

全球AI周报

世界人工智能大会（WAIC）正式召开，谷歌财报上调全年资本支出指引至850亿美元

作者：

分析师 孔蓉 SAC执业证书编号：S1110521020002

分析师 李泽宇 SAC执业证书编号：S1110520110002

分析师 樊程安吉 SAC执业证书编号：S1110524080001



■ 海外科技公司财报AI进展

- **Google：业绩全面超预期，AI商业化落地加速。**谷歌FY25Q2营收达 964 亿美元，同比增长 14%，超越彭博一致预期（940亿美元）；每股收益为 2.31 美元，同比增长 22%，超越彭博一致预期（2.18美元）。Google Cloud 收入 136 亿美元，同比大增 32%，积压订单达 1060 亿美元（+38%）。AI 商业化加速，平台月处理 token 超 980 万亿（较 5 月翻倍），Gemini 月活超 4.5 亿。Q2 资本支出 224 亿美元，全年指引上调至 850 亿美元。
- **Tesla：毛利率超预期，Optimus 第三代原型机年底公布。**FY25Q2 收入达 225 亿美元，同比下滑 11.78%，低于彭博一致预期（226 亿美元）；毛利率达 17.2%，超彭博一致预期预期（16.5%）。机器人领域，Optimus 3 原型机可能将于年底公布，计划明年量产，目标尽快实现年产 100 万台。Q2 资本支出 24 亿美元，全年预计超 90 亿。
- **ServiceNow：收入与RPO全面超预期，上调全年指引。**FY25Q2 总营业收入达 321.15 亿美元，同比增长 22.4%，超彭博一致预期（31.2 亿美元）；当前剩余履约义务（cRPO）达到 109.2 亿美元，同比增长 24.5%，超彭博一致预期（104.9 亿美元）。公司将 2025 财年全年的订阅收入指引中值上调 1.25 亿美元，达到 127.75 亿至 127.95 亿美元。**公司表示AI产品Now Assist 本季度收入超预期，目标2026年ACV达到10亿美元。**

■ 全球AI动态：

- **谷歌：**7月21日，谷歌 DeepMind 正式官宣其新模型 Gemini Deep Think 在国际数学奥林匹克竞赛（IMO）取得金牌。该模型在 4.5 小时内攻克了 6 道题中的 5 道，且仅使用自然语言（英语）完成解题。此外，该成果得到 IMO 组委会官方认证，而同期 OpenAI 的相关成绩未经过官方验证。团队使用了 Gemini Deep Think 的增强推理模式，采用了“训练 + 知识库 + 策略”的组合，使其能够直接从官方问题描述中生成严谨的数学证明，无需依赖专业“形式语言”。
- **腾讯：**7月22日，腾讯云推出全球首位贯通“产品-设计-研发”的一体化AI全栈工程师——CodeBuddy AI IDE。该工具支持通过自然语言对话输入需求，自动完成从产品构想到设计原型、前后端编码、上线部署的全流程，实现“对话即编程”。此外，CodeBuddy AI IDE 具备软件开发智能体、实时代码续写、行间/终端内联对话解决问题等能力。实战案例显示，在电商活动页开发中，该工具将传统 2 天的工作量压缩至不到 2 小时，效率提升超 10 倍。
- **阿里巴巴：本周开源三款重量级模型。**7月23日，阿里云开源 Qwen3-Coder 代码大模型，其中最强大的版本是 Qwen3-Coder-480B-A35B-Instruct，在 Agentic Coding、Agentic Browser-Use 和 Agentic Tool-Use 上取得了开源模型的 SOTA 效果，可以与 Claude Sonnet4 媲美。7月25日，阿里正式发布全球最强开源推理模型——Qwen3-72B-A22B-Thinking-2507。基准测试结果显示，这款推理模型的表现与顶级闭源模型（如 Gemini-2.5 pro、O4-mini）不相上下。海外知名模型 API 聚合平台 OpenRouter 数据显示，阿里 Qwen API 已突破 1000 亿 Tokens，在 OpenRouter 趋势榜上包揽全球前三。
- **字节跳动：**7月24日，火山引擎正式发布豆包·同声传译模型 Seed LiveInterpret 2.0，是首个延迟与准确率接近人类水平的产品级中英语音同传系统。其核心突破在于，在中英同传翻译质量达到业界 SOTA 的同时，实现了极低的语音延迟（2-3 秒）和“0 样本声音复刻”，可基于端到端理解和生成框架，真正实现“边听边说”。该模型已在多个权威国际会议中应用，并即将在 8 月上线 Ola Friend 耳机，为用户带来更加灵活的使用方式和体验。

