

人工智能在社交媒体品牌声誉管理中的应用研究

——以微博智搜为例

张珂玮

扬州大学新闻与传媒学院, 江苏 扬州

收稿日期: 2026年6月9日; 录用日期: 2026年7月5日; 发布日期: 2026年7月15日

摘要

随着社交媒体成为品牌舆情的主要发生场域, 人工智能技术在品牌声誉管理中的应用日益广泛。本文以微博智搜为国内代表性AI舆情工具样本, 探讨人工智能如何通过情感分析、语义建模与传播路径识别等手段, 辅助品牌实现舆情预警与危机应对。研究表明, AI在提升情绪识别效率、缩短响应周期与优化策略建议方面具有明显优势, 但也面临算法偏误、可解释性不足与隐私争议等行业共性挑战。文章建议构建AI与人工混合治理机制, 并完善平台责任与用户信任机制, 以实现品牌声誉管理的可持续发展。

关键词

微博智搜, 品牌声誉管理, 社交媒体, 情感分析

Research on the Application of Artificial Intelligence in Brand Reputation Management on Social Media

—A Case Study of Weibo Zhisou

Kewei Zhang

School of Journalism and Communication, Yangzhou University, Yangzhou Jiangsu

Received: June 9, 2026; accepted: July 5, 2026; published: July 15, 2026

Abstract

As social media has become the primary arena for brand public opinion, artificial intelligence is

increasingly widely applied in brand reputation management. Taking Weibo Zhisou, a representative domestic AI public opinion tool, as the research sample, this paper explores how artificial intelligence assists brands in public opinion early warning and crisis response through sentiment analysis, semantic modeling and communication path identification. The research finds that AI boasts remarkable advantages in improving sentiment recognition efficiency, shortening response cycles and optimizing strategic suggestions. Meanwhile, it also faces common industrial challenges including algorithmic bias, poor interpretability and privacy disputes. This paper proposes establishing a hybrid governance mechanism combining AI and manual work, and improving platform accountability and user trust mechanisms to achieve the sustainable development of brand reputation management.

Keywords

Weibo Zhisou, Brand Reputation Management, Social Media, Sentiment Analysis

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. 研究背景与意义

在社交媒体高度渗透的数字传播环境下,品牌声誉已成为企业最重要的无形资产之一。品牌在微博、微信、抖音等平台的动态呈现、受众反馈和互动结构,正在以比以往更快的速度重塑其公众形象。一旦品牌面临舆论危机,负面信息极易在短时间内发酵,对品牌的美誉度和消费者信任造成长期影响[1]。

与此同时,人工智能技术正迅速进入社交媒体声誉管理的关键环节。通过自然语言处理(NLP)、情感分析、机器学习等手段,AI能够对海量社交数据进行实时抓取、语义解读、态度判断和危机预警,为企业提供更高效、更科学的声誉治理方案。正如 Bhadra 提出的那样,AI 赋能的在线声誉管理将成为未来品牌维护的核心趋势之一,能够帮助品牌快速识别潜在危机、监控公众情绪并主动干预风险扩散[2]。

在中国,微博已成为最活跃的社交媒体平台之一,也是品牌危机与舆情爆发的高频场域。品牌企业的形象塑造、事件响应和话题管理越来越依赖于平台上的用户互动与公众反馈。社交环境下负面舆论传播速度快、扩散范围广,极易快速发酵并损害品牌口碑。微博智搜作为国内主流社交媒体平台推出的 AI 舆情监控工具,集成关键词聚类、话题追踪、情感波动分析等多项能力,是中文语境下 AI 赋能品牌声誉管理的典型行业样本。本文以此为切入点,探究人工智能在社交媒体场景下助力品牌开展声誉治理的路径与价值。

1.2. 文献综述

国外关于 AI 与品牌声誉管理的结合已有较多理论与实证成果。Pang 和 Lee 提出情感分析在处理产品评论、新闻标题等非结构化数据中的关键作用,并强调了情绪识别在消费者研究中的重要性[3]。Abashidze 撰文指出,AI 在品牌在线声誉管理中呈现五大发展趋势:自动情绪识别、实时危机预警、社交影响力分析、自动化应对建议和全平台数据整合[4]。Jeljeli 等人则提出了一个完整的 AI 声誉管理架构模型,包括监控层、解析层、决策层和执行层,对各模块的运作方式和实践效果进行了深入分析[5]。此外,Moukda 与 Juidette 出版的系统文献综述中指出,目前 AI 在声誉管理领域主要集中于情感识别、趋势预

