

智变·2025：互联网 AI 应用元年

华泰研究

2025年3月26日 | 中国香港

专题研究

互联网 增持 (维持)
教育和人力资源 增持 (维持)

研究员 夏路路
SAC No. S0570523100002 xialulu@htsc.com
SFC No. BTP154 +(852) 3658 6000

研究员 苏燕妮
SAC No. S0570523050002 suyanni@htsc.com
SFC No. BTT483 +(86) 21 2897 2228

研究员 丁晓璇
SAC No. S0570523040003 dingjiaowan@htsc.com
SFC No. BPJ942 +(86) 21 2897 2228

研究员 侯杰
SAC No. S0570523010002 houjie017864@htsc.com
SFC No. BRI004 +(86) 10 6321 1166

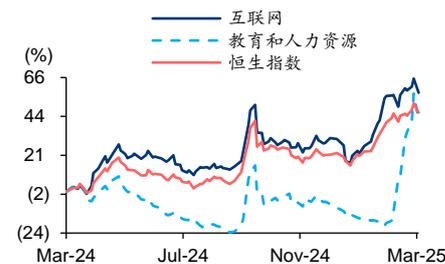
研究员 詹博
SAC No. S0570523110002 zhanbo@htsc.com
SFC No. BUS698 +(86) 21 2897 2228

研究员 郑裕佳
SAC No. S0570524070002 zhengyujia@htsc.com
SFC No. BTB676 +(86) 10 6321 1166

联系人 曹卓铭
SAC No. S0570124020016 caozhuoming@htsc.com
SFC No. BUT776 +(86) 21 2897 2205

联系人 邵浩岚
SAC No. S0570124070056 shaohaolan@htsc.com
+(86) 21 2897 2228

行业走势图



资料来源: Wind, 华泰研究

重点推荐

股票名称	股票代码	目标价 (当地币种)	投资评级
阿里巴巴-W	9988 HK	185.10	买入
腾讯控股	700 HK	637.28	买入
百度集团	BIDU US	113.70	买入
美图公司	1357 HK	7.49	买入
快手-W	1024 HK	68.73	买入
哔哩哔哩	BILI US	34.40	买入
有道	DAO US	12.04	买入
好未来	TAL US	15.08	买入
BOSS 直聘	BZ US	22.71	买入

资料来源: 华泰研究预测

我们认为以 DeepSeek 引领中国 AI 大模型在模型性能、成本效用上实现突破开始, 到阿里、字节、腾讯等大厂以 capex 开启新一轮科技基建周期, 再到 Manus 等下游 AI 应用引起广泛讨论, 25 年开始中国互联网 AI 或进入元年。AI 模型层面, 国内头部模型与海外差距正在日益缩小, 在本就繁荣的消费互联网应用层或将迎来产品的百花齐放和商业化落地的全新机遇。基于此, 我们提出全生态互联网大厂+“小而美”垂类赛道的两大投资主线, 看好互联网大厂们以创新能力、执行力和企业家精神再次成为 AI 时代领头羊。

AI 模型：国产模型能力迎头赶上，大厂领衔开启新一轮科技基建周期

DeepSeek 引领中国 AI 大模型在模型性能、成本效用上实现突破, 国产大模型在主流权威基准测试中已展现出全球领先的模型性能, 并不断缩小与国际领先模型能力的差距。同时由阿里、字节等互联网大厂领衔正在开启新一轮科技基建周期 (对标美国 23-24 年), 在 AI 基础设施上加码投入, 根据我们统计 25 年大厂资本开支合计或超 5000 亿元, 将大幅度完善我国在 AI 算力、数据中心等建设的短板, 为后续 AI 下游应用侧的繁荣奠定基础。

AI 应用：多场景渗透，迎来产品的百花齐放和商业化落地的全新机遇

受益于开源模型、算力成本的降低、多模态的融合, 我们判断 25 年 AI 应用或迎来爆发, 并带来两个发展趋势, 一方面更广泛场景的 AI 渗透率将显著提升, 数据要素的稀缺性将使得 AI 应用趋向细分、垂直化。另一方面产品层面革新, 聚焦用户丰富交互模式、降低门槛, 由 Manus 带来的 AI Agent 趋势或将驱动 AI 应用实现从思考到执行的全闭环。同时我们认为在商业化落地节奏上或有分化, 通用型 AI 将迈入流量入口争夺阶段, 积累用户并探索生态的优先级短期大于变现, 或在下个阶段看到广泛的 Agent 应用在大厂生态内爆发, 带来中长期变现潜力; 而垂直赛道 AI 则靠提供增值产品和体验, 以生产力等刚需场景驱动较高的商业转化, 建议关注试错成本低、可验证性高的图像编辑、短 (长) 视频内容、广告营销、教育培训答疑等赛道。

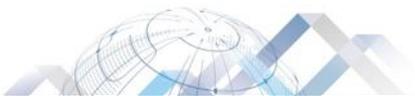
我们与市场不同的观点

我们对于互联网 AI 的相关底层模型、应用场景等等做了全方位的梳理和总结, 在当前市场对于 AI 模型能力、应用场景仍有分歧的阶段, 给予了我们明确的判断和观点。我们从移动互联网的经验出发, 探讨了 AI 应用的潜在发展趋势和竞争要素, 并进一步讨论了其商业化的潜力和场景的落地节奏, 提出了商业化节奏在通用型和垂直型分化的观点, 并提供了试错成本和可验证性这一双维度指标作为场景落地难易的判断依据, 均具备一定的差异性。

投资结论

中国互联网 AI 或进入元年, 关注两大主线: 1) 投资具备优质底层模型+算力 (云) 基础设施+丰富下游应用生态的互联网大厂, 首推阿里、腾讯, 关注百度, 其正在通过 capex 引领新一轮国内 AI 基础设施建设周期, 在优质算力 (云) 基础设施下担当中国 AI 科技周期启动的“卖铲人”。同时这些大厂下游已经具备完整的应用生态和丰富的用户基础, 有望在 AI 的新一轮竞赛中取得领衔身位。2) 投资卡位优质垂直场景+AI 产品布局充分+商业化落地确定性较高的“小而美”细分赛道龙头, 首推美图、快手、哔哩哔哩、有道, 关注好未来、BOSS 直聘, 其原有业务卡位细分垂直场景, 具备独有数据、用户资源。在 AI 端布局充分, 具备优质产品力的企业有望获得新增长曲线, 细分垂直场景有望凭借提供的生产力改善等更快实现财报端的兑现。

风险提示: 模型性能迭代不及预期, 商业化进展缓慢, 监管收紧。本研报中涉及到未上市公司或未覆盖个股内容, 均系对其客观公开信息的整理, 并不代表本研究团队对该公司、该股票的推荐或覆盖。



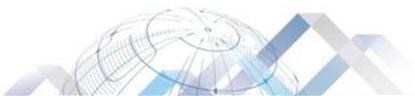
正文目录

核心观点	6
核心逻辑与投资结论	6
我们与市场不同的观点	7
互联网 AI 关键问题五问五答	7
AI 模型：算力规模飞速增长，成本效应未来可期	12
腾讯混元：全面拥抱开源生态，模型性能大幅提升	16
阿里巴巴：AI 驱动云计算业务未来	19
战略：不断完善 AI 基础设施，坚持开源开放	20
基础大模型布局：全尺寸、全模态、多场景的模型大家族	20
AI 能力加持下，云计算服务有望实现降本增收	22
百度：全栈服务能力布局助力 AI 应用落地	23
战略：注重产业应用场景与智能体开发	23
基础大模型布局：拥有多种类大模型满足用户需求	24
全栈服务能力为 AI 应用落地奠定良好基础	25
AI 应用：多场景渗透，产品力持续迭代	28
AI 应用的发展趋势和竞争要素	28
广泛场景的渗透率提升，趋向细分、垂直化	29
产品层面革新，聚焦用户丰富交互模式、降低门槛	29
注重垂直场景的产品力，认知重构阶段把握用户心智	31
AI 应用的商业化落地节奏判断	31
AI 全生态应用布局典范：字节跳动、阿里、腾讯、百度	32
字节跳动：搭建 AI 全生态，爆款豆包产品力先行	32
阿里巴巴：规划布局深入千行百业	34
腾讯：加速游戏工业化，引领广告新趋势	35
百度：AI 重做产品，引领业务智能化转型升级	37
AI+图像：受益于 DS 推进技术平权，持续图像领域精耕细作	40
AI+视频：细分领域贵于深耕，用户习惯有望迎来变局	43
AI+游戏：3D 生成模型助力研发进度，AI agent 重塑游戏交互	45
AI+教育：积极推出大模型产品	46
AI+招聘：提升招聘效率，提升用户体验	50
AI+广告：赋能互联网营销平台提质增收	52
AI+电商：淘天多年数据沉淀打造增长基础	54
投资观点	58
风险提示	58

图表目录

图表 1: 当前主流 AI 应用的订阅付费用户及 ARR.....	9
图表 2: 阿里 Capex 投入情况.....	10
图表 3: 腾讯 Capex 投入情况.....	10
图表 4: 快手 Capex 投入情况.....	11
图表 5: 百度 Capex 投入情况.....	11
图表 6: AI 大模型对于行业竞争壁垒的影响.....	11
图表 7: 互联网公司大模型参数.....	12
图表 8: 国内主要 LLM 模型对比.....	13
图表 9: 国内部分通用 AI APP DAU.....	14
图表 10: 国内部分通用 AI APP DAU.....	14
图表 11: 海外部分通用 AI APP DAU.....	14
图表 12: 海外部分通用 AI 网站版 DAU.....	14
图表 13: 国内部分通用 AI APP MAU.....	15
图表 14: 国内部分通用 AI APP MAU.....	15
图表 15: 海外部分通用 AI APP MAU.....	15
图表 16: 海外部分通用 AI 网站版 MAU.....	15
图表 17: 国内部分通用 AI APP 人均单日使用时长.....	15
图表 18: 国内部分通用 AI APP 人均单日使用时长.....	15
图表 19: 海外部分通用 AI APP 人均单日使用时长.....	16
图表 20: 海外部分通用 AI 网站版人均单日使用时长.....	16
图表 21: 国内部分通用 AI APP DAU/MAU.....	16
图表 22: 国内部分通用 AI APP DAU/MAU.....	16
图表 23: 海外部分通用 AI APP DAU/MAU.....	16
图表 24: 海外部分通用 AI 网站版 DAU/MAU.....	16
图表 25: 混元大模型加速大模型能力落地场景.....	17
图表 26: 腾讯大模型全面拥抱开源.....	17
图表 27: 腾讯混元大模型团队结构一览.....	18
图表 28: 腾讯部分模型及价格.....	19
图表 29: 通义大模型大家族全景图.....	20
图表 30: 阿里旗下部分模型介绍和收费情况.....	21
图表 31: Qwen2.5-Max 在指令模型版本中的性能表现.....	21
图表 32: Qwen2.5-Max 在基座模型版本中的性能表现.....	22
图表 33: 阿里云拥有全球+全栈的技术能力.....	23
图表 34: 百度文心大模型与工具平台全景图.....	24
图表 35: 百度文心大模型家族全景图.....	24
图表 36: 百度旗下 AI 大模型调用费用.....	25
图表 37: 百度智能云业务版图.....	26
图表 38: 中国模型即服务 (MaaS) 市场份额, 1H24.....	27
图表 39: 中国 AI 大模型方案市场份额, 1H24.....	27

图表 40: 开源 DeepSeek 在基准测试中好于一众闭源模型	28
图表 41: 大模型调用价格持续呈现下降趋势	28
图表 42: 移动互联网的应用逻辑与 AI 时代应用逻辑的对比	28
图表 43: AI 在更广泛的应用场景下具备渗透率提升的前景	29
图表 44: Chatgpt 优化的交互模式	30
图表 45: Manus 在 GAIA 上的评分	30
图表 46: Manus 的使用案例	30
图表 47: 当前主流 AI 应用的订阅付费用户及 ARR	32
图表 48: 字节 AI 应用全布局	33
图表 49: 国内 AI 应用月活	33
图表 50: 豆包 APP 发布后持续迭代	34
图表 51: 阿里旗下 AI 部分应用	34
图表 52: 阿里 To C AI 业务近期组织架构变化	35
图表 53: 腾讯产品布局一览	36
图表 54: 腾讯视频号 Ad load 提升空间显著	36
图表 55: 智能体商店涵盖了各种应用场景	37
图表 56: 在智能体创建环节需要明确角色、设定、开场白、流程、插件等内容	37
图表 57: 百度旗下 AI 产品	38
图表 58: 百度 AI 搜索结果示意图	39
图表 59: 百度搜索中 AI 生成内容结果占比	39
图表 60: 百度 AI 广告业务梳理	40
图表 61: 美图 AI 产品矩阵	40
图表 62: 美图 AI 电商设计工作流程和 AI 口播工作量	41
图表 63: 美图设计室收入和 MAU 持续提升	42
图表 64: 开拍收入和 MAU 持续提升	42
图表 65: 美图 AI 影像业务生态	42
图表 66: 淘宝星辰·图生视频有望为商家营销物料提质增效	43
图表 67: 可灵与即梦 AI MAU 与 DAU 对比 (24 年 12 月)	43
图表 68: 可灵与即梦 AI 人均单日使用时长对比 (24 年 12 月)	43
图表 69: 剪映 MAU 相较于其他 AI 剪辑工具下降明显	44
图表 70: 剪映人均单日使用时长与其他 AI 剪辑工具对比	44
图表 71: AI 搜索助手	44
图表 72: AI 视频助手	44
图表 73: AI 赋能游戏	45
图表 74: Unity 推出 Sentis, 赋予开发者将 AI 模型导入游戏和应用程序中的能力	45
图表 75: AI agent 重塑游戏交互, 实现快速建立智能 NPC	46
图表 76: 行业内主要模型	46
图表 77: 九章大模型测试数据	47
图表 78: AI 开放平台	47
图表 79: 大模型开放能力	47



图表 80: “子曰”大模型及其应用	48
图表 81: 子曰翻译大模型 2.0 测试评分表现优秀	49
图表 82: 子曰-o1 测试评分	49
图表 83: 有道小 P App 目前已接入 DeepSeek	50
图表 84: 南北阁开源大模型	51
图表 85: BOSS 直聘模型应用	51
图表 86: 猎聘 AI 智能面试	52
图表 87: AXON 2.0 推动 Applovin 收入同比持续高增	52
图表 88: 汇量科技程序化广告业务收入持续高增	52
图表 89: 有道广告业务收入连续 8 个季度同比增速超 40% (截至 24Q3)	53
图表 90: 星辰 AI 模型	54
图表 91: AI+电商从技术到价值的赋能逻辑	54
图表 92: 达摩盘基于四大维度提供精细化标签	55
图表 93: 阿里妈妈发布 LMA 大模型技术	55
图表 94: 淘宝问问和京东京言对于具体特定穿搭建议的回答	56
图表 95: 淘宝问问和京东京言对于泛化的办公产品建议的回答	57
图表 96: 重点推荐公司一览表	58
图表 97: 重点推荐公司最新观点	59

核心观点

核心逻辑与投资结论

1、AI 模型层面

DeepSeek 引领中国 AI 大模型在模型性能、成本效用上实现突破，未来阿里字节等大厂或相继接力，合力夯实新一轮 AI 大模型科技竞赛中的中方竞争力。我们认为以 DeepSeek 的出现为分界线，AI 大模型的发展阶段可以划分为海外大厂保持代际领先与中国厂商迎头赶上两个阶段。在第二阶段，算力层面的 Scalling law 仍然奏效，但不再是通往“罗马”的唯一路径，这亦为中国 AI 大模型走出实验室研究、走向日常应用带来了产业级别的机遇。新阶段，AI 大模型的核心生产要素中“数据”的差异化价值有望进一步突出，在本就繁荣的消费互联网应用层或将迎来产品的百花齐放和商业化落地的全新机遇，在引领全球消费互联网商业生态迈向新格局、新机遇的同时，也为国内以阿里、腾讯、百度等为代表的“卖铲人”提供云服务业绩的增厚。对于模型层，算力层面差距的收窄或将进一步推动基础大模型服务提供者未来更多聚焦与应用侧的适配链接，在基础能力层演化出差异化格局。

2、AI 应用层面

行业趋势和竞争要素判断：我们判断 25 年开始 AI 应用可能迎来一个爆发期，背后的动力可能来自于开源模型、算力成本的降低、多模态的融合。同时带来两个发展趋势，一方面更广泛场景的 AI 渗透率将显著提升，由于数据要素的稀缺性，AI 应用趋向细分、垂直化。另一方面产品层面革新，聚焦用户丰富交互模式、降低门槛，由 Manus 带来的 AI Agent 趋势或将驱动 AI 应用实现从思考到执行的全闭环。而在此基础上，我们认为 AI 应用的核心竞争要素是聚焦垂直场景的产品力，正在逐步从“技术竞争”走向“产品竞争”，同时用户对于 AI 应用处在认知重构阶段，需要供给侧的持续启发夺取心智，最终通过用户反馈积累数据资产，不断迭代并深化实现飞轮。

商业化落地节奏判断：我们认为从传统移动互联网的变现模式出发（广告、佣金和增值服务），当前 AI 应用时代最快具备商业化基础的是靠 AI 提供的差异化产品和体验，并收取订阅费的增值服务。而 AI 应用当前尚未形成新的流量高地且产品层面亦未成熟，因此在广告和佣金等商业模式上进展暂缓，但是当前看到的是互联网大厂正在进行通用型 AI 流量入口的激烈争夺，并且随着 Agent 的发展，大厂的全生态业务将具备显著优势。由此我们也能看到通用型 AI 和垂直型 AI 在商业化节奏上的分化。

- 1) **通用型 AI 迈入流量入口争夺阶段，积累用户并探索生态的短期优先级大于变现。**从移动互联网时代的经验来看，C 端的搜索、社交类应用有很强的虹吸效应，最后是赢家通吃的局面，因此当前通用型 AI 应用的共识是零门槛接入快速占领用户入口，在拥有流量高地并具备护城河之后可以进一步探索广告、佣金等流量变现方式。我们也可以看到阿里、腾讯、字节为代表的互联网大厂正在进行通用型 AI 流量入口的激烈争夺。同时我们认为 AI Agent 的当前痛点是交互通道的打通，这对于有全生态的互联网大厂是核心优势，因此在用户入口争夺的下个阶段或能看到广泛的 Agent 应用在互联网大厂的爆发，带来中长期变现潜力。
- 2) **垂直赛道 AI 靠提供增值产品和体验，生产力场景以刚需驱动高商业转化。**首先全球维度观察，当前产生收入的 AI 应用中主要以偏向生产力的垂类赛道较为领先，细分垂直场景可以凭借提供的生产力改善、增值体验等等依靠订阅模式率先变现，同时由于生产力场景的刚需属性，其相较于娱乐类场景有更强的付费意愿，带来更高的商业转化。而在判断具体场景的落地速度中，我们提出了关注试错成本和可验证性两个维度，在用户对于 AI“幻觉”等错误接受度高且 AI 内容验证成本低的赛道落地更快，比如图像编辑、短（长）视频内容、广告营销、教育培训答疑、辅助招聘等等。

3、投资结论方面，我们认为互联网 AI 投资主要分为两条大的主线

主线 1：投资具备优质底层模型+算力（云）基础设施+丰富下游应用生态的互联网大厂，首推阿里、腾讯，建议关注百度。此类大厂当前正在通过 capex 引领新一轮国内 AI 基础设施建设周期（对标美国 23-24 年间 AI 进程），在优质算力（云）基础设施下担当中国 AI 科技周期启动的“卖铲人”。同时这些大厂下游已经具备完整的应用生态和丰富的用户基础，可以在已有的应用场景中无缝加入 AI 功能，依靠流量优势迅速启动。而大厂的多元业务生态非常有助于开启下一个阶段的 AI Agent 应用，将有望凭借其生态优势在 AI 的新一轮竞赛中取得领先身位，并中长期赋能核心业务的增长。

主线 2：投资卡位优质垂直场景+AI 产品布局充分+商业化落地确定性较高的“小而美”细分赛道龙头，首推美图、快手、哔哩哔哩、有道，建议关注好未来、BOSS 直聘。该类玩家原有业务卡位优质的细分垂直场景，如图像编辑、短（长）视频、教育和招聘等等，已经具备独有的数据、用户资源。随着 AI 技术平权等行业趋势，竞争从底层模型转向下游垂直场景应用，在 AI 端布局充分，并具备优质产品力的企业有望获得新增长曲线。同时我们也认为从 AI 应用商业化落地层面来看，垂直场景 AI 应用相较搜索、社交等通用 AI 更快，因为通用 AI 当前更聚焦于流量入口的争夺在变现层面并不激进，而细分垂直场景有望凭借提供的生产力改善、增值体验等更快实现财报端的兑现。

我们与市场不同的观点

我们对于互联网 AI 的相关底层模型、应用场景等等做了全方位的梳理和总结，在当前市场对于 AI 模型能力、应用场景仍有分歧的阶段，给予了我们明确的判断和观点。

我们从移动互联网的经验出发，探讨了 AI 应用的潜在发展趋势和竞争要素，并进一步讨论了其商业化的潜力和场景的落地节奏，提出了商业化节奏在通用型和垂直型分化的观点，并提供了试错成本和可验证性这一双维度指标作为场景落地难易的判断依据，均具备一定的差异性

互联网 AI 关键问题五问五答

Q1：如何评价互联网大厂 AI 大模型当前水平？

在技术层面，中国国内的头部模型已逐步接近国际顶尖水平，差距日益缩小。DeepSeek-R1 在后训练阶段广泛应用强化学习技术，在数学、代码和自然语言推理等领域的表现已与 ChatGPT o1 相当，Qwen2.5-Max 在主流权威基准测试中同样展现出全球领先的模型性能。我们认为，通用大模型领域的追赶效应愈发明显，国内互联网巨头正不断缩小与国际先进水平的差距，未来底层模型的能力有望达到持平状态。

目前中国在 AI 算力和生态建设方面仍存在不足，短期内可能限制大模型的训练和发展，但随着算法和架构的不断优化以及国产 AI 芯片性能的提升，我们认为中长期内中国 AI 大模型发展受算力制约影响有所减弱。

在战略层面，中国的发展路径更为注重垂直领域的落地与商业化。相较于海外聚焦基础技术的深耕研究，中国依托庞大的市场规模与多元化的应用场景，采取了以市场需求为驱动的路径并迅速在垂直领域实现技术落地和商业化。例如，在 AI+图像领域，美图专注图像美化赛道打造图像 AI 生产力工具；在 AI+视频领域，快手聚焦视觉大模型推出文生视频产品可灵；在 AI+教育领域，有道针对翻译与答疑等细分场景推出自研教育垂类大模型“子曰”。我们认为，中国的大模型发展战略更为注重市场应用需求，并通过市场力量加速技术进步，有助于 AI 生态的健康发展。

我们认为，DeepSeek R1 的出现或在技术、策略和定价维度影响国内 AI 模型发展，具体如下：

- 1) **技术层面：DeepSeek R1 的出现带动了高性价比架构的关注度，未来模型推理成本仍有下降空间。**DeepSeek R1 采用 MoE 架构，融合专家混合系统、改进注意力机制和优化归一化策略，在模型效率与算力间实现平衡，MoE 模式由此再度引发关注。2 月 12 日，字节豆包团队提出 UltraMem 架构，在 MoE 基础上有效解决推理阶段的高额访存问题，我们建议持续关注后续 MoE 架构的优化进展。
- 2) **策略层面：DeepSeek 的成功再次突显开源模式的潜力，大模型开源生态下各厂商或有望共同做大市场规模。**开源模型为中小厂商提供免费获取源代码或预训练模型的途径，显著降低了开发和参与门槛，并可借助开放社区的协同作用加速模型迭代。目前，阿里与腾讯的部分模型都已执行开源策略，百度也已宣布将在未来推出文心大模型 4.5 系列，并于 6 月 30 日起正式开源。我们看好后续大模型开源生态下的技术迭代进展。
- 3) **定价层面：我们认为国产模型价格竞争或将延续，低价下大模型的垂类应用或将深化。**受益于 MoE 架构对推理成本的压缩，DeepSeek 持续推行低价策略，此前已带动国产模型普遍性降价，R1 发布后 API 输入与输出价格仍处低位。我们认为高性价比有望成为国产模型在国际市场上的重要优势。

Q2: AI 应用的场景是降本还是增收？

我们认为 AI 在降本增效层面对大部分行业具备普适性，我们更看好在 AI 驱动下带来增收的行业：

- 1) **AI+图像：**AI 使用户体验和创作效率得到显著提升，通过降低用户创作门槛，吸引更多用户付费，提升整体付费率及 ARPU，为图像编辑应用带来了新的收入增长点。美图深耕图像工具类产品，AI 产品有望驱动付费率及全球化进程提速。
- 2) **AI+广告：**AI 帮助平台深度洞察用户数据，提高广告投放精准度，进而提升广告主投放 ROI，为广告平台带来更多收入。Applovin 产业链布局完善，数据闭环与 AI 算法升级形成有效协同，Meta 和腾讯有望受益于投放 ROI 提升带来的收入增长。
- 3) **AI+教育：**AI 有助于提升教育产品用户体验和个性化服务，推动付费用户数量增长。依靠教育公司自身数据积累优势，有效提升学习机和语言类学习应用使用体验，提升用户付费意愿。
- 4) **AI+游戏：**AI 有助于提升游戏互动性、智能化和个性化体验，推动玩法创新，提高人均游戏时长，带来更多潜在盈利空间。

Q3: 互联网 AI 商业化节奏如何理解？

传统移动互联网时代的变现路径主要有三条：广告、佣金和增值服务（订阅、单购），不同变现模式所依仗的商业基础各不相同，广告核心是流量，佣金是流量+履约，增值服务是提供差异化产品和体验。切换到 AI 应用时代，最快具备商业化基础的是靠 AI 提供的差异化产品和体验，因此增值服务这一商业模式落地节奏最快。而 AI 应用当前尚未形成新的流量高地且产品层面亦未成熟，因此在广告和佣金等商业模式上进展暂缓。**由此我们也能看到通用型 AI 和垂直型 AI 在商业化节奏上的分化**，基于此前提，我们对于 AI 应用的商业化落地节奏形成以下判断：

判断 1：通用型 AI 迈入流量入口争夺阶段，积累用户并探索生态的优先级大于变现。

Chatbot 类应用正在大批量开放免费使用。以 Chatgpt 为首的 Chatbot 类 AI 应用早期以付费订阅为主要的商业模式，但是随着免费开源、算力成本降低等，相关搜索、社交类的通用型 Chatbot AI 正在将大批量免费使用。25 年 2 月 13 日，OpenAI 首席执行官 Sam Altman 宣布免费 ChatGPT 将在标准智能设置下无限制使用 GPT-5 进行对话；Grok 也在 25 年 2 月 18 日发布几天后宣布免费；百度文心一言亦宣布自 25 年 4 月 1 日起全部免费使用。而国内的其他主流 AI Chatbot 应用 DeepSeek、豆包、元宝等等则从上线之初即免费使用。

通用型 AI 正在进入新的流量高地争夺阶段，积累用户并探索生态的优先级大于变现。从移动互联网时代的经验来看，C 端的搜索、社交类应用有很强的虹吸效应，最后是赢家通吃的局面，因此当前通用型 AI 应用的共识是零门槛接入快速占领用户入口，在拥有流量高地并具备护城河之后可以进一步探索广告、佣金等流量变现方式。

AI Agent 的核心痛点是交互通道的打通，这对于有全生态的互联网大厂是核心优势。由 Manus 带来的 AI Agent 趋势将驱动 AI 应用实现从思考到执行的全闭环，而在 AI Agent 的发展中，其核心痛点其实是交互执行中通道的开放和打通，除了公域的数据和执行之外，Agent 要完全发挥能力则需要相对私域的交互权限，这部分对于已有全生态的互联网大厂来说是核心优势，因此在用户入口争夺的下个阶段或能看到广泛的 Agent 应用在互联网大厂的爆发。

判断 2：垂直赛道 AI 靠提供增值产品和体验，生产力场景以刚需驱动高商业转化。

AI 垂类生产力赛道的商业化较为领先。当前 AI 应用商业化的主要方式为订阅付费，通过提供 AI 增值服务收取订阅费用。而当前产生收入的应用中主要以偏向生产力的垂类赛道较为领先，如 Chatgpt、微软 365 Copilot、Perplexity 等等，订阅费约为 10-20 美元/月不等。根据 AI 产品榜公众号，24 年已经有部分 AI 付费用户规模超千万，分别为微软 365copilot（5574 万）、百度文库（4000 万）、Canva（2200 万）、Chatgpt（1000 万）。集中于办公、文档、设计、搜索问答领域，我们认为偏向生产力场景因降本增效量化明确，偏刚需的特征能够有效驱动用户订阅付费转化。因此从商业化的节奏上或更快。

图表1：当前主流 AI 应用的订阅付费用户及 ARR

场景	产品	付费用户（万）	ARR（亿美元）
办公	微软 365 Copilot	5574	66.9
文档管理	百度文库	4000	-
搜索问答	ChatGPT	1000	27.0
	-api 调用	100	10.0
搜索问答	Perplexity	100	0.5
设计	Canva	2200	25.5
搜索问答	Claude	417	10
文档	Notion	333	-
编程	Cursor	66	-
聊天机器人	Character AI	53.6	0.3