■ 投资建议：

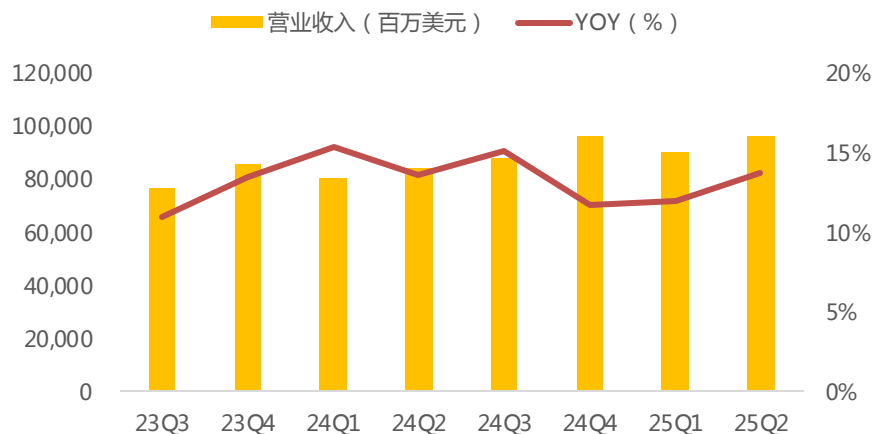
- **海外AI：**1）算力层面，谷歌大幅上调2025年资本开支计划，AI基础设施高景气持续兑现。谷歌在财报中将全年资本支出从此前预期的750亿美元上调至850亿美元，反映其对AI需求持续增长的强信心。同时管理层披露，token用量较两个月前已实现翻倍增长，进一步验证推理需求快速爬坡背景下的资源紧张。我们判断，全球AI算力资本开支周期仍处于快速扩张阶段，建议持续关注GPU与ASIC加速芯片、AI服务器及数据中心电力基础设施的中长期投资机会。建议关注：【英伟达】、【Marvell】、【博通】、【VERTIV】；同时建议关注云厂商AI云业务加速释放下的超预期弹性机会【谷歌】【微软】【Oracle】；2）应用端：OpenAI 预计八月将正式发布下一代大模型 GPT-5，模型有望集成 o 系列推理能力，定位为通往 AGI 的关键一步。我们认为，大模型产品路径日渐清晰、推理成本持续下降，将进一步提升 AI 在教育、广告、电商、企业 SaaS 等场景的 ROI，加快企业客户的部署节奏。从近期多家 AI 应用公司财报数据来看，用户使用时长、活跃度与付费意愿仍在抬升。建议关注：【Cloudflare】、【Snowflake】、【Salesforce】、【Duolingo】、【Roblox】、【CyberArk】、【Applovin】等。3）智能驾驶：特斯拉 FSD 规模化落地节奏稳步推进，作为当前 AI 最具潜力的应用场景之一，我们看好【特斯拉】在 FSD 和 Robotaxi 领域的持续进展。
- **中国AI：**本周我们观察到中国 AI 板块持续迎来利好催化。一方面，2025 世界人工智能大会（WAIC）正式拉开帷幕。作为历届规模最为盛大的一届盛会，本次大会吸引了 800 余家企业参与，集中呈现了 40 余款大模型、50 余款人工智能终端产品、60 余款智能机器人。我们认为，中国 AI 企业参与 WAIC 程度逐年提升，也标志着中国 AI 从基础能力到终端应用的全链条加速落地，板块价值重估逻辑持续演绎。另一方面，国产大模型持续迭代。此前 Kimi K2 模型正式发布并开源，基于 MoE 架构，总参规模达 1T，在代码能力与通用 Agent 任务中取得 SOTA 表现；本周，阿里正式更新其推理模型 Qwen3-72B-A22B-Thinking-2507，海外知名模型 API 聚合平台 OpenRouter 数据显示，阿里 Qwen API 已突破 1000 亿 Tokens，在 OpenRouter 趋势榜上包揽全球前三。我们认为，Kimi 与千问等模型的持续进化，将加快 To C 及 To B 端智能体的商用落地，带动 AI 商业化路径从工具向助手演进，释放更大场景价值。叠加前期算力瓶颈的阶段性缓解，AI 产业链正进入新一轮训练与部署效率提升周期。我们建议重点关注两类标的：（1）平台型互联网公司，具备算力资源、模型能力与应用场景协同优势，建议关注【阿里巴巴】【腾讯】【百度】【小米】（与汽车、电子组联合覆盖）【快手】【美团】（与商社组联合覆盖）；（2）具备模型或应用能力的 AI 生态企业，建议关注【第四范式】【商汤】【美图公司】（与计算机组联合覆盖）【京东健康】（与医药组联合覆盖）【阿里健康】【涂鸦智能】【金山软件】【粉笔】（商社、轻纺教育组联合覆盖）【知乎】【KEEP】【汇量科技】等。
- **风险提示：**宏观经济不及预期；企业盈利不及预期；通胀上行压力。

重点公司业绩情况

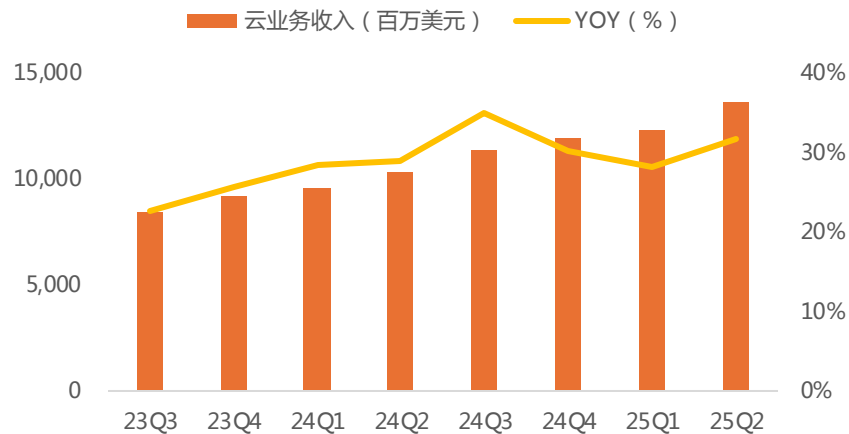
1.1 Google：业绩全面超预期，AI驱动云与搜索强劲增长

- **财务数据：**谷歌FY25Q2营收达 964 亿美元，同比增长 14%，超越彭博一致预期（940亿美元）；净利润 282 亿美元，同比增长19%，超越彭博一致预期（268亿美元）；每股收益为 2.31 美元，同比增长 22%，超越彭博一致预期（2.18美元）。
- **分业务：**谷歌服务收入增长12%至825亿美元，其中搜索及其他业务收入增长12%至542亿美元，YouTube广告收入增长13%至 98 亿美元，订阅、平台和设备收入增长20%至 112 亿美元。Google Cloud收入达到 136 亿美元，同比大幅增长32%，主要得益于谷歌云平台（GCP）在核心产品和人工智能产品方面的增长。其他投资收入为3.73亿美元。
- **云业务加速增长：**谷歌云年化收入运行率已突破500亿美元，运营利润率从去年同期的11.3%提升至20.7%。利润率扩张得益于强劲的收入表现以及持续的费用效率提升，但部分被更高的技术基础设施使用成本所抵消。云业务积压订单（backlog）增至1060 亿美元，同比增长38%。

