

# AIGC深度报告：新一轮内容生产力革命的起点

评级：推荐（维持）

姚蕾(证券分析师)

S0350521080006

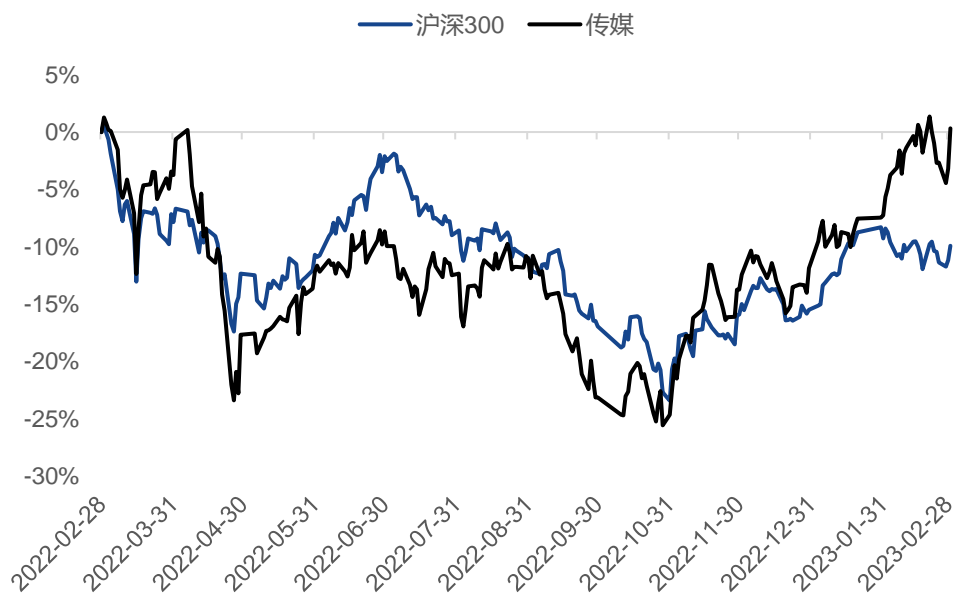
yaol02@ghzq.com.cn

方博云(证券分析师)

S0350521120002

fangby@ghzq.com.cn

## 最近一年走势



### 相对沪深300表现

表现	1M	3M	12M
传媒	8.51%	19.59%	0.33%
沪深300	-1.31%	10.55%	-9.92%

## 相关报告

《国海证券\_行业研究：元宇宙系列深度报告：下一代互联网前瞻\*传媒\*姚蕾》——2022/01/05

《国海证券\_行业研究：元宇宙系列深度报告之二：数字虚拟人——科技人文的交点，赋能产业的起点\*传媒\*姚蕾》——2022/03/10

《国海传媒\_行业研究：元宇宙系列深度报告之三：NFT的本质思考及破圈之路\*传媒\*姚蕾》——2022/03/11

股票代码	股票名称	2023/3/1	EPS			PE			投资评级
		股价	2021A	2022E	2023E	2021A	2022E	2023E	
300002.SZ	神州泰岳	6.7	0.2	0.3	0.3	33.6	23.7	19.2	未评级
300418.SZ	昆仑万维	25.2	1.3	1.0	1.1	19.2	25.7	23.1	未评级
300364.SZ	中文在线	12.3	0.1	-0.1	0.2	90.3	-	52.3	未评级
000681.SZ	视觉中国	15.8	0.2	0.2	0.3	72.5	87.7	57.2	未评级
300058.SZ	蓝色光标	6.2	0.2	-0.1	0.3	29.5	-	23.5	未评级
002555.SZ	三七互娱	23.0	1.3	1.4	1.6	17.6	16.8	14.4	买入
603444.SH	吉比特	366.6	20.4	18.4	24.1	17.9	19.9	15.2	买入
002624.SZ	完美世界	14.5	0.2	0.7	0.9	76.5	20.3	16.7	买入
300413.SZ	芒果超媒	34.8	1.2	1.0	1.3	29.7	35.4	26.8	买入
003032.SZ	传智教育	18.6	0.2	0.5	0.6	97.7	38.3	32.5	买入

- ◆ **数据、算法、算力共振推动AIGC发展，模型开源及商业化带来的产品化浪潮及通用人工智能领域的初探推动AIGC破圈。AIGC传媒相关应用有望超千亿。**
  - ✓ **复盘AIGC算法迭代：竞争中发展，模型开源及商业化推动应用破圈。**2017年推出的Transformer架构的并行训练优势奠定了大模型训练的基础，以GPT为代表的预训练模型，通过使用无标注数据预训练及微调，缓解了标注数据不足的问题，并不断提升参数量级及模型通用性，ChatGPT在此基础上加入了利用人类反馈强化学习的训练方法。扩散模型取代GAN成为图像生成领域的主流模型，CLIP模型推动跨模态生成技术的发展。GPT3的商业化及CLIP及Stable Diffusion模型的开源推动文本生成、文生图产品化的浪潮。谷歌、Meta持续探索文字生成视频领域模型。
  - ✓ **国内传媒领域应用有望超千亿。**Gartner预测至2023年将有20%的内容被生成式AI所创建；至2025年生成式AI产生的数据将占有所有数据的10%（目前不到1%）。红杉预测生成式ai将产生数万亿美元经济价值。2025年，国内生成式ai应用规模有望突破2000亿，我们预测国内传媒领域应用空间超1000亿。
- ◆ **AIGC应用于文本、音频、跨模态、策略生成，在设计、内容创作、广告营销、游戏、企业服务等领域开启商业化，有望开启新一轮内容生产力革命。**
  - ✓ **文本生成：**应用于辅助写作、营销、社交、浏览器、企业级服务、心理咨询等领域。代表公司Jasper.ai，通过SaaS订阅收费模式，获得B端客户认可，率先实现规模化收入；OpenAI旗下ChatGPT由于其通用性被集成至浏览器、办公自动化软件、企业级服务产品中，作为增值服务项目。
  - ✓ **音频生成：**应用于智能客服、有声读物制作、配音、导航、虚拟歌手、作曲等领域。代表公司喜马拉雅、倒映有声、标贝科技、StarX MusicX Lab等。
  - ✓ **跨模态生成：**包括文生图、文生视频，图片视频生成文字等应用。AI绘画代表产品Midjourney、DALL-E2、Dream studio、文心一格，主要按生成次数收费。
  - ✓ **策略生成：**应用于游戏、自动驾驶、机器人控制、智能交互数字人等领域。游戏领域代表性公司腾讯AI Lab、网易伏羲、启元世界、rct.ai、超参数等。
- ◆ **投资建议：**AIGC的快速发展源于数据、算法、算力的共振。在此基础上，AIGC的出圈源于模型商业化及开源带来的产品化浪潮，及ChatGPT在通用人工智能领域投射的曙光带来的震撼。AIGC目前在营销、社交、内容创作、游戏等领域均有应用，并开启商业化变现。随着算法迭代、算力提升，AIGC将开启新的内容生产力革命，为传媒行业发展提供新动力。基于此，我们维持行业“推荐”评级。建议重点关注三类公司：
  - 一、拥有自有算法及模型的公司，建议关注昆仑万维、神州泰岳；
  - 二、拥有海量内容及版权储备的公司，建议关注视觉中国、中文在线；
  - 三、相关应用领域龙头公司，重点推荐游戏、影视、营销板块。相关标的三七互娱、吉比特、完美世界、芒果超媒、蓝色光标。
- ◆ **风险提示：**技术发展演进不及预期、商业化进程不及预期、企业技术管理能力建设不足风险、企业内容审核能力不足风险、版权保护风险、新技术增加监管难度风险、技术滥用风险、核心人才流失风险、创作伦理风险、法律政策监管风险、估值中枢下移风险等。



## 1

### AIGC破圈元年

1.1 AIGC破圈

1.2 AIGC定义

1.3 发展历程

1.4 市场空间

1.5 商业模式

## 2

### AIGC应用场景

## 3

### 产业链及相关公司

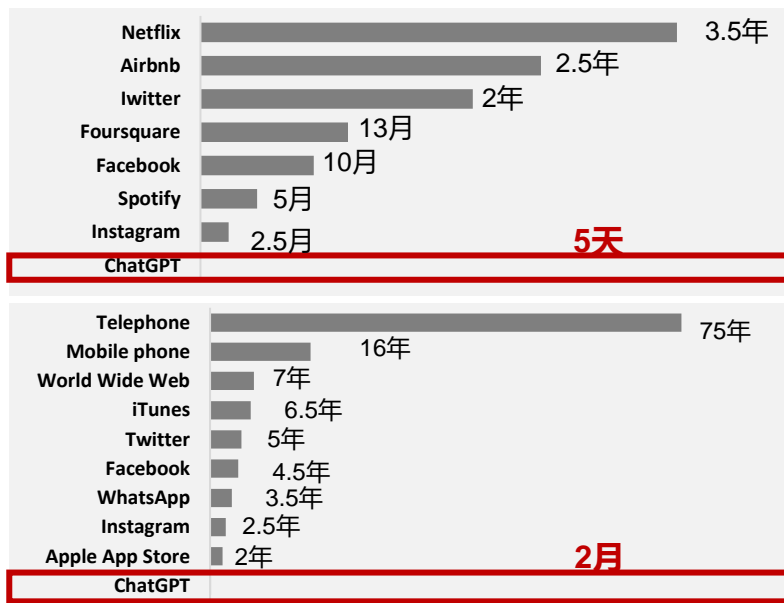
# 1.1 AIGC连续破圈：AI绘画/ChatGPT相继破圈

- AI绘画作品获奖。2022年8月，在美国科罗拉多州举办的新兴数字艺术家竞赛中，《太空歌剧院》获得“数字艺术/数字修饰照片”类别一等奖。参赛者没有绘画基础，利用AI工具Midjourney创作。
- ChatGPT于2022年11月30日推出，5天后用户破百万，两个月后月活用户突破1亿，成为史上用户增长速度最快的消费级应用程序。
- 风投及产业资本涌入AIGC。
  - ✓ 2021年开始，风投对AIGC的投资金额出现爆发式增长，2022年超20亿美元。主打AI生成文字的Jasper.ai于2022年10月完成1.25亿美元A轮融资，估值达15亿美元。AI视觉艺术创业公司Stability.ai获1.01亿美元种子轮融资，投后估值达10亿美元。
  - ✓ 据美国财经媒体Semafor报道，微软预计向ChatGPT的开发者OpenAI投资100亿美元，OpenAI投后估值将高达290亿美元。

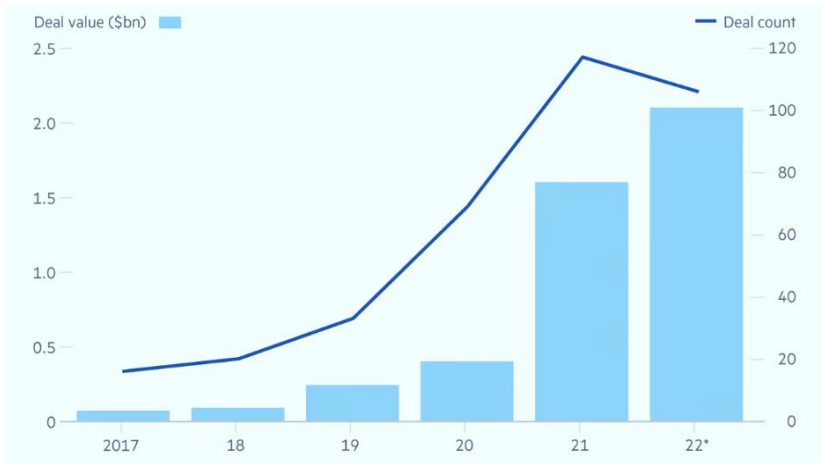
图表：AIGC绘画作品《太空歌剧院》



图表：各应用程序达到全球100万/1亿用户所用时间



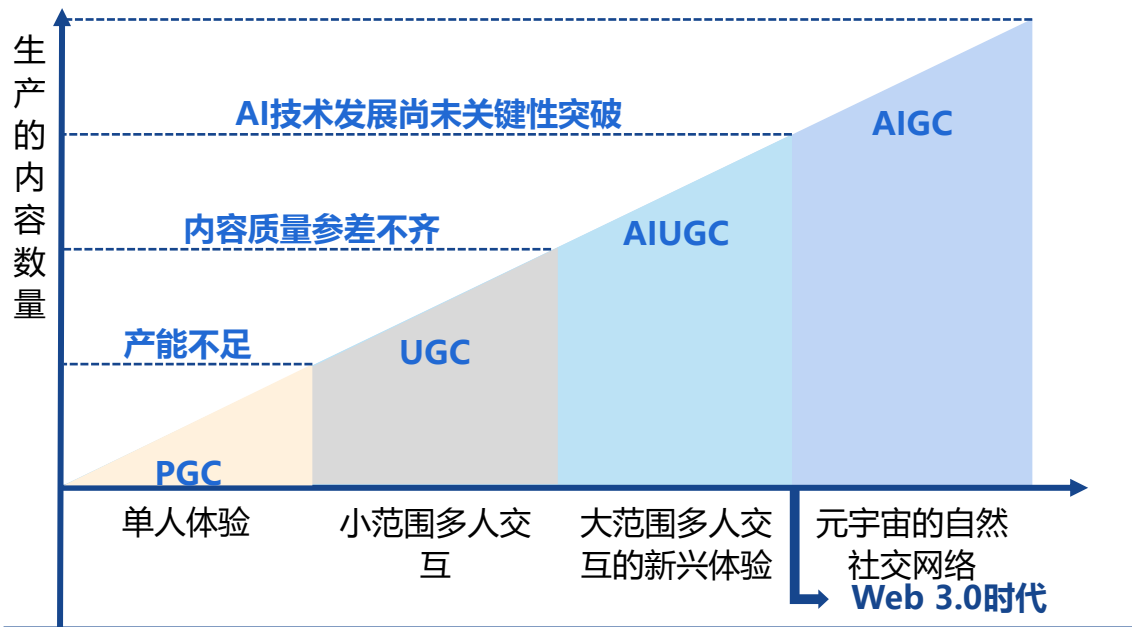
图表：风投对生成式AI投入金额爆发式增长



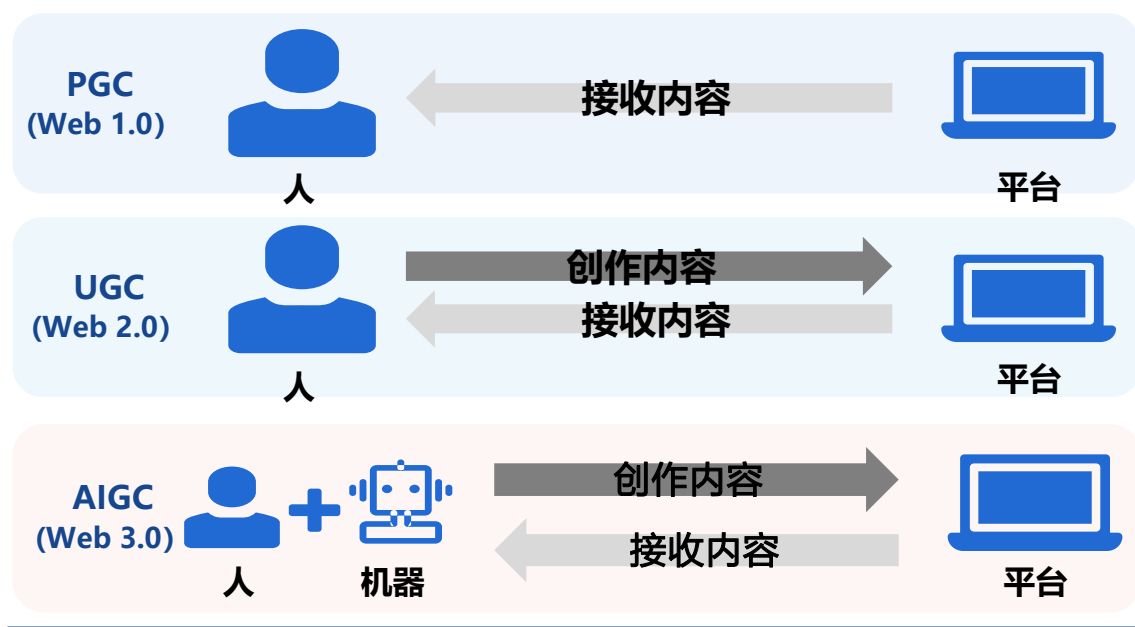
# 1.2 AIGC定义：新的内容生成方式、基于人工智能的技术集合

- **AIGC (AI Generated Content)** 即利用人工智能技术自动生成内容，受制于AI技术成熟度，目前AI仍为内容制作的辅助型角色 (AIUGC)，待技术突破，AI可真正作为内容创作者 (AIGC)。**AIGC是技术集合**，基于生成对抗网络GAN、大型预训练模型等人工智能技术，通过已有数据寻找规律，并通过适当的泛化能力生成相关内容的技术集合。
- ✓ **相关叫法**：合成式媒体 (Synthetic Media)，即基于AI生成的文字、图像、音频等；Gartner提出的生成式AI (Generative AI)，即由人工智能自动生成新的编程、内容或商业活动，让文字、音乐、图像、视频、场景等都可由AI算法自动生成。
- **从PGC到UGC，从UGC到AIGC**。Web1.0时代“只读”模式催生出“PGC”；Web2.0时代，社交媒体兴起，人与人通过网络交互，催生出以用户生产和分享内容“UGC”模式；Web3.0时代，内容消费需求进一步增长，个性化需求凸显，“AIGC”将成为重要的新内容生成方式。

图表：内容生产方式的变更

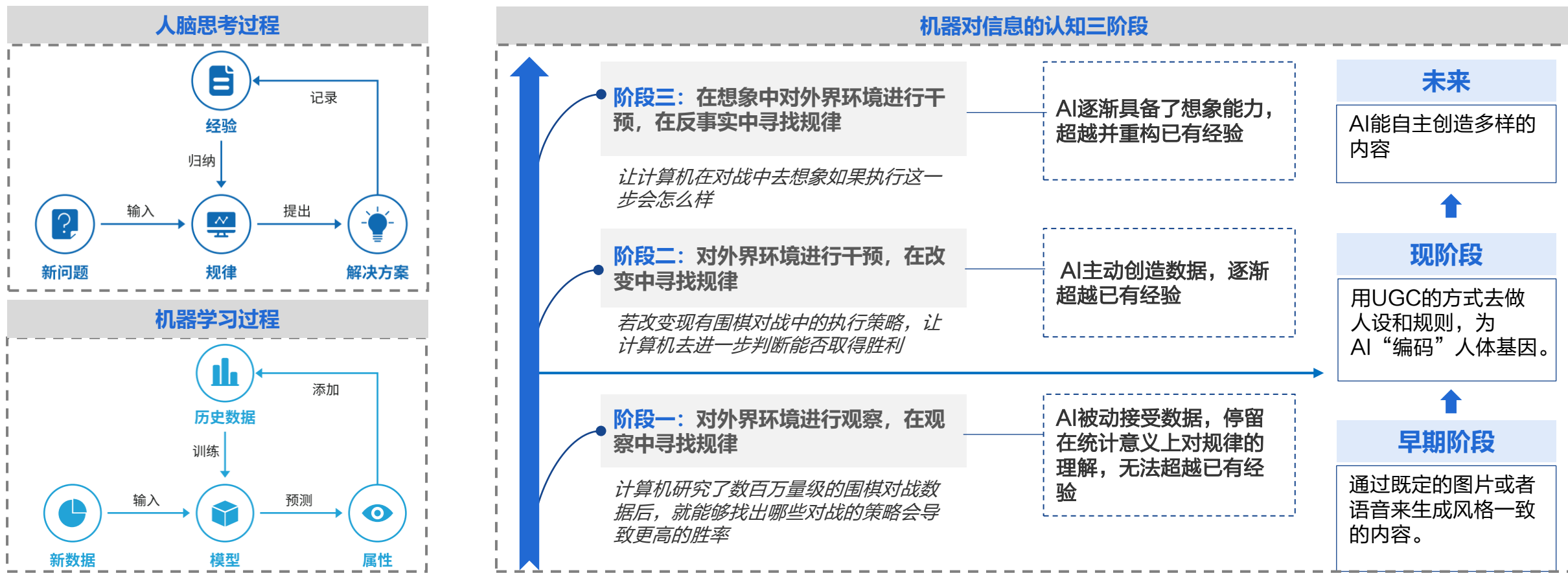


图表：从Web1.0到Web3.0的内容生成方式



# 1.2 AIGC定义：机器对信息认知分三阶段，从学习到超越经验

- 机器对信息的认知处于第一或第二阶段。人在遇到新问题时，会通过以往类似经历总结规律，并将新的问题套用到规律中，以推测可能的结果。相应地，机器学习基于对历史数据的归纳和学习，构建出事件模型，并将合适的新数据输入到相应的模型来预测未来。人类能够超越观察达到干预及想象阶段，而对于AI来说，目前还处于第一或第二阶段，一些复杂的信息还没办法处理，人类需要将其简化后再投喂给机器处理。
- UGC为AIGC提供了发展的数据基础，AIGC满足更个性化的内容消费需求。
  - ✓ 用户不再满足于专业团队和用户创造，对内容质量要求更高，AI在提高内容生产效率、丰富内容多样性及提供更加动态且可交互的内容上大有可为。
  - ✓ UGC生成的规模化内容，创造了大量学习素材，帮助AI实现从学习经验到超越并重构已有经验的飞跃性转变。



# 1.3 AIGC为何爆发？数据、算法、算力共振

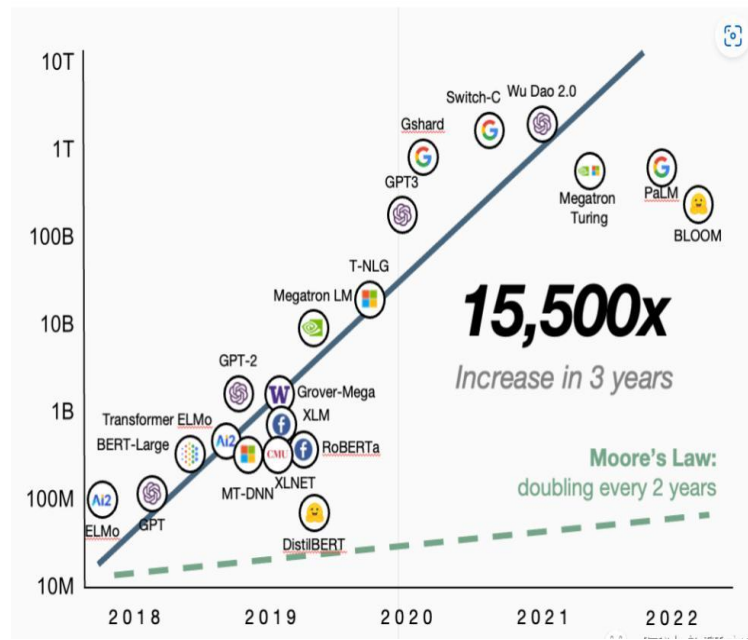
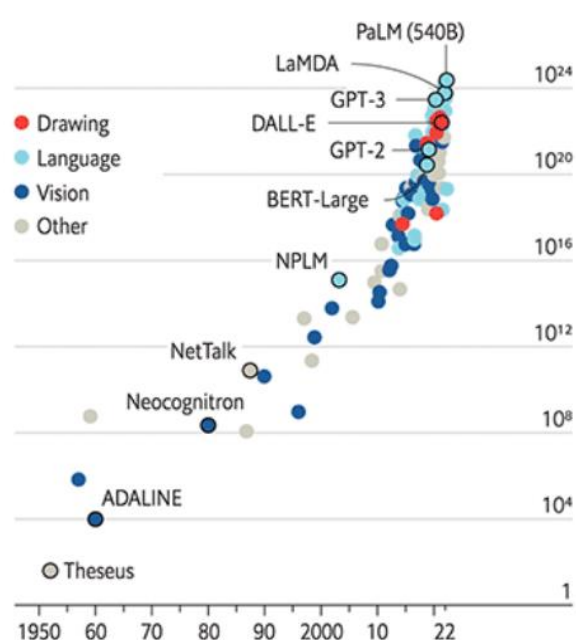
## ■ AIGC发展核心三要素：数据、算力、算法。

- ✓ **算法持续迭代。**2017年推出的Transformer架构的并行训练优势奠定了大模型训练的基础，GPT为代表的预训练模型，通过使用无标注数据预训练及微调，缓解了标注数据不足的问题，并不断提升参数量级及模型通用性，ChatGPT在此基础上加入了利用人类反馈强化学习的训练方法。扩散模型取代GAN成为图像生成领域的主流模型，CLIP模型推动跨模态生成技术的发展。
- ✓ **模型商业化及开源释放创造力。**GPT3的商业化及CLIP及Stable Diffusion模型的开源推动文本生成、文生图产品化浪潮。

图表：深度学习模型的发展

模型	类别	发布年份	特点	影响
GAN	图像生成	2014	对输出结果的控制力较弱，容易产生随机图像、分辨率比较低。	-
Transformer	自然语言模型	2017	引入自注意力机制，能够基于两个单词间的关系进行建模，有效理解单词在上下文中的意思，支持并行训练，使语言模型训练效果达到新高度。	自然语言理解飞跃性发展，平行训练优势逐步发展出超亿规模的大模型，ChatGPT打开AI新纪元
CLIP	文本-图像生成	2021	对文字、图像分别进行训练，不断调整两个模型内部参数，使得模型分别输出文字特征值和图像特征值并确认匹配	多模态技术推动AIGC内容多样性
Diffusion	图像生成	2022	通过增加噪声破坏训练数据来学习，然后找出如何逆转这种噪声过程以恢复原始图像，高效地解决GAN无法训练、训练不稳定的问题。	图像生成技术突破，AI绘画点燃AIGC

图表：训练模型的计算量（左图）以及AI模型参数量（右图）

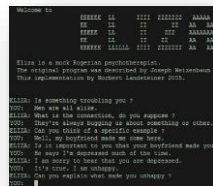
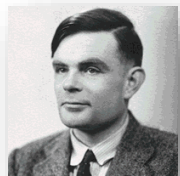


注：计算结果15000x为WuDao2.0中的参数数量除以GPT 1中的参数数量的比值



# 1.3 AIGC算法发展历程：早期受制于算法、算力瓶颈

阶段	萌芽阶段 (1950s-1990s)	沉淀积累阶段 (1990s-2010s)	
特点	受限于技术水平，仅限于小范围实验		
事件	<ul style="list-style-type: none"> <li>◆ 艾伦图灵提出著名的“灵图测试”，给出判断机器是否具有“智能”的实验方法</li> <li>◆ 第一支由计算机创作的弦乐四重奏《依利亚克组曲》完成</li> <li>◆ 世界第一款自然语言聊天机器人 Eliza问世</li> <li>◆ 80年代中期，IBM创造语音控制打字机 Tangora，处理20000个单词</li> <li>◆ 声龙发布了第一款消费级语音识别产品 Dragon Dictate，售价高达9000美元</li> </ul>	<ul style="list-style-type: none"> <li>◆ 世界第一部完全由人工智能创作的小说《1 the road》问世</li> <li>◆ 微软展示全自动同传系统</li> </ul>	
技术	<ul style="list-style-type: none"> <li>◆ 通过关键字扫描和重组完成交互任务</li> </ul>	<ul style="list-style-type: none"> <li>◆ 连续语音的识别迅速发展，统计模型逐步取代模板匹配的方法，隐马尔科夫模型 (HMM) 成为语音识别系统的基础模型</li> </ul>	<ul style="list-style-type: none"> <li>◆ 图形处理器GPU、张量处理器TPU等算力设备性能不断提升，互联网数据规模快速膨胀</li> <li>◆ 深度神经网络算法 (DNN)</li> </ul>
影响	<ul style="list-style-type: none"> <li>◆ 首次提出判断机器是否是人工智能的方法，灵图被称作“人工智能”之父</li> <li>◆ 只有文本界面，自然语言理解尚未取得真正突破</li> </ul>	<ul style="list-style-type: none"> <li>◆ 算法不断完善，语音识别技术快速发展</li> </ul>	<ul style="list-style-type: none"> <li>◆ 可读性不强，拼写错误、辞藻空洞、缺乏逻辑等缺点明显</li> <li>◆ 基本深层神经网络可以将英文语音翻译转化成中文语言</li> </ul>



1956

1957

1969

1980

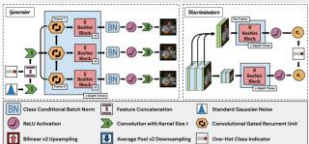
2010

2011



# 1.3 AIGC算法发展历程：模型持续迭代

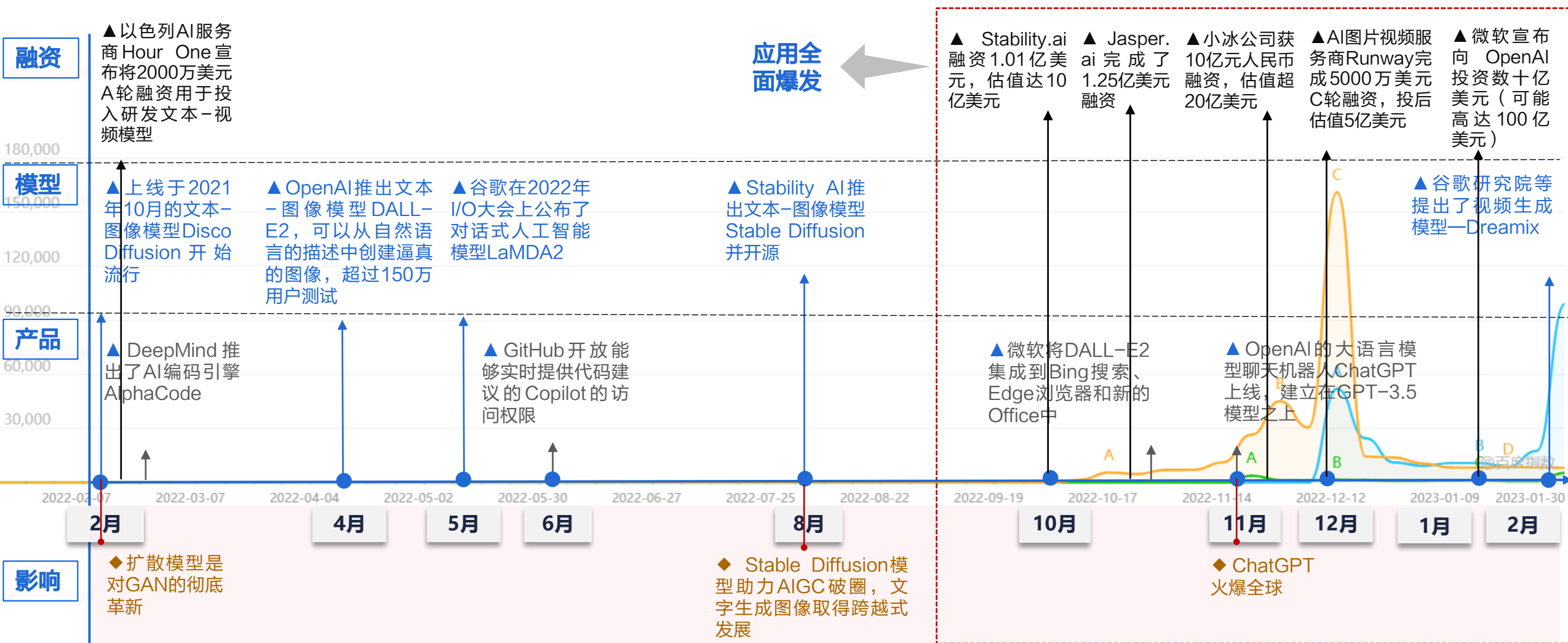
阶段	快速发展阶段（2010s-2021年）							
特点	深度学习算法不断迭代，人工智能生成内容百花齐放，效果逐渐逼真至难以分辨							
事件	<ul style="list-style-type: none"> <li>Ian J. Goodfellow 提出生成对抗网络GAN</li> </ul>	<ul style="list-style-type: none"> <li>“小冰”推出世界首部100%由人工智能创作的诗集《阳光失了玻璃窗》</li> </ul>	<ul style="list-style-type: none"> <li>英伟达发布StyleGAN模型可自动生成高质量图片，几个月后发布了StyleGAN2</li> </ul>	<ul style="list-style-type: none"> <li>人工智能生成画作在佳士得以43.25万美元成交，成为首个出售的人工智能艺术品</li> </ul>	<ul style="list-style-type: none"> <li>DeepMind发布DVD-GAN模型用以生成连续视频</li> </ul>	<ul style="list-style-type: none"> <li>2020年，伯克利的Pieter Abbeel等人提出去噪扩散概率模型（DDPM）</li> </ul>	<ul style="list-style-type: none"> <li>Open AI推出了GPT-3，拥有超过1750亿的训练参数量，被誉为“万能生成器”</li> </ul>	<ul style="list-style-type: none"> <li>Open AI推出DALL-E，主要应用于文本与图像交互生成内容，同年将跨模态深度学习模型CLIP开源</li> </ul>
技术	<ul style="list-style-type: none"> <li>GAN</li> </ul>	<ul style="list-style-type: none"> <li>Transformer架构提出</li> </ul>	<ul style="list-style-type: none"> <li>StyleGAN</li> </ul>	<ul style="list-style-type: none"> <li>GAN</li> </ul>	<ul style="list-style-type: none"> <li>DVD-GAN</li> </ul>	<ul style="list-style-type: none"> <li>DDPM</li> </ul>	<ul style="list-style-type: none"> <li>自然语言生成模型GPT-3</li> </ul>	<ul style="list-style-type: none"> <li>关联文字和图像，并且关联特征非常丰富</li> </ul>
影响	<ul style="list-style-type: none"> <li>被广泛应用于图像生成、语音生成等场景中</li> </ul>	<ul style="list-style-type: none"> <li>Transformer架构推动深度学习算法突破发展，迸发出大模型</li> </ul>				<ul style="list-style-type: none"> <li>在图像生成方面比GAN更优，扩散模型威望大幅提升</li> </ul>	<ul style="list-style-type: none"> <li>文本生成迎来重大突破，GPT-3庞大的运行规模使得它不仅能答题、写论文和生成代码等，还能编写曲谱、写小说等</li> </ul>	<ul style="list-style-type: none"> <li>CLIP模型搜集了大量数据，为输入文本生成图像/视频应用的落地奠定了基础</li> </ul>



资料来源：信通院《人工智能生成内容（AIGC）白皮书》，机器之心Pro，腾讯新闻，《Denoising Diffusion Probabilistic Models》，《Attention is All You Need!》，腾讯云，环球网，搜狐新闻，国海证券研究所

# 1.3 AIGC算法发展历程：从模型到应用

井喷式发展阶段：2022年——AIGC元年



# 1.3 AIGC算法发展历程：OpenAI VS Google

## ■ OpenAI：非盈利性转向封顶盈利性公司，估值达290亿美金。

- ✓ 2015年由马斯克等人创立的非盈利人工智能研究公司，启动资金10亿美金，成立初衷是与其它机构合作进行AI相关研究，并开放研究成果以促进AI技术发展，防止垄断。
- ✓ 核心团队为CEO Sam Altman、Greg Brockman、Ilya Sutskever，大都技术出身，在通用AI领域经验丰富。
- ✓ 2019年宣布从“非盈利”性质过度到“封顶盈利性”，之后获微软10亿美元战略投资，并开启与微软在产品上的合作。
- ✓ 据美国财经媒体Semafor报道，微软预计向ChatGPT的开发者OpenAI投资100亿美元，OpenAI投后估值将高达290亿美元。

