

被质疑的 OpenAI

本报记者 宋婧

最近,长期盘踞在人工智能 (AI) 舆论场中心的 OpenAI 烦心事有点儿多: 先是对标谷歌搜索的搜索引擎 SearchGPT 在演示中翻车,紧接着又经历了新一轮高层人事地震。屋漏偏逢连夜雨,特斯拉CEO马斯克再度将 OpenAI 告上法庭,指控其违反了“开发 AI 技术以造福人类”的创始使命。还有消息称 OpenAI 正面临 50 亿美元的亏损,现金流支撑不了一年。一系列事件引发业界质疑: 风光无限的 OpenAI 到底是怎么了?



盈利能力被质疑

作为崛起速度最快的 AI 初创企业之一, OpenAI 的盈利能力一直是业界关注焦点。有数据显示, OpenAI 今年运营总成本可能将达 85 亿美元,主要分为推理成本、训练成本和人工成本三大块。其中,推理和训练成本分别高达 40 亿美元和 30 亿美元。然而根据此前已公开信息推测, OpenAI 全年收入预估在 35 亿美元至 45 亿美元之间。也就是说, OpenAI 今年或面临高达 50 亿美元的巨额亏损。

实际上,以 OpenAI 的发展规模与速度而言,运营成本高并不意外。众所周知,训练一个 AI 大模型的成本非常高。AI 初创公司 Anthropic 首席执行官 Dario Amodei 指出,目前类似 GPT-4o 的模型训练成本约为 1 亿美元,而正在开发的 AI 大模型训练成本可能高达 10

亿美元。而且随着技术的进步和模型规模的扩大,训练成本呈指数级增长。Dario Amodei 预测,未来三年内, AI 大模型的训练成本将上升至 100 亿美元甚至 1000 亿美元。

除了在 AI 大模型上的投入, OpenAI 在服务器、GPU 等硬件以及电费上的投入,都是消耗性成本。根据 Digital Information World 发布的最新报告,数据中心为训练 AI 模型产生的能耗将为常规云工作的三倍。数据显示, OpenAI 训练 GPT-3 耗电为 1.287 吉瓦时,大约相当于 120 个美国家庭 1 年的用电量。而这仅仅是训练 AI 模型的前期电力,仅占模型实际使用时所消耗电力的 40%。

与高额的投入形成鲜明对比的是 OpenAI 的盈利表现。在最初的“尝鲜期”过后, ChatGPT 的用户访

问量正在持续下降。根据数据公司 Similarweb 的统计,从去年 5 月到今年 1 月这 8 个月里, ChatGPT 的总访问量有 5 个月的时间都在下降。今年 1 月 ChatGPT 的总访问量对比去年 5 月时的峰值,已经下降了 11%。与此同时, ChatGPT 移动端 App 的用户增长也在快速下滑,总用户数甚至比移动社交 APP Snapchat 的季度新增用户还要少。

究其原因,对于大多数普通用户来说, ChatGPT 本质上仍只是一个“聊天机器人”而已,尚未找到更有价值的应用场景。猎户星空董事长傅盛分析称,对于 OpenAI 而言,现在这么烧钱的情况下,用户量增长就是它的核心命脉,一旦不增长,融资下笔钱的时候估值不能再上一个台阶,就会很麻烦。

为了寻找新盈利点, OpenAI 开

OpenAI 今年运营总成本可能将达 85 亿美元,主要分为推理成本、训练成本和人工成本三大块。

始尝试做搜索引擎,结果 SearchGPT 还未推出就在公开演示中翻车。“就目前看到的产品形态来讲, OpenAI 的搜索产品并没有对其他搜索引擎产品形成碾压,形成 SO-TA (state-of-the-art, 最佳表现)。发布这个产品,可能会让用户有一点增长,但不会翻倍增长。”昆仑万维董事长兼 CEO 方汉坦言。

此前, OpenAI 还计划自研 AI 芯片,以解决其所依赖的 AI 芯片短缺以及成本高昂问题。而在芯片领域,英伟达已主导了全球 95% 的 AI 训练领域市场。业内人士判断, OpenAI 开发出能媲美英伟达服务器芯片的可能性很小,需要多年打磨才可能有成果。可见,无论是搜索引擎,还是芯片,想要从早已巨头林立的成熟市场中分得一杯羹谈何容易。

