

第 1 篇

AIGC 的蜕变



基于微软 BING 图像创建者和 PowerPoint 生成，2023

第 1 章

开启 AIGC 的魔盒

1.1 AIGC 的来历与预言

AIGC 是三个单词的缩写，即 Artificial Intelligence（人工智能）、Generated（生成的）和 Content（内容）。AI 指的是一种模拟人类智能的技术，可以通过机器学习、自然语言处理和计算机视觉等方法，实现从数据中学习和理解规律、做出决策和完成任务等功能。Generated 是动词 generate 的过去分词形式，表示这些内容是由 AI 程序生成的而非由人类创作。Content 意为“内容”，指文本、图像、音频、视频等各种形式的信息。AIGC 是指借助预训练（超）大规模的模型、生成式对抗网络（GAN）等 AI 技术，对已有数据进行规律性的提取和总结，最终通过释放泛化能力^①生成多样化内容。2022 年，除了越发流行的 ChatGPT 文本内容生成，在图像生成方面，Disco Diffusion、

① 泛化能力（Generalization Ability）是指机器学习算法对新鲜样本的适应能力。学习的目的是学到隐含在数据背后的规律，对具有同一规律的学习集以外的数据，经过训练的网络也能给出合适的输出。

Midjourney、DALL·E^①和Imagen AI等也是效果惊人，而这只是AIGC的一个缩影，还有用MusicLM生成音乐、用GET3D生成三维物体等。

如图1-1所示的两张图都是AIGC的杰作，图1-1(a)是AIGC工具NightCafe基于DALL·E创作的中国长城的美图，你会感觉它似乎有哪里不对，但总体的感觉又十分壮观。实际上，生成它的逻辑是基于人们对长城的各种典型概念和印象而形成的。图1-1(b)是2022年美国一家游戏公司的总裁杰森·艾伦(Jason Allen)基于包括Midjourney在内的AIGC工具创作的数字化作品“太空剧院”，这幅画在科罗拉多州博览会上获得数字类第一名。这是基于AIGC生成的艺术品对传统艺术的第一次重大挑战，关于它是否是真正的艺术创作的争论仍在进行，但丝毫不影响AIGC的崛起。



(a) 中国长城



(b) 太空剧院

图1-1 AIGC生成的“中国长城”与“太空剧院”

资料来源: NightCafe & The New York Times

简单回顾AIGC发展的历史，可以追溯到20世纪50年代，当时人们开始研究计算机语言和自然语言处理技术，这是AIGC的雏形。在接下来的几十年里，研究人员不断探索和改进各种机器学习算法，并应用于文本、图像、音频和视频等多个领域，随着计算能力的提高和深度学习算法的发展，AIGC取得了重大突破。回顾这些里程碑事件，包括：2014年，谷歌发布了一种基于深度学习的图像

^① DALL·E是OpenAI开发的一种图像生成模型，它通过120亿个参数版本的GPT3 Transformer模型来理解自然语言输入并生成相应的图片。它既可以生成现实的对象，也能够生成现实中不存在的对象。

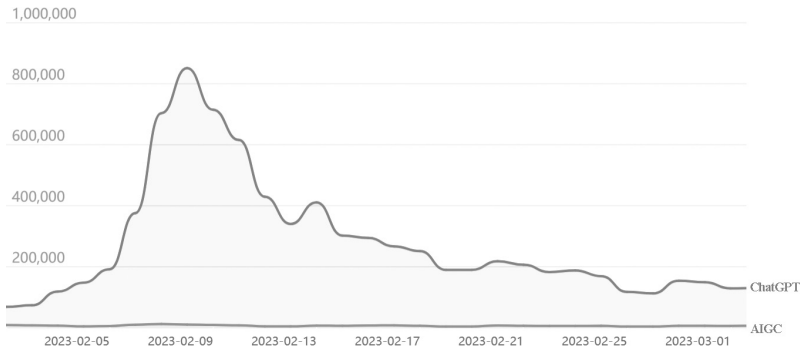
识别系统，大幅提升了计算机对图像的理解能力；同年，微软也发布了一种基于机器学习的自然语言处理系统，成功地生成了与人类写作相似的文章；2016年，OpenAI发布了一个名为GPT的自然语言处理模型，由175亿个参数组成，这是当时最大的模型之一（如今已超千亿参数），该模型可用于生成各种类型的文本内容，包括新闻报道、故事和诗歌等；2018年，Facebook的AI系统开发了自己的语言并与其他机器人进行了对话，虽然该事件被关闭，但是它激发了对于AI意识形态的讨论；2020年，OpenAI的GPT3模型推出，成为目前最具代表性的自然语言处理模型之一，其性能超越了以往所有的模型；2022年11月，OpenAI推出轰动世界的AI聊天机器人ChatGPT。现如今，AIGC已经广泛应用于各个领域，随着技术的不断发展，AIGC内容的应用前景也越来越广阔。根据国外商业咨询机构Acumen Research and Consulting预测，若考虑下一代互联网对内容需求的迅速提升，2030年AIGC市场规模将达到1100亿美元。Gartner则预计：到2025年，大型企业30%的对外营销信息将是合成的，在2022年这一比例还不到2%；到2030年，一部重要的大片上映时，影片中90%的文本及视频将由AI生成，而2022年则没有此类影片。

AIGC的两个代表性项目是ChatGPT和DALL·E，它们都是OpenAI的杰作，前者生成文本，后者生成图片，当然OpenAI还有音乐生成器MuseNet。这两个项目自推出以来得到了业界广泛的关注，并在社交媒体上表现突出，不少人将ChatGPT3比喻为推动AI走向工业时代的“蒸汽机”。如图1-2（a）展示了过去一年，ChatGPT与DALL·E在谷歌被搜索的趋势，显然ChatGPT受到极大的追捧。如图1-2（b）则展示了国内在2023年2月对ChatGPT和AIGC的搜索热度的暴增，数据显示，ChatGPT在此期间的整体日均搜索量达到了20万，尤其是在2月9日达到85万，整体环比增长970%。



(a) ChatGPT 与 DALL·E 在谷歌上的搜索趋势

图 1-2 ChatGPT、DALL·E 及 AIGC 的热度



(b) ChatGPT 与 AIGC 的百度搜索指数

图 1-2 (续)

资料来源：谷歌与百度搜索引擎生成，2023

最新版本的 DALL·E 2 于 2022 年 4 月公布。DALL·E 和 DALL·E 2 是深度学习模型，可以根据用户的自然语言提示生成图像。与 DALL·E 相比，DALL·E 2 可以生成分辨率更高的图像，更加逼真，并能结合各种风格。令人兴奋的是，2023 年 3 月，谷歌宣布会将类 ChatGPT 的 AI 整合到其办公套件 Workspace 中，微软则紧随其后宣布推出 Microsoft 365 Copilot，即在所有的微软办公套件中（Word、PPT、Excel、Outlook、Teams、Microsoft Viva、Power Platform）内置 GPT4 的功能，通过 Microsoft 365 Copilot 而不是需要安装额外独立的互联网或桌面软件，用户凭借最通用的界面和自然语言，就能轻松玩转 AI 工具，如图 1-3 所示。AI 平民化的时代来临了，其冲击就像个人电脑进入家庭时一样。

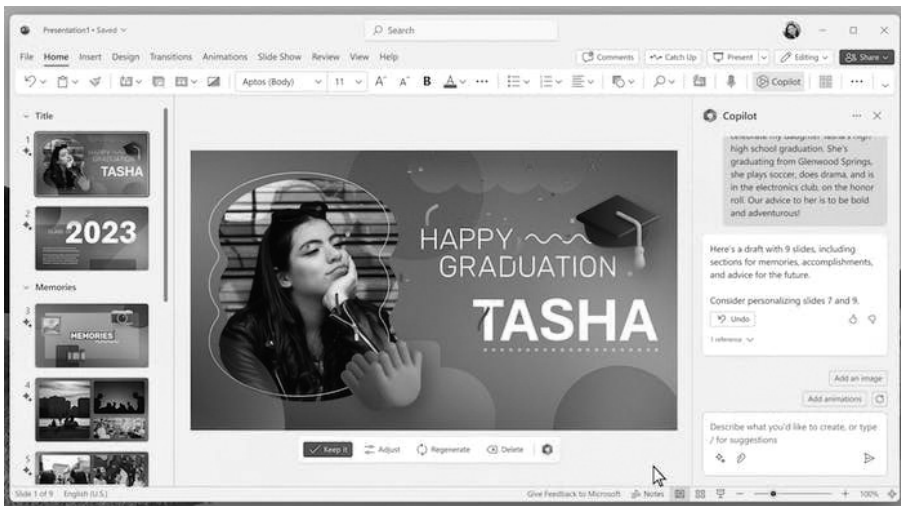


图 1-3 基于 ChatGPT4 智能生成的 PPT

资料来源：微软官网，2023

为了说明内容生成历史的若干阶段，诸多媒体把 AIGC 定义为一种全新的数字内容生成与交互的形态。它是继专业生成内容（Professional Generated Content, PGC）、用户生成内容（User Generated Content, UGC）之后的新范式。所谓范式是指公认的模式，AIGC 来得如此之快，以至于网上已到处可见 AIGC 生成的文本与图像资料。其实对于 AIGC 建立统一的定义还为时尚早，随着 AIGC 的边界不断扩大，通常由它所在领域的领导者提供更为精准的定义。正是由于技术的突破来得着实令人兴奋，业界大多数 AIGC 的描述更多强调它的技术优势，即围绕高效和高质的自动化与个性化来阐述，这与 AIGC 技术作为高科技、高估值的属性是相关的。事实上，所有的 AI 产品并非仅靠技术就能成功，关键还是应用，而应用就需要有庞大的内容与素材基础来训练。即使可以使用混合数据，但也是基于实际数据训练来建模、模拟生成的。

