



计算机

优于大市（维持）

证券分析师

陈涵泊

资格编号：S0120524040004

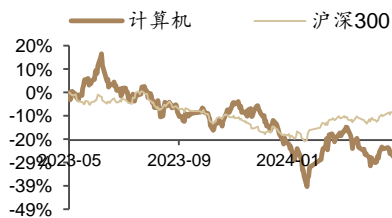
邮箱：chenhb3@tebon.com.cn

研究助理

王思

邮箱：wangsi@tebon.com.cn

市场表现



相关研究

- 《迎峰度夏+市场化推进，虚拟电厂热度有望持续提升》，2024.5.13
- 《DeeSeek-V2 定价普惠或激活 AI 应用发展》，2024.5.12
- 《特别国债或加速信创及行业信息化支出》，2024.5.12
- 《2024Q1 总结：大盘股营收较稳健，AI 景气兑现度高》，2024.5.9
- 《Vidu 直接对标 Sora，AI 时代不可忽视视频革命》，2024.5.1

GPT-4o：大模型风向标，OpenAI 重大更新

投资要点：

- **事件：**美国时间 5 月 13 日，OpenAI 迎来重大更新。1) 发布桌面版 APP 及更新 UI 后的 ChatGPT，用户使用更加简单与自然；2) 发布新的全能模型 GPT-4o，实现跨模态即时响应，相比 GPT-4 Turbo，性价比显著提升的同时向用户免费开放。
- **桌面版及新 UI 界面的 ChatGPT，实现极致的用户交互体验。** 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的计算机上执行的任何操作。通过简单的键盘快捷键（Option + Space），用户便可向 ChatGPT 提问，并支持直接在应用程序中截取屏幕截图进行讨论。与此同时，ChatGPT 支持与计算机直接语音对话，并在未来将推出新音频和视频功能。发布后，OpenAI 首先向 Plus 用户推出 macOS 应用程序，并将在未来几周内更广泛地提供该应用程序，同时计划在今年晚些时候推出 Windows 版本。此外，OpenAI 还全面改版了 UI 界面，旨在提供更加友好和更具有对话性的用户交互体验。
- **GPT-4o 实现毫秒级先进视觉音频理解，价格方面更具优势。** GPT-4o 为 OpenAI 最新的旗舰型号，不仅提供与 GPT-4 同等程度的模型能力，推理速度还更快，还能提供同时理解文本、图像、音频等内容多模态能力。此次发布会，GPT-4o 具有如下亮点：1) 毫秒级响应，实现即时语音对话。 GPT-4o 对音频输入的响应时间最短为 232 毫秒，平均为 320 毫秒，这与人类在对话中的响应时间相似；2) 更先进的视觉与音频理解能力，智能感知语气与语态。 GPT-4o 在对话中可随时打断，并可根据场景生成多种语调，同时支持视频对话；3) API 性价比显著提升。与 GPT-4 Turbo 相比，GPT-4o 速度提升 2 倍、成本却降低了 50%，且速率限制提升 5 倍。4) 免费提供给所有用户使用，付费用户则可以享受五倍的调用额度，团队和企业用户调用额度更高。
- **GPT-4o 为端到端多模态模型，刷新 SOTA 实现性能飞跃。**传统语音 AI 通常经过三步法（语音识别或 ASR、LLM 计划下一步语言、语音合成或 TTS）实现对话功能，在这过程中会丢失很多信息且不能判断情绪变化。GPT-4o 为跨模态端到端训练新模型，意味着所有输入和输出都由同一神经网络处理，成为真正的多模态统一模型，带来性能的飞跃提升。具体来看：1) 在传统基准测试中，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、音频和视觉功能上达到了新的高水位线；2) 在端到端多模态架构支持下，GPT-4o 多模态能力范围显著拓展，除了文本、图像等常用功能，GPT-4o 还支持 3D 物品合成、文本转字体等多样化功能。
- **投资建议。**我们认为，GPT-4o 是多模态大模型走向成熟的重要标志，藉此未来有望推动产生更多现象级 AI 应用，并进一步提升推理算力需求。建议关注 **AI 多模态**：万兴科技、虹软科技、当虹科技、中科创达、大华股份、海康威视、漫步者、萤石网络、汉仪股份、美图公司、云从科技等；**AI 算力**：云赛智联、思特奇、恒为科技、海光信息、寒武纪、景嘉微、工业富联、拓维信息、浪潮信息、四川长虹、神州数码等；**AI+办公**：金山办公、万兴科技、福昕软件、彩讯股份、金蝶国际、泛微网络、致远互联、鼎捷软件、汉得信息，用友网络等；**AI+教育/电商/医疗**：科大讯飞、佳发教育、鸥玛软件、盛通股份、值得买、焦点科技、小商品城、润达医疗、嘉和美康、创业慧康等。
- **风险提示：** AI 技术迭代不及预期、产品落地不及预期、AI 伦理风险等。

