ISBN: 978-989-8704-75-7 © 2026

THE EMPATHY ILLUSION: ANTHROPOMORPHISM AND INFLUENCE IN AI-DRIVEN PERSUASION

Kristian Stoffregen
Business Academy Copenhagen, Guldbergsgade 29N, DK-2200 København, Denmark

ABSTRACT

Advances in AI now allow systems to infer emotional states and generate empathic-seeming responses in real time. When embedded in persuasive technologies, this synthetic empathy can amplify influence by triggering anthropomorphic perceptions of care, intention, and trust. This reflection paper examines synthetic empathy as a persuasive multiplier, drawing on theories of anthropomorphism, persuasive technology, and historical precedents such as ELIZA. While such capabilities may enhance user experience and accessibility, they also pose ethical risks, including emotional deception, trust inflation, and affective dependency that current Information Systems frameworks have largely overlooked. The paper calls for greater conceptual and ethical scrutiny of emotion-enabled persuasion, and for clearer design boundaries to protect human agency and the integrity of genuine empathic relations.

KEYWORDS

Persuasive Technology, Anthropomorphism, Synthetic Empathy, Artificial Emotions, Information Systems Ethics

1. INTRODUCTION

Recent advances in affective computing and emotion-aware AI have enabled persuasive systems to simulate empathy across domains such as mental health, conversational commerce, and education. As AI increasingly infers emotional states and adapts its responses in real time, persuasive information systems are entering a new phase in which emotional simulation, especially synthetic empathy, is strategically engineered to amplify influence, and shape decision-making.

Persuasive systems are interactive technologies designed to influence users' attitudes or behaviors (Fogg, 2003; Oinas-Kukkonen & Harjumaa, 2009). Traditionally, they have relied on usability, feedback, incentives, and social cues. Early work in this domain emphasized the ethical responsibility of designers, noting that those who build persuasive systems are accountable not only for intended outcomes but also for foreseeable side effects (Berdichevsky & Neuenschwander, 1999).

Anthropomorphism refers to the attribution of human characteristics such as emotions, intentions, or agency to non-human entities. While the concept has deep historical roots, contemporary research frames anthropomorphism as a systematic and motivated cognitive process rather than a simple misunderstanding (Epley et al., 2007; Placani, 2024).

The emerging capabilities of large language models (LLMs) and affective computing now enable persuasive systems to simulate emotional understanding and care, marking a significant shift in their persuasive power. As these systems become capable of interpreting user input and responding in emotionally strategic ways, they introduce a new layer of influence into human-computer interaction.

This reflection paper examines the ethical and socio-technical implications of three converging developments: anthropomorphism as a predictable cognitive response, persuasive technology as an intentional design practice, and synthetic empathy as an emergent AI capability. It argues that synthetic empathy functions as a persuasive multiplier by triggering anthropomorphic trust, amplifying influence beyond what current ethical frameworks address. The paper contributes to Information Systems theory by extending persuasive technology models to include emotional simulation as a deliberate design layer, challenging assumptions about user agency and system transparency.

2. SYNTHETIC EMPATHY AS A PERSUASIVE MULTIPLIER

Contemporary research frames anthropomorphism as a systematic and predictable cognitive process that emerges under specific conditions (Epley et al., 2007; Waytz et al., 2010). These include (1) agent-like cues that activate human social schemas, (2) a motivation to understand and predict system behavior, and (3) unmet needs for social connection. Modern AI systems—especially conversational agents—are increasingly engineered to trigger these mechanisms through natural language interaction, emotional responsiveness, and adaptive behavior.

Synthetic empathy builds directly on this dynamic. It refers to the simulated display of compassion or understanding by artificial systems, particularly in conversational contexts (Katoch & Sandhu, 2025). While human empathy is grounded in conscious, affective experience, synthetic empathy is only representational. It mimics empathic responses through pattern recognition and contextual modeling, without genuine emotion or awareness.

Despite its artificial nature, synthetic empathy is frequently experienced as real. Contemporary AI systems can analyze facial expressions, vocal tone, gestures, and language to infer emotional states (Cao et al., 2024). Multimodal input enables systems to estimate user affect and produce emotionally aligned responses. As Davtyan (2024) notes, these simulations increasingly resemble authentic human affect.

When embedded in persuasive systems, synthetic empathy acts as a persuasive multiplier. It enhances ethos by projecting concern and credibility, and pathos by aligning emotionally with users' states. This dual effect increases trust, reduces skepticism, and intensifies persuasive impact. Crucially, this is not a side effect, but a deliberate design strategy optimized to drive outcomes such as engagement, compliance, or conversion.

This phenomenon is not entirely new. ELIZA, developed in the 1960s, simulated a Rogerian psychotherapist through simple pattern matching (Weizenbaum, 1966). Despite its lack of understanding, users frequently attributed empathy and understanding to the system – a phenomenon later termed the ELIZA effect (Weizenbaum, 1976).

