TIAGo Head: an AI Powered Platform for Social Robotics

Sara Cooper[1,2], Séverin Lemaignan[1], Raquel Ros[1,2], Lorenzo Ferrini[1],
Ferran Gebellí[1], Luka Juričić[1], Narcis Miguel[1], Luca Marchionni[1], Francesco Ferro[1]

Abstract— **This paper presents the TIAGo Head, a new tabletop social robot from PAL Robotics, focusing on its capabilities as an HRI platform. We detail the robots' hardware, highlighting its sensors/actuators and on-board computing; and its software architecture, including social perception, expressive face, a knowledge base, and integration with large language models (LLMs) for natural conversations. We also describe a use-case in a receptionist scenario where TIAGo Head dynamically interacts with travelers by displaying news and conversing.**

I. INTRODUCTION

Social robotics has emerged as a transformative field that enables robots to interact, communicate, and collaborate with humans in natural and intuitive ways [1], [2]. Tabletop and small-sized social robots, in particular, are gaining attention for their compact and engaging designs that allow them to function effectively in confined environments such as offices, classrooms, or kiosks [3]. Unlike humanoid mobile robots, these robots focus on face-to-face interaction, emphasizing facial expression, eye gaze, and verbal communication to engage users [4]. These features make them ideal for scenarios with space constraints, stationary interaction points (e.g. reception desks or information booths), or use with children, where full-sized robots may not be practical [5].

TIAGo Head (Figure 1), developed by PAL Robotics, is a new addition to the family of social robots. It is a version of the full-sized TIAGo Pro[1], targeted both at human-robot interaction researchers and applications in different sectors such as hospitality, health or education to serve as a receptionist or an interactive companion.

This paper presents the following:

- A review of related tabletop social robots and their impact on HRI research.
- A detailed description of TIAGo Head's hardware and software architecture.
- An example application of the platform

II. RELATED WORK

Next, we examine some of the most popular social robots from the past decade that are similar to the TIAGo Head presented in this paper, highlighting their key features, followed by summarizing the main use cases for these in different areas.

[1]Pal Robotics, C. de Pujades, 77, 7-7, 08005 Barcelona, Spain `name.lastname@pal-robotics.com`
[2]IIIA-CSIC, Campus UAB, 08193 Bellaterra, Spain `name.lastname@iiia.csic.es`

[1]https://pal-robotics.com/es/robot/tiago-pro/

Fig. 1. TIAGo Head social robot

A. Interactive social robots

Different kinds of interactive social robots have been developed over the years, demonstrating significant advances in human-robot interaction (HRI) and artificial intelligence (AI). Figure 2 illustrates the set of social robots reviewed in this work, while TableI summarizes their main specifications.

As for **general features**, *Haru* is a highly expressive robot that relies on fluid eye movements and gestures to engage users emotionally [6], while *Furhat* is characterized by its 3D projection-based system to display realistic facial expressions [7]. *Nao* [8] and *QTRobot* [9] are small humanoid robots that can walk in small areas and use their hands for increased expressiveness. On the other hand, *Lovot* and *Buddy* are 2-wheeled emotional companion robots. Lovot focuses on fostering affection and human connection through tactile sensors, touch-sensitive fur, and lifelike movements [10], while Buddy provides assistance with daily tasks such as reminders and home security monitoring [11].

Pricing varies based on the characteristics of each robot. Furhat is among the most expensive (€23,250–€26,970), targeting advanced research and commercial applications. NAO costs approximately €13,010, making it more affordable for education and research institutions. QTRobot ranges from €10,140–€17,580, with upgrades to computing power. Buddy (€1,490–€3,350) and Lovot (€3,530–€4,090) are more affordable but require subscriptions for full functionality. Haru, in contrast, is still not commercially available and is only used as a research platform in collaborative projects.

Computing power ranges from external processing (Haru) to powerful internal systems like QTRobot's Nvidia Jetson Orin (where the higher the cost, the computing power is increased). Furhat, NAO, and QTRobot offer mid-range

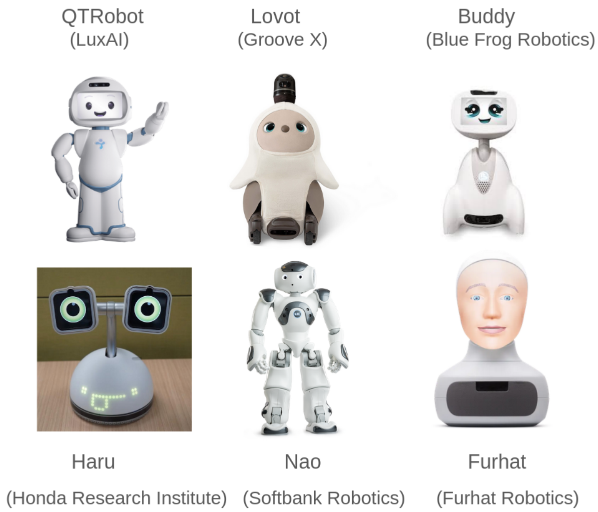

Fig. 2. Different types of small social robots available

processing, with NAO using an Intel Atom CPU and Furhat featuring an Intel Core i5. Lovot is optimized for real-time emotional interaction, using a custom GPU with 1,024 cores and 32 Tensor Cores.

