

# AI 人机交互，系统，模型与新商业

行业(海外科技)

评级:

增持

## ——AI 应用系列报告之人机交互



秦和平(分析师)



0755-23976666



qinheping027734@gtjas.com

登记编号 S0880523110003

### 本报告导读:

LLM 驱动全新多模态人机交互方式，Agent 作为人机交互的载体和入口，或改变端侧操作系统的生态，重塑商业模式。

### 投资要点:

- 投资建议:** LLM 驱动的多模态人机交互方式引领新一轮生产力革命。Agent 或改变端侧操作系统的生态，重塑商业模式。苹果将受益于 APPLE Intelligence 支持的 Siri 带来的新生态，软件收入逻辑或迎来颠覆。谷歌将受益于 Agent 重塑人机交互模式后带来的搜索调用量增长，以及交互闭环中谷歌应用全家桶的流量增长。微软 Copilot 依托于 Windows 系统及 Office 套件，变现进度较快。推荐标的：苹果 (AAPL.US)、谷歌 (GOOGL.US)、微软 (MSFT.US)。
- 人机交互方式的变革，显著降低 AI 应用门槛，可能带来新一轮生产力革命。** 由 LLM 驱动的多模态交互方式，大幅提升了自然语言理解能力，加入视觉感知能力，是对人机交互方式的又一次颠覆。我们认为，AI 应用门槛将随着人机交互方式的简化而降低，从而引领新一轮生产力革命。
- AI Agent 是交互的载体和入口，Agent 将集成、统一各类 APP 的入口，成为操作系统级别的“超级 APP”。** 以 Siri 为代表的人机交互能力变革，将改变操作系统的生态，搜索及调用的流量入口从各应用程序回归到硬件端。Siri 的屏幕感知、跨应用操作、行为智能能力，使其具备超级入口的潜质，未来 Agent 将成为操作系统的主导，而应用的形态和功能将被弱化乃至取代。1) 安卓：谷歌推出 Project Astra，具备 Agent 潜力。2) iOS：用 Apple Intelligence 重做系统，Siri 成为人机交互的入口。推出 iOS 18.1 开发者测试版，Apple Intelligence 开始落地。自研端侧模型+云侧模型，模型表现追平头部模型。3) Windows：Copilot+PC 已率先落地 PC 端并开启变现。
- 人机交互变革带来新商业模式。** 1) 苹果：Agent (Siri) 带来新生态，苹果软件收入逻辑或迎来颠覆。Apple Intelligence 的开放的 API 接口将吸引开发者加入，丰富苹果生态，提升苹果税收收入；Apple Intelligence 支持的原生应用可收取订阅费用；可向第三方应用收费，以提升 Siri 的调用量；云端模型可向用户收取调用费用。2) 谷歌：Astra 有望充分释放谷歌全家桶的潜力，复刻“Google Intelligence”。多模态输入、自然语言交互将大大降低软件应用门槛，谷歌应用全家桶的用户数有望突破性增长。相较于 AI 搜索，谷歌的优势在于具备流量入口 (AI 助手 Project Astra) 及全方位的应用生态，多模态交互式搜索将在谷歌生态圈内闭环完成。对于投资者担忧的 AI 搜索的冲击，我们认为因大模型幻觉导致的结果不可靠、用户粘性不强、成本高且变现率较低、无法建立用户反馈的循环机制等问题，AI 搜索短期内仍难以撼动传统搜索引擎。3) 微软：Copilot 赋能 Office，已率先实现变现落地。
- 风险提示:** 大模型技术进展不及预期；AI 应用渗透、用户接受度不及预期；变现进度不及预期；数据安全和隐私风险；政策监管风险。

请务必阅读正文之后的免责条款部分

### 交易数据

行业主要上市公司	市值(百万、美元)
苹果 (AAPL.US)	3,307,355.92
谷歌 (GOOGL.US)	1,997,789.90
微软 (MSFT.US)	3,023,834.34
市值合计	8,328,980.16

### 相关报告

《AI 手机，AI 产业革命的决定性力量》	2024.4.2
《AI 社交网络，人类的远大前程》	2024.7.16

## 目录

1. 多模态 LLM 进化, 颠覆性变革人机交互方式.....	3
2. 安卓、iOS、Windows 系统的人机交互载体: 多模态智能体助手(Agent)	5
2.1. 安卓: Astra 打通安卓体系, Screen AI 具备读屏能力 .....	5
2.2. iOS: Apple Intelligence 重做系统, Siri 成为人机交互的入口 .....	6
2.3. Windows: Copilot+PC 实现 Agent 构想.....	9
3. 多模态人机交互模式与新商业空间 .....	10
3.1. 苹果: Siri 或重塑系统生态, 软件收入逻辑迎来颠覆 .....	10
3.1.1. 短期推动存量设备换新.....	10
3.1.2. 长期看, 系统生态重塑, 打开软件收入空间.....	11
3.2. 谷歌: 交互门槛降低带来调用量增长, 人机交互在谷歌生态圈内闭环完成.....	15
3.3. 微软: Copilot 赋能 Office.....	20
3.4. 投资建议与估值.....	20
4. 风险提示 .....	22

## 1. 多模态 LLM 进化，颠覆性变革人机交互方式

人机交互方式的变革，显著降低 AI 应用门槛，可能带来新一轮生产力革命。回顾人机方式的变革：从命令行交互到图形交互，引爆了 Windows 等一系列操作系统的推广；从图形交互到触控交互，引爆了智能手机的推广。新一轮由 LLM 驱动的多模态交互方式，大幅提升了自然语言理解能力，加入视觉感知能力，是对人机交互方式的又一次颠覆。我们认为，AI 应用门槛将随着人机交互方式的简化而降低，从而引领新一轮生产力革命。

图1：人机交互方式变革历程



数据来源：《人工智能之人机交互》，智源社区，人人都是产品经理，国泰君安证券研究

**LLM 驱动的人机交互：** LLM 取代复杂的软件成为交互入口和中介，使用自然语言即可调用。传统的人机交互方式：根据需求情景、数据类型的不同，用户需要应用不同的软件。交互特点是交互接口繁琐而多样化，且具备一定的交互门槛，例如使用 Photoshop 编辑图像，对用户的专业技能有一定要求。而 LLM 驱动的人机交互方式：LLM 取代各类软件成为交互接口，站在了人机交互的中心位置，各类软件站在了 LLM 幕后，用户只需使用自然语言调用 LLM 即可。现阶段 LLM 表现对用户的 Prompt 较敏感，我们认为随着 LLM 理解能力的提升，最终人与 LLM 的对话门槛将进一步降低。

图2：传统人机交互：以多种软件为交互中介

传统人机交互范式：复杂繁琐且多样化的交互接口



数据来源：《自然语言交互：大语言模型带来的交互方式变革》

图3：LLM 驱动的交互：以 LLM 为交互中介

大模型时代的人机交互范式：统一便捷的自然语言接口



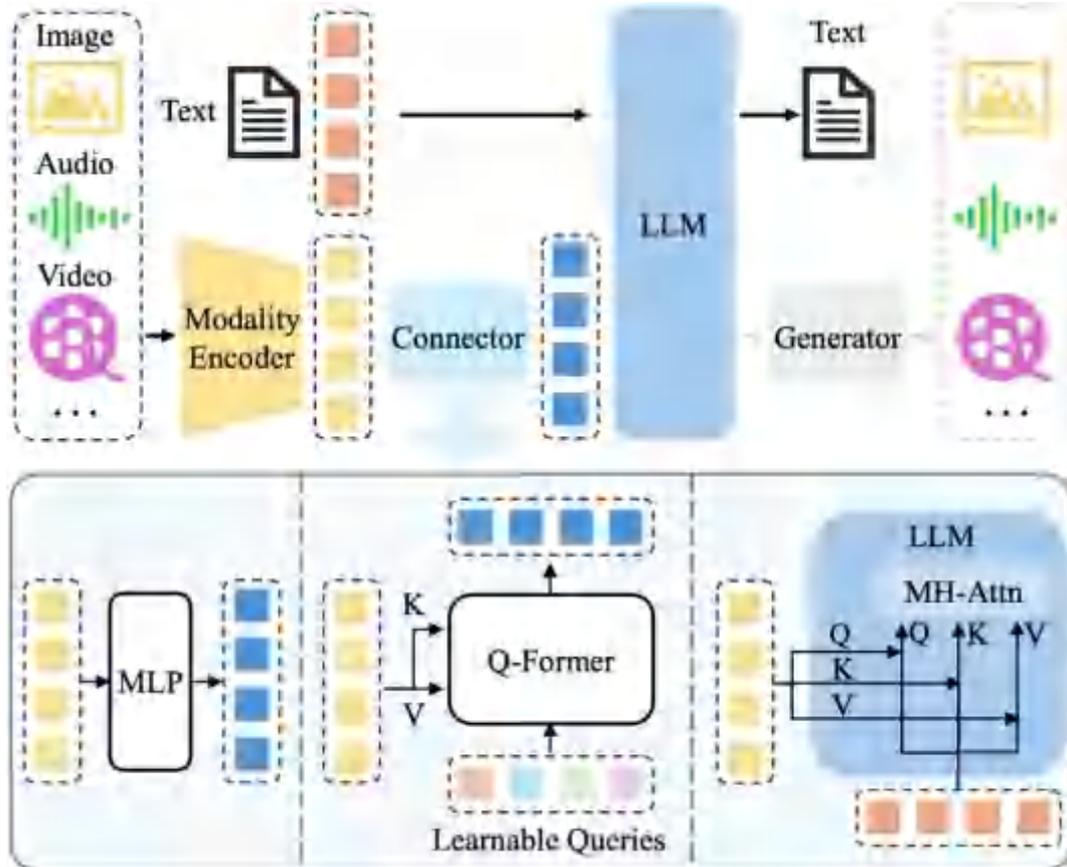
数据来源：《自然语言交互：大语言模型带来的交互方式变革》

**GPT-4o 支持文字、语音、图像三种交互方式，是实现 AI 时代人机交互的技术基底。**GPT-4o 可以接收文本、音频和图像的任何组合作为输入，并实时生成文本、音频和图像的任何组合输出，这种跨文本、音频和图像进行实时推理的能力，向更自然的人机交互迈出了重要一步。

