

更多AI工具可直接访问：<https://www.faxianai.com/>

希望 Sora 别走 GPT4 的老路



瑞泽洞察 2024-02-19 21:40 上海

原文地址：<https://mp.weixin.qq.com/s/p7fWCQzgGMXdWioqrayuWQ>

秉着让子弹飞一会儿的想法，开工第一天才来聊 Sora，一个直观感受是，这一波在国内 AI、科技圈的热度远不如去年初 GPT3.5 的问世。

但仍不得不说，Sora 的表现堪称惊艳，国内大模型要追赶上也仍需时日，只是动不动又说很快要颠覆、掀起惊涛骇浪、又大幅拉开中美差距之类的，对促进认知没有意义。

看了官方发布的几十个视频和论文中译版及专业解读，也看了一些算法专家的分析，文生视频和文生文是有较大不同，但总离不开基础理论、算法模型、数据源和工程化等几方面的迭代。

Sora 的出现，基础理论并没有重大创新，算法模型是基于现有几种的混合应用，数据源和质量没有披露，工程化细节也没有解密，当然这也是最具壁垒的部分。

工程化能力是近几年被高频提及的词，简单讲一种新材料可以在实验室的反复试错条件下获得极小部分颗粒，但要形成年产万吨的材料，则需建造一个占地百亩的化工园区，大到一套高达数十米的裂解装置，小到里面零零散散的精密元件，以及运营数据背后高度关联的一套套监测、控制软件。万分之一的细节不足都可能影响产品品质、生产效率乃至安全风险。

大模型领域的工程化能力自然不会涉及那么多大型装备、安装调试、工艺流程和更为复杂的项目管理，但一个不一样的壁垒在于，**这需要算法、机器视觉等领域极高的人才门槛和钻研深度，不是单纯的人才密度。**

从 Sora 的论文发布可以看出，核心研究团队就那么几个人，并且核心作者还是刚毕业的几个博士甚至本科学历。国内 top10 的计算机专业院校，每年也毕业数以千计的计算机博士，更别说算上硕士、本科学历，但为什么这些人可能都去不了达摩院，更别说被 OpenAI 以一两百万美金年薪的应届生招进去，是的，比 P10 薪资还高得多。

这就是为什么在国外诞生 Pika，在国内有 Minimax，Moonshot 的一个根本原因。有顶尖学历的人才不少，但其中**具备较高的产品思维能力、优秀的创造性问题解决能力以及极强的热爱精神的，少之又少。**这几个条件缺一个，都没办法用惊人的速度在工程化的道路上解决数以百计的熬人挑战。

就我认识的 AIGC 领域创业者、技术大拿都有个特质，为解决一个大问题可以连续一周通宵，平均每天睡 3 小时，其实是够睡的，因为人在极度亢奋情况下根本没睡意。纵观人类的科研历史已经反复证明，哪怕是顶尖聪明的科学家，也只有一定时期内处于高度亢奋状态，才可能突破 5% 以外的人类智力极限，尤其普遍发生在 35 岁前。

但另一面还得强调门槛，这批作者几乎都是全美顶尖计算机专业院校的尖子生，这种人才放到全世界，每年的产量恐怕也就上百人，其实是极为稀缺的。先得识别出这样的尖子生，还得给予去大厂 5-10 倍的应届生年薪，然后这批人还得满足以上几个特质，并且最终还得背靠极其烧钱的基础设施能力，**这是 OpenAI 对抗大厂的一个关键壁垒，不在于堆人。**

所以总有人问，我们赶上别人需要多久？**重点不在于多久，而在于别人为什么总能比你领先？**把这个背后的一整套逻辑看清楚，那么我们要解答的就不是多久，而是怎么去打造这样的创新环境和激励机制。

其实国内大厂早就对此有深刻的反思，毕竟也是一批次顶尖人才，虽然待在大厂发挥不了实力，但不代表看不懂别人为什么能干出来。不仅反思，也做了应对方案。譬如阿里，虽然堆了自研团队，人数很多，但既不够出挑，也不够卖命。达摩院的一些大牛走得也差不多了。所以还得买个保险，必须同时下注（投资）外部的独立创业团队。