相较于传统的 PGC、UGC，除了了解 AIGC 与它们存在着生成者主体的区别外，更为重要的是了解它们之间的关系。PGC 是最早的内容生成形式，在最初的发展阶段中，学会使用个人电脑完成一系列的注册和内容编辑和发布，并不是常人可以做到的。当 PGC 进入“百家争鸣”时代，这些核心观点或理念需要通过交互来形成热门话题，并形成反馈的热度和流量经济。这种广泛的讨论生成了新的内容，即把一个专业话题作为内容的“树干”，通过众人参与交互形成了一棵完整的“话题树”，从而最终形成一片“话题的树林”，这时关于各领域的内容素材已十分丰富。随着移动互联网的发展，特别是短视频的推出，使全民轻松互动成为现实，由此这个“话题的树林”飞速成长，相互交织，形成了“话题的森林”，甚至达到了信息爆炸的状态。另外，算力提升不仅促成了“话题的森林”，同时也助推了算法的进步。算力由“后摩尔定律”^①推动前行，即使有所放缓但依然增长迅猛。这时，AIGC 出现了，以基于千亿级参数的模型和算法及神奇的输出效果吸引了全球的眼光，机器以超越人类预期的水平与人类展开了内容的对话。我们看到，AIGC 的神奇实现固然归功于算力和算法的进步，而更为重要的是归功于它的要素来源——超大规模的内容池——“话题的森林”。例如，如今 AIGC 也不能生成任意的内容，在用户需要对特定图像进行再造时，依然需要向机器提供

① 摩尔定律是由英特尔创始人之一戈登·摩尔（Gordon Moore）提出来的。目前，业界认为单纯靠提升工艺来提升芯片性能的方法已经无法充分满足时代的需求，半导体行业也逐步进入了后摩尔时代，其技术路线按照两个不同的维度继续演进：①“More Moore”，继续延续摩尔定律的精髓，以缩小数字集成电路的尺寸为目的，同时器件优化重心兼顾性能及功耗；②“More than Moore”，更多依靠电路设计以及系统算法优化，同时，借助先进封装技术，实现异质集成。

素材。可以预见，PGC 和 UGC 并不会消失：一方面 AIGC 借助前两者的内容基础生成新内容；另一方面，PGC 和 UGC 会聪明地根据 AIGC 的模型特征产生新的内容，这些内容又会被 AIGC 捕获或采集，经过再次重组反馈给用户，这样就形成了一个循环。那么，为什么 PGC 和 UGC 会向 AIGC 定向生产内容呢？毕竟 AIGC 已是一种新的门户与入口（Portal），从市场营销的角度来看，任何希望提升品牌地位的组织或个人，都希望他们的信息出现在 AIGC 生成内容的首屏，那么新的内容就必须根据 AIGC 生成的规则进行优化，就像今天 AIGC 工具根据 Google 的搜索排名规则，来优化用户生成的营销方案的逻辑。总结一下，PGC、UGC 和 AIGC 的关系如图 1-4 所示，内容生成的四个阶段并非是相互取代的关系，它是一种升级，后者以前者为基础，前者则更多地融入后者中，而所有阶段综合产生的混合数据又更进一步地推动 AIGC 的发展。

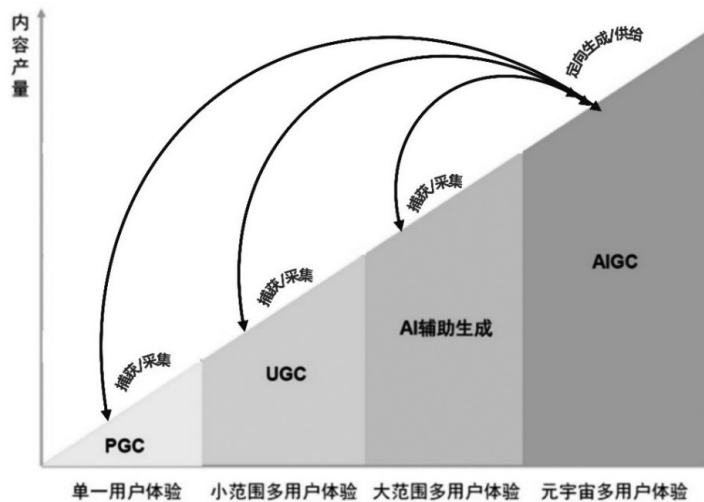


图 1-4 内容生成的四个阶段

资料来源：根据 a16z、ShineINFAITH、Muse Labs，2023 年提供的图片修改

如果我们需要区分 AIGC 与人类内容创作的区别，可以用“农夫山泉”的一句广告词来映射 AIGC——“我不是内容的生产者，我是内容的智能搬运工”，只是这种搬运确实越发智能了，不仅知道搬运什么，还知道在搬运的时候如何最优化“进销存”。既然 AIGC 是一种智能的内容搬运工，那么局限性也就容易理解了，混合数据的出现使 AIGC 不再依赖人类产生的数据，但并不意味着 AIGC 会产生人类创意的数据。虽然 AIGC 在某种程度上拟合出了人类未曾思考的领域或维度，但 AIGC 永远不会像人类那样创新。无论是哪种内容创作，其程序都包括生产者对信息的筛选、过滤、加工、整合等步骤。有些观点认为：在 PGC 时代，

这一系列的过程都是建立在创作者长期独立研究的基础上，花费了大量的时间和精力。当 PGC 和 UGC 的生产潜力被耗尽时，AIGC 或许可以弥补内容生态的不足。事实上随着科技的发展，专业的内容始终对人类有着极大的依赖，我们可以通过 AIGC 生成普适的内容，但不能指望它生成新的知识（AIGC 生成的随机的创新内容和专业的创新内容是两回事），而且 PGC 通过使用 AIGC 工具，会挖潜其巨大的生产力，而不是在 AIGC 面前瘫软下来而耗尽精力。AIGC 就像 AI 在诸多领域应用的场景，它确实可以突破很多人为的限制，包括不知疲倦地不断尝试，高效地生成与规定运作的反复执行，它也将逐步学习人类的创作方式，从而创造出更丰富、更多样的内容。理论上，AIGC 将实现内容生态的无限供给，考虑到生产效率和专业性，其内容质量将基于 PGC 和 UGC，在特定领域和场景下有可能超过 PGC 的部分内容，PGC、UGC、AIGC 之间的关系见表 1-1。

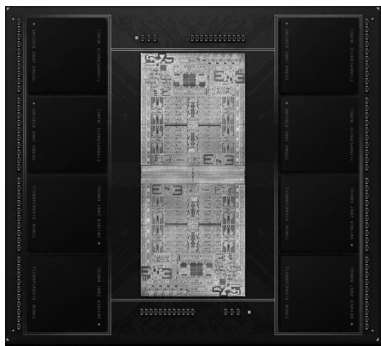
表 1-1 PGC、UGC、AIGC 之间的关系

	PGC	UGC	AIGC
创作主体	专业人士	广大用户	AI 和算法
特点区别	专业准确 有限生成 严谨可靠	创意丰富 无序生成 参差不齐	专业可控 批量生成 持续优化
相互关系	<ul style="list-style-type: none"> ● 三者并非完全隔离与竞争，就像家庭厨房、外卖与高档酒店在某种程度上存在着一定的竞争与替代关系，但更准确的说法是一种互相渗透的共存 ● PGC 需要 UGC 形成话题，无数的话题最终为 AIGC 提供丰富的素材。前二者互补相生，为第三者提供了内容基础 ● 最终三者融为一体，相互参照并互为输入与输出 		

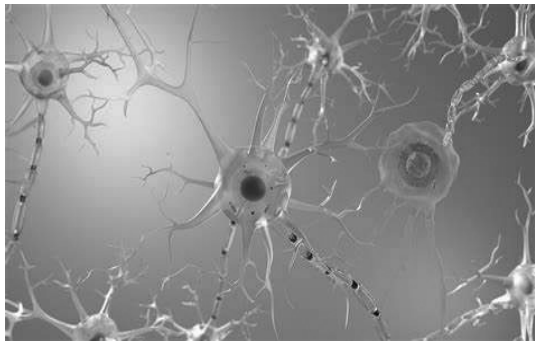
资料来源：Renaissance Rachel

基于硅基的二进制机器的智能生成，有可能在基于碳基的优秀人类的训练下，普遍超越基于碳基的一般人类的能力吗？作者认为有一部分是会的，就如通过搜索引擎查询信息，一定比人类在图书馆手动搜索来得强大一样，基于大模型算法在规范性内容的处理上已远超越人类。由 2022 年苹果公司推出、台积电代工的 M1 Ultra 芯片含有 1140 亿个晶体管，而人类大脑中仅有 130 亿~140 亿个细胞，如果将晶体管与细胞都看作一个微观世界的计算单元，那么在数量上，人脑显然已是相形见绌了。即便如此，机器与生命不是一回事，人类的智能的抽象性（这与 AIGC 基于已有内容生成抽象内容并不一样）和低能耗是机器不可比拟的，AIGC 代表的生成式 AI 会成为卓越的工具，但在各种应用场景中，其背后依然是基于人类智慧，并且为人类所用。因此，在 AI 走向人类智能的过程中，重要的突

破来自两者交叉科学的探索与发现，其主要趋势为：基于硅基物理架构支撑的系统向碳基生物计算突破。反之，也可以理解为，将更多碳基生物思考与运作的方式更多地应用于硅基系统。如图 1-5 展示了 M1 Ultra 芯片中的计算单元，其晶体管数量已远超人类大脑中的细胞数量，达到了后者的 7 倍之多。在我们欢呼人类在芯片产业创举的同时，另一拨科学家正在不断探索大脑的奥秘，例如美国哥伦比亚大学和西奈山伊坎医学院研究人员正在开展一个项目，该项目将生成整个人类大脑及其所有细胞的综合图集。这些数据可帮助揭示大脑的结构和组织如何在疾病和健康中产生行为、情感和认知，这都将为硅基计算提供新范式的参照。芯片技术与生物科技，必然与必定走到一起，对于 AI 的发展来说更是如此，即 AI 正在更多地学习人类，探索走向人类智能的道路。



(a) 含有 1140 亿个晶体管的 M1 Ultra 芯片



(b) 人脑神经元联系示意图

图 1-5 芯片的晶体管数量已远超人脑细胞的数量

资料来源：苹果官网 & Tamas 实验室

由于 AI 技术的进步和更复杂算法的出现，同时随着质量的上升和成本的下降，AIGC 正在不断普及。公司正越来越多地转向 AI 生成的内容，以减少人工创建内容的负担并节省成本。具体应用包括 AI 写作、AI 配乐、AI 视频生成、AI 语音合成以及近段时间火遍全网的 AI 绘画——只要简单输入几个关键词，几秒钟时间一幅画作就能诞生，这可以更有效地节省时间和资源，并提高内容的质量。除了图片的生成，更常见的是输入一篇文章的描述，或者只是故事的开头，让 AI 为你完成文章。凡是需要写作或内容创作的地方，它都有广泛的应用，如撰写财务报告、开发代码、创建销售 / 营销材料。AIGC 可以帮助人们更快地理解和分析复杂的信息，从而帮助他们做出更好的决定并产生重大价值。AIGC 还可以用于创建可量化的内容，并使用数据驱动的方法来给出建议，以提高内容的质量。此外，AIGC 可以更有效地分析和汇总数据，以便更好地理解客户的行为和需求，为客

