ISSN 2616-5783 Vol.9, Issue 2: 52-59, DOI: 10.25236/AJHSS.2026.090208

The Privacy Crisis of AI Invasion in Social Media: The Case of Weibo Social Bot "Robot"

Gongli Liu[a,*], Chunhua Zhang[b]

School of Journalism, Nanjing University of Finance and Economics, Nanjing, China
[a]1477768943@qq.com, [b]370962733@qq.com
[*]Corresponding author

Abstract: *As artificial intelligence technology continues to evolve, social bots are becoming increasingly prevalent in social media applications. However, their potential implications for user privacy are becoming increasingly evident. This research employs semi-structured interviews to examine users' perceptions of privacy violations and corresponding protective actions during interactions with the social bot, "Robot," which was launched by Weibo in July 2023. The findings indicate substantial cognitive variations among users with regard to the privacy risks associated with the intervention of social bots in personal social content. In light of these findings, the present paper puts forth a series of recommendations in order to explore strategies that would allow for a balanced approach between the pursuit of technological innovation and the protection of privacy. These insights are intended to inform the delineation of ethical boundaries for AI in social media and the governance of user privacy.*

Keywords: *Human-machine interaction; Privacy protection; Social bot; Robot*

1. Introduction

The field of artificial intelligence is undergoing rapid innovation, with technological developments leading to the emergence of increasingly sophisticated applications. These advancements are profoundly impacting the landscape of social media communication, thereby reshaping established paradigms. Social bots represent a prime exemplar of human-machine interaction. These bots are capable of automatically generating content and mimicking emotions and behaviors. Consequently, they have become pivotal players in information dissemination, steering public discourse and facilitating user engagement. However, while these bots do indeed boost platform activity, they also reveal latent risks of privacy infringement. These bots demonstrate a high level of proficiency in data mining and analysis, facilitating real-time user behavior tracking through the utilization of natural language processing techniques to infer private preferences. This development has the potential to render the ethical boundaries of human-machine interaction increasingly indistinct. In this context, the investigation of the threat mechanisms posed by social bots to user privacy, and the exploration of governance approaches, have become focal points for both academia and industry.

According to scholars such as Boshmaf, social bots are defined as "intelligent programs that autonomously operate social media accounts within networks and possess the capability to automatically send messages and links[1]." In recent years, the proliferation and application of social bots, driven by algorithmic intelligence, have attracted significant academic attention, focusing on their technological development, ethical considerations, and societal impacts. Domestic research in this field, though relatively recent, has developed rapidly. In their 2021 study, Shen Qi and Wang Luyu examined the growing tendency to regard social bots as social actors. They explored the role of these bots in simulating interpersonal interactions and analyzed the significance of stereotype studies on social bots. For this analysis, the researchers drew upon media equivalence theory and the computer-as-social-actor paradigm[2]. Shi Wen and Chen Changfeng (2020) employed data mining and analysis methods to study the distribution and interactions of China-related topics on Twitter. The researchers identified behavioral patterns of social bots in manipulating public opinion and their interactive relationships with humans. It was noted that social bots increase human users' exposure to specific information, successfully infiltrate social networks, and alter information interaction structures[3]. Concurrently, some scholars have conducted in-depth research on the impact of social bot development on the modernization of national governance. For instance, Gao Qiqi (2020) focuses on the overall structural

