

AI应用专题： 各大厂新模型持续迭代，重视AI应用板块投资机会

行业研究 · 行业专题
投资评级：优于大市（维持评级）

证券分析师：熊莉

021-61761067

xiongli1@guosen.com.cn

S0980519030002

证券分析师：云梦泽

021-60933155

yunmengze@guosen.com.cn

S0980525110001

➤**国外大厂AI应用落地，垂直场景深耕、技术革新：**医疗领域方面，OpenAI推出 ChatGPT Health，Anthropic 发布 Claude for Healthcare，均聚焦合规性与专业服务，覆盖健康管理、临床服务等核心场景；电商与消费电子方面，亚马逊推出AI退货看板优化跨境电商运营，谷歌模型嵌入智能电视、家电等多终端场景化服务；技术底座方面，英伟达发布全新Vera Rubin AI平台及DLSS 4.5，提升AI推理性能与终端体验。

➤**国内各大厂聚焦AI应用化场景融合：**阿里“蚂蚁阿福”升级健康服务、千问 APP 公测对标 ChatGPT；字节火山引擎成春晚独家AI云伙伴，豆包大模型赋能全链路；Deepseek 即将发布V4旗舰模型，革新架构与稀疏性技术；腾讯推出小程序成长计划，提供算力、资源等全周期支持；智谱、Minimax 相继在港上市，成为AI行业资本化标杆，上市后股价截至1月13日分别上涨 56.45%、212.12%。

