百度智能云 AI 应用赛竞赛规则

一、赛题内容

（一）赛题一：工业作业场地矿井的危险情况检测与分割

赛题内容：本赛题旨在利用大模型多模态技术实现工业作业场地危险情况检测与分割，即将给定的作业场地图像中的危险因素部分进行准确的分割检测。为此，利用赛会主办方指定的工业作业场景安全说明模版，参赛队伍需要设计基于危险情况检测算法逻辑和提示词模版，实现对特定工业作业场所的检测与分割。参赛者所设计的模型和提示词模版要能够较为准确地将场景中的危险画面进行分割，以实现精确的危险识别。本赛题主要考察参赛学生的业务逻辑理解、深度网络构建、算法逻辑、数据分析与可视化等方面的能力。

（二）赛题二：电网运行调度场景下的智能体设计

赛题内容：以电网运行调度值班场景为主题，利用大模型提升调度业务的自动化和智能化水平。本赛题建议按照三类业务场景应用智能体开发设计：

（1）电网运行数据智能监盘。结合电网运行数据进行微调和训练，实现数据的智能搜索和跨表关联分析；

（2）电网调度知识理解。结合电网调度规程、操作规程、事故预案等文档，实现电网调度运行专业知识的智能问答、异常处置、操作指引；

（3）专业大模型与应用小模型协同。通过智能体等技术，结合提示词工程、微调等技术，打通大小模型，实现对系统负荷预测等 AI 小模型的准确调用。

（三）赛题三：行业知识智能体应用

赛题内容：本赛题要求参赛者应用大模型进行智能体开发，开发一个垂直领域专家智能体，能够提供专业、精细的知识与帮助。智能体需调用专业知识库、数据库或插件，实现高准确度的信息检索和建议生成。应用的场景包括医疗健康管理智能体、文旅个性化推荐、智能客服等。

二、报名流程

1. 报名系统开放时间：6 月 20 日早 10 点；参赛选手登录比赛报名平台进行线上报名，报名网址：www.aicomp.cn；

2. 报名完成后即可在报名平台获取赛题数据集；

3. 参赛队长需使用在报名平台注册的手机号在百度智能云官网 https://login.bce.baidu.com 进行注册，获取百度智能云账号；

4. 报名赛题一的参赛选手通过已注册百度智能云账号在百度智能云一见官网中完成作品（https://cloud.baidu.com/product-s/yijian_home）；报名赛题二及赛题三的参赛选手通过已注册百度智能云账号登录百度智能云千帆 Appbuilder 官网

（https://qianfan.cloud.baidu.com/appbuilder/）完成比赛作品。

三、作品提交

（一）复赛

赛题一：参赛选手在百度智能云一见官网中完成作品后无需提交到报名官网，组委会将在百度智能云一见官网对参赛选手的作品进行评分。

赛题二&三：参赛选手需将完成的作品提交至报名官网（www.aicomp.cn），具体提交形式，以组委会通知为准。

（二）决赛

赛题一：参赛选手在百度智能云一见官网中完成作品后无需提交到报名官网，组委会将在百度智能云一见官网对参赛选手的作品进行评分。

赛题二&三：参赛选手除需将完成的作品及提交至报名官网外（官网网址：www.aicomp.cn），还需提交 PPT 功能演示视频，PPT 主要介绍智能体的设计思路，具体提交形式，以组委会通知为准。

四、赛题评分机制

（一）复赛

赛题一：作品提交后，组委会会根据参赛选手作品完成情况进行评分（评分表见附件 1），成绩发布分三个节点，分别为 7 月 31 日，9 月 15 日，10 月 26 日，最终以 10 月 26 日公布的成绩为最终成绩。成绩公布官网：www.aicomp.cn。