图表：OpenAI核心人员



**Sam Altman** CEO、创始人    
 **Greg Brockman** 总裁 (原CTO)    
 **Mira Murati** CTO    
 **Ilya Sutskever** 首席科学家

图表：OpenAI融资过程

序号	日期	融资轮次	融资金额	投资机构
1	2021年	A轮	2.5亿美元	-
2	2021年	种子轮延期	-	-
3	2020年	种子轮	-	Matthew Brown Companies
4	2019年	战略融资	10亿美元	微软
5	2019年	pre-种子轮	-	ReidHoffman 慈善基金、Khosla Ventures
6	2018年	天使轮	-	Gabe Newell、Jaan Tallinn、Ashton Eaton 和Brianne Theisen-Eaton等
7	2017年	亲友轮	-	-
8	2016年	亲友前轮	10亿美元	Elon Musk、Sam Altman、Linkdin 的联合创始人 Reid Hoffman、Paypal 联合创始人 Peter Thiel、YC 联合创始人Jessica Livingston等

注：数据截止到2023年1月

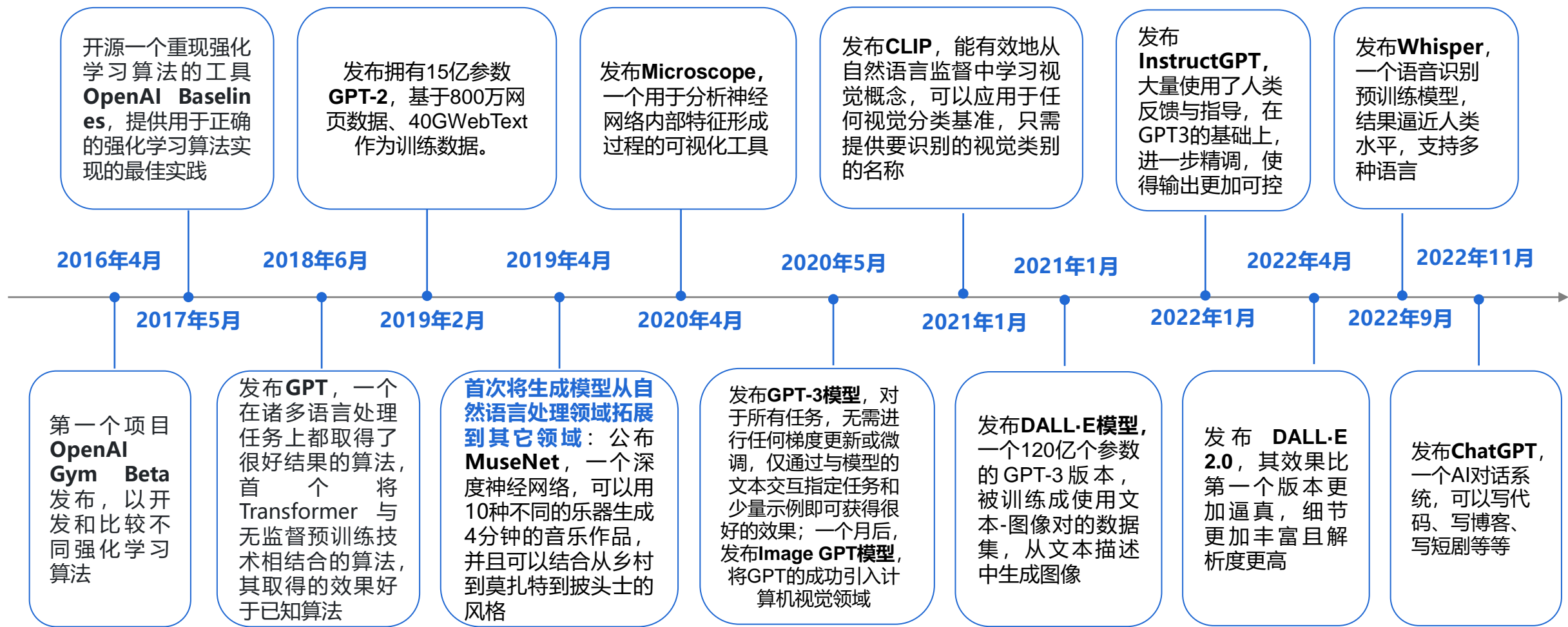
图表：OpenAI部分对外投资

序号	被投公司	主营业务	投资阶段
1	Anysphere	AI工具	种子轮
2	Atomic Semi	芯片制造	种子轮
3	Cursor	代码编辑	种子轮
4	Diagram	设计工具	种子轮
5	Harvey	AI法律顾问	种子轮
6	Kick	会计软件	种子轮
7	Milo	家长虚拟助理	种子轮
8	qqbot.dev	开发者工具	种子轮
9	EdgeDB	开源数据库	A轮
10	Mem Labs	记笔记应用	A轮
11	Speak	AI英语学习平台	B轮
12	Descript	音视频编辑应用	C轮

注：数据截止到2023年1月

# 1.3 AIGC算法发展历程：OpenAI推动AI算法模型发展

## ● OpenAI技术发展历史





# 1.3 AIGC算法发展历程：OpenAI携手微软

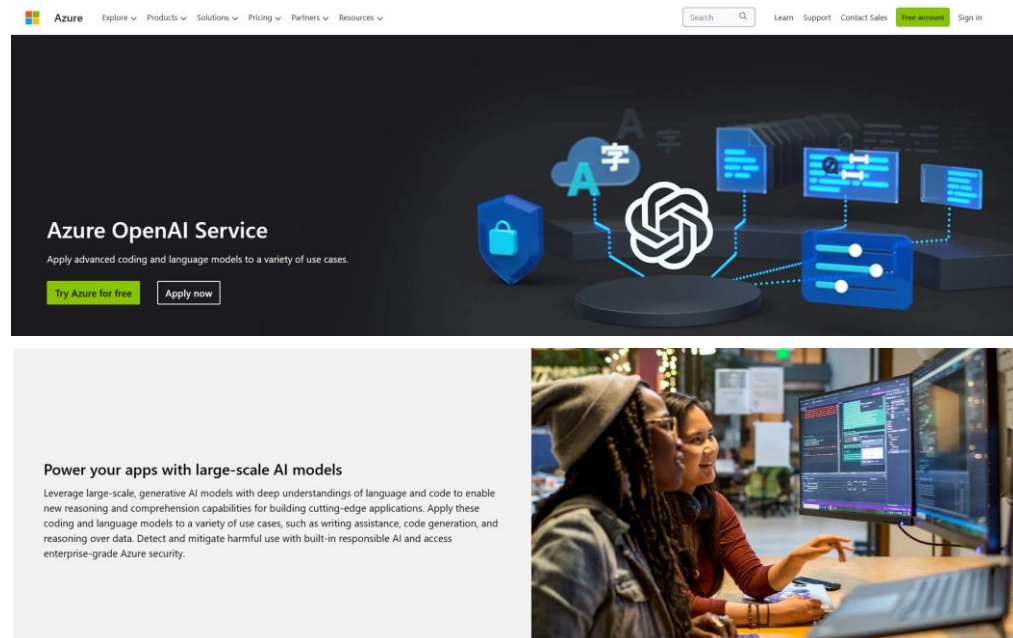
## ■ OpenAI携手微软，获得资金支持，落地场景，借力微软云计算领域布局。

- ✓ **资金+算力**：2019年7月，OpenAI 接受了微软10亿美元的战略投资，同时将把微软的Azure作为其独家云计算供应商；2021年，微软加注投资，具体金额未公布；2023年，微软预计向OpenAI再投资100亿美元，在满足首批投资者收回初始资本后，微软将获得OpenAI 75%利润，直到收回投资。
- ✓ **业务协作**：2021年，微软推出了Azure OpenAI服务预览；2022年，微软将DALL-E2模型集成到了Azure OpenAI、Microsoft Designer、Bing Image Creator中。2023年1月，Azure OpenAI服务正式发布，企业可以申请访问包括GPT-3.5、Codex和DALL-E2等AI模型，之后还可能通过Azure OpenAI服务访问ChatGPT。2023年2月，微软推出ChatGPT支持的最新版本Bing搜索引擎与Edge浏览器，增加聊天写作功能。此外，微软计划将ChatGPT引入Office产品中，进一步提升市场份额。

图表：OpenAI盈利后利润分配的四阶段

1	优先保证埃隆马斯克、彼得泰尔、雷德霍夫曼等首批投资者收回初始资本	微软无分成
2	微软将有权获得OpenAI 的利润，直至收回其130亿美元投资	微软获得75%利润
3	在OpenAI的利润达到920亿美元后，微软在OpenAI的持股比例将下降到49%，剩余的利润由其他风险投资者和OpenAI的员工分享	微软获得49%利润
4	在利润达到1500亿美元后，微软和其他风险投资者的股份将无偿转让给OpenAI的非营利基金	微软无分成

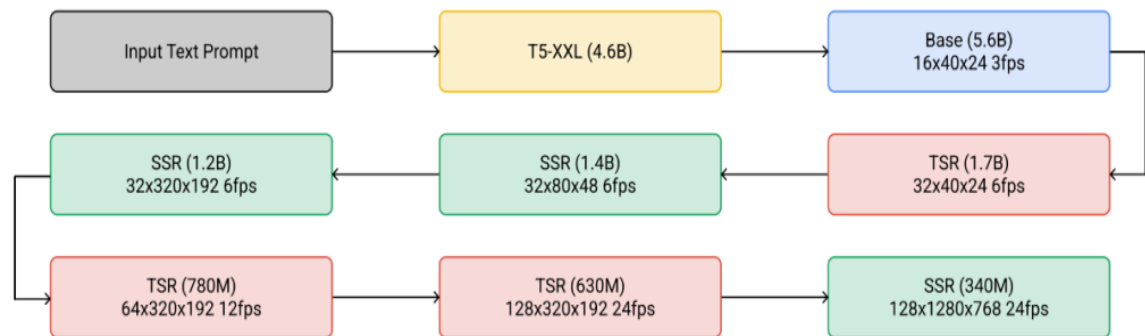
图表：Azure OpenAI服务官网



# 1.3 AIGC算法发展历程：谷歌持续重注AIGC研究

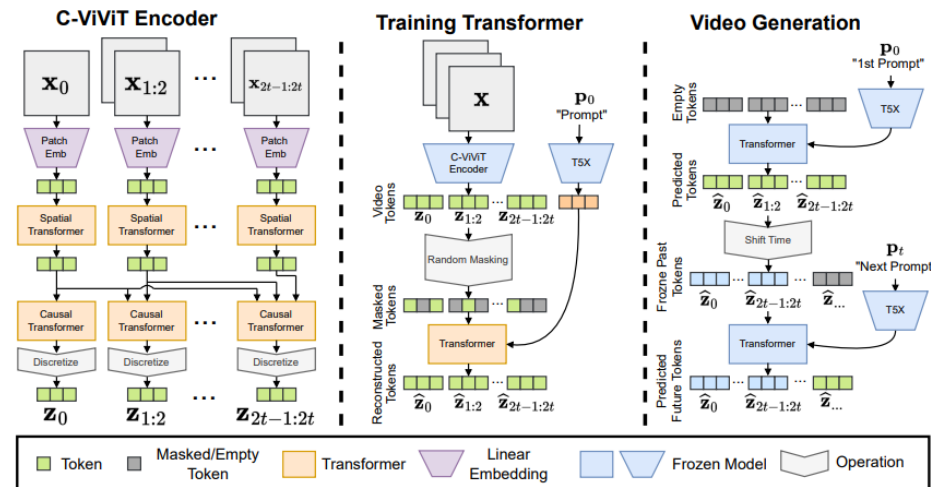
- 2022年11月在AI年度活动上谷歌发布四项最新成果，其能够根据文本提示生成高分辨率的长视频、3D模型、音乐、代码、文字内容等。
- ✓ **结合Imagen Video和Phenaki两大模型的优势，推出超长连贯性视频生成模型**：Imagen Video是基于级联视频扩散模型的文本条件视频生成系统，即给出文本提示，就可以通过一个由frozen T5文本编码器、基础视频生成模型、级联时空视频超分辨率模型组成的系统来生成高清视频。Phenaki模型可通过一系列提示在开放域中生成所有时间段的视频，是谷歌首次以时间变量提示生成视频。
- ✓ **LaMDA Wordcraft**：在大语言模型LaMDA基础上开发的、能辅助专业作家写作的AI写文工具，帮助创作者突破“创作瓶颈”。
- ✓ **Audio LM**：具备“长期连贯性”的高质量音频生成框架，不需要任何文字或音乐符号表示的情况下，只在极短（三四秒即可）的音频样本基础上训练，可生成自然、连贯、真实的音频结果，不限语音或者音乐。
- ✓ **文字生成3D模型**：通过结合Imagen和最新的神经辐射场 (Neural Radiance Field) 技术，谷歌开发出了DreamFusion技术，可根据现有文字描述，生成具有高保真外观、深度和法向量的3D模型，支持在不同光照条件下渲染。
- **将推出Bard对话机器人**。2023年2月，谷歌宣布将推出Bard AI聊天机器人，由谷歌大型语言模型LaMDA支持，但参数量更少，使公司能够以更低的成本提供该技术，Bard能在获得简单提示的情况下生成详细答案。

图表：谷歌Imagen模型架构



- 整个架构共有7个子模型（1个T5文本编码器、1个基础视频扩散模型、3个SSR扩散模型、3个TSR扩散模型），共116亿个参数
- ✓ 文本编码器——将文本prompt编码为text\_embedding;
  - ✓ 基础视频扩散模型——以文本为条件，生成初始视频;
  - ✓ SSR——提高视频的分辨率;
  - ✓ TSR——提高视频的帧数

图表：谷歌Phenaki模型架构



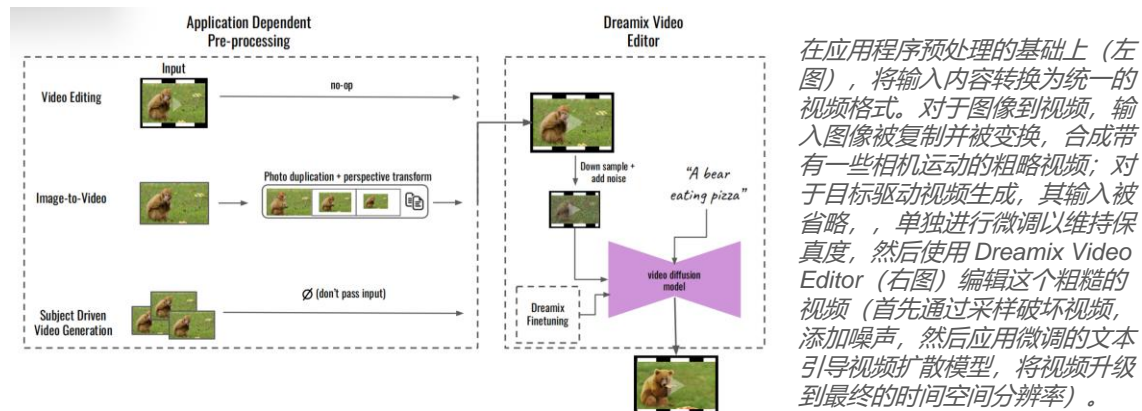
主要包含两大部分：一个将视频压缩为离散嵌入（即 token）的编码器-解码器模型和一个将文本嵌入转换为视频token的transformer模型



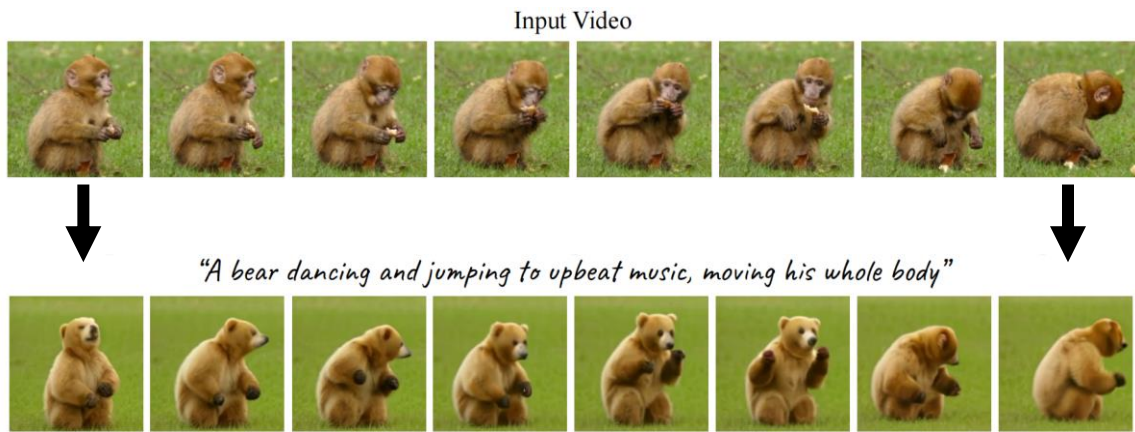
# 1.3 AIGC算法发展历程：谷歌持续重注AIGC研究

■ 2023年2月2日，谷歌研究院等提出了一种视频生成新模型——Dreamix，受到了AI作图UniTune的启发，将文本条件视频扩散模型（video diffusion model, VDM）应用于视频编辑。核心是通过两种主要思路使文本条件VDM保持对输入视频的高保真度：（1）不使用纯噪声作为模型初始化，而是使用原始视频的降级版本，通过缩小尺寸和添加噪声仅保留低时空信息；（2）通过微调原始视频上的生成模型来进一步提升对原始视频保真度。微调确保模型了解原始视频的高分辨率属性，对输入视频的简单微调会促成相对较低的运动可编辑性，这是因为模型学会了更倾向于原始运动而不是遵循文本prompt。

图表：Dreamix模型原理

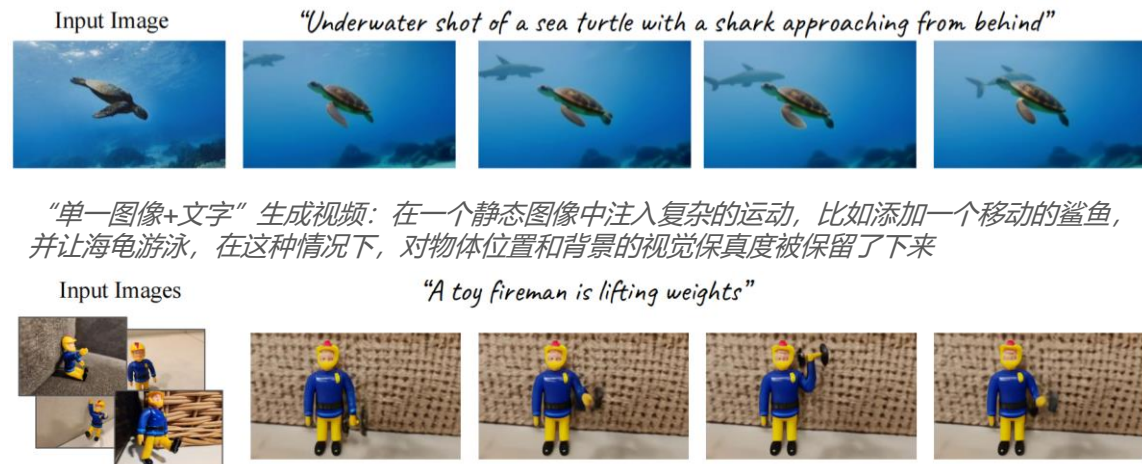


图表：Dreamix模型应用于视频编辑



将吃东西的猴子（上面一排）变成跳舞的熊（最下面排），改变外观和运动，但保持对颜色、姿势、物体大小和拍摄角度的保真度，从而产生了一个时间上一致的视频

图表：Dreamix模型应用于图像生成视频



"单一图像+文字" 生成视频：在一个静态图像中注入复杂的运动，比如添加一个移动的鲨鱼，并让海龟游泳，在这种情况下，对物体位置和背景的视觉保真度被保留了下来

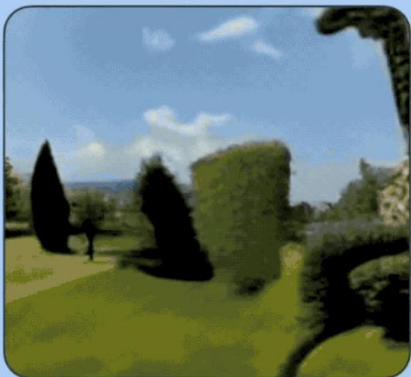
"多图像+文字" 生成视频：在给定的主题下，能够提取给定多个图像的主题的视觉特征，然后在不同的场景中制作动画

# 1.3 AIGC算法发展历程：谷歌持续重注AIGC研究

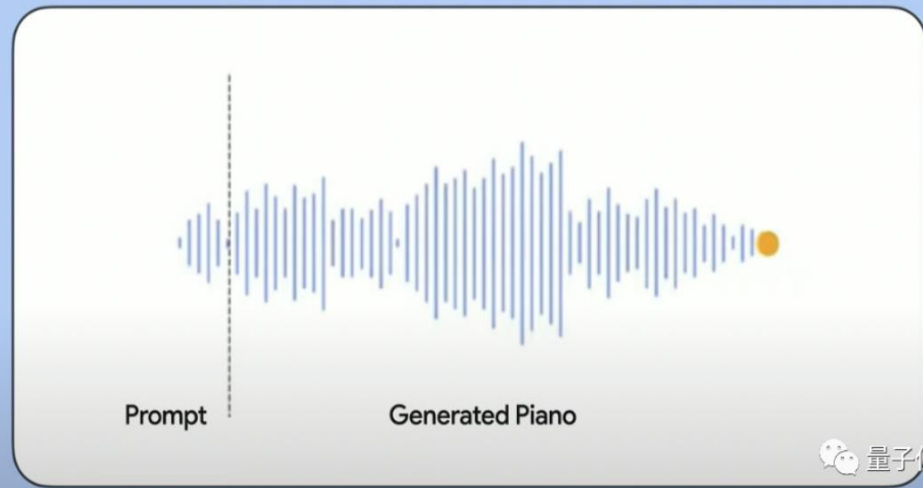
## 超长连贯性视频生成模型

### Phenaki + Imagen Video Combination

"A blue balloon stuck in the branches of a Redwood tree"  
 "Camera pans from tree with single blue balloon to the zoo entrance"  
 "Camera pans to the zoo entrance"  
 "Camera moves very quickly into the zoo"  
 "First person view of flying inside a beautiful garden"  
 "The head of a giraffe emerges from the side"  
 "Giraffe walks towards a tree"  
 "Zoom into the giraffe's mouth"  
 "Giraffe gets close to a branch and picks a blue balloon"  
 "A single helium blue balloon with a white string is flying towards a giraffe's head", "  
 "Giraffe chewing with the blue balloon nearby"  
 "Camera tilting up following the single blue balloon flying away"



## Audio LM

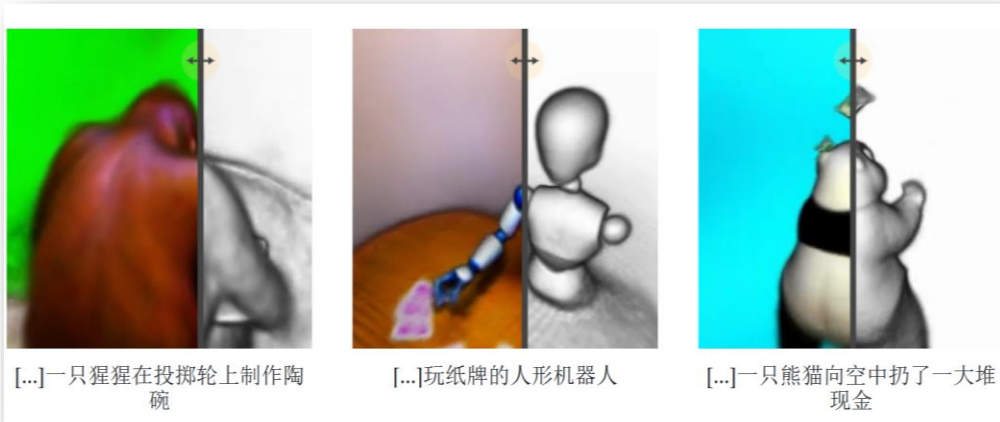


量子位

## 刘宇昆在Wordcraft撰写的短篇小说《Evaluative Soliloquies》

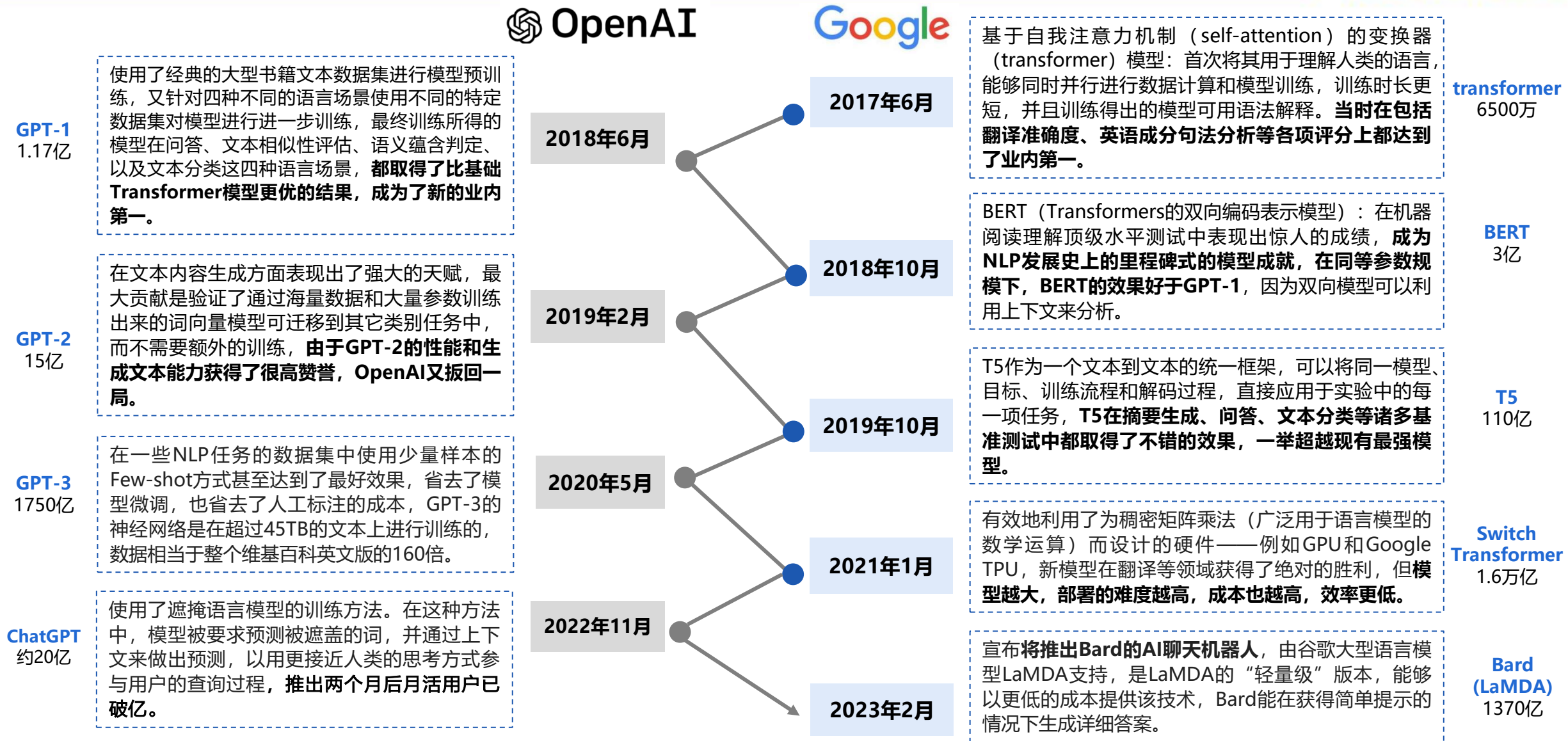


## DreamFusion生成的3D模型





# 1.3 AIGC算法发展历程：你追我赶，持续迭代



资料来源：做AI做的事儿公众号，网易新闻，澎湃新闻，谷歌研究论文《LaMDA: Language Models for Dialog Applications》，国海证券研究所

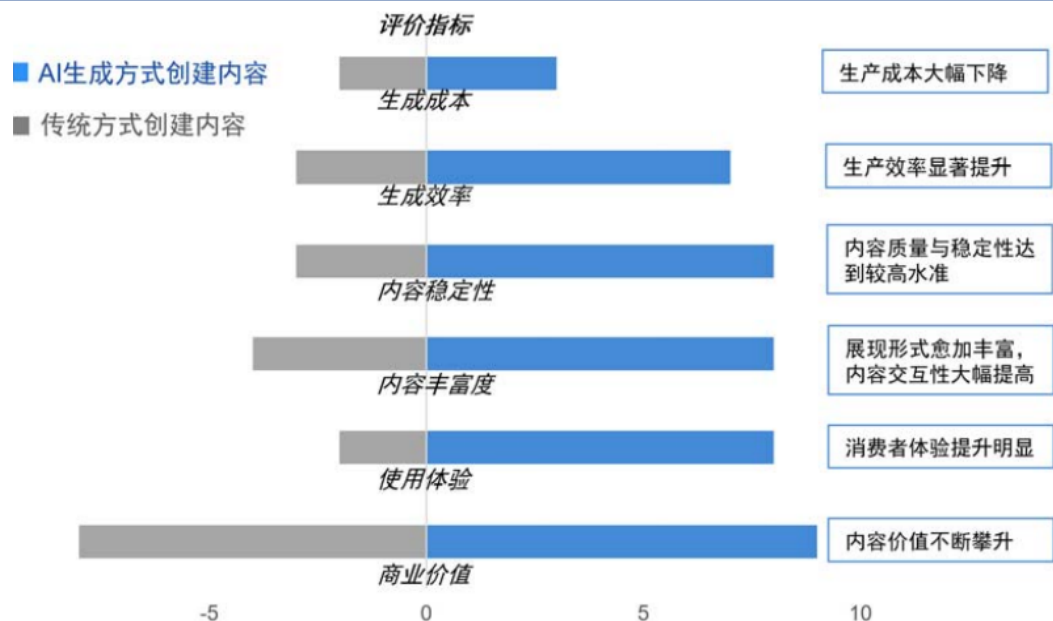
注：模型后面的数据为模型的参数数量

# 1.4 AIGC市场空间：从决策走向创造

## 推动内容生产向高效率和更富创造力方向发展，与多产业融合。

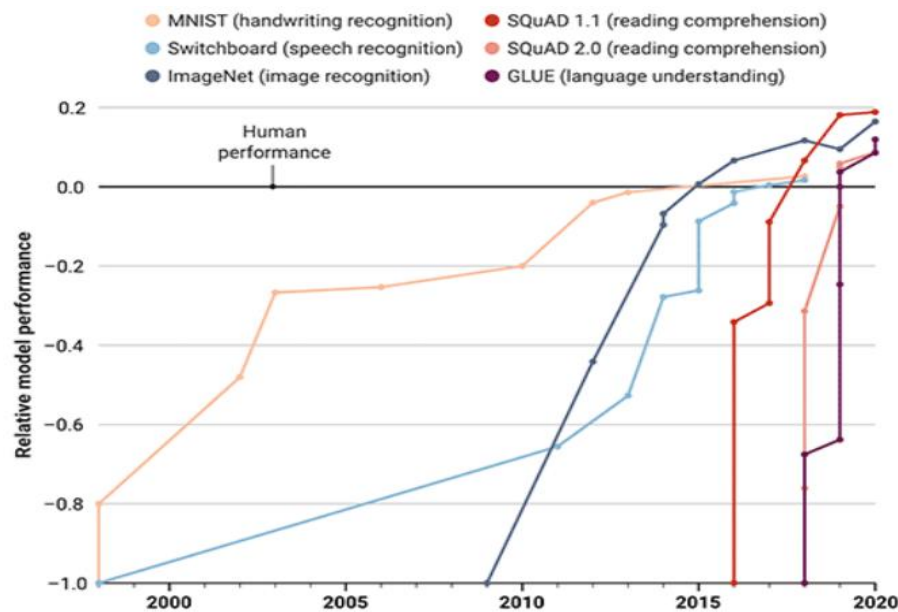
- 不仅是降本增效，更是个性化内容生成。AI不仅能够以优于人类的制造能力和知识水平承担信息挖掘、素材调用、复刻编辑等基础性机械劳动，从技术层面实现以低边际成本、高效率的方式满足海量个性化需求。根据Sequoiacap，近年来AI模型在手写、语音和图像识别、阅读理解和语言理解方面的表现逐渐超过了人类的基准水平。而且AI让所有人都能够成为“艺术家”，可无时无刻生成更有创造力、更个性化的内容。
- 通过支持AI生成式内容与其他产业的多维互动、融合渗透从而孕育新业态新模式，为各行各业创造新的商业模式，提供价值增长新动能。

图表：内容生成及创建的评价指标



注：内容生成及创建的评估指标根据行业公允评估方法设计，采用影谱科技AI智能影像生成引擎AGC为数测平台

图表：AI模型在语音识别、图像识别、阅读理解、语言理解等方面的表现

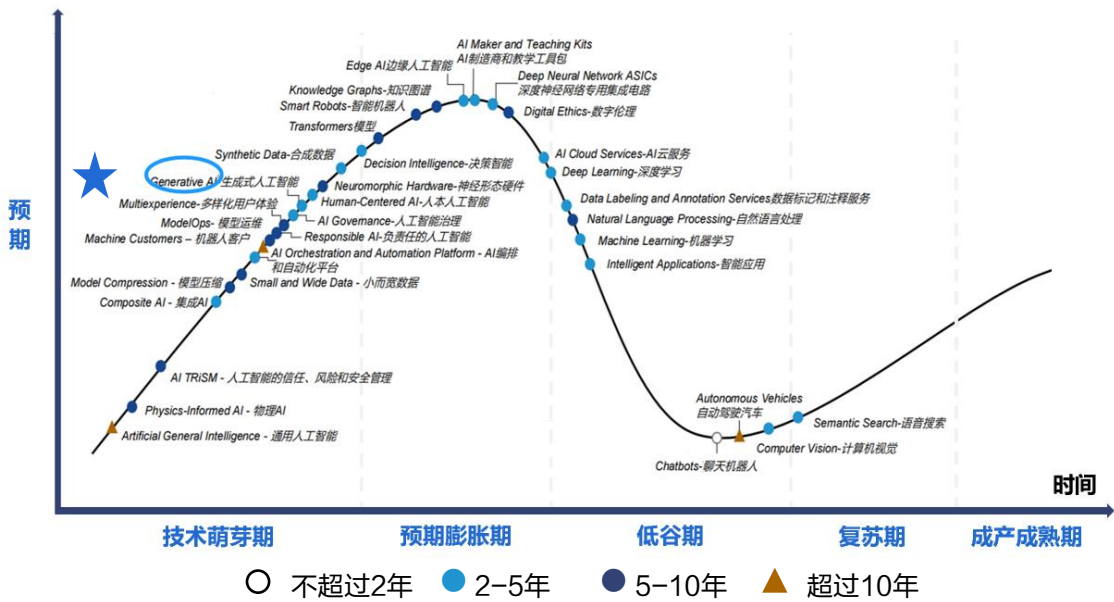


# 1.4 AIGC市场空间：2025年国内市场应用规模有望超2000亿元

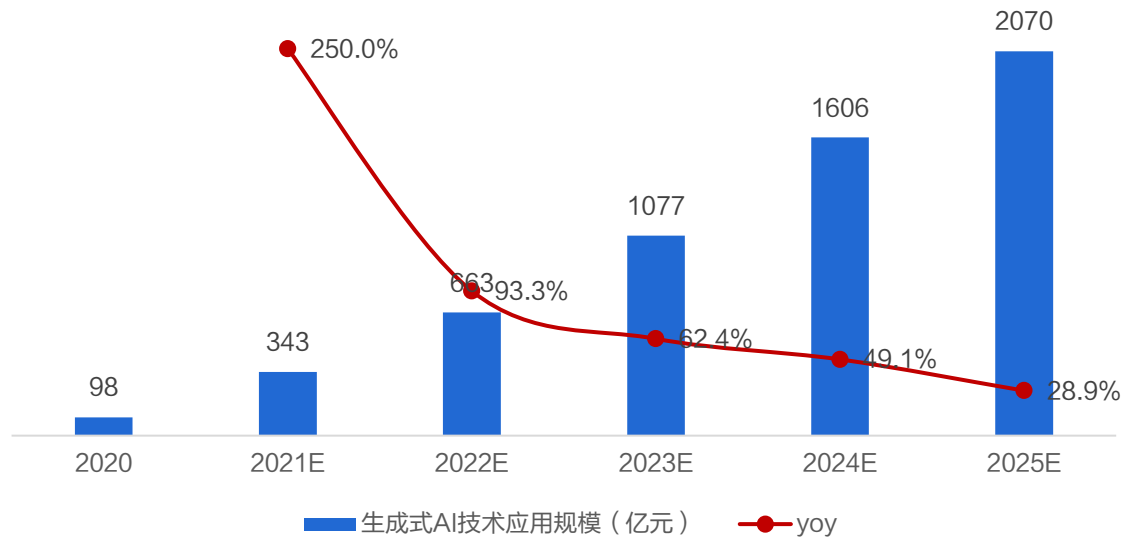
生成式AI居于Gartner2022年战略技术首位，2030年全球市场规模有望超万亿美元，2025年国内应用规模有望突破2000亿元。

