Edward Curry · Philip Piatkiewicz
Fredrik Heintz · Heike Vornhagen
Ahmed Nabil Belbachir · Emanuela Girardi
Marc Schoenauer · Juha Röning *Editors*

Artificial Intelligence, Data and Robotics

Foundations, Transformations and Future Directions

The AI Data Robotics Association

OPEN ACCESS

Springer

Artificial Intelligence, Data and Robotics

Edward Curry • Philip Piatkiewicz
Fredrik Heintz • Heike Vornhagen
Ahmed Nabil Belbachir • Emanuela Girardi
Marc Schoenauer • Juha Röning
Editors

Artificial Intelligence, Data and Robotics

Foundations, Transformations and Future Directions

Springer

Editors
Edward Curry
Insight Research Ireland Centre for Data Analytics, Data Science Institute
University of Galway
Galway, Ireland

Fredrik Heintz
Linköping University
Linköping, Sweden

Ahmed Nabil Belbachir
Norwegian Research Centre (NORCE)
Bergen, Norway

Marc Schoenauer
INRIA (National Institute for Research in Digital Science and Technology)
Rocquencourt, France

Philip Piatkiewicz
The AI, Data and Robotics Association (ADRA)
Brussels, Belgium

Heike Vornhagen
Insight Research Ireland Centre for Data Analytics, Data Science Institute
University of Galway
Galway, Ireland

Emanuela Girardi
The AI, Data and Robotics Association (ADRA)
Brussels, Belgium

Juha Röning
University of Oulu
Oulu, Finland

ISBN 978-3-032-10560-8 ISBN 978-3-032-10561-5 (eBook)
https://doi.org/10.1007/978-3-032-10561-5

This work was supported by the University of Galway.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

If disposing of this product, please recycle the paper.

A Companion Robot Platform for Exploring Technical and Ethical Aspects in Elderly Care

Lorenzo Boi, Silvia M. Massa, Diego Reforgiato Recupero, Daniele Riboni, Rubén Alonso, Michele Cardinali, Elena Ricci, and Alberto Pirni

Abstract Large language models (LLMs) are driving significant advancements across various sectors. Combined with robotics, they can lay the foundation for a new paradigm in healthcare, particularly in elderly care and companionship. For years, research has focused on the concept of carebots and robot companions. The latest capabilities of LLMs and robotics promise to address many challenges regard-

L. Boi · S. M. Massa · D. Riboni
Department of Mathematics and Computer Science, University of Cagliari, Cagliari, Italy
e-mail: lorenzo.boi@unica.it; silviam.massa@unica.it; riboni@unica.it

D. Reforgiato Recupero (✉)
Department of Mathematics and Computer Science, University of Cagliari, Cagliari, Italy

ICT and Robotics, R2M Solution s.r.l., Pavia, Italy
e-mail: diego.reforgiato@unica.it

R. Alonso
ICT and Robotics, R2M Solution s.r.l., Pavia, Italy
e-mail: ruben.alonso@r2msolution.com

M. Cardinali
Institute of Law, Politics and Development, Sant'Anna School of Advanced Studies, Pisa, Italy

Department of Human Studies, University of Macerata, Macerata, Italy

Ludes Campus, Lugano, Switzerland
e-mail: michele.cardinali@santannapisa.it

E. Ricci
Institute of Law, Politics and Development, Sant'Anna School of Advanced Studies, Pisa, Italy

Department of Human Studies, European University of Rome, Rome, Italy
e-mail: elena.ricci@santannapisa.it

A. Pirni
Institute of Law, Politics and Development, Sant'Anna School of Advanced Studies, Pisa, Italy
e-mail: alberto.pirni@santannapisa.it

E. Curry et al. (eds.), *Artificial Intelligence, Data and Robotics*,
https://doi.org/10.1007/978-3-032-10561-5_17

ing interaction and personalized experience faced by earlier generations of companion robots. However, the potential of these emerging technologies is accompanied by significant ethical, technical, and social challenges. This chapter presents the design of an elderly care robot that exploits the capabilities of the latest generation of LLMs and considers the ethical, technical, and social implications for improving human-robot interaction capabilities. We report the technical design of a novel robotic platform for elderly care, detailing the devised interaction scenarios, a prototype implementation with NAO and PEPPER robots, and the results of a technical validation that shows the efficiency and feasibility of our system. This study serves as a guide to identify future needs, explore the impact of companion robots on older adults, and understand acceptance of these new approaches.

Keywords Large language models · Artificial intelligence · Companion robots · Natural language processing · Trust · Roboethics

1 Introduction

LLMs have revolutionized the technological landscape, driving significant advancements across various sectors, including artificial intelligence (AI) and robotics. Built upon Transformers [1], an advanced neural network architecture, these models exhibit an unprecedented ability to interpret and generate text, surpassing the limitations of traditional natural language processing (NLP) systems [2]. Due to their ability to learn from large amounts of data, LLMs can perform numerous language tasks, such as automatic translation [3], content synthesis [4], answering questions [5], and code generation [6]. However, their impact is not limited to linguistic processing. Their integration with robotic systems has opened up new opportunities for developing autonomous and social robots, significantly enhancing the interaction between humans and machines [7].

The synergy between LLMs and robotics is transforming the relationship between technology and humans, promoting more personalized, accessible, and efficient healthcare [8]. Robots equipped with advanced language capabilities can engage with patients naturally, gathering symptom information, answering questions, and supporting healthcare professionals during diagnosis and treatment [9]. These advancements enhance medical resource utilization and offer a more personalized and reassuring experience [10].

In the past, many conversational agents in healthcare employed finite-state or frame-based dialogue management strategies, limiting the ability to have natural, adaptive conversations [11]. These limitations hindered customization and the ability to respond to the individual needs of users, particularly older adults. The advent of LLMs offers a promising avenue to overcome these weaknesses. However, the use of LLMs in companion robots for older adults is not without challenges. Irfan et al. [12] highlight several problems encountered when integrating LLMs into conversational robots. These include frequent interruptions in conversations, slow or repetitive responses, inconsistent interactions, language barriers, hallucinations, and out-of-date information. These problems can cause frustration, confusion, and worry in the elderly, potentially impeding the effectiveness of robotic companions.

By leveraging the advanced features of the latest versions of LLMs and robotics, a variety of innovative use cases have emerged across multiple domains. Emerging key applications include:

- Interaction and preliminary triage: Robots can collect details about patients' symptoms, record essential information, and respond promptly to common questions. This support lightens the burden on healthcare staff, improving efficiency during the initial stages of diagnosis [13].
- Elderly assistance: Humanoid robots equipped with LLMs monitor physical health, remind patients to take medications, provide emotional support, and send alerts in an emergency. These functionalities promise to improve the quality of life for the elderly, ensuring greater safety and autonomy [14].
- Rehabilitation: Specialized robots guide patients during rehabilitation exercises, providing personalized feedback to improve treatment effectiveness and engagement [15].
- Psychological support and autism therapy: Robots equipped with advanced language capabilities can provide assistance to individuals with psychological or cognitive needs by engaging in seemingly empathetic conversations. In particular, they can play a significant role in helping children with autism develop social skills [16].
- Feedback collection and monitoring: Robots equipped with LLMs can interview patients, collecting detailed data on their physical and mental state to optimize future treatments [17].

Companion robots and carebots represent a rapidly growing sector, designed especially to support and provide companionship to older adults [18].

The former provides emotional support, social interaction, and companionship; the latter assists with physical care tasks, medical monitoring, and well-being management. With their advanced language comprehension, these types of robots respond to users' emotional and daily needs, reducing feelings of isolation and promoting greater autonomy. Features such as empathetic conversations, personalized reminders, and daily health monitoring make them essential tools for improving quality of life in home settings. However, despite the progress, using LLMs in the healthcare sector raises important safety and ethical concerns. The sensitive data processed and the vulnerability of LLMs to manipulations or errors could result in undesirable behavior or misleading information. Therefore, it is essential to develop rigorous verification protocols and alignment strategies to ensure the safe and ethical use of these technologies [19, 20].

A major challenge in using companion robots for elderly care is ensuring their successful adoption. Indeed, older adults are often unfamiliar with new technologies, and their seamless interaction with robots may be disrupted by skepticism or cognitive and physical barriers. While encouraging trust and emotional bonds with robots is crucial for engagement, these connections carry significant risks. Indeed, over time, robotic companions may replace (rather than complement) human interaction, increasing social isolation. Additionally, excessive reliance on robots for emotional support and daily tasks may determine dependency, limit autonomy, and reduce opportunities for engagement with family, caregivers, and friends. In order to mitigate these risks, designs must look for an appropriate balance between technological assistance and human-centered care. The most appropriate designs are those that promote healthy technology habits, foster connections with real people, and support group and social activities.

