

2023年03月25日

# AIGC 从算法+算力已到场景！谁是“格局型”标的？

——计算机行业周报 20230320-20230324

看好

## 相关研究

"国产 AI 芯片的百倍算力需求！-AIGC 系列之七" 2023 年 3 月 24 日  
"一层是 AIGC 狂潮，深层是数字经济出海！- 计算机行业周报 20230313-20130317" 2023 年 3 月 18 日

## 证券分析师

刘洋 A0230513050006  
liuyang2@swsresearch.com  
洪依真 A0230519060003  
hongyz@swsresearch.com  
黄忠煌 A0230519110001  
huangzh@swsresearch.com  
宁柯瑜 A0230520070005  
ningky@swsresearch.com  
施鑫展 A0230519080002  
shixz@swsresearch.com  
胡雪飞 A0230522120002  
huxf@swsresearch.com

## 联系人

胡雪飞  
(8621)23297818x  
huxf@swsresearch.com

## 本期投资提示：

- **本周周报包括三部分：**1) 发布 AIGC 深度 7 《**国产 AI 芯片的百倍算力需求**》，探讨 AIGC 下芯片和算力需求。2) 我们认为，从 AI 算法、AI 算力，再到 AI 场景应用，会较为重要 (ChatGPT 中实现了对插件的初始支持是催化剂)。一图寻找到国内、国外的 AIGC 场景重要公司。3) 由于标的扩散，投资者关心，哪些才是 AIGC 浪潮的“格局型”标的？
- **首先，本周发布 AIGC 深度 7 《国产 AI 芯片的百倍算力需求》，探讨 AIGC 下芯片和算力需求。**1) **AI 芯片是大模型的基础。**AI 芯片主要分为训练和推理两类芯片，云和端对训练和推理芯片要求不同，目前主流架构包括 GPU、FPGA 和 ASIC 三类，通用性 GPU>FPGA>ASIC，性能功耗比 GPU<FPGA<ASIC。2) **四大 AI 芯片技术路线，均围绕打破英伟达 CUDA 生态垄断展开。**英伟达凭借 CUDA、cuDNN 和 TensorRT 等软件工具链以及和 Tensorflow 的深度绑定构筑了极高的生态壁垒，2021 年 GPU 市占率超 80%，高性能 AI 芯片 A100、H100 被禁止向中国出口后，快速推出 800 系列合法出口中国；寒武纪复制英伟达成长之路。3) AIGC 传导至潜在的百倍算力需求！**AI 芯片+AI 服务器，受益于 AIGC+类 GPT 等应用的“鲑鱼效应”。**4) **重点推荐：1) 直接受益 AI 芯片需求崛起的寒武纪**（思元 590，接近英伟达 A100 性能的国产 AI 训练芯片）、**海光信息**（预计 2023 年底量产的 DCU3 性能与 NVIDIA H100 性能持平）。2) **AI 服务器供应商：浪潮信息、中科曙光、中兴通讯（通讯）、工业富联（电子）。**也建议关注通信、电子公司。
- **其次，我们认为，从 AI 算法、AI 算力，再到 AI 场景应用，会较为重要。ChatGPT 中实现插件支持是催化剂。一图寻找到国内、国外的 AIGC 场景重要公司。**1) AIGC 场景中，2C 包括搜索/办公/游戏/NLP 等，2B 包括视觉 CV/医疗/教育/工业/汽车/金融等。海外巨头微软/谷歌/脸书/特斯拉/deepmind 等已分别布局，英伟达等提供算力。2) 依照场景，国内**金山办公+汉得信息**（办公）、**360+百度**（申万互联网，搜索）、**安防双杰+虹软**（视觉 CV 等）、**科大讯飞**等（NLP）、**润达医疗**等（申万医药等，医学）、**中控技术+汇川技术**（申万机械）、**同花顺+东方财富**（金融）、**广联达**（建筑）、**石基信息**（酒旅）等，都可能会是 AIGC 的核心公司。而底层算力的代表是**寒武纪+海光+浪潮**。3) 正文有详解图。
- **最后，谁是“AIGC 浪潮”的“格局型”标的？**讨论**金山办公、AI 安防双雄、三六零（申万传媒）、寒武纪等**（潜在**海光信息/浪潮信息**等）。也讨论潜在“格局型”标的，建筑领域世界领军**广联达**（数字建筑 AI 化的领军）、酒旅世界领军**石基信息**（云产品包含 AI 功能）。
- 数字经济领军：**金山办公、广联达**（云与 AI 结合）、**石基信息**（也是 AI）、**中控技术**（tmt&机械，**中东突破**）、**赛意信息、启明星辰、英方软件、AI 领军、恒生电子、德赛西威**。
- AIGC 算力：**寒武纪、海光信息、浪潮信息、中科曙光等**
- AIGC&数据弹性：**润达医疗**（医药）、**大华股份、AI 领军、广电运通、虹软科技、上海钢联、汉得信息、科大讯飞、深桑达、云赛智联、拓尔思**。
- 信创弹性：**太极股份、深信服、纳思达、深桑达、中国软件、海量数据、软通动力**
- **风险：**由于复工/宏观环境等扰动，2022-2023 年内存在业绩波动风险。



