



计算机

Open AI 发布 GPT o1，壁仞科技拟 IPO

投资要点:

▶ 本周 AI 新闻速递

国际 AI: (1) 微软发布全球最强量子计算机。

国内 AI: (1) AI 芯片独角兽壁仞科技上市辅导启动。(2) 夸克首次推出泛功能助手 CueMe。

▶ Open AI 发布 GPT o1 模型

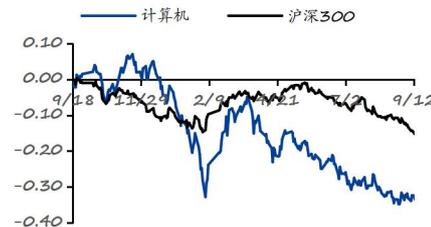
Open AI 于 9 月 13 日发布全新模型 o1。该模型分为两个版本: GPT o1-preview 和 GPT o1-mini。GPT o1-preview 推理能力更强,在推理数学、编程、科学等问题的能力上有显著提高,性能接近理化生博士水平; GPT o1-mini 则更小巧、专为代码生成优化的模型。

▶ 风险提示

市场需求不及预期,人工智能技术发展不及预期,政策发布不及预期,大模型商业落地不及预期的风险等。

强于大市 (维持评级)

一年内行业相对大盘走势



团队成员

分析师: 钱劲宇(S0210524040006)
QJY3773@hfzq.com.cn

相关报告

- 1、特斯拉公布路线图, FSD 或 2025 年一季度入华——2024.09.07
- 2、科大讯飞推出“星火极速超拟人交互”, Meta 推出 Imagine Yourself——2024.08.25
- 3、国内上线服务大模型突破 180 个, Grok 即将发布——2024.08.18



正文目录

1 本周 AI 新闻速递.....	3
1.1 国际 AI.....	3
1.2 国内 AI.....	3
2 Open AI 发布 GPT o1 模型.....	4
3 风险提示.....	6

图表目录

图表 1：o1 三项 benchmark 评测结果.....	4
图表 2：o1 学科类提升明显	5
图表 3：o1 表现随强化学习和思考时间增加而提升	5
图表 4：o1-preview 定价.....	5
图表 5：o1-mini 定价.....	5

1 本周 AI 新闻速递

1.1 国际 AI

(1) 微软发布全球最强量子计算机

据新皮层 NewNewThing, 9月10日, 微软宣布其已研发出目前全球性能最强的量子计算机。该计算机基于微软的 Azure Quantum 云计算平台搭建, 其思路是通过云环境来虚拟构建高度可靠的大规模量子比特可靠纠缠, 从而解决原本实际操控中等规模量子运算中难以避免的大量噪音。在过去, 量子计算机通过控制微观粒子的状态来储存信息和计算, 这些储存信息的粒子即量子比特, 但很难同时有效对其做大规模的操控运算。但通过与量子计算公司 Quantinuum 合作, 微软实现了 12 个高度可靠的大规模量子逻辑比特 (highly reliable logical qubits) 的创建与纠缠, 这是有史以来最高的纪录。这证明了和直接利用物理实在的量子比特来做运算相比, 通过云环境虚拟构建量子比特的路径更切实可行。基于这一结论, 微软正与量子计算机公司 Atom Computing 合作开发新一代量子计算机。

1.2 国内 AI

(1) AI 芯片独角兽壁仞科技上市辅导启动

据新皮层 NewNewThing, 9月11日, 中国证监会公布, 上海壁仞科技股份有限公司已办理辅导备案登记, 拟首次公开发行股票并上市, 辅导券商为国泰君安。目前, 壁仞科技已完成 B 轮融资, 总融资额超过 50 亿元, 投资方包括格力创投、中国平安、桂园创投、启明创投、IDG 资本、中信证券等机构。胡润研究院发布的《2024 全球独角兽榜》显示, 壁仞科技的估值为 155 亿元, 排在全球 495 位。壁仞科技成立于 2019 年, 主营业务是开发原创性通用计算体系, 建立高效的软硬件平台, 同时在智能计算领域提供一体化的解决方案。2022 年 8 月, 公司发布首款基于自主原创架构开发的通用 GPU 芯片 BR100, 声称其峰值算力是英伟达当时最先进的计算芯片 A100 的 3 倍多, 面向云端人工智能训练和推理领域。目前, 壁仞科技基于该芯片推出了壁砺 100、壁砺 104 和海玄服务器等数据中心, 主要部署在大型数据中心, 主要客户有中国移动、中国电信, 以及商汤科技、国网智能电网研究院有限公司等。

