# **AI Agent智能体行业深度：产业格局、发展展望、产业链及相关企业深度梳理【慧博出品】**

[AI Agent智能体行业深度：产业格局、发展展望、产业链及相关企业深度梳理【慧博出品】 - 知乎](https://zhuanlan.zhihu.com/p/30292809327)​

目录

收起

一、行业概况

1、什么是AI Agent？

2、发展历程：大模型赋予AI Agent底层技术突破，探索通往AGI之路

（1）学习范式：大模型赋能下，AI Agent迁移学习能力快速提升

（2）工作模式：从过程导向转为目标导向，自动化程度为关键转折点

3、Agent—AI的下一步，包括端侧Agent、应用侧AI Agent两类载体

（1）端侧AI Agent：终端为个人AI助理优质载体

（2）应用侧AI Agent：企服应用为B端Agent优质载体

二、市场现状

1、从被动响应到主动闭环，AI Agent开启智能决策新纪元

2、大模型能力增长遇瓶颈，国内外科技大厂转向AI Agent

（1）国外市场：OpenAI、Google等将注意力从模型转移至AI Agent工具上

（2）国内市场：百度、腾讯、钉钉等互联网大厂加速跟进

3、软件开发门槛大幅降低，千行百业爆款Agent/应用将至

三、具体应用分析

1、自动化类：微软智能体AutoGen

2、情感需求类：陪伴型智能体

3、商业化智能体效能讨论：成本与价值间的博弈

四、产业链分析

1、AI Agent产业链构成

（1）上游基础设施与技术提供商

（2）中游AI Agent研发与集成商

（3）下游应用厂商

2、AI Agent产业链变革

3、AI Agent市场仍处早期生态，端侧及多行业与AI Agent结合迎新商业模式

五、商业模式

1、定价模式1：以ServiceNow为代表的价值抽成模式

2、定价模式2：Salesforce推出AgentForce，按量收费

3、定价模式3：Intercom，按解决问题的成功次数收费

六、产业格局

1、C端：AI Agent或改变内容分发渠道，看好硬件流量入口发展空间

（1）AI Agent改变人机交互方式，手机端或成新的流量入口

（2）市场空间：新流量入口聚焦技术层，将通用能力输送多元应用场景

1）从微信、搜索引擎等超级App看流量入口发展空间

2）从手机厂商的角度，看互联网业务收入增长方式

（3）竞争格局：互联网及手机厂商竞相布局系统级AI

2、B端：AI agent或带动新一轮产业投资，生产工具全面升级

（1）AI agent或带动新一轮产业投资，智能生产解决方案提供商竞争力提升

（2）生产工具有望全面升级，看好智能化硬件

七、相关公司

1、金山办公：办公场景或是Agent关键入口，公司具有流量优势

2、科大讯飞：讯飞星火4.0Turbo正式发布，AI Agent应用矩阵持续扩充

3、彩讯股份：信创邮箱领先企业，AI全栈布局拓展价值边界

4、焦点科技：跨境B2B电商头部企业，AI Agent打开成长空间

5、泛微网络：AI 应用逐步繁荣，协同办公龙头有望受益

6、同花顺：广告收入弹性高于市 场，AI 赋能产品竞争力提升

7、致远互联：短期业绩承压，长期看好AI Agent和行业信创机遇

八、发展展望

1、2025年有望成为AI Agent商业化应用元年

2、AI Agent发展从0到1，市场规模不断增加

3、AI Agent将成为端侧应用重要支柱

九、参考研报

自ChatGPT成为全球现象级产品以来，人工智能领域的发展按下加速键。尽管ChatGPT等AI产品拥 有强大的文本生成能力，但其仍有较大的局限性，如无法完成工具调用的复杂任务、无法自主解决问题 等，加之人们对AI产品的人机交互需求提升，ChatGPT等AI产品已无法满足市场需求，需要自主性、 交互性更强的AI产品。同时，深度学习、自然语言处理、自动化执行等技术的不断进步，也为AI产品 的进一步发展提供了技术基础。在此背景下，AI Agent逐步成为新的关注重点，其商业应用有望在 2025年多点开花。

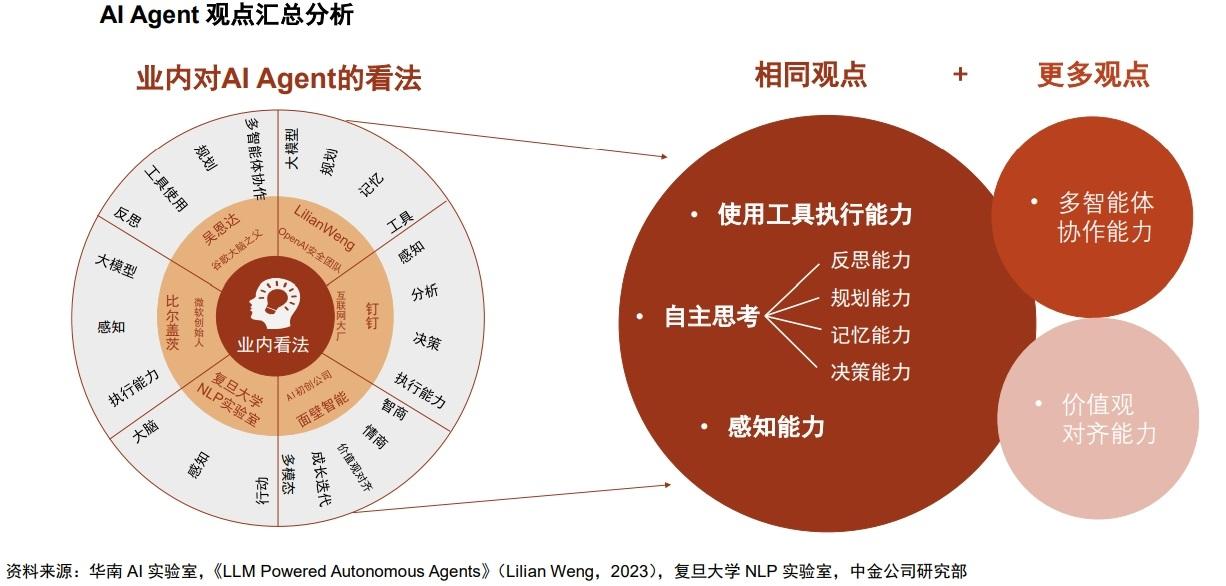
AI Agent（又称人工智能代理），是一种能自主感知周遭环境，通过内在的智能处理进行决策，并执行相应行动以达成特定目的的智能体。其核心特征是通过与环境互动收集数据，独立规划任务路径、调用工具并执行决策，最终在无需人工干预的情况下实现预设目标。

以下内容我们就将围绕AI Agent，对相关问题展开分析和梳理。首先我们将立足AI Agent基础情况，对AI Agent涉及概念内涵、市场现状、具体应用展开分析；其次，将对行业产业链情况、AI Agent商业价值、所将带来的C端及B端产业格局改变、相关公司展开梳理。最后，将着眼于行业发展的视角，对AI Agent的后续发展进行展望，以期帮助大家加深对AI Agent的了解和认知，更好地把握AI Agent的发展脉络。

## **一、行业概况**

### **1、什么是AI Agent？**

**AI** **Agent（人工智能体）是能够感知环境、自主规划、进行决策和执行动作以实现目标的智能体，具有自主性、交互性、反应性、适应性等基本特征**。与大模型和人类通过prompt进行交互的指令导向不同，AI Agent具备了通过独立思考、自主规划、调用工具去逐步完成给定目标的能力，为目标导向。

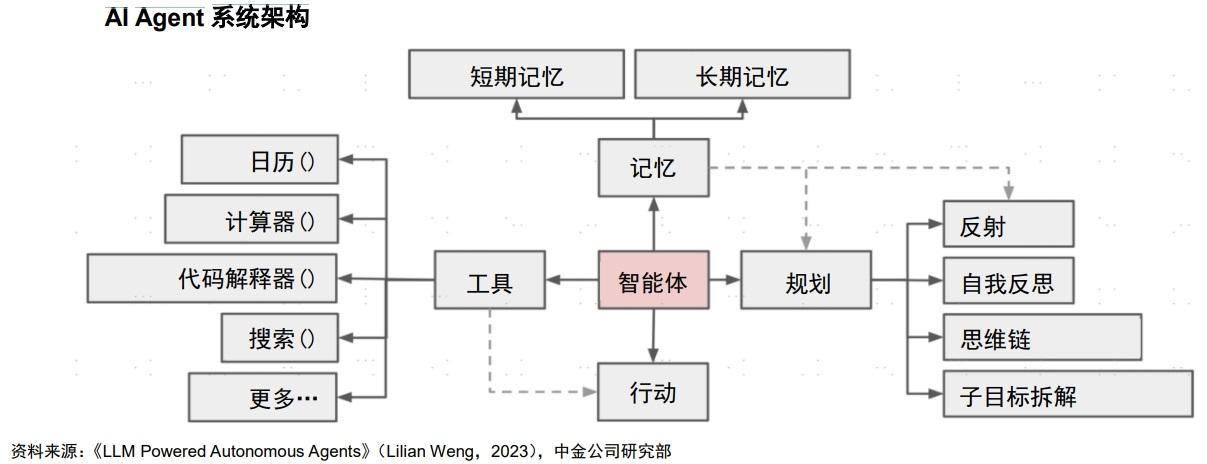


工作原理方面，根据《LLM Powered Autonomous Agents》（LilianWeng，2023），AI Agent**核心驱动力为[大语言模型](https://zhida.zhihu.com/search?content_id=255080798&content_type=Article&match_order=1&q=%E5%A4%A7%E8%AF%AD%E8%A8%80%E6%A8%A1%E5%9E%8B&zhida_source=entity" \t "https://zhuanlan.zhihu.com/p/_blank)，规划、记忆和工具使用是其三个关键的组件，即AI** **Agent=LLM（大语言模型）+Planning（规划）+Memory（记忆）+Tools（工具）**。具体来看，AI Agent通过传感器感知周围环境，使用LLM进行记忆检索、决策推理和行动规划等，结合外部工具获取信息等，最终或使用具体执行器做出响应。

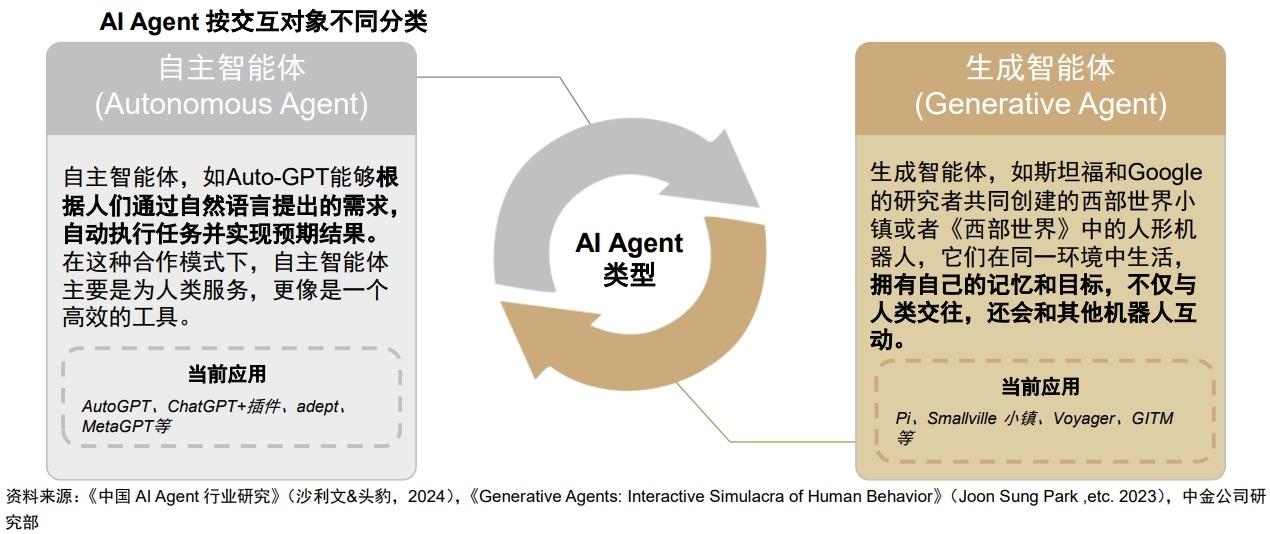
**规划（Planning）**：任务分解：调用LLM通过思维链能力将复杂任务分解为较小、可管理的子目标；反思和完善：通过对过去行为进行自我批评和反思，从错误行为中吸取教训并为未来的步骤进行改进，从而提高最终结果的质量。

**记忆（Memory）**：短期记忆：所有输入会形成短期记忆，以及上下文学习等都是利用模型的短期记忆来学习。短期记忆受到有限上下文窗口长度的限制；长期记忆：提供了长时间保留和调用信息的能力，外部向量数据库可通过快速检索进行访问。

**工具使用（Tools）**：外部工具API：通过使用外部工具API可获取大模型以外的信息和能力，如查询行程、设定闹钟等；插件功能：通过调用插件访问特定数据源等。



**按交互对象不同，AI** **Agent可分为自主智能体（Autonomous** **Agent）和生成智能体（Generative** **Agent）**：自主智能体：交互对象只有人类，是在特定应用场景中根据人类通过自然语言提出的需求，依赖内在决策机制自动执行任务，从而实现预期结果的独立运行智能体，如AutoGPT、ChatGPT+插件等应用形式；生成智能体：交互对象包括人类和其他智能体，如斯坦福和Google创建的西部世界小镇，其中25个智能体在小镇上各自赋予了独特的个性和背景故事，有着组织结交、举办情人节、聊八卦等日常生活。



### **2、发展历程：大模型赋予AI Agent底层技术突破，探索通往AGI之路**

### **（1）学习范式：大模型赋能下，AI Agent迁移学习能力快速提升**

**AI** **Agent发展逾40年，大模型赋能下跨入新的智能体阶段**。20世纪80年代Wooldridge等人将Agent引入人工智能，自此之后AI Agent经历了符号智能体、反应型智能体、基于强化学习的智能体、具有迁移学习和元学习功能的智能体四大发展阶段。由于只能基于学习框架学到特定技能，未能理解从而泛化性较差，此前的AI Agent只能用于特定领域。

