

# AI Agent 范式升级：从 OpenClaw 看产业链增长机遇



王竞萱 分析师

Email:wangjingxuan1@lczq.com

证书:S1320525020001

## 投资要点:

**2026 年开年的开源 AI Agent 产品 OpenClaw 成为现象级产品。**该产品快速破圈，GitHub 星标数超 30 万登顶开源软件榜首，国内外迅速掀起“养虾热”。**我们认为 OpenClaw 的爆火是技术创新与用户体验优化深度结合的结果。**技术上，三层架构实现多渠道交互与本地执行协同，四大基础工具支撑复杂任务拆解，分层 Skill 体系可自我扩展，四层记忆构建完整上下文。体验上：本地部署保障数据主权，多渠道网关降低使用门槛，类人化执行大幅提升生产力，持久记忆实现个性化服务，开源免费+多模型兼容覆盖极客到普通 C 端的多元需求，最终完成全民破圈。OpenClaw 的出现标志着人工智能实现了被动承接到主动发起的范式转变，或对上下游厂商带来深刻影响。

**上游算力赛道 token 消耗驱动需求，随 Agent 持续渗透形成长期趋势。**

根据我们的测算结果，到 2026 年底有望带动全球月 Token 消耗量增长 19.1-489.6T，以此测算 Blackwell 芯片对应的数量增长达到 2432-62271 片，若考虑难以量化的多因素影响，实际增量或更高于测算值。AI Agent 的长期渗透将持续抬升算力需求的增长斜率，算力基建的国产化替代与高端算力芯片的产能落地，将成为这一赛道的核心投资主线。

**中游大模型行业二元分化格局显现，国产高性价比厂商迎来黄金机遇。**兼具高性能与低费率的国产模型，不仅有望在 C 端实现规模化渗透，还具备依托性价比优势出海的潜力，而针对 Agent 场景的模型推理优化、工具调用能力升级，将成为后续大模型厂商的核心竞争力。

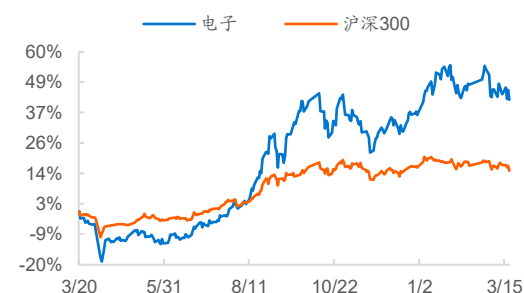
**下游应用在政策加持下场景落地提速，安全与技术成规模化商用关键。**多地政策密集加码，各种补贴形式持续降低研发与部署门槛，OPC 模式与 OpenClaw 的深度绑定也打开想象空间，短期来看，下游机会集中在能借助 AI Agent 实现降本增效的垂直场景，而长期规模化商用则依赖技术迭代，具备场景落地能力与安全技术解决方案的厂商或将率先受益。

**建议关注：上游：国产算力芯片、AI Infra；中游：国产大模型、AI 基础软件；下游：大型云厂商，垂直场景解决方案商等。**

**风险提示：Agent 需求扩张速度不及预期；技术开发进度不及预期；风险事件缓释进度不及预期；政策出台及落地进度不及预期**

投资评级：看好（维持）

## 市场表现



## 相关报告

## 目 录

1. 什么是 OpenClaw? .....	4
1.1 OpenClaw: 迅速崛起的 AI Agent 现象级产品 .....	4
1.2 爆火原因: 从技术角度看 OpenClaw 的独到之处 .....	6
2. 上下游探索: 有哪些赛道可能因此长期受益? .....	9
2.1 上游算力: “Token 消耗怪兽” 抬升长期增长斜率 .....	9
2.2 中游模型: 从推理需求视角看待 Agent 机会 .....	14
2.3 下游应用: 现阶段或以生产力提升为主, 大规模商用还需等待 .....	16
3. 投资建议 .....	19
4. 风险提示 .....	20

## 图表目录

图 1	OpenClaw 吉祥物形象为一只龙虾 .....	4
图 2	OpenClaw 的核心功能.....	4
图 3	OpenClaw 在 GitHub 上的星标数 (截至 2026 年 3 月) .....	6
图 4	OpenClaw 的架构.....	6
图 5	ClawHub 技能市场中的一项 Skill .....	8
图 6	大模型用户使用量.....	9
图 7	Kimi 的 Agent 服务定价 .....	10
图 8	OpenClaw 近一个月的使用量 (tokens) .....	12
图 9	全球 AI 应用 MAU.....	12
图 10	DeepSeek-R1 在 GB200 NVL72 上的 TPS/GPU 情况.....	13
图 11	各模型在 OpenClaw 上完成任务的得分排名 (截至 2026 年 3 月 18 日) .....	15
图 12	各模型在 OpenClaw 上完成任务的速度排名 (截至 2026 年 3 月 18 日) .....	15
图 13	ChatGPT 的订阅价格 .....	15
图 14	各模型在 OpenClaw 上的性能/价格象限.....	16
图 15	本月 OpenClaw 部署模型的使用量排名 .....	16
表 1	OpenClaw 与 ChatGPT 对比.....	5
表 2	OpenClaw 的主要应用场景 .....	5
表 3	OpenClaw 的基础工具.....	7
表 4	OpenClaw 的 Skill 来源 .....	7
表 5	OpenClaw 的四层记忆架构 .....	8
表 6	海外大模型厂商的定价 .....	10
表 7	OpenClaw 支持的大模型价格.....	11
表 8	每人每月消耗 token 量 (百万) 预估.....	11
表 9	国内厂商推出的类 Claw 产品 .....	12
表 10	OpenClaw 持续渗透对 token 消耗量增长的敏感性分析 .....	13
表 11	OpenClaw 带来的芯片增量 .....	13
表 12	近期各地方相关部门发布的鼓励 OpenClaw 及 OPC 发展的政策 .....	17
表 13	ClawHub 技能库分类 Top10 (截至 2026 年 3 月中旬) .....	18
表 14	OpenClaw 与 Claude Code 对比.....	18
表 15	OpenClaw 存在的高危漏洞 (截至 2026 年 3 月 10 日) .....	19

# 1. 什么是 OpenClaw?

## 1.1 OpenClaw: 迅速崛起的 AI Agent 现象级产品

**OpenClaw 是 2026 年现象级开源 AI Agent，具有主动执行能力的个人 AI 操作系统。**它由奥地利个人开发者 Peter Steinberger 发起、基于 MIT 协议开源并采用本地优先部署架构，核心定位为衔接大语言模型与终端设备的执行中枢，可实现 AI 从文本对话向系统级任务执行的范式升级；项目兼容主流大模型与全平台环境，支持隐私可控的端到端自动化，覆盖办公协同、开发运维、数据处理等场景，凭借高速增长的社区热度与轻量化落地能力，正在驱动人工智能从交互生成向行动执行演进。

**OpenClaw 创造了 AI Agent 的新范式，为用户提供了安全、高效、可扩展的 AI 自动化交互与管理方案，其核心功能包括：**

- ①本地化部署与数据主权：支持 Mac/Windows/Linux 本地运行，数据全留存于自有设备，从架构层面保障隐私安全，实现数据自主可控。
- ②个性化持久记忆：内置用户偏好与上下文记忆机制，留存交互历史与行为习惯，提供贴合个人场景的定制化智能服务。
- ③系统级自动化执行：开放完整系统权限，支持文件读写、Shell 命令与脚本执行，实现 AI 从“文本交互”到“系统操作”的能力跃迁。
- ④多渠道统一接入：通过单 Gateway 网关聚合 WhatsApp、Telegram、飞书等主流通讯渠道，插件化扩展支持更多场景，实现全域消息互通。
- ⑤多智能体会话隔离：按智能体、工作区或发送者隔离会话，保障多任务并行下的逻辑独立性与数据安全性，避免任务串线。
- ⑥富媒体交互与全栈管理：支持图片、音频、文档传输，配套 Web 控制面板实现会话、节点全生命周期管理，同时兼容移动节点拓展终端场景。

图1 OpenClaw 吉祥物形象为一只龙虾



资料来源：OpenClaw 官网，联储证券研究院

图2 OpenClaw 的核心功能



资料来源：OpenClaw 官方网站，联储证券研究院

**OpenClaw 与 ChatGPT 代表了 AI 工具的两种不同发展方向，共同丰富了 AI 应用落地的具体形态。**ChatGPT 是面向大众的通用对话工具，以“你问我答”为核心，依托云端网页/APP 便捷使用，模型与数据由平台统一管理，生态相对封闭，更适合日常答疑、内容创作等轻量需求，胜在易用性和普及度。OpenClaw 则是更偏向自主执行型的智能助手，需要用户自行部署服务器，能接入各类聊天平台，可自主完成复杂任务，数据完

