **架构：**Enlitecore · Forge（星熔炉）

技术栈：FastAPI + SQLAlchemy + PostgreSQL + Vue 3 + TypeScript

核心功能：管道检测记录、设备档案管理、AI辅助异常识别、报告自动生成

技术亮点：报告生成从“小时级”压缩至“分钟级”；支持多管线并行管理，数据结构化率达100%。



高升创新 — 半导体尾气设备管理系统

架构：Enlitecore · Forge（星熔炉） | **技术栈：**FastAPI + SQLAlchemy + PostgreSQL + Vue 3 + TS

核心功能：已售设备建档、零配件全生命周期管理、维保周期智能预警、商机主动发现。

技术亮点：支持预警时间窗口灵活配置，推动商机转化从“被动响应”转向“主动前置”，显著提升客户跟进效率。

附件 A.2：流量与消费分析类



小人国 — 人流分析系统

架构与技术: Enlitecore · Vision (星慧眼) | YOLOv8 + DeepSORT + 视觉与传感器融合 + 热力图生成

核心亮点: 支持同时追踪100+目标，视觉与传感器双源校验，统计误差率 <5%；热力图刷新间隔 ≤10秒，实现精准客流管控。

小人国 — 游客倾向预测系统

架构与技术: Enlitecore · Signal (星脉动) | 多源异构数据融合 + 旅游新闻NLP热点分析 + 长短期时序预测模型

核心亮点: 覆盖7/30天预测窗口，融合电商、投流等5类数据源；相比传统单源分析，趋势捕捉提前量提升2-3周，赋能精准营销。

附件 A.3：企业知识管理类



智练猿 — 企业知识培训平台

🔗技术架构 & 栈：Enlitecore • Forge (星熔炉) | FastAPI + SQLAlchemy + Vue 3 + TypeScript (SQLite物理隔离)

⚡核心功能：企业私有知识库 RAG 问答 / 员工培训管理 / 严格的多租户隔离体系

🌟关键亮点：每组织独立物理数据库，消除跨租户数据泄露风险；基于企业知识库的RAG回答准确率显著优于通用LLM，保障业务适用性。

PathCoach — AI 职业发展平台

Official Domain: pathcoach.ai

🔗技术架构 & 栈：Enlitecore • Protocol (星法则) | 自研 HEP (H1-H8) + DSIP + ABE + SDE | Flask + React Native 跨端开发

⚡核心功能：LLM 驱动的动态职业路径规划 / 智能行为契约执行 / 企业级多层协议管控

🌟关键亮点 & 研究：基于自研HEP八层协议实现对LLM行为边界的分层约束，确保AI决策全链路可追溯、可审计。目前聚焦研究方向：Behavioral Contract Enforcement in LLM-Driven Systems。

附件 A.4：电商选品

🏗️ 电商 海外选品系统

🔗 技术架构：Enlitecore · Signal（星脉动）

💻 技术栈：多平台官方API(Amazon/Shopee) + 补充爬虫 + 选品评分模型 + LLM趋势分析 + FastAPI + Vue 3

🌀 核心功能：多平台数据聚合、AI驱动选品评分、市场热点趋势分析

🚀 技术亮点：接入4+主流平台，LLM与结构化评分双引擎，分析耗时从“数天”压缩至“分钟级”。



轶德医疗 · AI手术显微镜系统

🏗️ 技术架构：Enlitecore · Vision (星慧眼) | YOLOv8 + SAM2 + Real-ESRGAN + ONNX Runtime

🌀 核心功能：图像增强(4倍清晰度) · 手术结构实时识别 · 视频关键帧分析 · LLM手术报告自动生成

⚡ 关键指标：推理延迟 <50ms (满足手术实时性要求)；YOLOv8+SAM2 协同，识别精度显著优于单模型。

附件 A.5（一）：RekonBio — 药物靶点发现系统

技术架构

Enlitecore • Nexus（星河图）
构建高性能的知识图谱底座，支撑复杂的生物学推理需求。

核心技术栈

GraphRAG + Neo4j + Qdrant + Celery
+ Redis + PostgreSQL + FastAPI +
WeasyPrint + Vue 3 + TypeScript

