

高能所机器学习研讨会 总结 – 大模型

李科

2024/10/18

大模型主题会场

同步辐射数据解析基座模型 C305, Main building	一鹤 庞 08:45 - 08:50
Dr. Sai物理分析智能体系统 C305, Main building	bolun zhang 08:50 - 08:55
智能值班助手 C305, Main building	Huangteng Hu 08:55 - 09:00
AI在星座控制里的应用及进展 C305, Main building	炼 陶 09:00 - 09:10
高能物理知识服务数据底座 & 智能科研助手 C305, Main building	法宇 姜 09:10 - 09:15
基于LLM的加速器智能问答机器人 C305, Main building	Diwei Chen 09:15 - 09:20
HEPS束线智能大模型 C305, Main building	一 张 09:20 - 09:25

不同中心的7个报告，大模型和多智能体

同步辐射数据解析基座模型

- 目标：以大语言模型为核心的同步辐射谱学数据解析

- 内容：

- 领域知识问答
- 线站信息问答
- 谱学数据解析
- 实验流程规划



研究方法

- 通用大语言模型提供基本人机对话服务
- 专用指令集微调(Supervised fine-tuning, SFT)注入领域术语知识
- 结合大语言模型(Retrieval-Augmented Generation, RAG)增强复杂问题回答及逻辑推理

- 现状：

- 基于LLaMa2微调->HEPS ChatBot-1.0
- 本地部署



通用大语言模型



领域大语言模型

同步辐射谱学数据解析模型

- 计划：

- 利用领域和线站知识提升模型性能
- 实验数据解析功能多样性
- 实验流程自动规划
- 提升数据解析精准度

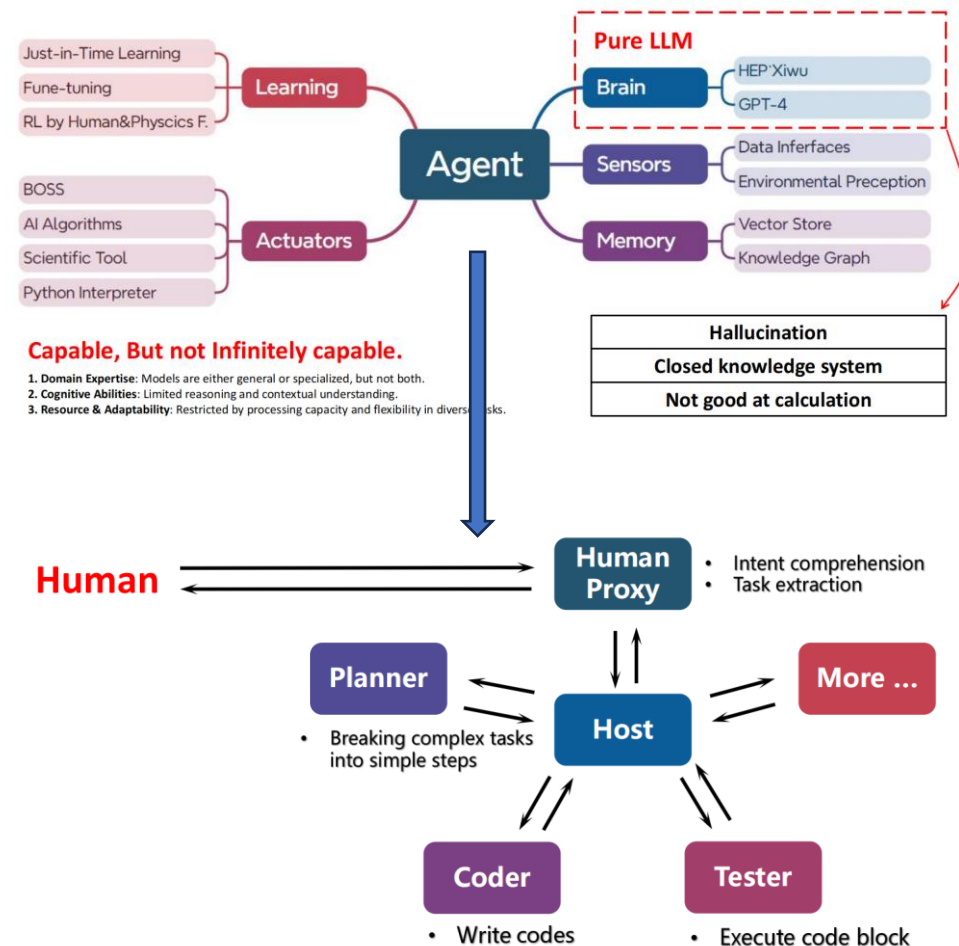
- 大量理论模拟数据预训练，挖掘信号关联映射模式
- 少量实验数据微调，实现端到端构效解析

