PMEC@Home 2025 Team Description Paper

José R. R. Teles   Alexandre C. F. Filho   Carlos E. R. Duarte
Daniel R. da Silva   Eduardo D. Peixoto   Gabriel P. O. Ruotolo
Igor D. Aguiar   João G. da S. Chaveiro   João V. C. Souza
José G. R. Teles   Lavínia R. Lima   Letícia L. Mendes
Luis F. F. de O. Freitas   Lygia M. Rebouças
Marcelo H. L. Ferreira   Marcelo R. Jordão

February 14, 2025

Abstract. This paper presents Miss Piggy, a service robot developed by Pequi Mecânico for the RoboCup@Home competition in 2025. The robot integrates technologies in natural language processing, computer vision, robotic manipulation, and SLAM to improve autonomy and human-robot interaction. Our goal is to advance service robotics and showcase these developments in the competition.

1 Introduction

Pequi Mecânico is a non-profit student organization at the Federal University of Goiás (UFG), bringing together undergraduate and postgraduate students from multiple disciplines to research and develop robotics. Since its founding in 2011, the team has participated in various robotics competitions, including IEEE SEK, IEEE Open, RoboCup Small Size Soccer, IEEE HRR, IEEE VSSS, and RoboCup Soccer Simulation 2D. In 2019, Pequi Mecânico expanded into the RoboCup@Home league, focusing on service robotics with the goal of developing intelligent and autonomous domestic robots. The team operates with support from UFG's School of Electrical, Mechanical, and Computer Engineering (EMC), the Institute of Informatics (INF), the Center of Excellence in Artificial Intelligence (CEIA), and the Advanced Knowledge Center for Immersive Technologies (AKCIT).

The PMEC@Home team consists of students from Electrical Engineering, Computer Engineering, Physical Engineering, and Artificial Intelligence, working together to advance service robotics. Since 2022, the team has competed at the national level in Brazil's RoboCup@Home competitions, continuously improving its robotic systems and AI capabilities. Through these experiences, the team has gained valuable insights into human-robot interaction, navigation, and object manipulation, striving to develop intelligent domestic robots capable of assisting in real-world environments.

Miss Piggy (see Figure 1) is the robotic platform developed by PMEC@Home to tackle challenges in service robotics. Designed as a research and development environment, the robot features a robust structure and a modular architecture, enabling continuous experimentation and evolution of its systems. Its design integrates sensors and advanced algorithms for autonomous navigation and environmental interaction, serving as a versatile foundation for testing new approaches in artificial intelligence and mobile robotics. Built on an open-source software ecosystem, Miss Piggy offers the flexibility needed for innovative implementations, enhancing the team's technical expertise and driving the development of real-world solutions.

Fig. 1. Front view of the Miss Piggy robot

Miss Piggy's modular architecture allows for hardware and software upgrades, enabling rapid adaptation to new tasks and environments. This flexibility enhances its applicability in real-world scenarios, making it a valuable platform for research in service robotics and human-robot interaction.

2 Miss Piggy's Hardware

The hardware setup includes the Jackal robotic platform, equipped with a differential drive system, LIDAR, IMU, and wheel encoders for localization and mapping. The 6-DOF robotic arm features an adaptive gripper for versatile object manipulation. AI processing is powered by NVIDIA Jetson Xavier AGX, optimized for deep learning and computer vision tasks. The system also incorporates cameras and microphones for perception, speakers and screen displays for interaction, and an ECS BOX 2.5 for additional computational power.

2.1 Clearpath Jackal UGV

The Clearpath Jackal UGV [1] serves as a robust and versatile mobile base, making it ideal for the @Home competition. Its open-source ROS interface enables seamless integration for advanced robotics development. The onboard computing efficiently processes sensor input, allowing the team to focus on autonomous functionalities while shifting hardware efforts toward designing a custom base. By encapsulating communication in Docker containers, the system enhances interface reliability and usability. This setup enables seamless sensor access via micro-ROS and ROS2 [2], linking the base to the systems managing Miss Piggy's intelligent autonomous behavior. The Jackal's adaptability and performance are crucial to the project's success.