➤**风险提示：**技术研发不及预期、市场需求不及预期等

[01] 国外大厂AI应用落地

[02] 国内大厂AI应用落地

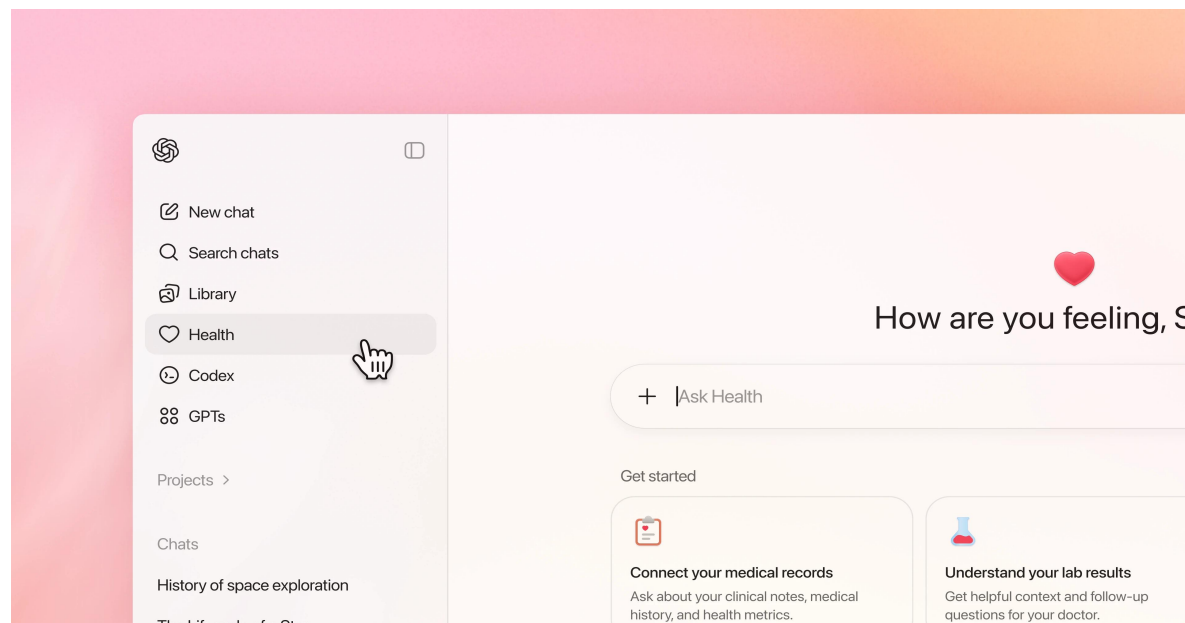
[03] 产业链梳理

[04] 风险提示

Open AI发布 ChatGPT Health 健康应用

- ◆ **ChatGPT Health正式落地医疗场景**，成为OpenAI首个垂直行业旗舰产品。据 OpenAI 透露，平台每周收到的健康与养生类咨询量已超 2.3 亿人次。其产品核心如下：
 - ◆ **数据交互能力**：支持用户直接连接苹果健康等主流健康类应用，实现健康数据一键导出与整合分析；
 - ◆ **合规性保障**：与实时健康数据连接平台b.well 合作，遵循行业最高标准的数据安全与隐私规范，且医疗记录仅向年满 18 岁的用户提供。同时，可以连接 Apple 健康信息以及其他健康应用，健康信息与记忆不会回流至主要聊天中，确保用户健康数据传输与存储的安全性。
 - ◆ **专业评估体系**：联合来自 60 个国家/地区、数十个专科的**260多名医生共同开发HealthBench**评估框架，在 30 个重点领域对模型输出提供了超过 60 万次反馈。**保障医疗建议的专业性与可靠性。**

图：ChatGPT Health

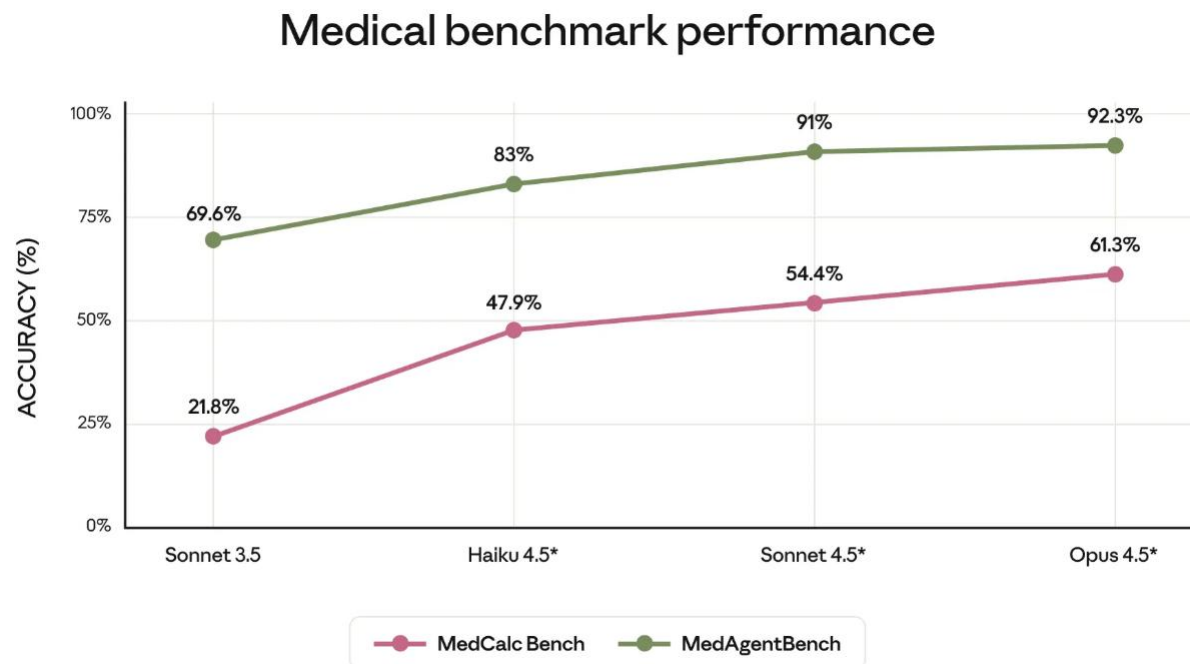


资料来源：Open AI 官网，国信证券研究所整理

Anthropic: Claude for Healthcare

- ◆ 基于去年10月发布的Claude for Life Sciences，Anthropic 于 2026年1月12日发布 Claude for Healthcare，一项符合美国《健康保险流通与责任法案》要求的医疗服务，面向医院、医疗机构和个人用户开放。
 - ◆ 全链路医疗 AI 支撑：覆盖临床服务、医保结算、科研协作、个人健康管理四大核心场景。
 - ◆ 建立多个连接器：
 - ◆ 链接行业标准系统与数据库：医疗保险与医疗补助服务中心（CMS）覆盖数据库、《国际疾病分类》第十版（ICD-10）、国家提供者标识符登记。
 - ◆ 代理功能：FHIR（医疗系统之间数据交换的现代标准）、一项示例事先授权审查技能。
 - ◆ 安全合规优先：基于 Anthropic Constitutional AI 框架，保障 AI 输出的无害性、诚实性，严格遵循医疗数据安全、国家提供者标识符登记标准。

