

Folk Theories of Ethical Agency on Reddit Threads: Negotiating Morality in AI Data Assemblages

Friday 19 September 2025 10:00 (15 minutes)

In this paper, we discuss how to understand and integrate the diverse ethical perspectives which emerge in the design, deployment, and application of AI systems. Drawing on qualitative content analysis of comments written on /r/Ethics, /r/Technology and /r/Futurology subreddits (n=145), written between July 2023 and June 2025, we identified three main folk theories regarding the ethical future of Artificial Intelligence: the first theory is that AI corporations manipulate emotions through computational performance, the second theory questions the moral limits of AI systems as well as their sentience, and the third theory highlights the tensions between ethical autonomy and technological advancements and stipulates that enhanced machine agency should take into account the respect of ethical benchmarks. All of these folk theories enact users with the necessary expertise to impose certain forms of human agency in relation to AI.

Author: OBREJA, Dragos (University of Bucharest)

Co-authors: Mr ROSNER, Daniel (National University of Science and Technology POLITEHNICA Bucharest); SIMINIUC, Rodica (Technical University of Moldova)

Presenter: OBREJA, Dragos (University of Bucharest)

Session Classification: Social Aspects of Networking Environment Today

Track Classification: Social Aspects of Networking Environment Today