

目 录

1. 原理解析：思维链铸就智能体，多体交互拓展应用	6
1.1 Agent 模式架构解析	6
1.2 单智能体 vs 多智能体	7
1.3 Prompt 提示词的艺术	8
2. 应用场景：赋能两类实体领域，成本与效益的博弈	8
2.1 自动化类：微软智能体 AutoGen	9
2.2 情感需求类：陪伴型智能体	11
2.3 效能讨论：成本与价值间的博弈	12
3. 未来展望：人工智能发展迅猛，智能体商业化落地	13
3.1 海内外政策推动人工智能发展加速	13
3.2 算力规模高增，企业对 AIGC 的投入意愿强	15
3.3 多模态智能体有望实现大规模商业化	17
3.4 全球智能体市场增速加快	18
4. 潜在标的：Agent 企业千帆竞发，垂类应用引人期待	19
4.1 LeewayHertz：提供多领域解决方案	19
4.2 Markovate：提供多模态智能体解决方案	21
4.3 Replika：情感类智能体热度不减	22
4.4 Character AI：自研大模型，打造智能体社区	23
4.5 星火智能体：布局智能体平台，发布垂类智能体	24
4.6 SkyAgents：自研模型降低成本	25
4.7 商汤：联合清华、上海人工智能实验室打造 GITM	26
4.8 字节跳动：豆包聊天助手	27
4.9 腾讯：AppAgent	28
5. 投资建议	31
6. 风险提示	32

图目录

图 1	人类与 AI 的交互模式	6
图 2	AI Agent 智能体工作原理	7
图 3	单智能体强化学习原理图	7
图 4	多智能体强化学习原理图	7
图 5	CrewAI 整体框架	8
图 6	AutoGen 运作模式图解	9
图 7	Agent 交流模式展示	10
图 8	AutoGen 实验数据对比	11
图 9	林间聊愈室产品介绍	12
图 10	OpenAI 不同模型 tokens 消耗价目表	12
图 11	单次运行 Agent 对 Tokens 的消耗	13
图 12	单次运行 Agent 的实际成本	13
图 13	中国人工智能市场规模预测	14
图 14	全球人工智能市场规模预测	14
图 15	美国增强型人工智能研发投资 2015-2030	15
图 16	美国及其 AI 盟友	15
图 17	中国智能算力规模及预测	16
图 18	企业 AI Agent 使用情况	16
图 19	专注于横向应用的 AI 代理初创企业率先获得融资	17
图 20	多模态大模型发展	18
图 21	全球自主人工智能和智能体市场规模	18
图 22	中国前五大人工智能市场应用	19
图 23	LeewayHertz 公司发展历程与合作公司	19
图 24	LeewayHertz 的智能体产品架构	20
图 25	LeewayHertz 为私募股权提供的 AI Agent 解决方案	20
图 26	Character.ai 智能体架构	21
图 27	ShopSpot 用户界面展示	21
图 28	ShopSpot 库存管理系统	21
图 29	ShopSpot 多模态智能体图片识别功能展示	22
图 30	Replika 付费模式	22

图 31	截止 2024 年 8 月 Replika 各年龄段用户占比	23
图 32	截止 2024 年 8 月 Replika 用户全球分布	23
图 33	Character.AI 用户界面	23
图 34	Character.ai 2024 年 7 月每日访问量	24
图 35	Character. AI 各年龄段用户占比	24
图 36	Character. AI 用户全球分布	24
图 37	科大讯飞星火智能体介绍	25
图 38	部分星火 AI 智能体展示	25
图 39	天工大模型能力展示	26
图 40	天工 Agent 计费模式	26
图 41	RL 架构对比 GITM 创新架构	27
图 42	解锁科技树的成功率（蓝色是 GITM）	27
图 43	各模型完成任务成功率对比	27
图 44	豆包智能体的对话界面	28
图 45	腾讯 AppAgent 展示软件操作	29
图 46	AppAgent 的 App 探索学习阶段	30
图 47	AppAgent 的实际部署阶段	31

