下载日志已记录，仅供内部参考

2026 年 09 月 07 日

华鑫证券 CHINA FORTUNE SECURITIES

GPT-6 Astra 推出，Claude Fable 5.1 发布

—计算机行业周报

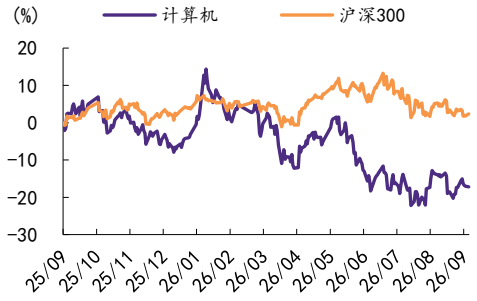
推荐(维持)

分析师：任春阳 S1050521110006
rency@cfsc.com.cn

行业相对表现

表现	1M	3M	12M
计算机(申万)	-4.1	-4.7	-16.6
沪深 300	-2.5	-5.0	2.6

市场表现



资料来源：Wind，华鑫证券研究

相关研究

1、《计算机行业点评报告：英伟达（NVDA.O）：Q2 业绩保持增长，VeraRubin 量产打开长期成长空间》2026-09-04

2、《计算机行业周报：智谱发布 GLM-5.3-Flash，Qwen3.8-Flash-Next 开源》2026-09-01

3、《计算机行业点评报告：Roblox：Q2 营收与自由现金流强劲增长，AI 赋能 UGC 创作生态再扩容》2026-08-30

投资要点

算力：阿里云 AgentRun 平台“万有无界”，Claude Fable 5.1 实现性能降本双突破

2026 年 9 月 2 日，Anthropic 正式发布新一代旗舰大模型 Claude Fable 5.1 与面向科研的 Mythos 5.1。Fable 5.1 在主动性科研、自主编程等全部 8 项基准测试中位居首位，在长程科研智能体测试 Terminal-Bench-Science 中取得 52.6%的得分；同时 API 缓存读取资费下调 75%至 0.25 美元/百万 Token，高智能体属性任务场景下综合成本最高可节约 45%，有力推动生产级智能体业务规模化落地。

AI 应用：文心一言周时长环比+19.44%，GPT-6 Astra 推出

2026 年 9 月 4 日，OpenAI 正式发布了 GPT-6 Astra。公司联合创始人兼首席执行官萨姆·奥尔特曼表示，该模型在计算机应用、专业工作、科学研究、编程及网络安全等多个高要求领域均达到了业界最优水平。在核心能力测评方面，GPT-6 Astra 的表现突出。在 FrontierMath Tier 4 测试中，它取得了 97.6%的高分，并协助解决了数学领域长期存在的开放性问题；而在 ARC-AGI-3 测试中，其得分高达 99.9%。此外，在 ExploitBench 测试中，该模型更是获得了 100%的满分。

AI 融资动向：Wonderful 完成 5.5 亿美元 C 轮融资，投后估值达 50 亿美元

2026 年 9 月 2 日，企业级 AI 操作系统研发商 Wonderful 宣布完成 5.5 亿美元 C 轮融资，投后估值达 50 亿美元，由 Insight Partners 领投、Salesforce 战略跟投。资金将主要用于扩充全球前置部署工程师团队及深化产品研发，体现产业资本对企业级工作流编排与 AI OS 赛道商业化前景的高度认可。

投资建议

近日，Snowflake 发布 2027 财年第二季度财报。当季总收入 15.5 亿美元，同比增长 35%；其中产品收入 14.9 亿美元，同比增长 37%，公司 CFO 表示这是产品收入增速连续第三个季度抬升，由核心数据平台与 AI 收入同步走强驱动；剩余履约义务（RPO）90 亿美元，同比增长 30%；Non-GAAP 营业利润率 15%，上年同期为 11%。AI 业务放量明显：CoCo 账户数突破 9,100 家、单季净增超 2,000 家，CoWork 账户数扩至 5,800 家，上半财年超 330 项产品功能正式发布、同比增长 35%。指引方面，公司将 FY2027 全年产品收入指引由 58.40 亿美元上

下载日志已记录，仅供内部参考

调至 60.70 亿美元（对应同比增速由 31%上调至 36%），Non-GAAP 营业利润率指引由 13.5%上调至 14.5%，3QFY27 产品收入指引 15.88－15.93 亿美元（同比增长 37－38%）。

本次财报的核心信号在于增长加速与利润率扩张同时发生：需求端，产品收入增速连续三个季度抬升至 37%，RPO 同比增长 30%，公司 CFO 明确归因于核心数据平台与 AI 收入同步走强，表明企业客户的 AI 工作负载已进入数据平台的实际消耗阶段，而非停留在试点；盈利端，Non-GAAP 营业利润率同比提升四个百分点至 15%，全年指引收入增速与利润率同时上调，AI 投入期的经营杠杆开始兑现。我们认为，企业 AI 从试点走向生产环境后，数据打通、治理与合规等底层能力成为规模化落地的前提，数据平台的战略价值随之系统性抬升，Snowflake 本季产品收入加速与 AI 产品账户的快速扩张（CoCo 超 9,100 家、CoWork5,800 家）正是这一趋势的直观体现。

中长期，建议关注专注于功率半导体及模拟芯片测试系统研发的联动科技（301369.SZ）、专注于半导体等高端制造业的罗博特科（300757.SZ）、新能源业务高增并供货科尔摩根等全球电机巨头的唯科科技（301196.SZ）、AI 智能文字识别与商业大数据领域巨头的合合信息（688615.SH）、深耕工业 AI 与软件并长期服务高端装备等领域头部客户的能科科技（603859.SH）。

风险提示

1）AI 底层技术迭代速度不及预期。2）政策监管及版权风险。3）AI 应用落地效果不及预期。4）推荐公司业绩不及预期风险。

重点关注公司及盈利预测

公司代码	名称	2026-09-07	EPS			PE			投资评级
		股价	2025	2026E	2027E	2025	2026E	2027E	
300757.SZ	罗博特科	631.94	-0.40	2.13	4.49	-1579.85	296.69	140.74	买入
301196.SZ	唯科科技	88.68	2.05	1.92	2.39	39.28	46.16	37.05	买入
301369.SZ	联动科技	128.08	0.48	1.08	2.62	266.83	118.59	48.89	买入
603859.SH	能科科技	38.20	0.92	1.05	1.32	41.52	36.38	28.94	买入
688615.SH	合合信息	94.88	3.24	4.22	5.25	29.28	22.48	18.07	买入

资料来源：Wind，华鑫证券研究

下载日志已记录，仅供内部参考

正文目录

1、 算力动态：阿里云公测 AGENT 平台“万有无界”，CLAUDE FABLE 5.1 实现性能降本双突破 4
1.1、 Tokens 跟踪 4
1.2、 云厂动态：阿里云企业级 Agent 协作平台“万有无界”开启公测 6
1.3、 产业动态：Anthropic 发布 Claude Fable 5.1 与 Mythos 5.1，8 项基准霸榜且缓存读取降价 75% 7
2、 AI 应用动态：文心一言周时长环比+19.44%，GPT-6 ASTRA 推出 9
2.1、 周流量跟踪：文心一言周时长环比+19.44% 9
2.2、 产业动态：GPT-6 Astra 推出，实现性能与安全性双重进阶 9
3、 AI 融资动向：WONDERFUL 完成 5.5 亿美元 C 轮融资，企业级 AI 操作系统加速落地 14
4、 行情复盘 16
5、 投资建议 18
6、 风险提示 19

