看似“人工智能”,实则“人为陷阱”

——揭开借助AI技术实施诈骗的新套路

新华社北京3月14日电　3月14日,《新华每日电讯》发表题为《看似“人工智能”,实则“人为陷阱”——揭开借助AI技术实施诈骗的新套路》的报道。

号称可提供ChatGPT服务,实际是冒牌AI;视频电话中熟悉的“亲友”,是不法分子AI换脸模拟而成;为博取流量,用AI技术编造虚假谣言,流量变“流毒”……

随着人工智能的发展迭代,生成式AI正以前所未有的速度重塑着日常生活,但由此衍生出的新骗局、新套路也在刷新人们对风险防范的认知。这背后不仅关乎广大消费者的财产安全,还潜藏着人身安全风险和隐患。记者梳理近期管理部门查处的AI相关案件,揭开人工智能“圈套”背后的真相。

提高警惕。新华社发 曹一作

9块9使用ChatGPT?其实是“套壳AI”

自人工智能产品ChatGPT爆火,一些企业看到无限“钱景”,强行“关联”。一时间,市场上涌现出不少与ChatGPT“沾亲带故”的服务产品。2023年2月,一个名为“ChatGPT在线”的公众号引起了上海市徐汇区市场监管部门的注意。

这个头像与ChatGPT原开发公司官方标识高度相似的公众号,在用户短暂免费体验后即需注册会员付费使用,支付9.9元可以对话20次,随着对话次数增加,其收费也逐步提高。该公众号仅用两个月就吸纳超36万人的粉丝量,累计注册付费用户4231人,经营额共计125385.44元。

执法人员调查发现,该公众号的运营公司与实际ChatGPT开发公司并无关联。所谓“ChatGPT在线”也并非“ChatGPT”产品本身。“当事人为实现销售目的,使用类似图像、名称及服务简介等多种手段实施复合性混淆行为,利用‘ChatGPT’热点进行攀附,混淆真实情况,谋取交易机会,获取不当利益。”上海市徐汇区市场监督管理局执法稽查科副科长张琦说,当事人行为违反了《中华人民共和国反不正当竞争法》第六条第(四)项的规定,市场监管部门已责令其停止违法行为,处以罚款,并要求其妥善处置相关消费纠纷。

记者发现,“冒牌AI”的现象并不少见。2023年,百度“文心一言”上市前夕,网上出现大量打着“文心一言”旗号的社交媒体账号。随后百度官方发文辟谣称“文心一言”尚未注册社交账号。

“人工智能大模型开源生态的建立,让AI发展加速的同时也让生成式AI更容易被滥用。”中国行政法学研究会常务理事、华东政法大学教授沈福俊认为,考虑到立法的滞后性,监管部门还需更加关注新技术发展对市场端的影响,加大事前监管力度,充分运用现有的法律法规资源实施有效监管,切莫使新问题隐匿在“监管盲区”。

张琦也表示,消费者出于对新兴技术的好奇,往往在不知不觉陷入“圈套”。“如果误用了别有用心的‘套壳AI’,甚至可能被不法分子套取个人信息,埋下安全隐患。”他建议,消费者在选择AI产品时需仔细甄别,避免被商家误导。“如果购买到了仿冒产品,要保留好相应的产品购买凭证,主动联系监管部门,维护自身合法权益。”

“AI换脸”诈骗　眼见不一定为实

只需一通视频电话,不法分子就骗走了430万元。2023年4月,福建省某科技公司法定代表人郭先生接到“好友”的微信视频通话,对方声称自己当下需要430万元保证金用以项目竞标,想借用郭先生公司的账户“走个账”。因为有先前的视频通话,加之对“好友”的信任,郭先生陆续给对方转账共计430万元。随后郭先生再次联系好友时才发现自己被骗。所幸在警方帮助下,成功止付拦截336.84万元。