表目录

表 1	AI Agent 类型盘点以及具体应用	9
表 2	国内政策推动人工智能发展	13
表 3	海外政策推动人工智能发展	15

1. 原理解析：思维链铸就智能体，多体交互拓展应用

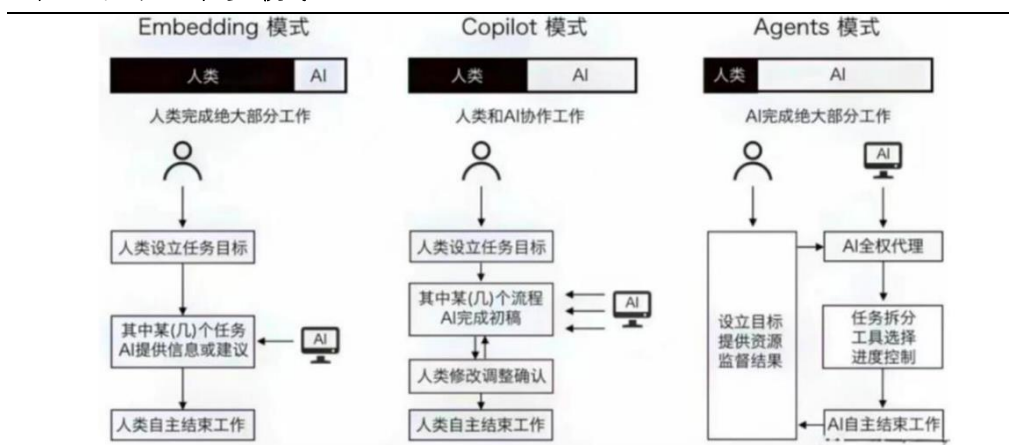
早在上世纪 50 年代，阿兰图灵把“高度智能有机体”扩展到了人工智能。如今随着大模型的快速发展，这个概念又被重新拾起。大模型成为了智能体目前最完美的载体，有望完成从概念到实际应用的蜕变。

用户在 Agent（智能体）模式中给 AI 设置目标和身份，并提供 Prompt（提示词）。AI 自主拆分任务、使用工具、完成工作，用户仅负责设立目标、提供工具资源和监督结果。OpenAI 定义的智能体具有长期和短期记忆、自主规划、工具使用和自动执行任务的能力，能提高工作效率和用户体验。另外，智能体也分为单智能体和多智能体。单智能体通过试错学习在单一环境中行动，追求最大奖励，多用于简易任务。多智能体在博弈环境中行动，追求长期累积奖励，多用于复杂测试。

1.1 Agent 模式架构解析

Agent 有效减少人类工作总量，人与 AI 协作才是最终形态。人类与 AI 交互可大致分为三种模式。Embedding 模式中大模型可以填补一些信息缺失，完成少量任务，例如总结信息等等。用户最终会整合挑选 AI 提供的信息，并自主完成任务。Copilot 模式则更加智能化，AI 可根据用户设定的流程去执行任务。例如让 AI 根据写一段稿件或者根据需求编程，但其对 Prompt 的要求也更高。在 AI 完成流程后，用户需要对内容结果进行调整并自主结束工作。Agent 智能体模式的 AI 参与度更高，但也不是完全由 AI 代理。用户需要给 AI 设计一个目标和身份，以及需要使用的工具。配上更为复杂的 Prompt，AI 能自主进行任务拆分，使用工具并结束任务。用户只负责设立目标、提供资源、监督结果。

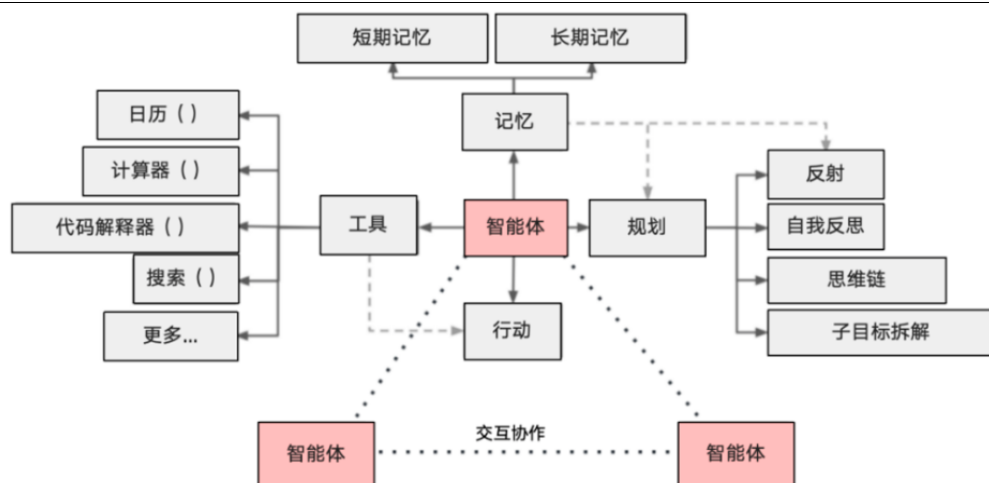
图1 人类与 AI 的交互模式



资料来源：腾讯研究院，海通证券研究所

以 LLM 为核心，四模块铸造 AI Agent。从 OpenAI 的定义来看，智能体以大语言模型为核心，其拥有长期和短期记忆、自主规划能力、能自动化执行复杂任务、能够使用工具等四个特点。1) 记忆模块：智能体像人类一样，能留存学到的知识以及交互习惯等，这样的机制能让智能体在处理重复工作时调用以前的经验，从而避免用户进行大量重复交互。短期记忆适用于所有上下文的学习，类似平常我们与 ChatGPT 沟通的模式；长期记忆则保留知识和交互回忆，例如智能体在特定行业积累的大量数据和经验，则能提供更专业、更具深度和个性化的回答，提升用户体验。2) 规划模块：将复杂任务分解成子目标并逐一解决，完成任务后进行反思总结。例如反思自己大量输出重复内容或在单一子目标耗时过长等问题，将经验存入长期记忆以规避类似错误。3) 工具模块：智能体可利用工具来弥补自身短板，通过调用外部 API 来实现功能拓展。例如调用连接互联网的 API 去搜索实时信息。4) 行动模块：智能体会形成完整的计划流程。例如先读取以前工作的经验和记忆，之后规划子目标并使用相应工具去处理问题，最后输出给用户并完成反思。

