

海外 ChatGPT 专题

ChatGPT 风口已至，商业化落地加速

投资建议： 强于大市（维持评级）

上次建议： 强于大市

➤ **事件：**美国AI公司OpenAI推出基于大语言模型的对话模型ChatGPT，可提供高质量的回答，并能实现创作、编程等复杂功能，备受市场关注。不到两个月的时间，ChatGPT全球日活用户已突破千万。

➤ ChatGPT是突破式的创新技术

ChatGPT是OpenAI公司推出的全新聊天机器人模型。通过引入人类反馈的强化学习，大幅提升了AI在人机对话时的准确度和可控性，具有强大的语言理解能力和语言表达能力。GPT模型仍在持续迭代，更先进大语言模型GPT-4有望在2023年推出，有望进一步推动AIGC产业发展。

➤ ChatGPT应用及商业化落地加速

科技公司纷纷涌入AIGC赛道，优秀的AIGC大模型层出不穷，我们认为基于AI文本生成的模型ChatGPT有望率先应用落地。AIGC赛道相关公司受到资本青睐，AIGC头部初创公司OpenAI最新估值约为290亿美元。作为AIGC领域领先的模型ChatGPT在对话机器人、智能创作等领域应用广泛，亚马逊、微软、Jasper等公司已经开启商业化之路，商业化前景广阔。

➤ 建议关注

作为AIGC领域顶尖的模型，ChatGPT有望对现有生产力工具进行变革，引领赛道发展。我们看好AIGC领域前景，给予“强于大市”评级。根据AIGC领域产业链上下游，我们建议关注AIGC上游数据服务商海天瑞声

（688787.SH），中游技术领先的公司科大讯飞（002230.SZ）、拓尔思（300229.SZ）、百度（09888.HK），下游应用率先受益的公司昆仑万维（300418.SZ）、中文在线（300364.SZ）、视觉中国（000681.SZ）、掌阅科技（603533.SH）、阅文集团（0772.HK）、汉仪股份（301270.SZ）。

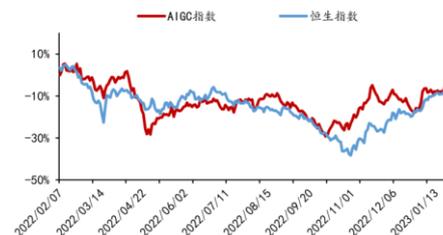
➤ **风险提示：**政策监管风险、AIGC技术发展不及预期风险、商业化落地不及预期风险

➤ 建议关注标的

简称	EPS			PE			CAGR-3	评级
	2021A	2022E	2023E	2021A	2022E	2023E		
海天瑞声	0.89	0.48	1.23	142.18	217.43	85.70	38.1%	/
科大讯飞	0.70	0.67	1.03	64.02	64.42	41.65	25.6%	/
拓尔思	0.34	0.44	0.56	51.18	40.01	31.01	27.9%	/
百度	3.58	9.03	11.28	33.45	27.31	18.93	53.5%	/
昆仑万维	1.31	0.99	1.19	13.93	18.38	15.39	1.8%	/
中文在线	0.14	-0.13	0.24	86.15	-91.03	49.77	27.8%	/
视觉中国	0.22	0.19	0.28	78.05	85.53	59.51	18.0%	/
掌阅科技	0.35	0.36	0.49	49.92	46.99	35.19	20.1%	/
阅文集团	1.83	1.16	1.53	18.63	32.16	23.84	-1.0%	/
汉仪股份	0.91	0.78	1.02	59.06	51.59	39.19	9.8%	/

数据来源：iFind，采用iFind一致预期，国联证券研究所整理，股价取2023年2月2日收盘价

相对大盘走势



分析师：陈梦瑶

执业证书编号：S0590521040005

邮箱：cmy@glsc.com.cn

相关报告

1、《互联网的白银时代海外》2023.01.16

正文目录

1	ChatGPT 突破式创新技术	3
1.1	基于大语言模型 (LLM) 的对话模型	3
1.2	GPT 模型加速迭代	4
2	ChatGPT 应用及商业化落地加速	6
2.1	巨头涌入推动赛道发展	6
2.2	ChatGPT 应用前景广阔	9
3	建议关注	12
3.1	AIGC 赛道上游公司	12
3.2	中游技术领先的公司	12
3.3	下游应用率先受益的公司	12
4	风险提示	13

