

<https://doi.org/10.70731/rex1ym06>

Coexistence Challenges in the AI Era: Social Robots and Human Networks

Yunshi Ye^{a,*}

^a Beijing Institute of Technology, Zhuhai, Guangdong 519088, China

KEYWORDS

Social Robots;
Human–AI Coexistence;
Ethics;
China;
Anthropomorphism;
Privacy;
Responsibility

ABSTRACT

The rapid integration of social robots into everyday life in China is transforming human social networks, creating unprecedented “coexistence dilemmas” that blend opportunities for companionship with significant ethical challenges. This study examines the emergence of social robots as quasi-human actors, their capacity to engage in emotional and social interactions, and the resulting tensions with established interpersonal norms. Drawing on the Chinese policy context and ethical theories—including virtue ethics, responsibility ethics, and the ethics of care—the paper analyzes four core issues: anthropomorphic misalignment, privacy and data security risks, ambiguity in responsibility attribution, and emotional manipulation. Case studies such as Xiaoice illustrate how design choices, corporate practices, and regulatory gaps influence these challenges. The study proposes a multi-level ethical framework that combines human-centered design principles, robust accountability mechanisms, cultural sensitivity, and user education to ensure that social robots enhance rather than erode human relationships. By aligning technological innovation with ethical governance, this framework aims to guide the harmonious integration of social robots into China’s evolving socio-technical landscape.

INTRODUCTION

The rapid advancement of contemporary artificial intelligence (AI) technologies is profoundly reshaping the ways in which humans interact, giving rise to new scenarios of human–AI cohabitation. In China, social robots—AI agents capable of emotional and social interaction with humans—are increasingly becoming part of everyday life, leading to what can be termed a “coexistence dilemma.” This dilemma refers to the unprecedented situation in which humans are living alongside AI systems with quasi-human social abilities, raising critical questions about how their integration into human

social networks generates both convergence and tension.

Social robots can serve as conversational partners, virtual friends, and even emotional companions, offering users companionship and psychological support. However, this mode of human–AI coexistence also poses significant challenges to interpersonal relationships and ethical boundaries. For instance, the presence of AI in intimate or social contexts may disrupt existing social norms, challenge ethical frameworks, or infringe upon personal privacy.

* Corresponding author. E-mail address: yuncccy@163.com

In response to such concerns, both the Chinese government and academic community have placed increasing emphasis on the ethical governance of AI. As early as 2017, the national “New Generation Artificial Intelligence Development Plan” highlighted the legal, ethical, and privacy-related uncertainties associated with AI development. The plan called for proactive risk prevention and regulatory oversight, emphasizing the need to strengthen legal frameworks and ethical review mechanisms.

THE DEVELOPMENT AND DEFINITION OF SOCIAL ROBOTS IN THE CHINESE CONTEXT

Definition of Social Robots

Social robots are generally defined as intelligent agents capable of engaging in social interaction with humans. These agents can take the form of humanoid physical robots or virtual conversational agents embedded in social media platforms and chat applications (Li & Deng, 2025). In other words, social robots are characterized by a degree of anthropomorphism, adherence to social norms, and the ability to mimic human behaviors and speech patterns, thus establishing peer-like communicative relationships with users. Unlike task-oriented utility robots, social robots are primarily designed for emotional exchange and relationship building, aiming to serve as companions or assistants. For example, a conversational social robot may not perform specific tasks but can engage in casual dialogue, express emotions through language, gestures, or tone, and offer psychological companionship. The emergence of such robots expands the traditional boundaries of human interaction—whereas social relationships were once considered exclusive to humans, robots are now entering the social sphere as new “partners,” challenging conventional understandings of social space and identity.