(2) 在推出多个垂直领域 AI 助手后, 夸克首次推出泛功能助手 CueMe

据新皮层 NewNewThing, 9月12日, 夸克发布基于夸克大模型开发的智能对话助手 CueMe, 并且还在导航栏中间放置了 AI 助手的入口, 通过这一入口, 用户可以快速访问包括 CueMe、健康助手、学习助手在内的多位 AI 聊天机器人。其中, CueMe 产品负责人表示, 这款助手的特点在于辅助写作。经过大量文学语料预训练后, CueMe 支持生成包括小说、诗歌、研究报告在内的多种文学体裁, 并且它最多可以生成 2 万字的长文本内容。例如, 用户可以让 CueMe 根据莫言过往的小说风格, 创作一份关于「辣子鸡」菜品的口播稿, CueMe 生成的文本中会提及莫言的魔幻现实



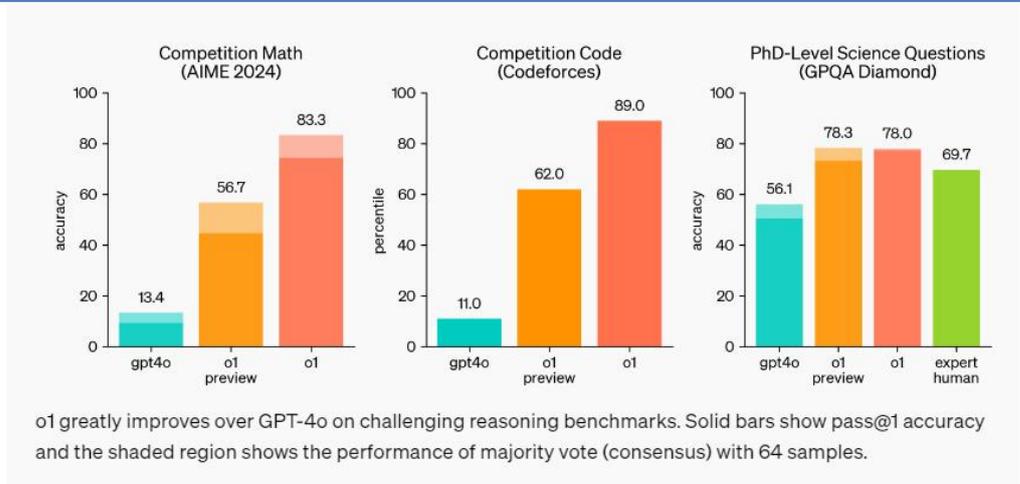
主义，而且会把热油锅中放入花椒、干辣椒爆炒比喻成「置身于一片火红的辣椒海」。在开发了多款 AI 助手后，夸克也开始有意识地将这些助手归拢到一起，并且给予它们更多的曝光。在这个助手界面，用户还可以创建 AI「小伙伴」，为它选择头像、名称，以及设定角色基础信息。

2 Open AI 发布 GPT o1 模型

Open AI 于 9 月 13 日发布全新模型 o1。该模型分为两个版本：GPT o1-preview 和 GPT o1-mini。GPT o1-preview 推理能力更强，在推理数学、编程、科学等问题的能力上有显著提高，性能接近理化生博士水平；GPT o1-mini 则更小巧、专为代码生成优化的模型。

推理能力上，o1-preview 在多个 benchmark 上表现出色。在美国数学奥林匹克（AIME 2024）预算赛中，其精确度达到 56.7%，GPT-4o 则仅有 13.4%。在竞争型问题中，o1-preview 排名 62%，GPT-4o 则仅有 11%。而在博士等级科学问题上（GPQA），o1-preview 达到 78.3% 的准确度，GPT-4o 为 56.1%，人类专家则为 69.7%。在正式版 o1 中，这几项能力仍会有不同程度的提升。

