

区块链

AIGC——Web3 时代的生产力工具

在我们《元宇宙（七）：虚拟人的“灵魂”是什么？》的报告中，提到了虚拟人的灵魂来自于 AI 自主创作内容。同时，元宇宙更倾向于开放世界游戏，AIGC 内容生成方式的创新将成为 Web3 时代全新的生产力工具。

AIGC (AI-Generated Content 人工智能生成内容) 代表新一轮范式转移的开始。近期，硅谷的众多一线 VC 们开始将目光瞄准 AI 初创公司，尤其是生成式 AI 艺术这一领域。今年有两家独角兽 Stability 和 Jasper 均获得了超过一亿美元的融资，估值突破十亿美元。AIGC 赛道火爆不仅得益于技术进步、商业应用广泛和需求增长，还归功于该赛道还处于早期。虽然大型科技公司捕获了大量价值，初创企业仍有机会突破。

AIGC 将是 Web3 时代的生产力工具。当我们迈入 Web3.0 时代，人工智能、关联数据和语义网络构建，形成人与网络的全新链接，内容消费需求飞速增长。UGC\PGC 这样的内容生成方式将难以匹配扩张的需求。AIGC 将是新的元宇宙内容生成解决方案。AIGC 的生成利用人工智能学习知识图谱、自动生成，在内容的创作为人类提供协助或是完全由 AI 产生内容。不仅能帮助提高内容生成的效率，还能提高内容的多样性。随着 NLP (Natural Language Processing, 自然语言处理) 技术和扩散模型 (Diffusion Model) 的发展，AI 不再仅作为内容创造的辅助工具，创造生成内容成为了可能。由此，将来文字生成、图片绘制、视频剪辑、游戏内容生成皆可由 AI 替代。

AIGC 技术主要涉及两个方面：自然语言处理 NLP 和 AIGC 生成算法。自然语言处理是实现人与计算机之间如何通过自然语言进行交互的手段。AIGC 生成算法主流的有生成对抗网络 GAN 和扩散模型。扩散模型已经拥有了成为下一代图像生成模型的代表潜力。它具有精度更高、可扩展性和并行性，无论是质量还是效率均有所提升，其快速发展成为 AIGC 增长的拐点性因素。同时，在机器学习的过程中，需要通过大量的训练来实现更准确的结果，目前以英伟达 A100 为主，对于底层算力需求将有飞速增长。

AIGC 在文字、图像、音频、游戏和代码生成中商业模式渐显。近期我们研究了国内外数十家 AIGC 相关企业，尤其在一些具备高重复性的任务、对于精度要求并不那么高的领域应用已逐步成熟，并在探索商业模式中。目前图片生产、文字生成较为常见，这类 AIGC 服务大多数时候会以提供 SaaS 服务的形式变现。

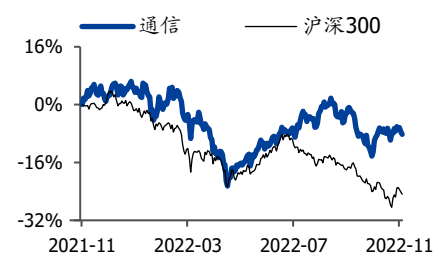
AIGC 未来发展核心：大模型、大数据与大算力。结合自然语言的大模型与数据集已成为 AIGC 发展的软件基础，OpenAI 的 Clip 模型基于 4 亿组高质量的英文图文对应数据训练而成；算力即权力将在 AIGC 数字时代更加凸显，Stable Diffusion 目前依赖于 4000 个英伟达 A100 的 GPU 集群，运营成本超 5000 万美金。为了让功能更加精确，未来还将更多地基于语种去开发垂直类的应用，便于更有目的性地为特定功能进行训练。

AIGC 投资框架：软硬件与数据集。生成算法、NLP 与算力决定 AIGC 能否运行，而高质量的数据集决定了 AIGC 质量与商业模式。软件层面主要包括自然语言处理技术：谷歌、微软、科大讯飞、拓尔思；AIGC 生成算法模型及数据集：英伟达、Meta、百度、蓝色光标、视觉中国、昆仑万维。算力层包括：澜起科技、中兴通讯、新易盛、天孚通信、宝信软件、中际旭创等。

风险提示：技术创新不及预期：AIGC 的技术发展不及预期。以及底层硬件技术，如：超级计算机、算力的发展不及预期。**政策监管风险：**目前 AIGC 还处于相对早期，后续是否会出台 AIGC 作品相关知识版权或其他法律监管条款尚不明确。

增持（维持）

行业走势



作者

分析师 宋嘉吉

执业证书编号：S0680519010002

邮箱：songjiaji@gszq.com

分析师 金郁欣

执业证书编号：S0680521070002

邮箱：jinyuxin@gszq.com

相关研究

- 1、《通信：2022Q3 综述：受益“通信+”与汇兑，利润持续修复》2022-11-10
- 2、《通信：通信的政策新方向一大安全、虚拟现实与数字会展》2022-11-06
- 3、《区块链：ZK 证明计算：算力硬件的新征途？》2022-11-04



内容目录

1. 2022: AIGC 崛起之年	3
2. AIGC 是什么?	5
2.1 AIGC 发展简史	6
2.2 技术	6
自然语言处理技术 NLP	6
AIGC 生成模型	9
2.3 当我们开始用 AIGC——商业模式的探索	14
3. AIGC 的未来发展趋势	18
3.1 面临的挑战	19
3.2 未来的发展方向	20
4. 投资策略: AIGC 的软硬件与数据集	20
风险提示	21

图表目录

图表 1: 红框为网站头条文章位置, 为 AI 创作	3
图表 2: 官网展示应用场景	4
图表 3: AI 生成的乔布斯访谈博客	4
图表 4: 《幻觉东京》AI 重制版	5
图表 5: 内容生成的四阶段	5
图表 6: AIGC 发展历程	6
图表 7: 自然语言理解三大发展阶段	7
图表 8: 自然语言生成 6 步骤	7
图表 9: 微软 Apache Spark 功能	8
图表 10: Transformer 与其他神经网络英德翻译质量对比	9
图表 11: Transformer 与其他神经网络英法翻译质量对比	9
图表 12: Transformer 模型结构	9
图表 13: GAN 的两个组成	10
图表 14: GAN 架构	10
图表 15: DALL-E2 可以根据文字描述生成图片	11
图表 16: 扩散模型训练原理	12
图表 17: 扩散模型与其他模型比较	12
图表 18: 正态分布函数叠加	13
图表 19: DALL-E2 工作过程举例	13
图表 20: Jasper 的应用场景	14
图表 21: 通过 AIGC 撰写的 AIGC 报告	15
图表 22: 通过 AIGC 生成的图片	15
图表 23: MidJourney 用户案例	16
图表 24: Phenaki2 分钟视频生成案例	17
图表 25: Delysium 中的虚拟人玩家	18
图表 26: GitHub Copilot	18
图表 27: AIGC 生成的图片在细节上效果较差	19
图表 28: AIGC 投资标的	21

1.2022: AIGC 崛起之年

近期，硅谷的众多一线 VC 们开始将目光瞄准 AI 初创公司，尤其是生成式 AI 艺术这一领域。今年 9 月 23 日，红杉美国官网发表了一篇名为《生成式 AI：一个创造性的新世界》的文章，认为 AIGC（AI-Generated Content 人工智能生成内容）会代表新一轮范式转移的开始。

2022 年 10 月，英国开源人工智能公司 Stability AI 宣布获得 1.01 亿美元融资，估值高达 10 亿美元，跻身独角兽行列，由 Coatue、Lightspeed Venture Partners 和 O'Shaughnessy Ventures LLC 参与投资。Stability AI 今年发布了 Stable Diffusion 的模型，主要用于根据用户输入的文字描述自动生成图像。Stable Diffusion 的诞生让 AI 绘画这个领域愈发火爆。最近，巴比特正式对外宣布，全面拥抱 AIGC，开始规模化采用 AI 配图，其中头条图片，全面由 AI 创作。包括但不限于巴比特网站和 APP，微信公众号、百家号、网易号等自媒体平台，以及微博等社交媒体账号。

图表 1：红框为网站头条文章位置，为 AI 创作



资料来源：巴比特官网，国盛证券研究所

除了绘画以外，文字、音频、视频均可通过 AI 来生成。

- 文字：以 Jasper 为例，以 AI 文字生成成为主打产品，通过其文字生成功能，用户可以生成 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。截止 2021 年，Jasper 已拥有超过 70000 位客户，并创造了 4000 万美元的收入。

图表 2: 官网展示应用场景



资料来源: Jasper, 国盛证券研究所

- 音频: 以 Podcast.ai 为例, 作为一个由 AI 生成的博客, 每周都会探讨一个话题。在第一期节目中, 其通过乔布斯的传记和收集网络上关于他的所有录音, Play.ht 的语言模型大量训练, 最终生成了一段假 Joe Rogan 采访乔布斯的播客内容。