The Italian project TRI-TECH ("TRust in Technology: How to Assess and Improve RoboT-User Interaction in Elderly Care Integrating EtHical, Technical and Social Variables") aims to address these challenges by adopting a multidisciplinary approach. In particular, the project aims to explore the impact of companion robots on the lives of elderly individuals, with a particular focus on assessing the ethical implications through various indicators. Various tailored functionalities are implemented, which aim at fully understanding the participants' requests and demonstrating empathy in every stage of the human-robot interaction. The goal is to break down traditional human-robot barriers, creating a more natural and human-like interaction that fosters smooth and reassuring communication, ultimately improving the overall user experience.

This chapter aligns with the core objectives of the AI, Data, and Robotics Partnership [21] by introducing a novel robotic platform designed to enable user-friendly interaction between robots and elderly individuals, leveraging LLMs and carefully designed interaction flows to promote human-centric and accessible healthcare solutions. Starting from a critical review of current robotic platforms and related open challenges, both on the technical and ethical side, we designed a system architecture that leverages innovative AI tools to stimulate the interaction between the older adult and the robot. In order to evaluate the usability and ethical considerations in the usage of robotic systems by elderly people, we designed different human-robot interaction scenarios, which were implemented using the PEPPER and NAO humanoid robots. We also conducted a technical validation of the system, which showed the efficiency and feasibility of our proposed solution. We also conducted a technical validation of the system, which showed the efficiency and feasibility of our proposed solution.

The main contributions of this work are the following:

1. We critically review and discuss the state of the art in human-robot interaction, considering practical, technical, and ethical challenges.
2. We introduce the architecture of a novel robotic platform specifically designed to explore the technical and ethical implications of human-robot interaction in elderly care.
3. We explain how we designed and implemented the interaction flow between the robot and the senior in different use case scenarios.
4. We present two prototype implementations using NAO and PEPPER robots, together with the results of a technical validation of our system.

The rest of this chapter is organized as follows: Sect. 2 reviews the related work about human-robot interaction. Section 3 highlights the key challenges related to the definition of a companion robot for personal assistance addressed to elderly people. Section 4 presents our system's architecture and implementation. Section 5 focuses on practical interaction and its implications. Section 6 discusses ethical issues and their impact. Future research directions are outlined in Sect. 7, where opportunities for further exploration are emphasized. Finally, Sect. 8 concludes the chapter and summarizes the main implications.

2 Related Works

In recent years, robotic systems have significantly advanced in their intelligence and ability to engage in natural interactions with humans [22]. The desired result would be a humanoid robot that communicates seamlessly, recognizes objects, understands actions, and adapts its responses to the user's needs. These capabilities have been significantly enhanced by integrating ontologies with NLP techniques [23–25]. For instance, the NAO robot was able to discuss various topics and execute commands through the work of different authors [26–28]. Furthermore, Fukuda's framework allowed robots to identify objects through verbal interaction, fostering a communication interface that is more intuitive and that resembles human interaction [29]. The idea that the interaction should be similar to human-to-human interaction is one of the primary goals of multimodal interaction [30]. In fact, the field of multimodal interaction is grounded in the understanding that human communication is inherently multimodal. Principles of multimodality not only facilitate fluidity and engagement but also help reduce errors in interaction, adapt to potential user limitations, and allow users to choose their preferred mode of interaction at any given moment. Another relevant work in the field of human-robot interaction was carried out by Markievicz, which used supervised machine learning to interpret natural language commands, allowing actions to be classified and voice instructions to be converted into executable tasks [31].

AI and robotics are revolutionizing their field and making their way into healthcare and elderly companionship. Research projects like Triage-Bot [32] illustrate how automation and AI can optimize healthcare delivery. Developed by Sharma et al., Triage-Bot leverages AI-driven tools like facial recognition, automatic speech recognition, and photoplethysmography to collect and process patient data. Integrated with Electronic Medical Records, it automatically classifies patients according to the Canadian Triage and Acuity Scale (CTAS), enhancing efficiency and access to care.

Moreover, in the field of social robotics, several projects have aimed to develop robotic companions capable of long-term interaction and emotional engagement with humans. The FLASH robot,[1] designed within the European LIREC project,[2] focused on integrating emotional expression, behavioral consistency, and modular design to enhance human-robot interaction [33]. The MARIO project[3] focused on active and healthy aging, designing a robot aimed at exploring its use as companionship and to reduce loneliness and social isolation in dementia patients [34]. These works demonstrated the potential for robotic companions to establish meaningful and enduring relationships with users by combining advanced hardware with affective control systems. Arunachalam et al., in 2024, presented a work that proposed an AI-driven human-robot interaction system using recurrent neural network (RNN) models to provide adaptive assistance and intelligent companionship for the elderly [32]. The significance of this work lies in enhancing the accuracy and responsiveness of robotic caregivers in geriatric settings by leveraging high-quality and diversified datasets. This technology aims to manage medication, monitor health, and improve the autonomy and quality of life of the elderly by integrating IoT and AI.

3 Technical and Design Challenges

Designing a companion robot for personal assistance for elderly people presents a range of technical and design challenges. In addition to meeting complex functional requirements, the robot must ensure a positive, empathetic, and personalized user experience. An impactful aspect, yet challenging, is the integration with an LLM, since it enhances NLP capabilities. However, the success of the robot also depends on its ability to consider and enhance user perception, which influences key factors such as acceptance, trust, and engagement. In the following, we illustrate how the robot integrates advanced technologies, such as LLMs, to provide real-time, empathetic, and personalized interactions. We consider challenges like optimizing computational efficiency and ensuring scalability through modular design. Finally, we explore strategies to enable multimodal communication for natural interactions and foster trust and acceptance through empathy and user feedback and adaptability to evolving needs.

In designing a carebot or a companion robot to assist elderly people, it is necessary to ensure that it provides practical support but also emotional engagement and adaptability. The key factors that should be considered during the development include:

- Execution of various functions: the robot must handle a wide range of tasks, from daily practical assistance (e.g., medication reminders) to entertainment features, such as simple games (e.g., riddles or "guess the word"), storytelling, and the possibility to play songs or videos. Moreover, the robot must include func-

[1] https://spectrum.ieee.org/strange-polish-robot-

[2] https://cordis.europa.eu/project/id/215554/it

[3] http://www.mario-project.eu/portal/

tionalities that encourage the elderly person to engage in beneficial activities, such as stretching routines or mindfulness exercises. Integrating an LLM allows for contextual and personalized understanding and response to the user's needs, balancing complexity and ease of use.

- Speed of execution and fluidity of interaction: a quick, natural, and efficient interaction is crucial for conveying empathy and responsiveness, key aspects for building trust and improving the user experience. The LLM represents the core of the robot's linguistic intelligence, but its processing requires significant computational resources. Optimizing the model's inference pipeline is essential to reduce latency and response times while ensuring smooth and scalable performance. The user perception is closely tied to these aspects; indeed, a system that responds quickly and accurately is perceived as more reliable, enhancing the user's sense of security and satisfaction.
- Scalability and integration of new features: to keep the robot resilient over time, a modular design is needed that allows for the addition of new functionalities and system customization for specific contexts. The ability of the robot to evolve according to the needs of the elderly person reinforces the sense of usefulness and engagement, fostering a long-term emotional bond and improving the user's perception.
- Multimodality of the interaction and adaptation to context and user's sensations: the robot must interpret the context and respond appropriately to the user's emotions. Multimodality enables a natural interaction while allowing the robot to supplement the information received through one mode with data from others (e.g., better understanding voice commands through gestural information) or to better understand the context of the interaction. The integration of the LLM with sensors (e.g., cameras or microphones) enables the combination of linguistic and physical cues for a more complete understanding. This adaptation not only improves the quality of interactions but also reinforces the perception of the robot as an empathetic and attentive companion, capable of responding to the elderly person's emotional needs.
- Demonstration of empathy: the robot's ability to demonstrate empathy is central to building an emotional bond with the user. Thanks to the LLM, the robot can generate responses that simulate an empathetic attitude. However, this function must be carefully designed to avoid biases or inappropriate responses, ensuring that the tone remains respectful and consistent with the user's expectations. Perceived empathy is what transforms the robot from a mere technological assistant into a life companion. Also, in relation to multimodality, multimodal outputs can be provided, offering more information through gestures, voice, or visual effects that allow the robot to seem more natural and facilitate communication.
- User feedback collection: collecting feedback is crucial for improving the robot's functionalities and personalizing its responses. However, the chosen method must respect the elderly person's cognitive and physical capabilities, avoiding intrusive interactions. Well-designed feedback mechanisms also enhance the perception of the robot as a system attentive to the user's needs, increasing trust and acceptance. Additionally, the robot provides empathetic feedback based on