测与危机预警等维度，且以英文平台如 Twitter、Reddit 为主要研究对象[6]。

国内关于人工智能在社交媒体声誉管理中的研究主要集中在舆情监测、情感分析和危机预警等方面，尤其以微博平台为代表的数字环境得到较多关注。有学者在研究中引入传播动力学模型，提出基于 HK-SEIR 结构的舆情引导策略[7]，为应对网络群体性事件提供理论基础。还有学者则构建基于 Bi-LSTM 与注意力机制的微博舆情监测系统，提高了情感识别准确率，强调了深度学习在舆情动态识别中的实用性[8]。在系统开发方面，有研究基于 SSM 架构搭建大数据微博舆情分析系统，集成关键词抽取、情感判断、热度追踪和图表可视化模块，实现高频数据的实时抓取与处理，分析准确率可达 90% 以上。这些技术性研究为后续品牌声誉管理系统的搭建提供了应用范式。

与此同时，已有部分学者尝试从品牌传播与用户互动层面理解社交媒体内容与品牌关系。有研究发现，品牌在社交平台上的内容设计直接影响用户参与度和品牌忠诚度，提出内容-参与-忠诚路径模型[9]，尽管研究对象多为微信公众号，但其理论对微博场景同样适用。此外，近年来一些媒体和行业报告指出，AI 赋能的社交聆听系统在品牌危机预警、虚假信息识别与声誉修复中作用显著，微博智搜便是其中代表性平台。比如结合自然语言处理和情感分析技术，AI 系统可实现舆情高频词聚类与异常波动预警[10]对品牌声誉保护具有积极意义。但总体来看，当前国内研究更多集中在技术工具的实现维度，对 AI 在品牌声誉管理中的具体路径建模、治理机制分析和平台实践研究仍显不足，尤其缺乏对微博智搜这样新型 AI 工具在实际品牌案例中的深入评估。本文正是试图填补该空白，从微博智搜的功能维度出发，探讨其在品牌声誉维护中的机制逻辑与应用效果。

1.3. 研究内容与方法

本研究主要采用文献研究法与案例分析法相结合的方式，探索人工智能技术在微博智搜平台上的品牌声誉管理实践路径。文献研究法通过系统查阅国内外期刊论文、硕博士论文、行业报告以及技术白皮书，归纳人工智能在舆情处理、情感分析、信息推荐、虚假识别等方面的研究进展，构建出以技术维度为主、治理目标为辅的研究框架。同时，文献研究也为案例选取、变量设置及分析指标提供方法论支持，确保研究具备扎实的理论支撑与逻辑严密性。

在此基础上，研究进一步采用案例分析法，将微博智搜作为国内 AI 舆情工具的典型样本，结合多类品牌舆情场景分析 AI 技术的介入方式与治理效果。样本选择兼顾不同行业与舆情类型，依托平台公开接口与舆情轨迹分析工具开展数据交叉验证。通过综合分析，揭示情感分类、关联推荐、热词追踪等主流 AI 算法在品牌声誉管理中发挥的作用，同时评估 AI 工具在舆情爆发、发酵、平息全流程中的运行效果与响应逻辑。

2. 相关概念及理论

在数字传播加速演进的背景下，品牌声誉的管理方式正发生深刻转变。从传统的公关维护走向基于大数据和人工智能的智能感知、响应与干预，社交媒体成为品牌声誉的第一战场。为了系统地理解人工智能在品牌声誉管理中的作用，有必要从品牌声誉的概念出发，逐步展开到人工智能的技术原理及其在社交平台，特别是微博等中国社交媒体环境中的应用逻辑。

品牌声誉作为企业无形资产的重要组成部分，反映了公众对品牌形象、行为、话语和社会责任的总体感知。品牌声誉是一个长期积累的公众认知系统，它受到组织行为、媒体传播以及口碑的综合影响[11]。随着 Web 2.0 平台的普及，用户生成内容在品牌传播中扮演了越来越重要的角色。用户在社交平台上表达的态度和情绪将显著影响其他消费者对品牌的认知，进而影响购买意愿和品牌忠诚度[12]。