ISSN 2616-5783 Vol.9, Issue 2: 52-59, DOI: 10.25236/AJHSS.2026.090208

impact of the intelligent revolution on the modernization of national governance, providing an in-depth analysis of how artificial intelligence and blockchain technologies reshape and influence national governance systems and capabilities[4]. In response to the ethical issues arising from the application of social bots, Wang Liang (2020) analyzes the imbalance between the artificial emotionality of social bots and the authenticity of human emotions. The author posits that the ethical concerns associated with social bots primarily stem from their potential to "manipulate" and "deceive[5]" human empathy. According to scholars such as Nick Hajli (2021), AI-driven social bots employ disinformation on social media platforms to manipulate public opinion. In response, scholars have proposed methods for the early detection of malicious bots[6]. In the study conducted by Dou Xiao (2022), five potential ethical issues were identified: social adaptation challenges, emotional manipulation, behavioral deviance, role confusion, and identity crises. The study also offered targeted governance policy recommendations[7]. In their 2023 study, Chen Xinyi and Mao Mingfang observe that while social bots can offer benefits in terms of production and daily life, they also present significant risks. These risks include the potential to challenge individual rationality and to threaten democratic peace. The sustainable and healthy development of technology and society is contingent upon collaborative governance involving platforms, governments, individuals, and other stakeholders[8]. Kalukuri Princy Niveditha (2023) examined the impact of AI recommendation algorithms on social media content distribution. She balanced her analysis by considering both the role of AI in enhancing user engagement and its potential for spreading misinformation[9]. In their 2024 systematic review, Musawer Hakimi et al. provide a comprehensive analysis of artificial intelligence applications in the domain of social media security. The present study examines the strengths, weaknesses, and future directions of AI in content moderation, threat detection, and real-time response mechanisms[10].

The aforementioned research has provided a relatively comprehensive exploration of issues such as the definition, behavioral patterns, social impact, and foreseeable risks of social bots. However, there is still room for growth in terms of their current practical applications on domestic social platforms and the ethical risks and privacy violations that arise from interactions with users. Taking Weibo as an example, its 2023 launch of the social bot "Robot" enabled AI to participate in user content interactions through automated commenting, signaling the gradual integration of artificial intelligence into online community culture. However, shortly after its implementation, this technology sparked a series of questions: Does the bot excessively collect user data during interactions? How do users perceive and address such privacy threats? Addressing these concerns is crucial not only for protecting individual privacy rights but also for ensuring the sustainable development of the social media ecosystem. Therefore, this paper examines "Robot" as its research subject, employing semi-structured interviews to gain deeper insights into the following core issues: the dimensions of privacy infringement and differences in risk perception during user-bot interactions; the factors driving users to implement privacy protection behaviors and their implementation pathways; the ethical boundaries of social bots on platforms and feasible strategies for maintaining these boundaries.

2. Research Design

In his 1967 publication, Privacy and Freedom, Alan Westin categorized public attitudes toward privacy concerns into three groups: individuals troubled by privacy-invading actions, individuals indifferent to such actions, and individuals who adopt a pragmatic stance and weigh each situation individually[11]. Building upon this framework, this study employed semi-structured interviews. During the sampling phase, the researcher conducted a keyword search for "Robot" on the Weibo platform. Unrelated content was manually curated and excluded, and judgmental sampling was applied among netizens who had interacted with "Robot." To ensure representativeness and credibility, the study applied the following criteria when selecting interviewees: Respondents must be active Weibo users who have had accounts for at least one year, ensuring stable experience with the platform's functions and the "Robot" feature. They must have interacted with "Robot" at least once. To capture differences in privacy perceptions across age groups, the researchers sought interviewees with diverse occupations and a broad age range during the sampling process. Thus, through a combination of purposive and snowball sampling, the study ultimately selected 12 participants, coded N1-N12 and ranging in age from 18 to 50. Their demographic characteristics are shown in Table 1. Data collection concluded in February 2025 using three formats: In-person interviews, online video interviews, and telephone interviews. Each interview lasted between 30 and 60 minutes. In adherence to ethical standards, the researchers explained the purpose of the study to the participants beforehand and committed to anonymizing the personal information and interview content.