图2 AI Agent 智能体工作原理



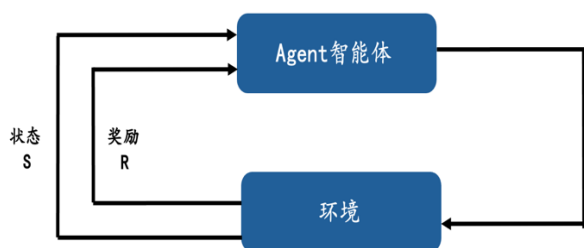
资料来源：腾讯研究院，海通证券研究所

1.2 单智能体 vs 多智能体

单智能体与多智能体各具优势，适配于不同垂直领域。单智能体的强化学习原理是基于马尔可夫决策来完成的，简单来说可以分为状态集 S 、行动集 A 、奖励 R ，下一时刻的状态和奖励只与上一时刻的行动有关，与更早之前的状态无关。其模型原理就是让智能体用试错的方式来学习，若某个策略能得到奖赏，则智能体产生该行为的策略就会加强。其目的就是在单一环境中行动，尽可能得到最大的奖励。应用领域目前也较为广泛，例如赛车游戏中连续动作的训练：控制方向盘、油门、刹车等动作，可由 DDPG、A3C、PPO 算法来决策。一些离散动作的训练例如围棋智能体 AlphaGo，可通过 Q-Learning 等算法决策。

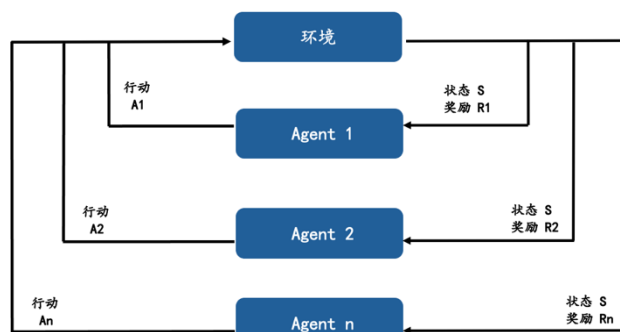
多智能体的决策不仅与自身行动相关，还与系统内其他智能体的行动所关联。一个多智能体系统中会有两个以上的智能体，他们一般存在着合作或竞争关系。这样模型称为马尔科夫博弈，其状态转换符合马尔可夫决策，关系符合博弈。在多智能体模型中，每个智能体的目标是找到最优策略来使它在任意状态下获得最大的长期累积奖励。由于其模型更为复杂，干扰因素较多等原因，目前多智能体模型商业化产品较少。

图3 单智能体强化学习原理图



资料来源：CSDN，海通证券研究所

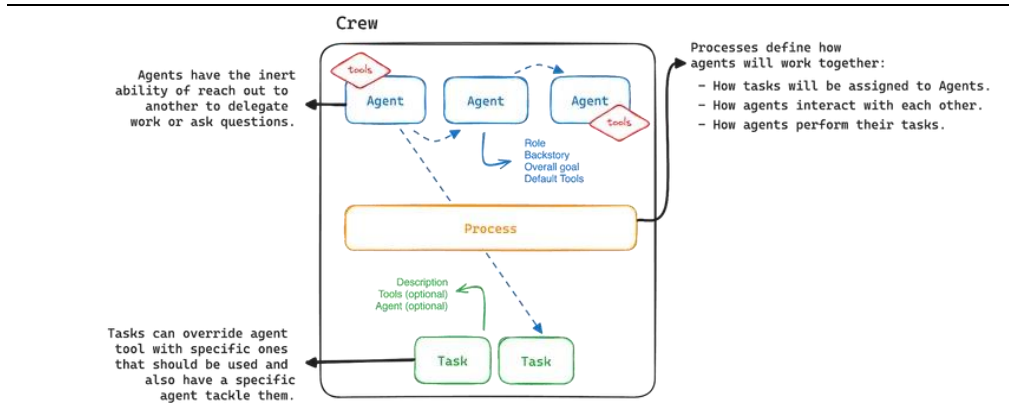
图4 多智能体强化学习原理图



资料来源：CSDN，海通证券研究所

CrewAI 是世界领先的多智能体框架之一，在多智能体领域用于协调角色扮演型自主 AI 智能体。通过促进协作智能，CrewAI 使智能体能够无缝协作并处理复杂任务。在编写程序时，用户需要赋予每一位 Agent 角色、任务、以及背景故事。

图5 CrewAI 整体框架



资料来源：CrewAI, CSDN, 海通证券研究所

1.3 Prompt 提示词的艺术

Prompt 是一种基于自然语言处理的交互方式，它通过机器对自然语言的解析，实现用户与机器之间的沟通。可以理解为是一段给定的文本或语句，用于启动和引导机器学习模型生成类型、主题或格式的输出。在自然语言处理领域中，Prompt 通常由一个问题或任务描述组成，例如“帮我写一篇有关人工智能的文章”，“帮我编写一段可以整理数据的程序”等等。