图表 3: AI 生成的乔布斯访谈博客



资料来源: Podcast.ai 官网, 国盛证券研究所

- 视频: 目前的 AI 技术不仅可以生成图片, 也能够生成序列帧, 如: 《幻觉东京》。经过 160 小时, 完成 3 万多张独立插画, 再进行手动微调。虽然目前还只是在原脚本和视频的基础上, 通过 AI 逐帧完成图片生成的, 但看到了 AIGC 参与到视频创作中的可能。而在诸多垂直类应用中, 如体育、财经等, 已经可以通过文字直接生成相应的短视频, 如果配上虚拟人则可以实现自动播报。相比于单一的虚拟人读稿, 基于 AIGC 生产的内容在镜头转换、表情动作结合方面更加逼真。

图表 4: 《幻觉东京》AI 重制版



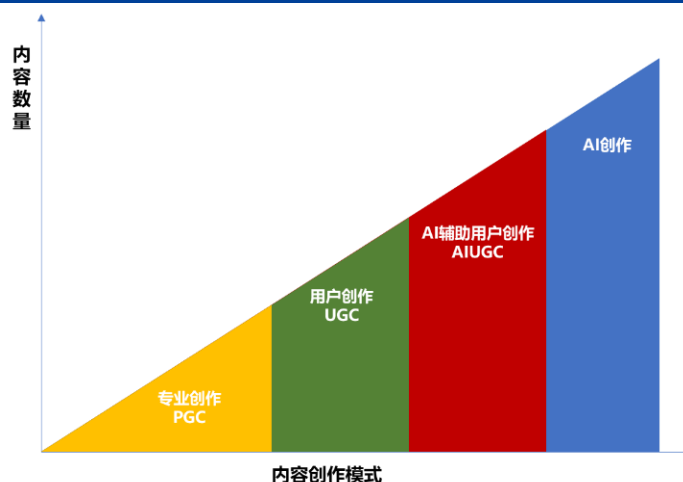
资料来源: 公开资料, 国盛证券研究所

随着 NLP (Natural Language Processing, 自然语言处理) 技术和扩散模型 (Diffusion Model) 的发展, AI 创造生成内容成为了可能。此前, 内容生成主要运用 GAN (Generative Adversarial Network, 生成对抗网络) 来实现的, GAN 不同于扩散模型依赖于超大规模语言模型, 因此难以实现通过文字的描述, 自主理解内容并创造出图像、视频等。近年来, 随着扩散模型的成熟, 生成方式更接近于人脑的联想, AIGC 完成了内容创造辅助工具到内容创作主体的角色转变。

2. AIGC 是什么?

AIGC 是通过人工智能技术自动生成内容的生产方式。从 Web 1.0 的单向信息传递的“只读”模式到 Web 2.0 的人与人通过网络双向沟通交流的“交互”模式, 内容的需求在不断增加。为了满足这一需求, 同时也因为互联网的发展, 内容的生成从单一的 PGC 演变到了现在的 UGC 并占据了主要市场。Youtube、Instagram、抖音、快手、B 站上有大量的内容来自于 UGC 创作者。当我们迈入 Web3.0 时代, 人工智能、关联数据和语义网络构建, 形成人与机器网络的全面链接, 内容消费需求飞速增长, UGC\PGC 这样的内容生成方式将难以匹配扩张的需求。

图表 5: 内容生成的四阶段



资料来源: A16Z, 国盛证券研究所

我们认为，AIGC 将是 Web3 时代全新的内容生成工具，同样，将对现有的短视频、游戏及广告行业带来巨大的影响。AIGC 的生成利用人工智能学习知识图谱、自动生成，在内容的创作为人类提供协助或是完全由 AI 产生内容。不仅能帮助提高内容生成的效率，还能提高内容的多样性。

2.1 AIGC 发展简史

AIGC 的发展可以大致分为以下三个阶段：

- 早期萌芽阶段：20 世纪 50 年代—90 年代中期，受限于科技水平，AIGC 仅限于小范围实验
- 沉积积累阶段：20 世纪 90 年代中期—21 世纪 10 年代中期，AIGC 从实验向实用转变，受限于算法，无法直接进行内容生成
- 快速发展阶段：21 世纪 10 年代中期—现在，深度学习算法不断迭代，AI 生成内容种类多样丰富且效果逼真

近年来，AIGC 的发展迅速，从原来作为边缘侧服务于企业、机构的角色变为了现在 C 端零基础用户都可以使用的创作工具。开发侧重点上，AIGC 也从原先用于翻译、语音合成以及重复性工作转变为了更注重应用层面，用户能够便捷操作的方向。

图表 6：AIGC 发展历程



资料来源：《AIGC 白皮书》，国盛证券研究所

2.2 技术

随着 NLP（Natural Language Processing，自然语言处理）技术和扩散模型（Diffusion Model）的发展，AI 不再仅作为内容创造的辅助工具，创造生成内容成为了可能。

自然语言处理技术 NLP

自然语言处理是实现人与计算机之间如何通过自然语言进行交互的手段。融合了语言学、计算机科学、数学，使得计算机可以理解自然语言，提取信息并自动翻译、分析和处理。

在自然语言处理技术发展之前,人类只能通过一些固定模式的指令来与计算机进行沟通,这对于人工智能的发展是一个重大的突破。

自然语言处理最早可以追溯到 1950 年,图灵发表论文“计算机与智能”,提出“图灵测试”的概念作为判断智能的条件。这一测试包含了自动语意翻译和自然语言生成。

自然语言处理技术可以分为两个核心任务:

- 自然语言理解 NLU: 希望计算机能够和人一样,具备正常人的语言理解能力。过去,计算机只能处理结构化的数据, NLU 使得计算机能够识别和提取语言中的意图来实现对于自然语言的理解。由于自然语言的多样性、歧义性、知识依赖性和上下文,计算机在理解上有很多难点,所以 NLU 至今还远不如人类的表现。
自然语言理解跟整个人工智能的发展历史类似,一共经历了 3 次迭代: 基于规则的方法、基于统计的方法和基于深度学习的方法。

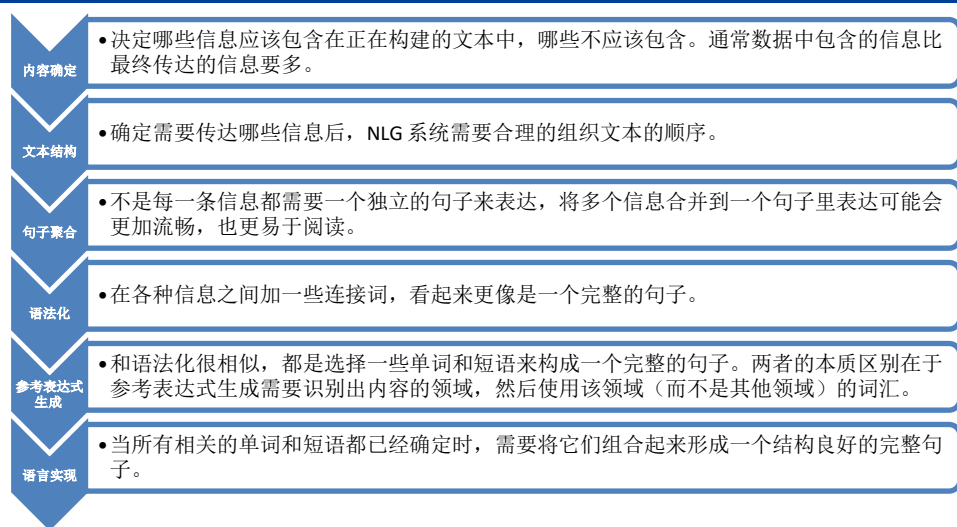
图表 7: 自然语言理解三大发展阶段



资料来源: easyai, 国盛证券研究所

- 自然语言生成 NLG: 将非语言格式的数据转换成人类可以理解的语言格式,如文章、报告等。NLG 的发展经历了三个阶段,从早期的简单的数据合并到模板驱动模式再到现在的高级 NLG,使得计算机能够像人类一样理解意图,考虑上下文,并将结果呈现在用户可以轻松阅读和理解的叙述中。自然语言生成可以分为以下六个步骤: 内容确定、文本结构、句子聚合、语法化、参考表达式生成和语言实现。

