

科技向善视域下生成式AI电商直播的风险治理研究

樊蕊, 张玥

江苏大学马克思主义学院, 江苏 镇江

收稿日期: 2026年7月27日; 录用日期: 2026年8月10日; 发布日期: 2026年9月20日

摘要

生成式AI在电商直播领域的广泛应用, 既是技术赋能商业创新的生动实践, 也是技术伦理与价值治理的重要场域。本文以“科技向善”为分析视域, 系统识别生成式AI电商直播中的价值偏离, 并探讨相应的治理路径。本文的理论贡献在于将“科技向善”从伦理倡导转化为可操作的分析视域, 构建了“偏离识别-价值回归”的分析链条, 为数字时代电商直播的伦理治理提供了分析工具。

关键词

科技向善, 生成式AI, 电商直播, 风险治理

Research on Risk Governance of Generative AI in E-Commerce Live Streaming from the Perspective of Technology for Good

Rui Fan, Yue Zhang

School of Marxism, Jiangsu University, Zhenjiang Jiangsu

Received: July 27, 2026; accepted: August 10, 2026; published: September 20, 2026

Abstract

The wide application of generative AI in e-commerce live streaming is not only a vivid practice of technology-empowered business innovation but also an important field of technological ethics and value governance. Taking “Technology for Good” as its analytical perspective, this paper systematically identifies the value deviations of generative AI in e-commerce live streaming and explores corresponding governance pathways. The theoretical contributions of this paper lie in transforming “Technology for

文章引用: 樊蕊, 张玥. 科技向善视域下生成式AI电商直播的风险治理研究[J]. 电子商务评论, 2026, 15(9): 665-671.

DOI: 10.12677/ecl.2026.1591044

Good” from an ethical appeal into an operational analytical perspective, constructing an analytical chain of “deviation identification - value return”, and providing an analytical tool for the ethical governance of e-commerce live streaming in the digital era.

Keywords

Technology for Good, Generative AI, E-Commerce Live Streaming, Risk Governance

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. 引言

1.1. 科技向善：数字时代的技术伦理转向

“科技向善”这一理念的提出，是对工业革命以来技术双重社会后果的深刻反思。技术发展在极大提升生产效率的同时，也可能带来伦理风险与社会异化。早在 19 世纪，马克思主义经典作家敏锐地指出，在资本主义制度下，技术往往异化为剥削劳动、控制人的工具，“技术的胜利，似乎是以道德的败坏为代价换来的” [1]。这一洞见揭示了技术发展的内在悖论，即技术本身并不天然地服务于人的福祉，其社会后果取决于制度安排与价值引导。科技向善理念正是对这一悖论的时代回应，其强调技术发展应回归“为人服务”的本质，使科学技术从资本的附庸回归为人类福祉的杠杆。近年来，负责任创新(Responsible Innovation)、伦理设计(Value-Sensitive Design)等学术思潮均体现了相似的取向[2]。

就本文而言，科技向善不仅是一种伦理倡导，更是构成核心的分析视域。这意味着我们不是将“向善”作为诸多规制措施之一，而是以其为价值坐标，审视生成式 AI 在电商直播中的应用实践，评判其价值偏离，探索其回归路径。这一视域帮助我们超越技术中心主义的盲区，穿透效率至上的迷思，在技术创新与人文关怀之间寻求平衡。

1.2. 生成式 AI 电商直播：技术赋能与价值风险的共生

生成式 AI 的爆发式发展正在深刻重塑电商直播的业态格局。以 AI 虚拟主播为代表的新形态，凭借全天候在线、超写实形象、实时互动能力等优势，拓展了传统直播的效能边界。已有研究指出，直播电商平台通过引入 AI 虚拟主播，能够有效降低运营成本并延长直播时长，对中小商家尤为显著[3]。AI 正从辅助工具转向直播场域的重要参与者，人机协同的多元共生新生态正在形成，这无疑是技术赋能的有力证明。

