

【赛事详解】“Agent开发哪家强？”——胜算云国产多模型实战PK大赛正式启动！

这不是一场简单的编程马拉松，这是一次对国产大模型应用能力的“全身体检”。

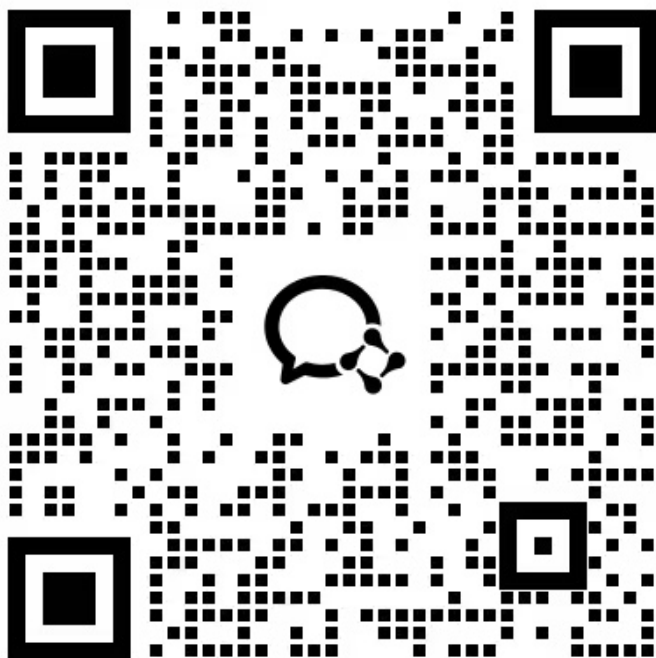
当所有人都使用相同的题目、相同的环境，唯一的变量是你选择的模型和你构建的Agent架构时，谁才是真正的“六边形战士”？



报名问卷地址：

<https://tcnqigfvmci2.feishu.cn/share/base/shrcnH566namZrQV3RpbyEpqp5T>

扫描加入大赛交流群



扫描添加客服领取报名福利



01 赛事背景与初衷

Background & Vision

2025年被称为“**Agent元年**”。我们见证了LLM从简单的对话（Chat）走向了复杂的任务执行（Action）。然而，在实际的工程落地中，开发者们面临着巨大的选择困难：

- **模型迷雾**：国产大模型百花齐放，哪一款能真的在推理、规划、LLM生成、反思、逻辑推理、工具调用上真正具备实战能力？
- **成本黑洞**：在完成相同agent执行命令任务时，如何平衡 Token 消耗与产出质量？
- **稳定性难题**：Demo 里的“神操作”，能否在生产级或高并发环境中稳定复现？

在大模型应用爆发的今天，Agent（智能体）已成为连接LLM与真实世界的关键桥梁。但面对百花齐放的国产大模型，究竟哪一款最适合构建复杂的Agent？在成本、效率与稳定性之间，开发者又该如何权衡？

胜算云联合Aitocoder、Mixlab无界社区、奇绩创坛等多家顶尖技术社区与创投机构，发起「**多模型PK · Agent 开发挑战赛**」。我们拒绝纸上谈兵，邀请全国开发者基于单一国产模型，在统一的评测框架下进行实战开发，角逐最强Agent开发者的桂冠！

02 核心玩法：单模型极限挑战

Core Gameplay

不同于以往允许“混合调用”或“外挂最强模型”的比赛，本次大赛制定了严格的“**单模型约束（Single-Model Constraint）**”规则：

1. **绑定机制**：参赛队伍在报名阶段需从指定的国产大模型列表中（如：Qwen系列、DeepSeek系列、GLM系列等）选定**一款**作为Agent运行的主力模型。
2. **全程单一**：在比赛的三道赛题中，你的 Agent **所有的** 推理、规划、LLM生成、反思步骤，都通过我们提供的免费额度key调用该特定模型的 API。
3. **禁止套娃**：严禁在后端偷偷转发调用 GPT-4 或 Claude 等非指定模型。我们将通过 测试 API 调用日志与代码审计进行核查。

为什么这么定规则？

我们要逼出国产模型agent应用的极限——通过优秀的 Prompt Engineering（提示词工程）、RAG（检索增强）、CoT（思维链）设计以及鲁棒的错误处理机制，弥补模型本身的短板，实现“小模型办大事”或“国产模型超越SOTA”的效果。

