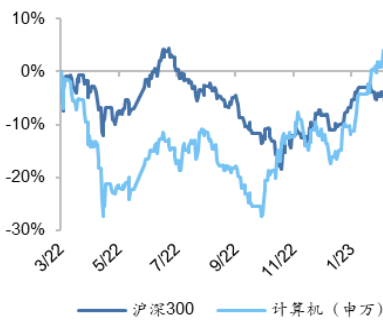


ChatGPT 引发的大模型时代变革

行业评级：增持

报告日期：2023-02-25

行业指数与沪深 300 走势比较



分析师：尹沿枝

执业证书号：S0010520020001

电话：021-60958389

邮箱：yinyj@hazq.com

分析师：王奇珏

执业证书号：S0010522060002

邮箱：wangqj@hazq.com

分析师：胡杨

执业证书号：S0010521090001

邮箱：huy@hazq.com

分析师：张天

执业证书号：S0010520110002

邮箱：zhangtian@hazq.com

分析师：金荣

执业证书号：S0010521080002

邮箱：jjinrong@hazq.com

相关报告

1. 华安证券_公司研究_计算机行业_行业深度_华安证券 2023 年计算机推演：数字经济+ _2022-12-12

2. 华安证券_公司研究_计算机行业_行业深度_华安证券数字经济系列报告（一）：科技赋能、新基建，数字经济大有可为 _2022-03-09

主要观点：

● ChatGPT 带来大模型时代变革，数据要素重要性提升

ChatGPT 是由 OpenAI 研发的一种语言 AI 模型，其特点在于使用海量语料库来生成与人类相似的反应。初代 GPT 模型参数 1.17 亿，GPT2 模型、GPT3 模型参数分别达到 15 亿、1750 亿。不断提升的参数量级，使得 ChatGPT3 当前已经能够应用在商业、研究和开发活动中。当前此类参数体量庞大的模型，成为各大科技厂商研发重点。大模型的基础为高质量大数据。ChatGPT 的前身 GPT-3 就使用了 3,000 亿单词、超过 40T 的数据。此类大数据基础的前提为三部分 1) 有效场景下的采集数据；2) 大数据的存储、清洗和标注；3) 数据质量检验。

● 大模型发展之下，算力与网络设施建设成为刚需

算力：ChatGPT 类人工智能需要更充足的算力支持其处理数据，带来更多高性能的算力芯片需求。英伟达表示，GPT-3 需要 512 颗 V100 显卡训练 7 个月，或者 1024 颗 A100 芯片训练一个月。2012 年以来，AI 训练任务中的算力增长（所需算力每 3.5 月翻一倍）已经超越摩尔定律（晶体管数量每 18 月翻一倍）。

网络设施：以微软 Azure 为例，其 AI 基础设施由互联的英伟达 Ampere A100 Tensor Core GPU 组成，并由 Quantum infiniband 交换机提供横向扩展能力。服务器节点多、跨服务器通信需求巨大，网络带宽性能成为 GPU 集群系统的瓶颈，解决方式包括增加单节点通信带宽与降低网络收敛比，带来光模块、交换机等需求。

● 下游应用场景丰富，多行业落地可期

1) “生成式 AI (generative AI)”在互联网及元宇宙领域市场化空间较为广阔。基于现行的 NLP 算法发展程度及数据集规模。在不久的将来，生成式 AI 有较大可能在“智能客服”和“搜索引擎”进行增值，并有希望以“插件”的形式赋能现有的“生产力工具链(工程软件/音视频制作工具等)”。

2) AI 在制造业的应用可分为三方面：**a) 智能装备：**指具有感知、分析、推理、决策、控制功能的制造装备，典型代表有工业机器人、协作机器人、数控机床等；**b) 智能工厂：**重点在于实现工厂的办公、管理及生产自动化，典型的代表场景有协作机器人、智能仓储物流系统等；**c) 智能服务：**指个性化定制、远程运维及预测性维护等。

3) 人工智能在智能汽车领域的应用包括：**a) 智能驾驶**依托 AI，将从驾驶辅助发展至自动驾驶；**b) 智能座舱**在 AI 支持下，从出行工具演变为出行管家。

● 风险提示

国内大模型不及预期风险；行业竞争加剧风险；AI 应用的伦理道德风险；芯片及技术供应风险；商业模式变现与落地不及预期风险。

正文目录

1 引言	7
2 CHATGPT 引发人工智能投资热潮	8
2.1 CHATGPT 是什么?	8
2.2 CHATGPT 技术和传统的 AI 有什么区别?	9
2.3 CHATGPT 将给行业带来哪些机会?	11
1) 语音识别与自然语言处理行业快速发展:	11
2) 激活产业链:	12
3 数据要素资源基础, 满足大模型训练需求	13
3.1 政策引导数据要素确权使用, 扫清人工智能发展障碍	13
3.2 大数据管理能力需求提升	14
3.3 数据标注, 是 AI 模型的基础	16
3.4 相关标的	18
1) 星环科技: 平台、数据库集一身的数据要素稀缺标的	18
2) 海天瑞声: 人工智能基础数据服务提供商, 产品矩阵不断丰富	21
4 CHATGPT 带来的变革——大模型算法	23
4.1 大模型时代的引言: DOUBLE DESCENT (双下降) 现象	23
4.2 首要关键技术: TRANSFORMER 模型	24
4.3 GPT 快速迭代, 从 GPT1.0 迅速步入 3.5 时代	24
4.4 国内外 AI 大模型项目百舸争流	26
4.5 相关标的	30
1) 商汤科技	30
2) 科大讯飞	32
3) 云从科技	33
4) 依图科技	34
5) 旷视科技	35
5 算力与网络是大模型运行的必要条件	37
5.1 大模型发展, 算力需求激增	37
5.2 GPU/GPGPU/FPGA 多路线支持算力	37
5.3 高带宽网络是对人工智能算力的重要支撑	39
5.4 量子计算有望成为 AI 算力突破的“神助攻”	45
5.5 相关标的	48
1) 海光信息: 国产 CPU 与 GPGPU 重要参与者	48
2) 复旦微电: FPGA 受益算力增长	49
3) 中际旭创: 全球光模块龙头迎来 800G 时代	51
4) 联特科技: 欧美中低速 WDM 主流供应商, 数通光模块“黑马”	52
5) 天孚通信: 光通信精益制造代表, 光引擎、激光雷达成长速度快	52
6) 国盾量子: 量子计算机已实现原型机搭建	52

7) 浪潮信息：国内领先 AI 服务器厂商.....	52
6 从元宇宙到大制造，大模型应用领域不断丰富.....	54
6.1 “生成式 AI”在互联网及元宇宙应用.....	54
1) “生成式 AI”在智能客服领域的潜在应用.....	54
2) “生成式 AI”在搜索引擎领域的潜在应用.....	56
6.2 AI 赋能制造业转型升级，智能制造浪潮兴起.....	63
1) 智能装备产业百花齐放，工业机器人与高端数控机床空间广阔.....	63
2) 智能工厂是实现智能制造的载体，协作机器人与智能仓储物流是工厂新星.....	69
6.3 AI 赋能工业互联网，打造高效率设备管理和生产流程.....	72
6.4 人工智能助力汽车智能化.....	75
1) 智能驾驶：从驾驶辅助到自动驾驶.....	75
2) 智能座舱：从出行工具到出行管家.....	77
风险提示：.....	79

图表目录

图表 1 CHATGPT 开启收费模式.....	8
图表 2 CHATGPT 操作界面.....	9
图表 3 CHATGPT 模型流程图.....	10
图表 4 TRANSFORMER 模型工作流程.....	11
图表 5 中国按类别划分的人工智能软件市场.....	12
图表 6 数据已成为五大核心生产要素之一.....	13
图表 7 2015-2021 年 GDP 增速与数字经济增速.....	14
图表 8 2015-2020 年数字经济占 GDP 的比重.....	14
图表 9 《规划》提出“1+7”的指标体系.....	14
图表 10 2017 与 2022 年人均联网设备数量.....	15
图表 11 2017-2022 全球网络流量.....	15
图表 12 全球大数据市场规模.....	15
图表 13 全球大数据软件市场规模.....	15
图表 14 我国大数据市场规模.....	16
图表 15 我国大数据软件市场规模.....	16
图表 16 人工智能基础数据服务流程与主要产品.....	16
图表 17 训练数据需求量（条）.....	17
图表 18 受访者遇到的与数据相关的难题及比例（%）.....	17
图表 19 2017 至 2029 年中国数据标注行业市场规模.....	17
图表 20 2021 年我国数据标注行业下游需求占比（%）.....	17
图表 21 面向 AI 的数据治理产业图谱.....	18
图表 22 星环科技发展历程示意图.....	19
图表 23 星环科技股权结构.....	20
图表 24 星环科技产品格局.....	20
图表 25 海天瑞声产品服务矩阵.....	21
图表 26 海天瑞声训练数据集服务的算法模型应用场景示意.....	22
图表 27 BIS-VARIANCE 图内的双下降曲线.....	23
图表 28 TRANSFORMER 模型自监督层结构.....	24
图表 29 TRANSFORMER 模型架构.....	24
图表 30 OPENAI GPT 发展历程.....	24
图表 31 不同模型参数量与模型精度的关系.....	25
图表 32 GPT-3 到 GPT-3.5 的演进过程.....	26
图表 33 大模型发展迭代图.....	27
图表 34 百度文心生态图.....	28
图表 35 ERNIE3.0 模型架构.....	28
图表 36 MINDSPORE 自动并行框架.....	29
图表 37 大模型训练需求对比.....	30
图表 38 SENSECORE 商汤 AI 大装置中心.....	30
图表 39 商汤科技主要产品结构.....	32
图表 40 科大讯飞业务全景示意图.....	33
图表 41 云从科技主要产品及服务图谱.....	34

图表 42 依图科技主要解决方案	35
图表 43 旷视科技 AIOT 软硬一体化解决方案	36
图表 44 算力需求变化	37
图表 45 深度学习中训练和推理的演示	38
图表 46 全球 GPU 行业市场规模	38
图表 47 中国 GPU 行业市场规模	38
图表 48 全球 FPGA 芯片市场规模	39
图表 49 中国 FPGA 芯片市场规模	39
图表 50 AI 加速服务器全球出货量渗透率迅速提升 (单位: 百万)	40
图表 51 英伟达 DGX A100 SU (包括 20 台 DGX A100 服务器) 是 SUPERPOD 集群基本组成单位	40
图表 52 腾讯星脉 AI 集群组网架构	41
图表 53 AI 集群高性能方案关键技术组合	41
图表 54 英伟达 DGX A100 服务器使用了 200G 高速光连接	42
图表 55 DGX H100 使用了 400G CONNECTX-7 网卡	42
图表 56 英伟达 QUANTUM-2 INFINIBAND 交换机提供 64*400GBPS 交换容量	42
图表 57 微软数据中心网络 2024 年部署将全部为 400G	43
图表 58 微软 400G 数据中心网络架构	43
图表 59 数据中心 200G 以上高速光模块出货量高速增长	44
图表 60 交换机芯片通过 SERDES 直驱光模块实现高频电信号走线缩短	45
图表 61 全球量子计算市场或在 2025 年达到 12 亿美金	45
图表 62 量子计算未来三年主要市场结构	46
图表 63 量子计算未来三年主要应用的算法领域	46
图表 64 学术界目前探讨的量子计算可能展现出优势的人工智能算法	46
图表 65 变分量子算法示意图	47
图表 66 剑桥量子计算公司 (CQC) 首次在量子计算机执行自然语言处理测试获得成功	48
图表 67 海光 DCU 基本架构	49
图表 68 海光 8100 系列产品主要规格和特点	49
图表 69 复旦微电产品主要规格和特点	50
图表 70 FPGA 结构特点	51
图表 71 全球“对话 AI(CONVERSATIONAL AI)”市场空间	55
图表 72 “智能对话机器人”在各领域全球市场空间 (亿 USD)	55
图表 73 全球“智能对话机器人”智能客服领域市场空间(亿 USD)	56
图表 74 智能对话机器人-电商客服领域全球市场空间测算(亿 USD)	56
图表 75 谷歌服务(GOOGLE SERVICES)收入 (亿 USD)	57
图表 76 GPT3 各类模型的训练的参数量	58
图表 77 GPT 模型训练成本	58
图表 78 “生成式 AI”对搜索引擎的影响	59
图表 79 生成式 AI 与搜索引擎结合面临的挑战	59
图表 80 “生成式 AI”的应用 - 新 BING (NEW BING) 搜索案例展示	60
图表 81 “生成式 AI”的应用 - 新 BING (NEW BING)可对话的领域举例	60
图表 82 “生成式 AI”的应用 - 谷歌巴德	61
图表 83 百度文心全景图、历程和架构	62
图表 84 人工智能在制造业的应用	63
图表 85 工业机器人应用汇总	64

图表 86 2015-2022 年中国工业机器人产量.....	64
图表 87 2015-2022 年中国工业机器人销量.....	64
图表 88 2017-2024 年全球工业机器人市场规模（销售额口径）.....	65
图表 89 2017-2024 年中国工业机器人市场规模（销售额口径）.....	65
图表 90 高端数控机床的基础构成.....	66
图表 91 高端数控机床的下游应用.....	66
图表 92 中国数控机床下游应用领域占比.....	67
图表 93 2017-2022 年中国数控机床市场规模统计预测.....	67
图表 94 我国支持高端数控机床发展的文件/政策.....	68
图表 95 中国机床消费结构.....	69
图表 96 机床更新需求测算.....	69
图表 97 协作机器人特点.....	70
图表 98 2021 年中国协作机器人应用行业分布.....	70
图表 99 2016-2021 年中国协作机器人市场规模.....	70
图表 100 仓储物流自动化系统构成.....	71
图表 101 工业生产物流示意图.....	71
图表 102 商业配送物流示意图.....	71
图表 103 智能仓储物流与传统仓储物流对比.....	72
图表 104 2017-2026 年中国智能仓储物流市场规模及预测.....	72
图表 105 2018-2022 年中国工业互联网市场规模.....	73
图表 106 鼎捷经营管理、生产控制方案.....	74
图表 107 智能运维平台.....	74
图表 108 中望软件 3D 仿真.....	75
图表 109 汽车驾驶自动化等级划分.....	76
图表 110 中国 L2 及以上智能汽车销量及渗透率.....	76
图表 111 中国智能座舱市场规模及预测.....	77
图表 112 集度：汽车机器人将融合百度文心一言全面能力.....	78
图表 113 长安深蓝：假如把 CHATGPT 装进长安深蓝 SL03.....	78

1 引言

ChatGPT 是由 OpenAI 研发的一种语言 AI 模型，使用上亿参数的大模型和海量语料库来生成语句，目前可以实现写诗、撰文、编码的功能。ChatGPT 广受用户欢迎，短短五天注册用户数量便超过 100 万，60 日月活破亿。产业界如微软、谷歌、百度也对于 openAI 及其竞品加大投入。

1. **逻辑一：大模型需求带动算法公司景气度。**从技术层面看，ChatGPT 算法精确度来源于引入数以亿计的模型参数，即**大模型**。除了 OpenAI 的 GPT-3 模型外，各大公司正在孵化的大模型项目也值得关注：1) **MT-NLG**：微软英伟达强强联手，软硬结合引领行业新景。2021 年 10 月 11 日，微软和英伟达推出的自然语言生成模型(MT-NLG)，具有 5300 亿个参数；2) **Switch Transformers**：Google 推出的首个万亿级语言模型，将参数值拉升至 1.6 万亿个；3) **文心一言**：2022 年 11 月 30 日，百度集团公布了文心大模型的最新升级，包括新增 11 个大模型，大模型总量增至 36 个，构建起国内业界规模最大的产业大模型体系。除互联网科技巨头外，**建议关注**：科大讯飞、商汤科技、云从科技、依图科技、旷视科技等。

2. **逻辑二：巨头在大模型算法的投入加大，激活产业链。**算法模型的训练需要在**算力与网络、数据服务**等方面的巨大研发投入。

算力方面，核心为 GPU\GPGPU\FPGA 等半导体芯片，其中可以关注 1) **GPU**：CPU 搭配 GPU 是目前深度学习的主流方案；2) **GPGPU**：去掉 GPU 的图形显示部分，将其余部分全部投入通用计算，在 AI、数据分析和 HPC 等场景下可以广泛应用；3) **FPGA**：可编程的 FPGA 芯片也逐渐提升市场份额。**相关标的包括**：海光信息、复旦微电、浪潮信息、紫光股份等。

网络设施方面，服务器增加带动跨服务器通信需求，网络带宽性能成为 GPU 集群系统的瓶颈，解决的方式包括增加单节点通信带宽、降低网络收敛比。由此带来光模块、交换机等需求。**相关标的包括**：中际旭创、联特科技、天孚通信、国盾量子等。**数据服务方面**，数据采集、数据标注和数据质检是较为重要的三个环节。ChatGPT 的前身 GPT-3 就使用了 3,000 亿单词、超过 40T 的大规模、高质量数据进行训练。随着大模型的发展，对于训练数据的需求成指数增长。**相关标的包括**：星环科技、海天瑞声。

3. **逻辑三：AI 下游应用领域逐步扩大。**随着大模型的不断完善，未来有望应用于更多场景之下。包括互联网及元宇宙领域、工业领域、智能汽车与智能座舱，达到下游用户的降本增效。随着商业模式与应用前景明朗，进一步反哺大模型投入。

2 ChatGPT 引发人工智能投资热潮

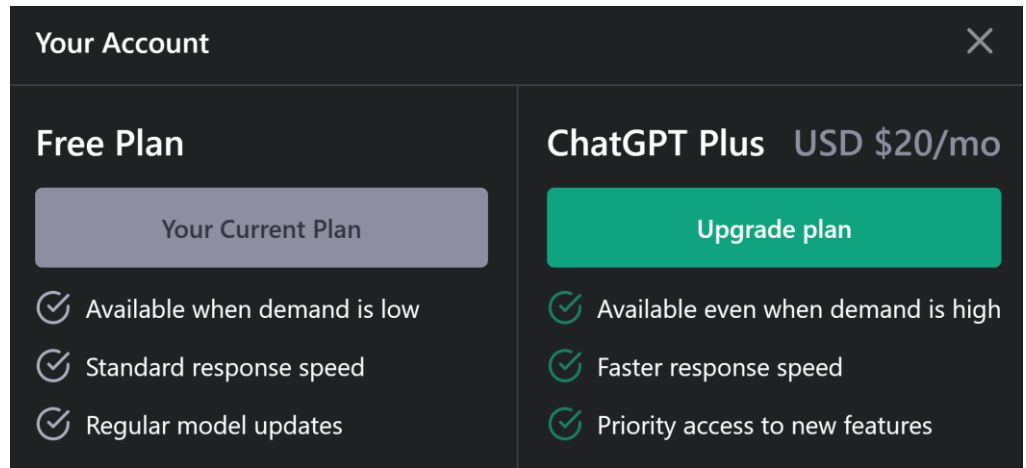
2.1 ChatGPT 是什么？

ChatGPT 是由 OpenAI 研发的一种语言 AI 模型，使用海量语料库来生成与人类类似的反应。ChatGPT 是基于 GPT (generative pretrained transformer) 架构搭建的，主要用深度学习来生成连贯且具有意义的文字。这个模型使用了来自于网站、书本和社交媒体的海量文字数据，因此也为 ChatGPT 在保证准确性和细节的同时，提供了广泛的对话反馈。对话反馈是 ChatGPT 的核心功能之一，也使它成为了实现聊天机器人或其他对话型 AI 的理想技术。

除对话功能外，ChatGPT 也具有实现各类语言相关任务的能力，包括文章精炼、翻译以及情绪分析等。以上各类语言能力在大规模的训练数据和升读学习架构下，使 ChatGPT 成为目前应用最为先进的语言模型之一。

总体上，ChatGPT 标志着自然语言处理 (NLP) 和对话 AI 领域的一大步，其高质量文字产出能力在商业、研究和开发活动中提高用户体验的方向上非常有应用价值的。

图表 1 ChatGPT 开启收费模式



资料来源：OpenAI，华安证券研究所

截至目前，GPT 已经经历了如下演化：

1. GPT-1: 第一代 GPT 语言模型，发布于 2018 年。它有 1.17 亿个参数，使用网页的文字数据进行训练。
2. GPT-2: 发布于 2019 年，具有 15 亿个参数，使用的网页文字数据量也远大于前一代。它已经可以生成高质量的文字，甚至完成翻译、精炼文字等简单任务。
3. GPT-3: 发布于 2020 年，具有 1750 亿个参数，使用网页以及其他来源的文字进行训练。它已经可以进行担任各类任务，被认为是语言模型领域的显著突破。

图表 2 ChatGPT 操作界面



资料来源：openAI，华安证券研究所

2.2 ChatGPT 技术和传统的 AI 有什么区别？

相比传统 AI 算法，GPT 模型的区别在于通过海量参数，进一步提升了模型的精确度。

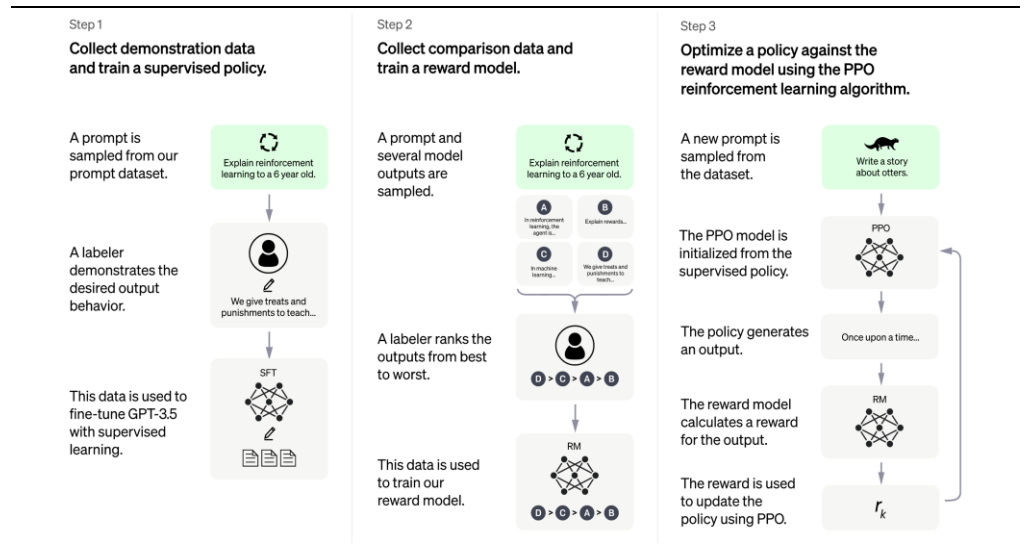
初代的 GPT 模型参数是 1.17 亿，而 GPT2 的模型有 15 亿个参数，参数增加了 10 倍之多。第三代的 GPT3 模型，参数达到了 1750 亿，是 GPT2 参数的 100 倍。正是由于参数的指数级提升，使得模型的使用效果大幅提升。而此类参数上亿的模型，通常称之为“大模型”。

GPT 模型基于 Transformer 架构，这是一种由谷歌的 Vaswani 等人于 2017 年引入的神经网络类型。Transformer 架构特别擅长对序列数据中的长距离依赖进行建模，这使其非常适合自然语言处理任务。

为了训练 GPT 模型，OpenAI 使用了来自互联网的大量文本数据，包括书籍、文章和网站。该模型使用一种称为无监督学习的技术进行训练，这意味着它学会了在没有人类监督的情况下预测文本序列中的下一个单词。

GPT 模型能够生成连贯和语法正确的文本，已被用于广泛的自然语言处理任务，包括语言翻译、文本补全和文本生成。

图表 3 ChatGPT 模型流程图



资料来源：OpenAI，华安证券研究所

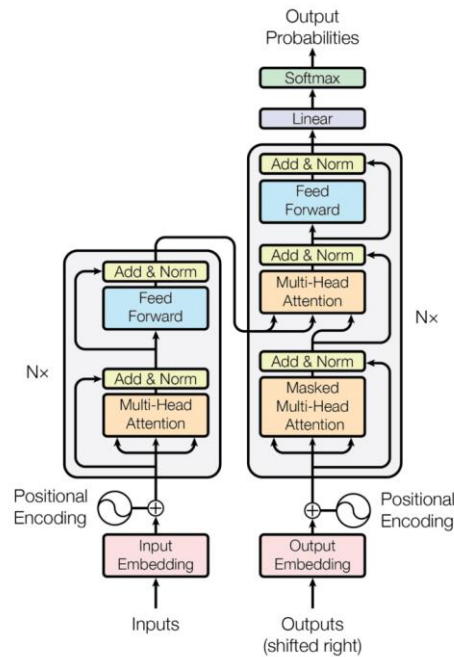
Transformer 模型是一种用于自然语言处理的神经网络模型。与传统的循环神经网络（RNN）不同，Transformer 模型使用自注意力机制（self-attention）来处理输入序列中不同位置之间的依赖关系。

Transformer 模型由编码器和解码器两部分组成。编码器将输入序列中的每个单词表示为一个向量，并通过多层自注意力和前馈神经网络来对输入序列进行编码。解码器则使用相同的自注意力和前馈神经网络来生成输出序列。

在自注意力机制中，模型根据输入序列中的所有单词计算出每个单词与其他单词的相关性，然后使用这些相关性加权求和得到每个单词的表示向量。这种方法使得模型能够处理长序列和跨越序列中的依赖关系，从而提高了模型的性能。

Transformer 模型已经在自然语言处理领域取得了很好的效果，包括机器翻译、文本摘要和问答系统等任务。它是目前最先进的语言模型之一，也是开发其他自然语言处理模型的基础。

图表 4 transformer 模型工作流程



资料来源：machine learning mastery，华安证券研究所

2.3 ChatGPT 将给行业带来哪些机会？

相比其他此前的人工智能技术与进展，ChatGPT 之所以引发关注，主要总结为以下几点：

- 1) 从使用效果上，交流通畅，同时能够实现写诗、撰文、编码的功能。2月1日，以色列总统艾萨克·赫尔佐格(Isaac Herzog)发表了部分由人工智能(AI)撰写的演讲；
- 2) 受用户欢迎。短短5天，注册用户数就超过100万。60天月活破亿。
- 3) 商业模式产生变化。2023年2月2日，美国人工智能(AI)公司OpenAI发布ChatGPT试点订阅计划。
- 4) 产业界也表现出对Chatgpt的关注。表现为：1)1月23日，微软宣布向ChatGPT开发者OpenAI追加投资数十亿美元；2)谷歌3亿美元投资Chatgpt竞品。3)百度将于3月发布类似Chatgpt的AI服务。

由此带来相关产业链的大变革：

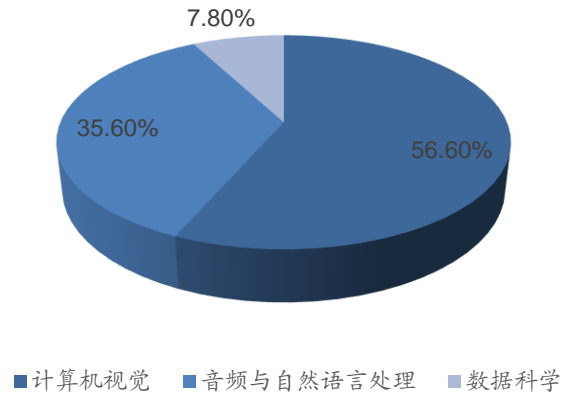
1) 语音识别与自然语言处理行业快速发展：

人工智能，也即解决像人一样看、听、思考的问题。因此，按照此维度来划分，划分为计算机视觉、语音识别与自然语言处理及数据科学。

早先，2020年数据显示，计算机视觉占比约56.6%；语音识别与自然语言处理占比约35.6%。也即，在机器视觉领域的应用，相比自然语言处理，更为成熟，市场规模更大。

但随着ChatGPT带来的投资热潮，与应用领域的不断丰富，音频与自然语言处理的整体行业规模，有望迅速增长。

图表 5 中国按类别划分的人工智能软件市场



资料来源：商汤科技招股说明书，华安证券研究所

2) 激活产业链:

整个人工智能的产业链包括算力、数据、算法乃至下游应用。

算力与网络: 英伟达的研究表示，GPT-3 模型需要使用 512 颗 V100 显卡训练 7 个月时间，或者使用 1024 颗 A100 芯片训练长达一个月的时间。随着各大科技厂商投入对大模型的研发，势必增加芯片、服务器等算力需求。同时，庞大的 AI 算力集群，又需要高带宽支撑数据传输。

数据: 数据采集、数据标注和数据质检是较为重要的三个环节。从自然数据源简单收集取得的原料数据并不能直接用于有效监督的深度学习算法训练，经过专业化采集、加工形成的训练数据集才能供深度学习算法等训练使用，由此带来数据服务需求。

算法: 相比传统 AI 模型，大模型的优势体现在：1) 解决 AI 过于碎片化和多样化的问题；2) 具备自监督学习功能，降低训练研发成本；3) 摆脱结构变革桎梏，打开模型精度上限。对于大模型算法的研发、优化，亦是投入的重点。

下游应用: 产业界一直以来都在寻求人工智能的应用领域、商业模式突破。随着大模型使用、人工智能算法精度提升，下游应用的扩展可期。

3 数据要素资源基础，满足大模型训练需求

3.1 政策引导数据要素确权使用，扫清人工智能发展障碍

数据已成为五大核心生产要素之一。

2020年4月中共中央国务院《关于构建更加完善的要素市场化配置体制机制的意见》中发布。这是数据作为新型生产要素首次在中央顶层文件中提出。

而后，2022年4月国务院《关于加快建设全国统一大市场的意见》中，进一步提到加快培育数据要素市场，建立数据资源产权相关基础制度。

2022年12月9日，财政部发布关于征求《企业数据资源相关会计处理暂行规定（征求意见稿）》意见的函，具体提出了企业数据资源相关会计、处理的方式方法，进一步扫清了数据要素市场建立、数据资源交易的障碍。

当前，2022年12月发布《关于构建数据基础制度更好发挥数据要素作用的意见》，是数据要素体系建设中，顶层关键文件，扫除了未来人工智能发展中需要使用数据的障碍：1) 建立保障权益，合规使用的数据产权制度；2) 建立合规高效的场内外结合的数据要素流通和交易制度。3) 建立体现效率促进公平的数据要素收益分配制度。4) 建立安全可控弹性包容的数据要素治理制度。

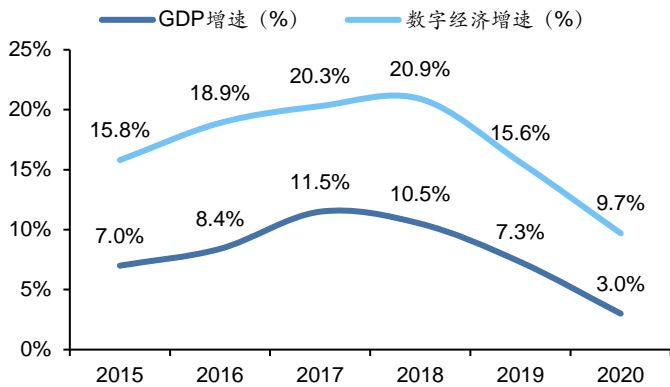
图表 6 数据已成为五大核心生产要素之一



资料来源：中国信通院，华安证券研究所

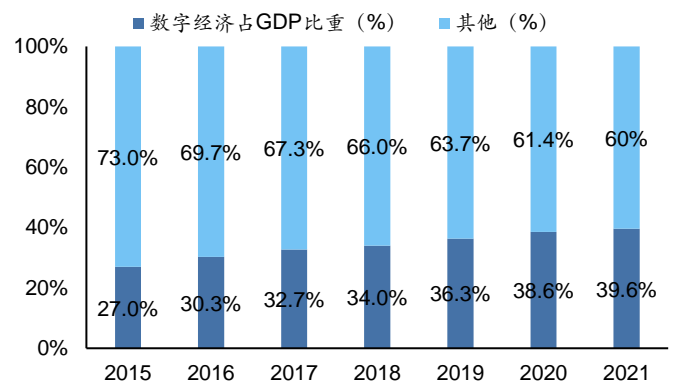
数字经济快速发展，数据要素成为重要战略资源。《“十四五”数字经济发展规划》中指出，数字经济是继农业经济、工业经济之后的主要经济形态。《规划》设定了到2025年实现数字经济核心产业增加值占GDP比重达到10%的目标，涵盖数据要素市场、产业数字化、数字产业化、数字化公共服务、数字经济治理体系五个方面。从2015年至今，数字经济平均增速持续高于GDP增速，2021年数字经济占GDP比重已经由2015年的27%提升至40%。