图：迭代模型的医学基准表现



资料来源：Anthropic 官网，国信证券研究所整理

图：Claude for Healthcare

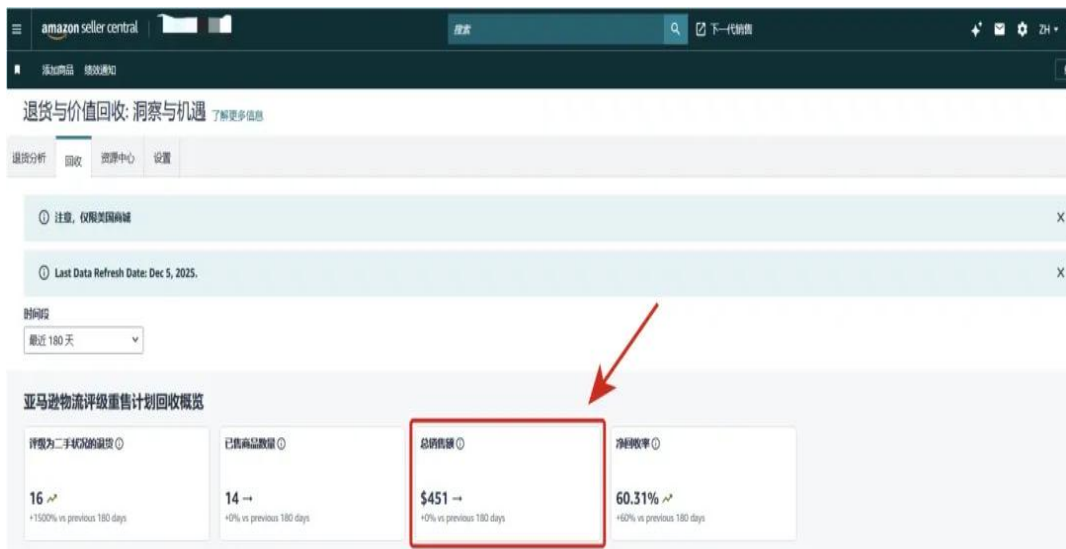


资料来源：Anthropic 官网，国信证券研究所整理

亚马逊：推出AI退货看板

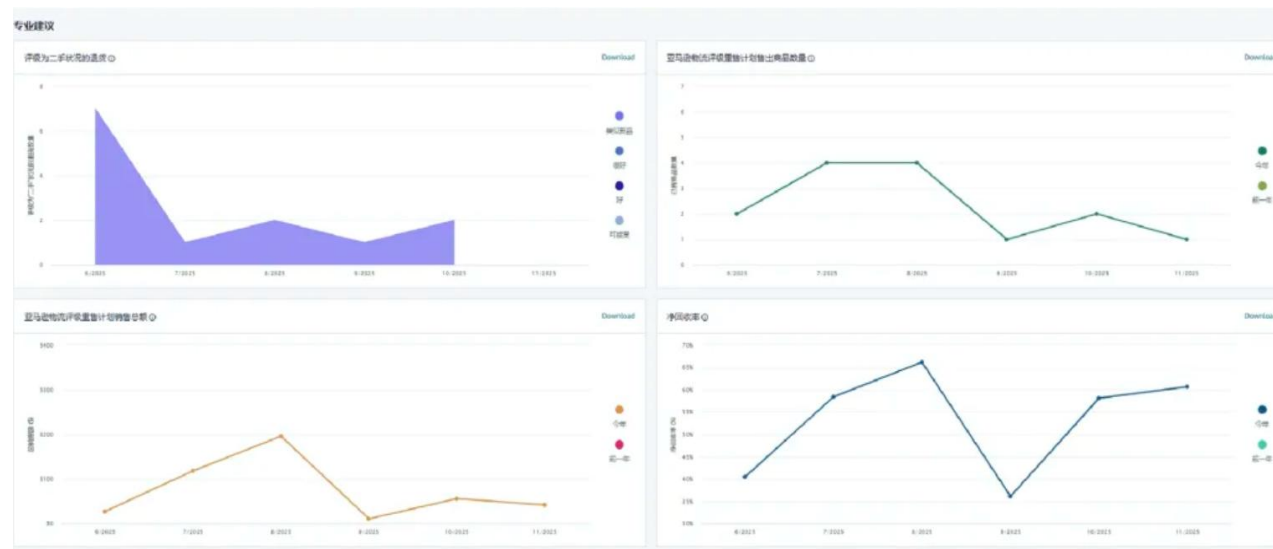
- ◆ 于2026年1月7日推出针对跨境电商卖家退货管理痛点的全链路智能解决方案，通过 AI 整合数据、定位问题、提供方案，打通“退货退款 - 商品评级 - 库存二次利用”全流程，帮助卖家从“被动处理退货”转向“主动优化运营”。
- ◆ 退货分析页面：核心功能页，全局数据可视化，可按照商品风险等级标记高退货率交易，AI分析归因并同步生成解决方案。
- ◆ 回收页面：实时监控退货商品的处置情况，展示“已售二手商品数量”“回收总销售额”“净回收率”等指标，提供180天回收数据趋势图。
- ◆ 资源中心页面：聚合所有退货管理所需工具，提供可下载报告入口。
- ◆ 设置页面：一站式提供全流程设置入口，与风险规避工具链接。

