

生成式人工智能投资的短期与长期策略

许英博

面对未来，人们总是习惯高估一年的变化，低估五到十年的变化。

2023年，美股科技股领涨全球。英伟达、特斯拉、微软、脸书等公司涨幅居前。体现全球投资人对于生成式人工智能（AIGC，AI Generated Content）的较高预期。以 ChatGPT 为代表的生成式人工智能打开资本市场想象力，其令人惊叹的自传播性，快速完成投资者教育，引发产业界和资本市场对人工智能（AI，Artificial Intelligence）的新一轮高度关注。人们重新审视和评估数字化、智能化在千行百业中可能带来的机遇和潜力。

笔者判断，本轮生成式人工智能将孕育长期可持续的投资机遇。我们将人工智能产业链划分为：基础算力、数据和模型、应用层等。长期看，生成式人工智能重构千行百业，应用层蕴藏规模最庞大、收益最丰厚的投资机会，但仍需时日等待 AI 原生爆款应用的出现。中期看，数据治理和模型精进，是确保人工智能应用能够大规模产业化落地的前提，MaaS（Model as a Service）+API（Application Programming Interface）调用模式或成为潜力方向。短期看，在爆款应用和商业模式不确定的情况下，算力是当前相对确定的投资机会。

生成式人工智能引领 2023 年资本市场上涨

生成式预训练模型（GPT，Generative Pre-Trained Transformer）持续升级迭代。2022年11月30日，美国科技公司 OpenAI 发布 ChatGPT，其注册用户数快速突破1亿，成为史上用户增长最快的



许英博，中信证券研究部董事总经理，首席科技产业分析师。2007年硕士毕业于清华大学汽车工程系，同年加入中信证券研究部，任汽车行业分析师。2016年，许英博创立中信证券前瞻研究团队，在中国证券市场上率先以跨行业的全球前沿科技为主要研究对象，研究内容覆盖互联网、科技硬件、云计算和人工智能、自动驾驶等。

消费者应用，引发资本市场、产业界和学界高度关注。2023年3月14日，升级版本 GPT-4 发布，进一步提升文字输入限制至 2.5 万字，准确率提升、识图水平提升。2023年9月，GPT 更新支持基于图像和语音进行分析和对话。

GPT 驱动 OpenAI 收入超预期。据媒体报道，受益于付费版 ChatGPT 订阅用户量增长，OpenAI 在 2023 年的收入规模或超过 13 亿美元，显著超越此前资本市场的预期。2022 年，公司的收入仅为约 2800 万美元。收入规模快速上涨，驱动公司在一级市场的估值显著提升。OpenAI 年初接受微软等公司超 100 亿美元的投资，投后估值约为 270-290 亿美元。媒体报道，近期公司计划出售员工持股，估值达近 900 亿美元。

微软持有 OpenAI 约 49% 的股权，公司业务和估值受益于此。微软 Copilot 人工智能助手将 GPT 整合嵌入至操作系统 Windows 11 和办公软件

Microsoft 365 Copilot，有助提升用户效率。微软采用“加量加价”模式，将 Copilot 功能定价于每月 30 美元。资本市场预期，人工智能应用将有助公司打开市场空间，微软股价年内最高上涨超 60%。

大模型训练需要规模庞大的 GPU (Graphics Processing Unit) 算力。英伟达凭借芯片的出色性能，以及 CUDA (Compute Unified Device Architecture) 软件生态，具有较强的市场竞争力。大模型企业的持续涌现、模型版本的持续迭代，科技巨头的算力军备竞赛，加剧 GPU 现阶段供需紧张状态。英伟达收入和订单快速增长，今年以来截至 11 月 30 日，公司股价上涨约 227%。

脸书推出开源大模型 Llama2，并开放商业

许可，获得市场关注，公司股价上涨约 162%。Transformer 亦可应用于汽车自动驾驶领域，特斯拉股价上涨约 122%。

人工智能助推美股领涨全球。今年以来截至 11 月 30 日，美股纳斯达克指数和费城半导体指数分别上涨 37% 和 49%。同期，中国人工智能和科技相关指数表现相对平淡，中信半导体指数上涨 0.6%，中信电子指数和中信计算机指数分别上涨 6.7% 和 5.8%；港股恒生科技指数下跌 8%。

