E-Commerce Letters 电子商务评论, 2025, 14(5), 1855-1860
Published Online May 2025 in Hans. https://www.hanspub.org/journal/ecl
https://doi.org/10.12677/ecl.2025.1451470

Hans汉斯

生成式AI电商客服虚假信息责任与技术中立抗辩边界

许阳成

贵州大学法学院，贵州 贵阳

收稿日期：2025年3月31日；录用日期：2025年4月15日；发布日期：2025年5月23日

摘 要

在数字化浪潮的推动下，生成式人工智能技术正迅速渗透到电商客服领域。然而，随着技术的广泛应用，虚假信息传播的责任归属和技术中立原则之间的抗辩边界日益模糊。本文以ChatGPT自动回复引发的合同纠纷案为切入点，深入探讨此问题。通过对相关案例的分析和法律条文的解读，阐述了电商平台、AI开发者和运营者等各方在虚假信息传播中的责任承担，以及“技术中立”抗辩在何种情况下不能成立，并提出解决方案，旨在为电商行业合理运用生成式人工智能技术提供法律参考，促进电商行业的健康发展。

关键词

生成式人工智能，虚假信息，民事责任，技术中立抗辩

Generative AI E-Commerce Customer Service False Information Liability and Technology Neutrality Defense Boundary

Yangcheng Xu

School of Law, Guizhou University, Guiyang Guizhou

Received: Mar. 31st, 2025; accepted: Apr. 15th, 2025; published: May 23rd, 2025

Abstract

Driven by the wave of digitalization, generative AI technology is rapidly penetrating into the field of e-commerce customer service. However, with the widespread use of technology, the boundaries between responsibility for the spread of disinformation and the principle of technical neutrality are increasingly blurred. This article takes the contract dispute case caused by ChatGPT's automatic reply

文章引用: 许阳成. 生成式 AI 电商客服虚假信息责任与技术中立抗辩边界[J]. 电子商务评论, 2025, 14(5): 1855-1860.
DOI: 10.12677/ecl.2025.1451470

as the starting point to explore this issue in depth. Through the analysis of relevant cases and the interpretation of legal provisions, this paper expounds the responsibilities of e-commerce platforms, AI developers and operators in the dissemination of false information, and the circumstances under which the defense of "technology neutrality" cannot be established, and proposes solutions, aiming to provide legal reference for the reasonable use of generative AI technology in the e-commerce industry and promote the healthy development of the e-commerce industry.

Keywords

Generative AI, False Information, Civil Liability, Technology Neutrality Defense

1. 问题的提出

在当今数字经济蓬勃发展的时代背景下，电商行业迎来了迅猛增长的态势，客服服务作为电商企业与消费者之间沟通交流的关键纽带，其重要性日益凸显。生成式人工智能技术的横空出世，为电商客服领域注入了全新活力，它具备快速且自动回应顾客咨询的能力，极大地提升了客服工作效率。不过，该技术在应用过程中也潜藏着诸多风险，尤其是虚假信息传播所引发的法律风险，世界经济论坛发布的《2024 年全球风险报告》称未来两年世界“十大风险”之首是获 AI 助力的“信息错误和虚假信息”[1]，这引发了广泛的社会关注。与以往的人工智能技术相比，生成式人工智能(如 ChatGPT)给人类社会带来的潜在风险更加多样、防范也更为紧迫，现有法律制度面临挑战，有必要做出法律回应[2]。以 ChatGPT 为例，这款极具代表性的生成式人工智能其自动回复功能在电商客服场景应用时，可能会触发一系列棘手的法律问题，就如近期备受瞩目的 ChatGPT 自动回复导致的合同纠纷案件：在这起备受关注的合同纠纷案中，涉及一家电商企业(以下简称“电商 A”)。电商 A 在其客服体系中全面引入 ChatGPT 作为客服工具，旨在提高客服效率和应对大量的顾客咨询。一位长期合作的企业客户(以下简称“客户 B”)在与 ChatGPT 客服交互过程中，询问某特定产品是否符合特定行业标准且能够在特定时间内交付足够数量以满足其大型项目需求。ChatGPT 的自动回复明确表示该产品完全符合行业标准并且能够按时足量交付。基于这一回复，客户 B 与电商 A 签订了一份价值不菲的合同。然而，在合同执行过程中，客户 B 发现产品并不完全符合行业标准，且交付数量也无法满足项目需求，导致客户 B 遭受了巨大的经济损失，包括重新寻找供应商的额外成本、项目延误的违约金等，客户 B 认为电商 A 提供了虚假信息，随后将电商 A 告上法庭，要求赔偿所有损失。

