



信达证券  
CINDA SECURITIES

Research and  
Development Center

# AI 音频软硬结合声动未来，AI 视频扬帆起航

传媒互联网及海外行业

## 证券研究报告

## 行业研究

## 行业专题研究（普通）

## 传媒互联网及海外行业

投资评级 看好

上次评级 看好

冯翠婷 传媒互联网及海外 首席分析师

执业编号：S1500522010001

联系电话：17317141123

邮箱：fengcuiting@cindasc.com

信达证券股份有限公司

CINDA SECURITIES CO., LTD

北京市西城区宣武门西大街甲127号金隅大厦

B座

邮编：100053

## AI 音频软硬结合声动未来，AI 视频扬帆起航

2024 年 10 月 16 日

## 本期内容提要：

- **OpenAI DevDay AI 音频更新亮眼，有望催化 AI 语音交互应用成熟。**  
2024 年 10 月 1 日，OpenAI 开发者大会上，公司公布了几大创新：实时 API、提示词缓存、模型蒸馏、视觉微调。与 ChatGPT 的高级语音模式类似，Realtime API 支持使用六种预设声音进行自然的语音对语音对话。文本输入令牌的价格为 5 美元/1M tokens，输出令牌的价格为 20 美元/1M tokens。音频输入的价格 100 美元/1M tokens，输出的价格为 200 美元/1M tokens。这相当于每分钟音频输入约 0.06 美元，每分钟音频输出约 0.24 美元。
- **海外 AI+音频代表性应用 Duolingo 股价历史新高，Duocon2024 发布独特虚拟 IP 视频通话、虚拟世界探险学玩、扩展音乐课程等功能，订阅付费渗透率逐渐提升至 8%以上。**公司财务和经营数据均表现较为亮眼。24Q2 公司实现营收 1.78 亿美元，比去年同期增长 41%；调整后的 EBITDA 为 4810 万美元，去年同期为 2090 万美元，24Q2 和 23Q2 调整后的 EBITDA 利润率分别为 27.0%和 16.5%；用户情况：24Q2 Duolingo 月活超过 1 亿，季度末付费用户总数达到 800 万，比去年同期增长 52%，即目前 Max 产品付费率超过 8%，本次产品更新有望进一步加速 Max 产品付费渗透率。
- **AI+视频迭代持续火热，Meta、字节发布新模型，美图 MOKI 一键成片产品正式上线，可灵、Pika 迭代新版本。**1) 10 月 4 日，Meta 发布了 Meta Movie Gen 模型，是一项针对图像、视频和音频的突破性生成式 AI 研究，Movie Gen 具有四种功能：视频生成、个性化视频生成、精确视频编辑和音频生成。2) 字节豆包视频大模型 PixelDance&Seaweed——主打时序性复杂动作指令和交互能力。字节跳动正式进军 AI 视频生成领域，其数据、技术和人才等储备深厚。9 月 25 日，视频大模型赋能的字节 C 端产品——即梦正式开启内测邀请。3) 2024 年 9 月 30 日，快手可灵 AI 视频模型产品发布产品更新，产品重点新增「对口型」功能；4) Pika 1.5 引入了一套创新效果，称为 Pikaffects，允许用户轻松大幅改变视频元素，例如：膨胀、爆炸、压碎、融化、挤压和蛋糕化；5) 美图 MOKI 一键成片正式上线开放。
- **投资建议：**

**AI 音频方面：**在语音实时对话效果不断提升的基础上，相比之前的虚拟陪伴、虚拟人物的语音互动，OpenAI Realtime API 更有望催生一大批基于纯语音交互的 AI 应用，例如旅行规划、点餐、出游、语音学习、电商购物、教育等诸多方向。建议关注：盛天网络、中文在线、汤姆猫、昆仑万维等；

**AI 视频方面：**远不止于视频生成，从 AI 生成到 AI 工作流，一站式

AI 视频生成+剪辑+故事创作有望成为产业核心发展方向。AI+视频时代来临，思考哪类公司存在商业化变现的可能性？我们认为，1）一站式平台型公司，如 Adobe、美图公司；2）AI+视频技术头部服务商转型产品类公司，如 Runway、商汤科技；3）视频剪辑类公司，如快手；4）广告营销类公司，如易点天下、蓝色光标、因赛集团、利欧股份；5）UGC 社区类公司，如 Bilibili；6）视频数据类公司，如捷成股份、华策影视、视觉中国、中广天择；7）IP 类公司，如上海电影、阅文集团、汤姆猫、中文在线、果麦文化；8）探索 AI 视频工作流及其他创作方向类公司，如博纳影业、超讯通信、柠萌影视。9）其他建议关注猫眼娱乐、光线传媒、芒果超媒、万达电影等。

➤ 风险因素：AI 大模型发展不及预期、AI 视频产品付费率提升不及预期；

表：相关上市公司估值表（截至 2024.10.15）

分类	证券简称	总市值 (亿元)	归母净利润（百万元）					市盈率	
			23A	24E	25E	26E	24E	25E	26E
一站式平台型	ADBE	15,840.1	53483.0	59276.0	65823.0	74153.0	26.7	24.1	21.4
	美图公司	104.7	368.3	548.5	787.9	1046.0	19.1	13.3	10.0
技术服务类	商汤-W	490.5	-6440.0	-3645.0	-2522.0	-1742.0	-	-	-
UGC 社区类	哔哩哔哩-W	608.0	-4822.32	-1734.0	-143.8	990.0	-	-422.9	61.4
视频剪辑类	快手-W	1,868.2	6396.0	15734.6	21711.7	27026.2	11.9	8.6	6.9
IP 类公司	阅文集团	248.5	804.9	1249.2	1400.7	1543.0	19.9	17.7	16.1
	上海电影	101.4	127.0	232.8	333.9	422.7	43.6	30.4	24.0
	汤姆猫	162.4	-864.6	200.0	300.0	400.0	81.2	54.1	40.6
广告营销类	易点天下	82.9	217.0	286.7	357.1	433.0	28.9	23.2	19.1
	蓝色光标	169.5	116.6	371.6	527.1	639.6	45.6	32.2	26.5
	因赛集团	59.5	41.6	/	/	/	/	/	/
	利欧股份	113.8	1966.0	/	/	/	/	/	/
视频数据训练类	华策影视	119.2	382.2	455.8	517.5	562.8	26.2	23.0	21.2
	视觉中国	86.7	145.6	169.2	197.9	224.1	51.3	43.8	38.7
	捷成股份	137.2	450.0	593.0	650.3	728.3	23.1	21.1	18.8
	中广天择	32.2	-8.7	/	/	/	/	/	/
AI 短剧/短片探索类	博纳影业	65.0	-552.6	314.2	498.9	576.7	20.7	13.0	11.3
	超讯通信	47.2	18.8	111.1	164.8	243.4	42.5	28.6	19.4
	柠萌影视	11.9	213.6	258.1	322.7	372.3	4.6	3.7	3.2
其他	光线传媒	232.0	417.8	1068.1	1212.0	1373.9	21.7	19.1	16.9
	万达电影	226.0	912.2	1338.7	1663.6	1931.3	16.9	13.6	11.7
	芒果超媒	438.5	3555.7	2040.4	2273.7	2524.2	21.5	19.3	17.4
AI 音频	汤姆猫	162.4	-864.6	200.0	300.0	400.0	81.2	54.1	40.6
	盛天网络	53.7	169.3	224.4	291.7	322.4	23.9	18.4	16.7
	昆仑万维	440.0	1258.4	864.7	962.0	1045.9	50.9	45.7	42.1
	中文在线	175.3	89.4	134.3	172.7	215.1	130.6	101.5	81.5

资料来源：iFind、Bloomberg，信达证券研发中心

## 目 录

一、AI+视频持续火热：Meta、字节发布新模型，可灵、Pika、美图迭代版本 .....	5
二、OpenAI DevDay AI 音频更新亮眼，有望催化 AI 语音交互应用成熟 .....	10
三、投资建议 .....	15
四、风险因素 .....	16

## 表 目 录

表 1: OpenAI Realtime 音频 API 价格和其他模型文本价格比较 .....	15
表 2: 相关上市公司估值表（截至 2024.10.15） .....	16

## 图 目 录

图 1: Meta Movie Gen 基础架构 .....	5
图 2: Meta Movie Gen 视频模型和其他模型量化指标比较 .....	6
图 3: Meta Movie Gen 视频模型生成效果比较 .....	6
图 4: Meta Movie Gen 模型视频编辑原理 .....	7
图 5: Meta Movie Gen 模型可生成带有音频的视频 .....	7
图 6: 字节 PixelDance 视频模型复杂 prompt 时序性指令生成 .....	8
图 7: 可灵 AI 0930 更新“对口型”功能 .....	9
图 8: Pika 1.5 官网更新视频效果 .....	9
图 9: 美图 MOKI 一键成片功能展示 .....	10
图 10: Speak 接入 OpenAI Realtime API 支持其角色扮演进行新语言联系功能 .....	11
图 11: Duolingo 2024 Duocon 更新功能（video call with Lily, Adventures, Math, Music） .....	12
图 12: 20Q3-24Q2 Duolingo 日活和月活用户数 .....	12
图 13: 20Q1-24Q2 Duolingo 付费订阅用户数 .....	13
图 15: 西湖心辰 Lingo AI 语音交互演示 .....	13
图 16: 盛天网络给麦 AI 功能迭代 .....	14
图 17: 昆仑万维 Mureka 创作页面 .....	15

