BEST SELLING 2024
中国邮政发行畅销报刊

老年日报

总9160期 今日8版

生活 SHENGHUOZHOUKAN 周刊

国内统一连续出版物号：CN23-0018
邮发代号：13-18

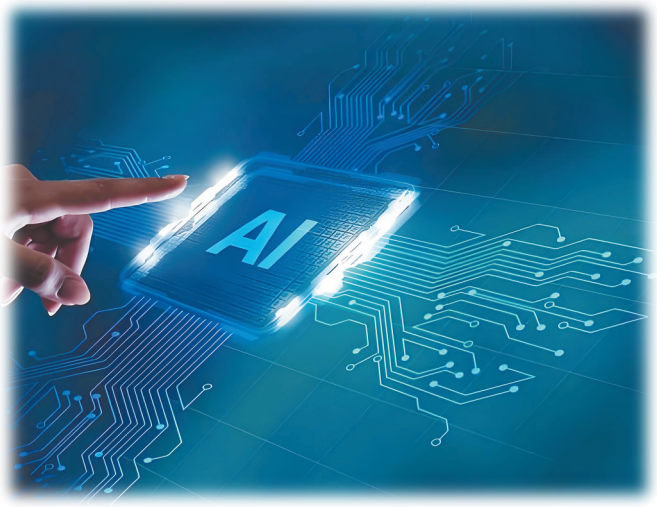
2024/07/20
星期六
甲辰年 六月十五

中国百强畅销报刊 年长者的精神家园

北京、广州、哈尔滨、西安、南京、济南、郑州、石家庄同步印刷

让造谣更简单、谣言更“科学”？

AI谣言乱象调查



“有图有真相，这都是专家认证过的。”

近日，天津市民李萌（化名）与母亲就一篇“科普文章”发生了激烈的争吵：母亲坚信文章中有视频、有图，还有各种所谓博士、医疗团队得出的研究结论，不可能是假的；李萌仔细辨别文章发现，文章为AI生成，平台也进行了辟谣，肯定是假的。

这篇文章的内容与猫有关——有一名女生抱着猫玩，得了一种绝症叫“毛病”，后期整个人变得面目全非。也正是因为这篇文章，李萌母亲坚决反对她养猫，因为害怕她也患上“毛病”。李萌对此哭笑不得，“真希望我妈能少上点网”。

被AI谣言坑骗的远不止李萌母亲。近期，多地公安机关发布了多起利用AI工具实施造谣的相关案件，如发布“西安突发爆炸”虚假新闻的账号所属机构，最高峰一天能生成4000至7000篇假新闻，每天收入在1万元以上，而公司实际控制人王某某经营着5家这样的机构，运营账号达842个。

专家指出，便利的AI工具大幅降低了造谣的制造成本，提升了谣言的数量级和传播力。AI造谣呈现出门槛低、批量化、识别难等特点，亟待加强监管，斩断背后利益链条。

用AI编造不实消息 传播迅速多人受骗

6月20日，上海警方发布通报，两名品牌营销人员为蹭热度，编造了“中山公园地铁站捅人”等不实信息，相关人员已被警方行政拘留。通报中，有个细节引人注意：一名造假者使用AI软件生成视频技术，编造了地铁行凶的虚假视频等不实信息。

近年来，利用AI造谣现象频发，且传播速度极快，一些谣言造成不小的社会恐慌和危害。

去年，在上海一女童走失事件中，一团伙以“标题党”“震惊体”方式，恶意编造炒作“女孩父亲系继父”“女孩被带往温州”等谣言。该团伙利用AI工具等生成谣言内容，通过114个账号矩阵，在6天内发布268篇文章，多篇文章点击量超过100万次。

公安部网安局近期公布一起案例。2023年12月以来，一条“西安市鄠邑区地下涌出热水”的信息频繁在网络上传播，出现如“地下出热水是因为发生了地震”“是因为地下热管道破裂”等谣言。经查，相关谣言是通过AI洗稿方式生成的。

近日，“济南一高层住宅楼起火，多人跳楼逃生”“晨练大爷在济南英雄山附近发现坟中活人”……这些离谱的“重磅消息”在网上广泛传播，引起大量关注。济南市委网信办第一时间通过济南互联网联合辟谣平台进行辟谣，但还是有不少人被“有图有真相”的表象迷惑。

清华大学新闻与传播学院新媒体研究中心今年4月发布的一份研究报告显示，近两年的AI谣言中，经济与企业类谣言占比最高，达43.71%；近一年来，经济与企业类AI谣言量增速高达99.91%，其中餐饮外卖、快递配送等行业更是AI谣言重灾区。