从大模型说起

生成式人工智能基于模型训练和人类反馈强化学习 (RLHF, Reinforcement Learning from Human

	港股		美股			A股			
	恒生指数	恒生科技	纳斯达克指数	纳斯达克中国金龙指数	费城半导体指数	上证指数	电子(中信)	计算机(中信)	半导体(中信)
2023-01-03	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
2023-01-31	8.42%	7.30%	11.53%	13.68%	16.83%	4.47%	6.83%	7.22%	5.16%
2023-02-28	-1.78%	-7.28%	10.29%	1.22%	18.29%	5.23%	7.05%	12.89%	3.54%
2023-03-31	1.26%	1.66%	17.67%	4.87%	29.18%	5.02%	14.83%	31.95%	14.97%
2023-04-28	-1.24%	-7.84%	17.71%	-5.53%	19.75%	6.63%	8.96%	23.02%	12.71%
2023-05-31	-9.49%	-14.33%	24.53%	-13.84%	38.07%	2.83%	10.63%	24.29%	12.33%
2023-06-30	-6.10%	-7.61%	32.74%	-5.72%	46.86%	2.75%	9.47%	25.83%	6.04%
2023-07-31	-0.33%	7.48%	38.12%	13.03%	54.40%	5.60%	8.00%	18.20%	4.05%
2023-08-31	-8.75%	-1.27%	35.12%	1.38%	46.78%	0.11%	3.60%	14.41%	-0.65%
2023-09-28	-13.76%	-10.75%	27.09%	-5.34%	36.79%	-0.19%	1.27%	8.36%	-5.32%
2023-10-24	-15.05%	-11.20%	23.72%	-10.41%	28.58%	-3.14%	5.79%	2.97%	0.04%
2023-11-30	-15.40%	-7.88%	36.96%	-7.08%	48.92%	-2.79%	6.65%	5.76%	0.57%

全球资本市场科技相关板块走势 (2023 年 1 月 - 2023 年 11 月)

港中美市场 AI 标的表现对比 (2023 年 1 月 - 2023 年 11 月)

资料来源: wind, 中信证券研究部

	美股					中概					A股				
	特斯拉	英伟达	脸书	微软	谷歌	阿里巴巴	百度	腾讯控股	美团	理想汽车	金山办公	科大讯飞	工业富联	中际旭创	海康威视
	TSLA.O	NVDA.O	META.O	MSFT.O	GOOGL.O	BABA.N	BIDU.O	0700.HK	3690.HK	LI.O	688111.SH	002230.SZ	601138.SH	300308.SZ	002415.SZ
2023-01-03	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
2023-01-31	60.24%	36.48%	19.42%	3.44%	10.91%	19.81%	13.05%	16.57%	-1.24%	18.68%	-0.93%	18.07%	2.15%	2.83%	4.55%
2023-02-28	90.30%	62.18%	40.24%	4.37%	1.05%	-4.56%	15.58%	4.80%	-23.02%	12.54%	2.49%	40.65%	3.11%	25.44%	11.93%
2023-03-31	91.91%	94.07%	69.91%	20.64%	16.39%	11.09%	26.69%	17.67%	-18.83%	18.92%	66.58%	85.28%	84.96%	108.13%	24.41%
2023-04-28	52.00%	93.88%	92.66%	28.57%	20.44%	-7.93%	1.24%	5.04%	-24.77%	12.01%	55.03%	76.32%	73.15%	218.02%	10.09%
2023-05-31	88.65%	164.34%	112.22%	37.71%	37.87%	-13.51%	3.12%	-4.59%	-37.67%	38.47%	52.14%	73.84%	93.88%	279.86%	3.71%
2023-06-30	142.16%	195.59%	130.06%	42.81%	34.31%	-9.38%	14.92%	1.86%	-30.83%	67.30%	66.31%	97.73%	170.68%	421.96%	-1.50%
2023-07-31	147.39%	226.52%	155.41%	40.87%	48.92%	11.07%	30.94%	8.86%	-17.25%	104.00%	43.22%	83.23%	145.81%	353.11%	8.05%
2023-08-31	138.74%	244.87%	137.21%	37.74%	52.79%	1.00%	19.89%	-0.17%	-27.32%	98.52%	39.20%	59.77%	141.19%	305.75%	4.95%
2023-09-28	127.92%	201.11%	143.67%	31.81%	48.46%	-7.00%	11.18%	-8.65%	-37.33%	64.59%	30.80%	47.62%	116.86%	309.92%	0.55%
2023-10-31	85.79%	184.98%	141.52%	42.09%	39.23%	-10.26%	-11.86%	-11.17%	-37.44%	61.15%	1.28%	32.76%	62.04%	210.63%	8.44%
2023-11-30	122.09%	226.84%	162.27%	59.56%	48.71%	-18.59%	-0.39%	0.45%	-48.76%	76.12%	6.45%	30.40%	66.22%	236.01%	3.50%

