

The multi-modal interface of Robot-Era multi-robot services tailored for the elderly

DI NUOVO, Alessandro <<http://orcid.org/0000-0003-2677-2650>>, BROZ, Frank, WANG, Ning, BELPAEME, Tony, CANGELOSI, Angelo, JONES, Ray, ESPOSITO, Raffaele, CAVALLO, Filippo and DARIO, Paolo

Available from Sheffield Hallam University Research Archive (SHURA) at:

<https://shura.shu.ac.uk/16739/>

This document is the Published Version [VoR]


Citation:

DI NUOVO, Alessandro, BROZ, Frank, WANG, Ning, BELPAEME, Tony, CANGELOSI, Angelo, JONES, Ray, ESPOSITO, Raffaele, CAVALLO, Filippo and DARIO, Paolo (2017). The multi-modal interface of Robot-Era multi-robot services tailored for the elderly. *Intelligent Service Robotics*, 11 (1), 109-126. [Article]

Copyright and re-use policy

See <http://shura.shu.ac.uk/information.html>

The multi-modal interface of Robot-Era multi-robot services tailored for the elderly

Alessandro Di Nuovo^{1,2}  · Frank Broz³ · Ning Wang⁴ · Tony Belpaeme^{4,5} · Angelo Cangelosi⁴ · Ray Jones⁶ · Raffaele Esposito⁷ · Filippo Cavallo⁷ · Paolo Dario⁷

Received: 2 January 2017 / Accepted: 12 August 2017
© The Author(s) 2017. This article is an open access publication

Abstract Socially assistive robotic platforms are now a realistic option for the long-term care of ageing populations. Elderly users may benefit from many services provided by robots operating in different environments, such as providing assistance inside apartments, serving in shared facilities of buildings or guiding people outdoors. In this paper, we

present the experience gained within the EU FP7 ROBOT-ERA project towards the objective of implementing easy-to-use and acceptable service robotic system for the elderly. In particular, we detail the user-centred design and the experimental evaluation in realistic environments of a web-based multi-modal user interface tailored for elderly users of near future multi-robot services. Experimental results demonstrate positive evaluation of usability and willingness to use by elderly users, especially those less experienced with technological devices who could benefit more from the adoption of robotic services. Further analyses showed how multi-modal modes of interaction support more flexible and natural elderly–robot interaction, make clear the benefits for the users and, therefore, increase its acceptability. Finally, we provide insights and lessons learned from the extensive experimentation, which, to the best of our knowledge, is one of the largest experimentation of a multi-robot multi-service system so far.

✉ Alessandro Di Nuovo
a.dinuovo@shu.ac.uk

Frank Broz
f.broz@hw.ac.uk

Ning Wang
ning.wang@plymouth.ac.uk

Tony Belpaeme
tony.belpaeme@plymouth.ac.uk

Angelo Cangelosi
a.cangelosi@plymouth.ac.uk

Ray Jones
ray.jones@plymouth.ac.uk

Raffaele Esposito
r.esposito@sssup.it

Filippo Cavallo
f.cavallo@sssup.it

Paolo Dario
paolo.dario@sssup.it

Keywords Socially assistive robotics · Human–robot interaction · User-centred design · Multi-modal user interface · Elderly care

1 Introduction

The growing number of older people living alone in need of care is one of modern society's great challenges. It has been estimated that, by the year 2050, there will be three times

- ¹ Sheffield Robotics, Department of Computing, Sheffield Hallam University, Sheffield, UK
- ² Faculty of Engineering and Architecture, University of Enna Kore, Enna, Italy
- ³ School of Mathematical and Computer Sciences, Heriot-Watt University, Edinburgh, UK
- ⁴ School of Computing, Electronics and Mathematics, Plymouth University, Plymouth, UK
- ⁵ IDLab, IMEC, Ghent University, Ghent, Belgium

- ⁶ School of Nursing and Midwifery, Plymouth University, Plymouth, UK
- ⁷ BioRobotics Institute, Scuola Superiore Sant'Anna (SSSA), Pisa, Italy

more people over the age 85 than there are today and more than 20% of the population of most developed countries (e.g. Japan, USA, Europe, Australia) will be over 65 [43]. This older population will be more affluent and keen to enjoy their third age [54]. This is affecting the growth of the robotics market and research focused on services for *ageing well*, with robots that are increasingly available to assist and accompany the elderly users [52]. More are designed to help with physical aspects such as housekeeping, walking, and social communication, than to manage other activities, for example, money, laundry, grooming [5]. Other examples are Mobile Robotic Telepresence (MRT) systems that incorporate video conferencing equipment onto mobile robot devices which can be steered from remote locations [34].

In this scenario, the new field of *Socially Assistive Robotics* (SAR) has been defined [24] to identify those service robots that provide assistance through social interaction but without the necessity of physical contact with the user. In the application domain of elderly care, SAR platforms are often embedded in Ambient Assisted Living environments [11], in which smart assistance methods and systems are designed and developed for a better and safer life at home, with the aim of enabling elderly people to stay in their preferred environment longer. Many robots may be employed to cover different user needs, and thus there is the need to integrate and enable synergistic cooperation among all these robots that can be deployed to assist the elderly inside or outside their home [1, 16].

Social robots are seen as one possible way to address human resource and economic pressures on health care systems, but, at the same time, several studies have stressed that they are seen as expensive and job stealers by the public and suggest that care workers should be assisted but not be replaced by robots [12]. Research based on public surveys towards using robots in elderly care and other applications showed a high acceptance for pet-like therapeutic robots, human-like care robots, and surveillance care robots [40]. But it also reported a rejection in the case of a bathing robot because of the judgement that the robot-based action would be inferior to human-based action and that it would take away jobs from human workers. More generally, older adults are willing to accept home technological assistance when they are able to live independently [23, 50], but only if the perceived benefits of using the technology are clear [39]. Overall, stakeholders and the general public seem to be sceptical or even against the use of robots for the care of people [14].

In order to increase their acceptability, such robotic applications must be carefully designed from the requirements of end-users and developed taking into account the particular needs of specific user populations in terms of physical limitations and the digital divide. In particular, how to help older people to make use of new technologies is an important

research and development area, where accessibility, usability and lifelong learning play a major role. Thus, the human–robot interface must be easy to use to make clear the benefits offered by the system, in order to overcome scepticism and facilitate the adoption and actual use of robots in elderly care.

This paper presents the user-centred design approach of a multi-modal user interface (MMUI), which targets a more intuitive human way of interacting with machines, by means of speech, gestures or other modalities, that may be preferred over unimodal interfaces by elderly users [31]. Multimodal interfaces have been demonstrated to offer better flexibility and reliability than other human/machine interaction means [20].

The work presented in this paper has been carried out within the framework of the EU FP7 Large Integration project Robot-Era [46], which developed, implemented and demonstrated the general feasibility, scientific/technical effectiveness and social/legal plausibility and acceptability by end-users of a plurality of complete advanced robotic services for the elderly. Services were performed by three robots, and therefore, a single flexible interface was designed to simplify the use of such a heterogeneous system [32] and be the link between the user and the assistive robotic system. It can provide the user with a better awareness of all system functionalities and the feedback needed to perceive the benefits and make the whole system usable and acceptable.

The rest of the paper describes the work done by the authors to develop a more inclusive and usable user interface to facilitate interaction with the robots in order to increase the acceptance and willingness to use the novel service system. Section 2 describes the user-centred design approach of the robots and services of the Robot-Era system, and the experimental environments, and introduces the implementation of MMUI and the software system for the prototype. Section 3 presents the motivation and the objectives of the two experimental studies, methods and the instruments used to analyse the data, along with the details of elderly participants. Section 4 reports and discusses results from the two experimentation sessions, including some insights and lesson learned from our extensive experimentation in realistic environments, which, to the best of our knowledge, is one of the largest experimentation of a multi-robot service system so far. Finally, Sect. 5 gives our conclusion.

2 User-centred design of the Robot-Era multi-modal interface

In designing for older people, a designer needs to deliberately discard assumptions that the people being designed for are similar to the designer. This makes involvement of

Table 1 Overview of the user-centred design approach for designing, testing and refining the MMUI for the elderly users of the Robot-Era system

Step	Research objective	Evaluation		
		Method	Metrics	Location
0	Define user requirements	Focus groups, pilot tests	Interactive workshops, questionnaires, interviews, thinking aloud	Conference rooms, laboratory
1	Design and implementation of the first Robot-Era system and MMUI prototype			
2	Usability and acceptance	Experimental testing with a functional prototype	Questionnaires (UTAUT, SUS, ad hoc) & Interviews	Realistic environments (<i>living labs</i>)
3	Refinement and update of the Robot-Era system and MMUI prototype			
4	Interaction modality preference & updates validation	Experimental testing with updated prototype	Questionnaires (SUS, ad hoc) & Video analysis (gaze, attention)	Laboratory

representative older people in the design process extremely important [27]. The steps of the user-centred design of the MMUI along with research objectives, methods and metrics are summarized in Table 1.

The user-centred design process included one preliminary study (step 0) to define user requirements for the design. To set the requirements, a series of focus groups and workshops were conducted with elderly people and stakeholders in the healthcare industry. Older participants of these preliminary studies were recruited with the same criteria used for the general experimentation, which are reported in Sect. 3.2. Based on stated needs and desires and their technical feasibility, a set of services were selected for design and implementation to provide realistic test scenarios for the potential users. Full details of step 0 are in the Robot-Era deliverable D2.2 [26], which is downloadable from the project website. The services tested in the experimental studies presented in this paper are: Shopping, Reminding, Garbage, Communication, Laundry and Food Delivery.