很多AI生成的谣言中，都夹杂着“据报道”“相关部门正对事故原因进行深入调查，并采取措施进行抢修”“提醒广大市民在日常生活中要注意安全”等内容，在网上发布后人们往往难辨真伪。

除了AI新闻，科普文章，图片、配音视频、换脸后模仿声音，这些都可以利用AI产生，在人工进行微调并融入一些真实的内容后，它们就会变得难以辨别。

中国人民大学新闻与社会发展中心研究员曾持说，“生成式AI”的拼接本质和谣言有很强的亲近性，二者都属于“无中生有”——创造看起来真实合理的信息。“网络平台可以用AI技术反向对图像和视频的拼接进行检测，但很难对内容进行审查。当下人们并没有能力完全拦截谣言，更何况有许多未经证实或无法证实、模棱两可的信息。”曾持说。

为获流量牟利造假 涉嫌构成多种罪名

有些AI软件的“造谣效率”非常惊人。比如，出现了“一天能生成19万篇文章”的造假软件。

据查获该软件的西安警方介绍，警方提取了该软件7天保存的文章，发现总量超过100万篇，涉及时事新闻、社会热点、社会生活等各方面。账号的使用者有组织地将这些“新闻”发布到相关平台，再利用平台的流量奖励制度等获利。目前，该案涉及的账号已被平台查封，有关软件和服务器也被关闭，警方还在对案件进行深挖。

在多起AI造谣事件背后，造谣者的动机主要源于引流和借此牟利。

“用AI批量生产爆款文案，突然你就变有钱了”“让AI帮我写推广文，1分钟搞定3篇”“图文创作，AI自动写文章，单号轻松日产500+，可多号操作，小白轻松上手”……很多社交平台上都有类似“致富”文章流传，评论区还有不少博主在推送。

今年2月，上海公安机关发现，在一家电商平台上出现了某艺人“命运多舛、含恨离世”等短视频，引发大量点赞和转发。

经查，该视频内容系伪造。视频发布者到案后交代，他在某电商平台上经营一家土特产网店。由于销量不佳，他便通过编造夺人眼球的虚假新闻给网店账号吸引流量。他不会视频剪辑，便利用AI技术生成文本和视频。

北京瀛和律师事务所合伙人张强表示，利用AI编造网络谣言，尤其是编造、故意传播虚假的险情、疫情、灾情、警情，可能涉嫌刑法编造、故意传播虚假信息罪。如果影响到个人或企业的名誉，可能涉嫌刑法诽谤罪和损害商业信誉、商业名誉罪。

持续完善辟谣机制 明确标注合成内容

为治理AI造假乱象，深化网络生态治理，近年来相关部门和平台出台多项政策和措施。

早在2022年，中央网信办等就发布了《互联网信息服务深度合成管理规定》，规定任何组织和个人不得利用深度合成服务制作、复制、发布、传播法律、行政法规禁止的信息，不得利用深度合成服务从事危害国家安全和利益、损害国家形象、侵害社会公共利益、扰乱经济和社会秩序、侵犯他人合法权益等法律、行政法规禁止的活动。深度合成服务提供者和使用者不得利用深度合成服务制作、复制、发布、传播虚假新闻信息。

今年4月，中央网信办秘书局发布《关于开展“清朗·整治‘自媒体’无底线博流量”专项行动的通知》，要求加强信息来源标注展示。使用AI等技术生成信息的，必须明确标注系技术生成。发布含有虚构、演绎等内容的，必须明确加注虚构标签。

针对疑似使用AI技术的内容，一些平台会在下方贴上“内容疑似AI生成，请谨慎甄别”的提示，并对包含虚构、演绎等环节的内容，明确加注虚构标签。部分大模型开发者也表示，会通过后台设置的方式，对通过大模型生成的内容打上水印，告知用户。

在张强看来，人们对于生成式AI还没有足够的了解，也缺乏应对经验，这需要有关部门在执法层面加大响应力度，对于通过AI造谣、诈骗等行为及时予以查纠。

中国传媒大学文化产业管理学院法律系主任郑宁认为，应进一步完善现有的辟谣机制，一旦某一条信息被甄别为谣言，要立刻进行标注，并且向浏览过该谣言的用户再次推送，进行辟谣提示，以免该谣言进一步传播，造成更大的伤害。

张守坤

出版：老年日报社 订阅投递查询热线：11185转0 发行部：(0451)84612023 广告部：(0451)84612293 办公室：(0451)84612031