此一案例凸显两大核心问题：一是虚假信息传播的责任归属：电商平台、AI 开发者与运营者如何划分责任？二是技术中立抗辩的适用边界：技术开发者能否以“中立性”为由免除法律责任？现有法律框架对生成式 AI 的规制存在空白，技术伦理与商业利益的冲突加剧了责任认定的复杂性。这一典型案例为深入探究生成式人工智能在电商客服领域的法律风险，提供了极具价值的样本和研究方向，促使人们进一步思考如何在享受技术红利的同时，有效规避潜在的法律危机，保障电商客服活动在合法合规的轨道上稳健前行。

2. 虚假信息传播责任归属与技术中立抗辩困境

(一) 虚假信息传播的多主体责任

生成式人工智能因缺乏人格尊严，被认定不具备法律主体资格[3]。传统侵权事件中，虚假信息的传播通常是人为的，行为人需对此承担相应的法律责任。然而，生成式人工智能生成虚假信息并非人为，而是由其内部的算法所导致。那么，在生成式人工智能发布虚假信息侵害公民权利的情况下，应当由谁来承担责任呢？

1、电商经营者责任

1) 信息审查义务

从我国《电子商务法》第 27 条规定可知，平台经营者的信息审查义务为一种第三方法定义务[4]，电商企业作为与顾客直接进行交易的主体，有义务对 ChatGPT 等生成式人工智能的回复进行审查。即使是自动回复，也不能免除其确保信息准确的责任。在上述案例中，电商 A 如果能够在使用 ChatGPT 回复客户 B 之前，对回复内容进行简单的核实，例如与产品部门确认产品的标准和供应能力，就可能避免提供虚假信息。如果电商企业未能履行审查义务，导致虚假信息传播，应当承担相应的法律责任。

2) 合同关系的考量

从合同关系角度看，客户 B 与电商 A 之间存在买卖合同或服务合同关系。电商 A 有责任按照合同约定提供准确的信息，保障客户 B 的权益。如果因为虚假信息导致客户 B 权益受损，电商 A 构成违约。在这个案例中，电商 A 提供的虚假信息直接影响了客户 B 的合同决策，导致客户 B 遭受损失，因此电商 A 需要承担违约责任。

2、电商平台责任

电商平台作为生成式人工智能客服工具的直接使用者和商业利益获得者，应当承担部分因虚假信息传播而产生的责任。依据《电子商务法》第 27 条的规定，电商平台有义务对所展示的商品或服务信息的真实性与合法性进行审查，而这一审查义务并不会因为信息是由人工智能生成而被免除。如果电商平台将类似 ChatGPT 这样的工具作为客服的主要组成部分，那么就需要构建与人工客服同等有效的审查机制。例如，当客户咨询涉及合同关键条款(如交付时间、质量标准等)时，平台应当制定“人工智能回复预审规则”，要求系统自动触发人工复核流程，或者进行数据交叉验证(例如比对库存系统与物流信息)。若电商平台仅以“技术不可控”为由推卸责任，那么这种行为将构成对法定义务的消极履行。