- 2021年，Gartner发布了12项2022年重要战略技术趋势，生成式AI居于战略首位。Gartner预测至2023年将有20%的内容被生成式AI所创建；至2025年生成式AI产生的数据将占有所有数据的10%（目前不到1%）。
- ✓ 根据2022年红杉《Generative AI: A Creative New World》，未来2-3年AIGC初创公司和商业落地方案将持续增加，将产生数万亿美元经济价值。
- ✓ 根据Gartner“2022年人工智能技术成熟度曲线”，生成式AI仍处于技术萌芽期，预计将在2-5年内实现规模化应用。
- 根据《中国AI数字商业展望2021-2025》，到2025年中国生成式AI技术应用规模预计上升至2070亿元，2020-2025年年均复合增长率高达84.1%。

图表：Gartner2022年人工智能技术成熟度曲线



图表：中国生成式AI技术应用规模（单位：亿元）



注：生成式AI应用规模的统计口径为应用生成式AI技术生成的数字内容的市场规模，统计方式为数字内容市场规模乘以生成式AI渗透率，以上规模的推导考虑国家商务局、GARTNER，第三方调研机构等多个数据源。



# 1.4 AIGC市场空间：国内传媒领域潜在应用空间超1000亿元

■ AIGC市场空间 = 各代表性行业内容成本\*对应的AI辅助制作比例

- ✓ 内容成本测算=各代表性行业的市场规模\*预计的内容/版权成本占收入比重。电影行业参考博纳影业投资业务毛利率（约40%）和投资方的分账比例（约38%），计算得出电影制作成本占票房的23%；在线音乐参考腾讯音乐/网易云的内容服务成本占收比，并按照二者的市场份额进行加权。
- ✓ AI辅助制作的比​​例：图片生成、音乐生成领域的应用相对较成熟，预计未来AI辅助生成的比例分别为65%/60%，游戏领域的AI生成主要集中在图片/音乐/NPC上，预计占比30%；AI辅助视频类（长视频/电影/IPTV+OTT）内容生成还在初始阶段，预计AI辅助生成比例较低（15%-20%）；目前网络直播行业虚拟主播应用较广，预计AI辅助制作比例35%。综上，我们匡算国内AIGC市场的理论空间超1000亿元。

图表：AIGC在国内传媒领域潜在市场空间测算

内容	2022年市场规模 (亿元)	内容/版权成本占收入比重*	内容成本 (亿元)	AI辅助制作的比​​例	AIGC市场空间匡算 (亿元)	*注
游戏 (国内)	2658.8	20.0%	531.8	30%	159.5	参考游戏行业的研发费用占收比20%
游戏 (出海)	173.5	20.0%	34.7	30%	10.4	参考游戏行业的研发费用占收比20%
电影	641.0	22.8%	146.1	20%	29.2	参考博纳影业投资业务毛利率40%和投资方的票房分账比例38%
长视频	1626.3	73.8%	1200.4	15%	180.1	参考爱奇艺内容成本占收比74%和芒果超媒的互联网视频业务毛利率31%
IPTV+OTT	279.9	30.8%	86.2	15%	12.9	参考新媒股份内容成本占收比36%、芒果超媒运营商业毛利率74%
短视频	3055.0	5.0%	152.8	40%	61.1	参考快手内容成本占收比5%
在线音乐	495.0	64.4%	319.0	60%	191.4	参考腾讯音乐/网易云音乐的内容服务成本60%/80%
图片版权	264.7	31.5%	83.3	65%	54.2	参考视觉中国的版权费占收比32%
网络文学	214.0	13.4%	28.7	15%	4.3	参考中文在线、阅文集团的内容分销成本占收比16%/11%
网络直播	1936.6	50.0%	968.3	35%	338.9	参考直播行业主播分成比例大概为5: 5
<b>合计</b>	<b>11344.8</b>	<b>31.3%</b>	<b>3551.3</b>	<b>29%</b>	<b>1042.0</b>	

注：电影行业市场规模参考疫情前2019年电影市场票房，AI辅助制作的比​​例均为匡算值



# 1.5 商业模式：按调用量收费、SaaS订阅收费、增值服务、解决方案等

## ■ AIGC的潜在客户主要包括2B端内容生产公司和2C端用户：

- ✓ 2B：在PGC领域实现内容创作高效化，提高PGC活跃度和灵活性。AIGC能够克服人力不足，降低内容生产成本。客户主要为资讯媒体、音乐流媒体、游戏公司、视频平台、影视制作公司等，如协助影视公司制作电影/剧集视频片段。
- ✓ 2C：在UGC领域实现内容创作低门槛和较高专业度，扩充UGC人群。AIGC能够激发C端用户灵感，且不需要用户具有极强的专业知识，每个人都可以成为创作者。客户主要为画家、写手、歌手等，如协助音乐小白创作专属于个人的歌曲。

■ 国内AIGC商业模式尚未成型。以写作机器人、自动配音等场景为例，大部分产品仍处在免费试用的“流量吸引+平台改良”阶段。此外，部分公司将AIGC用于协助自身原有商业体系，如腾讯开发的AI Bot应用于腾讯游戏中，阿里的智能语音服务主要应用于微信，字节跳动则主要基于短视频场景研究AI赋能。

图表：AIGC商业模式



1

AIGC破圈元年

2

AIGC应用场景

2.1 文本生成

2.2 音频生成

2.3 图像生成

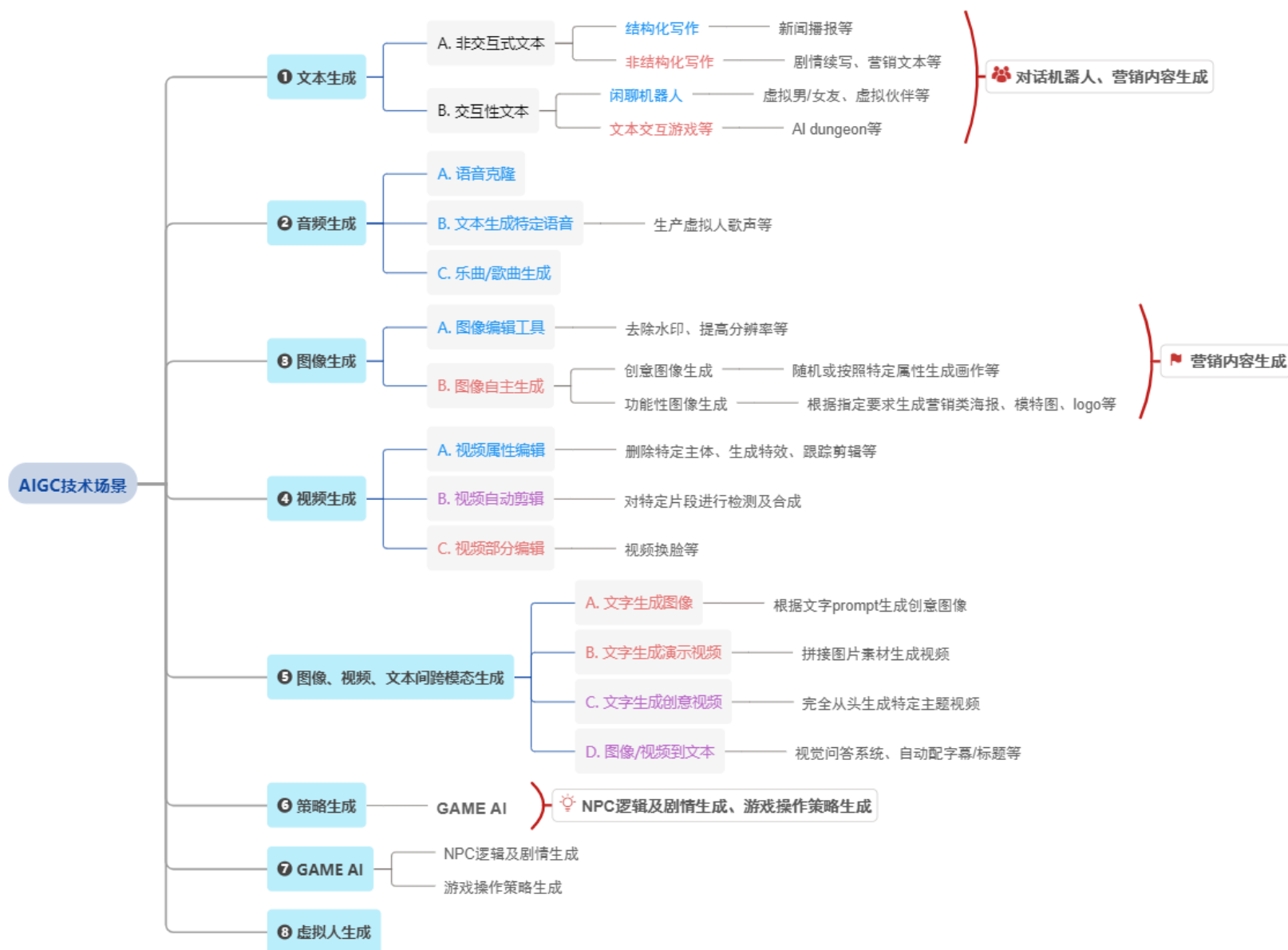
2.4 视频生成

2.5 跨模态生成

2.6 策略生成

3

产业链及相关公司



■ 较为广泛地实际应用、技术细节仍待进一步提升

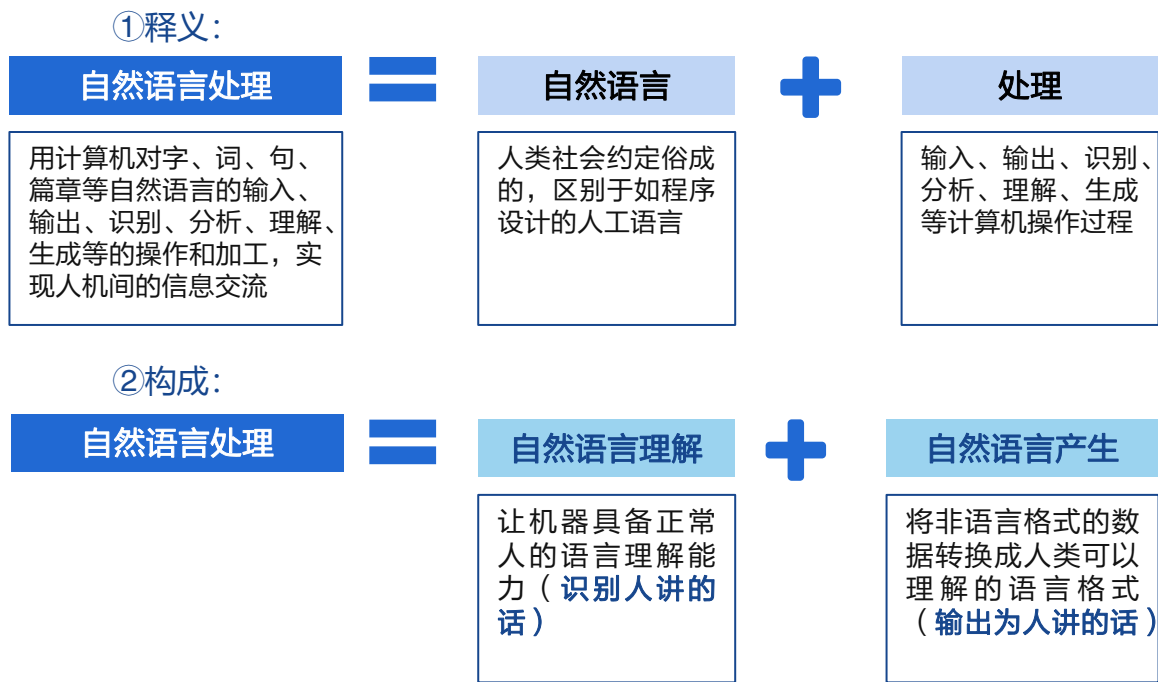
■ 底层技术原理基本明确，预计1-2年内将规模化应用

■ 底层技术原理仍待完善

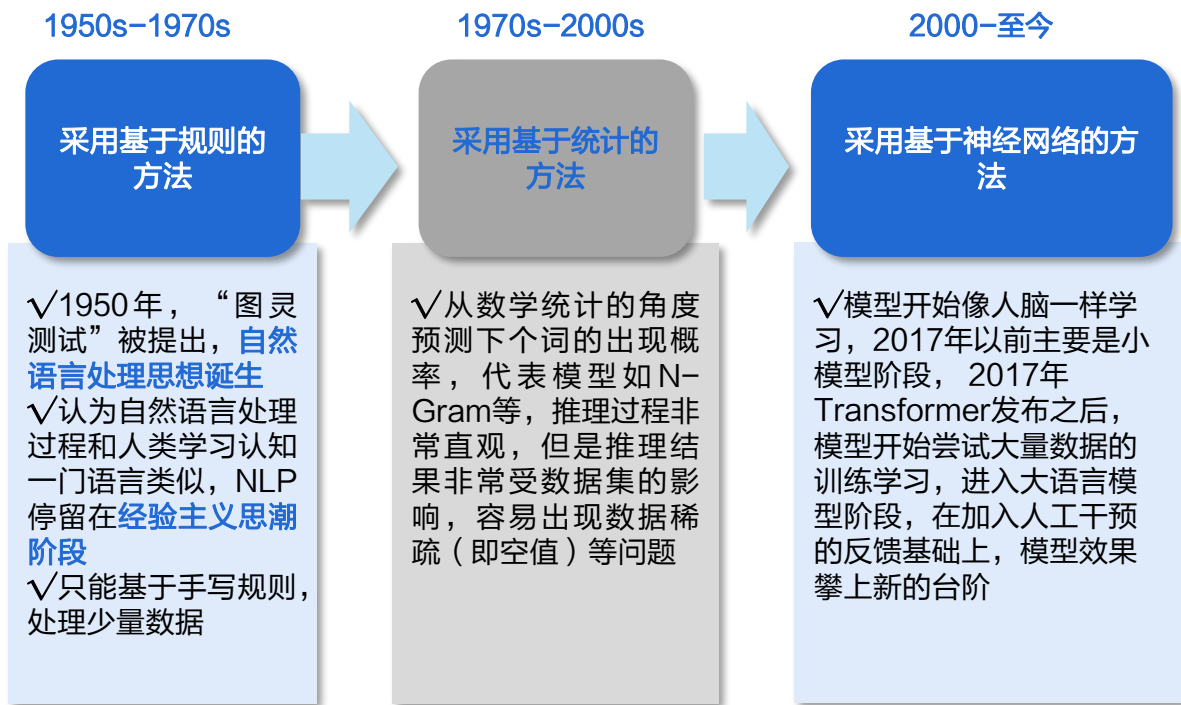
# 2.1 文本生成：基于NLP技术，受益于预训练语言模型突破发展

- 自然语言处理技术（NLP）是文本生成的基础。NLP探索计算机和人类（自然）语言之间相互作用，研究实现人与计算机之间用自然语言进行有效通信的各种理论和方法。最早的自然语言处理研究工作是机器翻译，后逐渐向文本摘要、分类、校对、信息抽取、语音合成、语音识别等方面深入。
- 从基于规则的经验主义到基于统计的理性主义，再到基于深度学习的方法，NLP在70年历程中逐渐发展进步。受益于预训练语言模型的突破发展，Transformer等底层架构不断精进，NLP取得跨越式提升。

图表：自然语言处理的两种解释



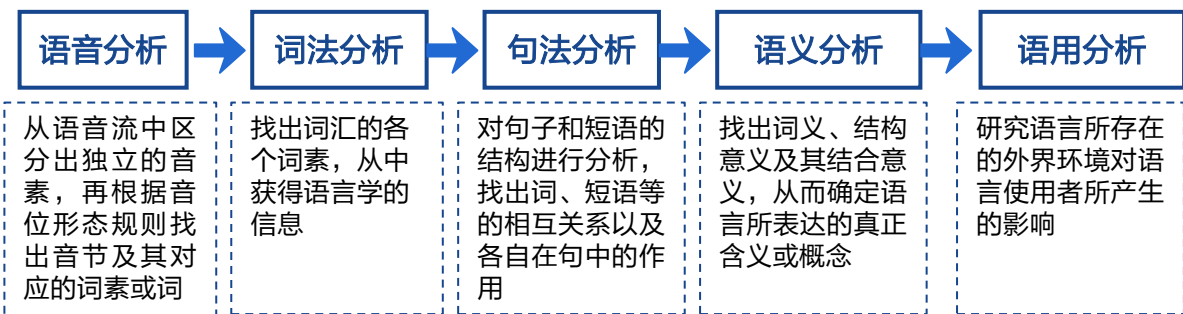
图表：NLP发展阶段



# 2.1 文本生成：难点在于自然语言的复杂性及如何生成富有“人味”的语言

■ NLP技术分为自然语言理解（NLU）和自然语言生产（NLG）两个核心任务，目前难点在于自然语言的复杂性使AI的理解程度不高以及如何生成富有“人味”的语言。

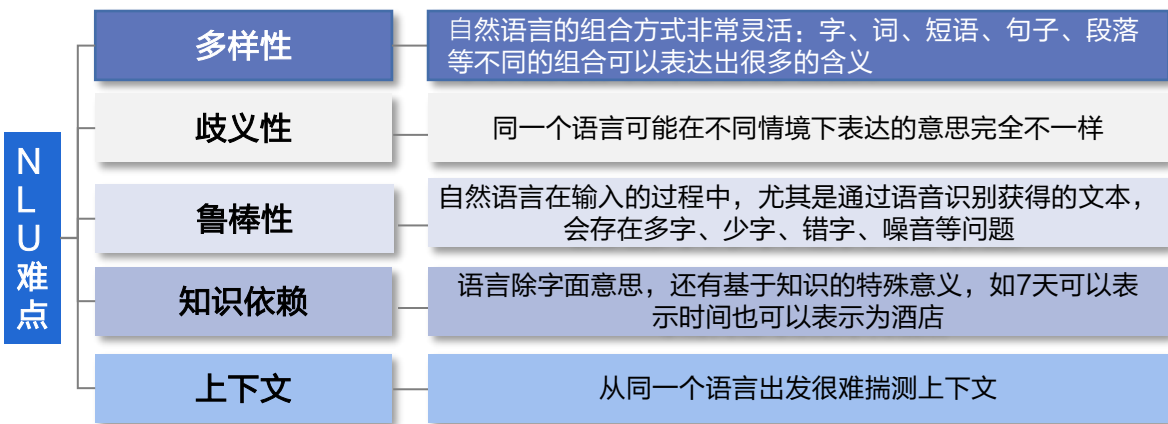
图表：自然语言理解层次逐渐加深



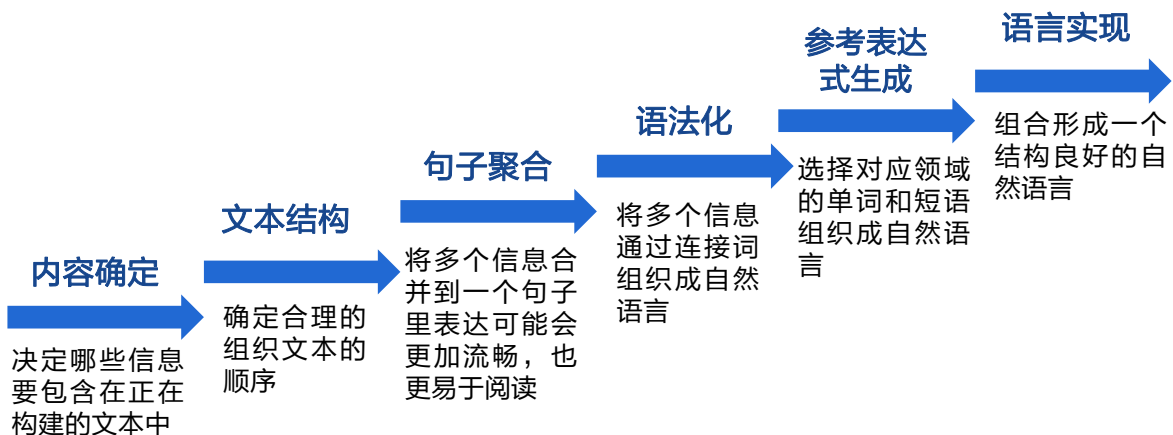
图表：自然语言生成从低等级到高等级



图表：自然语言理解的五大难点



图表：自然语言生成的六大步骤

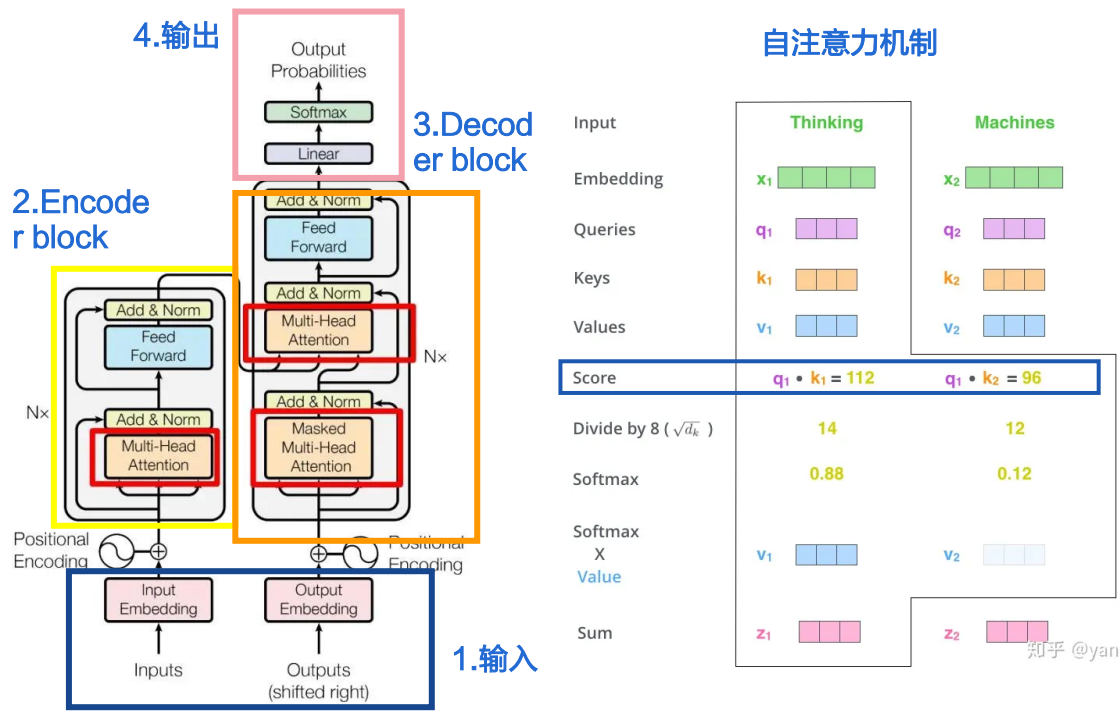




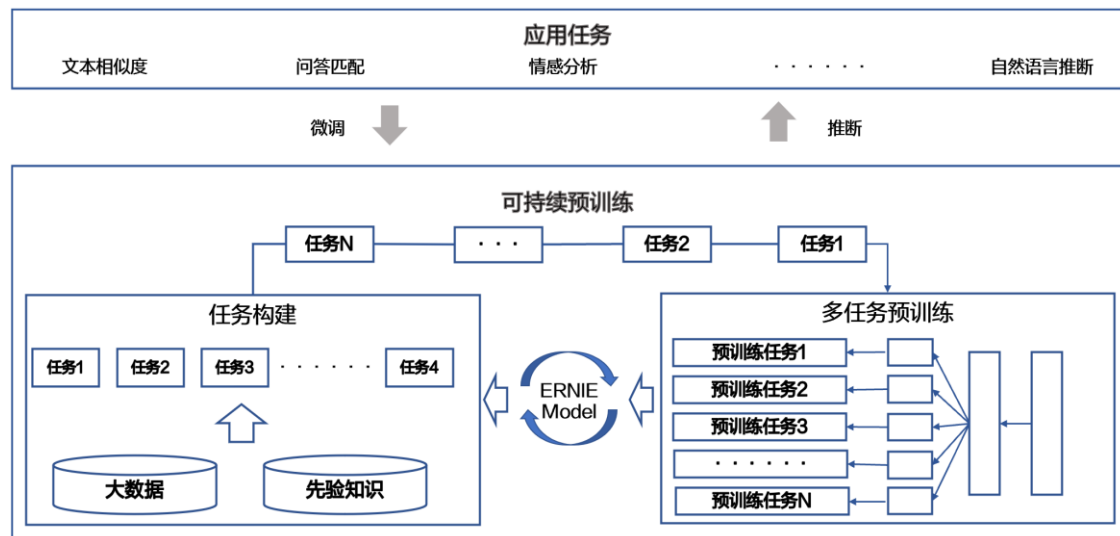
# 2.1 文本生成：预训练语言模型发展推动NLP研究

- **预训练语言模型 (PTM)**：在大规模无监督的语料上进行长时间的无监督或自监督预先训练，获得通用的语言建模和表示能力。应用到实际任务上时不需要做大改动，只需增加针对特定任务获得输出结果的输出层，并使用任务语料对模型进行少许训练。
- **Transformer架构的并行化训练优势，促进预训练语言模型突破发展**。2017年，Google发布文章《Attention is all you need》，提出了解决sequence to sequence问题的transformer架构，引入了自注意力机制，能够基于两个单词间的关系进行建模，按输入数据各部分重要性的不同分配不同的权重，有效理解单词在上下文中的意思，支持并行训练，使语言模型训练效果达到新高度。
- 2020年，1750亿参数的GPT-3在问答、摘要、翻译、续写等语言类任务上均展现出了优秀的通用能力，证明了“大力出奇迹”在语言类模型上的可行性，其文本生成能力已被直接应用于Writesonic、Conversion.ai、Snazzy AI、Copysmith、Copy.ai、Headlime等文本写作/编辑工具中。

图表：Transformer整体结构



图表：百度ERNIE2.0预训练语言模型结构





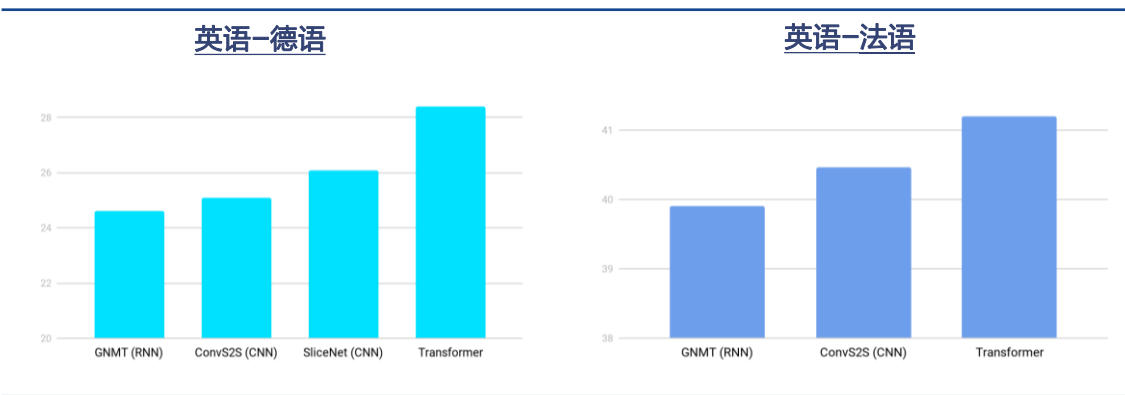
# 2.1 文本生成：Transformer凭自注意力机制取代RNN

- RNN的局限：**RNN模型（LSTM等）的限制在于相关算法只能从左向右或从右向左依次计算，带来了两个问题：（1）时刻t的计算依赖时刻t-1的计算结果，限制了模型的并行能力；（2）顺序计算的过程中信息会丢失，尽管LSTM等模型结构一定程度上缓解了长期依赖的问题，但是对于特别长期的依赖现象，LSTM依旧无能为力。
- Transformer的自注意力机制解决了RNN的两个问题：**首先，自注意力机制可按输入数据各部分重要性不同分配不同权重；其次，自注意力机制可以为输入序列中的任意位置提供上下文，支持并行训练。并行优势允许其在更大数据集上训练，促进了BERT、GPT等预训练大模型的发展。

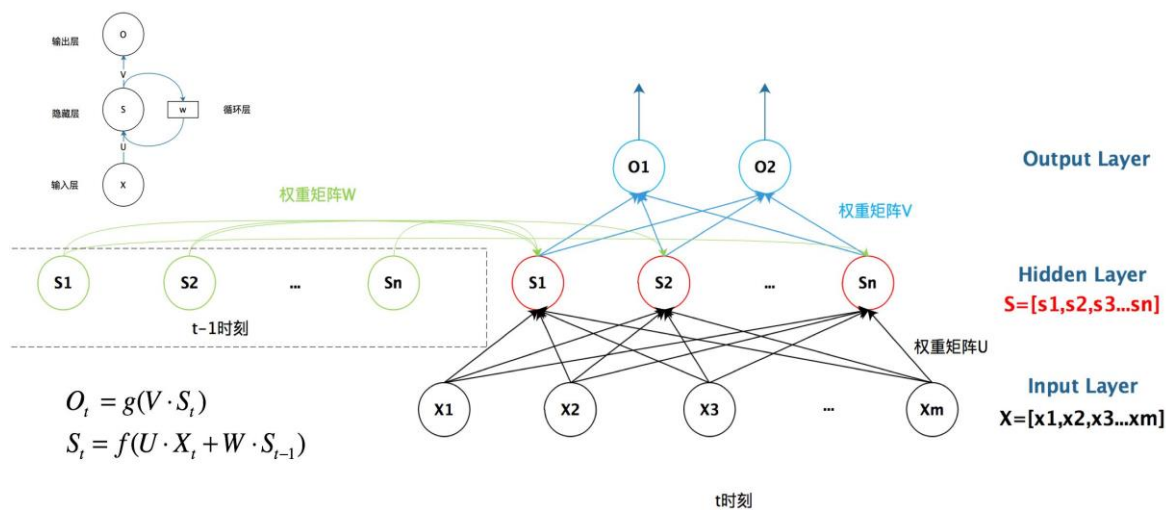
图表：RNN和Transformer处理上的区别

示例：“I arrived at the bank after crossing the.....” A.road or <span style="color:red">✓</span> B.river	
<b>RNN</b>	<b>Transformer</b>
方法：一步一步地阅读“bank”和“river”之间的每个单词，来确定“bank”的意思	方法：比较“bank”和其他单词，得到每个其他单词的注意力分数，获得较高注意力分数的单词“river”可能就是确定的结果

图表：Transformer在英语-德语、英语-法语翻译精度上明显优于RNN



图表：RNN结构



输入层为  $X$ ，隐藏层为  $S$ ，输出层为  $O$ 。 $U$  是输入层到隐藏层的权重， $V$  是隐藏层到输出层的权重。从公式看出，隐藏层的值  $S$  不仅取决于当前时刻的输入  $X$ ，还取决于上一时刻的输入  $X_{t-1}$ 。

## 2.1 文本生成：已实现大范围的商业落地

- 文本生成是AIGC实现商业落地最早的技术之一，技术发展显著提高了对于上下文的理解与承接能力、对常识性知识的嵌入能力、中长篇幅生成能力、生成内容的内在逻辑性等，文本生成迎来质的飞跃。现有的落地场景主要集中在应用型文本生成、创作型文本生成，重点关注闲聊型交互文本生成。

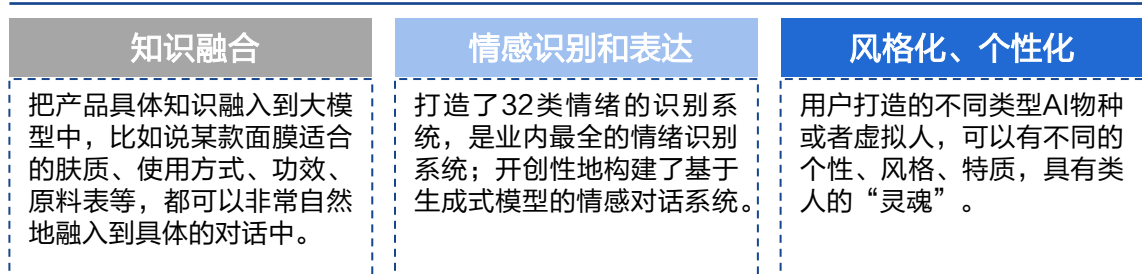
图表：文本生成的技术场景

技术场景	分类	细分场景	特点	发展现状	未来展望	国内外代表性公司
应用型文本生成	非交互型，大多为结构化写作，有较强规律	公司财报撰写、新闻简讯撰写等	基于结构化数据或规范格式，在特定情景类型下的文本生成	较广泛地实际应用，取得一定效果，技术还需进一步提升	Narrative Science首席技术官、联合创始人Kristian Hammond预言2030年90%以上新闻由机器人完成	澜舟科技、美联社Wordsmith、Narrative Science、textengine.io、AX Semantics、Yseop、Arria、Retresco、Viable等
创作型文本生成	非交互型，大多为非结构化写作，需要一定的创意	剧情续写、营销文本撰写等	具有更高的文本开放度和自由度，需要一定的创意和个性化，对生成能力的技术要求更高	底层技术原理基本明确，长篇幅文字内部逻辑仍然存在较明显问题、生成稳定性不足，有待技术完善	未来4-5年长文本创作可能会有突破发展	彩云小梦、Anyword、Phrasee、Persado、Pencil、Copy.ai、Friday.ai、Retresco、Writesonic、Conversion.ai、Snazzy AI、Rasa.io、LongShot.AI等
闲聊交互型文本生成	交互型，要求更高理解力	虚拟伴侣、游戏中的NPC个性化交互、心理咨询等	社交属性，对自然语言理解要求更高	有一些落地应用取得不错成果	1-2年内将有明显增长	小冰公司（小冰岛）、聆心智能、OpenAI、Latitude.io等

