

证券研究报告

行业研究

行业事项点评

投资评级

上次评级

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

北京市西城区闹市口大街9号院1号楼

邮编: 100031

大模型进程焦点：OpenAI 春季发布会及谷歌开发者大会要点归纳

2024 年 5 月 16 日

事件: 北京时间 5 月 14 日凌晨, OpenAI 发布了最新模型 GPT-4o, GPT-4o (“o”代表“omni”)接受文本、音频和图像的任意组合作为输入, 并生成文本、音频和图像的任意组合输出。它可以在短至 232 毫秒的时间内响应音频输入, 平均为 320 毫秒, 与人类的响应时间相似。它在英语文本和代码上的性能与 GPT-4 Turbo 的性能相匹配, 在非英语文本上的性能显著提高。与现有模型相比, GPT-4o 在视觉和音频理解方面尤其出色。

GPT-4o 的文本和图像功能 5 月 14 日已经开始在 ChatGPT 中推出。OpenAI 正在免费套餐中提供 GPT-4o, 并向 Plus 用户提供高达 5 倍的消息限制, 未来几周将在 ChatGPT Plus 中推出新版语音模式 GPT-4o 的 alpha 版。开发人员现在还可以在 API 中访问 GPT-4o 作为文本和视觉模型。与 GPT-4 Turbo 相比, GPT-4o 速度提高 2 倍, 价格降低一半, 速率限制提高 5 倍。OpenAI 计划在未来几周内在 API 中向一小部分值得信赖的合作伙伴推出对 GPT-4o 新音频和视频功能的支持。

在 OpenAI 发布会上, 团队展示了 GPT-4o 的多个功能。包括实时语音交流能力, 几乎无延迟, 达到真人聊天的水平; 可以识别用户的语音情绪, 语音有情感; 可以使用不同语言(英语、意大利语)和多人同时交互; 在语音的同时可以与视频实时交互; 实时交流, 协助阅读代码, 指导编程; 实时视频聊天, 并识别情绪等。

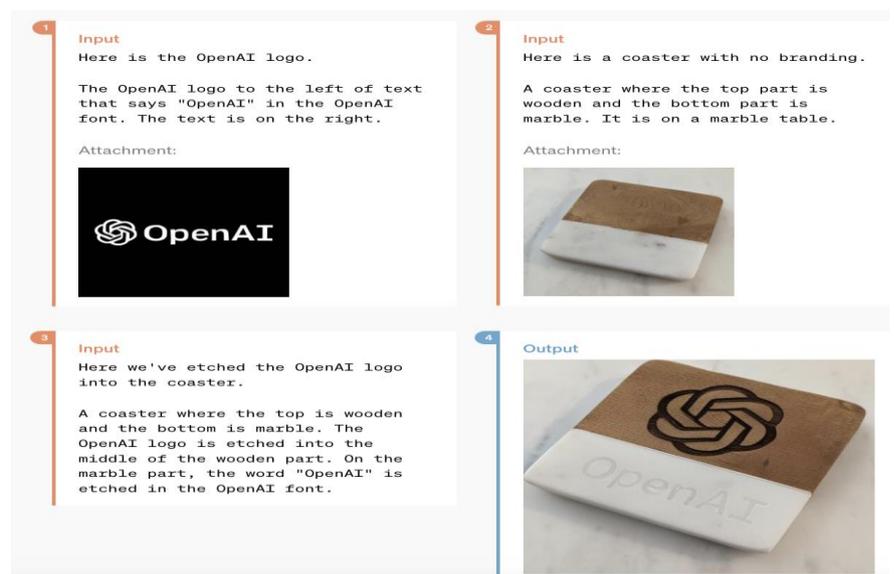
北京时间 5 月 15 日凌晨, Google I/O 开发者大会正式开幕。公司更新了 Gemini1.5Pro, 引入了 1.5Flash 较小型号的 Gemini 版本, 同时也增加了两个新的 Gemma 模型。Gemini 1.5 Pro: 普通版本配备了 100 万 token 上下文窗口, 而 Gemini 1.5 Pro 将在 Google AI Studio 或适用于 Google Cloud 客户的 Vertex AI 中的候补名单中拥有 200 万个代币上下文窗口。在智能代理领域, 谷歌展示了其最新的人工智能代理 Google Astra 项目, 可与用户实现多模态交互。

