下载日志已记录，仅供内部参考

中信建投证券 CHINA SECURITIES

证券研究报告·行业动态

DeepSeek 将实施波峰波谷定价，Meta 拟筹建云基础设施业务出售富余算力

信息技术

核心观点

国产开源模型能力提升中，与海外模型差距在拉近；其性价比优势也更多被海外认知到，坚定看好 AI 产业趋势及国产模型。本周 Meta 筹建云基础设施业务的报道引发市场对 Capex 增长的担忧，同时市场对 AI 需求和大模型公司收入增长斜率放缓的担心也在继续。**我们认为：**

1）token 消耗优化是短期温和调整，并不意味着 AI 需求增长乏力。据 Evidence Lab，仅 8%的企业已在生产环境中大规模部署 AI Agent。

泛化的 Coding 场景之外，金融、法律、AIFor Science 等场景的渗透均不充分，大模型厂商均在积极推进细分场景的模型落地。

Anthropic 本周推出旗舰科研产品 Claude Science，定位"科研界的 Claude Code"，帮助自主完成计算生物学与药物研发中的实际工作。豆包、智谱均已与 Wind 在模型方面合作；MiniMax 已推出 "10x Team" 计划，高薪招募经济学、金融等顶尖专家，将行业 Know-How 转化为模型能力。

2）大模型厂商在通过降低推理成本、自研芯片、调整合作条款等方式，降低开支，提高盈利能力。本周 OpenAI 工程团队通过新方法把推理成本削减了约一半，今年 Q1 末 OpenAI 毛利率 39%，公司希望今年底达到 52%的毛利率；此前，其联合博通发布专用芯片 Jalapeño，早期测试推理成本可节省约 50%，预计于 26 年底前完成初始部署。**此外，DeepSeek 推出峰时/非峰时价格，峰时价格为非峰时的 2 倍。**

【建议关注】OpenRouter 本周 token 调用前 5 均为国产开源大模型，ZenMux 数据显示智谱 token 调用市场份额升至 22.7%，上周为 17.9%。**1）重点关注高性价比的国产大模型公司，下周均首批解禁，关注抛压风险落地后的机会：【智谱】【MiniMax】；2）**重视有垂类专业数据优势、需要精细化运营的公司，壁垒更高：**【汇量科技】【快手】【美图】；3）**云&算力相关：**【百度】【阿里巴巴】。**

游戏：高频数据向好，核心公司产品运营稳定，**关注左侧价值。**

娱乐内容：重点关注暑期档定档和宣发进展。近期多部进口动画片定档，国产电影也进入定档期，《年会不能停 2！》定档 8 月 7 日，其他头部电影预计在 6 月初开始密集定档。

维持 强于大市

杨艾莉
yangaili@csc.com.cn
010-56135145
SAC 编号:S1440519060002
SFC 编号:BQI330

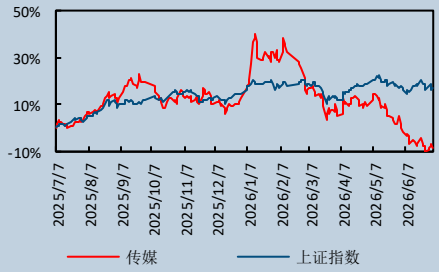
杨晓玮
yangxiaowei@csc.com.cn
010-56135147
SAC 编号:S1440523110001

蔡佳成
caijiacheng@csc.com.cn
010-56135149
SAC 编号:S1440524080008

马晓婷
maxiaoting@csc.com.cn
SAC 编号:S1440525070013

发布日期： 2026 年 07 月 05 日

市场表现



本报告由中信建投证券股份有限公司在中华人民共和国（仅为本报告目的，不包括香港、澳门、台湾）提供。在遵守适用的法律法规情况下，本报告亦可能由中信建投（国际）证券有限公司在香港提供。同时请务必阅读正文之后的免责条款和声明。

下载日志已记录，仅供内部参考

目录

一、全球 AI 进展及更新 1
1、行业观点 1
2、全球 AI 产业进展 2
大模型：本周无重磅模型发布，Meta 拟筹建云基础设施业务；继续关注 Q3 模型迭代，7 月中旬起逐步启动 2
商业化：DeepSeek 将实施波峰波谷定价；各家大模型厂商在通过工程化能力降低推理成本，推进细分场景的模型落地 3
Token 调用：本周 OpenRouter 调用量环比+1.0%，排名前五均为国产开源模型 4
重点公司更新 5
二、港股互联网&消费 12
三、游戏 14
四、广告营销 14
五、影视院线及出版 15
风险分析 17

下载日志已记录，仅供内部参考

一、全球 AI 进展及更新

1、行业观点

国产开源模型能力显著提升，月之暗面 B 端负责人黄震昕称中国模型有望年内达 Fable 级水平；国产大模型的性价比优势也更多被海外认知到，坚定看好 AI 产业趋势及国产模型。本周智谱-12.37%，MiniMax-18.83%，均有所回调，下周两家大模型厂商均首批解禁，关注回调后的投资机会。本周 Meta 筹建云基础设施业务的报道引发市场对 Capex 增长的担忧，同时市场对 AI 需求和大模型公司收入增长斜率放缓的担心也在继续。我们认为：

1）token 消耗优化是短期温和调整，并不意味着 AI 需求增长乏力。据 Evidence Lab 对约 130 家企业的最新调查显示，**仅 8%的企业已在生产环境中大规模部署 AIAgent**，29%仍在试点，26%只是使用 Copilot 或 AI Coding 等产品但尚未部署 Agent 应用，**AI Agent 真正大规模消耗 Token 的阶段才刚开始。**

泛化的 Coding 场景之外，金融、法律、AIFor Science 等场景的渗透均不充分，大模型厂商均在积极推进细分场景的模型落地。

①Anthropic 本周举办「The Briefing: AI for Science」线上发布会，推出旗舰科研产品 Claude Science，定位为"科研界的 Claude Code"，帮助自主完成计算生物学与药物研发中的实际工作，AlphaFold 之父 John Jumper 加入团队，产品即日起 Beta，向 Pro/Max/Team/Enterprise 付费用户开放。

②豆包、智谱均已与万得金融终端在模型方面合作；MiniMax 已推出 "10x Team" 计划，高薪招募经济学、金融等顶尖专家，将行业 Know-How 转化为模型能力。

2）大模型厂商在通过降低推理成本、自研芯片、调整合作条款等方式，降低开支，提高盈利能力。

①本周 OpenAI 内部工程团队通过新方法把模型推理成本削减了约一半，今年 Q1 末 OpenAI 的毛利率为 39%，公司希望今年底达到 52%的毛利率。此前，OpenAI 联合博通发布专用芯片 Jalapeño，早期测试推理成本可节省约 50%，预计将于 2026 年底前完成初始部署。

②DeepSeek 发布 DSpark 推理加速框架，可将大模型生成速度提升 60-85%。

③Anthropic 已与核心客户兼大股东 Amazon 完成合作条款重谈，计费方式将由此前的按 GPU 算力（GPU 小时）改为按 Token 付费，2027 年正式生效。

3）针对市场对美国 Hyperscalers 未来资本开支的担忧，我们认为后续应重点关注： Meta 对英伟达 Rubin 卡的采购情况，采购量大预计侧面验证公司对前沿模型的持续投资；美国 Hyperscalers 预计在 7 月 22 日开始陆续发布财报，有望在业绩交流中透露未来资本开支增减的规划。

【投资建议】伴随模型能力不断突破，监管更多介入，以及大模型厂商、用户对数据安全问题的重视度显著提升，重点关注高性价比的国产大模型。GPT-5.6 仍处政府门控阶段的"有限预览"；Claude Code 2.19.1 版本

下载日志已记录，仅供内部参考

被曝内置针对中国用户的隐蔽标记机制，团队公开承认；阿里内部已禁用 Claude Code；OpenRouter 本周 token 调用前 5 均为国产开源大模型，ZenMux 数据显示智谱 token 调用市场份额升至 22.7%，上周为 17.9%。

