

GPT-4 Turbo亮相,使用成本降低,GPT商店即将上线

## OpenAI生态野心显现:开发者大会将成新的AI“科技春晚”

本报记者 曲忠芳 北京报道

北京时间11月7日凌晨,ChatGPT推出还不到一年,OpenAI就举行了自己的首次开发者大会。在短短45分钟的时间里,OpenAI首席执行官Sam Altman宣布推出最新的大模型GPT-4

Turbo。turbo的意思是涡轮增压器,用以实现汽车的高性能化,而GPT-4 Turbo从文本长度、模型控制、知识库、多模态、定制化、高速率、价格下降等方面对GPT大模型进行了性能的大幅升级。

更为重要的是,继今年3月

推出ChatGPT插件功能,实现成功“联网”之后,在此次开发者大会上,OpenAI宣布将在11月底推出GPT商店,支持用户通过自然语言定制自己的GPT,上传至GPT应用商店后可获得收入分成。OpenAI鼓励企业使用其技术来构建基于

AI技术和聊天机器人ChatGPT的人工智能体(AI Agent)。这些举动被业内视为OpenAI对ChatGPT商业化的持续探索,更是其打造自身生态的“野心”显现。

根据OpenAI官方公布的数据,从2022年11月30日发布至

今,基于ChatGPT的API(应用程序编程接口)开发者大约200万人,超过92%的《财富》500强公司都在使用,每周大约有1亿活跃用户。

GPT-4 Turbo的发布引起了业内的广泛关注,不少媒体、评论员、创业者在评价时甚至使用了

“王炸”“震撼”“恐惧”等颇具夸张色彩的词来形容。抛开这些外在的修饰,OpenAI的新动态对于业界可能会带来哪些影响?对于创业者来说究竟是冲击,还是机遇?《中国经营报》记者对此采访了来自产业界、投资界等的多方人士。

## 性能亮点与生态野心

与其他新技术、新产品一样,GPT同样在朝着更好用、更易用、成本低的方向演进。

OpenAI首届开发者大会,尽管只有45分钟时长,但内容干货较多,总结起来,产品功能的更新大致包括三点:一是最新的大模型GPT-4 Turbo性能大幅升级,比如支持更长的上下文长度,从原先的8K扩展到128K大小,这相当于一本300页的图书,知识库也从原先的2021年9月更新至2023年4月;二是开放了多模态API,包括支持图像输入的GPT-4 Turbo、文本生成图像模型DALL·E3、文本转语音模型TTS,不久后还将支持自动语音识别模型Whisper V3;三是推出GPT定制化服务,任何人无须编程即可创建ChatGPT的定制版本,从而在日常生活、特定任务、公司工作或家庭应用中更加方便,并与其他人分享。

值得注意的是,GPT-4 Turbo在价格上做了大幅下调,用户输入1000 token(指语言模型表示文本的最小单位)比GPT-4便宜3倍,为0.01美元,输出1000 token则便宜了2倍,为0.03美元。同时,GPT-3.5 Turbo也进行了功能升级与各版本的价格下调,比如GPT-3.5 Turbo 16K输入降价至0.001美元,输出降至0.002美元;微调版本的输入费用和输出费用分别降至0.003美元、0.006美元。总体来说,与其他新技术、新产品一样,GPT同样在朝着更好



OpenAI召开首届开发者大会,推出用户自定义版ChatGPT。

视觉中国/图

用、更易用、成本低的方向演进。

除了性能跃升、价格下调之外,最值得关注的是OpenAI对于生态构建的“野心”。Sam Altman在现场表示,本月晚些时候将推出GPT商店,商店里将“展示最好的和最受欢迎的GPT”“我们将付钱给那些提供最有和最常用的GPT开发者以收入分成”“非常高兴能够培育一个充满活力的生态系统”。需要注意的是,苹果正是依靠建立起生态——软件硬件通吃、集聚全球广大开发者、实行收入分成政策的应用商店(App Store)取得了商业成功。显然,OpenAI的商业化探索正从ChatGP等AI产品的订阅付费、API接