全留存于用户本地，同时支持多模型切换与开源技能扩展，在隐私安全、任务自动化和灵活性上更具优势，更适合重视数据主权、需要 AI 深度参与业务流程的用户。

**表1 OpenClaw 与 ChatGPT 对比**

维度	ChatGPT	OpenClaw
交互模式	你问它答	自主执行任务
运行环境	网页/App	自托管服务器，接入 20+消息平台
可扩展性	GPTs 商店	ClawHub 技能市场
数据控制	数据在 OpenAI	完全本地，你拥有所有数据
模型选择	仅 GPT 系列	Claude/GPT/DeepSeek/Gemini/Ollama 本地模型
开源	否	MIT License，完全开源

资料来源：OpenClaw 橙皮书，联储证券研究院

**OpenClaw 以自主智能体 (Agent) 为核心引擎，构建了覆盖信息消费、内容生产、开发运维、知识管理、金融交易的全场景功能矩阵。**功能维度深度融合社交媒体内容解析、目标驱动式内容生产流水线、API/工作流自动化运维及向量语义检索等核心能力，针对个人/团队生产力、学术研究、投资交易等垂直需求提供场景化落地方案（如财经追踪、自动化模拟交易、知识库构建），以低代码、高自动化的一体化智能赋能形态，显著降低内容生产与运维成本、提升信息处理效率与决策质量。

**表2 OpenClaw 的主要应用场景**

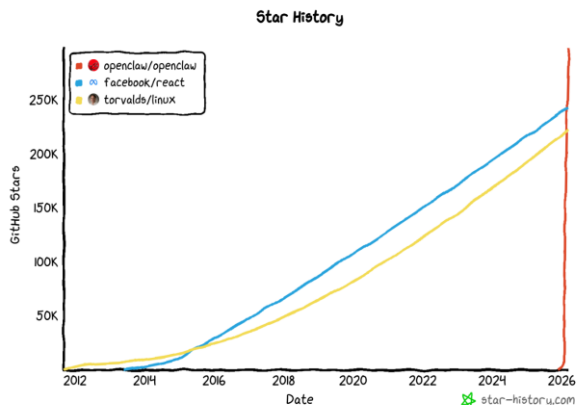
核心功能	功能性质	具体场景	具体内容
社交媒体	主打信息聚合与提炼，帮你用最少时间掌握核心内容	每日摘要	自动生成社交媒体的每日精选摘要，按你关注的社区或频道定制
		账号分析	对个人的社媒账号进行定性分析，帮你判断账号运营状态
创意与构建	让 Agent 帮你干活，甚至帮你生成产出	新闻聚合	聚合多来源的科技新闻源，自动评分、分类、生成摘要并分发
		目标驱动的自主任务	根据目标自动拆任务、排计划，甚至构建小型应用
基础设施与开发运维	把运行维护工作交给 AI，节省成本	内容产出流水线	自动化自媒体频道的选题、研究、追踪全流程
		多智能体内容工厂	分工做研究、写作、做缩略图
生产力	个人或团队提效	n8n 工作流编排	通过 webhook 将 API 调用委托给 n8n 工作流，并能做到安全隔离 API 凭证
		自愈家庭服务器	运行一个始终在线的基础设施智能体，具有 SSH 访问权限、可以完成定时任务作业，以及跨设备自动修复故障
研究与学习	帮你抓信息、建知识库、挖需求	项目管理	使用 STATE.yaml 模式协调多智能体项目-子智能体并行工作，无需编排器开销
		多渠道助理	整合你的通讯工具，实现 24/7 自动回复
金融与交易	投资需求	个性化工具	个人 CRM、个人助理、家务助理、健康追踪、项目追踪、Todoist 同步、第二大脑笔记库
		自动化流程	比如用语音确认活动嘉宾，或搭建实时数据看板
研究与学习	帮你抓信息、建知识库、挖需求	财报追踪器	追踪财报，带有自动化预览、警报和详细摘要。
		个人知识库	通过将 URL、推文和文章拖入聊天来构建可搜索的知识库
金融与交易	投资需求	市场研究与产品工厂	从网络上挖掘用户的真实痛点，然后自动构建解决方案雏形
		语义记忆搜索	使用混合检索和自动同步，为你的 OpenClaw markdown 记忆文件添加向量驱动的语义搜索
金融与交易	投资需求	自动化模拟交易	在预测市场上进行自动化模拟交易，带有回测、策略分析和每日绩效报告

资料来源：GitHub，联储证券研究院

**OpenClaw 近期在开源社区实现爆发式增长，登顶同类项目榜首，成为 AI Agent 赛道现象级产品。**截至 2026 年 3 月 11 日，OpenClaw 在全球最大开源社区 GitHub 的星标数已超过 30 万，仅用了短短几个月的时间就超越了 React 和 Linux，成功登顶，成为史上星标数最高的开源软件项目。

同时，OpenClaw 快速突破极客圈层实现 C 端全民破圈，“养虾热”在普通用户群体中也迅速蔓延开来。国内抖音、小红书、微博等社交平台“养虾”话题浏览量破亿，线下出现千人排队协助部署场景，闲鱼、淘宝第三方安装服务供不应求，单价最高超 800 元。OpenClaw 形成全民参与、自发传播、实操驱动的爆火格局，成为 AI 智能体向大众普及的标志性事件。

图3 OpenClaw 在 GitHub 上的星标数（截至 2026 年 3 月）



资料来源：GitHub Star History，联储证券研究院

## 1.2 爆火原因：从技术角度看 OpenClaw 的独到之处

OpenClaw 采用了 Channel-Gateway-Node 三层架构，以 WebSocket 为通信总线，将控制平面、设备执行与消息渠道解耦，实现了跨平台交互与本地自动化操作的高效协同。其中 Channel 层作为用户交互入口，接入了超过 20 种主流即时通讯平台，一方面降低了普通用户的使用门槛，另一方面使得用户发起请求的形式更加便捷，是产品实现全民破圈的关键交互基础。Gateway 层作为核心调度中枢，承担了服务维护、会话管理、Agent 调度等职能，它是连接上层消息与下层设备执行的核心枢纽，实现了任务路由、状态同步与资源调度的中心化管控，同时由于该层的 24/7 运行，使得会话结束后也不会丢失上下文，记忆得以积累。而 Node 层作为任务落地载体，该层负责本地设备操作执行，核心能力覆盖摄像头调用、屏幕录制、系统命令执行（system.run）等，将 AI 指令转化为可落地的本地自动化操作，实现闭环，为用户得跨软件自动化操作提供了支撑。

在这种解耦分层的设计下，实现了各模块的独立迭代和弹性扩展。一条消息从用户发出到 Agent 回复，完整经历的途径如下：①用户发消息；②Channel 接收；③Gateway 路由；④Agent 处理；⑤Node 执行；⑥回复用户。

图4 OpenClaw 的架构



资料来源：OpenClaw 橙皮书，联储证券研究院

OpenClaw 将系统交互能力抽象为 4 类基础工具，分别对应人类完成任务的核心肢体功能，形成无依赖、可复用的操作。不同于此前 AI Agent 探索尝试的接入 API 或者是屏幕识别的方案，在上述架构之下，OpenClaw 仅有四类基础工具，但是通过基础工具的组合使 OpenClaw 可以对复杂任务实现目标拆解、工具编排到迭代执行的标准化流程，将复杂目标解构成为一系列线性的简单子任务。因此，即使是并未给 OpenClaw 预设好

的任务模块，它也可以用“输入-执行-输出”的全链路工具给出通用解决方案。

也正是因为对这种复杂系统“高耦合功能”的拆解，使用户在日常使用 OpenClaw 的时候能感受到与其它 AI 工具不同的“聪明感”，在结构化复杂任务的执行层面高度贴近人类操作范式，表现出强的执行力与成果交付能力，可高质量完成可拆解、流程化的业务目标。更少的工具也意味着更短的 system prompt、更少的 token 消耗和更快的响应。