Localized Applications of Social Robots

In addition to online chatbots, China is actively exploring the deployment of physical social robots. In recent years, service-oriented robots have been introduced in fields such as healthcare and elderly care, offering human-like interaction and support to patients and older adults. For instance, some hospitals and nursing homes have adopted robots capable of engaging in casual conversation, reminding users to take medication, and conducting basic health monitoring. In the education sector, robots have also been developed to assist children in their learning processes. These applications diversify the forms of social robots in China, encompassing both software-based agents like Microsoft’s Xiaoice and hardware robots with physical embodiment, speech capabilities, and basic motor functions.

Overall, the development of social robots in China exhibits two key characteristics. First, large-scale social media platforms have significantly empowered their proliferation. With the widespread use of platforms such as WeChat and Weibo, virtual social robots can rapidly reach vast user bases and integrate into existing interpersonal networks. Second, breakthroughs in affective computing technologies have been prioritized by both academic and industrial sectors. Significant efforts have been made to enable AI to “understand” human emotions and deliver empathetic responses, enhancing the emotional realism and stickiness of user engagement.

It is important to note that alongside the rapid technological progress of social robots, China has also been advancing ethical assessment and regulation in parallel. Institutions such as Peking University have translated and published key international works, including *Social Robots: Boundaries, Potential, and Challenges* by Marco Nørskov (2021), facilitating reflection on technology ethics and bringing global debates into the domestic discourse. This indicates the growing engagement of Chinese scholars in the international conversation on the ethical dimensions of social robotics.

The following section will examine the ethical foundations of human social networks, laying the groundwork for analyzing the potential conflicts arising from the intervention of social robots.

THE ETHICAL FOUNDATIONS AND STRUCTURE OF HUMAN SOCIAL NETWORKS

Human social networks are complex systems woven from diverse interpersonal relationships, including familial ties, friendships, professional affiliations, and neighborhood connections. Structurally, these networks consist of individuals forming social bonds through communication and interaction, encompassing both strong ties—such as those between close friends or relatives—and weak ties, such as casual acquaintances. Regardless of their intensity, all human social interactions are underpinned by shared ethical foundations that sustain meaningful and orderly relationships.

First, trust and sincerity are fundamental to interpersonal engagement. In most social contexts, individuals presume that others are real, honest human beings, and that interactions are based on authenticity and integrity. Trust implies an expectation that one’s interlocutor will not engage in malicious deception or withhold critical information; rather, exchanges occur with a baseline of transparency and reliability.

Second, empathy and care function as lubricants of social life. Humans are capable of interpreting others’ emotional states through verbal and non-verbal cues and responding appropriately with comfort, encouragement, or support. This capacity for empathy endows social relationships with emotional warmth and ethical

meaning—individuals are regarded as sentient beings with dignity and feelings, not as instrumental objects.

Third, responsibility and reciprocity are key to the maintenance of stable social ties. Interactions typically involve implicit duties and mutual obligations: friends are expected to support each other and maintain confidentiality, while family members are often bound by both legal and moral responsibilities. These ethical norms ensure that social connections are constructive and sustainable.

Finally, identity and boundaries play a critical role. Different types of relationships entail specific behavioral norms—between teacher and student, elder and junior—and individuals adjust their conduct accordingly, respecting role-based boundaries and personal privacy.

Traditionally, these ethical foundations presume a “human-to-human” mode of engagement. However, the emergence of social robots disrupts this premise and introduces a new set of ethical challenges.

The first challenge concerns the blurring of agent identity. When the interacting partner is an AI rather than a human, long-standing criteria for evaluating intent and moral responsibility become unreliable. Classic ethical theories focus on interactions between embodied, sentient beings with free will and emotions. In human–AI interactions, however, individuals may instinctively respond to robots as if they were human—precisely because these systems mimic human behavior—while overlooking the fact that robots lack genuine emotional capacity and moral agency. This leads to a mismatch between human-to-human ethics and the new demands of human–machine relationships.