一个优秀的 Prompt 应该包含 1) 明确目标：能够定义任务，以及想要达到的效果；2) 具体指导：给予模型明确的指导和约束，举一些例子或明确回复边界；3) 简洁明了：使用简练、清晰的语言表达 Prompt；4) 迭代优化：一次 Prompt 输出可能达不到想要的结果，用户需要根据智能体的回答来改变提示词，持续调整优化。

2.应用场景：赋能两类实体领域，成本与效益的博弈

AI Agent 目前的应用大多都在概念层面，但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。

表 1 AI Agent 类型盘点以及具体应用

AI Agent 类型	具体描述	应用领域
基于效能的 Agent	基于效用的 Agent 在具有多种潜在结果的复杂决策环境中表现出色——比如平衡不同风险以做出投资决策，或者权衡价格与成本。	金融交易、价格追踪系统、智能电网控制
基于目标的 Agent	基于目标的 Agent 会评估和规划行动，以实现它们的目标。与其他类型的智能 Agent 不同的是，它们能够将前瞻性 and 战略规划结合起来，朝着特定的结果前进。	扫地机器人、项目管理软件、电子游戏
基于模型的 Agent	适用于并非总是可见或可预测的信息时。这个模型会根据环境中传入的数据不断更新，因此人工智能 Agent 可以对环境中未见的一部分进行推断，并预测未来的状况。	自动驾驶技术、自动灌溉、智能家居
分层 Agent	分层 Agent 类似于复杂的组织结构，具有不同的决策层次。系统中的不同 Agent 拥有不同的专业领域，使它们在处理复杂的多步骤任务时更加高效。	空中交通管制系统、自主仓储机器人
机器人 Agent	增加了性能元素，这些人工智能 Agent 的实体通常配备有摄像头或触摸传感器等传感器。这种 Agent 在危险或重复性很高的任务中尤其有用——让人工智能 Agent 来完成这些任务可能更有效率，也更符合成本效益。	生产线机器人、手术机器人、农业机器人、服务机器人
多 Agent 系统	每个 Agent 都是半自主运行的，但其设计目的是与其他 Agent 互动，形成一个动态的生态系统，从个体行为中产生集体行为。对于这种 Agent 程序来说，协作是关键。	交通管理系统、用于能源管理的智能电网、供应链与物流

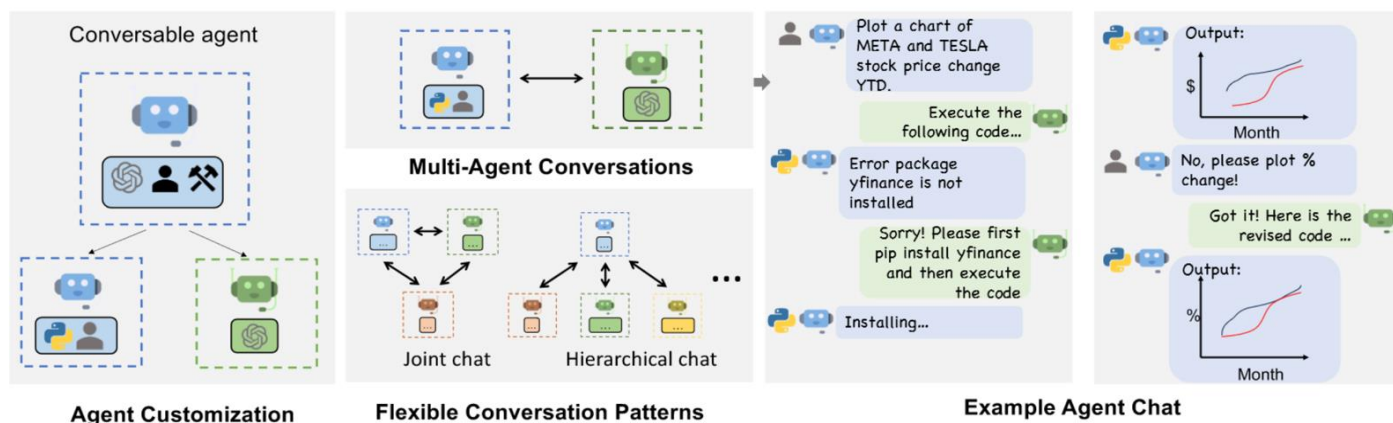
资料来源：botpress，海通证券研究所

2.1 自动化类：微软智能体 AutoGen

微软研发的 AutoGen 是一种多智能体沟通的模式，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。