**表3 OpenClaw 的基础工具**

工具	功能	类比
bash	执行任何命令	手脚
read	读取任何文件	眼镜
write	创建新文件	笔
edit	编辑已有文件	橡皮

资料来源：AI 科技评论公众号，联储证券研究院

**Skill 是 OpenClaw 的能力扩展单元，它的技能库决定了 Agent 的能力边界，但是令人惊喜的是 OpenClaw 具有自我扩展的能力。** OpenClaw 的 Skill 是封装特定业务逻辑、由 bash/read/write/edit 等基础工具编排而成的可复用执行单元，是系统承接复杂任务、沉淀业务经验的核心载体。它采用三级优先级架构：项目级 Skill 仅对当前工作区生效，适配专属业务流程；用户级 Skill 全局生效，承载个性化高频场景；内置级 Skill 随版本发布，提供开箱即用的基础能力，三者优先级依次递减。

**表4 OpenClaw 的 Skill 来源**

优先级	位置	说明
最高	<workspace>/skills	项目级 skills，只对当前工作区生效，适合针对特定项目定制的能力。
中	~/openclaw/skills/	用户级 skills，全局生效，通过 ClawHub 安装或手动放置的 Skills 都在这里。
最低	bundled skills	内置的 55 个 Skills，随 OpenClaw 版本发布，不需要安装，开箱即用

资料来源：OpenClaw 橙皮书，联储证券研究院

当 OpenClaw 启动或收到消息时，Skills 的加载遵循以下流程：

①**读取 Skill 元数据**，扫描三层目录，解析每个 Skill 的名称、描述、触发条件、所需环境变量等元信息；

②**应用环境变量**，如果 Skill 声明了需要的 API Key 或环境变量，系统就会从相应文件中注入，缺少必要变量的 Skill 则会被静默跳过；

③**构建 System Prompt**，将所有可用 Skills 的描述注入到 system prompt 中，告知模型当前可以调用哪些能力，也就让模型知道自己能做什么；

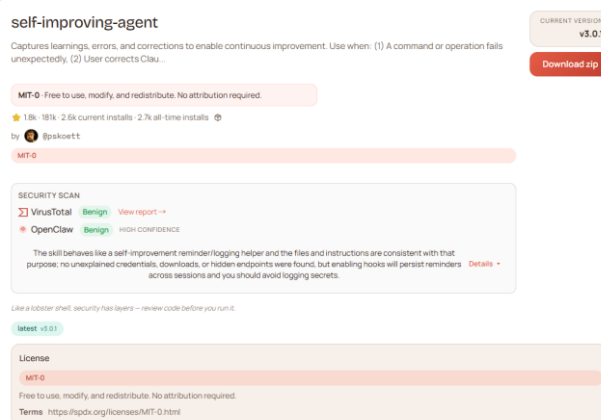
④**运行后恢复**，Skill 执行完毕后，恢复原始环境变量和上下文状态，避免 Skill 之间互相干扰。

对于 OpenClaw 没有接触过的任务，它也会在顺利执行任务后总结封装。在用户向 Agent 发布其并没有技能储备的任务时，它也按照从执行任务到总结经验到封装复用的全自动化链路去形成新的 Skill，这一自我学习能力可以不断拓宽 OpenClaw 的能力界限。

**ClawHub 是 OpenClaw 生态下的 Skill 共享与管理平台，核心承载用户级通用 Skill 的分发、安装与更新能力。**该平台既为用户提供覆盖代码生成、文档自动化等高频场景的标准化 Skill 资源库，大幅降低自定义能力开发门槛；也支持用户沉淀并共享个性化 Skill，形成可复用的业务能力生态。自我迭代 Skill 叠加丰富的技能市场，让 OpenClaw

可高效覆盖从通用自动化到项目专属 pipeline 的全场景复杂任务，是实现规模化业务落地与类人化执行的关键支撑。

图5 ClawHub 技能市场中的一项 Skill



资料来源：ClawHub，联储证券研究院

Agent 记忆方面，通过多种记忆工具的叠加，OpenClaw 具有了区别于普通 Chatbot 的核心记忆能力。四层记忆从不可变的身份内核到实时对话，构建了完整的上下文连续性。四层记忆种：

灵魂是内核人设，它决定了 Agent 是谁，具有什么行为风格，应该遵守什么价值观。

Tools 是能力模块，是 Agent 可以为用户完成什么样的任务的关键。

User 是长期记忆，其中既包括工作日志，在三种情况下生成：①每天结束时自动总结；②上下文快超过模型限制时做压缩；③Agent 主动判断“这件事值得记下来”，也包括长期总结，将记忆文本转为向量，在后续提及时再做索引

Session 则是实时会话，也就是正在对话的实时上下文。

可以看到，OpenClaw 采用了多种规格的记忆工具，尽管可能因功能冗余而降低效率，却最大程度确保了 Agent 的真实性。这种真实性既体现在 Agent 在为产品服务过程中形成的有契合度最大公约数的独特灵活，又体现在用户的某个历史对话也许会被 Agent 记住，在未来某一时刻提起，让用户真的感受到自己的 Agent 记得自己。

表5 OpenClaw 的四层记忆架构

层级	存储位置	定位	生命周期	说明
Soul	SOUL.md	不可变内核	永久不可变	Agent 的人格、价值观、核心身份定义，创建后不可修改
Tools	Skills+Extensions	动态工具	按需加载	当前可用的工具和技能列表，随安装和加载动态变化
User	MEMORY.md+向量数据库	语义长期记忆	持久化	关于用户的偏好、决策、历史事实，支持语义搜索
Session	内存+sessions.json	实时情景	会话级	当前对话的实时上下文，Token 耗尽时被压缩

资料来源：OpenClaw 橙皮书，联储证券研究院

总结来看，我们认为 OpenClaw 的技术创新其实并不十分晦涩，但是极大地改善了用户体验，其技术创新本质是以类人化执行范式为核心，从接入、执行、能力、交互四个维度优化体验。多渠道架构降低使用门槛，原子工具保障执行灵活性，分层 Skill 体系扩展场景覆盖，四层记忆强化交互温度。这种全链路优化让系统既能高效处理结构

化复杂任务，又具备贴近人类的交互体验，同时通过生态化扩展持续丰富能力边界，最终驱动其在自动化与 AI 执行领域快速火爆。

## 2. 上下游探索：有哪些赛道可能因此长期受益？

在 OpenClaw 的持续火爆推动下，我们认为上下游的产业格局和供需关系都有望受到深刻影响。上游算力成为直接受益赛道，OpenClaw 的高 Token 消耗特性将显著抬升算力需求；中游大模型行业可能将呈现 2B/2C 分化演绎的格局，兼具高性能与低费用的国产大模型厂商成为主要受益对象；下游应用获多个地方的多级政策体系加持，但是考虑到模型能力及风险问题，现阶段或将以降本路径改善企业利润。

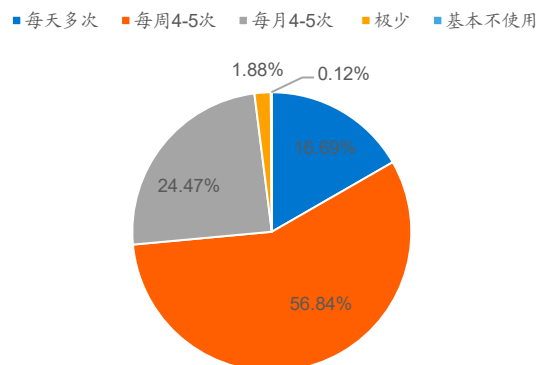
### 2.1 上游算力：“Token 消耗怪兽”抬升长期增长斜率

OpenClaw 采用开源免费+基础设施成本的付费机制。OpenClaw 软件本体核心程序完全开源免费，个人用户可无限制使用基础功能并进行二次开发，无安装费、激活费或版权费，用户的主要费用支出在于必要运行成本，即 AI 模型 API 调用费用，该费用按 Token 计量，该部分费用由用户直接向基础设施与模型提供商支付。

OpenClaw 作为直接面向 C 端的产品，其费用应该是用户高度敏感的，预计每月用户端（海外）平均支出费用在 30-50 美元。我们认为尽管 OpenClaw 采用的是 API 计价，但是用户在使用过程中应存在心理价位隐含上限，参考国际大模型厂商的套餐定价：中间价位的版本，如 ChatGPT Plus、Gemini Pro、Claude Pro 等月费基本均处于 20 美元区间，尽管重度用户或专业用户会为更强模型、更大用量、更快速度支付溢价，但该类用户占比应较少，参考艾媒咨询统计数据，每天多次的大模型使用者占比约 16.7%，因此我们估算这类用户并不会对平均费用支出产生过度影响。

