

生成式 AI 多领域落地，赋能传媒行业发展

——AI 产业系列深度报告（一）

强于大市（维持）

2024 年 06 月 28 日

行业核心观点：

生成式人工智能 (Generative AI)，指的是通过人工智能技术自动生成内容的生产方式。着重于创造新的、有创意的数据，其关键原理在于学习和理解数据的分布，进而生成具有相似特征的新数据。自“十三五”、“十四五”规划将人工智能作为重点任务，明确了战略发展目标之后，国家陆续出台指导性政策支持性政策推动生成式 AI 发展，目前在图像、文本、音频、视频等多种领域都有广泛的应用。

投资要点：

图像生成：Midjourney 稳住行业龙头地位，AI 文生图迈向多模态融合。

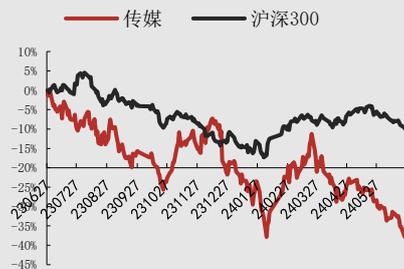
AI 文生图，即通过人工智能技术解析文字描述内容，提取关键信息，利用深度学习模型形成草图，再经过优化算法细化图像细节，增强图像质量，输出符合描述的视觉内容，是近年来快速发展的一项技术。

- 1) **自身发展**，自 2023 年起，图像生成已经在摄影艺术、设计、二次元、CG 艺术和影视制作等多个行业得到广泛应用，目前该技术不仅限于单一的图像生成模型，而是开始向多模态融合方向发展，市场规模高速增长，网民参与度显著提升；
- 2) **产品**，目前市场上的产品已经发展比较成熟，主流的 AI 文生图工具分别有 Midjourney V6、Adobe Firefly 3、Stable Diffusion 3、Dalle 3 等。Midjourney 产品迭代速度快于同行，整体图像产出质量显著提高，同时通过庞大的用户量及用户数据，使得积累的数据集具有独家性，并依托这些数据进行反复训练，来构建自我艺术风格，巩固自身竞争优势，稳住文生图龙头地位。

视频生成：新品推出受到市场关注，创新点赋能行业发展。 AI 文生视频的本质是能够根据输入的文本描述自动生成相应的视频内容。这项技术结合了自然语言处理 (NLP)、计算机视觉、图像生成和动画技术等多个领域的研究成果，通过训练大量文本和视频数据，依靠文本解析、图像生成、动态渲染等核心技术，将文字转化为图像，并赋予这些图像动态效果，全自动化形成视频文件。

- 1) **自身发展**，技术更迭带动行业发展，真正兴起点在于将扩散模型在文生图领域经验拓展至文生视频，成为 AI 视频生成领域的主流技术路径，带动行业产品推出频率及数量显著提升，但由于存在缺乏相符性高的训练数据、计算成本高昂、信息冗杂难处理等挑战，目前市场仍处于初期阶段，潜力可期；
- 2) **产品**，目前市场上受到关注度较高的文生视频产品分别是来自的 Sora、Pika1.0、近期快手推出的可灵、Luma AI 推出的 Dream Machine 以及 Runway 推出的 GEN-3。新品的推出通常伴随着领域技术的突破，从目前的文生视频能力来看，已经具备了一定的产业生产力，

行业相对沪深 300 指数表现



数据来源：聚源，万联证券研究所

相关研究

踏 AI 之浪潮，扬新生态之帆

《黑神话：悟空》开启预购登顶 Steam 全球热销榜，6 月 15 款进口游戏版号过审丰富种类+“剧转影”+好莱坞多项优势加持，端午档票房有望维持较好表现

分析师：夏清莹

执业证书编号：S0270520050001

电话：075583223620

邮箱：xiaqy1@wlzq.com.cn

分析师：李中港

执业证书编号：S0270524020001

电话：02032255208

邮箱：lizg@wlzq.com.cn

未来将赋能影视产业链，为其降本增效提供助力。

音频生成：音乐生成类市场潜力大，企业探索视频生音频技术。AI 音频利用先进的人工智能技术和复杂算法来创造音频内容，包括语音合成、音乐制作和声音效果合成等。通过融合机器学习和深度学习算法，AI 音频生成技术能够精确地模仿人类的语音、音乐节奏和声音效果，生成自然且逼真的音频，广泛应用于娱乐、广告、教育和新闻传播等多个领域。

- 1) **自身发展**，AI 音频生成行业的产业链以中游 AI 音频算法和服务平台为核心，上下游协同发展，从应用场景来看，语音合成占主导，音乐生成潜力大，语音识别应用成熟；
- 2) **产品**，标杆性的音乐生成类音频 AI 是由 Suno AI 研发的 Suno 系列产品，于今年 5 月推出了 V3.5 版本，维持了与 Open AI 合作，一次性完成歌词、演唱、编曲、配乐等全部流程，领先于市面上单纯生成音乐的同类型产品，同时关注到 ElevenLabs 及 Google DeepMind 均宣布了自动匹配画面 AI 生成音效产品，即无需人工输入提示词也可以为视频配音，实现真正的 AI 音频及 AI 视频合作，这将带领文生视频迈入有声时代，实现行业再突破，双向促进市场繁荣发展。

传媒行业应用：AI 赋能传媒子行业，助力技术变革。在影视、游戏等多个传媒互联网的细分领域进行广泛应用，从而提升内容生产效率，降低内容生产成本，助力行业变革，建议关注已有生成式 AI 应用落地的公司。

- 1) **游戏方面**，生成式 AI 能压缩游戏整体项目的研发周期与人员规模，大幅降低游戏制作成本，降本增效；对已有的游戏进行产品更新升级，实现真正的人与 AI 的互动，优化游戏体验；对电竞行业进行数字化创新；营销买量制作周期缩短，缩减营销成本；
- 2) **广告营销方面**，生成式 AI 内容生产能力契合广告营销需求，从获取广告创意、进行内容生成、后期进行广告投放，最后实现高效的消费转化，有望显著提升营销内容生产效率，同时降低优质营销内容生产门槛；
- 3) **影视方面**，生成式 AI 从剧本制作、影视制片、导演拍摄、后期制作以及宣发营销五大环节对整体影视制作提供助力，大幅缩短创意落地实现的时间成本，也降低了影片创作的人力成本，显著提高工作效率。

风险因素：政策环境变化；市场竞争加剧；创新技术应用不及预期；AI 版权及知识侵权风险；市场需求的不确定性；违法违规风险。

正文目录

1 生成式 AI：快速发展赋能多领域，助力技术变革	5
1.1 发展历程：21 世纪加速人工智能领域探索，生成式 AI 四领域落地	5
1.2 政策：国家陆续出台指导性政策支持性政策，推动行业发展	5
2 细分领域：技术及产品更新迭代频率加快，未来发展可期	7
2.1 图像生成：Midjourney 稳住行业龙头地位，AI 文生图迈向多模态融合	7
2.1.1 发展历程：迈向多模态融合，逐渐成为大模型核心模态之一	7
2.1.2 市场规模：市场规模高速增长，网民参与度显著提升	8
2.1.3 产品：产品发展成熟，Midjourney 稳住行业龙头地位	9
2.2 视频生成：新品推出受到市场关注，创新点赋能行业发展	9
2.2.1 发展历程：技术更迭带动行业发展，扩散模型成为主流技术路径	10
2.2.2 市场规模：发展存在一定挑战，市场潜力可期	10
2.2.3 产品：新品推出受到市场关注，创新点赋能行业发展	10
2.3 音频生成：音乐生成类市场潜力大，企业探索视频生音频技术	12
2.3.1 产业链：协同发展，中游 AI 音频算法和服务平台为核心	12
2.3.2 市场分类：语音合成占主导，音乐生成潜力大，语音识别应用成熟	13
2.3.3 产品：Suno V3.5 实现功能再突破，Google、ElevenLabs 探索视频生	13
音频技术	13
3 传媒行业应用：AI 赋能传媒子行业，助力技术变革	15
3.1 AI 赋能游戏行业：优化游戏制作流程，推动迈入制作运营新阶段	16
3.2 AI 赋能广告营销行业：契合广告营销需求，带动行业变革	18
3.3 AI 赋能影视行业：助力影视制作阶段，缩减实现成本	19
4 投资建议	20
5 风险因素	20
图表 1：生成式人工智能发展历程	5
图表 2：生成式 AI 相关政策文件	6
图表 3：AI 文生图发展历史	8
图表 4：AI 文生图市场规模及网民参与度	8
图表 5：市面上主流 AI 文生图产品对比	9
图表 6：AI 文生视频发展历程	10
图表 7：Sora、Dream Machine 以及可灵对比示例	11
图表 8：文生视频相关产品	11
图表 9：AI 音频产业链	13
图表 10：AI 音频生成分类	13
图表 11：Suno 试用	14
图表 12：ElevenLabs Riffusion 试用	14
图表 13：AI 音频相关产品	14
图表 14：生成式 AI 应用内容示意图	16
图表 15：AI 赋能游戏行业	17
图表 16：《Generated Adventure》	17
图表 17：Smallville AI 小镇	17
图表 18：《王者荣耀》绝悟 AI	18