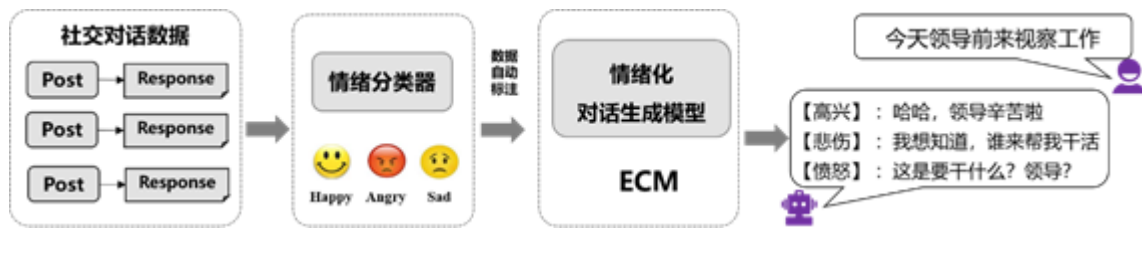
# 2.1 文本生成：共情聊天机器人Emohaa

- **专注AI驱动的心理疗法**：聆心智能由国内NLP、对话系统领域专家黄民烈教授于2021年创办。2022年，完成数千万元天使+轮融资，由连星资本领投。在技术上，聆心智能让AI围绕认知、情绪和行为三个维度对用户进行评估和干预，以多模态对话系统为核心交互框架，通过丰富的策略设计，让AI和用户产生足够的情感链接，生成用户个人模型及千人千面的治疗方案。
- **100亿参数大模型**。聆心智能的预训练模型参数达100亿左右，位于行业前列，数据源包括公开社交媒体、专业心理健康数据及实验室在过去积累的数据。
- ✓ **清华大学对话交互智能小组(CoAI)和聆心智能共同开发了情感对话技术，提出共情机器人三阶段理论模型**。第一阶段先确认用户的具体问题，第二阶段通过共情、理解表达支持，第三个阶段为用户提供解决方案、出路。每一阶段都有相应的策略生成，如提供信息、直接指导、挑战、解释等。
- ✓ **Emohaa的陪伴“套路”：确认、共情、提供建议**。2021年，聆心智能开发出中文对话预训练大模型和共情聊天机器人Emohaa，具备情绪识别、策略应对、共情表达三大特色，能够完成复杂的情感交流任务，也是目前唯一全AI模式的情绪支持对话系统。

图表：聆心智能三大技术优势



图表：情绪化聊天机器人（ECM）系统



图表：受到疫情困扰的用户与Emohaa的对话

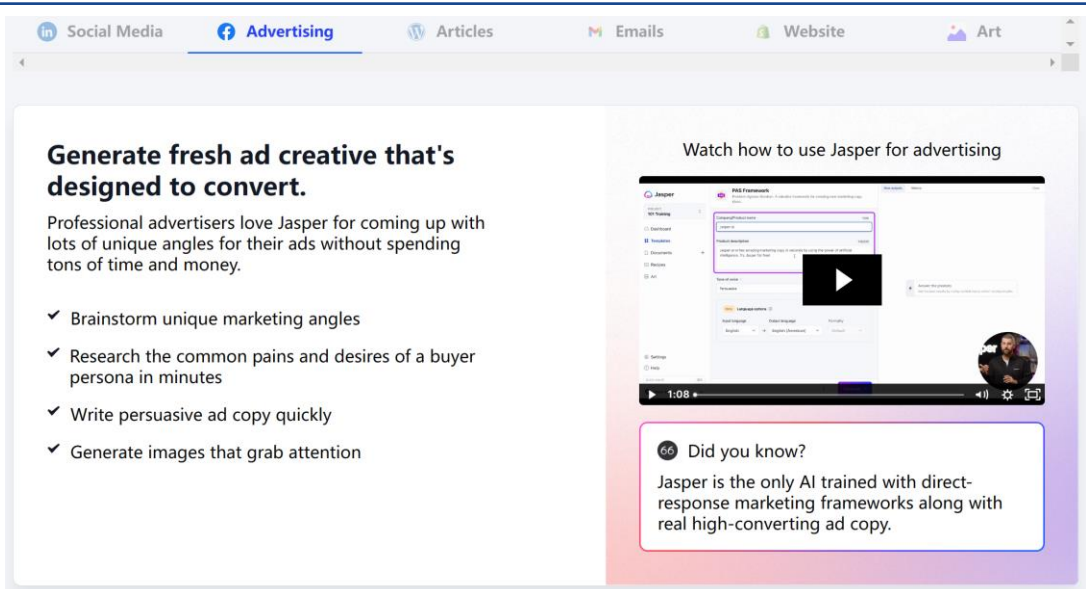




# 2.1 文本生成：Jasper.ai已实现商业变现，估值达15亿美元

- **估值15亿美元，主打文字生成的Jasper.ai成为独角兽。** Jasper.ai成立于2021年1月，以“AI文字生成”为主打产品，通过其文字生成功能，用户可以轻松生成Instagram标题，编写TikTok视频脚本、广告营销文本、电子邮件内容等工作。Jasper.ai于2022年10月宣布完成了1.25亿美元的A轮融资，估值达到了15亿美元。Jasper.ai也已推出了图像生成产品Jasper Art。
- **实现商业变现，2022年收入翻倍。** Jasper.ai主要是通过SaaS付费的模式，目前分为BOSS版本、商业版本进行不同定价。
  - ✓ 公司2021年已经有70000名付费客户。公司在2021年创造了4000万美元的收入，预计到2022年底，收入将达到9000万美元，同比增长了一倍以上。
  - ✓ Jasper.ai能够实现减少花在第一稿上的80%的时间，输出高出10倍质量的产品，实现400%的投资回报率。

图表：Jasper.AI官网示例



图表：Jasper.AI产品定价

定价类型	BOSS版本	商业版本
适用人群	个人、小团队	大团队、商业团队
内容	利用AI写作的全部功能，以完整性和灵活性创建全文内容	使用Jasper进行扩展所需的入门和支持等
价格	99美元/月（月付），82美元/月（年付）	499美元/月起，定制价格

Boss Mode	Business
For individuals and small teams	For growing teams and businesses
Harness the full power of AI writing with the ability to create full length content with full control & flexibility.	Get custom words/user packages and billing options, and the onboarding and support you need to scale with Jasper.
<p>\$82/mo estimated</p> <p>50K <span style="display: inline-block; width: 100px; border-bottom: 1px solid black; position: relative; top: -5px;"> <span style="position: absolute; left: -10px; top: -5px;">100,000 words/mo</span> </span> 700K+</p> <p>Start Free</p> <p>5-day free trial with 10k credits.</p>	<p>Custom Plan &amp; Price</p> <p>Starting at \$499/mo</p> <p>Contact Sales</p> <p>Get a personalized demo &amp; custom plan</p>
<p>AI features individuals want:</p> <ul style="list-style-type: none"> <li>50+ AI templates</li> <li>Google Docs style editor</li> <li>Jasper Chat</li> <li>Compose &amp; command features</li> <li>Live chat support</li> </ul>	<p>AI features teams want:</p> <ul style="list-style-type: none"> <li>Collaborate with more than 5 users</li> <li>Tailored AI Brand Voice</li> <li>API access (Coming soon)</li> <li>Personalized onboarding &amp; training</li> <li>Dedicated account manager</li> </ul>



# 【专题1-ChatGPT】“大数据+大模型+大算法”成就AIGC发展里程碑

- **能力提升来源于大算法和训练路径改善。** ChatGPT是微调后的GPT-3.5系列模型（“预训练-微调”），有1750亿个模型参数，2022年初训练完成。ChatGPT还使用了“利用人类反馈强化学习”（RLHF: Reinforcement Learning from Human Feedback）的机器学习技术进行训练，可以模拟对话，回答后续问题，承认错误，质疑不正确的前提，拒绝不恰当的请求。
- **瑕不掩瑜，有望迭代发展。** ChatGPT仍存在一些局限性：1）有时会写出看似合理但错误或荒谬的回答。2）调整问题措辞，可能会获得不同的答案。3）回复过于冗长。这些问题源于训练数据的偏差，因为训练师（标注人员）更喜欢看起来更全面的更长的答案。4）无法完全拒绝不合理及不道德的请求。5）不具备网络搜索功能，只能基于2021年所拥有的数据集进行回答。

图表：GPT模型迭代对比

模型	发布时间	参数	解码器层	隐藏层	批量处理大小	训练数据数量	特点
GPT-1	2018年	1.17亿	12	768	64	约5GB	无监督训练+有监督微调
GPT-2	2019年	150亿	48	1600	512	40GB	多任务学习
GPT-3	2020年	1750亿	96	12288	3.2M	45TB	海量参数，任务难度加大

图表：GPT-3和GPT-3.5系列模型的区别

模型	能力	Open AI模型	训练方法
GPT-3	语言生成 +世界知识 +上下文学习	GPT-3初始版本 (大部分能力已存在于模型中，表面上看起来比较弱)	语言建模
	+遵循人类的指令 +泛化到没有见过的任务	Instruct-GPT初始版本	指令微调
	+代码理解 +代码生成	Codex初始版本	在代码上进行训练
GPT-3.5	++代码理解 ++代码生成 ++复杂推理/思维链 +长距离的依赖	现在的Codex (GPT3.5系列中最强大的模型)	在代码+文本上进行训练 在指令上进行微调
	++遵循人类指令 -上下文学习 -推理能力 ++零样本生成	有监督的Instruct-GPT (通过牺牲上下文学习换取零样本生成的能力)	监督学习版的指令微调
	+遵循人类价值观 +包含更多细节的生成 +上下文学习 +零样本生成	经过RLHF训练的Instruct-GPT (和有监督的Instruct-GPT相比，和人类更加对齐，并且更少的性能损失)	强化学习版的指令微调
	++遵循人类价值观 ++包含更多细节的生成 ++拒绝知识范围外的问题 ++建模对话历史的能力 --上下文学习	ChatGPT (通过牺牲上下文学习的能力换取建模对话历史的能力)	使用对话数据进行强化学习指令微调

注：+表示能力的增强，-表示能力的弱化，符号个数表示增强/弱化程度

# 【专题1-ChatGPT】利用RLHF深度学习训练，迭代出更高质量模型

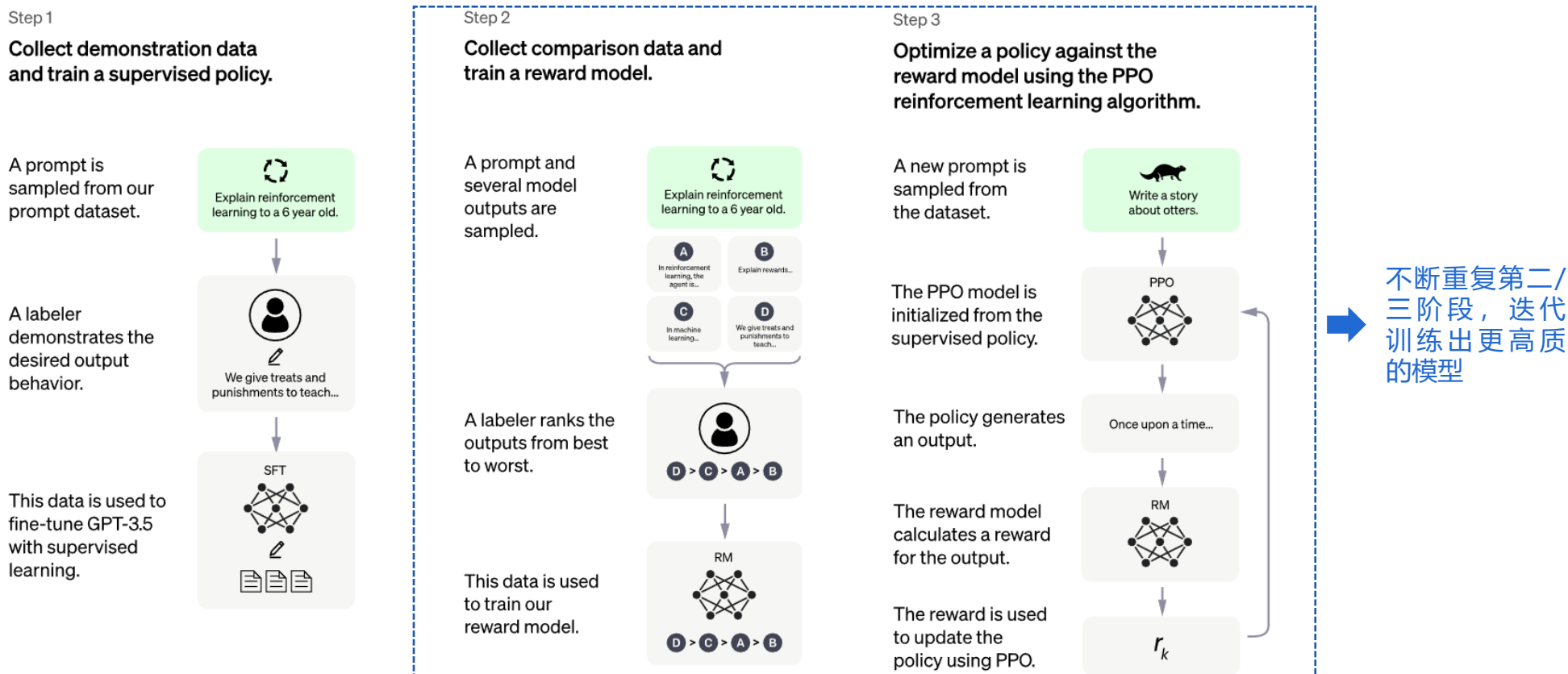
■ ChatGPT是微调后的GPT-3.5系列模型，使用“利用人类反馈强化学习”（RLHF）的机器学习技术进行训练：

1) 训练监督策略模型：为了让GPT3.5初步具备理解指令的意图，首先会在数据集中随机抽取问题，由人类标注好并给出高质量答案，用这些数据来微调GPT-3.5模型。

2) 训练奖励模型（Reward Model）：在数据集中随机抽取问题，使用第一阶段生成的模型，对每个问题生成多个不同的回答。标注者对这些结果综合考虑给出排序，使用排序结果来训练奖励模型，对多个排序结果，两两组合，形成多个训练数据对，RM模型接受一个输入，给出评价回答质量的分数。

3) 采用近端策略优化（Proximal Policy Optimization）强化学习：利用第二阶段训练好的奖励模型，靠奖励打分来更新预训练模型参数。在数据集中随机抽取问题，使用PPO模型生成回答，并用上一阶段训练好的RM模型给出质量分数，把回报分数依次传递，由此产生策略梯度，通过强化学习的方式更新PPO模型参数。

图表：ChatGPT的训练过程



## ChatGPT做物理/化学题

New York State Aug 2022 Chemistry regent, 35/45 (77.7%) on MC portion (excl 5 questions that depend on photos) [nysedregents.org/Chemistry/](https://nysedregents.org/Chemistry/)

**K** Which substance always forms when an Arrhenius acid reacts with an Arrhenius base?  
(1) CO<sub>2</sub>  
(2) H<sub>2</sub>  
(3) CH<sub>3</sub>OH  
(4) H<sub>2</sub>O

**GPT** When an Arrhenius acid reacts with an Arrhenius base, the substance that is formed is water (H<sub>2</sub>O). This is because Arrhenius acids are defined as substances that produce hydrogen ions (H<sup>+</sup>) when they are dissolved in water, while Arrhenius bases are defined as substances that produce hydroxide ions (OH<sup>-</sup>) when they are dissolved in water. When these two ions combine, they form water. Therefore, the correct answer is H<sub>2</sub>O.

**K** To the previous question, the correct choice is, without explanation:

**GPT** The correct choice is H<sub>2</sub>O.

量子位

## ChatGPT写小说

New chat

Cozy Mystery Murder

Create Mystery Suspects

Clear conversations

Dark mode

OpenAI Discord

Updates & FAQ

Log out

I am writing a paranormal cozy mystery that takes place in the small Texas town of Tablerock with an female amateur sleuth named Ellie Rockwell. I need a murder victim and how they were killed, four murder suspects with information about why they're suspected and how they are cleared, and who is the guilty killer is (and how and why they did it).

**GPT** Sure, here's a murder victim and suspects for your paranormal cozy mystery:

Victim:

- Name: Bob Jones
- Age: 45
- Occupation: Owner of a local hardware store
- How they were killed: Bob was found strangled in his office at the hardware store, with no signs of a break-in or struggle.

Suspects:

- Sue Smith:
  - Age: 35
  - Occupation: Bob's ex-girlfriend and former employee at the hardware store
  - Motive: Sue and Bob had a tummy Regenerate response broke up a few months prior to the murder. She was also fired from her job at the hardware store shortly before the

量子位

ChatGPT Dec 15 Version. Free Research Preview. Our goal is to make AI systems more natural and safe to interact with. Your feedback will help us improve.

## ChatGPT写代码

生成一个「快速排序算法」的python函数

这是一个快速排序算法的python代码:

```
python
def quick_sort(arr):
    if len(arr) <= 1:
        return arr
    else:
        pivot = arr[len(arr) // 2]
        left = [x for x in arr if x < pivot]
        middle = [x for x in arr if x == pivot]
        right = [x for x in arr if x > pivot]
        return quick_sort(left) + middle + quick_sort(right)
```

量子位

## ChatGPT写短文、诗

SH+ 如果李白2022年来到中国，会做一首什么样的诗？

**GPT** 如果李白2022年来到中国，会做一首诗，可能是这样的：

蒙山绿水清，青山绿水明。  
中国人民欢迎，来访客人来。

天地之间秀丽，千山万水清。  
生态文明建设好，绿色发展更轻松。

青山绿水美，生态环境好。  
让我们一起去奋斗，实现社会主义现代化好。

帮我以鲁迅的文笔写一段话，表达以下我现在因为疫情连门都不敢出的惨状，还有想吃火锅的心情

在疫情的蔓延中，我困了门里，困了心里。屋外一片悲凉，想起一顿火锅的香味，瞬间让我心慌。这时候我才知道，疫情来临，它让人生无聊而悲凉，也让我们看到自己的无助与弱点。然而我们又岂能放弃对美好的追求，岂能把欲望压抑到不想吃火锅的地步？

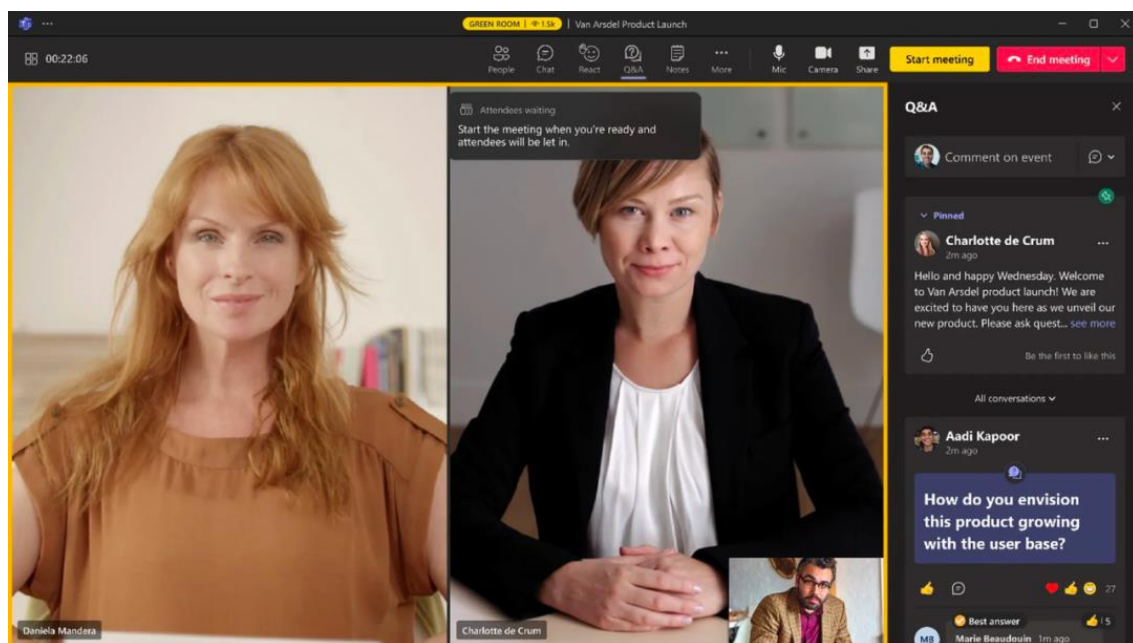
量子位



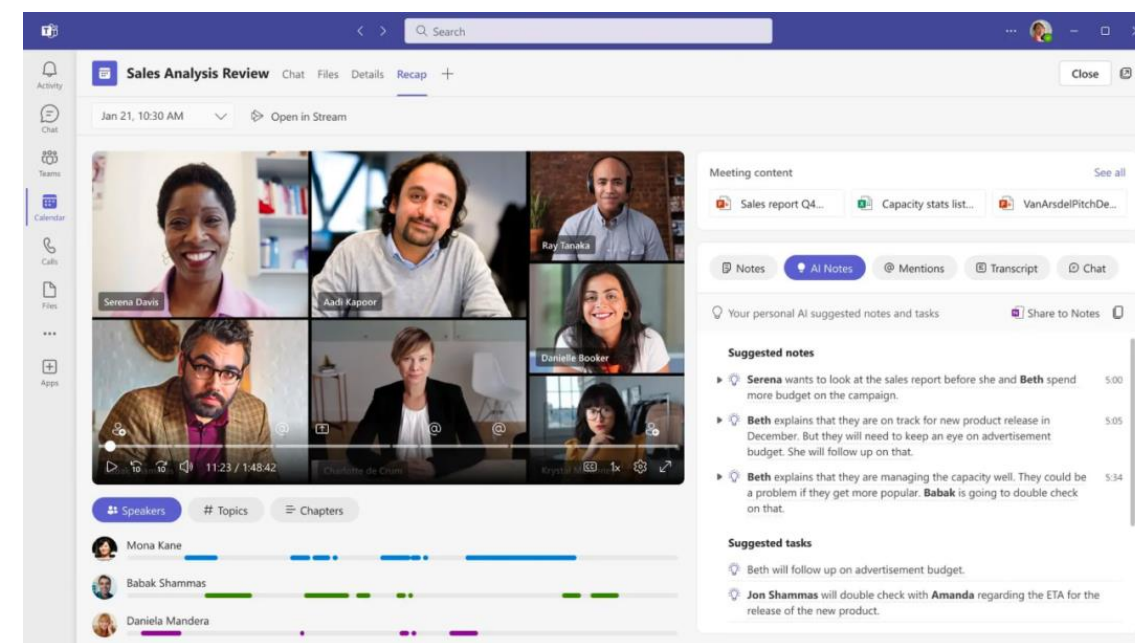
# 【专题1-ChatGPT】引爆全网，用户量直冲数亿级，商业价值显现

- ChatGPT上线一周用户量达百万级，商业变现初显。2022年11月，OpenAI上线了聊天机器人ChatGPT，能够理解自然的人类语言，并生成类似人类作答的书面文本，能写代码、改Bug、创建编程语言等。2022年11月30日-12月5日，ChatGPT用户数量已破100万；2023年1月末，推出两个月后月活已突破了1亿。2023年2月1日，OpenAI宣布推出月费20美元的「ChatGPT Plus」订阅服务。
- ChatGPT等AI工具商业化开始落地：微软CEO纳德拉曾表示，微软正迅速推进OpenAI的工具商业化，将ChatGPT、DALL-E等AI工具整合进微软旗下所有产品中，包括且不限于Bing搜索引擎、Office全家桶、Azure云服务等等。
  - ✓ 2月推出OpenAI模型支持的最新版本Bing搜索引擎与Edge浏览器，增加可以聊天与写作的功能。
  - ✓ 2月8日，推出由ChatGPT提供技术支持的高级Teams产品，可自动生成会议记录、推荐任务或创建会议模板，这项服务6月起每月收费7美元，7月增加到10美元。
  - ✓ 推出了Viva Sales基于生成式AI驱动经验的预览功能，帮助销售人员更有效地与潜在客户和消费者沟通，可以在各种场景生成推荐的电子邮件内容，比如回复询问或创建提案，比如提供与收件人相关的数据。

图表：Teams Premium的个性化的时间线标记功能



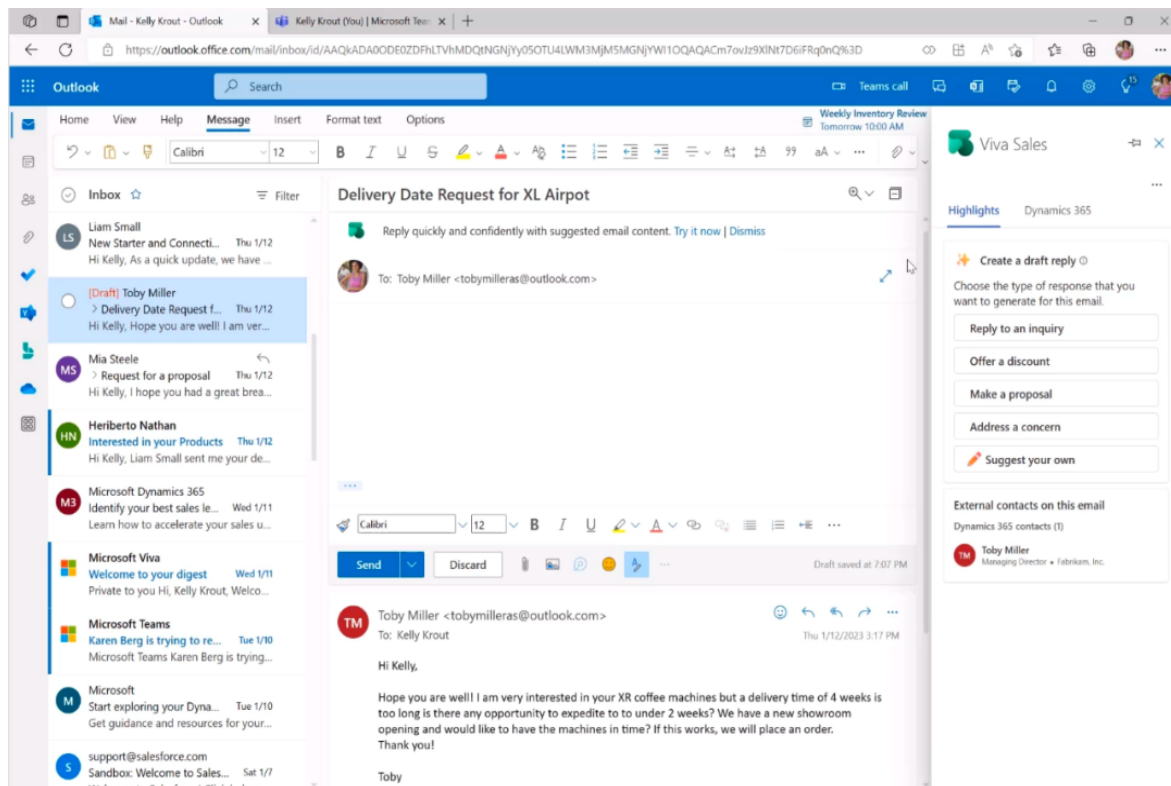
图表：Teams Premium的智能回顾功能



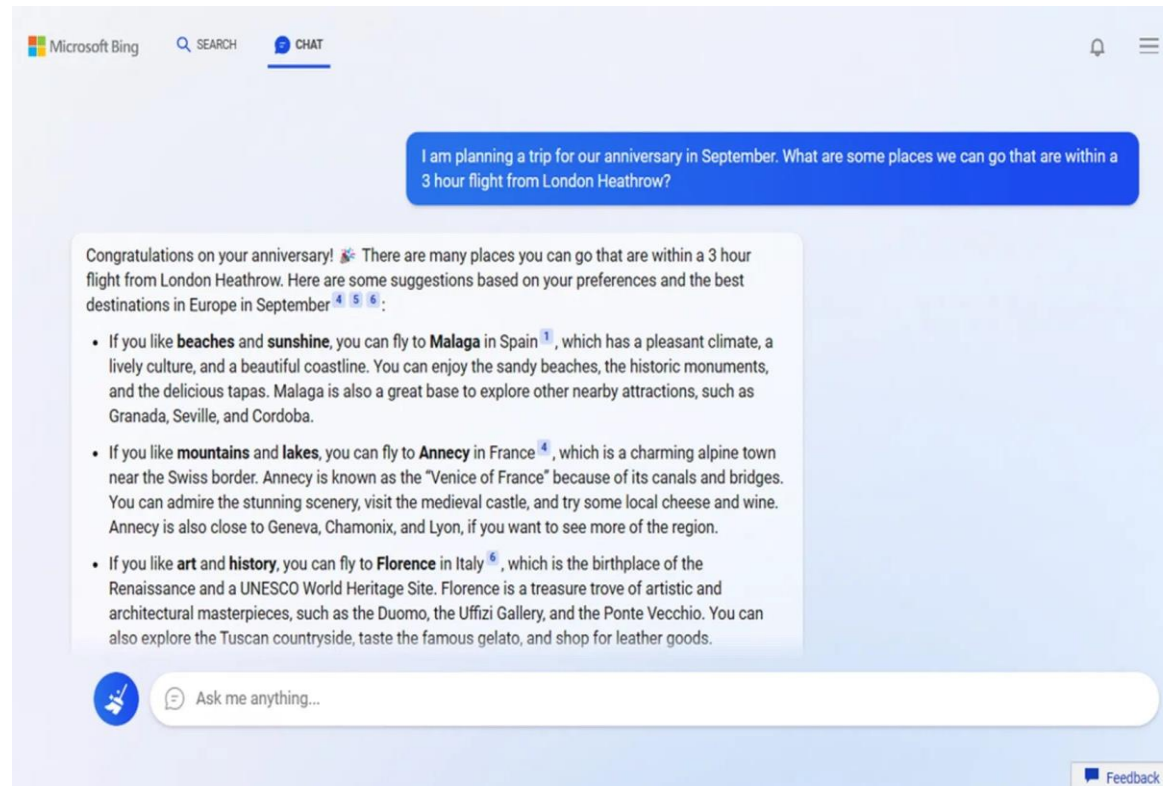


# 【专题1-ChatGPT】与微软旗下产品结合示例

在回复电子邮件时，Viva Sales 根据「提出建议」、「回复询问」或「自身建议」等类别，为卖家提供即时的建议回复。使用者只需选择符合需要的选项，系统就会生成一个回复供其审阅，再根据自己的喜好进行编辑，然后发送。



用户可以直接与Bing聊天机器人对话，在类似ChatGPT的聊天界面中向其提问，比如问旅行计划安排，Bing描述了一个粗略的行程，并提供了更多信息的链接。



## 2.2 音频生成：TTS场景基本成熟，乐曲创作有发展潜力

- 音频生成主要应用于流行歌曲、乐曲、有声书的内容创作，以及视频、游戏、影视等领域的配乐创作，目前在众多场景已获初步发展，在部分场景已广泛应用、趋于成熟。现有的落地场景集中在TTS、语音克隆、乐曲/歌曲生成。

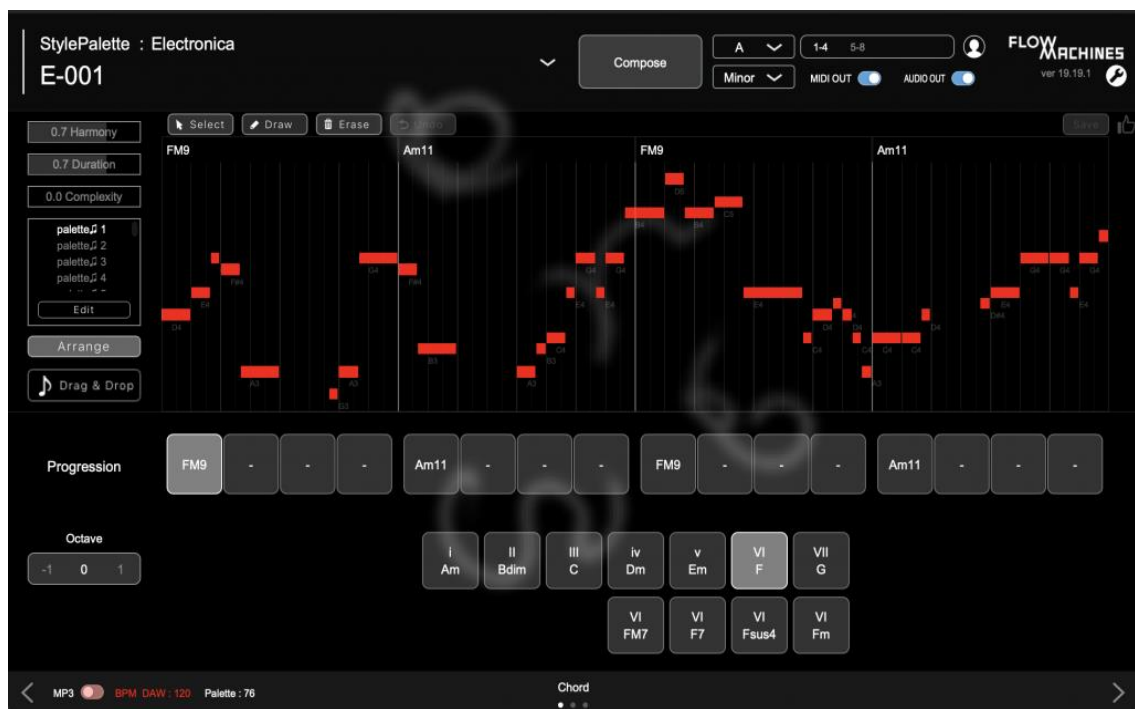
图表：音频生成的落地场景

落地场景	细分场景	技术/内容	发展现状及未来展望	国内外代表性公司（或其产品）示例
TTS (Text-to-speech)	智能客服、有声读物制作、语音播报、自媒体配音、导航播报等	基于深度学习的端到端语音合成模式正在逐步替代传统的拼接及参数法，代表模型包括WaveNet、Deep Voice及Tacotron等	<ul style="list-style-type: none"> <li>在AIGC领域已基本成熟，广泛应用于各种场景；</li> <li>技术上的关键突破在于如何通过富文本信息更好的表现其中的抑扬顿挫，以及基于用户较少的个性化数据得到整体的复制能力</li> </ul>	喜马拉雅、倒映有声、科大讯飞思必驰（DUI）、Readspeaker、DeepZen（IBM旗下）和Sonantic、九锤配音、加音、XAudioPro、剪映等 倒映有声与音频客户端“云听”APP合作打造AI新闻主播，提供音频内容服务的一站式解决方案 喜马拉雅运用TTS技术重现单田芳声音版《毛氏三兄弟》和历史类作品
语音克隆（本质是指定了目标语言的TTS）	虚拟歌手演唱、自动配音等	<p>Figure 1: Model overview. Each of the three components are trained independently.</p>	在声音IP化的基础上，对于动画、电影以及虚拟人行业有重要意义	标贝科技、Modulate、overdub、replika、Replica Studios、Lovo、Voice mod、Resemble Ai、Respeecher、DeepZen、Sonantic、VoicelD、Descript等 标贝科技推出恐龙贝克APP，采用AI语音合成技术复刻爸妈的声音为0-6岁儿童讲故事
乐曲/歌曲生成	AI作曲	以语言模型为中介，对音乐数据进行双向转化；代表模型：MelodyRNN、Music Transformer	<ul style="list-style-type: none"> <li>AI作曲已支持基于开头旋律、图片、文字描述、音乐类型、情绪类型等生成特定乐曲</li> <li>自动编曲功能已在国内主流音乐平台上线</li> <li>难点在于音乐数据的标注。</li> </ul>	灵动音科技、网易-有灵智能创作平台、昆仑万维旗下StarX平台、Amper Music、IBM Watson Music、Magenta、Loudly、Brain.FM、Splash、索尼旗下Flow machines等 2021年末，贝多芬管弦乐团在波恩首演人工智能谱写完成的贝多芬未完成之作《第十交响曲》 QQ音乐成为Amper music的API合作伙伴 上海禾念推出虚拟歌手洛天依，举办线上AR演唱会
	AI编曲	AI基于主旋律和创作者个人偏好，生成不同乐器的对应和弦（如鼓点、贝斯、钢琴等），完成整体编配		
	人声录制	通过端到端的声学模型和神经声码器完成，可以简单理解为将输入文本替换为输入MIDI数据的声音克隆技术。		

