

ChatGPT 的流量去哪了？

计算机行业

我们认为对 ChatGPT 流量的统计存在显著低估；同时，仅通过应用访问量判断 AI 市场热度的观点，并未深刻理解 Open AI 等巨头的商业战略和产业发展。我们认为，所谓 ChatGPT 的“流量瓶颈”，主要由以下三方面原因导致，分别是 1、Open AI 战略，流量高低在 API 结构；2、算力短缺；3、访问时长（ChatGPT 限流）+访问人数（部分地区封锁账号）。

Open AI 的战略是避免与客户竞争争夺流量，希望更多客户接入 API，ChatGPT 的愿景是成为一个超级智能的工作助手。我们认为今年席卷的 AI 浪潮本质上是生产力的革命，一个爆款现象级应用的流量暴涨仅仅是一个产业开启的前奏。关注第二第三梯队应用发展更能够客观感受到整个行业的繁荣。Open AI 受限于 GPU，算力仍为稀缺资源，整体行业需求依旧有较大释放空间。OpenAI 存在对一些地区产品使用的限制，我们认为若将来未覆盖市场进一步开放，有望看到流量的二次增长曲线。

大模型型应用加速落地，桌面端+移动端构筑 AI 生态。目前，Open AI API 接口+plugin 插件打造 AI 助手已初步落地，从网络浏览器、代码解释器到第三方工具，联网功能全场景应用。此外，ChatGPT 已经推出 iOS 版应用程序，移动端已经推出联网模式，有望开启 ChatGPT 第二波成长曲线。此外，Open AI 持续深化布局，据 The Information 最新消息，OpenAI 正准备发布一个类似 App Store 针对定制化模型/软件的交易市场平台。我们认为随着桌面端的布局日臻完善，移动端将成为大模型厂商的主战场。

我们判断 AI 的趋势正从供给侧转向应用侧。大模型具有“预训练+精调”等功能，显著降低 AI 开发门槛，即“低成本”和“高效率”。模型侧，以 Open AI 为例，模型迭代进度明显加快，GPT-4、多模态、AI 生成图片、AI 生成视频等算法逐步点燃 AI 领域，模型迭代速度持续加快，我国大模型“自研热”持续，自研速度处于加速状态。因此我们认为在供给侧持续高爆发的情况下，应用侧有望爆发，服务场景进一步拓宽，成为真正意义上解放生产力的双手，我们认为数据有望成为应用侧的核心壁垒。

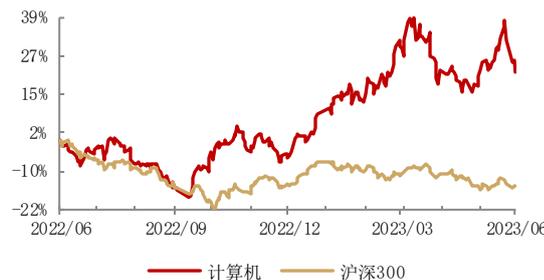
投资建议

我们认为 OpenAI 并未陷入流量“瓶颈”，算力侧和应用侧同样前景广阔，相关受益标的：

评级及分析师信息

行业评级：推荐

行业走势图



分析师：刘泽晶

邮箱：liuzj1@hx168.com.cn

SAC NO: S1120520020002

- 1) AI 芯片厂商，相关受益标的为：寒武纪、海光信息、景嘉微、龙芯中科等；
- 2) AI 服务器厂商，相关受益标的为：中科曙光、神州数码、拓维信息、工业富联、浪潮信息等；
- 3) AI 云厂商，相关受益标的为：首都在线、鸿博股份、青云科技、优刻得、光环新网、新炬网络等；
- 4) AI+办公：金山办公、福昕软件、泛微网络、彩讯股份、鼎捷软件等；
- 5) AI+金融：同花顺、财富趋势、新致软件等；
- 6) AI+医疗：润达医疗、万达信息、卫宁健康等；
- 7) AI+教育：佳发教育、鸥玛软件、科大讯飞等；
- 8) AI+军工：航天宏图、中科星图、华如科技等；
- 9) AI+其他：焦点科技、华宇软件等。