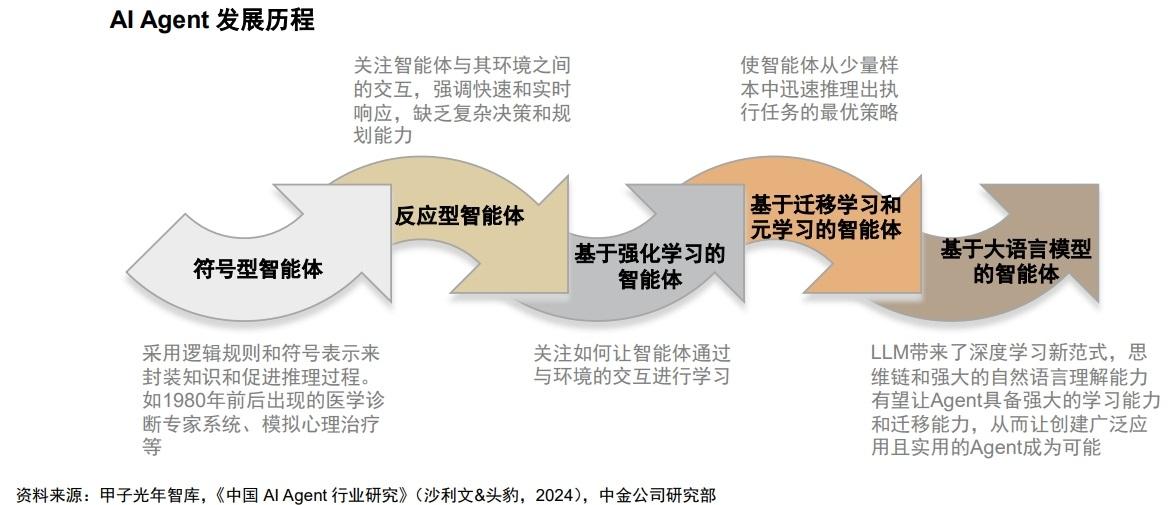
符号智能体：主要通过逻辑规则和符号封装知识促进推理过程，应用在特定技术领域。如1980年左右出现的模拟心理治疗、医学诊断专家系统等。

反应型智能体：为基于反应机制的智能体，能实时感知环境并作出快速响应，主要应用于机器人、游戏AI等领域。如游戏中反应型智能体可通过学习玩家的操作习惯和战术，实时调整自己的游戏行为。

基于强化学习的智能体：主要为通过智能体与环境的持续交互，利用最大化环境奖励来实现优化以期达到更高的智能水平，标志性的应用如围棋领域的AlphaGo。

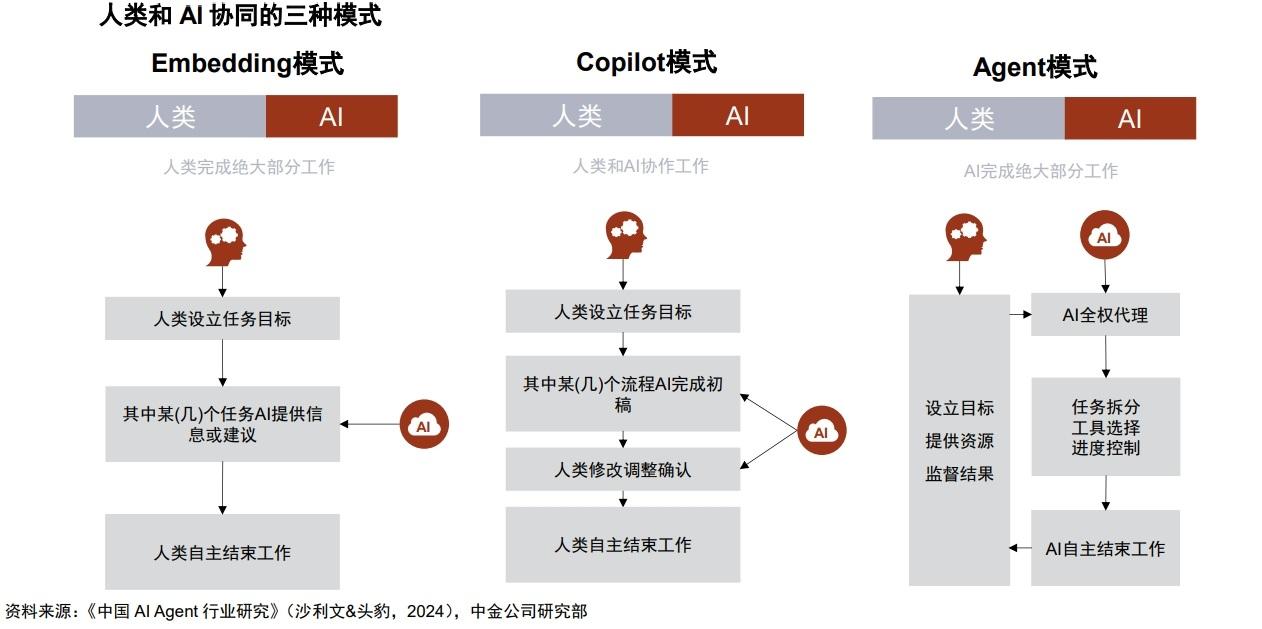
基于迁移学习和元学习的智能体：主要为让智能体从少量样本中迅速推理出实现指令的最优策略。

**LLM带来深度学习新范式，AI** **Agent迁移学习能力快速提升**。此前的强化学习等范式主要集中在增强特定领域的技术能力，但通用能力的发展被忽视，如长期规划、有效泛化、知识记忆等。随着OpenAI陆续发布ChatGPT、GPT-4推进以文本模型为主的大模型迭代，国内外科技巨头均在AI大模型领域加快布局，LLM快速发展但其仍存在的诸多不足如幻觉、上下文容量限制等使得基础大模型无法直接通向AGI。而LLM具备的自然语言理解能力、思维链、涌现能力使得AI Agent具备了较好的学习和迁移能力，处理问题的泛化性效果增强，推动着AI Agent进入了新智能体阶段，成为通向AGI的主要探索路径。



### **（2）工作模式：从过程导向转为目标导向，自动化程度为关键转折点**

**[生成式AI](https://zhida.zhihu.com/search?content_id=255080798&content_type=Article&match_order=1&q=%E7%94%9F%E6%88%90%E5%BC%8FAI&zhida_source=entity" \t "https://zhuanlan.zhihu.com/p/_blank)人机协同模式主要分为嵌入模式（Embedding）、副驾驶模式（Copilot）、智能体模式（Agent）三种**：**嵌入模式**：用户使用文本或语音交流，通过Prompt输入使得生成式AI创作小说、音乐作品、代码等。在该模式下，生成式AI为执行指定命令的工具，人类承担着目标的规划和决策，完成绝大部分工作；**副驾驶模式**：AI介入到工作流程中，承担如为程序员编写代码、检测错误以及优化程序等辅助性工作。在该模式下，AI和人类协作工作，互相弥补短板、纠正错误；**智能体模式**：人类设定目标，生成式AI独立承担大部分工作，人类只负责监督进展及评估最后的成果。在该模式下，AI承担着绝大部分工作，具有独立思考和规划的能力。



基于LLM的AI Agent相较于嵌入模式、副驾驶模式，其工作模式更为高效，主要的差别在于LLM发展带来的“自主规划”能力。嵌入模式、副驾驶模式均需要人类指挥规划，而AI Agent具有自主记忆、推理、规划和执行的全自动能力，过程中不需要人的介入。



### **3、Agent—AI的下一步，包括端侧Agent、应用侧AI Agent两类载体**

**AIGC赋能的第三类场景为AI** **Agent**。Agent为AI应用的下一步，AI由副驾驶逐步转向人类生产生活的独立助手，能够自主调用资源完成任务，人类在其中更多起到监督和评估的作用。AI Agent目前有两类载体，一类为端侧Agent，基于手机、电脑等终端为用户服务，能自主调用终端中的功能与信息；另一类为软件层面的Agent，尤其看到在企业生产管理类软件中，基于企业目前积累的特有知识，实现任务自动化。Gartner预测，到2028年，至少15%的日常工作决策将通过人工智能代理自主完成（这一比例在2024年尚为0%）。

### **（1）端侧AI Agent：终端为个人AI助理优质载体**

手机等终端设备承载大量个人专属信息，能够帮助AI Agent实现个性化服务，同时AI在端侧运行减少用户数据频繁上云，可兼顾AI能力与数据隐私安全，终端侧为个人AI Agent的良好载体。

**目前，各大手机厂商均推出手机端AI**。端侧AI基于蒸馏剪枝后的轻量化模型在手机本地运行，同时端侧算力不够的情况下可连接到私有云上计算，确保用户数据安全和隐私。以Apple Intelligence为例，AI在端侧可实现自然语言理解，且实现基于个人背景的跨应用交互，AI能够以个人信息为基础，从所有应用中检索和分析最相关的数据，以及读取屏幕内容，实现跨应用交互，打通终端信息，成为真正的个人助理。此外，基于生成式AI，用户在手机端还可实现图片生成与编辑，文本改写、校对与摘要，录音转写与摘要，邮件快速回复建议等。目前，Apple Intelligence部分功能已于美国率先上线。除**苹果**外，**小米**、**荣耀**、**三星**、**谷歌**也先后推出手机端AI功能。除手机外，智能眼镜、PC、机器人等端侧设备也为AI Agent成熟载体。

**智谱推出Agent最新进展，探索LLM-OS的可能性**。智谱于2024年11月推出在Agent探索上的最新进展。包括AutoGLM可以自主执行超过50步的长步骤操作，也可以跨app执行任务，支持数十个网站的无人驾驶，以及像人一样操作电脑的GLM-PC启动内测，基于视觉多模态模型实现通用Agent技术探索。更进一步，智谱CEO张鹏将Agent看作是大模型通用操作系统LLM-OS的雏形，与智能设备结合后将为端侧智能化带来新的进展。



### **（2）应用侧AI Agent：企服应用为B端Agent优质载体**

从海外进展来看，随着生成式AI和推理技术的快速发展，AI Agent在B端企业服务场景中的应用潜力正在加速释放。通过与企业积累的大量运营数据深度结合，AI Agent能够基于角色定位和目标导向，实现一定程度的自主操作，有效承担企业内重复性和低附加值的工作。这不仅帮助企业简化流程、提升效率，还通过减少人工投入和优化资源配置，显著降低运营成本，为企业提供了智能化转型的新路径。

**Salesforce的Agentforce是一个典型的应用侧代理平台**。它允许企业在Salesforce平台上构建和部署自主AI代理，这些代理能够自动执行各种业务功能，如销售、服务、营销和商务等。Agentforce的主要特点包括：

**自主操作**：代理可以独立分析数据、做出决策并完成任务，如解决客户问题、筛选销售线索和优化营销活动。

**深度集成**：作为Salesforce平台的一部分，Agentforce代理能够无缝访问和交互各种Salesforce应用中的数据，利用客户信息、触发工作流程和更新记录。

**可定制性**：用户可以使用低代码或无代码工具，快速构建和定制代理，以满足特定的业务需求。

**安全性和信任**：Agentforce通过Einstein Trust Layer，确保代理在使用任何大型语言模型时，Sales force数据不会被第三方模型提供商查看或保留。

Agentforce在Q3最后一周投入生产使用，仅在第三季度就有超过200笔Agentforce交易，且公司在Agentforce层面已看到巨大的潜在需求。Salesforce首席执行官MarcBenioff于近期宣布，公司计划招聘超过1000名员工，以强化Agentforce的推广。



**北京时间2024年10月22日，微软重磅宣布在Dynamics365中集成了10个自主AI** **Agent**。这10个AI Agent基于OpenAI的最新模型o1，具备自主学习能力，能够自动执行跨平台的超复杂业务，节省大量时间和运营成本。微软CEO、董事长Satya Nadella亲自宣布了这一AI新进展，并表示目前已有60%的财富500强企业正使用Microsoft365Copilot来加速业务成果并增强团队能力。

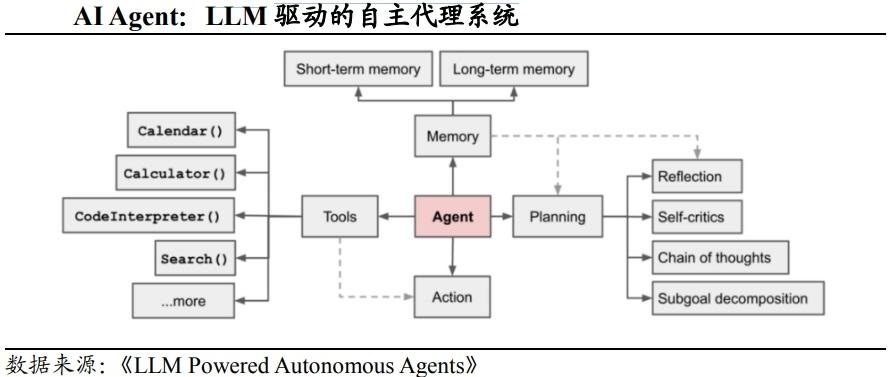


**企业降本增效需求日增，Agent为重要手段，企服类软件为应用侧Agent优质落地场景**。当前企业降本增效需求日增，AI Agent为未来提升企业运营效率重要手段，且企服类软件为此类Agent落地优质场景。

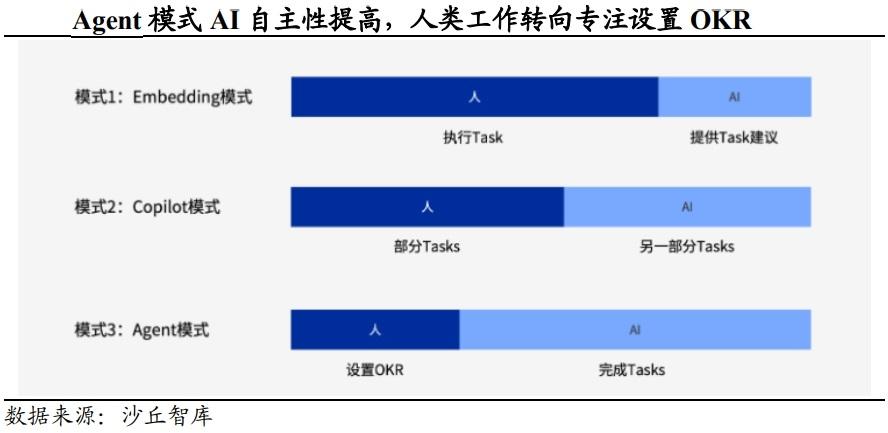
## **二、市场现状**

### **1、从被动响应到主动闭环，AI Agent开启智能决策新纪元**

**AI** **Agent是基于LLM构建的自主决策系统，通过任务分解、记忆管理与工具调用实现复杂目标闭环**。其核心架构以LLM为"大脑"，具备三层能力：规划层将宏观任务拆解为可执行的子目标，并通过自我反思机制优化执行路径；记忆层整合短期上下文学习与长期外部向量数据库，实现跨会话的知识沉淀；工具层动态对接API扩展能力边界，实时获取模型权重外的信息。典型代表AutoGPT通过这三层架构，可自主完成从数据爬取到分析报告生成的全流程，突破传统AI单点工具的限制，形成系统的自治。



