

英特尔财报佐证 AI 供不应求, 巨头 AI 应用进展喜人

分析师及联系人

孟灿

SAC: S1130522050001

mengcan@gjzq.com.cn

李忠宇

SAC: S1130524100002

lizhongyu01@gjzq.com.cn

核心要点

产业前沿

欧盟委员会公布《欧盟网络安全法》修订草案，预计存储芯片 2026-2027 年仍将是供不应求状态，2028 年才会改善，工信部印发《关于全面开展算力态势感知自动化监测工作的通知》；OpenAI、Anthropic 25 年年化收入已突破 200、90 亿美元；OpenAI 计划未来几周在 ChatGPT 的“免费版”和“Go 版”中测试广告，26 年下半年推出首款 AI 设备；字节跳动 AI Agent 平台“扣子”宣布 2.0 升级。

资本风向

英特尔 25Q1 财报表现优异，但 26Q1 预期不佳，阿里考虑将平头哥分拆上市；Meta 新成立的人工智能实验室已于本月向内部交付首批重要 AI 模型，过去一年谷歌 Gemini API 调用量增长逾一倍，达 850 亿次，Gemini 企业增长至 800 万订阅用户，千问大模型全球下载量超 10 亿，衍生模型数量突破 20 万；苹果计划在 6 月的发布全新 AI 聊天机器人，并于 9 月正式推出，ServiceNow 计划利用 OpenAI 模型构建语音对话智能体，阿里健康 AI 产品“氢离子”已完成内测并开放下载，钉钉、高德、支付宝联手推出“AI 差旅”。

本周观点

天孚通信等上市公司业绩快报/预告，直观表明光通信、覆铜板、存储等 AI 上游高景气，也佐证 AI 行业的高景气度。平头哥上市有助于增强阿里在 AI 领域的竞争力，持续看好国产 AI 芯片前景。OpenAI、苹果、字节、阿里等巨头在 AI 应用方面动作频繁，大模型落地斜率陡峭，有望形成 AI 投入产出的正循环。

内容目录

一、产业前沿	2
AI 基础层：欧盟修订网络安全法，存储涨价有望继续	2
AI 模型层：OpenAI、Anthropic 最新年化收入达 200 亿、90 亿美元以上 ..	3
AI 应用层：OpenAI、字节等推动广告、AI 硬件、Agent 等应用探索	4
二、资本风向	5
AI 基础层：英特尔财报和预期显示 AI 需求强劲，供给紧张	5
AI 模型层：谷歌、阿里旗下大模型使用高增长	8
AI 应用层：苹果、阿里等巨头推动 AI 聊天机器人、医疗、差旅等应用 ..	9
三、本周观点	10
四、产业链数据更新	10
国内大模型公司上市，收入端快速增长	10
风险提示	11

图表目录

图表 1：英特尔收入及增速（百万美元）	6
图表 2：英特尔毛利率	6
图表 3：智谱 2022-2025H1 收入及净利润（亿元）	11
图表 4：Minimax 2022-2025Q3 收入及净利润（亿元）	11

一、产业前沿

AI 基础层：欧盟修订网络安全法，存储涨价有望继续

欧盟委员会于当地时间 1 月 20 日公布《欧盟网络安全法》修订草案，计划在包括 5G 通信、半导体、电力系统、自动驾驶、医疗设备等 18 个关键领域，逐步淘汰所谓“高风险供应商”的组件和设备。此次修订的核心变化在于，将此前作为自愿性指南的“供应链工具箱”，正式纳入具有法律约束力的强制执行框架，并把涉及的行业范围扩大。修订后的《欧盟网络安全法》仍需在未来几个月内，与成员国政府和欧洲议会展开多轮磋商，最终在三方达成一致后方能生效。草案中有关“高风险供应商”的定义、名单制定机制及实施节奏，仍存在调整空间。

