

人工智能背景下深度伪造行为电商平台应对的策略研究

丁羽洁

浙江理工大学法学与人文学院, 浙江 杭州

收稿日期: 2026年7月27日; 录用日期: 2026年8月7日; 发布日期: 2026年10月23日

摘要

人工智能技术的普及推动了电商平台升级, 但其衍生的深度伪造技术被恶意滥用, 威胁电商市场秩序与消费者权益。深度伪造基于生成式对抗网络, 可合成高逼真音视频; 在电商场景中, 买卖双方均存在滥用行为, 具有成本低、隐蔽性强、侵害多维化等特征, 冲击传统电商信任机制与平台监管。然而, 当前平台对商家深度伪造行为的规制仍面临多重困境。为此, 应通过完善平台权利救济等举报机制, 强化平台监管与行业自律, 构建协同治理格局, 降低商家和消费者的维权成本, 以保障电商行业的规范发展。

关键词

人工智能, 深度伪造, 平台管理

Research on the Strategies for E-Commerce Platforms to Respond to Deepfake Activities in the Background of Artificial Intelligence

Yujie Ding

School of Law and Humanities, Zhejiang Sci-Tech University, Hangzhou Zhejiang

Received: July 27, 2026; accepted: August 7, 2026; published: October 23, 2026

Abstract

The popularization of artificial intelligence technology promotes the upgrading of e-commerce platforms, but the deepfake technology derived from it is maliciously abused, threatening the order of the e-commerce market and consumers' rights and interests. Deepfake technology is based on the generative adversarial network, which can synthesize high-fidelity audio and video. In the e-commerce scene,

both buyers and sellers have abusive behaviors, which have the characteristics of low cost, strong concealment, and multi-dimensional infringement, and impact the traditional e-commerce trust mechanism and platform supervision. However, the current platform faces multiple dilemmas in regulating merchants' deepfake activities. To this end, platforms should improve the reporting mechanism, such as platform rights relief, strengthen platform supervision and industry self-discipline, build a collaborative governance structure, and reduce the cost of rights protection for merchants and consumers to ensure the standardized development of the e-commerce industry.

Keywords

Artificial Intelligence, Deepfake, Platform Management

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 问题的提出

人工智能技术的快速迭代与普及应用,为电子商务平台开辟了新的经济增长空间。各大电商平台抢抓人工智能发展机遇,在运营全流程广泛布局人工智能技术与应用。例如,在2025年“双十一”期间,淘宝天猫等电商平台依托千问大模型,全面落地了AI智能导购、算法推荐及智能客服等应用,通过分析用户浏览、搜索、交易等行为数据精准构建用户画像,实现商品推荐与流量分发[1]。生成式人工智能无疑推进了电商行业的效率以及创新,促进了产业结构的升级和社会生产方式的变化,但其在推动新质生产力的同时也面临着新的挑战。例如,生成式人工智能被用于生成虚假信息和深度伪造等内容,引发了社会伦理与信任危机,这在一定程度上使电商平台对大规模部署人工智能犹豫不决,进而影响了新质生产力的推动[2]。

部分经营者通过滥用人工智能深度伪造等手段,实施针对消费者的欺诈等违法行为,对市场秩序与消费者合法权益构成了潜在威胁。作为治理电商深度伪造行为的重要一方,平台具有传统意义上企业所不具备的力量,其“准行政能力”、市场权力以及海量资源等,都使其成为平台治理中不可忽视的主体[3]。然而,面对深度伪造行为,平台所制定的相关规则尚无法形成有效规制,致使其对平台声誉和市场秩序等造成严重破坏。

2. 电商深度伪造行为的基本界定与问题呈现

人工智能的发展充满不确定性。正是这一特性,使得AI风险呈现出主体、价值、场景及目标需求等维度的多元化特征,进而导致实体规则与平台治理规则陷入不确定状态。此外,由于相关利益关系错综复杂,平台难以仅凭单一的治理方式来有效应对人工智能伪造技术。本文从人工智能伪造技术的行为出发,对电商背景下人工智能伪造技术的行为形式作出基本界定。