Section 2.3 presents the implementation steps of the Food Delivery service as an example of the user-centred design approach used to improve the services following the feedback from elderly participants to our experimentation. For additional information, a step-by-step example of the Shopping service has been provided in [17], while a video presentation of all Robot-Era services can be found on YouTube [47].

Two experimental sessions were carried out with elderly users to collect feedback and verify the usability and acceptability of the system (steps 1 and 3). The implementation and refinement sessions took advantage of several pilot experiments with testers of any age to verify the effectiveness of the system before the actual experimentation.

Section 2.1 gives an overview of the Robot-Era system. Next, Section 2.2 describes the MMUI technical implementation with details of its design and guidelines followed. Additional technical details and examples, including some preliminary evaluations of specific aspects of the Robot-Era system, can be found in [17–19, 56].



Fig. 1 The three Robot-Era robotic platforms: Outdoor (*left*), Condominium (*centre*) and Domestic (*right*)

2.1 Introduction to the Robot-Era system: robotic platforms, software system and living labs

Within the EU FP7 project Robot-Era, three different robotic platforms were developed and optimized for the services and for usability and acceptability by elderly people in the home (*domestic*), communal shared living areas (*condominium*), and *outdoor* environments. Robots are represented in Fig. 1.

Robotic platforms used in the project are a RoboTech Dustcart (Outdoor Robot: ORo), and two specifically developed for Robot-Era on the base of Metralabs SCITOS G5 platforms: Domestic Robot, DoRo, and Condominium Robot, CoRo. Robots can be seen in Fig. 1). The appearance of the domestic and condominium platforms was studied in the preliminary studies with elderly and stakeholders and specifically designed for the Robot-Era project [9]. RoboTech Dustcart was the output of the DustBot project, which studied and developed the platform to have a friendly aspect that proved to be suitable for an urban environment [25].

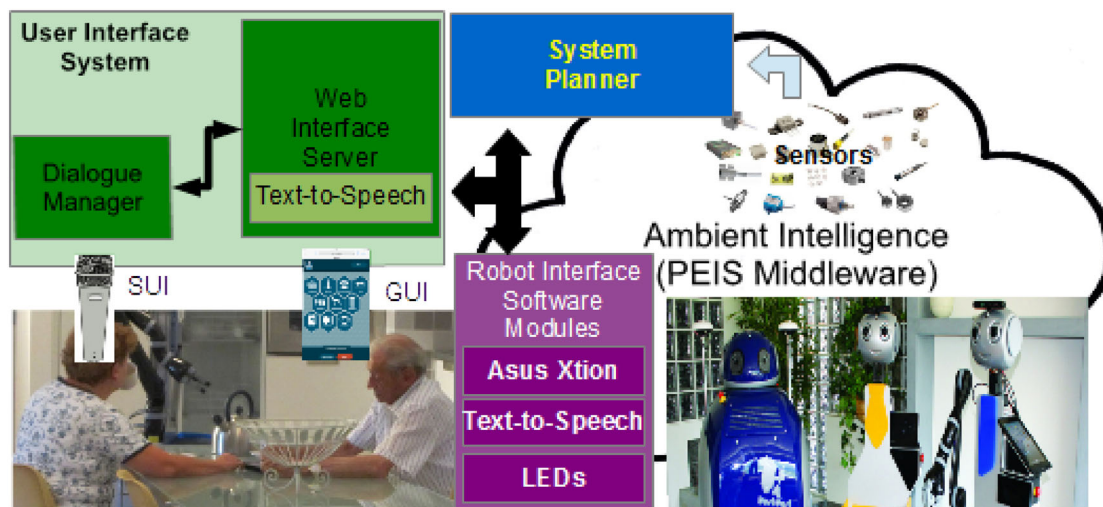


Fig. 2 The Robot-Era hardware and MMUI software system

The navigation sensors of DORO and CORO consist in a laser scanner (SICK S300) positioned on the front of the robot and a rear laser (Hokuyo URG-04LX) to have a 360 field of view, to be able to avoid an obstacle and for self-localization. The navigation stack relies on CogniDrive, a proprietary software of MetraLabs, and it is linked to ROS middleware [44], which is used for the interconnection with all the rest of software. DoRo and CoRo navigate autonomously with a maximum speed of 0.5 m/s (for safety), and the battery life is approximately 18 h. DoRo localizes the user position in the house via a Wireless Sensor Networks using RSSI and Zig-Bee technology [4]. Furthermore, the DoRo and CoRo are equipped with two RGB cameras (1024×768 , up to 30 fps) and one Asus Xtion (Kinect like) RGB-D camera (640×480 depth-image, up to 30 fps). These cameras are used to the perception capabilities of platforms and in particular for implementing the object manipulation. The overall strategy is that objects, providing enough optical features, are detected using the SIFT algorithm [36]. Finally, a Kinova Jaco arm is mounted on DoRo. The control algorithm for the manipulation tasks makes use of the ROS MoveIt! Framework, which collects a large set of algorithms for collision-aware motion planning [51] and has quickly gained popularity within the research community.

With regard to ORO, the sensors for obstacle detection consist of a laser scanner (Hokuyo UTM-30LX) positioned on the front of the robot and of infrared and ultrasound sensors used to detect steps, sidewalks, road gaps and any other common obstacle in the urban environment. Navigation is achieved using the navigation stack of ROS and localization makes use of both GPS and AMCL (Adaptive Monte Carlo Localization) which provides an accurate position and orientation of the robot with an accuracy of 5 cm and 1° with differential correction applied. ORO navigates autonomously

with a maximum speed of 0.8 m/s and the battery life is 8 h on a single charge. More details can be found in [10].

The Robot-Era software system is composed of several modules as shown in Fig. 2. Modules are interconnected each other using the PEIS Middleware [49] in order to collect information from the environment (i.e. sensors of the apartment and robots), to plan the robot's action and to allow users to interact with the system. Each robot runs ROS and the Metralabs MIRA [21] middlewares, the first is used for connecting the additional devices mounted on the robots (e.g. Asus Xtion, LEDs), while the second supports the CogniDrive navigation software that moves the robots. The Robot-Era software system also has a *Configuration Planner* Module (CPM) that is responsible for the planning of robot activities and for High-Level Reasoning. Details on CPM can be found in [19].

In the design for the Robot-Era Indoor and Condominium robots, we decided to adopt a detachable tablet because it has been found that elderly users are very receptive to tablets [57]. Indeed, they consider touch interfaces easier to use than other forms of interaction (e.g. classic keyboards and mouse) as shown in [35]. The tablet is physically attached to the robots, but mounted on a magnetic frame, which makes it removable. During our experiments, all the users were able to easily detach it when needed. This solution also demonstrated how to overcome the limitation of the short distance needed to interact with the robot when a fixed touch screen is adopted [42]. In addition, the graphical interface has been developed using web technology, which makes it accessible from any device (laptop, mobile phone, etc.) equipped with a web browser and virtually accessible from anywhere.

A total of three test sites were available to test the Robot-Era system. The two main test sites of the project were located in Peccioli, Italy, and Ängen, Örebro, Sweden, which are two

“living labs” that realistically emulate a home environment equipped with an ambient intelligence infrastructure, i.e. a wireless sensor network used to exchange information among the system artificial agents and localize the robot and the user. The third test site was an office room, specially fitted for the experiment to resemble a sitting room, within the facilities of Plymouth University, UK. Shopping, Reminding, Garbage and Communication were available in the Italian test site Peccioli, while Communication, Reminding, Laundry and Food delivery were tested in the Swedish test site Ängen. Not all the services were experimented on both test sites because the different availability of facilities: a laundry room was present in Ängen but not in Peccioli, and thus the related laundry service was tested only in Ängen; the pavement in front of the building entrance at Ängen was not suitable for the outdoor robot, and thus the services that use this robot (i.e. Shopping and Garbage) were not tested there.

2.2 MMUI design and implementation

In MMUI design, we had to extend the general principles and examples of recent MMUI for SAR in AAL Environments. In particular, the majority of the recent experimental projects on robots for elderly care focused on a single platform that was operating inside the user’s home [38]. Many projects used a wheel drive platform with a fixed touch screen as the main Graphic User Interface (GUI), but a few also implemented other modalities such as a Speech User Interface (SUI).

The requirements identified from the preliminary studies (step 0) suggested new design solutions to increase usability and acceptability such as the detachable tablet. With regard to the MMUI, tablet and robot voices were selected by the elderly according to the robot roles’ stereotypes: the domestic robot had a female voice as this was associated to a female maid, while condominium and outdoor have male voices as they were associated with a male warden/concierge and a delivery boy/postman. Other requirements from preliminary studies were on how to improve the graphical design, e.g. to add a specific button to go back to the main page because the usual logo icon link was not recognized by the elderly. Also, participants reported a clear preference for a menu with a higher number of buttons, with larger icons and without a grouping of functionalities into separate pages that the user must navigate to access services [17]. Finally, our preliminary studies suggested particular emphasis on the types and quality of feedback that the robot interface provides in order to make the system more intuitive for elderly users [8] and how often this feedback is given [7].

The Robot-Era MMUI software accepts commands through two main interchangeable modalities: they are a GUI, typically running on the tablet attached to the indoor robot, and a SUI, with a noise cancelling wearable microphone on the user. Indeed, two main software modules implement the

MMUI: the *Web Interface System* that includes the graphic user interface (GUI) and the text-to-speech (TTS) software; the *Dialog Manager* that implements the Speech User Interface (SUI) with the Automatic Speech Recognition (ASR) software.

