

人工智能也能提供“情绪价值”了

新版ChatGPT,加强语音、文本、音视频交互功能

北京时间5月14日凌晨,全球知名人工智能研究机构OpenAI在其春季发布会上宣布,正式推出全新多模态大模型GPT-4o。这一革命性的产品标志着生成式AI领域的新突破,为用户带来前所未有的交互体验。GPT-4o以其强大的文本、音频、图像处理能力,以及快速响应和免费开放的特点,迅速引发业界的广泛关注。

现代快报/现代+记者 龙秋利 综合



◀和GPT-4o(小图)一起过生日的演示视频

响应速度仅232毫秒,聊天堪比真人

GPT-4o作为OpenAI的最新旗舰产品,其名称中的“o”代表“omni”,意为“全能的”。这款模型不仅支持文本输入,还能接受音频和图像的任意组合作为输入,并生成相应的文本、音频和图像输出。这一特性使得GPT-4o在人机交互方面展现出极高的灵活性和适应性。

发布会上,OpenAI展示了GPT-4o的实时互动能力。无论是语音输入还是图像识别,GPT-4o都能在短时间内给出准确的回应。用户能够像与助手互动一样自然地使用ChatGPT,也可以在ChatGPT回答问题时中断它。而且,新模型能够捕捉到用户声音中的情感,并以不同的情感风格生成语音,如同

真人一般。

音频输入方面,GPT-4o的响应速度仅232毫秒至320毫秒,与人类对话反应速度相当。这种近乎实时的交互体验,使得GPT-4o在语音助手、智能客服等领域具有巨大的应用潜力。

在OpenAI释出的演示视频中,当使用者将镜头对准生日蛋糕和蜡烛时,GPT-4o能快速反应出他们在过生日。当使用者提出希望听到生日祝福歌时,GPT-4o能够像真人一样唱歌,且语气俏皮不生硬。

值得一提的是,GPT-4o将对所有用户免费开放,这大大降低了AI技术的使用门槛,让更多人能够体验到AI技术带来的便利。同时,OpenAI还为Plus用户提供了额外的福利,包括5倍的调用额度等。

GPT-4o的强大功能得益于其跨文本、视觉和音频的端到端训练

方式。这意味着所有输入和输出都由同一个神经网络处理,从而实现了高效的信息整合和生成。这种训练方式不仅提高了模型的性能,还使得GPT-4o在图像和音频理解方面的表现尤为出色。

CEO山姆·奥特曼表示,最初的ChatGPT展示了语言界面的雏形,而新的ChatGPT则给人截然不同的感觉。“它快速、智能、有趣、自然,且有用。”

“对我来说,与电脑交谈从来没有真的自然过,而现在却很自然。我真的看到了一个令人兴奋的未来,在这个未来,我们可以使用计算机做比以往更多的事情。”奥特曼说。

GPT-4o或引发AI应用新风口

这款多模态大模型在文字、语

音和视频方面取得显著突破,极大地提升了AI的应用潜力。GPT-4o的推出或将加速AI应用的落地进程,推动AI技术向更广泛的领域渗透。同时,GPT-4o的多模态交互能力也将为AI技术带来全新的发展机遇和挑战。

硅谷连续创业者、Traini创始人及CEO孙邻家在接受第一财经记者采访时表示,OpenAI的最新发布将交互方式进行了升级、变革。“语音正在成为一个较稳定的接口,拓宽人类与产品交互的物理边界。而且与GPT实时对话时,GPT能富有情感地回应,这个很酷,之前AGI是缺乏情感的。”他还表示,人工智能(AI)领域正在往AI应用方向发展,将给生活创造真实价值。

然而,也有人认为GPT-4o的未来发展持谨慎态度。他们认为,虽然GPT-4o在技术上取得重大突

破,但在实际应用中仍需要解决一些问题和挑战。例如,如何确保GPT-4o在处理用户数据时的隐私安全、如何避免GPT-4o的误判和误导等。

对此,OpenAI表示,为GPT-4o“创建了新的安全系统,为语音输出提供保护”。GPT-4o还与社会心理学、偏见和公平性以及错误信息等领域的70多名专家进行了广泛的外部评估,以确定新增加的模型会带来或放大哪些风险。

针对可汗学院的可汗使用GPT-4o辅导他儿子做数学题的示例,人工智能软件开发人员麦凯·瑞格利(Mckay Wrigley)在社交平台上写道:“这个演示太疯狂了。学生与GPT-4o共享iPad屏幕,人工智能与他们对话,实时帮助他们学习。想象一下,如果世界上每个学生都能这样学习,未来是如此的光明。”

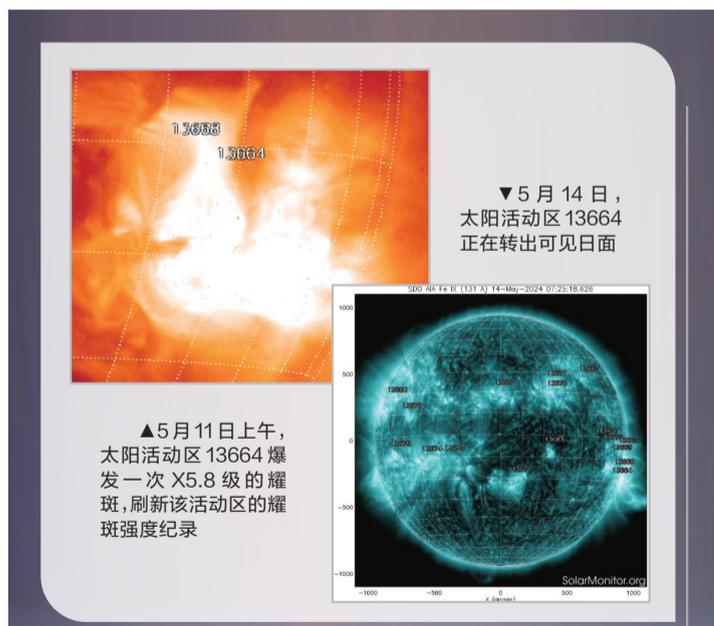
有网友觉得OpenAI进一步拉开了和苹果的差距,认为GPT-4o“功能秒杀Siri”。