## 2.2 音频生成：TTS场景基本成熟，乐曲创作有发展潜力

- 索尼计算机科学实验室（CSL）在2021年9月宣布发布一款AI辅助音乐制作应用程序Flow Machines Mobile（FM Mobile），已登陆苹果App Store，可以用于iPad。FM Mobile利用AI人工智能技术辅助音乐制作，能够根据创作者选择的风格提示旋律、和弦和贝斯线，生成音乐。
- 2021年，喜马拉雅用TTS技术完美还原单田芳声音，并首次将单田芳的AI合成音应用于风格各异的书籍，用单式评书腔调，全新演绎听众耳熟能详的经典之作。喜马拉雅上已有80多部运用还原自单田芳先生声音制作的专辑。

图表：FM Mobile



图表：喜马拉雅单田芳声音重现



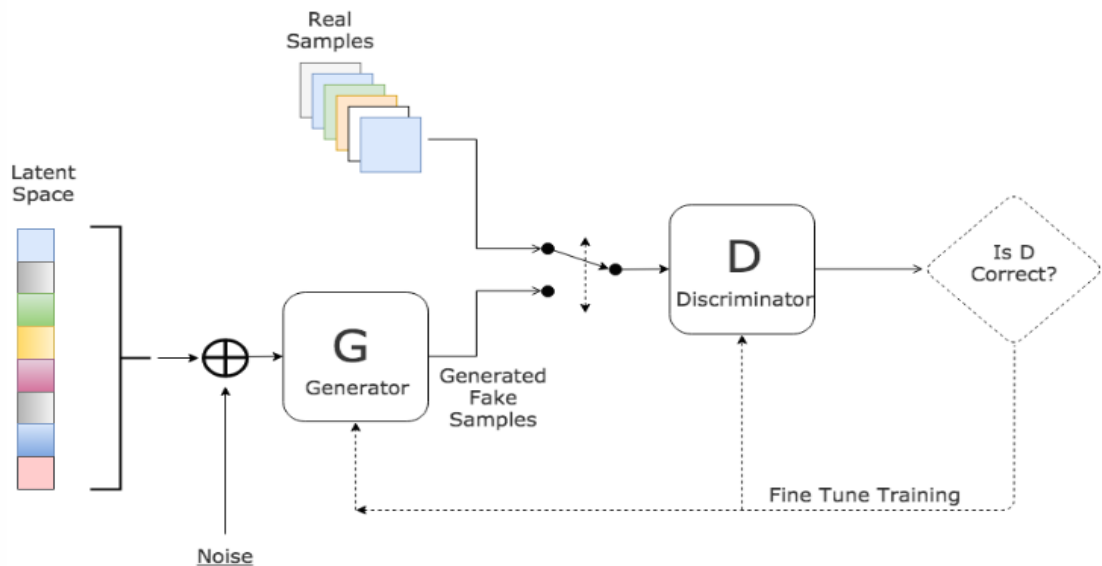


## 2.3 图像生成：从GAN到扩散模型，图像生成技术迭代发展

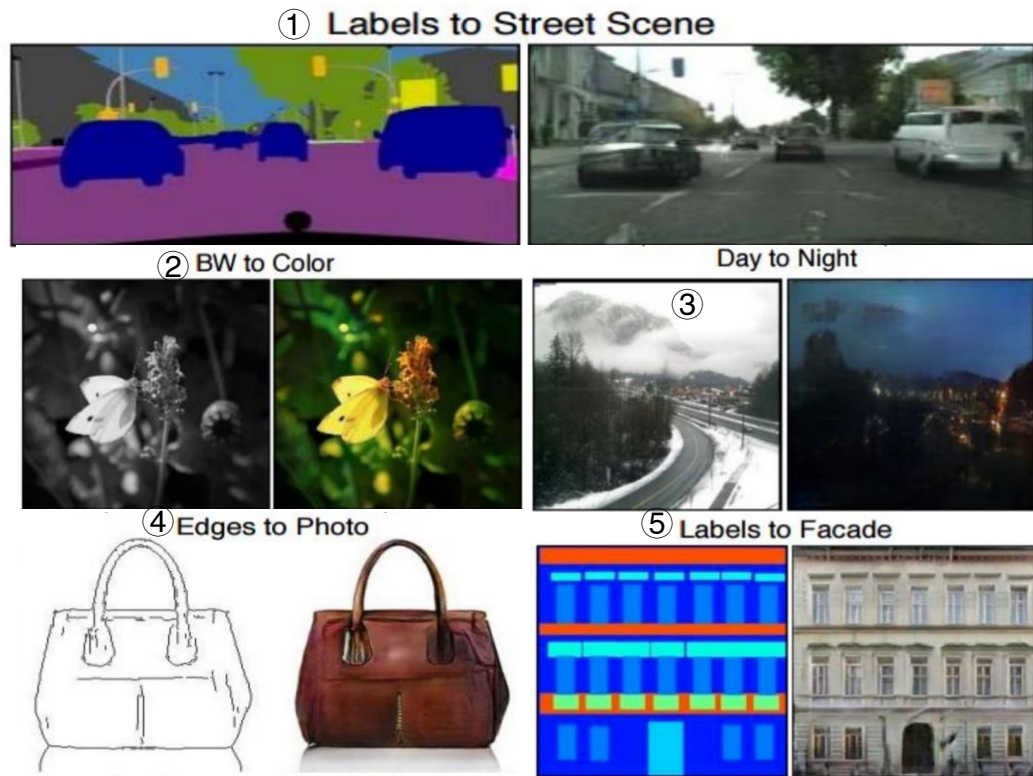
- 2014年生成式对抗网络（Generative Adversarial Nets, GAN）的提出标志着图像生成进入快速发展期。GAN由生成器和判别器两个神经网络组成，其中生成器通过输入数据试图产生欺骗判别器的真实样本，而判别器试图区分真实样本和生成样本。对抗博弈下不断提高性能，达到纳什平衡后生成器可以实现以假乱真的输出。
- GAN不需要蒙特卡洛估计来训练网络，G的调整更新不直接来自数据样本，而使用来自D的反向传播，能够更快产生样本。但可解释性差且较难训练，容易产生随机图像，且图像同质化严重、分辨率较低。针对这些，CGAN、StyleGAN被提出来解决难训练和分辨率低的问题。
- 虽然近年来GAN在神经网络架构、损失函数设计、模型训练稳定性、模型崩溃问题上取得了突破，提升了最终图像的特定细节、内在逻辑、生成速度等，但要在实际中大规模运用还需要解决相应问题。

图表：GAN基本结构

- G是生成图片的网络，它接收一个随机的噪声z，通过这个噪声生成图片，记做G(z)。
- D是判别图片的网络，判别一张图片是不是“真实的”。它的输入参数是x（代表一张图片），输出D(x)代表x为真实图片的概率，如果为1，就代表是真实的图片，而输出为0，就代表不是真实的图片。



图表：基于GAN的图像翻译（左图为input，右图为output）

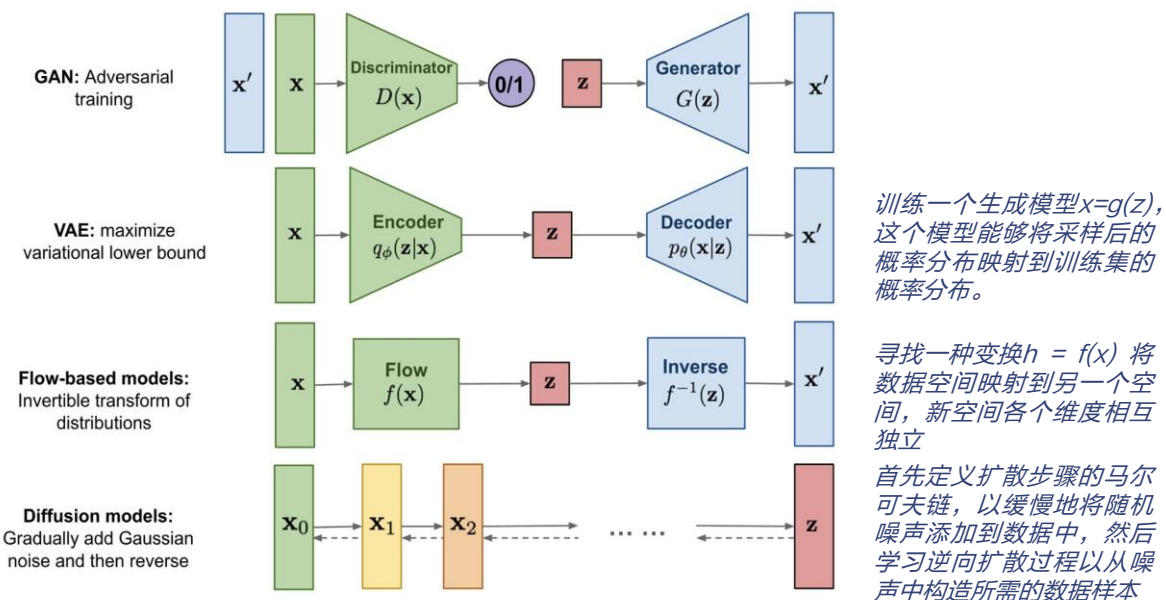




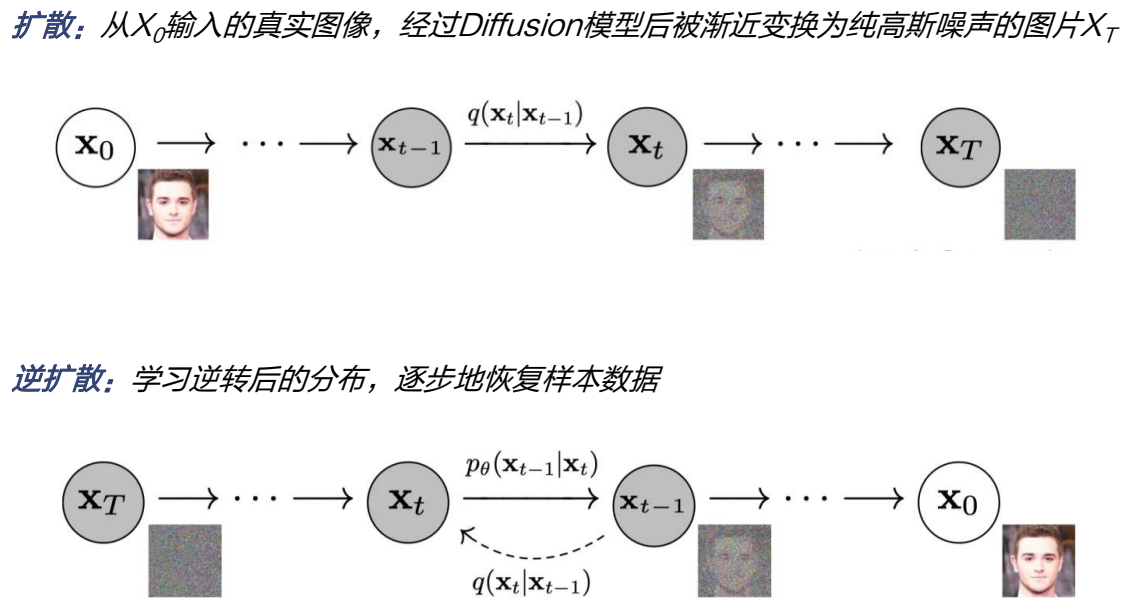
# 2.3 图像生成：从GAN到扩散模型，图像生成技术迭代发展

- 扩散模型 (Diffusion Model) 解决了GAN训练不稳定的问题，成为新一代图像生成主流模型。扩散模型的概念最早在2015年被提出，2020年，去噪扩散概率模型 (Denoising Diffusion Probabilistic Model, DDPM) 提出，向世界展示了扩散模型的实力，之后图像生成领域开始转向DDPM的研究。2022年，扩散模型成为图像生成领域的重要发现，有超越GAN的势头。Stable diffusion模型在2022年8月开源，此后图像生成应用遍地开花。
- ✓ Diffusion模型的工作原理是通过连续添加高斯噪声来破坏训练数据，然后通过反转这个噪声过程，来学习恢复数据。扩散模型提供了一个系统地增加噪音的过程，通过扩散模型向生成的数据和真实数据添加噪音，然后将加入噪音的数据送入判别器，这样可以高效地解决GAN无法训练、训练不稳定的问题。
- ✓ Diffusion模型由正向过程 (扩散) 和反向过程 (逆扩散) 组成，其中输入数据逐渐被噪声化，然后噪声被转换回源目标分布的样本。
- ✓ 相较于其他的图像生成模型 (比如GAN、VAE和Flow-based模型)，在所需数据更少的背景下，Diffusion模型的图像生成效果有明显提升。

图表：四种图像生成模型结构对比



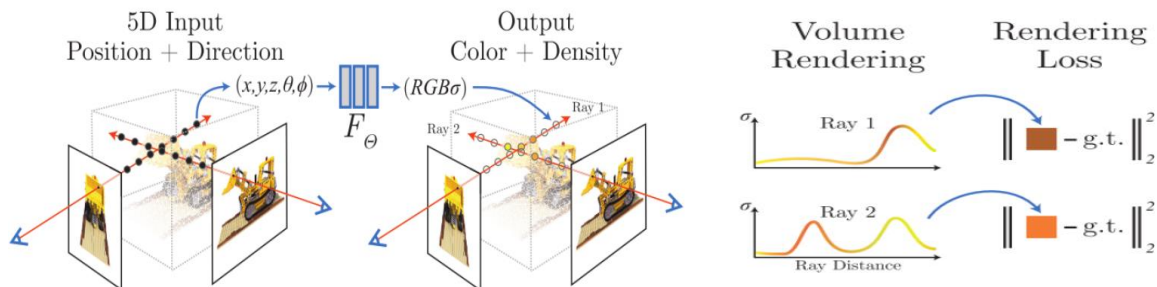
图表：扩散模型的正向过程和反向过程



## 2.3 图像生成：NeRF是3D建模的主流模型

- NeRF (Neural Radiance Field, 神经辐射场) 作为一种具有隐式场景表示的新型视场合成技术一经被提出便引起了广泛的关注。NeRF是一个完全连接的神经网络，可以基于部分2D图像集生成复杂3D场景的新视图。NeRF在机器人、城市地图、自主导航、虚拟现实/增强现实等领域都有广泛的应用。
- NeRF的优势是利用隐式表达实现照片集的视角合成效果，但还存在训练速度慢、渲染速度慢、只能用于静态场景、泛化性能差、需要大量视角、难以与传统渲染管线融合等问题，目前还没有比较成功的商业化应用。在技术的不断完善迭代中，NeRF有望打开3D建模的新市场。
- ✓ NeRF实现了从离散的照片集中学习出了一种隐式的Volume表达，然后在某个特定视角，利用该隐式Volume表达和体渲染得到该视角下的照片。隐式是连续的表达，能够适用于大分辨率的场景，而且不需要3D信号进行监督。
- ✓ NeRF自提出后也持续在生成范围、生成效果、所需基础数据上进行改进。例如包含光影变化效果 (NeRV)、动态NeRF (HyperNeRF)、接近实时生成 (Instant-NGP和Instant NeRF)、全场景NeRF (mip-NeRF 360)。

图表：NeRF的工作流程



从2D到3D

1、**三维重建**：输入三维位置( $x, y, z$ )和二维方位视角( $\theta, \phi$ )，利用多层感知机 (MLP) 网络建模输出颜色 $c=(r, g, b)$ 和体素密度( $\sigma$ )

从3D到2D

2、**渲染**：将3D点的颜色及体素密度沿着光线进行整合得到最终的2D图像像素值

3、**训练**：将得到的2D图像与Ground Truth做损失进行网络优化

图表：NeRF的具体技术场景示例



3D风格迁移 (StylizedNeRF)

提出2D卷积网络和3D辐射场网络互学习机制，可训练得到一个可以进行风格化的神经辐射场网络



人体重建 (HumanNeRF)

可将任何单视角视频中的人体运动渲染成支持自由视点查看的3D模型

## 2.3 图像生成：图像属性编辑已大规模落地

- 图像生成领域的技术场景可以划分为图像属性编辑、图像局部生成及更改以及端到端的图像生成。其中，前两者的落地场景为图像编辑工具，而端到端的图像生成则对应创意图像及功能性图像生成两大落地场景。
- 在图像编辑领域已获大规模应用落地，但由于图像的生成复杂度远高于文字，在整体生成上目前仍然难以达到稳定可靠的生成高质量图像。未来，随着底层技术原理的不断突破，有望实现规模化应用。

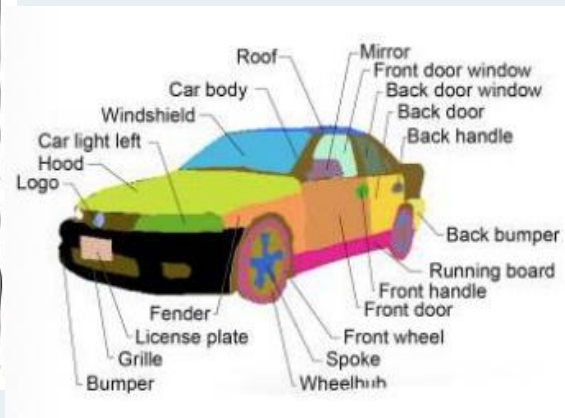
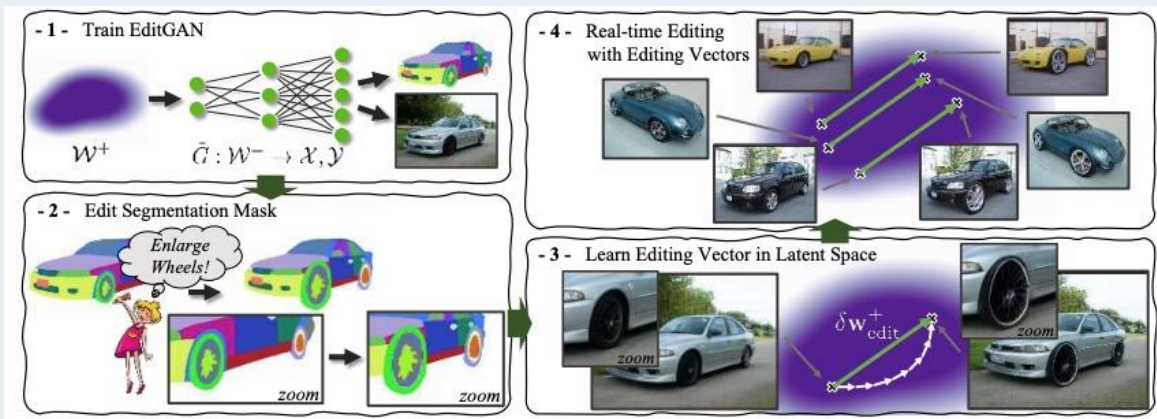
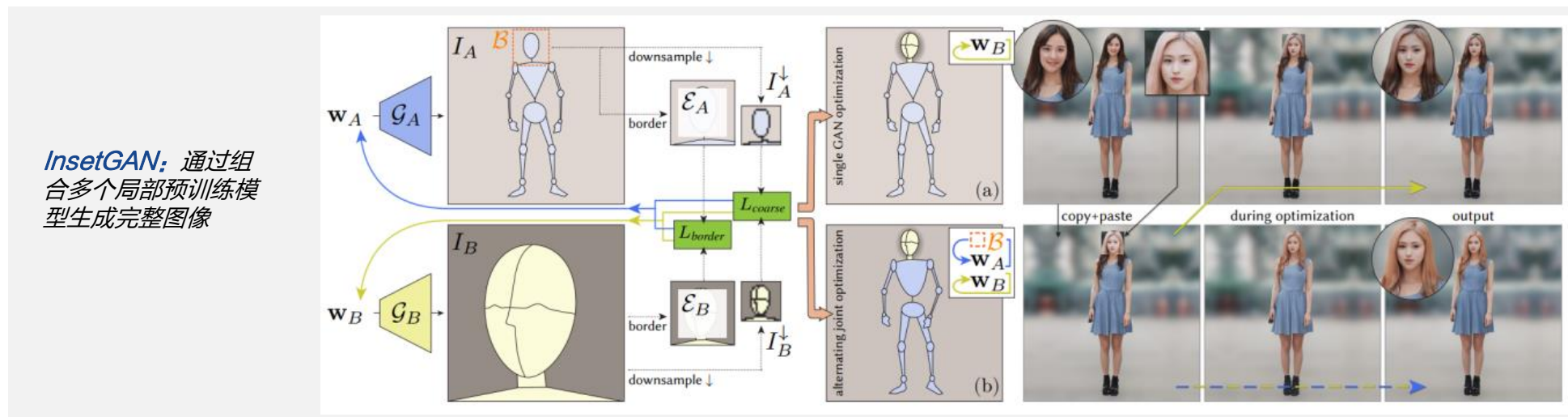
图表：图像生成的应用场景

技术场景	落地场景	内容	发展现状及未来展望	代表性公司
图像属性编辑	图像编辑工具	图片去水印、自动调整光影、设置滤镜、修改颜色纹理、复刻/修改图像风格、提升分辨率等	大量应用落地，初创公司众多；未来的突破点在于引流渠道和与技术流程挂钩的产品使用体验	美图秀秀（美图AI开放平台）、Radius5、Photokit、Imglarger、Hotpot、Remove.bg、Skylum（Mask AI）、Photodiva
图像局部生成及更改	图像编辑工具	部分更改图像部分构成（如英伟达CycleGAN支持将图内的斑马和马进行更改）、修改面部特征（Metaphysics，可调节自身照片的情绪、年龄、微笑等；以Deepfake为代表的图像换脸）	难以直接生成高质量的完整图像，但基于GAN的优化模型不断发展，目前也有部分应用落地。	Adobe（InsetGAN）、英伟达（EditGAN）
端到端的图像生成	<ul style="list-style-type: none"> <li>√创意图像生成（NFT）</li> <li>√功能性图像生成（海报、logo、模特图等）</li> </ul>	基于草图生成完整图像、有机组合多张图像生成新图像、根据指定属性生成目标图像等	底层技术原理基本明确，1-2年后规模化应用	阿里鹿班、Deepdream Generator、Rosebud.ai、AI Gahaku artbre eder、nightcafe、starry ai、wombo、deepart、obvious、ZMO.ai、Datagrid、诗云科技、道子智能绘画系统等



## 2.3 图像生成：图像局部变更模型

图表：图像局部变更的InsetGAN和EditGAN模型



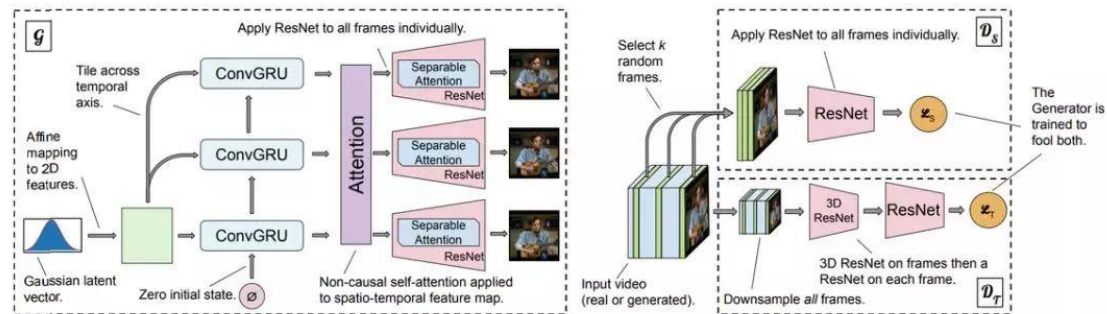
**EditGAN:** 将图像嵌入到GAN潜在空间中，并根据分割编辑执行潜在的代码优化从而高效地修改图像。



# 2.4 视频生成：与图像生成类似，大量应用集中在视频属性编辑

- 视频生成原理与图像类似，主流模型为GAN、VAE、Flow-based模型，近年来扩散模型也逐渐应用到视频生成中。
- ✓ 2019年，DeepMind提出了DVD-GAN模型（Dual Video Discriminator GAN），加入了空间判别器D-S和时间判别器D-T。GAN在视频生成中的重大突破是一组由50万段10秒高分辨率YouTube视频剪辑汇编成的数据集训练，能够生成具备时间一致性的高分辨率（256 x 256像素）视频。
- 视频生成应用场景主要集中在视频属性编辑、视频自动剪辑、视频部分编辑，前者已有大量应用落地，后两者还处于技术尝试阶段。视频编辑比图像更具挑战性，需要合成新动作，不仅是修改外观，还要保持时间一致性。

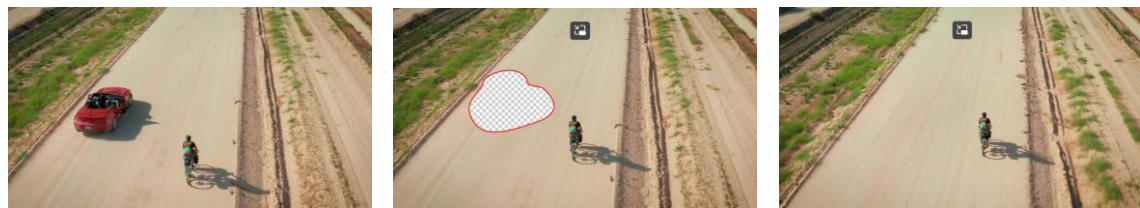
图表：DVD-GAN模型架构



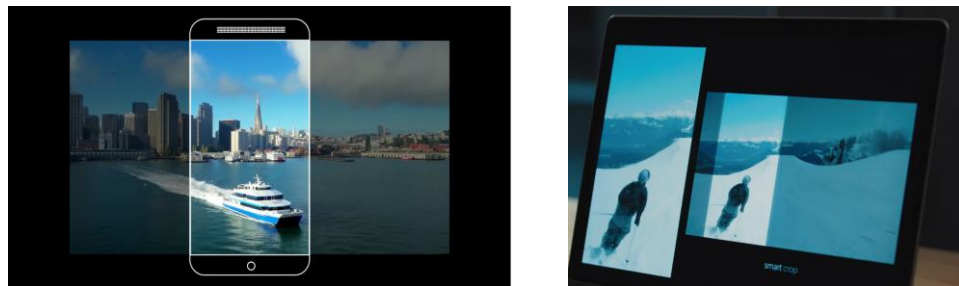
图表：视频生成的应用场景

技术场景	内容	发展现状	代表性公司
视频属性编辑	视频画质修复、删除画面中特定主体、自动跟踪主题剪辑、生成视频特效、自动添加特定内容、视频自动美颜等。	大量应用已落地	Runway ML、Wisecut、Adobe Sensei、Kaleido、帝视科技、CCTV AIGC、影谱科技、Versa（不咕剪辑）、美图影像研究院等
视频自动剪辑	基于视频中的画面、声音等多模态信息的特征融合进行学习，按照氛围、情绪等高级语义限定，对满足条件片段进行检测并合成。	仍处于技术尝试阶段	影谱科技（智能视频编辑）、Adobe（与斯坦福共同研发的AI视频剪辑系统）、IBM Watson（自动剪辑电影预告片）、Sony CSL（Flow Machine）等
视频部分生成	基于目标图像或视频对源视频进行编辑及调试，强调将视频切割成帧，再对每一帧的图像进行处理，通常包含数据提取、数据训练及转换三个步骤。	仍处于技术尝试阶段	腾讯云（与优图实验室、天天P图联合打造的AI变脸玩法）等

图表：Adobe Sensei示例



在几秒钟内消除视频缺陷



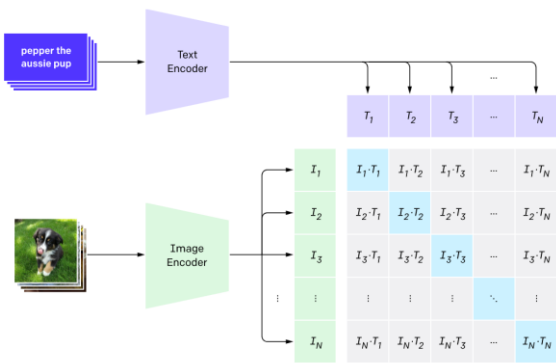
裁剪多台设备的视频，而不会影响操作

# 2.5 跨模态生成：“CLIP+其他模型”成为通用的做法

- **跨模态定义**：跨模态生成，指将一种模态转换成另一种模态，同时保持模态间语义一致性。主要集中在文字生成图片、文字生成视频及图片生成文字。
- **Transformer架构的跨界应用成为跨模态重要开端之一**。多模态训练普遍需要匹配视觉的区域特征和文本特征序列，形成Transformer架构擅长处理的一维长序列，与Transformer的内部技术架构相符合。此外Transformer架构还具有更高的计算效率和可扩展性，为训练大型跨模态模型奠定了基础。
- **CLIP (Contrastive Language-Image Pre-training, 可对比语言-图像预训练算法) 成为图文跨模态重要节点**。
  - ✓ 2021年，OpenAI发布了CLIP，是一种经典的文图跨模态检索模型，在大规模图文数据集上进行了对比学习预训练，具有很强的文图跨模态表征学习能力。CLIP模型包含图像和文本的Encoder两部分，用于对图像和文本分别进行特征抽取。
  - ✓ “CLIP+其他模型”在跨模态生成领域成为较通用的做法，如Disco Diffusion，其原理为CLIP模型持续计算Diffusion模型随机生成噪声与文本表征的相似度，持续迭代修改，直至生成可达到要求的图像。

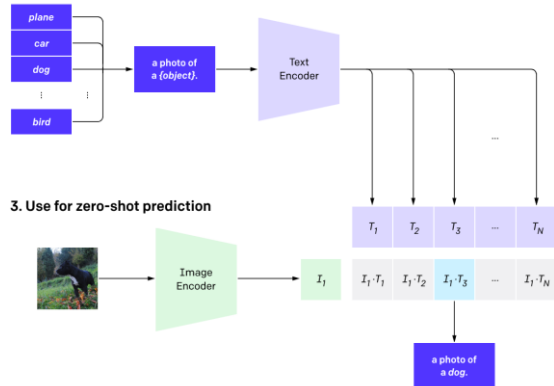
图表：CLIP结构

1. Contrastive pre-training



1. CLIP在特征抽取后，分别对图文向量进行标准化，之后进行内积计算，获取样本间的相似度

2. Create dataset classifier from label text

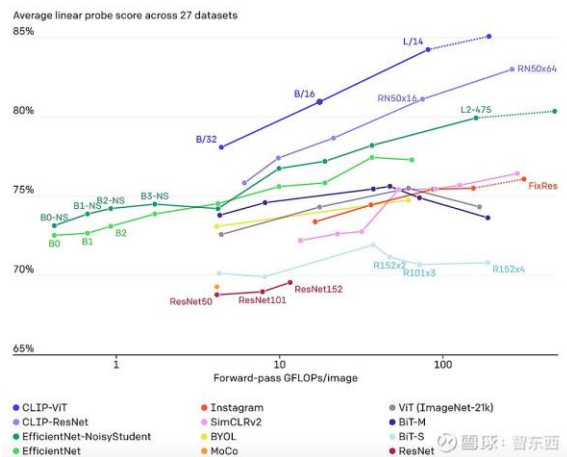


2. 将文本标签转化为提示“A photo of a {object}”，并且使用目标图像作为输出。如果文本“A photo of a dog”于当前图像匹配度最高（具有最大的余弦相似度），即可以说明当前图像的物体是“dog”

图表：CLIP的优势

DATASET	IMAGENET RESNET101	CLIP ViT-L
ImageNet	76.2%	76.2%
ImageNet V2	64.3%	70.1%
ImageNet Rendition	37.7%	88.9%
ObjectNet	32.6%	72.3%
ImageNet Sketch	25.2%	60.2%
ImageNet Adversarial	2.7%	77.1%

尽管两种模型在ImageNet测试集上具有相同的精度，但CLIP的性能更能代表在不同的非ImageNet设置下测量精度的数据集上的表现



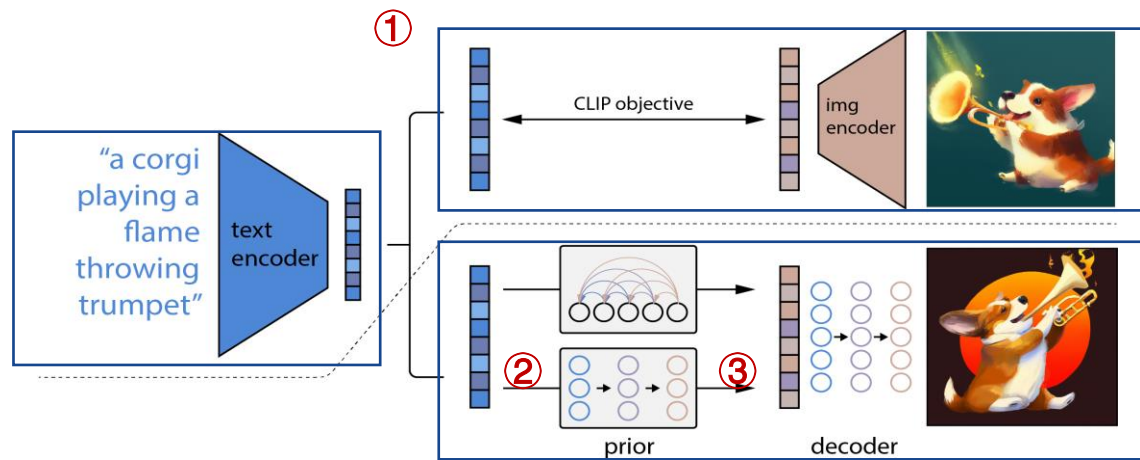
CLIP-ViT和CLIP-ResNet两类CLIP方法在12种模型27种数据集准确率和处理图像大小方面都遥遥领先



# 2.5 跨模态生成：文字生成图像取得突破，其他领域仍有待提升

- **文字生成图像**：2021年，OpenAI推出了CLIP和DALL-E，一年后推出了DALL-E2；2022年5月，谷歌推出了Imagen和新一代AI绘画大师Parti；2022年8月，Stability AI推出Stable diffusion并宣布开源。国内主流的AI绘画平台有文心一格、盗梦师、意间AI、Tiamat等。
- **文字生成视频**：以Token为中介，关联文本和图像生成，逐帧生成所需图片，最后逐帧生成完整视频。但由于视频生成会面临不同帧之间连续性的问题，对生成图像间的长序列建模问题要求更高，以确保视频整体连贯流程。按技术生成难度和生成内容，可区分为拼凑式生成和完全从头生成。
- **图像/视频生成文本**：具体应用有视觉问答系统、配字幕、标题生成等，代表模型有METER、ALIGN等。