The MMUI software is also responsible for providing feedback to users via all the available modalities to increase understanding thanks to the redundancy. Indeed, feedback to the user is given with both graphic and voice (from robot or mobile device), and also using an array of LEDs placed into the domestic robot eyes. This was implemented by making robotic platforms and web graphic interface able to produce sounds, including speech, at the same time. In addition, visual feedback is given showing specific text messages on the GUI and changing LEDs’ colour.

2.2.1 The graphic user interface: a web-based solution

The GUI for the Robot-Era services was implemented with a server-client architecture in order to provide remote control through mobile devices. It uses web technology to support the widest range of devices that can be connected to the system network, including tablets and smartphones that are preferred by the elderly [57]. A main menu index page introduces all the services and allows the user to navigate between them by clicking on the corresponding icon. All service interfaces are directly available from the main menu and contained on one page only, in order to avoid going back and forth through the pages to search for functionalities, as this could confuse the elderly user [53]. Colours were selected to maximise contrast according to the guidelines and design recommendations about web and tablet interfaces for the elderly provided by the World Wide Web Consortium (W3C), known as “*Web Accessibility for Older Users*” [13]. The icons are meant to be bold and simple. Where possible, we used a retro version of a technology so that they are easily recognizable by drawing on long engrained ideas of the elderly as to how things should look, e.g. to identify the communication service an old rotary dial phone was used. The *Acapela Voice As A Service* (VAAS) is used for TTS. This is a web service that receives any text with voice parameters and responds with an audio file containing the speech.

Information from the robots or the ambient sensor network is also made available to the user via notifications and warnings. The interface is paired to speech control of the robot. The two modalities are usually switchable, except for communication services where the GUI is required by the video calling, for which Skype was integrated using its web API. The GUI can run on any web browser and platform (e.g. PC, tablet, smartphones), but the graphic design aims to maximize the integration with the host device, in order to give the impression of a real product and, moreover, to provide people that have previous knowledge of the device with the

basic commands (e.g. volume and brightness controls) that they already know.

2.2.2 The multi-language speech user interface

The SUI was implemented in two steps during the Robot-Era project. First, a basic speech recognition system was implemented to provide the user with simple commands such as call the robot, start and confirm the robotic services by using one-word commands. The first version was intended to be simple in order to perform a first test with elderly participants and evaluate acceptance, in particular of the wearable microphone. After the success of the general experimentation and the positive user feedback, a refined speech recognition system and a more complex dialogue manager were developed for the final version that was tested in the focused study.

Speech recognition is implemented using the Nuance Communication SDK and is based on a set of restricted grammars. Nuance ASR was preferred because it supports all the languages used in our experimentation (Italian, Swedish and English, which was used also for debugging and demonstration). The recognition grammars are loaded dynamically to change what input the system is “listening for” based on the context and stage of the verbal interaction.

The main upgrade to the original speech interaction architecture was incorporating a more flexible and efficient dialogue flow control mechanism, as well as a more powerful dialogue manager. Context-aware models were implemented to improve recognition accuracy and system efficacy. The dialogue manager was based on the open-source Olympus dialogue management architecture. The Ravenclaw dialogue manager, part of Olympus, simplifies the authoring of complex dialogues, has general support handling speech recognition errors, and can be extended to support multi-modal input and output [3]. The main task in achieving a context-dependent spoken dialogue system was to design dialogue task specifications according to user expectations and service requirements. We did this by following three steps: (i) user expectation exploitation; (ii) service-specific grammar design; (iii) context-aware grammar flow switch.

The speech recognition is performed out-of-the-box, as no training session was required. Two thresholds are set according to the recognition confidence. All utterances below the lower threshold are discarded, while those above the higher threshold are accepted. A confirmation is requested to the user for all those recognitions between the two confidence thresholds. This strategy was adopted to avoid the need for frequent confirmation or error-recovery dialogues that could frustrate the user in the case of low speech recognition accuracy.

Users begin verbal interaction with the robot by calling the robot by name (e.g. Doro for the domestic robot). The robot’s name is defined as a “wake-up word” which must be

recognized before a service request interaction is initiated by the speech interface. This avoids service requests from being issued based on false positives, which could otherwise happen in situations where the user is speaking to another agent (real or artificial) rather than the robot. The keywords used to identify each service are specified in the grammars and may be said alone or recognized as part of a longer natural language phrase. After the user has called the robot, the dialogue proceeds in a system-initiative manner. In the initial service request interaction, the user can request any service. The following interaction will be determined by which service was selected. The speech interface is designed to generate short, naive, command-oriented dialogues with the user. In the case of services which require complex or extended user input, e.g. generating a shopping list or adding an entry for a reminder, the SUI suggests the user to follow the messages on the GUI display, for instance to double check items in the shopping list. The SUI also offers the robot to read aloud lists, such as the shopping list or the items in the food menus.

The Acapela web service was used also for robot speech production, via a ROS module on robots controlled via the dialogue manager. Voices were selected according to the preliminary selections.

2.3 Design and implementation example: food delivery

In this section, we will give a step-by-step example of the food delivery service to better present the functionalities offered by the Robot-Era MMUI and the changes implemented because of the user-centred design approach.

The food delivery service implements the meals on wheels service. The user can order a set menu among 3 choices using the MMUI. The system will then deliver the meal at a pre-set time. It employs the condominium robot only. The interaction starts when users tap on the corresponding button or call the robot to ask for delivery service using the SUI. Then, the

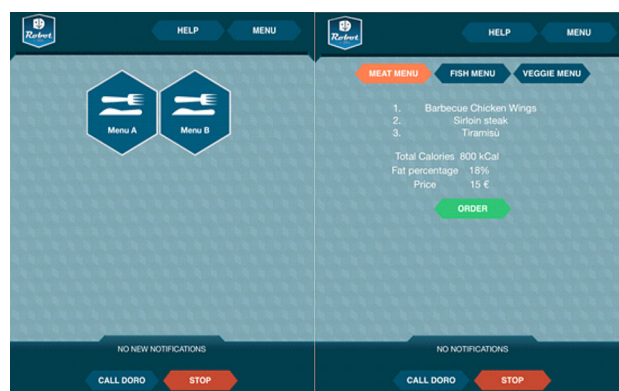


Fig. 3 GUI interface development example: (Left) First version of the food delivery page. (Right) Updated food delivery page with menu details after the user feedback



Fig. 4 Food delivery experimentation example: (left) After the user called the domestic robot via SUI, it turned to allow to detach the tablet and the user is reading the menu details via the GUI. (right) The user collects the goods from the condominium robot at the entrance door

food delivery page will be displayed on the GUI (Fig. 3). If users start the service via SUI, they can detach the tablet and use this device to read the menus as shown in Fig. 4.

For the first experimentation, the interface to this service was initially implemented as a simple choice between two menus (*A* or *B*), while the details were supposed to be given by the service provider using other means (e.g. printed on paper). After the first experimentation, many participants expressed the requirement to see the details on the tablet, and thus the interface was extended to allow the service provider to upload the details of the menus on the system. Thanks to the feedback from the general experimentation, the MMUI has been updated and upgraded, including a pilot study with a few elderly participants to confirm that the modifications went in the right direction. The updated design allows the user to browse three different options: meat, fish and vegetables. Each option has 3 courses, and the total calories are shown. A price is also shown to be more realistic. The user can also use the SUI to navigate among the menus and read the items aloud. After deciding which option they prefer, users can select it using the SUI or press the *order* button to proceed with the food order. Figure 3 presents the two versions of the GUI, while Figs. 5 and 6 report the dialogue flow of the first and second versions, respectively.

3 Material and methods

3.1 Motivation and objectives

The first study presented in this paper was part of the user-centred design and aimed to evaluate the first system prototype after the requirements set by the preliminary focus groups and pilot experiments. To this end, we recruited

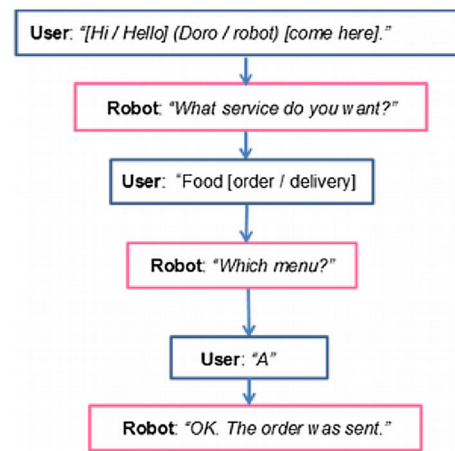


Fig. 5 Dialogue flow for the food delivery service with the first version of the SUI

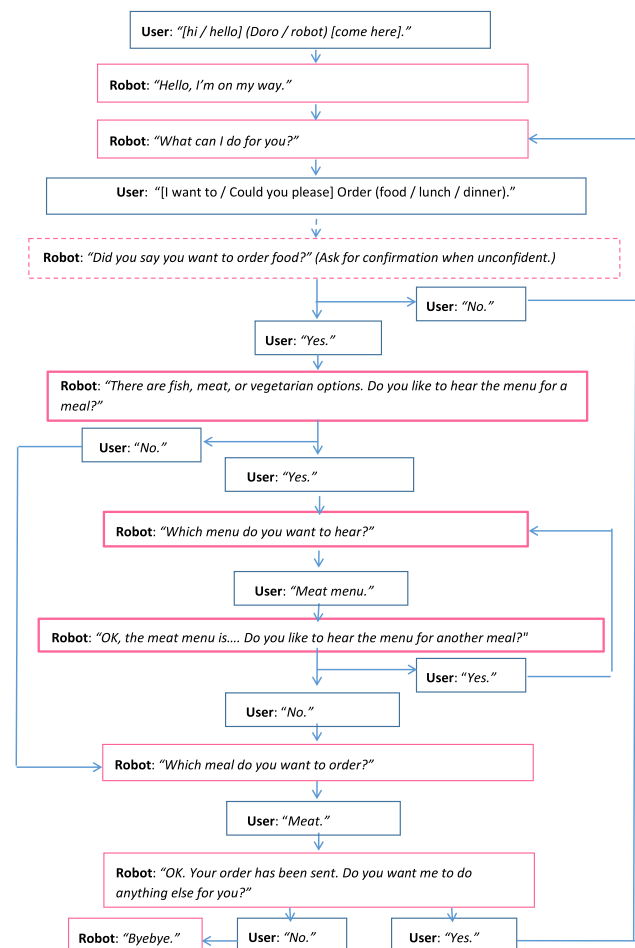


Fig. 6 Dialogue flow with the upgraded version of the SUI

elderly participants and asked them to realistically experiment some of the services in our test sites, because older people generally have difficulties to make the imaginative leap to seeing fictional demonstrations as representing an

actual application [27]. This realistic experimentation was designed to measure usability and acceptability of all the Robot-Era system parts, including the interface. An additional objective was to investigate the perception of the system between less and more technology-experienced participants. In particular, we tested the hypotheses that the evaluation of the GUI could influence the perceived ease of use of the entire system only in the less experienced, while the more experienced could relate their opinion to their attitude towards the use of the robotic system.

