



Sora 颠覆

Sora 到底是何方神圣？

Sora 某种程度上是 ChatGPT 的又一次翻版：它把视频数据统一了，然后用大语言模型的方法理解视频，最后“涌现”出了对物理世界的构建能力。

□ 记者 | 陈 冰

龙年伊始，Sora 横空出世，震惊众人。这个自称是“世界模拟器”的新工具，作为 OpenAI 发布的首个 AI 文（图）生视频模型，带给人们的感受真的只能用炸裂来形容——根据输入的文字提示，就能生成效果逼真的 60 秒视频，而且是可以一镜到底的那种。

仅仅在一年前，人们还在嘲笑 AI 生成的威尔·史密斯（Will Smith）吃意大利面的视频是人工智障。在那段视频里，威尔·史密斯用一种诡异、尴尬的姿势将意大利面铲进一张抽搐、不似人类的嘴里。

可是现在，人们对着 Sora 输入以下文字：

一位时尚的女士穿着黑色皮夹克、长红裙和黑色靴子，手拿黑色手袋，在东京一条灯光温暖、霓虹灯闪烁、带有动感城市标志的街道上自信而随意地行走。她戴着太阳镜，涂着红色口红。街道潮湿而有反光效果，色彩缤纷的灯光仿佛在地面上

创造了镜面效果。许多行人在街上来往。

在这段刷屏的视频中，提示词中的全部细节都得到了体现，而且无论是人物脸上的雀斑，还是水中的倒影都显得极其逼真。以前人们会怀疑，这是假的吧？现在人们会惊叹，这是实拍的吧？不像是 AI 生成。这种从文本到图像再到视频的进阶之路实在走得太快，以至于一年之前大佬还在预言 AGI（通用人工智能）来临之前，还有很长的路要走，现在纷纷开始改口称：AIG 的实现将大大提速。360 集团创始人、董事长周鸿祎近日就宣称，AIG 将在一到两年内实现。

Sora 背后的技术架构是怎样的，Sora 到底有没有理解物理世界的能力，它的出现是否意味着 AGI 将在一年内实现？以 ChatGPT、Sora 为代表的 AI 技术将如何颠覆人类社会既有的种种规则？我们又将如何应对？

关于 Sora，我们有太多的未解之问。

Sora 生成的红裙女视频，提示词中的全部细节都得到了体现。