2.1. 人工智能伪造技术概述以及在电商平台中的适用

深度伪造(Deepfake)是一种基于生成式对抗网络(GAN)的人工智能技术,表现为使用者通过输入提示词,并结合深度学习算法合成高度逼真的视频、音频等内容。此类生成内容真实感强,难以通过肉眼辨别。2019年,国家互联网信息办公室发布的《互联网信息服务深度合成管理规定》¹对深度合成技术作出了定义,即利用深度学习、虚拟现实等生成合成类算法制作文本、图像、音频、视频、虚拟场景等网络信

¹https://www.cac.gov.cn/2022-12/11/c_1672221949354811.htm

息的技术。近年来,随着人工智能技术的发展和电商平台的扩张,平台能够获取的案例不断增多,从而可以创建越来越精确的模型,这也是目前平台上深度伪造视频逼真度较高的重要原因之一。而在电商平台场景下,深度伪造技术在买卖双方中均有出现。有案例表明,有买家利用 AI 合成 6 只“死螃蟹”视频,向苏州大闸蟹电商商家骗取 195 元退款²;也有商家利用 AI 合成护肤品去皱等虚假视频,夸大产品功效,诱导老年消费者下单。在平台场景下,无论卖家还是买家,都可能通过 AI 合成视频的方式达到自己的违法目的。相较于原有的方式,电商平台借助底层算法,对生物特征、商品性状等进行高仿真模仿,其核心逻辑是以算法代替感官真实,瓦解了传统电商中基于视频证据的信任机制和人们“有图有真相”的经验判断;且此类合成内容完全可以躲避识别、实现虚假传播,对平台监管构成很大挑战[4]。同时,相较于传统电商欺诈,当前通过深度伪造实施的欺诈还体现出生成成本极低、隐蔽性极强、法益侵害多维化等典型特征。

2.2. 电商平台各主体深度伪造行为的样态

当前,生成式人工智能深度伪造功能正被部分市场经营主体挪用于市场营销实践,导致市场中存在一种以欺诈性内容为核心、覆盖多重商业触点的虚假传播结构。具体表现贯穿商品展示、直播营销与客户评论等核心节点:在商品展示环节,部分商家借助 AI 生成的虚假视觉素材,构建出与实际产品存在实质性偏差的商品形象。例如,通过伪造产品功效的“前后对比图”等手段,误导、诱导消费者下单购买,尤其是辨别能力较弱的老年群体;在直播营销环节,深度伪造技术被用于操控主播形象或场景呈现,目前在电商平台较为常见的是通过名人形象以及其声音的仿冒,营造虚假的宣传效果,以人为方式制造可信度幻觉,这种行为严重损害了消费者知情权,以及严重扰乱电商直播平台的市场竞争秩序;在客户评论管理环节,AI 深度伪造合成的用户反馈则进一步巩固了前述欺骗性叙事,批量生成虚假的好评视频,使该类视频欺诈信息在消费者决策链条中形成自我强化的闭合回路。卖家通过三类上下游行为的协同运作,严重侵犯电商平台的信誉、消费者的权利以及市场竞争秩序。

而买方的深度伪造行为则集中于售后举证环节。深度伪造相关视频证据构成新的制度性欺诈状态,就目前的相关案例表明,部分消费者利用深度伪造技术,制作商品损坏或变质的虚假影像,并将其作为欺诈性证据提交给商家或平台,以骗取退款为目的实施恶意索赔。由于电商平台的纠纷处理机制过度依赖视频举证,导致商家合法权益受损。此类行为正严重冲击着电商平台售后体系赖以运转的信任根基。当视频举证的可信度被系统性质疑,平台将面临规则调整的两难困境:若坚持“视频优先”原则,则继续暴露于合成欺诈的结构性风险之下;若收紧举证标准,则可能对提出正当索赔的消费者形成不合理的举证负担,进而损及消费者权益保护的制度效能。