申万宏源研究微信服务号

本周周报包括三部分：

- 1) 发布 AIGC 深度 7 《国产 AI 芯片的百倍算力需求》，探讨 AIGC 下芯片和算力需求。
- 2) 我们认为，从 AI 算法、AI 算力，再到 AI 场景应用，会较为重要(ChatGPT 中实现了对插件的初始支持是催化剂)。一图寻找到国内、国外的 AIGC 场景重要公司。
- 3) 由于标的扩散，投资者关心，哪些才是 AIGC 浪潮的“格局型”标的？

## 1、AI 三要素之算力

本周发布 AIGC 深度 7 《国产 AI 芯片的百倍算力需求》，探讨 AIGC 下芯片和算力需求。

表 1:AIGC 系列多篇报告与 AI 大模型系列深度报告

序列	时间	标题	备注
	2022 年 11 月		
AIGC 系列一		《游戏政策面转暖，元宇宙 AIGC 催化》	游戏 AIGC
AIGC 系列二	2023 年 1 月	《微软投资 openAI 一小步，AI 与 chat GPT 一大步》	AI 算法/商业原因 大模型/算法/场景等 FAQ 问题
AIGC 系列三	2023 年 2 月	《ChatGPT/AIGC 九问九答》	
AIGC 系列四	2023 年 2 月	《微软产品融合 ChatGPT，全球巨头布局梳理》	微软和全球巨头
AIGC 系列五	2023 年 3 月	《跨模态：更多应用场景出现，中国公司得到更大机会》	大模型/跨模态
AIGC 系列六	2023 年 3 月	《百度文心：一个符合预期的起点》	百度文心
AIGC 系列七	2023 年 3 月	《国产 AI 芯片的百倍算力需求》	AI 芯片/AI 算力
序列	时间	标题	备注
	2022 年 12 月		
大模型深度		《大模型技术突破，AI 多场景落地——2023 AI 行业投资策略》	大模型技术突破
	2021 年 12 月		
大模型深度		《领军路径分歧：更大的模型还是更低的成本？》	大模型深度

资料来源：Wind,申万宏源研究

**AI 芯片是大模型的基础。**AI 芯片主要分为训练和推理两类芯片，云和端对训练和推理芯片要求不同，目前主流架构包括 GPU、FPGA 和 ASIC 三类，通用性 GPU>FPGA>ASIC，性能功耗比 GPU<FPGA<ASIC。

**四大 AI 芯片技术路线，均围绕打破英伟达 CUDA 生态垄断展开。**英伟达凭借 CUDA、cuDNN 和 TensorRT 等软件工具链以及和 Tensorflow 的深度绑定构筑了极高的生态壁垒，2021 年 GPU 市占率超 80%，高性能 AI 芯片 A100、H100 被禁止向中国出口后，快速推出 800 系列合法出口中国；寒武纪复制英伟达成长之路。

**GPT-4 参数量高达 100 万亿，是 GPT-3 的 500 倍以上！**即使考虑到 AI 芯片能力的提升（从当前主流的 A100 升级至 H100），仍然需要巨量的额外增量投资。

**AI 芯片+AI 服务器，受益于 AIGC+类 GPT 等应用的“鲶鱼效应”。**原有英伟达等供给有限，所以国产 AI 芯片有理论上的爆发弹性，AI 服务器有成长空间。由于 AIGC、类 GPT 有“鲶鱼效应”，带来约百倍算力需求。而英伟达等供给解决需求有瓶颈（根据 IDC 咨询，预测 2025 年 AI 服务器市场空间仅仅 318 亿美元，21-25 年预计 CAGR 仅仅 19.5%），因此国产 AI 芯片在逻辑上有爆发弹性，此外 AI 服务器也有成长空间。

**寒武纪思元 590 将是最早实现商业应用的接近英伟达 A100 性能的国产 AI 训练芯片。**目前华为昇腾 910 性能超越英伟达 V100，但未达到 A100 水平，壁仞科技 7nm 通用 GPU 芯片 BR100 称其可与被禁售的英伟达 H100 一较高下，但尚未量产上市。寒武纪思元 590 芯片面积 800mm<sup>2</sup>，和 A100 一样。内存带宽 2.7T，是 A100 1.8T 的 1.5 倍。HBM2 使用海力士，功耗达 350W-550W，FP32 算力到 80TFLOps，目前已经客户送样测试阶段，在高性能国产 AI 芯片中进程最快，最有机会承接国内 AI 算法商对英伟达 A100、H100 的需求。

**重点推荐：1）直接受益 AI 芯片需求崛起的寒武纪**（思元 590 将是最早实现商业应用的接近英伟达 A100 性能的国产 AI 训练芯片）、**海光信息**（预计 2023 年底量产的 DCU3 性能与 NVIDIA H100 性能持平）。2）**AI 服务器供应商：浪潮信息、中科曙光、中兴通讯（通讯）、工业富联（电子）。**

**TMT 领域算力基建逻辑，也建议关注通信、电子相关公司。**

**通信**——天孚通信、中际旭创、紫光股份、锐捷网络、新易盛等。此外也持续关注流量基建主线的数据中心等环节供需变化，包括奥飞数据、润泽科技等，以及伴随算力增长的温控散热环节，英维克等。

**电子**——GPU：景嘉微、好利科技；视觉芯片及 IP：思特威、富瀚微、芯原股份；服务器存储及接口芯片：澜起科技、聚辰股份；CPU：海光信息、龙芯中科；FPGA 芯片：复旦微、安路科技；数通/服务器 PCB：沪电股份、深南电路、生益科技、胜宏科技；光通信芯片：源杰科技。