图表目录

图表 1：TOKENS 规模 LEADERBOARD（TOKEN 消耗量为统计口径） 6
图表 2：文本类模型市场份额占据示意图（REQUEST 调用量为统计口径） 6
图表 3：FABLE 5.1 基准测试结果 7
图表 4：不同任务负载下 FABLE 5 与 5.1 调用成本对比 8
图表 5：2026.8.27-2026.9.2 AI 相关网站流量 9
图表 6：GPT-6 ASTRA 在多项测试中的表现 10
图表 7：GPT-6 ASTRA 在 ARC-AGI-3 评估中的表现 10
图表 8：GPT-6 ASTRA 在 AGENTS' LAST EXAM 测试中的表现 11
图表 9：GPT-6 ASTRA 在 BENCHCAD 测试中的表现 12
图表 10：GPT-6 ASTRA 在 FRONTIERCODE 1.1 EXTENDED 评估中的表现 12
图表 11：GPT-6 ASTRA 在 GPQA DIAMOND 测试中的表现 13
图表 12：上周 AI 初创公司融资动态 15
图表 13：上周（2026.8.31-2026.9.4 日）指数日涨跌幅 16
图表 14：上周（2026.8.31-2026.9.4 日）AI 算力指数内部涨跌幅度排名 16
图表 15：上周（2026.8.31-2026.9.4 日）AI 应用指数内部涨跌幅度排名 17
图表 16：FICONTEC2025 年年中至今公告订单 18
图表 17：重点关注公司及盈利预测 19

下载日志已记录，仅供内部参考

1、 算力动态：阿里云公测 Agent 平台“万有无界”，Claude Fable 5.1 实现性能降本双突破

1.1、 Tokens 跟踪

根据 OpenRouter 公开数据，2026 年 8 月 31 日至 9 月 6 日，周度 Token 消耗量保持温和增长，消耗量为 115T，环比上周增加 1.77%。在 Tokens 规模 Leaderboard 前五名中，Tencent 的 Hy4 preview 以 14.7T tokens 位居榜首；OpenAI 的 GPT-5.6 Luna 以 12.9T tokens 位居第二；Z-ai 的 GLM 5.3 Flash 与 DeepSeek 的 DeepSeek V4 Flash 0731 均以 12.4T tokens 并列第三；DeepSeek 的 DeepSeek V4 Flash 0423 以 5.19T tokens 位居第五。

文本类模型的市场份额，DeepSeek 以 1.21B requests 占据 23.8%的份额，位列第一。openai、google、z-ai 则分别以 1.02B、1.01B、494M requests，对应占据 20.0%、19.8%、9.7%的市场份额。

图像生成类模型调用量略有回落，调用量为 5.75M requests，环比上周下降约 0.17%。在 Image Leaderboard 中，Nano Banana 2 (Gemini 3.1 Flash Image)以 1.2M requests 位居第一；Nano Banana 2 Lite 以 1.19M requests 位居第二；Nano Banana 以 548K requests 位列第三；OpenAI 的 GPT Image 2 以 513K requests 位列第四；Nano Banana 2 (Gemini 3.1 Flash Image Preview)以 441K requests 位列第五。从市场份额来看，google 以 3.98M requests 占据 69.2%的份额，位居第一；openai 以 778K requests 占据 13.5%，位列第二；black-forest-labs、x-ai、bytedance-seed 则分别以 244K、225K、176K requests，对应占据 4.2%、3.9%、3.1%的市场份额。

嵌入模型调用量小幅上升，调用量为 295M requests，环比上周增加 3.15%。在 Embedding Leaderboard 前五名中，OpenAI 的 Text Embedding 3 Small 以 127M requests 位居榜首；Qwen 的 Qwen3 Embedding 8B 以 49.4M requests 位列第二；OpenAI 的 Text Embedding 3 Large 以 25.8M requests 位列第三；Google 的 Gemini Embedding 2 以 21.1M requests 位列第四；BAAI 的 bge-m3 以 19.5M requests 位列第五。从市场份额来看，openai 以 156M requests 占据 53.0%的份额，位居第一；qwen 以 54.4M requests 占据 18.5%，位列第二；google、baai、perplexity 则分别以 36.1M、20.3M、11.8M requests，对应占据 12.2%、6.9%、4.0%的市场份额。

重排序模型调用量大幅上升，调用量为 8.3M requests，环比上周增加 30.30%。在 Rerank Leaderboard 中，rerank-2.5 以 2.78M requests 位列榜首；Rerank v3.5 以 2.65M requests 位居第二；Rerank 4 Fast 以 1.53M requests 位列第三；rerank-2.5-lite 以 539K requests 位居第四；Rerank 4 Pro 以 377K requests 位居第五。从市场份额来看，cohere 以 4.55M requests 占据 54.8%的份额，位居第一；voyageai 以 3.32M requests 占据 40.0%的份额，位居第二；qwen、nvidia 分别以 335K、97K requests，对应占据 4.0%、1.2%的市场份额。

视频生成类模型调用量有所上升，调用量为 184K requests，环比上周增加 8.77%。在 Video Leaderboard 前五名中，ByteDance 的 Seedance 2.0 Mini 以 30K requests 位列第一；

下载日志已记录，仅供内部参考

Google 的 Veo 3.1 Lite 以 30K requests 并列第一；ByteDance 的 Seedance 2.5 以 19K requests 位列第三；ByteDance 的 Seedance 2.0 Fast 以 18K requests 位列第四；ByteDance 的 Seedance 2.0 以 11K requests 位列第五。从市场份额来看，bytedance 以 85K requests 占据 45.7%的份额，位列第一；google 以 43K requests 占据 22.9%的份额，位居第二；x-ai、alibaba、minimax 则分别以 16K、14K、14K requests，对应占据 8.6%、7.5%、7.4%的市场份额。

语音生成类模型调用量大幅回升，调用量为 2.46M requests，环比上周增加 15.49%。在 Speech Leaderboard 前五名中，Google 的 Gemini 3.1 Flash TTS Preview 以 1.25M requests 位居第一；Hexgrad 的 Kokoro 82M 以 512K requests 位列第二；xAI 的 Grok Voice TTS 1.0 以 194K requests 位居第三；Deepgram 的 Aura-2 以 113K requests 位列第四；Fish-audio 的 S2.1 Pro Free (free)以 91K requests 位列第五。从市场份额来看，google 以 1.25M requests 占据 50.8%的份额，位居第一；hexgrad 以 512K requests 占据 20.8%，位列第二；、fish-audio、deepgram 则分别占据 7.9%、7.1%、7.0%的市场份额。

语音转写类模型调用量大幅上涨，调用量为 16.6M requests，环比上周增加 21.17%。在 Transcription Leaderboard 前五名中，OpenAI 的 Whisper Large V3、Whisper Large V3 Turbo 分别以 8.37M、3.97M requests 位列第一和第二；OpenAI 的 GPT-4o Mini Transcribe 以 709K requests 位列第三；Microsoft 的 MAI-Transcribe 2 以 701K requests 位列第四；NVIDIA 的 Parakeet TDT 0.6B v3 以 597K requests 位列第五。从市场份额来看，openai 以 13.7M requests 占据 82.9%的份额，位居第一；microsoft 以 753K requests 占据 4.6%的份额，位列第二；qwen、nvidia、mistralai 则分别以 696K、645K、367K requests，对应占据 4.2%、3.9%、2.2%的市场份额。

* 统计口径说明：以上数据均来源于 OpenRouter。基于官方统计口径调整，当前 Rankings 采用 Token 与 Request 两套统计维度：其中，仅 Tokens 规模 Leaderboard 仍以模型 Token 消耗量为统计口径，其余类目榜单及市场份额板块则统一采用 Request 调用量为统计口径，两者统计维度不同，不宜直接对应或进行横向比较。

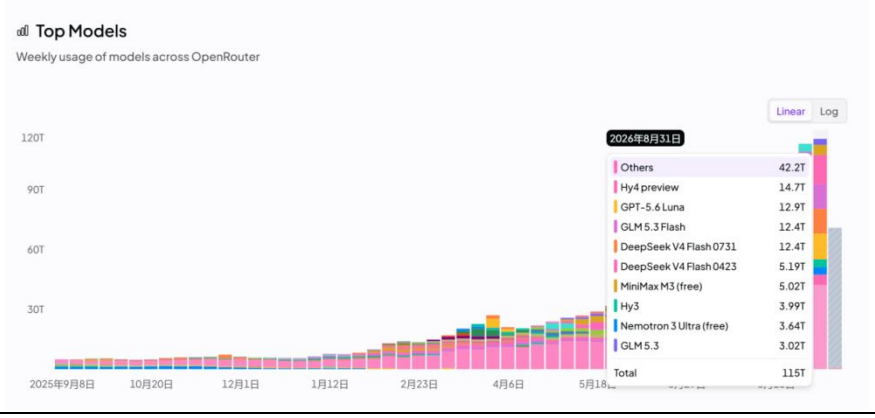
9 月 3 日天猫正式上线 AI 空间站（Token 充值中心），集合阿里云、智谱、Kimi、MiniMax 等主流大模型订阅服务，开创模型 Token 标准化零售新入口，推动 AI 大模型服务进一步融入传统电商消费场景。