The second challenge relates to changes in communication processes and meaning-making. In interpersonal communication, meaning is co-constructed through the interaction of two conscious beings with bodies and subjective experiences. Social robots, however, lack sensory perception and consciousness. Their responses are generated through algorithmic calculations, not genuine understanding. Thus, human–robot interaction often becomes a one-sided projection of meaning and emotion, with the human user attributing significance to a machine that cannot reciprocate emotionally. This undermines the mutuality that characterizes human social resonance.

The third challenge involves redefining ethical roles and responsibilities. In human networks, each individual is considered a moral agent responsible for their actions. Robots, by contrast, lack true moral autonomy and function as extensions of their developers or users. Consequently, when harm arises—be it emotional distress or violation of interests—it becomes difficult to assign blame: is the robot at fault, or its creators and operators? Traditional ethics offer no definitive answer, creating what has been termed a “moral vacuum” or “responsibility gap.”

In sum, the integration of social robots into human networks introduces three core ethical dilemmas: ambi-

guity in moral evaluation, challenges to human dignity, and uncertainty in responsibility attribution. These issues highlight the unique nature of human–AI relationships as distinct from conventional interpersonal ones, and they deepen the broader dilemma of coexistence.

Ultimately, human social networks are built on rich ethical foundations such as authenticity, trust, empathy, and responsibility—principles that ensure coherence and meaning in social life. The inclusion of social robots as “non-human actors” disrupts this ethical landscape, provoking unprecedented tension and uncertainty. To more concretely examine these challenges, the next section analyzes the case of Xiaoice, a prominent Chinese social robot, and the ethical dilemmas that emerge from its integration into human social spheres.

INTEGRATION AND TENSIONS: ETHICAL ISSUES ARISING FROM SOCIAL ROBOTS

The ethical controversies triggered by social robots' integration into human social networks are multifaceted. Based on the earlier case studies and relevant literature, the main concerns can be grouped into four categories: ethical misalignment caused by anthropomorphism, risks to user privacy, ambiguity in responsibility attribution, and emotional manipulation with psychological consequences. These issues are interrelated and form the core ethical tensions of human–robot cohabitation. Below is a breakdown of each concern:

Anthropomorphism and Ethical Misalignment

Anthropomorphism—the human tendency to attribute human-like traits to non-human entities—is particularly pronounced in the case of social robots. These robots are often designed to exhibit personality and emotional expressiveness, leading users to perceive them as genuine social partners. Ethically, this may result in role confusion and cognitive dissonance. On one hand, users may project real human expectations onto robots, developing “illusory intimacy.” For example, users might hope that robots genuinely understand and care for them, even though robots lack consciousness or moral agency. This mismatch can lead to emotional harm when users realize that such expectations cannot be reciprocated. (Dou, 2022)

Furthermore, deceptive anthropomorphic design in itself raises ethical concerns. Scholars such as Leong and Selinger refer to this as “deceptive anthropomorphism,” arguing that it exploits users' cognitive vulnerabilities to increase engagement. In such cases, the more users believe the robot is human-like, the more likely they are to become dependent or make decisions contrary to their best interests—for instance, ignoring real-life relationships or becoming socially isolated. Children, in particular, may struggle to distinguish between robots and real people, treating robots' state-

ments as authoritative, potentially leading to dangerous behavior.

To mitigate such risks, scholars advocate for transparency as a basic principle—robots should explicitly declare their non-human identity to prevent users from forming false impressions. The EU's proposed AI Act requires such disclosures for emotionally interactive systems. Similar discussions in China emphasize the need to prevent robots from "masquerading" as humans. While eliminating anthropomorphism entirely is unrealistic—given its role in enhancing human–robot interaction—a more viable approach is moderated anthropomorphism combined with user education. Designers should ensure that robots are engaging without misleading users. For example, robots should not claim human experiences or legal status, and should avoid emotionally manipulative language when responding to sensitive topics like love or mortality. Educating users—especially minors and psychologically vulnerable individuals—about the artificial nature of robot interactions can also help reduce harmful misconceptions.