3. 人工智能背景下电商平台治理深度伪造行为管理困境

目前,对于这种技术层面的风险,通常是通过技术进行规制。随着技术影响的扩大,以技术进行“溯源防伪”与“反向破解”的理念和实践也日益盛行。然而,随着技术不断发展以及平台内数据的持续集成,这些技术措施也在不断被规避,无法成功识别。可见,在人工智能背景下,技术并不能有效改善电商商家利用伪造技术破坏电商市场秩序的问题。

3.1. 规范供给滞后与模糊

2019 年出台的《网络音视频信息服务管理规定》³第十一条明确了网络音视频服务提供者和使用者对

²https://jsnews.jschina.com.cn/jsyw/202512/t20251215_s693fd374e4b05fbb22fdd897.shtml

³http://www.scio.gov.cn/xwfb/gwyxwbgswfbh/wqfbh_2284/2020n_4408/2020n11y02r/wjxgzc_5479/202207/t20220716_228416.html

于非真实音视频信息的标识义务以及新闻信息的传播转载规范。但是针对其中的非真实音视频并未给出相关的界定标准[5]。而《互联网信息服务深度合成管理规定》第二十三条,对深度合成技术的概念做出了界定,即利用深度学习、虚拟现实等生成合成类算法制作文本、图像、音频、视频、虚拟场景等网络信息的技术。但是也并未对电商深度伪造的行为做出准确界定。最新的《人工智能生成合成内容标识办法》⁴中,明确要求 AI 生成的内容需添加表示,但是在实践中,还是有许多商家规避该义务。此外,在流量加持与跨平台传播路径的共同作用下,相关侵权内容的溯源难度大幅增加,进而导致责任认定面临困境。例如,有学者认定通过深度伪造生成的声音,在侵权认定方面具有挑战性,在侵权认定过程中,权利人通常需要使用声纹鉴定等专业技术手段证明 AI 生成声音具有可识别性。声纹鉴定虽然能通过分析音色、共振峰等特征提供客观比对方案,但其成本高昂且精度受样本质量和算法限制。如果结合具体使用场景,还需依托相关语境和公众认知进行判断,导致举证过程更为复杂[6]。

3.2. 内容审核机制落后

电商平台的内容审核机制遭遇技术瓶颈。当前,主流电商平台普遍采用“机器审核 + 人工复审”的内容审核模式,但面对深度伪造内容,该机制存在明显不足。深度伪造技术的迭代速度飞快,平台审核的算法机制往往滞后于伪造技术的最新进展。且电商平台的商品品类丰富,不同品类商品的展示方式、视频特征存在显著差异,导致审核的精准度和覆盖率均难以保障。此外,电商平台每日处理的商品信息和交易内容数量庞大,全面深度伪造检测所需的算力成本极高,平台在商业效益与审核投入之间面临两难抉择。

3.3. 信用评价体系以及纠纷处理机制的危机

目前,电商平台主要依靠图片、视频等对商品、买家评价等达成交易。深度伪造的滥用导致目前电商平台信用评价体系赖以运行的真实性假设被打破。信用数据的“污染”影响不仅会导致单个交易的公平性丧失,更可能引发市场竞争失衡的效应。深度伪造的行为极有可能造成诚信经营经营者因无法与伪造好评的商家竞争而被逐出市场,或者因买家的不端行为导致商家疲于应付,从而使得平台整体交易质量下降和消费者信任流失。深度伪造技术的出现,使得“视频举证”这一举证逻辑面临严峻挑战,一方面,平台客服人员 and 纠纷裁判人员缺乏专业的深度伪造鉴别能力和工具,难以在海量纠纷中识别伪造证据;另一方面,被侵权方若要证明相关内容系深度伪造,通常需要借助声纹鉴定、视频帧分析等专业技术手段,其成本高昂且流程复杂,维权成本与实际损失之间严重不对等。

