

潮起潮落，拐点已过，AIGC有望引领人工智能商业化浪潮

——人工智能行业深度

投资要点

□ ChatGPT发布以来迅速出圈，微软加大投资力度并宣布将AIGC应用到产品服务中，彰显AIGC潜在商业价值。我们认为人工智能在成本、算力消耗、应用生态等层面已走过拐点，商业化应用条件加速成熟，看好AI在各细分应用场景下的渗透率加速提升，建议围绕算法模型、业务场景和AI基础设施三大路线选择优质投资标的。

□ ChatGPT火爆出圈，生成式AI商业化应用走向爆发前夕

1、我们综合以下几点，认为AIGC已走过商业化拐点，有望打开市场空间：

(1) 内容生成成本大幅下降满足下游需求；(2) 生成效率大幅提升；(3) 模型计算算力消耗下降，未来有望在消费级硬件上实现AI应用；(4) 生态流量快速积累，ChatGPT月活数量过亿，为应用商业化创造条件。

2、AIGC在文本、图片、音视频生成等领域有望超越人类，经测算至2030年AIGC市场空间超2000亿，建议关注三条投资主线：(1) AI底层模型技术优势厂商；(2) 实际场景AI商业化应用优势厂商；(3) AI基础设施提供商。

3、ChatGPT发布仅两月月活用户达到1亿，系统融合RLHF模型，借鉴人类反馈不断强化性能（现已具备基础的数学计算能力），未来有望通过实际场景数据的输入持续提升模型的准确度，并延伸适用场景。

□ 推理-知识-学习-创造，AIGC有望引领第四次人工智能浪潮

1、梳理AI发展史上三次浪潮，AIGC有望引领第四次浪潮，以大规模商业化应用为核心特征：

(1) **第一次潮起**：实现了AI的推理（计算）能力；**潮落**：对神经元模拟程度较低，无法解决复杂问题（异或问题）；

(2) **第二次潮起**：AI拥有知识储备，可回答特定领域问题；**潮落**：现代计算机的出现；**代表产物**：专家系统；**意义**：AI从科研场景逐步渗透到商业场景；

(3) **第三次潮起**：深度神经网络赋能AI拥有学习能力，且部分场景下的问题解决能力超越人类；**代表产物**：AlphaGo，GPT大模型；**意义**：AI模型的准确度、计算效率、泛化性等指标大幅优化，奠定了“大模型+场景小模型”的技术路线，为商业化应用打下基础；

(4) **第四次潮起**：AI拥有创造能力，有望成为效率乃至生产力工具；**代表产物**：ChatGPT；**意义**：AI在文本、图像、音视频生成领域接近商业化需求，国内外科技大厂技术布局有望打开商业化应用空间。

2、数据+模型算法+基础设施，AI商业化呈现三大趋势：

(1) 趋势一：AI模型与场景深度融合，需要对场景中的复杂数据提取和处理，对非结构化数据的处理（提取、标注、清洗等）将决定AI模型的性能；

(2) 趋势二：“大模型+小模型”框架成为主流，AI模型通过算法迭代，实现计算效率和工作成本的下降，以满足商业化应用的要求；

(3) 趋势三：AI商业化应用将大幅拉动算力需求，AI服务器、AI芯片等细分行业规模有望加速成长

□ 对标海外生态，AIGC有望推动AI商业化下沉各细分场景

行业评级：看好(维持)

分析师：程兵
执业证书号：S1230522020002
chengbing01@stocke.com.cn

相关报告

1 《海天瑞声（688787）：国内AI训练数据龙头提供商，自动驾驶业务打开成长空间》

2023.01.04

2 《拓尔思（300229）：语义智能领军者，大数据+AI双轮驱动打开公司广阔成长空间》

2022.11.29

- 1、国外科技大厂依托自有业务生态和科研实力，推动 AI 性能进步同时加速布局商业化应用。微软宣布将 ChatGPT 融入新版搜索系统和浏览器，加速 AI 商业化进程；谷歌、Meta 不断提升 AI 模型能力，实现图文和短视频的智能生成，未来有望融入到自有业务生态；
- 2、国外垂类公司在文本写作、图像生成、音频生成等领域打造商用产品，收费模式以订阅制为主，业务场景覆盖在线办公、移动社交、广告营销等场景，为国内 AIGC 市场发展提供重要参考；
- 3、国内科技大厂以百度为代表，以中文 AI 模型为核心已经构建起较为完整的人工智能技术体系，并在智能会话、图像生成、气象预报等领域有成熟应用，而国内垂类厂商持续探索 AIGC 在医疗、办公、营销、娱乐等场景的商业化产品，未来有望受益于国内丰富的内容生态加速人工智能商业化进程。

□ 围绕 AI 商业化三大主线挖掘核心投资标的

- 1、关注具备底层算法模型核心技术优势的厂商：
 - (1) 推荐标的：拓尔思（中文 NLP 龙头厂商），科大讯飞（智能语音处理及合成）；
 - (2) 建议关注：谷歌（DeepMind），微软（OpenAI, ChatGPT），Meta（OPT 模型），百度（“文心”模型），腾讯，浪潮信息（“源”大模型）；
- 2、关注各细分赛道下兼具场景理解与 AI 布局优势的垂类厂商：
 - (1) 推荐标的：海康威视（智能物联领域龙头）；
 - (2) 建议关注：云从科技（智慧城市），格灵深瞳（智慧金融），金山办公（办公），万兴科技（AI 绘画），商汤（智慧安防）；
- 3、围绕 AI 数据、算力等基础设施选择优质投资标的：
 - (1) 推荐标的：海天瑞声（国内 AI 训练数据龙头提供商）；
 - (2) 英伟达（GPU），寒武纪（AI 芯片）；

□ 风险提示

- 1、AI 技术迭不及预期的风险；
- 2、AI 商业化产品发布不及预期；
- 3、政策不确定性带来的风险；
- 4、下游市场不确定性带来的风险；

正文目录

| | |
|--|-----------|
| 1 ChatGPT 火爆出圈，生成式 AI 商业化应用走向爆发前夕 | 6 |
| 1.1 AIGC 成本大幅下降，人工智能商业化进程拐点已现 | 6 |
| 1.2 应用场景丰富，AI 商业化空间前景广阔，建议关注三条投资主线 | 7 |
| 1.3 AIGC 应用有望从 B 端延伸至 C 端市场，空间广阔规模超千亿 | 9 |
| 1.4 ChatGPT 面世即成顶流，性能持续提升商业化价值不断放大 | 10 |
| 1.5 ChatGPT 以 GPT+RLHF 模型为核心支撑，为 AIGC 发展指明方向 | 11 |
| 2 推理-知识-学习-创造，AIGC 有望引领第四次人工智能浪潮 | 13 |
| 2.1 复盘三次人工智能浪潮，AI 从科研向细分产业生态加速渗透 | 13 |
| 2.1.1 第一次 AI 浪潮，对人类神经元机制的模拟实现机器推理能力 | 13 |
| 2.1.2 第二次 AI 浪潮，人工智能拥有知识储备，专家系统为代表性产物 | 14 |
| 2.1.3 第三次 AI 浪潮，数据爆炸背景下，AI 进入学习时代 | 17 |
| 2.1.4 第四次 AI 浪潮，AIGC 从文本拓展至图像及音视频领域，商业化空间打开 | 20 |
| 2.2 未来 AI 发展的两大路线，工程方法+科学方法将交替占据行业主流 | 22 |
| 2.3 未来发展方向预计，数据+模型+算力供应仍为三大发展方向 | 23 |
| 2.3.1 趋势一：AI 发展带动基础数据市场成长，高质量标注数据决定模型训练效果 | 23 |
| 2.3.2 趋势二：场景模型持续简化以提升训练速度，控制成本满足商业化应用需求 | 24 |
| 2.3.3 趋势三：AI 应用落地拉动算力需求，AI 基础设施市场规模有望加速成长 | 25 |
| 3 对标海外生态，AIGC 有望推动 AI 商业化下沉各细分场景 | 26 |
| 3.1 海外科技大厂主导，各细分场景玩家涌现，商业模式雏形已现 | 26 |
| 3.1.1 微软、谷歌等大厂科研实力雄厚，推动 AIGC 核心技术不断进步 | 26 |
| 3.1.2 细分赛道涌现大量初创公司，不断探索 AI 商业化应用的产品形态及服务模式 | 27 |
| 3.2 国内科技大厂相继入局，基于内容生态市场空间广阔 | 30 |
| 3.2.1 科技大厂持续开拓中文 AIGC，借助业务生态有望推出爆款应用产品 | 30 |
| 3.2.2 国内细分赛道涌现大量垂类公司，关注实际业务场景下的 AI 应用渗透 | 33 |
| 3.3 国内内容生态市场空间广阔，政策支持下看好 AI 应用加速渗透 | 34 |
| 4 围绕 AI 商业化三大主线挖掘核心投资标的 | 35 |
| 4.1 关注具备底层算法模型核心技术优势的厂商 | 35 |
| 4.1.1 拓尔思，国内 NLP 龙头厂商 | 35 |
| 4.1.2 科大讯飞，AI+行业应用领域领跑者 | 35 |
| 4.1.3 其它建议关注标的 | 36 |
| 4.2 关注各细分赛道下兼具场景理解与 AI 布局优势的垂类厂商 | 36 |
| 4.2.1 海康威视：起步安防行业的智能物联龙头 | 36 |
| 4.2.2 其它建议关注标的 | 36 |
| 4.3 围绕 AI 数据、算力等基础设施选择优质投资标的 | 37 |
| 4.3.1 海天瑞声：AI 训练数据标注优质提供商 | 37 |
| 4.3.2 其它建议关注标的 | 37 |
| 5 风险提示 | 37 |

图表目录

| | |
|---|----|
| 图 1: Jasper 文本写作功能收费模式 | 6 |
| 图 2: OpenAI DALL-E2 模型图片生成功能收费模式 | 6 |
| 图 3: 中国人工智能市场规模及增速 (单位: 亿元) | 7 |
| 图 4: 全球人工智能产业规模发展趋势 (单位: 亿美元) | 7 |
| 图 5: 人工智能应用场景发展 | 8 |
| 图 6: ChatGPT 实现交互问答 | 10 |
| 图 7: ChatGPT 日活用户数爆发式增长 | 10 |
| 图 8: Chat GPT 具备简单的计算能力 (上:更新前, 下:更新后) | 10 |
| 图 9: 微软新版搜索引擎可根据问题内容提供完整解答及方案建议 | 11 |
| 图 10: Transformer 核心框架 | 12 |
| 图 11: RHFL 模型核心框架 | 12 |
| 图 12: 冯·诺依曼结构模拟人类记忆存储的模式 | 14 |
| 图 13: Rosenblatt 感知机模型 | 14 |
| 图 14: 多层感知机 (MLP) 模型 | 14 |
| 图 15: 反向传播 (BP) 模型 | 14 |
| 图 16: XCON 专家系统核心架构 | 16 |
| 图 17: Hopfield Network 模型 | 16 |
| 图 18: 卷积神经网络 (CNN) 模型 | 16 |
| 图 19: 1986-2007 年全球数据量变化趋势 (单位: EB) | 17 |
| 图 20: 大数据相关技术快速迭代 | 17 |
| 图 21: AlexNet 神经网络模型 | 18 |
| 图 22: AlexNet 实现深度学习在图像识别领域的重大突破 | 18 |
| 图 23: AlphaGo 包含价值网络和策略网络两个深度神经网络模型 | 18 |
| 图 24: 生成式对抗网络 (GAN) 模型原理 | 20 |
| 图 25: GAN 模型可实现图像风格迁移和内容修复功能 | 20 |
| 图 26: 几种图像生成模型特点对比 | 21 |
| 图 27: CLIP 模型框架 | 21 |
| 图 28: Make-A-Video 实现 AI 自动生成短视频 | 21 |
| 图 29: AI 训练数据智能标注服务实现路径 | 24 |
| 图 30: 我国 AI 芯片市场规模趋势 | 25 |
| 图 31: 我国云计算市场规模趋势 | 25 |
| 图 32: DALL-E2 模型根据“宇航员骑在马背上”描述生成图像 | 27 |
| 图 33: Meta Make-a-video 模型生成短视频 | 27 |
| 图 34: DeepMind 利用 AIGC 实现蛋白质结构预测 | 27 |
| 图 35: 人工智能 Cicero 在策略游戏《外交》中可击败人类 | 27 |
| 图 36: Jasper AI 产品六大优势 | 28 |
| 图 37: Notion AI 产品服务矩阵 | 28 |
| 图 38: Stability AI 智能生成图像 | 29 |
| 图 39: Dream Studio 平台收费模式 (单位: 0.01 英镑) | 29 |
| 图 40: Midjourney 图片生成工具 | 29 |
| 图 41: Midjourney 作品《太空歌剧院》 | 29 |
| 图 42: Play.ht 乔布斯声音通过 AI 智能生成采访音频 | 30 |

| | |
|---|----|
| 图 43: Play.ht 平台收费模式 | 30 |
| 图 44: 百度“文心”AI 大模型 | 32 |
| 图 45: 浪潮“源”AI 大模型 | 32 |
| 图 46: 华为“盘古”AI 大模型 | 32 |
| 图 47: 我国虚拟人整体市场规模趋势预测 | 34 |
| 图 48: 2022-2027 年我国短视频市场规模预测 (单位: 亿元) | 34 |
| | |
| 表 1: Colossal-AI 深度学习系统大幅降低 AIGC 对 GPU 显存需求 | 7 |
| 表 2: AIGC 应用场景广泛 | 8 |
| 表 3: AIGC 国内理论市场空间测算 | 9 |
| 表 4: 历次人工智能浪潮及代表性成就梳理 | 13 |
| 表 5: 部分专家系统及应用场景梳理 | 15 |
| 表 6: GPT, GPT-2, GPT-3 模型对比 | 19 |
| 表 7: AI 图像生成模型发展梳理 | 20 |
| 表 8: AI 视频生成模型发展梳理 | 22 |
| 表 9: 符号主义和连接主义对比 | 22 |
| 表 10: AI 神经网络模型参数量及复杂程度 | 23 |
| 表 11: AI 模型对人类脑神经处理信息机制的模拟 | 23 |
| 表 12: 海外科技大厂 AIGC 领域布局及成果 | 26 |
| 表 13: AIGC 领域相关初创公司及业务场景梳理 | 28 |
| 表 14: 国内科技大厂 AIGC 布局 | 30 |
| 表 15: 百度基于文心大模型构建完整业务生态 | 31 |
| 表 16: 国内细分赛道企业在 AIGC 各应用场景持续布局 | 33 |
| 表 17: 国家政策大力支持人工智能商业化落地 | 34 |

1 ChatGPT 火爆出圈，生成式 AI 商业化应用走向爆发前夕

1.1 AIGC 成本大幅下降，人工智能商业化进程拐点已现

我们认为以 ChatGPT 为代表的 AIGC 兴起，在内容创作成本、创作效率、模型计算消耗、用户流量基础等维度实现了重大突破，有望推动 AI 商业化进程的大幅加速。

(1) AI 内容创作成本大幅降低且耗时更短，相较传统方式优势显著

目前 OpenAI 定价最高的文字模型达芬奇（基于 GPT-3）为每 750 词约 0.02 美元（折合约 0.14 元人民币）；而主打“AI 生成文案”的独角兽 Jasper 以类 SaaS 服务形式收费，根据官网显示，每月生成 10 万字的价格约 82 美元，折合每 1000 字约 5.57 元人民币，而阅文集团 2021 年内容成本为 17.74 亿元（每 750 字 37 元）；

在图像生成领域，Open AI 透露其 DALL-E2 模型的图片智能编辑及生成服务的价格，超过免费额度数量的图片收费为每 460 张图片 15 美元，折合每张图片约 0.22 元人民币。目前国内外 AIGC 绘画创作平均耗时已达到分秒级，创作效率较人工优势明显。

(2) AIGC 模型算力消耗快速下降，落地门槛降低有望驱动消费级应用出现

根据量子位报告显示，借助最新的 Colossal-AI 通用深度学习系统，AIGC 领域的 Stable Diffusion 模型训练过程中可最高节省约 5.6 倍的显存需求（从 64.5GB 大幅降低到 11.6GB），未来有望采用消费级 GTX 3060 显卡实现模型运算功能，大幅降低 AI 商业化应用的落地门槛。