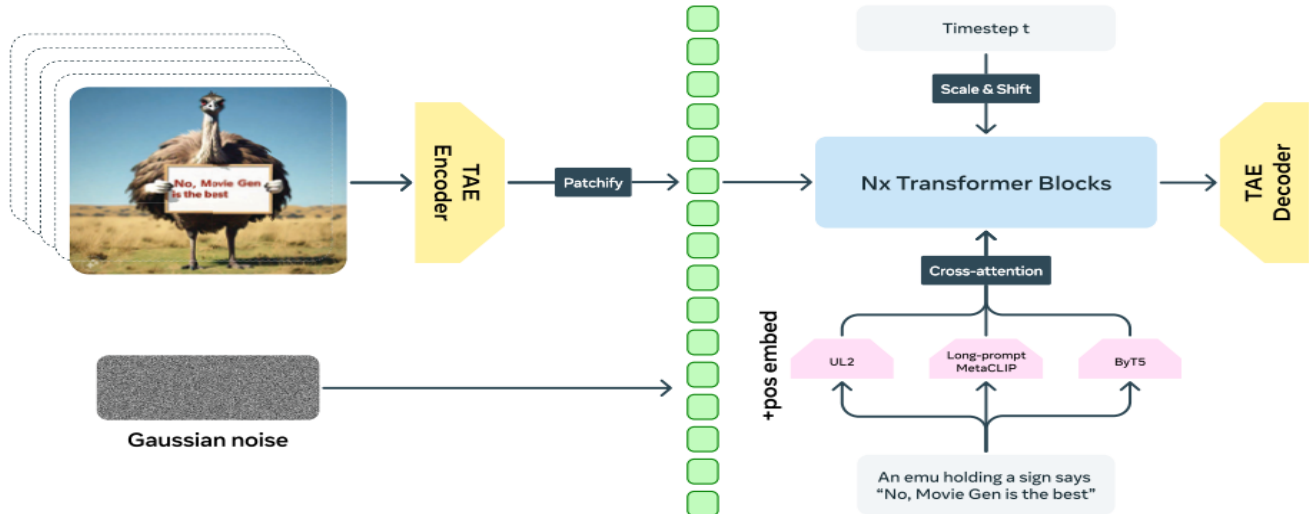
## 一、AI+视频持续火热：Meta、字节发布新模型，可灵、Pika、美图迭代版本

### （一）Meta Movie Gen——主打精确视频编辑和音视频同步生成

2024 年 10 月 4 日，Meta 发布了 Meta Movie Gen 模型，是一项针对图像、视频和音频的突破性生成式 AI 研究，Movie Gen 具有四种功能：**视频生成、个性化视频生成、精确视频编辑和音频生成**。1) 视频生成：该 30B 参数转换器模型能够以每秒 16 帧的速度生成长达 16 秒的视频。2) 个性化视频生成：将一个人的图像作为输入，并将其与文本提示相结合，以生成包含参考人物和文本提示所告知的丰富视觉细节的视频。3) 视频编辑：同一基础模型的编辑变体以视频和文本提示作为输入，精确执行任务以生成所需的输出。它将视频生成与高级图像编辑相结合，执行局部编辑（例如添加、删除或替换元素）以及全局更改（例如背景或样式修改）。4) 音频生成：13B 参数音频生成模型，该模型可以接收视频和可选的文本提示，生成长达 45 秒的高质量高保真音频，包括环境声音、音效 (Foley) 和乐器背景音乐，所有这些都与视频内容同步，在音频质量、视频到音频对齐和文本到音频对齐方面总体上实现了一流的性能。

Meta 首先通过一个时间自动编码器模型（TAE）训练了一个时空压缩的隐空间，然后再基于此训练了一个生成模型。模型架构上，Meta 采用了 Transformer，整体位于 Llama 3 的设计空间中。推理阶段，Meta 的一个创新思路是首先使用 Llama 3 对用户输入的提示词进行重写，将其扩展成更加详细的版本。实践表明该方法确实有助于提升生成结果的质量。

图 1：Meta Movie Gen 基础架构



资料来源: Meta 官网, 信达证券研发中心

**个性化视频生成：**基于 Meta 发布的 30B Movie Gen 视频模型，公司拓展到了个性化视频的生成，用户只需提供人物图像输入和对应的文本提示词，就能生成包含该人物以及文本描述的细节的视频。Meta 表示 Movie Gen 生成的个性化视频在保留人类身份和运动方面做到了 SOTA。训练过程先是进行预训练（分为身份注入、长视频生成、提升自然度三个阶段），然后执行监督式微调。

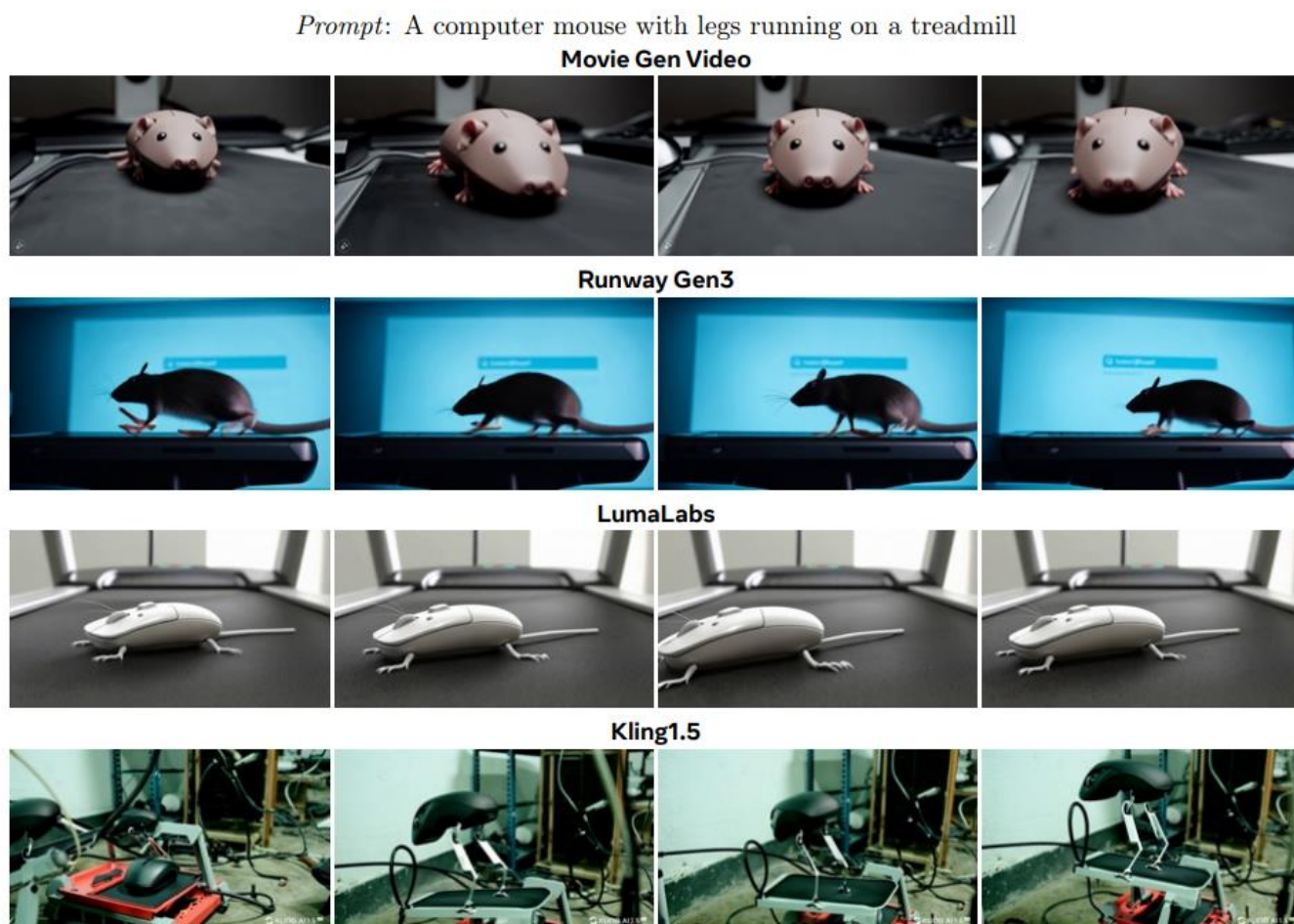


图 2: Meta Movie Gen 视频模型和其他模型量化指标比较

	MOVIE GEN VIDEO net win rate vs. prior work				$\sigma$
	Runway Gen3	LumaLabs	OpenAI Sora	Kling1.5	
Overall Quality	35.02	60.58	8.23	3.87	$\pm 5.07$
Consistency	33.1	42.14	8.22	13.5	$\pm 4.08$
Motion Naturalness	19.27	29.33	4.43	0.52	$\pm 3.98$
Motion Completeness	-1.72	23.59	8.86	-10.04	$\pm 1.68$
Text-alignment	10.45	12.23	17.72	-1.99	$\pm 3.74$
Realness	48.49	61.83	11.62	37.09	$\pm 2.52$
Aesthetics	38.55	48.19	6.45	26.88	$\pm 4.84$

