

# ChatGPT 跨行业专题报告

## AIGC 发展大年，推动新一轮产业革命

近年全球科技龙头持续投入 AIGC 领域以及近期 ChatGPT 应用落地加速，我们认为 2023 年 AIGC 大模型与产业将迎来快速发展期，赋予人工智能大规模落地场景，为底层算法和芯片等领域带来巨大需求，同时加速 AIGC 应用渗透和商业化进程。在本篇报告中，我们对 ChatGPT/ AIGC 算法层、算力层和应用层的发展和产业链进行深度分析，建议重点关注上游算法商汤和科大讯飞，中游芯片景嘉微和寒武纪，及下游应用微软、谷歌、Meta、百度、腾讯、阿里和字节。

- **算法层：ChatGPT 底层的大模型技术演进。**大模型是在大规模无标注数据上进行训练，学习出特征和规则，泛化能力强。大模型技术的兴起，有效降低了 AI 应用研发门槛，也能解决在特定应用场景有效数据不足导致模型精度低的问题。现阶段的大模型主要是面向 NLP 领域，包括 ChatGPT 背后的 GPT-3，而在 CV 领域，多个公司也在积极加大技术储备。针对 5 年来中美发布的主要大模型，我们从两个维度进行了比较：1) 在模型参数量上，模型体量随时间呈指数扩张趋势，美国同等参数量级的大模型比中国领先 1-2 年；2) 在 NLP 专利数量上，微软和谷歌的持有数已经在几千个的量级，而国内领先的 BAT 持有数则在 600 个左右。
- **算力层：大模型驱动 AI/内存芯片需求全面加速。**AI 芯片为人工智能应用提供所需的基础算力，主要分为 GPU、FPGA 和 ASIC，ChatGPT 基础模型上大规模预训练需要强大算力作为支持，而大算力需要更多高算力和内存芯片去支撑算法运行。目前 GPU 为主导 AI 芯片，因具有强大计算能力和高通用适用性，广泛应用于 AI 和机器学习领域。近年多家国内半导体公司积极布局 GPU 领域，受中美科技摩擦影响，预计国产 GPU 芯片替代将快速推进。
- **应用层：搜索+智能客服先行，内容创作+文娱生活场景渗透加速，长期看好 B 端产业机会。**进入“数智化时代”，看好 ChatGPT/AIGC 应用场景持续渗透及商业化加速（包括搜索引擎、智能客服、教育培训、金融等），助力行业降本增效的同时，开拓创新变现增量。我们认为 ChatGPT 较难取代搜索引擎，更多体现在赋能优化，在于离线非实时数据库、训练模型下答案非正确、成本过高。中短期，基于内容创作及文娱生活导向的 AIGC 场景应用更为集中，包括文娱、传媒、电商等，长期产业升级机会可期。预计 2030 年 AIGC 市场规模达千亿美元，商业模式从按量付费、订阅延伸至 MaaS。
- **受益标的：**在 ChatGPT/AIGC 应用层面，建议重点关注拥有海量数据、大模型能力及 AI 内容创作产品丰富（文/音/图/视）的头部平台，包括 OpenAI、微软、谷歌、Meta、百度、腾讯、阿里、字节；可同时关注近期推出类 ChatGPT 服务或泛 AI 产品的垂类腰部平台，如网易、知乎、阅文、昆仑万维等。我们看好百度在 ChatGPU 及 AIGC 的全栈技术储备、搜索引擎协同及产品管线，受益于：1) 文心一言待推出，数字人+智能创作应用矩阵丰富；2) 文心大模型技术支持及海量数字资产优势（语料库、参数量大）。在 AI 底层算法领域，建议关注拥有 CV 大模型的商汤及在智能语音技术领先的科大讯飞。在芯片算力领域，建议关注 GPU 领先企业景嘉微、寒武纪、海光信息和龙芯中科，以及 FPGA 龙头紫光国微和复旦微电。

伍力恒

(852) 3900 0881

alexng@cmbi.com.hk

黄群

(852) 3900 0889

sophiehuang@cmbi.com.hk

颜宇翔

(852) 3916 3719

marleyngan@cmbi.com.hk

杨天薇博士

(852) 3916 3716

lilyyang@cmbi.com.hk

李汉卿

lihanqing@cmbi.com.hk

刘梦楠

claudialiu@cmbi.com.hk

许义鑫

easonxu@cmbi.com.hk

李博文

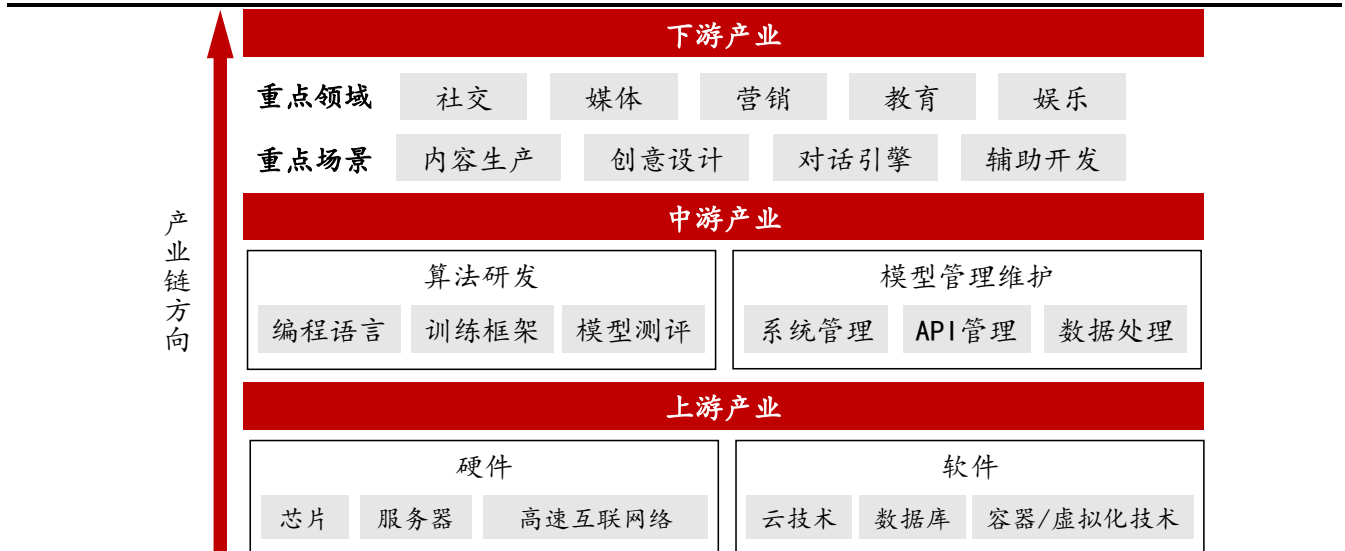
libowen@cmbi.com.hk

相关报告：

中国互联网行业 - ChatGPT & AIGC 在中国市场的发展前景  
([link](#))

## 大模型产业链及潜在受益标的

图 1：大模型产业链梳理



资料来源：智源研究院，招银国际环球市场

图 2：大模型产业链投资标的

产业链	国外潜在受益标的	国内潜在受益标的
应用	微软、谷歌、Facebook	百度、腾讯、阿里巴巴、字节跳动、华为、商汤、科大讯飞、思必驰
AI 模型、算法	深度学习框架：谷歌(TensorFlow)、Facebook (PyTorch)、微软 (CNTK)	百度 (PaddlePaddle)、华为 (MindSpore)、商汤 (SenseParrots)
	大模型：微软、谷歌、Facebook	百度、腾讯、阿里巴巴、商汤
基础计算设施	CPU：Intel	
	GPU：Nvidia, AMD	华为昇腾，燧原科技，壁仞科技
	FPGA：Intel, Xilinx	
	ASIC	百度，复旦微电，寒武纪
	服务器周边	澜起科技、中际旭创
	数据中心	万国数据、秦淮数据

资料来源：招银国际环球市场

## 受益标的估值比较

图 3: 受益标的估值比较

公司	代码	产品	股价	市值	P/E		P/B		ROE
			(当地货币)	(百万美元)	FY23E	FY24E	FY23E	FY24E	FY23E
<b>云计算</b>									
微软	MSFT US	AI 云计算	266.73	1,985,486	28.6	24.8	9.3	7.3	36.4
谷歌	GOOG US	AI 云计算	100	1,276,670	18.2	15.2	4.5	3.9	24.6
百度	BIDU US	AI 云计算	152.34	52,641	16.5	14.6	1.6	1.7	7.3
腾讯	700 HK	AI 云计算	395.6	481,702	23.4	20.0	3.2	2.8	14.4
阿里巴巴	BABA US	AI 云计算	105.11	278,282	13.4	13.3	2.2	2.3	8.7
<b>AI 算法</b>									
商汤	20 HK	AI 算法及应用	2.78	11,853	n.a.	n.a.	3.2	3.5	(11.7)
科大讯飞	002230 CH	AI 算法及应用	46.4	15,911	51.5	38.5	5.4	5.0	11.3
创新奇智	2121 HK	AI 算法及应用	20	1,425	n.a.	n.a.	4.8	5.2	(14.6)
虹软科技	688088 CH	AI 算法及应用	28.73	1,721	42.3	29.5	3.8	3.3	7.8
云从科技	688327 CH	AI 算法及应用	31.7	3,465	n.a.	n.a.	12.0	12.2	(26.7)
<b>芯片</b>									
英伟达	NVDA US	GPU	222.05	546,243	68.1	51.3	25.6	23.0	31.1
超威半导体	AMD US	GPU	84.69	136,550	27.1	19.7	2.4	2.3	7.4
寒武纪	688256 CH	GPU	79.07	4,677	n.a.	n.a.	7.0	7.4	(15.9)
景嘉微	300474 CH	GPU	82.67	5,545	80.3	58.2	10.5	9.1	12.7
海光信息	688041 CH	CPU	53.51	18,353	82.3	56.3	7.6	6.7	10.9
龙芯中科	688047 CH	CPU	113	6,687	101.5	76.4	8.9	8.0	8.9
莱迪思	LSCC US	FPGA	81.41	11,158	40.3	34.3	n.a.	n.a.	50.7
安路科技	688107 CH	FPGA	73.03	4,312	308.1	138.8	18.1	16.3	5.8
紫光国微	002049 CH	FPGA	126.85	15,903	26.3	19.9	7.6	5.5	30.5
复旦微	1385 HK	FPGA	33.7	6,942	19.1	17.3	4.7	4.0	22.0
航锦科技	000818 CH	FPGA	34.06	3,413	n.a.	n.a.	n.a.	n.a.	n.a.
兆易创新	603986 CH	DRAM	118.08	11,623	28.5	25.4	4.5	3.9	15.3
<b>数据中心</b>									
易昆尼克斯	EQIX US	IDC	730.99	67,644	79.4	67.7	6.1	6.2	8.0
万国数据	GDS US	IDC	25.03	4,770	n.a.	n.a.	1.6	1.7	(6.4)
秦淮数据	CD US	IDC	8.84	3,243	26.1	21.0	2.0	1.8	9.7

资料来源: 彭博、招银国际环球市场

# 目录

大模型产业链及潜在受益标的.....	2
受益标的估值比较.....	3
追本溯源：ChatGPT 的概述及现状.....	5
什么是 ChatGPT?.....	5
ChatGPT 商业化落地的限制?.....	6
从小模型到大模型的兴起.....	7
自然语言处理 (NLP) 大模型.....	8
计算机视觉 (CV) 大模型.....	8
多模态模型.....	9
算力层：大模型驱动 AI/内存芯片需求全面加速.....	10
AI 芯片.....	10
内存芯片.....	14
应用层：AIGC 万亿市场可期，大模型+内容生态为关键.....	15
场景应用：搜索+智能客服先行，内容创作+文娱生活加速落地，长期 B 端产业升级可期.....	15
变现空间：AIGC 万亿市场可期，订阅付费加速商业化，MaaS 探索长期增长点.....	16
中美对比：美国底层技术领先，中国内容场景多点开花.....	18
中美引领全球预训练大模型发展，美国底层技术领先.....	18
技术面：数据及大模型参数量差距较小，算力和模型迭代仍需精进.....	19
应用面：美国商业化加速、赋能 2B；中国变现早期，侧重生活互娱 C 端场景.....	20
潜在受益标的.....	22
微软 (MSFT US) - 深度融合 OpenAI 工具，技术和应用并行.....	23
谷歌 (GOOG US) - 发布“Bard AI”及开放生成语言 API，LaMDA 大模型积淀深厚.....	24
百度 (9888 HK) - 全栈式 AI 技术先锋，文心一言+数字人+AI 创作平台矩阵丰富.....	25
腾讯 (700 HK) - 混元 AI 大模型赋能多业务场景落地、实现技术提效.....	27
阿里巴巴 (BABA US) - 通义大模型凸显开源，构建 AI 统一底座.....	29
商汤 (20 HK) - 拥有 300 亿参数视觉大模型.....	31
科大讯飞 (002230 CH) - AI 语音市场龙头.....	32
思必驰 (未上市) - “云+芯”渗透 AI 语音语义市场.....	33
英伟达 (NVIDIA, NVDA US) - 全球 AI 计算平台龙头.....	34
超威半导体 (AMD, AMD US) - CPU+GPU 双芯片行业龙头.....	34
莱迪思 (Lattice, LSCC US) - 全球领先 FPGA 供应商.....	35
景嘉微 (300474 CH) - 国产 GPU 龙头企业.....	35
寒武纪 (688256 CH) - 国内 AI 芯片领先者.....	36
海光信息 (688041 CH) - 中国微处理器领军者.....	37
龙芯中科 (688047 CH) - 国产 CPU 龙头企业.....	37
紫光国微 (002049 CH) - 特种集成电路龙头.....	38
航锦科技 (000818 CH) - 军工/电子/化工三大业务板块驱动成长.....	39
安路科技 (688107 CH) - 国内领先的 FPGA 芯片供应商.....	40
复旦微电 (688385 CH) - FPGA 平台型公司.....	40
兆易创新 (603986 CH) - DRAM 龙头企业，自有品牌广泛应用.....	41

## 追本溯源：ChatGPT 的概述及现状

### 什么是 ChatGPT?

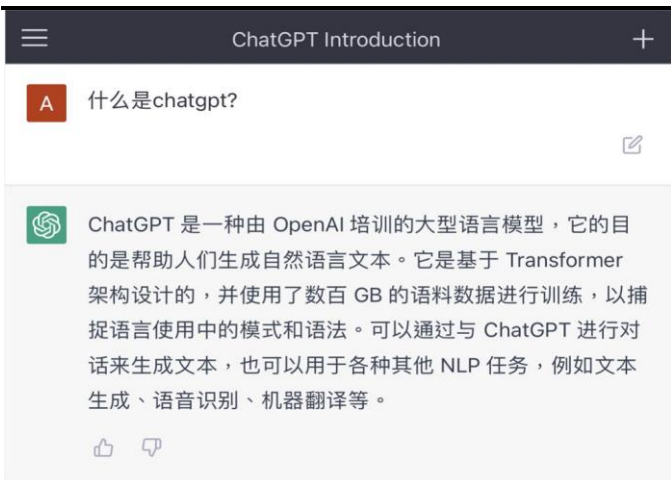
ChatGPT 是 OpenAI 于 2022 年 11 月推出的一款聊天机器人 (AI Chatbot)，它是基于 GPT-3.5 (GPT-3 的改进版) 模型的变体。ChatGPT 可以在对话中根据上下文形成类似人类的文本响应，与其他使用预定义的响应或规则生成文本的聊天机器人不同，ChatGPT 可以根据接收到的输入生成响应，从而生成更自然、更多样化的响应。ChatGPT 不仅是聊天机器人，还能进行撰写邮件、文案、代码等任务。从 2022 年 11 月上线以来，在 ChatGPT 上的累积用户已经达到 1 亿。

图 4: OpenAI 及 ChatGPT 的发展历程

年份	大事记
2015	<ul style="list-style-type: none"> <li>作为 AI 研究实验室，OpenAI 在美国正式成立</li> <li>Y Combinator 的 Sam Altman 和 Elon Musk 作为 OpenAI 的首届联合主席</li> </ul>
2018	<ul style="list-style-type: none"> <li>Elon Musk 离开 OpenAI，Sam Altman 成为了 OpenAI 的 CEO</li> <li>发布通用 NLP 模型 GPT，含 1.17 亿参数</li> </ul>
2019	<ul style="list-style-type: none"> <li>微软 (MSFT US) 向 OpenAI 投资 10 亿美元</li> <li>发布 GPT-2，含 15 亿参数</li> </ul>
2020	<ul style="list-style-type: none"> <li>OpenAI 发布 GPT 3，含 1,750 亿参数</li> <li>微软取得 GPT-3 的独家授权</li> </ul>
2021	<ul style="list-style-type: none"> <li>发布 CLIP，匹配图像和文本的预训练大模型</li> <li>发布 DALL-E，文字生成图像 AI 系统</li> <li>发布 DALL-E 2</li> </ul>
2022	<ul style="list-style-type: none"> <li>推出 ChatGPT 聊天机器人，5 天内注册用户超过 100 万</li> <li>微软宣布旗下所有产品将全线整合 ChatGPT，除了搜索引擎 Bing 和 Office 外，微软将在 Azure 云平台整合 ChatGPT，Azure 的 OpenAI 服务将允许开发者访问 AI 模型</li> </ul>
2023	<ul style="list-style-type: none"> <li>微软宣布将向 OpenAI 追加投资数十亿美元</li> </ul>

