

外媒评 GPT

最近外媒评价GPT-4，认为这款人工智能系统标志着该领域一个巨大的飞跃。其强大的自然语言处理能力和先进的学习机制使其在处理自然语言方面比以往的人工智能系统更出色，甚至能够像人类一样进行推理和概括。同时，GPT-4也引领了技术道德规范的又一重大转变，提醒人们在人工智能领域中注重技术的慎重使用和规范化。总之，GPT-4的发展显示了人工智能系统未来的巨大潜力，但我们也需要持续关注和控制这一技术，确保其发展方向符合人类的利益和安全。

OpenAI 早期上线的聊天机器人 ChatGPT 用自动生成的流畅文字获得了大量关注，也让公众对其杜撰论文、虚构剧本的能力感到不安。要知道，ChatGPT 用的还是一年多前就已经落伍的老一代技术 GPT-3。



相比之下，最先进的 GPT-4 模型不仅能够自动生成文字，还能够根据用户的简单要求描述图像。例如，当向 GPT-4 展示拳击手套挂在一端有球的木制跷跷板上的照片时，用户问如果手套掉下来会发生什么，GPT-4 会回答说它会撞到跷跷板，让球飞起来。

早期测试者曾声称，GPT-4 在推理和学习新事物的能力方面非常先进。微软周二也透露，上个月发布的必应人工智能聊天机器人一直在使用 GPT-4。

开发者周二表示，这项技术将进一步彻底改变人们的工作和生活。但同时这也让公众担心，人工该如何与这种精密得可怕的机器竞争，人们该如何相信自己在网上看到的内容。

OpenAI 高管们表示，GPT-4 跨文本和图像的“多模式”使其在“高级推理能力”方面远超过 ChatGPT。由于担心功能被滥用，公司推迟发布 GPT-4 的图像描述功能，而且 GPT-4 支持的 ChatGPT Plus 服务订阅用户只能用文本功能。

OpenAI 政策研究员桑迪尼·阿加瓦尔 (Sandhini Agarwal) 称，公司之所以还没发布这一功能，是为了更好地了解潜在风险。OpenAI 发言人尼科·费利克斯 (Niko Felix) 表示，OpenAI 正计划“实施保障措施，防止图像中的个人信息被识别”。

OpenAI 还承认，GPT-4 仍会有“产生幻觉”、出现无稽之谈、延续社会偏见，以及提供糟糕建议等惯常错误。

微软已经向 OpenAI 投资几十亿美元，希望人工智能技术能成为旗下办公软件、搜索引擎和其他在线产品的杀手锏。该公司将这项技术宣传成超级高效的合作伙伴，可以处理重复性工作，让人们专注于创造性工作，比如说帮助软件开发人员完成整个团队工作。

但有关关注人工智能的人士表示，这些可能只是表象，人工智能可能会导致没人能预料商业模式和风险。

人工智能的快速发展，加上 ChatGPT 的爆火，导致业内公司围绕争夺人工智能领域主导地位展开激烈竞争，竞相发布新软件。

这种狂热也引发了不少批评。很多人认为这些公司急于推出未经测试、不受监管且不可预测的技术，可能会欺骗用户，破坏艺术家的作品，并给现实世界带来伤害。

由于设计上就是为了生成令人信服的措辞，人工智能语言模型经常会提供错误答案。而且这些模型要用网络上的信息和图像接受训练，也学会了模仿人类偏见。

OpenAI 研究人员在一份技术报告中写道，“随着 GPT-4 和类似的人工智能系统被广泛采纳”，它们“会强化固有见解”。

OpenAI 前研究员、开源人工智能公司 Hugging Face 政策总监艾琳 索莱曼 (Irene Solaiman) 认为，这种技术进展的速度需要全社会对潜在问题及时做出反应。

她进一步表示，“作为社会一员，我们已经可以在一些不应该由模型造成的危害方面达成广泛共识”，“但许多伤害是微妙的，主要影响的是少数群体。”她补充说，那些有害的偏见“不能成为人工智能性能的次要考虑因素。”

最新的 GPT-4 也并不完全稳定。当用户祝贺人工智能工具升级为 GPT-4 时，它的回应是“我仍然是 GPT-3 模型”。然后被纠正后它为此道歉，并说：“作为 GPT-4，我感谢你的祝贺！”随后用户又开玩笑称，它实际上仍然是 GPT-3 模型，人工智能再次道歉，并说自己“确实是 GPT-3 模型，而不是 GPT-4”。

OpenAI 发言人费利克斯表示，公司研究团队正在调查问题出在哪里。

周二，人工智能研究人员批评 OpenAI 披露的信息不足。公司并没有公布有关 GPT-4 存在哪些偏见的评估数据。热切的工程师们也失望发现，关于 GPT-4 模型、数据集或训练方法的细节信息很少。OpenAI 在技术报告中表示，考虑到面对的“竞争格局和安全影响”，该公司不会透露这些细节。

GPT-4 所处的多感官人工智能领域竞争激烈。谷歌母公司 Alphabet 旗下人工智能公司 DeepMind 去年发布了一个名为 Gato 的全能模型，可以描述图像和玩视频游戏。谷歌本月发布了一款多模式系统 PaLM-E，能将人工智能视觉和语言分析整合到单臂机器人中。例如，如果有人让它去取一些芯片，它可以理解这个请求，转向抽屉，选择合适的物体。

类似系统激发了人们对这项技术潜力的无限乐观，一些人甚至从中看到几乎与人类同等的智能水平。然而，正如批评者和人工智能研究人员认为那样，这些系统只是在重复训练数据中发现既定模式和固有关联，并没有清楚理解其中之意。

GPT-4 是自 2018 年 OpenAI 首次发布以来的第四个“生成式预训练转换器”，基于 2017 年开发的突破性神经网络技术“转换器”。这种通过分析网络文字和图像进行“预训练”的系统让人工智能系统在如何分析人类语音和图像的方式方面取得了飞速进展。

多年来，OpenAI 也从根本上改变了向大众发布人工智能工具的潜在社会风险。2019 年，该公司曾拒绝公开发布 GPT-2，称虽然人工智能表现非常好，但他们担心会出现使用它的“恶意应用程序”。

但去年 11 月，OpenAI 公开上线了基于 GPT-3 微调版本的 ChatGPT。在发布后的短短几天内，用户就超过了 100 万。

ChatGPT 和必应聊天机器人的公开实验表明，如果没有人工干预，这项技术离完美表现还很远。在出现一系列奇怪对话和错误答案后，微软高管承认，就提供正确答案而言，人工智能聊天机器人仍然不值得信赖，

但表示正在开发“信心指标”来解决这个问题。

GPT-4 有望改进一些缺点，科技博主罗伯特 斯科博 (Robert Scoble) 等人工智能倡导者认为，“GPT-4 比任何人预期的都要好。”

OpenAI 首席执行官萨姆 奥尔特曼 (Sam Altman) 曾试图缓和人们对 GPT-4 的期望。他在 1 月份表示，对 GPT-4 能力的猜测已经达到了不可能的高度，“关于 GPT-4 的谣言是荒谬的”，“他们会失望的。”

但奥尔特曼也在推销 OpenAI 的愿景。在上个月的一篇博客文章中，他表示，公司正在计划如何确保“所有人类”都受益于“通用人工通用能” (AGI)。这个行业术语指的是一种仍然不切实际的想法，要让超级人工智能和人类一样聪明，甚至比人类更聪明。

本文链接：<https://dqcm.net/zixun/167885315112260.html>