**1) 语音方面:**相较于当下的语音助手(如未引入 Apple Intelligence 前的 Siri), GPT-4o 在三方面具有颠覆性提升: 一是用户可以随时打断 GPT-4o 的话, 不用等它说完就可以继续下一轮对话; 二是模型延迟极低, 对音频输入的响应时间平均为 320 毫秒, 与人类在对话中的响应时间相似 (GPT-3.5 平均延迟 2.8 秒, GPT-4 平均延迟 5.4 秒); 三是模型能够充分理解人类的情感和状态 (累不累, 喘气粗不粗), 自己也能输出各种感情。

**2) 视觉方面:**当前 GPT-4o 的视觉理解能力体现在对图像中物体和复杂场景的精准识别与理解、对人脸表情的识别与情绪检测。例如可以解决手写的方程式、识别人面部表情等。

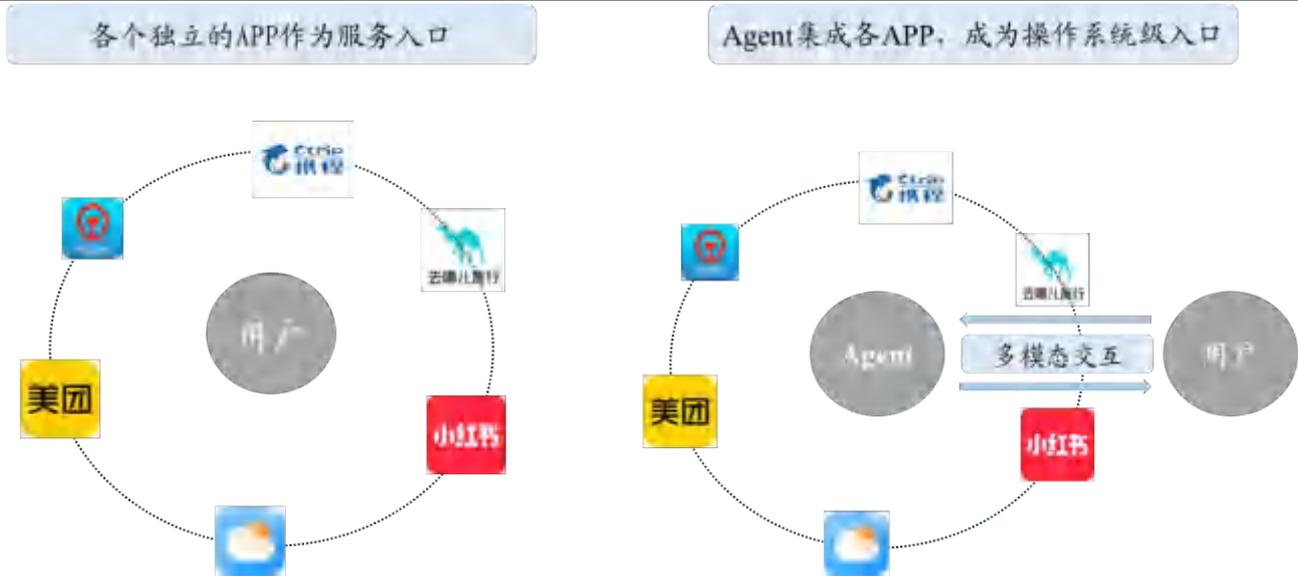
图4: 多模态大模型可支持文字、语音、图像等交互方式



数据来源: 《A Survey on Multimodal Large Language Models》

**AI Agent 是交互的载体和入口，Agent 将集成、统一各类 APP 的入口，成为操作系统级别的“超级 APP”。**LLM 是人机交互的技术底座，而 AI Agent 是交互的载体和入口，用户直接与 Agent 发生联系。当前用户的各类需求由各个独立的 APP 分别承接，以旅行为例: 订票需要 OTA 平台，通勤需要火车票/机票/打车平台，饮食需要美团/抖音等团购平台，制定攻略需要小红书等内容社区，查看天气需要天气类 APP。而 Agent 将作为一个统一的入口，集成各类需求和功能，完成一体化式调度。再以旅行为例: 用户将目的地、预算、时长等信息以自然语言输入给 Agent 后，Agent 会结合用户的个性化偏好，统一调度各个独立的 APP，自主完成订门票、订机酒、选餐厅、制作旅游攻略、穿衣提示等操作。Agent 将成为操作系统级别的超级 APP，用户只需与 Agent 建立联系，由 Agent 打通各独立的 APP。

图5: Agent 将作为未来人机交互的入口和载体



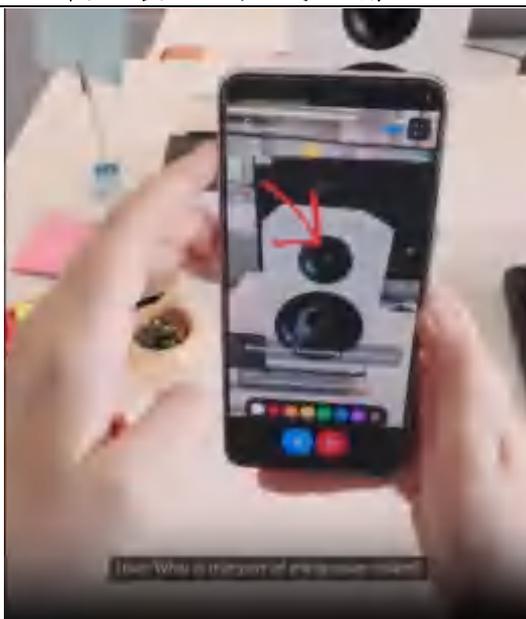
数据来源: 国泰君安证券研究

## 2. 安卓、iOS、Windows 系统的人机交互载体: 多模态智能体助手 (Agent)

### 2.1. 安卓: Astra 打通安卓体系, Screen AI 具备读屏能力

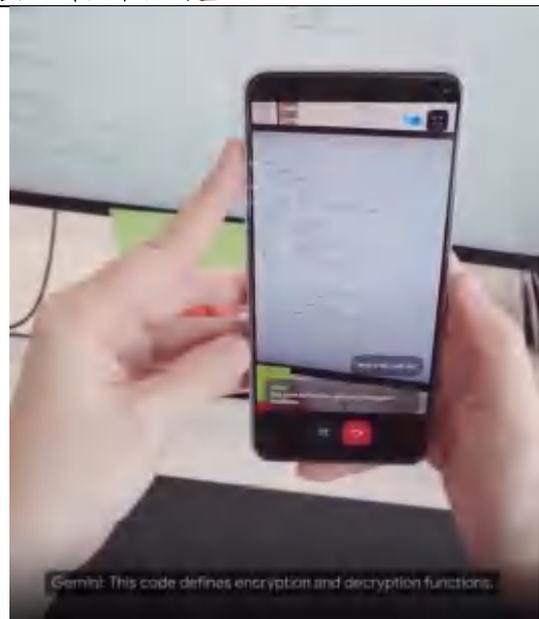
**Android 系统: 谷歌推出 Project Astra, 具备 Agent 潜力。** Astra 是谷歌推出的 AI 智能体, 支持视觉与语音交互。Astra 可理解现实物理世界的复杂信息, 能够横跨物理、编程、文学、地理等多个维度, 与用户在现实世界中进行实时交流。例如, 在与用户交互的过程中, Astra 可识别桌上摆放的音箱, 能根据用户进一步的指向, 对应解释音箱上的高频扬声器的功能; 能看懂电脑屏幕上的算法代码; 识别著名建筑并推理出当前所在位置等。我们认为, Astra 进化出的多模态能力, 是其成长为 Agent 的技术基础。