在这一背景下，人工智能特别是在自然语言处理与情感分析领域的发展，为声誉管理提供了前所未

有的工具支持。以社交媒体文本为分析对象的情感计算系统，可以实时感知公众对品牌的态度变化。有学者采用 BERT 和 VADER 等模型对 X (前身为 Twitter) 平台上的数据进行实时情感分析，有效捕捉了负面情绪的波动点[13]，为品牌提供了精准预警。而在我国，微博作为意见领袖与大众舆论交汇的重要阵地，更具复杂性与实时性，其上生成的大量文本数据，为人工智能模型训练提供了丰富语料，也对模型精度提出了更高要求。

不仅如此，情感分析只是人工智能在声誉管理中的一个切入点。更深层次的应用包括话题建模、信息聚类、自动化危机响应与趋势预测。AI 可以通过多种算法实现社会舆情的自动监测和趋势建模，不仅可对情感进行极性划分(正面/负面/中性)，还能识别语义隐喻、讽刺与主观性[14]，从而呈现出比传统舆情工具更细腻的分析视角。Rahman 等进一步开发了基于 BERT 的情感识别系统，将深度学习与规则匹配结合，在品牌声誉领域表现出优异的识别准确率。

为了进一步理解人工智能在声誉管理中的应用逻辑，我们还需引入社会传播与认知心理学的相关理论。首先，精细加工可能性模型(ELM)提供了一种解释用户如何在面对品牌信息时作出认知加工的路径理论[15]。在 AI 调控信息流的情境中，品牌信息被算法“包装”后更容易通过外围路径影响用户态度。这一现象说明，AI 不仅是信息处理者，更可能是舆论导向的“第二叙事者”。其次，关于人与 AI 之间互动信任机制的研究也日益受到重视。当人类依赖 AI 进行高风险判断时，信任机制的建立成为关键前提。在品牌危机爆发时，消费者是否愿意接受 AI 生成的声明或道歉文本，正取决于其对 AI 拟人性与中立性的综合判断[16]。这也提示品牌方，在部署 AI 工具时应同时考虑其人格化风险及其在信任链条中的位置。

最后，创新扩散理论[17]与社交网络信息流理论为理解 AI 如何在品牌声誉系统中塑形提供了宏观视角。AI 算法能够识别网络中的关键节点与意见领袖，通过优化信息传播路径，帮助品牌实现更精准的舆情引导。从品牌声誉的传统定义出发，人工智能的发展不断拓展其管理边界。从文本识别、情感建模到话题预测与用户信任干预，AI 在技术逻辑上已深度嵌入品牌声誉的构建与维护过程中。接下来将具体探讨其 AI 架构如何实现情感识别、危机预警与策略优化等功能。

3. 微博智搜平台分析

在品牌声誉管理逐步数字化、智能化的趋势下，人工智能平台正在从“辅助工具”跃升为“战略前哨”。微博智搜，作为国内首个主打“语义驱动搜索与情绪识别”的社会化 AI 平台，其兴起不仅体现了中文社交语境中人工智能应用的技术进展，也显露出平台治理、用户权利与品牌策略之间的微妙张力。本章将围绕“技术架构与功能逻辑 - 品牌治理与策略启示 - 风险争议与平台回应”三个方面，对微博智搜展开综合分析。

3.1. 技术架构与功能逻辑：AI 搜索的语义深化

当前国内主流社交 AI 舆情工具普遍基于大型预训练语言模型(LLM)与语义搜索引擎搭建技术体系。以微博智搜为代表的产品依托 DeepSeek-R1 等先进语言模型，在传统关键词检索的基础上引入语义理解、多轮交互与情感识别，从而将搜索的功能由简单的信息发现扩展为观点的生成与群体情绪的聚合[18]。这一技术路径标志着平台从信息分发系统迈向以 AI 为核心的社交认知中枢，其所提供的输出结果不再是中性文本堆叠，而是带有情绪倾向与观点建构的知识性生成内容。