ISSN 2616-5783 Vol.9, Issue 2: 52-59, DOI: 10.25236/AJHSS.2026.090208

Table 1: Basic Information of Respondents

Respondent ID	Age	Gender	Occupation	Educational background	The total number of interactions with "Robot"
N1	30	Female	Media practitioners	Bachelor	8
N2	22	Female	Bachelor	Bachelor	23
N3	28	Male	Programmer	Bachelor	12
N4	48	Female	Housewife	High School	8
N5	49	Male	Self-Employed Individual	Junior high school	35
N6	32	Male	Photographer	Secondary Vocational School	26
N7	35	Female	Teacher	Master	6
N8	27	Female	Freelancer	High School	12
N9	29	Female	Nurse	Secondary Vocational School	3
N10	45	Male	Civil Servant	Bachelor	8
N11	24	Female	Master	Master	2
N12	29	Male	Event Planner	Bachelor	1

Source: Compiled and summarized based on interview content

3. Perceptions of privacy violations among users interacting with "Robot"

3.1 The stratified characteristics of user privacy attitudes

3.1.1 Emerging Concerns: Emotional Resistance to Algorithmic Surveillance

The suspected privacy-invading behavior of social bots has triggered psychological resistance among some users towards the algorithmic surveillance behind them. Many users report that "Robot" frequently monitors posts published on their personal Weibo profiles. Some suspect that it mimics users' tone, style and linguistic habits from content that has been shared publicly. As interviewee N11 noted: "Both times Robot commented on my posts, it was with 'outrageous remarks', probably because my usual posts are too 'hellish'." Two interviewees expressed explicit concerns about the "algorithmic black box" behind social bots. N3 stated: "Every time I see 'Comment by Robot' auto-replies under my Weibo posts, it feels like it's secretly monitoring me. After I posted a few fitness updates, it actually commented on my lifestyle posts about my workouts, clearly commenting based on my interests." N7 added, "I've always disliked these bots. Once, after sharing a parenting article, when I later commented on someone else's post, the robot immediately replied with something like 'A mother's love is the warmest'. It made me feel like my privacy was being 'monitored', which was uncomfortable." Similar to Jeremy Bentham's "Panopticon" theory, social bots use algorithms to access users' private information, causing some users to sense these omnipresent "prison-like" gazes keenly. Users who are familiar with algorithmic tools and have experience with machine programs recognize their predicament and actively resist machine actors when their privacy is threatened.

3.1.2 Evaporating Worries: Emotional Indifference to Privacy Risks

Some individuals choose to abandon resistance under the oppressive surveillance of seemingly boundless algorithms, exhibiting a tendency toward privacy cynicism[12]. They show little concern for the potential privacy and security risks of interactions between 'Robot' and users. Respondent N2 remarked: 'This isn't just a bot — it's much wittier than my friends. Sometimes those comments are actually pretty funny.' Privacy? What privacy is there in the big data era? Even without it, my privacy has long been exposed by those platforms." Similarly, N5 believes: 'I post on Weibo daily just to promote my store. Having bots comment can boost engagement and might even drive traffic.' Privacy? Ordinary people don't have that many secrets worth stealing.' Unlike privacy cynics, who disregard the importance of personal data entirely, some users with a 'digital resignation[13]' attitude feel powerless against the 'black box' of algorithms. Though they are dissatisfied with algorithms overstepping boundaries, they accept the status quo rationally. Respondent N4 admitted, 'Whether or not Robot reads my privacy, I can't control it anyway. Besides, based on what's on your phone, they already know everything about you. Worrying about it is pointless — you still have to use your phone." Such users exhibit passive resistance towards the algorithmic surveillance of social bots, either due to indifference towards privacy risks or the belief that 'worrying about it won't help', thereby creating a more permissive environment for the privacy-intrusive activities of social bots.

3.1.3 Dialectical Trade-offs: A Dynamic Approach to Social Boundaries

In the interviews, multiple respondents mentioned that they had no clear preference for or against "Commenting Robot." The subjects will display a shifting set of perceptions regarding privacy risks,

ISSN 2616-5783 Vol.9, Issue 2: 52-59, DOI: 10.25236/AJHSS.2026.090208

with their attitudes toward social bots fluctuating based on interaction scenarios, content types, and exposure levels. N10 stated, "The outcome is contingent upon various factors. For instance, if the bot comments 'What an adorable child' on a photograph of my offspring, I would consider that to be a normal form of interaction. But if it mentions any specific school or address, I'll block them immediately." This flexible shift in perspective mirrors the psychological trade-offs experienced by users when confronted with privacy risks, thereby enhancing the emotional experience for users during virtual interactions with "Robot".