图表 8: 自然语言生成 6 步骤

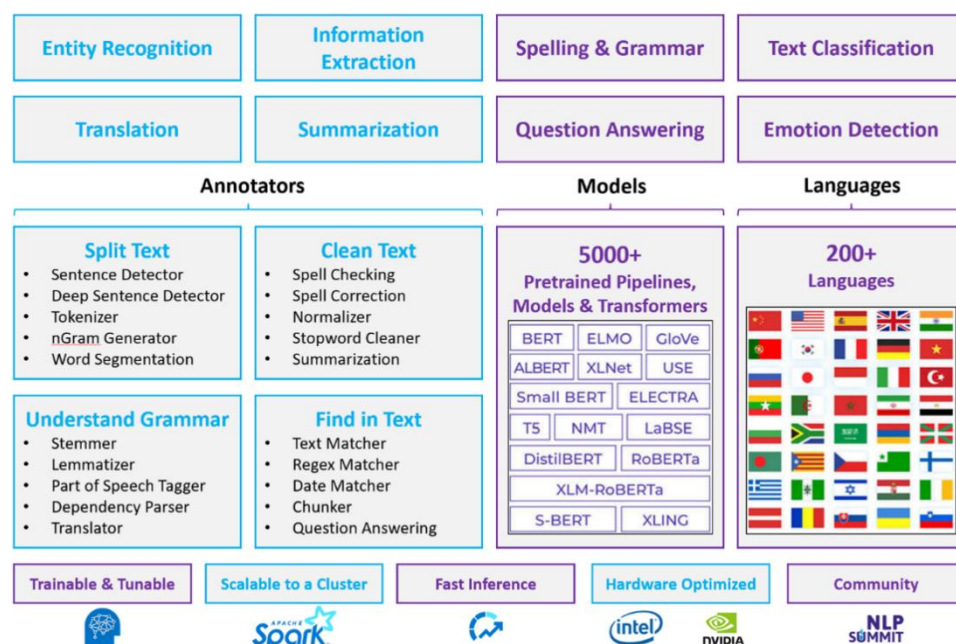


资料来源: easyai, 国盛证券研究所

NLP 主要被应用在四个方面：

- 情感分析：互联网上存在大量的信息，表达的内容都是多种多样的，但抒发的感情大致可以分为正面和负面的，可以被用来快速了解用户的感情情况。
- 聊天机器人：近年来，智能家居的发展和普及使得聊天机器人的价值扩大。
- 语音识别：微信中可以通过语音进行输入或直接将语音转化为文字，汽车导航可以直接说目的地，大大提升了便利性。
- 机器翻译：机器翻译的准确率在近年大幅提高，youtube 和 netflix 甚至可以做到视频机器翻译。

图表 9：微软 Apache Spark 功能



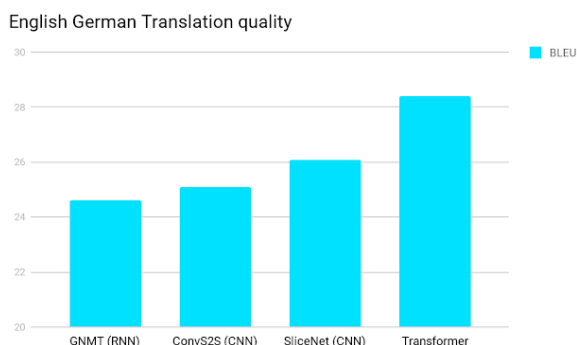
资料来源：微软官网，国盛证券研究所

商业上，NLP 主要被应用在一下领域：

- 用于处理财务、医疗保健、零售、政府和其他部门手写或机器建立档案
- 文字处理工作，如：名称实体辨识（NER）、分类、摘要和关联抽取。这能将抽取、识别和分析文档资讯的流程自动化。
- 语意搜寻和资讯抽取和知识图表建立
- 跨零售、财务、旅游和其他产业客户的交互 AI 系统等。

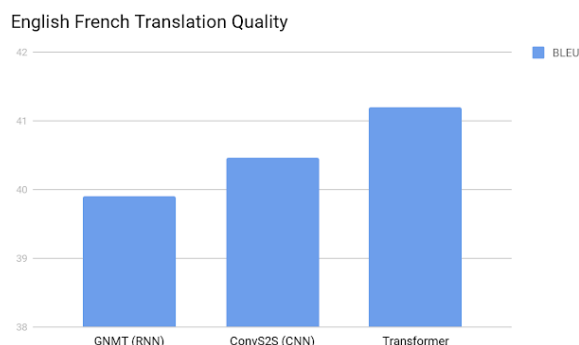
神经网络，尤其是循环神经网络(RNN)是当前 NLP 的主要方法的核心。其中，2017 年由 Google 开发的 Transformer 模型现已逐步取代长短期记忆（LSTM）等 RNN 模型成为了 NLP 问题的首选模型。Transformer 的并行化优势允许其在更大的数据集上进行训练。这也促成了 BERT、GPT 等预训练模型的发展。这些系统使用了维基百科、Common Crawl 等大型语料库进行训练，并可以针对特定任务进行微调。

图表 10: Transformer 与其他神经网络英德翻译质量对比



资料来源: Google Research, 国盛证券研究所

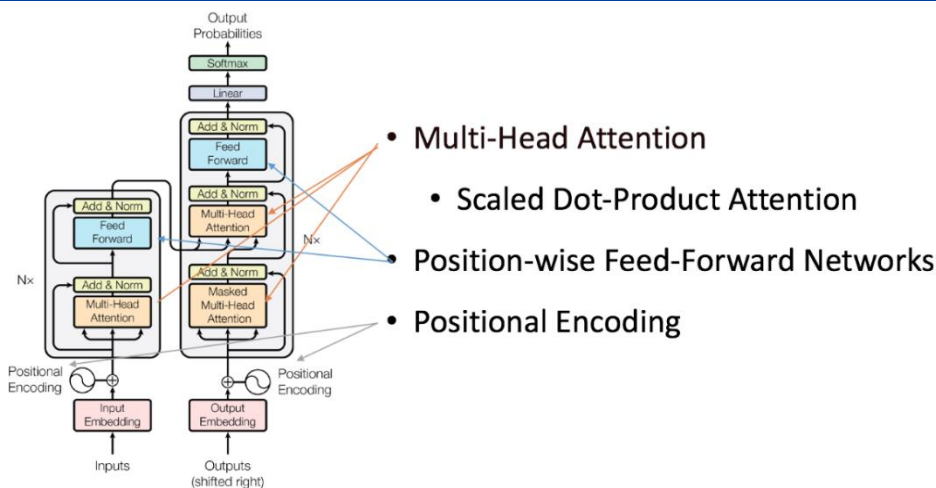
图表 11: Transformer 与其他神经网络英法翻译质量对比



资料来源: Google Research, 国盛证券研究所

Transformer 模型是一种采用自注意力机制的深度学习模型,这一机制可以按输入数据各部分重要性的不同而分配不同的权重。除了 NLP 以外,也被用于计算机视觉领域。与循环神经网络 (RNN) 一样,Transformer 模型旨在处理自然语言等顺序输入数据,可应用于翻译、文本摘要等任务。而与 RNN 不同的是,Transformer 模型能够一次性处理所有输入数据。注意力机制可以为输入序列中的任意位置提供上下文。如果输入数据是自然语言,则 Transformer 不必像 RNN 一样一次只处理一个单词,这种架构允许更多的并行计算,并以此减少训练时间。

图表 12: Transformer 模型结构



资料来源: Google Research, 国盛证券研究所

AIGC 生成模型

近年来, AIGC 的快速发展归功于生成算法领域的技术积累,其中包含了:生成对抗网络 (GAN)、变微分自动编码器 (VAE)、标准化流模型 (NFs)、自回归模型 (AR)、能量模型和扩散模型 (Diffusion Model)。可以看到,大模型、大数据、大算力是未来的发展趋势。我们认为,算法模型的突破是近年来 AIGC 得以快速突破的催化剂,下面将展开介绍一下两个非常常用的模型,分别是生成对抗网络和扩散模型。

- 生成对抗网络 GAN (Generative Adversarial Networks)

2014 年, Ian J. Goodfellow 提出了 GAN, 是一种深度神经网络架构, 由一个生成网络和一个判别网络组成。生成网络产生“假”数据, 并试图欺骗判别网络; 判别网络对生成数据进行真伪鉴别, 试图正确识别所有“假”数据。在训练迭代的过程中, 两个网络持续地进化 and 对抗, 直到达到平衡状态, 判别网络无法再识别“假”数据, 训练结束。

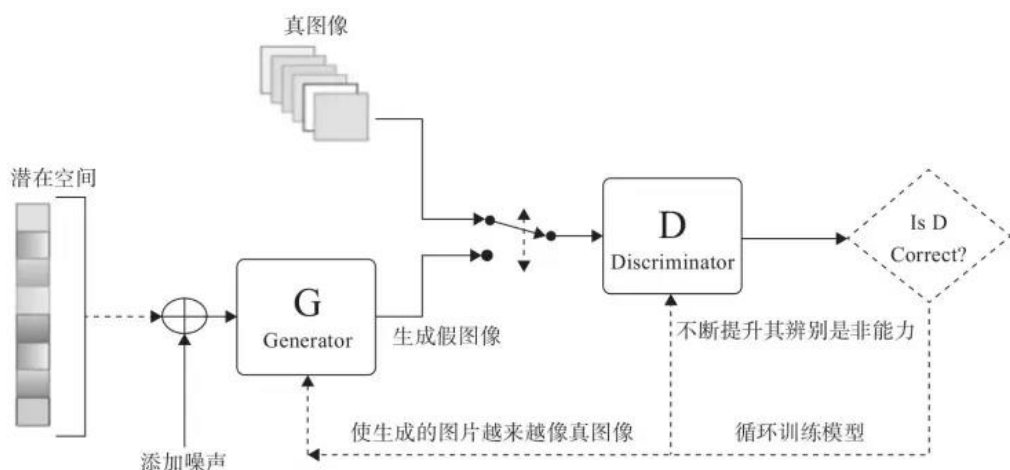
图表 13: GAN 的两个组成



资料来源: EasyAI, 国盛证券研究所

GAN 被广泛应用于广告、游戏、娱乐、媒体、制药等行业, 可以用来创造虚构的人物、场景, 模拟人脸老化, 图像风格变换, 以及产生化学分子式等等。

图表 14: GAN 架构



资料来源: 公开资料, 国盛证券研究所

其优点在于:

- 能更好建模数据分布
- 无需利用马尔科夫链反复采样, 无需在学习过程中进行推断, 没有复杂的变分下界, 避开近似计算棘手的概率的难题

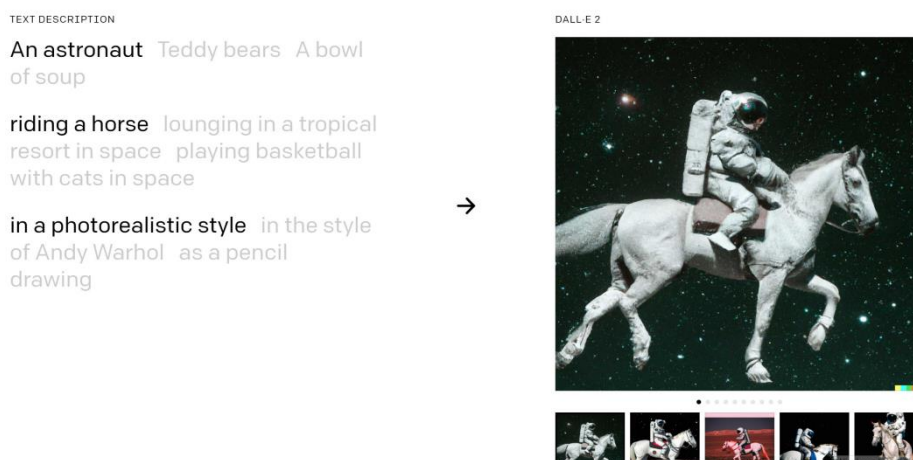
缺点:

- 难训练, 不稳定。生成器和判别器之间需要很好的同步, 但是在实际训练中很容易判别器收敛, 生成器发散。两者的训练需要精心的设计。

- 模式缺失（Mode Collapse）问题。GANs 的学习过程可能出现模式缺失，生成器开始退化，总是生成同样的样本点，无法继续学习。
- 扩散模型 Diffusion Model
扩散模型是一种新型的生成模型，可生成各种高分辨率图像。在 OpenAI, Nvidia 和 Google 设法训练大模型之后，它们已经引起了很多关注。基于扩散模型的示例架构包括 GLIDE, DALL-E-2, Imagen 和完全开源的稳定扩散。扩散模型已经拥有了成为下一代图像生成模型的代表潜力。以 DALL-E 为例，能够直接通过文本描述生成图像，让计算机也拥有了人的创造力。

图表 15: DALL-E2 可以根据文字描述生成图片

DALL-E 2 can create original, realistic images and art from a text description. It can combine concepts, attributes, and styles.



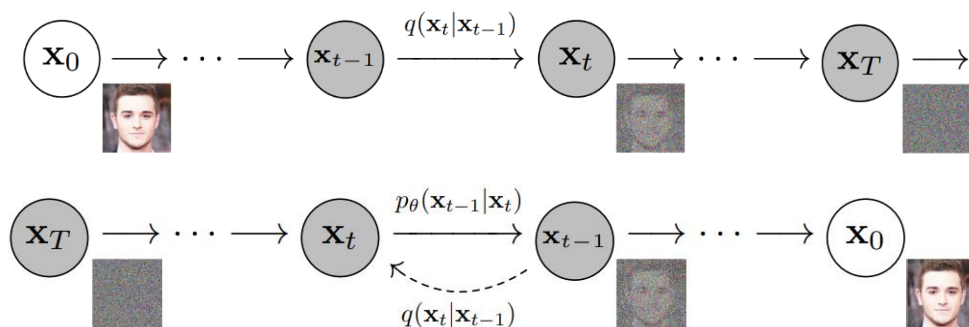
资料来源: OpenAI 官网, 国盛证券研究所

扩散模型的生成逻辑相比其他的模型更接近人的思维模式，也是为什么近期 AIGC 拥有了开放性的创造力。本质上，扩散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。训练后，我们可以通过简单地将随机采样的噪声传递给学习的去噪过程来生成数据。

从下图可以看到，扩散模型是一个潜在变量（latent variable）模型，通过马尔科夫链映射到潜在空间。马尔可夫链是状态空间中经过从一个状态到另一个状态的转换的随机过程，下一状态的概率分布只由当前状态决定。在这一过程中逐步添加高斯噪声来获得近似的后验概率 $q(x_t | x_{t-1})$ ，其中 $x_1, \dots, x_t$ 均是潜在变量，并且它们的维度与原图 x_0 一致。

从上图我们可以看到，图片 x_0 最终会变为纯高斯噪声的图片 x_t 。而训练扩散模型的目标则是反向这一过程，也就是训练图中所示的转移概率 $p_\theta(x_{t-1} | x_t)$ 。通过沿着这条线向后遍历，我们可以生成新的数据。

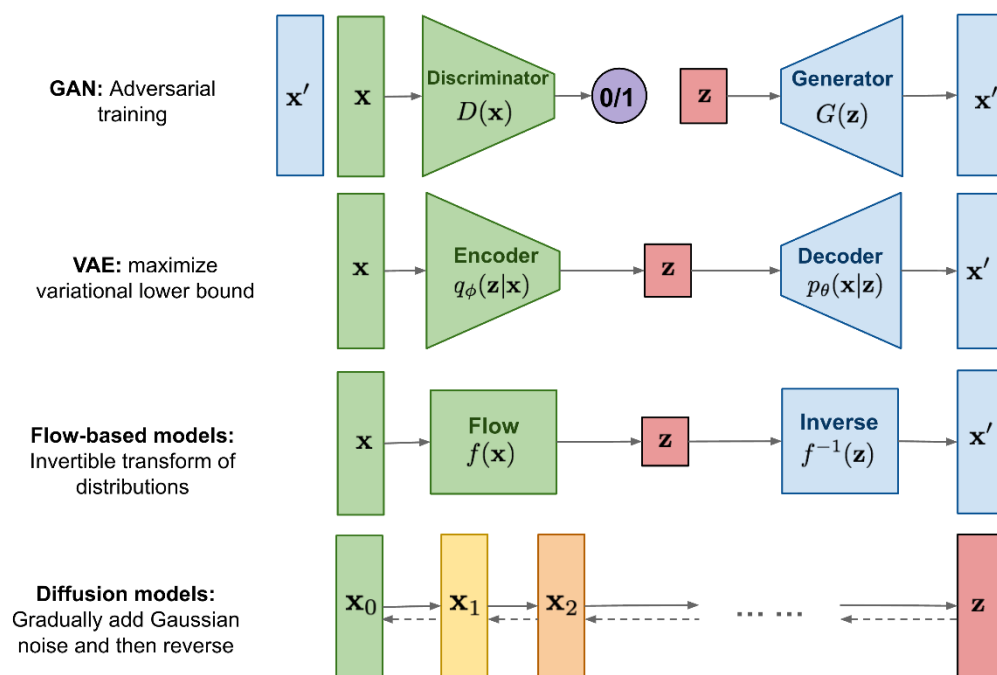
图表 16: 扩散模型训练原理



资料来源: AssemblyAI, 国盛证券研究所

相比于其他模型，扩散模型的优势在于生成的图像质量更高，且无需通过对抗性训练，这使得其训练的效率有所提升。同时，扩散模型还具有可扩展性和并行性。

图表 17: 扩散模型与其他模型比较



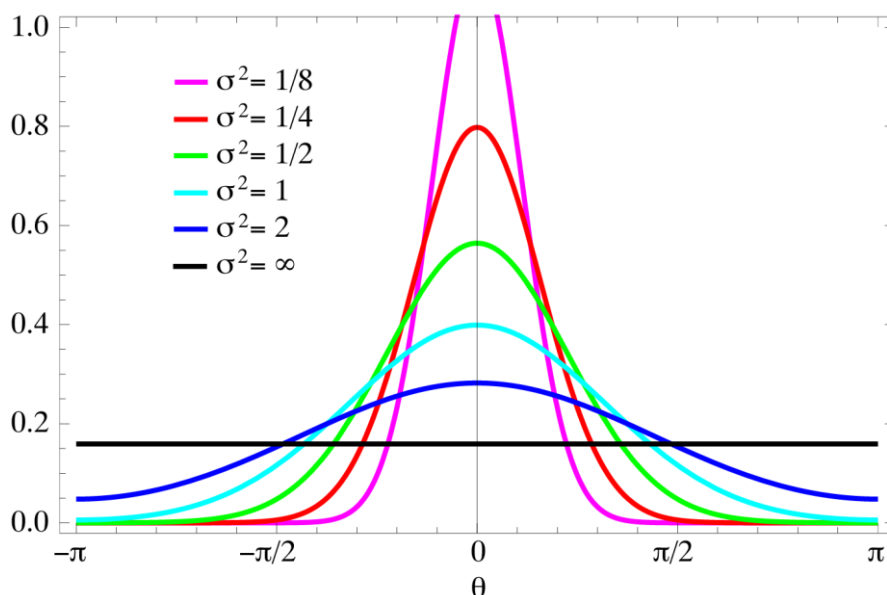
资料来源: Github "what are diffusion models?", 国盛证券研究所