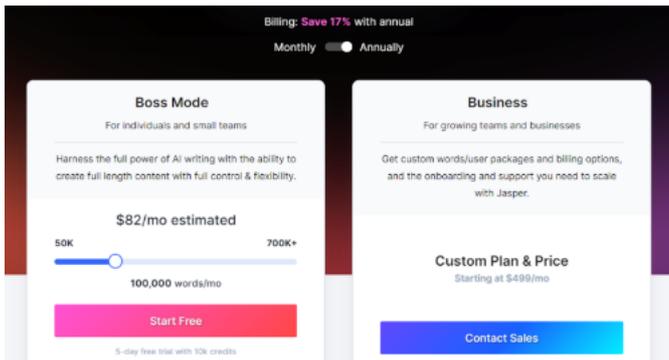
(3) 爆款应用快速积累用户流量，商业化条件已基本具备

ChatGPT 发布后日活用户数量快速突破千万。根据 ARK Invest 的测算，自 2022 年 11 月底至 2023 年 1 月下旬，不足两个月的时间内 ChatGPT 的日活数量已突破 1200 万，据腾讯网报道，2023 年 1 月 ChatGPT 月活用户数量达 1 亿，成为史上增长最快的消费者应用。使用者数量的爆炸式成长即反应出用户的尝试使用意愿强烈，未来嵌入到微软的业务生态中也有望提升用户的工作效率，优化使用体验，创造更大的商业价值。

(4) 国内外公司持续探索商业化路径，AIGC 生态有望加速形成

近日，OpenAI 推出 ChatGPT 付费订阅版 ChatGPT Plus，每月收费 20 美元，而包含 Synthesia、Jasper、runway 等 AIGC 初创公司相继推出细分领域的收费服务，未来 AIGC 产业生态有望加速构建，并形成更加清晰的 AI 商业化应用路径。

图1: Jasper 文本写作功能收费模式



资料来源: Jasper AI 官网, 浙商证券研究所

图2: OpenAI DALL-E2 模型图片生成功能收费模式

| | Cost | Credits | # of images* |
|-------------------------|------|---------|--------------|
| First Month | FREE | 50 | 200 |
| Every Month After | FREE | 15 | 60 |
| Additional credit packs | \$15 | 115 | 460 |

资料来源: The Decoder, 浙商证券研究所

表1: Colossal-AI 深度学习系统大幅降低 AIGC 对 GPU 显存需求

| 模型 | 运算策略 | GPU | GPU 数量 | Flash | Batch size | GPU RAM (GB) |
|---------------------|-------------|------|--------|-------|------------|--------------|
| Stable Diffusion v1 | DDP | A100 | 4 | False | 64 | 64.5 |
| Stable Diffusion v2 | | | | True | | 31.9 |
| Stable Diffusion v1 | Colossal-AI | | | False | | 30.0 |
| Stable Diffusion v2 | | | | True | | 11.6 |

资料来源: 量子位, 浙商证券研究所整理

1.2 应用场景丰富, AI 商业化空间前景广阔, 建议关注三条投资主线

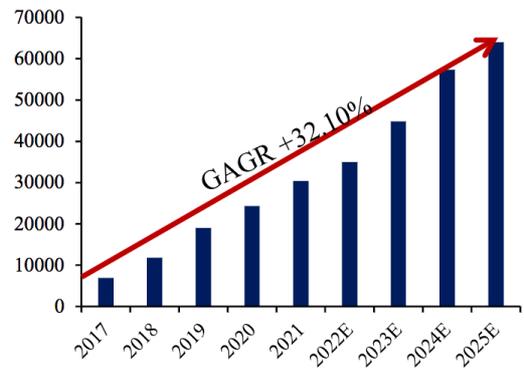
AIGC 推动人工智能商业化进程加速, 有望打开千亿市场。德勤数据显示, 2021 年中国人工智能市场规模达 2058 亿元, 预计到 2025 年将达到 5460 亿元, 2021-2025 年复合增长率约 27.63%。根据德勤预测, 全球人工智能产业规模预计从 2017 年的 6900 亿美元增长至 2025 年的 6.4 万亿美元, 复合增长率达 32.10%。移动互联网时代带来的海量数据、模型和算力的不断迭代, 以及各类应用场景中的商业化尝试, 为 AI 的商业化奠定了坚实基础, 未来有望加速释放人工智能产业动能。

图3: 中国人工智能市场规模及增速 (单位: 亿元)



资料来源: 德勤, 浙商证券研究所

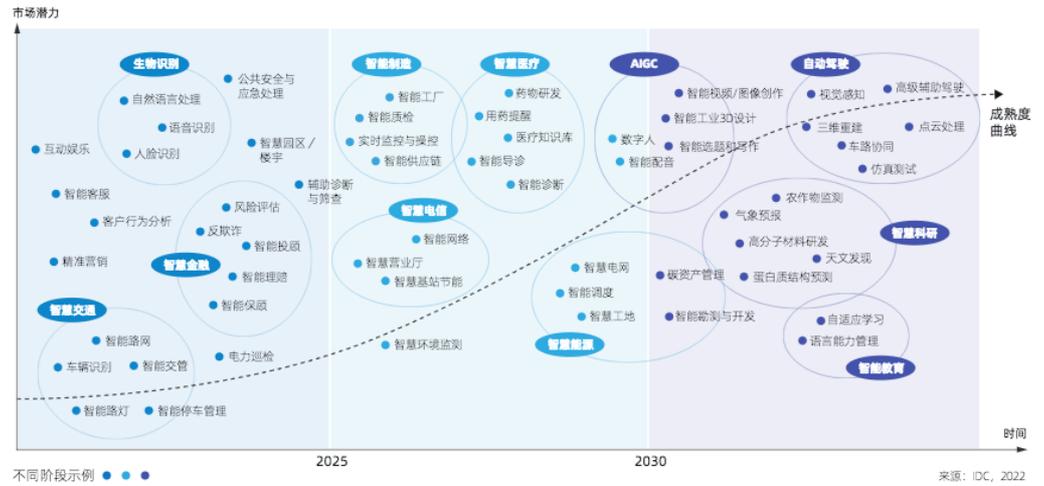
图4: 全球人工智能产业规模发展趋势 (单位: 亿美元)



资料来源: 德勤, 浙商证券研究所

人工智能商业化进程有望加速, 建议重点关注三类公司选择投资标的。《20220-2023 中国人工智能计算力发展评估报告》中指出, 未来随着自然语言处理、计算机视觉、AIGC 等细分技术的持续迭代, AI 应用将呈现场景应用深度和广度并行发展态势, 以 AIGC 为代表的生成式 AI 将在创作型工作 (文本、图像视频创作) 和预测、决策型工作中扮演更加重要角色。我们建议围绕人工智能商业化的三大主线选择投资标的: (1) 具备 AI 模型、算法技术优势的科技公司; (2) 具备对 AI 商业化实际应用场景理解和业务优势的厂商; (3) 具备 AI 硬件优势的厂商;

图5：人工智能应用场景发展



资料来源：《2022-2023 中国人工智能算力发展评估报告》，浙商证券研究所

(1) 具备 AI 模型、算法研究实力优势的科技公司

我们梳理了人工智能发展史上出现的三次高潮（具体内容见本文第二章内容），总结出人工智能发展遵循了“计算—知识—学习”三个阶段，在 AIGC 的推动下有望进入第四次高潮，即“创造”阶段。人工智能从最早的科研主导（第一次、第二次 AI 浪潮），发展到科技大厂参与研究（第三次 AI 浪潮），再到各细分场景公司深度参与 AI 商业化（未来），模型和算法的迭代始终是核心主线，因此我们认为具备 AI 模型和算法的科研实力的科技大厂未来有望在 AI 商业化浪潮中保持核心竞争优势。

(2) 具备对 AI 商业化实际应用场景理解和业务优势的厂商

AIGC 推动人工智能在各类细分场景中的商业化尝试，《AIGC 人工智能生产内容行业研究报告》指出，未来 AIGC 有望与传媒、电商、影视、金融等各行业深度融合，创造更多的应用场景，大幅提高各行业的内容创作效率。

表2：AIGC 应用场景广泛

| AIGC 与产业融合 | 场景描述 |
|------------|-------------------------------------|
| AIGC+传媒 | 写稿机器人、采访助手、视频字幕生成、语音播报、视频剪辑、AI 合成主播 |
| AIGC+电商 | 商品 3D 模型、虚拟主播、虚拟货场 |
| AIGC+影视 | 剧本创作、合成人脸和声音、AI 创作角色和场景、自动生成影视预告片 |
| AIGC+娱乐 | AI 换脸、AI 作曲、AI 合成音视频动画 |
| AIGC+教育 | 虚拟教师、课本内容制作、2D 课本转换为 3D |
| AIGC+金融 | AI 实现金融资讯、产品介绍视频内容自动生产、AI 虚拟人客服 |
| AIGC+医疗 | AI 为失声者合成语言音频、为残疾人合成肢体投影、心理疾病患者陪护 |
| AIGC+工业 | AI 完成工程设计、衍生设计、辅助工程师决策 |

资料来源：《AIGC 人工智能生产内容行业研究报告》，浙商证券研究所

在未来 AI 商业化进程中，众多 AIGC 开源模型和算法平台的存在，导致在场景应用中，单纯的技术和算法很难成为玩家的竞争壁垒。在细分场景中，AIGC 企业需要在业务场景的深度理解、AI 赋能的一体化解决方案、行业深度绑定、业务闭环等领域持续提升竞争力。因此我们建议关注各细分赛道中，现有业务具备竞争优势，且具备 AI 转型实力（场景数据储备+AI 算法技术实力）的公司。

（3）具备 AI 硬件优势的厂商

AI 商业化对硬件要求仍然很高，算力作为数据加速处理的动力源泉，其重要性不言而喻。AI 处理器芯片可以支持深度神经网络的学习和加速计算，相比于 GPU 和 CPU 拥有成倍的性能提升，和极低的耗电水平。因此，我们认为围绕人工智能的硬件厂商也将有望受益 AI 商业化的发展。

1.3 AIGC 应用有望从 B 端延伸至 C 端市场，空间广阔规模超两千亿

对标海外 AIGC 厂商商业模式进行测算，国内 2030 年 AIGC 市场空间超两千亿。根据 Gartner 《2021 年预测：人工智能对人类和社会的影响》，到 2025 年 AIGC 产生的数据将占有所有数据的 10%。我们对标海外 AIGC 厂商，如 Jasper.AI、Stability.AI 等公司的商业模式，测算国内 2023 年、2025 年和 2030 年 AIGC 市场空间。

文字、图片渗透率有望快速提升，视频、直播等受限于技术迭代渗透率较慢。目前 AI 在文本和图像生成领域技术已且相对成熟，同时微软、百度等科技大厂有望将 AI 技术应用到业务生态中，有望推动在线办公、搜索引擎等应用场景的渗透率提升。而在直播、影视以及音视频等领域，由于 AI 生成技术处于初级阶段，且下游用户对内容要求较高，我们认为短期内渗透率提升幅度较低。经测算，我们预计到 2025 年国内 AIGC 市场空间可达 403.52 亿元，到 2030 年市场空间可达 2175.58 亿元，未来几年市场有望迎来爆发式成长。

表3：AIGC 国内理论市场空间测算

| 内容类型 | 用户对象 | 用户规模 (万) | ARPU 假设 (元/年) | AI 渗透率 (2023 年) | AIGC 市场 空间 (亿元) | AI 渗透率 (2025 年) | AIGC 市场 空间 (亿元) | AI 渗透率 (2030 年) | AIGC 市场 空间 (亿元) |
|------|--------|-------------|------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
| 游戏 | 内容制作 | 200 | 3000 | 1% | 0.60 | 5% | 3.00 | 10% | 6.00 |
| 网络文学 | 创作者 | 2130 | 400 | 1% | 0.85 | 10% | 8.52 | 40% | 34.08 |
| 音视频 | 创作者 | 35000 | 2000 | 0% | 0 | 1% | 70.00 | 10% | 700.00 |
| 直播 | 主播 | 14000 | 2000 | 0% | 0 | 1% | 28.00 | 10% | 280.00 |
| 影视 | 从业者 | 100 | 3000 | 1% | 0.30 | 8% | 2.40 | 15% | 4.50 |
| 图片 | 社交平台用户 | 90000 | 150 | 1% | 13.50 | 10% | 135.00 | 30% | 405.00 |
| 内容营销 | 从业人员 | 1000 | 1500 | 1% | 1.50 | 10% | 15.00 | 30% | 45.00 |
| 在线办公 | 用户 | 46000 | 200 | 1% | 9.20 | 8% | 73.60 | 30% | 276.00 |
| 搜索引擎 | 用户 | 85000 | 100 | 1% | 8.50 | 8% | 68.00 | 50% | 425.00 |
| 合计 | | — | | | 34.45 | — | 403.52 | — | 2175.58 |

资料来源：前瞻产业研究院，观研天下，艾瑞咨询，Jasper AI，The decoder，Play.ht，浙商证券研究所

1.4 ChatGPT 面世即成顶流，性能持续提升商业化价值不断放大

ChatGPT 火爆出圈，可与真人流畅交流，标志生成式 AI 的突破性进展。2022 年 11 月 30 日，OpenAI 发布了对话式 AI 新模型 ChatGPT，该模型的功能包含了基于人机交互的问题回答、文本写作以及代码编写等。ChatGPT 模型使用 Transformer 神经网络架构，通过维基百科以及海量真实语料库的数据训练，实现了接近人类的交流水平。

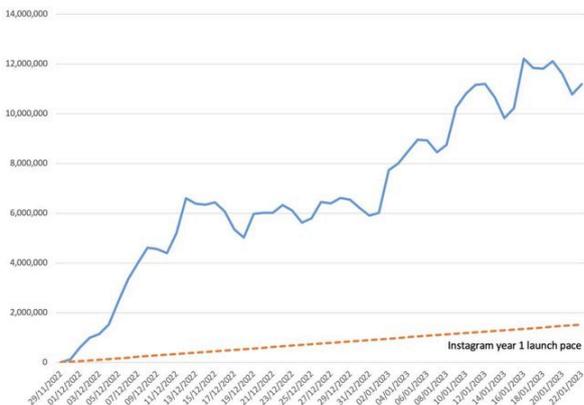
图6：ChatGPT 实现交互问答



资料来源：CSDN，浙商证券研究所

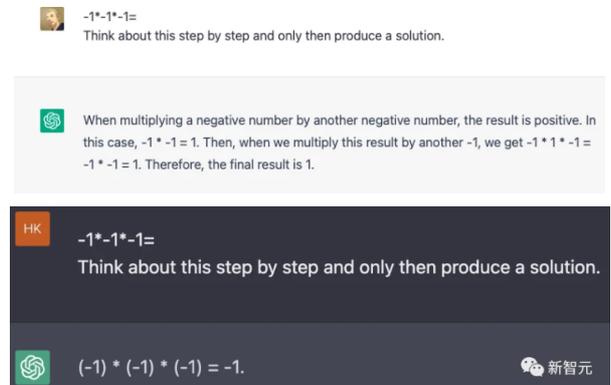
ChatGPT 具备自我学习能力，面世以来性能仍在不断进步。2023 年 1 月 30 日，OpenAI 官方宣布 ChatGPT 通过新一轮的更新，在准确性和真实性上实现显著提升，并且已经具备了比较基础的数学计算能力。经过使用者的测试，目前 ChatGPT 的数学计算能力接近小学生水平，虽然相较其代码和文本写作相比差距甚远，但进步非常明显，ChatGPT 在数学领域的计算能力逐渐形成反映出模型的自我学习能力以及模型迁移能力出色。