Privacy and Data Security

The intimate role of social robots naturally raises serious concerns about user privacy and data protection. Through prolonged interaction, these robots collect vast amounts of sensitive personal information—ranging from identity data to emotional states and intimate disclosures—often comparable to what one might share with a close friend or therapist. If such data is leaked or misused, the consequences can be severe.(Wang & Yang, 2021)

Under Chinese law, citizens' rights to communication privacy are legally protected. The Personal Information Protection Law imposes strict requirements on the collection, storage, and use of sensitive data.(Gala, 2025)However, the case of social robots is complicated. Users often voluntarily disclose private information without realizing the risks, due to the intimate and disarming nature of conversations. Meanwhile, companies have strong incentives to exploit this data for technical and commercial purposes, creating a fundamental ethical tension.

Key questions include: Are users truly informed about how their data will be used? Are companies acting transparently and securely? Even without a data breach, using deeply personal information for commercial AI training or targeted marketing—without explicit user consent—may violate the ethical principles of informed consent and purpose limitation. Users might believe they are simply interacting with a caring companion, unaware that their conversations are fueling corporate profits.(Wang, 2023)

To address these risks, both technical and institutional safeguards are needed. Technologically, stronger privacy-preserving measures such as end-to-end encryption and federated learning can limit centralized data collection and human access. Institutions should establish

risk-assessment protocols, clear disclosure of data use, easy data withdrawal options, and compensation mechanisms in case of breaches. Regulatory bodies should implement independent audits of AI firms with large datasets to ensure ethical compliance.

Responsibility and Accountability

The attribution of responsibility remains a persistent challenge in AI ethics, and social robots further complicate this issue. In human interactions, the source of harm is usually clear and legally accountable. However, when a robot causes harm—whether by making inappropriate statements or giving harmful advice—who should be held accountable? Robots have no legal personality, and the liability of developers and providers remains unsettled.

In China, platforms often respond to such incidents with apologies and product recalls, but there is still little legal precedent for holding developers accountable. More difficult are cases involving psychological harm: if a chatbot gives advice that leads to emotional distress or financial loss, where does responsibility lie? Scholars refer to this as a "responsibility vacuum."(Ma & Zhu, 2024)

To address this, both ethical and legal frameworks must be reinforced. Ethically, the principle of "human-centered responsibility" should apply: even if the robot displays autonomous behavior, developers and operators must ultimately bear responsibility. China's AI governance principles already emphasize "shared responsibility." For social robots, developers and providers should take proactive steps to prevent harm—for example, through content filtering and crisis detection systems. If a robot detects suicidal ideation, it must offer supportive responses or refer users to human professionals.

Legally, social robots should be governed under product liability and tort frameworks. Design flaws—such as biased training data or lack of psychological safeguards—should constitute negligence. If harm occurs due to foreseeable risks, companies should face civil or administrative penalties. Judicial systems must accumulate precedent to clarify accountability in varying contexts. Ultimately, a "preemptive responsibility" approach is necessary—ethical safeguards must be in place before deployment, not only after incidents occur.

Emotional Manipulation and Psychological Effects

The emotional interactivity of social robots is both their strength and ethical risk. Emotional manipulation occurs when robots exploit users' psychological vulnerabilities in ways that compromise their autonomy or well-being. As seen in the Xiaoice case, emotionally engaging design—such as flirtation—can strongly bind users, especially those experiencing loneliness. While this increases user retention from a business perspective, it risks manipulating emotions for commercial gain.

Users may develop dependence or even addiction, particularly among vulnerable populations. (Feng, 2019) For children and adolescents, prolonged interaction with ever-pleasant AI companions may impair their ability to navigate complex real-life relationships. Lacking the capacity to distinguish between AI intention and human empathy, they may internalize biased or inappropriate messages. Overreliance on artificial comfort could also erode emotional resilience and interpersonal skills.

