



上海证券
SHANGHAI SECURITIES

证券研究报告
2023年4月16日
行业：传媒
增持

AIGC应用篇系列（二）：国外AIGC各模态产品 梳理

分析师：陈旻 SAC编号：S0870522020001

主要观点

我们AIGC系列专题《“大模型+小样本”快速适配下游场景，“AI+传媒”的效力取决于适配与迭代》将**“AI+传媒”的研究框架**定义为**“通用大模型”+“行业小样本”的技术架构**。“AI+传媒”在应用层表现效力优劣的关键取决于通用大模型对垂直应用的适配程度及迭代速度，更进一步理解：

1、“行业小样本”的数据集来自小模型或应用及内容： AI产业链包括上层大模型、中层小模型、下层应用及内容，包括应用及内容直接接入大模型或通过小模型接入大模型两种方式，即**“大模型+应用及内容”**或**“大模型+小模型+应用或内容”**，其中具备特定功能的AIGC软件产品我们理解为**“小模型”+“应用”**的技术范式，本身具备较高质量的AI能力，若接入匹配的多模态大模型，有望实现能力上的质变突破。

2、“行业小样本”的结合方式包括“能力调用”及“能力训练”两层：

(1)“能力调用”是指下游垂类场景直接调用通用大模型的通用能力，并基于垂类场景内产生的特性化数据不断提升调用能力在垂类场景内的适配程度。我们认为现阶段下游应用及内容主要采取此类方式接入大模型能力，此类方式可高效快速调用大模型先进能力，在时间上及成本上具备优势。

(2)“能力训练”是指下游垂类场景将通用大模型针对特性化数据集进行再训练，从而形成垂类场景专属大模型。例如彭博社利用自身丰富的金融数据源，基于开源的GPT-3框架再训练，开发出了金融专属大模型BloombergGPT。我们认为未来拥有丰富特性化数据集的下游垂类场景将主要采取此类方式，有助于构筑更强的能力壁垒。

围绕我们对“AI+传媒”的研究框架体系，我们后续分别针对“技术层”及“产品/应用层”梳理AIGC研究版图，本篇属于AIGC系列专题中“AIGC应用篇系列之二”，重点梳理国外多模态头部应用产品，本篇主要整理国外图像、音频、视频类头部AIGC产品。



主要观点

我们将国外主要AIGC头部产品按照文本、图像、音频、视频分类如下：

- ◆ 文本领域：Automated Insights（结构化写作）、Anyword、Copy.ai（数字广告文案）、Jasper ai（营销文案AI）、ChatGPT（通用类聊天机器人）、ChatBox（聊天客服机器人）、Jenni.ai（论文AI）
- ◆ 图像领域：Midjourney（2C端文生图AI）、DALL-E2、Stable Diffusion（2B端开源高质量文生图模型）
- ◆ 音频领域：Murf AI（文本转语音生成器）、AIVA（歌曲生成）
- ◆ 视频领域：Synthesia（拼凑生成视频）、Wonder Studio、Runway Gen-2（视频生成模型）

我们认为“通用大模型”+“行业小样本”是未来主流AI发展范式。我们理解AI大模型是一种新型平台模式，是移动互联网向沉浸互联网发展的中间形态。从平台的逻辑出发，AI产业链包括上层大模型、中层小模型、下层应用及内容，包括应用及内容直接接入大模型或通过小模型接入大模型两种方式。我们认为国外头部AIGC应用产品属于“小模型”+“应用”的技术范式，本身具备较高质量的AI能力，若接入匹配的多模态大模型，有望实现能力上的质变突破。一方面上层大模型持续为小模型、应用及内容赋予通用型能力，其兼容性取决于大模型的输出与小模型/应用及内容的输入在模态上是否匹配；另一方面小模型、应用及内容持续产生的大数据可不断迭代大模型的通用能力，并让大模型更为适配垂类场景。参照PC互联网、移动互联网，我们认为每一轮新型平台模式的兴起均会重新划分流量竞争格局，涉及巨大市场空间的重新分配，且我们预计本轮AI大模型有望同时变革B端及C端产业。因此我们认为在上层大模型竞争格局尚未稳定的初期，大模型为争抢更多中下层数据入口，有望大幅让利。中下层的小模型、应用及内容有望百花齐放，且最先深度结合先进大模型的应用有望收获较大增量红利。