3、ChatGPT 开发者责任

1) 开发过程中的注意义务

生成式人工智能的开发者在模型构建过程中，负有对数据真实性、准确性进行严格审查的法定义务，并应持续优化算法逻辑以最大限度降低虚假信息生成的风险。若开发者因主观疏忽(如采用错误数据源或存在算法缺陷)导致虚假信息传播，则需依法承担相应责任。若 ChatGPT 开发者在数据采集阶段未能对产品信息进行有效核实，或算法在处理特定问题时存在技术漏洞，进而生成误导性回复，开发者亦应被认定为责任主体之一。这一责任认定不仅符合技术伦理，也与《生成式人工智能服务管理暂行办法》中关于数据质量与算法安全的要求相一致。部分开发者出于成本和效益的考量存在利用用户去训练模型的倾向，给予了普通人对人工智能进行目的性训练的能力，从而给虚假信息传播创造了空间，形成“人为操纵型风险”。生成式人工智能在服务用户过程中所产生和收集的数据同样会成为其训练数据的一部分，若就某些训练样本量不足的人物或事件对人工智能进行特定训练，则极有可能生成虚假信息[5]。由于 AI 引发的损害往往由多方面原因导致，各方均应采取合理注意义务避免损害发生，因此，AI 设计者、AI 使用者都应当采取适当注意义务，否则就应承担责任[6]。

2) 产品说明与警示义务

产品说明与警示义务通常来说就是产品缺陷的警示缺陷，对于人工智能产品来说，无论是产品缺陷的认定，还是因果关系的证明，都需要立法更新认定标准[7]。开发者有义务对生成式人工智能的功能、

可能存在的风险进行准确的说明，并向使用者(如电商企业)提供必要的警示。如果没有履行这些义务，在虚假信息传播引发法律问题时，也应当承担相应的责任。开发者没有向电商 A 说明其回复可能存在一定的误差范围，尤其是在涉及商业交易的关键信息方面，那么在出现问题时，开发者也不能完全免责。

(二) “技术中立”抗辩的局限性

1、技术并非完全中立

尽管 ChatGPT 开发者声称技术中立，但在实际的技术架构搭建与实现流程里，开发者的主观取向不可避免地会融入其中。以数据层面为例，数据来源的选取以及筛选机制的设定，直接关乎生成结果的精准度；再看算法维度，算法的独特设计架构，会对回复内容呈现出明显的导向性。就拿本案例来说，倘若 ChatGPT 的预训练数据囊括了诸多未经严谨核实，亦或是已然过时的该产品相关信息，那么生成不准确回复的概率便会大幅攀升。而且，若算法设计之初，未能充分将商业应用场景下对准确性的严苛标准纳入考量范畴，同样极易炮制出误导性十足的回复内容，给电商客服环节带来潜在隐患。

2、合理技术风险与不可接受风险难以区分

随着互联网技术、人工智能等科技的迅猛发展，法律与科技之间的难题不断凸显，司法中关于技术之定位的疑难案件也反复出现。技术中立原则常常被用于反对法律对技术的监管，或者为技术服务者免责[8]。

1) 合理的技术风险

在技术发展过程中，必然存在一定的风险，例如由于数据的不完整性或者算法的局限性导致的偶尔的不准确回复。对于这些合理的技术风险，如果开发者和电商企业已经采取了合理的措施进行防范，如定期更新数据、优化算法、进行人工审查等，那么在一定程度上可以作为减轻责任的依据。在上述案例中，如果 ChatGPT 开发者定期更新数据以确保信息的准确性，并且电商 A 也对 ChatGPT 的回复进行了一定程度的人工审查，但仍然出现了偶尔的不准确回复，这种情况下可以考虑适当减轻责任。

2) 不可接受的风险

在风险社会理论的框架下，风险的不可预测性要求法律体系能够适应并管理不同社会子系统间的相互作用，以建立旨在实现“可接受”安全水平的制度。在生成式人工智能(AI)的背景下，系统性地产生虚假信息是一个显著的风险。这种风险可能源于算法设计上的偏见，或者是因为使用了大量未经验证的数据。在这种情况下，所产生的风险是不可接受的，不能简单地以技术的中立性作为免责的借口。例如，如果在某个案例中发现，为了提高回答速度，ChatGPT 的算法牺牲了信息的准确性，或者依赖了大量来源不可靠的数据，导致了对产品信息的误导性回复，那么无论是电商平台 A 还是 ChatGPT 的开发者，都不能以技术中立为由来规避责任。