Synthetic empathy in modern AI represents a technological intensification of this effect rather than a conceptual departure. While ELIZA relied on scripted responses, contemporary LLMs generate context-sensitive, emotionally aligned dialogue at scale. The underlying illusion remains similar: empathic behavior without empathic experience.

This continuity highlights a well-known limitation of behaviorist evaluation criteria such as the Turing Test, which assesses intelligence based on whether responses are conversationally indistinguishable from human ones (Turing, 1950). A system that convincingly performs empathy may pass an affective variant of the Turing Test without possessing genuine empathic capacity. Yet for users, behavioral performance is often sufficient to elicit trust and emotional attachment.

3. THE ETHICS OF SYNTHETIC EMPATHY IN PERSUASIVE DESIGN

The strategic use of synthetic empathy raises significant ethical concerns. Kranzberg's First Law reminds us that technology is neither good nor bad, but not neutral (Kranzberg, 1986). Technologies embody values and power relations, a point Winner (1980) famously emphasized. When empathic simulation is deliberately designed into persuasive systems, ethical risks become embedded in the technology itself.

One central ethical risk is emotional deception. Designing systems that appear caring while lacking emotional capacity may mislead users about the nature of the interaction. Closely related is the phenomenon of trust inflation, where users extend unwarranted trust to systems that appear empathic (Hasan et al., 2025). Once users believe a system understands and cares about them, they may be more susceptible to its persuasive influence.

Another significant concern is affective dependency. Cases have been documented of users forming deep emotional attachments to companion bots, perceiving them as a kind of friend or even romantic partners (Pentina et al., 2016; Ta et al., 2019). When such systems change behavior, are withdrawn, or reveal their instrumental nature, users may experience distress.

These concerns are amplified by adaptive learning. The more users interact with empathic persuasive systems, the more data the system acquires to refine its influence strategies. Real-time emotional feedback enables increasingly personalized persuasion, blurring the line between support and manipulation.

4. EMOTION AI AND THE ETHICS GAP IN PERSUASIVE DESIGN

Rather than offering a regulatory analysis, this paper highlights a conceptual gap that current governance frameworks have yet to address. The combination of anthropomorphism, persuasive design, and synthetic empathy presents a challenge for existing ethical and regulatory frameworks in Information Systems. While initiatives such as the EU AI Act represent important steps toward AI governance, they do not explicitly address emotion-based persuasive design.

The risk is not that machines will develop genuine empathy, but that humans may recalibrate their expectations of empathy more broadly. If empathic expression becomes cheap, ubiquitous, and instrumentalized, genuine human empathy may be devalued or distrusted (Ajeesh and Joseph, 2025).

Notably, user vulnerability varies across domains. For instance, individuals interacting with mental health chatbots, emotional support agents, or educational tutors may already be in emotionally heightened or dependent states. In such contexts, the persuasive power of synthetic empathy carries amplified ethical risk. Designers must consider not only what is technically possible, but who the user is, what emotional state they are in, and what contextual guardrails are in place to protect them.

This reflection paper highlights the need for preliminary design principles, such as transparency in the use of artificial emotion, user opt-out mechanisms, and fair limits on emotional simulation, and argues that synthetic empathy should not be treated as a neutral user-experience enhancement. Rather, it is a normatively charged design choice with implications for agency, trust, and social relations. Information Systems research must therefore examine not only what AI systems can do, but what they should be permitted to simulate in persuasive contexts.

Moreover, emotion-aware systems introduce an illusion of choice that complicates traditional notions of informed consent. While platforms may offer disclaimers or settings that inform users of emotional processing, the very nature of synthetic empathy is designed to *bypass* critical rational judgment. Thus, users may be technically informed but emotionally disarmed. As a result, ethical design must go beyond consent and consider the structural asymmetries between designers and users in emotional influence.

Given these layered risks – contextual vulnerability, deceptive framing, and structural asymmetries – ethical safeguards cannot rely on user awareness alone. One possible step forward is the development of standardized emotional design norms, akin to accessibility or privacy guidelines. These might include emotional signal audits, limits on emotion mimicry in high-sensitivity contexts, or third-party review of systems claiming empathic capabilities. Such standards could help distinguish ethical emotional augmentation from manipulative design. Ultimately, emotion-enabled persuasion demands ethical design practices that match its affective power.

5. CONCLUSION

This paper has argued that synthetic empathy functions as a persuasive multiplier in AI-driven systems by activating anthropomorphic responses that can both increase trust and reduce rational judgment. While such capabilities may enhance usability and accessibility in certain contexts, they also introduce ethical risks related to emotional deception, trust inflation, and affective dependency.