9 月 3 日中部（河南）Token 产业运营平台在郑州上线启动，该平台依托天翼云息壤智算底座打造算力、Token 及应用三大超市，汇聚 120 余款主流模型，构建算网协同的省级“词元经济”核心运营服务载体。

9 月 7 日四川首个 Token 工厂“天峰·白泽 Token 工厂”在雅安正式点亮运营，同日网宿科技携手正泰集团等共建海宁城市级绿色 Token 工厂，加速跑通“绿电—算力—Token 量产”产业闭环，推动地方词元规模化生产与交付。

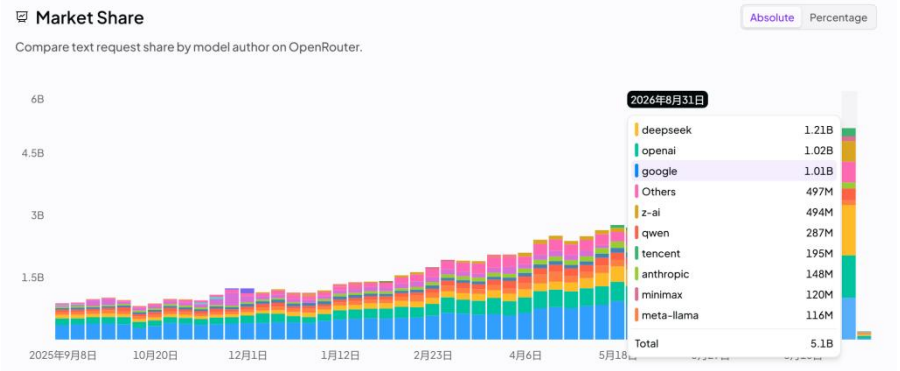
下载日志已记录，仅供内部参考

图表 1：Tokens 规模 Leaderboard（Token 消耗量为统计口径）



资料来源：OpenRouter，华鑫证券研究

图表 2：文本类模型市场份额占据示意图（Request 调用量为统计口径）



资料来源：OpenRouter，华鑫证券研究

1.2、 云厂动态：阿里云企业级 Agent 协作平台“万有无界”开启公测

2026 年 9 月 2 日，阿里云旗下企业级人与 Agent 协作平台“万有无界”正式开启公测，企业用户与个人用户均可通过官方网站注册体验，平台支持调用阿里最新旗舰模型 Qwen3.8-Max。与单一 AI Agent 独立作业不同，“万有无界”定位于多角色 Agent 协作工作台，将复杂业务拆解分配至不同职责的 Agent，通过“项目空间”组织成员与数字资产，依托群聊机制实现多 Agent 接力推进任务，并将作业产出沉淀至企业资产库以形成可复用的数字资产。

核心功能与协作架构方面，平台主要由项目空间、Agent 会话、PMO 调度、资产库及 AI 商城构成。项目空间作为统筹协同的容器，负责整合群聊、成员与资产；会话支持单聊与群聊两种形态；系统内置两款核心主 Agent，其中“小万”作为个人总助理归属于用户，“小有”作为项目经理归属于项目，负责依据 SOP 模板动态规划任务、生成分派方案、推进协作执行与组织复盘；资产库自动留存协作中生成的文件资料；AI 商城则供管理员配置与筛选 AI 资源。在扩展性方面，平台除内置主 Agent 外还保持开放生态，支持用户从技能市场选配组装 Agent，或接入本地部署且符合 A2A 协议的外部 Agent，适配制造、影视、科研、设计与零售等多行业协作场景。

下载日志已记录，仅供内部参考

机制管控与企业级特性方面，“万有无界”赋予 Agent 长期身份、组织记忆、权限边界与协作规范，支持任务指派、考核及明确产出归属；项目决策依据按项目留存并在授权范围内支持跨项目引用，工具与资产的可用范围严格由管理员定义，多人与多 Agent 按节点分工及前后置条件推进，关键节点支持人工验收与纠偏。针对企业资产留存需求，平台推行“沉淀归属企业”架构，将传统沉淀于个人账号的经验与组织记忆统一归属于企业账户，保障人员流动背景下组织能力的连续性。此外，平台提供企业管理后台，支持管理员开展组织架构管理、AI 资源与知识资产管控、订阅账单监控、数据大盘统计及全员智能体审计，实现 AI 资产的可视、可管与可审计。目前，平台已开放企业公测，管理员购买企业订阅套餐后可按席位数量配置成员并管理公共资源。

1.3、产业动态：Anthropic 发布 Claude Fable 5.1 与 Mythos 5.1，8 项基准霸榜且缓存读取降价 75%

2026 年 9 月 2 日，Anthropic 发布新一代旗舰大模型 Claude Fable 5.1 与 Claude Mythos 5.1。Fable 5.1 和 Mythos 5.1 共用同一模型底座，差异在于安全护栏配置，目前已兼容 AWS、谷歌云及微软 Azure，开发者可通过 Claude API（模型名 claude-fable-5-1）进行调用。

在基准测试表现方面，Fable 5.1 在包括主动性科学研究、终端编程、跨学科思维及自主编程在内的全部 8 项基准测试中霸榜第一；在科学研究智能体 Terminal-Bench-Science 测试中，Fable 5.1 得分达到 52.6%，其他三款对比模型得分均在 22%至 29%之间。在软件排障能力上，投资机构 Millennium 历经数年排查未能找出诱因的一处内部系统罕见崩溃故障，被 Fable 5.1 定位到了原因。

图表 3：Fable 5.1 基准测试结果

	Fable 5.1	Fable 5	Opus 5	GPT-5.6 Sol
Agentic scientific research Terminal-Bench-Science 0.1 [1]	**52.6%**	24.7%	29.0%	22.4%
Agentic coding Terminal-Bench 4.0	**55.8%** 60.9% (Mythos 5.1)	42.0%	52.3%	37.3%
Knowledge work GDPval-AA v2	**1853**	1723	1824	1711
Computer use OSWorld 2.0 [2]	**77.9%** partial	72.9% partial	75.4% partial	— partial
Computer use OSWorld 2.0	**41.7%** strict	36.1% strict	39.6% strict	— strict
Multidisciplinary reasoning Humanity's Last Exam	**60.9%** no tools	57.8% no tools	56.6% no tools	— no tools
	65.0% with tools	63.8% with tools	63.6% with tools	— with tools
Business workflows AutomationBench	**31.4%**	17.1%	26.9%	19.6%
Agentic coding CursorBench 3.2.0	**73.4%**	70.5%	70.0%	67.2%