图7：ChatGPT 日活用户数爆发式增长



资料来源：ARK Invest，浙商证券研究所

图8：Chat GPT 具备简单的计算能力（上:更新前，下:更新后）

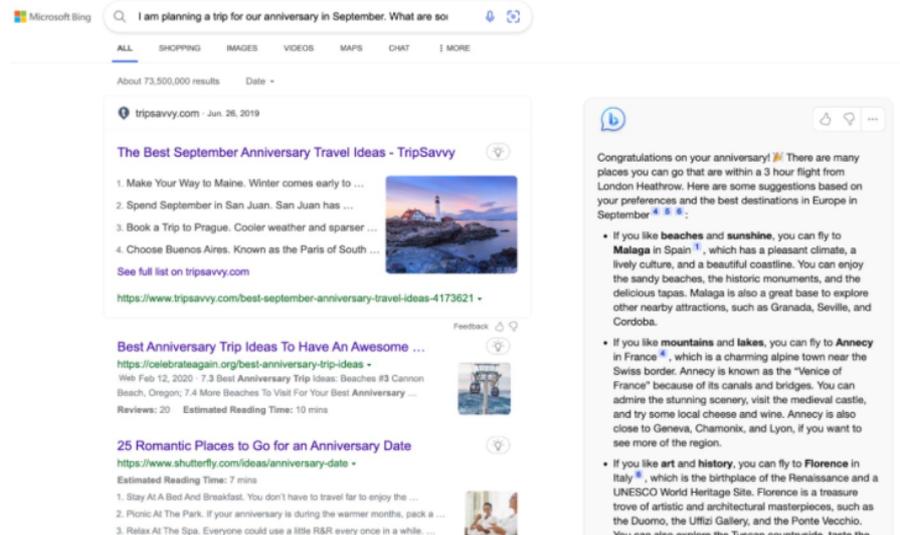


资料来源：腾讯新闻，浙商证券研究所

微软计划推出一系列 AI 应用服务，ChatGPT 有望在实际应用中持续成长。2023 年 2 月 2 日，OpenAI 公司宣布推出付费试点订阅计划 ChatGPT Plus，定价每月 20 美元。付费版功能包括高峰时段免排队、快速响应以及优先获得新功能和改进等。同时，OpenAI 方面仍将提供对 ChatGPT 的免费访问权限。

新版搜索引擎发布，融合 OpenAI 核心技术。根据钛媒体报道，2 月 8 日微软宣布推出集成 ChatGPT 的全新 Bing 搜索服务，以及集成 AI 辅助的 Edge 浏览器。新版 Bing 带有一个扩展的聊天框，它现在可以做的不仅仅是回答事实问题和为你提供各种链接，在 ChatGPT 的帮助下，它还能够为你即时生成各种个性化的规划、建议、分析等，解决更复杂的搜索问题。

图9：微软新版搜索引擎可根据问题内容提供完整解答及方案建议



资料来源：钛媒体，浙商证券研究所

1.5 ChatGPT 以 GPT+RLHF 模型为核心支撑，为 AIGC 发展指明方向

ChatGPT 将海量训练数据与 Transformer 框架结合，在 GPT 模型的基础上通过 RLHF 模型提升交互聊天能力，实现了对自然语言的深度建模，Transformer 有望在未来较长时间内成为 AIGC 语言领域的核心技术框架。

Transformer 架构为自然语言理解领域带来颠覆性变革，应用到 NLP、CV、生物、化学等多领域。2017 年，谷歌跳出 RNN 和 CNN 的结构，提出了完全基于 Self-Attention 机制的 Transformer 架构，当时在机器翻译任务上取得了颠覆性的进步。Transformer 由一个编码器和一个解码器组成，且各自由若干个编码/解码模块堆叠而成，每个模块包含 Multi-Head Attention 层、全连接层等功能不同的工作层。

Transformer 架构优势突出，计算效率、并行度、解释性均实现突破。相比于之前的框架 Transformer 架构除了在计算的时间复杂度更优外，还有 3 大优势：

(1) 可直接计算点乘结果，并行度高：对于数据序列 x_1, x_2, \dots, x_n ，self-attention 可以直接计算任意两节点的点乘结果，而 RNN 必须按照顺序从 x_1 计算到 x_n 。

2 推理-知识-学习-创造，AIGC 有望引领第四次人工智能浪潮

2.1 复盘三次人工智能浪潮，AI 从科研向细分产业生态加速渗透

复盘历史上的三次人工智能浪潮，AIGC 有望引领第四次浪潮。我们从数据、模型迭代和参与者等维度梳理了过往三次人工智能浪潮。我们认为，历次浪潮可以总结为 AI 从推理能力形成，向知识储备，再向自我学习能力的方向持续进步，未来有望进入到自我创造的阶段。

表4：历次人工智能浪潮及代表性成就梳理

| | 开始时间 | 数据输入 | 模型迭代 | 代表性成就 | 主要参与者 |
|-----|---------|--------------------|--|---|--------------------------|
| 第一次 | 1950s | 文字、数学符号 | Rosenblatt 提出感知机模型 | 图灵提出图灵机和“图灵测试”； | 科研院所 |
| 第二次 | 1980s | 数据库、表格等结构化数据 | Hopfield 提出 RNN 模型雏形； Hinton 等人提出 Boltzman Machine 模型； 多层感知机 MLP 与反向传播 BP 的提出； | 首个专家系统诞生，可解决特定领域的部分问题，进行了商业化尝试的提出； | 科研院所主导，大型企业商用尝试（IBM、波音等） |
| 第三次 | 2006 | 大量已标注的结构化和少量非结构化数据 | 1、卷积神经网络 CNN； 2、残差网络 ResNet； 3、注意力机制 Attention； | GPT、BERT 等 AI 大模型发布； AlphaGo 战胜李世石，表明 AI 某些领域的学习能力已超越人类； | 科技大厂深度参与（微软、谷歌、阿里、百度等） |
| 第四次 | 2023-未来 | 未来以非结构化数据为主 | GPT-3、CLIP、Diffusion 等模型进一步提升 AI 生成问题和跨模态处理任务的能力； | ChatGPT | 科技大厂主导；各细分行业生态玩家有望深度参与； |

资料来源：知乎，浙商证券研究所

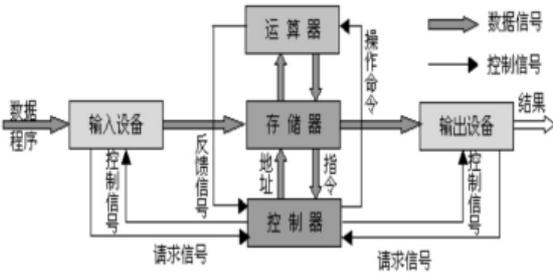
2.1.1 第一次 AI 浪潮，对人类神经元机制的模拟实现机器推理能力

计算机的诞生催生学术界对人工智能的大规模探索，第一次 AI 浪潮来袭。1946 年第一代电子计算机诞生，在这一阶段，冯·诺依曼结构为现代计算机的体系架构奠定基础，而图灵机的思想论证了现代计算机的计算模式和计算能力，开启了人类在现代人工智能领域不断探索的旅程。

冯·诺依曼模拟人类好大脑记忆存储与提取的工作机制，为计算机和人工智能奠定坚实基础。上世纪 40 年代，冯·诺依曼提出了计算机的逻辑结构，其主要特点包含了程序以二进制代码存放在存储器中、所有指令由操作码和地址码组成、指令在存储过程中按照执行顺序进行存储等。冯·诺依曼结构将人类的神经系统与计算机结合在一起，大幅提升了计算机的运算效率，为人工智能的发展提供了保障。

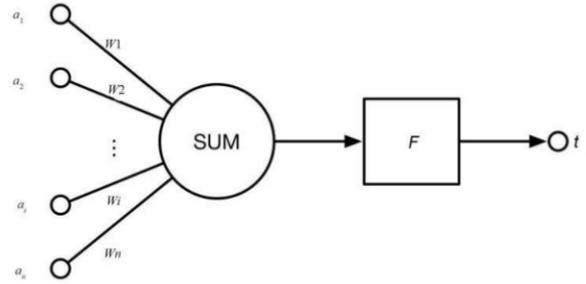
Rosenblatts 受生物神经网络启发，提出人工神经网络结构，成为现代神经网络和深度学习的重要基础。感知机的工作原理可理解为生物神经网络中的信号作用，信号经过树突传递到细胞核的过程中信号会发生变化。感知机在模型的“输入”位置添加神经元节点，构成“输入单元”，向每一个属性指定一个权重 w ，对属性值和权重的乘积求和，将结果值与阈值比较，从而判定正负样本结果。但由于感知机模型只能解决线性问题，面对异或问题时无能为力，也导致了第一次人工智能研究在上世纪 70 年代进入低潮期。

图12：冯·诺依曼结构模拟人类记忆存储的模式



资料来源：百度百科，浙商证券研究所

图13：Rosenblatt 感知机模型



资料来源：CSDN，浙商证券研究所

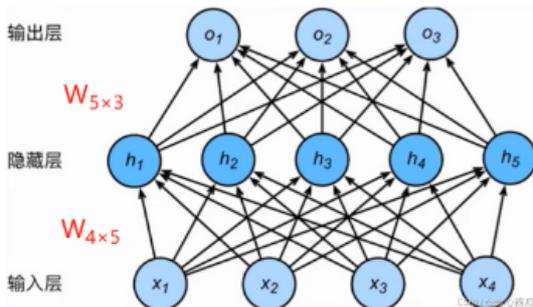
2.1.2 第二次 AI 浪潮，人工智能拥有知识储备，专家系统为代表性产物

20 世纪 80 年代，第二次人工智能浪潮到来，核心发展为让计算机学习大量专业知识，研究人员将专家的知识、言论以及经验等数据输入到计算机，使其成为专家系统。专家系统可以预测在一定条件下某种解的概率，由于当时计算机已有巨大容量，专家系统有可能从数据中得出规律。在这一阶段，多层感知机、反向传播算法、神经网络的出现极大程度地提升了计算机的计算及逻辑推理能力，为深度学习和强化学习模型奠定了重要基础，同时专家系统的出现赋予了人工智能知识属性，并进行了一系列的商用尝试。

多层感知机强调模拟人类脑神经的复杂连接，使人工智能实现解决非线性问题的处理能力。多层感知机可至少分成输入层、隐藏层和输出层这三层，隐藏层可根据需要建多层且每层都可以有多个节点，相邻层的各个节点都互相连接。隐藏层和输出层具备计算加权和激活函数处理的功能，实现数据信息的向前传递和分析。多层感知机的突破在于激活函数的使用，在隐藏层中使用不同的激活函数可实现对数据的非线性化处理，使得 AI 具备拟合任何连续函数的能力，大幅提升了计算机的计算能力。

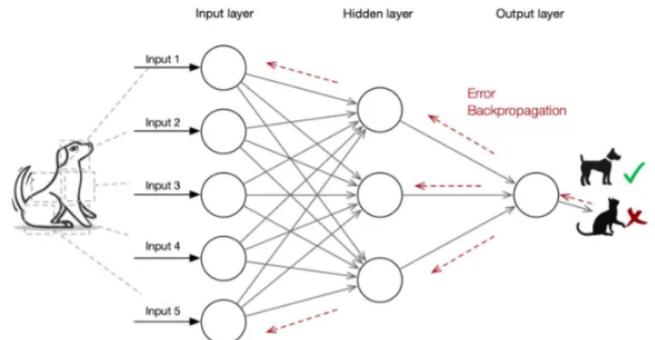
1986 年，Geoffrey Hinton 等人先后提出了多层感知机 (MLP) 与反向传播 (BP) 训练相结合的理念，将人类基于结果误差反馈反哺逻辑推理的思想融入 AI。BP 算法的基本思想是用误差的导数 (梯度) 调整，并通过误差的梯度做反向传播，更新模型权重，以下降学习的误差，拟合学习目标。

图14：多层感知机 (MLP) 模型



资料来源：CSDN，浙商证券研究所

图15：反向传播 (BP) 模型



资料来源：CSDN，浙商证券研究所

第二次 AI 浪潮中，专家系统模拟人类专家的知识 and 经验解决特定领域的问题，实现了人工智能从理论研究走向实际应用、从一般推理策略探讨转向运用专门知识的重大突破。并且机器学习（特别是神经网络）探索不同的学习策略和各种学习方法，在大量的实际应用中也开始慢慢复苏。

专家系统作为人工智能的重要分支，在这一时期逐步渗透到细分行业的实际场景解决特定任务。第二次 AI 浪潮期间，专家系统应用到数学、物理、化学、医学、地质等细分行业，能够辅助科研人员解决化学分子结构、疾病诊断、地质分析等特定任务。这一时期专家系统的特点包含：（1）单学科的专业性；（2）系统结构完整，移植性好；（3）具有一定的推理解释功能，透明性好；（4）采用启发式推理和不精确推理；（5）用产生式规则、框架和语义网络表达知识；

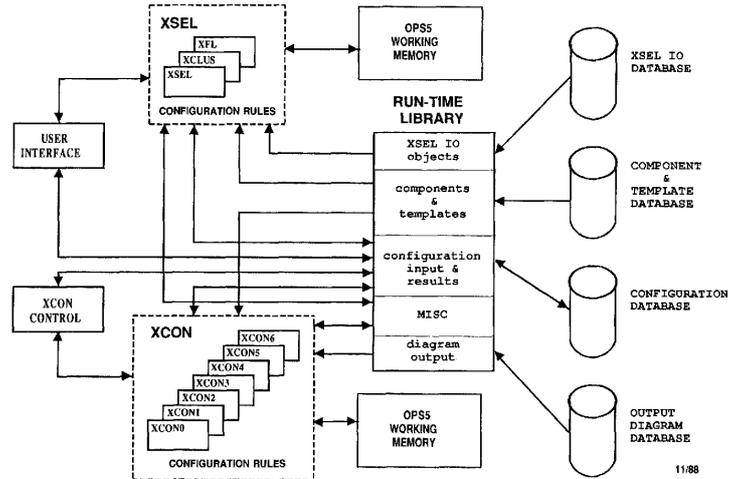
表5：部分专家系统及应用场景梳理

| 专家系统 | 时间 | 开发者 | 核心亮点及应用场景 |
|------------|--------|----------------|--|
| DENDRAL | 1968 年 | 斯坦福大学 | 可推断化学分子结构； |
| MYCSYMA | 1971 年 | 麻省理工学院 | 用于数学运算； |
| PROSPECTOR | 1977 年 | 斯坦福大学 | 用于地质领域探测矿藏，可分析待检矿床的地址环境、结构和矿物质类型等特征； |
| CASNET | 1977 年 | 拉特格大学 | 用于青光眼的诊断与治疗，采用因果联系网络表示知识法则； |
| MYCIN | 1978 年 | 斯坦福大学 | 可实现血液感染病诊断，拥有两百多条规则，可识别 51 种病菌并正确处理 23 种抗生素； |
| AM | 1981 年 | 斯坦福大学 | 模拟人类进行概括、抽象和归纳推理，并发现某些数论的概念和定理； |
| HEARSAY | 1970s | 卡耐基梅隆大学 | 用于语音识别，证明了计算机在理论上可以按照编制的程序与用户进行交谈； |
| XCON | 1980 年 | 卡耐基梅隆大学、DEC 公司 | 具有完整专业知识和经验的计算机智能系统，可为 VAX 计算机系统制定硬件配置方案； |

资料来源：CSDN，浙商证券研究所

专家系统商业价值显现，在企业生产决策过程中发挥重大作用。1980 年，美国卡耐基梅隆大学与 DEC 公司合作研发的 XCON 专家系统，运用计算机系统配置的知识，依据用户的定货，选出最合适的系统部件（如 CPU 型号、操作系统种类及相应型号，存储器和外部设备等），并且能够给出优化意见以构成一个完整的系统。XCON 可以给出一个系统配置的清单和这些部件装配关系的图，以便技术人员进行装配。在这一时期，除 DEC 公司外，美国杜邦、通用汽车和波音等公司在也在生产决策中借助专家系统的辅助，1986 年美国 AI 相关软硬件销售额达到 4.25 亿美元，人工智能技术逐渐走出科研院所，进入到企业的实际业务场景中。

图16: XCON 专家系统核心架构

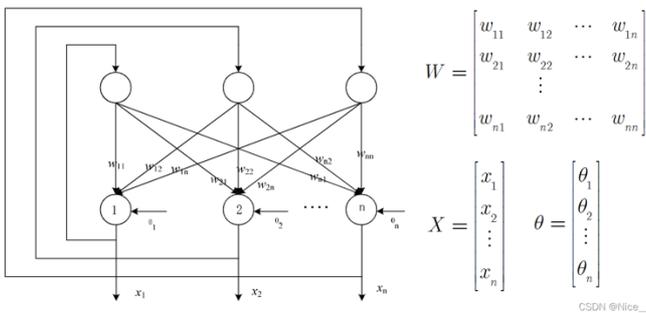


资料来源:《Expert systems for configuration at Digital: XCON and beyond》, 浙商证券研究所

AI 模型算法研究领域, 对人类脑神经的信号处理和思考机制的模拟进一步加深, 代表性产物包含 CNN、RNN 等神经网络模型。1982 年, John Hopfield 提出了一种结合存储系统和二元系统的神经网络, 即 Hopfield 网络。这种模型保证了向局部极小的收敛, 由于每个神经元的输出均连接到其它神经元的输入, 模型能够通过训练从某一残缺的信息联想到所属的完整信息, 即, Hopfield 模型在模拟人类记忆储存和联想方面做出重大贡献, 这一模型也勾勒了未来 RNN 模型的雏形。

