AI-Powered Large Language Model Framework for Natural Language Robot Control (LLM-NLRC)

Rohollah Moghadam, Ph.D, SMIEEE
Assistant Professor of Electrical and Electronic Engineering Department, College of Engineering and Computer Science

SACRAMENTO STATE
COLLEGE OF ENGINEERING & COMPUTER SCIENCE

Introduction

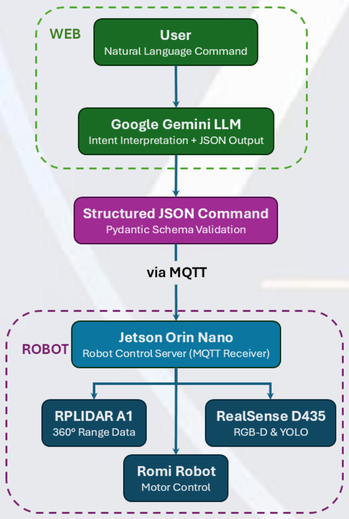
- Robotics education traditionally requires students to master programming, microcontroller interfaces, motor drivers, and sensor communication before they can operate a robot, creating a steep entry barrier.
- Large Language Models can interpret natural language and produce structured, machine-readable commands, offering a new way for students to begin engaging with robotics.
- The LLM-NLRC project connects a conversational AI interface to a working mobile robot, allowing users to control the robot using everyday English instead of code.
- This lowers the barrier for first-generation, transfer, and novice students while preserving deeper learning pathways into embedded systems and AI.

Project Objectives

- Create a natural language interface for mobile robot control.
- Integrate the interface with a fully working robot platform, not a software-only demo.
- Design a modular architecture suitable for instruction and independent study.
- Support EEE 187, EEE 225, and the Competitive Robotics Club.

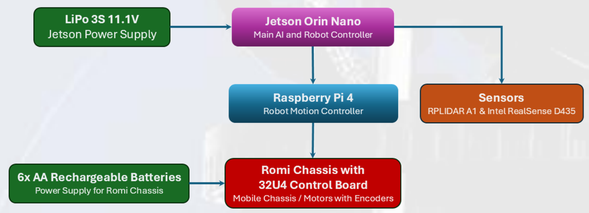
System Architecture

- ARIA uses a layered pipeline: a browser-based interface sends natural language to Google Gemini, which returns a validated JSON command. The Jetson Orin Nano receives it over MQTT and dispatches it to the robot's motors and sensors.



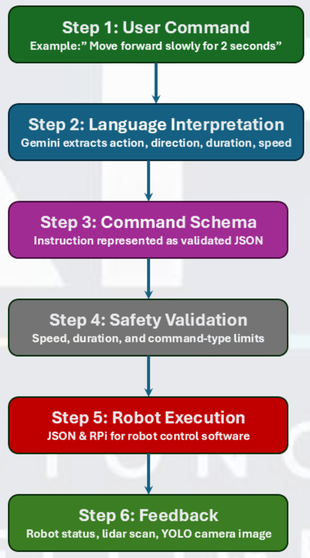
Hardware Setup

- The platform balances capability, affordability, and educational value. The LiPo battery powers the Jetson Orin Nano; the Romi chassis runs independently on 6× AA rechargeable batteries.



Software and Implementation

- **Streamlit Interface:** A browser-based control panel lets users type natural language commands without installing software.
- **AI Interpretation:** Google Gemini converts the request into a structured RobotCommand object using constrained JSON output.
- **Pydantic Validation:** Speed (≤0.3 m/s), duration (≤3.0 s), and angle (≤180°) are enforced at runtime; invalid commands default to stop.
- **MQTT Transport:** A lightweight publish/subscribe broker decouples the web interface from the robot, simplifying classroom deployment.
- **Vision Pipeline:** YOLOv8s-WorldV2 (open-vocabulary, small footprint) detects objects in RealSense RGB-D frames with median-patch depth sampling for robust distance estimates.

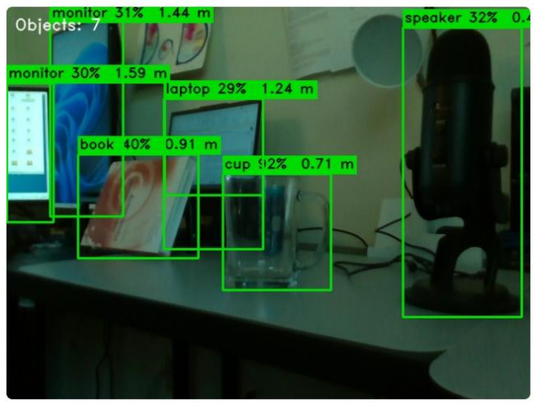


Meet ARIA

ARIA (Autonomous Robot Intelligence Assistant) connects conversational AI to a physical robot for classroom demonstration and student research.

Experimental Demonstration

- Command: "Take a photo" → get_image
- The RealSense D435 captures an RGB-D frame; YOLOv8s-WorldV2 detects objects (e.g., Speaker, Monitor, Laptop, Book, Cup) with confidence scores and depth-based distance estimates.



YOLO Camera Output

Detected Objects

Object	Confidence (%)	Distance (m)
cup	91.900000	0.710000
book	40.000000	0.910000
laptop	33.000000	1.240000
speaker	32.000000	0.490000
monitor	31.000000	1.440000
monitor	30.400000	1.590000
laptop	28.600000	1.240000

Educational Impact

- Provides a scaffolded learning path: interaction → command structure → software → hardware → advanced sensing and AI.
- Suitable for EEE 187 Robotics, EEE 225 Advanced Robotics and Control, independent study, and capstone projects.
- Supports the Sacramento State Competitive Robotics Club through workshops, demonstrations, and outreach events.
- Lowers the barrier to robotics participation for first-generation, transfer, and novice-programming students, advancing the university's diversity and inclusion goals.
- Makes AI concrete and observable: students see a language model directly producing physical, real-world robot behavior.

Project Outcomes

- Fully assembled, working robotics platform integrating AI computing, perception, motor control, and mobile power.
- Complete three-file software stack: app.py (web interface), aria_mqtt_receiver.py (robot-side handler), and camera_yolo_module.py (vision pipeline).
- Platform supports different validated command types with live LiDAR and YOLO sensor feedback.

Conclusion

- The LLM-NLRC project successfully developed ARIA, a deployable educational robotics platform that connects conversational AI with physical robot control through a complete, validated software and hardware stack. By pairing a Pydantic-validated command schema with a decoupled MQTT communication layer, the system protects classroom hardware while giving students an accessible, observable bridge between natural language and embodied AI — every architectural decision, from the modular file structure to the singleton camera-management pattern, doubles as a concrete teaching example for courses in embedded systems, robotics, computer vision, and AI.
- The project directly fulfills the goals of the Probationary Faculty Development Grant by strengthening instructional innovation, supporting faculty professional development, and creating durable educational materials, while establishing a platform for future external funding proposals in human-robot interaction, AI-enhanced pedagogy, and embedded AI systems.