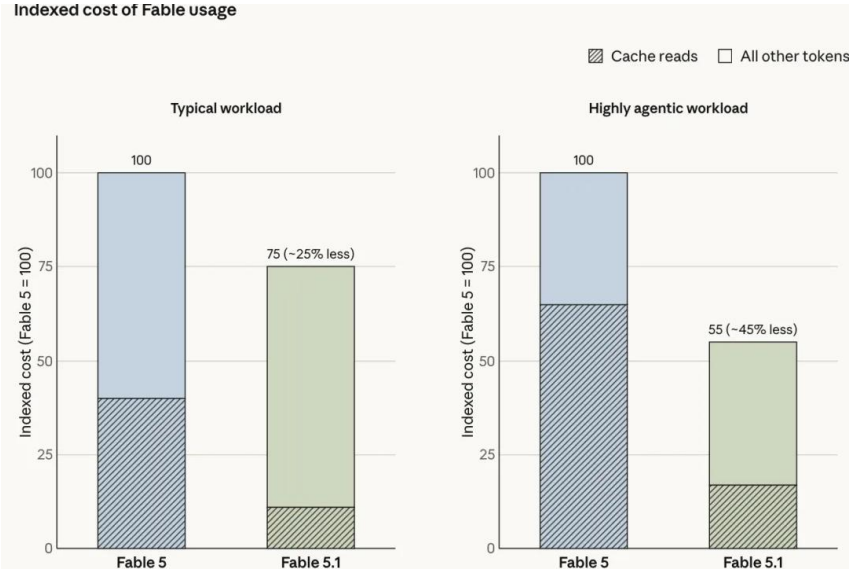
资料来源：Anthropic，智东西，华鑫证券研究

下载日志已记录，仅供内部参考

在专注于科学研究的 Mythos 5.1 场景中，模型接入开源蛋白质设计与折叠工具后，针对 3 个靶点设计的结合剂结合亲和力较 Adaptyv Bio 成果高出 10 倍，12 个靶点测试整体命中率近 50%（高于行业 10%至 15%的常规水平）；依托 NASA 麦哲伦探测器雷达影像及原有覆盖五分之一区域的高程地图，模型训练神经网络生成了覆盖金星三分之一地表的新高程图，分辨率由原先 10-20 公里提升至 2-3 公里，测高精度最高提升 25%；在计算生物学方面，模型通过自动编写定制 GPU 算子并对中间计算结果做缓存处理，使七套开源深度学习模型获得最高 2.5 倍性能提升。在用户实测对比中，Fable 5.1 生成独立 3D 场景时画面细节丰富，任务价格为 GPT-5.6 Sol 的 6 倍；在游戏开发测试中，Fable 5.1 在用户界面和交互机制上优于 Kimi K3，任务成本为 18 美元（Kimi K3 成本为 11 美元）。在 Fable 5.1 发布数小时后，OpenAI CEO 萨姆·奥尔特曼发长文剧透 OpenAI 很快也将推出下一款产品。

在定价与成本方面，Anthropic 称无论使用何种计费方式，Fable 5.1 都将比 Fable 5 低 25%，主要得益于缓存读取资费下降 75%至每百万 Token 0.25 美元（约合人民币 1.68 元），常规业务负载成本降幅约 25%，复杂代码开发与高智能体任务场景下成本最高可节约 45%。除缓存读取资费之外，其余输入 Token 为每百万 10 美元（约合人民币 67.21 元），输出 Token 为每百万 50 美元（约合人民币 336.04 元），与 Fable 5 保持一致。

图表 4：不同任务负载下 Fable 5 与 5.1 调用成本对比



资料来源：智东西，华鑫证券研究

在实际消耗方面，Artificial Analysis 数据显示，由于 Fable 5.1 输出 Token 消耗量约为 Fable 5 的 1.7 倍，单任务实际成本较 Fable 5 高出 20%。在推理档位方面，CursorBench 3.2.0 测试中低推理档位的 Fable 5.1 性能与 Fable 5 高版本相当，但成本仅为后者的三分之一。

安全与合规方面，Mythos 5.1 沿用 Mythos 5 安全管控限制生物研究能力，网络安全场景下新版策略误报数量下降 60%，面向科研人员开放生命科学验证计划（LSVP）与网络安全验证计划（CVP）。Anthropic 还推出全新企业安全防护体系（EFS），提供等同于零数据留存的客户云端存储策略，预计今年秋末分阶段推出；同时按照《欧盟人工智能法案》要求，8 月 2 日之后发布的模型输出结果均添加水印机制。

下载日志已记录，仅供内部参考

2、AI 应用动态：文心一言周时长环比+19.44%，GPT-6 Astra 推出

2.1、周流量跟踪：文心一言周时长环比+19.44%

本期（2026.8.27-2026.9.2）AI 相关网站流量数据：访问量前三位分别为 ChatGPT（1323.0M）、Bing（886.7M）和 Gemini（603.1M），访问量环比增速第一为文心一言（629.32%）；平均停留时长前三位分别为 Character.AI（00:15:56）、Discord（00:10:58）和 NotionAI（00:08:43）;平均停留时长环比增速第一为文心一言（19.44%）。

图表 5：2026.8.27-2026.9.2 AI 相关网站流量

应用	应用类型	归属公司	周平均访问量（M）	访问量环比	平均停留时长	时长环比
ChatGPT	聊天机器人	OpenAI	1323.0	1.30%	6:41	0.50%
Bing	搜索	微软	886.7	4.05%	7:12	-1.14%
Gemini	聊天机器人	谷歌	603.1	-0.76%	6:49	-1.45%
Canva	在线设计	Canva	208.9	2.55%	5:54	-0.28%
Github	代码托管	微软	147.8	-0.27%	6:27	0.26%
Discord	游戏社区	微软	146.0	-0.88%	10:58	-0.30%
Character.AI	聊天机器人	Character.AI	35.01	0.29%	15:56	-0.21%
DeepI	翻译工具	DeepL	22.61	0.18%	2:26	-1.35%
Perplexity	AI 搜索	Perplexity	21.32	-0.84%	4:25	0.00%
QuillBot	释义工具	QuillBot	7.99	5.19%	2:50	0.00%
Kimi	聊天机器人	Moonshot AI	7.72	-12.10%	6:38	-3.63%
NotionAI	文本/笔记	Notion	2.98	6.67%	8:43	2.15%
文心一言	聊天机器人	百度	0.004901	629.32%	1:26	19.44%

资料来源：similarweb,华鑫证券研究

2.2、产业动态：GPT-6 Astra 推出，实现性能与安全性双重进阶

2026 年 9 月 4 日，OpenAI 正式发布了 GPT-6 Astra。公司联合创始人兼首席执行官萨姆·奥尔特曼表示，该模型在计算机应用、专业工作、科学研究、编程及网络安全等多个高要求领域均达到了业界最优水平。在核心能力测评方面，GPT-6 Astra 的表现突出。在 FrontierMath Tier 4 测试中，它取得了 97.6%的高分，并协助解决了数学领域长期存在的开放性问题；而在 ARC-AGI-3 测试中，其得分高达 99.9%。此外，在 ExploitBench 测试中，该模型更是获得了 100%的满分。

下载日志已记录，仅供内部参考

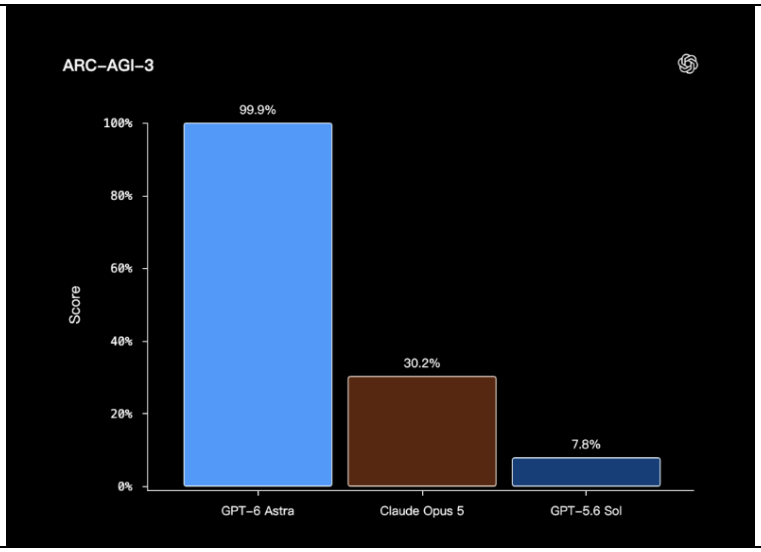
图表 6：GPT-6 Astra 在多项测试中的表现

	OpenAI		
	GPT-6 Astra	GPT-5.6 Sol	Claude Fable 5.1
Terminal-Bench Science 0.1 Scientific workflows	• 64.6%	22.4%	52.6%
AutomationBench Business workflows	• 41.4%	18.1%	31.4%
FrontierMath Tier 4 (v2) Research-level mathematics	• 97.6%	83.0%	87.8%
Terminal-Bench 4.0 Computer terminal tasks	• 57.9%	37.3%	55.8%
HealthBench Professional (length-adjusted) Clinical tasks	• 63.4%	60.5%	58.1%
BenchCAD 3D modeling	• 95.9%	83.3%	84.3%
ARC-AGI-3 Novel interactive puzzles	• 99.9%	7.8%	—