资料来源: Meta 官网, 信达证券研发中心

图 3: Meta Movie Gen 视频模型生成效果比较

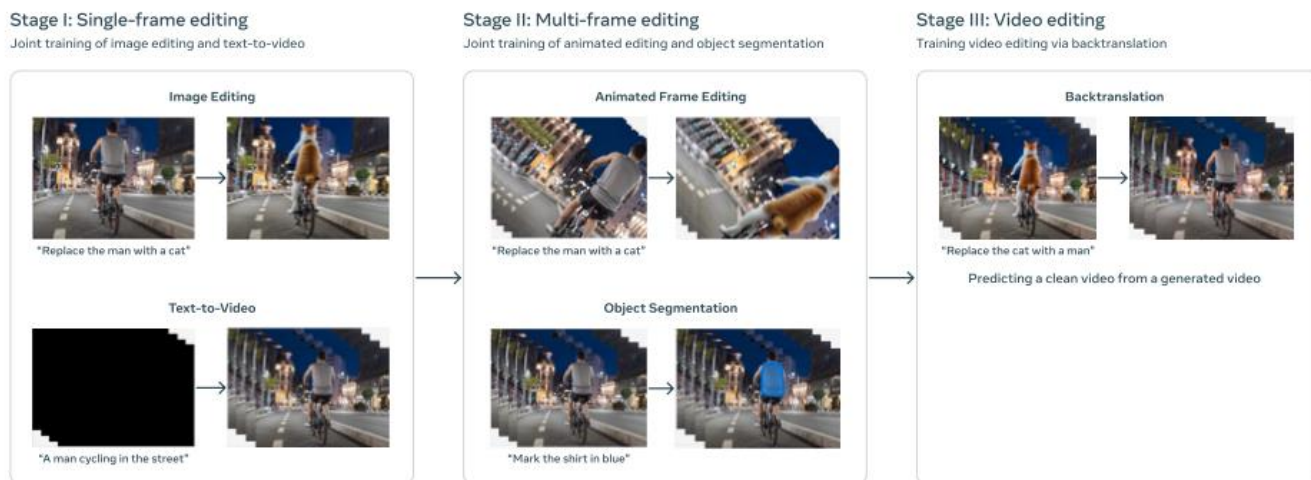


资料来源: Meta 官网, 信达证券研发中心

**视频精确编辑:** 给定一段视频和文本提示词, 模型可以生成符合要求的经过修改的输出, 其中包括一些非常高阶的编辑功能, 比如添加、移除和替换元素, 修改背景和风格等全局要素。Meta 团队同样采用了一种多阶段方法:

首先执行单帧编辑，然后进行多帧编辑，之后再整体编辑视频，效果图中展示较好。

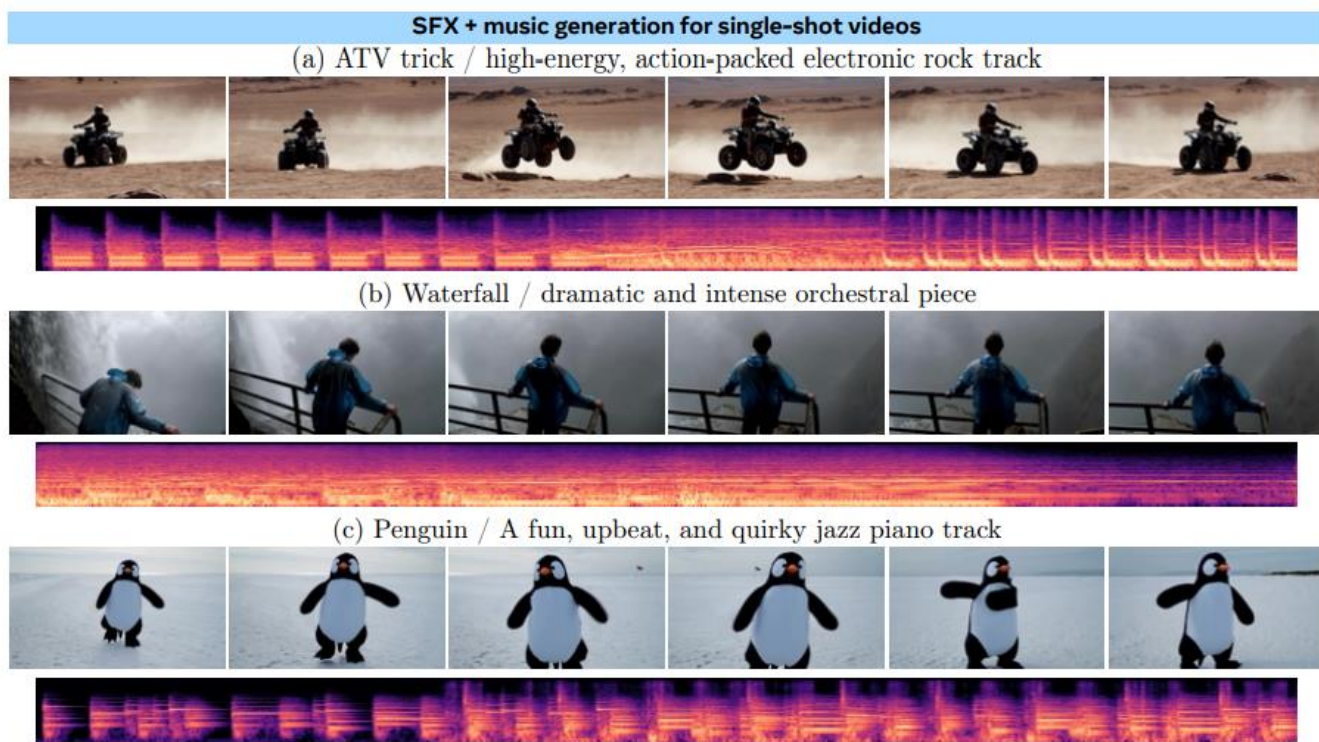
图 4: Meta Movie Gen 模型视频编辑原理



资料来源: Meta 官网, 信达证券研发中心

**音频同步生成:** Meta 训练了一个 13B 参数的音频生成模型 **Movie Gen Audio**。该模型可以基于视频和可选的文本提示词生成长达 45 秒的高质量高保真音频，包括环境声音、音效 (Foley) 和背景音乐 —— 所有这些都与视频内容同步。公司还提出了一种音频扩展技术，可以为任意长度的视频生成连贯的音频，用户后续可以直接生成带有背景音乐的 AI 视频，相较目前 AI 视频竞品来看，技术端有了较为不错的进步。

图 5: Meta Movie Gen 模型可生成带有音频的视频



资料来源: Meta 官网, 信达证券研发中心



## （二）字节豆包视频大模型 PixelDance&Seaweed——主打时序性复杂动作指令和交互能力

字节跳动正式进军 AI 视频生成领域，其数据、技术和人才等储备深厚。2024 年 9 月 24 日，2024 火山引擎 AI 创新巡展在深圳举办，发布了两款豆包视频生成大模型——PixelDance 和 Seaweed，开启了邀测阶段。9 月 25 日，视频大模型赋能的字节 C 端产品——即梦正式开启内测邀请。

1) PixelDance V1.4 是基于 DiT 结构的视频生成大模型，同时支持文生视频和图生视频，能够一次性生成长达 10 秒的精彩视频片段。模型可完成时序性多拍动作，支持多主体复杂交互，还拥有丰富的运镜效果，多风格多比例兼容性强，能快速生成优质的视频片段，赋能影视创作，广告传媒，短视频，直播，电商等多个场景。

2) Seaweed 视频生成大模型基于 Transformer 架构，支持文生视频和图生视频。利用时空压缩的潜空间进行训练，模型原生支持多分辨率生成，适配横屏、竖屏，并能够根据用户输入的高清图像分辨率进行适配和保真。默认输出为 720p 分辨率、24fps、时长 5 秒，并可动态延长至 20-30 秒。

图 6: 字节 PixelDance 视频模型复杂 prompt 时序性指令生成(男人走进画面, 女人转头看着他, 互相拥抱, 背景周围的人在走)