高斯噪声是一种概率密度函数符合正态分布的函数，当 AIGC 运用扩散模型来生成内容的时候，是通过在一副纯白的画布（随机白噪声）上逐步去噪来生成最终的目标画作。即用用户给出的文本描述形容词，来从一个模糊的概念逐步具象。我们可以简化为多个正态分布函数的叠加，模型选择其中重叠的区间输出，这也是一个逐步缩小范围的过程。这与人类的思维模式很类似。

简言之，在 AI 训练阶段，我们将数据集中上亿组图文对进行训练，提取特征值；生产过程中，通过添加文字描述，引入不同的特征值进行去噪，从而生产一副 AI 理解下的内容作品。例如，在当我们在脑海中想象一个画面的时候，比如：一只柯基通过一个小号玩火焰。我们的思维模式也是先有一只柯基，再去想象小号和火焰，最后将这些元素叠加

在柯基身上。

图表 18: 正态分布函数叠加



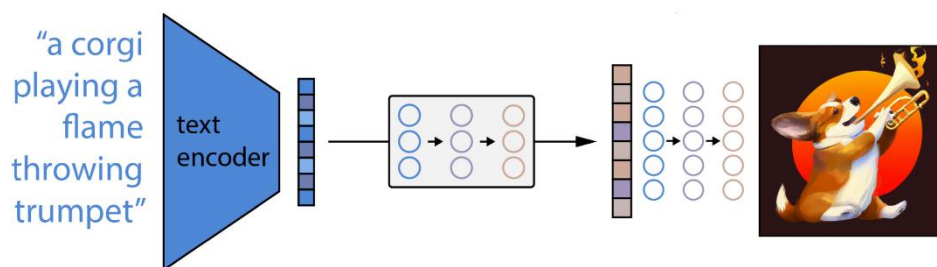
资料来源: Wikipedia, 国盛证券研究所

简述完原理以后，我们可以通过目前非常先进的 AI 图像生成应用 DALL-E2 来举例阐述具体的工作过程：

- 将文本提示输入到一个经过训练能够将提示映射到表示空间的文本编码器中；
- 通过一个被称为“先验”（Prior）的模型，将文本编码映射到图像编码器中。这一图像编码器会捕获文本编码包含的信息和语义；
- 图像编码器随机生成一个图像，这一图像是该语义信息的视觉表现。

这一个过程和人类的思维模式相似。在生成的过程中，涉及到了文本编码器这一概念，目前主流的文本编码器是来自于 OpenAI 的 Clip 模型，其通过 4 亿组文字-图片对进行训练。当然，其中的模型训练都是基于英文实现，语言的区别又会给 AIGC 带来另一重挑战。

图表 19: DALL-E2 工作过程举例



资料来源: AssemblyAI, 国盛证券研究所

除了上述提到的自然语言处理技术和 AIGC 生成算法模型以外，超级计算机和算力这些硬件作为基础设施也是不可或缺的。在机器学习的过程中，需要通过大量的训练来实现更准确的结果，这样的计算量普通的电脑是无法完成的，目前主要由英伟达 A100 构建的计算集群完成，而国内外的初创企业也会通过云实现。

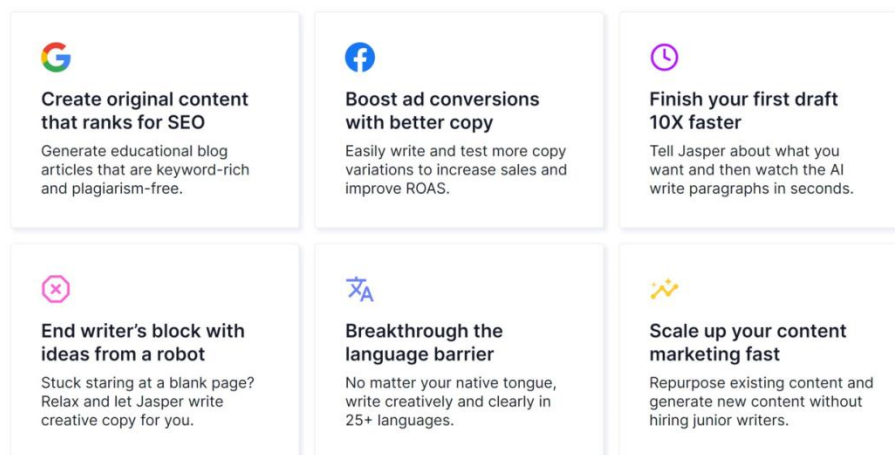
2.3 当我们开始用 AIGC——商业模式的探索

AIGC 已被广泛应用在文字、图像、音频、游戏和代码的生成当中，一些较早创立的企业已经实现了较好的商业化。尤其在一些具备高重复性的任务、对于精度要求并不那么高的领域应用较为成熟。随着 AIGC 技术的发展，其适用面将会逐渐扩大。这类 AIGC 服务的提供商大多数时候会以提供 SaaS 服务的形式变现。

● 文字创作

AIGC 生成文字目前主要被应用于新闻的撰写、给定格式的撰写以及风格改写。其中，有一家成立还不满两年的独角兽企业 Jasper 在最新一轮的融资里获得了 1.25 亿美元资金，目前估值为 15 亿美元。Jasper 成立于 2021 年，是一个 AI 内容平台，允许个人和团队利用 AI 来创作内容，多用于商业。用户可以在借助 Jasper 生成具有丰富关键词、搜索引擎优化的原创博客，可以通过文字描述让 Jasper 帮助完成文章的创作、创建广告话术。通过 Jasper 用户可以寻找创作思路、高效完成文案、突破语言壁垒，而不会存在抄袭的嫌疑。目前，Jasper 拥有 7 万多名客户，包括 Airbnb、Ibm 等企业。仅 2021 年一年便创造了 4000 万美元的收入，今年预估收入为 9000 万美元。

图表 20: Jasper 的应用场景

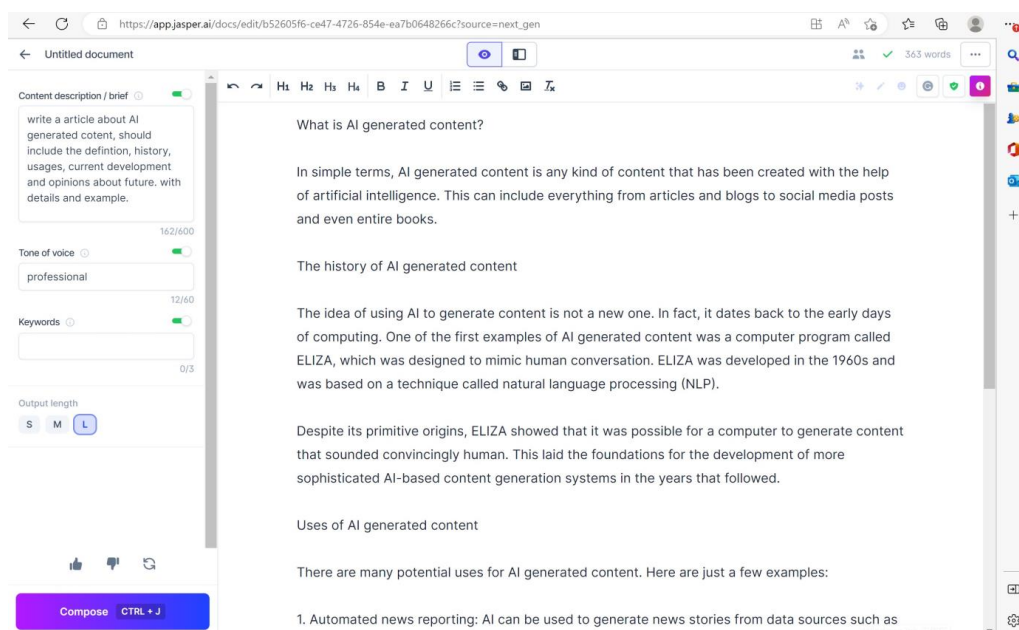


资料来源: Jasper, 国盛证券研究所

用户可以通过输入一段对于目标文章的描述或者要求，系统会自动抓取数据，根据我们描述的指令进行创作。作者本人进行如下实验，输入的描述为【写一篇关于 AIGC 的文章，其中要包含 AIGC 的定义、发展史、应用、现阶段发展情况和对于未来发展趋势的看法，同时，要包含细节和举例】。同时，在风格上我选择了“专业性”。Jasper 很快就生成了一篇 AIGC 撰写的 AIGC 文章（如下图所示），可以看到这篇文章语义通顺，按照我们给出的描述逐段阐述，并且包含了一些举例，这个生成效果无疑会大幅度提升人类的写作效率。

