

Sora火爆上线 谁能打赢“翻身仗”？

春节期间，美国人工智能团队OpenAI发布“文字生成视频”大模型Sora，几乎碾压同行的能力，引发全球震撼。

与此同时，Runway、Pika等国外大模型团队正准备招兵买马，与Sora正面竞争。发布短短96小时内，国内已有AI团队火速上线相关“文字生成视频”大模型，但与国外主流AI视频模型能力相仿，均无法超越Sora的水平。而上海作为国内人工智能产业的重镇，也已经着手开始AI生成视频的研发工作。



碾压同行

2月16日，谷歌发布新一代多模态大模型Gemini 1.5 Pro，在性能上超越OpenAI的GPT-4 Turbo，堪称业界最强的大模型。然而，OpenAI在同日发布Sora大模型后，风头直接盖过谷歌，仅需通过文本即可自动生成视频，这也是继文本模型ChatGPT和图片模型Dall-E之后，又一极具颠覆性的大模型产品。

Sora并非首个文本生成视频的大模型。据不完全统计，截至去年底，全球能实现文生视频的大模型包括Runway、Pika、Stable Video Diffusion等超过20个产品。然而，Sora的诞生仍给业界带来了不小的震撼，它所展现出来的能力几乎可用“碾压”来形容。用户仅需输

入简短一句话，Sora可以生成一段长达60秒的视频，远远超过市面同类型级别的AI视频生成时长。

60秒的视频时长，已经超过抖音等短视频平台的平均视频时长，Sora的诞生也为以后短视频平台的内容生产提供了更大的可能性。

记者体验此前的AI生成视频时发现，即便是几秒钟的视频也并不连贯，有较强的拼凑感。而Sora几乎能做到60秒视频一镜到底，OpenAI的演示视频显示，一名时尚女性走过喧闹繁华的东京街头，无论是背景建筑和街道以及人像的逼真程度，都保持了一致性，哪怕是各种镜头的“蒙太奇”，都没有出现明显的失真情况。

OpenAI在官方博客中写道，Sora不仅可以理解用户的需求，还知道这些事物在现实世界如何存在。而这离不开OpenAI在文本和图像方面的长期积累。

上海市人工智能行业协会秘书长钟俊浩解释，Sora的训练依赖于大量带有文本标题的视频数据。OpenAI利用了DALL-E 3中的标题生成技术，为视频生成高质量的文本标题，以此来提高文本和视频数据之间的对齐度。同时，在大规模训练过程中，Sora展示了一些“涌现的模拟能力”，如三维一致性、长距离连贯性、物体持久性以及与世界互动等。这些能力表明，Sora能够在一定程度上模拟物理世界和数字世界中的人物、动物和环境。

漏洞不少

虽然技术令人惊艳，但Sora的视频生成能力并不完美。根据现已披露的视频显示，不少素材仍会呈现出“一眼假”，不符合物理学规律等AI生成的Bug（漏洞）也不少。

以“幼狼玩耍”的视频为例，Sora生成的内容中出现了巨大漏洞：在嬉戏打闹中，幼狼的数量从3只到5只来回变换，虽然动作依然连贯，但“戏法”般的效果让人摸不到头脑。另外在“一个老人吹蜡烛”的指令下，Sora生成视频却显示，老人对着生日蛋糕的蜡烛吹气，却没有一根蜡烛熄灭；在呈现“红酒杯在

桌上摔碎”的镜头中，杯子在摔碎之前，红酒已经洒满了桌子，“先摔碎后洒酒”的逻辑，Sora显然没有理解。

或许是考虑性能、安全等问题，Sora并未向公众开放。

在钟俊浩眼中，Sora面临的技术挑战还有不少，特别是与文本对话和图片生成相比，训练成本高昂、高质量数据集的缺乏以及视频描述的模糊性都将成为Sora需要跨越的门槛。