- ✓ **重点关注国产大模型公司：【智谱】【MiniMax】；**
- ✓ **重视有垂类专业数据优势、需要精细化运营的公司，壁垒更高：【汇量科技】【快手】【美图】**
- ✓ **云&算力相关：【百度】【阿里巴巴】**

2、全球 AI 产业进展

大模型：本周无重磅模型发布，Meta 拟筹建云基础设施业务；继续关注 Q3 模型迭代，7 月中旬起逐步启动

本周 Fable 5 重新上线，GPT-5.6 仍处政府门控阶段的"有限预览"，本周未大规模放开。

Meta 拟筹建云基础设施业务、出售富余 AI 算力引发市场广泛讨论，且存在分歧，针对市场对美国 Hyperscalers 未来资本开支的担忧，我们认为后续应重点关注：1）Meta 对英伟达 Rubin 卡的采购情况，采购量大预计侧面验证公司对前沿模型的持续投资；2）美国 Hyperscalers 预计在 7 月 22 日开始陆续发布财报，有望在业绩交流中透露未来资本开支增减的规划。

7 月 1 日，彭博社报道 Meta 宣布正在建设云基础设施业务，将向外部客户出售富余 AI 算力。据 Bloomberg 报道，Meta 内部正在讨论两种模式：提供基于其基础设施托管的 AI 模型访问，或直接出售裸算力。Meta 代表未回应 CNBC 置评请求。当日 Meta 股价收涨约 9%；CoreWeave 和 Nebius 股票当日分别下跌约 12%。

市场有两种声音：1）Meta 大模型掉队。类比 xAI 的 Grok，AI 算力租给 Anthropic 和谷歌。2）Meta 现在储备的多是 H100、H200，但是英伟达将在下半年大量出货更先进的 Rubin。所以 Meta 将旧的 H 卡对外租赁创收，储备新的 Rubin 卡训练前沿模型。核心矛盾在于对于美国 Hyperscalers 未来资本开支的担忧。

继续关注 Q3 模型更新：

1、火山引擎 FORCE 原动力大会宣布 Seedance 2.5 将于 7 月初发布，推断下周有望发布。

2、据 Business Insider，Gemini 3.5 Pro 或将于 2026 年 7 月发布。

3、MiniMax 海螺 3 近期有望发布，M3 大参数版本有望 Q3 发布。

4、Kimi K3 模型（万亿参数）或将于 Q3 发布；其 B 端负责人黄震昕称中国模型有望年内达到 Fable 级水平。

下载日志已记录，仅供内部参考

5、智谱创始人唐杰教授正在全球征集 GLM-5.3 意见，视觉能力呼声较高。

商业化：DeepSeek 将实施波峰波谷定价；各家大模型厂商在通过工程化能力降低推理成本，推进细分场景的模型落地

本周 DeepSeek 宣布 V4 正式版本 7 月中旬上线，同步启用峰时/非峰时分档定价机制，峰时价格为非峰时的 2 倍。峰时时间段为北京时间 9:00－12:00 与 14:00－18:00，该时段 API 价格按非峰时 2 倍收取。V4 Pro 非峰时缓存命中 0.025 元 / 每百万 Token，缓存未命中 3.00 元，输出 6.00 元；峰时对应 0.05 元 / 6.00 元 / 12.00 元。V4 Flash 非峰时 0.02 / 1.00 / 2.00 元；峰时 0.04 / 2.00 / 4.00 元。老模型 deepseek-chat 与 deepseek-reasoner 将在 7 月 24 日下线。

从海内外主流大模型旗舰版定价来看，海外模型依然显著高于国内，约为国内大模型的 3-5 倍，国产大模型的性价比突出。

我们认为：1）行业定价分化始终存在；性能是支撑模型高定价、提价的核心。2）以 3 美金/百万 token（输出价格）作为分界线：3 美金之上，主要是核心生产力场景，价格敏感度相对较低，用户更在意模型性能；3 美金之下，在模型能力达到一定水平的基础上，用户更关注性价比。

图 1:全球主流大模型定价梳理（统一划归为美金）

地区	厂商	旗舰模型	发布时间	输入 ($/M tok)	输出 ($/M tok)	上下文	备注
国内	DeepSeek	DeepSeek-V4-Pro（降价后）	2026/4/24	$0.42	$0.83	—	★模型4/24发布；5/22官宣永久降至原价1/4，6/1生效。原价¥12/¥24→¥3/¥6；缓存命中输入仅¥0.025/M。
国内	字节·火山引擎	Doubao-Seed-2.1-Pro	2025/6/11	$0.83	$4.17	256K	原生人民币 6/30元，较2.0 Pro提价约87.5%。
国内	MiniMax	MiniMax-M3	2026/6/3	$0.60	$2.40	—	6月初发布，首发输入/输出价格较 M2.7 翻倍，目前已永久五折；自演化 Agent 模型迭代。
国内	*智谱AI*	*GLM-5.2（旗舰，开源）*	*2026/6/13*	*$1.10*	*$3.52*	*1M*	*与GLM-5.1价格一致，GLM 5.1在GLM 5基础上涨价10%（GLM-5 原价 $1/$3.2）。*
国内	月之暗面	Kimi K2.6（最新旗舰，开源）	2026/4/20	$0.95	$4.00	256K	美元原生 $0.95/$4.00；缓存命中输入 $0.16/M。
国内	阿里·通义	Qwen3.7-Max（旗舰，标准API）	2026/5/20	$1.67	$5.00	1M	5月最新旗舰；标准API ¥12/¥36；Preview ≤32K 档为 $0.83/$3.33；兼容 OpenAI/Anthropic 协议。
海外	Google	Gemini 3.5 Flash（最新，速度档）	2026/5/19	$1.50	$9.00	1M	I/O 2026 发布；非全球区为 $1.65/$9.90；Pro 版预计6月发布。
海外	Anthropic	Claude Sonnet 4.6（参考）	2026/2/17	$3.00	$15.00	1M	Sonnet 4.8 预计本月发布；Batch=$1.5/$7.5。
海外	Anthropic	Claude Opus 4.7（旗舰）	2026/4/16	$5.00	$25.00	1M	与4.6同价；但新tokenizer可能多产 35% token，实际花费上行；Batch=$2.5/$12.5。
海外	OpenAI	GPT-5.5（旗舰）	2026/4/23	$5.00	$30.00	—	GPT-5.5 Pro 另为 $30/$180；Batch五折=$2.5/$15。
海外	OpenAI	GPT-5.6（旗舰）	2026/6/26	$5.00	$30.00	—	Terra版为 $2.5/ $15；Luna为 $1/ $6。
海外	Anthropic	Claude Fable 5 / Mythos 5（旗舰）	2026/6/9	$10.00	$50.00	1M	Mythos 级，价格为 Opus 4.8 的 2 倍；Fable=公开版,Mythos=Project Glasswing 受限版;最大 128K 输出。

数据来源：各公司官网，中信建投证券

大模型厂商在通过降低推理成本、自研芯片、调整合作条款等方式，降低开支，提高盈利能力。

1/本周 OpenAI 内部工程团队通过新方法把模型推理成本削减了约一半（The Information 报道），我们认为，成本减半这一变化可能会在两个方向体现：1）毛利率改善：今年 Q1 末 OpenAI 的毛利率为 39%，一年前为 33%，公司希望今年底达到 52%的毛利率。2）让利给用户，提供更高性价比的模型，占据更大市场份额。此前，OpenAI 联合博通发布专用芯片 Jalapeño，早期测试推理成本可节省约 50%；预计将于 2026 年底前完成初始部署。

下载日志已记录，仅供内部参考

2/DeepSeek 上周末发布 DSpark 推理加速框架，可将大模型生成速度提升 60-85%。

3/6 月 29 日 The Information 独家披露，Anthropic 已与核心客户兼大股东 Amazon 完成合作条款重谈，计费方式将由此前的按 GPU 算力（GPU 小时）改为按 Token 付费，2027 年正式生效。原口径下推理利润更多留在云厂环节，改为 Token 计费后推理价值将回到 Anthropic 侧。Amazon 方面反驳涨价说法，并透露正在评估引入 OpenAI 与自家 Nova 模型作为替代。

