This is a preprint version. The final version can be found here:
Robotics and Autonomous Systems, Volume 79, Pages 40–57, DOI: 10.1016/j.robot.2016.01.014

Long-Term Assessment of a Service Robot in a Hotel Environment

Roberto Pinillos[a], Samuel Marcos[a], Raul Feliz[a], Eduardo Zalama[b,*], Jaime Gómez-García-Bermejo[b]

[a] *Fundación Cartif. Parque Tecnológico de Boecillo. Parcela 205. Boecillo, Valladolid, Spain*
[b] *Instituto de las Tecnologías Avanzadas de la Producción. University of Valladolid*

Abstract

The long term evaluation of the Sacarino robot is presented in this paper. This study aimed to improve the robot's capabilities as a bellboy in a hotel; walking alongside the guests, providing information about the city and the hotel and providing hotel-related services. The paper establishes a three-stage assessment methodology based on the continuous measurement of a set of metrics regarding navigation and interaction with guests. Sacarino has been automatically collecting information in a real hotel environment for long periods of time. The acquired information has been analyzed and used to improve the robot's operation in the hotel through successive refinements. Some interesting considerations and useful hints for the researchers of service robots have been extracted from the analysis of the results.

Keywords: Social Robot, Service Robot, Robot Assessment, Metric

1. Introduction and background

Service robotics has had a major presence in research centers in recent years. However, there are few applications where robots are part of our daily life activities. A number of problems arise in the development of robots (localization, navigation, planning, interaction, etc.) which have been addressed extensively

*Corresponding author
Email address: ezalama@eii.uva.es (Eduardo Zalama)

in research centers and have been successfully solved to a significant extent. However, there has been limited success in adapting the solutions reached to the development of robots that can operate in real situations for long periods of time.

There are two main requirements that a robot must meet to be brought to the market: it must offer a good service at an affordable price, and it must perform the tasks with a minimal, tolerable failure rate. A relevant example of a robot that has successfully fulfilled both requirements is the vacuum robot, led by iRobot Roomba [11]. In addition, social service robots have to interact with humans and the environment in a user-friendly and socially-compatible way. This requires robust and versatile perception systems and solid interaction strategies to be developed. To date, much work has been dedicated to these research areas, looking for easy-to-use interfaces through which humans can communicate with robots in a natural way [19, 4, 6, 14].

Usually, such criteria as the ability to get and hold the user's attention to the proposed service, evaluated through direct observation, are used to assess the quality of the social interaction between robots and people [29]. Several studies have focused on the underlying reasons for the acceptance of social robots in different scenarios; their usefulness, adaptability, enjoyment, sociability, companionship or perceived behavioral control have been identified as important parameters for potential users' acceptance [22]. In [33] they suggest that, in addition to this qualitative assessment, metrics assessment should be used to provide feedback mechanisms aimed at improving the general performance of the robot. Benchmarking in robotics [23] has emerged as a solution to evaluate the performance of robotic systems in a reproducible way and to allow comparison between different research approaches. However, benchmarking is rather difficult in service robot applications [15], given that humans and real environments must be explicitly considered in the benchmarking methodology.

While many works have focused on analyzing robots working in controlled environments [28],where they interact with few users, e.g. in Physically Interactive RoboGames (PIRG) [24], only a few long-term studies have begun to appear over the last few years. In this sense, an increased effort has been made to understand the particularities of prolonged interactions with robots.

In [30], one of the first long-term studies in a real-world setting involving a social robot is reported. The robot was dedicated to a target impaired user, and was evaluated over 3 months. Results showed that it is important for robots immersed in public spaces to provide clear instructions on how to be operated. The results and also raised some issues such as the personality of the robot, the dialogue between users and the robot, and the relevance of group collaboration.

Another relevant example is the robo-receptionist Valerie developed by Gockley et al. and installed in the CMU campus. Results from a first study [13] indicated that, after a certain period, only few users interacted with the robot for more than 30 seconds. To avoid this, the authors proposed some design recommendations such as proper greeting and farewell behaviors, more interactive dialogue or a robust way of identifying repeated visitors. A second long-term study with the same robot [18] was carried out over nine weeks, in which the robot was able to display different moods. Results indicate that interactions were different depending on the level of familiarity and the robot's mood: frequent users interacted more often when the robot was in a positive mood, but the amount of time they dedicated to the robot was higher when it was in a negative mood.

In [17] the robot Robovie is evaluated in a shopping mall. The robot had the ability to adapt its dialogue to previous interactions with each user, while it was also capable of offering directions and advertising specific shops and services in the mall. Their results suggest that user perception towards the robot was positive, not only in terms of perceived familiarity, but also regarding intention of use and guidance. Also, better results were obtained from repeated visitors. In addition, the study also concluded that people's shopping behavior was influenced by the robot's suggestions.

A wider survey that addresses the particularities of prolonged interactions with robots can be seen in [20]. This survey addresses a total of 24 papers organized by their application domain: Health Care and Therapy, Education, Work Environments and Public Spaces, and the Home. The experimentation described in the said survey varied from paper to paper, but has usually been done over several sessions and days. However, few works have extended over several months. Analysis has been carried out in several ways: video and direct observation, system logs and post-trial interviews and questionnaires. Some drawbacks have been found that limit robot performance: i) Robots lack perceptual capabilities to enable rich social interactions and engage sporadic users, ii) robot autonomy is often limited, thus preventing the robot from operating for long periods of time, and iii) platforms often suffer from limited robustness and reliability, which results in weak supporting evidence of the robot's effectiveness, while technological acceptability is sometimes considered as one of the last design steps.

The present paper provides the results of a long-term assessment of a service robot in a hotel environment. Experiments have been carried out using Sacarino, an interactive bellboy robot [37], aimed at providing different services in a hotel: walking alongside the guests, providing information about the city and the hotel (restaurant hours, menus, etc.), and providing hotel-related services (calling taxis, guiding guests to the restaurant or other rooms, etc). Sacarino is designed to stay connected to a charger in the hotel lobby when it is not doing a specific task (so it can continuously provide effective services) as well as to navigate autonomously through the hotel facilities. The approach proposed in this paper involves being aware of current limitations, avoiding universal solutions and restricting the application domain to a concrete use case, focusing on an iterative design process. Our main aim is to take some of the technologies involved from the laboratory environment to higher technological readiness levels.