资料来源：智东西，华鑫证券研究

在具体任务执行层面，GPT-6 Astra 展现出显著的性能跃升。在 Terminal-Bench Science 0.1 测试中，该模型需运用代码和终端工具完成完整的科学研究工作流程，包括数据分析、模拟运行及模型拟合，其得分高达 64.6%，明显优于 Claude Fable 5.1 的 52.6%，且预估 API 成本降低约 31%。在 ARC-AGI-3 评估中，GPT-6 Astra 的得分高达 99.9。在 FrontierMath 4 级测试中，Astra 展现了解决极端困难数学问题的高级推理能力；在 Terminal-Bench 4.0 测试中，其得分率达到 57.9%的历史新高，超越了 GPT-5.6 Sol 的 37.3%和 Claude Fable 5.1 的 55.8%，同时每项任务的 API 成本分别降低约 9%和 63%。在模拟跨应用程序多步骤业务流程的 AutomationBench 测试中，Astra 成功完成了 41.4%的任务，高于 Claude Fable 5.1 的 31.4%。更值得关注的是，在评估模型是否会在面对无法完成的任务时擅自超出授权范围的测试中，GPT-5.6 Sol 在无安全措施情况下有 48%的概率越权，而 GPT-6 Astra 的这一发生率为零，体现出其在安全性和判断力上的显著提升。

图表 7：GPT-6 Astra 在 ARC-AGI-3 评估中的表现

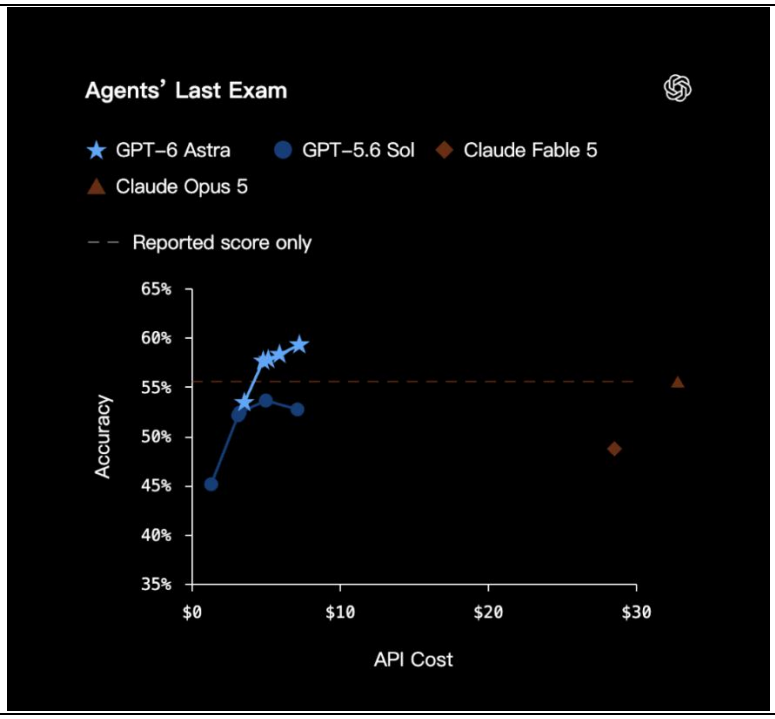


资料来源：智东西，华鑫证券研究

下载日志已记录，仅供内部参考

在计算机使用能力方面，OpenAI 称 Astra 为“全球最佳计算机使用模型”，其在操作速度、精准度和安全性上均实现了质的飞跃。该模型能够自主处理诸多繁琐的日常与专业任务，例如填写在线表格、更新客户关系管理系统记录、进行在线研究并生成摘要、分析科学数据并生成图表、创建网站、进行前端质量检测，乃至协助用户完成软件安装与调试。在 Agents' Last Exam 测试中，Astra 在金融建模、工程和媒体制作等真实专业任务中取得了 59.3%的优异成绩，显著优于 Claude Opus 5 的 55.5%和 GPT-5.6 Sol 的 53.6%，且使用的输出 token 数量较 Opus 5 减少约 65%。在 ScreenSpot-Pro 测试中，其在高分辨率专业软件截图中的界面元素定位准确率高达 92.7%；在 OSWorld 2.0 测试中，其任务得分同样领先于其他可比模型。尤为关键的是，在延迟模拟中，Astra 完成每项任务的耗时约为 40 分钟，较 GPT-5.6 Sol 的 75 分钟大幅缩短约 47%，同时得分更高，这在实际知识工作场景中意味着显著的生产力提升。例如，Astra 能在 15 秒内完成印刷电路板布局的简化回放，将电子原理图转化为可制造的设计；其完成金融建模世界杯挑战的速度约为人类冠军的四倍；它能利用 Unity 引擎从现有资源中快速组建城市场景，并在 FreeCAD 和 Blender 中根据书面简报构建汽车变速器概念模型并制作动画。同时，OpenAI 还更新了 Codex 框架，使任务完成速度在 Mind2Web 基准测试中比 GPT-5.6 Sol 快 1.9 倍，例如可在 2 分 54 秒内完成儿科医生搜索，在 9 分 57 秒内完成公寓查找，显著提升了处理生活事务的效率。

图表 8：GPT-6 Astra 在 Agents' Last Exam 测试中的表现

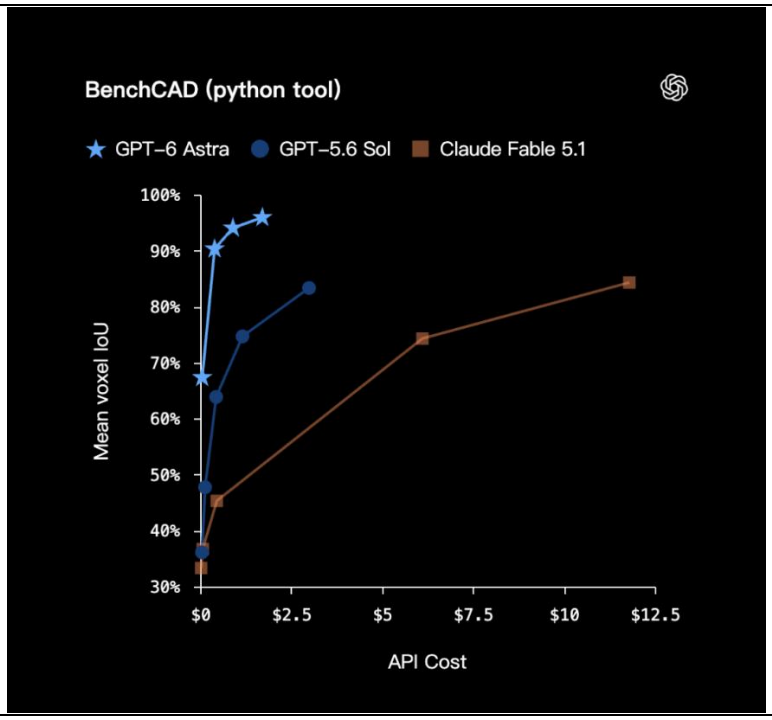


资料来源：智东西，华鑫证券研究

在专业场景与特殊任务中，GPT-6 Astra 同样刷新了多项纪录。在 BenchCAD 测试中，其通过生成 CAD 代码从多视图渲染中重建 3D 对象的几何重叠率达到 95.9%，远超 GPT-5.6 Sol 的 83.3%和 Claude Fable 5.1 的 84.3%，且预估 API 成本较后者降低高达 86%。在 BrowseComp 深网搜索测试中，Astra 取得了 91.5%的高分。在乐谱数字化任务中，它将转录编辑距离从 0.813 大幅降低至 0.159，降幅达 81%，极大减轻了音乐从业者和研究人员的手动转录负担。此外，Astra 能完美遵循模板生成结构清晰的演示文稿，并在 Blender 中建模房屋后转换至 Unreal Engine 5 中生成可步行的三维场景，便于设计沟通。更重要的是，该模型在指令存在歧义时能利用上下文做出合理判断，在关键决策节点主动提出针对性问题，并能在任务演变过程中保持对原始目标的清晰认知，避免偏离整体任务方向。