(三) 技术中立抗辩不成立情形

第一，明知或应知虚假信息的存在而未采取措施，电商平台或 AI 开发者明知或应知生成式 AI 传播的信息属于虚假信息，却未采取必要措施加以制止或纠正，那么“技术中立”抗辩将不能成立。电商平台在发现 ChatGPT 自动回复的产品性能信息与实际不符后，仍未对回复内容进行修改或提示消费者注意，此时电商平台不能以“技术中立”为由免责。第二，未履行合理的注意义务，电商平台和 AI 开发者在使用和运营生成式 AI 时，应履行合理的注意义务，包括对生成内容的审核、监测和管理。如果因未履行或未充分履行注意义务，导致虚假信息传播，给消费者造成损害，那么“技术中立”抗辩也不能成立。电商平台未对 ChatGPT 的回复内容进行定期审核，导致其长期传播虚假的售后服务信息，电商平台应对此承担相应责任。第三，滥用生成式人工智能技术，电商平台或 AI 开发者为了追求商业利益或其他目的，故意滥用生成式人工智能技术，导致虚假信息传播，那么“技术中立”抗辩将无法成立。例如，电商平台为了吸引消费者购买产品，故意使用 ChatGPT 生成虚假的促销信息或好评，这种行为明显超出了“技术中

立”的范畴，应当承担相应的法律责任。

3. 应对生成式 AI 电商客服虚假信息传播民事责任风险的建议

(一) 加快人工智能领域的专门立法

欧盟、美国和中国作为全球 AI 发展的主要推动者，也是 AI 风险治理者。美国的治理模式为产业导向的分散化立法，欧盟为统一立法并制定了全球首个综合性 AI 治理法案即《人工智能法案》，中国则为行政主导的敏捷治理。以欧盟为例，其于 2025 年 2 月 4 日公布了《人工智能法案》的指导方针草案，对雇主、网站、平台等主体合规利用 AI 的方式和情景作出了一系列解释说明。目前我国也在积极探索人工智能领域的立法，《人工智能法》(学者建议稿)《人工智能示范法 3.0》(专家建议稿)的陆续发布，意味着我国人工智能治理体系建设即将迈入新阶段。

完善法律适用与法律责任分配机制化解法律风险，包括健全与生成式人工智能技术发展相适应的法律体系，以多方责任共担机制确认生成式人工智能的法律责任等[9]。各国人工智能治理模式法案中都明确责任原则[10]，针对生成式人工智能在电商客服中的应用，明确电商企业和开发者在虚假信息传播中的责任主体地位，以及各自的责任范围。根据不同的过错程度、损害结果等因素，确定具体的赔偿责任，制定具体的法规条款，明确在类似电商 A 与客户 B 的案例中，电商企业和开发者分别应承担的责任比例，根据各自的过错程度进行合理分配。制定明确的规则来规范“技术中立”抗辩的适用条件，避免其被滥用，规定在何种情况下可以认定为合理的技术风险，以及需要满足哪些条件才能进行技术中立抗辩。可以设立专门的评估标准，对于技术风险的可接受程度进行量化评估，如规定数据的误差率在一定范围内属于合理风险，超过该范围则不能以技术中立抗辩。