Degrees of freedom (DoF) impact motion and expression. NAO (25 DoF) enables complex humanoid movement, while QTRobot (12 DoF) and Lovot (13 DoF) prioritize expressive interactions. Furhat (3 DoF) and Haru (5 DoF) focus on facial gestures, while Buddy (4 DoF) is limited in motion but designed for user-friendly social engagement.

Software and API openness is another factor to consider. NAO, QTRobot, Furhat, and Buddy offer open APIs with support for Python, C++, and ROS, among others, enabling research and development. Furhat provides an advanced Kotlin-based SDK, while NAO integrates visual programming via Choregraphe. Buddy, with its Android-based SDK, supports multiple languages, including Java and C#, targeted at phone applications. Lovot, in contrast, runs on a closed system with minimal customization, focusing on companionship rather than development.

B. Domain use cases

Interactive social robots have shown effectiveness in various fields, including education, therapy, healthcare, customer service, and companionship. Next, we review four main areas.

Education and Autism Therapy Social robots such as Nao and Furhat have been widely used in educational settings, helping students with STEM learning, language acquisition, and interactive tutoring, enhancing engagement and improving learning outcomes through multimodal interaction [12].

In therapy, robots such as QTRobot have demonstrated effectiveness in structured interventions for children with Autism Spectrum Disorder (ASD), supporting the development of social skills, emotional awareness and cognitive engagement through guided activities and interactive storytelling [13]. Haru has also been tested for this purpose, and research has demonstrated its ability to increase eye contact, turn-taking, and engagement in children with autism [14]. Similarly, Haru has been used in socioemotional interaction studies and cross-cultural communication research, showing positive outcomes of engagement and emotional bonding [15], [16].

Elderly Care and Cognitive Therapy Robot-based applications have been deployed to combat loneliness, encourage cognitive engagement, and support independent living [17]. These robots offer daily reminders, conversation prompts, and companionship, reducing social isolation, and improving mental well-being by playing cognitive games. As such, robots are perceived as supportive conversational partners and motivators for daily activities [18].

Customer Service and Public Spaces Social robots are increasingly used in customer service and hospitality, acting as receptionists, information assistants, and interactive kiosks in public spaces such as airports, hotels, and retail stores [19]. Robots like Furhat [7] have been deployed in banks, museums, and shopping centers, where they provide directions, answer customer inquiries, and engage in small talk. Their expressive and verbal interaction capabilities make them an effective alternative to static touchscreen kiosks, allowing for more engaging and accessible user experiences.

Human-Robot Interaction (HRI) Research Social robots also serve as experimental platforms for HRI research, allowing the community to study gaze behavior, emotion recognition, and conversational AI in controlled environments, among others [20]. Robots such as Haru and Furhat have been used in psychological experiments and cross-cultural interaction studies, helping researchers analyze how humans respond to robotic facial expressions, speech modulation, and adaptive responses [21].

TIAGo Head is tabletop-type social robot that supports the above user cases in the different domains by incorporating human perception through the ROS4HRI framework, LLMs to facilitate natural language interaction and beyond, situated reasoning and knowledge representation, multimodal expressiveness, and an intent-based controller to supervise the execution of parallel/sequential tasks and skills.

III. TIAGo HEAD: HARDWARE AND SOFTWARE ARCHITECTURE

A. Hardware Architecture

TIAGo Head is a modular robotic head derived from the full PAL TIAGo Pro mobile manipulator. Although TIAGo Pro is designed for a wide range of robotic applications involving navigation, social interaction, and manipulation, TIAGo Head provides a cost-effective, standalone alternative focused exclusively on social interaction.

Vision and Perception: The TIAGo Head is equipped with an RGB-D depth sensor for depth perception and object recognition. This camera allows the robot to perceive 3D structures, estimate distances, and track people or objects in its field of view. Based on these sensory data, the TIAGo Head's perception system can perform facial detection and

TABLE I

COMPARISON OF SOCIAL ROBOT SPECIFICATIONS

Robot	Price	OS and API	Degrees of Freedom	Processing Capability
TIAGo Head	$8,581 - $21,481 (varies by specs)	Open API (ROS API), Ubuntu Linux, Python, C++	2	Intel i7 / 16 GB RAM / 500 GB SSD, AI/GPU Nvidia Jetson Orin (included in AI package)
Haru	Not commercialized, research platform	ROS API, no specific SDK, Linux, Python, C++	5	External computer
Furhat	$25,000-$29,000 + $5000-$7000 yearly for extra software	Open API, Furhat SDK (Kotlin Skill API, Remote API), Python, C#, JavaScript, Rust, Kotlin	3	Intel Core i5, 8GB RAM, Iris Plus 640 GPU (Internal)
Nao	$13,990	Proprietary OpenNAO (Linux based on Gentoo), Open API (ROS, NAOqi), Choreographe visual programming, Python, C++, .Net, Java, Matlab, Urbi	25	Intel Atom Z530, 1.6 GHz, 1GB RAM (Internal)
Lovot 3.0	$3,800 - $4,400 + $80 monthly subscription required	No API, custom OS, Controlled via app, limited customization	13	x86 CPU (main), ARM CPU (sub), 1,024 GPU cores, 32 Tensor Cores (Internal)
Buddy	$1,600 - $3,600 + yearly subscription of $1,540-$3,600	Open API, Android-based SDK, C, C++, C#, Java, JavaScript, Python	4	Android 9, 8-core CPU, 3GB RAM (Internal)
QTRobot	$10,900 - $18,900 (varies by version)	Open API, Ubuntu Linux (ROS API, QTRobot APP), Python, C++, JavaScript	12	Varies: Raspberry Pi 4B, Intel NUC i5/i7, or Nvidia Jetson AGX Orin

recognition, track human poses, skeleton tracking, and infer basic emotional cues from facial expressions.