这里有个细节，有人鼓吹说大厂全靠接 GPT4 接口的说法，不认同。国内搞大模型的 5 家大厂和 5 家独立团队，如果有 1-2 家完全靠接 GPT4 接口，或许是可信的，但把国内 top 大模型说的如此一文不值，不符合逻辑。

这几个大厂，如果过去整整一年，整个大模型团队的产出都靠接别人的，首先，这 token 量无比之大，OpenAI 对单账户是有明确限制且 2023 年新账号一度暂停申请，这种规模的使用量，怎么可能撑住一年？其次，你是字节、阿里、百度的大模型一号位，这种大规模用量的套壳操作，你敢？

充其量不过是小规模使用下，学习友商的一种策略。哪个企业追赶对手，不把对手的产品大卸八块、反复研究的？至于基于别家的开源大模型，除了李开复的团队说的不明不白的，已经被圈内吐槽吃相难看了，还有哪几家没说清楚的？用开源不丢人，关键得说清楚。

所以不必问追赶还需多久，只需想想，OpenAI 的竞争优势，我们怎么弥补。META 首席科学家、AI 大神 Lecun 最近还就中国的学术环境发表评论，他说咱们的目标导向有点问题、激励机制不是太完善，翻译过来就是，唯利是图、氛围糟糕。

不仅在 AI 领域，硬科技走到深水区后也是如此，绝大部分院士级的硬科技项目，哪个教授每天只睡 3 小时，然后把剩余时间 100% 投入创业的？我看到的都是，教授既不从高校离职，也没全情投入，所谓研发就是包给研究生干苦力，自己占大股。

做得出来是自己的，做不出来继续享受学术圈的座上宾，一个不拉。真搞不出来，普通投资人敢动院士一根毫毛？这就是咱们该反思的，这种基础科学的学术环境下，怎么可能在源头保持竞争力？所以真正好人才，一半都去国外了。

接下来进入正题，Sora 足够惊艳毫无疑问，但关键问题仍需探讨。

首先，Sora 团队并没有明确的说，已经能实现基于成熟的世界因果关系模型，而是巧妙地用了世界模拟器这种描述，但**共生关系和复杂世界运行背后的连续因果关系是两个层次**。顶尖学术圈是很严谨的，即便奥特曼是资本运作和营销大师，也不会随便乱说。

就在 Sora 发布后的 2-3 天，Lecun 在世界政府峰会上就差指名道姓点出自己对 Sora 的不看好，指出这并不是真正的世界模型，并且仍会面临 GPT4 的巨大瓶颈。同样的，这样的学术大神也是爱惜自己

名声的，如果真牛逼，人家不会吝啬，更不会如此坚定的不看好。

涌现这个词也被用烂了。去年 GPT 刚出现涌现能力时，说实话哪怕大神也得掂量几分，坐等后续，结果呢？伴随而来的是海量、持续、难以解决的幻觉、精度问题，到今天依然无法解决。

文生文、文生图片到底在全世界颠覆了什么产业？好像放眼全球也没看到哪个能解决过去长期人类痛点的全球规模化的商业应用，有么？2C 的没有，2B 的更没有。GPT4 用周鸿祎的说法，影响力已经从原子弹变成茶叶蛋。

放到文生视频，目前只是测试阶段，还未放开。既然 GPT4 至今仍无法解决幻觉严重问题，换到文生视频工具上，**任何复杂视频的生成也无不牵涉到物理世界数以万级的因果关系和逻辑推理**，现在还只是几只猫、几条船翻来覆去绕，大规模开放使用后，有没有可能在修复当前 bug 的同时，也出现越来越多且难以消除的幻觉以及 corner case 问题？

一旦放开使用，有没有可能出现数以百万个逻辑漏洞百出的视频全世界满天飞？然后和当初 GPT 横空出世一样，全世界用好玩的心态玩一遍，最后发现大规模商业化并不容易。

马斯克这两天还对 Sora 进行了补刀，他说**这个能力特斯拉一年前就有了并已经在用，FSD 的训练视频也有文生视频，对真实道路的视频模拟也达到了高颗粒度的物理世界还原**，只是都是些房子、车子、路桥、杂货铺子和来来往往的行人，发出来也很无趣。