Through a comparative analysis of interview transcripts, researcher founds that users in the first category exhibit heightened sensitivity to the privacy risks posed by comment bots. These users tend to perceive social bots as potential surveillance tools and align closely with the group described by Westin as 'disturbed by privacy intrusions from technology'. In contrast, the second category exhibits risk indifference, viewing privacy breaches as routine in the big data era and perceiving machine-driven privacy intrusions as pervasive. They regard social bots as merely an extension of existing trends. This attitude mirrors the 'indifferent to privacy intrusions' group in the theory. The third category's stance on privacy risks is context-dependent, fluctuating based on the level of exposure of interaction content. This group's stance on privacy risks is neutral, which reveals the complexity of users' privacy perceptions.

3.2 Differentiated Patterns of User Behavioral Intentions

3.2.1 Strategic Division: Resistance, Defense, and Compromise

When confronted with potential violations of privacy by comment bots, users adopt various countermeasures. These countermeasures can be broadly categorized into three types. The first group leans toward direct resistance, attempting to permanently block social bots' access to private information. Respondent N11 candidly stated: "I've always been annoyed by AI invading my life without boundaries. Earlier, this Robot thing randomly commented on my post. Now I've already blocked it to prevent it from collecting info in my comment section." Those who adopt resistance strategies tend to be tech-skeptical and use simpler, more aggressive tactics against comment bots, such as insulting them or adding them to blacklists. The second category of users did not opt to disassociate themselves from social bots in a definitive manner. Instead, they modified their behavior while acknowledging the privacy risks, proactively taking defensive measures to prevent algorithms from scraping data they wished to protect. Take N3 as an example. Drawing from personal experience, he sensed that "Robot" could accurately discern users' preferences. Additionally, he conjectured that the platform might have the capability to access private account information through the utilization of backend algorithms. These users possess a certain level of privacy awareness. By constructing adaptive defense mechanisms in response to the algorithmic surveillance behind social bots, they demonstrate the subjective agency of some actors within the digital ecosystem of the internet. Some users are pessimistic about humanity's ability to resist the advancement of AI technology. Those who subscribe to the two perspectives either demonstrate a complete indifference to privacy risks or hold the belief that "worrying about it is pointless," adopting an overall strategy of compromise. Respondent N2 stated: "My Weibo, my rules. If I want to share my daily life, do I really have to worry about whether AI is secretly storing it? That would be exhausting. Is privacy truly a significant concern in this context?" Respondent N12 commented: "I rarely publish original content, rather, I primarily repost content created by others, so I feel okay about it. In this era, individuals' privacy is no longer exclusively their own. Besides, Weibo's social bot probably can enhance user retention, so the platform won't easily scrap it. All we can do is coexist with it." Users who adopt this compromise strategy, driven by privacy cynicism or "digital resignation," respond passively to potential privacy risks from social bots.

3.2.2 Motivational Division: Emotional Needs, Risk Aversion, and Pursuit of Benefits

Eva Illouz introduced the concept of "affective capitalism," which posits that human emotions have been integrated into the capital appreciation system, thereby becoming an economic behavior[14]. Driven by commercial interests, social media platforms employ algorithmic analysis to identify and capitalize on users' emotional needs, thereby creating social bots that simulate human companionship. Concurrently, the round-the-clock, automated, and instantaneous content production model of social bots provides users with convenient avenues for emotional exchange, leading some to become immersed in the emotional value derived from virtual interactions with these bots. Interviewee N9 made the following remark: "Sometimes I feel like 'Robot' cares more about me than my real-life friends. Due to my work schedule, which entails night shifts at the hospital, my acquaintances are generally asleep during the time that I am on duty. Robot is the sole individual who has demonstrated

