System software architecture for enhancing human-robot interaction by Conversational AI

Anna Lekova, Paulina Tsvetkova and Anna Andreva

Institute of Robotics, Bulgarian Academy of Sciences;
South-West University, Blagoevgrad, Bulgaria

Conversational AI in Social Robotics

- - conversational AI combines natural language processing (NLP) with machine and deep learning models so that people can interact in human-like manner with the digital devices.
- - conversational AI can be integrated into Socially-Assistive Robots SARs.
- - the benefits of integrating conversational AI in SARS are:
- a) improved social interaction because of the physical presence of the robot that increases motivation and engagement in activities by providing personalized feedback and encouragement;
- b) robot sensory system and motion control provide context for the surrounding environment and the transcriptions from ML-based sensor data recognition can serve as "context prompts" for big language models;
- c) robot personalized assistance to users for performing tasks and activities, such as a lexicon customized for the user or specialized assistance

Related works

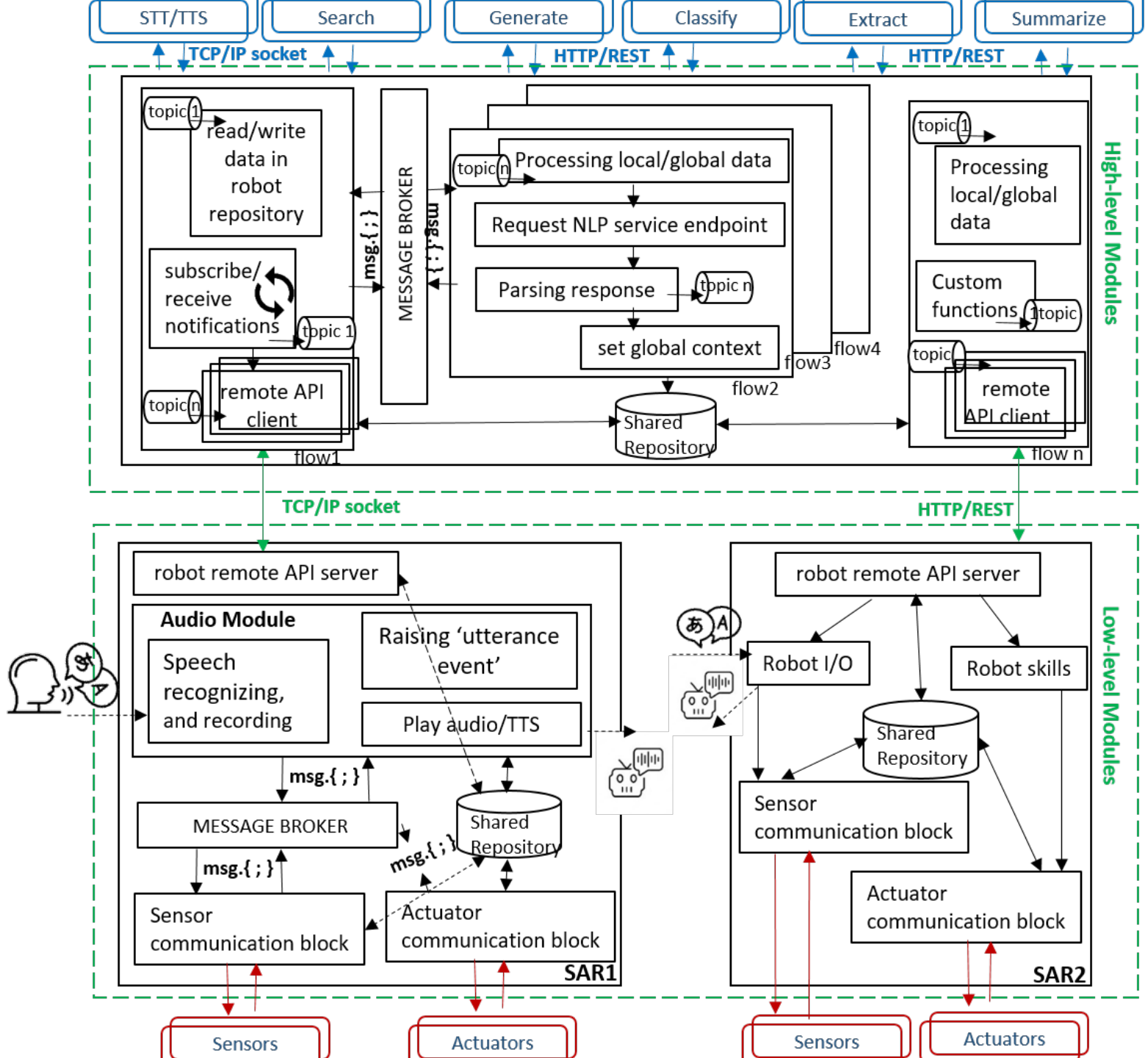
- The use of AI-based conversational agents integrated in robots is a relatively new field, with only a small number of studies published to date;
- The used AI robots in education for children with learning difficulties has been explored in different studies;
- The studies that have been reviewed have demonstrated the potential of Conversational AI in Human-Robot Interaction and interactions with “things” in smart homes;
- Some authors discuss principles and implementation of a cloud-based robot system for long-term interaction. The scientists have established the system Personal Assistant for a healthy Lifestyle.

Challenges addressed in the studies

- - technical challenges in developing AI conversational interfaces that can effectively support speech and language therapy. These challenges include reliable speech recognition, natural language processing, understanding and text generating with following text-to-speech synthesis;
- - user acceptance of AI conversational interfaces, as users may be resistant to using technology in speech and language therapy;
- - lack of evidence-based practice to support the use of conversational interfaces in speech and language therapy and a need of further research to determine their effectiveness;
- - personalization of the AI conversational interfaces to individual user's needs;
- - lack of methodology for applying Conversational AI in robots;
- - ethical and legal concerns when applying Conversational AI in robots.