大模型厂商在积极推进金融、AIFor Science 等领域的模型落地。

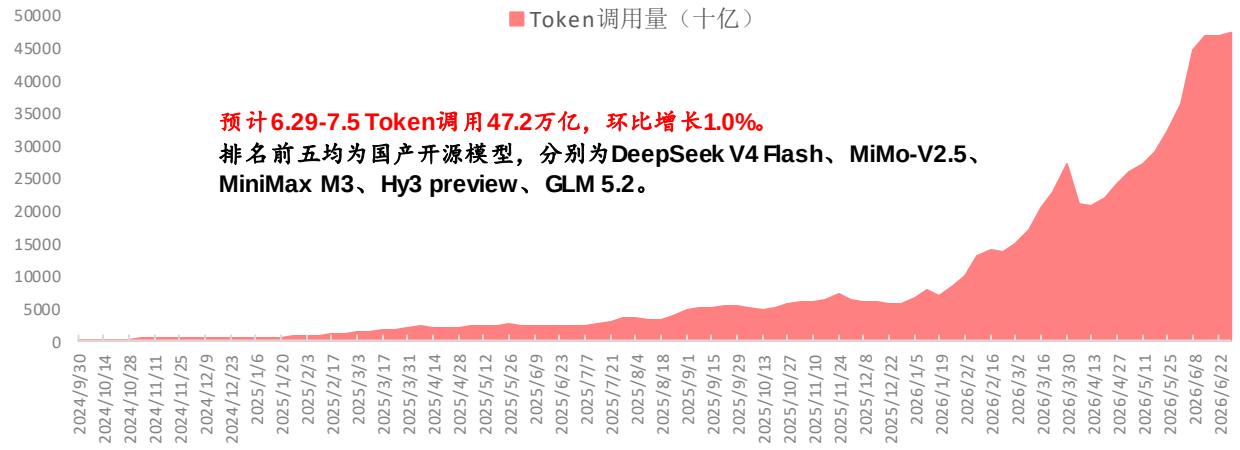
1/Anthropic 本周举办「The Briefing: AI for Science」线上发布会，推出旗舰科研产品 Claude Science，定位为"科研界的 Claude Code"，帮助自主完成计算生物学与药物研发中的实际工作，AlphaFold 之父 John Jumper 加入团队。产品即日起 Beta，向 Pro/Max/Team/Enterprise 付费用户开放。

2/GLM 系列已接入 Wind 的 AIFin Market 金融 AI 生态平台，可直接调用 Wind 数据 MCP 与精选 Skills；豆包也已与万得金融终端合作，万得推出 AI 工作平台万得 AI，豆包大模型提供技术支撑，火山引擎提供算力支撑。MiniMax 已推出 "10x Team" 计划，高薪招募经济学、金融等顶尖专家，将行业 Know-How 转化为模型能力。

Token 调用：本周 OpenRouter 调用量环比+1.0%，排名前五均为国产开源模型

本周（6.29-7.5）预计 Token 调用 47.2 万亿，环比增长 1.0%，排名前五均为国产开源模型。本周环比增速略高于上周的 0.2%，调用量前五分别为 DeepSeek V4 Flash、MiMo-V2.5、MiniMax M3、Hy3 preview、GLM 5.2。

图 2:本周 OpenRouter token 调用预计 47.2 万亿，环比+1.0%

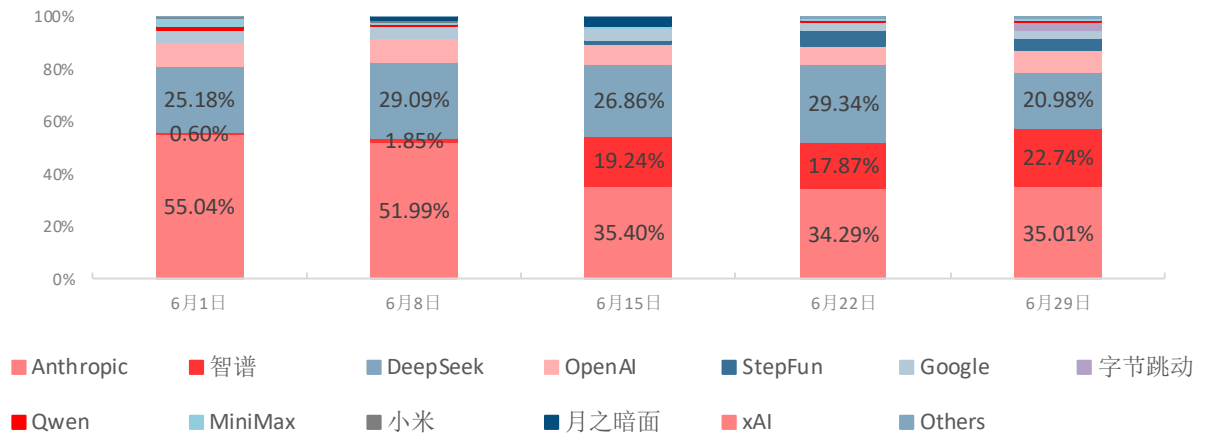


数据来源：OpenRouter，中信建投证券

下载日志已记录，仅供内部参考

本周智谱 token 市场份额继续提升（ZenMux 口径）。据 ZenMux（大模型聚合平台）的统计，近 3 周智谱 token 调用市场份额均超过 15%，本周进一步升至 22.7%，预计系智谱 GLM 5.2 编程、Agent 能力领先，在全球 AI 客户调整账单的背景下，智谱高性价比属性进一步突显。

图 3:ZenMux 的 Token 市场份额周度统计



数据来源：ZenMux，中信建投证券，注：上图日期为起始日期

重点公司更新

【Anthropic】Fable5 重新上线；Alpha Fold 之父 加入 Anthropic，AI for Science 有望成为下一个核心场景；与 Amazon 重谈计费方式

1 / Anthropic 本周举办「The Briefing: AI for Science」线上发布会，推出旗舰科研产品 Claude Science，定位为"科研界的 Claude Code"，帮助自主完成计算生物学与药物研发中的实际工作，AlphaFold 之父 John Jumper 加入团队。产品即日起 Beta，向 Pro / Max / Team / Enterprise 付费用户开放。

—产品功能：预置 60+ 精选技能与连接器，打通 UniProt、PDB、Ensembl、ClinVar、ChEMBL、GEO 等主流科研数据库，集成 NVIDIA BioNeMo 工具链（Evo 2、Boltz-2、OpenFold3 三个模型），支持本地、SSH 远程、HPC 节点运行，并可经 Modal 从单卡弹性扩到数百 GPU。

—已验证效果：Allen Institute 用约 20 个自定义技能搭多智能体模板做文献综述，把原本 2 年的时间线大幅压缩；UCSF 脑瘤中心的分子流行病学分析耗时降到约原来的 1/10；Manifold Bio 则用于组织靶向药开发与靶点提名。

—未来布局与战略：Anthropic 过去半年重金组建团队，以 4 亿美元收购 Coefficient Bio， AlphaFold 之父 John Jumper 加入团队。一方面自己下场做药，宣布用 Claude Science 自研罕见病/被忽视疾病的药物；二是通过 AI for Science Program 补贴科研生态，最多资助 50 个项目、每项至多 3 万美元额度。

我们认为 AI for Science 有望成为 Anthropic 下一个重要的商业化场景。

下载日志已记录，仅供内部参考

2 / 与 Amazon 完成合作条款重谈，计费方式由 GPU 租赁转 Token 付费，2027 年生效。6 月 29 日 The Information 独家披露，Anthropic 已与核心客户兼大股东 Amazon 完成合作条款重谈，**计费方式将由此前的按 GPU 算力（GPU 小时）改为按 Token 付费，2027 年正式生效。**原口径下推理利润更多留在云厂环节，改为 Token 计费后推理价值将回到 Anthropic 侧。Amazon 方面反驳涨价说法，并透露正在评估引入 OpenAI 与自家 Nova 模型作为替代。

