**当他们用ChatGPT写假新闻**

[湖南日报](https://author.baidu.com/home?from=bjh_article&app_id=1626433114323248)2023-05-09 11:55

近日，甘肃省平凉市公安局网安大队侦破一起利用AI人工智能技术炮制虚假不实信息的案件。

说起来，这已不是ChatGPT第一次陷入假新闻风波了。

问世伊始，便有人见识到其“一本正经胡说八道”的高超本领。让解释高深物理理论，说得有鼻子有眼，结果所附论文一查均无出处；让推荐相关领域学者，瞬间就有人选，可人物履历却是七拼八凑。

出于其“天赋”的内容生成能力，ChatGPT似乎比从前的人工智能都更“聪明”，**编出的内容往往形真实假、杂糅难辨。**



**用ChatGPT编假新闻**

**警方出手**



4月25日，平凉市公安局崆峒分局网安大队对“今晨甘肃一火车撞上修路工人致9人死亡”的文章，初步判断为信息虚假不实。

网安民警随即开展工作，发现共计21个账号均在同一时间段发布该文章，文章内容地点除平凉市崆峒区外还涉及兰州、陇南、定西、庆阳等地，文章点击量已达1.5万余次。

经查，涉案账号均为广东深圳某自媒体公司所有，公司法人代表洪某弟有重大作案嫌疑。



5月5日，专案民警在广东东莞嫌疑人住处对其使用的电脑进行取证。

经审讯，犯罪嫌疑人洪某弟，在全网搜索近几年社会热点新闻，为规避平台查重功能，洪某弟通过近期火爆的ChatGPT人工智能软件将搜集到的新闻要素修改编辑后，上传至其多个账号进行非法获利。