2.2 Manipulator

The manipulator integrated with Miss Piggy is the ViperX 300s 6DoF (Trossen Robotics) [3], featuring eight Dynamixel servos (XM540 and XM430) arranged to achieve six degrees of freedom and a maximum payload capacity of 750g. The automation strategy adapts to the task at hand, employing either traditional control techniques or reinforcement learning algorithms as needed.

2.3 Jetson Xavier AGX

The Jetson AGX Xavier [4] is a state-of-the-art AI computing platform designed for high-performance and energy-efficient processing by NVIDIA. It features an 8-core ARM v8.2 64-bit CPU, a 512-core Volta GPU with Tensor Cores, and a dedicated Deep Learning Accelerator, enabling it to handle complex AI and deep learning tasks with ease. Equipped with 32GB of LPDDR4x memory and a high bandwidth of 137GB/s, it ensures fast and seamless data processing. Its compact and power-efficient design makes it ideal for embedded applications. In Miss Piggy, the Xavier is dedicated to AI applications such as inference in large language models (LLMs), speech-to-text, text-to-speech, computer vision, and reinforcement learning models.

2.4 Auxiliar Components

The auxiliar components are of huge importance to our robot functionality, they enhance the experience and help to achieve a better performance. They consist in:

- **HRI display**: 15" screen, located on the robot's front for User Interaction displaying Miss Piggy's expressions and a subtitle function.
- **LOG display**: also, a 15" screen on the robot back for debugging and to tasks overview and companion.
- **RGB-D Sensors**: Realsense cameras (d435i and d455), on the robot's top and lower front as sensors helping with recognition tasks gathering RGB-D images and using it as subjects and for navigation tasks employing the images to create point clouds to help with all the algorithms.

– **Lidar 2D**: RPLIDAR A2, on the robot's lower back to help with the navigation and localization systems.
– **Microphone**: Rode Videomic GO, on top of the robot to receive directional audio input to be used as data for tasks and operations.
– **Speaker**: LG XBOOM XG5s to reproduce the text-to-speech output.

2.5 ECS BOX

The power distribution of a robotic system, such as Miss Piggy, constitutes a fundamental component of its functionality. To address this issue, a decision was made to develop a unified distribution system that retains a degree of modularity, thereby ensuring the system's adaptability to potential future modifications to the robot's internal components and adhering to electrical specifications as illustrated in the following diagrams on Figure 2.

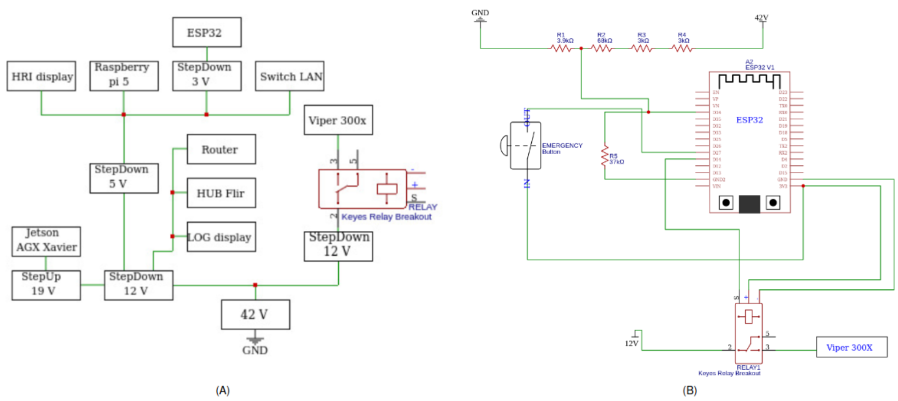

Fig. 2. (A) ECS BOX Electric Diagram. (B) ECS BOX Electronic Diagram.