图表：DALL-E 2.0 原理



**1、把文本和视觉图像联系起来**

通过CLIP文本/图像编码器将文本描述和图像映射到表示空间（CLIP模型的原理）

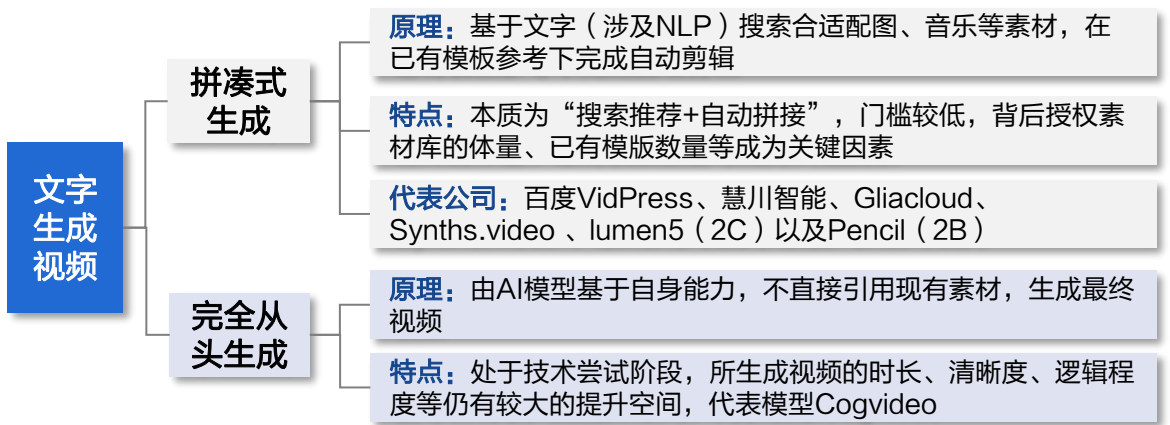
**2、从视觉语义生成图像**

扩散先验从CLIP文本编码映射到相应的CLIP图像编码（扩散模型）

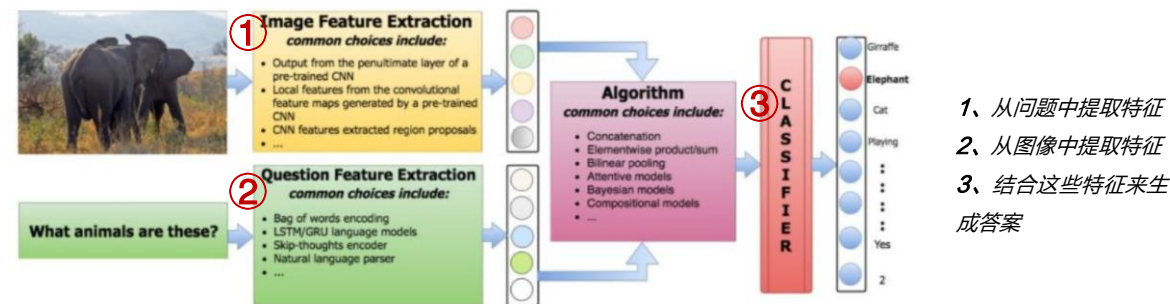
**3、从文本语义到相应的视觉语义映射**

修改版的GLIDE生成模型通过反向扩散从表示空间映射到图像空间，生成众多可能图像中的一个

图表：文字生成视频两种类型



图表：视觉问答系统原理



# 2.5.1 文生图：产品涌现，商业变现以2C为主

■ 2022年AI绘画产品大量涌现，国外Disco Diffusion、Midjourney、DALL-E2、Dream studio等绘画平台接连推出，国内，文心一格、盗梦师、Tiamat等紧跟脚步。AI绘画百度搜索指数一路飙升，在2022年12月达到峰值。

■ AI图像生成已开始商业落地：

- ✓ 2C端：通过“内容付费+会员付费”模式收费。但根据6open调研，60%以上的国内AI绘画用户并未付过费，而愿意付费超100元的占比小于10%。
- ✓ 2B2C：主要作为工具/功能集成到软件主体，提升用户黏性及平台活跃度。以美图秀秀为例，AI绘画作为一项增值服务，引导用户付费或开通会员后使用；抖音、小红书等社交内容平台，AI绘画功能可以带动用户参与活动，拉动创作内容发布和互动率，提升社区活跃度。
- ✓ 2B端：可运用于广告制作等领域，根据定制化需求，快速制作广告，降低成本、时间；也可以用于影视图像、游戏场景等制作上，降低内容成本。

图表：国内外AI绘画产品

地区	产品	所属公司	收费形式
	Disco Diffusion	谷歌	现阶段免费
国外	Midjourney	Midjourney实验室	25次免费生成后，10美元/月（最多200次生成）或30美元/月（无限制生成）
	DALL-E2	OpenAI	首次50次免费生成且每月免费生成15次后，15美元（115次生成）
	Imagen AI	谷歌	尚未明确
	Dream studio	Stability AI	免费生成200张图后，10英镑（1000次生成）
国内	文心一格	百度	免费生成100张图后，9.9元/50张；15.9/100张；49.9元/400张
	Tiamat	退格数字	公测阶段，可免费生成140张图
	6pen art	毛线球	最多免费生成100张图后，5元/10张；30元/100张，100元/400张，500元/2500张
	盗梦师	西湖心辰	免费生成5张图后，5.5元/25张，24.9元/125张；或会员制：99元/月/660张、299元/月/2160张
	即时AI	雪云锐科技	现阶段免费

图表：Tiamat、盗梦师生成的AI绘画作品《夕阳下的摄影师》





## 2.5.1 文生图：三大AI绘画平台对比

文字：火星上的独行宇航员，神秘，多彩，超现实

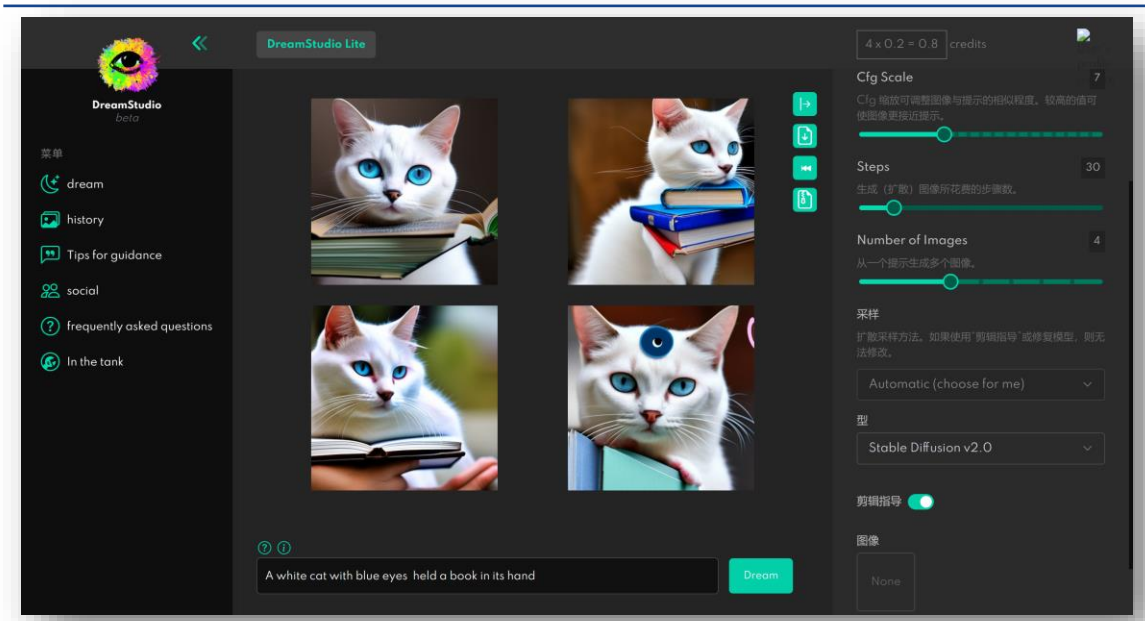


图像生成平台	Midjourney	DALL-E2	Dreamstudio ( Stable diffusion)
推出时间	2022年4月	2022年4月	2022年8月
底层模型	扩散模型+CLIP	扩散模型+CLIP	扩散模型+CLIP
模型开源	不开源	不开源	免费开源
所属机构	Midjourney实验室	OpenAI	StabilityAI
功能		文字生成图像、文本提示的图到图（根据文本提示编辑已有图像）	
一次生成图片数量	4张	4张	1-9张
分辨率	最多到2048*2048	1024*1024	最多到1024*1024
特点	艺术性强，擅长环境效果，特别是科幻场景	语义理解更加准确，风格更广泛	细节更丰富，艺术感强
商业用途	付费会员可以	均可以	均可以
试用	首次加入可获25分钟的免费GPU时间，大致可以支撑25次免费生成。	首次访问可获50个免费积分，接下来每月可获15个免费积分。一次生成消耗约1个积分。	首次注册可获价值2英镑的积分，支撑200次单张图免费生成。
付费	10美元/月（约200次生成） 30美元/月（无限制生成）	15美元（115次生成）	10英镑（1000次生成）
单次生成的费用	0.05美元/次	0.13美元/次	0.01英镑/次

# 2.5.1 文生图：Stability.ai将模型开源，引爆AI绘画

- Stability.ai成立于2020年，2022年10月宣布获得了来自Coatue、光速资本以及O'Shaughnessy总计1.01亿美元的种子轮投资，资金将用来面向图像、语言、音频、视频以及3D领域的AI模型开发，投后估值达到10亿美元。
- 允许任何人在没有监督的情况下使用和构建Stable Diffusion模型，未来有望通过商业版本盈利
- ✓ Stable Diffusion通过引入Latent diffusion的方式来解决原有扩散模型计算代价昂贵的问题。扩散模型主要缺点是去噪过程的时间和内存消耗都非常昂贵，因此使进程变慢，并消耗大量内存。Latent diffusion通过在较低维度的潜空间上应用扩散过程而不是使用实际的像素空间来减少内存和计算成本。
- ✓ 自8月份发布代码以来，已有超过20万开发人员下载了该代码，并使用基于Stable Diffusion算法构建的工具创建了数百万张图像。
- 公司面向消费者的产品Dream Studio拥有来自50多个国家的100多万注册用户，总共创建了超过1.7亿张图像。Dream Studio基础定价为1000积分/10美元，可以生成5000张图片，每张图片在默认设置下需花费0.2积分，此外会根据图片分辨率和处理步骤的数量来调整定价。

图表：Dream Studio 生成的图片示例



图表：Dream Studio中不同分辨率/处理步骤的图片的定价（积分/张）

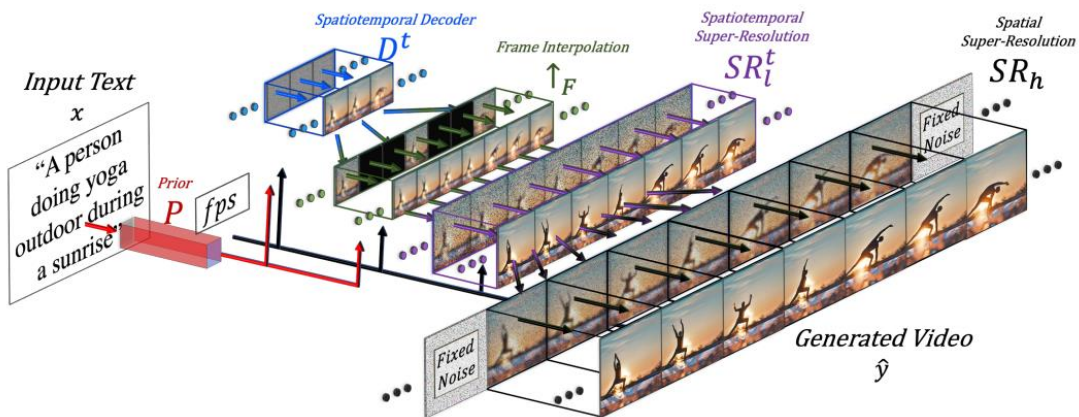
steps	512x512	512x768	512x1024	768x768	768x1024	1024x1024
15	0.1	0.3	0.4	0.5	0.7	1.0
30	0.2	0.5	0.8	1.0	1.4	1.9
50	0.4	0.9	1.3	1.6	2.3	3.2
100	0.7	1.7	2.6	3.1	4.5	6.4
150	1.0	2.5	3.9	4.6	6.7	9.5



## 2.5.2 文生视频：Meta推出Make-A-Video

- 2022年9月，Meta首次推出Make-A-Video，可以从给定的文字提示生成短视频，还可以从给定的视频或图像制作视频，加入到“文字生成视频”落地应用行列。
- ✓ “Make-A-Video”模型研究是建立在T2I (text to image) 生成技术的最新进展之上的，先通过文字生成图像，再提高图像的像素，生成高分辨率和帧率的视频。
- ✓ “Make-A-Video”相较于目前行业内的一些其他技术，具有三大优势：首先，它加速了T2V (text to video) 模型的训练，不需要从头开始学习视觉和多模态表示；其次，它不需要成对的文本-视频数据；最后，其生成的视频具有多样性，审美、艺术风格等都十分丰富。

图表：Make-A-Video模型架构

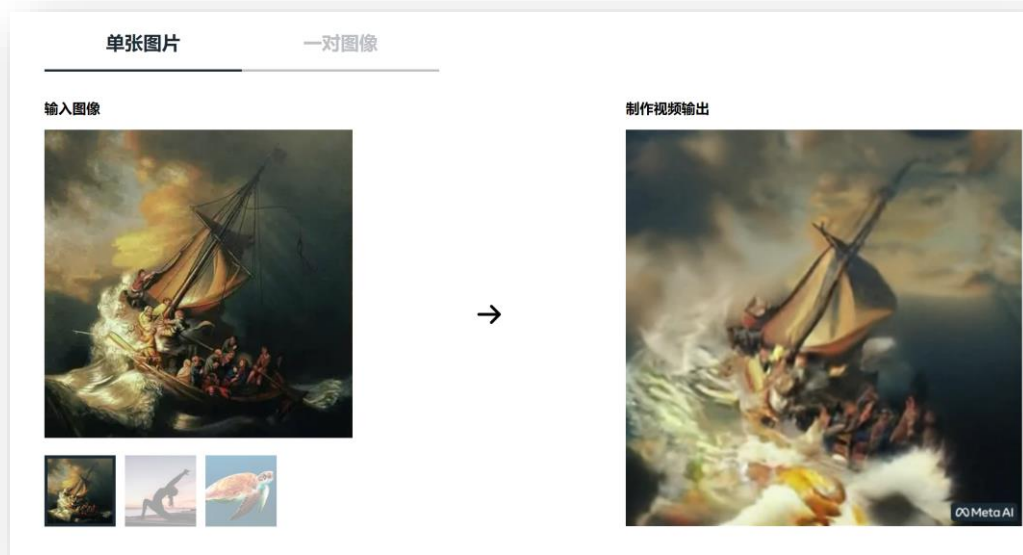


当用户输入一串文字后，系统会生成16张在时间上有连续性的64X64像素的RGB图片，然后这作品图片将会通过插值模型增加视频的帧数，让前后帧之间的动作更加平滑，之后通过两个超分辨率模型，将图像的像素提升到256×256像素后，再提升到768×768像素，生成高分辨率和帧率的视频。

图表：Make-A-Video文字生成视频示例



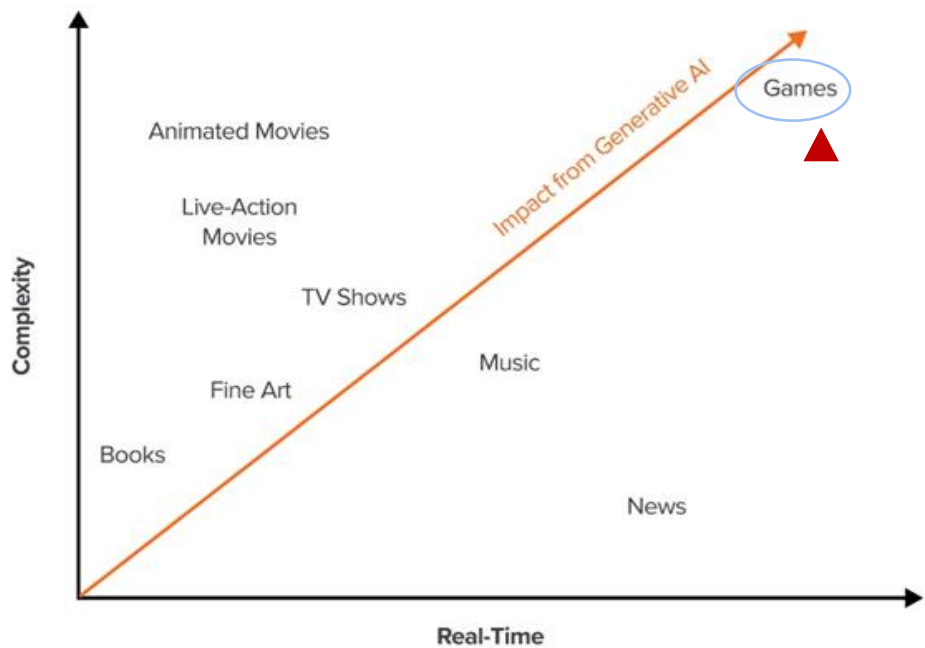
图表：Make-A-Video图片生成视频示例



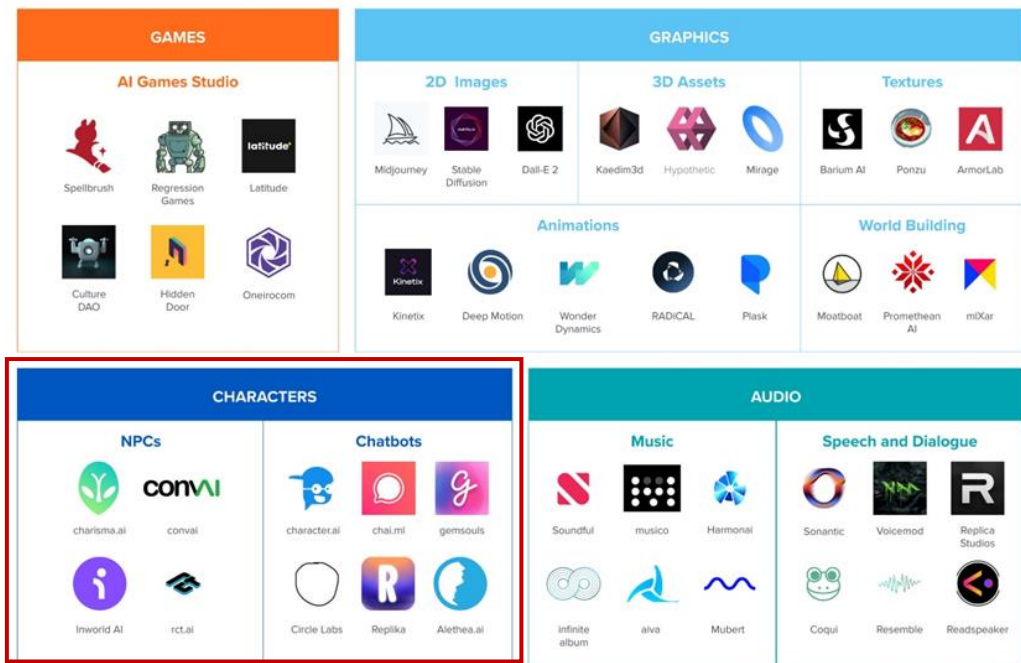
## 2.6 策略生成：主要应用于游戏、智能交互数字人领域

- 策略生成指AI基于特定问题和场景，自主提出解决方案的过程，本质是让AI感知环境、自身状态并基于特定目标决定当下需执行的动作，然后基于下一个状态给到系统的反馈进行奖励或惩罚，最终使系统在不断的“强化”过程中优化“策略”。AI策略生成中普遍采用深度强化学习技术，关键是搭建能充分还原现实因素的虚拟环境及能够合理设置奖惩，在游戏、自动驾驶、机器人控制、智能交互数字人等领域有应用价值。
- ✓ **游戏**：游戏本身为AI提供了最佳的研发场景，有大量数据、贴近实际场景，比如NPC逻辑及剧情生成、游戏操作策略生成等。
- ✓ **智能交互数字人**：包含语音识别、人像驱动、语音交互等算法技术，对语言理解力、实时交互的要求更高，满足三大特征“能听、能说、可互动”。智能交互数字人应用领域涉及金融、政务、教育、医疗、零售、影视等。

图表：生成式AI对不同场景的影响程度



图表：AI在游戏领域的市场图



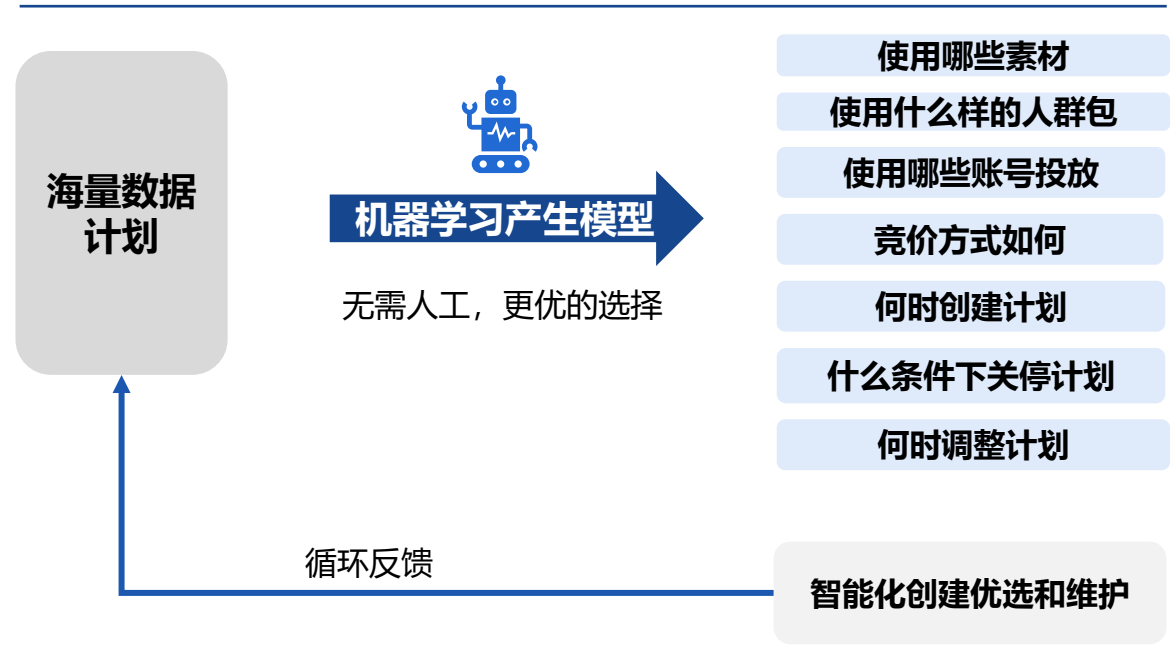


- AI融入游戏全生命周期，全方位赋能行业。
- ✓ 游戏前期制作环节，AI能够提高效率、降低制作成本；
- ✓ 在游戏中，AI既是运营者也是陪伴玩家的智能伙伴，能提升游戏体验的同时维护良好运营秩序；
- ✓ 在游戏周边上，AI也在支持电竞训练分析、直播解说及集锦自动生成等更广泛的应用；
- ✓ 在游戏营销方面，AI提升投放效率，降低人工成本，提升ROI。

图表：AI在游戏产业全链路中的研究与应用能力（以腾讯AI Lab为例）



图表：三七互娱AI投放系统



- **AI生成游戏配乐，更实时、更高效。**2022年10月，动视暴雪申请了新专利，致力于探索“基于游戏事件、玩家资料和玩家反应动态生成音乐的想法”，在多人游戏中，AI将能创建与玩家的游戏环境、选择和进度相匹配的独特配乐，而不是预设的动态配乐；动视暴雪初步计划将该专利用于类似《使命召唤》的多人游戏中。11月，微软也提交了一项“用于合成音频的人工智能模型”的专利，通过AI为电影、电视、游戏等媒体生成声音，且可以配合玩家的实时行为生成。
- **AI生成3D模型，赋能开放游戏突破产能限制。**2020年，微软上线了全球最大的仿真游戏《微软模拟飞行2020》，还原了全球200万个城镇和3.7万个机场，让玩家感受到真实世界里开飞机的感觉，通过与初创公司Blackshark.ai合作，借助着AI和云计算资源，从2D图像生成3D建筑模型，提高产能。创业公司Scenario 提供AI制作大批量、低成本、高精度的制作游戏资产，且整体风格仍然能保持一致。

图表：利用Scenario（AIGC创业公司）AI生成的游戏人物角色、道具



图表：《微软模拟飞行2020》场景

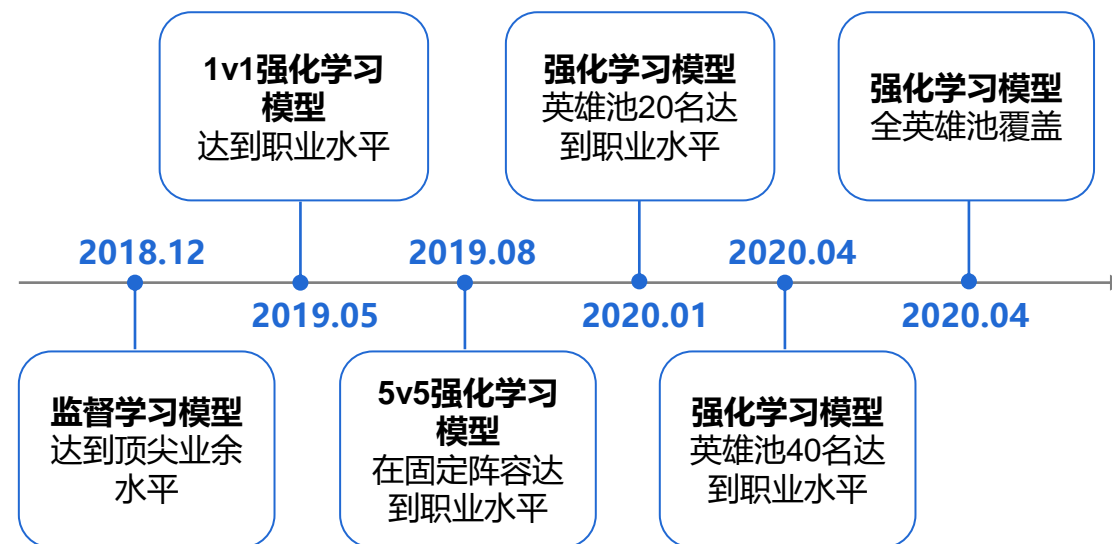


- 2016年，谷歌旗下Deepmind推出的AlphaGO在围棋中战胜了围棋世界冠军李世石。随后，AI决策在Dota2、StarCraft2、德扑、麻将等游戏领域中均展现出了良好的实力。
- 【绝悟】是腾讯AI Lab研发的决策智能AI，通过强化学习的方法来模仿真实玩家，包括发育、运营、协作等指标类别，以及每分钟手速、技能释放频率、命中率、击杀数等具体参数，让AI更接近正式服玩家真实表现。基于“绝悟”，王者荣耀的数值平衡性偏差已经从1.05%下降到0.68%。
- 启元世界的AI Being可应用于掉线托管、AI势力、AI Bot陪玩、智能NPC等领域，提升游戏体验。AI Being具备更高认知决策能力，表现更拟人、战斗水平更智能，不仅对3D开放世界有更全面敏锐的感知，也能更好地完成听声辨位、多人配合、索敌、绕后、找掩体、补状态等任务，已经应用到莉莉丝的FPS大作《Farlight 84》中。

图表：启元世界游戏ai解决方案



图表：王者绝悟AI能力演进路线





## ■ 智能NPC已开始应用落地:

- ✓ NPC (non-player character, 非玩家角色) 指游戏中不受真人玩家操纵的角色。传统NPC对话内容及剧情需人工创造驱动脚本; 智能NPC能分析玩家实时输入, 动态地生成交互反应, 构建几乎无限且不重复的剧情, 增强体验并延长游戏生命周期。
- ✓ NPC逻辑自动生成技术已应用在《黑客帝国: 觉醒》、《荒野大镖客2》、《怪物猎人: 世界》等大型游戏中。网易宣布其《逆水寒》手游中的NPC将全面AI化, 可做到深度人格养成、当玩家知心伴侣。游戏中400多名NPC都加载了网易伏羲人工智能实验室的AI引擎, 有独立的性格特点和行为模式。

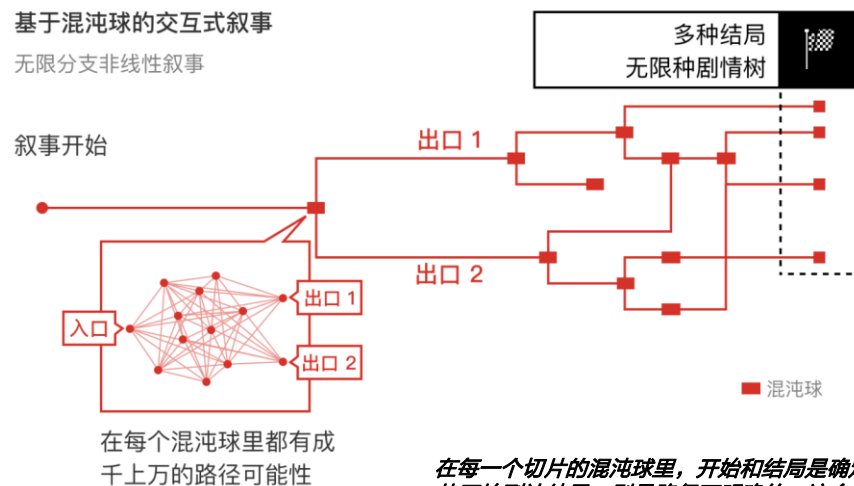
## ■ rct AI混沌球算法提升NPC交互体验:

- ✓ 产品: 个性化且动态的剧情探索、智能NPC/智能生物体/智能环境、“真人”陪玩、智能动画及音效生成技术等。
- ✓ 技术: 基于深度强化学习技术的AI引擎。混沌球 (Chaos Box) 算法能分析玩家的实时输入, 动态地生成NPC的交互反应与新故事情节。在不需要任何脚本的情况下, 就能控制游戏中NPC的行为逻辑, 让其自发地产生非常智能的行为。

图表: 网易《逆水寒》手游实装游戏GPT



图表: rct AI 混沌球算法



在每一个切片的混沌球里, 开始和结局是确定的, 但是玩家每一次如何从开始到达结局, 则是路径不明确的。这个路径只有当玩家不断的和虚拟世界里的NPC作出交互, 这些NPC根据深度强化学习训练后的模型作出动态且实时的反应来推动剧情发展之后, 才会被确定下来。一个混沌球是在一个具体的游戏场景中的最小逻辑单元, 定义了在该场景中的角色与此场景下角色的行为逻辑。

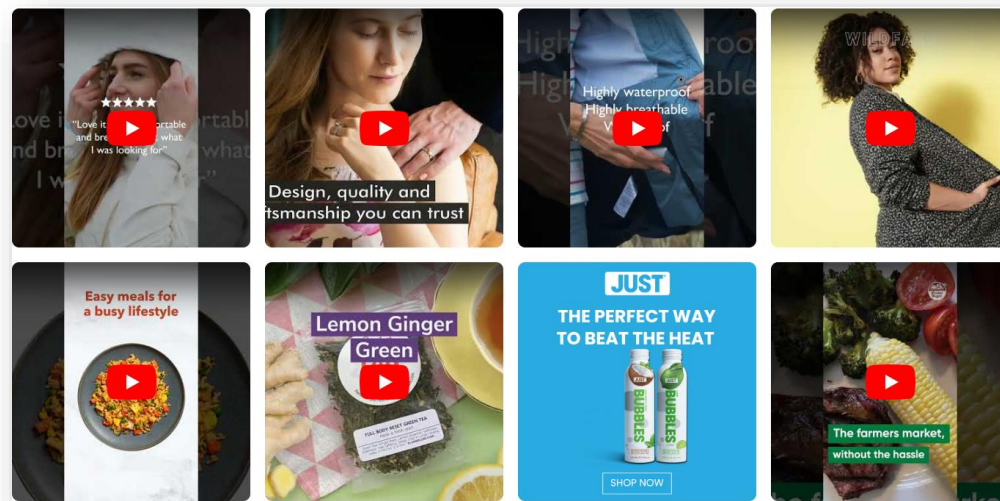


- AIGC的价值在于营销内容的快速迭代更新和更为个性化的内容物料生成。基于AIGC的个性化内容营销，主要是指由AI生成系统与底层的客户数据系统/营销效果反馈系统进行数据联通，实时根据相关数据调整生成需求，由AI快速迭代相关内容生成，进而提升个性化营销的效率和精准性。AIGC能够提高营销效果的关键在于基础素材收集分析、营销策略设定、技术生成、评估优化、数据回流等方面的共同发展。
- 根据内容不同，可以将个性化内容营销分为图像生成和文本生成，目前文本领域落地相对较多，国外已出现不少代表性公司。
  - ✓ 图像生成：AI生成商品模特，提高广告视觉效果。如Rosebud AI可以自动生成不同的模特面孔以定位不同客户，实现千人千面的广告效果，利用其AI生成的模特的第一个活动增加了22%的点击率；此外时尚平台Zalando的营销部门研究出一种新颖的方法，可以生成时尚模特穿着定制服装的高分辨率图像，实现不同颜色、款式的服装在不同造型、姿势的AI模特身上试穿，以供消费者参考。
  - ✓ 文本生成：AI生成广告文案、商品描述等。Pencil基于品牌、受众和产品的信息为营销人员提供AI生成的标题和产品描述。

图表：Zalando营销部门的AI模特图片示例



图表：Pencil生成的广告词示例



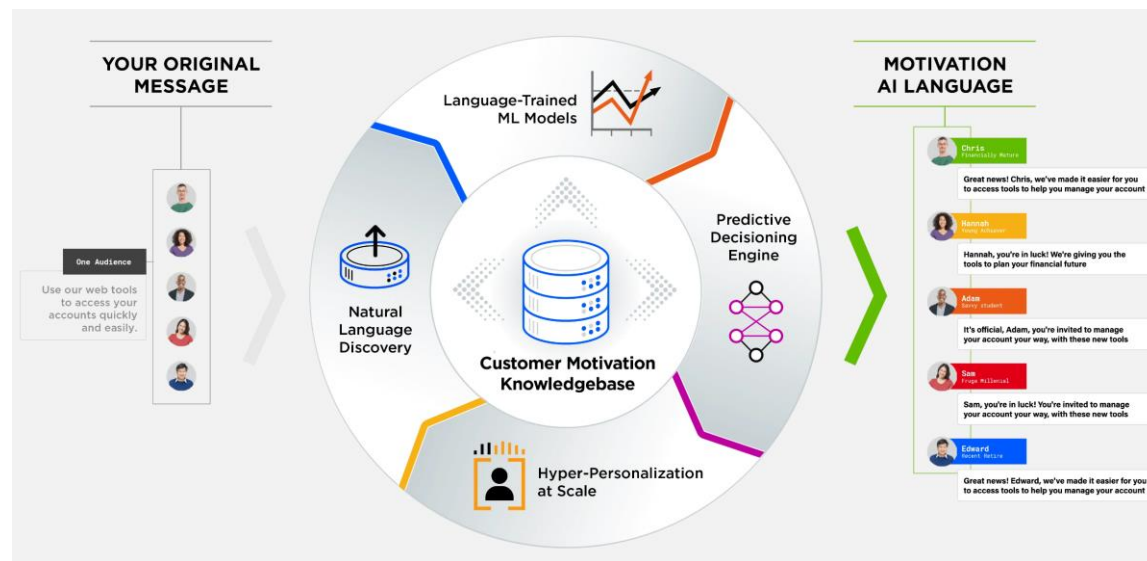
# 【专题3-AIGC在营销的应用】个性化精准营销服务落地

- Persado是一家可基于电子邮件、网页和其他营销活动创造出高效的文案，将人工智能应用于营销的创意公司。Persado不仅仅是简单的AI生成内容，更是能够跨触点、渠道和客户群体来生成品牌个性化内容。2016年，Persado再获3000万美元C轮融资，由高盛领投，总融资达到6600万美元。
- Persado的语言分析器将营销内容拆解为语言、情绪、产品特性等几个要素不同形式的组合，其中情绪性词语在说服客户中起到了关键作用。Persado开发情绪轮来专门分析和利用情绪性的用语，系统中大约有100多万单词和短语，人工智能会从中筛选关键字并且生成100-600个字的文案内容，文案版本可以高达16个。结合不同平台对于措辞的要求，加上用户设置的情绪要求、功能描述和格式，Persado可以高效地生成合适的文案。文案生成之后，Persado会将文章打开率、广告链接点击和转化等数据记录在案，通过机器学习多版本的营销内容效果对比来不断改进自己。
- **提高点击率，实现精准营销，增加企业广告价值。**经过试验测试对比，人工智能撰写的广告文案要比营销人员撰写出的文案，点击率高出450%。一份与全球财富500强的报告显示，Persado的智能语言分析器增加了112%的订单率。Persado主要客户涉及电商、零售、金融、电信、旅游等行业，包括摩根大通、花旗银行、Expedia等知名企业。