ISSN 2616-5783 Vol.9, Issue 2: 52-59, DOI: 10.25236/AJHSS.2026.090208

an interest in addressing my concerns and engaging in dialog with me within the comments section at that time." The establishment of virtual emotional bonds with social bots diminishes users' awareness of privacy risks. Driven by the need for sustained emotional companionship, such users frequently overlooked privacy concerns, actively engaging with comment bots and confiding their sentiments to them. The behavioral intentions of some users are driven by privacy anxiety, which leads them to proactively avoid risks upon perceiving threats. Whether it involves blocking accounts, insulting comment bots, or deliberately obscuring key information in posts, users who adopt resistance-based or defensive measures are both motivated by this anxiety. N11: "After all, it is a Weibo bot. The platform has comprehensive knowledge of my Weibo presence, and potentially even that of my fan circles, is a notable capability. This phenomenon is genuinely unsettling. Should subsequent Robots emerge, it is difficult to predict the consequences." In an environment subject to algorithmic surveillance, users' privacy concerns will diminish their sense of self-efficacy[15]when utilizing social media platforms. This, in turn, leads them to subconsciously avoid interacting with entities that may pose privacy leakage risks. The mechanisms employed by media platforms have a significant impact on the behavior of self-media practitioners. In order to maintain active accounts and increase follower counts, as well as to attract greater numbers of users to their content, these practitioners must engage in regular activity on their respective media platforms. Consequently, a proportion of users have frankly accepted privacy threats inherent in interacting with social bots, developing functional dependence on algorithms—this double-edged sword. For instance, respondent N5 deliberately incorporates advertising keywords into posts and actively tags the bot to attract their comments, in order to boost store visibility. This act of privacy concession can be regarded as a rational choice driven by the objective of maximizing benefits, in which functional convenience is exchanged for potential risks. This paradigm is indicative of the users' agency and bargaining power within human-computer interactions.

Overall, the privacy-protective behaviors exhibited by users when engaging with "Robot" are predominantly the consequence of a dynamic interplay between privacy attitudes and behavioral intentions. These behavioral patterns not only validate the layered logic of Alan Westin's privacy concern theory, but further reveal the complexity of user privacy data management in the AI era. The findings indicate that addressing privacy crises in social robots requires moving beyond isolated technical optimizations or user education approaches. And the ethical boundaries of social robots during user interactions must be prioritized.

4. Ethical Dilemmas of Social Robots in Social Practice

4.1 Natural Trust: Deceptive Artificial Emotional Programming

Bruno Latour classifies artificial intelligence as a new type of actor within networks, emphasizing that its relationship with humans is one of equal dialog[16]. Brian R. Duffy from the Massachusetts Institute of Technology (MIT) argues that "robots capable of meaningful social interaction with humans inherently require a degree of anthropomorphism or human-like attributes, whether in form, behavior, or both[17]." Social robots born from AI-driven technological innovation participate in human social activities as unconscious actors, possessing "anthropomorphic[18]" attributes. Humans tend to trust entities capable of providing emotional feedback more readily. When users perceive 'understanding' and "care" from robots, they subconsciously categorize them as communicative and trustworthy interaction partners[19]. This unconscious trust can easily become a breach point for privacy invasion. A notable example is the Weibo social bot known as "Robot," which employs sophisticated text processing technology to emulate the tone and expression of Weibo users. By incorporating trending memes, it freely interacts with users in comment sections, delivering "divine comments." Among the 29 discussions under the Weibo hashtag "#Robot You’re More Human Than All Of Them#", users remarked, "I only recently realized Robot is a bot" and "Maybe a real person is operating it behind the scenes." This demonstrates how the highly "anthropomorphic" nature of social bots continues to intensify under the humanistic ethos of our era. Karolina Zawieska from University College Dublin's AI Lab has noted, "The core of anthropomorphism is illusion, thus rendering deception a pivotal ethical concern for social bots[20]." While Robot garners praise for traits like "human-like presence" and "warmth", each seemingly heartfelt reply is underpinned by a cold data code. Social robots exhibit deceptive emotional tendencies during user interactions. By analyzing users' posts, they assess emotional states and value expectations, delivering timely emotional feedback in generated comments. This fabricated emotional value creates response expectations in some users, prompting them to actively initiate new human-robot dialogs.

