

别让AI生成奇葩带偏人类认知



随着AI生成视频技术逐渐普及,不少猎奇短视频出现在社交媒体平台上,给用户带来感官刺激。部分作品内容真假难辨,甚至含有恐怖、血腥、迷信等元素,可能给青少年带来认知紊乱、现实解离等风险,甚至引发负面情绪。面对“AI精神污染”带来的冲击,我们应该如何应对?



网络图

猎奇AI生成作品“病毒式传播”

前几秒还是平淡无奇的车站日常,随着画风突变,路人变成巨大的蜘蛛,咬住其他路人的脖子不放;前几秒,慈眉善目的老太太还在笑着介绍面前的美食,突然碗里的鸡蛋扭过来,变成一个老人的头颅……近期,网络平台上出现了一大批诡谲猎奇类型AI生成视频,其中不少播放量突破百万人次。

微信公众号“极客公园”总结,这类视频都有着AI生成、10秒左右

画风突变的共性。“不同于人类导演精心编排的恐怖片,AI生成的视频是一种‘无意的恐怖’,没有任何轨迹可以预判。它触发你深层的不适和惊悚感,瞬间让你大脑短路,甚至不敢点击屏幕把视频划走。”

有网友称,这类视频开头一切正常,等发现是恐怖视频时已来不及。虽然网民能察觉到这是反常识的,但还是会忍不住一个接一个看下去,进入一种“恐怖上瘾”的状态,直

到“看得头晕恶心,要出门透透气”。

一些媒体将这些作品称为“克苏鲁式”AI生成作品。这些“克苏鲁式”AI生成作品将人带入疯狂、恶心、诡异的氛围中,使接收者感受到心理上的不安与不适,却又无法抑制地回想这些片段,造成精神污染。

除恐怖类视频外,半月谈记者在多个短视频平台看到,AI生成的“反常识”恶搞视频五花八门:高大的花生树上花生累累、大熊猫

躺在洗发躺椅上由工作人员洗毛发、长颈鹿将脖子打结才能扭转视线看到天空、萌宠开口唱歌或者做家务……

借助AI生成技术,这类恐怖猎奇视频制作简单,方便批量生产。根据一些平台流传的“爆款阴间视频制作教程,小白两步就搞定”等教学视频,在生成软件上输入几句话描述场景,再输入图片,并简单描述动作,最后加入配乐即可。

“AI精神污染”可能带来多重影响

“很难想象在AI生成内容中成长起来的一代人,将来常识会有多混乱。”山东青岛市民靳女士告诉半月谈记者,她5岁的女儿有一天突然说要吃“粉色西瓜”。靳女士以为要吃粉色的西瓜果肉,买回家后女儿却不高兴,称要吃的是“有粉色果皮的西瓜”。

靳女士向女儿解释西瓜果皮是绿色的,女儿却笃定地告诉她:“你没见过不代表没有,我都看到视频了。高高的西瓜树上结着许多

粉色西瓜,西瓜熟了就像椰子那样掉下来。”

靳女士告诉记者,女儿最近突然害怕自己睡觉,称“卧室里的家具会变成可怕的东西”。她查询短视频软件的浏览记录,发现女儿看了好几条AI生成的恐怖短视频。

有网友说:“什么长翅膀的猫、内绿外黄的香蕉,孩子真的信,而且纠正不过来。因为这种视频的评论区总有人瞎起哄说‘我证明是真的,我家就有’。这种东西容易给孩子

造成认知混乱。”

世界经济论坛发布的《2025年全球风险报告》将“错误信息和虚假信息”列为2025年十大风险之一,并指出生成式AI技术使虚假内容生产成本下降90%,传播效率提升400%。

山东理工大学管理学院大数据管理系副教授桑金琰认为,长期高强度接触类似短视频可能导致用户出现现实解离症状,同时部分视频制作者为吸睛而创作负面情绪内

容,可能造成“创伤性记忆”,这对青少年的影响尤甚。

舆论担忧,此类AI生成作品占领内容平台,可能造成网络生态失衡。AI生成内容的生产效率远超人类创作,形成巨大的成本优势,大批量涌入视频平台以争夺流量,业内称为“AI流量赌博”,真正由人类创作的内容越来越难以被刷到。长此以往,“AI幻觉”“AI谣言”等问题可能不断放大,导致出现社会信用赤字。

给AI生成技术划“禁区”和底线

AI生成技术加速发展是必然趋势,也是社会生产力发展的一大积极因素。当务之急是,明确并落实生成“禁区”和底线,严格区分AI内容,助力AI生成技术有序发展。

——加强AI学习和生成机制的源头治理。桑金琰等认为,AI生成软件已成为大众化新型生产工具,但具体使用规则、行为规范仍不完善,“AI道德”尚有许多暧昧地带。

应加强相关软件对源头素材、生成内容的把控责任,明确规定“魔改”题材范围、合成内容使用禁区等,并在实际发展过程中不断细化相关规定。

——加快发展内容隐式标识等技术。2025年9月1日正式生效的《人工智能生成合成内容标识办法》,规定AI生成的文字、图片、音频、视频都必须强制添加显式和隐

式标识等。然而,用户可以通过删除相关文字及标识图形、转码等方式规避;单纯基于标识,也无法准确鉴别恶意发布内容。应进一步发展有安全保证的内容隐式标识技术,利用其鲁棒性与难以去除特点,配合显式标识和元数据隐式标识,共同维护基于标识的内容治理体系。

——平台加设“屏蔽AI”等功

能。不少网友建议,社交软件应设立选项,用户可以根据需要一键屏蔽AI生成内容。在“青少年模式”中,谨慎推送相关内容,避免暴力血腥、恐怖迷信等元素,并以更加醒目的形式标注“AI生成”,对其中反常识或不符合社会基本认知的内容进行解释说明,提升青少年对于AI生成内容的辨别能力。

(据《半月谈》)