然而技术是一把双刃剑。从“AI 张文宏”卖蛋白棒，到 AI 生成虚假测评、AI 换脸仿冒名人，生成式 AI 应用失范问题频频进入公众视野。这些现象表明，技术赋能与技术失范并非先后关系，而是一体两面的共生存在。这一共生关系的深层本质，是工具理性与价值理性的冲突在人工智能时代的集中体现。技术以其强大的效率优势攻城略地，却可能在不经意间消解诚信的根基、模糊真实的边界、侵蚀信任的纽带。因此，效率提升与价值守护之间的张力，正是当下电商直播治理必须面对的核心难题。如何在技术发展中坚守科技向善的价值坐标，成为亟待回应的时代命题。

2. 科技向善视域下价值偏离的多重维度

2.1. 责任维度：责任主体与归责机制之困

科技向善的首要要求是“责任性”，即技术应用主体明确，责任链条可追溯。唯有如此，当技术对人

造成损害时, 权利救济才能落到实处。然而在生成式 AI 电商直播的实践中, 这一要求正遭遇前所未有的挑战。特别是 AI 虚拟主播的出现, 从根本上颠覆了传统直播的责任结构。在传统模式下, 主播是明确的法律主体, 其言行责任清晰可辨。但是当 AI 成为直播的主角, 问题就变得复杂起来。例如, AI 虚拟主播涉嫌虚假宣传时, 责任主体究竟是开发 AI 的技术公司、提供 AI 工具的平台、使用 AI 的商家, 还是运营 AI 账号的 MCN 机构? 当 AI 生成虚假评价误导消费者时, 谁来负责? 当 AI 换脸技术侵犯名人肖像权时, 责任该如何追究?

这些并非理论上的假设, 在现实中已有典型案例。在“AI 张文宏”直播间售卖蛋白棒事件中¹, 消费者基于对张文宏医生的信任下单, 事后才发现主播是 AI 合成的虚假形象。当消费者试图维权时, 面临的却是责任主体的模糊不清, 即商家声称其只是使用了 AI 工具, 技术公司认为其只提供了技术服务, 平台则主张已尽到了审核义务。消费者在多方推诿中维权无门, 这也反映出当前法律框架应对 AI 侵权的力不从心。

从电商经济视角看, 责任模糊不仅直接抬高了交易成本, 损害了平台声誉, “而且会损害消费者与直播商家的长期关系质量”[4]。从科技向善的视角审视, 责任维度的偏离本质上是“责任性”原则的落空。责任性意味着技术开发者、平台、商家等各参与方都应承担与自身角色相匹配的责任。然而在生成式 AI 电商直播中出现责任模糊的现象, 使得责任链条断裂、归责机制失灵。当技术造成损害却无人担责, 消费者的权益保障就沦为一句空话, 技术发展的“向善”导向也就失去了最基本的制度保障。

2.2. 信任维度：信息真实与消费权益之困

科技向善的“人民性”维度, 要求技术发展应增进消费者福祉、保障用户权益。在电商直播场域, 这意味着技术应用应当维护市场诚信、保障消费者知情权与公平交易权。然而生成式 AI 的某些应用, 正在反其道而行之。众所周知, AI 生成技术的核心能力是基于数据建模生成符合人类认知的新内容, 以数字手段创造高度逼真的视觉、听觉内容。这一能力在为内容创作带来便利的同时, 也为信息误导打开了方便之门。在电商直播中, AI 渲染技术可以让廉价面料呈现昂贵质感、AI 图像生成可以塑造“九头身”的理想化身材、AI 文案生成可以批量炮制“用户好评”等等。消费者面对这些高度仿真的内容, 往往难以辨别真伪, 知情权和选择权受到严重侵害。

已有研究表明, 电商直播中的信息不对称是导致消费纠纷的主要原因之一, 而 AI 合成内容加剧了这种不对称[5]。科技向善要求技术发展应服务于消费者的切身利益, 而生成式 AI 电商直播的“超真实”误导, 恰恰消解了消费者对真实性的判断能力, 侵蚀了市场诚信的根基。当技术成为欺诈的推手, 消费者权益被置于商业利益之下, 技术发展便从“为人服务”的初衷异化为侵害消费者权益的现实, 这无疑是对科技向善理念的背离与消解。