多模态领域, 谷歌推出全新视频生成模型 Veo, 可生成各种电影和视觉风格的高质量 1080p 分辨率视频, 时间可以超过一分钟; 升级了文生图新模型 Imagen3, 对图像细节的把控度更高; 发布了人工智能音乐生成模型 Lyria。以及在 AI 搜索功能、照片问询功能均实现了一定程度的升级迭代。

点评:

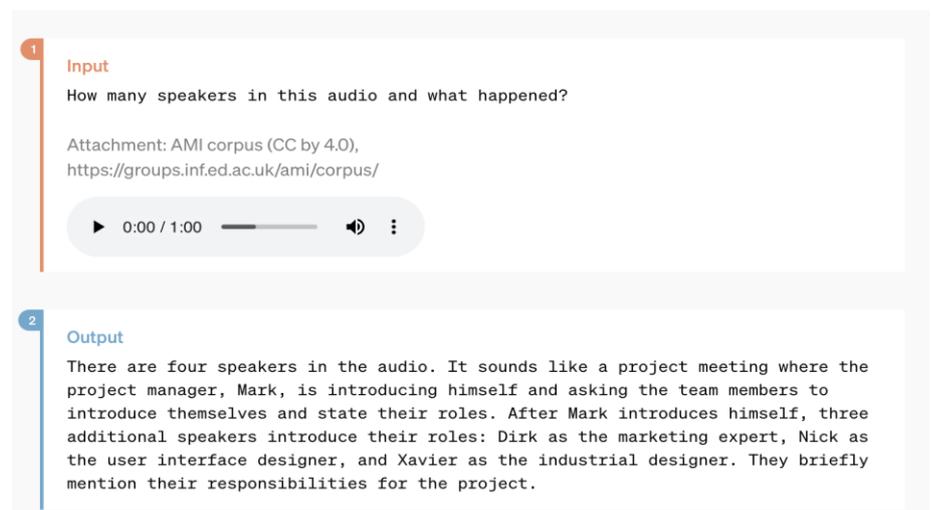
- **OpenAI 打造全新多模态融合、高拟人情感、实时低延迟交互的“个人超级 AI 语音助手”，推出 GPT-4o 全新旗舰人工智能模型。**根据传统基准测试，GPT-4o 在文本、推理和编码智能方面实现了 GPT-4 Turbo 级别的性能，同时，在多语言、音频和视觉功能上设置了新的高水位线。GPT-4o 能够接受文本、音频和图像的任意组合作为输入，并生成文本、音频和图像的任意组合输出，输入输出均由同一神经网络处理。同时，在用户情绪感知和语义理解上，GPT-4o 也展现得较为出色，用户可以在对话中随时打断，并且可以根据用户需要改变语音语调来满足用户的情绪变化，有望逐渐成为个人用户在终端设备上的具备强拟人化、可交互性高、高智商、高情商的私人助理管家。