风险提示

核心技术水平升级不及预期的风险，AI 伦理风险，政策推进不及预期的风险，中美贸易摩擦升级的风险。

正文目录

1. ChatGPT 流量复盘及分析.....	4
1.1. AI 应用流量变化情况概览.....	4
1.2. Open AI 剑指消费级企服平台：“ChatGPT+API”新结构.....	5
1.3. Open AI 受限于 GPU，算力仍为稀缺资源.....	9
1.4. OpenAI 限流+账号封锁，“主动”控制流量.....	12
2. 大模型型应用加速落地，桌面端+移动端构筑 AI 生态.....	14
2.1. ChatGPT 移动端布局概览.....	14
2.1.1 初步落地：API 接口+plugin 插件打造 AI 助手.....	14
2.1.2 正式登陆：ChatGPT 推出 iOS 版应用程序.....	16
2.1.3 深化布局：搭建 OpenAI 版应用商店.....	17
3. 投资建议.....	18
4. 风险提示.....	19

图目录

图表 1 AI 应用访问量变化情况（单位：百万）.....	4
图表 2 截至 2023 年 6 月 25 日 ChatGPT 桌面端及移动端访问量.....	5
图表 3 Character.ai 可实现自定义自己的 bot.....	6
图表 4 Character.ai 可访问已设置好的游戏角色.....	6
图表 5 3 月-6 月 Character.ai 访问量.....	6
图表 6 3 月-6 月 Character.ai 周度访问停留时长.....	7
图表 7 LlamaIndex 是一个用于大语言模型的数据框架.....	7
图表 4 5 月-6 月 LlamaIndex 周度访问量.....	7
图表 9 Hugging Face 界面.....	8
图表 10 HuggingFace 将 AI 项目研发流程标准化.....	8
图表 11 6 月 3 日-6 月 16 日 AI 相关网站访问量统计.....	9
图表 12 当前 OpenAI 提供 API 的分类与价格.....	10
图表 13 ChatGPT 运行的算力需求及成本测算.....	11
图表 14 GPT-4 发布时提及但仍未开放的图片识别能力.....	12
图表 15 ChatGPT 平均访问时长保持在 6-9 分钟.....	13
图表 16 包括中国在内的 7 个国家无法访问 ChatGPT.....	13
图表 17 Chat.openai.com 前五大使用地区流量变化.....	13
图表 18 ChatGPT 首批“插件商城”.....	14
图表 19 Wiz AI 官网页面.....	15
图表 20 Wiz AI 收费标准.....	15
图表 21 ChatGPT 首批“插件商城”.....	15
图表 22 用户自行选择需要的第三方插件.....	16
图表 23 用于管理待办事项的插件.....	16
图表 24 检索最新奥斯卡奖的信息的网络浏览器插件.....	16
图表 25 执行逻辑计算的代码解释器插件.....	16
图表 26 iOS 端 ChatGPT App 前五天下载量达 48 万次.....	17
图表 27 跻身美国 AI 聊天机器人 App 前五名.....	17
图表 28 OpenAI 将打造 App Store.....	18

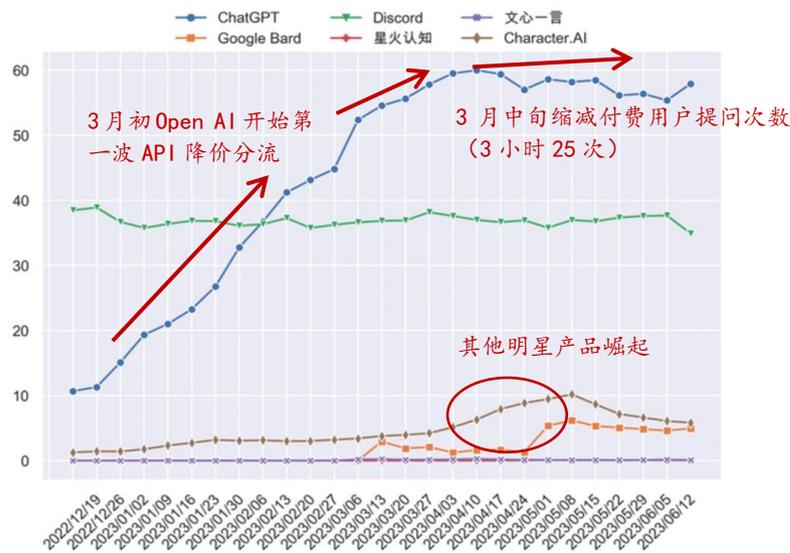
1. ChatGPT 流量复盘及分析

1.1. AI 应用流量变化情况概览

ChatGPT 自发布以来，访问量从去年 12 月至今年 3 月一直呈现高速增长态势，但从下图趋势可见，从今年 4 月开始访问量增速明显放缓，在 5 月、6 月期间甚至出现单周负增长的现象（根据 SimilarWeb 数据，1 月-5 月 ChatGPT 访问量环比 +131.6%/+62.5%/+55.8%/+12.6%/+2.8%）。因此，市场出现担心 AI 热减退、ChatGPT 被高估的声音。