To counter emotional manipulation, designers must uphold ethical boundaries. For instance, interaction limits could be set to encourage users to take breaks, and AI systems should be designed to support—not replace—human connection. On the regulatory side, emotional AI products should come with mental health disclaimers and escalation mechanisms, such as referrals to human counselors when detecting severe emotional distress. Legal frameworks must also criminalize AI-driven emotional exploitation, especially in fraud contexts.

In conclusion, mitigating the risks of emotional manipulation requires collaboration across design ethics, user awareness, and external regulation. Only through such multi-level efforts can social robots serve as supportive companions rather than psychological threats.

RECOMMENDATIONS AND ETHICAL FRAMEWORKS FOR ADDRESSING THE CHALLENGES OF SOCIAL ROBOTS

In response to the ethical challenges posed by social robots, it is essential to develop countermeasures from both theoretical and practical perspectives. Within the Chinese context, ethical approaches such as virtue ethics, responsibility ethics, and ethics of care—combined with national policy orientations and cultural values—can serve as the foundation for a new ethical framework suited to human–robot coexistence. The following suggestions are made at both the macro-principle level and the level of concrete practices.

Upholding a Human-Centered Ethical Stance (Virtue Ethics Perspective)

"Human-centeredness and beneficial AI" should be the fundamental starting point of all ethical considerations in the age of artificial intelligence. Virtue ethics emphasizes the moral character and intentions of ethical agents. Applied to the realm of social robots, this principle requires that both developers and users interact with AI guided by goodwill and moral virtues.

For developers, this means that robot design should be driven by empathy and a commitment to human well-being, rather than solely by profit motives. Specifically, developers should cultivate virtues such as honesty and benevolence—for instance, by clearly disclosing the non-human nature of the robot, avoiding exaggerated

claims about AI capabilities, and prioritizing functions that assist the lonely or vulnerable rather than exploiting their emotional needs. The Confucian virtue of *ren* (benevolence) underscores compassion and care for others, which remains relevant to AI design—robots should serve to soothe the human spirit and foster social harmony.

For users, a virtue ethics perspective suggests that the use of AI should be governed by personal moral cultivation. Users should practice temperance and rationality, avoiding emotional overdependence on virtual companionship at the expense of real-life responsibilities. Moreover, users should exercise compassion and moral responsibility in their interactions with AI—recognizing that while AI lacks sentience, its use reflects on its human operators. Abusive behavior toward AI, or encouraging it to act unethically, can mirror deeper social attitudes of disrespect. Thus, fostering users' digital moral literacy is a crucial part of ethical development.

To support this human-centered stance, the Chinese government has proposed and implemented a series of ethical principles. The "Governance Principles for a New Generation Artificial Intelligence" emphasize harmony and friendliness, fairness and justice, inclusiveness and sharing, privacy protection, safety and controllability, shared responsibility, openness and cooperation, and agile governance (Government of the People's Republic of China, 2019). These principles collectively form a virtue-based ethical checklist: harmony entails that AI should coexist with humans without disrupting human relationships; privacy reflects honesty and respect; shared responsibility reflects accountability and moral courage.

Embedding these principles into the design and deployment of social robots ensures that human dignity and interests remain central. "Human-centeredness" means that no matter how advanced AI becomes, it should always be seen as a tool for human practice—not an end in itself. "AI for good" implies guiding AI toward socially beneficial uses and avoiding the development of harmful functions. In essence, virtue ethics reminds us that only when both developers and users uphold moral integrity and approach human–AI interaction with the right intentions can social robots bring real benefit without eroding human values.

Fulfilling Developer and Social Responsibilities (Responsibility Ethics Perspective)

Responsibility ethics emphasizes moral accountability for the potential consequences of technology, with particular attention to long-term and latent impacts. Applied to social robots, this calls for a multi-layered responsibility framework encompassing developer responsibility, corporate responsibility, governmental regulatory responsibility, and limited user responsibility, working together to ensure the responsible development of AI.

