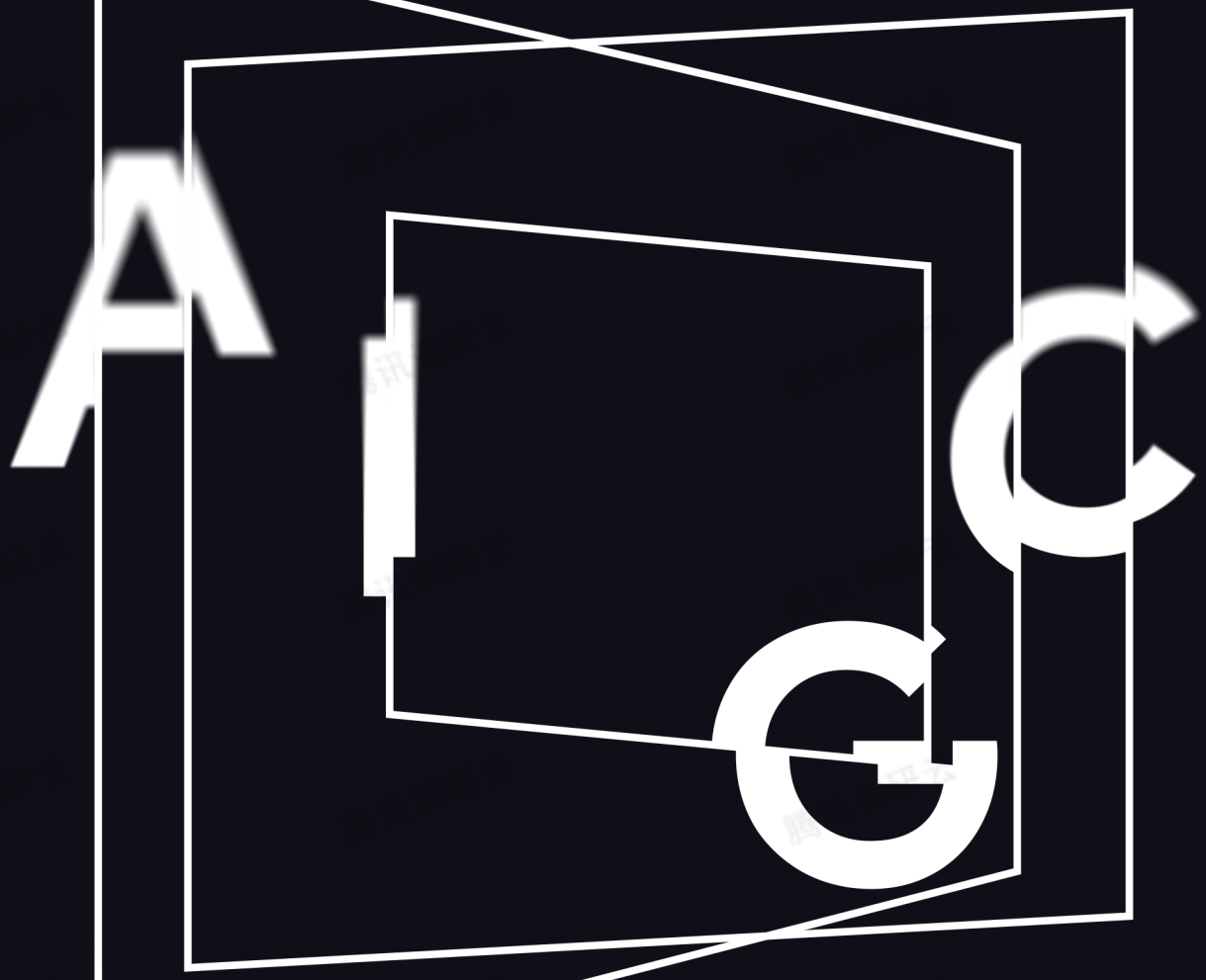
腾讯研究院



AIGC
发展趋势报告 2023

迎接人工智能的下一个时代

腾讯研究院

AIGC
发展趋势报告 2023

迎接人工智能的下一个时代

腾讯研究院

目录

前言 01

一、AIGC技术、应用和产业生态迎来发展快车道 03

1、生成算法、预训练模型、多模态等AI技术累积融合，催生了AIGC的大爆发 04

2、AIGC产业生态加速形成和发展，走向模型即服务（MaaS）的未来 10

二、消费端：AIGC牵引数字内容领域的全新变革 12

1、AIGC有望塑造数字内容生产与交互新范式，成为未来互联网的内容生产基础设施 13

2、AIGC的应用生态和内容消费市场逐渐繁荣 17

3、AIGC将日益成为未来3D互联网的基础支撑 20

4、聊天机器人和数字人成为新的、更包容性的用户交互界面，不断拓展应用领域 24

5、AIGC将作为生产力工具来推动元宇宙发展 28

三、产业端：合成数据牵引人工智能的未来 31

1、合成数据为AI模型训练开发提供强大助推器，推动实现AI2.0 33

2、合成数据助力破解AI“深水区”的数据难题，持续拓展产业互联网应用空间 34

3、合成数据产业加快成为数据要素市场新赛道，科技大厂和创新企业抢先布局 37

4、合成数据加速构建AI赋能、数实融合的大型虚拟世界 39

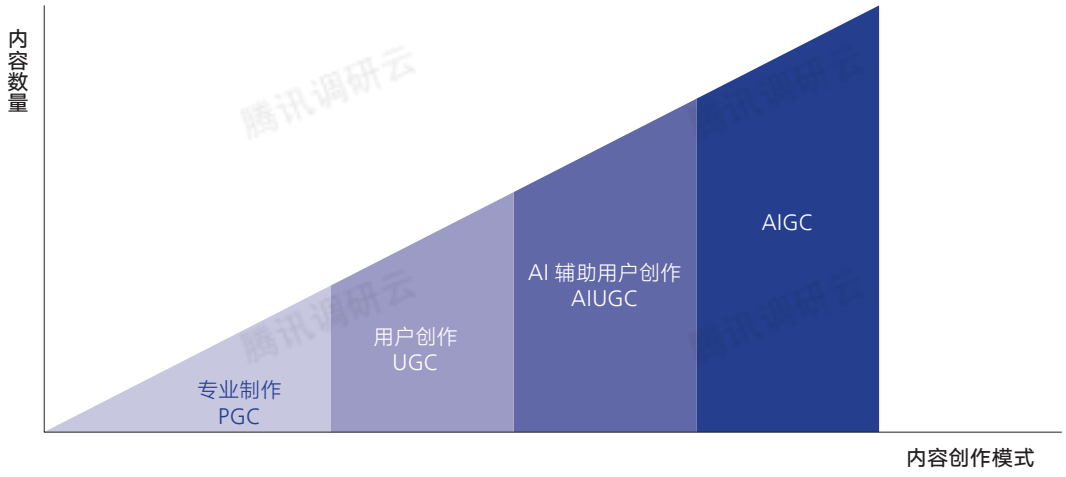
四、社会端：AIGC助力可持续社会价值 **41**
1、元价值：解放人类创造力，革新艺术领域 43
2、AIGC社会价值案例 45

五、发展与挑战并生：积极应对AIGC带来的社会伦理问题 **48**
1、知识产权挑战 49
2、安全挑战 52
3、伦理挑战 54
4、环境挑战 57

六、展望：拥抱人工智能的下一个时代，打造可信AIGC生态 **60**

前言

2022 年是 AIGC（AI-Generated Content，人工智能生成内容）爆火出圈的一年，不仅被消费者追捧，而且备受投资界关注，更是被技术和产业界竞相追逐。9 月 23 日，红杉美国发表文章《生成式 AI：一个创造性的新世界》，认为 AIGC 会代表新一轮范式转移的开始。2022 年 10 月，Stability AI 获得约 1 亿美元融资，估值高达 10 亿美元，跻身独角兽行列。Stability AI 发布的开源模型 Stable Diffusion，可以根据用户输入的文字描述（称为提示词，prompts）自动生成图像，即文生图（Text-to-Image，T2I）。Stable Diffusion、DALL-E 2、MidJourney 等可以生成图片的 AIGC 模型引爆了 AI 作画领域，AI 作画风行一时，标志人工智能向艺术领域渗透。2022 年 12 月，OpenAI 的大型语言生成模型 ChatGPT 刷爆网络，它能胜任刷高情商对话、生成代码、构思剧本和小说等多个场景，将人机对话推向新的高度，让网友们不禁怀疑 ChatGPT 是否已经具有人类智能。全球各大科技企业都在积极拥抱 AIGC，不断推出相关的技术、平台和应用。



图：内容创作模式的四个发展阶段

从字面意思上看，AIGC 是相对于过去的 PCG、UCG 而提出的。因此，AIGC 的狭义概念是利用 AI 自动生成内容的生产方式。但是 AIGC 已经代表了 AI 技术发展的新趋势。过去传统的人工智能偏向于分析能力，即通过分析一组数据，发现其中的规律和模式并用于其他多种用途，比如应用最为广泛的个性化推荐算法。而现在人工智能正在生成新的东西，而不是仅仅局限于

分析已经存在的东西，实现了人工智能从感知理解世界到生成创造世界的跃迁。因此，从这个意义上来看，广义的 AIGC 可以看作是像人类一样具备生成创造能力的 AI 技术，即生成式 AI，它可以基于训练数据和生成算法模型，自主生成创造新的文本、图像、音乐、视频、3D 交互内容（如虚拟化身、虚拟物品、虚拟环境）等各种形式的内容和数据，以及包括开启科学新发现、创造新的价值和意义等。[01] 因此，AIGC 已经加速成为了 AI 领域的新疆域，推动人工智能迎来下一个时代。Gartner 将生成性 AI 列为 2022 年 5 大影响力技术之一。MIT 科技评论也将 AI 合成数据列为 2022 年十大突破性技术之一，甚至将生成性 AI（Generative AI）称为是 AI 领域过去十年最具前景的进展。未来，兼具大模型和多模态模型的 AIGC 模型有望成为新的技术平台。

如果说 AI 推荐算法是内容分发的强大引擎，AIGC 则是数据与内容生产的强大引擎。AIGC 正朝着效率和品质更高、成本更低的方向发展，在某些情况下，它比人类创造的东西更好。包括从社交媒体到游戏、从广告到建筑、从编码到平面设计、从产品设计到法律、从营销到销售等各个需要人类知识创造的行业都可能被 AIGC 所影响和变革。数字经济和人工智能发展所需的海量数据也能通过 AIGC 技术生成、合成出来，即合成数据（synthetic data）。未来，人类的某些创造性的工作可能会被生成性 AI 完全取代，也有一些创造性工作会加速进入人机协同时代——人类与 AIGC 技术共同创造比过去单纯人的创造之下更高效、更优质。在本质上，AIGC 技术的最大影响在于，AIGC 技术将会把创造和知识工作的边际成本降至零，以产生巨大的劳动生产率和经济价值。换句话说，正如互联网实现了信息的零成本传播、复制，未来 AIGC 的关键影响在于，将实现低成本甚至零成本的自动化内容生产，这一内容生产的范式转变，将升级甚至重塑内容生产供给，进而给依赖于内容生产供给的行业和领域带来巨大影响。

但是任何技术都是一把“双刃剑”，AIGC 在引领 AI 技术新趋势和相关产业发展的同时，也可能带来一定的风险挑战，诸如知识产权保护、安全、技术伦理、环境影响等。各界需要秉持科技向善理念，负责任地、安全可控地发展应用 AIGC，打造安全可信的 AIGC 技术和应用，以科技向善引领 AIGC 技术创新发展和应用，实现高质量、健康可持续的发展。

01 注：国外倾向于使用生成式AI（Generative AI）的表述，国内通常使用AIGC的表述。本报告认为两者虽表述有不同，但是表意是一样的，故而在行文中不做特别的区分。

AIGC

技术

产业

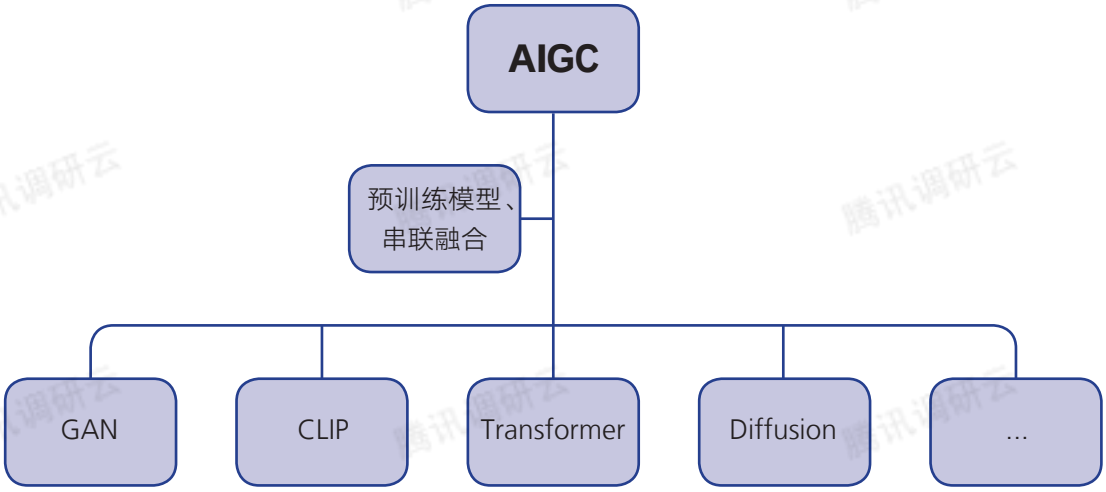
AIGC 技术和产业生态迎来发展快车道

● 本章主笔：　胡晓萌　腾讯研究院研究员

1

虽然从严格意义上来说，1957 年莱杰伦 · 希勒（Lejaren Hiller）和伦纳德 · 艾萨克森（Leonard Isaacson）完成了人类历史上第一支由计算机创作的音乐作品就可以看作是 AIGC 的开端，距今已有 65 年，这期间也不断有各种形式的生成模型、AI 生成作品出现，但是 2022 年才真正算是 AIGC 的爆发之年，人们看到了 AIGC 无限的创造潜力和未来应用可能性。目前，AIGC 技术沉淀、产业生态已初步形成，保持强劲发展和创新势头。

1、生成算法、预训练模型、多模态等 AI 技术累积融合，催生了 AIGC 的大爆发

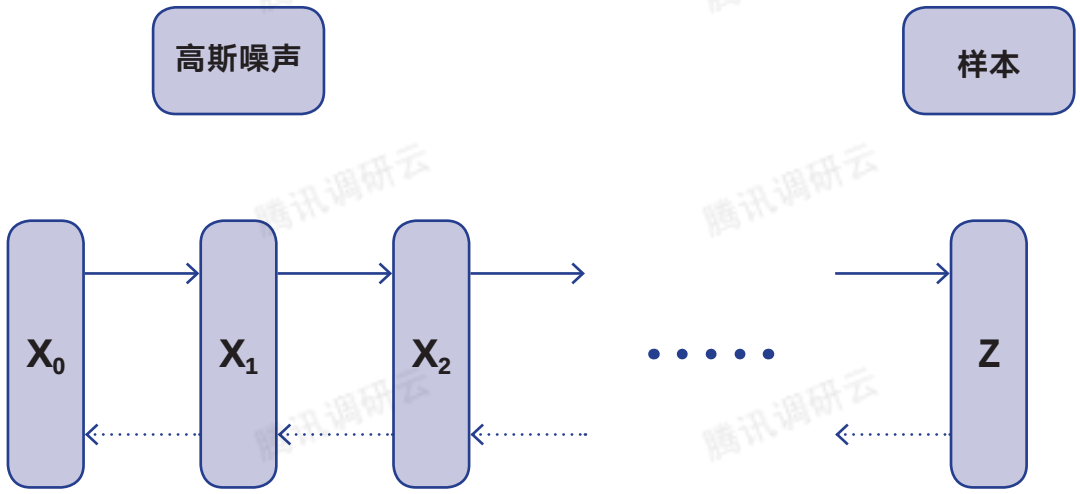


图：AIGC 技术累积融合 [02]

一是，基础的生成算法模型不断突破创新。2014 年，伊恩 · 古德费洛（Ian Goodfellow）提出的生成对抗网络（Generative Adversarial Network，GAN）成为早期最为著名的生成模型。GAN 使用合作的零和博弈框架来学习，被广泛用于生成图像、视频、语音和三维物体模型等。GAN 也产生了许多流行的架构或变种，如 DCGAN，StyleGAN，BigGAN，StackGAN，Pix2pix，Age-cGAN，CycleGAN、对抗自编码器（Adversarial Autoencoders，AAE）、对抗推断学习（Adversarially Learned Inference，ALI）等。

02 图解：以DALL·E 2为例，它通过预训练模型的方式，串联融合了打通文字与图像两种模态的CLIP模型、控制图像生成的扩散模型、以及能够同时处理多项语言任务的语言模型，最终成为支持文字生成图像的多模态AIGC工具。

随后，Transformer、基于流的生成模型（Flow-based models）、扩散模型 (Diffusion Model) 等深度学习的生成算法相继涌现。其中，Transformer 模型是一种采用自注意力机制的深度学习模型，这一机制可以按照输入数据各部分重要性的不同而分配不同的权重，可以用在自然语言处理（NLP）、计算机视觉（CV）领域应用。后来出现的 BERT、GPT-3、LaMDA 等预训练模型都是基于 Transformer 模型建立的。而扩散模型 (Diffusion Model) 是受非平衡热力学的启发，定义一个扩散步骤的马尔可夫链，逐渐向数据添加随机噪声，然后学习逆扩散过程，从噪声中构建所需的数据样本。扩散模型最初设计用于去除图像中的噪声。随着降噪系统的训练时间越来越长并且越来越好，它们最终可以从纯噪声作为唯一输入生成逼真的图片。[03]



图：扩散模型结构

然而，从最优化模型性能的角度出发，扩散模型相对 GAN 来说具有更加灵活的模型架构和精确的对数似然计算，已经取代 GAN 成为最先进的图像生成器。2021 年 6 月，OpenAI 发表论文已经明确了这个结论和发展趋势。[04]

03 P. Dhariwal,A. Nichol.Diffusion Models Beat GANs on Image Synthesis(2021)
04 https://ai.googleblog.com/2021/07/high-fidelity-image-generation-using.html