图表 19:	虚拟裁判“奈昔”.....	18
图表 20:	AI 赋能广告营销行业	18
图表 21:	淘宝全 AI 制作短片	19
图表 22:	可口可乐‘实拍+3D+Stable Diffusion’制作广告	19
图表 23:	AI 赋能电影、影视行业	19
图表 24:	芒果 TV AI 角色对话及 AI 导演爱芒.....	20
图表 25:	央视 AI 制作微短剧	20

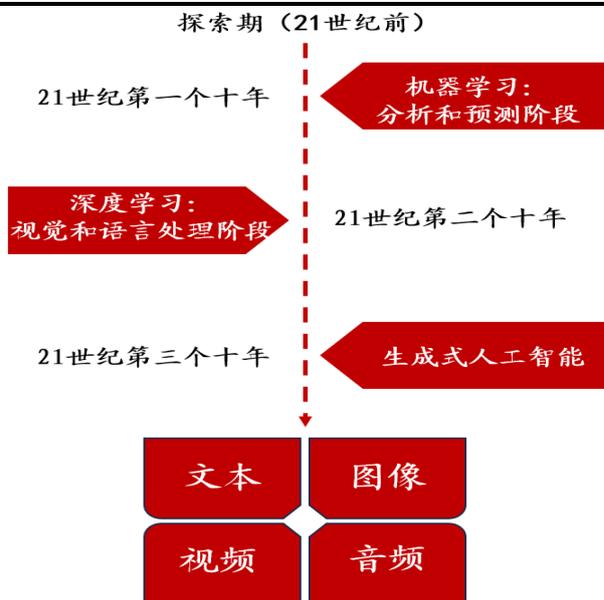
1 生成式 AI：快速发展赋能多领域，助力技术变革

生成式人工智能 (Generative AI)，指的是通过人工智能技术自动生成内容的生产方式。通过训练模型来生成新的、与训练数据相似的内容。与传统类型的 AI 主要关注识别和预测现有数据的模式不同，生成式 AI 着重于创造新的、有创意的数据，其关键原理在于学习和理解数据的分布，进而生成具有相似特征的新数据，在图像、文本、音频、视频等多种领域都有广泛的应用。

1.1 发展历程：21 世纪加速人工智能领域探索，生成式 AI 四领域落地

21 世纪加速人工智能领域探索，生成式 AI 四领域落地。在 21 世纪之前，人工智能处于缓慢探索期，该概念最早可以追溯到 20 世纪 40 年代，在 1943 年沃伦和沃尔特提出了人工神经网络的概念以及艾伦·图灵发表了著名的论文《计算机器与智能》，提出了图灵测试，用以判断机器是否具有智能，从而开启对于人工智能的早期探索，但限于计算机的计算能力和数据处理能力有限，以及人工智能研究过于注重理论而忽视了实际应用，发展速度有所减缓；直到进入 20 世纪 90 年代，互联网的兴起为人工智能的发展带来了新的机遇，在 21 世纪第一个十年间，搜索引擎、推荐系统和机器翻译等应用的出现以及支持向量机、随机森林等算法的提出，为人工智能应用提供了更加强大的工具，进入机器学习阶段；2010 年以来，随着大数据的出现和计算能力的提升，深度学习成为人工智能研究的热点。深度神经网络在图像识别、语音识别和自然语言处理等领域取得了令人瞩目的成绩；在经历了前二十年的技术积累后，生成式 AI 在 21 世纪第三个十年诞生，由 OpenAI 开发的 GPT-4 语言模型，标志着基于语言的人工智能应用程序迈入了崭新的功能阶段，并在图像生成、文本生成、音视频生成上取得显著成果。

图表1：生成式人工智能发展历程



资料来源：前瞻产业研究院、万联证券研究所

1.2 政策：国家陆续出台指导性及支持性政策，推动行业发展

国家陆续出台指导性及支持性政策推动生成式 AI 发展。从政策发展的角度来看，中国的生成式 AI 是在国家层面的人工智能政策推动下得到发展的。“十三五”规划首次将人工智能作为重点任务，并明确了新一代人工智能的发展战略目标，这标志着人工智能上升到了国家战略的高度。在“十四五”规划中，进一步强调了在人工智能关键技

术领域实现突破的重要性。为了确保生成式AI的健康发展和规范应用，同时保护国家安全和公共利益，2023年7月，国家互联网信息办公室、国家发展和改革委员会、教育部、科学技术部、工业和信息化部等部门联合发布了《生成式人工智能服务管理暂行办法》。此后，各相关部门陆续出台了指导性及支持性的政策文件，以加速生成式AI的发展，并鼓励其在各个领域的应用。

图表2: 生成式 AI 相关政策文件

发布时间	发布部门	政策名称	重点内容解读	政策性质
2024年2月	中国科学院	《关于推动未来产业创新发展的实施意见》	推动下一代移动通信、卫星互联网、量子信息等技术创新突破，加速类脑智能、群体智能、大模型等深度赋能加速培育智能产业。加快突破GPU芯片、集群低时延互连网络、异构资源管理等技术，建设超大规模智算中心，满足大模型迭代训练和应用推理需求。	指导类
2023年12月	自然资源部	《关于部署开展国土空间规划实施监测网络建设试点的通知》	以建设服务数字生态文明的数字生态基础设施为使命，以生成式人工智能等先进技术在国土空间规划领域的应用研发为突破口，推进相关算法重构模型重构、标准重构和感知系统重构，着力提升国土空间规划实施监测网络“智慧”能力。	支持类
2023年12月	工信部	《关于加快传统制造业转型升级的指导意见》	推动产业园区和集群整体改造升级。探索建设区域人工智能数据处理中心，提供海量数据处理、生成式人工智能工具开发等服务，促进人工智能赋能传统制造业。探索平台化、网络化等组织形式，发展跨物理边界虚拟园区和集群，构建虚实结合的产业数字化新生态。	支持类
2023年9月	国家广播电视总局	《关于开展广播电视和网络视听虚拟现实制作技术应用示范有关工作的通知》	研究基于人工智能方式的虚拟场景生产技术，开展基于人工智能方式的剧本创作、故事板生成、三维数字资产建模、智能语音生成、短视频生成、动作驱动等场景应用，提升虚拟场景生产效率、降低虚拟场景生产成本。	支持类
2023年7月	国家互联网信息办公室等	《生成式人工智能服务管理暂行办法》	鼓励生成式人工智能技术在各行业、各领域的创新应用，生成积极健康、向上向善的优质内容，探索优化应用场景，构建应用生态体系。支持行业组织、企业、教育和科研机构、公共文化机构、有关专业机构等在生成式人工智能技术创新、数据资源建设、转化应用、风险防范等方面开展协作。	支持类
2023年4月	工信部等	《关于推进IPV6技术演进和应用创新发展的实施意见》	推动IPV6与5G、人工智能、云计算等技术的融合创新，支持企业加快应用感知网络、新型IPV6测量等"IPV6+"创新技术在各类网络环境和业务场景中的应用。	指导类