**风险提示：AI 芯片迅速迭代，开启价格战；AIGC 行业落地需求不及预期；供应链不稳定**

## 1.1 AI 芯片：大模型的基础

**人工智能芯片主要分为“训练（Training）”芯片和“推理（Inference）”芯片。**

“训练芯片”主要用于人工智能算法训练，即在云端将一系列经过标记的数据输入算法模型进行计算，不断调整优化算法参数，直至算法识别准确率达到较高水平。“推理芯片”主要用于人工智能算法推理，即将在云端训练好的算法模型进行裁剪优化变“轻”之后，进入“实战”阶段，输入数据直接得出准确的识别结果。

**不同用途（训练 or 推理）、不同应用场景（端-边-云）对 AI 芯片有着不同的要求。**首先，训练芯片追求的是高计算性能（高吞吐率）、低功耗，但是推理芯片主要追求的是低延时（完成推理过程所需要的时间尽可能短）、低功耗。其次，“端-边-云”三个环节

对 AI 芯片的不同要求见下表——其中端和边上进行的大部分是 AI “推理”，因此用于端和边的 AI 芯片性能要求和上述推理芯片一致；大部分的训练过程是在云和数据中心进行，训练过程对时延没有什么要求，因此需要保证 AI 芯片在尽可能保证较高算力的情况下，功耗尽可能低，另外许多推理过程也是在云端进行。

**表 2：“端-边-云”对 AI 芯片的算力和功耗、延时性有不同的要求**

	端			边			云
	耳机电话	智能手机	个人电脑	网络摄像机 IPC	边缘服务器	数据中心	
算力	20MOPS	100GOPS	1-10TOPS	10-20TOPS	10-20TOPS	10-500TOPS	200+TOPS
功耗	1 mW	10 mW	1-2W	3-10W	3-10W	10-300W	200+W
模型大小	10KB	100KB	10MB	10-100MB	10-100MB	100+MB	300+MB
延时	<10 ms	~10ms	10-100ms	10-500ms	10-500ms	ms~s	ms~s
是否推理	Y	Y	Y	Y	Y	Y	Y
是否训练	N	N	Y	Y	Y	Y	Y
昇腾系列	Nano	Tiny	Lite	Mini	Mini	Multi-Mini or Max	Max or Multi Mini

资料来源：ofweek 电子工程网，华为全联接大会，申万宏源研究

**图 1：云端训练、云端推理、设备端推理三个细分市场的 AI 芯片竞争格局**



资料来源：雷锋网，申万宏源研究

## 1.2 AI 芯片四大技术路线，寒武纪复制英伟达

由于 AIGC、类 GPT 应用有鲰鱼效应，带来约百倍算力需求。而英伟达等供给解决需求有瓶颈，因此国产 AI 芯片有逻辑上需求弹性，AI 服务器也有空间。根据 IDC 数据，2021 年全球 AI 服务器市场规模为 156 亿美元，预计到 2025 年全球 AI 服务器市场将达到 318 亿美元，预计 21-25 年 CAGR 仅仅 19.5%。AI 服务器的增长和规模总额恐怕无法满足类 GPT 类应用的百倍需求（例如生产地域、供应商产能、工人等限制），因此 AI 芯片可能会大量爆发，其次是 AI 服务器。

近期的行业领袖创业潮，会加速这种趋势。2012-2014 年 AI 创业潮，造就 2015-2017 年 AI 机会。2022H2-2023 新一轮 AI 大模型创业潮。

目前 AI 芯片主要玩家应对英伟达塑造的 AI 生态壁垒，选取了不同的商业策略：

- （1）英伟达 AI 芯片依然是 AI 训练和推理最佳选择；
- （2）寒武纪在走英伟达的路线；
- （3）AMD 在走部分兼容 CUDA 的路线；
- （4）谷歌、华为、百度走的是“深度学习框架+AI 芯片”自研路线。

## 1.3 AIGC 传导至潜在的百倍算力需求！

当前市场关注 AIGC 所拉动的增量算力需求及相应投资机会。此前我们指出，AI 必然伴随海量计算，算力是完成数据运算的基础，chatGPT 一共有 1750 亿个参数，训练过程中消耗大量算力，其训练模型中就已导入至少 1 万颗英伟达 GPU。AI 发展的基础是算力提升，AI 应用的拓展反向推动算力基建的部署。此前 ChatGPT 宣布因满负荷运行暂停服务，算力基建部署值得关注。

推理侧，ChatGPT 海量的访问需求带来了巨大的计算成本。我们根据公开数据对此进行估计：

根据 OpenAI 首席执行官 Sam Altman 提供的信息，ChatGPT 单次聊天的成本约为 0.01 美元。ChatGPT 目前有超过 2000 万日活且在持续上行，假设未来日活数量达到 5000 万、平均每个日活进行 10 条对话，则每日消耗的成本为 500 万美元，推测其中大部分为 AI 服务器成本（折旧、维护等）及对应的电费。

图 2：OpenAI 首席执行官 Sam Altman 谈 ChatGPT 每次聊天的成本



资料来源：Twitter，申万宏源研究

训练侧，由于大型语言模型（LLM）参数量极高，训练成本与参数量正相关，因此大模型路线势必带动大量额外的算力需求。参考海外最新研究《The Economics of Large Language Models》<sup>1</sup>，可以有如下估计：