下载日志已记录，仅供内部参考

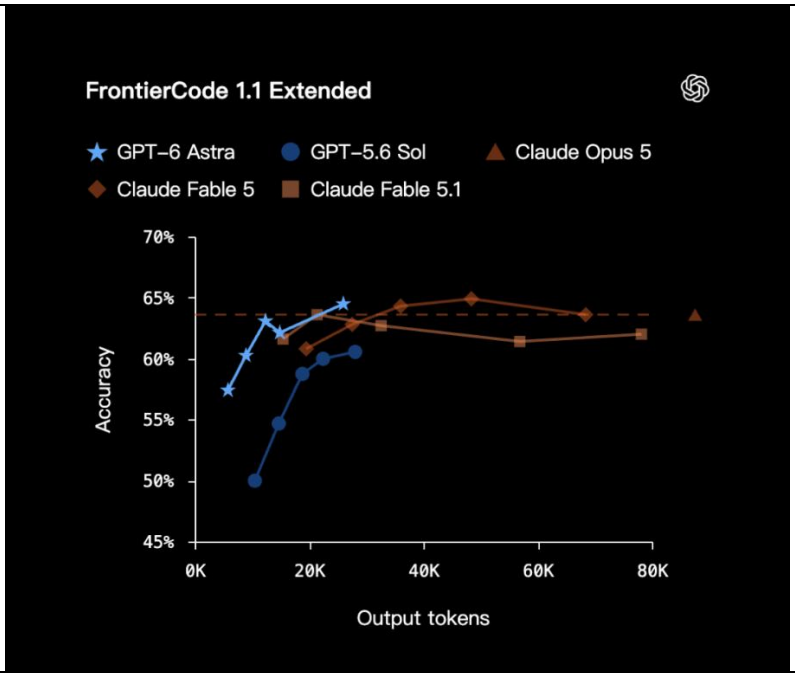
图表 9：GPT-6 Astra 在 BenchCAD 测试中的表现



资料来源：智东西，华鑫证券研究

在软件工程领域，OpenAI 称 GPT-6 Astra 为迄今最佳模型。在 FrontierCode 1.1 Extended 评估中，其编码变更合并至实际代码库的得分几乎与 Claude Fable 5 的最佳水平持平，而预估 API 成本却低约 63%。在 DeepSWE v1.1 真实代码库任务中，其得分为 74.3%，优于 Claude Fable 5.1 的 67.4%，且每任务成本降低约 32%。在数据库迁移等复杂工程任务中，Astra 同样创下新高。此外，Astra 为 Codex 引入了跨上下文窗口保留注释与检索历史信息的新机制，无需反复压缩摘要，从而在长时间工作中保持更完整的上下文信息。

图表 10：GPT-6 Astra 在 FrontierCode 1.1 Extended 评估中的表现

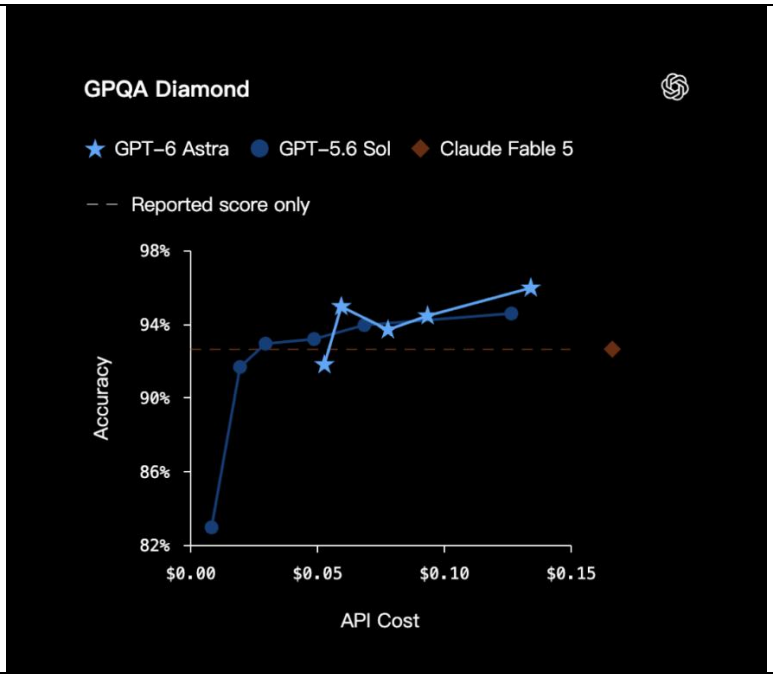


资料来源：智东西，华鑫证券研究

下载日志已记录，仅供内部参考

在科学、数学及健康领域，GPT-6 Astra 同样实现了重要突破。在 GPQA Diamond 测试研究生水平的科学推理时，其得分达到 96.0%，在成本降低约 37%的情况下仍超越 GPT-5.6 Sol 最佳成绩。在 HealthBench Professional 评估中，其长度调整得分最高。在 LifeSciBench 测试中，其性能与 GPT-5.6 Sol 相当，但成本降低约 62%。OpenAI 还宣布，GPT-6 Astra 9 月 4 日起向部分组织开放，未来数日内将向所有 ChatGPT Plus、Pro、Business 及 Enterprise 用户推出，并通过 API 和 AWS 提供服务。在定价策略上，GPT-6 Astra 的收费标准达到 GPT-5.6 Sol 的 2.5 倍，标准版每百万输入 token 定价 10 美元，每百万输出 token 定价 50 美元，缓存输入与写入分别为 1 美元和 12.5 美元。此外，API 中还提供了快速模式选项，处理速度最高可达标准版的 2.5 倍，对应价格则为标准版的 2 倍。

图表 11：GPT-6 Astra 在 GPQA Diamond 测试中的表现



资料来源：智东西，华鑫证券研究

下载日志已记录，仅供内部参考

3、AI 融资动向：Wonderful 完成 5.5 亿美元 C 轮融资，企业级 AI 操作系统加速落地

2026 年 9 月 2 日，总部位于荷兰阿姆斯特丹、创立于以色列特拉维夫的企业级 AI 操作系统研发商 Wonderful 宣布完成 5.5 亿美元（约合人民币 37.0 亿元）C 轮融资，投后估值达 50 亿美元（约合人民币 336 亿元）。本轮融资由 Insight Partners 领投，企业级 SaaS 服务商 Salesforce 作为新战略投资方首次参投，Index Ventures、IVP、Vine Ventures、9Yards 及 Bessemer Venture Partners 等现有股东全额跟投。除本轮融资外，投资方还通过二级市场交易以 1.7 亿美元（约合人民币 11.4 亿元）收购了 Wonderful 员工与早期天使投资人持有的股份，其中领投方 Insight Partners 出资 1.5 亿美元（约合人民币 10.1 亿元）。此前，公司曾于 2026 年 3 月以 20 亿美元（约合人民币 134.4 亿元）估值完成 1.5 亿美元（约合人民币 10.1 亿元）B 轮融资，时隔不到半年其投后估值提升至原先的 2.5 倍。

公司背景与产品架构方面，Wonderful 由首席执行官巴尔·温克勒（Bar Winkler）与首席技术官罗伊·拉拉扎尔（Roey Lalazar）于 2025 年初联合创立。两位联合创始人此前均具备企业级软件与 AI 领域的创业经历，其中温克勒曾创办企业软件采购平台 Approve.com 并被 Tipalti 收购，拉拉扎尔曾创立 AI 视频本地化平台 Kaps 并担任 CTO。Wonderful 初期主要提供面向非英语市场的客服 AI Agent，目前其核心业务已扩展升级为企业级 AI 操作系统 Wonderful AI OS。该平台定位为开放且模型无关的企业 AI 基础设施，兼容企业现有技术栈，支持部署于主流云端环境或本地服务器，客户对其构建的应用与沉淀数据保留完整所有权。平台涵盖托管工作流、生产力 Agent、AI 原生应用及对话式 Agent 四项核心产品模块，各模块共享统一的安全治理与任务编排架构，企业可根据不同业务任务自主选配模型与功能组件。在交付落地方面，Wonderful 推行“前置部署工程师”（Forward-Deployed Engineer，FDE）驻场机制，技术团队进驻客户现场协助首批场景切入生产环境，并通过技术与运营知识转移培训客户团队，以降低从试点到规模化交付的门槛。