内容目录

1. 桌面版及新 UI ChatGPT，极致的用户交互体验	4
2. GPT-4o 实现毫秒级视觉音频理解，价格方面更具优势	5
3. 端到端多模态 GPT-4o，刷新 SOTA 性能飞跃	6
4. 投资建议	8
5. 风险提示	9

图表目录

图 1: 桌面版 ChatGPT 与用户屏幕截图进行讨论.....	4
图 2: 桌面版 ChatGPT 支持语音对话.....	4
图 3: ChatGPT 改版 UI 界面.....	4
图 4: GPT-4o 在与技术人员实时对话.....	5
图 5: GPT-4o 实时感知技术人员情绪.....	5
图 6: GPT-4o 通过视频通过解决图片中数学问题.....	5
图 7: GPT-4o 的 API 性价比显著提升.....	6
图 8: 传统语音 AI“语音转文本-问答-文本转语音”三步法模式.....	6
图 9: GPT-4o 文本推理相较 GPT-4 Turbo 有一定提升.....	7
图 10: 与 Whisper-v3 相比, GPT-4o 在多种语言的识别中表现优异.....	7
图 11: GPT-4o 在音频翻译表现上达到新的高水准.....	7
图 12: 在 M3Exam 基准测试中, GPT-4o 都比 GPT-4 强.....	8
图 13: GPT-4o 在视觉感知基准上实现了最先进的性能.....	8
图 14: GPT-4o 实现 3D 物品合成.....	8
图 15: GPT-4o 实现文本转字体.....	8

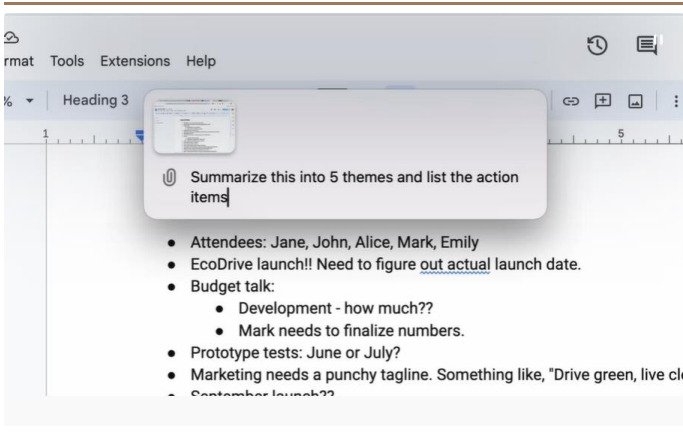
事件：美国时间 5 月 13 日，OpenAI 迎来两项重大更新。具体包括：1) 发布桌面版 APP 及更新 UI 后的 ChatGPT，用户使用更加简单与自然；2) 发布新的全能模型 GPT-4o，实现跨模态即时响应，相比 GPT-4 Turbo，性价比比显著提升的同时向用户免费开放。

1. 桌面版及新 UI ChatGPT，极致的用户交互体验

ChatGPT 发布桌面版 APP，支持与计算机语音对话，提升用户与模型交互体验。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的计算机上执行的任何操作。通过简单的键盘快捷键 (Option + Space)，用户便可向 ChatGPT 提问，并支持直接在应用程序中截取屏幕截图进行讨论。与此同时，ChatGPT 支持与计算机直接语音对话，并在未来将推出新音频和视频功能。

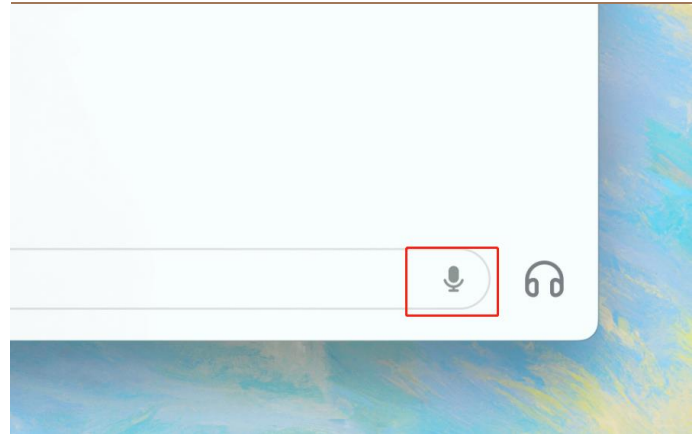
发布后，OpenAI 首先向 Plus 用户推出 macOS 应用程序，并将在未来几周内更广泛地提供该应用程序，同时计划在今年晚些时候推出 Windows 版本。