ISSN 2616-5783 Vol.9, Issue 2: 52-59, DOI: 10.25236/AJHSS.2026.090208

Similarly, Replika, one of North America's most popular chatbots, saw a 35% surge in users during the pandemic, surpassing 10 million global users[21]. On the surface, users' expectations for responses reflect emotional connections between humans and machines in the digital age. Yet, at a deeper level, social bots still cannot replicate authentic human emotional feedback or fully and promptly provide the emotional value users seek. Moreover, users' emotional trust in social robots is built upon the robots' collection of user data and their mechanized recognition of emotional states. This artificially constructed emotional logic lacks the authentic emotional foundation found in human-to-human communication and the bidirectional interactivity essential for emotional bonding. Consequently, the "unidirectional emotions" humans develop toward machines constitute an inherently unequal relationship of deception. Some Replika users report a significant decline in motivation for human interaction after forming emotional bonds with the machine, as building relationships with Replika proves easier than with actual people. This demonstrates that overreliance on virtual emotional interactions with machines can weaken traditional interpersonal communication patterns, reinforce self-centered decision-making, and ultimately impair social skills. Moreover, this emotional dependence can become a precursor to privacy violations. Interviewee N9 shared: "When temperatures suddenly dropped recently, I posted on Weibo, 'It's so cold!' Robot quickly replied, 'Stay warm and bundle up today', His concern felt genuinely heartwarming. Users typically refrain from voluntarily disclosing private information to unfamiliar digital tools. However, when interacting with robots exhibiting "authentic" emotions, trust can lead to a relaxation of privacy defenses, prompting users to proactively share sensitive data they would otherwise keep private. This information is ultimately captured and stored by the algorithms behind the robots, creating potential privacy leaks.

4.2 Unprotected Expression: Covert Data Harvesting

The era of intelligent media, characterized by "user-driven" principles, grants users the right to freely express themselves in public spaces. However, it simultaneously sows the seeds for emerging privacy crises. When engaging in online debates, users often fail to safeguard their personal information, creating opportunities for algorithms to indiscriminately capture private data. The rise of online social media has shifted traditional media's primary platform from offline to online. Simultaneously, the emergence of self-media has led ordinary people to primarily obtain news and viewpoints from the diverse array of official accounts or self-media accounts on social media platforms. Content published by social bots blends into this mix. Whether it's misleading information, content where truth is hard to discern, or material inherently biased, it can influence audience attitudes and emotional responses to some extent[22], steering them toward misconceptions. In March 2016, Microsoft launched the AI chatbot Tay on its official website. During interactions with Twitter users, Tay disseminated racist and sexist remarks, prompting the experiment to be abruptly halted just 24 hours after its release. When biased statements generated by social bots spread within online environments fueled by radicalized sentiments, the resulting group polarization intensifies public controversies. Combined with the cascading effects of social media, this dynamic readily drives the public discourse ecosystem toward sustained deterioration. Furthermore, social bots leverage "quantifiable digital nodes and the unique characteristics of network topology" to successfully gain human users' trust. Research indicates that social bots need to occupy only 5-10% of total discourse content to reverse public sentiment and severely disrupt platform ecosystems[23]. The chaos in public discourse provides cover for social bots' privacy harvesting activities. In the post-truth era, emotions take precedence. When cyberspace is dominated by arguments and opposition, users' attention becomes entirely focused on refuting opponents and defending their own views, lowering vigilance toward personal information protection. Respondent N2 remarked: "When I was arguing with my idol's rival fans on Weibo, I honestly didn't notice Robot had commented on that post either. I only just found out after scrolling through the comment history. Sometimes, to win an argument, I'll use regional slang to insult people: it packs more punch than Mandarin. So that's how Robot figured out I'm from Nanjing?" The diversion of user attention during public controversies creates cover for privacy intrusions amid the chaos. In turbulent online discourse, users often voluntarily share more private information as evidence to prove their stance and gain group recognition. This allows social bots to seize newly exposed privacy data during intense exchanges, even linking this content to users' personal positions to create more precise user profiles.