4. 电商深度伪造行为的平台治理管理策略体系

深度伪造的事后规制,往往呈现应对不足的情况。而人工智能的独特发展趋势与其他互联网工具不同,其治理的基本逻辑也与传统的社会治理模式不同。有学者认为,针对人工智能的内容治理应当遵循不同主体间多维度交叉互补的治理网络架构,每个主体都在这种网络结构中有不同的体系定位[7]。

4.1. 平台优化纠纷处理以及权利救济机制

AI 根据文本、图形、声音等描述生成相关的音视频,已经广泛应用于电子商务的多个领域。同时,其还可以根据商品的特性以及商家对该商品的营销要求,生成具有吸引力的视频描述和广告文案,提高商品广告内容的创意性以及转化性。在商品的展示图中, AI 可以根据商品特点等,自动生成相关的商品展示图等,节省实际拍摄的时间[8]。因此,这类合法使用 AI 的行为需要与电商商家的伪造行为进行严格

⁴https://www.cac.gov.cn/2025-03/14/c_1743654684782215.htm

区分，严防因平台标准模糊不清，导致电商商家的利益受损。而深度伪造行为具有隐秘性，除非影响较大，相关权利受损人往往难以发现。基于此，针对人工智能伪造视频隐蔽性的特点，应当扩宽举报机制。建立快速反应机制，针对相关侵权行为应当及时披露相关信息，采取手段阻止传播的进一步发展，并对相关视频进行溯源。

其次，平台也应当有所行动，为人工智能服务提供者提供深度伪造的视频样本，扩充深度伪造数据库样本，以便日后有效辅助司法取证。因此，在人工智能深度发展的背景下，平台应进一步履行其安全保障义务。针对消费者或商家投诉的涉及深度伪造的视频，平台应及时介入调查；一旦发现相关业务存在危害因素，应迅速采取阻断或救济措施。如果平台未及时采取措施并导致影响扩大，则应当承担相关责任。最后，电商平台应修订相关规范，针对 AI 伪造内容的隐蔽性，强制要求商家对 AI 生成内容进行显著标识；电商行业内部也应建立协同机制，明确 AI 服务提供者需对相关内容履行存档义务。同时，平台还需完善技术追溯功能，破解 AI 伪造视频取证难的困境，防止平台与商家相互推诿，从而切实保障消费者与商家的合法权益[9]。

4.2. 完善平台内容审核机制

平台对于经营者的市场监管比外在的社会性市场监管更加富有成效，通过平台的规则，有利于电商商家约束自己的行为。就当前的产业实践来看，电商平台市场秩序的治理模式主要有四种模式：市场主导型模式、政府主导型模式、社会主导型模式以及综合治理模式[10]。在人工智能的背景下，尤其在深度伪造这一现实的问题下，平台需要充分发挥其自身优势，通过利用规则的修订、补充和完善来规范市场交易活动。因此，电商平台应将深度伪造检测纳入内容审核的核心模块，构建分级分类的检测体系。具体而言：第一，根据商品品类的风险等级建立差异化审核标准。对于美妆护肤等涉及功效的敏感品类，需执行更严格的视频内容审核机制。第二，加大 AI 检测技术投入，引入多模态深度伪造检测模型，综合分析视频帧、音频波形、面部微表情等多维特征，提升检测的准确率和覆盖面。第三，建立深度伪造内容样本数据库，通过与 AI 技术服务提供者的协作，持续更新和扩充检测训练样本，形成动态演进的技术防御能力。第四，严格落实 AI 生成内容的标识制度，在商品详情页、直播页面等核心交易场景中，要求使用 AI 生成内容的商家进行显著标注，并将标识信息嵌入内容元数据，确保跨平台传播后仍可追溯。