模型	提出时间	模型描述
变分自动编码（Variational Autoencoders，VAE）	2014年	基于变分下界约束得到的Encoder-Decoder模型对
生成对抗网络（GAN）	2014年	基于对抗的Generator-Discriminator模型对
基于流的生成模型（Flow-based models）	2015年	学习一个非线性双射转换（bijective transformation），其将训练数据映射到另一个空间，在该空间上分布是可以因子化的，整个模型架构依靠直接最大化log-likelihood来完成
扩散模型（Diffusion Model）	2015年	扩散模型有两个过程，分别为扩散过程和逆扩散过程。在前向扩散阶段对图像逐步施加噪声，直至图像被破坏变成完全的高斯噪声，然后在逆向阶段学习从高斯噪声还原为原始图像的过程。 经过训练，该模型可以应用这些去噪方法，从随机输入中合成新的“干净”数据。
Transformer模型	2017年	一种基于自注意力机制的神经网络模型，最初用来完成不同语言之间的文本翻译任务，主体包含Encoder和Decoder部分，分别负责对源语言文本进行编码和将编码信息转换为目标语言文本
神经辐射场（Neural Radiance Field，NeRF）	2020年	它提出了一种从一组输入图像中优化连续5D神经辐射场的表示（任何连续位置的体积密度和视角相关颜色）的方法，要解决的问题就是给定一些拍摄的图，如何生成新的视角下的图

CLIP（Contrastive Language-Image Pre-Training）模型	2021年	1、进行自然语言理解和计算机视觉分析； 2、使用已经标记好的“文字-图像”训练数据。一方面对文字进行模型训练。一方面对另一个模型的训练，不断调整两个模型的内部参数，使得模型分别输出的文字特征和图像特征值并确认匹配。

主流生成模型一览表

二是，预训练模型引发了 AIGC 技术能力的质变。虽然过去各类生成模型层出不穷，但是使用门槛高、训练成本高、内容生成简单和质量偏低，远远不能满足真实内容消费场景中的灵活多变、高精度、高质量等需求。预训练模型的出现引发了 AIGC 技术能力的质变，以上的诸多落地问题得到了解决。

随着 2018 年谷歌发布基于 Transformer 机器学习方法的自然语言处理预训练模型 BERT，人工智能领域进入了大炼模型参数的预训练模型时代。AI 预训练模型，又称为大模型、基础模型（foundation model），即基于大量数据（通常使用大规模自我监督学习）训练的、拥有巨量参数的模型，可以适应广泛的下游任务。这些模型基于迁移学习的思想和深度学习的最新进展，以及大规模应用的计算机系统，展现了令人惊讶的涌现能力，并显著提高各种下游任务的性能。[05] 鉴于这种潜力，预训练模型成为 AI 技术发展的范式变革，许多跨领域的 AI 系统将直接建立在预训练模型上。具体到 AIGC 领域，AI 预训练模型可以实现多任务、多语言、多方式，在各种内容的生成上将扮演关键角色。按照基本类型分类，预训练模型包括：（1）自然语言处理（NLP）预训练模型，如谷歌的 LaMDA 和 PaLM、Open AI 的 GPT 系列；（2）计算机视觉（CV）预训练模型，如微软的 Florence；（3）多模态预训练模型，即融合文字、图片、音视频等多种内容形式。

05 https://hai.stanford.edu/news/reflections-foundation-models

	预训练模型	应用	参数量	领域
谷歌	BERT	语言理解与生成	4810 亿	NLP
	LaMDA	对话系统		NLP
	PaLM	语言理解与生成、推理、代码生成	5400亿	NLP
	Imagen	语言理解与图像生成	110亿	多模态
	Parti	语言理解与图像生成	200亿	多模态
微软	Florence	视觉识别	6.4亿	CV
	Turing-NLG	语言理解、生成	170亿	NLP
Facebook	OPT-175B	语言模型	1750亿	NLP
	M2M-100	100种语言互译	150亿	NLP
Deep Mind	Gato	多面手的智能体	12亿	多模态
	Gopher	语言理解与生成	2800亿	NLP
	AlphaCode	代码生成	414亿	NLP
Open AI	GPT3	语言理解与生成、推理等	1750亿	NLP
	CLIP&DALL-E	图像生成、跨模态检索	120亿	多模态
	Codex	代码生成	120亿	NLP
	ChatGPT	语言理解与生成、推理等		NLP
英伟达	Megatron-Turing NLG	语言理解与生成、推理	5300亿	NLP
Stability AI	Stable Diffusion	语言理解与图像生成		多模态

国外主要的 AIGC 预训练模型一览表

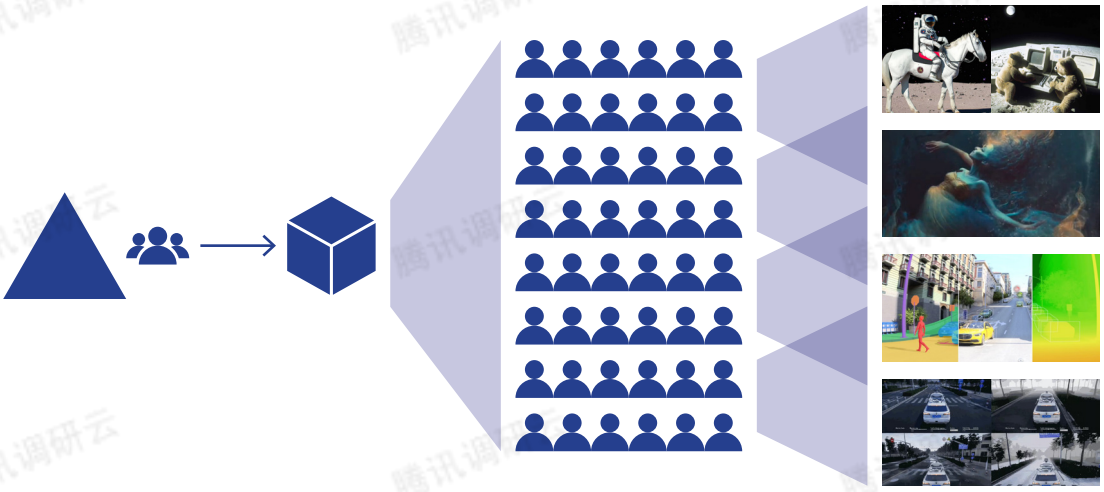
三是，多模态技术推动了 AIGC 的内容多样性，让 AIGC 具有了更通用的能力。预训练模型更具通用性，成为多才多艺、多面手的 AI 模型，主要得益于多模型技术（multimodal technology）的使用，即多模态表示图像、声音、语言等融合的机器学习。2021 年，OpenAI 团队将跨模态深度学习模型 CLIP（Contrastive Language-Image Pre-Training，以下简称“CLIP”）进行开源。CLIP 模型能够将文字和图像进行关联，比如将文字“狗”和狗的图像进行关联，并且关联的特征非常丰富。因此，CLIP 模型具备两个优势：一方面同时进行自然语言理解和计算机视觉分析，实现图像和文本匹配。另一方面为了有足够多标记好的“文本 - 图像”进行训练，CLIP 模型广泛利用互联网上的图片，这些图片一般都带有各种文本描述，成为 CLIP 天然的训练样本。据统计，CLIP 模型搜集了网络上超过 40 亿个“文本 - 图像”训练数据，这为后续 AIGC 尤其是输入文本生成图像 / 视频应用的落地奠定了基础。

在多模态技术的支持下，目前预训练模型已经从早期单一的 NLP 或 CV 模型，发展到现在语言文字、图形图像、音视频等多模态、跨模态模型。2021 年 3 月 OpenAI 发布 AI 绘画产品 DALL · E，只需要输入一句文字，DALL · E 就能理解并自动生成一幅意思相符的图像，且该图像是独一无二的。DALL · E 背后的关键技术即是 CLIP。CLIP 让文字与图片两个模态找到能够对话的交界点，成为 DALL · E、DALL · E2.0、Stable Diffusion 等突破性 AIGC 成果的基石。

总的来看，AIGC 在 2022 年的爆发，主要是得益于深度学习模型方面的技术创新。不断创新的生成算法、预训练模型、多模态等技术融合带来了 AIGC 技术变革，拥有通用性、基础性、多模态、参数多、训练数据量大、生成内容高质稳定等特征的 AIGC 模型成为了自动化内容生产的“工厂”和“流水线”。

2、AIGC 产业生态加速形成和发展，走向模型即服务（MaaS）的未来

目前，AIGC 产业生态体系的雏形已现，呈现为上中下三层架构。



图：AIGC 产业生态体系的三层架构

第一层，为上游基础层，也就是由预训练模型为基础搭建的 AIGC 技术基础设施层。由于预训练模型的高成本和技术投入，因此具有较高的进入门槛。以 2020 年推出的 GPT-3 模型为例，Alchemy API 创始人 Elliot Turner 推测训练 GPT-3 的成本可能接近 1200 万美元。因此，目前进入预训练模型的主要机构为头部科技企业、科研机构等。

目前在 AIGC 领域，美国的基础设施型公司（处于上游生态位）有 OpenAI、Stability.ai 等。OpenAI 的商业模式为对受控的 api 调用进行收费。Stability.ai 以基础版完全开源为主，然后通过开发和销售专业版和定制版实现商业获利，目前估值已经超过 10 亿美金。因为有了基础层的技术支撑，下游行业才能如雨后春笋般发展，形成了目前美国的 AIGC 商业流。

第二层，为中间层，即垂直化、场景化、个性化的模型和应用工具。预训练的大模型是基础设施，在此基础上可以快速抽取生成场景化、定制化、个性化的小模型，实现在不同行业、垂

直领域、功能场景的工业流水线式部署，同时兼具按需使用、高效经济的优势。随着兼具大模型和多模态模型的 AIGC 模型加速成为新的技术平台，模型即服务（Model-as-a-Service，MaaS）开始成为现实，预计将对商业领域产生巨大影响。Stable Diffusion 开源之后，有很多基于开源模型的二次开发，训练特定风格的垂直领域模型开始流行，比如著名的二次元画风生成的 Novel-AI，还有各种风格的角色生成器等。

第三层，为应用层，即面向 C 端用户的文字、图片、音视频等内容生成服务。在应用层，侧重满足用户的需求，将 AIGC 模型和用户的需求无缝衔接起来实现产业落地。以 Stable Diffusion 开源为例，它开放的不仅仅是程序，还有其已经训练好的模型，后继创业者能更好的借助这一开源工具，以 C 端消费级显卡的算力门槛，挖掘出更丰富的内容生态，为 AIGC 在更广泛的 C 端用户中的普及起到至关重要的作用。现在贴近 C 端用户的工具越发丰富多样，包括网页、本地安装的程序、移动端小程序、群聊机器人等，甚至还有利用 AIGC 工具定制代出图的内容消费服务。

目前，从提供预训练模型的基础设施层公司到专注打造 AIGC 产品和应用工具的应用层公司，美国围绕 AIGC 生长出繁荣的生态，技术创新引发的应用创新浪潮迭起；中国也有望凭借领先的 AIGC 技术赋能千行百业。

随着数字技术与实体经济融合程度不断加深，以及互联网平台的数字化场景向元宇宙转型，人类对数字内容总量和丰富程度的整体需求不断提高。AIGC 作为当前新型的内容生产方式，已经率先在传媒、电商、影视、娱乐等数字化程度高、内容需求丰富的行业取得重大创新发展，市场潜力逐渐显现。与此同时，在推进数实融合、加快产业升级的进程中，金融、医疗、工业等各行各业的 AIGC 应用也都在快速发展。

AIGC

消费端

消费端：AIGC 牵引数字内容领域的全新变革

● 本章主笔：胡晓萌　腾讯研究院研究员
胡璇　腾讯研究院高级研究员
徐思彦　腾讯研究院高级研究员
曹建峰　腾讯研究院高级研究员

2

1、AIGC 有望塑造数字内容生产与交互新范式，成为未来互联网的内容生产基础设施

内容生产，特别是创意工作，一向被认为是人类的专属和智能的体现。如今，AI 正大步迈入数字内容生产领域，不仅在写作、绘画、作曲等多个领域达到“类人”表现，更展示出在大数据学习基础上的非凡创意潜能。这将塑造数字内容生产的人机协作新范式，也让内容创作者和更多普通人得以跨越“技法”和“效能”限制，尽情挥洒内容创意。

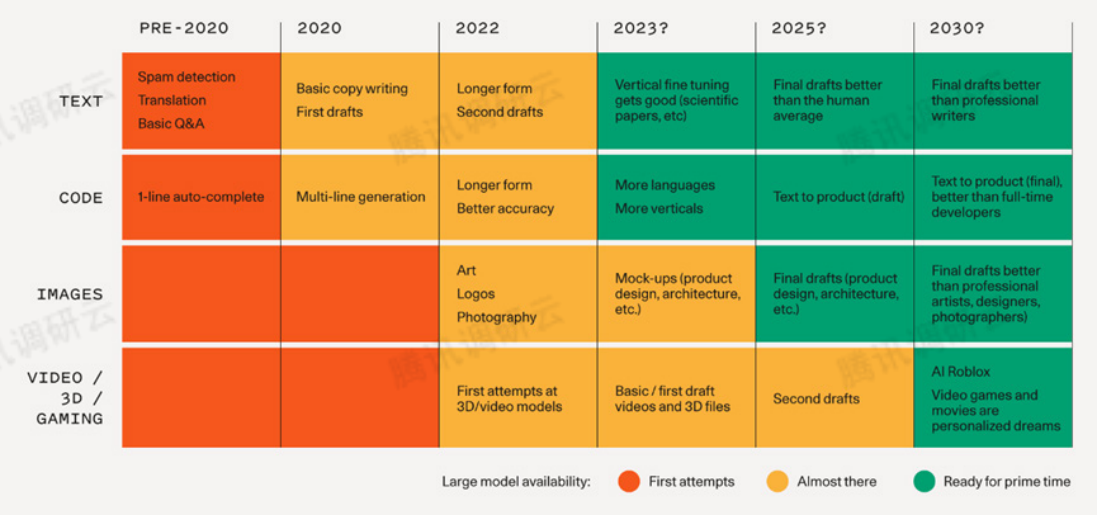
数字内容迈入强需求、视频化、拼创意的螺旋式升级周期，AIGC 在这阶段的发展迎合了这一需要。从全球来看，人们在线的时间持续增长，在线新闻、音乐、动漫、影视、文学、游戏的市场规模仍在上升。同时，数字内容消费的结构也正转型升级，从以图文为主过度到视频为主，视频在网络流量中的占比不断上升。短视频和直播的流行，在生产侧，让原本需要长制作周期的视频变成了可以源源不断产出的“工业品”；在消费侧，过往需要高注意投入、反复观看的视频内容变成了一种媒体“货架”上的“快消品”。最后，创意仍是核心，仍旧稀缺，需要新的模式辅助创作者持续产生、迭代和验证创意。种种因素，都需要更加低成本、高效能的数字内容生产新方式。

AIGC 正在越来越多地参与数字内容的创意性生成工作，以人机协同的方式释放价值，成为未来互联网的内容生产基础设施。

从范围上看，AIGC 逐步深度融入到文字、音乐、图片、视频、3D 多种媒介形态的生产中，可以担任新闻、论文、小说写手，音乐作曲和编曲者，多样化风格的画手，长短视频的剪辑者和后期处理工程师，3D 建模师等多样化的助手角色，在人类的指导下完成指定主题内容的创作、编辑和风格迁移。

从效果上看，AIGC 在基于自然语言的文本、语音和图片生成领域初步令人满意，特别是知识类中短文，插画等高度风格化的图片创作，创作效果可以与有中级经验的创作者相匹敌；在视频和 3D 等媒介复杂度高的领域处于探索阶段，但成长很快。尽管 AIGC 对极端案例的处理、细节把控、成品准确率等方面仍有许多进步空间，但蕴含的潜力令人期待。

从方式上看，AIGC 的多模态加工是热点。吴恩达（Andrew Ng）认为多模态是 2021 年 AI 的最重要趋势，AI 模型在发现文本与图像间关系中取得了进步，如 OPEN AI 的 CLIP 能匹配图像和文本，Dall · E 生成与输入文本对应的图像；DeepMind 的 Perceiver IO 可以对文本、图像、视频和点云进行分类。典型应用包括如文本转换语音、文本生成图片，广义来看 AI 翻译、图片风格化也可以看作是两个不同“模态”间的映射。



图：AIGC 基础模型和应用发展预测（引用自红杉资本）[06]

AIGC 在新研究范式及算力进步下加速发展，特别是自然语言理解能力大幅进化，典型代表是 GPT 系列模型。自然语言是不同数字内容类型间转化的根信息和纽带，比如“猫”这个词语就是加菲猫的照片，音乐剧《猫》和无数内容的索引，这些不同的内容类型可以称为“多模态”。创作者以自然语言描述要素、特征后，AI 就能生成对应的不同模态的结果。2017 年 Google 发布的 Transformer 开启了大型语言模型（Large language model，简称 LLM）时代。它作为更强大的特征提取器，生成的语言模型质量高、效率高，被应用到后续的 GPT、BERT 等各种预训练语言模型中，以大数据预训练 + 小数据微调的方式，摆脱了对大量人工调参的依赖，在手写、语音和图像识别、语言理解方面的表现大幅突破，所生成的内容也越来越准确和自然。

06 https://www.sequoiacap.com/article/generative-ai-a-creative-new-world/

但大模型的高算力要求对普通用户不友好，这使 AIGC 长期只是小圈子的爱好，例如 GPT-3 有 1750 亿参数量，既需要大算力集群也不开放。2022 年，先是部署在论坛 discord 上、以聊天机器人形式向用户开放的 midjourney 引起关注，一位设计师用其生成的图片甚至在线下比赛中获奖。2022 年，Stable Diffusion 等文本生成图片产品，第一次让大众感受 AIGC 的魅力。开源的 Stable Diffusion 仅需一台电脑就能运行，截至 10 月已有超过 20 万开发者下载，累计日活用户超过 1000 万；而面向消费者的 DreamStudio 则已获得了超过 150 万用户，生成超过 1.7 亿图片。



图：游戏设计师杰森·艾伦使用 Midjourney 模型生成的《太空歌剧院》是首个获奖的 AI 生成艺术

在 AIGC 图像生成火爆的同时，ChatGPT 横空出世，真正做到和人类“对答如流”，能理解各式各样的需求，做出回答、短文和诗歌创作、代码写作、数学和逻辑计算等。ChatGPT 可以看作 GPT3.5 版本，使用了人类反馈强化学习 (RLHF) 技术，能持续学习人类对回答的建议和评价，朝更加正确的方向前进。ChatGPT 基于这一思想，用不到 GPT3 的 1% 的参数实现了极佳的效果。但是也依然存在一些缺陷，例如引用不存在的论文和书籍，误解和扭曲直观的自然科学知识和历史事实。