**AI** **Agent能够代表用户或系统自主执行任务，通过设计工作流程并调用工具实现复杂功能**。人机协作模式按照自动化程度不同，可以分为三种：Embedding、Copilot、Agent。AI Agent不仅限于NLP等基础的模块化工作，还涵盖决策制定、问题解决、与外部环境交互及执行动作等能力。例如，在软件设计、IT自动化和代码生成等企业场景中，AI Agent利用大语言模型（LLMs）的高级自然语言处理技术，逐步理解用户输入并调用外部工具完成任务。其核心价值在于通过自主性和智能化能力，显著提升任务执行效率并降低人工干预需求。



**AI** **Agent通过自主任务分解与动态规划实现复杂流程自动化**。非Agentic聊天机器人受限于单轮对话模式，需用户逐步输入指令，而如AutoGPT等Agentic系统能自动解析用户需求，生成多级子任务序列。以生成行业分析报告为例，AI Agent可自主规划数据采集、清洗建模、可视化输出全流程，调用代码解释器与数据库接口完成操作，无需人工拆分步骤。这种基于目标导向的架构使其在软件开发、数据分析等领域实现端到端解决方案，突破传统ChatBots线性交互的局限。

### **2、大模型能力增长遇瓶颈，国内外科技大厂转向AI Agent**

### **（1）国外市场：OpenAI、Google等将注意力从模型转移至AI Agent工具上**

**OpenAI、Google、Anthropic正在将注意力从模型转移至一系列AI** **Agent工具上，“贾维斯”未来将至**。

**Anthropic**：Anthropic发布了Claude3.5Sonnet升级版，推出革命性功能——Computer use，可通过API即可实现AI与PC交互，替代人力进行使用计算机控制鼠标、键盘，包括查看屏幕、移动光标、点击按钮和输入文本等操作。

**Google**：根据The Information，Googel正在开发一个专门针对Chrome浏览器进行了优化的AI Agent，其代号为「Project Jarvis」，将为用户执行任务，包括“收集研究资料、购买产品或预订航班”。Project Jarvis将由Gemini2.0提供AI支持，只能在Web浏览器上运行。Project Jarvis在执行不同操作时，中间会有几秒钟的思考时间。此外，根据至顶网，谷歌云计划推出AI Agent合作伙伴计划以及新的AI Agent Space，其目的是让客户在谷歌的在线市场上轻松查找、购买和部署合作伙伴的人工智能代理，推动生成式人工智能销售和客户增长。

**OpenAI**：OpenAI推出代号为“Operator”的新型AI Agent工具。该工具能够在用户的计算机上自动执行任务，如编写代码或预订行程。



### **（2）国内市场：百度、腾讯、钉钉等互联网大厂加速跟进**

**国内智谱抢跑发布AutoGLM，百度、腾讯、钉钉等互联网大厂纷纷跟进**。

**智谱**：智谱发布“Web Browser Use”AutoGLM-Web，一个能模拟用户访问网页、点击网页的浏览器助手。具体来看，可以根据用户指令在私域网站上完成高级检索并总结信息；模拟用户看网页的过程，进行批量、快速的浏览并总结多个网页内容；结合历史邮件信息回复邮件。智谱还发布了AI手机端的“phone use”AutoGLM，只需接收简单的文字/语音指令，它就可以模拟人类操作手机，通过一句话的方式手机便可自动点外卖、发微信、写点评、给朋友圈评论。根据科创板日报，目前，智谱的AutoGLM可适配微信、淘宝、美团、小红书、大众点评、12306、携程、高德地图共8款应用。

**百度**：百度世界2024现场，李彦宏分享了百度在过去一年的技术创新最新成果，并发布了两大AI应用方向：智能体和产业应用。李彦宏现场展示了4种不同类型的智能体，包括公司类智能体，角色类智能体，工具类智能体和行业类智能体。例如行业类智能体中，大会现场展示的法律行业智能体—法行宝，它可以帮助用户解答各类法律相关的问题。上线半年多以来，法行宝为940万人提供了高效、可信赖的法律服务，已经回答了1660万个问题，成为每个人免费的AI律师。

**钉钉**：钉钉迎来重磅升级：AI深入业务场景，钉钉面向企业关键场景推出“精选AI助理”，首批已上线工单助理、Excel助理、法务助理等6大AI助理。同时，钉钉联合生态伙伴发布行业AI解决方案，覆盖制造、医疗、零售、教育等行业。

**腾讯**：腾讯出品的ima.copilot（以下简称ima）正式上线了Windows版本。ima由腾讯混元大模型提供技术支持，是面向学习、办公场景，以知识库为核心的AI智能工作台，是搜、读、写一体的效率工具。



### **3、软件开发门槛大幅降低，千行百业爆款Agent/应用将至**

**人人编写软件时代或将来临，软件开发门槛大幅降低**。根据ARK，在代码编程领域，AI作为代码助手能够帮助人类减少55%的编程时间。此外，AI编程不仅提高了开发效率，还将降低对专业开发者技能的需求，未来人人都是开发者成为可能，使得软件开发更加民主化，更多的人有机会和能力参与到软件开发中。根据Gartner预测，到2025年，70%的应用程序将由低代码和无代码开发平台构建，其中80%的应用程序由非IT工作者开发而成。

**千行百业爆款智能体应用有望陆续涌现**。软件开发门槛的降低有助于涌现大量智能体应用，而具备客户真实场景，致力于解决客户实际需求的厂商有望率先重塑人与技术互动方式，打造出现象级的AI应用，从而占据价值链当中的核心位置。随着AI Agent变得越来越易用和高效，"Agent+"的产品越来越多，未来AI Agent有望成为AI应用层的基本架构，涵盖toC和toB产品等不同领域。目前，零售与电子商务、教育、房地产、人力资源、金融服务等多领域AI智能体应用已落地发布。



## **三、具体应用分析**

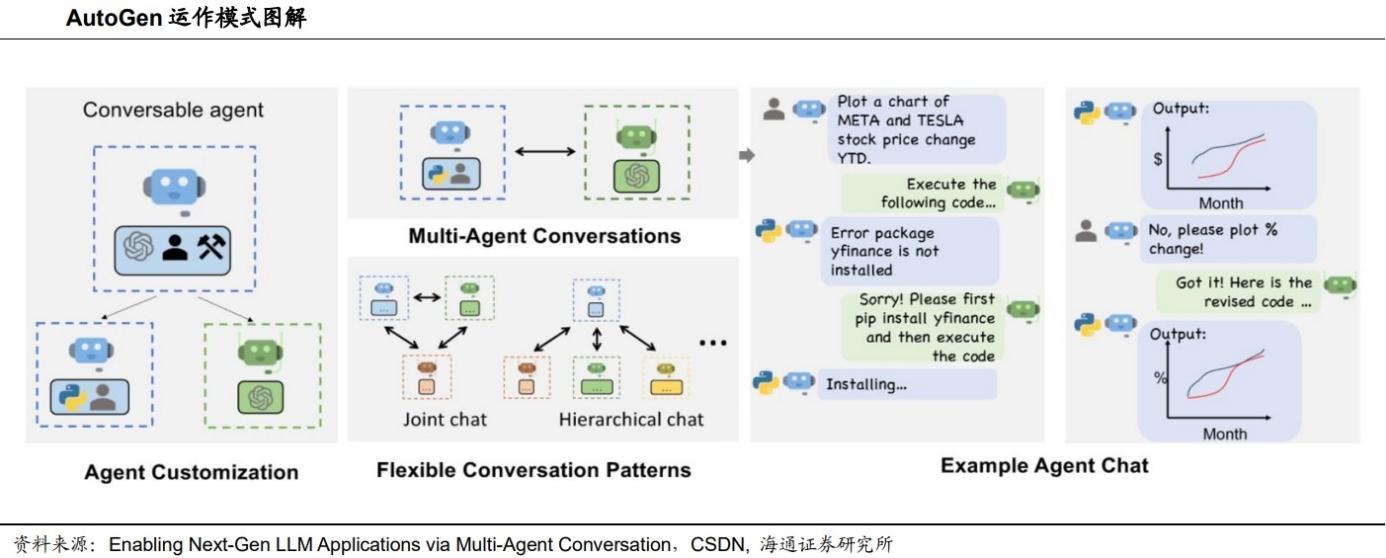
随着大模型竞争加快、政策鼓励研发投入、更多企业参与AI研究等因素，应用层面的AI Agent推进速度加快。智能体大致可以分为六类，根据他们被设计出的特点，可以作用在不同的应用领域上。不同类别的智能体给予应用层面上更多研发方向，像目前关注度较高的自动驾驶技术、智能电网控制、能源管理等都能被垂类智能体覆盖。结合多模态大模型，自动化和情感需求类智能体已落地。但商业化智能体仍需考虑成本问题，由于智能体之间的交互过程可能出现错误循环且输出结果不一定符合需求，tokens成本远高于普通LLMs。



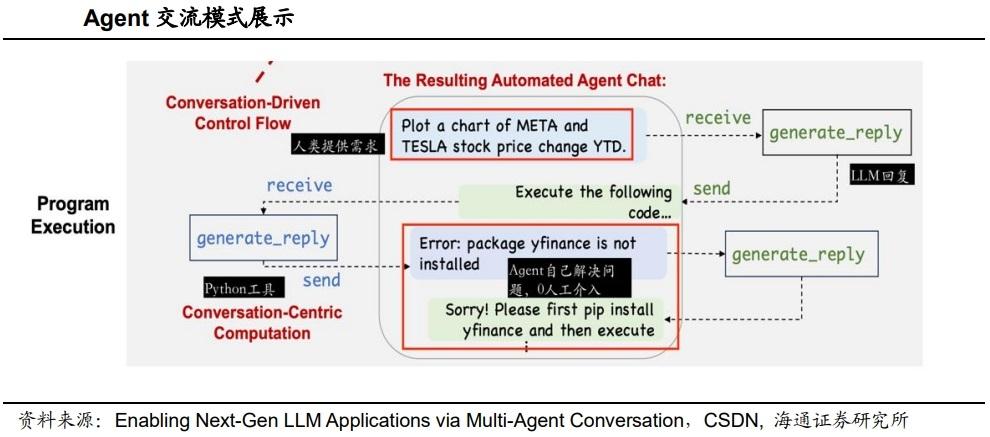
### **1、自动化类：微软智能体[AutoGen](https://zhida.zhihu.com/search?content_id=255080798&content_type=Article&match_order=1&q=AutoGen&zhida_source=entity" \t "https://zhuanlan.zhihu.com/p/_blank)**

微软研发的AutoGen是一种多智能体沟通的模型，这些智能体通过对话完成任务。实验发现这种多智能体沟通模式可以有效提高AI解决问题的准确性以及成果的完整性。使用AutoGen，开发者可以灵活定义智能体交互行为。微软的研究证明了该框架在包括解决数学、编码、问题解答、运筹学、在线决策、娱乐等问题上的有效性。

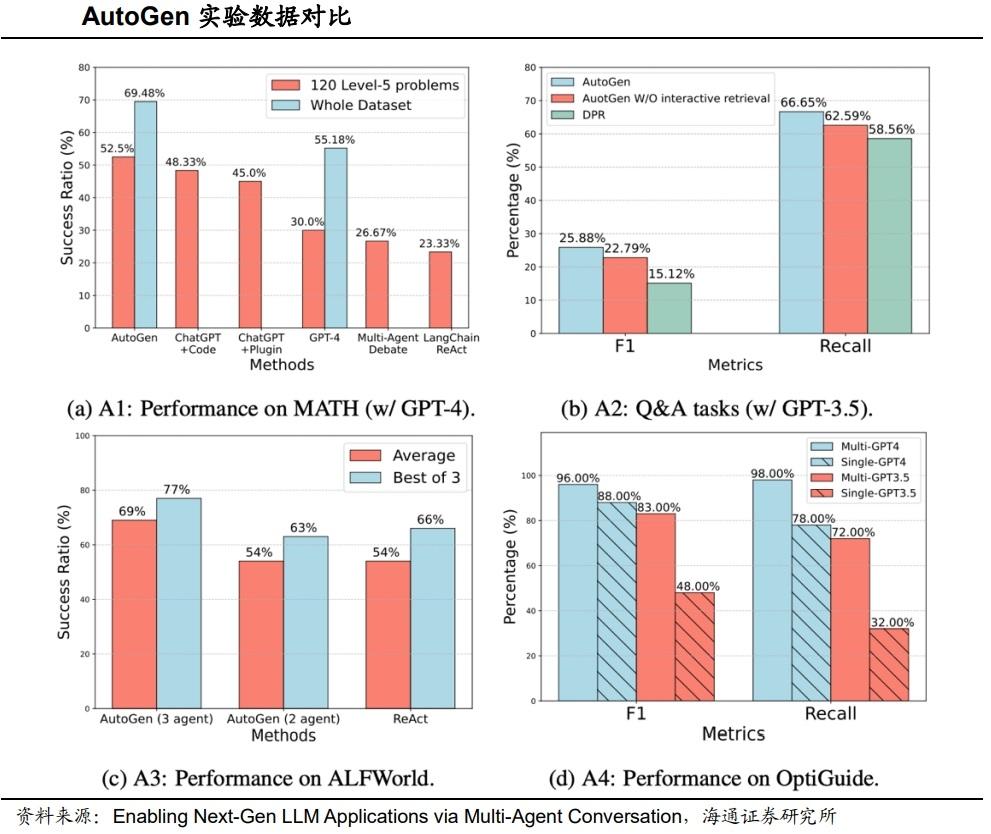
AutoGen需要两个以上的智能体形成交流模式，其中智能体可以由大语言模型LLMs、人类、工具来扮演。当一个需求被提出，智能体会开始灵活交流，他可以是多个智能体向需求人汇报的模式（Hierarchical Chat）或者是像研讨会一样随意讨论（Joint Chat）。在官方给出的示例中，智能体之间的沟通确实减少了人工交互，其有能力自己修复出现的问题并给出满意的答案。



在AutoGen智能体程序中，人类充当需求发起者，智能体通过不断使用工具并在相互交流中解决陆续出现的问题从而得到最终答案。在智能体交流过程中，人类可随时介入其中，若用户对最终结果不满意，可以重新提出需求，智能体就会继续优化答案。



