

VOICE-CONTROLLED ROBOTIC ASSISTANT

Context-aware human-robot interaction through natural language LLM-based reasoning with vision, action, and feedback



Pro²Future

Elias Andrieu¹, Tobias W. Rutar¹, Amer Kajmaković¹, Georg Macher²

Pro2Future GmbH¹, TUG-ITI²

¹ Sandgasse 34, 8010 Graz

² Inffeldgasse 16/1, 8010 Graz



MOTIVATION & GOALS

Current robotic systems often require **technical interfaces, predefined commands, or expert knowledge** to operate them effectively. This limits their accessibility and flexibility in dynamic environments where users need intuitive interaction with robotic systems.

The SHIELD demonstrator investigates **natural and context-aware human-robot interaction** by combining speech, visual perception, LLM-based reasoning, and robotic manipulation. The goal is to allow users to describe tasks in natural language while the system interprets the request, perceives the workspace, determines suitable robot actions, and executes them.

A further objective of SHIELD is to investigate how **safeguarding mechanisms** can be integrated throughout the interaction and control pipeline to support safe LLM-driven robotic systems.

Project FactBox

Project Name SHIELD
Project ID StratP PROD.S
Duration 24 Months

Area 4.1
Cognitive Products

Project Lead
DI Dr. Amer Kajmaković

APPROACH

The demonstrator combines **natural-language interaction, camera-based perception, LLM reasoning, and robotic execution**. A spoken task request is interpreted together with visual information from the workspace. A **VLM and image segmentation** are used to detect relevant objects and estimate their positions, enabling robot actions to be derived from the current scene rather than predefined object locations.

CONTRIBUTION

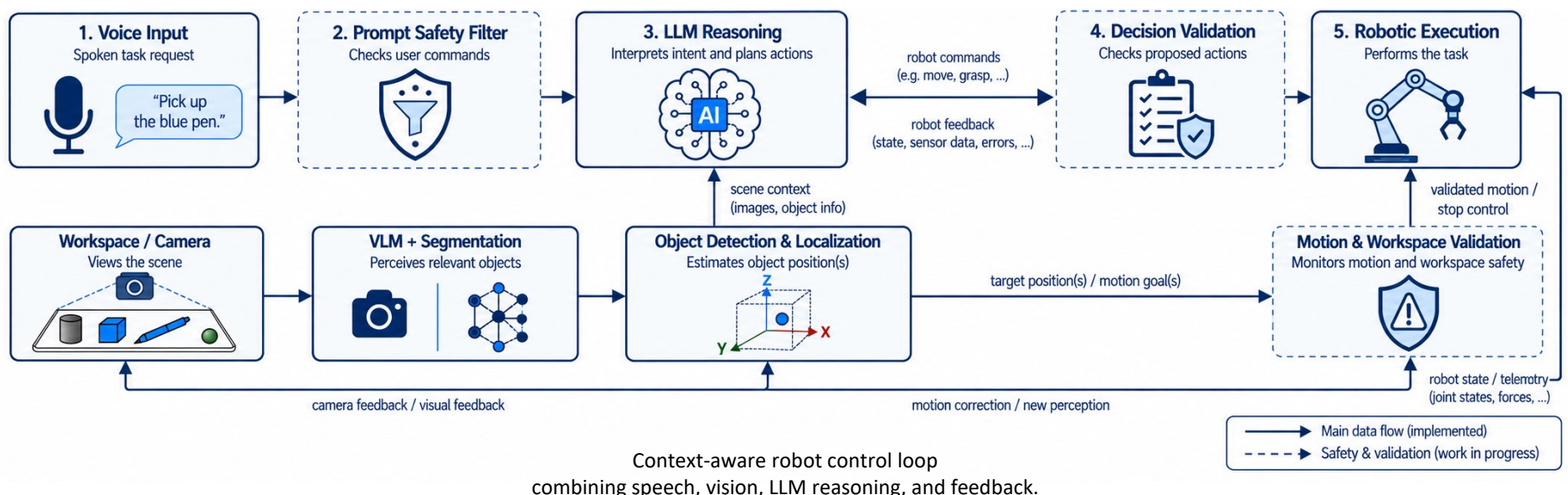
Scientific contribution

The demonstrator connects LLM-based task reasoning with visual workspace perception and robot control. It shows how user intent, object availability, and robot feedback can be combined for context-aware robotic assistance.

Economic contribution

The demonstrator shows how robotic assistants can become easier to use in work environments through natural language control. This can reduce training effort, support flexible workflows, and improve process reliability in industrial, medical, and assistive applications.

SYSTEM ARCHITECTURE



Contact: Elias Andrieu, Pro2Future GmbH, elias.andrieu@pro2future.at

Acknowledgement: This work was supported by Pro2Future II (FFG, 911655).