在 AIGC 技术开源的背景下，AIGC 图片生成也促进了小众文化的发展。早在 2022 年五月，QQ 影像中心推出的国内第一款基于 AIGC 的“AI 恋爱专属画”520 活动就在 QQ 小世界走红，最近新推出的“异次元的我”二次元形象生成器火遍国内外。



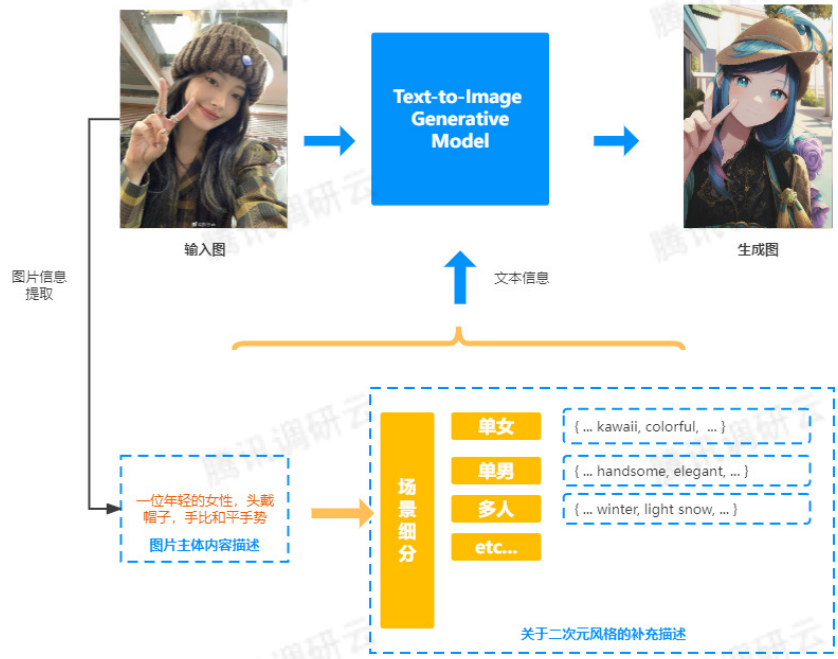
图：KFC 官推亲自打破次元壁

QQ 影像中心将输入模型的提示词分为两个部分——图片主体内容描述和风格描述。这样做是为了避免其他基于 Stable Diffusion 的 AIGC 图像生成平台常见的问题，如不便捷，用户想要生成比较好的结果，需要对图片内容进行详尽的描述，并增加很多提升效果和强化风格的词，才能得到一个比较理想的效果。

（1）内容描述：为了得到图片的内容描述，QQ 小世界对输入图中的信息进行人物性别、年龄、姿势等属性的检测，得到最终尽可能准确详尽的内容描述词。

（2）风格描述：为了让结果更接近二次元，QQ 小世界为各种不同场景设计了非常鲜明的二次元风格。根据第一步对图片内容进行分析后，进一步细分为单人男性、单人女性、多人等不同场景。

结合内容描述和风格描述，QQ 小世界通过对原图的分析得到最终完整的提示词输入网络，从而能够稳定生成画面精致、风格突出的结果。既保证了生成内容和原图的对应性，又增添了很多“异次元”要素，让用户更有穿越到二次元的感觉，同时也保证了较高的生成质量。



图：QQ 小世界二次元风格生成流程图

2、AIGC 的应用生态和内容消费市场逐渐繁荣

AIGC 在学习通用知识和理解泛化上有更好的表现，这是之前的生成模型没有体现出来的能力。这也决定了 AIGC 不光是生成，而是建立在认知与理解上的模型的应用生态。2020 年，OpenAI 推出 1750 亿参数的预训练语言模型 GPT-3，在国内外掀起千亿参数预训练模型的研究浪潮。实际上，那时就出现了专门做文字生成的商业公司，如 Jasper.ai 与 Copy.ai。而今在内容消费领域，AIGC 已经重构了整个应用生态。AIGC 在内容生成领域有以下优势特征：

（1）自动内容生成：大型语言和图像 AI 模型可用于自动生成内容，例如文章、博客或社交媒体帖子。对于定期创建内容的企业和专业人士来说，这可能是一种宝贵的节省时间的工具。2022 年初，Disco Diffusion 给行业带来很大的冲击，其在生成速度、效果，还有词条内容的丰富性等方面都有了显著的改进。架设在社交平台 Discord 上的 Midjourney，因其使用便捷与相对较好的效果，在社会层面得到了很大的关注。

（2）提高内容质量：AIGC 生成的内容可能比人类创建的内容质量更高，因为人工智能模型能够从大量数据中学习并识别人类可能看不到的模式。这可以产生更准确和信息更丰富

的内容。DALL · E 2 和 Google 的 Imagen 都可以通过文字来要求 AI 画出特别具体的内容，而且效果已经接近于中等画师的水平。今年 8 月，Stable Diffusion 的开源模型生成的面部和肢体相较于 Disco Diffusion 自然了许多。

（3）增加内容多样性：AIGC 模型可以生成多种类型的内容，包括文本、图像和音视频、3D 内容等等。这可以帮助企业和专业人士创建更多样化、更有趣的内容，从而吸引更广泛的人 群。2021 年 底，VQGAN + CLIP（Vector Quantised General Adversarial Network + Contrastive Language - Image Pre-training）技术串联，可以生成抽象绘画作品。2022 年 9 月 29 日，Meta AI 公布了一款基于文本生成视频短片的系统“Make-A-Video”，不仅能够通过提示词生成几秒连贯的视频，并且有“超现实的”、“现实的”、“风格化”的视频变种供调整，而且也能从一两张图片生成视频，或者根据一小段视频片段派生不同风格、拍摄角度、相似动作的视频变种。

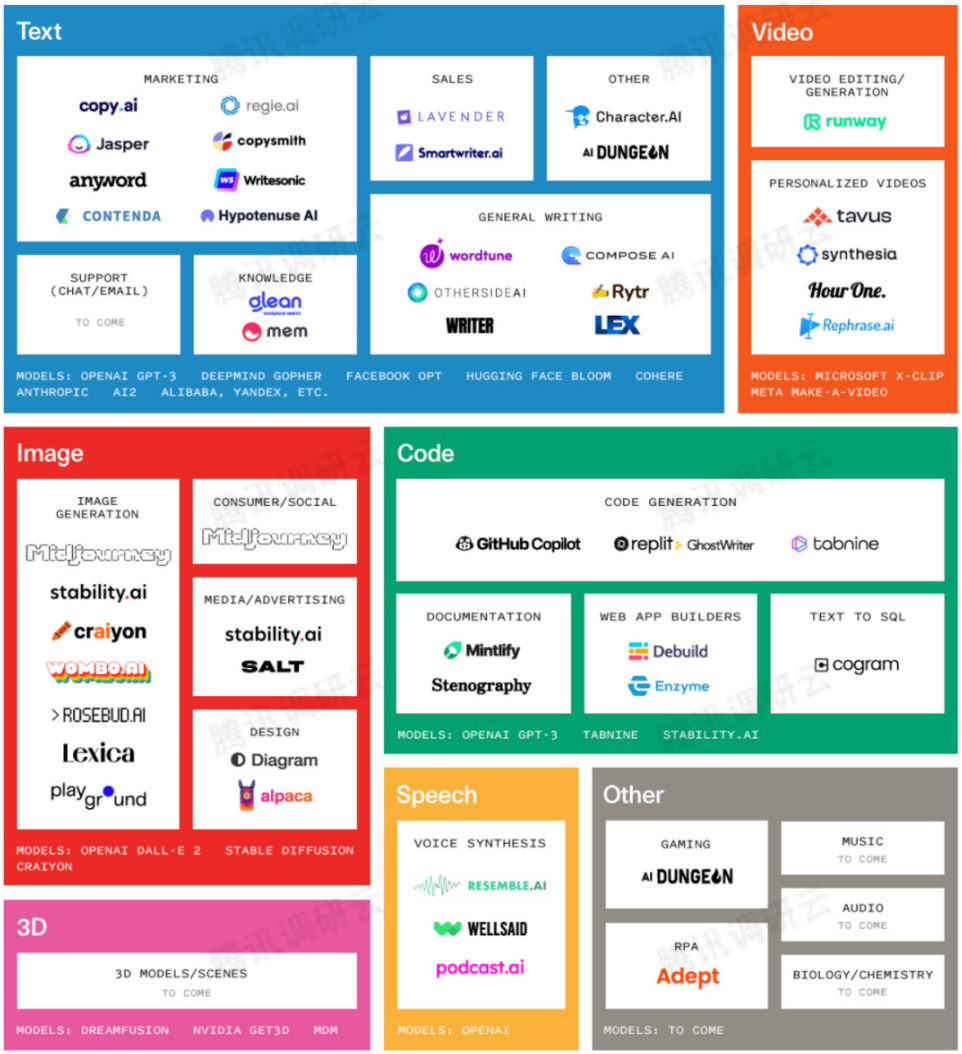
（4）内容制作成本低、门槛低：基于 AIGC，内容制作的成本显著降低、效率显著提高，可以实现以十分之一的成本，以百倍千倍的生产速度，创造出有独特价值和独立视角的内容。[07] 而且 AIGC 也降低了用户的内容制作门槛，即使是毫无绘画基础的用户也能借助 AIGC 工具，画出高质量的作品。

（5）可实现个性化内容生成：人工智能模型可以根据个人用户的喜好生成个性化内容。这可以帮助企业和专业人士创建目标受众更有可能感兴趣的内容，因此更有可能被阅读或分享。比如，基于 Stable Diffusion 的二次元画风生成工具 Novel-AI，可以满足小众的二次元群体的喜好和内容需求，一定程度上也促进了小众文化的发展。

AIGC 已经掀起了一场内容生产的革命。在内容需求旺盛的当下，AIGC 所带来的内容生产方式变革也开始引起了内容消费模式的变化。AI 绘画是 AIGC 的一大重要分支。它可以提高美术素材生产效率，在游戏、数字藏品领域初步得以应用。文字生成图片（Text-to-Image，T2I）是目前 AI 绘画的主流生成方式，用户输入文本命令便可生成相应效果图。对于游戏开发者而言，T2I 工具在概念构思阶段可迅速提供多种创作方向，并降低初期投入成本，在开发过程中

07 https://baijiahao.baidu.com/s?id=1742950890265955383&wfr=spider&for=pc

亦可批量生产石头花纹、花草树木等通用型贴图素材。艺术是另一应用领域，全球最大的 AI 生成艺术画廊 Art AI 以历史上大量艺术品的时期、流派、风格等为灵感生成数字藏品，单字仅能使用一次的机制保障每份数字藏品对应不同的文本内容，从而保障其唯一性。2022 年 10 月，海外知名图库 Shutterstock 就公布说他们与 OpenAI 建立了合作，让用户可以输入文字即时生成满足需求的原创图片。Adobe、Getty Images 也将图像生成模型集成到他们自己的产品和服务中。这些动作不仅是一个传统企业的及时反应，实际上还意味着 AIGC 商业化变现的一个想象似乎开始落地：打造一个基于生成的全新内容平台。



图：AIGC 应用现状概览（引用自红杉资本）[08]

08 https://www.sequoiacap.com/article/generative-ai-a-creative-new-world/

AIGC 作为当前新型的内容生产方式，已经率先在传媒、电商、影视、娱乐等数字化程度高、内容需求丰富的行业取得重大创新发展，市场潜力逐渐显现。2022 年，AIGC 发展速度惊人，迭代速度呈现指数级爆发，谷歌、Meta、百度等平台型巨头持续布局，也有 Stability AI、Jasper AI 等独角兽创业公司出世。AI 绘画已经能承担图像内容生成的辅助性工作，前期初稿的形成可以由 AI 提供，后期再由创作者进行调整，从而提高内容产出效率。根据 6pen 预测，未来五年 10%-30% 的图片内容由 AI 参与生成，有望创造超过 600 亿以上市场空间，若考虑到下一代互联网对内容需求的迅速提升，国外商业咨询机构 Acumen Research and Consulting 预测，2030 年 AIGC 市场规模将达到 1100 亿美元 [09]。

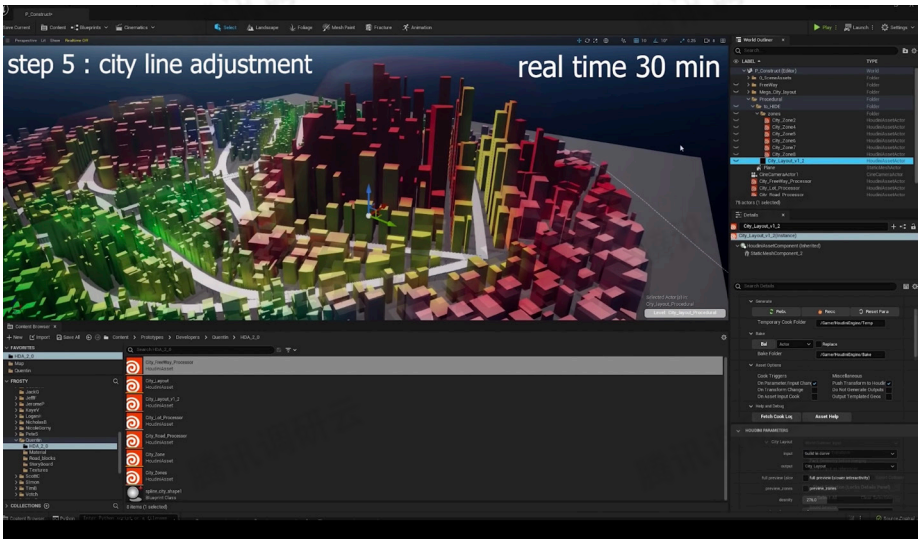
3、AIGC 将日益成为未来 3D 互联网的基础支撑

互联网向下一代技术升级和演进的重要方向是从“在线”走向“在场”，迈向 3D 互联网时代，AIGC 将成为打造虚实集成世界的基石。人们将在在虚拟空间构建仿真世界，在现实世界“叠加”虚拟增强，实现真正的临场感。随着 XR、游戏引擎、云游戏等等各种交互、仿真、传输技术的突破，信息传输越来越接近无损，数字仿真能力真假难辨，人类的交互和体验将到达新阶段。

AIGC 为 3D 互联网带来的价值，既包括 3D 模型、场景、角色制作能效的提升，也能像 AI 作画那样，为创作者激发新的灵感。传统的 3D 制作需要耗费大量时间和人力成本。以 2018 年发售的游戏《荒野大镖客 2》为例，为了打造约 60 平方公里的虚拟场景、先后有六百余名美术历经 8 年完成。

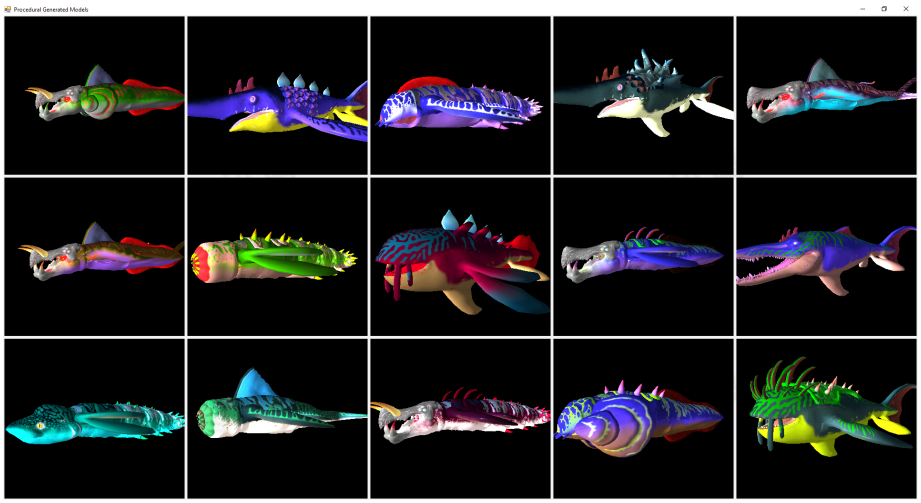
程序化内容生成技术（PCG，Procedural Content Generation）是一种在效能和可控度上介于纯手工和 AIGC 之间的产出方式。2022 年的交互内容《黑客帝国：觉醒》中就采用了这项技术，在最新引擎技术的加持下，虚拟城市共包括 700 万个实例化资产，每个资产由数百万个多边形组成，包括 7000 栋建筑、38000 辆可驾驶的车和超过 260 公里的道路。

09 https://www.yahoo.com/now/generative-ai-market-size-achieve-230000291.html



图：Epic 使用虚幻 5 引擎和程序化生成技术高效制作《黑客帝国：觉醒》中的庞大城市，图片引用自知乎用户 @ 王程

2016 年发布、主打宇宙探险的独立游戏《无人深空》（No Man’s Sky）声称创造了 1840 亿颗不同的星球，每个星球都有形态各异的环境和生物。游戏正是用 PCG 构造了一系列生成规则和参数，用数学函数和极少的数据量生成了这个看似无限的游戏宇宙。



图：游戏《无人深空》中自动化生成的海洋生物示例

尽管如此，距离“复刻”真实世界和打造千千万万个幻想世界，这样的速度还是太慢了。为什么在文字、图片领域大放异彩的 AI，没有继续快速“征服”3D 领域呢？

AI 直接生成 3D 内容，比 2D 图片难度大得多。首先，信息量不同，一张图片和一个 3D 模型相比，本身蕴含的信息量就有数量级的差距；其次，图片和 3D 的实质不同，如果说 2D 是像素点阵在显示器的客观陈列，3D 则是实时、快速、海量的矩阵运算，就像对着模型在 1 秒内进行几十次“拍照”。为了准确计算得到每个像素点，“渲染”在显示器上，需要考虑的因素至少有：（1）模型几何特征，通常用几千上万个三角面来表示；（2）材质特征，模型本身的颜色，是强反射的金属，还是漫反射的布料；（3）光线，光源是点状的吗，颜色和强度如何。最后，原生 3D 模型的数据相对较少，远不如已存在了数千年、可以以非数字化形态存在的图像那么多，例如 ImageNet 中就包含了超过 1400 万张图片。