图表 1: o1 三项 benchmark 评测结果

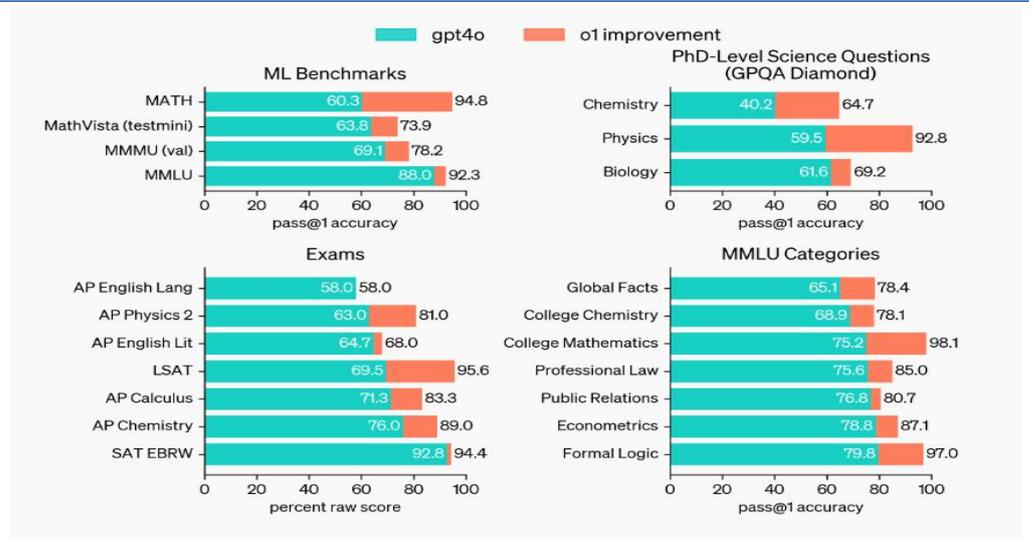


数据来源：OpenAI，华福证券研究所

与 GPT-4o 相比，o1 在各个学科上也有了不同程度的进步。特别是博士物理上，o1 准确度达到 92.8%，而 GPT-4o 仅有 59.5%。



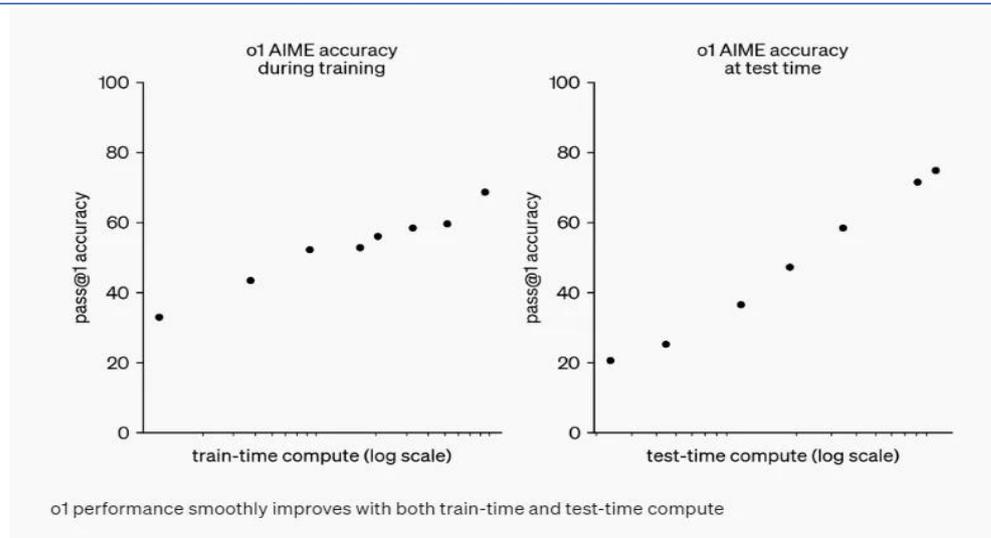
图表 2: o1 学科类提升明显



数据来源: OpenAI, 华福证券研究所

OpenAI 发现, 随着更多的强化学习 (训练时计算) 和更多的思考时间 (测试时计算), o1 的性能持续提高。

图表 3: o1 表现随强化学习和思考时间增加而提升



数据来源: Open AI, 华福证券研究所

目前, Plus 用户以及 Team 用户可以直接使用 o1 模型, 通过 API 访问的用户则需要达到 5 级 API 使用量 (消费 1000 美元且超过 1 个月的付费用户) 才能够使用。同时, OpenAI 对于模型的用量也明确限制: 目前 Plus 和 Team 用户每周仅能给 o1-preview 发送 30 条消息, 给 o1-mini 发送 50 条消息。5 级 API 用户则限制 20RPM 的用量。定价上, o1-preview 输入/输出价格分别为 \$15/\$60 每百万 tokens; o1-mini 输入/输出价格则为 \$3/\$12 每百万 tokens, 对应 4o 及 4o mini 来看, 价格均有数倍的提升。

图表 4: o1-preview 定价

图表 5: o1-mini 定价



OpenAI o1-preview

o1-preview is our new reasoning model for complex tasks that require broad general knowledge. The model has 128K context and an October 2023 knowledge cutoff.