2.3. 人文维度：情感连接与人的主体性之困

科技向善的“可持续性”维度, 要求技术创新应与人文关怀相向而行, 使技术发展始终锚定于人的主体性, 而非在效率的追逐中消解社会关系的根基。然而在生成式 AI 电商直播的实践中, 技术对人的异化正在悄然发生。一方面, AI 主播的“关怀”本质上是算法预设的情感劳动。当用户与 AI 主播互动时, 系统会根据情绪识别算法调整话术、根据用户画像选择表达方式, 营造出“被理解”“被关心”的错觉。但这种关怀缺乏真实共情的基础, 其本质是算法的计算结果。问题在于, 观众明知是 AI, 却在长期互动中对其产生情感依赖, 形成一种“准社会关系”, 而这种关系是单向的、想象的情感连接, 使得消费者在

¹央视网. AI 换脸张文宏推销蛋白棒? 律师: 侵犯肖像权. AI 软件或涉“助罪” [EB/OL]. <https://news.cctv.com/2024/12/18/ARTIXgfusekuYhvDrjIGNCud241218.shtml>, 2024-12-18.

不知不觉中丧失警惕, 更容易被诱导消费。这种准社会关系在算法推荐环境中被进一步强化, 当下“直播带货”技术逻辑和营销逻辑冲击着个体的有限理性, 导致消费者更多地沉浸在主播营造的话语氛围中, 而常常忽略了所“秒杀”的商品是否合格、售后服务是否满意等不确定性风险[6]。另一方面, 直播内容呈现同质化趋势。AI 生成的内容往往采用标准化模板, 无论是视觉呈现还是话术表达, 都高度雷同。“全网最低价”“限时秒杀”“错过等一年”等口号被 AI 反复生成、反复使用, 直播内容从创意表达沦为流水线生产。这不仅降低了用户体验, 也削弱了电商直播作为差异化的营销渠道的竞争力。对“完美拟像”的统计学追求剥离了生物性瑕疵, 在触发“恐怖谷”效应的同时, 也模糊了真实与虚构的边界。当 AI 合成方言、少数民族语言时, 还可能产生刻板印象或文化误读, 进一步加剧文化表达的扁平化。

在电商直播这一典型的数字消费场景中, 情感连接与内容创意是驱动用户购买和品牌忠诚的核心要素。从科技向善的视角来审视, 人文维度的偏离是对“可持续性”原则的背离。科技向善强调技术创新应与人文关怀相协调, 技术应服务于人的全面发展。然而生成式 AI 的“伪情感互动”异化了真实的人际连接, 让消费者在算法编织的情感幻象中渐失主体性; 内容的同质化生产则消解了直播作为创意表达的多样性。当技术替代了人、算法控制了情感, 技术发展便从服务用户转向操控用户, 这显然偏离了科技向善的价值轨道。

2.4. 秩序维度: 制度供给与技术创新之困

科技向善要求技术发展应与制度创新协同推进, 只有制度与技术的动态适配, 才能确保技术创新始终沿着“向善”的方向前进。然而在生成式 AI 电商直播领域, 制度供给与技术创新的错位日趋显著, 困境日益凸显。具体而言, AI 技术的迭代周期以月计, 大模型版本更新、新功能上线几乎每周都有新变化; 而相关法律制度的修订周期以年计, 一部法规从起草到出台往往需要数年时间。这种时间差导致了制度供给与制度需求之间形成巨大缺口。当新技术已经广泛应用、新问题已经大量出现时, 监管规则却还在酝酿之中, 监管真空由此形成。这一错位带来的具体问题集中体现在三个方面: 一是在资质准入层面, AI 主播是否需要“持证上岗”? 其从业资格如何认定? 二是在内容审核层面, AI 实时生成的海量内容, 平台审核能力如何匹配? 当 AI 每秒生成数千条评论、数百个产品展示时, 传统的人工审核显然力不从心。三是在跨境监管层面, 境外开发的 AI 主播在国内平台直播, 如何管辖? 这些问题在现行制度框架中都缺乏明确答案。