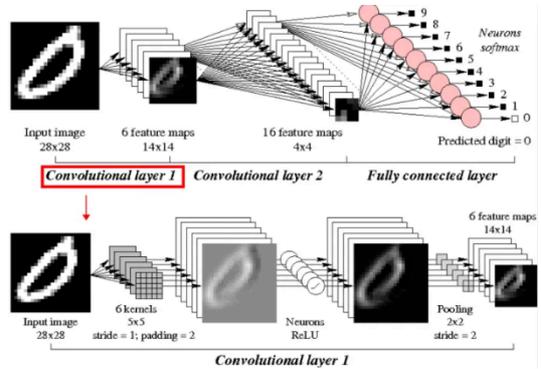
Yann LeCun 提出 CNN 模型, 并成功验证其商用的可行性。1989 年, LeCun 结合反向传播算法与权值共享的卷积神经层发明了卷积神经网络 (Convolutional Neural Network, CNN), CNN 通常由输入层、卷积层、池化 (Pooling) 层和全连接层组成。卷积层负责提取图像中的局部特征, 池化层用来大幅降低参数量级, 全连接层则类似传统神经网络部分输出想要的结果。卷积神经网络 (CNN) 成功应用到美国邮局的手写字符识别系统中, 在此基础上于 1998 年诞生了稳定可商业应用的 CNN 模型 LeNet-5。

图17: Hopfield Network 模型



资料来源: CSDN, 浙商证券研究所

图18: 卷积神经网络 (CNN) 模型



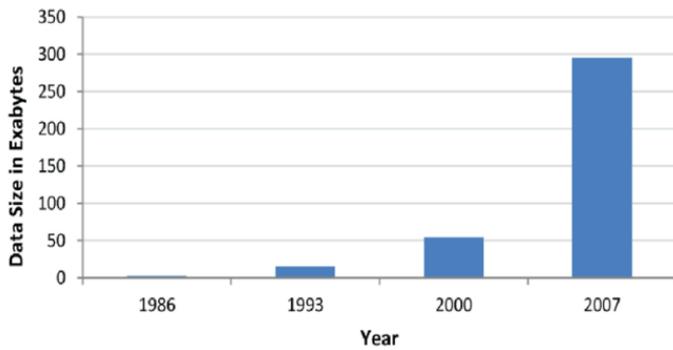
资料来源: 知乎, 浙商证券研究所

2.1.3 第三次 AI 浪潮，数据爆炸背景下，AI 进入学习时代

人类进入互联网时代，数据量爆炸以及大数据技术的成熟，为人工智能的发展注入新动能。PC 互联网以及移动互联网时代的到来驱动了人类活动数据量的爆炸式成长，为 AI 模型的训练提供有价值的原料，同时围绕大数据、云计算等领域的技术迭代使得人工智能模型的计算效率、成本等均有大幅优化，为商业化应用打开了空间。

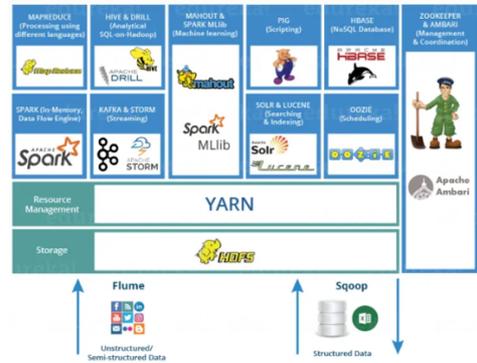
根据海外学者论文《Modeling and Simulation Strategies for Performance Evaluation of Cloud Computing Systems》，进入互联网时代后，全球产生的数据量呈现爆炸式成长，于 2000 年突破 50EB，到 2007 年已接近 300EB。在针对大规模的数据计算领域，2003 年，Google 公布了 3 篇大数据奠基性论文，为大数据存储及分布式处理的核心问题提供了思路：非结构化文件分布式存储（GFS）、分布式计算（MapReduce）及结构化数据存储（BigTable），并奠定了现代大数据技术的理论基础。

图19：1986-2007 年全球数据量变化趋势（单位：EB）



资料来源：《Modeling and Simulation Strategies for Performance Evaluation of Cloud Computing Systems》，浙商证券研究所

图20：大数据相关技术快速迭代

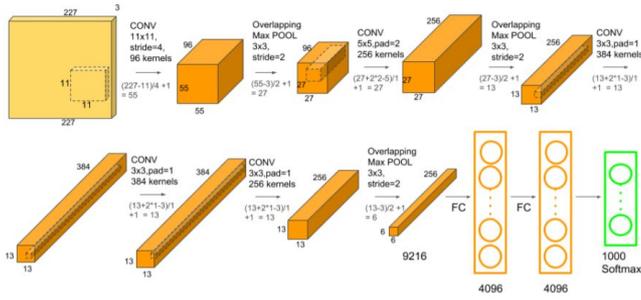


资料来源：知乎，浙商证券研究所

Geoffrey Hinton 等人提出深度学习（Deep Learning）概念，掀起第三次人工智能热潮。深度学习的本质是使用多个隐藏层网络结构，通过大量的向量计算，学习数据内蕴信息的高阶表示。相较于传统的神经网络，深度学习模型具有更优异的特征学习能力，学习得到的特征对数据有更本质的刻画，有利于可视化或分类；其次，深度神经网络在训练上的难度，可以通过“逐层初始化”（layer-wise pre-training）来有效克服，因此模型精调的效率得以大大提升。深度学习的出现标志着计算机实现了基于数据的学习和挖掘分析能力，AI 能够挖掘复杂信息中包含的关联关系。

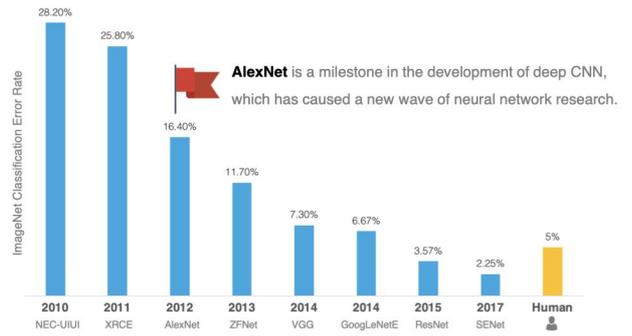
2012 年 AlexNet 神经网络获得 ImageNet 大赛冠军，标志着 AI 在图像识别领域的突破性成就。2012 年，在计算机视觉领域的竞赛 Image Net 中，AlexNet 以提升 10% 的错误率的进步力压第二名以 SIFT+FV、LBP+FV、GIST+FV、CSIFT+FV 等组合特征的算法。AlexNet 基于卷积神经网络 CNN 模型，在数据、算法及算力层面均有较大改进，创新地应用了 Data Augmentation、ReLU、Dropout 和 LRN 等方法，并使用 GPU 加速网络训练。之后图像识别领域 AI 模型均在此基础上进行优化，并且识别准确度超越了人类水平。

图21: AlexNet 神经网络模型



资料来源: 知乎, 浙商证券研究所

图22: AlexNet 实现深度学习在图像识别领域的重大突破

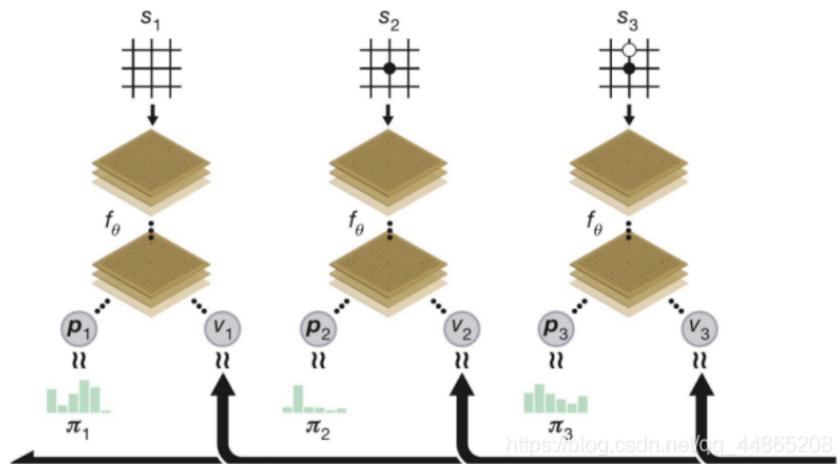


资料来源: 知乎, 浙商证券研究所

AlphaGo 具备深度学习能力, 可结合人类专家比赛中学到的棋谱, 以及在和自己下棋 (Self-Play) 中进行强化学习。2016 年, 人工智能 AlphaGo 在围棋竞技中击败李世石, 标志着 AI 具备了自我迭代和学习强化的能力。AlphaGo 的深度学习模型可分为策略网络和价值网络, 分别解决了围棋中的“下一手最佳策略”和“下一手后的胜率”两个问题, 并以此为核心通过学习人类棋谱和自我对弈, 不断调整策略提升棋力。AlphaGo 较好地模拟了人类棋手对于围棋落子的直觉判断, 即, 由于围棋存在的广度 (每一步的可能性) 和深度 (棋局从开始到结束的步数) 导致计算机无法模拟计算所有的可能性, AlphaGo 通过策略网络减少搜索广度, 即每一步搜索少量节点, 而通过价值网络减少搜索深度, 无需将棋局下到底即可得出当前局势的判断。

在 AlphaGo 基础上, 谷歌公司后续推出了迭代版本 AlphaGo Zero, 增强了 AI 在面对人类棋谱中未出现过的局势时的自我推理和学习能力, 标志着 AI 的自我学习具备了一定的迁移能力。

图23: AlphaGo 包含价值网络和策略网络两个深度神经网络模型



资料来源: CSDN, 浙商证券研究所

Transformer 架构横空出世, Open AI 发布的 GPT 模型, AI 进入“大模型”时代。GPT (Generative Pre-Training, GPT) 模型利用更大规模的文本数据以及更深层的神经网络

模型学习更丰富的文本语义表示，打破了自然语言处理各个任务之间的壁垒，使得搭建一个面向特定任务的自然语言模型不再需要非常多的任务北京，**只需要根据任务的输入和输出形式应用预训练模型便可达到满意效果**。NLP 模型的搭建只需要“无监督预训练+有监督的精调”两阶段：

(1) 生成式预训练：在大规模文本数据上训练一个高容量的语言模型，以学习更丰富的上下文信息；

(2) 判别式任务精调：将预训练好的模型适配到下游任务中，并使用有标注数据学习判别式任务；

从 GPT 到 GPT-2 和 GPT-3，参数提升和训练数据积累提升模型工作性能。到了 GPT-2 阶段，Open AI 将自然语言模型定义为要完成零样本无监督多任务学习的目标，即预训练好的模型可以直接用于下游任务。GPT-2 模型在学习更广泛的数据信息，增加了模型的维度参数后，不需要针对特定的任务修改模型，将任务也作为提示词，和文本一起输入预训练好的模型。

承接 GPT-2 的核心思想，到了 GPT-3，模型参数量达到了 1750 亿，在不进行微调的情况下可以在多个 NLP 基准上达到最先进的性能。GPT 模型的演进为国内外 AI 大厂提供了重要的借鉴意义，即通过优质训练数据输入和模型复杂度的增加来提升模型性能，而未来 AI 以少样本乃至无样本学习为核心目标。

表6：GPT, GPT-2, GPT-3 模型对比

| | GPT-1 | GPT-2 | GPT-3 |
|-------|--|-------------------------------------|---|
| 训练数据集 | BooksCorpus: 7000 本未发表的书籍 | WebText: 清洗过的 Reddit 数据, 40GB 文本数据 | 5 种语料库: Common Crawl、WebText2、Books1、Books2、Wikipedia |
| 模型框架 | Transformer decoder, 每层由 masked self attention + feed forward layer 组成 | | |
| 模型参数 | 12 层, 768 维, 1.17 亿参数 | 48 层, 1600 维, 15 亿参数 | 96 层, 12888 维, 1750 亿参数 |
| 模型表现 | 在问答、阅读理解、文本总结等任务上表现出色 | 参数增多实现原有任务的表现进一步优化 | 在少样本学习甚至零样本学习任务上取得出色表现 |
| 目标函数 | 无监督的语言模型、有监督的精调分类模型 | 引入任务条件函数, 可根据输入和任务预测输出, 实现少样本到零样本学习 | |

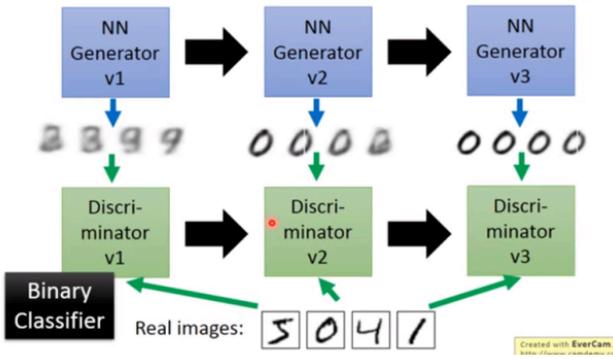
资料来源：知乎，浙商证券研究所整理

AI 大模型在任务准确度和场景泛化性方面提升显著，在此基础上针对不同实际场景提炼出针对性的 AI 小模型以完成特定任务，可大大降低 AI 商业化部署的成本，提高训练速度。其中，Transformer 架构和 GAN 模型为现在的语言文本生成和图像生成奠定了基础。

GAN 模型参考博弈学习理论，通过内部两个模型的互相博弈，实现图像视频的 AI 自动生成。生成式对抗网络（GAN, Generative Adversarial Networks）是近年来复杂分布上无监督学习最具前景的方法之一。模型通过框架中两个模块：生成模型（Generative Model）和判别模型（Discriminative Model）的互相博弈学习产生输出。判别模型负责判断一个实例是真实的还是模型生成；而生成模型负责生成实例来骗过判别模型。两个模型对抗中最达到平衡，即生成模型生成的实例与真实的没有区别，判别模型无法区分输入数据是生成的还是原始真实的数据。

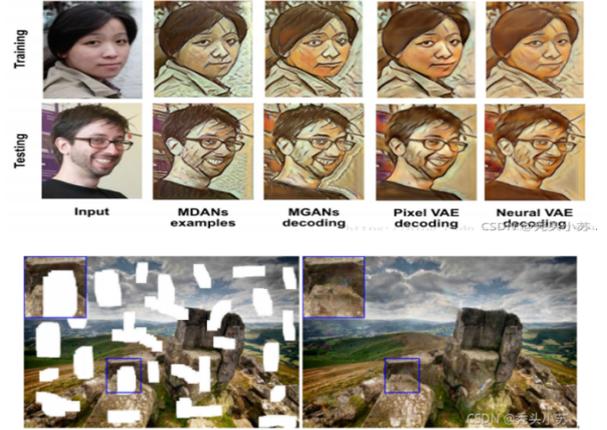
GAN 模型应用广泛，成为 AIGC 重要底层框架但早期存在问题。GAN 模型应用包含逼真图片生成、图片翻译、3D 建模等，但由于基础的 GAN 模型存在收敛不稳定、容易受到部分图片样本欺骗等问题，初期生成的图像可能存在失真问题，适用范围相对局限。

图24：生成式对抗网络（GAN）模型原理



资料来源：ELEMENT UI，浙商证券研究所

图25：GAN 模型可实现图像风格迁移和内容修复功能



资料来源：CSDN，浙商证券研究所

2.1.4 第四次 AI 浪潮，AIGC 从文本拓展至图像及音视频领域，商业化空间打开

AIGC 在图片、音视频领域的模型在大模型基础上持续优化，性能趋向商用化需求，空间有望打开。我们认为以 ChatGPT 为代表，AI 将进入创造（创作）时代，引领第四次人工智能浪潮。AIGC 以自动化生产和高效为两大特点，目前已经可以自动生成文字、图片、音频、视频，甚至 3D 模型和代码。未来大量数字原生内容有望由 AI 协助完成创作。