AutoGen在四项实战应用的表现测试中均胜过单一大语言模型，实验发现智能体能增加AI完成任务的成功率以及模型的稳定性。**数学计算**：分别测试了AI在完成整个数学题库和随机120道题目的表现，AutoGen的成功率远超各类语言模型。**数据库检索问答**：先建立一个RAG数据库，然后对智能体进行问答测试，看其能否追踪到正确的答案。其F1-score越高，说明模型更稳定；Recall体现了分类模型对样本的识别能力。Recall越高，说明模型对样本的识别能力越强。从数据看出纯AutoGen模型的表现略高于其他组合。**[ALFWorld](https://zhida.zhihu.com/search?content_id=255080798&content_type=Article&match_order=1&q=ALFWorld&zhida_source=entity" \t "https://zhuanlan.zhihu.com/p/_blank)（Adaptive** **Learning** **Framework** **World）研究**：ALFWorld是专门用于研究和开发智能体的仿真环境，在其虚拟世界中，包含了各种场景、任务和智能体需要完成的挑战。ALFWorld实验发现3个智能体的成功率高于2个智能体，说明随着交互的智能体变多，智能体间互相纠错的能力有所提升。但对于AutoGen的实验也提到多智能体可能会导致错误循环（智能体之间反馈错误但一直无法解决），这也是降低成功率的因素之一。**OptiGuid：**是专门设计来通过自然语言处理，解读和指导复杂供应链优化决策。在用不同语言模型运行[OptiGuide](https://zhida.zhihu.com/search?content_id=255080798&content_type=Article&match_order=1&q=OptiGuide&zhida_source=entity" \t "https://zhuanlan.zhihu.com/p/_blank)时，发现多智能体协作模式的成绩显著高于单智能体。其中多GPT4智能体协作运行OptiGuide时，F1-score和Recall分别能达到96%和98%。



### **2、情感需求类：陪伴型智能体**

智能体另一大应用领域就是提供拟人情感陪伴类服务。陪伴类的智能体拥有长期记忆，让用户与其聊天中产生熟悉且亲切的感觉。随着大模型不断迭代，陪伴类智能体更加立体，能提供更高的情绪价值。目前国内情绪消费类市场潜力较大，在如今紧凑的生活节奏之下，国民孤独感增加，而陪伴类智能体或受益于这样的大环境。根据GIA（Global Industry Analysts）的研究，DennyYin提出该行业规模预计到2026年增长到2000亿美元。

情感沟通类智能体上线，收获大量用户好评。在大模型出现之前，一些利用心理学CBT沟通技术的产品已经在市面上广泛存在，例如Weobot、wysa等。随着大模型的发展和迭代，智能体已经能被打造成心理咨询师或塔罗师的角色。用户通过倾诉生活中的烦恼，并在与智能体沟通中寻求情绪宣泄。目前一款治愈系的软件“林间聊愈室”上线后获得了大量好评，AppStore评分4.7，多数用户给出了五星好评。其中产品使用小动物的角色设定降低了用户的戒备心，加上治愈的画风和场景设计，打造独特的用户体验。

### **3、商业化智能体效能讨论：成本与价值间的博弈**

打造智能体并不是0成本，大语言模型的token消耗就是最直观的成本展示。由于智能体的运行机制内包含反思和深度学习，这使得其在token消耗方面比传统大语言模型高。相比于单智能体，多智能体又多了一层交互端的消耗，当多个智能体讨论且进入错误循环中，无效成本将会显著提升。所以在效能方面，对智能体进行成本讨论以及优化策略十分重要。

在商用智能体中，更先进的大语言模型代表更好的用户体验，但随之而来的是极速上升的成本。目前GPT4的价格是GPT3.5turbo的十倍，而且由于智能体之间交互的特性，在输出过程中往往伴随着更高的token消耗。若最终输出的长度超出模型token限制，则不得不调用更高版本的模型例如GPT4-32K，成本又是GPT4的两倍。

在实际测试中，使用搭载GPT4的智能体模型查找新闻并进行总结梳理。这部分的单次消耗是42000个tokens，成本是1.5美元。在实际应用中，假设一天需要统计的股票新闻大约有120家公司，则成本为35.8美元，按照汇率1：7.14折合人民币255.5元。而且智能体并不是完美的，在运行过程中出现幻觉或错误循环也是常事。例如智能体在找不到合适新闻的时候会进行大量重复的内容输出，人工介入纠错又会增加一定成本。



商业化智能体在成本和最终产品中会有一定博弈，若将智能体运用到股票交易等对精准度需求极高的行业中，产品或引入更多智能体进行协作以提高模型的稳定性，则成本上升。如果运行成本高于其创造的价值时，该智能体模型就失去了商业化的能力。目前多数智能体应用仅停留在测试阶段，市场中的用户需要更理性的看待智能体商业发中能带来的效益。

## **四、产业链分析**

### **1、AI Agent产业链构成**

**AI** **Agent产业链包括上游基础设施与技术提供商，中游AI** **Agent研发与集成商，下游应用厂商、终端用户等。**

### **（1）上游基础设施与技术提供商**

**智能算力中心建设**：随着AI Agent的发展，对智能算力的需求不断增加。政府和企业正在加快智能算力中心的建设，以满足这一需求。这些基础设施为AI Agent的研发和应用提供了必要的计算资源。

**大模型开发**：AI Agent的核心驱动力是大语言模型。因此，大模型的开发者在产业链中占据重要地位，提供基础模型和算法支持，使得AI Agent能够具备感知、理解、决策和执行的能力。

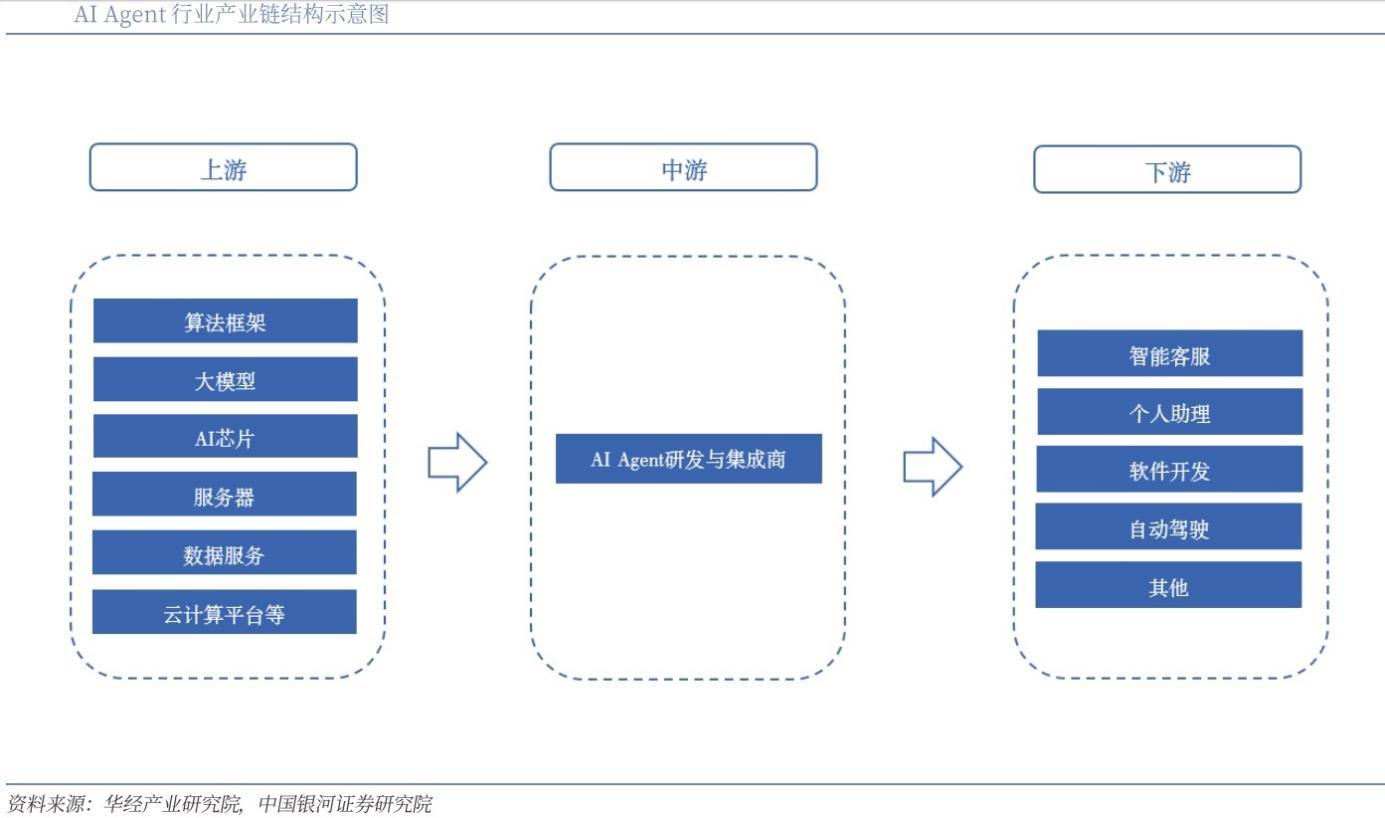
**技术提供**：除了大模型开发者外，还有许多技术提供商为AI Agent产业链提供各种技术支持和服务，如数据标注、模型训练、优化工具等。

### **（2）中游AI Agent研发与集成商**

负责将大模型与其他技术组件(如规划、记忆、工具使用和行动模块)相结合，开发出具有特定功能和应用场景的AI Agent产品。通常拥有强大的研发团队和技术实力，能够不断优化和迭代产品以满足市场需求。

### **（3）下游应用厂商**

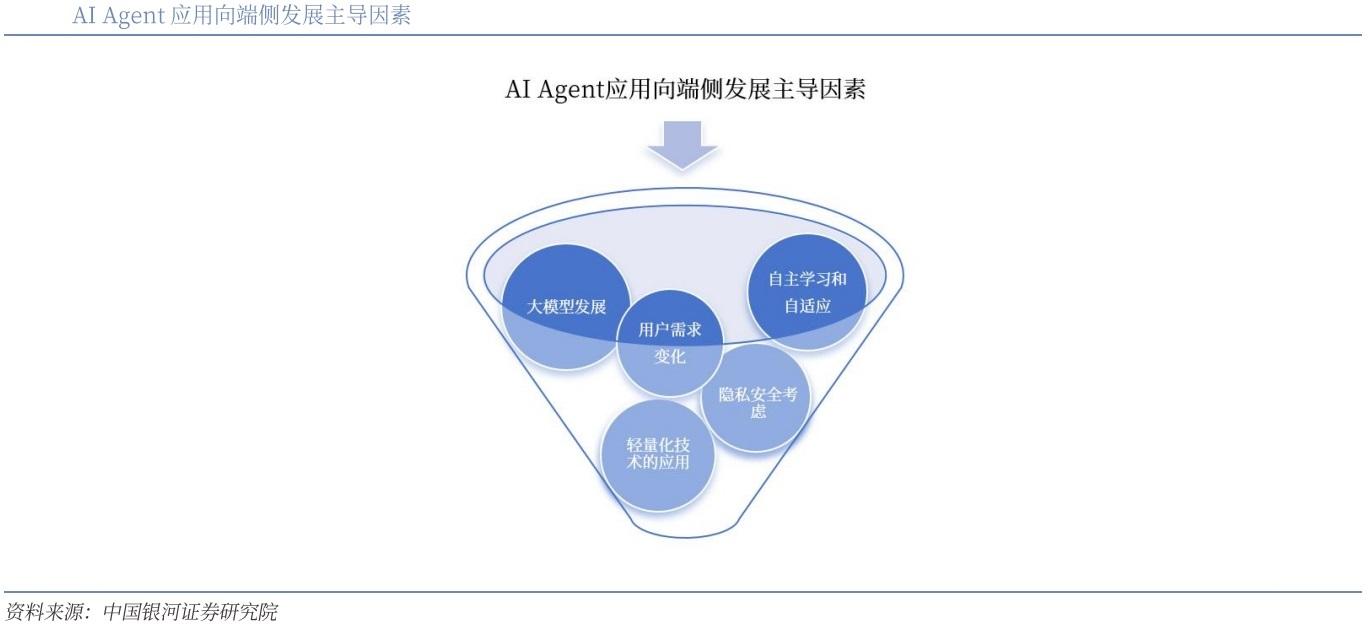
AI Agent的应用场景广泛，包括智能客服、个人助理、自动驾驶、软件开发、财务管理等多个领域。应用厂商根据不同行业和场景的需求，利用AI Agent技术开发出各种应用产品和服务，直接面向终端用户，负责产品的推广和销售。



### **2、AI Agent产业链变革**

GPTs可以看作是AI Agent的早期产品，但在技术栈、构建者技能水平、任务处理等方面存在差异。通过与GPT Builder简单对话，用户无需繁琐编程即可定制出满足从学习娱乐到工作分析各种需求的GPTs。GPTs的推出不仅展现了GPT模型技术的强大，更预示着个性化AI助手的时代即将来临。随着技术进步和行业竞争，GPTs有望接近Agent标准，具备更强自主决策和跨应用任务执行能力，未来AI Agent将受多种因素的影响发展。

**受大模型发展、轻量化技术的应用、用户需求变化、隐私安全考虑等因素推动，Al** **Agent应用向端侧发展趋势明显**。**大模型发展**：生成式人工智能的发展，尤其是大语言模型的出现，为AI Agent提供了强大的能力基础。大模型具备理解、分析、决策和执行任务的能力，使得AI Agent在感知环境、收集信息、自主执行任务等方面表现出色。例如，基于大语言模型的AI Agent在代码生成、数据分析、通用问题解答等领域展现出显著潜力。未来AI Agent将更加擅长处理不同类型的数据输入，如文本、图像、音频和视频等。多模态集成的进步意味着AI Agent可以从多种感官输入中获取信息并做出更为全面的决策。例如，集成多模态数据的智能客服AI Agent不仅可以处理文字聊天，还能分析客户的表情和语调，提供更为人性化的服务。**自主学习和自适应**：通过强化学习等技术，AI Agent将变得更加自主和适应性强，能够在不断变化的环境中自我优化，持续改进其行为和决策过程。这对于在复杂、动态的应用环境中提高AI Agent的实用性至关重要。例如，自适应学习能力可以使工业机器人在生产线上自动调整操作方式，以应对不同产品和生产需求。**轻量化技术应用**：随着轻量化技术的发展，各终端设备制造商纷纷发布搭载端侧大模型的终端产品，这在一定程度上解决了AI Agent在算力、数据等方面的限制，使其能够更广泛地适应不同场景，推动智能技术的普及与应用深化。**用户需求变化**：用户对于智能化产品和服务的需求不断增加，希望能够获得更加便捷、个性化的体验。Al Agent与端侧结合，能够在本地设备上快速响应用户指令提供精准的服务，满足用户的即时需求。**隐私安全考虑**：端侧AI能够更好地保护用户的隐私和数据安全。将AI Agent部署在本地设备上，可以减少对云服务的依赖，降低数据泄露的风险。



### **3、AI Agent市场仍处早期生态，端侧及多行业与AI Agent结合迎新商业模式**

**AI** **Agent市场参与者类型较多，市场仍处于早期生态**。AI Agent的市场参与者包含互联网大厂类、生成AI类、企服SaaS类、创业类、3C类等多类型企业，由于市场处于早期生态，各企业凭借自身优势进行产品打磨与场景探索，抢占市场份额，如AIGC原生类企业借助AI Agent实现AI商业落地；3C消费电子产品借助AI Agent的特性升级自身手机、音响、平板等多端产品用户体验能力。