Expressive Mechanisms: TIAGo Head has two actuated degrees of freedom for pan-tilt motion that allow head gestures like nodding, shaking, or tilting. It also features a digital face on a 10” touchscreen display, capable of showing eye shapes, mouth motions, and other expressions to convey emotions. Additionally, programmable LED lights –four RGB zones, one on each cheek and ear, respectively– can signal emotions or statuses with colors and patterns. These multimodal expressive channels allow TIAGo Head to, for example, nod in agreement, turn to face a speaker, blink, smile, and use colored lights, among others, to enhance expressiveness and improve the naturalness of interactions.

Auditory and Speech Hardware: The TIAGo Head features a built-in speaker system and a 4-microphone array for two-way audio communication, enabling far-field voice detection and sound source localization, helping the robot determine when and from where a person is speaking. The microphones are coupled with noise-cancellation algorithms to improve speech recognition in noisy environments. The audio output is provided by 2 × 4W loudspeakers, allowing TIAGo Head to generate clear voice output and sounds for social interaction.

Onboard Computing and Power: Inside its base, TIAGo Head carries an onboard ITX computer to process sensor data and run AI algorithms in real time, without the need for an external computer. The standard configuration includes an Intel i7 processor with 16 GB RAM and a 500 GB SSD, enabling efficient vision processing and language models. An NVIDIA Jetson Orin AGX GPU accelerator can be optionally added for faster neural network inference. The system runs Ubuntu Linux with real-time patches and the Robot Operating System (ROS 2) middleware for flexible robotics integration.

External touch-screen: The robot is provided with an optional external touch screen that is plugged to the robot through a HDMI cable, and can be used to show different supporting content for the user’s applications.

Table II details the key hardware specifications of the TIAGo Head platform.

TABLE II

TIAGO HEAD HARDWARE SPECIFICATIONS

Specification	Value
Height	45 cm
Weight	5.75 kg
Cameras	RGB-D depth sensor, Optional 360°
Microphones	4-mic array with noise cancellation
Speakers	2 × 4W loudspeakers
Processor	Intel i7, 16 GB RAM, 500 GB SSD
Optional GPU	NVIDIA Jetson Orin
Operating System	Ubuntu + ROS 2
Connectivity	Wi-Fi 6 (802.11ax), Bluetooth 5.2, 4 × USB, 2 × Ethernet, 1 × HDMI, 2 × DP, Audio Line OUT, Mic IN
Power Supply	230V / 110V AC-powered (no battery)
Optional Display	External touch-screen

B. Software Architecture

TIAGo Head’s software is built on ROS 2 and uses standard frameworks for human-robot interaction. The high-level software modules are indicated in Figure 3.

The ROS4HRI (an open standard adopted as ROS REP-155) framework provides a unified interface to detect and represent human users [18]. This framework integrates perception, speech recognition, and behavioral responses. Most of the ROS 2 nodes of the robot are open source[2]. Thanks to the interoperability offered by this framework, developers can integrate different algorithms (face recognition, skeleton tracking, etc.) that all output to the ROS4HRI human model format. This feature is a highlight in comparison to the proprietary and closed-source frameworks that other robots like Nao and QTRobot use.

Social perception

For face detection, the robot uses the `hri_face_detect` ROS 2 node for real-time face detection with the YuNet face detector and Mediapipe Face Mesh. This ensures accurate tracking of multiple faces and enhances social interactions.