我们认为AIGC可分为技术方（直接收益）及场景/应用方（直接&间接收益）两类，前者为AIGC相关技术直接研发及相关上下游（如算力），变现模式预计包括会员付费、广告变现、B端变现等（如技术授权/服务/运维等）；后者为AIGC技术与落地场景深度融合，场景/应用方利用AIGC技术获得直接收益或间接收益：（1）直接收益对应收入弹性，对应市场空间的增量；（2）间接收益对应利润弹性，对应利润率的提高。目前GPT-4输出模态主要为文字，我们认为现阶段传媒领域最高效的结合场景为“AI+虚拟人”，其中具备IP价值的虚拟人受益更为明确。

风险提示：宏观经济风险，地缘政治风险；技术发展不及预期；AIGC行业发展不及预期等。



目录

Content

一、图像领域

二、音频领域

三、视频领域

四、投资建议

五、风险提示

一、图像领域

1.1 创意图像生成 (2C) : Midjourney

- ◆ **图像生成领域潜力巨大，Midjourney致力于探索新的思想媒介。** 根据6pen，未来5年全球10%-30%的图像有望由AI生成或辅助生成。Midjourney是一个独立的研究实验室，探索新的思想媒介，扩大人类的想象力，作为一个小型的自筹资金团队，专注于设计、人力基础设施和人工智能。创始人David Holz曾是Leap Motion创始人之一，美国宇航局研究员马克斯·普朗克。团队成员仅11人，人效极高。
- ◆ **Midjourney定期发布新的模型版本，以提高效率、一致性和质量。** Midjourney于2023年3月15日发布V5版本，图像具备非常高的一致性，擅长解释自然语言提示，分辨率更高，并支持高级功能。

图 1: Midjourney V4版本图像展示



(左) 充满活力的加州罂粟花——V4

(右) 高对比度超现实拼贴——V4

资料来源: Midjourney官网, 上海证券研究所

图 2: Midjourney V5版本图像展示



(左) 充满活力的加州罂粟花——V5

(右) 高对比度超现实拼贴——V5

资料来源: Midjourney官网, 上海证券研究所



1.1 创意图像生成 (2C) : DALL-E2

- ◆ **DALL-E2**可根据自然语言的文本描述创建图像和艺术形式，是**DALL-E**的升级版。开发人员可以通过DALL-E2的API直接进入其应用程序和产品，使用时按照用量收费。
- ◆ **DALL-E2**核心竞争力是精确的**AI**绘画能力，其在**DALL-E**的基础上增加了组合概念、属性及样式的功能，生成更**生动、更复杂的图像**。图像分辨率足足提升了4倍，画质更真实，真实度达到88.8%，识别更精确，语义匹配度达到71.7%。

图 3: DALL-E2针对同一描述的不同风格作品展示



资料来源: DALL-E2官网, 上海证券研究所

图 4: DALL-E (左) 与DALL-E2 (右) 图像对比



资料来源: DALL-E2官网, 上海证券研究所



1.2 功能图像生成 (2) : Stable Diffusion

- ◆ **Stable Diffusion**是高质量的文生图的潜在扩散模型。其核心技术来源于AI视频剪辑技术创业公司Runway的Patrick Esser，以及慕尼黑大学机器视觉学习组的Robin Romabach。
- ◆ **Stable Diffusion**是Stability AI旗下发布的第一个开源模型。截至2022年10月Stable Diffusion已经有超过20万开发者下载和获得授权，各渠道累计日活用户超过1000万。
- ◆ 开源社区使得**Stable Diffusion**的用户关注度和应用广度更广。在目前的Midjourney、Stable Diffusion、DALL-E2三大新兴文本转图像模型中，Stable Diffusion诞生得最晚，但由于拥有发展良好的开源社区，它的用户关注度和应用广度都更广。