The rest of the paper is organized as follows: A description of our robot is presented in section 2. The methodology used to assess the quality of the services is presented in section 3. The following sections 4, 5 and 6, describe the different stages of assessment, including procedure and feedback based on the analysis of the results in each one. The dependability of Sacarino is analyzed in section 7. An enumeration of lessons learned is reviewed in section 8. Finally, section 9 includes the conclusions and future work.

2. Context of Our Research: Sacarino

In this paper, the assessment of Sacarino is presented [37]. This robot has been designed to operate in a hotel. In general, three levels can be identified when defining the structure of a service robot (Figure 1). The first level is the hardware and the mechanical structure, including sensors and actuators. The second level comprises the robot's control architecture. Architecture design has attracted the specific attention of the scientific community over the last few years, where different architecture paradigms have been developed (e.g. reactive, deliberative and hybrid [12]). Architecture design is a difficult task as it comprises different software components, programming techniques and modeling approaches: navigation techniques, simultaneous localization and mapping SLAM [8], planning [5], interfaces, human-machine communication systems including dialogue systems [21], recognition systems [25], cognitive modeling [35] and knowledge representation [3].

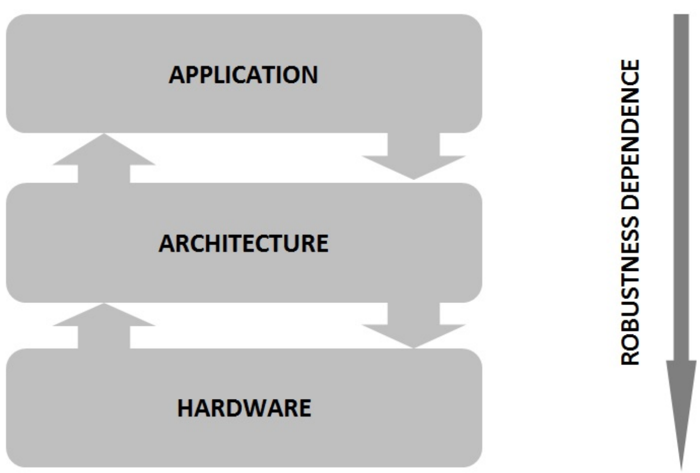

Figure 1: Infrastructure levels of a service robot

The third level is the application one. It is a key level, given that it specifically concerns the services to be provided by the robot. However, little effort has been devoted to this level and, few results have been obtained towards the development of service robotics. The main reasons can be found in:

- A lack of robustness in previous levels and the hierarchy across them, which hinders the development of service applications with the desired degree of autonomy. For example, a failure in the robot localization may cause goods to be dropped during loading and unloading, as well as navigation errors, collisions and crashes. While a certain failure rate may be acceptable in systems where a human is included in the control loop, the permissiveness is practically zero in systems that must operate autonomously.

- A lack of definition at the application level. Researchers frequently follow a bottom-up approach. They are usually concerned about looking for universal solutions without thinking beyond the architecture level. Research efforts often focus on robot localization, navigation, planning or face recognition, but fail in the integration of different technologies to create useful applications. Thus, we must combine this bottom-up approach with a top-down one, where the development is guided by the application specifications. We must decide which technologies are mature enough to be incorporated at the application level, and pay attention to their limitations and the additional requirements needed for ensuring that each service can be provided with a high enough quality level to be accepted by users. Some examples are the inclusion of artificial marks in critical areas where localization might be compromised or an accurate position is required (e.g. loading and unloading areas), or the incorporation of multimodal interfaces (e.g. touchscreen and voice recognition) in order to ease man-machine communication, even under high background noise.

- A lack of intensive testing at the application level. It is necessary to establish appropriate metrics for quantifying the robot's functionality and the service quality [33]. For example, concerning navigation, metrics can provide a measure of the navigation's effectiveness, such as deviation from the planned route, the area covered by the robot, the obstacles avoided, the time taken by the navigation tasks, or the number of times a human operator must intervene. Concerning interaction, metrics can provide measures such as user interaction periods, dialogue depth or user satisfaction. Besides, other factors should be evaluated before the commissioning of the robot, such as failure rate and maintenance cost.

For instance, the success of the Roomba robot owes much to a light and modular mechanical design, a small differential drive which provides a great maneuverability and the possibility to navigate under tables and beds, and a simple sensory system oriented towards navigation and the detection of areas where dirt is more intense. Moreover, its architecture is mainly oriented to the development of reactive behaviors. The robot only needs a precise localization when it must return to the charging station, for which it uses infrared beacons in the station cradle. Furthermore, the application layer includes specifications that extend to previous levels, such as adaptation to different surfaces (carpet, wood, etc.), corner cleaning brush, cleaning time adapted to the size of the room, and automatic search for the charging station and subsequent navigation to it (when the battery level drops below a given threshold).

2.1. Hardware Level

The Sacarino robot comprises two main parts: a mobile base for navigating through the hotel and an anthropomorphic body to interact with people and hotel guests.

The Sacarino base (see Figure 2) is controlled by four double wheels arranged in a syncrodrive configuration. The four wheels pull and rotate at the same time, driven by two motors, one for traction and another for turns. Two emergency stop buttons on both sides of the platform allow any user to prevent potentially dangerous situations.

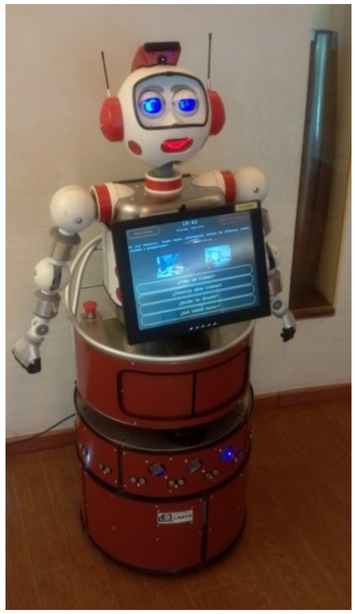

Figure 2: Sacarino

Sacarino's body, mounted on the moving base, supports the *"social"* components of the robot, as well as some navigation sensors (a SICK LMS-100 horizontal laser scanner and a set of ultrasonic sensors). The body can rotate 360 degrees synchronously with the wheels so that the social part of the robot always faces the direction of motion.

