

观点
扫描



闹剧已经散场,但反思不应停止。为何这么多人被流量冲昏头脑?“谁蹭谁红”的逻辑怎么还能大行其道?还会出现下一场“群魔乱舞”?……一系列问号需要拉直,无不提醒我们必须重新审视“流量为王”“流量至上”。



要对抗“AI代写”的黑灰产,需要监管者更快升级技术工具,同时,建立更加严格的行业性规范。一方面,用“AI来检测AI”;另一方面,要对“AI工具如何使用”建立明确的标准。



小天才电话手表粗暴设置敏感词并非孤例,它实际上代表了一种普遍现象。近年来,诸多社交平台加大了利用人工智能进行内容审核的力度,而这些人工智能又不够聪明,结果导致笑料频出。

无底线“蹭流量”该消停了

随着郭有才从一个草根歌手一炮而红成为顶流主播,菏泽南站被“带火”得一塌糊涂——一大批自媒体蜂拥而至,有的奇装异服、扮丑搞怪,有的表演低俗、哗众取宠,称之为“群魔乱舞”也不为过。

这场“流量的狂欢”在舆论场引发轩然大波,也并非无迹可循。从“拉面哥”家门口被围堵,到全红婵老家被打卡者“包围”,再到炮制“秦朗丢作业”热点等,近年来线上线下闻流量而动的自媒体大有人在,各种热搜话题也是轮番登场。冲着流量和变现,一些自媒体博主无所不用其极,在道德和法律的边缘疯狂试探。这次还有人搞出了“北大破格录取郭有才”的谣言,简直“滑天下之大稽”。

虽然如此,把板子全部打到郭有才或菏泽当地头上,也失之偏颇。在察觉到事态或走向失控之后,5月20日,菏泽南站发布公告称,鉴于考试比较集中,将不再举办各类文娱活动,并在现场增加了执法人员维持秩序,对直播的非正能量网红进行劝阻劝离。5月22日,郭有才在直播后宣布,将

停播三天“去学习学习”,网友点赞“这才是正确态度”。

闹剧已经散场,但反思不应停止。为何这么多人被流量冲昏头脑?“谁蹭谁红”的逻辑怎么还能大行其道?还会出现下一场“群魔乱舞”?……一系列问号需要拉直,无不提醒我们必须重新审视“流量为王”“流量至上”。无论是个体还是城市甚至整个社会,一旦丧失理性思考、降低自我要求、放松底线原则,必将掉入“流量陷阱”,带来难以逆转的影响。

人们乐见线下空间的规范、清理,但还有一个绕不过的问题是,在这些事件中的直播、短视频平台又该如何应对?今年4月开展的“清朗·整治‘自媒体’无底线博流量”专项行动,明确提出加强重点平台和重点环节管理,加强自媒体账号全流程管理,以及完善流量管理措施,侧面反映出在优化算法模型和推荐机制方面,平台大有可为,做得还远远不够。对于那些纯粹为了博眼球的自媒体,也是时候拿出真招实招,让它们消停消停了。

陆夷 来源:南方日报

应对AI论文造假,不妨“以AI检测AI”

论文代写,一直是学术作弊黑灰产中的隐疾。而随着人工智能时代到来,AI又成了其中得力的生产工具。近日,据新京报卧底调查,在论文代写产业链中,黑灰产团队已经开始通过AI工具来批量编造论文,并在各大平台接单生产售卖。

报道中值得注意的细节是,有代写团队成员向暗访记者表示,因为AI代写工具的出现,使得论文代写的门槛降低,甚至出现了“职高生用AI工具帮博士生代写论文”的现象。

在AIGC(人工智能生成内容)工具出现之初,行业普遍认为人工智能将能够提升内容生产效率,但或许令人意想不到的,AI工具暂时还没有来得及在严肃的学术研究生产中得到广泛应用,却率先成为了黑灰产业提高产量的工具。

而要对抗“AI代写”的黑灰产,需要监管者更快升级技术工具,同时,建立更加严格的行业性规范。一方面,用“AI来检测AI”;另一方面,要对“AI工具如何使用”建立明确的标准。

当前,国内外已经开始研发专门针对“AIGC内容”的检测工具。在部分内容平台上,已经开始对疑似“AI生成内容”进行了异常标注。从数据检测识别的逻辑看,这类AI造假内容的特征相对更加明显,通过对文本结构、用词与语言模型等进行检测,应当可以识别察觉出AI的创作痕迹。

因此,在传统的抄袭查重系统之外,至少目前,可以尽早将针对“AI内容”的检测工具也应用到学术论文的检测流程之中,并对疑似AI造假的内容进行标注,辅助审核人员进行判断。

在技术性防范手段之外,针对AI论文造假,学术界也可以迅速行动起来,为AI工具的应用设定清晰的边界。例如,可以使用AI辅助制作插图,但不得进行任何文本、数据生成。

当然,随着AI被应用于检测预警,以及行业共同规范的建立与调整,AI造假的漏洞终将被迅速填平。从长期来看,人工智能仍然是创新的朋友,而不是造假的同伙。马尔文 来源:新京报

别让儿童电话手表违禁词“逼疯”家长

据报道,近日,有部分家长反映,自己通过小天才电话手表App跟孩子聊天时,“我是你妈”等表述被系统提示“可能含有不适宜内容”,令人哭笑不得。

对于有经验的网友来说,遭遇这些毫不奇怪,肯定是因为家长在使用过程中,其发布的内容触碰了电话手机系统后台设置的敏感词,导致被屏蔽。然而,即便在保护儿童这个超级正确的目标下,出现这种明显荒诞的结果,仍显得不正常。

厂商对儿童手表的内容进行干预可以理解,但干预总要有个限定边界。厂商并非监管部门,这种明显影响到用户日常使用的干预行为,已经侵犯了消费者的合法权益。

对此,厂商官方客服表示,“是想创造一个绿色的沟通环境,一些敏感词汇或者不文明用语都是发送不出去的。”但从实际情况看,或许“绿色”在表面上实现了,但“沟通环境”却遭到严重破坏,未免本末倒置。

就像一些网友评论所说,把内容审核都交给无脑屏蔽系统,厂商一可以节约成本,二可以推卸责任,但牺牲的却是用户体验。

而小天才电话手表粗暴设置敏感词并非孤例,它实际上代表了一种普遍现象。近年来,诸多社交平台加大了利用人工智能进行内容审核的力度,而这些人工智能又不够聪明,结果导致笑料频出。

内容审核事关内容安全和用户体验两个重要方面,以目前的AI技术水平,想靠技术手段简单解决的结果很可能就是两个方面都做不到。

从目前情况看,市场主体对互联网内容进行干预必不可少,但在干预的同时应兼顾互联网用户体验,同时亦应时刻牢记干预的目的所在。如果忘却初心而将之变成一种应付程序和形式主义,类似发生在儿童手表上的笑话肯定还将层出不穷。

信海光 来源:新京报