不过,据外媒报道,苹果公司已接近与OpenAI达成协议,今年将后者的部分技术引入iPhone,提供由ChatGPT支持的“聊天机器人”作为iOS 18中人工智能功能的一部分。

不过,苹果同时还与谷歌就授权Gemini聊天机器人进行了谈判。而谷歌将在OpenAI此次春季更新发布会之后一天举行I/O开发者大会,谷歌在大会的官方微博文中曾提到“Gemini时代”,预计将会发布Gemini大模型的最新动态。

7天爆发至少9个X级耀斑,13664谢幕

超强太阳风暴会卷土重来吗



▼5月14日,太阳活动区13664正在转出可见日面

▲5月11日上午,太阳活动区13664爆发一次X5.8级的耀斑,刷新该活动区的耀斑强度纪录

5月8日以来,太阳活动区13664几近“疯狂”,爆发了至少9次X级耀斑活动,产生超强太阳风暴袭击地球,引发最高等级的超大地磁暴。但随着太阳的自转,我们终于要和它说再见了。

现代快报/现代+记者 储希豪 部分图片来自国家预警中心、SolarMonitor、微博

小黑点的逆袭成长

太阳活动区13664最早形成于5月2日,起初只是个不起眼的小黑点,没人会去过多地关注它。当时,人们的目光大概都集中在它的“前辈”——太阳活动区13663上。

5月3日“嫦娥六号”月球探测器发射前,太阳活动区13663突然爆发X1.6级大耀斑,引发关注。随后的6天时间里又爆发4次X级耀斑,但并没有伴随明显的日冕物质抛射和太阳质子事件。简单来说,虽然厉害,但只有一轮攻击——耀斑。

而天文学家将太阳13664归为 $\beta-\gamma-\delta$ 类型,表明这一活动区的磁场结构极为复杂,极易产生大的耀斑爆发。果然,当“前辈”即将转出可见日面之际,太阳活动区13664立马接过“接力棒”,在5月8日下午和5月9日早上爆发了两次X1.0级大耀斑。此时的它,已经从小黑点逆袭成为可见日面上极为庞大的太阳黑子群。

大耀斑爆发时,这一活动区靠近可见日面中心,伴随有多个全晕状日冕物质抛射,这场太阳风暴直接命中地球的概率很高,于是国家空间天气监测预警中心发布大地磁暴预警。此时人们才意识到,太阳活动区13664绝非等闲之辈。

引发超大地磁暴事件

事实上,这场太阳风暴远比预计的更强烈。5月10日晚至11日凌晨,地球出现3小时超大地磁暴、6小时特大地磁暴,而且近来地磁活动一直较为活跃。

这次超大地磁暴事件给地球上的短波通信、导航定位带来影响,我国大部分地区电离层出现扰动。同时也使电网系统的稳定和安全受到威胁,新西兰国家电网就因安全原因暂时关闭了8个变电站。此外,地磁暴期间高层大气密度显著增加,导致部分低轨卫星出现轨道衰减。

“上一次发生如此规模的地磁暴,是2003年10月‘万圣节事件’。”中国科学院紫金山天文台滕伟霖说,当时正处于第23太阳活动周极大期,一系列爆发事件给地球的空间环境带来灾害性的影响,瑞典5万人的电力供应中断,珠峰探险队通信中断,全球卫星通信受到干扰……

当然,这场超大地磁暴事件也为全球中高纬度地区带来了20年来最强的极光活动。在我国黑龙江漠河、内蒙古锡林浩特、新疆阿勒泰地区等地,以及英国、挪威、加拿大、美国、新西兰等国家,都出现绚丽的极光。

还会再相见吗?

5月11日,太阳活动区13664继续发力,爆发了当前第25太阳活动周第二强的X5.8级大耀斑。

频繁又剧烈的太阳活动,为天文学家提供了很好的研究样本,尤其是对研究太阳活动的发生机制和能量释放过程、太阳磁场与太阳耀斑和日冕物质抛射的关系、太阳磁场的演化历程,以及太阳耀斑和日冕物质抛射背后诱因等研究都具有重要意义。

随着太阳的自转,这一活动区来到了可见日面的西侧边缘,告别的方式与“前辈”如出一辙,在5月14日上午再爆发一次X1.7级耀斑,但比“前辈”更强烈,彰显个性。

从5月8日以来,太阳活动区13664已经发生至少9次X级太阳耀斑,爆发强度和爆发频次实属罕见,小到中等耀斑仍在持续爆发。但从它转出可见日面那一刻起,对地球的影响就要暂时告一段落了。

为什么是“暂时”?难道还会相见?据了解,太阳自转一圈需要25.05天,考虑到地球公转的速度,大约再过两周,太阳活动区13664所在位置又将出现在可见日面上,但会不会再次相遇,还无法预测。滕伟霖表示,即使这一黑子活动区继续存在,也会被赋予新的编号。