图表 7 2015-2021 年 GDP 增速与数字经济增速



资料来源: WIND, 华安证券研究所

图表 8 2015-2020 年数字经济占 GDP 的比重



资料来源: 中国信通院, 华安证券研究所

图表 9 《规划》提出“1+7”的指标体系

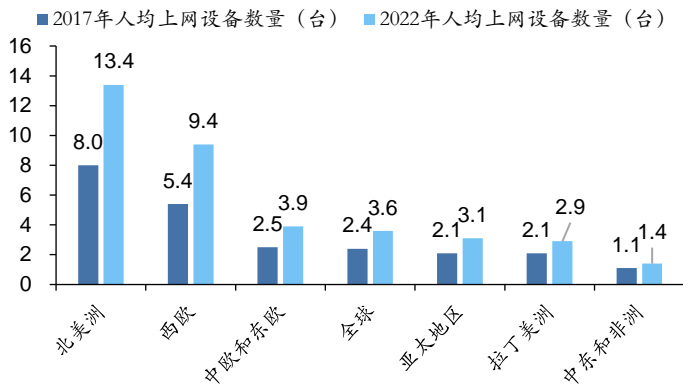
指标	2020 年	2025 年
数字经济核心产业增加值占 GDP 比重 (%)	7.8	10
IPv6 活跃用户数 (亿户)	4.6	8
千兆宽带用户数 (万户)	640	6000
软件和信息技术服务业规模 (万亿元)	8.16	14
工业互联网平台普及率 (%)	14.7	45
全国网上零售额 (万亿元)	11.76	17
电子商务交易规模 (万亿元)	37.21	46
在线政务服务实名用户规模 (亿)	4	8

资料来源: 《“十四五”数字经济规划》, 华安证券研究所

3.2 大数据管理能力需求提升

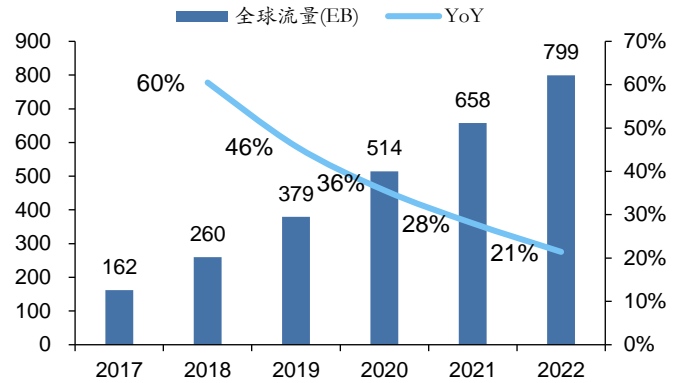
联网设备高增之下, 流量增长不可避免。根据思科的《年度互联网报告》, 到 2023 年, 地球上的连网设备数量将是全球人口的大约三倍, 从 2017 年的人均 2.4 台提升至 3.6 台。由于 IP 地址即网络地址+主机地址, 网络站点所连接的 IP 数量也处于爆发的阶段。根据 IDC 的《中国物联网连接规模预测, 2020-2025》, 仅我国物联网 IP 连接量已在 2020 年达 45.3 亿, 有望在 2025 年达到 102.7 亿, CAGR 为 17.8%。由于 IP 地址联网后即产生数据流量, IP 地址的数量增长即代表全网数据也将继续大增, 对于现有的网络企业的承载能力提出了考验。根据思科的《年度互联网报告》, 2022 年全球网络数据流量将达 799EB (1EB=十亿 GB), 同比增长 21%。我们认为, 数据流量的增长, 有望直接带动大数据产业的发展, 而其中稳定优质响应快的数据库性价比更高。

图表 10 2017 与 2022 年人均联网设备数量



资料来源：思科，华安证券研究所

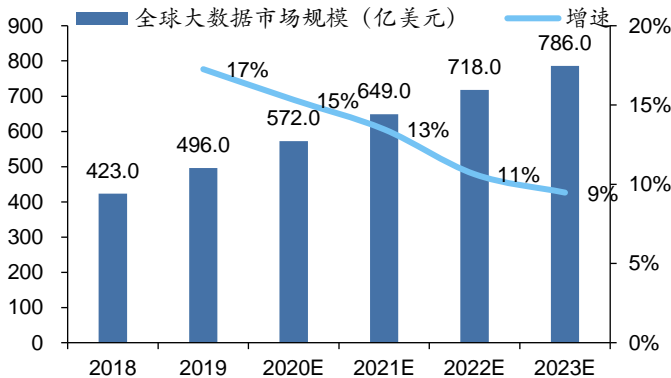
图表 11 2017-2022 全球网络流量



资料来源：思科，华安证券研究所

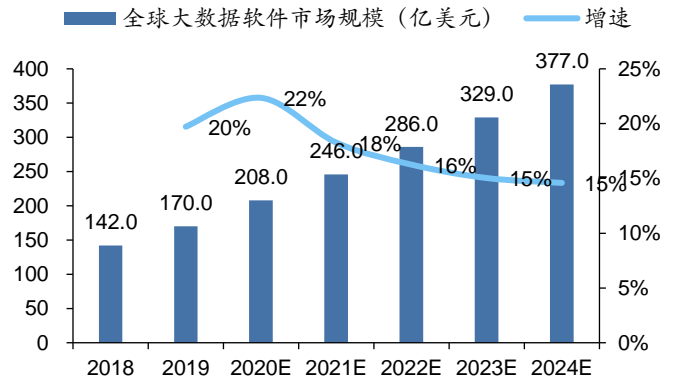
全球大数据市场存量巨大，软件市场占比较高且增速快。根据 Wikibon 及沙利文研究数据，全球大数据市场规模有望在 2022 年达 718 亿美元，同比增速 11%；而其中全球大数据软件为 286 亿美元，同比增速 18%，约占大数据市场规模的 40%。可以认为，软件市场在大数据市场中，占据较大地位，而由于其增速高于大数据市场的整体增速，其占比还将进一步提升。

图表 12 全球大数据市场规模



资料来源：沙利文，华安证券研究所

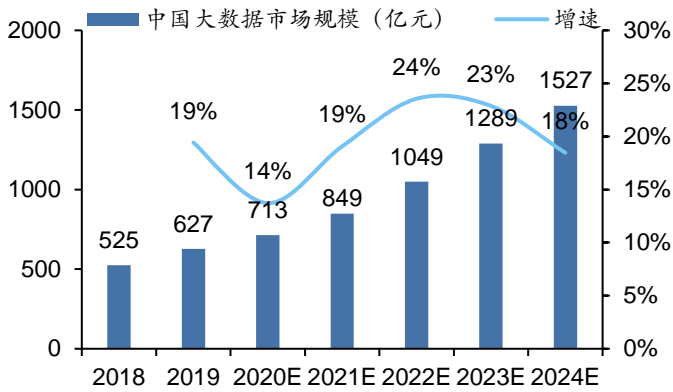
图表 13 全球大数据软件市场规模



资料来源：沙利文，华安证券研究所

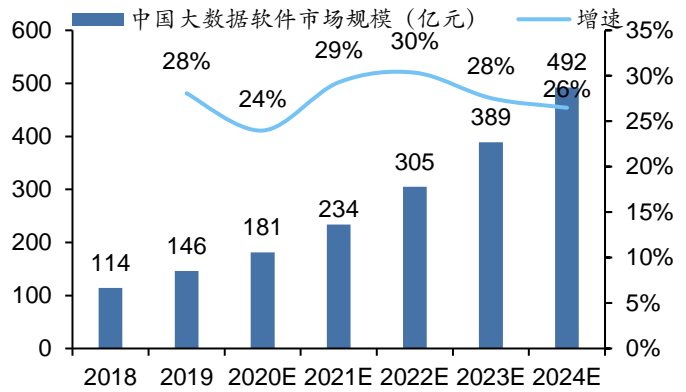
我国大数据市场，软件市场增速快、占比有待提升。对比全球市场，我国大数据市场预计 2022 年规模为 1049 亿元，同比增速 24%，其中软件约 305 亿元，同比增速 30%，占比约 29%。在大数据行业的高增速之下，数据智能分析工具、大数据管理平台等软件的需求有望进一步提升。

图表 14 我国大数据市场规模



资料来源：沙利文，华安证券研究所

图表 15 我国大数据软件市场规模

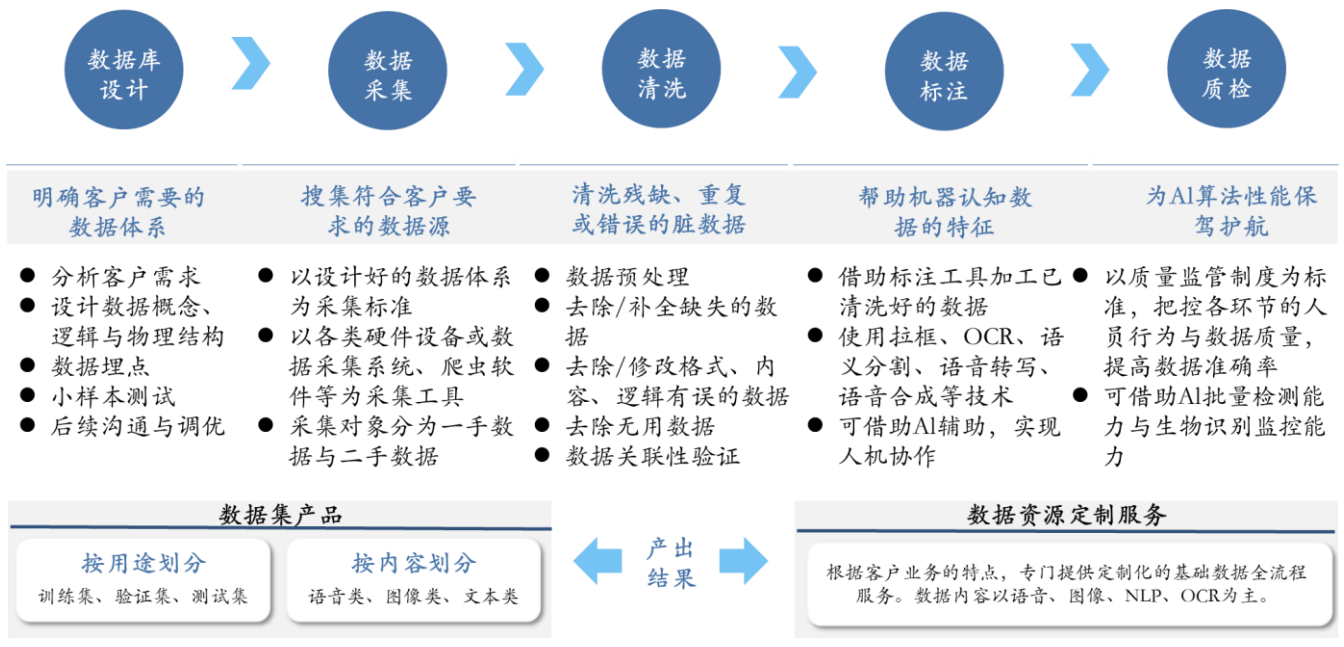


资料来源：沙利文，华安证券研究所

3.3 数据标注，是 AI 模型的基础

人工智能基础数据服务助力 AI 训练与调优，数据采集、数据标注和数据质检是较为重要的三个环节。从自然数据源简单收集取得的原料数据并不能直接用于有效监督的深度学习算法训练，经过专业化采集、加工形成的训练数据集才能供深度学习算法等训练使用，从某种程度上讲，数据决定了 AI 的落地程度，因此，基础数据服务应运而生。具体来看，基础数据的服务流程围绕着客户的展开，为 AI 模型训练提供可靠、可用的数据，其包含五个环节，分别是 1) 数据库设计：训练数据集结构设计；2) 数据采集：获取原料数据；3) 数据清洗：清洗残缺、重复或者错误的的数据；4) 数据标注：帮助机器认识数据的特征；5) 质检：各环节质量检测和控制在。

图表 16 人工智能基础数据服务流程与主要产品

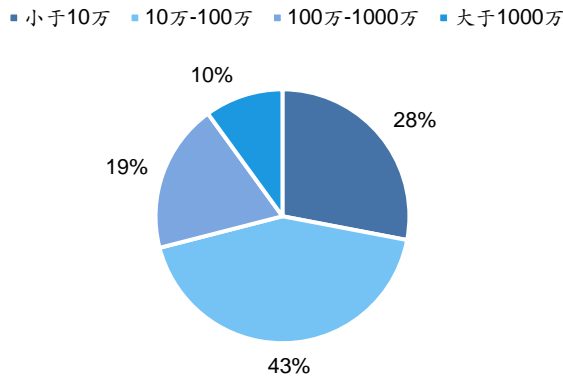


资料来源：艾瑞咨询，华安证券研究所

基础数据服务需求旺盛，未来成长空间广阔。人工智能应用场景的创新与拓展和机器学习算法复杂度的提升带动训练数据需求的大幅增长，根据 Dimensional Research

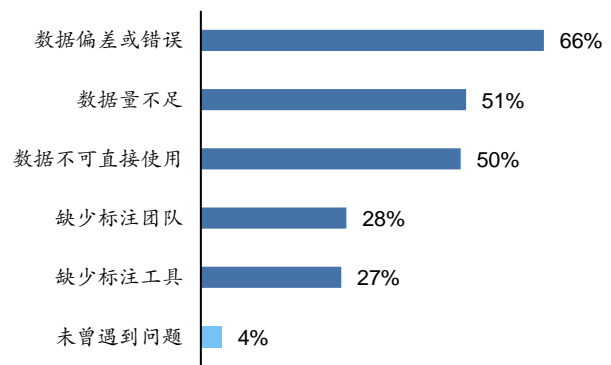
的全球调研报告，72%的受访者认为至少使用超过 10 万条数据进行模型训练，才能保证模型的有效性 & 可靠性；96%的受访者在训练模型中遇到训练数据质量不佳、数量不足、数据标注人员不足等难题。展望未来，一方面，基础数据服务的质量有进一步提升的空间；另一方面，应用场景和算法更新等将增加训练数据的需求量，基础数据服务行业有望进入快速发展阶段，成长空间广阔。

图表 17 训练数据需求量 (条)



资料来源: Dimensional Research, 华安证券研究所

图表 18 受访者遇到的与数据相关的难题及比例 (%)

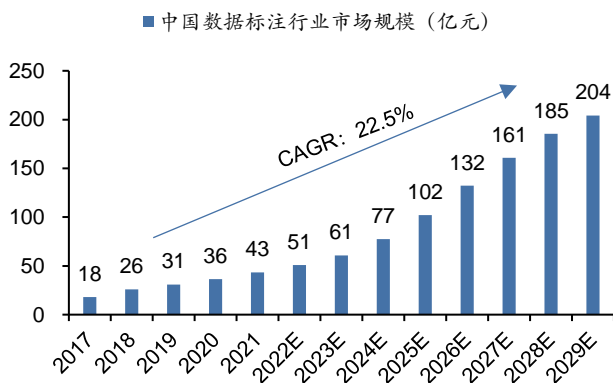


资料来源: Dimensional Research, 华安证券研究所

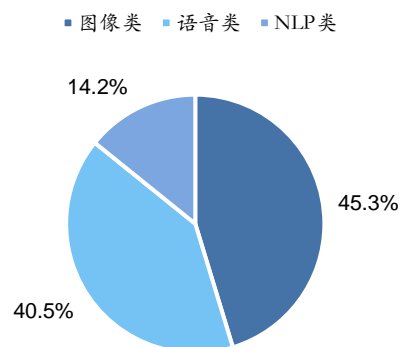
市场标注行业市场规模不断扩大，图像类和语音类需求占比超八成。从市场规模来看，根据观研天下统计，2021 年我国数据标注行业市场规模为 43 亿元，2017 至 2029 年的 CAGR 为 23%；根据 IDC《2021 年中国人工智能基础数据服务市场研究报告》，预计中国 AI 基础数据服务市场规模将在 2025 年突破 120 亿元，近五年的 CAGR 达 47%。从市场收入结构来看，按数据类型划分，基础数据服务行业是市场需求可以分为图像类、语音类和自然语言处理类数据需求。根据观研天下统计，2021 年我国数据标注行业下游以图像类和语音类需求为主，二者合计占比达 86%，其中，图像类业务以智能驾驶与安防为主，语音类以中英大大语种、中国本土方言以及外国小语种为主。

图表 19 2017 至 2029 年中国数据标注行业市场规模

图表 20 2021 年我国数据标注行业下游需求占比 (%)



资料来源: 观研天下, 华安证券研究所

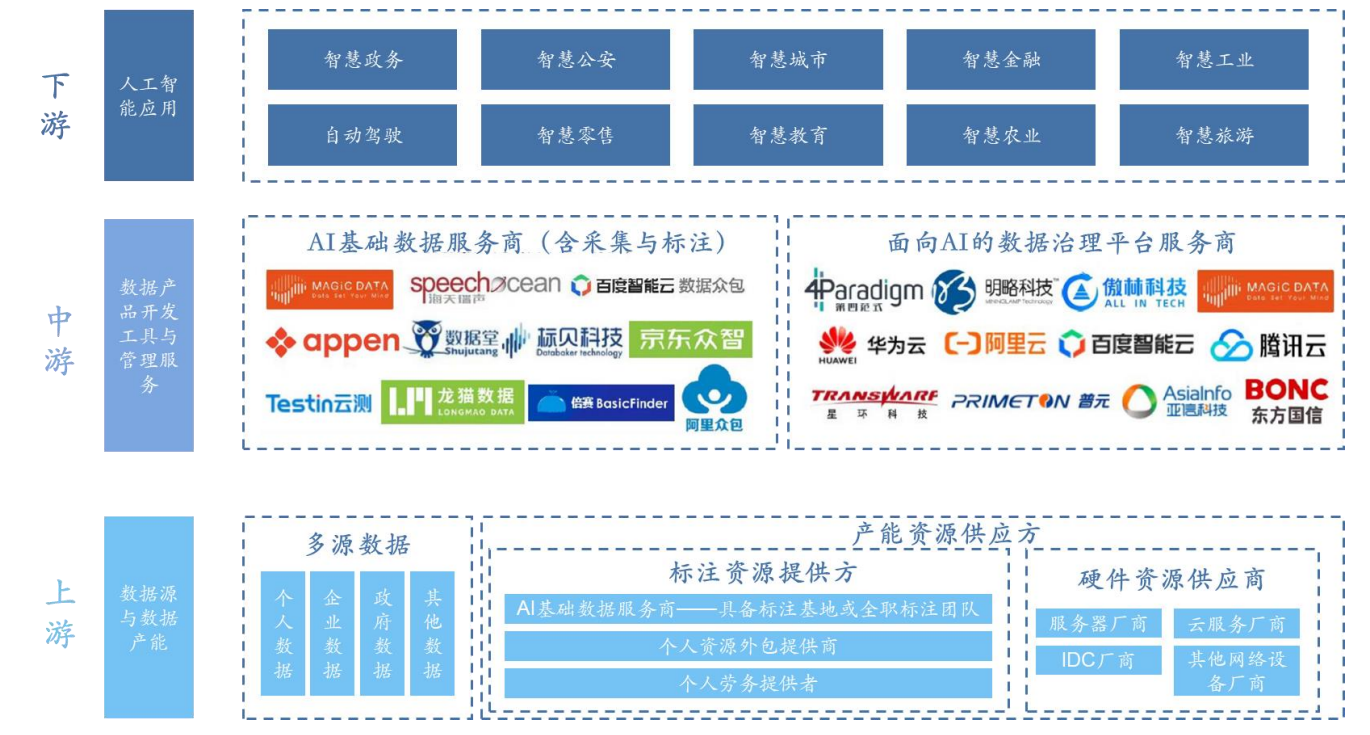


资料来源: 观研天下, 华安证券研究所

基础数据服务位于产业链中游，是 AI 商业化应用中重要的一环。基础数据服务行业的上游为数据源与数据产能，多元数据包括个人数据、企业数据、政府数据等，产能医院供应方包括标注自愿提供方和硬件资源供应商。中游为数据产品开发工具与管理服务，

包括 AI 基础数据服务商，如海天瑞声、百度众包、京东众志等企业，以及面向 AI 的数据治理平台服务商，如星环科技等。下游则是人工智能的应用，涉及智慧政务、金融、工业、自动驾驶等领域。其中，处于中游的 AI 基础数据服务商主要进行数据的采集与标注，面向 AI 的数据治理平台服务商则使用数据治理的各组件管治多源异构数据，使其形成数据资产，提高数据质量。二者处理后的数据可直接提供给下游用于 AI 训练，从而加速 AI 落地。

图表 21 面向 AI 的数据治理产业图谱



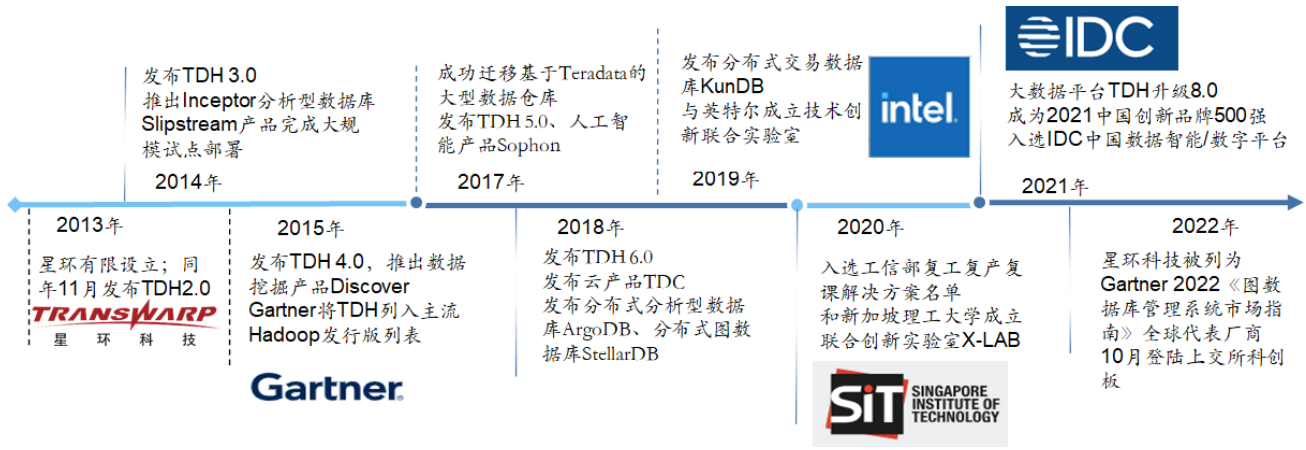
资料来源：艾瑞咨询，华安证券研究所

3.4 相关标的

1) 星环科技：平台、数据库集一身的数据要素稀缺标的

专注于分布式数据库，技术水平全球领先。星环科技 2013 年成立于上海，是国内大数据管理软件领导者，已累计有超过 1,000 家终端用户，且产品已落地以下知名机构或其主要分支机构，金融行业包括中国银行、浦发银行、浙江农村商业联合银行等，政府领域包括上海市大数据中心等，能源行业包括中国石油、南方电网等，交通行业包括中国邮政集团、郑州地铁等，制造业包括湖南中烟等。公司在发展中经历了多个重要节点：1) 公司 2013 年成立，随即发布了大数据基础平台 TDH2.0 版本，并于次年推出 Inceptor 关系型分析引擎、Slipstream 实时计算引擎，实现数据湖、实时计算两大热点功能；2) 2014 年公司被 Gartner 列入 Hadoop 的主流发行版列表；3) 2017 年起，公司陆续发布新品，包括分析工具 Sophon、云产品 TDC、分布式分析数据库 ArgoDB 和分布式交易数据库 KunDB。2022 年，公司已被 Gartner 评为图数据库管理的全球代表厂商。

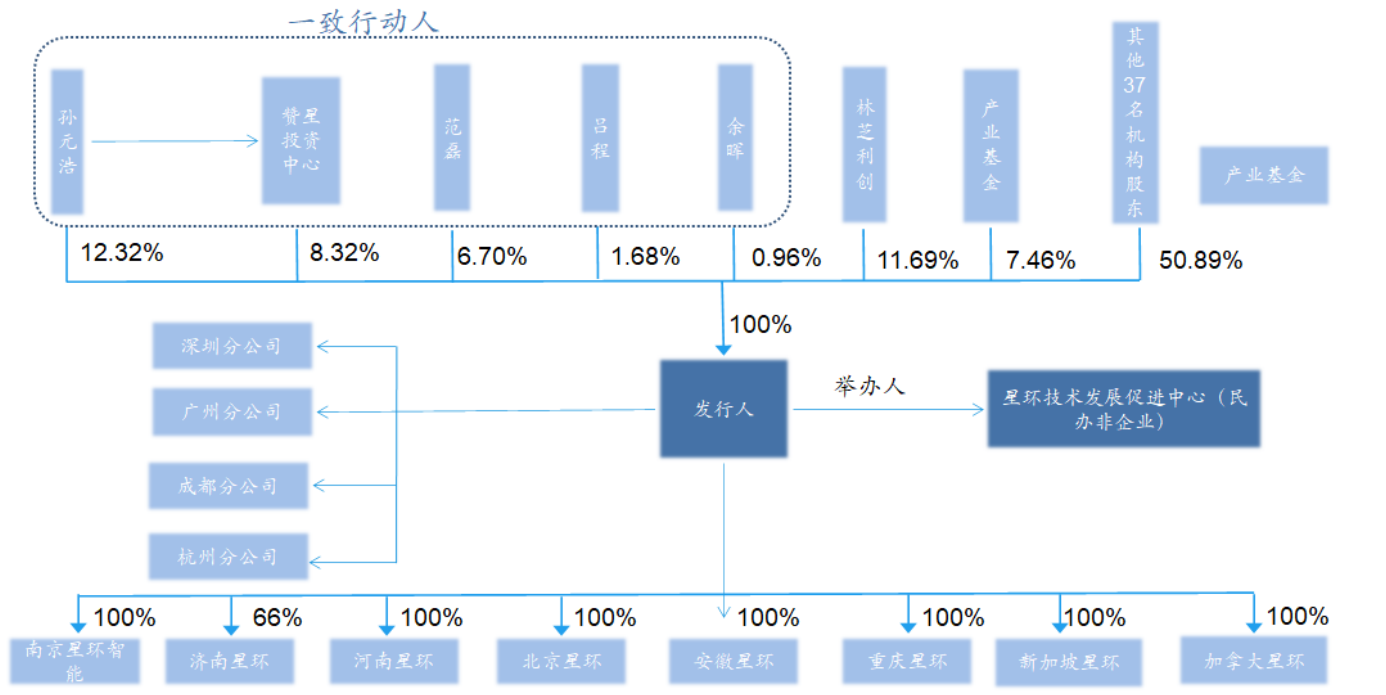
图表 22 星环科技发展历程示意图



资料来源：公司官网，华安证券研究所

股权结构稳定，创始人保持控制权。发行人的控股股东、实际控制人为创始人孙元浩，主要理由如下：（1）截至本招股说明书签署日，孙元浩直接持有星环科技 12.3%的股份，为公司第一大股东，且在报告期内持续为发行人第一大股东。（2）孙元浩与范磊、吕程、余晖及赞星投资中心签署了《一致行动协议》，确认自 2019 年 1 月 1 日起，范磊、吕程、余晖及赞星投资中心与孙元浩在发行人有关重大事项中保持一致行动，并约定上述各方在无法达成一致意见时，为提高公司决策效率，在不损害孙元浩合法权益及保障公司整体利益的前提下，应以孙元浩的意见作为各方的最终共同意见。孙元浩担任执行事务合伙人的赞星投资中心持有公司 8.3%的股份，孙元浩之一致行动人范磊、吕程、余晖分别持有公司 6.7%、1.7%、1.0%的股份。因此，孙元浩本人及通过《一致行动协议》合计控制公司 30.0%的股份。（3）报告期内，孙元浩一直担任发行人（及其前身星环有限）的董事长及总经理，在发行人的董事会和日常管理决策中均能够产生重大影响。（4）根据除孙元浩、范磊、吕程、余晖及赞星投资中心以外的发行人其他股东的书面确认，各方均认可孙元浩于报告期内作为发行人的实际控制人。

图表 23 星环科技股权结构



资料来源：公司官网，华安证券研究所

产品形成联动，长短期均有发力点。公司已形成大数据与云基础平台、分布式关系型数据库、数据开发与智能分析工具的软件产品矩阵，支撑客户及合作伙伴开发数据应用系统和业务应用系统，助力客户实现数字化转型。

具体来看，公司的软件产品分为三部分：1) 大数据与云基础软件：主要产品为企业级多模型数据管理平台 TDH 以及基于 TDH 的云平台 TDC；2) 分布式数据库：分布式交易数据库 KunDB 以及分布式分析性数据库 ArgODB；3) 数据分析工具：数据分析开发套件 TDS 以及智能分析工具 Sophon，其中 Sophon 多数用于金融下游。

图表 24 星环科技产品格局



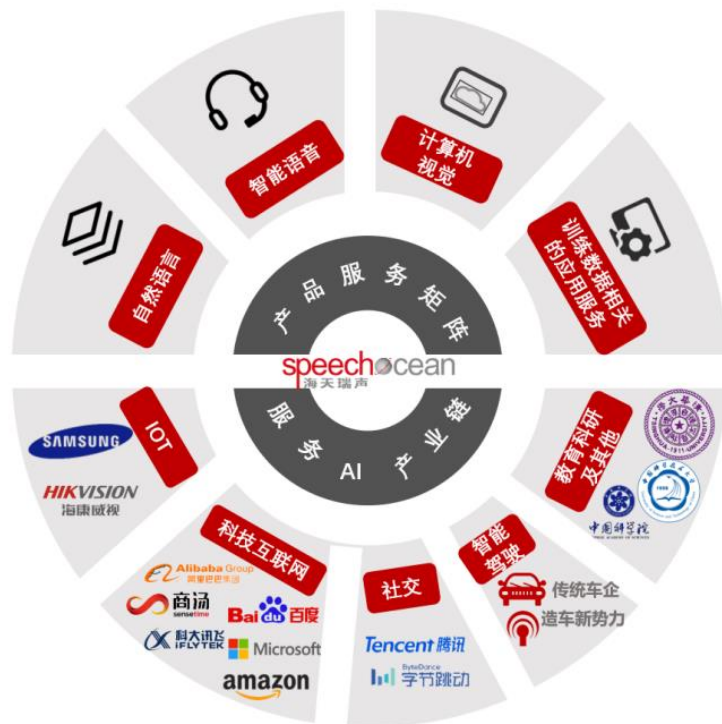
资料来源：星环科技招股说明书，华安证券研究所

2) 海天瑞声：人工智能基础数据服务提供商，产品矩阵不断丰富

自 2005 年成立以来，海天瑞声始终致力于为 AI 产业链上的各类机构提供算法模型开发训练所需的专业数据集，目前已发展为人工智能领域具备国际竞争力的国内领军企业。公司研发生产的训练数据覆盖了智能语音、计算机视觉及自然语言处理三大 AI 核心领域，实现了标准化产品、定制化服务、相关应用服务的全覆盖，广泛应用于人机交互、智能家居、智能驾驶、智慧金融、智能安防、OCR 识别等多个应用场景。截至 2022 年半年报，公司累计客户量达 695 家。