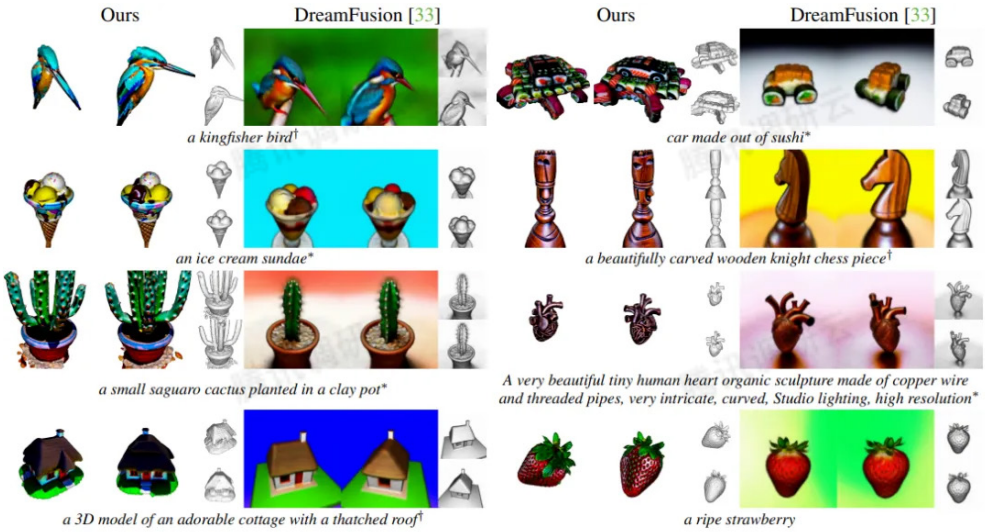
现阶段，AIGC 在 3D 领域的突破，主要是从图片、现实中存在的物体取材后，再由 AI 介入进行转化。计算机视觉技术的进步，为 3D 另辟蹊径：我们可以通过图片来推测物体在 3D 空间中的形态，尽可能还原每个像素点原本的深度信息。采用激光雷达等设备扫描物理世界，能快速得到大量 3D 模型。当然，这些素材要经过简化处理才能更好使用。因为它们携带了海量的坐标点和颜色信息，呈现起来耗费的计算资源多。想要应用在影视、游戏、元宇宙等场景中，和运行在电脑、手机等设备中，需要经过一系列的后处理，比如让模型的几何表示更简洁，还原材质信息、尽可能去除拍摄时的光照影响等。处理后的模型，也会具有更强的可编辑性，比如更改它原本的形状、颜色等。

目前 AIGC 在 3D 模型领域还处于探索阶段，从基本原理和使用的模型来看，仍然以扩散模型为基础，基于文字 - 图片 -3D 路径进行处理。谷歌和英伟达在这一领域较为领先，先后在 2022 年发布了自己的文字生成 3D 的 AI 模型。[10] 但从生成效果看，距离现在人工制作 3D 内容的平均质量还有距离；生成速度也未能尽如人意。

10 月，谷歌率先发布了 DreamFusion，它能够使用文本到图像模型先生成 2D 图像，然后优化为体积 NeRF（神经辐射场）数据。但其缺点也很显著，首先扩散模型仅对 64x64 的图像生效，导致生成 3D 的质量不高；其次场景渲染模型使基于 Mip-NeRF 360 的大型 MLP，不仅需要海量样本，也在计算上费时费力，导致生成速度较慢。

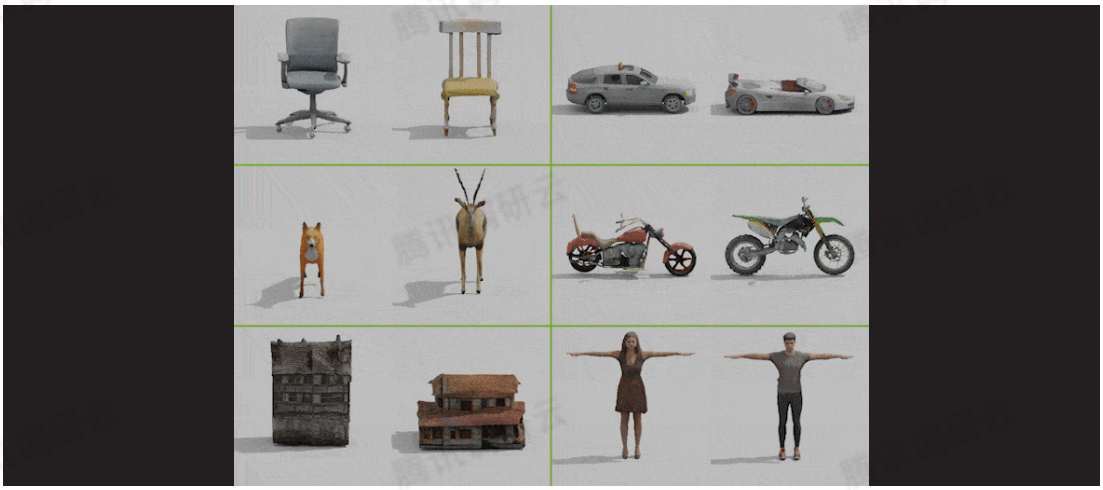
10 NVIDIA 研究团队构建 AI 模型，为虚拟世界填充 3D 物体和人物，NVIDIA英伟达微信公众号，https://mp.weixin.qq.com/s/ZYSEou1ki0a4JVY2Nv8_SA，2022-9-26

随后，英伟达发布了 Magic3D，同样是从文字描述中生成 3D 模型的 AI 模型。在输入诸如“一只坐在睡莲上的蓝色毒镖蛙”这样的提示后，Magic3D 在大约 40 分钟内生成了一个 3D 网格模型，并配有彩色纹理。Magic3D 的转换分为两步，先将低分辨率生成粗略模型，再优化为高分辨率的精细模型，而且使用 Instant NGP 的哈希特征编码，高分辨率图像特征表示的计算成本大大降低，因此比谷歌 DreamFusion 的 3D 生成速度更快、效果更好。同时，与文本到图片的扩散模型类似，Magic3D 可以在几代生成图像中均保留相同的主题（一般称为“一致性”），并将 2D 图像（如立体派绘画）的风格应用于 3D 模型。研究人员表示，模型完善后有望使 3D 合成技术进一步普及，并激发每个人在 3D 内容上的创造力。所产生的技术可以加速游戏和 VR 应用的开发，可能最终会在电影和电视的特效制作中实现落地应用。



图：Magic3D（第 1、3 列）与 DreamFusion（第 2、4 列）对比

借助 AI 从图片直接生成 3D 也是一条切实可行的路径。英伟达在 12 月的 NeurIPS 上展示了 AIGC 模型 GET3D（Generate Explicit Textured 3D 的缩写），具备生成具有显示纹理的 3D 网格的能力，可根据其所训练的建筑物、汽车、动物等 2D 图像类别，即时合成 3D 几何体。AIGC 生成的物体具有高保真的纹理和复杂的几何细节，并且以常用图形软件应用中所使用的三角网格格式创建，这使得用户可以十分轻松地将这些几何体导入 3D 渲染器和游戏引擎，以进行后续编辑。可用于构建为游戏、机器人、建筑、社交媒体等行业设计的数字空间，比如建筑物、户外空间或整座城市的 3D 表达。该模型是在 NVIDIA A100 Tensor Core GPU 上使用从不同相机角度拍摄的约 100 万张 3D 几何体的 2D 图像训练而成，每秒可生成约 20 个物体。



图：NVIDIA GET3D 基于 AI 生成的具有光影效果的 3D 物体模型

结合团队的另一项技术，AI 生成的模型的可编辑性会大幅加强。6 月发布的成果显示，在分离 3D 物体中的材质和光照方面取得了重大进展，这反过来又提高了 AI 生成几何体的可编辑能力，创造者们能够轻松地对在场景中移动的物体替换材质或者调整光照。

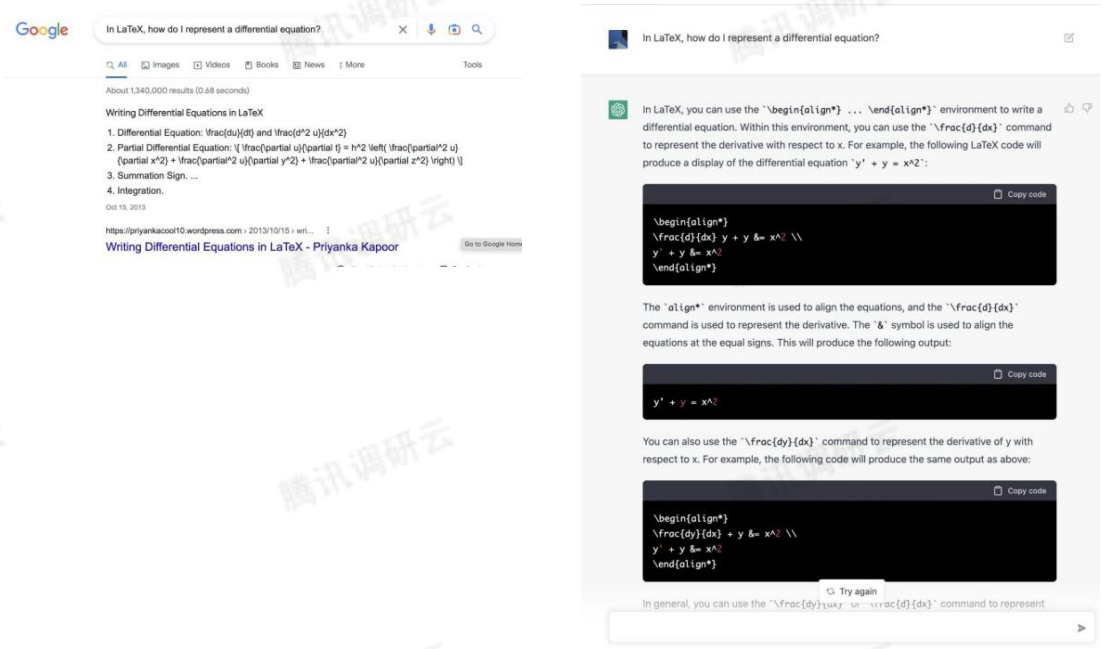
4、聊天机器人和数字人成为新的、更包容性的用户交互界面，不断拓展应用领域

（1）聊天机器人：2022 年 12 月 1 日，美国人工智能研究公司 OpenAI 研发的 AI 聊天机器人产品 ChatGPT 正式对外开放，一经推出便火遍全网，截至 12 月 12 日已拥有超过 100 万名用户。而 ChatGPT 正是典型的文本生成式 AIGC。

ChatGPT 采用 Web 浏览器上的对话形式交互，不仅能够满足与人类进行对话的基本功能，能够回答后续问题、承认错误、质疑不正确的前提和拒绝不适当的请求。而且，根据各领域的用户在使用后反馈，ChatGPT 还可以驾驭各种风格和文体，且代码编辑能力、基础脑力工作处理能力等一系列常见文字输出任务的完成程度也大大超出预期。比如根据关键词或提问，生成剧本、发言稿等多种文书。因此，业内人士宣称 ChatGPT 已经大幅超越过去的 AI 问答系统。[11]

11 https://mp.weixin.qq.com/s/Yr-8HSVFQj26X-n75cyKvw

未来随着性能的进一步提升，对话式 AIGC 在搜索、知识传播等领域有很大的应用空间。正如 Open AI 的 CEO Sam Altman 在 Twitter 上说过，AIGC 的最终目标是做一个类似于像新时代的搜索引擎。[12] 目前，从 ChatGPT 展示出来的内容输出质量和内容覆盖多维度，已经可以直面“搜索引擎”与“问答社区”。根据《纽约时报》报道，谷歌及其母公司 Alphabet 的首席执行官 Sundar Pichai 参加了几次围绕谷歌人工智能战略的会议，并指示公司的许多团队重新集中精力解决 ChatGPT 对其搜索引擎业务构成的威胁。近日，谷歌并为 ChatGPT 带来的威胁发布“红色警报”，着手进行紧急应对。[13]



图：有网友将 Google 和 ChatGPT 进行了对比，分别在 Google 搜索和 ChatGPT 中提出了同一个问题

（2）数字人：数字人、虚拟机器人是数字智能体（digital agent），作为新的交互形式，目前已有很多应用，包括元宇宙应用中的 NPC 虚拟角色、用户虚拟替身 / 虚拟形象（avatar），主要分为真人驱动和 AI 驱动两种。这里主要阐述由 AI 驱动的数字人。

12 https://m.huxiu.com/article/739676.html?type=text&visit_source=home_page&collection_id=139
13 https://www.thepaper.cn/newsDetail_forward_21282873



图：应用了 AIGC 的腾讯 2D 数字人自动化管线生产流程（引用自腾讯研究院《数字人产业发展趋势报告》）

首先，AIGC 大大提升了数字人的制作效能。用户可以上传照片 / 视频，通过 AIGC 生成写实类型的数字人，具有生成过程较短、成本低、可定制等特点。基于 AIGC 的 3D 数字人建模已经初步实现产品化，目前精度可以达到次世代游戏人物级别，优势是可以开放程序接口，对接各种应用，因此潜力较大，范围较广，特别是 C 端的应用，可以通过便捷化的上传图片即可生成 3D 数字人面部模型。未来发展方向是通过算法驱动提高精度，优化建模效果，比如：偏移矫正、阴影修复等。在用户 avatar 生成、创建方面，AI 引擎可以分析用户的 2D 图片或 3D 扫描，然后形成高度逼真的仿真渲染，同时结合脸部表情、情绪、发型、年龄特征等因素让用户的虚拟形象更具活力。目前，Meta、英伟达等众多科技公司已经在利用 AI 技术帮助用户在虚拟世界打造虚拟化身，例如英伟达的 omniverse avatar 可以生成、模拟、渲染可互动的虚拟形象。

其次，AIGC 支撑了 AI 驱动数字人多模态交互中的识别感知和分析决策功能，使其更神似人。自然语言处理是数字人的大脑，直接影响交互体验，而计算机视觉决定了数字人面部表情和肢体动作。目前主流的方式是围绕 NLP 能力通过文本驱动，本质是通过 ASR-NLP-TTS 等 AI 技术进行感知 - 决策 - 表达的闭环来驱动数字人交互。计算机视觉（CV）目前数字人声唇同步技术相对完善，在游戏中已经大量应用。

	技术	阶段	作用和目的
语音理解	ASR	感知阶段	将人的语音转换为文本
	NLP	决策阶段	处理并理解文本，以对话能力为核心，为数字人的大脑
	TTS	表达阶段	将需要输出的文本合成为语音
动作合成	AI驱动嘴形动作	表达阶段	建立输入文本到输出音频与输出视觉信息的关联映射，主要是对采集到的文本到语音和嘴形视频（2D）/ 嘴形动画（3D）的数据进行模型训练，得到相关模型，并智能合成
	AI驱动其他动作	表达阶段	动作是采用随机策略或者脚本进行预设，需要人工配制描述性的数据或者标签

表：AIGC 支撑 AI 驱动数字人多模态交互

数字人进化的趋势，一是融入会话式 AI 系统（Conversational AI），给传统的 Siri 等虚拟助手、智能客服等聊天机器人以一个具象化、有亲和力的人类形象，提升交流中情感的连接，有望给这一领域带来更大的市场前景。据机构统计，2021 年会话式 AI 的全球市场规模为 68 亿美元，预计到 2026 年将增长到 184 亿美元。随着线上空间日益丰富，更多普通用户也希望拥有自己的个性化虚拟形象，因此，数字人进化的第二个方向是制作工具更丰富、更易用。例如 Epic 在虚幻引擎中集成的虚拟人工具 Metahuman，用户可以在系统提供的基础形象模板上修改参数，仅用 30 分钟就能“捏”成独一无二的形象。可调节内容既包括整体的肤色、身材，也包括细节的面庞轮廓、五官大小等。未来，对话式 AI 系统、先进的实时图形处理等技术的结合，将使得数字人、虚拟助手、虚拟伴侣、NPC 等数字智能体（digital agent）能够逼真地模仿人类的音容笑貌，变得更加智能化、人性化。这将带来更复杂的、自然交互的 AI 虚拟角色，除了模仿人类的语言表达，还具有表情、肢体语言、情绪甚至物理交互等能力，给用户提供更直

观的、更具沉浸感的数字化体验。可以说，数字人等新型 AI 角色将决定 VR/AR、元宇宙等未来互联网应用的体验质量和吸引力。

5、AIGC 将作为生产力工具来推动元宇宙发展

元宇宙最令人兴奋的特征之一是它的可扩展性潜力，它将极大扩展人类在虚拟世界的生存空间。但是问题在于，谁来建造这个巨大的世界？它将如何填充内容？如今，创建高保真 3D 模型或虚拟人通常需要大量的计算机科学家和数字艺术家。面对这个空间所需要的大量的数字内容，单靠人工来设计和开发根本无法满足需求，而 AIGC 将是新的元宇宙内容生成解决方案。近期，LinkedIn 做了一个调查，向 LinkedIn 用户征集他们最希望了解哪个元宇宙相关的技术主题，76% 的受访者选择了 AIGC。[14] 只有通过 AIGC，元宇宙才可能以低成本、高效率的方式满足海量用户的不同内容需求。



图：扎克伯格在元宇宙的岛屿上，可以通过发出语音命令变换天气、生成创建海滩等不同场景

14 https://www.linkedin.com/pulse/how-generative-ai-impact-metaverse-info-tech-research-group/?trk=public_post

首先，AIGC 为构建沉浸式的元宇宙空间环境提供了核心基础设施技术，将成为元宇宙的生产力工具。元宇宙在理论上可以用于工作、会议、游戏以及生活社交等各种人类活动，因此需要在元宇宙空间中创建各种各样的活动对象、建筑和活动环境。但是，在过去，为了构建这些数字环境，需要开发团队半手工地创建每一个部分，从山丘到大海，用鼠标拖动来放置树木或家具，而现在 AIGC 已经可以实现创建逼真的 3D 虚拟空间环境、虚拟人物，并且效率和成本可以满足大规模的元宇宙空间环境创建。

在元宇宙发展的过程中，游戏是虚拟世界重要的载体，将需要大量数字原生的场景。AIGC 可以极大提高数字原生内容的开发效率，降低游戏开发成本。具体来说，剧情、角色、头像、道具、场景、配音、动作、特效、主程序未来都可以通过 AIGC 生成。通过 AIGC 加速复刻物理世界、进行无限内容创作，从而实现自发有机生长。例如，当用户朝任何方向行走时，游戏 Minecraft 的世界地图会通过 AIGC 生成模型无限扩展。NVIDIA 也开发了 AIGC 生成模型，仅通过将 2D 视频输入来创建语义标记和可修改的 3D 环境。正如 Unity 人工智能高级副总裁 Danny Lange 所言：“市政府想要通过城市的数字孪生来做一些规划。而城市数字孪生很难聘请大量的技术开发人员来编写代码。而此时就是 AIGC 的用武之地。”[15]