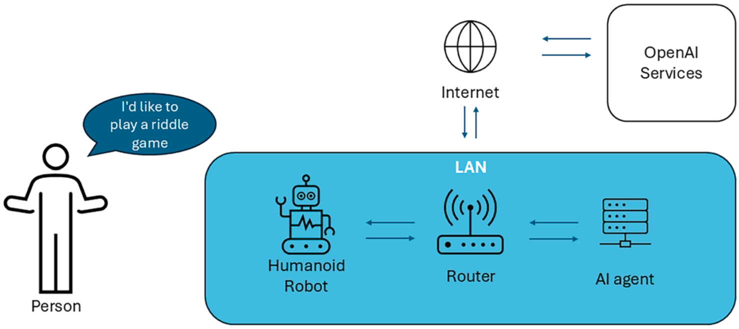

Fig. 1 System architecture

context, using personalized movements that combine body language, eye color, and voice tone. This approach makes interactions even more natural, human, and responsive, strengthening the emotional bond between the robot and the elderly person.

4 System Architecture

The architecture of our proposed human-robot interaction system is based on the employment of LLMs and programmable robots. This configuration allows robots to understand and respond to complex commands, both physical and cognitive, in a natural and intuitive manner. The architecture is illustrated in Fig. 1.

In this section, we will describe its main components and their functioning.

4.1 Main Components

The system architecture is composed of four fundamental elements:

- Humanoid robots: For this study, we selected the NAO[4] and PEPPER[5] humanoid robots. Both robots are equipped with a variety of sensors, including gyroscopes, microphones, cameras, and tactile sensors, enabling multimodality and an advanced perception of their surroundings. Additionally, both robots are fully programmable and capable of performing complex movements, making them ideal for the intended purpose. The main difference between the two lies in their physical and functional configuration. PEPPER lacks legs but moves using

[4] https://corporate-internal-prod.aldebaran.com/en/nao

[5] https://corporate-internal-prod.aldebaran.com/en/pepper

wheels and features an integrated touchscreen tablet. In contrast, NAO is equipped with legs and walks but does not include a tablet.

- Language model: We conducted preliminary evaluations to determine the most suitable LLM for our system's specific requirements. In our informal benchmarking, we considered several state-of-the-art models, including GPT-4 (OpenAI), Gemini (Google), LLaMA (Meta), and Mistral. Our primary selection criteria included support for the Italian language, fluency in generating natural and contextually appropriate responses, fast response times, and ease of integration via APIs. Based on these factors, GPT-4 demonstrated the most reliable performance, particularly in handling Italian inputs and delivering coherent, user-friendly replies. Hence, we adopt OpenAI's GPT-4 language model [35] to process user requests and generate appropriate responses or actions. GPT-4 can understand natural language and respond to general questions or interpret specific commands, adapting to different linguistic variants and contexts. Importantly, the system's architecture is model-agnostic by design: the underlying LLM can be replaced with minimal effort by any other model that supports OpenAI-compatible APIs, such as Gemini via a wrapper, or even local models via Ollama. For well-known alternatives that do not natively support this interface, only minor code adjustments are required, typically limited to a few lines to adapt the API endpoints and response parsing logic.
- AI agent: The AI agent is designed to leverage the advanced capabilities of an LLM model, ensuring efficient and seamless management of interactions between the user and the robot. The architecture consists of the following main modules:

  1. Functionality request recognition module: It analyzes user requests to identify and activate the most suitable functionalities available in the system.
  2. Robot gesture selection module: It processes contextual data to select and execute the most relevant movements according to the specific situation.
  3. User data tracking module: It manages user-related information, including preferences, interaction history, and details helpful in personalizing the experience.
  4. Log tracking module: It records and monitors all system activities, including user-robot interactions, providing valuable data for debugging, optimization, and performance analysis.

4.2 Detailed Architecture and Data Flow

In this section, we provide a detailed description of the data flow within the system architecture:

1. Reception of voice input: The interaction begins with a voice instruction provided by the user to the humanoid robot (e.g., "I would like to play a guessing

game”). The audio signal is captured through microphones integrated into the robot (NAO or Pepper), as depicted in Fig. 2.
2. Audio transmission and textual transcription: After receiving the audio signal, the robot transmits it over the local network (LAN) for transcription. The specific transmission flow depends on the robot model used:

   (a) NAO robot: The robot sends the audio signal to the AI agent server, which forwards it to the Speech-to-Text (STT) transcription service utilizing OpenAI’s Whisper model, as represented in Fig. 3.
   (b) Pepper robot: The Pepper robot sends the audio signal directly to the STT service, receives the transcribed text, and subsequently forward it to the AI agent, as illustrated in Fig. 4.

3. Intent classification and response generation: Upon receiving the textual transcription of the voice input, the AI agent employs a LLM, such as OpenAI’s GPT-4, to perform a detailed semantic classification of the user’s intent. The LLM analyzes the received text by comparing it with a predefined set of available functionalities (e.g., play games, listen to music, perform exercises, answer general questions) and selects the functionality that best matches the user’s request as illustrated in Fig. 5.

The response generation can follow two distinct modalities:

1. Static response (predefined): The AI agent selects a pre-existing, hardcoded response stored in the robot. The most appropriate response is selected based on the recognized intent, as depicted in Fig. 6.
2. Dynamic response: When a personalized and original response is required, the AI agent utilizes LLM, such as OpenAI’s GPT, to generate content specifically tailored to the interaction context, as illustrated in Fig. 7.
3. Transmission of response to the robot and user interaction: The generated response, whether static or dynamic, is transmitted back to the humanoid robot, which converts it into vocal output using an integrated text-to-speech (TTS) system. Verbal communication is accompanied by appropriate body movements from the robot to enhance interaction and natural communication with the user.

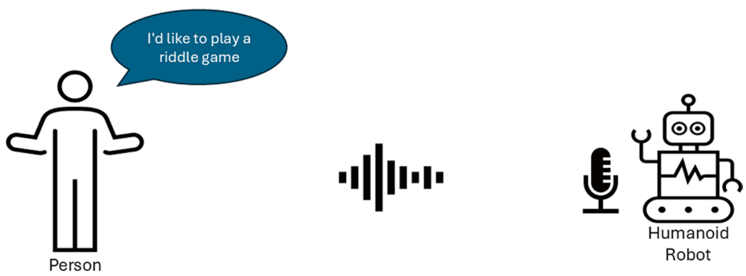

Fig. 2 Reception of voice input

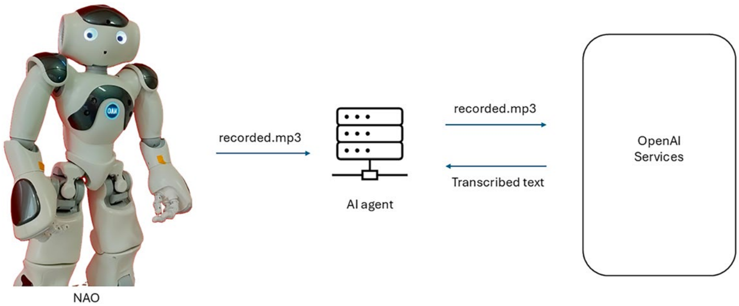

Fig. 3 Transcription of vocal input on NAO

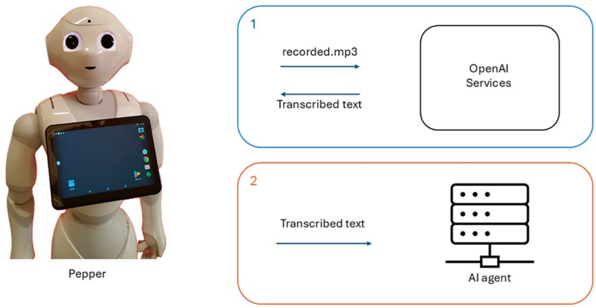

Fig. 4 Transmission of the transcribed text to the AI agent

Fig. 5 Intent classification

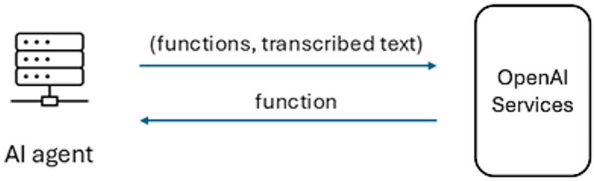

Fig. 6 Sending a static response to the robot

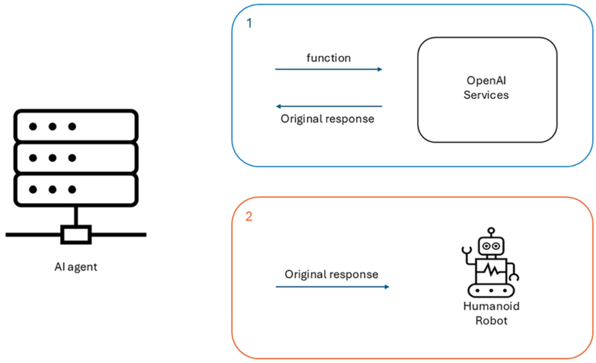

Fig. 7 Generating and sending a static response to the robot

4.3 Integration and Communication

The communication between the various components of the system occurs through a local area network (LAN), where the robots interact with a central unit responsible for managing the core functions and integrating with GPT-4 and the AI agent. The exchange of requests and responses is implemented using the REST architecture over the HTTP protocol, ensuring efficient data transmission. This approach promotes smooth and natural robot-user interaction while significantly reducing perceived conversation latency.