图6: 用户可以对照片提问、使用自然语言提问, 搜索门槛大幅降低, 实现“对话式”搜索



数据来源: Google I/O 2024

图7: Project Astra 可解读“看”到的代码, 多模态能力是实现科技平权的基础

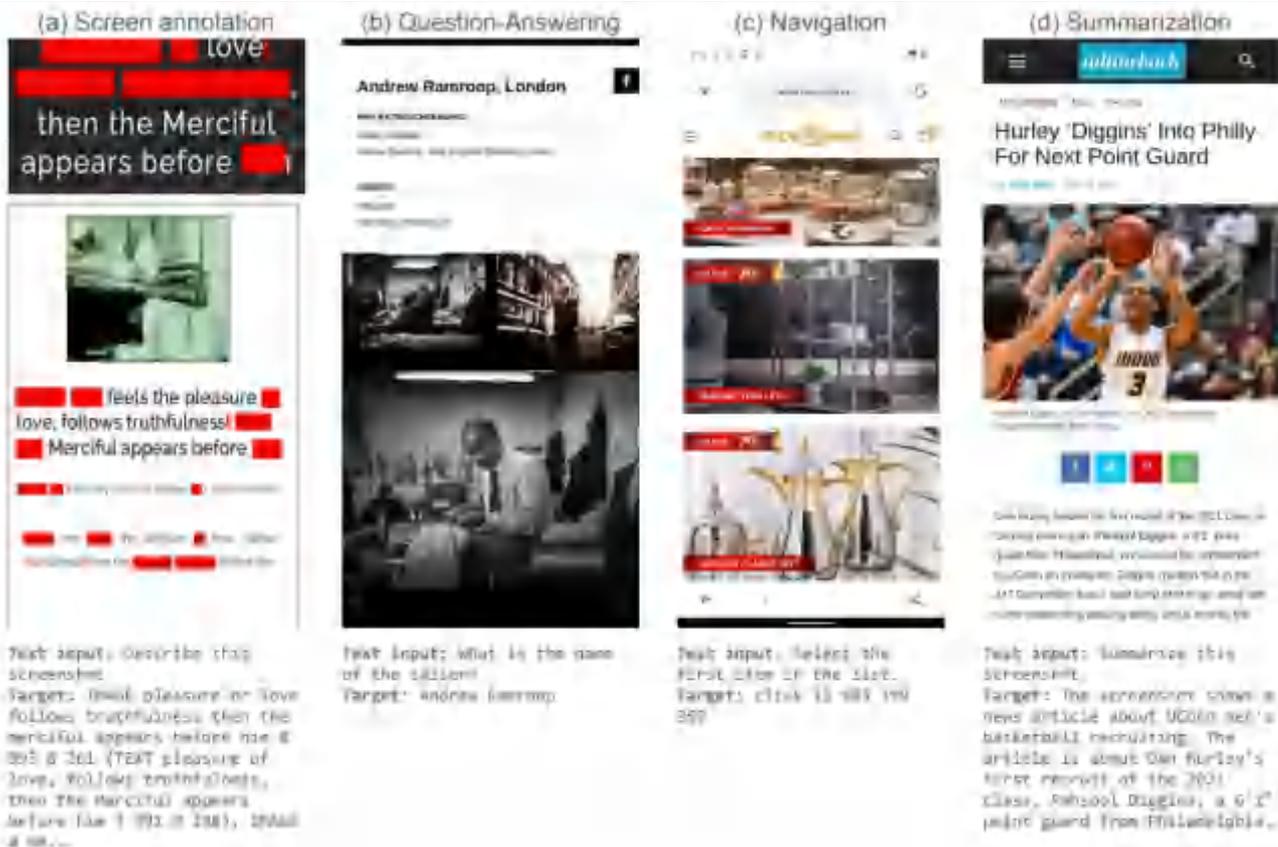


数据来源: Google I/O 2024

Screen AI 具备“读屏”能力, 支持 Agent 进行系统级操作。屏幕用户界面

(UI) 和信息图表(例如图表、图解和表格)在人机交互中发挥着重要作用,但由于其复杂性和多种呈现格式,对大模型的识别和理解能力要求较高。谷歌团队推出 Screen AI,是一种理解用户界面和信息图表的视觉语言模型,具备三种能力:一是问答能力,模型回答与截图内容有关的问题,例如餐厅什么时候开门;二是屏幕导航,模型将自然语言表达转换为屏幕上可执行的操作,例如单击搜索按钮;三是屏幕摘要,模型用一两句话总结屏幕内容。我们认为,UI 对屏幕的理解能力支持 Agent 进行系统级操作,是 Agent 落地端侧设备、颠覆人机交互体验的前提。

图8: Screen AI 的读屏能力: 屏幕注释, 屏幕内容问答, 屏幕导航, 屏幕摘要



数据来源:《ScreenAI: A Vision-Language Model for UI and Infographics Understanding》

## 2.2. iOS: Apple Intelligence 重做系统, Siri 成为人机交互的入口

**iOS 系统: 用 Apple Intelligence 重做系统, Siri 成为人机交互的入口。** Apple Intelligence 的推出,将 AI 内置于操作系统层面,而 Siri 作为语音交互助手,具备屏幕感知能力、跨应用操作能力、行为智能能力,是理想的 Agent、人机交互入口。

**1) Siri 具备屏幕感知能力:** 具备屏幕感知能力的 Siri 将能够理解用户在众多 APP 中的内容,并据其进行相应操作。例如,当用户填写表格需要输入驾照号码时, Siri 可以自动感知、自动搜寻电子驾照并提取相关号码,自动化完成表单填写。

**2) Siri 具备跨应用操作能力:** 结合 Apple Intelligence, Siri 有能力在各类 Apple 原生及第三方 APP 中完成数百种操作。例如,“照片”应用直接编辑用户正在浏览的图片,并在短信或第三方社交 APP 内发送,全过程由 Siri 代为操作。跨应用操作能力,是 Agent 处理复杂多任务的基础。

**3) Siri 可深层次理解用户任务,具备“行为智能”:** 当前 AIGC 涌现的功能主要体现为“内容智能”,即根据指令生成内容,而 Apple Intelligence 支持下的 Siri 涌现出“行为智能”,即基于用户交互行为、理解操作、完成任务的智能,能够对用户的行为进行分析和判断,在产生更深层次的理解后驱动设备及应用完成任务。例如,用户可以提问:“我朋友的飞机几点落地?”

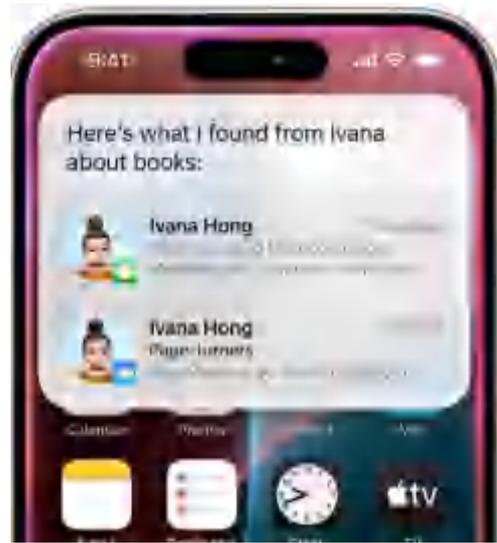
Siri 会自行查询航班详情，然后与航班的实时执飞轨迹交叉对比，给出抵达时间，参考在备忘录中写的用餐地址，帮助计算能否按时到达餐馆。

图9: 屏幕感知能力: Siri 识别信息中的地址, 并自动添加至联系人名片



数据来源: APPLE 官网

图10: 跨应用操作能力: Siri 在信息 APP 和邮件 APP 内搜寻相关信息



数据来源: APPLE 官网

推出 iOS 18.1 开发者测试版, Apple Intelligence 开始落地。7月30日, 苹果发布了 iOS 18.1、iPadOS 18.1 和 macOS Sequoia 15.1 的开发者测试版(目前都仅向每年支付 99 美元的注册苹果开发者开放), 本次更新添加了此前 WWDC 大会上预告的部分 Apple Intelligence 功能, 后续 Apple Intelligence 的进一步更新(如 Apple Intelligence 加持的 Siri) 预计将于 2025 年推出。我们认为本次 iOS 测试版更新标志着 Apple Intelligence 开始落地, 后续 iPhone 16 系列搭载的正式版 iOS 18 将进一步颠覆人机交互体验。

表1: iOS 18.1 开发者测试版更新: 部分 APPLE Intelligence 功能落地

本次更新功能 (AI 相关)	
Siri	在不同请求之间保持上下文关联; 可理解用户说话磕磕绊绊, 或中途改变内容; 回答有关设备功能和设置的问题。
写作工具	在苹果原生应用和第三方应用程序中, 调用 Apple Intelligence 进行文本总结、校对、重写、摘要。
邮件	邮件内容摘要; 置顶重要邮件; 智能回复。
相册	自然语言搜索功能, 只需描述所找照片的画面内容; 查找视频片段中的特定时刻; 只需简单描述即可创建记忆影片, 自动挑选相关照片和歌曲。
尚未实现功能 (此前 WWDC 预告的)	
Siri	调用个人信息的能力; 应用程序内操作的能力; 跨应用操作能力。
图像	图像生成; Genmoji 表情生成。
相册	照片自动清理。
外部大模型	集成 ChatGPT。

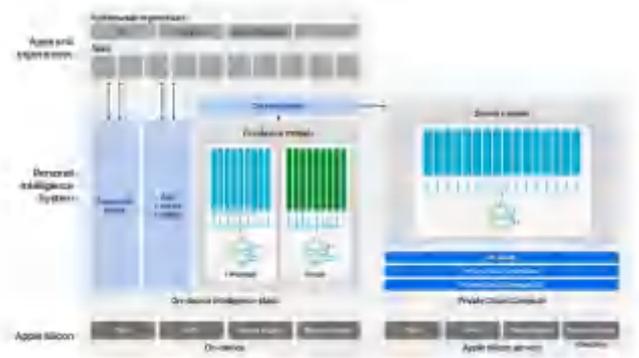
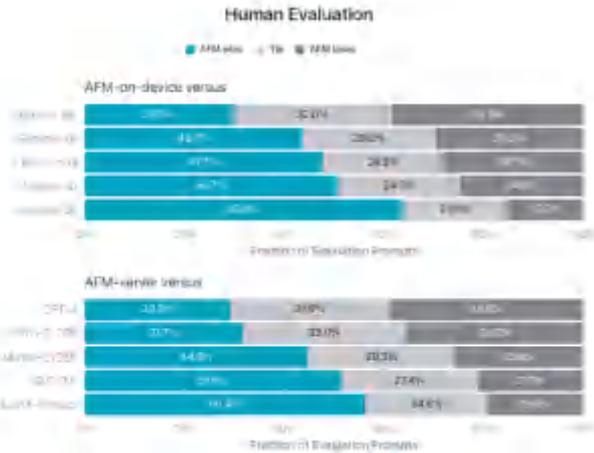
数据来源: 量子位, CNBC, cnBETA, 国泰君安证券研究