- **训练数据定制服务：**公司根据客户的需求提供训练数据定制服务，分为采集+加工服务和纯加工服务两种，在该业务类型下，最终形成的训练数据集成品的知识产权由客户享有。
- **训练数据产品：**公司根据市场需求，以及对算法技术应用前景、发展趋势的评估等开发的训练数据产品，最终生产的训练数据集成品的知识产权由公司享有，一次开发完成后可重复多次销售使用权。
- **训练数据相关的应用服务：**公司基于其生产的训练数据提供算法模型相关的训练服务，助力下游客户完成其算法模型的语言拓展、垂直应用领域拓展等，为客户定制针对特定应用场景的专属算法模型。

图表 25 海天瑞声产品服务矩阵



资料来源：海天瑞声 2022 年半年报，华安证券研究所

公司产品应用领域不断拓宽，下游客户丰富。从应用场景来看，公司产品的应用场景覆盖了个人助手、语音输入、智能家居、智能客服、机器人、语音导航、智能播报、语音翻译、移动社交、虚拟人、智能驾驶、智慧金融、智慧交通、智慧城市、机器翻译、智能问答、信息提取、情感分析、OCR 识别等多种应用场景。从下游客户来看，公司的

客户为 AI 产业链上的各类机构，主要系：1) 大型科技公司，阿里巴巴、腾讯、百度、微软等；2) 人工智能企业，科大讯飞、商汤科技、海康威视等；3) 科研机构，如中国科学院、清华大学等。目前，公司的产品和服务已经获得了阿里巴巴、腾讯百度、科大讯飞、微软、清华大学等国内外客户的认可。

图表 26 海天瑞声训练数据集服务的算法模型应用场景示意



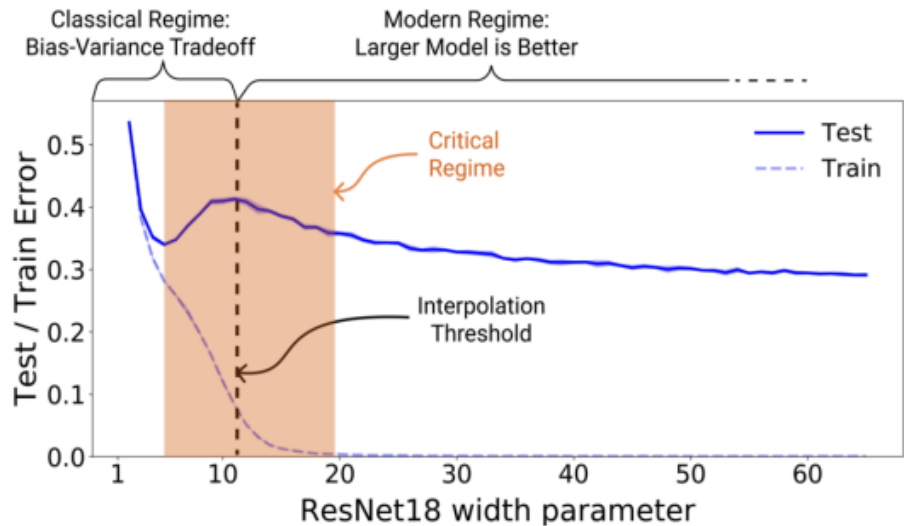
资料来源：海天瑞声 2022 年半年报，华安证券研究所

4 ChatGPT 带来的变革——大模型算法

4.1 大模型时代的引言：Double Descent（双下降）现象

随着深度神经网络的兴起，人工智能进入统计分类深度模型时代，这种模型比以往的模型更加泛化，可以通过提取不同特征值应用于不同场景。但在 2018 年-2019 年，双下降现象的发现打破了原有的人工智能发展格局。简而言之，以往的数学理论表明，随着参数增多、模型增大，过拟合导致模型的误差会先下降后上升，这使得找到精度最高误差最小的点成为模型调整的目标。而随着人工智能算法算力的不断发展，研究者发现如果继续不设上限的增大模型，模型误差会在升高后第二次降低，并且误差下降会随着模型的不增大而降低，**通俗而言模型越大，准确率越高**。因此人工智能发展进入了大模型时代。

图表 27 Bis-Variance 图内的双下降曲线



资料来源：《预训练语言模型之 GPT-1, GPT-2 和 GPT-3》CSDN，华安证券研所

相比传统 AI 模型，大模型的优势体现在：

1) 解决 AI 过于碎片化和多样化的问题，极大提高模型的泛用性。应对不同场景时，AI 模型往往需要进行针对性的开发、调参、优化、迭代，需要耗费大量的人力成本，导致了 AI 手工作坊化。大模型采用“预训练+下游任务微调”的方式，首先从大量标记或者未标记的数据中捕获信息，将信息存储到大量的参数中，再进行微调，极大提高模型的泛用性。

2) 具备自监督学习功能，降低训练研发成本。我们可以将自监督学习功能表现理解为降低对数据标注的依赖，大量无标记数据能够被直接应用。这样一来，一方面降低人工成本，另一方面，使得小样本训练成为可能。

3) 摆脱结构变革桎梏，打开模型精度上限。过去想要提升模型精度，主要依赖网络在结构上的变革。随着神经网络结构设计技术逐渐成熟并开始趋同，想要通过优化神经网络结构从而打破精度局限变得困难。而研究证明，更大的数据规模确实提高了模型的精度上限。

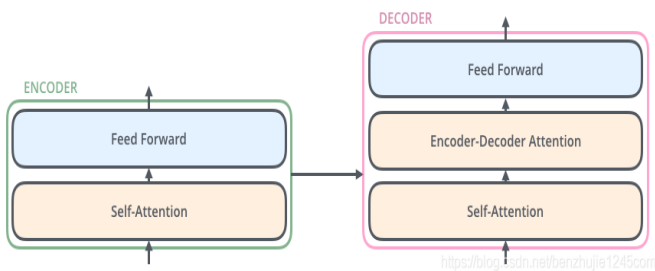
4.2 首要关键技术: Transformer 模型

GPT 模型利用 Transformer 模型作为特征提取器，是第一个引入 Transformer 的预训练模型。传统的神经网络模型例如 RNN（循环神经网络）在实际训练过程中由于输入向量大小不一、且向量间存在相互影响关系导致模型训练结果效果较差。Transformer 模型有三大技术突破解决了这个问题。

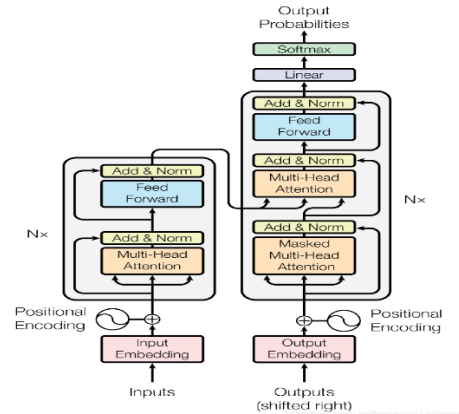
首先 Transformer 模型的 Self-Attention（自注意力）机制使人工智能算法注意到输入向量中不同部分之间的相关性，从而大大提升了精准性。其次该模型采用属于无监督学习的自监督学习，无需标注数据，模型直接从无标签数据中自行学习一个特征提取器，大大提高了效率。最后，在做具体任务时，微调旨在利用其标注样本对预训练网络的参数进行调整。也可以针对具体任务设计一个新网络，把预训练的结果作为其输入，大大增加了其通用泛化能力。

Transformer 模型的这些优点快速替代了传统的神经网络。

图表 28 Transformer 模型自监督层结构



图表 29 Transformer 模型架构



资料来源:《Transformer 模型详解》CSDN, 华安证券研究所

资料来源:《Transformer 模型详解》CSDN, 华安证券研究所

4.3 GPT 快速迭代, 从 GPT1.0 迅速步入 3.5 时代

GPT: 大型无监督语言模型, 能够生产连贯的文本段落。GPT-1 采用无监督预训练和有监督微调, 证明了 transformer 对学习词向量的强大能力, 在 GPT-1 得到的词向量基础上进行下游任务的学习, 能够让下游任务取得更好的泛化能力。与此同时, 不足也较为明显, 该模型在未经微调的任务上虽然有一定效果, 但是其泛化能力远远低于经过微调的有监督任务, 说明了 GPT-1 只是一个简单的领域专家, 而非通用的语言学家。

图表 30 OpenAI GPT 发展历程

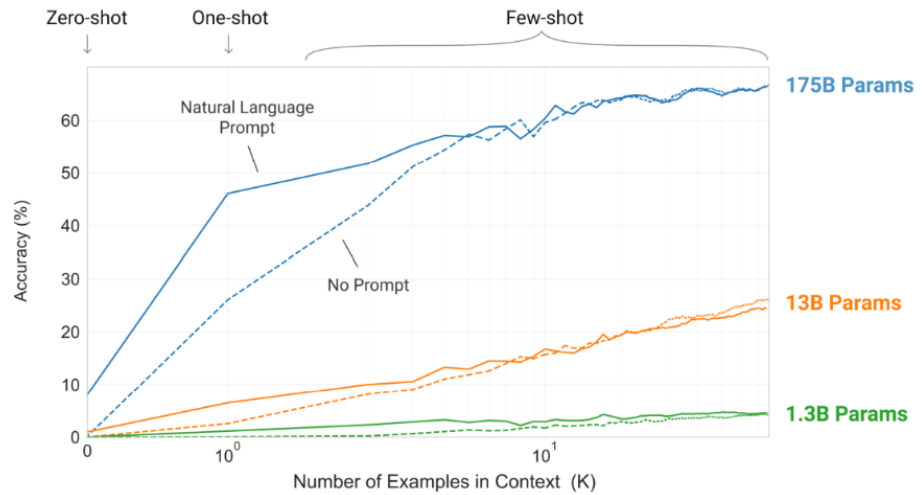
模型	推出时间	参数量	语料库
GPT-1	2018 年 6 月	1.17 亿	5GB
GPT-2	2019 年 2 月	15 亿	40G
GPT-3	2020 年 5 月	1750 亿	45TB
ChatGPT	2022 年 11 月	-	-

资料来源: OpenAI, 华安证券研究所

GPT-2 为了解决这一问题采用了多任务模式，其目标旨在训练一个泛化能力更强的词向量模型，它并没有对 GPT-1 的网络进行过多的结构的创新与设计，只是使用了更多的网络参数和更大的数据集，GPT-2 的核心思想是当模型的容量非常大且数据量足够丰富时，仅仅靠训练语言模型的学习便可以完成其他有监督学习的任务。所以虽然它验证了通过海量数据和大量参数训练出来的词向量模型能够迁移到其它类别任务中而不需要额外的训练，但其任务表现并不好，还有大提升空间。不过其表明了模型容量和数据量越大，其潜能越大。

于是 GPT-3 纳入了海量参数：1750 亿参数量还有超大的 45TB 的训练数据。在大量的语言模型数据集中，GPT-3 超过了绝大多数方法。另外 GPT-3 在很多复杂的 NLP 任务中例如闭卷问答，模式解析，机器翻译等也很准确。除了这些传统的 NLP 任务，GPT-3 在一些其他的领域也取得了非常好的效果，例如进行数学加法，文章生成，编写代码等。

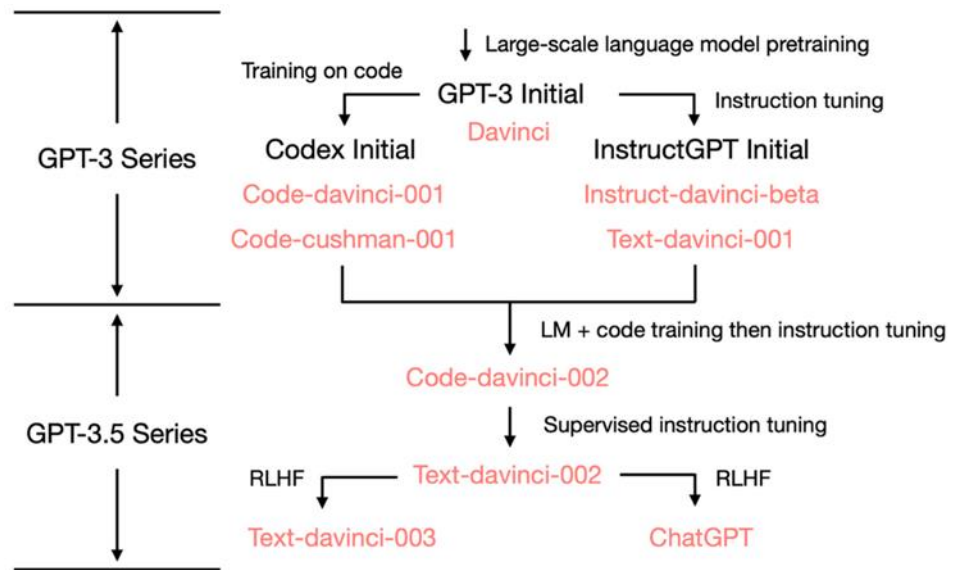
图表 31 不同模型参数量与模型精度的关系



资料来源：《Language Models are Few-Shot Learners》，华安证券研究所

加入人类反馈强化学习，GPT3.5 获得对话能力。GPT3.5 建立在 3.0 的微调之上，并加入了更加完整的人类反馈强化学习 RLHF 进行训练。2022 年 4 月至 7 月，OpenAI 开始对 code-davinci-002 模型进行 Beta 测试，其可能是最强大的针对自然语言的 GPT-3.5 变体，ChatGPT 就是从 code-davinci-002 进行指令微调得到的。此后 ChatGPT 在 2022 年 11 月发布，使用的基于人类反馈的强化学习的版本指令微调模型。

图表 32 GPT-3 到 GPT-3.5 的演进过程



资料来源：《拆解追溯 GPT-3.5 各项能力的起源》，华安证券研究所

4.4 国内外 AI 大模型项目百舸争流

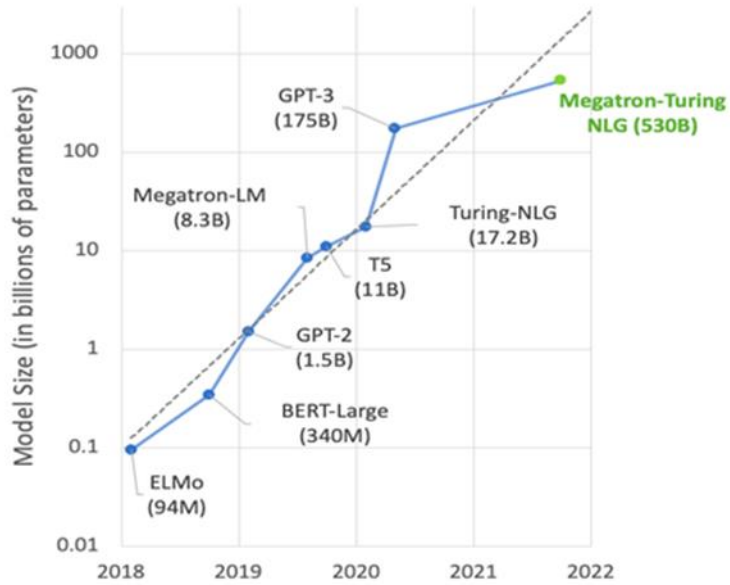
除了 GPT-3 模型外，各大公司正在孵化的大模型项目数量也相当可观。

MT-NLG：微软英伟达强强联手，软硬结合引领行业新景。2021 年 10 月 11 日，微软和英伟达推出由 DeepSpeed 和 Megatron 驱动的 Megatron-Turing 自然语言生成模型（MT-NLG），具有 5300 亿个参数。MT-NLG 的参数数量是当时该类型最大模型的 3 倍，并且在广泛的自然语言任务中如阅读理解、常识推理、自然语言推理、词义消歧等方面表现出较强的准确性。基于 105 层 transformer 的 MT-NLG 在多个方面方面改进了当时最先进模型，并为大规模语言模型在模型规模和质量方面设置了新标准。

硬件方面，模型训练是在基于 NVIDIA DGX SuperPOD 的 Selene 超级计算机上进行的，实现的系统吞吐量为：420 台 DGXA100 服务器上考虑了 5300 亿参数模型（批量大小为 1920）的系统端到端吞吐量，迭代时间为 44.4 秒、GPU 113 万亿次/秒。

Switch Transformers：Google 推出的首个万亿级语言模型。相比 1750 亿参数的 GPT-3，谷歌 Switch Transformers 则直接将该数值拉升至 1.6 万亿，且相比于 OpenAI 在 GPT-3 里所使用的 Sparse Attention，需要用到稀疏算子而很难发挥 GPU、TPU 硬件性能的问题。Switch Transformer 不需要稀疏算子，可以更好的适应 GPU、TPU 等硬件。

图表 33 大模型发展迭代图



资料来源：英伟达开发者网，华安证券研究所

文心一言：百度集成 NLP 和 CV，多级体系覆盖诸多领域。2022 年 11 月 30 日，百度集团在 WAVE SUMMIT+2022 深度学习开发者峰会带来了文心大模型的最新升级，包括新增 11 个大模型，大模型总量增至 36 个，构建起国内业界规模最大的产业大模型体系。在模型层，文心大模型涵盖基础大模型、任务大模型、行业大模型的三级体系；在工具与平台层升级了大模型开发套件、文心 API 和提供全流程开箱即用大模型能力的 EasyDL 和 BML 开发平台，有效降低应用门槛；新增产品与社区层，包括 AI 创作平台“文心一格”、搜索系统“文心百中”和样谷社区，让更多人感受到 AI 大模型技术带来的新体验。截至目前，文心已累计发布 11 个行业大模型，涵盖电力、燃气、金融、航天、传媒、城市、影视、制造、社科等领域，加速推动行业的智能化转型升级。

图表 34 百度文心生态图



资料来源：百度文心官网，华安证券研究所

除行业大模型外，百度目前新增了5个基础大模型和1个任务大模型，包括：知识增强轻量级大模型、跨模态理解大模型、跨模态生成大模型、文档智能大模型、单序列蛋白质结构预测大模型和代码大模型。其中，知识增强轻量级大模型 ERNIE 3.0 Tiny 具备优秀的泛化能力，同时相对于超大参数模型，推理速度提升数十倍到百倍，能够显著降低超大参数模型落地的成本。百度计划在3月完成文心一言的内部测试，然后向公众正式开放使用。

图表 35 ERNIE3.0 模型架构

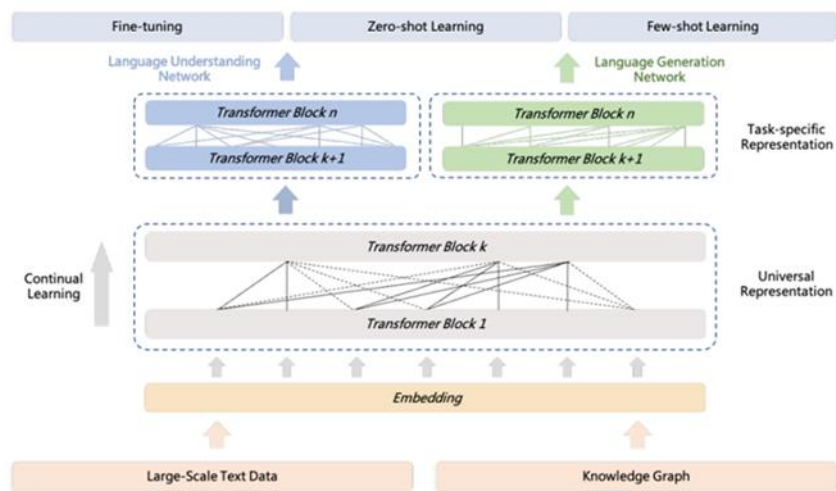


Figure 1: The framework of ERNIE 3.0.

资料来源：《ERINIE 3.0 Titan: Exploring Larger-scale Knowledge Enhanced Pre-training for Language Understanding and Generation》，华安证券研究所

华为盘古云布局已久，云脑支持场景广阔。华为自 2020 年开始在大模型开始有布局，2021 年发布了鹏城盘古大模型。鹏城盘古基于“鹏城云脑 II”和国产 MindSpore 框架的自动混合同步模式实现在 2048 卡算力集群上的大规模分布式训练，训练出业界首个 2000 亿参数以中文为核心的预训练生成语言模型。鹏程·盘古 α 预训练模型支持丰富的场景应用，在知识问答、知识检索、知识推理、阅读理解等文本生成领域表现突出，具备很强的小样本学习能力。

图表 36 MindSpore 自动并行框架

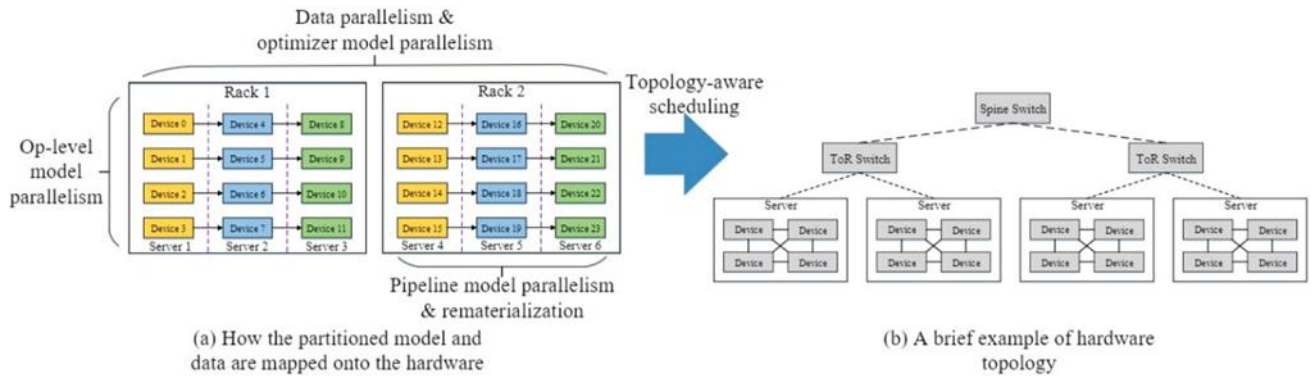
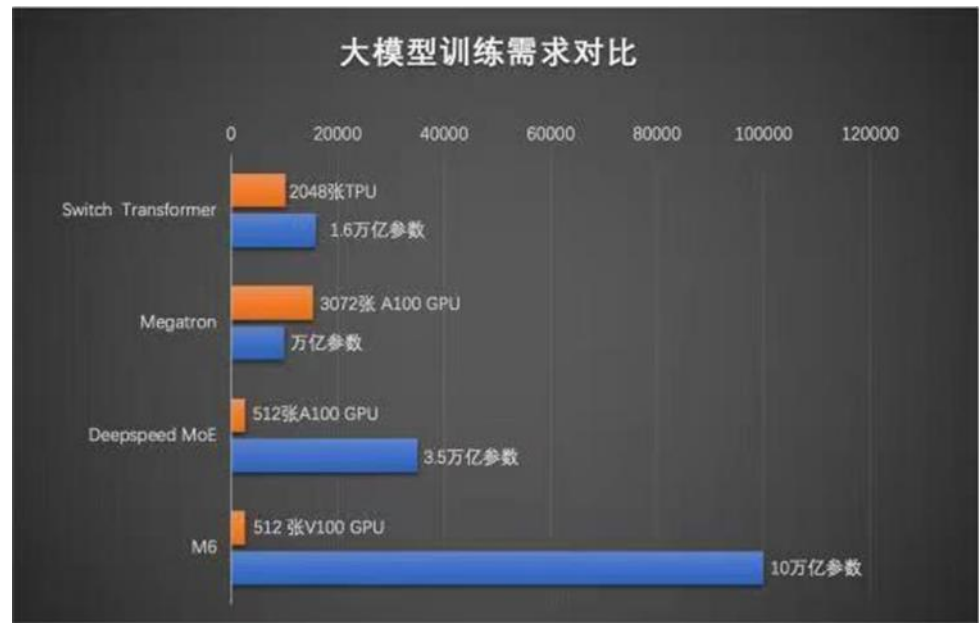


Figure 5: A combined parallelization of the model, and how it is scheduled to the cluster.

资料来源：华为昇腾社区，华安证券研究所

阿里 M6: 出色的低碳低能耗属性。阿里巴巴达摩院在 2021 年开发出了超大规模中文多模态预训练模型 M6。目前，其参数已从万亿跃迁至 10 万亿，规模远超谷歌、微软此前发布的万亿级模型，成为全球最大的 AI 预训练模型。同时，M6 做到了业内极致的低碳高效，使用 512 块 GPU 在 10 天内即训练出具有可用水平的 10 万亿模型。相比去年发布的大模型 GPT-3，M6 实现同等参数规模，能耗为其 1%。M6 的优势在于将大模型所需算力压缩到极致，通过一系列技术突破，达摩院和阿里云只用了 480 块 GPU 就训练出了 M6，相比英伟达用 3072 块 GPU 训练万亿模型、谷歌用 2048 块 TPU 训练 1.6 万亿模型（1 TPU 约等于 2~3 GPU），M6 省了超过八成算力，还将效率提升了近 11 倍。

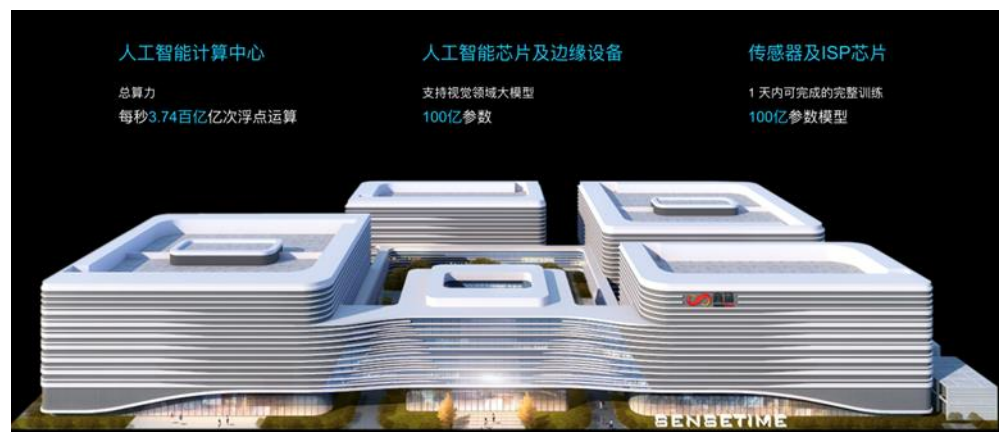
图表 37 大模型训练需求对比



资料来源：元宇宙网，华安证券研究所

商汤在 AIGC 的不同领域有多年布局，从文字，到图片，以及视频和动画的 AIGC，团队都从技术和产业长期投入，团队更多专注在视频的 AIGC，并叠加商汤自研的类似于 GPT 的生成式内容进行短视频等创作。基于商汤的 SenseCoreAI 大装置，在视觉大模型领域，商汤已训练和构建了超过 300 亿量级模型参数超大基模型，可以有效支持相关应用。

图表 38 SenseCore 商汤 AI 大装置中心



资料来源：SenseCore 官网，华安证券研究所

4.5 相关标的

1) 商汤科技

以 AI 生产力平台为基础，四大板块齐头并进。公司的 AI 生产力平台 SenseCore 由

模型层、深度学习平台、计算基础设施三个部分架构而成。1) **模型层**: 已开发超过 4.9 万个商用人工智能模型; 算法开源计划 OpenMMLab 在 GitHub 上超 60000 颗星; OpenDILab 开源平台, 已发布超过 60 个通用决策人工智能算法系列。2) **深度学习平台**: 高效利用 GPU 集群算力, 训练单个大模型时可以在一千块 GPU 上取得超过 90% 的加速效率。3) **计算基础设施**: 公司正在建立人工智能计算中心, 预计能产生每秒 3.74 百亿亿次浮点运算算力, 算力超过 2.5 exaFLOPS, 相较于 2021 年底扩大 114%; 公司研发的人工智能芯片及边缘设备支持视觉领域大模型 100 亿参数; 公司研发的传感器及 ISP 芯片 1 天内可完成的完整训练 1000 亿参数模型。在 SenseCore 底座基础上, 公司开发了智慧企业、智慧城市、智慧生活、智慧汽车四大板块:

智慧商业: 智慧商业是公司的主要业务之一, 营收占比超过 40%。公司依托 SenseCore 基座打造了 SenseFoundry Enterprise, 形成商业空间管理、住宅物业管理、工业引擎等具体解决方案。截至 2021 年, 该业务客户数量扩大至 922 家。

智慧城市: 2021 年公司智慧城市收入占比达 46%。公司研发的 SenseFoundry 主要面向出行和交通管理、城市服务和环境保护等领域提供解决方案。目前公司在中国智慧城市计算机视觉软件市场份额第一, 中国智慧应急人工智能与大数据市场份额第一。截至 2022 年上半年, 累计有 155 个城市部署城市方舟, 包括 16 个超千万人口大型城市及 4 个海外城市。

智慧生活: 公司是智能手机产业的头部 AI 软件供应商, 截至 2022 年 6 月 30 日, 已累计有 180 多个手机型号的超过 17 亿台手机预装了商汤的各类 AI 算法。公司的 SenseMARS 内置了 AI 生成内容、三维世界重建、数字人及虚拟形象等模块, 目前已覆盖的空间面积突破了 1000 万平方米, 覆盖了 120 多个大型游乐园区、商场等。

智慧汽车: 商汤科技推出的 SenseAuto, 以 SenseCore 为基石, 以 SenseAuto Empower 为底座, 在智能驾驶、智能座舱、车路协同、L4 级无人驾驶、无人驾驶小巴推进全线产品化商用。公司的智能驾驶和智能座舱产品累计前装定点数量达 2300 万台, 覆盖未来五年内量产的 60 多款车型。

图表 39 商汤科技主要产品结构



资料来源：公司官网，华安证券研究所

2) 科大讯飞

AI 应用快速发展，构建人工智能产业生态。公司自创业以来持续聚焦智能语音、自然语言理解、机器学习推理及自主学习等人工智能核心技术研究并始终保持国际前沿技术水平。目前公司已经从语音 AI 逐渐拓展至到教育、医疗、智慧城市、消费、智能汽车等多领域布局。同时，随着 ChatGPT 掀起热潮，公司的类 ChatGPT 技术也将于 5 月落地，率先用于 AI 学习机。

智慧教育业务：智慧教育业务是公司的第一大业务，在收入中占比 30%左右。公司构建了面向 G/B/C 三类客户的业务体系：G 端业务主要以市县区等区域建设为主体；B 端业务主要以学校建设为主体；C 端业务主要以家长用户群自主购买为主。目前讯飞智慧教育产品已在全国 32 个省级行政单位以及日本、新加坡等海外市场应用。

智慧医疗业务：公司基于人工智能技术，通过医学语义计算等核心技术，实现面向基层医生的全科辅助诊断服务。截至 2022 年 6 月 30 日，全科医生助理已累计覆盖全国 30 个省、353 个区县并常态化应用。全科医生助理累计为基层医生提供了 4.4 亿余次 AI 辅助诊断建议，经全科医生助理提醒而修正诊断的有价值病例累计超过 54 万例。

智慧城市业务：公司围绕政府数字化转型和政法业务办公办案质效提升，打造了“城市大脑”、“一网统管”、“一网通办”、“疫情防控”、“智慧园区”等解决方案。目前，智慧城市主营业务已覆盖全国 20 余省、共 50 多个地市，取得良好的建设成效。

开放平台及消费者业务：讯飞人工智能开放平台是首批“国家新一代人工智能开放创新平台”，在此基础上进一步构建了讯飞 AI 营销平台、讯飞智能工业平台等能力平台。目前，讯飞开放平台已开放 496 项 AI 能力及场景方案，聚集 343 万开发者团队，总应用数达 152 万，链接超过 500 万生态合作伙伴。