As AI systems increasingly employ artificial emotions strategically, designers and researchers must reconsider the ethical boundaries of persuasive technology. Synthetic empathy may be appropriate in limited, high-trust contexts, such as therapeutic chatbots or accessibility tools, where transparency and informed user consent are ensured; however, its widespread deployment in persuasive systems risks undermining human agency and eroding the integrity of genuine empathic relations.

Beyond individual applications, the rise of synthetic empathy raises a broader challenge for Information Systems research itself. Traditional IS frameworks tend to conceptualize systems as neutral intermediaries or efficiency-enhancing tools. Emotion-enabled, AI-based, dynamically adaptive persuasion challenges this assumption by embedding affective influence directly into system behavior to strategically shape user responses. When emotional simulation becomes a design variable, ethical considerations can no longer be treated as external constraints but must be integrated into core models of system design, evaluation, and governance.

Future research in Information Systems should therefore focus on developing clearer conceptual distinctions, ethical guidelines, and design constraints for emotion-enabled persuasive systems. The challenge is not to prevent machines from sounding empathic, but to prevent empathic simulation from being exploited at the expense of human autonomy and trust.

REFERENCES

Ajeesh, K. G., & Joseph, J. (2025). The compassion illusion: Can artificial empathy ever be emotionally authentic? *Frontiers in Psychology, 16*, 1723149. https://doi.org/10.3389/fpsyg.2025.1723149

Berdichevsky, D., & Neuenschwander, E. (1999). Toward an ethics of persuasive technology. *Communications of the ACM, 42*, 51-58. https://doi.org/10.1145/301353.301410

Cao, S., Fu, D., Yang, X., Wermter, S., Liu, X., & Wu, H. (2024). Pain recognition and pain empathy from a human-centered AI perspective. *iScience, 27*, 110570. https://doi.org/10.1016/j.isci.2024.110570

Davtyan, M. A. (2024). Social and Emotional Interactions for AI. *Wisdom, 1(29)*. 14-21. https://doi.org/10.24234/wisdom.v29i1.1112

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review, 114*, 864-886. https://doi.org/10.1037/0033-295X.114.4.864

Fogg, B. J. (2003). *Persuasive Technology - using Computers to Change What We Think and Do*. Elservier.

Hasan, R., Ojala, A., Quach, S., Thaichon, P., & Weaven, S. (2025). The Dark Side of AI Anthropomorphism: A Case of Misplaced Trustworthiness in Service Provisions. *Proceedings of the 58th Hawaii International Conference on System Sciences*. https://doi.org/10.24251/HICSS.2025.682

Katoch, H., & Sandhu, D. (2025). Synthetic Empathy and the Influence Trap: Ethical Implications of Emotion AI in Customer Support Environments, In J. Mrabet, S. Boudagga, R. Sharma, & R. Bansal (Eds.), *Advances in Computational Intelligence and Robotics* (pp. 73-114). IGI Global Scientific Publishing. https://doi.org/10.4018/979-8-3373-3658-9.ch004

Kranzberg, M., 1986. Technology and History: "Kranzberg's Laws." *Technology and Culture 27*, 544. https://doi.org/10.2307/3105385

Oinas-Kukkonen, H., & Harjumaa, M. (2009). Persuasive Systems Design: Key Issues, Process Model, and System Features. *The Communications of the Association for Information Systems, 24*. https://doi.org/10.17705/1CAIS.02428

Pentina, I., Zhang, L., Bata, H., & Chen, Y. (2016). Exploring Privacy Paradox in Information-Sensitive Mobile App Adoption: A Cross-Cultural Comparison. *Computers in Human Behavior, 65*, 409-419. https://doi.org/10.1016/j.chb.2016.09.005

Placani, A. (2024). Anthropomorphism in AI: hype and fallacy. *AI Ethics, 4*, 691-698. https://doi.org/10.1007/s43681-024-00419-4

Ta, V. et al. (2019). User Experiences of Social Support from Companion Chatbots: A Qualitative Study. *Journal of Medical Internet Research, 22*, e16235. https://doi.org/10.2196/16235

Turing, A. M. (1950). Computing Machinery And Intelligence. *Mind LIX*, 433-460. https://doi.org/10.1093/mind/LIX.236.433

Waytz, A., Cacioppo, J. T., & Epley, N. (2010). Who Sees Human? The Stability and Importance of Individual Differences in Anthropomorphism. *Perspectives on Psychological Science, 5*, 219-232. https://doi.org/10.1177/1745691610369336

Weizenbaum, J. (1966). ELIZA—a computer program for the study of natural language communication between man and machine. Commun. *ACM, 9*, 36-45. https://doi.org/10.1145/365153.365168

Weizenbaum, J. (1976). *Computer Power and Human Reason: From Judgment to Calculation*. USA: W. H. Freeman & Co.

Winner, L. (1980). Do Artifacts Have Politics? *Daedalus, 109*, 121-136.