Moral Norm Variation in Large Language Models

Aida Ramezani

Computational Linguistics Group

The influence of morality in our lives is indisputable. Nonetheless, our morals can change on different levels, and analyzing these changes is a requirement for understanding the dynamics of psycho-social behaviour.

Recent advances in Natural Language Processing show that language can be a window to people's moral values. Yet, how language changes can inform us of the changes in morality has not been fully explored.

In my work, I develop computational tools to analyze language change in both temporal (e.g., change throughout history and change throughout child development), and cultural (e.g., language change between different countries) levels and show how these changes correspond to the changes in people's moral beliefs, values, and behaviour.

My research connects theories of morality and moral psychology to natural language processing and language models to gain a deeper understanding of the role of morality in shaping people's social lives.

ARIA
APPLIED RESEARCH IN ACTION