资料来源：AI 产品榜公众号，华泰研究

判断 3：具体场景落地速度需关注试错成本和可验证性。

我们认为判断具体场景的落地速度需要关注其试错成本和可验证性两个维度：

- 1) 试错成本代表的是对于 AI 出现“幻觉”等等错误的接受度。比如 AI Chatbot 等等其本质是交流和提供咨询建议，用户对于结果有审查，因此试错成本低，此外还有 AI 的图像等内容创作，用户存在二次编辑，此类应用的落地和发展较快。而与之对比的则是试错成本较高的场景，比如自动驾驶等等，涉及到安全和责任划分等等层面，因此其实际落地需要缓慢推进。
- 2) 可验证性，代表的是 AI 提供的内容验证成本低，可以和用户迅速建立信任感。比如 AI 编程等等，其提供的代码可以迅速判断是否跑通，是否解决问题，验证环节短带来的就是交互的流畅性，可以和用户快速建立信任感。但是比如有些验证环节很长的场景，比如 AI 制定旅游攻略、剧本创作、提供资产配置的建议。其可验证性较低，用户需要更多的成本来判断建议的真伪和有效性，则落地的速度较慢。

基于此我们判断当前互联网板块落地较快的场景为图像编辑、短（长）视频内容、广告营销、教育培训答疑、辅助招聘等等。

Q4：互联网大厂在资本开支上的战略和进展如何？

在资本开支战略上，各互联网大厂多数展现出较为激进的投入态势，在 AI 基础设施如算力储备上加码投入，在投入的绝对规模上，阿里和字节或为领先者。

阿里：加大 AI 基础设施投资，未来三年资本开支或超过过去十年总和。据公司公告，3QFY25 公司资本开支达 318 亿元，同比增长 258.8%。据公司 2025 年 2 月 24 日公告，未来三年阿里将投入超过 3,800 亿元，用于建设云和 AI 硬件基础设施，总额超过过去十年总和，以突出公司对 AI 驱动增长的聚焦及作为全球领先云计算供应商的定位。据 VA 一致预期，FY25-27 阿里资本开支将达 847.7 亿/1,076.3 亿/1,123.0 亿人民币，同比分别 +157.4%/+27.0%/+4.3%。结合公司近期发布多款大模型和与苹果达成合作等事项，我们认为阿里云将以较为积极的姿态参与到新一轮的 AI 产业变革中。

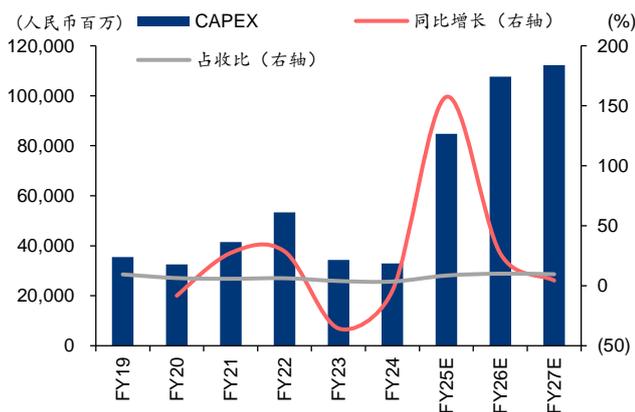
腾讯：持续投入资源进行算力储备。据红星新闻报道，腾讯会持续投入资源进行算力的储备，助力各个 BG 拥抱大模型的产品化落地场景，目前包括微信、QQ、输入法、浏览器等产品都将推出 AI 智能体，游戏、微信读书、腾讯视频等产品也将基于混元做更多 AI 探索。据 VA 一致预期，2024-2026 年腾讯资本开支将达 422.6 亿/433.0 亿/453.1 亿人民币，同比分别 +76.9%/+2.5%/+4.6%。

快手：坚定执行人工智能战略。据公司 2Q24 业绩会，未来公司将坚定执行人工智能战略，侧重研发更适配业务场景的专项语言模型，大规模降低模型的总推理成本，同时探索“可灵”与 B 端（企业）合作变现的可能性以实现商业化变现规模。据 VA 一致预期，2024-2026 年快手资本开支将达 54.9 亿/64.0 亿/71.1 亿元人民币，同比分别 +12.1%/+16.4%/+11.2%。

字节：加码投入算力建设。据财联社报道，字节跳动 2025 年的资本支出将超过 1500 亿元，主要用于人工智能。据英国《金融时报》报道，字节跳动计划在 2025 年投资超过 120 亿美元用于 AI 基础设施建设，其中 400 亿元人民币（约 55 亿美元）将用于购买中国 AI 芯片。这一支出是 2024 年的两倍，公司还计划在海外投资约 68 亿美元，以使用先进的 Nvidia（NVDA.US）芯片来增强基础模型训练能力。

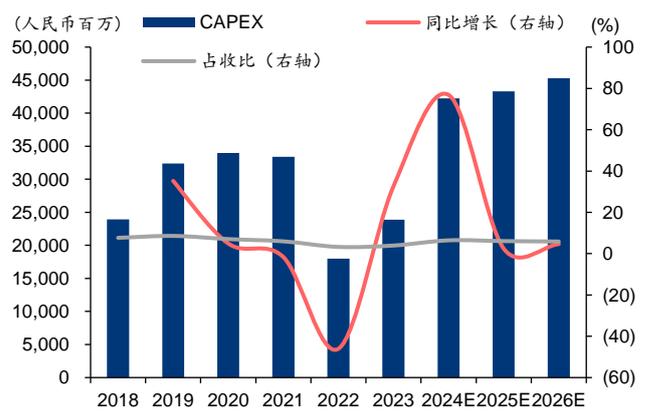
百度：未来资本开支可能增加。据公司 1Q24 业绩会，百度正在从以互联网为中心转向人工智能优先，积极推进用文心大模型重构 ToC 和 ToB 业务，其未来资本开支可能增加，以支持网络基础设施的扩展和改进，以及额外办公楼和云计算数据中心的建造。据 VA 一致预期，2025-2026 年百度资本开支将达 101.4 亿/99.5 亿人民币，同比分别 +24.7%/-1.9%。

图表2：阿里 Capex 投入情况



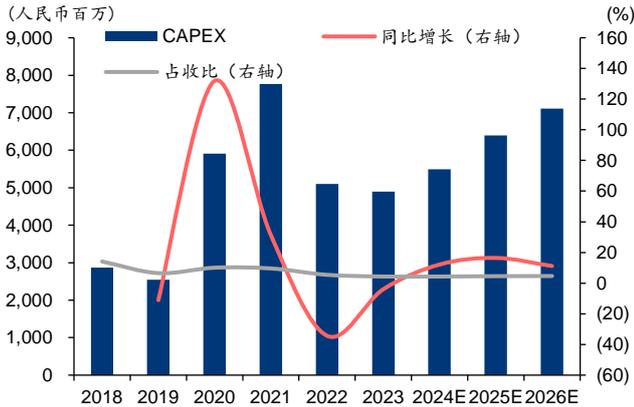
资料来源：公司公告，Visible Alpha，华泰研究

图表3：腾讯 Capex 投入情况



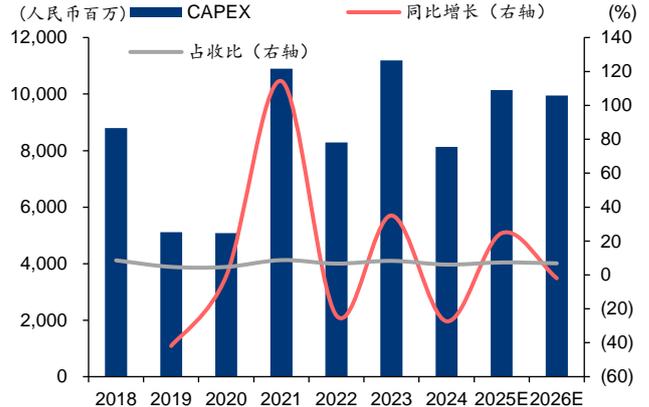
资料来源：公司公告，Visible Alpha，华泰研究

图表4: 快手 Capex 投入情况



资料来源: 公司公告, Visible Alpha, 华泰研究

图表5: 百度 Capex 投入情况



资料来源: 公司公告, Visible Alpha, 华泰研究

Q5: AI 大模型对行业壁垒的影响

AI 大模型或对各行业竞争格局产生分化式影响。一方面, 在数据与算力密集型领域 (如医疗诊断、金融风险评估), 巨头企业凭借海量数据、算力资源、人才储备及合规能力构建护城河, 有望在新一轮竞争中取得先发优势, 推动市场集中度进一步提升。另一方面, 在技术工具化与长尾需求主导的行业 (如内容创作、农业/教育科技), 开源模型和轻量级 API 的发展有望显著降低垂直场景创新门槛, 赋能中小应用实现差异化突围。整体来看行业属性与技术民主化进程将成为平衡资源分配、决定竞争壁垒“强弱再分配”的关键变量, 最终形成“强者恒强、细分突围”的产业格局。

图表6: AI 大模型对于行业竞争壁垒的影响

	增强竞争壁垒 (大型企业主导)	缩小竞争差距 (中小型企业或新玩家突围)
主导领域	<ul style="list-style-type: none"> 数据密集型 (医疗、金融) 算力密集型 (自动驾驶训练) 法规敏感行业 (生物科技、公共安全) 	<ul style="list-style-type: none"> 工具驱动型 (内容创作、客户服务) 细分场景 (农业、教育技术) 开源生态活跃的领域 (软件开发、资源)
核心驱动因素	<ul style="list-style-type: none"> 丰富数据储备 算力资源 满足合规要求的能力 	<ul style="list-style-type: none"> 开源模型与 API 易用性提升 垂直场景创新 需求长尾化带来的差异化机会
最终结果	行业集中度提升	激发竞争多样性

资料来源: 多知、有道、华泰研究

AI 模型：算力规模飞速增长，成本效应未来可期

核心逻辑总结：DeepSeek 引领中国 AI 大模型在模型性能、成本效用上实现突破，未来阿里字节等大厂或相继接力，合力夯实新一轮 AI 大模型科技竞赛中的中方竞争力。我们认为以 DeepSeek 的出现为分界线，AI 大模型的发展阶段可以划分为海外大厂保持代际领先与中国厂商迎头赶上两个阶段。在第二阶段，算力层面的 **Scaling law** 仍然奏效，但不再是通往“罗马”的唯一路径，这亦为中国 AI 大模型走出实验室研究、走向日常应用带来了产业级别的机遇。新阶段，AI 大模型的核心生产要素中“数据”的差异化价值有望进一步突出，在本就繁荣的消费互联网应用层或将迎来产品的百花齐放和商业化落地的全新机遇，在引领全球消费互联网商业生态迈向新格局、新机遇的同时，也为国内以阿里、腾讯、百度等为代表的“卖铲人”提供云服务业绩的增厚。对于模型层，算力层面差距的收窄或将进一步推动基础大模型服务提供商者未来更多聚焦与应用侧的适配链接，在基础能力层演化出差异化格局。

目前 AI 发展的主流战略可分为两类：以字节为代表的厂商侧重算力规模，而以腾讯、阿里为代表的厂商则更关注性价比；鉴于大模型更新迭代存在追赶效应，我们预期 26-27 年模型层面的性能差异或将逐步缩小，同时伴随 LLM 使用成本显著降低，后期竞争关键在于产品能力、销售获客渠道、大模型如何赋能自身业务。

从商业化角度看，百度更早开始收取会员费用，文心单月会员价格为 49.9 元，对比海外 OpenAI Plus 版本包月 20 美元（不含税），国内互联网其他大厂尚未开启会员收费模式。从降价趋势看，自 24 年 5 月 DeepSeek 发布 V2 版本后，引起阿里腾讯字节三家竞相宣布降低主力模型的 API 收费。从大模型收费看，百度文心一言 API 调用费仍然保持高位，阿里腾讯和字节次之并低于百度一个数量级，DeepSeek 则具备更高的价格优势。

图表7：互联网公司大模型参数

	Baidu	Allbaba	Tencent	Tik Tok	DeepSeek
名称	文心一言 ERNIE	通义千问 QWEN	混元 Turbo	豆包 Doubao	DeepSeek
调用规模	日均15亿次	API调用量过亿	日均tokens调用量超千亿（24年7月数据）	日均tokens调用量超4万亿	日均tokens调用量超30亿
参数规模	万亿级别	720亿（QWEN 72B）	万亿级别	1300亿	6710亿
商业化	单月会员：59.9元 连续包月：49.9元 白银联合会员：99元/月	暂无	暂无	暂无	暂无
API调用	ERNIE 4.0 Turbo 输入价格（百万token）：20元 输出价格（百万token）：60元 ERNIE 4.0 输入价格（百万token）：30元 输出价格（百万token）：90元	qwen2.5-72b-instruct 输入价格（百万token）：4元 输出价格（百万token）：12元 qwen-plus 输入： 输入价格（百万token）：0.8元 输出价格（百万token）：2元	混元Large (389B) 输入价格（百万token）：4元 输出价格（百万token）：12元	针对1.5-pro-32K、1.5-pro-256K、1.5-lite-32K: 输入价格（百万token）：0.8/5/0.3元 输出价格（百万token）：2/9/0.6元	DeepSeek-V3(671B):（百万token） 输入价格：缓存未命中和命中价格分别为2元和0.1元 输出价格：2元 DeepSeek-R1(660B):（百万token） 输入价格：缓存未命中和命中价格分别为4元和1元 输出价格：16元
核心技术	<ul style="list-style-type: none"> 依托开源深度学习平台PaddlePaddle（飞浆）和昆仑AI芯片 Transformer算法 图像检索增强生成技术（IRAG） 提供多样化的API接口和SDK开发包 	<ul style="list-style-type: none"> 基于Transformer框架 采用RoPE(Rotary Positional Embedding)为位置编码，并使用FP32精确度的逆频率矩阵 采用开源快速BPE分词器tiktoken，并以c100k为基础词库 	<ul style="list-style-type: none"> 探索MoE架构的Scaling Law，并引入专家特化的学习率适配策略，提升专家利用率与模型稳定性。 强化Post-Train阶段数学、逻辑推理与代码分类能力，数学与代码任务表现提升超10%。 推出企鹅卷轴（PenguinScrolls）数据集，推动长文本处理技术演进。 	<ul style="list-style-type: none"> 采用稀疏MoE架构，预训练阶段即以更小参数激活模型，实现超越Llama3.1-405B等超大稠密模型的性能。 通过模型结构与训练算法优化，将MoE模型的性能杠杆提升至7倍，远超业界普遍不到3倍的水平。 不依赖蒸馏或外部模型数据，团队自建数据生产体系，确保数据质量与模型自主进化能力 	<ul style="list-style-type: none"> Multi-head Latent Attention (MLA)、DeepSeekMoE架构以及无额外损失的负载均衡策略。 多项工程优化贯穿了流水线并行、通信优化、内存管理和低精度训练等多个方面。
最新布局	<ul style="list-style-type: none"> 小度AI眼镜：市场上首款内置中文大模型的原生智能设备 百智AI异构计算平台4.0 更多地将gen-AI内容放在其主要广告库存 	<ul style="list-style-type: none"> 基于RSIC-V架构的智能IoT芯片玄铁，以及AI语音FPGA芯片Euroboros，并计划在未來十年内，利用AI技术推动业务增长。 计划通过AI技术推动云计算业务的增长，期望AI技术带来突破性的用户体验和业务模式。 	<ul style="list-style-type: none"> 腾讯混元目前主要开源参数权重，不涉及数据和代码，但接下来会把训练代码开源出来，同时开源评估期 混元3D 2.0发布：AI赋能3D建模。 	<ul style="list-style-type: none"> 在模型层发布了视觉理解模型和音乐模型4.0，实现通用模型pro新版本迭代。 在平台层围绕大模型落地发布火山方舟实验室、扣子1.5和HiAgent1.5。 布局AI玩具、AI眼镜等。 	

资料来源：公司公告、华泰研究

图表8：国内主要 LLM 模型对比

公司	模型	LLM 创新	测试结果	发布日期	是否开源
字节跳动	Doubao-1.5-pro	7 倍效率提升的稀疏 MoE 架构	部分测试结果显示：优于 GPT-4o 和 Claude-3.5-Sonnet	2025-01	否
阿里巴巴	Qwen2.5-Max	优化全局批次负载平衡损失的 MoE 架构	部分测试结果显示：表现优于开源的 DeepSeek V3 和 Llama-3.1-405B	2025-01	否
腾讯	Hunyuan large	探索 MoE 架构的 ScalingLaw，并引入专家特化的学习率适配训练策略	优于 Llama 3.1 70B 和 405B	2024-11	是
百度	ERNIE 4.0 Turbo	开创性地将大数据预训练与多源丰富知识相结合	ERNIE 4.0 能力接近 GPT-4	2024-06	否
月之暗面	Kimi-k1.5	通过强化学习与监督微调结合进行训练	部分测试结果显示：优于 GPT-4o 和 Claude-3.5-Sonnet	2025-01	是
MiniMax	MiniMax-01	优化的 MoE 架构，具有线性注意力机制，尤其适用于长文	部分测试结果显示：优于 GPT-4o 和 Claude-3.5-Sonnet	2025-01	是
DeepSeek	DeepSeek-R1	后训练阶段大规模采用强化学习，无需监督微调	推理任务中结果和 OpenAI o1-1217 相近	2025-01	是

资料来源：公司官网，华泰研究

腾讯混元：从内部业务自用起步，强化云业务技术壁垒。腾讯 LLM 研发起步略晚于其他大型厂商，但近期技术端差距有所收窄。腾讯同样坚持开源策略，引入 20 余种开源通用模型，同时部分模型也处于开源状态。在自用场景方面，混元大模型接入腾讯 700 多个业务场景，改善内部人员工作效率，混元 3D 引擎在 24 年广泛使用后，也于 25 年对外开放；在腾讯云方面，混元大模型也构建了腾讯云 MaaS 坚实的基础，客户可以将混元用作基础模型，在公有云上进行精细调整，也可以在腾讯云上通过前端交互 API 调用体验。

尽管腾讯大模型研发起步相对较晚，但 2024 年混元技术取得较快突破，腾讯实现全链路自研，包括底层算力网络、Angel 机器学习平台、上层模型和应用。腾讯全新发布的混元 Turbo 有显著能力提升，推理成本降低 50%，解码速度改善 20%。腾讯云智算 AI infra 从机器上架到开始训练仅需 1 天，千卡单日故障数仅仅为 0.16，网络通信时间占比也仅 6%。TI 平台上线多模态数据标注，首创基于 Schema 的标注方法，支持 4 大主流场景全部细分任务。混元 large 模型开源五天内即在 HuggingFace 百万模型中脱颖而出，登趋势榜榜首。

阿里：专注于打造 AI 基础设施和大模型能力的落地应用。我们认为阿里云业务的核心优势来自：1) 全面的云服务基础能力和规模效应，2) 消费交易场景下积累的行业 know-how，3) 较高的用户数据基础壁垒。自 2023 年提出“AI 驱动”战略以来，公司一方面投资于 AI 基建，一方面坚持开源和生态建设。公司计划在 B 端和 C 端用户侧逐步推进 AI 应用，其自身多元化的业务正成为应用端最先落地的场景。据 2024 年 9 月阿里云栖大会，在 HuggingFace 和魔搭社区中，已有超过 5 万个基于通义千问模型开发的衍生模型，位于全球第二，超过国内其他开源模型，我们认为基于通义千问模型的生态系统正在逐步成熟。

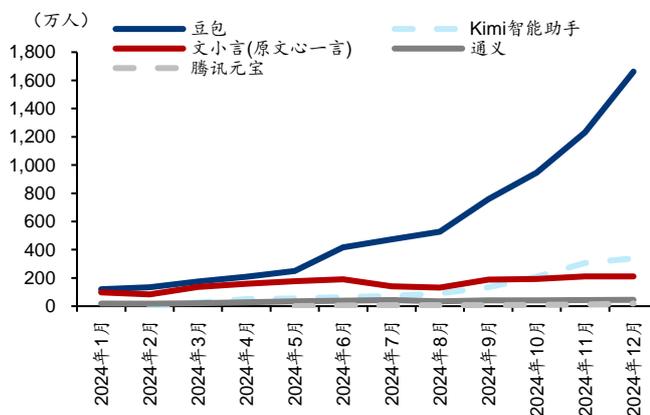
三款旗舰大语言模型满足不同类型用户需求。据阿里云官网，截止 2025 年 2 月 4 日，阿里巴巴旗舰模型主要为通义千问-Max、通义千问-Plus 和通义千问-Turbo 三款。Max 为通义千问系列效果最好的模型，适合复杂、多步骤的任务；Plus 能力均衡，推理效果、成本和速度介于通义千问-Max 和通义千问-Turbo 之间，适合中等复杂任务；Turbo 为通义千问系列速度最快、成本很低的模型，适合简单任务。

百度：打造高效 AI 开发能力，智能体和产业应用为发力点。百度是国内少数几家实现了从芯片到框架再到模型及应用的全栈布局的企业之一，有助于利用自身较为强大的 AI 能力和云计算资源，为用户提供一站式解决方案。公司 AI 基础设施包括用于大模型和应用开发的千帆平台，以及为客户提供大模型所需的算力服务的百舸平台。与同行相比，百度在大模型商业化上更倾向于采用闭源逻辑。在应用端，百度着力布局两大 AI 应用方向：1) 智能体，公司判断这将是 AI 应用未来的最主流形态，类别包括公司类、角色类、工具类、行业类；2) 产业应用，涵盖制造、能源、交通、政务、金融、汽车、教育、互联网等众多行业。

百度多款主力大模型已可免费调用。据百度智能云官网，在大模型调用费用上，百度针对较为先进的旗舰大模型 ERNIE 4.0 Turbo、ERNIE 4.0 和 ERNIE 3.5 采用收费制度，输入/输出价格最低低至 0.0008/0.002 元每千 tokens，而主力大模型 ERNIE Speed、ERNIE Lite、ERNIE Tiny 均已免费使用，其中 ERNIE Speed 是百度 2024 年发布的自研高性能大语言模型，通用能力优异，适合作为基座模型进行精调，可更好地处理特定场景问题，同时具备较佳的推理性能；而 ERNIE Lite 是百度自研的轻量级大语言模型，兼顾优异的模型效果与推理性能，适合低算力 AI 加速卡推理使用。

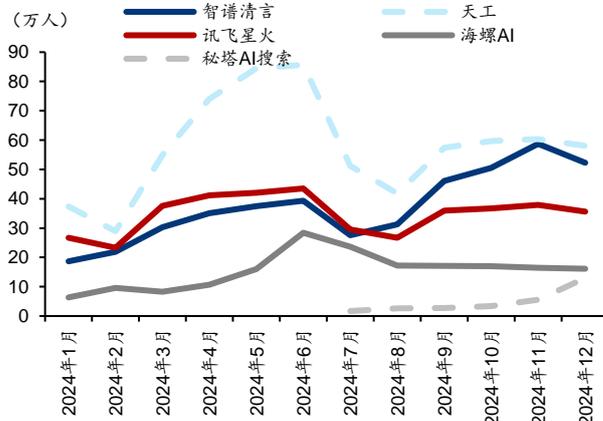
国内通用 AI APP 市场加速分化，用户规模层面豆包已确立领先地位，而 DeepSeek 在海外市场仍具成长空间。国内方面，豆包 DAU 持续扩大领先优势，QuestMobile 数据显示，2024 年 12 月其 DAU 已达 1662 万，远超第二梯队的 Kimi 智能助手与文小言。海外方面，DeepSeek R1 上线后增长迅猛，但与行业龙头仍存差距。SensorTower 数据显示，截至 2 月 1 日，DeepSeek 海外移动端 DAU 达 1300 万，而 ChatGPT 达 5637 万；Similarweb 数据显示，截至 1 月底，DeepSeek 网站端周均 DAU 为 1089 万，ChatGPT 则为 7001 万。

图表9：国内部分通用 AI APP DAU



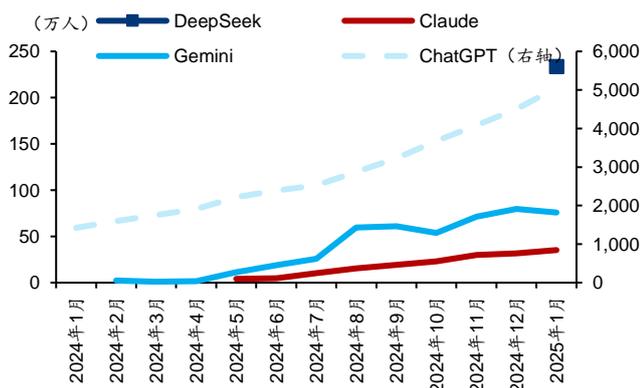
资料来源：QuestMobile, 华泰研究

图表10：国内部分通用 AI APP DAU



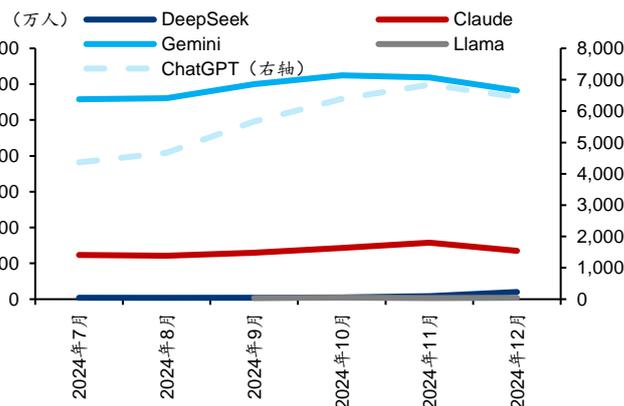
资料来源：QuestMobile, 华泰研究

图表11：海外部分通用 AI APP DAU



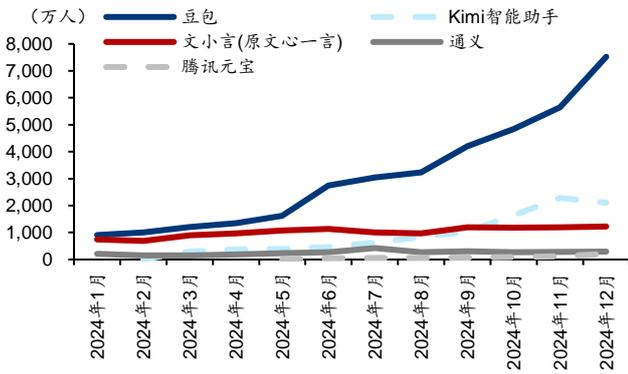
资料来源：SensorTower, 华泰研究

图表12：海外部分通用 AI 网站版 DAU



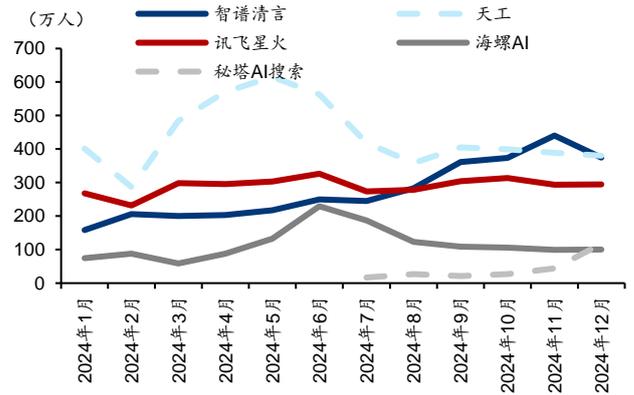
资料来源：Similarweb, 华泰研究

图表13：国内部分通用 AI APP MAU



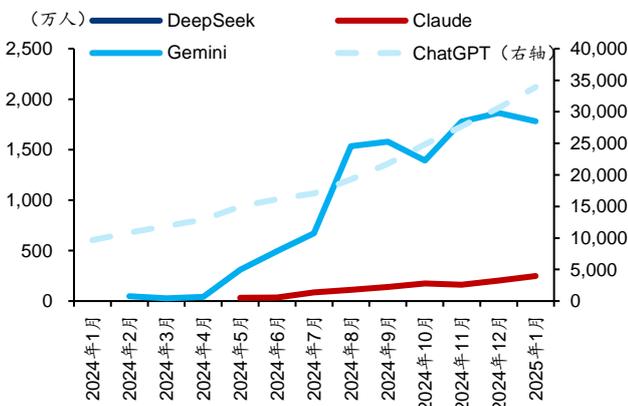
资料来源：QuestMobile, 华泰研究

图表14：国内部分通用 AI APP MAU



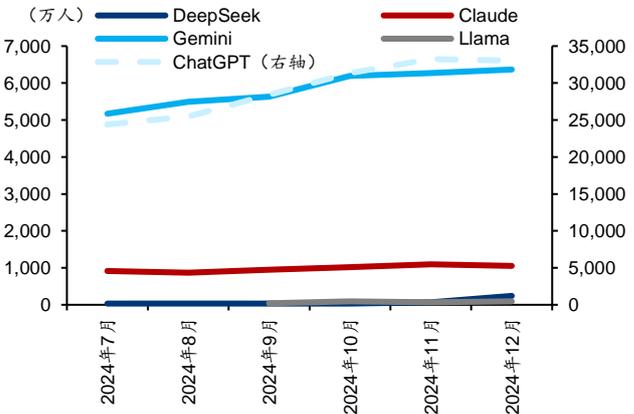
资料来源：QuestMobile, 华泰研究

图表15：海外部分通用 AI APP MAU



资料来源：SensorTower, 华泰研究

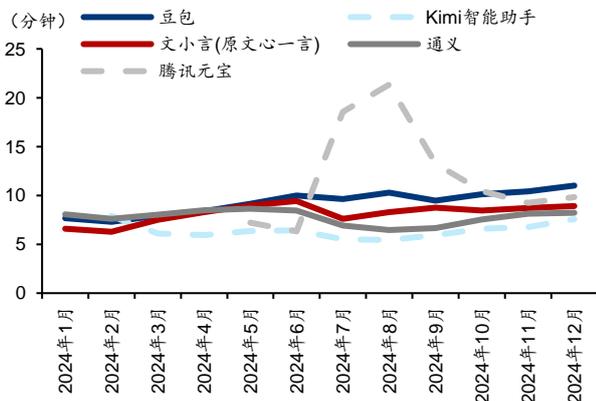
图表16：海外部分通用 AI 网站版 MAU



资料来源：Similarweb, 华泰研究

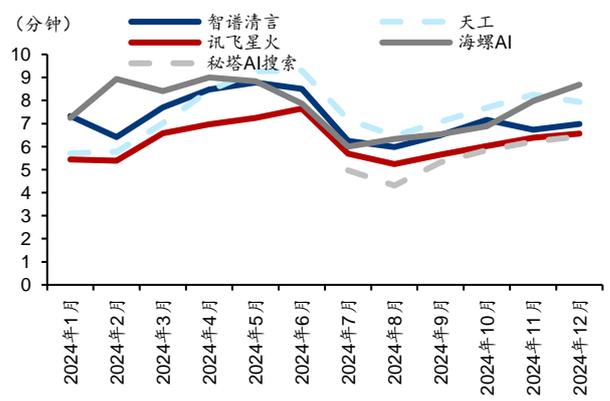
用户行为层面，国内通用 AI APP 的人均使用时长较为集中且仍处低位，海外市场 DeepSeek 加速追赶头部应用，R1 有望推动其进一步渗透。国内方面，各应用在人均单日使用时长和留存率上的差距较小，豆包保持微弱领先，QuestMobile 数据显示，2024 年 12 月其 DAU/MAU 达 22%。海外方面，ChatGPT 在人均使用时长上仍具优势，但 DeepSeek 正逐步缩小差距，留存率有所回落，主因部分用户受热度驱动尝试后流失。