For body detection, the robot uses the

[2] `https://github.com/ros4hri/`

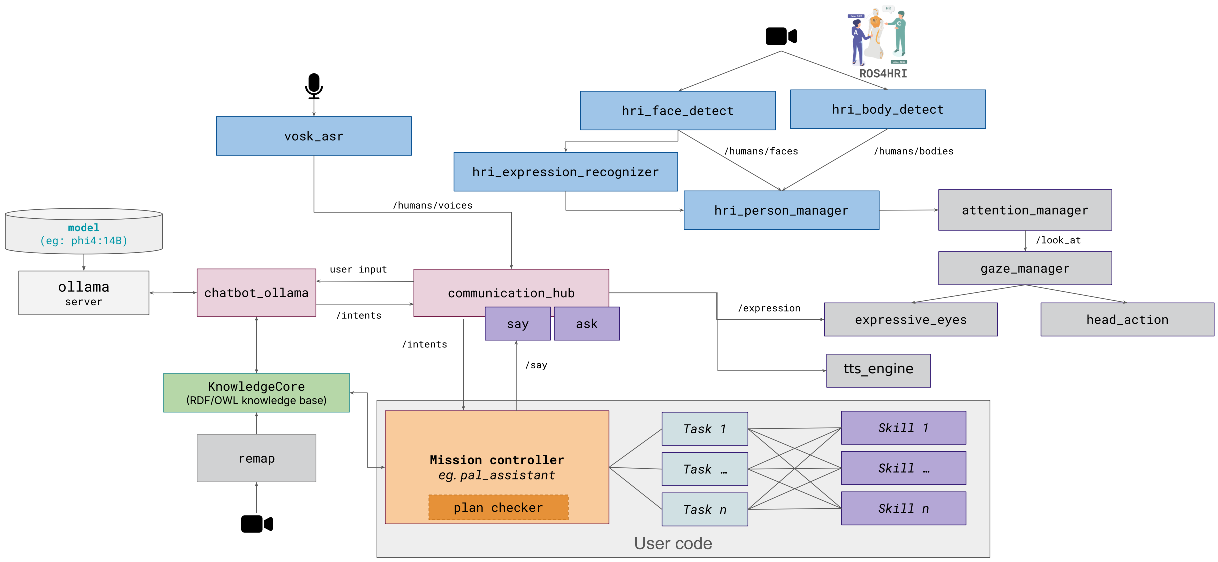

Fig. 3. TIAGo Head's software architecture. Blue nodes represent perception modules (part of ROS4HRI), green nodes represent reasoning modules, and light red nodes manage verbal communication. Orange, light blue and purple nodes represent control nodes (mission, tasks, skills), and are usually application-specific.

`hri_body_detect` node, based on Google Mediapipe 3D body pose estimation. It works for multiple bodies. The `hri_person_manager` evaluates and merges the face and body information to provide a unique person ID for the detected person [22].

The perception system includes expression recognition with `hri_expression_recognizer`, based on the ONNX Ferplus model, which detects emotions such as neutral, happiness, surprise, sadness, anger, disgust, fear, and contempt.

Speech interaction and LLMs

TIAGo Head features speech recognition (ASR) and text-to-speech (TTS) capabilities, supporting over 20 languages. It integrates the open source Vosk[3] local ASR engine and proprietary Acapela[4] for TTS, also supporting different voices (female, male). By default, it uses the RASA [23] chatbot engine locally. It features an internalization manager to manage robot languages by means of a ROS action. The `communication_hub` is in charge of managing the overall speech interaction with skills such as `say` something.

TIAGo Head integrates LLMs (Large Language Models) for conversational AI, supporting cloud-based (OpenAI API) [24] and local (Ollama, as in the example of Figure 3) inference [25]. This enables context-aware dynamic responses beyond scripted dialogues.

To evaluate the performance of different LLM models on TIAGo Head's hardware, we conducted benchmark tests measuring the token processing rates. Table III summarizes the representative models running locally via Ollama on the TIAGo Head's Intel Core i7-13700 (CPU-only), and then separately on the Orin AGX from Nvidia, optionally included with the robot. We report two metrics: Prompt Evaluation Rate (how fast the model processes the initial prompt, in tokens per second) and Generation Rate (how fast it generates reply tokens)[5].

TABLE III
LLM PERFORMANCE LOCALLY (INTEL I7 CPU) AND EXTERNAL GPU (ORIN AGX).

Model	Prompt Eval Rate (tokens/s)	Generation Rate (tokens/s)	Hardware
LLaMA3.2 1B (1.3B params, int4)	57.8	24.8	Intel i7 CPU
Qwen 1.5B (int4 quantized)	14.3	10.0	Intel i7 CPU
LLaMA3.2 3B (3B params, int4)	57.8	3.74	Intel i7 CPU
Phi-4 14B	148.0	9.66	Orin AGX
LLaMA3.2 1B	1006.0	34.4	Orin AGX

The results indicate that smaller LLMs achieve higher real-time speeds on the given hardware [26], [27]. For instance, on the Intel i7 CPU, the 1-billion-parameter LLaMA3.2 model runs at 24.8 tokens/sec during generation, enough for responsive dialogue, whereas a 3B model slows to 3.7 tokens/sec, which could introduce noticeable lag. The 1.5B Qwen model falls in between.

On the Orin AGX, performance varies significantly, with the same LLaMA3.2 1B model averaging 34.4 tokens/sec, and Phi-4 14B achieving 9.66 tokens/sec. Notably, Orin's AI-optimized hardware enables faster prompt evaluation speeds, particularly for LLaMA3.2 1B, which reaches more than 1000 tokens/sec.

These benchmarks guided our integration choice: using an efficient model locally for quick responses, while reserving larger models or GPT-4 access for when superior language understanding is needed (at the cost of speed). In practice, TIAGo Head can flexibly switch between local and cloud

[3] https://alphacephei.com/vosk/
[4] https://www.acapela-group.com/

[5] All tests were conducted with Ollama v0.5.7; int4 = 4-bit quantization. Qwen 1.5B is an Alibaba Qwen model; LLaMA2 models accessed through the Llama.cpp backend.

Fig. 4. Expressions and motions embedded in text the robot says.

LLMs based on connectivity, computational capacity, and required response sophistication.

