

# 5分钟出两首，AI写歌“杀疯了”

## 入门门槛很低，人人可以娱乐；模仿能力十足，但难创造经典

听过唐僧、孙悟空唱歌吗？

最近，在各大视频网站上，“3A音乐视频”风靡。所谓“3A音乐”即AI作词、AI作曲、AI制作画面，形成的音乐视频。比如，bilibili的创作者（UP主）漫游会议室，就创作了一系列《西游记》相关的作品，其中由孙悟空和黑熊精“演唱”的歌曲，都获得了超500万的播放量。

记者翻阅了几个热度较高的视频，发现无论是唱功，还是词曲的水平，AI作品完全不输真人。有网友说：“AI学会作词唱歌后，在音乐区‘杀疯了’”，也有网友提出疑问：“这样有深度的词曲，真的是AI能够自主完成的吗？”

近日，记者也尝试使用了下相关的AI音乐制作软件。

### 记者实测： 输入关键词，5分钟速成两首歌

记者输入了关键词：“一首关于记者、潮水、激情的带有科技感的流行歌曲”。不到5分钟，就生成了2首歌词相同、曲风不同、长度在3分40秒左右的歌曲。

在歌词部分，仔细区分了前奏、主歌、副歌、间奏以及结尾。每个部分对创作思路和风格进行了描述，比如其中一首的副歌就标注了“旋律：节奏紧凑，带有电子音乐的律动感，突出记者的忙碌与坚定。”

记者将歌曲发给了几位同事，倾听后，他们纷纷表示作品非常完整，“乍一听，和成熟的商业作品很像。不过仔细再听，还是有明显的AI感。”

“AI创作，已经广泛参与到音乐创作领域中。”浙江音乐学院音乐工程系音乐设计与制作教研室负责人沈佶表示，目前学校已经开设了AI创作相关的课程，并在今年第四届数贸会期间主办了AI作曲大赛，“AI创作已经成为了新生代音乐从业者必须掌握的一门技能，可以大幅度降低产业时间成本。我们也鼓励学生使用，利用AI赋能音乐创作。”

在沈佶的记忆中，其实十多年前，人机协作的音乐创作软件就已经开始慢慢发展。

“比如当时最火的Band-in-a-Box，就是通过一些事先设定好的音乐模型，去实现生成自动伴奏，人为干预和操作的占比更高。”沈佶解释道，随着生成式大模型不断发展，现在最快2~3分钟就能自动生成一首歌，“可以说入门门槛很低，只要输入基础的关键词就可以，人人都可以娱乐一下。”

### 专家解释： AI能模仿金曲，但难创经典

当下AI技术发展迅猛，沈佶觉得，很多作品对于非专业的人来说，已经很难区分是真人还是AI。

不过，他也提到，AI创作对于音乐创作者个人审美、音乐素养、知识储备、辨析能力等要求很高，“就拿关键词来说，什么样的描述是最合适和贴切的，什么样的质量或者效果是贴合画面的……这些都需要经验，一首成熟的AI音乐作品，往往需要2~3天的调教过程。”

日常的教学生活中，沈佶也会观看一些AI音乐作品，“AI音乐作品，给短视频行业带来了不小的冲击。各种创意十足的作品，也给我们提供了很多灵感和方向。”

涉及到AI创作，版权风险是绕不过去的坎。沈佶说，现在很多AI音乐创作平台和唱片公司合作，建立自己的AI音乐素材库，“未来，相信整个行业一定会越来越正规。”

对于大家担心的AI音乐“泛滥”会不会导致原创歌曲减少，甚至艺术创作水平下降的问题，沈佶持乐观态度，“当下AI创作的更多是功能性歌曲，通过对现有歌曲的学习、模仿，创作形成新的歌曲。但无法创作出新的艺术风格，也无法突破至新的艺术领域。”

沈佶觉得，随着AI音乐应用越来越广泛，一定程度上对音乐从业者提出了更高的要求。“泛娱乐化时代，一首歌如何能传播更广、更久，更值得我们思考。”

据《钱江晚报》

## 杭州人登山 开“外挂”了



“嚯，这么轻松！”刚穿上智能助行外骨骼，70岁的张海云试着走了两步就忍不住惊叹，腰背处的深灰色“外挂”，引得路过的登山者纷纷好奇地停下脚步。

这套形似“束腰护膝”的深灰色装备仅重2.4公斤，却暗藏玄机——内置AI随动算法，能实时感知使用者动作，精准分配助力——背着它登山至少能节省30%体力。

“膝盖别用力，咱们慢慢走，先适应适应。”一旁工作人员不放心地叮嘱。结果张海云已经健步如飞地跑了起来，顺着半山国家森林公园的斜坡一路往山上去。

近日，在杭州市半山国家森林公园南门，杭州首个提供登山助力常态化共享租赁的服务点刚刚启用。现场摆放的30台供市民体验的智能助行外骨骼机器人，不到1小时已经借出去一半。