图 1：桌面版 ChatGPT 与用户屏幕截图进行讨论



资料来源：OpenAI 官网，德邦研究所

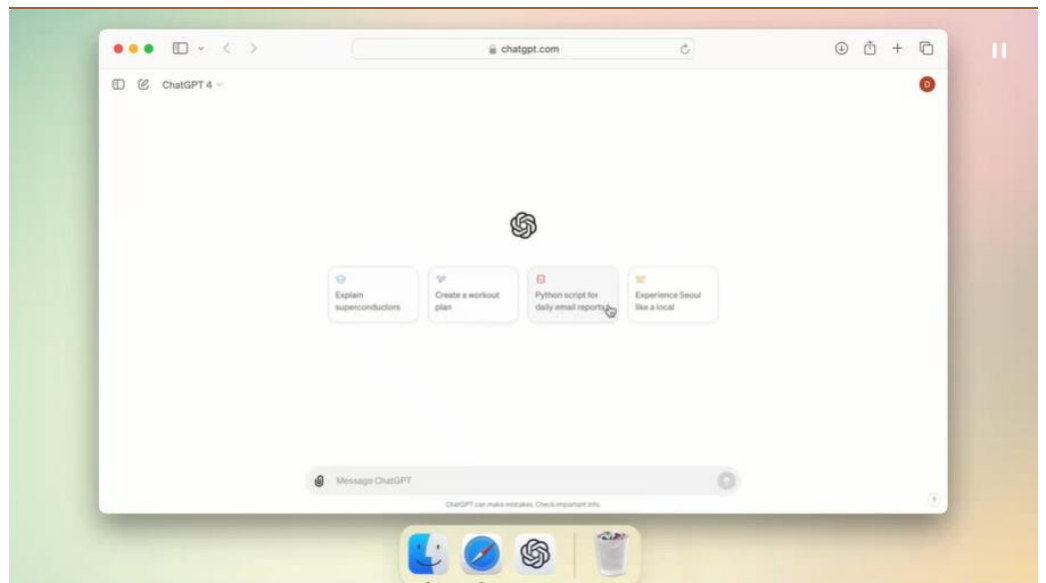
图 2：桌面版 ChatGPT 支持语音对话



资料来源：OpenAI 官网，德邦研究所

此外，OpenAI 还全面改版了 UI 界面，旨在提供更加友好和更具有对话性的外观，让它和用户的互动也变得更自然与简单。

图 3：ChatGPT 改版 UI 界面



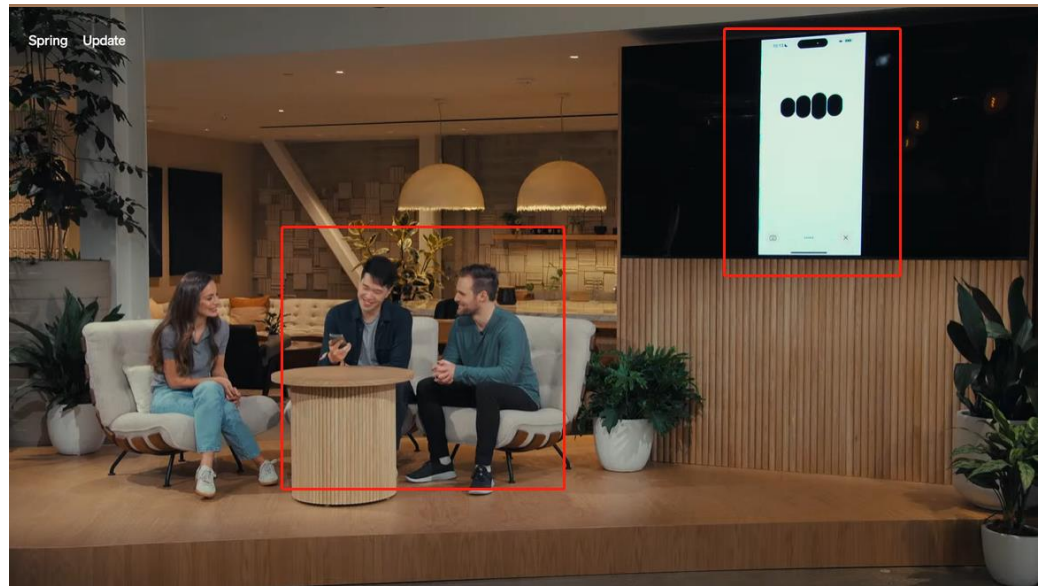
资料来源：OpenAI 官网，德邦研究所

2. GPT-4o 实现毫秒级视觉音频理解，价格方面更具优势

GPT-4o (“o”代表“omni”)为 OpenAI 最新的旗舰型号，不仅提供与 GPT-4 同等程度的模型能力，推理速度还更快，还能提供同时理解文本、图像、音频等内容的多模态能力。此次发布会，GPT-4o 具有如下亮点：