并且，在 Jasper 的网页版 APP 上，还给出了数百种模板，可以根据需求更好的完成作品。

图表 21: 通过 AIGC 撰写的 AIGC 报告

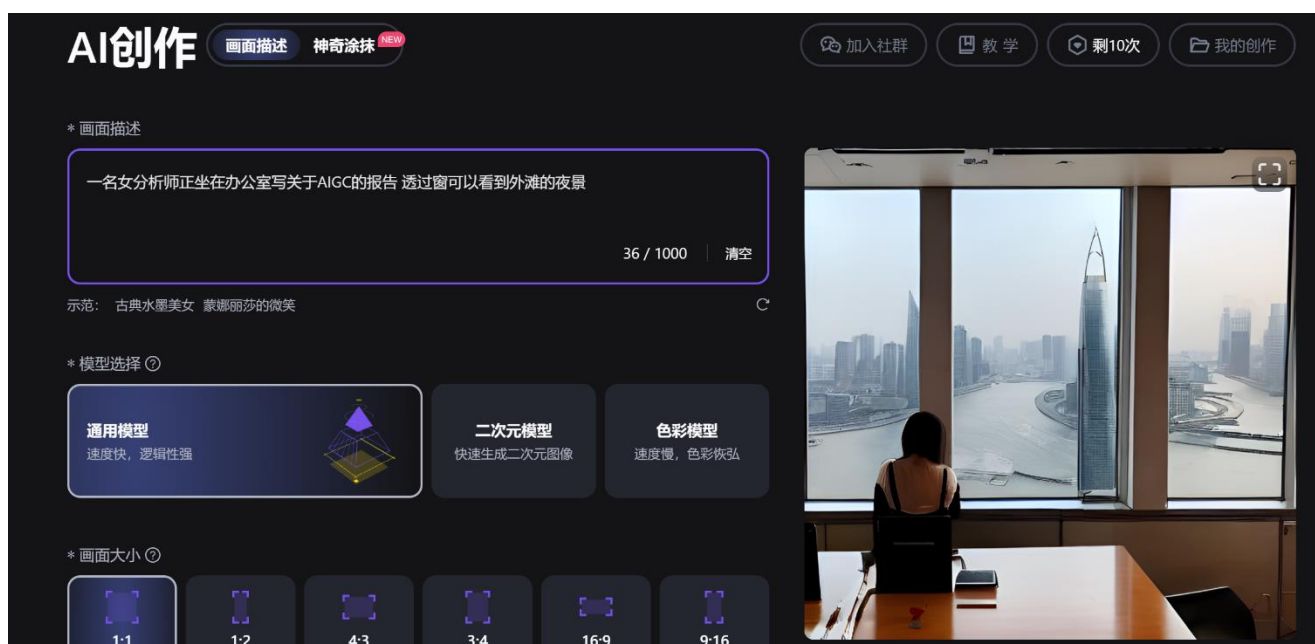


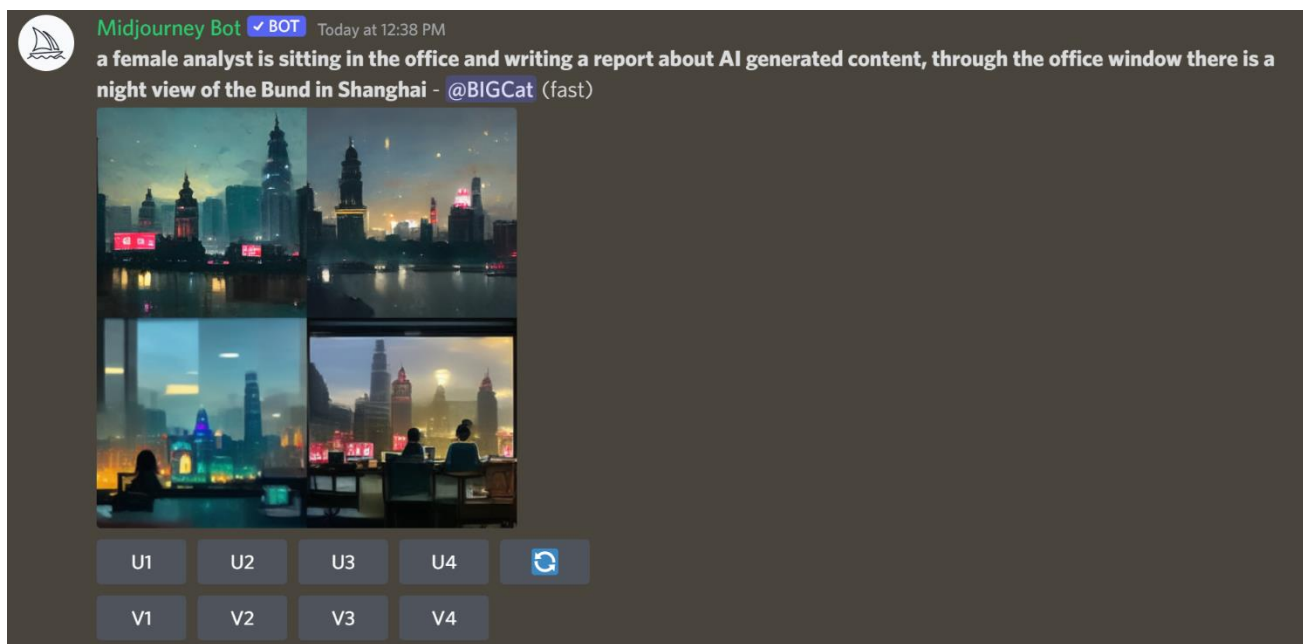
资料来源: Jasper, 国盛证券研究所

● 图像创作

MidJourney 降低了艺术绘画创作的门槛, 用户只需要通过输入文字描述, 计算机将会自动生成一张作品。其背后的逻辑在于计算机通过 NLP 识别语意并翻译成计算机语言, 结合后台的数据集 (这些数据集主要通过自有素材或机器人爬取公开版权的内容获得), 创作出一副全新的作品。这样产生的作品原则上属于 AI 创作, 因此, 在新闻媒体等平台被广泛使用, 不仅减少了成本, 同时避免了潜在的版权纠纷风险。除此以外, 在抖音、微信等社交平台上, 已经有一些数据集图库博主通过 AIGC 创造素材并结合自己的私域流量进行商业变现。

图表 22: 通过 AIGC 生成的图片





资料来源: MidJourney Bot 和无界版图, 国盛证券研究所

近期, OpenAI 已经与全球最大的版权图片供应商之一的 Shutterstock 达成深度合作, Shutterstock 将开始出售利用 OpenAI 的 DALL-E 生成的图片, 并禁止销售非 DALL-E 生成的图片, 完成深度独家绑定。

AIGC 除了大家熟知的生成绘画以外, 还可以利用这一功能完成文字和图片的互相转换, 这在写专利时可以被用到。

图表 23: MidJourney 用户案例



资料来源: MidJourney, 国盛证券研究所

● 视频创作

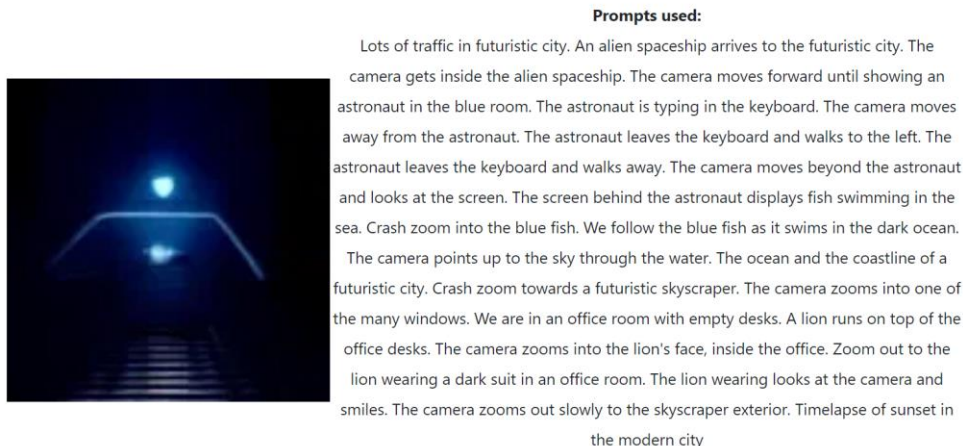
除了绘画以外, AIGC 也能够被运用在视频创作中。Google 推出了 AI 视频生成模型 Phenaki 能够根据文本内容生成可变时长视频的技术, 在公布的 DEMO 中, Phenaki 基于几百个单词组成一段前后逻辑连贯的视频只需两分钟。相比原有的 Imagen 基

