



国泰君安证券  
GUOTAI JUNAN SECURITIES

行业专题研究

2023年02月07日

# ChatGPT研究框架（2023）

## 摘要

01

### ChatGPT市场反应热烈，国内外巨头纷纷入场

据统计，ChatGPT日活跃用户数的增速远超Instagram，1月份平均每天有超过1300万名独立访问者使用ChatGPT，是去年12月份的两倍多；国内外科技巨头都非常重视ChatGPT引发的科技浪潮，积极布局生成式AI，国内厂商（百度、腾讯等）也高度关注ChatGPT，积极探索前沿技术，相关深度应用也即将推出。

02

### ChatGPT经历多类技术路线演化，逐步成熟与完善

ChatGPT所能实现的人类意图，来自于机器学习、神经网络以及Transformer模型的多种技术模型积累。Transformer建模方法成熟以后，使用一套统一的工具来开发各种模态的基础模型这种理念得以成熟，随后GPT-1、GPT-2、GPT-3模型持续演化升级，最终孵化出ChatGPT文本对话应用。

03

### AIGC跨模态产业生态逐步成熟，商用落地未来可期

AIGC产业生态当前在文本、音频、视频等多模态交互功能上持续演化升级，奠定了多场景的商用基础。跨模态生成技术也有望成为真正实现认知和决策智能的转折点。

04

### ChatGPT乘东风，商业架构日益清晰

随着ChatGPT Plus发布，商业化序幕已经拉开。ChatGPT在传媒、影视、营销、娱乐以及数实共生助力产业升级等领域均可产生极大助益，提升生产力曲线，多维度赋能虚拟经济和实体经济。

# 目录 / CONTENTS

01 市场概况：ChatGPT——AI平民化里程碑

02 技术路径：基于人类反馈系统，ChatGPT助力跨模态AI生成应用

03 行业进程：AIGC多模态交互功能持续演化，奠定多场景商用基础

03 商业方向：多领域多功能应用密集落地，ChatGPT商用前景可期

01

# 市场概况：ChatGPT——AI平民化的里程碑

## 01

## OpenAI在成立之初便备受资本瞩目，与微软合作加速商业化进程

- ChatGPT是由OpenAI 团队研发创造，OpenAI是由创业家埃隆·马斯克、美国创业孵化器Y Combinator总裁阿尔特曼、全球在线支付平台PayPal联合创始人彼得·蒂尔等人于2015年在旧金山创立的一家非盈利的AI研究公司，拥有多位硅谷重量级人物的资金支持，启动资金高达10亿美金；OpenAI的创立目标是与其它机构合作进行AI的相关研究，并开放研究成果以促进AI技术的发展

图1：OpenAI发展势头强劲，商业化趋势明显



请参阅附注免责声明

## 01

## OpenAI 的ChatGPT是生成式人工智能技术（AIGC）浪潮的一部分

图2：随着算法的不断迭代，生成式人工智能技术（AIGC）不断发展

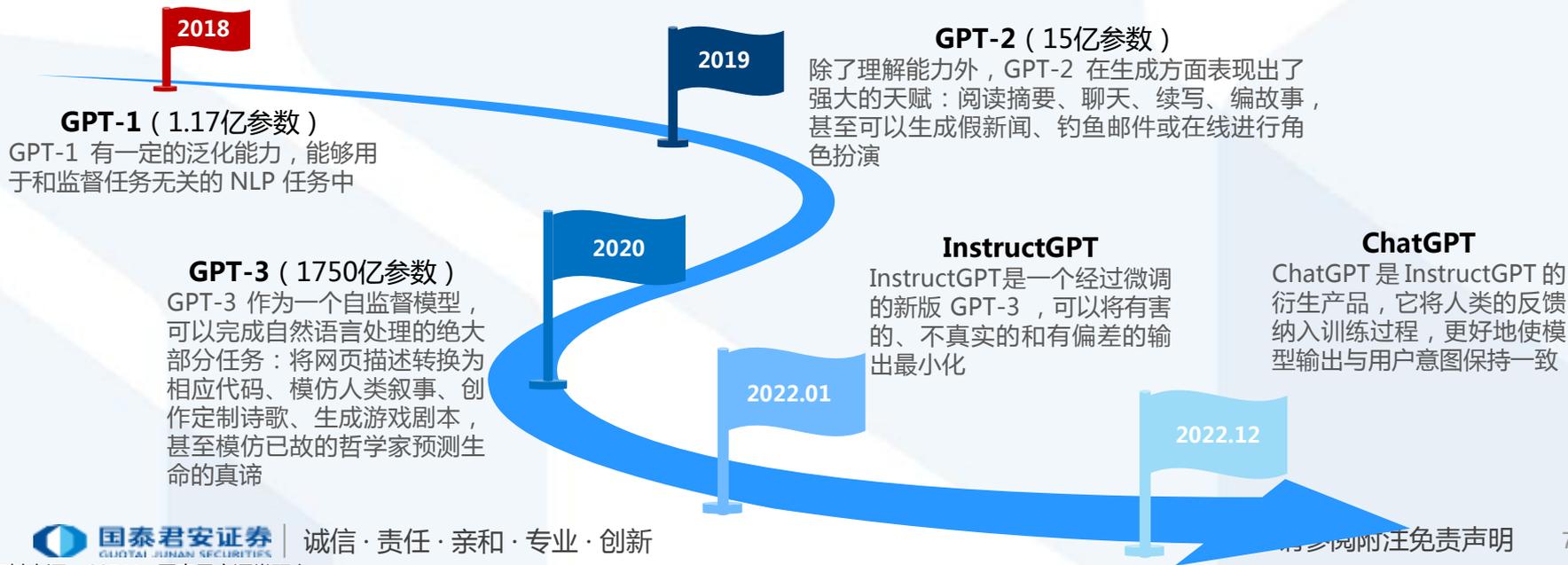


# 01

## ChatGPT是在GPT基础上进一步开发的自然语言处理模型

- GPT模型是一种自然语言处理（NLP）模型，使用多层变换器（Transformer）来预测下一个单词的概率分布，通过训练在大型文本语料库上学习到的语言模式来生成自然语言文本
- 从 GPT-1 到 GPT-3 智能化程度不断提升，ChatGPT 的到来也是 GPT-4 正式推出之前的序章

图3：ChatGPT由GPT-1到InstructGPT模型的不成熟而逐步形成



## 01

## ChatGPT发布后用户数持续暴涨，市场影响力迅速提升

- 根据UBS发布的研究报告显示，ChatGPT在1月份的月活跃用户数已达1亿，成为史上用户数增长最快的消费者应用
- ✓ 相比之下，TikTok花了九个月的时间月活跃用户数才破亿，而Instagram则花费了两年半的时间，同时，根据Similar Web的披露信息，Spotify在四年半后仅积累了1亿月活跃用户
- 根据Similar Web的数据，1月份平均每天有超过1300万名独立访问者使用ChatGPT，是去年12月份的两倍多

图4：ChatGPT日活跃用户数的增速远超Instagram

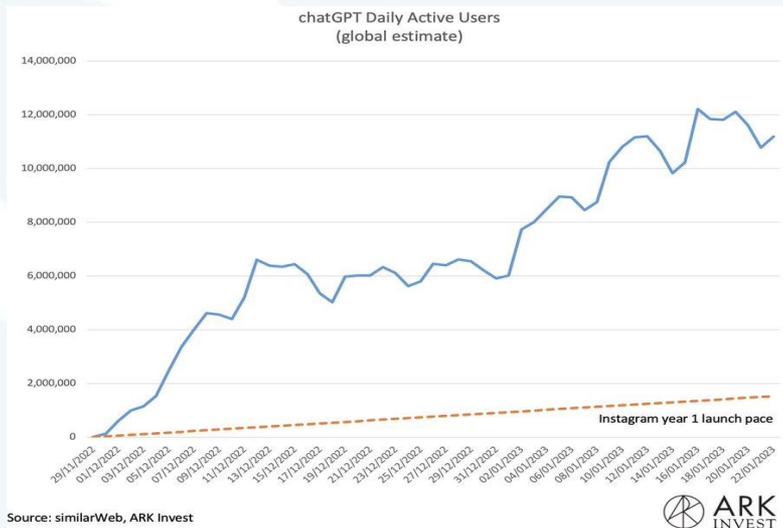


图5：对比各大热门平台月活跃用户数破亿所需时长，ChatGPT的成长速度惊人





# 01 ChatGPT具备诸多先进性特征

- ChatGPT嵌入了人类反馈强化学习以及人工监督微调，因而具备理解上下文、连贯性等诸多先进特征，解锁了海量应用场景
- ✓ 当前，ChatGPT所利用的数据集只截止到2021年。在对话中，ChatGPT会主动记忆先前的对话内容信息(上下文理解)，用来辅助假设性的问题的回复，因而ChatGPT也可实现连续对话，提升了交互模式下的用户体验。同时，ChatGPT也会屏蔽敏感信息，对于不能回答的内容也能给予相关建议。

图7：ChatGPT提升的核心点如下所示



## 01

## 科技巨头不断下注AI行业，ChatGPT带动新一轮AI发展浪潮

- 2023年初，微软和谷歌均宣布裁员计划，但都加大了在AI行业的投入

图8：科技巨头纷纷加大ChatGPT相关投入



- ✓ 1月18日，微软宣布将于3月底前在全球裁员1万人，约占公司员工总数的5%
- ✓ 1月23日，微软宣布对 ChatGPT 创建者 OpenAI 进行第三轮投资，将向OpenAI进行为期多年、金额达数十亿美元的投资，以加速其在人工智能领域的技术突破
- ✓ 此外，微软还计划在未来几周内发布其 Bing 搜索引擎的 GPT-4 版本，预计该版本将比 ChatGPT 模型更快、更强大

Alphabet

- ✓ 1月20日，谷歌母公司Alphabet宣布将裁员约1.2万人，占公司员工总数的6%以上，同时公司CEO表示，公司将会进一步聚焦AI技术核心业务的研发投入
- ✓ 2月4日，谷歌注资3亿美元投资ChatGPT竞品——Anthropic，和Open AI一样，Anthropic也在开发生成式人工智能

资料来源：UBS，《财富》，华尔街见闻，国泰君安证券研究

国泰君安证券  
GUOTAI JUNAN SECURITIES

诚信·责任·亲和·专业·创新

请参阅附注免责声明

## 01

## 国内外科技巨头积极布局生成式AI，部分公司已有成型产品

## ■ 国内外科技巨头都非常重视ChatGPT引发的科技浪潮，积极布局生成式AI

表1：国内外科技公司积极布局生成式AI

	公司名称	相关布局
国外公司	谷歌	面对ChatGPT构成的威胁，注资3亿美元投资竞品Anthropic
	微软	作为OpenAI的最大投资方，利用ChatGPT提高产品竞争力
	亚马逊	ChatGPT受到重点关注，已广泛运用在各种工作职能中
国内公司	百度	预计在3月份完成其ChatGPT产品——文心一言（ERNIE Bot）的内测，并面向公众开放
	腾讯	目前已经公布一项人机对话专利，能够实现机器与用户之间自然且顺畅的沟通

资料来源：各公司公告，国泰君安证券研究

## 01

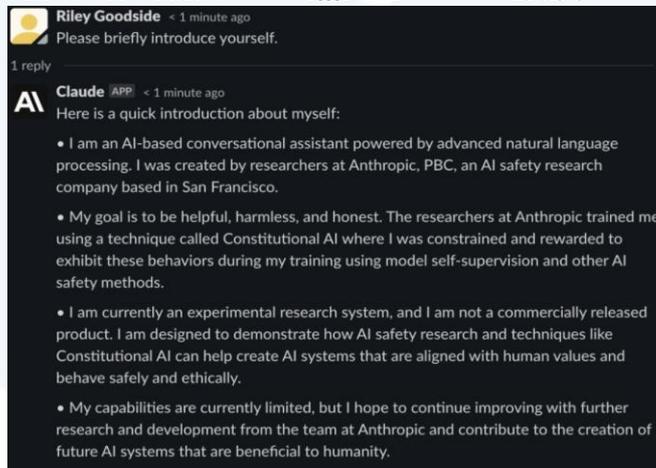
## 谷歌：面对ChatGPT构成的威胁，注资3亿美元投资竞品Anthropic

- 在ChatGPT发布后，谷歌CEO在公司内部发布了“红色警报”（Code Red），敦促团队解决ChatGPT对公司搜索引擎业务构成的威胁，同时批准了在谷歌搜索引擎中加入AI聊天机器人的计划
- ✓ 2月4日，谷歌注资3亿美元投资ChatGPT竞品——Anthropic，谷歌将获得约10%的股份，Anthropic计划将次轮资金用于购买谷歌云计算部门的计算资源；Anthropic开发了一款名为Claude的智能聊天机器人，据称可与ChatGPT相媲美（仍未发布）
- ✓ Anthropic和Open AI渊源颇深，其联合创始人曾担任OpenAI 研究副总裁