资料来源：前瞻产业研究院、万联证券研究所

2 细分领域：技术及产品更新迭代频率加快，未来发展可期

从传媒领域的覆盖情况来看，更注重内容端成果落地，因此我们着重探讨图像生成、视频生成及音频生成发展情况及相关产品运营情况。

2.1 图像生成：Midjourney 稳住行业龙头地位，AI 文生图迈向多模态融合

AI文生图，即通过人工智能技术解析文字描述内容，提取关键信息，利用深度学习模型形成草图，再经过优化算法细化图像细节，增强图像质量，输出符合描述的视觉内容，是近年来快速发展的一项技术。

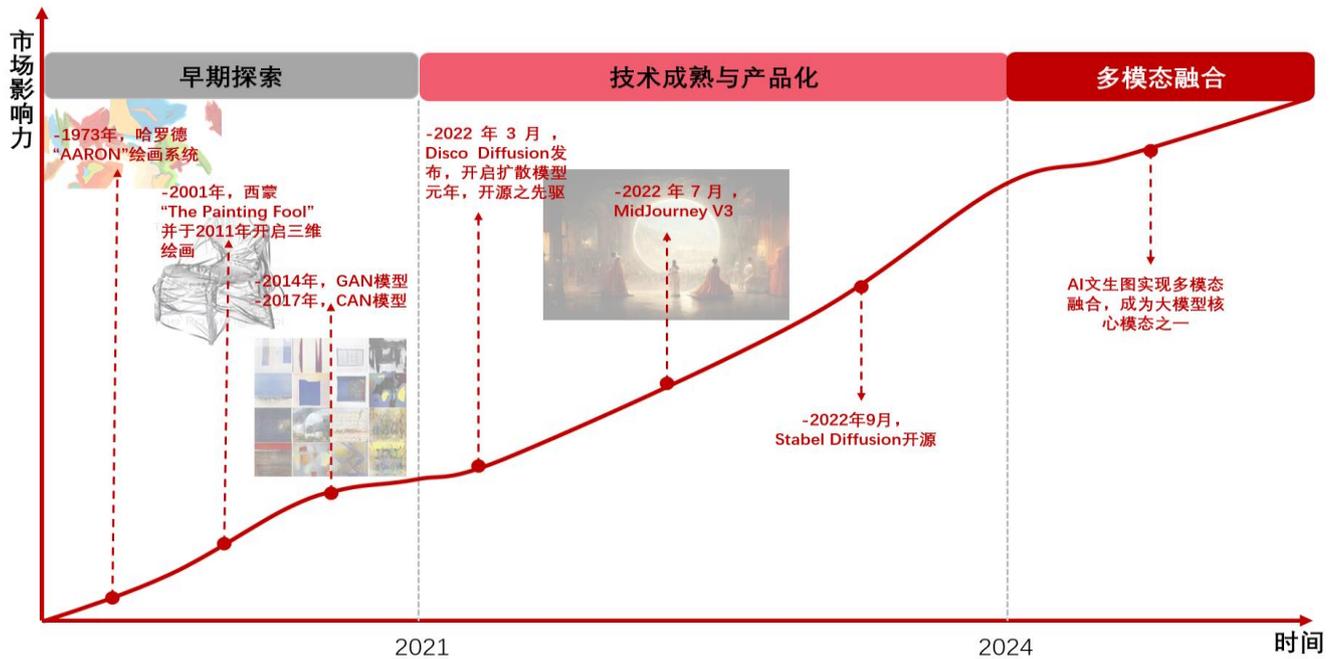
2.1.1 发展历程：迈向多模态融合，逐渐成为大模型核心模态之一

早期探索阶段：以1973年哈罗德创造智能AI绘画系统“AARON”为起点，系统的绘画方式是建立在哈罗德本人对绘画的理解之上，将其绘画风格与技术用层序编码形式展现出来，是最初代的AI绘画系统；2001年西蒙开始研发智能图形软件，创造“The Painting Fool”，可以根据照片里的色块分布，使用现实中的绘画材料进行创作，并在2011年开发3D建模能力，开创三维绘画时代；在2014年，深度学习模型对抗生成网络GAN(Generative Adversarial Network)提出，奠定了早期AI绘画实现技术的基础，在2017年罗格斯大学实验室基于GAN制作出创造性对抗网络CAN(Creative Adversarial Networks)，带动行业更进一步发展。

技术成熟及产品化阶段：2022年初，由独立开发者Somnai开发的Disco Diffusion在谷歌Colab云服务上正式对世界开放使用，标志着行业迈向技术成熟及产品化阶段；2022年7月，MidJourney V3上线，也是目前主流的AI文生图产品之一，且9月份AI文生图中核心技术Stable Diffusion以开源底层代码的形式面向大众，促使市场相关应用程序大量开发涌现，使AI文生图的艺术质量呈现出指数级的进化速度。

多模态融合阶段：自2023年起，它已经在摄影艺术、设计、二次元、CG艺术和影视制作等多个行业得到广泛应用。AI文生图技术不仅限于单一的图像生成模型，而是开始向多模态融合方向发展，成为大型AI模型如GPT4-V和Dalle-3的核心组成部分。这些模型能够处理包括文本、图像、音频和视频在内的多种数据类型，实现跨媒体的理解和生成。随着技术的持续进步和应用场景的扩展，预计AI文生图技术将在未来获得更广泛的应用和更深入的集成。

图表3: AI 文生图发展历史

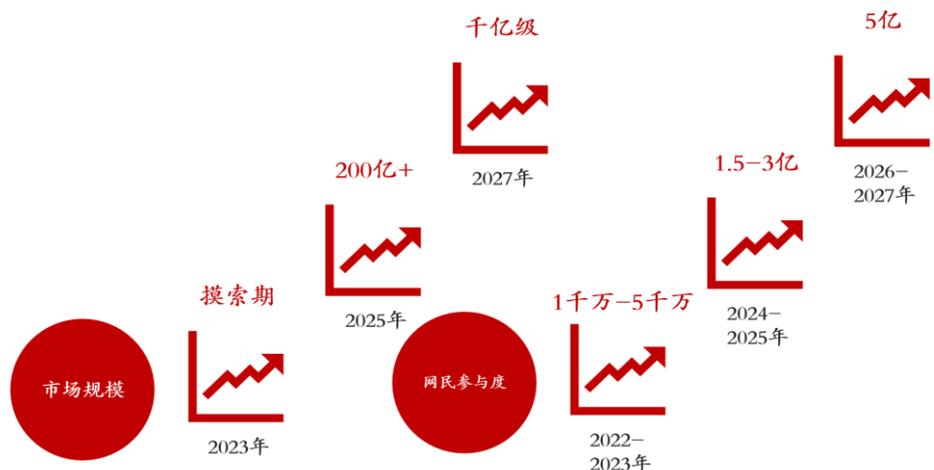


资料来源: 朗艺美术馆、《中国AIGC文生图产业白皮书》、元宇宙新声、万联证券研究所

2.1.2 市场规模: 市场规模高速增长, 网民参与度显著提升

市场规模高速增长, 网民参与度显著提升。 市场规模方面, 根据《中国AIGC文生图产业白皮书》数据显示, 2023年, 中国AI文生图市场多数企业处于融资初期或A轮阶段。预计到2024年, 行业将出现有影响力的企业, 推动市场市值显著增长, 但规模仍在百亿人民币内。2025年, 市场将进入商业模式成熟期, 预计至少3家企业成为独角兽, 市场规模有望突破200亿。2026年起, 市场将进入快速增长期, 预计2027年接近千亿规模。**网民参与度方面**, 2022年约有1000万网民参与AI文生图, 到2023年底, 这一数字将增至5000万。随着商业模式的成熟, 2024至2025年, 中国AI文生图网民有望达到1.5至3亿, 其中许多人将深度参与市场建设。如果2026至2027年市场爆发, 网民数量可能达到5亿。"参与网民"包括所有接触、了解、使用AI绘画功能、平台、工具或产品的用户。随着市场的发展, 用户参与度将显著提高, 推动AI文生图行业的快速成长。