Dr. Sai物理分析智能体系统

- 目标：粒子物理实验（BESIII）数据分析自动化系统
- 内容：
 - 领域大模型-溪悟
 - 多智能体系统
 - 分析流程自动化
- 现状：
 - 基于LLaMa3微调->Xiwu2
 - Dr.Sai 初始版正在BESIII内部测试
- 计划：
 - 升级为LLaMa3-70B
 - 利用思维链提升多智能体交互逻辑



多智能体系统

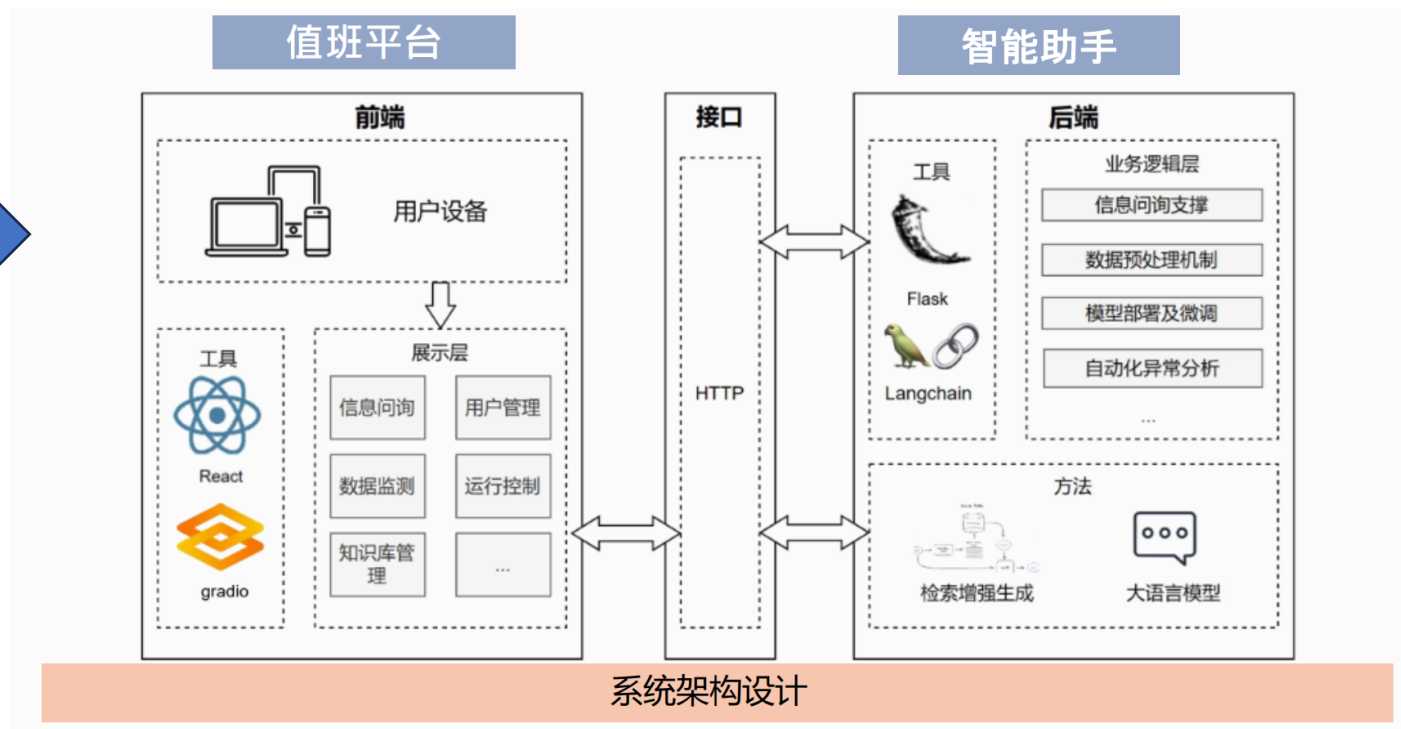


台中微子实验TAO智能值班助手

- 目标：远程稳定的智能化值班助手
- 内容：
 - 远程在线监控实验数据
 - 实验故障及异常自动告警
 - 智能化地分析各类异常
- 现状：
 - 基于Qwen2-7B
 - 知识问答功能
 - 自动化异常分析测试
- 计划：
 - 性能优化
 - 系统性评估