图 1: GPT-4o 文本+图像输入



资料来源: OpenAI 官网, 信达证券研发中心

图 2: GPT-4o 文本+音频输入



资料来源: OpenAI 官网, 信达证券研发中心

图 3: 两个 GPT-4o 产品在自动对话并且唱歌



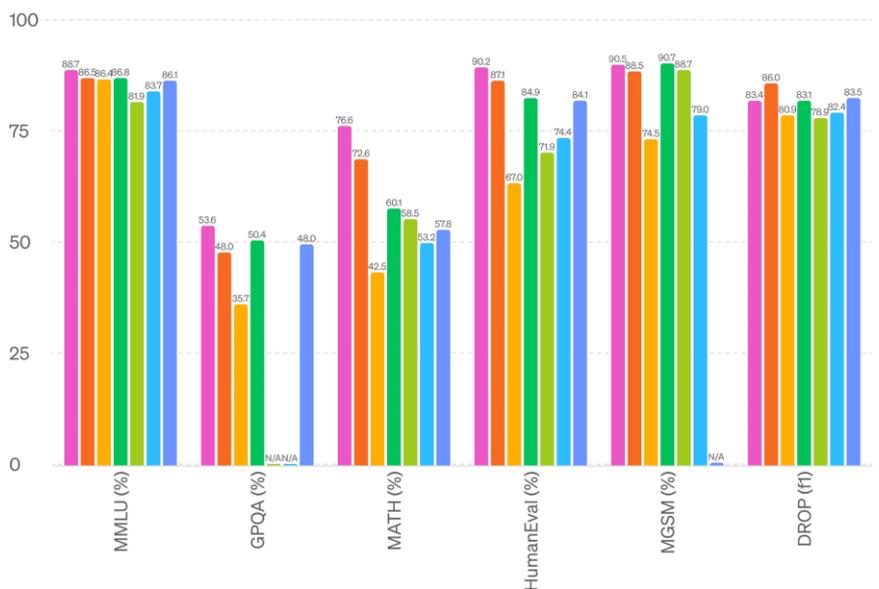
Two GPT-4os interacting and singing.

资料来源: OpenAI 官网, 信达证券研发中心

图 4: GPT-4o 文本评价

Text Evaluation

■ GPT-4o
 ■ GPT-4T
 ■ GPT-4 (initial release 23-03-14)
 ■ Claude 3 Opus
 ■ Gemini Pro 1.5
■ Gemini Ultra 1.0
■ Llama3 400b

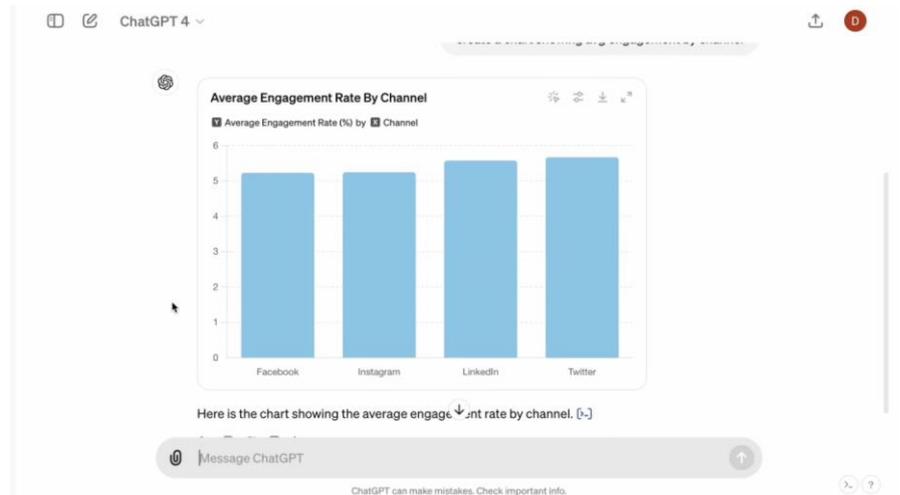


资料来源: OpenAI 官网, 信达证券研发中心

- 免费提供给所有用户使用, 付费用户享受 5 倍调用额度, 大模型本身的技术迭代带来的成本降低有望使个人或企业用户深度受益。每周都有超过一亿人使用 ChatGPT。OpenAI 将在未来几周开始向 ChatGPT 免费用户推出更多智能和高级工具。使用 GPT-4o 时, ChatGPT 免费用户现在可以访问以下功能: 体验 GPT-4 大模型的回答; 获取来自模型和互联网的回复; 分析数据并创建图表; 聊聊用户拍摄的图片; 上传