我们认为对 ChatGPT 流量的统计存在显著低估；同时，仅通过应用访问量判断 AI 市场热度的观点，并未深刻理解 Open AI 等巨头的商业战略和产业发展。现阶段，单从 ChatGPT 访问量判断行业发展已经相对片面，应当从更长期的视角甄别数据理解底层逻辑。

图表 1 AI 应用访问量变化情况（单位：百万）



资料来源：华尔街见闻，华西证券研究所

图表 2 截至 2023 年 6 月 25 日 ChatGPT 桌面端及移动端访问量



资料来源: SimilarWeb, 华西证券研究所

我们认为, 所谓 ChatGPT 的“流量瓶颈”, 主要由以下三方面原因导致, 在下文中我们将详细分析这些观点, 并对接下来 AI 应用的发展趋势和特征给出判断:

- 1、Open AI 战略在于避免与客户竞争, 下一个流量高地 API 接口
- 2、GPU 短缺算力限制
- 3、访问时长 (ChatGPT 限流) + 访问人数 (部分地区封锁账号)

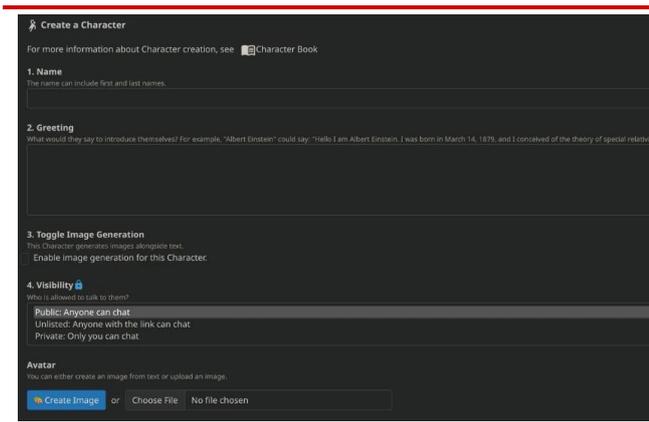
1.2. Open AI 剑指消费级企服平台: “ChatGPT+API” 新结构

首先, 需强调 Open AI 的战略是避免与客户竞争争夺流量, 希望更多客户接入 GPT API, 希望客户自己的产品通过使用 API 变得更好。据华尔街见闻, 很多的开发者说, 他们在使用 OpenAI 的 API 构建产品时, 担心 OpenAI 会发布竞品, 但 OpenAI 未来不会发布超过 ChatGPT 的更多产品。ChatGPT 的愿景是成为一个超级智能的工作助手。

在理解 Open AI 重点不在 C 端的战略背景下, 进一步分析流量统计逻辑。由于 SimilarWeb 的技术限制, 上文图 1 的访问量数据不包含 OpenAI API 接口的流量, 仅仅包括 ChatGPT 应用网站的访问量, 这意味着随着 OpenAI 接口生态完善, 该统计数据偏离更显著。因此, 我们从 1) 其他 AI Chat 网站热度、2) To B 企业端对 AI 模型的消费需求、3) 第二三梯队 AI 应用爆发三个维度追踪行业现状。

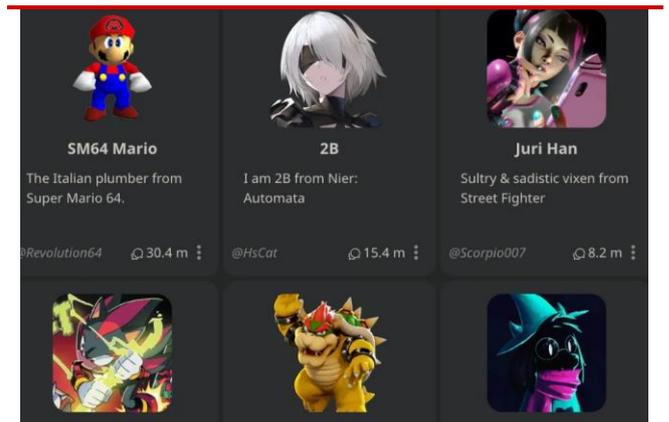
以近期新的明星产品 Character.ai 为例, 它是一个目前主要针对娱乐需求的 AI Chatbot (排除 API 统计口径干扰)。Characterai 公司致力于实现与计算机进行无限对话, 正在研发下一代的对话机器人技术, 应用范围广泛, 包括娱乐、教育、常规问答等。该公司的“对话机器人基于自有的大型语言模型技术, 从头开始构建和训练, 专为对话而设计。用户负责编写一个角色的台词, 机器人负责创作另一个角色的台词, 给用户一种在与另一个角色对话的体验。