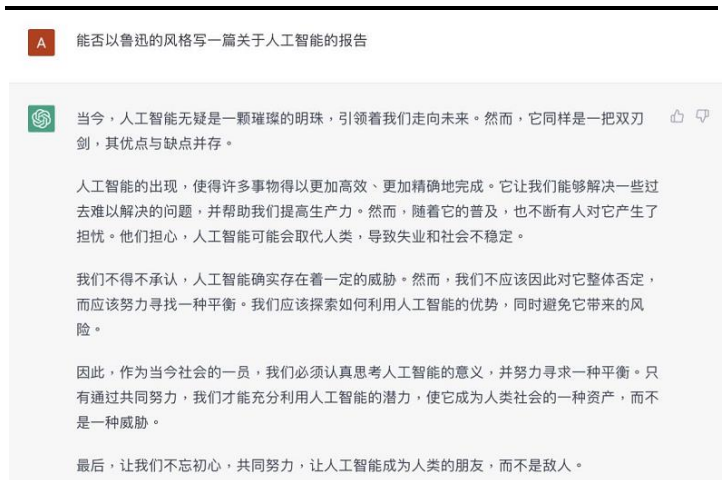
资料来源：OpenAI，微软，招银国际环球市场

图 5: ChatGPT 使用示意



资料来源：ChatGPT，招银国际环球市场

图 6: ChatGPT 使用示意



资料来源：ChatGPT，招银国际环球市场

## ChatGPT 商业化落地的限制？

虽然 ChatGPT 的应用空间广泛，但短期内它仍面对几个关键技术瓶颈。

### 1) 回答的真实性

ChatGPT 可能会写出看似合理但不正确的答案。这主要是因为模型进行强化训练 (Reinforced Learning) 时，缺乏真实信息来源。另外，为了提高模型的严谨性，模型会拒绝回复部分它能正确回答的问题。同时，监督训练会误导模型，由于模型产生的答案是基于模型的认知而不是人类标注师的认知。

### 2) 数据量级和算力门槛影响模型的实时性

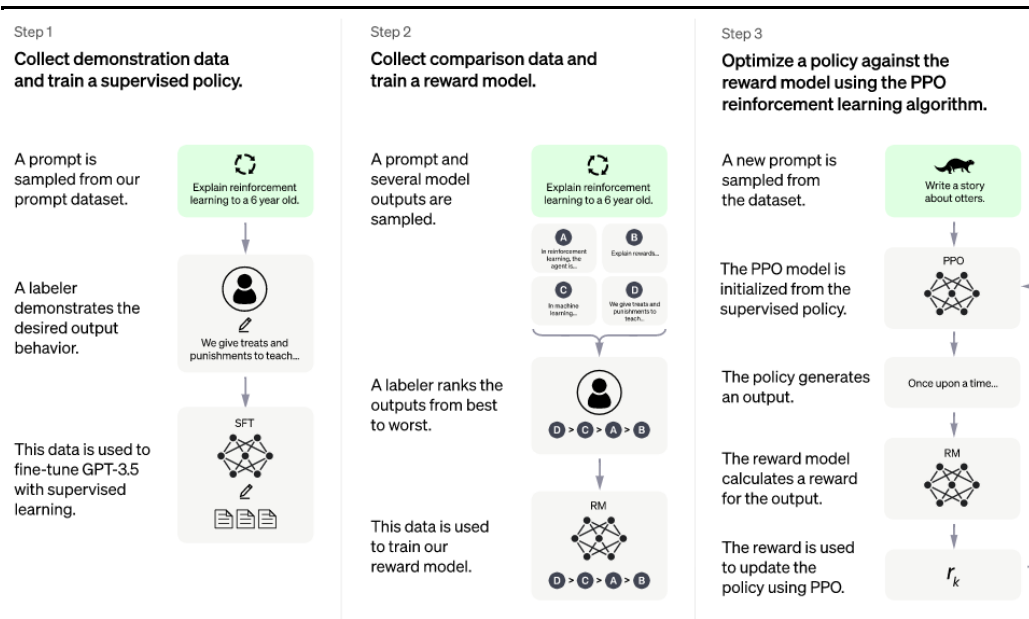
ChatGPT 学习的数据库停留在 2021 年，由于模型的数据量级大 (GPT-3 1750 亿参数，使用 800GB 数据进行训练)，加入标注数据训练让实时数据的引入非常困难，重新预训练模型的成本也高 (单次几百万美元以上)，导致 ChatGPT 对于 2021 年后发生的事情的了解有限。

### 3) 回答内容存在偏见或有害指令

ChatGPT 使用了“人类反馈强化学习” (RLHF) 的训练方法。首先，预训练的模型在少量已标注的数据上进行调优，随后让预训练好的模型 (SFT) 针对新问题列表生成若干条回答，并让人类标注师对这些回答进行投票，以创建一个由比较数据组成的新数据集。在此数据集上训练新模型 (Reward Model, RM)。RM 模型用于进一步调优和改进 SFT，并将最后两个步骤反复迭代以获得最终的模型。

这个方法有明显的局限性，用于 fine-tuning 模型的数据会受到人工标注者的偏好和偏差影响，标注者通常对模型输出的排名持不同意见，RLHF 的方法假设所有人都有相同的价值观，这导致 ChatGPT 导出的内容可能存在偏见。

图 7: ChatGPT 基于人类反馈对语言模型进行强化学习



资料来源: OpenAI, 招银国际环球市场

## ChatGPT 的底层技术发展演进

### 从小模型到大模型的兴起

AI 模型最初是针对特定应用场景需求进行训练（即小模型）。小模型的通用性差，换到另一个应用场景中可能并不适用，需要重新训练，这牵涉到很多调参、调优的工作及成本。同时，由于模型训练需要大规模的标注数据，在某些应用场景的数据量少，训练出来的模型精度不理想的情况，这使得 AI 研发成本高，效率低的情况。

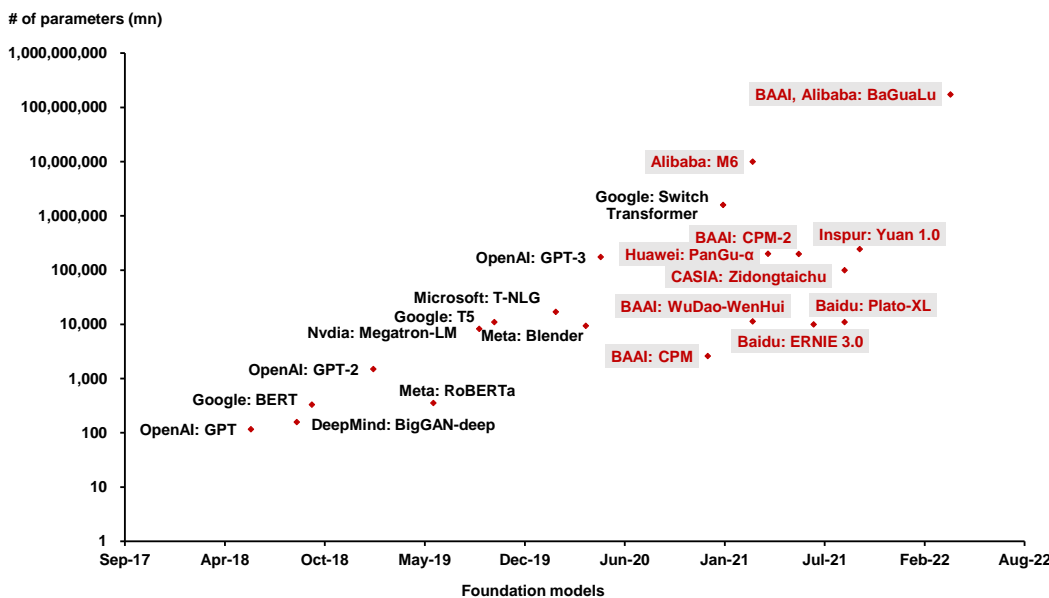
过去 5 年，随着数据，算力及算法的提升，AI 技术也有了变化，从过去的小模型到大模型的兴起。大模型（也称为基础模型 Foundation Models）是在大规模无标注数据上进行训练，学习出特征和规则。大模型泛化能力强，美国 OpenAI, 谷歌，微软，Facebook 的机构过去几年发布了千亿或万亿参数量级的大模型，这些大模型主要是在自然语言领域（NLP），包括 ChatGPT 背后的 GPT-3, Switch Transformer, BERT 等。

基于大模型进行应用开发，将大模型进行微调（在下游特定任务上的小规模有标注数据进行二次训练），或者不进行微调，可以完成多个应用场景的任务。这有效降低 AI 应用研发门槛，也能解决在应用场景数据量少，模型精度低的问题。

在图 8，我们列出了过去 5 年中美主要大模型及其参数量的规模，在发布时间的差距上，我们可以看到中国的大模型是集中在 2021 年到 2022 年之间，而美国同等参数量级的模型比中国领先 1-2 年。另外，我们看到模型的规模也呈现指数级扩张。

但需要注意的是，模型的参数规模不一定是越大越好，大模型的可解析性和可控性比较薄弱，复杂的模型使他们比普通的神经网络更难以理解，在检测和减轻这些 blackbox 模型的偏差会变得更加困难。

图 8：国内外企业发布的超大深度学习模型



资料来源：Github，招银国际环球市场

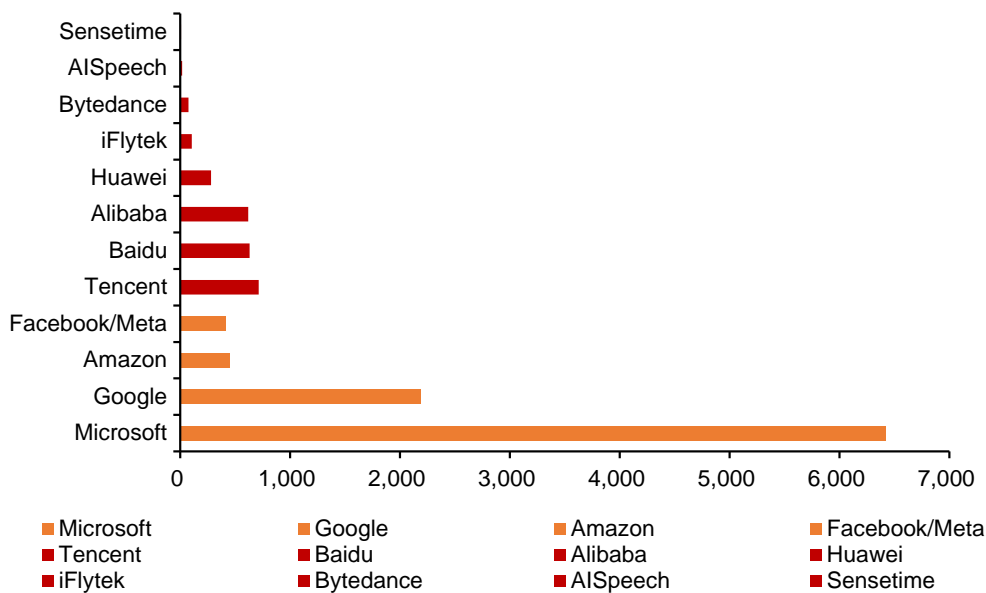
## 自然语言处理 (NLP) 大模型

自然语言处理 (NLP) 主要通过学习通用语言，使得模型具备语言理解和生成能力。在 AI 的感知层 (识别能力)，目前机器在语音识别 (Speech Recognition) 的水平基本达到甚至超过了人类的水平。然而，机器在处理自然语言时还是非常困难，主要是因为自然语言具有高度的抽象性，语义组合性，理解语言需要背景知识和推理能力。

在 2018 年以来，以 BERT 和 GPT 为代表的语言大模型，弥补了自然语言处理标注数据的缺点，促进了 NLP 技术的发展。从技术的角度，这些大模型的训练，透过事先遮住一些文本片段，让 AI 模型通过自监督学习，通过海量语料库的预训练，逐步掌握上下文语境，把这些被遮住的片段，尽可能合乎逻辑的方式填上去。

现阶段只有微软和谷歌正式发布了基于 NLP 大模型的可交互式应用，暂时没有办法评判每家公司大模型在实际产品中的表现。我们尝试比对国内及国外比较领先的 AI 公司在 NLP 领域的专利数，根据 WIPO 国际专利分类，我们参考了 G06F17/20 项 (Handling natural language data)，国内比较领先的是互联网企业，前三分别是腾讯，百度和阿里巴巴 (专利数在 600 个以上)，而专注于语音方面的 AI 企业科大讯飞只有 100 多个专利。国内 AI 企业和国际巨头相比，在 NLP 领域的专利数量差距还是比较明显，像微软及谷歌等，他们在 NLP 的专利数已经在几千个的量级。

图 9: NLP 领域专利数量比较



资料来源: WIPO, 招银国际环球市场 \*比较 G06F17/20 项

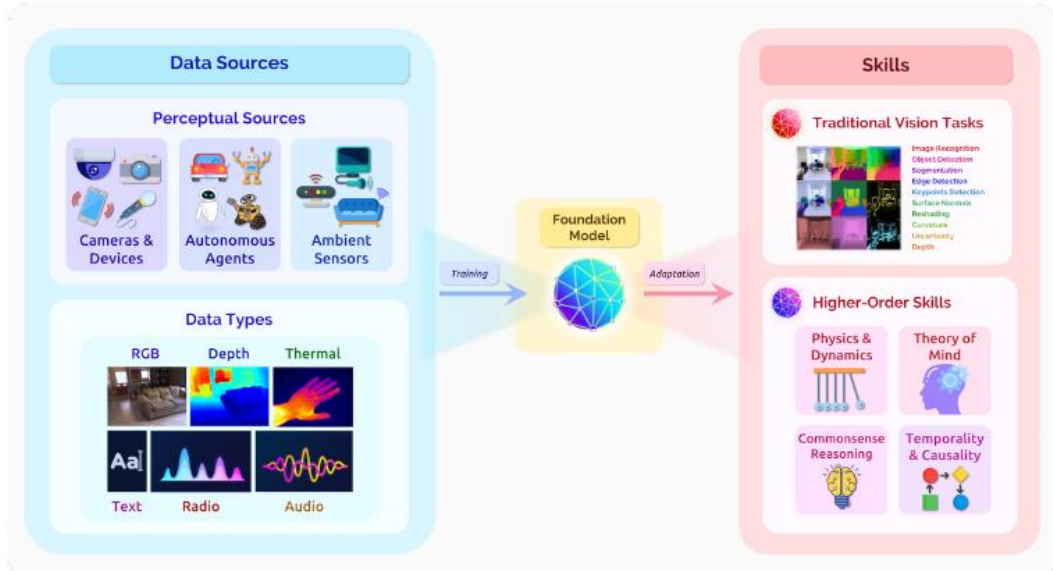
## 计算机视觉 (CV) 大模型

在早期的计算机领域 (CV)，模型的训练需要透过提取特征工程。视觉大模型是通过学习大量的图像和视频数据，形成具备视觉通用能力。相较于上千亿参数的 NLP 模型相比，CV 模型规模要小两三个数量级，目前 CV 大模型的发展还是存在一些挑战，包括 1) 在 CV 领域可用于训练的有效数据比 NLP 领域有不少差距，2) CV 领域的学习方法还需要突破，3) 不同的视觉应用仍需要依赖于不同的模型，如何建立通用的视觉模型还是未知，4) 供训练的图像尺寸越来越大，较小的模型也可能有很大的计算量。



现阶段比较有名的 CV 大模型包括微软的 swin-transformer 系列，谷歌的 ViT 系列以及有 150 亿参量的 V-MOE 模型。国内方面，截止 2021 年，商汤训练了 300 亿参数的 CV 模型，2022 年，华为也发布了 30 亿参数的盘古系列 CV 模型。

图 10: CV 大模型



资料来源：“On the Opportunities and Risks of Foundation Models”，招银国际环球市场

### 多模态模型

多模态机器学习 (MultiModal Machine Learning, MMML) 是指通过机器学习的方法实现处理，理解及融合多源模态信息的能力，模态是指某种类型的信息，例如文本、图像、视频、音频等，目前比较热门的研究方向是文本-图像的模型及应用，这里面比较有名的是 OpenAI 发布的 CLIP 和 DALL·E。

CLIP (Constrastive Language-Image Pre-training) 是一个基于对比图片-文本学习的跨模态预训练模型，通过自然语言的监督学习来有效理解视觉概念，用户只需要提供要识别的视觉类别标签，CLIP 就能完成视觉分类任务。

基于 CLIP 模型技术，OpenAI 在 2022 发布 DALL·E 2，它是一个 AI 程序，可以通过文本描述生成图像。例如，当文本提示为“生成牛油果形状的躺椅”，DALL·E 可以理解牛油果及躺椅的概念，并生成各种牛油果形状的，绿色的躺椅。在输出图像是，DALL·E 采用 CLIP 进行排序，从中选取最优结果。

图 11: DALL·E 演示



资料来源：OpenAI，招银国际环球市场

## 算力层：大模型驱动 AI/内存芯片需求全面加速

ChatGPT 运行需具备三大条件：数据上，获取训练数据，然后进行清洗和分离，获得不同用于训练和测试调优的数据集；算法上，在基础模型上大规模预训练需要大量算力，获取基于模型调优后不断解锁和进化；算力是核心基础和关键所在，数据获取和模型调优等需大量算力支持；而大算力需更多高端、通用或专用算力芯片和 AI 芯片去支撑算法运行。

随着 AIGC 典型应用 ChatGPT 发展推动企业和个人服务多模型和技术向 AI 方向构建，同时还将驱动底层领域包括半导体、HPC 和云计算等的协同发展，预计未来对 AI 和内存芯片需求将快速提升。我们建议关注 AI 芯片相关标的：1) GPU：海外包括英特尔、英伟达和 AMD；国内包括景嘉微、寒武纪、海光信息、龙芯中科和航锦科技；2) FPGA：海外包括英特尔、AMD；国内包括紫光国微、复旦微电和安路科技；3) ASIC：海外包括英伟达、谷歌和英特尔；国内包括百度和寒武纪。内存芯片：建议关注海外标的包括三星、美光和 SK 海力士；国内关注兆易创新。

