



信息技术产业行业研究

买入（维持评级）

行业月报

证券研究报告

国金证券研究所

分析师：孟灿（执业 S1130522050001）

mengcan@gjzq.com.cn

分析师：樊志远（执业 S1130518070003）

fanzhiyuan@gjzq.com.cn

分析师：张真桢（执业 S1130524060002）

zhangzhenzhen@gjzq.com.cn

AI 上游持续景气，关注原生多模态背景下的商业化机会

投资逻辑

- **计算机行业观点：**近期，AI 核心人才加盟国内大厂新闻持续发酵，叠加国产 AI 访问量环比增速快于全球 AI 产品、上市公司 AI 商业化变现进入 1 到 10 规模化阶段，或进一步增强“关键人才→爆款产品→商业化落地”的 AI 应用发展路径。8 月，全球/国内 Top20 AI 产品访问量 MoM+3.5%/+11.9%/+1.3%；8 月，国内大模型中标金额 YoY+234%；2025 半年度，部分计算机上市公司 AI 收入占比已提升至 10-30%。1、三季度进入国内外头部大模型厂商密集更新迭代周期，市场开始担忧大模型直接延伸到应用端，对于原有软件厂商的影响，我们认为应该一分为二的来看：1）对于产品用户相对不够多，用户体验和粘性不够深，切入用户生活或者业务不够深，比较标准化的功能，相对更容易被大模型厂商延伸到该环节去形成可能得替换，本身我们也看到国内外 AI 大模型厂商有一部分从大模型竞争中逐步退出，走向应用端，有一定的模型基础能力优势；2）对于产品用户基数大，市场占有率高，用户粘性强，与用户日常生活或工作业务流结合度高的产品，有数据积累、场景理解、业务认知、商务关系等多层积累的壁垒，从而相对更难被大模型厂商轻易替代，或者退一步讲，即便大模型厂商看好这块，自我投入的成效或者可能性不太大，更多可能是通过入股或者并购的方式切入，预计大概率也会支付相对市价一定幅度的溢价，这个在过去互联网+时代、第一轮 AI+时代都有案例。此外，软硬结合的部分，由于涉及到供应链、产品设计、开发，渠道建设和管理等环节，相对比线上纯软更复杂，因此大模型厂商下场延伸的概率更低。2、从目前 ai 应用落地的进展看，一方面我们看好国内在 AI+软硬结合落地的优势，软件端我们看好 AI 在 C 端和出海方向落地，美图和快手的 ARR 也初步验证了这一点，此外我们还看好 AI 在 B 端与企业服务软件、制造业信息化等环节的结合，今年半年度预计陆续披露相关进展，下半年预计环比有积极增长，明后年预期收入占比进一步提高。推荐海康威视、大华股份、禾赛、迈富时、九方智投等。
- **电子行业观点：**博通 AI 业绩超预期，ASIC 增长强劲，下游云厂商加大 AI 基础设施投资。Meta 计划到 2028 年之前，在美国建设数据中心和其他基础设施方面的投入将“至少达到 6000 亿美元”，Meta 在 2025 年资本开支预计为 660-720 亿美元，意味着未来三年将大幅增加资本开支。OpenAI 首席执行官 Sam Altman 近期表示，鉴于其最新模型 GPT-5 的需求不断增长，该公司正在优先考虑计算能力问题，并计划“在未来五个月内”将其计算集群的规模扩大一倍。我们认为下游推理需求激增，带动 ASIC 需求强劲增长，谷歌、亚马逊、Meta 等公司 ASIC 芯片快速发展，我们预测 2026 年三家公司 ASIC 芯片的数量将超过 700 万颗，OpenAI 及 xAI 等厂商也在大力推进 ASIC 芯片，未来几年也有望大幅放量。
- **通信行业观点：**1）英伟达发布 Rubin CPX，推理效率翻倍推动硬件升级需求。Rubin CPX 是首款面向百万 token 推理的 CUDA GPU，单卡算力达 30 PFLOPS。Vera Rubin NVL144 CPX 平台集成 144 颗 GPU 与 100 TB 高速内存，单机架 AI 性能达 8 EFLOPS，约为 GB300 NVL72 的 7.5 倍。推理效率翻倍推动硬件升级需求，我们看好 1.6T 光模块需求持续提升，建议关注光模块供应商。2）甲骨文 Q1 RPO 达 4550 亿美元，同比+359%，环比增加 3170 亿美元，AI 云合同订单大幅增长。已签约客户包括 OpenAI、xAI、Meta 等，推动 OCI 全年收入指引上调至+77%，未来四年有望达 1440 亿美元。建议关注甲骨文供应链投资机会。3）阿里、百度已在自研芯片上训练 AI 大模型，阿里芯片性能接近英伟达 H20，百度则用昆仑 P800 训练新版文心模型；两家公司仍部分依赖英伟达芯片开发最尖端模型。阿里同步发布 Qwen3-Next 架构并开源 80B 模型，并拟发行 32 亿美元零息可转债加强云基础设施建设。国产算力链建设有望加速。世纪互联宣布获领先互联网公司 40MW 批发订单，部署于固安 IDC 园区，分阶段交付以支持核心业务，建议关注 IDC 投资机会。

风险提示

- 国内外需求环境波动，技术落地不及预期，竞争加剧影响利润率等。



内容目录

1. AI 应用月度数据洞察	3
1.1 谁在吸引 AI 产品用户的注意力?	3
1.2 什么样的 AI 产品变现能力最强?	9
1.3 国内 AI 大模型招投标需求如何?	11
1.4 上市公司 AI 商业化进度如何?	12
2. 全球 AI 本月重大新闻.....	14
2.1 海外大模型及 AI 应用进展.....	14
2.2 国内 AI 行业动态及重点公司新闻	15
3. 投资建议.....	15
4. 风险提示.....	18

图表目录

图表 1: 全球访问量 Top20 AI 产品排名情况	3
图表 2: 2024 - 2025 年全球访问量 TOP20 月度排名表	3
图表 3: 全球访问量环比增速 Top20 AI 产品情况	4
图表 4: 国内访问量 Top20 AI 产品排名情况	5
图表 5: 2024 - 2025 年国内访问量 TOP20 月度排名表	5
图表 6: 国内访问量环比增速 Top20 AI 产品情况	6
图表 7: 出海访问量 Top20 AI 产品排名情况	7
图表 8: 2024 - 2025 年出海访问量 TOP20 月度排名表	7
图表 9: 出海访问量环比增速 Top20 AI 产品情况	8
图表 10: 8 月海外/国内/出海赛道访问量环比增速 Top 产品表.....	8
图表 11: 全球 AI 产品年经常性收入 Top20 排名情况.....	10
图表 12: 国内 AI 产品年经常性收入 Top20 排名情况.....	10
图表 13: 出海 AI 产品年经常性收入 Top20 排名情况.....	11
图表 14: 2025 年 1-8 月国内大模型中标项目数量情况	11
图表 15: 2025 年 1-8 月国内大模型中标项目披露金额情况	12
图表 16: AI 应用公司 25 上半年 AI 产品相关收入.....	12
图表 17: 服务器指数 (8841058.WI)走势	17
图表 18: 光模块(CPO)指数 (8841258.WI)走势	17
图表 19: IDC 指数 (866052.WI)走势	18

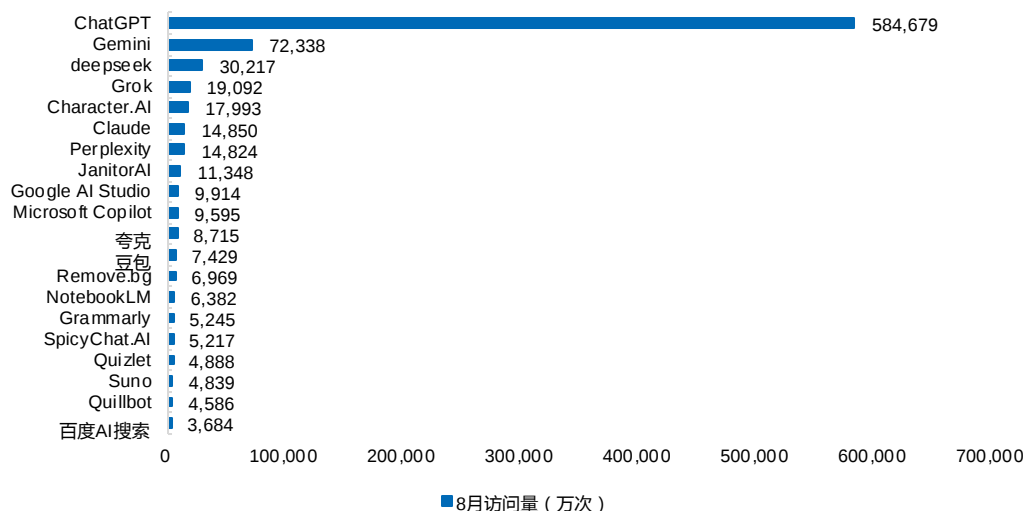


1. AI 应用月度数据洞察

1.1 谁在吸引 AI 产品用户的注意力？

根据 SimilarWeb 统计及非凡产研整理，2025 年 8 月全球访问量 Top20 的 AI 应用以头部互联网公司和 AI 大模型龙头为主，月访问量过亿的应用共 8 款。访问量排名前五的应用分别为来自 OpenAI 的聊天机器人 ChatGPT、来自谷歌的聊天机器人 Gemini、国产大模型产品 deepseek、xAI 的个人助理 Grok、情感陪伴类工具 Character.AI。

图表1：全球访问量 Top20 AI 产品排名情况



来源：SimilarWeb，非凡产研，国金证券研究所

据全球 AI 应用访问量 Top20 月度排名，ChatGPT 在整个观测期内始终稳居榜首。排名第二的产品在不同阶段呈现激烈竞争，Bing 曾在 2024 年 6 月至 2025 年 1 月稳居第二，但随后被 deepseek、Gemini 等应用取代，其中 deepseek 自 2025 年 2 月起稳定跻身前三，也是观测期内唯一长居全球前三的中国 AI。Canva 也曾长期位列前三，但在 2025 年 2 月后逐步被新兴应用超越。值得关注的是，“豆包”自 2025 年 3 月起进入榜单并保持稳定，腾讯元宝、百度 AI 搜索等国内应用也逐渐崭露头角，显示出本土 AI 产品在全球范围内的渗透力提升。整体来看，全球 AI 应用市场头部格局相对稳定，但中下游竞争较为激烈。