图表4: AI 文生图市场规模及网民参与度



资料来源: 《中国AIGC文生图产业白皮书》、万联证券研究所

2.1.3 产品：产品发展成熟，Midjourney 稳住行业龙头地位

产品发展成熟，Midjourney 稳住行业龙头地位。目前市场上的产品已经发展比较成熟，主流的AI文生图工具分别有Midjourney V6、Adobe Firefly 3、Stable Diffusion 3、Dalle 3等。从数字生命卡兹克的数据来看，分别从细节质量、审美及语义理解三个方面进行产品测评，Midjourney V6均占据较为显著的优势，主要原因在于产品迭代速度快于同行，整体图像产出质量显著提高，同时通过庞大的用户量及用户数据，使得Midjourney积累的数据集具有独家性，并依托这些数据进行反复训练，来构建自我艺术风格，巩固自身竞争优势，稳住文生图龙头地位。

图表5： 市面上主流 AI 文生图产品对比



资料来源：数字生命卡兹克、万联证券研究所

注：左上角Midjourney V6、右上角Adobe Firefly 3、左下角Stable Diffusion 3、右下角Dalle 3。

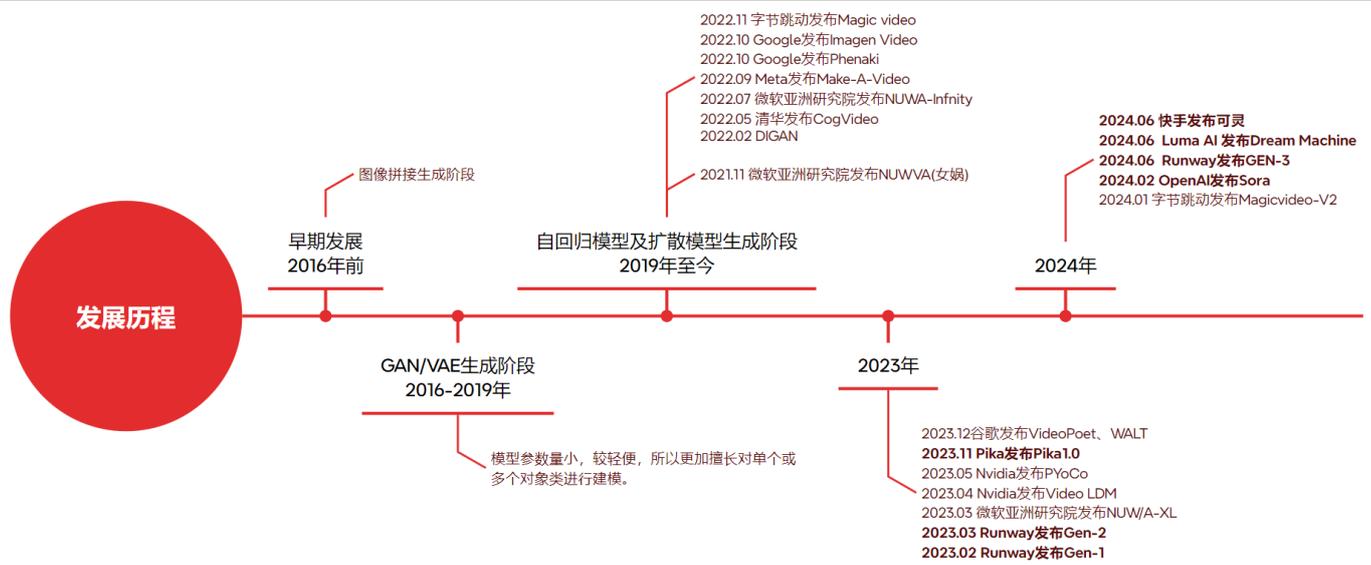
2.2 视频生成：新品推出受到市场关注，创新点赋能行业发展

AI文生视频的本质是能够根据输入的文本描述自动生成相应的视频内容。这项技术结合了自然语言处理（NLP）、计算机视觉、图像生成和动画技术等多个领域的研究成果，通过训练大量文本和视频数据，依靠文本解析、图像生成、动态渲染等核心技术，将文字转化为图像，并赋予这些图像动态效果，全自动化形成视频文件。文生视频技术的应用可以大幅降低内容创作的门槛，使得短视频、影视镜头、广告等内容的制作更加便捷。

2.2.1发展历程：技术更迭带动行业发展，扩散模型成为主流技术路径

技术更迭带动行业发展，扩散模型成为主流技术路径。在发展早期阶段，主要依靠图像进行拼接形成视频；在2016年GAN及VAE开始兴起，给文生视频的发展奠定了基础，这两者均是生成模型，原理是学习数据的概率分布并生成类似于训练数据的新数据，而GAN能够生成更高质量的数据，但也伴随着稳定性较差、生成图像缺乏多样性问题显现，导致应用范围有限，行业发展速度缓慢；2019年自回归模型推出，相较于GAN，具有明确的密度建模和稳定的训练优势，而行业真正兴起点在于将扩散模型在文生图领域经验拓展至文生视频，成为AI视频生成领域的主流技术路径，带动行业产品推出频率及数量显著提升，例如Runway的GEN-3、Pika1.0、Open AI的Sora、快手的可灵等。

图表6: AI 文生视频发展历程



资料来源：Z计划、SpaceTools、Runway、快手、万联证券研究所

2.2.2市场规模：发展存在一定挑战，市场潜力可期

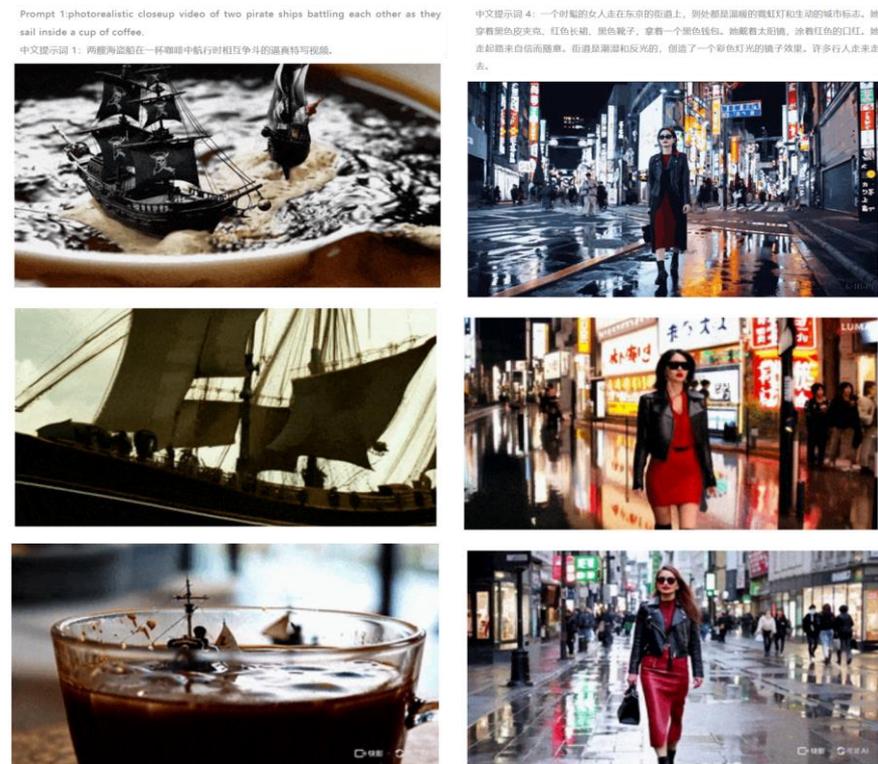
相较于AI文生图来说，文生视频发展较为缓慢，目前市场仍处于初期阶段，主要挑战在于：1) 缺乏相符性高的训练数据，视频与文本相符性高的大规模数据集较少，获取成本及定点标注成本较为高昂，模型训练将会受到阻碍；2) 计算成本高昂，视频制作时间长短决定了帧数多少，为确保帧间空间和时间逻辑保持一致性，从而需要大量的计算资源，训练模型的成本高昂；3) 信息冗杂难处理，视频具有多场景、多任务、视觉动态等特性，信息冗杂，处理难度较高。根据GIR (GlobalInfo Research) 调研，2023年全球文生视频大模型收入大约为720万美元，预计2030年达到22.19亿美元，2024至2030期间，年复合增长率CAGR有望达到56.6%，市场潜力可期。