### AI 芯片

AI 芯片能为人工智能应用提供所需的基础算力；按技术架构主要分为 GPU、FPGA 和 ASIC。ChatGPT 有着大量复杂计算需求的 AI 模型，AI 芯片专门用于处理人工智能应用中的大量计算任务，是不可或缺的底层硬件。随着 AI 技术快速发展，AI 应用场景扩大至智能制造、家居、教育、医疗和金融等领域；而智能产品种类不断丰富，包括 ChatGPT 等 AIGC 产品持续升级对芯片算力提出更高要求，AI 芯片算力和需求旺盛增长。据 Frost&Sullivan 预计，2022-2026 年全球 AI 芯片市场规模将以 26.3% 复合增速增长。

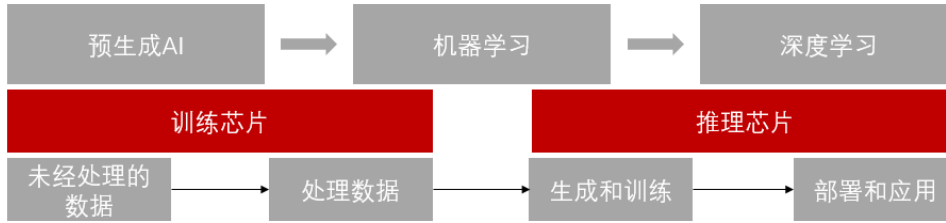
从基础预生成 AI 到机器学习，最终达到深度学习，主要依赖于审阅庞大的数据集、处理数据、生成算法模型、训练算法模型的能力，而 AI 芯片就是从处理数据到终端赋能的算力基础。AI 芯片主要分为三种，分别是 GPU, FPGA 与 ASIC 芯片。不同类别的 AI 计算芯片有各自突出的优势和适用的领域，贯穿 AI 训练与推理阶段。目前 CPU 在人工智能领域中的应用有限，主要受限于 CPU 在 AI 训练方面的计算能力不足。

图 12：不同类型的人工智能计算芯片

	CPU	GPU	FPGA	ASIC
<b>定义</b>	中央处理器	图像处理器	现场可编程逻辑门阵列	专用处理器
<b>算力与能效</b>	算力最低、能效比差	算力高、能效比中等	算力中，能效比优秀	算力高，能效比优秀
<b>上市速度</b>	快，产品成熟	快，产品成熟	快	慢、开发周期长
<b>成本</b>	用于数据处理时，单价成本最高	用于数据处理时，单价成本高	较低的试错成本	成本高，可复制，量产规模生产后成本可有效降低
<b>性能</b>	最通用	数据处理通用性强	数据处理能力较强、专用	AI 算力最强，最专用
<b>适用场景</b>	广泛应用于各个领域	广泛应用于各种图形处理，数值模拟，机器学习算法领域	适用成本要求较低的场景，如军事、实验室、科研等	主要满足场景单一的消费电子等高算力需求领域
<b>芯片公司</b>	英伟达 英特尔 Xeon 谷歌 AMD Ryzen 莱迪思	Telsa, Titan, Xavier Gaudi, Greco Addebaran	Arria, Stratix Versal, Alveo Lower Power	NNP, NCS2 Cloud TPU, Edge TPU

资料来源：公司，招银国际环球市场

图 13: 人工智能学习与涉及芯片类型



资料来源: OpenAI, 招银国际环球市场

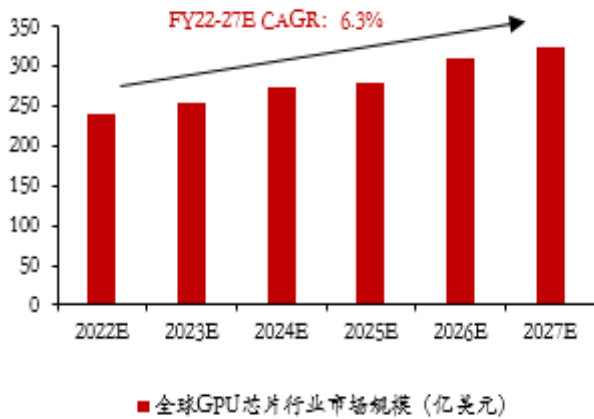
1) GPU (图形处理器)

GPU 作为显示芯片是可在个人电脑、工作站、游戏机和智能手机/平板电脑上做图像和图像相关运算工作的微处理器。随着互联网用户和网络应用兴起, 数据体量逐渐庞大, 数据中心对计算需求快速增加, GPU 凭借其相对通用灵活和适应并行计算等特性成为主要选择, GPU 也从最初专用于图像处理制作后逐渐应用于计算, 是核心计算资源基础。

目前来看, GPU 仍然是 AI 应用的主导芯片, 主要因为它具有强大的计算能力和高通用适用性, 广泛应用于各种图形处理、数值模拟及人工智能算法领域。较低的开发成本也时期能够快速在各个垂直下游领域被广泛应用, 加速优化拓展。

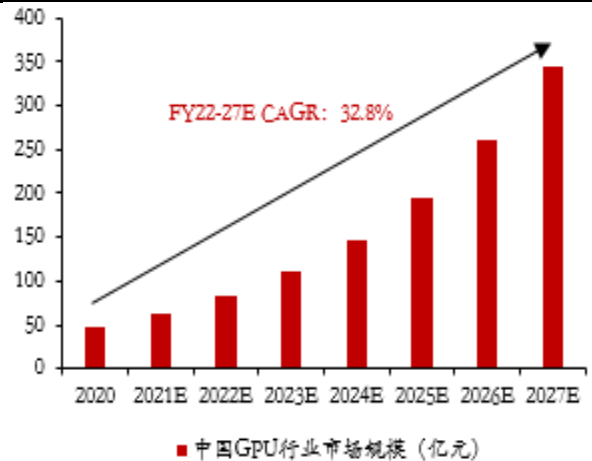
ChatGPT 作为有着大量复杂计算需求的 AI 模型, GPU 芯片成为能帮助其完成计算任务的关键硬件之一; 随着深度学习技术发展, GPU 成为 AI 算法训练中重要硬件组成部分。同时, 随着云计算技术发展, 云端对 GPU 算力要求在不断提升。据机构 JPR 预测显示, 2022-2026 年全球 GPU 出货量将以 6.3% 复合增速增长; 预计 2027 年行业市场规模将超 320 亿美元。

图 14: 全球 GPU 芯片行业市场规模预测



资料来源: JPR, 前瞻产业研究院, 招银国际环球市场预测

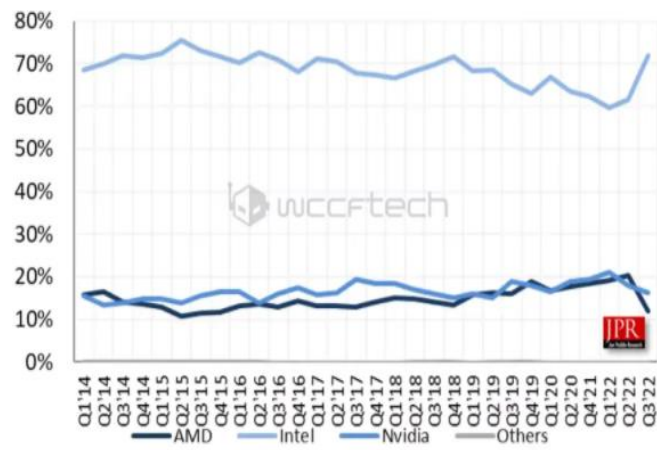
图 15: 中国 GPU 芯片行业市场规模预测



资料来源: Huaon, 招银国际环球市场预测

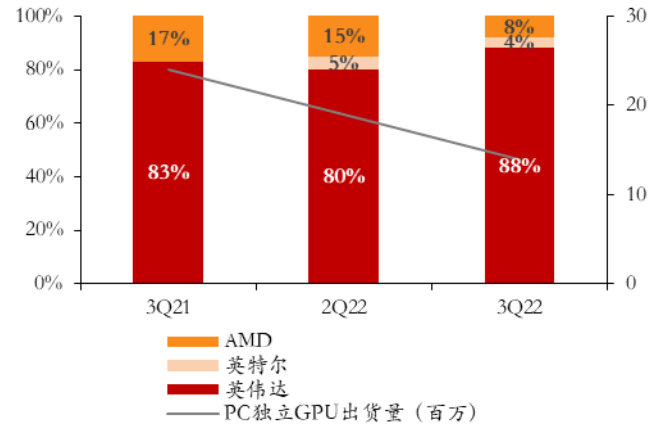
全球 GPU 芯片行业竞争格局中英特尔占主导地位, 但在独立 GPU 领域, 英伟达高端 GPU 占据较大份额 (超过 70%)。英伟达 GPU 因其高性能和对 CUDA 的支持而广泛应用于 AI 和机器学习领域 (CUDA 是 NVIDIA 的并行计算平台和编程模型)。因此, OpenAI 主要使用 NVIDIA GPU (图形处理单元) 来训练和运行其 AI 模型, 如 GPT-3 和 ChatGPT。据公开数据显示, ChatGPT 已导入至少 1 万个英伟达高端 GPU。

图 16: 1Q14-3Q22 全球电脑显卡市场份额变化



资料来源: JPR, 招银国际环球市场预测

图 17: 全球独立 GPU 市场份额及出货量



资料来源: JPR, 招银国际环球市场预测

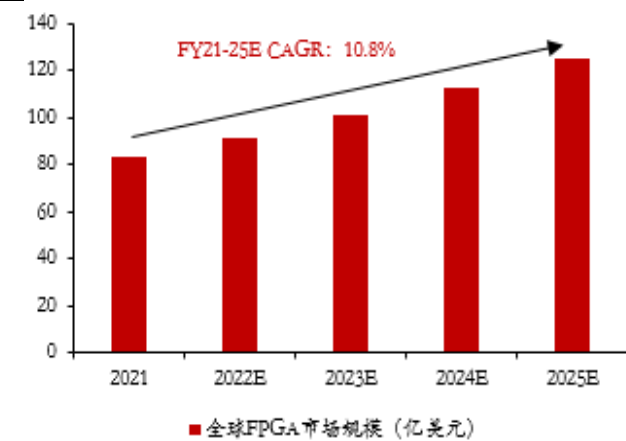
国内 GPU 主要上市公司中, 景嘉微通过自主研发 GPU 并规模化商用, 是国内显示控制市场龙头企业。此外, 海光信息、寒武纪等公司均有布局 GPU 产品。受美国对中国 GPU 芯片出口管制影响, 预计国产 GPU 芯片替代将快速推进。

**2) FPGA (现场可编程门阵列)**

FPGA 是支持 AI 算力芯片的更优解, 优点主要在于低延迟、优秀的能效比、较短的开发周期以及高度可编程能力, 可以供给不同垂直行业使用。FPGA 是一种硬件可重构的集成电路芯片, 通过在硅片上预先设计实现具有可编程特性, 可通过软件重新配置芯片内部的资源来实现不同功能, 广泛应用于数据中心、航空航天工程、人工智能、工业、物联网以及汽车等领域。FPGA 拥有软件的可编程性和灵活性、兼具硬件的并行性和低延时性, 相较于 ASIC、GPU 等处理器更为灵活, 具备上市周期和成本优势。在 5G 通信、人工智能等迭代升级周期频繁、技术不确定性较大的领域, FPGA 是较为理想的解决方案。

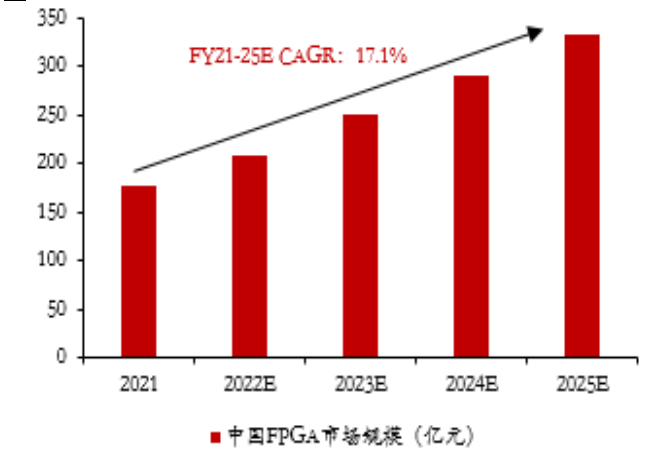
随着全球新一代通信设备部署以及人工智能与自动驾驶技术等新兴市场领域需求的不断增长, 前瞻产业研究院预计全球 FPGA 市场规模将从 2021 年的 82.9 亿美元增长至 2025 年的 125.2 亿美元, 年均复合增长率约为 10.8%, 中国 FPGA 市场规模快速增长, 预计 2021-2025 年规模复合增速 17.1%。

图 18: 全球 FPGA 芯片市场规模预测



资料来源: 前瞻产业研究院, 招银国际环球市场预测

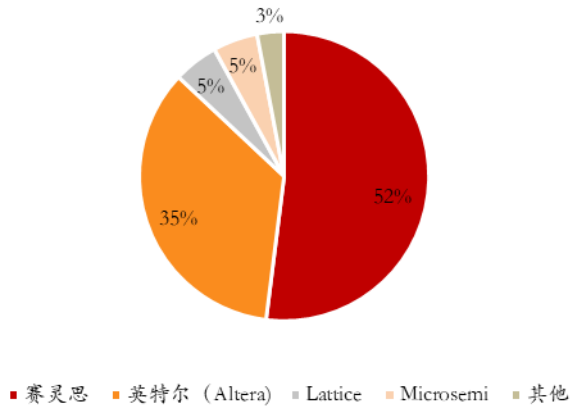
图 19: 中国 FPGA 芯片市场规模预测



资料来源: Huaon, 招银国际环球市场预测

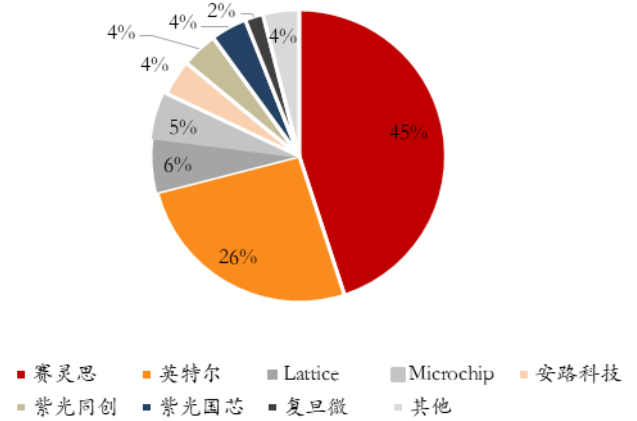
全球 FPGA 市场主要被赛灵思和 Altera 占据，目前市占率分别为 52%和 35%；其次为 Lattice 和 Microsemi，份额均为 5%。中国 FPGA 厂商中紫光国微、复旦微电和安路科技在 2021 年市占率超过 15%。受益于国产化加速推进，中国 FPGA 厂商将拥有巨大成长空间。

图 20: 2021 年全球 FPGA 市场竞争格局



资料来源: ASKCI, 招银国际环球市场预测

图 21: 2021 中国 FPGA 市场竞争格局



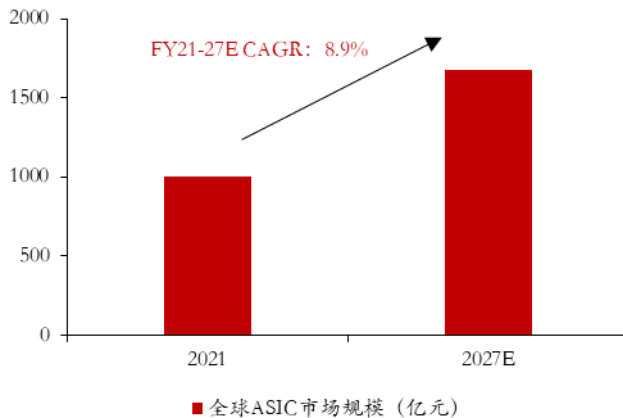
资料来源: 安路科技公司公告, 招银国际环球市场预测

### 3) ASIC (专用集成电路)

ASIC 是为特定用途而定制集成电路，具有高性能、低能耗的特点，专用化程度最高，其特点同时适合 AI 训练和推理阶段的使用。ASIC 芯片包括 DPU 和 NPU 类别，其中，DPU 主要承担网络、存储和安全的加速处理任务，旨在满足网络侧专用计算需求，适用于服务器量多、对数据传输速率要求严苛的场景。ASIC 应用领域广泛，包括医疗、工业、军事、航空等。其中虚拟币是主要应用领域之一，因相对 GPU、FPGA，ASIC 能效更高。

目前 ASIC 渗透率依然较低，主要由于其较高的研发成本与研发壁垒，因此市场份额显著小于 GPU。随着人工智能行业愈加成熟，蓬勃发展的下游应用对于上游研发成本的优化力不断增高，ASIC 芯片依赖于它的卓越性能及低功耗，将成为人工智能技术的首选。根据贝哲斯咨询预测，到 2027 年，全球 ASIC 市场规模预计将达 1677.49 亿元，2021-2027 年均复合增长率为 8.9%。

图 22: 全球 ASIC 芯片市场规模预测



资料来源: 贝哲咨询, 招银国际环球市场预测

全球 ASIC 主要竞争者包括英伟达、谷歌和英特尔。其中：1) 英伟达：在 2020 年推出 BlueField-2 DPU，将其定义为继 CPU 和 GPU 之后“第三颗主力芯片”；2) 谷歌：在 I/O 2022 活动中发布其新一代张量处理器 TPU v4 集群，新的算力集群被称为 Pod，包含 4096 个 v4 芯片，可提供超过 1 exaflops 的浮点性能，公司将在位于俄克拉荷马州的数据中心部署 8 个 TPU v4 集群，合计实现约 9 exaflops 的性能；3) 英特尔：在英特尔 FPGA 技术大会上，英特尔发布了全新可定制解决方案英特尔 eASIC N5X，帮助加速 5G、人工智能、云端与边缘工作负载的应用性能。英特尔 eASIC N5X 通过 FPGA 中的嵌入式硬件处理器帮助客户将定制逻辑与设计迁移到结构化 ASIC 中，带来更低的单位成本，更快性能和更低功耗等好处。目前英伟达的定制化 ASIC 芯片价格为 2.48 美元/小时，而最初的设计者谷歌的定制化 ASIC 芯片价格为 8 美元/小时。