随着“蜜月期”热情消退, OpenAI 与微软试探性地把触角探入了彼此的地盘。

与微软的关系被质疑

近日,微软在最新年度报告中赫然将合作伙伴 OpenAI 列为其在 AI 产品、搜索、新闻广告等领域的竞争对手,引发广泛关注。尽管双方都声称合作关系并未改变,但业内人士普遍认为,这一变化标志着两家公司的关系进入新阶段。

OpenAI 与微软的联盟曾一度被科技圈认为是“天作之合”。作为 OpenAI 最大投资方和独家云服务供应商,微软在 OpenAI 身上的投入已超 130 亿美元,为后者的技术研发提供了充足的“弹药”。而 OpenAI 也投桃报李地将 ChatGPT 嵌入微软云服务、搜索引擎以及 Office 全家桶等一系列核心产品与服务中,帮助微软成为了这波 AI 浪潮的引领者。

然而,随着“蜜月期”热情消退, OpenAI 与微软试探性地把触角探入了彼此的地盘,这也让二者的关系发生了一些微妙的变化。

OpenAI 突然进军搜索引擎市场,让微软始料未及,也彻底点燃了这轮“战火”。OpenAI 最新推出的这款名为 SearchGPT 的生成式 AI 搜索引擎原型,底座是 GPT-4 系列 AI 模型。而微软旗下的搜索引擎产品 New Bing 也在基于 GPT-4 进行 AI 升级。

作为互联网流量重要入口,搜索历来是兵家必争之地。尽管谷歌盘踞搜索市场龙头地位多年,微软却始终没有放弃,反而投入了大量物力、财力和人力试图把 Bing 推上舞台。整合了 GPT-4 的 New Bing 更是被微软寄予厚望。对于微软而言, OpenAI 推出 SearchGPT 的行为无异于“偷塔”。不顾合作关系,直接将 OpenAI 列为竞争对手足以说明微软的态度。

与此同时,云服务也成了双方角力的“战场”。外媒称,与 OpenAI 展

开合作后,微软的云服务 Azure OpenAI 可以直接调用 OpenAI 模型,包括 ChatGPT、Codex 以及 DALL·E,其客户无须通过 OpenAI 就能在聊天机器人、搜索引擎等产品中使用这些 AI 大模型。微软的一份内部文件透露, Azure 销售人员告诉客户,微软能比 OpenAI 提供更多、更安全的服务。这种说辞确实帮助微软云服务吸引了更多客户,比如营销软件初创公司 Touchcast 就因为安全性和性能保证,才选择了通过 Azure 接入相关的大模型服务。

OpenAI 则试图通过微软云服务触达更多大客户,希望与其建立直接联系,甚至采用延迟授予微软大模型接入权限的方式来制造“时间差”,从而与客户达成合作。比如 GPT-4 语言模型发布后, Duolingo 和 Stripe 等公司已经直接向 OpenAI 付费之后,微软的云服务才获得了

GPT-4 的接入权。

值得一提的是,在大模型方面, OpenAI 也并非微软的唯一选择。据悉,微软正在研发一款名为 MAI-1 的最新 AI 大模型。微软员工表示, MAI-1 是一款全新的大语言模型,拥有约 5000 亿个参数,在性能上或能与谷歌的 Gemini 1.5、Anthropic 的 Claude 3 和 OpenAI 的 GPT-4 等知名大模型相匹敌。

“与微软的复杂关系对 OpenAI 而言,既是生命线也是潜在的威胁。”业内专家 Edward Zitron 分析称,“OpenAI 在很大程度上依赖于微软提供的云计算资源和资金支持,这可能影响 OpenAI 的独立性和自主决策能力。它需要巧妙地平衡这种关系,确保能够从微软那里获得支持,同时确保在技术共享和市场销售方面的独立性,同时减少对单一合作伙伴的依赖。”

