

## AI 智能体辅助高能物理数据生产与分析项目介绍

一、 导师与团队基本信息	
导师姓名	丁雪峰
所在院所名称	中国科学院高能物理研究所
职称/职务	特聘青年研究员/副研究员
导师或课题组介绍网址:	<a href="https://dingxf.cn/">https://dingxf.cn/</a>
电子邮箱和联系电话	<a href="mailto:dingxf@ihep.ac.cn">dingxf@ihep.ac.cn</a> 18602755475
一级学科	物理学
二级学科	粒子物理与原子核物理
主要研究方向	中微子物理数据分析、深度学习和智能体开发
日常协助指导人	导师本人、博士后 2 名、博士生 1 名、硕士生 2 名

二、 项目基本情况	
项目名称	面向 JUNO 实验刻度与重建数据生产、验证及数据质量检查的科学智能体构建与应用研究
项目科学意义	江门中微子实验 (JUNO) 是我国重大基础科学装置, 其刻度、重建数据生产与数据质量检查是后续物理分析的重要基础。当前这些工作中仍包含大量重复性、流程化且依赖经验的操作, 如作业提交与查询、输出检查、基础 validation 和数据质量巡检等, 消耗较多

<b>二、项目基本情况</b>	
	<p>人力。大语言模型与科学智能体的发展，为复杂科研流程的自动化、标准化和知识沉淀提供了新的可能。</p> <p>本项目拟围绕 JUNO 分析基础任务，构建由任务规格、执行流程和技能库驱动的科学智能体，探索其在数据生产、验证和质量检查中的应用，为大科学装置智能化运行、科研效率提升以及青年学生交叉能力培养提供支撑，契合人工智能赋能科研的前沿方向。</p> <p>江门中微子介绍：<a href="http://www.ihep.cas.cn/dkxzz/juno/">http://www.ihep.cas.cn/dkxzz/juno/</a></p>
<b>使用的实验方法、仪器设备、数据软件</b>	<p>本项目以计算与数据分析为主，不涉及高风险实验操作。依托 JUNO 实验已有软件环境、实验数据和模拟数据，在 Linux 服务器/集群上开展研究。主要使用 Python、C++、ROOT、Git 等工具进行数据处理、日志解析、结果可视化和代码管理；使用 Markdown 文档沉淀任务规格（spec）、执行流程（runbook）和技能文档（skill）；结合大语言模型/科学智能体框架与 JUNO 计算环境、作业提交和查询工具，完成简单 production、基础 validation 以及数据质量检查等任务。</p>
<b>本科生研究任务</b>	<p>本科生将在导师和研究生指导下，参与以下具体工作：</p> <ol style="list-style-type: none"> <li>1. 阅读 JUNO 实验分析基础任务、数据质量检查以及科学智能体相关文献，完成简要综述和技术调研。</li> <li>2. 梳理刻度、重建数据生产和基础 validation 中的典型流程，整理任务输入、输出、判断标准和异常处理逻辑。</li> <li>3. 编写和维护相关 Markdown 文档，包括任务规格、执行流程和技能文档。</li> <li>4. 编写 Python 脚本，完成日志解析、文件检查、基础统计、绘图和自动化报告生成。</li> <li>5. 参与智能体原型测试，比较人工执行与智能体执行在耗时、正确率和稳定性方面的差异。</li> <li>6. 参与基础数据质量检查，如事件数、运行状态、关键物理量分布和异常样本排查等。</li> </ol>
<b>大致时间安排</b>	<p><b>6月：</b> 文献调研，了解 JUNO 实验分析基础任务和科学智能体基本概念，完成开发环境搭建。</p> <p><b>7-8月：</b> 梳理 1—2 类典型任务流程，完成初版任务规格、执行流程和技能文档，编写基础辅助脚本。</p> <p><b>9月-27年3月：</b> 在高能所或远程环境中开展集中实践，测试智能体完成简单作业提交、查询、结果检查和基础 validation。</p> <p><b>27年4-5月：</b> 参与数据质量检查案例分析，完善代码和文档，完成总结报告、答辩材料或展示海报。</p>

## 二、 项目基本情况

预期目标和成果形式	预期完成 1—2 类 JUNO 分析基础任务的智能化原型，形成一套可复用的任务规格、执行流程和技能文档；完成若干用于日志解析、基础统计和数据质量检查的脚本工具；形成项目总结报告、展示 PPT 或海报。表现优秀的学生可在后续毕业设计或长期科研训练中继续深入，并有机会获得课题组进一步推荐与支持。
实践地点	中国科学院高能物理研究所玉泉路园区（北京）

## 三、 对学生的要求与保障措施

拟接收人数	<=3 人，视情况而定
专业知识要求	具有基本的物理、数学和计算机基础；能够使用 Python 进行简单编程；具备基本英文文献阅读能力。有 Linux、Git、ROOT、数据分析或机器学习基础者优先，但不是必须条件。
工作时间要求	学期中建议每周投入 6—8 小时，能够参加课题组例会和阶段汇报；暑期如条件允许，可来高能所进行 2—4 周集中实践。项目过程中需按阶段提交代码、文档或总结材料。
课题组支持条件	课题组将提供必要的软件环境、服务器/集群使用条件、数据样本、代码仓库和文档模板；安排导师和研究生进行日常指导与进展跟踪；为学生提供参与真实科研问题训练的机会。对表现优秀、投入度高的学生，可在后续科研训练、毕业设计或推荐申请中给予进一步支持。