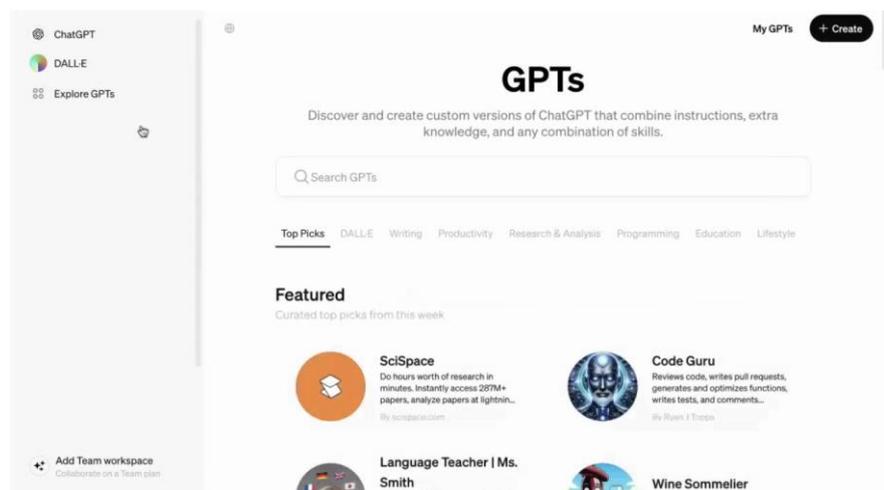
文件并提供帮助、总结、写作或分析；发现并使用 GPT 和 GPT 商店；使用 Memory 打造更有帮助的体验。

图 5: 上传数据文件, GPT-4o 进行数据分析



资料来源: OpenAI 官网, 信达证券研发中心

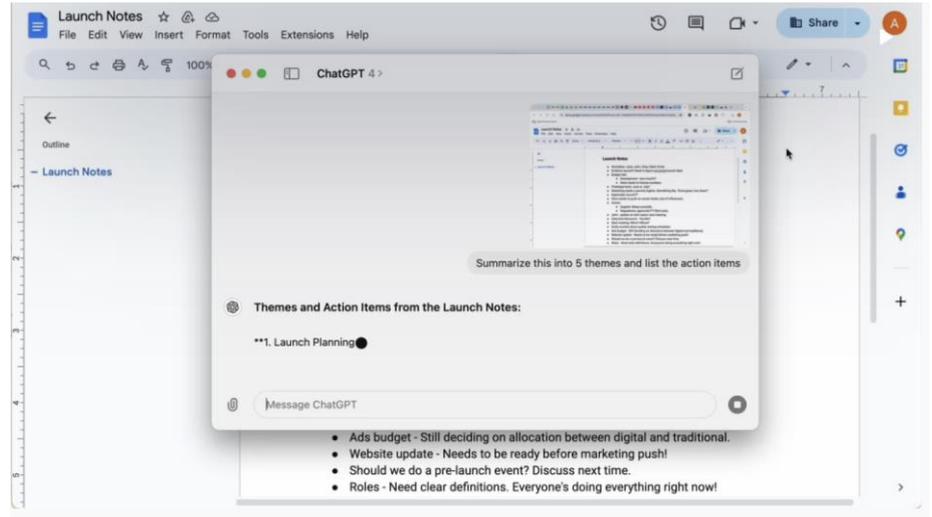
图 6: GPTs 和 GPTs 有关应用推荐



资料来源: OpenAI 官网, 信达证券研发中心

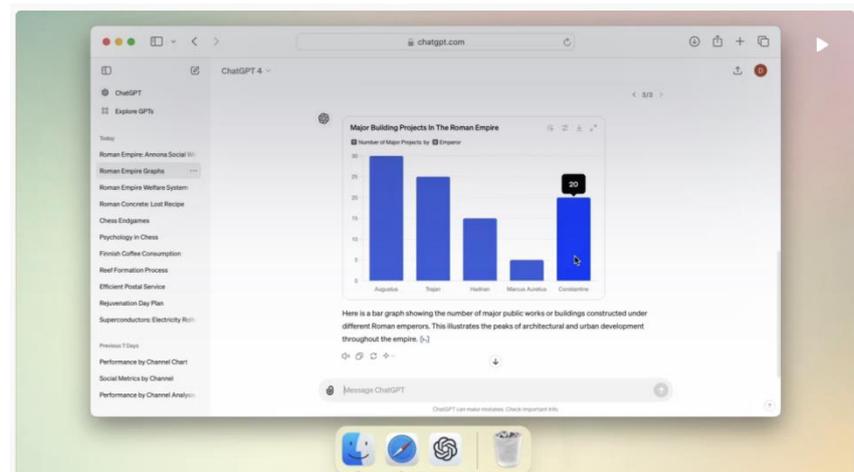
- **OpenAI 推出新的桌面应用程序, 简化用户的工作流程。**对于免费和付费用户, OpenAI 推出了适用于 macOS 的新 ChatGPT 桌面应用程序, 该应用程序旨在无缝集成到用户在计算机上执行的任何操作。通过简单的键盘快捷键 (Option+Space), 用户既可向 ChatGPT 提问, 还可以直接在应用程序中截取屏幕截图并进行讨论。同时, 用户可以直接从计算机与 ChatGPT 进行语音对话, 目前暂时提供 ChatGPT 发布时提供的语音模式, 新的音频和视频功能将在未来推出。同时, 从发布会当日开始, OpenAI 将向 Plus 用户推出适配于 macOS 的应用程序, 并将在未来几周内更广泛地提供该应用程序, Windows 版本计划在今年晚些时候推出。应用程序简化了 ChatGPT 的界面外观, 整体 UI 设计更加友好具备交互性。