The second study was focused on the evaluation of the MMUI only. It was motivated by the outcomes of the first study, where both interaction modalities were well perceived by participants, but we were not able to retrieve significant information on preferences and if different modalities could be more effective in different conditions. Thus, the primary objective of the second study was to evaluate the user gaze as an indirect measure of the attention towards the different interaction options (GUI and SUI) during the different phases of the human–robot interaction. To this end, we designed a specific experiment that was focused on one service only, the new version of the Food Delivery as described in Sect. 2.3. Food Delivery is the service that required a more sophisticated interaction with several phases, which allowed the study of the user behaviour in different conditions. In addition, it was the best candidate to test the improvements of the user-centred design approach as it was the one that received the lowest score during the first experimentation.

Finally, both studies were motivated by the need to demonstrate the feasibility and effectiveness of the user-centred design approach in this application domain. Indeed, both studies had the secondary objective of validating improvement and optimization of the system updates according to the user requirements and feedback.

3.2 Elderly participants

During the experimentation phases of the Robot-Era prototypes, more than 100 elderly volunteers were involved and 82 were selected to participate in the realistic experimentation according to the following inclusion criteria: (i) retired and over 65 years old; (ii) able to live independently; (iii) with no, low or mild health impairments, cognitive and/or motor deficits; (iv) living alone or with relatives but without a dedicated carer. These requirements were motivated by the preliminary discussions with the stakeholders and a market analysis which identified this group as the most likely to benefit from the introduction of robotic services and, therefore, to adopt it in the near future. We excluded people with severe health problems as these are usually full-time assisted by specialized personnel in dedicated facilities. We decided not to include cases where the actual user is not the elderly person but the care worker as this introduces further con-

straints related to medical device legislation. Finally, these were motivated by the interest in investigating the interaction between technology and the elderly and its integration in their daily living, in order to support carers and not replace the relationship between older people and their family and caregivers. Only 72 participants were able to participate in all sessions and to complete the entire experimentation. Withdrawals were usually motivated by health-related issues after the first session.

Our participants comprised 31 males and 51 females with average age of 77.6 years (range 63–97). The average education level was secondary school education, only 12% ($N = 10$) had a university degree. All socio-demographic data and familiarity with the current technology were self-reported through a preliminary questionnaire. If we split the participants according to the test sites, the three sub-samples have different distributions in terms of age (Swedish were older), education (English had a higher level of education) and experience with technology devices (more Swedish and English owned and used smartphones and/or tablets regularly).

We know from the literature that the perceived ease of use is related to the previous experience with computers, but not to age and education [28], which is confirmed by our correlation analysis. To further investigate the possible differences in user behaviour according to the previous experience with technology, we created two subgroups:

- The *LOW* subgroup comprises those that have no or low usage of smartphones and tablets ($N = 26$).
- The *HIGH* subgroup is formed by those with high experience with a technological device (either smartphone or tablet) ($N = 25$).

Each subgroup is representative of the population, and they are otherwise similar. Indeed, if we consider the other basic descriptors (age, gender, education), the only significant difference is age (75.7 for Low, 83.3 for High). Note that participants ($N = 31$) that do not complete all the services or reported average experience with both devices were excluded from these subgroups.

3.3 Experimental procedure

The selected participants went to one of the testing sites, in which they actually experienced the Robot-Era services and interacted with the robots by means of predefined use cases. First, participants took part in a brief training phase in which they watched an instructive video clip about the potential real-world application of the Robot-Era system, including a demonstration of all services. There was a free question time in which volunteers' doubts were clarified in order to avoid the risk of failure. Participants did not receive any preliminary

explanation on how to use the interface. Then, participants experienced one by one the Robot-Era services available at the test site. Participants were asked to follow a use case, different for each service, which consisted in one or more tasks the participant should fulfil by using the robot. After all the tasks were completed, they had time for free interaction and additional personal testing of the interface. They were assisted by the interviewer only if needed. The interviewer is always behind the user and he/she did not interact with the participant unless explicitly prompted by them.

The scenarios were proposed by the interviewer in order to exemplify a possible realistic application to the user. The scenarios used for the services tested in the experiments presented in this paper are:

- **Shopping** The user wants to have breakfast with a visitor. While preparing breakfast, he realizes that bread is missing. Thus, he has to add bread to the shopping list and ask the Robot-Era system to go and get the bread for him.
- **Communication** In the first task, the user is in bed or has a mobility problem so he/she calls the robot to be connected rapidly with the family, for daily communication or service providers, in case of need. After the first task is completed, a simulated gas leak is triggered by the system so that the robot goes to the user to alert them and to start a call with a maintenance staff.
- **Laundry** In care or assisted living facilities as well as in accommodation where washing machines are located outside the apartments, the laundry has to be carried from users' apartments to the laundry room. This is time consuming and can be physically demanding. Thus, the robot could pick up the laundry at the user's apartment, carry it to the laundry room, put it into the washing machine, activate the machine and carry the laundry back to the user.
- **Reminding** A doctor, the caregiver or the user himself can set an event (take medicine, phone call or generic) to be reminded at a specific date and time. When it is the time, the domestic robot goes to the user and repeats the reminder until the user acknowledges that they have understood. This requires the domestic robot only.
- **Garbage** It is a very cold and windy day and the user has the flu, but the garbage bin is full and smells bad. Therefore, the user would use the Robot-Era service for the garbage collection.
- **Food Delivery.** The user wants a meal to be delivered to his/her apartment at a specified time (as offered by "meals on wheels"-service of the caring facility). The user calls the robot to select the menu for the next meal. The condominium robot delivers a meal to the user's apartment at the predefined time.



Fig. 7 Setting of the focused experiment. An elderly participant is seated in front of the robot with a table in between. The tablet is detached and placed on a table for easy reaching by the user

There was no fixed time for completing these services. Users were free to interrupt the experiment and take breaks if needed. During the usability and acceptability study, participants had to attend several experimental sessions, in which they experienced one or more services among the ones available at the testing site. Not all participants experienced all the services available. Further details on the experimental protocol can be found in the Robot-Era Deliverable D8.1 [37].

The procedure explained above applies also to the second experiments. The exceptions are that the elderly participant sits in front of the robot, while the tablet is detached and positioned on the table easily available to the participant. In the second experiment, participants had to come to the University only once to test exclusively the food delivery service using the same scenario described above. Then, we asked some specific questions to better investigate their preference of interaction modality between GUI and SUI. Meanwhile, we reduced the total number of items to answer by removing all the UTAUT questions. This was because the acceptability was not part of the evaluation and the participants of the first experiment often reported too many questions to answer. Figure 7 shows the experimental setting for the focused experiment.

To further analyse the interaction behaviour of participants during the service, we identified the following 7 steps related to the dialogue phases:

1. *Waking up* The participant wakes up the robot calling it by name "Johnny" or just "Robot".
2. *Begin ordering food* The robot asks "how can I help?", the user starts the food delivery service.
3. *Which menu to hear?* The robot proposes to read aloud the items of the three available menus. The user can select one menu or ask to read them all one by one.
4. *Reading menu* The robot declaims the items of the selected menu. When the robot is reading a menu, the tablet is automatically showing the items on the screen.
5. *Which meal to order?* The participant selects a meal.

Table 2 UTAUT questionnaire constructs [29]

Code	Construct	Definition
ANX	Anxiety	Evoking anxious or emotional reactions when using the system
ATT	Attitude	Positive or negative feelings about the appliance of the technology
FC	Facilitating conditions	Objective factors in the environment that facilitate using the system
ITU	Intention to use	The outspoken intention to use the system over a longer period in time
PAD	Perceived adaptability	The perceived ability of the system to be adaptive to the changing needs of the user
PENJ	Perceived enjoyment	Feelings of joy or pleasure associated by the user with the use of the system
PEOU	Perceived ease of use	The degree of ease associated with the use of the system
PS	Perceived sociability	The perceived ability of the system to perform sociable behaviour
PU	Perceived usefulness	The degree to which a person believes that using the system would enhance his or her daily activities
SI	Social influence	The user's perception of how people who are important to him think about him using the system
SP	Social presence	The experience of sensing a social entity when interacting with the system
TRU	Trust	The belief that the system performs with personal integrity and reliability

6. *Making order* The participant confirms the order or restarts from “which menu to hear?”.
7. *Ending the task* The robot asks if it can do anything else. The participant closes the interaction by saying “no”.