所以有人说马斯克退出 OpenAI 董事会肯定相当后悔，那是你想多了。这个人过去 20 年把世界上不可想象的极难、伟大任务都完成了，领导世界顶尖的 AI 团队，自己还有 2,500 亿美金的身价。今天只要马斯克觉得值得优先干的事，他为什么不干？同理，他面对全世界的自信满满，你觉得是逗着玩？

奥特曼应该找马斯克好好聊下：兄弟，没必要老怼我，怎么说也有 OpenAI 创业缘分，再说，不为 OpenAI 考虑，也得为 50 万亿美元的美股市场考虑下，OpenAI 搞不成没关系，美股怎么办？世界经济怎么办？

依然用开头的话来总结：Sora 很惊艳，国内追赶尚需时日，问题咱们自己也清楚，问多久赶上不如问我们差在哪？同时，Sora 自身仍充满挑战，较大可能走到 GPT4 的大瓶颈，全球级的学术大神、产研巨佬均不看好或不屑一顾。

最近说点看了一堆文章都没分析却不得不分析的论证。

Sora 的文生视频之惊艳，确实让人想到是不是在电影、短视频、游戏、后期制作、广告等领域有最直接的应用价值。**大部分人只是直觉认为行，却不深入思考为什么能行。**

当论证一项新的应用技术怎么颠覆传统产业时，最起码要搞明白传统产业是怎么运行的，比如究竟什么是创意内容、什么是短视频、什么是流量分发、什么是电影工业和广告。

先做个假设，即使 Sora 很成熟，成本和使用门槛也足够低，且 99% 情况下没 bug，那 Sora 的文生视频工具，到底能撬动多大的市场？我们粗略把内容分成长视频、短视频，以及不同精美程度下的各类内容题材。

首先就内容创意领域。很多人不知道，每年各种小电影、微电影、网络电影乃至动辄几千万投入的院线电影加起来有几万部，并且大部分还是专业人士进行编导的，最终能到院线排片的就几百部，能出圈的就十几部。

国内全职的编导从业者就有几十万人，细分到文艺、喜剧、悬疑、科幻等不同内容题材，里面既有阳春白雪，也有精神垃圾，应有尽有，凭什么人家靠吃饭的家伙都只能勉强过活，而所谓有创意天赋的人，加上好的生产工具就能超越这几十万专业人士？

好内容离不开传播平台，再讲内容分发。

照理说有短视频平台的加持，这每年几万部电影不应该有出圈机会么？实际上哪个导演、编剧是靠短视频产生了可持续的商业模式的？如果短视频上能跑通长视频，那爱优腾早就倒闭了，这是一个基本的互联网产品思维，也是为什么抖音上最火的电影都是精华剪辑风格，即便严重侵犯了版权。

短视频平台走不通的话，那制作一批精美的长视频内容，走院线、爱优腾渠道呢？那更是异想天开。群众的审美是有限的、院线的档期是有限的、明星的稀缺是有限的、资本的力量是无限的，宁浩加上刘德华都搞不定的市场，岂是内容优劣决定的？

一部全程靠宣传减肥的电影，获得了 1 亿骂声和 1 亿叫好，也获得了几十亿票房，贾玲无疑是极其成功的商业片导演，但成功背后是为什么？真实的传统电影市场及其产业链，远比随便 YY 复杂的多，岂是一个有满腹创意的普通人，加上一套神级制作工具，能掀翻的？

说完长视频，再说短视频内容。抖音早年能一骑绝尘，靠的那些搞笑、生活视频，并没有多大的制作成本，放到今天也就上千到几千块的单个视频成本，但是，视频脚本的逻辑性以及人物演绎张力背后的一连串细微因果关系，并不简单。本来就上千的制作成本，用了 Sora 恐怕还得反复微调，那 Sora 到底能起到什么颠覆作用？

