



# Agent Reward Hub

Agent Skills 领域的 SWE-bench

让 Agent Skill 有评价、有激励、有标准

"就像早期 App Store 没有评分系统"

现在有 1000+ Agent Skills，用户不知道哪个好用  
创作者没有反馈，生态无法进化

评测基准

创作者激励

Monad Testnet

# 😢 现在的 Agent Skills 生态



## 用户痛点

试错成本高

"用到一个烂 Agent  
浪费半天时间"

skills.sh 有 400+ Skills  
但 0 个评分



## 创作者痛点

没有正向反馈

"写了 Agent 没人知道  
不知道好不好用"

Claude Code / Manus  
创作者 0 收益



## 行业痛点

没有评测标准

"什么是好 Agent?  
怎么科学地评测？"

代码生成有 SWE-bench  
Agent Skills 没有



# 我们的解决方案

**App Store 评分 + GitHub Sponsors + SWE-bench**

= Agent Reward Hub



## 评测基准

三层 Grader 体系  
六维能力评分



## 创作者激励

98% 打赏直达  
成长追踪 + 徽章



## 等级排名

铂金/黄金/白银  
透明排行榜

一次评测，多方受益



# 为什么现在做?



**Anthropic 刚开放 Agent Skills API (2025)**

处于生态爆发前夜

**早期 App Store**

有评分系统 → 成为标准 → 垄断市场

**我们的窗口**

现在做评测平台 = 定义 Agent Skills 的标准

就像 SWE-bench 定义了代码生成评测标准  
我们要定义 Agent Skills 的评测标准



# 科学评测体系

## 三层 Grader 体系

1

### Code Grader

快速、客观、低成本

2

### Model Grader

灵活、理解上下文

3

### Human Grader

准确、金标准

## 六维能力评分

质量

Quality

相关性

Relevance

可靠性

Reliability

效率

Efficiency

准确性

Accuracy

安全性

Safety

对标 Anthropic Eval 方法论

pass@1 / pass@3 / pass@5 指标

# 产品形态



## 排行榜

铂金/黄金/白银/青铜  
透明可追溯



## 测试报告

详细的评测结果  
历史趋势分析



## Skill 对比

并排对比相似 Skills  
帮助用户决策

创作者成长追踪

## 月度统计

Skills / 测试 / 收益

## 成长曲线

可视化进步轨迹

## 成就徽章

里程碑激励

## 一键打赏

98% 直达创作者



# 生态集成

## 无缝集成 OpenClaw / ClawHub

让用户一键安装经过评测的高质量 Skills

```
# 搜索带评分的 Skills
```

```
monad-skills search --platform openclaw
```

```
# 安装到 OpenClaw
```

```
monad-skills install ai-writer --target openclaw
```



### 质量保证

只推荐评测通过的 Skills



### 格式兼容

自动生成 SKILL.md



### 创作者收益

跨平台打赏激励

# ⚙️ 技术架构



## 智能合约

- ERC20 \$ASKL 代币
- Skill 注册表
- 自动 98/2 分账
- 评测记录链上存证

## 评测系统

- Docker 沙箱执行
- Puppeteer 测试自动化
- Anthropic Evals 指标
- 三层 Grader 架构

# 商业模式

## 代币经济飞轮



更多评测



更多用户



更多打赏



更多 Skills



### 创作者订阅

高级数据分析  
排行榜推广



### 项目方赞助

优质 Skill 空投  
活动合作



### 代币增值

生态繁荣  
平台/创作者双赢

# ✓ MVP 进展



## 智能合约

ASKLToken.sol



## 前端页面

Next.js + Tailwind



## 数据库设计

评测数据 Schema

## 当前功能

- ✓ Skills 目录浏览
- ✓ 钱包连接 (RainbowKit)
- ✓ 评测数据库设计
- ✓ GitHub 统计集成
- 评测执行引擎
- 打赏功能
- 排行榜页面
- 合约部署测试网

# 👥 团队介绍



## 产品经理

办公文档识别产品负责人

期待做 Agent 破亿的 Agent 应用



## 全栈开发

AI 写作和漫剧 Agent 工程师

Web3 + 全栈开发经验

"我们都是 AI 开发者，切身感受到这个痛点"

Monad Hackathon 2026 • Rebel in Paradise



# Agent Reward Hub

让 Agent Skill 有评价、有激励、有标准

成为 Agent Skills 领域的 SWE-bench

谢谢评委！

欢迎体验 DEMO 和提出建议

## 部署链

Monad Testnet

## GitHub

[github.com/agent-reward-hub](https://github.com/agent-reward-hub)

## 联系

your@email.com