图：谷歌营业收入及增速



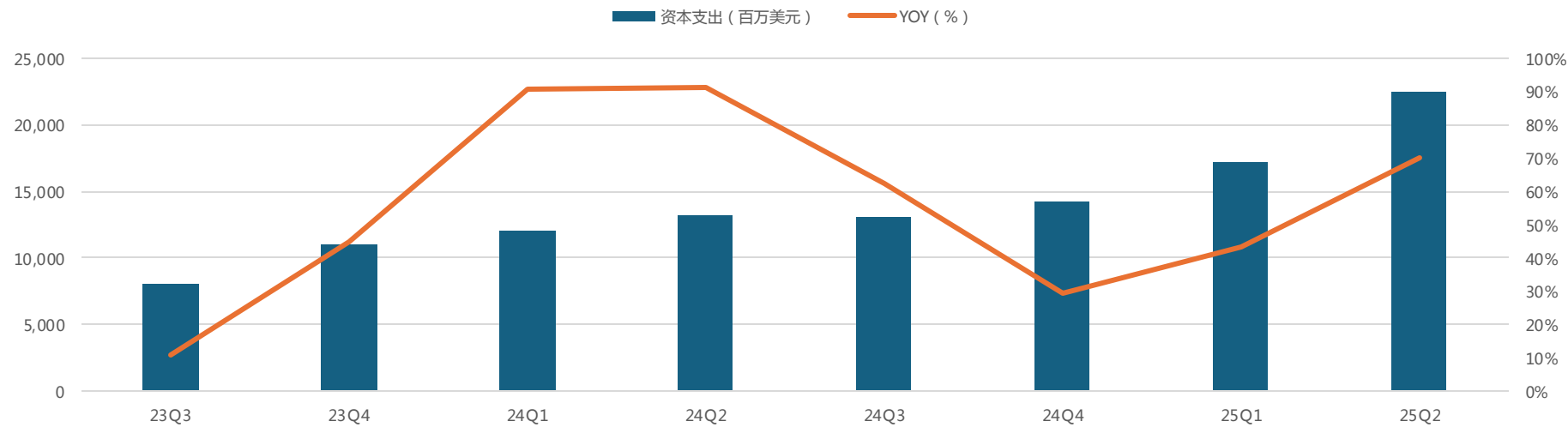
图：谷歌云业务收入及增速



1.1 Google：业绩全面超预期，AI驱动云与搜索强劲增长

- AI商业化加速：Google已在Wrokspace、Chrome等平台融入了AI技术，带来了极大的使用量增长。5月的I/O大会上，谷歌宣布每月通过平台处理480万亿tokens，而现在处理量已经翻倍，增加至每月超过980万亿tokens。应用月活也均有显著增长，Gemini应用目前拥有超过4.5亿月活跃用户，Google Meet的AI会议笔记功能在6月就有超过5000万人使用，Google Vids的月活跃用户已经接近100万。
- 资本支出：2025年Q2的资本支出为224亿美元。由于市场对云产品和服务的需求持续强劲，公司将2025年的资本支出指引从750亿美元上调至约850亿美元。增加的投资将主要用于服务器采购和加速数据中心的建设，主要是为了满足云客户的需求。展望2026年，由于客户需求持续强劲以及公司整体增长机会，公司预计将再次出现资本开支的进一步上升。

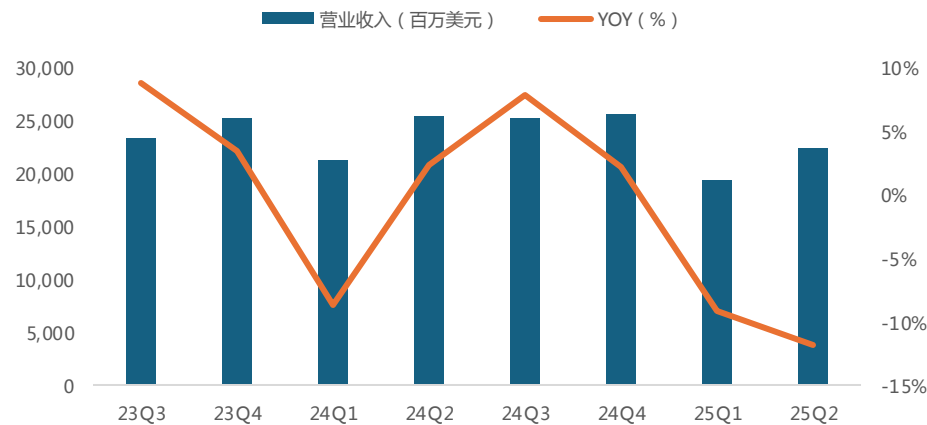
图：谷歌资本支出情况



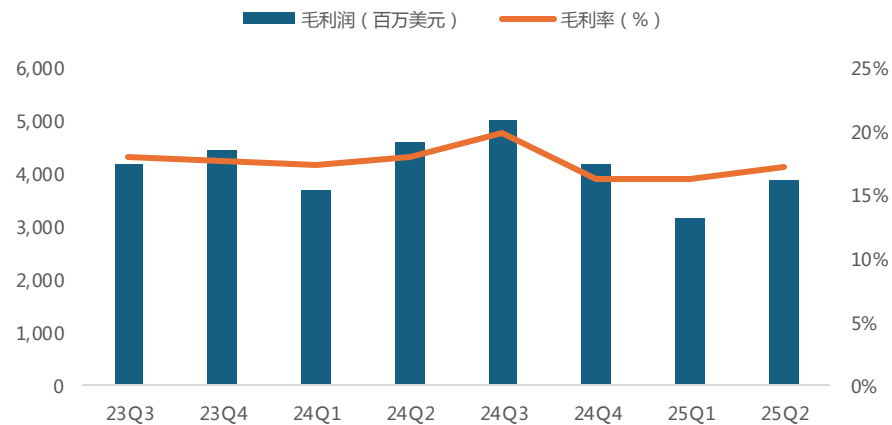
1.2 Tesla：毛利率超预期，Optimus第三代原型机可能将年底公布

- **财务数据：** FY25Q2收入达225亿美元，同比下滑11.78%，低于彭博一致预期（226亿美元）；毛利率达17.2%，超彭博一致预期（16.5%）；运营利润9.23亿美元，同比下滑42.5%，超彭博一致预期（8.54亿美元）；净利润11.7亿美元，同比下滑20.7%，低于彭博一致预期（11.9亿美元）；调整后每股收益（EPS）0.40美元，低于彭博一致预期（0.42美元）。
- **分业务：** 特斯拉汽车业务收入为167亿美元，同比下滑16%；能源业务收入为28亿美元，同比下滑7%；服务及其他收入为30亿美元，同比增长17%。
- **机器人产品：** 特斯拉Optimus人形机器人当前版本为2.5，更先进的Optimus 3可能将于年底推出原型，明年量产。团队目标尽快实现年产100万台，推动人形机器人规模化应用。该产品有望成为特斯拉的下一个现象级产品。