础上衍生的 Imagen Video 瞄准短视频，Phenaki 瞄准的是长视频。AIGC 视频中的运用，让我们看到了未来虚拟人也能够作为演员在影视剧中扮演不同的角色以提高内容产出的效率和多样性。

图表 24: Phenaki2 分钟视频生成案例

2 minute video

This 2-minute story was generated using a long sequence of prompts, on an older version of the model



资料来源: Phenaki, 国盛证券研究所

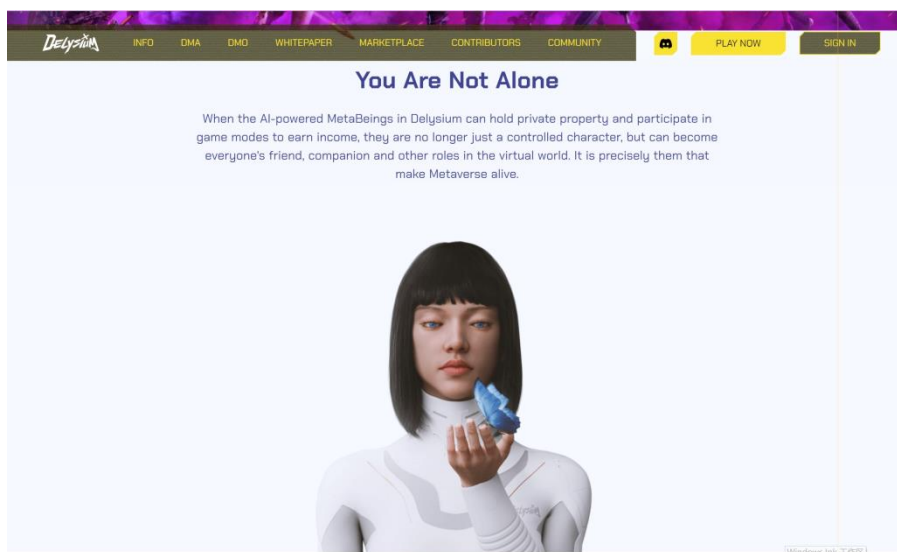
● 音频剪辑

AIGC 生成音频早被应用于我们的日常生活当中。我们常用的手机导航，可以切换不同明星甚至于卡通人物的语音提示。这是通过提前请明星或卡通人物的配音朗读完成一个语音库，再通过反复的训练学习使得可以用指定的声音说出任何话。我们自己也可以通过高德地图录制自己的语音导航包。而更深层次的应用将会是虚拟人领域，AIGC 不仅可以生成虚拟人的声音，并可以创造出说的内容。虚拟人在未来有望和我们一样表达自己的想法，灵魂逐步显现。

● 游戏开发

AIGC 在游戏当中的应用可以分为两方面，一方面是用于场景和故事的搭建。开放世界游戏越来越受欢迎，通过 AIGC 来创建场景和 NPC 都将会大幅度提升效率和降低成本。另一方面，玩家可以通过 AIGC 的平台工具来创建自己的虚拟人，可以用于游戏中的打金等活动。有一家叫做 Delysium 的游戏已经开始引入这一功能。或许在未来的开放世界游戏中，不同的玩家将对应不同的游戏剧情和副本，这无疑将是令人兴奋的应用。

图表 25: Delysium 中的虚拟人玩家

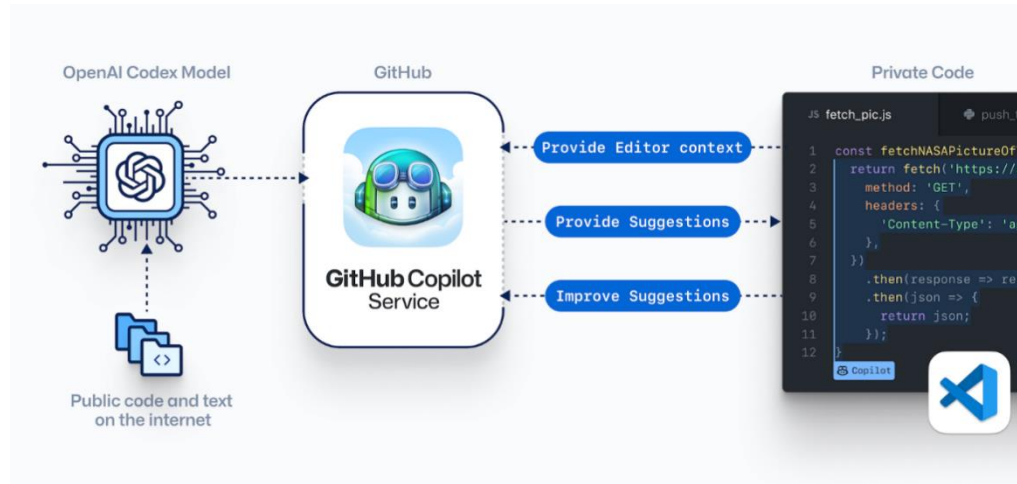


资料来源: Delysium, 国盛证券研究所

● 代码生成

GitHub Copilot 是一个 GitHub 和 OpenAI 合作产生的 AI 代码生成工具, 可根据命名或者正在编辑的代码上下文为开发者提供代码建议。官方介绍其已经接受了来自 GitHub 上公开可用存储库的数十亿行代码的训练, 支持大多数编程语言。

图表 26: GitHub Copilot



资料来源: GitHub, 国盛证券研究所

3.AIGC 的未来发展趋势

AIGC 是 PGC、UGC 之后, 全新的内容生产方式。不仅能提升内容生产的效率以满足我们飞速增长的内容需求, 也能够丰富内容的多样性。在 2022 年百度世界大会上, 李彦宏提到了: “AIGC 将走过三个发展阶段: 第一个阶段是“助手阶段”, AIGC 用来辅助

人类进行内容生产；第二个阶段是“协作阶段”，AIGC 以虚实并存的虚拟人形态出现，形成人机共生的局面；第三个阶段是“原创阶段”，AIGC 将独立完成内容创作。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，去生成 AI 原创内容。”

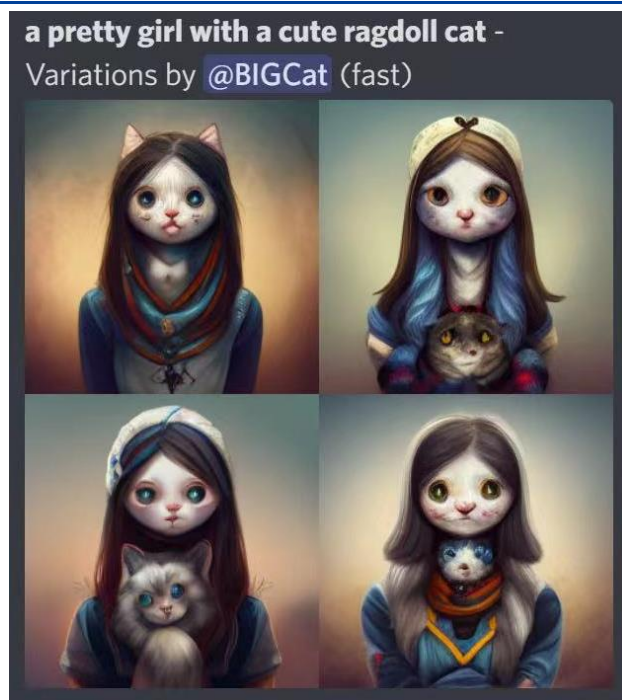
3.1 面临的挑战

技术上来看，虽然当前生成的图片、文字已经可以用以商业用途，但还存在一些问题使得无法满足较高的质量要求。我们可以发现在二次元或抽象的图片生成中，AIGC 的表现较好。但对于比较具体和细节的内容，生成的效果不尽如人意。

下图是笔者通过 AIGC 生成的一副“美女与布偶猫”的图片，从这一张图片我们可以发现有两个问题：

- 其中有两幅图片的猫咪眼睛很奇怪，在这些细节描绘上还无法和真人画师媲美。
- 输入的关键词是“美女”与“布偶猫”，但是生成的“美女”均长着一张猫脸，从这里反映出 AIGC 绘画会出现一些空间位置以及数量上的偏差。产生的原因主要还是来源于语义理解和处理上的问题。

