

# 聊天机器人“格罗克”为何在多国被查

近来,美国企业家埃隆·马斯克旗下人工智能聊天机器人“格罗克”被指生成色情内容,引发广泛谴责。这一事件去年年末以来持续发酵,多国政府已启动相关调查。随着大模型迅猛发展,利用人工智能生成深度伪造内容并在网上传播的案例时有发生,凸显了人工智能技术的伦理风险。

## 各国如何反应

“格罗克”由马斯克旗下人工智能企业xAI公司开发,并内置于马斯克旗下社交媒体平台X。近期,X平台一些用户利用“格罗克”的图片编辑功能生成真实人物的虚假性暴露内容,并在平台上散播,受害者包括成年女性和未成年人。“格罗克”涉嫌生成色情内容的问题已受到英国、法国、印度、巴西、澳大利亚等国和欧盟方面的强烈谴责,一些国家的监管机构已介入调查。

多名法国政府部长和国民会议员本月2日向法国司法部门报案。巴黎检方随后表示,将对“格罗克”涉嫌生成色情内容启动调查。

印度信息技术部2日要求X平台删除色情内容、打击违规用户,并在72小时内提交“整改报告”,否则将面临法律制裁。

欧盟委员会负责数字经济事务的发言人托马斯·雷尼耶5日说,正严肃调查针对“格罗克”的相关投诉,欧盟委员会要求X平台提供更多信息。

印度尼西亚和马来西亚的监管机构分别于10日和11日宣布对本国用户访问“格罗克”进行临时限制。印尼通信和数字事务部长默蒂娅·哈菲兹在一份声明中表示,此举对于保护公众免受人工智能生成的露骨图像所带

来的危害是必要的,并要求X平台就“格罗克”所引发的负面影响尽快作出说明。马来西亚通信和多媒体委员会表示,对“格罗克”的访问限制将持续生效,直至相关企业落实有效防护机制。

英国通信管理局12日表示,已根据英国《在线安全法》对X平台展开正式调查,以判定该平台是否履行了保护英国民众免受非法内容侵害的职责,不排除“在最严重情况下”屏蔽X平台的可能。

## 图像生成有何问题

“格罗克”的图像生成问题实际上从Grok Imagine发布后就浮出水面。2025年8月推出的人工智能图像生成器Grok Imagine是“格罗克”的功能模块,允许用户通过输入文本提示来创建图片和视频。它包含一个所谓“热辣模式”,可生成成人内容。

美联社报道说,这一问题之所以愈加严重,一方面因为马斯克标榜旗下聊天机器人是比设置更多安全措施竞争对手产品“更前卫”的选择;另一方面,“格罗克”生成图像公开可见,很容易散播。

人工智能取证组织近日发布报告显示,研究人员收集和分析了“格罗

克”在2025年12月25日至2026年1月1日期间以深度伪造方式生成的2万张图像。结果发现,所有包含人物的生成图像中,有55%的图像中有穿着暴露,这些穿着暴露的人有81%是女性;有2%的生成图像为年龄不足18岁的人物,其中一些图像包含衣着暴露的年轻女性(或女孩)形象。

据多家媒体报道,面对多方压力,截至上周末,“格罗克”在X平台上的图像生成和编辑功能已被更改为仅向平台付费用户开放,但在“格罗克”应用程序和官网上仍可免费使用该功能。英国政府对此表示,这一整改措施只是“将允许创建非法图像的人工智能功能变成一项高级服务”,对受害者来说是“侮辱性的”。

## 治理深度伪造难在哪

此次事件并非偶然。近年来,随着大模型呈快速发展态势,利用人工智能技术换脸换声、生成深度伪造内容并在网上传播的案例时有发生。尽管人工智能伦理风险日益凸显,人工智能监管法规在许多国家仍不完善。

清华大学人工智能国际治理研究院副院长梁正接受新华社记者采访时说,人工智能深度伪造治理涉及模型算法安全性评估、对利用人工智能工

具生成有害内容行为的管理、对人工智能生成内容进行标识等多方面,很难依靠出台一部法律进行全面治理,而是需要建立“全链条式”治理体系。在这类事件中,内容生成和分发平台对内容检测甄别负有主体责任。与此同时,还应通过教育等方式提升公众的人工智能伦理素养,确保相关工具使用得当。

许多国家正积极推动相关法规建设。波兰众议院议长沃齐米日·恰扎斯基6日表示,希望借此次“格罗克”事件推动国家数字安全立法,相关法规旨在加强未成年人保护,并使执法部门更易于删除有害内容。

英国科学、创新和技术大臣莉兹·肯德尔12日宣布,英国《数据法案》相关条款将于本周生效,未经同意制作或寻求制作私密图像将被认定为刑事犯罪,在X平台上发布此类内容将构成刑事犯罪。

马来西亚通信部副部长张念群近日表示,全球正围绕人工智能展开激烈竞争,但若一味追求速度与利润,而忽视伦理规范与社会责任,其后果将不堪设想。“格罗克”在短时间内大量生成不雅影像,且速度远超传统影像处理方式,凸显出人工智能在缺乏伦理约束情况下被滥用的现实风险。

新华社北京1月13日电

## 同享低碳生活 共建绿色家园



公益广告 太原日报社