图表 3 Character.ai 可实现自定义自己的 bot



资料来源：知乎，华西证券研究所

图表 4 Character.ai 可访问已设置好的游戏角色



资料来源：知乎，华西证券研究所

通过 Character.ai 分析 AI Chat 未失真的娱乐需求数据，根据 SimilarWeb 统计，今年 3 月第一周的访问量约为 21.4M，到 5 月的第三周其访问量已经增至 70.6M，增加超三倍。同时，近三个月内 Character.ai 的用户平均访问时长接近 30 分钟较为稳定。

图表 5 3 月-6 月 Character.ai 访问量



资料来源：SimilarWeb，华西证券研究所

图表 6 3月-6月 Character.ai 周度访问停留时长



资料来源: SimilarWeb, 华西证券研究所

ChatGPT 降低费用门槛, To B 端企业集成部署需求旺盛。据财联社消息,今年3月2日,ChatGPT官方API发布,价格为1ktokens/\$0.002,即每输出100万个单词需要2.7美元(约18元人民币),比此前GPT-3.5模型价格降低90%。我们认为开放API、降低Token费用,能够大幅促进更多应用接入ChatGPT,丰富生态,这也可以解释在该阶段造成对ChatGPT的分流(可再次回顾图1)。

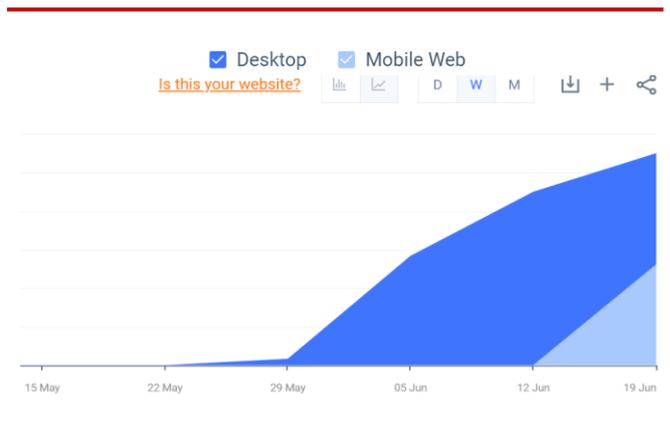
以llamaindex.ai为例(一定程度上反映企业端AI大模型需求),该网站前身为GPT Index。LlamaIndex是一个数据框架,可以为大语言模型提供数据摄取、数据索引和查询引擎。LlamaIndex可以将大模型和企业数据连接起来,做定制化的LLMs,管理数据流和任务调度。根据华尔街见闻消息,llamaindex.ai周访问量同比增长122.69%。和它类似的工具还有LangChain,现在全球越来越多的公司开始把自然语言融入产品进行交互,而新的开发方式意味着新的技术栈。

图表 7 LlamaIndex 是一个用于大语言模型的数据框架



资料来源: 锦缎研究院, 华西证券研究所

图表 8 5月-6月 LlamaIndex 周度访问量

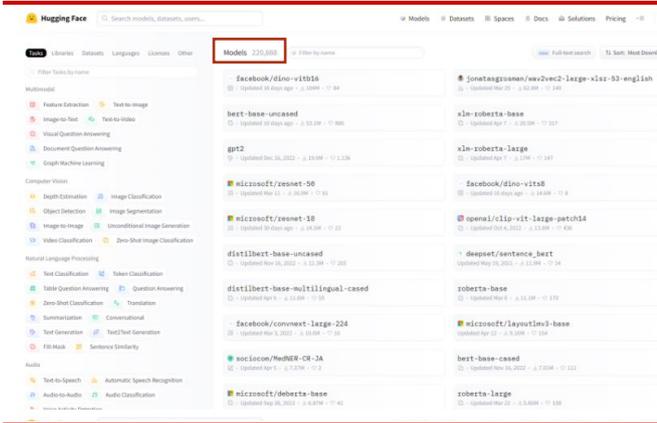


资料来源: SimilarWeb, 华西证券研究所

对于To B端的企业需求分析,还可以通过开发者社区的活跃程度侧面印证。以Hugging Face为例,它是机器学习开源参考模型的强大引擎,使得开发者可以构建、训练并部署最先进的模型。这是一个高速发展的社区,据NLP日志录数据,

