



内蒙古自治区消费者权益保护服务中心指导

你刷的短视频，标注“AI生成”了吗？

“AI生成画面”“疑似使用了AI生成技术，请谨慎甄别”“内容为AI技术制作”……不知你在浏览信息、刷短视频时是否注意到，一些图文、视频内容加上了这样的标识信息。

前不久施行的《人工智能生成合成内容标识办法》（以下简称《标识办法》）提出，对生成合成内容强制添加显式和隐式标识等规范要求，即用AI生成的每一段文字、每一张图片、每一条音频、视频，都必须强制标明“数字身份证”。

亮明身份

“给用AI生成的内容加上标识，我觉得很有必要。”来自湖北武汉的刘女士表示，她每天都会花上一两个小时来刷短视频，以获取资讯、娱乐等。之前，她看过一些内容怀疑其真实性，但又无法核实，最近，刘女士发现她刷到的一些短视频在画面下部加上了AI生成的提示信息，“这就非常清楚，一目了然。”

中国互联网络信息中心最新发布的《生成式人工智能应用发展报告（2025）》显示，截至2025年6月，我国生成式人工智能用户规模达5.15亿人。随着AI技术的迅猛发展和生成式人工智能的快速迭代，各种纷杂的人工智能生成信息让人眼花缭乱，有时难辨真假，这也给监管部门和平台治理提出了新的挑战。

今年9月1日，国家网信办、工业和信息化部、公安部、国家广播电视总局4部门联合发布的《标识办法》正式施行，该办法明确所有AI生成的文字、图片、视频等内容都要“亮明身份”。作为配套，强制性国家标准《网络安全技术 人工智能生成合成内容标识方法》（以下简称《标识方法》）也于9月1日同步实施。

具体怎么“亮明身份”呢？《标识办法》明确，人工智能生成合成内容标识主要包括显式标识和隐式标识两种形式。

显式标识是指在生成合成内容或者

交互场景界面中添加的，以文字、声音、图形等方式呈现并可以被用户明显感知到的标识，就像人脸一样可以快速区分，让普通用户“一眼可见”。如文本内容需在开头、结尾或关键位置标注“人工智能生成”或“AI生成”，字体需清晰可辨，不得刻意缩小或模糊等。

隐式标识更像指纹，是指采取技术措施在生成合成内容文件数据中添加的，不易被用户明显感知到的标识。它侧重于技术层面的追溯，要求在AI生成内容的文件内部嵌入元数据，且字段名称必须包含“AIGC”标识符号。元数据需涵盖AI生成确认状态、生成服务提供方信息、内容传播平台名称、唯一编号以及数字签名或哈希校验值等关键信息，为内容溯源与责任认定提供技术支撑。

中国电子技术标准化研究院副院长范科峰认为，《标识办法》和《标识方法》协同发布，开创了技术标准与管理制度配合实施的新型治理模式。作为我国人工智能领域首部“办法+强标”组合拳，其在管理维度明确生成、传播、分发全链条主体责任，在监管维度建立审核、监测、追责的闭环机制，在技术维度规范显隐标识体系并预留创新空间，这种三位一体的治理架构为人工智能安全治理提供了清晰的实施路径。

打通链路

专家指出，《标识办法》的另一重意义在于，打通全链条治理路径，科学设定主体责任义务，对不同主体设置了不同义务。

例如，传播平台需要核验文件元数据中是否含有隐式标识，对于含有隐式标识的，应当采取适当方式在发布内容周边添加显著的提示标识。应用分发平台则需在应用程序上架或上线审核时，核验服务提供者是否按要求提供生成合成内容标识功能。

在中国信息通信研究院政策与经济研究所副所长李强治看来，《标识办法》的内容标识要求覆盖设计开发、上架分

发、用户使用、内容传播和责任追溯等全链条，依据生成合成内容服务提供者、内容传播服务提供者、互联网应用程序分发平台、用户等不同主体在潜在风险构成中的作用和合规能力，科学划分责任义务，实现生成合成内容上下游链条的全覆盖。

据了解，许多大模型厂商已经开始对AI生成合成内容添加标识。如DeepSeek发布公告称，已在平台内对AI生成合成内容添加标识，并提醒用户相关内容由AI生成。用户不得恶意删除、篡改、伪造、隐匿此类生成内容标识，不得利用AI制作、传播虚假信息、侵权信息以及从事任何违法违规活动。

今日头条、快手、小红书、哔哩哔哩、喜马拉雅等平台也发布公告，创新探索内容标识制度落地方案，落实双标识、标识检测等义务，并提供“作者声明”等功能，激励人工智能向善应用，保障正常合规使用AI工具。

以抖音为例，目前平台已经上线两项核心功能：一是AI内容标识功能，协助创作者为AI内容添加提示标识，方便用户区分辨别；二是AI内容元数据标识读写功能，可识别并写入AI内容的元数据信息，为内容溯源提供技术支撑。

在抖音注册的账号若制作、发布AI内容，需使用抖音提供的标识功能主动添加显式标识；同时，抖音将核验、检测站内未主动标识的内容是否由AI技术生成，对疑似AI内容或元数据中已包含AI内容标识的作品，补充添加显式标识。此外，对于在抖音发布的所有AI内容，平台还将同步为其添加或更新隐式标识，该标识包含内容传播要素等关键信息，进一步完善AI内容的全链路管理。