Feedback)，建立在超大规模数据学习基础上，可用于文本、图片、视频、代码等多模态信息生成和交互。生成式人工智能通过分析字词间的统计学关联，预测句子中将要出现的下一个词或者空缺的词的概率，不断计算空缺单词的概率分布，最终生成完整的句子。

2017年谷歌 Bert 为代表的预训练+微调的 Transformer 模型成为当时人工智能的主流方向。OpenAI 在其基础上改进，提出了 GPT-2 模型，使其更适合生成式任务。早期的 GPT-2 模型并不完善，基准成绩不算优秀。经过模型持续迭代，人工智能开始在语言任务的部分场景中追平人类。

2020年发布的 GPT-3 参数量高达 1750 亿，较 GPT-2 的 15 亿参数量显著提升。2022年初，OpenAI 推出 InstructGPT，引入人类反馈强化学习 RLHF 来进一步调整模型输出结果。尽管只有 13 亿参数，InstructGPT 的输出效果却有显著提升，能够生成类似人类对话的结果，同时可遵循公共价值取向。

2022年11月底推出的 ChatGPT 使用与 InstructGPT 相同的人类反馈强化学习训练方法，并对数据收集优化，互动效果优秀。ChatGPT 上线仅 5 天后，其注册用户数即超过 100 万；上线两个月后，

注册用户数破亿。

今年 3 月，OpenAI 正式推出多模态大模型 GPT-4，支持图像和文本的输入输出，拥有强大的识图能力。根据官网信息，GPT-4 面对复杂问题的能力大幅提升，在一些专业问题和学术基准上表现已经和人类持平，如在 AP 考试、GRE 考试等测试中表现优异。9 月，GPT 更新支持基于图像和语音进行分析和对话。

OpenAI 从开源走向闭源，加速商业化。与以往不同，OpenAI 既没有公开发布 GPT-4 相关论文，也没有提供详细的框架说明，仅提供了一份 98 页的技术文档，主要描述模型能力以及相关评测的得分，几乎没有任何技术细节。通过提供对应的 API 接口，OpenAI 闭源加速商业化进程。

看到 ChatGPT 及其演进版本的效用和市场反响，其他竞争对手快速跟进。7 月 19 日，脸书发布其最新开源可商用版本 Llama2 大语言模型。模型层面的最大亮点在于开源了与 ChatGPT 相同的监督微调（SFT, Supervised Fine-Tuning）与人类反馈强化学习 RLHF，成为目前开源类别中最接近 ChatGPT 的大语言模型。脸书的论文向业内拆解了 SFT 以及 RLHF 等的技术过程，并给出了可以直接使用的版本。我们判断，开源模型迭代进程将加速。