存储芯片在 2026 年一季度大幅涨价（一般型 DRAM 合约价季增 55-60%，NAND Flash 季增 33-38%）之后，二季度涨价有望持续，Counterpoint 预计，Q2 存储芯片价格有望环比上涨 20% 左右。从全球存储芯片需求（AI 需求持续强劲）及大厂扩产的情况来看，TrendForce 预估存储芯片产业 2026 年将达到 5516 亿美元（同比增长 134%），2027 年则将再创高峰达 8427 亿美元，同比增长 53%，预测 2026 年 HBM 需求量增长可能高于 70%，DRAM 三大制造商持续把先进制程产能分配给 HBM 和高阶服务器 DRAM，已导致消费级 DRAM 供给受限，美光 HBM 产量每增加一个单位，传统 DRAM 供给就会减少三个单位；三星正在扩张现有工厂产能，SK 海力士已宣布将大举投资新产能，美光 1000 亿美元在纽约州 Onondaga 兴建超大晶圆厂，1 月 16 日动工，中国存储芯片大厂长存、长鑫也在积极扩产，但是这些新产能投产基本都在 2028 年放量，研判存储芯片 2026-2027 年仍将是供不应求的状态，2028 年才会出现改善。

1月22日消息，工业和信息化部办公厅近日印发《关于全面开展算力态势感知自动化监测工作的通知》，要求到2026年底，实现全国31个省(自治区、直辖市)及重点算力企业算力资源数据的自动化监测，基本建成覆盖全国、标准统一、智能高效的算力态势感知自动化监测体系，监测数据质量、智能分析能力、监测结果应用水平有效提升。下一步，工业和信息化部将加强《通知》的宣贯解读，指导各地区、各企业稳妥推进算力态势感知自动化监测工作，持续推动优化算力资源合理部署，构建高速敏捷算力支撑网络，推进算力资源精细化、智能化、高效化供需匹配，不断提升算力资源配置效率。

AI 模型层:OpenAI、Anthropic 最新年化收入达 200 亿、90 亿美元以上

OpenAI CFO 称公司 2025 年年化收入已突破 200 亿美元。OpenAI 首席财务官莎拉·弗莱尔 (Sarah Friar) 表示，该公司 2025 年年化收入已突破 200 亿美元，较 2024 年的 60 亿美元大幅增长，增长与计算能力的扩张密切相关。该公司上周表示，将开始在 ChatGPT 中向部分美国用户投放广告，加大力度获取收入，以支付开发相关技术的高昂成本。

据媒体援引知情人士称，OpenAI 首席执行官 Sam Altman 近期一直在与中东的顶级投资者会面，为公司新一轮融资铺路，该轮融资规模可能至少达到 500 亿美元，对应的公司估值约为 7500 亿至 8300 亿美元。目前谈判仍处于早期阶段，融资金额仍有可能发生变化。

1月24日消息，据知情人士透露，人工智能公司 Anthropic 的收入规模在过去半年内显著增长，其年化收入已从 2025 年夏季的约 40 亿美元提升至 2025 年底的 90 亿美元以上。与此同时，Anthropic 正在推进新一轮大规模融资。多名知情人士称，该公司正与 Coatue Management、新加坡政府投资公司 GIC 以及

Iconiq Capital 等机构洽谈融资事宜，单笔投资规模预计均在 10 亿美元以上。本轮融资需求原计划约为 100 亿美元，但目前已出现超额认购迹象，整体融资规模或超过最初目标。

月之暗面开启新一轮融资，多模态大模型近期有望更新。大模型独角兽月之暗面（Kimi）正在进行一轮全新融资。本轮融资投前估值接近 48 亿美元，而不到一个月前，月之暗面完成 5 亿美元 C 轮融资，投后估值 43 亿美元。据了解，月之暗面预计将在近期发布较为重要的模型技术更新，或为外界关注的 K2-VL 多模态大模型。