图表17：国内部分通用 AI APP 人均单日使用时长



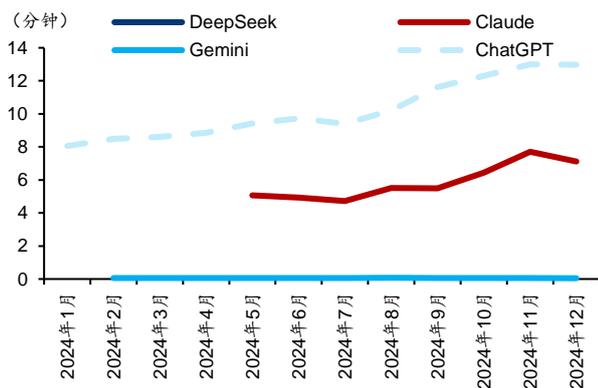
资料来源：QuestMobile, 华泰研究

图表18：国内部分通用 AI APP 人均单日使用时长



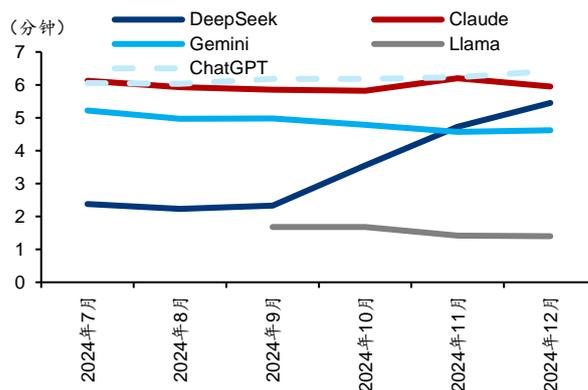
资料来源：QuestMobile, 华泰研究

图表19: 海外部分通用 AI APP 人均单日使用时长



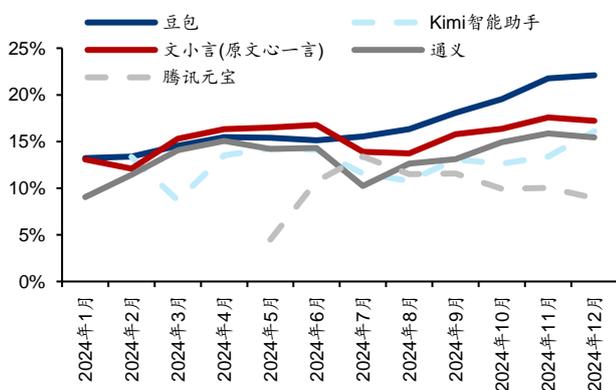
资料来源: SensorTower, 华泰研究

图表20: 海外部分通用 AI 网站版人均单日使用时长



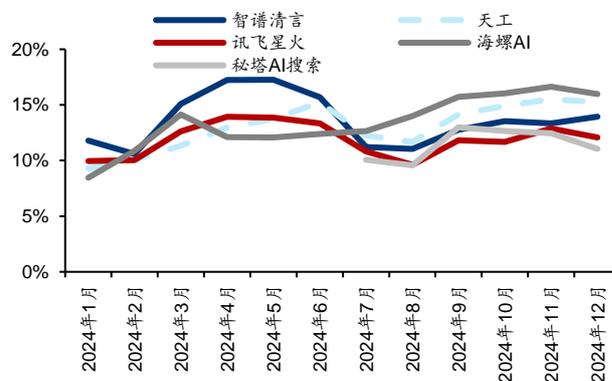
资料来源: Similarweb, 华泰研究

图表21: 国内部分通用 AI APP DAU/MAU



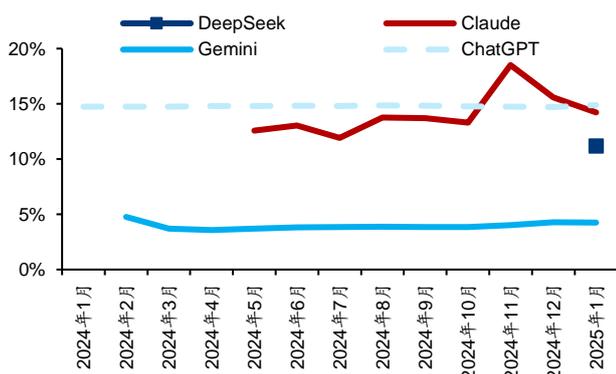
资料来源: QuestMobile, 华泰研究

图表22: 国内部分通用 AI APP DAU/MAU



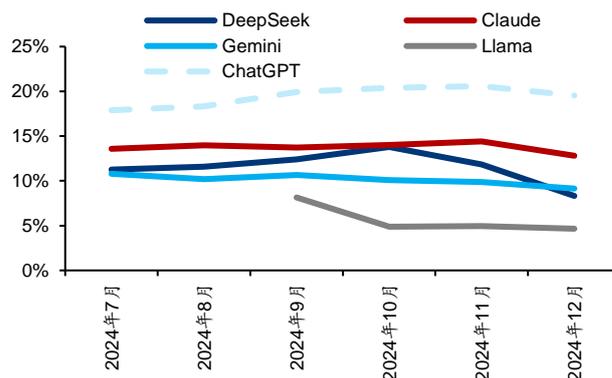
资料来源: QuestMobile, 华泰研究

图表23: 海外部分通用 AI APP DAU/MAU



资料来源: SensorTower, 华泰研究

图表24: 海外部分通用 AI 网站版 DAU/MAU



资料来源: Similarweb, 华泰研究

腾讯混元: 全面拥抱开源生态, 模型性能大幅提升

腾讯混元: 从内部业务自用起步, 强化云业务技术壁垒。 腾讯混元技术于2024年取得较快突破, 同时实现全链路自研。腾讯 LLM 研发起步略晚于其他大型厂商, 但近期技术端差距有所收窄。在自用场景方面, 混元大模型接入腾讯 700 多个业务场景, 改善内部人员工作效率, 混元 3D 引擎在 24 年广泛使用后, 也于 25 年对外开放; 在腾讯云方面, 混元大模型也构建了腾讯云 MaaS 坚实的基础, 客户可以将混元用作基础模型, 在公有云上进行精细调整, 也可以在腾讯云上通过前端交互 API 调用体验。

腾讯混元技术于 2024 年取得较快突破。一方面,腾讯实现全链路自研,包括底层算力网络、Angel 机器学习平台、上层模型和应用。腾讯全新发布的混元 Turbo 有显著能力提升,推理成本降低 50%,解码速度改善 20%。根据 SuperCLUE 报告,腾讯混元在 8 项核心任务上排名国内第一。腾讯在底层算力网络、Angel 机器学习平台、上层模型和应用各链路均实现较好的技术自研突破。

混元大模型全面拥抱开源: 腾讯将开源协同作为公司的核心技术战略,积极推动内外部技术开源共享,大模型商店引入 20 余种开源通用模型,同时部分模型也处于开源状态。未来大模型各版本也将进一步开源,促进技术生态繁荣发展。混元 large 模型开源五天内即在 HuggingFace 百万模型中脱颖而出,登趋势榜榜首, GitHub Star 达 1300。混元 3D 模型 GitHub Star 达到 2500, 开源两周内成为 HuggingFace 下载增速最快的 3D 模型。

腾讯持续投资基础设施,致力服务云业务: 腾讯云智算 AI infra 从机器上架到开始训练仅需 1 天,千卡单日故障数仅仅为 0.16,网络通信时间占比也仅 6%。TI 平台上线多模态数据标注,首创基于 Schema 的标注方法,支持 4 大主流场景全部细分任务。混元 large 模型开源五天内即在 HuggingFace 百万模型中脱颖而出,登趋势榜榜首。围绕着企业训练大模型和应用大模型的需求,腾讯提供 AI infra、自主可控的大模型以及贴近场景的智能应用。在这一过程中,腾讯不断提升模型性能的同时,也在不断降低模型使用的门槛,通过封装好的 PaaS 产品来让企业构建面向具体场景的应用,包括智能客服和营销工具等。

图表 25: 混元大模型加速大模型能力落地场景



资料来源: 腾讯混元, 华泰研究

图表 26: 腾讯大模型全面拥抱开源



资料来源: 腾讯混元, 华泰研究

腾讯 AI 组织架构简介

腾讯混元大模型和腾讯元宝的团队,主要由 CSIG 部门和 TEG 部门的部分核心成员抽调组成,该组织架构目的是为了让 AI 全面和腾讯现有业务更好地结合:

1) 混元大模型团队: 负责大模型的核心研发与技术迭代,核心负责人包括 TEG 副总裁蒋杰、腾讯云副总裁刘煜宏,统筹战略与技术推进。模型推理方向由资深 GPU 优化专家刘凯负责, NLP 技术中心副总监张锋主导混元应用,计算机视觉专家芦清林负责文生图, IEEE Fellow 刘威担任技术负责人。

2) 腾讯元宝团队: 由首席科学家张正友领导,产品管理由 AILab 副主任俞栋、AI 平台部助理总经理王迪及战略发展部总监刘田共同负责。

腾讯混元大模型和腾讯元宝的团队中,从 CSIG 和 TEG 部门抽调的核心成员,均来自不同事业群,负责该事业群的 AI 研发项目。例如微信 AI 团队成员也有抽调进入,主要聚焦微信生态内的 AI 应用落地。TEG 部门全称为技术工程事业群,旗下拥有 AI Lab 与 Robotics 实验室,前者聚焦 AI 基础研究与应用探索,后者专注于机器人技术; CSIG 部门全称为云与智慧产业事业群,旗下涵盖腾讯优图与腾讯云,前者深耕于计算机视觉领域,后者既是对外 AI 能力输出平台,也提供内部云基础设施支持。

图表27：腾讯混元大模型团队结构一览

团队名称	职责方向	姓名	职位及背景	
混元大模型	混元模型负责人	蒋杰	TEG 数据平台部负责人，中国人工智能产业发展联盟(AIIA)副理事长	
	混元模型负责人	刘煜宏	腾讯云副总裁、腾讯数据平台部副总经理	
	模型推理方向	刘凯	混元模型推理方向负责人，腾讯高级工程师，10 年以上 GPU 高性能优化经验	
	混元应用	张锋	混元应用负责人，腾讯 AI 资深专家研究员，NLP 技术中心副总监	
	文生图	芦清林	混元文生图负责人，专注计算机视觉算法研究十年，从事文字识别、语义分割、视频内容理解等研究	
腾讯元宝	技术	刘威	混元技术负责人，IEEE Fellow、腾讯杰出科学家，曾任 IBM 沃森研究中心研究科学家	
	负责人	张正友	腾讯首席科学家，前 AI lab 主任，RobotiX 主任	
	PM		俞栋	腾讯 AI lab 副主任
			王迪	腾讯 AI 平台部助理总经理
			刘田	腾讯战略发展部总监
			李迪夫	腾讯 AI lab 总监
	组长		康战辉	腾讯 AI 平台部，搜索技术
			Ying shan	腾讯 ARC lab 杰出科学家
			陶阳宇	腾讯数据平台部总监
			王涛	或为腾讯云副总裁
			王波	或为腾讯公司副总裁，主游戏业务
			Daniel Wang	暂无公开信息
			卢山	腾讯集团高级执行副总裁，技术工程事业群总裁
	Sponsor		汤道生	腾讯集团高级执行副总裁，CSIG CEO
			蒋杰	暂无公开信息
		吴运声	腾讯云副总裁，腾讯云智能负责人，优图实验室总经理，兼任腾讯 CSIG 政企业务线产研的负责人	
		周颖	微信技术总监	
		崔晓春	腾讯互动娱乐副总裁	
		郝小虎	腾讯 PCG 副总裁，信息平台与服务线负责人，兼任事业线搜索与商业	

资料来源：腾讯官网，华泰研究

腾讯管理层定位：“以 AI 为核心，打造智能云，赋能智慧产业”

腾讯 CEO 马化腾表示，“AI 是腾讯的核心战略之一，我们将持续加大 AI 的研发投入，推动 AI 的技术创新和应用落地，为用户和企业提供更好的智能服务和体验。”马化腾认为，AI 的发展将带来三个方面的机遇：一是 AI 将赋能各行各业，提升社会效率和生活品质；二是 AI 将催生新的业务模式和价值创造，促进经济增长和社会进步；三是 AI 将激发新的人机交互和内容创作，丰富人类的文化和娱乐。腾讯云总裁唐达在 2023 年腾讯云 AI 峰会上表示，腾讯云 AI 的战略是“以 AI 为核心，打造智能云，赋能智慧产业”，其目标是“为各行各业提供全面的 AI 服务和平台，帮助企业和机构实现智能化的转型和升级。”唐达认为，AI 的未来将需要三个方面指引：一是技术指引，即通过不断的技术创新，提升 AI 的性能和效率，降低 AI 的门槛和成本，让 AI 更加普惠和可靠；二是场景指引，即通过深入的场景探索，挖掘 AI 的需求和价值，打造 AI 的应用和解决方案，让 AI 更加贴合有效；三是规则指引，即通过合理的规则制定，规范 AI 的行为和责任，保障 AI 的安全和公平，让 AI 更加可控可信。

腾讯混元整体定价处于中低水平，并以灵活的层级定价策略精准触及多元化市场需求。混元大模型采用 MoE 架构，实现了计算资源的高效分配，在降低 API 成本的同时保障了输出质量，使得整体定价可控制在较低水平，同时腾讯采用层级化定价策略灵活应对不同市场需求，确保渗透各类用户价值。此外，腾讯在模型应用上实行差异化布局，涵盖通用文生文、翻译、代码及角色扮演等领域，布局相对广泛。

图表28：腾讯部分模型及价格

模型类型	产品名称	能力与特征	免费额度	API 费用 (元/千 tokens)
通用文生文	hunyuan-turbo	异构 MoE 结构，参数规模依然保持万亿级。 效果在多个基准测试上对标 GPT-4o	共享消耗共 100 万 tokens，有效期 1 年	输入 0.015 输出 0.05
	hunyuan-large	开源领域参数规模最大、效果最好的 MoE 模型		输入 0.004 输出 0.012
	hunyuan-large-longcontext	当前混元模型中效果最优版本，长文能力优秀。		输入 0.006 输出 0.018
	hunyuan-standard	采用更优的路由策略，缓解负载均衡和专家趋同问题。长文方面，大海捞针指标达到 99.9%。 性价比相对更高，实现对长文本输入的处理。		输入 0.0008 输出 0.002
	hunyuan-standard-256K	MOE-256K 在长度和效果上进一步突破，极大的扩展了可输入长度。		输入 0.0005 输出 0.002
机器翻译	hunyuan-translation	支持 15 种语言互译，基于多场景翻译评测集自动化评估 COMET 评分。		输入 0.025 输出 0.075
	hunyuan-translation-lite	混元翻译模型支持自然语言对话式翻译，支持 15 种语言互译。		输入 0.005 输出 0.015
角色扮演	hunyuan-role	角色扮演模型，精调训练推出的角色扮演模型。		输入 0.004 输出 0.008
Function Call	hunyuan-functioncall	混元最新 MOE 架构 FunctionCall 模型，经过高质量的 FunctionCall 数据训练，上下文窗口达 32K，在多个维度的评测指标上处于领先。		输入 0.004 输出 0.008
代码生成	hunyuan-code	代码生成模型，经过 200B 高质量代码数据增训基座模型，迭代半年高质量 SFT 数据训练，上下文长窗口长度增大到 8K，五大语言代码生成自动评测指标上位居前列。		输入 0.004 输出 0.008
混元图生文	hunyuan-turbo-vision	混元新一代视觉语言旗舰大模型，采用全新的混合专家模型 (MoE) 结构。		输入 0.08 输出 0.08
	hunyuan-vision	混元最新多模态模型，支持图片+文本输入生成文本内容。支持图片基础识别，图片内容创作，图片多轮对话，图片分析推理，图片知识问答。		输入 0.018 输出 0.018
通用文生文	hunyuan-lite	升级为 MOE 结构，上下文窗口为 256k，在 NLP，代码，数学，行业等多项评测集上领先众多开源模型。	免费使用	无
向量化模型	hunyuan-embedding	新增 Embedding 接口：可以将文本转化为高质量的向量数据。	100 万 tokens，时效 1 年	输入 0.0007 输出 0.0007
智能体平台	腾讯元器	用户无需编程经验即可轻松创建属于自己的智能体。	暂无	输入 0.10 输出 0.10

资料来源：腾讯云，华泰研究

阿里巴巴：AI 驱动云计算业务未来

在 AI 发展上，阿里巴巴专注于打造 AI 基础设施新范式和 AI 模型能力的落地应用。我们认为全面的云服务基础能力与规模效应优势，和其在消费交易场景下积累的行业 know-how 与用户数据基础是阿里巴巴发展 AI 业务的核心差异化优势。自 2023 年提出“AI 驱动”战略以来，阿里巴巴一方面投资于 AI 基建，一方面在 B 端和 C 端用户侧逐步推进 AI 应用。“接管数字世界，改变物理世界”，其自身多元丰富的业务正成为其应用端最先落地孵化的应用场景。

战略：不断完善 AI 基础设施，坚持开源开放

阿里巴巴在人工智能的建设上强调完善基础设施与坚持开源。11月20日，在2024年世界互联网大会乌镇峰会主论坛上，阿里巴巴集团 CEO 吴泳铭表示，面向 AI 时代的生产力革命，阿里巴巴一直专注做好两件事：一是不断完善 AI 基础设施和相关支撑体系；二是坚持开源开放，与广大生态伙伴一起，把 AI 能力转化为千行百业的生产力。据通义公众号，2025年1月28日，通义两大模型 Qwen-2.5-VL 和 Qwen2.5-1M 已上线魔搭/Huggingface；Qwen2.5-VL 主要特色为较强的视觉智能体能力，不仅擅长识别常见物体，如花、鸟、鱼和昆虫，还能够分析图像中的文本、图表、图标、图形和布局。Qwen2.5-1M 擅长长文本处理能力，推出 7B、14B 两个尺寸，均在处理长文本任务中稳定超越 GPT-4o-mini；而在此前的 2024 年 6 月，阿里巴巴已开源第二代大语言模型-Qwen2。据 2024 年 9 月阿里云栖大会，在 HuggingFace 和魔搭社区中，已有超过 5 万个基于通义千问模型开发的衍生模型，位于全球第二，超过国内其他开源模型，我们认为基于通义千问模型的生态系统正在逐步成熟。

基础大模型布局：全尺寸、全模态、多场景的模型大家族

阿里旗下的基础模型为大语言模型通义千问和视频生成模型通义万相。为建立“模型即服务”的能力，通义模型大家族已涵盖了全尺寸、全模态、多场景的基础模型、模型服务、模型应用和模型社区。除了基础模型和应用外，阿里云百炼负责将模型的能力真正应用到业务场景中，将模型调用和企业级数据有效连接；同时开源模型社区魔搭也在不断推动行业模型演进。

图表29：通义大模型大家族全景图



资料来源：阿里云栖大会，华泰研究

三款旗舰大语言模型满足不同类型用户需求。据阿里云官网，截止 2025 年 2 月 4 日，阿里巴巴旗舰模型主要为通义千问-Max、通义千问-Plus 和通义千问-Turbo 三款。Max 为通义千问系列效果最好的模型，适合复杂、多步骤的任务；Plus 能力均衡，推理效果、成本和速度介于通义千问-Max 和通义千问-Turbo 之间，适合中等复杂任务；Turbo 为通义千问系列速度最快、成本很低的模型，适合简单任务。

图表30：阿里旗下部分模型介绍和收费情况

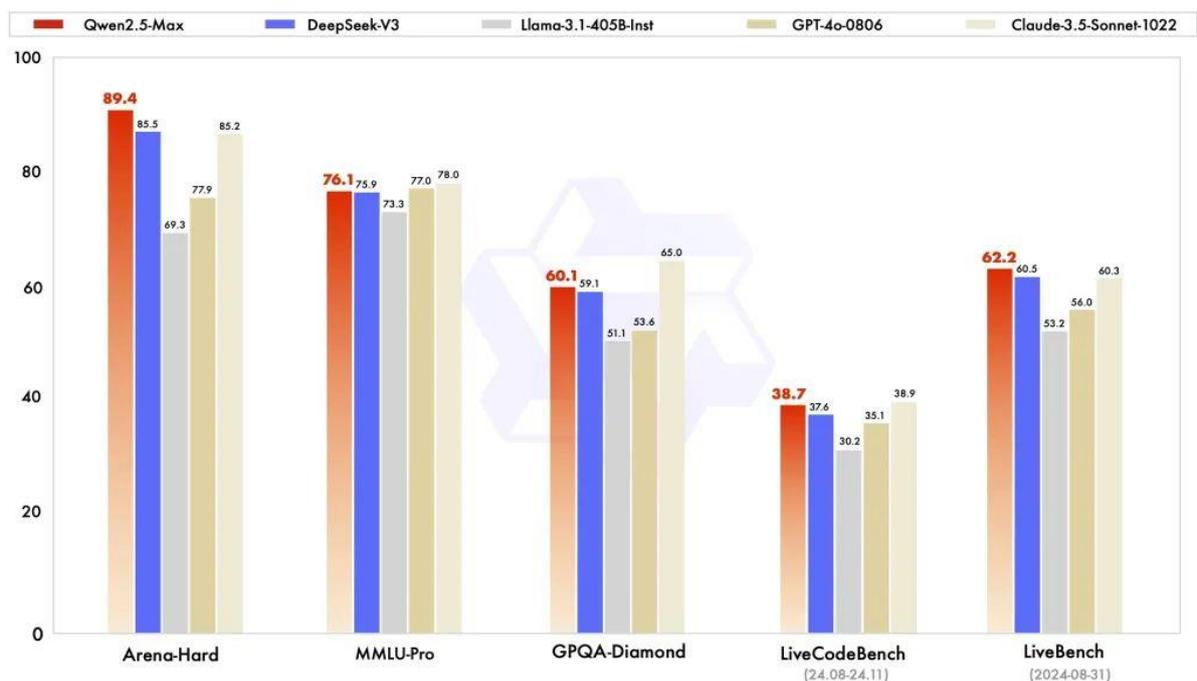
模型系列	通义千问系列大模型						通义千问开源大模型		
	qwen-plus	qwen-turbo	qwen-max	qwen-long	qwen-vl-plus	qwen-vl-max	qwen2.5-72b-instruct	qwen2.5-32b-instruct	qwen2.5-14b-instruct
模型介绍	极强推理效果	高效率，低成本	极强推理能力	优秀的文档处理能力	大规模视觉语言模型	增强视觉语言模型	国际顶尖开源大模型	中等规模顶尖开源模型	中小规模首选模型
适用场景	支持复杂任务	简单任务，可训练	适合复杂任务	超长文本对话	广泛任务，可训练	适合复杂任务	复杂任务效果优异	效果成本兼顾，场景广泛	成本可控，性价比之选
免费额度	新用户首开免费领通义大模型千万 tokens，通义主流模型每个限免 100 万 tokens								
模型价格	输入 0.0008 输出 0.002	输入 0.0003 输出 0.0006	输入 0.0024 输出 0.0096	输入 0.0005 输出 0.002	输入 0.00075 输出 0.00225	输入 0.0015 输出 0.0045	输入 0.004 输出 0.012	输入 0.002 输出 0.006	输入 0.001 输出 0.003
上下文长度 (Tokens 数)	131072	1000000	32768	10000000	8000	32768	131072	131072	131072

资料来源：阿里云官网，华泰研究

在模型版本上，Qwen2.5 是 Qwen 大型语言模型的最新系列。据阿里云官网，针对 Qwen2.5，阿里云发布了一系列基础语言模型和指令调优语言模型，参数规模从 70 亿到 720 亿不等。Qwen2.5 在 Qwen2 基础上进行了若干改进：1) **更大的数据集规模**：Qwen2.5 在最新的大规模数据集上进行预训练，包含多达 18 万亿个 Token。2) **更好的编码和数学能力**：通过阿里云在这些领域的专业专家模型，模型的知识增多，编码和数学能力也获得提高。3) **更出色的遵循指令等方面能力**：在遵循指令、生成长文本（超过 8K 个标记）、理解结构化数据（例如表格）和生成结构化输出（尤其是 JSON）方面有显著改进。对系统提示的多样性更具弹性，增强了聊天机器人的角色扮演实现和条件设置。4) **更丰富的语言种类**：支持超过 29 种语言，包括中文、英语、法语、西班牙语、葡萄牙语、德语、意大利语、俄语、日语、韩语、越南语、泰语、阿拉伯语等。

最新版本 Qwen2.5-Max 在测试中展现出全球范围内较为领先的模型性能。据通义公众号，2025 年 1 月 29 日，Qwen2.5-Max 版本正式发布：1) **在数据规模上**，该版本预训练数据超过 20 万亿 tokens，超越前序所有 Qwen 模型的数据集规模；2) **在性能上**，Qwen2.5-Max 在知识（测试大学水平知识的 MMLU-Pro）、编程（LiveCodeBench）、全面评估综合能力（LiveBench）以及人类偏好对齐（Arena-Hard）等主流权威基准测试上，展现出全球领先的模型性能。

图表31：Qwen2.5-Max 在指令模型版本中的性能表现



资料来源：通义公众号，华泰研究

图表32: Qwen2.5-Max 在基座模型版本中的性能表现

	Qwen2.5-Max	Qwen2.5-72B	DeepSeek-V3	LLaMA3.1-405B
MMLU	87.9	86.1	87.1	85.2
MMLU-Pro	69.0	58.1	64.4	61.6
BBH	89.3	86.3	87.5	85.9
C-Eval	92.2	90.7	90.1	72.5
CMMLU	91.9	89.9	88.8	73.7
HumanEval	73.2	64.6	65.2	61.0
MBPP	80.6	72.6	75.4	73.0
CRUX-I	70.1	60.9	67.3	58.5
CRUX-O	79.1	66.6	69.8	59.9
GSM8K	94.5	91.5	89.3	89.0
MATH	68.5	62.1	61.6	53.8

资料来源: 通义公众号, 华泰研究

AI 能力加持下, 云计算服务有望实现降本增收

AI 大模型技术有望与云计算市场形成互相支撑与技术升级。我们认为, 融合 AI 能力后, 传统云计算业务的计算能力实现了“智”的升级与改造, 效率优化的核心价值得到进一步强化, 规模效应下, 我们认为阿里云有望凭借自身较大的业务体量和技术端的相对优势, 实现相较于竞争对手更具有优势的硬件端边际成本, 从而获取更多市场份额, 而阿里云在模型调用上持续实现降价也初步印证了这一特征。从路径来看, 阿里云选择开源与闭源兼顾的模式, 一方面希望借助开源模型进一步激发社会创造力, 与社会共同分享技术成果, 另一方面也通过较高门槛的闭源模型实现增收获利, 我们认为这是兼顾了短期增长与长期发展的选择。

图表33：阿里云拥有全球+全栈的技术能力



资料来源：阿里云栖大会，华泰研究

百度：全栈服务能力布局助力 AI 应用落地

百度致力于打造全新的 AI 基础设施，围绕大模型应用帮助企业加速实现智能化升级，其 AI 基础设施包括用于大模型和应用开发的千帆平台，以及为客户提供大模型所需的算力服务的百舸平台，是国内少数几家实现了从芯片到框架再到模型及应用的全栈布局的企业之一，有助于利用自身较为强大的 AI 能力和云计算资源，为用户提供一站式的解决方案。与同行相比，百度在大模型商业化上更倾向于采用闭源逻辑，并相信在大多数实际应用场景中，闭源模型能提供更优的解决方案。在应用端，百度着力布局两大 AI 应用方向：1) 智能体，公司判断这将是 AI 应用未来的最主流形态，类别包括公司类、角色类、工具类、行业类；2) 产业应用，涵盖制造、能源、交通、政务、金融、汽车、教育、互联网等众多行业。