图表2：2024 - 2025 年全球访问量 TOP20 月度排名表

	2024/6	2024/9	2024/12	2025/1	2025/2	2025/3	2025/4	2025/5	2025/6	2025/7	2025/8
Top 1	ChatGPT	ChatGPT	ChatGPT	ChatGPT	ChatGPT	ChatGPT	ChatGPT	ChatGPT	ChatGPT	ChatGPT	ChatGPT
Top 2	Bing	Bing	Bing	Bing	deepseek	deepseek	deepseek	Gemini	Gemini	Gemini	Gemini
Top 3	Canva	Canva	Canva	Canva	Gemini	Gemini	Gemini	deepseek	deepseek	deepseek	deepseek
Top 4	Gemini	Gemini	Brave	Gemini	Character.AI	Grok	Grok	Character.AI	Character.AI	Grok	Grok
Top 5	Character.AI	DeepL	Gemini	deepseek	Perplexity	Character.AI	Character.AI	Grok	Grok	Character.AI	Character.AI
Top 6	DeepL	Character.AI	Character.AI	DeepL	Q-Chat	Perplexity	Perplexity	Perplexity	Perplexity	Perplexity	Claude
Top 7	Notion	Notion	DeepL	Character.AI	JanitorAI	JanitorAI	JanitorAI	JanitorAI	Claude	Claude	Perplexity
Top 8	百度网盘	Yandex	Yandex	Intuit	Claude	Claude	Claude	Claude	JanitorAI	JanitorAI	JanitorAI
Top 9	Intuit	Q-Chat	Shop	Yandex	Quillbot	Q-Chat	Q-Chat	Microsoft Copilot	Google AI Studio	Microsoft Copilot	Google AI Studio
Top 10	Shop	Intuit	Notion	Notion	Grammarly	Quillbot	Google AI Studio	Quizlet	Microsoft Copilot	Google AI Studio	Microsoft Copilot
Top 11	Salesforce	Freepik	Freepik	Shop	Grok	Google AI Studio	Quillbot	Google AI Studio	豆包	豆包	夸克

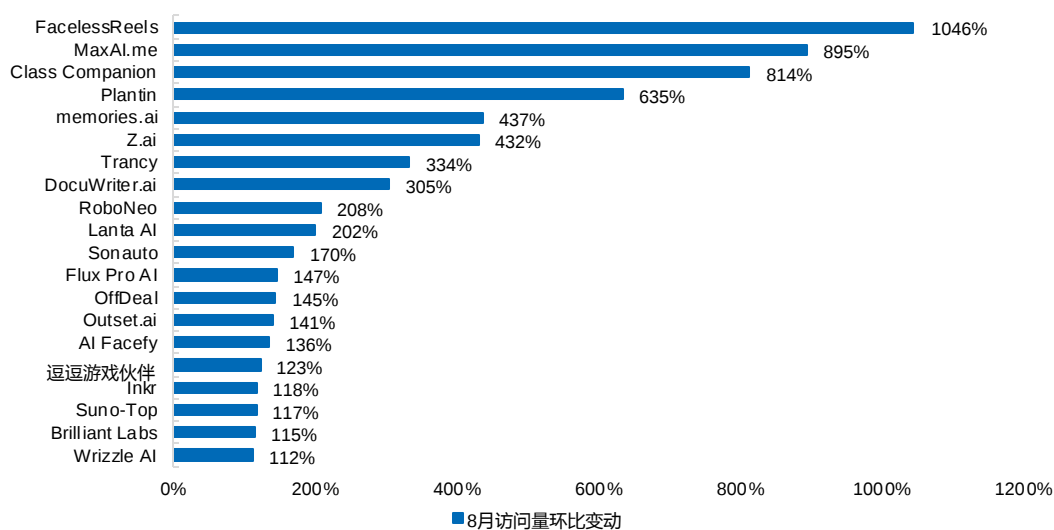


Top 12	Q-Chat	Salesforce	Intuit	Freepik	Quizizz	Grammarly	Grammarly	NotebookLM	NotebookLM	NotebookLM	豆包
Top 13	Freepik	Shop	Perplexity	Salesforce	Suno	豆包	豆包	豆包	Quizlet	SpicyChat.AI	Remove.bg
Top 14	Perplexity	Quillbot	Salesforce	Perplexity	豆包	Quizizz	Suno	Quillbot	SpicyChat.AI	Suno	NotebookLM
Top 15	Claude	Perplexity	JanitorAI	JanitorAI	Kimi	Suno	NotebookLM	Suno	Suno	Grammarly	Grammarly
Top 16	Shutterstock	Remove.bg	Q-Chat	Q-Chat	Google AI Studio	Qwen	Qwen	Grammarly	Grammarly	Quizlet	SpicyChat.AI
Top 17	JanitorAI	Claude	Claude	Claude	NotebookLM	Kimi	Quizizz	SpicyChat.AI	Quillbot	Quillbot	Quizlet
Top 18	Suno	Grammarly	Remove.bg	Duolingo	SpicyChat.AI	NotebookLM	SpicyChat.AI	Quizizz	lovable	Kimi	Suno
Top 19	Grammarly	Shutterstock	Shutterstock	Remove.bg	Hugging Face	SpicyChat.AI	Kimi	Qwen	Qwen	Miro	Quillbot
Top 20	Remove.bg	Suno	Quillbot	Shutterstock	Miro	腾讯元宝	lovable	lovable	Miro	PolyBuzz	百度 AI 搜索

来源：SimilarWeb，非凡产研，国金证券研究所（注：红色表示排名上升、蓝色表示排名下降、黑色表示排名不变。）

2025 年 8 月全球访问量增速最快的应用分别为 FacelessReels、MaxAI.me 以及 Class Companion，访问量分别同比增长 1046%、895%、814%。其中，FacelessReels 是一款用 AI 自动生成并自动发布“无脸短视频”的订阅式工具，在选定领域与视频格式、风格与音乐后，系统会按频率自动生成并发布短视频，覆盖 TikTok、Instagram、YouTube，目标是“省时、省钱、可规模化”；MaxAI.me 通过浏览器扩展，让用户可以在任何网页的侧边栏直接使用 AI 进行对话、写作、总结和翻译，无需切换标签页；Class Companion 面向中小学及高校教师，提供作业的即时反馈、AI 一对一辅导和学情洞察。它能显著减轻教师批改作业负担，同时为学生提供个性化的学习支持。

图表3：全球访问量环比增速 Top20 AI 产品情况

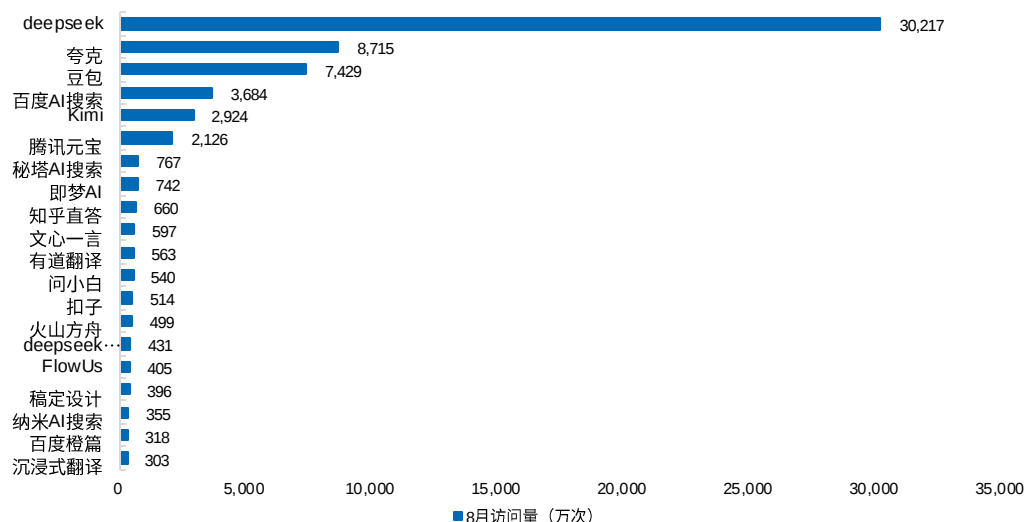


来源：SimilarWeb，非凡产研，国金证券研究所

根据 SimilarWeb 统计及非凡产研整理，2025 年 8 月国内访问量 Top20 的 AI 应用中有 13 款的访问量超过 500 万次。访问量排名前三的应用分别为 deepseek、夸克、豆包，访问量分别为 3.0 亿次、8,715 万次、7,429 万次。



图表4：国内访问量Top20 AI产品排名情况



来源：SimilarWeb，非凡产研，国金证券研究所

据国内 AI 应用访问量 Top20 月度排名，国内市场已形成“头部双强”格局：deepseek 自 2025 年 2 月起长期稳居第一，豆包稳定位列前二，展现出强劲的用户粘性与市场影响力；Kimi、百度 AI 搜索及腾讯元宝等腰部应用紧随其后，其中百度 AI 搜索在大厂生态加持下排名稳步提升；秘塔 AI 搜索保持稳定表现，而即梦 AI、知乎直答、阿里通义等则波动明显，体现出中腰部竞争激烈、替代性强。此外，教育、办公、翻译等实用型工具仍维持稳定需求，显示 AI 正加速嵌入多场景工作流。边际看，8 月夸克升至 Top2。

图表5：2024 - 2025 年国内访问量TOP20 月度排名表

	2024/6	2024/9	2025/2	2025/3	2025/4	2025/5	2025/6	2025/7	2025/8
Top 1	CapCut	Kimi	deepseek	deepseek	deepseek	deepseek	deepseek	deepseek	deepseek
Top 2	Kimi.ai	文心一言	豆包	豆包	豆包	豆包	豆包	豆包	夸克
Top 3	CrushOn.AI	豆包	Kimi	Kimi	Kimi	Kimi	百度 AI 搜索	Kimi	豆包
Top 4	文心一言	Vidu	deepseek 开放平台	腾讯元宝	腾讯元宝	百度 AI 搜索	Kimi	百度 AI 搜索	百度 AI 搜索
Top 5	yodayo	阿里通义	百度 AI 搜索	百度 AI 搜索	百度 AI 搜索	腾讯元宝	腾讯元宝	腾讯元宝	Kimi
Top 6	Fotor	秘塔 AI 搜索	硅基流动	秘塔 AI 搜索	文心一言	秘塔 AI 搜索	秘塔 AI 搜索	秘塔 AI 搜索	腾讯元宝
Top 7	Cutout Pro	C 知道	秘塔 AI 搜索	文心一言	秘塔 AI 搜索	文心一言	问小白	即梦 AI	秘塔 AI 搜索
Top 8	秘塔 AI 搜索	海螺 AI	腾讯元宝	阿里通义	即梦 AI	知乎直答	即梦 AI	知乎直答	即梦 AI
Top 9	SeaArt AI	知乎直答	文心一言	纳米 AI 搜索	纳米 AI 搜索	问小白	文心一言	文心一言	知乎直答
Top 10	PixAI.Art	智谱清言	纳米 AI 搜索	问小白	问小白	即梦 AI	知乎直答	问小白	文心一言
Top 11	PicWish	360AI 搜索	阿里通义	deepseek 开放平台	知乎直答	有道翻译	纳米 AI 搜索	有道翻译	有道翻译
Top 12	阿里通义	360 苏打办公	火山方舟	知乎直答	阿里通义	纳米 AI 搜索	沉浸式翻译	扣子	问小白
Top 13	Hix AI	可画	知乎直答	火山方舟	deepseek 开放平台	火山方舟	有道翻译	火山方舟	扣子
Top 14	MaxAI.me	稿定设计	有道翻译	有道翻译	有道翻译	扣子	火山方舟	纳米 AI 搜索	火山方舟
Top 15	有道翻译	百度 Chat	即梦 AI	即梦 AI	扣子	百度橙篇	扣子	FlowUs	deepseek 开放平台
Top 16	Gauth	沉浸式翻译	问小白	硅基流动	火山方舟	FlowUs	FlowUs	deepseek 开放平台	FlowUs