2026 年 2 月施行的《直播电商监督管理办法》²明确提出, 使用 AI 技术生成的人物图像、视频从事直播电商活动, 必须按照国家有关规定进行标识, 这是制度创新迈出的重要一步。但具体实施细则仍在探索中, 标识的标准、范围、方式等有待统一。多家平台签订了《促进 AI 技术规范应用承诺书》³, 可各家的执行标准却不尽相同, 缺乏统一的行业规范。

在电商直播这一快速迭代的新兴业态中, 技术创新与制度供给之间的动态匹配是市场健康运行的关键前提。从科技向善的视角来审视, 秩序维度的偏离是对技术发展内在规律的忽视。科技向善不仅要求技术创新本身“向善”, 更要求制度环境为“向善”提供保障。当制度滞后导致监管真空, 技术创新就可能偏离“向善”轨道, 甚至被滥用为侵害权益的工具。制度与技术的协同演进, 是科技向善得以实现的必要条件。

综观上述四重价值偏离, 它们并非孤立存在, 而是相互关联、相互强化的系统性问题, 对电商直播行业的健康发展构成现实挑战。责任模糊增加了消费者维权成本和平台合规风险, 也为责任逃避创造了

²国家市场监督管理总局、国家互联网信息办公室. 直播电商监督管理办法[EB/OL].

<https://www.xinhuanet.com/fortune/20260107/d70f1ad6d89548e0b9324fc678624698/c.html>, 2026-01-07.

³北京市市场监督管理局. 新京报: 8 家平台企业签订全国首份促进 AI 技术规范应用承诺书[EB/OL].

https://scjgj.beijing.gov.cn/zwxx/mtjj/202512/t20251231_4382392.html, 2025-12-04.

空间; 信息误导直接侵蚀消费者信任, 损害平台商誉和品牌商家的长期利益; 情感异化与内容同质化削弱了电商直播的核心竞争力, 导致用户粘性下降和流量内卷加剧; 制度滞后则使前三重问题难以得到及时规制, 市场失序风险上升。在生成式 AI 深度嵌入电商直播产业链的进程中, 技术创新若缺乏清晰的伦理指引和适配的制度规范, 将可能偏离负责任创新的初衷, 损害电商生态的公平性、透明度和可持续性。

3. 科技向善视域下的价值回归与治理路径

为应对科技向善视域下的多重价值偏离, 本部分从法律规制、平台治理、价值嵌入、社会共治四个维度, 探讨生成式 AI 电商直播风险治理的具体路径, 将“科技向善”理念转化为可操作的制度安排与行业规范。

3.1. 法律规制: 以归责清晰重塑责任链条

生成式 AI 电商直播中普遍存在的责任模糊现象, 不仅暴露出现行法律框架应对 AI 侵权的能力不足, 也直接影响了电商交易中的消费者信任和市场秩序。因此, 法律规制应从以下几个方面完善制度供给: 其一, 完善侵权认定标准。明确技术提供方、内容生成者、平台、商家的责任划分, 这是破解归责困境的基础。《直播电商监督管理办法》已确立“直播间运营者首责”原则, 即使用 AI 技术生成的人物图像、视频从事直播电商活动出现违法行为的, 由管理或使用该人物图像、视频的直播间运营者承担责任。这一规定初步明确了责任主体, 但技术提供方的过错责任仍需进一步细化。当 AI 系统本身存在设计缺陷或算法偏见时, 技术公司是否应当承担责任? 当平台提供的 AI 工具被商家滥用时, 平台是否应当承担审核不力的责任? 这些问题需要在未来的立法和司法实践中逐步明确。其二, 落实标识义务。标识是保障消费者知情权的基本手段。《直播电商监督管理办法》明确要求, AI 生成内容必须按照国家有关规定进行标识。标识不仅是形式要求, 更是对技术透明度的制度保障, 当消费者知晓对话对象是 AI 时, 其情感投入和消费决策会趋于理性, 被算法操控的风险也随之降低。标识的标准、范围、方式需要进一步统一, 确保“AI 生成”或“拟真场景”等提示内容在显著位置清晰呈现, 让消费者一目了然。其三, 建立身份备案制度。对 AI 主播实行“数字人”身份备案, 是源头治理的有效路径。每一名 AI 主播都应有唯一的身份标识, 其开发主体、运营主体、使用主体等信息应当可查询、可追溯。唯有如此, 才能确保 AI 主播“可解释、可审查、可追溯”, 从源头消除责任模糊的空间, 使技术应用始终在透明、可监督的轨道上运行。