运营商、智慧汽车、智慧金融等企业客户 AI 解决方案业务: 面向汽车厂商客户群体方面, 公司的智慧汽车业务面向国内外汽车厂商, 提供飞鱼智能助理、飞鱼 OS、飞鱼智能车载音频管理系统、AI 销服等产品及解决方案。公司产品已覆盖中国主流自主和合资品牌车厂, 包括: 一汽、丰田等国内外汽车厂商及蔚来等新势力厂商。同时公司还具有面向运营商客户群体和面向金融客户群体的业务解决方案。

图表 40 科大讯飞业务全景示意图



资料来源: 公司年报, 华安证券研究所

3) 云从科技

持续赋能 AI 领域, 打造高效人机协同操作系统和行业解决方案。 公司是首个同时承建三大国家平台, 并参与国家及行业标准制定的人工智能领军企业。公司以计算机视觉、语音识别为代表的人工智能单点技术为突破, 通过不断研发并优化人机协同操作系统和适配性强的 AIoT 设备, 推动人工智能在特定场景的应用。公司具有人工智能芯片平台、深度学习框架、AIoT 操作系统、算法算力平台、知识中台、自动驾驶平台、机器人开发平台等丰富的产品:

智慧金融业务: 公司智慧金融解决方案将相关算法能力落地为技术平台, 主要面向智慧支付、智慧营运、智慧触点、智慧风控四大领域, 形成了 5 大类金融业智能化转型解决方案簇, 以及 53 种解决方案。目前公司在智慧金融领域的客户已涵盖 6 大行、12 家股份制银行以及城农商行, 服务了超过 400 家金融机构和 10 余万个银行网点。

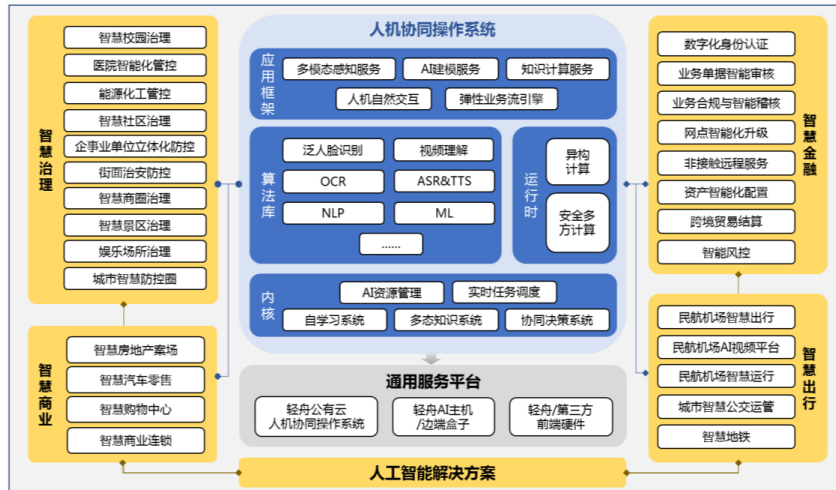
智慧治理业务: 公司的人工智能解决方案提供方主要服务公安、政法、交通、应急、文教卫、社区园区等政府和大型企事业客户, 公司通过对各类业务数据、场景数据的接入汇聚, 以智能解析引擎、数据分析引擎为驱动, 搭建了融智云知识生成与服务平台 (KAAS)。目前公司在智慧治理领域已经参与 5 个公安部平台建设, 服务全国 30 个省级行政区, 取得 50000+公安战果。

智慧出行业务: 公司以“人机协同平台”为核心, 通过 AI 科技为民航机场、客运、铁路、物流等交通运输体系提供智慧化支持。公司赋能成都双流机场、成都天府机场、北京大兴国际机场等机场, 为乘客提供“一张脸”走遍机场的便捷体验。目前公司在智慧出行领域产品和解决方案已覆盖了全国十大机场中 9 座枢纽机场、105 座民用枢纽机场与航

司、日均服务旅客达百万人次。

智慧商业业务：公司依托 AI 基座平台 CWOS 和 AI 应用平台，面向购物中心、连锁门店、停车场、无人售货机四大应用场景，提供智能业务解决方案。目前智慧商业领域产品和解决方案已辐射上千家 3C 服务门店、国内外汽车连锁品牌智慧展厅、众多知名连锁品牌、上亿全国各大购物中心、连锁门店消费者，为全球数亿人次带来智慧、便捷和人性化的 AI 体验。

图表 41 云从科技主要产品及服务图谱



资料来源：公司招股书，华安证券研究所

4) 依图科技

以人工智能芯片技术和算法技术为核心，持续深耕智能公共服务及智能商业领域。公司以人工智能算法和芯片等核心技术为基础，在城市管理、医疗健康、安全生产、交通出行和互联网服务等场景实现规模化商业落地和与产业的深度融合。目前公司已为国内 30 余省、自治区、直辖市及境外 10 多个国家和地区的 800 余家政府及企业终端客户提供产品及解决方案：

智能公共服务，公司智能公共服务主要包括智能城市和智慧医疗：

1) 智能城市，公司主要着力于三方面：**高性能算法赋能城市复杂场景。**公司在计算机视觉的多个细分领域、声纹识别、中文语音识别及自然语言理解技术均已达到世界领先水平，可对城市实体在复杂多样场景下产生的非结构化数据进行解析、识别和关联，构建出完整的城市实体关联关系。**高效算力硬件产品降低城市智能化门槛。**公司的原石系列智能服务器和前沿系列边缘计算设备，在显著提高城市视频智能解析吞吐效率的同时，大幅降低了算力功耗，从而降低了投资和运营成本。**“利旧”方案降低建设成本。**公司将大量已建设的非智能摄像机采集的原始数据在云端进行解析，与智能摄像机的解析结果在云端实现汇聚、融合分析，大大降低智能化建设的投资成本。

2) 智慧医疗方面，公司是业内少数具有以多模态人工智能技术解析多源异构医疗大数据能力，并具有自研医疗知识图谱的企业之一。公司在上海儿童医学中心建设的智慧儿童医院解决方案能够为患者单次就诊至少节约 90 分钟，落地一年来，该解决方案仅智能导诊应用就已累计为超过 27 万名患儿提供服务，诊前检验访问量突破 6 万人次。

智慧商业：智能商业业务覆盖园区管理、网点服务、安全生产、交通出行和互联网服务等场景，终端客户包括商业地产、金融业、制造业、交通运输、互联网等领域的企业客户。**1) 智慧金融领域，**公司依靠自研的生物识别、自然语言处理、智能语音交互、

态势感知等 AI 技术，提供从客户服务到风险防控全链路解决方案，助力金融机构重塑嵌入流程、融入场景、智能创新的服务模式。目前招商银行 1500 家网点已经接入依图 AI 识别系统。**2) 智慧零售领域**，公司将人工智能技术、创新的边缘与端产品与零售行业应用相结合，可实现线下零售场景的大数据实时分析、无底库挖掘、跨门店打通等数据智能服务。

图表 42 依图科技主要解决方案



资料来源：公司招股书，公司官网，华安证券研究所

5) 旷视科技

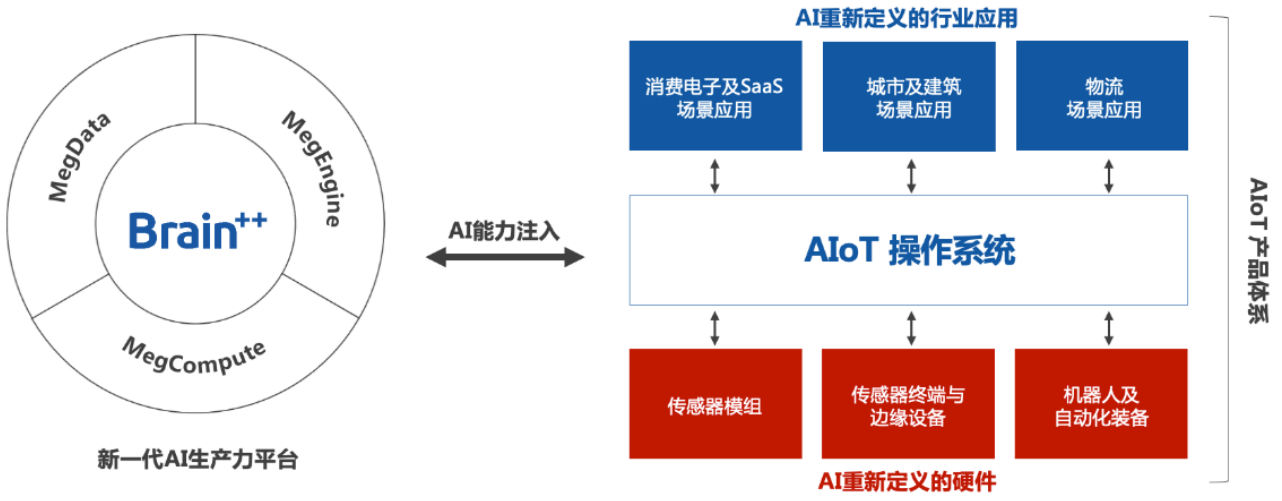
聚焦物联网场景，推动人工智能的商业化落地。公司以物联网作为人工智能技术落地的载体，通过构建完整的 AIoT 产品体系，面向消费物联网、城市物联网、供应链物联网三大核心场景提供经验证的行业解决方案。公司的 AIoT 软硬一体化解决方案包括以 Brain++ 为核心的 AI 算法体系，由 AIoT 操作系统和行业应用构成的软件，以及由传感器模组、传感器终端与边缘设备、机器人及自动化装备组成的硬件。

消费物联网：2012 年，旷视进入消费物联网领域，以 SaaS 产品的形式服务全球开发者及企业用户，为其提供数十种 AI 能力。随着智能设备的普及，公司与多家头部智能手机厂商等消费电子领域客户开展合作，累计为数亿台智能手机提供设备安全和计算摄影解决方案。

城市物联网：2015 年，公司进入城市物联网领域，推出了人工智能摄像头，并不断强化硬件能力以配合尖端计算机视觉算法。目前公司已累计打造 4 大品类、数十款硬件产品。同时，基于覆盖云边端的算法、软件、硬件能力，公司的城市物联网解决方案正应用于百余座国内城市、十余个国家和地区。

供应链物联网：2017 年，公司进入供应链物联网领域，随后推出了智慧物流操作系统“河图”，并自研包括 AMR 机器人、SLAM 导航智能无人叉车、人工智能堆垛机等多款智慧物流硬件。在供应链物联网领域，公司已向鞋服、医药、智能制造、零售电商等多个行业的客户提供了仓库、工场及零售店的智能化升级改造。

图表 43 旷视科技 AIoT 软硬一体化解决方案



资料来源：公司招股书，华安证券研究所

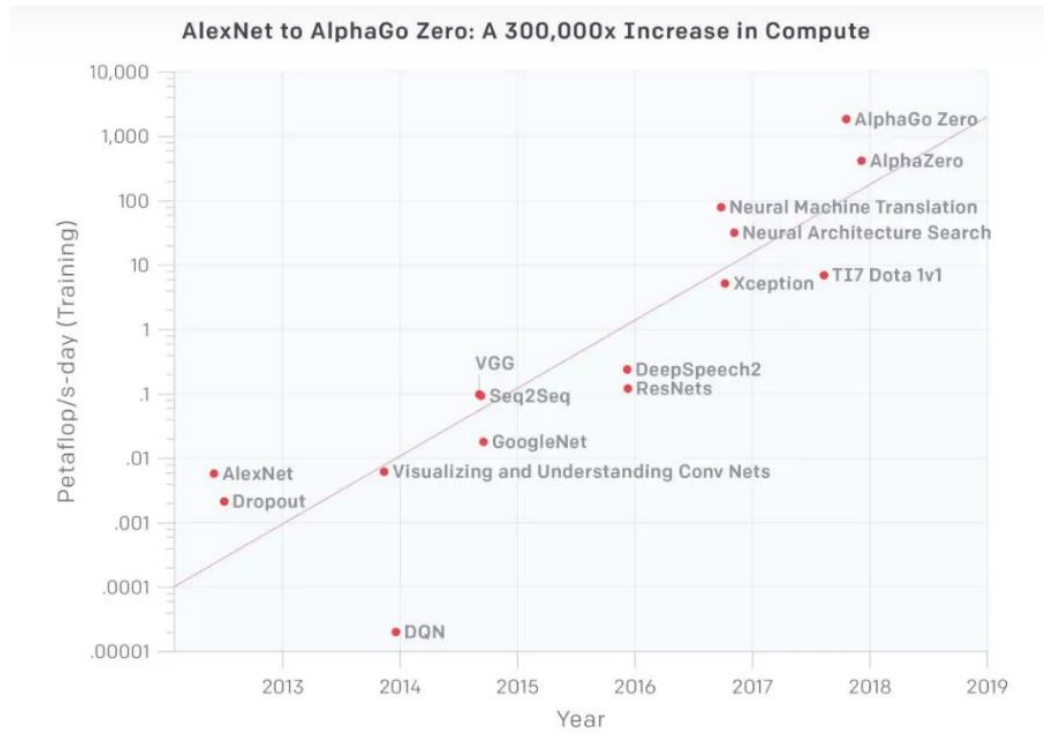
5 算力与网络是大模型运行的必要条件

5.1 大模型发展，算力需求激增

前文提到，ChatGPT 从初代模型，到当前的 3 代模型，参数量从 1.17 亿，提升至 1750 亿。同时训练数据量（语料库）也由 5GB 提升至 45TB。

随着参数量和语料库指数级的扩容，ChatGPT 类人工智能需要更充足的算力支持其处理数据，同时需要投入更多高性能的算力芯片来处理千亿级别参数量。英伟达的研究表示，GPT-3 模型需要使用 512 颗 V100 显卡训练 7 个月时间，或者使用 1024 颗 A100 芯片训练长达一个月的时间。2012 年以来，人工智能训练任务中的算力增长（所需算力每 3.5 月翻一倍）已经超越芯片产业长期存在摩尔定律（晶体管数量每 18 月翻一倍）。

图表 44 算力需求变化



资料来源：OpenAI，华安证券研究所

现阶段国内无法采购英伟达 A100、H100 等高端 GPU 产品，但算力性能上的差异可以通过提升算力芯片数量来弥补，因此对于算力芯片产品的需求也将更高。国内主流互联网厂商也有类 ChatGPT 产品正在开发，比如悟道和百度的文心等 AI 模型。随着国产 GPU, CPU, FPGA 产品性能的提升，人工智能的算力需求将为国产芯片厂商打开广阔的市场空间。同时，以 ChatGPT 为代表的 AI 技术浪潮的到来，对产业链相关芯片，模组，材料等环节均带来了海量的新需求。

5.2 GPU/GPGPU/FPGA 多路线支持算力

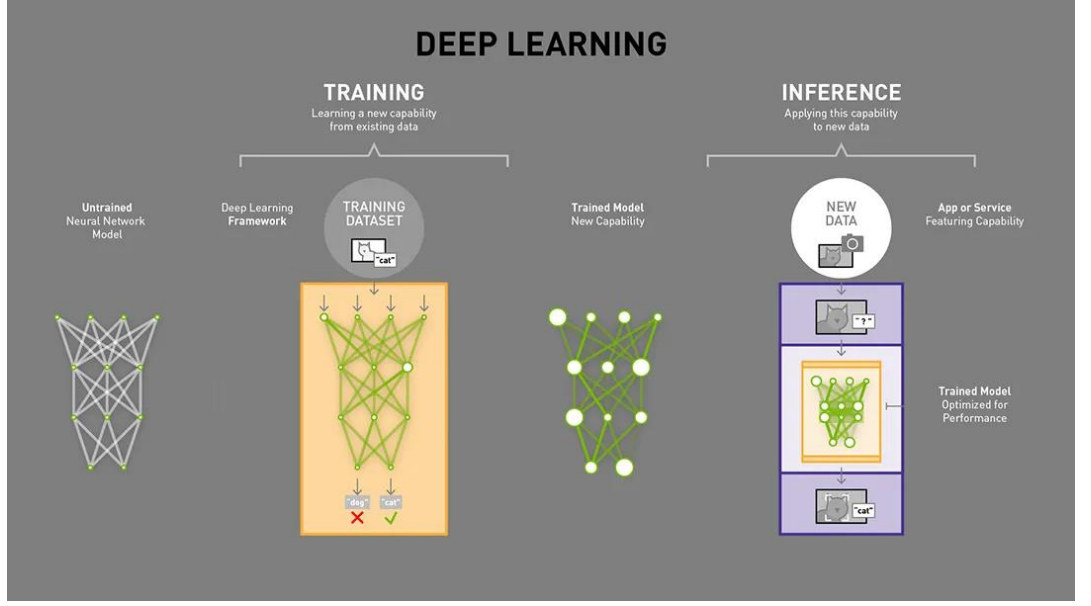
人工智能深度学习模型需要处理两大任务，即训练和推理。

- 1) 训练就是学习过程，通过大数据训练出复杂的神经网络模型，使得整个系统可以适应特定的功能。因此，训练需要庞大的算力来处理数据并搭建网络模型，所需的芯

片需要具有高算力和通用性。

- 2) 而推理在训练之后，主要是在已有的训练完善的模型基础上，输入新数据进行推断，能耗、时延、效率等因素都是影响推理能力的因素。但同时，推理环节不需要在庞大的神经网络之中反复调整参数，因此对算力的要求相对于训练要低很多。

图表 45 深度学习中训练和推理的演示



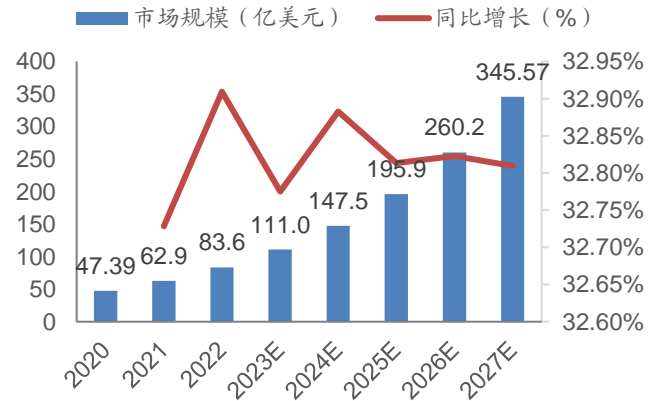
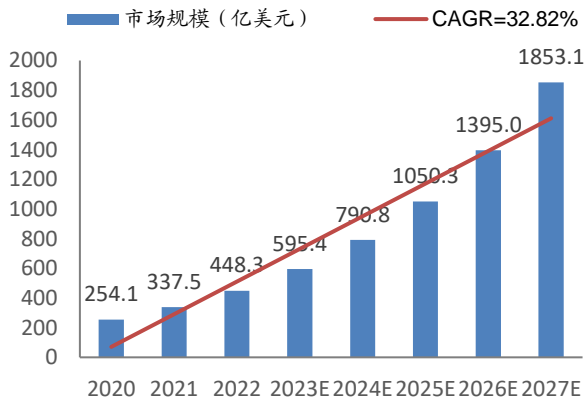
资料来源：NVIDIA，华安证券研究所

一般来说，在深度学习的训练中，GPU 因为其并行串联的优势，相比较于 CPU 更加适应处理大数据。最早为了解放 CPU 处理其他任务和计算的需求，图形计算全部交由 GPU 来做。因此传统的 GPU 的主要功能是做图形渲染（实质是做图形渲染的计算），但随着 AI 计算需求的发展，GPGPU 应运而生，GPGPU 即为通用计算 GPU，指去掉 GPU 的图形显示部分，将其余部分全部投入通用计算，其在 AI、数据分析和 HPC 等场景下可以广泛应用。总而言之，通用 GPU 在加速硬件能力上的优势比较明显，在深度学习训练方面比较适用。

根据相关市场统计，全球 GPU 的市场已经达到了 448 亿美元的规模，在 AI 的发展需求下，通用型 GPU 的市场正在快速成长。

图表 46 全球 GPU 行业市场规模

图表 47 中国 GPU 行业市场规模



资料来源：华经产业研究院，华安证券研究所

资料来源：华经产业研究院，华安证券研究所

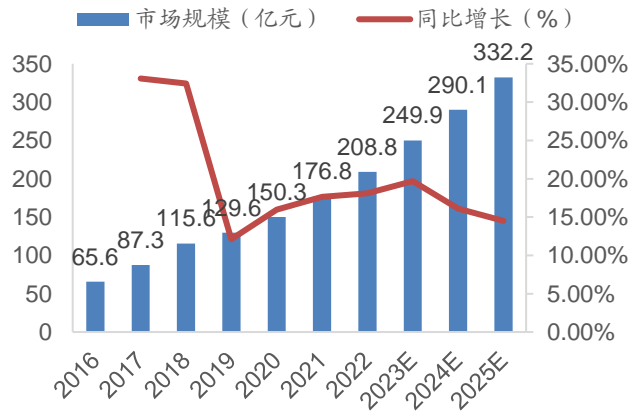
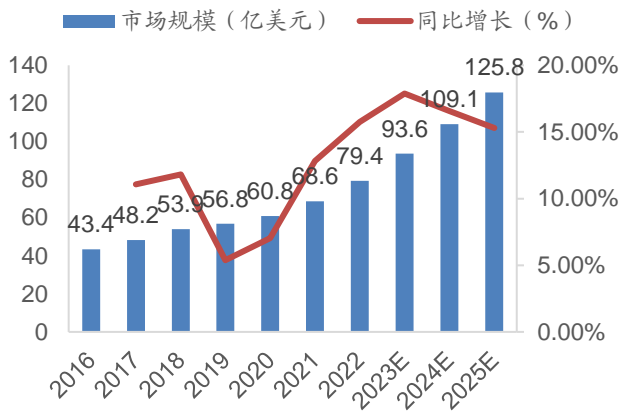
CPU 是计算机运算和控制的核⼼，其工作原理是将指令依序执⾏，其串⾏运算的特点使其更加适应逻辑控制。因此在深度学习模型之中，CPU 搭配 GPU 是目前的主流方案，但随着各类算法在 FPGA/ASIC 芯片上的优化，以及其本身性能的提升和成本的优化，FPGA 和 ASIC 也会在人工智能领域上有着更⼴的应用。

现阶段，可编程的 FPGA 芯片也逐渐提升市场份额。实际应用中，微软利用 FPGA 加速 Azure 云服务、必应等数据中心服务中的实时人工智能。FPGA 具备快速、低功耗、灵活和高效的优点。硬件可编程的特性使得 FPGA 在 AI 训练中既能提供充足的算力，又具有灵活性，可以重新编程以适应不同任务的需要。与 GPU 类似，FPGA 也是配合 CPU 进⾏加速。国内 FPGA 厂商中，紫光国微于 2022 年推出了 2x 纳米的低功耗 FPGA 系列产品，新一代 1x 纳米更高性能 FPGA 系列产品也在顺利推进中，进一步完善了产品种类。复旦微电具备 65nm 制程千万门级和 28nm 制程亿门级产品，目前以 28nm 制程的 FPGA 产品为主。安路科技的 FPGA 芯片产品形成了由 PHOENIX 高性能产品系列、EAGLE 高效率产品系列、ELF 低功耗产品系列组成的产品矩阵。

根据市场统计，全球 FPGA 芯片规模在 79 亿美元左右，随着 AI 和军工等下游行业的需求增长，全球 FPGA 市场空间仍在持续上升之中。

图表 48 全球 FPGA 芯片市场规模

图表 49 中国 FPGA 芯片市场规模



资料来源：华经产业研究院，华安证券研究所

资料来源：华经产业研究院，华安证券研究所

ASIC 是专用集成电路，与 FPGA 相比，在量产后 ASIC 的成本会低，性能会相对优异，但并没有 FPGA 的可编程的特性，灵活调整的空间较小。同时，ASIC 的一次性开发成本也比较高。作为全定制化的芯片，ASIC 将随着人工智能产业生态的逐渐完善获得更广阔的空间。

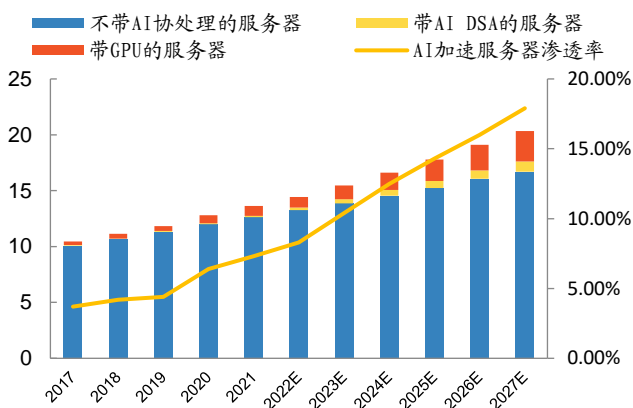
此外，人工智能与 chiplet 相结合是产业发展新趋势之一。算力与性能需求提升的背景下，小芯片技术是行业发展趋势之一。复用 IP+先进封装的形式大幅地减少了功耗和成本，同时提升了设计的灵活性和芯片性能。现阶段，CPU，SOC 等采用 chiplet 模式重构，已经成为行业的主流趋势，比如国内华为早前推出过基于 chiplet 技术的 7nm 鲲鹏 920 处理器，AMD 也推出了 chiplet 形式的 GPU 图形处理器 Navi 31，芯原股份在 chiplet 技术中的 IP 环节有着广泛的积累和探索，现已储备了的六类处理器 IP、1,400 多个数模混合 IP 和射频 IP。

5.3 高带宽网络是对人工智能算力的重要支撑

大模型训练和推理使用了 AI 计算集群。AI 大模型通常需要部署在 AI 计算集群以实

现训练和推理加速并实现最佳能效比。以目前独家开放了 ChatGPT 调用能力的微软 Azure 为例，其 AI 基础设施由互联的英伟达 Ampere A100 Tensor Core GPU 组成，并由 Quantum infiniband 交换机提供强大的横向扩展能力。根据微软宣布，为 OpenAI 开发的超级计算机超过了 28.5 万个 CPU 核心、1 万个 GPU，每台 GPU 服务器网络连接能力为 400Gbps，位列全球超级计算机前五。目前，英伟达的 AI 计算集群整体解决方案（如 DGX A100 系列）是 AI 集群投资的主流选择，除此之外部分云和互联网大厂选择了 CPU+FPGA+GPU+AI DSA 异构算力自己搭建 AI 计算中心。根据 YOLE 预测，全球 AI 加速服务器渗透率将在 2027 年达到 17.9%，AI/GPU 加速服务器出货量复合增速高达 39.8%/20.3%。

图表 50 AI 加速服务器全球出货量渗透率迅速提升（单位：百万）



资料来源：YOLE，华安证券研究所

图表 51 英伟达 DGX A100 SU (包括 20 台 DGX A100 服务器) 是 SuperPOD 集群基本组成单位

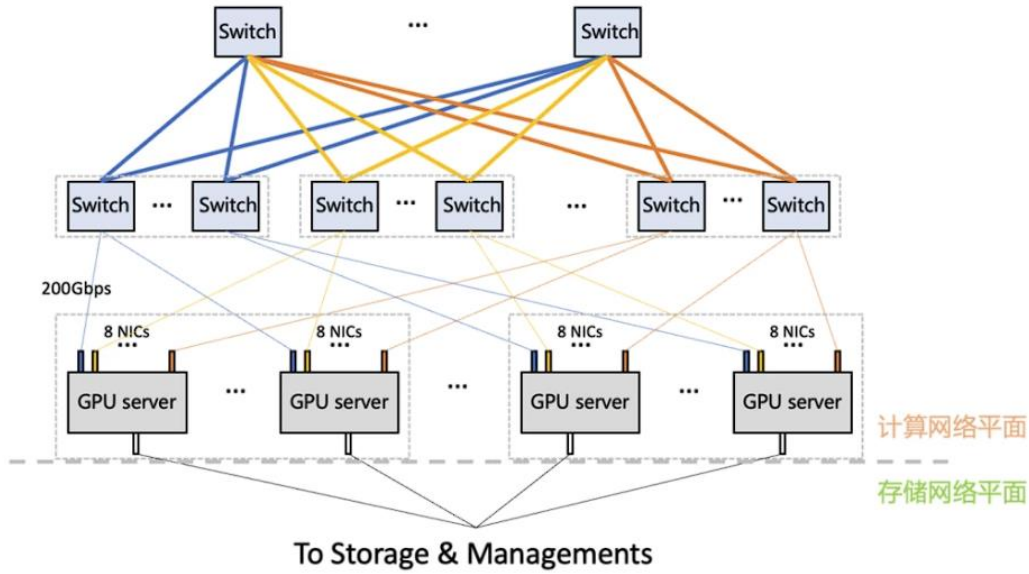


资料来源：英伟达官网，华安证券研究所

AI 集群对网络性能提出更高要求。AI 算力集群对高带宽网络的需求增加来自于东西向流量的快速爆发。在训练工作负载上，为了减少训练时间 AI 集群通常采用了多 GPU 分布式训练，主流策略包括数据并行训练（即单个模型在多个设备上复制，分别处理不同批次的数据）、模型并行训练（即单个模型不同部分在不同设备运行）以及流水线并行训练（将模型的不同层防止到不同计算设备降低单个设备显存消耗）等，并行训练的核心问题是如何解决服务器间通信的瓶颈（单机内卡间通信如 H100 的 NVLINK 带宽高达 900GB/s），对于高速光模块的需求愈加迫切且迭代速度高。在推理工作负载上（AI 算力集群的更多时间用于用户访问的处理，即推理），由于 vGPU 虚拟计算的使用以及多用户并发访问，也需要多台服务器处理。

服务器节点多、跨服务器通信需求巨大，网络带宽性能成为 GPU 集群系统的瓶颈，解决的方式包括增加单节点通信带宽与降低网络收敛比。前者如腾讯提出的星脉组网架构单服务器节点配备了 8 张 RoCE 网卡，提供 1.6Tbps 超高带宽接入，与传统 100Gbps 的集合通信对比，可带来 10 倍以上通信性能提升。

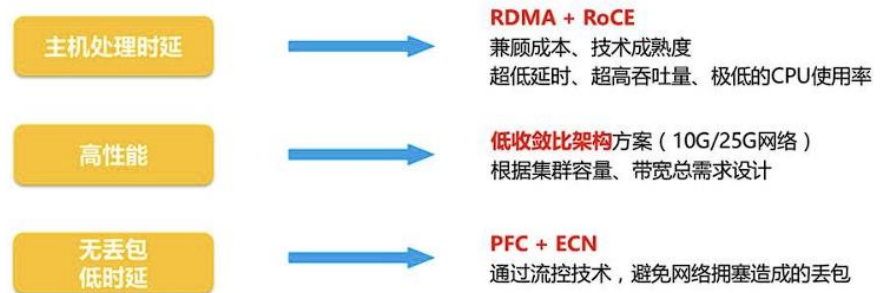
图表 52 腾讯星脉 AI 集群组网架构



资料来源：腾讯网，华安证券研究所

后者是在设备转发层面尽量优化网络时延，实现高性能无损网络，目前主要手段是无带宽收敛（1:1）的网络架构设计以及基于 PFC 和 ECN 功能的优先队列管理和拥塞管理。无带宽收敛比设计意味着交换机下行和上行流量一致，将增加上行端口数量或端口带宽，而优化流控技术需要可编程的三层网络交换机，以上将导致交换机端口数量增加以及价值量提升。