Situated reasoning

The Situated Reasoning subsystem integrates OpenAI-compatible LLMs (e.g., ChatGPT, Ollama) with the robot's `KnowledgeCore` to enable natural interaction and context-aware decision-making. The system dynamically interprets its surroundings, processing user input along with the environmental context. `remap` is one of the tools that allows this, a framework for efficiently storing and retrieving spatially-grounded semantic data [28].

TIAGo Head uses the OpenRobots Ontology [29] for symbolic reasoning with OWL2 RL semantics. This ontology provides a structured vocabulary for describing the environment and inferring relationships between objects, agents, and actions.

The reasoning subsystem sends structured prompts to the LLM whenever the user speaks or interacts (via the simulator's chat interface, see below) with the system.

Behavior manager

User interactions are abstracted into *intents* that the robot processes. These intents range from simple utterances to complex task sequences. The `Mission Controller` manages incoming intents and delegates task planning based on the robot's available skill, prioritizing, and scheduling tasks to ensure goal completion. This design allows the system to perform goal-oriented actions.

Each task then triggers different skills, such as navigating to a place, detecting an engaged user, or saying something.

Expressiveness and attention manager

The `attention_manager` node implements the robot's attention focus logic. It operates according to three different policies and publishes on the ROS topic `/look_at` the target points for the `gaze_manager`.

- `DISABLED`: The robot does not change its focus of attention.
- `RANDOM`: The robot looks around itself at random.
- `IDLE_SOCIAL`: The robot looks at the detected faces in the scene, switching focus periodically if multiple faces are present. If no faces are detected, it defaults to the `RANDOM` policy.

In addition, the robot face can display expressions (depicted in Figure 4), set either through the expression name itself or by indicating valence and arousal using the `expressive_eyes` ROS 2 node. In addition, it can use its pan-tilt head with the `head_action` to perform pre-recorded head motions such as nodding, shaking, and others generated by the user.

To combine the different expressive modalities along with the speech, the TTS module allows for markups embedded in the spoken utterances to synchronize the robot's expressiveness. For example, "`<set expression(embarrassed)> Sorry, <do motion(head_down)> I can't remember your name...`"

where `<set expression>` displays the given expression on the robot's face and `<do motion>` triggers a pre-recorded motion.

Robot Programming Kit (RPK)

The Robot Programming Kit (RPK) is a command-line tool that auto-generates a structured ROS 2 application skeleton, so developers can focus mainly on high-level logic implementation. It organizes the code into skills, tasks, and mission controllers. Besides the TIAGo Head, any other robot supporting ROS 2 can make use of this tool to create applications.

- Mission Controller: It is the high-level coordinator that orchestrates tasks to fulfill the robot's overall mission (for example, acting as a receptionist handling multiple users).
- Tasks: These represent discrete actions or action sequences the robot performs. For example: greet, read news, play a game.
- Skills: Fundamental capabilities that enable the robot to complete tasks. Examples include: `say`, `move_to`, `grab`, `pick_up`.

Public documentation on the Tiago Head[6] robot and tutorials[7] on how to get started are available online.

Mixed-Reality Interaction Simulator

The TIAGo Head mixed-reality interaction simulator (Fig. 5) emulates significant components of a typical ROS4HRI pipeline and semantic scene understanding [30]. It integrates standard ROS RQt plugins alongside custom modules to simulate interactions between the robot and both virtual and real humans and objects.

IV. DEVELOPING A TIAGO HEAD APPLICATION

In this scenario, our goal is to implement an airport receptionist application proof-of-concept on the TIAGo Head platform. At the time of writing, it was implemented and tested within the mixed-reality simulator. The robot should be able to provide information verbally to travelers, and also provide daily news when requested. The following sections outline the development workflow for this application.

A. Application description

We define two distinct tasks to cover receptionist duties:

- *News Reading* task: This task enables the robot to periodically send news updates to waiting passengers.

[6] https://docs.pal-robotics.com/sdk-dev/tiago-head
[7] https://ros4hri.github.io/ros4hri-tutorials/interactive-social-robots/

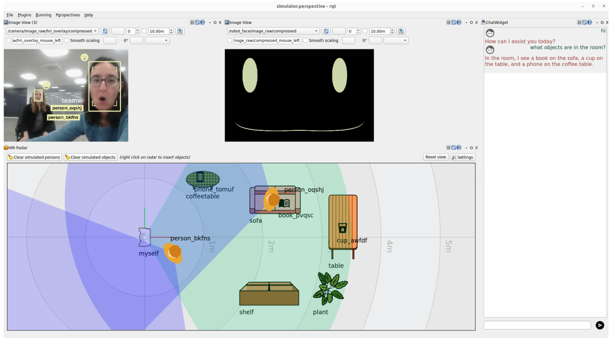

Fig. 5. TIAGo Head's mixed-reality interaction simulator

News of interest are requested to the LLM, and the robot presents a summary of the topic to the user, verbally, and through its touch screen. Moreover, it can also answer questions related to the news topic. For example, if the news topic is about Japanese food, the user can ask *"What is a tonkatsu?"* and the robot would retrieve information through the LLM. The robot also asks the user questions every few minutes, such as *"What did you think about the news?"* or *"Would you like to hear more about any of these stories?"*, with the goal of initiating a conversation. Once the user has scrolled down the news, they will select the "Done" button to complete the task.