户提供更个性化的内容。AIGC 可以用于创建丰富的多媒体内容，如图像、视频和音频。它可以用于创建自动化的新闻报道或分析，以及动态的广告内容。此外，AIGC 还可以用于创建和管理大量的内容，包括社交媒体内容，以满足客户对个性化与时效性等方面更高的需求。

如表 1-2 所示，展示了 AIGC 与人类角色的特点，AIGC 首先是人类的智能助理，而人类承担了真正的知识管理工作，AIGC 是人类知识管理工作的一部分，需要人类的训练、监督和优化。AIGC 体现在上述内容的诸多专业的领域，而人类代表了通用智能，AIGC 正在更多地学习人类，向人工通用智能进化；AIGC 是完成了知识编纂，而更复杂、非常规以及在互动过程中的隐性知识的传递是人类完成的。最后，AIGC 是拿来即可，它试图回答知识“是什么（what）”及“如何（how）”的问题，“如何（how）”是指人类在处理（大）数据时容易忽略的维度，总有一些影响的过程与环节容易被疏漏，但人类善于想象、比喻、解释、推论、预测，培养下一代专家并承担人类特有的职责。

表 1-2 AIGC 和人类在管理知识方面的共生关系

AIGC		人类角色	
个人智能助理	<ul style="list-style-type: none"> ● 帮助解决信息过载问题 ● 增加认知“带宽” ● 对信息资源进行过滤 	知识管理工作人员	<ul style="list-style-type: none"> ● 培训和个性化智能助理 ● 监控和严格评估智能助理的绩效
专业智能	<ul style="list-style-type: none"> ● 在约定边界的环境中提供专业智能 ● 以任务为中心的智能（不易在不同任务环境中转移） ● 管理知识内容 	通用智能	<ul style="list-style-type: none"> ● 运用知识进行战略层面的思考 ● 跨语境翻译知识 ● 通过自我反思辨别知识背景（悟）
知识编纂	<ul style="list-style-type: none"> ● 使低层次、高容量知识流程更为顺畅 ● 促进人与人之间的联系并产生专业知识 	知识协作	<ul style="list-style-type: none"> ● 处理复杂的、非常规的知识流程 ● 通过社会互动传递隐性知识
关于知识“如何”和“是什么”	<ul style="list-style-type: none"> ● 在（大）数据中发现被忽视的范式 ● 通过制定自己的规则来扩展现有知识资源之外的专用技术 	知道为什么	<ul style="list-style-type: none"> ● 想象、比喻 ● 解释、推论并证明建议的合理性 ● 培养潜力人才或专家并获得组织支持 ● 承担责任

资料来源：根据人工智能和知识管理，Mohammad Hossein Jarrahi 等，2023 改编

随着与用户的交互越来越多，每个用户都在所使用的 AIGC 中形成了用户画

像，这使 AIGC 的个性化成为可能，AIGC 得以从用户的问题中判断其性别、性格、年龄、职业、兴趣爱好，甚至目前的工作重点和职业发展需求，这都使 AIGC 提供充分满足个人需求的自动化服务流程，提供更为精准的深度定制。定制又使 AIGC 生成的内容更有吸引力，如互动故事、互动视频和互动游戏。当然，AIGC 可以使用深度学习算法来创建更加自然的内容，更好地模仿人类语言表达，这样就可以改善和完善客户对产品的体验，并在更多渠道上推广产品与服务，吸引更多的用户。

AIGC 能够实现如此强大的功能，主要是因为它由以下三大前沿技术支撑：

- **数字内容孪生能力：**构建现实世界到虚拟世界的映射，孪生能力包括智能增强与转译技术，其中增强技术弥补内容数字化过程中的信息损失，转译技术在理解基础上对内容进行多种形式呈现。
- **数字编辑能力：**打通现实世界到虚拟世界的交互通道。编辑能力包括智能语义理解与属性控制，语义理解帮助实现数字内容各属性的分离解耦，属性控制则在理解基础上对属性进行精确修改、编辑与二次生成，最终反馈于现实世界，形成孪生—反馈闭环。
- **数字创作能力。**从数据理解向数据创作转变，创作能力可分为基于模仿的创作与基于概念的创作，前者基于对某一类作品数据分布进行创作，后者从海量数据中学习抽象概念，并基于概念创作出现实世界不存在的内容。

AI 内容生成器的工作方式是通过自然语言处理^①和自然语言生成^②方法生成文本。这种形式的内容生成有利于为企业提供数据，根据用户行为定制内容并提供个性化的产品描述。文本生成模型一般通过无监督的预训练，语言转换模型从大量的数据集中学习并捕捉无数有价值的信息。在这样的海量数据上进行训练，可以使语言模型动态地生成更准确的矢量表示和带有上下文信息的单词、短语、句子和段落的概率。由于梯度消失问题^③，传统的递归神经网络（RNN）深度学习模型在长期建模的背景下步履蹒跚。转换器（Transformers）则不同，它克服了梯度

① 自然语言处理（Natural Language Processing，NLP）是计算机科学领域与 AI 领域中的一个重要方向，一门融语言学、计算机科学、数学为一体的科学。它的研究方向是实现人与计算机之间用自然语言进行有效沟通。

② 自然语言生成（Natural Language Generation，NLG）是研究使计算机具有人一样的表达和写作的功能，即能够根据一些关键信息及其在机器内部的表达形式，经过一个规划过程，来自动生成一段高质量的自然语言文本。

③ 梯度消失问题是指在深度多层前馈网络或递归神经网络，无法将信息从模型的输出端传播回模型输入端的附近层，导致具有多层的模型普遍不能在给定的数据集上进行训练，或者过早地满足于一个次优的解决方案。

消失问题，因为语言模型随着数据和架构规模的扩大而扩大，转换器实现了并行训练并能够捕获更长的序列特征，使更全面有效的语言模型的突破成为可能，因此转换器正在迅速成为 NLG 的主导架构。今天，像 GPT3/4 这样的 AI 系统被设计为生成类似人类创造力和写作风格的文本，而大多数人类一般难以分辨。这样的 AI 模型也被称为生成式 AI，即可以为广泛地使用案例创建新颖的数字媒体内容和合成数据的算法。生成式 AI 的工作原理是生成一个物体的许多变体，并对结果进行筛选，以选择那些具有有用的目标特征的变体。AIGC 基于算法生成内容，而算法是在大量的数据上训练出来的。然后根据这些数据和一点点用户的输入返回输出结果，关键是，内容是新的，而且是自动生成的。生成式 AI 算法最常见的例子是 ChatGPT。GPT3/4 在互联网中的大量人类文本上进行训练，并“教导”语言模型在与用户互动时如何回应。其他生成式 AI 程序的工作方式与此类似，它们被训练来发展一套知识体系，并使用这些知识来创造新的输出。目前许多商业生成式 AI 产品都是基于 OpenAI 的生成式 AI 工具，如 ChatGPT 和 Codex，下面我们大致了解一下不同内容的生成形式和目前所处的阶段。

- **文本与代码：**文本是最先进的领域。自然语言曾经是一项巨大的挑战，它对质量要求很高。今天，这些模型在通用的短篇/中篇写作方面还算不错（但即便如此，它们通常也是用于迭代或写出一个初稿）。随着时间的推移，模型越发健壮，高质量的成果正在输出——质量更高、内容更长以及更好地面向用户特定市场或需求的调整。如 GitHub CoPilot 所示，代码生成在短期内可能会对开发者的生产力产生很大影响。它也将使非开发人员更容易获得对代码的创造性使用。
- **图形图像：**相对于文本来说是一个较新的应用，但它的娱乐性显然好过文本，因此像病毒一样传播开来，这些应用和 ChatGPT 一样易于使用，创意力非常惊人。在 Twitter 上分享生成的图像比文字要有趣得多，可以看到具有不同审美风格的图像模型，以及编辑和修改生成图像的不同技术。例如，设计一个 logo，未来的大部分人，包括设计师第一时间都会通过 AIGC 辅导设计，这并不是说机器真的具备智慧，但算力的构建能力可以在人脑出现疲惫的时候持续给出质量稳定的呈现。
- **语音合成：**它的出现已有近十年时间了（还记得“你好，Siri！”吗？），但消费者和企业应用才刚刚开始。对于像电影和播客这样的高端应用来说，要想获得听起来非机械化的、具备一次性完整表达能力的、类似人类质量标准的语音是非常困难的。但未来是令人期待的，就像图像生成，今

天的语音合成模型将为持续完善以达到高质量输出的商业性应用提供一个快速发展的起点。

- **视频和 3D 建模：**这个更为复杂的领域正在迅速崛起。人们为这些模型通过大量生成创意内容所能创造出的新型市场的潜力感到兴奋。例如，电影、游戏、VR、建筑和物理产品设计。研究性组织正在发布基础的视频和 3D 模型，以便于形成新的产业链以衍生出不同应用，未来的大部分视频作品均可能借助 AIGC 生成。

如图 1-6 所示，我们通过 AI 图片自动生成引擎 Midjourney 在十几秒的时间内分别生成了上海春天的冬天的鲜花，图片清晰地呈现了城市和季节的特征，其美感和质量已超过了很多人 PS 出来的效果。我们可以看到，当我们给出不同的指令（极其简单的英文关键词命令，通常以 /imagine 呈现）之后，类似的图片会持续生成出来，新的图片视角不同，但美感依然。对于那些十分模糊的指令，通常图片的生成是虚幻的，但并不是没有内容，一个模糊的指令会对应一个模糊的图片渐变地产出，图片呈现的内容可能什么都不像，却有一种神秘感。其实面对这种模糊，可以不断增加清晰的关键词，直到做出自己满意的图片，比如不再用关键字，而是用一段文字，比如“朝阳冉冉升起，一条弯弯的河流穿过一条中国特色的古镇，河流上飞过几只欢快的鸟儿”。



