

# 真正的龙虾平权，是零门槛的全民生产力

爆火时，地铁里有人在刷教程，朋友圈里有人在晒养虾战绩，全民养虾的热情，让Token经济学以一种近乎病毒传播的方式完成了对全市场的教育。

然而不到三个月时间过去，所有讨论迅速收敛回极客圈内部，全球范围内龙虾相关的谷歌指数、百度指数、微信指数集体暴跌。

难道是一夜之间，大家才突然意识到部署龙虾需要注册GitHub，要有英文阅读能力；Clone项目需要懂Git；配置环境需要会写Python；申请API Key需要信用卡；调试Skills需要理解权限和路径；如果是Windows系统还要额外处理OpenClaw的各种兼容问题。

龙虾降温背后，问题究竟出在哪里？

AI坟场，堆满了养不起的赛博龙虾

要回答龙虾为何降温这个问题，我们不妨先算一笔账。

一个典型的养虾失败案例是这样的：某办公室白领，在朋友圈刷到“安装这个龙虾，节省了我XX%的办公时间”“借助龙虾，人人都是一人公司”的文章后热血沸腾，花了600块找技术服务商完成了本地部署，又花了几千块购买高配显卡电脑。

然后呢？

然后就没了然后了。

因为养虾的实际成本比很多人想的要高得多：

大部分普通电脑没有针对本地Agent运行设计权限保护机制，因此需要将养虾本与日常工作办公的设备区分，相关支出7000+起步。

但真正的成本大头，其实在后续的使用与维护。要想虾努力工作，就要使用Claude Sonnet 4.7在内的SOTA模型，月均Token消耗成本700-2000元不等。

因为龙虾为代表的Agent与普通Chat AI的核心区别在于其多轮深度推理能力。

一个简单的报表整理任务，可能就需要拆解需求、调用工具、校验数据、输出结果，每个环节都需要调用模型推理，堪称Token消耗黑洞。

也就是说，在Agent模式下，完成一个任务，其模型调用次数会是日常问答的十多倍。在此基础上，又会带来Token消耗近乎百倍的指数级增长。

Transformer架构模型，采用的自注意力机制会要求每个Token与

全量Token关联计算，这就导致算力消耗与上下文长度呈平方级增长。上下文翻倍，算力会翻4倍；上下文从1K增长到10K，token消耗就会暴涨100倍。

当然，成本问题并非无解。

目前行业一个典型做法，就是将部分任务放在本地完成，解决云端大模型反复调用的成本困境。

拿荣耀近期将发布的搭载自研YOYO Claw技术的MagicBook轻薄本来说，通过为本地硬件搭载行业首个端侧Token智能调度引擎，相比于OpenClaw来说，可以为养虾带来50%以上的综合Token成本节约。

具体来说，Token智能调度引擎做的第一件事是本地分流。基于用户任务理解，它会借用丰富的本地调用工具，比如用机控制、语义搜索、生图识图项目模型、文件管理等等，将部分相对简单、重复、好解决的问题直接放在本地处理，做到云端零Token消耗。

在此基础上，Token智能调度引擎还会针对用户的提问，做精细化的上下文压缩以及skills匹配。

在上下文环节，很多龙虾会一股脑将所有的历史会话与信息全部喂给大模型，Token调度引擎则会对本轮上下文信息进行压缩，然后通过向量数据库做跨会话的相似性检索，召回最相关的部分，来降低整体的上下文输出长度。

做Skills调用时也是相同的逻辑，基于任务意图定向调用Skills，避免全量技能加载的无效消耗。

也就是说，哪怕一半的上下文压缩，都能带来四倍的算力节约。而荣耀官宣的50%Token节约，甚至还只是一个保守数据。

荣耀YOYO Claw技术成本解决的另一个秘诀则在于绝对的安全管控。过去为了避免龙虾误授权导致出现删库、调取摄像头、访问机密文件，用户不得不为了养虾部署云服务或者购买单独的硬件设备。

而YOYO Claw技术在系统层面，专门部署了一个独立安全龙虾，作为所有AI操作的守门人，保证用户身份、记忆数据全本地处理，专属Skills通过硬件加密存储，全链路权限管控杜绝越权。一旦检测到高危操作，安全龙虾就会实时拦截并告警，从底层根除安全隐患。如此一来，用户就可以放心的用它作为日常的办公本、生活本，安心养虾。