自研端侧模型+云侧模型, 模型表现追平头部模型。苹果公布了支持 Apple Intelligence 的两个基础模型: 端侧模型 AFM-on-device (AFM 即 Apple Foundation Model) 约 30 亿参数, 于设备端运行; 云侧模型 AFM-server 为基于服务器的云端模型(参数未披露), 使用私有云计算 (Private Cloud Compute) 系统在苹果的云基础设施中运行, 以保护用户数据。苹果基于大

量的数据优势,对于常用任务,引入了 Adapters,专门针对工具调用、书写、指令理解、摘要等方面进行优化,显著提升在邮件、信息、提醒方面的实际表现。在人类测试中,AFM-on-device 超越了 Phi-3-mini、Mistral-7B、Gemma-2B 等开源模型,接近 Llama-3-8B 的水平,AFM-server 超越 Llama-3-70B、Mixtral-8x22B、GPT-3.5 等闭源模型,接近 GPT-4 的能力。

图11: 模型能力: AFM-on-device 接近 Llama-3-8B, AFM-server 接近 GPT-4

图12: APP、硬件、大模型 (AFM-on-device 和 AFM-server) 的交互逻辑



数据来源:《Apple Intelligence Foundation Language Models》

数据来源:《Apple Intelligence Foundation Language Models》

Ferret UI 具备读屏能力,类似 Screen AI,支持 AI Agent 实现系统级操作。苹果发布的论文中, Ferret UI 是为移动 UI 定做的大模型(类似于谷歌的 Screen AI),可理解手机屏幕并执行任务,具备引用、定位和推理能力。Ferret UI 的屏幕理解能力将辅助 Agent 实现系统级操作,是 Agent 落地端侧设备、颠覆人机交互体验的前提。

图13: Ferret UI: 具备引用、定位和推理能力



数据来源:《Ferret-UI: Grounded Mobile UI Understanding with Multimodal LLMs》

Ferret UI 在三方面做出优化,大幅提升可用性,而可用性是苹果设备领先

于安卓的最大优势。Ferret UI 采用 MLLM 技术路线，针对可用性做出了三方面优化：一是 Ferret UI 能够识别和处理不同宽高比的屏幕，可以理解同一个页面不同尺寸的适配方案；二是增强了识别小尺寸图像的能力，比如 icon、按钮、文本等等，能够更准确地识别和定位屏幕上的各个元素及其功能；三是在识别图像后更进一步，能够理解对话和推断任务，也即理解屏幕内容变化与用户操作结果之间的关系，更好地响应用户的指令。

图14: Ferret UI 能够识别和处理不同宽高比的屏幕



数据来源:《Ferret-UI: Grounded Mobile UI Understanding with Multimodal LLMs》

图15: Ferret UI 能够理解屏幕内容变化与用户操作结果之间的关系



数据来源:《Ferret-UI: Grounded Mobile UI Understanding with Multimodal LLMs》

以 Siri 为代表的人机交互能力变革，将改变操作系统的生态，搜索以及调用的流量入口从各个应用程序回归到硬件端。Siri 的屏幕感知能力、跨应用操作能力、行为智能能力，使其具备超级入口的潜质，未来 Agent 将成为操作系统的主导，而应用的形态和功能将被弱化乃至取代。或许应用只需以 API 的形式接入 Agent 中，以供 Agent 调用，而无需以独立 APP 的形态出现。形态的变化实际意味着话语权的更迭，Agent 以及操作系统将掌握流量的分发权。

应用程序转变为“服务提供商”，向 Agent 开放服务接口。新的人机交互范式下，Agent 直接服务用户，而应用程序为 APP 提供服务，应用展现出原子化的趋势，存在形式或由独立 APP 转变为接入 Agent 的 API 接口。

表2: Agent 与应用程序的博弈: Agent 占据流量分发权, 应用程序形态或改变

应用分类		典型产品	Agent 的威胁	存在形式
社交类应用	内容平台	抖音、小红书	Agent 无法替代用户获得社交体验; 数据隐私性较强	独立的第三方应用
	社交平台	微信、Facebook		
	有线下生态	美团、饿了么	本身有存在价值, 矛盾在于流量分发	独立的第三方应用, 向 Agent 提供调用权
工具类应用	在细分领域有特色	粘土滤镜 Remini	Agent 根据用户的小众需求, 按需调用	原子化组件, 向 Agent 开放接口
	纯工具类	翻译、文本修饰、修图	端侧模型即可完成	被淘汰
	无特色			

数据来源: 国泰君安证券研究

### 2.3. Windows: Copilot+PC 实现 Agent 构想

Windows: Copilot+PC 率先实现 Agent 构想。从交互方式来看，用户只需轻点 Copilot 按键即可快速交互，进行自然语音对话。从功能看，Recall 功能可搜索和回忆用户在 PC 上看到或做过的任何事情，Cocreator 功能可在

设备上近乎实时地生成和优化 AI 图像，并通过实时字幕消除语言障碍，将多种语言的音频翻译成英语；Windows 系统界面里任何东西都可以拖向 Copilot 窗口里，与 LLM 交互。

图16: Copilot 具备实时互动引导、辅助玩家游戏的交互能力



数据来源: Microsoft Copilot

### 3. 多模态人机交互模式与新商业空间

#### 3.1. 苹果: Siri 或重塑系统生态, 软件收入逻辑迎来颠覆

##### 3.1.1. 短期推动存量设备换新

类似 5G 升级驱动的换机潮, 本轮 Apple Intelligence 或将驱动新一轮换机潮。苹果的 AI 端侧运算对终端设备的硬件规格要求较高, 需要搭载 A17 pro/M1 以上规格芯片、内存达 8G 以上才可使用 iOS 18 的 AI 功能, 目前手机机型中仅 iPhone 15 pro 和 iPhone 15pro max 以上机型才能支持 AI 功能, 或将拉动存量用户换新。

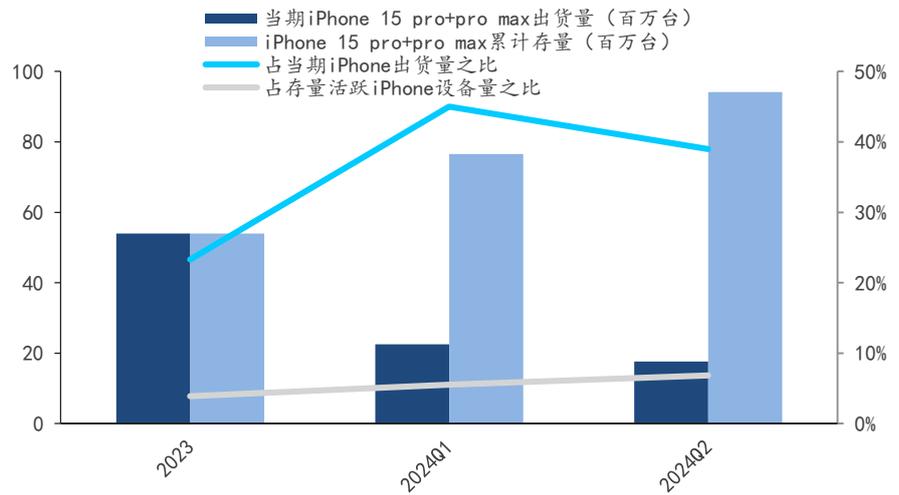
图17: Apple Intelligence 或将复刻 4G 升级 5G 带来的换机潮



数据来源: IDC, 国泰君安证券研究

截至 2024 Q2 仅 7% 的存量设备支持 Apple Intelligence，换新空间广阔。支持 Apple Intelligence 的 iPhone 15 pro + pro max 在 2023 年/2024 Q1/2024Q2 合计出货 5400 万台/2255 万台/1763 万台，占当期 iPhone 出货量的 23%/45%/39%。从存量角度来看，截至 2023 年/2024 Q1/2024Q2 支持 Apple Intelligence 的存量设备数量（即 iPhone 15 pro + pro max）仅占当期 iPhone 存量设备总数的 4%/6%/7%，仍存在大量旧机型无法支持 Apple Intelligence，换新空间广阔。

图18: 支持 Apple Intelligence 的机型仅占存量设备的 7%，有望拉动旧机型换机



数据来源: Statista, Canalsy, CIRP, IDC, 国泰君安证券研究

Apple Intelligence 硬件的升级带来成本上涨及相应的 ASP 上调。以高通骁龙 8 Gen4 为例，台积电 3nm 技术成本上涨（3nm 工艺下每片晶圆的价格大约比 5nm 高出 25%），导致骁龙 8 Gen 4 报价预计比骁龙 8 Gen 3 高出约 25%~30%。我们预计搭载最新 A18 芯片及 8GB 以上内存的 iPhone 16 系列将提振 ASP 重回 900 美元/台，扭转 iPhone 15 系列因降价促销而导致 ASP 下滑的趋势。

图19: iPhone 15 系列因降价促销导致 ASP 同比下滑较大，iPhone 16 系列有望提振 ASP



数据来源: 公司公告, 国泰君安证券研究

### 3.1.2. 长期看，系统生态重塑，打开软件收入空间

端侧 AI 演进下，设备本身集成了 AI 功能，或取代第三方 APP。以苹果

2024 WWDC 公布的最新功能为例，多项 AI 功能集成于操作系统和原生应用中，并可在不同的应用场景灵活调用，将强烈冲击第三方 APP。我们认为，Apple Intelligence 重做原生应用后，设备本身就集成了 AI 能力，或冲击和取代第三方垂类 APP。长期来看，当部分第三方垂类 APP 出清后，原生应用或可实现货币化。