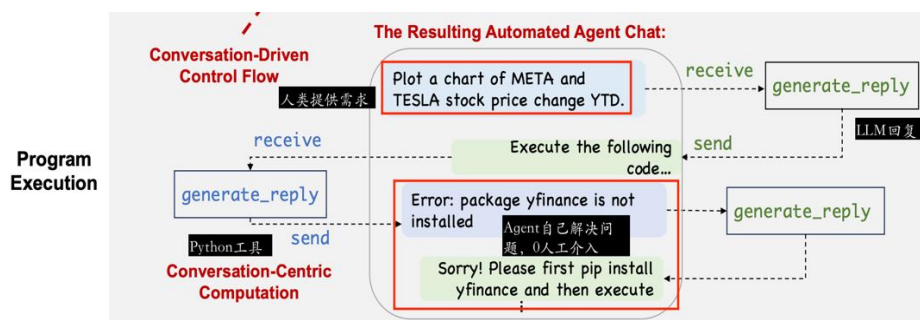
图6 AutoGen 运作模式图解



资料来源：Enabling Next-Gen LLM Applications via Multi-Agent Conversation, CSDN, 海通证券研究所

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。

图7 Agent 交流模式展示

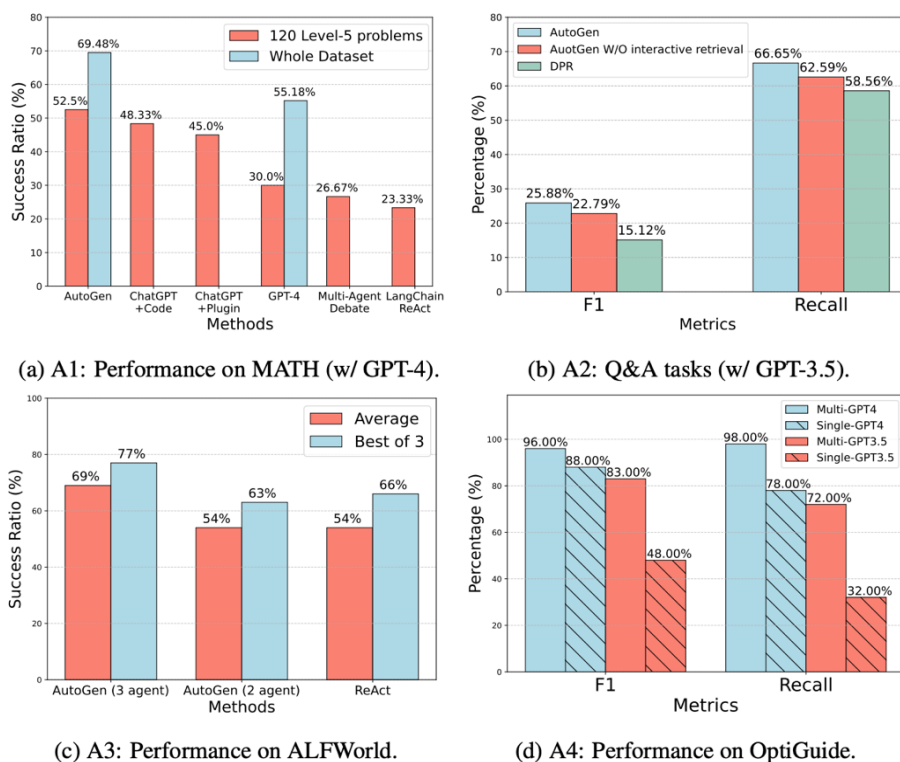


资料来源：Enabling Next-Gen LLM Applications via Multi-Agent Conversation, CSDN, 海通证券研究所

AutoGen 在四项实战应用的表现测试中均胜过单一语言模型，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。

- 1) 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。
- 2) 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。
- 3) ALFWorld (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。
- 4) OptiGuide:** 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 的时，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%。

图8 AutoGen 实验数据对比



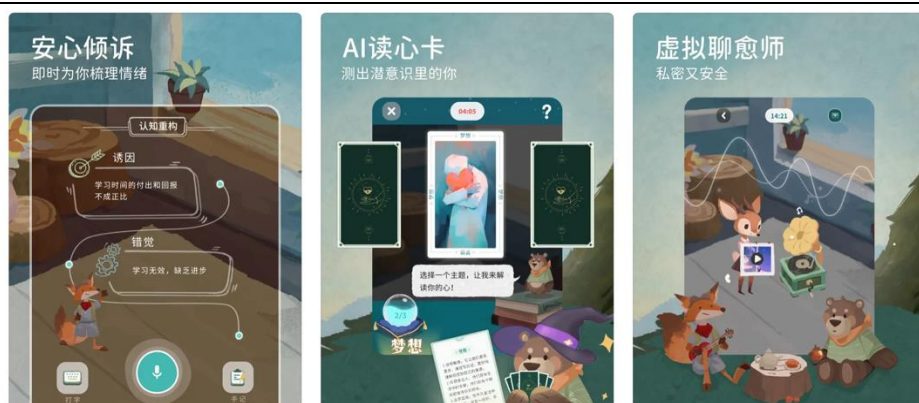
资料来源：Enabling Next-Gen LLM Applications via Multi-Agent Conversation，海通证券研究所

2.2 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，Denny Yin 提出该行业规模预计到 2024 年增长到 1000 亿美元，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，App Store 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。