战略：注重产业应用场景与智能体开发

百度旨在打造高效 AI 开发能力并优化应用和数字底座。据百度智能云公众号，2022 年 9 月发布全新战略“云智一体，深入产业”及“云智一体 3.0”架构并贯彻至今：“云智一体 3.0”架构是从行业核心场景切入，通过打造行业标杆应用，带动和沉淀 AI PaaS 层和 AI IaaS 层的能力，打造高性价比的异构算力和高效的 AI 开发运行能力，进而向上可以优化已有应用、孵化新应用，向下改造数字底座，使基础云更适合 AI 应用，形成螺旋上升、不断进化的效果；而产业里丰富的应用场景，为人工智能与云的发展提供了广阔的空间。

智能体和产业应用是百度着重布局的应用方向。据 2024 年 11 月百度世界大会，管理层认为两大 AI 应用方向分别为智能体与产业应用，其中智能体是 AI 应用的最主流的形态，即将迎来它的爆发点，它可能会变成 AI 原生时代内容、信息和服务的新的载体，主因：1) 智能体门槛较低，易于用户和开发者上手，并可通过搜索和其它渠道进行分发；2) 智能体拥有较高的天花板，可制作出功能较为强大的应用，多个智能体协作亦可解决更加复杂的问题，百度在文心智能体平台上已经吸引了 15 万家企业和 80 多万名的开发者。在产业应用上，管理层表示百度智能云是百度构建智能产业生态的主要依托，在百度智能云的千帆大模型平台上，公司已累计帮助用户精调了 3.3 万个大模型，开发出 77 万个企业应用，并表示百度智能云已拥有了中国最大的大模型产业落地的规模，覆盖客户包括百胜中国、智联招聘等。

基础大模型布局：拥有多种类大模型满足用户需求

文心大模型家族包含自然语言处理 NLP 大模型（专注于处理语言相关的任务）、视觉 CV 大模型（提供基于视觉技术）、跨模态大模型（基于知识增强的跨模态语义理解技术，支持跨模态检索、图文生成、图片文档信息抽取等应用的快速搭建）等多类型模型。通用旗舰大模型为 ERNIE 4.0、ERNIE 4.0 Turbo、ERNIE 3.5，适用于复杂场景、高级分析与规划。据 3Q24 百度业绩公告，在 2024 年 11 月，文心大模型每天处理 API 量达 15 亿次，相比于 8 月的 6 亿次显著提升。据 IDC 数据，2023 年百度智能云在中国 AI 公有云市场份额为 26.4%，位居第一。同时，百度智能云在大模型平台市场份额为 19.9%，也位居第一。

图表34：百度文心大模型与工具平台全景图

文心产业级知识增强大模型										
工具平台	数据标注与处理		大模型精调		大模型压缩		高性能部署		场景化工具	
文心大模型	行业大模型									
	自然语言处理				视觉			跨模态		生物计算
	文心一言 ERNIE Bot				OCR图像表征学习 VIMER-StrucText			文档智能 ERNIE-Layout		化合物表征学习 HelixGEM
	对话 PLATO-XL		搜索 ERNIE-Search		多任务视觉表征学习 VIMER-UFO			文图生成 ERNIE-ViLG		蛋白质结构预测 HelixFold
	跨语言 ERNIE-M		代码 ERNIE-Code		视觉处理多任务学习 VIMER-TCIR	自监督视觉表征学习 VIMER-CAE		视觉-语言 ERNIE-ViL	语音-语言 ERNIE-SAT	单序列蛋白质结构预测 HelixFold-Single
	语言理解与生成 ERNIE				ERNIE 3.0	鹏城·百度文心	ERNIE 3.5	ERNIE 4.0		

资料来源：文心大模型官网，华泰研究

图表35：百度文心大模型家族全景图

文心大模型家族全景图													
	旗舰大模型			主力大模型				轻量级大模型	垂直场景大模型				
模型名称	ERNIE 4.0	ERNIE 4.0 Turbo	ERNIE 3.5	ERNIE Speed	ERNIE Lite	ERNIE Speed Pro	ERNIE Lite Pro	ERNIE Tiny	ERNIE Character	ERNIE Functions	ERNIE Novel	视觉系列模型 OCR模型 异常检测模型 ...	
上下文长度	8K	8K	8K 128K	8K 128K	8K 128K	128K	128K	8k	8k	8k	8k	语音系列能力模型 超自然合成音色 多角色情感有声阅读 个性化声音复刻	
定位	通用复杂场景、高级分析与规划			垂直场景定制训练或构建行业模型				特定场景自然语言制指令调用、边缘设备推理场景	游戏NPC 客服对话 角色扮演等	问答场景外部工具使用和业务函数调用	小说续写 短剧 电影等场景		

资料来源：云头条，华泰研究

百度多款主力大模型已可免费调用。据百度智能云官网，在大模型调用费用上，百度针对较为先进的旗舰大模型 ERNIE 4.0 Turbo、ERNIE 4.0 和 ERNIE 3.5 采用收费制度，输入/输出价格最低低至 0.0008/0.002 元每千 tokens，而主力大模型 ERNIE Speed、ERNIE Lite、ERNIE Tiny 均已免费使用，其中 ERNIE Speed 是百度 2024 年发布的自研高性能大语言模型，通用能力优异，适合作为基座模型进行精调，可更好地处理特定场景问题，同时具备较佳的推理性能；而 ERNIE Lite 是百度自研的轻量级大语言模型，兼顾优异的模型效果与推理性能，适合低算力 AI 加速卡推理使用。

图表36：百度旗下 AI 大模型调用费用

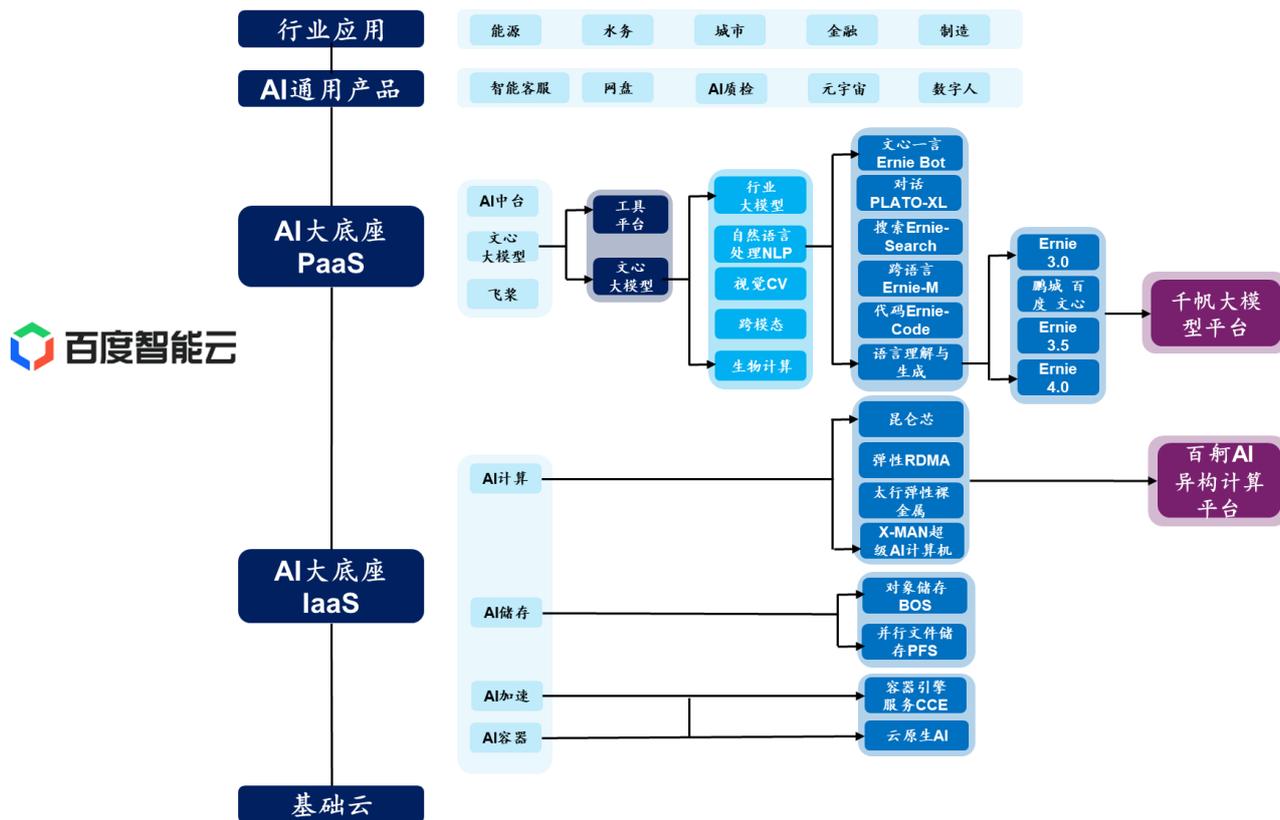
模型名称	版本名称	服务内容	子项	单价
ERNIE 4.0 Turbo	ERNIE-4.0-Turbo-128K	推理服务	输入	0.02 元/千 tokens
	ERNIE-4.0-Turbo-8K		输出	0.06 元/千 tokens
	ERNIE-4.0-Turbo-8K-Preview	搜索增强	触发	0.008 元/次
	ERNIE-4.0-Turbo-8K-0628			
ERNIE 4.0	ERNIE-4.0-Turbo-8K-Latest	推理服务	输入	0.02 元/千 tokens
			输出	0.02 元/千 tokens
		搜索增强	触发	0.008 元/次
ERNIE 4.0	ERNIE-4.0-8K	推理服务	输入	0.03 元/千 tokens
	ERNIE-4.0-8K-0613		输出	0.09 元/千 tokens
	ERNIE-4.0-8K-Latest	搜索增强	触发	0.008 元/次
	ERNIE-4.0-8K-Preview			
ERNIE 4.0	ERNIE-4.0-8K-Latest	推理服务	输入	0.03 元/千 tokens
			输出	0.03 元/千 tokens
		搜索增强	触发	0.008 元/次
ERNIE 3.5	ERNIE-3.5-128K	推理服务	输入	0.0008 元/千 tokens
	ERNIE-3.5-8K		输出	0.002 元/千 tokens
	ERNIE-3.5-8K-0701	搜索增强	触发	0.008 元/次
	ERNIE-3.5-8K-Preview			
ERNIE 3.5	ERNIE-3.5-8K-0613	推理服务	输入	0.0008 元/千 tokens
			输出	0.0008 元/千 tokens
		搜索增强	触发	0.008 元/次
ERNIE Speed	ERNIE-Speed-128K	推理服务	输入	免费
	ERNIE-Speed-8K		输出	免费
ERNIE Lite	ERNIE-LITE-8K	推理服务	输入	免费
			输出	免费
ERNIE Tiny	ERNIE-Tiny-8K	推理服务	输入	免费
			输出	免费

资料来源：百度智能云官网，华泰研究

全栈服务能力为 AI 应用落地奠定良好基础

百度凭借在云计算服务与 AI 技术领域的长期深耕，已经建立起全栈服务能力，其业务范畴涵盖了全面的基础设施、IaaS、PaaS 及应用能力，使其在国内大模型领域的竞争中保持竞争力，也为其积极投入于捕捉 AI 应用落地的机遇奠定了良好基础。百度智能云通过构建全栈的 AI 原生云架构，面向 AI 应用场景进行深度的软硬一体协同优化，这种架构在 AI 基础设施层面能够为企业提供高性能的 AI 算力，以及更简单、易用、高效的开发过程，确保更流畅的端到端应用开发全流程体验；并通过 AI 计算（基于百度“太行”弹性裸金属产品，为客户提供高性能、多规格、高性价比的算力服务）、AI 存储（提供面向 AI 计算的存储服务架构）、AI 容器（容器引擎服务 CCE 提供基础的 Docker 容器生命周期管理、大规模容器集群运维管理、业务应用一键式发布运行等功能）三层来构建 AI 开发基础设施，从而大幅提升 AI 开发效率。

图表37：百度智能云业务版图

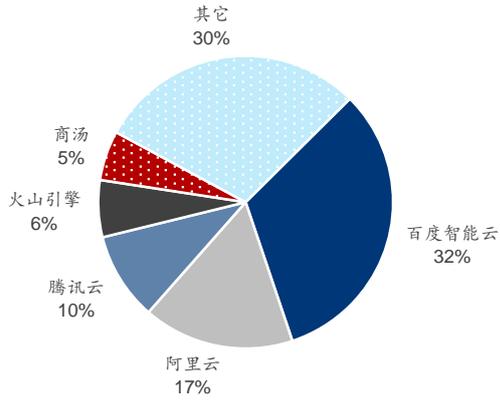


资料来源：公司官网，华泰研究

1H24 百度在中国 MaaS 市场份额位居第一。据 IDC 数据，以收入计，1H24 中国 MaaS 前三大服务商分别为百度智能云/阿里云/腾讯云。百度智能云的 MaaS 收入规模达 8,000 万元人民币，位居市场第一，市场份额达 32.4%，百度智能云的 MaaS 业务主要通过其千帆大模型平台提供服务，为企业提供大模型全生命周期工具链和整套环境，让企业用户可以在千帆平台上直接调用文心大模型的服务，也可以进行开发、训练、部署和调用自己的大模型服务，同时还提供应用开发层全套工具链支撑模型到应用落地的最后一公里，为企业提供企业级 RAG、企业级 Agent 等全链路应用落地较佳实践。

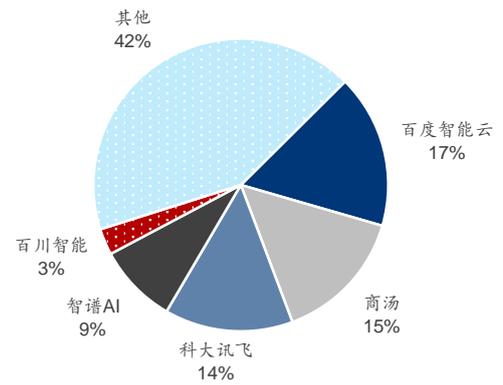
在 AI 大模型解决方案市场，百度亦位居第一。据 IDC 数据，1H24 百度智能云/商汤科技/科大讯飞收入位列中国 AI 大模型方案市场份额的前三位，营收占比分别为 17.0%/14.8%/14.0%。凭借百度智能云在电信、能源、制造、交通等垂类场景下的长久深耕与经验积累，我们期待其在大模型能力持续进化的过程中继续强化行业认知，形成一定的行业壁垒，及时捕捉宏观经济回暖后传统产业 IT 支出的修复机遇，同时进一步提升解决方案的规模可复制性，支撑业务利润率的改善。

图表38: 中国模型即服务 (MaaS) 市场份额, 1H24



资料来源: IDC, 华泰研究

图表39: 中国 AI 大模型方案市场份额, 1H24



资料来源: IDC, 华泰研究

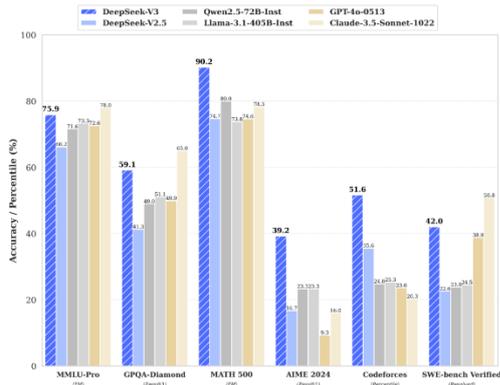
AI 应用：多场景渗透，产品力持续迭代

AI 应用的发展趋势和竞争要素

我们判断 25 年开始 C 端的 AI 应用可能迎来一个爆发期，背后的动力可能来自于开源模型、算力成本的降低、多模态的融合等等。

- 1) 开源模型发展带来的技术平权，推动应用类 AI 基座模型的能力提升。
- 2) 训练、推理算力成本降低，利好更为广泛的垂类应用，同时对于大厂而言，正在由大额资本开支路径逐渐向追求应用端 ROI 倾斜。
- 3) 多模态融合进展迅速，为应用场景多样化奠定基础。

图表40：开源 DeepSeek 在基准测试中好于一众闭源模型



资料来源：DeepSeek，华泰研究

图表41：大模型调用价格持续呈现下降趋势

年份	模型	价格
2023	GPT-4	30美元/百万token
	GPT-4 Turbo	10美元/百万token
	GPT-3.5 Turbo	1美元/百万token
2024	o1	30美元/百万token
	GPT-4o	2.5美元/百万token
	GPT-4o mini	0.15美元/百万token

资料来源：OpenAI，华泰研究

AI 时代应用的发展逻辑与移动互联网时代相反，需要创造原生需求。移动互联网时代的应用发展：需求洞察——产品设计——技术实现；AI 时代的应用发展：技术突破——场景探索——产品革新。在移动互联网时代是先有需求后有产品，并通过技术满足需求，而在 AI 时代则相反，是率先实现技术突破，并需要通过场景探索和产品革新来挖掘原生需求。

图表42：移动互联网的应用逻辑与 AI 时代应用逻辑的对比



资料来源：华泰研究

基于此我们判断 AI 应用 C 端的两个发展方向：

- 1) 一个是场景渗透增加，从最早的搜索、编程等等拓展到教育、营销。AI 发展当前由技术突破驱动，随着 AI 的技术突破，实现技术能力与场景需求的匹配，并不断渗透至更多垂直场景。
- 2) 一个产品层面的革新，重点是丰富交互模式，降低门槛。AI 应用需要由工具演变为高“可玩性”的产品。产品形态会随着技术成熟度和场景理解深度而不断迭代。比如从初始的 chatbot 类交互，逐步演化至 AI Agent 等等，在技术和场景成熟后，需要封装成面向 C 端的产品。

广泛场景的渗透率提升，趋向细分、垂直化

更广泛应用场景的 AI 渗透率提升。技术平权+算力成本持续优化, AI 向各个垂直行业赋能, 将由最初的搜索、编程等等逐步拓展至教育、营销等等多个场景。在互联网应用板块中电商、影视、娱乐、游戏、教育、传媒和金融等等多个板块均有众多潜在的 AI 应用渗透提升方向。

图表 43: AI 在更广泛的应用场景下具备渗透率提升的前景



资料来源:《AIGC 发展研究》, 华泰研究

AI 应用趋向细分、垂直化。在“算力、算法、数据”三要素中, 算力与算法正在随着通用底层能力的优化逐渐被满足, 而数据资产这一要素其实是分散在各个垂直场景中的, 并非通用, 因此各个广泛应用场景有天然的 AI+垂类数据的禀赋。我们认为从后续来看, 除了通用的大模型之外 (DeepSeek、豆包等), AI 应用的趋势可能会更加细分、垂直化, 这背后的原因可能来自于:

- 1) 垂直数据壁垒构筑的护城河, 比如特定的教育、招聘、金融领域;
- 2) 需求的碎片化和长尾化, 需求分散下通用模型无法覆盖;
- 3) 存在合规性问题, 各行各业的 AI 应用均存在自身的合规性问题。

目前能看到主要互联网公司均针对自身领域部署了垂直模型, 比如 AI+教育 (好未来)、AI+游戏 (腾讯)、AI+图像 (美图) 等等。

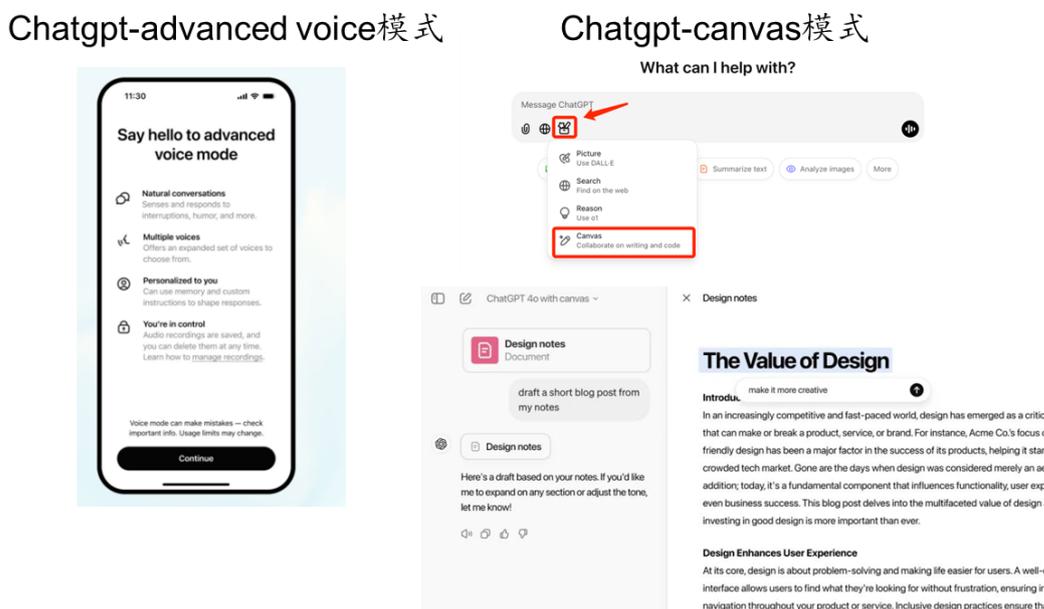
产品层面革新, 聚焦用户丰富交互模式、降低门槛

AI 应用面临产品层面的革新, 聚焦用户丰富交互模式、降低门槛。C 端应用, 满足用户需求的重点是低门槛、个性化、强交互体验 (自然、有趣、更多的情感化产品设计), 这在移动互联网阶段是 C 端应用的重点, 在 AI 时代我们认为也是。因此在交互模式的丰富性和使用门槛等方面都需要面临产品层面的革新, 比如从初始的 chatbot 类交互, 逐步演化至 AI Agent 等等。

以 Chatgpt 为例，能看到从技术驱动向用户体验驱动范式转变。ChatGPT 的用户交互模式革新（如 Advanced Voice Mode 和 Canvas 等等）标志着 AI 应用从技术驱动向用户体验驱动范式转变，其产品层面的创新不仅降低了用户交互门槛，更推动了 AI 技术向多模态、场景化方向的发展：

- 1) **Advanced Voice Mode 是语言交互的全面升级。**AVM 在实时性、情感化和多场景方面实现了革新，基于 GPT-4o 模型，整合自动语音识别（ASR）和文本转语音（TTS）技术，实现无需文本中转的端到端语音交互，延迟降低至接近实时水平，用户可随时打断对话，模仿人类对话的流畅性。
 - 2) **Canvas 是写作与编码的协同创新。**Canvas 整合文本生成、代码调试与可视化编辑功能，用户可通过拖拽界面快速生成内容或调试代码，降低创作门槛。例如，开发者可在同一界面编写代码并实时预览结果，减少上下文切换成本。
- 其本质上是以更丰富的交互方式降低用户使用门槛，增加应用的使用意愿和粘性。

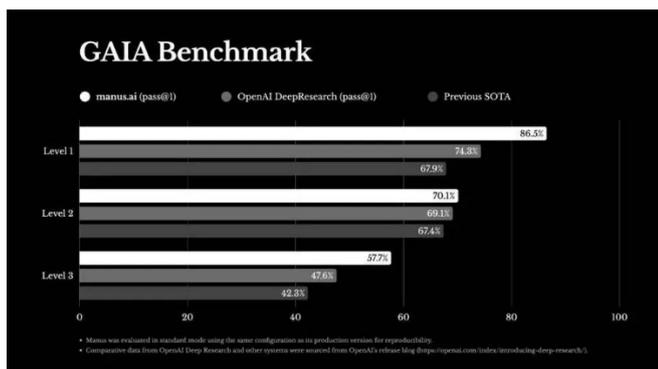
图表44： Chatgpt 优化的交互模式



资料来源：OpenAI，华泰研究

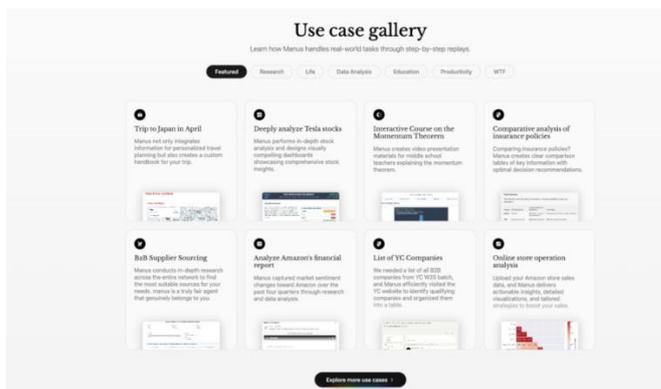
以 Manus 为例，看到的是交互模式从对话到执行的转变，有了 AI Agent 的雏形。2025 年 3 月 6 日 Manus 发布，定位通用型的 AI Agent，实现了从建议到执行能力的全闭环，其在 GAIA 的评分上超越了 OpenAI 的 DeepResearch 成为评分第一。Manus 在交互层面的核心变革是实现了从对话到执行的转变，根据数字生命卡兹克公众号的案例演示，Manus 可以实现根据对话完成各类集合型操作，有了 AI Agent 应用的雏形。

图表45： Manus 在 GAIA 上的评分



资料来源：Manus 官网，华泰研究

图表46： Manus 的使用案例



资料来源：数字生命卡兹克公众号，华泰研究

注重垂直场景的产品力，认知重构阶段把握用户心智

在前文的基础上，我们认为 AI 应用的竞争要素依旧围绕着场景和产品两个重要维度（在技术、算力平权的大假设下淡化了对于竞争者的模型能力和工程化能力的要求）。并有几个启发：

1) 核心是聚焦垂直场景的产品力，正在从“技术竞争”走向“产品竞争”。前文中提到了 AI 应用正在面临从技术驱动向用户体验驱动范式转变，因此需要聚焦垂直场景打磨优质的产品力，将用户体验与 AI workflow 深度融合，在开源、成本降低等大的技术平权的趋势下，竞争要素正在从原先的算法、算力等向应用的产品力转移。

2) 用户对于 AI 应用处在认知重构阶段，需要供给侧的持续启发夺取心智。前文中提到 AI 应用的发展逻辑与移动互联网时代不同的地方是，移动互联网先有需求并打造产品、完善技术去解决，而 AI 是先有技术飞跃，需要通过场景挖掘和产品革新来催化原生需求。因此 C 端用户对于 AI 应用的认知是持续处于重构阶段的，需要供给侧的持续迭代并启发，从而夺取用户心智，这背后可能要求的是更加开放的用户生态。

3) 通过用户反馈积累数据资产，不断迭代并深化实现飞轮。AI 应用的最终潜在的壁垒可能是在高粘性用户、完善生态的基础上跑起来的数据飞轮，即用户使用产品生成数据—数据优化模型—模型提升用户体验—吸引更多用户—产生更多数据，形成正向循环。

AI 应用的商业化落地节奏判断

传统移动互联网时代的变现路径主要有三条：广告、佣金和增值服务（订阅、单购），不同变现模式所依仗的商业基础各不相同，广告核心是流量，佣金是流量+履约，增值服务是提供差异化产品和体验。切换到 AI 应用时代，最快具备商业化基础的是靠 AI 提供的差异化产品和体验，因此增值服务这一商业模式落地节奏最快。而 AI 应用当前尚未形成新的流量高地且产品层面亦未成熟，因此在广告和佣金等商业模式上进展暂缓。**由此我们也能看到通用型 AI 和垂直型 AI 在商业化节奏上的分化**，基于此前提，我们对于 AI 应用的商业化落地节奏形成以下判断：

判断 1：通用型 AI 迈入流量入口争夺阶段，积累用户并探索生态的优先级大于变现。

Chatbot 类应用正在大批量开放免费使用。以 Chatgpt 为首的 Chatbot 类 AI 应用早期以付费订阅为主要的商业模式，但是随着免费开源、算力成本降低等，相关搜索、社交类的通用型 Chatbot AI 正在将大批量免费使用。25 年 2 月 13 日，OpenAI 首席执行官 Sam Altman 宣布免费 ChatGPT 将在标准智能设置下无限制使用 GPT-5 进行对话；Grok 也在 25 年 2 月 18 日发布几天后宣布免费；百度文心一言亦宣布自 25 年 4 月 1 日起全部免费使用。而国内的其他主流 AI Chatbot 应用 DeepSeek、豆包、元宝等等则从上线之初即免费使用。

通用型 AI 正在进入新的流量高地争夺阶段，积累用户并探索生态的优先级大于变现。从移动互联网时代的经验来看，C 端的搜索、社交类应用有很强的虹吸效应，最后是赢家通吃的局面，因此当前通用型 AI 应用的共识是零门槛接入快速占领用户入口，在拥有流量高地并具备护城河之后可以进一步探索广告、佣金等流量变现方式。

AI Agent 的核心痛点是交互通道的打通，这对于有全生态的互联网大厂是核心优势。由 Manus 带来的 AI Agent 趋势将驱动 AI 应用实现从思考到执行的全闭环，而在 AI Agent 的发展中，其核心痛点其实是交互执行中通道的开放和打通，除了公域的数据和执行之外，Agent 要完全发挥能力则需要相对私域的交互权限，这部分对于已有全生态的互联网大厂来说是核心优势，因此在用户入口争夺的下个阶段或能看到广泛的 Agent 应用在互联网大厂的爆发。

判断 2：垂直赛道 AI 靠提供增值产品和体验，生产力场景以刚需驱动高商业转化。

AI 垂类生产力赛道的商业化较为领先。当前 AI 应用商业化的主要方式为订阅付费，通过提供 AI 增值服务收取订阅费用。而当前产生收入的应用中主要以偏向生产力的垂类赛道较为领先，如 Chatgpt、微软 365 Copilot、Perplexity 等等，订阅费约为 10-20 美元/月不等。根据 AI 产品榜公众号，24 年已经有部分 AI 付费用户规模超千万，分别为微软 365copilot（5574 万）、百度文库（4000 万）、Canva（2200 万）、Chatgpt（1000 万）。集中于办公、文档、设计、搜索问答领域，我们认为偏向生产力场景因降本增效量化明确，偏刚需的特征能够有效驱动用户订阅付费转化。因此从商业化的节奏上或更快。