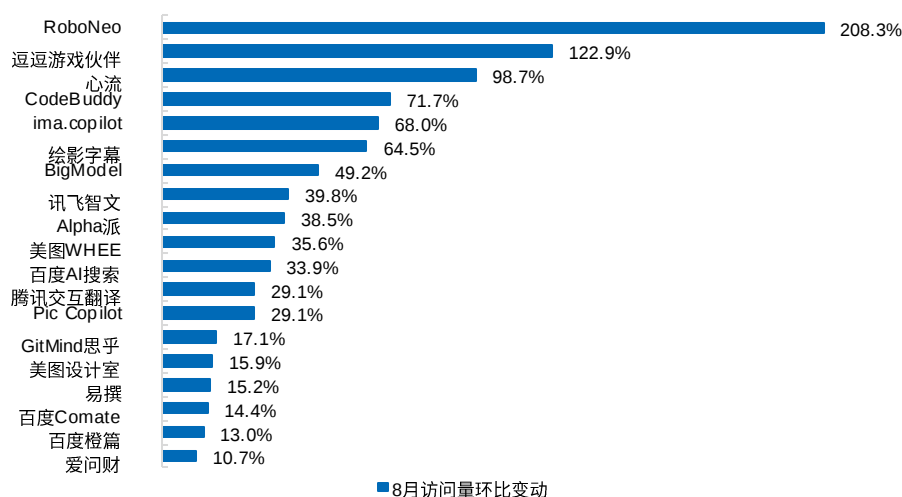


Top 17	Openart	deepseek	FlowUs	扣子	FlowUs	deepseek 开放平 台	deepseek 开放平 台	稿定设计	稿定设计
Top 18	Coze	火山方舟	LiblibAI	LiblibAI	沉浸式翻译	沉浸式翻译	百度橙篇	LiblibAI	纳米 AI 搜索
Top 19	豆包	熊猫办公	沉浸式翻译	FlowUs	C 知道	C 知道	LiblibAI	沉浸式翻译	百度橙篇
Top 20	OpusClip	讯飞星火	扣子	C 知道	LiblibAI	LiblibAI	稿定设计	扣子	沉浸式翻译

来源：SimilarWeb，非凡产研，国金证券研究所（注：红色表示排名上升、蓝色表示排名下降、黑色表示排名不变。）

2025 年 8 月国内访问量 Top20 的 AI 应用中增速前三的应用分别为 RoboNeo、逗逗游戏伙伴、心流，分别同比增长 208.3%、122.9%、98.7%。其中，RoboNeo 是一款美图旗下的影像与设计 AI 助手，能通过简单的指令，帮助用户快速完成图片编辑（如换衣服、发型、修背景）、设计（如海报、物料、LOGO）以及视频制作（如生成短片、修复画质）等任务，实现高效创作；逗逗游戏伙伴是一个带有“陪玩 + 陪伴”功能的 AI 游戏社交应用，内置多个 AI 角色，能在你玩游戏时提供战术建议、角色互动、实时语音开黑等；心流是一款专为科研人员、高校教师、学生及职场人士等专业知识管理需求用户打造的智能工具。它集成了智能搜索、知识问答、智能阅读、辅助创作等多项功能，助力用户高效获取、管理与应用知识。

图表6：国内访问量环比增速 Top20 AI 产品情况

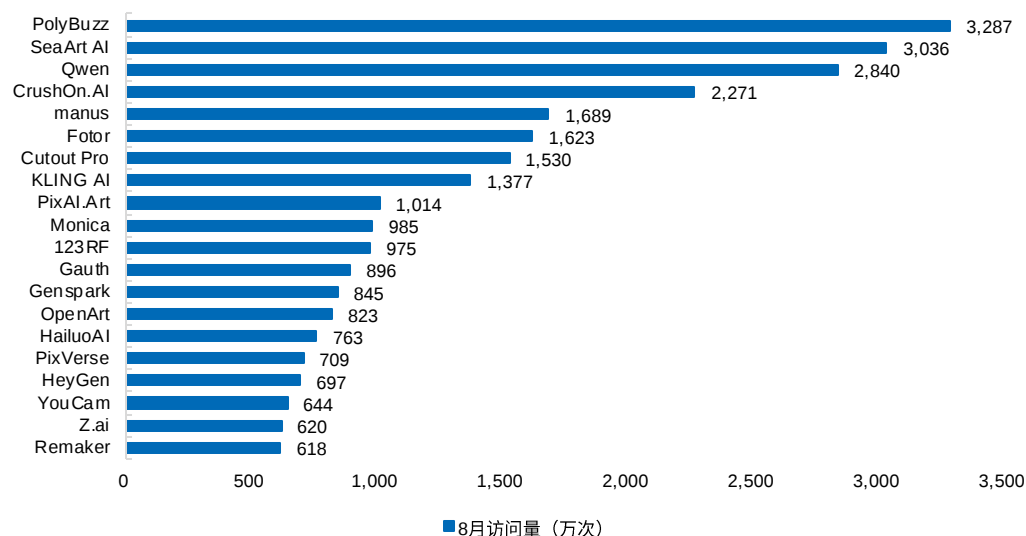


来源：SimilarWeb，非凡产研，国金证券研究所

出海类 AI 应用中，2025 年 8 月访问量排名前三的应用分别为 PolyBuzz、SeaArt AI、Qwen，访问量分别为 3,287 万次、3,036 万次、2,840 万次。其中，PolyBuzz 可以模拟动漫、影视、游戏或原创角色的语言与性格，支持文本、语音、图像多模态交互，让用户沉浸式体验角色扮演与故事共创；SeaArt AI 利用多风格生成模型，快速生成高质量动漫、3D、油画等图片，具备文生图、图生图、智能补图，适用于创作和设计多场景；Qwen 系列大模型由阿里巴巴研发，集成“规划—执行—总结”三阶段任务流程，覆盖 119 种语言及方言，具备自动检索资料、结构化整理信息与生成多格式报告的能力，支持多轮搜索与多任务并行。



图表7：出海访问量 Top20 AI 产品排名情况



来源：SimilarWeb，非凡产研，国金证券研究所

据出海 AI 应用访问量 Top20 月度排名，PolyBuzz 自 2025 年 4 月起快速攀升并在 7-8 月稳居首位；SeaArt AI、Qwen 及 CrushOn.AI 亦长期位列前五，。中游应用如 manus、Fotor、KLING AI、Cutout Pro 等保持相对稳定，满足特定场景或区域需求；Top20 下游应用则更替频繁，Monica、Gauth、123RF、HeyGen 等间歇上榜。整体来看，中国 AI 应用出海已形成相对稳定的头部梯队与多元竞争生态。

图表8：2024 - 2025 年出海访问量 TOP20 月度排名表

	2024/9	2025/4	2025/5	2025/7	2025/8
Top 1	CapCut	Qwen	Qwen	PolyBuzz	PolyBuzz
Top 2	KLING AI	SeaArt AI	CrushOn.AI	SeaArt AI	SeaArt AI
Top 3	Fotor	CrushOn.AI	PolyBuzz	Qwen	Qwen
Top 4	Gauth	PolyBuzz	SeaArt AI	CrushOn.AI	CrushOn.AI
Top 5	Cutout Pro	KLING AI	KLING AI	manus	manus
Top 6	CrushOn.AI	manus	manus	Fotor	Fotor
Top 7	SeaArt AI	Fotor	Hailuo AI	KLING AI	Cutout Pro
Top 8	123RF	Hailuo AI	Fotor	Cutout Pro	KLING AI
Top 9	MaxAI.me	Gauth	Cutout Pro	Hailuo AI	PixAI.Art
Top 10	Monica AI	Clipto.AI	Monica	Remaker	Monica
Top 11	PicWish	Cutout Pro	Gauth	Monica	123RF
Top 12	Udio	Monica AI	Clipto.AI	PixAI.Art	Gauth
Top 13	yodayo	Zeemo	Zeemo	123RF	Genspark
Top 14	Talkie	PixVerse	123RE	Genspark	OpenArt
Top 15	Otter AI	123RF	PixVerse	Openart	Hailuo AI
Top 16	Media.io	insMind	PiXAI.Art	Gauth	PixVerse
Top 17	Sider	PixAI.Art	Genspark	HeyGen	HeyGen
Top 18	YouCam	Openart	HeyGen	PixVerse	YouCam
Top 19	PixAI.Art	Genspark	Sider	Clipto.AI	Z.ai
Top 20	Openart	Otter AI	TensorArt	Dreamina	Remaker

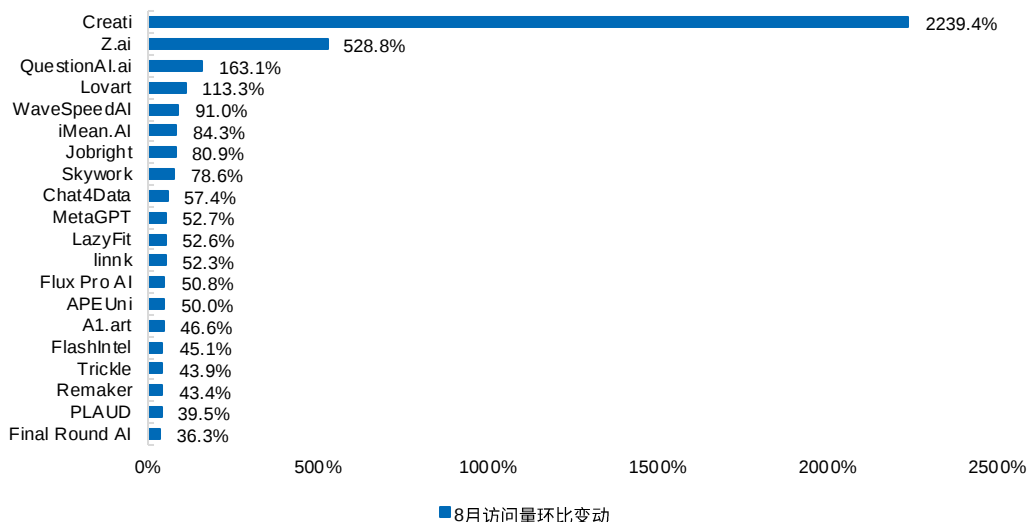
来源：SimilarWeb，非凡产研，国金证券研究所（注：红色表示排名上升、蓝色表示排名下降、黑色表示排名不变。）

