

电子行业点评报告

国产算力认知强化! Tokens 消耗——AI 需求侧核心逻辑正式向多模态大模型延展

增持（维持）

2025 年 10 月 08 日

证券分析师 陈海进

执业证书: S0600525020001

chenhj@dwzq.com.cn

研究助理 李雅文

执业证书: S0600125020002

liyw@dwzq.com.cn

投资要点

■ 为什么我们认为这一轮多模态大模型的发布具有“划时代”意义？

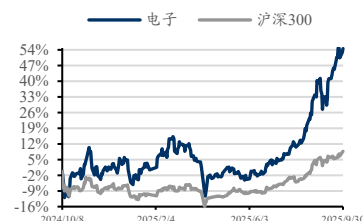
■ 1、海外——Sora 2 打样多模态交互入口，多模态能力突破带动 C 端付费意愿提升。OpenAI 方面，10 月 1 日上线 Sora 2/Pro App，支持最长 15s 的文生视频，物理运动更精准、并支持用户“客串”生成内容，发布三天后登上美国 App Store 第一名。10 月 7 日的开发者大会上宣布 ChatGPT 可直接调用第三方应用，又推出 GPT-5 Pro、Sora 2 API 等开发者工具，标志着 OpenAI 从单一对话工具向 AI 应用及社交平台转型。xAI 方面，10 月 6 日推出视觉生成模块“Imagine”，可从文字生成高质量图像与视频，Grok 由问答助手升级为多模态创作 AI，构建“社交+搜索+创作”一体化平台。Anthropic 方面，9 月 30 日发布 Claude Sonnet 4.5 编程模型，强调其在构建“生产就绪级”AI Agent 上的能力，推动智能体从原型验证迈向实际部署。我们认为，从 Sora 2 的 App 端表现来看，当前多模态理解与生成技术已突破具备商业化价值的临界点。AI 生成的视频中动作可控性与内容观赏度均显著提升，用户的主动创作与付费意愿大幅提升。在此基础上，海外头部厂商正一边持续突破多模态与智能体等模型边界，一边强化商业化路径与 C 端生态布局。AI 正从生产力工具加速演进为内容与交互的基础平台，带动多模态生成、实时交互等高频应用涌现，整体推高模型调用与 Tokens 消耗，为算力与平台层开启新一轮增长周期。

■ 2、国内——多模态能力崭露头角。快手可灵 2.5 Turbo 于 10 月 2 日在全球视频生成模型榜单中登顶图文视频与文生视频双榜首，体现其在视频生成与内容质量上的国际领先水平。字节跳动于 10 月 2 日与 UCLA 推出 Self-Forcing++ 视频生成技术，可生成 4 分 15 秒高清视频，视觉稳定性提升至 2.6 倍；旗下豆包 1.6-Vision 模型于 9 月 30 日更新，通过工具调用将图像融入思维链，提升视觉理解精度。腾讯 9 月 28 日推出并开源混元图像 3.0，作为业界首个开源工业级原生多模态生成模型，快速登顶 Hugging Face 热榜并保持领先。我们认为，当前 AI 应用落地的竞争核心从单一语言智能转向多模态生成与理解能力。国内头部厂商正快速补齐这一关键环节并在国际基准测试中崭露头角。

■ 国产算力投资逻辑：从供给侧向需求侧延展，AI 应用需求成为“国产算力”新引擎。国产算力崛起后的典型标志——适配工作渗透到 AI 大模型研发的更早期阶段。DeepSeek 于 9 月 30 日发布 DeepSeek-V3.2-Exp，DSA 架构使得推理成本更低，支持更长上下文；且模型支持国产开发生态 tilelang 语言，并实现与寒武纪、海光、昇腾等芯片的首日适配，展示了开放协同的算力兼容能力。智谱于 9 月 30 日开源 GLM-4.6，在 Agentic Coding 等关键能力上显著提升，代码生成性能已对齐 Claude Sonnet 4，并完成与国产 GPU 的联合优化，强化了自主算力生态。阿里巴巴于 10 月 4 日发布并开源的 Qwen3-VL 系列多模态模型也实现昇腾芯片的 ODay 适配，加速国产硬件生态落地。我们认为，不同于以往因 H20 限售所驱动的“供给侧替代逻辑”，本轮投资逻辑演进为：多模态能力突破所带来的需求侧拉动——AI 应用自身增长的牵引形成了国产算力的原生成长路径。随着多模态生成与实时推理场景不断丰富，国产算力有望进入内生驱动的新一轮成长周期。

■ 投资建议：云端算力：推荐寒武纪、海光信息、芯原股份、盛科通信、兆易创新，建议关注翱捷科技、裕太微等。端侧算力：推荐晶晨股份、瑞芯微、恒玄科技，建议关注乐鑫科技等。