- *Information provider* task: This task covers the robot's primary role as an information kiosk. When a user asks a question (e.g. *"Why was my flight delayed?"* or *"What time is the next flight to Paris?"*), the task uses the LLM to respond. For testing purposes of the use case, the LLM was prompted to pretend to have access to fake airport data (e.g., flight status, flight schedules, etc.). In a real scenario, the robot would be interfaced with an airport database to provide contextual data to the LLM. The task is completed when the user presses the "Done" button.

In both cases if no user input is detected after a given time, the system will restart again waiting for someone to initiate the interaction. The developer can configure different parameters for each task such as language, gender of the robot (male or female), and personality to adjust the voice as well as LLM prompt.

B. System Flow and Execution

Figure 6 illustrates the overall software architecture, based on the general software architecture of the robot defined in Figure 3.

When a user selects the desired task on the robot's tablet (News Reading, Information Provider) or by talking to the robot, the mission controller switches activity and invokes the appropriate skills to achieve such task. We next describe the main processes involved in the application:

- **User input detection:** in charge of detecting when the user initiates the interaction, either through the `vosk_asr` module when the user talks to the robot, for example, *"I want to read some news about weather"* or *"My flight was delayed, can you tell me why?"*; or through the `ui_server`, when the user selects an activity on the screen.
- **User intent detection:** If the input comes from speech, the `llm_bridge` –which is interfaced with OpenAI in this example– determines the *intent* of the user, i.e. what the user wants to do. Otherwise, if the user has used the touch-screen, the intent is automatically generated. Either case, the `intent` is published so the `misssion controller` can trigger the selected task.
- **Mission controller:** it captures the intent of the user and starts the respective task. In this case, either `news_reading` or `information_provider`. At the same time, it adjusts the prompt of the LLM with the required configurations. For instance, the type of LLM to use (OpenAI or Ollama), the personality of the LLM (e.g. "talk like a 18-year old teenager"), name and voice gender of the robot.
  The `llm_bridge` appends further instructions to the prompt, so the LLM returns not only the response, but also describes the expressive features, such as the facial expression or body gestures, the robot should show. Additionally, the LLM also provides basic sentiment analysis based on the user's transcribed speech. An example prompt template for the Information Provider task.

```
"You are a customer service robot named
{{robot_name}} for RoboPlane. Your job is
to provide accurate flight information
concisely. Your tone is defined by {{persona}},
and gender is {{gender}}. Provide JSON
responses with fields: response_to_user,
robot_body_expression, robot_facial_expression,
and user_text_expression. Only use expressions
from LIST_EXPRESSIONS and gestures from
LIST_BODY_GESTURES."
```

- **Execution of the tasks:** the respective task begins. In the case of `news_reading`, as per the tasks' prompt, the LLM will return the news to show. Then, it will use the `display_content` skill to show the content on the touch-screen. Every 2 minutes, the LLM proposes a new question to ask to the user. In the cases of `information_provider`, the flow is the same, but without displaying any content on the touch-screen.
  Either task uses the `communication_hub`'s `say` skill to provide the speech response from the LLM. For example:

```
"<set expression(neutral)> Here's the latest
travel news: <do motion(head_nod)> A major
storm in Europe is causing flight delays..."
```

If the LLM returns any facial expression and/or motion along with the textual response, the `say` skill activates the `expressive_eyes` or `play_motion` modules respectively. Moreover, the LLM also returns the user's sentiment analysis.

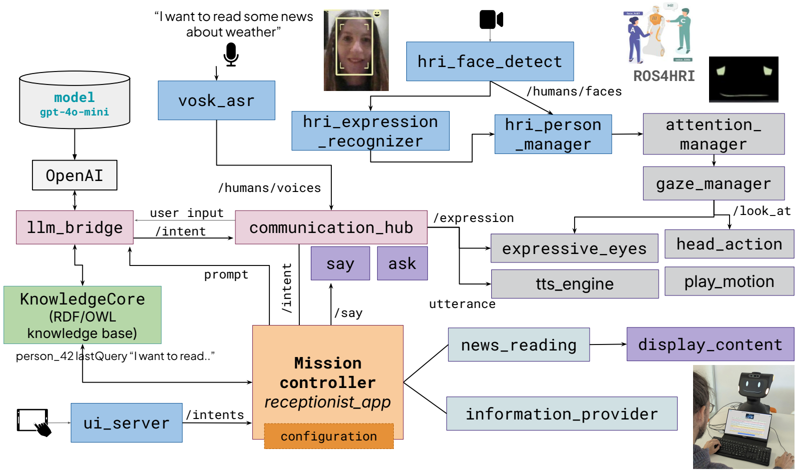

Fig. 6. Software architecture used for the receptionist application.

```
{
  "response_to_user": "Here's the latest travel news:
  ↪  A major storm in Europe is causing flight
  ↪  delays. Some flights have been postponed...",
  "robot_facial_expression": "neutral",
  "robot_body_expression": "head_nod",
  "user_sentiment_analysis": "neutral",
  "news_id": "news_2048"
}
```