The Energy Center System (ECS) is composed of two primary components: a board that receives all the required tension regulators that are properly connected to distribute the primary power source, a 42 V lithium-ion battery, to all the eight interfaces mounted on both sides. Each interface is equipped with a switch for manual control of the power, a fuse to protect against overloads, and an XT60 connector that outputs the correct electric configuration to the specific component connected, as is shown in Figure 3.

Fig. 3. ECS BOX power management system.

It is important to note that this ECS BOX version is still in the prototype stage and has yet to fully meet its initial objectives. However, it has demonstrated the potential to upgrade its primary function, which is to monitor the battery state, generate diagnostics, and handle various emergency situations. The prototype is equipped with an ESP32, a Micro-ROS framework that communicates with a ROS2 environment, a buzzer to alert when the battery reaches a low point, a relay assembled with an emergency button to cut he power distribution and prevent the robot to harm people or to damage its structure or component.

3 Miss Piggy's Software

Miss Piggy's software stack integrates advanced algorithms and frameworks to enable autonomous navigation, object manipulation, and seamless human-robot interaction. The system leverages real-time mapping, deep learning-based perception, and motion planning to operate efficiently in dynamic environments.

3.1 Autonomy

The primary function of Miss Piggy's autonomy stack is to facilitate autonomous navigation in both familiar and unfamiliar environments. To address this challenge, the autonomy team addresses three interdependent domains that must function in unison to ensure safe and efficient navigation:

The mapping domain employs RTAB-Map [ref] for 2D mapping, leveraging data from the RGB-D Auxiliary Camera and the 2D Lidar mounted on the robot's posterior.

Localization is achieved through odometry and IMU fusion, processed with an Extended Kalman Filter (EKF), with the help of pose updates related to the map also provided by RTAB-Map.

Navigation is managed by Navigation 2 (ROS2), which includes route planning and obstacle detection using LIDAR and RGB-D sensors.Several cost map layers for global planning and a dynamic local layer for adjusting trajectories and avoiding collisions are used. The SMAC Local Planner further optimises trajectory planning.

We are currently experimenting with the various approaches provided by both the mapping and navigation frameworks and, based on the limitations encountered, we instigating their research to solve the proposed challenges.

3.2 Perception

Miss Piggy's perception system integrates advanced computer vision techniques to ensure robust and efficient recognition of both people and objects. For object detection and human pose estimation, the system leverages YOLO11 [5], a state-of-the-art deep learning model known for its speed and accuracy. To achieve reliable individual identification, DeepFace [6], in conjunction with MTCNN [7], is employed for precise face recognition. Additionally, person tracking is handled using ByteTrack [8], which enables stable and consistent tracking of individuals across frames, even in dynamic environments.

To maintain real-time performance, all inference processes are optimized using TensorRT [9], significantly accelerating computations while reducing latency. This optimized pipeline enables Miss Piggy to swiftly and accurately perceive its environment, making it well-suited for dynamic and interactive applications.

Ongoing improvements focus on enhancing the perception system further. The object detection pipeline is being improved by refining data acquisition, annotation processes, and model training, ensuring higher-quality datasets and more accurate detections. Additionally, the person recognition approach is also being expanded to enable identification without reliance on explicit facial features, increasing adaptability in real-world scenarios.

3.3 Human Robot Interaction

To enable interaction with Miss Piggy, as shown in the Figure 4, the system integrates automatic speech recognition (ASR), natural language processing (NLU), and text-to-speech (TTS) components. For speech processing, we utilize the Riva Toolkit [10], provided by Nvidia, which allows the implementation of a Conformer model for highly efficient and high-quality speech-to-text (STT) conversion. Additionally, for speech synthesis, we employ FastSpeech2 [11], which is also part of the Riva Toolkit. At the core of the system, the NLU module is based on the LLaMA 2 [12] base model, leveraging the efficiency provided by Ollama [13]. All these components are seamlessly integrated through ROS2 [2], enabling flexible and customizable interactions tailored to specific tasks.