当然，导致龙虾退潮，成本只是诸多原因中最好解决的一个，对大部分跟风的人来说，龙虾不会布，布了又不知道怎么用，什么时



候用，才是最大的问题。

原生OpenClaw整个产品的设计逻辑都建立在用户懂技术的前提下。但技术社区觉得理所当然的操作，对普通人来说难如登天。

针对这个问题，荣耀直接将过去养虾需要的繁琐部署安装调试简化到扫码登录即用。用户打开MagicBook，开机扫码登录，就能一键打通微信、飞书两大国民办公工具，后续不管有什么需求，只需要对话交互就能搞定。

但这还只是基础，过去行业最大的误区，是把产品部署等同于产品可用乃至产品好用。市场上存在大量非技术背景但对AI有刚需的普通人，比如家长想用AI辅导孩子但不会配置；职场人想让AI帮忙写周报但找不到入口；银发族想用AI辅助健康管理但不知如何开始。他们正是AI技术走向规模化落地必须跨越的最大群体。而一个不会帮助他们解决问题的龙虾，注定只是一个闲置的算力玩具。

为此，荣耀把场景覆盖做到了最后一公里。借助YOYO Claw技术，荣耀直接跳过用户自学的环节，为学生、办公人群预置了教育、办公、学术、创作、智能辅助五大场景专属龙虾。写论文时自动触发学术虾，做报表时唤醒办公虾，整个过程无需用户手动调用，系统会基于场景主动匹配能力。

而对于父母或者小朋友这样的养虾小白用户，一句“你能干什么”，就能唤醒全技能清单，接下来，用户点选即可执行。

甚至，对那些对龙虾有定制需求的深度用户他们也能通过在“我的技能栏”中自动导入技能包，做定制化开发，解决个性化需求。从而零代码完成个性化龙虾搭建，把极客的定制能力，下放给了普通人。

而在人人能用的基础上，荣耀还更进一步，让龙虾做到了人人好用。

传统AI Agent是静态的：用户

配置什么，它就一直做什么。而荣耀的方案加入了Agent Memory（智能体记忆），让AI会从与用户的交互中学习，逐步理解你的偏好、习惯、风格、解决问题的方式，并将其转化为专属Skills，让你的龙虾越用越好用。

而对于一些交互结果中可能会被复用的结果，荣耀还创新性地加入任务快照、执行记录功能，直接复用历史结果，规避了相同问题重复推理的消耗。

甚至，通过简单定制，家庭场景下，父母健康数据异常可同步至子女端Agent，自动提醒体检、用药，让AI能力从个人生产力，延伸到家庭服务。

大模型向左，终端向右：AI Agent的入口到底在谁手里？

2026年的养虾热潮中，玩家大致可分为四类：独立开发者、互联网大厂、模型厂商、硬件派。看似群雄逐鹿，实则各自有难以突破的短板。

比如，独立开发者的优势是灵活度拉满，能做深度定制，但这只是极客小众游戏。

模型厂商，比如OpenAI、Anthropic、deepseek作为行业上游卖水人，擅长把AI做强，却不擅长把AI做易。它们的核心商业模式是卖API、按Token收费。相应的，用户用得越多，厂商赚得越多。完全没有做成本优化的动力。更不用提按照云端Token用量方式调用，就意味着用户的所有隐私数据都要上传云端，会带来用户使用中的极大不安感。

互联网大厂做到了低门槛，却始终困于生态孤岛，飞书、钉钉、微信互不打通，跨国办公的Zoom、Google生态更是无法覆盖。很多AI智能体需要的系统级权限（文件管理、跨应用操作），App级别的嵌入根本做不到。也因此，AI智能体在他们手里，更像是App的一个

插件，而不是一个独立的产品。

唯有硬件厂商，能集齐所有优势：软硬一体兼顾本地部署与云端灵活调用，系统级打通突破生态壁垒，零门槛体验对标互联网大厂，定制能力适配极客需求。

但硬件厂商众多，为什么率先完成龙虾普惠的，是荣耀？

答案很简单，荣耀用十年时间，2600人规模的AI研发团队，打造了覆盖端侧模型、云端模型、智能体全栈技术涵盖在内的AI全场景开放生态。

荣耀是行业少有的全场景硬件玩家，此前早已实现手机、平板、PC的无缝数据流转，以及AI能力适配。2016年，荣耀在第一代Magic机型上，就开启了AI与硬件融合的探索；此后到了2023-2024年的百模大战全民拼参数、拼榜单时，荣耀逆势发布魔法大模型1.0、2.0，深耕端侧模型与硬件协同，死啃优化意图理解与应用调度；并在2024年的Magic7系列率先落地端侧AI智能体，完成硬件Agent的技术验证。