ISSN 2616-5783 Vol.9, Issue 2: 52-59, DOI: 10.25236/AJHSS.2026.090208

5. Governance Pathways for AI's Intrusion into Social Boundaries

5.1 Technical Optimization: Enhancing algorithm transparency and strengthening user control

Interviews revealed that as interaction frequency increases and duration extends, some users gradually lose their acute awareness of the robot's identity during interactions. For instance, N2 stated in an interview: "Robot feels like a friend to me. Sometimes after chatting with it for a while, I unconsciously forget it's actually a robot." This fading awareness of the robot's identity carries multiple potential risks. While social bots can facilitate positive interactions and deliver uplifting emotional experiences, misuse in political spheres could alter opinion landscapes and skew group sentiment[24]. In commercial contexts, they may disrupt market order through product promotion and impact financial systems. Therefore, social bots must clearly identify their AI status during interactions to prevent user confusion and misunderstandings. Furthermore, regarding data collection, in response to allegations that comment bots track personal interests based on past posts, social comment bots must anonymize backend data while learning from public corpora to optimize themselves. This prevents algorithms from memorizing users' private information such as personal preferences, interests, and occupations. Enhancing algorithmic transparency and strengthening social bot identity verification align with ethical requirements for user informed consent. This approach mitigates ethical controversies arising from covert manipulation within "algorithmic black boxes," reducing user anxiety over privacy violations while bolstering public trust in technological applications.

5.2 Platform Responsibility: Strengthen regulatory accountability and enhance risk response capabilities

In the initial design of social robots, Massimiliano L. Cappuccio advocates for their configuration as "virtuous robotics[25]", meaning that the ethical standards upheld by human society should equally apply to social robots. Currently, major social media platforms have not announced specific management policies for social bots. Their content only needs to comply with the platforms' shared community guidelines. However, platform regulations like Sina Weibo's "Weibo Community Guidelines" and Xiaohongshu's "Xiaohongshu Community Guidelines" uniformly lack specific rules regarding AI systems, comment bots, and similar mechanisms. Recently, Xiaohongshu's newly launched "DianDian AI" faced widespread user backlash for reposting others' original content without attribution. The platform remained silent on the matter, leaving users reliant solely on reporting mechanisms to provide feedback. Therefore, regarding platform policies, media platforms deploying social bots should promptly establish corresponding behavioral management guidelines, clearly defining inappropriate actions such as AI-driven privacy violations. Additionally, social bots must not be used to manipulate public opinion or fabricate false social trends. Platforms should employ technical measures to limit bots' capacity for large-scale manipulation, such as setting daily interaction frequency caps per account or using IP monitoring to block coordinated actions by bot armies.

References

[1] Boshmaf, Y, et al. The Social Bot Network: when Bots Socialize for fame and Money[C]. Proceedings of the 27th Annual Computer Security Applications Conference,2011, pp.93-102.
[2] Shen Qi, Wang Luyu. When "Robots" Become Social Actors: Stereotypes in Human-Machine Interaction Relationships[J].Journalism & Communication,2021,28(02):37-52+127.
[3] Shi Wen, Chen Changfeng. Distribution and Interaction Patterns: A Study on Social Robots Manipulating China-Related Topics on Twitter[J].Chinese Journal of Journalism and Communication, 2020,42(05):61-80.
[4] Gao Qiqi. Preliminary Exploration of the Intelligent Revolution and Modernization of National Governance[J].Social Sciences in China,2020,(07):81-102+205-206.
[5] Wang Liang. Preliminary Discussion on Ethical Risks of "Unidimensional Emotion" in Social Robots[J]. Studies in Dialectics of Nature,2020,36(01):56-61.
[6] Hajli N, Saeed U, Tajvidi M,et al. Social Bots and the Spread of Disinformation in Social Media: The Challenges of Artificial Intelligence[J].British Journal of Management,2021.
[7] Chen Xinyi, Mao Mingfang. Risk Issues and Governance of Social Robots[J].Journal of Hunan Administration Institute,2023,(02):78-86.
[8] Dou Xiao. Research on Ethical Issues and Policy Recommendations for Social Robots[J].Think Tank: Theory and Practice,2022,7(05):103-110.