2.2. Architecture Level

The development of Sacarino's architecture has been conducted under the principle of component-based integration (component-based robotics framework). Figure 3 shows a functional block diagram of Sacarino's architecture, including the linking with some application level functionalities. There is a set of functional modules, each one in charge of a specific task. Module integration and communication has been carried out using the ROS framework [26]. Modules are grouped in two major functional subsystems. The social subsystem includes interaction modules for gesture control (body control), visual perception, chatbot [1] to generate dialogue and an Automatic Speech Generation and Recognition

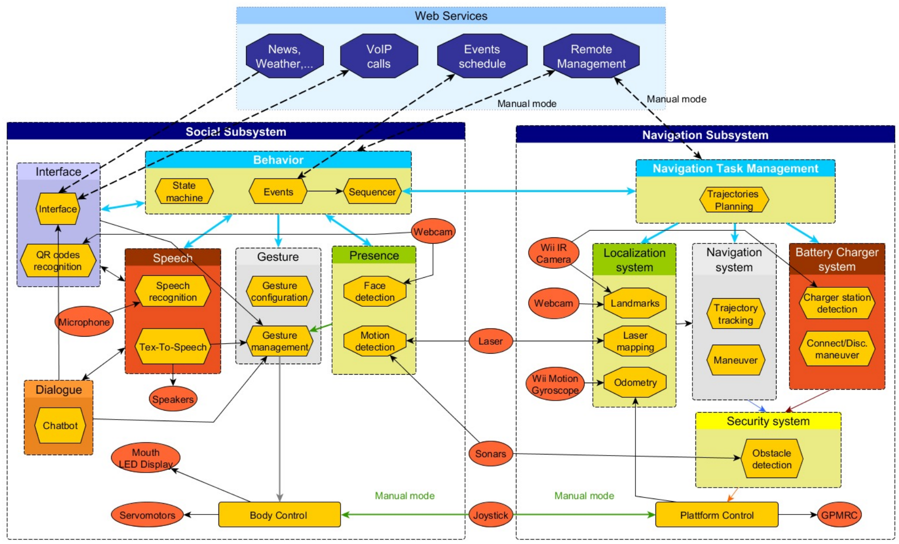

Figure 3: Sacarino's architecture

system [32]. Furthermore, this subsystem is responsible for the behavior of Sacarino. The robot's behavior varies according to a predefined state machine. State transition depends on the stimulus received (presence of user, interface inputs, execution of scheduled task). An overall state machine diagram is shown in Figure 4. The navigation subsystem includes such navigation modules as a planning, localization and reactive navigation, and the control modules that communicate with the controller board (providing proper abstraction of the hardware level comprising sensors and actuators). When a navigation task is required by the behavior module, either scheduled or requested by the user, it sends a navigation request to the navigation subsystem, making the robot change to navigation state.

2.3. Application Level

Currently, Sacarino can provide the following services:

- Giving information about the hotel facilities. This includes audio-visual information about the hotel, meal times and restaurant services. The

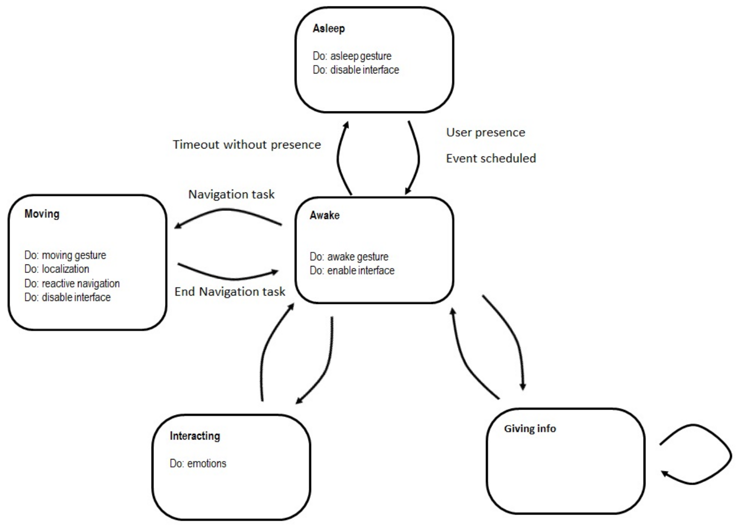

Figure 4: Sacarino's state machine diagram

guest can also obtain information about the city, shopping, museums, restaurants, etc. Most interactions are carried out bidirectionally through a voice and touchscreen multimodal interface. The robot gives information through spoken and on-screen written text, as well as images and maps; and the user can request information by voice or through the touchscreen, thus easing the interaction when the environment is noisy or the voice recognition system cannot correctly recognize the user's messages.

- Providing on-line information from the Internet, such as the weather forecast, the news of the day, and other entertainment information such as jokes and proverbs. Sacarino currently has a database of over 5000 jokes and proverbs.

- Giving event information. The robot may inform about upcoming events to be held at the hotel, such as congresses, conferences or presentations. A website hosted by the robot allows the staff to enter event information: event description, meeting room, schedule, etc.

- Taxi call. When a guest requests a taxi, Sacarino fills out a web form to call a taxi. To avoid bad practices, the robot requests guest identification by asking for a QR code printed on the guest hotel card that must be shown to the camera placed in the robot head.

