

- [1] Spatola, Nicolas, and Olga A. Wudarczyk. Ascribing emotions to robots: Explicit and implicit attribution of emotions and perceived robot anthropomorphism. *Computers in Human Behavior* 124 (2021).
- [2] Chinmulgund, Avanti, et al. Anthropomorphism of AI based chatbots by users during communication. *2023 3rd International Conference on Intelligent Technologies (CONIT)*, IEEE (2023).
- [3] Kolomaznik, Michal, et al. The role of socio-emotional attributes in enhancing human-AI collaboration. *Frontiers in Psychology* 15 (2024).
- [4] Schopenhauer, Arthur. *Arthur Schopenhauer: The World as Will and Presentation: Volume I. Vol. 1*. Routledge (2016, originally published 1819).
- [5] Leibniz, Gottfried Wilhelm. *Theodicy: Essays on the Goodness of God, the Freedom of Man, and the Origin of Evil*. Translated by E.M. Huggard. London: Routledge & Kegan Paul (1951, originally published 1819).
- [6] Nietzsche, Friedrich. *The Joyous Science*. Translated by Walter Kaufmann. New York: Vintage Books (1974, originally published 18882).
- [7] Norem, Julie K. Defensive pessimism, anxiety, and the complexity of evaluating self-regulation. *Social and Personality Psychology Compass* 2.1: 121-134 (2008)
- [8] Kahneman, Daniel. *Thinking, Fast and Slow*. Farrar, Straus and Giroux (2011).
- [9] Bortolotti, L., Antrobus, M., & Sullivan-Bissett, E. *The epistemic innocence of optimistically biased beliefs* (2019).
- [10] Khalid, Iraj, et al. Effects of experiencing the COVID-19 pandemic on optimistically biased belief updating. *eLife* 13 (2025).
- [11] Liu, Yang, et al. "Do Large Language Models Possess Sensitive to Sentiment? *arXiv preprint arXiv* (2024).
- [12] Picard, Rosalind W. *Affective Computing*. MIT Press (1997).
- [13] Mei, Katelyn, Sonia F., & Aylin C. Bias against 93 stigmatized groups in masked language models and downstream sentiment classification tasks. *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency* (2023).
- [14] Huang, Po-Sen, et al. Reducing sentiment bias in language models via counterfactual evaluation. *arXiv preprint arXiv* (2019).
- [15] Subhangy G. No more Mr. Nice Guy: Say goodbye to the sycophantic ChatGPT. *Business Insider* (2025).
- [16] OpenAI. Sycophancy in GPT-4o: what happened and what we're doing about it (2025).
- [17] Carro, María Victoria. Flattering to Deceive: The Impact of Sycophantic Behavior on User Trust in Large Language Model. *arXiv preprint arXiv* (2024).
- [18] Asha, P., et al. Efficient mental health therapist chatbot assisted by artificial intelligence. *2024 10th International Conference on Advanced Computing and Communication Systems (ICACCS)*. Vol. 1. IEEE (2024).
- [19] Abd-Alrazaq, Alaa A., et al. An overview of the features of chatbots in mental health: A scoping review. *International journal of medical informatics* 132 (2019).
- [20] Hanji, Bhagyashri R., et al. Self-Heal: Conversational Therapy Bot with AI Enhanced Features for Mental Health. *2024 8th International Conference on Computational System and Information Technology for Sustainable Solutions (CSITSS)*. IEEE (2024).
- [21] Kallivalappil, Neave, et al. Empath. ai: a context-aware chatbot for emotional detection and support." *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*. IEEE (2023).
- [22] Norem, Julie K., and Nancy Cantor. "Anticipatory and post hoc cushioning strategies: Optimism and defensive pessimism in "risky" situations." *Cognitive therapy and research* 10, 347-362 (1986).
- [23] Norem, Julie K. "Defensive pessimism, anxiety, and the complexity of evaluating self-regulation." *Social and Personality Psychology Compass* 2.1: 121-134. (2008).
- [24] Sah, Jyoti, et al. "Risk Management of Future of Defi Using Artificial Intelligence as a Tool." *AI-Driven Decentralized Finance and the Future of Finance*. IGI Global, 252-272. (2024).
- [25] Schuett, Jonas. "Three lines of defense against risks from AI." *AI & SOCIETY* 1-15 (2023).
- [26] Nick, Bostrom. *Superintelligence: Paths, dangers, strategies*. (2014).
- [27] Ouyang, Long, et al. Training language models to follow instructions with human feedback. *Advances in neural information processing systems* 3, 27730-27744 (2022).
- [28] Philipov, Daniel, et al. Simulating User Agents for Embodied Conversational-AI. *arXiv preprint arXiv:2410.23535* (2024).