4.4 Technical Evaluation

We tested our robotic infrastructure by measuring the response times of its various modules to evaluate the system's efficiency and responsiveness. The most computationally expensive modules of the system are the following:

- Speech-to-text, for transcribing the words spoken by the elderly
- Text generation, used in cases where it is necessary to produce a response to the elderly
- Input classification, that is, the analysis of the speech-to-text output to understand the elderly's intentions

To evaluate the system's performance, the response times of each of these components were measured during real-world interaction with the robot. The collected data includes repeated measurements over 20 sessions of interaction. Table 1 reports the average and standard deviation of the recorded times for each module. Results show that the execution times are feasible for our application scenarios. The most expensive task is text generation, which may take considerable time for particular

Table 1 Response times of system components

	μ (s)	σ (s)
Input classification	1.43	1.10
Speech to text	1.47	0.47
Text generation	4.58	1.50

tasks, especially when the LLM is asked to invent a story based on a theme chosen by the elderly.

Additionally, we conducted preliminary tests involving colleagues and individuals external to the project. This approach allowed us to gather unbiased feedback, free from implementation-related influence, which helped identify potential issues and improve the overall interaction with the robot.

4.5 Advantages of the Proposed Architecture

The proposed architecture offers several advantages:

- Precise natural language processing and classification: Integration with an LLM enables the system to extract relevant information from user inputs typically expressed in non-technical language, ensuring effective system operation.
- Scalability and adaptability: Due to its modular design, the system allows for easy integration of new features and modifications to the AI agent's parameters, facilitating future adjustments to the interaction and enabling continuous system evolution.
- Multimodality: Given the capabilities of robots and the flexibility of LLMs, the architecture supports the customization of different forms of user communication. Multimodality also streamlines the robot's ability to respond in a more versatile and engaging way.
- Relationality: The architecture is designed to foster meaningful human-robot relationships, prioritizing the development of connections that address emotional and social needs of elderly persons. By promoting interactions that prevent the isolation often experienced in caregiving contexts, robots would enhance psychological and relational well-being. This focus on relational engagement ensures that the technology transcends its role as a tool, becoming a trusted companion that integrates seamlessly into the social fabric, ultimately improving the quality of life for individuals in need of care.
- Stability of ethical framework: The proposed architecture consolidates a robust ethical framework, with trust as its cornerstone value. Designed with context-related usage in mind, it ensures that every deployment of the system is preceded by careful consideration of the situational needs and potential impacts. By embedding trust as a central tenet, it supports ethical decision-making while fostering reliable and meaningful interactions, establishing a standard for responsible and inclusive innovation. This framework embodies a twofold role. On the

one hand, it can be considered as a foundational-orientative guide for the work of robotic engineers, shaping the programming process to align with ethical principles. On the other, it offers the opportunity for stabilizing a common ground of inter-disciplinary cooperation and fostering trans-disciplinary achievements, from both a technological and social point of view.

5 User Interaction

The system operates through four main phases to ensure smooth and personalized interaction between the user and robots. Each phase is accompanied by complete tracking of conversations through the Log Tracking Module, which records and monitors every exchange of information between the robot and the user. The four main phases that ensure the system's proper functioning are discussed below.

5.1 User Basic Information

The first phase involves the robot acquiring personalized information through a series of questions posed to the user. These questions primarily concern the user's name, the name the user wishes to give to the robot, and their preference regarding the formality of the conversation (more or less formal). The collected parameters are stored in the User Data Tracking Module, allowing the system to adapt to the user's needs. Based on these preferences, the robot will begin addressing the user by name and use appropriate language, creating a more natural and engaging experience. Figure 8 illustrates the flowchart of the User Basic Information Configuration process, highlighting the sequential steps from posing questions to the user to adapting the robot's behavior based on the collected preferences.

5.2 Actual Interaction

Once the information are set, the system begins actively interacting with the user, inviting them to perform various activities. The Functionality Request Recognition Module analyzes the user's requests and activates the most suitable features. Specifically, the robot can:

- Play songs or videos (only in the case of PEPPER)
- Propose simple games, such as riddles or "guess the word," allowing the user to test the robot
- Tell themed stories
- Guide the user through basic physical exercise routines or mindfulness sessions

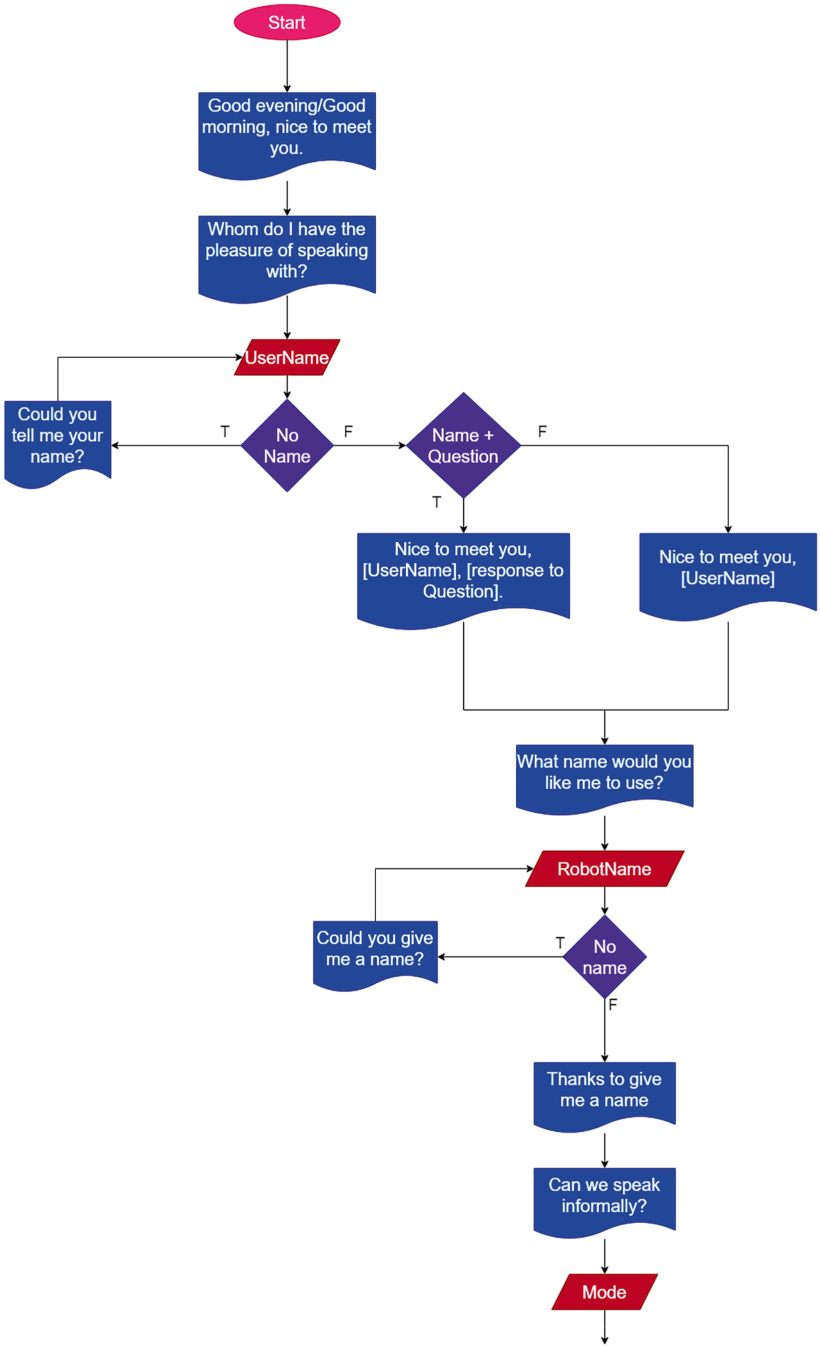

Fig. 8 User basic information configuration flowchart

- Respond to general questions or comments from the user (e.g., “Who was Napoleon?” or “Tell me a joke”)

During each activity, the robot will ask the user whether they enjoyed the experience, fostering engagement and empathy. In Fig. 9, the feedback request can be observed when the user asks the robot to tell a story. After the story is told, the question “Did you enjoy the story?” is asked.