(a) AI 生成图片：上海春天的鲜花

(b) AI 生成图片：上海冬天的鲜花

图 1-6 AI 生成图片：上海的鲜花

资料来源：Midjourney

除如上几个经典的应用场景，许多领域都在进行基础模型的研发，从音频和

音乐到生物和化学。这些应用五花八门，不胜枚举。

- **文案写作：**市场对个性化网络和电邮内容的需求日益增长，它可以用来促进销售和营销策略以及客户支持，这些都是语言模型的完美应用。公司在市场竞争中面临时间和成本的压力，更为快速、简明扼要且具有个性风格的言语表达，在应用中已出现。
- **细分市场的写作助手：**今天的大多数写作助手都是横向的，或称大众化的，AIGC 正在为特定的专业领域建立更好的生成性应用，从法律合同写作到编剧，这里的产品差异化在于对特定工作流程的模型和用户体验模式进行微调。
- **代码生成：**目前的应用为开发者提供“涡轮增压”的助力引擎，使他们的工作效率大大提升。GitHub Copilot 现在在安装它的项目中生成了近 40% 的代码。但更大的机会可能是为消费者打开了编码的通道。
- **艺术的产生：**整个艺术史和流行文化被编码在大型模型中，允许任何人随意探索以前需要用一生才能掌握的主题和风格。
- **媒体 / 广告：**想象一下将代理工作自动化的潜力，并为消费者优化广告文案和创意。这里有多模式生成的巨大机会，将销售信息与互补的视觉效果配对。
- **设计：**数字和实体产品的原型设计是一个劳动密集型和反复的过程。粗略的草图和提示的高保真渲染已经成为现实。随着 3D 模型的出现，下一代的 iPhone 到新款的高尔夫球鞋都可能由机器设计的。
- **游戏。**梦想是使用自然语言来创建复杂的场景或可操纵的模型。最终状态可能还很遥远，但有一些更直接的选择在短期内可操作，如生成纹理和天空盒艺术。

为了发挥一下 ChatGPT 的功能，我们总结了在过去一年多的时间中，诸多知名研究院发布的 AIGC 报告，根据发布的时间顺序让 ChatGPT 来对这些资料做一下介绍，见表 1-3。与 ChatGPT 的问答形式很简单，通常通过以下列出的几种方式让 ChatGPT 做出回答（这些报告中比较重要的图表由笔者整理）。


- (1) 输入资料名称；
- (2) 介绍一下 ×××；
- (3) ××× 的主要内容是什么？
- (4) 再试一下；
- (5) 请补充更多内容。

当 ChatGPT 找不到内容而回答“抱歉”时，可以重复同样的问题。

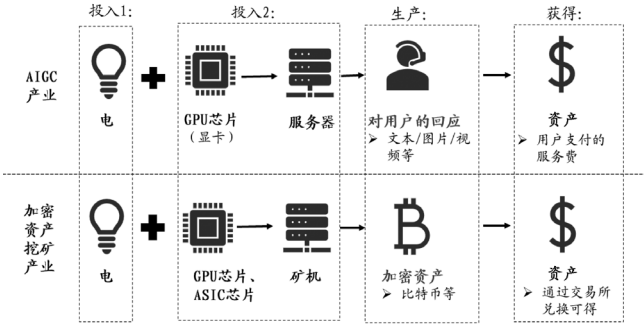
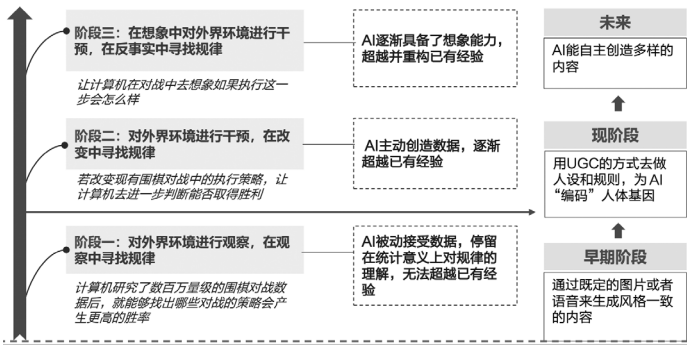
表 1-3 关于各大媒体或研究院发表 AIGC 报告的总结

资料来源	ChatGPT 回答 + 人工校对
<p>中国信通院联合京东探索研究院</p> <p>《人工智能生成内容 (AIGC) 白皮书》</p> <p>2022 年 9 月</p>	<p>该白皮书介绍了人工智能生成内容技术的现状、发展趋势和应用场景，并对该技术的道德、法律和安全等方面进行了讨论。白皮书还提出了相关技术的规范和标准，以促进该领域的健康发展</p> <p>主要内容包括：</p> <ul style="list-style-type: none"> ● 人工智能生成内容技术的现状和发展趋势，以及其对社会和经济的影响 ● AIGC 技术的应用场景和商业模式，包括广告、媒体、教育、医疗等行业 ● AIGC 技术所面临的道德、法律和安全等方面的挑战，以及如何规范和监管该技术的应用 ● 推动 AIGC 技术健康发展的建议和措施，包括相关技术的规范和标准，促进产学研合作等 <p>该白皮书旨在为人工智能生成内容技术的发展提供指导和规范，同时也强调了该技术的重要性和潜力</p>  <p>AIGC 的发展历程与概念演化</p>
<p>浙商证券</p> <p>《AIGC 之 AI 绘画：技术与应用双突破，生产力变革在即》</p> <p>2022 年 11 月</p>	<p>该报告主要探讨了基于 AIGC 技术的 AI 绘画在艺术领域的应用和前景，并重点介绍了该技术的核心原理、发展历程和应用场景。此外，报告还分析了 AI 绘画在推动数字文化创意产业升级和提高生产力方面的潜力，并探讨了相关的商业模式和投资机会。该报告旨在为读者提供有关 AIGC 技术在艺术创作领域的最新进展和未来趋势的信息，以及如何参与相关领域的投资活动等方面的建议</p> <p>AI 绘画是 AIGC 重要的应用分支。近两年包含扩散模型在内的关键技术取得突破，技术可用性显著提高，图像生成效果效率均显著提升，技术转化为生产力的契机产生。AI 绘画应用推动技术进入民用领域，几月内产出效果便有直观改善，被设计师、游戏工作室等广泛用于辅助生产，生产力变革来临。随着 Stable Diffusion 等应用破圈，用户接受度和参与度持续提高，适用行业不断拓展。AI 绘画是基于深度学习技术的。生成式对抗网络 GAN 配合可对比语言—图像预训练算法 CLIP，解决了跨模态问题，支持文本生成</p>

续表

资料来源	ChatGPT 回答 + 人工校对																		
浙商证券 《AIGC 之 AI 绘画：技术与应用双突破，生产力变革在即》 2022 年 11 月	<p>图像。AI 绘画的实操可分为四个环节：加噪点、去噪点、复原图片和作画。随着图像掩码建模（MIM）、特征处理器（Transformer）、扩散模型（Diffusion Model）和神经辐射场（NeRF），在上述四个环节中发挥作用，AI“画技”显著提升。未来，我们预计深度学习领域将有两大主要前进趋势：大模型和人工通用智能</p> <table border="1" data-bbox="417 447 1025 633"> <thead> <tr> <th>行业</th> <th>应用方式</th> <th>实例</th> </tr> </thead> <tbody> <tr> <td>广告行业</td> <td>借助 AI 制作宣传素材</td> <td>《时尚芭莎》36 周年纪念 AI 海报</td> </tr> <tr> <td>设计行业</td> <td>AI 辅助包装设计、服装设计出图</td> <td>阿里助力犀牛智造推出 AI 设计平台</td> </tr> <tr> <td>营销定制行业</td> <td>借助 AI 生成营销素材</td> <td>百家号自动生成视频功能</td> </tr> <tr> <td>漫画/动漫行业</td> <td>配文自动生成漫画、插画、视频</td> <td>AI 绘制《浣河舞厅》MV</td> </tr> <tr> <td>游戏行业</td> <td>AI 辅助角色、道具、场景设计</td> <td>日本游戏开发者使用 Midjourney 生成素材开发横版射击游戏</td> </tr> </tbody> </table> <p style="text-align: center;">AI 绘画应用行业示例</p>	行业	应用方式	实例	广告行业	借助 AI 制作宣传素材	《时尚芭莎》36 周年纪念 AI 海报	设计行业	AI 辅助包装设计、服装设计出图	阿里助力犀牛智造推出 AI 设计平台	营销定制行业	借助 AI 生成营销素材	百家号自动生成视频功能	漫画/动漫行业	配文自动生成漫画、插画、视频	AI 绘制《浣河舞厅》MV	游戏行业	AI 辅助角色、道具、场景设计	日本游戏开发者使用 Midjourney 生成素材开发横版射击游戏
行业	应用方式	实例																	
广告行业	借助 AI 制作宣传素材	《时尚芭莎》36 周年纪念 AI 海报																	
设计行业	AI 辅助包装设计、服装设计出图	阿里助力犀牛智造推出 AI 设计平台																	
营销定制行业	借助 AI 生成营销素材	百家号自动生成视频功能																	
漫画/动漫行业	配文自动生成漫画、插画、视频	AI 绘制《浣河舞厅》MV																	
游戏行业	AI 辅助角色、道具、场景设计	日本游戏开发者使用 Midjourney 生成素材开发横版射击游戏																	
招商证券 《AIGC：打开元宇宙大门的新魔法棒》 2022 年 11 月	<p>主要介绍了人工智能、区块链、物联网等技术在元宇宙领域的应用，探讨了元宇宙的概念、发展历程、产业布局和价值体系，并对元宇宙相关领域的投资机会进行了分析。报告认为，AIGC（AI+ 区块链+ 物联网+ 图形计算）是打开元宇宙大门的新魔法棒，具有重要的战略意义和巨大的市场潜力，值得关注和投资</p> <p>报告认为，元宇宙是一个集虚拟现实、人工智能、物联网、区块链等多种技术于一体的综合型虚拟世界，将成为未来数字经济和数字社会的核心场景。在元宇宙中，人们可以通过虚拟身份、虚拟资产、虚拟社交等方式参与其中，实现虚拟现实和现实世界间的无缝衔接。报告指出，在 AIGC 技术的支持下，元宇宙将具备高度智能化、高度安全性、高度互联性等特点，为数字经济、文化产业、社交娱乐、教育培训等领域带来广阔的发展空间。报告还列举了一些元宇宙应用场景，如虚拟商城、虚拟竞技、虚拟旅游等，并提出了相应的投资机会，包括虚拟资产、元宇宙平台、虚拟内容等相关产业</p> <p>总的来说，该报告对于当前的科技热点——元宇宙的发展趋势、技术特征和投资机会进行了详尽的分析和研究，对于关注这一领域的人士具有一定的参考价值</p> <div data-bbox="370 1221 1063 1603"> <p style="text-align: center;">代表应用</p>  <pre> graph LR AIGC[AIGC 技术场景] --- B[文字生成] AIGC --- C[图像生成] AIGC --- D[音频生成] AIGC --- E[视频生成] AIGC --- F[虚拟人生成] </pre> <p>文字生成</p> <p>国内：新华社的快笔小新、彩云小梦、小冰、九歌、中文在线、腾讯的梦幻写手及新兴的写作猫、Gilloso 写作机器人等 国外：Jasper.ai、Automated Insights、Narrative Science、textengine.io、Anyword、Phrasee 等</p> <p>图像生成</p> <p>国内：ZMO.ai、图宇宙、百度的文心一格、视觉中国、蓝色光标的创意画廊等 国外：Disco Diffusion、DALL·E 2、Make-A-Scene、Midjourney、Stable Diffusion 等</p> <p>音频生成</p> <p>国内：昆仑万维的 StarX MusicX Lab、网易的网易天音、灵动音科技的 DeepMusic 等 国外：Mubert API 等</p> <p>视频生成</p> <p>国内：字节跳动的剪映、快手的云剪、影谱科技、蓝色光标的蓝标分身等 国外：Meta 的 Make-A-Video、谷歌的 Imagen Video 和 Phenaki 等</p> <p>虚拟人生成</p> <p>国内：倒映有声等 国外：Stable Diffusion、英伟达的 Stylegan 等</p> </div> <p style="text-align: center;">AIGC 的代表性应用</p>																		