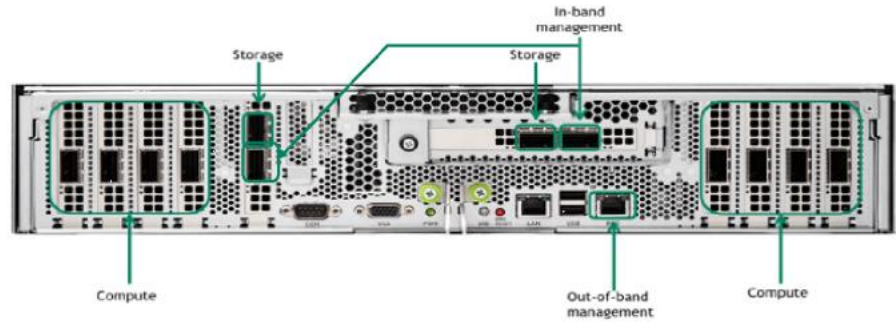
图表 53 AI 集群高性能方案关键技术组合



资料来源：51CTO，锐捷网络，华安证券研究所

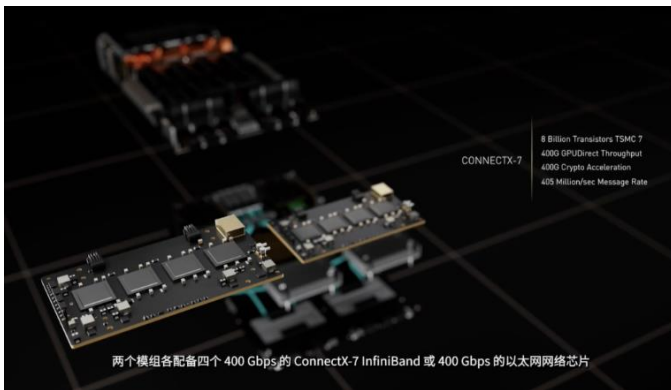
部署 AI 的云大厂使用了高端网卡和高速光模块。以英伟达 DGXA100 服务器为例，其网络拓扑包括四大部分：计算网络、存储网络、带内管理和带外管理，其计算网络每台服务器使用了 8 块 200Gb/s 的 ConnectX-6 HCAs 智能网卡，存储网络使用了 2 块 ConnectX-6 HCAs，带内管理使用了 2 个 100Gbps 光口，带外管理使用千兆以太网接口，单台服务器满配需要 10 个 200G AOC (或 2*100G)/多模光模块和 2 个 100G AOC/多模光模块。而在最新发布 DGX H100 集群中，每台 H100 通过 8 枚 ConnectX-7 400Gbps 网卡互联，通过 4 个 800G 光模块连接到 Quantum-2 交换机（可接入 8 台 H100）。

图表 54 英伟达 DGX A100 服务器使用了 200G 高速光连接



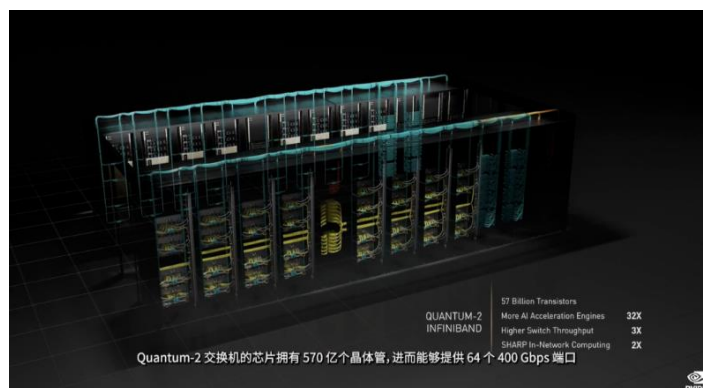
资料来源：英伟达《DGX SuperPOD 白皮书》，华安证券研究所

图表 55 DGX H100 使用了 400G ConnectX-7 网卡



资料来源：英伟达 GTC 2022 演讲，华安证券研究所

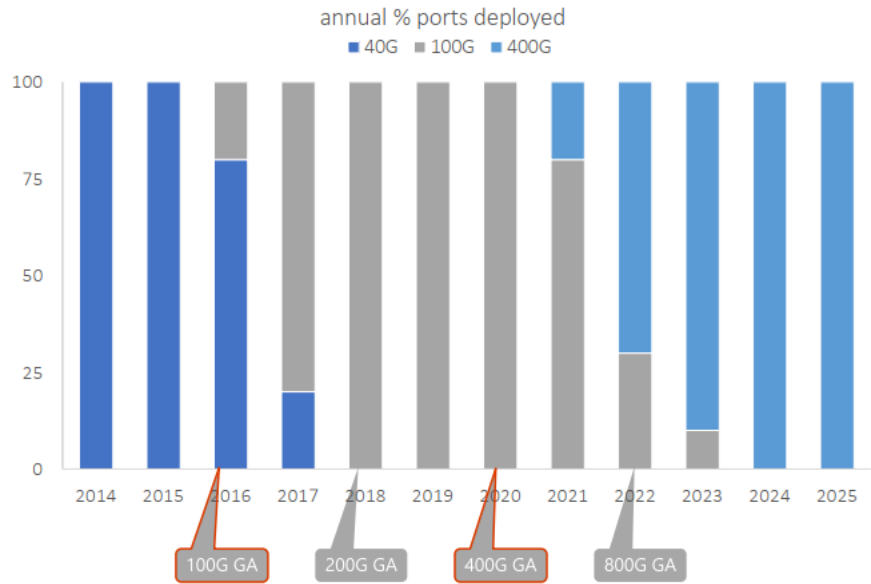
图表 56 英伟达 Quantum-2 InfiniBand 交换机提供 64*400Gbps 交换容量



资料来源：英伟达 GTC2022 演讲，华安证券研究所

在数据中心集群网络方面，北美 Top4 云厂商已经全面进入 400G 部署阶段。以微软数据中心为例，其在 2016 年开始批量部署 100G，由于 AI 算力的快速增长，2021 年开启 400G 部署，并预计 2024 完成 400G 的全面部署，加速向 800G 迈进。而对于 400G 部署更早的亚马逊和谷歌，预计在 2023 年开启小规模 800G 部署，更大的端口带宽将带来更高的端口密度，节约空间的同时每 Gbps 成本下降。根据 LightCounting 预测，全球数据中心以太网光模块市场将在 2027 年突破 100 亿美金，5 年 CAGR 11.5%，其中 800G CAGR 达 72%。

图表 57 微软数据中心网络 2024 年部署将全部为 400G



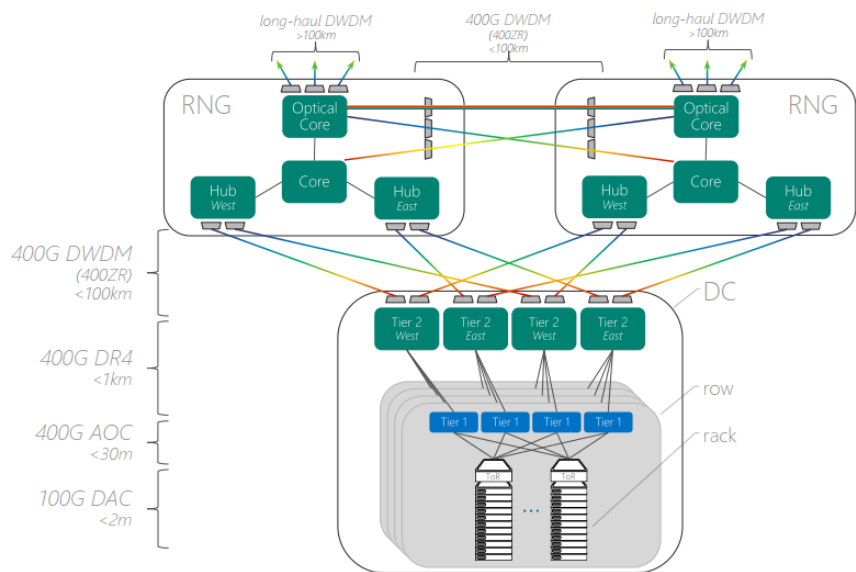
资料来源：微软《The race to 800G》，华安证券研究所

单机网卡数增加和网络收敛比降低增加了交换机和光模块用量。根据数据中心光模块需求量计算公式（流量法）：

- 1) 服务器到 TOR 交换机光模块：服务器网卡端口数*2
- 2) TOR 到 LEAF 交换机光模块：服务器流量/一级收敛比/TOR 上联端口速率*2
- 3) LEAF 到 SPINE 光模块：LEAF 流量/二级收敛比/LEAF 上联端口速率*2

假设普通 Hyperscale 数据中心和 AI 数据中心分别有 X 台服务器，每台服务器网卡速率为 100Gbps，数量分别为 2/10，网络收敛比分别为 2:1、1:1，交换机上联端口速率为 400G，计算得到普通 Hyperscale 数据中心光模块（包括 AOC/DAC）用量为 4.75X，而 AI 数据中心光模块用量为 23.75x，光模块用量大幅提升。

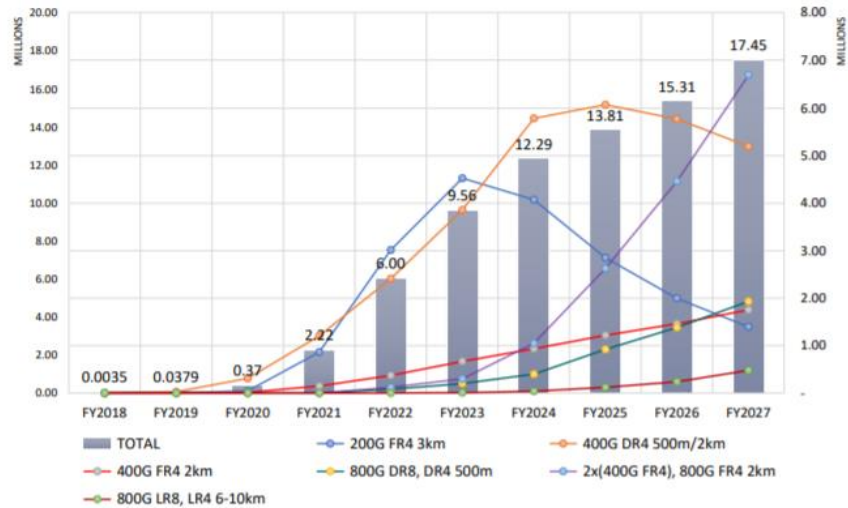
图表 58 微软 400G 数据中心网络架构



资料来源：微软《The race to 800G》，华安证券研究所

结论：AI 计算中心光模块量价齐升。根据以上探讨，我们看到 AI 算力的增长带来 AI 计算集群网络架构的改变，高速率光模块用量大幅提升，同时端口带宽速率加速迭代，单端口价值量不断提升，驱动数据中心光模块市场规模高速增长。根据 LightCounting 2021 年的报告，全球 200G 以上以太网光模块出货量将在 2027 年达到 1745 万只，5 年 CAGR 23.8%。

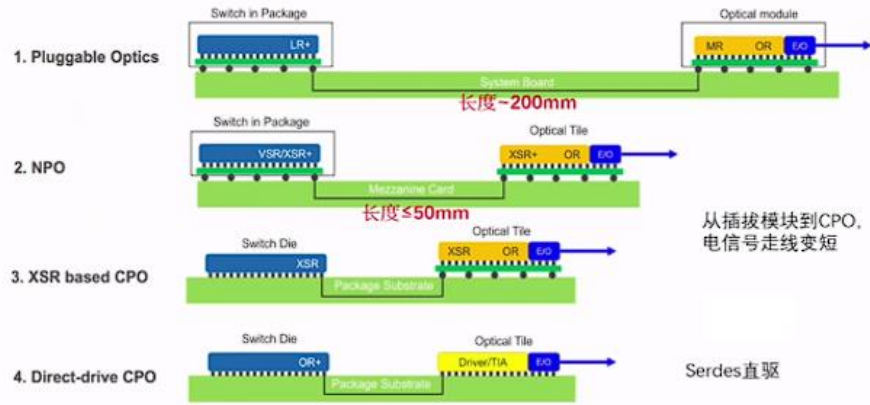
图表 59 数据中心 200G 以上高速光模块出货量高速增长



资料来源：剑桥科技 2021 年报，LightCounting，华安证券研究所

CPO 探讨：一系列技术问题的解决和产业链结构的重塑。CPO（光电共封装）的主要形态为交换芯片与光引擎封装在一块基板上，交换芯片与光引擎通过 XSR SerDes 直联。我们认为 CPO 可能是数据中心交换机端口发展到 1.6Tbps 以上一种可能的光模块形态，虽然 AI 算力将加速数据中心交换机带宽的增长，但 CPO 的渗透仍将是个缓慢的过程。CPO 主要解决的是高速 SerDes 信号衰减和功耗问题（1.6T 光模块可能使用 200G SerDes）、1.6T 光模块多通道设计和良率难题以及相应带来的成本高企，但目前也有一些问题需要解决，比如光源设计问题（如设计在交换机内光引擎附近容易出现热失效，一种可能的思路为外置光源 ELS 但相应也会带来功耗的增加和布线的成本大幅提升）、与交换机芯片的联调问题、以及将来的替换维护问题（需拆机维护）。我们认为，CPO 或将重塑数通产业链结构，话语权可能会向交换芯片厂商、交换机厂商倾斜，同时光引擎封装、硅光芯片、保偏光纤、CW 激光器、封装基板等环节也将成为新增投资机会。

图表 60 交换机芯片通过 SerDes 直驱光模块实现高频电信号走线缩短



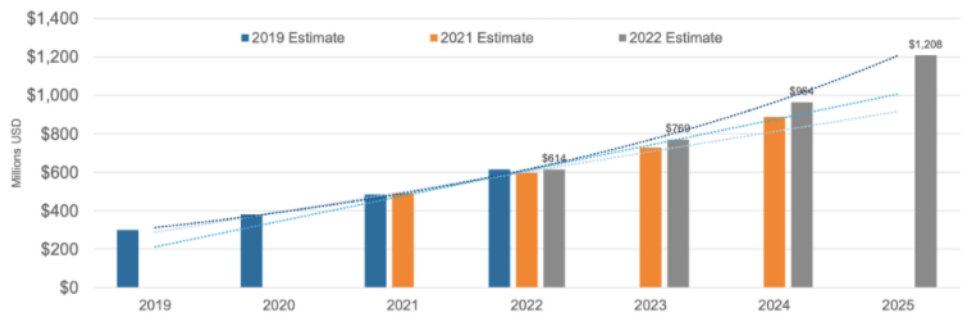
资料来源：华工科技，华安证券研究所

5.4 量子计算有望成为 AI 算力突破的“神助攻”

量子计算产业蓬勃发展。量子计算机基于量子力学原理构建，量子态叠加原理使得量子计算机的每个量子比特（qubit）能够同时表示二进制中的 0 和 1，相较经典计算机算力呈指数级爆发式增长。目前量子计算机已被证明在特定计算任务上具备指数加速能力，即实现所谓的“量子霸权”，目前量子计算产业化需要解决的三大目标一是开发大规模可容错的量子计算机以尽快实现可编程计算；二是开发精妙的量子计算算法以在特定场景的商业化价值；三是解决量子计算资源稀缺性难题，通过云平台提升综合性普惠服务能力。

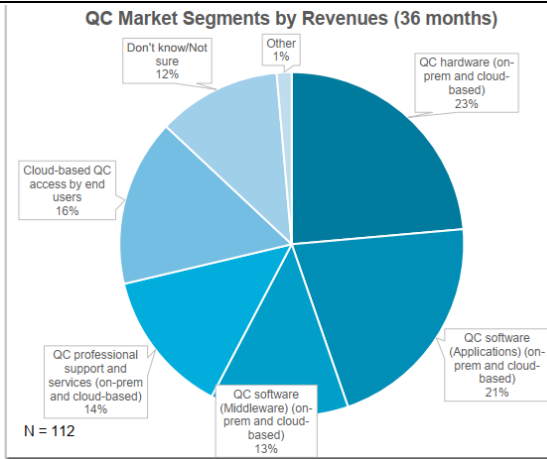
根据 Hyperion Research 最新展望，全球量子计算市场 2022 年市场规模 6.14 亿美元，预计到 2025 年达到 12.08 亿美元，CAGR 25%，其中机器学习市场占比 25%，被认为是最有潜力的应用市场之一。

图表 61 全球量子计算市场或在 2025 年达到 12 亿美金



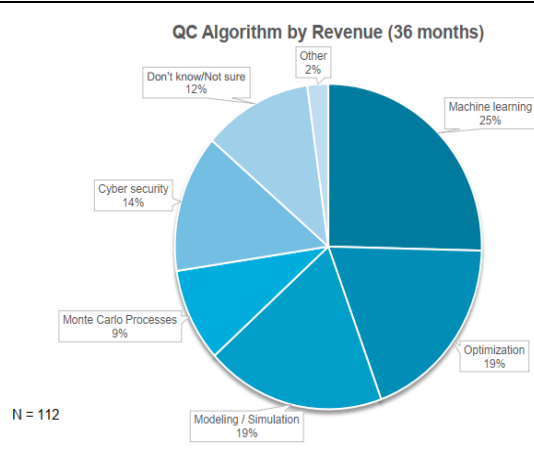
资料来源：Hyperion Research，华安证券研究所

图表 62 量子计算未来三年主要市场结构



资料来源: Hyperion Research, 华安证券研究所

图表 63 量子计算未来三年主要应用的算法领域



资料来源: Hyperion Research, 华安证券研究所

量子计算在 AI 的应用处于探索初期, 发现部分算法可能有优势, 但还没实现量子霸权。神经网络和机器学习系统的发展依赖于算力的进步以及庞大的训练数据量积累, 随着摩尔定律演进的放缓以及“内存墙”等经典计算机架构原因, 算力增长正在放缓, 而隐私问题&数据标注成本也阻碍大数据的无限度获取。目前, 产业界认为量子计算可能会成为机器学习发展的“神助攻”或开辟崭新的量子机器学习领域。

一是量子计算机本身的工作架构可被看做一个神经网络 (N 个 Qubit 可代表 2^N 的 N 次方个神经元), 麻省理工大学物理学家 Lloyd 估计, 60 个 Qubit 量子计算机可以编码的数据量就相当于人类一年生成的所有数据, 而神经网络运算相当于同时对这些“神经元”对矩阵运算。二是传统神经网络初始层导入输入的样本将被中间层生成不同组合形式的输入, 而采用量子比特编码的数据集数量和多样性都有可能扩大和丰富从而可能更好地训练模型。目前学术界探讨的可能展现出优势的人工智能算法包括决策问题、搜索问题、博弈理论、自然语言处理、贝叶斯网络、模式识别等, 但由于目前还无法成功解决模型的初始参数输入以及准确的结果测量手段等问题, 量子计算机在以上人工智能算法上还未展现出“量子霸权”优势。

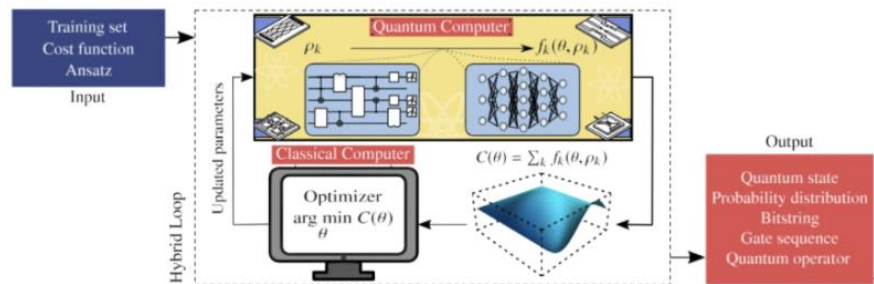
图表 64 学术界目前探讨的量子计算可能展现出优势的人工智能算法

量子人工智能算法	算法依据	量子计算可能的优势
决策问题	许多决策问题可以用决策树表示	基于哈密顿演化的量子算法能以指数级速度解决由一类决策树表示的决策问题
量子搜索	许多人工智能问题可以简化为搜索, 如计划、调度、定理证明和信息检索	Grover 算法表明量子计算机可以比传统计算机更快完成搜索任务
量子博弈理论	博弈论被广泛应用在多智能体系统和分布式人工智能中	量子博弈理论提供了新的工具解决人工智能的一些问题
自然语言处理	人工智能在自然语言语义分析中使用的数学结构与量子力学中使用的数学结构相似	例如通过量子叠加来表达歧义可能会大大加快自然语言的处理速度
量子贝叶斯网络	贝叶斯网络是对概率信息进行表示和推理的图形模型在人工智能得到广泛应用	量子系统本质是概率性的, 可以通过一些局部操作实现量子贝叶斯网络推理
模式识别	对经典物体的识别和辨别是关于量子态和量子运算识别的特例	通过对量子测量分辨的研究可以设计一些最优的协议以最少的查询来实现量子操作识别

资料来源: 自然杂志:《量子计算与人工智能》, 华安证券研究所整理

变分量子算法 (VQA) 有望成为助力量子机器学习落地利器。变分量子算法的实现需要量子计算机和经典计算机的配合，其主要想法是对问题的解进行编码的损失函数，对参数进行训练以最小化损失的 **ansatz**，这将通过量子计算机来有效估计损失（或其梯度）。这些信息被输入到一台经典计算机中，该计算机利用优化器能力解优化问题，一旦满足终止条件，VQA 将输出问题解决方案的估计值，输出形式可能包括量子态、门序列、比特串、概率分布等。VQA 在量子计算机上运行参数化量子电路，然后将参数优化外包给经典优化器，量子神经网络或成为 VQA 重要应用领域。

图表 65 变分量子算法示意图

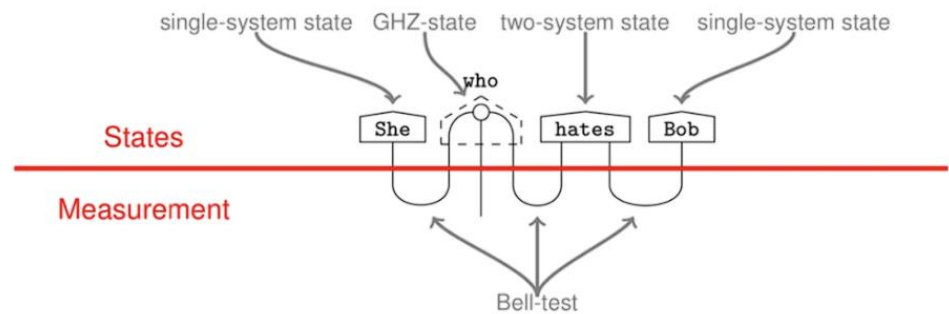


资料来源：光子盒，华安证券研究所

量子计算加速 NLP 成为最被看好的前沿应用领域。作为目前爆火的 AI 大模型，ChatGPT 是自然语言处理 (NLP) 的创新成果，结合了语言学、计算机和人工智能，以理解和模仿人类如何使用语言。量子计算在 NLP 大型复杂数据集处理中可能具有价值，Omdia 首席量子计算分析师 Sam Lucero 认为量子计算将在 NLP 中发挥作用，最终在 ChatGPT 和 AIGC 中发挥作用，该研究分支被称作 QNLP。

根据启科量子研究，QNLP 可能存在以下优势：1) NLP 的主要任务即相关搜索或任务分类的算法加速；2) 指数级的量子状态空间适用于更加复杂的语言结构；3) 运用密度矩阵的新型意义模型自然地模拟了诸如下义关系和语言歧义等语言现象；4) 可以大大提高训练效率，用更少的训练数据达到相同的能力水平。2020 年初，剑桥量子计算公司 (CQC) 宣布首次在量子计算机上执行自然语言处理测试并获得成功，他们通过将语法句子翻译成量子电路，然后在量子计算机上实现生成的程序执行问答，发现了在含噪声中等规模量子计算时代获得量子优势的途径。

图表 66 剑桥量子计算公司 (CQC) 首次在量子计算机执行自然语言处理测试获得成功



资料来源：CQC，华安证券研究所

量子计算目前主要用于科研实验，未来可用于对经典计算机的异构加速。目前的量子计算机仅能被证明在一些特定领域（如高斯玻色采样、随机量子线路采样）实现“量子霸权”优势，距离产生商用价值还很远。根据目前学术界的探索，量子计算机模拟电子/光子/原子等基本粒子运动具有先天优势，未来可能率先在生物制药、化工材料、光学仿真等领域商用落地，但目前由于量子退相干特性的存在以及量子态存储技术尚未实现，量子计算机需要和经典计算机配合（如作为某些特定任务的加速器）才能实现计算任务。目前量子计算机距离通用化可编程还要很大距离，需要解决量子比特的拓展性、连通性、保真度、电子穴等芯片设计、低温物理难题，其次更需要算法的进化，以解决量子计算机最难的“可纠错”问题。业内普遍估计符合一定学术条件时，当 3000 量子比特左右的量子计算机出现时，可能代替经典计算机做一些自然演化实验，而实现等效逻辑比特后量子计算的发展可能会加速。

我们认为，量子计算目前处于“从 0 到 1”的发展阶段，目前产业链主要价值在于量子计算原型机的集成、相关器件供应链，未来将以云平台形式对特定领域的科研/模拟形成商业价值。我们看好量子算法在 AI 加速的潜力，未来可能出现针对特定领域的量子 AI 解决方案公司。

5.5 相关标的

1) 海光信息：国产 CPU 与 GPGPU 重要参与者

海光信息是一家以 CPU 和 DCU 产品为主的芯片设计厂商，公司多款产品性能达到了国际同类型主流高端处理器的水平。CPU 类产品兼容国际主流 x86 处理器架构和技术路线以及国际上主流操作系统和应用软件，软硬件生态完善，可靠性和安全性较高，得到终端客户的认可，已经广泛应用于运营商、金融、互联网、教育等重要行业及领域。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、商业计算等应用领域。

在 ChatGPT 等应用场景下的所指的 GPU 即为 GPGPU，去掉 GPU 为了图形处理而设计的加速硬件单元，保留了 GPU 的 SIMT 架构和通用计算单元。所以对于 ChatGPT 场景下的 AI 训练、矩阵运算等通用计算类型的任务仍然保留了 GPU 的优势，即高效搬运，运算，重复性的有海量数据的任务。

图表 67 海光 DCU 基本架构



资料来源：海光信息招股说明书，华安证券研究所

公司的 DCU 产品（8000 深算系列产品）归属的 GPGPU 范畴，可广泛用于包括 ChatGPT 所处的人工智能及大数据等领域。DCU 系列深算一号已于 2021 年实现商业化应用，深算二号处于研发阶段。海光 8100 能够充分挖掘应用的并行性，发挥其大规模并行计算的能力，并正快速开发高能效的应用程序。海光 8100 采用先进的 FinFET 工艺，典型应用场景下，其性能指标可以达到国际同类型高端产品的同期水平。

图表 68 海光 8100 系列产品主要规格和特点

海光 8100	
产品图片	
典型功耗	260-350W
典型运算类型	双精度、单精度、半精度浮点数据和各种常见整型数据
计算	① 60-64 个计算单元（最多 4096 个计算核心） ② 支持 FP64、FP32、FP16、INT8、INT4
内存	① 4 个 HBM2 内存通道 ② 最高内存带宽为 1TB/s ③ 最大内存容量为 32GB
I/O	① 16 Lane PCIe Gen4 ② DCU 芯片之间高速互连

资料来源：海光信息招股说明书，华安证券研究所


基于较好的性能表现，海光 DCU 在未来可以广泛应用于大数据处理、人工智能、商业计算等高算力需求领域。现阶段海光信息已有相关互联网行业的终端客户，互联网领域的市场空间十分广阔。

2) 复旦微电：FPGA 受益算力增长

复旦微电的 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，公司在国内较早的推出了亿门级 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 FPGA 和 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。

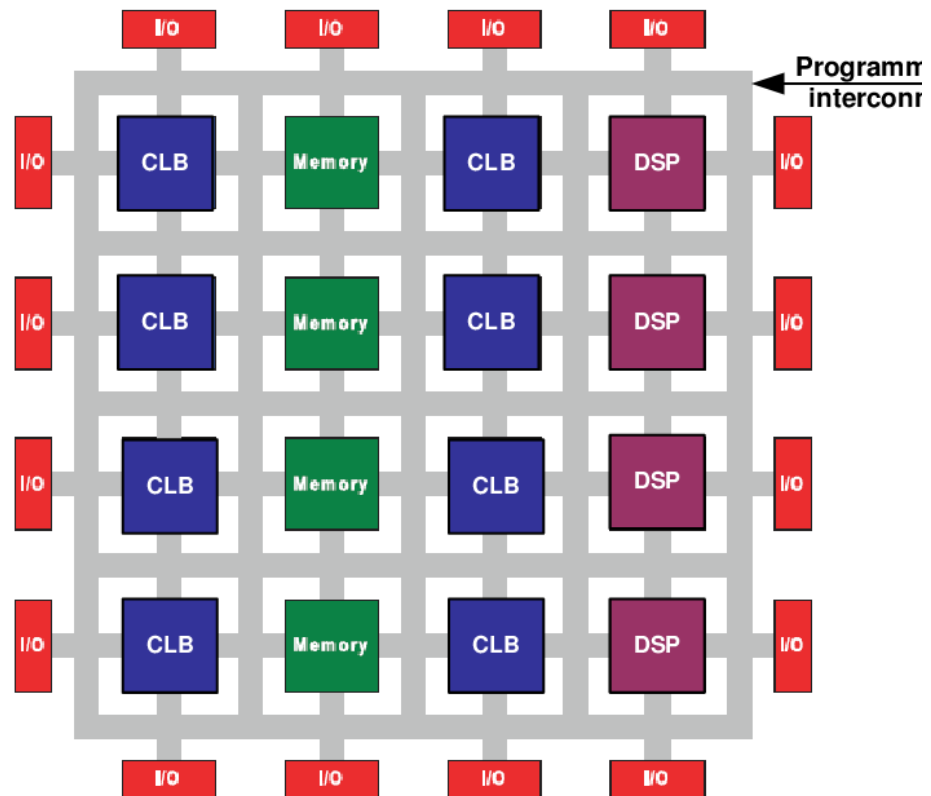
图表 69 复旦微电产品主要规格和特点

产品类型	产品介绍	应用领域	产品或终端样图
千万门级 FPGA 芯片	采用 65nm CMOS 工艺，是一系列高性能、高性价比 SRAM 型 FPGA 产品	适用于网络通信、信息安全、工业控制、高可靠等高性能、大规模应用	
亿门级 FPGA 芯片	采用 28nm CMOS 工艺，是一系列高性能、大规模的 SRAM 型 FPGA 产品	适用于 5G 通信、人工智能、数据中心、高可靠等高性能、大带宽、超大规模应用	
嵌入式可编程器件 PSoC	采用 28nm CMOS 工艺，是一系列嵌入式可编程片上系统产品	适用于视频、工控、安全、AI、高可靠等应用	

资料来源：复旦微电招股说明书，华安证券研究所

FPGA 在人工智能加速卡领域应用广泛。FPGA 通过与 CPU 搭配，CPU 的部分数据运算转移给 FPGA，最终 FPGA 起到加速作用。无论是赛灵思还是英特尔，其 FPGA 在数据中心运算方面的产品形态均为加速卡，在服务器中与 CPU 进行配合。人工智能领域属于加速计算的一个分支，如阿里云、腾讯云之类的数据中心均对加速卡存在大量需求。根据人工智能的不同应用领域，可将各种算力需求和控制逻辑用最合适的资源组合实现，在这一过程中，FPGA 在其中起到了关键作用。总体而言，FPGA 具备快速、低功耗、灵活和高效的优点。硬件可编程的特性使得 FPGA 在 AI 训练中既能提供充足的算力，又具有灵活性，可以重新编程以适应不同任务的需要。

图表 70 FPGA 结构特点



资料来源：ICTlearn，华安证券研究所

复旦微电同时逐步完善 FPGA 的开发软件工具。公司推出的可用于完整可编程器件开发流程的工具软件 Procise™ 可以支持公司全系列可编程器件，实现了 FPGA 配套 EDA 软件的多项核心关键技术的突破。相关技术包含 FPGA 架构技术、可编程器件编译器技术、多协议超高速串行收发器技术、异构智算架构技术、高可靠可编程器件技术、超大规模可编程器件配套全流程 EDA 技术等关键技术，复旦微电在 FPGA 和 PSoC 领域形成了明显的技术集群优势。

2022 年前三季度复旦 FPGA 及其他芯片收入约为 6.00 亿元，维持较高增长。公司新一代十亿门级 FPGA 产品研发进展顺利，其各方面性能对比上一代产品均有较大幅度的提升。公司新一代配置有 APU、GPU、VPU、eFPGA、AI 引擎的异构智能 PSoC 产品也取得了阶段性的研发成果，将进一步丰富公司的可编程产品系列谱系，不断满足各应用领域客户的需求。