图：亚马逊AI回收看板



资料来源：亚马逊官网，国信证券研究所整理

图：亚马逊AI看板数据可视化结果与归因分析示例



资料来源：亚马逊官网，国信证券研究所整理

◆ 最新发布：

- ◆ **Gemini 3 Pro**：于2025年11月18日发布，谷歌当前最先进原生多模态推理模型，支持文本、图像、音频、视频、代码等多模态输入，最高1M tokens上下文+64K tokens输出，采用稀疏MoE+Transformer架构，可实现整库级代码理解、超长文档分析，在多模态理解、复杂推理、Agent能力上显著提升。
- ◆ **Gemini 3 Flash**：于2025年12月18日发布，轻量化高性能模型，运行速度达Gemini 2.5 Pro的3倍，成本仅为Gemini 3 Pro的1/4，保持接近Pro版的推理能力，成为谷歌搜索AI模式预设系统，定价为每百万输入令牌0.5美元、输出令牌3美元。

表：Gemini场景化应用落地

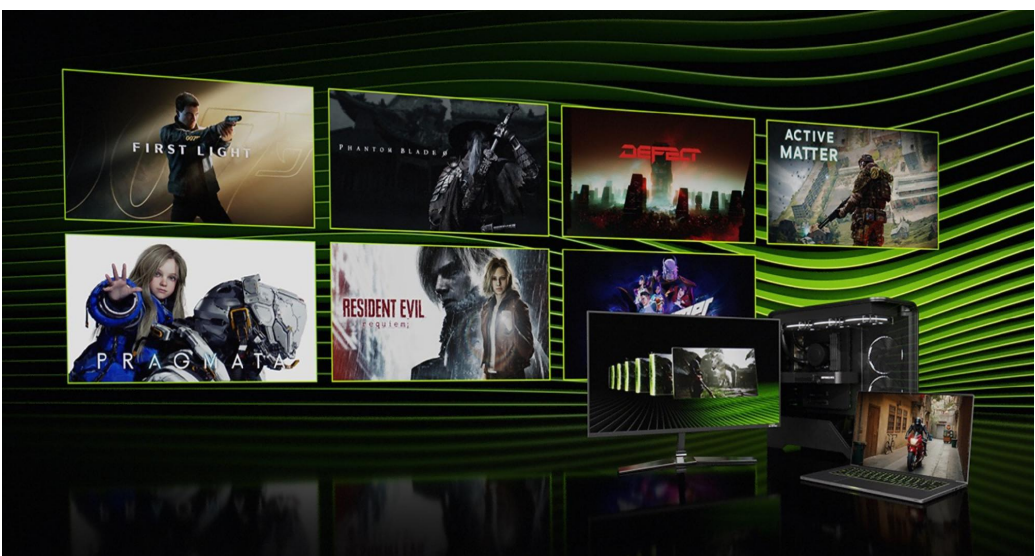
应用场景	终端	AI核心能力	典型功能示例
智能电视	Google TV 与合作品牌电视（如 TCL）	自然语言交互、视觉理解、多模态反馈	通过语音调整电视设置，动态提供信息与知识解释，访问相册、内容推荐、生成图片/视频等互动体验
智能家电/厨房电器	Bespoke AI 冰箱、AI 葡萄酒柜、其他厨房电器	视觉识别、物体感知与分类	自动识别冰箱内食材种类并管理清单；识别酒瓶标签并更新库存信息
移动消费电子	手机、平板等设备	语音助手、系统级智能服务	扩大 AI 功能在移动端部署规模