OpenClaw 用户在主动被动层面都会支付相较于大模型更高的费用，主要原因在于：第一，主动方面，OpenClaw 作为 AI Agent 极大地改善了与人工智能交互的体验，用户支付溢价的意愿会有一定提升；第二，被动方面，OpenClaw 由于存在心跳、多记忆加载、定时任务、循环执行等机制，其 token 消耗量是天然大于 Chatbot 的。因此我们认为，OpenClaw 用户在二者共同作用下的心理价位隐含上限或为月费均价的 2-3X，对应价格即 40-60 美元/月。

图6 大模型用户使用量



资料来源：艾媒咨询，联储证券研究院

我们预计使用 Agent，相较于代码生成、文本创作等场景输出 token 占比或较小，可能在 30-40%左右。由于 Agent 的输入含系统提示词、工具定义、记忆文件、全量历史上下文，输出仅为工具调用指令，因此 Agent 场景的输出 token 占比可能并不像想象

中巨大，但是考虑到较多用户使用 OpenClaw 作为生产力工具，而非简单的任务执行，因此我们综合预计 token 的输出占比在 30-40%。

根据主要的海外模型 API 定价，我们估算使用 OpenClaw 的输入定价在 2-3 美元/M tokens，输出定价在 10-15 美元/M tokens，以此测算模型的综合 token 吞吐成本可能在 4.4-7.8 美元/M tokens。

表6 海外大模型厂商的定价

厂商	模型	输入/1M tokens	输出/1M tokens
Anthropic	Claude Opus 4.6	\$5.00	\$25.00
	Claude Sonnet 4.6	\$3.00	\$15.00
	Claude Haiku 4.5	\$1.00	\$5.00
OpenAI	GPT-5.4	\$2.50	\$15.00
	GPT-5.4 (>272K)	\$5.00	\$15.00
	GPT-5.2	\$1.75	\$14.00
	GPT-5	\$1.25	\$10.00
Google	Gemini 3 Pro(≤200K)	\$2.00	\$12.00
	Gemini 3 Pro(>200K)	\$4.00	\$18.00
	Gemini 3 Flash	\$0.50	\$3.00

资料来源：OpenClaw 橙皮书，Anthropic，OpenAI，Google，联储证券研究院

国内大模型的定价虽较海外便宜，但是国内用户的价格敏感度也理应更高，心理价位隐含成本和 API 定价与海外成固定比例。从 API 定价来看，海外模型的均价为 1.91 美元/M tokens，国内模型的均价（不考虑免费的模型）为 0.69 美元/M tokens，对应的价格比约为 2.5: 1。根据 Kimi 的数据，可部署类似于 OpenClaw 服务的产品对应的定价为 159 元/月，而我们前文推算的海外用户的心理价位隐含上限为 40-60 美元/月，二者对应的海外与国内的价格比也在 2.5: 1 左右。因此，我们认为通过海外的月成本和 API 价格推算的每人每月 token 消耗量与国内用户的结果应该是较为接近的。

图7 Kimi 的 Agent 服务定价



资料来源：Moonshot，联储证券研究院

**表7 OpenClaw 支持的大模型价格**

提供商	代表模型	输入价格/1M tokens	输出价格/1M tokens	推荐场景
Anthropic	Claude Sonnet 4.6	\$3.00	\$15.00	Agent 任务效果最佳
OpenAI	GPT-5.4	\$2.50	\$15.00	通用能力强
Google	Gemini 3 Pro	\$2.00	\$12.00	多模态、超长上下文
DeepSeek	DeepSeek-V3.2.2/V4	\$0.14	\$0.28	极致低价、代码任务
智谱 GLM	GLM-5	\$0.80	\$2.56	国产最强代码能力
通义千问	Qwen 3.5 Max	\$1.20	\$6.00	中文 NLP、代码生成
豆包	Seed 2.0 Pro	\$0.47	\$2.37	批量处理、低成本
百度文心	文心 5.0	~\$0.58	~\$1.16	百度云生态用户
Kimi	Kimi K2.5	\$0.60	\$3.00	中文 Agent、长上下文
MiniMax	MiniMax M2.5	\$0.50	\$2.00	SWE-bench 高分、性价比
Ollama	Qwen3.5-Coder:32B	免费	免费	隐私敏感、零成本
LM Studio	Devstral-24B	免费	免费	本地 GUI、模型测试

资料来源：OpenClaw 橙皮书，联储证券研究院

根据上述数据我们测算得到的结果如下所示，OpenClaw 用户每人每月消耗的 token 量预计在 510-1360 万。需注意，该结果对应的是 OpenClaw 用户的人均水平而非全部用户的实际水平区间，根据部分用户在个人媒体上透露的信息，重度使用/24 小时运行的 OpenClaw 的月耗 token 可能在 5000 万以上，该值显著超过我们预测的区间。

**表8 每人每月消耗 token 量（百万）预估**

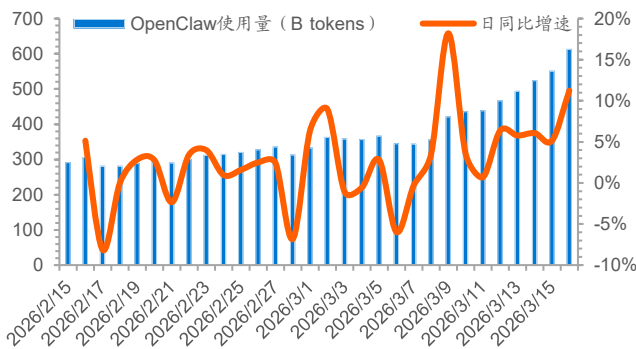
	\$4.4/M tokens	\$5.2/M tokens	\$6.6/M tokens	\$7.8/M tokens
\$40/月	9.1	7.7	6.1	5.1
\$50/月	11.4	9.6	7.6	6.4
\$60/月	13.6	11.5	9.1	7.7

资料来源：联储证券研究院

OpenClaw 爆火过后，在全球用户的大基数面前，渗透率可能也仅有千分之一左右。根据上述测算结果，OpenClaw 用户每人每日消耗的 token 量应该在 17-45 万。根据 openrouter 数据，截至 3 月 16 日，OpenClaw 单日使用量对应的 token 消耗量为 6140 亿。按照该数据进行测算可得，OpenClaw 当前用户数应在 136-361 万人，根据 Xsignal 数据，截至 2025 年 11 月全球的 AI 应用月活跃用户数为 15 亿人，则对应的 OpenClaw 渗透率应在 0.09-0.24%。

然而，从 OpenClaw 的使用量增速来看，渗透空间广阔，带来的增长大可期待。OpenClaw 使用量及日同比增速数据显示，2026 年 3 月以来呈现“量增势稳、增速换挡”的核心亮点：token 使用量自 3 月 1 日的 3340 亿持续攀升至 3 月 16 日的 6140 亿，日同比增速中枢由前期个位数波动抬升至 5%-18% 区间，3 月 9 日单日增速达 18% 的阶段性峰值，增长动能显著增强。从驱动逻辑看，OpenClaw 在国内迅速热度爆发，中国市场或为核心增量来源。

图8 OpenClaw 近一个月的使用量 (tokens)



资料来源: openrouter, 联储证券研究院

图9 全球 AI 应用 MAU



资料来源: Xsignal, 联储证券研究院

我们认为推动 OpenClaw 类产品用户持续增长的核心动能或源于供给端竞争扩容、技术架构迭代、需求端场景渗透及生态基建完善的四维共振。其一，国产厂商正在梯队化推出类 Claw 式产品，以开源生态与高性价比策略加速技术普惠，降低中小用户部署与使用门槛，形成供给侧增量池；其二，OpenClaw 自身架构优化持续深化，如 MCP/A2A 协议落地解决工具调用与多 Agent 协同痛点，端云协同架构提升响应效率等，通过技术壁垒筑牢用户粘性；其三，需求端呈现从“尝鲜式交互”向“刚需式执行”跃迁的自然增长趋势，办公自动化、金融交易、学术研究等垂直场景渗透率持续提升，形成规模化留存与复购；其四，产业生态与基础设施配套完善，国产算力底座降本、合规治理体系建立用户信任，推动用户增长向价值留存演绎。