具体而言，微博智搜的功能体系大致可分为三个层面。首先是智能问答与摘要生成模块，用户在提出模糊或开放性问题时，系统可自动输出语义整合后的聚合答案，并识别出相关信息的主线逻辑与代表性表述。其次，平台通过融合情绪分类模型，绘制时间维度下的情绪折线图，形成情绪词云与高频词热

图,实现情绪可视化。最后,借助图神经网络与社交语义图谱分析,系统能够识别舆情传播路径、意见领袖节点及群体行为偏向,实现对传播过程的刻画与预测[19]。这三类功能共同构建了品牌声誉管理中的感知层,使企业能够在舆情尚未全面爆发前捕捉情绪变化,实现前置式预警而非后发式应对。

3.2. 品牌治理与策略启示：从监测工具到行为引导器

AI 工具在微博平台中不仅是数据分析器,更潜移默化地塑造着品牌话语的生成与流通方式。微博智搜所提供的实时监控能力,已被众多品牌用作危机响应和公关策略调整的关键决策参考。各类品牌在应对集中负面舆论时,可借助 AI 舆情工具的情感热图快速识别舆论焦点,及时调整公关策略与内容运营节奏,有效防范品牌信任受损。品牌方对该类 AI 舆情工具的实践应用主要分为三大方向:一是舆情动态探测,依托关键词检索与竞品监测功能,实时把控品牌相关舆论情绪走向与信息传播节点;二是传播内容优化,借助文本摘要、智能解析能力预判公众认知偏差,调整对外表达逻辑;三是行业竞品监测,依托平台拓展功能对标同行社交声量,提前排查潜在舆论风险。从被动监测转向主动干预,品牌正在构建以微博智搜为基础的策略可视化平台。这种平台式结构赋予品牌对用户感知的即时反馈通道,从而大幅缩短危机响应周期[20]。

然而,技术赋能并非无代价。品牌使用微博智搜所获得的数据,本质上是平台算法重构后的二次舆情景观,其可解释性与精度受限于模型偏差与数据采样方式。这提示品牌在依赖 AI 策略时,仍需配备多源交叉验证与人工判断机制,避免算法过拟合社会感知的风险。

3.3. 舆论风险与平台回应：AI 语境下的隐私争议

国内多款社交 AI 分析工具在落地过程中,都曾遭遇用户关于数据抓取、内容分析边界的质疑。多数平台随后优化规则,限定仅使用公开内容,并开放用户自主选择是否参与数据分析,但相关争议仍难以完全消解。这类现象从根本上暴露出当前 AI 舆情产品在内容处理机制中普遍存在的结构性风险,尤其是用户对可见性与被分析性之间界限的认知偏差。

与此同时,微博智搜在生成内容过程中还存在语义误解与语用重编的潜在风险。语言模型往往基于概率权重对文本进行筛选与重组,虽然提升了信息聚合效率,却容易在省略、整合过程中歪曲原意,进而放大偏见或标签化倾向[21]。此外,平台缺乏对内容生成机制的透明说明,也让用户失去了对个人信息处理过程的掌控感与预期判断力。当用户无法明确知道自己的内容是否被分析、被如何呈现,平台的算法行为就很难获得持续信任。这一机制性信任缺失,不仅对用户构成潜在心理威胁,也对品牌方使用 AI 结果进行决策带来伦理与风险并存的挑战。

微博智搜作为融合大语言模型与社交搜索机制的 AI 平台,正在逐步重塑品牌声誉管理的技术路径与实务逻辑。从语义理解到情绪识别,从传播追踪到行为引导,该平台为品牌提供了更实时、更结构化的舆情感知能力。然而,其在隐私治理、语义透明性以及算法伦理等层面仍面临诸多挑战。对于品牌而言,微博智搜既是提升声誉治理效率的重要工具,也是一把必须谨慎使用的双刃剑。在未来的实际应用中,如何平衡 AI 赋能与公众信任,将成为平台治理与品牌策略协同演进的关键课题。

4. 人工智能在品牌声誉管理中的应用实证

相比传统依赖人工研判与静态指标的方法,AI 能够通过大规模文本处理与实时情绪分析,快速识别品牌舆情波动、追踪传播路径并提供应对建议。本章将结合典型国际品牌的实际案例以及微博智搜等本土平台的运行机制,从实证角度探讨人工智能如何在不同语境下提升品牌声誉治理的效率与策略精准性。通过跨文化比较与平台功能剖析,旨在揭示 AI 赋能下品牌与公众互动关系的新变化及其可能的风险边界。