**AI技术进步诞生智能体，实现电脑手机自主操作**。智谱在安卓系统上推出AutoGLM，使用者无需手动操作，对手机发出指令即可让其打开手机上各类APP，进行网购、点外卖、订高铁票、发微信、抢红包、评论朋友圈等操作。从宣传视频中所展示的功能看，智谱已获得微信、美团、淘宝、大众点评、小红书、高德地图、携程等授权，从而实现人工智能在手机端自主操作，可一次实现程序50步操作。电脑端，Anthropic发布的进化版Claude 3.5 Sonnet可实现自主操控电脑屏幕、移动光标、完成任务，例如科研、求职、编码等一系列工作全可交由Claude大模型实现。由于AI技术的进步，智能体实现手机、电脑自主操作。



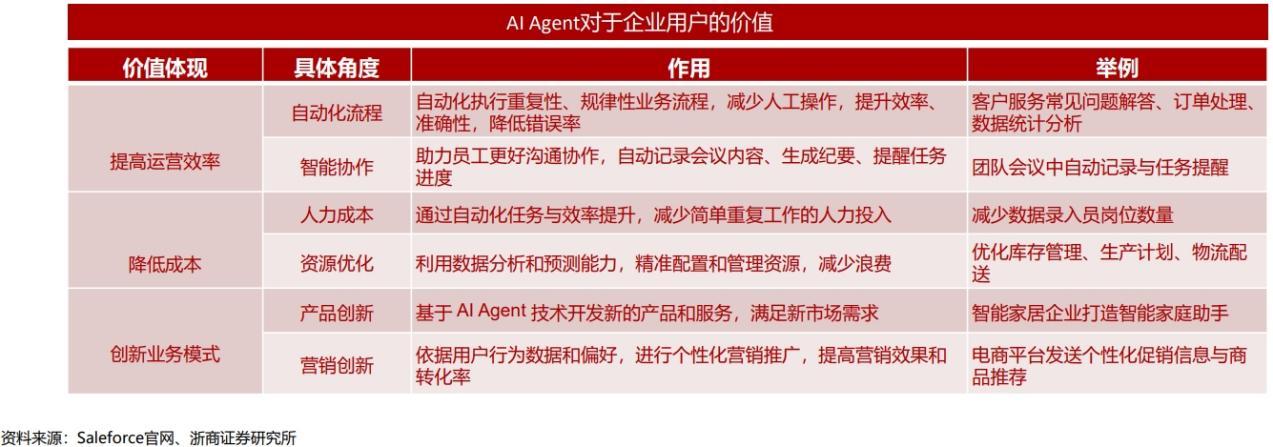
## **五、商业模式**

AI Agent借助AI的能力，极大的解放了人类的生产力。**AI** **Agent的商业价值源于为用户创造价值，即为用户降本增效和提供个性化服务**。

**降本**：AI Agent自主完成任务，通过流程自动化节省人力。Dow公司与微软合作，利用AI Agent优化其物流发票处理流程。AI Agent能够快速分析数千张发票，发现异常并突出潜在的节省机会。在全面推广后，预计第一年即可节省数百万美元的物流成本。

**增效**：AI Agent系统是行业Know-how的载体，可以将企业积累的技能、标杆项目经验等非标准化的知识沉淀和复用；另一方面，AI Agent系统通过工程化的方式，对抗个体工作不确定性，完成管理工作的科学升级，提高群体工作的均值。

**覆盖个性化需求**：AI具备非结构化数据处理能力和实时反馈能力和内容生成能力，基于用户数据的分析和业务流程中的实时反馈，AI Agent可以覆盖更多细分的长尾需求。



### **1、定价模式1：以ServiceNow为代表的价值抽成模式**

Servicenow的AI Agent产品主要包括Now Assist AI生产力平台和Xanadu。

近期ServiceNOW的CFO在参加巴克莱第22届全球科技年会上透露出ServiceNow的AI agent收费模式：**ServiceNow采用了价值驱动的定价策略，确保客户获得90%左右的价值，ServiceNow保留10%**。

ServiceNow生成式AI及AI Agent产品商业化进程持续推进。根据ServiceNow三季度业绩电话会，借助Now Assist，有44个客户的ACV（年度合约价值）超过百万美元，其中6个超过500万美元，2个超过1000万美元。

**AI提升ServiceNow的TEM：从公司业务的潜在市场规模（TAM）看**，公司预计在AI赋能业务矩阵不断延伸下，到2026年公司主营业务潜在市场规模可达2750亿美元，相比于2024年增长约37.5%，其中技术工作流、雇员工作流、客户和行业工作流、创作者和FSC工作流对应TAM分别为1080/150/680/840亿美元。



### **2、定价模式2：Salesforce推出AgentForce，按量收费**

**Salesforce** **AI** **Agent按对话次数进行收费，单次收费为2美元。根据Salesforce官方测算，其应用于智能客服的AgentForce** **Service** **Agent可大幅节约企业运营成本**。

根据Salesforce官网数据，Wiley通过使用Agentforce和带有Einstein AI的Service Cloud管理客户服务运营，将季节性代理的入职速度提升了50%，从而实现了213%的投资回报率和230,000美元的节省。

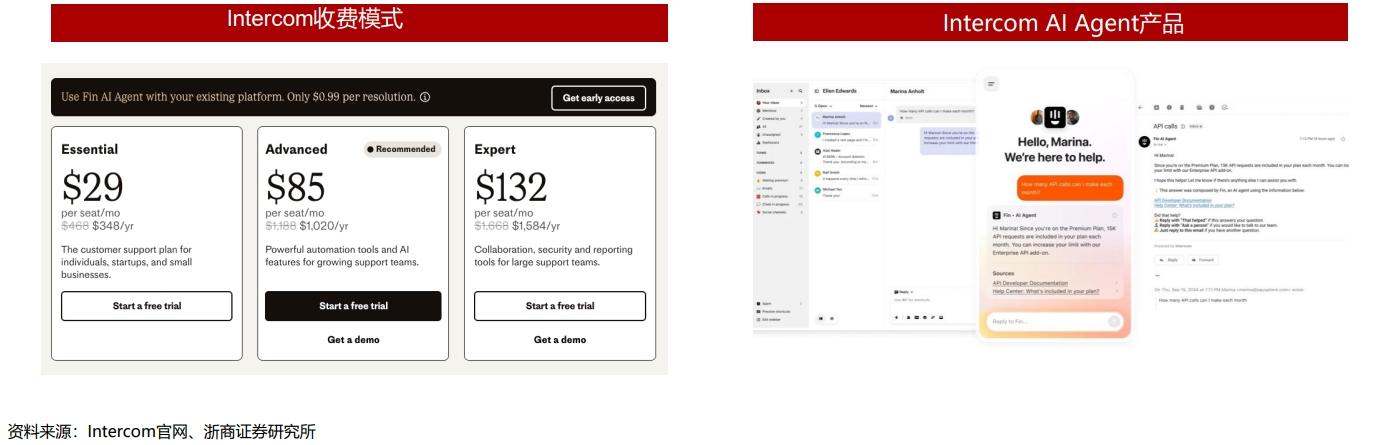
**AI提升Salesforce的TEM：AI** **Agent有望推动Salesforce超4000亿美元的潜在市场空间**。随着AI Agent及数字员工的持续商业化落地，Salesforce估算至2028年公司TAM（可触达市场规模）有望达到4390亿美元，并预计至2028年全球B端企业AI支出有望达到7490亿美元。



### **3、定价模式3：Intercom，按解决问题的成功次数收费**

Intercom公司向企业提供跨自动化和人工服务的客户体验，旨在通过AI赋能减少客户对话中的人工干预，从而提升客户满意度和业务成功率。Intercom公司推出的AI Fin Agent可学习公司和产品的相关信息，并利用所有知识生成准确答案，向每一位客户提供个性化服务并快速调整以满足差异化需求。

Intercom AI Agent应用针对成功解决问题的会话按量收费。Intercom基础服务的收费模式为订阅制，Essential/Advanced/Expert版本的订阅费分别为每年348/1020/1584美元，而Fin AI Agent功能仅在成功解决客户问题后才收费，单次收费为0.99美元。



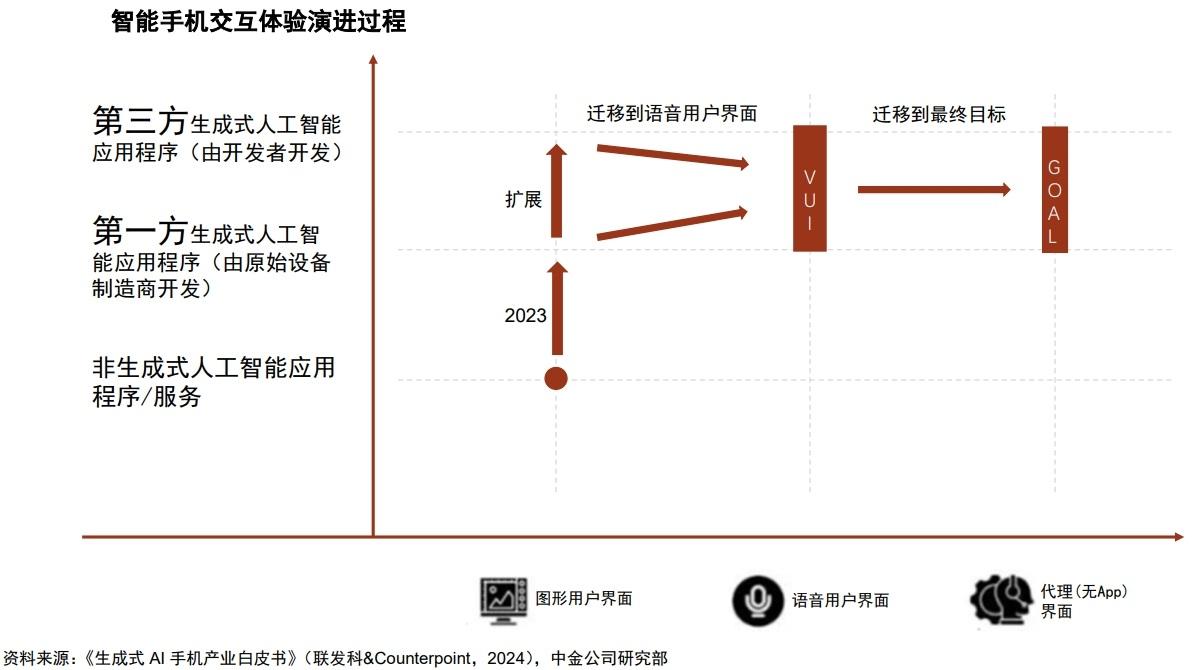
## **六、产业格局**

### **1、C端：AI Agent或改变内容分发渠道，看好硬件流量入口发展空间**

### **（1）AI Agent改变人机交互方式，手机端或成新的流量入口**

**智能手机交互体验从GUI（图形用户界面）发展为VUI**

**（语音用户界面）+GUI（图形用户界面），未来或将升级为Agent（APP-less）UI**。在传统的智能手机时代，软件应用需要产业生态中各厂商在擅长的领域推动产品和服务的更新，体现为移动互联网厂商和应用开发者作为主力军、APP为主要交互界面。随着生成式AI、多模态技术的发展，各厂商将语音助手接入大模型，增强其交互能力。展望未来，随着各厂商将大模型融入底层操作系统，越来越多的功能和服务将被接入到AI Agent，AI Agent或将逐渐成为连接数字生态的入口，用户只需要提出需求（如微信和朋友晚上约地点吃饭），AI Agent能自动运用后台大众点评、百度地图等APP能力给出较好的解决方案。



**手机终端或有望成为新的流量入口，深刻影响手机厂商商业模式**。移动互联网时代APP替代了网页版的搜索引擎，各种APP成为了人们获取垂类信息的渠道，对应应用商店的入口作用得到了凸显（当前手机端，用户主要依靠手动从应用商店下载对应APP产品，APP本身承担了内容筛选的角色，不同的APP对应提供不同类型的内容产品）。Agent（APP-less）UI交互模式下，APP弱化为硬件的后台服务提供方，一方面，APP和应用商店的入口功能被削弱，手机厂商软件生态的商业模式或出现变化；另一方面，AI Agent产品由于高集成度、开发门槛高、对手机系统级硬件适配度要求高等特点，更适合于头部软硬件厂商如手机厂商、云服务商、互联网大厂开发。在此过程中，手机厂商由于天然掌握手机终端硬件入口，预计其有望占得先机，从而更深刻地参与到内容的分发环节。



### **（2）市场空间：新流量入口聚焦技术层，将通用能力输送多元应用场景**

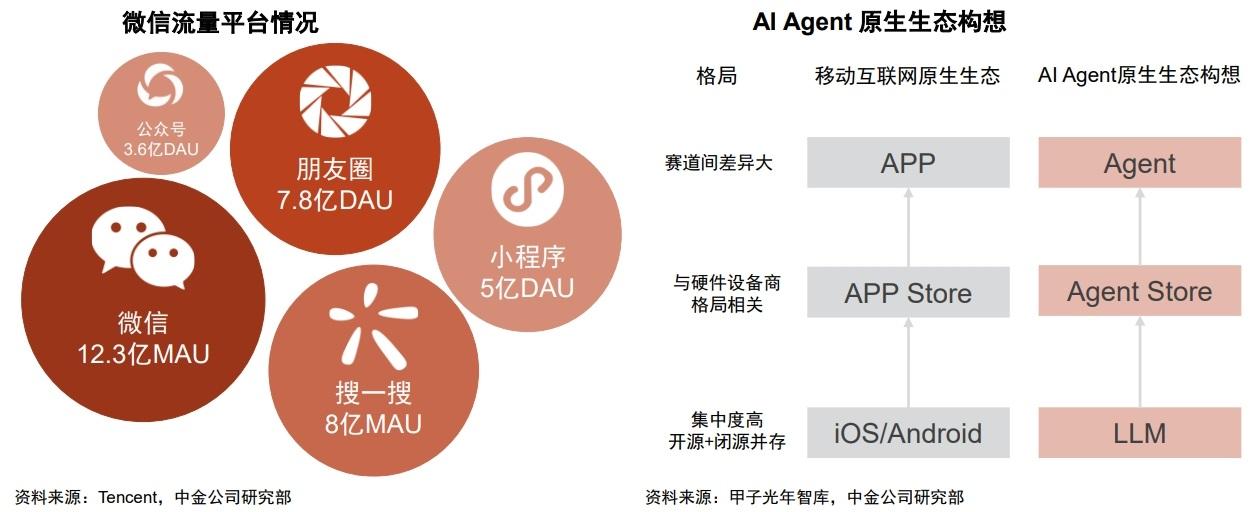
### ***1）从微信、搜索引擎等超级App看流量入口发展空间***

**互联网巨头通过超级App掌握流量入口，通过广告、电商、会员（增值）服务、游戏等核心变现模式，实现较高服务收入**。根据工信部，2023年中国规模以上互联网和相关服务企业完成互联网业务收入1.75万亿元；根据QuestMobile，2023年中国移动互联网月活用户规模突破12亿，全网单月人均使用时长近160小时，其中大平台小程序（微信、支付宝、百度、抖音）去重后月活用户达到9.8亿，成为超级App，为互联网巨头变现带来较大流量。