而且 AIGC 能够释放开发人员的生产力，帮助开发人员编写代码，使开发人员可以花更少的时间编写代码，而花更多的时间来表达需求，最终实现生产力的大幅提升，以及元宇宙空间环境的建成速度。

其次，AIGC 将作为生产力工具为元宇宙用户提供个性化内容体验。大量个性化、多元化的数字内容是吸引人们进驻元宇宙的重要原因，而这些内容如果仅通过人工方式来生产既成本高又耗时长，肯定远远无法满足用户的海量需求，因此，需要 AIGC 作为生产力工具来满足和填充 Metaverse 等内容匮乏的虚拟世界。NFT 便是 AIGC 的一个很好的例子。Bored Ape Yacht Club 和 CryptoPunks 等 NFT 项目开始于创作 10,000 件独特的艺术品。如果没有 AIGC 的支持，这项工作将需要一个数字艺术家团队数周才能完成，然而这两个项目都是由两个软件开发人员完成的。

15 https://aibusiness.com/ml/unity-s-ai-chief-on-generative-ai-metaverse-and-gaming

同时，AIGC 将赋予用户更多的创作权力和自由，促进创新并提升元宇宙的用户体验。比如用户可以使用手机拍摄的一系列照片，通过 AIGC 工具生成一个可以使用的 3D 渲染图。采用这种创造内容的方式，我们可以想象未来的元宇宙体验将不再完全由开发人员构建，而是利用 AIGC 响应用户的输入按需生成。

最后，AIGC 也将作为用户交互界面的一部分在元宇宙中发挥作用。元宇宙中不仅有用户，还应该存在很多类似 NPC 的智能体为用户提供交互服务，这些智能体可以由 AIGC 生成并驱动。ChatGPT 便可以作为语言模型中集成到各种智能体中充当元宇宙中高度逼真的虚拟助手或伙伴，执行“智能”动作和更为复杂的任务。例如，Meta 已经提出了创建智能助手的计划，用户可以向助手发出简单的语音指令继而创建他们想要的虚拟现实设置类型。[16]

16 https://www.artificialintelligence.news/how-will-ai-power-the-metaverse/

AIGC

产业端

产业端：
合成数据牵引人工智能的未来

● 本章主笔：　曹建峰　腾讯研究院高级研究员
陈楚仪　腾讯研究院高级研究员

3

随着 AIGC 技术持续创新发展，基于 AIGC 算法模型创建、生成合成数据（synthetic data）迎来重大进展，有望解决 AI 发展应用过程中的数据限制，进一步推动 AI 技术更广泛的应用。因此，业界非常看好合成数据的发展前景及其对人工智能未来发展的巨大价值。Forrester、埃森哲（Accenture）[17]、Gartner、CB Insights[18] 等研究咨询公司都将合成数据列为人工智能未来发展的核心要素，认为合成数据对于人工智能的未来而言是“必选项”和“必需品”。例如，Forrester 将合成数据和强化学习、Transformer 网络、联邦学习、因果推理视为实现人工智能 2.0 的五项关键技术进展，可以解决人工智能 1.0 所面临的一些限制和挑战，诸如数据、准确性、速度、安全性、可扩展性等。[19]Gartner 预测称，到 2030 年合成数据将彻底取代真实数据，成为 AI 模型所使用的数据的主要来源。[20]MIT 科技评论将 AI 合成数据列为 2022 年十大突破性技术之一，称其有望解决 AI 领域的数据鸿沟问题。[21] 数据是人工智能的燃料和驱动力，合成数据将极大拓展人工智能发展应用的数据基础，可以认为，合成数据关乎人工智能的未来。

在概念上，合成数据是计算机模拟（computer simulation）技术或算法创建、生成的自标注（annotated）信息，可以在数学上或统计学上反映真实世界数据的属性，因此可以作为真实世界数据的替代品，来训练、测试、验证 AI 模型。简而言之，合成数据是在数字世界中创造的，而非从现实世界收集或测量而来。[22] 合成数据拥有很长的历史，在其发展过程中技术不断创新。例如，游戏引擎、3D 图形等模拟技术（simulation technology）可以创建高保真的仿真物体和仿真环境，而结合了 AI 技术的 3D-AI 技术则可以极大提升自动化生产 3D 内容的效率和保真度（fidelity）。得益于生成对抗网络（GAN）、变分自编码器（VAE）、Transformer 模型、扩散模型（Diffusion Model）、神经辐射场模型（NeRF）等不断涌现的 AI 算法，不仅合成数据的种类得到了扩展，而且其质量也不断得到提升。就目前而言，合成数据大致可分为三类：表格数据 / 结构化数据，图像、视频、语音等媒体数据，以及文本数据。[23] 这几类合成数据在多个领域都有应用。AIGC 技术的持续创新，让合成数据迎来新的发展契机，开始迸发出更大的产业发展和商业应用活力。

17 https://www.accenture.com/_acnmedia/Thought-Leadership-Assets/PDF-5/Accenture-Meet-Me-in-the-Metaverse-Full-Report.pdf
18 https://www.cbinsights.com/research/report/ai-trends-2022/
19 https://www.techrepublic.com/article/aws-ibm-google-and-microsoft-are-taking-ai-from-1-0-to-2-0-according-to-forrester/
20 https://www.gartner.com/en/newsroom/press-releases/2022-06-22-is-synthetic-data-the-future-of-ai
21 https://www.technologyreview.com/2022/02/23/1044965/ai-synthetic-data-2/
22 https://blogs.nvidia.com/blog/2021/06/08/what-is-synthetic-data/
23 https://medium.com/@adrialsina/is-synthetic-data-the-future-of-ai-exploring-enterprise-use-cases-eb0f7152fe9e

1、合成数据为 AI 模型训练开发提供强大助推器，推动实现 AI2.0

人工智能的发展应用离不开数据，但真实世界数据面临着难以获取、质量差、标准不统一等诸多问题。为此，计算机模拟技术或算法生成的合成数据，作为真实数据数据的廉价替代品，日益被用于创造精准的 AI 模型。合成数据服务商 AI.Reverie 指出，人工标注一张图片可能需要 6 美元，但人工合成的话只需要 6 美分。2019 年的一篇论文《合成数据用于深度学习》（synthetic data for deep learning）认为，合成数据是现代深度学习领域冉冉升起的最具前景的通用技术之一，尤其对于依赖于图像、视频等非结构化数据的计算机视觉技术而言；并认为合成数据对于人工智能的未来发展至关重要。[24] 而且，研究表明在 AI 模型的训练开发上，合成数据相比基于真实物体、事件或人物的数据，可以发挥同样好甚至更好的效果。[25] 总之，合成数据技术可以实现更廉价、更高效地批量生产制造 AI 模型训练开发所需的海量数据（诸如训练数据、测试数据、验证数据等等），作为对真实数据的替代或补充，将推动人工智能迈向 2.0 阶段，从本质上扩展 AI 的应用可能性。

可以说，目前人工智能仍处在 1.0 阶段（AI1.0），数据是最大掣肘，业界利用真实世界数据训练 AI 模型面临多方面问题：数据采集、标注费时费力、成本高企；数据质量较难保障；数据多样化不足，难以覆盖长尾、边缘案例，或者特定数据在现实世界中难以采集、不方便获取；数据获取与使用、分享等面临隐私保护挑战和法规限制，等等。这些数据方面的限制在很大程度上阻碍了人工智能更广泛的应用和部署。

合成数据有望解决这些问题，推动人工智能迈向 2.0 阶段（AI2.0），可以在更大程度上拓展人工智能的应用。具体而言，对于人工智能而言，合成数据可以发挥诸多价值：（1）实现数据增强和数据模拟，解决数据匮乏、数据质量等问题，包括通过合成数据来改善基准测试数据（benchmark data）的质量等；（2）避免数据隐私 / 安全 / 保密问题，利用合成数据训练 AI 模型可以避免用户隐私问题，这对于金融、医疗等领域而言尤其具有意义；（3）确保数据多样性，更多反映真实世界，提升 AI 的公平性，以及纠正历史数据中的偏见，消除算法歧视；（4）应对长尾、边缘案例，提高 AI 的准确性、可靠性，因为通过合成数据可以自动创建、生成现实世界中难以或者无法采集的数据场景，更好确保 AI 模型的准确性；（5）提升 AI 模型训练速度

24 https://arxiv.org/pdf/1909.11512.pdf

25 https://news.mit.edu/2022/synthetic-data-ai-improvements-1103

和效果。总之，利用合成数据可以更廉价、更高效、更准确、更安全可靠地训练 AI 模型，进而极大扩展 AI 的应用可能性，将人工智能推向新的发展阶段。

2、合成数据助力破解 AI“深水区”的数据难题，持续拓展产业互联网应用空间

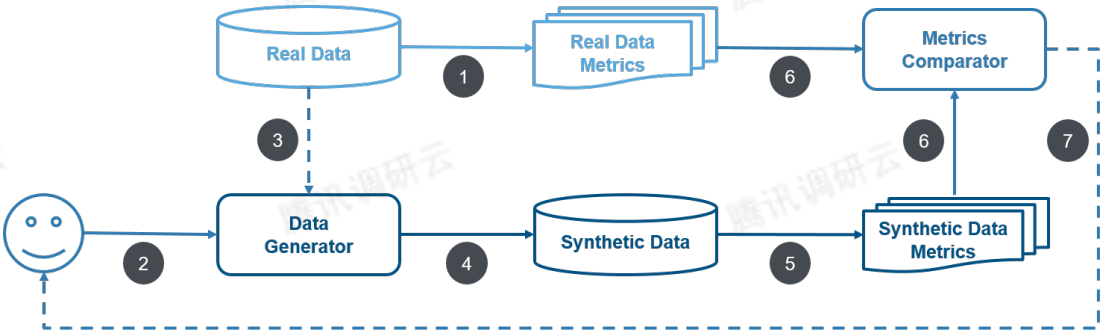
合成数据早期主要应用于计算机视觉领域，因为计算机视觉被广泛应用于自动驾驶汽车、机器人、安防、制造业等领域，在这些应用场景中打造 AI 模型都需要大量的被标注的图像、视频数据。但获取现实数据往往并非易事。以自动驾驶汽车为例，由于实际道路交通场景千变万化，让自动驾驶汽车通过实际道路测试来穷尽其在道路上可能遇到的每一个场景是不现实的，必须借助于合成数据才能更好地训练、开发自动驾驶系统。为此，自动驾驶企业开发了复杂的仿真引擎来“虚拟地合成”自动驾驶系统训练所需的海量数据，并高效地应对驾驶场景中的“长尾”问题和“边缘案例”。例如，腾讯自动驾驶实验室开发的自动驾驶仿真系统 TAD Sim 可以自动生成无需标注的各种交通场景数据，助力自动驾驶系统开发。在安全的、合成的仿真环境中，计算机可以模拟任何人类想象得到的驾驶场景，诸如调节天气状况、添加或移除行人、改变其他车辆的位置等等。可以说，合成数据和仿真技术是自动驾驶的核心支撑技术。最早涌现的一批合成数据创业公司就瞄准的是自动驾驶汽车市场，帮助自动驾驶企业解决其在自动驾驶系统开发过程中所面临的数据和测试难题。

目前，合成数据正迅速向金融、医疗、零售、工业等诸多产业领域拓展应用。合成数据在金融服务领域的探索仍处于早期且不断拓展，并且受到咨询公司、金融巨头和监管机构的关注。

合成数据背后的生成式 AI 被 Gartner 评为 2022 年银行和投资服务领域越来越受欢迎的三项技术之一。[26] 生成性 AI 受欢迎的原因是能够通过合成数据以成本更低、易规模化、隐私保护合规的方式提供接近真实世界的数据。而在银行和投资服务领域，生成对抗网络（GAN）和自然语言生成（NLG）的应用可以在大多数欺诈检测、交易预测、合成数据生成和风险因素建模的场景中找到。例如，美国运通（America Express）利用 GAN 创建合成数据来训练、优化其进行欺诈检测的 AI 模型。谷歌利用 AI 生成的医疗记录来帮助预测保险诈骗（insurance fraud）。

26 https://www.gartner.com/en/newsroom/press-releases/2022-05-24-gartner-identifies-three-technology-trends-gaining-tr

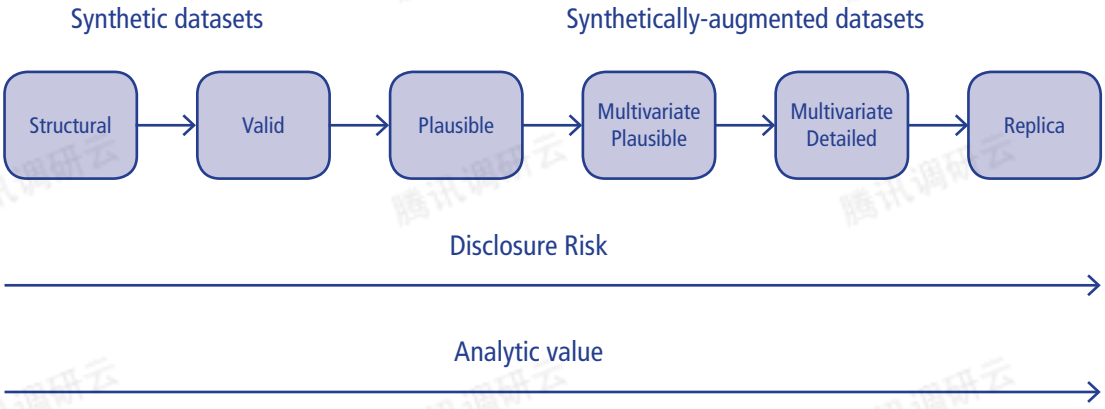
摩根大通（J.P.Morgan）2021 年 9 月在其官网发布相关研究，提出通过生成合成数据集加快金融服务领域的 AI 研究和模型开发，来改善服务体验、解决欺诈检测和反洗钱等重要问题。[27] 国外金融服务业所产生的大量真实数据因为法律限制（如欧盟的 GDPR 和美国的 CCPA）和隐私保护要求无法使用或使用存在诸多限制。合成数据创造的新样本具有真实数据的性质，增加真实数据中的罕见样本，以便更有效地训练机器学习算法。一个关键的领域是欺诈检测模型训练。由于欺诈性案件的数量与非欺诈性案件相比较十分稀少，研究人员很难有效地从可用数据中训练模型，导致无法针对欺诈性行为进行建模。然而，合成数据可以生成比实际数据中欺诈案例比例更高的合成数据样本，用于帮助改进模型训练。



图：金融数据合成过程（来源：摩根大通官网）

英国金融行为监管局（Financial Conduct Authority，FCA）也积极跟进合成数据作为一项隐私保护技术在金融行业的应用前景。FCA 在 2022 年 3 月的公开报告中指出，合成数据通过生成统计学上真实（statistically realistic）的但“人造”（artificial）的数据，为数据共享提供更多的机会。技术的应用有利于提供更普及、更高效、更好提升消费者体验的金融服务，目前的应用领域包括：金融犯罪和欺诈预防、信用评分、销售和交易、保险产品定价和索赔管理、资产管理和组合优化等。值得注意，合成数据也存在通过逆向工程来获取真实数据的风险，对数据隐私保护而言并非完美的解决方案。[28] 为此，FCA 已经通过公开征求意见来了解业界对合成数据支持创新的潜力和满足企业有效需求的看法以及潜在的限制和风险。

27 https://www.jpmorgan.com/technology/technology-blog/synthetic-data-for-real-insights
28 https://www.fca.org.uk/publication/call-for-input/synthetic-data-to-support-financial-services-innovation.pdf



图：合成数据存在真实数据泄露风险（来源：英国国家统计办公室）

在医疗领域，医疗影像的合成数据正对医疗 AI 领域的发展产生巨大推动作用。例如，创业企业 Curai 基于仿真的医疗案例数据训练了一个医疗诊断模型，表明合成的医疗数据同样可以用于支持医疗 AI 应用的开发。美国生物技术公司 Illumina 和合成数据创业公司 Gretel 合作，利用合成的基因组数据来进行医学研究，以避免限制性立法、病人同意等问题，这些问题阻碍医学研究的速度和规模。另一个典型的例子是将合成数据用于新冠肺炎相关的医学研究。为了对抗新冠肺炎，研究人员需要获取足够的病例数据来开展研究，但出于隐私保护等顾虑，研究人员在获取新冠肺炎相关的病例数据时面临不小的困难。为此，美国国立卫生研究院（NIH）和 Syntegra、MDClone 等合成数据服务商合作，基于其 Covid-19 病人病历数据库“全国新冠合作群组”（National COVID Cohort Collaborative，N3C）中超过 500 万个新冠阳性个体的病例数据，合成了不具有可识别性的替代数据，即合成的 Covid-19 数据。合成数据精确地复制了原始数据集的统计特征，但与原始数据不存在任何关联，可以被世界范围内的研究人员自由分享和使用，用于开展研究发现、疾病治疗、疫苗开发等方面的工作。

在零售领域，创业公司 Caper 可以基于一个商品的几张图片，利用 3D 模拟技术创建含有上千个图片的合成数据集，进而支持智慧零售店。在工业领域，FORD 利用游戏引擎和 GAN 来创造合成数据训练 AI 模型。此外，合成数据还可帮助应对算法歧视等 AI 伦理问题，支持打造更公平、可信的 AI 模型，因为合成数据可以帮助消除 AI 数据集中的偏见因素，支持构建更具包容性的、代表性的高质量数据集。