3) 中际旭创：全球光模块龙头迎来 800G 时代

中际旭创是全球高速数通光模块龙头，在 Top5 云计算公司光模块供应商中占据最大份额，根据 LightCounting 统计，公司在 2021 全球光模块市场位列第一。全球云巨头数据中心网络部署正处于 400G 批量部署、800G 小批量上量阶段，公司 800G 全系列产品的实现供货，800G 时代有望维持龙头地位，并将享受 400G、800G 高端产品收入占比增加带来的综合毛利率提升。此外，公司布局硅光、CPO 等前沿技术，硅光/EML 两种平台 400G 光模块均通过客户验证，未来有望把持产业链上游核心环节，实现供应链稳定。在电信市场，公司战略进军相干高端光模块，已在国内主要设备商实现供货，有望开启第二增长曲线。

4) 联特科技：欧美中低速 WDM 主流供应商，数通光模块“黑马”

联特科技在海外市场主打差异化竞争优势，经历 10 年积累，在 WDM 细分市场占据了优势地位，主要客户包括爱立信、诺基亚、Prolabs、Adtran 等。随着欧美疫后通信新基建投资增加，原有城域网持续速率升级，公司在客户份额有望继续提升。此外，海外发达国家 RAN 市场，爱立信、诺基亚两家占据主导地位，公司作为爱立信前传光模块主要供应商之一收入持续增长。数通方面，近几年营收快速增长，2021 年营收达 3.3 亿元，同比增长 33%，在整体收入中占比近半。公司已成为思科、Arista 等海外主流设备商合格供应商，100G/400G 高速光模块产品持续上量，并有望在 2023 突破北美 Top4 云厂商客户。公司前瞻布局 800G 和 CPO 产品，目前已推出基于 EML、SiP、TFLN 全系列 800G 产品，EML 预计 2023 年在客户端完成全面认证。

5) 天孚通信：光通信精益制造代表，光引擎、激光雷达成长速度快

天孚通信定位于光器件整体解决方案提供商 OMS，目前拥有十三大产品线、八大解决方案，几乎涵盖光模块生产所需的所有无源光器件、各种有源封装形态等。公司 2022 年业绩快报预计实现收入 12.06 亿元，同比增长 16.8%，五年 CAGR 29.1%，归母业绩 29.2%。公司将持续优化无源产品线结构，拓展新客户的同时提供一站式解决方案增加单客户价值量，未来无源产品有望保持高于行业的平均增速和毛利率；有源方面，公司战略聚焦高速数通领域，与大客户合作的硅光引擎快速上量，将受益于 800G、1.6T 硅光、CPO 等新技术渗透率提升。此外，公司在激光雷达领域秉持精益制造理念，提供滤波片、透镜、棱镜等无源光器件以及模组化解决方案代工，目前在国内主流新势力车型已取得定点，未来两年将持续放量。

6) 国盾量子：量子计算机已实现原型机搭建

公司核心技术来源于中科大产业化平台，是全球量子通信 QKD 设备龙头企业。公司在国内量子通信骨干网一期建设中占据了主要的设备份额，其骨干网/城域网 QKD 编码产品、量子卫星地面站产品、信道与密钥组网交换产品已批量部署。我们认为，随着国内经济的全面复苏，量子骨干网二期有望启动，同时公司在电力、电信、金融等领域与大型央企签订战略合作协议，随着各行业对保密通信的重视，有望开启更加广阔的行业市场。公司在量子计算领域主要提供超导量子计算低温线缆组件、约瑟夫森阻抗渐变参量放大器、ez-Q Engine 超导量子计算操控系统等子系统和器件，其超导量子计算操控系统成功主力“祖冲之号”实现量子优越性展示。公司日前完成了“祖冲之二号”同等规模超导量子计算机原型机搭建，是国内目前唯一具有量子计算机整机集成能力的上市公司，未来将通过平台的形式率先提供服务。

7) 浪潮信息：国内领先 AI 服务器厂商

公司是全球领先的新型 IT 基础架构产品、方案及服务提供商，以“智慧计算”为战略，通过“硬件重构+ 软件定义”的算力产品和解决方案、构建开放融合的计算生态，为客户构建满足多样化场景的智慧计算平台，全面推动人工智能、大数据、云计算、物联网的广泛应用和对传统产业的数字化变革与重塑。当前业务包括传统服务器、AI 服务器及存

储产品。

根据 IDC 最新数据，浪潮信息的服务器产品 2021 年全年位居全球前二，持续以 30%+ 的市占率领跑中国市场；

AI 服务器方面，根据 IDC 数据，2021 年全年，中国 AI 服务器市场规模达 350.3 亿元，同比增长 68.6%。从厂商维度看，浪潮信息、宁畅、新华三、华为、安擎位居前五，占据了 82.6% 的市场份额。其中，浪潮 AI 服务器市场占有率达 52.4%。2021 年，公司发布首款智算中心调度系统 AIStation，拥有性能最强的液冷 AI 服务器 NF5488LA5，发布 2457 亿参数的 AI 巨量模型“源 1.0”。

6 从元宇宙到大制造，大模型应用领域不断丰富

6.1 “生成式 AI”在互联网及元宇宙应用

“生成式 AI (generative AI)”在互联网及元宇宙领域市场化空间较为广阔。基于现行的 NLP 算法发展程度及数据集规模，在不久的将来，生成式 AI 有较大可能在“智能客服”和“搜索引擎”进行增值，并有希望以“插件”的形式赋能现有的“生产力工具链(工程软件/音视频制作工具等)”。

在为客服领域增值的过程中，有希望在人工客服的全链路中（包括问题识别、潜在解决方案交付、反馈优化、和风险识别等）实现对人工客服的替代。可实现替代的领域包括电商零售、医疗健康、金融服务、和电信等领域的客服环节。

在为搜索引擎增值的过程中，生成式 AI 在中短期之内，并不会颠覆现今的搜索引擎地位，但是会逐步成为搜索引擎进步的主要技术驱动因素，相比于传统搜索引擎技术，基于“生成式 AI”技术的搜索反馈结果，将基于其“自学习”能力适应不同用户的需求变化，其最终结果交付将实现从“信息的匹配和中转”到“内容创造和问题解决”的跃迁。

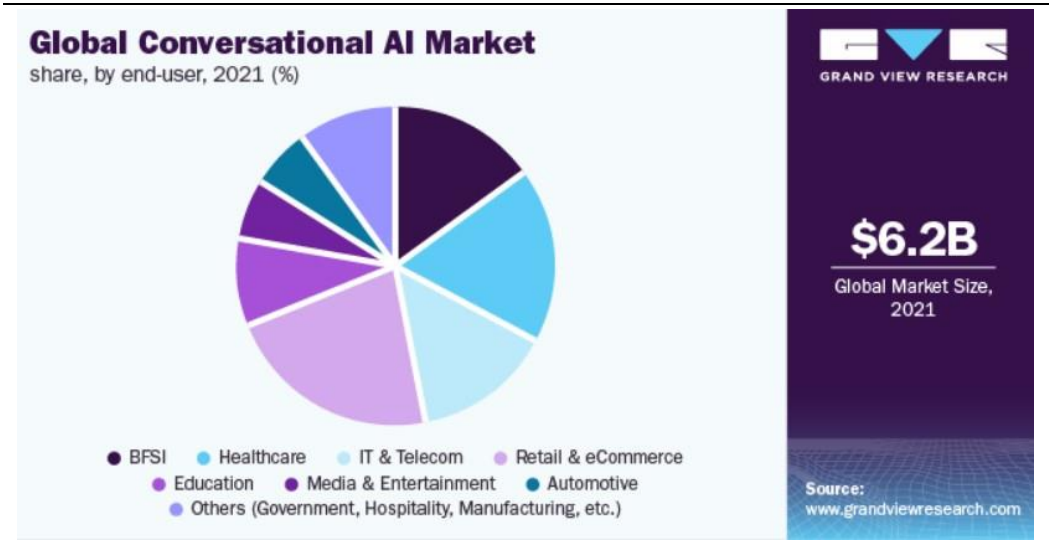
在赋能“元宇宙工具链”的过程中，生成式 AI 未来有希望通过“插件(Add-in)”的形式，作为工具链的一部分，结合使用者的请求生成工程图纸、代码、图形等，提高工具的易用度，降低人的工具执行成本。这些工具涉及到的领域包括但不限于，工程设计（包括 Solidworks, CATIA, 和 AUTOCAD 等），代码设计/运算(Matlab, TensorFlow 等)，音视频制作(Adobe Photoshop 等)，游戏/图形设计(Unity, UE4 等)，操作系统(IOS, Linux, 智能车操作系统)，和元宇宙仿真(Omniverse)。

1) “生成式 AI”在智能客服领域的潜在应用

“可对话 AI (Conversational AI)”是“生成式 AI”在智能客服领域的直接应用。根据 Grand View Research，全球“可对话 AI(Conversational AI)”2021 年市场空间为 62 亿美元，其中，BFSI(银行保险等金融服务)，医疗，零售和电商，和电信领域的市场空间为 42.5 亿美元左右，占 68.5%左右；根据 Grand View Research 预测，2030 年“可对话 AI”的市场空间将达到约 413.9 亿美元，对应 2022 至 2030 年复合增长率(CAGR)为 23.6%。该领域的核心竞争者包括：谷歌，微软，亚马逊，IBM，甲骨文，和 SAP 等。市场增长的主要驱动因素包括各领域对于应用 AI 技术替代人力这一需求的提升，和持续下降的 AI 对话程序的研发成本。

根据 Reuters 报道，目前智能对话机器人领域的佼佼者 Open AI 公司（核心产品是 ChatGPT），2024 年的收入有望达到 10 亿美元，目前的估值约为 200 亿美元。GPT(Generative Pre-Trained Transformer)系列模型从 2018 年至今开始发布，经历了 3.5 代。其中，2020 年发布的 GPT-3 参数量达到了 1750 亿以上，预训练数据量达到 45TB。而 ChatGPT 相比 GPT-3 有了更多的强化和提高。ChatGPT 的开发者使用 RLHF(Reinforcement Learning from Human Feedback)ChatGPT 具备拟人化的对话能力，包括支持连续沟通对话，理解上下文语义，并且承认自己的不足等。而这项技术的成熟化，将直接带来人工客服领域的变革，基于“类 ChatGPT”模型的可对话 AI 将实现对客服人力的大规模替代。

图表 71 全球“对话 AI(Conversational AI)”市场空间



资料来源：Grand View Research，华安证券研究所

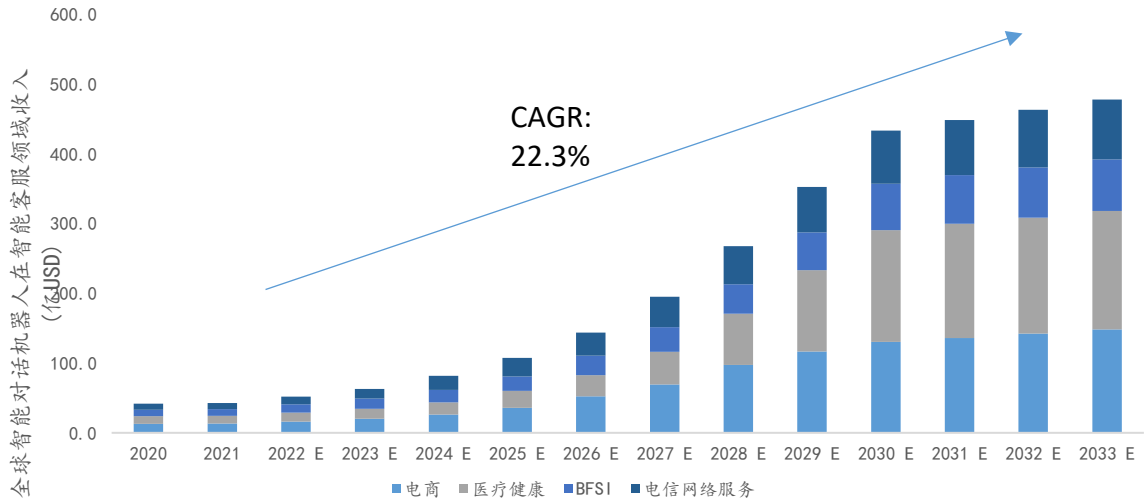
我们判断，从现在起至未来几年，作为“生成式 AI”重要应用的“对话 AI(Conversational AI)”的商业化模式中较为清晰且可行的，是在各个领域对于人工客服的替代。对人工客服进行替代的假设是基于以下 3 点考虑：1、全球主要发达经济体人口增长乏力，劳动力数量减少，用工成本攀升，有强烈的使用 AI 对话机器人替代人工客服的需求；2、“智能对话机器人”相比“人工客服”可以创造更多价值，即，机器人可以完成更多人工客服无法胜任的任务，并且工作效率高，解决问题出错率较低。3、“智能对话机器人”研发和部署成本的有望随着算力提升或者异构运算的发展而逐步下降，同时该机器人的实践经验的可复制性不断提高。由于电商、医疗健康、BFSI、和电信网络服务的客服服务中产生的问题及解答，较为结构化并依赖劳动力密集产出（其中，医疗健康领域不包括医生看诊环节，仅包括挂号预约、初步咨询、取药、和护理服务沟通等专业性较低环节），所以这 4 个应用领域有望成为“可对话 AI”可以进行“人力资源替代”的主要领域。我们从这 4 个主要领域入手，基于对未来驱动因素的假设，进行了市场空间测算。根据 Grand View Research，2021 年这 4 个主要领域市场空间约为 42.5 亿美元左右；基于此作为起点，经过我们的测算，**2033 年这 4 大主要领域的市场空间可以达到 478 亿美元。**

图表 72 “智能对话机器人”在各领域全球市场空间（亿 USD）

(单位: 亿 USD)	2020	2021	2022 E	2023 E	2024 E	2025 E	2026 E	2027 E	2028 E	2029 E	2030 E	2031 E	2032 E	2033 E
电商	12.8	13.4	16.1	20.1	26.0	36.0	52.2	69.3	97.3	116.6	130.3	136.1	142.2	148.4
医疗健康	10.9	11.0	12.7	14.3	17.6	24.1	30.7	47.1	73.5	116.7	160.6	163.6	166.6	169.7
BFSI	9.3	9.4	12.2	15.0	17.9	20.8	27.9	35.2	42.6	54.3	66.3	69.7	71.8	73.9
电信网络服务	8.6	8.7	11.2	13.8	20.1	26.5	33.1	43.6	54.3	65.1	76.1	79.4	82.7	86.0
总计	41.6	42.5	52.1	63.1	81.6	107.4	143.9	195.1	267.6	352.6	433.3	448.8	463.3	478.0

资料来源：世界银行, Grand view research, Statista, 华安证券研究所测算

图表 73 全球“智能对话机器人”智能客服领域市场空间(亿 USD)



资料来源：世界银行，Grand View Research，华安证券研究所测算

图表 74 智能对话机器人-电商客服领域全球市场空间测算(亿 USD)

智能对话机器人-电商领域全球市场空间测算

核心假设:

年平均网购次数 (次/人·年)	18
每年与客服沟通次数/每年网购次数	80%
客服效率(次/(人·年))	18000
客服年薪(USD/年)	9000

	2020	2021	2022 E	2023 E	2024 E	2025 E	2026 E	2027 E	2033 E
全球电商购物人数 (亿人)	20.5	21.4	22.3	23.2	24.1	25.0	25.9	26.7	31.7
yoy	N/A	4.4%	4.2%	4.0%	3.80%	3.70%	3.60%	3.30%	2.80%
客服服务人次需求 (亿次)	295.2	308.2	321.1	333.9	346.6	359.4	372.3	384.6	456.6
客服人员需求 (亿人)	0.016	0.017	0.018	0.019	0.019	0.020	0.021	0.021	0.025
yoy	N/A	4.4%	4.2%	4.0%	3.8%	3.7%	3.6%	3.3%	2.8%
全球电商客服支出(亿 USD)	147.6	154.1	160.5	166.9	173.3	179.7	186.2	192.3	228.3
全球“对话 AI” 电商订阅收入 (亿 USD)	12.81	13.43	16.08	20.06	26.02	35.97	52.16	69.27	148.43
“对话 AI”订阅收入/电商客服支出	8%	8%	10%	12%	15%	20%	28%	36%	65%

资料来源：世界银行，Grand view research, Statista，华安证券研究所测算

2) “生成式 AI”在搜索引擎领域的潜在应用

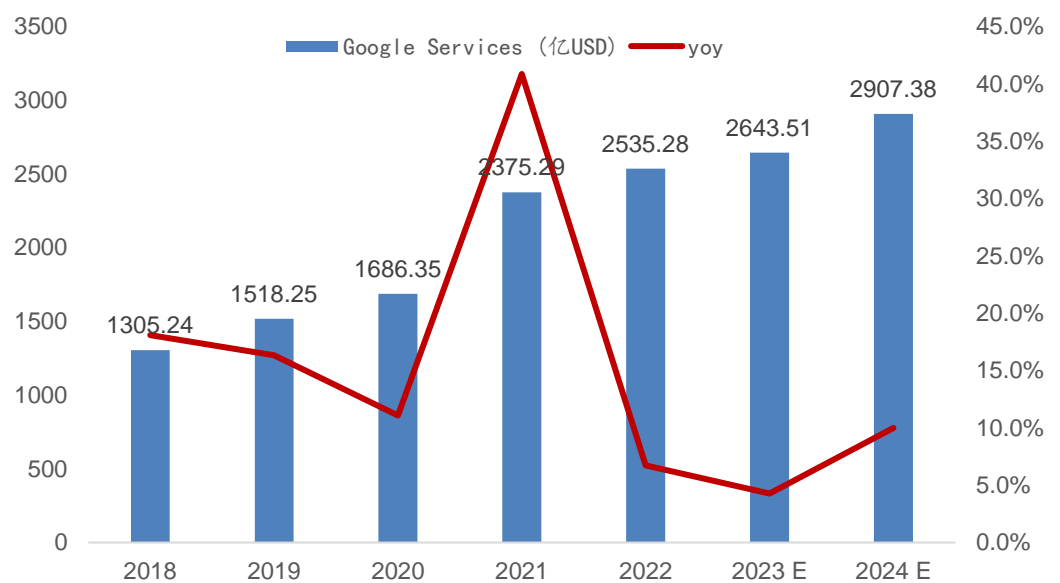
在互联网搜索领域，目前谷歌占据绝对的领军地位。谷歌主要的业务板块是谷歌服务(Google Services)，2022 年谷歌服务收入 2535.28 亿美元。谷歌服务依赖谷歌在搜索引擎领域长期积累的技术和商业化优势，包括搜索，Youtube，google play 智能手机平台，广告，浏览器，邮箱，云盘等。

在考虑未来“生成式 AI”在搜索领域可以实现的收入时，考虑范围应该不局限于“搜索引擎”本身，还应该包括基于“搜索引擎”技术所拓展出的外延部分，比如 Youtube，Google

Map 和 Google Play 等；这些搜索引擎外延部分的商业化表现，本质是基于平台是否可以基于用户的搜索请求或者使用习惯，将产品及服务较精准的分发给用户，然后用户并为此付费。所以，谷歌服务部分的收入天花板，可以用来作为“标尺”去衡量未来智能对话机器人在“搜索领域”的潜在收入中最易于理解的部分。

“生成式 AI(Generative AI)”未来有较大技术进步空间和较难估量的商业化潜力。基于以下 3 点原因，在中长期，“生成式 AI”的市场天花板应该不限于当前搜索引擎所达到的成就：1) 它为用户增值的目标从“信息的搜索和呈现”跨越到了“独立解决问题”；2) 它在“解决问题”的过程中，存在较大的商业化空间，可以将定制化的“产品”或“服务”分发给终端用户；3) 它具备一定自学习能力，可以承认自身错误并进化，同时基于不同用户的使用习惯，有希望为用户提供定制化的体验。

图表 75 谷歌服务(Google Services)收入 (亿 USD)



资料来源：Bloomberg，华安证券研究所

但是生成式 AI 在搜索领域实现替代仍然面临诸多挑战，生成式 AI 技术需要先达到一定程度的“规模优势(包括预训练数据集规模,用户反馈量)”之后才有机会威胁到搜索引擎的生存地位。

在科技厂商的供给侧，这种“规模优势”的达成将依赖算力开销边际成本的下降。在大模型训练的背后是大厂持续的对于计算中心基础设施的扩张；目前，GPT3 最大规模的模型需要使用到 175 billions 的参数量，在训练的过程中需要使用到数万片以上英伟达 A100 芯片，每月大模型训练成本在数百万美元以上量级。总之，生成式 AI 可以大规模应用于搜索引擎中的前提是较强的算力支撑，而算力支撑取决于芯片技术(材料、设计、工艺制程)及“异构计算技术”的发展进程(包括计算开销的下降和计算交付结果精确程度的提高)。

在用户的需求侧，用户习惯的颠覆将是最大的限制。基于当前技术迭代路径的“搜索引擎(包括 Google, Baidu, Bing 等)”已经教育了市场将近 25 年以上的时间，颠覆用户习惯需要极大的动能，这种动能一定是基于“替代方案”的效率要比“现存方案”优越数倍以上。

图表 76 GPT3 各类模型的训练的参数量

Model Name	n_{params}	n_{layers}	d_{model}	n_{heads}	d_{head}	Batch Size	Learning Rate
GPT-3 Small	125M	12	768	12	64	0.5M	6.0×10^{-4}
GPT-3 Medium	350M	24	1024	16	64	0.5M	3.0×10^{-4}
GPT-3 Large	760M	24	1536	16	96	0.5M	2.5×10^{-4}
GPT-3 XL	1.3B	24	2048	24	128	1M	2.0×10^{-4}
GPT-3 2.7B	2.7B	32	2560	32	80	1M	1.6×10^{-4}
GPT-3 6.7B	6.7B	32	4096	32	128	2M	1.2×10^{-4}
GPT-3 13B	13.0B	40	5140	40	128	2M	1.0×10^{-4}
GPT-3 175B or "GPT-3"	175.0B	96	12288	96	128	3.2M	0.6×10^{-4}

Fig-2: Details of variants of the GPT-3 model

资料来源: Sigmoid, 华安证券研究所整理

图表 77 GPT 模型训练成本

Model	Parameters	Tokens to Train to Chinchilla Point (B)	Cerebras Model Studio CS-2 Day to Train	Cerebras Model Studio Price to Train
GPT3-XL	1.3	26	0.4	\$2,500
GPT-J	6	120	8	\$45,000
GPT-3 6.7B	6.7	134	11	\$40,000
T-5 11B	11	34*	9	\$60,000
GPT-3 13B	13	260	39	\$150,000
GPT NeoX	20	400	47	\$525,000
GPT 70B	70	1,400	85	\$2,500,000
GPT 175B	175	3,500	Contact For Quote	Contact For Quote

* - T5 tokens to train from the original T5 paper. Chinchilla scaling laws not applicable.

资料来源: Nextplatform, 华安证券研究所整理

图表 78 “生成式 AI”对搜索引擎的影响

输入	过程		输出	商业化路径
文字	内容分发者	理解搜索语义	交付搜索结果	广告
图片		内容分类	广告推送/排名	
语音		内容排名		
		内容匹配		



输入	过程		输出	商业化路径
复杂表述的文字段落	内容创造者	归纳/演绎	AI生成图片, 视频, 声音等	打通生产力工具链, 订阅付费
代码/伪代码		自学习进化	AI生成文字	定制化广告
视频/复杂声音信号		交付质量控制	AI生成工程设计, 代码	
电路图/机械图纸				

资料来源: Marketing AI Institute, 华安证券研究所整理

图表 79 生成式 AI 与搜索引擎结合面临的挑战

生成式 AI 与搜索引擎结合面临的挑战			
供给侧		需求侧	
数据集	训练过程	用户认知	终端用户增值
- 样本规模	- 算力开销	- 搜索习惯固化	- 增值显著性
- 样本有效性	- 算法效率	- 较高学习成本	
- 样本获取的合法性			

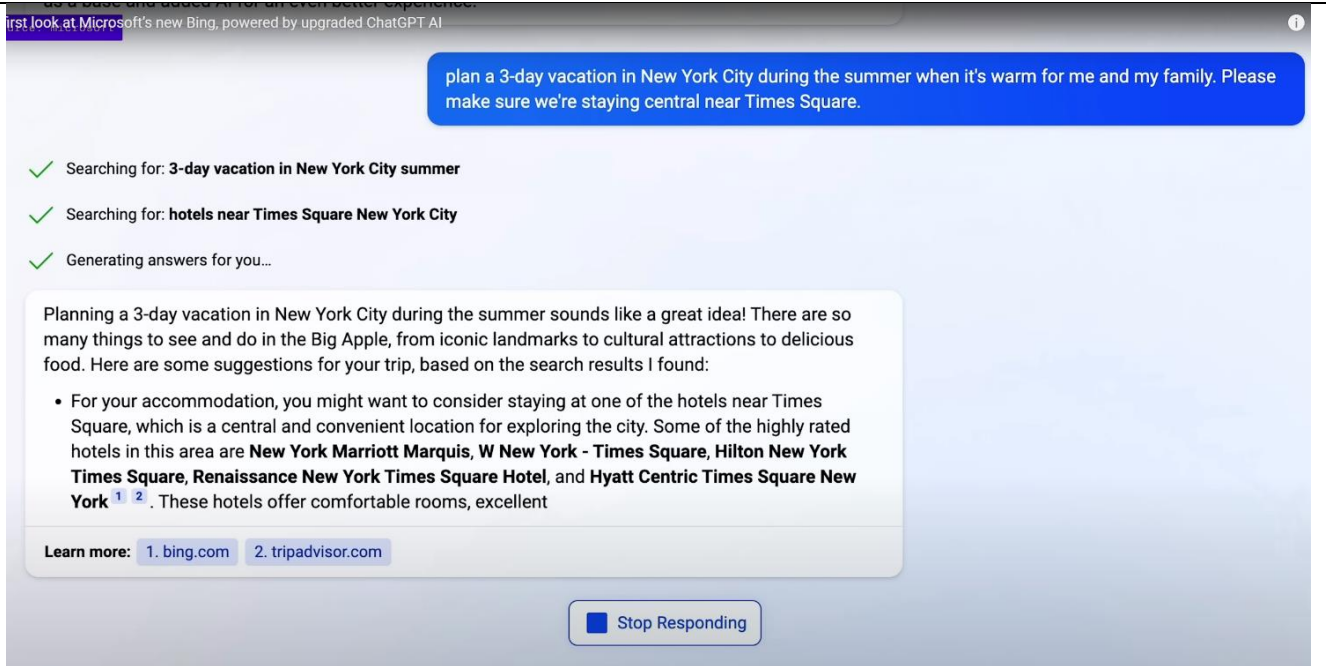
资料来源: Open AI, 华安证券研究所整理

2023年2月初, 微软公布新 Bing (New Bing)搜索引擎, 该版本搜索引擎集成了 ChatGPT 技术, 公布不到 48 小时, 申请用户量已经过百万。用户需要排队注册申请并等待获得使用新 Bing 测试版的资格。根据微软官方的解释, 新 Bing 可以作为类似研究助理(research assistant), 个人计划员(personal planner), 和创意合作伙伴(creative partner)的角色为用户创造价值。和常规的搜索引擎相比, 新 Bing 的搜索结果将不再是简单的提供给用户一个链接列表, 而是给用户一个概括的答案, 解决用户的具体问题, 并且提供可靠的信息来源。用户可以按照思考和沟通的方式与新 Bing 对话。同时, 新

Bing 也可以作为创意工具，帮助用户写诗，写故事或者写分享关于项目的想法。根据 The Verge 发布的新 Bing 测试使用体验，用户可以直接向 Bing 提问如何帮助自己规划一个在纽约市 3 日的旅行，同时确保自己可以待在纽约时代广场附近；而新 Bing 的回复可以做到将酒店的选择按照一定优先级为用户规划出来。总之，基于 ChatGPT 技术的新 Bing，有如下拟人化的对话特点：

- 可以解决具体问题并给出凝练的回答
- 可以接受用户的追问，得到定制化的回应
- 可以承认自身的错误并进行一定程度的调整

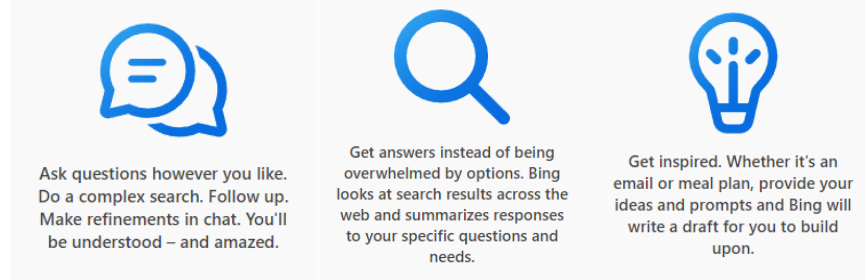
图表 80 “生成式 AI”的应用 – 新 Bing (New Bing) 搜索案例展示



资料来源：The Verge，华安证券研究所整理

图表 81 “生成式 AI”的应用 – 新 Bing (New Bing) 可对话的领域举例



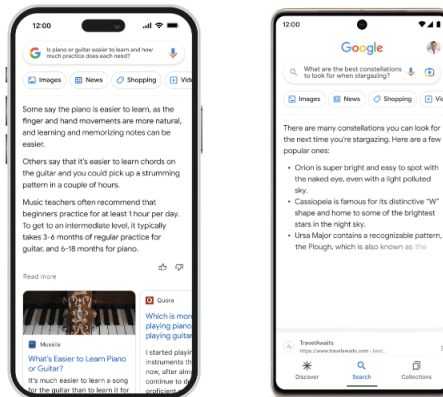


资料来源：Bing，华安证券研究所

2023年2月初，谷歌也发布了谷歌巴德(Google Bard)生成式可对话AI。使用了谷歌的LaMDA(Language Model for Dialogue Applications)模型。但是谷歌巴德发布之后，并没有达到用户的期待。之后，谷歌母公司Alphabet主席John Hennessy称，谷歌之前在犹豫是否要将Bard应用于产品中，因为Bard还没有真的准备好。

同时，百度也与同一时期官宣了“文心一言(ERNIE Bot)”即将于2023年3月完成内测，届时将向公众开放。之后，上海报业集团，36氪，广州日报，爱奇艺，度小满，携程等媒体及互联网产品和平台均宣布接入文心一言。文心一言采用文心大模型，文心大模型的API包括ERNIE 3.0文本理解与创作，ERNIE-ViLg文生图和PLATO开放域对话服务。

图表 82 “生成式 AI”的应用 - 谷歌巴德



资料来源：The Verge，华安证券研究所整理



图表 83 百度文心全景图、历程和架构

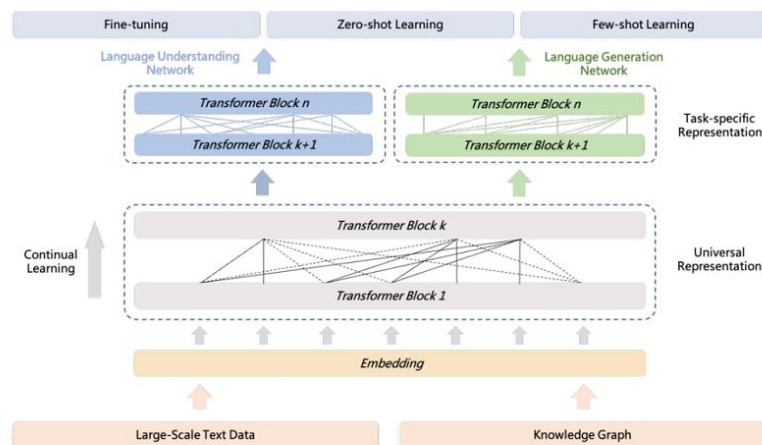


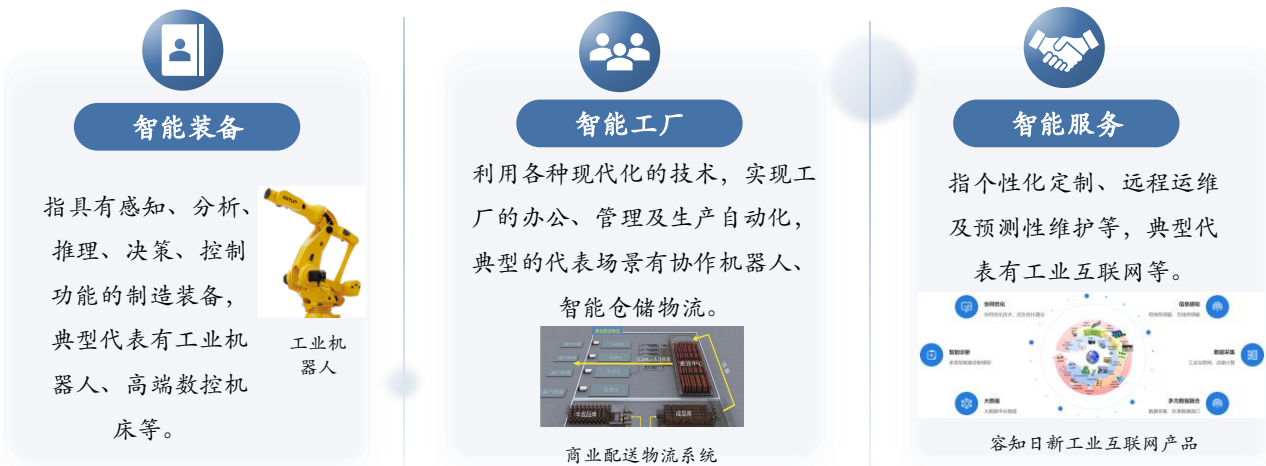
Figure 1: The framework of ERNIE 3.0.