“视频生成需要处理和预测长时段跨度的动作和场景变化，这增加了模型的复杂度和训练的计算需求。”他认

为，与文本和图像相比，适用于视频生成的多模态高质量数据集相对较少，这限制了模型学习复杂动作语义的能力。不仅如此，描述视频内容比描述静态图像更加复杂，因为它需要包括时间上的变化、动作以及可能的情感或故事线索。

值得一提的是，在Sora发布的同时，OpenAI宣布一项要约售股交易，由此OpenAI整体估值高达800亿美元，一年内飙升两倍多。以国外调研公司CB Insights统计，OpenAI已经成为全球估值最高的创业公司之一，仅次于字节跳动和SpaceX。

难言颠覆

短短4天内，中信建投、国泰君安、申万宏源、招商证券等10家券商在研报中均表示，Sora是人工智能发展进程中的“里程碑”，预示AGI（通用人工智能）将加速到来，众多行业将迎颠覆式变革。

比如天风证券认为，内容创作 workflow有望被颠覆，下一个亿级用户的互联网平台雏形出现。2000亿美元的短视频创作生态有望率先被颠覆，生成式AI在视频创作和世界模型的大踏步进步将实现对视频、3D、游戏等下游应用场景的渗

透；招商证券指出，当前展示的效果来看，Sora显著领先于其他文生视频模型，推动AI视频生成进入了一个全新的时代；中信证券从投资角度判断，Sora背后的涌现能力为自动驾驶、设计等需要现实世界建模的行业提供了明确方向，硬件端的需求也必然会随着多模态的技术进步而不断提高，仍然持续看好AI算力。

“Sora让人震惊，很大原因是AI的发展速度远远超过了原来的预期。但不意味着短时间能形成生产力。”钟俊

浩认为，Sora现在能力还不足以颠覆行业，特别是影视行业 and 游戏行业高度依赖创意、情感投入和人类故事讲述的领域。AI视频生成技术可以提供帮助，比如降低制作成本、加速内容生成、甚至在某些情境下提供创意灵感，但它们目前还不能完全替代人类在创作过程中的独特视角和情感深度。“Sora的角色更多的是赋能创作者，而不是替代他们提供实现创意愿景的新途径。”

本报综合

@微语录

@余华：人总会有选择自己态度和行为方式的自由，即使被命运残酷对待，依然要选择温柔地对待这个世界。

@白岩松：行走在人群中，我们总是感觉有无数穿心掠肺的目光，有很多飞短流长的冷言，最终乱了心神，渐渐被缚于自己编织的一团乱麻中。其实你是活给自己看的，没有多少人能够把你留在心上。

@毕淑敏：只要你自己不倒，别人可以把你按倒在地上，却不能阻止你满面灰尘遍体伤痕地站起来。

@廖一梅：轻易听信别人告诉你的，让禁忌阻碍你的视野，给自己定下条条框框，过约定俗成的生活，我把这叫做二手生活。

@桐华：如何从容自信地爱不完美的自己，是一门比如何爱别人更深奥的学问。

@微趣图



猫咪：“如果你表现好，我就和你一块玩。”

@微笑话

@苗苗：门诊来了位年轻妈妈，抱着女儿来打疫苗。女儿天真地问：“妈妈，我们来干嘛？”妈妈：“打针啊。”女儿：“干嘛要打针，针做错什么事了吗？”10分钟后女儿开始咆哮：“这是针打我！这是针打我！”

@红色高跟鞋：两只乌鸦站在树枝上张望，看到脚下的麦田里站着个稻草人。一只乌鸦害怕地问另一只乌鸦：“那是个人吗？”另一只乌鸦很不屑地说：“别紧张，那只是稻草人。”“你怎么知道？”“你没发现吗？他没看手机！”

@太阳的宝宝：我给同桌讲数学题，教了半天，同桌还是一窍不通。我还是没有放弃，找了几道同类型的数学题，启发我同桌：“这几道题，都有一个相同点，你仔细看看。”我同桌看了许久，恍然大悟：“哦，这个相同点就是我都做不出来。”我：“……”

据新浪微博