3.4 Instruments and statistical analyses

At the end of the testing of each service, participants completed the final questionnaires. Attitude, usability, acceptance and quality of life are the metrics that will constitute the kernel of each protocol, and they were evaluated all along the testing phases by means of qualitative ad hoc questions and standardized tools. Here, we report those that are relevant for the analysis of the MMUI.

3.4.1 System usability scale

Participants completed the System Usability Scale (SUS). This is a reliable, lightweight usability scale that can be used for global assessments of technological systems usability. SUS was developed by Brooke in 1996 [6], it is a simple, ten-item, five-point attitude Likert scale giving a global view of subjective assessments of usability. SUS yields a single score on a scale of 0–100. The SUS has been widely used in the evaluation of a range of systems. Bangor, Kortum and Miller [2] have used the scale extensively over a 10-year period and have produced normative data that allow SUS ratings to be positioned relative to other systems. According to them, products which are at least passable have SUS scores above 70, with better products scoring in the high 70s to upper 80s. Best products score better than 90.

3.4.2 UTAUT questionnaire

Then, each participant filled a questionnaire based on the model of the UTAUT [55]. We adopted the UTAUT model as proposed by [29], which has been widely used for the evaluation of robotic platforms and has been found to be highly reliable in several previous studies with elderly (e.g. [15, 29]). This model uses a structured questionnaire, in which each construct is represented by multiple questions (from 2 to 5). The UTAUT questionnaire comprises 36 items and 13 constructs, and definition and acronyms used here are in Table 2. The original questionnaire was translated from English into Italian and Swedish using the back-translation process employing bilingual speakers. This procedure ensures meaning and linguistic nuances are not lost, while the translated version remains as true to the original construct as possible.

3.4.3 Ad hoc questionnaire

In addition to the standard instruments, an ad hoc questionnaire was administered to evaluate some aspects of the human robot interaction for each service. As in the other questionnaires, participants could indicate their level of agreement to the statements on a five-point Likert scale including verbal anchors: “totally disagree” (1) “disagree” (2) “neither agree nor disagree” (3) “agree” (4) “totally agree” (5). The ad hoc questionnaire has three constructs, one for each modality (GUI, SUI) and one cumulative (HRI), which comprises all the questions. As a preliminary requisite, we tested the reliability of the ad hoc questionnaire constructs by means of Cronbach's alphas analyses.

It consisted of the following questions:

1. I feel the robot understood what I wanted to do. (HRI)
2. I could clearly hear what the robot said to me. (HRI)
3. I found the tablet easy to use to perform the service (GUI–HRI).
4. I could clearly read the messages on the tablet (GUI–HRI).
5. I understood what buttons I needed to press to perform the service (GUI–HRI).
6. I found it easy to speak to the robot to perform the service (SUI–HRI).
7. I understood what I could say to the robot to perform the service (SUI–HRI).

The Graphic User Interface (GUI) construct comprises questions 3, 4 and 5, while SUI construct is formed by question 6 and 7. The HRI construct comprises all the questions.

Cronbach's alpha is a measure of internal consistency, that is, how closely related a set of items are as a group. Here, we consider a solid construct when Alpha is at least 0.7, for details, see [33]. Cronbach's alphas calculated on the entire sample ($N = 82$) were: GUI 0.827; SUI 0.870; HRI 0.752.

3.4.4 Statistical analyses and subgroups

In data analyses of questionnaire constructs, we considered the following descriptive statistics: Average score (Avg), Median score (Med), and Standard Deviation (Std). In presenting the results, scores of negative items were reversed, and thus the higher the better in all cases.

To study the differences between the two subgroups, we also calculated statistical correlations to identify possible relations between constructs. On the basis of the correlation results, we hypothesized links between constructs and built two different models according the previous experience with the technology. The first model (LOW experience subgroup) relates the Perceived Ease of Use (PEOU) with the GUI, while the second (HIGH experience subgroup) shows that the PEOU is linked to the user attitude (ATT). Finally, we used the regression analysis to verify if the hypothesized model fitted the data collected. As part of the analysis, we established the R^2 value of the regression, which can be used as an indication of the predictive strength.

To statistically validate the comparison between the LOW and HIGH subgroups, the Mann–Whitney U (otherwise known as the Mann–Whitney–Wilcoxon) test has applied descriptive statistics and regression analysis to test the independence of the samples. The U test was preferred because it has more general applicability and, thus, reliability when samples have different distribution, size and variance. We set the statistical significance level to 0.05.

3.4.5 Video analysis of user behaviour

During the interaction sessions, a high-quality video was recorded to qualitatively analyse user behaviour in terms of interaction modalities. The camera was out of sight of the subject and mounted on a tripod in such a way that it provided an overview of the experimental area (see Fig. 7 as an example). The video was segmented into one-second intervals, and each second was analysed using two general categories: eye gaze and speech. The eye gaze category included the participants' gaze direction as a measure of attention focus: the robot, the tablet, and other objects in the room (including the interviewer). The speech category identified who was speaking: the participant, the robot, none. Behaviour could be logged simultaneously into different categories (for example, when the participant was looking at the robot and speaking). The evaluation was made with the use of a record sheet divided into seconds. Two observers separately compiled the results. One of the observers was the researcher that carried out the experiments, while the second observer was not involved in the project, and he was not aware of the results of the first observer. Results were compared for verification: the inter-rater reliability index is 92% and no statistically significant differences between the behaviour lengths rated by the two observers were found (Kolmogorov–Smirnow test).

3.5 Limitations and possible biases

We recognize that the experiments suffered some limitations that are intrinsic to this kind of research that involves both elderly people and sophisticated robotic platform prototypes within complex ambient intelligence environments. The flats were two realistic and unique environments that could not be reproduced elsewhere. For this reason, the recruitment was carried out by advertising the project and inviting potential users from nearby towns to come to one of the two labs to experience the Robot-Era system. Meanwhile, the majority of the potential users suffered from age-related problems that reduced their mobility and limited their availability. These circumstances affected the total number of people volunteering for the experimental phase, which was lower than planned, and biased their representation compared to the typical population [30]. In particular, the Italian test site is located in a new development not easy to reach from the nearby villages; for this reason, we can hypothesize that volunteers were those with the highest attitude and/or participant to a strong social influence. Indeed, these two factors had very high scores from the Italian sample.

The participant experience was limited to the real implementation of the services, which were conceived to demonstrate the service rather than provide full functionalities (e.g. food menus were fixed, and no personalization was allowed). The performance of the system could also be a limit as the

level of service was not constant for all participants. As an example, some participants had to repeat orders several times via the speech interface as the system was not able to recognize their particular voice or orders were mistaken. However, given the few number of critical system failures, we did not find a significant relation between technical faults and participants' opinion, which could still be affected.

4 Experimental results and discussion

This section presents the analyses of the data gathered from the experiments with elderly participants. Two experimental studies are discussed in the following subsections. The first study comprised all the services, and it was carried out in the living labs in Italy and Sweden. The second study focused on one service to confirm the improvement of the MMUI updates and upgrades after the user feedback of the first study and to perform a behavioural study in a smaller and more controlled environment. We conclude this section with a discussion of possible limitations and bias of our studies.

4.1 First study: usability and acceptability study

To study acceptability and usability of the system by the users, we analysed the data from SUS and the ad hoc questionnaires for usability of the MMUI, and from UTAUT for both acceptability and usability of the entire Robot-Era system. No significant differences were recorded between the participants of the two test sites. In particular, we underline that even if different services were experienced by the two groups, a similar Perceived Usefulness was scored in the UTAUT questionnaire with no statistical difference.

4.1.1 Usability results

Table 3 reports the descriptive statistics of the scores given by participants to the ad hoc questionnaire and SUS constructs.

In the ad hoc questionnaire, the greater majority of the users evaluated positively (4 or 5) all services experienced, average score is in almost all cases above 4, and median values are all 4 and often 5. The lower average score is for question n. 3 ("I found the tablet easy to use to perform the service") for the shopping service. This is due two main reasons: It was the very first attempt to use the system for all of participants, and it was difficult to recognize the shopping items because of the size and the lack of labels. Participants suggested adding labels and increasing the size of the icons representing the shopping items to improve the system. Indeed, we report that, after we updated the system according to these suggestions, a small focused study with 5 elderly participants confirmed that with these updates the service was fully usable for them. In the SUS questionnaire,

Table 3 Results of usability constructs for each service and overall

Service		HRI	GUI	SUI	SUS
Shopping	Med	5	4	5	88
	Avg	4.30	3.71	4.75	80.93
	Std	0.85	1.24	0.57	15.79
Communication	Med	5	5	5	85
	Avg	4.33	4.23	4.39	81.05
	Std	1.02	1.14	1.02	17.25
Laundry	Med	5	5	4	79
	Avg	4.30	4.54	3.78	75.83
	Std	0.85	0.72	0.95	20.48
Reminding	Med	5	5	5	78
	Avg	4.48	4.15	4.85	73.47
	Std	0.79	1.14	0.37	22.73
Garbage	Med	5	5	5	95
	Avg	4.49	4.33	4.49	90.00
	Std	0.88	1.12	0.80	14.74
Food	Med	4	4	4	75
	Avg	4.16	4.34	3.98	74.60
	Std	0.99	0.79	1.07	18.44
Overall	avg	4.34	4.22	4.37	79.31
	Std	0.90	1.02	0.80	18.24

participants scored all the services well, in particular Shopping, Garbage and Communication have a median score of 85 and above (meaning a superior result in terms of usability for the elderly participants).