到了AI agent时代，这种全栈能力再次完成质变：PC端龙虾学会的Skills，沉淀的Agent Memory，能同步到手机、平板全终端。你在户外用手机发一句指令，家里的PC就能自动完成报表整理、论文查重、文件归档，搞定跨终端的Agent协同。

这种单设备→多设备→全场景的深度覆盖，是单一硬件厂商、软件厂商都无法复刻的壁垒。

结尾

2026年3月，中国AI产业悄然跨越了一道分水岭，日均Token调用量突破140万亿。

这一数字意味着什么？对比来看，2024年初这一指标尚在千亿美元级，也就是说，两年增长超过千倍，中国AI产业，正式告别技术验证期，迈入规模化应用阶段。

如果说Token经济崛起的上半场，行业在拼消耗、拼规模、拼调用量；到了下半场，竞争的核心早已切换，谁能降低Token成本，谁能保障数据安全，谁能抹平使用门槛，谁就能掌控AI普及的主动权。

在这个过程中，荣耀是行业第一个把Token效率落地的企业，也是第一个把AI Agent从极客玩具，做成全民生产力的企业。

而历史上，无数的商业变迁也无数次证明了，所有成功的技术革命，最终指向的都是科技平权。

## 马斯克突然出手：SpaceX 600亿美元“截胡”Cursor，微软退出竞购



两名知情人士透露，在SpaceX本周宣布获得以600亿美元收购人工智能编程初创公司Cursor的排他性权利前，微软曾评估过收购该公司的可能性。

因谈判细节未公开而要求匿名的消息人士表示，微软近期正提升其人工智能工具的普及率，以应对相关工具市场的快速扩张，但最终决定不参与对Cursor的

竞购。微软方面对此拒绝置评。Cursor发言人亦未回复置评请求。

当前的AI编程辅助市场由Cursor、Anthropic及OpenAI主导。尽管微软通过旗下的GitHub Copilot在开发者群体中维持了一定规模，但其在该赛道的核心角色仍侧重于云服务提供商与财务投资者。此前，微软已向OpenAI

和Anthropic投入数十亿美元，作为交换，这两家企业承诺在微软Azure云平台上进行大规模的算力采购。

资本对代码生成工具的狂热正在急剧升温。据CNBC本月早些时候报道，部分风险投资机构原计划以500亿美元的估值参与Cursor的新一轮融资。

然而，由埃隆·马斯克(Elon Musk)控制的SpaceX于周二在社交平台X上发布声明称，已与Cursor达成协议，取得在今年年底以前以600亿美元收购该公司的权利；若交易最终未能成行，SpaceX将向其支付100亿美元的解约金。

另一名知情人士补充称，SpaceX的介入发生在Cursor此轮融资的尾声，这一交易令潜在投

资者措手不及。在正式宣布交易的前几周，SpaceX就已开始向Cursor提供算力资源支持。

针对该项合作，SpaceX在声明中表示：“SpaceXAI与Cursor现正密切合作，旨在开发全球领先的编程与知识处理AI模型。”Cursor首席执行官迈克尔·特鲁尔(Michael Truell)随后在X平台上回应称，期待与SpaceX团队合作，共同提升其AI模型Composer的性能边界。

此次收购动议发生在全球科技巨头密集整合AI资产之际。今年2月，马斯克将SpaceX与其名下的AI初创企业xAI合并，合并后实体估值达1.25万亿美元。目前，他正筹备IPO，按现有估值计算，这有望成为全球资本市场规模最大的上市交易。

相比之下，微软在资本市场的表现近期承压。今年以来，该公司股价累计下跌约10%，整体走势弱于大盘及主要的云计算竞争对手。在业务层面，微软首席执行官萨提亚·纳德拉(Satya Nadella)曾于今年1月向分析师披露，GitHub Copilot的付费订阅用户量达到470万，较去年同期增长75%。

行业内其他竞争者的数据亦显示出该领域的扩张速度。OpenAI首席执行官萨姆·奥尔特曼(Sam Altman)周二在X平台上透露，其代码生成工具Codex的活跃用户数在突破300万后，仅用时不到两周便增至400万。此外，受旗下Claude Code服务推动，Anthropic本月已实现年化收入300亿美元的里程碑。