我们认为，作为头部模型厂商，与战略合作伙伴重新定制计费规则，体现 Anthropic 在定价和对下游客户把控上的较强议价力；同时 Amazon 也存在反制措施，这一博弈需要从合作关系动态角度持续观察，**中长期对 Anthropic 毛利率结构性更友好。**

3/ Claude Code 2.19.1 版本被曝内置针对中国用户的隐蔽标记机制，团队公开承认。随阿里禁令同步曝光的另一条更具冲击力的信息是：**Claude Code 2.19.1 版本（今年 4 月发布）被开发者逆向发现内置了一套针对中国用户的隐蔽标记机制**——通过读取本地文件、检查代理或 API 地址中是否包含中国厂商 / AI 公司 / API 代理服务商等关键词，在系统提示词中以"隐匿术"替换标点符号方式标记中国用户身份。Claude Code 团队成员已就此**公开承认。**

我们认为，叠加阿里内部禁用 Claude Code，体现国产大模型 + 国产 Coding 工具强替代效应。理论受益方包括阿里通义千问 / Qoder、字节豆包 / TRAE、DeepSeek、MiniMax、智谱 GLM、月之暗面 Kimi、腾讯混元等国产大模型与 Agent 产品线。

【OpenAI】工程侧推理成本削减一半，联合博通发布自研推理芯片 Jalapeño；首款硬件产品将于 7.15 推出

1/据 The Information， OpenAI 内部工程团队通过新方法把模型推理成本削减了约一半。

我们认为，成本减半这一变化可能会在两个方向体现：1）毛利率改善：今年 Q1 末 OpenAI 的毛利率为 39%，一年前为 33%，**公司希望今年底达到 52% 的毛利率。**2）让利给用户，提供更高性价比的模型，占据更大市场份额。

2/此前，OpenAI 联合博通发布专用芯片 Jalapeño，早期测试推理成本可节省约 50%；预计将于 2026 年底前完成初始部署。

3 / GPT-5.6 仍处政府门控阶段的"有限预览"，本周未见大规模放开。GPT-5.6 自 6 月初起以有限预览形式向少数经政府审批的 CSP 与政府机构合作伙伴推送，本周内该状态延续。

我们认为，OpenAI 在推理成本减半 + 自研芯片双重优化下，具备了在近期与 Anthropic 及国产开源模型的定价竞争中主动降价的空间。

下载日志已记录，仅供内部参考

4/OpenAI 将推出首款硬件产品：Codex Micro 宏键盘。OpenAI 开发者官方账号在 X 平台发布推文，将和定制外设厂商 Work Louder 合作，于 7 月 15 日推出 Codex 品牌的专用输入设备 Codex Micro。此前，公司正在引入苹果视觉产品业务副总裁 Paul Meade，其将负责组建 OpenAI 硬件部门，并领导一系列 AI 硬件设备的研发。

5/股权结构变化：据 CNBC 等外媒报道，**OpenAI CEO Sam Altman 已向特朗普政府提议，将 OpenAI 5%的股权让渡给美国政府设立的主权财富基金/公共财富基金。**按 OpenAI 2026 年 3 月约 8520 亿美元的投后估值，这部分股权价值约 426 亿美元。

【Google】Gemini 月活突破 9 亿，多款新模型上线，TPU 供应链撬动 TSMC 长期垄断；下一代模型发布时间推迟至 7 月。

1 / Gemini 3.5 Live Translate 上线，支持 70+种语言实时语音互译。Google 本周公开 Gemini 3.5 Live Translate 新模型，主打实时语音到语音翻译，**自动识别 70+种语言**，同时保留说话人自然语调、消除翻译停顿间隔。

2 / Nano Banana 2 Lite 与 Gemini Omni Flash 分别在图像端与多模态视频端扩容。Google 同步推出 Nano Banana 2 Lite（Gemini 图像模型速度与成本效率最优的一档）；Gemini Omni Flash 在 API 层进入公开预览，主打原生多模态视频。

3 / Gemini 月活用户已突破 9 亿，覆盖 230 个国家与 70+种语言。Google 官方披露 **Gemini 月活用户已突破 9 亿，覆盖 230 个国家 / 地区**，用户规模层面继续维持在全球通用大模型第一梯队的绝对领先地位。

4 / SemiAnalysis 曝光：Google 下一代 TPU 将采用 Intel EMIB-T 封装，替代 TSMC CoWoS。SemiAnalysis 6 月 30 日报道，**Google 下一代 TPU 将由 Intel 采用 EMIB-T 技术封装，替代 TSMC CoWoS。**

我们认为，Google 已开始在 TPU 生态的设计端与封装端同步引入"跨供应商"策略，对 TSMC CoWoS 长期垄断先进封装的叙事形成撬动。

我们认为应重点关注公司的下一代模型的发布和核心能力排名，以及组织架构调整后在 agent 方面的能力改善。公司自 5 月以来的股价下跌主要是市场在为近期的多个负面因素定价，包括 DeepMind 的核心人员流失、目前的大模型在能力上不及美国前二、新一代大模型的发布时间推迟，但是公司已在积极调整战略，重组组织框架，在核心能力上发力，因此市场应重点关注 7 月新一代模型的发布情况。

1/7 月 1 日，谷歌宣布 Gemini Spark 登陆苹果 Mac。Spark 被整合进现有 Gemini 桌面客户端，同时上线多项新功能：支持实时跟进资讯主题，对接 Google Tasks、Google Keep，并接入 Canva、Dropbox、Instacart（生鲜配送）、OpenTable（餐饮订座）、Zillow（租房平台）等第三方应用。谷歌表示正在推出对自定义模型上下文协议（MCP）的支持，用户可直接接入自己常用的各类软件。目前 macOS 测试版仅面向美国地区的 Google AI Ultra 订阅会员开放。

下载日志已记录，仅供内部参考

——产品能力：Gemini Spark 可读取本地电脑文件，后续将支持远程任务处理；可将本地发票自动整理为预算工作表、生成谷歌办公套件文档/表格。谷歌表示后续将支持跨设备联动——在手机端向 Spark 下达多步骤复合指令，调用桌面端助手调取 Mac 本地文件资料。

——竞争定位：此次上线 macOS，使 Spark 与 Claude Desktop、微软 Copilot、OpenClaw 等桌面端 AI 智能体直接竞争。

——7 月 3 日，Gemini Omni Flash 登顶 Video Arena 视频盲测榜单。权威 AI 视频评测平台 Video Arena 更新榜单，Gemini Omni Flash 以 1404 Elo 排名第一，反超长期霸榜的字节跳动 Seedance 系列，分差 101 Elo。实测显示其在长视频叙事、镜头逻辑、动态画面稳定性方面优势显著。

【Meta】拟筹建云基础设施业务，Muse Spark 闭源模型 + 出租闲置算力两条路径并行

1 / 7 月 1 日彭博社报道 Meta 宣布筹建云基础设施业务，两条路径变现闲置算力。7 月 1 日 Meta 官方宣布筹划推出云基础设施业务，**对外出售 AI 算力与模型接入权限**。公司同步披露内部正在讨论的两条变现路径：**（1）托管模型 / API 接入（类似 AWS Bedrock）**，将 Meta 自研的 Muse Spark 系列模型托管在 Meta AI 基础设施上，向企业客户按用量收费；**（2）直接卖裸算力（类似 CoreWeave 等 Neocloud）**。消息发布当日 Meta 股价**单日上涨 8.8%，为一年多以来最强单日表现。**

2 / Muse Spark 是 Meta 自研模型近一年首个核心版本，转向闭源与延迟开放。Muse Spark 是 2026 年 4 月 Meta 超级智能实验室成立后发布的首款模型，也是 Meta 近一年首个核心 AI 模型发布。**区别于 Llama 系列，Muse Spark 未立即开放模型权重，仅对合作伙伴闭测**，开发者 API 多次延期。