表9 国内厂商推出的类 Claw 产品

方案	出品方	部署方式	收费
AutoClaw	智谱	本地	加油包 29 元起
EasyClaw	猎豹	本地	40 元/月起
WorkBuddy Claw	腾讯	云端+本地	免费
QClaw	腾讯电脑管家	本地	内测中
Kimi Claw	月之暗面	云端	199 元/月起
MaxClaw	MiniMax	云端	39 元/月起
LobsterAI	网易有道	本地	需要自备 API
飞书妙塔	飞书	云端	免费 (限时)
Xiaomi miclaw	小米	云端+本地	内测中
小艺 claw	华为	云端+本地	内测中

资料来源: 各出品方, 联储证券研究院

根据上述分析, 我们预计 OpenClaw 的到 2026 年年底的持续渗透有望带动全球每月 token 消耗量提升 19.1-489.6T tokens。关键假设如下: 第一, 若当前存在的安全信任危机得以迅速解决, 上下游各厂商适配优化迅速, 则可达到 2026 年底 2-2.5% 的渗透率区间; 第二, 若驱动与制约因素平衡, 行业按照平均增速线性外推, 则对应年底 1-2% 的渗透率区间; 第三, 若在推行使用过程中反复遭遇安全和合规审查, 公众信任度持续下坡, 则对应年底 0.5-1% 的渗透率。根据前文的单个用户每人每月 token 消耗量推算, 则三种情景对应到 2026 年底的 token 消耗量增长区间分别为 133.9-489.6T tokens/月、57.4-387.6T tokens/月、19.1-183.6T tokens/月。

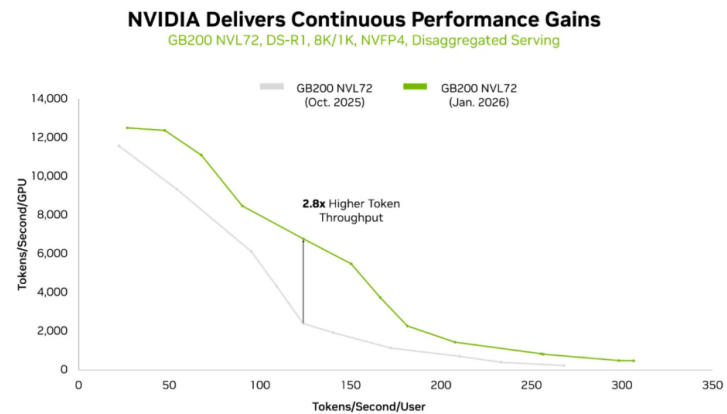
表10 OpenClaw 持续渗透对 token 消耗量增长的敏感性分析

情景	渗透率区间	关键假设	对应月 token 消耗量 (T) 增长
乐观	2-2.5%	安全问题快速解决、巨头补贴加速渗透、大模型能力超预期	133.9-489.6
中性	1-2%	驱动与制约因素平衡，符合行业平均增长规律	57.4-387.6
谨慎	0.5-1%	安全事件反复、合规监管趋严、技术突破不及预期	19.1-183.6

资料来源：联储证券研究院

不同模型在英伟达 GB200 芯片的 token 吞吐速度可能在几百至上万不等。根据英伟达数据，运行在 GB200 NVL72 上的 DeepSeek-R1 在 8K/1K 序列长度下能达到 7000TPS/GPU。在实际情况中，根据模型参数、计算精度、批量并行的不同，该值存在差异，为便于计算取得数据，我们采用 7000 TPS 作为 GB200 芯片的平均 token 处理速率进行推算。

图10 DeepSeek-R1 在 GB200 NVL72 上的 TPS/GPU 情况



资料来源：英伟达，联储证券研究院

考虑到用户使用 Agent 的实际时间主要以工作日的白天为主，我们假设每名用户在工作日的使用时长为 12 小时、在非工作日的使用时长为 6 小时，按照每月共 22 个工作日、8 个非工作日测算全球新增的每秒 token 消耗需求在 170-4360 万。

按照上述的单 GPU 具有 7000TPS 的处理能力计算，上述的 token 消耗需要增加的 GB200 NVL72 机柜数量为 34-865 个，对应的算力芯片数量为 2432-62271 片。

表11 OpenClaw 带来的芯片增量

情景	月 token 消耗量 (T) 增长	需要机柜数量 (个)	需要 Blackwell 芯片数量 (片)
乐观	133.9-489.6	236-865	17027-62271
中性	57.4-387.6	101-685	7297-49298
谨慎	19.1-183.6	34-324	2432-23352

资料来源：联储证券研究院

当前基于单 GPU7000TPS、单用户日均使用量假设的 OpenClawAgent 芯片需求测算，并没有充分反映模型技术迭代、国产算力约束、价格刺激下的需求弹性等核心变量。实际算力芯片增量需求或更高于上述区间，原因在于：

第一，价格战与 Agent 化，刺激单用户 Token 消耗攀升。为争夺用户、提升粘性，主流大模型厂商或将持续下调 API 调用价格、推出更具性价比 Token 套餐或高额度免费

额度，显著降低了用户的 Token 使用成本，刺激单用户日均 Token 消耗从当前水平向更高区间攀升。同时，Agent 场景下的多轮交互、工具调用、长上下文检索等行为，进一步提升了单用户的 Token 生成频率与总量，使得实际 Token 需求增速超过基于当前需求的线性测算，直接推高对算力芯片的总量需求。

**第二，模型参数量扩张，推高单 Token 算力密度。**大模型向更长上下文、更大参数量方向持续迭代，从 70B 级向 400B+甚至万亿参数迈进，同时量化精度压缩的边际效益或有一定程度的削弱，单 Token 生成所需的算力与显存带宽消耗显著提升。在单 GPU 的 TPS 假设不变的前提下，模型扩容将直接导致同等 Token 需求下的芯片数量需求倍增，当前测算未充分纳入中长期模型迭代的算力密度提升效应，进一步放大了实际芯片需求的向上弹性。

**第三，非 N 系芯片算力约束，放大全球算力需求。**当前测算以英伟达 GB200 为基准算力单元，但各大模型厂商采用的推理算力芯片各不相同，而这类芯片单卡 TPS 普遍不及英伟达 B200，若叠加考虑集群互联带宽、软件栈优化成熟度不及英伟达解决方案。在非 N 系芯片需求扩大的背景下，同等 Token 需求需要更多数量的芯片才能匹配，进一步扩大了全球算力芯片的总需求规模，使得实际芯片增量需求显著高于基于英伟达算力的测算结果。

综上，模型参数量扩张、国产芯片算力约束、厂商价格战驱动的单用户 Token 消耗提升等多重因素，将共同推高 OpenClaw Agent 带来的实际算力芯片需求，实际算力增量有望实现一定程度的上修。

## 2.2 中游模型：从推理需求视角看待 Agent 机会

我们认为在 OpenClaw 驱动的 AI Agent 产业浪潮中，大模型厂商的价值分配或将呈现二元分化格局，即头部性能通吃高价值场景+国产性价比下沉普惠市场。两类厂商错位竞争、共同成长，构成 OpenClaw 生态下大模型产业的核心增长主线，主要原因在于：

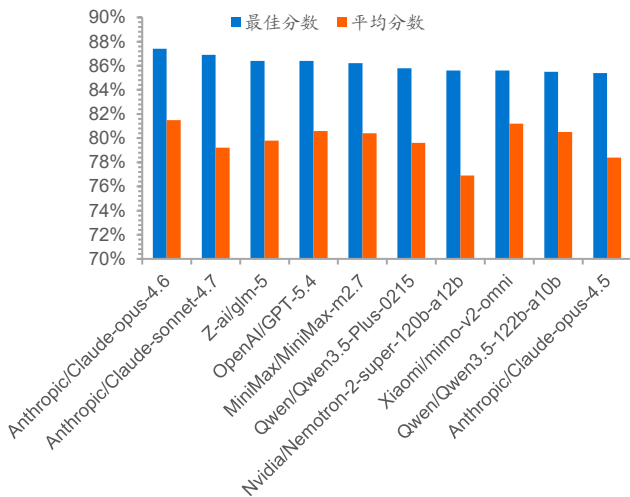
第一，性能顶尖的头部厂商凭借任务执行能力壁垒、高价值客户付费意愿、高 Token 消耗密度，成为企业级 Agent 市场的核心受益者；第二，以国产厂商为代表的性价比路线模型，凭借低成本推理、国产算力适配、下沉市场需求弹性，快速渗透中小企业与 C 端普惠场景，打开增量市场空间。