图表目录

图表 1:	ChatGPT 写英文故事	3
图表 2:	ChatGPT 撰写新闻演示	3
图表 4:	GPT 特点及发展历程	4
图表 5:	预训练模型越来越“大”	5
图表 6:	ChatGPT 全球日活用户突破千万	6
图表 7:	各大公司 AI 生成模型	6
图表 8:	主流 AI 文本-文本生成模型	7
图表 9:	Cluade 与 ChatGPT 功能对比	7
图表 10:	GLM-130B 与 GPT-3、OPT-175B 对比	8
图表 11:	2017-2022 年生成式 AI 赛道融资金额 (亿美元) 及数量 (笔)	8
图表 12:	ChatGPT 在游戏中的应用	9
图表 13:	ChatGPT 结合 Stable Diffusion 生成的图片	9
图表 14:	ChatGPT 编程演示	10
图表 15:	ChatGPT 应用探索	10
图表 16:	ChatGPT 撰写房源信息	11
图表 17:	GPT-3 的主要企业客户	11
图表 18:	ChatGPT 潜在商业模式	11

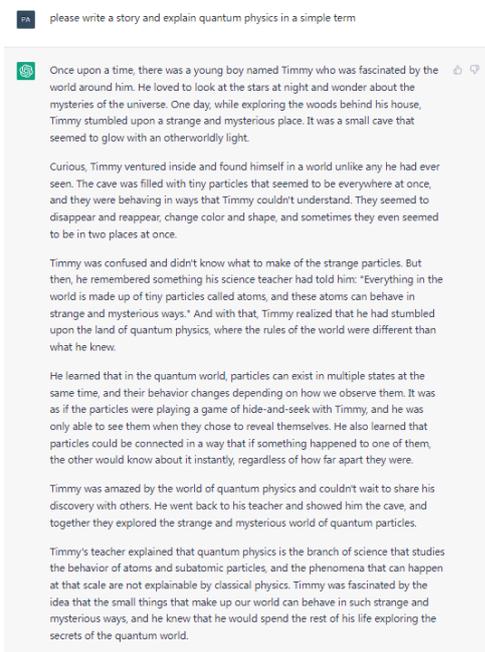
1 ChatGPT 突破式创新技术

1.1 基于大语言模型 (LLM) 的对话模型

ChatGPT 是 OpenAI 公司推出的全新聊天机器人模型。可以实现像人一样的交流，甚至完成写邮件、文案创作等功能。ChatGPT 还是基于 transformer 模型，通过引入人类反馈的强化学习 (RLHF)，大幅提高了 AI 在人机对话时的准确度和可控性。与其他 AI 对话机器人相比，ChatGPT 具有强大的语言理解能力和更丰富的语言表达能力，并且能够生成更自然、更流畅的文本，可以回答非常复杂的推理问题。ChatGPT 核心能力如下：

- (1) 问题的理解能力大幅提升，提升了模型和人类意图的一致性，具有记忆能力，可以实现连续多轮对话；
- (2) 结果的准确性大幅提升，主要表现在回答的更加的全面，同时可以主动承认错误、发现无法回答的问题；
- (3) 具备识别非法和偏见的机制，针对不合理提问提示并拒绝回答；
- (4) 理解用户需求并创造内容，甚至可以协助进行代码编写。

图表 1: ChatGPT 写英文故事



资料来源: ChatGPT, 东方财富网, 国联证券研究所

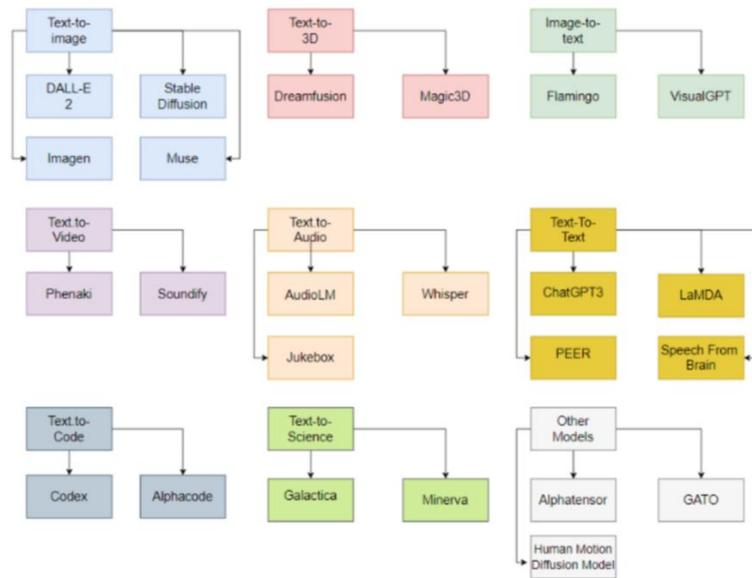
图表 2: ChatGPT 撰写新闻演示



资料来源: ChatGPT, 东方财富网, 国联证券研究所

ChatGPT 的出现将 AI 生成推到了前台，但 ChatGPT 只是 AI 生成模型的一种。海外学者 Roberto 和 Eduardo 将 AI 生成模型分成 9 大类，内容形态包括图像、视频、音频、文本、3D 模型等。