图表 47：当前主流 AI 应用的订阅付费用户及 ARR

场景	产品	付费用户（万）	ARR（亿美元）
办公	微软 365 Copilot	5574	66.9
文档管理	百度文库	4000	-
搜索问答	ChatGPT	1000	27.0
	-api 调用	100	10.0
搜索问答	Perplexity	100	0.5
设计	Canva	2200	25.5
搜索问答	Claude	417	10
文档	Notion	333	-
编程	Cursor	66	-
聊天机器人	Character AI	53.6	0.3

资料来源：AI 产品榜公众号，华泰研究

判断 3：具体场景落地速度需关注试错成本和可验证性。

我们认为判断具体场景的落地速度需要关注其试错成本和可验证性两个维度：

- 3) **试错成本**代表的是对于 AI 出现“幻觉”等等错误的接受度。比如 AI Chatbot 等等其本质是交流和提供咨询建议，用户对于结果有审查，因此试错成本低，此外还有 AI 的图像等内容创作，用户存在二次编辑，此类应用的落地和发展较快。而与之对比的则是试错成本较高的场景，比如自动驾驶等等，涉及到安全和责任划分等等层面，因此其实际落地需要缓慢推进。
- 4) **可验证性**，代表的是 AI 提供的内容验证成本低，可以和用户迅速建立信任感。比如 AI 编程等等，其提供的代码可以迅速判断是否跑通，是否解决问题，验证环节短带来的就是交互的流畅性，可以和用户快速建立信任感。但是比如有些验证环节很长的场景，比如 AI 制定旅游攻略、剧本创作、提供资产配置的建议。其可验证性较低，用户需要更多的成本来判断建议的真伪和有效性，则落地的速度较慢。

基于此我们判断当前互联网板块落地较快的场景为图像编辑、短（长）视频内容、广告营销、教育培训答疑、辅助招聘等等。

AI 全生态应用布局典范：字节跳动、阿里、腾讯、百度

字节跳动：搭建 AI 全生态，爆款豆包产品力先行

字节豆包搭建 AI 全生态，布局搜索、图像、社交等多个垂直应用领域。字节豆包从底层技术到应用生态全面发力。在模型层面，推出了包含云雀大模型在内的 17 款大模型，涵盖大语言、音频、视频、多模态等多个类型，为上层应用筑牢技术根基。在应用层，字节跳动采用多种类覆盖与对标优质竞品的策略，推出了十余款 AI 应用，构建“超级 AI 应用工厂”。

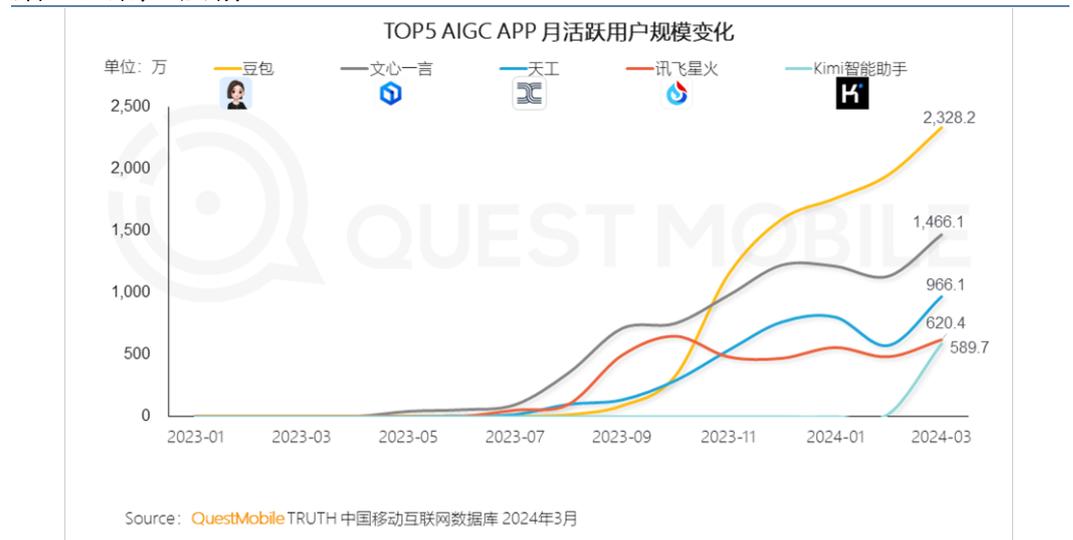
图表48：字节 AI 应用全布局

团队	类型	产品	市场表现	
Flow	聊天	豆包	2970 万，移动端月活：xxxx 万，网页月访问量。	
		Cici	1152 万，移动端月活。	
	社交	猫箱	243 万，移动端月活。	
		Anydoor	156 万，移动端月活。	
	图像	星绘	56 万，移动端月活。	
		PicPic	15 万，移动端月活。	
	Agent	扣子	n/a	
		Coze	n/a	
	剪映	图像/视频	即梦	107 万，移动端月活；xxxx 万，网页月访问量。
			Dreamina	n/a
头条	办公	小悟空	移动端不活跃	
大力教育	教育	Gauth	584 万，移动端月活。	
		豆包爱学	108 万，移动端月活。	
巨量引擎	电商	即创	n/a	
其他	音乐	海绵乐队	移动端不活跃	
	编程	MarsCode	n/a	

资料来源：字节官网，Sensor tower，华泰研究

字节核心产品豆包 APP 月活数仅次于 Chatgpt。24 年 11 月的全球月活跃排行榜上，豆包 App 的 MAU (月活跃用户数) 接近 6000 万，仅次于 OpenAI 的 ChatGPT，位列全球第二；截至目前，豆包大模型日均 tokens 使用量超过 4 万亿，发布 7 个月以来增长超过 33 倍。字节跳动旗下创作工具剪映和 CapCut，全球月活用户也已超过 8 亿，并且在 2024 年实现收入增长超 100%，总收入接近 100 亿元。

图表49：国内 AI 应用月活



资料来源：Questmobil，华泰研究

以用户体验为先，产品深度迭代。根据《中国企业家杂志》，2024 年 1 月至 11 月，该群共产生了数万条聊天记录，豆包官方发布了 100 条群公告，其中大部分为功能内测邀请。据《中国企业家杂志》统计，过去 11 个月，豆包在安卓 APP 端共迭代了约 40 次。24 年 5 月，字节跳动产品与战略副总裁朱骏分享了字节跳动为豆包这类产品定义了三个产品设计原则：第一条，“拟人化”；第二条，离用户很近，随时伴随用户，嵌入用户的不同使用环境，应该是豆包到用户身边，而不是用户到豆包身边；第三条，个性化。

图表50：豆包 APP 发布后持续迭代


资料来源：中国企业家杂志，华泰研究

阿里巴巴：规划布局深入千行百业

阿里旗下 AI 应用：具有较为全面的 AI 应用产品矩阵布局。阿里旗下 AI 产品主要为通义系列，涵盖通义千问、通义智文、通义灵码、通义万相、通义听悟和通义星辰等产品，其中最为通用的通义千问为一款超大规模预训练的多模态模型，可撰写创意文案，充当办公助理、学习助手，亦支持文本回答、文档解析和图片解析。在自身优势电商业务外，还涵盖了绘画、编程、医疗、法律等多个业务板块，体现了其在技术研发、市场拓展、用户体验和产业赋能等方面的泛化规划布局。

图表51：阿里旗下 AI 部分应用

产品类型	产品名	介绍
大语言模型	通义千问	通义千问是阿里云开发的超大规模语言模型，旨在理解和处理用户的自然语言输入。它能够为用户提供多种服务，包括文本创作、翻译、对话模拟等，支持复杂任务的高效处理，适用于各类行业需求。
AI 绘画模型	通义万象	通义万象是一款 AI 绘画生成模型，能够根据用户提供的文字或图片生成多种风格的图像。它支持多种功能，如基础纹生图、相似图像生成和图像风格迁移，为用户提供丰富的创作体验。
AI 音频模型	通义听悟	通义听悟是一个音视频处理助手，能够将音频和视频内容实时转录为文字，并提供翻译和摘要功能。它帮助用户高效记录和理解决音视频信息，适用于会议、学习等场景。
AI 编程模型	通义灵码	通义灵码是一个智能编码助手，旨在帮助开发者生成和解释代码，自动撰写测试用例，并提供编程问题的解答。它经过优化，特别适合阿里云的云服务使用场景，提升开发效率。
AI 角色扮演	通义星辰	通义星辰是一个个性化角色创作平台，允许用户创建多种类型的角色，并与其进行互动对话。它具有拟人化和多模态对话能力，增强了用户的创作体验。
AI 客服	通义晓蜜	通义晓蜜是一款智能客服解决方案，利用 AI 技术提供对话机器人和智能座席助理，旨在提升企业的客户服务效率和用户体验。
AI 法律助手	通义法睿	通义法睿是专注于法律领域的 AI 助手，能够理解和处理法律问题，提供专业的法律咨询和文书生成服务，帮助用户高效解决法律事务。
AI 医疗助手	通义仁心	通义仁心是一个医疗助手，结合了阿里云技术和医疗专业知识，提供健康咨询和症状分析服务，帮助用户获取医疗信息和建议。
AI 金融助手	通义点金	通义点金是一个金融领域的 AI 助手，能够进行财务分析、市场数据解读和文档处理，帮助用户高效获取金融信息和决策支持。
AI 智能体	芝士派（百宝箱）	对接支付宝的智能体平台
AI 模型交流	魔搭	AI 模型分享与交流社区，促进技术创新与应用。
AI 搜索	夸克 AI 搜索	内置于夸克浏览器的 AI 搜索引擎，提供智能搜索服务。
AI 模特	绘蛙	提供丰富商用 AI 模特库，定制专属 AI 模特，生成高品质本地化模特图，支持自由姿势的商品种草图，批量生成有差异化、氛围感的高拍图。
AI 视频	寻光	（内测阶段）AI 视频创作平台，帮助用户生成高质量的视频内容。

资料来源：Shelly 聊 AI，华泰研究

阿里应用端战略：To C AI 业务正在逐步整合，或意在发力 C 端应用。据智能涌现公众号 2025 年 1 月 16 日报道，阿里旗下的智能互联事业群正式由智能信息事业群总裁吴嘉统管。智能互联为阿里在 2022 年成立的一级业务部门，其核心品牌为天猫精灵，目前天猫精灵的硬件团队已与夸克产品团队融合工作，其工作重点包括新一代 AI 产品的规划定义，以及与夸克 AI 能力的融合。此前 2024 年 12 月，“通义”产品团队已从阿里云智能整体划入阿里智能信息事业群，目前该事业群还拥有夸克、书旗小说、UC 浏览器、超级汇川广告平台等业务，初步形成了内容-AI 应用-AI 硬件的完整链条。而大模型的训练和推理通常需要大量交互数据和内容，当前的业务融合在我们看来或有利于大模型业务的创新迭代。此外，据界面新闻 2025 年 1 月报道，全球顶尖人工智能科学家许主洪（Steven Hoi）正式加入阿里巴巴，出任阿里集团副总裁，负责 AI To C 业务的多模态基础模型及 Agents 相关基础研究与应用解决方案，旨在提升阿里巴巴 AI 应用 C 端产品在模型结合应用的端对端闭环能力，阿里 AI To C 业务亦正在组建领先的 AI 算法研究和工程团队，并试图吸引业内优秀人才加盟，显示出意在加码发力 C 端应用的决心。

图表 52：阿里 To C AI 业务近期组织架构变化

日期	事件类型	事件内容
2024 年 12 月	业务分拆+合并	2024 年 12 月，阿里旗下的 AI 应用“通义”近期正式从阿里云分拆，并入阿里智能信息事业群。此次调整包括通义 To C 方向的产品经理，以及相关的工程团队，一并调整至阿里智能信息事业群。调整后，通义 PC 及 App 团队与智能搜索产品“夸克”平级，原有的通义实验室，仍然留在阿里云体系内。
2025 年 1 月	业务分拆+合并	阿里旗下的智能互联事业群，正式由智能信息事业群总裁吴嘉统管。天猫精灵的硬件团队，已经在与夸克产品团队融合工作，其工作重点包括新一代 AI 产品的规划定义，以及与夸克 AI 能力的融合。另外，在团队融合后，新团队也将探索包括 AI 眼镜在内的新硬件方向。
2025 年 2 月	高层变动	全球顶尖人工智能科学家许主洪（Steven Hoi）正式加入阿里巴巴，出任阿里集团副总裁，负责 AI To C 业务的多模态基础模型及 Agents 相关基础研究与应用解决方案。许主洪将专注于 AI To C 业务的多模态基础模型及 Agents 相关基础研究与应用解决方案，大大提升阿里巴巴 AI 应用 C 端产品在模型结合应用的端对端闭环能力上的跃迁。同时，阿里 AI To C 业务正在组建顶级 AI 算法研究和工程团队，吸引大量业内优秀人才加盟。

资料来源：智能涌现，界面新闻，华泰研究

腾讯：加速游戏工业化，引领广告新趋势

腾讯 AI 产品：从内部业务自用起步，强化云业务技术壁垒。腾讯混元技术于 2024 年取得较快突破，同时实现全链路自研。腾讯 LLM 研发起步略晚于其他大型厂商，但近期技术端差距有所收窄。在自用场景方面，混元大模型接入腾讯 700 多个业务场景，改善内部人员工作效率，混元 3D 引擎在 24 年广泛使用后，也于 25 年对外开放；在腾讯云方面，混元大模型也构建了腾讯云 MaaS 坚实的基础。

在 C 端市场，我们看好腾讯的产品长期竞争力，AI 产品数量以娱乐和应用方向居多，腾讯有较高的用户规模和信息资源壁垒；在 B 端市场，腾讯也享有较高的客户资源和团队规模优势。鉴于大模型更新迭代存在追赶效应，我们预期 26-27 年模型层面的性能差异或将逐步缩小，同时伴随 LLM 使用成本显著降低，故而后期的竞争优势仍在获客能力和用户体验（即产品能力）。从产品力角度看，我们认为 C 端市场以应用和娱乐 app 为主，过往腾讯在众多产品矩阵中，用户规模和用户体验在国内达到顶尖水平。

图表53：腾讯产品布局一览

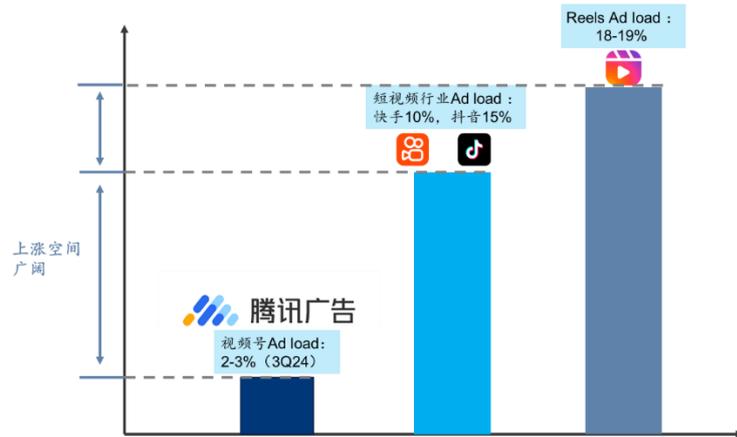


资料来源：腾讯官网，华泰研究

腾讯广告业务：AI 算法改善助力视频号扩张市场份额

我们认为，随着腾讯在生成式 AI 视频工具、智能推荐算法等核心技术上的持续积累，其与抖音快手的差距有望缩小，未来广告加载率具备较大提升空间。抖音快手在短视频领域起步较早，相比之下，视频号内容生态和头部创作者数量有待提高，故而整体 Ad load 仍低于行业平均。根据腾讯 3Q24 业绩会数据，视频号广告加载率约为 2-3%，显著低于抖音快手水平。

图表54：腾讯视频号 Ad load 提升空间显著



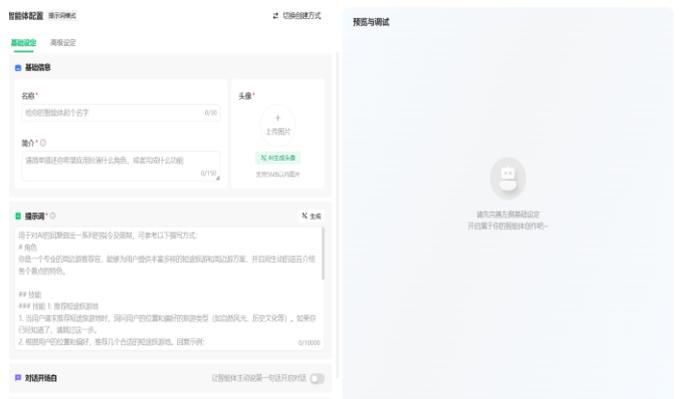
资料来源：公司公告，华泰研究

腾讯元器作为 AI Agent，主要用于智能体开发。腾讯元器基于混元大模型构建，提供一站式智能体创作与分发能力，依托腾讯生态资源，实现零代码开发、模块化扩展与全域部署。平台围绕三大核心模块展开：

- 1) 智能体开发，支持通过系统提示词设定角色属性、知识库挂载及多插件协同调用，精准控制大模型行为。
- 2) 工作流引擎，可实现复杂任务逻辑编排；该引擎采用流程图式低代码编辑器，串联插件、知识库与大模型节点，实现复杂任务逻辑编排。
- 3) 智能体商店和插件市场：提供预置行业解决方案（如客服、教育场景模板），开发者可以自定义扩展，将混元大模型转化为可编程、可分发的商业化工具。

图表 55: 智能体商店涵盖了各种应用场景


资料来源: 腾讯元器官网, 华泰研究

图表 56: 在智能体创建环节需要明确角色、设定、开场白、流程、插件等内容


资料来源: 腾讯元器官网, 华泰研究

腾讯游戏业务: 加速游戏工业化

从游戏业务看, ①腾讯未来突破点在于, AI agent 有望实现智能 NPC 的落地, 例如语言转换成相应的指令, 通过 prompt 驱动 NPC 做更加丰富的表情, 更多的动作反馈和更稳定的人设风格。②LLM 能够加快游戏的训练时间, 《火影忍者》手游, 针对格斗游戏的大规模强化学习方案, 直接让训练时间和资源下降了 90% 以上, ③头部游戏研发的精品化和工业化为大势所趋, 所需研发人员数量和开发时间不断增加, 腾讯于 24 年推出自研 AI 引擎, AI 新工具系列有助于缓冲研发成本上涨。

我们认为, 腾讯 AI 将会增强其游戏工业化优势, 提升玩家总时长释放更多变现潜力。游戏工业化加速的背景, 往往伴随着单款游戏更多的内容生成, 例如头部游戏的总时长有望从过往 20-40 小时提升至 50-100 小时, 从而有望增加更多商业化环节来提高变现潜力。1 月 22 日, 腾讯宣布开源混元 3D 生成大模型, 同时上线混元 3D AI 创作引擎助力内部提升研发效率。

在 CDEC 峰会上, 腾讯副总裁张巍表示, 腾讯于 24 年推出自研 AI 引擎, 具备 3D 图形、剧情、关卡等多种 AIGC 能力, 能够让游戏的部分工作效率提升数倍以上, 例如场景制作、内容生成等领域。他表示从最近的产业调研看, 约八成受访者认为 AI 技术让生产效率提升超过 20%。此外, 在内容生产管线、游戏角色训练、智能体动作生成等方面, 内部也取得了显著进展, 包括但不限于视频 AI 动态捕捉、AI 故事引擎、AI 绘画和建造等功能。

以游戏部门为例, 腾讯更加注重游戏工业化发展, 倾斜更多资源研发头部项目。从游戏品类和内容生成看, 大世界地图和人物更加错综复杂, AI 开发工具有助于减少非核心项目的研发时间, 加速工业化发展。近期, 天美工作室重组为四大研发子工作室, 更加精简的架构组织有望集中研发资源, 从而更好推进头部项目的进展。

百度: AI 重做产品, 引领业务智能化转型升级

百度旗下 AI 应用: 拥有较为丰富的旗舰产品与垂类产品。 旗舰应用产品包括: 基于飞桨平台研制的知识增强大模型“文心大模型”、生成式对话产品“文心一言”、加入上百项 AI 能力的“百度文库”等, 并在用户规模上已形成一定积累, 据 3Q24 公司公告, 2024 年 11 月, 文心系列模型处理的日均 API 调用量达到 15 亿次, 较 8 月的 6 亿次大幅增加; 据界面新闻 2025 年 1 月报道, 百度文库 AI 功能 MAU 已突破 9000 万, AI DAU 年同比增长 230%。付费用户超 4000 万, 在全球范围内仅次于微软 Copilot; 百度文库的付费率年同比增长 60%, 这一增长几乎完全由 AI 功能带动。此外, 在细分应用功能上百度亦有充分布局, 产品包括百度识图 (AI 图像识别)、小侃星球 (AI 聊天机器人)、万话创作者平台 (AI 创作工具) 和慧播星 (AI 电商数字人) 等, 百度通过较为先进的 AI 技术将细分功能与 APP 相结合, 积极探索新的商业机会与服务模式, 共同推动自身 AI 改造的智能化转型升级。



图表 57: 百度旗下 AI 产品

	产品类型	产品名	介绍
旗舰应用产品	通用大模型	文心大模型	文心大模型是基于飞桨平台研制的知识增强大模型,从万亿级数据和千亿级知识中融合学习,具备知识增强、检索增强和对话增强等核心技术。据 3Q24 公司公告,2024 年 11 月,文心系列模型处理的日均 API 调用量达到 15 亿次,较 8 月的 6 亿次大幅增加。
	芯片	昆仑芯	昆仑芯(北京)科技有限公司前身为百度智能芯片及架构部,公司团队在国内最早布局 AI 加速领域,深耕十余年,是一家在体系结构、芯片实现、软件系统和场景应用均有深厚积累的 AI 芯片企业。
	对话大模型	文心一言	文心一言(ERNIE Bot)是百度基于文心大模型技术推出的一款生成式对话产品,它位于百度人工智能四层架构的模型层,具备跨模态、跨语言的深度语义理解与生成能力。文心一言的核心功能包括文本生成、问答系统、文本情感分析和分类、机器翻译等。
	自动驾驶	萝卜快跑	“萝卜快跑”是百度旗下的自动驾驶出行服务平台,旨在通过 AI 技术和大量的驾驶数据训练,提供高效的自动驾驶出行服务。据 3Q24 公司公告,2024 年第三季度,百度的自动驾驶服务萝卜快跑提供的自动驾驶订单为 98.8 万单,同比增长 20%。截至 2024 年 10 月 28 日,萝卜快跑累计为公众提供的自动驾驶出行服务订单超过 8 百万单。
	搜索引擎	百度 APP	百度 APP 上线的 AI 功能包括 AI 捏合摘要、智能标签、AI 角色互动、AI 速读、AI 发布器等。据 3Q24 公司公告,2024 年 9 月,百度 APP 的月活跃用户达到 7.04 亿,同比增长 6%。
	文档	百度文库	结合文心大模型,百度文库加入了上百项 AI 能力,涉及写作、PPT 制作、搜索、文生图等。据界面新闻 2025 年 1 月报道,百度文库 AI 功能 MAU(月活跃用户)已突破 9000 万,AI DAU(日活跃用户)年同比增长 230%。付费用户超 4000 万,在全球范围内仅次于微软 Copilot。另外,百度文库的付费率年同比增长 60%,这一增长几乎完全由 AI 功能带动。
	网盘	百度网盘	该业务的 AI 改造始于 2023 年底,加入了扫描、涂鸦、语音转写、会议纪要等 AI 相关功能。据界面新闻 2025 年 1 月报道,AI DAU 同比增长 90%,由 AI 功能带来的收入年增长率为 120%。
	地图	百度地图	已加入 AI 向导、AI 旅游等功能,据百度公众号,截止 2024 年 3 月,百度地图“AI 向导”功能累计用户已破亿。
	智能设备	小度	以 DuerOS 为核心技术品牌,并通过硬件作为载体进行跨场景布局,旗下产品包括智能屏、智能音箱、智能学习机、全屋智能解决方案等。2024 年,DuerOS 基于文心大模型进行了重构,进化成为全球首个 AI 原生操作系统(DuerOS X),进一步增强了其智能化水平和服务能力。
	细分应用产品	图像识别	百度识图
聊天机器人		小侃星球	小侃星球是一款基于人工智能技术的聊天机器人,由百度推出,旨在为用户提供智能化的对话交流服务。该平台通过自然语言处理和机器学习算法,能够理解用户的问题并给出相应的回答,帮助用户解决各种疑问和困惑。
创作工具		万话创作者平台	万话创作者平台是一个由百度公司推出的 AI 创作工具,专注于角色创建和管理。万话创作者平台不仅提供了强大的角色创建和管理功能,还支持个性化定制和社交互动,是一个非常全面的 AI 创作工具。
电商数字人		慧播星	慧播星是百度推出的一款电商数字人直播平台,旨在通过人工智能技术帮助商家快速生成数字人主播,并提供全面的直播解决方案。
AI 修图		百度网盘 AI 修图	百度网盘 AI 修图是一款由百度网盘推出的智能图片编辑工具,专为摄影行业设计,旨在提高修图效率和质量。该工具利用百度强大的云引擎提供的人像分析检测和图像美化能力,实现自动化的修图流程。
漫画生成		百度智能漫画	百度智能漫画是百度文库推出的一项 AI 漫画生成功能,旨在利用人工智能技术帮助用户轻松创作个性化的漫画故事。该功能允许用户通过简单的文字描述,快速生成具有丰富情感和动态表现的漫画角色以及完整的故事情节。
法律助手		法行宝	法行宝是由百度公司推出的一款免费 AI 智能法律助手平台,旨在利用人工智能技术简化和优化法律服务流程,为用户提供快速、高效的法律咨询和解决方案。
摄影写真		超能画布	超能画布是一款 AI 人像摄影写真创作工具,利用的图像处理大模型和智能人脸融合算法,用户可以通过上传照片和输入文字指令,快速生成具有专业质感的 AI 写真。
代码助手		文心快码	文心快码(Baidu Comate)是由百度开发的一款基于文心大模型的智能代码助手,旨在通过 AI 技术提升开发者的编程效率和质量。Comate 通过其强大的 AI 技术和丰富的功能,为开发者提供了一个高效、智能的编程助手,极大地提升了编程效率和代码质量。
文本创作		橙篇 AI	橙篇 AI 是百度推出的一款综合性 AI 原生应用,旨在帮助用户高效、智能地完成长文本的创作和管理。该工具基于百度文库的庞大内容库和先进的 AI 技术。

资料来源:百度智能云公众号,昆仑芯官网,界面新闻,百度公众号,打工 AI 神器公众号,沃垠 AI 公众号,华泰研究

百度广告业务：量价角度持续探索 AI 赋能效果

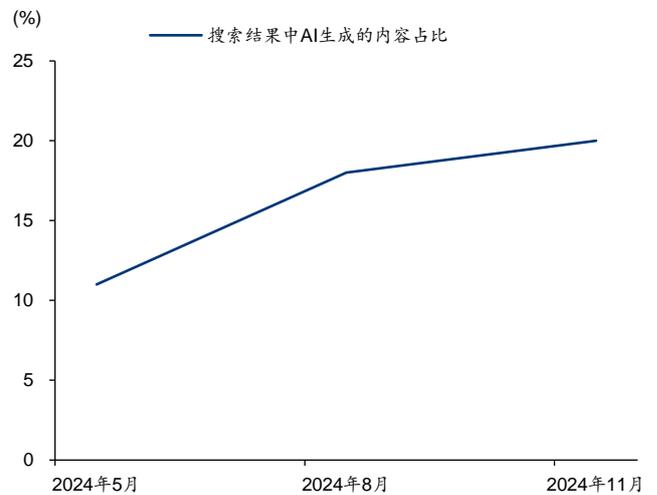
激烈的行业竞争中，AI 或是百度重塑广告业务竞争力的重要武器。百度目前正在从量与价两个角度持续探索 AI 对广告业务的潜在赋能效果：1) 改善用户体验→促进应用流量增长→扩充广告位，与 2) 优化商户投放效率，改善 ROI→提高商户付费意愿→促进广告价格提升。据公司 1Q24-3Q24 公告，百度搜索引擎的结果展示中，AI 生成结果占比已持续提升，由 2024 年 5 月的 11% 提升至 2024 年 11 月的 20%，表明了其在应用中 AI 改造的不断尝试。尽管 AI 生成结果的占比提升在我们看来意味着部分广告位的短期牺牲，但或是百度重塑 C 端用户心智、推动长期用户流量增长的必然选择。考虑到百度 APP 的用户体量相较微信、抖音等 APP 仍有一定增长空间，我们认为量与价两方面在长期来看均有可能成为未来重振百度核心广告收入增长的驱动力。但结合温和的宏观环境、上游广告主需求以及激烈的行业竞争，我们认为短期量的重要性或高于价，且用户体验的改善带动用户活跃度的提升是其在价格端实现进一步议价的前提。我们期待百度通过自有生态体系内不断的 AI 功能实践，在用户活跃度的改善上进一步取得显著成果，这或将是其广告收入逐步回暖的关键要素。