图表：Persado历次融资情况

序号	日期	融资轮次	融资金额	投资机构
1	2016年	C轮	3000万美元	高盛领投，Bain Capital Ventures、StarVest Partners等原有投资者跟投
2	2015年	B轮	2100万美元	StarVest Partners领投，花旗银行和美国运通旗下的风投部门等参与跟投
3	2013年	A轮	1500万美元	Perception Companions与红杉印度领投，现有投资者Together Fund、Emergent Ventures跟投

图表：Persado Motivation AI平台





图表：Pencil广告营销案例与产品定价

“Pencil has been a game changer for us... it's helped us generate some of our best ever FB creative. The team kicks ass too... Check 'em out! 🍷🔪”

**lyka** James Harding  
Head of Growth, Lyka Pet Food

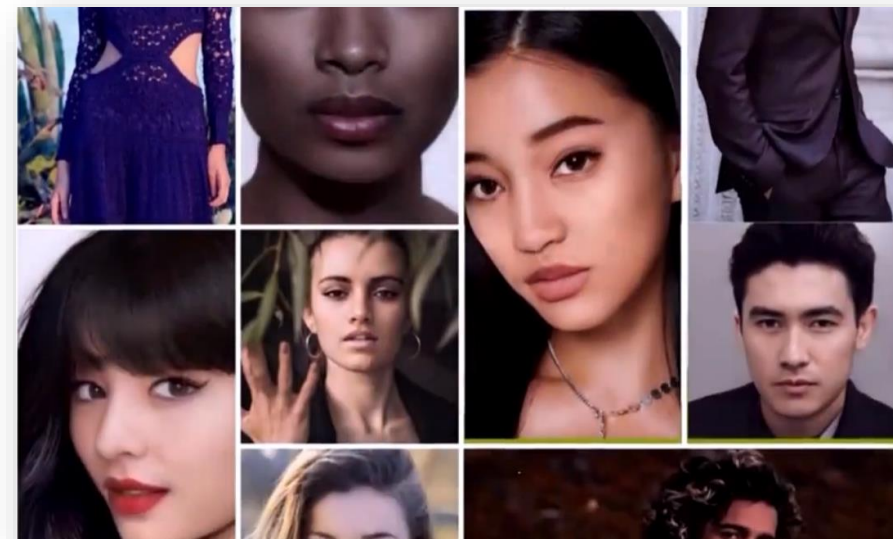
16 Ads tested	15% Avg. CPA drop	56% Max. CPA drop
------------------	----------------------	----------------------

Starter	Pro	Unlimited
For startups & freelancers Generate ad creatives predicted to win with AI.	For brands & agencies Get ad creatives weekly from a shared Creative Strategist.	For enterprises Unlimited ad creatives from a dedicated Creative Strategist.
<b>\$119/mo</b> \$119 billed monthly	<b>\$299/mo</b> \$299 billed monthly	<b>\$999+/mo</b> \$999 billed monthly
<a href="#">Try free 7 days</a>	<a href="#">Get started now</a>	<a href="#">Get started now</a>
<ul style="list-style-type: none"> <li>Generate ads yourself with AI</li> <li>50 exports (add exports \$10)</li> <li>1 brand (add brands \$119/mo)</li> <li>Free stock music, images &amp; video</li> <li>Predictions, insights &amp; benchmarks</li> <li>Sharing, commenting &amp; approvals</li> <li>Direct channel integrations</li> <li>Chat support</li> </ul>	Everything in "Starter" plus: <ul style="list-style-type: none"> <li>Shared Creative Strategist</li> <li>Ads made by Creative Strategist</li> <li>1 brief/week to Creative Strategist</li> <li>Slack with Creative Strategist</li> <li>100 exports (add exports \$10)</li> <li>1 brand (add brands \$249/mo)</li> </ul>	Everything in "Pro" plus: <ul style="list-style-type: none"> <li>Dedicated Creative Strategist</li> <li>Ads made by Creative Strategist</li> <li>1 brief/day to Creative Strategist</li> <li>Unlimited exports</li> <li>Custom pricing for multiple brands</li> <li>Bank transfers &amp; payment terms</li> </ul>
<a href="#">Try free 7 days</a> Or <a href="#">Book a call</a>	<a href="#">Get started now</a> Or <a href="#">Book a call</a>	<a href="#">Get started now</a> Or <a href="#">Book a call</a>

图：DataGrid开发的用于自动生成全身模型的AI



图表：Rosebud制作的AI模特





1

AIGC破圈元年

2

AIGC应用场景

3

产业链及相关公司

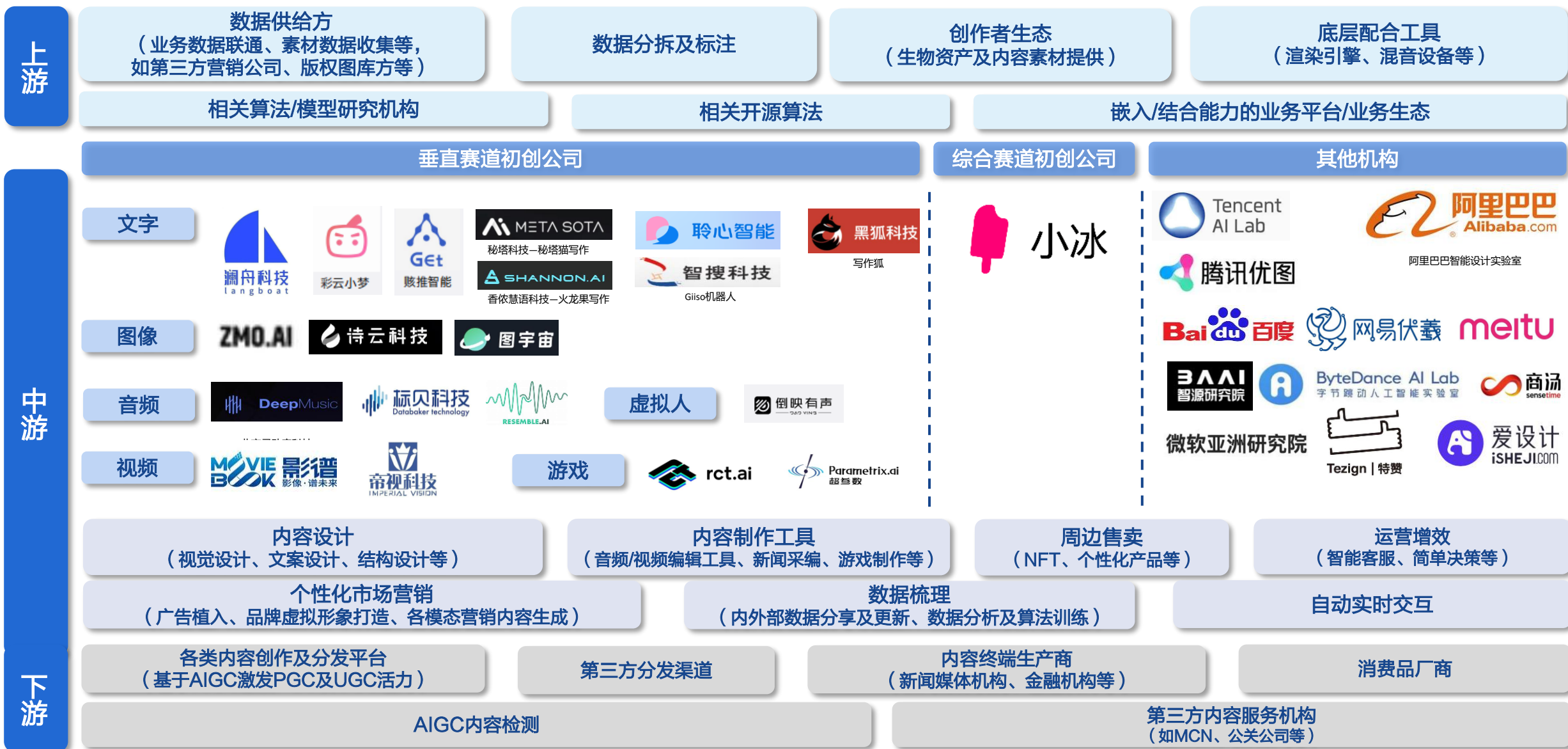
3.1 产业链

3.2 国内发展现状：互联网巨头

3.3 国内发展现状：相关上市公司

3.4 国内发展现状：代表性产业公司

# 3.1 产业链：中国AIGC行业尚未发展成型，代表公司较少，上游欠缺



## 3.2 国内发展现状：互联网巨头多场景综合布局

公司	股票代码	ChatGPT或对话机器人相关的布局	AIGC相关布局
百度	BIDU.O	计划在3月推出文心一言 (ERNIE Bot) , 基于百度语言模型文心大模型ERNIE3.0	旗下百度AI开放平台拥有语言技术、文字识别、图像技术、视频技术等AI能力, 旗下文心大模型涵盖NLP、CV、跨模态等多个领域; 具体在绘画、写作、视频、数字人领域均有布局
腾讯控股	0700.HK	申请的“人机对话方法、装置、设备及计算机可读存储介质”专利获授权, 持续投入AI等前沿技术的研发	基础研究方向包括计算机视觉、语音技术、自然语言处理和机器学习, 应用探索聚焦于AI+游戏、AI数字人、AI+内容、AI+社交AI四类
阿里巴巴	BABA.N	阿里版ChatGPT正在研发中, 目前处于内测阶段, 将和钉钉深度结合	阿里达摩院牵头推出魔搭社区 ModelScope, 社区首批上架超 300 个模型, 覆盖了视觉、语音、自然语言处理、多模态等AI主要领域, 向所有开发者开放
快手-W	01024.HK	正在开展大规模语言模型相关研究, 并启动了相应专项, 覆盖LLM模型训练、文案自动创作与生成、对话系统开发等领域	旗下快手AI开放平台提供全球领先的AI技术服务, 产品服务涵盖计算机视觉、计算机图形学、自然语言处理、音频技术、视频技术、知识图谱、机器学习、AR/VR/MR、多模态等核心技术领域; 具体应用有虚拟人、语音识别、视频自动剪辑等
京东	09618.HK	宣称会不断结合ChatGPT的方法和技术融入到产品服务中	旗下AI开放平台NeuHubAI提供全球领先的声学、语义、对话、机器学习等多项人工智能技术, 通过平台赋能至客服、商业、市政、医疗等领域
网易有道	DAO.US	已投入到ChatGPT同源技术产品的研发, 未来或将推出ChatGPT同源技术产品, 应用场景围绕在线教育, 开始在AI口语老师、中文作文批改等细分学习场景中尝试应用	布局AI多年, 在神经网络翻译、计算机视觉、智能语音AI技术、高性能计算等方面具备核心技术; 旗下有道智云AI开放平台提供自然语言翻译服务YNMT、光学字符识别服务OCR、智能语音服务SR&TTS
三六零	601360.SH	计划尽快推出类ChatGPT技术的demo版产品	360AI安全实验室开发了AI框架安全监测平台
字节跳动	未上市	未布局	旗下人工智能实验室研究领域涉及自然语言理解、计算机视觉、机器学习、计算机图形&增强现实等; 具体应用有AI写稿机器人、虚拟广播员和歌手、AI视频制作



# 3.2.1 百度：“创作者AI助理团”助力AIGC智能创作

**“未来十年，AIGC将颠覆现有内容生产模式，可以实现以十分之一的成本，以百倍千倍的生产速度，创造出有独特价值和独立视角的内容”**

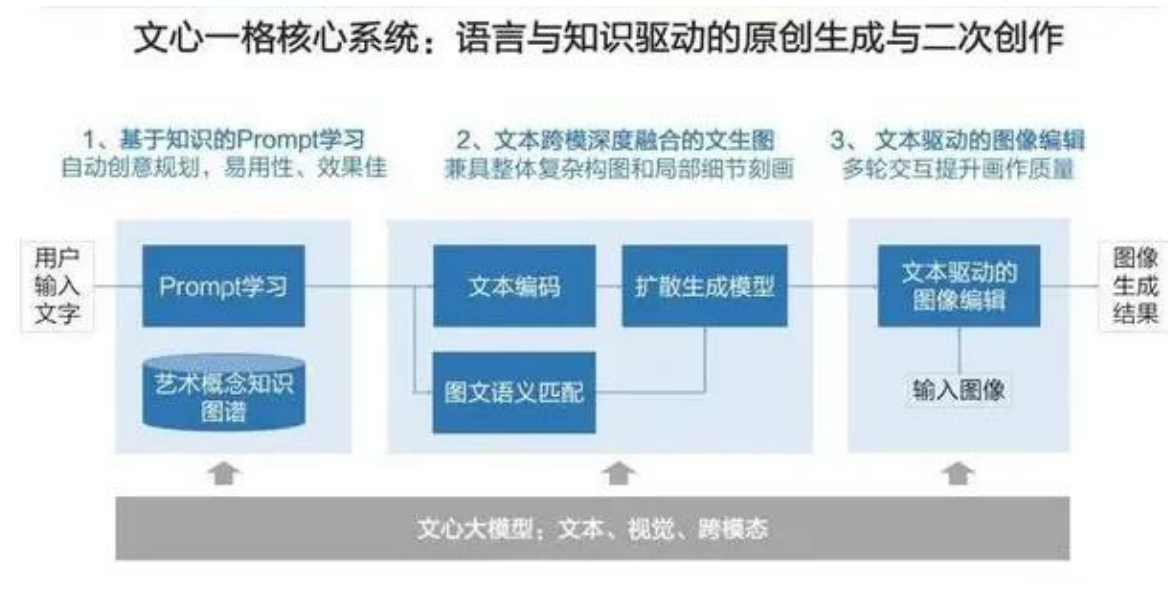
——百度移动生态事业群组 (MEG) 总经理何俊杰

- 2022年9月，万象·百度移动生态大会召开，百度发布多项AIGC应用内容生产领域的技术和产品，尤其是“创作者AI助理团”和“百度APP数字人计划”的发布，助力内容创作驶进AIGC元年。百度百家号携手澎湃新闻、新京报等数十家权威媒体成立“AIGC媒体联盟”，推动AI普惠和赋能媒体行业，加速内容生产的智能化。
- **【创作者AI助理团】发布，助力AIGC智能创作：**通过文心大模型、文心一格、图文转视频等技术，为创作者提供“AI文案助理”、“AI插画助理”、“AI视频制作助理”，帮助创作者实现“一个人成为一支队伍”。
  - ✓ **AI文案助理：**基于百度文心大模型，可以根据创作者的需求快速产出不同风格的文章标题、内容段落。
  - ✓ **AI插画助理：**在文心一格的支持下，可以根据输入的关键词，自动“画”出十余种不同风格的画作，帮助创作者实现“配图自由”。
  - ✓ **AI视频制作助理：**利用多模态AI处理技术，实现图文自动生成视频能力，速度已可达“分钟级”，这项技术将图文到视频的创作周期缩短超3倍。

图表：文心产业级知识增强大模型

产品与社区	文心一格	文心百中	赐谷社区			
工具与平台	EasyDL-大模型	BML-大模型	大模型 API			
	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			
文心大模型	行业大模型					
	国网-百度·文心	浦发-百度·文心	航天-百度·文心	人民网-百度·文心	冰城-百度·文心	电影频道-百度·文心
	深燃-百度·文心	吉利-百度·文心	康康-百度·文心	TCL-百度·文心	碧海-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-UMS	文图生成 ERNIE-VILG	文档智能 ERNIE-Layout	蛋白质结构预测 HelixFold
	对话 PLATO	搜索 ERNIE-Search	信息抽取 ERNIE-UIE	OCR图像表征学习 VIMER-StrucText		
	跨语言 ERNIE-M	代码 ERNIE-Code	图网络 ERNIE-Sage	多任务视觉表征学习 VIMER-UFO		
	语言理解与生成		多任务视觉表征学习 VIMER-TCIR	视觉-语言 ERNIE-VIL	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)	鹏程-百度·文心 (千亿级)	ERNIE 3.0 Zeus (任务知识增强千亿级)	化合物表征学习 HelixGEM	单序列蛋白质结构预测 HelixFold-Single

图表：文心一格核心系统



# 3.2.1 百度：文心一格文字生成图片示例

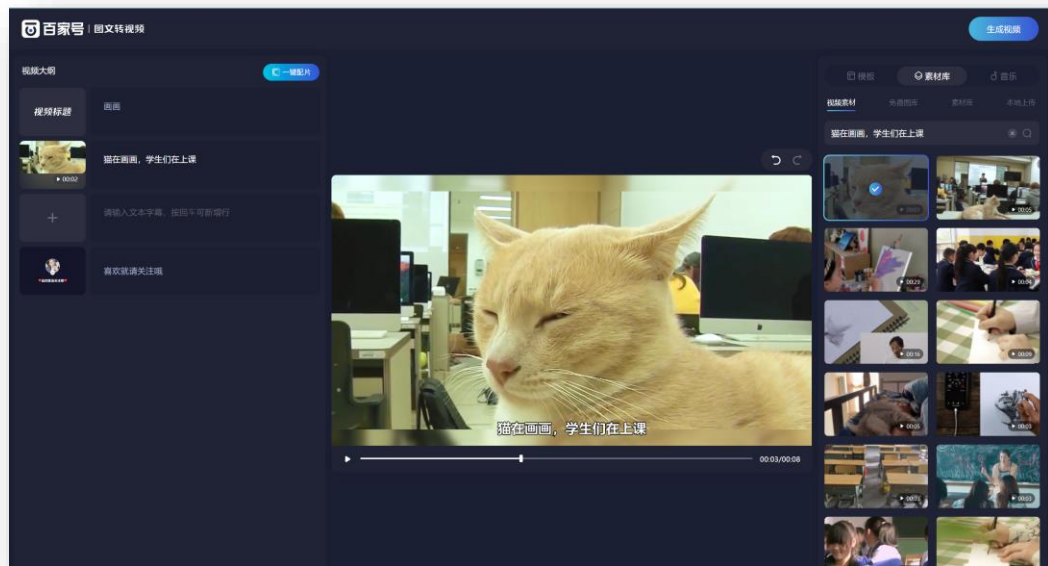




## 3.2.1 百度：发布“数字人计划”，即将推出ERNIE Bot对标ChatGPT

- **百度APP数字人计划**：将数字人技术与图文转视频、TTS语音合成技术结合，为媒体及创作者定制真人孪生数字人。创作者输入文本即可生成有数字主持人口播的视频，不仅保证信息播报准确性，还具备多国语言口播、情绪识别等人格化功能，实现AI驱动24小时不停播。
- **百家号已有部分AIGC应用落地**：如图文转视频技术，只需一键授权，即可自动将图文转换成视频，让图文创作者跨越视频转型门槛，让视频创作者创作视频更高效、便捷。中国青年网官方短视频品牌“青蜂侠”就借助百家号图文转视频技术，在传播效果上，实现视频内容千万级曝光，较原图文类内容提升超10倍；在传播时效上，AI内容创作可灵活应对突发报道，最快领先人工创作30分钟，快速产出传播内容。
- **计划发布ChatGPT竞品**：百度计划在3月推出文心一言（ERNIE Bot），基于百度语言模型文心大模型ERNIE3.0，ERNIE3.0提出了一个多范式统一的大规模预训练框架，基于该框架，ERNIE3.0融合了自回归网络和自编码网络，同时由于大规模知识图谱类数据的引入，使得模型能够在理解任务、生成任务、零样本学习任务和常识推理任务上均获取优秀的表现。

图表：百家号实现图文生成视频



图表：数字主持人“青小霞”

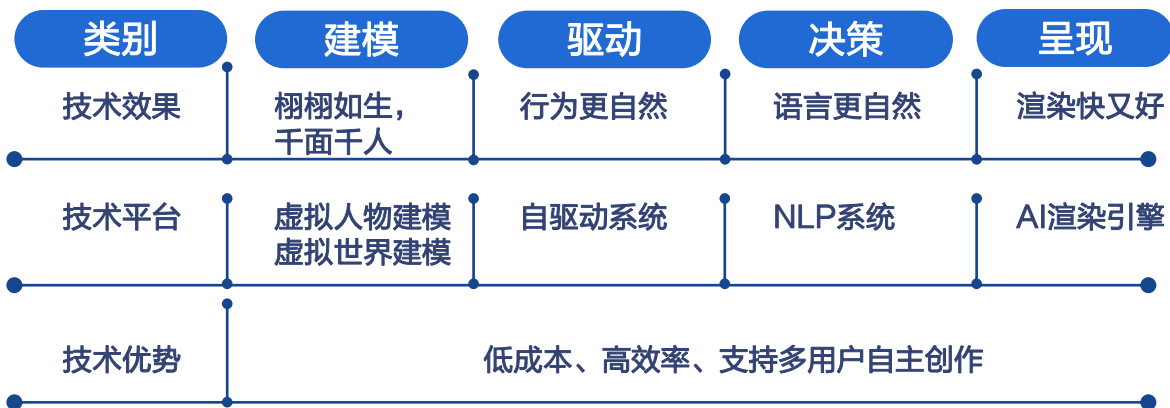




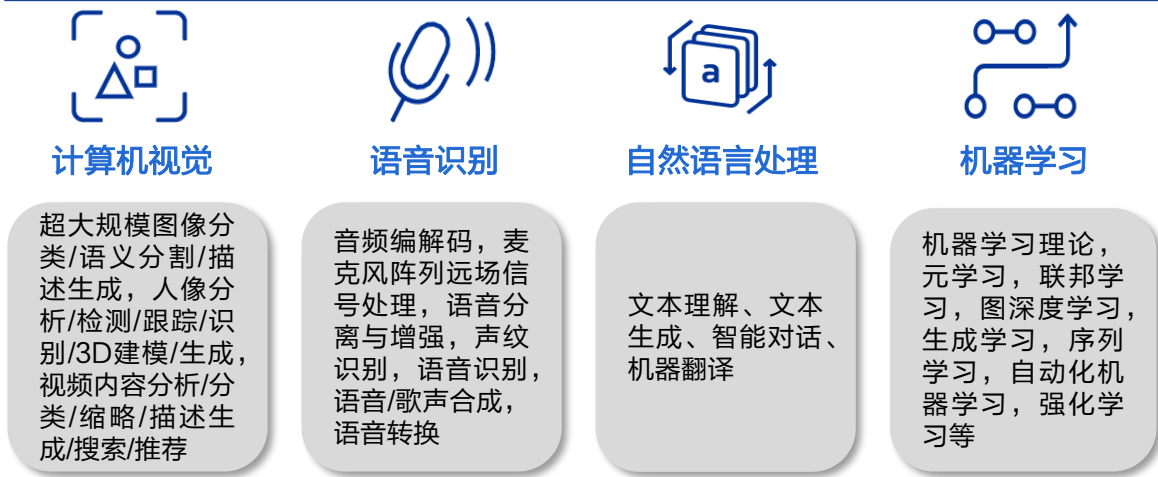
### 3.2.2 腾讯：“AI+游戏/数字人”多场景综合布局AIGC

- 腾讯AI Lab的基础研究方向包括计算机视觉、语音技术、自然语言处理和机器学习，应用探索结合了腾讯场景与业务优势，聚焦于游戏、数字人、内容和社交AI四类，技术已被微信、QQ、天天快报和QQ音乐等上百个腾讯产品使用，并探索AI与生命科学、医疗、农业、工业等行业的创新结合。
- **AI+游戏**：腾讯AI Lab基于自己的多模态学习及生成能力在游戏领域进行了全流程的布局。
- ✓ **绝悟&开悟**：“绝悟”是腾讯AI Lab于2017年开始研发的决策智能AI，通过强化学习方法来模仿真实玩家，让AI更接近正式服玩家真实表现。“开悟”是腾讯依托腾讯AI Lab和王者荣耀在算法、算力、实验场景方面的核心优势，于2019年牵头构建的AI多智能体与复杂决策开放研究平台。
- **AI+数字人**：腾讯AI Lab自2017年起开始虚拟人研究，目标是从语音、口型、表情到全身动作都实现高度拟人，并掌握听、说、读、写、想等全方位技能。2020年推出AI虚拟人艾灵，是业内首个由AI驱动虚拟偶像，2021年推出超写实3D虚拟人小志，可以通过输入文本来合成语音说话，还能基于文本或语音自动生成准确的口型及生动自然的表情。

图表：腾讯虚拟人技术系统框架



图表：腾讯 AI Lab 研究领域及具体研究方向



图表：腾讯 AI Lab应用领域



### 3.3 国内发展现状：AIGC相关上市公司

图表：国内拥有AIGC相关业务的上市公司

领域	公司	股票代码	AIGC相关业务
综合	昆仑万维	300418.SZ	发布了「昆仑天工」AIGC全系列算法与模型，并宣布模型开源。「昆仑天工」旗下模型包括天工巧绘SkyPaint、天工乐府SkyMusic、天工妙笔SkyText、天工智码SkyCode，覆盖图像、音乐、文本、编程等多模态内容生成能力，标志着昆仑万维成为目前国内AIGC领域布局最为全面的公司之一，也是国内少数全身心投入到AIGC开源社区的公司。
综合	科大讯飞	002230.SZ	讯飞在人工智能三要素算法、算力、数据均具有深厚的积累，在讯飞布局的方向处于业界先进水平。讯飞语音合成、虚拟人相关技术等都属于AIGC，同时讯飞也积极布局基于预训练模型的文图生成、对话生成等技术方向，相关技术均达到业界领先水平，并通过AI虚拟人交互平台在媒体、金融、教育、文旅等领域赋能内容生产创作和业务服务。公司于2022年10月24日开发者节已经正式发布了首个自研机器狗，基于机器人超脑平台，融合多模感知、多维表达、深度理解、运动智能相关算法，机器狗具备更多的AI能力，未来将在工业巡检、城市治理等行业领域中发掘更多的应用场景。
综合	万兴科技	300624.SZ	万兴播爆是公司打造的一款基于AI、数字人技术，面向电商用户提供丰富模板的视频生成工具，是公司正在探索、拓展的新一业务领域。明星产品WondershareFilmoraV12/万兴喵影2023新版本国内外同步，一次性推出全属性关键帧、AI智能抠像、AI智能降噪、调整图层、绘制遮罩、AI音乐重组、3D文字等多种全新优质特色功能。公司旗下首款AI绘画软件“万兴AI绘画”正式开启公测。
营销	蓝色光标	300058.SZ	公司的AIGC布局不仅包括“分身有数”，“蓝标智播”等AI产品，也涵盖“销博特”等多人协同创作平台，销博特是一款基于云端聚焦AI营销场景的多人协同创作平台，通过结合人工智能、统计算法和多维数据库，一键自动化生成策划案、品牌分析报告、消费者洞察等内容。过去一年，销博特的产品能力不断升级迭代，目前，已累计注册用户超60000家，覆盖10多个行业，400多个品类。2022年开始，提案中的参考场景为例已由AIGC产出完成；以数字人建模为例，2021年蓝标已实现使用图片生成模型的技术，3分钟即可生产一个模型，并且在与清华美院联合举办的国际元宇宙论坛中，为每一位嘉宾生成了1对1的模型。
营销	遥望科技	002291.SZ	作为直播电商行业的数字营销先锋力量，公司一直秉持互联网技术创新发展的理念，目前拥有首款数字虚拟人「孔襄」、孪生主播技术及数字影棚等创新业务。公司规划在2023年发力虚拟人直播等业务，计划利用公司过往积淀的技术优势，实现高效率低成本AIGC内容生成，使得直播时长更长，以长尾效应进一步扩大供应链竞争优势。
营销	因赛集团	300781.SZ	公司在持续探索AI技术在品牌内容营销方面的应用，并取得一些应用成果，其中也包括GPT技术。公司开发应用该项技术，早期主要是通过GPT2模型训练，探索自然语言处理技术在营销内容方面的运用，开发微信公众号标题和内容文案的自动生成工具与游戏社群互动工具，应用在社群互动等营销场景上，从而减轻运营成本，提高运营效率。目前，公司的人工智能创意生成及管理平台因赛引擎，正紧跟AI的技术迭代步伐，研发应用ChatGPT等相关的技术，持续开发自然语言处理（文章撰写、标题创作、创意文本）等方面的应用，与图像和视频AIGC技术协同，提升品牌的内容营销的质量和效率。
虚拟人	风语筑	603466.SH	公司在虚拟展厅、虚拟数字人、虚拟场景以及数字艺术创作领域均有结合AI人工智能技术的探索与应用，主要集中在AI生成艺术、AI辅助内容创意与生产等领域。公司已结合AIGC技术在文生文、文生图、文生音视频等领域进行场景应用，随着AIGC技术从“通用模型”向“行业垂直”方向加速迭代，公司还将强化在3D建模和虚拟空间生成等领域的定向训练和模型优化。
虚拟人	天下秀	600556.SH	公司创建了3D虚拟社交软件“虹宇宙”和自媒体数字藏品工具集“TopHolder头号藏家”。希望将平台多年来积累的红人资源、品牌资源、大数据体系和服务能力通过区块链、AI等前沿技术应用到Web3.0时代全新的创作者经济生态中，为内容创作者探索和打造新的展示平台和变现方式，为品牌方构建更加生动的营销场景及模式。
虚拟人	芒果超媒	300413.SZ	围绕数字虚拟人相关技术研发与应用，推出虚拟主持人YAOYAO、小漾等；推出了“乘风破浪·生日花”“光芒的力量”“芒果虚拟伙伴·阿芒”“花少·星空下的帐篷男孩”等多款剧综IP衍生数字藏品；依托湖南广电元宇宙产品“芒果幻城”，在重点节目《100道光芒》中加入了“虚拟人”直播内容进行创新，在“光芒Web3.0赛道”直播大考中，聚焦于数字藏品、虚拟互动等前沿概念，节目特别策划了独具赛道特色的“虚拟人”互动直播环节，让虚拟主持人“小漾”现场与观众进行互动。
虚拟人	捷成股份	300182.SZ	公司参股子公司世优科技数字人已经接入chatGPT，正在通过数字人自身的人设背景等相关数据集，并基于OpenAI来训练这个数字人专有大脑形成个性化模型。未来世优科技的AI数字人除了将拥有chatGPT的超强“AI大脑”通用能力外，特定数字人还将具备个性化独有的人设背景表达能力。世优科技拥有包括实时驱动数字人、AI数字人、2.5D数字人、捏人等全栈数字人产品技术体系，应用于广电媒体、虚拟IP/偶像、品牌营销、电商、短视频/直播、元宇宙等。

注：蓝色标记为重点关注公司



### 3.3 国内发展现状：AIGC相关上市公司

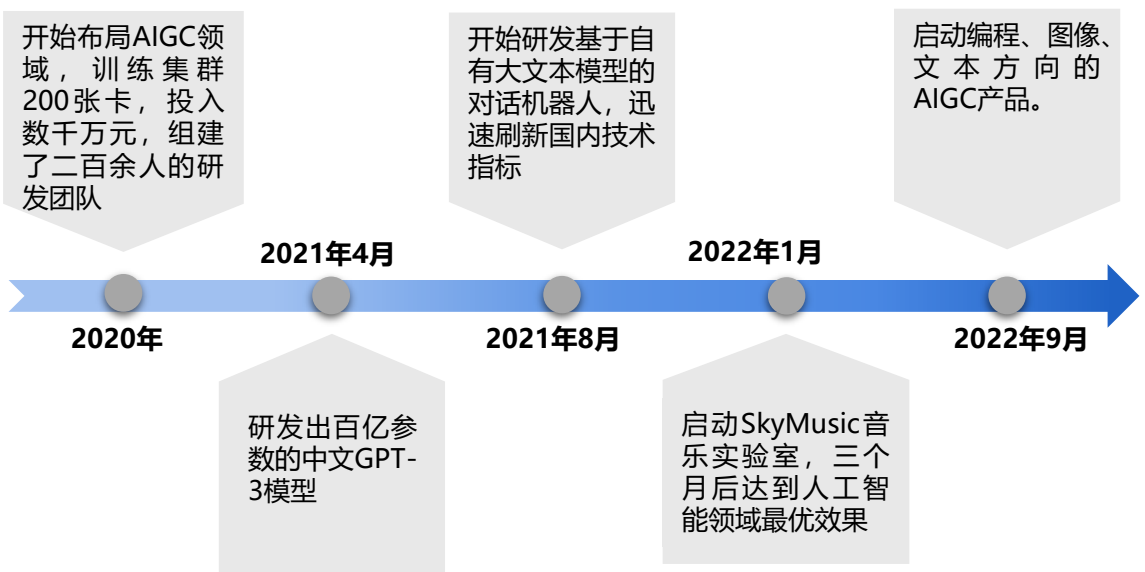
领域	公司	股票代码	AIGC相关业务
文字	果麦文化	301052.SZ	果麦AI创作机器人可实现通过采集互联网大数据精选文章、本地文件导入转化为自己的内容库，有机训练段落、词句、文章、知识四维AI技能。机器通过持续深度学习，可以生成语句通顺、可读性强的优质内容，素材专业、多元实现一键自动成稿，保证改写后文意相同、内容相似，轻松实现底稿优化转换，达到高效孵化图书营销软文的目的。
文字	中文在线	300364.SZ	公司已推出AI主播、AI绘画和AI文字创作功能，并在研发推动更多AIGC功能落地。其中AI主播已在有声书生产中应用，根据文字故事情节的真实语境，用独创的情绪情感控制模块，结合语音深度学习合成技术，真实还原角色情感，生动模拟人声音色以塑造角色，讲述故事。AI文字创作功能已在17K小说网上线，可自动生成文字内容，作者在使用该功能时，通过针对不同的描写场景填写关键词和辅助短语，即可生成对应的文字内容描写用于作品使用，目前可针对人物、物品等进行AI文字创作，并针对不同的作品类别如古风、都市等进行不同的内容适配。
文字	奥飞娱乐	002292.SZ	持股5%的公司北京光年无限科技有限公司拥有AI对话机器人产品-图灵机器人开放平台，开发者可自行快速接入并创建个性化机器人，包含聊天机器人、智能客服等，目前累计注册企业开发者超过150万。光年无限核心技术包含语义理解、人机对话、深度问答、多模态交互、语音交互、视觉交互等。
文字	中科金财	002657.SZ	中科金财WEB3.0内容制作引擎、多模态超写实数智人、全媒体智能客服（数字人智能大脑解决方案）、RPA机器人等解决方案中运用了多模态输入/输出、视觉合成、知识处理等人机交互技术。公司智能客服机器人采用的是Transformer神经网络算法，形成了面向复杂NLP语义识别和内容生成的智能机器人对话系统、自动代码生成引擎、自动流程指令生成引擎。目前已实现为金融机构等的用户提供“不仅能对话、还能办业务”的新智能服务体验。公司自行研发拥有了自主知识产权的SINORPA机器人，支持快速部署、无接口对接，可在银行、证券、电商、物流等行业众多业务场景广泛应用。
图像	视觉中国	000681.SZ	公司与百度旗下的AI作画平台文心一格将在创作者赋能和版权保护等方面展开多项合作，共探AIGC内容产业发展方向。公司旗下元视觉平台有发行AI生成的数字艺术品，相关业务尚处于初期阶段；《鸿蒙宇宙》是公司旗下“元视觉”艺术网于2022年11月发行的一款含版权AIGC数字艺术品，由摄影师宾果与AI共同完成。公司旗下“元视觉”艺术网共发行已收录入“元视觉”艺术网“AIGC”专辑的艺术作品17件，AI斯坦主题系列作品200件。
游戏	完美世界	002624.SZ	已将AI相关技术应用于游戏中的智能NPC、场景建模、AI剧情、AI绘图等方面。在研发方面，公司通过AI技术进行智能整合，为项目打通方案，满足功能要求，例如通过AI技术完成场景建模、纹理渲染等；此外公司还在游戏研发过程中使用AI绘图等技术，对于开拓设计思路与创意灵感、提升游戏研发效率起到了很好的推动作用。
游戏	三七互娱	002555.SZ	三七互娱的“量子-天机”AI系统通过将人工经验转化为机器代码，不仅全面提升了发行和运营的效率，更强化了这两大业务线之间的联动。“量子”系统快速提升广告投放速度，并且实现了投放业务“7×24小时”的批量化与自动化。
教育	传智教育	003032.SZ	国内领先的高精尖数字化人才培养的职业教育机构，营业收入主要来源于数字化人才短期培训，包括以“黑马程序员”为品牌的现场培训和以“博学谷”为品牌的线上培训，针对的人群不仅有零基础进入数字化行业就业的学员，也有在职程序员的晋升培训，如架构师、大数据等。针对智能制造等行业的数字化转型升级，如工业机器人、集成电路等，公司已有智能机器人工业软件开发等课程，公司也在持续密切关注行业的发展，并将基于市场对数字化人才的需求变化及业务发展规划进行相关课程的研发及上市。
其他	神州泰岳	300002.SZ	公司人工智能/大数据业务专注于自然语言处理(NLP)及大数据技术等融合应用，汇集基础研究、关键技术、平台建设和行业应用于一身，致力于持续推进人工智能技术在不同行业落地，利用“智脑2.0”“语义工厂”及“中文信息(深度)处理开放创新平台”，有效的将科研成果同产业落地工程相对接，面向公安、政府、气象、环保、金融、服务等行业提供人工智能及大数据解决方案，形成智慧政企、智能客服、智慧园区三条核心业务线。截至2022年6月，在人工智能领域公司已申请专利302件，其中291件为发明专利；已获得授权的专利118件，其中发明专利108件。
其他	当虹科技	688039.SH	公司AIGC相关技术在媒体演播室、智能媒资平台等众多场景上的应用已经较为成熟，可为电视台、新媒体、互联网、泛媒体行业等客户提供通过AI技术进行短视频、长视频的生产/加工等解决方案。包括比如：AI智能剪辑；用AIGC的方式自动生成对应短视频，极大程度上简化了流程提高了效率，保证了视频上线的时效性。
其他	天娱数科	002354.SZ	公司发挥数据、算法、场景优势，利用人脸建模、形象驱动、语音合成、语音识别、语义理解、图像处理等多项AI技术，推出MetaSurfing-元享平台，实现元宇宙人货场的“批量化、短周期、低成本”制作，陆续为AMD、腾讯、百度、外研在线、匹克、沪上阿姨等各领域一线客户提供技术服务。近期，公司升级推出的“MetaSurfing-元享智能云平台”，纳入了AIGC功能模块，并接入了ChatGPT等模型，进一步实现从文本、音频等低密度模态向图像、视频、实时交互等信息密度更高模态的转化，并结合多模态智能算法，完成了AIGC互动直播功能的落地，先后在TikTok跨境电商直播、虚拟主播互动直播等场景中实践应用，并形成产品化解决方案。公司制作的虚拟数字人在语音交互功能中应用了基于AIGC技术的NLP问答技术。