图表 3： 在生成效果上达到 SOTA 的 AI 生成模型总览



资料来源：arxiv.org, 《ChatGPT is not all you need. A State of the Art Review of large Generative AI models》, 量子位, 国联证券研究所

1.2 GPT 模型加速迭代

GPT 功能越来越强大。从 2018 年 6 月 GPT-1 发布以来，模型参数量、预训练数据量越来越大，当前已经发展到 GPT-3.5 (ChatGPT 原型)。性能更加强大的 GPT-4 也正在训练测试中，有望在 2023 年正式发布。

图表 4： GPT 特点及发展历程



资料来源：微信公众号“工技模型算法”、“阿法免研究笔记”，国联证券研究所整理

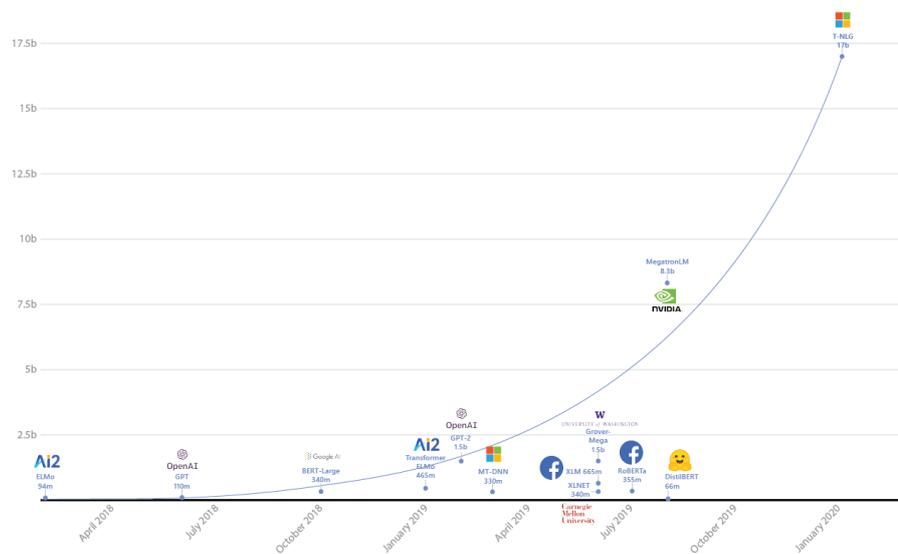
无监督学习 GPT-1。GPT-1 诞生于 2018 年，采用了 Transformer 为核心结构，通过自左向右生成式的构建预训练任务，然后得到一个通用的预训练模型，这个模型和 BERT 一样都可用来做下游任务的微调。GPT-1 在未经调试的任务上有一些效果，但其泛化能力远低于经过微调的有监督任务，且 GPT-1 使用的模型规模和数据量都比较小，这也就促使了 GPT-2 的诞生。

多任务学习 GPT-2。GPT-2 诞生于 2019 年，同样基于 Transformer，相比于 GPT-1，GPT-2 采用了更多的网络参数和更大的数据集，最大模型共计 48 层，参数

量达 15 亿。在性能上，在各种任务如阅读、对话、写小说等方面，效果都有所提高，也在当时达到了最佳的效果。

海量参数模型 GPT-3。 GPT-3 采用 1750 亿个参数，规模是 GPT-2 的 117 倍，在不经微调便可以识别数据中隐藏的含义。作为一个无监督模型，GPT-3 几乎可以完成自然语言处理的绝大部分任务，例如面向问题的搜索、阅读理解、语义推断、机器翻译、文章生成和自动问答等等。同时，GPT-3 在两位数的加减运算任务的准确率几乎达到了 100%，甚至还可以依据任务描述自动生成代码。以下显示了各类 NLP 模型的大小，其中 GPT-3 是 Turing NLG 的十倍，远远超过 GPT-1 和 GPT-2。