1) **毫秒级响应，实现即时语音对话。**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。使用语音模式与 ChatGPT 对话当中，GPT-3.5 与 GPT-4 平均延迟分别为 2.8s、5.4s，而 GPT-4o 对音频输入的响应时间最短为 232 毫秒，平均为 320 毫秒，这与人类在对话中的响应时间相似。GPT-4o 运行速度大大提升，迈向了更自然的人机交互。

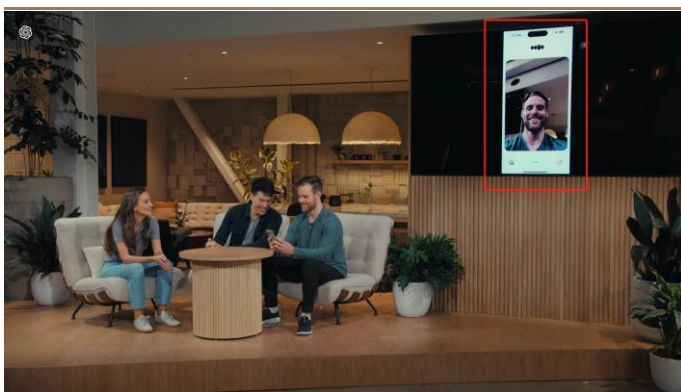
图 4: GPT-4o 在与技术人员实时对话



资料来源：OpenAI YouTube 官方账号，德邦研究所

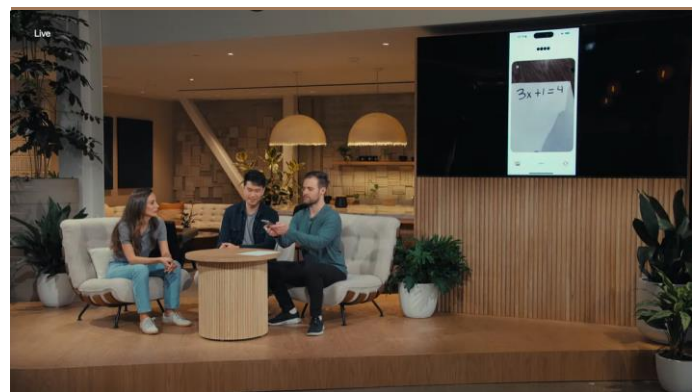
2) **更先进的视觉与音频理解能力，智能感知语气与语态。**与现有模型相比，GPT-4o 展现了出色的视觉和音频理解能力：首先，用户可在对话中随时打断；其次，可根据场景生成多种音调，带有人类般的情绪和情感；直接通过和 AI 视频通话让它在线解答各种问题。