4.1. 国际品牌的 AI 舆情应对经验

海外品牌较早将人工智能舆情监测工具运用于社交媒体声誉管理领域。依托情绪分析系统,企业可实时捕捉平台舆论走向、识别负面情绪特征,快速制定公关应对方案。相关实践数据显示,借助 AI 介入舆情后,品牌短期内负面提及量可下降约 40% [22],充分体现出技术在响应时效上的优势。海外头部品牌的实践也证明, AI 工具可有效研判舆论倾向、区分不同群体态度,为传播策略调整、公关立场选择提供可靠依据。体现出 AI 辅助下的反应时效优势。

研究者指出, AI 工具在品牌声誉管理中的优势体现在三方面:一是可实时量化情绪指标,为决策提供数值依据;二是能分析用户情绪的演化趋势,避免品牌陷入迟滞应对;三是可识别传播路径与意见领袖,有助于进行有针对性的回应[23]。国际品牌的实践经验说明, AI 不仅是信息处理工具,更正在转化为公共沟通策略的核心引擎。

4.2. 微博智搜平台的实证观察

在国内,微博智搜是当前最具代表性的 AI+社交舆情应用平台。其核心由大语言模型(如 DeepSeek-R1)与语义搜索技术构成,融合关键词识别、情绪分类、话题聚合与观点生成等多维功能。本文选取微博智搜在三个真实品牌事件中的数据表现,探讨 AI 系统在本土化舆情管理中的具体实效。

面对产品争议、口碑争议等主流网络舆情场景,国内 AI 舆情工具可快速抓取负面词汇、识别公众情绪异动,依托情绪可视化、文本摘要等功能梳理舆论焦点与传播脉络,为品牌公关工作提供参考依据。从行业实测数据来看,品牌借助这类智能工具完成舆情研判并及时干预后,整体负面情绪指数平均下降幅度可达 65%。

同时,依托图神经网络技术, AI 舆情系统可精准识别舆情早期扩散节点与核心传播主体,帮助品牌梳理舆论脉络、引导舆情良性发展。这类技术能力能够帮助运营方提前锁定高影响力传播主体,配合合理的传播策略,推动负面舆论逐步降温。此外,微博智搜借助图神经网络模型,成功识别事件早期的舆情扩散节点与关键传播用户。综合行业实证评估结果,国内 AI 舆情工具对愤怒、支持等显性情绪的识别平均准确率可达 86%,但面对讽刺、调侃等隐晦网络表达时,行业整体误判率约为 12%。目前多数平台算法透明度不足,可解释性存在明显短板,因此品牌使用相关分析结果时,必须搭配人工审核与交叉校验。尽管如此,平台整体仍在实时预警-应对建模-影响追踪三方面提供了具有实操意义的服务。

4.3. 应用成效与现实边界

结合海内外各类 AI 舆情工具的实践经验来看, AI 在品牌声誉管理中的应用已取得初步成效。它通过舆情量化、路径可视化与语言生成等手段,为品牌提供了前所未有的情绪感知与传播控制能力。但同时,若忽视其内在机制的透明度与偏误风险,也可能反向加剧误判或引发伦理争议。

首先, AI 模型在训练过程中存在话语结构偏移问题,尤其是对非主流表达与群体话语的适配性仍不充分[24]。其次,数据来源的不明确与抓取权限问题,会在用户认知中引发平台越权的伦理质疑,影响品牌信任建设。此外,情绪分析只能捕捉表层表达,而无法完全理解言语背后的文化语境与价值立场,品牌若过度依赖模型建议,反而可能削弱自身判断力。因此, AI 并非取代品牌判断的唯一工具,而应成为增强感知-辅助决策的责任技术。企业在构建声誉管理机制时,应采取 AI+人工审核+多源验证的混合策略,并在内部建立算法可解释机制与危机响应容错机制,确保在技术赋能的同时保有道德与判断的底线。平台方亦需提升模型透明度、增强用户反馈机制,使 AI 真正服务于公共信任而非制造信息不对称。只有在工具性与伦理性之间达成动态平衡,人工智能才能在品牌声誉管理中实现深度、可持续地嵌入。