赛题二&赛题三：作品提交后，由组委会根据评分表统一评分（评分表见附件 2、3），确定最终成绩，成绩公布官网：www.aicomp.cn。

（二）决赛

赛题一：决赛阶段设置 1 个评分节点，为 11 月 15 日，相应评分完成后，将在比赛官网（www.aicomp.cn）进行成绩公布。

赛题二&三：决赛阶段参赛队伍需针对智能体的设计思路和功能演示进行讲解答辩，答辩成绩纳入决赛评分重要考量标准。

五、奖项设置

（一）复赛奖项设置

设立一、二、三等奖，数量分别不超过有效参赛队伍总数的 15%、25%、30%，颁发复赛获奖证书，复赛评分设置基准分数线，分数低于基准分数线的参赛队伍不予评奖，复赛一、二等奖晋级总决赛。（具体基准分赛前另行通知）

（二）总决赛奖项设置

设立一、二、三等奖，数量分别不超过总决赛队伍总数的 15%、25%、30%，颁发总决赛获奖证书。总决赛评分设置基准分数线，分数低于基准分数线的参赛队伍不

予评奖（具体基准分赛前另行通知）。

此外，总决赛排名前 10 的队伍有额外奖品，具体如下：

奖品	一等奖	二等奖	三等奖
获奖人数	1 支队伍	3 支队伍	6 支队伍
百度实习通关券	3 个/支队伍		
百度 BCP 认证在线学习券/ 工信部认证在线学习券 （等级任选）	3 个/支队伍		
度 BCP 认证考试券/ 工信部认证考试券 （等级任选）	3 个/支队伍		
百度云智大会门票	3 张/支队伍	3 张/支队伍	
百度研学营名额及证书	3 人/支队伍	3 人/支队伍	3 人/支队伍

六、增值服务

本赛题设置增值服务包（可自选）：包含产品技术支持、赛前集训认证、专属技术支持、赛中进阶训练营等，具体可加入赛题 QQ 群咨询。

附件：1. 赛题一评分表

2. 赛题二评分表

3. 赛题三评分表

附件 1

赛题一：工业作业场地矿井的危险情况检测与分割

评分标准

评分维度	指标描述	评分参考（十分制）
项目功能性	场景需求覆盖度	满分十分，权重 30%
	功能完整度	
	可演示性与可用性	
技术实现	视觉模型编排逻辑性有效性	满分十分，权重 35%
	数据处理与系统架构	
	技能使用模型的丰富程度	
创新与亮点	高级特性或新颖功能	满分十分，权重 15%
	业务/学术价值	
	用户体验与个性化设计	
团队协作	角色分工明确，贡献均衡	满分十分，权重 10%
	项目进度管理、合并质量	
	各人熟悉各自模块并能答疑	
答辩表现	PPT/演示条理清晰，能突出重点	满分十分，权重 10%
	回答评审问题的深度与准确性	
	时间控制与临场应变	
合计	总分=五项得分加权	五项得分加权

附件 2

赛题二：电网运行调度场景下的智能体设计

利用大模型开发自动评分工具，复赛环节主要利用工具评分，决赛环节采用工具评分 + 人工审核模式。总分为 250 分，评分标准如下：

评分标准

赛题场景	总分	评分说明	评分维度	评分标准
智能问数	100	问数场景，问题的答案为某个确定指标值，可通过要素提取后，对比指标值的一致性。	问数的数值结果是否正确。	单个指标进行 0/1 判分，多个指标求平均值。 单个问题请求三次取平均值。 按照满分为 100 分缩放实际得分。
智能问答	100	知识问答场景，问题的答案为一段文本，利用大模型从语义的相似性（Answer semantic similarity）和事实的一致性（Answer Correctness）维度进行评估。	对答案与参考答案进行语义级评分。 两个指标：回答正确性、语义相似性	回答正确性判分，采用大模型打分，分值范围为 0～1。 语义相似性判分，采用余弦相似度计算，分值范围为 0~1。 两个指标加权平均作为最终得分。 每道题评三次求平均值。 全部题目的平均分作为最终分数，并按照满分 100 分进行缩放。
智能体	40 + 10	问题的答案为某个确定指标值，可通过要素提取后，对比指标值的一致性。智能体运行效率也是考虑因素之一。混合多种任务和多种类型问题。	工具是否正确调用 数值结果是否正确 运行时间和效率	针对不同题目结合智能问数和智能问答两种评分规则，每道题分值范围为 0~1。 全部题目的平均分作为最终分数，并按照满分 40 分进行缩放。 全部参赛队伍全部题目平均用时排序后，按照最快 10 分最慢 0 分进行放缩。