**表3: 苹果推出的原生 APP，具备系统级 AI 功能，冲击第三方 APP**

原生应用	功能	潜在取代对象
Writing Tool	写作助手是操作系统的内置 AI，可应用在手机原生短信或邮件中，也支持所有需要输入的第三方应用，提供文本校对、文本改写、语气切换等功能。	Grammarly 等第三方写作语法辅导软件。无需将文本拷贝至第三方软件，在原生应用场景即可调用 Writing Tool
Math Notes	Apple Intelligence 驱动的计算器，键入或手写数学表达式后系统自动计算结果，绘图功能支持插入对应的图像，单位换算可转换长度、重量、货币等单位	第三方计算器软件
Genmoji	原创表情包，可根据文本描述或真人照片自动生成表情，直接应用在任何消息文本内	Midjourney (文生图)、Remini (黏土滤镜) 等第三方多模态生成式 APP
Image playground	端侧图像生成应用，可选择动画、插画、手绘三种样式，既可直接内置在信息、备忘录、Keynote、无边记、Pages 文稿等原生 app，也可在集成了 Image Playground API 的第三方 app 中使用	

数据来源：WWDC 2024，APPLE 官网，国泰君安证券研究

**Agent 带来新生态，苹果软件收入逻辑或迎来颠覆。**长期来看，Agent 颠覆了当前的应用生态，Agent 成为操作系统的入口和主导，为苹果的软件收入注入了新的逻辑。我们从两个角度：苹果税收入（当前存在的商业模式）和非苹果税收入（新商业模式），分别测算 Apple Intelligence 及 Agent 落地带来的潜在商业空间。

**从苹果税（APP Store 收入）的角度来看：**苹果向开发者开放了 Apple Intelligence 的 API 接口，开发者能够通过 App Intents API 在自己的应用程序中引入 Apple Intelligence，接入到苹果的 AI 生态内，供 Siri 调用。开放的 API 接口将吸引开发者加入，大大丰富苹果生态，APP 数量和付费场景数量都将得以提升，进而提升苹果税收入。我们对 2025 年 APP Store 收入（苹果税收入）做敏感性测试，中性情景下（APP Store 交易额同比增长 15%、Take rate 25.30%）2025 年苹果税收入 288 亿美元，同比 2023 年增长 29%。我们认为 Apple Intelligence 丰富 APP Store 的生态将长期体现在苹果税的收入提升中。

图20: APP Store 收入 (苹果税) 近四个季度同比增速稳定在 12%



数据来源: 公司公告, Sensor Tower, 国泰君安证券研究

表4: 2025 年苹果税收入敏感性测试: 中性情景下 288 亿美元

2025 年 APP Store 收入(\$,mn)		APP Store 交易额同比增长				
		10%	12%	15%	20%	25%
take rate	25.50%	27789	28294	29052	30315	31578
	25.30%	27571	28072	28824	30077	31331
	25.10%	27353	27850	28596	29840	31083

数据来源: 公司公告, Sensor Tower, 国泰君安证券研究

从非苹果税的角度来看: 我们认为, Agent 重塑生态系统逻辑后, 苹果服务收入主要有三个来源的增量。

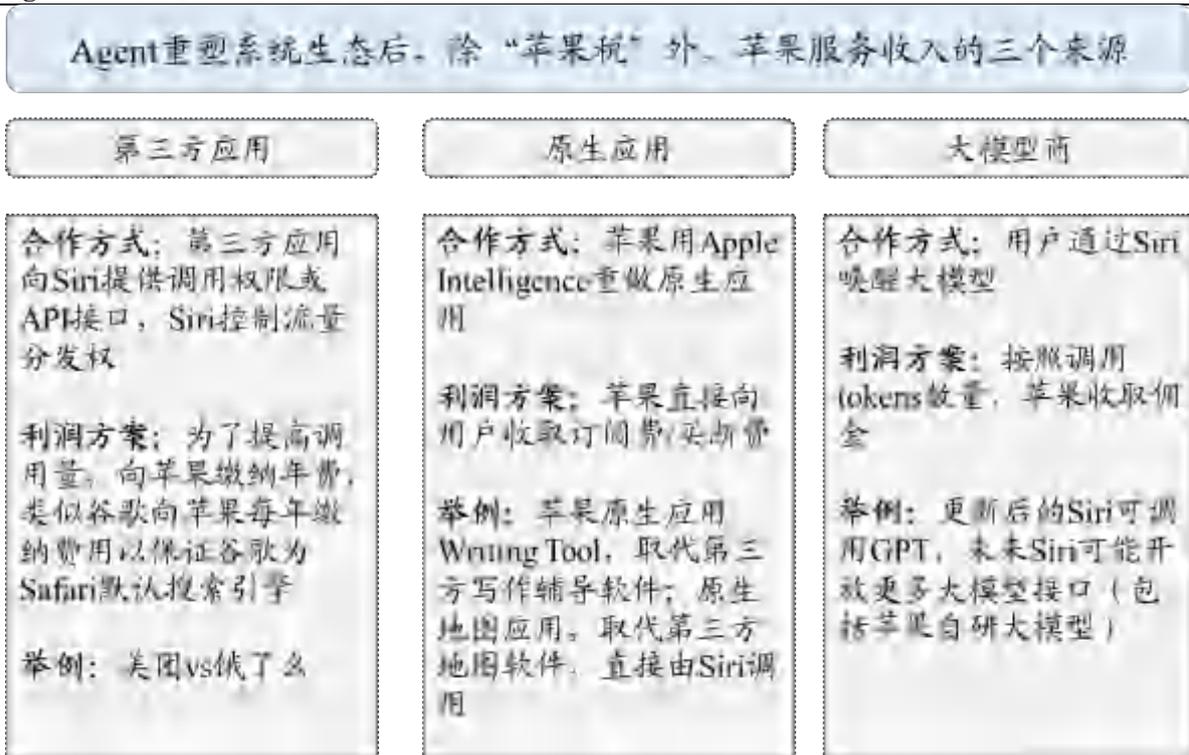
1) 第三方应用: 第三方应用向 Siri 提供调用权限或 API 接口, 对于同质化的服务, 由 Siri 控制流量分发权。比如用户“点一份轻食外卖”而未指定外卖平台, 此时由 Siri 自行选定接入哪一个外卖平台。第三方应用向苹果缴纳推广费及收入分成, 以提升调用量和流量分发权重, 类似谷歌向苹果缴纳广告收入分成以保证谷歌为默认搜索引擎。

2) 原生应用: 如上文所述, Apple Intelligence 重做的原生应用将是苹果变现的另一个来源, 苹果可直接收取订阅费或买断费。我们推测, 苹果可能推出两种收费模式: 一是分别对单个应用收费, 二是对 Apple Intelligence 功能进行打包统一收费。

3) 大模型商: 用户通过 Siri 可唤醒大模型, 除了苹果自研的端侧/云侧大模型外, 也可接入外部模型, 苹果与 OpenAI 达成合作, 用户可通过 Siri 直接调用 GPT。当前的生态下, 并未出现统一的大模型方案, 苹果 22 亿存量设备将会是大模型厂商争夺的富矿, 苹果或可根据调用量收取佣金。

需要指出的是, 短期(约 3 年内) Apple Intelligence 或不会收费, 因 Apple Intelligence 功能尚未完善、消费者习惯和心智需要培养, 短期内无法体现于业绩, 变现将是长期过程。

图21: Agent 重塑生态系统, 苹果可以通过第三方应用、原生应用、大模型商三种渠道变现



数据来源: 国泰君安证券研究

**1) 第三方应用:** 在端侧 AI Agent 与第三方应用之间的博弈中, 我们认为端侧 Agent 为较为强势的一方。一方面, 对于用户模糊不清的指令, Agent 可以结合用户的历史习惯进行自主决策, 例如 Siri 自主选择美团下单外卖。另一方面, Agent 具有基于用户需求和个性化数据进行自主学习的能力, Siri 会主动向用户推送特定的应用和功能, 例如 Siri 在下班时间会主动弹出打车界面并询问是否需要打车。无论何种情况, Siri 都占据了流量入口、拥有流量分发权。我们认为, 新的生态秩序的建立需要较长时间, 长期来看, 变现机会体现在第三方应用为争夺流量而向 Siri 付费, 可参考谷歌为保证成为 Safari 默认搜索引擎而每年向苹果缴纳收入分成 (2022 年谷歌向苹果支付了 200 亿美元)。

**2) 原生应用:**

我们首先从单个应用角度测算潜在变现空间。以原生的 Writing tools 功能与第三方写作工具 Grammarly 做对比。相较于免费版, Grammarly 会员版具有重写句子、调整语气、使用提示词 (prompts) 进行本文生成等进阶功能, 与 Apple Intelligence 支持的 Writing tools 较为相似。我们分别使用自上而下、自下而上两种方法, 测算 Writing tools 潜在的变现空间, 在中性情况下, 两种方法分别测算得出年收入为 2.34/2.37 亿美元, 乐观情况下可达约 10 亿美元。需要指出的是, Apple Intelligence 将重做所有的原生应用并可能开发出新的原生应用, 长期变现空间极为广阔。

表5: Writing tools 潜在变现空间: 自上而下看, 约 2.34 亿美元/年

Writing tools 年收入(\$,mn)		Writing tools 市占率				
		20.00%	25.00%	30.00%	35.00%	40.00%
Rev/DAU (美元/人)	1	62	78	94	109	125
	1.5	94	117	140	164	187
	2.5	156	195	234	273	312
	5	312	390	468	546	624
	8	499	624	748	873	998