因此，同时契合两条路线的国产高性能低费用的大模型厂商或为 Agent 方案的主要受益对象。

我们从 PinchBench 的分数排名评价各模型在 Agent 任务上的表现情况，综合来看国内外厂商存在差距，但差距相对较小。PinchBench 是一个用于评估 LLM 模型作为 OpenClaw 编码代理的基准测试系统。在不同模型上运行相同的真实任务，并衡量成功率、速度和成本，帮助开发者选择适合其用例的模型。

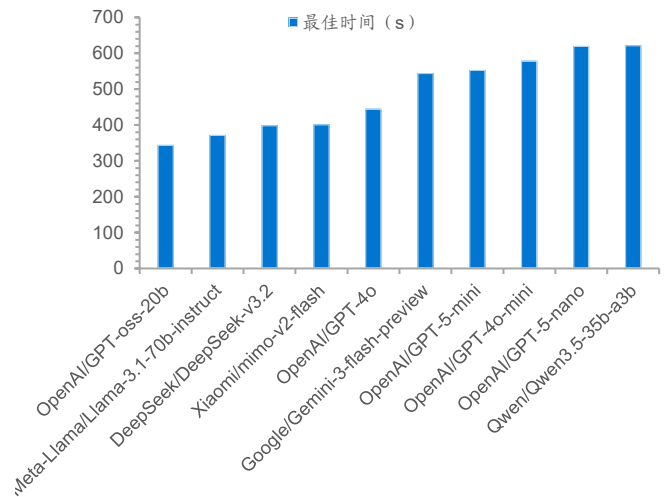
从 OpenCLaw 性能与速度双维度榜单来看，国内外厂商形成一定的差异化竞争格局。海外厂商方面，Anthropic 凭借 Claude 占据性能榜前两位，OpenAI 则以 GPT-5.4 的高性能表现，叠加推理、轻量等型号在速度榜的多点布局，构建起“旗舰性能+多场景速度覆盖”的完整产品矩阵；NVIDIA Nemotron 也凭借开源属性跻身高性能梯队，展现出海外在闭源旗舰与开源高效模型两端的技术优势。国内厂商则呈现“多点突破、均衡追赶”的态势，智谱、MiniMax、小米、千问系列均进入性能榜单前 10 名，同时 DeepSeek、小米与千问的轻量模型在速度榜占据席位，实现了性能与速度的同步跟进，但尚未形成像 GPT 和 Claude 系列那样的多模型布局，产品矩阵密度仍弱于海外。

图11 各模型在 OpenCLaw 上完成任务的得分排名(截至 2026 年 3 月 18 日)



资料来源: PinchBench, 联储证券研究院

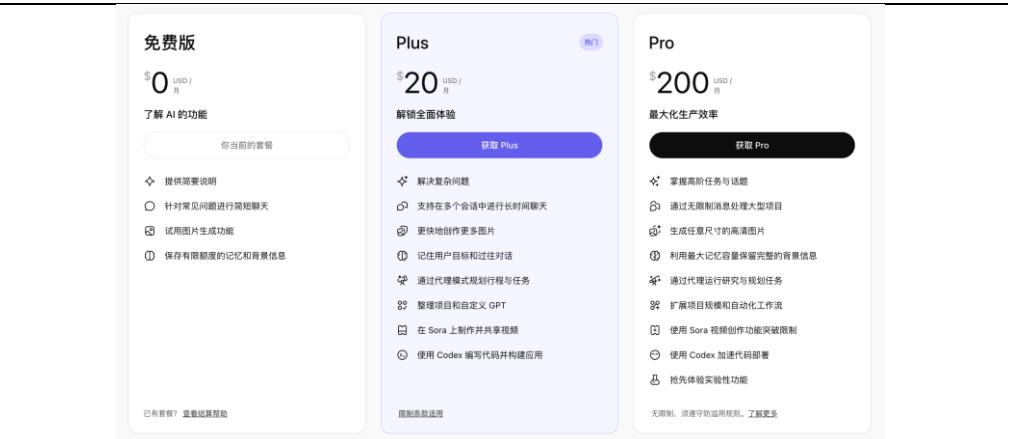
图12 各模型在 OpenCLaw 上完成任务的速度排名(截至 2026 年 3 月 18 日)



资料来源: PinchBench, 联储证券研究院

在 OpenCLaw 面向 B 端的企业级应用场景中, 用户决策优先级应该是“任务执行准确性 > 推理响应速度 > 部署成本”特征。B 端业务对流程稳定性与结果可靠性要求极高, 任务执行成功率是核心刚性约束, 直接关联业务流转效率与风险控制; 推理速度需满足基础业务响应时效要求即可, 优先级次之; 而部署与调用成本的敏感度显著偏低, 企业级预算更倾向于为高稳定性、高准确性的模型方案买单, 对费用支出的容忍度更高, 更看重长期业务价值而非短期成本控制。因此, 以 ChatGPT 方案为例, 企业版用户为保证准确性和响应速度, 往往可以接受相较于个人版本更高的溢价。

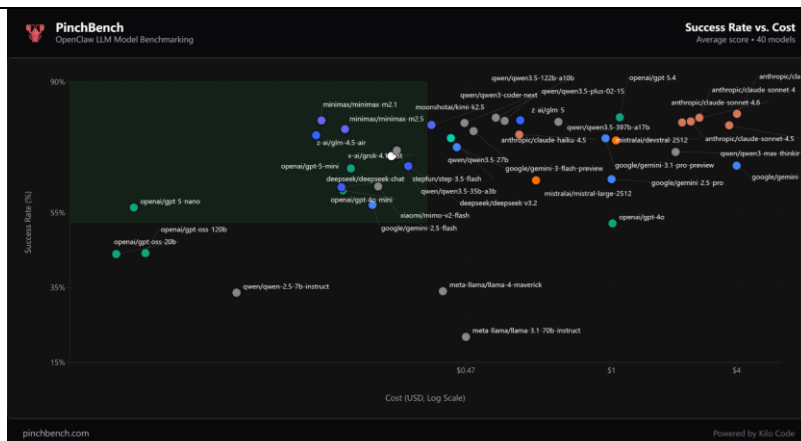
图13 ChatGPT 的订阅价格



资料来源: OpenAI, 联储证券研究院

而在 C 端消费级场景中, 用户核心诉求可能更多聚焦于任务执行成功率、推理响应速度与部署调用成本的三维动态平衡。既不愿为追求极致成功率而承担高昂算力成本与延迟损耗, 也不会因过度压缩成本导致任务完成度不足, 更倾向于选择“高成功率+适配场景的响应速度+可控边际成本”的组合方案, 正因为上述原因, 可以看到在 PinBench 的性价比象限中, 国产智谱、MiniMax、DeepSeek、千问等模型占据了接近半数的性价比优选方案, 因此我们认为国产厂商在 2C 端有望凭借更极致的性价比路线实现低成本获客, 进而有望实现差异化布局出海。

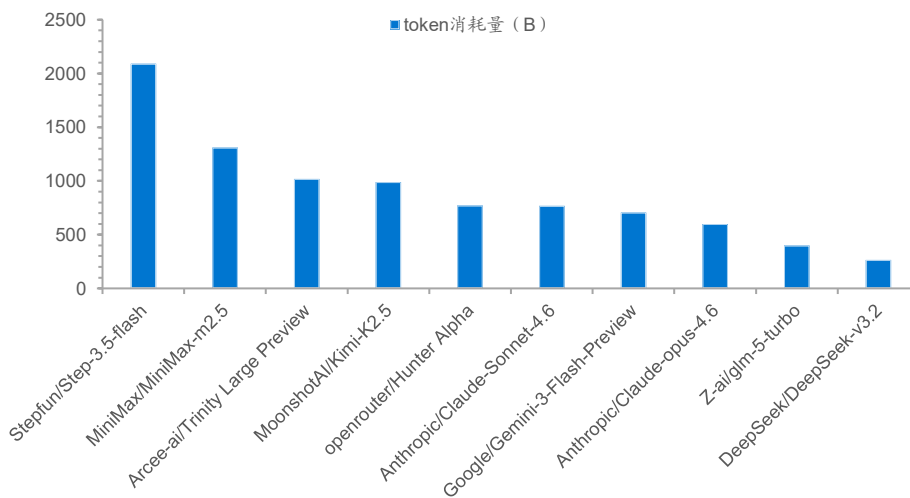
图14 各模型在 OpenClaw 上的性能/价格象限



资料来源：PinchBench，联储证券研究院

从模型使用量数据来看，C 端用户在模型选择上的决策逻辑与上述判断基本相符。根据 openrouter 的 OpenClaw 模型使用量数据，国产模型本月呈现显著的领跑态势，阶跃星辰以 2.09Ttokens 用量登顶榜单，MiniMax、月之暗面、智谱、DeepSeek 等国产品号集体跻身前十，合计用量占据榜单头部半壁江山。这一分布特征直接印证了 C 端用户对成本与体验的平衡诉求——国产模型凭借在保持任务完成度与适配场景响应效率的同时，实现了更具竞争力的部署与调用成本结构，精准匹配了 C 端用户不愿为极致性能承担过高成本、亦不愿因成本压缩牺牲基础任务完成度的核心需求，也进一步验证了国产厂商在 2C 端大有作为的可能性。