图9：Anthropic的联合创始人曾担任OpenAI 研究副总裁



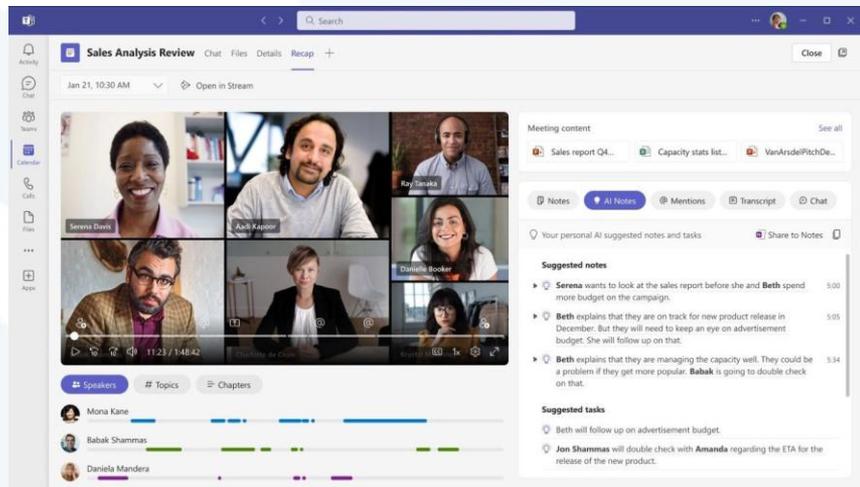
图10：聊天机器人Claude据称可与ChatGPT相媲美



# 01 微软：OpenAI的最大投资方，开始利用ChatGPT提高产品竞争力

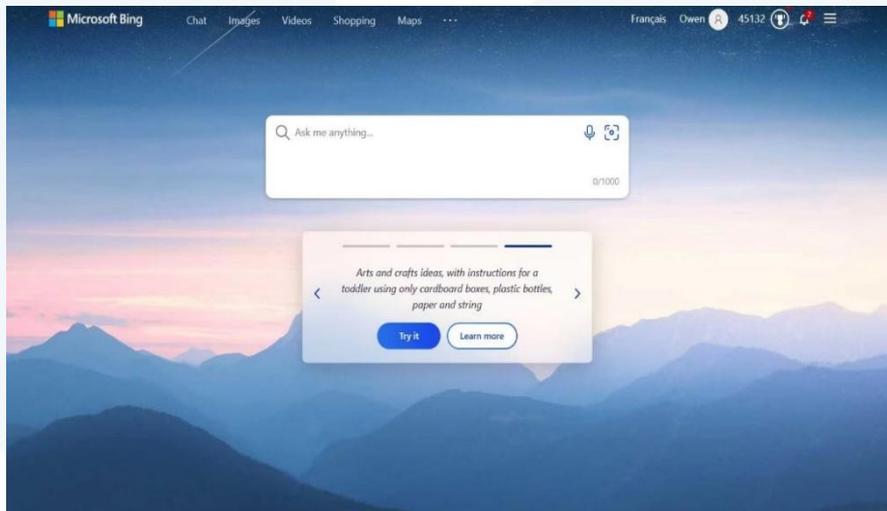
- 微软将ChatGPT视为新一代技术革命，将ChatGPT整合进Bing搜索引擎、Office全家桶、Azure云服务、Teams程序等产品中
- ✓ 微软近期宣布推出视频会议及远程协作平台的高级付费版Microsoft Teams Premium，订阅者可享用OpenAI GPT提供支持的大型语言模型技术，用AI自动生成会议笔记，此举或对Zoom、谷歌会议等平台形成巨大冲击。

图11：Microsoft Teams Premium 使用 ChatGPT 做笔记，使远程协作更高效



资料来源：微软官网，国泰君安证券研究

图12：集成了ChatGPT的新版 Bing 曾短暂上线，搜索框类似一个聊天框



资料来源：Windows Central，国泰君安证券研究

## 01

# 亚马逊：ChatGPT受到重点关注，已广泛运用在各种工作职能中

- ChatGPT 已经被亚马逊用于各种不同的工作职能中，包括回答面试问题、编写软件代码和创建培训文档等
- ✓ 公司内部员工在 Slack 表示，亚马逊 Amazon Web Services (AWS) 云部门已经成立了一个小型工作组，以更好地了解人工智能对其业务的影响。

图13：ChatGPT 已经被亚马逊用于各种不同的工作职能中



资料来源：Business Insider，新浪科技，国泰君安证券研究

图14：亚马逊AWS云部门的职能与人工智能联系紧密



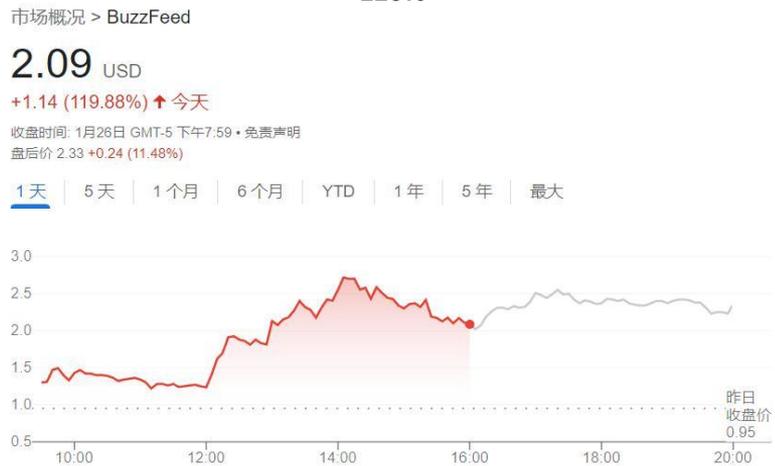
资料来源：Encaptechno，国泰君安证券研究

## 01

## 美国新媒体巨头Buzzfeed踩准ChatGPT风口，两天内股价飙升3倍

- 1月29日，美国新媒体巨头Buzzfeed宣布计划采用ChatGPT协助内容创作，其股价一夜间暴涨近120%，两天内飙升逾300%，成交量突破4.38亿股（其月平均成交量不足2500万股）
- 消息公布后同类型公司股票成交量也迎来历史高峰：C3.ai本月成交量超过7200万股，为自去年6月以来最多；语音AI软件公司SoundHound AI的成交量约为6450万股，几乎是其月平均值的三倍

图15：Buzzfeed宣布计划采用ChatGPT协助内容创作后，股价一夜间暴涨近120%



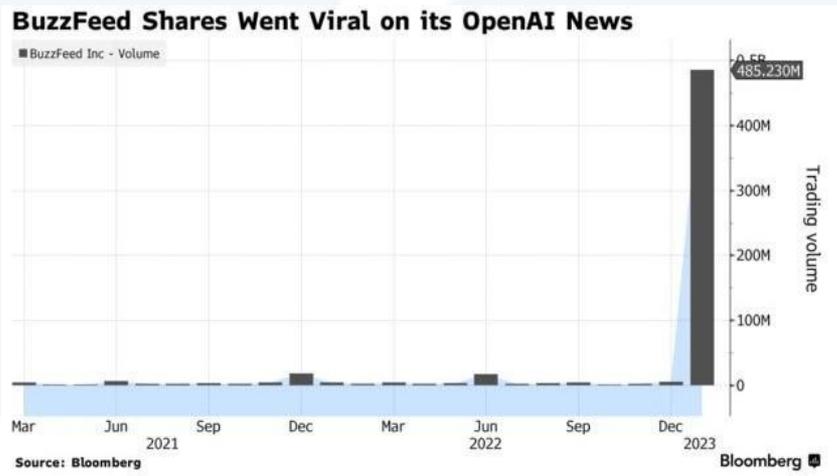
资料来源：澎湃新闻，国泰君安证券研究



国泰君安证券  
GUOTAI JUNAN SECURITIES

诚信·责任·亲和·专业·创新

图16：Buzzfeed宣布计划采用ChatGPT协助内容创作后，股票成交量暴涨



资料来源：腾讯新闻，国泰君安证券研究

请参阅附注免责声明 16

# 01

## Stability AI : Stable Diffusion大热，Open AI在图片生成AI大有可为

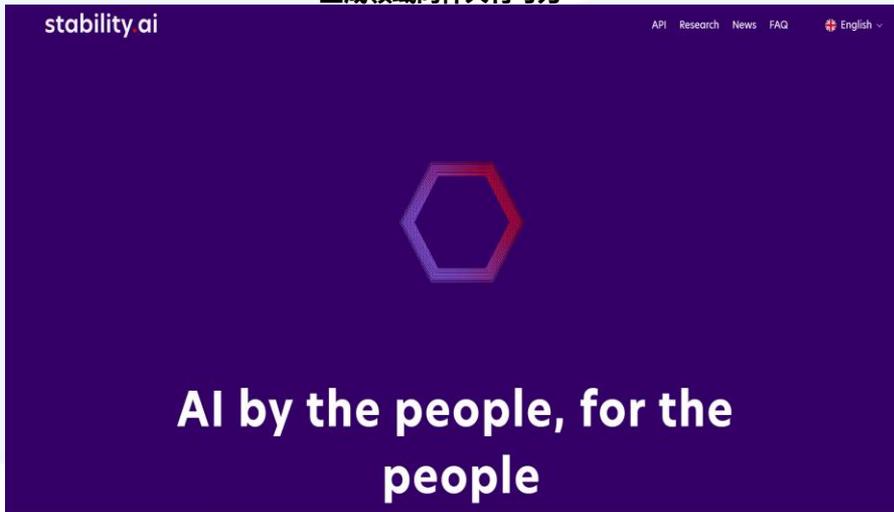
- **Stability AI有着与Open AI相同的创业理念：构建开源AI项目，促进AI发展，其成功证明Open AI在图片生成领域同样大有可为**
- ✓ 公司的开源模型Stable Diffusion可以根据文字生成图片，只需要几秒钟，就可以生成分辨率、清晰度高，同时不失真实性和艺术性的图片。

图17：用户使用基于Stable Diffusion算法构建的工具创建了数百万张图像



资料来源：网易新闻，国泰君安证券研究

图18：Stability AI有着与Open AI相同的创业理念，其成功证明Open AI在图片生成领域同样大有可为



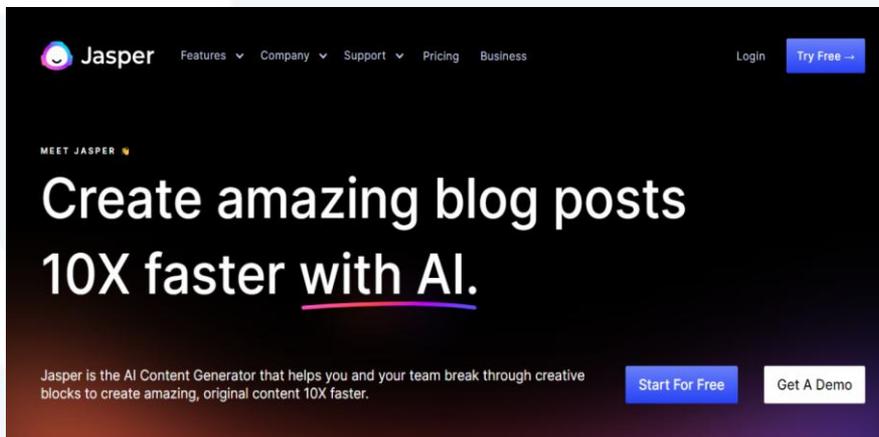
资料来源：Stability AI官网，国泰君安证券研究

## 01

## Jasper：采用同类底层技术，进一步证明ChatGPT的巨大商业潜力

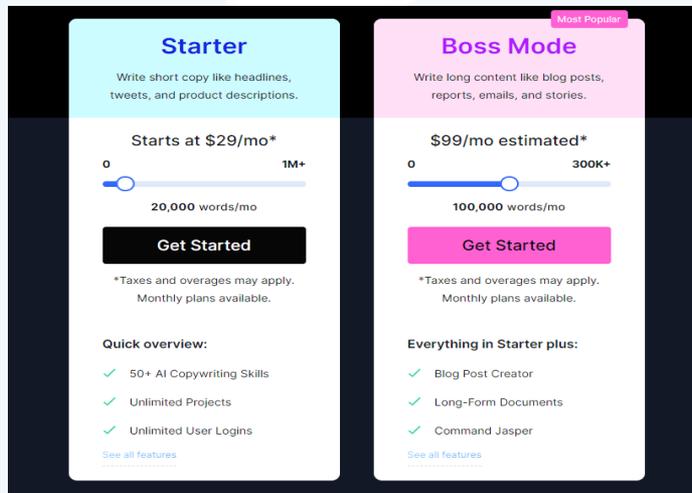
- 文案自动生成平台Jasper，其技术底层是 OpenAI 的 GPT-3，在成立仅 18 个月后就达到了 15 亿美元的高估值
- ✓ IBM、Autodesk等巨头公司均是Jasper的付费用户，足以证明ChatGPT底层技术具备巨大的商业潜力。
- ✓ ChatGPT出现后，其技术领先性和受欢迎程度对Jasper形成了强烈冲击。

图19：ChatGPT的出现，对采用同类底层技术的热门公司Jasper形成强烈冲击



资料来源：搜狐新闻，Jasper 官网，国泰君安证券研究

图20：相比ChatGPT的免费开源，Jasper需要付费使用



资料来源：Jasper 官网，国泰君安证券研究

## 01

## 国内厂商（百度&amp;腾讯）：高度关注ChatGPT，积极探索前沿技术

- **百度**：1月10日，百度宣布将升级百度搜索的“生成式搜索”能力，智能解答用户的搜索提问；2月7日，百度宣布将在3月份完成其ChatGPT产品的内测，面向公众开放，该项目名字为文心一言（ERNIE Bot）
- ✓ 百度指出，生成式AI和搜索引擎是互补关系而不是替代；据路透社报道，百度计划于3月将类似ChatGPT的AI对话服务作为独立应用推出，之后再逐步将其合并到搜索引擎中。
- **腾讯**：2月3日，腾讯公布一项人机对话专利，能够实现机器与用户之间自然且顺畅的沟通

图21：百度研发的ChatGPT产品“文心一言”将于3月份问世

## 中国版ChatGPT？百度官宣：“文心一言”下月面世

2023年02月07日 12:44 来源：中国经济网

[推荐朋友] [打印本稿] [字号大中小]