3.2. 平台治理: 以责任内化替代效率优先

平台是技术应用的关键枢纽, 也是风险防控的第一道防线。当前, 平台在审核机制、技术防控、惩戒措施等方面仍存在诸多不足, 因此, 平台治理的逻辑应从“效率优先”转向“责任内化”, 将科技向善理念嵌入运营全流程。其一, 在技术防控层面, 可以引入 AI 内容检测系统, 识别仿冒名人形象、未标识 AI 内容, 以技术手段防范技术滥用。各大传播平台已普遍采用过滤和拦截技术, 防止侵权视频传播。作为内容提供者的 AI 视频生成平台, 在侵权违法内容的审查中应承担更高责任, 将技术防控前置化、智能化。其二, 在审核机制层面, 建立 AI 主播资质审核机制, 要求商家上传 AI 技术授权材料、AI 生成内容合规承诺书, 将风险拦截在准入环节。其三, 在惩戒机制层面, 建立商家信用档案, 将 AI 违规使用与流量扶持、佣金比例等权益挂钩, 形成“守信激励、失信惩戒”的市场导向。对于以流量为核心的电商模式, 流量管控能立竿见影地影响违规经营者收入来源, 具有传统罚款无法比拟的威慑力。与此同时, 监管部门可将违规直播间信息通报平台, 平台需采取限制流量、暂停直播乃至关闭账号等措施。其四, 在交互熔断层面, 对用户使用时长与情感依赖指数超阈值者, 系统自动中断 AI 对话并提示风险, 以技术手

段体现人文关怀,防止消费者在算法编织的情感幻象中丧失主体性。这一机制在效率与关怀之间寻求平衡,体现了科技向善中“以人为本”的价值取向。

3.3. 价值嵌入：以内生自觉超越外部监管

法律规制和平台治理皆属于外部约束,但科技向善的可持续性最终取决于价值理念是否真正内化为企业行为自觉。为扭转人文维度的情感连接与人的主体性之困,需要在算法设计、行业自律、人才培养三个层面协同发力,实现从外部监管到内生自觉的跃升。其一,在算法设计阶段,将“诚信”“公正”“以人为本”等价值观转化为可操作的算法伦理准则。算法从来不是价值中立的,其设计理念、数据选择、参数设置都蕴含着价值判断。因此,应当在算法设计之初就嵌入伦理考量,使“向善”成为算法的内在基因。已有研究表明,“营销者可以通过改变对道德消费的叙述框架策略来影响消费者对算法推荐的信任”[7],进而增强消费者的购买意愿。其二,在行业自律层面,建立AI主播“伦理认证”机制,对符合伦理标准者给予流量扶持或“绿色直播间”认证标识。认证机制可以发挥市场激励作用,使“向善”成为竞争优势而非成本负担。北京市消费者协会推动的《促进AI技术规范应用承诺书》是行业自律的重要探索。其三,在人才培养层面,培育既懂技术又具人文素养的复合型人才,使科技向善理念在代际传承中不断深化。电商直播行业的持续创新需要既熟悉技术逻辑又理解消费者心理的专门人才。应当推动高校、平台企业和行业组织多方协同,在技术教育中融入商业伦理与消费者权益保护内容,培养具备价值自觉的从业者。