行业走势



相关研究

《OpenAI 战略重构驱动 AI 基建再定价，A 股算力链景气度跃升》

2025-10-07

《Switch 芯片研究框架（一）：GPU-GPU 互连，从 Scale-Up 到 Scale-Out 的格局如何？》

2025-09-30

■ 风险提示：需求不及预期、技术迭代风险、供应链风险等。

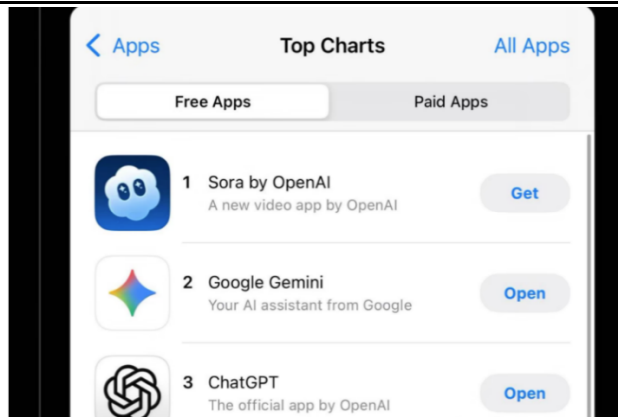
表 1：重点公司估值

代码	公司	总市值 (亿元)	收盘价 (元)	EPS (元/股)			PE			投资评级
				2024A	2025E	2026E	2024A	2025E	2026E	
688256	寒武纪-U	5,543.14	1,325.00	-1.08	4.07	9.62	-1,225.44	325.55	137.73	买入
688041	海光信息	5,871.28	252.60	0.83	1.23	1.85	304.06	205.37	136.54	买入
688521	芯原股份	962.06	183.00	-1.14	-0.19	0.39	-160.11	-963.16	469.23	买入
603986	兆易创新	1,423.30	213.30	1.65	2.48	3.36	129.09	86.01	63.48	买入
688099	晶晨股份	468.18	111.18	1.95	2.52	3.35	56.96	44.12	33.19	买入
603893	瑞芯微	948.95	225.55	1.41	2.53	3.44	159.52	89.15	65.57	买入
688608	恒玄科技	500.89	297.50	2.73	5.14	7.39	108.78	57.88	40.26	买入

数据来源：Wind，东吴证券研究所预测

注：收盘价截至 2025/9/30

图1：Sora 登顶 App Store 美区



数据来源：人工智能产业链，东吴证券研究所

图2：Grok 更新支持文生图应用



数据来源：AI 普瑞斯，东吴证券研究所

图3：可灵 2.5 Turbo 构图光影突破性升级



数据来源：创新创业中关村，东吴证券研究所

图4：DeepSeek 发布 V3.2-Exp



数据来源：算泥，东吴证券研究所

免责声明

东吴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。

本研究报告仅供东吴证券股份有限公司（以下简称“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，本报告中的信息或所表述的意见并不构成对任何人的投资建议，本公司及作者不对任何人因使用本报告中的内容所导致的任何后果负任何责任。任何形式的分享证券投资收益或者分担证券投资损失的书面或口头承诺均为无效。

在法律许可的情况下，东吴证券及其所属关联机构可能会持有报告中提到的公司所发行的证券并进行交易，还可能为这些公司提供投资银行服务或其他服务。

市场有风险，投资需谨慎。本报告是基于本公司分析师认为可靠且已公开的信息，本公司力求但不保证这些信息的准确性和完整性，也不保证文中观点或陈述不会发生任何变更，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。

本报告的版权归本公司所有，未经书面许可，任何机构和个人不得以任何形式翻版、复制和发布。经授权刊载、转发本报告或者摘要的，应当注明出处为东吴证券研究所，并注明本报告发布人和发布日期，提示使用本报告的风险，且不得对本报告进行有悖原意的引用、删节和修改。未经授权或未按要求刊载、转发本报告的，应当承担相应的法律责任。本公司将保留向其追究法律责任的权利。

东吴证券投资评级标准

投资评级基于分析师对报告发布日后 6 至 12 个月内行业或公司回报潜力相对基准表现的预期（A 股市场基准为沪深 300 指数，香港市场基准为恒生指数，美国市场基准为标普 500 指数，新三板基准指数为三板成指（针对协议转让标的）或三板做市指数（针对做市转让标的），北交所基准指数为北证 50 指数），具体如下：

公司投资评级：

买入：预期未来 6 个月个股涨跌幅相对基准在 15%以上；

增持：预期未来 6 个月个股涨跌幅相对基准介于 5%与 15%之间；

中性：预期未来 6 个月个股涨跌幅相对基准介于-5%与 5%之间；

减持：预期未来 6 个月个股涨跌幅相对基准介于-15%与-5%之间；

卖出：预期未来 6 个月个股涨跌幅相对基准在-15%以下。

行业投资评级：

增持：预期未来 6 个月内，行业指数相对强于基准 5%以上；

中性：预期未来 6 个月内，行业指数相对基准-5%与 5%；

减持：预期未来 6 个月内，行业指数相对弱于基准 5%以上。

我们在此提醒您，不同证券研究机构采用不同的评级术语及评级标准。我们采用的是相对评级体系，表示投资的相对比重建议。投资者买入或者卖出证券的决定应当充分考虑自身特定状况，如具体投资目的、财务状况以及特定需求等，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。

东吴证券研究所

苏州工业园区星阳街 5 号

邮政编码：215021

传真：（0512）62938527

公司网址：<http://www.dwzq.com.cn>