资料来源：CES大会，国信证券研究所整理

英伟达：推出全新Rubin平台及更新DLSS 4.5

- ◆1月5日推出全新 AI 平台“Vera Rubin”，首个采用极致协同设计、集成六款芯片的 AI 平台，现已全面量产。由Vera CPU、Rubin GPU等六款芯片组成。涵盖Rubin GPU、Vera CPU、NVLink 6等组件。该平台预期将在 2026 下半年进入市场，并提供更高推理性能、更低运行成本。
- ◆1月6日推出“DLSS 4.5”，引入第二代超分辨率和“6x 多帧生成”能力，提升画质与高帧率表现。支持超过 250 款游戏/应用，并计划在新系列显卡上春季发布完整功能。

图：DLSS 4.5



资料来源：英伟达官网，国信证券经济研究所整理

表：NVIDIA Rubin 平台技术突破

技术突破	核心内容	关键能力提升
第六代 NVIDIA NVLink 互连技术	GPU–GPU 高速互连，每 GPU 带宽达 3.6 TB/s，NVL72 机架总带宽 260 TB/s	支撑超大规模 MoE 模型的高速通信，显著缩短训练时间与推理延迟
第三代 Transformer 引擎	支持硬件级自适应压缩与 NVFP4 精度计算	在保持模型效果的同时大幅降低推理计算量与 token 成本
NVIDIA Vera CPU 架构	面向智能体推理设计的高能效 CPU，支持 NVLink-C2C	提升 CPU–GPU 协同效率，支撑大规模 AI 工厂级部署
第三代 NVIDIA 机密计算	在 CPU、GPU、NVLink 域间提供端到端数据保护	保障超大规模专有模型与训练/推理数据安全
第二代 RAS 引擎	覆盖 CPU、GPU 与互连的可靠性与可维护性设计	实时健康监测、容错与主动维护，显著提升系统稳定性与可用率

资料来源：英伟达官网，国信证券经济研究所整理

[01] 国外大厂AI应用落地

[02] 国内大厂AI应用落地

[03] 产业链梳理

[04] 风险提示

阿里巴巴：AQ健康品牌升级为蚂蚁阿福

◆2025年12月15日更名阿福，代言人何炅。月活用户据悉超1500万，以国内第一大健康APP为目标。

◆核心功能升级：

- ◆健康陪伴：支持接入苹果、华为、荣耀等9大品牌设备、提供饮食计划等指导
- ◆健康问答：支持对话、拍照，模拟医生主动追问
- ◆健康服务：提供线上问诊、吸纳下就医云陪诊、预约挂号等；支持医保实时结算、商保码自动赔付等。

◆权威合作：已和8省市卫健委、18省市医保局“依保儿”合作，推出 200 家医院就医智能体，并与16家机构 AI 联合研究。

图：蚂蚁阿福



资料来源：蚂蚁阿福官网，国信证券研究所整理

阿里巴巴：千问APP公测上线

- ◆ 于2025年11月17日推出千问 App，基于2025年4月阿里通义千问团队推出的Qwen3 系列开源模型，包含 2 个 MoE（混合专家）模型、6 个 Dense 模型，参数量覆盖 0.6B-235B，适配手机端至云端全场景部署打造,全面对标 ChatGPT。
 - ◆ 高效架构：MoE 架构（如 Qwen3-235B-A22B）总参数量 2350 亿，仅激活 220 亿参数，显存占用为同类模型 1/3，4 张 H20 显卡即可部署满血版本，大幅降低成本。
 - ◆ 性能突破：依托 36T tokens 预训练数据（覆盖 119 种语言），数学推理（AIME25 竞赛正确率近 100%）、编程（LiveCodeBench 全球前三）、256K 长文本理解能力比肩 GPT-4o、Gemini 2.5-Flash。
 - ◆ 开源生态：通过 Amazon Bedrock、FireworksAI 等平台全球部署，吸引开发者共建生态，反哺 App 快速迭代。