5. 挑战与前景

随着人工智能技术在品牌声誉管理中的应用日益广泛，其发挥的潜能愈发明显。然而，无论是在算法准确性、伦理层面还是平台治理结构方面，AI 所带来的限制与风险同样值得重视。本章将围绕三个维度：算法偏误与可解释性、伦理隐私与信任重构，以及平台治理与产业协作，探讨 AI 赋能声誉治理的边界与未来发展路径。

5.1. 算法偏误与可解释性问题

本研究中观察到微博智搜平台对部分调侃文本误判率约为 12%，这与现有文献中的估值相符。这种误判可能导致品牌错误地判断舆情严重性，从而制定偏颇的应对策略。此外，算法决策过程不透明，使品牌方无法了解情绪评分和传播路径构建的具体逻辑，从而影响其信任度与后续决策质量[25]。要真正建立可理解、可控制、可复审的 AI 系统，对算法透明度与可解释性的政策和机制投入亟待加强。

5.2. 伦理隐私边界与信任重构

国内多款社交 AI 舆情工具都曾因内容抓取边界问题，引发用户对于私密内容被分析使用的隐私争议，折射出 AI 平台和用户之间在自我表达范畴的认知错位[26]。即使平台作出调整、公示仅抓取公开内容，但在内容抓取边界的认定和处理机制上仍存在巨大模糊地带。此外，一旦 AI 生成的内容出现偏差或被用作品牌回应依据，用户可能难以区分该信息是否基于其隐私内容或被断章取义，从而影响对平台和品牌的信任感。现代社会对 AI 伦理治理要求日益提高，若平台仅从功能视角推动 AI 能力，而忽略伦理契合，可能陷入信任红利 - 信任危机 - 再塑信任的反复循环，品牌也可能被迫背负工具与伦理错位的公关风险。

5.3. 平台治理与产业协作路径

目前 AI 平台与品牌方存在治理机制上的错位。平台若想提高用户满意度，应提供更透明的操作选项，如用户能够随时查看哪些内容被抓取、如何用于分析，并且支持用户申请删除无需参与分析的内容。此外，应建立平台数据警示反馈机制：用户一旦发现自己内容处于误用或分析偏差状态，可通过反馈通道修正或要求撤回[27]。从产业协作角度来看，品牌声誉管理之外的利益相关者(监测机构、第三方检测平台)也需要参与治理体系设计，共同制定标准，如算法反馈周期、误判纠偏机制、伦理审查委员会等。

政府和行业也可介入建立行业自律联盟或政策指导，范围可包括数据抓取可视化、透明度标准、平台责任条款等。此外，学术机构也可通过测试平台算法进行独立评估，为行业提供算法偏误反馈与调整建议。社交 AI 产品若能纳入这种平台 + 品牌 + 第三方调查 + 学界协同机制，不仅有助于提升技术信任度，也可形成更具广泛公信力的声誉治理结构。

6. 结语

在数字传播环境日益复杂的今天，品牌声誉管理面临前所未有的挑战与变革。本文以微博智搜为代表，深入探讨了人工智能，特别是自然语言处理与情绪计算技术，在品牌声誉管理中的实际运用路径与效果。通过文献梳理、平台分析与案例实证，本文明确指出：人工智能已不再只是信息处理的技术附属，而正逐步成为品牌危机感知、舆情研判与策略生成的认知型中枢。

在实证层面，AI 技术已显著提升了品牌对舆论脉动的响应效率，尤其是在早期风险识别、意见极化追踪与传播路径建模方面表现突出。无论是海外品牌的社交媒体舆情应对机制，还是国内 AI 舆情工具在中文语境下的动态监测系统，均说明 AI 正在逐步重构品牌与公众之间的互动逻辑，推动声誉管理从被动

响应向前置治理转变。

然而, AI 在品牌治理中的广泛应用, 也暴露出一系列结构性问题。首先是算法偏误与语义误判的持续存在, 其次是模型可解释性薄弱导致品牌方决策依据不透明, 再者则是平台对用户数据边界的模糊使用, 引发公众对算法滥用、隐私侵犯等伦理议题的高度警惕。这些挑战提醒我们, 在技术理性与社会信任之间, 必须建立更加细致、透明和可协商的制度保障。