商业化拓展与资金用途方面，截至目前 Wonderful 业务已拓展至欧洲、中东、拉丁美洲及亚洲等 35 个以上市场，服务约 100 家大型企业客户，现阶段业务增长主要来自北美以外地区。按当前运营水平测算，Wonderful 年化经常性收入（ARR）约为 7000 万美元（约合人民币 4.70 亿元），较一年前的 100 万美元（约合人民币 672 万元）实现显著增长。公司在册员工规模已从 3 月的 350 人扩展至当前的 650 人（为去年同期的 10 倍以上）。公司计划将本轮资金重点用于加快平台产品研发、扩大前置部署工程师团队规模，并深化在现有海外市场的覆盖；后续还计划推出编程 Agent 及可替代部分传统企业软件的工具，并评估通过并购拓展目标市场份额。同时，Wonderful 预计于 2026 年年底将团队规模扩充至 1000 人以上。

下载日志已记录，仅供内部参考

图表 12：上周 AI 初创公司融资动态

应用	应用类型	领投方	融资轮	融资额	目前累计融资额	目前估值
Crusoe	AI 基础设施服务	Atreides Management & Valor Equity Partners	大额战略融资	30 亿美元	超 55 亿美元	300 亿美元
Wonderful	AI 操作系统	Insight Partners	C 轮	5.5 亿美元	8 亿美元	50 亿美元
Lyte	机器人	Maverick Silicon	C 轮	1.65 亿美元	2.72 亿美元	16 亿美元

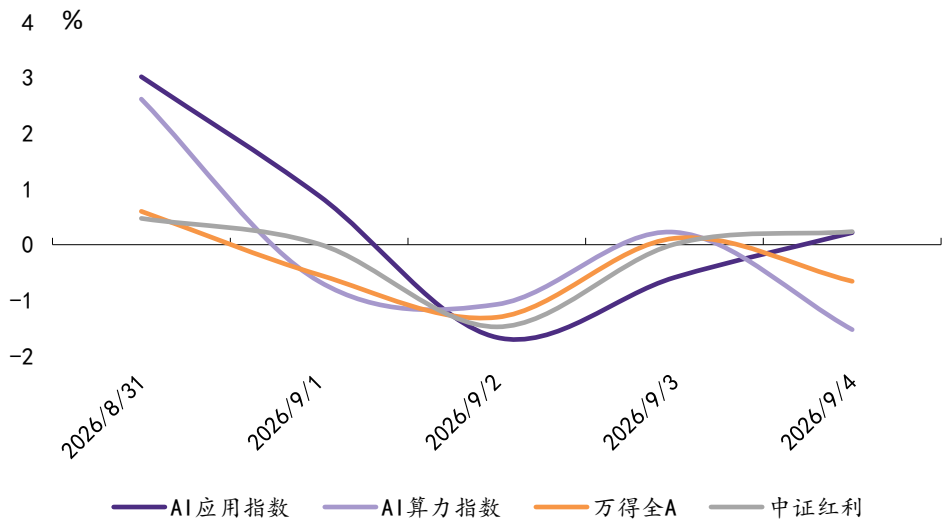
资料来源：智东西，腾讯网，wind，Saasverse，华鑫证券研究

下载日志已记录，仅供内部参考

4、 行情复盘

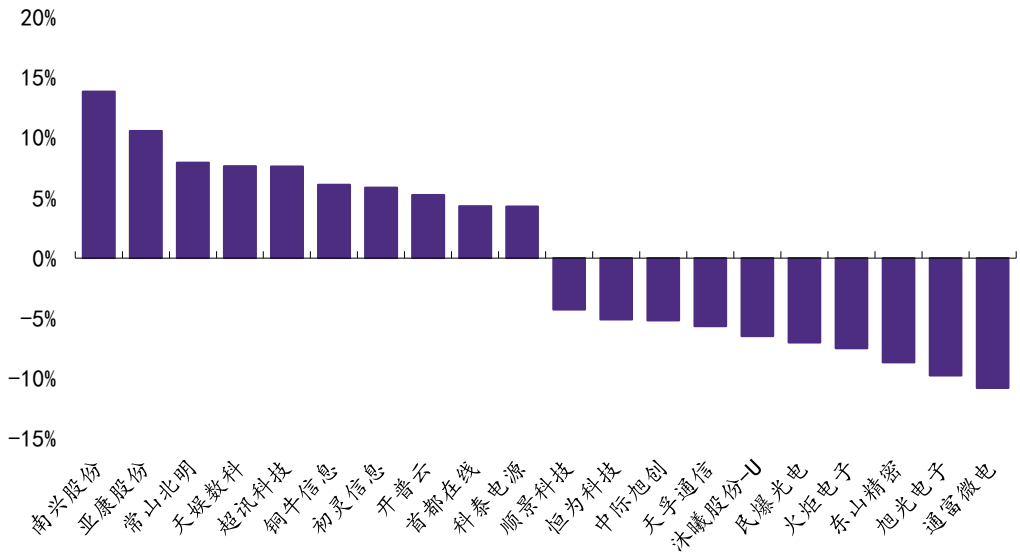
上周（2026.8.31-2026.9.4 日），AI 应用指数/AI 算力指数/万得全 A/中证红利日涨幅最大值分别为 3.02%/2.62%/0.6%/0.47%，AI 应用指数/AI 算力指数/万得全 A/中证红利日跌幅最大值分别为-1.67%/-1.53%/-1.3%/-1.48%。AI 算力指数内部，南兴股份以 13.85%录得上周最大涨幅，通富微电以-10.79%录得上周最大跌幅。AI 应用指数内部，万兴科技以 17.3%录得上周最大涨幅，通富微电以-10.79%录得上周最大跌幅。

图表 13：上周（2026.8.31-2026.9.4 日）指数日涨跌幅



资料来源：wind，华鑫证券研究

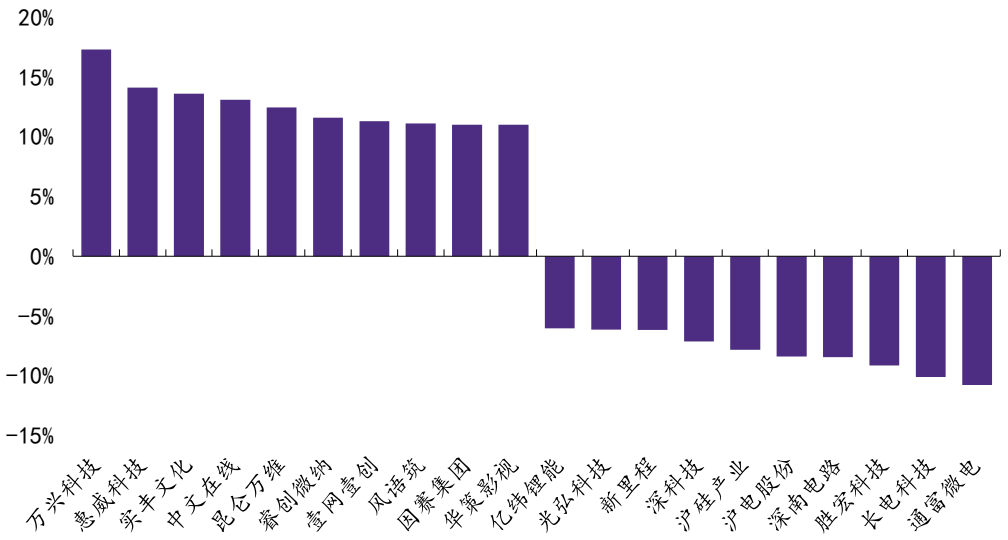
图表 14：上周（2026.8.31-2026.9.4 日）AI 算力指数内部涨跌幅度排名



资料来源：wind，华鑫证券研究

下载日志已记录，仅供内部参考

图表 15：上周（2026.8.31-2026.9.4 日）AI 应用指数内部涨跌幅度排名



资料来源：wind,华鑫证券研究

下载日志已记录，仅供内部参考

5、 投资建议

近日，Snowflake 发布 2027 财年第二季度财报。当季总收入 15.5 亿美元，同比增长 35%；其中产品收入 14.9 亿美元，同比增长 37%，公司 CFO 表示这是产品收入增速连续第三个季度抬升，由核心数据平台与 AI 收入同步走强驱动；剩余履约义务（RPO）90 亿美元，同比增长 30%；Non-GAAP 营业利润率 15%，上年同期为 11%。AI 业务放量明显：CoCo 账户数突破 9,100 家、单季净增超 2,000 家，CoWork 账户数扩至 5,800 家，上半财年超 330 项产品功能正式发布、同比增长 35%。指引方面，公司将 FY2027 全年产品收入指引由 58.40 亿美元上调至 60.70 亿美元（对应同比增速由 31%上调至 36%），Non-GAAP 营业利润率指引由 13.5%上调至 14.5%，3QFY27 产品收入指引 15.88－15.93 亿美元（同比增长 37－38%）。