高层震荡不断,加剧了业界对这家明星 AI 创企运营稳定性和发展前景的担忧。

公司稳定性被质疑

高层震荡往往会对公司运营的连续性与稳定性造成影响。 ChatGPT 走红以来, OpenAI 已经历了两轮令人瞩目的高层人事变动。去年年底, OpenAI 灵魂人物 Sam Altman 被董事会开除,而 Altman 本人半小时之前才得到通知。几小时后, OpenAI 董事长、联合创始人 Greg Brockman 公开宣布,自己也已经辞职。后来,两位核心人物又戏剧性地加盟了微软,短短几日后再次重返 OpenAI 担任原职,才结束了此轮风波。

而仅仅过去了半年, OpenAI 再次因高层集体出走引发热议。此轮风波涉及三位核心高管,分别是总裁 Greg Brockman (宣布长期休假)、联合创始人 John Schulman (转投 OpenAI 的强力竞争对手 Anthropic) 以及产品副总裁 Peter Deng。

值得一提的是,经历了此轮变动之后, OpenAI 最开始的 11 位联合创始人中,只剩下 Sam Altman 和负责 OpenAI 语言和代码的 Wojciech Zaremba 还留在岗位上。此外,还有多位知名度较高的安全负责人也纷纷退出了公司,其中几位公开声称公司将新产品的迭代置于安全之上。

高层震荡不断加剧了业界对这家明星 AI 创企运营稳定性和发展前景的担忧。而事实证明,这一担忧并非无的放矢。 OpenAI 最新发布的 GPT-4o 从产品创新性来看,并没太多实质性飞跃,模型能力和 GPT-4 没有本质区别,哪怕是和 GPT 的老对手 Anthropic 的 Claude 3 相比也没有拉开差距。近期更有大量用户反映,该模型出现了“惰性”的特点,其回应速度变慢、回答

问题敷衍且拒绝回答并中断对话等现象频现。 OpenAI 解释称是由于长时间未对 GPT 模型进行更新所致,引发诸多质疑。

有观点认为, GPT-4o 的发布更像是一种抢跑,展现一种领先的姿态,也是维持热度刺激购买需求的一种方式。也有观点认为,业内对 OpenAI 的期望值过高,除非 GPT-5“震撼”登场,否则很难满足大家的期待。

然而,被千呼万唤的 GPT-5 却在不断延迟发布。 OpenAI 近日确认,今年的开发者大会 (DevDay) 期间不会发布其下一代大模型 GPT-5。 Altman 公开表示,开发 GPT-5 还需要一些时间,目前 OpenAI 持乐观态度,但还有很多工作要做。他坦言, GPT-5 相关的技术仍处于早期阶段,存在数据和算法方面的

问题。

此外,在视频模型方面, Sora 最近也没有新消息传出。 AI 搜索方面,最新推出的 SearchGPT 剑指谷歌、微软,却在公开演示中出现低级错误,遭受群嘲。何时能够正式上线尚未可知。一系列的迹象显示,曾经“遥遥领先”的 OpenAI 似乎变“慢”了。

“表面看似似乎是高层震荡影响了 OpenAI 前进的速度,实际上 OpenAI 变‘慢’的核心原因在于路线选择引发的分歧。”一位在 AI 大厂工作多年的从业者对记者说道。一边是顶着盈利风险专注于实现 AGI (通用人工智能) 梦想,另一边是不惜一切代价加速商业化, OpenAI 似乎想从这两条路上寻找一个平衡点。但现实远比理想残酷,留给 OpenAI 的时间不多了。

2024 世界机器人大会

将于 8 月 21 日举行

本报讯 记者王伟报道: 记者从 8 月 13 日举办的 2024 世界机器人大会新闻发布会上获悉, 2024 世界机器人大会将于 8 月 21 日至 25 日在京举行, 同期还将举办 2024 世界机器人博览会, 届时将有 20 余款人形机器人整机亮相本届博览会现场。

人形机器人是近年来讨论度最高的机器人品类。20 余款人形机器人整机将亮相本届博览会现场, 创历届之最。这 20 余款人形机器人各具特色, 有的能跑跳越障, 有的能执行精巧操作, 还有的能表演才艺。