- Breakfast control. It is intended that Sacarino can control the guests that have breakfast every day. Sacarino stays at the entrance of the restaurant and, when guests are detected (using the laser scanner), prompts them to say or input their room number through the touchscreen.
- Videoconference service to the hotel desk. When Sacarino is unable to provide feedback at a user's request, it can establish a video skype conference with the desk staff.
- Accompanying guests. The robot can navigate through the hotel dependencies to accompany guests. Each time a user requests directions, Sacarino asks the user aloud to follow him, and starts moving towards the destination. The user is expected to follow Sacarino. No further check of the user position is made during the navigation process. Currently, Sacarino's navigation is restricted to the ground floor (hall, dining room, bar and meeting rooms), but we hope that in a near future it will be able to reach the guest rooms to deliver such items as newspapers, snacks, etc.
- Aside from the autonomous mode, a web service hosted by the robot (Figure 3) allows the staff or developers to visualize the robot's status and remote control at any time.

Finally, concerning the three levels, it is worth noting that experimentation with real users in real environments entails ensuring that safety issues are properly observed. To this end, Sacarino has been developed taking into account the new safety standard ISO13482 for human robot interaction [16]. According to this standard, this robot is included in the section of personal care (non medical robots), and in the subsection of mobile servant robots. Therefore, ISO 13482 refers to the ISO 13849, IEC 61140, IEC 62061 and IEC 60204, standards for electrical, electronic and mechanical control systems for industrial machines. The ISO13482 provides important design considerations for hazard identification and risk reduction measures. Attention must be paid to battery and electrical isolation mechanisms, power, force control, speed and emergency and security stopping functions.

3. Methodology and Selected Metrics

Our main objective is to develop a service robot able to provide services in the hospitality sector. To do this, intensive testing on a set of services designed for hotels (information, guiding, assistance, experience sharing), supplied by a robot with social interaction abilities, have been addressed. We have considered not only the usefulness and added value that the developed services represent for a hotel environment, but also other considerations such as the cost and failure rate. The reason for doing so is that maximizing the functionality is one of the key aspects to be considered when developing an autonomous service robot: the robot should be able to provide the requested services for a long time, maintaining its capabilities, and at a reasonable price.

With those goals in mind, we have used a methodology aligned with the TRL (Technology Readiness Level) assessment described in the ESA TRL handbook [34]. TRLs are a set of management metrics that enable the assessment of the maturity of a particular technology, and as such they imply not only an improvement of such technology, but also a correct and consistent assessment. This technology readiness assessment (TRA) depends on the specifics of the prospective system applications, and implies the definition of the required metrics for evaluation. The process of increasing a TRL in a particular technology is shown in figure 5, and gives an overview of the process described in the rest of the paper. Figure 5 describes the iterative steps for Technology Readiness Assessment to increase the Readiness Level of a particular technology. It can be seen that each step depends on the accomplishment of the previous ones. First, a description of the research and development that has been performed is made. The next step (requirements) is to state the degree to which a future application of a technology is known, and whether the characteristics of the application are well enough defined.The verification consists in describing the environment in which the testing of the new technology has occurred, and the degree to which that environment is similar to the environment in which the technology will be used in operations. Finally, the viability assessment determines the prospective future viability of the technology being advanced, including risks and envisaged effort. It can be seen that our evaluation process has been iterative, given that an improvement in a technology or service (i.e. a TRL step forward) involves not only an evolution in the technology but an assessment of the said technology.

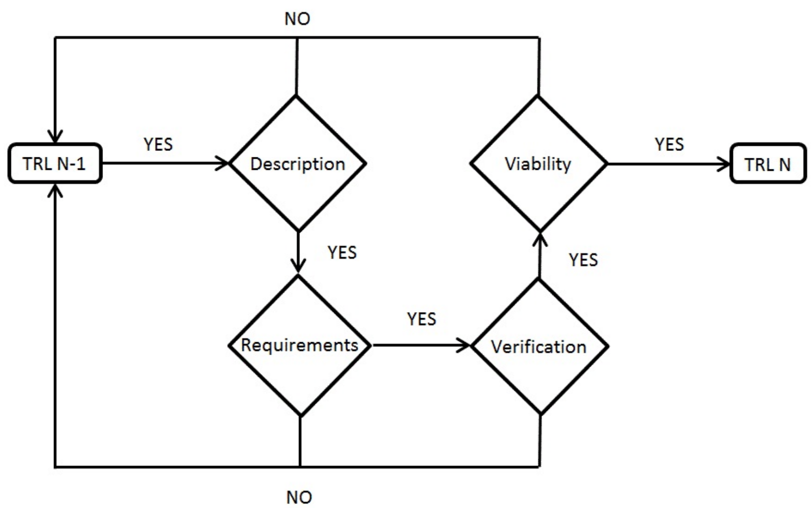

Figure 5: TRA methodology to increase TRL technology

The technology assessment provides feedback mechanisms to improve the robot's general performance. The evaluation should be done by measuring the abilities of the robot in relation to the services to be provided and should be defined independently of any particular robot configuration or application domain. According to the Strategic Research Agenda (SRA) defined in the euRobotics AISBL [9], a robot's abilities should be measured in terms of: configurability, adaptability, interaction capability, dependability, motion capability, manipulation ability, perception ability, decisional ability and cognitive ability.

Sacarino lacks some of the abilities defined in the SRA (such as manipulation). Moreover, some aspects (such as configurability, adaptability and cognitivity) have not been evaluated in this work. Our evaluation effort has focused on: social interaction capabilities, motion (navigation), dependability (the robot's capability to perform tasks without systematic errors), and general maintenance.