据郭先生回忆,由于在视频电话中确认了对方的面孔和声音,所以毫不怀疑对方身份有诈。不仅如此,犯罪分子还精准了解郭先生与好友的关系,并成功盗取好友微信账号实施诈骗,令人不寒而栗。

AI技术的迭代升级让不法分子借助智能AI换脸和拟声技术,就可轻松实现远程视频诈骗。记者调查发现,最近多地出现的AI换脸诈骗案件,均具备定制性、迷惑性等特征。

其中,香港警方近期披露一起涉案金额高达2亿港元的多人AI换脸诈骗案尤为典型。据媒体报道,某公司职员受邀参加公司“多人视频会议”时,先后将2亿港元分别转账到5个本地银行账户内。据警方调查,这场视频会议中除了受害者外,其余均为AI换脸后的诈骗人员。

不少业内专家表示,随着文生视频大模型Sora等多模态人工智能的探索和出现,人们可能陷入“眼见也不一定为实”的困局。“一直以来,银行等部门将实时视频用做检验身份的手段之一,如今其可靠性将面临巨大挑战。”沈福俊说,“随着人工智能技术的迭代升级,这类违法行为还可能演变出更多形态。”

沈福俊坦言,对此类AI换脸和AI拟声的恶性诈骗案件,若只依靠传统监管手段已不能防堵,监管部门在提升自我科技能力储备的同时,有必要引入新兴技术,探索用AI技术监管AI的可能。

从消费者角度出发,沈福俊建议提升全民个人信息保护意识,谨防隐私泄露。“不管是在互联网上还是社交软件上,尽量避免过多地暴露自己的信息,在涉及转账交易等行为时,可以多角度询问身份信息,反复验证对方是否为本人。”

警惕“AI造谣”　小心“流量”变“流毒”

当前AIGC技术已在文本生成、图片创作等方面广泛应用,输入几个关键词即可由AI快速生成一张画或一篇文章。但一些用户为博眼球、蹭热度,却将人工智能技术用以编“伪消息”,造“假通报”。

2023年6月,一条名为《浙江工业园现大火浓烟滚滚,目击者称有爆炸声!》的视频在网络上流传,引发网友关注。后经相关部门核实为谣言。据调查,当事人为给自己账号涨粉引流,获取更多利益,通过非法渠道购买了AI视频生成软件。该当事人将网络热门话题通过AI自动生成视频产品,并上传至多个热门视频平台。

截至案发,相关当事人发布的虚假视频多达20余条,涉及浙江、湖南、上海、四川等多个省市,累计阅读观看量超过167万次。目前浙江绍兴上虞法院已开庭审理并当庭宣判了这起案件,两名被告均被判处有期徒刑。

但“AI谣言”仍时有发生。2024年1月,广西东兴市骆某某为博取流量,将其他地区的抗洪、救灾视频,经AI软件自动编辑,编造新疆乌什县发生地震的虚假视频信息。同年1月,四川一网民在某平台发布“贵州女婴被弃”的谣言文章。经调查,该涉谣文章由AI系统生成并发布,相关言论及照片均为不实信息。

技术的普及带动自媒体产业的繁荣。但一些网民为获取流量不惜利用技术手段编造生成虚假视频,让“AI谣言”在网络上传播,不仅给网络安全带来严峻挑战,也严重扰乱社会秩序。

2023年,最高人民法院、最高人民检察院、公安部联合发布的《关于依法惩治网络暴力违法犯罪的指导意见》中规定,对“利用‘深度合成’等生成式人工智能技术发布违法信息”的情形,依法从重处罚。

对此,沈福俊等多位专家表示,一方面,要提高广大自媒体经营者的法律意识,从源头减少此类“AI谣言”的产生;另一方面,监管部门也要加强相关案件的宣传推广,提高网民对“AI谣言”的鉴别力。

印刷:茂名市名升印务有限公司　地址:茂名市电白区(原茂港区)进港大道128号　广告热线:2963992 2963993　茂名日报社广告发布登记通知书编号:440900100009