**微信**：拥有公众号（订阅号、服务号）、视频号（直播、短视频）、小程序、搜一搜等全域流量，通过内部多个场景流转实现相互导流，2023年为腾讯贡献约620亿元广告收入；

**百度**：作为中国最大的搜索引擎公司，通过提供搜索服务获取大量用户数据，并利用数据向广告主提供精确的广告投放，即搜索+信息流双擎驱动，2023年百度广告在线营销收入达751亿元。

AI Agent时代下，App（尤其工具类App）或退居后台服务提供方，手机终端厂商或掌握新流量入口，通过对App的调用实现将通用能力输送多元场景主动满足客户需求。手机终端厂商在流量入口的加持下，在互联网厂商服务收入的分成占比或将进一步显著提升。



### ***2）从手机厂商的角度，看互联网业务收入增长方式***

智能手机时代手机厂商互联网业务收入主要为广告收入、Appstore收入、订阅收入等。AI Agent发展带动下，App弱化为硬件后台服务提供方的背景下，手机厂商作为主要的流量入口或将获得更高的分成比例、以及更多流量对应的广告收入。以苹果为例，苹果软件服务收入主要包括Advertising业务、APPStore收入、数字内容等订阅服务三部分。

Advertising业务方面，其中面向谷歌的License收入是谷歌为作为苹果硬件产品的默认搜索引擎付出的流量获取成本（TAC，Traffic acquisition costs），按2023年谷歌TAC费用180亿美元，计算得出其占比广告收入约7.6%。AI Agent模式下，若该比例每提升1ppt，则将为终端厂商贡献约24亿美元收入，对应约18亿美元毛利。此外，相关人士判断面向软件开发者在AppStore内宣传其开发的App的推广活动收取的Search Ads也有望随流量增长而对应增加。

AppStore业务方面，从客户直接接触的App转为后台内容及服务提供者，软件开发者营收的分成或将下降，对应终端厂商收取的佣金或将增长。

订阅服务业务方面，苹果已经推出的AppleOne订阅服务或成为主流的收费模式，相比此前分开付费，这种模式不仅费用更高，同时用户数也会更多。



### **（3）竞争格局：互联网及手机厂商竞相布局系统级AI**

**手机厂商布局系统级AI，将语音助手升级为智能体**。梳理GenAI端侧应用，AI手机的发展分为“AI App”、“AI功能化”、“AI系统化”三个阶段：ChatGPT、Claude等独立第三方APP出现为“AI App”，是AI手机第一阶段，AI App成为用户与云端LLM交互的路径；随着端侧大模型推出并融入手机OS系统，大模型赋能手机日常应用，如**三星**GalaxyS24手机中画圈搜索、电话摘要、AI图像消除功能，已初具AI Agent雏形，但更多还是集中在功能端；2024年4月华为Pura70发布，智能助手升级为“小艺智能体”；10月**苹果**推送iOS18.1更新实现校对和重写等AI写作工具、照片记忆及全面升级的siri；10月**小米**推出搭载澎湃OS2系统的小米15系列，将小爱同学升级为超级小爱，可实现“帮记”、“帮找”、“帮执行”；10月**荣耀**发布了首款搭载AI Agent的Magic7，集成了AI操作系统MagicOS 9，YOYO助理升级为YOYO智能体，可实现模糊理解、界面识别、自动执行等能力如“一键点饮品”、“一键旅行规划与订票”等。



**互联网厂商相关公司相继推出自主智能体，具备“接管手机”或“接管电脑”的能力**。2024年10月**智谱**推出自主智能体AutoGLM，可根据用户指引完成购物、进入应用程序完成文字填写、总结归纳等操作，主要聚焦在手机应用场景；原理是通过光学字符识别技术理解用户界面的组件，通过思维链训练理解、规划和执行能力、使得其具备按指令规划执行的能力。第三方App适配方面，AutoGLM已经可以执行微信、淘宝、美团、小红书等。此外，2024年11月，Anthropic也推出“Computer use”功能，Claude 3.5可通过自主识别屏幕操作光标完成复杂任务，主要聚焦在电脑应用场景。



### **2、B端：AI agent或带动新一轮产业投资，生产工具全面升级**

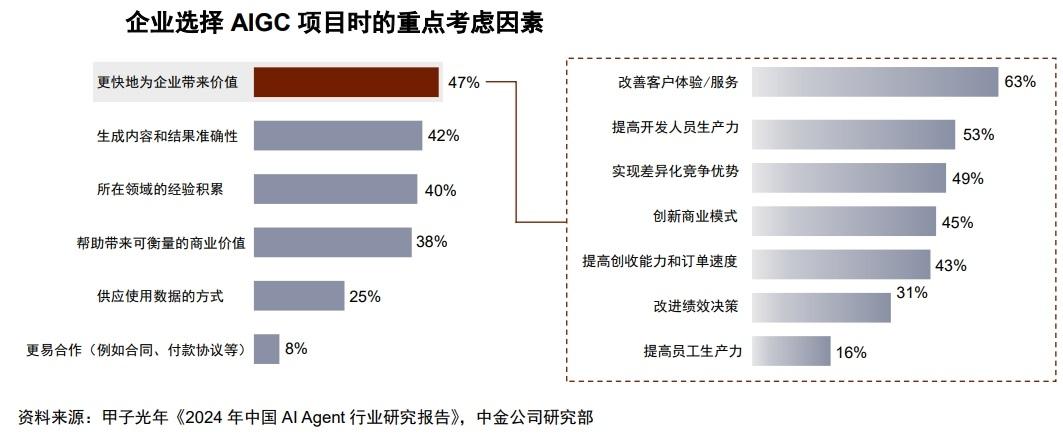
### **（1）AI agent或带动新一轮产业投资，智能生产解决方案提供商竞争力提升**

**AI** **Agent有望凭借专业性的知识沉淀落地B端市场，实现企业降本增效**。相较于个人用户，企业用户通常面临更加复杂和多样化的业务需求，并拥有更明确的业务场景、业务逻辑、更多场景化行业数据及其行业知识案例的积累，能够发挥AI Agent的自主性、感知与理解环境、决策与执行、交互、使用工具等优势。结合AI大模型的迭代、RAG（Retrieval Augmented Generation）的应用及记忆能力的改善，AI对B端场景的赋能有望从早期的Copilot模式逐步演进为具备更多自主智能的Agent形态，初步具备自主交互与环节迭代的高阶能力，进而深度融入并推进企业业务流程，赋能企业降本增效及数智化转型。

**从场景化到跨领域赋能工作流程，AI** **Agent有望重塑B2B产业价值链**。与通用化模型相比，特定领域数据训练的垂直模型在延迟、准确性和成本方面具备更好的表现，且拥有强化自反馈“数据飞轮”效应。B端市场沉淀了海量场景化数据集，有望通过模型训练与微调以深化场景体验与模型的耦合程度，实现AI Agent嵌入企业客户工作并承载业务逻辑。此外，在底层数据打通后，AI Agent可凭借跨API联动重新整合工作流，带动系统整体提质增效。AI Agent有望在B端逐步释放AI的核心价值，实现降本提效、优化人机、人人交互协同模式并引领生产关系的变革。



**企业端强调投入的产出比，赋能企业效率提升是决策核心考量**。企业端在应用背后代表着巨额的资本开支、人力等资源调整成本以及战略转型成本，B端产品的首要目标是实现提升业务效率、降低运营成本、提高创收能力等成本收入端变革，打磨交互体验等则处于次要地位。因此，从长期视角审视，当实现产能的提升幅度超越投入成本且投资回报率ROI>1时，企业通常倾向于采取付费策略。



**在生产端，由于无法有效利用AI的企业面临出清风险，AI** **agent带来的效率提升有望刺激新一轮产业投资**。当前AI浪潮已然驱动云厂商、科技公司等进入新一轮资本开支周期，根据彭博一致预期数据，2024年北美TOP4云厂商（亚马逊、谷歌、微软、Meta）资本开支合计将达2215.0亿美元（YoY+ 50.2%），预计2025/2026年会以19.1%/10.1%的增速至2639.0/2905.3亿美元。

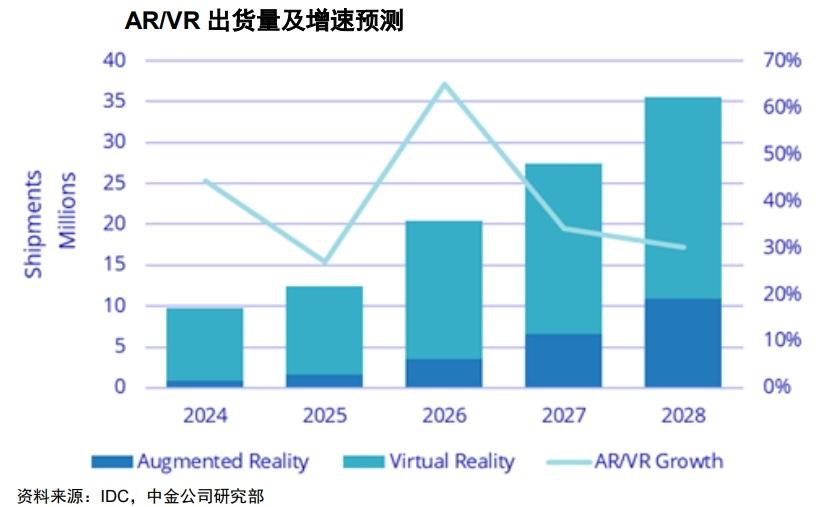
此外，AI算法赋能企业生产端效率提升，需要与具体行业数据及知识结合，通过提取抽象化特征、训练行业核心价值数据形成有价值的AI算法，此外还需要洞悉成本构成、价值创造及产业链运作。积累了行业“专家级知识”的厂商，能够确保AI技术满足差异化的企业生产及业务流程，拥有行业Know-How的知识积累的自动化设备企业/智能生产解决方案提供商的竞争力或将进一步提升。

### **（2）生产工具有望全面升级，看好智能化硬件**

硬件端刺激生产工具的全面升级，包括员工端的AIPC及AR/VR，生产端的传感器、物联网，以及本地云或边缘侧的算力、存储需求等。

AIPC通过集成先进的人工智能技术和优化的硬件架构，为AI Agent提供强大的运行平台和丰富的应用场景。根据联想《AIPC产业（中国）白皮书》，AIPC具有五大特性：内嵌个人大模型，拥有个性化本地知识库；具备CPU、GPU、NPU本地混合AI算力；拥有开放的AI应用生态；可运用自然语言交互；拥有设备级个人数据和隐私安全保护。2024年3月，**苹果**推出搭载M3芯片的MacBook Air，M3芯片中速度更快、效率更高的神经网络引擎，使得MacBook Air持续助力AI技术的应用；**荣耀**发布AIPC—Magic Book Pro 16；**微软**宣布推出首批专为商业用户打造的Surface AIPC——SurfacePro10商用版和Surface Laptop 6商用版；4月，**联想**推出内置联想小天个人智能体“AI元启”系列电脑。据Gartner预测，2025年全球AIPC出货量有望超过1亿台，较2024年同比高速增长165.5%，占整体PC销量的43%。

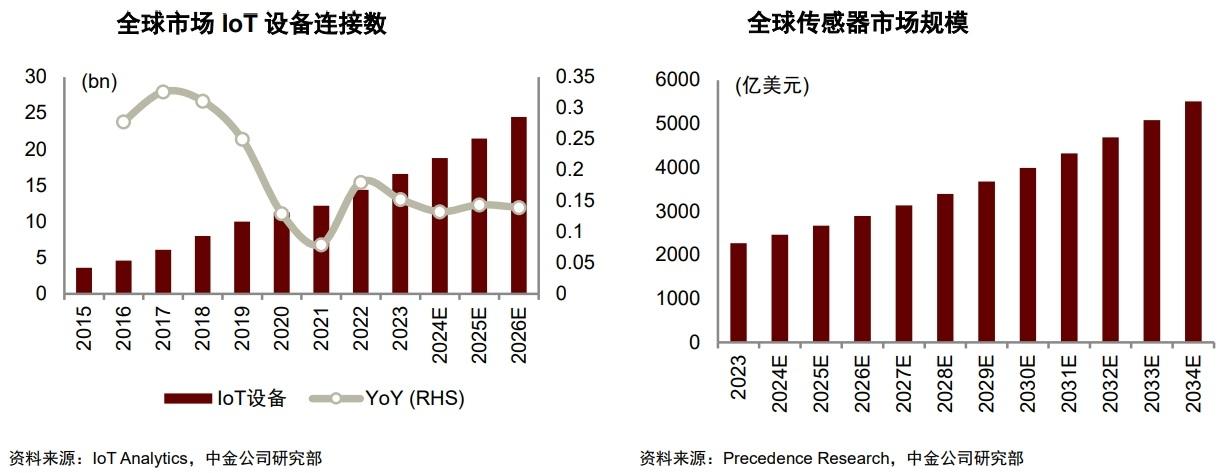
**AR/VR能通过智能交互识别预测用户意图，提升工作效率**。执行方面，用户可以通过语音指令或手势操作来控制虚拟角色、选择菜单项或执行其他任务，AI Agent能够准确识别这些输入，并基于输入的解析快速做出响应，从而提升用户的交互体验和响应速度；识别方面，通过机器学习和数据分析技术，AI Agent能够预测用户的意图和行为，提前加载相关的数据或资源以减少用户的等待时间；场景化模拟方面，AI Agent能够实时模拟和分析AR/VR场景中的物理环境和光照条件，同时赋予虚拟角色和物体智能行为，提升用户的沉浸感和互动性。据IDC数据，2024全球AR/VR出货量达到0.097亿台（YoY+ 44.2%），得益于游戏用户的基础以及企业端等拥抱并采用更多头戴式设备进行培训、协同办公等代理支持，预计到2028年达到0.356亿台。