出海榜单中，2025 年 8 月国内访问量 Top20 的 AI 应用中增速最快的应用分别 Creati、



Z.ai、QuestionAI.ai 分别同比增长 2239.4%、528.8%、163.1%。Creati 通过 AI 模型将爆款视频转化为病毒视频模板，在商家上传产品图片和 URL 后，基于产品特性和定位对故事线、人物主题、叙事背景等进行调整，最终生成个性化视频；Z.ai 是国内大模型厂商智谱面向全球推出的模型体验平台，整合了 32B 基座、推理、沉思三类 GLM 模型，后续将作为智谱最新模型的交互体验入口；Question.AI 是作业帮面向海外市场推出的 AI 教育产品，基于自研大模型和 AI 算法，为学生提供覆盖全学段、全学科的解决方案。

图表9：出海访问量环比增速 Top20 AI 产品情况



来源：SimilarWeb，非凡产研，国金证券研究所

8 月海外/国内/出海赛道访问量环比增速 Top 产品中，多模态应用产品占 1/3，泛情感陪伴类占 2/9，Qwen 在模型更新之后关注度阶段性快速提升，面向垂类场景的应用开发仍有爆款机会。

图表10：8 月海外/国内/出海赛道访问量环比增速 Top 产品表

产品名称	发布时间	产品定位	产品亮点	潜在竞品
Faceless Reels	/	视频生成	订阅式 AI 无脸短视频自动生成与发布工具，支持用户选定领域、视频格式、风格及音乐后，按频率自动生成内容并同步至 TikTok、Instagram、YouTube 等平台，核心优势是 5 分钟快速创建视频，批量生产提升效率，采用“三步走”简易流程（选题材 — 定风格 / 音乐 — 连账号自动上传）	Pictory.ai
MaxAI.me	2023 年 12 月	个人助理	浏览器扩展 AI 工具，可在任意网页侧边栏直接实现对话、写作、总结、翻译等功能，无需切换标签页，支持 GPT、Gemini、Claude 等多种 AI 模型，安装便捷，能读取和处理网页数据，帮助用户在浏览过程中高效完成各类文本相关任务。	Monica
Class Companion	2023 年发布， 2025 年 4 月被 Panorama Education 收购	教育	面向 K12 及高校教师的 AI 教学辅助平台，支持教师创建作业（可自定义格式、内容）并邀请学生完成，通过 AI 实现作业即时反馈、一对一辅导及学情洞察，减轻教师批改负担，为学生提供个性化学习支持，学生需通过教师提供的 join code 注册，流程简洁（添加作业 — 邀请学生 — 查看提交与洞察）。	School AI
RoboNeo	2025 年 6 月	图片生成	影像与设计类 AI 助手，能通过简单指令快速完成图片编辑（换衣、修背景等）、设计（海报、LOGO 等）及视频制作（生成短片、修复画质等），支持制作 emoji 动态小人、表情包等，创作耗时仅 0.5-1 分钟，聚焦高效创意内容产出。	Clipdrop



逗逗游戏伙伴	2023 年 10 月发布测试版 2025 年 8 月发布 1.0 正式版	游戏生成	专注游戏领域的 AI 社交应用，内置多个二次元 AI 角色，具备“陪玩 + 陪伴”功能，借助 VLM 技术和实时语音通话，能“看懂”游戏画面、记住共同经历，提供战术建议、角色互动、实时语音开黑等服务，陪伴用户游戏并建立共同回忆。	Character.AI
心流 AI	2024 年 7 月	智慧搜索	智能知识管理工具，集成智能搜索、知识问答、智能阅读、辅助创作等功能，可检索论文并提供 AI 精读、论文图谱、引文直达等服务，支持自定义知识库问答、对话保存为笔记及生成播客，满足专业场景下高效获取与管理知识的需求。	Scite.ai
PolyBuzz	/	情感陪伴	AI 角色扮演在线聊天工具，支持模拟动漫、影视、游戏及原创角色的语言与性格，提供文本、语音、图像多模态交互，用户可与虚拟角色沉浸式对话、共创故事，具备角色自定义创建、图片生成、艺术画廊收藏等功能，拥有海量角色库，还原角色风格与声线。	Character.AI
SeaArt AI	2023 年 6 月	图片生成	AI 艺术创作应用，具备文生图、图生图、智能补图、视频生成等功能，支持动漫、3D、油画等多风格创作，拥有丰富模型库与虚拟角色库，可与虚拟角色对话互动，提供一站式创作社区，用户能上传图片 / 视频并描述需求生成内容，兼顾创作与社交属性。	Midjourney
Qwen	2023 年 8 月	聊天机器人	阿里巴巴研发的大模型系列，集成“规划 — 执行 — 总结”任务流程，覆盖 119 种语言及方言，具备自然语言处理、多模态交互（视觉、音频理解）、工具使用、角色扮演等能力，支持自动检索资料、结构化整理信息、生成多格式报告，可处理文本生成、翻译、视觉问答等任务，支持多轮搜索与多任务并行。	Chat-GPT

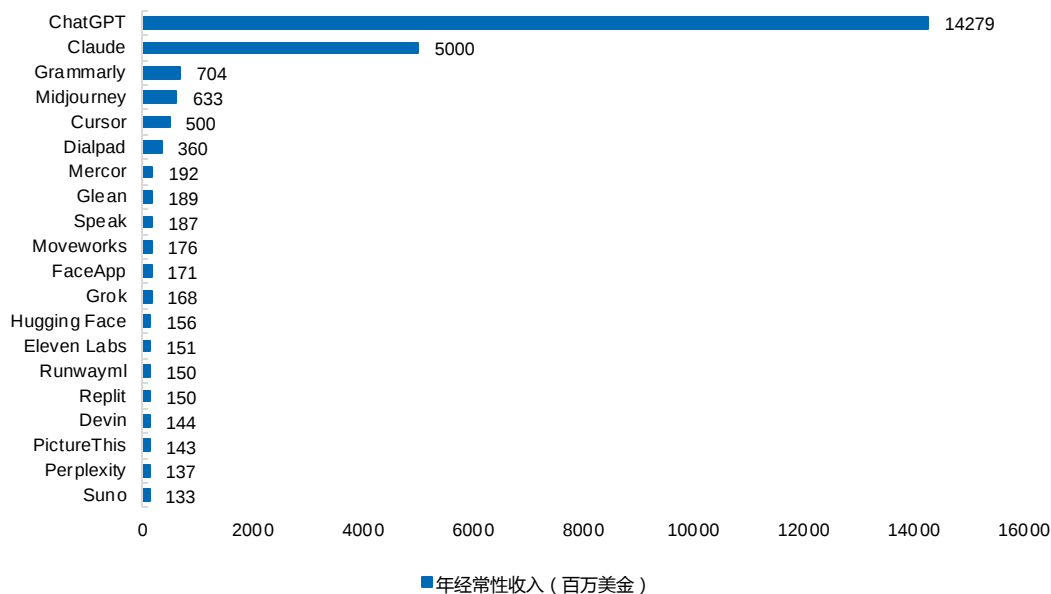
来源：SimilarWeb，非凡产研，国金证券研究所

1.2 什么样的 AI 产品变现能力最强？

根据 SimilarWeb 统计及非凡产研整理，截至 2025 年 8 月，全球 AI 产品年经常性收入（ARR）Top20 主要集中于头部互联网科技公司和 AI 大模型厂商。其中共有 6 款产品年收入超过 2 亿美元，排名前五的分别是 OpenAI 的 ChatGPT、Anthropic 的 Claude、Grammarly、Midjourney、Anysphere 的 Cursor，年经常性收入分别为 142.79 亿美元、50 亿美元、7.04 亿美元、6.33 亿美元、5 亿美元。其中，Midjourney 支持文本或语音生成、编辑、动画化多风格图像，并具风格迁移、模型切换、任务管理，实现高效个性化视觉创作；Cursor 基于聊天式 AI 编程，将自然语言转为代码生成、调试与项目管理，支持多模态上下文、Agent 执行和 MCP 接入。



图表11：全球 AI 产品年经常性收入 Top20 排名情况

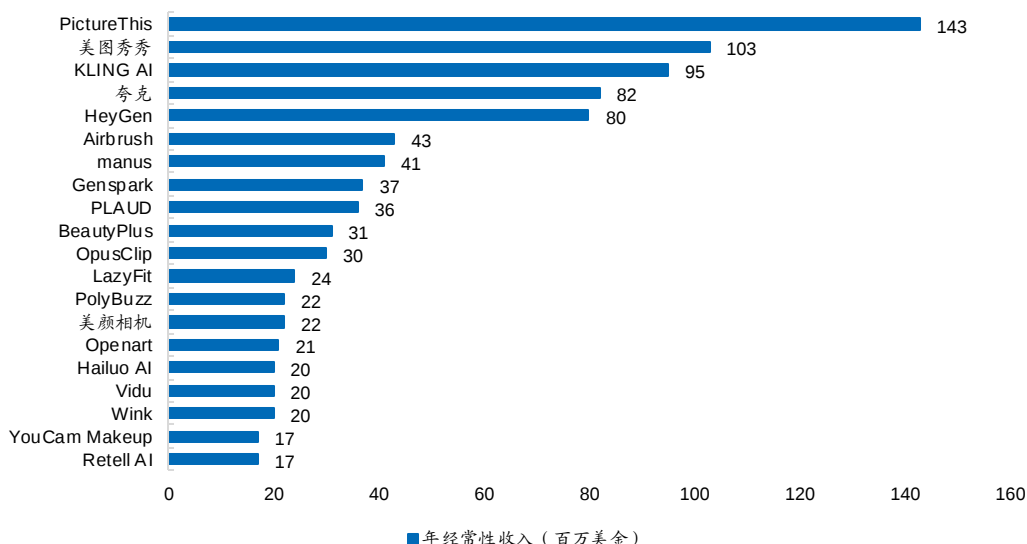


来源：SimilarWeb，非凡产研，国金证券研究所

说明：AI 产品年经常性收入 (ARR) 为产品 Web 端与 App 端合计的年度化经常性收入

根据 SimilarWeb 统计及非凡产研整理，截至 2025 年 7 月，国内 AI 产品年经常性收入 (ARR)Top20 中，有 2 款产品年收入超 1 亿美元。收入排名前三的应用分别为 PictureThis、美图秀秀、KLING AI，年经常性收入分别为 1.43 亿美元、1.03 亿美元、0.95 亿美元。其中，PictureThis 集成 AI 识别与植物数据库，支持多语种互动，快速识别植物并提供养护信息与社区交流；KLING AI 融合全新 DiT 架构，实现高精度视频图像生成、动态捕捉与多模态创作，助用户轻松打造电影级视觉内容。