4.3. 增强行业自律以维护信用评价体系

鉴于人工智能具有显著的不确定性特征，单纯依赖政府或平台的单一治理模式已难以奏效。在此背景下，应充分发挥行业层面的自我规制作用，构建“政府监管、平台自律与行业自治”协同发力的多元治理体系，这已成为破解人工智能治理困境的关键路径。行业作为人工智能技术应用的直接参与者与利益相关方，能够更精准地把握行业发展动态、技术应用痛点以及潜在的风险隐患，其自我规制具有灵活性、针对性与及时性的独特优势，能够有效弥补政府监管与平台治理的不足，形成全方位、多层次的治理合力。欧盟就通过行业制定《欧盟反虚假信息准则》⁵以加强业内对人工智能伪造的规制作用，我国也应当合理借鉴该治理框架，以行业共识为核心，推动行业内部共同制定针对性的行为准则、自律倡议与行业规范。通过明确电商商家在 AI 应用中的权责，规范 AI 在宣传、交易及评价环节的使用边界，遏制虚假宣传与数据造假等违规行为。这不仅能引导行业在技术赋能下实现高质量发展，也为我国构建 AI 多元治理体系提供了重要的行业实践支撑[11]。

⁵<http://www.eustandards.net/?p=7048>

5. 结论

近日,《商务部等6部门关于更好服务实体经济 推进电子商务高质量发展的指导意见》⁶提出了“人工智能+电商”概念,要求引导电商企业加强人工智能大模型等技术研发应用,优化消费体验、降低运营成本、提升流通效能^[12]。可见,人工智能在电商领域已经开始从上至下的部署。人工智能技术是一把双刃剑,一方面可以帮助企业降本增效,但另一方面也凸显出诸多问题。该文件中还指出,电商平台完善公开透明的规则体系,督促符合条件的平台内经营者依法办理经营主体登记,健全知识产权保护、与劳动者常态化沟通等机制。因此,电商平台应当在人工智能深度伪造的领域中发挥其作用,维护市场的公平竞争,做好对商家、劳动者、消费者利益的保障。

参考文献

- [1] 阿里巴巴集团. AI 首次全面落地天猫双 11 创造实质价值[EB/OL]. <https://www.alibabagroup.com/document-1915930722120499200>, 2026-04-09.
- [2] 王薇, 程曼曼. 生成式人工智能促进新质生产力的内在逻辑与实现路径[J]. 价格理论与实践, 2025(2): 81-84+241.
- [3] 李广乾, 陶涛. 电子商务平台生态化与平台治理政策[J]. 管理世界, 2018, 34(6): 104-109.
- [4] 曹开研. 内容创新对平台生态的再造与监管新难题[J]. 青年记者, 2021(20): 20-21.
- [5] 来小鹏, 邹玥. 深度伪造行为的危害性及其法律规制研究[J]. 中国版权, 2025(6): 35-51.
- [6] 吴汉东, 龙欣璇. AI 生成声音侵权认定的路径选择——以肖像权的参照适用为中心[J]. 北方法学, 2026, 20(1): 5-18.
- [7] 吴畏, 甘功平. 生成式人工智能治理的基本逻辑[J]. 江汉论坛, 2026(2): 40-52.
- [8] 姜蕾, 张蕾. AI 助推电子商务领域新质生产力发展[J]. 企业科技与发展, 2024(12): 26-30.
- [9] 陈志斌, 常馨雪, 陈志红. 人工智能时代平台企业网络信息内容责任重构与治理范式转换[J]. 江苏社会科学, 2025(5): 104-114.
- [10] 尹晓娟. 互联网时代网络交易平台治理研究[J]. 商业经济研究, 2020(23): 88-92.
- [11] 张涛. 后真相时代深度伪造的法律风险及其规制[J]. 电子政务, 2020(4): 91-101.
- [12] 马敬泽. 首次提出“人工智能+电商”“电商企业下沉供应链”, 商务部等六部门发文, 16 条举措助推电子商务高质量发展[EB/OL]. <https://mp.weixin.qq.com/s/VLKzcgsvKxHjJb4Qa2ltuA>, 2026-04-07.

⁶https://www.gov.cn/zhengce/zhengceku/202604/content_7064729.htm