- 每个 token ( 1000 token 约等于 750 个单词 ) 的训练成本通常约为 6N ( 而推理成本约为 2N ) ，其中 N 是 LLM 的参数数量；
- 假设在训练过程中，模型的 FLOPS 利用率为 46.2%，与在 TPU v4 芯片上进行训练的 PaLM 模型（拥有 5400 亿参数）一致；
- 1750 亿参数模型的 GPT-3 是在 3000 亿 token 上进行训练的，成本参考谷歌在 GCP TPU v4 芯片上训练 PaLM 模型。
- 计算结果：GPT-3 训练成本为 140 万美元左右。

图 3：GPT-3 训练成本估算

$$\text{cost of training} = \frac{\overbrace{175\text{B model parameters} \times \frac{6 \text{ FLOPs}}{\text{token} \cdot \text{model parameter}}}_{\text{FLOPs per token for training the model}}}{\underbrace{\frac{4 \text{ TPUs}}{\text{machine}} \times \frac{275 \text{ peak TFLOPs}}{\text{TPU}}}_{\text{realizable FLOPs per second for each machine}}} \times \frac{\text{realizable TFLOPs}}{\text{peak TFLOPs}} \times \frac{\$8.12}{\text{hour} \cdot \text{machine}} \times \frac{1 \text{ hour}}{3600 \text{ seconds}} \times 300\text{B tokens} = \$1.398\text{M}$$

资料来源：《The Economics of Large Language Models》，申万宏源研究

将上述计算框架应用于其他的 LLM 模型中，可以估算出对应的训练成本。其中参数量最大的谷歌 PaLM 模型，训练成本超过 1000 万美元。

图 4：主流 LLM 模型训练成本估算

<sup>1</sup> <https://sunyan.substack.com/p/the-economics-of-large-language-models>

	GPT-3 (OpenAI)	Gopher (Google DeepMind)	MT-NLG (Microsoft/Nvidia)	PaLM (Google Research)
Model Parameters	175B	280B	530B	540B
FLOPs/Token/Model Parameter			6	
TPUs/Machine			4	
Peak FLOPs/TPU			275T	
FLOPs Utilization			46.2%	
Cost/Machine/Hour (1-year reserved)			\$8.12	
Seconds/Hour			3600	
Training Cost/1000 Tokens	\$0.0047	\$0.0075	\$0.0141	\$0.0144
Train Tokens	300B	300B	270B	780B
Training Cost	\$1,398,072	\$2,236,915	\$1,810,744	\$11,216,529

资料来源：《The Economics of Large Language Models》，申万宏源研究

考虑到未来 LLM 不断升级演进，训练/推理成本将持续上行。相应的成本估算框架如下图所示。可见成本与模型参数数量、模型训练数据集大小、Cost/FLOP 效率成正比，并与计算资源和硬件利用率成反比。

图 5：LLM 训练/推理成本测算框架



资料来源：《The Economics of Large Language Models》，申万宏源研究

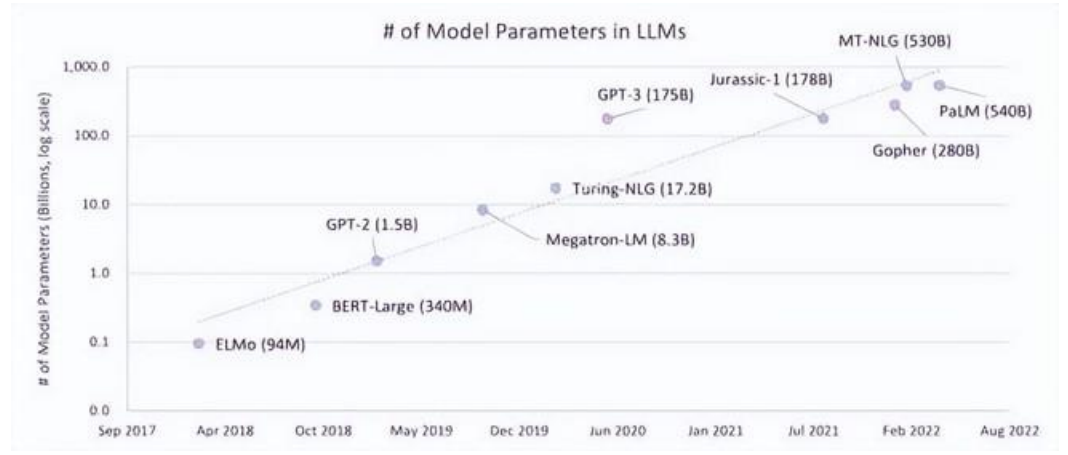
超越“安迪-比尔”定律<sup>2</sup>的算力需求爆发！模型参数量每年 10 倍增长，即使 AI 芯片的性能提升仍然遵从摩尔定律<sup>3</sup>，计算资源仍然可能难以满足需求。过去 5 年中，LLM 模型参数呈现指数增长态势。2018 年，OpenAI 推出了 1.17 亿参数的 GPT-1，谷歌推出了 3 亿参数的 BERT；接下来的两年（2019、2020 年），OpenAI 陆续推出参数更大的迭代版本 GPT-2、GPT-3，前者有 15 亿参数，后者有 1750 亿参数。

GPT-4 参数量高达 100 万亿，是 GPT-3 的 500 倍以上！即使考虑到 AI 芯片能力的提升（从当前主流的 A100 升级至 H100），仍然需要巨量的额外增量投资。