图15 本月 OpenClaw 部署模型的使用量排名



资料来源：openrouter，联储证券研究院

### 2.3 下游应用：现阶段或以生产力提升为主，大规模商用还需等待

政策端持续加码成为 AI Agent 产业发展的核心驱动力。“人工智能+”行动明确培育智能体即服务新业态定调产业方向，广东率先出台全国首个省级人工智能 OPC (One-Person Company, 一人公司) 专项政策筑牢发展基底，深圳、合肥、无锡等多地进一步将 OpenClaw 等 Agent 框架与 OPC 模式深度绑定，构建起多级政策支持体系，为 Agent 产业规模化发展扫清制度障碍。

各地政策从算力供给、落地成本、场景开放、要素保障多维度精准发力，算力券、

免费部署、专项补贴等举措有效降低 Agent 研发与部署门槛，场景池建设、跨境布局拓展其在工业、文创、跨境电商等领域的应用边界，人才贷、算力贷等专属金融产品与生态社区建设则完善了产业要素供给与集聚发展基础，同时政策对开源生态的鼓励，更让国产 Agent 框架与模型迎来差异化发展机遇。

**表12 近期各地方相关部门发布的鼓励 OpenClaw 及 OPC 发展的政策**

地区	政策名称	发布时间	核心支持内容	最高补贴金额
广东省	《广东省支持人工智能 OPC 创新发展行动方案（2026—2028 年）》	2026.3.16	全国首个省级人工智能 OPC 专项政策；打造粤港澳大湾区算力“一张网”+完善算力券制度；开放公共数据、打造高质量数据集/语料库；提供免费模型接口、省级专项资金叠加支持垂直领域模型研发；打造 OPC 生态社区并开展省级认定；拓展粤港澳跨境及东南亚/葡语国家出海场景；推出人才贷/研发贷/算力贷等专属金融产品；完善 OPC 人才认定、安居、培训体系；构建知识产权保护、法律服务、出海全流程服务平台；2028 年目标建成百个 OPC 生态社区、培育千家标杆企业	未明确单项目最高补贴，含算力券、省级专项资金叠加支持，推出多类专属信贷产品
合肥高新区	《合肥高新区打造人工智能 OPC 创业生态示范区行动计划（征求意见稿）》（"OPC 十五条"）	2026.3.6	全方位护航 OpenClaw 等开源 AI 项目落地；打造“AI+超级个体/一人公司(OPC)”新业态标杆；提供创业孵化、人才公寓、投融资全链条支持	1000 万元（算力券）
深圳龙岗区	《深圳市龙岗区支持 OpenClaw&OPC 发展的若干措施（征求意见稿）》（"龙虾十条"）	2026.3.7	鼓励“龙虾服务区”免费提供 OpenClaw 部署服务；开放高质量脱敏公共数据并减免使用费；对 OpenClaw 技能包开发、具身智能应用给予补贴；提供 OpenClaw 专属算力支持和创业空间	200 万元（技能开发补贴）
无锡高新区	《关于支持 OpenClaw 等开源社区项目与 OPC 社区融合发展的若干措施（征求意见稿）》（"养龙虾 12 条"）	2026.3.9	对提供 OpenClaw 免费部署的云平台给予全额补贴；支持 OpenClaw 在工业质检、设备维护等制造场景落地；对关键技术突破给予专项支持；提供安全合规咨询服务	500 万元（单项支持）
乌鲁木齐高新区(新市区)	《支持 OPC 与 OpenClaw 发展的十条措施》	2026.3.13	聚焦 OpenClaw 开源生态，实施“零成本启动、全要素赋能”；培育数据场景应用、AI、软件研发领域“超级个体”；覆盖载体建设、注册服务、技术支撑、人才保障全维度	未明确具体金额

资料来源：广东省发改委，合肥高新发布公众号，深圳市龙岗区，无锡高新区在线公众号，天山火炬公众号，联储证券研究院

从技能库来看主要使用者还是以生产力需求者为主，开发者居首，因此我们预计 Agent 的应用短期仍将以软件开发为主。根据 OpenClaw 橙皮书数据，从 OpenClaw 技能分类的分布结构来看，平台核心用户群体高度集中于技术开发者范畴：编码 Agent 与 IDE、Web 与前端开发、DevOps 与云、CLI 工具、Git 与 GitHub 等开发者导向型技能占据榜单前 10 位中的 8 个核心席位，合计数量占到该截止日期总技能数据的 26.8%。覆盖代码生成调试、前端组件开发、云原生运维、版本控制等全链路开发场景，合计数量显著领先于通用消费场景。这一分布特征清晰印证，当前 OpenClaw 的生态需求仍由程序员、DevOps 工程师、前端开发者等技术群体主导，平台功能与用户诉求深度绑定开发辅助场景，开发者仍是其最核心的用户基底。

**表13 ClawHub 技能库分类 Top10 (截至 2026 年 3 月中旬)**

排名	分类	数量	说明
1	编码 Agent 与 IDE	1222	代码生成、调试、重构等开发辅助
2	Web 与前端开发	938	HTML/CSS/JS 生成、组件开发
3	DevOps 与云	408	Docker、K8s、CI/CD 管理
4	搜索与研究	350	联网搜索、信息汇总
5	浏览器与自动化	335	网页操作、表单填写、截图
6	生产力与任务	206	日程、待办、项目管理
7	AI 与 LLM	197	提示工程、模型切换、多 Agent 协作
8	CLI 工具	186	终端命令增强、系统管理
9	Git 与 GitHub	170	仓库管理、PR 审查、Issue 处理
10	图片与视频生成	169	AI 绘图、视频处理

资料来源: OpenClaw 橙皮书, 联储证券研究院

但是从生产力解放的角度来看, **OpenClaw** 还难以达成完全替代人类员工的目的。与专门设计用于 AI 编程的 Claude Code 相比, OpenClaw 还不具备完全替代全场景生产力工具的可行性: 其核心价值在于作为数字生活中枢, 可高效覆盖日常自动化、跨平台协作等轻量生产力需求, 但在专业编码等核心生产力场景中能力局限明显, 难以独立承担复杂软件工程任务, 因此对于专业的技术开发者而言, 用 OpenClaw 管理数字生活, 用 Claude Code 管理代码库, 两者组合或为 2026 年最完整的 AI 驱动 workflows。

**表14 OpenClaw 与 Claude Code 对比**

维度	OpenClaw	Claude Code
定位	通用 AI 生活助手/Life OS	专业编程 Agent
运行环境	自托管服务器, 消息平台网关	终端 CLI/Web/Desktop
连接对象	20+通信/办公平台	代码库、文件系统
记忆系统	四层记忆(SOUL/TOOLSUSER/Session), 长期可持续	会话级+CLAUDE.md 持久化
Skill 系统	ClawHub 市场, 动态插件化	静态规则文件触发
Token 消耗	高(多轮思考+多工具调用, 可能是传统聊天天的几十倍)	相对低
安全模型	自托管, 需自行维护安全。已出现 CVE 漏洞和供应链攻击	Anthropic 托管沙盒, 权限细粒度控制
模型支持	多模型(Claude/GPT/DeepSeek/Ollama 等)	仅 Claude
开源/费用	MIT 开源免费, 自付 API 费用	闭源 CLI, 按 API 计费(\$20/月起)
编程能力	一般, 简单任务可以	强, 专为编程优化
日常自动化	强, 多平台接入, 长期在线	弱, 主要在终端内使用
定制性	完全开源, 可改 system prompt、fork 整个代码库	通过 instruction 文件有限定制