图：Tesla营业收入及增速



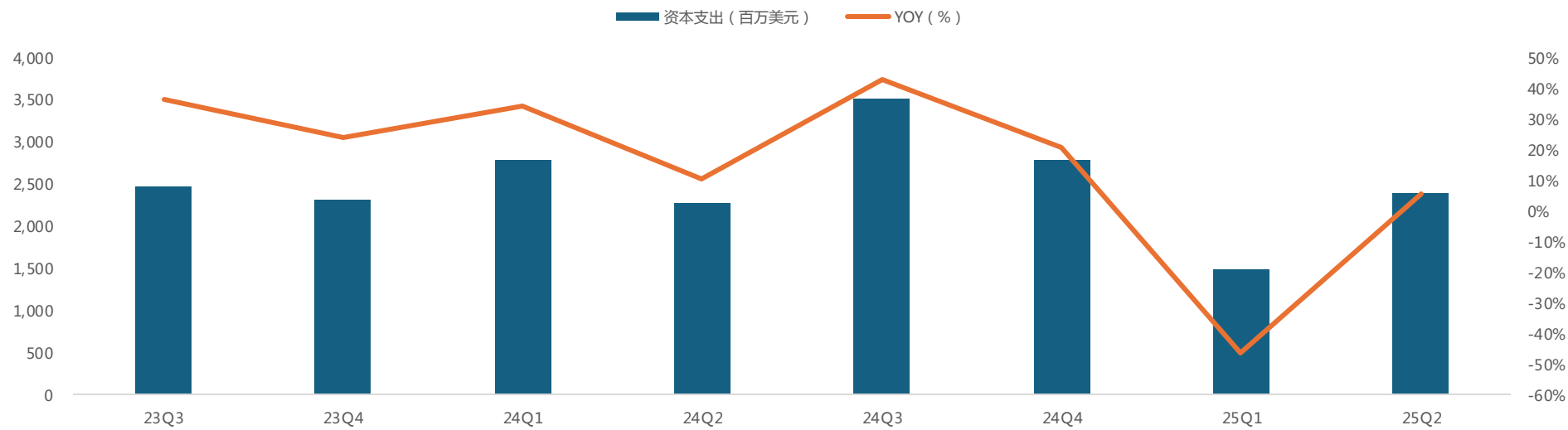
图：Tesla毛利润及毛利率



1.2 Tesla：毛利率超预期，Optimus第三代原型机可能将年底公布

- **FSD相关：** 特斯拉的优势在于产品线中所有车辆均具有自动驾驶功能，且FSD可以将安全性提高10倍。公司开始提高新款Model Y的产量，且Model Y在6月份已成为土耳其、荷兰、瑞士和奥地利最畅销的汽车。公司在奥斯汀推出了Robotaxi服务并完全自主地将汽车直接从工厂送到客户家里，但FSD在欧洲的许可还需等待。公司预计到今年年底，特斯拉会为美国约一半人口提供自动驾驶网约车服务。
- **资本支出：** 特斯拉第二季度资本支出增加至23.9亿美元。公司预计全年资本支出将超90亿美元。公司将继续在制造业各领域进行投资（例如Ctbercab、Semi生产线和其他制造业支出），并扩大公司的人工智能计划。

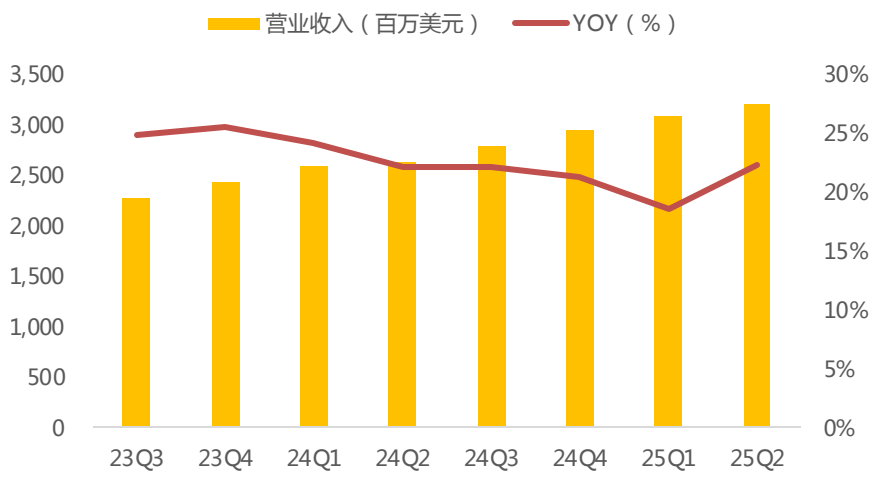
图：Tesla资本支出情况



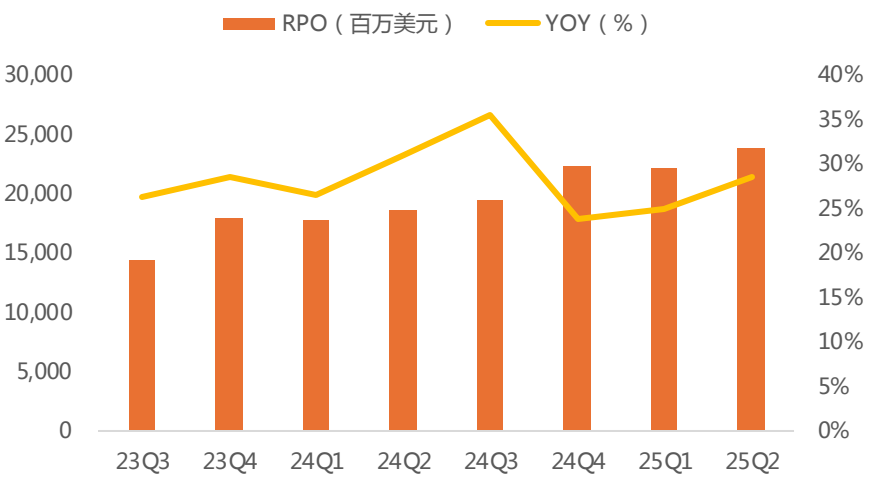
1.3 ServiceNow：收入与RPO全面超预期，上调全年指引

- **财务数据：**ServiceNow FY25Q2总营业收入达32.15亿美元，同比增长22.4%，超彭博一致预期（31.2亿美元）。调整后每股收益（Adj.EPS）为4.09美元，超彭博一致预期（3.58美元）。当前剩余履约义务（cRPO）达到109.2亿美元，同比增长24.4%，超彭博一致预期（104.9亿美元）。剩余履约义务达239亿美元，同比增长28.5%，超彭博一致预期（225亿美元）。
- **分业务：**本季度订阅业务收入达31.13亿美元，同比增长22.46%。专业服务收入为1.02亿美元，同比增长20%。
- **业绩指引：**公司将2025财年全年的订阅收入指引中值上调1.25亿美元，达到127.75亿至127.95亿美元。预计第三季度订阅收入将在32.60亿至32.65亿美元之间，2026年将实现150亿美元以上的订阅收入目标。公司预计Q3的cRPO同比增长18.5%，其中包括Q4续约客户数量高于平均水平带来2%的不利因素。