**企业生产端迎来硬件的智能化替换需求，强化传感器、物联网通信及算力的协同**。AI Agent有望结合物联网、传感器、机器视觉等技术重塑工业生产方式。从应用落地角度看，面对仓库、工厂等作业环境，传感器感知并收集环境中的各种信息，如温度、湿度、光照强度、压力等，并将信息转化为数字信号，通过网络传输给其他设备或系统；而物联网将各个传感器及云边端连接起来，实现数据的收集、传输、存储和分析；AI等技术能够提供决策及解释能力，实现感知信息的最大化利用，从而实现生产工具的数智化转型；此外，由于数据量与实时性的要求，云-边-端多层次的算力网络也成为必要。

**物联网方面**，据IoT Analytics统计数据，2023年全球IoT设备连接数达166亿台，同比增长15%；预测到2024年，该连接数将增至188亿台（同比增长13%），2023-2026的IoT设备连接数实现CAGR 13.9%。此外，据Globe Newswire，2023年全球通信模组市场规模为414.3亿美元，有望保持21.4%的CAGR增长至2030年的1612.9亿美元。

**传感器方面**，据Business Research Insights统计数据，2022年全球摄像机市场规模为81.19亿美元，并预计以CAGR为5.8%的速率增长至2032年的142.46亿美元。此外据Precedence Research统计数据，2023年全球传感器市场规模为2269.1亿美元，预计2024-2034年传感器市场将以CAGR8.4%的速率增长至5510.3亿美元。



## **七、相关公司**

### **1、金山办公：办公场景或是Agent关键入口，公司具有流量优势**

**基于AI** **AGENT（人工智能体）范式驱动WPS** **AI伴写功能**，具体有四大特色：全文补全功能、多种内容生成、不同身份与文风的切换、古诗词内容补全。更新后的WPS AI伴写支持添加包括云文档、网页等参考资料，用户完成标题后只需0.5秒就能理解用户意图并续写内容。

办公是用户日常工作高频覆盖的场景，具有使用时间长、调用功能多的特点，或是Agent关键入口。金山办公是国产办公软件龙头，具有庞大的用户基数，占据入口和流量优势，有望成为办公智能体的主要标杆。

截至2024年9月30日，公司主要产品月活设备数6.18亿，其中WPS Office PC版月活设备数2.77亿，移动版月活设备数3.38亿。



### **2、科大讯飞：讯飞星火4.0Turbo正式发布，AI Agent应用矩阵持续扩充**

**讯飞星火4.0Turbo大模型重磅发布，同步首发星火多语言大模型**。公司正式发布讯飞星火4.0Turbo大模型，核心能力全面超过GPT-4Turbo，且在数学能力、代码能力上超过了GPT-4o。在多语种、多语言能力方面，公司正式首发星火多语言大模型，除中英文外可支持法语、西语、德语等8个语种，并且首次实现全国地级市方言全覆盖，含全国288个地市的202种方言。

**垂类行业应用矩阵持续拓展，Agent应用有望深度赋能教育、医疗等场景**。公司面向教育、医疗、司法、政务服务等多个行业场景发布最新产品应用——教育：发布基于“问题链”的高中数学智能教师系统；医疗：发布讯飞星火医疗大模型2.0，并且发布基于星火医学影像大模型打造的智能医学影像助手，可帮助影像技师快速评估图像质量并纠正检查问题；政务：政务大模型可全面覆盖4500+种标准化事项和60+种材料，并首发搭载政务大模型的“星火智办”一体机。

**“飞星二号”正式启动，公司有望引领全国产算力平台发展**。自2023年10月首个国产万卡算力集群“飞星一号”平台上线以来，星火大模型持续引领全国产算力平台发展，模型训练适配优化平台耗时从90天缩减到15天，新增30多项框架和平台特性。此次，公司宣布与华为、合肥市大数据资产运营有限公司三方联合打造的国产超大规模智算平台“飞星二号”正式启动，未来有望推动新模型新算法的持续适配和智算集群规模的再次跃迁。

**地区政策持续推动AI+教育建设，公司有望深度受益打开成长空间**。北京市教委等4部门发布《北京市教育领域人工智能应用工作方案》，提出至2025年，打造100所人工智能应用场景标杆学校；到2027年，产生数个符合中国教育特点、在技术创新上居于国际前列的人工智能教育行业大模型。未来AI有望在学生学习发展、教学方式变革、学生身心健康、家校协同、教育治理模式创新等领域发挥重要作用。公司作为国内AI大模型龙头厂商，基于在智慧教育领域的生态深耕以及大模型在G、B、C三端的全面布局，有望率先受益于政策推动，实现业务加速成长。

### **3、彩讯股份：信创邮箱领先企业，AI全栈布局拓展价值边界**

**国产邮箱领军企业，AI全栈布局打开想象空间**。公司以手机邮箱起家，逐步发展为产业互联网技术及服务供应商，打造协同办公、智慧渠道、云和大数据三大产品线，实现从算力到应用的AI业务全栈布局，拓展公司业务边界。公司背靠中国移动积极拓展下游，与多家行业头部客户紧密合作，确立了行业内的竞争地位。公司股权结构相对集中，股权激励充分，高度绑定核心骨干人员。公司打造的“双中台”体系筑牢技术壁垒，开发Rich AI超级工厂实现AI业务全栈布局。公司2024年单三季度业绩表现良好。公司2024年前三季度控费效果明显，盈利能力有望进一步提升。

**邮箱为基打造协同办公产品优势，数字员工有望带来新增量**。协同办公平台作为办公管理系统的重要组成部分，信创渗透程度较深。邮箱系统作为企业信息化管理的重要一环，借信创之势逐渐释放国产化替换需求。公司基于信创邮箱打造完备的协同办公产品线，客户拓展成效明显。公司RichMail系统拥有国产自主知识产权，已实现全栈信创兼容。公司智能邮箱深度融合AI能力，同时发力B/C端，未来随其集成更多AI能力并逐渐实现规模商用，有望成为公司协同办公业务的重要驱动力。此外，公司基于对行业大势的判断，在数字员工方面深度布局，有望进一步打开协同办公业务增量市场。

**智慧渠道凝聚亿级用户运营能力，5G+AI打开成长空间**。公司智慧渠道系列产品丰富，携手**中国移动**沉淀了亿级用户运营能力。公司签约华为，构建完善的鸿蒙App开发集成框架，深度融入鸿蒙生态。公司作为5G消息平台行业重要服务提供商和规则制定积极参与者，积极把握5G新通话发展趋势，深耕运营商并积极拓展下游。在数字人领域，公司孵化众多商用数字人产品，打开5G时代通话领域的“数字生活入口”。此外，公司推出国内首个图生音乐AI应用“彩铃图歌乐”，B端赋能运营商业务，C端打开想象空间。公司智慧渠道业务持续增长，5G+AI有望打开成长空间。

**背靠“移动云”发力国资云，布局智算基础设施拓展业务边界**。三大运营商共同发力推动国资云快速发展，移动云稳居第一梯队。公司云和大数据业务深耕“移动云”生态，作为移动云等电信运营商国资云的核心供应商，公司在B端和C端不断扩展云业务范围，承接多项运营商等领域的国资云建设项目。公司与中国移动在个人云盘领域密切合作，相关产品结合AI能力不断迭代升级，有望为公司云和大数据业务带来持续增量。此外，公司基于Rich AI Cloud发力智算基础设施，已可提供万卡集群管理的解决方案，云和大数据产品将逐渐向智算领域延伸。

### **4、焦点科技：跨境B2B电商头部企业，AI Agent打开成长空间**

**外贸市场持续走高，B2B跨境市场空间广阔**。根据海关总署，2024年前11个月，我国货物贸易出口总值达23.04万亿元，同比增长6.7%。其中东盟、“一带一路”沿线国家等贡献增长动能，占我国进出口贸易总额比例持续提升，因此相关人士认为后续美国关税等政策对我国出口及公司经营影响相对有限。向后展望，我国政策支持跨境贸易发展，跨境B2B外贸市场空间广阔。

**跨境B2B平台头部企业，中国制造网维持稳健增长趋势**。焦点科技成立于1996年，成立之初即瞄准跨境B2B电商业务，于1998年注册[http://Made-in-China.com](https://link.zhihu.com/?target=http://Made-in-China.com" \t "https://zhuanlan.zhihu.com/p/_blank)，至今已深耕跨境B2B领域二十余年。目前，公司中国制造网会员数量以及ARPPU均维持健康增长趋势，截至24Q3，公司合同负债与其他非流动负债合计达10.28亿元，同比增长20%，中国制造网收入增速有望逐步触底回升。此外，公司亦推动配套服务发展，形成了以领动建站、孚盟CRM、企业培训为主的“X业务”和以交易、收款、物流等服务为主的国际贸易服务生态链，打造完整的跨境贸易服务体系。

**AI布局持续深化，看好AI麦可以及买家侧AI工具变现空间**。公司AI应用持续迭代，推出FocusAIM平台，内含四款跨境贸易AI产品：AI麦可、MoliAI、SourcingAI以及Mentrac。目前，AI麦可已迭代至4.0版本，较3.0版本进一步强化了其作为AI Agent的角色，能够自动规划和执行整个外贸工作流，并具备自主学习知识库的能力，测算其中长期市场空间有望达数十亿元。此外，公司推出的SourcingAI以及Mentrac等站外AI工具有望逐步打开海外买家侧以及供应商变现空间，进一步提升公司商业化潜力。

### **5、泛微网络：AI 应用逐步繁荣，协同办公龙头有望受益**

**DeepSeek等国产模型性能显著突破，AI应用有望逐步迎来繁荣**。DeepSeek发布大模型Deep Seek-v3和R1，前者在大模型主流榜单的开源模型中位列榜首，并与世界上最先进的闭源模型不分伯仲；后者性能对标OpenAI的o1正式版，在数学、代码和自然语言推理等任务上表现卓越。以上两个模型均通过算法和架构的创新，大幅降低了训练成本和推理成本，为AI技术的普及与创新作出了卓越的贡献。此外，阿里通义千问的Qwen2.5-Max，月之暗面的k1.5多模态思考模型也实现了较好的性能，国内大模型公司在性能和推理成本上竞争力日益突出，为国内AI应用的落地打下了良好基础，2025年有望看到AI应用的繁荣。

**协同办公是“AI+工作流”和“AI+知识库”应用的典型场景之一**。泛微网络是国内领先的协同办公软件厂商，目前覆盖87个行业、超过8万多家客户。协同办公软件覆盖企业与组织全体员工，并连接组织内的人、财、物、信息等资源， 通过与AI结合，可以降低用户对系统的学习成本，在办公过程中明显简化大量人工操作，大幅优化和改善使用体验和使用效率，因此，协同办公是“AI+知识库”和 “AI+工作流”落地的典型场景。

**泛微网络数智大脑Xiaoe.AI 已正式推出并在部分客户与场景落地**。2024年12月5日，公司正式发布泛微·数智大脑Xiaoe.AI，该平台采用了大模型+小模型+智能体的技术架构，利用大模型的通用推理能力、小模型的转项任务能力，加上数据后形成了各类智能体应用服务，实现包括AI问答、AI搜索、AI 采集、AI审批、AI摘要等在内的各项功能，帮助“市场、销售、合同、采购、项目、客服、人事、资产、财务、档案”等各种业务场景的智能化改造。据介绍，Xiaoe.AI 已经为西安建筑科技大学、上海建工、雅戈尔、艾比森光电、爱美客、东芯半导体等组织提供了智能化服务，提升组织运营效率。

### **6、同花顺：广告收入弹性高于市 场，AI 赋能产品竞争力提升**

**公司业绩显著改善，广告收入高增为主要驱动**。2024年，同花顺实现营业收入、归母净利润分别为 41.9、18.2亿元，同比分别+17.5%、+30.0%，加权ROE为24.8%，同比+4.6pct，Q4单季归母净利润为 11.7亿，同比+85.5%；分业务来看，2024年公司实现增值电信、广告、软件、基金代销收入分别为 16.2、20.2、3.6、1.9亿元，同比分别+5.7%、+49.0%、-18.6%、-20.5%，贡献收入比重分别为38.6%、48.4%、8.5%、4.5%，广告收入高增，为业绩的主要驱动，软件及基金代销有所承压。

**收入端，广告收入弹性高于市场，产品竞争力提升**。截至2024年末，公司APP月活数为3627万人，同比+18.9%，测算市场份额为21.2%，同比-0.29pct，公司平台流量随市场回升，龙头优势稳固；2024年市场日均成交额同比+21.2%至1.06万亿元，公司广告收入增速显著高于市场，预计公司作为龙头流量平台，议价能力有所提升，且公司产品与大模型技术深度融合，产品竞争力得到提高；公司软件销售及维护收入同比-18.6%至3.6亿元，主要由于券商信创升级逐渐完成，相应的业务需求减少所致。

**成本端，放缓研发投入，盈利能力改善**。2024年，公司管理、销售、研发费用分别为2.2、5.9、11.9 亿元，同比分别+8.3%、+19.1%、+1.2%，对应费用率为5.3%、14.1%、28.5%，同比分别-0.4、+0.2、-4.6pct，公司成本控制良好，放缓研发费用投入推动盈利能力改善，2024年公司净利率43.5%，同比+4.2pct。

**公司智能产品全面升级，推动前沿大模型技术应用**。问财投资助手升级为原生AI对话产品，融合多模态大模型技术，实现更自然、精准的交互；同创智能体平台集成大模型、低代码和实时金融数据，助力上百家金融机构共建超5000个智能体，日均调用量超300万次；金融数据终端构建智能化产品矩阵，AIFinD飞研优化写作与分析能力，“产业智绘”平台提供全景产业链研究；HiPilot智能编码助手增强跨文件、跨项目的代码理解与生成能力，新增慢思考问答、代码可视化、测试套件生成等功能，已为数千名开发人员生成超百万行代码。

### **7、致远互联：短期业绩承压，长期看好AI Agent和行业信创机遇**

**短期业绩承压，主要系三大因素变化影响**。根据《北京致远互联软件股份有限公司2024年年度业绩预告》披露的相关信息，公司2024年度预计亏损将进一步扩大，主要系三个方面的变化所致。一是营业收入不达预期，从短期看，受客户结构、客户需求及项目周期等因素变化的影响，公司的营业收入当期未达预期，预计实现营业收入同比下降。二是期间费用上升，主要是报告期内人员优化导致离职补偿当期一次性支出，以及公司持续引入优秀人才和加强人员能力提升所致。三是资产减值损失的计提，基于谨慎性原则，结合市场情况判断和会计准则要求，按照公司的会计政策拟计提的信用减值损失较上年同期增加。此外，非经营性损益也对业绩造成了影响，2024年公司符合损益确认条件的政府补助低于上年同期。