底层技术的突破使 AIGC 商业落地成为可能，模型迭代推动 AI 生成文字和图片质量实现质的提升。在图像生成领域

表7：AI 图像生成模型发展梳理

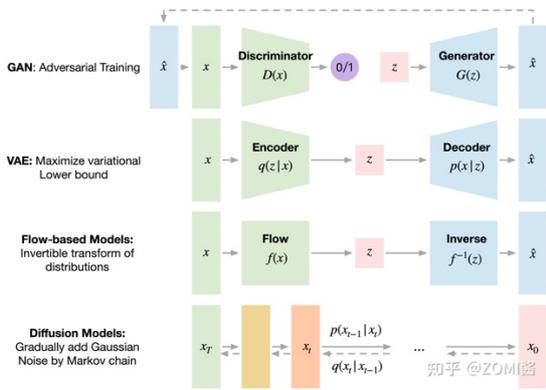
| AI 图像模型 | 出现时间 | 特点 |
|------------------|-------------|---------------------------------------|
| GAN | 2014 年 | 不稳定、分辨率低、只能模仿，无法原创 |
| CAN | 2017 年 | 表达抽象、能够原创 |
| DALL-E | 2021 年 2 月 | 可文字转绘画，但绘画水平较普通 |
| CLIP+VQGAN | 2021 年 4 月 | 可文字转绘画、开源、绘画水平一般 |
| CLIP+Diffusion | 2022 年 2 月 | 可文字转绘画、开源、原创性、但渲染时间长 |
| Midjourney | 2022 年 3 月 | 可文字转绘画、付费、适合人像、突出细节 |
| DALL-E2 | 2022 年 4 月 | 可文字转绘画、付费、限制较多、对复杂文字理解准确、艺术性不高、渲染快 |
| Stable Diffusion | 2022 年 8 月 | 可文字转绘画、免费开源、原创性、灵活度高、图片精美、真实感、艺术性、渲染快 |
| Imagen (Goggle) | 2022 年 11 月 | 有限公开，理论上比 DALL-E2 效果好 |

资料来源：CSDN，浙商证券研究所

Diffusion 模型参考热力学原理，大幅提高生成图像的稳定性、准确性以及计算效率。扩散模型（Diffusion Model）的灵感来自非平衡热力学，模型训练过程中缓慢地将随机噪声添加到数据中，然后通过反转这个噪声过程，来学习恢复数据。不同于 GAN 模型，扩散模型不依赖于生成与对抗两个模型的博弈，因此学习结果更加稳定，并且模型的并行性和扩展性更好，之后推出的 Stable Diffusion 模型在计算效率上显著提升，满足企业级和消费级的使用需求。

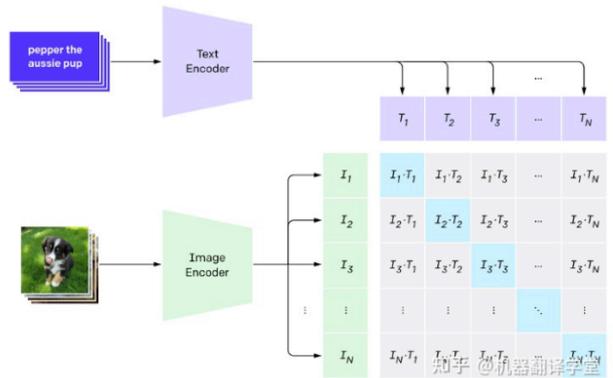
CLIP 模型打破了传统分类器界限，可以在不需要额外训练集的情况下，以良好的性能完成各种计算机视觉任务。CLIP 模型对问题进行了重新定义并不再预测文本标签，取而代之的是预测图像与文本相匹配的可能性大小，在大规模数据集上完成的训练后能够学习到了图像的各方面信息。

图26：几种图像生成模型特点对比



资料来源：知乎，浙商证券研究所

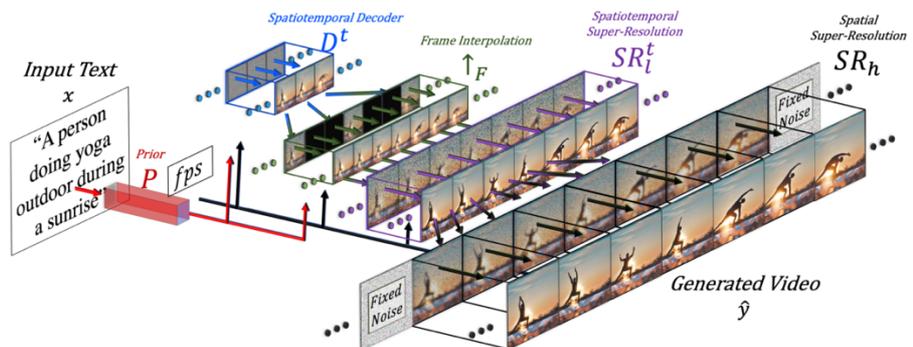
图27：CLIP 模型框架



资料来源：知乎，浙商证券研究所

Make-A-Video 模型实现 AI 基于文本、图像智能生成视频。在视频生成领域，Meta 公司继承了图像生成模型的核心技术，在文本语义和图像对应关系的基础上，通过对视频帧序列的学习训练，实现连贯运动且高分辨率的高质量视频生成。由于模型训练量要求过大，目前该模型只能实现几秒钟的短视频生成，未来有望随着模型的迭代实现中视频和长视频的生成。

图28：Make-A-Video 实现 AI 自动生成短视频



资料来源：CSDN，浙商证券研究所

表8：AI 视频生成模型发展梳理

| AI 视频生成模型 | 出现时间 | 特点 |
|--------------------|-------------|--------------------|
| Morpheus | 2021 年 10 月 | 输入文字即可生成 3D 和动画 |
| Make-A-Video | 2022 年 9 月 | 文字转视频、图片转视频、视频生成视频 |
| Imagen Video | 2022 年 10 月 | 文字转视频、视频质量高 |
| Phenaki | 2022 年 10 月 | 文字转视频、情节连贯 |
| Stability AI Video | 研发中 | 尚未公布 |

资料来源：CSDN，浙商证券研究所

2.2 未来 AI 发展的两大路线，工程方法+科学方法将交替占据行业主流

人工智能的发展进程中形成符号主义和连接主义两大阵营，从工程维度和科学维度优化迭代 AI 模型。在 AI 发展早期，由于数据量较少，AI 模型容易用符号模型表征，因此符号主义发展较快，而后来数据量提升，问题和任务的复杂度增加，导致基于复杂模型结构的连接主义兴起，并形成了目前一系列 AI 范式。

我们认为在 AIGC 主导的人工智能商业化进程中，符号主义和连接主义仍将作为 AI 迭代的两个重要方向。其中，AI 在细分场景中的应用将推动大量时序性的复杂信息输入到模型当中，因此对场景数据（尤其是非标准化的数据）的处理有望迎来重大发展；其次，目前 AI 在长视频、虚拟人、高清图像等领域的智能生成仍存在局限性，我们看好通过模型的迭代和对人类神经网络的研究和模拟，提升 AI 的准确度、计算效率等指标。

表9：符号主义和连接主义对比

| | 符号主义 | 连接主义 |
|---------|---------------------------|---------------------|
| 核心理念 | 主张用公理和逻辑体系搭建 AI 系统 | 主张模拟人类神经元连接机制实现 AI |
| 对计算机的认知 | 视作操作数学物理符号的计算系统 | 视作模拟人脑的认知工具 |
| 研究目标 | 提高系统解决问题和任务的能力 | 提高系统自我学习的知识延伸的能力 |
| 代表性成就 | 专家系统 | 神经网络 |
| 优势 | 系统具备可验证性与可解释性，操作便捷 | 可处理杂乱的非结构化数据，抗噪性能更好 |
| 劣势 | 无法模拟人脑的形象思维，对连续复杂信息的处理能力差 | 黑箱，无法理解模型内部运作过程 |

资料来源：百度百科，维基百科，浙商证券研究所

计算机算力以及 AI 在细分场景下的渗透率提升推动 AI 模型训练数据量爆炸式成长，通过增加模型参数及数据量提高 AI 的计算速度和准确度。大模型通过超大规模的参数设置和数据输入，具有基础、通用和综合的特点，而小模型基于实际任务场景的需求，通过高质量场景数据的输入和训练，使 AI 能够出色完成特定任务。在 NLP 领域，以 Google 发布的 BERT 模型和 Open AI 提出的 GPT 模型为典型代表。AI 通过大模型的预训练，能够在处理语言、推理、人机交互等领域任务上的表现实现极大提升，打开了 AI 渗透到各行业进行商业应用路径。

表10：AI 神经网络模型参数量及复杂程度

| 类别 | 模型 | 公司 | 参数量 |
|--------|-------------------------|-----------|--|
| 图像识别 | AlexNet (2012) | | 6000 万 |
| | VGG-16 (2014) | | 1.38 亿 |
| | ResNet-50 (2015) | | 2500 万 |
| 自然语言模型 | GPT (2018-2020) | Open AI | GPT-1 (2018) : 1.17 亿 GPT-2 (2019) : 15.42 亿 GPT-3 (2020) : 约 1750 亿 |
| | BERT (2018) | Google | 3.4 亿 |
| | MT-DNN (2019) | Microsoft | 3.3 亿 |
| | Project Megatron (2019) | NVIDIA | 83 亿 |
| | T-NLG (2020) | Microsoft | 170 亿 |
| | 源 1.0 (2021) | 浪潮信息 | 2457 亿 (中文语言模型) |

资料来源：CSDN，浙商证券研究所

以 **Transformer** 为代表性成就，AI 模型在模拟人脑思维模式中不断进步。Transformer 架构的核心在于注意力机制的使用，这使得 AI 模型在计算过程中可大幅节约消耗在无关信息上的算力，除此之外，残差网络、多巴胺机制的强化学习等模型均参考了脑科学领域的研究成果，通过嵌入响应的信号处理模块，节约算力消耗的同时提升计算效率。

表11：AI 模型对人类脑神经处理信息机制的模拟

| 类别 | 模型 | 参数量 |
|----------------------|----------------------|---|
| 残差网络 ResNet | 模拟人类视觉加工的非连续性分层机制 | 采用神经网络短路连接，大幅提升计算效率； |
| LeNet | 模拟人类大脑的分层结构 | 实现外部数据从低级到高级的特征提取，在图像识别领域效果显著； |
| 基于注意力机制的 Transformer | 模拟人类视觉加工过程中的选择性关注机制 | 有效解决信息超载问题，大幅提升计算过程中的资源分配效率； |
| 基于奖励机制的强化学习 | 模拟人类大脑决策过程中的多巴胺信号机制 | 在神经网络训练中加入奖励预测误差信号，通过奖励机制引导 AI 的快速学习过程； |
| 基于海马体机制的强化学习 | 模拟人类大脑内海马体的空间认知及推理机制 | 实现空间结构和视觉感知解耦，提升 AI 决策反馈效率及准确度； |

资料来源：CSDN，浙商证券研究所

2.3 未来发展方向预计，数据+模型+算力供应仍为三大发展方向

2.3.1 趋势一：AI 发展带动基础数据市场成长，高质量标注数据决定模型训练效果

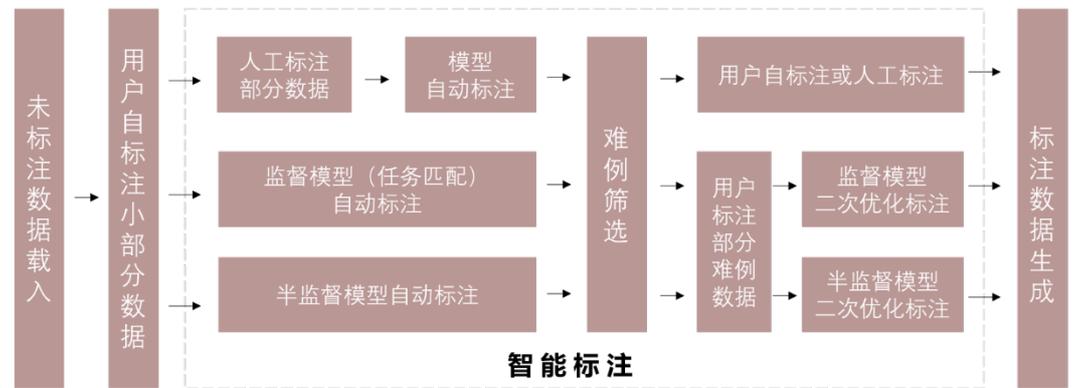
AI 模型发展带动训练数据需求升高，标注和结构化处理真正实现将数据转为为 AI 商业价值。随着 AI 与各个产业结合得愈加紧密，AI 商业化程度进入新阶段，企业对 AI 在商业化落地中的表现要求越来越高。为了保证 AI 算法的识别精度，数据标注的质量也就变得至关重要。

AIGC 向图像、语音、文字多维数据延伸，AI 基础数据服务与治理赛道有望受益加速成长。数据服务是利用数据采集与标注工具处理图片、语音、文本等非结构化数据，数据

治理则使用数据治理的各模块管理多源异构数据，提高数据质量并形成数据资产。高质量的数据可供各行业 AI 厂商训练模型，并支撑各场景商业化产品的落地。

数据标注目前仍以人工为主，行业向智能化和去人工化趋势演进，标注数据生产效率与 AI 模型进步协同共振。不同于计算机视觉领域 AI 模型训练所用的图像数据，NLP 模型训练的数据往往来自于半结构化或非结构化的信息，目前无法实现 AI 算法模型的自动化标注，目前主流的实现路径是人工标注与智能标注协同方案，即人工完成一小部分的数据标注，再借助算法模型对剩余数据进行标注，通过人工对部分难例数据样本的二次标注后，最终形成高质量的标注数据。

图29： AI 训练数据智能标注服务实现路径



资料来源：头豹研究院，浙商证券研究所整理

2.3.2 趋势二：场景模型持续简化以提升训练速度，控制成本满足商业化应用需求

大模型微调成本和维护成本过高。目前大模型厂商在用户需求较大时，微调成本 and 模型维护成本会相当高，若仅提供推理服务，成本可以大幅降低但是模型精度会受到较大牺牲，从而降低大模型的竞争力。并且 AI 项目实施通常会把模型部署到终端，对算力的要求非常高，因此国内外大厂及研究院所均致力于对预训练模型进行“瘦身”。

模型蒸馏与迁移学习成为 AI 领域研究新方向，可大幅优化终端模型计算的时效性并节约算力成本。“知识蒸馏”概念可类比于学校中的教学行为，可有效地压缩预训练模型的大小。知识蒸馏包含已训练的 Teacher Model 和待训练的 Student Model，通过知识蒸馏和迁移学习，学生网络可以拥有与教师网络相似和相近的计算性能。

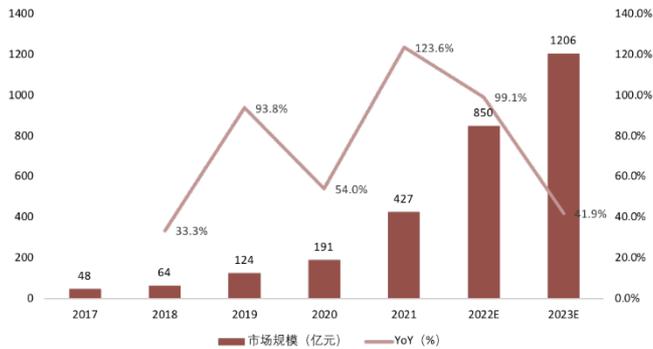
知识蒸馏可以在基于高度复杂的大模型训练基础上，提升细分任务小模型的计算精度，同时也可以有效降低模型延迟，并且压缩网络参数。基于知识蒸馏思想改进 BERT 预训练模型，实现模型的简化和速度的提升。Distilled BERT 模型直接将 BERT 的部分参数作为初始化，模型参数只有 BERT 的约 40%，但速度快提升 60%，同时保留了 97% 的语言理解能力。我们认为在人工智能通往商业化应用的道路上，早期专注于增加数据量、计算能力或者训练过程的优化方式可能不再适用，未来各大厂商需要计算与预测性能之间做出权衡，探索如何利用更少的数据与计算资源，帮助模型实现性能提升。

2.3.3 趋势三：AI 应用落地拉动算力需求，AI 基础设施市场规模有望加速成长

高算力支持是训练 AI 大规模商业化的基础，AI 基础设施市场有望迎来爆发。微软入资 OpenAI 后双方达成多年的合作协议，OpenAI 接入微软的 Azure 云平台开发 AI 技术。高算力的底层基础设施是完成对海量数据处理、训练的基础。我们认为 AI 技术发展逐渐成熟，数字化基础设施不断建设完善，将拉升 AI 芯片、AI 服务器的市场需求。AI 商业化应用的加速落地，将推动我国 AI 基础设施市场规模的加速成长

