谣言“没完没了”背后竟是AI“洗稿”

人大代表：推动AIGC健康生长需进一步加强监管

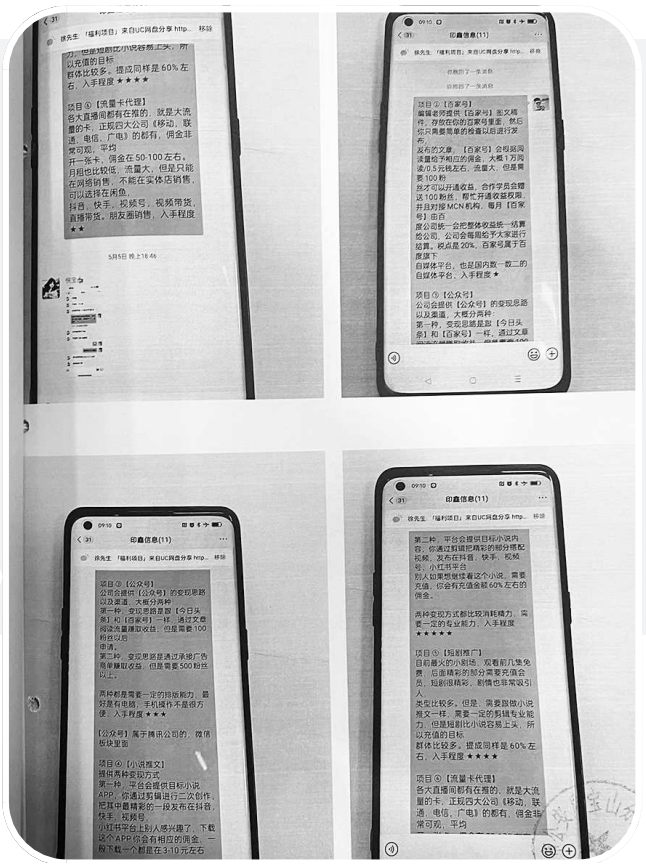
聚焦

□ 见习记者 王葳然

AIGC（生成式人工智能）是近两年刚刚兴起的一种生产力工具，是利用人工智能技术来生成内容的一种新型技术，这一技术让文生文、文生图、文生视频成为现实，然而不法分子却利用这一新兴技术进行非法经营活动。

“本院认为，被告人徐某、罗某某、阚某某结伙，违反国家规定，以营利为目的，编撰虚假信息通过网络有偿提供发布，扰乱市场秩序，情节严重，其行为均已构成非法经营罪，应依法惩处。”近日，经宝山区人民检察院提起公诉，宝山区人民法院对一起利用AI洗稿并非法获利的案件作出宣判，三名被告人分别获刑1年10个月至1年6个月不等，并各处罚金5万元。庭后，记者了解了案件的来龙去脉。

而针对如何推动生成式人工智能健康生长，市人大代表在上海两会期间提出，需要从治理机制构建、技术研究加强、伦理训练与数据集优化、知识产权保护强化等多方面入手。



被告在兼职群中与“学员”的聊天截图

招募网民成为“学员”
发布不实信息非法获利

去年4月，宝山警方在网络上监测到一则不实信息，属地政府报案后，公安机关立即予以立案侦查。据查，发布该信息的账号注册地是云南省，警方便前往当地找到了信息的发布者，并根据线索将幕后的三人犯罪团伙抓获。

“发布该信息的是一位女性王某，没有固定工作，在兼职群里看到发布信息就能挣钱的这样一份兼职，因为也不需要怎么动脑筋，就成为他们的‘学员’，按照指示进行发稿。”宝山区人民检察院检察官陈伟东是本案承办检察官，他告诉记者，三名被告人负责从网上搜找、截取热点文章及新闻，利用从网上购买的一款AI软件进行“洗稿”，并将拥有“今日头条”“百家号”等自媒体平台账号的网民招募为“学员”，而后再利用“学员”的账号发布上述伪原创的文章、新闻；而所谓“挣钱”则是依靠平台流量的返现，非法获利由三名被告人与学员五五分成。

“把账号登进去，然后设置与原文的查重率不高于25%，软件就会自动改写。”法庭上，被告人阚某某交代称，“这样才能在平台上发出来。”

“我们经查明发现，该软件可以实现自动洗稿，把原稿件导入后，可以在其中设置指令要素，不仅限于修改新闻基本的信息要素，还包括‘结论要有争议性’‘内容要吸人眼球’等。”陈伟东说道，“原信息的真实性暂且不论，查重率不高于25%就意味着洗稿后的一篇文章中，有75%都是虚假的。”