入等商业模式向应用商店的生态模式拓展。

谦询智库合伙人龚斌认为,GPT应用商店比之前的插件商店迈进了一大步,变成了GPTs(指各种定制的GPT应用服务)的分发平台,如果它成功构建起一套应用生态的话,OpenAI显然也会变成一个应用平台和AI系统,一个AI时代的产品分发入口。“没错,非常熟悉的味道,跟苹果App Store类似的封闭性甚至垄断性。”

海银资本创始合伙人王煜全预测,未来几年,OpenAI的开发者大会可能会替代苹果发布会,成为AI时代最引人注目的“科技春晚”。

## 哪些业务模式或受冲击?

“未来一家企业的价值,要么拥有大模型,要么拥有用户、场景和数据,如果两者都没有掌控的公司,其生存空间和价值会急剧减小。”

GPT商店能否将苹果App Store生态复制成功,显然还有待观察与验证。不过,数位受访的投资者、创业者向记者普遍提到,GPT-4 Turbo的更新对于原本那些做“GPT套壳”的项目会带来直接的冲击,比如面向B端、C端做知识库、问答、咨询类的项目,还有一些做中间件的创业机会。其中一位受访的投资者坦言,本来今年中期已在看几个相关项目,但GPT更新之后觉得投资风险大而感到担忧,所以决策会更加谨慎。

根据OpenAI的介绍,用户无需编写任何代码就可以为自己的特定任务创建定制版本的GPT,还可以将GPT连接到API以执行任务——比如管理数据库、发送短信、回复电子邮件等。OpenAI现场演示了策划编程课程、生成设计图、自动化流程等应用场景。

龚斌指出,OpenAI推出的“助手API”(Assistant API)使开发者能够更轻松地构建辅助AI应用,并调用模型和工具,显然是在AI框架及底座方面的充实。此举或将挤压许多基于大语言模型的开源框架的生存空间。未来一家企业的价值,要么拥有大模型,要么拥有用户、场景和数据,如果两者都没有掌控的公司,其生存空间和价值会急剧减小。

王煜全也认为,如果有的创业公司花了大量时间和成本却只是做模型调优,或者在模型基础上做功能叠加,那自然会被取代。但另一方面,这一轮的生成式AI变革也会带来新的机遇。

观看了OpenAI整场开发者大会的北京高勤律师事务所律师王源律师指出,GPT-4的更新将更多赋能的功能集成在底层平台中,去掉了中介,进一步消除了信息的不对称性。对于企业来说,开发及使用的门槛都大幅降低,各种细分领域、行业、场景中或将涌现出各种GPTs,工作效率会提升。但提醒企业注意的是知识产权问题,这是AIGC(生成式人工智能)领域一直绕不开的问题,无论是生成的图片、文本、语音等内容,在商用时都会面临知识产权归属的顾虑。需要指出的是,Sam Altman在开发者大会上表示,针对ChatGPT企业版和API开发人员,如果面临版权侵权的法律诉讼,OpenAI将为客户辩护并支付所产生的费用。不过,使用ChatGPT免费版本的用户不包括在内。据记者了解,此类法律保护行为此前已被谷歌、微软、亚马逊等科技企业所采用,面向其生成式AI用户有条件地提供。

另有业内人士提醒,除了知识产权的风险之外,开发者还要

考虑和注意平台内化吞噬的风险,如果价值最高、最热门的应用均由OpenAI自己做,变成“全家桶”模式,那么第三方应用积累的知识能力和容易被GPT所内化吞噬,变成GPT能力的一部分。

那么,作为通用人工智能的代表,GPT迭代升级,是否会垂直类的大模型冲击呢?几位来自国产大模型的人士表示,通用大模型聚焦的仍是通用领域,从目前的数据、算力及算法来看,AI大模型并不是万能的,在产业落地方面仍然受到很多制约,加之在一些国家和地区市场,ChatGPT产品还没有获得监管层面的认可;垂直类大模型在落地解决对精度要求较高、对泛化能力要求较低的场景中可能会更适合,从我国的市场环境来看,通用大模型和行业垂直类的大模型共存将是一个常态。