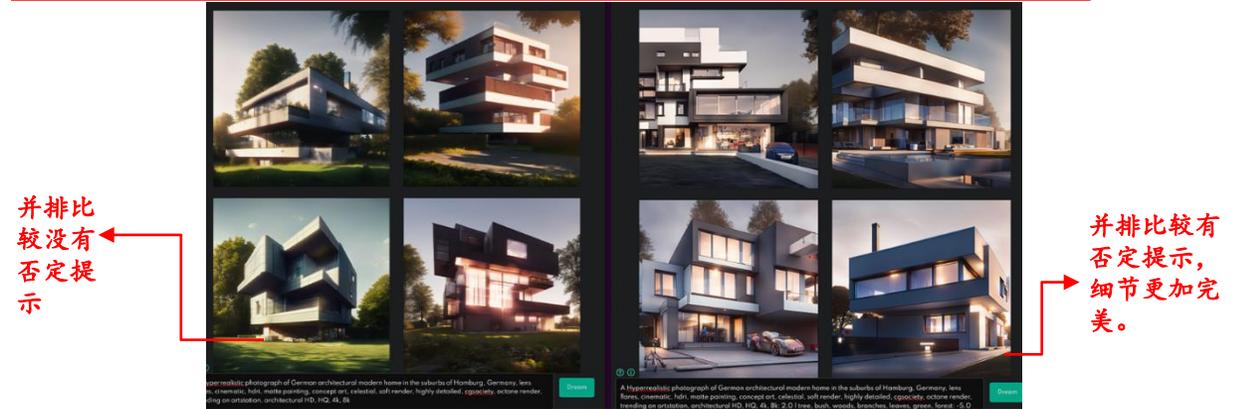
图 5: Stable Diffusion 2.1版本图像能力提升



资料来源：新智元，上海证券研究所



图 6: Stable Diffusion否定提示案例



并排比较没有否定提示

并排比较有否定提示，细节更加完美。

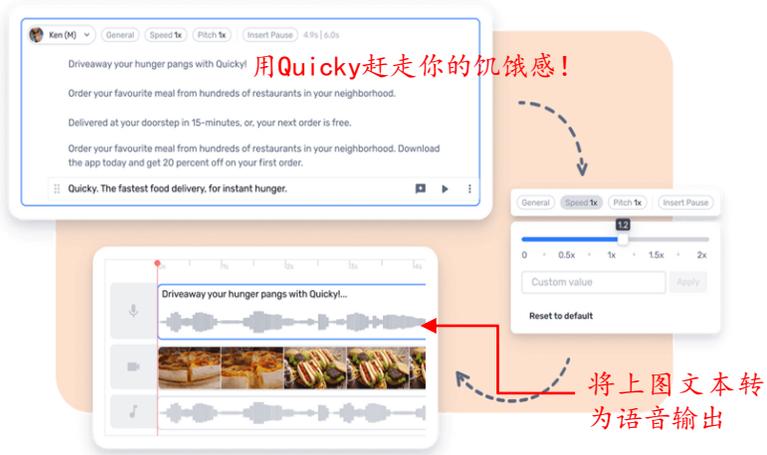
资料来源：新智元，上海证券研究所

二、音频领域

2.1 TTS场景: Murf ai

- ◆ **Murf ai: 2020年10月创办的AI语音生成器，可以利用机器学习和深度学习将文本转换为自然语音。**
- ◆ **Murf ai可以在短时间内输出高质量画外音，用户可以使用AI语音克隆功能来满足自己对语言情感更多元化的需求。**其可以通过调整音调、音量、语言和速度来提供栩栩如生的发音与全方位的人类情感以此提供全方位服务。
- ◆ **Murf ai拥有可调用的AI功能，以此来确保特定术语能够准确无误的传达。** Murf ai提供多合一语音发生器，操作界面简单便捷，能够与谷歌幻灯片共享编辑，可以在AI语音中选择最合适的演示文稿语音。