OpenAI 发展路径及 GPT 产品迭代



资料来源：OpenAI，中信证券研究部

中国AGI主要参与者			
互联网巨头	高校研究院	创业公司	传统AI企业
百度—文心一言4.0	智源研究院（北大黄铁军）—悟道	搜狗王小川—百川智能	云从科技—从容1.5
阿里—通义千问2.0	IDEA研究院（郭建、谢育涛等）	创新工厂李开复—Project AI	商汤科技—日日新
腾讯—混元大模型3.0	中科自动化所—紫东太初	微软亚研院周明—澜舟科技	云知声—山海
字节跳动—研发中	清华（唐杰团队）—智谱AI 证券	京东周伯文—街远科技 03:18:41	科大讯飞—星火3.0
快手—快意	清华（刘知远、曾国洋等）—面壁智能	商汤闫俊杰—MiniMax	第四范式—式说3.0
华为—盘古大模型3.0	清华（岂凡超、李潇翔等）—深言科技	奇点智源	
奇虎360—智脑	复旦（邱锡鹏团队）—MOSS	快手李岩—元石科技	
知乎—知海图AI	西大（蓝振忠团队）—西湖心辰	字节王长虎—爱诗科技	
京东—言犀	人大（卢志武等）—智子引擎 00:22:00	京东梅涛—智象未来 03:18:41 中信证券	
美团—光年之外		Gitlab陈冉—OpenCSG	
小红书—研发中			

资料来源：企查查、各公司官网，中信证券研究部

12月6日，谷歌发布新一代大模型 Gemini。模型包括 Ultra、Pro、Nano 三个版本，能够在从数据中心到移动端的多设备高效运营，多模态能力强大。谷歌在人工智能领域有着深厚基础，最早提出 Transformer 模型，但在大模型产业化训练过程中不够坚定，错失先机。在 ChatGPT 推出近一年之后，谷歌持续打磨技术和产品，追赶 OpenAI 和微软。

此外，海外的初创模型和应用公司中也有如 Stable Diffusion、Midjourney、Anthropic、Cohere、Adept、Pika 等市场参与者。

在中国市场上，大量科技互联网巨头、初创公司参与到竞争当中。腾讯混元大模型、百度文心大模型、阿里通义大模型、快手快意大模型等互联网背景大模型具有较大的用户基础、较大规模的研发支出，快速开展通用模型建设。商汤、讯飞等人工智能公司，凭借此前积累的技术和资源优势，快速跟进迭代。以智谱、Minimax 为代表的初创公司，在一级市场获得较高关注。清华系等高校背景大模型公司亦从科研走向产业化。此外，字节跳动大模

型亦值得期待。

生成式人工智能驱动算力需求快速提升

生成式人工智能训练对算力的要求持续提升。OpenAI 预计，人工智能科学研究要想取得突破，所需消耗的计算资源每 3-4 个月就要翻一倍。

英伟达是全球高性能计算龙头公司。2022 年英伟达新推出的 H100 芯片搭载 Transformer 引擎，使用每层统计分析来确定模型每一层的最佳精度，在保持模型精度的同时实现最佳性能，相较于上一代产品提供 9 倍的训练吞吐量，性能提升 6 倍。

过去十年，英伟达驱动人工智能计算能力持续提升。以单芯片 Int8 指标为例，2012 年推出的主力芯片算力约为 3.94Tops，2022 年推出的 H100 芯片算力约为 3958Tops。如综合考虑处理器、系统、连接等系统级因素，英伟达创始人黄仁勋表示，过去十年中人工智能处理加速和推进了一百万倍。同时他表示，相信未来十年人工智能处理将再加速一百万倍。

生成式人工智能浪潮引爆后，大模型公司持续

涌现、大模型版本持续训练迭代，高性能 GPU 需求爆发。核心厂商英伟达收入快速调整。2023 年第二季度，公司收入达 135 亿美元，同比增长 101%，环比第一季度增长 88%。同时，公司指引第三季度的收入预期可达 160 亿美元。综合考虑市场需求和产能周期等因素，英伟达 GPU 产能紧张状态或持续至 2024 年。

一枝独秀的同时，英伟达在算力市场也面临挑战。谷歌、亚马逊、微软等科技巨头和需求大户，纷纷开展人工智能芯片自研计划，特斯拉亦推出 Dojo，希望用于人工智能训练。AMD 和 Intel 亦积极拓展人工智能芯片市场。