Furthermore, the robot will react emotionally and gesturally based on the context: for instance, if the user expresses sadness, the robot will respond with appropriate gestures, thanks to the Robot Gesture Classification Module.

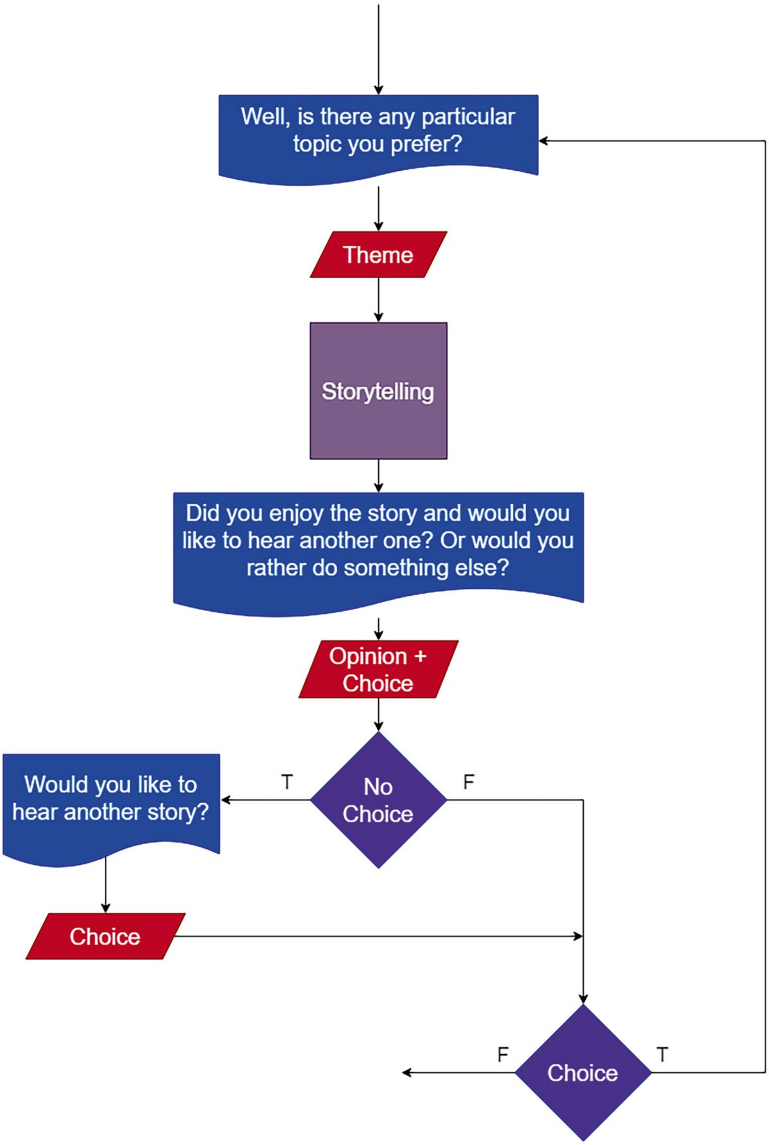

Fig. 9 Storytelling flowchart

5.3 Event Reminders

The system is designed to track important events and remind the user about them, encouraging the elderly user to perform activities necessary for their well-being. For example, periodically, the robot will remind the user to drink water or do some physical exercise, ensuring that the user follows daily routines and does not forget essential tasks.

5.4 Continued Interaction

After the initial interaction phase, the robot will continue encouraging the user to engage in activities, maintaining active involvement. If the user shows no interest or provides no feedback to continue, the robot will wait for new requests. The user can stop using the robot anytime, interrupting the interaction. The system is designed to be flexible and respect the user's wishes, ensuring a personalized and non-intrusive experience. This modular and interactive approach allows the system to dynamically respond to the user's needs, enhancing the overall interaction experience and fostering continuous, respectful, and engaging communication.

5.5 System at Work

The TRI-TECH project aims to evaluate and enhance the interaction between elderly individuals and companion robots, focusing on key aspects such as trust, acceptance, and perceived utility of these technologies. The proposed architecture will be employed in an experimental setting, involving 1-h usage sessions with groups of elderly participants recruited from care facilities and healthcare environments. Recruited seniors will interact with two robots, shown in Fig. 10: NAO and PEPPER. A video showcasing the Pepper robot interacting with a user, narrating a story, and utilizing the mentioned architecture is available at https://youtu.be/kaR-vO7YFEGs. A video showcasing NAO proposing a set of physical exercises in available at https://youtube.com/shorts/rcd99D44Wpo.

Participants will follow a structured process: initially, they will complete a preliminary questionnaire to assess their expectations and predispositions. Subsequently, after their experience with the robot, they will fill out a second questionnaire to provide detailed observations and feedback on their perceptions and satisfaction levels. During the sessions, qualitative and quantitative data on the interaction between participants and the robot will be collected. These data, derived from textual input provided by users, will include patterns of verbal communication, emotional expressions, and the level of engagement evident in their responses. The methodological approach also involves simulating real-life scenarios, such as

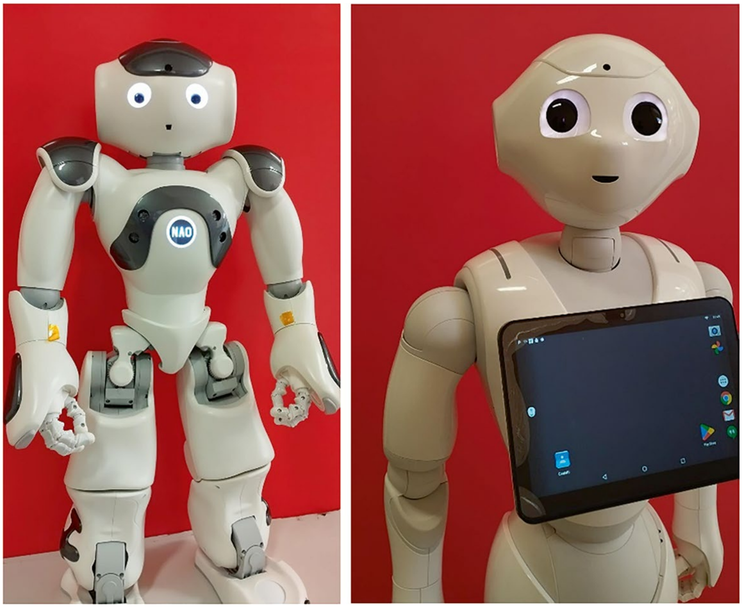

Fig. 10 The NAO and PEPPER robots used in our prototypes

companionship and health monitoring, to ensure that the analyzed interactions reflect realistic and meaningful situations for the target demographic. This element ensures that the findings can be applied to practical and relevant contexts in the daily lives of elderly individuals. The integration of ethical, technical, and social variables is essential for a comprehensive analysis of the impact of robots on the quality of life of the elderly. The collected data will be used to investigate critical aspects such as trust in the robot, the degree of acceptance, and the effectiveness of interaction modalities in promoting user well-being and comfort. The results of this study will guide the optimization of the design, functionality, and communication methods of companion robots, fostering their large-scale adoption. Ultimately, the project aims to improve elderly care by providing reliable, useful technologies that seamlessly integrate into their daily lives, thereby contributing to significant advancements in care and social inclusion.

6 Ethical Issues

In the sections above, we examined the technical aspects involved in the project; we now turn our attention to the ethical implications and concerns. The rapid advancement of digital technologies has resulted in the increasing utilization of intelligent systems and autonomous machines across diverse domains, encompassing healthcare, industry, education, and transportation. Within these contexts, trust has been identified in the scientific literature as a crucial factor [36] significantly influencing the acceptance, implementation, and efficacy of robots and computerized systems in general [37].