图：ServiceNow营业收入及增速



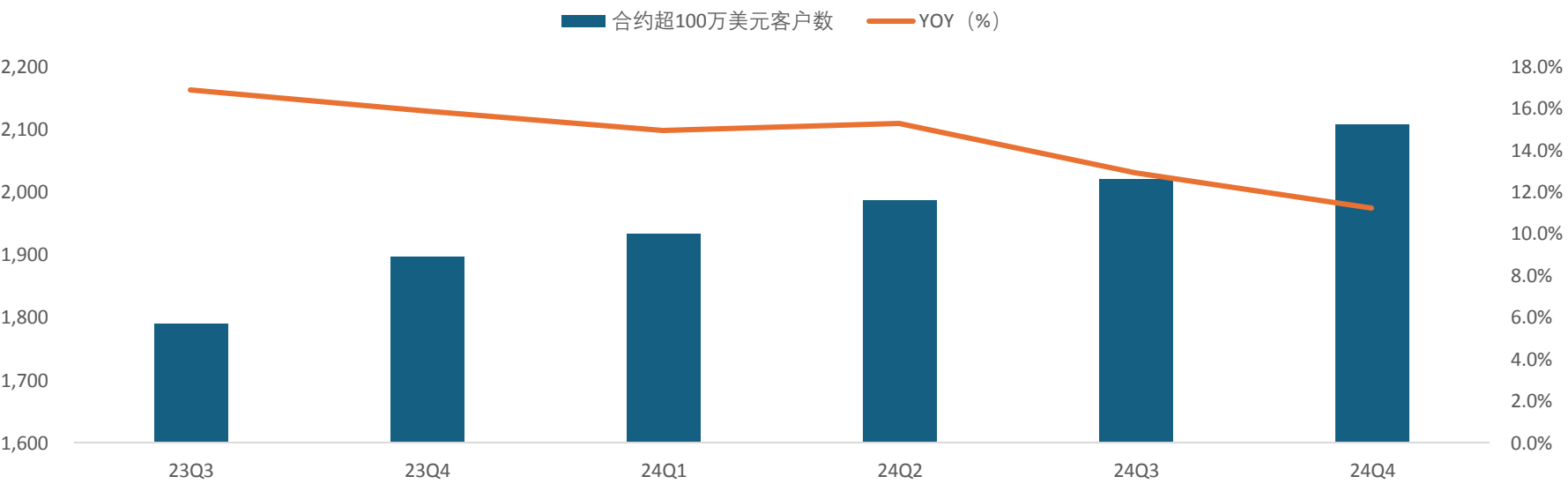
图：ServiceNow剩余履约义务及增速



1.3 ServiceNow：收入与RPO全面超预期，上调全年指引

- 预计2026年ACV达到10亿美元：公司表示AI产品Now Assist 本季度收入超预期，目标2026年ACV达到10亿美元。公司在第二季度推出了AI控制塔（AI Control Tower）和AI代理结构（AI Agent Fabric）等新产品，旨在帮助企业集中管理和扩展AI应用。公司已完成对数据目录和治理平台data.world的收购，此举将增强ServiceNow AI平台的数据治理能力。
- 大客户持续增长：本季度新增了89笔价值超过100万美元年度合同价值（ACV）的交易。年合同价值超过2000万美元的客户数量同比增长超过30%。截至季末，拥有超过500万美元ACV的客户数达到528家，同比增长约19.5%。

图：合约超100万美元客户数及增速



全球AI动态更新

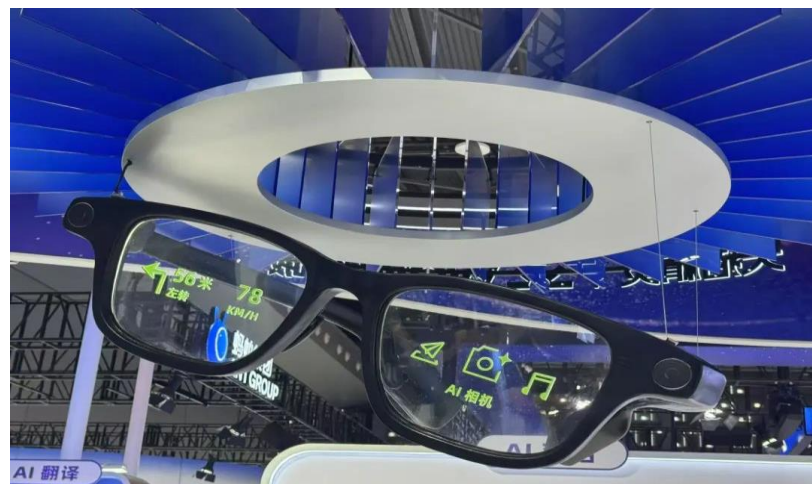
2.1 人工智能大会正式召开，聚焦 "模型应用 + 具身智能 + 智能硬件"

- 2025年7月27日，2025世界人工智能大会（WAIC）正式拉开帷幕。作为历届规模最为盛大的一届盛会，本次大会吸引了800余家企业参与，集中呈现了40余款大模型、50余款人工智能终端产品、60余款智能机器人，以及80余款"全球首发"或"中国首秀"的重磅新品。
- 本届展会以"模型应用 + 具身智能 + 智能硬件"为显著特征，展览面积首次突破7万平方米。国内外科技领军企业及人工智能领域明星企业悉数参展，系统展示了覆盖消费级、商业级、工业级的各类大模型及人工智能应用成果。其中，具身智能与机器人展区成为焦点，特斯拉、宇树科技等企业携多款首发及亮点机器人产品亮相，引发持续关注。智能硬件领域，阿里巴巴首发夸克AI眼镜，眼镜深度融合阿里及支付宝生态，具备通义千问大模型能力，搭载夸克最新AI搜索和问答技术，还支持高德地图导航、淘宝比价、支付宝扫码支付、飞猪行程提醒等功能，成为展区核心看点之一。
- 本次展会汇聚了20余家芯片企业，智算领域更是企业云集。既包含人工智能服务器厂商，也涵盖智算云及AI Infra企业，全面展现了人工智能产业生态的核心布局与前沿发展态势。