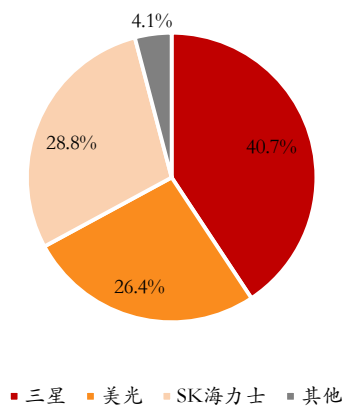
国内主流的 ASIC 上市公司包括百度和寒武纪：1) 百度 (9888 HK)：百度从 2011 年起布局 AI 加速领域，昆仑芯团队深耕 FPGA 到 ASIC 芯片已 10 余年，在体系结构、芯片实现、软件系统和场景应用均有深厚积累。2) 寒武纪 (688256 CH)：以深耕 ASIC 为主，其思元 270-S4 型号芯片专注于为数据中心提供服务。

## 内存芯片

高性能计算芯片需要搭配的 DRAM 内存数量大幅增加，DRAM 特征是读写速度快、延迟低，常用于计算系统的运行内存。DRAM 集成度高、功耗低、存取速度慢。从存储能力来看，DRAM 所能提供的存储容量更大，访问时间较长。随着 ChatGPT 采用高性能计算芯片提升算力，我们认为需要搭配的 DRAM 内存数量将大幅增加。

全球 DRAM 市场高度集中，三星、美光、SK 海力士 DRAM 出货量在 3Q22 合计占比约 96%。受行业周期性低迷的影响，三星电子、SK 海力士和美光正扎堆减产、应对库存问题、节约资本开支，并推迟先进技术的进展，以应对存储器需求的疲软态势。据集邦咨询研究，由于消费需求疲弱，存储器卖方库存压力持续，仅三星(Samsung)在竞价策略下库存略降。

图 23: 3Q22 全球 DRAM 出货量市占率



资料来源：Omdia, 招银国际环球市场预测

兆易创新 (603986 CH) 为国内 DRAM 龙头企业，与长鑫密切合作，推动自有品牌 DRAM 多行业应用，实现了从设计、流片，到封测、验证的全国产化，助力国产自主供应生态圈的发展构建，该产品主要面向消费类、工业控制类及汽车类等市场领域。公司与长鑫存储的紧密合作关系，为公司 DRAM 产品提供稳定产能保障。同时，依托于多年积累的、完善的销售网络和技术团队，公司能够为客户提供快速的本地化服务响应和技术支持。

## 应用层：AIGC 万亿市场可期，大模型+内容生态为关键

进入“数智化时代”，看好 ChatGPT/AIGC 应用场景持续渗透及商业化加速 (包括搜索引擎、智能客服、教育培训、金融等)，助力行业降本增效的同时，开拓创新变现增量。

### 场景应用：搜索+智能客服先行，内容创作+文娱生活加速落地，长期 B 端产业升级可期

我们看好 ChatGPT 在搜索引擎的协同和能力加强，多项回复展示更好满足用户信息获取，智能创作、编程、分析等能力助于提高搜索转化率。考虑当前 ChatGPT 信息基于离线非实时数据库、训练模型下答案非正确、成本过高，我们预计 ChatGPT 较难取代搜索引擎，更多体现在赋能优化。电商领域，ChatGPT 有望在智能客服领域通过大模型通用智能能力完善人机对话丝滑度和专业答疑。期待 ChatGPT 未来在代码机器人、翻译、文学创作、媒体、教育培训、家庭陪护、金融等领域加速落地。文字之外，AIGC 也将在内容创作、数字经济、产业升级层面带来变革，中期电商、传媒、娱乐、影视等领域加速渗透机会显现，长期看好 AIGC 推动 B 段医疗、工业、金融等产业升级 (例如：合成数据)。

图 24：AIGC 应用场景

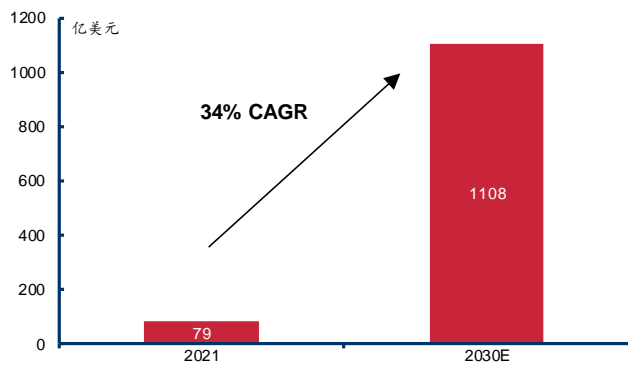
AIGC 应用场景：前期内容创作+文娱生活场景导向，长期看好 B 端产业机会					
娱乐	传媒	影视	电商	B 端产业机会	
<b>全民娱乐</b> <ul style="list-style-type: none"> <li>人脸美妆/融合</li> <li>黑白图像上色</li> <li>图像风格转换</li> <li>人像属性变换</li> </ul>	<b>采集</b> <ul style="list-style-type: none"> <li>采访助手</li> </ul>	<b>前期制作</b> <ul style="list-style-type: none"> <li>人脸美妆/融合</li> <li>黑白图像上色</li> <li>图像风格转换</li> <li>人像属性变换</li> </ul>	<b>商品展示</b> <ul style="list-style-type: none"> <li>智能商详</li> <li>商品 3D 模型生成</li> </ul>	<b>教育</b> <ul style="list-style-type: none"> <li>合成历史人物</li> <li>视频虚拟教师</li> <li>线上课堂音视频处理</li> </ul>	<b>金融</b> <ul style="list-style-type: none"> <li>营销视频合成</li> <li>智能金融客服</li> <li>VR 金融场景构建</li> </ul>
<b>社交互动</b> <ul style="list-style-type: none"> <li>虚拟歌姬</li> <li>虚拟网红</li> </ul>	<b>编辑</b> <ul style="list-style-type: none"> <li>写稿机器人</li> <li>智能视频拆条视频锦集</li> <li>视频字幕生成</li> </ul>	<b>中期拍摄</b> <ul style="list-style-type: none"> <li>数字复活已故演员</li> <li>高难度动作合成</li> <li>演员角色年龄跨越</li> <li>虚拟物理场景合成</li> </ul>	<b>主播打造</b> <ul style="list-style-type: none"> <li>品牌虚拟主播</li> </ul>	<b>工业</b> <ul style="list-style-type: none"> <li>辅助工程设计</li> <li>加速数字孪生系统构建</li> </ul>	<b>医疗</b> <ul style="list-style-type: none"> <li>医学图像处理</li> <li>智能病例录入</li> <li>合成肢体投影</li> <li>合成医护陪伴</li> </ul>
<b>偶像养成</b> <ul style="list-style-type: none"> <li>用户数字化身</li> </ul>	<b>播报</b> <ul style="list-style-type: none"> <li>AI 主播</li> <li>智能播报</li> </ul>	<b>后期制作</b> <ul style="list-style-type: none"> <li>替换“劣迹艺人”</li> <li>多语言译制片音画同步</li> <li>影视作品修复影视预告片生成</li> <li>影视内容 2D 转 3D</li> </ul>	<b>交易场景</b> <ul style="list-style-type: none"> <li>虚拟商城构建</li> </ul>		

资料来源：中国信息通信研究院，京东探索研究院，招银国际环球市场整理

## 变现空间: AIGC 万亿市场可期, 订阅付费加速商业化, MaaS 探索长期增长点

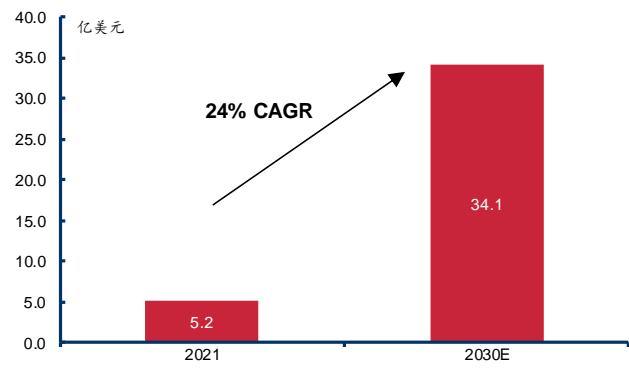
下一代互联网业态对内容产出效率、内容消费的蓬勃需求, 将驱动 AIGC 快速发展。结合 Gartner 及 Acumen Research and Consulting 预测, 2025 年 AIGC 渗透率将达 10%, 至 2030 年 AIGC 市场规模达 1108 亿美元(对应 2021-2030 年 CAGR 达 34%)。其中, 对话式 AI 预计 2030 年市场规模达 34 亿美元(vs. 2021 年 5.2 亿美元)。当前 AIGC 企业变现模式主要为: 订阅付费(每月订阅)、按量收费(数据请求量、计算量、图片张数、模型训练次数等), 未来 MaaS(模型即服务)有望复制 SaaS 企服路线迎来需求爆发奇点(海外 Hugging Face 已成型)。

图 25: 全球 AIGC 市场规模



资料来源: Acumen Research and Consulting, 招银国际环球市场

图 26: 对话式 AI 市场规模



资料来源: Acumen Research and Consulting, 招银国际环球市场

**受益标的:** 在 ChatGPT/AIGC 应用层面, 建议重点关注拥有海量数据、大模型能力及 AI 内容创作产品丰富(文/音/图/视)的头部平台, 包括 OpenAI、微软、谷歌、Meta、百度、腾讯、阿里、字节; 可同时关注近期推出类 ChatGPT 服务或泛 AI 产品的垂类腰部平台, 如网易、知乎、阅文、昆仑万维等。我们看好百度在 ChatGPU 及 AIGC 的全栈技术储备、搜索引擎协同及产品管线, 受益于: 1) 文心一言待推出, 数字人+智能创作应用矩阵丰富; 2) 文心大模型技术支持及海量数字资产优势(语料库、参数量大)。



图 27: ChatGPT/AIGC 相关受益标的 (应用及场景方向)

公司	股票代码	ChatGPT/AIGC 相关业务情况
微软	MSFT US	1) ChatGPT 聊天机器人; 2) 全新 Bing 搜索服务 (新增一个扩展聊天框, 在回答问题的同时帮助即时生成个性化分析和建议); 3) Edge 浏览器 (是新增聊天和内容创作两项功能)
谷歌	GOOGL US	1) Bard 聊天机器人 (基于 LaMDA 大模型); 2) AI 生成音频 (Magenta); 3) AI 生成图像 (Imagen, Parti, NeRF, NeRV); 4) AI 生成视频 (Imagen Video, Phenaki)
Meta	META US	1) AI 生成文字 (VX2Text); 2) AI 图文: Make a Scene; 3) AI 视频制作: Make A Video
百度	BIDU US	1) 深度学习开源框架, 飞桨 Paddle; 2) 全栈 AI 基础设施, 百度 AI 大底座; 3) AI 应用场景全覆盖大模型, 文心大模型 (NLP、CV、跨模态、生物计算及各行业大模型)。
阿里巴巴	BABA US	1) 基础设施: 通义大模型; 2) 类 ChatGPT 产品研发内测中; 3) AIGC 应用于智能客服、视觉生成、在线设计、在线试穿、数字人等
腾讯	700 HK	1) HunYuan-NLP-1T (混元 AI 大模型, 涵盖 NLP、CV、多模态、文生图等大模型); 2) AI 生成文字 (梦幻写手 Dreamwriter); 3) AI 生成游戏 (绝悟 AI); 4) AI 绘画 (腾讯 QQ 小世界 AI 画匠); 5) 数字人 (AI Lab 数字人 PaaS 方案, 创建的数字人“艾灵”) 等
网易	NTES US	1) 网易有道未来或将推出 ChatGPT 同源技术产品, 应用场景围绕在线教育; 此前已应用 AI 口语和作文批改。2) AI 生成声音: 有灵、网易天音; 3) 数字人
知乎	ZH US	作为中文最大的问答社区, 拥有海量可训练素材
阅文	772 HK	联合微软 (亚洲) 互联网工程学院开启 AI 赋能网络文学 (NLP)
美图	1357 HK	美图秀秀 AI 绘画功能 (CV)
万兴科技	300624 CH	Filmora、万兴喵影、Pixso (CV)
拓尔思	300229 CH	1) TRS 小思智能回答机器人; 2) 语义智能技术平台“智拓” (NLP)
昆仑万维	300418 CH	1) 多模态内容生成功能; 2) “昆仑天工”旗下模型 (包括天工巧绘 Sky Paint、天工乐府 Sky Music、天工妙笔 Sky Text、天工智码 SkyCode)
视觉中国	000681 CH	发售数轮 AIGC 数字藏品 (与百度文心一格达成战略合作协议, NLP、CV)
中文在线	300364 CH	1) AI 文字辅助创作功能; 2) AI 辅助绘画功能已推出; 3) “AI 主播”和复活 3D 模型虚拟数字人 (NLP、CV)
汉仪股份	301270 CH	字体生成, AI 造字
天娱数科	002354 CH	1) 多模态智能算法等; 2) 升级发布“Meta Surfing-元享智能云平台” (在原有功能基础上纳入 AIGC 功能模块)
风语筑	603466 CH	1) 虚拟展厅; 2) 虚拟数字人; 3) 虚拟场景以及数字艺术创作领域 (CV、数字人等)

资料来源: 公司资料, 招银国际环球市场整理

## 中美对比：美国底层技术领先，中国内容场景多点开花

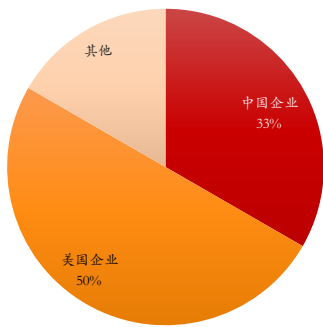
### 中美引领全球预训练大模型发展，美国底层技术领先

大模型底层创新始于美国：Google 在 2017 年首次提出 NLP 模型 Transformer 及 OpenAI 的 GPT 3.0；中国于 2021 年后加速追赶，重点发力 CV 大模型和多模态融合。

**当前中美两国引领全球预训练大模型发展。**据国内 OpenBMB 开源社区统计，目前，在全球超千亿参数大模型中，中国企业或机构贡献了 1/3，美国贡献了 1/2，世界其他国家或组织贡献了剩下的 1/6。截至 2022 年 10 月，在拥有大模型数量及参数量前十名的组织中，中/美分别占据 4/6 席。

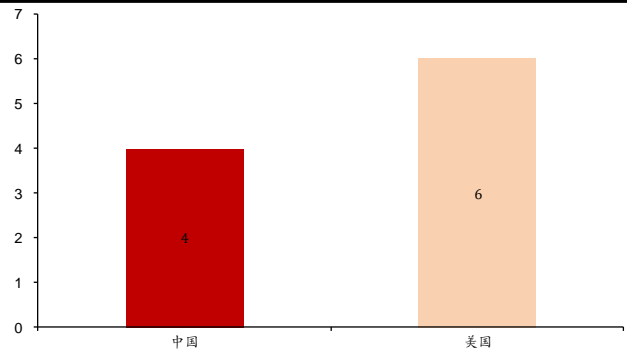
我们认为，目前美国 ChatGPT 的技术领先国内公司 2-3 年，其中 2020 年 OpenAI 发布 GPT-3.0 后中美差距有所拉大。中美技术差距并非由于中国缺乏大数据和参数量（百度文心大模型参数量达 2600 亿，vs. ChatGPT 1750 亿），而在于：1) 中国原创模型匮乏，算法深刻见解待提升；2) 2022 年 8 月美国限制部分高端 GPU 出口中国，一定程度影响模型训练速度。随着中国公司在算法 AI 的持续投入、巨头积极布局，底层技术处于跟随和追赶态势，部分算力约束通过国产 GPU/大规模分布式 CPU 替代。看好中国强大的基建能力、用户规模及大数据优势，有望依托大模型在 AIGC 应用落地、产业升级爆发巨大增长潜力。

图 28：全球千亿参数模型贡献比例



资料来源：OpenBMB，招银国际环球市场

图 29：全球大模型数量及参数量 TOP10 分布



资料来源：OpenBMB (2022.10)，招银国际环球市场

## 技术面：数据及大模型参数量差距较小，算力和模型迭代仍需精进

美国凭借在模型迭代和算力上的先发优势，聚焦大模型 API 的基础设施层开发，带动信息产业和科学行业发展，并加速商业化。中国公司在 AI 数据及大模型参数量层面与美国差距较小，但模型迭代和训练仍有较大提升空间，大部分平台更加聚焦 AIGC 内容和应用端。

美国主要大模型包括 OpenAI 的 GPT-3、Anthropic 的 Claude、英伟达与微软的 Megatron Turing-NLG 以及 Meta 的 OPT 等。其中，OpenAI 预计在今年年中或者下半年上线 GPT-4，参数量将达到万亿级。Google 把 GPT 系列模型作为红线，并全力投入 LLM（大语言模型）模型的构建。

中国大模型主要包括百度文心、华为盘古、阿里通义、腾讯混云以及商汤的视觉模型等。百度推出文心大模型，基本实现跨模态和多下游多领域的应用，且预计在今年 3 月上线文心一言（ERNIE Bot），从参数量来看可与 ChatGPT 媲美。华为通过模型泛化，解决传统 AI 作坊式开放模式下不能解决的 AI 规模化和产业化难题。

整体来看，美国在模型上仍具有先发优势，源于长时间的技术迭代和数据累积。从 2018 年发布 GPT 1.0，到 2021 年发布 GPT-3.5 和 Instruct GPT，OpenAI 已形成了一系列海量数据积累和训练模型，或将于今年年中/下半年发布 GPT-4 版本(万亿级参数量)。虽然当前部分中国大模型参数量能达到千亿级乃至万亿级，但数据质量及训练模型尚需时间精进。