6.1 Mapping Trust in Human-Robot Interaction

Generally speaking, trust constitutes a cornerstone of interpersonal relationships, encompassing both human-to-human and human-to-robot interactions, particularly when these interactions involve elements of reciprocity (where applicable) or shared experiences. The presence of trust exerts a significant influence on the quality of relationships and reveals a complex dynamic: as trust increases, relationships tend to become more robust, authentic, and resilient to dialectical tensions in a generative manner, facilitating the resolution of conflicts without descending into exclusionary or adversarial dynamics. Conversely, in the absence of trust, relationships often assume a defensive profile characterized by caution or suspicion, frequently necessitating precautionary or contractual measures to mitigate perceived risks. These considerations assume even greater significance in interactions with artificial intelligence systems or robotic entities [38].

Fostering trust can become particularly critical when it comes to robots because, unlike human beings, robots lack the innate social and emotional cues that naturally elicit trust. Humans are predisposed to trust other human beings due to shared experiences, emotional understanding, and the capacity for reciprocal empathy [39]. These qualities are deeply rooted in our evolutionary and cultural frameworks and make human-to-human trust more intuitive and accessible.

In contrast, robots often lack these relatable attributes. They do not share our biological makeup, cultural experiences, or emotional depth, which makes it harder for us to perceive them as trustworthy agents. It is no coincidence that robots are often designed to resemble humans in appearance and behavior. It has been proven that anthropomorphism plays a crucial role in fostering trust and facilitating interaction [40, 41].

6.2 Preliminary Ethical Framing of Emotional States

Research in human-robot interaction has shown that the more a robot appears capable of understanding and replicating human emotional states, the more likely users are to perceive it as reliable and approachable [42]. This suggests that we tend to feel more comfortable with and trust things that resemble us as human beings.

Moreover, robots are perceived as operating due to complex, often opaque, systems of rules and algorithms that are not immediately understandable to most users. This “black box” nature of AI systems can further erode trust, as users may struggle to predict or fully comprehend a machine’s actions and decisions [43–45]. It should also be noted that the perceived trustworthiness of these systems is often contingent on the reliability of their information, performance, and capabilities as demonstrated in practice. Furthermore, from the user's perspective, a robotic entity will be considered "worthy of trust" only to the extent that it is not perceived as threatening, obstructive, or deceptive [46, 47]. Instead, it must be seen as a mechanical entity capable of promoting advantageous and productive outcomes while remaining sustainable from the perspective of human relationships.

6.3 Considering Intercultural Ethical Complexity

Finally, an additional layer of complexity arises from the cultural context. For example, the elderly population in Europe has historically exhibited a cautious attitude toward artifacts such as robots and artificial intelligence systems, contrasting with more accepting attitudes observed in Asian or Eastern populations [48]. Historical, political, socio-cultural, symbolic, anthropological, mythical, and religious factors shape the perception of robots, thereby influencing the degree of trust accorded to them [49]. For instance, some individuals may perceive a robot as a mere service tool, enabling the delegation of burdensome tasks, thereby inspiring a level of trust comparable to that traditionally reserved for a tool like a hammer.

As previously discussed, numerous factors have the potential to undermine human-robot trust, which becomes particularly significant in the context of companion robots. Examples include elderly care, assistance for individuals with disabilities, and the promotion of general well-being. Indeed, some scholars emphasize that the relationship between trust and use can be captured by the phrase "no trust, no use," underscoring its fundamental importance [50]. This presents a significant challenge: if care robots are designed to combat issues such as isolation and loneliness but fail to inspire trust, they risk becoming underutilized despite their potential benefits.

For this reason, the TRI-TECH project places a strong emphasis on investigating this construct, which, in the context of personalized companion robots (PCRs), can be conceptualized as encompassing more than merely an emotion, a perception of reliability, or technical competence in performing specific tasks. Rather, these dimensions are interrelated and necessitate a holistic, integral approach that transcends a narrow focus on the immediate goals the system achieves. Given these complexities that characterize human-robot interactions, strategies to enhance trust in these systems are crucial and can take various forms. These include investigating how humans perceive robots or identifying feasible modes of interaction, designing robots guided by the principles of "Ethics by Design," and ensuring that robots embody the very properties that constitute trust.

6.4 Enabling Multifaceted Trust

The TRI-TECH project integrates several properties identified as critical to the value of trust while addressing these dimensions. Specifically, by refining communicative exchanges, it pays attention to aspects such as proxemics, gestures consistent with semantic content, and ensuring that robots perform "checks" to verify the accuracy of their actions. We expect these characteristics will enhance the trust that elderly users may place in these systems. More concretely, we focus on designing actions embedded with specific sub-values that constitute trust and programming them into robots.

The selected values included autonomy and vulnerability—concepts previously explored in existing literature [51–53]—as well as relationality and meaningfulness, which are recognized by the Ethics of Care as indispensable components of human and caregiving contexts. Integrating these values into robot behavior aims to create interactions that inspire trust and strengthen user-robot relationships.

In summary, fostering trust in intelligent systems, particularly in sensitive and vulnerable environments, requires a multidimensional and culturally aware approach. Trust is not merely an auxiliary attribute but a fundamental element that influences the efficacy and sustainability of human-robot interactions. By incorporating design principles that align with human values and needs, we can establish a foundation for robotic systems that are not only technologically reliable but also socially and ethically integrated into the fabric of human relationships. This integration is essential to ensure their widespread acceptance and effective utilization, ultimately maximizing their potential to enhance users' well-being.

7 Implementation Barriers and Future Research Directions

The integration of robotics and LLMs as a tool for healthcare presents several challenges that need to be addressed, along with opportunities for further improvement. At the time of writing, LLM models often present problems such as generating false, obsolete, general, or unreliable information in response to questions. Currently, while the interaction flow is predefined by the systems designers, the dialogue of the robot is generated by an LLM based on a set of defined prompts. We have carefully crafted the prompts in order to minimize unexpected outputs or hallucinations from the LLM. Despite our efforts, occasional unexpected or hallucinated outputs remain a known limitation of current LLMs. In order to mitigate this issue, in future work, we plan to investigate the use of Retrieval Augmented Generation methods to drive the LLM outputs by means of curated external knowledge [54]. This approach would ensure more accurate, relevant, and contextually appropriate responses. During the experiments, a variety of interaction information is collected, enabling the identification of weaknesses in the robot's current functionality and the areas that require improvement. By analyzing user interaction patterns, it will be possible to refine responses, incorporate the most required functions, and optimize overall performance. Advanced machine-learning techniques will allow the robot to adapt and evolve based on real-world usage.

A further technical issue in the dialogue interaction with the robot regards the variability in elderly speech along with privacy and data security concerns. Indeed, in order to optimize speech-to-text, it is useful to exploit powerful AI tools that are commonly executed on cloud infrastructures due to their computational complexity. Of course, relying on external infrastructures for speech-to-text transcription, as done in our current implementation, exposes the sensitive communications of the elderly to unintended access by third parties. In future work, we will investigate the use of compact yet effective AI models for executing speech-to-text transcription on a dedicated secure server inside our system infrastructure. However, in scenarios where local computational resources or a private cloud infrastructure are available, the proposed system can be adapted seamlessly to operate within these environments, thereby ensuring greater control over data privacy and security.

Future work also includes seamless integration with various medical devices, such as smartwatches, blood pressure monitors, and smart scales.

These devices can provide valuable health information, enabling the robot to assess the health status of elderly users more accurately. By incorporating such devices, the robot can offer personalized health advice, fostering a more profound sense of trust and strengthening the relationship between the user and the companion robot. This integration is a significant step toward creating a more holistic and supportive caregiving solution.