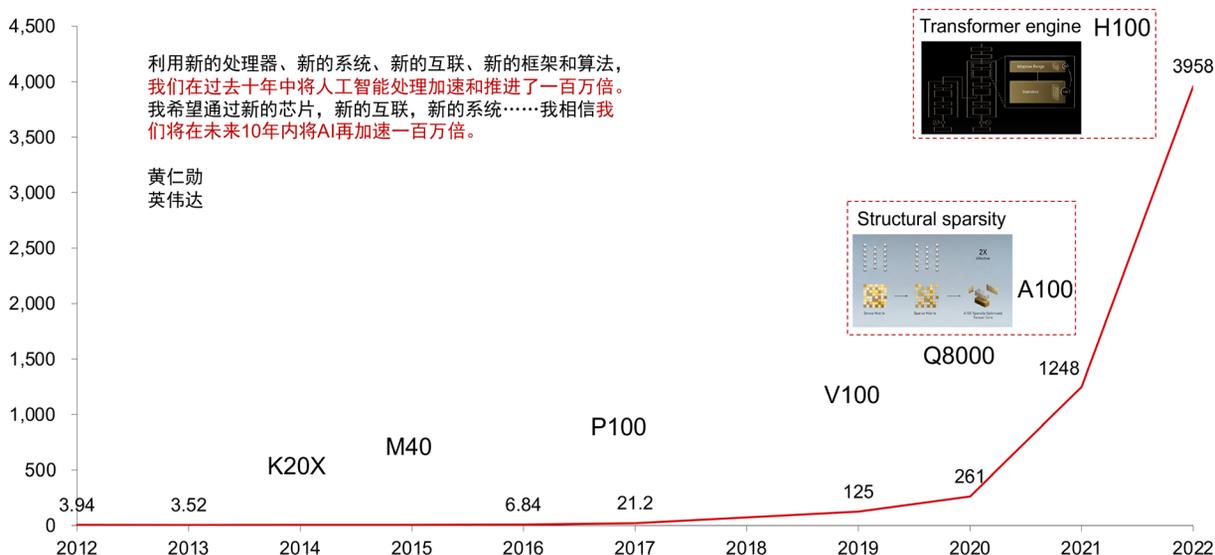
中国企业是英伟达 GPU 重要的需求方，但美国政府的禁令将导致英伟达无法向中国客户出售人工智能芯片。10 月 17 日，美国商务部推出针对人工智能和半导体领域新的禁令。在新规下，面向中国市场出口高性能 GPU 如 A100、H100、A800、H800、L40、L40S、4090 等均会受到影响。此举将影响英伟达 2024 年在华销售规模和总收入预期。中国企业将被迫寻找替代方案。

应用端的机遇: AI 原生应用和下一代科技硬件

互联网和移动互联网等科技大潮的经验表明，人们倾向于高估新技术带来的短期变化，低估技术浪潮的长期潜力。个人电脑的年销量峰值约 3.5 亿台，建立在此基础上的互联网产业巅峰市值约千亿美元量级。智能手机的年销量峰值约 14 亿部，建立在此基础上的移动互联网产业市值以万亿美元计。人工智能将带来千行百业数字化和智能化，长期看有望打开万亿美元潜在市场空间和更大规模的市值空间，甚至孕育下一代科技硬件。