AI 应用层：OpenAI、字节等推动广告、AI 硬件、Agent 等应用探索

本周，海外聊天助手类应用活跃度中，多数应用活跃度持续恢复，其中 Gemini 访问量续创新高。国内多数聊天类应用活跃度保持稳定。

当地时间 1 月 16 日，OpenAI 表示，计划未来几周在 ChatGPT 的“免费版”和“Go 版”中测试广告。同日，该公司宣布将把月费 8 美元的“Go 版”从美国市场推广至所有 ChatGPT 覆盖地区。OpenAI 承诺，将始终把用户信任和透明度放在首位，ChatGPT 的回复不会受到广告影响——广告将独立显示并附带有清晰的标签。除此之外，OpenAI 声称，Plus、Pro 等付费较高的 ChatGPT 服务不会收到广告，包含健康、心理健康或政治等敏感话题的聊天也将不适用于广告。

据报道，OpenAI 将在 2026 年下半年推出其首款人工智能设备。这一消息由 OpenAI 全球事务主管克里斯·莱恩透露，他将“设备”列为公司 2026 年值得期待的重大动向之一。此前多方爆料显示，这款首款由 ChatGPT 驱动的设备可能具备以下特点：口袋大小、能感知用户周围环境与生活情境，并且完全无屏幕。据

报道，OpenAI 正考虑在未来几年推出多款 AI 产品，可能包括一支笔和一款“家居式设备”。因此，本次爆料的产品可能与 OpenAI 首发的设备并非同一款。

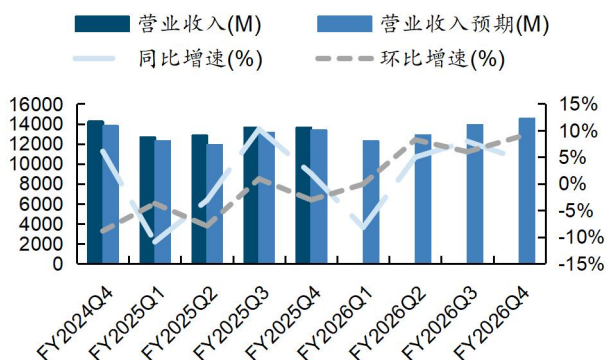
IT 之家 1 月 19 日消息，字节跳动旗下 AI Agent 平台“扣子”宣布 2.0 品牌升级，集成了 Agent Skill、Agent Plan、Agent Coding、Agent Office 能力。通过封装领域知识、标准化操作流程或集成专用工具，扣子能够将通用 AI 的认知能力与特定任务需求相结合，更贴合实际应用场景中多元化、高标准的任务要求，保证输出稳定性；扣子可根据任务场景，智能调用或组合多个已安装技能，让 Agent Skills 提供综合性解决方案；用户还可利用扣子提供的工具，为自己量身定制私有技能，还可以将个人经验沉淀封装为可复用的模块。