图 30：美国大模型类型及参数量

	预训练模型	应用	参数量	领域
谷歌	BERT	语言理解和生成	4810 亿	NLP
	LaMDA	对话系统	1370 亿	NLP
	PaLM	语言理解与生成、推理、代码生成	5400 亿	NLP
	Lmagen	语言理解和图像生成	110 亿	多模态
	Parti	语言理解和图像生成	200 亿	多模态
微软	Florence	视觉识别	6.4 亿	CV
	Turing-NLG	语言理解和生成	170 亿	NLP
Meta	OPT-175B	语言模型	1750 亿	NLP
	M2M-100	100 种语言互译	150 亿	NLP
Deep Mind	Gato	多面手的智能体	12 亿	多模态
	Gopher	语言理解与生成、推理、代码生成	2800 亿	NLP
	AlphaCode	代码生成	414 亿	NLP
Open AI	GPT3	语言理解与生成、推理等	1750 亿	NLP
	CLIP&DALL-E	图像生成、跨模态检索	120 亿	多模态
	Codex	代码生成	120 亿	NLP
	ChatGPT	语言理解与生成、推理等	1750 亿	NLP
Anthropic	AnthropicLM v4-s3	语言理解与生成、推理等	520 亿	NLP
	Claude	语言理解与生成、推理等	520 亿	NLP
英伟达	Megatron-Turing NLG	语言理解与生成、推理	5300 亿	NLP
Stability AI	Stable Diffusion	语言理解和图像生成		多模态

资料来源：腾讯，网上公开信息整理，招银国际环球市场

图 31：中国大模型类型及参数

	预训练模型	应用	参数量	领域
百度（文心）	百度·文心	文本理解与创作	2600 亿	NLP
	VIMER-UFO	视觉识别	170 亿	CV
	ERNIE-ViLG 2.0	语言理解和图像生成	AI 作图全球最大（目前）	多模态
	Helix	生物计算		科学计算
华为（盘古）	NLP 大模型	文本理解与创作	千亿	NLP
	CV 大模型	视觉识别	30 亿	CV
	多模态大模型	语言理解和图像生成	千亿	多模态
	科学计算大模型	气象预测		科学计算
阿里巴巴	通义-M6	语言理解、图像生成和制造业	十万亿	多模态
	通义-AliceMind	文本理解与创作	270 亿	NLP
	通义-视觉	视觉识别		CV
腾讯	混元 AI 大模型	语言理解和图像生成（理解广告内容）	千亿	多模态
	精排大模型	精准匹配广告和人群	千亿	科学计算
商汤科技	视觉模型	视觉识别	300 亿	CV

资料来源：网上公开信息整理，招银国际环球市场

## 应用面：美国商业化加速、赋能 2B；中国变现早期，侧重生活互娱 C 端场景

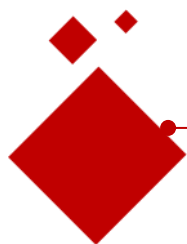
美国企业 AIGC 商业化进程加速，且服务和收费方式日渐丰富。例如，Open AI 今年预计实现 2 亿美元收入(Fortune 报道)，且宣布推出 ChatGPT Plus 付费订阅套餐(20 美元/月)。订阅用户可在免费服务基础上享受高峰时段免排队、快速响应、优先获取新功能等额外权益。此外，DALL-E 和 Stable Diffusion 等其他主流应用亦开启商业变现，企业根据使用量、内容产出量、提供服务以及使用属性等方面进行收费。美国 AIGC 场景落地侧重 2B 产业升级（如：合成数据等）。中国的商业化进程尚在早期，大部分产品处于免费试用阶段或根据使用量收费，且主要集中在语音、写作、数字人等 2C 场景应用。在 2B 方面，头部公司在金融、物流、政务层面亦在延伸。

图 32：中美大模型部分应用对比

公司	模型	领域	场景/效果
谷歌	PaLM	医疗	可为多项选择提供数据集，并对医疗专业人员和非专业人士提出的问题做出更详细的回答。根据一组医生的数据，92.6% 的 Med-PaLM 反馈与临床医生生成的答案 (92.9%) 相当。
	BERT	信息	BERT 可以从大量非结构化数据中提取信息，可用于为任何类型的图书馆创建搜索界面。
微软	Turing-NLG	信息	模型可以帮助用户对文档和邮件进行摘要化处理，同时能够应用在 Office 系列软件中，提供写作帮助和疑难解答。为下一代聊天机器人和数字助理铺平了道路，有望被用于客服、销售等方向。
	Florence	信息	通过整合图像-文本数据的通用视觉语言表示能力，Florence 可以轻松适用于各种计算机视觉任务，如分类、目标检测、VQA、看图说话、视频检索和动作识别，在多种类型的迁移学习中均表现出色。
百度	文心大模型	电影	测试数据显示，使用电影频道-百度·文心大模型每天可修复视频 28.5 万帧，解决了绝大部分画面的修复问题。即便是需要进一步精修，修复速度也能提升 3-4 倍。
	文心大模型	金融	金融文本命名实体识别模型的 F1 值提升约 3%，有效提升授信报告的处理效率；在金融语义相似度，金融事件主体抽取等一系列公开数据集上相对通用大模型取得显著提升。

文心大模型	汽车	<p>共建模型的智能客服知识库扩充任务人工评估可利用率相对基线提升 24.37%，能大幅提高问答的泛化水平，提高客服系统的智能化体验；除此以外，在车载语音系统短答案生成、汽车领域知识库构建等生成任务上的效果均有 20%-35% 的提升，可辐射至多个汽车行业场景。</p>
NLP 大模型	舆情	<p>在金融、电商，政务领域实现精准舆情分析、企业运营软件分析。关键技术：文档信息抽取、情感分类、文档自动摘要等。</p>
华为	营销	<p>取代目前逐渐上升的人力成本和降低成功率智能客户与营销系统。关键技术：文本匹配，问答和对话系统、意图识别等。</p>
CV 大模型	巡检	<p>进行无人机智能巡检，并解决巡检系统中数据标注工作量大和缺陷种类繁多问题。数据筛选效率提升 30 倍、筛选质量提升 5 倍，模型的精度比普通模型提升 18%。</p>
	物流	<p>1 个模型覆盖 9 种物流场景，监测收货、入库、在库和出库全流程。关键能力：大模型具备强泛化能力，一个模型同时做行为异常，轨迹异常检测；利用大模型的时空对齐能力，综合多摄像头数据，解决局部遮挡等问题，准确估计物体轨迹异常样本种类多、数量少的场景可利用小样本数据取得高精度结果。</p>

资料来源：网上公开信息整理，招银国际环球市场

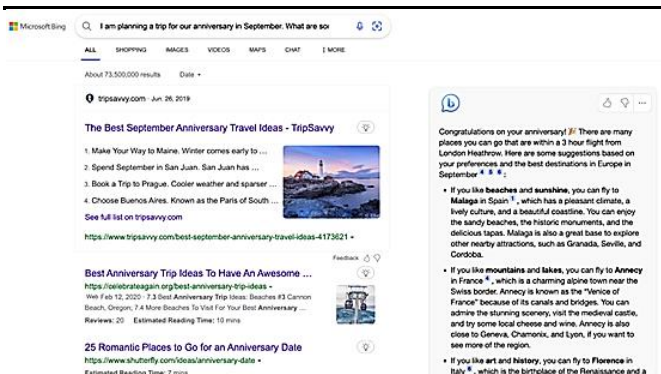


## 潜在受益标的

## 微软 (MSFT US)：深度融合 OpenAI 工具，技术和应用并行

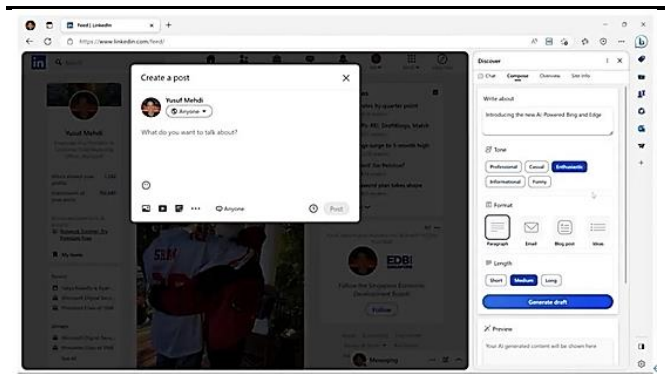
投资 OpenAI, 全线接入 OpenAI 应用。作为 OpenAI 股东(持股 49%), 微软将全线接入 OpenAI 应用, 包括将 ChatGPT、DALL-E 等整合进搜索、浏览器、Office、云服务、Teams 聊天程序等。2月8日, 微软推出集成 ChatGPT 的全新 Bing 搜索服务, 以及集成 AI 辅助的 Edge 浏览器。新版 Bing 搜索引擎新增一个扩展聊天框, 在回答问题的同时帮助即时生成个性化分析和建议(如做方案、写故事、写代码、做菜单等)。Edge 浏览器则是新增聊天和内容创作两项功能, 可帮助读报告(AI 助手整合摘要和关键点、做对比分析)、写邮件、发帖等。

图 33: 新版 Bing 搜索引擎 (集成 ChatGPT)



资料来源: Bing, 招银国际环球市场

图 34: Edge 浏览器帮助进行 LinkedIn 发帖



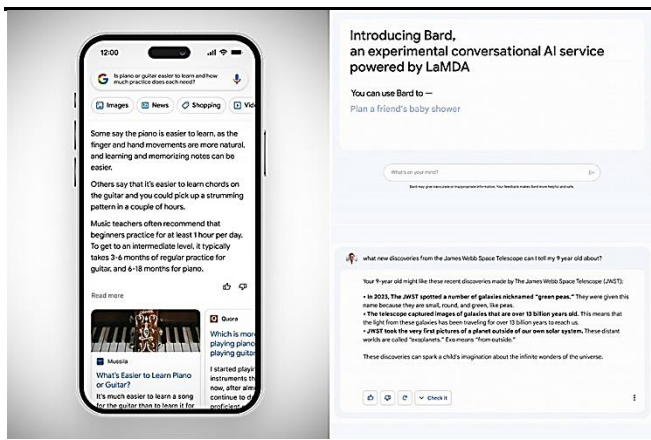
资料来源: Edge, 招银国际环球市场

大模型+云计算基础设施领先, AIGC 布局广泛。微软和英伟达合作的 Megatron Turing NLG 超大语言模型具备高参数优势(5300 亿), 叠加 Azure 云平台接入, 为微软 AIGC 布局及 OpenAI 训练开发提供强大支持。微软在 AIGC 应用层面布局广泛, 涵盖 AI 编程 (Copilot)、AI 生成图像(小冰框架)、AI 视频(GODIVA)、数字人(小冰)等。

## 谷歌 (GOOG US) : 发布“Bard AI”及开放生成语言 API, LaMDA 大模型积淀深厚

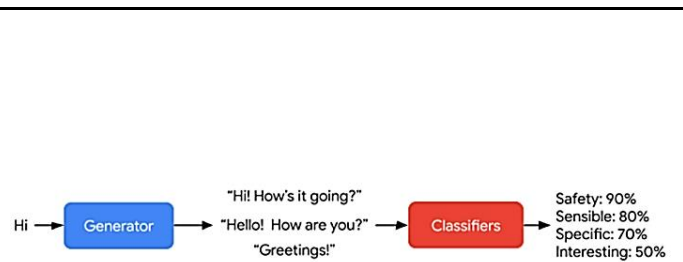
谷歌推出类 ChatGPT 产品: 基于 LaMDA 大模型的 Bard 聊天机器人。与 ChatGPT 相比, Bard 能够融合互联网最新事件信息, 而 ChatGPT 则是基于 2021 年之前语料库。虽然 2 月 8 日谷歌 Bard AI 在演示中给出了错误回答(市场反馈负面), 但我们认为训练模型局限下类似产品都存在回答并非准确的缺陷, 后续效果仍待观察。谷歌 LaMDA 大模型参数量(1370 亿)略低于 OpenAI, 其特色为开放域互动对话能力。此外, 谷歌计划 3 月将向企业、开发者和创造者开放生成语言 API, 鼓励开发者通过 API 打造自身产品。谷歌 AIGC 产品包括 AI 生成音频(Magenta), AI 生成图像(Imagen, Parti, NeRF, NeRV)和 AI 生成视频(Imagen Video, Phenaki)。

图 35: 谷歌 Bard 聊天机器人



资料来源: 谷歌, 招银国际环球市场

图 36: LaMDA 大模型训练方式



资料来源: 谷歌研究院, 招银国际环球市场



## 百度 (9888 HK) : 全栈式 AI 技术先锋, 文心一言+数字人+AI 创作平台矩阵丰富

百度作为国内首个“All in AI”头部平台, 以“文心大模型+深度学习”为技术底座, 在 ChatGPT 及 AIGC 领域拥有天然技术储备和全栈式布局优势。相较于行业小模型路线, 大模型泛化能力强、智能程度领先、研发更标准(高参数量数据训练+“填喂”优化), 极高的前期开发成本构筑壁垒, 且多模态多任务优势凸显、长尾投入低。相较于当前部分头部平台“卷参数”、“比落地”的大模型, 百度文心大模型更加侧重产业化(to B)、产品化(“文心百中”, “文心一格”等)及知识增强。目前文心拥有 36 个大模型(业界规模最大), 涵盖能源、金融、航天、制造、传媒、城市、社科、影视等领域, 且参数量达 2600 亿(高于 OpenAI ChatGPT 1750 亿)。期待文心大模型应用到搜索引擎、智能云的业务协同, 同时对外将加速产业渗透及多下游应用。

图 37: 百度文心大模型全景图

产品与社区	文心一格 AI艺术和创意辅助平台		文心百中 大模型驱动的产业级搜索系统		赐谷社区 大模型创意与探索社区			
工具与平台	EasyDL-大模型 零门槛 AI 开发平台		BML-大模型 全功能 AI 开发平台		大模型 API			
	大模型套件							
	数据标注与处理	大模型精调	大模型压缩	高性能部署	场景化工具			
文心大模型	行业大模型							
	国网-百度文心	浦发-百度文心	航天-百度文心	人民网-百度文心	冰城-百度文心	电影频道-百度文心		
	深燃-百度文心	吉利-百度文心	泰康-百度文心	TCL-百度文心	辞海-百度文心			
	NLP 大模型		CV 大模型		跨模态大模型		生物计算大模型	
	医疗 ERNIE-Health	金融 ERNIE-Finance		商品图文搜索表征学习 VIMER-UMS		文图生成 ERNIE-ViLG	文档智能 ERNIE-Layout	化合物表征学习 HelixGEM
	对话 PLATO	搜索 ERNIE-Search	信息抽取 ERNIE-UIE	OCR图像表征学习 VIMER-StrucText				蛋白质结构预测 HelixFold
	跨语言 ERNIE-M	代码 ERNIE-Code	图网络 ERNIE-Sage	多任务视觉表征学习 VIMER-UFO		视觉-语言 ERNIE-ViL	语音-语言 ERNIE-SAT	地理-语言 ERNIE-GeoL
	语言理解与生成			视觉处理 多任务学习 VIMER-TCIR	自监督视觉 表征学习 VIMER-CAE			单序列蛋白质结构预测 HelixFold-Single
	ERNIE 3.0 Tiny (亿级)	ERNIE 3.0 (百亿级)	鹏城-百度文心 (千亿级)	ERNE 3.0 Zeus (任务级千亿级)				

资料来源: 百度

**憧憬“文心一言”内测推出, 赋能搜索引擎提效。**百度宣布“文心一言”(ERNIE Bot) 计划于今年 3 月完成内测并开放, 以对标海外“ChatGPT”。预计产品落地将融入百度搜索, 提供多答案回复、智能生成等功能。在文心大模型加持下, 我们期待“文心一言”上线后用户迅速起量(中国有海量用户基数), 功能赋能搜索引擎后有望提高搜索转化率, 短期财务影响有限, 但未来商业化空间可期(当前 ChatGPT 推出付费订阅模式: US\$20/每月)。当前市场存在对文心一言“中文文本质量”、“监管内容审查”部分担忧, 但我们认为高参数量加持下部分中文质量影响有限, 且模型自我审查能力加强降低敏感内容回答风险。

**数字人: 打造数字人家族, 多平台合作推动 AIGC 加速渗透。**百度智能数字人平台依托百度智能云及 AI 能力, 提供 2D/3D 的服务型及演艺型数字人形象生产和创作(输入文本后可直接生成数字人口型匹配的视频, 数字人口型合成准确率达 98.5%, 基于百亿参数对话大模型 PLATO), 服务金融、运营商、互娱/品牌商、媒体/广电、MCN 等行业。例如, 百度数字人“度晓晓”自 2020 年上线后, 多次主持百度世界大会、创作歌曲、画作、直播等, 能够以 40 秒完成 40 篇高考作文。未来随着 AI 算法突破, 我们看好数字人制作周期缩短、成本下降下产业的加速渗透。

图 38: 百度数字人产品架构



资料来源：百度，招银国际环球市场

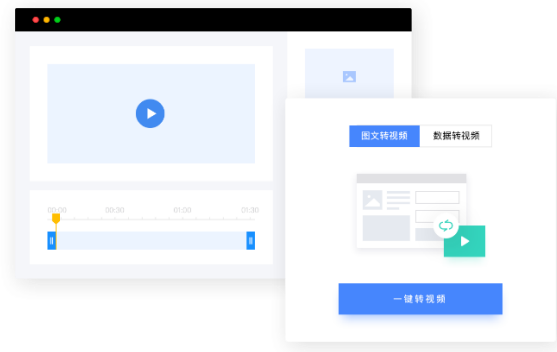
**智能创作平台：视频创作+智能写作赋能内容生产。**百度智能创作平台提供一站式AI创作服务和工具(于2022年4月正式商用)，涵盖视频创作、图片创作和智能写作，赋能媒体、金融、政企等多行业，自动创作文章/短视频 200万+/15万+，并和超20家媒体落地合作。其中，视频AI创作主要基于多模态创作技术，创作者可高效、低成本、低门槛的实现“图文转视频”、“数据转视频”、“智能配音”、虚拟主播等功能(13种工具)。智能写作则是集合庞大语料库和资讯素材，提供自动写作的同时具备纠错、润色、标题推荐等服务。