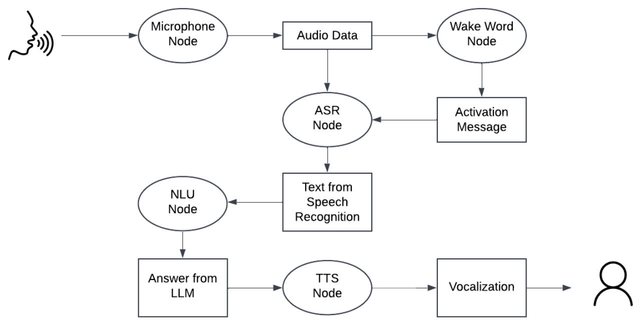

Fig. 4. A diagram of the current architecture of Miss Piggy Human Robot Interaction.

Furthermore, the entire system operates through a wake word pipeline, trained using OpenWakeWord [14], ensuring that the model responds only when explicitly activated. Additionally, our Retrieval-Augmented Generation (RAG) system enables the robot to access information beyond its base training, enhancing its contextual understanding and response capabilities. These advancements collectively contribute to a more intelligent, responsive, and adaptable interaction system, improving the overall user experience and practical applicability.

Given the team's extensive background in artificial intelligence, we have explored various areas of human-robot interaction. Currently, our research focuses on advancing multiple aspects of this field. One key area of development is the enhancement of the wake word pipeline, where we aim to extend Miss Piggy's activation beyond its designated name to include emergency words and movement commands.

Additionally, to contribute to the scientific advancement of speech technologies in Brazil, we are working on the development of ASR/TTS models in Portuguese, enhancing embedded speech capabilities within a domestic environment.

Finally, we are continuously innovating our natural language understanding architecture in response to the rapid advancements in large language models. To keep pace with these developments, we are designing a novel graph-based architecture [15], enabling greater customization and generalization of our applied model.

3.4 Manipulation

The approach to utilizing the ViperX300S is based on the open-source ROS2 stack provided by Trossen Robotics [16] combined with MoveIt [17] for configuring object capture routines, supported by the CHOMP planner [18] for manipulation. Object position identification is achieved using the RealSense D455, which, through perception algorithms, estimates the spatial location of artifacts in a three-dimensional environment for grasp pose estimation.

4 Conclusions and future work

In this work, we presented Miss Piggy, our service robot developed for RoboCup @Home 2025. The robot integrates state-of-the-art technologies in artificial intelligence, computer vision, speech processing, and autonomous navigation to improve its performance in human-robot interaction and domestic assistance tasks. Through a modular hardware and software architecture, Miss Piggy enables continuous development and adaptation to new challenges in service robotics.

For future work, we aim to refine perception and human-robot interaction, enhancing real-time recognition and wake word detection. Additionally, we are expanding our efforts in scientific research, fostering undergraduate research initiatives and increasing our production of publications in robotics and AI. As we prepare for RoboCup@Home 2025, we remain committed to advancing both our technical developments and academic contributions to the field.