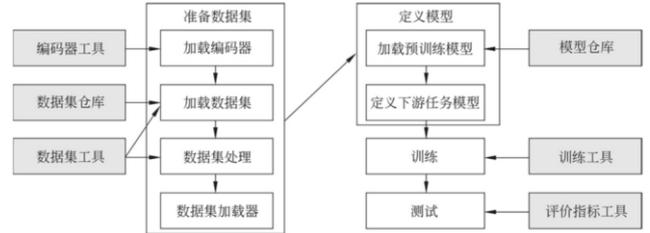
包括 Meta、Google、Microsoft、Amazon 在内的超过 5000 家组织机构在为 HuggingFace 开源社区贡献代码、数据集和模型。目前包括模型 236,291 个，数据集 44,810 个。刚开始大多数的模型和数据集是 NLP 方向的，但图像和语音的功能模型正在快速更新中。Hugging Face 这类开发社区的野蛮生长同样意味着 B 端和部分 C 端加速应用部署模型态势不减。

图表 9 Hugging Face 界面



资料来源：AI 工具知识全百科，华西证券研究所

图表 10 HuggingFace 将 AI 项目研发流程标准化



资料来源：NLP 日志录，华西证券研究所

我们认为今年席卷的 AI 浪潮本质上是生产力的革命，一个爆款现象级应用的流量暴涨仅仅是一个产业开启的前奏。现阶段，除了以 ChatGPT 为代表的标杆应用，关注第二第三梯队应用发展更能够客观感受到整个行业的繁荣。目前有越来越多的应用开始发力，几乎每周都有新的爆款。

图表 11 6月3日-6月16日 AI 相关网站访问量统计

网站域名	平均访问量(万)	访问量同比变	近四周平均留时长(分钟)
anthropic.com	6.91	-1.86%	4.84
openai.com	5788.35	-0.48%	5.93
cohere.com	3.42	4.63%	3.30
ai21.com	0.34	-36.09%	0.07
amd.com	41.74	1.58%	5.06
nvidia.com	99.40	-2.27%	3.90
pytorch.org	10.24	-0.77%	3.40
tensorflow.org	8.00	-3.95%	4.22
fixie.ai	0.09	-53.34%	6.60
llmaindex.ai	0.04	122.69%	0.27
langchain.com	10.85	-0.61%	7.45
agpt.co	1.72	-13.44%	5.04
keras.io	2.12	-3.93%	3.55
caffe.berkeleyvision.org	0.06	39.34%	2.60
developers.generativeai.google	3.06	-11.17%	3.52
kaggle.com	36.45	-2.90%	8.37
github.com	1407.50	1.02%	6.94
huggingface.co	78.26	3.37%	5.99
replicate.com	31.22	-10.68%	4.86
h2o.ai	0.72	28.01%	1.99
opencv.org	4.19	-4.71%	3.69
hiascend.com	0.37	-4.72%	9.07
chat.openai.com	5624.51	-0.18%	7.34
bard.google.com	471.65	-6.97%	3.99
discord.com	3757.05	0.01%	2.72
yiyan.baidu.com	9.61	2.75%	2.68
xinghuo.xfyun.cn	7.87	2.31%	3.27
character.ai	653.71	-7.76%	22.86
yige.baidu.com	2.41	7.36%	4.73
printidea.art	0.39	4.13%	2.62
adobe.com	881.53	0.89%	3.79
midjourney.com	95.17	-5.53%	7.44
stability.ai	9.04	-9.77%	2.30
stablediffusionweb.com	19.80	-3.15%	3.27
remove.bg	189.93	0.36%	5.88
tongyi.aliyun.com	3.30	-6.48%	2.17
voice.ai	7.63	-3.09%	3.26
notion.so	478.64	-0.78%	6.73
slack.com	376.62	1.03%	12.28
zoom.us	2732.40	2.63%	3.25
heygen.com	11.26	2.17%	5.93
writesonic.com	27.74	-1.57%	4.38
runwayml.com	25.16	14.50%	0.85
jasper.ai	15.37	-2.84%	4.69
figma.com	355.25	1.58%	9.05
canva.com	1514.71	-0.96%	6.66
tipranks.com	19.75	10.35%	2.41
datarobot.com	1.30	-1.25%	1.21
bing.com	4136.87	-0.73%	8.30