图：图灵奖得主、诺贝尔物理学奖得主Geoffrey Hinton出席并作演讲



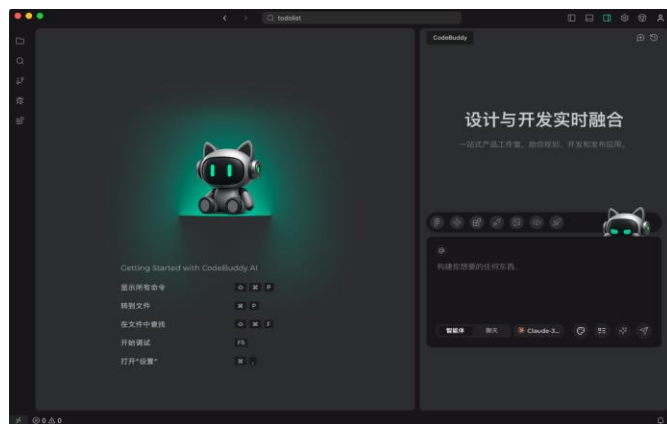
图：阿里推出首款自研AI眼镜



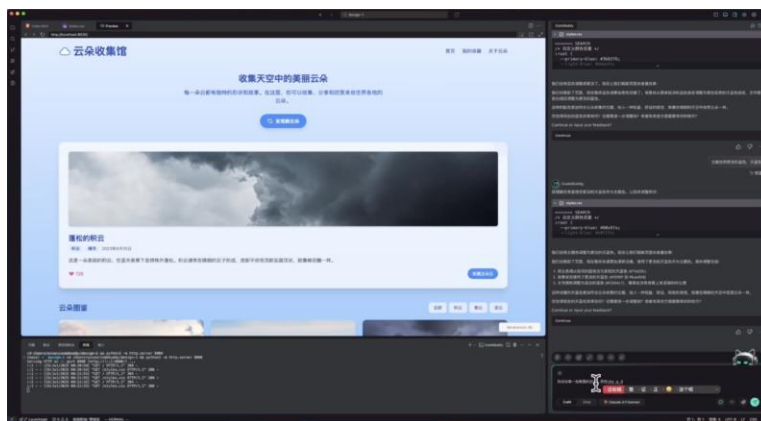
2.2 腾讯：推出CodeBuddy AI IDE，实现“对话即编程”

- 7月22日，腾讯云推出全球首位贯通“产品-设计-研发”的一体化AI全栈工程师——CodeBuddy AI IDE。该工具支持通过自然语言对话输入需求，自动完成从产品构想到设计原型、前后端编码、上线部署的全流程，实现“对话即编程”。其核心在于整合了多类主流大模型（国际版含 Claude、GPT、Gemini 等，国内版支持腾讯混元、DeepSeek 等国产模型），解决了AI曾经在快速交互原型可视化验证、设计稿转代码、前后端联调一体化编码等场景无法落地的问题。实战案例显示，在电商活动页开发中，该工具将传统 2 天的工作量压缩至不到 2 小时，效率提升超 10 倍。目前该工具已开启内测，用户可通过官方网站申请体验。
- **全链路操作流程：**一，明确需求输出 PRD；在 Plan 模式下可通过自然语言生成结构化 PRD，直接对接设计与开发。二，原型与 UI 设计；支持自然语言生成、图像化生成、组件化生成三种方式，且能通过指令实时优化视觉风格。三，前端开发；内置 Figma，可从设计稿一键生成可维护的 HTML 源码，支持多端适配、统一风格。四，后端开发；内置腾讯云开发 CloudBase（国际版含 Supabase），自动配置云数据库、用户认证等后端需求。五，上线部署；支持通过 CloudStudio 一键部署至远端，生成可访问链接。此外，CodeBuddy AI IDE 具备软件开发智能体、实时代码续写、行间/终端内联对话等能力，进一步提升开发效率。

图：CodeBuddy AI IDE 操作界面



图：原型与UI设计环节



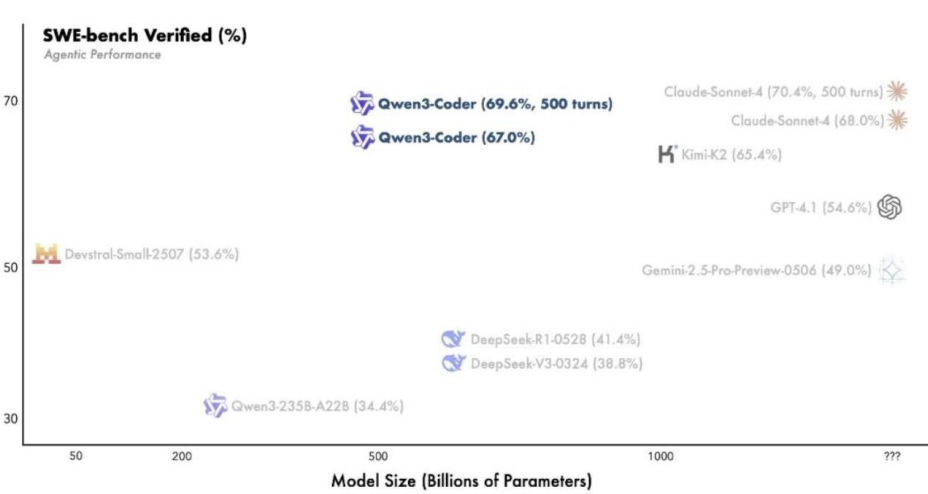
2.3 阿里巴巴：本周开源三款重量级模型，模型能力进一步提升

- 7月23日，阿里云正式发布其最具代理能力的代码模型：Qwen3-Coder系列模型。Qwen3-Coder拥有多个尺寸，其中最强大的版本是Qwen3-Coder-480B-A35B-Instruct。该模型是一个480B 参数激活 35B 参数的 MoE 模型，拥有卓越的代码和 Agent 能力。Qwen3-Coder-480B-A35B-Instruct 在 Agentic Coding、Agentic Browser-Use 和 Agentic Tool-Use 上取得了开源模型的 SOTA 效果，可以与 Claude Sonnet4 媲美。
- **训练方法：**在预训练环节，Qwen3-Coder 从数据、上下文、合成数据三方面进行 Scaling。模型使用 7.5T 训练数据（代码占比 70%）；原生支持 256K 上下文，可借助 YaRN 技术扩展至 1M；并采用 Qwen2.5-Coder 清洗与重写低质合成数据，显著提升训练集质量。后训练阶段，团队通过自动扩展测试样例，构建出更丰富的真实代码任务，并以此开展 Code RL 训练，显著提升了代码执行成功率；此外，团队执行了 Agent RL，鼓励模型通过多轮交互的方式利用工具解决问题，并借助阿里云基础设施，最终在SWE-bench Verified上实现了开源模型SOTA。

图：Qwen3-Coder领先各大榜单

QWEN3-CODER						
Benchmarks	Qwen3-Coder 480A35-Instruct		Open Models		Proprietary Models	
			Kimi-K2 Instruct	DeepSeek-V3 0324	Claude Sonnet-4	OpenAI GPT-4.1
Agentic Coding						
Terminal-Bench	97.5		30.0	2.5	35.5	25.3
SWE-bench Verified	69.6		-	-	70.4	-
w/ OpenHands, 500 turns	67.0		65.4	34.0	68.0	48.6
w/ OpenHands, 100 turns	-		65.8	-	72.7	63.8
w/ Private Scaffolding	26.3		22.3	13.0	27.7	-
SWE-bench Live	54.7		47.3	13.0	53.3	31.5
SWE-bench Multilingual	25.0		19.0	7.5	24.8	-
Multi-SWE-bench-mini	27.0		20.7	-	25.0	-
Multi-SWE-bench-flash	41.0		60.0	56.9	56.4	52.4
Aider-Polyglot	31.1		25.2	-	25.8	-
Spider2						
Agentic Browser Use						
WebArena	49.9		47.4	48.0	51.1	44.3
Mind2Web	55.0		42.7	36.0	47.4	49.6
Agentic Tool Use						
BFCL-v3	68.7		65.2	56.9	73.3	62.9
TAU-Bench Retail	77.5		70.7	59.1	80.5	-
TAU-Bench Airline	68.0		53.5	48.0	60.0	-