图 39: 百度AI智能写作



资料来源：百度，招银国际环球市场

图 40: 百度AI视频创作

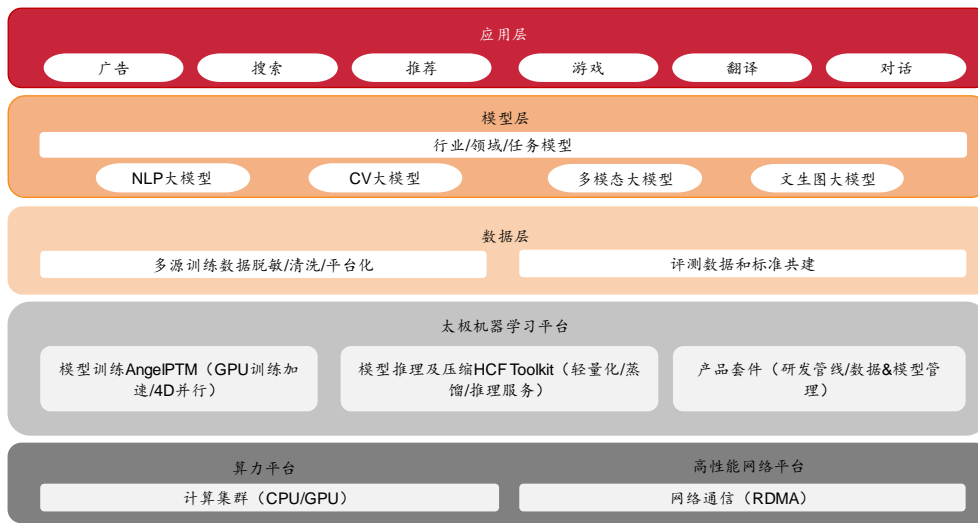


资料来源：百度，招银国际环球市场

## 腾讯 (700 HK): 混元 AI 大模型赋能多业务场景落地、实现技术提效

基于海量社交数据、机器学习、NLP 等技术积累，腾讯于 2022 年 12 月推出了万亿中文 NLP 预训练模型 HunYuan-NLP-1T (混元 AI 大模型)，是目前国内首个低成本、可落地的 NLP 万亿大模型，且登顶自然语言理解任务榜单 CLUE。在模型架构层面，和其他大模型相比，混元首创“层级化”表征，对输入文字/视频做分层后再进行关联性检索，大幅提升检索准确度。

图 41: 腾讯混元 AI 大模型



资料来源: 腾讯, 招银国际环球市场

图 42: 腾讯混元大模型在五大跨模态视频检索数据集榜单排名首位

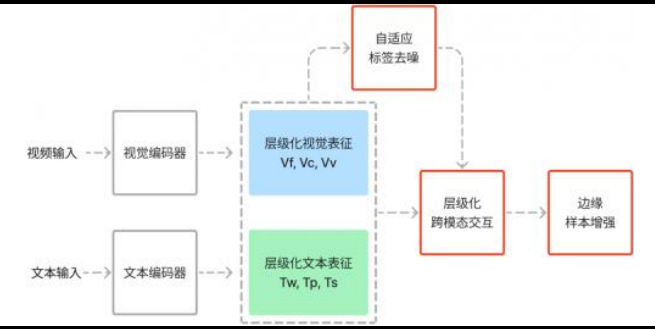
Method	Hunyuan_tvr	DRL (阿里)	MDMMT-2 (华为)	腾讯 PCG (CLIP2video)	CAMoE (快手)	CLIP3Clip (微软)
MSR-VTT	55.0%	53.3%	48.8%	52.9%	48.8%	44.5%
MSVD	58.2%	50.0%	56.8%	47.0%	49.8%	46.2%
LSMDC	29.7%	26.5%	26.9%	-	25.9%	22.6%
DiDeMo	52.1%	49.0%	-	-	-	41.4%
ActivityNet	57.3%	46.2%	-	-	-	41.4%

资料来源: 腾讯

注释: MSR-VTT、MSVD、LSMDC、DiDeMo、ActivityNet 是行业内最具权威性的五大跨模态视频检索数据集榜单，主办单位包括微软、加州大学伯克利、阿卜杜拉国王科技大学等

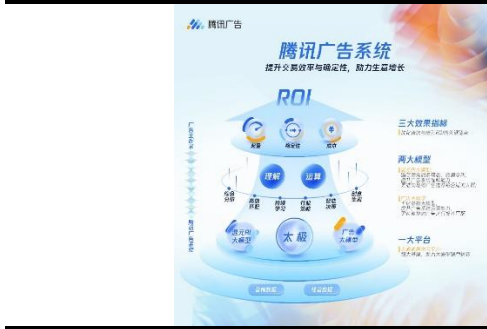
**大模型多产品协同，广告降本提效显著。**混元涵盖 NLP、CV、多模态、文生图等大模型，对内已和腾讯广告、微信、QQ、游戏等产品实现协同，并通过腾讯云对外商业化。有别于百度文心与搜索的融合，腾讯大模型在广告提效层面更为显著：依托于太极机器学习平台，腾讯广告通过两大模型训练(混元 AI 大模型与广告大模型)强化内容理解、精准匹配、文案创意生成等能力，有效提升投放 ROI。同时，在大模型广告智能创作下，文案转视频功能亦大幅降低视频广告制作成本。

图 43: 混云模型架构示意图



资料来源：腾讯，招银国际环球市场

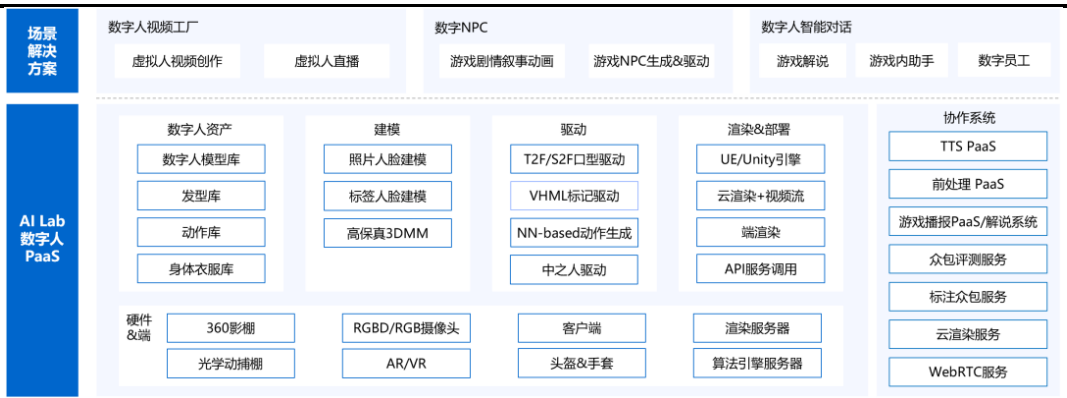
图 44: 腾讯广告系统搭载“一大平台两大模型”



资料来源：腾讯，招银国际环球市场

实验室矩阵支撑 AIGC 能力建设，围绕“内容+社交+游戏”多元布局。腾讯依托 AI Lab、优图实验室、微信 AI 等多个 AI 实验室，积极布局 AIGC 产品，包括 AI 生成文字(梦幻写手 Dreamwriter)、AI 生成游戏(绝悟 AI)、AI 绘画(腾讯 QQ 小世界 AI 画匠)、数字人(AI Lab 数字人 PaaS 方案，创建的数字人“艾灵”)等。例如，腾讯 2022 年 12 月由 AI Lab 研发的智能创作助手文涌 (Effidit) 更新至 2.0 版本，新增风格化智能续写功能。

图 45: 腾讯 AI Lab 数字人全栈技术管线和场景解决方案



资料来源：腾讯，招银国际环球市场

## 阿里巴巴 (BABA US): 通义大模型凸显开源, 构建 AI 统一底座

通义大模型构建 AI 统一底座, 类 ChatGPT 产品研发内测中。阿里达摩院于 2021 年发布使用 512 卡 V100 GPU 实现全球最大规模 10 万亿参数多模态大模型 M6, 并于 2022 年发布最新「通义」大模型系列。通义大模型注重开源开放, 首次通过“统一范式”实现多模态、多任务、多结构的运行, 并通过模块化设计实现高效率高性能。其中, 通义 - AliceMind (深度语言模型体系) 在中文语言理解测评基础 CLUE 上三个维度排名首位。基于大模型技术储备、强大的算力及云服务支持, 阿里目前在积极研发类 ChatGPT 产品 (处于内测阶段), 或通过 AI 大模型与钉钉合作。

图 46: 阿里通义大模型架构



资料来源: 阿里巴巴, 招银国际环球市场

图 47: 阿里 AliceMind 模型结果在中文语言理解测评基础 CLUE 上获得了三榜第一

排行	模型	研究机构	测评时间	Score 1.1	认证	AFQMC	TNEWS 1.1	IFLYTEK	OCNLI_50K	WSC 1.1	CSL	CMRC 2018	CHID 1.1	C3 1.1
1	HUMAN	CLUE	19-12-01	86.678	已认证	81.0	71.0	80.3	90.3	98.0	84.0	92.4	87.1	86.0
2	通义-AliceMind	达摩院NLP	22-08-31	85.738	待认证	83.4	71.6	65.7	83.3	94.1	94.9	86.8	99.2	92.7
3	Wenjin	Meituan NLP	22-06-24	84.900	待认证	83.5	69.6	64.4	85.6	92.3	93.2	86.3	98.9	90.4
4	HunYuan_ntp	腾讯 TEG	22-05-11	84.730	待认证	83.4	64.0	66.6	85.2	92.3	93.9	87.9	98.5	90.8
5	ShenNonQ	云小微AI	21-12-01	84.351	待认证	82.6	65.6	64.4	86.0	94.2	91.2	86.5	97.9	90.8
6	Shenzhou	QQ浏览器实验室	21-09-19	83.873	待认证	80.6	65.4	67.7	86.4	89.1	91.0	87.9	97.9	89.1
7	OBERT	OPPO-小布助手	22-06-01	81.383	待认证	77.9	65.8	63.9	83.4	91.0	87.3	84.1	94.7	84.4
8	CL-BERT	CL-BERT	22-04-06	81.288	待认证	82.4	64.0	64.8	82.8	88.6	91.5	82.0	98.5	77.0
9	Mengzi	澜舟科技-创新工场	21-09-14	81.092	待认证	81.8	65.2	65.1	82.6	86.5	89.9	84.0	95.1	79.8
10	PAI-EasyNLP BERT	wjn1996	22-07-12	80.570	待认证	77.1	61.2	61.2	82.8	87.1	94.2	77.3	97.7	86.6

资料来源: 阿里巴巴

图 48: 通义大模型文生图效果展示: 文艺复兴时期皇家花栗鼠肖像画



资料来源: 阿里巴巴, 招银国际环球市场

**AIGC 智能客服先行, 视觉生成、在线设计、在线试穿等多点开花。**相较于 AIGC 在百度搜索、腾讯广告中的应用侧重, 我们认为阿里 AIGC 落地场景更多为智能客服(未来 ChatGPT 融入后对话将更为类人及智能)及广告, 助于提升电商消费体验并降低运营成本。当前阿里云智能客服已全面应用通义大模型, 在表格问答层面交付准确率达到 90%+, 对外服务近 20 个行业。此外, 阿里在 AI 生成文学(阿里 DT 稿王)、AI 视觉(阿里鹿班)、AI 设计

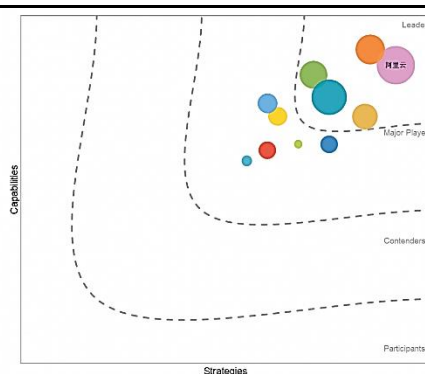
(Lubanner)、数字人(AI 模特塔玑、超写实数字人 AYAYI)等领域多点开花。同时，阿里巴巴 FashionAI 项目中将 AIGC 应用于在线试穿、个性化智能搭配、虚拟货场构建。

图 49: FashionAI 个性化智能搭配



资料来源: 阿里巴巴, 招银国际环球市场

图 50: IDC 智能客服评估中阿里云两大维度排名第一



资料来源: IDC MarketScape <2022 年新一代智能客服市场>, 招银国际环球市场

## 商汤 (20 HK) - 拥有 300 亿参数视觉大模型

商汤是国内最大的计算机视觉软件公司，根据 IDC 发布的《1H22AI 及软件市场追踪报告》，商汤在国内计算机视觉市场分额为 20.7%。截止到 2022 年 6 月 30 日为止，公司累计商用模型 4.9 万个，专利资产数达到 12,502 个。公司从 2018 年至今累计研发投入超过 100 亿元人民币，拥有 4,093 个研发人员。

### 技术储备：自研训练框架，自建超算集群

和其他 AI 软件公司相比，公司的模型都是在自建的超级计算集群上做训练。公司建立了超过 23 个超算集群，拥有超过 20,000 个 GPU，同时，公司有自研的深度学习训练框架 (SenseParrots)，比其他只支持固定计算路径的训练框架，SenseParrots 有内置的基石编译器可以实时快速编译计算路径。这些底层的算力基础设施及训练框架使公司在 AI 模型的训练上更高效。

### 技术路径：打造视觉基模型，更快更高效开发特定场景 AI 软件

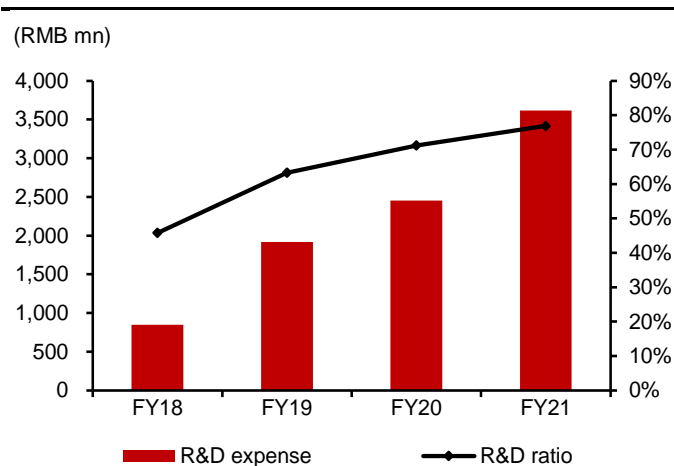
商汤的技术储备让公司在开发 AI 软件上选了不一样的技术路径。公司已经训练具有高达 300 亿个参数的计算机视觉 (CV) 基模型。商汤的超大基模型在 ImageNet 大规模识别任务中实现超 90% 精度，与谷歌、微软共同跻身世界前三。基于大模型的自动数据生产线实现了数据自动标注速度相较业界传统人工标注提升 600 倍，成本降低至 1/500，且标注精度跟人工水平一致。

基模型聚集来自众多场景海量数据的通用知识，使得在开发针对特定场景的 AI 软件时，公司可以利用已经训练好的基模型，以更少的训练数据生成特定场景模型，这个能有效处理长尾场景训练数据不足的问题。

### 变现空间：订阅模式的 AI-as-a-Service

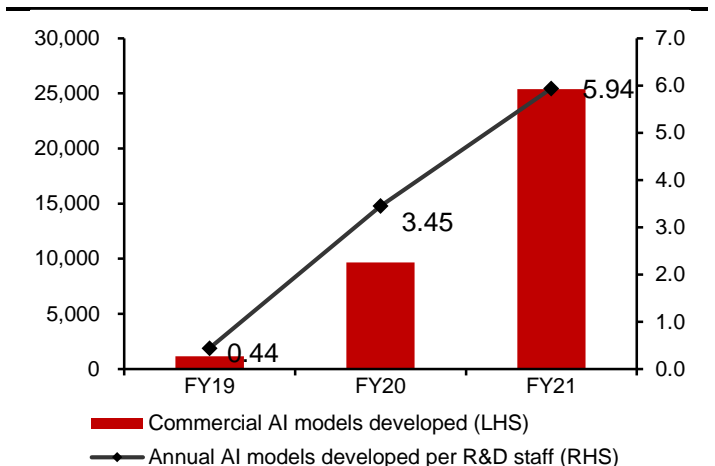
商汤已建成并启用的 AI 超算算力规模超过 2.5 exaFLOPS，其中在上海临港的 AI 智算中心 (AIDC) 上线了 1.745 exaFLOPS，其中对外服务的算力突破 1 exaFLOPS，客户包括学术机构和企业（特别是车企客户）来订阅。

图 51：商汤研发投入持续增加



资料来源：公司资料，招银国际环球市场

图 52：视觉超大模型带动 AI 模型产量大幅提升



资料来源：公司资料，招银国际环球市场

## 科大讯飞 (002230 CH) - AI 语音市场龙头

科大讯飞是中国人工智能语音龙头，根据 IDC 发布的《2022 上半年中国 AI 软件及应用市场研究报告》，公司在中国人工智能语音语义市场份额占比达 12%，排名第一。公司依托讯飞开放平台，覆盖教育、医疗、智慧城市、企业数智化转型、办公等赛道。公司五年来研发投入持续提升，2021 年达到 29.4 亿元人民币。2021 年公司研发人员合计 8,367 人，占公司员工总数达 58.5%。

### 认知智能领域技术领先

科大讯飞在认知智能相关领域已经具备了长期深厚的积累，体现在 1) 世界领先的多语言理解、跨语言迁移和知识图谱推理等核心技术，2022 年已经在 CommonsenseQA 2.0、OpenBookQA 等 12 项认知智能领域权威评测中位列第一；2) 公司的开源中文预训练语言模型在 Github 平台的星标数位列业界第一，应用场景覆盖了 6 大类、超过 40 个通用领域。