2.2.3产品：新品推出受到市场关注，创新点赋能行业发展

新品推出受到市场关注，创新点赋能行业发展。目前市场上受到关注度较高的文生视频产品分别是来自OpenAI的Sora、Pika1.0、近期快手推出的可灵、Luma AI推出的Dream Machine以及Runway推出的GEN-3，从创新点来看，快手的可灵可为已生成的视频（含文生视频及图生视频）提供便捷的一键续写和连续多次续写，将视频最长可延伸至约3分钟，在确保新生成部分与原视频间的运动连贯性及物理合理性的同时，还能巧妙融入大幅度的动作变化，提升视频的生动性；Luma AI的Dream Machine则是支持物理模拟，生成具有真实物理特性的视频，如重力下落、碰撞和光影变化等，从而

确保视频在视觉和物理行为上都具有真实性和连贯性等等。新品的推出通常伴随着领域技术的突破，从目前的文生视频能力来看，已经具备了一定的产业生产力，未来将赋能影视产业链，为其降本增效提供助力。

图表7: Sora、Dream Machine 以及可灵对比示例



资料来源：机器之心、万联证券研究所

注：从上到下依次是Sora、Dream Machine以及可灵。

图表8: 文生视频相关产品

模型名称	研发机构	视频长度	竞争优势	局限性
Sora	OpenAI	60 秒	<ul style="list-style-type: none"> -在生成视频的时长上出现了飞跃式的进步，从原有的单次 2-4s 左右提升至最大 60s，使 AI 生成视频真正具有了一定的生产力价值； 	<ul style="list-style-type: none"> -物理原理和时间推移的模拟方面还不完善； -一些物理过程仍不能完美模拟（如玻璃破碎）； -有时不能准确模拟出视频中人和物的交互； -较长时间的视频中“穿帮”镜头也偶有出现。
GEN-3	Runway	10 秒	<ul style="list-style-type: none"> -通过文本、图像、视频混合的训练模式，提高了生成视频的质量和一致性，同时还降低了训练消耗； -Runway 提出了延时扩散模型，在预训练的扩散模型中引入时序层，使得模型在推理阶段具有更高的时间一致性； -相比于 GEN-2，时长升级至 10 秒， 	<ul style="list-style-type: none"> -操作界面相对复杂，对于初学者可能有一定的学习曲线； -对于复杂语句的理解可能出现偏差。

			艺术性审美更强，光线及光影把控更加细节。	
Pika1.0	PIKA LABS	3-7 秒	-生成的视频在逻辑连贯性、流畅性、画质和风格准确性上有明显优势； -Pika labs 1.0 能够支持对于视频实时编辑和修改。	-少数视缺少真实感。
可灵	快手	5 秒	-可为已生成的视频（含文生视频及图生视频）提供便捷的一键续写和连续多次续写，将视频最长可延伸至约3分钟； -融入了提示词融合技术，深化了对图像语义与用户指令的理解整合。这意味着，模型能够根据用户提供的不同文本指令，精妙地变换视频中的动态表现。	-人物运动过大时，会出现脸部和身体的失真； -细节表现不太完美。
Dream Machine	Luma AI	5 秒	-支持物理模拟，生成具有真实物理特性的视频，如重力下落、碰撞和光影变化等； -在运动流畅度和电影及戏剧效果上表现出色。	-可控性有限：在演员表现、摄影机调度和画面控制方面的可控性有限，尤其是在处理复杂情节和多人物交互场景时； -风格局限性：Luma AI 在某些特定风格（如国风、玄幻风格）的表现上不尽人意，且稳定性有待加强。

资料来源：甲子光年、乐晴智库、SpaceTools、AiBot机器人对话、去玩AI、万联证券研究所

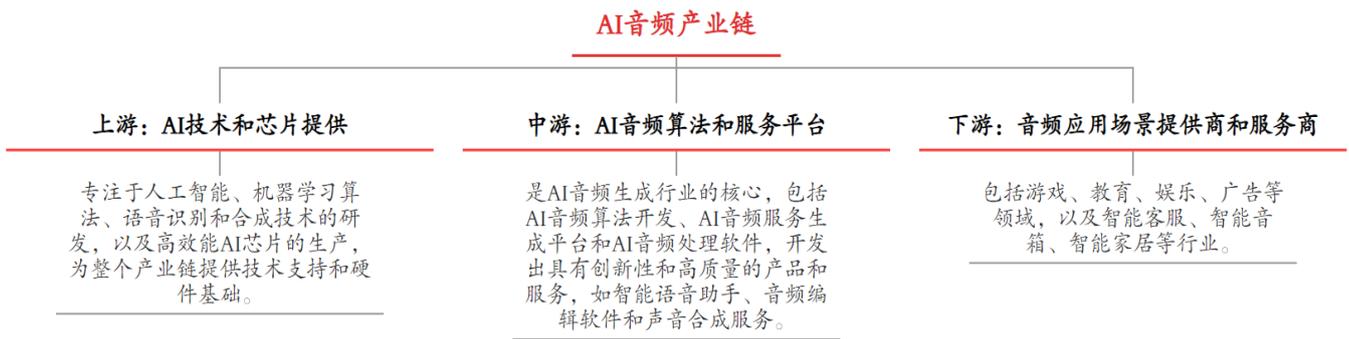
2.3 音频生成：音乐生成类市场潜力大，企业探索视频生音频技术

AI音频生成行业，作为生成式AI的关键应用领域，正快速成为技术创新的领先领域。该行业利用先进的人工智能技术和复杂算法来创造音频内容，包括语音合成、音乐制作和声音效果合成等。通过融合机器学习和深度学习算法，AI音频生成技术能够精确地模仿人类的语音、音乐节奏和声音效果，生成自然且逼真的音频，广泛应用于娱乐、广告、教育和新闻传播等多个领域。

2.3.1 产业链：协同发展，中游AI音频算法和服务平台为核心

产业链协同发展，中游AI音频算法和服务平台为核心。在AI音频生成行业的产业链中，上游主要由AI技术和芯片提供商组成，他们负责AI技术的研究与开发以及芯片的生产，为整个产业链提供必要的技术支持和硬件基础。中游则是AI音频生成行业的核心，包括AI音频算法的开发、AI音频服务生成的平台以及AI音频处理软件，这些环节对技术的要求极高，是推动行业发展的关键。下游则是具体的应用领域，涉及游戏、教育、娱乐、广告等多个场景，以及智能客服、智能音箱、智能家居等产品，这些提供商和服务商将AI音频技术应用于实际场景中，满足不同用户的需求。

图表9: AI 音频产业链



资料来源: 千际投行, 万联证券研究所

2.3.2 市场分类: 语音合成占主导, 音乐生成潜力大, 语音识别应用成熟

语音合成占主导, 音乐生成潜力大, 语音识别应用成熟。AI音频生成技术根据不同的应用场景主要分为三个类别: 语音合成、音乐生成和语音识别。**语音识别**在早期的智能音箱和语音助手等产品中已经得到了广泛的应用,市场较为成熟;**语音合成**则是基于深度学习算法,能够精确地模拟人类的语音特征,包括音色、音调和语调,将文本信息转化为口语化的语音输出,在朗读软件和语音导航等领域得到了广泛应用,并且目前占据了整个AI音频市场的近70%份额;**音乐生成**是随着AI技术不断深入发展而兴起的一个重要应用领域,它利用机器学习和深度学习算法来模拟人类的音乐创作过程,具有在音乐创作、游戏音效制作和电影配乐等方面的巨大潜力。随着AI技术的不断进步和创新,音乐生成领域的应用前景非常广阔,有望为音乐产业带来革命性的变化。