续表

<p>资料来源</p> <p>国盛证券《区块链行业深度 AIGC, Web 3.0 时代的生产力工具》 2022 年 11 月</p>	<p style="text-align: center;">ChatGPT 回答 + 人工校对</p> <p>这是一份关于区块链技术在 Web 3.0 时代应用和发展趋势的研究报告。该报告主要从行业背景、技术特点、应用前景等方面对区块链技术进行了分析，重点讨论了其在身份认证、数字资产交易、供应链管理等领域的应用前景，并介绍了各类区块链应用案例。同时，该报告还提到了人工智能、物联网等技术与区块链的融合趋势，以及可能出现的挑战和风险。这份报告旨在为投资者和从业者提供有价值的市场分析和决策参考</p>  <p style="text-align: center;">AIGC 与加密货币挖矿产业链</p>
<p>国海证券《AIGC：内容生产力的革命》 2022 年 12 月</p>	<p>该报告主要探讨了人工智能与内容产业的结合，认为 AIGC 技术将对内容生产力产生深远影响。报告指出，在 AIGC 技术的支持下，内容生产过程中的多项环节将实现自动化和智能化，从而提高制作效率、降低成本、优化用户体验，并大幅度提升内容价值和品质。此外，报告还介绍了 AIGC 技术在文学、电影、游戏、音乐等领域应用的具体案例，并分析了 AIGC 技术对相关行业的投资机会和前景。该报告旨在为读者提供更全面的 AIGC 技术与内容产业的融合分析，以促进双方的深入合作和发展</p>  <p style="text-align: center;">机器对信息认知的三个阶段</p>
<p>中信证券《从 ChatGPT 看生成式 AI 产业机遇与落地场景》 2023 年 1 月</p>	<p>该报告主要探讨了生成式 AI 技术在各个领域的应用前景，并且结合 ChatGPT 等相关技术分析了 AIGC 的发展趋势和市场机遇。此外，报告还详细介绍了 AIGC 在不同行业的应用场景及其商业模式，并提出了一些投资建议。该报告旨在为读者提供有价值的信息，帮助他们更好地了解 AIGC 技术的应用前景和商业机会，以及如何参与相关领域的投资活动</p>

续表

<p>资料来源</p> <p>中信建投证券 《从 ChatGPT 到生成式 AI (Generative AI): 人工智能新范式, 重新定义生产力》 2023 年 1 月</p>	<p>ChatGPT 回答 + 人工校对</p> <p>主要探讨了 ChatGPT 和生成式 AI 的概念、应用以及对生产力的影响。文章介绍了 ChatGPT 技术的基本原理和功能, 并强调了它在自然语言处理领域的广泛应用。此外, 文章还介绍了生成式 AI 的特点和应用场景, 如文本生成、图像生成、视频生成等。同时, 文章也探讨了这些新技术对生产方式的影响, 并提出了人工智能将如何改变生产方式的讨论。最后, 文章呼吁政府、企业和社会各方共同努力, 推动人工智能技术的健康发展, 并充分利用这些新技术的潜力, 为人类创造更美好的未来</p> <p>AI 正在推动第四次工业革命</p>
<p>西部证券 《AIGC 行业专题报告: 2023 年有望成为 AIGC 的拐点》 2023 年 1 月</p>	<p>该报告主要分析了当前 AIGC 技术的发展现状和未来趋势, 并指出 2023 年将是 AIGC 技术应用的一个转折点, 预计在智能制造、智慧城市、医疗保健、金融科技等领域有着广阔的应用前景。此外, 报告还介绍了 AIGC 技术相关产业链和价值链的构成以及重点企业的布局和投资机会。该报告旨在为读者提供更深入的 AIGC 技术研究和产业投资参考, 以促进 AIGC 技术在各行各业的推广和应用</p> <p>AIGC 产业链</p>
<p>腾讯研究院 《AIGC 发展趋势报告 2023》 2023 年 1 月</p>	<p>ChatGPT 直接反馈: 经过检索, 没有找到腾讯研究院发布的《AIGC 发展趋势报告 2023》。但是腾讯研究院曾发布过类似主题的报告, 例如《2020 人工智能白皮书》等, 这些报告主要分析了人工智能技术的发展现状和未来趋势, 并探讨了人工智能在商业、医疗、教育等领域的应用和前景。这些报告旨在为政府、企业和公众提供更详尽的人工智能发展情况和应用建议, 以推动人工智能在各个领域的创新和发展</p>

续表

资料来源	ChatGPT 回答 + 人工校对
<p>腾讯研究院 《AIGC 发展趋势 报告 2023》 2023 年 1 月</p>	<p>由于 ChatGPT 反馈其知识库中尚不包括这篇内容，因此笔者分段输入了报告的文本，得到了以下的总结内容：</p> <p>AIGC 技术通过生成算法模型、预训练模型和多模态技术的突破创新，实现了文字、代码、图像、语音、视频、3D 物体等类型的内容和数据生成。其中，预训练模型技术极大地提升了 AIGC 模型的通用化能力和工业化水平。同时，AIGC 领域正在加速形成三层产业生态并持续创新发展，正走向模型即服务 (MaaS) 的未来。未来，随着算法的不断进步，AIGC 领域将迎来更多激动人心的应用，多模态 AI 将日益成为主流，推动 AIGC 领域的蓬勃发展</p> <p>AIGC 产业生态体系呈现上中下三层架构，分别为基础层、中间层和应用层。基础层以预训练模型为基础搭建 AIGC 技术基础设施层，在国外以 OpenAI、Stability.ai 为代表，通过受控 API、开源等方式输出模型能力。中间层是在预训练模型基础上，通过专门的调试和训练，快速抽取形成垂直化、场景化、定制化的小模型和应用工具层，例如 Novel-AI 等二次元画风生成模型。应用层依托底层模型和中间层的垂直模型，各厂商进一步开放面向 C 端和 B 端用户的各种各样的 AIGC 产品和服务，例如群聊机器人、文本生成软件、头像生成软件等。目前，美国围绕 AIGC 生成繁荣的生态，中国也有望凭借领先的 AIGC 技术赋能千行百业。未来，预计模型即服务 (MaaS) 将对商业领域产生巨大影响</p> <p>首先，AIGC 有望成为新型的内容生产基础设施，塑造数字内容生产与交互新范式，持续推进数字文化产业创新。其次，AIGC 的商业化应用将快速成熟，市场规模会迅速壮大。最后，AIGC 还将作为生产力工具，不断推动聊天机器人、数字人、元宇宙等领域发展。这些发展趋势都预示着 AIGC 在未来将会有广泛的应用和重要的作用</p> <p>基于 AIGC 技术的合成数据在产业互联网领域有重大发展。合成数据作为真实世界数据的廉价替代品，为 AI 模型训练开发提供强大助推器，推动实现 AI 2.0，并助力破解 AI “深水区”的数据难题，持续拓展产业互联网应用空间。此外，合成数据加速构建 AI 赋能、数实融合的大型虚拟世界，将成为 AI 数实融合的关键载体。由于合成数据对人工智能未来发展的巨大价值，合成数据正加速成为一个新产业赛道，涌现了合成数据即服务 (SDaaS) 这一全新商业模式</p> <div data-bbox="358 1237 1117 1652" style="text-align: center;"> <p>基础层：预训练模型</p> <p>中间层：垂直化、场景化、个性化模型</p> <p>应用层：图像、语音和文字生成等 各种各样的 AIGC 的应用</p> <p>代表企业：OpenAI, stability ai, 小冰, Google DeepMind, Meta AI, NVIDIA, NovelAI, 意间AI绘画</p> </div> <p>基于 AIGC 产业生态体系三层架构改编</p>

