

Silesian University of Technology

Faculty of Automatic Control, Electronics and Computer Science

Department of Automatic Control and Robotics

Abstract of the Doctoral Dissertation

Study and analysis of the sentiment of statements in the artificial intelligence of social robots

Author: Kamil Skowroński, MSc, Eng

Supervisor: Prof. Adam Gałuszka, PhD, DSc, Eng.

Auxiliary Supervisor: Eryka Probierz, PhD

Gliwice, September 2026

Abstract

Artificial intelligence (AI) has also found applications in robotics, especially in its service and therapeutic areas, known as social robotics. Its foundation is human–robot interaction (HRI), aimed at achieving increasingly better mutual understanding. The first stage of this process is recognising the intentions and emotional state of the person involved, as the same statement may require a different response depending on the speaker's attitude.

Emotions reach the recipient through several channels at the same time. The most obvious is facial expression, but emotional content is also conveyed by words, tone of voice, and biomedical signals that remain outside conscious control. However, AI algorithms using these channels are developed mainly for English and are rarely evaluated with regard to the hardware limitations of a robot.

The dissertation concerns sentiment analysis of statements in the AI of social robots, with particular focus on the Polish language. The text-based research began with an original dataset generated with the help of a large language model and verified by humans, and then progressed from recurrent architectures to fine-tuning the largest Polish transformer model. Response time evaluation showed that models with the lowest requirements achieve results comparable to the best-performing models while responding orders of magnitude faster. Speech research, based on the only Polish-language dataset of recordings with transcriptions, demonstrated the need to analyse the audio signal, as the content of a statement alone does not contain information about emotion, while the trained architectures outperformed a much more demanding general-purpose audio model. Biomedical signals from wearable devices were examined as a supporting channel. After combining the modalities, the results showed that the key factor is not the combination itself, but checking their consistency and withholding a decision when the available evidence is contradictory. The final part concerns the conditional activation of hate speech detection based on the relationship between hate speech content and negative polarity and negative emotions.

The research forms a coherent set of solutions for the Polish language — a dataset, text and acoustic models, a tool for analysing cross-modal consistency, and a conditional activation architecture — that can be deployed on the hardware of a social robot. The greatest potential for further development lies in naturalistic, multimodal datasets, while the research path presented in this dissertation provides a ready basis for transferring these solutions to a real-world system.