二、资本风向

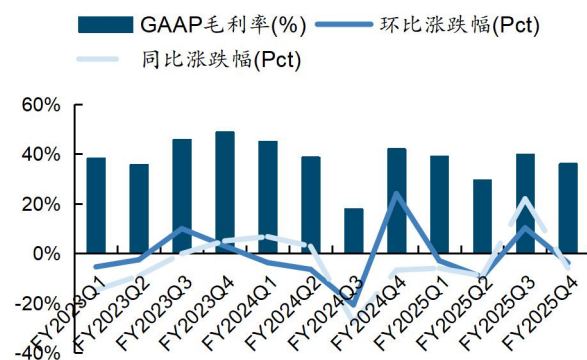
AI 基础层：英特尔财报和预期显示 AI 需求强劲，供给紧张

英特尔本季财报表现优异，但 26 年第一季度预期不佳。英特尔 2025 年第四季度整体呈现出“需求强、供给短”背景下的超指引交付：公司实现营收 137 亿美元、非 GAAP 毛利率为 37.9%、非 GAAP EPS 0.15 美元，同时第四季度经营性现金流 43 亿美元、调整后自由现金流转正至 22 亿美元，反映出在费用纪律与现金管理层面的阶段性改善；但管理层也明确指出，行业与公司内部的供应约束显著限制了对强劲市场需求的捕捉，且这一约束在 2026 年第一季度最为严峻，因此给出 2026Q1 营收 117 - 127 亿美元、非 GAAP 毛利率约 34.5%、EPS 接近盈亏平衡的指引。毛利率阶段性下探的核心原因在于收入回落叠加固定成本结构，以及 Panther Lake 在产品结构中占比提升但成本端仍具稀释效应。

图表 1: 英特尔收入及增速(百万美元) 图表 2: 英特尔毛利率



来源: 路透、国金证券研究所



来源: 路透、国金证券研究所

展望 2026 年, 公司判断自第二季度起可用供给将逐季改善, 主要依靠良率与产出效率提升来释放供给弹性, 并对服务器业务在 AI 推理与数据中心升级驱动下的全年增长前景持更为积极的看法; 相对而言, 代工业务的兑现节奏更偏中长期, 公司对 14A 的产能投入保持高度纪律, 强调在获得明确客户承诺前不会大规模扩产, 并将风险量产与规模量产分别放在 2027 年后段与 2028 年。

总的来说, 库存紧缺是当前 AI 服务器出货节奏的一个实质性制约因素, 其影响会随着供给恢复程度的推进而逐步减弱, 但在短期内已成为影响英特尔服务器业务和市场节奏的重要变量。我们认为, 未来几个季度的关键验证点集中在: Q2 以后供给改善能否如期兑现、18A/Series 3 的成本曲线能否推动毛利率从 Q1 低位逐步修复至更可接受区间, 以及在 DRAM、NAND、基板等关键组件持续紧张并可能涨价的环境下, 客户端需求与盈利能力是否面临进一步压制。

1 月 21 日盘后, 天孚通信发布 2025 年度业绩预告, 公司预计 2025 年实现归属于上市公司股东的净利润约 18.81 亿元-21.5 亿元, 同比增长 40.00%-60.00%; 预计 2025 年实现扣非净利润约 18.29 亿元-21.08 亿元, 同比增长 39.19%-60.4%, 但逊于 Wind 一致预期。天孚通信指出, 人工智能行业的加速发展以及全球范围

内数据中心建设的推进，带来了高速光器件产品需求的持续稳定增长。这一趋势有效抵消了部分宏观因素的不利影响，不仅带动了营收规模扩大，也验证了行业的高景气度。

1月21日晚间，腾景科技公告称，公司收到某客户C子公司的采购订单，总金额为1280万美元（约为8915万元人民币），产品为二维准直器阵列。据介绍，本次订单供货产品二维准直器阵列为公司在光通信领域高端光器件的重点产品之一，可满足下游OCS光交换机厂商对于不同端口数量规格的准直器阵列产品的定制化需求，为AI算力中心及超大规模光互联网络建设提供关键元器件支持。

剑桥科技1月17日发布业绩预增公告，预计2025年度实现归属于上市公司股东的净利润2.52亿元至2.78亿元，较上年同期增加8,531.88万元至1.11亿元，同比增长51.19%至66.79%。公告显示，公司预计2025年度实现归属于上市公司股东的扣除非经常性损益后的净利润为2.49亿元至2.75亿元，与上年同期相比增加9,774.57万元至1.24亿元，同比增幅达64.62%至81.81%。公告指出，公司三大核心业务——高速光模块、宽带接入及无线接入业务的协同发展是本期业绩持续增长的直接驱动力。

金安国纪发布2025年业绩预告，预计2025年业绩2.8-3.6亿元，同比增长655-871%，如果按照业绩指引中值测算，2025Q4业绩环比2025Q3增长43.4%，公司表示，业绩大幅增长的主要原因是覆铜板市场需求有所好转，公司覆铜板产销量同比增长、销售价格有所回升。