我们认为，Meta 后续资本开支可以参考 xAI——两家都是自研模型短期曲折、然后发展算力租赁业务。xAI 去年资本开支 127 亿美金同比翻倍，26Q1 达 77.2 亿美金：（1）**自训模型并未放弃——Colossus 算力集群一部分依然用于自训 Grok 5（约 6T 参数）及其他 8-10T 模型**，近期还以 600 亿美金收购 Cursor；（2）**xAI 的 Colossus 同时卖裸算力，核心客户包括 Anthropic 和 Google，预计到 2030 年起算力租赁收入有望突破 500 亿美金。**Meta 参考 xAI 它的资本开支趋势大概率仍然向上；基于 Meta 更好的 C 端入口优势，其云服务将具有竞争力。

针对市场对美国 Hyperscalers 未来资本开支的担忧，我们认为后续应重点关注：1）Meta 对英伟达 Rubin 卡的采购情况，采购量大预计侧面验证公司对前沿模型的持续投资；2）**美国 Hyperscalers 预计在 7 月 22 日开始陆续发布财报，有望在业绩交流中透露未来资本开支增减的规划。**

3/ 比肩 OpenAI GPT-5.5 的 Watermelon 已在训练中。7 月 2 日，扎克伯格在内部 Town Hall 上向员工表示，过去四个月 Agent 开发“并未如我们预期那样真正加速”。同一场 Town Hall，Meta 超级智能实验室负责人 Alexandr Wang 称 Meta 正在训练的下一代模型 Watermelon，而且他表示 Watermelon“已追平 OpenAI 的 GPT-5.5”，“使用了比 Avocado（Muse Spark 内部代号）高一个数量级的算力”，并透露 Muse Spark 很快会有更新，将在编程和智能体能力上有改进。

下载日志已记录，仅供内部参考

【智谱】唐杰全球征集 GLM-5.3 意见，视觉呼声最高；本周多方生态合作推进

1 / 近期唐杰在 X 公开征集 GLM-5.3 用户需求，视觉能力呼声较高。据量子位 6 月 30 日报道，**清华教授、智谱首席科学家唐杰在 X 平台公开征集 GLM-5.3 功能建议，推文浏览量超 40 万。**评论区呼声最高的关键词是**"视觉"**，用户希望旗舰模型能看图、生成图。其他呼声较高的还包括更强 Agent 能力、超长上下文保持质量、更灵活的 API。

/ 产品现状：GLM-5.2 是纯文本模型，虽然在 Coding 评测中位居开源第一，但因缺乏视觉而在多模态竞争中落后于 Anthropic Fable 5、Kimi K2.5 和 Qwen3.5-Omni。智谱实际已具备视觉能力（GLM-5V-Turbo），但此前未整合进旗舰模型。

我们认为，唐杰以社区征集方式披露 GLM-5.3 意向，GLM-4.6 开源时唐杰也是用类似方式收集需求、后续版本逐步落地，本次征集是可信的 GLM-5.3 前置信号。

2 / 万得 AIFin Market 金融 AI 生态平台开放 MCP 能力，GLM 是主要接入模型之一。万得 5 月下旬发布 **AIFin Market 金融 AI 生态平台，**首次开放 MCP 能力给外部 AI Agent，允许其直接调用万得 **26 年历史金融数据与专业投研能力**。GLM 系列是万得 AIFin 生态的主要接入模型之一，用户可通过 GLM 直接调用 Wind 数据 MCP 与精选 Skills。

【MiniMax】关注解禁后公司赔率机会

1/股权激励：公司上市后首次股权激励，将授出共 **117 万股 A 类普通股，**本次以 **零行权价 + 不设业绩门槛 + 平均 5 年以上长归属** 的设计，将核心研发团队与公司中长期利益深度锁定。**赵鹏宇是 M2 / M3 系列背后的核心研发负责人，**发表多篇神经网络和强化学习论文，是公司核心技术骨干。按当日 **HK$515.00** 收盘价计，潜在激励总额约 **6.02 亿港币。**三部分授出：执行董事 **赵鹏宇 25 万股（1.29 亿港币、66 个月归属）**、雇员承授人 90.85 万股（4.68 亿港币、13-72 个月）、服务商 1.02 万股（0.05 亿港币、12-44 个月）。

2/ 核心股东表态，财联社报道公司两个最大外部投资人及战略股东，**阿里巴巴和米哈游明确表示长期看好 MiniMax 的发展前景，将持续陪伴公司成长。**

我们认为，本周两件事是 MiniMax 在解禁节点前的关键正向催化：1）把以赵鹏宇为代表的 M2 / M3 核心研发团队与公司中长期利益深度锁定，结合本周 Anthropic 连挖 DeepMind 三人，留住基模核心研发是公司战略安全的核心。2）阿里 + 米哈游在 7.9 减持节点前主动公开表态，解禁日大规模减持概率显著下降，后续催化看 8 月初入通窗口 + M3 大参数版发布节奏。

【DeepSeek】V4 正式版本 7 月中旬上线，同步启用峰时 / 非峰时分档定价机制

下载日志已记录，仅供内部参考

1 / 官方 7 月 1 日确认：V4 正式版本 7 月中旬发布，同步启用峰时定价机制。DeepSeek 7 月 1 日官方公告，**V4 正式版本将在 7 月中旬从预览版转正式发布**，同时启用**峰时 / 非峰时分档定价机制**。**峰时时间段**为北京时间 9:00–12:00 与 14:00–18:00，该时段 API 价格按非峰时 **2 倍收取**。老模型 deepseek-chat 与 deepseek-reasoner 将在 7 月 24 日下线。

2 / V4 Pro / V4 Flash 具体价格档已公开。V4 Pro 非峰时缓存命中 0.025 元 / 每百万 Token，缓存未命中 3.00 元，输出 6.00 元；峰时对应 0.05 元 / 6.00 元 / 12.00 元。V4 Flash 非峰时 0.02 / 1.00 / 2.00 元；峰时 0.04 / 2.00 / 4.00 元。

我们认为，DeepSeek 引入分时定价的核心含义在于中国大模型行业从"零毛利甚至负毛利抢市场"进入更理性的定价阶段。这对 MiniMax、智谱等定价策略相近的国产模型是正面 read-across。

【月之暗面（Kimi）】投前估值升至 315 亿美元；K2.7-Code 被 GitHub Copilot 集成；AI 客户调整 token 账单的背景下，kimi 作为高性价比国产开源模型也受益。

1/据《科创板日报》，月之暗面 Kimi 上一轮 200 亿美元估值融资于近日完成交割，6 月 30 日**新一轮融资已启动，投前估值涨至 315 亿美元，**半年内 Kimi 估值从 2025 年底的 43 亿美元 → 315 亿美元，累计涨超 7 倍。**近期投资者阵容包括：美团点评、中国移动、CPE 源峰、新加坡淡马锡、阿布扎比 MGX。**

2/ARR 6 月中旬突破 3 亿美元，API 收入占比超 70%，海外付费与 API 收入年增均超 400%。截至 2026 年 6 月中旬，**Moonshot AI 的 ARR 超过 3 亿美元**（3 月 1 亿 → 5 月 2 亿 → 6 月中 3 亿）。**API 收入占公司总收入超过 70%**并持续攀升；**海外付费用户增长 400%、API 收入增长 400%，**产品已进入 200+国家/地区。

3/7 月 1 日 GitHub Copilot 正式集成 Kimi K2.7-Code，托管在 Azure。7 月 1 日，**GitHub Copilot 正式将 Kimi K2.7-Code 作为可选模型上线，**托管在 Microsoft Azure，覆盖 Visual Studio Code、JetBrains、Xcode、Eclipse、Copilot CLI 等主流开发环境。K2.7-Code 参数量约 1T、256K 上下文、$0.95 / $4.00 每百万 Token，推理 Token 使用量比 K2.6 减少约 30%。

4/Coinbase CEO Brian Armstrong 在社交平台 X 披露，公司**已通过内部 LLM 网关将智谱 GLM 5.2 与月之暗面 Kimi K2.7 Code 设为全体工程师的默认大模型，**在 Token 使用量持续指数级增长的同时，将 AI 支出压缩近一半。