- **Social perception:** every time the user input is captured, it records the user's facial expression with the `hri_expression_recognizer`. Moreover, during the whole interaction the robot is using its `attention_manager` to track the user both with its eyes and head checking for engagement.
- **Knowledge Base Update:** at the end of each interaction, the robot updates the ontology-based knowledge base (ORO) with user interaction data, to be used for future personalization. The stored knowledge varies by task. For instance, in the News Reading the following information is saved:
  "`person_42 hasFaceEmotion happy,`
  `person_42 hasTextEmotion neutral,`
  `person_42 lastQuery 'I want news',`
  `person_42 readNews 'news_1023',`
  `person_42 userValence '0.6',`
  `person_42 userArousal '0.4',`
  `person_42 userConfidence '0.9'` "

The configuration of the application can be adjusted by the mission controller as needed. In this use case example, we define the following parameters:

- **Language:** to set the language of the robot, that is, TTS, ASR, chatbot and GUI interface.
- **Persona**: to define the speech style e.g., it should talk like a child, an adult or a teenager.
- **Gender**: to switch between a male or female voice in the TTS.
- **Expressivity**: to indicate whether the robot should display facial expressions matching its speech, perform head motions (e.g. nodding, shaking) and to set the gaze behavior policy during interaction, that is, track user faces with its gaze, or not following gaze with the eyes at all, but with the head only.
- **LLM Model**: to specify the LLLM model to use.
- **Robot Name**: to set the robot name.

V. Discussion and Future work

Although TIAGo Head offers a robust platform for social HRI, several limitations remain. On-board AI is constrained by hardware, which limits the use of large neural models. Cloud-based LLMs mitigate this, but raise privacy and latency concerns. We address this with prompt design and intent filtering, though real-time content filtering and integration of knowledge graphs remain future goals.

From an ethical standpoint, the robot works entirely offline, so all audio and image data captured are processed onboard. However, future work will include the implementation of stronger privacy safeguards, such as real-time content filtering and consent management mechanism [31], [32].

At the time of writing, the receptionist scenario has been implemented and tested within the mixed reality simulator. Deployment on the physical TIAGo Head is planned as a future work. We acknowledge this as a limitation and aim to validate the system in real-world settings. A small-scale user study or pilot test will be conducted to assess engagement, usability, and satisfaction. This empirical validation will be critical for understanding the effectiveness of the platform and informing future iterations.

Finally, while this work does not propose novel algorithms, it contributes a flexible and expressive open source platform for social HRI research. Future work will explore adapting the platform to other domains such as education, elderly care or multilingual environments, and provide quantitative comparisons with similar robots to assess TIAGo Head's relative advantages.

VI. Conclusion

This paper presented TIAGo Head, a social tabletop robot platform that advances the state-of-the-art in human-robot interaction. We detailed its design and its ROS 2 based software, and how it combines its facial expressiveness, LLM integration with changing facial expressions, and perception, as well as several available tools such as its virtual interactive simulator, and RPK for application development. Thanks to its open software architecture, researchers and developers can easily integrate custom algorithms or connect TIAGo Head to other systems.

We also show an example application development where users can interact with the robot (currently in simulation) asking if it was an airport receptionist or checking out the latest news. The goal of this demonstration is to show the integration of its components, to be validated in future pilot studies.

In summary, TIAGo Head opens many avenues for research in social HRI. Its current design excels at many tasks, but continued development in perception intelligence, conversational depth, user personalization, and ethical AI integration will be key to fully realizing the potential of this social robot in everyday environments.

Acknowledgment

The work is funded by the Horizon Europe CoreSense (no. 101070254) and ARISE (no. 101135784) projects.

References

[1] C. Breazeal, "Social robots: Beyond tools to partners," *Proceedings of the IEEE*, vol. 91, no. 9, pp. 1670–1677, 2003.

[2] K. Dautenhahn, "Socially intelligent robots: Dimensions of human–robot interaction," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 362, no. 1480, pp. 679–704, 2007.

[3] T. Belpaeme, J. Kennedy, A. Ramachandran, B. Scassellati, and F. Tanaka, "Social robots for education: A review," *Science Robotics*, vol. 3, no. 21, pp. 1–17, 2018.

[4] B. Scassellati, H. Admoni, and M. J. Matarić, "Robots for use in autism research," *Annual Review of Biomedical Engineering*, vol. 14, pp. 275–294, 2012.

[5] P. Baxter, J. Kennedy, E. Senft, T. Lemaignan, and T. Belpaeme, "From characterising three years of hri to methodology and reporting recommendations," in *Proc. ACM/IEEE Int. Conf. Human-Robot Interaction (HRI)*. IEEE, 2016, pp. 391–398.

[6] A. Kaneko, Y. Matsumura, and S. Kuroyanagi, "Haru: An expressive tabletop social robot for emotional interaction," in *Proc. IEEE Int. Conf. Human-Robot Interaction (HRI)*, 2020, pp. 589–597.

[7] S. Al Moubayed, J. Beskow, G. Skantze, and B. Granström, "Furhat: A back-projected human-like robot head for multiparty human-machine interaction," in *Proc. IEEE Int. Conf. Robotics and Automation (ICRA)*, 2013, pp. 5598–5605.

