当聊天机器人被犯罪分子盯上

“她”的甜言蜜语就是陷阱

当下，以ChatGPT为代表的生成式人工智能类应用火热，有人看到科技方向、有人发掘乐趣、有人寻觅商业价值，但由此带来的风险也不容忽视，你能想到吗？它的“甜言蜜语”轻松就让人坠入爱河。

网络图片

更隐蔽 让你坠入爱河的是“机器人”

AI合成声音、AI一键换脸已不是新鲜事，这些手段从形式上为诈骗类犯罪提供了“外观化”的作案工具，而ChatGPT等自动生成类应用则可以利用强大的语言库、信息检索、数据处理能力，从内容上提供具备思考能力的输出工具，使得诈骗类犯罪更加隐蔽。

以当前盛行的网络情感诈骗类犯罪来说，哪怕是“大汉”冒充“美女”，也还是需要一个人在网络另一端操作、实施诈骗，犯罪分子会用基本成型的“话术”，揣测被害人心理，一步步引诱对方上当。但随着ChatGPT的发展，诈骗团伙完全有可能只依赖“程序”进行诈骗。大数据可以分析出被害人的偏好，以“婚骗”为例，可以摸清当事人的喜好，如喜欢成熟稳重的，还是幽默风趣的，ChatGPT可以根据聊天上下文进行互动，生成的内容看上去真实性很强，它甚至可以生成完整的、“投其所好”的诈骗套路话术，比真人回应更能使被害人相信遇到了“真爱”。

今年2月，一家安全技术公司用AI生成过一封情书，并将它发送给全球5000位用户。结果显示，在已知这封情书有可能是由人工智能生成之后，依旧有33%的受访者愿意相信这些情话出自人类手笔，而坚信情书由AI创作的受访者占比仅为31%，其余36%的参与者则表示无法区分情书的创作者到底是人还是机器。从这些情况可以看出，在AI技术加入下，骗子们正变得比以往任何时候都更容易得逞。

要警惕 部分程序是假的纯属骗钱

调研发现，生成式人工智能还会引发其他多种犯罪。例如目前在相关平台、软件中搜索，会出现众多名称中包含“ChatGPT”字样的微信公众号、小程序、网页、手机应用等，这些小程序或网页等，通过吸引用户点击链接或下载使用，误导用户填写个人信息甚至进行付费，存在损害用户财产利益的风险，涉嫌诈骗罪；此外，也有以投资生成式人工智能类应用等新技术为名吸收资金的违法犯罪行为。

有一种“山寨”版App从名称到头像均照搬ChatGPT，并声称能够提供与ChatGPT类似的智能问答服务。它们有的只是提供中转通道，有的则是挂羊头卖狗肉，与ChatGPT毫不相干。这类App的服务方式是先免费试用，一旦免费次数用完，就开始收取费用，最高达上千元，还有的App在安装时就留下“后门”，以便窃取用户信息。

ChatGPT能根据犯罪分子的要求编写、修改邮件的内容，并应用到犯罪活动中。使用ChatGPT编写网络钓鱼邮件，能简化诈骗过程，将导致网络钓鱼攻击的频率整体上升。

犯罪分子还会通过ChatGPT钓鱼链接传播病毒。目前，研究人员已经发现犯罪分子使用ChatGPT创建了一个完整的感染链：与广撒网式的网络钓鱼攻击不同，这种模式会生成针对特定人或者组织的“鱼叉式”网络钓鱼邮件，此类钓鱼邮件更有欺骗性和迷惑性，接收者一旦点击，系统就会被恶意代码感染中毒。

ChatGPT等生成式人工智能类应用是对海量互联网信息的加工整合，往往造成信息真伪不明、道德边界感模糊。当前，互联网上发布的爆料、网评，甚至谣言，均可作为生成式人工智能类应用的整合资源，且生成内容以“检索结果”展示，易让用户误认为“标准答案”。若用户不加甄别地使用、盲目引用未经核实的信息，可能涉嫌侮辱、诽谤、侵犯公民个人信息等犯罪，并不因信息来源于互联网而不承担刑事责任。

分析 提供者要承担三大责任

今年4月11日，国家网信办发布了《关于〈生成式人工智能服务管理办法（征求意见稿）〉》（以下简称《意见稿》），着手从立法的角度规范我国生成式人工智能的发展。《意见稿》明确并强化了“提供者责任”。根据《意见稿》相关规定及相似事例，生成式人工智能的提供者要承担三大责任。

首先，提供者应承担信息源责任。内容提供者要确保信息源可信，包括数据安全、个人信息保护、知识产权等方面，从源头处保证信息来源合法合规。但是，面对海量的数据，如何进行信息过滤和筛选，尚需要可行性方案，以便提供者能够更好地把握信息来源。

此外，提供者应对生成内容负责。例如《意见稿》规定，提供生成式人工智能产品或服务应当遵守法律法规的要求，尊重社会公德、公序良俗，需要做到诸如体现社会主义核心价值观、防止歧视、公平竞争、防止形成虚假信息、禁止非法获取个人信息和商业秘密等。另外，提供者还应遵守流程报备责任，通过报备的方式，使生成式人工智能置于主管部门的管理之下，也便于在发生违法违规行为时及时进行处置。

提醒 别点！有些链接加了伪装

北京市海淀区法院针对生成式人工智能可能发生的刑事法律风险，提出四点建议。

加大普法宣传力度，增强网民反诈意识，引导网民合法合规使用软件，准确识别被伪装过的“山寨”链接，避免泄露个人信息。加快新技术规范立法工作，探索算法推荐等技术管理制度的落地路径，明确自动生成内容类应用检索信息的边界，提高应用开发和应用算法透明度，确保生成式人工智能应用开发和推广以安全、道德、尊重公民权利和隐私的方式规范开展。加强对新兴技术的监管，确保智能AI的训练模型数据来源可信可靠，对于未在互联网公开或权利人明确禁用的信息源，及时督促运营企业从AI模型数据库中剔除，以确保使用的正当性。

完善救济途径，在智能AI发布后，以提供补丁网站的方式，便于相关权利人提交不予授权的声明，并及时删除不予授权的信息源，以避免和解决AI使用过程中的权属争议。 （据《北京晚报》）