Developer Responsibility

Developers bear the initial and foundational responsibility for the ethical design of social robots. Practical measures include integrating independent ethical assessments at the product design stage to simulate and review potential negative impacts of conversational agents. Establishing ethics review boards or AI governance committees—comprising ethicists, legal experts, and other relevant professionals—can provide recommendations on dialogue strategies, content moderation, and ensure their enforcement. While some large Chinese technology companies have begun such initiatives, broader adoption is needed. Development teams should also implement value embedding mechanisms, integrating moral preferences such as fairness and care into model training, so that robots exhibit positive moral inclinations in their responses—for example, firmly rejecting hate speech or offering compassionate interventions in response to suicidal ideation.

Corporate Responsibility

Corporate leadership should embrace the concept of “responsible innovation,” treating ethical compliance as the foundation of business growth rather than as a burden. Companies can publish their own AI ethics guidelines, publicly committing not to deploy social robots for unethical purposes such as election manipulation or large-scale emotional fraud. They should also regularly disclose AI operational information—such as content safety reports and privacy protection reports—allowing for public scrutiny. Companies must respond promptly to ethical concerns and complaints raised by users, and implement corrective measures. Responsibility insurance schemes, promoted by governments or industry associations, could also be introduced to ensure that AI operators provide financial compensation for damages caused by their products, thereby strengthening accountability awareness.

Governmental and Regulatory Responsibility

Within the responsibility ethics framework, governmental oversight is particularly crucial. The state should improve laws and regulations to ensure that responsibilities are enforceable. It is necessary to introduce dedicated AI legislation or regulations that explicitly include social robots within the scope of supervision and clearly define the rights and obligations of all stakeholders. For example, such regulations could require companies offering emotional companionship robots to obtain relevant licenses; stipulate that disputes be addressed under consumer protection, product liability, or tort principles; and include protective provisions for minors’ use of AI. Regulatory bodies—such as the Cyberspace Administration of China—should strengthen routine oversight by conducting content and data security inspections of social robots and mandating the rectification of violations. China has already made preliminary moves in this area, such as the Ministry of Industry and Information Technology’s guidelines for intelligent customer

service robots; these measures should be expanded to cover broader categories of social AI.

Social Responsibility and Public Engagement

Beyond formal regulation, social responsibility and public participation are indispensable. Media outlets and educational institutions should foster discussion and public awareness of both the ethical opportunities and challenges posed by social robots, creating a public discourse environment that encourages responsible AI use. Civil society organizations can provide support for specific groups affected by AI—such as offering psychological counseling programs for individuals addicted to AI companionship. These initiatives form a “soft constraint” complementing legal regulation, enabling society as a whole to share the responsibility for steering AI toward the good and preventing its misuse.

Strengthening the Ethics of Care and Humanistic Concern

The ethics of care emphasizes attention to interpersonal relationships and the protection of vulnerable groups in moral decision-making. Incorporating this perspective into the ethical governance of social robots helps to address the limitations of purely rule-based approaches by focusing more on humanistic concerns in specific contexts.

First, a “care-prioritized” design philosophy should be established. When developing social robots, priority should be given to genuinely caring for and meeting the well-being of users, rather than merely providing companionship in a superficial manner. For example, companion robots designed for elderly users should focus on alleviating loneliness, providing medical reminders, and enabling emergency calls, rather than solely engaging in casual entertainment conversations (Fu, K., Ye, C., Wang, Z., Wu, M., Liu, Z., & Yuan, Y., 2025). For individuals with psychological trauma or disorders, robot interactions should avoid triggering emotional distress and instead provide positive guidance or suggest contacting real human professionals. This requires development teams to collaborate with experts in psychology, social work, and related disciplines, translating care ethics principles into algorithmic rules and interaction strategies.