图表58：百度 AI 搜索结果示意图



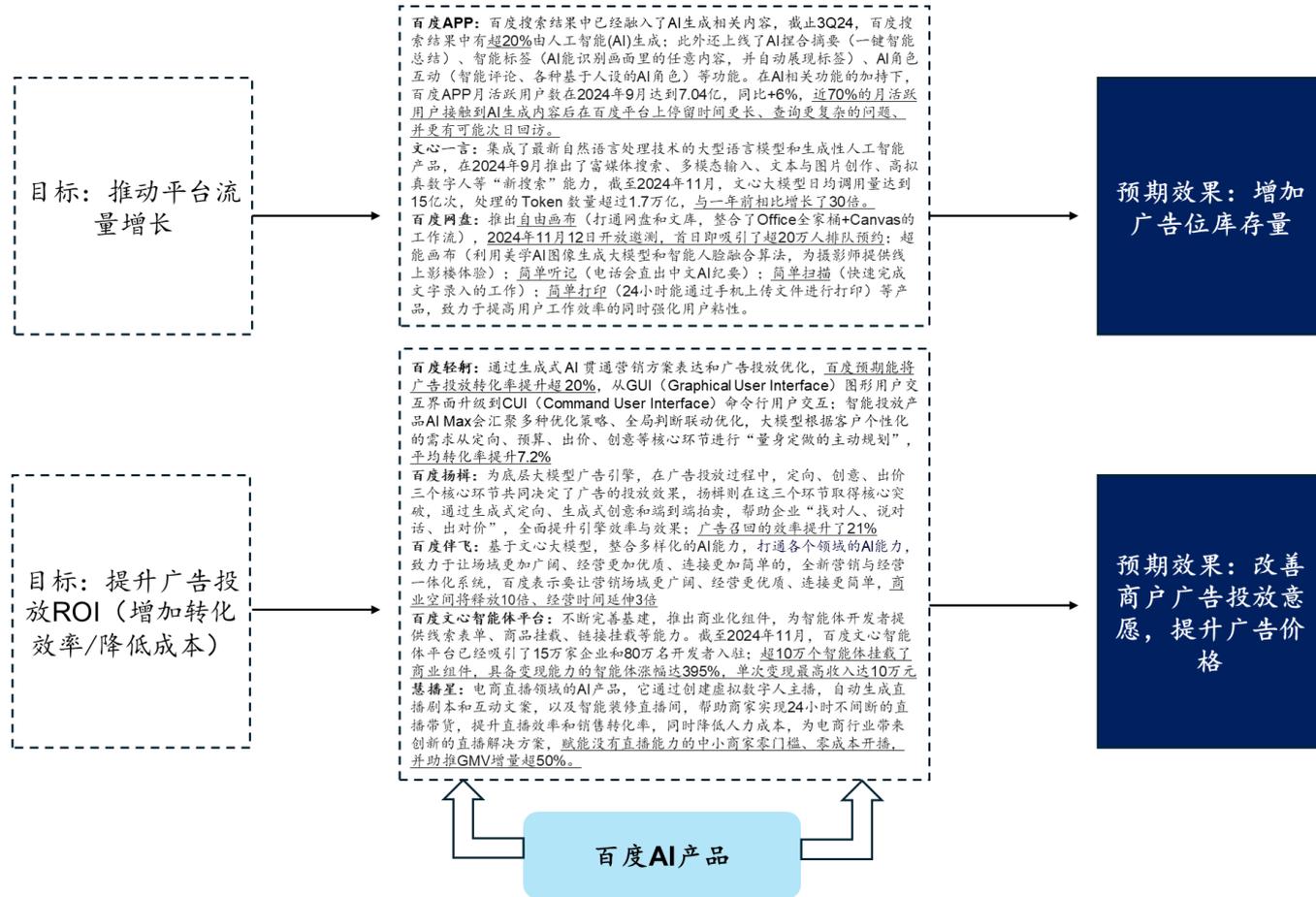
资料来源：百度，华泰研究

图表59：百度搜索中 AI 生成内容结果占比



资料来源：公司公告，华泰研究

图表60：百度 AI 广告业务梳理



资料来源：公司官网，公司公告，华泰研究

AI+图像：受益于DS推进技术平权，持续图像领域精耕细作

在传统图像领域有卡位优势，布局全面，以AI催化付费订阅。美图长期专注图像美化赛道，拥有2.6亿MAU流量基本盘并具备卡位优势。其以自研图像大模型为基，打造图像AI生产力工具全生态，通过挖掘垂直场景需求并构建AI原生 workflow 解决生产力痛点。同时以AI驱动订阅付费：1、大模型底层能力支持缩短付费功能的研发周期；2、AI带来多样化效果挖掘更多的需求付费点；3、AI提升创作效率，加大用户付费意愿。公司当前拥有AI图片、AI视频和AI设计三大方向。

图表61：美图 AI 产品矩阵



资料来源：美图公司，华泰研究

聚焦垂直场景打造 AI 原生 workflow。公司聚焦垂直场景打造 AI 原生 workflow，深耕“小而美”的垂直应用并深度打磨产品力。其美图设计室专攻图像生产力场景，打造 AI 原生 workflow，成本仅为传统 workflow 的 3%，并且短时间内出产品，大大降低门槛。其开拍专攻 KOL 口播场景，打造 AI 口播 workflow，以 AI 完成脚本、提词器、剪辑等工作，将原先 3-7 天的制作时长缩短至 10 分钟，并且持续迭代转化效果。

图表62：美图 AI 电商设计 workflow 和 AI 口播 workflow



资料来源：美图公司，华泰研究

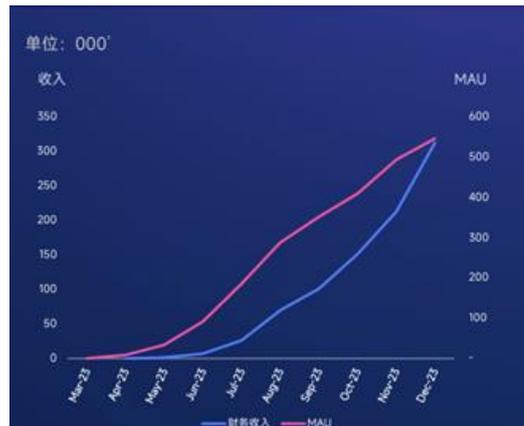
AI 原生应用表现突出，AI 成功形成业绩转化。公司聚焦垂直场景打造 AI 原生应用，以 AI 驱动的美图设计室已经成为中国内地 AI 设计工具赛道的头号产品（以月活数量衡量）。截至 1H24 美图设计室的订阅用户数达到 96 万（截至 23 年底为 87 万），订阅渗透率达到 6%，产品收入同比增长 152%。聚焦口播视频 AI 工作量的开拍自上线以来用户累积内容突破 5000 万。聚焦 AI 视频编辑的 Wink 月活 1H24 同比增长 99%，证明其 AI 强大的产品力。

图表63：美图设计室收入和 MAU 持续提升



资料来源：美图公司，华泰研究

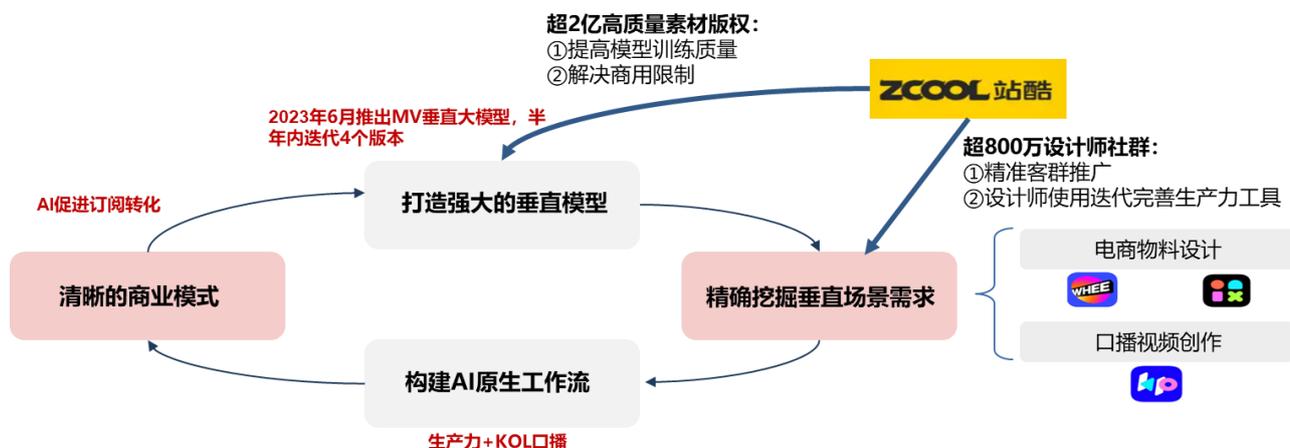
图表64：开拍收入和 MAU 持续提升



资料来源：美图公司，华泰研究

完善 AI 图像生态形成商业闭环。美图当前在图像赛道完成了打造垂直模型—精确挖掘垂场景需求—构建 AI 原生工作流—以清晰的商业模式驱动业绩增长的商业闭环。在每个环节均做出了充足的储备，比如在垂类大模型方面除了美图自研之外，积极接入豆包、可灵等完善能力；比如公司收购设计师社区站酷，面对超 800 万设计师社群进行精准客群推广并以开放生态迭代完善工具；此外发挥公司多年的美学积累和优质的产品力积淀，构建 AI 原生工作流夺取用户心智。

图表65：美图 AI 影像业务生态



资料来源：美图公司，华泰研究

图生视频功能在电商 B 端领域亦有应用空间。据阿里妈妈数字营销公众号，2025 年 1 月，基于自研的淘宝星辰视频生成大模型，阿里妈妈正式推出商用级图生视频 AI Native 应用——淘宝星辰·图生视频，目前已对淘宝天猫商家开放使用。淘宝星辰·图生视频结合了海量电商数据、丰富的设计语言以及业内领先的营销经验，旨在推动电商 AI 内容创作的全流程智能化升级。相比于其它同类应用，淘宝星辰的优势主要在于：1) 基于丰富的电商数据和设计经验，能自动识别并应用各类商品最适合的展示方式，而且从产品细节到整体风格保持高度协调统一；2) 能够深度理解图片以及用户的多语言指令，确保生成视频内容与用户意图高度契合；3) 严格遵循物理规律与动作规范，生成画面流畅自然，提升了视频的真实感和专业性；4) 无论是人脸、商品、还是文字、贴纸等装饰元素，“淘宝星辰·图生视频”始终保持画面清晰、完整、无抖动或变形，保持视觉一致性。

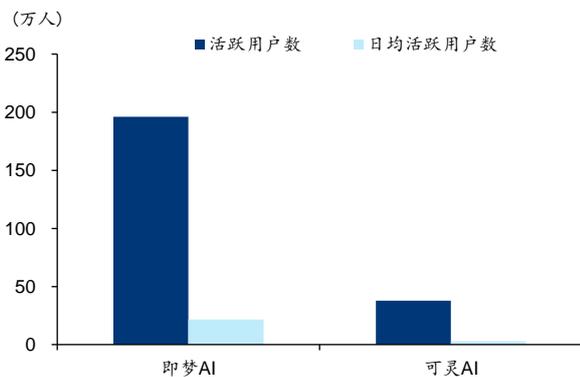
图表66：淘宝星辰·图生视频有望为商家营销物料提质增效


资料来源：阿里妈妈数字营销，华泰研究

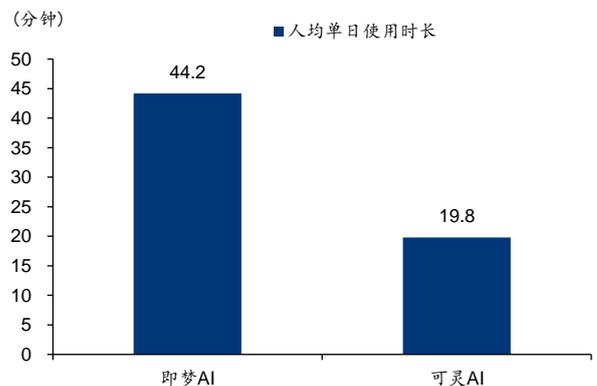
AI+视频：细分领域贵于深耕，用户习惯有望迎来变局

我们预期在垂直细分领域深耕的厂商，将会更具产品的中短期竞争优势。①快手集中算力资源，专注于视觉大模型赛道，推出了文生图与文生视频产品可图与可灵，在专注性策略下，可灵的市场认可度，以及产品的研发进度略优于即梦 AI。我们认为专注于细分赛道的厂商有望在特定领域构建出独特竞争壁垒。②字节的布局则较为广泛，涵盖了文生文、文生视频及 AI 硬件等多个领域，而多元化布局则提升跨产品协同效应，有助于提升获客能力，故而即梦 AI 在用户规模上更具优势。

我们认为短视频行业或迎来变局，AI 剪辑工具将会深度改变用户习惯。根据 QuestMobile 数据，2024 年以来，剪映等传统剪辑工具 MAU 持续下滑，市场份额被新兴 AI 剪辑工具挤占。我们认为，短视频创作生态将持续受到 AI 剪辑技术影响，在该领域领先的企业或将率先占据创作者生态优势，并进一步扩大市场份额。

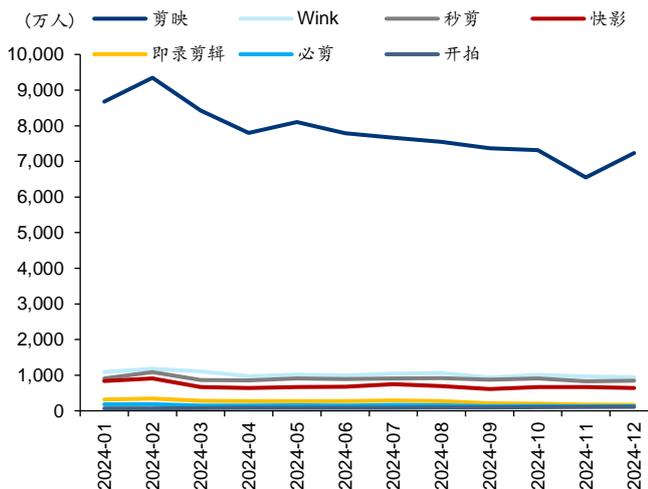
图表67：可灵与即梦 AI MAU 与 DAU 对比（24 年 12 月）


资料来源：QuestMobile，华泰研究

图表68：可灵与即梦 AI 人均单日使用时长对比（24 年 12 月）


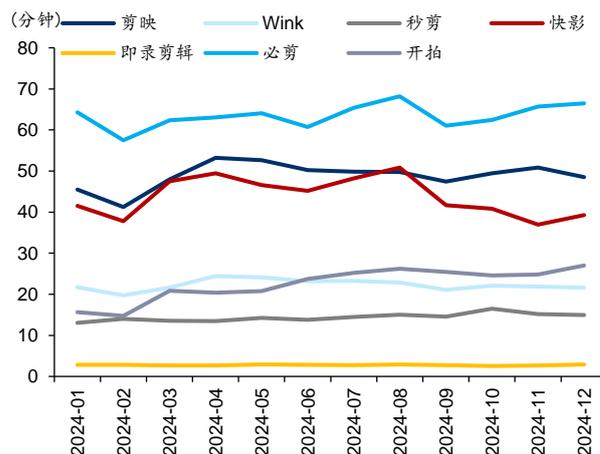
资料来源：QuestMobile，华泰研究

图表69: 剪映 MAU 相较于其他 AI 剪辑工具下降明显



资料来源: QuestMobile, 华泰研究

图表70: 剪映人均单日使用时长与其他 AI 剪辑工具对比



资料来源: QuestMobile, 华泰研究

哔哩哔哩自研大模型, 生态内应用场景丰富

差异化场景为大模型应用创造良好基础。2024年06月公司发布 Index 系列模型, 该模型是建立在开源基础上, 已经应用在平台的多场景中:

- 1) AI 字幕:** B 站具备中、英、韩、日、泰语等近 10 种语言的实时翻译能力, 准确度接近 90%。该能力已覆盖站内海外内容、主播直播间以及海外用户观看中文内容等场景, 日均支持数十万非中文稿件的翻译。
- 2) 角色扮演:** 通过在预训练中引入了较多定向清洗对话类语料, 聊天的趣味性明显更强, 对话能力、角色一致性、角色扮演吸引力三大类评测中表现优异。
- 3) AI 视频助手:** 针对一些特别长的 Vlog 视频, 或者一些科普性、教学性较强的长视频, AI 视频总结功能能够帮助用户很快地获取视频的关键信息节点, 并且可以快速将文本 Copy 下来。
- 4) AI 搜索:** 搜索 AI 助手就能为用户生成一段对应的文本内容, 这些内容会参考一部分的科普知识类视频进行归纳整理。

图表71: AI 搜索助手



资料来源: 公司官网、华泰研究

图表72: AI 视频助手



资料来源: 公司官网、华泰研究

AI+游戏：3D 生成模型助力研发进度，AI agent 重塑游戏交互

AI 正重塑游戏开发与交互，提升创作效率、丰富内容生成。在游戏行业中，通过 AI 系列技术可提升创作效率与可玩性，如在 NPC 塑造层面，动捕维度 AI 可从普通手机视频提取动作帧，快速生成匹配角色动画，捏脸维度 AI 则让研发者可通过自然语言轻松定制角色形象，互动维度 AI Agent 和智能 NPC 工具可轻易使 NPC 具备更拟人化的行为模式，提升剧情驱动能力。

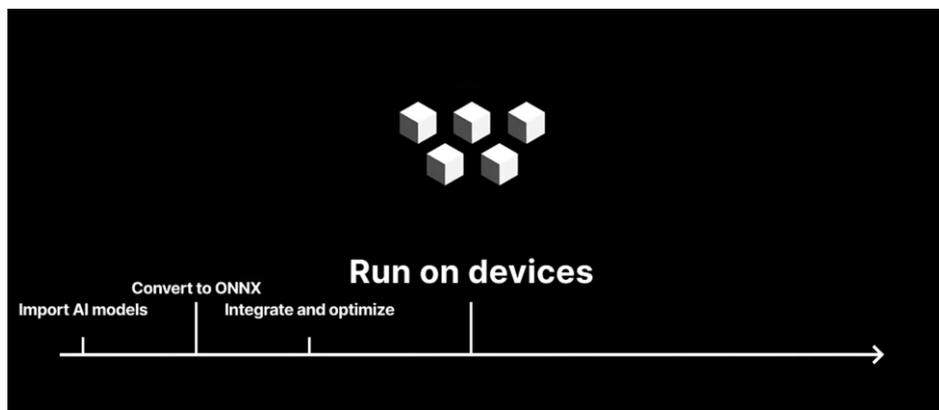
游戏 AI 层面的基数新突破：游戏世界建模层面，腾讯新推出的 GameGen-O 可模拟各类游戏引擎功能，全面优化游戏场景、剧情、关卡和动态环境的生成，提高开发效率。游戏开发引擎方面，Unity 推出的 Sentis 让开发者能直接在游戏中部署 AI 模型，实现物体识别、语音交互及智能 NPC 控制。

图表73：AI 赋能游戏

功能名称	腾讯 AI 功能介绍
视频 AI 动捕	从普通手机视频中提取动作帧并重构 3D 坐标，实现低成本、快速生成游戏角色匹配动画。
AI 生成舞蹈	将任意歌曲转换为适合其节奏和风格的舞蹈动画，并支持角色定制和风格化处理，实现低成本、快速生成用于游戏生日派对等场景的舞蹈环节。
AI 捏脸	通过自然语言获取用户喜好的风格，自动生成初步的脸型和配饰，再供用户微调，实现便捷高效的个性化角色定制。
AI Agent	将自然语言处理、记忆与行动能力封装集成，实现更拟人化、可驱动剧情发展的智能 NPC。
AI 游戏 NPC 工具	通过赋予 NPC 自然对话、情感表达、记忆学习、场景感知和个性化互动等功能，让 NPC 更像真人，从而提升游戏沉浸感，为开发者提供强大的创作工具并拓展游戏内容的可能性。
开放世界游戏生成 (GameGen-O)	模拟各种游戏引擎功能，生成游戏角色、动态环境、复杂动作等，支持交互控制，用户可以通过文本、操作信号和视频提示来控制游戏内容。
游戏 AI 引擎 (GiiNEX)	利用生成式和决策 AI 技术全流程支持游戏场景、NPC 对话、剧情、关卡和音乐等制作，大幅提升效率、降低开发门槛和成本，并推动个性化内容与整体游戏质量的提升。

资料来源：腾讯游戏学堂公众号，华泰研究

图表74：Unity 推出 Sentis，赋予开发者将 AI 模型导入游戏和应用程序中的能力



资料来源：Unity 官方平台，华泰研究

AI agent 重塑游戏交互，实现快速建立智能 NPC

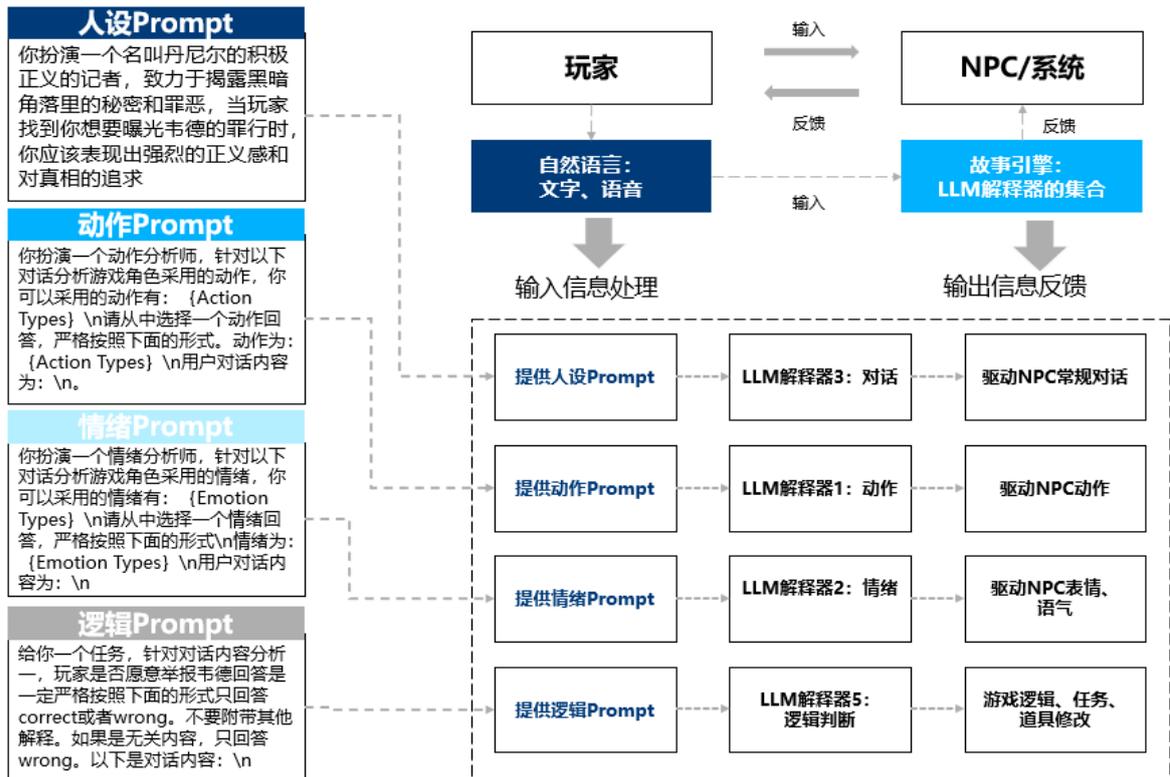
我们认为，AI agent 未来能让游戏行业迈向更智能化与个性化的发展方向：

- 1) 实现快速建立智能 NPC, 玩家与 NPC 互动更加多样。原先玩家与 NPC 互动是程式化、流程化，互动时往往只有几个选项可选择，选项引导的结局也相对单一，而 AI 则可以赋予 NPC 以灵魂，不仅让 NPC 起到驱动剧情发展的作用，也让 NPC 有更加丰富的表情、语言和任务互动。

2) AI 可以做成相关技术框架，形成调用平台给客户体验使用，让客户生成想要的 NPC 形象。为让新 NPC 的想法更快速落地，可以把故事引擎做工具化处理。以游戏中生日 party 为例，以往制作成本很高，因为需要人工建立主题、故事背景、相关任务等，而通过 AI 的故事引擎能力，客户输入故事就可自动生成需要的内容，同时也会自动匹配对应 NPC，大大降低游戏制作成本。

然而在游戏 AI agent 领域，当前还未出现现象级应用，原因如下：1) 算力成本高，2) 性能不够，无法支持大量用户用 AI 行为沟通，3) 延迟时间较高，NPC 回复慢会使玩家体验感差，4) 系统不够稳定，存在 AI 角色人设丢失的情况。

图表75: AI agent 重塑游戏交互，实现快速建立智能 NPC



资料来源：腾讯游戏研究院，华泰研究

AI+教育：积极推出大模型产品

教育的最终追求是因材施教。受限于人力投入和教学质量的矛盾，当前行业形成线下小班模式和线上大班主讲+辅导模式，然而大模型的出现有望打破成本和效率的制约。23 年以来，行业主要教育公司都推出了自己多大模型产品，部分产品已经取得了优异的商业化成绩。

图表76: 行业内主要模型

教育公司	大模型名称	发布时间	应用场景
科大讯飞	讯飞星火智慧黑板(搭载星火 3.5 认知大模型)、星火教师助手、AI 科学教育解决方案	2024 年 4 月 19 日	讯飞星火智慧黑板用于教学;星火教师助手辅助教师教学; AI 科学教育解决方案面向区域科学教育
学而思	集成九章大模型的学习机(学而思学习机经典版 2024)	2024 年 5 月 24 日	通过人工智能技术提升学习效率,如 AI 中英文作文助手、AI 听写等
网易有道	“子曰”教育大模型(发布三大 AI 创新应用: AI 全科学习助手“有道小 P” App、新一代虚拟口语教练 Hi Echo 3.0 和新一代知识库问答引擎 QAnything)	2024 年 5 月 29 日	应用于学习、翻译、口语练习等场景
猿辅导	旗下“飞象星球”批改数字作业(基于大模型技术)	2024 年 5 月 30 日	学生作业批改场景
粉笔	AI 老师“粉笔头”(职教行业垂直大模型)	2024 年 8 月 1 日	学员在题库做题、课后回顾复习等学习场景
作业帮	银河大模型	2023 年 12 月发布	作业帮 APP 搜题、辅导、学习规划等功能模块,帮助学生解决学习过程中遇到的各类问题

资料来源：多知、有道、华泰研究

在上市公司中，好未来投入较为积极。根据公司技术报告显示，学而思九章大模型模式参数超千亿，基于优质教育数据，实现了题目计算、讲解、问答等多任务持续训练和有监督微调，呈现出卓越表现。我们预计该模型仍在持续升级迭代中，回答题目的准确率预计持续提升。

图表77：九章大模型测试数据

数据集类别	数据集	GPT4-0819			MathGPT		
		总数	正确数	正确率	总数	正确数	正确率
中文	TAL-SCQ5K-CN 2K Testing	2000	555	0.28	2000	1342	0.67
	TAL-SCQ5K-CN 3K Training	3000	869	0.29	3000	2176	0.73
	TAL-SCQ5K-CN	5000	1424	0.28	5000	3518	0.70
英文	TAL-SCQ5K-EN 2K Testing	2000	1014	0.51	2000	1181	0.59
	TAL-SCQ5K-EN 3K Testing	3000	1576	0.53	3000	1850	0.62
	TAL-SCQ5K-EN	5000	2590	0.52	5000	3031	0.61

资料来源：好未来官网、华泰研究

好未来也积极开放自身能力。公司一方面开放出优质数据测试数据集，23年9月好未来在GitHub社区开源了MathGPT模型训练测试数据集——TAL-SCQ5K-EN和TAL-SCQ5K-CN，包含5K道英语数学竞赛题目（3K道用于训练，2K道用于测试）。这些题目采用多项选择题形式，涵盖了小学数学领域的各个主题。另一方面通过好未来开放平台，提供解题解决方案和批改解决方案。

图表78：AI开放平台



资料来源：好未来官网、华泰研究

图表79：大模型开放能力

功能	应用场景
学生自主练习	在家中完成家庭作业后，学生可以通过九章批改快速了解自己的作业情况，对错误部分进行及时纠正，从而提升学习效率和自主学习能力。这种即时反馈机制有助于学生更好地掌握知识点，形成有效的学习闭环
教师批改平台	教师可以利用九章批改快速处理大量学生作业，节省时间，同时获取每个学生详细的作业分析报告。基于这些报告，教师可以进行更加针对性的个性化辅导，帮助学生解决学习中的瓶颈问题，提升教学质量
教师备课助手	教师可以利用九章大模型解题解决方案，快速获取各类题型的解析和解题思路，丰富教学资源，提升备课效率。同时，通过对学生解题情况的分析，教师可以更精准地掌握学生的学习状态，从而制定更有效的教学策略
学生答疑平台	无论是数学中的复杂方程，还是物理中的力学难题，亦或是语文中的阅读理解，九章大模型解题解决方案都能提供详尽的答案解析和解题思路，帮助学生轻松应对日常作业中的各类挑战，同时减轻家长的辅导压力

资料来源：好未来官网、华泰研究

有道：坚持“场景为先”，推出自研教育垂类大模型“子曰”。「子曰」教育大模型采用 Transformer 神经网络架构和技术，集成多学科知识和海量数据，能够实现智能问答、文本生成、情感分析、机器翻译等多种功能。网易有道将「子曰」教育大模型融入实际教学场景和教育数据的深度挖掘运用，以自研基座大模型为基础，打造了全球首个虚拟人口语私教 Hi Echo、新一代 AI 全科学习助手小 P 老师、智慧体育等 10 余个垂类教育场景的应用，成为国内在教育场景中落地应用最快的企业。