限制与自由？

1. 核心模型限制：

Agent 的核心逻辑（包括推理、规划、文本生成、代码生成）**必须且只能**调用参赛时选定的那一款国产大模型。严禁在后端“套壳”调用其他 LLM（如 GPT, Claude 等）。

2. 多模态调用自由：

Agent 在执行任务时，可自由接入**图片、视频、音频**等非 LLM 类的多模态模型作为工具。

- 胜算云模型市场已聚合大量多模态接口，参赛者可直接调用，无需重复造轮子。

3. 编程辅助不设限：

参赛者在编写代码的过程中，可自由使用各类 AI 模型辅助编程（部分题目限定编程工具）。另外我们将为参赛者开通所选主模型的免费额度，欢迎大家在开发过程中体验其代码生成能力。尝试用你选择的主力模型来辅助写代码。

题目设置：三大实战关卡

比赛设置三道统一题目（题目及评分详情见下），所有参赛队伍在相同的环境、评测标准下进行解答。

评分维度：全方位评估



不仅仅看结果，更看重工程化落地能力。评分将综合考量：

-  **任务完成度**：Agent 是否解决了问题？
-  **效率与性能**：响应速度如何？
-  **成本控制**：Token 消耗与模型性价比。
-  **稳定性**：在多次运行中的表现是否一致。
-  **可观测性**：决策链是否清晰可追溯。

03 赛程安排与奖项设置

赛程时间线Timeline

从报名到决赛，我们为你准备了充实的5周开发之旅，包含技术大牛直播、上手实战营与路演对接。

-  **报名预热期：12月14日 - 12月26日**
 - 12月14日GDPS全球开发者先锋大会首发开启报名。
 - 全网报名开启，组队招募。
-  **启动与上手营：12月26日 - 1月2日**

- **Kick-off 启动会：**赛题详解与参赛规则说明。
 - **7天开发上手营：**每日任务打卡 + 技术大牛晚间直播教学，助你快速跑通流程。
 - 🖥️ **主力开发周：1月2日 - 1月9日**
 - 高强度开发阶段，实时反馈开发难点，专业导师帮助攻克。
 - 优秀选手技术分享会。
 - 🏁 **冲刺与评审：1月9日 - 1月16日**
 - 开发终稿冲刺，提交最终作品。
 - 专家组进行投稿复核。
 - 🏆 **决赛路演：1月16日 - 1月23日**
 - 决赛线上路演 Demo Day。
 - 颁奖典礼 & 招聘/投融资对接会。
-

04 报名奖励与资源支持

Registration Rewards & Resources

为了让每一位开发者都能无后顾之忧地尽情创造，我们准备了丰厚的“军火库”支持。**越早报名，福利越多；组队出征，优势更大！**

🎁 个人报名福利

只要报名成功，即可解锁以下四重好礼：

1. 💰 **30元 平台算力金：**直接发放到胜算云账户，可用于比赛之外的模型调用体验。
2. 🗝️ **开发周无限畅用权：**比赛开发期间，指定主力模型免费调用（详情见下文）。
3. 📖 **独家技术手册：**获赠《Agent 开发独家实战手册》PDF，内含高阶 Prompt 技巧与案例解析。
4. 💬 **专家答疑直通车：**加入官方专属技术交流群，技术大牛在线解惑。

🚀 组队报名福利

推荐！2-4人/队，抱团作战优势更强

1. ⚡ **抢跑优势：**全队模型免费权限**提前 2 天**开启，先人一步调试 Prompt！
2. 👑 **队长双倍激励：**队长将获得 **双倍算力金 (60元)**，致敬带队领袖。
3. ❤️ **全员同享：**所有队员均可获得上述“个人报名福利”中的全部内容。

Free Model Access Policy

为了支持高强度的开发测试，我们将为参赛者开通胜算云主力模型的**免费使用权限**（不消耗账户余额）。

- **开启时间:**

- [illegible]

- **账号规则：**

- 免费权限将开通至报名时填写的**主账号**（团队参赛请指定队长账号）。
- **支持分Key**：团队可共用该主账号，通过创建多个 API Key（分Key）供不同队员同时调用，协作无障碍。