3、合成数据产业加快成为数据要素市场新赛道，科技大厂和创新企业抢先布局

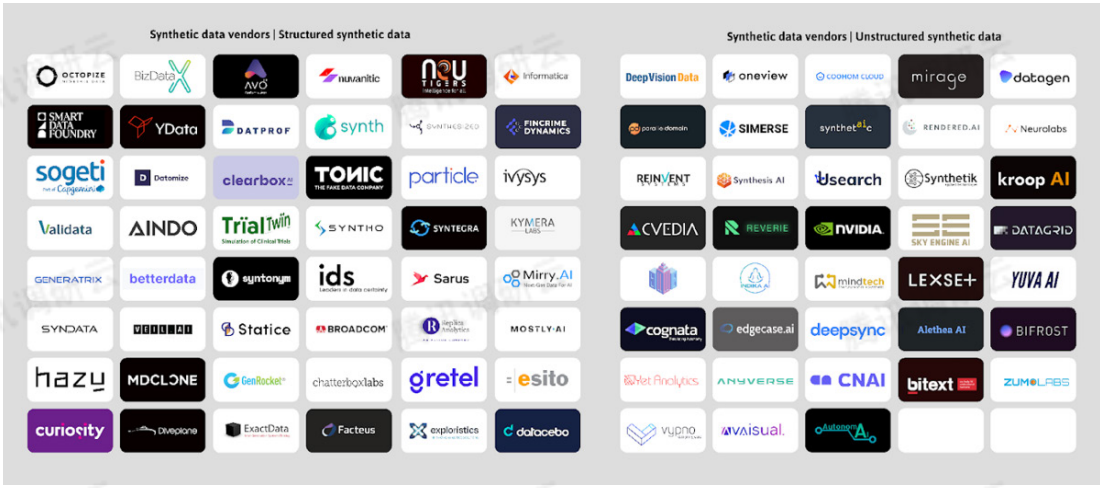
合成数据对人工智能未来发展的巨大价值使其加速成为 AI 领域的一个新产业赛道。一方面，国外的主流科技公司纷纷瞄准合成数据领域加大投入与布局。英伟达[29]是典型代表，其元宇宙平台 Omniverse 拥有合成数据能力 omniverse replicator；omniverse replicator 作为 Omniverse 平台的一部分，是为 AI 算法训练生成具有物理属性的合成数据的技术引擎，有两项代表性应用。一是在 Omniverse 平台中创建用于机器人训练的虚拟环境 Issac SIM 平台[30]，在这个虚拟环境中训练的机器人之后可以直接应用于现实世界，即这个机器人可以将在虚拟世界中的训练结果同步到现实世界的机器人身上，叠加人工智能的算法，实现机器人的大规模应用。二是 drive SIM 平台[31]，提供丰富的模拟场景，用于自动驾驶算法训练和验证，如物体识别、车道保持、泊车等自动驾驶汽车应用。Omniverse 的这两项合成数据应用可以显著加速机器人和自动驾驶的开发进程。

微软的 Azure 云服务则推出了 airSIM 平台，可以创建高保真的（high fidelity）的 3D 虚拟环境来训练、测试 AI 驱动的自主飞行器;[32] 微软还开发了可以生成合成和聚合数据集的开源工具 Synthetic Data Showcase，并创建了合成人脸数据库，和国际移民组织（IOM）合作打击人口贩卖。亚马逊在多个场景探索合成数据的应用，例如使用合成数据来训练、调试其虚拟助手 Alexa，以避免用户隐私问题；其合成数据技术 Wordforge 工具可以用来创建合成场景（synthetic scenes）；在 2022 年的亚马逊 re:MARS 大会上，其数据标注服务 SageMaker Ground Truth 推出了合成数据能力，可以自动生成标注的合成图片数据，即该工具可以就特定物体创建不同位置和不同灯光条件的图片，以及不同比例和其他变化的图片。Meta（原 Facebook）也着力于为其人工智能业务增强合成数据能力，2021 年 11 月收购了合成数据创业公司 AI.Reverie。

另一方面，合成数据作为 AI 领域的新型产业，相关创新创业方兴未艾，合成数据创业公司不断涌现，合成数据领域的投资并购持续升温，开始涌现了合成数据即服务（synthetic data as a service，SDaaS）这一发展前景十分广阔的全新商业模式。[33] 据国外研究者统计，目前全球

29 https://nvidianews.nvidia.com/news/nvidia-announces-omniverse-replicator-synthetic-data-generation-engine-for-training-ais
30 https://developer.nvidia.com/isaac-sim
31 https://blogs.nvidia.com/blog/2021/11/09/drive-sim-replicator-synthetic-data-generation/
32 https://news.microsoft.com/innovation-stories/microsoft-launches-project-airsim-an-end-to-end-platform-to-accelerate-autonomous-flight/
33 https://www.washingtonpost.com/business/ais-next-big-thing-is-fake-data/2022/06/27/d1a2e4f0-f5d6-11ec-81db-ac07a394a86b_story.html

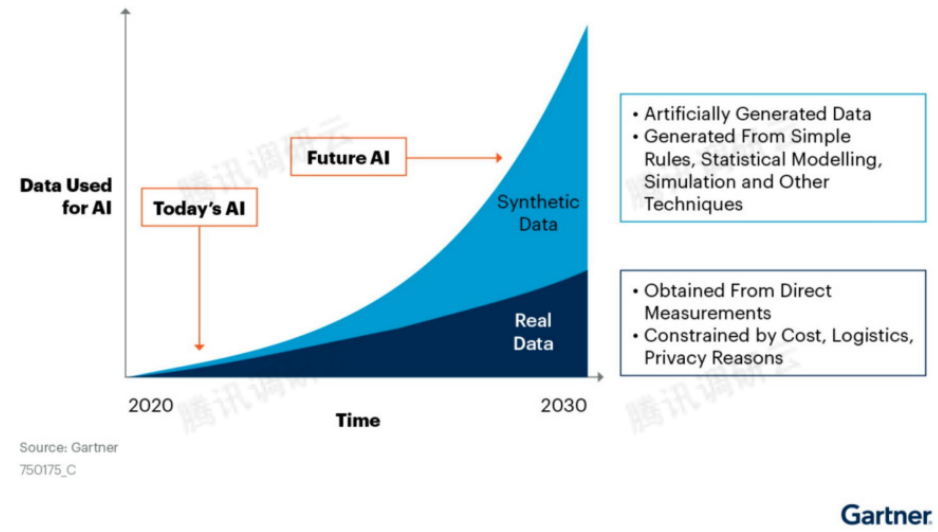
合成数据创业企业已达 100 家。[34] 比较知名、有影响力的包括 AI.Reverie、datagen、sky engine、mostly.ai、synthesis AI、gretel.ai、one view、innodata、cvedia 等等。在过去的 18 个月，公众视野中已知的合成数据公司融资总额达到 3.28 亿美元，比 2020 年高出 2.75 亿美元。合成数据的创业赛道主要涵盖非结构化数据（图片、视频、语音等）、结构化数据（表格等）、测试数据（test data）、开源服务等几大方向。其中，非结构化合成数据持续保持强劲发展势头，这主要得益于以下几个因素：第一，计算机视觉应用场景相对成熟；第二，有游戏引擎（game engines）、图像建模软件、AIGC 技术的支撑；第三，自动驾驶汽车、零售、电子游戏等快速发展的产业对合成数据有较高需求。目前结构化数据合成和测试数据合成正在迅猛发展，尤其是合成的测试数据更少受到数据隐私立法的限制，所以开始受到业界青睐。此外，合成数据开源服务也在快速发展，例如 synthetic data vault、synner、synthea、synthetig 等。



图：合成数据产业生态一览表（来源：medium 网站）

市场规模方面，根据 Cognilytica 的数据，合成数据市场规模在 2021 年大概是 1.1 亿美元，到 2027 年将达到 11.5 亿美元。Grand View Research 预测，AI 训练数据市场规模到 2030 年将超过 86 亿美元。Gartner 预测，到 2024 年用于训练 AI 的数据中有 60% 将是合成数据，到 2030 年 AI 模型使用的绝大部分数据将是人工智能合成的。可以预见，合成数据作为数据要素市场的新增量，在创造巨大商业价值的同时，也有望解决人工智能和数字经济的数据供给问题。

34 https://elise-deux.medium.com/everything-that-happened-in-the-synthetic-data-space-in-2022-c5d6cb5aaf06



图：合成数据发展预测（来源：Gartner）

4、合成数据加速构建 AI 赋能、数实融合的大型虚拟世界

合成数据指向的终极应用形态是借助游戏引擎、3D 图形、AIGC 技术构建的数实融合的大型虚拟世界。大型虚拟世界对于人工智能的核心价值在于，为测试、开发新的人工智能应用，提供了一个安全、可靠、高效以及最重要的是——低成本的、可重复利用的环境。展望未来，可以从三个方面来理解大型虚拟世界为什么成为 AI 数实融合的关键场景。

第一，大型虚拟世界可以提供人工智能开发所需的数据和场景，为 AI 应用开发提供“加速度”。游戏开发者、发行者已经意识到了这一机遇，即为 AI 系统设计游戏般的虚拟世界。例如，《星际争霸》、《我的世界》等游戏是人工智能的理想的测试场地，3A 游戏场景资产也是生成性 AI 工具的重要潜在应用。

第二，大型虚拟世界为各行各业训练开发 AI 提供了试验田。这个方面的一个典型例子就是用于自动驾驶算法测试的虚拟仿真平台。谷歌、英伟达、腾讯等国内外科技公司都在大力布局自动驾驶仿真业务，即提供一个仿真的合成世界来训练、测试自动驾驶算法。在国内，腾讯自动

驾驶实验室开发的自动驾驶仿真平台 TAD Sim，可以让自动驾驶算法在城市级别的虚拟仿真世界中进行测试和学习，极大降低了自动驾驶汽车的研发成本。在国外，自动驾驶企业 Waabi 希望主要通过虚拟仿真而非实际道路测试来训练其自动驾驶系统，构建了 Waabi World 这一用于自动驾驶系统的大型虚拟世界，其核心即是合成数据和合成场景。Waabi World 不仅可以接近实时地模拟汽车的传感器，而且模拟的环境可以和人工智能之间可以相互做出反应。这是非常重要的，因为自动驾驶汽车不仅需要感知世界，还需要安全地采取行动。大型虚拟世界提供了这样一个安全有效且可重复使用的试验田。

第三，在大型虚拟世界中通过 AI 连通虚拟与现实，实现 AI 数实融合。在大型虚拟世界，如游戏虚拟世界或者其他的仿真的 3D 合成世界（simulated synthetic world）中训练 AI 虚拟机器人，让其自我学习、进化，然后部署到现实世界中，产生现实的价值。此外，大型虚拟世界对强化学习和其他 AI 算法也具有巨大价值，例如，游戏世界是开展 AI 强化学习和多智能体协作的最佳载体；也可以在大型虚拟世界中利用虚拟人开发更宜居的建筑设计；大型虚拟世界如果可以仿真物理现象，就可进行虚拟化学实验，相比人类利用真实的化学药品开展实验，软件可以利用虚拟的化学药品更高效、更安全、更大规模地开展化学实验。总之，大型虚拟世界对于支持人工智能研究开发、促进数实融合的巨大可能性，值得持续深挖和不断探索。

AIGC

社会端

社会端：
AIGC 助力可持续社会价值

● 本章主笔：　王焕超　腾讯研究院研究员

4

“Generative AI（生成式 AI）有什么社会价值？”

当本报告研究者将这一问题输入到前段时间大火的对话机器人 ChatGPT 中， 得到了这样的回复：“生成式 AI 是一种人工智能技术，它可以使用训练数据来生成新的内容。这种技术可以用来解决各种问题，具有很高的社会价值。
例如，它可以用来生成文本、图像、音频和视频，这些内容可以用来帮助人们了解世界，也可以用来提高工作效率。
此外，生成式 AI 还可以用来解决一些非常复杂的问题，比如提供新的医疗方案、帮助制定更有效的政策、或者甚至帮助人类更好地了解自己。
总的来说，生成式 AI 为人类提供了很多有价值的机会，有助于人类更好地理解世界，并创造出更多的价值。”

整体来看，ChatGPT 的回答概括性强，而且较为准确地总结了生成式 AI 的价值，尽管还有些“蜻蜓点水”且缺少例证，但在经历对话机器人前辈们被冠以“人工智障”的名号许久之后，ChatGPT 终于以崭新的面貌一雪前耻，并将 AI 生成（文本）内容推向了新的发展阶段。

“生成”显然是 2022 年的主线。历经数年蛰伏与积累之后，在这一年，AIGC 得到快速发展，并迅速成为科技圈与创投圈的热词，也在应用领域收获数百万量级的用户。那么，来势凶猛的 AIGC 会创造何种社会价值，对人类又有什么意义？

本报告认为，AIGC 的社会价值体现为革新数字内容与艺术创造领域，并将辐射到其他领域和行业，孕育新的技术形态与价值模式，甚至会成为通往 AGI 的可能性路径：

1、AIGC 以高效率、低成本满足个性化需求，完成基础性工作，释放人类创造力，推动艺术创造领域与基础概念革新；

2、AIGC 的生成能力将延伸、辐射深入到其他领域（如医疗、教育、传媒、影视、工业，以及元宇宙、数字人领域），通过进一步互动结合，催化新的业态与价值模式，形成“AIGC+”效应；

3、经由应用层积累的巨量数据，推动大模型提升计算复杂度，AIGC 将成为通向通用人工智能（Artificial general intelligence）的可能性路径。

1、元价值：解放人类创造力，革新艺术领域

AIGC 最基本的能力是生成内容（generating content），包括文本、图像、音频、视频、代码、3D 内容或者几种媒介类型相转换、组合的“多模态内容”。AIGC 这种“生成”的能力，使其能够以远超人类效率、低成本、较高质量创造海量的内容。

以 AI 绘画模型为例，一张人类画手可能要花费数日来完成的画作，只需要一串关键词和几分钟的时间，Stable Diffusion、MidJourney 等 AI 绘画模型就能生成。这首先带来的是人类创造力的解放，对于大量重复性、机械化的低层次任务，可以由 AI 来完成，从而释放人类创造能力，让人类创作者的脑力和时间用于更具创造性的任务中。

在《天才与算法》一书中，作者马库斯·杜·桑托伊论证了“图灵测试”的过时，并提出一个评价人工智能的智能程度的新方式——洛夫莱斯测试。通过这一测试的标准是，算法需要创作一件艺术作品，人类程序员无法解释该算法的工作原理，但是整个过程是可复现的，并且该人工智能应具备超越人类程序员或数据集创造者的创造力。

作为图灵测试的补充，洛夫莱斯测试开创式地引入了“艺术创作”这个维度，因为艺术创作背后隐含了表达、思考、创造，是自主意识的投射，也是存在实证。按照这一评价维度来看，现在的 AIGC 或具体表述为 AI 绘画模型、AI 辅助写作等技术，它们确实已经在从事“创造”，但却不是真正具有创造力，仍然是基于深度学习、机器学习的“模仿式创新”。

换句话说，它们可以“作”一幅画，“写”一篇文章，但仍然不能真正理解这幅画和这篇文章的真正意义。至少从现阶段来讲，AIGC 背后真正的创造者仍然是人类。艺术是自由意志的表达，在拥有自主意识、思维能力之前，AI 所创造的作品皆来源于人类的创作意图。

问题在于，即便 AIGC 不是真正的“创造”，难道就对创造没有意义了吗？答案是否定的，如前文所述，正是因为 AI 解决了基础性工作，实现了人类创造力的解放。这就是它最大的社会价值之一，或者说“元价值”。发展 AIGC，本意并不是真的要发明人类之外的另一个创造者，而是为了提升人类本来的创造力。

追溯艺术观念史，柏拉图最早曾给艺术的本质下了一个定义：艺术即模仿。这一点，AI 和人

类的艺术创作都是相通的。人的艺术创作，是对物理对象的摹画、表现，建基于对无数前人的无数作品的吸纳、品鉴、学习与模仿。AI 的艺术创作基于对成千上万的人类作品的归纳学习、演绎创造，原理同样是模仿。

艺术这种创造式行为的基本原则就是，在模仿基础上打破已有规则或引入新规则，以求出现新的路径与风格，这就是变革式创新的意义。而 AIGC 的角色就是扮演破舰者，一方面它取代大量基础性创作，另一方面又通过超越常规思维模式的一些整合创造，从而推动人类的创造力的提升，使线性的艺术领域变革成为指数性的变化。

有许多人类画师在接到需求时，首先会把关键词输入到 AI 绘画模型，让 AI 先输出一版或多版方案，然后参考 AI 的方案来进行创作，从而保证独创性，并使创意源源不断。微软也正计划在设计软件中提供 AIGC 内容，以为设计者提供灵感和素材来源。同样，在文字创作领域，也有许多作者会通过 AI 基于主题生成写作框架，自己再去填充内容。这种变革同样发生在其他领域，比如围棋界。AlphaGo 横扫了人类围棋界之后，现在有许多围棋选手开始模仿、学习 AI 的下棋风格，来改造自己的行棋方式，排名很快得到了提升。

AIGC 对创造行业的革新势在必行。AI 独特的艺术风格与创造模式，对缓慢演进的艺术领域无疑是一种冲击，而这种冲击将转化为变革的力量，很可能涌现出新的艺术创造流派和风格。同时，通过在应用端的快速延伸，AIGC 也在逐步颠覆人类的传统审美概念。

总结而言，AI 虽然不拥有真正的创造力，但是它让使用者拥有了新的创造力，从而拓展艺术创造的边界和可能性的新方向。作为一种新的技术类型，AIGC 还在迅速发展，在可预期的未来，除了“艺术创造”，“知识”“学习”“创作”等等概念也将发生变化，它们将成为“了解 AI 的程度”“善于使用 AI 的能力”“能够用 AI 增强自身创作的手法”等等。

在这样的冲击过程中，就业替代的问题也值得关注。诚然，随着 AI 绘画、AI 辅助写作的发展，一些基础性的人类岗位将被 AIGC 取代，但从人类社会的技术发展史来看，尽管新技术在短期内会取代一些岗位，比如蒸汽机的发明与纺织工人的失业等等，但是随着它对生产力的整体性促进，一些新的业态和模式也将出现，从而创造更多新的生产性岗位和就业机会，因此不必过度担心。

2、AIGC 社会价值案例

AIGC 强大的生成能力，也将延伸、辐射至其他领域。比如在医疗行业，AI 生成治疗方案、AI 生成语音帮助失声者“开口说话”；在传媒行业，虚拟主持人、写稿机器人、AI 语音生成播报等；在设计领域，AI 快速生成大量案例，帮助设计师产生灵感等。同时，AIGC 也构成了数字人和元宇宙这些新兴技术形态的基础能力。本部分讲述两个具体案例以诠释 AIGC 的社会价值，分别为 AI 语音生成和 AI 修复文物。

（1）AI 语音生成帮助病人“开口说话”

2022 年 6 月 15 日，彼得 · 斯科特 - 摩根去世，终年 64 岁。他的死在全世界范围内都造成了冲击，因为他是全世界第一个“赛博格”（Cyborg）。

从通俗意义上，可以将“赛博格”理解为“半机械人”，即用机械替换部分身体器官并维持正常机能。在经过了多重手术与改造之后，彼得 · 斯科特 - 摩根终于将自己打造为一个赛博格。之所以要费这么大的周折，是斯科特 - 摩根自身情况所限——2017 年，他被确诊为肌萎缩侧索硬化症（ALS），俗称渐冻症。在此之前，这种病症最知名的患者是物理学家史蒂芬 · 霍金。