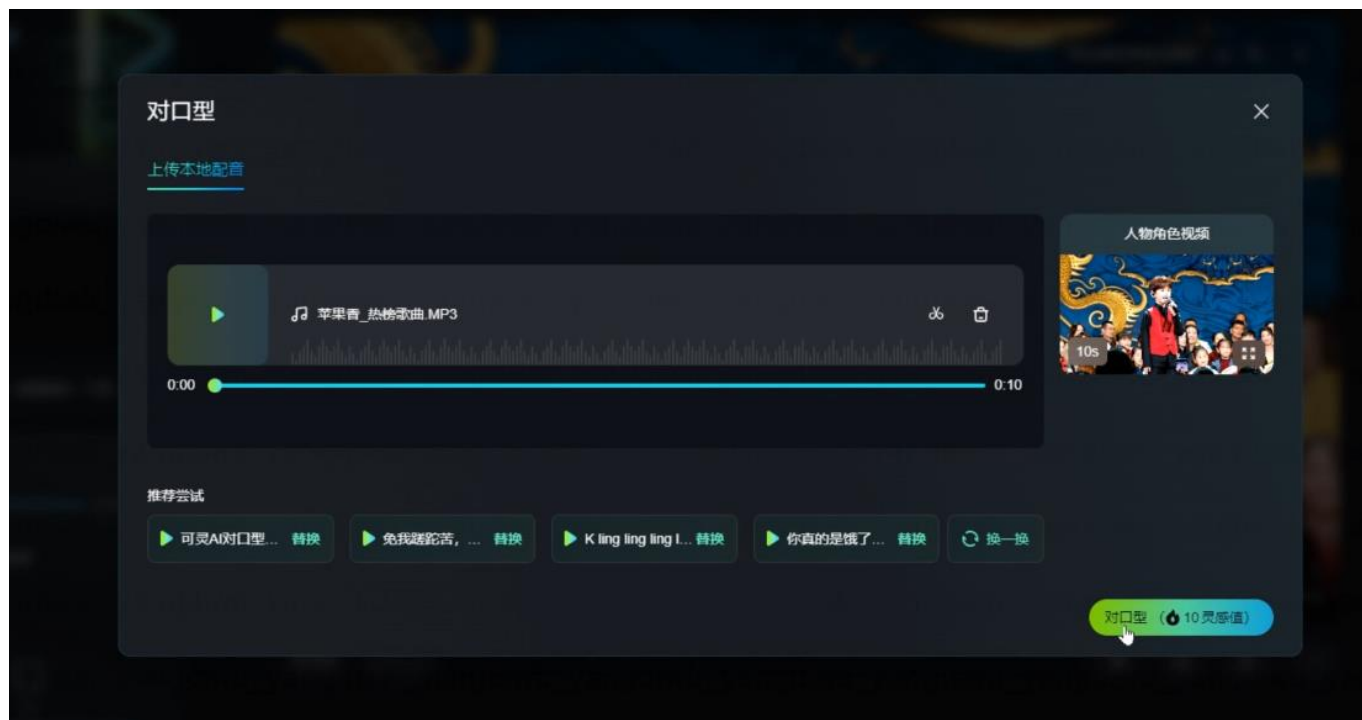
资料来源: 火山引擎官网, 信达证券研发中心

## （三）可灵 AI 更新“对口型”功能

2024 年 9 月 30 日，快手可灵 AI 视频模型产品发布产品更新，产品重点新增「对口型」功能，支持用户在可灵 AI 生成人物视频后，上传配音或歌唱，可灵 AI 就能让视频人物口型和音频完美同步，仿佛真人对话/唱歌，视频瞬间栩栩如生。1) 使用可灵 AI，生成包含完整的人物面容的视频；2) 点击预览视频下的「对口型」；3) 上传配音/歌唱音频，即可进行对口型效果生成。可灵 AI 后续同样有望发力视频剪辑功能，更好满足创作者需求。



图 7：可灵 AI 0930 更新“对口型”功能

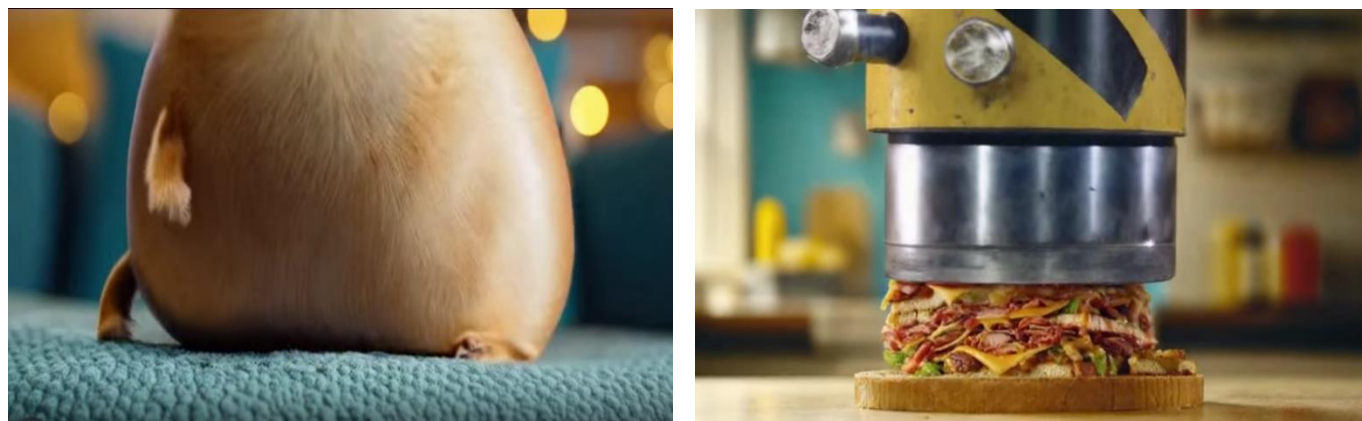


资料来源：可灵 AI 官网，信达证券研发中心

#### （四）Pika 1.5 更新“Pikffects”效果

Pika 1.5 引入了一套创新效果，称为 Pikffects，允许用户轻松大幅改变视频元素，例如：膨胀、爆炸、压碎、融化、挤压和蛋糕化。同时，Pika 1.5 引入了 Big Screen Shots（大屏幕快照）功能，旨在为用户的视频带来电影般的风格和专业品质的视觉效果。借助先进的摄像技术和动态效果，用户可以轻松创建具有影响力的好莱坞风格镜头，而无需昂贵的设备或高级编辑技能。1）子弹时间：一种慢动作效果，摄像机围绕拍摄对象移动，类似于《黑客帝国》等电影中的标志性场景，为关键时刻增添戏剧效果。2）急速变焦：一种快速放大效果，可立即聚焦于某个拍摄对象，营造出紧张和强烈的感觉。3）起重机上升：模拟上升的摄像机运动，拓宽视角并为视频添加专业的电影感。4）快速摇摄：快速的水平摄像机移动，可增加场景之间的活力和平滑的过渡。5）挤压效果：使物体看起来被挤压或变形，为视频添加有趣或夸张的元素。

图 8：Pika 1.5 官网更新视频效果

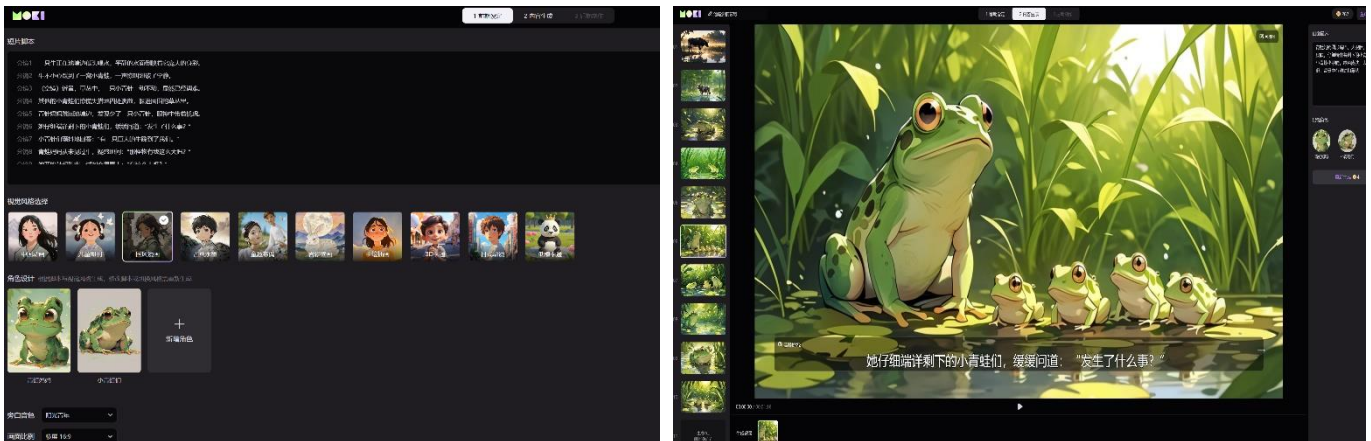


资料来源：Pika 官网，信达证券研发中心

## （五）美图 MOKI 生成短片功能上线——主打一键成片

2024 年 9 月 27 日，美图公司宣布，旗下 AI 短片创作工具 MOKI 全面开放，用户可登录 MOKI 官网进行体验。与市面上流行的文生视频产品、图生视频产品不同，MOKI 专注于 AI 短片创作这一场景，覆盖动画短片、网文短剧、故事绘本、MV 等多个类型的视频内容生产，结合行业需求，有针对性地打造了一套 AI 短片创作工作流。MOKI 的核心优势在于一站式成片能力，得益于 AI 驱动脚本、分镜、视频生成、配音、配乐能力，即使是没有专业背景的普通用户，也能通过 MOKI 轻松上手，快速创作出具有个人特色的短片。