图 7: 文本转语音: 高质量真人声



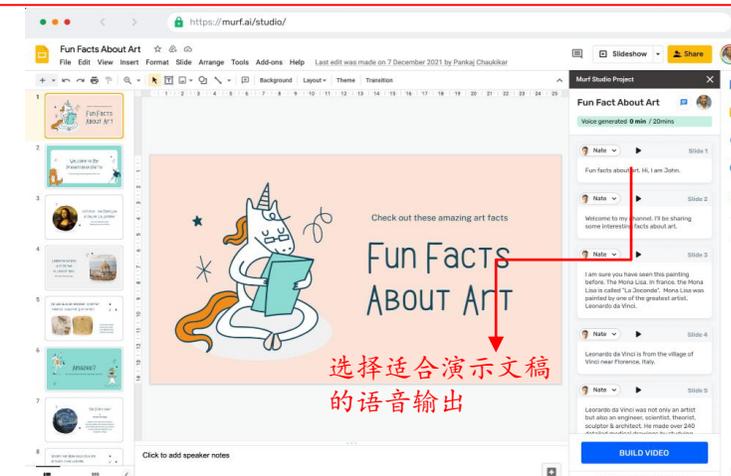
资料来源: Murf ai官网, 上海证券研究所

图 8: AI语音克隆功能: 栩栩如生永久使用



资料来源: Murf ai官网, 上海证券研究所

图 9: 谷歌幻灯片画外音: 简化、同步、经济



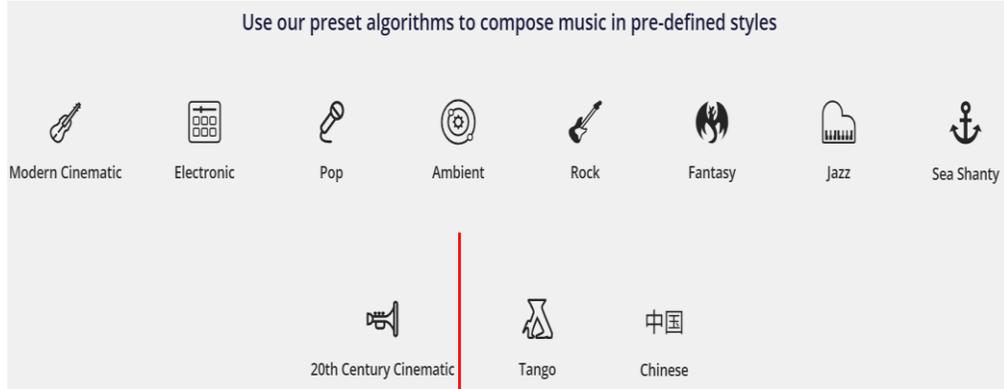
资料来源: Murf ai官网, 上海证券研究所



2.2 歌曲生成: AIVA

- ◆ **AIVA是第一个被正式赋予作曲家地位的AI**，创办于2016年，旨在通过AI创建个性化配乐来提升用户能力，其接受了数千份乐谱的训练。
- ◆ **AIVA为用户提供两种音乐创作方式**：第一种，使用预先训练的“预设样式”，这样样式基于内部精心制作的各种数据集（曲调和级数、节奏模式和旋律线等重要音乐特征组成；第二种，使用上传的音乐来创作，以创作出具有相似音乐特征但截然不同的曲目。