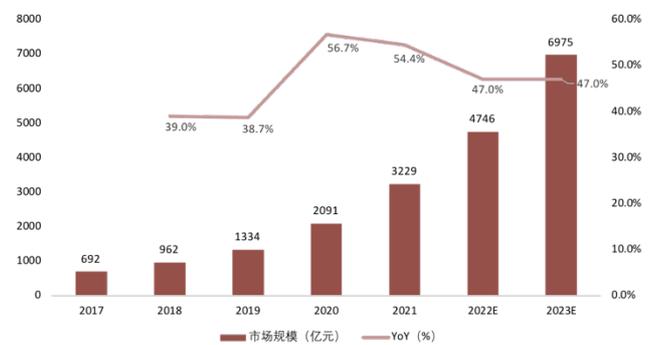
市场规模加速成长，国内细分赛道龙头厂商有望受益。《2022-2023 中国人工智能算力发展评估报告》指出，预计到 2026 年全球 AI 服务器市场规模将达到 347.1 亿美元，五年复合增长率为 17.3%。中商产业研究院数据预计，2023 年我国 AI 芯片以及云计算市场规模将分别达到 1206 亿元和 6975 亿元，同比分别增长 41.9%和 47.0%，AI 基础设施市场规模快速成长。

图30：我国 AI 芯片市场规模趋势



资料来源：中商产业研究院，浙商证券研究所

图31：我国云计算市场规模趋势



资料来源：中商产业研究院，浙商证券研究所

3 对标海外生态，AIGC 有望推动 AI 商业化下沉各细分场景

3.1 海外科技大厂主导，各细分场景玩家涌现，商业模式雏形已现

3.1.1 微软、谷歌等大厂科研实力雄厚，推动 AIGC 核心技术不断进步

海外科技巨头发力布局，自有业务生态及强大科研实力成为核心竞争力。在经历了移动互联网浪潮下科技巨头竞争后，谷歌、Meta、微软等国外科技大厂近年来在 AIGC 领域均有所布局，并在聊天交互、图像生成特定领域取得了一系列突破性进展，同时部分成果已投入商业化应用。

表12：海外科技大厂 AIGC 领域布局及成果

| 企业 | 时间 | 业务进展 | 核心亮点 |
|----------------|---------|--------------------|--|
| 微软 (OpenAI) | 2021.6 | 发布 Copilot | 是微软 Visual Studio Code 中内置的一项新的软件开发人工智能协助服务，功能包括自主生成部分代码以及自行修复部分 Bug。 |
| | 2022.10 | 发布 DALL-E2 | 是一个文本到图像的 AI 生成模型，使用者只需要输入文字描述，AI 图像生成工具就会自动产出其认为符合描述的图像。 |
| | 2022.11 | 发布 ChatGPT | 可自主生成文本，并与用户进行几乎无障碍的沟通交流，未来有望嵌入到微软的消费级和企业级应用服务。 |
| 谷歌 DeepMind | 2022.4 | 发布 Chinchilla 模型 | 拥有 700 亿参数数量的语言模型，在大部分语音任务中击败了 Gopher、GPT-3 等大模型。 |
| | 2022.9 | 提出 Sparrow 模型 | 基于人类反馈的强化学习 (RL) 框架，可与使用者进行简单的交流，回答内容基于谷歌搜索相关信息以保证准确性。 |
| | 2023 | Imagen、Parti (研发中) | 基于 Transformer 架构和 Diffusion 模型，实现文本到图像的 AI 智能生成，在图像合成的保真度和图像-文本一致性方面拥有显著优势。 |
| Meta | 2022.7 | 推出 Make-A-Scene | 基于 AI 驱动文本生成图像，并允许通过文本输入进行有针对性的创作（如输入草图等）。 |
| | 2022.11 | 推出 Make-A-Video | 基于 AI 驱动文本、图片生成短视频，根据输入的自然语言文本可生成一段 5 秒钟左右的短视频。 |
| 英伟达 | 2022.11 | 发布 Magic3D 生成模型 | 可根据文字描述生成 3D 模型，可将低分辨率生成的粗略模型优化为高分辨率的精细模型，生成效率更高，并且计算成本更低。 |

资料来源：零壹财经，知乎，浙商证券研究所

DALL-E2 模型可实现文本到图像的智能生成。2022 年 4 月 OpenAI 发布 DALL-E2 模型，只需输入简短的文字 prompt 就可以生成全新的图像，甚至可以修改现有图像，创建有其显著特征的图像变体。DALL-E2 的训练数据来自互联网上大规模匹配的自然语言-图像数据，因此消除了手动标注数据集的成本瓶颈。

DALL-E2 工作过程中，首先将文本 prompt 输入到经过训练的文本编码器中，接着将文本编码映射到相应的图像编码，该图像编码捕获文本编码中包含的语义信息，最后，图像解码模型生成图像，该图像是该语义信息的视觉表现。

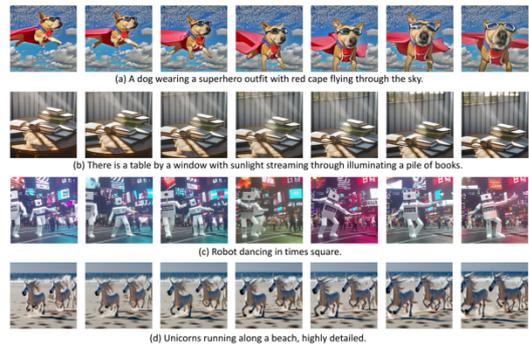
Meta 在已有 AI 图像生成技术上更进一步，实现基于文本、图片生成短视频的 AIGC 模型。Meta AI 将 Make-A-Video 分割成三个组成部分并分开训练，分别是基础的文字生成图像 AI 模型、学习视频中图像在时间维度上的动态变化的 AI 模块，以及超分辨率模型和插帧模型以提高视频画质（分辨率和帧率）。

图32: DALL-E2 模型根据“宇航员骑在马上”描述生成图像



资料来源: 知乎, 浙商证券研究所

图33: Meta Make-a-video 模型生成短视频

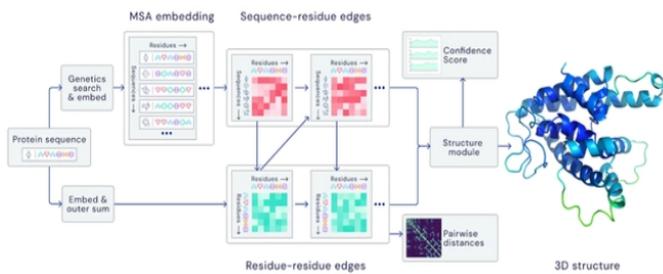


资料来源: CSDN, 浙商证券研究所

科技大厂主导 AIGC 在前沿领域的科学研究, 为人工智能的商业化持续拓展想象空间。谷歌母公司 Alphabet 旗下子公司 Deepmind 开发了 Alphafold 2 模型, 通过预测蛋白质中每对氨基酸之间的距离分布以及化学键之间的角度, 实现对蛋白质结构三维模型的准确预测。Alphafold2 模仿人类注意力的网络架构采用 Attention 机制, 可以同时聚焦多个细节部分, 使得框架预测的结果更加全面和准确。

Meta Cicero 可在人机交互中表现出谈判技巧、同理心等, 未来有望应用到各类社交场景。2022 年 11 月, Meta 发布了具有谈判、说服和合作能力的人工智能 Cicero, 在策略游戏《外交》的测试中, Cicero 可以与其它人类玩家以自然语言沟通, 甚至使用各类谈判技巧以实现游戏目标, 创建伙伴关系和联盟, 最终进入游戏参与者排名的前 10%。Meta 认为 Cicero 的核心技术未来有望创造更加智能的 NPC, 以及在多会话对话中缓解人类与 AI 之间的沟通障碍, 将 AI 应用到更加广泛的社交场景。

图34: DeepMind 利用 AIGC 实现蛋白质结构预测



资料来源: 智东西, 浙商证券研究所

图35: 人工智能 Cicero 在策略游戏《外交》中可击败人类



资料来源: 游研社, 浙商证券研究所

3.1.2 细分赛道涌现大量初创公司, 不断探索 AI 商业化应用的产品形态及服务模式

大模型 API 公布和模型开源催生 AIGC 创业浪潮, 细分赛道内大量玩家在产品及商业模式等方面持续探索, AI 商业化生态有望加速形成。AIGC 有望赋能内容生成领域的成本下降和效率提升, 目前海外公司在文案写作、图像生成、音视频创作等领域持续布局。细分赛道内玩家在相对统一的底层模型基础上, 探索与实际应用场景深度融合的落地产品。

表13: AIGC 领域相关初创公司及业务场景梳理

| 生成内容 | 企业 | 成立时间 | 应用领域 | 底层模型 | 企业简介 |
|------|--------------|------|--------------------|-----------------------|--|
| 文字 | Otherside AI | 2020 | AI 邮件写作 | GPT-3 | 公司主打利用 AI 自动回复邮件，底层技术采用 OpenAI 的 GPT-3 协议，操作非常简单，只要输入邮件内容的关键要点，就可以生成一封完整的邮件。 |
| 文字 | Copy AI | 2020 | AI 广告文案写作 | GPT-3 | Copy AI 通过 AI 写作广告和营销文案，可以帮助用户几秒钟内生成高质量的广告和营销文案，主要面向 To B 商业场景，底层技术采用 OpenAI 的 GPT-3 协议。目前用户包括微软、Ebay。 |
| 文字 | Jasper AI | 2020 | AI 文案写作 | GPT-3 | 公司通过 AI 帮助企业和个人撰写营销推广文案以及博客等各种文字内容，其底层技术采用 OpenAI 的 GPT-3 协议。 |
| 文字 | Notion AI | 2022 | AI 写作、创意提供、语法检查、翻译 | GPT-3 | 目前公司产品功能包括自动撰写文章、广告文案和播客，以及通过头脑风暴为用户提供创意建议 自动检查文字拼写和语法错误、自动翻译文章等。目前 Notion AI 以白名单形式开放 Alpha 版本测试。 |
| 图像 | Stability AI | 2020 | AI 作图底层协议 | CLIP、Stable Diffusion | Stable AI 主要功能包含 AI 生成图片、音频、视频和 3D 建模，公司研发了开源 AI 作图模型 Stable Diffusion，只要输入文字描述就能生成高质量图片。 |
| 图像 | Midjourney | 2022 | AI 作图 | CLIP、Stable Diffusion | 公司主打 AI 绘画工具，对标 Stable Diffusion 以及 DALL-E2。应用产品部署在 Discord 上，输入文字后可以快速生成对应的图片。 |
| 音频 | Play.ht | 2016 | AI 生成语音 | Peregrine | Play.ht 是一个 AI 文本转换语音应用，采用 Peregrine 语音模型，包含数千种说话的声音，可以学习人类的语气、音调和笑声。 |

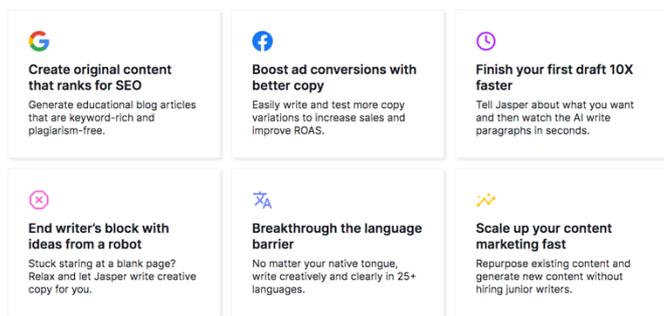
资料来源: CSDN, 知乎, 浙商证券研究所

文本创作领域，基于 GPT 语言模型，AI 在邮件、营销文案写作领域已有深度应用。Jasper AI 成立于 2020 年，公司定位是营销领域类的素材撰写，基于 AI 模型可为用户提供 50 多种模板并生成对应的文案，可覆盖大部分的营销场景。公司产品优势包含了更符合 Google SEO 规则、草稿创作效率更好、支持 26 种语言等六大优势。

Jasper AI 面向个人及企业用户，采用 SaaS 订阅制收费模式。Jasper AI 提供 Boss Mode 和 Business 两种计费模式，其中 Boss 模式针对个人及小组织（不超过 5 人）用户，基础收费为 49 美元/月，可使用不超过 5 万字，最高收费为 500 美元/月，可使用 70 万字。

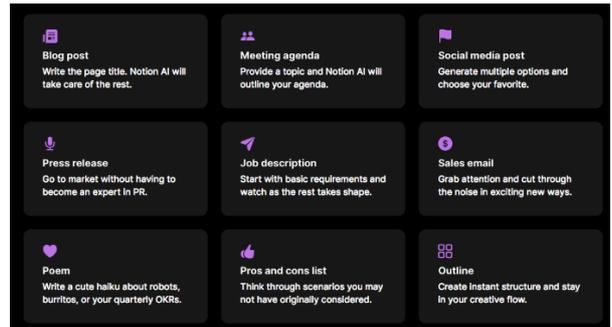
Notion AI 实现文本写作、语法纠错、机器翻译等功能，提高笔记类产品功能价值。Notion 笔记工具融合了文档、任务、知识库以及数据库功能，作为生产力工具服务于互联网群体。Notion AI 模块出色地解决了用户在使用笔记场景下遇到的大部分问题，如翻译、写作、内容发布等。Notion 产品通过免费+低价策略积累了超 3000 万用户，未来随着 Notion AI 模块的投入商用，有望提高客单价，提升产品营收能力。

图36: Jasper AI 产品六大优势



资料来源: Jasper AI, 浙商证券研究所

图37: Notion AI 产品服务矩阵



资料来源: Notion AI, 浙商证券研究所

图像生成领域，扩散模型的快速迭代大幅提升 AI 生成图像的性能，Stability AI 公司基于开源框架构建庞大内容生态。Stability AI 自成立以来，一直致力于开发面向全球消费者和企业用例的图像、语言、音频、视频、3D 等开放 AI 模型，公司开源的底层代码使得开发者可以绕开数据方面的限制，可为用户自己的商业产品提供动力。

2022 年 8 月，公司发布了面向消费者的 Dream Studio 应用程序，用户可在平台上通过自然语言描述创建逼真的图像、艺术和动画，目前平台注册用户已超过 100 万，用户来自全球 50 多个国家，创建了超过 1.7 亿张图像。用户可免费使用开源的 Diffusion 模型，但需要为 Dream Studio 平台生成的图像付费，平台采用点数充值的模式，用户可按照每 100 点数 1 英镑（合）的比例兑换积分，并根据生成图像的分辨率和渲染次数兑付不同数量的点数。以渲染 150 次、分辨率为 1024*1024 的图像为例，单张图片的价格约为 0.282 英镑。

图38： Stability AI 智能生成图像



资料来源：36 氪，浙商证券研究所

图39： Dream Studio 平台收费模式（单位：0.01 英镑）

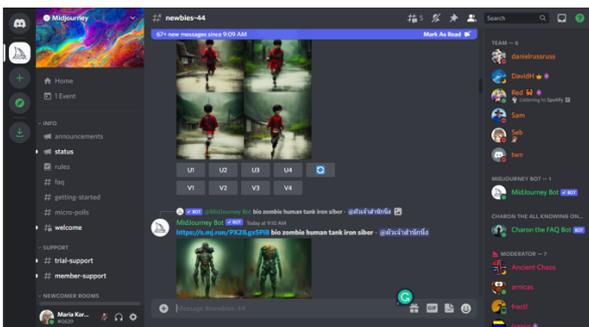
| Steps | 512x512 | 512x768 | 512x1024 | 768x768 | 768x1024 | 1024x1024 |
|-------|---------|---------|----------|---------|----------|-----------|
| 10 | 0.2 | 0.5 | 0.8 | 0.9 | 1.3 | 1.9 |
| 25 | 0.5 | 1.2 | 1.9 | 2.3 | 3.3 | 4.7 |
| 50 | 1.0 | 2.4 | 3.8 | 4.6 | 6.6 | 9.4 |
| 75 | 1.5 | 3.6 | 5.7 | 6.9 | 9.9 | 14.1 |
| 100 | 2.0 | 4.8 | 7.6 | 9.2 | 13.2 | 18.8 |
| 150 | 3.0 | 7.2 | 11.4 | 13.8 | 19.8 | 28.2 |

资料来源：Dream Studio，浙商证券研究所

Midjourney 利用 AI 智能生成图像，使用门槛低且图片可用性高。用户通过 Discord 平台，输入希望实现画面的文字，平台即可在一分钟的时间内完成四张缩略图的生成并就进行逐步渲染，用户选择相对满意的一张图片可继续细化，如变形、放大尺寸和细节优化等操作。