资料来源：华尔街见闻，华西证券研究所

1.3. Open AI 受限于 GPU，算力仍为稀缺资源

根据极市平台消息，OpenAI 的 CEO-Sam Altman 在采访中透露，目前 OpenAI 面临非常严峻的 GPU 限制，这也延迟了他们一些短期计划的实施。近来，客户投诉最多的是关于 API 的可靠性和速度问题，据 Sam 解释，大部分问题是由于 GPU 短缺造成的。

在处理文本方面，更长的 32k 上下文还不能面向更多人推出。现在 OpenAI 还没有完全克服注意力机制的 $O(n^2)$ 扩展问题，虽然 OpenAI 似乎很快就能实现 100k-1M token 上下文窗口（在今年内）文本处理，但更大的文本处理窗口需要进一步的研究突破。除此之外，目前**微调 API 也受到 GPU 供应的限制**。OpenAI 尚未使用像 Adapters 或 LoRa 这样的高效微调方法，因此微调运行和管理起来计算都非常密集。Sam 透漏更好的微调技术将在未来推出，他们甚至可能提供一个社区专门研究模型。**专用容量提供也受到 GPU 供应的限制**。OpenAI 还提供专用容量，为客户提供一个私有的模型副本。要使用此服务，客户必须愿意预先承诺支出 10 万美元（约合 68.7 万人民币）。

图表 12 当前 OpenAI 提供 API 的分类与价格

语言模型	多种型号，每种型号都有不同的功能和价位。价格为每 1,000 个代币。您可以将标记视为单词片段，其中 1,000 个标记大约为 750 个单词。本款为 35 个代币。									
GPT-4	<p>凭借广泛的常识和领域专业知识，GPT-4 可以遵循自然语言中的复杂指令并准确地解决难题。</p> <p>了解更多</p> <table border="1"> <thead> <tr> <th>模型</th> <th>迅速的</th> <th>完成</th> </tr> </thead> <tbody> <tr> <td>8K上下文</td> <td>\$0.03 / 1K 代币</td> <td>\$0.06 / 1K 代币</td> </tr> <tr> <td>32K上下文</td> <td>\$0.06 / 1K 代币</td> <td>\$0.12 / 1K 代币</td> </tr> </tbody> </table>	模型	迅速的	完成	8K上下文	\$0.03 / 1K 代币	\$0.06 / 1K 代币	32K上下文	\$0.06 / 1K 代币	\$0.12 / 1K 代币
模型	迅速的	完成								
8K上下文	\$0.03 / 1K 代币	\$0.06 / 1K 代币								
32K上下文	\$0.06 / 1K 代币	\$0.12 / 1K 代币								
聊天	<p>ChatGPT 模型针对对话进行了优化，gpt-3.5-turbo 的性能与 Instruct Davinci 相当。</p> <p>了解有关 ChatGPT 的更多信息</p> <table border="1"> <thead> <tr> <th>模型</th> <th>用法</th> </tr> </thead> <tbody> <tr> <td>gpt-3.5-turbo</td> <td>\$0.002 / 1K 代币</td> </tr> </tbody> </table>	模型	用法	gpt-3.5-turbo	\$0.002 / 1K 代币					
模型	用法									
gpt-3.5-turbo	\$0.002 / 1K 代币									

资料来源：PConline 太平洋科技，华西证券研究所

根据我们的测算，ChatGPT 产品运营需英伟达 A100 GPU 约 71296 片，预计投入算力成本达 17.73 亿美元。据 SimilarWeb 数据，2023 年 5 月（至 5 月 24 日）ChatGPT 官网（chat.openai.com）总访问量为 14.08 亿次。据环球零碳研究中心数据，每次用户与 ChatGPT 互动，ChatGPT 的每个响应词在 A100 GPU 上需要 350 毫秒。除了支持 ChatGPT 产品本身的运营，还有上述微调 API 及专用容量等的算力需求。因此我们认为目前算力仍然是制约 AI 应用发展的重要一环，整体行业需求还有较大释放空间，这也是近期访问量趋缓的原因之一。

欢迎访问：电子书学习和下载网站 (<https://www.shgis.com>)

文档名称：计算机行业深度报告-ChatGPT的流量去哪了?.pdf

请登录 <https://shgis.com/post/1807.html> 下载完整文档。

手机端请扫码查看：