权威外部数据源

整合全球顶级生物医学知识库：
DisGeNET / PubMed / STRING /
Reactome

核心功能模块

- 五步GraphRAG推理引擎 & 4维评分体系
- 支持多跳关系推理与语义相似搜索
- 自动化生成专业PDF靶点分析报告

技术与效率亮点

- 融合“疾病-基因-蛋白-通路-文献”五类知识图谱，支持5跳复杂推理
- 效率质变：传统人工数周的靶点初筛工作，系统仅需数分钟即可完成并输出报告。

学术与科研价值

- 打造基于GraphRAG的工业级药物靶点发现系统标杆
- 相关技术成果已在整理中，拟投递并发表高水平学术论文，推动行业方法论发展。

附件 A.5（二）：RekonBio · Omics（星千寻）— 多组学靶点挖掘系统



技术架构与技术栈

架构：Enlitecore · Nexus (星河图)
栈：Scanpy + DESeq2 + TCGA/GEO接口 + AlphaFold DB + Neo4j + Qdrant + FastAPI + Vue3 + TypeScript



多维公开数据源

- TCGA：覆盖33种癌症类型，11,000+样本
- GEO：整合500万+基因表达样本
- 补充：EBI单细胞图谱、AlphaFold 2亿+蛋白结构



全流程分析功能

- 自动化scRNA-seq分析与细胞聚类
- 差异基因与通路富集分析，自动提取候选靶点
- 蛋白成药性预评估与多组学报告一键生成



技术核心优势

- 时间领先：从实验数据直接挖掘，较文献发表早1-2年
- 链条延伸：结合结构生物学，将分析从基因表达延伸至“可成药性”评估
- LLM赋能：将复杂生信结果转化为研发决策语言



与星河图深度协同

- 数据流转：星千寻输出的候选靶点，可自动导入星河图进行知识图谱验证与评分。
- 闭环构建：形成“数据挖掘-知识验证”上下游闭环，共同支撑AI药物发现平台。



商业与学术价值

目标客户：寻求靶点发现提速的药企研发部门，及需扩展技术能力的CRO公司。
学术规划：探索多组学驱动的新发现方法论，计划与星河图系统整合发表高水平论文。

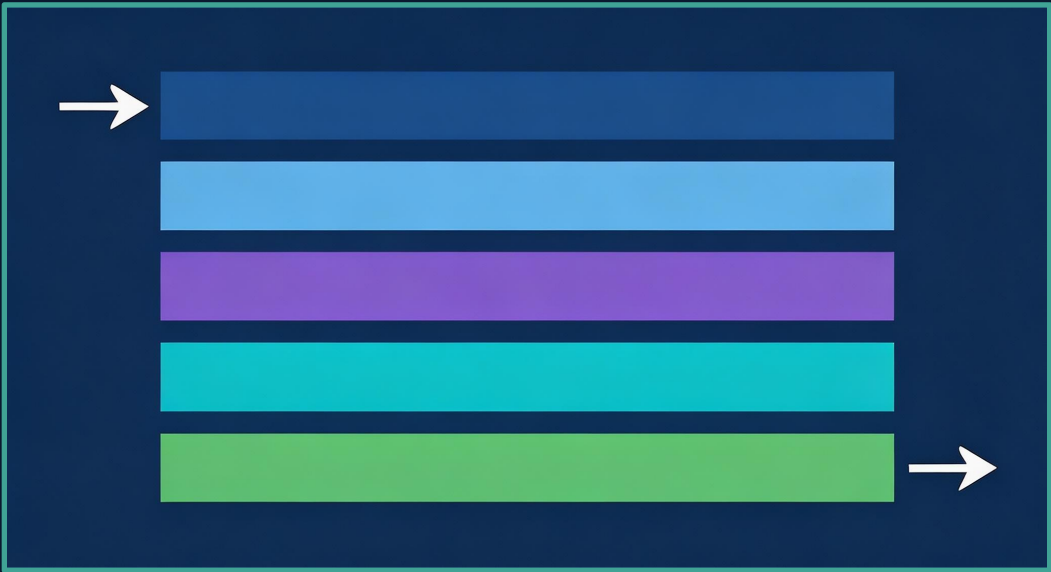
附件 A.6：AI创作类 + 结语

AIW-mbai — AI视频创作平台

技术架构：Enlitecore · Studio（星创世）

技术栈：Seedance 1.5 Pro + Flask + 五层导演/故事/角色/内容/生成Profile架构

核心能力突破：支持首尾帧中继生成、30–45秒长视频拼接与多模型路由。突破单次生成时长限制，实现单段5秒扩展至连贯叙事视频；五层架构实现对风格、角色、镜头语言的精细控制。



▲ 五层导演/故事/角色/内容/生成Profile 架构示意
实现视频创作的全链路可控性