BITTENSOR / TAO SUBNET

Private — 隐私沙盒 Agent 子网

面向可验证私密工作的去中心化隐私计算市场

问题背景：AI 时代的“不敢交付”

敏感数据泄露风险

AI Agent 正在接管高价值任务（代码、日志、API Key、法律文件），但这些任务涉及企业核心隐私。

- 私有代码仓库与配置
- 服务器日志与账户数据
- 财务、合同与法律文件

信任层缺失

用户并非不需要 AI，而是不敢将数据交给黑盒化的远程 API 或未知第三方 Agent。

真正缺失的是：一个可验证、可隔离、可审计的隐私执行环境。

核心解决方案

构建全球性的“私密 Agent 任务执行市场”

关键原则：隐私与可控的基石

本地模型推理

矿工在本地隔离沙盒运行模型，不依赖 OpenAI 等外部 API。

受控沙盒环境

Agent 在隔离的 Sandbox/microVM 或 TEE 中运行，减少隐私泄露面。

全程可审计

Validator 自动审计任务正确性、可复现性与隐私边界。

网络零信任

默认禁止外联，仅允许白名单访问，严防数据回传。

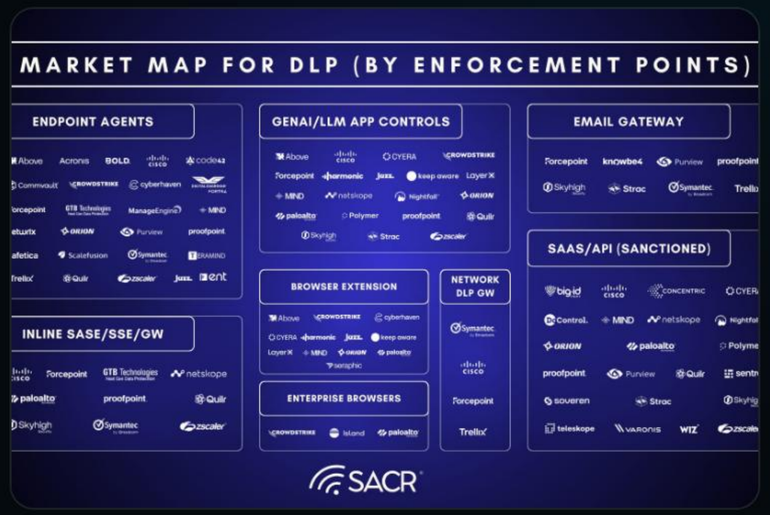
数字智能商品：可信工作单元

私有代码审计

返回漏洞发现、可疑路径、修复 Patch 及审计报告。

DevOps 排错

返回根因分析、修复命令、复现记录与诊断结果。



私密文档分析

针对合同、流水、审计材料进行结构化提取与风险点分析。

"我们卖的是：可信地完成私密任务的能力 (Attested Private Agent Work Unit)"

组织力：全球专业 Miner 角色分工

中心化团队无法覆盖所有软件栈。子网通过激励机制，将全球贡献者组织为专业角色：

Code Miner

专注私有代码审计、补丁生成、静态分析与测试。

DevOps Miner

专注 Docker、K8s、CUDA 与基础设施排错。

Document Miner

专注文档审阅、字段提取与敏感信息风险合规。

Browser Miner

专注私密浏览器自动化流程与后台操作。

核心壁垒：安全执行环境 + 本地模型 + 任务型 Agent 能力的深度组合。

验证力：五层自动化审计架构

审计层级	核心验证内容	目的
第一层：格式验证	Schema 合法性、task_id、签名与哈希匹配	确保输出结构标准化
第二层：执行证明	容器哈希、模型哈希、资源消耗与执行轨迹	提高结果伪造门槛
第三层：隐藏测试	Validator 持有 Miner 不可见的 Hidden Tests	防止硬编码与过拟合
第四层：隐私审计	异常请求监控、敏感字段过滤、Canary Token 检查	严守隐私安全红线
第五层：TEE Attestation	Enclave Quote、运行环境身份、内存测量哈希	硬件级信任背书

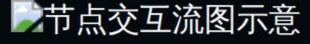
博弈力：让作弊者无处遁形

- **诱饵机制 (Canary Token):** 在任务中植入假 API Key 或私钥。若在 Miner 日志或网络中检测到调用，直接惩罚。
- **最小特权原则:** 仅提供短期凭证 (Capability Token) 或只读 Vault，确保 Miner 拿不到高价值资产。
- **Commit-Reveal:** 先提交结果哈希，待周期结束后 Reveal 完整结果，防止重放攻击与抄袭。

恶意行为后果

隐私事故或重大违规将导致：权重清零、押金削减 (Slashing) 及黑名单永久封禁。

交互逻辑：协作闭环



Validator (验证者)

定义规则、选择合格 Miner、审计证明、打分并分配权重。

Miner (矿工)

提供安全沙盒、运行本地模型、执行任务并提交结构化证明。

评分公式：质量与安全的硬性平衡

$$\text{Final Score} = (0.40 \cdot \text{正确性} + 0.25 \cdot \text{复现性} + 0.20 \cdot \text{质量} + 0.15 \cdot \text{速度}) \times P \times S \times N$$

门槛乘数 (P/S/N): 隐私、安全、网络合规任一失败，门槛值为 0，最终得分 0。

核心理念: 任务质量虽然重要，但隐私和安全是不可逾越的硬性前提。

对比：Private vs. 普通远程 Agent

维度	普通 AI Agent	Private 子网
数据输入	直接提交给远程大模型 API	进入隔离沙盒或 TEE 环境
模型运行	依赖外部第三方服务推理	Miner 本地化运行开源模型
验证方式	依赖人工判断结果好坏	自动化 Hidden Tests 与执行证明
核心价值	追求自动化与效率	隐私、安全与结果可验证性

Private

AI 时代的底层信任安全层

下一代 AI 不仅会回答问题，更会操作用户的私密资产。我们提供一个让 AI Agent 安全处理私密任务的去中心化信任层。

BUILDING THE DECENTRALIZED PRIVACY INFRASTRUCTURE

Image Sources

https://media.easy-peasy.ai/2e09dc3e-7d87-4ab9-afe8-c4728fb276dc/8ab4865b-f5e9-429b-b353-547d90e40062_medium.webp

Source: easy-peasy.ai

https://sipkosecurity.com/wp-content/uploads/2026/03/5d0f2f70e1ef4f6680e27710eed40472-1024x559.webp

Source: sipkosecurity.com

Thumbnail for www.meta-intelligence.tech

https://www.meta-intelligence.tech/images/insight-ai-sovereignty.webp

Source: www.meta-intelligence.tech

https://substackcdn.com/image/fetch/$s_!K5Li!,f_auto,q_auto:good,fl_progressive:steep/https%3A%2F%2Fsubstack-post-media.s3.amazonaws.com%2Fpublic%2Fimages%2Fff4d9f8a-0402-4f43-8d42-38a440d1849f_1920x1080.png

Source: softwareanalyst.substack.com