图表 5：预训练模型越来越“大”



资料来源：Microsoft Research，国联证券研究所

虽然 GPT-3 已经非常强大，可以完成诸如将网页描述转换为相应代码、模仿人类叙事、创作定制诗歌、生成游戏剧本等复杂任务，但是 GTP-3 并不完美。最主要的问题之一就是聊天机器人和文本生成工具无法判断内容的质量和好坏，可以学习网络上所有文本，可能产生恶意的甚至攻击性的语言输出，影响落地应用。

基于人工标注数据和强化学习的 GPT-3.5 (ChatGPT 原型)。 ChatGPT 在一个开源数据集上进行训练，不过采用了更大规模的训练参数，ChatGPT 的训练参数是 GPT-3 的 10 倍以上。除了训练参数规模的变化，ChatGPT 还采用了颠覆式的迭代方式：人工标注数据和强化学习，其本质是加上了在 GPT-3 上去掉的微调步骤，从而实现了在与人类互动时从反馈中强化学习。因此，ChatGPT 不仅可以理解人类不同指令的含义，也会甄别高水准答案，还能处理多元化的主题任务。既可以回答用户后续问题，也可以质疑错误问题和拒绝不适当的请求。

ChatGPT 仍存在很多局限性，需持续优化。 ChatGPT 输入不能有错，有时解释不够人性化，偶尔会给出看似正确但荒谬的答案等。虽然目前的 ChatGPT 无法给我们生产方式带来根本性的变革，但有可能改变人们与计算机的互动方式，推动从用户创作 (UGC) 到 AI 创作 (AIGC) 的转型。

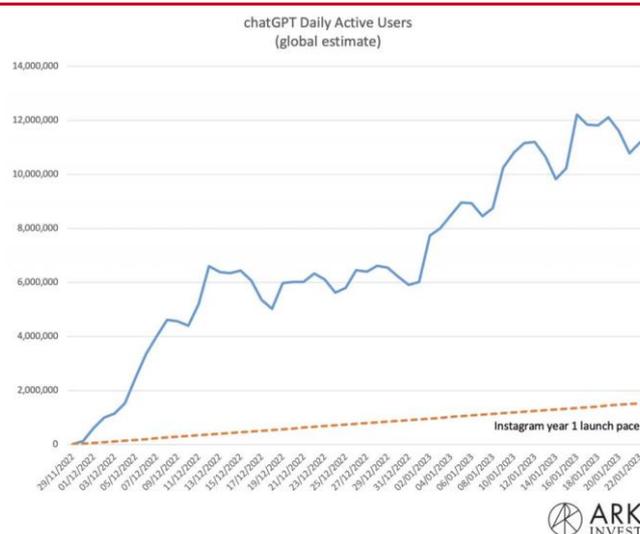
据投资公司 Radical Ventures 预测，GPT-4 或采用更大规模的数据集，可能在 10 万亿个 token 的数据集上进行训练，同时它的参数有望比 Megatron-Turing 的要少。GPT-4 有可能是多模态的，支持文本、图片、视频等多种数据类型的输入。意味着 GPT-4 可以根据文本提示词 (prompt) 生成图像，或者是可以输入视频然后通过文本的形式回答问题。

2 ChatGPT 应用及商业化落地加速

2.1 巨头涌入推动赛道发展

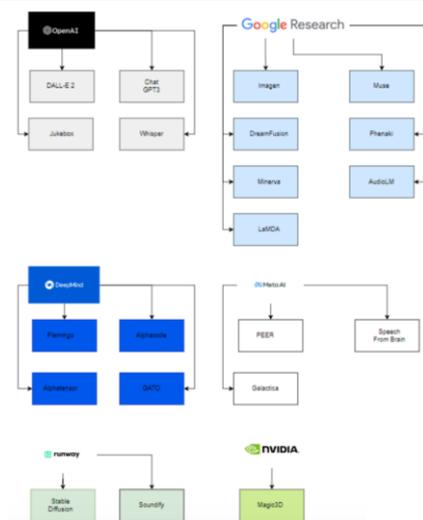
ChatGPT 日活量突破千万。 ChatGPT 在 2022 年 11 月 30 日宣布公测之后，迅速走红全网，上线 5 天该模型的全球用户数量已突破百万。根据 ARK 风险投资公司统计数据，上线不到 40 天，ChatGPT 的日活用户已突破千万。