From observation during the experiments and recorded videos, we saw that the subjects, who were not used to technology, had initial difficulties using the tablet, they often failed the first attempt, and sometimes they needed help from the interviewer to go on with the service. For the SUI, some subjects had difficulties because they did not say the right keyword to activate the system. Initial problems were mainly due to the lack of training, because usually after a short adaptation time subjects were able to use the interface independently with success. Indeed, in completing the questionnaire at the end of the experiment, they usually scored well the interface usability as reported in Table 3.

Overall, there is no correlation between the interface evaluation, which is equally positive for all services, and the SUS score. This is because the different perception of the usability that elderly users have according to their previous experience with the technology, as shown in the next subsection.

4.1.2 Influence of previous experience with the technology and analysis of the construct interrelations

The UTAUT questionnaire was completed at the end of the experimental session, and it referred to the entire system to

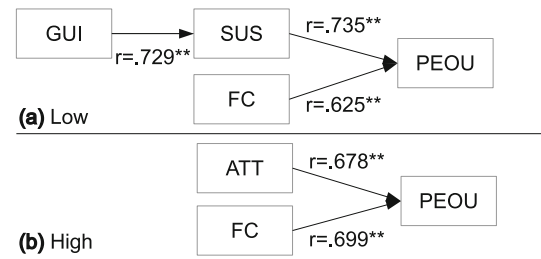
Table 4 Comparison of questionnaires' results according to the level of technology experience

Construct	LOW (<i>N</i> = 26)			HIGH (<i>N</i> = 25)			(Sig.)
	Avg	Med.	Std.	Avg	Med.	Std.	
ITU	4.00	4.00	1.23	2.95	3.00	1.42	< 0.01
PEOU	4.53	5.00	0.64	3.78	2.80	0.67	< 0.01
PU	3.81	4.00	1.20	3.44	4.00	1.29	0.29
ATT	3.78	4.00	0.86	3.79	4.00	1.10	0.75
PAD	4.42	4.33	0.48	4.07	4.33	0.89	0.15
ANX	4.56	4.75	0.66	4.39	4.50	0.94	0.59
FC	3.98	4.00	0.95	3.46	3.50	1.14	0.09
PENJ	4.41	4.70	0.81	4.03	4.20	0.65	0.01
SI	4.15	4.25	0.90	3.44	3.00	1.01	0.02
TRU	4.92	5.00	0.31	3.72	4.00	1.23	< 0.01
SP	2.61	2.60	0.64	2.44	2.20	0.82	0.40
PS	3.45	3.50	1.10	2.95	3.00	0.88	0.06
SUS	82.72	85.83	13.65	79.11	77.78	11.87	0.21
GUI	4.11	4.19	0.57	4.28	4.33	0.72	0.16
SUI	4.54	5.00	0.65	4.08	4.33	0.83	0.04
HRI	4.39	4.50	0.42	4.24	4.39	0.54	0.46

Statistically significant results ($p < 0.05$) are in bold

capture general feedback on the usability and acceptability of the Robot-Era system. Descriptive statistics are reported in Table 4, where the two groups (LOW and HIGH experience) are separated so we can see different evaluations from them. Statistical significance (Sig.) of the Mann–Whitney U test for the independence of groups is also reported. Table 4 reports also the comparison of the ad hoc and SUS questionnaires for the two subgroups.

Both groups demonstrated a very low anxiety (ANX), a good attitude (ATT) and the majority of them agreed that the system is perceived as useful (PU) without statistical difference between the two groups. With regard to PU, if we are stricter in the subgroupings so that only those with high experience in both tablet and smartphones are in the HIGH subgroup, we can see that the average for HIGH is 3.00 and the difference is statistically significant ($p < 0.05$). For the purposes of our study, the main substantial difference between the two groups is in the willingness to use the system (ITU); indeed the majority of the LOW were positive, while HIGH were neutral. This result can be explained with the Almere model developed by Heerink et al. [29] to assess the acceptance of assistive social agent technology by older adults. Indeed, according to this model, we can suppose that the result is due to the significantly lower Perceived Ease of Use (PEOU), Social Influence (SI) and Trust of the robot by the more technology-experienced participants as these relations were confirmed by a correlation analysis. The highest difference is seen in Trust, in terms of reliability, which is clearly due to the different experience

**Fig. 8** Relations among constructs according to the user experience with technology: **a** no experience; **b** high experience. Double asterisk indicates statistical significance.

between the two groups. Indeed, HIGH experienced participants were more able to recognize errors and glitches of the prototype behaviours because they compared them with the standard service of similar computer/mobile device applications. However, Trust construct scores are positive for the majority in both groups: 100% of LOW participants had an average score of 3.5 or more for this construct, while in the HIGH group 84% had at least 3 and 60% had 4 or more.

Of these constructs, the only one related to the MMUI is PEOU, which, according to the Almere model, should be influenced by Anxiety (ANX) and Perceived Enjoyment (PENJ). But we found no statistically significant correlations between these constructs. This can be because the UTAUT questionnaire and the Almere model were to evaluate a simple robot companion (iCat), and they may not fit all the characteristics of a multi-robot complex system like Robot-Era.

Correlation analysis of the sub-samples suggests that the usability of the system (SUS) is related to the perception of the ease of use (PEOU) only when the user has no or low experience, while expert users' perception is more related to their attitude (ATT) towards the robot, i.e. to their open mind to the use of a robot in their home. Two different models can be derived to explain this and they are presented in Fig. 8. The Facilitating Conditions (FC) construct is among the two main factors that influence the PEOU according the Almere Model [29]. For participants that have no previous experience with smartphones and tablets, the second factor is SUS, which is linked to the GUI. Conversely, for the more experienced users PEOU is related to their Attitude (ATT) towards the robot. The relations among these constructs are made explicit by the regression model analyses shown in Table 5. This result can be explained by the fact that older adults are more likely to call upon past experience (i.e. crystallized intelligence) [45]. For this reason, the elderly are *selective* users, i.e. they learn and use only what they really need [48]. Thus, they are reluctant to learn new procedures to complete the same task if they do not see a decisive improvement. As an example, the communication service is appreciated by those that do not use smartphones, but it is less attractive to experienced users who had already learnt how to use a mobile.

Table 5 PEOU regression model analyses

Experience Lvl	Predictor variables	Target variable	Beta	<i>t</i>	Sig.	<i>R</i> ²
Low	Constant	PEOU		2.407	0.024	0.626
	SUS		0.551	3.825	0.001	
	FC		0.356	2.477	0.021	
Low	Constant	SUS		1.701	0.098	0.522
	GUI		0.723	6.186	<0.001	
High	Constant	PEOU		5.590	0.014	0.621
	ATT		0.412	3.229	0.021	
	FC		0.548	3.767	0.017	

Statistically significant results ($p < 0.05$) are in bold

Table 6 Usability analysis and comparison

	First prototype			Second prototype			Sig.
	Avg	Med	SD	Avg	Med	SD	
HRI	4.16	4	0.99	3.99	4	0.58	0.140
GUI	4.34	4	0.79	4.13	4	0.45	0.110
SUI	3.98	4	1.07	3.90	4	0.65	0.485
SUS	74.60	75	18.44	80.67	80	9.04	0.229

4.2 Study on interaction modality preference

This section presents the results of a second experimentation that focused on the Food Delivery service with 15 elderly volunteers only. This focused study was planned after the first experiment with the double objective to validate the updates and investigate possible preferences between the two main interaction modalities (GUI and SUI).

Table 6 presents descriptive statistics of the usability questionnaires (ad hoc and SUS) scores given by participants of the focus experiment. Scores of the first version of the Food Delivery service are also reported for comparison. The usability score improved as expected as the SUS median increased 5 points with a lower variance. Results of the ad hoc questionnaire are similar with the same median scores and slightly lower averages but lower variance. The Mann–Whitney U test confirms that there is no statistically significant difference between the two groups.

It should be noted that the participants positively answer to the question “I am satisfied with the conversation with the robot about the food delivery service”, which was added to the ad hoc questionnaire. The average score for this question was 3.73, median 4, minimum 3, maximum 5.

Results of the video analysis are reported in Fig. 9, which reports for all the fifteen participants to the focused experiment. From the gaze analysis, we see that almost all (14/15) participants focused their attention on the technological devices. In detail, 12 participants (80%) spent the more

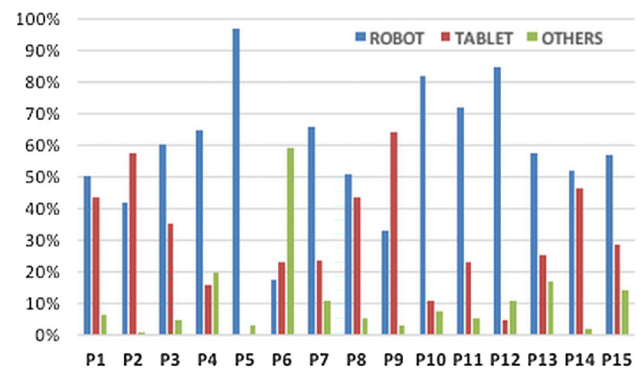


Fig. 9 The figure presents percentages of time spent by each participant (P1–15) looking at: the *robot*, the *tablet*, and *other* objects in the room

than 50% of the time looking at the robot, while just a two, P2 and P9, favoured the tablet. From the post-experiment interviews, participants told that their preference was not influenced by physical conditions or previous experience. Only P5 never considered the tablet because her physical condition, indeed she has a severe visual impairment and she did not have her magnifying glasses during the experiment. However, she liked to speak with the robot and she fully agreed to the question “I’m confident that I can use only speech (no tablet) to complete the food delivery service”. Vice versa, P12 did not pay much attention to the tablet as a personal preference as she had no vision problem.