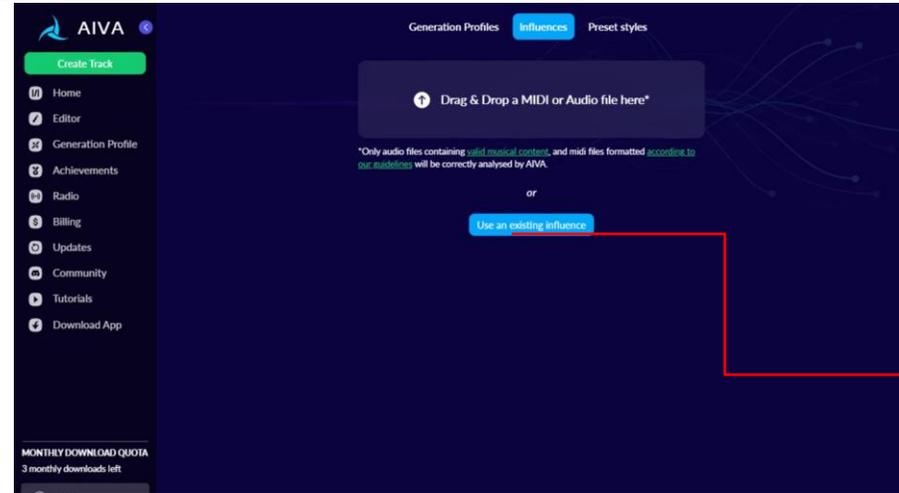
图 10: 预定义曲风多样



可以使用AIVA的预设算法以预定义的风格创作音乐

资料来源: AIVA官网, 上海证券研究所

图 11: 使用临时轨道控制合成过程



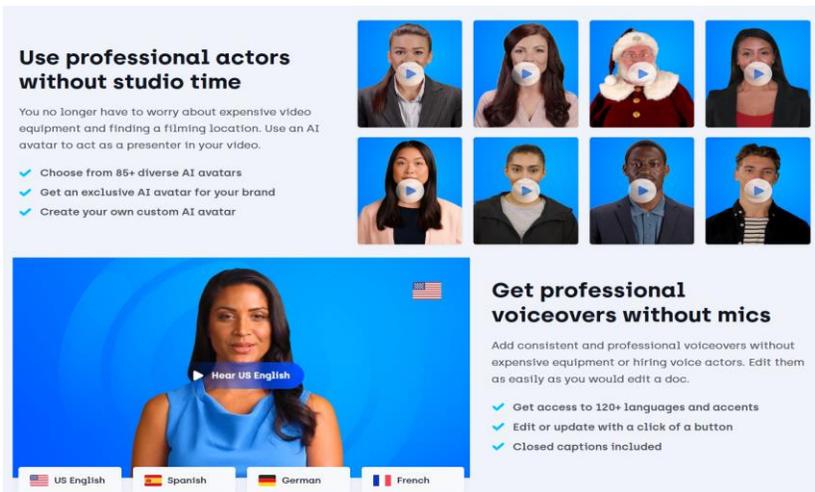
可以上传自己的MIDI文件来影响AIVA的作曲过程，从而得到一个与现有乐谱具有相似情感影响的原始乐谱

资料来源: 《AAIVA为短视频打造专属音乐背景》，上海证券研究所



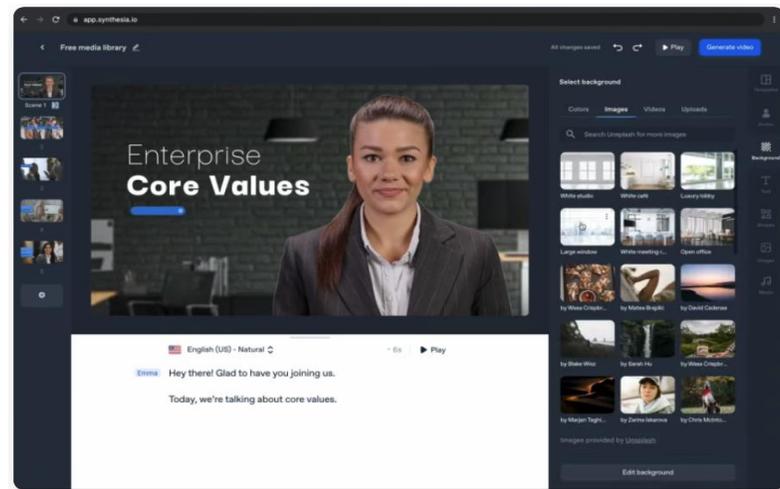
- ◆ Synthesia由来自伦敦大学学院、斯坦福大学、慕尼黑工业大学和剑桥大学的AI研究人员和企业家团队于2017年创立。2020年推出的AI视频创作产品“Synthesia”已有数千家公司在使用。
- ◆ 1) 利用AI技术创建和定制数字孪生模型从而生成视频。用户可以从现有演员库中选择或上传自己的视频来创建AI形象，然后输入脚本让AI配音、拼接素材就能输出视频，节省高达80%的时间和预算。
- ◆ 2) 主要运用于企业传播、数字视频营销和广告本地化。此产品大幅减少真人出镜录制等环节，节省制作费用和周期，同时能一键生成多国语言视频，便于本土化服务。