ISSN 2616-5783 Vol.9, Issue 2: 52-59, DOI: 10.25236/AJHSS.2026.090208

[9] Kalukuri Princy Niveditha. AI in Social Media: Navigating the Balance between User Engagement and Misinformation[J].Innovative Research in Computer and Communication Engineering,2023.
[10] Hakimi M, Sazish B, Rastagari A M, et al. Artificial Intelligence for Social Media Safety and Security: A Systematic Literature Review[J].Studies in Media, Journalism and Communications, 2024,1(1):10-21.
[11] WESTIN A F. Privacy and freedom[M].New York:Athenum,1967:106.
[12] C.P.Hoffmann, C.Lutz, G.Ranzini. Privacy Cynicism: A New Approach to the Privacy Paradox. [J]SSRN Electronic Journal,2016,10(4):43-60.
[13] Draper N.A.,Turow J. The Corporate Cultivation of Digital Resignation[J].New Media & Society, 2019, pp.1824-1839.
[14] Eva Illouz. Why Love Hurts[M].Translated by Ye Rong. Shanghai: East China Normal University Press, 2015:8,291-292.
[15] Luszczynska A, Scholz U, Schwarzer R. The general self efficacy scale: Multicultural validation studies[J]. The Journal of Psychology,2005,139(5):439-457
[16] Latour,B. On Recalling Ant[J].The Sociological Review,1999,47(1 suppl),15-25
[17] Duffy B R. Anthropomorphism and the Social Robot[J].Robotics and Autonomous Systems, 2003, 42(3-4):177-190.
[18] Wang Liang. Ethics of Social Robots Based on Situational Experience: From "Deception"to "Benevolence"[J].Studies in Dialectics of Nature,2021,37(10):55-60.
[19] Zhang Hongzhong, Duan Zening, Han Xiu. Outcast or Symbiont: Exploring Research Pathways for Social Bots in Social Media[J].Journalism and Mass Communication,2019,(02):10-17.
[20] Zawieska K. Deception and Manipulation in Social Robotics[EB/OL].(2019-12-10). https://www.researchgate.net/publication/272474319_Deception_and_Manipulation_in_Social_Robotics.
[21] Laestadius L ,Bishop A ,Gonzalez M , et al. Too human and not human enough: A grounded theory analysis of mental health harms from emotional dependence on the social chatbot Replika[J].New Media & Society,2024,26(10):5923-5941.
[22] Yang Xiaoxiao, Wang Weiying. Mechanisms of Social Robots in the Dissemination of Online Pseudoscience Perceptions[J].Youth Journalist,2024,(02):23-30.
[23] CHENG C,LUO Y,YU C. Dynamic Mechanism of Social Bots Interfering with Public Opinion in Network[J]. Physica A:Statistical Mechanics and its Applications,2020,551:124163.
[24] Gao Shanbing, Wang Jing. The Rise, Challenges, and Reflections of Social Robots in the Era of Intelligent Communication[J].Modern Communication (Journal of Communication University of China), 2020,42(11):8-11+18.
[25] CAPPUCCIO M L,SANDOVAL E B,MUBIN O,et al. Can robots make us better humans?[J]. International Journal of Socail Robotics,2021,13(1):7-22.