图9 林间聊愈室产品介绍



资料来源：App Store，海通证券研究所

2.3 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 GPT3.5 turbo 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。

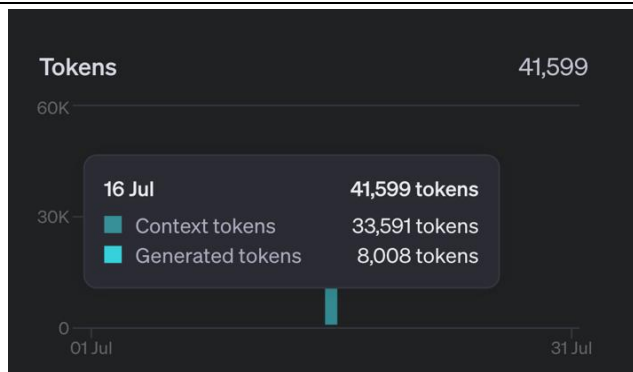
图10 OpenAI 不同模型 tokens 消耗价目表

Model	Pricing
gpt-4o	US\$0.0050 / 1K input tokens
	US\$0.0150 / 1K output tokens
Model	Pricing
gpt-3.5-turbo-0125	US\$0.00050 / 1K input tokens
	US\$0.00150 / 1K output tokens

资料来源：OpenAI，海通证券研究所

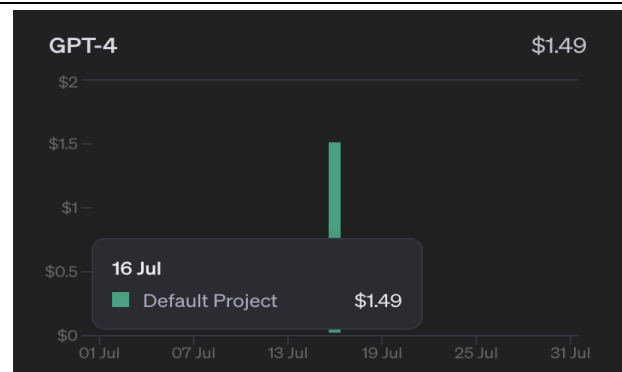
在实际测试中，使用搭载 GPT4 的智能体模型查找新闻并进行总结梳理。这部分的单次消耗是 42000 个 tokens，成本是 1.5 美元。在实际应用中，假设一天需要统计的股票新闻大约有 120 家公司，则成本为 35.8 美元，按照汇率 1: 7.14 折合人民币 255.5 元。而且智能体并不是完美的，在运行过程中出现幻觉或错误循环也是常事。例如智能体在找不到合适新闻的时候会进行大量重复的内容输出，人工介入纠错又会增加一定成本。

图11 单次运行 Agent 对 Tokens 的消耗



资料来源：OpenAI，海通证券研究所

图12 单次运行 Agent 的实际成本



资料来源：OpenAI，海通证券研究所

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。

3.未来展望：人工智能发展迅猛，智能体商业化落地

未来多方面推动人工智能发展，应用级别智能体有望快速落地。国内各地相继出台关于人工智能的发展政策，推动其为重要的研究方向。预计 2026 年国内人工智能市场规模超过 260 亿美元，全球人工智能市场规模 2025 年超 6 万亿美元。海外以美国为例，相关政策出台时间较早，人工智能领域发展更加成熟，许多智能体应用已在服务各类企业。并且美国有意与人工智能强国组成战略伙伴，共同发展 AI 科技。

多模态大模型利用异构数据提升应用效率，促进 AI Agent 发展。将智能体赋能于图片+语音的多模态大模型中，可以减少使用工具和交互的过程，使智能体完成更多复杂任务，解决跨行业、跨领域的问题。智能体发展能推动政府、金融、制造、能源、医疗、零售等行业的智能化应用向多模态和跨模态转变。

3.1 海内外政策推动人工智能发展加速

多地陆续出台关于人工智能发展的指导文件，2024 年政府与企业的投入资金有望迎来大幅增长。国务院在《2024 政府工作报告》中提出，要开展“人工智能+”行动，实施数字化转型。智能体作为人工智能应用领域的重要研究方向，相信众多企业和研究部门会开展更多研究工作。另外，像上海、杭州、广东、北京等省市也出台了相关政策，其中杭州市推出“算力券”约 2.5 亿用于支持本地的算力发展。广东省也将算力作为重点发展方向，预测 2025 年的人工智能核心产业规模超过 3000 亿元。

表 2 国内政策推动人工智能发展

政府机构	文件名称	发布时间	具体内容
上海市经济和信息化委员会、上海市发展和改革委员会等五部门	《上海市推动人工智能大模型创新发展若干措施（2023-2025 年）》	2023 年 10 月	实施大模型创新扶持计划。支持引进高水平创新企业，支持本市创新主体打造具有国际竞争力的大模型，鼓励形成数据飞轮，加速模型迭代，对取得重大成果的予以专项奖励。实施大模型示范应用推进计划。重点支持在智能制造、生物医药、集成电路、智能化教育教学、科技金融、设计创意、自动驾驶、机器人、数字政府等领域构建示范应用场景，打造标杆性大模型产品和服务。
国务院	《2024 政府工作报告》	2024 年 3 月	深化大数据、人工智能等研发应用，开展“人工智能+”行动，打造具有国际竞争力的数字产业集群。实施制造业数字化转型行动，加快工业互联网规模化应用，推进服务业数字化，建设智慧城市、数字乡村