不难预测,OpenAI对于生态建设的布局,无疑将引发应用分发巨头的利益争夺,无论是作为OpenAI投资方的微软,还是谷歌、亚马逊、苹果等,还是超级富豪埃隆·马斯克旗下的X等,显然并不会坐视OpenAI赢家通吃的局面形成。AI领域的争夺战才刚刚开始,围绕新一轮生成式AI的创业大幕也才刚刚拉开,未来市场将如何变化值得业界持续关注。

## 阿里欲转型“科技平台” AI时代瞄准基础设施

本报记者 李立 上海报道

2023年世界互联网大会乌镇峰会上,阿里巴巴集团CEO吴泳铭提出阿里巴巴的最新定位:AI时代,阿里巴巴要成为一家服务

## 从平台到“科技平台”

在经济学概念里,平台是一种虚拟或真实的交易场所,平台本身不生产产品,但可以促成双方或多方供求之间的交易,收取适当的费用或赚取差价而获得收益。

与平台相对应,平台经济(Platform Economics)是基于数字技术,由数据驱动、平台支撑、网络协同的

## 打造AI基础设施,开放生态

11月9日上午,吴泳铭进一步分享了对“AI技术革命”的看法,在他看来,AI计算正在成为整个数字世界的基石,成为计算资源的调度者和中枢神经。

围绕这个核心观点,他分享了一些思考。一是AI技术将从根本上改变知识迭代和社会协同的方式。以大模型为代表的AI技术携带巨大的算力,对生产力和生产关系、数字世界和现实世界产生全面的深远影响。

二是我们正处在传统计算和AI计算重要级别切换的节点上,AI计算将会接管以前由程序员管理的传统计算资源;此外,AI将直接理解用户需求,让软件具备自主优化升级的能力。基于以上判断,吴泳铭认为,发展AI的关键是创建强健的基础设施和开放的生态。

全社会AI创新的、开放的“科技平台企业”。

在“科技平台企业”的新定位下,阿里巴巴将坚持做好两件事,“一是提供稳定高效的AI基础服务体系,二是创建开放繁荣的AI生

经济活动单元所构成的新经济系统。

作为平台经济的典型代表之一,阿里巴巴的新定位进一步强调了科技定位。吴泳铭预判:“会有更智能的下一代产品进入我们的生活。更多中小企业将通过AI化协同,灵活替代一部分目前只有大企业才能提供的服务。”

接近阿里巴巴的观察人士认为,这是吴泳铭首次详细阐述阿里巴巴的新定位以及具体策略。在吴泳铭宣布上任的两个半月内,AI与阿里巴巴各子业务的融合在加速。

在刚刚落幕的云栖大会上,吴泳铭和阿里巴巴集团董事会主席蔡崇信双双现身,蔡崇信更在开幕式上提出阿里云最新战略,“希望成为AI时代一朵最开放的云”,据其现场透露,“中国80%的科技企业和一半的大模型公司都跑在阿里云上。”

此番吴泳铭更进一步指出,“AI+云计算”的双轮驱动,是阿里云面向未来、支撑AI基础设施的底层能力。据他透露,阿里巴巴即将开源720亿参数大模型。这将是国内参数规模最大的开源大模型。

作为阿里巴巴核心业务的基本盘,AI也在加速渗透电商业务。记

者从淘天集团了解到,在今年的“双11”,淘宝天猫平台上商家已经开始规模化使用AI工具,调用AI次数已超过15亿次。

值得注意的是,这并不是吴泳铭第一次提及AI的重要性。9月12日,吴泳铭就任阿里巴巴集团CEO,就在内部信中明确两大战略重心:用户为

在这样的新定位下,阿里巴巴要做两件事,“一是提供稳定高效的AI基础服务体系,特别是强大的云计算能力,为全行业训练AI、全社会使用AI打造坚实的基础底座。二是创建开放繁荣的AI生态。”吴泳铭表示。

为什么新定位要“把平台能力从商品交易延伸到科技”?吴泳铭

先,AI驱动。

接近阿里巴巴的观察人士认为,无论是今年的“双11”还是云栖大会,在过去两个月,AI与阿里电商、云计算等子业务正在加速融合。

进一步解释,“回望这20年,我们打造了覆盖交易、物流、支付、生产等环节的数字商业基础服务,为电商行业发展提供了有力支撑。而随着AI成为中国数字经济创新突破的关键,阿里巴巴要做一家开放的科技平台企业,为千行百业的AI创新和转型提供基础设施。”