<sup>2</sup> 原话是 “Andy gives, Bill takes away. (安迪提供什么，比尔拿走什么。)” 安迪指英特尔前 CEO 安迪·格鲁夫，比尔指微软前任 CEO 比尔·盖茨，这句话的意思是，硬件提高的性能，很快被软件消耗掉了。

<sup>3</sup> 集成电路上可以容纳的晶体管数目在大约每经过 18 个月到 24 个月便会增加一倍。

图 6：过去 5 年中，LLM 模型参数呈指数增长



资料来源：《The Economics of Large Language Models》，申万宏源研究

## 1.4 标的：若仅考虑赛道，寒武纪/海光等卡位佳

目前，国内 BAT、字节、京东等大厂均投入 AIGC 大模型研发，模型的应用场景也从文字对话进一步扩展至绘画、视频等。多玩家、多场景、丰富的下游应用带来百倍的增量 AI 算力需求，算力基建直接受益！计算机行业核心关注 AI 芯片、服务器等环节。

### AI 芯片建议关注：

**寒武纪**：寒武纪思元系列产品适配 TensorFlow、Pytorch、Caffe 深度学习框架。2019 年开始适配 AI 巨头，峰值时刻合作开发团队有 70-80 人（公司派出 20-30 人），思元 290 与商汤在 CV 层面深度合作，NLP 领域在讯飞、百度语音都有出货。**寒武纪思元 590 将是最早实现商业应用的接近英伟达 A100 性能的国产 AI 训练芯片。**

**海光信息**：作为 AI 服务器中的核心，海光的 DCU 是 GPGPU 的一种。根据公司公开发布会，海光 DCU 同时支持全精度和半精度训练，有稀缺性。预计 2023 年底量产的 DCU3 性能与 NVIDIA H100 性能持平。已量产的 DCU2 约等于 70% NVIDIA A100，双方在模拟效率和精度都处于同一代际。算力需求提升叠加国产替代要求，共同构成催化。

### AI 服务器建议关注：

**浪潮信息**：国内最大的 AI、边缘和云服务器供货商，根据 IDC 数据，浪潮 AI 服务器出货量基本占据国内一半的市场份额，并且多年稳定。chatGPT 为代表的 AIGC 大模型风靡，对于算力的要求会明显提升，AI 服务器作为算力的基础、核心基建，深度受益。浪潮信息的主要客户也是互联网和 AI 巨头，在经历弱周期后，大厂算力竞争有望带来 AI 服务器投入显著回升。此外，根据国家信息中心和浪潮信息联合发布的报告，目前我国有超过 30 个城市正在或提出建设智算中心，提供增量。

**中科曙光**：在国内超算市场中处于领先地位，超算作为国家级重点项目的支撑，有望在算力的竞赛中承担重要角色，东数西算在内的政策形成催化，预计曙光也将在 AI 服务器以及超算中心加速中受益。



此外，中兴通讯（通信）、工业富联（电子）也具备 AI 服务器能力。

TMT 领域算力基建逻辑，也建议关注：

**通信**：天孚通信、中际旭创、紫光股份、锐捷网络、新易盛等。此外也持续关注流量基建主线的数据中心等环节供需变化，包括奥飞数据、润泽科技等，以及伴随算力增长的温控散热环节，英维克等。

**电子**：GPU：景嘉微、好利科技；视觉芯片及 IP：思特威、富瀚微、芯原股份；服务器存储及接口芯片：澜起科技、聚辰股份；CPU：海光信息、龙芯中科；FPGA 芯片：复旦微、安路科技；数通/服务器 PCB：沪电股份、深南电路、生益科技、胜宏科技；光通信芯片：源杰科技。

## 2、AI 三要素之场景

我们认为，从 AI 算法、AI 算力，再到 AI 场景应用，会较为重要。一图寻找到国内、国外的 AIGC 场景重要公司。

### 2.1 AIGC 场景一览图

AIGC 场景中，2C 包括搜索/办公/游戏/NLP 等，2B 包括视觉 CV/医疗/教育/工业/汽车/金融等，也应当包括酒旅与建筑。

其中，海外巨头微软/谷歌/脸书/特斯拉/deepmind 等已分别布局，英伟达等提供算力支撑。

依照场景，国内的金山办公+汉得信息（办公）、360+百度（申万互联网，搜索）、安防双杰+虹软（视觉 CV 等）、科大讯飞等（NLP）、润达医疗等（申万医药等，医学）、中控技术+汇川技术（申万机械，工业）、同花顺+东方财富（金融）、广联达（建筑）、石基信息（酒旅）等，都可能会是 AIGC 的核心公司。而底层算力的代表是寒武纪+海光+浪潮。