图表12：国内 AI 产品年经常性收入 Top20 排名情况

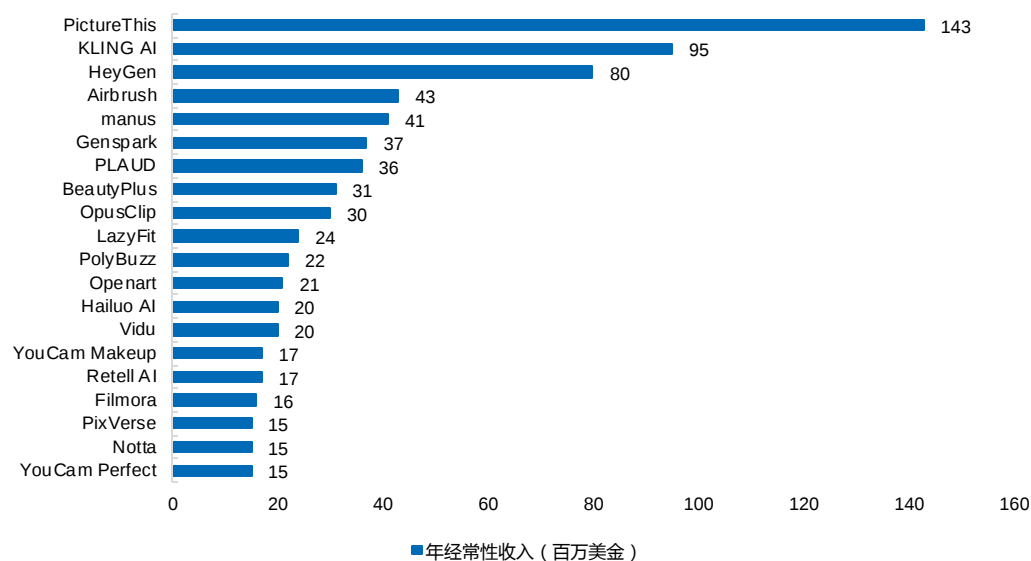


来源：SimilarWeb，非凡产研，国金证券研究所

出海类 AI 应用年经常性收入 Top20 中，排名前三的应用分别为 PictureThis、KLING AI、HeyGen，收入分别为 1.43 亿、0.95 亿、0.8 亿美元。其中，HeyGen 能理解语音的节奏、语气和情绪，自动生成拟人化表情与动作，实现从“一张照片+一句话”到“超逼真数字人视频”的端到端生成，可支持正脸、侧脸、全身多角度输入，广泛应用于动画制作、虚拟角色配音、播客可视化等场景。



图表13：出海 AI 产品年经常性收入 Top20 排名情况

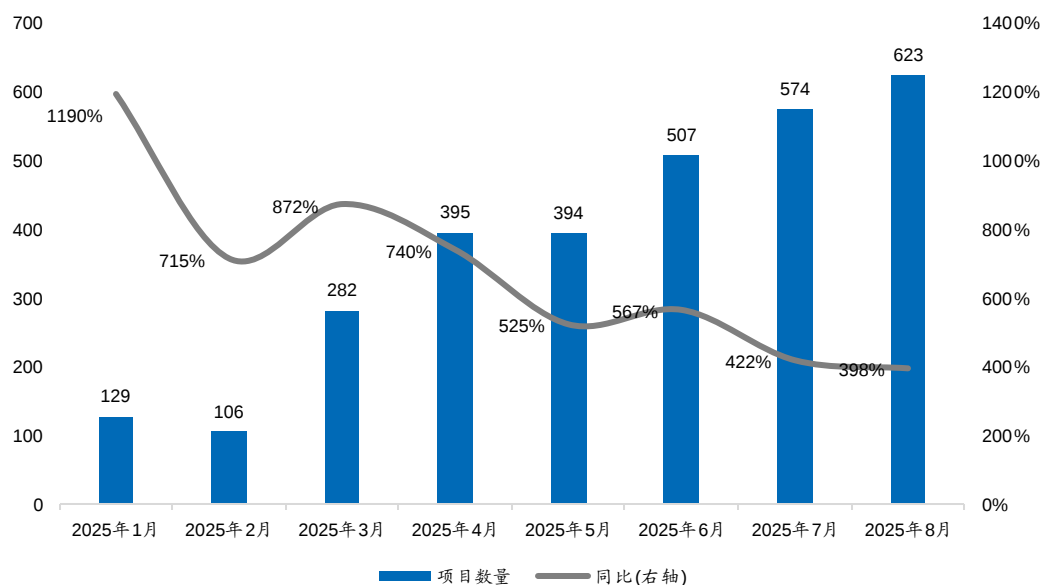


来源：SimilarWeb，非凡产研，国金证券研究所

1.3 国内 AI 大模型招投标需求如何？

2025 年 1-8 月，国内大模型中标项目数量实现逐月快速增长，1 月同比增速达到 1190%，其后同比增速虽有所下降，但始终维持在高位，证明大模型技术在国内市场的接受率大幅提速，落地速度加快。

图表14：2025 年 1-8 月国内大模型中标项目数量情况

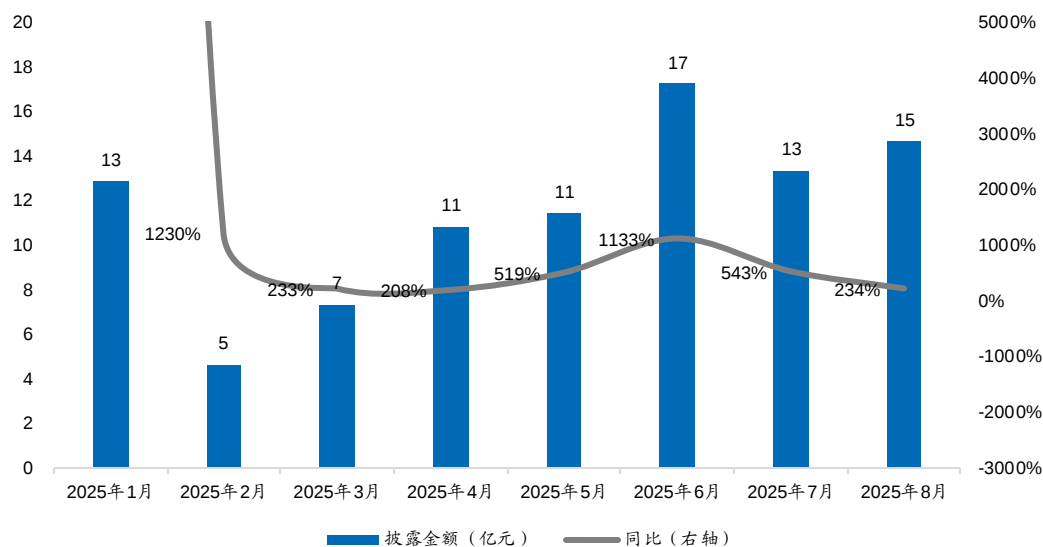


来源：智能超参数，国金证券研究所

根据中标项目披露的中标金额数据，2025 年 1-8 月中标项目披露金额与其同比增长率在 1 月达到峰值，2 月后中标项目披露金额逐月回升，8 月中标项目的披露金额为 14.7 亿元，相比 7 月的 13.4 亿元有所上升。



图表15：2025 年 1-8 月国内大模型中标项目披露金额情况



来源：智能超参数，国金证券研究所

1.4 上市公司 AI 商业化进度如何？

众多计算机上市公司在 AI 商业化领域已有积极进展。在收入贡献上，部分公司 AI 业务收入占比表现亮眼，如能科科技 AI 业务实现收入 2.1 亿元，占营收比达 28.5%；彩讯股份 AI 与算力相关业务收入为 1.77 亿元，占营收 19.2%。同时，不少公司聚焦 AI 技术在各行业的落地应用以推动商业化，如税友股份围绕税务场景，基于 AI 能力开发智能体；汉得信息将 AI 融入供应链、财务等业务领域，推动企业智能化升级。

图表16：AI 应用公司 25 上半年 AI 产品相关收入

股票简称	25H1 总营收 (亿元)	2025 年 H1 业务情况 (按订单或收入金额排序)	占收入比	AI 口径
能科科技	7.38	AI 业务实现收入 2.1 亿元，占总收入比例达 30%。公司聚焦工业领域，以“灵系列”AI 产品为核心，通过强化中台能力与场景化应用，推动 AI 技术在工业生产与管理中的规模化落地与复用。	28.5%	收入
彩讯股份	9.20	AI 与算力相关业务收入为 1.77 亿元，占总收入 19%。公司重点协助中国移动、南方电网、中国联通、南方航空、中国铁塔等大型企业客户探索 AI 平台开发、AI 技术及产品应用等智能化升级。	19.2%	收入
税友股份	9.22	实现 AI 驱动的产品及服务收入为 1.65 亿元。税友股份以财税 AIBM 生态服务平台为核心，通过 AI 驱动打造合规税优、财经管理等 Agent 应用，助力财税服务机构及中小微企业实现数智化转型与价值提升。	17.9%	收入
致远互联	3.38	签约的 AI 关联合同订单金额为 4,669.11 万元。AI-Agent 开发平台 CoMi Builder，推出了一系列的预制智能体，如公文智能体、会议智能体、合同智能体等，推出的一站式企业 AI 普及服务平台致慧泉，以及智慧表单 iForm。	13.9%	订单
新点软件	6.70	AI 相关产品实现营业收入 9000 万元，其中政务 AI 实现营业收入 2500 万元。AI 应用已从基础的客服、问答场景，拓展至建立流程自动化、优化客户服务质量和辅助决策：产品包括 AI 助手和 Aagent 平台，同时将 AI 能力融入产品中，如智能问答、智能办理、智能问数能力。	13.4%	收入
迈富时	9.28	AI 及 AI Agent 收入达 1.1 亿元，当前在手订单 1.1 亿元，年度经常性收入 (ARR) 为 1.9 亿元。AI Agentforce 智能平台 1.0 及 2.0 全面落地，覆盖营销与销售全链路场景，形成标准化可复用能力，并深度应用于消费零售、汽车、医疗、金融、文旅及制造等行业，推出 AI 外贸数字员工、销售陪练系统、智能招聘系统等多款 AI Agent 产品。	11.9%	收入