据介绍, 大会将举办 6 场主论坛及 25 场专题论坛, 416 位国内外嘉宾将畅谈机器人领域的前沿发展趋势。同期举办的 2024 世界机器

人博览会吸引了 ABB、KUKA、SMC、TESLA 在内的 160 余家国内外机器人头部厂商、知名公司、行业新锐参展, 60 余款新品将在展会上首发亮相。本届博览会首次设立前沿创新展区, 部分高校、科研院所将集中展示一批实验研发阶段的前沿创新成果, 如六足导盲机器人、飞行机器人等, 为机器人未来应用创新方向提供思路。

据介绍, 2024 世界机器人大会北京锦标赛也将同步举办, 作为机器人界的“奥林匹克”, 锦标赛集专业性、互动性、观赏性于一体, 旨在不断激发机器人行业的科技研发潜力。将有来自全球十余个国家和地区的 7000 余支赛队、 13000 余名参赛选手现场对决。

我国首个机器人多模态模型

及双臂协作系统亮相



本报讯 记者王伟报道: 记者近日从成都人形机器人创新中心获悉, 中国首个机器人多模态模型 RRMM (Raydicularious Robot Multimodal Model) 及双臂协作系统 RTACS (Raydicularious Two-Arm Cooperation System) 正式发布, 上述成果推动了我国人形机器人核心技术的突破。

此次成都人形机器人创新中心发布的多模态模型能使机器人理解推理抽象的语义指令, 并调度双臂协

作系统执行任务, 让人形机器人从“预设摆拍”“遥控操作”进化到拥有更强的“大脑”和自主执行能力。

据了解, 多模态 (即多种异构模态) 数据协同推理, 能融合两种或两种以上不同感知来源, 进行综合推理, 并完成复杂任务。多模态应用于人形机器人领域, 使人形机器人能融合图像、语义、力感知、环境感知等多种因素, 综合判断、生成任务并执行任务, 也是人形机器人具有自主思考能力的关键核心技术。

银河麒麟操作系统

发布 AI 版本

本报讯 近日, 在中国操作系统产业大会上, 国产操作系统银河麒麟发布了首个人工智能版本, 这标志着中国在自主操作系统领域实现了新的突破。

据了解, 操作系统是计算机之魂, 承接上层软件生态与底层硬件资源, 为 AI 算法、模型与应用的运行提供支撑环境, 在 IT 国产化中发挥重要作用。目前, 国产操作系统的安全性、稳定性已有跨越式发展, 实现了生态数量的“百万级”跃升, 形成了包括办公、经营、业务等一系列成熟应用场景。

据介绍, 此次发布的银河麒麟桌面操作系统 AI 版通过多项技术创新实现了人工智能与操作系统的

深度融合, 是我国首款国产操作系统和人工智能技术深度融合的产品, 具备人工智能集成能力、智能化功能、高效能计算等特点。该版本还可以让离线状态下的计算机也可具有人工智能能力, 填补了国产操作系统长期以来在这方面能力研发的空白。

当前, 人工智能技术已经成为全球信息产业技术争夺的新高峰。这为我国操作系统的发展带来新机遇。数据显示, 2023 年, 我国平台软件市场高速增长, 规模达 816.6 亿元, 同比增长 17.4%。我国操作系统市场增速进一步加快, 高达 23.2%。

(张维佳)

汉王科技推出智能新品

内置最新一代 NPU

本报讯 近日, 汉王科技发布三款新品——首款搭载 8 核处理器的智能手写电纸本 N10 Pro、全面升级的扫描王 APP 和汉王语音王 APP。

N10 Pro 手写电纸本内置最新一代 NPU, 其能提供 6TOPS AI 算力, 并支持大模型架构相关算子。此外, 该产品还配备业界领先的 300PPI 快刷墨水屏、自研高端数字电磁笔以及 6500mAh 超大容量电池等顶级配置。

N10 Pro 标配汉王自制的数字

式电磁笔 B5。B5 采用自研无线无源免充电技术, 内置自研专用低功耗 ASIC 芯片, 提笔即写, 流畅无迟, 同时提供专业的 8192 级精细压感, 笔触可随力度细腻变化。N10 Pro 也是汉王科技首款搭载 Android 14 操作系统的墨水屏设备, 相比上一代安卓系统, Android 14 在性能、电池、数据隐私、安全性以及兼容性等方面都做了优化, 让用户在操作体验、安全性、性能和多样化等方面表现更加出色。

(智文)