图：千问



资料来源：千问官网，国信证券研究所整理

表：千问助手功能一览

功能	能力描述
对话问答	专业领域问答能力突出，主动推荐图表 / 视频辅助理解；支持中文语境深度理解
全能相机	万物精准识别、以图搜图、超强视觉推理
办公效率工具	AI PPT、AI 文库、万字会议纪要上传后自动提炼要点
场景化服务	接入高德地图后，可生成含推荐餐厅、酒店、地理位置及出行方案的可视化决策卡片；未来将接入淘宝、饿了么等生态
多模态创作	10 秒内生成视频，支持 K 线图分析、“生熟食砧板分用”等专业知识解读

资料来源：IT之家，国信证券研究所整理

字节：官宣成为央视春晚独家AI伙伴

- ◆ 2026年春晚的独家AI云合作伙伴已经官宣为火山引擎，豆包大模型将深度协同，上线多种AI互动玩法。这是字节第三次加盟春晚，首次以AI助手类产品亮相。
- ◆ 验证火山引擎在超高并发、极致弹性、高稳定性场景下的AI云能力。
- ◆ 借助春晚IP提升ToB品牌力，结合豆包形成“技术+场景”的AI生态闭环。
- ◆ 覆盖节目制作、线上互动、全球直播保障，赋能春晚全链路。

图：春晚独家AI云合作伙伴发布



资料来源：澎湃新闻，国信证券经济研究所整理

表：字节春晚深度参与场景

参与场景	技术支持主体	具体内容
节目制作	火山引擎（AI 云）	AI 赋能节目创意策划，提供多模态内容生成建议； AI 特效制作，提升舞台视觉效果
线上互动	火山引擎（AI 云底座）+ 豆包（AI 助手）	“AI 同桌看春晚”实时互动等玩法； 全球用户互动数据实时分析与反馈
视频直播	火山引擎（云计算）	全球直播流分发与加速； 直播画面 AI 实时优化（降噪、增强、超分）
技术底座	火山引擎（AI 云基础设施）	多模态大模型推理服务； 数据存储与计算资源弹性调度
用户触达	豆包（AI 助手）	春晚专属 AI 互动入口； 多语言春晚内容实时翻译

资料来源：万德，国信证券研究所整理

Deepseek：预计2月中旬发布V4旗舰模型

- ◆ 2025年12月31日发布《mHC: 流形约束超连接》，提出全新训练架构，解决大模型扩容时的“信号爆炸”与“训练崩溃”难题；1月4日更新86页R1论文，详尽披露完整训练管线；2026年1月13日发布《Conditional Memory via Scalable Lookup: A New Axis of Sparsity for Large Language Models》，提出Engram条件记忆模块，构建“查—算分离”新机制，开启大模型稀疏性新维度。
- ◆ 预计2月中旬发布V4旗舰模型：
 - ◆ 架构级革命：独创mHC（流形约束超连接）技术，成功解决传统Transformer架构参数扩容时的“信号爆炸”与“模型坍塌”难题
 - ◆ 超长上下文实用化落地：基于NSA（原生吸出注意力）机制优化，可连贯处理百万级Token文本，解码速度较前代提升11.6倍，能一次性完成中型项目代码库的架构分析、跨文件漏洞检测与重构建议。
 - ◆ 编程能力全面反超国际顶尖模型：在Design2Code基准测试中以92.0分刷新纪录，实现从函数级补全到项目级架构设计的全栈服务。
- ◆ V4的发布或能助力DeepSeek重振市占率，应对AI行业向头部收敛的竞争格局，有望推动从国产算力芯片到AI应用融合的整条产业链发展。

腾讯：推出“AI应用及线上工具小程序成长计划”