图 9：美图 MOKI 一键成片功能展示



资料来源：MOKI 官网，信达证券研发中心

## 二、OpenAI DevDay AI 音频更新亮眼，有望催化 AI 语音交互应用成熟

2024 年 10 月 1 日，OpenAI 开发者大会上，公司公布了几大创新：实时 API、提示词缓存、模型蒸馏、视觉微调。

- 1) Realtime API：让开发者可以构建近乎实时的“语音转语音”的体验，并可以选择使用 OpenAI 提供的六种声音。
- 2) 提示词缓存：类似于 Anthropic 几个月前推出的缓存功能，允许开发人员在 API 调用之间缓存常用上下文，从而降低成本并改善延迟。
- 3) 模型蒸馏：让开发人员可以使用较大的 AI 模型（例如 o1-preview 和 GPT-4o）来微调较小的模型（例如 GPT-4o mini）。
- 4) 视觉微调：使开发人员能够使用「图像」结合「文本」，来微调他们的 GPT-4o 应用程序，“看”的能力被极大强化。
- 5) 新广场玩法：提示词新的框架、用法等，以及结构化输出新的突破，这些让开发变得前所未有的容易。

OpenAI 推出了 Realtime API 的公开测试版，让所有付费开发者都能在其应用中打造低延迟、多模式体验。与 ChatGPT 的高级语音模式类似，Realtime API 支持使用六种预设声音进行自然的语音对语音对话。OpenAI 还在 Chat Completions API 中引入了音频输入和输出（在新窗口中打开）以支持不需要 Realtime API 的低延迟优势的用例。通过此更新，开发人员可以将任何文本或音频输入传递到 GPT-4o，并让模型以他们选择的文本、音频或两者做出响应。

比如：1) Healthify 是一款营养和健身指导应用程序，它使用实时 API 实现与其 AI 教练 Ria 的自然对话，同时在需要个性化支持时让人类营养师参与进来。2) Speak 是一款语言学习应用程序，它使用 Realtime API 来支持其角色扮演功能，鼓励用户用新语言练习对话。

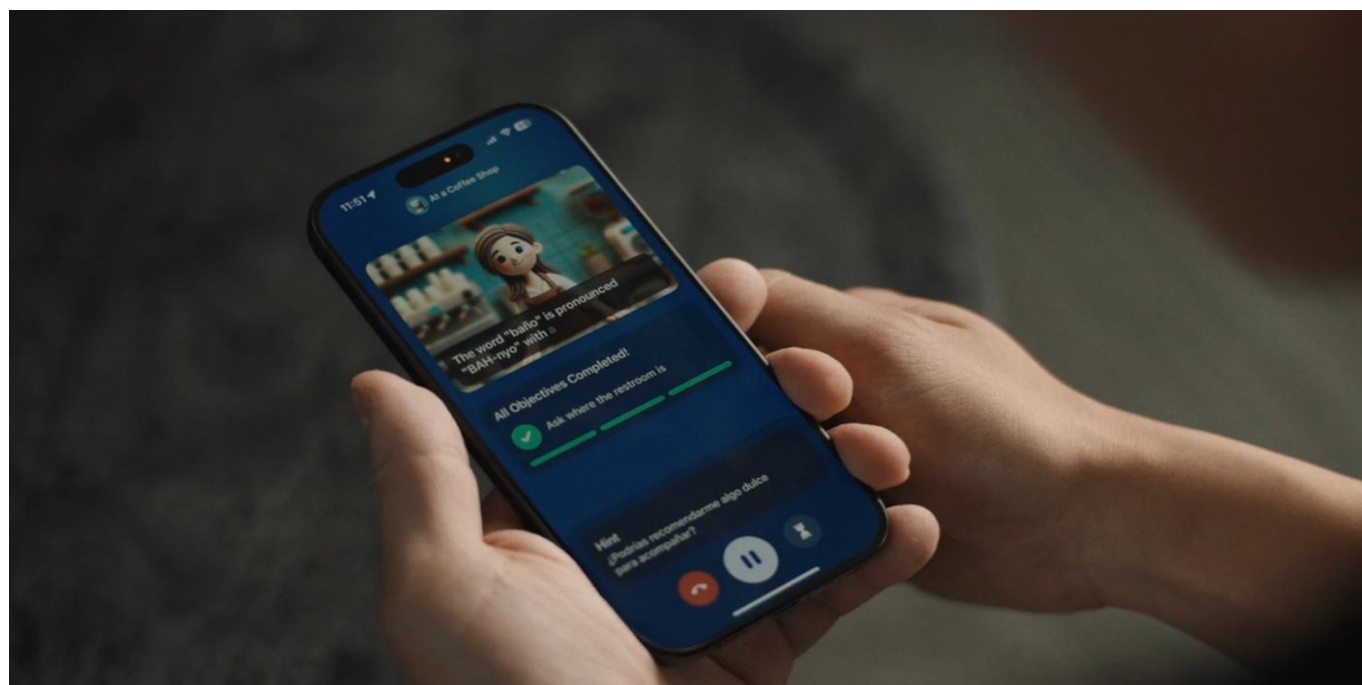
**Realtime API 收费模式和价格：**文本输入令牌的价格为 5 美元/1M tokens，输出令牌的价格为 20 美元/1M tokens。音频输入的价格为 100 美元/1M tokens，输出的价格为 200 美元/1M tokens。这相当于每分钟音频输入约 0.06 美元，每分钟音频输出约 0.24 美元。

**表 1: OpenAI Realtime 音频 API 价格和其他模型文本价格比较**

	Realtime API	GPT-4o	GPT-4o mini
输入	文本: 5 美元/1M tokens	2.5 美元/1M tokens	0.15 美元/1M tokens
	音频: 100 美元/1M tokens, 0.06 美元/分钟		
输出	文本: 20 美元/1M tokens	10 美元/1M tokens	0.6 美元/1M tokens
	音频: 200 美元/1M tokens, 0.24 美元/分钟		

资料来源: OpenAI 官网, 信达证券研发中心

**图 10: Speak 接入 OpenAI Realtime API 支持其角色扮演进行新语言联系功能**



资料来源: OpenAI 官网, 信达证券研发中心

## 海外 AI+音频代表性应用公司及应用场景：Duolingo（语言及其他课程学习）

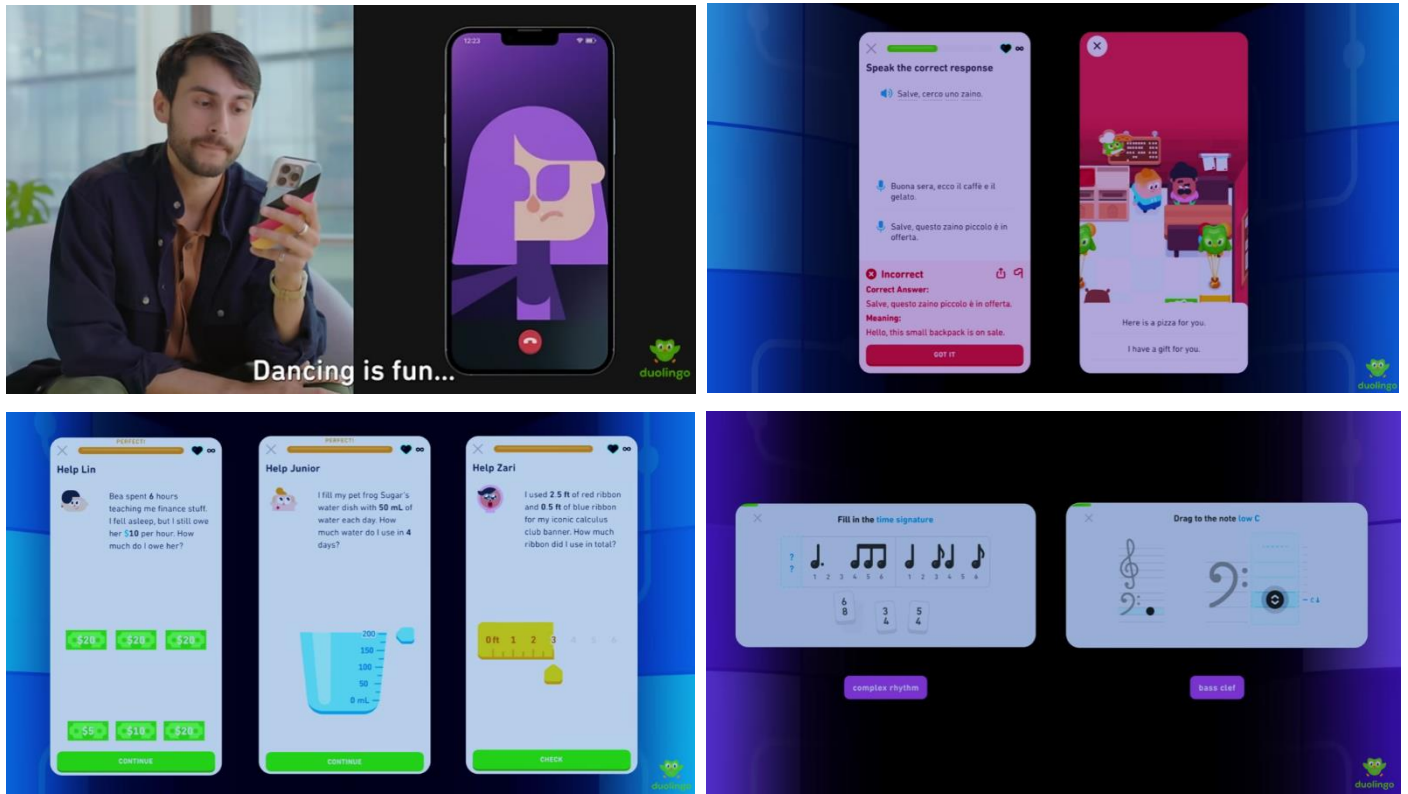
**2024 年 9 月 24 日**，全球领先的移动学习平台 **Duolingo** 在其年度会议 **Duocon 2024** 上宣布了重大产品创新，主要包括：独特虚拟 IP 视频通话、虚拟世界探险学玩、扩展音乐、数学课程等更新。