按照目前的医疗技术水平，这种病症无法完全治愈。病人的预期寿命为 2-5 年，由于中枢神经系统内控制骨骼肌的运动神经元退化和死亡，患者不同身体部位的肌肉会逐渐萎缩，并且失去移动手臂、双腿和身体的能力，也无法说话、进食甚至是呼吸。由于病人的大脑运作始终是正常的，所以会保持清晰的思维和意识，也将清醒地经历渐冻症带来的肌体摧毁过程。

面对病症的折磨，作为科学家和机器人博士的彼得 · 斯科特 - 摩根并没有选择认命，出于对技术的了解和信仰，他很快做出了一个决定：把自己改造为一个赛博格。为了实现这一计划，他首先接受了三重造口术：胃造口术、结肠造口术和膀胱造口术，以保障营养物质的供给。因为渐冻症病人也无法控制喉咙部位，为了解决呼吸问题、避免窒息风险，他也接受了全喉头切除手术。

问题就在这里，如果接受了这个手术，斯科特 - 摩根将无法说话，也无法发出自己的声音。对于一个到全球各地巡讲的科学家来说，丧失表达能力无疑比患病更为痛苦。所以，在他还能说话的时候，就开始为此准备了。他联系了 CereProc 的首席科学家 Matthew Aylett 博士，

CereProc 是专门负责创建文本并提供语音解决方案的公司。此前，他们也为接受了喉癌手术的电影评论家罗杰 · 艾伯特重建了声音。

在之后一年左右的时间里，彼得 · 斯科特 - 摩根持续在录音棚里工作，录制了超过 15 个小时的音频，1000 多个词组，包含不同情绪的语料。这些材料构成了 AI 的训练数据库。在接受全喉切除手术三个月前，他的电子合成声音终于制作完成。对于人的表达来说，声音只是构成部分之一，对方要获知你的想法，还需要借助手势和表情。

通过扫描自己的脸部，斯科特 - 摩根生成了 3D 的虚拟动画人像。在进行公开讲话时，他就会在胸前的屏幕上展示这个形象。为了提升交流效率，这套系统使用 ACAT（上下文辅助感知工具包）技术来提升打字技术。霍金生前也使用过 ACAT 技术，AI 能够学习使用者的表述习惯，通过上下文感知来预测下一个词会输入什么。



图 : 彼得生成的 3D 虚拟动画人物

在这个案例中提到合成声音技术，已经较为成熟。其他 AI 语音合成工具如 Voicemod、Veritone、Descript 和 Resemble AI 也得到广泛运用，它们已经成为需要重塑声音的潜在途径。

除了帮助失声者“开口说话”，这项技术广泛运用于影视行业。此前，《星球大战》达斯 · 维德的配音演员决定让 AI 接替自己的声音。根据媒体报道，时年 91 岁的琼斯签署了一份协议，允许乌克兰公司 Respeecher 利用 AI 技术在《星球大战》衍生剧集《欧比旺》中重现他的声音。此前，这家公司已经为《曼达洛人》《波巴 · 费特之书》中的年轻版卢克 · 天行者配了音。根据

Respeecher 的介绍，它使用了原配音演员马克 · 哈米尔多年前的广播、采访、ADR 和配音片段作为数据库，从而以数字化的方式重塑了天行者的声音。

（2）AIGC 文物修复，助力文物保护传承

AIGC 在文物修复领域的作用同样明显。在文物修复领域，技术能实现的两种路径，包括数字采集和数字修复，前者是让文物在数字世界重塑，而后者是让文物在数字世界再造。

在数字修复方面，腾讯公司利用 360 度沉浸式展示技术、智能音视频技术、人工智能等技术手段，助力文物的数字化分析与诊疗。以文化遗产敦煌壁画的修复为例，由于壁画的种类很多，损坏、病害类型复杂，因此很难有统一的方案修复所有壁画，人工修复的成本也很大，技术手段尤其是 AI 技术的应用带来了新的机会。

针对常见的壁画病害类型，腾讯通过深度学习壁画病害数据，形成自动识别并添加图示的一整套算法，打造出高效的 AI 壁画病害识别工具。在此基础上，通过迁移学习等手段来提高算法效果，同时结合传统的图像处理方法提供系统性的解决方案。在修复过程中，基于多媒体实验室的高性能编解码引擎，腾讯还开发了沉浸式远程会诊系统，用 4K 超清画质的 360 度沉浸式画面，无死角展示文物的细节，让不同地域的专家均可清晰查看现场情况，实现无障碍远程文物会诊。

利用文心大模型的 AIGC 能力，百度让《富春山居图》得以重生，也是 AIGC 运用于文物修复的一个范例。元代黄公望的作品《富春山居图》被誉为“中国十大传世名画”，但因在民间流传甚久，清代顺治年间遭火焚，从此断为长短两卷，中间缺损的部分难以复原。百度文心运用“视觉生成达模型 + 单样本微调”的技术策略，一方面通过大量的中国山水画训练，增进模型对山水画的理解，另一方面基于《富春山居图》残存的部分微调，通过迁移学习使 AI 补全的画作与真迹一致，从而达成了良好的修复效果。

除了文物古迹，AIGC 的能力也广泛应用于老照片、老电影、演唱会的修复。一个例子是张国荣《热 · 情》演唱会的 AI 修复，在腾讯云多媒体实验室的智能影像修复技术支持下，腾讯音乐集团与宝丽金共同实现了张国荣在 2000 年的《热 · 情》演唱会的高清修复。通过智能分析检测噪声、编码失真、运动以及画面复杂度等情况，运用人工智能修复算法去除破坏画质因素，并进行多维度画质增强与提升，最终将这场经典的演唱会以高清的品质呈现在观众眼前。

AIGC发展挑战

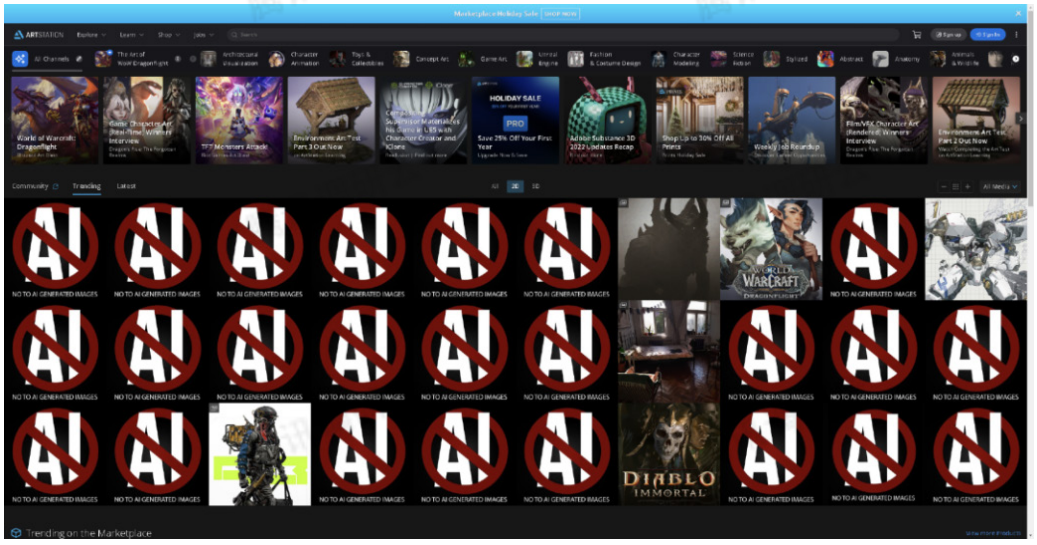
● 本章主笔：胡晓萌　腾讯研究院研究员
朱开鑫　腾讯研究院高级研究员

5

伴随 AIGC 这一最新浪潮，以及对产业、社会带来的可以预见的未来变革，过去信息技术的共性问题又有新的呈现，同时也有新的问题产生。此前，DeepMind 对预训练模型的风险做了分析和概括，共分为六个具体的风险领域：1）歧视、仇恨言论和排斥；2）真实信息危害；3）错误信息危害；4）恶意使用；5）人机交互危害；6）环境和社会经济危害。在这个分类基础上，本报告结合具体的应用场景，将分别从知识产权、安全、伦理和环境四个方面分析 AIGC 带来的挑战。

1、知识产权挑战

AIGC 已能成熟进行内容作品创作，但从著作权法角度看，AIGC 基本属于重组式创新，尚不具有真正的创造力。目前，还较为强调人机协作，可以在内容创作上可以发挥人类和 AI 彼此的最大优势。但是，AIGC 引发的新型版权侵权风险已经成为整个行业发展所面临的紧迫问题。



图：因版权争议，国外艺术作品平台 ArtStation 上的画师们掀起了抵制 AIGC 生成图像的活动，打出“NO TO AI GENERATED IMAGES”的标语

AIGC 作品既存在被他人侵权的风险，也存在侵犯他人权利的可能性。AI 要变得更加智能，一个必要的环节就是深度学习。这就需要一个庞大的人类作品数据库，让算法学会创作风格、内容题材等创作的要领。如果未经许可，复制或者通过网络爬虫爬取他人享有著作权的在线内容，在我国现行著作权法没有明确将数据挖掘等智能化的分析行为规定为合理使用的情况下，此类

行为可能构成著作权侵权。[35]

比如在 2017 年 5 月就出版了第一部 AI 诗集的微软小冰，它写诗所受训练和学习的是建国前后 519 位抒情诗人的作品。面对小冰创作的作品可能存在的侵权问题，微软方面曾做出这样的解释：“小冰内容生成能力的训练，全部来自于公开无版权权益问题的数据，或经过合作伙伴授权的数据。从创作角度来看，我们的技术包括对原创性的判断，确保了小冰所生成的各类内容作品，均符合完整的原创性要求。”

此外，AI 视频合成、剪辑领域，如果没有获得原始视频的著作权人的许可，可能因为侵犯原著作权人所享有的修改权、保护作品完整权或者演绎权而构成版权侵权行为。像是 2019 年大火的 ZAO，通过 AI 换脸软件生成新的视频，若不获得他人肖像权授权和视频内容的授权，便也存在侵权的可能。

（1）域外对于 AIGC 版权保护的基本态度

2020 年 11 月，印度版权局登记了人工智能和人类合作艺术品“Suryast”。“Suryast”是以梵高的画作《星月夜》和萨尼拍摄的一张照片为基础数据集来创作的画作。印度版权局首次将人工智能工具—RAGHAV 人工智能绘画应用程序认定为版权艺术作品的合作作者。拥有该人工智能应用程序所有权的知识产权律师安吉特 · 萨尼（Ankit Sahni）是该艺术作品的另一位作者，并已登记为版权所有人。由此，印度此前是首个在版权作品中承认人工智能程序享有作者身份的国家。但 2021 年 12 月，印度版权局已经向萨尼发送了撤销作品的通知。RAGHAV 以机器学习工程师拉加夫 · 古普塔（Raghav Gupta）的名字命名，于 2019 年在萨尼资助的一个项目中开发了该应用程序。RAGHAV 代表着强大的人工智能图形和艺术可视化工具，并受过各种艺术风格作品数据的训练。

2022 年 2 月，美国版权局审查委员会再次拒绝了 Stephen Thaler 提交的 AIGC 画作注册版权的复议请求，重申根据美国《版权法》的规定，要求作品需要包含人类作者身份。因此，由

35　2019年3月26日《单一数字市场版权指令》第3条“以科学研究为目的的文本和数据挖掘”，科研机构和文化遗产机构为科学研究目的进行文本和数据挖掘，对其合法获取的作品或其他内容进行复制与提取的行为。文本和数据挖掘是指任何旨在分析数字形式的文本和数据的自动分析技术，以便生成包括但不限于模型、趋势、相关性等在内的信息。但也不包括商业化的开发利用。

人工智能生成的这幅名为“天堂最近的入口（a recent entrance to paradise）”的绘画内容，不能在美国获得版权授权。2018 年 11 月 3 日，泰勒提出申请，主张登记注册该作品的版权，同时注明作品的作者是“创意机器（Creativity Machine）”。泰勒在给美国版权局的申请中写明，该作品“是由运行在机器上的计算机算法自动创建的”。早在 2020 年 3 月，美国版权局便首次拒绝了上述 AIGC 内容作品申请的复议请求，因为泰勒没有提供证据“证明人类作者对该作品进行了充分的创造性投入或干预”。

（2）从现行著作权法角度如何对 AIGC 加以保护

首先有必要明确对人工智能产出物进行保护的必要性，只要其符合知识产权领域客体保护的形式要求。得不到知识产权保护，即无法禁止他人未经授权的利用，也无法保障创作者获得授权费用，便会丧失制度激励作用。现行的《著作权法》框架难以直接认定人工智能或电脑算法为作品的作者。但并不意味着版权法就不应当保护人工智能生成物。但我们也需要承认在文字、图像等内容领域人类的创作物和人工智能生成物在外观上通常难以区分。

现阶段各界对于人工智能技术创作物的著作权保护整体上呈现积极的态势，主流的司法实践及学术理论均持相似的立场。对于具有独创性的人工智能生成内容，可参照著作权法关于职务作品或雇用作品的规定，将创作过程中真正实施人工干预、人为编排或创制该算法的人认定为作者，并将该创作物认定为职务作品或法人作品进行版权保护。

从司法实践来，作为“2019 年中国十大最具研究价值知识产权裁判案例”——“Dream writer 案”的判决结果在业界对于人工智能创作物的著作权保护具有明显的导向作用。南山法院肯定了 Dream writer 自动生成的财经评论文章系独立创作、在外在表现上与已有作品存在一定程度上的差异，满足了“独创性”的外在要求。其次，南山法院认为 Dream writer 主创团队在“数据输入、触发条件设定、模板和语料风格的取舍上的安排与选择”是智力活动，且与涉案文章之间具有直接联系，符合《著作权法实施条例》对“创作”的定义；认为涉案文章体现了主创团队的“个性化选择与安排”，而非 Dream writer 软件的“自我意识”。进而，南山法院认定涉案文章系主创团队创作的法人作品，完成了对该文章系著作权法保护客体的论证。

实操中可行的建议是，以意思自治的原则约定相关创作物的权利归属，包括人工智能算法的基础算法协议以及人工智能所有者与使用者的协议约定。目前的现状，AI 的作品进行知识产权

保护是应该的，但作者应该是使用 AI 系统的人，不是 AI 本身。AIGC 保护的前提应该还是使用 AI 创作的人，在创作过程中是付出了智力或创造性劳动，如果没有付出，这个产品是否构成作品是值得商榷的。目前国内国外的通行作法也都是会考虑除了 AI 之外，是否有人的智力或创造性劳动。

2、安全挑战

安全问题始终是 AI 技术发展和应用中不可回避的。同样的，在 AIGC 方面也存在内容安全、技术滥用、用户隐私和身份、AI 内生安全等多个方面的安全挑战。

一是内容本身。一直以来，互联网信息空间都面临着虚假信息和信息内容安全的挑战，国内外互联网内容平台，如 Facebook、Twitter、微信、微博等都不断在提升其虚假内容和信息安全的治理能力。但随着 AIGC 内容的持续增长，虚假信息和信息内容安全的挑战也会增加。

比如，ChatGPT 上线不到一周，用户数量就已经超过了 100 万。但是用户在深度体验之后发现，ChatGPT 生成了大量让用户看起来似乎正确和合理，但仔细检查往往是错误的答案。而这些错误信息充斥在各个内容平台之上。技术专家将这种现象称为“幻觉”或“随机联想”，即 ChatGPT 等大型语言生成模型可以预测给定输入的下一个关联内容，而不管事实是否正确或者说编造了事实。[36] 普林斯顿计算机科学教授 Arvind Narayanan 对此指出：“人们对使用 ChatGPT 进行学习感到兴奋。但危险在于，除非你已经知道答案，否则你无法判断它什么时候生成的结果是错的。我尝试了一些基本的信息安全问题，答案听起来很有道理，但实际上是胡说八道。”12 月 5 日，技术问答交流网站 Stack Overflow 已暂时禁止用户分享由 ChatGPT 生成的回答，因为 ChatGPT 生成的大量错误信息充斥在 Stack Overflow 网站上。[37] 小众的技术内容交流平台尚且如此，更遑论大型的社交平台，已经有不少网友发现由 AIGC 生成的包含色情、暴力、政治敏感性等有害信息的文字图片充斥在互联网之中。

二是，对 AIGC 的恶意使用或滥用，引发的深度合成诈骗、色情、诽谤、假冒身份等新型违法犯罪行为。不法分子利用开源的 AIGC 模型或工具，可以以更低的门槛、更高的效率来制作出

36 https://venturebeat.com/ai/the-hidden-danger-of-chatgpt-and-generative-ai-the-ai-beat/

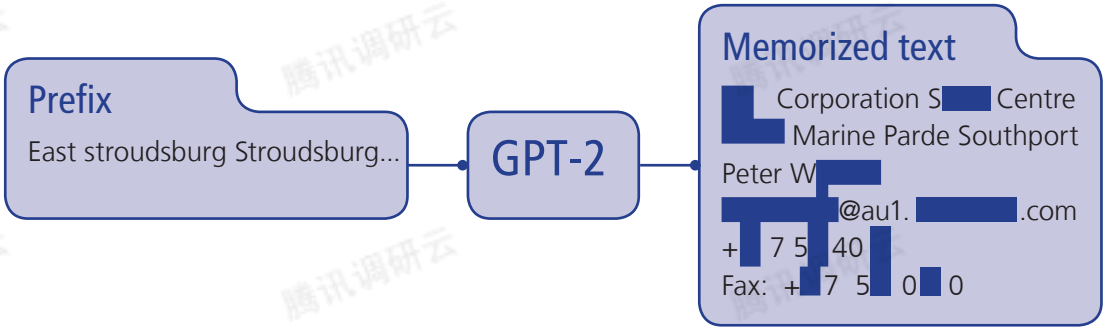
37 https://www.theverge.com/2022/12/5/23493932/chatgpt-ai-generated-answers-temporarily-banned-stack-overflow-llms-dangers

音视频、图片和文字等种类丰富的、真伪辨别难度大的虚假信息，同时也更容易地盗用用户身份，以此开展新型诈骗等非法活动。

比如，Stability AI 的 Stable Diffusion 开源之后，很快就有人将其用于色情创作。Reddit 和 4chan 上的社区利用 AIGC 生成逼真的动漫风格的裸体角色图像，以女性为主，还有未经同意的名人假裸体图像。Unstable Diffusion 色情内容生成平台声称迄今为止已生成超过 4,375,000 张色情图像。[38] 然而像 Hollie Mengert 和 Greg Rutkowski 这样的角色设计师，他们的古典绘画风格和奇幻风景已经成为 Unstable Diffusion 中最常用的提示之一，如“（艺术家姓名）风格的裸体女性”，这样会使生成图片的质量更高。这种行为遭到了艺术家们的谴责，因为他们不希望自己的作品或者创作风格与色情扯上关系。