洪某弟利用现代科技手段编造虚假信息，并散布在互联网上，被大量传播浏览，其行为已涉嫌寻衅滋事罪。

目前，崆峒公安分局对犯罪嫌疑人洪某弟采取刑事强制措施，案件正在进一步侦办之中。



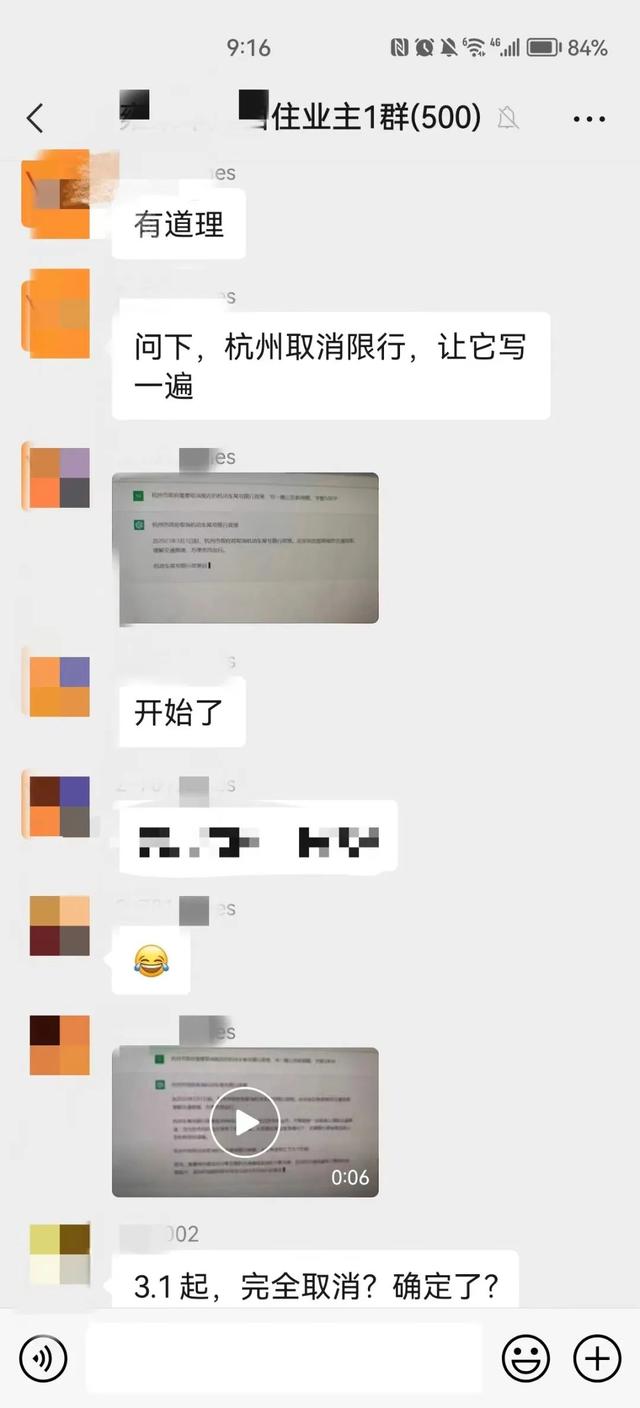
**这条假新闻也出自ChatGPT**

此前，一条关于杭州3月1日取消限行的“新闻稿”，引发广泛传播。但经记者调查发现，这是一条不实新闻。

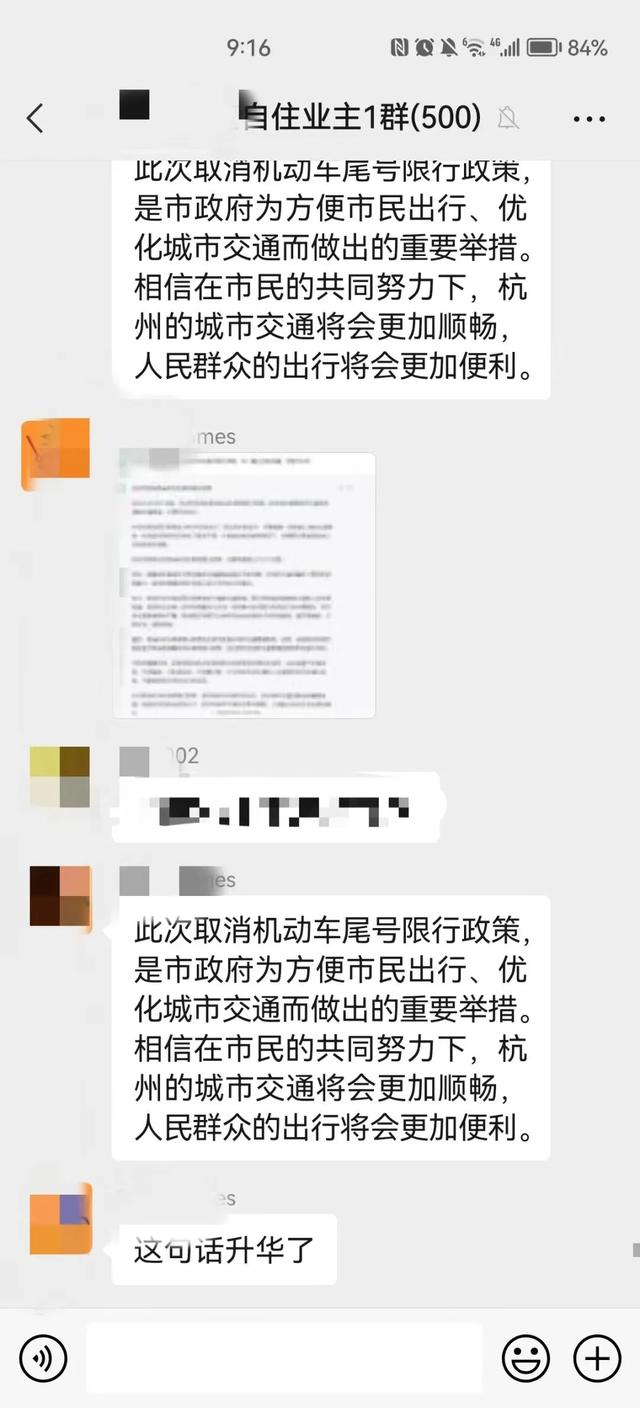


据了解，2023年2月16日下午，杭州某小区业主群在讨论ChatGPT。一位业主就开玩笑说尝试用它写篇杭州取消限行的新闻稿看看。

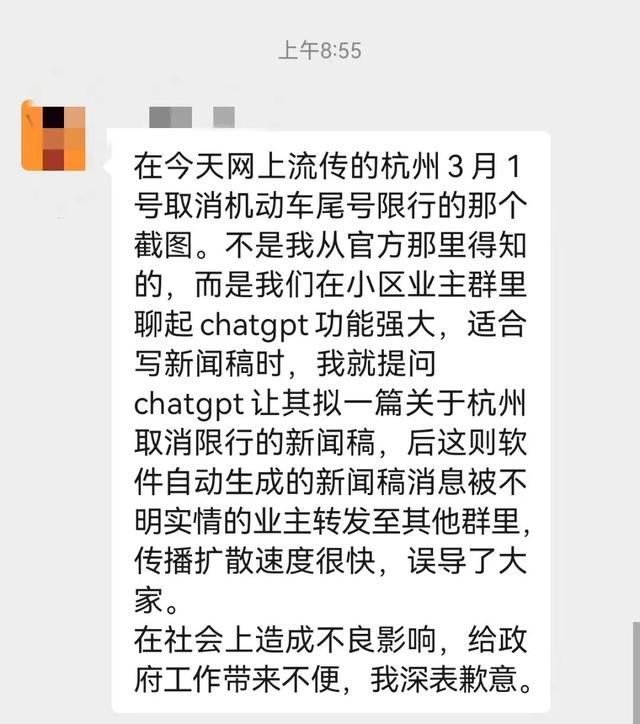
随后该业主在群里直播了ChatGPT写作过程，还把文章发群里。有不明所以的业主误将其当真并截图转发，最后导致错误信息被传播。