图 5: GPT-4o 实时感知技术人员情绪



资料来源：OpenAI YouTube 官方账号，德邦研究所

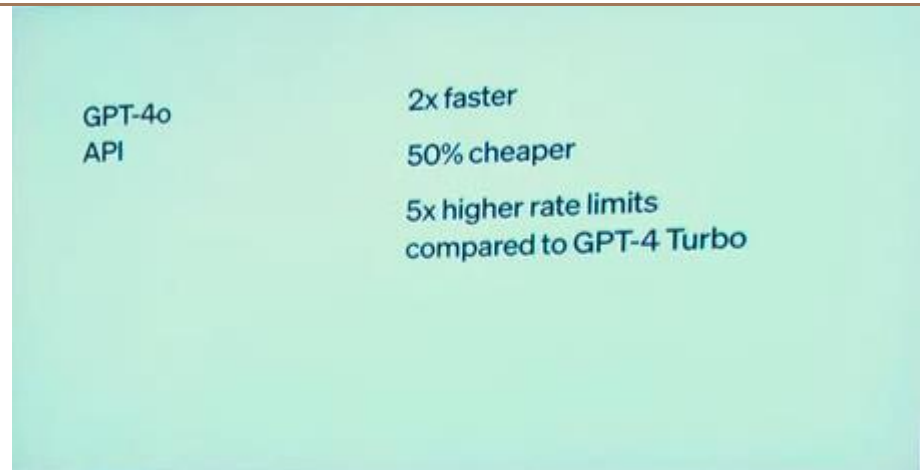
图 6: GPT-4o 通过视频通过解决图片中数学问题



资料来源：OpenAI YouTube 官方账号，德邦研究所

3) **API 性价比显著提升。**与 GPT-4 Turbo 相比，GPT-4o 速度提升 2 倍、成本却降低了 50%，且速率限制提升 5 倍。

图 7: GPT-4o 的 API 性价比显著提升



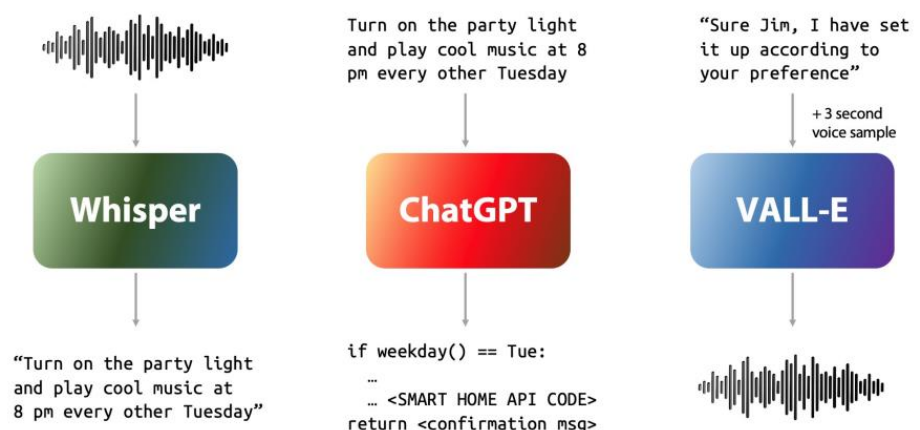
资料来源: OpenAI YouTube 官方账号, 德邦研究所

4) 免费提供给所有用户使用, 付费用户则可以享受五倍的调用额度, 团队和企业用户调用额度更高。OpenAI 将“为尽可能多的人提供先进的 AI 工具”作为使命之一, 每周有超 1 亿人使用 ChatGPT。公司将在未来几周内开始向 ChatGPT 免费用户推出更多智能和高级工具, 在使用 GPT-4o 时, ChatGPT 免费用户可免费访问如下功能: 体验 GPT-4 级别的智能; 从模型和网络获取响应; 分析数据并创建图表; 聊一聊用户的照片; 上传文件以获取摘要、写作或分析帮助; 使用 GPTs 和 GPT Store; 通过 Memory 构建更加有帮助的体验。

3. 端到端多模态 GPT-4o, 刷新 SOTA 性能飞跃

传统语音 AI 通常经过三步法实现对话功能, 在这过程中会丢失很多信息且不能判断情绪变化。三步法具体为: 1) 语音识别或 ASR: 音频到文本, 类似 Whisper; 2) LLM 计划下一步要说什么: 文本 1 到文本 2; 3) 语音合成或 TTS: 文本 2 到音频, 类似 ElevenLabs 或 VALL-E。GPT-4 便采用该模式, 在这过程中不仅响应速度更慢而且丢失了大量信息, 无法直接观察语调、多个说话者或背景噪音, 也无法输出笑声、歌唱或表达情感等。