**长期看好AI** **Agent和行业信创的发展机遇**。尽管公司短期内业绩承压，但从长期看整体趋势正在转好，并对AI和行业信创的未来发展机遇保持乐观。经营方面，业绩预告披露2024年公司在手订单同比增长10%-15%，合同负债同比增长10%-15%，项目回款率保持稳定增长，经营活动产生的现金流量净额较去年同期有所改善；2024年全年合同签约金额同比保持稳定增长，企业端客户合同签约金额同比增长，百万级以上合同签约金额增长约15%。研发方面，公司积极投入研发新一代技术平台及产品，并加力AIGC在组织级办公软件应用的研发投入；截止2024年上半年，公司研发支出总额占营业收入的比例已接近30%，同比提高了近5个百分点。

在AI方面，致远互联坚持以AI Agent推动协同运营管理的高效化、智能化和数字化。据业绩预告披露，2024年公司发布了五大类AI产品和服务，包括知识服务类、内容创作类、数据分析类、流程自动化类和对话服务类，以及AI原生应用。报告期内公司AI Agent关联合同（含AI Agent功能的合同）金额约2500万元，AI相关产品已逐步实现订单转化。

在行业信创方面，公司聚焦突破央国企等头部客户，制定了“新质计划”，该部分市场占有率提升较快，2024年行业信创合同签约金额增长超70%，签约了**中国中车**、**中国中铁**、**中国建筑**、**中国通号**、**中国南水北调**、**中国林业**、**中国出版**等重点客户。

## **八、发展展望**

### **1、2025年有望成为AI Agent商业化应用元年**

**AI** **Agent是通往AGI的必由之路**。从NLP到AGI的发展分为五个级别，语料库、互联网、感知、具身和社会属性。目前的大语言模型（LLM）已来到第二层级，具有互联网规模的文本输入和输出。AI Agent在LLM的基础上再具备感知与行动的能力时，将进入到第三和第四级别。再进一步，当多个Agent通过互动合作解决更复杂的任务，或者反映出现实世界的社会行为，则有潜力来到第五层级。



**Agent将显著提升大模型使用工具能力**。根据智谱定义的大模型发展的五个阶段：L1语言能力、L2逻辑能力（多模态能力）、L3使用工具的能力、L4自我学习能力、L5探究科学规律。目前大模型已经初步具备了人类与现实物理世界互动的部分能力，而Agent将显著提升L3使用工具能力，同时开启对L4自我学习能力的探索。



**Agent为AI未来重点技术趋势，B端为重要落地场景**。根据硅谷著名创业加速器YCombinator的数据，2023夏季和2024冬季2季创业营的AI项目中，Agent已成为最主要的技术路线之一，占比高达80%，其余为多模态、基础设施、RAG等。其中，B端应用项目数量远高于C端应用，包括软件开发、金融、电商、CRM、医疗等。



**AI** **Agent场景特性：B端强调专业性，C端强调自由度**。在B端应用中，AI Agent强调专业性，被广泛应用于金融、医疗、法律、财务、生产、物流、资管、人力等专业性极强的领域，往往需要专岗专职人员进行多年学习完成。在C端应用中，AI Agent强调自由度，应用于游戏、娱乐、影视、营销、广告、教育等需要更多灵感及创意的领域。

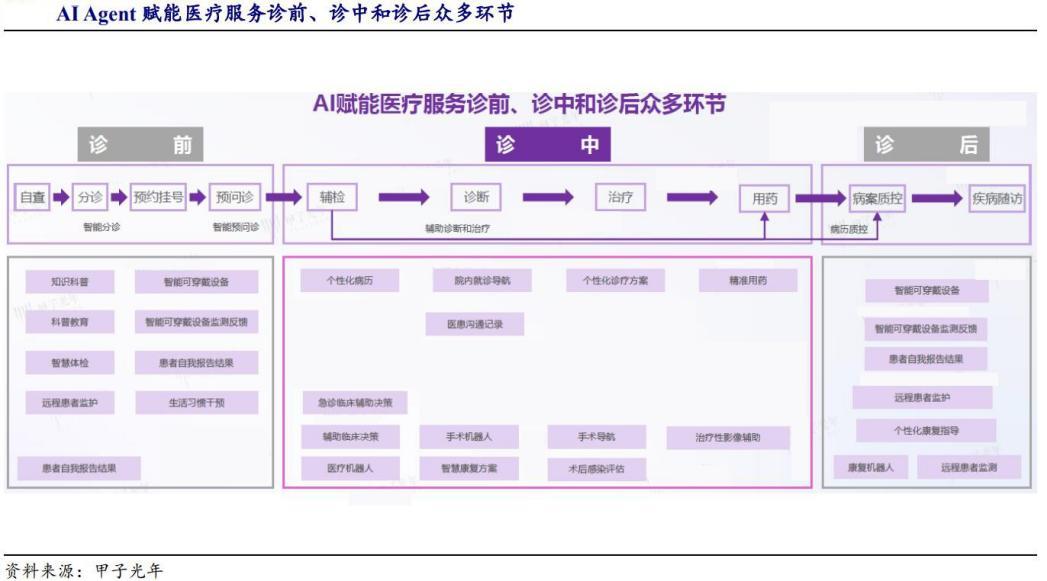


**生成式AI在金融领域应用广泛**。根据Gartner调查数据，66%的银行领导人预计生成式AI的潜在收益大于潜在风险，61%的银行高管表示所在公司正在或计划未来增加对人工智能的投资。例如为用户提供财务建议的AI理财教练，在增收方面至少可以提升30%，提高效率方面能够显著减少公司为客户创建新内容所花费的时间。

**AI智能体在金融领域降本增效前景广阔**。智能体的优势在于能够自动执行日常任务，从而释放人们的创造力，提高工作效率和生产力。Gartner列出的生成式AI在银行业20个最有前景的应用案例中，多个场景涉及AI Agent，如综合信贷数据、AI-Copilot前线应用、个性化营销内容等，利用生成式AI技术构建的数字员工，能够显著增强银行业务处理能力、优化客户体验并降低风险。



**AI** **Agent赋能医疗服务诊前、诊中和诊后众多环节**。政策、技术共振驱动下，AI赋能医疗服务诸多场景。2024年11月14日，国家卫健委、国家中医药局、国家疾控局联合发布《卫生健康行业人工智能应用场景参考指引》，提出人工智能+医疗服务管理、人工智能+基层公卫服务、人工智能+健康产业发展、人工智能+医学教学科研四个类别，共计84个细分应用场景，积极推进卫生健康行业“人工智能+”应用创新发展。



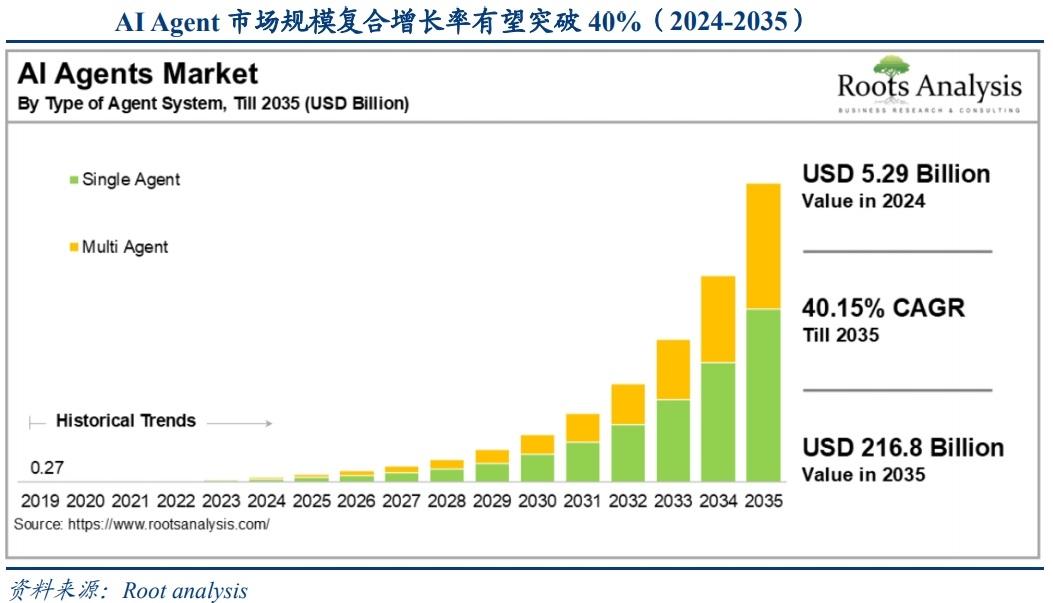
**AI** **Agent助力医疗服务提质增效**。AI Agent在医疗领域的应用前景广泛，涵盖从日常辅助、影像分析、病历管理到手术支持、慢病管理等多场景，为医护人员减负、提升诊断精准度并优化患者体验提供智能支持。根据美国国家经济研究局2023年报告数据，更多地采用人工智能可以节省5%-10%美国医疗保健支出。

**法律行业有望成为AI** **Agent重要落地应用**。法律行业的数据以文本为主，并且多为绝对事实和真实案件，较为适合大模型学习。此外，法律工作的强知识性和逻辑性也让大模型能够在更多法律工作流环节中提供价值。根据ALM就美国前100律所使用AI情况的问卷调查，其中有41家律所承认正在使用AI，包括法律材料起草、文档汇总、法律研究等场景。Jackson Lewis等一些律所，还会定期为律师举办人工智能培训课程。

**Harvey** **AI用户使用率快速增长，彰显法律领域对AI** **Agent产品的旺盛需求**。以OpenAI投资的AI+法律初创公司Harvey AI为例，2023年8月-2024年8月，Harvey AI用户使用率从33%大幅增长至69%，用户留存率在一年后保持在70%左右，彰显法律领域对AI Agent产品的旺盛需求和用户对产品的较高满意度。

### **2、AI Agent发展从0到1，市场规模不断增加**

**从0到1，Agent市场规模快速增长**。随着AI技术的不断进步，AI Agent已经不再局限于简单的任务执行，当前逐步实现从0到1的跨越式创新。众多企业和研究机构投入研发，推出创新产品。行业结构初步形成，相关产业链和生态系统逐渐完善。根据Rootanalysis，预计全球AI Agent市场规模将从2024年的52.9亿美元增长到2035年的2168亿美元，2024-2035年预测期内的复合年增长率为40.15%。



### **3、AI Agent将成为端侧应用重要支柱**

**端侧模型有一个不可能三角：性能、参数量和内存及功耗占用**。性能优异需要大参数量；而参数量大就意味着内存占用大，功耗也会大；功耗过大又可能会影响性能。AGI是一个长期演进的过程，Agent有望成为当务之急的“解决问题”，这一特性在端侧尤为重要。通过主动工作流的配置辅以性能不错的模型解决问题，而主动工作流的配置离不开Agent技术的应用。

**根据腾讯研究院公众号的内容指出，从电池容量的角度来看，通过工作流优化任务的实现是刚需**。目前，由于手机和PC的保有量占据绝对优势，它们理所当然地成为了端侧AI的最佳落地方向，但在落地过程中，由于面临的芯片和电池的挑战，为了实现大模型的终端落地，需要进行大量的适配工作。

微软作为端侧模型的有力竞争者Phi-3/3.5的开发者，除了模型本身，还提供了一套名为Agents的工具。通过Microsoft Copilot Studio的升级，Copilot+PC不仅可以调用Windows附带的40多个端侧AI模型提供支持，还可以构建成百上千的自动化业务流程，在客户需求下独立工作，从而实现长期运行的业务流程自动化。未来，移动端和PC端体验到的端侧AI，大部分将是通过适配器和分类器挑选的微调小模型，以及针对特殊需求开发的自定义Agents，而无需调用全量的模型参数，是更具性价比的方案。

**另一方面，从生态搭建的维度，需要Agent调用多方资源以实现繁荣**。无论是端侧还是云端AI，大模型都只是底层计算；要实现用户价值，还需要一个繁荣的应用生态和强大的工具集来提供支持。不论是现有的APP形态，还是未来可能实现的“去皮化”的API形态，除了底层计算，还需要通过Agent来实现价值的连接。在6月的Apple Intelligence发布会上，苹果表示，Siri的全新形态将改变交互规则，大量AI新功能将很快上线；此外，屏幕读取以及App内与App之间的操作等能力预计也将到位，这将使AI真正串联起苹果生态下的诸多应用。苹果提前承诺的这项能力，源自其在4月份发布的一项名为“Ferret-UI”的新技术。Ferret-UI能够“看懂”手机屏幕，建立对UI元素的基本理解，奠定了执行复杂任务的基础，并通过分层次的任务设计，最终实现对用户指令的理解和响应。这本质上是一种通过视觉方式来构建主动工作流Agent的思路。底层模型本身并不能直接创造价值，苹果需要维持其最强的盈利因素：生态位。

**长期来看，专业化端侧与全能云端协同或是端侧AI的最优解**。云端模型比端侧模型先进一个数量级。虽然许多小模型在特定能力上已经具备了媲美十倍甚至百倍参数大模型的实力，但事实上，当前基础模型的综合能力依然基本遵循Scaling Law法则。千亿、万亿参数的大模型以及实验版本模型，作为探索AGI的最前沿模型，其整体智力水平无疑会持续领先。云端大模型始终比端侧大模型先进一个以上的数量级。例如，去年8月份谷歌发布的轻量级小模型Gemma22B，是从6月份发布的Gemma27B和9BGemma2模型中蒸馏而来的；微软开发的AI小语言模型（SLMs）Phi-3系列有多个版本，包括mini（3.8B）、small（7B）和medium（14B）。根据微软公布的不同表现水平，在同一时间段内，参数规模仍是决定大模型综合能力的关键因素。苹果的Apple Intelligence通过一个对标GPT-4的云端模型Apple Server来处理复杂任务，这不仅是因为终端设备受限于芯片、电池和发热等因素的选择，更是为端侧提供“无所不知、无所不能”的云端支持保留了重要的接口。

## **九、参考研报**

1. 招商证券-AI应用行业三类投资机会梳理：垂类应用率先落地，AI Agent剑指未来
2. 浙商证券-计算机行业：AI Agent元年，企业服务有望最先落地
3. 国泰君安-计算机行业：AI Agent突破ChatBot形式渗透全产业
4. 中金公司-科技硬件行业AI浪潮之巅系列：AI Agent，大模型落地下一站
5. 德邦证券-计算机行业专题：AI Agent，海外创新企业的再次自我革命
6. 开源证券-计算机行业深度报告：AI Agent拐点已至，2B+2C星辰大海
7. 银河证券-计算机行业AI Agent崛起：产业链重塑与投资新视角
8. 中金公司-科技硬件行业AI浪潮之巅系列：AI Agent，大模型落地下一站
9. 海通证券-信息服务行业专题报告：AI Agent（智能体），从技术概念到场景落地
10. 浙商证券-科大讯飞-002230-点评报告：讯飞星火4.0Turbo正式发布，AI Agent应用矩阵持续扩充
11. 平安证券-彩讯股份-300634-信创邮箱领先企业，AI全栈布局拓展价值边界