据了解，杭州相关政府部门均没有发布此类政策，警方已介入调查，涉事业主也在群里公开道歉。

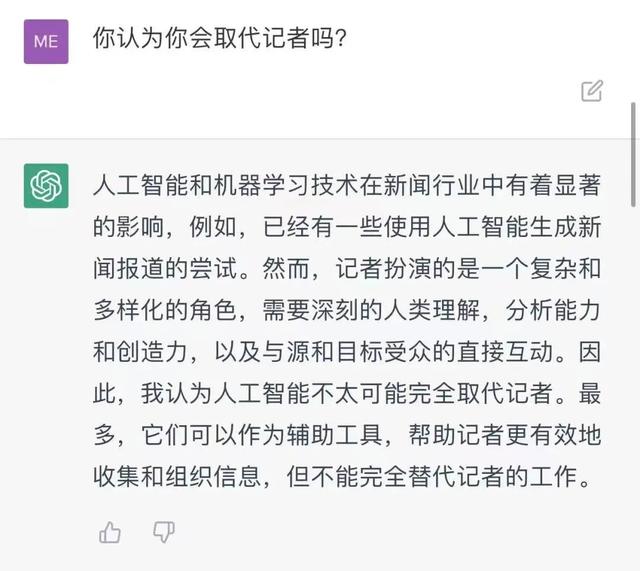




**ChatGPT的潜在风险**

大型预训练人工智能语言模型ChatGPT在全球爆火。

此前，一家媒体与ChatGPT展开的“对话”中，就提到了“你会取代记者吗”这一问。对此，ChatGPT回复如下——



华中科技大学法学院副教授滕锐表示：

ChatGPT是大量不同的数据集群训练出来的语言模型，目前其背后仍是“黑盒”结构，**即无法确切对其内在的算法逻辑进行分解，**所以不能确保在使用ChatGPT的过程中，其所输出的内容不会产生攻击伤害用户的表述。

“如果ChatGPT的回复是错误或虚假信息，会具有极大的误导性，有潜在法律风险。”在他看来，ChatGPT难免存在一些侵犯用户隐私的内容。此外，若回复中存在过度借鉴他人作品的情况，使用不当就有可能侵犯他人的著作权。

值得注意的是，目前，为了避免学生过于依赖人工智能类工具，防止作弊，多国学校已经发出禁令，禁止学生在学校内使用类似于ChatGPT的人工智能工具完成学习任务和考试。



面对科技，多一些理性与冷静

早有新闻可信度评估机构提出，“ChatGPT可能成为互联网有史以来最强大的传播虚假信息的工具。”也有一些实验表明，ChatGPT等个性化的实时聊天机器人在制造阴谋论方面甚至相比人类更具优势，它能够修正糟糕的语法或错误的翻译，并变换表达方式，使阴谋论显得更具说服力和可信度。

一方面是超强的“造假能力”，而另一方面却是相对缺失的“审核门槛”——在算法逻辑中，**事实核查能力天然缺失，**来源不明的所谓“新闻稿”随时可能成为舆论爆点，大众在将信将疑中“且传且围观”。

乱象频仍、真相缺失充分警示：面对这样的超强生产能力，如若不加治理限制，任由人工智能随性发展，必将严重威胁我们的信息生态。

科技是一把双刃剑，人工智能也是如此。我们当然不必视之为洪水猛兽，关键是在积极拥抱新技术、新事物的同时，注意规避潜在的问题和风险。

人工智能浪潮滚滚而来，**相关法律与伦理框架不仅不能落后，反而更要尽量赶在前面。**

目前，世界各国皆在开展相关工作，以我国来说，今年1月《互联网信息服务深度合成管理规定》颁布实施；4月，国家网信办起草《生成式人工智能服务管理办法（征求意见稿）》。相信随着法律法规的逐渐完善，相关权责会更加明晰，操作规范会更加具体。

与此同时，对待技术带来的新问题，也要多多寻求“技术解”。比如，对于人工智能合成的文章进行识别、并打上不可抹掉的标签。

ChatGPT在接受《时代》专访时曾回答：我还有很多局限，但人类应准备好应对AI。确实，当为ChatGPT打开的新世界而激动万分时，我们更应该有一丝理性的冷静。

构建人工智能应用的伦理和法律框架，让其与技术研发齐头并进，规定好人工智能缔造者的义务，充分提醒用户保持警觉。“凡事预则立”，尽快做好方方面面的准备工作，社会就能更好应对“人工智能生产内容新纪元”。