### 生成式预训练大模型仍在早期阶段

科大讯飞虽然在语音语义理解领域处于领先地位，但在 ChatGPT 所具备的自然语言数据处理及生成等技术方向上，仍与互联网巨头有明显差距。以 WIPO 上记录的 G06F17/20 类专利 (Handling natural language data) 数量为例，科大讯飞只有 100 多个，而 BAT 均持有 600 个以上。另外，公司的生成式预训练大模型在去年 12 月份才启动技术攻关，预计首先落地在教育等个别垂直领域的产品中。

### 开放 API 模式实现商业变现

在产品变现方面，公司除了教育、医疗、智慧城市等产业应用，还通过讯飞开放平台按 API 调用量实现收入。平台集成了语音合成、识别、自然语言处理、人机交互等模块，为开发者提供一站式的智能语音应用低代码开发能力。截至 2022 年 6 月 30 日，讯飞开放平台已开放 496 项 AI 能力及场景方案，开发者团队数量达 343 万，总应用数达 152 万，累计覆盖终端设备数 35.4 亿+，生态合作伙伴达 500 万。在自然语言处理能力方面，该平台目前已整合了 AIUI 人机交互、词法分析、语义依存分析、关键词提取等 17 项功能模块。

图 53: 讯飞开放平台能力概览



资料来源：公司资料，招银国际环球市场



## 思必驰（未上市）- “云+芯”渗透 AI 语音语义市场

思必驰是国内领先的 AI 语音公司。根据 IDC 发布的《2022 上半年中国 AI 软件及应用市场研究报告》，思必驰在中国 AI 语音语义市场份额中排名第四，次于科大讯飞、阿里云、百度智能云等巨头。思必驰于 2007 年在英国剑桥成立，随后在江苏注册。公司在成立之初就持续投入智能语音语义交互技术架构的研发，并于 2014 年推出首个对话智能云平台“对话工场”。2016 年，公司开始布局“云+芯”战略，并先后在 2017 年和 2019 年发布了全链路智能对话系统定制开发平台“DUI”和首款自研 AI 语音芯片 TH1520，通过软硬一体的方式提供自然语言交互解决方案。2022 年 7 月，公司向科创板递交 IPO 申请。

### “云”：DUI 平台加速大规模对话系统定制开发

思必驰是国内最早开展对话式 AI 定制技术和自动化工具研发的公司之一。公司的 DUI 平台（Dialog User Interface）内置丰富的 SDK 和控件资源，可以支持人机对话系统的低代码快速搭建。技术能力层面，基于 20+ 垂直行业的自然语言处理数据积累，DUI 平台的云端通用领域识别率达到了 97%+，场景语义理解率达到了 92%+。

### “芯”：自研智能语音处理辅助芯片，推进软硬一体产品落地

针对 DUI 平台的云端算法和模型特点，思必驰旗下芯片公司申聪智能对语音处理辅助芯片的架构进行的专用设计，并自研适配了神经网络计算指令，于 2019 年发布了首款自研 AI 语音处理辅助芯片 TH1520，显著优化特定场景处理性能，降低了计算资源占用，有助于公司软硬结合产品的规模化落地。公司的二代芯片 TH2609 于 2021 年 5 月发布，目前已完成流片，将面向智能家居、车载设备和智能办公等场景提供升级解决方案。

### 较小的公司体量难以支持 AI 大模型研发

2021 年公司的研发费用为 2.9 亿元人民币，研发人员 720 人，公司的研发投入和互联网巨头（像 BAT）或 AI 技术领先的公司（像商汤、科大讯飞）比较存在明显差距，虽然公司在语音语义方面的投入比较早，但我们认为思必驰在 NLP 或是大模型的训练能力上将难以追赶前面所述的技术领先企业。

图 54：主要产品结构

产品名称	产品类别	收费模式	典型案例	FY21 收入 (RMB mn)
智能人机交互 软件产品	智能家电语音解决方案	项目制	TCL 平板、欧瑞博中控	64.7
	智能汽车语音解决方案		理想汽车、小鹏汽车	
	数字政企智能助理		重庆农商行、广州地铁	
软硬一体化 AI 产品	AI 语音芯片	按销售数量	海信电视、钉钉拍	101.5
	AI 模组	按销售数量	当贝投影、华尔斯面板	
	AI 终端	项目制/按销售数量	五菱收音机、车萝卜 HUD	
对话式 AI 技术 服务	技术授权服务	买断/按调用时长	OPPO、中国移动	73.7
	定制开发服务	项目制	博泰车联网、美的空调	67.5

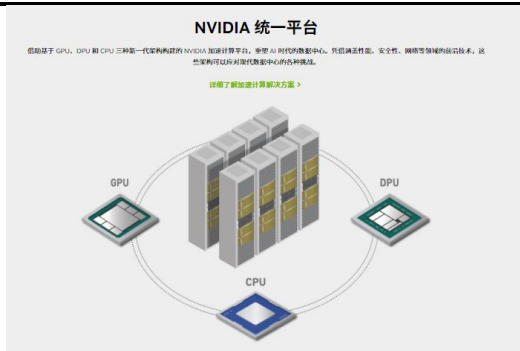
资料来源：公司招股书，招银国际环球市场

## 英伟达 (NVIDIA, NVDA US) - 全球 AI 计算平台龙头

英伟达目前是全球 AI 计算平台龙头，积累计算堆栈和神经网络各层级的完善产品矩阵，包含硬件、系统软件、平台软件和应用四层。**硬件方面**，最底层基于 CPU、GPU、DPU 三类芯片，形成三大算力支柱；**系统层面**，从云端到边缘部署广泛的多种系统，如 RTX（云端游戏服务器）、DGX（一站式 AI 解决方案）、HGX（AI 超级计算平台）、EGX（边缘 AI 计算平台）、OVX（数字孪生模拟）等，为开发者提供硬件基础设施。**平台软件层面**，英伟达为开发者提供 150 个加速库服务，涉及机器学习、神经网络、计算机视觉、基因检测等细分技术领域。**应用方面**，英伟达打造了一系列针对特定领域的 AI 应用框架，如 MODULUS（物理机器学习框架）、RIVA（语音 AI 应用）、Isaac（机器人平台）等，这些框架都基于 NVIDIA HPC、NVIDIA AI 和 NVIDIA Omniverse 平台构建。每一层级都对计算机制造商、服务提供商和开发者开放，软硬一体的解决方案适用于广泛下游领域。

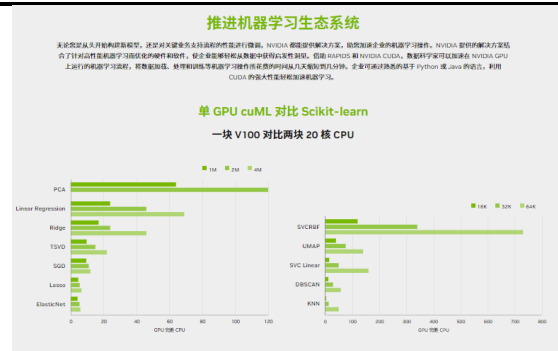
英伟达 GPU 因其高性能和对 CUDA 支持而广泛应用于 AI 和机器学习领域（CUDA 是 NVIDIA 的并行计算平台和编程模型）。OpenAI 主要使用英伟达 GPU 来训练和运行 AI 模型，如 GPT-3 和 ChatGPT，据公开数据显示，ChatGPT 已导入至少 1 万个英伟达 GPU。

图 55: 英伟达计算架构



资料来源：英伟达官网，招银国际环球市场

图 56: 英伟达机器学习生态系统



资料来源：英伟达官网，招银国际环球市场

## 超威半导体 (AMD, AMD US) - CPU+GPU 双芯片行业龙头

AMD 主要业务是 CPU 与 GPU，产品包括 x86 微处理器（PC、服务器、嵌入式）、独立和集成显卡、数据中心和专业显卡、加速器（APU）、主板芯片组和 SoC 等。作为全球 CPU 市场双雄之一，近年公司凭借全新的 Zen 架构以及台积电的先进工艺，在 PC CPU 产品领域不断夺取 Intel 市场份额。公司于 2022 年 2 月收购赛灵思，拓展 FPGA 领域，与 CPU、DSP 的融合将逐步成为市场主流。公司通过收购 Pensando，进军 DPU 领域，进一步拓展数字中心解决方案能力。

图 57: AMD AI 芯片产品系列



资料来源：超威半导体官网，招银国际环球市场

## 莱迪思 (Lattice, LSCC US) - 全球领先 FPGA 供应商

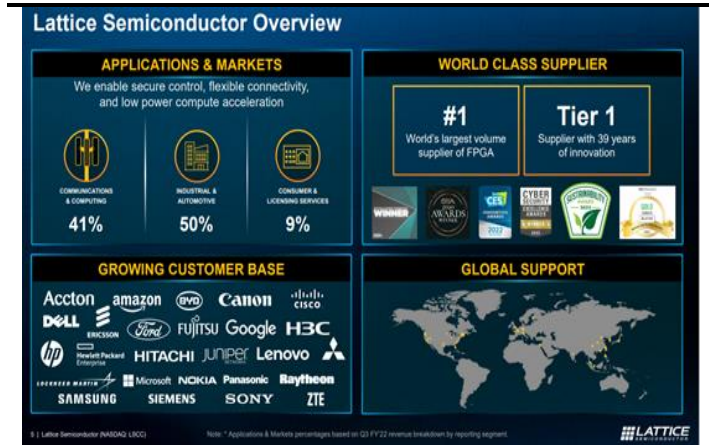
莱迪思是全球领先的低功耗 FPGA 供应商，下游客户覆盖全球一线互联网、通讯、数据中心、工业、汽车及消费级硬件公司，公司提供业界最广泛的现场可编程门阵列 (FPGA)、可编程逻辑器件 (PLD) 及其相关软件，包括现场可编程系统芯片 (FPSC)、复杂的可编程逻辑器件 (CPLD)，可编程混合信号产品 (ispPAC®) 和可编程数字互连器件 (ispGDX®)。公司两大主流 FPGA 产品线为 Lattice AVANT 与 Lattice NEXUS，并搭配有覆盖广泛下游应用场景的软件解决方案。

图 58: Lattice FPGA 产品矩阵



资料来源：莱迪思公司官网，招银国际环球市场

图 59: Lattice 公司概况



资料来源：莱迪思公司官网，招银国际环球市场

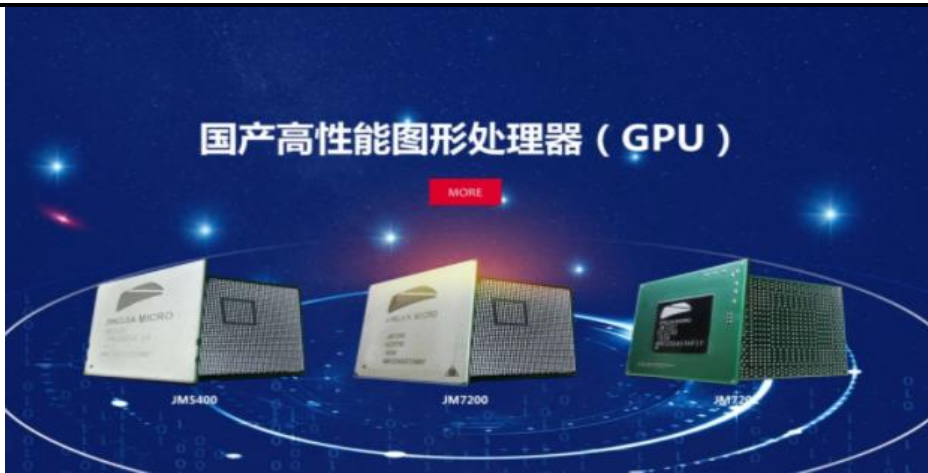
## 景嘉微 (300474 CH) - 国产 GPU 龙头企业

景嘉微产品主要涉及图形显控领域、小型专用化雷达领域和芯片领域。图形显控是公司现有核心业务，也是传统优势业务，小型专用化雷达和芯片是公司未来大力发展的业务方向。

在图形处理芯片领域，公司经过多年的技术钻研，成功自主研发了一系列具有自主知识产权的 GPU 芯片，是公司图形显控模块产品的核心部件并以此在行业内形成了核心技术优势。公司以 JM5400 研发成功为起点，不断研发更为先进且适用更为广泛的一系列 GPU 芯片，随着公司 JM7200 和 JM9 系列图形处理芯片的成功研发，公司联合国内主要 CPU、整机厂商、操作系统、行业应用厂商等开展适配与调试工作，共同构建国产化计算机应用生态，在通用领域成功实现广泛应用。

图形显控模块是信息融合和显示处理的“大脑”是公司研发最早、积淀最深、也是目前最核心的产品。公司成功研发了具有完全自主知识产权的系列 GPU 芯片，并以公司自主研发的 GPU 芯片为核心开发了系列图形显控模块产品，显著提升了公司产品竞争力。2022 年 5 月，公司 JM9 系列第二款图形处理芯片成功研发，可以满足地理信息系统、媒体处理、CAD 辅助设计、游戏、虚拟化等高性能显示需求和人工智能计算需求，可广泛应用于用于台式机、笔记本、一体机、服务器、工控机、自助终端等设备。

图 60: 景嘉微 GPU 产品系列



资料来源: 公司官网

### 寒武纪 (688256 CH) - 国内 AI 芯片领先者

寒武纪专注于人工智能芯片产品的研发与技术创新, 致力于打造人工智能领域的核心处理器芯片, 让机器更好地理解和服务人类, 主营业务为各类云服务器、边缘计算设备、终端设备中人工智能核心芯片的研发、设计和销售, 主要产品为云端智能芯片及加速卡、训练整机、边缘智能芯片及加速卡、终端智能处理器 IP 以及上述产品的配套软件开发平台。目前, 公司的主要产品线包括云端产品线、边缘产品线、IP 授权及软件。

2022 年 3 月, 公司发布新款训练加速卡 MLU370-X8。MLU370-X8 搭载双芯片四芯粒思元 370, 集成寒武纪 MLU-Link™ 多芯互联技术, 主要面向训练任务, 在业界应用广泛的 YOLOv3、Transformer 等训练任务中, 8 卡计算系统的并行性能平均达到 350W RTX GPU 的 155%。

图 61: 寒武纪最新 GPU 代表产品-思元 370 智能加速卡



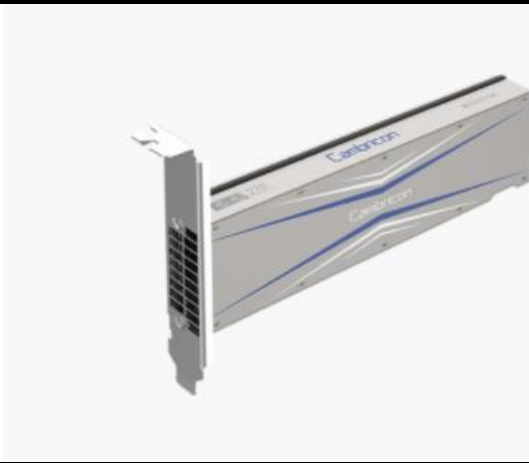
资料来源: 公司官网

图 62: 寒武纪思元 370 板卡与业内主流 GPU 性能对比



资料来源: 公司官网

图 63: 寒武纪 MLU270-S4 智能加速卡



资料来源: 公司官网

图 64: 寒武纪思元 270-S4 产品规格

思元270-S4 产品规格		
芯片型号	思元270 (MLUv02 架构)	
产品性能	INT8理论峰值/TOPS	128*
	INT4理论峰值/TOPS	256*
	INT16理论峰值/TOPS	64*
计算精度支持	低精度、混合精度	INT16, INT8, INT4, FP32, FP16
内存规格	内存容量	16GB DDR4, ECC
	内存位宽	256-bit
	内存带宽	102 GB/s
接口	PCIe接口	<16 PCIe Gen.3
功耗	最大热设计功耗	70w
	散热设计	被动散热
形态	半高半长, 单槽位	
尺寸	167.5mm × 68.9mm	
重量	310g	

\*为TOPS=1024GOP/S换算, 如INT8实际算力为131.072GOP/S或131TOPS。

资料来源: 公司官网

## 海光信息 (688041 CH) - 中国微处理器领军者

海光信息主营业务是服务器、工作站等计算、存储设备中的高端处理器, 产品包括海光通用处理器 (CPU) 和海光协处理器 (DCU)。海光 CPU 系列产品兼容 x86 指令集以及国际上主流操作系统和应用软件, 性能优异, 软硬件生态丰富, 安全可靠, 得到国内用户的高度认可, 已广泛应用于电信、金融、互联网、教育、交通等行业或领域。海光 DCU 系列产品以 GPGPU 架构为基础, 兼容通用的“类 CUDA”环境以及国际主流商业计算软件和人工智能软件, 软硬件生态丰富, 可广泛应用于大数据处理、人工智能、商业计算等领域。

公司研发出的第一代、第二代 CPU 和第一代 DCU 产品的性能均达到了国际上同类主流高端处理器的水平, 在国内处于领先地位。海光 DCU 目前已经实现商业化应用, 未来将广泛应用于大数据处理、人工智能、商业计算等应用领域。公司在高端处理器及相关领域开展了系统化的知识产权布局, 为公司保持技术的持续领先奠定了坚实的基础。