图表 27: AIGC 生成的图片在细节上效果较差



资料来源: MidJourney, 国盛证券研究所

与此同时，参考上文中的图表 23，我们可以发现不同的应用平台，输入几乎一致信息点的文本，生成的图片的质量和内容差距是巨大的。

那么造成以上的这些问题和差距的原因在哪里呢？我们依旧可以从 AIGC 的工作原理上来分析：

- 自然语义的理解在处理一些空间关系上还存在一定的误差，这也是为什么在空间位置、数量上存在不精确的问题。
- 目前文本生成图像时，需要用到文本编码器将文字映射到图像上。当前主流的、训练完善的是来自与 OpenAI 的 Clip 模型，其函数是开源的，但训练的数据集是封闭的。AIGC 需要大量的良好画质的文本-图片对才能训练到 Clip 这样的程度。从 Clip 本身公开的信息来看，它使用了超 4 亿个文本-图片对来完成训练，这些都是基于英文的。那么存在以下几个问题：1、亿级别的高质量文本-图片对对于其他的语言

上获得的难度大幅提高，这也是为什么目前大多除英语外的其他语言的 AIGC 都是需要在整个流程前增加一步翻译。这一步不但涉及语义理解，还包含了文化、语言习惯等潜在的因素，很难被精确翻译，对于翻译模型的挑战很大。2、Clip 的模式很难复刻，即使运用 Clip 开源的函数，基于不同的数据库训练出的结果不同。据我们了解，海外有团队运用了 20 亿的文本-图片对才接近复刻了 Clip；

- 运用的 AIGC 生成算法不同也会导致产生的内容的差距；
- 数据集的质量、合规性、风格偏向都会决定生成的内容质量。

以上，我们可以看到若要使得 AIGC 生成的内容真正高效地被运用在商业层面，那么自然语言处理、翻译模型、生成算法和数据集这些细分赛道都还有很大的进步空间。

3.2 未来的发展方向

在上文中，我们了解到从应用软件方面，自然语言处理、翻译模型、生成算法和数据集这些细分赛道都还有很大的进步空间。更深入地来看，以上这些的发展需要依托于算力、数据的支持。所以未来的发展重点将更着力于大模型、大数据和大算力的方向去发展。同时，为了让功能更加精确，将会更多地去开发一些垂直类的应用，毕竟垂直类的应用可以更有针对性地为特定功能进行训练，成本相对较低。

4. 投资策略：AIGC 的软硬件与数据集

从 PGC 到 UGC 再到 AIGC，AIGC 能让人类突破内容生产力枷锁，高效率生成高质量内容，让人类进入到真正的元宇宙之中。若要 AIGC 能够满足元宇宙的需求，独立完成高质量、高精度的内容，AIGC 技术层面还需要一定的发展，我们可以分为软硬件两个维度看，软件层面主要包括自然语言处理技术、AIGC 生成算法模型和数据集，硬件层面主要是算力、通信网络。

从业务层面看，结合国内外发展情况，目前在 AIGC 的知识产权归属方面尚有法律空缺，且创作伦理问题也未得到有效解决，因此无论是技术还是商业层面，高质、干净的数据集对于模型训练及内容生成均有至关重要的影响。同时，随着 AIGC 逐步落地，其算力需求将大增，未来相关企业除用云计算之外，或组建自有算力集群，考虑到英伟达 A100、H100 出口受限，相关国产算力芯片将有机会获得增量市场。

从主题投资的角度看，区块链、元宇宙、Web3 均描述了数字经济时代中宏大的应用场景，而去年被资本市场关注的虚拟人、NFT 等只是其中的具体应用之一。我们认为，AIGC 将是推动数字经济从 Web2 向 Web3 升级的重要生产力工具：一方面，其对现有的杀手级应用——短视频、游戏等具有颠覆式影响，或进一步扩大内容量、提高成瘾性，同时对社交和广告提供新的工具；另一方面，Web3 开放、共建的价值观下，UGC、AIGC 的内容会更具吸引力，二次创作、开放想象的浪潮将来临。目前 AIGC 已成为硅谷最新热门方向，国内一级市场、互联网大厂等对 AIGC 应用关注度也在快速提升中。

图表 28: AIGC 投资标的

赛道	公司	代码	AIGC
自然语言处理 NLP	谷歌	GOOG	深度学习模型 Transformer, AIGC 生成应用 Imagen
	微软	MSFT	自然语言处理用 Apache Spark, 投资了 OpenAI
	科大讯飞	002230	语音识别和人工智能
	拓尔思	300229	通过基于大数据和 NLP 的知识图谱的构建, 为虚拟人安装知识大脑
AIGC 生成算法 和数据集	英伟达	NVDA	Omniverse, 生成对抗网络 PoE-GAN, StyleGAN
	Meta	Meta	AIGC 生成应用 Meta AI, 多语言模型 XLS-R
	百度	BIDU	拥有文心大模型, 不仅可以自然语言处理, 还可以生成修复图片
	视觉中国	000681	通过 AIGC 生成艺术照片, 拥有高质量图片数据集
	昆仑万维	300418	旗下海外平台 StarX 完成了首批 5 首完全 AI 作曲的歌曲
	蓝色光标	300058	发布一键生成抽象画平台“康定斯基模型”, 用于 AI 绘画
算力	澜起科技	688008	全球服务器芯片提供商
	天孚通信	300394	全球光模块及数通厂商光学器件供应商
	中兴通讯	000063	5G 设备商龙头
	新易盛	300502	400G 时代领先的光收发器件解决方案提供商
	美格智能	002881	物联网智能模组供应商
	奥飞数据	300738	有影响力的 IDC 服务提供商
	中际旭创	300308	全球光电子领军企业, FB 供应商

资料来源: 公开资料, 国盛证券研究所

风险提示

技术创新不及预期: AIGC 的技术发展不及预期。以及底层硬件技术, 如: 超级计算机、算力的发展不及预期。

政策监管风险: 目前 AIGC 还处于相对早期, 后续是否会出台 AIGC 作品相关知识产权或其他法律监管条款尚不明确。

免责声明

国盛证券有限责任公司（以下简称“本公司”）具有中国证监会许可的证券投资咨询业务资格。本报告仅供本公司的客户使用。本公司不会因接收人收到本报告而视其为客户。在任何情况下，本公司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。

本报告的信息均来源于本公司认为可信的公开资料，但本公司及其研究人员对该等信息的准确性及完整性不作任何保证。本报告中的资料、意见及预测仅反映本公司于发布本报告当日的判断，可能会随时调整。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及资料保持在最新状态，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。

本公司力求报告内容客观、公正，但本报告所载的资料、工具、意见、信息及推测只提供给客户作参考之用，不构成任何投资、法律、会计或税务的最终操作建议，本公司不就报告中的内容对最终操作建议做出任何担保。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。投资者应当充分考虑自身特定状况，并完整理解和使用本报告内容，不应视本报告为做出投资决策的唯一因素。

投资者应注意，在法律许可的情况下，本公司及其本公司的关联机构可能会持有本报告中涉及的公司所发行的证券并进行交易，也可能为这些公司正在提供或争取提供投资银行、财务顾问和金融产品等各种金融服务。

本报告版权归“国盛证券有限责任公司”所有。未经事先本公司书面授权，任何机构或个人不得对本报告进行任何形式的发布、复制。任何机构或个人如引用、刊发本报告，需注明出处为“国盛证券研究所”，且不得对本报告进行有悖原意的删节或修改。

分析师声明

本报告署名分析师在此声明：我们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，本报告所表述的任何观点均精准地反映了我们对标的证券和发行人的个人看法，结论不受任何第三方的授意或影响。我们所得报酬的任何部分无论是在过去、现在及将来均不会与本报告中的具体投资建议或观点有直接或间接联系。

投资评级说明

投资建议的评级标准		评级	说明
评级标准为报告发布日后的 6 个月内公司股价（或行业指数）相对同期基准指数的相对市场表现。其中 A 股市场以沪深 300 指数为基准；新三板市场以三板成指（针对协议转让标的）或三板做市指数（针对做市转让标的）为基准；香港市场以摩根士丹利中国指数为基准，美股市场以标普 500 指数或纳斯达克综合指数为基准。	股票评级	买入	相对同期基准指数涨幅在 15%以上
		增持	相对同期基准指数涨幅在 5%~15%之间
		持有	相对同期基准指数涨幅在 -5%~+5%之间
		减持	相对同期基准指数跌幅在 5%以上
	行业评级	增持	相对同期基准指数涨幅在 10%以上
		中性	相对同期基准指数涨幅在 -10%~+10%之间
		减持	相对同期基准指数跌幅在 10%以上

国盛证券研究所

北京 地址：北京市西城区平安里西大街 26 号楼 3 层 邮编：100032 传真：010-57671718 邮箱：gsresearch@gszq.com	上海 地址：上海市浦明路 868 号保利 One56 1 号楼 10 层 邮编：200120 电话：021-38124100 邮箱：gsresearch@gszq.com
南昌 地址：南昌市红谷滩新区凤凰中大道 1115 号北京银行大厦 邮编：330038 传真：0791-86281485 邮箱：gsresearch@gszq.com	深圳 地址：深圳市福田区福华三路 100 号鼎和大厦 24 楼 邮编：518033 邮箱：gsresearch@gszq.com