- 1) **Video Call With Lily** 口语对练：这一功能已上线 iOS 端的英语、西班牙语及法语课程，供海外部分市场的 Duolingo Max 会员使用。允许 Duolingo Max 订阅者与 Duolingo 最受欢迎的角色之一 Lily 进行自发、真实的对话；
- 2) **Adventures**：一种创新的探索游戏式体验，用户首次可以探索动态设置和故事情节，通过结合上下文来与拽姐、奥斯卡等高人气角色对话，完成任务。
- 3) **Music**：Duolingo 宣布与领先的便携式乐器制造商 Loog 合作，打造专为 Duolingo Music 课程设计的紧凑型便携式数码钢琴，课程售价为 249 美元；



4) 数学课程: 多邻国新增四款益智游戏, 模拟更加生活化的教学场景, 满足不同群体的学习需求。

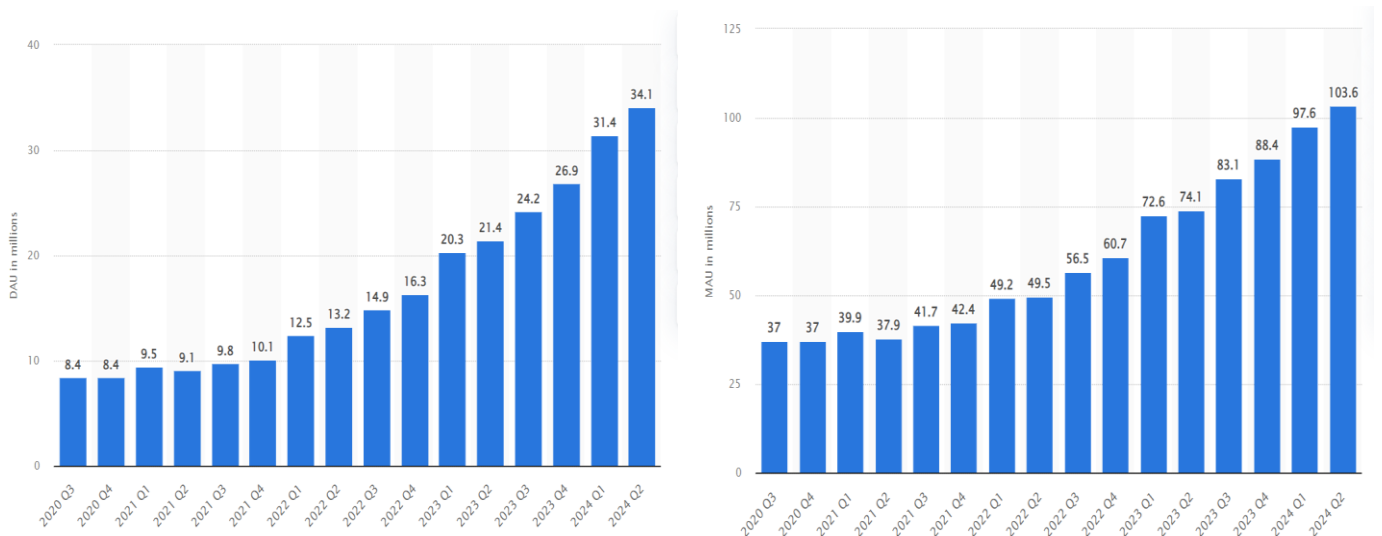
图 11: Duolingo 2024 Duocon 更新功能 (video call with Lily, Adventures, Math, Music)



资料来源: Duolingo 官网, 信达证券研发中心

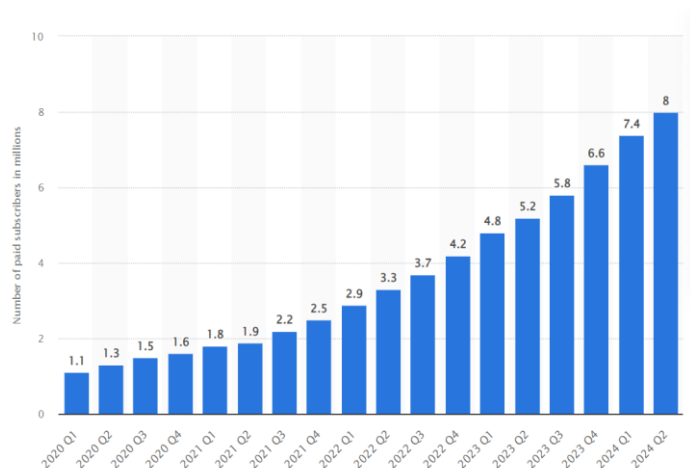
目前 Duolingo 股价历史新高, 受益于生成式 AI, 公司产品的用户付费率逐渐提升, 公司财务和经营数据均表现较为亮眼。24Q2 公司实现营收 1.78 亿美元, 比去年同期增长 41%; 调整后的 EBITDA 为 4810 万美元, 去年同期为 2090 万美元, 24Q2 和 23Q2 调整后的 EBITDA 利润率分别为 27.0%和 16.5%; 用户情况: **24Q2 Duolingo 月活超过 1 亿, 季度末付费用户总数达到 800 万, 比去年同期增长 52%, 即目前 Max 产品付费率超过 8%, 本次产品更新有望进一步加速 Max 产品付费渗透率。**

图 12: 20Q3-24Q2 Duolingo 日活和月活用户数 (单位: 百万人)

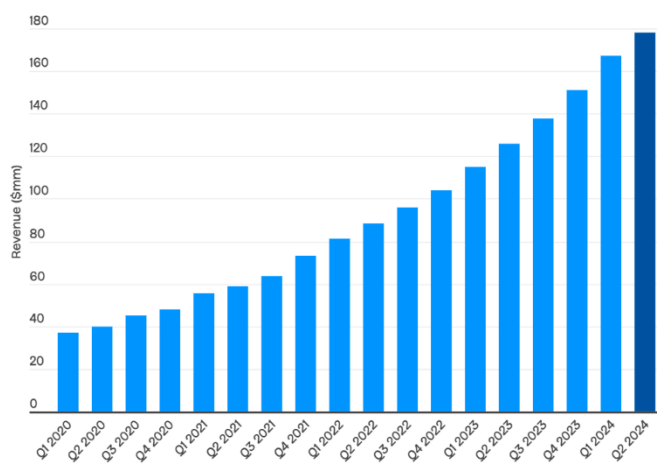


资料来源: Statista, 信达证券研发中心



**图 13：20Q1-24Q2Duolingo 付费订阅用户数（单位：百万人）**


资料来源: Statista, 信达证券研发中心

**图 14：20Q1-24Q2Duolingo 季度收入（单位：百万美元）**


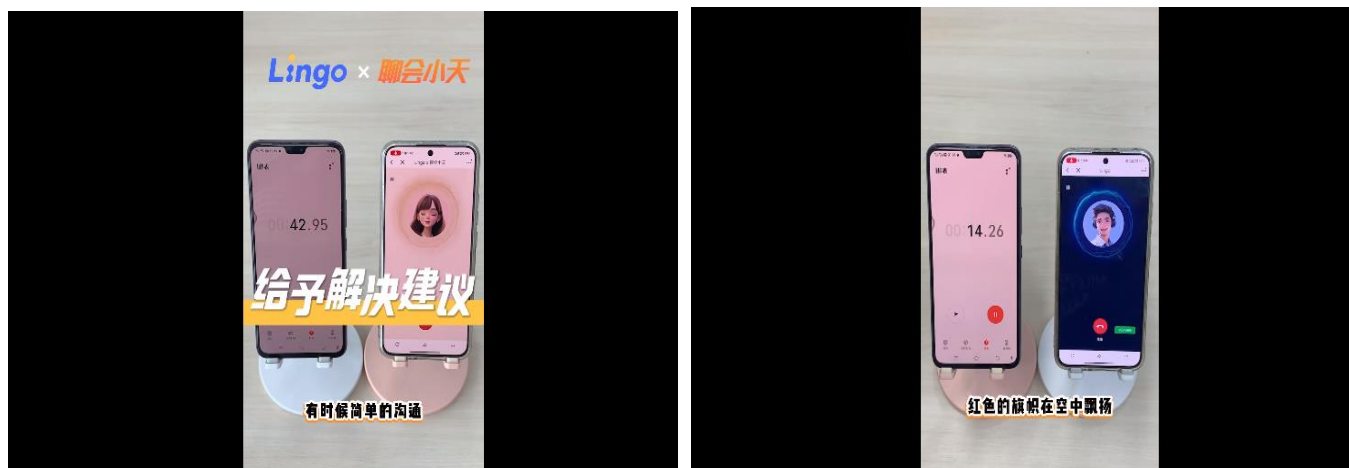
资料来源: Business of Apps, 信达证券研发中心

## 国内 AI+音频应用主要上市公司

### （一）汤姆猫

金科汤姆猫投资企业西湖心辰，其旗下心辰 Lingo 端到端语音大模型表现较优，在拟人语音交互方向较为领先。西湖心辰成立于 2021 年 7 月 9 日，已获得汤姆猫、蓝驰创投、凯泰资本、百度风投、西湖教育基金会可持续发展平台等知名机构数千万美金投资，作为国内首个端到端通用语音大模型，心辰 Lingo 在处理对话时可以直接理解语音，捕捉语气、节奏和情绪，并进行语音回复，减少了信息处理过程中的损失，让“机器”更懂人。未来，将优先在 **AI 销售**、**AI 心理咨询**、**AI 儿童陪伴**三大领域完成相关产品的落地。