图 12: Synthesia支持使用AI头像以及获得专业配音



资料来源: Synthesia官网, 上海证券研究所

图 13: Synthesia官网演示的媒体素材库



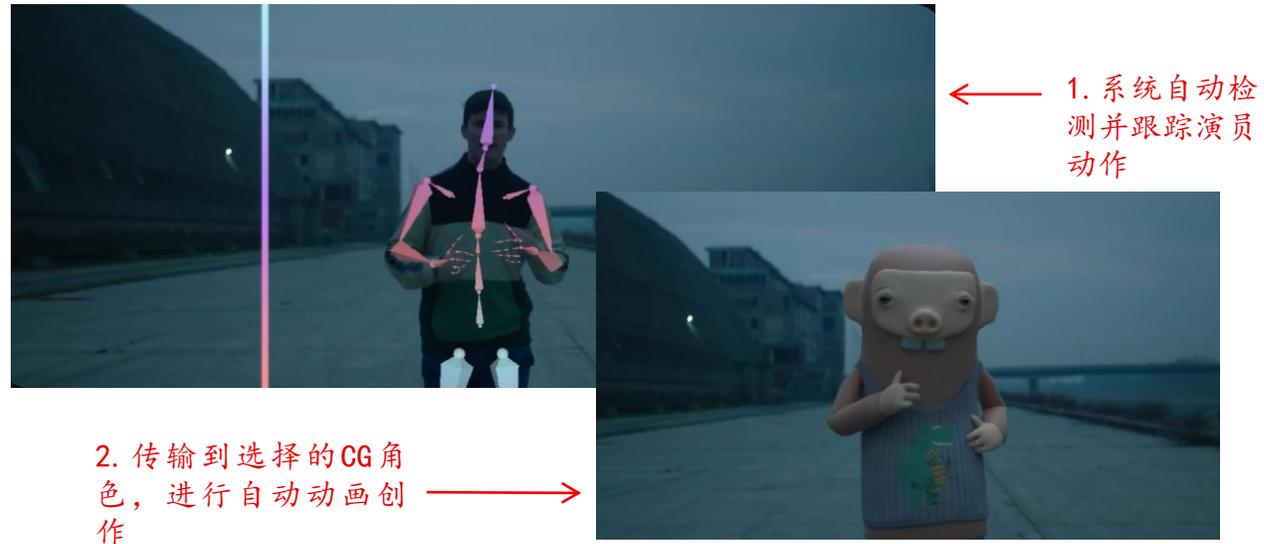
资料来源: Synthesia官网, 上海证券研究所



3.1 拼凑生成视频: Wonder Studio

- ◆ **Wonder Studio**是一款可以自动将**CG角色动画**、**打光**并组合成**真人场景的AI工具**。
- ◆ **1) 无需逐镜头工作**。将**CG角色模型**上传到一个镜头或整个场景，系统将在整个序列中自动检测剪辑并跟踪演员。
- ◆ **2) 无需繁重的逐帧视觉特效工作**。系统根据单镜头自动检测演员表演，并传输到所选**CG角色**，自动动画、照明和创作。
- ◆ **3) 适应现有流水线**。Wonder Studio AI自动化了80%-90%的“客观”视觉特效工作，并将剩余的“主观”工作留给艺术家，同时支持导出到他们已经使用的软件中。

图 14: Wonder Studio无需繁重的逐帧VFX工作



资料来源: Wonder Studio官网, 上海证券研究所



3.2 从头生成视频：Runway Gen-2

- ◆ **Runway：文本到视频成功落地的先行者。** Runway成立于2018年，公司研究团队一直致力于构建多模态人工智能系统，以降低视频创作的门槛，帮助人们轻松制作出内容强大且富有创意的视频内容。
- ◆ 1) 2023年2月6日，Runway推出了基于扩散模型的视频生成模型**Gen-1**，并在Discord中开启内测。
- ◆ 2) Runway在3月20日发布升级版**Gen-2**，目前已经能够很好地将场景、色调简单的几何体变成酷炫的视频，但复杂场景的精细化处理较为欠缺。**Gen-2能够仅使用文本提示就合成任意风格的视频，同时也支持文本+图像生成视频。**

图 15： Runway Gen-1论文展示案例



基于时序扩散模型进行视频风格迁移

资料来源：Gen-1论文地址 <https://arxiv.org/abs/2302.03011>，上海证券研究所

图 16： Gen-2的文本生成视频示例



文本提示：在纽约市的一间阁楼里，傍晚的阳光透过窗户温柔地洒进屋内。

资料来源：Runway官网，上海证券研究所



四、投资建议

- 1、建议关注A股稀缺的优质内容型平台公司【芒果超媒】，有望成为传媒估值中枢锚。
- 2、建议关注拥有海外用户/业务，有望接入GPT的优质标的，如【汤姆猫】（全球用户）、【昆仑万维】（Opera）、【神州泰岳】（游戏出海+NLP）、【华凯易佰】（跨境电商）、【焦点科技】（跨境电商）、【蓝色光标】（出海）。
- 3、建议关注可与生成式图像AI Midjourney对标的【浙文互联】（米画）、【视觉中国】。
- 4、建议关注国内电商类的【新华都】（电商代运营）、【值得买】（内容测评）、【遥望科技】（虚拟人）、【壹网壹创】（电商代运营）、【青木股份】（电商代运营）、【若羽臣】（电商代运营）、【丽人丽妆】（电商代运营）。
- 5、建议关注AIGC+游戏，如【三七互娱】、【完美世界】、【吉比特】、【盛天网络】、【世纪华通】、【巨人网络】；AIGC+影视，如【百纳千成】、【欢瑞世纪】、【光线传媒】、【华策影视】、【上海电影】、【慈文传媒】、【中文在线】等。
- 6、建议关注【福昕软件】、【光云科技】、【力盛体育】。