系统架构



AI在星座控制里地应用及进展

基于多智能体强化学习方法

- 目标：百颗卫星不间断、全天区、多目标、多参量观测

内容：

- 分布式星上自主决策方法
 - 各卫星共享策略

现状：

- 星上决策测试
 - 星上决策时间41秒，仅为地面~10倍，决策效率行业领先

计划：

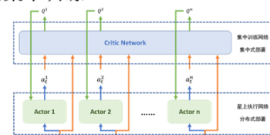
- CATCH星座在轨决策工具

观测规划的任务要求

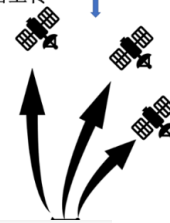
- 多个变源天体的**同时**观测
- 通过单颗卫星或者多星接力的方式，对变源天体进行**不间断**观测
- 尽量**避免**对星上已安排观测任务的**打断**
- 考虑后续可能的观测任务，**降低**后续观测任务对已安排观测任务的**打断概率**
- ToO (Target of Opportunity) 观测**响应速度**



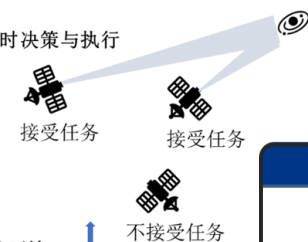
1、地面集中训练



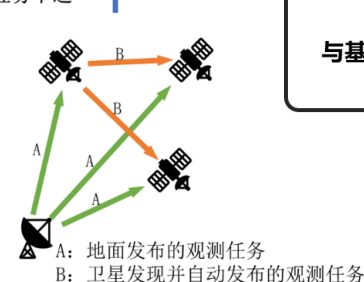
2、各星策略上传



4、分布式即时决策与执行



3、观测任务下达



效果

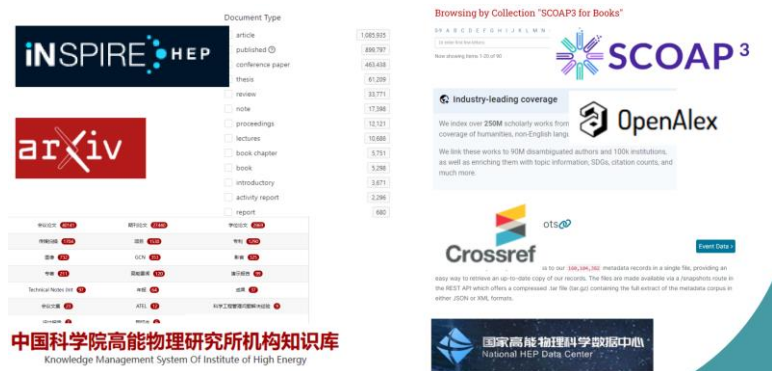
4.7%↑
整体收益
与基于协商的方法比较



高能物理知识底座&智能科研/文献助手

数据来源和自主平台建设

- 目标：助力科学研究
- 内容：
 - AI4S知识资源底座
 - 挖掘文献数据
 - 智能科研助手HaiAcademic
 - 文献情报智能助手



- 数据的分级管理
- 数据的大批量使用
- 新功能的加入
- 数据安全
- “以我为主”科学研究的支撑
- 国际合作话语权

智能科研助手



文献情报智能助手

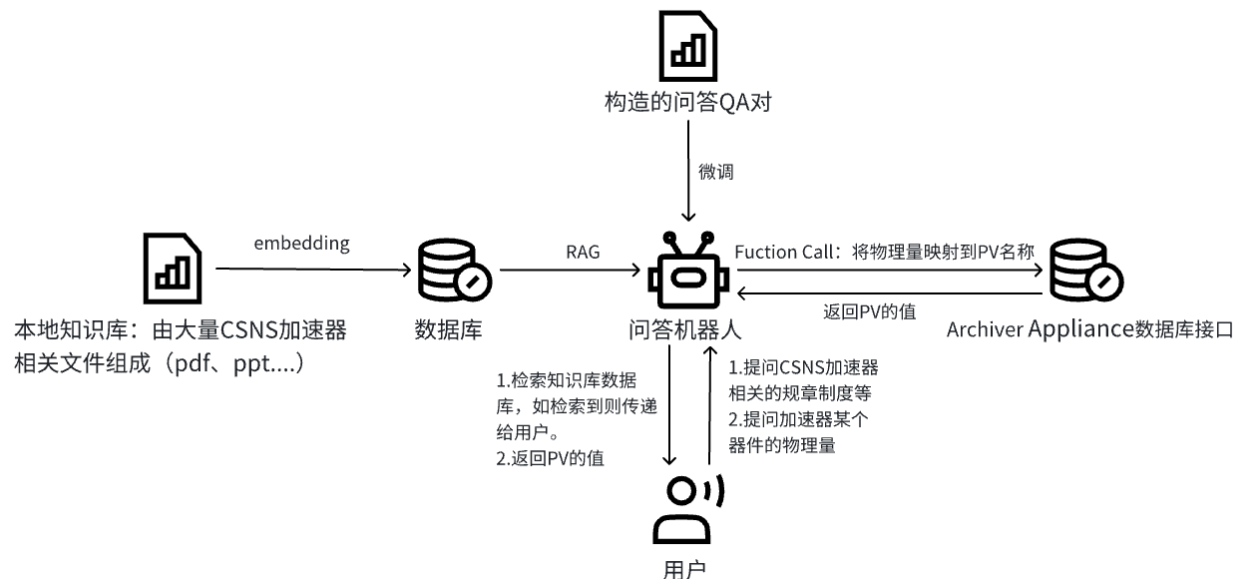


基于LLM的CSNS加速器问答机器人