- **风控限额：**

- 为防止程序死循环或恶意刷量，免费权限设置了 **每日 200元** 的消耗上限。
- 注：该额度足以支持极高强度的开发调试，请放心使用。如需增加用量，请联系胜算云客服说明情况。



05 奖项设置

Awards & Prizes

本次大赛不仅提供丰厚的现金与算力奖励，更致力于为开发者链接**顶尖创投机构与大厂就业机会**。

 金奖 (1名)




——授予全场综合评分最高的“最强 Agent”

- **现金大奖：**5000 ¥ 现金/队
- **算力/Token包：**价值 5000 ¥ 的高额模型调用额度
- **荣誉认证：**金奖实体奖杯 + 电子证书+胜算云推荐信
- **超级权益：**
 -  **奇迹创坛（MiraclePlus）** 等顶级VC直通车
 -  三家联合举办官媒的**顶级曝光**与作品展示

-  合作企业**免试直通卡**（面试直通终面）



银奖（3名）

—— 授予在技术实现与工程落地表现卓越的队伍

- 现金奖励：2000 ¥ 元/队
- 算力/Token包：价值 2000 ¥ 的模型调用额度
- 荣誉认证：银奖实体奖杯 + 电子证书+胜算云推荐信
- 核心权益：
 -  三家联合举办官媒的**顶级曝光**与作品展示
 -  投资机构项目案头推荐
 -  合作企业**优先内推资格**


铜奖（5名）


—— 授予具有优秀创意与完整度的参赛队伍

- 现金/礼品：¥[500] 元 或 等值极客装备
- 算力/Token包：价值 ¥[500] 的模型调用额度
- 荣誉认证：铜奖电子证书+胜算云推荐信
- 核心权益：
 -  三家联合举办官媒的**顶级曝光**与作品展示
 -  优秀开发者人才库入库，合作企业简历优先推荐

特别创新奖（1名）

—— 授予最具创新思维、商业潜力或在某一技术点有突破的“黑马”

- 极客大礼包：[如：显卡、键盘、现金奖励]（国产显卡合作）
- 额度：额外2000胜算云算力额度
- 荣誉认证：最具潜力奖电子证书
- 特别权益：
 - **奇迹创坛**等创坛顶级VC直通车
 -  全球开发者先锋大会等渠道的**顶级曝光**与作品展示

-  合作企业**免试直通卡**（面试直通终面）
- 获得技术专家/投资人 **1对1 深度Mentorship（指导）** 机会一次

参与奖（完赛即得）

—— 致敬每一位坚持到最后的开发者

只要提交完整作品并通过基础审核（跑通测试），即可获得：

- **完赛证书**：官方认证的电子完赛证明
- **阳光普照礼**：200胜算云平台额度券（可覆盖后续开发成本）
- **社群权益**：加入“Agent 开发者核心群”，长期获取前沿技术资料

补充说明

1. **奖金发放**：所有现金奖励均为税前金额。
2. **开源要求**：获得金、银、铜及潜力奖的队伍，**必须**保持项目代码在 GitHub 开源（遵循 MIT/Apache 2.0 协议），否则将取消获奖资格。
3. **作品归属**：参赛作品知识产权归参赛团队所有。
4. 特别创新奖由技术大牛和创投机构老师评选，可与银、铜、参与奖同时获选

06 赛题设置与评分标准

Tasks & Scoring criteria

本次比赛采用**统一评测框架**，确保公平性。

题目设置

- **Level 1：胜算云出题-智能客服监控Agent**
 - 题目详情文档：[📖 智能客服监控 Agent](#)

场景：智能问答 (RAG) & 状态监控

目标：异常告警 -> 文档同步

难点：问答与监控数据的实时联动

- **Level 2: Altocoder专属题目-小红书RedNote-Agent[本题仅限在AltoCoder编码工具中开发]**

- 题目详情文档：[📖 AitoCoder赛题](#)

场景：营销种草

目标：搜索 -> 分析 -> 写作闭环

工具：必须使用 AltoCoder

- **Level 3: Mixlab出题 - 产品经理agent**

- 题目详情文档：[📖 产品经理Agent：用户痛点和应用场景挖掘，输出需求文档](#)

场景：需求分析

目标：痛点挖掘 -> PRD输出

难点：逻辑推理 & 规划

07 作品提交物料清单详解

Submission Checklist

为了确保你的作品能被正确评测，请严格遵守以下提交规范：

1. 代码仓库 (GitHub Repo)[必交]