**图 7：AI 从算法、算力到场景应用。绘制了 AIGC+AI 下，各自场景的海外/中国代表性公司**

欧美		部分2C (请参照申万宏源互联网传媒)				2B							
欧美	基础设施	搜索	办公	游戏	NLP	视觉CV	医疗	教育	工业	汽车	金融	建筑	酒旅
英伟达等	算力 芯片+服务器	bing接入chatGPT	office接入chatGPT										
微软			新AI功能Copilot										
微软				Azure OpenAI服务									
微软					与openAI								
谷歌+脸书 deepmind						VIT+mae+ diffusion算法				Alphafold等	药物开发		
特斯拉						PaLM-E是视觉 语言模型							
特斯拉+谷歌										机器人+PaLM-大模型如BEV			
Adobe											AI模型集: Adobe Firefly		
Oracle													Oracle AI服
中国		部分2C (请参照申万宏源互联网传媒)				2B							
中国	基础设施	搜索	办公	游戏	自然语言处理	视觉CV	医疗	教育	工业	汽车	金融	建筑	酒旅
寒武纪+海光+浪潮等	算力 芯片+服务器												
360+百度		新一代搜索+文心一言											
金山办公等			探索结合AIGC、LLM										
汉得信息等			ERP接入百度+微软										
A股游戏公司				A股游戏公司									
科大讯飞等					NLP等技术								
安防AI双杰+虹软等						世界领军的CV环境							
润达医疗等						AI诊断检验							
中控技术+ 汇川技术(机械)等									中控工业AI InPlant IBD等				
造车新势力等													
同花顺/东方财富等										AI问答机器人 +金融数据智 能等			
广联达											AI平台+NLP		
石基信息													云PMS+AI

注 1：根据金山办公互动平台发布，探索结合 AIGC、LLM 技术的下一代人机交互体验，在文档翻译、听读、中英文校对、语音及音频转写、智能辅助写作及排版、表格数据智能分析、PPT 一键生成及美化等场景进行融合

注 2：根据科大讯飞互动平台发布：“N”是应用于教育、医疗、人机交互、办公、翻译、工业等多个行业领域的专用大模型

注 3：东方财富互动平台发布，已经陆续研发了东方财富金融数据 AI 智能化生产平台、多媒体智能资讯及互动平台系统等多个人工智能相关项目

注 4：领域橙色表明国家禀赋，即这个领域的应用优势大

资料来源：各公司历年年报，申万宏源研究

## 2.2 场景：办公+OA，chatGPT 上线插件是催化剂

3月23日，Open AI 推出 ChatGPT 插件功能。ChatGPT 能够连接到第三方应用程序并与开发人员定义的 API 进行交互，从而增强 ChatGPT 的功能并允许执行更广泛的操作，例如查询股票价格、查询公司文件、执行订机票等。第一批上线插件包括了 OpenAI 官方发布的三个应用，以及和第三方合作的 12 款应用。

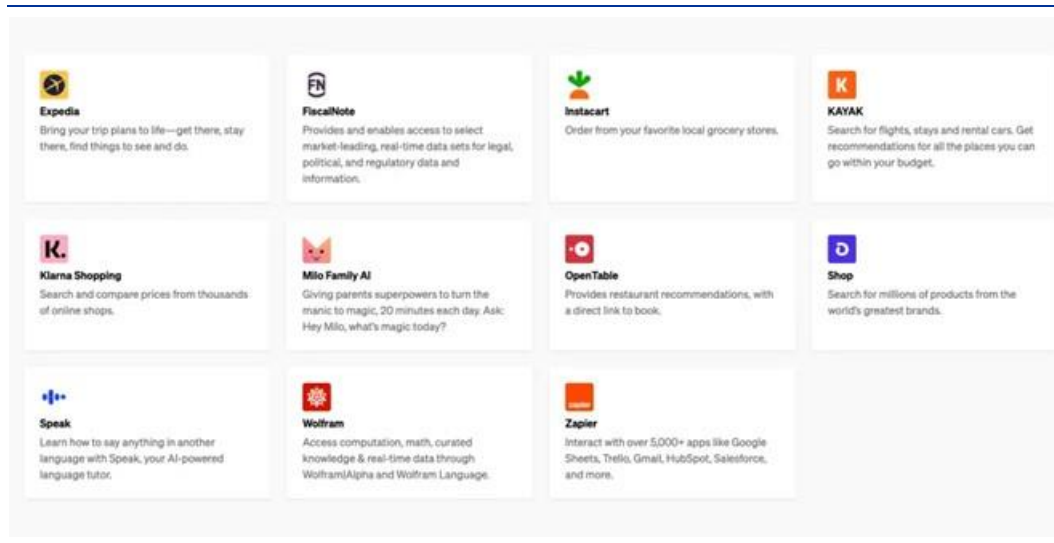
加入插件功能后，ChatGPT 的能力再次跃升。ChatGPT 插件主要由检索实时信息、检索知识库信息、代表用户执行操作三大功能组成。例如，搭载插件后只需要告诉 ChatGPT 要订哪个城市的酒店，入住时间和离店时间，它就会帮找到符合条件的酒店并预订。其他能力包括：

**检索实时信息：**查询体育比赛比分、查询股票价格、查询最新消息等。

**检索知识库信息：**检索公司文件、检索个人笔记等。

**代表用户执行操作：**自动订机票、订餐等。

**图 8：ChatGPT 上线的首批插件**



资料来源：OpenAI，申万宏源研究

**表 3：ChatGPT 首批插件及其功能**

插件名称	分类	功能
Expedia	商旅	将您的旅行计划变为现实 – 到达目的地，入住酒店，发现有趣的事情。
FiscalNote	实时数据集	提供并启用精选市场领先的实时数据集，用于法律，政治和监管数据和信息。
Instacart	消费	从您最喜欢的本地杂货店订购。
KAYAK	商旅	搜索航班，住宿和租车。获得有关您可以在预算范围内去哪些地方的推荐。
Klarna Shopping	消费	从数千个在线商店搜索和比较价格。
Milo Family AI	生活	赋予父母超能力，每天 20 分钟将忙碌变为魔术
OpenTable	消费	提供餐厅推荐，并直接链接预订。
Shop	消费	从世界上最伟大的品牌中搜索数百万产品。
Speak	学习	使用 Speak，您的 AI-powered 语言导师，学会如何用另一种语言说任何事情
Wolfram	学习	通过 Wolfram   Alpha 和 Wolfram Language 访问计算，数学，精选知识和实时数据。
Zapier	工具	与超过 5000 个应用程序交互，如 Google Sheets，Trello，Gmail，HubSpot，Salesforce 等