- 目标：具有领域知识的CSNS垂域大模型
- 内容：
 - 针对操作规程、故障处理的问答
 - 针对PV值得问答
- 现状：
 - 初步收集高质量数据
 - 基于Qwen2-7B
 - 测试了增强检索
- 计划：
 - 继续收集高质量数据集（如有标记的PV值）
 - 使用QWen2-70B
 - 微调+增强检索
 - 多LLM打分，结合人工评判进行微调



系统架构



HEPS束线智能大模型

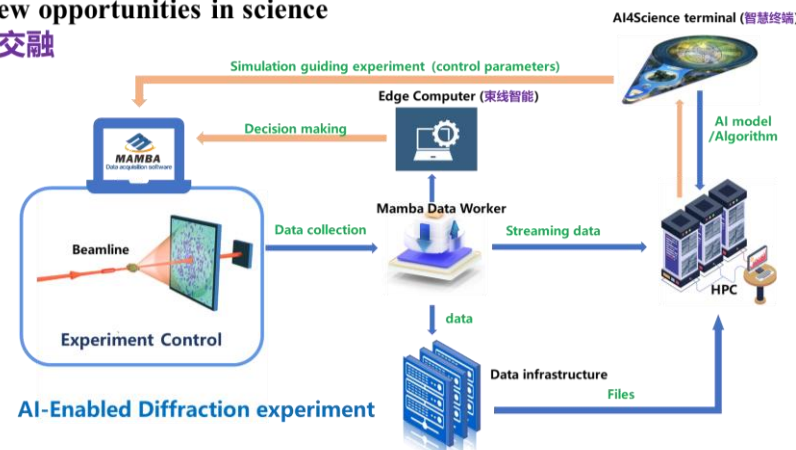
- 目标：光源控制的自动化和智能化
- 内容：
 - 新一代实验控制与数据采集软件系统Mamba
 - 智能调光
 - 流程控制
 - 科学预警
- 现状：
 - 开发Mamba多智能体
 - 配合增强检索提升准确率
- 计划：
 - 开发数据飞轮系统，不断迭代



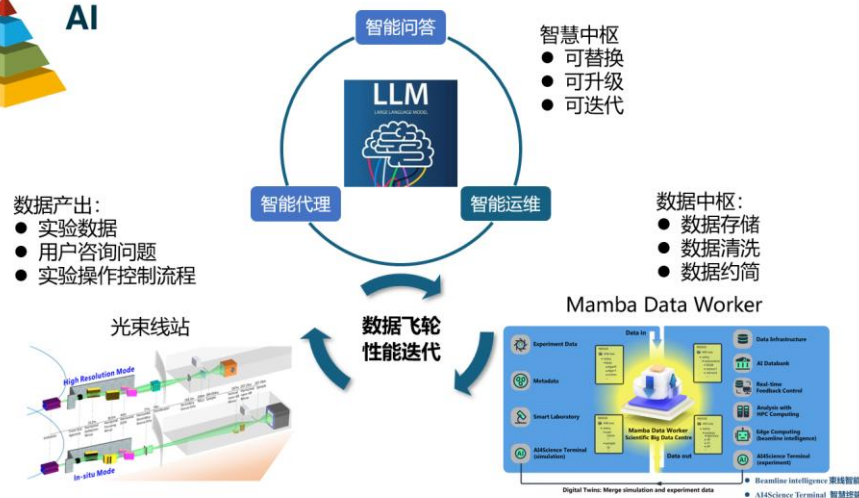
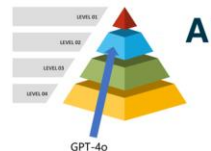
Mamba中多智能体协同系统

Open new opportunities in science

E: 理实交融



Fully leverage AI/ML and digital twin capabilities to extract information from big data streams, steer experiments, design experiments, and use on-demand data for ML-driven discovery



通用问题，个人思考和未来计划

- 领域知识如何嵌入
 - 利用高能所优势构造通用知识库 + 领域私有知识库？
 - 探索领域知识嵌入方式
- 实验数据如何解析
 - 特定实验数据开发专用解析模型？
- 实验流程如何建模
 - 从专家反馈数据中抽取思维链？

未来计划（仅供参考）：

- 探索高能所通用知识库
- 组织研讨会交流技术细节和经验
 - 大模型微调
 - 多智能体交互
 - 强化学习