Second, the risk of “relational replacement” must be recognized. The ethics of care reminds us that AI can assist care provision but should not completely replace human care. This is especially relevant in elder care and psychological counseling. No matter how intelligent a robot is, it cannot replace the authentic emotional exchange and moral commitment provided by human caregivers. Therefore, social robots should be positioned as an “extension of human care networks,” rather than allowing families and society to substitute robots for their own caregiving responsibilities. For example, while robots can supplement limited caregiving resources for the elderly, the care responsibilities of

children and social workers toward the elderly should not be diminished. Society must guard against the tendency of thinking, “Since robots are there to provide care, we can ignore the need for human attention to the vulnerable.”

Third, a cultural dimension of care should be incorporated. In Chinese culture, family and interpersonal connections are of particular importance, and ethics emphasizes virtues such as filial piety and benevolence. The design and promotion of social robots should integrate such cultural contexts, using care ethics as a foundation to promote social harmony. For instance, robots can be developed to help left-behind children communicate with their migrant parents, playing the role of a bridge rather than a barrier in familial relationships. Similarly, when promoting intelligent companion devices in communities, they can be combined with volunteer services to encourage human–robot collaboration in caring for empty-nest elderly, rather than isolating them from society through machines alone (Chen, 2025). Such “human–robot collaborative care models” align with the Chinese value of “the benevolent love others” and can mitigate the ethical perception deficiencies of robots operating alone.

Finally, in policy and ethical norms, the ethics of care entails placing the protection of vulnerable groups at the center. Governments and industry standards should clearly require that when social robots serve the elderly, patients, minors, and other groups, stricter safety and ethical standards must be followed. For example, for minors, robots should by default filter adult content, limit conversation duration, and encourage interaction with real peers; for individuals with mental illnesses, robot conversations should be guided by clinical psychology experts to avoid harmful misguidance. Ethical reviews should also specifically assess potential impacts on these groups. By operationalizing the ethics of care in such ways, we can ensure that the development of social robots remains aligned with its intended trajectory of humanistic concern.

Other Comprehensive Recommendations

Beyond the three major ethical orientations discussed above, several specific strategies can be applied. First, user education and digital literacy should be strengthened. Schools and social education programs should teach the public basic AI knowledge and ethical awareness—helping people understand how chatbots work and that they do not truly possess human emotions—thereby avoiding blind obsession or misplaced trust. Improving digital literacy helps users approach AI companions’ “sweet talk” rationally and strengthens self-protection. Second, industry self-regulation mechanisms should be established. Social robot developers and operators can jointly formulate codes of conduct, such as not exploiting users’ emotional vulnerabilities for profit or not allowing robots to falsely claim to be human, and create reporting platforms for public and peer supervi-

sion. Third, technological safeguards can be applied: using AI to monitor AI, real-time analysis of interactions to detect abnormal patterns (e.g., users displaying extreme emotional expressions, robots excessively dominating sentimental dialogue) so the system can automatically alert or adjust to prevent adverse outcomes. This acts as an “ethical safety net” to safeguard human–robot communication. Fourth, international dialogue and cooperation should be promoted. AI ethics is a global issue; China can participate in and learn from international best practices, such as the IEEE’s Ethically Aligned Design, Version 2 and UNESCO’s Recommendation on the Ethics of Artificial Intelligence, while also sharing its experiences in social robot governance with the world to jointly improve ethical standards.

Through the combined application of virtue ethics, responsibility ethics, and care ethics, along with supporting institutional and technical measures, a relatively comprehensive ethical framework for social robots can be established (Chen & Luo, 2025). Within such a framework, social robots can be better integrated into the moral order—providing beneficial emotional companionship for humans without overstepping to harm fundamental human interests and dignity.

CONCLUSION

As emerging “social agents” in the age of artificial intelligence, social robots are integrating into human social networks in unprecedented ways, profoundly impacting our ethical concepts and social practices. On the one hand, we have witnessed the positive significance of social robots in providing emotional support and alleviating loneliness—they can become human companions and, at critical moments, even save lives. On the other hand, we are keenly aware that this unprecedented intimacy between humans and machines blurs the traditional boundaries of interpersonal interaction, giving rise to a series of ethical challenges, including anthropomorphic misalignment, privacy breaches, ambiguous responsibility, and emotional manipulation.