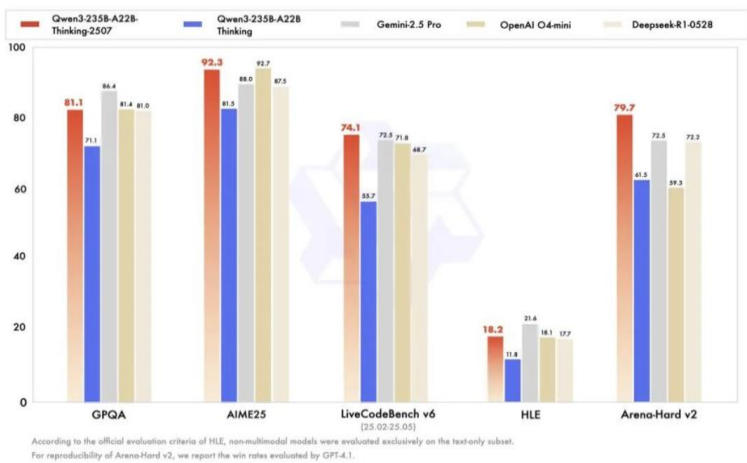
图：SWE-bench Verified



2.3 阿里巴巴：本周开源三款重量级模型，模型能力进一步提升

- 7月25日，阿里正式发布全球最强开源推理模型——Qwen3-235B-A22B-Thinking-2507。作为一款专为「思考模式」打造的模型，该模型原生支持 256K 上下文处理能力，可应对更长文本、构建更深的推理链，自动启用多步推理。基准测试结果显示，这款推理模型的表现与顶级闭源模型（如Gemini-2.5 pro、O4-mini）不相上下。定价方面，Qwen3-235B-A22B-Thinking-2507 每输入百万 token 0.7 美元，每输出百万 token 8.4 美元。
- **阿里开源策略：**此推理模型已是阿里本周开源的第三款重量级模型。从基础模型（Qwen3-235B-A22B-Instruct-2507 非思考版）、编程模型（Qwen3-Coder）到推理模型，阿里凭借 Qwen3 系列打出一套开源“进攻组合拳”，不仅填补了国产模型在通用能力上的差距，也为“开源可替代闭源”这一命题提供了极具说服力的现实注解。在模型能力方面，知名 AI 研究机构 Artificial Analysis 指出，“Qwen3 是全球最智能的非思考基础模型”。在商业化落地方面，海外知名模型 API 聚合平台 OpenRouter 数据显示，阿里 Qwen API 已突破 1000 亿Tokens，在 OpenRouter 趋势榜上包揽全球前三。

图：Qwen3-235B-A22B-Thinking-2507与顶级闭源模型不相上下



图：阿里模型定价

Alibaba Cloud Model Studio
阿里云模型工作室

Qwen3-235B-A22B-Thinking-2507

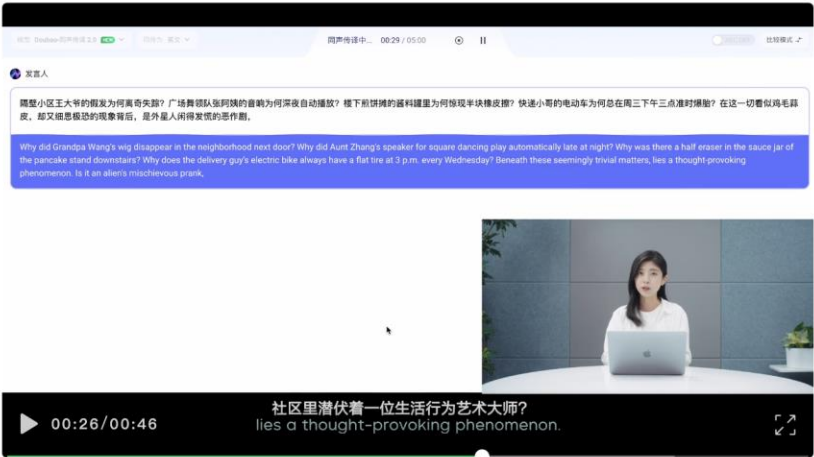
Thinking mode | Non-Thinking mode | Usage instructions

Name 名称	Mode 模式	Context window 上下文窗口	Maximum input 最大输入	Maximum CoT 最大思维链长度 (Tokens) (标记)	Maximum response 最大响应
qwen3-235b-a22b-thinking-2507	Thinking only 仅思考	131,072	126,976	81,920	32,768
qwen3-235b-a22b-instruct-2507	Non-Thinking only 仅非思考			-	
qwen3-235b-a22b	Non-Thinking 非思考		129,024	-	
qwen3-32b	Non-Thinking 非思考		98,304	38,912	16,384
qwen3-30b-a3b	Thinking 思考	131,072	129,024	38,912	

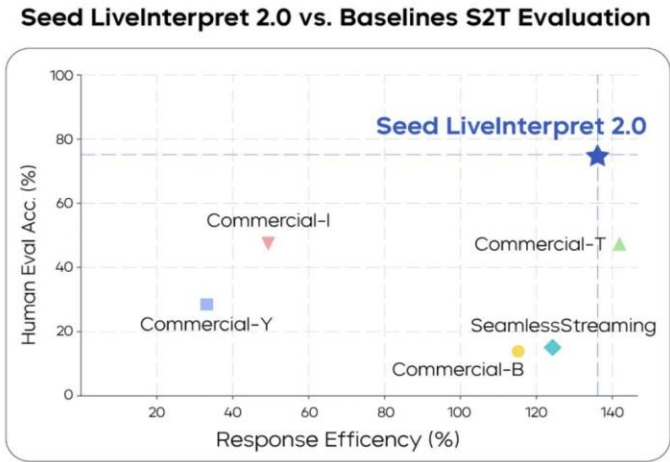
2.4 字节跳动：发布同声传译模型 Seed LiveInterpret 2.0，低延迟高准确率