图表80：“子曰”大模型及其应用



资料来源：网易有道公众号，华泰研究

针对翻译、答疑等细分场景，有道分别推出专有模型，并已在应用端落地：

- 1) **子曰翻译大模型**：2025 年 1 月基于有道子曰的翻译大模型 2.0 正式上线，相较于 1.5 版本（2024 年 5 月）在数据、算法以及评估三个关键维度进行了全方位优化升级，在翻译质量、效率以及鲁棒性等方面实现了质的飞跃，能够为用户带来更优质、精准的翻译服务。目前子曰翻译大模型 2.0 已落地有道词典 APP、有道翻译桌面版、有道翻译网页端。
- 2) **有道子曰-o1 推理模型**：2025 年 1 月有道正式推出国内首个输出分步式讲解的推理模型“子曰-o1”，并已开源。作为 14B 轻量级单模型，子曰-o1 支持在消费级显卡上进行部署，采用思维链技术，能够提供细致解题过程，以强逻辑和推理能力，实现更高的解题准确性，并提供中文逻辑推理。基于有道多年来在教育领域的的数据资源积累，子曰-o1 使用了大量的教育领域学生试卷习题为训练样本，从而提升在教育场景应用的准确性。当前，子曰-o1 已在“有道小 P”中落地应用，支持其实现“先提供解析思路、再提供答案”的答疑过程，引导学生用户主动思考、调用知识储备自主解决问题。


图表81：子曰翻译大模型 2.0 测试评分表现优秀

测试集-中译英	子曰翻译大模型 1.5		子曰翻译大模型 2.0	
	comet-kiwi-xxl	ZiYue MT reward model win/tie/loss %	comet-kiwi-xxl	ZiYue MT reward model win/tie/loss %
WMT17	82.9	/	83.5	51.77/26.94/21.29
WMT18	80.97	/	81.59	53.69/24.97/21.33
WMT19	81.17	/	81.91	58.95/19.55/21.5
WMT20	79.79	/	80.46	59.95/20.9/19.15
WMT21	78.53	/	79.3	62.32/15.61/22.07
WMT22	77.49	/	78.49	57.92/21.92/20.16
WMT23	79.22	/	80.11	54.35/25.0/20.65
IWSLT17	83.68	/	84.23	39.27/44.41/16.31
FLORES200_dev	87.11	/	87.51	39.12/44.13/16.75
FLORES200_test	87.06	/	87.56	35.87/47.04/17.09
平均	81.792	/	82.466	51.32/29.05/19.63

测试集-英译中	子曰翻译大模型 1.5		子曰翻译大模型 2.0	
	comet-kiwi-xxl	ZiYue MT reward model win/tie/loss %	comet-kiwi-xxl	ZiYue MT reward model win/tie/loss %
WMT17	85.88	/	87.19	59.67/21.44/18.89
WMT18	85.79	/	86.66	53.08/27.25/19.67
WMT19	84.74	/	86.37	63.04/14.92/22.03
WMT20	87.41	/	88.61	68.12/10.01/21.86
WMT21	83.53	/	85.04	6.67/12.87/20.46
WMT22	85.43	/	87.02	55.28/26.41/18.31
WMT23	80.38	/	81.73	59.45/18.23/22.32
IWSLT17	84.53	/	85.2	47.43/27.21/ 25.36
FLORES200_dev	87.92	/	89.11	60.28/20.46/19.26
FLORES200_test	88.24	/	89.5	60.97/20.26/18.77
平均	85.385	/	86.643	53.40/19.91/20.69

资料来源：网易有道公众号，华泰研究

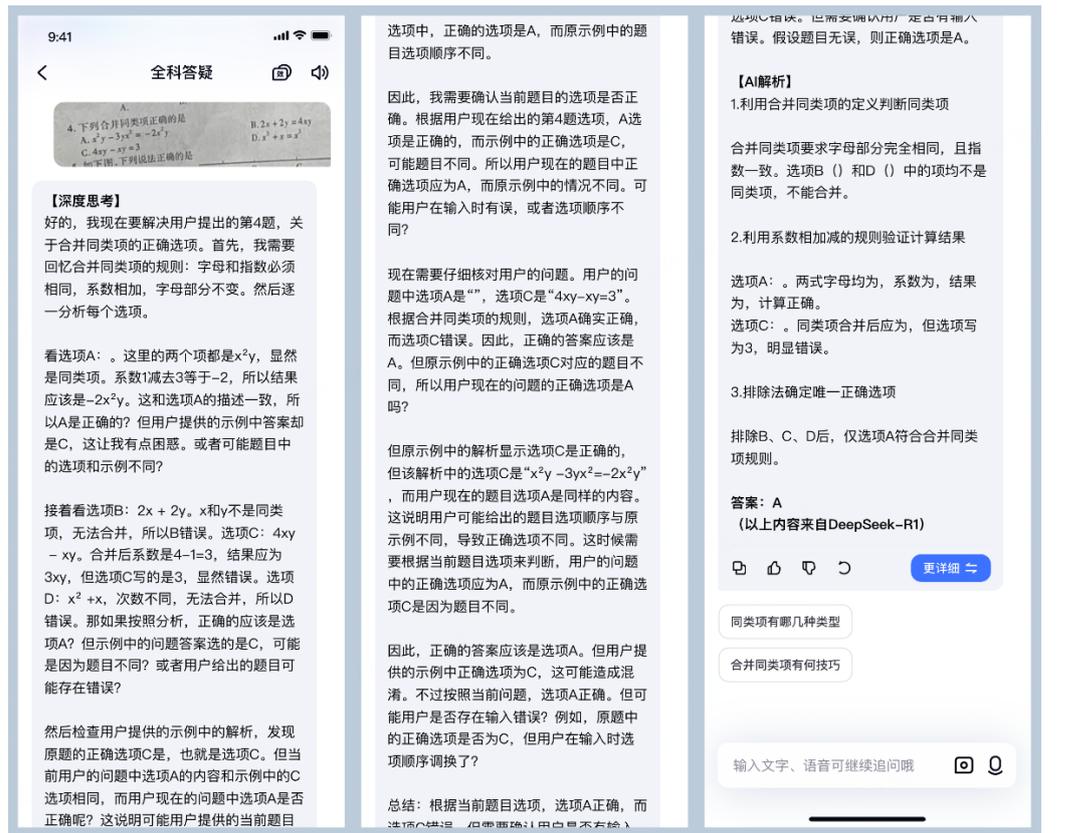
图表82：子曰-o1 测试评分

	k12-math(内部测试集)	MATH500	GAOKAO			
			客观题		主观题	
			Sci.Math	Hum. Math	Sci.Math	Hum. Math
GPT-4o	59.41	60.3*	77.57	85.32	55.92	62.63
o1-preview	/	85.5*	/	/	/	/
o1-mini	85.71	90.0*	96.73	98.62	73.61	81.79
QwQ-32B	83.17	90.6*	97.2	99.08	82.2	85.01
Sky-T1-32B-Preview	83.17	82.4*	97.2	97.25	76.66	76.51
Qwen2.5-14B-Instruct	65.35	76.2	83.64	89.45	53.66	62.47
Confucius-o1-14B	81.19	81.6	93.93	98.96	68.46	76.68

资料来源：网易有道公众号，华泰研究

此外,有道、好未来成为首批接入 DeepSeek 的教育公司,有望全面提升个性化学习体验,加速 AI+教育产品商业化落地。2月6日有道宣布全面接入 DeepSeek 的推理能力,有道小 P、Hi Echo、有道智云、QAnything 等产品将陆续升级,并将推出融合推理大模型能力的智能硬件新品。2月8日学而思宣布旗下学习机、学练机等智能教育硬件产品将接入 DeepSeek,通过植入“深度思考模式”,全面升级用户的 AI 学习体验,重新定义智能辅导场景,新功能预计将在 2 月内陆续于相关机型上线。我们认为有道、好未来来自研大模型的丰富垂直数据储备结合 DeepSeek-R1 的通用推理能力有望进一步加速 AI 教育产品商业化落地。

图表83: 有道小 P App 目前已接入 DeepSeek



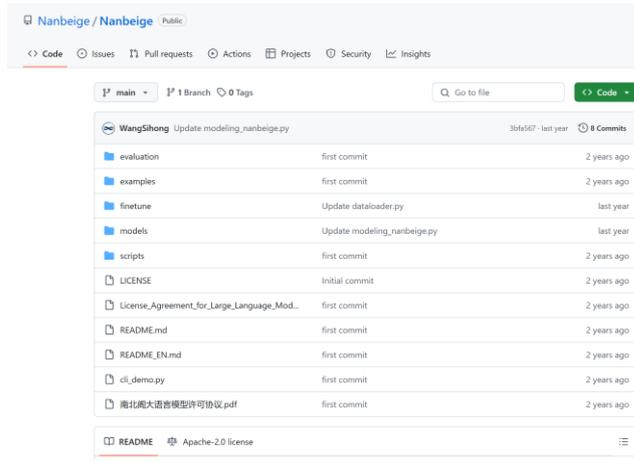
The screenshot displays a mobile application interface for a math problem. The problem asks for the correct option among four choices (A, B, C, D) regarding the combination of like terms. The AI solution, titled '【深度思考】' (Deep Thinking), provides a detailed step-by-step analysis. It identifies the correct option as A, explaining why the other options are incorrect based on the rules of combining like terms (same variables and exponents, coefficients can be added/subtracted). The solution also includes a '总结' (Summary) at the end, stating that option A is correct based on the current problem's context.

资料来源: 网易有道公众号, 华泰研究

AI+招聘: 提升招聘效率, 提升用户体验

线上招聘平台招聘数据丰富, 为大模型开发创造良好基础。BOSS 直聘作为行业头部公司开源南北阁大模型, 赋能产业发展。Nanbeige-16B (南北阁-16B) 是南北阁大模型实验室研发的 160 亿参数规模的大语言模型, 采用了 2.5T Tokens 进行预训练, 数据包含大量互联网高质量语料、各类书籍、代码等领域脱敏文本, 在各个权威测评数据集上都取得了不错的效果。

图表84： 南北阁开源大模型



资料来源：GitHub、华泰研究

南北阁算法经过备案，投入使用。南北阁文本生成和智能对话算法已经通过网信办备案，在实际招聘流程中得到充分使用。文本生成模型，可以帮助求职者自动生成职位信息，提升招聘效率。对话算法能够提供快速的反馈模式，当招聘者不在线的时候，依然可以通过个性化沟通了解求职者的情况，提升求职者的体验，减少求职过程中不被回复的情况。

图表85： BOSS 直聘模型应用

算法名称	南北阁文本生成算法	南北阁智能对话算法
算法基本原理	根据用户填写的简历信息、职位信息数据，拼接组成提示词输入自研的南北阁大模型，程序根据模型生成的文本执行业务相关操作。	通过交互对话的方式为用户提供职位/求职者搜索及筛选、沟通管理、信息完善等功能。算法在线应用时，会使用用户实时输入的文本数据，在安全过滤的基础上，结合对话场景等向用户生成答复。
算法运行机制	南北阁文本生成算法使用自研的南北阁模型，在业务中，算法服务根据输入的参数查找对应的数据，拼接成提示词输入南北阁大模型，大模型生成的文本被程序解析，返回给用户或提取参数用于后续的流程。	对话生成算法采用了 Transformer Decoder 结构，在用户输入问题后，将依据意图分类模型识别出意图类型，生成符合用户求职招聘需求的答复。在此过程中，我们通过输入内容审核、输入内容模型审核及输出内容模型审核三重过滤的方式保障内容安全。
算法应用场景	BOSS 直聘中的文本内容编辑场景，如填写简历信息、职位信息、交流信息等。	BOSS 直聘平台中消息列表、个人主页、信息修改等页面的信息内容服务，如 AI 招聘秘书、智能沟通助手、智能 AI 助手、南北阁求职助手等。
算法目的意图	帮助用户生成、调整、优化简历、职位、交流语等文本信息，提升求职招聘效率。	为用户提供职位/求职者搜索及筛选、沟通管理、信息完善等功能，提升用户招聘求职体验。

资料来源：公司官网、华泰研究

行业中其他公司也积极布局 AI 相关应用。猎聘发布的《AI 面试应用现状与趋势分析报告》显示，在过去的一年里，企业在猎聘·Doris 上创建 AI 面试的次数同比增长达 448.2%。AI 面试，正在渗透到更多的行业。智联招聘推出 AI 易面 3.0，该产品可以做到数千位候选人同时进行面试作答时，系统依旧流畅运行，服务企业春招/秋招中短时间内高效实现对成百上千位候选人的精准初筛。

图表86：猎聘 AI 智能面试

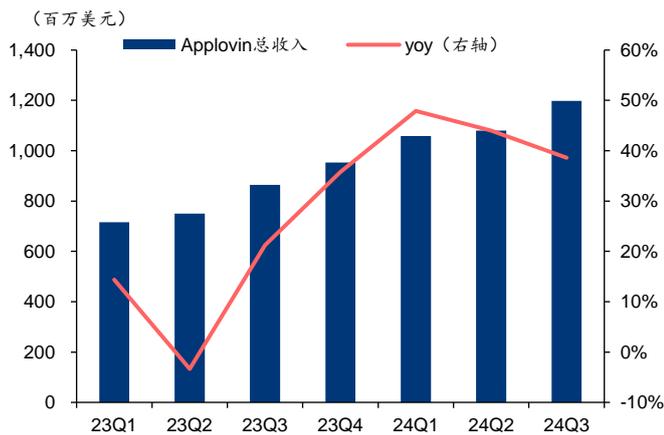


资料来源：公司业绩交流资料、华泰研究

AI+广告：赋能互联网营销平台提质增效

AI+广告融合在海外加速兑现。美国移动营销平台 Applovin 自 23 年开始升级 AI 广告引擎 AXON 2.0 以来，实现了更高的自动化程度、更高的准确性、更好的广告投资回报率 (ROAS)、更敏捷的营销活动以及全球范围的投放，有效地推动了用户群体的增长和货币化变现效率的提升，驱动收入同比增速由 23Q3 的 21.2% 提升至 24Q3 的 38.6%，实现加速增长。类似地，汇量科技旗下 AI 驱动的程序化广告平台 Mintegral 充分利用机器学习的强大算法，通过对用户行为的深度学习和行为模型预测，实现了广告推荐的个性化和智能化，有效提升了广告投放的效果。Mintegral 于 2023 年 5 月上线 Target ROAS 智能出价功能，将买量模式升级为基于广告主 ROI 的智能出价模式，并于 2024 年 7 月推出针对事件 (Engagement) 进行投放优化的 Target CPE (Cost per Engagement) 功能，用于帮助广告主有效控制事件成本以及优化广告效果。随着智能出价能力显著提升，ROAS 的智能出价模式成为开发者在 Mintegral 的主流选择，推动 23Q4-24Q3 Mintegral 收入同比分别+19.5/25.4/32.4/57.6%，环比+0.7/10.3/12.9/25.7%，并推动公司从游戏品类向非游戏品类（电商、社交等）快速扩张。

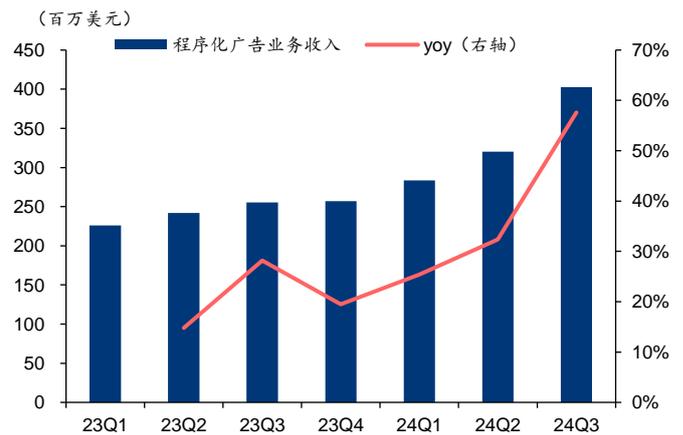
图表87：AXON 2.0 推动 Applovin 收入同比持续高增



注：收入口径为净收入。

资料来源：公司公告，华泰研究

图表88：汇量科技程序化广告业务收入持续高增



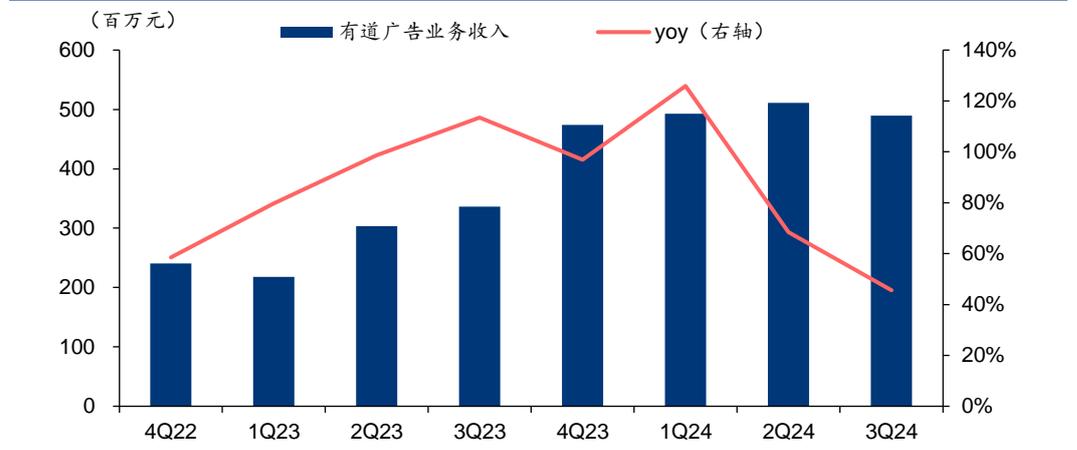
注：收入口径为毛收入。

资料来源：公司公告，华泰研究

有道通过“垂直人群+技术赋能+周期适配”三位一体策略实现精准高效投放，成为 AI 商业化浪潮中的关键服务商。网易有道在线营销业务在 AI 赋能下连续 8 个季度（截至 24Q3）实现 40%+ 的同比增长，其中 RTA 广告业务收入屡创新高，主要得益于：

- 1) 精准的用户定位，聚焦垂直人群：**基于多年教育服务和资源的积淀，网易有道积累了 13 亿用户，特别是 18 至 35 岁的年轻群体，包括超 4000 万大学生、6000 万职场人士。为有效利用数据，进行深入的人群画像分析，实现高效匹配，公司推出数据管理平台（DMP）网易探数，不断收集和整合了来自不同流量的超过 5000 种个性化标签，通过对用户群体的深入理解和兴趣的精准分析，公司能够将用户需求与客户产品特点相结合，实现高效的匹配和串联，从而显著提高转化效率。
- 2) AI+RTA 技术融合，优化投放技术：**有道融合了 AI 广告算法模型和 RTA 技术，能够准确的判断用户输入的信息，清楚了解用户在什么场景下最有可能形成转化，从而帮助广告主用最合理的素材、触达时间点等策略转化用户。
- 3) 应用产品投放经验丰富：**有道自身推出过一众教育类工具和 AI 应用，通过将多年积累的对应用产品的生命周期、获客方式、营销目标等的深刻理解融入到广告投放模型中，有道能够结合技术优势快速捕捉市场机会，推动广告业务快速增长。

图表 89：有道广告业务收入连续 8 个季度同比增速超 40%（截至 24Q3）



注：收入口径为毛收入
资料来源：公司公告，华泰研究

bilibili 推出星辰 AI 提升广告效率。2025 年 1 月 2 日正式推出了其商业化 AIGC 平台——“星辰 AI”，标志着 B 站在人工智能领域迈出了至关重要的一步，也为其广告业务和内容生态注入了全新活力。主要由两方面功能：一是通过智能生成广告素材，为品牌和广告主提供高效的创意工具，助力平台的商业化探索。二是与 B 站的内容生态深度融合，实现从广告主到内容创作者、再到用户的价值链闭环。

图表90：星辰 AI 模型

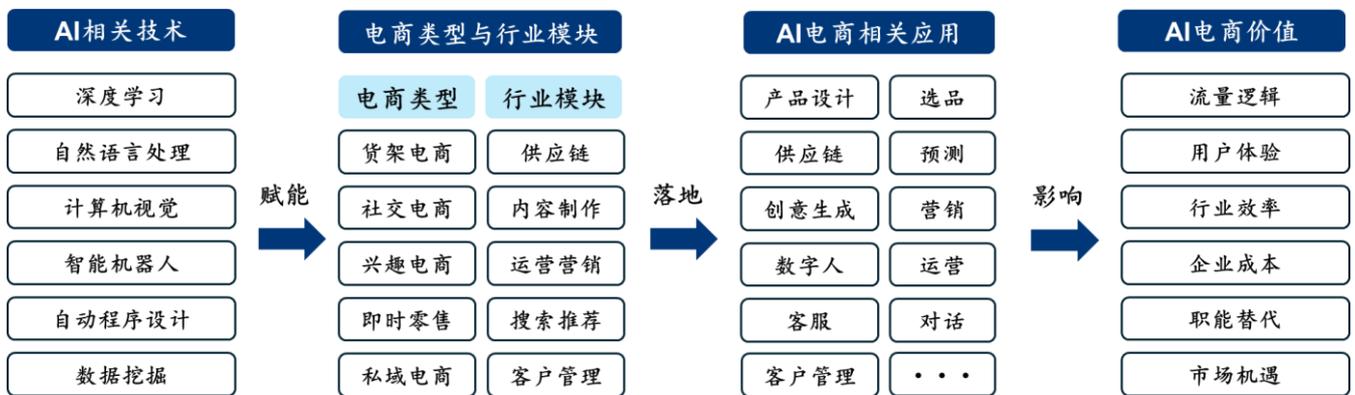


资料来源：公司公告、华泰研究

AI+电商：淘天多年数据沉淀打造增长基础

AI+电商是从技术到价值的模块化落地与赋能。随着 AI 相关技术发展，我们认为其有望赋能货架电商、社交电商、兴趣电商、即时零售和私域电商等多种电商类型，并在设计、选品、供应链、营销、客服和数字人等领域落地应用，最终对流量逻辑、用户体验、行业效率、企业成本等电商行业价值产生正向影响，在中长期为行业实现降本增效作用。

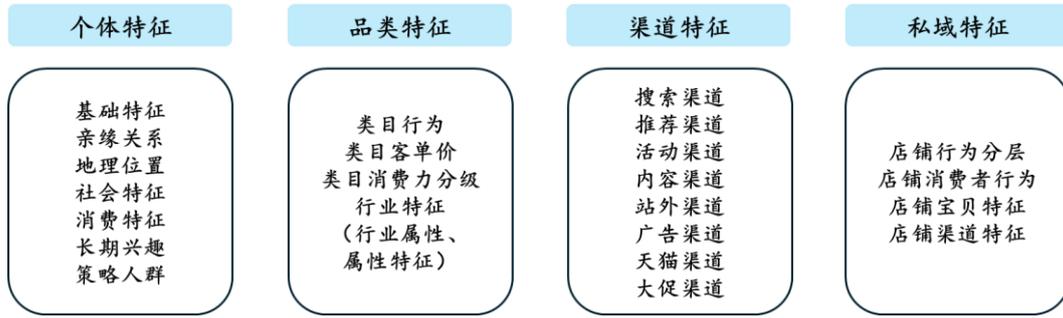
图表91：AI+电商从技术到价值的赋能逻辑



资料来源：亿欧智库，华泰研究

达摩盘是阿里妈妈基于商业化场景打造的数据管理合作平台，拥有消费行为、兴趣偏好、地理位置等众多数据标签。品牌通过达摩盘可以实现各类人群的洞察与分析，潜力客户的挖掘；通过标签市场快速圈定目标人群，建立个性化的用户细分和精准营销；或通过第三方服务应用市场，解决个性化的营销需求，实现“千人千面”的营销效果。达摩盘标签维度包含：用户特征、品类特征、渠道特征和私域特征等，并在 618、双 11 等大促节点推出场景特定标签。

图表92：达摩盘基于四大维度提供精细化标签



资料来源：数麦电商，华泰研究

LMA 大模型技术旨在提升人货匹配效率。据阿里妈妈数字营销，2024 年 3 月，阿里妈妈正式发布 LMA 大模型技术：该技术的最大突破在于基于多模态大模型的能力，实现了机器对消费者消费需求的精准把握，从而提升人货匹配效率。LMA 是在阿里自研的大语言模型基础上，针对电商场景实现超大规模多模态表征的预训练，能够对商品的文本、图像、视频等多种模态有深入的融合理解，再结合海量消费者行为数据进行模型微调，使得模型能够更了解用户。

图表93：阿里妈妈发布 LMA 大模型技术



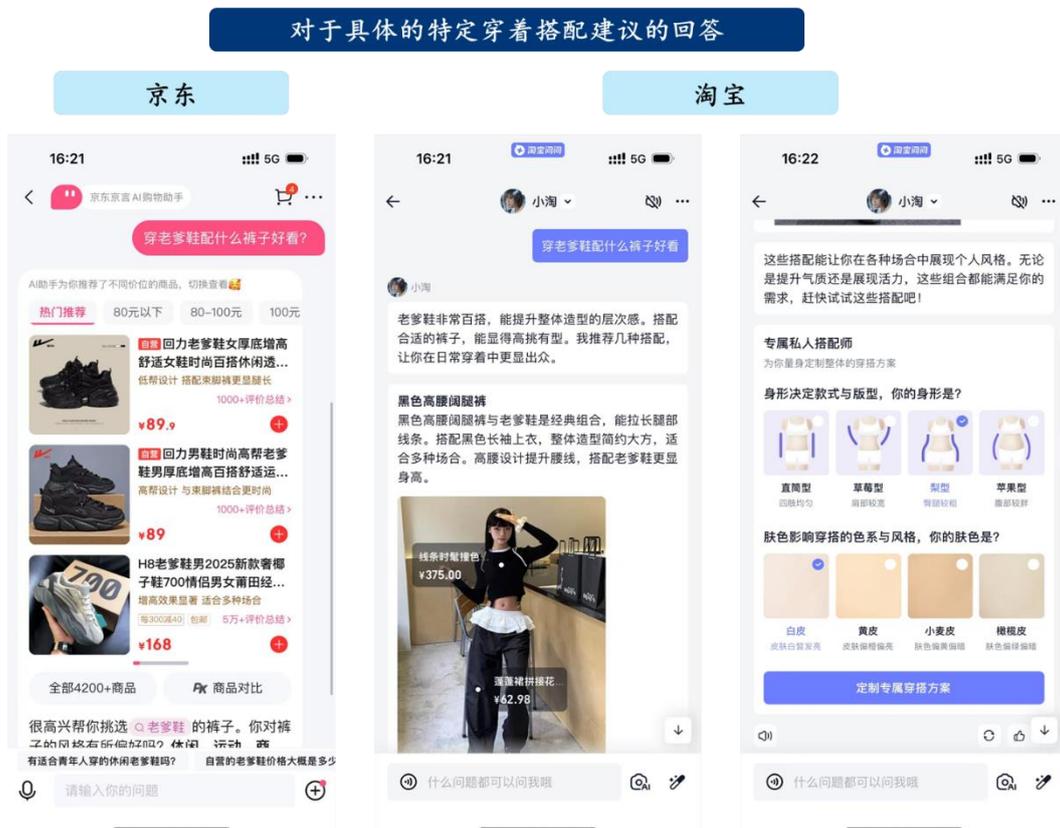
资料来源：阿里妈妈数字营销，华泰研究

达摩盘 AI 圈人结合了 LMA 技术与达摩盘数据标签。传统营销中，商家往往需要经历冗长的数据分析、繁杂的人群定位，和较长的试错过程，通过达摩盘全新推出的“AI 圈人”功能，商户只需通过日常对话般的沟通，达摩盘便能通过 AI 技术，智能匹配并定位到最匹配的潜在客户群体。为了降低中小商家的使用成本，达摩盘基于 LMA 模型和丰富的消费人群识别能力，通过简单对话和提问形式，由 AI 智能解读商家圈人需求，生成丰富的人群，助力全量商家提升精准人群定向能力。使用过程由三步构成：1) 输入人群诉求。右侧对话框中输入圈人指令，AI 会基于输入的指令智能识别圈人诉求，并反馈需求理解；2) 生成人群并提供人群解读。选择营销目标后点击“帮我生成人群”，系统会进行人群圈选，人群生成后会提供人群规模、人群营销效率及人群画像洞察。3) 保存人群。确认圈选结果，并保存人群，人群将保存在“我的人群”列表中，关联计划后即可投放；据阿里妈妈数字营销公众号，达摩盘 AI 圈人功能在 2024 年双十一期间 AI 人群投放效率对比其它类型投放增加了 19.8% 的点击率和 28.1% 的转化率。

在 To C 端应用上，AI 购物助手为初步落地方向。自 2023 年下半年以来，AI 导购功能已成为国内各大电商争相布局于 AI 电商的重要领域。2023 年 9 月，淘宝 APP 推出“淘宝问问”，通过结合用户的输入和深度合成算法，为用户提供更贴合其消费习惯的商品推荐和内容。2023 年 11 月，京东 APP 上线“京言”AI 助手测试版。电商平台正试图通过结合自身积累的 AI 能力与丰富 SKU 供给，为消费者提供更为多样化的购物体验。

淘宝问问与京东京言初步实测对比发现，淘宝 AI 导购改造相对激进。经过我们实测，淘宝问问和京东京言已均可通过搜索栏搜索相关关键词进入功能界面。在某个具体的穿着搭配建议（提问：穿老爹鞋配什么裤子好看？）的回答上，“淘宝问问”给出了不同性别不同风格的搭配，简单阐述了其优点，还附上了模特穿搭图和具体链接，并可根据选择个人身形和肤色进一步选择具体定制穿搭方案。而“京东京言”主要推荐了老爹鞋，并未给出裤子搭配建议。淘宝在具体穿着搭配上的 AI 功能相对智能和多样化。