图表10: AI 音频生成分类



资料来源: 千际投行, 万联证券研究所

2.3.3 产品: Suno V3.5 实现功能再突破, Google、ElevenLabs 探索视频生音频技术

Suno V3.5实现功能再突破, Google、ElevenLabs探索视频生音频技术。从具体的产品来看,标杆性的音乐生成类音频AI是由Suno AI研发的Suno系列产品,于今年5月推出了V3.5版本,维持了与**AI**一次性完成歌词、演唱、编曲、配乐等全部流程,领先于市面上单纯生成音乐的同类型产品,相较于V3版本,主要升级在最长片段长度扩展至4分钟;最长2分钟的歌曲拓展;歌曲结构的显著改进。这些均表示AI音频已升级至可以更轻松地制作出完整的歌曲,提供连贯且高质量的音乐输出,并使生

成的音乐更加流畅和自然，也使得音乐的情感表达更加丰富和细腻，同时关注到 ElevenLabs及Google DeepMind均宣布了自动匹配画面AI生成音效产品，即无需人工输入提示词也可以为视频配音，实现真正的AI音频及AI视频合作，后者Google的产品也可以根据提示词控制输出所需的音频或避免出现不需要的声音，自由度更高。随着相关产品功能的完善，这将带领文生视频迈入有声时代，实现行业再突破，双向促进市场繁荣发展。

图表11: Suno 试用



资料来源: SunoV3、万联证券研究所

图表12: ElevenLabs Riffusion 试用



资料来源: ElevenLabs、万联证券研究所

图表13: AI 音频相关产品

模型名称	研发机构	分类	介绍
SunoV3	Suno AI	音乐生成	其能够根据用户的简单提示，基本的创作需求，如音乐风格、情感氛围等，一次性完成歌词、演唱、编曲、配乐等全部流程，得益于与 Open AI 合作，Suno 模型负责生成音乐，而 ChatGPT 则负责生成歌词及标题。相较于 V3 版本，最长片段长度扩展至 4 分钟，增加了对 2 分钟歌曲的拓展功能，在歌曲结构的处理上进行了重大改进，使得生成的音乐更加流畅和自然。
Stable Audio 2.0	Stability AI	音乐生成	通过文本至音频和音频至音频的双重提示功能，用户能够制作旋律、伴奏轨、音轨分离和音效，从而提升创作流程。与其他顶尖模型相比，Stable Audio 2.0 的独特之处在于其能够生成最长达三分钟的完整歌曲，包括具有引入、发展和尾声的结构化作曲，以及立体声音效。
Riffusion (试用)	ElevenLabs	音乐生成 (匹配视频)	利用人工智能技术，可以为视频添加逼真的音效，从而增强视频的沉浸感。Riffusion 基于扩散模型，通过分析视频中的视觉元素，如物体运动和场景变化，自动生成与之匹配的音效。用户只需输入一段视频，Riffusion 就能识别视频内容，并生成同步的音效，如脚步声、开门声、雨声等。
V2A (Video-to-Audio) (未公开发布)	Google DeepMind	音乐生成 (匹配视频)	能根据画面内容或者手动输入的提示词直接为视频配音，可以为任何视频输入生成无限数量的音轨。导入视频自动音频之后，也能够根据提示词生成所需的音频，使用者可以通过输入“正面提示词”来引导模型输出所需的声音，或输入“负面提示词”来引导其避免出现不需要的声音，无需人工对齐音频视频，而是可

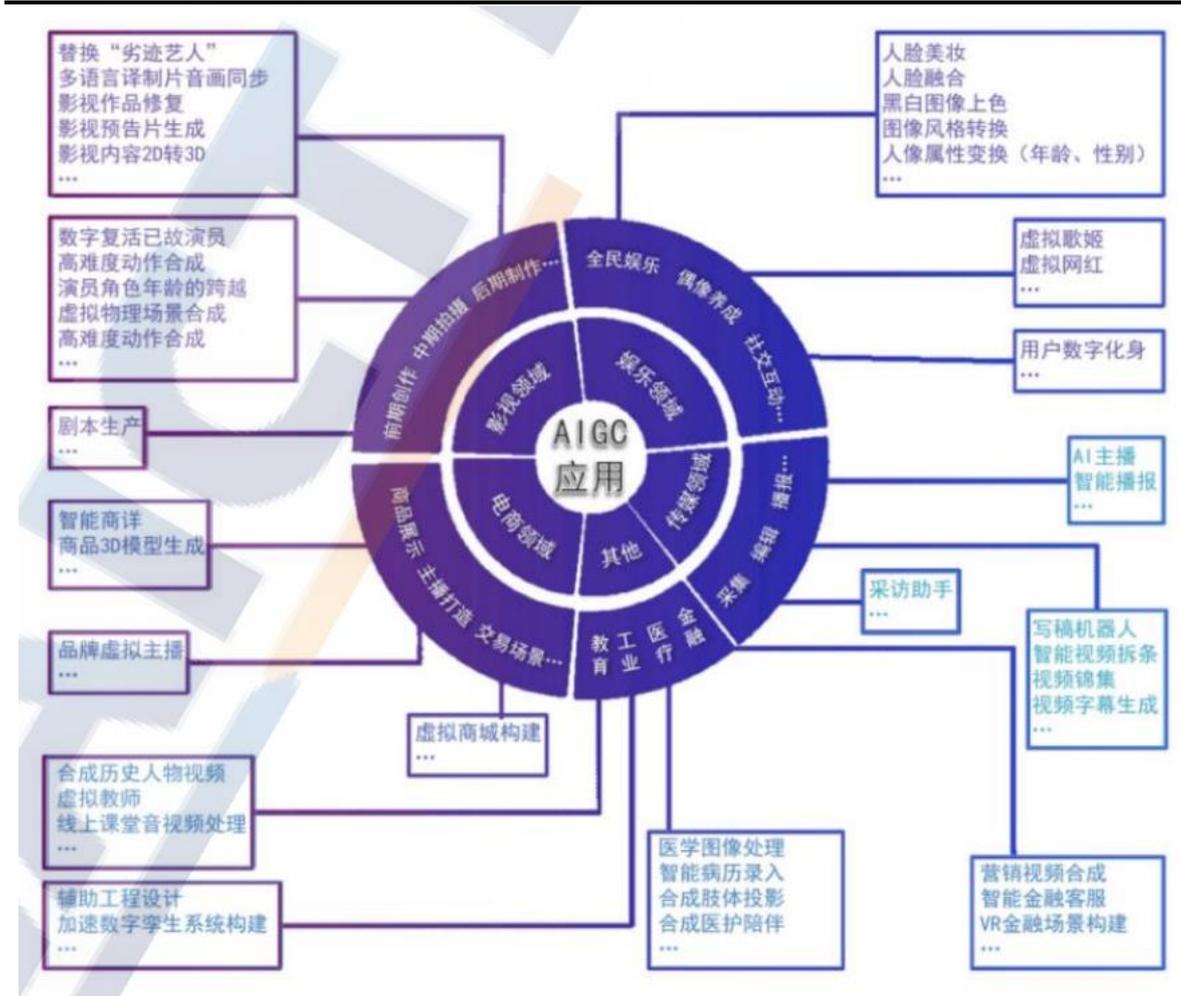
			以直接自动将音频与画面对齐。
Tacotron2	Google	语音合成	由 Google 研发的神经网络架构，主要用于语音合成。它是 Tacotron 的升级版，它结合了上一代 Tacotron 和 WaveNet（一种时域波形生成模型，可模拟任何人类声音的语音），不需要输入复杂的语法和语音特征，用经过训练的神经网络模型，即可根据文本生成类似人类的语音。
Resemble AI	Resemble AI Inc	语音合成	可以生成人类声音的人工智能工具。专注于生成逼真的、类似人类的语音。它可以从你提供的文本到语音、语音到语音、神经音频编辑、语言配音、情感、实时语音克隆、能够从少量的语音样本中捕捉到个体声音的独特之处，并创建出高质量的人声克隆效果。
So-Vits-Svc	rcell	语音合成	是一个开源的人工智能歌声合成器项目。能模拟真人歌手嗓音，能够将一个人的歌声转换为另一个人的风格，同时保留歌曲的旋律和节奏。通常被用于音乐制作、歌声合成和虚拟偶像的开发等领域，例如 AI 孙燕姿。
SkyMusic	昆仑万维	语音合成、音乐生成	基于昆仑万维强大的「天工 3.0」超级大模型，采用先进的 Sora 模型架构，特别优化了视频、音频和音乐的处理能力。这一技术突破，使得「天工 SkyMusic」能够生成长达 80 秒、采样率高达 44100Hz 的高品质音乐，同时在人声合成技术上达到了行业领先的 SOTA（技术前沿）标准。
讯飞听见	科大讯飞	语音识别	是科大讯飞推出的一款智能语音识别与转写软件，它能够将语音内容实时转化为文字，并支持多种语言识别。这款工具不仅广泛应用于各种会议和讲座场景，更在教学领域展现出巨大的潜力。