规范发展

生成式人工智能、深度合成等新技术快速发展，为生成合成文本、图片、音视频等提供了便利工具，在促进经济发展的同时，也造成虚假信息传播、破坏网络

生态等问题。

例如，一些未明确标识的AI生成合成的视频正扮演着“数字孙子”角色，通过提供情绪价值，将许多不明真相的老人哄得心花怒放，部分老人不顾家人劝阻，执意购买这些“孙子”在视频里推荐的产品，最终陷入骗局。

国家网信办有关负责人表示，《标识办法》聚焦人工智能“生成合成内容标识”关键点，通过标识提醒用户辨别虚假信息，明确相关服务主体的标识责任义务，规范内容制作、传播各环节标识行为，以合理成本提高安全性，促进人工智能在文本对话、内容制作、辅助设计等各应用场景加快落地。同时，减轻人工智能生成合成技术滥用危害，防范利用人工智能技术制作传播虚假信息等风险行为，推动人工智能健康有序发展。

专家指出，作为人工智能治理的重要一环，《标识办法》既与《中华人民共和国网络安全法》一脉相承，又与《互联网信息服务算法推荐管理规定》《互联网信息服务深度合成管理规定》等有效衔接。《标识办法》的实施有助于推动AIGC行业的规范化发展，倒逼企业建立合规体系，通过标识这一技术合规前置的制度设计，促使企业将伦理考量融入技术研发全流程；同时通过建立可信的内容标识体系，提高内容使用的透明度。

从国际上看，各国正积极探索建立AI生成合成内容的标识制度，以应对新形势下信息内容治理带来的挑战，生成合成内容标识已成为国际惯例。

浙江大学教授王春晖表示，鉴于AI生成合成技术可大规模产出与人类创作高度相似的内容，多数国家通过立法强制性要求企业履行内容标识义务，通过技术溯源以及显式和隐式标识，确保公众能快速识别AI生成合成内容的来源，抑制虚假信息的泛滥。

（据《人民日报》记者/叶子）

留心网购中的“数据刺客”

当前，数字化浪潮席卷全球，电子零售新业态蓬勃发展，只需使用手机线上选物下单，通过人脸识别完成支付验证，就能获得商品送货上门、即买即用等服务，使日常生活更加高效便捷。消费者在享受便利的同时，也应提高安全防范意识，提防“数据刺客”暗中窃密的多重手段。

安全隐患需留心

——寄送定位关联点位画像。通过APP、小程序等渠道，由手机下单将物品配送到指定位置的服务正逐渐流行。然而，部分线上商超在提供配送服务时，为获取买家的精确配送位置，可能会在配送地图上自动标注出一些敏感区域和点位，并根据消费频次、商品偏好、社交图谱等信息进行关联分析，形成该地区或点位的

用户画像。这些看似普通、用于正常经营的信息数据，一旦被窃取，可能成为境外间谍情报机关或个别别有用心之人进行窃密渗透、技术攻击等不法行为的突破口。

——刷脸支付窃取生物信息。随处可见的自动贩卖机，可通过扫码、“刷脸”等方式购买商品，即买即取、操作简便。然而，此类设备通常搭载录音录像、生物识别模块，并配置有动态人脸识别、红外成像等技术，能够采集周边往来人员面部特征和使用者的虹膜纹路等生物识别信息。个别设备开启刷脸功能获取人脸数据后，还会进一步要求用户提供手机号、性别、出生年月等个人隐私信息，若其后台连接的信息系统存在安全漏洞，可能被攻击者入侵，进而引发数据泄露风险。

筑牢数据安全防线

电子零售行业的快速发展让便利生活触手可及，但面对其中潜在的风险隐患，我们应当提高警惕，自觉筑牢安全防线，对埋伏其中的“数据刺客”“见招拆招”，悉心守护数据隐私和国家安全。

——对合法经营划底线。电子零售行业从业者要时刻以规章制度为底线，明确数据采集的合法边界，杜绝以用户授权、促销需要为名过度收集地理位置、生物特征等敏感信息，严禁将用户数据信息用于非必要场景。同时，要提升技术防护能力，定期开展数据安全审查，防止数据被截取或植入后门程序。

——对可疑要求说拒绝。消费者在租赁、购买电子零售产品时，要增强隐私保护意识，对于产品供应方提出的不合理

要求，如过度采集生物特征信息等，以及含有“信任此设备”“开启USB调试”等内容的可疑授权提示时，应当谨慎操作。

——对敏感数据多留心。要强化对数据安全的敏感性，面对日常生活中一些看似寻常的数据信息，如分布于重点位周边的坐标指向、人员分布、面部信息等，都应引起高度重视，从源头消除危害数据安全的风险隐患。

广大人民群众在进行电子零售商品的租赁购买时，如遇到可能危害国家安全的情况，可通过12339国家安全机关举报受理电话、网络举报平台（www.12339.gov.cn）、国家安全部微信公众号举报受理渠道或者直接向当地国家安全机关进行举报。

（据国家安全部微信公众号）