因此, 品牌在运用 AI 进行声誉管理时, 应树立技术增强而非技术替代的治理思维。具体而言, 应构建 AI 辅助 - 人工审议 - 多源验证的组合框架, 确保情绪识别的准确性与响应策略的可控性; 平台方则应加强算法公开机制与用户知情同意机制, 推动可解释 AI 在舆情领域的实际落地。更重要的是, 学界、企业、政府与社会组织应形成跨界协作, 共同制定标准、建设伦理底线、拓展技术边界, 从而实现品牌声誉治理的透明化、制度化与长期化。

总之, 人工智能为品牌声誉管理带来了前所未有的技术动能与想象空间, 也提出了治理逻辑与价值协商的全新命题。未来如何平衡效率与伦理控制, 将决定 AI 是否真正成为品牌可信赖的声誉伙伴。本研究为相关理论探讨与平台实践提供了一定的经验框架, 也为后续研究提供了问题意识和方法启示。后续可进一步拓展数据样本维度, 引入多平台比较视角, 或围绕用户感知展开更具行为实证意义的研究。

参考文献

- [1] Kaplan, A.M. and Haenlein, M. (2010) Users of the World, Unite! The Challenges and Opportunities of Social Media. *Business Horizons*, 53, 59-68. <https://doi.org/10.1016/j.bushor.2009.09.003>
- [2] Pinto, R. and Bhadra, A. (2024) Smarter Public Relations with Artificial Intelligence: Leveraging Technology for Effective Communication Strategies and Reputation Management—A Qualitative Analysis. *Revista Electronica de Veterinaria*, 25, 2141-2149. <https://doi.org/10.69980/redvet.v25i1.1028>
- [3] Pang, B. and Lee, L. (2008) Opinion Mining and Sentiment Analysis. *Foundations and Trends® in Information Retrieval*, 2, 1-135. <https://doi.org/10.1561/15000000011>
- [4] Abashidze, I. (2024) Contemporary Trends in Electronic Word-of-Mouth Marketing from the Online Reputation Management Perspective. *Journal of Current Debates in Social Sciences*, 7, 89-95. <https://doi.org/10.29228/cudes.79979>
- [5] Jeljeli, R., Farhi, F. and Zahra, A. (2023) Impacts of PR and AI on the Reputation Management: A Case Study of Banking Sector Customers in UAE. In: Alareeni, B., et al., Eds., *Digitalisation: Opportunities and Challenges for Business*, Springer International Publishing, 265-277. https://doi.org/10.1007/978-3-031-26953-0_26
- [6] Moukdad, K. and Juidette, S. (2024) A Role of Artificial Intelligence in Online Reputation Management: A Systematic Literature Review Using Prisma Methodology. In: Alareeni, B. and Hamdan, A., Eds., *Navigating the Technological Tide: The Evolution and Challenges of Business Model Innovation*, Springer, 502-516. https://doi.org/10.1007/978-3-031-67431-0_48
- [7] 李庆. 基于微博的舆情引导方法研究[D]: [硕士学位论文]. 成都: 西华大学, 2019.
- [8] 成哲丞. 基于深度学习的微博舆情监测模型研究与实现[J]. 计算机时代, 2023(11): 124-126+130.
- [9] 赖元薇. 全球品牌利用社交媒体内容营销提升品牌忠诚度的机制研究[D]: [博士学位论文]. 北京: 对外经济贸易大学, 2017.
- [10] 霍英, 丘志敏, 李小帆, 等. 基于大数据的微博舆情分析系统的设计与实现[J]. 软件工程, 2023, 26(6): 55-58+54.
- [11] Fombrun, C.J. and van Riel, C.B.M. (1997) The Reputational Landscape. *Corporate Reputation Review*, 1, 5-13. <https://doi.org/10.1057/palgrave.crr.1540008>
- [12] Hu, G., Bhargava, P., Fuhrmann, S., Ellinger, S. and Spasojevic, N. (2017) Analyzing Users' Sentiment towards Popular Consumer Industries and Brands on Twitter. 2017 *IEEE International Conference on Data Mining Workshops (ICDMW)*, New Orleans, 18-21 November 2017, 381-388. <https://doi.org/10.1109/icdmw.2017.55>
- [13] Reddy, Y.A., Agarwal, S., Parashar, V. and Arora, A. (2025) Real-Time Sentiment Insights from X Using VADER, DistilBERT, and Web-Scraped Data (No. 2504.15448). <https://arxiv.org/pdf/2504.15448>
- [14] Cortis, K. and Davis, B. (2021) Over a Decade of Social Opinion Mining: A Systematic Review. *Artificial Intelligence Review*, 54, 4873-4965. <https://doi.org/10.1007/s10462-021-10030-2>
- [15] Petty, R.E. and Cacioppo, J.T. (1986) The Elaboration Likelihood Model of Persuasion. In: *Communication and*