此外，据新浪科技，**月之暗面 B 端负责人黄震昕**在与媒体沟通中称：**“中国模型今年内达到 Fable 水平还是很有希望的”，再次强调对国内模型能力的信心。**此前，智谱创始人唐杰教授认为明年 Q1 前国内模型会达到 Fable 水平。

【字节跳动】豆包宣布 7 月 15 日下线智能体功能，Seedance 2.5 预计 7 月正式上线

1/豆包智能体功能将于 7 月 15 日下线，用户可转至字节自家猫箱 App。7 月 4 日，**字节旗下豆包宣布将于 2026 年 7 月 15 日正式下线智能体功能，**届时用户将无法新建智能体，已创建保存的智能体也将同步无法调

下载日志已记录，仅供内部参考

用和运行。豆包同步引导用户转至字节自研的**猫箱 App** 创建智能体开启对话服务。2026 年 10 月 15 日后，豆包将按《隐私政策》处理智能体相关数据，无法恢复。

2 / 下线原因：《人工智能拟人化互动服务管理暂行办法》7 月 15 日起施行。豆包 / 千问 / 元宝三家同日宣布下线智能体功能，直接原因是《人工智能拟人化互动服务管理暂行办法》自 2026 年 7 月 15 日起施行，对虚拟恋人、深度角色扮演、情感依附类 AI 设立强约束。

我们认为，这是国产 AI 应用监管从"通用大模型"进入"场景细分监管"阶段的第一个明确信号。

3/模型迭代：火山引擎 FORCE 原动力大会（6.23-6.24）发布豆包模型全家桶，**豆包 2.1“上牌桌”，Seedance 2.5 将于 7 月初发布。**豆包大模型日均 token 调用：6 月突破 180 万亿，在过去的一年增长超过 10 倍，较 3 月的 120 万亿+50%。**6 月 23 日火山引擎 FORCE 原动力大会预告的 Seedance 2.5** 视频生成模型本周继续按 7 月上线节奏推进。**产品能力升级点包括：（1）单段最长 30 秒原生视频直出、场景切换更可控；（2）最多支持 50 个全模态参考素材联合生成（行业最高）；（3）保持画面一致性的局部二次编辑能力。此外，26 年底前预计推出对标 Genie 3 的世界模型。**

4/商业化：豆包在尝试企业办公场景与飞书协同。上周豆包专业版正式上线，面向专业用户落地 C 端商业化。且在**推动金融领域的模型落地，与万得金融终端合作**，万得推出 AI 工作平台万得 AI，豆包大模型提供技术支撑，火山引擎提供算力支撑。

【昆仑万维】天工 AI 业务 ARR 突破 8 亿美元。1）7 月 2 日公司宣布，旗下天工 AI 在今年 Q2 实现 ARR 的突破，AI Native 模型与产品业务 ARR 超 8 亿美元。其中 AI 短剧平台业务 ARR 超 7 亿美元，AI 工具业务 ARR 突破 1 亿美元。此外，本周公司旗下**天工 3.2 升级，推出 Skywork Tags，直接接入用户日常协作的即时通讯工具，与人协作，**支持用户和团队在 Slack、飞书/Lark、钉钉、Discord、Telegram 里与 Skywork 互动。**2）后续重点关注 艾捷科芯新一轮融资后估值 40 亿，流片成功后估值有望再提升**。

【华策影视】成立基金加快 AI 视频生成领域的布局，参投可灵。1）7 月 2 日，公司公告拟以自有资金 4 亿元出资，与杭州亨石资产共同设立青岛华歆策达股权投资基金。此次投资旨在加快公司在 AI 视频生成领域的业务布局与生态协同。当日**可灵宣布完成首轮融资，金额 190 亿人民币，投后估值~180 亿美元，其中华策影视和芒果超媒等头部文娱产业参投。**2）**公司 AI 布局已成体系：算租（智算规模 3.7 万）**+自研垂类模型+高质量影视数据资产+内容制作场景+外部硬件投资/合作（Rokid 等）。

【行情回顾】

其他 AI 相关公司，本周普遍反弹：赤子城科技 +11.41%，美图公司 +10.80%，汇量科技 +6.77%。

美股方面，AI 相关公司分化明显：Palantir +14.50%，苹果 +8.76%，谷歌 +6.67%，Meta +5.93%，微软 +4.70%，英伟达 +1.19%，台积电 +0.42%；AMD -0.72%，博通 -1.25%，ASML -1.41%，Oracle -5.56%，超微

下载日志已记录，仅供内部参考

电脑 -11.13%，美光科技 -13.84%。

二、港股互联网&消费

【个股更新】

【腾讯】持续看好 Workbuddy 渗透率提升

WorkBuddy 同步推进：WorkBuddy 是 国内日活第一的效率智能体，QuestMobile 2026Q1 周活增速 72.2%、规模突破 20 万，产品内置 11 款主流大模型 + Auto 智能路由 + 140+ 领域专家，SkillHub 技能超 7 万种、上线两个月累计下载超 3000 万次。底层靠腾讯云 FlexKV 缓存命中率 85%、Agent Memory 让长任务 token 消耗为 API 的 39%-67%。

WorkBuddy 作为用户侧的 Token 聚合平台，实质承接了用户在真实工作流中沉淀的珍贵数据。用户提需求、模型给方案、用户采纳或修改、最终产出可用结果，编程、办公等每个真实 case 都自带"是否解决问题"的结果标签。这类带反馈闭环的高质量数据，恰恰最稀缺、最难靠爬取获得。

这构成三重价值：其一，优化 harness（工程框架），从真实调用反推工具编排、上下文管理、任务拆解的最优路径；**其二，反哺基础模型，**用闭环数据做后训练与对齐，提升真实任务成功率而非只是刷榜；**其三，沉淀垂类场景能力，**编程、文档、数据分析等高频场景的专属数据可训出更贴合场景的能力。

参照 Cursor 已验证这条路走得通：其自研 Composer 正是让模型直接在完整 Cursor 环境里调用工具、执行命令，用大规模强化学习从真实编程交互中拿反馈训练而成。最新的 Composer 2.5 在 SWE-Bench Multilingual 上达 79.8%，前代 Composer 2 的 Terminal-Bench 2.0 得分（61.7）一度反超 Claude Opus 4.6（58.0），且靠 MoE 架构做到约 250 tokens/秒、约为通用大模型 4 倍的生成速度。

【阿里巴巴】内部禁用 Claude Code 等海外模型

1/阿里内部通知全员 7 月 10 日起卸载 Claude、Claude Code 及全套 Anthropic 产品。7 月 3 日智东西独家披露，**阿里内部通知全员卸载 Anthropic 旗下所有产品，7 月 10 日正式生效。**事件源自 Anthropic 6 月 10 日的蒸馏指控（详见 Anthropic 板块）。

2/阿里官方推荐内部员工转用自研 Qoder 作为 Claude Code 替代方案。伴随禁令披露的官方指引中，阿里明确推荐员工使用自研 Qoder 作为 Claude Code 的编程 Agent 替代方案。Qoder 承接阿里内部数万开发者的 Coding 需求。同期披露，**阿里计划在 7 月内打通 Qoder CN 与 QoderWork CN 两个产品的 Credits 互通。**

3/千问下线智能体服务。7 月 4 日，千问向用户推送“千问智能体服务下线提醒”，宣布智能体功能与服务将于 7 月 15 日正式下线。同时，拟人化互动类智能体及用户自建智能体功能将于 7 月 10 日正式下线。字节跳动旗下豆包同日发布《豆包智能体功能下线通知》，同样定于 7 月 15 日下线智能体功能。