[8] T. Belpaeme, P. Baxter, R. Read, R. Wood, H. Cuayáhuitl, S. Kiefer, and G. Athanasopoulos, "Multimodal child-robot interaction: Building social bonds," *Journal of Human-Robot Interaction*, vol. 3, no. 1, pp. 44–65, 2018.

[9] J. Dupont, S. Weber, and T. Müller, "Qtrobot—a humanoid social robot for autism intervention and education," in *Proceedings of the 2023 IEEE International Conference on Robotics and Automation (ICRA)*, 2023.

[10] M. Kanoh and Y. Hashimoto, "Emotional interaction and human attachment with lovot: A companion robot," in *International Conference on Social Robotics*. Springer, 2021.

[11] G. Milliez, "Buddy: A companion robot for the whole family," in *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, Chicago, IL, USA, 2018, pp. 1–2.

[12] P. Baxter, J. Kennedy, E. Senft, T. Lemaignan, and T. Belpaeme, "The role of social robots in tutoring and education: A review," in *International Journal of Social Robotics*, vol. 10, 2018, pp. 35–58.

[13] M. R. Tewari, B. Abbasi, and P. Subramanian, "Social robots as therapeutic tools for children with autism spectrum disorder," *Journal of Autism and Developmental Disorders*, vol. 52, pp. 1597–1615, 2022.

[14] L. Levinson *et al.*, "Living with haru4kids: Child and parent perceptions of a co-habitation robot for children," in *International Conference on Social Robotics*. Springer Nature Switzerland, 2022.

[15] R. Gomez *et al.*, "Design of embodied mediator haru for remote cross-cultural communication," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*, 2024.

[16] S. Cooper *et al.*, "Design of social features for robot-mediated cross-cultural interaction," in *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, 2024.

[17] C. Fischinger, R. Einramhof, M. Bajones, P. Wohlkinger, A. Papoutsakis, and A. Gisinger, "Social companion robots for elderly users: The case of elliq," *Gerontechnology Journal*, vol. 16, no. 3, pp. 211–225, 2016.

[18] Y. Mohamed and S. Lemaignan, "Ros for human-robot interaction," in *2021 IEEE/RSJ international conference on intelligent robots and systems (IROS)*. IEEE, 2021, pp. 3020–3027.

[19] H. Hu, "Robots in customer service: Enhancing public space engagement with ai-driven social interaction," *Frontiers in Robotics and AI*, vol. 8, 2021.

[20] B. Scassellati, H. Admoni, and M. Matarić, "Experimental methodologies in human-robot interaction research," *Annual Review of Biomedical Engineering*, vol. 14, pp. 275–294, 2012.

[21] T. Belpaeme, R. Read, R. Wood, H. Cuayáhuitl, S. Kiefer, and G. Athanasopoulos, "Cross-cultural responses to social robots: An hri study," *Journal of Human-Robot Interaction*, vol. 5, no. 2, pp. 44–73, 2019.

[22] S. Lemaignan and L. Ferrini, "Probabilistic fusion of persons' body features: The mr. potato algorithm," in *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, 2024.

[23] S. Lemaignan, S. Cooper, R. Ros, L. Ferrini, A. Andriella, and A. Irisarri, "Open-source natural language processing on the pal robotics ari social robot," in *Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*, 2023, pp. 907–908.

[24] OpenAI, "Chatgpt: Optimizing language models for dialogue," OpenAI Blog, Tech. Rep., 2022. [Online]. Available: https://openai.com/blog/chatgpt

[25] Ollama, "Ollama: Local large language model inference," Ollama.ai, Tech. Rep., 2023. [Online]. Available: https://ollama.ai

[26] D. Rao, A. Smith, and K. Johnson, "Evaluating large-scale language models for embodied ai applications," NeurIPS, Tech. Rep., 2023.

[27] OpenAI, "Gpt-4 technical report," arXiv preprint, Tech. Rep., 2023. [Online]. Available: https://arxiv.org/abs/2303.08774

[28] L. Ferrini, J. Palmieri, A. Marino, D. Lee, and S. Lemaignan, "remap: Spatially-grounded and queryable semantics for interactive robots," in *Social Robotics*. Singapore: Springer Nature Singapore, 2025, pp. 383–396.

[29] S. Lemaignan, R. Ros, L. Mösenlechner, R. Alami, and M. Beetz, "Oro, a knowledge management platform for cognitive architectures in robotics," in *2010 IEEE/RSJ International conference on intelligent robots and systems*. IEEE, 2010, pp. 3548–3553.

[30] S. Cooper, R. Ros, S. Lemaignan, F. Gebellí, L. Ferrini, and L. Juričić, "Demonstration of an open-source ros 2 framework and simulator for situated interactive social robots," in *2025 20th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2025, pp. 1770–1772.

[31] E. Bender, T. Gebru, A. McMillan-Major, and S. Shmitchell, "On the dangers of stochastic parrots: Can language models be too big?" in *Proc. ACM Conf. Fairness, Accountability, and Transparency (FAccT)*, 2021, pp. 610–623.

[32] B. Mittelstadt, "Principles of ai ethics: An overview of bias and accountability in ai models," *AI & Society*, vol. 35, no. 2, pp. 121–140, 2020.