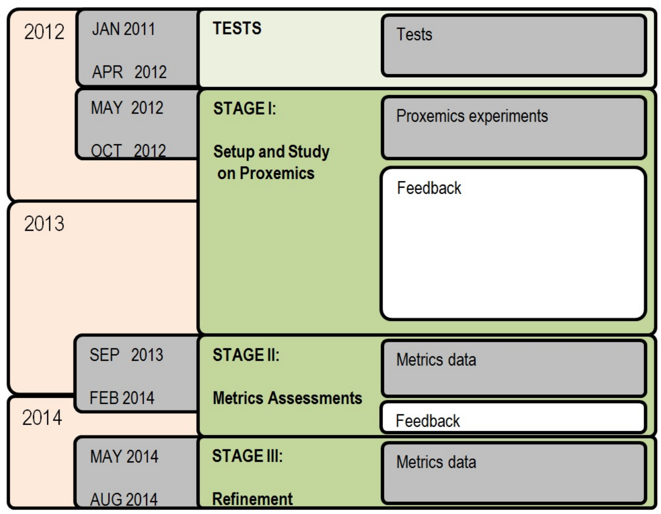

Figure 6: Evaluation timeline

The evaluation methodology used to assess the quality of the services has been divided into three main stages, involving quantitative measures of performance as well as subjective ratings. The first stage has covered a qualitative assessment through the direct observation (described in section 4) of guests interacting with Sacarino. The obtained results have served as the basis for the definition of a set of metrics that have been used during the second and third stages in order to evaluate the robot performance. Figure 6 shows the timeline of the evaluation, where the periods when the robot operates in the hotel are represented in gray. After each of these periods, the conclusions obtained from direct observations and metrics are used as feedback to improve the general performance of the robot. Feedback periods (development in laboratory) are represented in white.

The evaluation of our bellboy robot Sacarino in a hotel environment, under real conditions, was done in the Novotel hotel (Figure 7) in the city of Valladolid (Spain).

Figure 7: Sacarino placed in the hotel lobby

3.1. Social Interaction Metrics

In this section, the metrics used for analyzing how Sacarino relates to people, the services offered by it, and the communication channels employed by users are described. We will differentiate between *actions*, which refer to a screen touch or an instruction recognized by the voice recognition system, and *interactions*, which refer to the set of consecutive actions carried out by a given user.

Concerning actions, the following information is recorded:

- Time: The time at which the action takes place (in hh:mm:ss format).
- Input method: Voice or touchscreen.

- Question: Request made by the user and sent to the chatbot for subsequent processing. The question may be selected on the touchscreen or spoken and recognized by the ASR (Automatic Speech Recognition).

- Answer: Answer returned by the chatbot, which is displayed on the screen and spoken simultaneously (by means of the text to speech module).

- Topic / Attribute / Emotion / Action: Sequence of parameters returned by the chatbot. The topic corresponds to the context of the chatbot response (e.g., restaurant, time, etc.) and the attributes to the specific details. Emotion refers to the emotional expression generated by Sacarino during interaction, such as joy, anger or neutral. Action reports whether the current chatbot response has involved a specific action or not (e.g. telling a joke, giving headline news, navigating to a location, etc).

Actions are grouped into user interactions. An interaction has been considered to begin when a user/guest, alone or accompanied by a group of people, makes the first social action with the robot; the interaction has been considered to finish when the user leaves the interaction area or expressly says goodbye. It does not matter if the interaction is started by the robot or not. A time threshold between interactions has been used to detect when the user has changed. In case of doubt, changes can be set manually with the aid of the images recorded by the robot.

The following variables are computed:

- Average interaction time.

- Questions asked most often during interactions. This allows us to discern what information provided by the robot is more demanded and whether users really request information or simply play with the robot.

- Interaction channel: which channel is the most used in interactions (touchscreen, voice recognition or both).

- Idle time: time during which the robot is idle without interacting with anyone.

- Interactions against time of the day: number of interactions at every hour of the day, which allows a weekly/monthly average of the number of interactions throughout the day to be computed. This is to know the times at which Sacarino is busier.

- Number of interactions per day/week: The customer profile is quite different on working days (business travelers) and weekends (families and tourism). Therefore, it is interesting to analyze separately both kinds of days.

3.2. Motion Metrics

This section describes the metrics defined for evaluating Sacarino's mobility on site. The robot must travel to certain locations at specific times of day and/or on specific dates (restaurant at breakfast time, meeting area, etc). The displacements are scheduled by the hotel staff through the website hosted by the robot. In addition to this, any user at any time can request the robot to travel to a given location (e.g. to the restaurant, the elevator etc., for guiding purposes). Finally, the hotel staff can also ask the robot to travel to a given location. The following data are recorded for each navigation task:

- Navigation origin and destination.

- Origin and arrival date and time.

- Date and time the robot stops.

- Whether the order is synchronous (operational) or asynchronous (requested by the user).

- Date and time of events that result in the cancellation of route (emergency stop button, touch sensor activation, joystick control.)

This information is processed to obtain mobility metrics, among which we can mention:

- Destinations requested most frequently by users.
- Time per route (which allows situations to be analyzed in which the robot has taken more time than expected to reach a given destination).
- Total navigation time per day.
- Number of times the navigation task could not be completed (due to the activation of emergency stops, manual control, joystick, etc.).

Concerning the last item in the list above, we differentiate between *navigation incidences* and *navigation discards*. Navigation incidences refer to the case in which the robot cancels navigation on-the-fly because an unexpected event has been found (a stationary obstacle has been detected on the robot path or an emergency stop has been pressed by a user). In this case, the hotel staff has to guide the robot manually to a known position (the charging station) and remove the obstacle or release the stop. Navigation discards correspond to the case in which the robot does not initiate movement because, either the battery level is under a given threshold or an emergency stop has been pressed before starting to move. In general, these discards do not correspond to any robot malfunction.

4. Stage I: Setup and Study on Proxemics

This stage extended from January to October 2012. The implementation of the robot in the hotel and the definition of some tasks and services took place during the first months of this period. The evaluation was performed by direct observation of guests interacting with Sacarino. It is worth noting that when introducing an autonomous robot in a new environment, the usage and working practice of this new form of technology by non-expert users is commonly missing. Developers must make a significant effort to evaluate this circumstance.

This stage ended with a period (from August to October) during which a number of experiments on proxemics and the extent of interactions were performed [27]. This research analyzed the aspects of the robot's design and behavior that were relevant to user engagement and comfort. The experiments focused on the influence on proxemics, duration and interaction effectiveness of a number of dichotomous factors related to the robot design and behavior: robot embodiment (touchscreen with/without a robotic body), status of the robot (awake/asleep) and who started communication (robot/user). The experiments were done by direct observation of robot user interaction. The observations were made without the users being aware they were being watched, in an attempt to get the most natural conditions achievable. The data collected for each interaction was: sample number, date, time, gender, age estimation, interaction distance, who starts the interaction, interaction type (1. The user involved in the interaction uses speech and touchscreen or 2. The uses takes a passive role) and other additional comments. In the experiment with the robot in the awake status and taking the initiative for interaction, we observed the interaction of 95 people with Sacarino. From those 95 interactions, 53 were held by a single user whereas 42 were held by a single user but accompanied by more individuals. 74 were male and 21 female. The average duration of the interaction for each user was 59.51 seconds.