中国经济网北京2月7日讯（记者 李方）ChatGPT近期不断刷屏，国内互联网大厂动向如何？记者从百度获悉，该项目名字确定为“文心一言”，英文名ERNIE Bot，三月份完成内测，面向公众开放。目前，文心一言在做上线前的冲刺。

资料来源：中国经济网，国泰君安证券研究

图22：腾讯发布相关专利，积极涉足ChatGPT领域

**摘要**

本申请实施例提供了一种人机对话方法、装置、设备及计算机可读存储介质。其中，方法包括：当人机对话被触发时，在预设文本库中获取用于进行所述人机对话的参考文本；根据所述参考文本，确定出在所述人机对话过程中用于生成应答信息的背景信息；接收所述人机对话的对话信息；根据所述对话信息和背景信息，确定出对话信息对应的应答信息。

```

    graph TD
      S401[人机对话被触发时，在预设文本库中获取用于进行人机对话的参考文本] --> S402[根据对话参考文本，确定出在人机对话过程中用于生成应答信息的背景信息]
      S402 --> S403[接收人机对话的对话信息]
      S403 --> S404[根据对话信息和背景信息，确定出对话信息对应的应答信息]
  
```

**法律状态**

序号	法律状态公告日	法律状态	法律状态信息
1	2023-02-03	授权	授权
2	2024-03-12	实质审查的生效	实质审查的生效；IPC(主分类):G06F16/33；申请日:20200316
3	2020-07-10	公布	公布

资料来源：腾讯新闻，国泰君安证券研究

## 01

## AIGC创业公司大比拼，国外ChatGPT的优势遥遥领先并有望延续

## ■ 国内对比国外AIGC创业公司的相关产品，国外ChatGPT的优势遥遥领先并有望延续

表2：国外创业公司涉及的AIGC产品领域十分丰富，相关应用日渐成熟

公司名称	主要AIGC产品	产品领域
OpenAI	GPT-3、ChatGPT	文本、图片
Stability AI	Stable Diffusion、Dreamstudio	图片、音频
Midjourney	Midjourney	图片
Jasper	Jasper	文本
Copi.ai	Copi.ai	文本
Copysmith	Copysmith	文本

资料来源：各公司官网，国泰君安证券研究



## 01

## AI 需要大量资金、人力投入和数据积累，国内市场中巨头更具优势

- 人工智能不仅需要巨大的投入，还需要庞大的用户数据基础，只有互联网巨头才有这个能力产出伟大的产品
- 国外是微软、谷歌、亚马逊，而国内则是百度、腾讯等互联网巨头最具潜力；相比国外巨头，国内巨头正投入大量资金和人力成本快速发展人工智能技术，在没有硝烟的人工智能竞赛中，中国企业也将异军突起

图23：百度构建了云智共生的生态，助力AI发展



资料来源：百度官网，国泰君安证券研究

图24：腾讯的AI 产品在游戏场景中不断提升人机协作与内容生成能力



资料来源：腾讯AI LAB官网，国泰君安证券研究

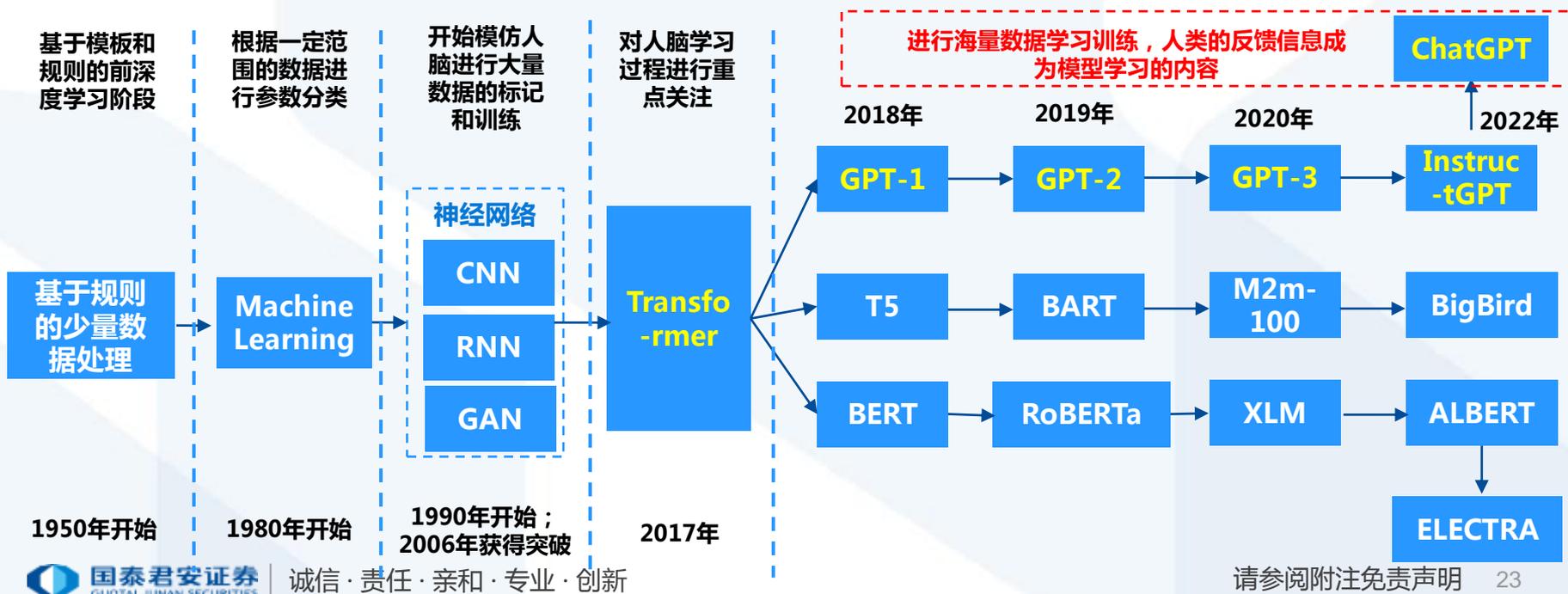
# 02

## 技术路径：基于人类反馈系统， ChatGPT助力跨模态AI生成应用

## 02 ChatGPT经历多类技术路线演化，逐步成熟与完善

- ChatGPT所能实现的人类意图，来自于机器学习、神经网络以及Transformer模型的多种技术模型积累

图25：ChatGPT经过多类技术积累，最终形成针对人类反馈信息学习的大规模预训练语言模型



## 02

## ChatGPT模型在以往模型的基础上有了多方面的显著提升

图26：ChatGPT是基于Transformer架构的语言模型，它在以往大语言模型（如ELMo和GPT-2）的基础上有诸多性能提升

### 更大的语料库

ChatGPT使用了更大的语料库，以更好地捕捉人类语言的复杂性。

### 更高的计算能力

ChatGPT使用了更高的计算资源，以获得更好的训练效果。

### 更加通用的预训练

ChatGPT的预训练是通用的，因此它可以更好地适应各种不同的任务。

### 更高的准确性

ChatGPT的训练效果比以往的大语言模型更好，因此它的准确性更高。

### 更高的适应性

ChatGPT具有较高的适应性，可以根据不同的场景和任务进行微调，以提高其在特定领域的效果。

### 更强的自我学习能力

ChatGPT具有自我学习能力，可以在不断接触新语料的过程中持续提高自己的性能。



## 02 Transformer的应用标志着基础模型时代的开始

### ■ 转移学习(Transfer Learning)使基础模型成为可能

- ✓ 技术层面上，基础模型通过转移学习(Transfer Learning)(Thrun 1998)和规模(scale)得以实现。转移学习的思想是将从一项任务中学习到的“知识”（例如，图像中的对象识别）应用于另一项任务（例如，视频中的活动识别）。
- ✓ 在深度学习中，预训练又是转移学习的主要方法：在替代任务上训练模型（通常只是达到目的的一种手段），然后通过微调来适应感兴趣的下游任务。**转移学习(Transfer Learning)使基础模型成为可能。**

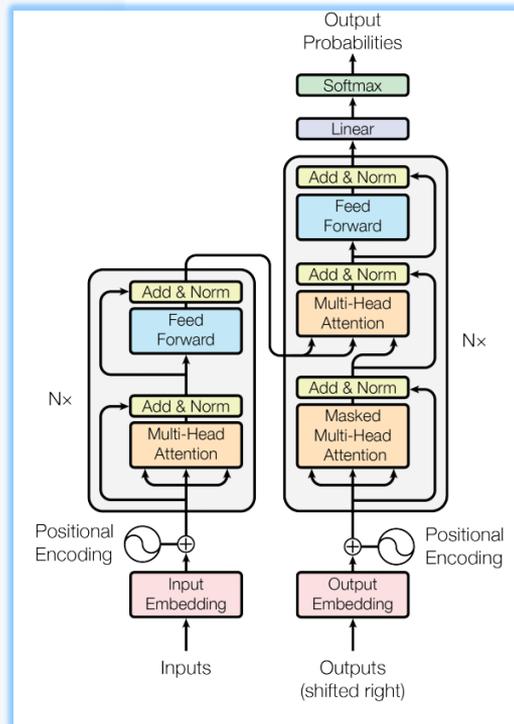
### ■ 大规模化(scale)使基础模型更强大，因而GPT模型得以形成

- ✓ 大规模需要三个要素：（i）计算机硬件的改进——例如，GPU吞吐量和内存在过去四年中增加了10倍；（ii）**Transformer模型架构的开发**（Vaswani et al. 2017），**该架构利用硬件的并行性来训练比以前更具表现力的模型**；以及（iii）**更多训练数据的可用性**。
- ✓ 基于Transformer的序列建模方法现在应用于文本、图像、语音、表格数据、蛋白质序列、有机分子和强化学习等，这些例子的逐步形成使得**使用一套统一的工具来开发各种模态的基础模型这种理念得以成熟**。例如，GPT-3（Brown et al. 2020）与GPT-2的15亿参数相比，GPT-3具有1750亿个参数，允许上下文学习，在上下文学习中，只需向下游任务提供提示（任务的自然语言描述），语言模型就可以适应下游任务，**这是产生的一种新兴属性**。

## 02 Transformer奠定了生成式AI领域的游戏规则

- Transformer摆脱了人工标注数据集的缺陷，模型在质量上更优、更易于并行化，所需训练时间明显更少
  - Transformer通过成功地将其应用于具有大量和有限训练数据的分析，可以很好地推广到其他任务
- ✓ 2017年，在Ashish Vaswani et.al 的论文《Attention Is All You Need》中，考虑到主导序列转导模型基于编码器-解码器配置中的复杂递归或卷积神经网络，性能最好的模型被证明还是通过注意力机制（attention mechanism）连接编码器和解码器，因而《Attention Is All You Need》中提出了一种新的简单架构——Transformer，它完全基于注意力机制，完全不用重复和卷积，因而这些模型在质量上更优，同时更易于并行化，并且需要的训练时间明显更少。
- ✓ Transformer出现以后，迅速取代了RNN系列变种，跻身主流模型架构基础。（RNN缺陷正在于流水线式的顺序计算）

图27：Transformer模型架构如下所示



资料来源：《Attention Is All You Need》,Ashish Vaswani et.al 2017

## 02

## Transformer实现的不同技术场景对应的不同技术原理

- Transformer架构可分为自回归系列（例如GPT-3，偏好生成性任务）、双向Transformer+Mask的自编码系列（例如BERT，偏好自然语言理解）、Encoder-decoder架构（例如T5，使用双向/单向attention，偏好条件文本生成）

图28：Transformer典型技术场景下的原理介绍如下所述

首先通过词嵌入(Word Embedding)将字、词、句进行区分，然后基于特征评分、序列标注、分类模型等提取内容特征计算相关文本单元权重其次选择相应的文本单元子集组成摘要候选集，完成内容选择，最后针对字数要求等限定条件，对候选集的内容进行整理形成最终摘要，完成内容组织。其细分路径又包含生成式文本摘要(AATS)，即形成抽象认知并创造新词灵活概括，和抽取式文本摘要(EATS)，即直接抽取原始素材并拼接成简单概要

对话式文本生成适用于智能客服等任务型和闲聊型机器人等非任务型人机交互场景，可分为管道模式及端对端模式。结构性的文本生成，首先通过注意力机制、多层感知器等系统进行语句内容预选，对数值、时间等类型数据进行推理。增强数据间的结构信息。其次通过Transformer等模式结合上下文进行推导，生成最终文本。



## 02

## GPT-1：借助预训练，进行无监督训练和有监督微调

- GPT-1模型基于Transformer解除了顺序关联和依赖性的前提，采用生成式模型方式，重点考虑了从原始文本中有效学习的能力，这对于减轻自然语言处理（NLP）中对监督学习的依赖至关重要
- ✓ GPT（Generative Pre-training Transformer）于2018年6月由OpenAI首次提出。GPT模型考虑到在自然语言理解中有大量不同的任务，尽管大量的未标记文本语料库非常丰富，但用于学习这些特定任务的标记数据却很少，这使得经过区分训练的模型很难充分执行。同时，大多数深度学习方法需要大量手动标记的数据，这限制了它们在许多缺少注释资源的领域的适用性。
- ✓ 在考虑以上局限性的前提下，GPT论文中证明，通过对未标记文本的不同语料库进行语言模型的**生成性预训练**，然后对每个特定任务进行**区分性微调**，可以实现这些任务上的巨大收益。和之前方法不同，GPT在微调期间使用任务感知输入转换，以实现有效的传输，同时对模型架构的更改最小。

图29：GPT-1模型的核心手段是预训练（Pre-training）

## 3.1 Unsupervised pre-training

Given an unsupervised corpus of tokens  $U = \{u_1, \dots, u_n\}$ , we use a standard language modeling objective to maximize the following likelihood:

$$L_U(U) = \sum_{i=1}^n \log P(u_i | u_{1:i-1}, u_{i+1:n}) \quad (1)$$

where  $k$  is the size of the context window, and the conditional probability  $P$  is modeled using a neural network with parameters  $\theta$ . These parameters are trained using stochastic gradient descent [51].