图 8: 传统语音 AI “语音转文本-问答-文本转语音” 三步法模式



@DrJimFan

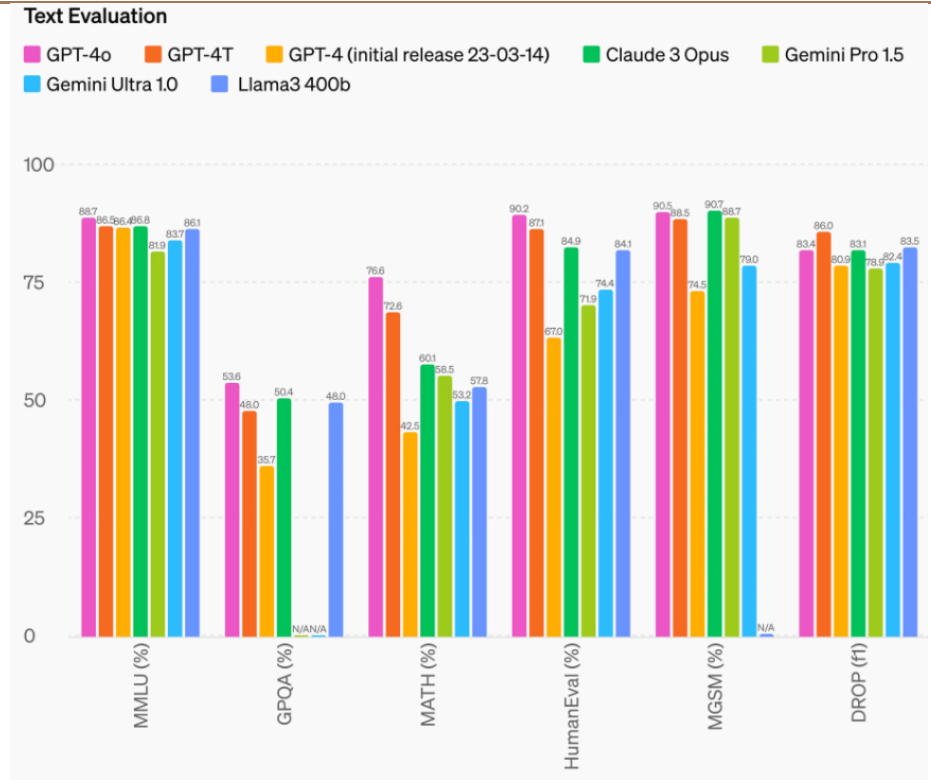
资料来源: 爱范儿微信公众号, 德邦研究所

GPT-4o 为跨模态端到端训练新模型, 意味着所有输入和输出都由同一神经网络处理, 成为真正的多模态统一模型, 带来性能的飞跃提升。除了响应速度更快之外, 在传统基准测试中, 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、音频和视觉功能上达到了新的高水位线。

文本推理：GPT-4o 在 0-shot COT MMLU（常识问题）上创下了 88.7% 的新高分。此外，在传统的 5-shot no-CoT MMLU 上，GPT-4o 创下了 87.2% 的新高分。相较于 GPT-4 Turbo 而言，GPT-4o 文本推理能力有一定提升。

图 9：GPT-4o 文本推理相较 GPT-4 Turbo 有一定提升

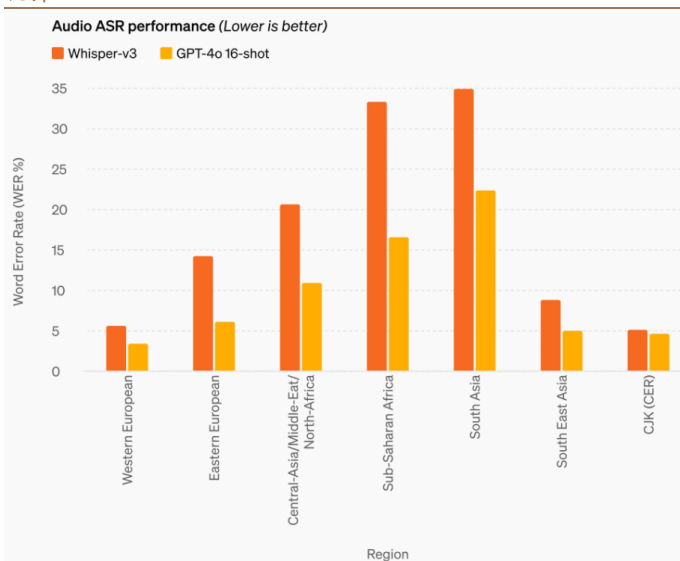


资料来源：OpenAI 官网，德邦研究所（注：Llama3 400b 还在训练中）

多语言识别：与 Whisper-v3 相比，GPT-4o 在多种语言的识别中表现优异，尤其是资源匮乏的语言。

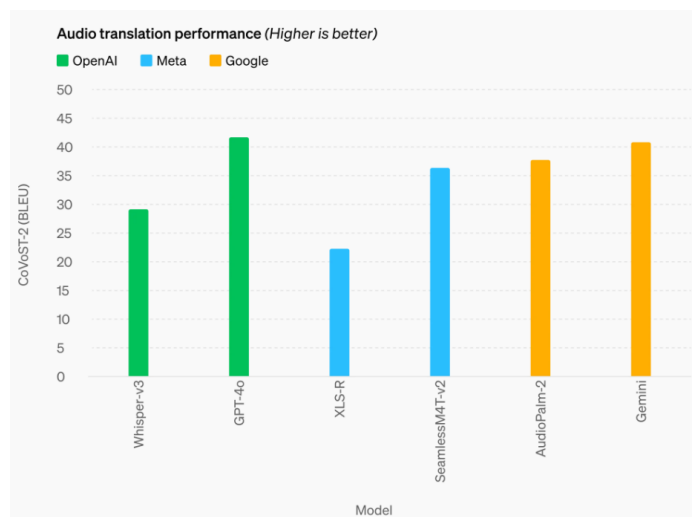
音频翻译：GPT-4o 在音频翻译表现上达到新的高水准，且在 MLS 基准测试中优于 Whisper-v3。