数据来源: StartupGeek, Grammarly, 贝哲斯咨询, 国泰君安证券研究

测算逻辑与关键假设: Grammarly 2022 年市占率约 28%, 单 DAU 贡献收入约 1.5 美元/人, 2022 年写作增强助手市场的 DAU 规模约 2.3 亿人次, 假设 2025 年写作增强助手市场的 DAU 达 3.1 亿人次 (yoy+10%)。测算 Writing tools 在不同的市占率水平和不同的单 DAU 收入贡献水平下, 收入的弹性。

**表6: Writing tools 潜在变现空间: 自下而上看, 约 2.37 亿美元/年**

Writing tools 年收入(\$,mn)		付费用户渗透率				
		1.00%	1.50%	2.00%	3.00%	5.00%
年费(美元/ 人)	60	49	74	99	148	247
	108	89	133	178	267	445
	144	119	178	237	356	593
	180	148	222	297	445	742
	240	198	297	396	593	989

数据来源: StartupGeek, Grammarly, 贝哲斯咨询, IT 之家, 国泰君安证券研究

测算逻辑与关键假设: 假设 2025 年苹果活跃设备数量 22.9 亿台, 其中 12% 的设备可运行 Apple Intelligence, Writing tools 渗透率为 30%。参考 Grammarly 付费用户渗透率约 1.14%, 年费为 144 美元/年 (连续包月), 测算 Writing tools 在不同的付费用户渗透率水平和不同的年费水平下, 收入的弹性。

我们再从 Apple Intelligence 整体角度测算变现空间。据 Counterpoint Research 预测, 苹果或对 Apple Intelligence 功能收取 10~20 美元/月的费用。参考苹果目前现有的 APPLE ONE (包括 iCloud、Music、TV+、Arcade) 的订阅费用为 19.95 美元/月, 我们合理假设 Apple Intelligence 收费标准为 20 美元/月, 并保守假设苹果付费用户数 10 亿人 (CY 23Q2 数据)。针对支持 Apple Intelligence 的设备的比例、Apple Intelligence 付费渗透率, 进行敏感性测算, 中性情形下 Apple Intelligence 或带来 600 亿美元/年的收入。

**表7: Apple Intelligence 潜在变现空间: 中性假设下, 约 600 亿美元/年**

Apple Intelligence 年收入(亿美元)		支持 Apple Intelligence 的设备的比例								
		10%	20%	30%	40%	50%	60%	70%	80%	90%
Apple Intelligence 付费渗透率	20%	48	96	144	192	240	288	336	384	432
	30%	72	144	216	288	360	432	504	576	648
	40%	96	192	288	384	480	576	672	768	864
	50%	120	240	360	480	600	720	840	960	1080
	60%	144	288	432	576	720	864	1008	1152	1296
	70%	168	336	504	672	840	1008	1176	1344	1512
	80%	192	384	576	768	960	1152	1344	1536	1728

数据来源: 公司公告, 公司官网, Counterpoint Research, CIRP, 国泰君安证券研究

测算逻辑与关键假设: 假设 Apple Intelligence 收费标准为 20 美元/月, 假设苹果付费用户数 10 亿人 (CY 23Q2 数据)。中性情形下, 约 50% 的设备可支持 Apple Intelligence 运行, Apple Intelligence 付费渗透率可达 50% (作为对比, iCloud 付费渗透率达 60%)。

**3) 大模型:** 用户通过 Siri 调用端侧和云侧的大模型, 我们认为, 端侧模型收入贡献或体现在 ASP 提升, 购机时一次性收取; 云侧模型收入按调用次数收取或采取按月订阅制。我们此处仅考虑文本大模型的潜在变现空间 (不考虑多模态大模型)。参考豆包和文心一言的每日 tokens 使用量, 参考 GPT-4o tokens 输入价格和输出价格, 假设 20% 推理在云端完成、80% 推理在端侧完成, 并假设 2025 年活跃设备中有 12% 可支持 Apple Intelligence, 测算可得用户调用云端大模型可为苹果带来约 0.82 亿美元/天的收入。

### 3.2. 谷歌: 交互门槛降低带来调用量增长, 人机交互在谷歌生态圈内闭环完成

**Astra 有望充分释放谷歌全家桶的潜力, 复刻 “Google Intelligence”。**谷歌的原生应用全家桶与 Agent 可顺利融合。当前 AI Agent 应用落地的一大阻

碍是，因为 Agent（如 Siri）可自主调用各应用，有流量分配权（例如 Siri 选择调用美团点外卖，而不是饿了么），因此部分第三方应用为了自主把握流量入口权，可能拒绝将调用权限放给手机的 Agent。谷歌拥有以搜索、地图、邮件、YouTube 为代表的原生应用，自家的 AI Agent 可顺利获取应用的调用权，因此 Agent 和应用之间的融合度更高。

图22: Google Astra 作为流量入口，具备打造生态级 AI 的条件

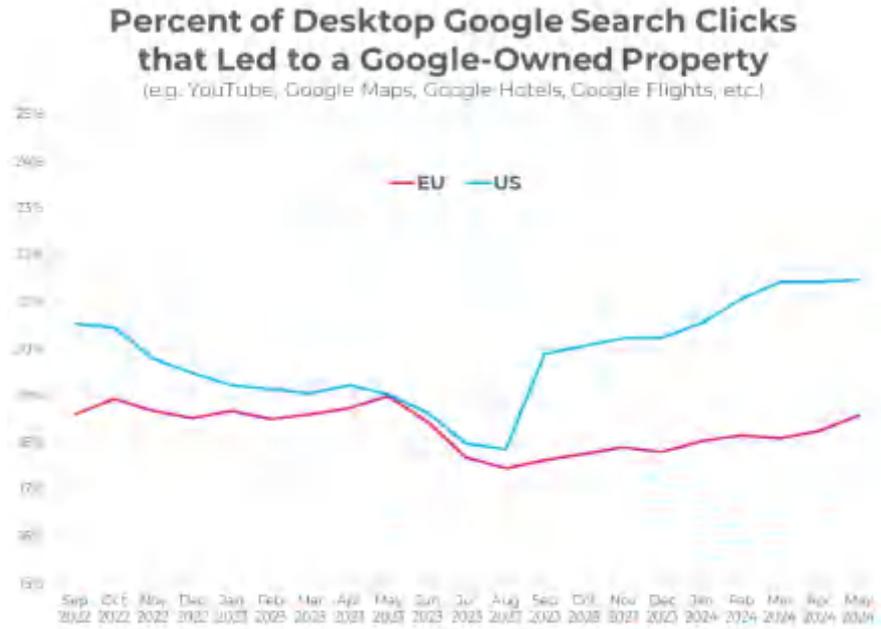


数据来源: 国泰君安证券研究

全新人机交互模式下，多模态输入、自然语言交互将大大降低软件应用门槛，谷歌应用用户数有望突破性增长。我们认为，AI 助手（Agent）的多模态交互能力、语义理解能力，将彻底颠覆人机交互模式。未来的搜索方式，将从过去的打字输入、关键词提问，迭代为语音/视觉输入、自然语言提问，大大降低搜索门槛，实现“对话式”搜索。除搜索外，谷歌的应用全家桶生态都将受益于全新的人机交互模式，例如与地图 APP 导航的语音交互将解放司机双手，与邮件 APP 的语音交互将大幅降低邮件编辑的繁琐性，将图片或视频导入 YouTube 进行编辑只需语音发出指令。我们认为，软件应用门槛大幅降低后，谷歌应用全家桶将迎来爆发式的用户增长、流量增长，推动广告收入增长。

相较于 SearchGPT、Perplexity 等 AI 搜索，谷歌的优势在于具备流量入口（AI 助手 Project Astra）及全方位的应用生态，多模态交互式搜索将在谷歌生态圈内闭环完成。部分投资者担忧 SearchGPT、Perplexity 等 AI 搜索工具将瓜分谷歌流量。Perplexity 获得成功的原因在于，其本质上是对话式答案引擎，而非算法赋能的传统搜索引擎，改进了交互模式，降低了搜索门槛。用户直接输入一句话，无需绞尽脑汁地思考关键词，且 Perplexity 输出具备逻辑性的答案，而非匹配一堆无关的网页。我们认为，谷歌搜索相较于 Perplexity 这类单纯的 AI 搜索产品，优势在于：一方面谷歌的 AI 助手 Project Astra 根植于安卓操作系统中，是掌握流量分配权的“大脑”，用户直接与 AI 助手交互，再由 AI 助手调用搜索引擎，谷歌掌握了流量入口，整个多模态交互式搜索的过程都将在谷歌生态圈内闭环完成；另一方面，谷歌除搜索引擎以外，具备全方位的产品矩阵，可在搜索结果中引用、调用、引流其他各产品，搜索生态更为完善。

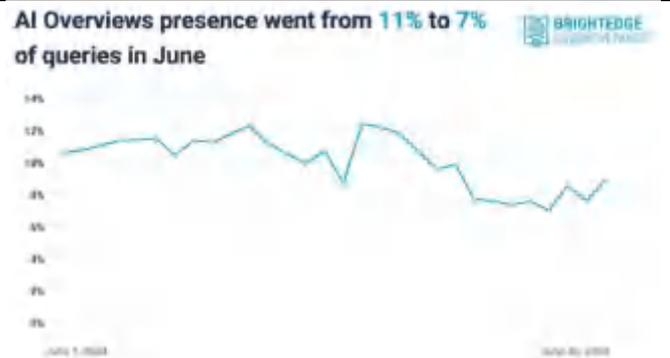
图23: 谷歌将搜索结果引流向自有产品 (YouTube、谷歌地图等), 比例环比提升