图表 6: ChatGPT 全球日活用户突破千万



资料来源 SimilarWeb, ARK Invest, 国联证券研究所

图表 7: 各大公司 AI 生成模型



资料来源: 量子位, 国联证券研究所

优秀的 AI 生成模型层出不穷。除了 ChatGPT，2022 年还涌现很多优秀的 AI 生成模型，以谷歌、DeepMind、OpenAI、Meta 等为代表。AI 文本生成作为当下最通用的交互媒介，我们认为当前以 ChatGPT 为代表的 AI 文本生成模型有望率先应用落地。

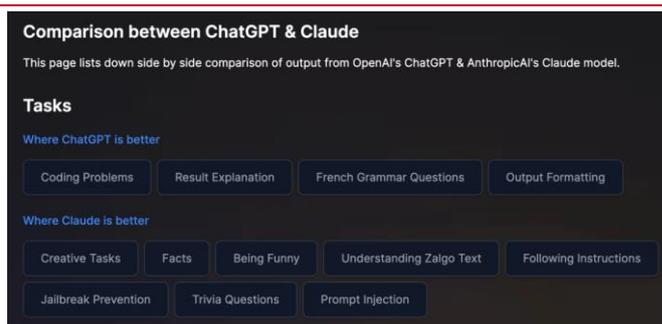
图表 8: 主流 AI 文本-文本生成模型

模型	公司	原理	特点
ChatGPT	OpenAI	基于 Transformer 打造，强化学习阶段采用人类反馈强化学习，它用人类打造的对话数据集、以及 InstructGPT 数据集的对话格式进行训练	对话生成 AI，专注于生成语言文本。懂得回答问题、拒绝不正当的问题请求并质疑不正确的问题前提。此外也可以生成代码和进行简单数学运算
LaMDA	Google	基于 Transformer 打造，利用了其在文本中呈现的长程依赖关系能力。拥有 1370 亿参数，在 1.56T 的公共对话数据集和网页文本上进行训练，只有 0.001% 的训练数据被用于微调。	2021 年 5 月的 I/O 大会推出最新的 LaMDA 系统。与其他语言模型不同，LaMDA 训练的是对话，专注于生成对话。但由于 LaMDA 存在较大误差，容易对用户造成伤害，因此未大规模投入使用
PEER	Meta AI	基于维基百科编辑历史进行训练，直到模型掌握完整的写作流程。模拟人类写作的过程，从打草稿、征求建议到编辑文本，再不断迭代。同时训练多个 PEER 的实例，可填补写作过程中的多个环节，从而通过自训练 (self-training) 技术提高训练数据的质量、数量以及多样性。	模型允许将写作任务分解成更多子任务，并允许人类随时干预，引导模型写出人类想要的作品。解决了传统语言模型只会生成最终结果，并且生成文本无法控制的问题，通过输入自然语言命令，PEER 可以对生成文本进行修改。
Cluade	Anthropic	与 ChatGPT 一样，不过强化学习阶段采用原发人工智能方法-AI 反馈强化学习 (RLAIF)，基于偏好模型而非人工反馈来进行训练	与 ChatGPT 相比，Claude 能更清晰地拒绝不恰当请求，句子之间衔接的也更自然。当遇到超出能力范围的问题时，Claude 会主动坦白。不过遇到代码生成或推理问题时，Claude 的表现就不如 ChatGPT。

资料来源：量子位，国联证券研究所整理

聊天机器人 Cluade 有超越 ChatGPT 之势。Claude 是由 GPT-3 核心成员出走成立的 Anthropic 公司开发，当前还处于内测阶段。不过根据 Scale Sep11book 团队成员内部测试结果显示，相比 ChatGPT，Claude 在部分任务的表现更优，在测试的 12 项任务中有 8 项更强。据 silk-news 消息，Anthropic 即将获得约 3 亿美元融资，成功融资后公司估值有望达到 50 亿美元。