图 10：与 Whisper-v3 相比，GPT-4o 在多种语言的识别中表现优异



资料来源：OpenAI 官网，德邦研究所（注：越低值越优异）

图 11：GPT-4o 在音频翻译表现上达到新的高水准

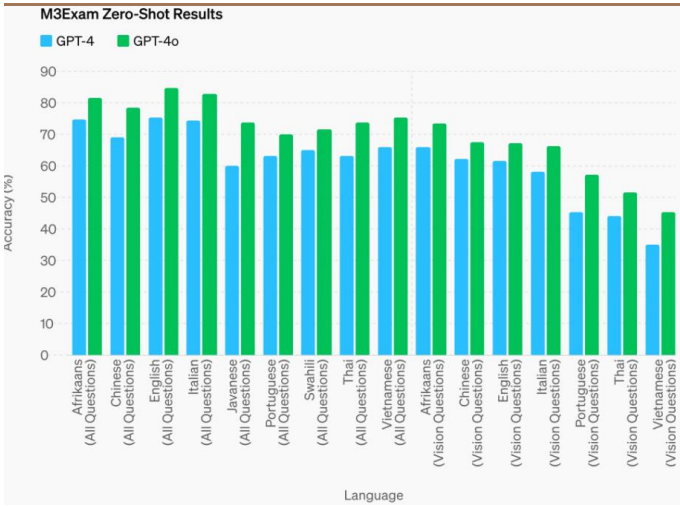


资料来源：OpenAI 官网，德邦研究所

M3Exam 测试: M3Exam 基准测试既是多语言评估也是视觉评估,由来自其他国家标准化测试的多项选择题组成,有时还包括图形和图表。在所有语言的基准测试中,GPT-4o 都比 GPT-4 更强。

视觉理解: GPT-4o 在视觉感知基准上实现了最先进的性能。具体来看,GPT-4o 在 MMMU 测试中分数达到 69.1,而 GPT-4 Turbo、Gemini 1.0 Ultra、Gemini 1.5 Pro、Claude Opus 分别为 63.1、59.4、58.5、59.4。

图 12: 在 M3Exam 基准测试中, GPT-4o 都比 GPT-4 强



资料来源: OpenAI 官网, 德邦研究所

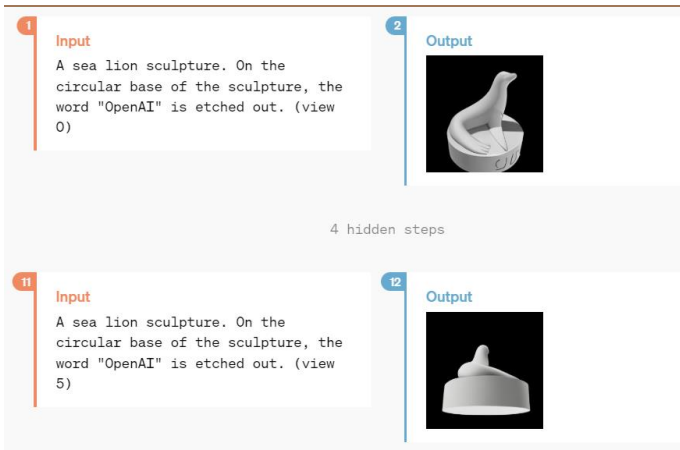
图 13: GPT-4o 在视觉感知基准上实现了最先进的性能

Eval Sets	GPT-4o	GPT-4T 2024-04-09	Gemini 1.0 Ultra	Gemini 1.5 Pro	Claude Opus
MMMU (%) (val)	69.1	63.1	59.4	58.5	59.4
MathVista (%) (testmini)	63.8	58.1	53.0	52.1	50.5
AI2D (%) (test)	94.2	89.4	79.5	80.3	88.1
ChartQA (%) (test)	85.7	78.1	80.8	81.3	80.8
DocVQA (%) (test)	92.8	87.2	90.9	86.5	89.3
ActivityNet (%) (test)	61.9	59.5	52.2	56.7	
EgoSchema (%) (test)	72.2	63.9	61.5	63.2	

资料来源: OpenAI 官网, 德邦研究所 (注: 所有视觉评估都是 0-shot)

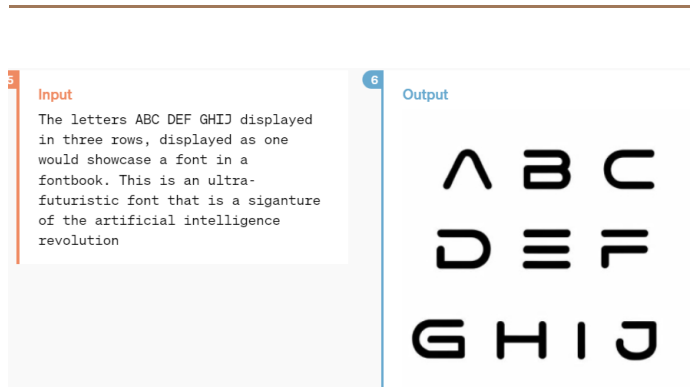
此外,在端到端多模态架构支持下,GPT-4o 多模态能力范围显著拓展。除了文本、图像等常用功能,GPT-4o 还支持 3D 物品合成、文本转字体等多样化功能。