- 参赛作品的知识产权归参赛者（或参赛团队）所有。所有参赛作品须遵循开源协议（建议 MIT 或 Apache 2.0），并托管于 GitHub 公开仓库中。
- 结构要求：
 - `/agent`：核心逻辑代码。
 - `/requirements.txt`：依赖包清单。
 - `/main.py` (或入口文件)：启动入口。
- **复现性**：评委将 Clone 你的仓库并在全新环境中运行，请确保没有硬编码的本地路径。

2. 评测输出 (Outputs)[必交]

- 对应三道赛题，需提交运行后的结果文件（如 `result_task1.json`，`generated_code_task3.py` 等）。

3. 演示视频 (Demo Video) [必交]

- **时长**：3分钟以内。
- **内容**：
 - **0-30s**：一句话介绍你的 Agent 架构与所选模型。
 - **30s-2m**：实录 Agent 运行过程（可加速播放），展示关键的思考/工具调用步骤。
 - **2m-3m**：展示最终成果及亮点（如：极低的成本、独特的自我修正机制）。

4. 技术报告 (Technical Report)[可选] - 加分项

- **PDF格式**，建议包含：
 - **agent架构设计**：你用了哪个agent架构？有什么调整的地方？
 - **Prompt 策略揭秘**：你如何写 Prompt 让模型更听话？
 - **Bad Cases 分析**：开发过程中遇到了哪些坑？是如何解决的？
 - **模型评价**：基于实战，客观评价你所选模型的优缺点。

通过一份清晰技术报告PDF文档，能够快速让评审老师了解你的agent项目。

08 评审机制与评审团队

Evaluation & Jury

本次大赛采用“专业视角 + 大众共识”的双维评审机制，既确保技术硬核与商业落地的专业性，也充分考量产品的市场热度与用户喜爱度。

评分权重

总分 (100%) = 专家评审 (60%) + 公共投票 (40%)

1. 专家评审维度（60分）

由技术大牛与投资人等专家组成的评审团，将从以下四个维度进行打分：

- **参赛完成度 (5分)**：
 - Demo 是否可流畅运行
 - 功能是否完整，而非仅仅是一个概念。

- 是否提交了技术报告PDF
- **技术实现与工程质量 (25分):**
 - Agent 架构设计的合理性与稳定性。
 - 对所选国产大模型能力的挖掘深度（Prompt Engineering/Fine-tuning 技巧）。
 - 代码规范性与开源完整度。
- **创意创新性 (15分):**
 - 切入点是否新颖，是否解决了未被满足的痛点。
 - 是否展现了 Agent 独特的交互方式或解决思路。
- **商业潜力与应用价值 (15分):**
 - 应用场景的市场规模与变现潜力。
 - 是否具备转化为实际创业项目的可能性。

2. 公共投票维度（40分）

Public Voting (40 Points)

我们将开启全网公众投票通道，将产品的评判权交给市场和用户。最终得分完全取决于**得票数排名**，排名越高，得分越高。

- **投票机制:**
 - 所有入围作品将在大赛官网/指定H5页面进行公开展示。
 - 用户需登录/验证后方可投票，以确保真实性。
- **计分规则（按排名赋分）：**
 - 🏆 **第 1 名：40分（满分）**
 - 🥈 **第 2 名：38分**
 - 🥉 **第 3 名：36分**
 - 📊 **第 4 名及以后：动态梯度计分**
 - 依据最终提交作品的总数量设定分值递减梯度（Step）。
 - 最低保底分10分

⚠️ 严正声明:

组委会将部署严格的**反刷票监控系统**。一旦发现使用机器刷票、购买僵尸粉投票等作弊行为，将直接**清空该项目的所有票数**，并取消其评奖资格。

我们邀请了来自**技术界、创投界、学术界**的领军人物组成“超级评审团”，为参赛者提供最顶级的反馈与机会。

创投与商业导师 (VC & Business Mentors)

负责评估商业模式、技术市场潜力，并发放“直通车”资格。

- **[周政军] | [胜算云] 创始人 / CEO 连续成功创业者、操作系统及云计算专家**

曾主导100万并发支撑能力的大规模GPU集群，共享式GPU渲染计算任务调度系统的开发，并成功商用。胜算云是根据过往成功经验积累将此多云并发GPU计算技术应用于大模型推理计算中。中星微初创团队核心成员，“星光中国芯”工程获国家科技进步一等奖。