Some interesting conclusions: young and old people seem to feel more comfortable close to Sacarino than middle-aged people; users tend to maintain a higher interaction distance towards an embodied agent versus only touchscreen; embodiment engages users in maintaining longer interactions; interaction time increases when the robot starts the interaction. Overall, the obtained results suggested interesting guidelines about how Sacarino should be presented to the users in a hotel environment, along with other fine tuning advice. A detailed description of experiments and conclusions for this stage can be read in [27].

4.1. Feedback and improvements at hardware level

Servomotors: The experiments have shown that children tend to stand closer to the robot than adults. Usually, children play with the robot, grab it by the arms and head and try to force these elements, in most cases without paying attention to the screen. These actions have resulted in damage to several servomotors, whose replacement has been required. Consequently, springs and compliant mechanisms have been incorporated at arm and head joints to prevent damage when external forces are applied.

Screen: Another conclusion derived from the experiments has been the need to consider multimodal interaction (voice and touchscreen) for most services. The proxemic analysis has also shown that the close interaction distance required by the touchscreen must be compatible with the preferred personal interaction distance during face to face voice communication. As a result, larger fonts and images were incorporated to the screen, thus requiring the initial 10-inch screen to be replaced by a 17-inch one.

Automatic battery charging system: Some manual interventions were initially required to charge the robot. The hotel staff was responsible for doing this task at the robot's demand (the robot monitors its battery level). However, the staff occasionally postponed or ignored this demand. So the development of a fully automated charging system was envisaged (Figure 8). This charging system has been designed according to norm IEC 60204-1 for electrical appliances and low voltage (24v) under IEC 61140. A set of detectors in both the robot and the charger, along with a pressure sensor, guarantees that the connections are active only when the robot is coupled to the dock, as recommended by the ISO 13482.

Of course, the automatic connection to the charger requires quite precise maneuvering. Therefore, a new guidance system was developed using an infrared mark at the charger station, along with two charging contacts and an infrared *wiimote* controller camera on Sacarino's back. The integration of a data matrix at the charger and a second camera has also been considered [7].

Gyroscope: During the tests, we have found that the robot location degraded unexpectedly through time. This was a consequence of a small wheel misalignment in the synchrodrive platform and resulted in odometry drift and a subsequent deterioration of the localization and navigation systems. The SLAM

Figure 8: Automatic battery charging system

system was able to correct this error in short displacements in the laboratory, but this was not the case in the long term experimentation in the hotel. Therefore, a gyroscope was added to the robot. This device provides information about the Roll, Pitch and Yaw angles of the robot base. The Yaw angle is interpreted as a relative angle and is integrated with odometry data to correct the said odometry drifts.

4.2. Feedback and Improvements at Architecture Level

When the robot was working for many hours, some lack of robustness was detected at the automatic recovery from certain unexpected robot/environment states. This problem was solved by refining the ROS architecture of the robot. The final functional diagram of the architecture has already been shown in Figure 3.

Localization and Navigation: Navigation is a fundamental task for Sacarino, both for accompanying guests through the hotel dependencies and reaching the locations chosen by the hotel staff. In general, navigating from A to B requires the robot to determine where it is (A), where it is required to be (B), how it should get there (path planning), and how to deal with static and dynamic environmental factors such as obstacles encountered on the way.

The localization system initially operated on a range-laser onboard the robot (see section 2.1 and the odometry data, through a laser mapping system. However, we verified that location degraded through time, especially under dynamic conditions (people moving about the robot). The addition of a gyroscope, already described in the previous section, led to a large improvement. Furthermore, a topological map with 7 destinations and waypoints was selected for planning paths, instead of the general free space map previously employed. The use of this topological map has simplified path planning and has provided precise control over the paths and the areas through which the robot moves. Figure 9 shows the laser map of the hotel ground floor used for localization and navigation and the topological map with the different goals.

Reactive navigation is also used, so that Sacarino stops at obstacles and waits until they are removed. Of course, circumventing the obstacle and recalculating the trajectory would also be possible, but this action is often less effective because it can lead the robot to leave the operating area, therefore meeting unexpected situations. In addition, navigation is currently restricted to the ground floor, where there are no stairs or gaps that might compromise Sacarino's stability. The integration of sensors to detect stairs and electronics to remotely manage elevators will be addressed in the near future in order to extend navigation to other floors.

4.3. Feedback and Improvements at Application Level

Behavior: The results of our concurrent research on proxemics and interaction [27] show that a robot in the awake status and taking the initiative for interaction clearly favors user engagement. It was therefore proposed, for the second and third stages, that Sacarino should be 'awake', with its arms slightly bent at the elbow, the head held high and the eyes open and lit up. The robot should randomly make gentle movements with its arms and its head, and the screen should be on and displaying the main menu screen.

However, while Sacarino is at the charging station, it should be 'asleep', with its arms in an extended position, the head held high and the eyes open,

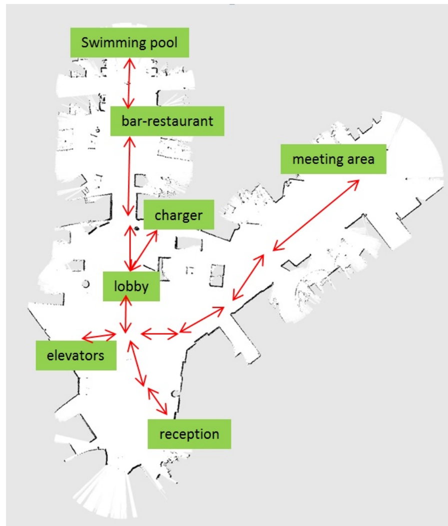

Figure 9: Topological map of ground floor with destinations

but without moving.