From an ethical standpoint, two aspects warrant particular attention as key future research directions: the emotional and relational needs of elderly users and the potential risks associated with the use of companion robots. With regard to the first issue, it should be noted that the use of companion robots necessitates a critical evaluation of the modes of reception or expectations that arise from users, as well as the ways in which the presence of robots reconfigures daily practices or the meanings they assume within the experiences of the elderly population. On one hand, the presence of assistive and companion robots, as demonstrated in the literature, can reduce levels of loneliness, stress, and depression, offering users communicative and interactive exchanges that mitigate feelings of isolation. Conversely, there is a tangible risk of delegating the entirety of relational tasks to robots or assuming that they can replicate the complexity of human relationships. Human connections are inherently multifaceted, comprising components of value, semantics, perception, and shared worldviews, as well as profound emotional resonance. To address these challenges, several measures warrant consideration. For example, companion robots should be designed with the ability to recognize and adapt to the emotional and relational needs of elderly users, as well as their individual preferences. Specifically, advancements in technology should enable companion robots to be engineered with the capacity to identify and respond to the nuanced emotional and relational requirements of each user. This could encompass mechanisms for detecting subtle emotional cues or anticipating specific needs, thereby ensuring that their responses are both appropriate and aligned with their preferences. Moreover, when a robot is unable to meet the subject's needs, it should be equipped to effectively communicate this limitation to the relevant individuals, such as family members or caregivers, thereby facilitating the efficient transfer of responsibility for further support. A second consideration pertains to the design and deployment of companion robots and related artificial intelligence technologies, ensuring that these systems are firmly grounded in ethical principles that prioritize transparency and fairness. For instance, it would be crucial to identify the most effective strategies for informing elderly users about potential risks while critically assessing the contexts in which such information is both necessary and genuinely beneficial for the individual. These risks may include misinterpretations or unintended deception that could arise during interactions with companion robots. Proactively addressing these challenges is essential to managing expectations realistically and preventing ethical oversights, thereby ensuring the preservation of dignity and autonomy. Moreover, companion robots should not function in isolation but should be integrated into a comprehensive network of care that encompasses human caregivers, family members, and social connections. Implementing this integration necessitates meticulous understanding and planning to ensure that robots are seamlessly incorporated into the caregiving ecosystem. Through this approach, robots can complement human relationships rather than supplant them, fostering collaboration rather than engendering dependence.

Finally, there is a need for longitudinal studies to comprehensively evaluate older adults' perceptions and experiences with the companion robot.

These studies will help assess the emotional and functional impact of the robot over long periods. By addressing these challenges and undertaking these initiatives, it would be possible to develop a truly reliable, adaptable, and user-centered robotic companion that caters to the unique needs of older adults.

8 Conclusions

Integrating LLMs with robotics represents a considerable opportunity to enhance healthcare and elderly care, providing personalized, empathetic, and efficient assistance. Several initiatives have shown that the development of companion robots can improve the quality of life of older people by addressing their emotional, physical, and social needs. The TRI-TECH project aims to assess and address the ethical issues related to human-robot interaction by creating a companion robot capable of performing various tasks, from daily reminders to advanced multimodal interactions and adapting to users' preferences, creating meaningful, humanlike connections. The proposed system architecture, which combines the advanced NAO and PEPPER humanoid robots with the capabilities of LLM GPT-4, enables continuous communication and personalized interaction and lays the foundation for future advances in robotic assistance.

Key challenges, such as latency optimization, user trust enhancement, and ethical and safe use assurance, emphasize the importance of continuous research and development. Ethical considerations are crucial in this context, as the interaction between robots and users involves sensitive data and complex emotional dynamics. Future research should prioritize this aspect along with efforts to perfect technical capabilities, ensuring that robots not only assist but also respect and support their users.

The continued evolution of companion robots will require interdisciplinary collaboration across technical, medical, and social domains. By leveraging advanced technologies such as *RAG* and integrating with wearable medical devices, the proposed system will be able to provide increasingly reliable and contextually relevant care. Long-term studies will be crucial to assess user acceptance and identify areas for improvement, contributing to the development of truly adaptive and user-centered solutions.

In conclusion, the combination of robotics and LLM paves the way for a new era of intelligent care. If carefully designed and ethically deployed, companion robots can redefine elderly care, fostering independence, improving well-being, and creating a more inclusive and humane healthcare system. By addressing the challenges and following future research directions, we can realize the vision of robots as reliable partners in improving the lives of the elderly and transforming the wider healthcare landscape.

Acknowledgments This work represents a preliminary outcome of the TRI-TECH project ("Trust in Technology: How to Assess and Improve RoboT-User Interaction in Elderly Care Integrating Ethical, Technical and Social Variables," coord. by A. Pirni), funded under the National Recovery and Resilience Plan (PNRR).

The project, identified by CUP B83C22004800006, is supported within Mission 4 "Education and Research," Component 2 "From Research to Business," Investment 1.3, financed by the European Union, NextGenerationEU. Funding details are specified in the official directives: D.D. 343/2024 (pdf), D.D. 365/2024 (rectification of Article 3.3, Table 1, Annex D, pdf), and D.D. 468/2024 (pdf), which extended the submission deadline to April 10, 2024, at 12:00.

The authors collaborated closely in discussing, planning, writing, and revising all parts of the chapter. Specifically, Lorenzo Boi, Silvia M. Massa, Diego Reforgiato Recupero, Daniele Riboni, and Rubén Alonso contributed to the writing of Sects. 1, 2, 3, 4.1, 4.2, 4.3, 5, and 7. Michele Cardinali, Alberto Pirni, and Elena Ricci focused on Sects. 4.3, 6, and 7.

References

1. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. In I. Guyon et al. (Eds.), *Advances in neural information processing systems* (Vol. 30). Curran Associates, Inc..
2. Obinwanne, T., & Brandtner, P. (2024). Enhancing sentiment analysis with gpt-a comparison of large language models and traditional machine learning techniques. In A. K. Nagar et al. (Eds.), *Intelligent sustainable systems* (pp. 187–197). Springer.
3. Zhang, B., Haddow, B., & Birch, A. (2023). Prompting large language model for machine translation: A case study. In *Proceedings of the 40th International Conference on Machine Learning, ICML'23*. Retrieved from JMLR.org

4. Zhang, T., Ladhak, F., Durmus, E., Liang, P., McKeown, K., & Hashimoto, T. B. (2024). Benchmarking large language models for news summarization. *Transactions of the Association for Computational Linguistics, 12*(1), 39–57.
5. Kamalloo, E., et al. (2023). Evaluating open-domain question answering in the era of large language models. In A. Rogers et al. (Eds.), *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics, Toronto, Canada, July 2023* (Long Papers) (Vol. 1, pp. 5591–5606). Association for Computational Linguistics.
6. Wang, J., & Chen, Y. (2023). A review on code generation with LLMS: Application and evaluation. In *2023 IEEE International Conference on Medical Artificial Intelligence (MedAI)* (pp. 284–289).
7. Kim, C. Y., Lee, C. P., & Mutlu, B. (2024). Understanding large-language model (LLM)-powered human-robot interaction. In *2024 19th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 371–380).
8. Kim, K., Windle, J., Christian, M., Windle, T., Ryherd, E., Huang, P.-C., Robinson, A., & Chapman, R. (2024). Framework for integrating large language models with a robotic health attendant for adaptive task execution in patient care. *Applied Sciences, 14*(21), 9922.
9. Wang, D., & Zhang, S. (2024). Large language models in medical and healthcare fields: Applications, advances, and challenges. *Artificial Intelligence Review, 57*, 299.
10. Browne, R., et al. (2024). Reflective dialogues with a humanoid robot integrated with an LLM and a curated NLU system for positive behavioral change in older adults. *Electronics, 13*(22), 4364.
11. Laranjo, L., et al. (2018). Conversational agents in healthcare: A systematic review. *Journal of the American Medical Informatics Association, 25*(9), 1248–1258.
12. Irfan, B., et al. (2025). Between reality and delusion: Challenges of applying large language models to companion robots for open-domain dialogues with older adults. *Autonomous Robots, 49*(1), 9.
13. Frosolini, A., Catarzi, L., Benedetti, S., Latini, L., Chisci, G., Franz, L., Gennaro, P., & Gabriele, G. (2024). The role of large language models (LLMS) in providing triage for maxillofacial trauma cases: A preliminary study. *Diagnostics, 14*(8), 839.
14. Cavallaro, A., Perillo, F., Romano, M., Sebillo, M., & Vitiello, G. (2024). Social robot in service of the cognitive therapy of elderly people: Exploring robot acceptance in a real-world scenario. *Image and Vision Computing, 147*, 105072.
15. Albanese, G. A., et al. (2024). Robotic systems for upper limb rehabilitation in multiple sclerosis: A SWOT analysis and the synergies with virtual and augmented environments. *Frontiers in Robotics and AI, 11*, 1335147.
16. Peca, A. (2016). Robot enhanced therapy for children with autism disorders: Measuring ethical acceptability. *IEEE Technology and Society Magazine, 35*(6), 54–66.
17. Han, S., Min, X., Lao, J., & Liang, Z. (2023). Collecting patient feedback as a means of monitoring patient experience and hospital service quality—Learning from a government-led initiative. *Patient Preference and Adherence, 17*(2), 385–400.
18. Weiss, A., & Hannibal, G. (2018). What makes people accept or reject companion robots? A research agenda. In *Proceedings of the 11th PErvasive Technologies Related to Assistive Environments Conference, PETRA'18* (pp. 397–404). Association for Computing Machinery.
19. Jain, N., Schwarzschild, A., Wen, Y., Somepalli, G., Kirchenbauer, J., Chiang, P. Y., Goldblum, M., Saha, A., Geiping, J., & Goldstein, T. (2023). Baseline defenses for adversarial attacks against aligned language models.
20. Zou, A., Wang, Z., Carlini, N., Nasr, M., Zico Kolter, J., & Fredrikson, M. (2023). Universal and transferable adversarial attacks on aligned language models.
21. Curry, E., Heintz, F., Irgens, M., Smeulders, A. W., & Stramigioli, S. (2022). Partnership on AI, data, and robotics. *Communications of the ACM, 65*(4), 54–55.
22. Umbrico, A., Orlandini, A., & Cesta, A. (2020). An ontology for human-robot collaboration. *Procedia CIRP, 93*, 1097–1102. In *53rd CIRP Conference on Manufacturing Systems 2020*.