赛意信息	9.02	AI 相关业务今年上半年订单达 1.03 亿元，其中约六成来自泛 ERP 领域，探索企业管理流程作业的自动化和智能化执行；智能制造领域 AI 占比约 40%，解决方案采用垂类 AI 模型，来自投资者关系活动记录表。	11.4%	订单
用友网络	35.81	第二季度 BIP 产品实现 AI 相关合同签约金额超 3.2 亿元，部分单个合同签约金额突破千万。“用友 BIP 企业 AI”具有“统一企业数智底座、融入企业核心业务、成果可靠、安全合规”的特性，实现 AI、数据、流程的原生一体化。	8.9%	订单
万兴科技	7.6	AI 原生应用收入超过 6,000 万元，同比增长超 2 倍，占公司软件业务总体收入超过 8%。上半年公司完成了 AI 云平台架构和服务的全面升级，构建起一站式 AI 创新能力平台，提供包括内外算法集成、工作流编排、演示体验中心、运营配置工具等全套能力体系，加强对算法模型等数据资产的管理	8.0%	收入
金桥信息	2.73	公司基于 AI 技术搭建多元解纷平台，上半年该平台调解相关业务实现收入超 2,100 万元。公司与阿里云智能集团签署合作框架协议，共建“金桥 - 阿里司法行业人工智能联合创新基地”，联手打造司法行业人工智能大语言模型，深化 AI 领域技术与生态合作。	7.7%	收入
索辰科技	0.57	开物系列(物理 AI 产品线)收入为 375.47 万元，自公司发布物理 AI 产品以来，公司快速拓展下游应用并形成收入。“开物”产品线以“物理 AI”为核心，涵盖物理 AI 模拟引擎、物理 AI 训练一体化平台（智能工坊）等。目前重点聚焦物理 AI 在低空经济、具身智能、新能源电池等战略性新兴领域。物理 AI 技术与 CAE 技术融合，赋能航空航天、汽车制造、能源动力等领域复杂装备研发，提升客户设计效率与优化能力。	7.0%	收入
汉得信息	15.75	AI 应用业务实现了重大突破达到约 1.1 亿元收入。公司持续深化“得灵”B 端 AI 应用软件及服务，推动营销、供应链、财务、工业等领域的智能化升级，并成功发布物流、矿山、锂电行业大模型及代码大模型。	7.0%	收入
金蝶国际	31.92	金蝶在报告期内 AI 相关合同金额已超 1.5 亿元，通过推出苍穹 AI Agent 平台及多款 AI 智能体产品，助力企业实现管理智能化转型。	4.7%	订单
鼎捷数智	10.45	公司 AI 业务收入同比增长 125.91%。公司发布的业内首款融入 AI 技术的 PLM 产品全生命周期管理系统，通过多轮验证、迭代，已获近百家客户签约、数十家客户成功部署。此外，公司发布 AI 智能套件加速推进 AI 与 ERP 的融合创新，推出指令式操作、多模态生单、智能附件、报告助手、敏捷问数等功能。	/	收入
科大讯飞	109.11	企业 AI 解决方案收入达 4.4 亿元，同比增长 350%。公司依托自研大模型技术，为运营商、金融及能源等行业提供智能化解决方案，助力企业数字化转型，并在多个重点领域成为 AI 应用落地的领军企业	4.0%	收入
京北方	23.61	人工智能及大数据创新产品实现收入 6400 万元，同比增长 73.4%，公司以人工智能及大数据创新产品线为核心，聚焦金融科技领域，提供 AI 风险管理与大模型平台解决方案，推动收入快速增长并助力整体毛利率提升。	2.7%	收入
焦点科技	9.19	上半年 AI 业务收入达 2174 万元，购买 AI 麦可的会员累计超过 13,000 位(不含试用体验包客户)。公司在尝试为买家推出 AI Agent 产品。公司预计于 2025 年下半年在中国制造网平台推出全新升级的 Sourcing AI，以精准响应海外买家需求。	2.4%	收入
昆仑万维	37.33	上半年，公司 AI 软件技术收入为 6527 万元。公司在二月份正式开源中国首个面向 AI 视频创作的开源视频生成模型——SkyReels-V1，四月份发布并开源全球首个基于“扩散强迫(Diffusion-forcing)”框架的无限时长电影生成模型——SkyReels-V2。此外，公司还同步开源国内首个 SOTA 级别、基于视频基座模型的表情动作可控算法——SkyReels-A1。	1.7%	收入
广联达	28.00	AI 相关订单金额超过 4000 万元,公司 AI 业务以行业 AI 平台为核心，支撑 AI 应用开发与模型服务，在 LLM、VLM 等关键技术实现突破，并聚焦设计、成本、施工三大场景落地智能产品，推动建筑行业数字化与智能化转型。	1.4%	订单



亚信科技	25.98	上半年公司完成大模型交付收入 2600 万元，另有 7000 万元大模型交付订单及 4000 万元框架订单。公司依托“基础模型+应用服务”模式，聚焦企业端大模型应用与交付，通过技术工具链和行业深耕，为客户提供高质量 AI 解决方案并构建竞争壁垒。	1.0%	收入
快手	676.54	可灵 Q2 收入超 2.5 亿元人民币，QoQ+66.7%，Q1 为 1.5 亿元。7 月公布付费用户数达到 4500 万。2025 年 5 月推出了可灵 2.1 系列，7 月做了重大升级。在 AI 大模型赋能内容和商业生态上，推出了 OneRec 推荐大模型，已经用于短视频内容推进。电商场景上使用了多模态大模型提升商品信息提炼能力。	0.4%	收入
美图公司	18.21	AI 技术赋能影像与设计产品（如“AI 闪光灯”“AI 消除”“AI 换装”“画质修复”功能、AI 设计助手 RoboNeo）是核心收入驱动力。另外，七月份推出 AI 设计助手 RoboNeo，通过自动化智能设计，用于创意生成与智能编辑。	/	/
金山办公	26.57	截至 2025 年 6 月底，WPS AI 月活跃用户数达到 2951 万，较 2024 年底的 1968 万实现显著提升。作为中国办公软件行业的开拓者与引领者，公司发布 WPS AI 3.0 并推出原生 Office 办公智能体，率先定义并开启我国办公的智能体时代，引领中国 AI 办公最佳实践。	/	/

来源：各公司 2025 半年报，各公司投资者关系记录表，国金证券研究所

2. 全球 AI 本月重大新闻

2.1 海外大模型及 AI 应用进展

· 谷歌 DeepMind 发布世界模型 Genie 3

8 月 5 日，谷歌 DeepMind 发布了通用世界模型 Genie 3，可通过一句话生成可实时交互的 3D 虚拟世界。该模型在动态一致性、物理交互模拟和多样化环境构建方面显著升级，此外，模型支持基于文本的动态事件生成功能，允许用户在固定场景中添加移动元素。

· OpenAI 发布 GPT-5

8 月 8 日，OpenAI 发布 GPT-5，作为自 2023 年 3 月 14 日 GPT-4 发布以来的重要升级版本，该模型在编程能力、减少幻觉和自动判断回复等方面的更新尤为显著，模型还引入了“安全完成”功能，旨在防止不当内容的生成。模型在 SWE-bench Verified 基准测试中以 74.9% 的成绩领先 Claude Opus 4.1、Gemini 同类模型。

· 英伟达推出 Cosmos 系列推理

8 月 11 日，英伟达发布一系列 AI 技术成果及工具，包括 Cosmos Transfer-2 世界模型、“推理”视觉语言模型 Cosmos Reason 模型和新的神经重建库与机器人 workflows 服务器。Cosmos Reason 应用于物理 AI 和机器人技术，Cosmos Transfer-2 加速合成数据生成，神经重建库的渲染技术可 3D 模拟现实世界，新服务器为机器人开发提供支持。

· Perplexity 计划以 345 亿美元收购 Chrome 浏览器

8 月 13 日，AI 浏览器公司 Perplexity 向谷歌母公司 Alphabet 发出书面要约，拟以 345 亿美元现金——几乎是 Perplexity 7 月最新一轮融资后 180 亿美元估值的两倍，收购 Chrome 浏览器。

· 谷歌上线图像生成与编辑模型 Gemini-2.5-Flash-Image

8 月 26 日，谷歌发布图像生成与编辑模型 Gemini 2.5 Flash Image (代号 nano-banana)。Gemini 2.5 Flash Image 此次更新了保持角色一致性、完全基于提示的图像编辑、原生世界知识和多图像融合四个能力。模型的价格为每百万输出 token 30 美元，每张图片为 1290 个输出 token，即每张图片需 0.039 美元。模型目前可在 Gemini API 和 Google AI Studio 中预览。

· 英伟达发布机器人计算平台 Jetson Thor

8 月 26 日，英伟达发布机器人计算平台 Jetson Thor，该平台基于最新的 Blackwell GPU 架构，使 AI 算力大幅提升至 2070 TFLOPS，比上一代提高了 7.5 倍，能效也提高了 3.5 倍，此外 Jetson Thor 还配 128GB 的超大内存。在软件优化方面，Jetson Thor 原生支持 NVIDIA Isaac、Isaac GR00T、Metropolis、Holoscan 等平台，支持从云到边缘的统一开发和部署路径。通过 FP4 量化和推测解码技术，某些模型的性能可再提升 2 倍。



2.2 国内 AI 行业动态及重点公司新闻

· 傅利叶发布全尺寸人形机器人 Care-bot GR-3

8月6日，傅利叶智能发布第三代全尺寸人形机器人 Care-bot GR-3，面向社交陪伴与辅助陪护两大应用场景。Care-bot GR-3 配备全感交互系统，通过视觉、听觉和触觉构建全新交互体验，能够进行眼神交流、回应触摸，并实现多种拟人行走姿态，在行进过程中能同步完成弯腰、下蹲等动作，满足日常交互、导览、陪伴等场景的需求。

· 字节布局 AI 硬件

8月19日，字节跳动被曝研发 AI 手机“豆包手机”（计划年底或明年初推出，初期内部测试）。字节负责大模型及部分系统工作，硬件由中兴 ODM 代工；主力团队为 AI 硬件团队 Ocean（含锤子、PICO、Ola Dance 及荣耀背景成员），隶属 Flow 部门并向朱骏汇报。此前字节尝试与手机厂商合作植入豆包模型受阻，转而自研集成硬件。

· DeepSeek 发布 V3.1-Base

8月20日，DeepSeek 发布开源基础模型 V3.1-Base，参数规模 685B，上下文长度扩展至 128K，支持 BF16 至 FP8 多精度格式。新增原生搜索功能（支持四个特殊 Token），架构去除“R1”标识或预示混合架构方向。

· 智谱推手机智能体

8月20日，智谱发布手机通用智能体 AutoGLM2.0，通过语音/文字指令操控手机或云端自动完成任务。采用云端原生方案（用户操作与 AI 物理解耦），结合在线强化学习自适应复杂环境，支持抖音、小红书等 40 余款高频应用，并开放 API 申请及开发者生态计划。

· 港大快手攻克视频生成难题

8月21日，港大与快手可灵联合发布交互式视频生成模型 Context as Memory，解决长视频场景一致性难题。核心通过记忆检索模块筛选历史帧，与待预测帧拼接后输入自回归视频扩散 Transformer，利用 Unreal Engine 5 构建含 100 视频、7601 帧/视频、12 场景风格的数据集（摄像机运动限二维平面）。

· 钉钉发布钉钉 8.0 版本

8月25日，钉钉发布钉钉 8.0 版本(AI 钉钉 1.0 版本)，该版本包含五大 AI 新产品，以及针对 AI 客服、AI 营销服和 AI 教育的三大 AI 新场景方案。AI 钉钉 1.0 的配套产品包含 Agent 驱动的工作信息流钉钉 ONE、AI 硬件 DingTalk A1.AI 听记、AI 搜问和 AI 表格。