- **[周雪] | [胜算云] 投融资总监VP**

擅长资源整合及PR媒体传播，统筹企业形象建设。具备丰富投融资沟通经验，擅长政企合作，项目申报。

技术与工程专家 (Technical Experts)

负责评估代码质量、模型应用深度及技术架构。

- **[shadow] | [Mixlab无界社区创始人] 程序员、设计师、连续创业者**

毕业于上海交通大学、同济大学蚂蚁集团高级前端架构师，负责服务日活上亿用户的支付宝智能UI引擎，数字人创新项目。Mixlab无界社区创始人ComfyUI中文爱好者社区发起人(10w+用户)

- **[陈明] | [Aitocoder] 智码酷达金牌教练**

企业架构师，《AI编程宣言》联合发起人，深耕软件与AI领域多年，兼具技术实战沉淀、权威资质背书与教育传播经验，资深AI编程教育者，出版《秒懂AI编程》一书。

- **[陶文龙] | [胜算云] 技术解决方案专家**

智能体架构与大模型全链路专家，专注于高性能、可扩展的智能体系统设计。具备从大模型开发、训练至部署调优的完整实战经验，善于在复杂业务场景中推动技术落地。始终关注前沿技术与产业融合，致力于促进智能体技术的创新与生态构建。

学术与生态顾问 (Academic & Community)

- **[谭中意] |**

开源专家、《AI编程宣言》发起人、中国开源推进联盟常务副秘书长、开放原子开源基金会TOC主席、国际内源基金会董事、《DeepSeek生态深度探索》主编

- **[薛志荣] | 设计开源社区发起人设计师、程序员**

华为人机交互实验室人机交互研究员、小鹏资深交互设计师、百度高级交互设计师。《AI改变设计》书籍作者《前瞻交互》书籍作者《写给设计师的技术书》书籍作者《智能座舱HMI设计》书籍作者

- **[叶彦辛] | [云司科技]**

解决方案架构师,央企与高校AI培训专家,全国教学数字化大赛一等奖获得者。微软、阿里、AWS认证 GenAI专家。AI布道师,智能体开发资深专家。

🚫 一票否决制 (Veto Power)

为了维护比赛的公平性与核心宗旨，评审团拥有“一票否决权”：

1. **模型违规**：经技术核查，发现 Agent 核心逻辑未调用指定的国产大模型，或后端“套壳”其他模型者，直接取消成绩。
2. **抄袭作弊**：发现代码或创意存在严重抄袭、侵权行为者，直接取消成绩。

为什么你不能错过？

Why Join?

- **实战练兵**：不再是写 Demo，而是面对真实的工程挑战。这是检验你 Agent 开发能力的最佳试金石。
 - **大厂直推**：你的代码和报告将直接呈现在 **Agent 厂商** 的招聘负责人面前。
 - **真金白银**：除了奖金，我们提供 **开发周免费 Token** 和 **算力支持**，让你零成本探索 AI 边界。
 - **圈层破壁**：加入官方群，与 Mixlab、Aitocoder 社区的顶尖开发者切磋，结识未来的创业合伙人。
-

❓ 常见问题 (FAQ)

Q: 比赛收费吗？

A: 完全免费。不仅不收费，我们还提供开发期间的免费模型额度与算力支持。

Q: 我是AI编程新手，可以参加吗？

A: 当然可以！我们在12月26日-1月2日设有“7天开发上手营”，有每日教程和直播带练，是新手入门Agent开发的绝佳机会。

Q: 必须组队吗？

A: 支持个人参赛，也支持2-4人组队。我们建议组队参赛，不仅能获得更多福利（如提前解锁Key），还能分工协作，提高胜率。

Q: 只能用国产模型吗？

A: Agent 的核心逻辑（包括推理、规划、文本生成、代码生成）**必须且只能**调用参赛时选定的那一款国产大模型。其余的**多模态调用、AI辅助编程不设限制**。

最终解释权归胜算云大赛组委会所有

胜算云大赛组委会成员：胜算云、Aitocoder、Mixlab、奇绩创坛、Cline-Chinese (Leo)