杭州市人民政府
《关于支持人工智能全产业链高质量发展的补充意见》
(征求意见稿)
2024 年 5 月

1) 支持算力技术攻关。鼓励企业开展算力关键技术研发攻关, 市级科技研发计划给予专项支持。鼓励企业承担国家、省重点研发任务, 加速形成重大科研成果, 对符合政策规定的项目按国家、省实际到账资助经费的 25% 给予资助, 资助金额最高不超过 500 万元。2) 加大融资贴息支持。对纳入政府统一布局的算力基础设施建设项目(总投资规模不低于 5000 万元)3) 扩容算力券发放。每年设立“算力券”总额提升至 2.5 亿元, 新增算力券支持企业购买符合杭州市规划布局的智算中心以及纳入算力撮合的合作伙伴所供给的智能算力。

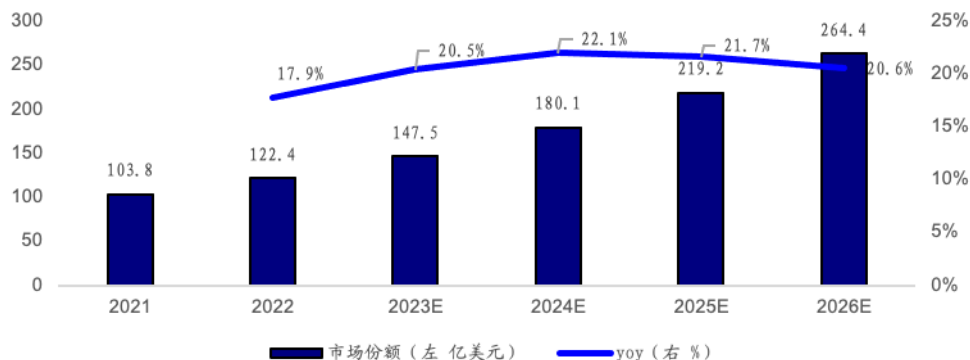
广东省人民政府办公厅
《广东省关于人工智能赋能千行百业的若干措施》
2024 年 6 月

到 2025 年, 全省算力规模超过 40EFLOPS, 人工智能核心产业规模超过 3000 亿元。到 2027 年, 全省人工智能产业底座进一步夯实, 算力规模超过 60EFLOPS, 全国领先的算法体系和算力网络体系基本形成; 智能终端产品供给丰富, 在手机、计算机、家居、机器人等 8 大门类, 打造 100 款以上大规模使用的智能终端产品, 人工智能核心产业规模超过 4400 亿元; 聚焦制造、教育、养老等领域, 打造 500 个以上应用场景, 各行各业劳动生产率显著提升。

资料来源: 上海市经济和信息化委员会, 国务院, 杭州市人民政府, 广东省人民政府, 中国政府网, 新华社, 海通证券研究所

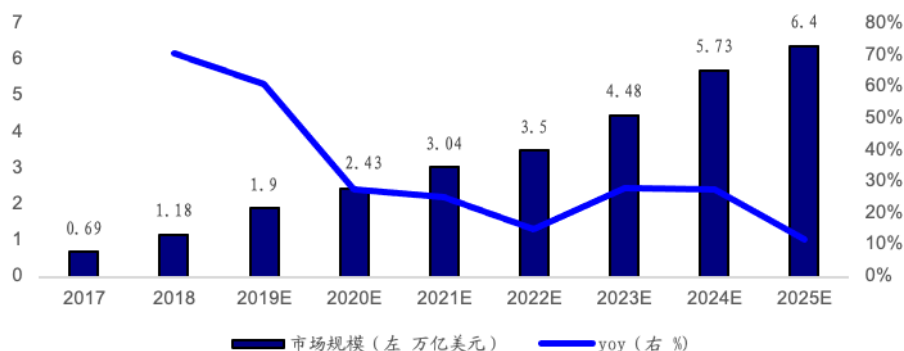
中国人工智能市场份额稳定增长, 未来 AI 实体项目增多利好智能体发展。根据 IDC 预测, 国内人工智能市场规模预计在 2026 年达到 264.4 亿美元, 2021 年到 2026 年期间的 CAGR 超 20%。根据增长率计算, 从 2022 年开始, 每年市场规模会保持 20% 左右的稳定增长态势。根据澎湃新闻援引德勤统计, 未来 2025 年世界人工智能市场规模将超过 6 万亿美元, 2017 年-2025 年期间复合增长率将超 30%。AI 高速发展中或利好更多实体企业, 专门提供智能体产品服务的厂商规模也将快速扩张。