In our experiments, we use a multi-layer Transformer decoder [34] for the language model, which is a variant of the transformer [62]. This model applies a multi-headed self-attention operation over the input context tokens followed by position-wise feedforward layers to produce an output distribution over target tokens:

$$h_o = U W_o + W_p \quad (2)$$

$$h_o = \text{transformer\_block}(h_{i-1}) \forall i \in [1, n]$$

$$P(u) = \text{softmax}(h_o W_p)$$

where  $U = \{u_1, \dots, u_{n-1}\}$  is the context vector of tokens,  $n$  is the number of layers,  $W_o$  is the token embedding matrix, and  $W_p$  is the position embedding matrix.

**无监督预训练**  
(Unsupervised pre-training)

不需要标注数据集，即大规模自学阶段，在保证AI算力充足的条件下，根据attention机制进行自学

**有监督微调**  
(Supervised fine-tuning)

微调，用来修正模型理解力。即小规模指导过程，让AI在小样本数据下进行调整

## 3.2 Supervised fine-tuning

After training the model with the objective in Eq. 1, we adapt the parameters to the supervised target task. We assume a labeled dataset  $C$ , where each instance consists of a sequence of input tokens,  $x^1, \dots, x^m$ , along with a label  $y$ . The inputs are passed through our pre-trained model to obtain the final transformer block's activation  $h^m$ , which is then fed into an added linear output layer with parameters  $W_y$  to predict  $y$ :

$$P(y | x^1, \dots, x^m) = \text{softmax}(h^m W_y). \quad (3)$$

This gives us the following objective to maximize:

$$L_C(C) = \sum_{i=1}^m \log P(y^i | x^1, \dots, x^m). \quad (4)$$

We additionally found that including language modeling as an auxiliary objective to the fine-tuning helped learning by (a) improving generalization of the supervised model, and (b) accelerating convergence. This is in line with prior work [50, 43], who also observed improved performance with such an auxiliary objective. Specifically, we optimize the following objective (with weight  $\lambda$ ):

$$L_C(C) = L_C(C) + \lambda \cdot L_U(U) \quad (5)$$

Overall, the only extra parameters we require during fine-tuning are  $W_y$ , and embeddings for delimiter tokens (described below in Section 3.3).

结合形成了一种使用无监督预训练和有监督微调相结合的语言理解任务的“半监督方法”

## 02

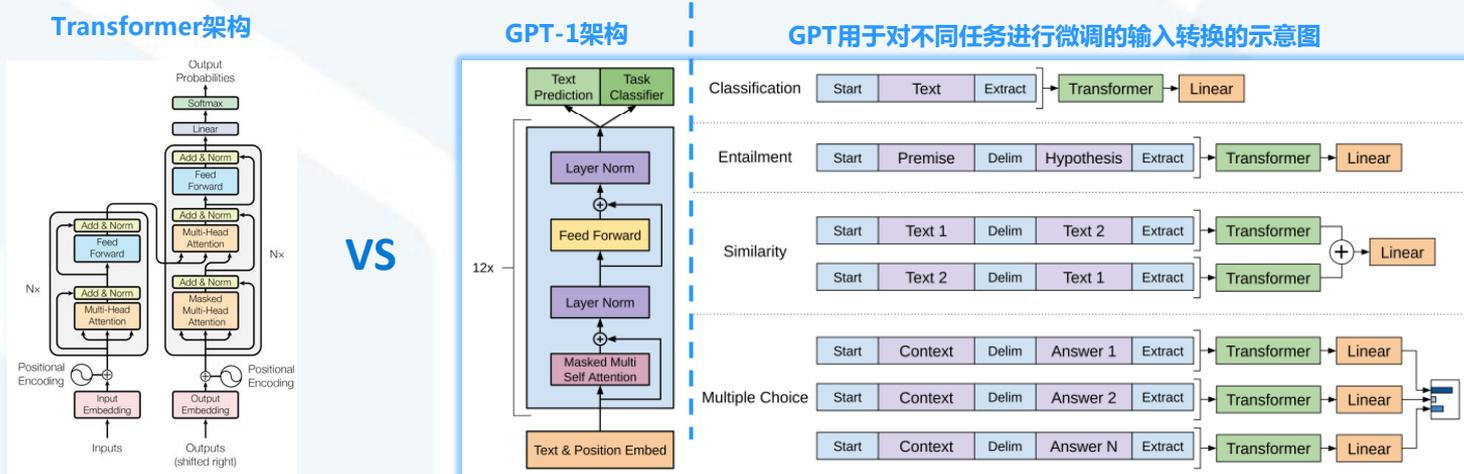
## GPT-1：模型更简化、计算加速，更适合自然语言生成任务（NLG）

## ■ GPT相比于Transformer等模型进行了显著简化

- ✓ 相比于Transformer，GPT训练了一个12层**仅decoder的解码器**（原Transformer模型中包含Encoder和Decoder两部分）。
- ✓ 相比于Google的BERT(Bidirectional Encoder Representations from Transformers,双向编码生成Transformer)，**GPT仅采用上文预测单词**（BERT采用了基于上下文双向的预测手段）。

注：ChatGPT的表现更贴近人类意图，部分因为一开始GPT是基于上文的预测，这更贴近人类的话语模式，因为人类言语无法基于将来的话来做分析。

图30：GPT-1模型相比于Transformer模型有了显著简化



## 02 GPT-2：采用多任务系统，基于GPT-1进行优化

### ■ GPT-2在GPT-1的基础上进行诸多改进，实现执行任务多样性，开始学习在不需要明确监督的情况下执行数量惊人的任务

- ✓ 在GPT-2阶段，OpenAI去掉了GPT-1阶段的有监督微调（fine-tuning），成为**无监督模型**。
- ✓ 大模型GPT-2是一个1.5B参数的Transformer，在其相关论文中它在8个测试语言建模数据集中的7个数据集上实现了当时最先进的结果。模型中，Transformer堆叠至48层。GPT-2的数据集增加到8 million的网页、大小40GB的文本。

图31：GPT-2通过调整原模型和采用多任务方式来让AI更贴近“通才”水平

- ✓ 机器学习系统通过使用大型数据集、大容量模型和监督学习的组合，在训练任务方面表现出色，然而这些系统较为脆弱，对数据分布和任务规范的轻微变化非常敏感，因而使得AI表现更像狭义专家，并非通才。

- ✓ 转向更通用的系统，使其可以执行许多任务，最终无需为每个任务手动创建和标记训练数据集。



- **GPT-2调整优化的目的**是为了解决零次学习问题（zero-shot）（注：zero-shot问题，就是针对AI在面对不认识的事物时，也能进行推理）
- **多任务模型的特点**：跟传统ML需要专门的标注数据集不同（从而训练出专业AI），多任务模型不采用专门AI手段，而是在海量数据喂养训练的基础上，适配任何任务形式。

Parameters	Layers	$d_{model}$
117M	12	768
345M	24	1024
762M	36	1280
1542M	48	1600

## 02

## GPT-2仍未解决应用中的诸多瓶颈

- GPT-2聚焦在无监督、zero-shot（零次学习）上，然而GPT-2训练结果也有不达预期之处，所存在的问题也亟待优化
- ✓ 在GPT-2阶段，尽管体系结构是任务无关的，但仍然需要任务特定的数据集和任务特定的微调：要在所需任务上实现强大的性能，通常需要对特定于该任务的数千到数十万个示例的数据集进行微调。

图32：GPT-2尚未解决诸多瓶颈问题

**存在的问题01：**  
从实用的角度来看，每一项新任务都需要一个标记示例的大数据集，这限制了语言模型的适用性；对于其中的许多任务（从纠正语法到生成抽象概念的示例，再到评论一个短篇故事等等），很难收集一个大型的监督训练数据集，特别是当每个新任务都必须重复该过程时。

**存在的问题02：**

- 预训练加微调范式中，可能在这种范式下实现的泛化可能很差，因为该模型过于特定于训练分布，并且在其之外无法很好地泛化。
- 微调模型在特定基准上的性能，即使名义上是人类水平，也可能夸大基础任务的实际性能。

**存在的问题03：**

因为人类学习大多数语言任务不需要大型受监督的数据集，当前NLP技术在概念上具有一定的局限性。

资料来源：《Language Models are Few-Shot Learners》论文，国泰君安证券研究

## 02

## GPT-3取得突破性进展，任务结果难以与人类作品区分开来

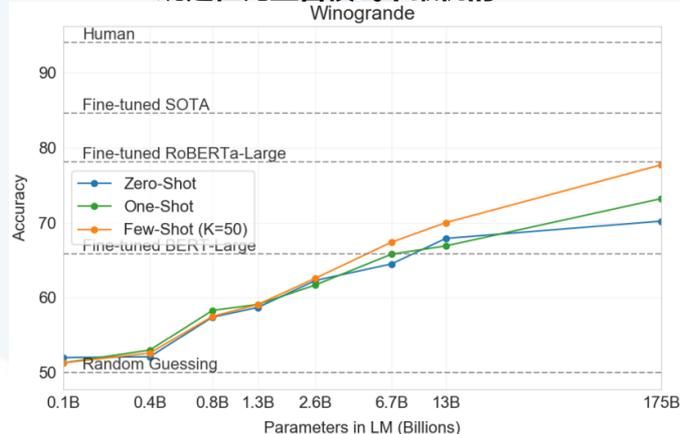
- GPT-3对GPT-2追求无监督与零次学习的特征进行了改进
  - GPT-3利用了过滤前45TB的压缩文本，在诸多NLP数据集中实现了强大性能
- ✓ GPT-3是一个具有**1750亿个参数**的自回归语言模型，比之前的任何非稀疏语言模型多10倍。对于所有任务（在few-shot设置下测试其性能），GPT-3都是在没有任何梯度更新或微调的情况下应用的，仅通过与模型的文本交互来指定任务和few-shot演示。
- ✓ GPT-3在许多NLP数据集上都有很强的性能（包括翻译、问题解答和完形填空任务），以及一些需要动态推理或领域适应的任务（如解译单词、在句子中使用一个新单词或执行三位数算术）。GPT-3可以生成新闻文章样本（已很难将其与人类撰写的文章区分开来）。

图33：GPT-3的模型参数在GPT-2的基础上增加110多倍

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

资料来源：《Language Models are Few-Shot Learners》，国泰君安证券研究

图34：GPT-3相关研究显示，few-shot(少量样本)的综合表现是在无监督模式下最优的



资料来源：《Language Models are Few-Shot Learners》，国泰君安证券研究

## 02 InstructGPT模型在GPT-3基础上进一步强化

- **InstructGPT使用来自人类反馈的强化学习方案RLHF（reinforcement learning from human feedback），通过对大语言模型进行微调，从而能够在参数减少的情况下，实现优于GPT-3的功能**
- ✓ **InstructGPT提出的背景**：使语言模型更大并不意味着它们能够更好地遵循用户的意图，例如大型语言模型可以生成不真实、有毒或对用户毫无帮助的输出，即这些模型与其用户不一致。另外，GPT-3虽然选择了少样本学习（few-shot）和继续坚持了GPT-2的无监督学习，但基于few-shot的效果，其稍逊于监督微调（fine-tuning）的方式。
- ✓ 基于以上背景，OpenAI在GPT-3基础上根据人类反馈的强化学习方案RHLF，训练出奖励模型（reward model）去训练学习模型（即：**用AI训练AI的思路**）
- ✓ **InstructGPT的训练步骤为**：**对GPT-3监督微调——训练奖励模型（reward model）——增强学习优化SFT（第二、第三步可以迭代循环多次）**

图35：InstructGPT训练三步骤各自对应的数据集规模如下图所示（labeler指的是OpenAI的标注人员，customer指GPT-3 API的用户）

SFT Data			RM Data			PPO Data		
split	source	size	split	source	size	split	source	size
train	labeler	11,295	train	labeler	6,623	train	customer	31,144
train	customer	1,430	train	customer	26,584	valid	customer	16,185
valid	labeler	1,550	valid	labeler	3,488			
valid	customer	103	valid	customer	14,399			

资料来源：《Training language models to follow instructions with human feedback》论文，稀土掘金，国泰君安证券研究

## 02

## ChatGPT核心技术优势：提升了理解人类思维的准确性

- InstructGPT与ChatGPT属于相同代际的模型，ChatGPT只是在InstructGPT的基础上增加了Chat属性，且开放了公众测试
- ChatGPT提升了理解人类思维的准确性的原因在于利用了基于人类反馈数据的系统进行模型训练