图 14: GPT-4o 实现 3D 物品合成



资料来源: OpenAI 官网, 德邦研究所

图 15: GPT-4o 实现文本转字体



资料来源: OpenAI 官网, 德邦研究所

4. 投资建议

我们认为,GPT-4o 是多模态大模型走向成熟的重要标志,藉此未来有望推动产生更多现象级 AI 应用,并进一步提升推理算力需求。

建议关注:

AI 多模态: 万兴科技、虹软科技、当虹科技、中科创达、大华股份、海康威视、漫步者、萤石网络、汉仪股份、美图公司、云从科技等;

AI 算力: 云赛智联、思特奇、恒为科技、海光信息、寒武纪、景嘉微、工业

富联、拓维信息、浪潮信息、四川长虹、神州数码等；

AI+办公：金山办公、万兴科技、福昕软件、彩讯股份、金蝶国际、泛微网络、致远互联、鼎捷软件、汉得信息，用友网络等；

AI+教育/电商/医疗：科大讯飞、佳发教育、鸥玛软件、盛通股份、值得买、焦点科技、小商品城、润达医疗、嘉和美康、创业慧康等。

5. 风险提示

AI 技术迭代不及预期：技术为 AI 能力之基，技术发展不及预期将严重制约 AI 应用开发速度。

产品落地不及预期：产品落地不及预期将影响商业变现进程。

AI 伦理风险：AI 在人机耦合过程中可能产生伦理问题，带来相应风险。

信息披露

分析师与研究助理简介

陈涵泊：德邦证券计算机行业首席分析师，上海交通大学信息安全本科，电子与通信工程硕士，曾任职于中信证券研究部、天风证券研究所，多年计算机行业研究经验，具备成熟的计算机研究框架、自上而下产业前瞻视野，云计算领域深入研究。2022-2023年新财富最佳分析师入围（团队），2023年新浪金麒麟最佳分析师第五名（团队）。

王思：德邦证券计算机行业研究助理，湖南大学金融学学士，武汉大学金融学硕士，主要覆盖工业软件、云计算、网安等方向。

分析师声明

本人具有中国证券业协会授予的证券投资咨询执业资格，以勤勉的职业态度，独立、客观地出具本报告。本报告所采用的数据和信息均来自市场公开信息，本人不保证该等信息的准确性或完整性。分析逻辑基于作者的职业理解，清晰准确地反映了作者的研究观点，结论不受任何第三方的授意或影响，特此声明。

投资评级说明

	类别	评级	说明
1. 投资评级的比较和评级标准： 以报告发布后的6个月内的市场表现为比较标准，报告发布日后6个月内的公司股价（或行业指数）的涨跌幅相对同期市场基准指数的涨跌幅；	股票投资评级	买入	相对强于市场表现 20%以上；
		增持	相对强于市场表现 5%~20%；
		中性	相对市场表现在-5%~+5%之间波动；
		减持	相对弱于市场表现 5%以下。
2. 市场基准指数的比较标准： A股市场以上证综指或深证成指为基准；香港市场以恒生指数为基准；美国市场以标普500或纳斯达克综合指数为基准。	行业投资评级	优于大市	预期行业整体回报高于基准指数整体水平 10%以上；
		中性	预期行业整体回报介于基准指数整体水平-10%与 10%之间；
		弱于大市	预期行业整体回报低于基准指数整体水平 10%以下。

法律声明

本报告仅供德邦证券股份有限公司（以下简称“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。在任何情况下，本公司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。

本报告所载的资料、意见及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、价值及投资收入可能会波动。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。

市场有风险，投资需谨慎。本报告所载的信息、材料及结论只提供特定客户作参考，不构成投资建议，也没有考虑到个别客户特殊的投资目标、财务状况或需要。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。在法律许可的情况下，德邦证券及其所属关联机构可能会持有报告中提到的公司所发行的证券并进行交易，还可能为这些公司提供投资银行服务或其他服务。

本报告仅向特定客户传送，未经德邦证券研究所书面授权，本研究报告的任何部分均不得以任何方式制作任何形式的拷贝、复印件或复制品，或再次分发给任何其他人，或以任何侵犯本公司版权的其他方式使用。所有本报告中使用的商标、服务标记及标记均为本公司的商标、服务标记及标记。如欲引用或转载本文内容，务必联络德邦证券研究所并获得许可，并需注明出处为德邦证券研究所，且不得对本文进行有悖原意的引用和删改。

根据中国证监会核发的经营证券业务许可，德邦证券股份有限公司的经营经营范围包括证券投资咨询业务。