23. Maynard, D., Li, Y., & Peters, W. (2008). NLP techniques for term extraction and ontology population. In *Ontology learning and population: Bridging the gap between text and knowledge* (Vol. 167(6), pp. 107–127).
24. Recupero, D. R., & Boi, L. (2024). Towards seamless human robot dialogue through a robot action ontology. In B. Sartini et al. (Eds.), *Joint Proceedings of the ESWC 2024 Workshops and Tutorials Co-located with 21st European Semantic Web Conference (ESWC 2024), Hersonissos, Greece, May 26–27, 2024* (CEUR Workshop Proceedings) (Vol. 3749). Retrieved from CEUR-WS.org
25. Recupero, D. R., & Spiga, F. (2020). Knowledge acquisition from parsing natural language expressions for humanoid robot action commands. *Information Processing & Management, 57*(6), 102094.
26. Alonso, R., Bonini, A., Recupero, D. R., & Spano, L. D. (2022). Exploiting virtual reality and the robot operating system to remote-control a humanoid robot. *Multimedia Tools and Applications, 81*(11), 15565–15592.
27. Alonso, R., et al. (2020). A flexible and scalable social robot architecture employing voice assistant technologies. In B. N. De Carolis et al. (Eds.), *Proceedings of the Workshop on Adapted intEraction with SociAl Robots, cAESAR 2020, Cagliari, Italy, March 17, 2020* (CEUR Workshop Proceedings) (Vol. 2724, pp. 36–40). Retrieved from CEUR-WS.org
28. Kobayashi, S., et al. (2011). Intelligent humanoid robot with Japanese Wikipedia ontology and robot action ontology. In *Proceedings of the 6th International Conference on Human-Robot Interaction, HRI'11* (pp. 417–424). Association for Computing Machinery.
29. Fukuda, H., Mori, S., Kobayashi, Y., Kuno, Y., & Kachi, D. (2013). Object recognition for service robots through verbal interaction based on ontology. In G. Bebis et al. (Eds.), *Advances in visual computing* (pp. 395–406). Springer.
30. Oviatt, S. (1996). Multimodal interfaces for dynamic interactive maps. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 95–102).
31. Markievicz, I., Kapociute-Dzikiene, J., Tamosiunaite, M., & Vitkute-Adzgauskiene, D. (2015). Action classification in action ontology building using robot-specific texts. *Information Technology and Control, 44*, 6. https://doi.org/10.5755/j01.itc.44.2.7322
32. Sharma, D., Rashno, E., Zulkernine, F., El Khodary, E., Beninger, M., Almeida, R., Tao, J., Alaca, F., & Elgazzar, K. (2024). Triage-bot: An assistive triage framework. In *2024 IEEE International Conference on Digital Health (ICDH)* (pp. 138–140).
33. Kedzierski, J., Kaczmarek, P., Dziergwa, M., & Tchon, K. (2015). Design for a robotic companion. *International Journal of Humanoid Robotics, 12*(1), 1550007.
34. D'Onofrio, G., Sancarlo, D., Raciti, M., Burke, M., Teare, A., Kovacic, T., Cortis, K., Murphy, K., Barrett, E., Whelan, S., et al. (2019). Mario project: Validation and evidence of service robots for older people with dementia. *Journal of Alzheimer's Disease, 68*(4), 1587–1601.
35. OpenAI, Achiam, J., Adler, S., Agarwal, S., et al. (2024). GPT-4 Technical Report
36. Dario, P., Ciuti, G., Pirni, A., Capasso, M., & Bisconti, P. (2022). Social robots between trust and deception: The impact on institutions and practices. *Frontiers in Artificial Intelligence and Applications, 366*, 677–682.
37. Sanders, T., Kaplan, A., Koch, R., Schwartz, M., & Hancock, P. A. (2019). The relationship between trust and use choice in human-robot interaction. *Human Factors, 61*(4), 614–626.
38. Fabris, A. (2020). Can we trust machines? The role of trust in technological environments. In A. Fabris (Ed.), *Trust* (pp. 123–135). Springer International Publishing.
39. Decety, J. (2014). *The neuroevolution of empathy and caring for others: Why it matters for morality* (pp. 127–151). Springer International Publishing.
40. Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y. C., de Visser, E. J., & Parasuraman, R. (2011). A meta-analysis of factors affecting trust in human-robot interaction. *Human Factors, 53*(5), 517–527.
41. Hancock, P. A., Kessler, T. T., Kaplan, A. D., Brill, J. C., & Szalma, J. L. (2021). Evolving trust in robots: Specification through sequential and comparative meta-analyses. *Human Factors, 63*(7), 1196–1229.

42. Kolomaznik, M., Petrik, V., Slama, M., & Jurik, V. (2024). The role of socio-emotional attributes in enhancing human-AI collaboration. *Frontiers in Psychology, 15*, 1369957.
43. Belisle-Pipon, J.-C., Monteferrante, E., Roy, M.-C., & Couture, V. (2023). Artificial intelligence ethics has a black box problem. *AI and Society, 38*(4), 1507–1522.
44. Hassija, V., Chamola, V., Mahapatra, A., Singal, A., Goel, D., Huang, K., Scardapane, S., Spinelli, I., Mahmud, M., & Hussain, A. (2024). Interpreting black-box models: A review on explainable artificial intelligence. *Cognitive Computation, 16*(1), 45–74.
45. von Eschenbach, W. J. (2021). Transparency and the black box problem: Why we do not trust AI. *Philosophy & Technology, 34*(4), 1607–1622.
46. Akalin, N., Kristoffersson, A., & Loutfi, A. (2022). Do you feel safe with your robot? Factors influencing perceived safety in human-robot interaction based on subjective and objective measures. *International Journal of Human-Computer Studies, 158*, 102744.
47. Hancock, P. A., Billings, D. R., & Schaefer, K. E. (2011). Can you trust your robot? *Ergonomics in Design, 19*(3), 24–29.
48. Brohl, C., Nelles, J., Brandl, C., Mertens, A., & Nitsch, V. (2019). Human-robot collaboration acceptance model: Development and comparison for Germany, Japan, China and the USA. *International Journal of Social Robotics, 11*(5), 709–726.
49. Haring, K. S., Mougenot, C., Ono, F., & Watanabe, K. (2014). Cultural differences in perception and attitude towards robots. *International Journal of Affective Engineering, 13*(3), 149–157.
50. Schaefer, K. E., Chen, J. Y. C., Szalma, J. L., & Hancock, P. A. (2016). A meta-analysis of factors influencing the development of trust in automation: Implications for understanding autonomy in future systems. *Human Factors, 58*(3), 377–400.
51. Pirni, A., Balistreri, M., Capasso, M., Umbrello, S., & Merenda, F. (2021). Robot care ethics between autonomy and vulnerability: Coupling principles and practices in autonomous systems for care. *Frontiers in Robotics and AI, 8*, 654298.
52. Pirni, A., & Carnevale, A. (2013). The challenge of regulating emerging technologies: A philosophical framework. In *Law and technology: The challenge of regulating technological development* (Vol. 1, pp. 59–75). Pisa University Press.
53. Pirni, A., et al. (2017). Sostenibilità etica dei personal care robot. Linee per un inquadramento preliminare. In A. Pirni (Ed.), *Il post-umano realizzato. Orizzonti di possibilita e sfide per il nostro tempo* (Nuova Corrente) (Vol. LIX, pp. 133–151). gennaio-giugno.
54. Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., Dai, Y., Sun, J., Wang, M., & Wang, H. (2024). Retrieval augmented generation for large language models: A survey.