References

1. Clear Path Robotics. Jackal unmanned ground vehicle. https://clearpathrobotics.com/jackal-small-unmanned-ground-vehicle/. Accessed: 2025-02-13.
2. ROS. Robot operating system. https://www.ros.org/. Accessed: 2025-02-13.
3. Trossen Robotics. Viperx 300 s, our largest and most capable research manipulator arm. https://www.trossenrobotics.com/viperx-300. Accessed: 2025-02-13.
4. Nvidia Corporation. Jetson agx xavier family. https://www.nvidia.com/pt-br/autonomous-machines/embedded-systems/jetson-agx-xavier/. Accessed: 2025-02-13.
5. Ultralytics. Yolo11. https://docs.ultralytics.com/pt/models/yolo11/, 2024. Accessed: 2025-02-13.
6. Y. Taigman, M. Yang, M. Ranzato, and L. Wolf. Deepface: Closing the gap to human-level performance in face verification. In *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1701–1708, 2014.
7. Jia Xiang and Gengming Zhu. Joint face detection and facial expression recognition with mtcnn. In *2017 4th International Conference on Information Science and Control Engineering (ICISCE)*, pages 424–427, 2017.
8. Yifu Zhang, Peize Sun, Yi Jiang, Dongdong Yu, Fucheng Weng, Zehuan Yuan, Ping Luo, Wenyu Liu, and Xinggang Wang. Bytetrack: Multi-object tracking by associating every detection box. In *Computer Vision – ECCV 2022: 17th European Conference*, page 1–21, 2022.
9. NVIDIA Corporation. Tensorrt. https://docs.nvidia.com/deeplearning/tensorrt/archives/tensorrt-825/pdf/TensorRT-Developer-Guide.pdf, 2024. Accessed: 2025-02-13.
10. NVIDIA Corporation. Nvidia riva sdk user guide. https://docs.nvidia.com/deeplearning/riva/user-guide/. Accessed: 2025-02-13.
11. Yi Ren, Chenxu Hu, Xu Tan, Tao Qin, Sheng Zhao, Zhou Zhao, and Tie-Yan Liu. Fastspeech 2: Fast and high-quality end-to-end text to speech. *arXiv*, 2020.
12. Jiaxiang Duan, Zhuohan Li, Sheng Shen, Xiaoxia Wu, Yi Ding, Eric Xing, Joseph E. Gonzalez, and Ion Stoica. Fastserve: Efficient model serving for foundation models. *arXiv*, 2023.
13. Ollama. Ollama: Get up and running with llama 3.3, deepseek-r1, phi-4, gemma 2, and other large language models. https://github.com/ollama/ollama. Accessed: 2025-02-13.
14. Daniel Scripka. openwakeword: Open-source wake word detection. https://github.com/dscripka/openWakeWord. Accessed: 2025-02-13.
15. Jialin Wang and Zhihua Duan. Agent ai with langgraph: A modular framework for enhancing machine translation using large language models. *arXiv*, 2024.
16. Trossen Robotics. Ros 2 open source packages. https://docs.trossenrobotics.com/interbotix_xsarms_docs/ros2_packages.html. Accessed: 2025-02-13.
17. MoveIt. Moveit 2 documentation. https://moveit.picknik.ai/main/index.html. Accessed: 2025-02-13.
18. MoveIt. Using chomp planner. https://moveit.picknik.ai/main/doc/how_to_guides/chomp_planner/chomp_planner_tutorial.html. Accessed: 2025-02-13.

Miss Piggy's Hardware Description

Mechanical and eletric specifications of Miss Piggy Robot Platform:

– Base: Clearpath Jackal UGV
– Torso: Cage of aluminum profiles(40x40mm)
– Arm: Attached to aluminum cage with base perpendicular to ground(6DoF)
– End Effector: Standard Trossen Gripper (750g Maximum payload)
– Gimbal: Dynamixel XM540(Pan), Dynamixel XM430(Tilt)
– Robot Dimensions: Width: 0.44m, Length: 0.5m, Height: 1,33m
– ECS-BOX: The power distribution system of Miss Piggy Platform
– Jetson AGX Xavier: Focused Computer on AI Algorithms
– HUB Flir USB: Hub to manage USB Connections
– Raspberry Pi 5: Computer focused on Display Interaction and Logging
– Ethernet Switch: Internet interface to connect all computers.

Fig. 5. Robot Miss Piggy

Sensoring on Miss Piggy Platform: :

– RealSense d455: Person Recognition and Object Detection
– RealSense d435i: Mapping and Navigation
– Rode VideoMicGo Microphone: Human Interaction Voice Capture
– LG XBOOM X5: Human Interactive Speaker
– Robopeak A2(Lidar 2D): Rear mounted on Miss Piggy

Miss Piggy's Software Description

For our robot we are using the following software:

– Operating System: Ubuntu 22.04
– Middleware: ROS2 Humble
– Navigation: ROS2 navigation stack
– Localization: EKF
– Mapping: RTABMap
– Object Recognition: Yolo11
– Face Detection: Yolo11
– Human Detection: Yolo11
– Face Recognition: DeepFace
– Speech Synthesis: FastSpeech
– Speech Recognition: Conformer-CTC
– Natural Language Understanding: Mistral