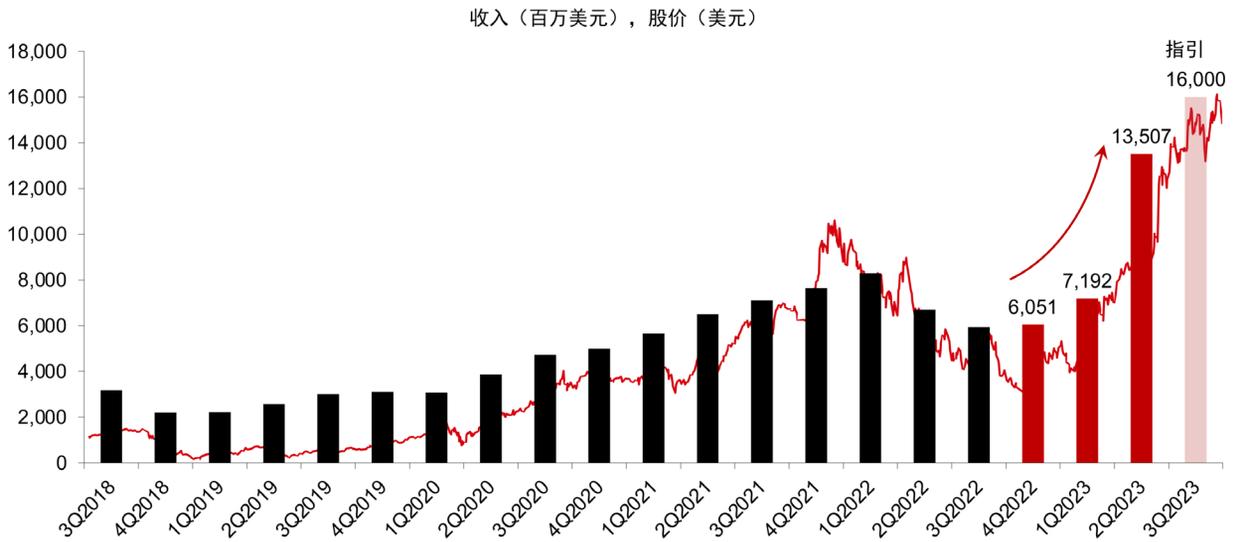
苹果 Vision Pro 首次提出“空间计算”等概念，未来“AI+ 科技硬件”的结合有望带来更多科技创新的可能。苹果 Vision Pro 是基于现实空间进行计算，经人工智能技术辅助，用户看到的虚拟画面可以理解为在真实世界基础上叠加了实时计算结果。交互方式亦从二维转换为三维，更符合人类现实生活场景的交互，不再借助手柄等外设，直接使用眼睛、双手、语音进行人机交互。AI+AR，带来中长期更多创新可能。

英伟达单芯片推理能力 (Int8 Tops 为例)



资料来源: 英伟达, 中信证券研究部

英伟达季度收入（含预测）与股价走势



资料来源: Bloomberg, 英伟达 (含预测), 中信证券研究部

生成式人工智能的潜力仍有待释放。在移动互联网时代，我们看到了专业内容创作 (PGC, Professional Generated Content) 向用户创作 (UGC, User Generated Content) 平台转变带来的潜力，字节跳动旗下抖音在全球快速斩获用户并保持高速增长。未来，生成式人工智能 AIGC 或将进一步改变内容产业，为我们提供创作门槛更低的丰富创作思路。Gartner 预计，到 2025 年，生成式人工智能将占有生成数据的 10%。

AI 原生应用很可能是人工智能中长期最重要的机遇。SaaS (Software as a Service) 的流量优势显现，叠加人工智能技术提升用户体验，可以显著提升用户粘性。微软 Copilot 是近期最值得关注的人工智能应用，如果其订阅数据保持持续稳定增长，意味着 SaaS+AI 商业模式被用户接受，具有显著的示范效应。Adobe 和 Salesforce 也推出基于人工智能的服务。以 Adobe 为例，Firefly 通过独立网址被访问期间，在今年 3 月 21 日到 6 月 8 日的两个多月内帮助用户生成超过 2 亿张图片，被整合进入 Photoshop 应用后，在 5 月 23 日到 6 月 8 日的两周多内帮助用户生成图片数量已高达 1.5 亿张。未来，

基于 MaaS 模式的大模型 API 调用方式，可能进一步改变软件开发和应用形态，AI 原生有望带来新的投资机会。

人工智能整体进展显著，算力、基础大模型快速迭代，但在垂直领域的数据和算法方面仍面对诸多挑战。数据方面，传统行业的数据治理仍处于早期。数据清洗、数据标注、数据结构化、数据规模化既需要时间，更需要传统产业决策者的认知提升。算法方面，需要开发者针对不同的垂直场景的数据进一步微调模型，处理“幻觉 (Hallucination)”带来的挑战。幻觉指大语言模型在回答时产生的大量虚假内容、甚至虚构事实。为了减轻幻觉的问题，需要开发者针对不同的垂直场景的数据进一步微调模型。

传统行业决策者和资本市场对人工智能期待较高，市场预期提升过快，但人工智能大规模产业化落地仍需要时间，从产业试水到真正贡献收入和利润需要更长时间。此外，中美人工智能产业客观差距仍存，中国亦面对算力硬件的瓶颈。

尽管如此，展望未来，我们依然坚信，人工智能将重构千行百业，带来长期投资机遇。🌟