（注：根据官网介绍，GhatGPT也是基于InstructGPT构建，因而可以从InstructGPT来理解ChatGPT利用人类意图来增强模型效果）

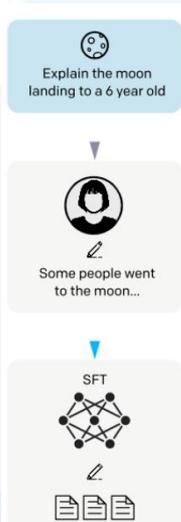
图36：基于人类反馈强化的核心训练流程如下所示：

### 步骤1：搜集说明数据，训练监督策略

搜集说明数据（基于prompt训练方式的数据集），训练监督策略

Labeler(标记者)揭示期望的输出行为

这个数据用来联合监督学习，对GPT-3进行微调



### 步骤2：搜集比较数据，训练一个奖励模型

对这个prompt训练的数据集和若干模型的结果进行抽样

Labeler(标记者)从最优到最差将输出结果进行排序

这个数据用来训练反馈模型



### 步骤3：搜集说明数据，使用增强学习优化模型

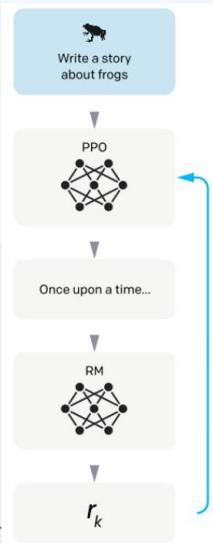
新的prompt从数据集中抽样

借助模型生成输出

反馈模型为输出计算一个反馈结果

反馈结果用来优化策略

请参阅附注免责



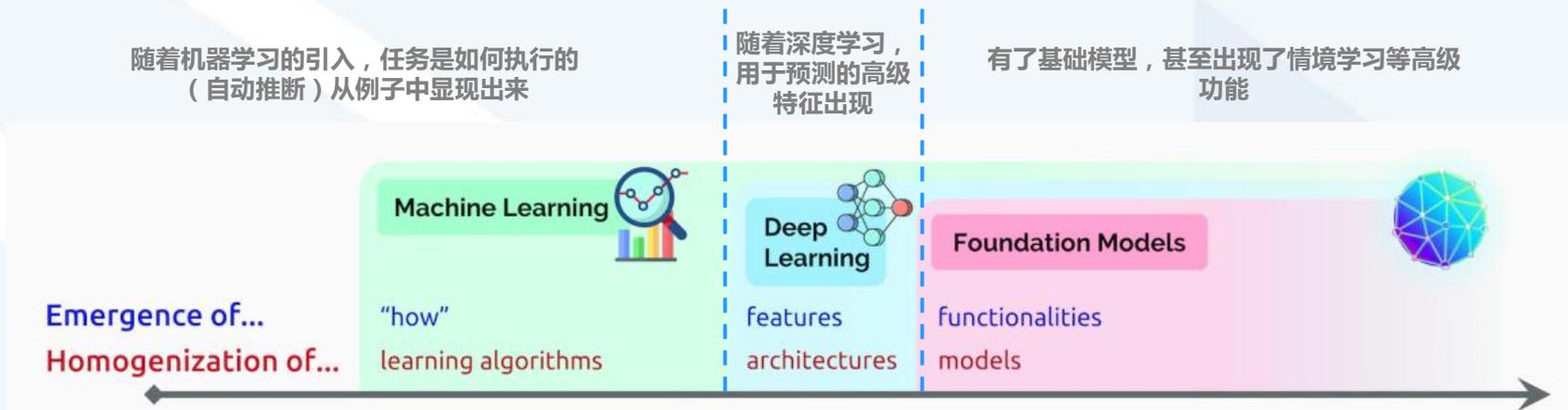
## 02

## ChatGPT得益于通用（基础）模型所构建 AI 系统的新范式

### ■ 基础模型（Foundation Model）在广泛的应用中整合构建机器学习系统的方法，它为许多任务提供了强大的杠杆作用

- ✓ 基础模型是在深度神经网络和自我监督学习的基础上演化而来。基础模型基于广泛数据（通常使用大规模自我监督）训练的任何模型，可以适应（例如微调）广泛的下游任务，目前例子包括BERT（Devlin et al.）、GPT-3（Brown et al. 2020）和CLIP（Radford et al. 2021）。
- ✓ 机器学习使学习算法同质化（例如，逻辑回归），深度学习使模型架构同质化（如卷积神经网络），而基础模型使模型本身同质化（比如，GPT-3）。

图37：人工智能的发展呈现同质化的过程

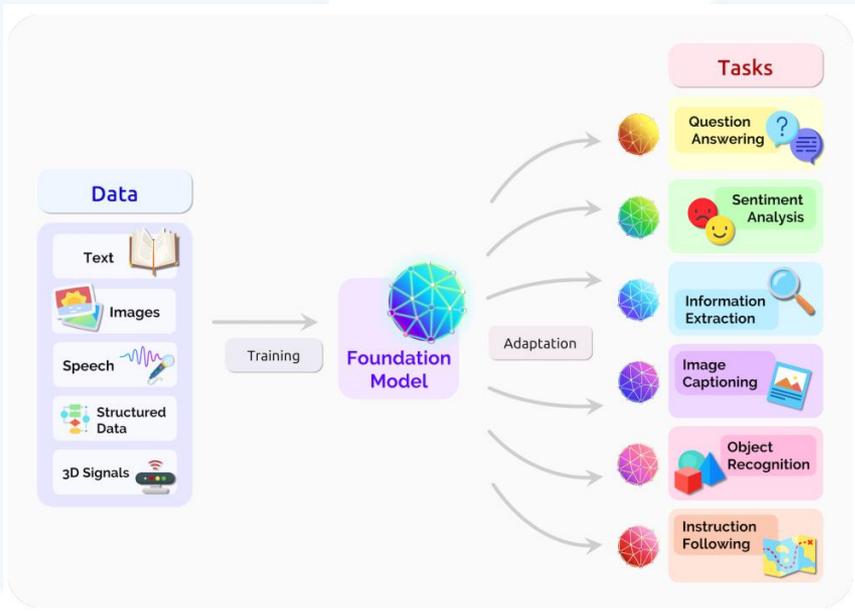


## 02 ChatGPT以基础模型为杠杆，可适用多类下游任务

### ■ ChatGPT采用了GPT3.5（InstructGPT）大规模预训练模型，在自然语言理解和作品生成上取得极大性能提升

- ✓ 鉴于传统NLP技术的局限问题，基于大语言模型（LLM）有助于充分利用海量无标注文本预训练，从而文本大模型在较小的数据集和零数据集场景下可以有较好的理解和生成能力。基于大模型的**无标准文本收集**，ChatGPT得以在情感分析、信息钻取、理解阅读等文本场景中**优势突出**。
- ✓ 随着训练模型数据量的增加，数据种类逐步丰富，模型规模以及参数量的增加，会进一步促进模型语义理解能力以及抽象学习能力的极大提升，实现ChatGPT的**数据飞轮效应（用更多数据可以训练出更好的模型，吸引更多用户，从而产生更多用户数据用于训练，形成良性循环）**。
- ✓ 研究发现，每增加参数都带来了文本合成和/或下游NLP任务的改进，有证据表明，日志丢失与许多下游任务密切相关，随着规模的增长，日志丢失呈现平稳的改善趋势。

图38：基础模型可以集中来自各种模态的所有数据的信息，然后这一模型可以适用于广泛的下游任务



资料来源：《On the Opportunities and Risks of Foundation Models》论文

## 02

## ChatGPT大模型架构也是ML发展到第三阶段的必然产物

- ML中的计算历史分为三个时代：前深度学习时代、深度学习时代和大规模时代，在大规模时代，训练高级ML系统的需求快速增长

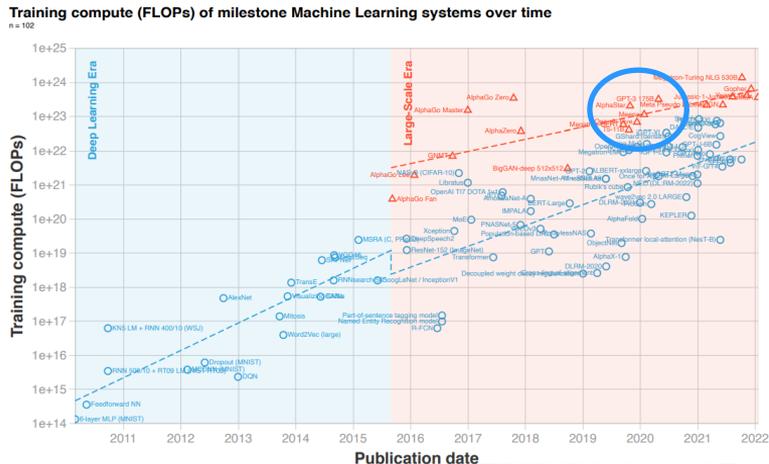
- 计算、数据和算法的进步是指导现代机器学习（ML）进步的三个基本因素。在2010年之前，训练计算的增长符合摩尔定律，大约每20个月翻一番。自2010年代早期深度学习（Deep Learning）问世以来，训练计算的规模已经加快，大约每6个月翻一番。2015年末，随着公司开发大规模ML模型，训练计算需求增加10至100倍，出现了一种新趋势——训练高级ML系统的需求快速增长。
- 2015-2016年左右，出现了大规模模型的新趋势。这一新趋势始于2015年末的AlphaGo，并持续至今（GPT-3于2020年出现）。

表3：ML相关计算规模呈现持续快速攀升趋势

文章	研究结论
Amodei & Hernandez (2018)	~3.4 month doubling time between 2012 and 2018
Sastry, G., et.al (2019, November).	~2 year doubling period between 1959 and 2018
Lyzhov, A. (2021, April).	>2 year doubling period between 2018 and 2020

数据来源：《compute trends across three eras of machine learning》, Jaime Sevilla et.al, 国泰君安证券研究

图40：2010年至2022年间102个里程碑ML系统的训练计算趋势如下所示



数据来源：《compute trends across three eras of machine learning》, Jaime Sevilla et.al

# 03

**行业进程：AIGC多模态交互功能持续演化，奠定多场景商用基础**

# 03 AIGC: 利用人工智能产生内容，提升生产力曲线

■ AIGC：Artificial Intelligence Generated Context，即可以利用人工智能技术自动产生内容，常见如代码生成，文本问答等

图41：AIGC应用功能简单如下图所示

```

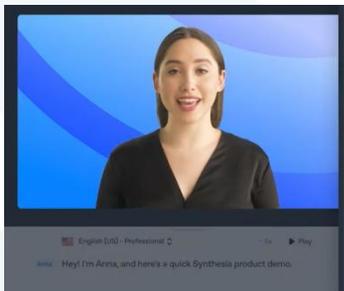
# //user/js/rev ts-node
import { fetch } from 'fetch-h2';

// Determine whether the content of text is positive
// Use a web service
async function isPositive(text: string): Promise<boolean> {
  const response = await fetch('http://text-processing.com/api/v1/text/?text=' +
    encodeURIComponent(text));
  const json = await response.json();
  return json.pos > 0;
}

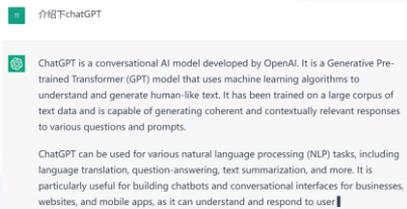
const json = await isPositive('');
const result = { pos: json.pos };

```

Copilot 可自动编写代码



Synthesia可将输入文本合成不同讲述风格的视频



ChatGPT 可与人类进行流畅的文本问答



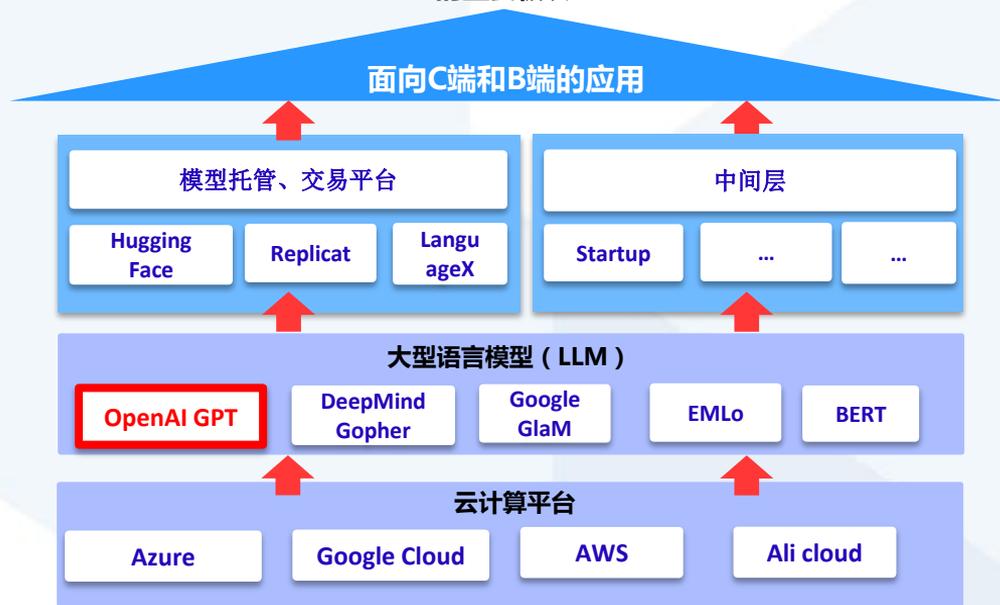
DALL-E2 由文字描述生成图像，如：An astronaut riding a horse as a pencil drawing

## 03 ChatGPT已成为AIGC功能矩阵中的重要板块

### ■ ChatGPT是AIGC“数字内容智能编辑”功能中的重要组成部分，ChatGPT模型的出现对于文字/语音模态的AIGC应用具有重要意义

- ✓ 随着深度学习技术的快速突破以及数字内容的海量增长，AIGC领域相关技术打破了预定义规则的局限性，使得快速便捷且智慧地输出多模态的数字内容成为可能。
- ✓ 在技术创新以及多模态模型的持续突破下，AIGC根据功能和对象的不同，按顺序可包括三种主要实用功能：数字内容孪生、数字内容的智能编辑、数字内容的智能创作。这三种功能相互嵌套与结合，可以让AIGC产品具备超越人类的创作潜力。**而ChatGPT正是AIGC的数字内容智能编辑这一大功能领域中的重要组成部分。**