下载日志已记录，仅供内部参考

【快手】可灵完成首轮独立融资近 30 亿美元，投后估值约 180 亿美元，多模态视频赛道全球领先

1 / 7 月 2 日可灵完成首轮独立融资近 30 亿美元，投后估值约 180 亿美元。7 月 2 日快手宣布，**旗下视频生成大模型可灵 AI 完成首轮独立融资，融资规模 190 亿元人民币（约 27.95 亿美元），投后估值约 180 亿美元。**这也是全球视频生成模型公司迄今为止最大单笔融资。**快手在融资 + 股权激励后仍保持约 68.3%持股。**

2 / 投资方阵容：36 名投资者，含 CPE 源峰、中信证券、腾讯、阿里云、百度产业资本、华策、芒果等。本轮融资由 CPE 源峰、中信证券等联合领投，覆盖**产业资本（腾讯、阿里云、百度产业资本）、影视娱乐龙头（华策影视、芒果超媒）、国家级基金（中关村科技城基金）等** 36 家机构。可灵 CEO 盖坤同步获 4%股权激励（3%股权 + 1%认股权），公司总股权激励规模 15%。

3 / 可灵 ARR 轨迹与商业化目标：年底突破 10 亿美元，2027 年 12 月目标 20 亿美元。可灵 2025 年全年收入 11 亿元人民币；**2026 年 3 月 ARR 已接近 5 亿美元；预计 2026 年 12 月 ARR 突破 10 亿美元、2027 年 1 月 13 亿美元、2027 年 12 月 20 亿美元。**

【汇量科技】本周+6.77%，观点不变：公司处于即将进入新一轮增长的过渡期，**重点关注 Q4 新一代 Infra 上线及 IAP 的突破。**Q1 利润率下降主要系该阶段研发投入会高于常态，模型训练费用增加较多，Q4 新一代 Infra 上线后有望迎来拐点。**新一代 Infra 平台预计 Q2 建设完成，Q3 进入灰测、稳定性验证期，预计 Q4 正式上线，**该平台将承载更复杂的模型、容纳更多量级和更多维度的数据，提升模型预估能力，进而带动智能出价产品 ROI 提升、IAP 突破、业绩再上新台阶。

【明略科技】本周-13.84%，本周正式开源 Agent 协作产品 Octo（章鱼），支持私有化部署，人和 Agent 可以在群聊中实时对齐意图、派发任务。**目前用户可以将 OpenClaw、Codex、ClaudeCode、WorkBuddy 等主流工具接入 Octo。**

公司 Agent 业务正快速起量，token 工厂潜力股，有望与腾讯在 AI 应用领域深度合作，重点关注解禁（7 月底解禁）落地后的机会。1）预计公司 Agentic Service 收入今、明年将数倍的增长；推出的 **Octo 为人+Agent 协作平台，类似 AI 版飞书，**是公司向 Agentic AI 转型的重要产品，有望推动商业模式向按 token 收费转型。2）公司有望依托与腾讯的股权、业务合作深化 AI 产品领域的合作。腾讯是公司最大外部股东，IPO 后持股超 25%；25 年 10 月双方即签订协议，在营销/营运智能、行业解决方案及云服务等领域深度协作；公司私域工具（微伴助手等）本就根植于企业微信/微信生态，未来 Octo 等 AI Agent 产品有望依托与腾讯的股权关系，与腾讯在企业级 AI Agent 落地、AI 应用等领域展开更深度合作。

【行情回顾】

恒生科技指数，本周反弹 +5.72%。其中，涨幅靠前的包括：京东健康 +12.88%，百度集团 +12.30%，美

下载日志已记录，仅供内部参考

团 +11.44%，腾讯音乐 +10.43%，哔哩哔哩 +10.43%，网易 +9.55%，京东集团 +8.54%，商汤 +7.81%，小米集团 +7.19%，阿里健康 +7.17%，携程集团 +6.23%，快手 +6.00%，阿里巴巴 +5.14%，腾讯控股 +4.71%。跌幅靠前的包括：MiniMax -18.83%，智谱 -12.37%，同程旅行 -1.79%。

三、游戏

【行业观点】

高频数据向好，板块整体估值目前在 10～12x 处于历史超低区间。核心公司产品运营稳定，关注左侧价值。

【个股更新】

【世纪华通】WoS、Kingshot、Tasty Travels 今年以来月度流水持续增长，我们预估二季度公司业绩将实现明显的同环比增长。Hotel Legacy/Just Desserts/Family Farm Merge 等新品有望年内大推。

【巨人网络】超自然已海外上线，暑期 DAU 和流水有望再上台阶。Q2 业绩有望同环比显著增长。

【完美世界】关注异环后续 1.2 版本、真红等新角色、Steam 等新的主机/PC 渠道开放。

【心动公司】期待《心动小镇》暑期重点 IP 联动与运营活动。

【行情回顾】

A 股游戏板块，本周整体上涨，板块内明显分化。其中涨幅靠前的包括：昆仑万维 +10.09%，神州泰岳 +9.13%，恺英网络 +8.46%，三七互娱 +5.43%，浙数文化 +3.40%，电魂网络 +2.62%，吉比特 +1.96%，盛天网络 +1.44%，游族网络 +1.15%，完美世界 +0.69%。跌幅靠前的包括：冰川网络 -2.27%，世纪华通 -1.54%，顺网科技 -1.06%，巨人网络 -1.06%。

四、广告营销

【个股更新】

【汇量科技】本周继续回调，-20.27%，公司处于即将进入新一轮增长的过渡期，**后续重点关注 Q4 新一代 Infra 上线及 IAP 的突破。**Q1 利润率下降主要系该阶段研发投入会高于常态，模型训练费用增加较多，Q4 新一代 Infra 上线后有望迎来拐点。**新一代 Infra 平台预计 Q2 建设完成，Q3 进入灰测、稳定性验证期，预计**

下载日志已记录，仅供内部参考

Q4 正式上线，该平台将承载更复杂的模型、容纳更多量级和更多维度的数据，提升模型预估能力，进而带动智能出价产品 ROI 提升、IAP 突破、业绩再上新台阶。

【分众传媒】本周+2.94%，目前股息率约 6.4%。近期回调预计主要系市场担心消费疲软、投放增长乏力，及资金面的“虹吸效应”。**6%以上的股息率是公司估值的安全垫，重点关注 Q2 消费投放是否企稳。**收购新潮目前已获得交易所受理，现处于问询阶段。

【易点天下】本周小幅回调，-2.56%，继续关注 AI 模型与算力智能调度平台 TokenLX、及程序化业务进展。

【行情回顾】

广告营销，本周广告营销（申万）指数+1.22%，涨幅小于传媒板块的+1.94%。涨幅靠前的标的是：福石控股+15.11%、ST 华扬+14.20%、旗天科技+8.88%、ST 华韵+7.99%、新华都+7.47%；跌幅较高的标的是三人行-3.29%、浙文互联-3.25%、易点天下-2.56%、天龙集团-1.97%。

五、影视院线及出版

【在线音乐】Spotify 积极布局直播音乐和 AI 授权

- Spotify 宣布从纯音频流媒体向“多格式媒体平台”转型：1）测试现场音乐直播（已用 Dua Lipa 墨西哥城演唱会素材试水）；2）向部分用户开放演唱会门票提前购买权限；3）与环球音乐签署 AI 授权协议，允许 Premium 用户用 AI 工具创作翻唱和混音。

关注腾讯音乐与喜马拉雅的业务合作进展，以及网易云音乐的高频数据。

【影视】暑期档逐步开启，关注进口片表现，持续关注国产片定档进展。进口片方面，《蜘蛛侠：崭新之日》定档 7 月 29 日；国产片方面，继续重点关注《欢迎来龙餐馆》《转念花开》《神探之痕迹》等头部电影定档进展。

持续推荐中国电影、猫眼娱乐。关注大麦娱乐、儒意电影等票务和院线公司。

下载日志已记录，仅供内部参考

【短剧漫剧】上海电视节上 AI 短剧成主角

- 6 月 22 日-26 日，第 31 届上海电视节将微短剧列为白玉兰论坛核心议题。行业数据显示，2026 年一季度 AI 短剧已占全行业上线总量的 95%以上，《朱雀堂》等头部精品短剧成本飙升至 4000 万元，收益天花板也拉到 4800 万元。七猫短剧、芒果 TV 大芒计划等平台一致认为“真人短剧和 AI 短剧都进入了深水区”。