· 群核科技开源空间语言模型 SpatialLM 1.5 和空间生成模型 SpatialGen

8月26日，群核科技开源空间语言模型 SpatialLM 1.5 和空间生成模型 SpatialGen，主要用于解决视频模型的空间一致性问题。SpatialLM 1.5 是基于大语言模型训练的空间语言模型，主要解决“理解与交互”问题，该模型支持用户通过对话交互系统 SpatialLM-Chat 进行可交互场景的端到端生成。

3. 投资建议

■ **计算机行业：**1、三季度进入国内外头部大模型厂商密集更新迭代周期，市场开始担忧大模型直接延伸到应用端，对于原有软件厂商的影响，我们认为应该一分为二的来看：1）对于产品用户相对不够多，用户体验和粘性不够深，切入用户生活或者业务不够深，比较标准化的功能，相对更容易被大模型厂商延伸到该环节去形成可能得替换，本身我们也看到国内外 AI 大模型厂商有一部分从大模型竞争中逐步退出，走向应用端，有一定的模型基础能力优势；2）对于产品用户基数大，市占率高，用户粘性强，与用户日常生活或工作业务流结合度高的产品，有数据积累、场景理解、业务认知、商务关系等多层积累的壁垒，从而相对更难被大模型厂商轻易替代，或者退一步将，即便大模型厂商看好这块，自我投入的成效或者可能性不太大，更多可能是通过入股或者并购的方式切入，预计大概率也会支付相对市价一定幅度的溢价，这个在过去互联网+时代、第一轮 AI+时代都有案例。此外，软硬结合的部分，由于涉及到供应链、产品设计、开发，渠道建设和管理等环节，相对比线上纯软更复杂，因此大模型厂商下场延伸的概率更低。2、从目前 ai 应用落地的进展看，一方面我们看好国内在 AI+软硬结合落地的优势，软件端我们看好 AI 在 C 端和出海方向落地，美图和快手的 ARR 也初步验证了这一点，此外我们还看好 AI 在 B 端与企业服务软件、制造业信息化等环节的结合，今年半年度预计陆续披露相关进展，下半年预计环比有积极增长，明后年预期收入占比进一步提高。推荐海康威视、大华股份、禾赛、迈富时、



九方智投等。

■ 电子：博通 AI 业绩超预期，ASIC 增长强劲

博通 AI 业绩超预期，ASIC 增长强劲，下游云厂商加大 AI 基础设施投资。Meta 计划到 2028 年之前，在美国建设数据中心和其他基础设施方面的投入将“至少达到 6000 亿美元”，Meta 在 2025 年资本开支预计为 660-720 亿美元，意味着未来三年将大幅增加资本开支。OpenAI 首席执行官 Sam Altman 近期表示，鉴于其最新模型 GPT-5 的需求不断增长，该公司正在优先考虑计算能力问题，并计划“在未来五个月内”将其计算集群的规模扩大一倍。我们认为下游推理需求激增，带动 ASIC 需求强劲增长，谷歌、亚马逊、Meta 等公司 ASIC 芯片快速发展，我们预测 2026 年三家公司 ASIC 芯片的数量将超过 700 万颗，OpenAI 及 xAI 等厂商也在大力推进 ASIC 芯片，未来几年也有望大幅放量。

博通、Credo AI 业绩超预期，ASIC 增长强劲

博通 FY25Q3 半导体收入中 AI 半导体收入为 52 亿美元，同比增长 63%，环比增长 8 亿美元（高于上季度 51 亿美元的指引）。XPU 业务收入占 AI 半导体收入比例为 65%。公司预计 FY25Q4 AI 收入达到 62 亿美元，环增提升至 10 亿美元。博通表示目前客户需求向上，公司总在手订单达到 1100 亿美元。除公司之前宣布今年量产的三家客户外，公司某新客户已经锁定 100 亿美元的 AI 订单，有望贡献 FY26 增量，公司 FY26 AI 收入增速有望超过 FY25。

FY26Q1 Credo 实现营收 2.23 亿美元，同比+274%，环比+31%，Non-GAAP 毛利率为 67.6%，Non-GAAP 净利润为 9828 万美元，同比+1296%，环比+51%。公司预计 FY26 全年营收增长为 120%，FY26Q3、FY26Q4 营收均实现中双位数环比增长。FY26Q1 公司营收占比超 10% 的客户超三家，营收占比分别为 35%、33%、20%。除了已有客户外，公司预计将有另外两家超大云厂商客户在 FY26 进入出货爬坡阶段，其中一家 FY26 收入占比有望超 10%。我们认为公司 AEC 未来空间广阔，除了新客户拓展外，当前 AEC 正在向 100G 单通道产品迭代，有望迎来价值量增长。

英伟达推出 Rubin CPX，面向大规模推理市场

NVIDIA 宣布推出 NVIDIA Rubin CPX，一款专为大规模情境处理而打造的新一代 GPU。Rubin CPX 使得 AI 系统能以突破性的速度与效率，处理百万个 token 的软件编码与影片生成。

Rubin CPX 与全新 NVIDIA Vera Rubin NVL144 CPX 平台中的 NVIDIA Vera CPU 和 Rubin GPU 协同工作。这款整合式 NVIDIA MGX 系统在单一机架配置下拥有 8 exaflops 的 AI 运算能力，可提供比 NVIDIA GB300 NVL72 系统高出 7.5 倍的 AI 效能，同时配备 100 TB 快速内存及每秒 1.7 PB 的内存带宽。NVIDIA 也将提供一个专属的 Rubin CPX 运算托盘 (compute tray)，以满足客户希望重复利用现有 Vera Rubin 144 系统的需求。

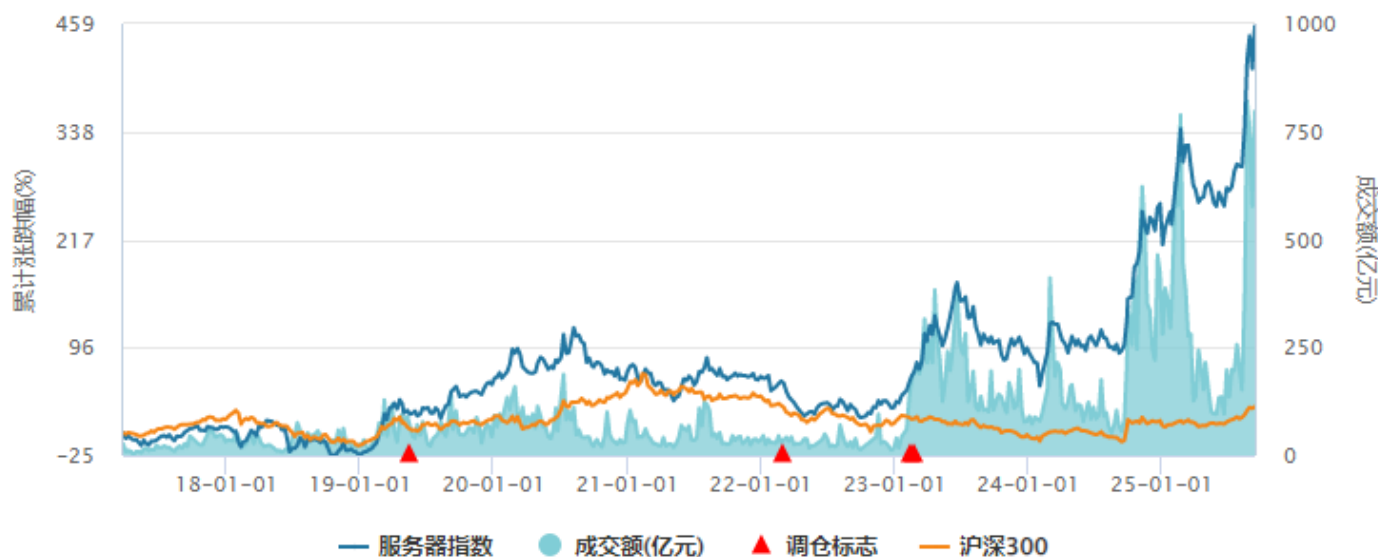
NVIDIA 创办人暨执行长黄仁勋表示：「Vera Rubin 平台将标志着 AI 运算领域的又一次跃进，不仅推出新一代 Rubin GPU，更创造名为 CPX 的全新处理器类别。正如 RTX 为图形与物理 AI 领域带来革命性的改变，Rubin CPX 是首款专为大规模情境 AI 打造的 CUDA GPU，能让模型同时处理具备数百万个词元的推理任务。」

■ 通信行业：英伟达发布 Rubin CPX，阿里发行可转债加强云基础设施建设

服务器：9.8-9.12 服务器指数+9.26%，9.1-9.12 服务器指数+1.91%。Nebius 与微软签订多年期 AI 云计算合作协议，基础合同金额 174 亿美元，并保留追加采购，总金额最高可达 194 亿美元，覆盖 2025-2031 年。Nebius 将在新泽西新数据中心为微软提供专属 GPU 算力，最快今年晚些时候开始交付。微软正持续加大资本支出以缓解 AI 云算力短缺，该交易有望显著提升其供给能力。建议关注服务器等北美数据中心建设环节的投资机会。世纪互联宣布获领先互联网公司 40MW 批发订单，将在固安 IDC 园区分期交付，助力客户核心业务。公司受益 AI 训练及推理需求增长，股价大涨超 16%，建议关注其 AIDC 机会。