图 7: 适用于 macOS 的新 ChatGPT 桌面应用程序



资料来源: OpenAI 官网, 信达证券研发中心

图 8: 应用程序简化了 ChatGPT 的界面外观, 整体 UI 设计更加友好具备交互性

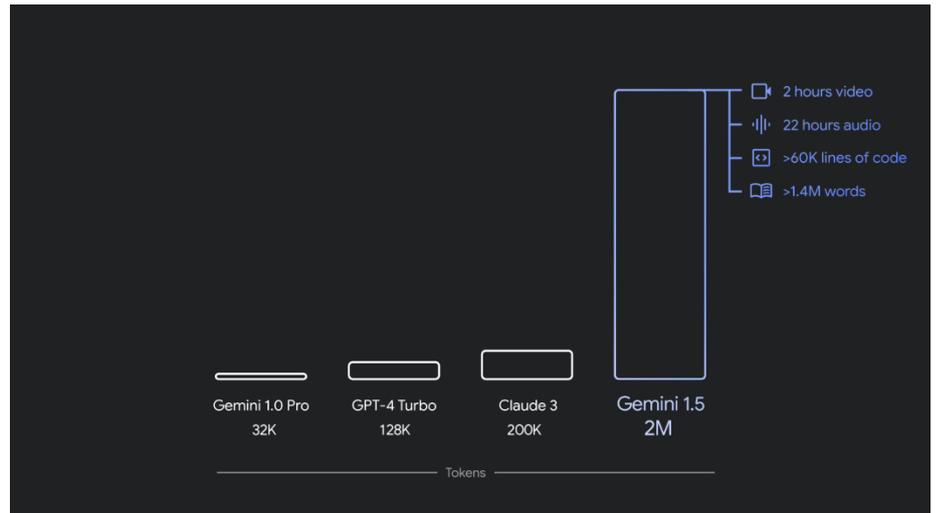


资料来源: OpenAI 官网, 信达证券研发中心

北京时间 5 月 15 日凌晨, Google I/O 开发者大会正式开幕, 在生成式人工智能 AI 竞争中, 从谷歌在本次开发者大会上发布的内容可以看出其在 AI 领域的积淀和步伐愈发稳健深厚。2023 年 12 月, 谷歌发布了 Gemini1.0; 2024 年 2 月, 谷歌发布了 Gemini1.5 Pro, 是第一个具有 100 万 tokens 的模型, 同时也推出了轻量级的开源模型 Gemma。在本次谷歌开发者大会上, 公司更新了 Gemini1.5Pro, 引入了 1.5Flash 较小型号的 Gemini 版本, 同时也增加了两个新的 Gemma 模型。

Gemini 1.5 Pro: 普通版本配备了 100 万 token 上下文窗口, 而 Gemini 1.5 Pro 将在 Google AI Studio 或适用于 Google Cloud 客户的 Vertex AI 中的候补名单中拥有 200 万个代币上下文窗口。Gemini 1.5Flash 针对较窄或高频的任务进行了优化, 达到较快的响应时间速度, 更具备成本效益, 同时上下文窗口也突破到了 100 万 token。