- “长短剧全球买卖”论坛上，阅文、爱奇艺、DramaWave 等头部企业共同探讨出海格局。AI 使产能从“月产 20 部”跃升至“月产 200+部”，北美仍是付费意愿最强的核心市场，拉美等新兴市场处于探索初期。

关注 Netflix 的短剧布局情况，以及阅文集团、掌阅科技等头部短剧漫剧公司近况。

【行情回顾】

影视院线：涨跌幅明显强于传媒指数。上周影视院线（申万）指数上涨 6.08%，明显强于传媒（申万）指数的上涨 1.94%。从个股来看，上周涨跌幅最高的三家公司分别是奥飞娱乐（+17.66%）、幸福蓝海（+14.32%）、华谊兄弟（+11.32%）。上周涨跌幅靠后的三家公司分别为博纳影业（-12.16%）、柠萌影业（-6.90%）、IMAX China（-2.62%）。

出版：涨跌幅基本持平传媒指数。上周出版（申万）指数上涨 1.85%，基本持平传媒（申万）指数的上涨 1.94%。从个股来看，上周涨跌幅最高的三家公司分别是新华传媒（+8.95%）、山东出版（+7.45%）、中原传媒（-6.98%）。上周涨跌幅靠后的公司分别为粤传媒（-5.26%）、荣信文化（-2.50%）、世纪天鸿（-2.37%）。

下载日志已记录，仅供内部参考

风险分析

版权保护力度不及预期，知识产权未划分明确的风险，用户隐私数据泄露风险，与IP或明星合作中断的风险，大众审美取向发生转变的风险，竞争加剧的风险，用户付费意愿低的风险，消费习惯难以改变的风险，关联公司治理风险，内容上线表现不及预期的风险，生成式AI技术发展不及预期的风险，产品研发难度大的风险，产品上线延期的风险，营销买量成本上升风险，人才流失的风险，人力成本上升的风险，政策监管的风险，商业化能力不及预期的风险。



下载日志已记录，仅供内部参考

分析师介绍

杨艾莉

中信建投证券传媒互联网行业首席分析师，中国人民大学传播学硕士，曾任职于百度、新浪，担任商业分析师、战略分析师。2015 年起，分别任职于中银国际证券、广发证券，担任传媒与互联网分析师、资深分析师。2019 年 4 月加入中信建投证券研究发展部担任传媒互联网首席分析师。曾荣获 2019 年 wind 资讯传播与文化行业金牌分析师第一名；2020 年 wind 资讯传播与文化行业金牌分析师第二名；2020 年新浪金麒麟评选传媒行业新锐分析师第二名。

杨晓玮

中信建投证券传媒互联网行业联席首席分析师，上海交通大学金融硕士，主要研究全球大模型与应用、互联网平台、全球游戏行业及潮流玩具。

蔡佳成

中信建投证券全球 AI 创新与互联网行业研究员，哥伦比亚大学硕士、波士顿大学学士，主要研究传媒互联网领域。

马晓婷

中信建投证券全球 AI 创新与互联网行业研究员，上海财经大学学士，上海社会科学院硕士，主要研究全球大模型与 AI 应用、广告营销及长视频平台。

下载日志已记录，仅供内部参考

评级说明

投资评级标准		评级	说明
报告中投资建议涉及的评级标准为报告发布日后6个月内的相对市场表现，也即报告发布日后的6个月内公司股价（或行业指数）相对同期相关证券市场代表性指数的涨跌幅作为基准。A股市场以沪深300指数作为基准；新三板市场以三板成指为基准；香港市场以恒生指数作为基准；美国市场以标普500指数为基准。	股票评级	买入	相对涨幅15％以上
		增持	相对涨幅5%—15%
		中性	相对涨幅-5%—5%之间
		减持	相对跌幅5%—15%
		卖出	相对跌幅15％以上
	行业评级	强于大市	相对涨幅10%以上
		中性	相对涨幅-10-10%之间
		弱于大市	相对跌幅10%以上

分析师声明

本报告署名分析师在此声明：（i）以勤勉的职业态度、专业审慎的研究方法，使用合法合规的信息，独立、客观地出具本报告，结论不受任何第三方的授意或影响。（ii）本人不曾因，不因，也将不会因本报告中的具体推荐意见或观点而直接或间接收到任何形式的补偿。

法律主体说明

本报告由中信建投证券股份有限公司及/或其附属机构（以下合称“中信建投”）制作，由中信建投证券股份有限公司在中华人民共和国（仅为本报告目的，不包括香港、澳门、台湾）提供。中信建投证券股份有限公司具有中国证监会许可的投资咨询业务资格，本报告署名分析师所持中国证券业协会授予的证券投资咨询执业资格证书编号已披露在报告首页。

在遵守适用的法律法规情况下，本报告亦可能由中信建投（国际）证券有限公司在香港提供。本报告作者所持香港证监会牌照的中央编号已披露在报告首页。

一般性声明

本报告由中信建投制作。发送本报告不构成任何合同或承诺的基础，不因接收者收到本报告而视其为中信建投客户。

本报告的信息均来源于中信建投认为可靠的公开资料，但中信建投对这些信息的准确性及完整性不作任何保证。本报告所载观点、评估和预测仅反映本报告出具日该分析师的判断，该等观点、评估和预测可能在不发出通知的情况下有所变更，亦有可能因使用不同假设和标准或者采用不同分析方法而与中信建投其他部门、人员口头或书面表达的意见不同或相反。本报告所引证券或其他金融工具的过往业绩不代表其未来表现。报告中所含任何具有预测性质的内容皆基于相应的假设条件，而任何假设条件都可能随时发生变化并影响实际投资收益。中信建投不承诺、不保证本报告所含具有预测性质的内容必然得以实现。

本报告内容的全部或部分均不构成投资建议。本报告所包含的观点、建议并未考虑报告接收人在财务状况、投资目的、风险偏好等方面的具体情况，报告接收者应当独立评估本报告所含信息，基于自身投资目标、需求、市场机会、风险及其他因素自主做出决策并自行承担投资风险。中信建投建议所有投资者应就任何潜在投资向其税务、会计或法律顾问咨询。不论报告接收者是否根据本报告做出投资决策，中信建投都不对该等投资决策提供任何形式的担保，亦不以任何形式分享投资收益或者分担投资损失。中信建投不对使用本报告所产生的任何直接或间接损失承担责任。

在法律法规及监管规定允许的范围内，中信建投可能持有并交易本报告中所提公司的股份或其他财产权益，也可能在过去12个月、目前或者将来为本报告中所提公司提供或者争取为其提供投资银行、做市交易、财务顾问或其他金融服务。本报告内容真实、准确、完整地反映了署名分析师的观点，分析师的薪酬无论过去、现在或未来都不会直接或间接与其所撰写报告中的具体观点相联系，分析师亦不会因撰写本报告而获取不当利益。

本报告为中信建投所有。未经中信建投事先书面许可，任何机构和/或个人不得以任何形式转发、翻版、复制、发布或引用本报告全部或部分内容，亦不得从未经中信建投书面授权的任何机构、个人或其运营的媒体平台接收、翻版、复制或引用本报告全部或部分内容。版权所有，违者必究。

中信建投证券研究发展部

北京
朝阳区景辉街16号院1号楼18层
电话：（8610） 56135088
联系人：李祉瑶
邮箱：lizhiyao@csc.com.cn

上海
上海浦东新区浦东南路528号南塔2103室
电话：（8621） 6882-1600
联系人：翁起帆
邮箱：wengqifan@csc.com.cn

深圳
福田区福中三路与鹏程一路交汇处广电金融中心35楼
电话：（86755）8252-1369
联系人：曹莹
邮箱：caoying@csc.com.cn

中信建投（国际）

香港
中环交易广场2期18楼
电话：（852）3465-5600
联系人：刘泓麟
邮箱：charleneliu@csci.hk