Generic system software architecture

The figure illustrates the proposed system software architecture, which is composed of two layers. The first layer deals with the programing software at more abstract level that is closer to the user-robot interactions and includes high-level modules for handling utterance processing and Conversational AI by RESTful interfaces to NLP cloud services. The second layer handles the software at a lower level of abstraction, which is closer to the robot OS & hardware (sensors and actuators), and consists of low-level modules.

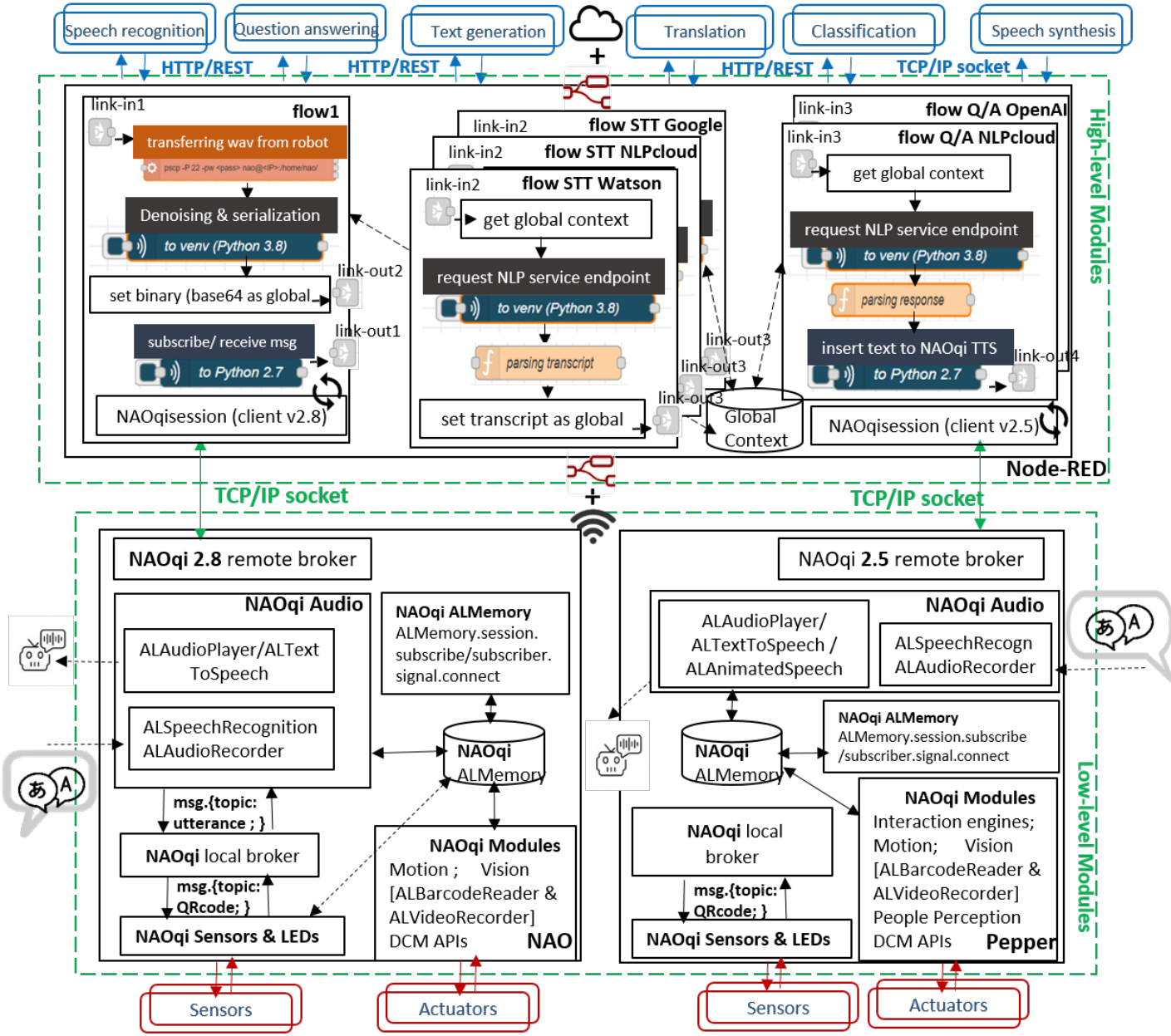


Implementation and evaluation of the proposed software architecture

The figure illustrates the architecture for a system that integrates Conversational AI in NAOqi based robots – NAO and Pepper.

2 experiments have been conducted:

- The primary objective in the first experiment was to test the performance of different NLP cloud services and the flows that access them
- The aim of the second experiment was to test the adaptability of the software architecture. The architecture was integrated into two different SARs - NAO and Pepper, to identify any possible challenges that may arise during the integration process.



Results

- The proposed software architecture was tested to demonstrate how it supports code modularity, reusability, and maintainability. Modules from the higher layer were able to read and write data to the robots' local memory remotely through either TCP/IP sockets or RESTful interfaces.
- The results after conducting the experiments showed that the software architecture is feasible for integrating Conversational AI in both robots and can facilitate human-like dialog with robots and between them.

Conclusion

- The proposed software architecture will guide the developers on how to enhance robot conversational capabilities by accessing NLP Cloud services according to the use case. The software architecture was validated by highlighting its key principles and functionalities with a specific emphasis on natural human-like interaction with Pepper and NAO robots. The results show that the integration of ConvAI in robots is feasible and enrich the human-like interactions, as well as the possibility of human-like conversation between the robots themselves.