资料来源：爱范儿，申万宏源研究

**OA 系统有望在接入类 ChatGPT 的大模型后，功能大幅强化。**从 ChatGPT 首批插件看，面向商旅的应用是一大亮点，我们认为这会对 OA 领域带来显著变化。以出差为例，传统 OA 系统中员工的出差申请、机票预定、酒店预订、报销审批等大多需要分别操作，效率不高。但在结合 ChatGPT 能力后，员工只需要说出出差时间、目的地和偏好，即可由 AI 辅助完成行程规划和预定等一系列操作，节约了大量时间。

**此外，上周微软发布划时代的 copilot 产品：**

3月16日晚，微软举办了名为“The Future of Work with AI”的线上活动，展示了其在人工智能领域的最新成果。会上正式发布由AI驱动的Microsoft 365 Copilot，让Word、PPT、Excel、Outlook、协同办公软件的生产力倍增。

图9：微软发布颠覆传统办公方式的Microsoft 365 Copilot



资料来源：微软，申万宏源研究

微软 CEO 纳德拉表示：“这是人类与计算机共生的新时代开始。”。在 GPT-4 的支持下，微软最新发布的 Word 中，AI 能秒出草稿，并根据用户要求增删文字信息和配图；PowerPoint 中，AI 能快速将文字转换成专业水准的 PPT；Excel 中，AI 将数据分析变得轻松高效，能快速提炼出关键趋势；Outlook 中，AI 能给邮件分类加精，并自动撰写回复内容；协同办公时，AI 能总结规划成员的工作进展、调取分析数据、做 SWOT 分析、整理会议核心信息。同时，Copilot 可以在整个 Office 中调用，所有 Word、Excel、PPT 之类的办公软件都是互通的。

表 4：Microsoft 365 Copilot 利用 GPT-4 的 AI 能力，全面赋能办公软件

办公软件	能力	案例
Word	写文章、改文字、做总结；根据简短的提示创建初稿，并根据需要从整个组织中调入信息；正确把握文章中的语气，给出各种建议；改进写作，建议加强论证的方法，或者消除不一致的地方 在 Word 中调用其他软件的素材写文；根据 Word 内容做出 Excel，数据分析自动完成	根据「一份文件」和「一份电子表格」的数据，起草一份两页的项目建议书。 使第三段更简明。改变文件的语气，使其更加随意。 根据这个粗略的提纲，起草一页草稿。
PPT	根据想法自动做出一整套 PPT，美观的页面自动设计好；根据写好的 word，自己生成一份 PPT；浓缩冗长的 PPT，并使用自然语言命令去调整布局、重新编排文本、甚至卡点完美的动画时间	在 Word 文档的基础上创建一个 5 张幻灯片的演示文稿，并包括相关的图片资料。 将这个演示文稿合并成一个 3 张幻灯片的摘要。 将这 3 条摘要重新编排成三列，每列有一张图片。
Excel	生成清晰的表格，完成数据处理；根据数据即时地创建出基于 SWOT 的分析，或者各种数据透视表；发现数据的相关性，提出假设方案，以及根据用户提出的问题给出公式的建议，甚至生成新的模型	按类型和渠道给出销售明细，插入一个表格。 预测[一个变量的变化]的影响，并生成一个图表来帮助可视化。 模拟[变量]增长率的变化将如何影响我的毛利率。

Outlook	写邮件，根据开头自动补完，润色文字；汇总邮箱信息；用户可以指定它写什么语气的邮件，写多少个	总结一下我上周外出时错过的邮件，标记所有重要的项目。起草一份回复，在表达感谢的同时，询问第二和第三点的更多细节；缩短这份草稿的长度，并使用更加专业的语气。邀请大家来参加下周四中午关于新产品发布的「午餐和学习」活动，现场会提供午餐。
Teams	实时总结项目，同时执行任务，让整个会议效率大大提升；在开会时自动转录；日程自动提醒；Teams 上的信息随时可以同步	总结一下我在会议中错过的内容。到目前为止提出了哪些观点？我们在这个话题上哪里有分歧？为[正在讨论的话题]建立一个优势和劣势的表格。在做决定之前，我们还应该考虑什么？作出了哪些决定，有哪些建议的下一步行动？
Business Chat	汇集了所有来自 word、PPT、邮件、日历、笔记和联系人的数据；聊天记录总结；邮件和项目计划写作	总结一下昨晚发生的关于[客户]升级的聊天记录、电子邮件和文件。 [项目]的下一个里程碑是什么，有没有发现任何风险？帮我头脑风暴一下，列出一些潜在的缓解措施。 按照[文件名 A]的风格写一个新的计划概述，包含[文件名 B]中的计划时间表，并结合[某人]的电子邮件中的项目清单。
Power Platform	通过 Power Apps 和 Power Virtual Agents 这种低代码工具，让新手很容易地开发各种应用	