续表

资料来源	ChatGPT 回答 + 人工校对												
招商证券 《ChatGPT 快速流行，重构AI商业模式》 2023年2月	<p>具备生成创造能力的 AIGC 技术已经成为 AI 技术发展的新趋势。ChatGPT 不仅代表人 AI 交互类的聊天工具，通过其底层逻辑，我们判断 ChatGPT 技术可以应用于多个领域。微软在官方博客宣布持续加大对 ChatGPT 母公司 Open AI 投入，后续投资金额将高达数十亿美元，预计未来公司旗下所有产品有望全线整合 ChatGPT。此外，亚马逊、BuzzFeed 等互联网大厂同样宣布与 ChatGPT 展开合作。AIGC 模型有大量参数及数据参与训练，因此对算力基础设施有较高要求，云计算服务商基于此逻辑布局 AIGC 产业链。模型必须具备大量语料素材、强大的 AI 团队以及充足资金支持，因此具备 AIGC 大模型构建能力的企业通常为海内外科技巨头</p> <p>AIGC 领域头部企业将逐渐构建自身壁垒，剩余的初创公司将依附头部企业去做应用，产业链未来将会分为技术提供方和应用开发者两种</p> <table border="1" data-bbox="316 666 1132 979"> <thead> <tr> <th data-bbox="316 666 602 687">AIGC 变现模式</th> <th data-bbox="602 666 1132 687">案例/收费方式</th> </tr> </thead> <tbody> <tr> <td data-bbox="316 687 602 733">作为底层平台接入其他产品对外开放，例如，GPT3 对外提供 API 接口，按照数据请求量和实际计算量计算</td> <td data-bbox="602 687 1132 733">采用的四种模型分别采用不同的按量收费方式</td> </tr> <tr> <td data-bbox="316 733 602 778">按产出内容收费</td> <td data-bbox="602 733 1132 778">例如，包括 DALL·E、Deep Dream Generator 等 AI 图像生成平台大多按照图像张数收费</td> </tr> <tr> <td data-bbox="316 778 602 824">模型训练费用，适用于 NPC 训练等个性化定制需求较强领域</td> <td data-bbox="602 778 1132 824">根据具体训练量收费</td> </tr> <tr> <td data-bbox="316 824 602 869">根据具体属性收费</td> <td data-bbox="602 824 1132 869">例如，版权授予（支持短期使用权、长期使用权、排他性使用权和所有权多种合作模式，用友设计图案的版权）、是否支持商业用途（个人用途、企业使用、品牌使用等）、透明框架和分辨率等</td> </tr> <tr> <td data-bbox="316 869 602 979">直接对外提供软件（SaaS）</td> <td data-bbox="602 869 1132 979">例如，个性化营销文本写作工具 AX Semantics 以约 1900 元/月的价格对外出售，并以约 4800 欧元/月的价格提供支持定制的电子商务版本。大部分 C 端 AIGC 工具以约 80 元/月的价格对外出售，最新 ChatGPT 订阅计划售价为 20 美元/月</td> </tr> </tbody> </table> <p style="text-align: center;">AIGC 商业模式梳理</p>	AIGC 变现模式	案例/收费方式	作为底层平台接入其他产品对外开放，例如，GPT3 对外提供 API 接口，按照数据请求量和实际计算量计算	采用的四种模型分别采用不同的按量收费方式	按产出内容收费	例如，包括 DALL·E、Deep Dream Generator 等 AI 图像生成平台大多按照图像张数收费	模型训练费用，适用于 NPC 训练等个性化定制需求较强领域	根据具体训练量收费	根据具体属性收费	例如，版权授予（支持短期使用权、长期使用权、排他性使用权和所有权多种合作模式，用友设计图案的版权）、是否支持商业用途（个人用途、企业使用、品牌使用等）、透明框架和分辨率等	直接对外提供软件（SaaS）	例如，个性化营销文本写作工具 AX Semantics 以约 1900 元/月的价格对外出售，并以约 4800 欧元/月的价格提供支持定制的电子商务版本。大部分 C 端 AIGC 工具以约 80 元/月的价格对外出售，最新 ChatGPT 订阅计划售价为 20 美元/月
AIGC 变现模式	案例/收费方式												
作为底层平台接入其他产品对外开放，例如，GPT3 对外提供 API 接口，按照数据请求量和实际计算量计算	采用的四种模型分别采用不同的按量收费方式												
按产出内容收费	例如，包括 DALL·E、Deep Dream Generator 等 AI 图像生成平台大多按照图像张数收费												
模型训练费用，适用于 NPC 训练等个性化定制需求较强领域	根据具体训练量收费												
根据具体属性收费	例如，版权授予（支持短期使用权、长期使用权、排他性使用权和所有权多种合作模式，用友设计图案的版权）、是否支持商业用途（个人用途、企业使用、品牌使用等）、透明框架和分辨率等												
直接对外提供软件（SaaS）	例如，个性化营销文本写作工具 AX Semantics 以约 1900 元/月的价格对外出售，并以约 4800 欧元/月的价格提供支持定制的电子商务版本。大部分 C 端 AIGC 工具以约 80 元/月的价格对外出售，最新 ChatGPT 订阅计划售价为 20 美元/月												
财通证券 《技术奇点已至，多模态时代开启》 2023年3月	<p>全面多模态将实现文本、音频、图像、视频等全方位信息源的输入/输出，实现在政务、金融、教育、医疗、传媒等垂直行业的商业化落地，AIGC 在各个垂直应用领域应用的时代已然开启。未来已至，人工智能将进入奇点时刻：</p> <ul style="list-style-type: none"> ● 算力是 AI 模型的能源，将最直接受益于 AI 的普及 ● 开发海外应用和国内基础层的公司将在中短期受益于行业“从 1 到 10”的快速拓荒阶段 ● 拥有底层语言模型及机器学习算法框架开发能力的公司有望作为行业边界的开拓者长期受益于产业趋势的浪潮 												

资料来源：ChatGPT 及各大机构的报告

在过去的几年里，对 AIGC 的投资大幅增长。根据 CB Insights 的一份报告，自 2015 年以来，全球 AIGC 领域的投资规模一直呈增长趋势。AIGC 初创公司的总投资资本从 2018 年的 73 亿美元增加到 2020 年的 199 亿美元，增长了近两倍。2019 年全球 AIGC 领域的投资金额达到了近 440 亿美元，2020 年受新冠疫情影响有所下降，但仍然接近 350 亿美元。这种增长主要是由于对 AI 和自动化产生的内容的需求增加，以及自然语言处理（NLP）和自然语言生成（NLG）技术的进步。AIGC 领域的主要参与者，如 Persado、Automated Insights 和 Narrative Science，在

过去几年都有大量投资。从投资类型来看，种子轮和天使轮投资仍占据着较大的比重，但随着行业的发展，中后期和上市融资的投资也在逐步增加。从投资领域来看，图像识别、语音识别、自然语言处理等技术是最受关注和投资的领域。此外，医疗健康、金融、教育等行业的应用也成为投资的热点。亚洲地区在全球AIGC领域的投资中日益发挥着重要的作用。中国、印度等亚洲国家的AIGC企业和初创公司正在吸引越来越多的国内外投资者，其中，中国是全球最大的AIGC投资市场之一，2019年中国AIGC领域的投资规模约为150亿美元。

除了AIGC领域的专业投资机构之外，越来越多的传统行业企业也开始涉足AIGC领域的投资。这些企业希望通过投资AIGC技术和企业，加速自身数字化转型的进程，提升自身的竞争力。另外，政府在推进数字经济发展的过程中，也将AIGC技术视为重要支撑。因此，在一些国家和地区，政府会设立专项资金用于支持AIGC领域的研究和应用，并鼓励民间投资。需要注意的是，尽管AIGC领域的投资呈现出较快的增长态势，但由于该领域的技术和商业模式还比较复杂，投资具有一定的风险性。因此，投资者需要通过深入的调研和分析，选择具备实力和潜力的AIGC企业进行投资。同时，监管部门也需要制定相应的政策和标准，规范AIGC领域的投融资活动，维护市场的稳健运行。

如图1-7所示，2023年1月23日，微软宣布与OpenAI展开全新合作，未来将追加投资数十亿美元，2月2日，微软宣布旗下产品将全线整合ChatGPT，对于微软而言，ChatGPT在拟人化交流、即时生成内容等方面对必应（Bing）的赋能有望助其突破谷歌的桎梏，作为回应，谷歌即刻投资Anthropic并计划推出类似ChatGPT的大型语言模型。除了加码AI文本、代码生成以外，海外巨头如Meta、Netflix亦着力布局音频、视频等内容生成，未来人机协同或是大势所趋。

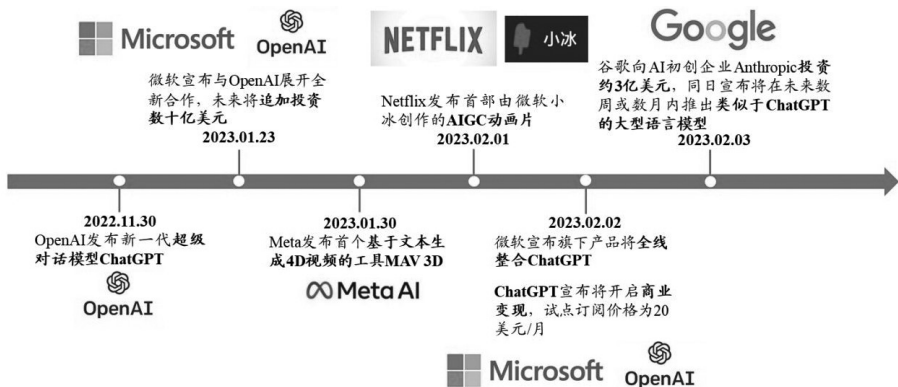


图 1-7 AI 海外巨头争相加码 AIGC

资料来源：慧博智能投研

麦肯锡的报告显示，全球数字化劳动力市场规模将迅速扩大，2030年有望达到1.7万亿元，其中交互应用、企业流程优化、工业应用、特殊应用规模分别达6247亿元、5213亿元、3215亿元、2583亿元。AIGC是数字化劳动力的又一次爆发和最主要的组成部分，其领域的投资规模也在不断扩大。未来，随着技术的进一步成熟和应用场景的拓展，AIGC领域的投资规模还将继续增长。