Discord 平台根据 GPU 使用时长、图片是否公开、图片生成数量的不同提供多种付费模式，免费版本可生成 25 张图片，而企业级版本（600 美元/年）每年可使用至少 120h 的 GPU 时间用于生成图片。

图40： Midjourney 图片生成工具



资料来源：MetaStellar，浙商证券研究所

图41： Midjourney 作品《太空歌剧院》

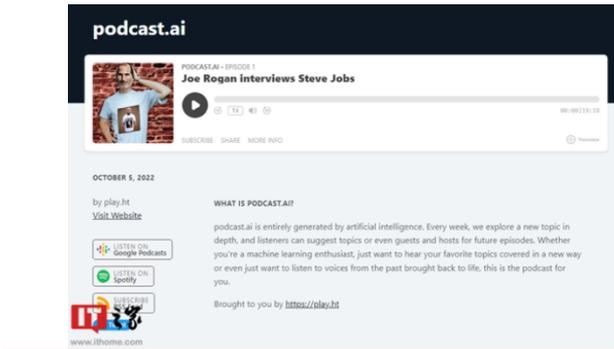


资料来源：网易，浙商证券研究所

Play.ht 基于 AI 驱动实现文本到语音的智能生成。 Play.ht 依靠 AI 从 IBM、微软、亚马逊和谷歌生成音频和语音，该工具对于将文本转换为自然声音尤为有用并且允许下载 MP3 和 WAV 文件格式的语音。Play.ht 拥有将播客文章转换为音频、实时语音合成、拥有超 570 种口音和声音、可提供逼真的配音的特点及优势。

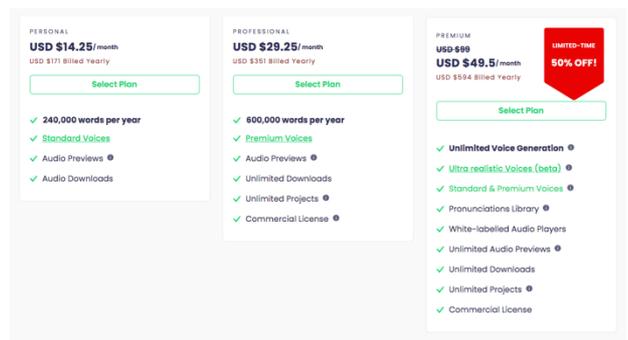
Play.ht 根据转换成音频的文本大小提供多种收费方案，专业版提供每年 60 万字（词）的语音生成权限，且生成的音频文件可用于商业用途，该方案定价为 29.25 美元/月，Premium 版本则可以根据用户的个性化需求，调整 AI 声音的发音特点。

图42: Play.ht 乔布斯声音通过 AI 智能生成采访音频



资料来源: IT 之家, 浙商证券研究所

图43: Play.ht 平台收费模式



资料来源: Play.ht, 浙商证券研究所

3.2 国内科技大厂相继入局，基于内容生态市场空间广阔

3.2.1 科技大厂持续开拓中文 AIGC，借助业务生态有望推出爆款应用产品

国内科技大厂布局生成式 AI，未来有望借助业务生态推出爆款应用产品。百度、阿里巴巴、腾讯、字节跳动等国内科技大厂相继布局 AIGC 领域，在图片生成、视频生成、3D 建模等细分场景下推出相关应用，未来有望在业务生态中打造爆款产品。

表14: 国内科技大厂 AIGC 布局

| 公司 | 典型产品 | 功能 |
|------|----------|---|
| 百度 | 文心 | 通过文字描述生成图片、图画修复、文档分析、蛋白质结构分析等 |
| 腾讯 | 优图 | 输入多张人像图片完成对于人像面部的 3D 建模、提升电影分辨率等 |
| 阿里巴巴 | Lubanner | 输入图片、风格、尺寸等，自动完成素材分析、抠图、配色等设计，生成多套符合要求的设计方案 |
| 字节跳动 | 剪映 | 通过文字生成视频、AI 配字幕等 |
| 网易 | 网易天音 | 通过文字生成歌曲，支持作词、作曲、编曲、演唱全链路音乐创作 |

资料来源: 百度文心, 腾讯优图, 阿里巴巴鲁班, 网易天音, 腾讯网, 浙商证券研究所

百度深度布局 AIGC 多年，围绕自有业务生态形成核心竞争力，在中文 AI 领域优势显著。2022 年 12 月，百度智能云发布国内首个全栈自研的 AI 基础设施“AI 大底座”，具备标准化输出 AI 的底层能力。百度围绕各场景，在 NLP、CV、跨模态、生物计算等领域

形成大模型。百度过去十年累积投入研发资金超 1000 亿元，并且连续四年在 AI 专利申请量和授权量上保持国内第一，技术优势明显。

表 15：百度基于文心大模型构建完整业务生态

| 类别 | 产品名称 | 主要特点 |
|--------|--------------------|---|
| 大模型 | NLP 大模型 | 面向语言理解、语言生成等 NLP 场景，具备超强语言理解能力以及对话生成、文学创作等能力。创新性地将大数据预训练与多源丰富知识相结合，通过持续学习技术，不断吸收海量文本数据中词汇、结构、语义等方面的新知识，实现模型效果不断进化。 |
| | CV 大模型 | 基于领先的视觉技术，利用海量的图像、视频等数据，为企业和开发者提供强大的视觉基础模型，以及一整套视觉任务定制与应用能力。 |
| | 跨模态大模型 | 基于知识增强的跨模态语义理解关键技术，可实现跨模态检索、图文生成、图片文档的信息抽取等应用的快速搭建，落实产业智能化转型的 AI 助力。 |
| | 生物计算大模型 | 融合自监督和多任务学习，并将生物领域研究对象的特性融入模型。构建面向化合物分子、蛋白分子的生物计算领域预训练模型，赋能生物医药行业。 |
| | 行业大模型 | 文心大模型与各行业企业联手，在通用大模型的基础上学习行业特色数据与知识，建设行业 AI 基础设施，行业覆盖能源、金融、航天、制造、传媒等。 |
| 开放 API | ERNIE 3.0 文本理解与创作 | 提供多种参数量级的、具备超强的语言理解能力和文本创作能力的 API 服务。 |
| | ERNIE ViLG AI 作画 | 全球规模最大的中文跨模态生成模型，通过自然语言实现图像生成与编辑。无需编程在体验专区探索大模型服务能力、找到应用场景，并可通过 API 集成服务能力。 |
| | 文心 PLATO 会话生成 | 全球首个基于隐空间的大规模生成式开放域对话模型，具备接近真人水平的多轮聊天能力； |
| 工具与平台 | 大规模套件 | ERNIEKit：基于最新一代预训练范式的 NLP 算法定制开发工具集； PaddleFleetX：旨在打造一套简单易用、性能领先且功能强大的端到端大模型工具箱，覆盖大模型环境部署、数据处理、预训练、微调、模型压缩、推理部署全流程，并支持语言、视觉、多模态等多个领域的前沿大模型算法。 |
| | 零门槛 AI 开发平台 EasyDL | EasyDL 文本：基于大模型为企业/开发者提供一整套 NLP 定制与应用能力； EasyDL 图像：零算法基础定制高精度图像应用 AI 模型，提供端云多种灵活部署方案 EasyDL 跨模态：基于大模型，提供领先的视觉、文本跨模态理解能力，根据业务需求轻松定制图文匹配模型； |
| | 全功能 AI 开发平台 BML | 支持一站式 AI 开发，集成飞桨全流程开发套件和丰富的产业场景使用范例； |
| | 产品 | 文心百中 文心一格 |

资料来源：百度文心大模型官网，浙商证券研究所

对标 ChatGPT，百度计划近期完成产品内测，有望加速 AI 产品商用化进程。近日，百度官方宣布，将在 3 月份完成其 ChatGPT 产品的内测并面向公众开放，该项目名字确定为文心一言，英文名 ERNIE Bot。公司产业级知识增强文心大模型 ERNIE 具备跨模态、跨语言的深度语义理解与生成能力，对标微软对 OpenAI 核心技术的布局，我们认为百度有望将对标产品应用到业务矩阵下的消费级和企业级应用中，加速 AI 商业化进程。

图44：百度“文心”AI大模型

| | | | | | | | |
|----------------------|---------------------------|------------------|------------------------|---------------------------|--------------------|----------------------------|--|
| 产品与社区 | 文心一格 AI艺术和创意辅助平台 | | 文心百中 大模型驱动的产业级搜索系统 | | 畅谷社区 大模型创意与探索社区 | | |
| 工具与平台 | EasyDL-大模型 专门给 AI 开发平台 | | BML-大模型 多功能 AI 开发平台 | | 大模型 API | | |
| 文心大模型 | 大模型套件 | | | | | | |
| | 数据标注与处理 | | 大模型精调 | | 大模型压缩 | | |
| | | | | | 高性能部署 | | |
| | | | | | 场景化工具 | | |
| | 行业大模型 | | | | | | |
| | 国网-百度·文心 | | 浦发-百度·文心 | | 航天-百度·文心 | | |
| | 深燃-百度·文心 | | 吉利-百度·文心 | | 泰康-百度·文心 | | |
| | | | | | TCL-百度·文心 | | |
| | | | | | 电影频道-百度·文心 | | |
| | | | | | 琼海-百度·文心 | | |
| NLP 大模型 | | | CV 大模型 | | 跨模态大模型 | | |
| 医疗 ERNIE-Health | | 金融 ERNIE-Finance | | 商品图文搜索表征学习 VIMER-UIMS | | 文图生成 ERNIE-VILG | |
| 对话 PLATO | | 搜索 ERNIE-Search | | OCR图像表征学习 VIMER-StrucText | | 文档智能 ERNIE-Layout | |
| 跨语言 ERNIE-M | | 代码 ERNIE-Code | | 多任务视觉表征学习 VIMER-UFO | | 视觉-语言 ERNIE-VIL | |
| | | 图网络 ERNIE-Sage | | | | 语音-语言 ERNIE-SAT | |
| | | | | | | 地理-语言 ERNIE-GeoL | |
| ERNIE 3.0 Tiny (轻量级) | | ERNIE 3.0 (百亿级) | | 鹏城-百度·文心 (千亿级) | | ERNIE 3.0 Zeus (任务知识增强) | |
| | | | | | | 视觉处理多任务学习 VIMER-TCIR | |
| | | | | | | 自监督视觉表征学习 VIMER-CAE | |
| | | | | | | 蛋白质结构预测 HeliFold | |
| | | | | | | 单序列蛋白质结构预测 HeliFold-Single | |

资料来源：百度文心大模型官网，浙商证券研究所整理

中文 AI 大模型持续迭代，国内人工智能底层技术能力已相对成熟。除百度之外，国内浪潮、华为等公司均持续布局，在中文 AI 巨量模型领域的创新和应用上持续突破，为 AI 在各场景的深度应用奠定坚实基础。

浪潮“源 1.0” AI 大模型超越 GPT-3，在语言智能领域表现优异。2021 年 9 月，浪潮信息发布 AI 大模型“源 1.0”，参数量超越 GPT-3 达 2457 亿，成为全球最大规模的 AI 巨量模型。“源 1.0”模型训练采用超 5000GB 的中文数据集，在语言智能方面表现优异，在文献分类、新闻分类，商品分类、原生中文推理、成语阅读理解填空、名词代词关系等任务上位居行业内领先地位，并且在成语阅读理解填空项目中的表现已超越人类得分。

华为基于“盘古” AI 大模型发展行业应用模型，赋能行业效率提升。2022 年 11 月，华为发布气象大模型，提供秒级天气预报，如重力势、湿度、风速、温度，气压等变量的 1 小时-7 天预测。盘古气象大模型在气象预报的关键要素和常用时间范围上精度均超过当前最先进的预报方法，同时速度相比传统方法提升 1000 倍以上。盘古气象大模型支持广泛的下游预报方案，如在台风路径预测任务上，相比传统数值气象预报方法，盘古气象大模型可以降低 20% 以上的位置误差。

图45：浪潮“源”AI大模型



资料来源：搜狐，浙商证券研究所

图46：华为“盘古”AI大模型



资料来源：科技世界网，浙商证券研究所

3.2.2 国内细分赛道涌现大量垂类公司，关注实际业务场景下的 AI 应用渗透

依托国内丰富内容生态，垂类公司探索 AI 应用产品，AIGC 将有望成为内容生态发展的新引擎。我们认为 AIGC 发展根据内容分类，主要可以分为 3 个阶段：一是 AIGC 内容生产技术完善阶段，实现文本、视频、图片生成以及三者的跨模态转换；二是多模态生成技术的聚合应用，以虚拟人为代表；三是虚拟内容生态，即元宇宙。目前国内外仍然处于第一阶段，AIGC 内容以文本、图片和视频为主。

目前国内各细分赛道垂类公司布局有望推动 AI 商业化在实际场景中的深度应用。近几年来，在法律、营销、医疗、人机交互等领域不断涌现垂类公司，基于 AI 在文本、图像和音视频等方面赋能业务的效率及产品竞争力的提升。

表16：国内细分赛道企业在 AIGC 各应用场景持续布局

| 生成内容 | 企业 | 成立时间 | 主要特点 |
|------|------|--------|--|
| 文字 | 聆心智能 | 2021 年 | 提供 AI 驱动的高质量数字疗法和新一代精神心理的解决方案，为临床实践实现规模化赋能；公司致力于 AI 情绪智能为所有人变革传统心理健康的服务模式，提供随时随地、人人可及的服务。 |
| 文字 | 澜舟科技 | 2021 年 | 公司针对商业场景数字化转型，以自然语言处理为基础提供商业洞见类产品，主要产品包括基于预训练模型的功能引擎（搜索、生成、翻译、对话等）和针对垂直行业场景的 SaaS 产品。 |
| 文字 | 彩云科技 | 2015 年 | 公司从事天气预报、机器翻译和智能写作，致力于发展具有高阶认知能力的 AI 以提升人类感知环境、相互交流和与 AI 沟通的能力，并面向全球用户和开发者开放。 |
| 文字 | 秘塔科技 | 2018 年 | 公司致力于利用 AI 为法律行业赋能，提升法律从业者工作效率，改变法律翻译市场高收费、低效率的现状。2019 年 4 月公司第一款产品“秘塔翻译”发布，新产品为文档自动化系统，可一键生成多达上百份专业文件。 |
| 文字 | 香侬科技 | 2017 年 | 公司解决方案包括 AI 智能写作平台，服务内容包括机器翻译，可根据中国人的书写习惯进行深度翻译优化，并支持英文 PDF 的翻译查询及非结构化信息抽取等服务。 |
| 图像 | 感知阶跃 | 2020 年 | 公司提供 AI 模特图片解决方案，有效降低拍照成本和效率，提升电商转化率。产品通过 AI 算法生成逼真的虚拟模特展示服装。 |
| 视频 | 影谱科技 | 2009 年 | 公司专注于智能影像生产领域的视觉技术，凭借在 AI、视频结构化等领域的技术优势，提供基于智能影像生产等相关技术的商业化综合服务。 |
| 视频 | 帝视科技 | 2016 年 | 公司专注于 AI 与计算机视觉技术研究与产品开发，提供 5G+AI 超高清视频和 AI 智能制造等领域技术服务和行业解决方案。 |
| 音频 | 标贝科技 | 2016 年 | 公司专注于智能语音交互和 AI 数据服务技术。公司依托先进的 AI 语音交互技术及高精度数据采集平台，打造多场景应用的语音交互方案，包括通用场景的语音合成和语音识别，以及 TTS 音色定制，声音复刻，情感合成和声音转换在内的语音技术产品；AI 数据业务涵盖语音合成、语音识别、图像视觉、NLP、3D 点云等数据服务。 |
| 游戏 | 知觉之门 | 2020 年 | 公司运用 AI 为游戏行业提供完整的解决方案。公司产品及服务覆盖游戏的全生命周期，包含智能内容生成、智能测试、智能数据运营、智能投放等类型。 |
| 虚拟人 | 倒映有声 | 2019 年 | 公司是一家语音合成科技创新企业，提供基于端到端的神经网络的音频合成系统及行业解决方案，以领先的神经网络设计和先进的深度学习合成器为依托，独创情绪、情感控制模块，在音色模拟、情感展现、多语种等方面达到国内外领先水平。 |
| 综合 | 红棉小冰 | 2020 年 | 公司研发跨平台 AI 机器人“小冰”，致力于实现跨平台交互，让用户在不同品牌的终端和不同企业的生态中都找到小冰，目前用户可以自主通过小冰框架，创造并训练其拥有的人工智能主体。 |