这是简单内容的短视频生成，那如果是自动生成精致、酷炫的短视频内容呢？我挺爱刷那些好莱坞科幻电影集锦的，但至今重复刷到的概率也很低，就这批好莱坞水准的内容供给量也已经够够了。抖音是要啥有啥，和源源不断的内容供给相比，流量才是稀缺且宝贵的，这是平台的命根子。

即便再有 5 倍量的靠 Sora 生成的科幻微电影，凭什么流量要给你，而不是直接用正儿八经重金投入的大片剪辑？短视频内容从题材来看是两极化的，用户既需要精神垃圾，也需要大片视觉和阳春白雪，无论哪种在供给侧都不是稀缺，而是有些过剩了。凭什么在流量还要分给直播和商城板块时，还要溢价招募新的 AI 内容创作者？挺讲不通的，这不是生意逻辑。

再插一下关于科幻电影、3D 游戏及其背后的后期制作软件市场，这应该是 Sora 在克服一系列技术挑战后，能产生直接商业化影响的领域。

科幻电影制作比 3D 游戏制作需要用到更多软件，剪辑、合成、跟踪、模型、特效几个环节经常用到的软件包括 Nuke，Houdini，Adobe Premiere & AE，游戏制作常用软件就更知名，包括市值 200 亿美元的 Unity 和 Unreal。

但这属于 AI 在 2B 市场的应用，和 2C 领域是完全不同的商业竞争逻辑。即便 Sora 成功了，能不能长期形成绝对壁垒，确保其他友商的水平永远差得远？如果不是，那最终结果无一例外是把产品价格打下来、市场规模打下来。

因此，假设 Sora 能让好莱坞大片、重度 3D 游戏的后期制作成本和周期大幅下降，但这并没有根本性改变电影、游戏行业的格局，也不能左右院线的排片、明星的档期、发行的渠道、资本的偏好。**倒是把后期制作软件市场干萎缩了。**

最后说广告制作，这更没什么可讲。随便问下你身边在广告公司干过的人，算一算该公司一年营收里，用于广告制作的成本是多少，广告制作的成本结构又是怎样的。

当我们说一则大片广告花了 1 亿预算时，大部分是流量投放成本，制作一般在几百万，尽管不算低，但其中大头并不是广告公司赚的，是给广告导演赚走了，创意赚不了几个钱。

当 Sora 足够好用时，能影响的也不过是这批赚的盆满钵满的广告导演的报价从几百万砍到几十万，这又能给广告行业带来多大冲击呢？本身就是个苦命的行当，钱都给流量平台拿走了，还能怎么颠覆？

当然，还会有其他场景的商业化应用可能性，不一一举证，说这些的目的无非是觉得每次谁放大招，能看的文章不是一味夸上天，各种马上要颠覆，就是东搬西抄，稍微有点独立、深度、辩证思考和价值判断的分析少之又少。

说错也没关系，就像我们上一篇卫星文章，洋洋洒洒几千字，但犯了一个比较大的论点错误被人指出，后面会重新论证。AI 投资圈已经够惨了，大家还是切实际一些，反倒能发现一些新出路、创造一些新机会。

最后补充一点，目前生成式 AI 领域还没迎来真正的竞争，大家都在试错和探索阶段，仍要时刻关注类似 META/Google 等大厂的切实进展，尽管大厂有大厂的组织效率问题，但毕竟仅这两家就超过 1,200 亿美金的利润，**谁知道这些大厂有没有暗地里两条腿走路呢？**只是当下犯不着和一家擅长营销、营收还在十几亿的初创企业干架。

还得留意一点，OpenAI 这个神奇公司，去年上演了全球惊叹的宫斗戏后，首席科学家 Ilya 已经很久没发声，这次 Sora 发布也就他没出声。此外，另一位初创成员、技术大牛 Andrej Karpathy 恰巧在 Sora 发布前宣布离职，他之前是特斯拉自动驾驶的负责人，对 FSD 的发展起到关键作用。

他们出去后会做什么？联合做点事还是一起去特斯拉帮衬？这么看，OpenAI 接下来除技术突破外，好戏应该也免不了，拭目以待。