数据来源: Datos, SparkToro

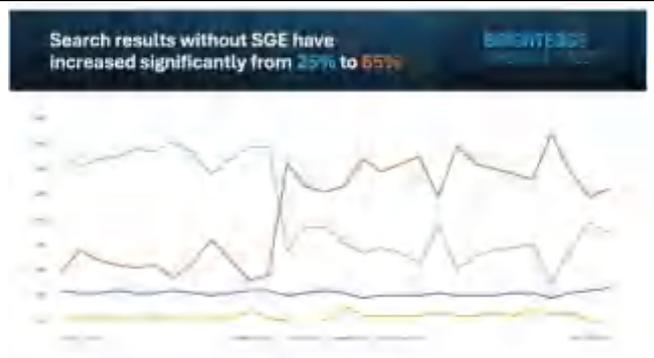
大模型幻觉导致搜索结果不可靠, AI 搜索的渗透率仍处低位。AI 搜索并非知识源, 而是从网络上抓取曾由人类发布的信息, 并进行汇总以生成答案。但 AI 搜索难以分辨检索到的某个前排信息的正确性, 导致结果不可靠, 例如谷歌 AI Overview 错误引用了 Reddit 网友的调侃性言论, 而给出了“添加胶水来防止奶酪从披萨上滑落”的错误答案。BrightEdge 的数据显示, 非 SGE (非搜索生成体验, 即非 AI 搜索) 的 Google 搜索查询量显著增加, 从 2024 年 4 月初的 25% 上升到 4 月底的 65%, 同时 SGE 结果在屏幕上所占的空间也从 1200 像素减少到 1050 像素以下, 表明用户更喜欢传统搜索结果而不是 AI 生成的内容。据 Search Engine Land 数据, AI Overview 出现在搜索结果中的比例由 5 月份的 80% 下降到 6 月份的 7%。

图24: 谷歌 AI Overview 的调用率从 5 月的 80% 下滑到 6 月最低 7%



数据来源: BrightEdge

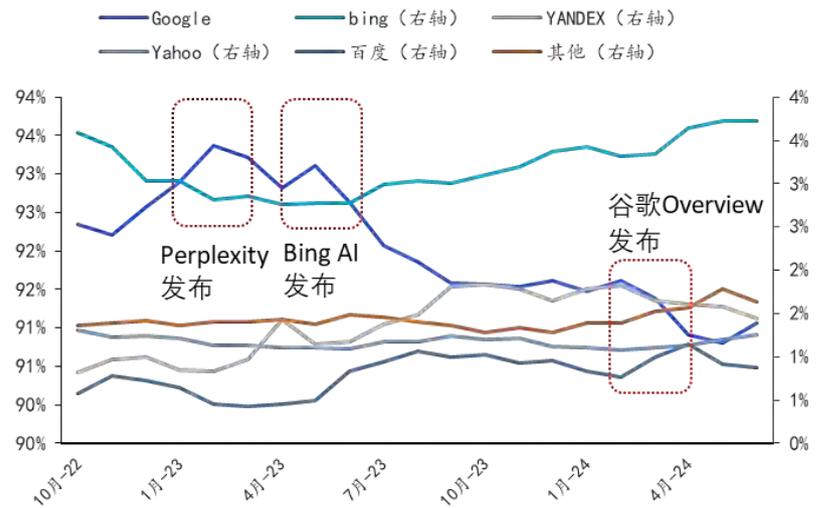
图25: 非 SGE (非搜索生成体验) 的查询量从 4 月底 25% 快速提升至 4 月底的 65%



数据来源: BrightEdge

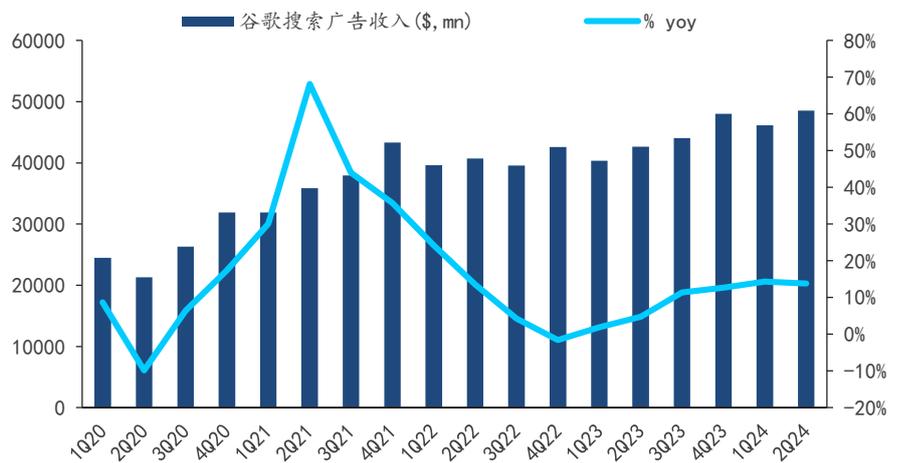
Perplexity 和 Bing AI 搜索的推出并未影响谷歌搜索的市占率, 短期内 AI 搜索难以撼动搜索引擎。从 AI 搜索产品的发布时间节点来看, 2022 年 12 月 Perplexity 发布、2023 年 2 月 Bing AI 搜索发布、2024 年 5 月谷歌 Overview 发布, 均没有对谷歌搜索市占率产生大幅冲击, 2024 年 6 月谷歌搜索市占率 91.1%, 同比 2022 年 11 月仅下滑 1.1%。我们认为当前的 AI 搜索存在结果不可靠 (大模型幻觉问题)、用户粘性不强、成本高且变现率较低、无法建立用户反馈的循环机制等问题, 短期内仍难以撼动传统搜索引擎。

图26: 搜索引擎市占率: 谷歌保持在 91%以上, 受 AI 搜索冲击较小



数据来源: StatCounter, 国泰君安证券研究

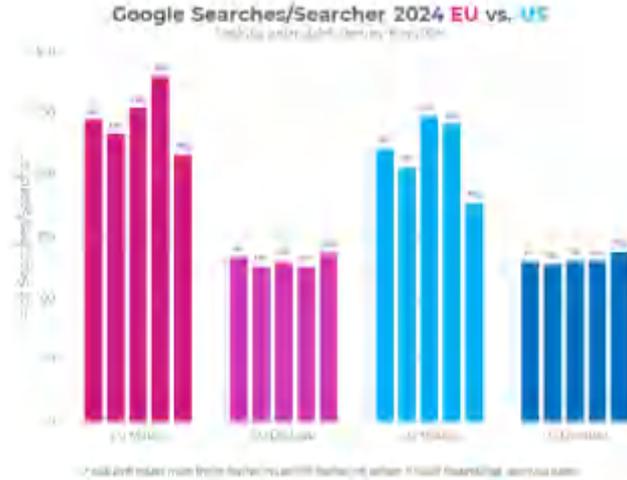
图27: 谷歌搜索的广告收入延续高增, 受 AI 搜索影响较小



数据来源: 公司公告, 国泰君安证券研究

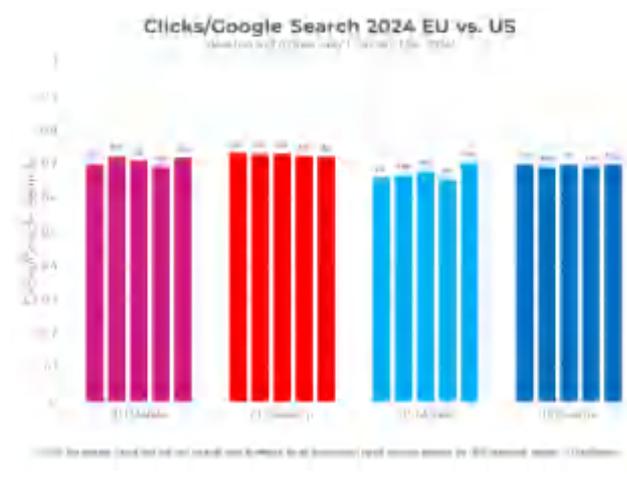
AI 搜索并未显著改变搜索者的行为模式, 并未显著改变搜索量和点击量。谷歌五月在美国推出 AI Overview (而非欧盟), 因此我们对比五月美国和欧盟的谷歌搜索数据, 可以观察 AI Overview 如何改变搜索者的行为模式。从单用户的搜索量看, 5月美国和欧盟的移动端搜索量都环比大幅下滑、桌面端搜索量都环比小幅提升; 从单次搜索的点击量来看, 5月美国和欧盟的移动端点击量都环比小幅提升、桌面端点击量都环比持平。美国和欧盟的搜索数据走向一致, 可以说明 AI Overview 并没有显著改变消费者的行为模式 (搜索量和点击量)。

图28: 五月 AI Overview 在美国推出后, 并未显著影响搜索量的变化趋势 (对比未推出 AI Overview 的欧盟)



数据来源: Datos, SparkToro

图29: 五月 AI Overview 在美国推出后, 并未显著影响点击量的变化趋势 (对比未推出 AI Overview 的欧盟)



数据来源: Datos, SparkToro

即使在 AI 搜索下, 广告的商业模式也仍能跑通。搜索引擎的广告收入逻辑在于, 通过竞价系统来排列搜索结果页上广告的展示顺序以及每次广告点击的费用。AI 搜索下, “排序” 的流程被省略, 直接给出答案, 投资者担忧广告商业模式不复存在。我们认为, AI 搜索模式下, 广告仍能跑通, 主要投放点在于: (1) 首页的热门词条推荐、热门新闻排行榜; (2) 答案生成的先后顺序; (3) 引用链接、延伸阅读等内容的排列顺序; (4) 搜索结果页的图文和视频广告。从落地情况来看, 国内的 360 AI 搜索已初步试水商业化; 百度的简单搜索 APP 的搜索结果依然直接提供购物链接; 谷歌于 5 月在其 AI 搜索引擎 AI Overviews 中投放广告, 用户在查询如何去除衣服上的褶皱时, 搜索结果会展示 Downy 和 Bounce 等品牌的除皱产品的赞助轮播广告。我们认为, AI 重塑搜索引擎生态后, “交互体验改善、搜索门槛降低→用户数提升、搜索量提升→广告曝光量和触达量提升→广告收入提升” 的逻辑链依然成立。