建滔积层分别在2025年8月、12月涨价三次，其他公司也纷纷涨价，生益科技在2026年1月5日与东莞松山湖相关部门签署了45亿元的投资意向书，旨在加大对高性能覆铜板的投入。

1月22日，彭博社援引知情人士消息称，阿里巴巴集团正考虑将其芯片设计子公司平头哥（T-Head）分拆为独立公司并推动上市。根据相关计划，阿里拟先将平头哥重组为一家员工部分持股的独立实体，后续再视市场情况启动首次公开募股（IPO），具体时间表尚未最终确定。

AI 模型层：谷歌、阿里旗下大模型使用高增长

1月21日，据路透社报道，Meta Platforms 首席技术官安德鲁·博斯沃思（Andrew Bosworth）表示，公司新成立的人工智能实验室已于本月向内部交付首批重要 AI 模型。他在瑞士达沃斯世界经济论坛年会期间的媒体简报会上称，这些模型由 Meta “超级智能实验室” 团队开发，“表现非常好” 且展现出较大的发展潜力。这是 Meta 首次公开确认该新 AI 团队已完成具有代表性的模型交付，但博斯沃思并未具体说明交付的模型类型或应用场景。

1月20日财联社消息，据了解 Gemini 销售的人士透露，过去一年，谷歌销售 Gemini AI 模型的业务激增。GeminiAPI 调用量已增长逾一倍，达到 850 亿次；Gemini 企业已增长至 800 万订阅用户。知情人士表示，这有望提升 Google Cloud 服务器销售这一核心业务的收入，因为客户用于 AI 领域的资金往往会导致他们在其他 Google 产品上产生额外支出。

阿里千问大模型全球下载量超 10 亿，衍生模型数量突破 20 万。1月21日，阿里千问系列模型下载量突破 10 亿次，平均每天被下载 110 万次。同时，全球最大 AI 开源社区 Hugging Face 最新数据显示，阿里千问衍生模型数突破 20 万个，成为全球首个达成此目标的开源大模型。

阿里通义千问开源 Qwen3-TTS 语音生成模型家族，采用 Dual-Track 双轨架构实现 97ms 极致低延迟，支持 3 秒音色克隆、10 种语言/方言迁移及 Voice Design

指令设计功能，通过 1.7B 与 0.6B 双版本兼顾高性能表现与边缘侧轻量化部署效率。

AI 应用层：苹果、阿里等巨头推动 AI 聊天机器人、医疗、差旅等应用

1 月 21 日，据彭博社援引知情人士透露，苹果计划在 6 月的全球开发者大会上发布代号为 Campos 的全新 AI 聊天机器人，并于 9 月正式推出。新版 AI 聊天机器人将取代现有界面，用户仍可以通过语音指令“Siri”或按住设备侧键来召唤服务，但功能将远超当前版本，具备类似 ChatGPT 或谷歌 Gemini 的对话式交互能力。据报道，全新 AI 聊天机器人将成为 iOS 27、iPadOS 27（代号 Rave）和 macOS 27（代号 Fizz）的主要新增功能；新版聊天机器人将采用苹果设计的用户界面，但依赖谷歌 Gemini 团队开发的定制 AI 模型。

ServiceNow 在周二宣布，已与 OpenAI 达成一项多年合作协议，将为其客户提供前沿模型能力的直接访问权限。利用 OpenAI 的先进模型，ServiceNow 计划构建语音对话智能体，让用户能够使用其首选语言进行交流，实时触发诸如创建工单或执行审批等操作。ServiceNow 中央产品管理高级副总裁 John Aisien 在邮件中表示，尽管 OpenAI 模型已经对 ServiceNow 客户开放，但两家公司将在“接下来几个月”内合作解决特定用例，并根据企业需求定制体验。