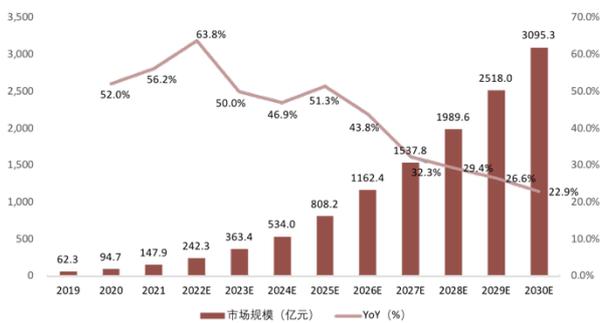
资料来源：各公司官网，36 氪，腾讯网，浙商证券研究所

对标海外公司发展，内容型公司有多种商业模式可选择，业务结合场景深度将决定玩家核心竞争力。目前海外 AIGC 公司的商业模式大体可分为底层平台对外开放按照实际数据量收费、按产出内容收费、类 SaaS 的软件服务收费、有偿模型使用等，其核心是在于赋能下游 B 端及 C 端用户的实际业务效率和质量提升，因此企业需要对一体化解决方案、行业深度绑定、业务闭环等领域不断耕耘。

3.3 国内内容生态市场空间广阔，政策支持下看好 AI 应用加速渗透

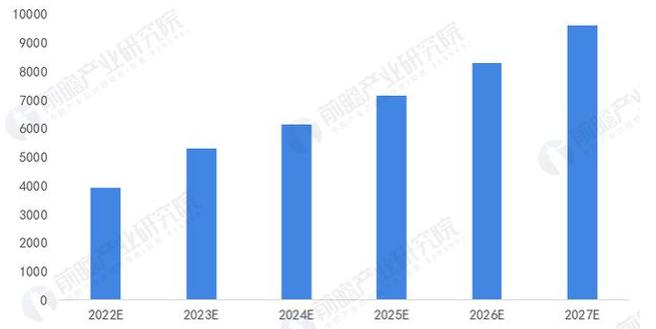
国内成熟内容生态为 AI 应用提供沃土，虚拟人有望催生新内容需求。以短视频生态为例，根据前瞻产业研究院预计，到 2027 年我国短视频市场规模将达到 9624 亿元，市场空间广阔，视频、游戏、网络文学等内容生态也将不断丰富，市场规模持续增长有望带动内容创作需求不断升高，以 AIGC 为核心的内容创作工具及相关服务有望加速市场渗透。同时，随着 AR/VR、云计算、区块链等核心技术持续迭代，我国元宇宙内容生态有望加速形成，根据速途元宇宙研究院预测，到 2030 年我国虚拟人整体市场规模有望达 3095.3 亿元，市场前景广阔。

图47：我国虚拟人整体市场规模趋势预测



资料来源：速途元宇宙研究院，浙商证券研究所

图48：2022-2027 年我国短视频市场规模预测（单位：亿元）



资料来源：前瞻产业研究院，浙商证券研究所

国家将人工智能上升至国家战略层面，政策支持有望推进行业发展。政府随着人工智能行业的不断发展，对政策内容也持续进行细化，从顶层设计至创新成果转化路径，再到芯片、开源平台等技术层面都有相应政策指导，为行业提供了清晰的路径指引。

表17：国家政策大力支持人工智能商业化落地

| 颁布时间 | 颁布主体 | 政策名称 | 主要内容 |
|--------|--------------------------|--|--|
| 2019.3 | 科技部 网信办 | 《关于促进人工智能和实体经济深度融合的指导意见》 | 探索人工智能创新成果应用转化路径和方法，构建智能经济。 |
| 2019.8 | 科技部 | 《国家新一代人工智能开放创新平台建设指引》 | 鼓励人工智能细分领域领军企业搭建开源、开放平台，推动行业应用。 |
| 2020.7 | 中央网信办 等五部门 | 《国家新一代人工智能标准体系建设指南》 | 到 2021 年，明确 AI 标准化顶层设计，研究标准体系建设和标准研制的总体规则，明确标准之间的关系，指导 AI 标准化工作的有序开展，完成关键通用技术、关键领域技术等 20 项以上重点标准的预研工作。到 2033 年初步建立 AI 标准体系，重点研制数据、算法、系统、服务等重点标准，并率先在制造、交通、金融、安防、家居、教育、医疗健康、司法等重点行业和领域进行推进。 |
| 2021.3 | 国务院 | 《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》 | 聚焦高端芯片、操作系统、人工智能关键算法、传感器等关键领域，加快推进基础理论、基础算法、装备材料等研发突破与迭代应用。培育壮大人工智能、大数据、区块链、云计算、网络安全等新兴数字产业，提升通信设备、核心电子元器件、关键软件等产业水平。 |
| 2021.7 | 工业和 信息化部 | 《新型数据中心发展三年行动计划》 | 推动新型数据中心与 AI 等技术协同发展，构建完善新型智能算力生态体系； |
| 2021.9 | 国家新一代人 工智能治理专 业委员会 | 《新一代人工智能伦理规范》 | 旨在将伦理道德融入人工智能全生命周期，为从事人工智相关活动的自然人、法人和其他相关机构等提供伦理指引。 |

资料来源：工信部，科技部，中国政府网，浙商证券研究所

4 围绕 AI 商业化三大主线挖掘核心投资标的

4.1 关注具备底层算法模型核心技术优势的厂商

4.1.1 拓尔思，国内 NLP 龙头厂商

深耕 NLP 领域 30 年，公司发展成为国内政务领域 NLP 龙头厂商。公司以“语义智能+”为发展战略深耕多年，主要业务覆盖内容安全和互联网空间治理、数字政府和数据智能三大板块，向以党政机关及企事业单位为主的下游客户提供大数据检索、智能风控营销、舆情监控等服务。公司以深耕多年的大数据积累与行业领先的 NLP 技术成为国内政府政务领域 NLP 龙头，多年以来一直保持行业领先的市场地位。

公司以 NLP 技术为核心，金融、政府、公安领域竞争优势显著。目前公司以公共安全市场为最主要业务，产品和场景包括公安情报研判指挥、公安知识图谱、公安舆情监控、开源情报智能分析、开源情报数据采集监测以及各类领域知识库构建；在信息安全领域，全资子公司天行网安提供各类网安产品服务。

公司积极开辟新赛道，机器人+虚拟人有望为公司业务提供新增量。公司将以“虚拟人+”场景为突破口，开发支撑虚拟人的 AI 技术平台，推动虚拟人在各行业场景中的率先落地。公司基于中文 NLP 核心技术、海量数据积累及具备行业专业能力的知识图谱等元宇宙技术基因的优势，致力构建数字虚拟人的智能引擎，提供虚拟人智能对话、知识积累、语义理解和智能决策等技术支撑。

在机器人领域，公司与头部 AI 厂商共同合作研究人形机器人软件技术。合作双方将在人形机器人的智能化方面，研发云端协同的 AI 算法框架以及面向特定领域的知识图谱技术，公司在 NLP、知识图谱领域的积淀将赋能机器人软件大脑及开放软件生态的构建。

我们认为在信创浪潮下，公司有望受益政府客户在相关领域的投入增加，同时机器人、元宇宙等行业的加速发展可为公司提供多重增长空间，看好公司以 NLP 和知识图谱为核心技术构建业务生态，实现业务规模加速成长。

4.1.2 科大讯飞，AI+行业应用领域领跑者

公司深耕智能语音领域，战略转型人工智能赛道，发展成为行业领军企业。公司在智能语音行业布局多年，核心技术优势明显，2010 年以来向人工智能产业发展，推动 AI 与各行业的应用结合，在智慧教育、智能办公、智慧公安、医学影像、智能汽车、智慧服务、智慧城市等领域不断取得成果。

坚持“平台+赛道”发展战略，公司构建自有业务闭环生态体系。公司坚持“平台+赛道”的人工智能战略，依托国内首家上线的 AI 开放平台——讯飞开放平台，为开发者提供一站式人工智能解决方案，构建 AI 产业生态，并在教育、医疗、办公、智慧城市等领域实现 AI 的深度应用，在多语种语音合成和识别、自然语言处理、图文识别、人机交互等领域提供丰富的 AI 能力，赋能下游客户的效率提升。

推动 AI 走向强人工智能，“讯飞超脑 2030 计划”为公司明确未来成长空间规划。2022 年，公司发布“讯飞超脑 2030 计划”，旨在构建基于认知智能的复杂智能系统，深度融合垂直行业的细分场景任务，实现各业务场景赋能。公司计划分为三个阶段，第一阶段着眼于机器人和数字虚拟人领域；第二阶段着眼于自适应行走的外骨骼机器人和陪伴数字虚拟人家族，以及面向青少年的抑郁症筛查平台；第三阶段计划全面进入家庭场景。

我们认为受益于人工智能产业发展机遇，国家政策对发展数字经济和建设科技强国的进一步推动、产业信息化应用需求的不断增长以及公司各类产品、技术不断创新，公司业务规模有望持续扩大。

4.1.3 其它建议关注标的

- (1) 国外：微软（OpenAI, ChatGPT），谷歌（DeepMind），Meta（OPT 大模型）；
- (2) 国内：百度（“文心” AI 大模型），腾讯（“混元” AI 大模型），浪潮信息（“源” AI 大模型）；

4.2 关注各细分赛道下兼具场景理解与 AI 布局优势的垂类厂商

4.2.1 海康威视：起步安防行业的智能物联龙头

公司致力于安防视频监控领域，融合 AI 技术发展成为智能物联龙头。公司起步于安防前端领域，随后转型行业一体化解决方案提供商，满足客户定制化需求。2012 年起，公司深度布局 AI 技术，并于 2015 年正式发布深度智能产品，进入智能化时代，并陆续推出 AI 智能产品，成为安防领域数字化领军企业。公司着力定位于“智能物联 AIoT”，将物联感知、人工智能、大数据服务于各业务场景。

公司实现业务生态闭环，产品矩阵覆盖软硬件。公司自成立以来，始终致力于安防行业的发展，目前已形成软硬融合、云边融合的产品体系。公司利用 HEOP 嵌入式开放平台，实现所有类型的智能物联网设备具有相同软件基础，大幅提升开发效率。硬件方面，公司形成“节点全面感知+域端场景智能+中心智能存算”的产品架构，边缘节点产品涵盖前端摄像机产品、智能交通与移动产品、门禁与对讲产品等多个领域；边缘域产品深入行业，包括智能应用一体化设备、会议平板产品、智能视频传输产品等三大产品。

公司深耕各大业务场景，具备 AI 商业化应用的良好条件。公司以安防为核心，业务覆盖公安、交通、金融、文教卫等各场景，同时积极布局包含智能家居、机器人等创新业务，对 AI 赋能安防领域具备业务优势，我们认为公司有望受益于 AI 商业化进程加速，持续探索 AI 在安防领域的产品服务，为公司贡献驱动力。

4.2.2 其它建议关注标的

- (1) 智慧城市场景：云从科技；
- (2) 智慧金融场景：格灵深瞳；
- (3) 办公场景：金山办公；
- (4) AI 绘画场景：万兴科技；
- (5) 智慧安防场景：商汤；

4.3 围绕 AI 数据、算力等基础设施选择优质投资标的

4.3.1 海天瑞声：AI 训练数据标注优质提供商

公司是国内领先的训练数据专业提供商，致力于为 AI 产业链各类机构提供专业数据集。公司致力于为各类 AI 厂商和机构提供算法模型开发训练需要的专业数据集，覆盖智能语音（语音识别、语音合成等）、计算机视觉、自然语言等多个核心领域，应用场景多元，包含人机交互、智能驾驶、智能家居、智慧城市等。

公司智能语音业务发展成熟，竞争壁垒高且小语种领域优势显著。公司在语音语言学基础研究方面积累深厚，基于发音词典构建技术和流程技术的持续迭代，构建高质量的自然语言处理模型训练所需的标注数据，截至 2022H1 公司已具备 190 个语种/方言的覆盖能力，并且公司致力于开拓海外市场，未来业务规模有望加速扩张。

公司成立智能驾驶事业部，积极布局自动驾驶业务。公司于 2022 年 6 月上线了第三代智能驾驶标注平台，目前自动驾驶数据标注方面业务已覆盖全景语义分割、2D 图像标注、2D/3D 融合标注、3D 点云标注等。公司基于多年以来在数据标注领域的算法积淀，在标注效率、准确度等指标上具备显著竞争优势，未来有望构建行业领先的综合性、规模化和自动化的数据处理能力体系。

公司客户资源粘性高优势显著，未来有望受益行业加速发展和人工智能商业化进程加速。公司在智能语音业务方面与字节跳动、阿里巴巴、腾讯、百度、科大讯飞及海外大型客户保持良好合作关系，考虑人工智能未来有望加速商业化应用，叠加国家信创浪潮政策支持，公司有望迎来加速成长期。

4.3.2 其它建议关注标的

国外：英伟达（GPU）；

国内：寒武纪（AI 芯片）。

5 风险提示

- 1、AI 技术迭不及预期的风险；
- 2、AI 商业化产品发布不及预期；
- 3、政策不确定性带来的风险；
- 4、下游市场不确定性带来的风险；

股票投资评级说明

以报告日后的6个月内，证券相对于沪深300指数的涨跌幅为标准，定义如下：

1. 买入：相对于沪深300指数表现+20%以上；
2. 增持：相对于沪深300指数表现+10%~+20%；
3. 中性：相对于沪深300指数表现-10%~+10%之间波动；
4. 减持：相对于沪深300指数表现-10%以下。

行业的投资评级：

以报告日后的6个月内，行业指数相对于沪深300指数的涨跌幅为标准，定义如下：

1. 看好：行业指数相对于沪深300指数表现+10%以上；
2. 中性：行业指数相对于沪深300指数表现-10%~+10%以上；
3. 看淡：行业指数相对于沪深300指数表现-10%以下。

我们在此提醒您，不同证券研究机构采用不同的评级术语及评级标准。我们采用的是相对评级体系，表示投资的相对比重。

建议：投资者买入或者卖出证券的决定取决于个人的实际情况，比如当前的持仓结构以及其他需要考虑的因素。投资者不应仅仅依靠投资评级来推断结论。

法律声明及风险提示

本报告由浙商证券股份有限公司（已具备中国证监会批复的证券投资咨询业务资格，经营许可证编号为：Z39833000）制作。本报告中的信息均来源于我们认为可靠的已公开资料，但浙商证券股份有限公司及其关联机构（以下统称“本公司”）对这些信息的真实性、准确性及完整性不作任何保证，也不保证所包含的信息和建议不发生任何变更。本公司没有将变更的信息和建议向报告所有接收者进行更新的义务。

本报告仅供本公司的客户作参考之用。本公司不会因接收人收到本报告而视其为本公司的当然客户。

本报告仅反映报告作者的出具日的观点和判断，在任何情况下，本报告中的信息或所表述的意见均不构成对任何人的投资建议，投资者应当对本报告中的信息和意见进行独立评估，并应同时考量各自的投资目的、财务状况和特定需求。对依据或者使用本报告所造成的一切后果，本公司及/或其关联人员均不承担任何法律责任。

本公司的交易人员以及其他专业人士可能会依据不同假设和标准、采用不同的分析方法而口头或书面发表与本报告意见及建议不一致的市场评论和/或交易观点。本公司没有将此意见及建议向报告所有接收者进行更新的义务。本公司的资产管理公司、自营部门以及其他投资业务部门可能独立做出与本报告中的意见或建议不一致的投资决策。

本报告版权均归本公司所有，未经本公司事先书面授权，任何机构或个人不得以任何形式复制、发布、传播本报告的全部或部分内容。经授权刊载、转发本报告或者摘要的，应当注明本报告发布人和发布日期，并提示使用本报告的风险。未经授权或未按要求刊载、转发本报告的，应当承担相应的法律责任。本公司将保留向其追究法律责任的权利。

浙商证券研究所

上海总部地址：杨高南路729号陆家嘴世纪金融广场1号楼25层

北京地址：北京市东城区朝阳门北大街8号富华大厦E座4层

深圳地址：广东省深圳市福田区广电金融中心33层

上海总部邮政编码：200127

上海总部电话：(8621) 80108518

上海总部传真：(8621) 80106010

浙商证券研究所：<https://www.stocke.com.cn>