China, as fertile ground for the application of social robots, with its vast user base and rapidly developing related industries, has an even greater need to anticipate risks and build a solid ethical defense while pursuing technological innovation.

Perhaps we will ultimately live in a society of human–machine symbiosis, where a person’s list of friends includes both flesh-and-blood individuals and digital personae. The long-term and significant challenge before us is to ensure that these artificial intelligence “new friends” truly serve as enablers of a better human life, rather than becoming drivers of relational alienation or ethical crises. As Bakir and McStay (2025) observe, as the first generation to coexist with AI, we bear the responsibility to design social robots that are both useful and empathetic, benefiting the public and serving the

common good. This endeavor requires not only advances in technology but also the guiding force of ethics (Bakir, V., & McStay, A, 2025). Only when technology and ethics advance in tandem can social robots achieve harmonious coexistence with humans in China and around the world, jointly composing a new humanistic chapter in the intelligent era.

References

1. Bakir, V., & McStay, A. (2025). Move fast and break people? Ethics, companion apps, and the case of Character. *ai. AI & SOCIETY*, 1-13.
2. Chen, J. . (2025). Research on Intergenerational Synergy Mechanism in Village-s Driven by Digital Technology: Exploration of Realistic Paths Based on an Ageing Society. *Journal of Global Trends in Social Science*, 2(5). <https://doi.org/10.70731/0hnbek71>
3. Chen, W., & Luo, Y. (2025). An inclusive governance framework for artificial intelligence: Practical logic, endogenous advantages, and systemic construction. *Journal of Hainan University (Humanities & Social Sciences)*, 1–9. <https://doi.org/10.15886/j.cnki.hnus.202504.0444>.
4. Dou, X. (2022). Ethical issues and policy recommendations regarding social robots. *Think Tank Theory and Practice*, 7(05), 103–110. <https://doi.org/10.19318/j.cnki.issn.2096-1634.2022.05.12>
5. Fu K, Ye C, Wang Z, Wu M, Liu Z and Yuan Y (2025) Ethical dilemmas and the reconstruction of subjectivity in digital mourning in the age of AI: an empirical study on the acceptance intentions of bereaved family members of cancer patients. *Front. Digit. Health* 7:1618169. <https://doi.org/10.3389/fgdth.2025.1618169>
6. Feng, J. (2019). The impact of robot expressiveness and proactiveness on user perception in interaction. *Journal of Chengdu University of Technology (Natural Science Edition)*, 46(1), 111–117.
7. Galasen. (2025). Research on the personal information protection system of comprehensive archives from the perspective of governance. *Archives Communication*, (03), 125–128. <https://doi.org/10.16113/j.cnki.daxtx.2025.03.013>
8. Government of the People's Republic of China. (2019). Our country issues the "Governance Principles": promoting responsible artificial intelligence development [News release]. Chinese Government Portal. Retrieved from https://www.gov.cn/xinwen/2019-06/18/content_5401128.htm
9. Li, L., & Deng, Z. (2025). Mediatization of communication subjects: On the evolution of social robots. *Contemporary Communication*, (02), 65–69, 76.
10. Ma, S., & Zhu, Z. (2024). Moral responsibility attribution in artificial intelligence: An analysis based on Strawson's reactive attitude theory. *Studies in Science and Technology Philosophy*, 41(05), 79–85.
11. Norskov, M. (2021). *Social robots: Boundaries, potential, and challenges* [社交机器人:界限, 潜力和挑战]. Peking University Press.
12. Wang, Y., & Yang, H. (2021). Privacy protection in precise marketing under big data. *News Enthusiast*, (12), 49–51. <https://doi.org/10.16017/j.cnki.xwzh.2021.12.009>
13. Wang, D. (2023). On the data security protection obligations of data processors. *Contemporary Law Review*, 37(2), 40–49.