3.4. 社会共治：以多元协同弥补单向监管

技术治理不能仅依赖政府和平台,更需要平台企业、行业组织、消费者和媒体等多元主体的共同参与。单靠自上而下的监管难以跟上技术迭代的速度,因此需要构建“平台负责、行业自律、公众参与、政府兜底”的协同治理格局,以多元主体的集体智慧弥补单一监管的不足。其一,消费者教育是基础。提升公众对AI生成内容的辨别能力,是抵御风险的第一道防线。例如,浙江省消保委建议消费者“在短视频或直播中要仔细观察主播的微表情是否自然、声音和口型是否同步、是否出现人物或背景模糊扭曲等异常画面;对异常低价或反常促销行为保持警惕”[8]。只有消费者自身具备足够的数字素养,才能在海量直播内容中识别风险、理性决策,减少因信息不对称导致的消费纠纷。其二,行业自律是关键。行业协会应推动建立AI直播规范,发布负面清单,鼓励企业签署“AI内容使用承诺书”,承诺不利用AI进行虚假宣传。行业自律能够弥补法律的滞后性,形成快速响应技术变化的柔性治理机制,降低全行业的合规成本。其三,媒体监督是补充。曝光典型AI违规案例,形成舆论压力。例如,2026年“3·15”期间,多家媒体集中报道AI仿冒名人带货、大数据杀熟等新型网络消费陷阱,推动形成社会共识,倒逼平台和商家规范经营。其四,投诉机制是保障。完善消费者投诉举报渠道,对AI相关投诉设立绿色通道,降低维权成本,确保消费者的声音能够有效进入治理体系。当技术侵权发生时,便捷高效的救济渠道是维护市场公平和消费者权益的最后屏障。

4. 结论

生成式AI在电商直播领域的广泛应用,既提升了商业运营效率,也带来了技术伦理与市场治理的新挑战。本文以“科技向善”为分析视域,系统识别了生成式AI电商直播中的价值偏离,并探讨了相应的治理路径。本文的理论贡献在于将“科技向善”从伦理倡导转化为可操作的分析视域,构建了“偏离识别-价值回归”的分析链条,为电商直播领域的AI治理提供了分析工具。技术发展的历史表明,每一次重大技术变革都会带来效率提升,也伴随着治理规则的重新调整。生成式AI在电商直播中的应用,需要

在技术创新与消费者权益保护之间寻求动态平衡。当算法效率与人文关怀、制度建设与市场自律形成合力, 电商直播有望成为兼具商业活力与社会信任的数字消费场景。

基金项目

2025 年度江苏省社科应用研究精品工程高校思想政治教育专项课题“人工智能赋能思想政治教育研究”(25SZC-019); 江苏大学人才启动金资助项目“新时代思想政治教育文化价值实现研究”(24JDG003)。

参考文献

- [1] 中共中央马克思恩格斯列宁斯大林著作编译局. 马克思恩格斯文集: 第2卷[M]. 北京: 人民出版社, 2009: 580.
- [2] Stilgoe, J., Owen, R. and Macnaghten, P. (2013) Developing a Framework for Responsible Innovation. *Research Policy*, 42, 1568-1580. <https://doi.org/10.1016/j.respol.2013.05.008>
- [3] 史嘉颖. AI 虚拟主播对消费者道德产品购买意愿的影响研究[J]. 现代营销, 2026(9): 16-18.
- [4] 彭宇泓. 直播电商在线服务失败影响消费者负面口碑评价的复合机制[J]. 财经论丛, 2024(5): 93-102.
- [5] 王振宇. 平台经济背景下消费者权益保护的困境与出路[J]. 中国流通经济, 2023(4): 116-128.
- [6] 徐颖. 平台经济视域下网络直播平台的治理困境研究——以网红直播带货为例[J]. 西安石油大学学报(社会科学版), 2020, 29(4): 37-43.
- [7] 徐岚, 陈全, 等. “共赢” vs. “牺牲”: 道德消费叙述框架对消费者算法推荐信任的影响[J]. 心理学报, 2024, 56(2): 179-193.
- [8] 网游、教培、短视频带货等领域投诉多发 省消保委发布提醒[N]. 浙江法治报, 2025-07-15(002).