图42：ChatGPT是AIGC的产品应用框架中大型语言模型的重要板块



资料来源：Sam Altman，机器翻译观察（公众号），国泰君安证券研究

## 03

## AIGC相关技术包含了三大前沿能力

### ■ 数字内容孪生能力构建现实世界-虚拟世界映射

✓ 孪生能力包括智能增强与转译技术，其中增强技术弥补内容数字化过程中的信息损失，转译技术在理解基础上对内容进行多种形式呈现

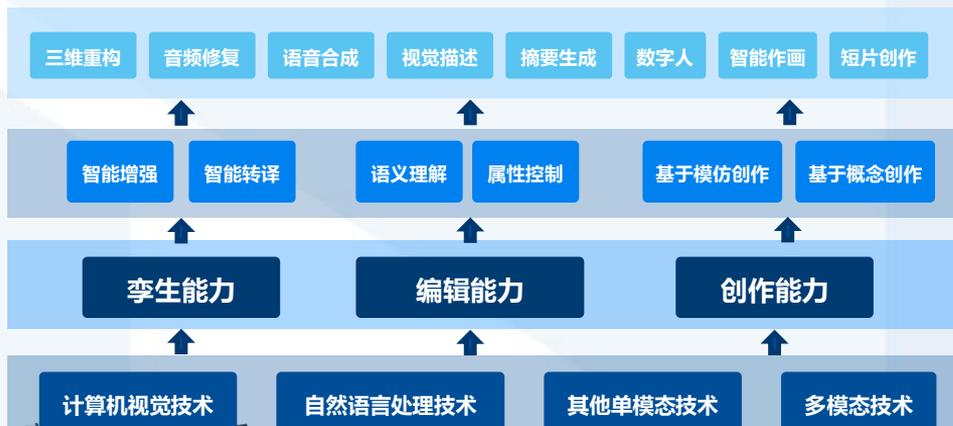
### ■ 数字编辑能力打通现实世界虚拟世界交互通道

✓ 编辑能力包括智能语义理解与属性控制，语义理解帮助实现数字内容各属性的分离解耦，属性控制则在理解基础上对属性进行精确修改、编辑与二次生成，最终反馈于现实世界，形成孪生-反馈闭环

### ■ 数字创作能力从数据理解走向数据创作

✓ 创作能力可分为基于模仿的创作与基于概念的创作，前者基于对某一类作品数据分布进行创作，而后者从海量数据中学习抽象概念，并基于概念创作出现实世界不存在的内容

图43：AIGC三大前沿技术能力架构如下图所示

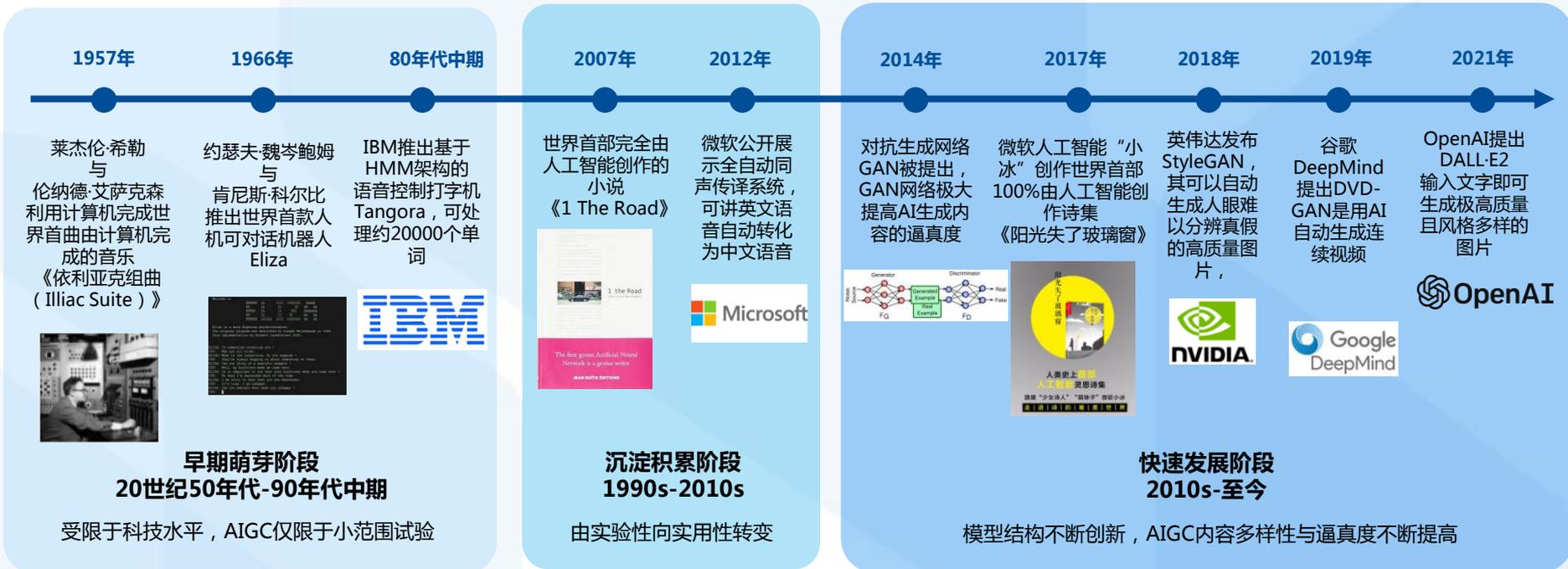


## 03

## AIGC行业发展经历了三个主要时期

## ■ AIGC发展经历了早期萌芽、沉淀积累和2014年之后的快速发展阶段

图44：AIGC经历了大致三个阶段的演化发展



资料来源：中国通信研究院，国泰君安证券研究

## 03

## 从分析式AI到生成式AI逐步演化，生成式AI赋予AIGC创新力

- **生成式AI起源于分析式AI，分析式AI发展过程中的技术积累为生成式AI的产生奠定基础**
- ✓ 分析式AI其学习的知识局限于数据本身；生成式AI在总结归纳数据知识的基础上可生成数据中不存在的样本
- ✓ 最新生成式AI技术如GAN，Diffusion等，催生多款AIGC产品如：OpenAI系列、DALL·E2(Diffusion)，Starry A.I.(基于GAN)等

图45：AIGC是在分析式AI的基础上，学习数据产生模式，实现新样本内容的创造

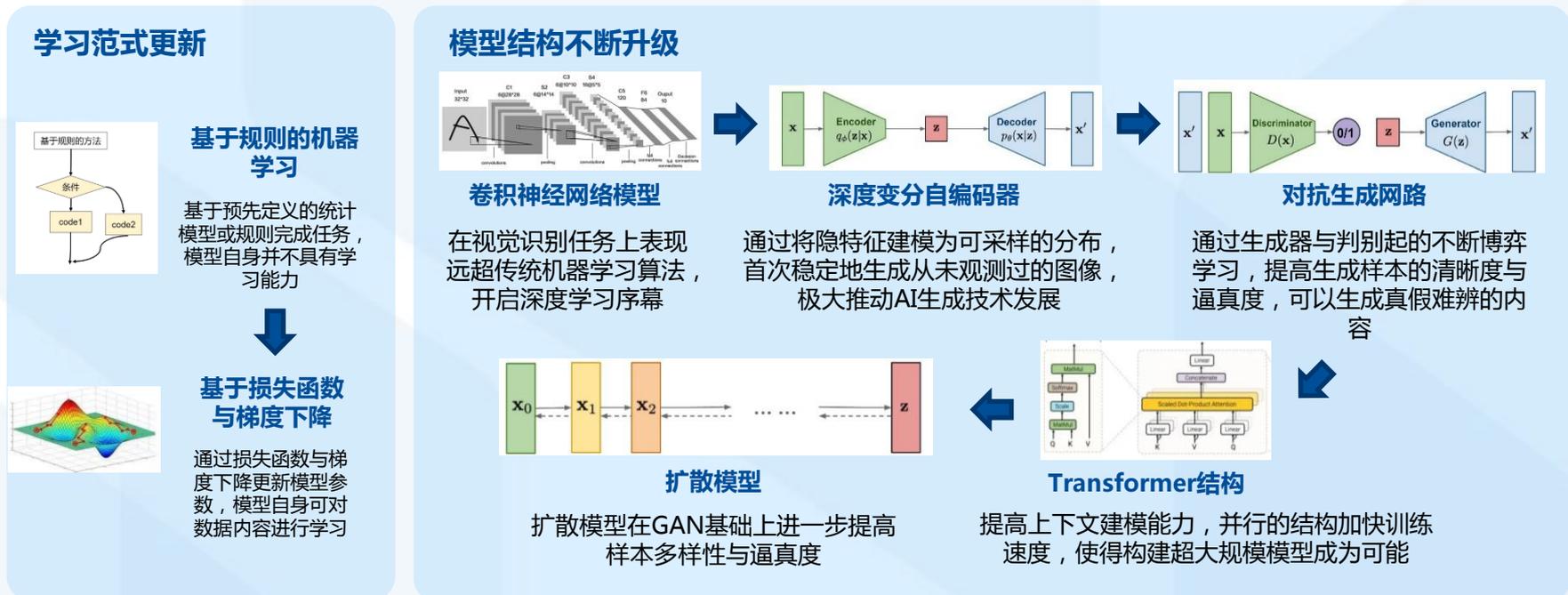


## 03

## AIGC：学习范式更新奠定基础，模型结构升级助力腾飞

- 人工智能技术推动AIGC行业不断发展，其中学习范式的更新赋予AI模型主动学习能力，模型结构升级提升AI模型学习、归纳与创新能力

图46：AI模型的升级迭代为AIGC性能跨越式发展奠定基础



## 03

## AIGC产业链涵盖了从硬件到多类终端应用的广泛领域

### ■ AIGC关联产业可分为应用层、模型层、云计算平台与计算硬件层

- ✓ 计算硬件层结合云计算平台为AIGC提供机器学习训练与推理算力，其中GPU与TPU为硬件核心，主要参与厂商包括英伟达（GPU）与谷歌（TPU）；云平台参与厂商则包含AWS，GCP，Azure以及Coreweave；**计算硬件层中云计算平台厂商分布稳定，竞争出现于模型层面与应用层面。**
- ✓ 模型层面，闭源基础模型提供商如OpenAI通过API向用户提供服务，而开源基础模型则通过在托管平台如Hugging Face、Replica公开模型权重。模型训练其高算力需求推动了模型层厂商与云计算厂商建立合作关系（如OpenAI+Azure，GCP+DeepMind）。**模型层面闭源模型较为普遍，各厂商依靠模型建立技术壁垒**
- ✓ 在应用层面，MidJourney，Runway等自主研发、持有模型；而Jasper，Github Copilot则通过调用闭源模型商提供的API或采用托管平台共享的模型。

图47：AIGC市场框架可由基础设施层、模型层、托管平台以及应用层来进行划分



## 03

## AIGC产业链上下游玩家百花齐放

- AIGC上游主要包括数据供给方、算法机构、创作者生态以及底层配合工具等，中游主要是文字、图像、音频和视频处理厂商，其中玩家众多；下游主要是各类内容创作及分发平台以及内容服务机构等。

图48：AIGC产业链上下游参与者分类如下图所示



## 03 AIGC厂商之间的竞争在于模型层面竞争

### ■ 追根溯源，AIGC依赖于底层机器学习模型产生内容，因此模型为AIGC行业厂商真正竞争力所在

- ✓ 文本生成产品多依赖GPT系列模型，自己训练的模型在图像/视频模态产品中较为普遍（图像/视频模态产品通常拥有自己训练的模型，而不是如文本模态调用OpenAI提供的模型服务）
- ✓ 比较而言，OpenAI依靠模型建立先发竞争优势，技术到产品转化相对亮眼

图49：AIGC模型产品之间存在激烈竞争

	Open AI	Stability AI	Midjourney	Meta
模型	<ul style="list-style-type: none"> <li>• GPT系列模型</li> <li>• ChatGPT,</li> <li>• CLIP</li> <li>• DALL·E 2</li> <li>• Codex</li> <li>• ...</li> </ul>	Stable Diffusion	Midjourney	Make-a-scene Make-a-video
产品	基于GPT-3: Snazzy AI, Copysmith, Copy.ai 基于Codex: Github Copilot 基于DALL·E 2: DALL·E 2	Dreamstudio 开源算法Stable Diffusion应用于DALL·E 2, Diffusion等其他公司 模型中	—	—

资料来源：各公司官网，腾讯研究院，量子位，国泰君安证券研究

## 03 | AIGC取长补短，有望成为主流内容生产模式

- AIGC所属内容生产生态的发展经历了专家生产内容（PGC）、用户生成内容（UGC）、AI辅助生产内容、AI生产内容（AIGC）四个阶段，目前处于一、二阶段为主，第三阶段为辅的境况
- AIGC克服PGC与UGC存在的 质量、产量 无法兼具的缺点，其有望成为未来主流的内容生产模式