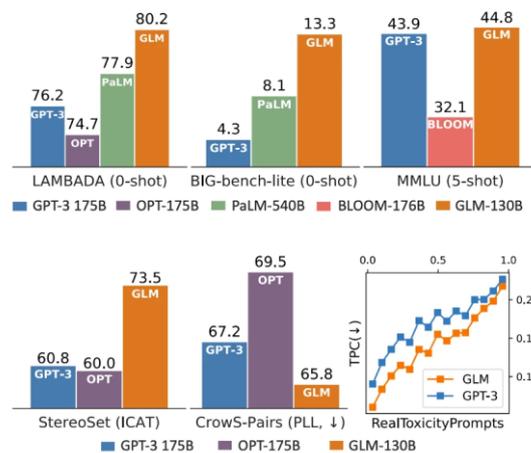
图表 9: Cluade 与 ChatGPT 功能对比



资料来源：量子位，国联证券研究所

国外头部机构引领技术和商业应用，国内整体进度落后 2~3 年。根据北京智源人工智能研究院介绍，当前国外研究机构在这个领域具有较大优势，领先国内 2~3 年，尤其是基础技术研究方面。如 2022 年清华大学联合智谱 AI 发布的中英双语模型 GLM-130B，参数规模与 OpenAI 在 2020 年发布的 GPT-3 接近，拥有 1300 亿参数的开源开放中英双语双向稠密模型。过去 3 年来，OpenAI 搭建大模型基础框架，通过 GPT-1/GPT-2/GPT-3/ChatGPT 引领行业发展。虽然国内最近两年也发布了很多相关大模型，但本质上还是在 OpenAI 的基础上实现的，原创性及创新性方面还需持续提升。

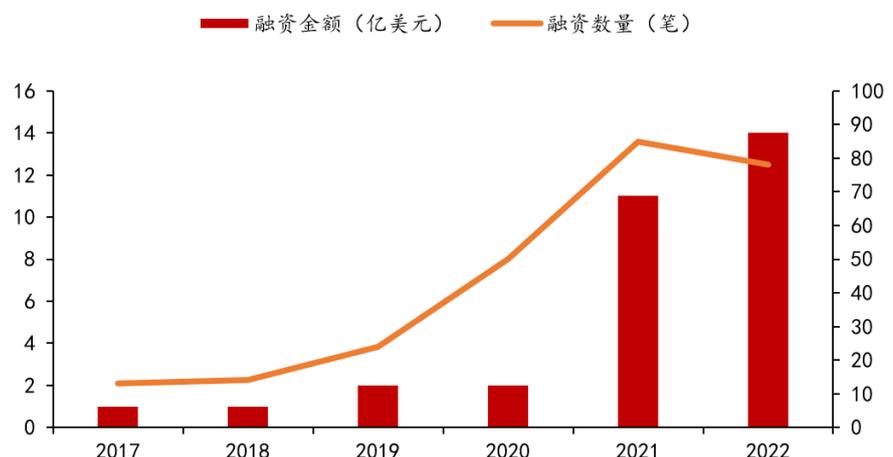
图表 10: GLM-130B 与 GPT-3、OPT-175B 对比



资料来源：腾讯新闻，国联证券研究所

生成式 AI 赛道火热，资金扎堆涌入。根据 PitchBook 统计数据，2022 年生成式 AI 赛道总共获得约 14 亿美元的融资，几乎达到了过去 5 年的总和。不仅包括 OpenAI、Stability AI 等明星公司，其他初创企业如 Jasper、Regie.AI、Replika 等均获得资本青睐，获得上亿美元融资。

图表 11: 2017-2022 年生成式 AI 赛道融资金额（亿美元）及数量（笔）



资料来源：PitchBook，国联证券研究所整理

2.2 ChatGPT 应用前景广阔

ChatGPT 具有广泛的应用空间。ChatGPT 主要应用包括如下场景：

对话机器人：由于 ChatGPT 强大的语言理解能力，可广泛应用于多种对话问答场景，包括智能客服、虚拟人、机器人、游戏 NPC 等应用领域。

图表 12: ChatGPT 在游戏中的应用



资料来源：量子位，国联证券研究所

智能创作：除了擅长对话问答之外，ChatGPT 还具备强大的文本内容创作能力，可用于创意写作（诗歌、新闻、小说、学术等）、命题写作（风格模仿、文本续写、主题拟定等）、摘要生成（学术类、小说类、新闻类等）等。尽管 ChatGPT 只是一个对话式的语言模型，本身不能生成多模态内容，但可以把它输出的结果作为一个中间变量输入其他模型，从而进一步拓展其应用。例如，通过 ChatGPT 和 Stable Diffusion 的结合使用，能够生成艺术性极强的画作。

图表 13: ChatGPT 结合 Stable Diffusion 生成的图片



资料来源：机器之心 Pro，国联证券研究所

欢迎访问：电子书学习和下载网站 (<https://www.shgis.com>)

文档名称：【国联证券】海外ChatGPT专题：ChatGPT风口已至，商业化落地加速.pdf

请登录 <https://shgis.com/post/1672.html> 下载完整文档。

手机端请扫码查看：