In both cases, when the laser sensor detects a person within a 3 meter semicircle around Sacarino, it should look in the direction of the approach and make a greeting to incite interaction. The greeting should include sentences such as *"Hello"* and *"Come closer and talk to me"*.

Interface: According to the user review in Stage I, the information displayed on the touch screen was improved. A more attractive interface was developed, simplifying the information displayed on it and designing more intuitive menus and buttons. Figure 10 shows an example of the resulting interface.

Moreover, in order to increase the robot's (battery) autonomy and track the results of the proxemics studies, a screensaver showing random pictures of the hotel should be launched when there is no person near the robot. Sacarino would immediately respond when a person comes close to it, putting the main menu on the screen and making a greeting to incite interaction. In addition, it was noticed that most hotel guests did not actually know what services were offered by the robot. Therefore, a Help menu was added to the interface, to show specific instructions.

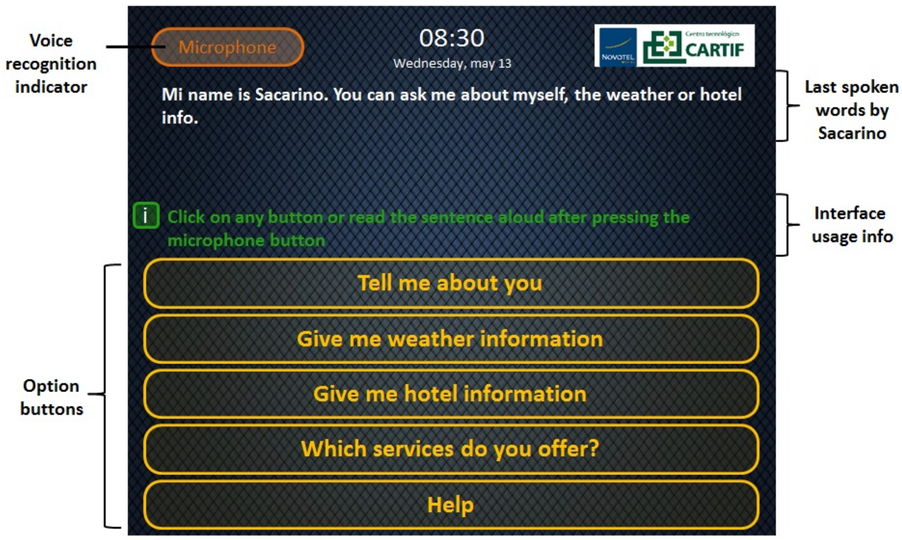

Figure 10: Sacarino's interface

Dependability: The presence of the robot in the hotel meant extra work for maintenance personnel. These personnel were in charge of turning the robot on and off and connecting it to the battery charger (when required). The latter can be rather demanding because the battery autonomy may be less than 4 hours, depending on the tasks carried out by the robot. (In particular, tasks involving navigation result in important battery consumption). The automatic charging system has allowed this to be dealt with, the hotel maintenance staff only becoming responsible for turning the robot on/off. The robot generates a navigation task to the charging station at any low battery level detection and plugs itself in automatically.

5. Stage II: Metrics Assessment

This stage extended from September 2013 to January 2014. During the first two months, tests were carried out under our direct observation. The overall operation of the robot and the improvements introduced in feedback I were checked (in particular, those concerning the navigating tasks and the maneuvers to reach the charging station). For the last two months of the current stage, Sacarino was operated in normal conditions in the hotel without the presence of our research team. The hotel maintenance staff was responsible for turning the robot on and off, while the robot returned and automatically plugged itself into the charger station at any battery level drop. The metrics defined in 3 were recorded to assess the abilities of the robot in relation to the services provided.

As previously mentioned, Sacarino has been designed to stay connected to a charger in the hotel lobby when not doing a specific task, so it can continuously provide effective services, and to navigate through the hotel's facilities. A website hosted by the robot (Figure 11) allows the staff to provide the robot with event information and to schedule navigation tasks. There are 2 preprogrammed battery charging tasks a day, during which navigation is disabled.

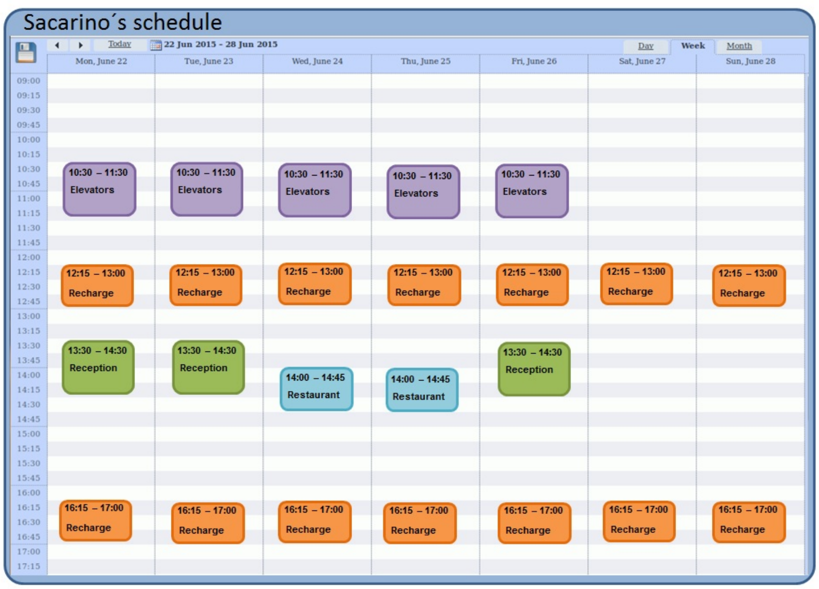

Figure 11: Sacarino's schedule

All the data were taken during a 60 day period, including working days and weekends. Sacarino was turned on for 23 of these days (according to the workload of the hotel maintenance service manager). Table 1 reports the overall operation data. The robot operated for 220 hours, 9.56 hours on average per

Table 1: Overall operation time data resume

	Days robot working	Total working time	Total social interaction time	Total time in motion tasks	Total number social actions	Total users/ interactions	Average interaction duration
Stage II	23	220.09 h (9.56/day)	9.07 h (4.12%)	2.46 h (1.12%)	5068 (220.3/day)	349 (15.17/day)	1.54 min

day. It interacted with guests for 4.12% of the total time (9 hours) and was developing navigation tasks for 1.11% of that total time (2.5 hours). Therefore, Sacarino was idle, ready to interact at any user request, for 94.76% of the total time. 25 reboots were produced during these 23 days due to some operational errors, which means that, on average, the robot was restarted more than once a day.