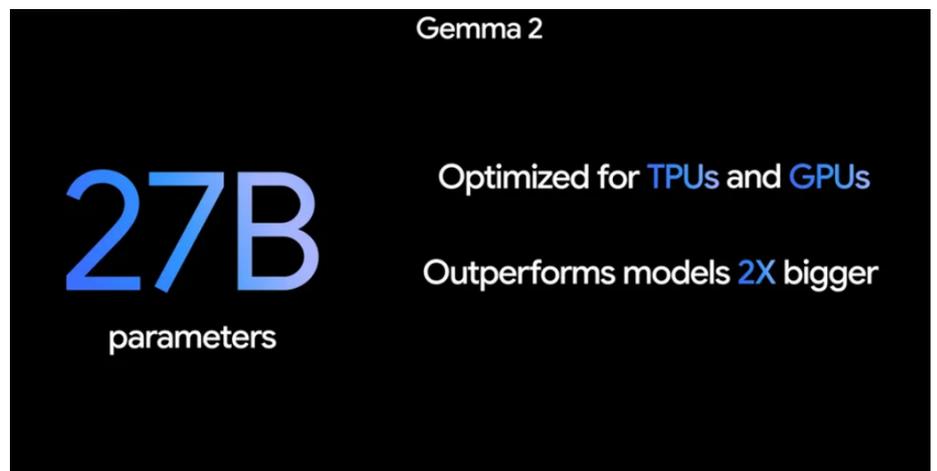
图 9: Gemini1.5 Pro 升级到 200 万 Tokens 上下文窗口



资料来源: 谷歌官网, 信达证券研发中心

谷歌即将于 2024 年 6 月推出 Gemma2 大模型, 模型参数量提高至 270 亿, 其性能优于一些尺寸超过其两倍的模型, 并且可以在 GPU 或 Vertex AI 中的单个 TPU 主机上高效运行。

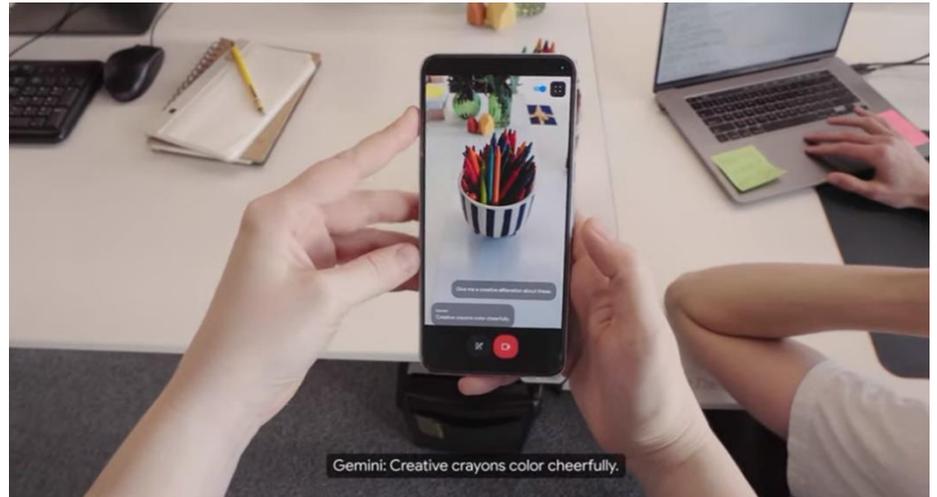
图 10: Gemma2 附带 27B 参数



资料来源: 谷歌官网, 信达证券研发中心

人工智能代理 Google Astra 项目: 作为智能体, 需要像人类一样理解和响应复杂且动态的世界, 并吸收并记住所看到和听到的内容, 以了解上下文并采取行动。它还需要主动、可教和个性化, 以使用户可以自然地与它交谈, 没有滞后或延迟。2024 年谷歌开发者大会上, DeepMind 团队展示了多模态理解和实时对话的能力, 希望构建一个在日常生活中有用的通用代理。和 OpenAI 刚发布的 AI 语音助手 GPT-4o 相比, Astra 在相应的延时、语气情绪的感知以及可打断性来看还有提升空间。