最后，我们利用 ChatGPT 来挖掘一下目前关于 AIGC 报告内容的共性。

(1) 技术原理：介绍 AIGC 的技术原理，包括机器学习、深度学习、自然语言处理等相关技术。

(2) 应用领域：探讨 AIGC 在不同领域的应用情况，如广告、新闻、文学创作、音乐、视频等。

(3) 优缺点分析：对于 AIGC 的优势和局限性进行分析，包括生成效率、内容质量、版权问题、道德风险等方面。

(4) 行业趋势：展望 AIGC 未来的发展趋势，探讨其可能带来的变革和影响，包括人类创作角色、版权保护、产业链变革等方面。

(5) 投资价值：评估 AIGC 的商业化前景和投资价值，对相关公司及市场进行分析和预测。

总之，现在关于 AIGC 的报告内容主要围绕着技术原理、应用领域、优缺点分析、行业趋势和投资价值这几个方面展开，旨在为读者提供全面的了解和参考。

1.2 AIGC 与 Web 3.0 及数字孪生

Web 3.0 是下一代互联网的体系架构，数字孪生是 Web 3.0 中一种重要表现形式——即在虚拟世界中可以映射现实世界中的对象，这些虚拟的对象能够像现实世界中的对象一样产生交互。AIGC 是 Web 3.0 与数字孪生的内容生产基础设施，是数字孪生世界中虚拟对象的“生成器”和“活化剂”。我们无法想象没有 AIGC 赋能的数字孪生，因为内容生成的工作量实在庞大，这些内容不仅是文字、图片、音频、视频，还包括太多个性化的内容，例如面容表情甚至表达的手势与语气。在工业元宇宙（工业界更愿意用元宇宙一词，而不是元宇宙）中，情境则更为复杂，需要呈现的精度与准确度可能是平常要求的几十倍之上。所以也有人说，AIGC 按下了元宇宙和数字孪生实操落地的快进键！

Web 2.0 在过去 20 年里为我们提供了很好的服务，但由于其广泛的隐私和安全问题，我们已步入一个被称为 Web 3.0 的新数字时代的初级阶段。Web 3.0 也被

称为语义网，是互联网发展的新阶段，有望带来更多的用户授权、开放和隐私保护。Web 3.0 应用程序旨在通过建立去中心化的基础设施来实现更高的安全性和互操作性，从而消除了对中央服务器的需求。以下列出了构成 Web 3.0 的五个关键构成要素：

(1) 语义网。语义网使用 AI 来理解用户或客户可能的意思或意图。它旨在根据搜索词的实际含义而不是关键词或数字，给搜索提供更准确的理解。

(2) AI。Web 3.0 的 AI 旨在更好地理解某人正在搜索的内容，以提供相关性更强的结果。

(3) 3D 图形和空间网络。虚拟现实 (VR) 头盔和现实图形的使用使网站在本质上变得更加逼真，为元宇宙建立一个 3D 渲染的、无边无际的虚拟世界。

(4) 区块链和加密货币。Web 3.0 的去中心化的关键是使用区块链和加密货币，这消除了中间商，使各方之间能够直接交易。

(5) 无处不在的连接。由于宽带、5G、Wi-Fi 和物联网，Web 3.0 应用的特点是不断连接。

在完全实现的 Web 3.0 空间网络中，物理世界中每个建筑的每个元素都将完全数字化。每个人都会有虚拟的化身，人们将能够在虚拟的工作或会议场所漫游。表 1-4 呈现了 Web 1.0、Web 2.0 和 Web 3.0 的特征及相互之间的区别。

表 1-4 Web 1.0、Web 2.0 和 Web 3.0 之间的区别

Web 1.0	Web 2.0	Web 3.0
通常是只读的	强可读写性	读写互动
拥有的内容	共享的内容	合并的内容
视觉 / 交互式网络	可编程网络	关联数据网络
主页	维基和博客	直播
网页	网络服务端点	数据空间
HTML/HTTP/URL/ 门户	XML/RSS	RDF/RDFS/OWL
页面浏览量	每次点击成本	用户参与
文件 / 网络服务器、搜索引擎、电邮、P2P 文件共享、内容和企业门户网站	即时通信, Ajax 和 Java Script 框架, Adobe Flex	个人智能数据助理、本体论、知识库、语义搜索功能
目录	给用户打上标签	用户行为
专注于公司	专注于社区	专注于个人
《大英百科全书》在线	维基百科	语义网
横幅广告	互动广告	行为广告

资料来源: Simplilearn, John Terra, 2023

尽管 Facebook 在 2021 年将其公司名称改为 Meta，但元宇宙并不是 Facebook 首创，它只是元宇宙的积极参与者与推动者。元宇宙是一个独立于设备和供应商的集体虚拟空间，由虚拟增强的物理和数字现实的融合创造。它有自己的虚拟经济，由数字货币和不可伪造的代币（NFTs）促成。教育、医疗、零售和虚拟活动可以提供更加沉浸的体验，它们不需要创建自己的基础设施。元宇宙将提供框架，虚拟活动以呈现更多的综合产品，而零售业可以提供更多的沉浸式购物体验。在这方面，元宇宙与 Web 3.0 的空间网络并无二致，它是一种专注于 3D 图形和真实世界图像的虚拟沉浸式体验，而不是像当前网络体验那样的 2D 图形和文本。用户不是通过单击一个网站的链接来访问信息，而是在虚拟空间中行走、观察与互动。

Web 3.0 和元宇宙这两种技术之间最大的区别是，人们使用 Web 3.0 来访问元宇宙，就像汽车使用道路一样。Web 3.0 是关于去中心化的所有权和控制权，并将网络放在其用户和社区的手中。另外，元宇宙是一个共享的数字现实，使用户能够相互联系，建立经济并进行实时互动——它并不关心谁拥有它。Web 3.0 也建立在区块链和加密货币的基础上，而元宇宙使用 AR/VR 和数字货币等技术。这是由于 Web 3.0 是去中心化的，没有大公司的影响或控制。两者在使用方式上也有所不同，Web 3.0 是一套关于如何使用和管理互联网的新标准，元宇宙是关于游戏、社交媒体、零售和其他体验。它们的共性在于，元宇宙将继续存在于表层和深层网络中，尽管 Web 3.0 仍然经常被称为是去中心化的，然而，在社会媒体控制平台方面，元宇宙仍将是中心化的，两者都建立在先进的技术，并将不断发展。语义网是元宇宙和 Web 3.0 的共同点。AI——这两种技术的另一个关键组成部分——将是建立一个复杂的用户界面不可或缺的。在技术意义上，与区块链一起取得的许多进展在两种技术中创造了共同点。每一个新的区块链概念都被评估为整合 Web 3.0 引擎的潜在模块，该引擎将为元宇宙的商品和服务提供动力。Web 3.0 和元宇宙都处于早期阶段。最终的产品将在几年后出现，而且可能会有很大的不同，因为技术并不总是与它的使用愿景相匹配。

2010 年以来 Web 3.0 月度活跃开发者数量趋势如图 1-8 所示。

Web 3.0 应用程序将在分布式区块链和云网络上运行。在设想的数字领域中，机器可以直接与其他机器和用户交流，然而这需要机器理解数字内容。AIGC 为此提供了解决方案，也正在成为 Web 3.0 的基本构件之一，通过更好的内容生成、推荐和改善人机互动来改善用户体验，由此 AIGC 也可以理解为 Web 3.0 的认知层，提供深度学习算法和分析能力，使机器与机器和用户之间“理解”在线内

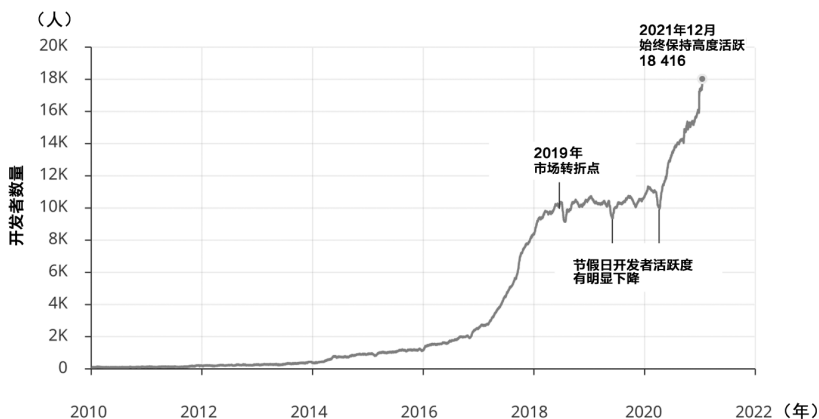


图 1-8 2010 年以来 Web 3.0 月度活跃开发者数量趋势

资料来源: Electric Capital

容。从本质上讲，超大规模算法模型将训练 AIGC 识别不同类型的内容，并为其赋予意义。这样一来，搜索引擎就不会只是推荐最受欢迎的内容类型，而是对其有一个全面深入的理解，据此更友好地与人类互动，以帮助改善整体的用户体验。

在 Web 2.0 中，AIGC 已经在各个领域开始了广泛的探索，在 Web 3.0 中自然会有很多应用方向。在与文本生成相关的 AI 工具方面已取得了突破性进展，AIGC 在文本创作中的应用包括编码、翻译和写作。文本创作本质上是对语言的使用。由于编程语言对 AI 来说相对更有结构性，更容易学习，但人类语言需要结合语境、语义等，因此，文本生成最成熟的应用场景是编码，代表性作品如 AI 出品的 GitHub Copilot。用户用文字输入代码逻辑，它可以快速理解，并根据海量的开源代码生成子模块，供开发者使用。现在，GitHub Copilot 生成的代码中有近 40% 是由 AI 编写的。虽然 Web 3.0 中的 SDK 等模块化插件提高了开发者的编程速度，但未来随着 AIGC 技术的普及，加密协议的开发效率可能会进一步提升。理想情况下，AIGC 可以自动检测市场需求或空缺，然后独立编程并生成新协议。在人类语言的内容创造方面，AIGC 也取得了很大的进展。目前，翻译的发展已经取得了很大的领先优势。Roblox 通过机器学习将英文开发的游戏自动翻译成其他 8 种语言，包括中文、德文、法文等（如图 1-9 所示）；腾讯开发的 Dreamwriter 新闻写作系统可用于 22 个规范的写作场景，平均发稿速度快至 0.46 秒；红杉资本的《生成式 AI：创意新世界》一文中，部分内容由 GPT3 自然语言模型撰写，阅读体验较好，还兼顾了行文流畅、逻辑清晰等写作要求。

AIGC 也将为 Web 3.0 的文本创作做出巨大贡献。Web 3.0 的新闻媒体和研究机构正面临着内容生态的双边困境。例如，虽然 CoinDesk 和 Messari 的产出质量