资料来源: SunoV3、Renee、智东西、AI电气行、AI智谷、科大讯飞、Aihomie、万联证券研究所

3 传媒行业应用：AI 赋能传媒子行业，助力技术变革

AIGC赋能传媒行业，助力技术变革。以生成式AI为核心理念的AIGC（即AI Generated Content）是指利用人工智能技术来生成内容，也被认为是继UGC（用户生成内容）、PGC（专业生产内容）之后的新型内容生产方式，囊括了人工智能生成的内容、生成内容所用的算法以及厂家提供的产品或服务。目前，AIGC已在广告、影视、游戏、娱乐等多个传媒互联网的细分领域进行广泛应用，从而提升内容生产效率，降低内容生产成本，助力行业变革。

图表14: 生成式AI应用内容示意图

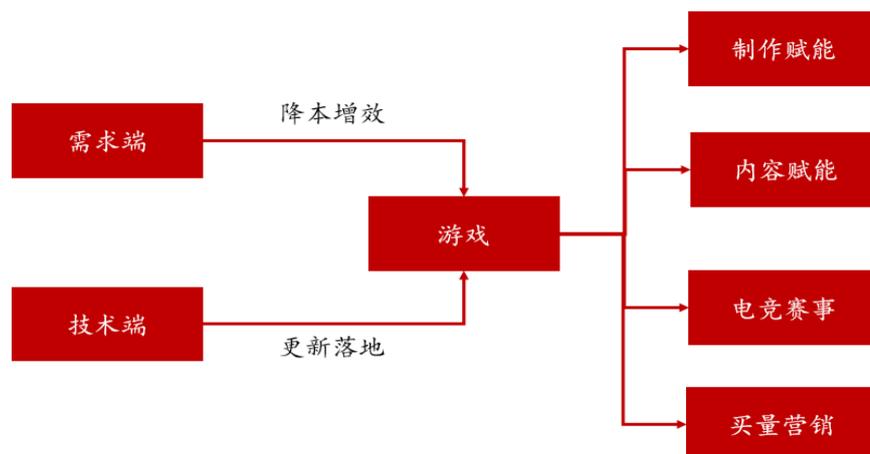


资料来源: 中国信通院, 万联证券研究所

3.1 AI 赋能游戏行业: 优化游戏制作流程, 推动迈入制作运营新阶段

需求端、技术端双重促进生成式AI赋能游戏领域。随着科技不断进步, AI已经逐步从弱人工智能向强人工智能迈进, 数字内容生产方式也愈发高效, 而游戏领域涉及的细分内容复杂且多, 开发时间较长, 开发商难以同时满足成本、质量、效率, 为保证质量, 需投入大量人员或延长制作周期, 开发成本和效率难以控制, 生成式AI的投入使用, 能够优化游戏制作全流程, 提高开发人员制作效率, 降低游戏研发成本, 并赋能产品内容创新, 并助力游戏行业运营、发行等领域实现变革, 全方位推动行业迈入制作运营新阶段。

图表15: AI 赋能游戏行业



资料来源: 万联证券研究所

生成式AI助力游戏制作、内容创新,降本增效,优化游戏体验。从游戏开发方面来说,主要分为游戏制作及内容创新。从游戏制作来看,生成式AI能够根据文本生成语音、自助或协助编写代码、根据主题设计生成游戏地图、根据二维图像生成三维模型等,极大提升游戏的策划、音频、美术、程序等环节的生产力,压缩游戏整体项目的研发周期与人员规模,大幅降低游戏制作成本,例如《Generated Adventure》是一款点击冒险类游戏,其游戏的美术、UI、音乐等均由Midjourney、ChatGPT等AI工具所打造,耗时仅72小时,虽然游玩体验较为粗糙,但仍是一次全新的技术尝试,未来随着技术不断发展进步,AI全线制作游戏的可行性将大幅提升;从内容创新来看,对已有的游戏进行产品更新升级,实现真正的人与AI的互动,例如在游戏《逆水寒》中引入ChatGPT,打破固定交流模式,实现高自由度地游戏互动体验;斯坦福大学构造了一个虚拟小镇,将25个AI智能体投入到沙盘环境中,利用GPT来生成其行为和语言,在设定的身份及性格框架下,赋予AI记忆和情感,会影响他们日后的所有决策,具有高度自由化,如同真实的现实小镇。

图表16: 《Generated Adventure》



资料来源:《Generated Adventure》官网、万联证券研究所

图表17: Smallville AI 小镇



原:《Generative Agents: Interactive Simulacra of Human Behavior》, 万联证券研究所

电竞行业进行数字化创新，营销买量制作周期缩短，实现降本增效。在游戏运营方面，电竞行业在赛事分析、电竞解说、虚拟对手训练等方面进行了技术升级，生成式AI可以广泛应用于赛前阵容分析、赛后数据分析、赛事回放、虚拟裁判以及通过AI数据搜集计算，生成虚拟对手进行实战演练，增强选手赛前训练以及赛后复盘效率。在游戏发行方面，随着生成式AI投入使用，近年来游戏买量素材生命周期变短、素材创意不足的问题得到有效解决，同时生成式AI能够分析数据，进行决策选择最优投放渠道及精准投放目标群体，提升投放效率及产出比，缩减游戏厂商买量营销成本。

图表18: 《王者荣耀》绝悟 AI



资料来源:《王者荣耀》、IT之家、万联证券研究所

图表19: 虚拟裁判“奈昔”

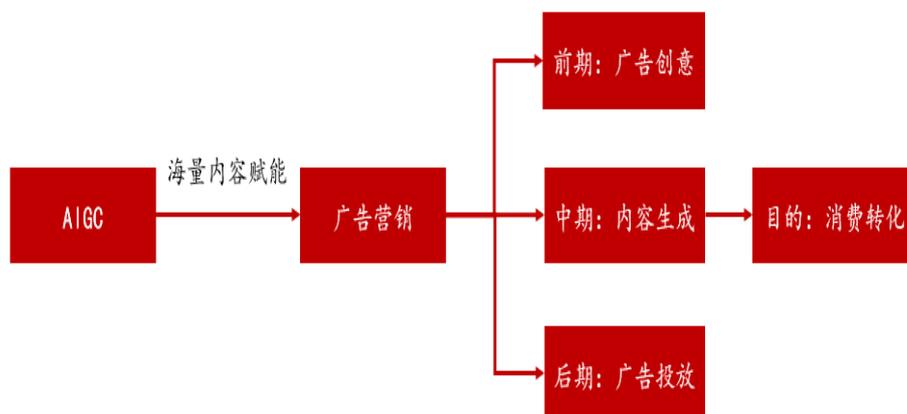


资料来源: 网易电竞NeXT、万联证券研究所

3.2 AI 赋能广告营销行业：契合广告营销需求，带动行业变革

生成式AI内容生产能力契合广告营销需求。生成式AI通过其强大的内容生成能力赋能广告营销各个阶段，从获取广告创意、进行内容生成、后期进行广告投放，最后实现高效的消费转化，有望显著提升营销内容生产效率，同时降低优质营销内容生产门槛，丰富广告服务功能，长期来看带动广告营销实现行业变革。

图表20: AI 赋能广告营销行业



资料来源: 万联证券研究所

前期广告创意阶段，通过ChatGPT等AI工具协助完成市场调研，帮助快速搜集资料并高效整理，并通过AI强大的计算能力分析市场主要关注内容，提供创意方向，深度挖

掘创意潜力，激发新的灵感来源，减轻内容创作压力，实现快速完成前期准备工作；**中期内容生成阶段**，生成式AI其强大的内容生成能力能够帮助营销公司从营销文案、拍摄素材、后期配音、视频剪辑等多个方面实现降本增效；**后期广告投放阶段**，生成式AI能够通过对大量用户数据的学习，更加精确地匹配消费者偏好，针对不同细分市场、不同细分用户做更多适配性的设计，提高广告投放效果。