阿里健康旗下 AI 产品“氢离子”已完成内测并已开放下载。阿里健康旗下 AI 产品“氢离子”已完成内测并已开放下载。该产品面向临床、科研领域的医生群体，定位是医生版“GPT”，这也是阿里健康自研医学大模型的首个落地应用。

钉钉、高德、支付宝联手推出“AI 差旅”。1 月 20 日，钉钉更新 8.2.5 版本，由钉钉、高德、支付宝合作的“AI 差旅”产品正式上线，所有企业用户在最新的钉钉内搜索“AI 差旅”、“差旅用车”、“机票”、“酒店”等相关关键词，

即可零门槛、免费开通这一服务，无需垫资或支付服务费。在该版本中，AI 印、AI 听记同声传译等能力也全量上线。

三、本周观点

天孚通信等上市公司业绩快报/预告，直观表明光通信、覆铜板、存储等 AI 上游高景气，也佐证 AI 行业的高景气度。阿里正考虑将其芯片设计子公司平头哥分拆为独立公司并推动上市，但具体时间表尚未最终确定。我们认为平头哥上市有助于增强阿里在 AI 领域的竞争力，持续看好国产 AI 芯片前景。OpenAI 在硬件、AI 应用及融资方面动作频繁，该轮融资规模可能至少达到 500 亿美元，对应的公司估值约为 7500 亿至 8300 亿美元。Anthropic 的年化收入已从 2025 年夏季的约 40 亿美元提升至 2025 年底的 90 亿美元以上，大模型应用落地斜率陡峭。OpenAI 表示，计划未来几周在 ChatGPT 的“免费版”和“Go 版”中测试广告，将在 2026 年下半年推出其首款人工智能设备。苹果计划在 6 月的全球开发者大会上发布代号为 Campos 的全新 AI 聊天机器人，并于 9 月正式推出。字节跳动旗下 AI Agent 平台“扣子”宣布升级，阿里健康旗下 AI 产品“氢离子”已完成内测并已开放下载，钉钉、高德、支付宝联手推出“AI 差旅”，巨头 AI 应用落地积极，有望形成 AI 投入产出的正循环。

四、产业链数据更新

国内大模型公司上市，收入端快速增长

2022-2025H1，智谱营收分别为 0.57、1.25、3.12、1.91 亿元，归母净利润分别为-1.43、-7.88、-29.56、-23.51 亿元。2022-2025Q3，MiniMax 营收分别

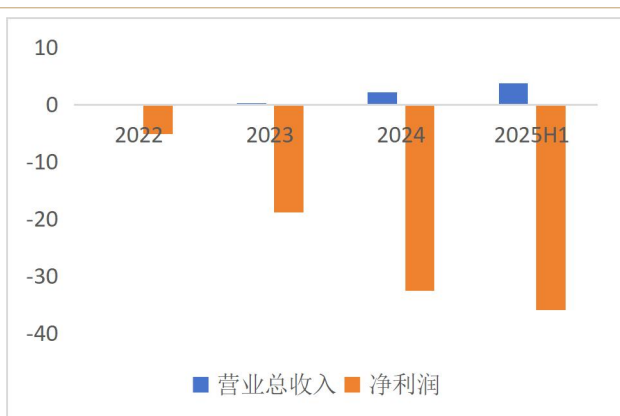
为 0、0.25、2.19、3.8 亿元，归母净利润分别为-5.13、-19.07、-33.44、-36.38 亿元。

图表 3: 智谱 2022-2025H1 收入及净利润 (亿元)



来源: Wind、国金证券研究所

图表 4: Minimax 2022-2025Q3 收入及净利润 (亿元)



来源: Wind、国金证券研究所

风险提示

算力供不应求的风险: 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。

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,使其难以在诸多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。

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。

并购整合不及预期的风险：部分公司进行并购扩大业务布局，如并购整合不能顺利完成，可能影响公司整体业绩表现。

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。

国金证券股份有限公司) 所有，未经事先书面授权，任何机构和个人均不得以任何方式对本报告的任何部分制作任何形式的复制、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。