图30: 360 AI 搜索: 首页的热门推荐是潜在的广告展示位



数据来源: 360 AI 搜索

图31: 360 AI 搜索: 结果页的图文视频展示位、引用链接排列顺序、引申阅读排列顺序等, 均可跑通竞价模式



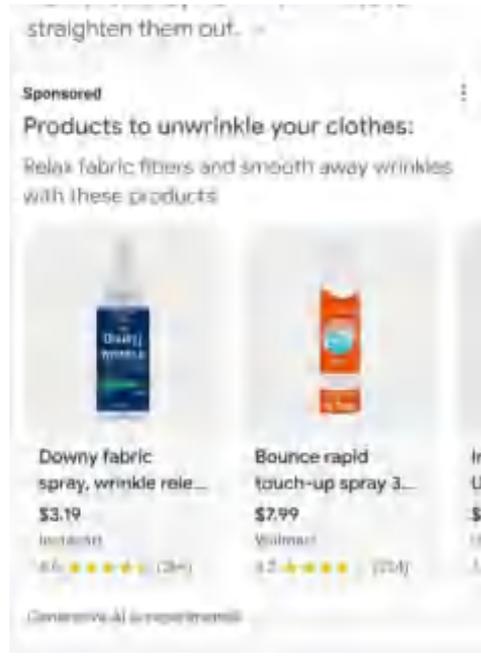
数据来源: 360 AI 搜索

图32: 百度简单搜索 APP: 搜索结果仍然直接展示购物链接



数据来源: 简单搜索 APP

图33: 谷歌于 AI Overview 中加入广告



数据来源: 谷歌 Marketing Live 展示会

### 3.3. 微软: Copilot 赋能 Office

Copilot 渗透带来的增量收入或高于 Office 365 当前存量收入。我们测算 Copilot 渗透率提升为微软带来的增量收入。保守假设 Office 365 用户数与 CY24 Q1 持平 (4 亿), 当前 Copilot 每月订阅价格 30 美元, 我们测算不同的折价溢价率、不同渗透率下, Office 365 Copilot 的年收入。中性假设下, Office 365 Copilot 或带来 648 亿美元的增量收入, 已高于 FY 2024 Office 365 收入体量 (约 485 亿美元)。

表8: Office 365 Copilot 收入敏感性测试: 当前定价、45%渗透率下, 年收入约 648 亿美元

Office 365 Copilot 年收入 (亿美元)		Copilot 渗透率								
		5%	15%	25%	35%	45%	55%	65%	75%	85%
Copilot 订阅价格折价/溢价率	70%	50	151	252	353	454	554	655	756	857
	80%	58	173	288	403	518	634	749	864	979
	90%	65	194	324	454	583	713	842	972	1102
	100%	72	216	360	504	648	792	936	1080	1224
	110%	79	238	396	554	713	871	1030	1188	1346
	120%	86	259	432	605	778	950	1123	1296	1469
	130%	94	281	468	655	842	1030	1217	1404	1591

数据来源: 微软业绩电话会, 微软官网, 国泰君安证券研究

### 3.4. 投资建议与估值

LLM 驱动的多模态人机交互方式, 显著降低 AI 应用门槛, 引领新一轮生产力革命。Agent 作为人机交互的载体和入口, 或改变端侧操作系统的生态, 重塑商业模式。苹果将受益于 APPLE Intelligence 支持的 Siri 带来的新生态, 软件收入逻辑或迎来颠覆。谷歌将受益于 Agent 重塑人机交互模式后带来的搜索调用量增长, 以及交互闭环中谷歌应用全家桶的流量增长。微软 Copilot 依托于 Windows 系统及 Office 套件, 变现进度较快。推荐标的: 苹果 (AAPL.US)、谷歌 (GOOGL.US)、微软 (MSFT.US)。

**图34: 重点公司估值表**

代码	名称	EPS (美元)				PE (x)			
		FY2023	FY2024E	FY2025E	FY2026E	FY2023	FY2024E	FY2025E	FY2026E
AAPL.US	苹果公司	6.2	6.7	7.4	7.8	27.9	32.3	29.2	27.6
GOOGL.US	Alphabet公司	5.8	7.9	8.9	9.9	22.8	20.7	18.5	16.5
MSFT.US	微软公司	9.7	11.8	13.3	15.4	34.8	34.4	30.6	26.4
	中证数	6.2	7.9	8.9	9.9	27.9	32.3	29.2	26.4
	平均值	7.2	8.8	9.8	11.1	28.5	29.1	26.1	23.5

数据来源: Bloomberg, 国泰君安证券研究

注: 预测值为 Bloomberg 一致预期, 数据截至 2024/08/12

## 4. 风险提示

### 大模型技术进展不及预期:

实现人机交互建立在大模型多模态能力进化的前提下，技术进展可能不及预期。

### AI 应用渗透、用户接受度不及预期:

用户对于全新人机交互方式以及 AI Agent 的接受度可能不及预期。

### AI 变现进度不及预期:

当前处于 AI 变现的早期阶段，潜在变现方式和空间存在不确定性。

### 数据安全和隐私风险:

AI Agent 需要调用用户的隐私数据，可能导致数据安全风险。

### 政策监管风险:

AI Agent 落地可能遭到反垄断的政策监管风险。

## 本公司具有中国证监会核准的证券投资咨询业务资格

### 分析师声明

作者具有中国证券业协会授予的证券投资咨询执业资格或相当的专业胜任能力，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，本报告清晰准确地反映了作者的研究观点，力求独立、客观和公正，结论不受任何第三方的授意或影响，特此声明。

### 免责声明

本报告仅供国泰君安证券股份有限公司（以下简称“本公司”）的客户使用。本公司不会因接收人收到本报告而视其为本公司的当然客户。本报告仅在相关法律许可的情况下发放，并仅为提供信息而发放，概不构成任何广告。

本报告的信息来源于已公开的资料，本公司对该等信息的准确性、完整性或可靠性不作任何保证。本报告所载的资料、意见及推测仅反映本公司于发布本报告当日的判断，本报告所指的证券或投资标的的价格、价值及投资收入可升可跌。过往表现不应作为日后的表现依据。在不同时期，本公司可发出与本报告所载资料、意见及推测不一致的报告。本公司不保证本报告所含信息保持在最新状态。同时，本公司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。

本报告中所指的投资及服务可能不适合个别客户，不构成客户私人咨询建议。在任何情况下，本报告中的信息或所表述的意见均不构成对任何人的投资建议。在任何情况下，本公司、本公司员工或者关联机构不承诺投资者一定获利，不与投资者分享投资收益，也不对任何人因使用本报告中的任何内容所引致的任何损失负任何责任。投资者务必注意，其据此做出的任何投资决策与本公司、本公司员工或者关联机构无关。

本公司利用信息隔离墙控制内部一个或多个领域、部门或关联机构之间的信息流动。因此，投资者应注意，在法律许可的情况下，本公司及其所属关联机构可能会持有报告中提到的公司所发行的证券或期权并进行证券或期权交易，也可能为这些公司提供或者争取提供投资银行、财务顾问或者金融产品等相关服务。在法律许可的情况下，本公司的员工可能担任本报告所提到的公司的董事。

市场有风险，投资需谨慎。投资者不应将本报告作为作出投资决策的唯一参考因素，亦不应认为本报告可以取代自己的判断。在决定投资前，如有需要，投资者务必向专业人士咨询并谨慎决策。

本报告版权仅为本公司所有，未经书面许可，任何机构和个人不得以任何形式翻版、复制、发表或引用。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，并注明出处为“国泰君安证券研究”，且不得对本报告进行任何有悖原意的引用、删节和修改。

若本公司以外的其他机构（以下简称“该机构”）发送本报告，则由该机构独自为此发送行为负责。通过此途径获得本报告的投资者应自行联系该机构以要求获悉更详细信息或进而交易本报告中提及的证券。本报告不构成本公司向该机构之客户提供的投资建议，本公司、本公司员工或者关联机构亦不为该机构之客户因使用本报告或报告所载内容引起的任何损失承担任何责任。

### 评级说明

投资建议的比较标准	评级	说明
投资评级分为股票评级和行业评级。  以报告发布后的12个月内的市场表现为比较标准，报告发布日后的12个月内的公司股价（或行业指数）的涨跌幅相对同期的当地市场指数涨跌幅为基准。	增持	相对当地市场指数涨幅15%以上
	谨慎增持	相对当地市场指数涨幅介于5%~15%之间
	中性	相对当地市场指数涨幅介于-5%~5%
	减持	相对当地市场指数下跌5%以上
行业投资评级	增持	明显强于当地市场指数
	中性	基本与当地市场指数持平
	减持	明显弱于当地市场指数

## 国泰君安证券研究所

	上海	深圳	北京
地址	上海市静安区新闻路669号博华广场20层	深圳市福田区益田路6003号荣超商务中心B栋27层	北京市西城区金融大街甲9号金融街中心南楼18层
邮编	200041	518026	100032
电话	(021) 38676666	(0755) 23976888	(010) 83939888
E-mail:	gtjarsresearch@gtjas.com		