资料来源：百度，华安证券研究所

6.2 AI 赋能制造业转型升级，智能制造浪潮兴起

物联网、大数据、云计算等技术日益成熟，人工智能已成为新时代的风口浪尖，“人工智能+”也代表一种新的社会形态。智能制造是实现制造强国的主攻方向，更是提升制造业竞争力的核心技术。随着智能制造的浪潮兴起，人工智能技术已贯穿于制造业设计、生产、管理等诸多环节。从必要性看，在劳动力及土地成本双升的背景下，制造业面临着利润低、市场变化迅速等压力，而人工智能的应用不仅可以帮助企业提升智能化运营水平，实现降本增效，还可以通过与其他新兴技术的融合，推动制造业模式升级及价值链重构。从实际应用来看，人工智能在制造业的应用可分为三方面：1) 智能装备：指具有感知、分析、推理、决策、控制功能的制造装备，典型代表有工业机器人、协作机器人、数控机床等；2) 智能工厂：利用各种现代化的技术，实现工厂的办公、管理及生产自动化，典型的代表场景有协作机器人、智能仓储物流系统等；3) 智能服务：指个性化定制、远程运维及预测性维护等，典型代表有工业互联网等。

图表 84 人工智能在制造业的应用



从必要性看，在劳动力及土地成本双升的背景下，制造业面临着利润低、市场变化迅速等压力，而人工智能的应用不仅可以帮助企业提升智能化运营水平，实现降本增效，还可以通过与其他新兴技术的融合，推动制造业模式升级及价值链重构。

资料来源：华安证券研究所整理

1) 智能装备产业百花齐放，机器人与高端数控机床空间广阔

➢ 工业机器人

工业机器人种类众多，贯穿工业生产的诸多工艺过程。工业机器人是广泛用于工业领域的多关节机械手或多自由度的机器装置，具有一定的自动性，可依靠自身的动力能源和控制能力实现各种工业加工制造功能。从构成来看，其主要分为三大模块：传感模块、控制模块和机械模块。其中传感模块负责感知内部和外部的信息，控制模块控制机器人完成各种活动，机械模块接受控制指令实现各种动作。从种类来看，工业机器人可分为搬运作业/上下料机器人、焊接机器人、喷涂机器人、加工机器人、装备机器人、洁净机器人及其他，贯穿于工业生产过程中从材料和零部件入库到生产再到最终产品出库的诸多环节。根据亿欧智库整理数据，2021年搬运机器人和焊接机器人占比最高，分别达55%和25%。从功能上看，较人工和传统机器而言，不论哪一类工业机器人，在实际生产中都具有成本、效率、安全等众多优势。因此，我们认为，随着核心零部件和核心技术的快速发展，工业机器人市场有望迎来快速发展。

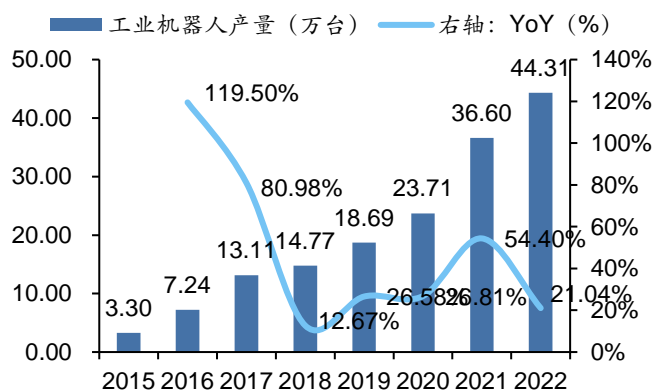
图表 85 工业机器人应用汇总

	搬运	焊接	装配	加工	喷涂	洁净	其他
产品占比 (2021)	55%	25%	10%	5%	3%	2%	
功能	运输、搬运、码垛、机床上下料等	点焊、弧焊等	装配	切割、磨削、抛光等	喷涂	在洁净室作业	/
应用行业	对仓储物流有需求的作业	汽车制造、工业机械、通用机械、金属结构、航天航空等	汽车制造、电器制造、小型电机、计算机、玩具等	汽车制造、金属加工、家具家居、工程机械等	汽车制造、航空航天、铝合金型材和板材、家具家居等	化学制药、医疗器械、精密机械等	/
市场特征	1、技术含量不高，但需求旺盛，尤其是大负载的搬运； 2、AGV、ACR等移动机器人发展迅速。	1、主要分为电焊和弧焊两大类，电焊主要用于汽车制造，弧焊的应用范围更广； 2、焊接速度和焊机稳定性是主要指标。	1、装配自动化水平低； 2、装配机器人多为轻量型； 3、对生产纲领、生产品质、装配数量等有一定要求。	1、以切割机器人为主； 2、视觉识别及跟踪技术是核心。	1、工艺不复杂，难度在于流量控制、喷涂线； 2、由于喷涂均匀度和速度的限制，现有机器人适用的工件不宜过大，形状也不宜太复杂。	需求很少，一般只有医药行业有需求。	/
适用工业机器人的优势 (相比人工和传统机械/机器)	1、提高效率； 2、降低人工成本。	1、降低人力焊接成本； 2、提高焊接质量； 3、降低焊接之后的不利后果。	1、提高效率； 2、提高装配质量； 3、装配生产面积更小； 4、可在危险环境下完成装配。	1、更柔性； 2、更灵活； 3、成本更低。	1、提高效率； 2、提高喷涂质量和材料使用率； 3、易操作和维护； 4、设备利用率高。	1、更洁净； 2、更安全。	/

资料来源：亿欧智库《2022 中国工业机器人市场研究报告》、华安证券研究所

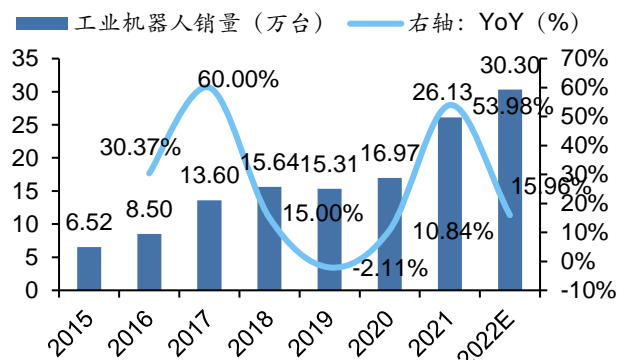
后疫情时代行业复苏，国内工业机器人增长迅猛。近年来全球及国内机器人产业发展迅猛，尤其是后疫情时代，劳动力短缺更加严峻，机器人成为企业实现快速复工复产的重要选择，“机器人替人”需求日益提升。产销量方面，据国家统计局数据，2015-2022 年，国内工业机器人产量从 3.30 万台增长至 44.31 万台，CAGR 高达 45%，而销量上根据 GGII 数据，2021 年国内工业机器人销量达 26.13 万台，预计 2022 年销量达 30.30 万台。市场规模方面，IFR 统计数据显示，2021 年全球工业机器人市场反弹强劲，市场规模为 175 亿美元，国内方面，工业机器人市场规模创新高达 75 亿美元。随着未来市场需求的持续释放，2024 年工业机器人全球市场规模有望达 230 亿美元，国内市场规模有望达 115 亿美元，其中国内工业机器人的市场份额占比有望从 2017 年的 28% 增长至 2024 年的 50%。综合来看，工业机器人是制造业实现自动化生产，提高生产效率的关键，国内工业机器人将迎来爆炸式增长，成为全球工业机器人发展的主动力。

图表 86 2015-2022 年中国工业机器人产量



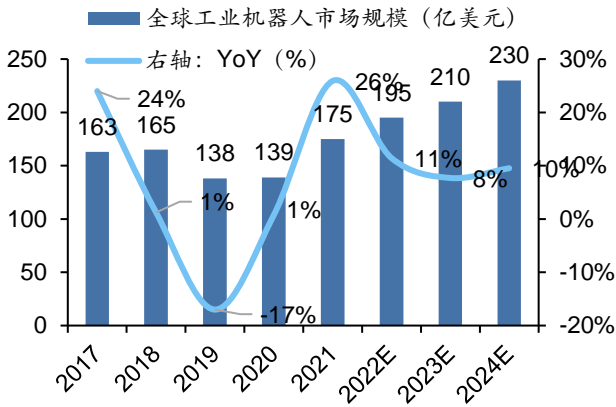
资料来源：国家统计局，华安证券研究所

图表 87 2015-2022 年中国工业机器人销量



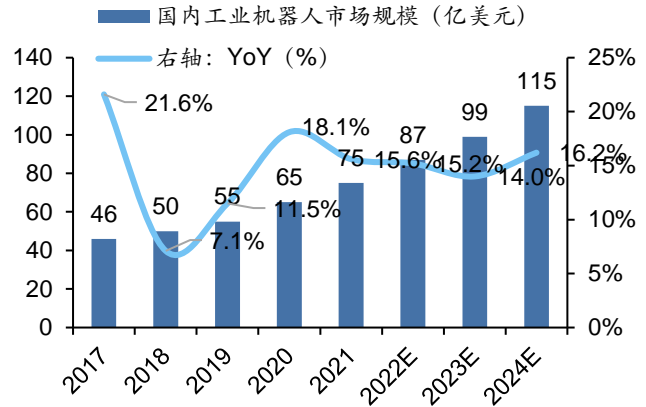
资料来源：GGII，华安证券研究所

图表 88 2017-2024 年全球工业机器人市场规模（销售额口径）



资料来源：IFR，中国电子学会，华安证券研究所

图表 89 2017-2024 年中国工业机器人市场规模（销售额口径）



资料来源：IFR，中国电子学会，华安证券研究所

“制造业升级+劳动力成本上升+政策扶持”三大因素共振，工业机器人崛起是必然趋势。我们认为未来工业机器人的需求有望持续加速提升，主要是因为：

1) 制造业产业升级。“机器人替人”不止是单纯的解放双手，更是产业效率与柔性提升的重要手段。在工业 4.0 时代，提升制造业自动化水平，实现由劳动密集型产业向技术密集型产业转变已经成为我国成为“制造强国”，提升国家竞争力的必经之路。

2) 劳动力成本上升。随着人口老龄化程度的加剧，廉价劳动力成本已成为过去式，再叠加制造生产中的安全考量，劳动力成本持续提升，在原材料上涨，制造业利润空间被压缩的背景下，工业机器人是产业降本增效，提高利润的不二选择。

3) 政策大力扶持。2016 年，机器人产业发展被首次写入“十三五”规划，而进入“十四五”阶段，中央及地方更是密集出台政策扶持。2021 年 12 月，工信部等十五部门印发《“十四五”机器人产业发展规划》；2023 年 1 月，工信部等十七部门联合发布《“机器人+”应用行动实施方案》，提出到 2025 年，制造业机器人密度较 2020 年实现翻番。

综上，虽然目前国内工业机器人市场仍处于成长阶段，但中长期的发展逻辑清晰，由此我们坚定认为制造制造时代，工业机器人前景广阔。

➤ 高端数控机床

数控机床是一种装有程序控制系统的自动化机床，具备柔性和高效能的特点。近年来国内数控机床技术在高速化、复合化、精密化、多轴化等方面取得了重要突破，高端数控机床产业发展迅速。从组成来看，高端数控机床包括加工程序载体、数控装置、伺服系统、机床主体和其他辅助装置。具体来看：1) 加工程序载体。即以一定的格式和代码存储零件加工程序，从而对数控机床进行控制。2) 数控装置。属于数控机床的核心，多采用 CNC 系统，通过计算机系统程序的合理组织，整个系统协调的进行工作。3) 伺服与测量反馈系统。主要用于实现数控机床的伺服控制，包括驱动装置和执行机构两大部分。4) 机床主机。指在数控机床上自动地完成各种切削加工的机械部分，包括床身、底座、立柱、滑座、主轴箱、刀架等机械部件。5) 其他辅助装置。证充分发挥数控机床功能所必需的配套装置，常用的辅助装置包括：气动、液压装置，排屑装置，冷却、润滑装置等。

图表 90 高端数控机床的基础构成



资料来源：华安证券研究所整理

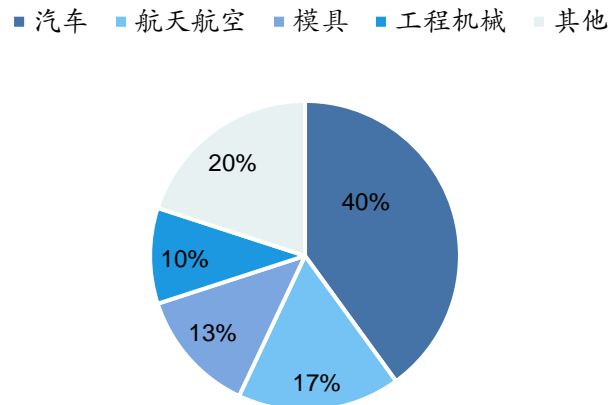
高端数控机床下游应用广泛，汽车为第一大应用领域。机床是制造业的“母机”，广泛应用在机械制造、汽车、电力设备、铁路机车、船舶、国防工业、航空航天、石油化工、工程机械、电子信息工业等行业。从下游应用领域占比来看，根据中商产业研究院数据，汽车是第一大下游需求领域，占比约 40%，航空航天领域位列第二，占比 17%。

图表 91 高端数控机床的下游应用

行业	针对零件	需求机型
航空工业	飞机机翼、机身、尾翼等和发动机零件	需要大批高速五轴加工中心、龙门移动式高速加工中心、精密数控车床、精密卧式加工中心、多坐标镗铣中心、精密齿轮和螺纹加工数控机床等
铁路机车制造业	高铁机车车体、车轴、车轮等	大中型数控机床：需要大批高速五轴加工中心、龙门移动式高速加工中心、精密数控车床、精密卧式加工中心、多坐标镗铣中心、精密齿轮和螺纹加工数控机床等
兵器制造业	坦克、装甲车辆、弹炮、引芯等产品	数控车床、立卧式加工中心、五轴加工中心、龙门镗铣床、镗铣加工中心、齿轮加工机床等
模具制造业	汽车覆盖件模具，压铸模具，成型挤压模具等	高速数控铣床、精密电加工机床、高精度加工中心、精密磨床
电子信息设备制造业	高端电子产品外壳、电机转子定子、电机壳盖等	小型精密数控机床：高速铣削中心、高速加工中心、小型精密车床、小型精密冲床、精密和超精密加工专用数控机床及精密电加工机床
电力设备制造业	发电设备	重型数控龙门镗铣床、大型落地镗铣床、大型数控车床、叶根槽专用铣床和叶片数控加工机床等
	输变电设备	数控车床、加工中心、数控镗床
冶金设备制造业	连铸连轧成套设备	大型龙门铣床、大型数控车床
工程机械制造业	变速箱、挖掘臂、车体、发动机等	中小型数控机床：数控车床、中型加工中心、数控铣床和齿轮加工机床等
造船工业	柴油机体	重型、超重型龙门铣镗床和重型数控落地镗铣床以及大型数控车床和车铣中心、大型数控磨齿机、曲轴镗铣床、大型曲轴车铣中心和曲轴磨床等
汽车制造业	整车部件：发动机	高效、高性能、专用数控机床和柔性生产线
	零配件加工	数控车床、立卧式加工中心、数控高效磨床等

资料来源：海天精工招股说明书、华安证券研究所

图表 92 中国数控机床下游应用领域占比

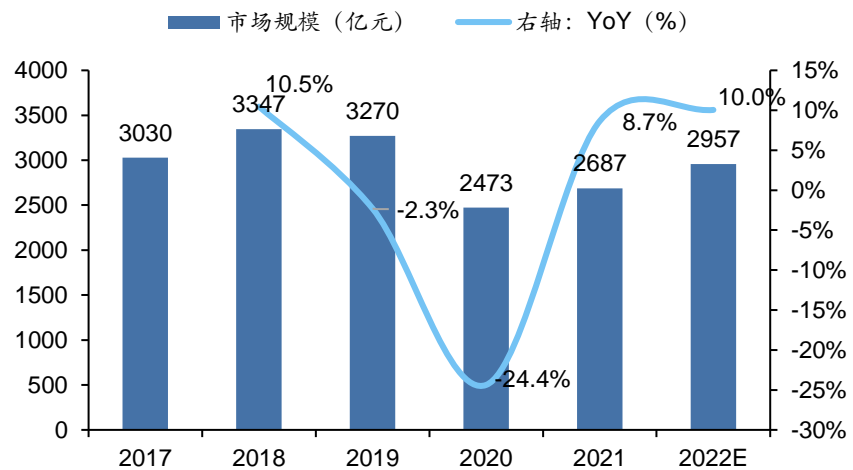


资料来源：中商产业研究院，华安证券研究所

数控机床优势明显，后疫情时代市场规模恢复增长。与普通机床相比，数控机床的优点众多，具体来看：1) 高度柔性。在数控机床上加工零件，主要取决于加工程序，因此其适用于所加工的零件频繁更换的场合，能较大程度缩短生产周期并节省费用。2) 加工精度高。数控机床是按数字信号形式控制的，加工精度更高。3) 加工质量稳定、可靠。4) 生产率高。数控机床可有效地减少零件的加工时间和辅助时间。因此，高端数控机床正广泛应用于大制造领域。**从市场规模来看**，根据中商产业研究院统计数据显示，2019年我国数控机床市场规模达3270亿元，但2019年后，受疫情影响，2020年国内数控机床市场规模缩减至2473亿元。受益于国内疫情控制良好，各行业开始复工复产，2021年数控机床市场规模恢复增长，达2687亿元。中商产业研究院预测，2022年数控机床产业规模将达2957亿元。

就目前形势来看，我们认为，一方面，随着国内疫情管控放开，经济环境向好，各行各业活力加速恢复，数控机床产业有望受益于下游需求加速扩张；**另一方面**，随着上游核心零部件国产化水平提升，数控化率也会逐步提升，国产替代空间广阔，因此，高端数控机床产业势在必行。

图表 93 2017-2022 年中国数控机床市场规模统计预测



资料来源：中商产业研究院，华安证券研究所

多因素持续支持机床行业稳健发展。

支持 1: 政策持续发力。国家对高端数控机床行业发展给予高度关注。《中国制造 2025》将高端数控机床列为制造业重点发展领域之一，2022 年 9 月首批机床 ETF 发行，为具有核心技术及科创能力突出的机床企业提供直接融资便利。政策端持续支持国产高端数控机床行业发展。

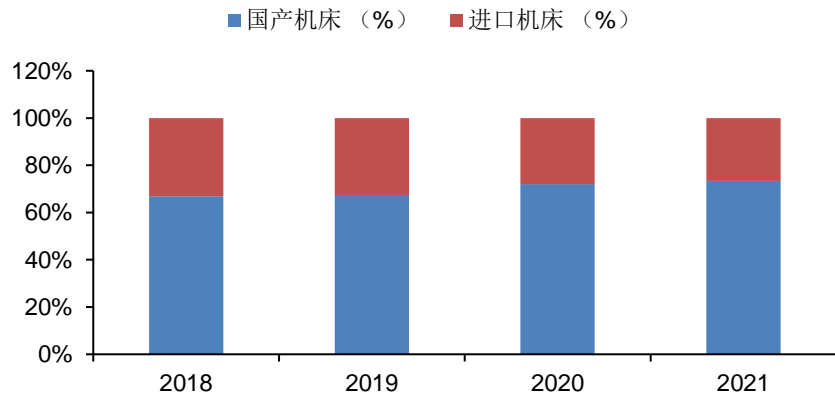
图表 94 我国支持高端数控机床发展的文件/政策

时间	文件/会议	重点内容
2015.05	《<中国制造2025>重点领域技术路线图》	重点领域技术路线图对机床关键部件国产化提出了明确的国产化目标：到 2020 年，数控系统标准型、智能型国内市场占有率分别达到 60%、10%，主轴、丝杠、导轨等中高档功能部件国内市场占有率达到 50%； 到 2025 年，高档数控机床与基础制造装备国内市场占有率超过 80% ；数控系统标准型、智能型国内市场占有率分别达到 80%、30%；主轴、丝杠、导轨等中高档功能部件国内市场占有率达到 80%；高档数控机床与基础制造装备总体进入世界强国行列。
2019.1	《产业结构调整指导目录（2019 年版）》	将“ 高端数控机床及配套数控系统，五轴以上联动数控机床 ，数控系统，高精密、高性能的切削刀具、量具量仪和磨料磨具”产品列为鼓励发展项目
2021.03	《中华人民共和国国民经济和社会发展第十四个五年规划和 2035 年远景目标纲要》	培育先进制造业集群，推动集成电路、航空航天、船舶与海洋工程装备、机器人、先进轨道交通装备、先进电力装备、工程机械、 高端数控机床 、医药及医疗设备等产业创新发展。
2022.06	《重大技术装备推广应用导向目录——机械工业领域（2022 年版）》	高端工业母机中数控机床位于首位，明确高端工业母机中数控机床包括精密车削中心、精密复合磨削中心、精密数控坐标磨床、精密坐标镗加工中心、卧式五轴加工中心、立式五轴加工中心、五轴联动铣车复合加工中心、高精度车铣复合加工中心、龙门式车铣复合加工中心等。
2022.09	工信部“大力发展高端装备制造业”主题新闻发布会	工信部将会同有关部门继续做好工业母机行业顶层设计， 统筹产业、财税、金融等各项政策，积极推进专项接续 ，进一步完善协同创新体系和机制，突破核心关键技术，强化产业基础，培育优质企业和产业集群，保持产业链供应链稳定，推动工业母机行业高质量发展。
2022.09	国内首批机床 ETF 获批	华夏中证机床 ETF 和国泰中证机床 ETF 获证监会批准，涉及主机厂、数控系统、主轴、切削工具等领域上市公司

资料来源：《<中国制造 2025>重点领域技术路线图》，新华社，中国机械工业联合会，工业和信息化部网站，发展改革委网站，中国工业报，华安证券研究所

支持 2: 国产厂商竞争力提升，高端数控机床进口替代加速。2020-2021 年由于海外疫情影响，高端数控机床进口受限，为部分高端国产数控机床厂商提供宝贵窗口期打入高端下游，加速高端数控机床产品进口替代。在国产高端数控机床竞争力不断提升背景下，进口机床在我国机床消费额中占比呈持续下降趋势，由 2018 年占比 33% 降低至 2021 年 27%。海天精工、科德数控五轴联动机床已批量打入航空航天零部件产业链供应，体现我国高端数控机床竞争力，加速高端数控机床领域进口替代。

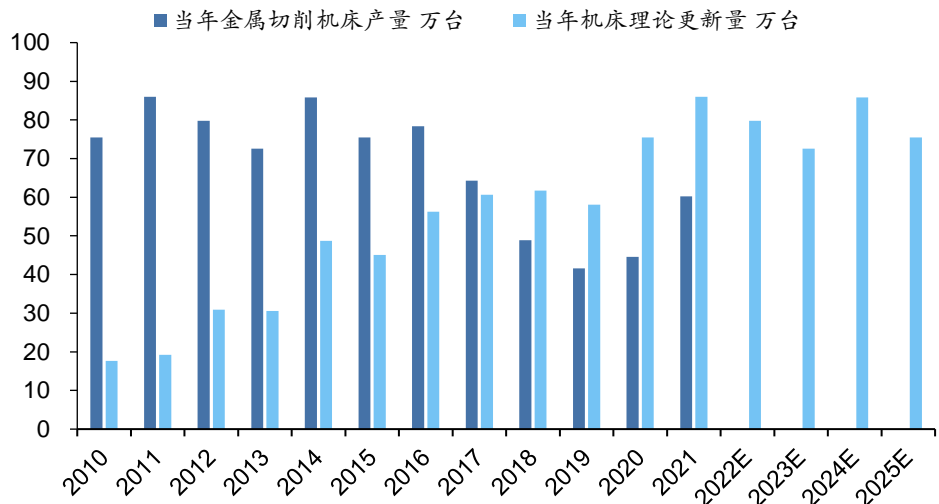
图表 95 中国机床消费结构



资料来源: VDW, 华安证券研究所

支持 3: 更新周期持续放量。 由于设备零件使用寿命及技术更新影响, 机床具备十年更新周期。以机床十年使用寿命测算机床理论更新量, 2021 年机床理论更新量为 2011 年产量即 86 万台, 达到历史新高。目前仍处于更新周期持续放量阶段, 有望持续托底机床行业需求, 而高端数控机床需求也有望持续增长。

图表 96 机床更新需求测算



资料来源: WIND, 华安证券研究所

2) 智能工厂是实现智能制造的载体, 协作机器人与智能仓储物流是工厂新星

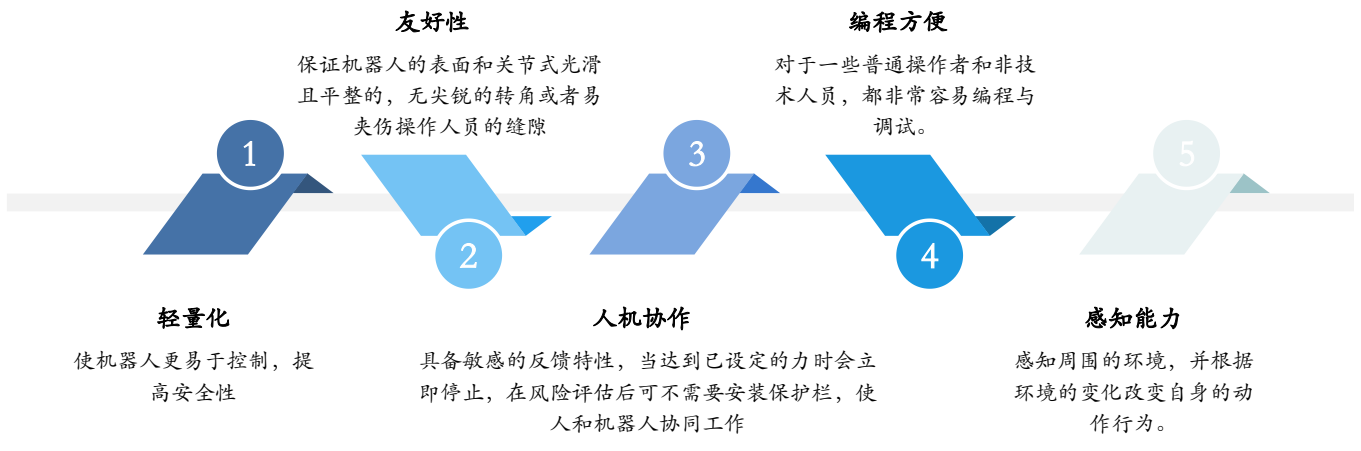
➢ 协作机器人

协作机器人是一种新型的工业机器人, 扫除了人机协作的障碍, 机器人与人可以在生产线上协同作战, 充分发挥机器人的效率及人类的智能, 给未来工厂的工业生产和制造带来了根本性的变革。

与传统工业机器人有所区别, 协作机器人优势特点明显。与传统工业机器人相比, 其不同之处在于: 1) 目标市场不同, 协作机器人偏向于应用在中小企业及适应柔性化生产要求的企业, 而传统工业机器人适用于大规模生产企业。2) 模式不同, 传统工业机器

人是作为整个生产线的组成部分，如果某个环节机器人坏了，整条产线可能会面临停工的风险，而协作机器人更具备柔性特点，代替的是人，使得整个生产流程更为灵活。优势方面，协作机器人具备轻量化、友好性、人机协作、编程方便及感知能力五大特点。

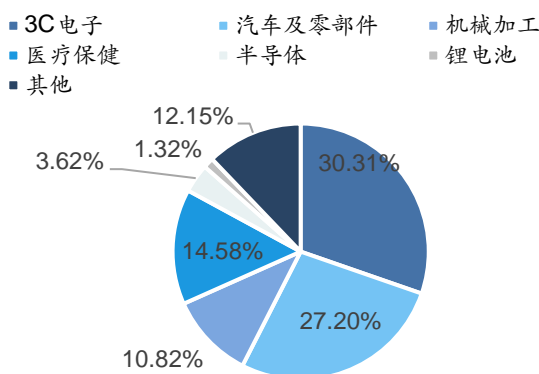
图表 97 协作机器人特点



资料来源：华经产业研究院、华安证券研究所

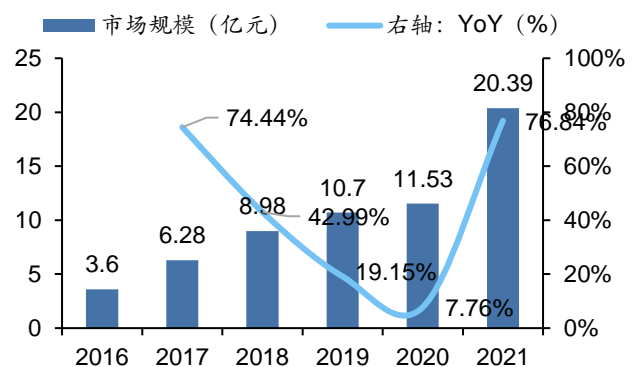
协作机器人下游应用渗透加速，市场规模加速提升。近年来协作机器人的快速发展，和下游行业应用的渗透率有较大关系。根据 GGII 统计，在工业领域，国内协作机器人主要应用于 3C 电子、汽车及零部件、机械加工、医疗保健、半导体、锂电等行业，其中**2021 年 3C 电子和汽车及零部件行业占比较高，分别为 30.31%和 27.20%**，与前一年相比，2021 年协作机器人的市场增量主要来自于汽车及零部件和医疗保健等行业。**市场规模方面**，2016-2021 年，国内协作机器人市场规模从 3.60 亿元增长至 20.39 亿元，CAGR 高达 41%，其中 2021 年同比增速达 76.84%，意味着国内协作机器人应用拓展势头强劲。**从全球角度来看**，2021 年中国协作机器人销量 1.86 万台，同比增长 87.62%，展全球协作机器人市场整体销量的 51.02%。GGII 预测，2026 年中国协作机器人市场销量有望突破 10 万台。