图表17：服务器指数 (8841058.WI) 走势

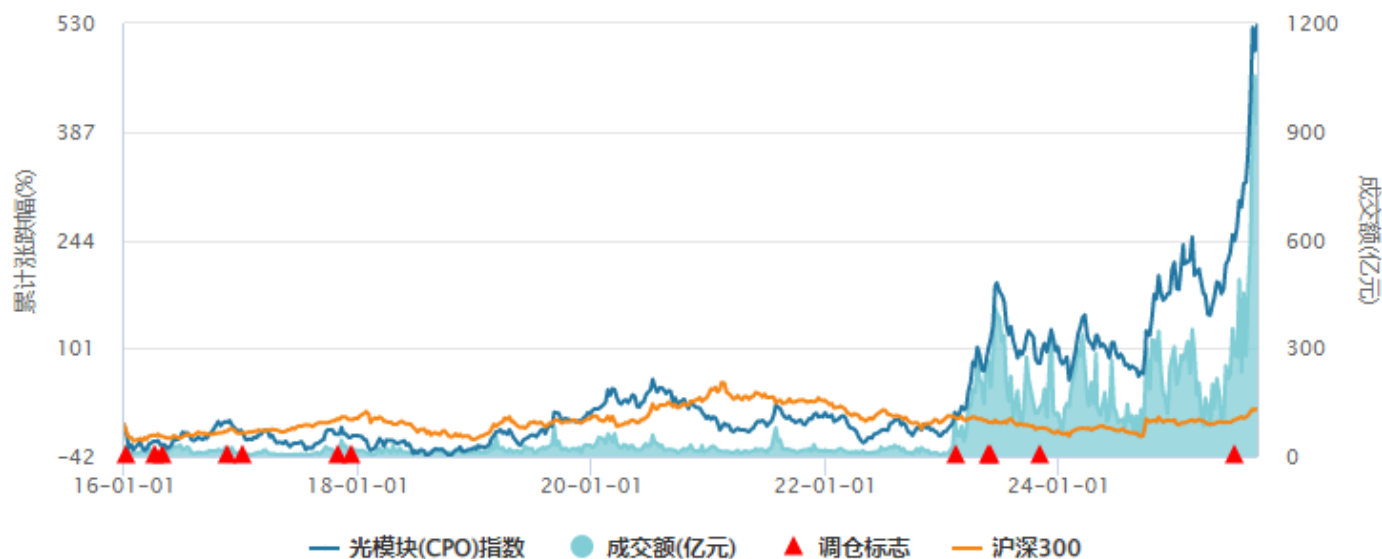


来源：wind，国金证券研究所

光模块：9.8-9.12 光模块指数+5.69%，9.1-9.12 光模块指数+0.46%。英伟达正式发布 Rubin CPX，成为首款面向百万 token 推理的 CUDA GPU，单卡算力 30 PFLOPS，注意力计算加速 3 倍。其 Vera Rubin NVL144 CPX 平台集成 144 颗 GPU 与 100TB 高速内存，单机架 AI 性能达 8 EFLOPS，约为 GB300 NVL72 平台的 7.5 倍。英伟达预计每投入 1 亿美元算力可带来高达 50 亿美元收入，显著提升 AI 视频生成、代码推理等应用的性能表现。推理效率翻倍将推动大规模硬件换代，拉动高带宽网络与光互连需求，看好 1.6T 光模块渗透率持续提升，建议重点关注高速光模块及硅光供应链厂商。

甲骨文公布第一财季 RPO 达 4550 亿美元，同比+359%，单季新增 3170 亿美元，创历史纪录。确认已与 OpenAI、xAI、Meta 等签署大额合同，OCI 正成为前沿 AI 模型训练的重要平台。基于此，公司预计 OCI 本财年收入增长 77%至 180 亿美元，并指引未来四年收入持续翻倍至 1440 亿美元。为满足激增的云需求，甲骨文计划将本财年资本开支提升至 350 亿美元，主要用于可立即产生收入的算力设备，建议关注甲骨文云供应链（嘉元科技等）。

图表18：光模块 (CPO) 指数 (8841258.WI) 走势



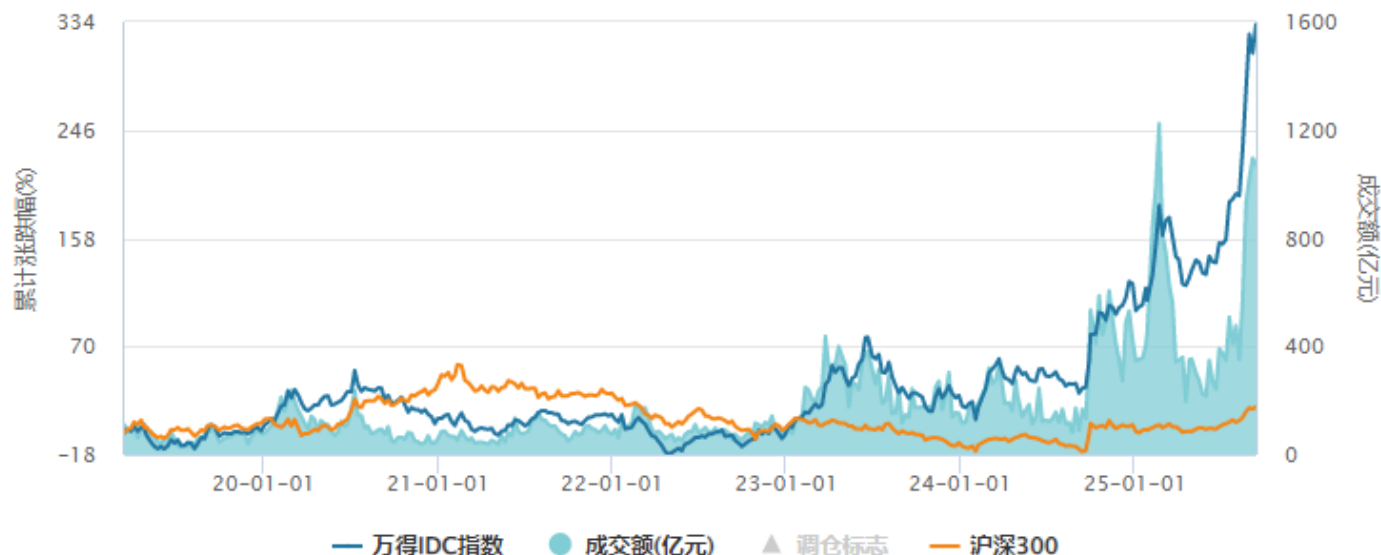
来源：wind，国金证券研究所

IDC：9.8-9.12 IDC 指数+5.65%，9.1-9.12 IDC 指数 1.83%。据 The Information 报道，



阿里、百度已开始采用自主设计的芯片训练 AI 大模型，部分替代英伟达 H20。阿里芯片已应用于轻量级模型训练，性能接近 H20；百度则试验昆仑 P800 芯片训练新版文心模型。两家公司仍使用英伟达芯片训练最前沿模型。阿里还发布下一代基础模型架构 Qwen3-Next 并开源 80B 模型，显示其在 AI 基础模型领域的技术自信。同时，阿里拟发行 32 亿美元零息可转债以扩建云基础设施、升级数据中心。国产算力建设加速，建议关注奥飞数据、世纪互联、万国数据、光环新网、数据港、科华数据等。

图表19：IDC 指数 (866052.WI) 走势



来源：wind，国金证券研究所

4. 风险提示

■ 底层大模型迭代发展不及预期

若底层大模型迭代发展不及预期，可能会影响 AI 应用落地的深度，使其难以在金融、教育、游戏等领域进行更加深入的应用。若底层大模型的蒸馏剪枝发展不及预期，可能会使其难以在边缘硬件上充分发挥性能。

■ 算力不足的风险

若训练或推理算力芯片不足，可能会推升现有芯片价格，降低 AI 应用的性价比，使得 AI 应用推广不及预期。

■ 国际关系风险

若出于国际关系原因，OpenAI 等海外大模型的调用或其他软硬件的进口受到影响，有可能使得国内 AI 应用的发展不及预期。

■ 监管风险

AI 生成内容在版权归属、个人隐私、伦理等方面存在较大争议，面临一定监管风险，可能对技术及应用层也会带来一定影响。



行业投资评级的说明：

买入：预期未来 3—6 个月内该行业上涨幅度超过大盘在 15%以上；

增持：预期未来 3—6 个月内该行业上涨幅度超过大盘在 5%—15%；

中性：预期未来 3—6 个月内该行业变动幅度相对大盘在 -5%—5%；

减持：预期未来 3—6 个月内该行业下跌幅度超过大盘在 5%以上。



特别声明：

国金证券股份有限公司经中国证券监督管理委员会批准，已具备证券投资咨询业务资格。

任何形式的复制、转发、转载、引用、修改、仿制、刊发，或以任何侵犯本公司版权的其他方式使用。经过书面授权的引用、刊发，需注明出处为“国金证券股份有限公司”，且不得对本报告进行任何有悖原意的删节和修改。

本报告的产生基于国金证券及其研究人员认为可信的公开资料或实地调研资料，但国金证券及其研究人员对这些信息的准确性和完整性不作任何保证。本报告反映撰写研究人员的不同设想、见解及分析方法，故本报告所载观点可能与其他类似研究报告的观点及市场实际情况不一致，国金证券不对使用本报告所包含的材料产生的任何直接或间接损失或与此有关的其他任何损失承担任何责任。且本报告中的资料、意见、预测均反映报告初次公开发布时的判断，在不作事先通知的情况下，可能会随时调整，亦可因使用不同假设和标准、采用不同观点和分析方法而与国金证券其它业务部门、单位或附属机构在制作类似的其他材料时所给出的意见不同或者相反。

本报告仅为参考之用，在任何地区均不应被视为买卖任何证券、金融工具的要约或要约邀请。本报告提及的任何证券或金融工具均可能含有重大的风险，可能不易变卖以及不适合所有投资者。本报告所提及的证券或金融工具的价格、价值及收益可能会受汇率影响而波动。过往的业绩并不能代表未来的表现。

客户应当考虑到国金证券存在可能影响本报告客观性的利益冲突，而不应视本报告为作出投资决策的唯一因素。证券研究报告是用于服务具备专业知识的投资者和投资顾问的专业产品，使用时必须经专业人士进行解读。国金证券建议获取报告人员应考虑本报告的任何意见或建议是否符合其特定状况，以及（若有必要）咨询独立投资顾问。报告本身、报告中的信息或所表达意见也不构成投资、法律、会计或税务的最终操作建议，国金证券不就报告中的内容对最终操作建议做出任何担保，在任何时候均不构成对任何人的个人推荐。

在法律允许的情况下，国金证券的关联机构可能会持有报告中涉及的公司所发行的证券并进行交易，并可能为这些公司正在提供或争取提供多种金融服务。

本报告并非意图发送、发布给在当地法律或监管规则下不允许向其发送、发布该研究报告的人员。国金证券并不因收件人收到本报告而视其为国金证券的客户。本报告对于收件人而言属高度机密，只有符合条件的收件人才能使用。根据《证券期货投资者适当性管理办法》，本报告仅供国金证券股份有限公司客户中风险评级高于 C3 级（含 C3 级）的投资者使用；本报告所包含的观点及建议并未考虑个别客户的特殊状况、目标或需要，不应被视为对特定客户关于特定证券或金融工具的建议或策略。对于本报告中提及的任何证券或金融工具，本报告的收件人须保持自身的独立判断。使用国金证券研究报告进行投资，遭受任何损失，国金证券不承担相关法律责任。

若国金证券以外的任何机构或个人发送本报告，则由该机构或个人为此发送行为承担全部责任。本报告不构成国金证券向发送本报告机构或个人的收件人提供投资建议，国金证券不为此承担任何责任。

此报告仅限于中国境内使用。国金证券版权所有，保留一切权利。

上海	北京	深圳
电话：021-80234211	电话：010-85950438	电话：0755-86695353
邮箱：researchsh@gjzq.com.cn	邮箱：researchbj@gjzq.com.cn	邮箱：researchsz@gjzq.com.cn
邮编：201204	邮编：100005	邮编：518000
地址：上海浦东新区芳甸路 1088 号 紫竹国际大厦 5 楼	地址：北京市东城区建国内大街 26 号 新闻大厦 8 层南侧	地址：深圳市福田区金田路 2028 号皇岗商务中心 18 楼 1806



【小程序】
国金证券研究服



【公众号】
国金证券研究