- ◆ 2026年1月5日推出AI应用及线上工具小程序成长计划，提供云开发资源、AI算力、数据分析、广告变现及流量激励等全方位支持，覆盖2026年全年，全周期陪伴开发者成长。
- ◆ 面向所有线上类小程序开发者，涵盖个人及企业开发者，适用场景包括线上工具、AI原生应用、娱乐类小程序等。申请者可登录电脑端微信小程序后台，进入行业能力栏目，选择该计划提交申请即可参与。
- ◆ 该计划通过全链路资源扶持，降低AI应用开发门槛，激励开发者将创意转化为可落地、可破圈、可盈利的小程序产品，推动微信小程序生态从工具平台向AI服务市场升级，巩固腾讯在AI应用生态的竞争优势。

图：微信AI应用及线上工具小程序成长计划



资料来源：央广网，国信证券经济研究所整理

表：计划核心扶持内容

扶持维度	具体内容
免费云开发资源	新开发者可免费创建6个月个人版云开发环境，含数据库、云函数等完备后端能力，支持对接AI大模型及Agent； 已有云开发环境的开发者可领取大额抵扣券
免费AI算力支持	提供1亿Token的腾讯混元2.0文生文模型额度； 提供1万张混元文生图模型额度； 后续混元最新模型将持续上架
运营增长支持	免费开放We分析专业版一年，提供深度数据洞察
商业化变现支持	开放安卓、iOS、鸿蒙全终端虚拟支付与会员订阅能力； 系统自动推荐广告位、支持实时预览上线

资料来源：腾讯云官网，国信证券研究所整理

- ◆ 智谱在 2026 年 1 月 8 日正式在港交所主板挂牌上市（02513.HK），是中国首家以大语言模型为主营业务在港公开上市的公司之一。上市前完成8轮市场化融资，累计募资约83.63亿元，2025年5月B6轮融资后投后估值达243.77亿元，较首轮增长数十倍。
- ◆ **MiniMax** 于 2026 年 1 月 9 日在香港上市（0100.HK），距公司成立仅约 4 年，成为目前“全球 AI 最快上市案例”。股东阵容豪华，含阿里巴巴、腾讯、小红书等互联网平台，高瓴资本、IDG、红杉中国等顶级机构，以及阿布扎比投资局、新加坡GIC、挪威央行等国际资本。

表：智谱与minimax上市行情

公司	当前市值	上市时间	截至今天收盘涨幅
智谱	800.34亿	2026-1-8	56.45%
MiniMax	1230.84亿	2026-1-9	212.12%

资料来源：Wind，国信证券研究所整理

[01] 国外大厂AI应用落地

[02] 国内大厂AI应用落地

[03] 产业链梳理

[04] 风险提示

应用方向	标的
多模态	昆仑万维、万兴科技、合合信息、 星环科技、三六零、当虹科技、虹软科技
医疗	卫宁健康、创业慧康、美年健康、嘉和美康、 讯飞医疗科技、迪安诊断、晶泰控股、中科信息、艾可蓝
漫剧	中文在线、华策影视、欢瑞世纪、掌阅科技
游戏	盛天网络、三七互娱、巨人网络、掌趣科技
GEO	蓝色光标、易点天下、汇量科技、视觉中国、人民网、值得买、 引力传媒、天娱数科、浙文互联、因赛集团、恒为科技、迈富时、天龙集团
电 商	光 云 科 技
教 育	科大讯飞、豆神教育、佳发教育、竞业达
金融	恒生电子、同花顺、金证股份、中金金财、信雅达、 南天信息、高伟达、京北方、天阳科技、宇信科技
办公	金山办公、用友网络、泛微网络、汉得信息、金蝶国际、百融云、 第四范式、深信服、卓易信息、税友股份、捷顺科技、金桥信息、彩讯股份
工业	宝信软件、中控技术、能科科技、东土科技、鼎捷数智
军工	航天宏图、拓尔思、中科星图
农业	托普云农