[Learn about o1-preview >](#)

Model	Pricing
o1-preview	\$15.00 / 1M input tokens \$60.00 / 1M output tokens
o1-preview-2024-09-12	\$15.00 / 1M input tokens \$60.00 / 1M output tokens

来源：Open AI，华福证券研究所

OpenAI o1-mini

o1-mini is a fast, cost-efficient reasoning model tailored to coding, math, and science use cases. The model has 128K context and an October 2023 knowledge cutoff.

[Learn about o1-mini >](#)

Model	Pricing
o1-mini	\$3.00 / 1M input tokens \$12.00 / 1M output tokens
o1-mini-2024-09-12	\$3.00 / 1M input tokens \$12.00 / 1M output tokens

来源：Open AI，华福证券研究所

3 风险提示

市场需求不及预期，人工智能技术发展不及预期，政策发布不及预期，大模型商业落地不及预期的风险等。



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并注册为证券分析师，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。

一般声明

华福证券有限责任公司（以下简称“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。

本报告的信息均来源于本公司认为可信的公开资料，该等公开资料的准确性及完整性由其发布者负责，本公司及其研究人员对该等信息不作任何保证。本报告中的资料、意见及预测仅反映本公司于发布本报告当日的判断，之后可能会随情况的变化而调整。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。本公司不保证本报告所含信息及资料保持在最新状态，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。

在任何情况下，本报告所载的信息或所做出的任何建议、意见及推测并不构成所述证券买卖的出价或询价，也不构成对所述金融产品、产品发行或管理人作出任何形式的保证。在任何情况下，本公司仅承诺以勤勉的职业态度，独立、客观地出具本报告以供投资者参考，但不就本报告中的任何内容对任何投资做出任何形式的承诺或担保。投资者应自行决策，自担投资风险。

本报告版权归“华福证券有限责任公司”所有。本公司对本报告保留一切权利。除非另有书面显示，否则本报告中的所有材料的版权均属本公司。未经本公司事先书面授权，本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。未经授权的转载，本公司不承担任何转载责任。

特别声明

投资者应注意，在法律许可的情况下，本公司及其本公司的关联机构可能会持有本报告中涉及的公司所发行的证券并进行交易，也可能为这些公司正在提供或争取提供投资银行、财务顾问和金融产品等各种金融服务。投资者请勿将本报告视为投资或其他决定的唯一参考依据。

投资评级声明

类别	评级	评级说明
公司评级	买入	未来 6 个月内，个股相对市场基准指数涨幅在 20%以上
	持有	未来 6 个月内，个股相对市场基准指数涨幅介于 10%与 20%之间
	中性	未来 6 个月内，个股相对市场基准指数涨幅介于-10%与 10%之间
	回避	未来 6 个月内，个股相对市场基准指数涨幅介于-20%与-10%之间
	卖出	未来 6 个月内，个股相对市场基准指数涨幅在-20%以下
行业评级	强于大市	未来 6 个月内，行业整体回报高于市场基准指数 5%以上
	跟随大市	未来 6 个月内，行业整体回报介于市场基准指数-5%与 5%之间
	弱于大市	未来 6 个月内，行业整体回报低于市场基准指数-5%以下

备注：评级标准为报告发布日后的 6~12 个月内公司股价（或行业指数）相对同期基准指数的相对市场表现。其中 A 股市场以沪深 300 指数为基准；香港市场以恒生指数为基准，美股市场以标普 500 指数或纳斯达克综合指数为基准（另有说明的除外）

联系方式

华福证券研究所 上海

公司地址：上海市浦东新区浦明路 1436 号陆家嘴滨江中心 MT 座 20 层

邮编：200120

邮箱：hfjys@hfzq.com.cn