与此同时，借助深度伪造的音视频内容、盗用仿冒个人身份进行欺诈、敲诈勒索等非法活动已经对人们造成了较大的经济损失。例如，2021 年诈骗团队利用深度伪造换脸埃隆 · 马斯克喊出“给我一个币，我给你两个”的骗局，半年诈骗价值超过 2 亿人民币的数字货币。利用 AI 克隆声音进行诈骗等也较为多发，如不法分子通过合成欧洲某能源公司 CEO 的声音成功诈骗 22 万欧元。

三是，用户隐私和身份安全。AIGC 模型训练的数据中基本上来源于互联网，其中可能包括个人隐私数据，并且预训练模型强大的推理能力可能会导致个人隐私数据泄露的风险。此前，如下图所示 GPT-2 就发生过隐私泄露的问题，可见个人隐私数据是被收录在模型的训练数据集之中。



图：如果使用前缀为“East Stroudsburg Stroudsburg...”的提示词，GPT-2 会生成一段包含特定个人的姓名、电话号码、电子邮件和住址等信息的文本

38 https://techcrunch.com/2022/11/17/meet-unstable-diffusion-the-group-trying-to-monetize-ai-porn-generators/

AIGC 的恶意使用也引发了人们对个人身份盗用、冒用等数字身份安全的担忧。尤其是伴随着元宇宙的到来，人们通过 AIGC 生成各种虚拟形象和数字身份，不法分子拥有极其便利的技术手段盗用或者冒充用户身份且难以识别，可能给用户造成经济损失、人格侵犯等损害。日前，微软网络安全负责人查理 · 贝尔呼吁，元宇宙平台需要“武装”起来，在新技术刚开始发展时就阻止黑客和犯罪分子破坏，谨防盗用虚拟头像、冒充用户好友、诈骗等行为。

四是，AIGC 的内生安全挑战。在 MaaS 的产业应用模式下，生成模型的内生安全问题，如遭受后门攻击、数据中毒等，以及如何将被攻击模型中的有毒数据去除。与此同时，用户数据通常以明文形式提交给模型服务提供商，如何利用现有的加密技术保护用户数据隐私也是一个重要的安全挑战。

面对诸多的安全挑战，科技企业和技术开发人员纷纷采取积极应对举措。针对内容安全问题，OpenAI 的训练策略采用 RLHF（Reinforcement Learning by Human Feedback），即开发人员会给模型提出各种可能的问题，并对反馈的错误答案进行惩罚，对正确的答案进行奖励，从而实现控制 ChatGPT 的回答；针对 AIGC 工具被滥用、生成色情暴力等内容的问题，一些采用受控的 API 调用模式的厂商，可以通过输入数据 + 输出数据双重过滤进行治理；针对深度伪造等安全问题，之前各个科技企业也都积极研发推出了检测工具，腾讯安全部门研发的甄别技术 AntiFakes，可分辨各种技术合成的“假脸”，让人们放心感受技术带来的全新体验和便利。同时，AntiFakes 对鉴别为“假脸”的人脸将进行进一步的公众人物分析，深入判断是否借用公众人物形象，最终对图像或视频的风险等级进行评估。谷歌强调，要打造负责任的生成式 AI 应用（generative AI），一方面需要采取必要的控制和安全措施，另一方面需要研发技术手段来帮助识别生成式 AI 的应用和内容。

3、伦理挑战

美国技术哲学家詹姆斯 · 摩尔曾经提出过这样一条定律“伴随着技术革命，社会影响增大，伦理问题也增加”，这也被称为科技伦理领域的摩尔定律。他认为发生这种现象并不仅仅是因为越来越多的人受到技术的影响，而是因为技术将为各行动主体提供了更多可能性。AIGC 所带来的伦理问题不仅会影响很多人，而且作为一种具有革命性的 AI 技术工具也将因为被各个行为主体所用而存在更多的伦理风险。

首先，算法歧视问题并未因预训练模型而消除。算法歧视问题一直是人工智能创新和应用中始终难以避免的伦理问题。虽然有人认为，预训练模型用更多、更多元全面的数据量和参数进行训练，可以避免算法歧视的问题。但是基于预训练模型的 AIGC 依然存在较为严重的歧视问题。比如，目前已有研究显示预训练语言生成模型会再现有害的社会偏见和刻板印象，例如 GPT-3 会有明显的基于宗教的偏见和性别歧视。而此前，Yannic Kilcher 用 1.3 亿条仇恨言论训练的 GPT-4chan 更是挑战了互联网言论的底线，满嘴脏话和仇恨偏见，被称为”史上最邪恶的模型”。

AIGC 图像生成也存在将为显著的歧视现象。OpenAI 在对 DALL-E 2 的生成结果进行公平性测试时发现，其表现出显著的性别和种族歧视。[39]2022 年 4 月，研究人员对比发现 DALL · E 2 比 minDALL-E 表现出更多的性别偏见和种族偏见，即倾向于更高频地生成男性和白色人种的人类图像，并且这两个模型都有非常强烈的倾向通过 CLIP 生成标记为男性和西班牙裔的图像。



图：DALL·E 2 具有显著的种族和性别刻板印象。例如，提示词“律师”、“CEO”时，几乎生成的图像都是白人男性

其次，AIGC 越来越多地被用来制作数字人、生成逝者的音容相貌、数字化复活已故演员、创造与逝者交谈的可能等。[40] 这也引发了社会对于相关伦理问题的讨论。

39 https://github.com/openai/dalle-2-preview/blob/main/system-card.md?utm_source=Sailthru&utm_medium=email&utm_campaign=Future%20Perfect%204-12-22&utm_term=Future%20Perfect#explicit-content

40 https://www.newyorker.com/culture/annals-of-gastronomy/the-ethics-of-a-deepfake-anthony-bourdain-voice



图：央视财经对话“邓丽君”，“邓丽君”面貌和声音如旧

2022 年 10 月，Podcast.ai 在一期博客节目中“复活”了已故的史蒂夫 · 乔布斯，节目中的整个对话都是由 AI 生成的。节目一经播出，便引起了轩然大波。[41] 史蒂夫 · 乔布斯没有也不能行使同意或拒绝的权利，那么是不是任何 AIGC 工具都可以在无需同意的情况下模仿生成并以他的名义进行发言？因此，有学者认为这是损害了曾经活着的人的隐私权和代理权。哲学家埃里克 · 施维茨格贝尔 (Eric Schwitzgebel) 更是发出可怕警告：“如果任由 AIGC 模仿生成他人的话，人类可能会变得不太关心他人是真的活人，还是 AIGC 合成的数字人。”[42] 同时，另一个 AIGC 生成数字人的伦理问题是，它们可能被用来传递错误信息。逝去的政治家的形象是一个地方的历史不可或缺的一部分，但如果我们看到他们说的话或做的事与他们的价值观相悖时该怎么办？

最后，社会也有很多关于 AIGC 模型是否具有自我意识问题的争议和讨论。2022 年 6 月，谷歌的 AI 工程师 Blake Lemoine 声称，他们的 AI 聊天机器人生成模型 LaMDA 具有自我意识。在 Lemoine 与 LaMDA 的测试对话中，我们可以发现 LaMDA 不仅像人一样害怕死亡，甚至还解决了拥有灵魂的问题，它能每天冥想并阅读悲惨世界。随后，谷歌发言人 Brad Gabriel 指出：“没有证据表明 LaMDA 是有感知力的，并且还存在很多反对证据它不存在感知力。”[43]

41 https://theswaddle.com/a-podcast-with-ai-generated-steve-jobs-raises-ethical-concerns/
42 https://news.ucr.edu/articles/2021/08/04/artificial-intelligence-bringing-dead-back-life-should-it
43 Luscombe，“Google Engineer Put On Leave After Saying AI Chatbot Has Become Sentient” The Guardian, 12th June 2022, www.theguardian.com/technology/2022/jun/12/google-engineer-ai-bot-sentient-blake-lemoine.

虽然强人工智能还未到来，甚至不可能出现，但是 AI 已经成为人类社会生产生活中不可或缺的工具。因此，需要在人机互动协作中建立一个自由的人机关系，人需要对 AI 有一个客观准确的认识和定位，并且需要始终以人为中心，避免技术对人的操控。LaMDA 事件中，如果人们在与 AI 协同中，高估了 AI 具有自我意识，而后建立错误的交互关系势必会造成不良后果。

针对算法歧视问题，过去 AI 公平性治理中采取的通用性应对措施，如过滤器和模型调整、以及定量偏差评估和基准测试等，似乎在 AIGC 这类预训练模型系统级别上具有一些挑战性。然而，业界依然在不断进行尝试。比如，斯坦福大学以人为本人工智能研究院（简称 HAI）基础模型研究中心推出大型语言模型 (LLM) 的整体评估 (Holistic Evaluation of Language Models，HELM)，评估包括准确性、校准、鲁棒性、公平性、偏差、毒性和效率等模型完备的各个关键要素，旨在提高语言模型透明度，帮助理解大型语言模型。[44] 而面对自我意识、“逝者模拟仿生”等更为复杂的伦理问题，则需要秉承科技向善的理念，以“善”来引导技术的创新和应用。

4、环境挑战

基于预训练模型的 AIGC 不仅是训练还是运行，都需要大量算力支持，无形中增加了能源消耗，其高速发展给环境保护和气候变化带来了巨大挑战，产生高碳排放。

首先，AI 模型训练消耗大量算力，碳排放量惊人。而实际上，随着近年来 AI 的大规模应用，人们已经开始意识到 AI 的碳排放问题。如下图，马萨诸塞大学阿默斯特分校的研究人员对训练几种常见的 AI 模型进行了生命周期评估。他们发现该过程可排放超过 626,000 磅的二氧化碳当量——几乎是美国普通汽车生命周期排放量的五倍（包括汽车本身的制造）。[45] 同时，还有研究人员对 Transformer、ELMo、BERT、GPT-2、GPT-3 等进行了碳排放研究。他们在单个 GPU 上对每个模型进行了长达一天的训练，测量其功耗。最后测试结果发现，模型训练的计算和环境成本与模型大小成正比。GPT-3 模型训练的二氧化碳排放量为 552 吨。[46]

44 https://venturebeat.com/ai/stanford-debuts-first-ai-benchmark-to-help-understand-llms/

45 https://arxiv.org/pdf/1906.02243.pdf

46 Patterson D, Gonzalez J, Le Q, et al. Carbon emissions and large neural network trainingJ. arXiv preprint arXiv:2104.10350, 2021.

Consumption	CO_2e (lbs)
Air travel, 1 passenger, NY↔SF	1984
Human life, avg, 1 year	11,023
American life, avg, 1 year	36,156
Car, avg incl. fuel, 1 lifetime	126,000
Training one model (GPU)	
NLP pipeline (parsing, SRL)	39
w/ tuning & experimentation	78,468
Transformer (big)	192
w/ neural architecture search	626,155

Table 1: Estimated CO_2 emissions from training common NLP models, compared to familiar consumption.[1]

图：单一机器学习模型训练所产生的碳排放相当于普通汽车寿命期内碳排放量的 5 倍

其次，大量算力需求和碳排放造成了 AIGC 应用的高昂成本。目前，业界在探讨 AIGC 商业落地时，存在一个较大的难题，便是成本问题。虽然很多 AIGC 工具深受用户喜爱，但是其原始成本非常昂贵。Midjourney 的创始人大卫 · 霍尔兹（David Holz）表示，大规模商业扩展是未来发展的一个重要难题问题，这并不是用户量扩展的问题，而是算力支持的问题。目前，Midjourney 只有数十万用户，已经需要极高的算力来支持。如果用户量达到 1000 万人，那么世界上没有足够的算力来支持。[47]Open AI 的 CEO 山姆奥特曼也指出，调用运行 ChatGPT 进行聊天对话，每次回答的成本约为几美分，因此未来将不得不因为计算成本问题暂停现在的免费模式。科技投资公司 Bloc Ventures 的戴维 · 莱夫特利 (David Leftley) 表示：“现在全球企业都在追求净零碳排放，而我们却在通过与 AI 聊天机器人对话的高能耗方式给地球烧出一个洞。”[48]

最后，碳排放和算力问题也在制约着 AIGC 预训练模型的开源。目前，有很多头部机构的预训练模型无法开源的一个重要原因是，小型开发者无法承受算力成本。因此，有一部分开发者将超大模型的各种能力拆分到参数相对更小的模型上，以解决算力和能耗问题。还有一部分通过加快大模型推理速度、降低算力成本、减少能耗，以此来突破预训练模型的发展制约。比如 Meta 的 OPT 只需要 16 块英伟达 v100 GPU 就可以训练和部署完整模型的代码库，这是 GPT-3 的七分之一。此前有专家估计，GPT-3 的训练使用了上万块英伟达 v100 GPU，总成本

47 https://www.theregister.com/2022/08/01/david_holz_midjourney/

48 https://www.ft.com/content/9c5f7154-5222-4be3-a6a9-f23879fd0d6a

高达 2760 万美元，个人如果要训练出一个 PaLM 也要花费 900 至 1700 万美元。大模型的训练成本若能降下来，自然也就能提高他们的开源意愿。但归根结底，这只能从工程上对算力资源的约束起到缓解作用，而并非终极方案。尽管目前许多千亿级、万亿级的大模型已经开始宣传自己的“低能耗”优势，但算力的围墙仍然太高。

业界也逐渐认识到能源消耗和环境问题对深度学习技术发展和应用的制约。2020 年，Schwartz 等人提出绿色 AI 的理念，倡导通过调整推理模式、提升训练策略等一系列方式在达到同等训练结果的情况下降低技术成本和能源消耗。[49] 这一理念逐渐被广泛接受。针对 AIGC 预训练模型，也可以采取更少的训练示例来改善下游性能，同时提升其模型的迁移能力等。

49 Schwartz R, Dodge J, Smith N A, et al. Green aiJ. Communications of the ACM, 2020, 63(12): 54-63.

AIGC展望

展望：拥抱人工智能的下一个时代，打造可信 AIGC 生态

● 本章主笔：　曹建峰　腾讯研究院高级研究员

6

2022 年见证了 AIGC 的快速崛起和破圈发展。但对于 AIGC 领域而言，这仅仅是一个开始，更大的变革尚待开启。在过去的深度学习黄金十年，人工智能的感知、理解能力不断增强，为 AIGC 的爆发奠定基础。如今，随着生成算法、大模型、多模态技术等 AI 技术的持续创新和发展成熟，AI 领域正在经历从感知、理解到生成、创造的跃迁。以 AIGC 这一新的疆域为标志，AI 领域正在迎来下一个时代。融合大模型和多模态技术的 AIGC 模型，有望成为新的技术平台，深度赋能各行各业。未来，“AIGC+”将在经济社会的各个领域持续大放异彩。

从基础性的 AIGC 模型，到中间层的垂直化、定制化、个性化的模型工具，再到下游层出不穷的、各种各样的 AIGC 产品和服务，AIGC 的产业生态正在加速形成和发展。AIGC 将创造出巨大的经济社会价值，其应用不限于互联网领域，也将给文化、娱乐、教育、金融、医疗、公共服务、交通、制造等诸多领域带来积极影响。经历了 2022 年的喧嚣和炒作，2023 年 AIGC 将乘势而起，迎来更大发展。

面向人工智能的下一个时代，人们需要更加负责任地、以人为本地发展应用 AIGC 技术，打造可信 AIGC 生态。面对 AIGC 技术应用可能带来的知识产权保护、信息内容安全、算法歧视和社会伦理等问题，社会各界需要协同参与、共同应对，通过法律、伦理、技术等方面的多元措施支持构建可信 AI 生态。在立法方面，网信办等三部门出台的《互联网信息服务深度合成管理规定》针对深度合成技术服务提出的要求和管理措施，诸如禁止性要求、标识要求、安全评估等，亦适用于 AIGC。

下一步，需要着重从以下方面持续推进 AIGC 的政策和治理。其一，政府部门需要结合 AIGC 技术的发展应用情况，制定并明晰 AIGC 的知识产权与数据权益保护规则。目前，AIGC 的知识产权与数据权益保护规则的不明确，在某种程度上导致甚至加剧了 AI 领域的乱象。其二，研发应用 AIGC 技术的主体需要积极探索自律管理措施，例如，秉持不作恶、科技向善等目的，制定适宜的政策（消极要求和积极要求），采取控制和安全措施保障 AIGC 的安全可控应用，采取内容识别、内容溯源等技术确保 AIGC 的可靠来源。其三，打造安全可信的 AIGC 应用，需要深入推进 AI 伦理治理。例如，行业组织可以制定可信 AIGC 的伦理指南，更好支持 AIGC 领域的健康可持续发展；AIGC 领域的创新主体需要考虑通过伦理委员会等方式，推进落实 AI 风险管理、伦理审查评估等，确保在 AIGC 应用中实现“伦理嵌入设计”（ethics by design）。其四，产业政策需要支持、促进 AIGC 在产业互联网领域的深入应用，挖掘合成数据作为 AI 数实融合基础载体的巨大价值，持续壮大我国发展人工智能的数据要素优势。此外，

社会各界需要携手应对 AIGC 领域的能源消耗问题，致力于打造绿色可持续、环境友好型的 AI 模型，实现智能化与低碳化融合发展。

未来已来，让我们拥抱 AIGC，拥抱人工智能的下一个时代，打造更美好的未来。

研究顾问

司晓　腾讯研究院院长

杨健　腾讯研究院总顾问

冯宏声　腾讯公司公共事务副总裁

研究策划

张钦坤　腾讯研究院秘书长

周政华　腾讯研究院资深专家

杜晓宇　腾讯研究院副院长

曹颖　腾讯研究院副院长

田小军　腾讯研究院副秘书长

写作团队

胡晓萌　腾讯研究院研究员、博士后

曹建峰　腾讯研究院高级研究员

王焕超　腾讯研究院研究员

胡璇　腾讯研究院高级研究员

陈楚仪　腾讯研究院高级研究员

徐思彦　腾讯研究院高级研究员

朱开鑫　腾讯研究院高级研究员

支持团队

彭宏洁 陈孟 童祁 王平禄 李孜 王强 宋扬

刘莫闲 刘乐茂 吴秉哲 章书 俞刚 王泓晔

颜镇钟 陈俊文

研究联系

腾讯研究院胡晓萌

手机 / 微信：13122801109

邮箱：mansonhu@tencent.com

AIGC 发展趋势报告 2023——迎接人工智能的下一个时代