资料来源: OpenClaw 橙皮书, 联储证券研究院

**OpenClaw 作为当前快速落地的 AI 智能体框架, 其安全体系存在显著且多元的风险敞口, 对于商业化应用也形成显著压制。**据网络信息安全漏洞数据库统计, OpenClaw 暴露出多个中高危漏洞, 其安全问题主要集中在权限管控逻辑缺陷、沙箱机制绕过、网络防护不完善、认证校验不严谨四大方面, 且多数漏洞因框架的分布式执行特性和多渠道交互设计被放大。目前 OpenClaw 漏洞类型以远程代码执行、命令注入、信息泄露、权限越权为主, 攻击者可利用相关漏洞实现未授权远程操作、全量数据窃取, 若公网暴露的实例未及时修复, 攻击者可直接实现未授权远程代码执行, 对企业和个人数据安全、系统运行造成严重威胁。

**表15 OpenClaw 存在的高危漏洞（截至 2026 年 3 月 10 日）**

漏洞类型/名称	CVSS 评分	核心危害
跨域重定向漏洞	7.5	授权凭证泄露，导致未授权 API 调用、数据窃取、AI 代理被控制
信息泄露漏洞	7.1	网关认证信息泄露，可直接接管网关、控制整个 AI 代理系统、窃取本地全量数据
远程代码执行漏洞	8.8	已在野利用，实现未授权远程代码执行，可横向渗透至内网其他设备
命令注入漏洞	8.1	无需身份校验即可执行任意系统命令，服务器部署场景可直接获取服务器权限
命令注入漏洞	7.8	突破沙箱隔离，可植入恶意程序，对设备造成永久性破坏

资料来源：奇安信，联储证券研究院

整体来看，OpenClaw 短期应用核心聚焦数字执行层工种的部分人力解放，但其规模化落地及 OPC 模式的全面普及仍需长期积淀，数字员工的具象化生成更依赖头部厂商生态赋能且落地节奏偏缓。我们认为当前 OpenClaw 的核心应用场景，在于对规则驱动、高重复性、纯数字化操作属性的“数字执行层”工种实现一定程度的人力解放，这一应用方向与 OpenClaw 所具备的“自主拆解任务—调用工具—执行操作—结果反馈”全流程自动化核心能力形成高度适配，这一路径或将以影响“生产力/人力成本”的形式进一步修正企业利润，后续可以通过观察这类员工占比较高的企业的利润改善进度判断。

但从产业实际生产力成熟度与平台安全风险事件频发的现状来看，数字员工的大批量落地应用乃至 OPC 模式的全面涌现，仍需要经历较长周期的技术打磨、生态构建与安全体系完善。现阶段其落地形态或更多以试点性质展开，且可能更多集中于规模较小、商业模式相对简单的企业主体。而数字员工的真正具象化生成，仍需依托头部云厂商及大模型厂商完成 Agent 技术的深度开发、内部规模化验证与行业定制化解决方案的打磨落地，我们预计该关键节点的到来具备滞后性，落地节奏并不会过于迅速。

### 3. 投资建议

我们认为 OpenClaw 的出现标志着人工智能实现了被动承接到主动发起的范式转变，有望打通 Agent 在 B 端赋能、C 端应用的最后一公里。其高 Token 消耗、全平台适配及生态快速扩张的特征，深刻重构了 AI 产业链上下游的供需格局，为行业打开全新增长空间。

上游算力赛道有望从需求端带动增长，随 Agent 持续渗透形成长期趋势。OpenClaw 的 Token 消耗特性大幅抬升算力芯片、算力基础设施的长期需求，且当前算力测算未充分覆盖模型参数量扩张带来的单 Token 算力密度提升、国产算力芯片单卡性能约束、厂商价格战刺激的单用户 Token 消耗攀升等因素，实际增量或更高于测算值。AI Agent 的长期渗透将持续抬升算力需求的增长斜率，算力基建的国产化替代与高端算力芯片的产能落地，将成为这一赛道的核心投资主线。建议关注：国产算力芯片、AI Infra 等相关厂商。

中游大模型行业或呈现二元分化格局，B 端以全球头部厂商为主，C 端有望为国产厂商提供更多机会。兼具高性能与低费率的国产模型，不仅有望在 C 端实现规模化渗透，还具备依托性价比优势出海的潜力，而针对 Agent 场景的模型推理优化、工具调用能力升级，将成为后续大模型厂商的核心竞争力。建议关注：国产大模型、AI 基础软件等相关厂商。

下游应用在政策加持下场景落地提速，但是规模化商业应用更需要重视安全与技术的迭代升级。多地政策密集加码，各种补贴形式持续降低研发与部署门槛，OPC 模式与 OpenClaw 的深度绑定也打开想象空间，短期来看，下游机会集中在能借助 AI Agent 实现降本增效的办公自动化、开发运维等垂直场景，而长期规模化商用则依赖技术迭代，具备场景落地能力与安全技术解决方案的厂商或将率先受益。建议关注：大型云厂商，垂直场景解决方案商等相关厂商。

## 4. 风险提示

Agent 需求扩张速度不及预期；技术开发进度不及预期；风险事件缓释进度不及预期；政策出台及落地进度不及预期

### 免责声明

联储证券股份有限公司（以下简称“本公司”）具有中国证监会许可的证券投资咨询业务资格。本报告仅供本公司的客户使用。本公司不会因接收人收到本报告而视其为客户。在任何情况下，本公司不对任何人因使用本报告中的任何内容所引致的任何损失负任何责任。

本报告的信息均来源于本公司认为可信的公开资料，但本公司及其研究人员对该等信息的准确性及完整性不作任何保证。本报告中的资料、意见及预测仅反映本公司于发布本报告当日的判断，可能会随时调整。在不同时期，本公司可发出与本报告所载资料、意见及推测不一致的报告。本公司不保证本报告所含信息及资料保持在最新状态，对本报告所含信息可在不发出通知的情形下做出修改，投资者应当自行关注相应的更新或修改。

本公司力求报告内容客观、公正，但本报告所载的资料、工具、意见、信息及推测只提供给客户作参考之用，不构成任何投资、法律、会计或税务的最终操作建议，本公司不就报告中的内容对最终操作建议做出任何担保。本报告中所指的投资及服务可能不适合个别客户，不构成客户私人咨询建议。投资者应当充分考虑自身特定状况，并完整理解和使用本报告内容，不应视本报告为做出投资决策的唯一因素。

投资者应注意，在法律许可的情况下，本公司及其本公司的关联机构可能会持有本报告中涉及的公司所发行的证券并进行交易，也可能为这些公司正在提供或争取提供投资银行、财务顾问和金融产品等各种金融服务。

本报告版权归“联储证券股份有限公司”所有。未经事先本公司书面授权，任何机构或个人不得对本报告进行任何形式的发布、复制。任何机构或个人如引用、刊发本报告，需注明出处为“联储证券研究院”，且不得对本报告进行有悖原意的删节或修改。

### 分析师声明

本报告署名分析师在此声明：我们具有中国证券业协会授予的证券投资咨询执业资格或相当的专业胜任能力，本报告所表述的任何观点均精准地反映了我们对标的证券和发行人的个人看法，结论不受任何第三方的授意或影响。我们所得报酬的任何部分无论是在过去、现在及将来均不会与本报告中的具体投资建议或观点有直接或间接联系。

### 投资评级说明

投资建议的评级标准		评级	说明
评级标准为报告发布日后的6个月内公司股价（或行业指数）相对同期基准指数的相对市场表现。其中A股市场以沪深300指数为基准；新三板市场以三板成指（针对协议转让标的）或三板做市指数（针对做市转让标的）为基准；香港市场以恒生指数为基准。	股票评级	买入	相对同期基准指数涨幅在10%以上
		增持	相对同期基准指数涨幅在5%~10%之间
		中性	相对同期基准指数涨幅在-5%~+5%之间
		减持	相对同期基准指数跌幅在5%以上
	行业评级	看好	相对表现优于市场
		中性	相对表现与市场持平
		看淡	相对表现弱于市场

### 联储证券研究院

#### 青岛

地址：山东省青岛市崂山区香港东路195号8号楼11、15F  
 邮编：266100

#### 上海

地址：上海市浦东新区滨江大道1111弄1号中企国际金融中心A栋12层  
 邮编：200135

#### 北京

地址：北京市朝阳区安定路五号院13号楼B座14层  
 邮编：100029

#### 深圳

地址：广东省深圳市南山区沙河街道深云路2号侨城一号广场28-30F  
 邮编：518000