资料来源：微软，新智元，申万宏源研究

**AI 助手加速引入工作流程，释放巨大生产力。**无独有偶，此前在 GPT-4 发布的同一天，谷歌就针锋相对，宣布会将类 ChatGPT 的人工智能整合到自家办公「全家桶」Workspace 中。这些工具也同样具备让用户自动生成 PPT 等功能，未来还会覆盖更多语言和国家的用户。谷歌在全球有 30 多亿的 Google Workspace 用户，而 Office 365 被全球超过 100 万家公司使用。两家公司争相在办公软件中引入最新的 AI 技术，毫无疑问会加速推动办公效率的颠覆式升级、释放数亿人的生产力。

**图 10：谷歌近日宣布为其 Google Workspace 软件推出一系列 AI 工具**

**能提高工作生产力的文档处理线上软件**

除了提供文三处理、网站建立及应用程序所需的工具之外，更能多人同时审阅、同步处理，让生产力大大提升。



**文档**

方便团队协作处理文档

您无需安装专用软件，即可直接在浏览器中创建和编辑文本文档。多个用户可同时处理文件，且任何更改均会自动保存。



**表格**

安全便捷的协作式电子表格，为快速变化的单位提供支持

表格为满足敏捷型组织的需求而精心设计。借助基于人工智能技术的功能，您可以利用正确的信息，制定富有成效的业务决策。



**表单**

电子制作且人人皆可填写的调查和问卷

无需额外成本即可创建用于调查和问卷的自定义表单。所有数据都会收集汇总到电子表格中，您可以直接在 Google 表格中进行分析。



**幻灯片**

以团队协作方式处理文档

您无需安装软件，即可在浏览器中创建和编辑精美演示文稿。可多人同时处理一个文件，而且所有人看到的始终是最新的版本。



**记事本**

记录重要事项，提高工作效率

轻松记录灵感妙想和待办事项。与团队协作处理记事，所有内容都会在您的设备之间同步，让您随时随地掌控重要事务。



**协作平台**

轻松创建具有影响力的即时会议

协同合作，为您的团队、项目或活动创建引人入胜的优质网站。不论是桌面设备屏幕还是智能手机屏幕，显示效果都十分理想。

资料来源：谷歌，申万宏源研究

**综上，办公软件和协同办公领域，重点关注金山办公：**

**AI有望变革用户工作方式，公司加大投入力争一流。**据 Gartner 估计，到 2025 年人工智能生成数据占比将达到 10%，而如今这一比例还不到 1%。文档作为用户创作内容的主要载体之一，此次 AI 技术革命有望加速办公软件从“效率工具”到“生产工具”的变革，极大提升组织和个人用户的工作效率。公司表示将加大对 AI 技术领域的投入，充分利用 AIGC 及 LLM 等技术驱动办公产品的重构与模式创新，努力达成世界一流水平。

**AIGC 的融入，助力 C 端订阅收入规模向百亿量级进发。**2022 年公司国内个人订阅业务收入约 20 亿，年累计付费用户数约 3000 万人，人均付费金额接近 70 元。若公司加快 AI 研发速度，并推出类似微软 copilot 的强大 AI 助手，则有望大幅提升付费意愿和 ARPU。假设未来公司 PC 端月活数量达到 3 亿，AI 产品加持下付费率达到三分之一，人均付费金额较当前翻倍达到 140 元，则 C 端年订阅收入规模将达到 140 亿，为当前的 7 倍！

**国内机构订阅业务持续高增，同样将受益于 AI 能力的融入。**2022 年公司数字办公平台（云和协作产品）相关收入同比增长 57%，数字办公产品全年新增政企客户 3990 家。公司已打造出中国五矿、招商银行、建设银行、比亚迪、完美世界等数字化办公标杆项目，并推动标杆案例不断复制与下沉。公有云 SaaS 在期付费企业数同比增长 51%，付费企业续约率超 70%，金额续费率超 100%，带动公有云领域相关收入同比增长超 100%。

**ChatGPT 及 AI 大模型有望对 OA 领域带来重大革新，重点关注：**

**致远互联：**数智化协同运营平台及云服务领导厂商。近日宣布成为百度文心一言首批生态合作伙伴，公司新一代协同运营平台 COP 将全面体验并接入文心一言的能力。致远互联较早早已布局人工智能，并携手百度 AI 联合推出智能工作助手“小致语音助手”，通过融合自然语言处理（NLP）、语音识别与合成等人工智能技术，能快速将人、应用和数据进行匹配，实现多种协同应用场景下的人机对话、智能数据搜索和业务梳理等，让员工以更智能、更专业的方式，提高工作效率。

**泛微网络：**推出泛微智能办公平台，是基于智能 AI 技术和泛微 OA 场景相结合打造出的的一款高度交互性的智能办公平台，简化前台操作，去除操作界面，让软件通过前端的语音语义技术，自动识别 OA 场景的各类命令并帮助组织每一位成员完成数据查询、知识协助、出差、报销、审批等各类日常办公问题。

## 2.3 场景：CV 视频+图片

由于消费人口众多、安防完善，我国的 AI 视觉应用世界领先。2017 年 transformer 曾被认为不利于视觉。但随后一系列论文的发布和工业试验，论证该领域用大模型后效果提高。

代表公司：**大华股份、安防领军、虹软科技**等。

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/908126004007006053>