运营等7大场景,简单高效触达淘宝数亿消费者。

内容生产领域,AIGC使得内容创作门槛大幅降低,淘宝通过AIGC技术帮助商家创作内容,得到更好的获客和留存。据悉,目前在应用淘宝AI创意能力的商家及品牌数已突破400万;为商家产出图文创意日均超3亿件,视频创意超500万件。

“从平台经济转型科技平台企业,在AI浪潮推动的新一轮竞争中,阿里巴巴的基础定位并未发生转变”,上述观察人士指出,从PC时代到移动互联网时代,阿里巴巴都以传统电商业务为核心,搭建电商交易系统的基础设施,进入AI时代,阿里巴巴仍然希望成为基础设施,一旦转型成功,有可能继续分享AI时代的新一轮红利。

运营等7大场景,简单高效触达淘宝数亿消费者。

内容生产领域,AIGC使得内容创作门槛大幅降低,淘宝通过AIGC技术帮助商家创作内容,得到更好的获客和留存。据悉,目前在应用淘宝AI创意能力的商家及品牌数已突破400万;为商家产出图文创意日均超3亿件,视频创意超500万件。

“从平台经济转型科技平台企业,在AI浪潮推动的新一轮竞争中,阿里巴巴的基础定位并未发生转变”,上述观察人士指出,从PC时代到移动互联网时代,阿里巴巴都以传统电商业务为核心,搭建电商交易系统的基础设施,进入AI时代,阿里巴巴仍然希望成为基础设施,一旦转型成功,有可能继续分享AI时代的新一轮红利。

## 从“P图”到“P视频” AIGC合规路如何走?

本报记者 秦鼎 北京报道

近日,一段郭德纲外语说相声,知名歌星泰勒·斯威夫特、演员艾玛·沃森等外国名人说地道中文的视频在社交平台上疯传,该视频实际由AI合成,不仅能模仿演员原声、腔调,甚至口型也对得上,让不少观众觉得惊讶。

## 7个月收入百万美元

此类AI视频的爆发,吸引了众多网友开始效仿,其背后一款AI应用也浮出水面。

记者了解到,大部分的翻译视频出自一款叫做HeyGen的应用,其并没有具体的App,只能在官网使用。

记者通过体验了解到,用户需要使用邮箱注册并登录HeyGen官方网站,用户需要上传想要讲话的人像图片或视频,点击“create video”之后,再上传音频,便能生成图片中说话者说翻译后语言的视频。HeyGen虽然可以免费试用,但是生成视频需要排队等待,想要跳过排队或者享受更多的服务则需要充值,付费版需要24美元/月。

但从其他用户上传的视频来看,通过HeyGen“翻译”后重新配音的新视频不仅进行了较准确的翻译,而且连声线和情绪都近乎逼真,嘴型很贴合,难辨真假。

有博主在社交平台上表示:“在排队等待7000个视频

实际上,不仅仅是视频,AI正广泛运用在绘图、配音、剧本写作等方面,但随之而来的是AI在版权、肖像权、姓名权等方面的争议。《中国经营报》记者在采访中了解到,一些有AI产出的内容并没有获得授权,这种行为可以被认定为侵权行为的可能性较高。

HeyGen的前身是一款名为“Movio”的AI视频翻译平台,背后所依靠的模型是自家的多模态内容生成引擎“Surreal Engine”,由诗云科技(深圳)有限公司(以下简称“诗云科技”)开发。天眼查数据显示,诗云科技成立于2020年,目前已完成两轮融资,获得了来自IDG、红杉中国、真格基金累计近千万元的投资。

2022年7月,诗云科技推出多模态内容生成引擎Surreal Engine,将内容生产分为Understanding(理解)、Framing(视框化)、Rendering(渲染)三个步骤。

HeyGen的产品正式上线不到一年,诗云科技创始人徐卓便对外表示,HeyGen在7个月内实现了100万美元的ARR(年度经常性收入),并保持连续9个月50%的月环比增长率。

下转 C3