[01] 国外大厂AI应用落地

[02] 国内大厂AI应用落地

[03] 产业链梳理

[04] 风险提示

- 技术研发不及预期
- 市场需求不及预期
- 个股梳理仅基于产业链结构不涉及主观投资建议

国信证券投资评级			
投资评级标准	类别	级别	说明
报告中投资建议所涉及的评级（如有）分为股票评级和行业评级（另有说明的除外）。评级标准为报告发布日后6到12个月内的相对市场表现，也即报告发布日后的6到12个月内公司股价（或行业指数）相对同期相关证券市场代表性指数的涨跌幅作为基准。A股市场以沪深300指数（000300.SH）作为基准；新三板市场以三板成指（899001.CSI）为基准；香港市场以恒生指数（HSI.HI）作为基准；美国市场以标普500指数（SPX.GI）或纳斯达克指数（IXIC.GI）为基准。	股票投资评级	优于大市	股价表现优于市场代表性指数10%以上
		中性	股价表现介于市场代表性指数±10%之间
		弱于大市	股价表现弱于市场代表性指数10%以上
		无评级	股价与市场代表性指数相比无明确观点
	行业投资评级	优于大市	行业指数表现优于市场代表性指数10%以上
		中性	行业指数表现介于市场代表性指数±10%之间
		弱于大市	行业指数表现弱于市场代表性指数10%以上

分析师承诺

作者保证报告所采用的数据均来自合规渠道；分析逻辑基于作者的职业理解，通过合理判断并得出结论，力求独立、客观、公正，结论不受任何第三方的授意或影响；作者在过去、现在或未来未就其研究报告所提供的具体建议或所表述的意见直接或间接收取任何报酬，特此声明。

重要声明

本报告由国信证券股份有限公司（已具备中国证监会许可的证券投资咨询业务资格）制作；报告版权归国信证券股份有限公司（以下简称“我公司”）所有。 ， 本公司不会因接收人收到本报告而视其为客户。 未经书面许可，任何机构和个人不得以任何形式使用、复制或传播。任何有关本报告的摘要或节选都不代表本报告正式完整的观点，一切须以我公司向客户发布的本报告完整版本为准。

本报告基于已公开的资料或信息撰写，但我公司不保证该资料及信息的完整性、准确性。本报告所载的信息、资料、建议及推测仅反映我公司于本报告公开发布当日的判断，在不同时期，我公司可能撰写并发布与本报告所载资料、建议及推测不一致的报告。我公司不保证本报告所含信息及资料处于最新状态；我公司可能随时补充、更新和修订有关信息及资料，投资者应当自行关注相关更新和修订内容。我公司或关联机构可能会持有本报告中所提到的公司所发行的证券并进行交易，还可能为这些公司提供或争取提供投资银行、财务顾问或金融产品等相关服务。本公司的资产管理部门、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意见或建议不一致的投资决策。

本报告仅供参考之用，不构成出售或购买证券或其他投资标的的要约或邀请。在任何情况下，本报告中的信息和意见均不构成对任何个人的投资建议。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。投资者应结合自己的投资目标和财务状况自行判断是否采用本报告所载内容和信息并自行承担风险，我公司及雇员对投资者使用本报告及其内容而造成的一切后果不承担任何法律责任。

证券投资咨询业务的说明

本公司具备中国证监会核准的证券投资咨询业务资格。证券投资咨询，是指从事证券投资咨询业务的机构及其投资咨询人员以下列形式为证券投资人或者客户提供证券投资分析、预测或者建议等直接或者间接有偿咨询服务的活动：接受投资人或者客户委托，提供证券投资咨询服务；举办有关证券投资咨询的讲座、报告会、分析会等；在报刊上发表证券投资咨询的文章、评论、报告，以及通过电台、电视台等公众传播媒体提供证券投资咨询服务；通过电话、传真、电脑网络等电信设备系统，提供证券投资咨询服务；中国证监会认定的其他形式。

发布证券研究报告是证券投资咨询业务的一种基本形式，指证券公司、证券投资咨询机构对证券及证券相关产品的价值、市场走势或者相关影响因素进行分析，形成证券估值、投资评级等投资分析意见，制作证券研究报告，并向客户发布的行为。



国信证券
GUOSEN SECURITIES

国信证券经济研究所

深圳

深圳市福田区福华一路125号国信金融大厦36层

邮编：518046 总机：0755-82130833

上海

上海浦东民生路1199弄证大五道口广场1号楼12楼

邮编：200135

北京

北京西城区金融大街兴盛街6号国信证券9层

邮编：100032