除了老年人，不少年轻人也被这个此前只在网上刷到过的黑科技吸引，加入体验。“我以为会很笨重，结果穿上以后除了大腿绑带有轻微压迫感外，其他都很灵活。”小黄和朋友相约爬山，自我调侃是三步一歇的“运动废材”，今天大家都一口气爬到观景台，还发了张四个人穿着外骨骼登山的照片在朋友圈收获一圈点赞。

服务点上一位年轻人正麻利地帮老人穿外骨骼，他其实就是这款智能助行外骨骼的核心技术人员，也是太希智能科技的90后创始人梁林超。“目前在半山国家森林公园投放的这批登山助力外骨骼是成本较低、功能性较强的机器人，就像给双腿装了隐形助推器。”梁林超说，一块电池的续航能力足够在半山游步道打来回，后续将在杭州布设更多的租赁柜和换电柜，比如从半山进一步延伸至龙山、虎山、杭钢公园等景点，串联起10公里景观绿道，“让更多人亲身感受到智能设备在跑步、登山等运动功能上的助力。”

正说着，张海云已经从山上下来了：“不累，真的不累，这比坐景观车带劲。”他不是来还外骨骼的，而是来问价格，想再租一套外骨骼给同来的老伴穿上，一起登望宸阁。

接下来，除了半山国家森林公园这个试点，这款智能助行外骨骼机器人还将在雁荡山、泰山等全国多个重点山岳类景区展开场景应用试点。

据潮新闻

## 高拟真数字人直播带货有多强？

当直播时长逼近6个小时、大多数电商主播显露疲态时，百度直播间里的主播“罗永浩”和搭档“朱萧木”仍然可以情绪饱满地回答观众提问，时不时还能讲两句“段子”，吸引观众下单。这两位主播并非真人，而是百度借助剧本驱动多模协同数字人技术打造的数字人主播。

在2025年世界互联网大会乌镇峰会上，由百度自主研发的剧本驱动多模协同的高拟真数字人技术获得世界互联网大会领先科技奖。

随着人工智能技术快速发展，大模型向多模态加速演进，数字人成为融合大语言模型与多模态技术的创新应用，电商直播是数字人落地的绝佳场景。数字人技术使商家无需投入大量人力、物力进行现场直播，能显著降低场地租赁、设备采购、人员培训等费用。同时，数字人可以24小时不间断直播，进一步增加商品曝光时间和销售机会，提升经济效益。

但传统数字人生成技术常面临语音、语言、视觉多模态割裂的问题，表现为台词生硬、语音语调与台词情感匹配不佳、表情手势单一等。百度首席技术官王海峰说，针对数字人应用痛点，百度创新研发了剧本驱动多模协同的高拟真数字人技术。

互动性强是电商直播场景的一大特点。在与观众互动过程中，语音合成的自然度

是决定用户沉浸感的关键因素。观众希望听到主播有情绪、有起伏的声音，而不是生硬机械的朗读。王海峰介绍，针对这一需求，百度提出了“文本自控的语音合成”方案。文本自控的语音合成大模型不仅具备高复原的语音合成能力，还能够结合直播台词及主播个人特征，把这些文本内容转化为自然、有感染力的声音，让数字人不仅能发出声音，更能精准传递出调侃、得意、强调等细微情绪。

除了要和用户交互，数字人主播在直播时还要与商品和所在空间进行符合物理逻辑的互动。如何做到这一点？高一致性超拟真数字人长视频生成技术，能对输入的历史视频数据、剧本脚本、语音信息以及骨骼驱动等多模态信号进行分析与理解，并基于此分别生成高表现力片段、复杂“人一物一场”交互片段以及大动作大表情片段。系统能够对这些片段在较长时间跨度上进行统一调度，确保语音、口型、表情与动作始终保持高度一致且同步。

中国科学院信息工程研究所正高级工程师韩冀中说，人们在拥抱数字人技术的同时，也需设立清晰的边界，必须防止利用高拟真技术以假乱真进行欺诈或虚假宣传。技术的发展必须与法律、伦理的约束并行，确保创新在正确的轨道上行稳致远。

据《科技日报》

晚报进家门·征订季活动

周周有好礼第一波火热开启

集字领好礼首字：