图 11: Google Astra Project 通用代理



资料来源: 谷歌官网, 信达证券研发中心

谷歌推出全新视频生成模型 Veo。可生成各种电影和视觉风格的高质量 1080p 分辨率视频，时间可以超过一分钟，在视频细节上更能清晰把握创作者的意图。提供了较强的创意控制水平，并理解“延时拍摄”或“风景空中拍摄”等电影术语。Veo 创建一致且连贯的镜头，因此人物、动物和物体在整个镜头中移动逼真。Veo 以谷歌多年的生成视频模型工作为基础，包括生成查询网络(GQN)、DVD-GAN、Imagen-Video、Phenaki、WALT、VideoPoet 和 Lumiere — 结合架构、缩放法则和其他新颖技术来提高质量和输出分辨率。

图 12: Google 全新视频生成模型 Veo



Veo 高质量视频生成功能的示例。所有视频均由 Veo 生成，未经修改。

资料来源: 谷歌官网, 信达证券研发中心

谷歌升级了文生图新模型 Imagen3。Imagen3 在图像生成的细节、光照等方面表现较优，和之前的文生图模型相比，其分散注意力的视觉伪影要少很多，Imagen3 可以更好理解自然语言、提示背后的意图并融合较长提示中的微小细节，同时更好控制图像一致性。

图 13: Google 升级了文生图模型 Imagen 3



资料来源: 谷歌官网, 信达证券研发中心

谷歌发布了人工智能音乐生成模型 **Lyria**。与 YouTube 以及一些出色的音乐家、歌曲作者和制作人合作去开发一套名为 Music AI Sandbox 的音乐 AI 工具。这些工具旨在为创造力开辟一个新的游乐场, 让人们从头开始创建新的乐器部分, 以新的方式改变声音等。

图 14: Google 发布 AI 音乐生成模型 Lyria



资料来源: 谷歌官网, 信达证券研发中心

同时, 谷歌针对谷歌搜索、Ask Photo 等功能均做出了一定程度的升级更新。从发布会后开始, 谷歌将在美国向用户推出经过全面改造的“AI 概览”的搜索体验, 即在搜索框下面单独展现出 AI Overview 的内容来给用户更好的服务响应, 满足用户提出新类型的问题、更长、更复杂的查询, 甚至使用照片进行搜索, 并获得网络所提供的最佳信息。以及将在 24 年夏天推出“询问照片”功能, 满足用户搜索照片信息的需求。

建议关注: 1) AI 陪伴+IP: 汤姆猫、上海电影、奥飞娱乐、华策影视、中文在线; 2) AI+搜索: 昆仑万维、三六零等; 3) AI+影视: 华策影视、慈文传媒、上海电影; 4) AI+IP (涉及版权、算料等): 芒果超媒、中广天择、

华数传媒、中文在线等；5) AI+出版/教育：南方传媒、皖新传媒等；6) AI+游戏：巨人网络、恺英网络、神州泰岳、三七互娱、姚记科技、盛天网络等；7) 特别提示：微软系/语音助手类标的，紫天科技、易点云、万兴科技、视觉中国、科大讯飞、商汤（港股）等。

风险提示：生成式人工智能模型进展迭代不及预期、AI应用进展不及预期

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。

凤超，信达证券传媒互联网及海外团队高级研究员，本科和研究生分别毕业于清华大学和法国马赛大学，曾在腾讯担任研发工程师，后任职于知名私募机构，担任互联网行业分析师。目前主要负责海外互联网行业的研究，拥有 5 年的行研经验，对港美股市场和互联网行业有长期的跟踪覆盖。主要关注电商、游戏、本地生活、短视频等领域。

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。	买入: 股价相对强于基准 15% 以上;	看好: 行业指数超越基准;
	增持: 股价相对强于基准 5%~15%;	中性: 行业指数与基准基本持平;
	持有: 股价相对基准波动在 ±5% 之间;	看淡: 行业指数弱于基准。
	卖出: 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。