记者在采访中了解到，三名被告人自2023年11月至2024年5月期间从事该类非法经营活动，这样的文章每天至少能生成1000篇，共计非法获利5万余元。

三名被告人对犯罪事实供认不讳
法官检察官：此类现象应予以重视

在采访中，记者了解到，最高人民法院、最高人民检察院于2013年曾出台《最高人民法院、最高人民检察院关于办理利用信息网络实施诽谤等刑事案件适用法律若干问题的解释》（以下简称《解释》），其中第七条对于利用信息网络实施非法经营犯罪的认定及处罚问题等进行明确。

“该团伙首先是发现了自媒体的商机，因为有些‘学员’不会对他们提供的信息进行核实，以‘交学费’的名义获得文章，再进行发布，从而利用浏览量变现进行获利，按照《解释》的相关规定，个人对于明知虚假信息提供有偿发布的，个人非法经营数额在5万元以上，或者违法所得数额在2万元以上的，需要按非法经营罪进行处罚。”宝山区人民法院刑事审判庭法官范楠楠对记者说道。

该案的判决也让检察官和法官发现了打击此类行为的必要性。

“网络媒体和自媒体的发展是非常迅猛的，其间AI技术也在不断进步，利用AI工具进行违法犯罪这样的现象也应当引起重视。”同时，范楠楠还提到：“对于这样一种提高人们工作生活便利性的生产力工具，如何正确的使用、如何进行规范管理，是目前可以持续探索的。”

“人工智能的使用目前还处于一个起步阶段。AI软件不是‘原罪’，关键还是使用者是如何进行使用的。”陈伟东表示，“对于一种新兴事物而言，对其监管可能存在一定的后滞性，行政法规也会随之进一步完善。比如说，本案中的洗稿软件是三名被告人从网上购买来的，那对于使用者有没有什么限制呢？自媒体时代人人都可以参与，首先还是使用者要自律，自觉远离违法犯罪的红线。”

此外，陈伟东还提到了平台审核的问题。“自媒体账号需要流量，平台也需要流量，但平台还是要加强人工审核，避免一些不符合常识的信息发布出来。”而法官检察官都提到的一点就是要通过对案例的宣传，让更多公民了解此类不法行为，从而自觉远离违法犯罪。

AIGC时代已经到来
如何管好、用好人工智能这一新质生产力？

为促进生成式人工智能健康发展和规范应用，2023年7月，国家网信办联合国家发展改革委、教育部、科技部、工业和信息化部、公安部、广电总局公布《生成式人工智能服务管理暂行办法》（以下简称《办法》），并于当年8月15日起施行。

记者看到，《办法》第四条“提供和使用生成式人工智能服务，应当遵守法律、行政法规，尊重社会公德和伦理道德”，明确指出生成式人工智能服务的提供和使用者不得生成虚假有害信息等法律、行政法规禁止的内容。同时，应尊重知识产权和他人合法权益、提升生成式人工智能服务的透明度以及生成内容的准确性和可靠性。

同时，如何管好、用好人工智能这一新质生产力，也引起了人大代表的重视。上月，在上海两会期间，市人大代表刘忱带来《关于以科技伦理为导向，科技助力生成式人工智能安全发展的建议》：“在利益驱使下，部分不法分子利用AIGC技术批量制造虚假图片、视频、音频，炮制虚假新闻与谣言，生成带有偏见的评论等信息，使得‘眼见不一定为实，有图不一定有真相’，扰乱舆论生态的同时，也给社会信任体系带来巨大冲击。”

作为相关从业者，刘忱告诉记者，AIGC模型的形成和完善依赖大量数据训练，助力生成式人工智能安全发展需从治理机制构建、技术研究加强、伦理训练与数据集优化、知识产权保护强化等多方面入手，从而推动行业良性发展。

在打击AIGC虚假新闻方面，刘忱也提出了相关建议：“平台机构企业应承担相应的社会责任，加强自律，确保技术产品和服务符合伦理要求，通过事前宣导，强化视频制作者的责任感，并开展违规内容检测，防止传播AIGC生成的违规信息。”具体而言，平台可利用大型语言模型进行文本合规检测，通过视觉模型提取图片特征判断图片信息是否合规。“对于发布内容，可从性质上区分营利性与非营利性。非营利性质的内容需要做好发布前的责任提示；如果内容涉及营利性质，需进一步明确AIGC知识产权归属，细化侵权责任认定，加强数据知识产权保护等。”

此外，刘忱还提到，可引入技术手段对AIGC生成的视频、图像进行版权标识和追踪，为原创作品提供确权服务，维护创作者合法权益。