心辰 Lingo 的落地场景例如：具身智能融合、心理疗愈、客户服务、儿童教育、岁月档案等。心辰 Lingo 可以扮演不同的助手角色，根据用户的需求和指令，提供个性化的语音服务。在心理健康应用中，心辰 Lingo 可以根据用户的情感状态，模拟“朋友”以及“亲人”的沟通方式，通过定制语音提供安慰和鼓励，帮助用户缓解压力和焦虑；在客户服务场景中，心辰 Lingo 卓越的即时响应能力，能确保在与用户交流时提供无任何体感延迟的语音服务。

**图 15：西湖心辰 Lingo AI 语音交互演示**


资料来源: 金科汤姆猫官方视频号, 信达证券研发中心

## （二）盛天网络：AI 音乐+社交——“给麦”

2023 年 7 月，盛天网络与超拟人大模型开发商聆心智能开展合作，拟将人工智能科技与自有娱乐场景有机结合，共同打造、探索 AI+游戏社交互动新场景。在 2023 年推出“给麦”APP。

自 2023 年上线以来，“给麦”APP 不断迭代升级，目前公司已在多模态模型中的情感分析、语音、文字识别等领域实现应用突破，涉及图像、视频、语音、文本等融合内容，推出了包括 AI 歌声进化 1.0、给麦 AI 实验室+AI 歌声进化 2.0、AI 全民会说唱、AI 写歌等多个版本和专题内容，迅速积累了大量忠实用户，并在 iOS 音乐类排行榜上取得了最高第六的佳绩；未来，“给麦”将继续在 AI 音乐创作方向深耕，提升音色克隆和音频/歌曲合成算法模型，为用户提供更优质的 AI 翻唱体验。同时，平台也在积极探索 AI 音乐创作新领域，目前已上线 AI 写歌 2.0、AI 作品广场 2.0 版本，支持写词、作曲、写歌等，后续也将探索更多内容如 AI 专辑封面、AI 音乐 MV 等，满足用户对音乐创作的需求。在 AI 音乐社交方面，“给麦”计划开放 AI 虚拟人分身创建，支持更多角色如 AI-CP、AI 疗愈师等，拓展线上元宇宙社交的可能性。

图 16：盛天网络给麦 AI 功能迭代



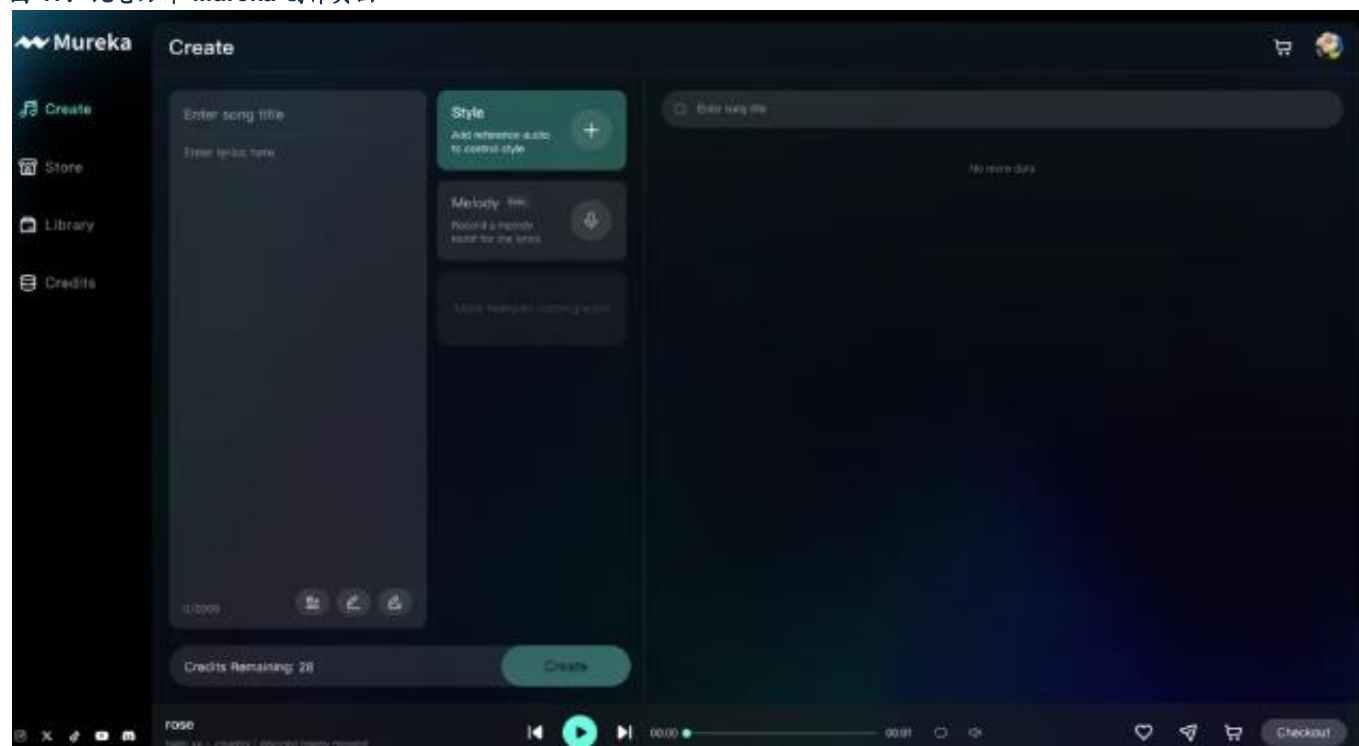
资料来源：盛天网络公告，信达证券研发中心

## （三）昆仑万维：

2024 年 8 月 14 日，昆仑万维正式发布全球首个 AI 流媒体音乐平台 Melodio，并同步推出 AI 音乐商用创作平台 Mureka。两款产品均搭载昆仑万维新款自研 DiT（Diffusion Transformer）架构音乐大模型 Skymusic 2.0，这是业内首个能够持续稳定生成特定风格歌曲的 AI 音乐大模型。相较于上一代模型，Skymusic 2.0 支持 500 字以上的歌词输入，生成长达 6 分钟双声道立体声歌曲。用户在 Melodio 音乐生成与播放过程中可以随时修改输入文案，改变音乐生成内容。根据 Statista 数据，2023 年全球纯流媒体音乐订阅收入为 193 亿美元。在全球音乐巨头 Spotify 平台上，近 50% 用户曾经收听过“心情/场景类”歌单，用户对于场景类的音乐需求强烈。

**商业模式：**Mureka 是一款革命性的 AI 音乐商用创作平台，专业艺术家与音乐爱好者均可在 Mureka 平台上创作专属音乐，并通过歌曲商店展示、收听、收藏、分享、下载，同时获得 AI 乐曲创作证明。用户还可以将满意的作品上架到 Mureka 平台进行出售，以此获得 AI 音乐带来的收益。在 Mureka 创作歌曲将消耗用户付费积分（credits）。歌曲创作完成后，用户可以下载创作证明，并申请在 Mureka 歌曲商店（Store）中进行展示出售。也可以付费购买歌曲的伴奏和分轨音频。用户在平台出售歌曲的收益，Mureka 会按既定比例进行抽佣。

图 17: 昆仑万维 Mureka 创作页面



资料来源: 昆仑万维公众号, 信达证券研发中心

### 三、投资建议

**AI 音频方面**, 在语音实时对话效果不断提升的基础上, 相比之前的虚拟陪伴、虚拟人物的语音互动, **OpenAI Realtime API** 更有望催生一大批基于纯语音交互的 **AI 应用**, 例如旅行规划、点餐、出游、语音学习、电商购物、教育等诸多方向。建议关注: 盛天网络、中文在线、汤姆猫、昆仑万维等;