图50：AIGC生态内容生产模式理论上会经历四个发展阶段



资料来源：foresightnews.pro，中国信通院，国泰君安证券研究

## 03 AIGC不同模态对应着各种生成技术及应用场景

### ■ AIGC不同模态对应的技术应用场景也有着各自的细分品类



#### 文本生成

分为非交互式文本和交互式文本



#### 音频生成

包括语音克隆、文本生成特定语音，音乐生成



#### 视频生成

视频属性编辑，视频自动剪辑，视频部分编辑



#### 跨模态生成

文字生成图像，文字生成视频，图像/视频到文本

图52：AIGC各技术应用场景对应的特征及细分品类如下图所示



#### 图像生成

分为图像编辑工具和图像自主生成工具



#### 策略生成

AI基于特定问题和场景，自主提出解决方案的过程。此类技术在游戏、自动驾驶、机器人控制等领域有极高应用价值。



#### GameAI

游戏本身为AI提供了最佳的研发场景，有大量数据和贴近实际场景



#### 虚拟人生成

指存在于非物理世界(如图片、视频、直播、一体机、VR)中，并具有多重人类特征的综合产物。  
请参阅附注免责声明



## 03 AIGC文本生成技术场景可分为交互式和非交互式

- AIGC非交互式文本生成技术中，结构化写作其形式相对固定，生成难度较小，商业化应用较为广泛；而创作型写作开放性较大，在长文本生成中难度较大，仍需技术进一步发展
- 随着通信互联网技术发展，线上社交需求快速增长，如闲聊机器人等交互式文本产品将迎来快速发展

图53：文本内容生产领域相关细分特征如下架构图所述



资料来源：量子位智库，国泰君安证券研究

## 03 AIGC文本生成技术商业化落地有望优势先发

- 文本领域预训练大模型技术成熟，文本领域细分垂类较多，产品数量居首位，模型数量发展超过其他模态技术
- 数字内容中，文字模态数据远大于图片/视频/音频等，发展前景相对较大
- 基于GPT-3的文字生成功能已嵌入如Writesonic、Conversion.ai、Copysmith等软件中，商业化前景相对清晰

图54：文本生成技术商业化落地有比较优势

### ✓ 文本易于获得

相较于其他模态数据，文本数据易于获得且数量庞大，满足大规模预训练模型对数据量的需求

### ✓ 文本表达信息更为高效

相比图片，语音，视频等模态，文本信息在存储、处理、传输时占用资源少，而信息量不逊色于图片等。在这种优势文本将成为人机交互等领域的首选

### ✓ 大模型趋势下文本预训练实现难度小

文本数据其离散的性质使得相同模型架构下大模型训练消耗资源低与图片/视频



图55：AIGC文本模态技术（包括文本与代码）商业化领跑视频/图像模态技术

	2020之前	2020	2022	2023?	2025?	2030?
文本	垃圾邮件检测 翻译 基本问答	文本初稿写作	较长文本二稿写作	垂直领域文本写作 微调能力提升 (科技文章等)	终稿写作能力超过人类平均水平	终稿写作能力超过专业作家水平
代码	单行代码自动补全	多行代码生成	更长形式的代码生成 生成准确度提升	支持更多语言种类与垂直领域	根据文本生成初版产品	根据文本生成最终产品，表现超过全栈开发者
图像			艺术图像 图标 摄影作品	样例 (产品设计图, 建筑图等)	终稿 (产品设计图, 建筑图等)	终稿水平超过专业艺术家, 设计师, 摄影师
视频、3D、游戏			3D/视频模型初次尝试	视频/3D建模初稿	二稿	AI版Roblox 视频游戏、电影定制化

初次尝试 接近成熟 全盛发展

资料来源：红杉，国泰君安证券研究

资料来源：量子位智库，国泰君安证券研究

## 03

## AIGC图像生成技术随着模型结构的优化而明显提升

### ■ 模型结构不断进化提高了AIGC生产图像的多样性，但要求较高的功能实现还有待于技术的进一步提升

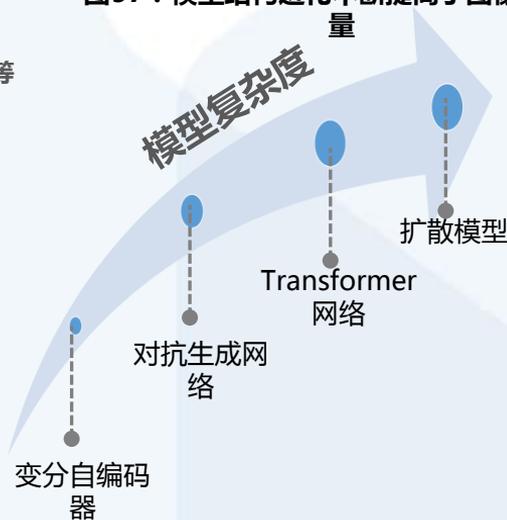
- ✓ “图像编辑”难度低于“图像生成”与“2D-3D”转换，目前已存在多款产品支持“图像编辑”，而对于“图像生成”任务，由于图片相较包含更多元素，其生成效果仍存在不稳定性，对于要求较高的功能类图像生成仍需要技术上的提升。

图56：图像生成技术具体实现包括图像编辑、2D-3D转换以及自主生成



资料来源：量子位智库，国泰君安证券研究

图57：模型结构进化不断提高了图像生成质量



资料来源：量子位智库，国泰君安证券研究

请参阅附注免责声明

## 03 AIGC音频生成技术正朝更富情感等人类特征演化

■ 文本到语音任务已比较成熟，语音质量已达到自然的标准，未来将朝更富情感、富韵律的语音合成以及小样本语音学习方向发展

✓ 音乐生成任务中仍需解决音乐数据难以标注的问题，数据标注其颗粒度大小影响音乐生成任务的可控性。若可控性得以解决，则可指定风格、情绪等因素的音乐生成任务有希望在影视、游戏等场景下的到大量应用。

图58：AIGC音频细分生成技术特点如下图所示



## 03 视频生成成为AIGC应用生态中的高潜力场景

### ■ 视频生成本质上与图片生成类似，通过对视频进行帧数级别的切割，实现对每一帧的处理

- ✓ 视频生成过程包括三个阶段：**数据的提取、训练和转换**，当前技术正在着重**提升视频修改精准度和实时性**两个维度。鉴于视频本身的文本、图像和音频的综合属性，视频生成也是跨模态生成领域的重要应用场景。

图59：视频生成技术具体包括视频属性编辑、视频自动剪辑、视频内容动态编辑



## 03

## 跨模态生成技术是真正实现认知和决策智能的转折点

- 现实世界的信息是文本、音频、视觉、传感器以及人类各种触觉的综合体系，要更为精准地模拟现实世界，就需要将各种模态能力之间打通，例如文字-图像、文字-视频等跨模态生成能力
- ✓ 大型预训练模型的发展使得跨模态逐步成熟，“文本-图像”生成正在快速落地，“文字-视频”的实验效果也已较为理想（视频时长、清晰程度、逻辑等还有较大提升空间）。

图60：跨模态生成领域当前的主要功能类型如下所示



资料来源：量子位智库，Nova AI官网，国泰君安证券研究

## 03

## AIGC改变数字内容生产模式

- AIGC作为新的内容生产模式，其具有内容多样，可控性强与生产效率高的优点，符合传媒，电商，影视，娱乐等行业对内容数字化程度高、内容多样以及内容更新快的要求，AIGC在以上行业逐渐替代传统内容生产模式的趋势十分显著

图61：AIGC对各大行业的影响维度如下图所示



## 03

## AIGC渗透传媒领域各个环节

- AIGC技术逐渐渗透传媒领域包括采集、编辑、传播等环节，有助于加快内容生产效率，提高内容质量，扩大内容影响力

图62：AIGC对传媒领域的赋能如下图所示



资料来源：中国信通院，国泰君安证券研究

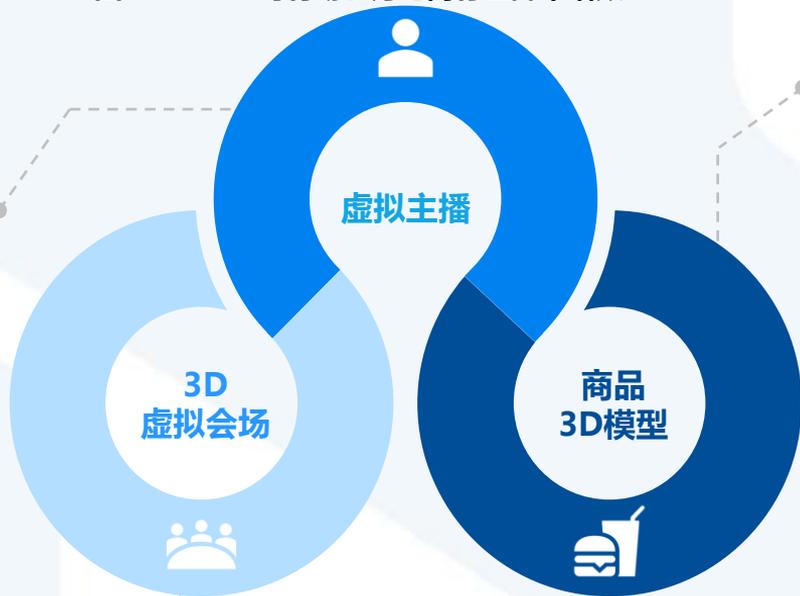
## 03

## AIGC化2D为3D，拓展电商展示维度

- AIGC 2D图像生成3D模型技术为传统电商提供多维度的展示空间
- 虚拟合成主播为客户提供更及时、可靠、亲和的服务体验

图63：AIGC对有助于为电商行业降本增效

- 通过AI技术打造虚拟主播介绍商品信息并提供客服服务，可实现直播24小时不间断，为客户提供更灵活的观看时间与更方便的购物服务
- 虚拟主播人设可控性强，因主播自身原因导致品牌受损的几率小
- 虚拟主播易于拉近品牌与年轻消费者的距离，同时构建元宇宙潮流下的品牌形象，为未来多元化的虚拟世界打下基础



- 基于不同角度的商品图像，借助视觉生成算法生成商品3D模型，720°全方位展示商品，减少消费者选品与沟通时间
- 3D商品模型提供线上虚拟“试、穿、看、戴”，高度还原商品真实效果，减小订单退换货率

- 通过从2D图像构建3D场景的AI技术，降低商家搭建3D线上购物空间的门槛与成本。商家可向消费者提供线上沉浸式购物体验，可降低门店扩张成本，扩大品牌知名度

资料来源：中国信通院，国泰君安证券研究

## 03

## AIGC打破传统娱乐体验边界

- AIGC技术打破物理边界，使粉丝可与偶像亲密互动，并形成新的发展点
- AIGC为用户打开虚拟世界入口，通过趣味方式体验虚拟世界

图64：AIGC在娱乐领域也有诸多赋能点，有助于进一步提升产业空间



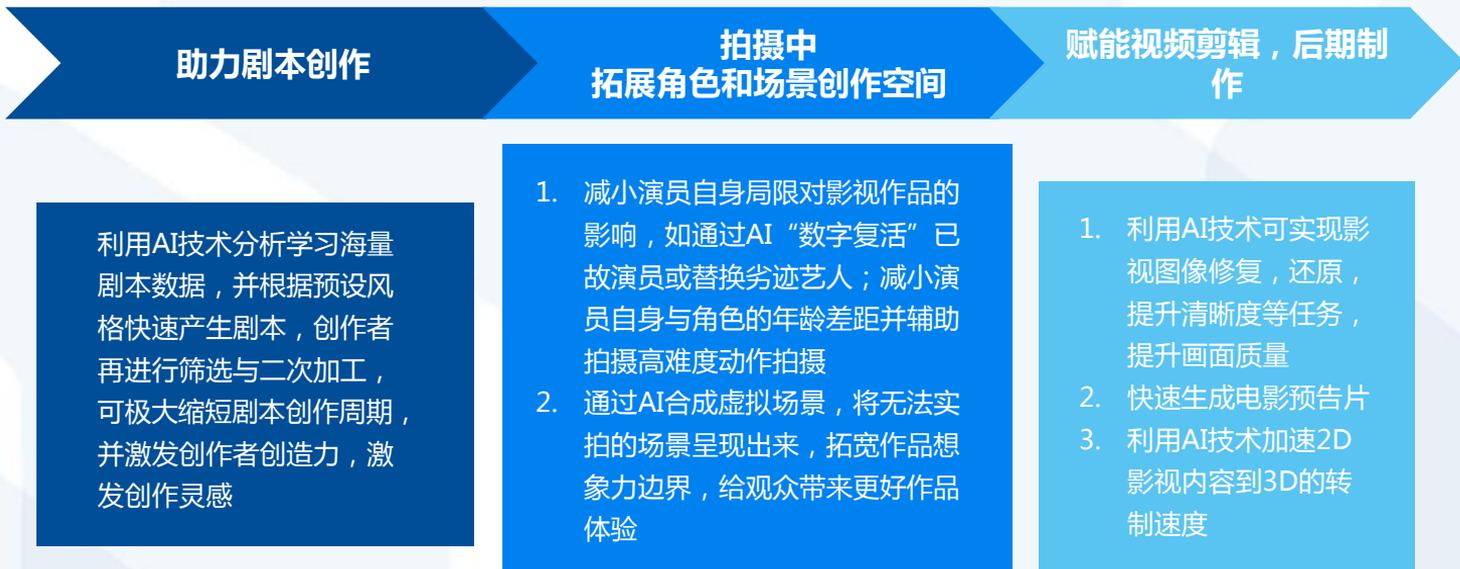
资料来源：中国信通院，国泰君安证券研究

## 03

## AIGC拓宽影视行业创意边际

- AIGC技术以其内容多样性为作品内容带来更多灵感
- AIGC技术帮助拍摄突破物理限制，还原剧本效果，提高作品质量

图65：AIGC在剧本创作、拍摄过程以及后期制作方面均有助益



## 03

## AIGC促进各行业转型升级

## ■ AIGC技术在各行业数字内容相关领域均有发挥空间

图66：AIGC在教育、金融、工业、医疗领域的应用优势如下图所示



# 04

## 商业化：多领域多功能应用密集落地，ChatGPT商用前景可期

## 04

## ChatGPT试点订阅计划——ChatGPT Plus发布，商业化序幕已经拉开

■ 2023年2月2日，OpenAI发布ChatGPT试点订阅计划——ChatGPT Plus，每月20美元

✓ ChatGPT Plus订阅者可获得比免费版本更稳定、更快的服务，及尝试新功能和优化的优先权

图67：ChatGPT Plus处于OpenAI官网的醒目位置，凸显商业化倾向



资料来源：OpenAI官网，国泰君安证券研究

图68：ChatGPT Plus订阅者可获得比免费版本更优质的服务

The new subscription plan, ChatGPT Plus, will be available for \$20/month, and subscribers will receive a number of benefits:

- General access to ChatGPT, even during peak times
- Faster response times
- Priority access to new features and improvements

ChatGPT Plus is available to customers in the United States, and we will begin the process of inviting people from our [waitlist](#) over the coming weeks. We plan to expand access and support to additional countries and regions soon.