本次财报的核心信号在于增长加速与利润率扩张同时发生：需求端，产品收入增速连续三个季度抬升至 37%，RPO 同比增长 30%，公司 CFO 明确归因于核心数据平台与 AI 收入同步走强，表明企业客户的 AI 工作负载已进入数据平台的实际消耗阶段，而非停留在试点；盈利端，Non-GAAP 营业利润率同比提升四个百分点至 15%，全年指引收入增速与利润率同时上调，AI 投入期的经营杠杆开始兑现。我们认为，企业 AI 从试点走向生产环境后，数据打通、治理与合规等底层能力成为规模化落地的前提，数据平台的战略价值随之系统性抬升，Snowflake 本季产品收入加速与 AI 产品账户的快速扩张（CoCo 超 9,100 家、CoWork5,800 家）正是这一趋势的直观体现。

中长期，建议关注专注于功率半导体及模拟芯片测试系统研发的联动科技（301369.SZ）、专注于半导体等高端制造业的罗博特科（300757.SZ）、新能源业务高增并供货科尔摩根等全球电机巨头的唯科科技（301196.SZ）、AI 智能文字识别与商业大数据领域巨头的合合信息（688615.SH）、深耕工业 AI 与软件并长期服务高端装备等领域头部客户的能科科技（603859.SH）。

图表 16：ficonTEC2025 年年中至今公告订单

签约日期	客户/描述	业务类型	金额	折合人民币
2025/6/20	美国某头部公司 A 及其子公司	光电子封测设备	约 1,710 万欧元	约 1.36 亿元
2025/7/11	美国某头部公司 B 及其子公司	光电子封测设备	约 1,418 万美元	约 0.98 亿元
2025/9/3	瑞士某头部公司 C 的子公司	全自动硅光子封装整线设备或服务	约 946.50 万欧元	约 0.75 亿元
2025/10/21	武汉驿路通科技股份有限公司	光纤预制及组装线相关自动化设备	约 900 万美元	约 0.62 亿元
2026/1/6	瑞士某头部公司 C 的子公司	第二条全自动 OCS（光交换机）封装整线设备及服务	约 770.00 万欧元	约 0.61 亿元
2025/9/24-2026/1/26	以色列的纳斯达克上市的头部公司 E	单面晶圆测试设备及服务	约 921.60 万美元	约 0.64 亿元

下载日志已记录，仅供内部参考

2026/3/13	暂未披露	双面晶圆测试设备及服务	约 608.09 万欧元	约 0.48 亿元
2026/3/19-2026/3/25	纳斯达克上市的公司 F 及其子公司	耦合设备及服务（可用于可插拔硅光高速光模块封装制程核心环节的量产）	约 6 亿元人民币	约 6 亿元
2026/4/1	纳斯达克上市的公司 F	耦合设备及服务（可用于可插拔硅光高速光模块封装制程核心环节的量产）	约 3,570 万美元	约 2.46 亿元
2026/4/8-2026/5/1	纽约证券交易所上市的公司 B 的子公司	耦合设备及相关服务	约 2680 万美元	约 1.83 亿元
2026/4/8-2026/5/1	纳斯达克上市的公司 F	视觉检测设备、高精度激光 bar 条封装设备及相关服务	约 3226 万美元	约 2.20 亿
2026/7/23	公司 H 的子公司	光学元器件量产的自动化制造设备及服务	约 1674 万欧元	约 1.29 亿
	总金额			约 19.22 亿元

资料来源：Wind，公司公告，华鑫证券研究

图表 17：重点关注公司及盈利预测

公司代码	名称	2026-09-07	EPS			PE			投资评级
		股价	2025	2026E	2027E	2025	2026E	2027E	
300757.SZ	罗博特科	631.94	-0.40	2.13	4.49	-1579.85	296.69	140.74	买入
301196.SZ	唯科科技	88.68	2.05	1.92	2.39	39.28	46.16	37.05	买入
301369.SZ	联动科技	128.08	0.48	1.08	2.62	266.83	118.59	48.89	买入
603859.SH	能科科技	38.20	0.92	1.05	1.32	41.52	36.38	28.94	买入
688615.SH	合合信息	94.88	3.24	4.22	5.25	29.28	22.48	18.07	买入

资料来源：Wind，华鑫证券研究

6、 风险提示

1）AI 底层技术迭代速度不及预期。2）政策监管及版权风险。3）AI 应用落地效果不及预期。4）推荐公司业绩不及预期风险。

下载日志已记录，仅供内部参考

中小盘&北交所组介绍

任春阳：华东师范大学经济学硕士，6 年证券行业经验，2021 年 11 月加盟华鑫证券研究所，从事计算机与中小盘行业上市公司研究

周文龙：澳大利亚莫纳什大学金融硕士

何春玉：金融学士、理学硕士，2023 年 8 月加盟华鑫证券研究所。

证券分析师承诺

本报告署名分析师具有中国证券业协会授予的证券投资咨询执业资格并注册为证券分析师，以勤勉的职业态度，独立、客观地出具本报告。本报告清晰准确地反映了本人的研究观点。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。

证券投资评级说明

股票投资评级说明：

	投资建议	预测个股相对同期证券市场代表性指数涨幅
1	买入	>20%
2	增持	10%—20%
3	中性	-10%—10%
4	卖出	<-10%

行业投资评级说明：

	投资建议	行业指数相对同期证券市场代表性指数涨幅
1	推荐	>10%
2	中性	-10%—10%
3	回避	<-10%

以报告日后的 12 个月内，预测个股或行业指数相对于相关证券市场主要指数的涨跌幅为标准。

相关证券市场代表性指数说明：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；美国市场以道琼斯指数为基准。

下载日志已记录，仅供内部参考

免责条款

华鑫证券有限责任公司（以下简称“华鑫证券”）具有中国证监会核准的证券投资咨询业务资格。本报告由华鑫证券制作，仅供华鑫证券的客户使用。本公司不会因接收人收到本报告而视其为客户。

本报告中的信息均来源于公开资料，华鑫证券研究部门及相关研究人员力求准确可靠，但对这些信息的准确性及完整性不作任何保证。我们已力求报告内容客观、公正，但报告中的信息与所表达的观点不构成所述证券买卖的出价或询价的依据，该等信息、意见并未考虑到获取本报告人员的具体投资目的、财务状况以及特定需求，在任何时候均不构成对任何人的个人推荐。投资者应当对本报告中的信息和意见进行独立评估，并应同时结合各自的投资目的、财务状况和特定需求，必要时就财务、法律、商业、税收等方面咨询专业顾问的意见。对依据或者使用本报告所造成的一切后果，华鑫证券及/或其关联人员均不承担任何法律责任。本公司或关联机构可能会持有报告中所提到的公司所发行的证券头寸并进行交易，还可能为这些公司提供或争取提供投资银行、财务顾问或者金融产品等服务。本公司在知晓范围内依法合规地履行披露。

本报告中的资料、意见、预测均只反映报告初次发布时的判断，可能会随时调整。该等意见、评估及预测无需通知即可随时更改。在不同时期，华鑫证券可能会发出与本报告所载意见、评估及预测不一致的研究报告。华鑫证券没有将此意见及建议向报告所有接收者进行更新的义务。

本报告版权仅为华鑫证券所有，未经华鑫证券书面授权，任何机构和个人不得以任何形式刊载、翻版、复制、发布、转发或引用本报告的任何部分。若华鑫证券以外的机构向其客户发放本报告，则由该机构独自为此发送行为负责，华鑫证券对此等行为不承担任何责任。本报告同时不构成华鑫证券向发送本报告的机构之客户提供的投资建议。如未经华鑫证券授权，私自转载或者转发本报告，所引起的一切后果及法律责任由私自转载或转发者承担。华鑫证券将保留随时追究其法律责任的权利。请投资者慎重使用未经授权刊载或者转发的华鑫证券研究报告。