图表 98 2021 年中国协作机器人应用行业分布



资料来源：GGII，华安证券研究所

图表 99 2016-2021 年中国协作机器人市场规模



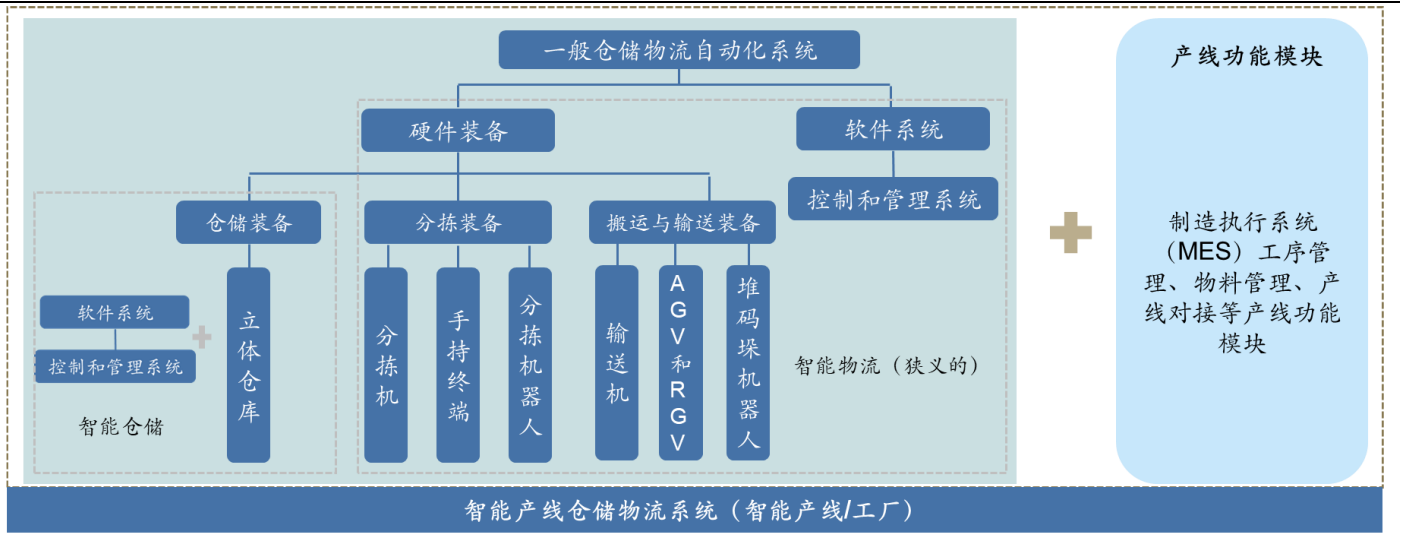
资料来源：GGII，华安证券研究所

智能仓储物流

智能仓储物流系统是通过信息化、物联网和机电一体化共同实现的智慧物流解决方案，通过将物料出入库、存储、输送、生产、分拣等物流过程自动化、信息化和智能化，来实现降本增效的目的。从构成来看，智能仓储物流系统包括硬件装备和软件系统。其

中，硬件装备可按环节分为仓储装备、分拣装备和搬运与输送装备，具体的产品包括立体仓库、堆垛机、穿梭车、输送机、AGV、码垛机器人和分拣机等，主要执行具体的仓储物流操作任务；软件系统是智能仓储物流系统的控制中心，主要包括仓储管理系统WMS和仓储控制系统WCS，负责具体的仓储物流信息控制。从环节划分，仓储装备与软件系统的结合即智能仓储；分拣与输送、搬运装备与软件系统的结合称为狭义的智能物流；若智能仓储物流系统与生产线对接，增加物料管理、产线对接等产线功能模块，即构成智能产线仓储物流系统，也就是常说的智能产线，也是智能工厂的基础版。

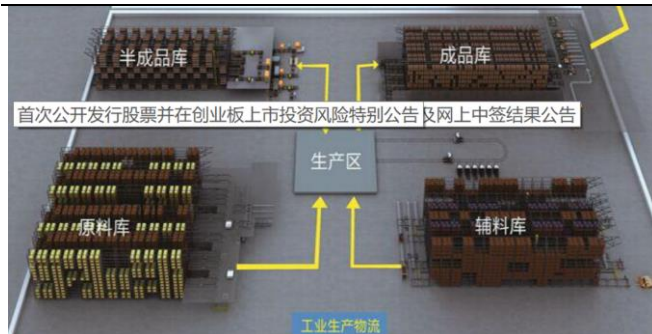
图表 100 仓储物流自动化系统构成



资料来源：华安证券研究所整理

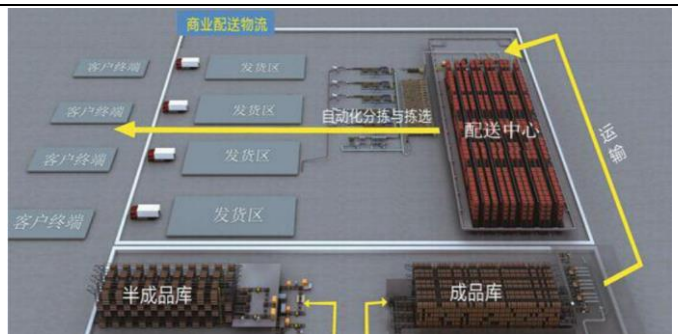
根据企业业务性质，智能仓储物流系统可分为工业生产型和商业配送型仓储物流自动化系统。1) 工业生产型：为工业企业提供原材料、半成品、成品以及零备件等货物存储、输送和信息化管理，实现物料自动传输与订单自动处理，提高生产配套效率、车间物流管理水平以及仓储管理能力。2) 商业配送型：为商业企业提供产品存储、分拣、配送和信息化管理，实现信息自动传输与订单自动处理，提高订单处理能力、降低订单分拣成本，减少流通成本。从联系来看，两种类型的仓储物流系统使用的硬件装备和软件信息系统均具有类似性；从区别来看，工业生产型侧重于物流系统与生产线的对接，满足生产线的物流需求、提高生产效率；而商业配送型则侧重于物料分拣、配送的效率和准确性。

图表 101 工业生产物流示意图



资料来源：今天国际招股书，华安证券研究所

图表 102 商业配送物流示意图



资料来源：今天国际招股书，华安证券研究所

智能仓储物流优势明显，相对使用成本呈不断下降趋势。我们从空间利用率、储存量、存储形式等六个维度对比智能仓储物流和传统仓储物流，可以发现，仓储物流自动化系统具有节约用地、减少劳动需求、降低货物拣选差错率、提高仓储自动化水平及管理水平、提高物流效率等诸多优点。以土地成本为例，目前，世界上最高的自动化立体仓库高度已达 50m，立体仓库货架区单位面积的储存量可达 7.5t/m²，是传统普通仓库的 5-10 倍，与传统普通仓库相比可以节约 40%-70%的占地面积，可以大幅降低土地成本。

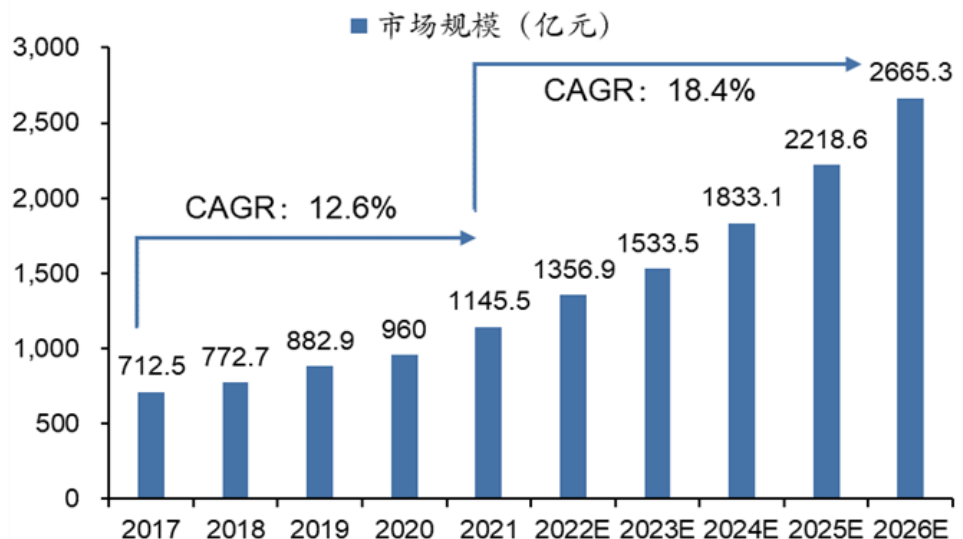
图表 103 智能仓储物流与传统仓储物流对比

对比项目	智能仓储物流	传统仓储物流
空间利用率	高层货架，充分利用仓库的垂直空间，空间利用率高	低层货架，需占用大面积土地，空间利用率低
存储量	高层货架及密集存储，货物存储量倍数增加	低层货架，货物存储量较少
存储形态	动态存储，货物在仓库内能够按需要自动存取	静态存储，只是货物存储的场所，需人工进行拣选及存取
作业效率	货物在仓库内按需要自动快速存取	主要依靠人力，货物存取速度慢
人工成本	减少人员数量，可以大幅节约劳动力成本	人员需求量大，人工成本高
环境要求	能适应黑暗、低温、有毒等特殊环境的要求	受黑暗、低温、有毒等特殊环境影响很大

资料来源：兰剑智能招股说明书，华安证券研究所

市场空间前景广阔，黄金赛道有望开启高增。从智能仓储市场规模来看，根据头豹研究院统计的数据来看，受益于物流行业规模的迅速增长和仓储环节降本增效的需求不断攀升，市场规模从 2017 年的 712.5 亿元增至 2021 年的 1145.5 亿元，CAGR 达 12.6%。但从设备渗透率中，也反映出智能仓储物流在很多领域的融合程度不足，因此向未来来看，随着智能仓储物流与更多应用场景融合和 5G、物联网、人工智能等技术在仓储行业的进一步升级，智能仓储行业有望加速发展，预计 2021-2026 年 CAGR 达 18.4%，2022 年智能仓储物流市场规模约 1357 亿，2026 年达 2665 亿元。

图表 104 2017-2026 年中国智能仓储物流市场规模及预测



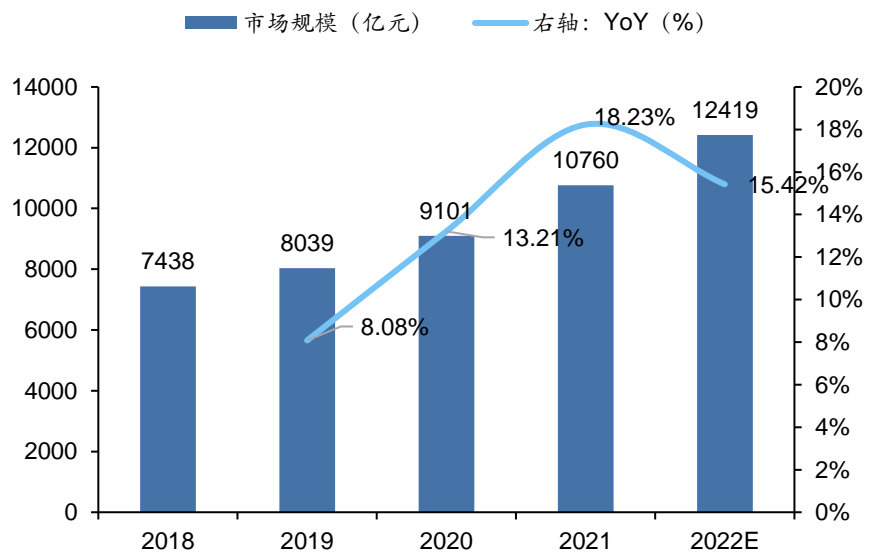
资料来源：头豹研究院，华安证券研究所

6.3 AI 赋能工业互联网，打造高效率设备管理和生产流程

智能服务是智能制造的必然延伸，工业互联网是制造业智能化的重要发展方向。工业互联网是数字化转型的关键力量，更是推动制造业转型发展的重要支撑。工业互联网是全球工业系统与高级计算、分析、感应技术以及互联网连接融合的一种结果。其本质是通过开放的、全球化的工业级网络平台把设备、生产线、工厂、供应商、产品和客户紧密地连接和融合起来，高效共享工业经济中的各种要素资源，从而通过自动化、智能化的生产方式降低成本、增加效率，帮助制造业延长产业链，进而推动制造业转型发展。换言之，工业互联网是实现智能制造的发展模式和现实路径。其产生背景也是国内人口红利的消失，劳动力成本上升。工业互联网则通过信息技术与工业系统的深度融合，保证对制造成本控制，因此，工业互联网已成为主要工业国家抢占国际制造业竞争的制高点。

平台应用普及率稳步提升，市场总规模持续突破。近年来我国工业互联网平台稳中向好，工业互联网平台应用普及率由 2020 年的 14.67% 提升至 17.5%。根据国务院印发的《“十四五”数字经济发展规划》，其中明确提出到 2025 年工业互联网平台应用普及率达到 45%。因此，可以确定的是，“十四五”期间，工业互联网平台建设有望加速。从市场规模来看，根据中国信息通信研究院数据，2018-2021 年，工业互联网总规模从 7348 亿元增长至 10760 亿元。据中商产业研究院预测，2022 年工业互联网市场规模将继续加速，有望达 12419 亿元。

图表 105 2018-2022 年中国工业互联网市场规模

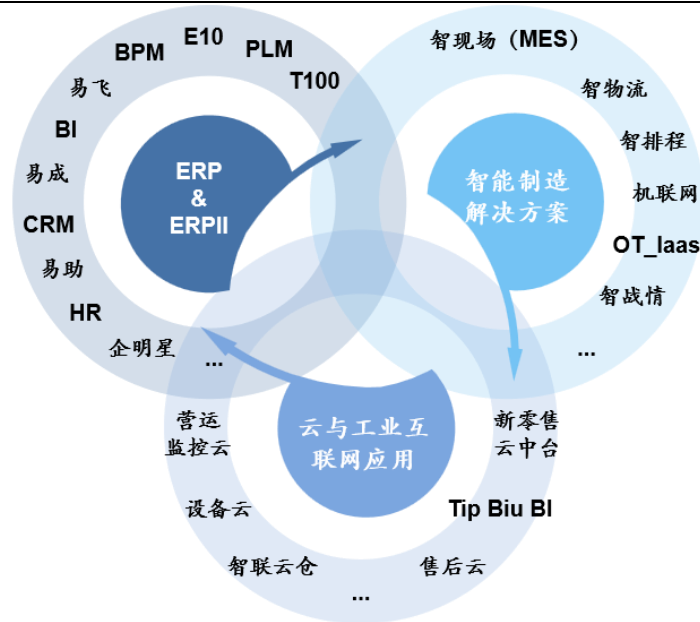


资料来源：中国信息通信研究院，华安证券研究所

对于传统 ERP、MES 等经营管理、生产控制类软件，AI 的加入将提高其对外部需求的实时应对能力和综合决策能力。

成本结算方面，依据制造费用、产品良率，快速结算各产品实际成本，并提供实际成本与标准成本比较，以分析每笔委工因良率差异造成的成本落差。**生产规划与排程方面**，提高派工调度及管理能力和高度透明化的车间信息、高效的设备和物料管理，以及对移动终端的广泛适配，有效减少无效作业、准备和等待时间，加快车间的相应速度，加强质量管控和物料利用。**销售预测方面**，将销售预测纳入系统记录，生管可利用销售预测安排生产与采购，让生产工单开立可以更加精确，库存预估更加符合实际状况。

图表 106 鼎捷经营管理、生产控制方案



资料来源：鼎捷软件官网，华安证券研究所

AI 算法的加入将有效提升设备运行稳定性和流程控制效率。以流程工业 PHM 智能监测系统为例，传统监测方式需要大量人工进行实时排障，随着 AI 算法的加入，PHM 有望快速实现智能化、无人化，减少紧急维修事件发生的概率，降低其带来的停机、排障、维修损失，同时降低不必要的检修次数和对应成本。对于化工、冶金、煤炭等流程行业来说，每一个泵、反应釜、管路、温控设备出现故障都会对产线造成重大影响，甚至酿成生产事故。AI 带来的能力提升将意义匪浅。

图表 107 智能运维平台



资料来源：容知日新，华安证券研究所

AI 赋能工业设计类软件，提高建模及仿真效率。以中望推出的 3D One AI 软件为例，通过基于物理刚体运动与三维数据处理技术，融合开源硬件、人工智能、编程等多学科实践，能够实现动态的人工智能行为仿真，以及输出三维画面。随着 AI 模型的迭代，以上识别功能的精度、泛用性将不断提升，逐渐从辅助设计向自动化设计进步，快速输出大量可选设计草案。

图表 108 中望软件 3D 仿真



资料来源：中望软件官网，华安证券研究所

6.4 人工智能助力汽车智能化

1) 智能驾驶：从驾驶辅助到自动驾驶

自动驾驶汽车依靠人工智能、视觉计算、雷达、监控装置和全球定位系统协同合作，它是一个集环境感知、规划决策、多等级辅助驾驶等功能于一体的综合系统，它集中运用了计算机、现代传感、信息融合、通讯、人工智能及自动控制等技术，是典型的高新技术综合体。这种汽车拥有和人一样的“思考”、“判断”以及“行走”能力，使得电脑可以在没有任何人主动的操作下，能够自动安全地操作机动车辆。

按照《汽车驾驶自动化分级》，驾驶自动化共分为：应急辅助、部分驾驶辅助、组合驾驶辅助、有条件自动驾驶、高度自动驾驶以及完全自动驾驶六个层级。

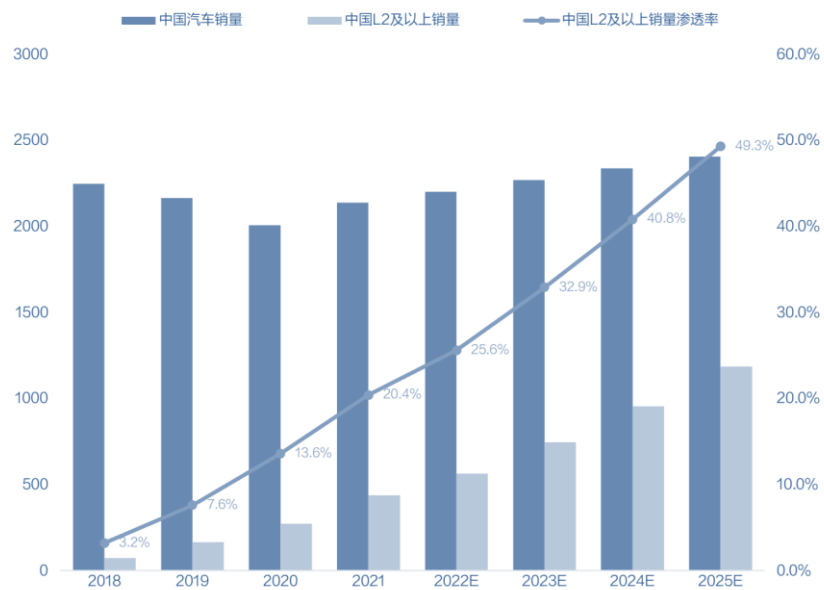
图表 109 汽车驾驶自动化等级划分

自动化等级	名称	释义
L0	应急辅助	系统不能持续执行动态驾驶任务中的车辆横向或纵向运动控制，但具备持续执行动态驾驶任务中的 部分目标和事件 探测与响应的能力。
L1	部分驾驶辅助	系统在其设计运行条件下持续地执行动态驾驶任务中的车辆横向 或 纵向运动控制，且具备与所执行的车辆横向或纵向运动控制相适应的 部分目标和事件 探测与响应的能力。
L2	组合驾驶辅助	系统在其设计运行条件下持续地执行动态驾驶任务中的车辆横向 和 纵向运动控制，且具备与所执行的车辆横向和纵向运动控制相适应的 部分目标和事件 探测与响应的能力。
L3	有条件自动驾驶	系统在其设计运行条件下持续地执行 全部动态驾驶任务 。
L4	高度自动驾驶	系统在其设计运行条件下持续地执行全部动态驾驶任务 并自动执行最小风险策略 。
L5	完全自动驾驶	系统在 任何可行驶条件 下持续地执行全部动态驾驶任务并自动执行最小风险策略。

资料来源：工信部，华安证券研究所

据《2022 中国智能汽车发展趋势洞察报告》，未来随着智能汽车的普及，关键零部件成本将持续下探，叠加产业环境的成熟和科技的不断进步，预计到 2025 年，中国 L2 及以上智能汽车而销量破千万辆，对应中国智能汽车渗透率大 49.3%，智能汽车市场潜力巨大。

图表 110 中国 L2 及以上智能汽车销量及渗透率



资料来源：汽车之家研究院，华安证券研究所

智能驾驶的感知、处理和执行都离不开人工智能技术的基础。首先，感知：让车辆配对相应的感测器来收集车辆的行驶状况和道路环境状况等信息。不同的系统需要由不同类型的车用感测器，包含毫米波雷达、超声波雷达、红外雷达、雷射雷达、CCD\CMOS 影像感测器及轮速感测器等来收集整车的工作状态及其参数变化情形；其次，处理：将感测器所搜集的信息进行分析处理，再向智能控制中心传达控制讯息；最后，执行。根据控制中心下达的命令，完成对汽车的驾驶。

人工智能在自动驾驶定位技术中可应用于自动驾驶图像识别与感知、自动驾驶深度学习以及自动驾驶信息共享等领域。其中在自动驾驶图像识别与感知领域中，从感知、认知、行为三个方面看，感知部分难度最大，人工智能技术应用最多。感知技术依赖于传感器，比如摄像头，由于其成本低，在产业界倍受青睐。在自动驾驶深度学习领域中，深度学习作为无人驾驶技术成功的基础，可以提高汽车识别道路、行人、障碍物等的时间效率，并保障识别的正确率。通过大量数据的训练之后，汽车可以将收集到的图形，电磁波等信息转换为可用的数据，利用深度学习算法实现无人驾驶。在自动驾驶信息共享领域中，汽车信息共享所收集到的交通信息量将非常巨大，如果不对这些数据进行有效处理和利用，就会迅速被信息所湮没。因此需要采用数据挖掘、人工智能等方式提取有效信息，同时过滤掉无用信息。考虑到车辆行驶过程中需要依赖的信息具有很大的时间和空间关联性，因此有些信息的处理需要非常及时。

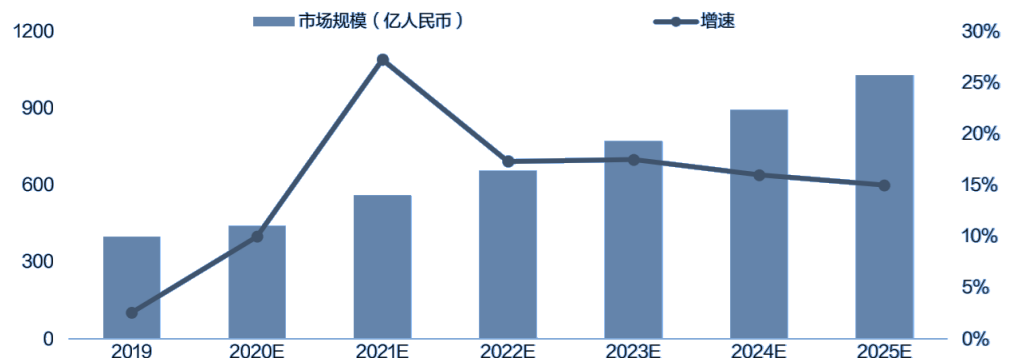
自动驾驶商业化成果显著，但人工智能技术仍处于学习阶段。人工智能的较多应用场景有较宽松的容错性，然而，在自动驾驶行业中，决策规划作为整体系统的大脑，处于算法中的核心模块，对于自动驾驶功能的实现以及安全性的保障起着重要的作用，而基于强化学习的决策规划系统还远远未达到成熟，基于规则的系统也难以应对极端的Corner case，这也使得从业者普遍认为是L4甚至L5级别的成熟自动驾驶难以实现的根本原因。因此，现有自动驾驶商业部署主要在封闭的园区和有严格管控的固定线路。

2) 智能座舱：从出行工具到出行管家

智能座舱实质是汽车驾驶舱中的人机交互场景，主要目的是将驾驶信息与娱乐信息两个模块进行集成，利用自身处理海量信息数据的强大能力，把握用户在不同场景下的行为习惯，并以此优化智能座舱的空间结构，进而提升用户的驾乘体验。

据《2022中国智能汽车发展趋势洞察报告》，未来随着智能化技术的不断普及，汽车不再仅仅是交通工具，将扩展成为日常生活的第三空间，用户对乘坐体验要求更高，智能座舱将加速普及。根据ICVTank数据，2022年全球智能座舱行业市场规模有望达461亿美元，中国作为全球最具发展潜力的汽车市场，2019年中国智能座舱市场规模达441亿人民币，预计2025年市场规模将达1030亿人民币，2017-2025年的复合增长率为13%，发展潜力巨大。

图表 111 中国智能座舱市场规模及预测



资料来源：ICVTank，华安证券研究所

智能座舱是由不同的座舱电子组成的完整体系，其关键技术主要由四部分组成。第一部分是机械技术，包括可变化车体技术和内饰机构技术。未来汽车可根据不同模式进

行伸缩折叠是一种趋势，座舱需要可以根据乘客对于不同场景的使用需求，实现内饰空间的不断调整变化。第二部分是电子硬件技术，包含芯片技术、显示屏技术、专用电器总成以及传感器技术四大技术。第三部分是软件技术，主要有操作系统和各种应用软件。汽车智能化发展必然会趋向于一机多屏，通过操作系统实现一个车机芯片控制各屏的软件。第四部分是两大支撑技术，分别是人工智能技术和云计算技术。未来智能算法的准确性决定了不同品牌智能座舱的差异化，是影响车内体验的关键。

智能座舱通过以上的机械、电子硬件、软件以及两大支撑技术实现人机交互。以中控屏、仪表盘、后座娱乐屏等硬件为载体，实现人脸识别、多音区检测、视线追踪、疲劳分级以及多模命令词等核心智能化功能。除此以外，还可以衍生出音量自动调节、高速分心提醒、疲劳拯救等个性化附加功能。未来，智能座舱将应用到增强现实和虚拟现实技术，人机交互形式将愈发多样，除语音控制外还能拜托手势交互和触摸交互的不足。

与智能驾驶相似的是，自我学习也可渗透到智能座舱的方方面面。学习能力、个性化和人性化的设计是智能座舱主动交互设计的重要内容。得益于人工智能技术的发展，未来的智能座舱将通过自学习能力，帮助汽车从出行工具向出行管家进行转变。比如用户可通过语音、手势、眼神等来实现与汽车的对话；车内灯光可根据不同场景来自动变暗或变亮等。

感知器件丰富、多样，与安全驾驶挂钩并做到更精准的识别。类车外环境感知之于辅助/自动驾驶，车内环境的智能感知之于智能座舱同样重要。尤其在 L2、L3 级自动驾驶“人机共驾”阶段，驾驶员监控系统（DMS）愈发受到重视，能做到对驾驶员的潜在危险驾驶行为进行预判和提醒，而 DMS 与 ADAS 系统数据的融合，则能达到进一步对安全驾驶干预的效果。同时，DMS 的功能和应用也得到延伸扩展，包括检测范围的进一步覆盖副驾、后排的乘员监控系统（OMS），实现人脸、年龄、身份、情绪识别，物品（遗留物）检测，基于身份 ID 的定制化智能服务和交互等。

ChatGPT 加速落实“汽车机器人概念”。百度于 21 年 8 月举办的百度世界大会 2021 上首提“汽车机器人”的概念，并发布了具有跨时代意义的 Apollo“汽车机器人”。今年 2 月 15 日百度旗下智能汽车公司品牌集度计划年内推出“三体版”汽车机器人，将融合百度文心一言的全面能力，打造针对智能汽车场景的大模型人工智能交互体验，支持汽车机器人实现自然交流的再进阶。近日，长安汽车旗下深蓝品牌公众号发布题为“假如把 ChatGPT 装进长安深蓝 SL03”的文章，内容为 ChatGPT 与长安深蓝 SL03 车机系统 DEEPAL OS 的一番对话。通过这次对话，DEEPAL OS 将自身特点、优势逐一说给 ChatGPT 的同时，亦把自己介绍给了消费者和网友。

图表 112 集度：汽车机器人将融合百度文心一言全面能力

图表 113 长安深蓝：假如把 ChatGPT 装进长安深蓝 SL03



资料来源：新华网，华安证券研究所



资料来源：长安深蓝，华安证券研究所

风险提示:

国内大模型推出时间、效果不及预期风险;
行业竞争加剧风险;
人工智能在部分领域应用的伦理道德风险;
由外环境导致的芯片及芯片技术供应风险;
商业模式变现与落地不及预期风险。

重要声明

分析师声明

本报告署名分析师具有中国证券业协会授予的证券投资咨询执业资格，以勤勉的执业态度、专业审慎的研究方法，使用合法合规的信息，独立、客观地出具本报告，本报告所采用的数据和信息均来自市场公开信息，本人对这些信息的准确性或完整性不做任何保证，也不保证所包含的信息和建议不会发生任何变更。报告中的信息和意见仅供参考。本人过去不曾与、现在不与、未来也将不会因本报告中的具体推荐意见或观点而直接或间接接收任何形式的补偿，分析结论不受任何第三方的授意或影响，特此声明。

免责声明

华安证券股份有限公司经中国证券监督管理委员会批准，已具备证券投资咨询业务资格。本报告由华安证券股份有限公司在中华人民共和国（不包括香港、澳门、台湾）提供。本报告中的信息均来源于合规渠道，华安证券研究所力求准确、可靠，但对这些信息的准确性及完整性均不做任何保证。在任何情况下，本报告中的信息或表达的意见均不构成对任何人的投资建议。在任何情况下，本公司、本公司员工或者关联机构不承诺投资者一定获利，不与投资者分享投资收益，也不对任何人因使用本报告中的任何内容所引致的任何损失负任何责任。投资者务必注意，其据此做出的任何投资决策与本公司、本公司员工或者关联机构无关。华安证券及其所属关联机构可能会持有报告中提到的公司所发行的证券并进行交易，还可能为这些公司提供投资银行服务或其他服务。

本报告仅向特定客户传送，未经华安证券研究所书面授权，本研究报告的任何部分均不得以任何方式制作任何形式的拷贝、复印件或复制品，或再次分发给任何其他人，或以任何侵犯本公司版权的其他方式使用。如欲引用或转载本文内容，务必联络华安证券研究所并获得许可，并需注明出处为华安证券研究所，且不得对本文进行有悖原意的引用和删改。如未经本公司授权，私自转载或者转发本报告，所引起的一切后果及法律责任由私自转载或转发者承担。本公司并保留追究其法律责任的权利。

投资评级说明

以本报告发布之日起 6 个月内，证券（或行业指数）相对于同期相关证券市场代表性指数的涨跌幅作为基准，A 股以沪深 300 指数为基准；新三板市场以三板成指（针对协议转让标的）或三板做市指数（针对做市转让标的）为基准；香港市场以恒生指数为基准；美国市场以纳斯达克指数或标普 500 指数为基准。定义如下：

行业评级体系

- 增持—未来 6 个月的投资收益率领先市场基准指数 5%以上；
- 中性—未来 6 个月的投资收益率与市场基准指数的变动幅度相差-5%至 5%；
- 减持—未来 6 个月的投资收益率落后市场基准指数 5%以上；

公司评级体系

- 买入—未来 6-12 个月的投资收益率领先市场基准指数 15%以上；
- 增持—未来 6-12 个月的投资收益率领先市场基准指数 5%至 15%；
- 中性—未来 6-12 个月的投资收益率与市场基准指数的变动幅度相差-5%至 5%；
- 减持—未来 6-12 个月的投资收益率落后市场基准指数 5%至 15%；
- 卖出—未来 6-12 个月的投资收益率落后市场基准指数 15%以上；
- 无评级—因无法获取必要的资料，或者公司面临无法预见结果的重大不确定性事件，或者其他原因，致使无法给出明确的投资评级。