The results obtained with the gaze analysis can be directly related to the modality of interaction preferred by the participants. Indeed, those that liked to interact with the SUI also spent the majority of the time looking at the robot, while the others mixed the two modalities. This confirms the preference expressed by participants of the previous experimentation.

Figure 10 reports average percentages of the total time spent by participants looking at the robot, tablet, and others, broken down for each step of the experiment. It can be easily seen that participants prefer to watch the menu items on the tablet while the robot is reading them aloud. Indeed, in the “*Reading menu*” step, 80% of participants (12 out of 15)

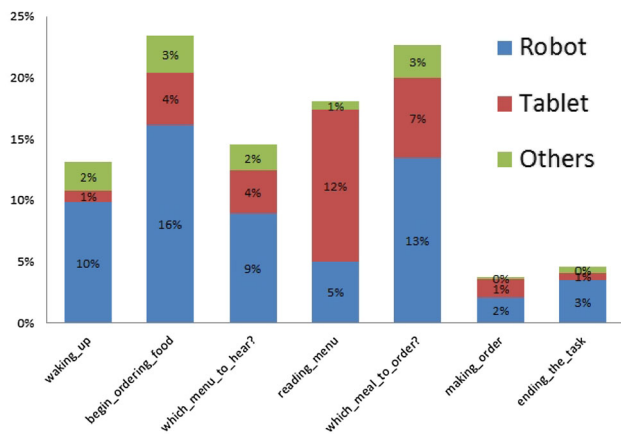


Fig. 10 Percentages of the total time of the experiment spent by all participants (P1–15) looking at robot, tablet, others during each step of the experiment

watched the tablet for the majority of the time (more than 50%). The possibility to switch between modalities is a clear advantage that allows elderly users to select the way they prefer to interact with the robot according to the different circumstances.

In fact, even if we see in the video a preference of the SUI as mean of interaction, the majority of the participants (8, 61%) gave the same score to the questions “I prefer to use the [tablet | speech] rather than [speech | tablet] for the food delivery service”. Moreover, as an additional confirmation of their preference of the multi-modal interaction capability of the system, all participants scored at least 4, with an average score of 4.14, “I like the idea of using speech and tablet together to complete the food delivery service”.

4.3 Insights and lesson learned

When it comes to the elderly and recent technology, one of the most common assumptions is that they need simplified tasks and more time to learn. On the contrary, in our experiments, even if we did not provide a preliminary training, all participants were able to complete all services on their own at least once. Indeed, we observed a quick learning performance and often they demanded more complex functionalities as shown in the Food Delivery service example. Participants with lowest experience tried to replicate gestures done by others; for instance, some attempted to push a button by sliding with a finger on the screen because they often saw this gesture being done by relatives when using their smartphones (e.g. to answer a call on Android or to unlock the screen on iOS).

In the first experimental study, 92.5% of participants spontaneously expressed the preference to speak to the robots during the unstructured interviews. This preference was independent of the technological experience and from the opinion of the GUI. It was motivated by the expectancy that the robot

could be a more than a simple servant, a real companion with which they want to have a conversation. In fact, many tried using more colloquial language, for instance by adding greetings like “my dear”, or even tried more complex answers than what was allowed, such as answering to the robot’s questions with jokes. This expectancy of a more natural conversation capability is likely to have influenced negatively the scores about the Social Presence of the robots in the UTAUT questionnaire, even if the appearance was very well evaluated according to the specific questionnaire (98% positive).

After the complexity of the dialogue was increased, in the second experiment, the majority of the participants (10/15) declared that they were satisfied with the conversation with the robot about the food delivery service.

Participants showed great tolerance to system malfunctions. The system usually performed well, but, in the cases when it required a restart because of critical failure, the participants were keen to retry. In fact, we did not observe any significant relation between errors and users’ opinions. Only once, one participant requested to stop the session because two consecutive system failures and he did not complete the questionnaires. We should remark that the speech recognition was improved by the use of a noise cancelling wearable microphone that was very well accepted by participants.

From the organizational point of view, the Robot-Era project carried out a large-scale evaluation of several robotic prototypes and services in realistic environments that, to the best of our knowledge, were not attempted before. However, as pointed out in the limitations, the participation was limited by the physical conditions and mobility opportunities of the elderly (e.g. not all of them are still able to drive). To reduce the number of visits to the labs, we organized sessions in which participants experienced more than one service (one after the other) and answered questionnaires and interviews. These sessions lasted usually from 2 to 3 h, including breaks that the participants were free to take when needed. Even if they were allowed to take breaks, some participants were visibly tired at the end of them. Also for this reason, in the second experiment, we focused on only one service and we reduced the number of questions. As a remark for future studies, we suggest to carefully consider the trade-off between the number of sessions and their length in order to take into account the physical limitations of this particular group of users. A good solution could be to organize a transport service or pay for taxis. Finally, if more sessions have to be scheduled, a larger sample should be initially recruited as we experienced a 12% dropout, usually because they experienced health-related issues after the first session.

We usually experienced some difficulties in recruiting elderly participants for experiments using passive advertising. Active advertising is the most effective strategy with the elderly. For instance, in recruiting for the second experiment, we organized a workshop in a sheltered accommodation

facility in Plymouth, in which we presented Robot-Era project, one robot and the potential use of robots in elderly care. At the end of the workshop, participants were very keen to sign up for the experiments.

5 Conclusion

In this paper, we presented the user-centred design approach, the technical implementation and the results of two experimentation studies in realistic and controlled environments of the robotic services tailored for elderly users by the Robot-Era project. To the best of our knowledge, the Robot-Era project partners carried out one of the largest experimentation of a multi-robot service system in realistic environments so far. The potential users were involved in the decision-making from the first stage of the hardware and software design process and invited to realistic tests of the robotic services.

Data from the experimental evaluation were analysed with a focus on the interface software system for multi-modal *elderly-robot* interaction. The results point out the positive evaluation of usability by potential users, especially those less experienced with technological devices (smartphones and tablets), who are those that can benefit more from the redundancy of the multi-modal interaction. On the other hand, more experienced users relate their perception of the ease of use to their attitude towards the use of the robotic system. This finding should drive the design of the future interfaces for *elderly-robot* interaction, because the number of elderly that possess and use technological devices is growing [41].

The positive acceptance showed by end-users and stakeholders involved in the Robot-Era project is of particular interest in Europe, because, just before the start of the project, an European survey indicated a positive attitude towards robots in general, but going into the specifics, 60% of respondents said that robots should be banned from the area of care for elderly, the children and the disabled [22]. Conversely, more than 90% of the participants and all the stakeholders contacted viewed positively the deployment of the Robot-Era services for taking care of the elderly in the near future.

An additional experiment on the user attention focus identified the common behaviour of the elderly participants to switch their attention between the different interaction modalities according to the current situation. This behaviour suggests that a MMUI facilitates a more personalized and flexible interaction, which could motivate the high acceptance and usability scores by the elderly participants of our experiments. Indeed, all participants showed to like the idea of using speech and tablet together to complete the service, suggesting that adoption of MMUI interfaces can be the key to overcome some age-related impairments. However, further research is required to identify the correct level of complex-

ity of the different modalities according to the user needs and expectations.

Finally, our experiments give additional evidence that multi-modality is an added value to the entire robotic system and it is a requisite for usable and widely accepted robotic services for people care. We strongly believe that a well-designed MMUI can play a decisive role in the adoption of future robotic systems as it can facilitate their use by those users that are less used to technology, who could potentially benefit more, and, thus, overcome the digital divide.

Acknowledgements The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n°288899.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Amirat Y, Daney D, Mohammed S, Spalanzani A, Chibani A, Simonin O (2016) Assistance and service robotics in a human environment. *Robot Auton Syst* 75(A):1–3
2. Bangor A, Kortum PT, Miller JT (2008) An empirical evaluation of the system usability scale. *Int J Hum Comput Interact* 24(6):574–594
3. Bohus D, Rudnicky AI (2009) The ravenclaw dialog management framework: architecture and systems. *Comput Speech Lang* 23(3):332–361
4. Bonaccorsi M, Fiorini L, Cavallo F, Esposito R, Dario P (2015) Design of cloud robotic services for senior citizens to improve independent living and personal health management. In: Ando B, Siciliano B, Marletta V, Monteriu A (eds) *Ambient assisted living*, Springer, pp 465–475
5. Broekens J, Heerink M, Rosendal H (2009) Assistive social robots in elderly care: a review. *Gerontechnology* 8(2):94–103
6. Brooke J et al (1996) SUS-a quick and dirty usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B (eds) *Usability evaluation in industry*, vol 189, no 194. CRC Press, pp 4–7
7. Broz F, Di Nuovo A, Belpaeme T, Cangelosi A (2015) Talking about task progress: towards integrating task planning and dialog for assistive robotic services. *Paladyn J Behav Robot* 6(1):111–118
8. Broz F, Nuovo AD, Belpaeme T, Cangelosi A (2012) Multimodal robot feedback for eldercare. In: *Workshop on robot feedback in human-robot interaction: how to make a robot readable for a human interaction partner at Ro-MAN 2012*, pp 1–4
9. Casiddu N, Cavallo F, Divano A, Mannari I, Micheli E, Porfirione C, Zallio M, Aquilano M, Dario P (2013) Robot interface design of domestic and condominium robot for ageing population. In: *ForITAAAL*
10. Cavallo F, Limosani R, Manzi A, Bonaccorsi M, Esposito R, Di Rocco M, Pecora F, Teti G, Saffiotti A, Dario P (2014) Development of a socially believable multi-robot solution from town to home. *Cognit Comput* 6(4):954–967
11. Cavallo F, Aquilano M, Bonaccorsi M, Limosani R, Manzi A, Carrozza MC, Dario P (2013) On the design, development and