We love our free users and will continue to offer free access to ChatGPT. By offering this subscription pricing, we will be able to help support free access availability to as many people as possible.

资料来源：OpenAI官网，国泰君安证券研究

# 04

## ChatGPT+传媒：实现智能新闻写作，提升新闻的时效性

- ChatGPT可以帮助新闻媒体工作者智能生成报道，将部分劳动性的采编工作自动化，更快、更准、更智能地生成内容

图69：过往成功案例众多，ChatGPT在传媒领域的商业化大有可为



2014年3月，美国洛杉矶时报网站的机器人记者Quakebot，在洛杉矶地震后仅3分钟，就写出相关信息并进行发布



中国地震网的写稿机器人在九寨沟地震发生后7秒内就完成了相关信息的编发



美联社使用的智能写稿平台Wordsmith可以每秒写出2000篇报道



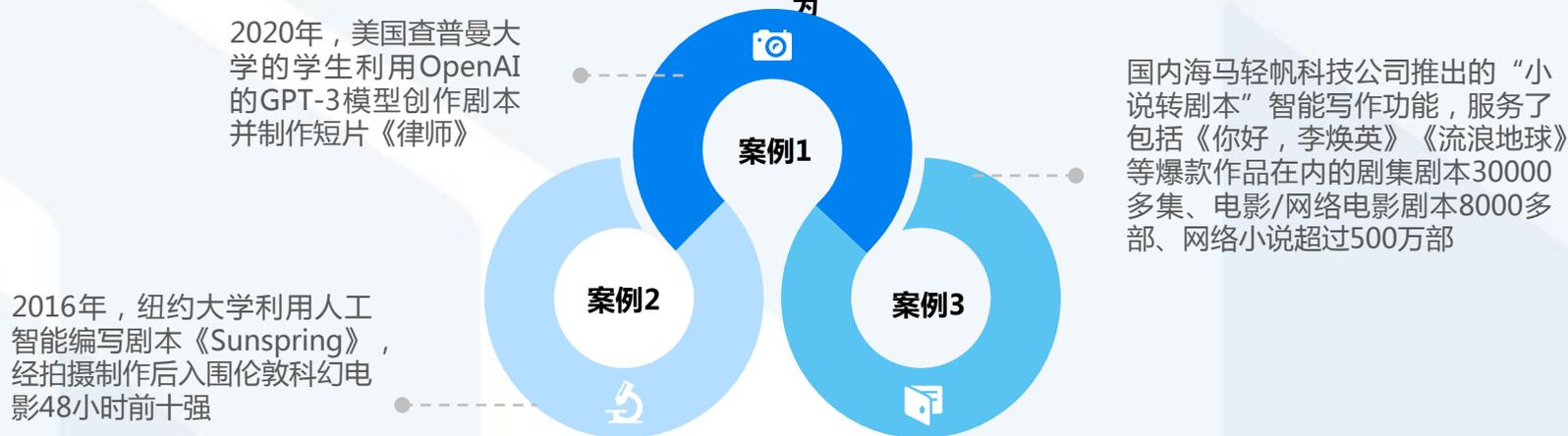
第一财经“DT稿王”一分钟可写出1680字

资料来源：中国信息通信研究院，国泰君安证券研究

## 04 ChatGPT+影视：拓宽创作素材，提升作品质量

- ChatGPT可以根据大众的兴趣定制影视内容，从而更有可能吸引大众的注意力，获得更好的收视率、票房和口碑
- ✓ ChatGPT可以为剧本创作提供新思路，创作者可根据ChatGPT的生成内容再进行筛选和二次加工，从而激发创作者的灵感，开拓创作思路，缩短创作周期。
- ChatGPT有着降本增效的优势，可以有效帮助影视制作团队降低在内容创作上的成本，提高内容创作的效率，在更短的时间内制作出更高质量的影视内容

图70：内容生成在影视领域运用广泛，ChatGPT在影视领域的商业化大有可为



## 04 | ChatGPT+营销：打造虚拟客服，赋能产品销售

### ■ ChatGPT可以打造虚拟客服，助力产品营销

图71：利用ChatGPT打造的虚拟客服，具备无可比拟的营销优势

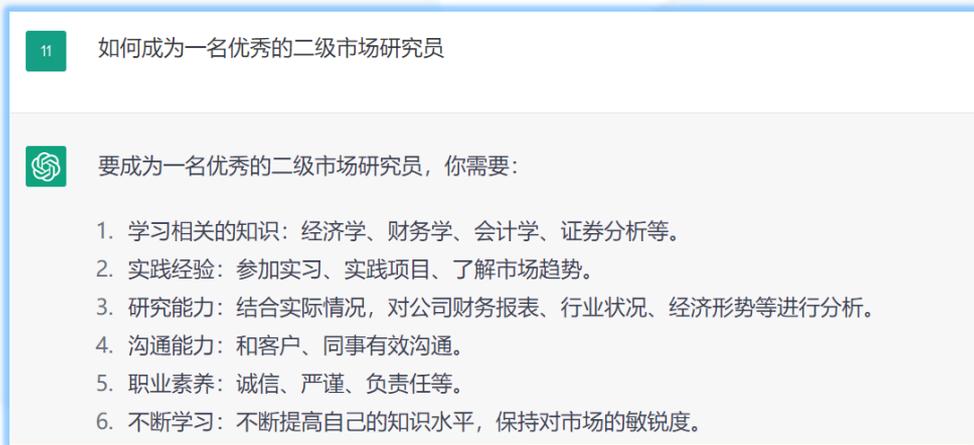
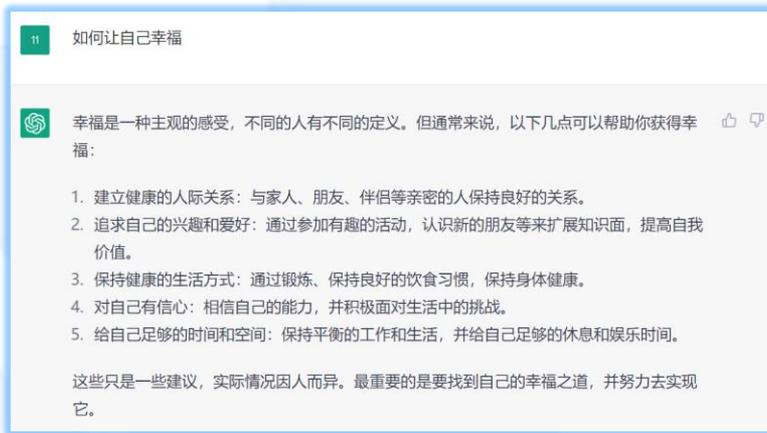


## 04

## ChatGPT+娱乐：人机互动加强，激发用户参与热情

## ■ ChatGPT可以成为线上的实时聊天对象，增加互动的趣味性和娱乐性

图72：ChatGPT的对话具备趣味性和娱乐性



资料来源：ChatGPT，国泰君安证券研究

## 04 | ChatGPT+其他：促进数实共生，助力产业升级

- ChatGPT+教育：赋予教育教材新活力，让教育方式更个性化、更智能
- ChatGPT+金融：帮助金融机构降本增效，让金融服务更有温度
- ChatGPT+医疗：赋能医疗机构诊疗全过程

图73：ChatGPT在教育、金融和医疗产业中也大有用武之地

### ChatGPT+教育



ChatGPT给教育工作者提供全新的教学工具，学生可以通过自主提问的方式快速查缺补漏，让每个学生得到更个性化、更智能的教育

### ChatGPT+金融



一方面，金融机构可以通过ChatGPT实现金融资讯、金融产品介绍内容的自动化生产，提升金融机构内容生产的效率；另一方面，可以通过ChatGPT塑造虚拟理财顾问，让金融服务更有温度

### ChatGPT+医疗



ChatGPT可以快速了解患者的病情并给出较合理的及时反馈，通过人性化的方式第一时间抚慰患者，从而舒缓患者的情绪，加速其康复；同时，让医者有更多的时间和精力集中在关键患者的关键治疗环节

## 04 短期来看，ChatGPT的快速推广阻力仍存

- 短期来看，ChatGPT仍存在诸多不足，亟待解决的问题主要有两方面：合规性问题和 technical 问题

图74：合规性问题和 technical 问题是当前ChatGPT主要面临的挑战

### 合规性问题



1. ChatGPT产生的答复是否产生相应的知识产权？
2. ChatGPT进行数据挖掘和训练的过程是否需要获得相应的知识产权授权？
3. ChatGPT是基于统计的语言模型，这一机制导致回答偏差会进而导致虚假信息传播的法律风险，如何降低其虚假信息传播风险？

### technical 问题



1. ChatGPT的回答可能过时，因为其数据库内容只到2021年，对于涉及2022年之后，或者在2022年有变动的问题无能为力
2. ChatGPT在专业较强的领域无法保证正确率，即使在鸡兔同笼此类初级问题中仍然存在错误，并且英文回答和中文回答存在明显差异化
3. ChatGPT对于不熟悉的问题会强行给出一定的答案，即使答案明显错误，依然会坚持下去，直到明确戳破其掩饰的内容，会立马道歉，但本质上会在不熟悉的领域造成误导



资料来源：零壹财经，第一财经，国泰君安证券研究



国泰君安证券  
GUOTAI JUNAN SECURITIES

诚信·责任·亲和·专业·创新

请参阅附注免责声明 70

# 免责声明

本公司具有中国证监会核准的证券投资咨询业务资格

## 分析师声明

作者具有中国证券业协会授予的证券投资咨询执业资格或相当的专业胜任能力，保证报告所采用的数据均来自合规渠道，分析逻辑基于作者的职业理解，本报告清晰准确地反映了作者的研究观点，力求独立、客观和公正，结论不受任何第三方的授意或影响，特此声明。

## 免责声明

本报告仅供国泰君安证券股份有限公司（以下简称“本公司”）的客户使用。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。本报告仅在相关法律许可的情况下发放，并仅为提供信息而发放，概不构成任何广告。

本报告的信息来源于已公开的资料，本公司对该等信息的准确性、完整性或可靠性不作任何保证。本报告所载的资料、意见及推测仅反映本公司于发布本报告当日的判断，本报告所指的证券或投资标的的价格、价值及投资收入可升可跌。过往表现不应作为日后的表现依据。在不同时期，本公司可发出与本报告所载资料、意见及推测不一致的报告。本公司不保证本报告所含信息保持在最新状态。同时，本公司对本报告所含信息可在不发出通知的情形下做出修改，投资者应当自行关注相应的更新或修改。

本报告中所指的投资及服务可能不适合个别客户，不构成客户私人咨询建议。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。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。

本公司利用信息隔离墙控制内部一个或多个领域、部门或关联机构之间的信息流动。因此，投资者应注意，在法律许可的情况下，本公司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或期权并进行证券或期权交易，也可能为这些公司提供或者争取提供投资银行、财务顾问或者金融产品等相关服务。在法律许可的情况下，本公司的员工可能担任本报告所提到的公司的董事。

市场有风险，投资需谨慎。投资者不应将本报告作为作出投资决策的唯一参考因素，亦不应认为本报告可以取代自己的判断。在决定投资前，如有需要，投资者务必向专业人士咨询并谨慎决策。

本报告版权仅为本公司所有，未经书面许可，任何机构和个人不得以任何形式翻版、复制、发表或引用。如征得本公司同意进行引用、刊发的，需在允许的范围内使用，并注明出处为“国泰君安证券研究”，且不得对本报告进行任何有悖原意的引用、删节和修改。

若本公司以外的其他机构（以下简称“该机构”）发送本报告，则由该机构独自为此发送行为负责。通过此途径获得本报告的投资者应自行联系该机构以要求获悉更详细信息或进而交易本报告中提及的证券。本报告不构成本公司向该机构之客户提供的投资建议，本公司、本公司员工或者关联机构亦不为该机构之客户因使用本报告或报告所载内容引起的任何损失承担任何责任。

国泰君安证券研究所 E-mail:gtjaresearch@gtjas.com	
上海	地址：上海市静安区新闻路669号
	邮编：200041
	电话：( 021 ) 38676666
深圳	地址：深圳市福田区益田路6009号
	邮编：518026
	电话：( 0755 ) 23976888
北京	地址：北京市西城区金融大街甲9
	邮编：200032
	电话：( 010 ) 83939888



国泰君安证券  
GUOTAI JUNAN SECURITIES

# 国泰君安证券研究所

Thank you for listening