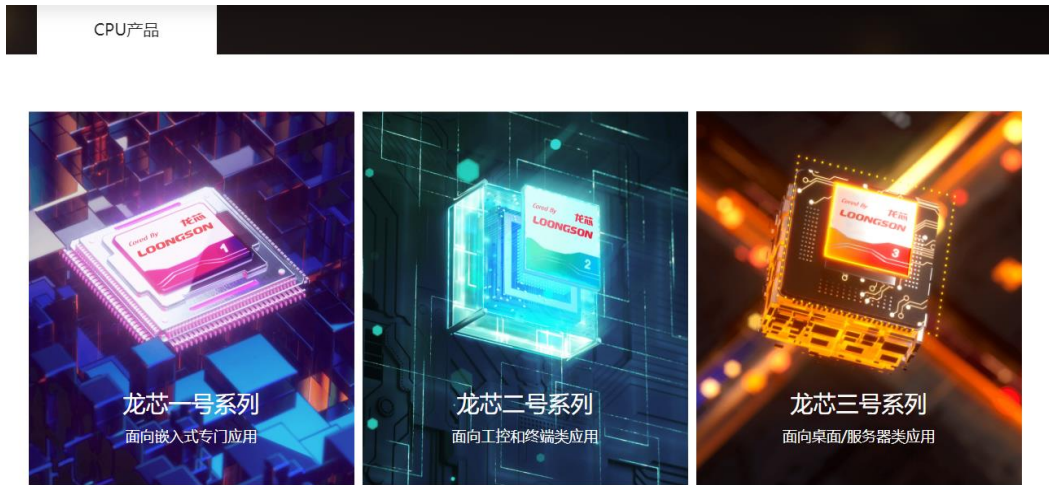
## 龙芯中科 (688047 CH) - 国产 CPU 龙头企业

龙芯中科主营业务为处理器及配套芯片的研制、销售及服务, 主要产品与服务包括处理器及配套芯片产品与基础软硬件解决方案业务。

**(1) 处理器及配套芯片产品:** 龙芯中科研制的芯片包括龙芯 1 号、龙芯 2 号、龙芯 3 号三大系列处理器芯片及桥片等配套芯片, 依据应用领域的不同可分为工控类芯片和信息化类芯片。在信息化领域, 国内数十家整机品牌推出基于龙芯 CPU 的台式机、笔记本、一体机与服务器设备, 已广泛应用于电子政务办公信息化系统, 并在金融、教育等应用中展开批量应用试点。在工控领域, 国内上百家主要工控和网络安全设备厂商推出了基于龙芯 CPU 的工控和网安产品, 包括工业 PC、工业服务器、工业存储设备、DCS (分布式控制系统)、PLC (可编程逻辑控制器)、交换机、路由器、防火墙等。为支持芯片销售及应用, 龙芯中科开发了基础版操作系统及浏览器、Java 虚拟机、基础库等重要基础软件, 持续优化改进, 并以两种方式免费提供给客户。

**(2) 解决方案:** 硬件模块业务的主要产品为基于龙芯处理器的开发板和验证模块。龙芯处理器通过在标准电路板上搭配配套芯片、内存条、电阻电容等不同类型电子元器件形成开发板和验证模块。技术服务业务主要是公司提供的基于龙芯处理器的系统产品开发服务, 目的是支持客户使用龙芯处理器。

图 65: 龙芯中科 CPU 主打产品系列



资料来源：公司公告

## 紫光国微 (002049 CH) - 特种集成电路龙头

紫光国微为国内主要的综合性集成电路上市公司之一，以智能安全芯片、特种集成电路为两大主业，同时布局半导体功率器件和石英晶体频率器件领域，为移动通信、金融、政务、汽车、工业、物联网等多个行业提供芯片、系统解决方案和终端产品。

主营业务分为：**1) 智能主要产品：**包括以通信 SIM 卡芯片、金融 IC 卡芯片、电子证照芯片等为代表的智能卡安全芯片和以金融支付终端芯片、非接触读写器芯片等为代表的智能终端安全芯片，同时可以为通信、金融、工业、汽车、物联网等多领域客户提供基于安全芯片的创新终端产品及解决方案。**2) 特种集成电路业务：**产品涵盖微处理器、可编程器件、存储器、网络总线及接口、模拟器件、ASIC/SoPC 等几大系列产品，500 多个品种，同时可以为用户提供 ASIC/SOC 设计开发服务及国产化系统芯片级解决方案。**3) 半导体功率器件业务：**产品涵盖 SJ MOSFET、SGT/TRENCH MOSFET、VD MOSFET、IGBT、IGTO、SiC 等先进半导体功率器件，在绿色照明、风力发电、智能电网、混合动力/电动汽车、仪器仪表、消费电子等多个领域形成系列成熟应用方案。**4) 石英晶体频率器件业务：**产品覆盖晶体谐振器、晶体振荡器、压控晶体振荡器、温补晶体振荡器、恒温晶体振荡器等主要品类，广泛应用于通讯设备、汽车电子、工业控制、仪器仪表等多个领域。

在特种集成电路领域，公司掌握了高可靠微处理器的体系结构设计、指令集设计和实现技术，建立了单片及组件总线产品的设计、验证和测试平台，在国内处于领先地位。目前已形成几大系列产品，核心产品在相关领域得到广泛应用，并以现场可编程技术与系统集成芯片相结合，成功推出具备现场可编程功能的高性能系统集成产品 (SoPC)，获得市场的广泛认可。

图 66: 紫光国微产品阵营

图 66 展示了紫光国微的产品阵营，分为两个主要部分：

**主要产品 (左侧):**

- 接触式IC卡芯片:** 包括智能终端安全芯片、半导体功率器件、超微晶体振荡器。
- 主要产品 (中间):** SIM芯片、eSIM芯片、NFC-SIM芯片、半导体功率器件。
- 主要产品 (右侧):** 智能终端安全芯片、USB-KEY芯片、mPOS芯片、石英晶体振荡器。

**主要产品 (右侧):**

- 智能终端安全芯片:** 提供安全解决方案。
- USB-KEY芯片:** 用于身份认证和数据存储。
- mPOS芯片:** 用于移动支付终端。
- 石英晶体振荡器:** 提供高精度时钟信号。

资料来源: 公司官网

### 航锦科技 (000818 CH) - 军工/电子/化工三大业务板块驱动成长

公司业务分军工、电子和化工三大板块。其中，军工板块以芯片产品为核心，产品涵盖存储芯片、总线接口芯片、模拟芯片、图形处理芯片、特种 FPGA、多芯片组件等，广泛应用于航空航天、兵器装备、机载雷达、舰载雷达、卫星通讯、电子对抗、雷达及末端制导、灵巧武器等。电子板块产品涵盖射频芯片、通信天线、北斗产品等，广泛应用于通信领域、医疗电子、工业控制领域、汽车电子、手机射频模组、导航、授时服务、地灾检测、高精度定位等。化工板块主要从事基础化工产业，主要产品以“烧碱、环氧丙烷、聚醚”为主，形成以三大产品为主的工业化生产格局。

长沙韶光依托承接军改后的体制外溢的优秀科研团队，瞄准特殊领域模拟市场，在各大军工集团等体制内科研院所林立的竞争中，公开竞标承揽了数十项国家科研项目，承接纵向课题获得经费和订单支持，同时形成成都、上海、沈阳三家研发中心，紧密贴近集团客户，为公司特种芯片产品库品类的充实奠定了坚实的基础。2021 年，公司专用数字芯片 (FPGA、GPU) 和通用类模拟芯片目录快速在体系内大单位院所客户放量。

图 67: 航锦科技业务板块架构

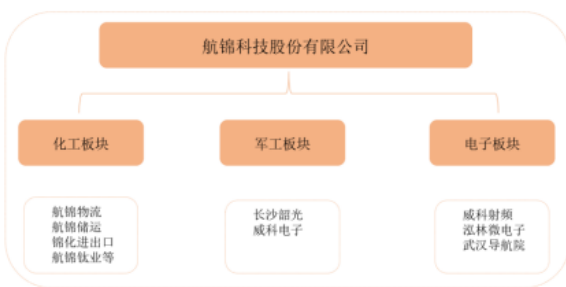


图 68: 寒武纪思元 270-S4 产品规格

产品系列	应用市场
总线控制类	设备类产品自主研发，可广泛应用于智能制造设备
网络控制类	网络核心器件，应用于工业设备的控制及传输领域
DSP 系列产品	可实现数字信号处理技术的集成电路，主要应用于雷达、导航等领域
FPGA 系列产品	灵活的逻辑门阵列，广泛应用于军事工业领域
配置存储器	配置存储器主要配合 FPGA 使用
CPUD	高性能低功耗高性能的数字集成电路，广泛应用于各个领域
AD	模拟信号转换为数字信号的电路，广泛应用于各个领域
DDS	直接数字频率合成电路
芯片产品	寒武纪芯片产品系列

资料来源: 公司官网

## 安路科技 (688107 CH) - 国内领先的 FPGA 芯片供应商

安路科技主营业务为 FPGA 芯片和专用 EDA 软件的研发、设计和销售，是国内领先的 FPGA 芯片供应商。公司主要向客户提供 FPGA 产品，包括 FPGA 芯片和专用 EDA 软件两部分。根据产品的性能特点与目标市场的应用需求，公司的 FPGA 芯片产品形成了由 SALPHOENIX 高性能产品系列、SALEAGLE 高效率产品系列、SALELF 低功耗产品系列组成的产品矩阵，FPSoC 产品包括面向工业和视频接口的低功耗 SALSWIFT 系列，实现了多种规格芯片和配套 EDA 软件的产品线覆盖，并持续致力于高容量高性能 FPGA 和高集成度 FPSoC 芯片的研发与拓展。

公司差异化定位的产品系列及不断丰富产品型号和应用参考设计使公司能够更好地覆盖下游客户的需求，产品已广泛应用于工业控制、网络通信、消费电子、数据中心等领域。

公司持续在 FPGA 硬件设计技术、FPGA 专用 EDA 软件技术、FPGA 测试技术、FPGA 应用技术等方面开展技术研发，相关技术申请发明专利 16 项。

图 69: FPGA 和 FPSoC 产品系列



资料来源：公司官网

## 复旦微电 (688385 CH) - FPGA 平台型公司

复旦微电是国内 FPGA 领域技术较为领先的公司之一，目前已可提供千万门级、亿门级 FPGA 芯片以及嵌入式可编程芯片 (PSoC) 等系列的产品。公司 FPGA 产品线拥有系列化超大规模异构融合可编程逻辑器件系列产品，率先研制成功了亿门级 FPGA 和异构融合可编程片上系统 (PSoC) 芯片，以及面向人工智能应用的融合现场可编程 (FPGA) 和人工智能 (AI) 的可重构芯片 (FPAI)。

公司累计向超过 500 家客户销售相关 FPGA 产品，在通信领域、工业控制领域等得到广泛应用。公司开发的致力于完整可编程器件开发流程的工具软件 ProciseTM，可支持公司全系列可编程器件，突破了 FPGA 配套 EDA 软件的多项核心关键技术。



## 兆易创新 (603986 CH) - DRAM 龙头企业，自有品牌广泛应用

兆易创新业务是多赛道多产品线的组合布局，目前主要是存储器、微控制器和传感器三大类；存储器又分为 SPI NOR、SLC NAND 和 DRAM，微控制器包括 ARM 核和 RISC-V 开源内核，传感器包括触控和指纹识别芯片。产品广泛应用于工业、消费类电子、汽车、物联网、计算、移动应用以及网络和电信行业等各个领域，助力社会智能化升级。

公司与长鑫存储密切合作，首款自有品牌 DRAM 产品 (4Gb DDR4) 已于 2021 年 6 月推出，实现了从设计、流片，到封测、验证的全国产化，在满足消费类市场强劲需求的同时，助力国产自主供应生态圈的发展构建。该产品主要面向消费类、工业控制类及汽车类等市场领域，应用于机顶盒、电视、监控、网络通信、智慧家庭、平板电脑、车载影音系统等诸多领域。在利基市场，公司 DRAM 产品在工艺制程上保持代差优势，有利于降低产品成本。公司与长鑫存储的紧密合作关系，为公司 DRAM 产品提供稳定产能保障。同时，依托于多年积累的、完善的销售网络和技术团队，公司能够为客户提供快速的本地化服务响应和技术支持。

图 70: 兆易创新首款自有品牌 DRAM 产品-4Gb DDR4



资料来源：公司官网

图 71: 针对利基型 DRAM 市场推出 DDR3L 产品



资料来源：公司官网

## 免责声明及披露

### 分析员声明

负责撰写本报告的全部或部分之分析员，就本报告所提及的证券及其发行人做出以下声明：（1）发表于本报告的观点准确地反映有关于他们个人对所提及的证券及其发行人的观点；（2）他们的薪酬在过往、现在和将来与发表在报告上的观点并无直接或间接关系。

此外，分析员确认，无论是他们本人还是他们的关联人士（按香港证券及期货事务监察委员会操作守则的相关定义）（1）并没有在发表研究报告 30 日前处置或买卖该等证券；（2）不会在发表报告 3 个工作日内处置或买卖本报告中提及的该等证券；（3）没有在有关香港上市公司内任职高级人员；（4）并没有持有有关证券的任何权益。

招银国际环球市场或其他关联机构曾在过去 12 个月内与报告内所提及发行人有投资银行业务的关系。

### 招银国际环球市场投资评级

买入	: 股价于未来 12 个月的潜在涨幅超过 15%
持有	: 股价于未来 12 个月的潜在变幅在-10%至+15%之间
卖出	: 股价于未来 12 个月的潜在跌幅超过 10%
未评级	: 招银国际证券并未给予投资评级

### 招银国际环球市场行业投资评级

优于大市	: 行业股价于未来 12 个月预期表现跑赢大市指标
同步大市	: 行业股价于未来 12 个月预期表现与大市指标相若
落后大市	: 行业股价于未来 12 个月预期表现跑输大市指标

### 招银国际环球市场有限公司

地址: 香港中环花园道 3 号冠君大厦 45 楼

电话: (852) 3900 0888

传真: (852) 3900 0800

招银国际环球市场有限公司("招银国际环球市场")为招银国际金融有限公司之全资附属公司(招银国际金融有限公司为招商银行之全资附属公司)

### 重要披露

本报告内所提及的任何投资都可能涉及相当大的风险。报告所载数据可能不适合所有投资者。招银国际环球市场不提供任何针对个人的投资建议。本报告没有把任何人的投资目标、财务状况和特殊需求考虑进去。而过去的表现亦不代表未来的表现，实际情况可能和报告中所载的大不相同。本报告中所提及的投资价值或回报存在不确定性及难以保证，并可能会受目标资产表现以及其他市场因素影响。招银国际环球市场建议投资者应该独立评估投资和策略，并鼓励投资者咨询专业财务顾问以便作出投资决定。

本报告包含的任何信息由招银国际环球市场编写，仅为本公司及其关联机构的特定客户和其他专业人士提供的参考数据。报告中的信息或所表达的意见皆不可作为或被视为证券出售要约或证券买卖的邀请，亦不构成任何投资、法律、会计或税务方面的最终操作建议，本公司及其雇员不就报告中的内容对最终操作建议作出任何担保。我们对因依赖本报告所载资料采取行动而引致之任何直接或间接的、疏忽、违约、不谨慎或各类损失或损害承担任何的法律或赔偿责任。任何使用本报告信息所作的投资决策完全由投资者自己承担风险。

本报告基于我们认为可靠且已经公开的信息，我们力求但不担保这些信息的准确性、有效性和完整性。本报告中的资料、意见、预测均反映报告初次公开发布时的判断，可能会随时调整，且不承诺作出任何相关变更的通知。本公司可发布其它与本报告所载资料及/或结论不一致的报告。这些报告均反映报告编写时不同的假设、观点及分析方法。客户应该小心注意本报告中所提及的前瞻性预测和实际情况可能有显著区别，唯我们已合理、谨慎地确保预测所用的假设基础是公平、合理。招银国际环球市场可能采取与报告中建议及/或观点不一致的立场或投资决定。

本公司或其附属关联机构可能持有报告中提到的公司所发行的证券头寸并不时自行及/或代表其客户进行交易或持有该等证券的权益，还可能与这些公司具有其他投资银行相关业务联系。因此，投资者应注意本报告可能存在的客观性及利益冲突的情况，本公司将不会承担任何责任。本报告版权仅为本公司所有，任何机构或个人于未经本公司书面授权的情况下，不得以任何形式翻版、复制、转售、转发及或向特定读者以外的人士传阅，否则有可能触犯相关证券法规。

如需索取更多有关证券的信息，请与我们联系。

#### 对于接收此份报告的英国投资者

本报告仅提供给符合(I)不时修订之英国 2000 年金融服务及市场法令 2005 年(金融推广)令("金融服务令")第 19(5) 条之人士及(II) 属金融服务令第 49(2) (a) 至(d) 条(高净值公司或非公司社团等)之机构人士，未经招银国际环球市场书面授权不得提供给其他任何人。

#### 对于接收此份报告的美国投资者

招银国际环球市场不是在美国的注册经纪交易商。因此，招银国际环球市场不受美国就有研究报告准备和研究分析员独立性的规则的约束。负责撰写本报告的全部或部分内容之分析员，未在美国金融业监管局("FINRA")注册或获得研究分析师的资格。分析员不受旨在确保分析师不受可能影响研究报告可靠性的潜在利益冲突的相关 FINRA 规则的限制。本报告仅提供给美国 1934 年证券交易法(经修订)规则 15a-6 定义的"主要机构投资者"，不得提供给其他任何人。接收本报告之行为即表明同意接受协议不得将本报告分发或提供给任何其他人。接收本报告的美国收件人如想根据本报告中提供的信息进行任何买卖证券交易，都应仅通过美国注册的经纪交易商来进行交易。

#### 对于在新加坡的收件人

本报告由 CMBI (Singapore) Pte. Limited (CMBISG) (公司注册号 201731928D) 在新加坡分发。CMBISG 是在《财务顾问法案》(新加坡法例第 110 章)下所界定，并由新加坡金融管理局监管的豁免财务顾问公司。CMBISG 可根据《财务顾问条例》第 32C 条下的安排分发其各自的外国实体，附属机构或其他外国研究机构编制的报告。如果报告在新加坡分发给非《证券与期货法案》(新加坡法例第 289 章)所定义的认可投资者，专家投资者或机构投资者，则 CMBISG 仅会在法律要求的范围内对这些人士就报告内容承担法律责任。新加坡的收件人应致电 (+65 6350 4400) 联系 CMBISG，以了解由本报告引起或与之相关的事宜。