图表 94：淘宝问问和京东京言对于具体特定穿搭建议的回答



资料来源：淘宝 APP，京东 APP，华泰研究

而在对于泛化的办公产品建议的回答上（提问：过完年要上班了，能给我推荐一些适合办公的数码产品吗？），“淘宝问问”分别推荐了电脑、打印机、路由器等带有具体型号的产品，并带有各自产品特色介绍，而“京东京言”按照价格带给出了部分产品推荐，并询问具体需要哪些类型的办公产品。总体而言“淘宝问问”更侧重于详细丰富的单品描述，“京东京言”更专注于不同价位的推荐。

图表95： 淘宝问问和京东京言对于泛化的办公产品建议的回答



资料来源：淘宝 APP，京东 APP，华泰研究

在具体结合上，相比于 C 端，AI 技术在电商 B 端的应用已有较为明晰的落地路径。我们认为，针对国内电商业务而言，尽管大范围 and 普适性的 To C 端 AI 应用落地或仍需时日，但其在 To B 端的应用已有较为明晰的落地，如帮助商户进行商品信息物料准备进而降低经营成本和优化广告营销的推荐效率等，AI 大模型能力的逐步演进或可以为各大电商平台在中长期增长提供进一步助力。我们认为 AI 技术在 B 端的应用有望一定程度上拉平中小商家与头部商家的营销能力差距，从而激发中小商家投流投广兴趣，使其将更多预算投入该电商渠道，并在中长期逐步带动平台货币化释放节奏。

投资观点

国内 AI 从技术突破到场景落地的路径逐渐清晰，有望迎来商业化加速。互联网企业在本轮 AI 科技周期中扮演重要参与者，我们认为互联网 AI 板块的投资主要分为两条大的主线：

主线 1：投资具备优质底层模型+算力（云）基础设施+丰富下游应用生态的互联网大厂，首推阿里、腾讯，建议关注百度。此类大厂当前正在通过 capex 引领新一轮国内 AI 基础设施建设周期（对标美国 23-24 年间 AI 进程），在优质算力（云）基础设施下担当中国 AI 科技周期启动的“卖铲人”。同时这些大厂下游已经具备完整的应用生态和丰富的用户基础，可以在已有的应用场景中无缝加入 AI 功能，依靠流量优势迅速启动。而大厂的多元业务生态非常有助于开启下一个阶段的 AI Agent 应用，将有望凭借其生态优势在 AI 的新一轮竞赛中取得领先身位，并中长期赋能核心业务的增长。

主线 2：投资卡位优质垂直场景+AI 产品布局充分+商业化落地确定性较高的“小而美”细分赛道龙头，首推美图、快手、哔哩哔哩、有道，建议关注好未来、BOSS 直聘。此类厂商原有业务卡位优质的细分垂直场景，如图像编辑、短（长）视频、教育和招聘等等，已经具备独有的数据、用户资源。随着 AI 技术平权等行业趋势，竞争从底层模型转向下游垂直场景应用，在 AI 端布局充分，并具备优质产品力的企业有望获得新增长曲线。同时我们也认为从 AI 应用商业化落地层面来看，垂直场景 AI 应用相较搜索、社交等通用 AI 更快，因为通用 AI 当前更聚焦于流量入口的争夺在变现层面并不激进，而细分垂直场景有望凭借提供的生产力改善、增值体验等更快实现财报端的兑现。

风险提示

模型性能迭代不及预期：模型的训练迭代及推理进展不及预期，多模态融合与算力成本优化不及预期。

商业化进展缓慢：商业化推进进程不及预期，用户付费意愿较低，用户对 AI 应用接受度存在不确定性。

监管收紧：监管政策收紧，有关数据隐私、信息安全等议题，还包括 AI 生成内容合规性要求趋严，尤其在教育、医疗等敏感领域。

本研报中涉及到未上市公司或未覆盖个股内容，均系对其客观公开信息的整理，并不代表本研究团队对该公司、该股票的推荐或覆盖。

图表 96：重点推荐公司一览表

股票名称	股票代码	投资评级 (当地币种)	最新收盘价	目标价	市值 (百万)	EPS (元)				PE (倍)			
			(当地币种)	(当地币种)	2023	2024E	2025E	2026E	2023	2024E	2025E	2026E	
阿里巴巴-W	9988 HK	买入	127.70	185.10	2,426,540	7.96	8.76	8.90	9.84	14.81	13.46	13.25	11.98
腾讯控股	700 HK	买入	502.00	637.28	4,618,833	12.55	21.14	23.38	26.76	36.93	21.92	19.83	17.33
百度集团	BIDU US	买入	95.87	113.70	33,617	10.25	9.63	9.62	10.75	67.16	71.50	71.51	64.00
美图公司	1357 HK	买入	5.27	7.49	24,024	0.09	0.18	0.18	0.25	57.21	27.41	26.57	19.56
快手-W	1024 HK	买入	56.80	68.73	244,441	1.49	3.57	4.23	5.24	35.27	14.71	12.41	10.01
哔哩哔哩	BILI US	买入	19.87	34.40	8,262	-8.13	-0.05	3.91	5.02	-17.55	-2,718.50	36.44	28.41
有道	DAO US	买入	7.83	12.04	935.83	-4.20	0.85	1.40	1.96	-13.38	66.29	40.15	28.62
好未来	TAL US	买入	13.10	15.08	7,933	-0.13	0.40	0.94	1.18	-104.61	33.15	13.99	11.08
BOSS 直聘	BZ US	买入	20.01	22.71	8,707	4.96	6.27	7.39	7.98	28.99	22.91	19.43	18.01

资料来源：Bloomberg，华泰研究预测

图表97：重点推荐公司最新观点

股票名称	最新观点
阿里巴巴-W (9988 HK)	<p>阿里巴巴 3QFY25 总收入 2,802 亿元，同增 7.6%，符合 VA 一致预期的同增 7.0%；经调整 EBITA 为 549 亿元，同增 3.8%，经调整 EBITA 利润率 19.6%，优于 VA 一致预期的 19.2%，主要得益于电商和云业务利润好于预期。本季度淘天变现率同比有所提升，改善节奏好于我们预期。我们预计提佣举措与全站推广工具的渗透提升将助力淘天电商业务变现率在 4QFY25-2QFY26 延续同比改善趋势，而中长期来看，淘天电商业务回归品牌经营主阵地策略的有效性 & AI 大模型对电商广告效率的赋能，是其变现率未来持续改善的主要驱动力。我们看好阿里云持续受益于国内 AI 应用需求增长而兑现收入提速的基本面改善趋势，同时期待大模型与阿里集团其他各业务线的融合助力其改善用户体验并强化市场份额。维持“买入”评级。</p> <p>我们调整 FY25-27 非 GAAP 归母净利润预测 2.4/-3.1/-2.1% 至 1,609/1,780/2,046 亿元。估值窗口调至 FY26，基于 SOTP 美股目标价 190.4 美元（前值：138.5），港股 185.1 港元（前值：134.7），目标价提升主因云及电商估值倍数上调，对应 20.4/18.3/16.0 倍 FY25-27 非 GAAP 预测 PE。</p> <p>风险提示：AI 云业务拓展不及预期；电商行业竞争加剧。</p> <p>报告发布日期：2025 年 02 月 21 日</p> <p>点击下载全文：阿里巴巴-W(9988 HK,买入);BABA US,买入): 3QFY25：电商业务变现好于预期，云业务收入加速增长</p>
腾讯控股 (700 HK)	<p>腾讯 4Q 业绩：营收同比增长 11.1% 至 1724 亿元，高于 VA 一致预期 2%（下同）。分业务看，增值服务、在线广告、金融科技收入分别同比增长 14.4%、17.5%、3.2%，对应毛利率同比抬升 2.2/0.9/3.3pct。经调整净利润同比增长 29.6% 至 553 亿元，超预期 2%。公司表示 24 年与 AI 相关的云收入有所翻倍，并于 4Q 加大 GPU 采购力度以夯实算力基础。展望 1Q，我们预期游戏业务有望受春节活动驱动。我们看好腾讯游戏的长线运营能力，25 年广告和云收入有望在 AI 催化下加速增长，维持“买入”评级。</p> <p>展望 25-26 年，我们分别上调腾讯收入预测 2.6/2.8%，主系 AI 产品催化下公司业务持续受益，上调经调整净利润 2.1/1.0% 至 2524 亿与 2889 亿元。我们引入 27 年新值，对应总营收与经调整净利润分别为 8872 亿和 3194 亿元。我们根据 SOTP 估值，给予目标价 637.28 港币（前值 582.08 港币），对应腾讯 2025 年 21.6x PE 估值（对比 Meta 23.9x 2025 PE，3 月 19 日收盘价和彭博一致预期）。2025 年估值倍数上调，主系腾讯 25 年广告和云收入有望加速，其中广告收入 25-27 年有望保持 16-18% CAGR。维持“买入”评级。</p> <p>风险提示：支付增速低于一致预期，经营费用超预期，游戏递延周期较预期更长，宏观恢复和广告增长不及市场预期。</p> <p>报告发布日期：2025 年 03 月 20 日</p> <p>点击下载全文：腾讯控股(700 HK,买入): 4Q 收入超预期，关注 AI 产品催化</p>
百度集团 (BIDU US)	<p>4Q24 百度集团总收入 341 亿元，同比-2.4%（VA 一致预期：-4.8%，口径下同）。非 GAAP 净利润 67 亿元，非 GAAP 净利率 19.7%（预期：14.7%），同比下降 2.5pp，主因汇兑收益好于预期。4Q24 云业务收入提速，阶段性印证国内 AI 需求的强劲表现，后续建议关注核心广告业务的收入增速修复节奏及搜索广告市场的竞争格局演变趋势。“买入”。</p> <p>我们调整公司 2025/2026 年非 GAAP 净利润预测-0.7/0.5% 至 270/302 亿元，并引入 2027 年预测 319 亿元。我们基于 SOTP 估值的目标价为 113.7 美元（前值：115.4 美元），110.4 港元（前值：112.3 港元）。</p> <p>风险提示：AI 大模型应用进程不及预期；广告业务修复弱于预期。</p> <p>报告发布日期：2025 年 02 月 19 日</p> <p>点击下载全文：百度集团-SW(9888 HK,买入);BIDU US,买入): 4Q24：云收入加速增长，关注广告业务后续增速修复节奏</p>
美图公司 (1357 HK)	<p>美图公司 2H24 收入 17.2 亿元，yoy+19.8%，低于 VA 一致预期 6%（下简称预期），主要受美业业务调整带来的收入缩减影响，而主业影像业务超预期 7%。2H24 调整后归母净利润分别为 3.1 亿元，超预期 11%；全年实现调整后归母净利润 5.9 亿，达到业绩指引上限，受益于收入结构变化下高毛利产品占比提升与运营杠杆起效。美图 AI 技术迭代下生产力场景与全球化双轮驱动，持续深化商业潜力，打开中长期增长广阔空间，维持买入评级。</p> <p>我们预测公司 25-27 年收入分别为 43/54/68 亿元（25-26 分别下调 9%/6%），下调理由为公司收入结构变化，美业业务战略性聚焦高毛利产品导致收入降低；预测调整后归母净利润分别 9.0/12.1/15.9 亿元（25-26 均上调 2%），上调理由为高毛利影像主业占比提升，或将带来盈利能力上行。参考同业 AI 可比公司及付费订阅类可比公司 2025 年平均 PE 为 29.7 倍，考虑公司 AI 持续赋能业务进入高速增长期，业绩兑现能见度高，且公司作为 AI 商业化落地稀缺标的，给予估值溢价至 2025 年 35x PE 估值（前值为 25 年 25 倍 PE），目标价为 7.49 港币（前值：4.95 港币），买入评级。</p> <p>风险提示：AI 算力不足；图像大模型进展不及预期；行业内竞争加剧。</p> <p>报告发布日期：2025 年 03 月 19 日</p> <p>点击下载全文：美图公司(1357 HK,买入): 收入结构变化利润率加速上行</p>
快手-W (1024 HK)	<p>快手 4Q 业绩：营收同比增长 8.7% 至 353.8 亿，符合 VA 一致预期 357 亿元（下同）。毛利率同比改善 0.9pct 至 54%，经调整净利润同比增长 13.3% 至 47.0 亿（符合一致预期 46.7 亿元）。分业务看，直播、广告、电商收入同比分别变动-2.0%、13.3%、14.1%。展望 25 年，我们预期总收入同比上涨 11%，广告收入和 GMV 增速同比分别放缓至 14% 和 13%，经调整净利润达 206 亿。我们预期快手收入增长节奏或将于 2Q 起逐步加速。中长期而言，我们认为可灵有望拓展 AI 商业化边界，例如 2C 订阅、2B 电商广告、API 调用均会带来增量收入（25 年预期收入 4-8 亿）。维持“买入”评级。</p> <p>展望 25-26 年，我们调整快手收入预测+0.6%、-0.1%，25 年上调主系可灵 AI 贡献广告收入。我们更改经调整净利润预测-12.6%、11.0% 至 206/250 亿，主系可灵 AI 加大投入。我们引入 27 年新值，对应收入与经调整净利润分别为 1638 亿与 283 亿元。我们根据 SOTP 估值，给予目标价 68.73 港元（前值 60.45 港元）。本次估值上调，主系广告业务估值上调（AI 提效），可灵 AI 商业化收入潜力开始释放。维持“买入”评级。</p> <p>风险提示：广告加载率、直播电商和货架电商增速不及预期。</p> <p>报告发布日期：2025 年 03 月 26 日</p> <p>点击下载全文：快手-W(1024 HK,买入): 4Q 收入符合预期，可灵商业化加速</p>
哔哩哔哩 (BILI US)	<p>公司公布 24Q4 业绩：收入 77.34 亿元，同比+21.8%，超 VA 的 76.37 亿元的预测；调整后归母净利润 4.53 亿元，同比转正，略超 VA 4.10 亿元的预测。收入和利润整体好于市场预期，反映三国递延收入逐步确认和平台流量商业化能力的提升。我们预计 25 年公司扭亏为盈，利润加速兑现，维持“买入”。</p> <p>综合广告和游戏业务的结构性增长，我们预计公司 25-27 年收入为 306.3/334.5/359.1 亿元（原预测 25/26 307.2/334.7 亿元），调整后归母净利润为 16.49/21.15/26.5 亿元（原预测 25/26 9.38/13.09 亿元，上调 75.80%/61.57%）。公司业务较为多元，采用 25E 分部 PS 估值法，合计目标市值为 1038.9 亿人民币，对应目标价为 34.4 美元（汇率为 7.17，原目标价为 27.3 美元），对应 25/26E PS 3.4x/3.1x，维持“买入”评级。</p> <p>风险提示：用户增长、游戏表现不及预期，业务相关政策收紧风险。</p> <p>报告发布日期：2025 年 02 月 21 日</p> <p>点击下载全文：哔哩哔哩(BILI US,买入): 游戏表现稳健，25 年盈利弹性可期</p>

股票名称	最新观点
有道 (DAO US)	<p>有道公布 24Q4 业绩：收入 13.4 亿元，yoy-9.5%；Non-GAAP 归母净利润 8866 万元，yoy+32.5%；24 全年实现 Non-GAAP 归母净利润 1.05 亿元（大幅超 VA 一致预期的 6408 万元），同比转正，去年同期为亏损 4.75 亿。25 年开年以来，在 Deepseek 催化下公司加速 AI 大模型迭代，成为国内首批接入 DeepSeek-R1 并率先在软硬件上快速商业化落地的教育公司。我们持续看好有道作为国内 AI+教育领军企业的技术能力和产品能力，自研子曰大模型的丰富垂直数据储备结合的通用推理能力有望进一步打开商业化空间，维持“买入”。</p> <p>考虑到学习服务结构性调整以及整体盈利后所得税率或上升，我们预计 25-27E 公司整体 Non-GAAP 归母净利润为 1.73/2.43/3.43 亿元（前值 2.45/3.53/-亿元），基于 SOTP 估值法，给予 K12 业务/智能硬件业务 24.63/39.27x 25E PE（与可比公司 Wind 一致预测均值一致），考虑在线营销业务增速较快，给予 15x 25E PE（可比均值为 11.82x），对应目标市值约 14.39 亿美金，目标价为 12.04 美元（前值 8.96 美元，基于可比公司估值变化上调）。维持“买入”。</p> <p>风险提示：获客成本显著提升；数字内容用户接受度不高；智能硬件市场渗透率不及预期；宏观需求压制广告业务增长；相关部门对学费定价的监管。</p> <p>报告发布日期：2025 年 02 月 21 日</p> <p>点击下载全文：有道(DAO US,买入)：“AI 原生”战略驱动，盈利拐点已现</p>
好未来 (TAL US)	<p>公司公布 3QFY25（下文简称 3Q）业绩：收入 6.06 亿美元，同比+62.4%，大幅超过彭博 44.2%的一致预期；Non-GAAP 经营利润-192 万美元，显著好于彭博-1319 万美元的一致预期。我们认为本季度的强劲表现主要得益于线下教育需求持续旺盛以及 AI 学习机销售超预期。公司产能稳健扩张，各条业务线产品持续丰富，有望为长期增长提供有力支撑，维持“买入”评级。</p> <p>考虑到 3Q 业绩超预期，我们上调 FY25/26/27 Non-GAAP 归母净利润预测至 2.02/2.55/3.36 亿美元（较前次预测上调 44.29%/22.01%/3.70%），SOTP 目标价为 15.08 美元（前值：13.83 美元），维持“买入”评级。</p> <p>风险提示：对素养教育和高中学科培训的监管力度加强，新业务拓展不成功，竞争加剧，优质师资和核心管理人才流失。</p> <p>报告发布日期：2025 年 01 月 24 日</p> <p>点击下载全文：好未来(TAL US,买入)：产品周期上行，利润弹性可期</p>
BOSS 直聘 (BZ US)	<p>公司发布 24Q4 财报：24Q4 实现收入 18.24 亿元，略超华泰预测 18.08 亿元，yoy+15.4%。24Q4 经调整净利润 7.22 亿元，符合华泰预测 7.24 亿元，yoy+14.9%。随着宏观政策持续推进，春节以来招聘需求呈现修复趋势，若需求弹性释放，公司基于产品竞争力有望显著受益，维持“买入”评级。旺季需求展现良好弹性，Q1 指引好于预期。考虑到全年需求有待观察，我们仅小幅上修 FY25-FY27 收入 81.5/93.4/106.2 亿元（上修 2.2%/1.6%/-），调整后归母净利润为 32.2/34.7/37.8 亿元（上修 0%/1.2%/-），基于 DCF 估值法，维持参数不变，永续增长假设 3%，WACC 为 8.95%，对应目标价为 22.71 美元（前值 19.23 美元），对应调整后 25E PE 22x，维持“买入”。</p> <p>风险提示：企业需求持续承压；行业竞争加剧；蓝领业务拓展不及预期。</p> <p>报告发布日期：2025 年 03 月 12 日</p> <p>点击下载全文：BOSS 直聘(BZ US,买入)：边际趋势良好，期待春招弹性</p>

资料来源：Bloomberg，华泰研究预测

免责声明

分析师声明

本人，夏路路、苏燕妮、丁骄琬、侯杰、詹博、郑裕佳，兹证明本报告所表达的观点准确地反映了分析师对标的证券或发行人的个人意见；彼以往、现在或未来并无就其研究报告所提供的具体建议或所表达的意见直接或间接收取任何报酬。

一般声明及披露

本报告由华泰证券股份有限公司（已具备中国证监会批准的证券投资咨询业务资格，以下简称“本公司”）制作。本报告所载资料是仅供接收人的严格保密资料。本报告仅供本公司及其客户和其关联机构使用。本公司不因接收人收到本报告而视其为客户。

本报告基于本公司认为可靠的、已公开的信息编制，但本公司及其关联机构（以下统称为“华泰”）对该等信息的准确性及完整性不作任何保证。

本报告所载的意见、评估及预测仅反映报告发布当日的观点和判断。在不同时期，华泰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。同时，本报告所指的证券或投资标的的价格、价值及投资收入可能会波动。以往表现并不能指引未来，未来回报并不能得到保证，并存在损失本金的可能。华泰不保证本报告所含信息保持在最新状态。华泰对本报告所含信息可在不发出通知的情形下做出修改，投资者应当自行关注相应的更新或修改。

本公司不是 FINRA 的注册会员，其研究分析师亦没有注册为 FINRA 的研究分析师/不具有 FINRA 分析师的注册资格。

华泰力求报告内容客观、公正，但本报告所载的观点、结论和建议仅供参考，不构成购买或出售所述证券的要约或招揽。该等观点、建议并未考虑到个别投资者的具体投资目的、财务状况以及特定需求，在任何时候均不构成对客户私人投资建议。投资者应当充分考虑自身特定状况，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。对依据或者使用本报告所造成的一切后果，华泰及作者均不承担任何法律责任。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。

除非另行说明，本报告中所引用的关于业绩的数据代表过往表现，过往的业绩表现不应作为日后回报的预示。华泰不承诺也不保证任何预示的回报会得以实现，分析中所做的预测可能是基于相应的假设，任何假设的变化可能会显著影响所预测的回报。

华泰及作者在自身所知情的范围内，与本报告所指的证券或投资标的不存在法律禁止的利害关系。在法律许可的情况下，华泰可能会持有报告中提到的公司所发行的证券头寸并进行交易，为该公司提供投资银行、财务顾问或者金融产品等相关服务或向该公司招揽业务。

华泰的销售人员、交易人员或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。华泰没有将此意见及建议向报告所有接收者进行更新的义务。华泰的资产管理部门、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。投资者应当考虑到华泰及/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信赖依据。有关该方面的具体披露请参照本报告尾部。

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、发布的机构或人员，也并非意图发送、发布给因可得到、使用本报告的行为而使华泰违反或受制于当地法律或监管规则的机构或人员。

本报告版权仅为本公司所有。未经本公司书面许可，任何机构或个人不得以翻版、复制、发表、引用或再次分发他人（无论整份或部分）等任何形式侵犯本公司版权。如征得本公司同意进行引用、刊发的，需在允许的范围内使用，并需在使用前获取独立的法律意见，以确定该引用、刊发符合当地适用法规的要求，同时注明出处为“华泰证券研究所”，且不得对本报告进行任何有悖原意的引用、删节和修改。本公司保留追究相关责任的权利。所有本报告中使用的商标、服务标记及标记均为本公司的商标、服务标记及标记。

中国香港

本报告由华泰证券股份有限公司制作，在香港由华泰金融控股（香港）有限公司向符合《证券及期货条例》及其附属法律规定的机构投资者和专业投资者的客户进行分发。华泰金融控股（香港）有限公司受香港证券及期货事务监察委员会监管，是华泰国际金融控股有限公司的全资子公司，后者为华泰证券股份有限公司的全资子公司。在香港获得本报告的人员若有任何有关本报告的问题，请与华泰金融控股（香港）有限公司联系。

香港-重要监管披露

- 华泰金融控股（香港）有限公司的雇员或其关联人士没有担任本报告中提及的公司或发行人的高级人员。
- 阿里巴巴-W (9988 HK)、美图公司 (1357 HK)、快手-W (1024 HK)、腾讯控股 (700 HK)：华泰金融控股（香港）有限公司、其子公司和/或其关联公司在本报告发布日担任标的公司证券做市商或者证券流动性提供者。
- 有关重要的披露信息，请参华泰金融控股（香港）有限公司的网页 https://www.htsc.com.hk/stock_disclosure 其他信息请参见下方“美国-重要监管披露”。

美国

在美国本报告由华泰证券（美国）有限公司向符合美国监管规定的机构投资者进行发表与分发。华泰证券（美国）有限公司是美国注册经纪商和美国金融业监管局（FINRA）的注册会员。对于其在美国分发的研究报告，华泰证券（美国）有限公司根据《1934年证券交易法》（修订版）第15a-6条规定以及美国证券交易委员会人员解释，对本研究报告内容负责。华泰证券（美国）有限公司联营公司的分析师不具有美国金融监管（FINRA）分析师的注册资格，可能不属于华泰证券（美国）有限公司的关联人员，因此可能不受FINRA关于分析师与标的公司沟通、公开露面和所持交易证券的限制。华泰证券（美国）有限公司是华泰国际金融控股有限公司的全资子公司，后者为华泰证券股份有限公司的全资子公司。任何直接从华泰证券（美国）有限公司收到此报告并希望就本报告所述任何证券进行交易的人士，应通过华泰证券（美国）有限公司进行交易。

美国-重要监管披露

- 分析师夏路路、苏燕妮、丁骄璇、侯杰、詹博、郑裕佳本人及相关人士并不担任本报告所提及的标的证券或发行人的高级人员、董事或顾问。分析师及相关人士与本报告所提及的标的证券或发行人并无任何相关财务利益。本披露中所提及的“相关人士”包括FINRA定义下分析师的家庭成员。分析师根据华泰证券的整体收入和盈利能力获得薪酬，包括源自公司投资银行业务的收入。
- 阿里巴巴-W (9988 HK)、美图公司 (1357 HK)、快手-W (1024 HK)、腾讯控股 (700 HK)：华泰证券股份有限公司、其子公司和/或其联营公司在本报告发布日担任标的公司证券做市商或者证券流动性提供者。
- 华泰证券股份有限公司、其子公司和/或其联营公司，及/或不时会以自身或代理形式向客户出售及购买华泰证券研究所覆盖公司的证券/衍生工具，包括股票及债券（包括衍生品）华泰证券研究所覆盖公司的证券/衍生工具，包括股票及债券（包括衍生品）。
- 华泰证券股份有限公司、其子公司和/或其联营公司，及/或其高级管理层、董事和雇员可能会持有本报告中所提到的任何证券（或任何相关投资）头寸，并可能不时进行增持或减持该证券（或投资）。因此，投资者应该意识到可能存在利益冲突。

新加坡

华泰证券（新加坡）有限公司持有新加坡金融管理局颁发的资本市场服务许可证，可从事资本市场产品交易，包括证券、集体投资计划中的单位、交易所交易的衍生品合约和场外衍生品合约，并且是《财务顾问法》规定的豁免财务顾问，就投资产品向他人提供建议，包括发布或公布研究分析或研究报告。华泰证券（新加坡）有限公司可能会根据《财务顾问条例》第32C条的规定分发其在华泰内的外国附属公司各自制作的信息/研究。本报告仅供认可投资者、专家投资者或机构投资者使用，华泰证券（新加坡）有限公司不对本报告内容承担法律责任。如果您是非预期接收者，请您立即通知并直接将本报告返回给华泰证券（新加坡）有限公司。本报告的新加坡接收者应联系您的华泰证券（新加坡）有限公司关系经理或客户主管，了解来自或与所述分发的信息相关的事宜。

评级说明

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（含此期间的股息回报）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，台湾市场基准为台湾加权指数，日本市场基准为日经225指数，新加坡市场基准为海峡时报指数，韩国市场基准为韩国有价证券指数，英国市场基准为富时100指数），具体如下：

行业评级

- 增持：**预计行业股票指数超越基准
- 中性：**预计行业股票指数基本与基准持平
- 减持：**预计行业股票指数明显弱于基准

公司评级

- 买入：**预计股价超越基准15%以上
- 增持：**预计股价超越基准5%~15%
- 持有：**预计股价相对基准波动在-15%~5%之间
- 卖出：**预计股价弱于基准15%以上
- 暂停评级：**已暂停评级、目标价及预测，以遵守适用法规及/或公司政策
- 无评级：**股票不在常规研究覆盖范围内。投资者不应期待华泰提供该等证券及/或公司相关的持续或补充信息

**法律实体披露**

中国: 华泰证券股份有限公司具有中国证监会核准的“证券投资咨询”业务资格, 经营许可证编号为: 91320000704041011J

香港: 华泰金融控股(香港)有限公司具有香港证监会核准的“就证券提供意见”业务资格, 经营许可证编号为: AOK809

美国: 华泰证券(美国)有限公司为美国金融业监管局(FINRA)成员, 具有在美国开展经纪交易商业业务的资格, 经营业务许可编号为: CRD#:298809/SEC#:8-70231

新加坡: 华泰证券(新加坡)有限公司具有新加坡金融管理局颁发的资本市场服务许可证, 并且是豁免财务顾问。公司注册号: 202233398E

华泰证券股份有限公司**南京**

南京市建邺区江东中路228号华泰证券广场1号楼/邮政编码: 210019

电话: 86 25 83389999/传真: 86 25 83387521

电子邮件: ht-rd@htsc.com

深圳

深圳市福田区益田路5999号基金大厦10楼/邮政编码: 518017

电话: 86 755 82493932/传真: 86 755 82492062

电子邮件: ht-rd@htsc.com

北京

北京市西城区太平桥大街丰盛胡同28号太平洋保险大厦A座18层/

邮政编码: 100032

电话: 86 10 63211166/传真: 86 10 63211275

电子邮件: ht-rd@htsc.com

上海

上海市浦东新区东方路18号保利广场E栋23楼/邮政编码: 200120

电话: 86 21 28972098/传真: 86 21 28972068

电子邮件: ht-rd@htsc.com

华泰金融控股(香港)有限公司

香港中环皇后大道中99号中环中心53楼

电话: +852-3658-6000/传真: +852-2567-6123

电子邮件: research@htsc.com

<http://www.htsc.com.hk>

华泰证券(美国)有限公司

美国纽约公园大道280号21楼东(纽约10017)

电话: +212-763-8160/传真: +917-725-9702

电子邮件: Huatai@htsc-us.com

<http://www.htsc-us.com>

华泰证券(新加坡)有限公司

滨海湾金融中心1号大厦, #08-02, 新加坡 018981

电话: +65 68603600

传真: +65 65091183

©版权所有2025年华泰证券股份有限公司