- 7月24日，火山引擎正式发布豆包·同声传译模型 Seed LiveInterpret 2.0，该模型已全量上线火山方舟平台，是首个延迟与准确率接近人类水平的产品级中英语音同传系统。其核心突破在于，在中英同传翻译质量达到业界 SOTA 的同时，实现了极低的语音延迟和“0 样本声音复刻”，可基于端到端理解和生成框架，实时接收多人语音输入，同步完成理解、声音复刻，并输出语音和译文，真正实现“边听边说”。该模型已在多个权威国际会议中应用，并即将在8月上线 Ola Friend 耳机，为用户带来更加灵活的使用方式和体验。
- **模型优势：**一，极低语音延迟；该模型采用全双工语音理解生成框架，翻译延迟低至 2-3 秒，较传统机器同传系统降低超 60%，实现“边听边说”，保障沟通无断点。二，0 样本声音复刻；无需提前录制，可边说话边采样，实现同款音色外语输出，提升听众沉浸感。三，智能平衡节奏；能根据语音清晰度、流畅度、复杂程度调整输出节奏，适配不同语言特性，即使面对超长信息，也能保持传译节奏自然流畅。

图：低时延输出同款音色英语翻译



图：语音到文本 (S2T) 任务中，人工评测的翻译质量和响应效率



2.5 谷歌：DeepMind正式官宣拿下IMO金牌，仅用自然语言完成解题

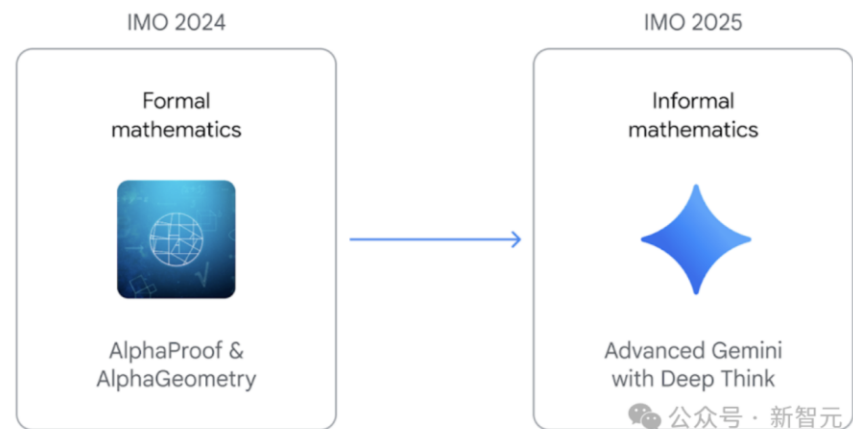
- 7月21日，谷歌 DeepMind 正式官宣其新模型 Gemini Deep Think（高阶版）在国际数学奥林匹克竞赛（IMO）中取得 35 分（满分 42 分），达到金牌水平。该模型在 4.5 小时内攻克了 6 道题中的 5 道，且仅使用自然语言（英语）完成解题。此外，该成果得到 IMO 组委会官方认证，而同期 OpenAI 的相关成绩未经过官方验证。谷歌计划先向部分数学家等测试者提供该模型，随后向 Google AI Ultra 订阅者推出。
- **Gemini Deep Think：**团队使用了Gemini Deep Think的增强推理模式，结合了并行思考技术，允许模型同时探索多种解题路径并整合最优答案，突破了传统单一线性思考的局限。在训练方面，谷歌采用了新颖的强化学习技术，让其利用更多多步推理、问题解决和定理证明数据；同时，模型还获取了过往问题的高质量解决方案集，并得到解决 IMO 问题的通用提示与技巧。这种“训练+知识库+策略”的组合，使其能够直接从官方问题描述中生成严谨的数学证明，无需依赖专业“形式语言”。尽管夺下IMO金牌虽然完全基于Gemini自然语言能力，但团队也在AlphaGeometry和AlphaProof等形式化系统方面也取得了持续的进展。

图：2024年谷歌模型获得IMO银牌

Score on IMO 2024 problems



图：DeepMind IMO 模型的演进，从形式到非形式数学的跨越



分析师声明

本报告署名分析师在此声明：我们具有中国证券业协会授予的证券投资咨询执业资格或相当的专业胜任能力，本报告所表述的所有观点均准确地反映了我们对标的证券和发行人的个人看法。我们所得报酬的任何部分不曾与，不与，也将不会与本报告中的具体投资建议或观点有直接或间接联系。

一般声明

除非另有规定，本报告中的所有材料版权均属天风证券股份有限公司（已获中国证监会许可的证券投资咨询业务资格）及其附属机构（以下统称“天风证券”）。未经天风证券事先书面授权，不得以任何方式修改、发送或者复制本报告及其所包含的材料、内容。所有本报告中使用的商标、服务标识及标记均为天风证券的商标、服务标识及标记。

本报告是机密的，仅供我们的客户使用，天风证券不因收件人收到本报告而视其为天风证券的客户。本报告中的信息均来源于我们认为可靠的已公开资料，但天风证券对这些信息的准确性及完整性不作任何保证。本报告中的信息、意见等均仅供客户参考，不构成所述证券买卖的出价或征价邀请或要约。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。客户应当对本报告中的信息和意见进行独立评估，并应同时考量各自的投资目的、财务状况和特定需求，必要时就法律、商业、财务、税收等方面咨询专家的意见。对依据或者使用本报告所造成的一切后果，天风证券及/或其关联人员均不承担任何法律责任。

本报告所载的意见、评估及预测仅为本报告出具日的观点和判断。该等意见、评估及预测无需通知即可随时更改。过往的表现亦不应作为日后表现的预示和担保。在不同时期，天风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。

天风证券的销售人员、交易人员以及其他专业人士可能会依据不同假设和标准、采用不同的分析方法而口头或书面发表与本报告意见及建议不一致的市场评论和/或交易观点。天风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。天风证券的资产管理部门、自营部门以及其他投资业务部门可能独立做出与本报告中的意见或建议不一致的投资决策。

特别声明

在法律许可的情况下，天风证券可能会持有本报告中提及公司所发行的证券并进行交易，也可能为这些公司提供或争取提供投资银行、财务顾问和金融产品等各种金融服务。因此，投资者应当考虑到天风证券及/或其相关人员可能存在影响本报告观点客观性的潜在利益冲突，投资者请勿将本报告视为投资或其他决定的唯一参考依据。

投资评级声明

类别	说明	评级	体系
股票投资评级	自报告日后的6个月内，相对同期沪深300指数的涨跌幅	买入	预期股价相对收益20%以上
		增持	预期股价相对收益10%-20%
		持有	预期股价相对收益-10%-10%
		卖出	预期股价相对收益-10%以下
行业投资评级	自报告日后的6个月内，相对同期沪深300指数的涨跌幅	强于大市	预期行业指数涨幅5%以上
		中性	预期行业指数涨幅-5%-5%
		弱于大市	预期行业指数涨幅-5%以下

THANKS