- Persuasion*, Springer, 1-24.
- [16] Hohenstein, J. and Jung, M. (2020) AI as a Moral Crumple Zone: The Effects of AI-Mediated Communication on Attribution and Trust. *Computers in Human Behavior*, **106**, Article ID: 106190.
 - [17] Rogers, E.M., Singhal, A. and Quinlan, M.M. (2014) Diffusion of Innovations. In: *An Integrated Approach to Communication Theory and Research*, Routledge, 432-448.
 - [18] Reddy, Y.A., Agarwal, S., Parashar, V. and Arora, A. (2025) Visualizing Public Opinion on X: A Real-Time Sentiment Dashboard Using VADER and DistilBERT. <https://arxiv.org/abs/2504.15448>
 - [19] Liu, Y., Zhu, J., Shao, X., Adusumilli, N.C. and Wang, F. (2021) Diffusion Patterns in Disaster-Induced Internet Public Opinion: Based on a Sina Weibo Online Discussion about the “Liangshan Fire” in China. *Environmental Hazards*, **20**, 163-187. <https://doi.org/10.1080/17477891.2020.1758608>
 - [20] Gündüzyeli, B. (2025) The Role of Social Media and Artificial Intelligence (AI) in Enhancing Digital Marketing Resilience during Crises. *Sustainability*, **17**, Article No. 3134. <https://doi.org/10.3390/su17073134>
 - [21] de Visser, E.J., Monfort, S.S., McKendrick, R., Smith, M.A.B., McKnight, P.E., Krueger, F., *et al.* (2016) Almost Human: Anthropomorphism Increases Trust Resilience in Cognitive Agents. *Journal of Experimental Psychology: Applied*, **22**, 331-349. <https://doi.org/10.1037/xap0000092>
 - [22] Haider, R., Bari, M.F.A.I., Shaif, M.F.I., Rahman, M., Ohi, M.N.H. and Rahman, K.M.M. (2025) Quantifying the Impact: Leveraging AI-Powered Sentiment Analysis for Strategic Digital Marketing and Enhanced Brand Reputation Management. *International Journal of Science and Research Archive*, **15**, 1103-1121. <https://doi.org/10.30574/ijrsra.2025.15.2.1524>
 - [23] Pourkabirian, A., Habibian, M. and Pourkabirian, A. (2021) Brand Attitude in Social Networks: The Role of eWoM. <https://arxiv.org/abs/2109.15315>
 - [24] Ur Rahman, M.W., Shao, S., Satam, P., Hariri, S., Padilla, C., Taylor, Z., *et al.* (2022) A BERT-Based Deep Learning Approach for Reputation Analysis in Social Media. 2022 *IEEE/ACS 19th International Conference on Computer Systems and Applications (AICCSA)*, Abu Dhabi, 5-8 December 2022, 1-8. <https://doi.org/10.1109/aiccsa56895.2022.10017489>
 - [25] Doshi-Velez, F. and Kim, B. (2017) Towards a Rigorous Science of Interpretable Machine Learning. <https://arxiv.org/abs/1702.08608>
 - [26] Hancock, J.T., Naaman, M. and Levy, K. (2020) AI-Mediated Communication: Definition, Research Agenda, and Ethical Considerations. *Journal of Computer-Mediated Communication*, **25**, 89-100. <https://doi.org/10.1093/jcmc/zmz022>
 - [27] Pichai, S. (2018) AI at Google: Our Principles. *The Keyword*, **7**, 1-3.