图表21: 淘宝全 AI 制作短片



资料来源: 淘宝、万联证券研究所

图表22: 可口可乐‘实拍+3D+Stable Diffusion’制作广告

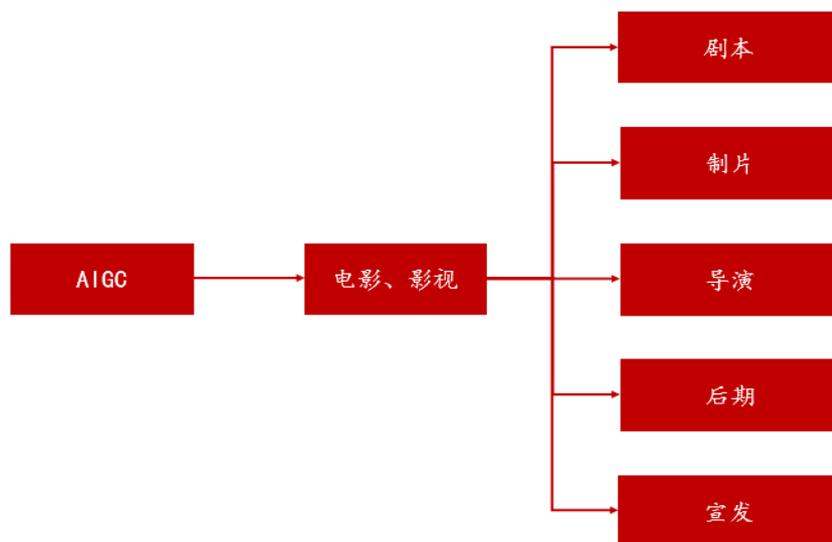


资料来源: 可口可乐、万联证券研究所

3.3 AI 赋能影视行业：助力影视制作阶段，缩减实现成本

生成式AI赋能影视制作各个阶段，大幅缩减创意落地实现成本。影视整体产业链中包括剧本制作、影视制片、导演拍摄、后期制作以及宣发营销五大环节，生成式AI将从不同角度对整体影视制作提供助力，大幅缩短创意落地实现的时间成本，也降低了影片创作的人力成本，显著提高工作效率。

图表23: AI 赋能电影、影视行业



资料来源: 万联证券研究所

剧本方面，AI能通过深度学习算法分析了所有先前的参考电影，根据预测结果进行故事情节点分析，按照预设风格生成剧本，激发编剧创作灵感，缩短创作周期，同时在剧本预演阶段，能够通过Midjourney等AI图像生成工具能快速把剧本内容具象化，以供

片方直观判断剧本影视化可能成效，减少沟通成本；**制片方面**，AI能够准确高效地辅助完成法务、制片统筹、剧组预算、日程安排等工作，让制片人将工作转移重心到沟通交流、商务洽谈、团队组建等人际关系处理中，实现个人能力最大化；**导演方面**，例如AI能从分镜绘制、设计镜头语言等方面进行辅助；**后期方面**，AI能够高效完成视频剪辑、字幕添加、颜色校正等重复性强、难度较低的任务，节省人力成本，还可通过AI进行面部识别进行数字建模，以达到替换劣迹艺人、调整演员微表情等目的，呈现更好的影视效果；**宣发方面**，AI可以通过数据搜集学习，完成不同风格的海报、文案等宣发材料制作，在影视剧集后续联动宣传，还可以通过生成式AI进行创新。

图表24: 芒果TV AI角色对话及AI导演爱芒



图表25: 央视AI制作微短剧



资料来源：芒果TV、湖南广电、万联证券研究所

资料来源：央视频、万联证券研究所

4 投资建议

21世纪加速人工智能领域探索，在国家指导性政策支持引导下，生成式AI在图像生成、文本生成、音视频生成等领域取得显著成果，AI文生图迈向多模态融合，逐渐成为大模型核心模态之一；AI文生视频新品推出受到市场关注，创新点赋能行业发展；AI音频产业链协同发展，企业探索视频生音频新技术。**行业应用**：在影视、游戏等多个传媒互联网的细分领域进行广泛应用，从而提升内容生产效率，降低内容生产成本，助力行业变革。**游戏方面**，生成式AI能压缩游戏整体项目的研发周期与人员规模，大幅降低游戏制作成本，降本增效；对已有的游戏进行产品更新升级，实现真正的人与AI的互动，优化游戏体验；对电竞行业进行数字化创新；营销买量制作周期缩短，缩减营销成本；**广告营销方面**，生成式AI生产能力契合广告营销需求，从获取广告创意、进行内容生成、后期进行广告投放，最后实现高效的消费转化，有望显著提升营销内容生产效率，同时降低优质营销内容生产门槛；**影视方面**，生成式AI从剧本制作、影视制片、导演拍摄、后期制作以及宣发营销五大环节对整体影视制作提供助力，大幅缩短创意落地实现的时间成本，也降低了影片创作的人力成本，显著提高工作效率。**建议关注已有生成式AI应用落地的公司。**

5 风险因素

政策环境变化；市场竞争加剧；创新技术应用不及预期；AI版权及知识侵权风险；市场需求的不确定性；违规违法风险。

行业投资评级

强于大市：未来6个月内行业指数相对大盘涨幅10%以上；

同步大市：未来6个月内行业指数相对大盘涨幅10%至-10%之间；

弱于大市：未来6个月内行业指数相对大盘跌幅10%以上。

公司投资评级

买入：未来6个月内公司相对大盘涨幅15%以上；

增持：未来6个月内公司相对大盘涨幅5%至15%；

观望：未来6个月内公司相对大盘涨幅-5%至5%；

卖出：未来6个月内公司相对大盘跌幅5%以上。

基准指数：沪深300指数

风险提示

我们在此提醒您，不同证券研究机构采用不同的评级术语及评级标准。我们采用的是相对评级体系，表示投资的相对比重建议；投资者买入或者卖出证券的决定取决于个人的实际情况，比如当前的持仓结构以及其他需要考虑的因素。投资者应阅读整篇报告，以获取比较完整的观点与信息，不应仅仅依靠投资评级来推断结论。

证券分析师承诺

本人具有中国证券业协会授予的证券投资咨询执业资格并登记为证券分析师，以勤勉的执业态度，独立、客观地出具本报告。本报告清晰准确地反映了本人的研究观点。本人不曾因，不因，也将不会因本报告中的具体推荐意见或观点而直接或间接收到任何形式的补偿。

免责声明

万联证券股份有限公司（以下简称“本公司”）是一家覆盖证券经纪、投资银行、投资管理和证券咨询等多项业务的全国性综合类证券公司。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，本报告中的信息或所表述的意见并不构成对任何人的投资建议。本报告中的信息或所表述的意见并未考虑到个别投资者的具体投资目的、财务状况以及特定需求。客户应自主作出投资决策并自行承担投资风险。本公司不对任何人因使用本报告中的内容所导致的损失负任何责任。在法律许可情况下，本公司或其关联机构可能会持有报告中提到的公司所发行的证券头寸并进行交易，还可能为这些公司提供或争取提供投资银行、财务顾问或类似的金融服务。

市场有风险，投资需谨慎。本报告是基于本公司认为可靠且已公开的信息撰写，本公司力求但不保证这些信息的准确性及完整性，也不保证文中的观点或陈述不会发生任何变更。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。分析师任何形式的分享证券投资收益或者分担证券投资损失的书面或口头承诺均为无效。

本报告的版权仅为本公司所有，未经书面许可任何机构和个人不得以任何形式翻版、复制、刊登、发表和引用。未经我方许可而引用、刊发或转载的引起法律后果和造成我公司经济损失的概由对方承担，我公司保留追究的权利。

万联证券股份有限公司 研究所

上海浦东新区世纪大道 1528 号陆家嘴基金大厦

北京西城区平安里西大街 28 号中海国际中心

深圳福田区深南大道 2007 号金地中心

广州天河区珠江东路 11 号高德置地广场