**AI 视频方面**, 远不止于视频生成, 从 **AI 生成到 AI 工作流**, 一站式 **AI 视频生成+剪辑+故事创作**有望成为产业核心发展方向。**AI+视频**时代来临, 思考哪类公司存在商业化变现的可能性? 我们认为, 1) 一站式平台型公司, 如 Adobe、美图公司; 2) AI+视频技术头部服务商转型产品类公司, 如 Runway、商汤科技; 3) 视频剪辑类公司, 如快手; 4) 广告营销类公司, 如易点天下、蓝色光标、因赛集团、利欧股份; 5) UGC 社区类公司, 如 Bilibili; 6) 视频数据类公司, 如捷成股份、华策影视、视觉中国、中广天择; 7) IP 类公司, 如上海电影、阅文集团、汤姆猫、中文在线、果麦文化; 8) 探索 AI 视频工作流及其他创作方向类公司, 如博纳影业、超讯通信、柠萌影视。9) 其他建议关注猫眼娱乐、光线传媒、芒果超媒、万达电影等。

表 2: 相关上市公司估值表 (截至 2024.10.15)

分类	证券简称	总市值 (亿元)	归母净利润 (百万元)				市盈率		
			23A	24E	25E	26E	24E	25E	26E
一站式平台型	ADBE	15,840.1	53483.0	59276.0	65823.0	74153.0	26.7	24.1	21.4
	美图公司	104.7	368.3	548.5	787.9	1046.0	19.1	13.3	10.0
技术服务类	商汤-W	490.5	-6440.0	-3645.0	-2522.0	-1742.0	-	-	-
UGC 社区类	哔哩哔哩-W	608.0	-4822.32	-1734.0	-143.8	990.0	-	-422.9	61.4
视频剪辑类	快手-W	1,868.2	6396.0	15734.6	21711.7	27026.2	11.9	8.6	6.9
IP 类公司	阅文集团	248.5	804.9	1249.2	1400.7	1543.0	19.9	17.7	16.1
	上海电影	101.4	127.0	232.8	333.9	422.7	43.6	30.4	24.0
	汤姆猫	162.4	-864.6	200.0	300.0	400.0	81.2	54.1	40.6

广告营销类	易点天下	82.9	217.0	286.7	357.1	433.0	28.9	23.2	19.1
	蓝色光标	169.5	116.6	371.6	527.1	639.6	45.6	32.2	26.5
	因赛集团	59.5	41.6	/	/	/	/	/	/
	利欧股份	113.8	1966.0	/	/	/	/	/	/
视频数据训练类	华策影视	119.2	382.2	455.8	517.5	562.8	26.2	23.0	21.2
	视觉中国	86.7	145.6	169.2	197.9	224.1	51.3	43.8	38.7
	捷成股份	137.2	450.0	593.0	650.3	728.3	23.1	21.1	18.8
	中广天择	32.2	-8.7	/	/	/	/	/	/
AI 短剧/短片探索类	博纳影业	65.0	-552.6	314.2	498.9	576.7	20.7	13.0	11.3
	超讯通信	47.2	18.8	111.1	164.8	243.4	42.5	28.6	19.4
	柠萌影视	11.9	213.6	258.1	322.7	372.3	4.6	3.7	3.2
其他	光线传媒	232.0	417.8	1068.1	1212.0	1373.9	21.7	19.1	16.9
	万达电影	226.0	912.2	1338.7	1663.6	1931.3	16.9	13.6	11.7
	芒果超媒	438.5	3555.7	2040.4	2273.7	2524.2	21.5	19.3	17.4
AI 音频	汤姆猫	162.4	-864.6	200.0	300.0	400.0	81.2	54.1	40.6
	盛天网络	53.7	169.3	224.4	291.7	322.4	23.9	18.4	16.7
	昆仑万维	440.0	1258.4	864.7	962.0	1045.9	50.9	45.7	42.1
	中文在线	175.3	89.4	134.3	172.7	215.1	130.6	101.5	81.5

资料来源: iFind、Bloomberg, 信达证券研发中心

## 四、风险因素

**AI 大模型发展不及预期:** AI 大模型升级迭代速度减缓, 多模态大模型升级不及预期;

**AI 视频产品付费率提升不及预期:** AI+视频产品力较弱, 用户付费意愿较低影响公司现金流。



## 研究团队简介

**冯翠婷，信达证券传媒互联网及海外首席分析师**，北京大学管理学硕士，香港大学金融学硕士，中山大学管理学学士。2016-2021 年任职于天风证券，覆盖互联网、游戏、广告、电商等多个板块，及元宇宙、体育二级市场研究先行者（首篇报告作者），曾获 21 年东方财富 Choice 金牌分析师第一、Wind 金牌分析师第三、水晶球奖第六、金麒麟第七，20 年 Wind 金牌分析师第一、第一财经第一、金麒麟新锐第三。

**刘旺，信达证券传媒互联网及海外团队高级研究员**。北京大学金融学硕士，北京邮电大学计算机硕士，北京邮电大学计算机学士，曾任职于腾讯，一级市场从业 3 年，创业 5 年（人工智能、虚拟数字人等），拥有人工智能、虚拟数字人、互联网等领域的产业经历。

**李依韩，信达证券传媒互联网及海外团队研究员**。中国农业大学金融硕士，2022 年加入信达证券研发中心，覆盖互联网板块。曾任职于华创证券，所在团队曾入围 2021 年新财富传播与文化类最佳分析师评比，2021 年 21 世纪金牌分析师第四名，2021 年金麒麟奖第五名，2021 年水晶球评比入围。

**白云汉，信达证券传媒互联网及海外团队研究员**。美国康涅狄格大学金融硕士，曾任职于腾讯系创业公司投资部，一级市场从业 2 年。后任职于私募基金担任研究员，二级市场从业 3 年，覆盖传媒互联网赛道。2023 年加入信达证券研发中心，目前主要专注于美股研究以及结合海外映射对 A 股、港股的覆盖。

## 分析师声明

负责本报告全部或部分内容的每一位分析师在此申明，本人具有证券投资咨询执业资格，并在中国证券业协会注册登记为证券分析师，以勤勉的职业态度，独立、客观地出具本报告；本报告所表述的所有观点准确反映了分析师本人的研究观点；本人薪酬的任何组成部分不曾与，不与，也将不会与本报告中的具体分析意见或观点直接或间接相关。

## 免责声明

信达证券股份有限公司（以下简称“信达证券”）具有中国证监会批复的证券投资咨询业务资格。本报告由信达证券制作并发布。

本报告是针对与信达证券签署服务协议的签约客户的专属研究产品，为该类客户进行投资决策时提供辅助和参考，双方对权利与义务均有严格约定。本报告仅提供给上述特定客户，并不面向公众发布。信达证券不会因接收人收到本报告而视其为本公司的当然客户。客户应当认识到有关本报告的电话、短信、邮件提示仅为研究观点的简要沟通，对本报告的参考使用须以本报告的完整版本为准。

本报告是基于信达证券认为可靠的已公开信息编制，但信达证券不保证所载信息的准确性和完整性。本报告所载的意见、评估及预测仅为本报告最初出具日的观点和判断，本报告所指的证券或投资标的的价格、价值及投资收入可能会出现不同程度的波动，涉及证券或投资标的的历史表现不应作为日后表现的保证。在不同时期，或因使用不同假设和标准，采用不同观点和分析方法，致使信达证券发出与本报告所载意见、评估及预测不一致的研究报告，对此信达证券可不发出特别通知。

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，也没有考虑到客户特殊的投资目标、财务状况或需求。客户应考虑本报告中的任何意见或建议是否符合其特定状况，若有必要应寻求专家意见。本报告所载的资料、工具、意见及推测仅供参考，并非作为或被视为出售或购买证券或其他投资标的的邀请或向人做出邀请。

在法律允许的情况下，信达证券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并进行交易，并可能会为这些公司正在提供或争取提供投资银行业务服务。

本报告版权仅为信达证券所有。未经信达证券书面同意，任何机构和个人不得以任何形式翻版、复制、发布、转发或引用本报告的任何部分。若信达证券以外的机构向其客户发放本报告，则由该机构独自为此发送行为负责，信达证券对此等行为不承担任何责任。本报告同时不构成信达证券向发送本报告的机构之客户提供的投资建议。

如未经信达证券授权，私自转载或者转发本报告，所引起的一切后果及法律责任由私自转载或转发者承担。信达证券将保留随时追究其法律责任的权利。

## 评级说明

投资建议的比较标准	股票投资评级	行业投资评级
本报告采用的基准指数：沪深 300 指数（以下简称基准）；  时间段：报告发布之日起 6 个月内。	<b>买入</b> ：股价相对强于基准 15% 以上；	<b>看好</b> ：行业指数超越基准；
	<b>增持</b> ：股价相对强于基准 5%~15%；	<b>中性</b> ：行业指数与基准基本持平；
	<b>持有</b> ：股价相对基准波动在±5% 之间；	<b>看淡</b> ：行业指数弱于基准。
	<b>卖出</b> ：股价相对弱于基准 5% 以下。	

## 风险提示

证券市场是一个风险无时不在的市场。投资者在进行证券交易时存在赢利的可能，也存在亏损的风险。建议投资者应当充分深入地了解证券市场蕴含的各项风险并谨慎行事。

本报告中所述证券不一定能在所有的国家和地区向所有类型的投资者销售，投资者应当对本报告中的信息和意见进行独立评估，并应同时考量各自的投资目的、财务状况和特定需求，必要时就法律、商业、财务、税收等方面咨询专业顾问的意见。在任何情况下，信达证券不对任何人因使用本报告中的任何内容所引致的任何损失负任何责任，投资者需自行承担风险。