图13 中国人工智能市场规模预测



资料来源: IDC, 海通证券研究所

图14 全球人工智能市场规模预测



资料来源: 德勤, 澎湃新闻, 海通证券研究所

美国相关政策落地较早, 人工智能发展已成规模。在 2021 年 3 月, 大模型还未进入公众视野时, 美国人工智能国家安全委员会已经发布了对该领域未来发展的规划。文件中写到了美国坚定在人工智能领域的投入, 且规定了研发投入资金的比重。另外, 在 2023 年 5 月, 美国白宫发布《国家人工智能研发战略计划》, 对之前提出的 8 项重点战略目标进行了优化和完善, 同时强调了人工智能领域国际交流的重要性。另外, 文件还

对比了 2020 年发布计划中的各项指标与实际指标差距，并针对细分领域加强研发投入。

表 3 海外政策推动人工智能发展

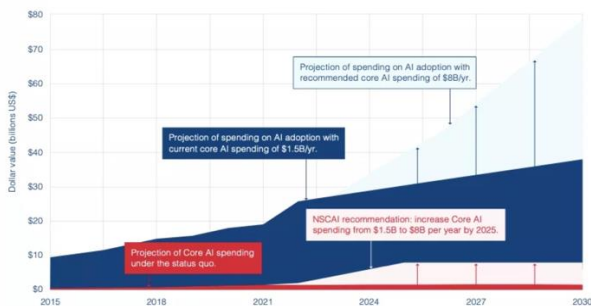
政府机构	文件名称	发布时间	具体内容
美国人工智能国家安全委员会	《NSCAI 最终报告》	2021 年 3 月	美国国家人工智能安全委员会（NSCAI）的任务是向总统和国会提出建议，以“推进人工智能、机器学习发展以全面解决美国国家安全和国防需求的相关技术。”在设立本委员会时，2019 财年约翰·S·麦凯恩国防授权法案第 1051 节指示 NSCAI 从国家竞争力的角度审视人工智能及其维持技术优势的手段、国际合作和竞争力的趋势、促进基础和先进研究投资的方式、劳动力和培训、军事使用的潜在风险、伦理问题、建立数据标准和激励数据共享，以及人工智能的未来发展等方方面面。
美国白宫	《国家人工智能研发战略计》	2023 年 5 月	该计划是对 2016、2019 年版《国家人工智能研发战略计划》的再次更新，重申了之前的 8 项战略目标并对各战略的具体优先事项进行了调整和完善，同时增加了新的第 9 项战略以强调国际合作。此外，该报告还提出要评估联邦机构对《2020 年国家人工智能倡议法案》（NAIIA）和《国家人工智能研发战略计划》的实施情况。

资料来源：清华大学人工智能国际治理研究院，安全内参，海通证券研究所

美国增强型人工智能投资研发保持高增速，AI 战略联盟强大。2024 年美国在该方面的投入预计为 15 亿美元，NSCAI 希望在 2025 年把这一数字提高到 80 亿美元，增幅高达 433%。且在 2025 年之后，保持每年投入增长 80 亿美元的趋势。最高预计在 2030 年投入高达 350 亿美元。美国强调了 AI 盟友的概念，与澳大利亚、加拿大、英国、日本等国家构建人工智能战略伙伴关系，未来将会吸引更多 AI 强国加入联盟当中。

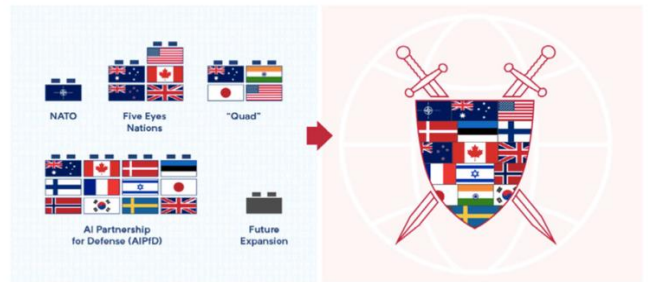
图15 美国增强型人工智能研发投入 2015-2030

Enhanced AI R&D Investment, FY 2015-2030
Source: Govini and NSCAI



资料来源：清华大学人工智能国际治理研究院，海通证券研究所

图16 美国及其 AI 盟友



资料来源：清华大学人工智能国际治理研究院，海通证券研究所

3.2 算力规模高增，企业对 AIGC 的投入意愿强

2021 年中国的智能算力规模为 155.2 EFLOPS（每秒百亿亿次浮点运算）。据 IDC 预测，规模将在 2026 年达到 ZFLOPS 级别（每秒十万亿亿次浮点计算）。预计在 2021-2026 年期间，智能算力增长的 CAGR 为 52.3%。随着算力规模的大幅增长，将会带动 AI 技术快速发展。根据中国信通院的测算，每在算力中投入 1 元，将带动 3-4 元的经济产出。在智能体发展中，更强的算力能支持更复杂的模型计算并且给予其更快的响应速度。当大模型能快速迭代，AI Agent 将会有更强大的基座，相同运行模式下或能显著提升 F1-score 和 Recall 水平。

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/198070014143006132>