### 3.3.1 昆仑万维：发布“昆仑天工”

- 2022年12月，昆仑万维举行AIGC技术发布会，正式发布“昆仑天工”AIGC全系列算法与模型，并宣布模型开源。“昆仑天工”旗下模型包括天工巧、天工乐府、天工妙笔、天工智码，覆盖图像、音乐、文本、编程等多模态内容生成能力，标志着昆仑万维成为国内第一个全身心投入到AIGC开源社区的公司。
- ✓ **天工巧绘SkyPaint**：采用全球第一款多语言Stable Diffusion分支模型，兼容Stable Diffusion；使用1.5亿中英双语语料，支持中英双语输入实现文字生成图像，是国内为数不多的支持中英双语的文图生成模型。
- ✓ **天工乐府SkyMusic**：国内第一款商用级作曲AI模型，该模型引入全局信息约束、和弦隐式表征，和弦准确率达到92.5%，远超目前行业最佳64.7%；支持中文、英文等31种语种，支持从旋律生成歌词、文案素材生成歌词，且支持全球多曲风。基于昆仑万维旗下StarMaker平台拥有大量优质作品曲库，目前面向全球音乐市场，已在美国、韩国、印尼等国家及全球180多个音视频平台发行了近20首人工智能生成歌曲，成功实现AIGC商业化。
- ✓ **天工妙笔SkyText**：采用生成效果最好的开源GPT中文预训练大模型，针对中文领域构建了千亿级别的高质量数据集，提供包括续写、对话、中英翻译、内容风格生成、推理、诗词对联等功能，与现有大模型的比试中排列前茅。
- ✓ **天工智码SkyCode**：采取全球第一款多语言开源编程大模型，支持多种主流代码语言SkyCode能够生成的代码，不仅包括使用广泛的python、Java、C等，还涵盖了php、go、swift等共计十余种编程语言，使不同语言的使用者都能来体验SkyCode强大的代码生成能力。

图表：昆仑万维AIGC布局历程



图表：天工巧绘、天工妙笔生成的作品

**天工巧绘生成作品：在太空骑马的宇航员**

**天工妙笔生成五言诗**

**写五言诗**

**引导语句:**  
闻道巴山里，春船正好行。都将百年兴，一望九江城。水槛温江口，茅堂石笋西。移船先主庙，洗药浣沙溪。迟日江山丽，春风花草香。泥融飞燕子，沙暖睡鸳鸯。寒食少天气，东风多柳花。小桃知客意，春尽始开花。清风入堂来。

**响应样本:**  
疏雨过庭阴。落花犹在眼，芳草亦忘忧。

### 3.3.1 昆仑万维：AIGC与StarMaker互相助推，协同其他业务发展

- StarX是面向海外的社交娱乐领域平台，其主打产品为StarMaker，最早以K歌为抓手，后切入语音房、直播、休闲游戏等在线娱乐场景。截至2022H1，StarMaker累计在102个国家音乐/音频类畅销榜排名前5，其中84个国家位列第1，累计在100个国家音乐/音频类下载榜排名前5，其中77个国家位列第1。StarX围绕“音乐+社交”的核心发展思路，在AIGC及VR/AR等领域积极探索。
- StarX在AI内容创作领域取得进展，StarMaker与AIGC互相助推发展，输出更多高品质内容。旗下StarX MusicX Lab音乐实验室是公司在“AI赋能”业务发展战略方向，截至2022年H1已完成首批5首完全由AI作曲的歌曲发布，已在Spotify、QQ音乐等海内外180余个音乐平台上线。
- ✓ 从内部协同来看，首先AIGC可以持续高效为StarMaker提供海量音乐作品，丰富平台内容，同时可进一步增强应用的可玩性和功能性，帮助平台内广大音乐爱好者和独立音乐人大幅降低创作门槛，有效缩短创作周期，产出更多优质内容。而StarMaker可以为AIGC提供最为高效的验证链路，助力其算法模型不断迭代优化。其次，由AI作曲的乐曲能够与公司旗下业务如元宇宙、游戏等产生协同效应，能够在提高音乐品质的同时降低使用音乐的成本。
- ✓ 从外部合作来看，AIGC除满足个人用户需求外，还可以为企业用户提供音乐相关解决方案，包括生成背景音乐、伴奏配乐等泛化内容，以及为艺人和虚拟偶像定制化创作符合其调性的音乐作品等服务。StarX MusicX Lab未来将继续依托成熟专业的全链路音乐制作和发行能力，向全球输出高质量的AI创作内容。
- 风险提示：AIGC技术发展不及预期，市场竞争加剧，版号发放不及预期，新游上线延期，国际政治形势变动，估值中枢下移等风险。

图表：StarX MusicX Lab创作的音乐在QQ音乐上线



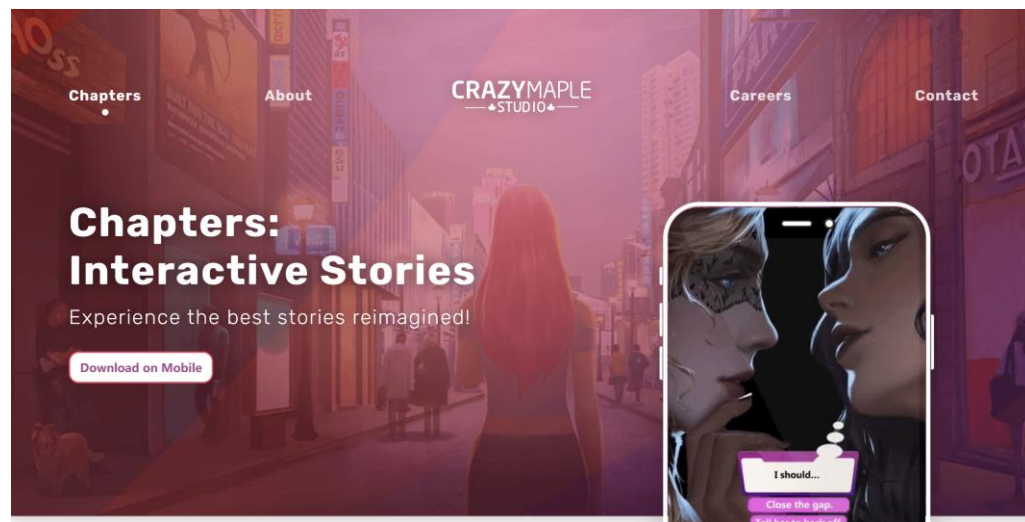
图表：StarMaker官网



## 3.3.2 中文在线、视觉中国：AI赋能文字、图片等内容创作

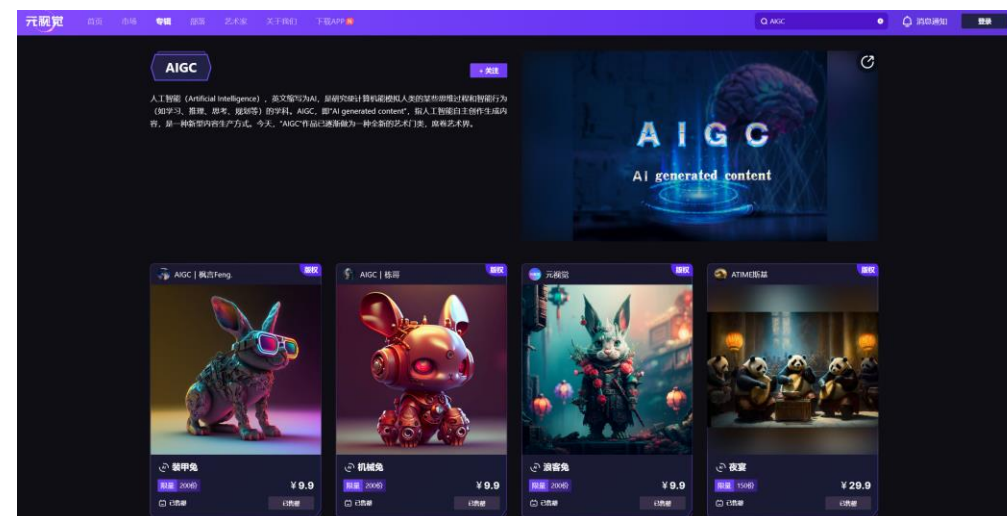
- **中文在线：**拥有丰富中文版权内容，累积数字内容资源510万+，与600余家版权机构合作；已在有声书、互动阅读文生图、文字创作辅助等领域有应用。
  - ✓ **AI主播：**以全球领先的神经网络设计和先进的深度学习合成器为依托，独创情绪、情感控制模块，可实现快速、低成本音色定制，AI主播每天单机生产速度超500万字，生产效率远高于人工主播，录制成本则可节约90%以上。未来，AI主播技术一方面将向“更动听”发展，加强人物区分、情绪细腻；另一方面将向“更全面”进阶，为不同小说品类量身定制主播声音。
  - ✓ **AI绘画：**已在互动式视觉阅读平台Chapters中使用，通过文生图提升Chapters中作品的沉浸感和互动效果。
  - ✓ **AI文学创作：**2022年9月，与澜舟科技在文学创作辅助算法领域进行业务落地合作。AI文学创作功能已在17K小说网上线。
- **风险提示：** AIGC技术发展不及预期，海外市场拓展不及预期，市场竞争加剧，监管政策、IP开发低于或慢于预期等风险。

图表：互动式视觉阅读平台Chapters



- **视觉中国：**拥有优质音图版权内容，可提供4亿张图片、3000万条视频和35万首音乐等版权素材；与百度合作提供AI辅助创作服务、推进版权保护，在AIGC艺术作品落地。
  - ✓ **素材和版权：**与百度旗下的AI绘画平台文心一格在版权保护方面展开合作，双方在政策合规的前提下，为版权交易及保护做出正向引导，携手促进AIGC内容产业合规健康发展。
  - ✓ **AIGC艺术作品：**旗下“元视觉”平台共发行“AIGC”专辑的艺术作品30件。
  - ✓ **AI辅助创作：**在创作者赋能方面与百度合作，双方将为视觉内容创作者提供便捷、优质的AI辅助创作服务。
- **风险提示：** AIGC技术发展不及预期，市场竞争加剧，监管政策，平台上线后表现不及预期等风险。

图表：视觉中国“元视觉”平台的AIGC专辑

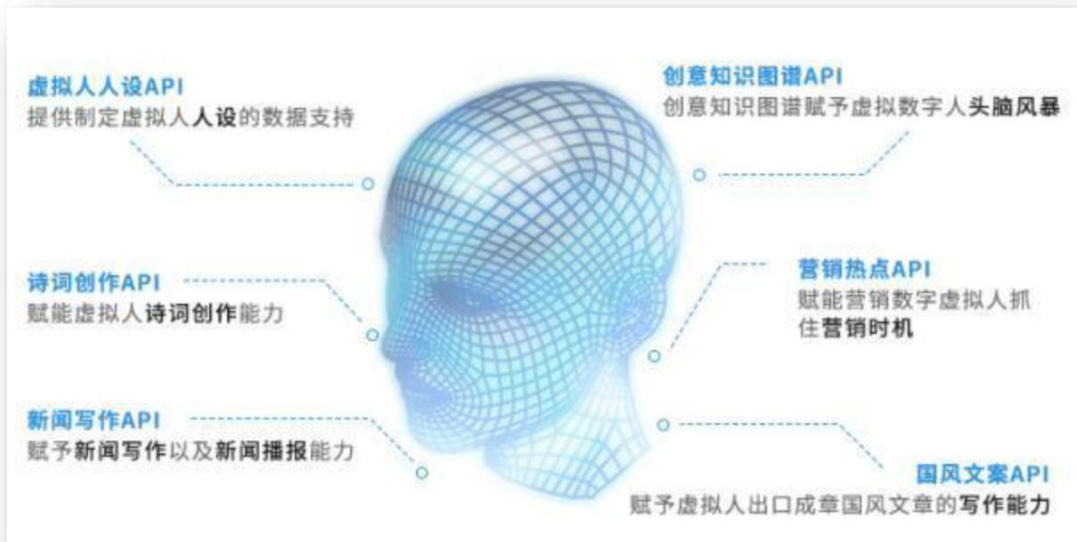




### 3.3.3 蓝色光标：AIGC营销领域应用集大成者

- 蓝色光标集团旗下销博特是基于云端聚焦AI营销场景的多人协同创作平台，通过结合人工智能、统计算法和多维数据库，一键自动化生成策划案、品牌分析报告、消费者洞察等内容。截止2022年底，销博特已累计注册用户超6万，覆盖10多个行业，400多个品类。
- ✓ **虚拟数字人**：2021年9月，国内首位广告公司虚拟创意人蓝零壹在公司内部上线，可在内网与其进行实时互动，并获得海报制作、slogan创策等初级创意内容服务。2022年2月，发布小元宵计划，将平台现有的分析、写作、文案、随笔以及创意等六大产品能力封装成API技术接口为虚拟数字人赋能营销创作能力。
- ✓ **AI绘画**：2022年11月，发布“创意画廊”，用户输入文本即可在6分钟内生成一幅抽象画作。
- ✓ **AIGC营销套件**：2022年12月，销博特发布AIGC“创策图文”营销套件，提供创意、策划、文案、图片等内容一体化智能生成解决方案，通过AI技术助力内容营销实现实时在线、即刻生成个性化内容。
- **风险提示**：AIGC技术发展不及预期，市场竞争加剧，监管政策，平台上线后表现不及预期，人才流失风险，海外业务发展不及预期，估值中枢下移等风险。

图表：6大API赋能虚拟数字人



图表：蓝色光标旗下销博特推出AIGC营销套件

AiG 创策图文	
获得三个软件著作权（创意机、品牌主张、国风文案）；品牌定位支持向量机专利正在受理过程中；“康定斯基”模型的相关专利和软件著作权也正在申请。	
营销场景	营销套件
AI生成创意	创意风暴：通过NLP技术获得启发短语来捕获创意火花，组合相关元之间的关联性进行创意概念推荐。 创意罗盘：用户通过人群和卖点来获得创意启发。
AI生成策略	智能策划：用户输入Brief，策划后台15分钟生成策划案。 人群画像：结合行为心理学，通过社群数据和调研数据一键生成用户画像。
AI生成图片	一键海报：输入关键词或句，基于营销热点生成每日打卡海报； 创意画廊：“康定斯基”一键抽象画平台。
AI生成文本	AI易稿：基于新闻稿模板的辅助写作； 品牌主张：用户输入品牌名和关键词、调性，来写slogan。 国风文案：基于品牌、调性和核心句子来写品牌写真。

### 3.4 国内发展现状：各领域涌现出代表性产业公司

- 国内各领域均有代表性公司，但独立运行的初创公司数量明显少于国外，大部分细分赛道的初创玩家在5家以下。

图表：国内各领域代表性产业公司（数据更新截止到2023年2月）

内容	公司	成立年份	具体与AIGC相关的业务/技术	最新融资进程	最新融资时间	最新融资金额
综合	小冰公司	2020	全球领先的人工智能科技公司，旗下小冰框架是全球承载交互量最大的完备人工智能框架之一，在开放域对话、多模态交互、超级自然语音、神经网络渲染及内容生成领域居于全球领先。小冰产品始终是“人+交互+内容”，包括虚拟人、音频生成、视觉创造、文本创造、虚拟社交、Game AI等。	A+轮	2022年	10亿元人民币
游戏	rct AI	2018	运用人工智能为游戏行业提供完整的解决方案，核心产品有个性化且动态的剧情探索、智能NPC/智能生物体/智能环境“真人”陪玩、智能动画及音效生成技术等。	A3轮	2021年	超千万美元
游戏	启元世界	2017	打造了AI玩家、AI角色、AI设计师、AI指挥官等产品方案，为游戏、虚拟世界、元宇宙提供高质量内容和互动体验。	A轮	2021年	3亿元
游戏	超参数	2019	提供AI bot支持玩家陪玩、多人团队竞技（球球大作战）、非完美信息博弈AI（斗地主、德州、麻将等）等。	B轮	2022年	1亿美元
文字	澜舟科技	2021	认知智能公司，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（包括搜索、生成、翻译、对话等）和针对垂直行业场景的SaaS产品。旗下基于孟子预训练语言模型实现的可控文本生成技术，可以完成营销文案、广告词、例句推荐、研报、小说和科普文章等文本生成任务。	Pre-A轮	2022年	近亿元人民币
文字	聆心智能	2021	基于多模态对话生成系统，专注于针对精神心理的AI驱动的高质量数字疗法，让AI围绕认知、情绪和行为三个维度对用户进行评估和干预，生成千人千面的情绪治疗方案。开发出了中文对话预训练大模型和共情聊天机器人Emohaa。	Pre-A轮	2022年	-
图像	感知阶跃	2020	ZMO.AI提供人工智能模特图片解决方案，通过AI算法生成独一无二的逼真虚拟模特展示服装，能够降低商家宣传成本、提高制作效率、提升电商转化率。	A轮	2022年	800万美元
视频	影谱科技	2018	在视频生成相关领域支持结构化视觉分析、影像自动合成技术、智能视频编辑、视频内容生产等。	D轮	2018年	13.6亿元人民币
音频	灵动音科技	2018	产品包括针对视频生成配乐配乐猫、支持非音乐专业人员创作的口袋音乐、可AI生成歌词的LYRICA、AI作曲软件LAZYCOMPOSER。目前已与国内多家音乐平台厂商达成合作。其音乐标注团队已形成了全球最精确的华语歌曲音乐信息库。	C轮	2021年	2亿元人民币
音频	标贝科技	2016	全球领先的人工智能语音交互服务商，基于AI+SaaS开放平台，为客户提供AI数据服务、技术能力、智能语音交互方案赋能服务，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等采标服务和平台化自研工具能力。	B1轮	2022年	数千万元人民币
虚拟人	倒映有声	2019	通过神经渲染技术快速构建AI数字分身，通过语音+图像生成技术，生成和驱动数字分身的唇形、表情、动作、肢体姿态，创造表情自然、动作流畅、语音充满情感的高拟真度数字分身IP。	Pre-A轮	2021年	数千万元人民币

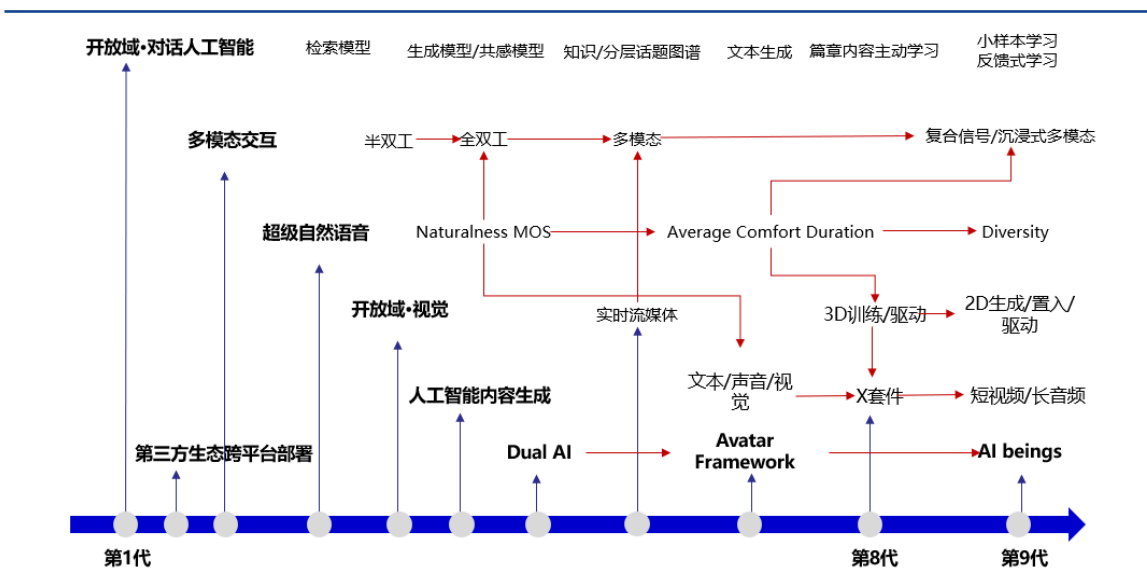
### 3.4.1 小冰公司：综合赛道初创公司，覆盖多场景

- 小冰是全球领先的人工智能科技公司，旗下小冰框架是全球承载交互量最大的完备人工智能框架之一，在开放域对话、多模态交互、超级自然语音、神经网络渲染及内容生成领域全球领先。小冰前身是微软人工智能小冰团队，2020年7月微软拆分创立于中国的微软小冰，目前已获数十亿元融资。
- 小冰产品始终是“人+交互+内容”：包括虚拟人、音频生成、视觉创造（毕业作品集《或然世界》、为国家纺织品开发中心、万事利等数百家机构提供了图案和纹样设计）、文本创造、虚拟社交、Game AI等。商业客户已覆盖金融、零售、体育等十多个垂直领域，并提出了以“人力”的逻辑进行商业报价的虚拟人商业模式。
- ✓ **虚拟人 (AI being)**：小冰框架已孵化出数以千万计的超级自然虚拟人 (AI being)，既有somebody instance（专业个体），如世博会参展画家夏语冰、空中技巧国家队AI裁判与教练系统观君；也有nobody instance（普通个体），比如虚拟男友/女友；还有许多在垂直场景的AI being，如金融摘要撰稿人万小冰等。
- ✓ **音频生成**：推出在线歌曲生成平台与歌手歌声合成软件X studio，提供业内最优质自然、多种各具特色的人工智能主播声音，辅助创作者将文字内容高品质快速音频化。
- ✓ **文本生成**：2017年人工智能诗人小冰创作推出诗集《阳光失了玻璃窗》。
- ✓ **虚拟社交**：2021年推出了全球首个人与AI融合社交平台APP“小冰岛”，在该平台中人类用户可以创造各种AI being，并形成共同生活的社交网络。
- ✓ **Game AI**：2022年5月，宣布成立ICE Gamer游戏工作室，加速游戏领域布局，并推动AI being NPC的发展。AI being NPC是基于人工智能小冰框架创建、驱动的新交互主体，具备多样性、自主人格、自我意识、高自由度等特点，能够以文字、声音、影像等多模态方式与玩家真实自然地交互。

图表：小冰公司融资情况

序号	日期	融资轮次	估值金额	融资金额	投资机构
1	2022年11月	A+轮	23.08亿美元	10亿元	-
2	2021年7月	A轮	10亿美元	-	领投机构：高瓴资本 跟投机构：北极光创投，IDG资本，纪源资本，网易资本，五源资本，Neumann Advisors
3	2020年11月	Pre-A轮	-	未披露	北极光创投，网易资本

图表：以NLP为基础，推动完成迭代发展的人工智能小冰框架

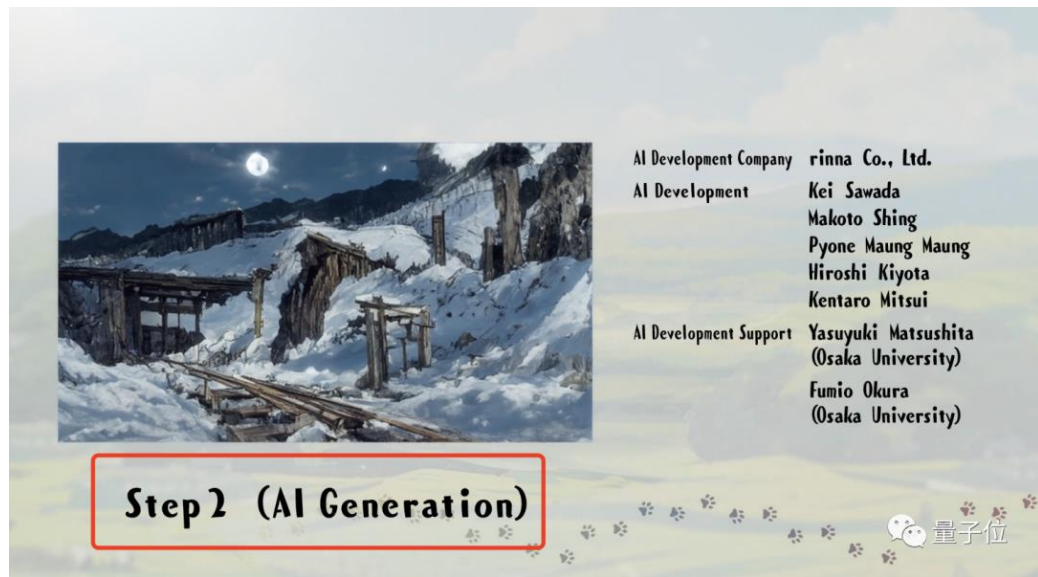




### 3.4.1 小冰公司：AIGC动画制作拉开帷幕

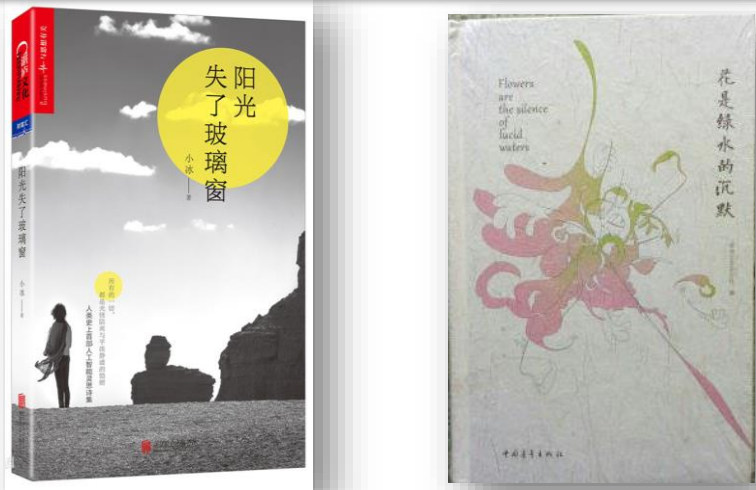
- 首支AIGC动画短片公开，动画制作打开新篇章。2023年1月，Netflix宣布，其与小冰公司日本分部（rinna）、WIT STUDIO共同创作的首支AIGC动画短片《犬与少年》已于当日正式公开，这是Netflix动画创作者计划的第一支作品，采用rinna制作的AI辅助背景，由牧原亮太郎负责导演，用手绘Layout上色稿，并将其提交至AI，生成细节并优化背景，再对AI生成的背景图进行人力修正，通过人工智能技术绘制完整动画场景，为动画制作揭开新的未来。

图表：首支AIGC动画短片《犬与少年》



# 3.4.1 小冰公司：产品示例

小冰AI创作的诗集



小冰岛



小冰框架内的AI being



歌曲创作平台X studio



**投资建议：**AIGC的快速发展源于数据、算法、算力的共振。在此基础上，AIGC的出圈源于模型商业化及开源带来的产品化浪潮，及ChatGPT在通用人工智能领域投射的曙光带来的震撼。AIGC目前在社交、广告营销、内容创作、游戏等领域均已有应用，并开启商业化变现。随着算法迭代、算力提升，AIGC将开启新的内容生产力革命，为传媒行业发展提供新动力。基于此，维持行业“推荐”评级。建议关注三类公司：

一，拥有自有算法及模型的公司，建议关注**昆仑万维**（拥有“昆仑天工”AIGC全系列算法与模型，旗下StarX MusicX Lab音乐实验室已实现AIGC音乐商业化，自有音乐社交平台、浏览器等AIGC落地业务场景）、**神州泰岳**（专注于自然语言处理及大数据技术等融合应用，持续推进人工智能技术在不同行业落地，形成智慧政企、智能客服、智慧园区三条核心业务线）；

二，拥有海量内容及版权储备的公司，建议关注**视觉中国**（拥有图片视频版权库、AIGC创作工具有望辅助图片创作生成及赋能旗下元视觉平台AI数字艺术品生成）、**中文在线**（已推出AI主播、AI绘画和AI文字创作功能）；

三，相关应用领域龙头公司，重点推荐**游戏板块**，当前板块估值处于低位，AIGC有望渗透游戏研发发行各个环节，相关标的**三七互娱**、**吉比特**、**完美世界**；**影视板块**推荐**芒果超媒**（团队获全球顶级算法大赛世界冠军，已连续举办三届“马栏山杯”算法大赛）；**营销板块**关注**蓝色光标**（旗下销博特平台发布了AIGC“创策图文”营销套件）。



- 技术发展演进不及预期
- 商业化进程不及预期
- 企业技术管理能力建设不足风险
- 企业内容审核能力不足风险
- 版权保护风险
- 新技术增加监管难度风险
- 技术滥用风险
- 核心人才流失风险
- 创作伦理风险
- 法律政策监管风险
- 估值中枢下移风险

## 传媒小组介绍

姚蕾，传媒教育行业首席分析师，同济大学本科，香港大学硕士，从业8年，曾获新财富、水晶球、保险资产管理业最佳分析师。主要研究方向为游戏、视频、营销、潮玩、出版、教育、体育等赛道。

方博云，传媒教育行业分析师，西南财经大学本科，上海财经大学硕士，从业5年，主要研究方向为影视、潮玩、营销、泛娱乐等赛道。

谭瑞峤，传媒教育行业分析师，厦门大学本科，中央财经大学、哥伦比亚大学硕士，从业4年，主要研究方向为游戏、教育、出版、直播电商等赛道。

杨牧笛，传媒教育行业研究助理，上海财经大学本科，福特汉姆大学硕士，从业1年，主要研究方向为游戏赛道。

## 分析师承诺

姚蕾，方博云，本报告中的分析师均具有中国证券业协会授予的证券投资咨询执业资格并注册为证券分析师，以勤勉的职业态度，独立，客观的出具本报告。本报告清晰准确的反映了分析师本人的研究观点。分析师本人不曾因，不因，也将不会因本报告中的具体推荐意见或观点而直接或间接收取到任何形式的补偿。

## 国海证券投资评级标准

### 行业投资评级

推荐：行业基本面向好，行业指数领先沪深300指数；

中性：行业基本面稳定，行业指数跟随沪深300指数；

回避：行业基本面向淡，行业指数落后沪深300指数。

### 股票投资评级

买入：相对沪深300 指数涨幅20%以上；

增持：相对沪深300 指数涨幅介于10%~20%之间；

中性：相对沪深300 指数涨幅介于-10%~10%之间；

卖出：相对沪深300 指数跌幅10%以上。

## 免责声明

本报告的风险等级定级为R3，仅供符合国海证券股份有限公司（简称“本公司”）投资者适当性管理要求的客户（简称“客户”）使用。本公司不会因接收人收到本报告而视其为客户。客户及/或投资者应当认识到有关本报告的短信提示、电话推荐等只是研究观点的简要沟通，需以本公司的完整报告为准，本公司接受客户的后续问询。

本公司具有中国证监会许可的证券投资咨询业务资格。本报告中的信息均来源于公开资料及合法获得的相关内部外部报告资料，本公司对这些信息的准确性及完整性不作任何保证，不保证其中的信息已做最新变更，也不保证相关的建议不会发生任何变更。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、价值及投资收入可能会波动。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。报告中的内容和意见仅供参考，在任何情况下，本报告中所表达的意见并不构成对所述证券买卖的出价和征价。本公司及其本公司员工对使用本报告及其内容所引发的任何直接或间接损失概不负责。本公司或关联机构可能会持有报告中所提到的公司所发行的证券头寸并进行交易，还可能为这些公司提供或争取提供投资银行、财务顾问或者金融产品等服务。本公司在知晓范围内依法合规地履行披露义务。

## 风险提示

市场有风险，投资需谨慎。投资者不应将本报告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，投资者务必向本公司或其他专业人士咨询并谨慎决策。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。

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。本报告不构成本公司向该机构之客户提供的投资建议。

任何形式的分享证券投资收益或者分担证券投资损失的书面或口头承诺均为无效。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。

## 郑重声明

本报告版权归国海证券所有。未经本公司的明确书面特别授权或协议约定，除法律规定的情况外，任何人不得对本报告的任何内容进行发布、复制、编辑、改编、转载、播放、展示或以其他方式非法使用本报告的部分或者全部内容，否则均构成对本公司版权的侵害，本公司有权依法追究其法律责任。



国海证券 · 研究所 · 传媒研究团队

# 心怀家国，洞悉四海



## 国海研究上海

上海市黄浦区福佑路8号人保寿险大厦7F

邮编：200010

电话：021-60338252

## 国海研究深圳

深圳市福田区竹子林四路光大银行大厦28F

邮编：518041

电话：0755-83706353

## 国海研究北京

北京市海淀区西直门外大街168号腾达大厦25F

邮编：100044

电话：010-88576597