(二) 加强技术监管

生成式人工智能技术和产品的创新发展方兴未艾，立法规制既要循序渐进，以情况清楚和证据充分为监管前提，避免不当阻碍新兴科技的发展；又要及时规制，防止大规模的社会风险发生，造成不可逆的损害[11]。生成式人工智能在电商客服中的应用需覆盖开发、应用及事后处置全链条，通过标准化流程与动态机制降低风险。对生成式人工智能的开发过程进行监管，包括对数据来源、数据质量、算法设计等方面的审查，确保开发者在开发过程中遵循相关的法律法规和伦理规范，减少虚假信息生成的可能性。监管部门还可以要求开发者对数据来源进行详细备案，定期检查算法的准确性和公正性。对电商企业使用生成式 AI 进行客服的情况进行监督，检查其是否履行了信息审查义务，是否存在滥用生成式 AI 导致虚假信息传播的情况。建立定期检查制度，对电商企业的客服记录进行抽查，发现问题及时责令整改。应用阶段的动态监测，对事实内容进行系统审核，包括第一：AI 辅助审核，电商平台需部署“双 AI 审核系统”，即主 AI 生成回复后，由辅助 AI (如基于规则引擎的审核模型)即时筛查敏感词、矛盾信息及逻辑漏洞，例如，当 ChatGPT 回复“产品支持 24 小时到货”时，辅助 AI 自动比对物流数据库，若发现与实际情况不符，则拦截该回复并触发人工复核。第二关键信息标记与追溯：对涉及合同条款、质量标准、交付时间等核心信息的 AI 回复，系统自动添加数字水印，记录生成时间、会话 ID 及所用数据版本，便于事后追责。而在事后处置环节，应当将重点放在消费者权益救济上，人工智能时代的消费者权益是否得到有效保护需要进一步检视[12]。可以考虑建立“AI 纠纷快速处理基金”，该基金由开发者和平台按照一定比例注入资金。消费者倘若遭遇因 AI 回复导致的权益受损情况，可凭借相关聊天记录申请先行赔付。基金在先行赔付给消费者之后，再依据记录的详细信息向责任方进行追偿，以此避免消费者陷入漫长的维权周期，切实保障消费者的合法权益。

(三) 提高企业自律

电商企业应当建立健全内部管理制度，加强对生成式人工智能客服的管理，设立专门的审查岗位，

对 ChatGPT 等生成式人工智能的回复进行人工审查，及时发现和纠正虚假信息。可以制定详细的审查流程和标准，确保审查工作的有效性。开发者也应当增强社会责任意识，在开发过程中积极采取措施防范虚假信息传播风险，建立数据验证机制，不断改进算法，提高生成式人工智能的准确性，开展行业内部的技术交流和合作，共同提高生成式人工智能的可靠性。

4. 结论

生成式人工智能在电商客服中的应用虽然带来了诸多便利，但也伴随着虚假信息传播等法律风险。通过对 ChatGPT 自动回复引发的合同纠纷案的分析，虚假信息传播责任和“技术中立”抗辩边界时存在着复杂性。为了应对这些法律风险，需要完善法律法规、加强技术监管以及提高企业自律，在充分发挥生成式人工智能优势的同时，保护消费者权益，维护电商市场的健康稳定发展。

参考文献

[1] 蒲泓宇, 李洪晨, 赵星. 人工智能生成虚假信息的内生安全治理框架[J]. 图书馆论坛, 2025, 45(5): 133-140.

[2] 郭金良. 生成式人工智能服务中“信息错误”的民事责任[J]. 政法论坛, 2025, 43(2): 25-35.

[3] 张劲松. 人是机器的尺度——论人工智能与人类主体性[J]. 自然辩证法研究, 2017, 33(1): 49-54.

[4] 秦红嫚, 张志力. 我国直播电子商务平台经营者信息审查义务的标准[J]. 浙江理工大学学报(社会科学版), 2022, 48(4): 442-450.

[5] 朱嘉珺. 生成式人工智能虚假有害信息规制的挑战与应对——以 ChatGPT 的应用为引[J]. 比较法研究, 2023(5): 34-54.

[6] 王利明. 生成式人工智能侵权的法律应对[J]. 中国应用法学, 2023(5): 27-38.

[7] 郑志峰. 人工智能产品责任的立法更新[J]. 法律科学(西北政法大学学报), 2024, 42(4): 3-17.

[8] 郑玉双. 破解技术中立难题——法律与科技之关系的法理学再思[J]. 华东政法大学学报, 2018, 21(1): 85-97.

[9] 吴宗宪, 张进帅. ChatGPT4.0 视域下的社会风险及治理路径探究[J]. 法治论坛, 2024(3): 87-101.

[10] 周辉, 许玖玖, 朱悦, 张心宇. 人工智能治理: 场景、原则与规则[M]. 北京: 中国社会科学出版社, 2022.

[11] 陈兵. 促进生成式人工智能规范发展的法治考量及实践架构——兼评《生成式人工智能服务管理暂行办法》相关条款[J]. 中国应用法学, 2023(4): 108-125.

[12] 冯子轩. 人工智能与法律[M]. 北京: 法律出版社, 2021.