- experimentation of the astro assistive robot integrated in smart environments. In: 2013 IEEE international conference on robotics and automation (ICRA), IEEE, pp 4310–4315
12. Coeckelbergh M, Pop C, Simut R, Peca A, Pintea S, David D, Vanderborght B (2016) A survey of expectations about the role of robots in robot-assisted therapy for children with asd: ethical acceptability, trust, sociability, appearance, and attachment. *Sci Eng Ethics* 22(1):47–65
 13. Consortium, W.W.W.: Web accessibility for older users. <http://www.w3.org/TR/wai-age-literature/>
 14. Conti D, Di Nuovo S, Buono S, Di Nuovo A (2017) Robots in education and care of children with developmental disabilities: a study on acceptance by experienced and future professionals. *Int J Soc Robot* 9(1):51–62
 15. de Graaf MM, Ben Allouch S (2013) Exploring influencing variables for the acceptance of social robots. *Robot Auton Syst* 61(12):1476–1486
 16. Di Nuovo A, Broz F, Cavallo F, Dario P (2016) New frontiers of service robotics for active and healthy ageing. *Int J Soc Robot* 8(3):353–354
 17. Di Nuovo A, Broz F, Belpaeme T, Cangelosi A, Cavallo F, Dario P, Esposito R (2014) A web based multi-modal interface for elderly users of the robot-era multi-robot services. In: Proceedings of IEEE conference on system, man and cybernetics, IEEE, pp 2186–2191
 18. Di Nuovo A, Wang N, Broz F, Belpaeme T, Jones R, Cangelosi A (2016) Experimental evaluation of a multi-modal user interface for a robotic service. In: Conference Towards autonomous robotic systems, TAROS, Springer, pp 87–98
 19. Di Rocco M, Sathyakeerthy S, Grosinger J, Pecora F, Saffiotti A, Bonaccorsi M, Cavallo F, Limosani R, Manzi A, Teti G, Paolo D (2014) A planner for ambient assisted living: from high-level reasoning to low-level robot execution and back. In: 2014 AAAI Spring symposium series
 20. Dumas B, Lalanne D, Oviatt S (2009) Multimodal interfaces: a survey of principles, models and frameworks. In: Human machine interaction, Springer, pp 3–26
 21. Einhorn E, Langner T, Stricker R, Martin C, Gross H (2012) Mira—middleware for robotic applications. In: Intelligent robots and systems (IROS), 2012 IEEE/RSJ international conference on, pp 2591–2598
 22. Eurobarometer S (2012) 382 ‘public attitudes towards robots’ European commission
 23. Ezer N, Fisk AD, Rogers WA (2009) More than a servant: self-reported willingness of younger and older adults to having a robot perform interactive and critical tasks in the home. In: Proceedings of the human factors and ergonomics society annual meeting, Vol 53. pp 136–140
 24. Feil-Seifer D, Mataric MJ (2005) Defining socially assistive robotics. In: Rehabilitation robotics, 2005. ICORR 2005. 9th international conference on, IEEE, pp 465–468
 25. Ferri G, Mondini A, Manzi A, Mazzolai B, Laschi C, Mattoli V, Reggente M, Stoyanov T, Lilienthal A, Lettore M, et al. (2010) Dustcart, a mobile robot for urban environments: experiments of pollution monitoring and mapping during autonomous navigation in urban scenarios. In: Proceedings of the IEEE international conference on robotics and automation (ICRA 2010) workshop on networked and mobile robot olfaction in natural, dynamic environments, IEEE
 26. Glende S, Nedopil C, Klemcke S, Marcellini F, Aquilano M, Cavallo F (2012) First report on the structures of the robot-era services. ROBOTERA Project Deliverable D2.2
 27. Hawthorn D (2007) Interface design and engagement with older people. *Behav Inf Technol* 26(4):333–341
 28. Heerink M (2011) Exploring the influence of age, gender, education and computer experience on robot acceptance by older adults. In: 6th ACM/IEEE international conference on human–robot interaction (HRI), 2011, pp 147–148
 29. Heerink M, Kröse B, Evers V, Wielinga B (2010) Assessing acceptance of assistive social agent technology by older adults: the almere model. *Int J Soc Robot* 2:361–375
 30. Hertzog C (1996) Research design in studies of aging and cognition. Academic Press, Cambridge
 31. Jian C, Shi H, Schafmeister F, Rachuy C, Sasse N, Schmidt H, Hoemberg V, Steinbchel N (2013) Touch and speech: multimodal interaction for elderly persons. In: Gabriel J, Schier J, Huffel S, Conchon E, Correia C, Fred A, Gamboa H (eds) Biomedical Engineering systems and technologies, communications in computer and information science, vol 357. Springer, Berlin Heidelberg, pp 385–400
 32. Kleinberger T, Becker M, Ras E, Holzinger A, Müller P (2007) Ambient intelligence in assisted living: enable elderly people to handle future interfaces. In: Universal access in human–computer interaction. Ambient interaction, Springer, pp 103–112
 33. Kline P (2000) The handbook of psychological testing, 2nd edn. Routledge, London
 34. Kristoffersson A, Coradeschi S, Loutfi A (2013) A review of mobile robotic telepresence. *Adv Hum Comput Interact* p 3
 35. Loureiro B, Rodrigues R (2011) Multi-touch as a natural user interface for elders: a survey. In: Information systems and technologies (CISTI), 2011 6th Iberian conference on, IEEE, pp 1–6
 36. Lowe DG (2004) Distinctive image features from scale-invariant keypoints. *Int J Comput Vis* 60(2):91–110
 37. Marcellini F, Bevilacqua R, Felici E, Klemcke S, Glende S, Esposito R, Cavallo F (2013) Experimental protocol for the first cycle of experiments of Robot-Era services. <http://www.robot-era.eu>. Robot-Era Project Deliverable D8.1
 38. Mayer P, Beck C, Panek P (2012) Examples of multimodal user interfaces for socially assistive robots in ambient assisted living environments. In: Cognitive infocommunications (CogInfoCom), 2012 IEEE 3rd international conference on, IEEE, pp 401–406
 39. Melenhorst AS, Rogers WA, Bouwhuis DG (2006) Older adults’ motivated choice for technological innovation: evidence for benefit-driven selectivity. *Psychol Aging* 21(1):190
 40. Moon A, Danielson P, Van der Loos HM (2012) Survey-based discussions on morally contentious applications of interactive robotics. *Int J Soc Robot* 4(1):77–96
 41. Ofcom (2014) Adults’ media use and attitudes report. Technical report
 42. Panek P, Edelmayer G, Mayer P, Beck C, Rauhala M (2012) User acceptance of a mobile led projector on a socially assistive robot. In: Wichert R, Eberhardt B (eds) Ambient assisted living, advanced technologies and societal change. Springer, Berlin Heidelberg, pp 77–91
 43. Plowman T, Prendergast D, Roberts S, Sherry J (2010) The global aging experience project. Ethnographic Research, Intel Research
 44. Quigley M, Conley K, Gerkey B, Faust J, Foote T, Leibs J, Wheeler R, Ng AY (2009) Ros: an open-source robot operating system. In: ICRA workshop on open source software, vol. 3.2
 45. Renaud K, Ramsay J (2007) Now what was that password again? a more flexible way of identifying and authenticating our seniors. *Behav Inf Technol* 26(4):309–322
 46. Robot-Era Project: Implementation and integration of advanced robotic systems and intelligent environments in real scenarios for the ageing population. <http://www.robot-era.eu>. FP7—ICT—Challenge 5: ICT for Health, Ageing Well, Inclusion and Governance. Grant agreement number 288899
 47. Robot-Era Project: Official video. <http://youtu.be/lv43z8YVQkY>
 48. Rousseau G, Rogers W (1998) Computer usage patterns of university faculty members across the life span. *Comput Hum Behav* 14(3):417–428

49. Saffiotti A, Broxvall M, Gritti M, LeBlanc K, Lundh R, Rashid J, Seo B, Cho YJ (2008) The peis-ecology project: vision and results. In: IEEE/RSJ international conference on intelligent robots and systems (IROS), IEEE, pp 2329–2335
50. Smarr CA, Prakash A, Beer JM, Mitzner TL, Kemp CC, Rogers WA (2012) Older adults preferences for and acceptance of robot assistance for everyday living tasks. In: Proceedings of the human factors and ergonomics society annual meeting, vol 56. pp 153–157
51. Sucan I, Chitta S (2015) Moveit!. <http://moveit.ros.org>
52. Tapus A, Maja M, Scassellatti B (2007) The grand challenges in socially assistive robotics. *IEEE Robot Autom Mag* 14(1):35–42
53. Torun M, van Kasteren T, Incel OD, Ersoy C (2012) Complexity versus page hierarchy of a gui for elderly homecare applications. In: *Computers helping people with special needs*, Springer, pp 689–696
54. Van Den Broek G, Cavallo F, Wehrmann C (2010) AALIANCE Ambient assisted living roadmap. IOS Press, Amsterdam
55. Venkatesh V, Morris M, Davis G, Davis F (2003) User acceptance of information technology: toward a unified view. *MIS Q* 27:425–478
56. Wang N, Broz F, Di Nuovo A, Belpaeme T, Cangelosi A (2016) A user-centric design of service robots speech interface for the elderly. In: *Recent advances in nonlinear speech processing*, Springer, pp 275–283
57. Werner F, Werner K, Oberzaucher J (2012) Tablets for seniors an evaluation of a current model (ipad). In: Wichert R, Eberhardt B (eds) *Ambient assisted living, advanced technologies and societal change*. Springer, Berlin Heidelberg, pp 177–184