图 1-9 Roblox 自动将英文游戏翻译成其他语言
资料来源: Roblox, Shine!NFAITH, Muse Labs

很高，但很难扩大生产规模。此外，受写作语言、翻译效率和准确性的限制，内容传播将进一步减少。另外，虽然推特上的内容很庞杂，但观点的质量无法保证。由于信息没有按照重要性和时效性等进行分类，因此呈现形式比较凌乱，没有分组、没有分类，也没有去重。显然，用户的需求并没有得到针对性的满足。同时，用户将面临信息过载的问题，在无效的内容上浪费了大量的时间。因此，Web 3.0 组织在平均生产规模和平均内容质量方面都明显落后于 Web 2.0 的同行。然而，Web 2.0 组织的规模和质量往往是基于众包策略，需要大量的初始投资。为了保证内容的质量，合格的分析师通常需要经过长期沉淀和强化培训，企业必须投入时间和培训成本。同时，为了保持产出规模，企业必须付出极高的人工成本进行大规模招聘。这类模式有两个明显的缺点：一个是成本过高；另一个是后期人才流失的风险，导致成本完全沉没。随着后续技术的进步，分析师至少可以节省总结标题和摘要的时间，而 AI 则能够通过理解全文直接生成。从长远来看，“合格的 AI 分析师”将迅速产生。Web 3.0 机构将大幅降低成本，同时提高内容生成的规模和质量，从而促进整个细分市场和整个行业的发展。信息协议、新闻协议或研究协议甚至可能出现在 Web 3.0 中。

AIGC 有可能引发 Web 3.0 音乐的新一轮创新。AIGC 开启了歌曲制作、歌词生成等方面的应用，互动性和实时性得到进一步加强。例如，自适应音乐平台 LifeScore 可以实时动态地安排音乐。一旦用户输入一系列的音乐素材，AI 就会对其进行改变、变形和重新混合，从而生成一场即时的音乐会。如图 1-10 所示，2020 年 5 月，LifeScore 为 Twitch 互动电视系列 *Artificial* 提供了自适应配乐，随着故事的展开，它可以根据观众的情绪状态影响配乐。系统提供了四种典型的情绪：快乐、紧张、神秘、悲伤。用户只需要通过口头语言或使用相关的表情就可以修改乐谱。从短期来看，AIGC 可以帮助创作者改编、再创作，或者直接辅助

音乐创作，大大减少他们的工作量，提高工作效率。从长远来看，一些音乐平台已经在 Web 3.0 中出现，随着 AIGC 技术的引入，协议可能会根据听众的个人喜好生成定制歌曲。该平台不仅可以极大地削减版权费用，而且用户还可以减少歌曲的支付。此外，用户还可以发布由 AI 创作的独家歌曲，为自己带来收益，从而增强 Web 3.0 音乐市场的创作者经济。

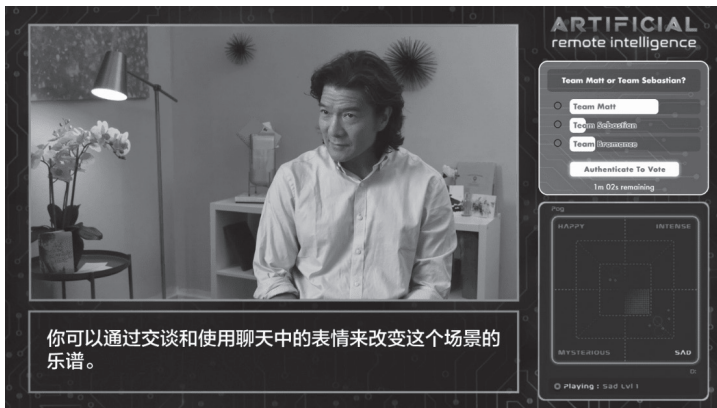


图 1-10 LifeScore 根据观众对剧情的感受，实时创建背景音乐
资料来源: Twitch

除了上述前沿方向外，AIGC 在其他 Web 3.0 市场领域也有很大的潜力，具体包括以下应用：

- NFT 的主体是图片或艺术作品。目前，许多 AI 模型已经收集了整个艺术史和流行文化的数据。任何用户都可以随意生成自己喜欢的 NFT。不同的 NFT 需要有不同的面孔、服装和情感特征。传统的生成方法承担着高成本和低效率，创作者需要进行原型设计、多次建模和渲染等工作，而 AIGC 可以帮助创作者在前期更有效地尝试草图，并在后期节省人力来完成画面的细节。在未来，AIGC 有可能实现 NFT 的低成本量产。此外，UGC 创作容易被复制和传播，侵权问题经常发生。然而，NFTs 具有唯一性、不可分割性和可交易性，可以解决资产防伪、确权、可追溯等问题，加强版权保护。
- AIGC 也在改进跨模状态下的生成，如文本生成图像/动画，反之亦然。
- AIGC 的进步也将促进 Web 3.0 社会市场领域的发展。真实的人不可避免地会有一些缺点，但 AI 可以创造出用户喜欢的虚拟人物，因为 AIGC 生成的虚拟人物将完全根据用户需求定制。用户可以定制或利用模板来定义人物的属性，如家庭、职业、年龄等。AI 将帮助虚拟人物在特定场景下在外观和动作上表现得更像真实的人，并赋予它们语言表达和互动的功

能，以体现一定的移情能力。此外，虚拟人物伴随着比人类更丰富的知识储备和更快的更新频率，不需要休息。因此，可以预见，虚拟人物在某些特定领域提供的娱乐和服务将与真人相媲美，甚至超越真人。例如，虚拟人物将通过与用户的交流继续学习，实现情感上的陪伴。参照 Web 2.0 中的 ACGN 群体和社交软件重度用户，Web 3.0 的社交市场在 AIGC 的支持下无疑会变得更大。AIGC 在 Web 3.0 教育中的应用可能产生意想不到的效果。由于 AI 的学习模式是相对结构化和条理化的，由 AIGC 制作的教科书和讲座可能会降低理解障碍，帮助受众更容易吸收知识。综上所述，AIGC 在 Web 3.0 领域的应用是相当广阔的。

自然语言处理（NLP）是语言学和人工智能最迷人的子领域之一，也将在语义网中发挥突出作用，使 AIGC 算法能够分析和逐步理解在线通信。NLP 在语义网中的实施对于创造一个更安全的在线环境和推进其一些最有前途的元素至关重要，例如，基于 AIGC 的聊天机器人来实现客户支持流程的自动化或更好的内容索引算法。AIGC 作为 Web 3.0 认知层的应用，可为用户带来以下好处：

- 产生更多个性化的建议。AIGC 驱动的推荐引擎已经在亚马逊和 Netflix 等流行的 Web 2.0 应用中被采用。AI 算法可能成为 Web 3.0 推荐引擎的核心，因为它们有能力分析大量的用户数据并在个人层面上创建预测模型。基于 AIGC 的推荐引擎将带来更好的导航性和用户体验，这要归功于它们能够真正“理解”用户的偏好并提供更多个性化的推荐。
- 更智能的 dApp 和 NFT。随着区块链开发者整合 AIGC，Web 3.0 将引入更智能的分散式应用（dApp），使其具有更先进的现实世界效用。一个引人入胜的早期例子是 Alethea AI 的 Alice——第一个具有自我学习能力的非可替换代币（NFT），当它从每一次新的互动中学习时，会改变与人的互动方式。Alice 有自己的个性，能够与互联网用户进行详尽的对话，并在这个过程中向他们学习。随着 AI 算法的进步，下一代 dApp 和 NFT 可能会从它们的功效中受益，因为分散的应用程序获得了更好的数据管理和分析能力，而下一代 NFT 将带有类似人类的行为。
- 更人性化的营销广告。虽然社交媒体巨头和营销公司已经在收集我们的浏览数据，以提供更多相关的广告，但这并不符合互联网用户的最佳利益。首先，互联网用户在没有意识到其真正价值的情况下，将自己的数据随意签给了暴利的第三方。其次，这些公司采用的侵入性数据挖掘做法在互联网用户中引起了广泛的隐私担忧。Web 3.0 的去中心化精神旨在通过让人

们重新控制自己的数据而带来更多的用户权利——这些数据可以选择与广告商共享，由广告商直接向用户付费。通过 AIGC 应用，广告公司将能够分析更大量的数据，并在个性化用户广告时考虑更多因素。因此，Web 3.0 用户将收到更多相关广告，而人工智能算法将不断适应每个人的喜好。

- 增强的内容安全性。内容安全是 Web 2.0 最具挑战性的问题之一，因为我们独特的主观视角，人类的内容安全索引往往导致元数据的不一致。AIGC 消除了人类思维的偏见，这可以使内容元数据更加一致和可靠。视频内容安全特别重要，因为每天上传的视频数量远远超过了人类的监督能力。然而，由于 AIGC 算法的大规模可扩展性，AI 有望彻底改变 Web 3.0 的内容安全。

最后，我们再与 ChatGPT 做一些互动。当我问它“你认为，到 2030 年会有多少书籍是 AI 生成的？”时，它的回答反映了 AIGC 的局限性，即目前 AIGC 的功能只能对已有的内容与数据进行提取，而无法像人类一样推论或预测，但它的回答也不是一无是处：“我作为一个语言模型，无法对未来做出确切的预测。但是随着 AI 技术的不断发展，AIGC 在出版业和其他领域中的应用正在增加。到 2030 年，可能会有大量的书籍是通过 AI 生成的，但具体数字取决于各种因素，如技术进步、市场需求、行业规范等。”所以我又问了一个具体的问题：“如何使用 ChatGPT 和 Midjourney 来完成一本书？”ChatGPT 的回答就非常有价值了：

(1) 确定书籍类型和主题。ChatGPT 可以帮助您了解特定领域的相关信息，以及目前市场上类似书籍的趋势。

(2) 开始规划书籍大纲。使用 Midjourney 的工具，如笔记、大纲和故事板，可以帮助您组织您的思想并开始编写您的书籍。

(3) 使用 ChatGPT 进行研究和阅读。ChatGPT 可以为您提供相关主题的资料和参考文献，以帮助您更深入地了解您所写的内容。

(4) 在 Midjourney 平台上开始编写书籍。在这个过程中，您可以使用 ChatGPT 作为一个资源，以获得灵感和帮助，确保您的内容准确、清晰和易于理解。

(5) 审查和编辑内容。使用 Midjourney 的审查和编辑工具，以及 ChatGPT 的建议和指导，确保您的书籍达到专业水平，并符合行业标准。

(6) 发布书籍。使用 Midjourney 的自助出版工具，将您的书籍发布到亚马逊等在线书店，从而向全世界的读者推广您的作品。

总之，ChatGPT 和 Midjourney 都可以为您提供宝贵的资源和工具，帮助您完成一本成功的书籍。但是，只有创造力和良好的写作技巧才是真正的关键。