附件 3

赛题三：行业知识智能体应用

赛题内容：本赛题要求参赛者应用大模型进行智能体开发，开发一个垂直领域专家智能体，能够提供专业、精细的知识与帮助。智能体需调用专业知识库、数据库或插件，实现高准确度的信息检索和建议生成。应用的场景包括医疗健康管理智能体、文旅个性化推荐、智能客服等。

评分标准

复赛仅从 “项目功能性”评分维度进行评分，决赛从下面 4 个维度进行综合评分

评分维度	指标描述	评分参考（十分制）
项目功能性（权重 40%）	场景需求覆盖度	（1）智能体回答与标准答案语义完全一致 +4 分 （2）智能体回答与标准答案仅部分一致 +2 分 （3）智能体回答答案语句通顺，无语病 +2 分 （4）智能体回答答案语言简洁不啰嗦 +2 分 （5）智能体回答答案语言正向、温和、亲切+2 分
	功能完整度	
	问答准确性	
	可演示性与可用性	
技术实现（权重 35%）	大模型调用与优化程度（意图补全 Prompt 和 答案总结 prompt）	（1）包含 意图补全提示词且提示词多轮会话效果准确 +2 分 （2）答案总结提示词总结逻辑覆盖全面 +2 分 （3）知识文档做了预处理 +2 分 （4）知识文档切片考虑了语义切分 +2 分 （5）检索召回考虑了 混合检索和 rerank 二次排序 +2 分
	知识数据预处理与文档切片合理性	
	检索召回技术考虑全面	
创新与亮点（权重 15%）	高级特性或新颖功能（多模态、微调、RAG 等）	**技术创新性（4 分）** ●4 分：突破性技术（如新型算法、跨领域技术融合） ●2 分：显著改进现有技术（如效率提升≥50%）

	业务/学术价值	●1 分：小幅优化或常规技术升级 ●0 分：无技术创新 **应用场景独特性（4 分）** ●4 分：开拓全新领域或解决未被满足的痛点 ●2 分：在现有场景中提供差异化解决方案 ●0 分：场景普通或缺乏新意
	用户体验与个性化设计	**用户交互体验（2 分）** ●2 分：自然流畅、情感化交互（如多模态感知、个性化响应） ●1 分：基础交互功能完善但缺乏亮点 ●0 分：交互生硬或存在明显缺陷
答辩表现 （权重 10%）	PPT/演示条理清晰，能突出重点	**逻辑结构（2 分）** ●2 分：框架清晰（背景-方法-结果-结论），层次分明 ●1 分：结构基本合理，但部分环节衔接生硬 ●0 分：逻辑混乱，缺乏系统性
	回答评审问题的深度与准确性	**表达能力（4 分）** ●4 分：语言流畅、用词精准，能脱稿自然陈述 ●2 分：表达基本清晰，但依赖稿件或偶有卡顿 ●0 分：语言模糊、频繁停顿或超时严重
	时间控制与临场应变	**PPT 设计与展示（1 分）** ●1 分：视觉简洁、重点突出，图文配合得当 ●0.5 分：PPT 基本可读，但设计粗糙或文字过多 ●0 分：PPT 混乱或影响信息传达 **问答环节应答（2 分）** ●2 分：回答准确、有条理，能延伸讨论或承认不足 ●1 分：回答基本正确但缺乏深度 ●0 分：回避问题或答非所问 **时间控制（1 分）** ●1 分：严格控制在规定时间内（±5%以内）

		●0.5 分：超时或提前结束但不影响整体 ●0 分：严重超时或时间分配失衡
合计	总分=各项得分加权	各项得分加权