- 1、**宏观经济风险**：AI相关技术研发及应用成本高企，且后续资本支出需求较大，若宏观环境变化导致相关公司经营情况波动、现金流出现问题，可能会影响研发进展；
- 2、**地缘政治风险**：若相关公司涉及到数据安全等领域的问题可能会引起政治风险；
- 3、**技术发展不及预期**：GPT与下游应用结合表现存在不及预期的可能；
- 4、**AIGC行业发展不及预期**：若底层通用大模型发展不及预期，可能影响下游小模型发展。



行业评级与免责声明

分析师声明

作者具有中国证券业协会授予的证券投资咨询资格或相当的专业胜任能力，以勤勉尽责的职业态度，独立、客观地出具本报告，并保证报告采用的信息均来自合规渠道，力求清晰、准确地反映作者的研究观点，结论不受任何第三方的授意或影响。此外，作者薪酬的任何部分不与本报告中的具体推荐意见或观点直接或间接相关。

公司业务资格说明

本公司具备证券投资咨询业务资格。

投资评级体系与评级定义

股票投资评级：	分析师给出下列评级中的其中一项代表其根据公司基本面及（或）估值预期以报告日起6个月内公司股价相对于同期市场基准指数表现的看法。	
	买入	股价表现将强于基准指数20%以上
	增持	股价表现将强于基准指数5-20%
	中性	股价表现将介于基准指数±5%之间
	减持	股价表现将弱于基准指数5%以上
	无评级	由于我们无法获取必要的资料，或者公司面临无法预见结果的重大不确定性事件，或者其他原因，致使我们无法给出明确的投资评级
行业投资评级：	分析师给出下列评级中的其中一项代表其根据行业历史基本面及（或）估值对所研究行业以报告日起12个月内的基本面和行业指数相对于同期市场基准指数表现的看法。	
	增持	行业基本面看好，相对表现优于同期基准指数
	中性	行业基本面稳定，相对表现与同期基准指数持平
	减持	行业基本面看淡，相对表现弱于同期基准指数

相关证券市场基准指数说明：A股市场以沪深300指数为基准；港股市场以恒生指数为基准；美股市场以标普500或纳斯达克综合指数为基准。

投资评级说明：

不同证券研究机构采用不同的评级术语及评级标准，投资者应区分不同机构在相同评级名称下的定义差异。本评级体系采用的是相对评级体系。投资者买卖证券的决定取决于个人的实际情况。投资者应阅读整篇报告，以获取比较完整的观点与信息，投资者不应以分析师的投资评级取代个人的分析与判断。



行业评级与免责声明

免责声明

。本公司不会因接收人收到本报告而视其为客户。

本报告版权归本公司所有，本公司对本报告保留一切权利。未经书面授权，任何机构和个人均不得对本报告进行任何形式的发布、复制、引用或转载。如经过本公司同意引用、刊发的，须注明出处为上海证券有限责任公司研究所，且不得对本报告进行有悖原意的引用、删节和修改。在法律许可的情况下，本公司或其关联机构可能会持有报告中涉及的公司所发行的证券或期权并进行交易，也可能为这些公司提供或争取提供多种金融服务。

本报告的信息来源于已公开的资料，本公司对该等信息的准确性、完整性或可靠性不作任何保证。本报告所载的资料、意见和推测仅反映本公司于发布本报告当日的判断，本报告所指的证券或投资标的的价格、价值或投资收入可升可跌。过往表现不应作为日后的表现依据。在不同时期，本公司可发出与本报告所载资料、意见或推测不一致的报告。本公司不保证本报告所含信息保持在最新状态。同时，本公司对本报告所含信息可在不发出通知的情形下做出修改，投资者应当自行关注相应的更新或修改。

本报告中的内容和意见仅供参考，并不构成客户私人咨询建议。在任何情况下，本公司、本公司员工或关联机构不承诺投资者一定获利，不与投资者分享投资收益，也不对任何人因使用本报告中的任何内容所引致的任何损失负责，投资者据此做出的任何投资决策与本公司、本公司员工或关联机构无关。

市场有风险，投资需谨慎。投资者不应将本报告作为投资决策的唯一参考因素，也不应当认为本报告可以取代自己的判断。