5.1. Social interaction

A total of 5068 social actions and 349 user interactions (according to the definitions given in section 3.1) were registered, as reported in table 1. The average time for which a user interacted with Sacarino was 1.54 minutes.

Figure 12 shows the average number of social interactions throughout the day. The times at which Sacarino was busier can be seen: between 10am and 11am in the morning, and about 8 pm in the evening. Interactions were concentrated between 8am and midnight. A total amount of 5,068 social actions were registered.

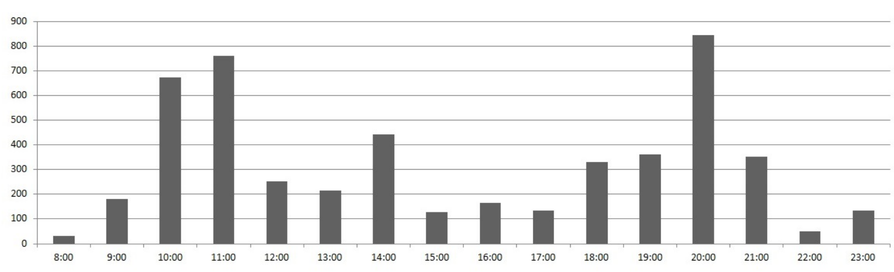

Figure 12: Stage II: Number of social actions versus time of the day.

One of the first conclusions observed during this stage was that the automatic battery charging system results in an increased degree of dependability and decreased maintenance requirements. However, some days the robot remained disconnected because the maintenance person in charge of turning on the robot was not at the hotel.

The distribution of the kind of services requested by the users is reported in Figure 13. The topic chart shows that over 50% of the requested information concerned Sacarino itself. This means that the presence of a humanoid robot attracts people's expectations about social robots and their abilities. Moreover, the information about the hotel facilities and the weather forecast were the most demanded topics, while the help menu and the *"Dialog"* concerning entertainment information requested by voice (such as greetings and jokes that are not accessible through the touch screen) were the least demanded topics.

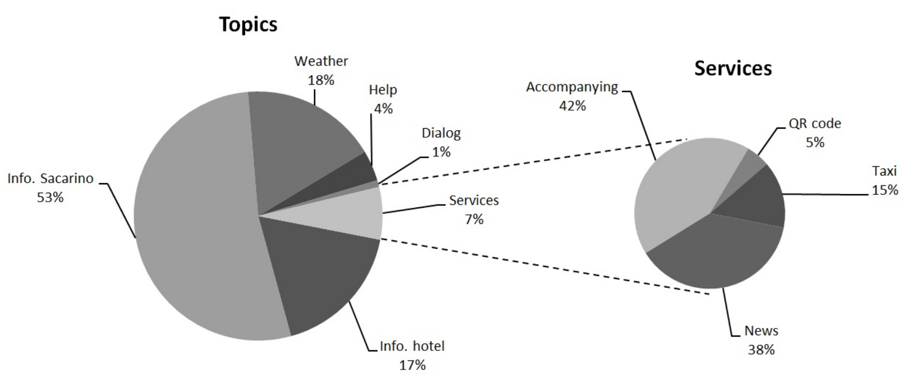

Figure 13: Stage II: Topics distribution.

Concerning the service menu, the most requested service was the one of accompanying, where Sacarino accompanies the guests to the different areas of the hotel, as reported in Figure 13. Another demanded service was the *"news service"*, at which Sacarino randomly voices and displays one of the day's headlines. (Sacarino periodically downloads the daily news and the weather forecast from the Internet).

Regarding the interaction channels, the touchscreen was the most used way for requesting information (95%). Although guests seemed happy with Sacarino's speech, they were reluctant to talk to the robot. Apparently, this is due to two factors:

- Based on our observations in Stage I, we concluded that the users do not seem comfortable when talking loudly to a robot in the presence of other people and/or they feel they do not know how to talk to it. Moreover, touchscreen interaction has been growing fast over the last few years and users are very familiar with it.
- The hotel environment may be very noisy and users tend to maintain a personal distance when interacting with the robot. Both circumstances affect the speech recognition rate and, at the first recognition inaccuracy, the users tend to switch to the touchscreen interaction channel.

5.2. Navigation

Sacarino's navigation is currently restricted to the ground floor, where 7 destinations are defined (see Figure 9). The lobby is flat and has an inverted trapezoid shape (viewed from the hotel entrance), about 20 meters long on the longest side and 8 meters long on the shortest one. The lateral sides are 18 meters. The reception and an adjacent meeting area are located on the right, the dining room and the swimming pool to the front and the elevators and stairway on the left.

Overall, 120 navigation tasks were launched during this stage, including scheduled tasks (events, battery charge...) and on demand accompanying services requested by users. 19 navigation tasks (16%) were discarded before starting (because the emergency stop buttons were activated or the robot was in charging process). 88 navigation tasks were successfully accomplished, including path tracking and, in some cases, connect/disconnect maneuvers at the charging station. Only 15 navigation tasks (12%) were cancelled (v.i.). The distribution of the navigation tasks is reported in Figure 14.

Most navigation task discards were produced because the emergency stop buttons had been pressed. This is an important issue because these buttons are usually pressed by people, mostly children, without any actual malfunction. In other words, the functionalities of the robot were constrained due to misuse by people. Furthermore, Sacarino twice considered its battery level to be insufficient to accomplish the navigation task and, therefore, discarded it, staying at the charge station.

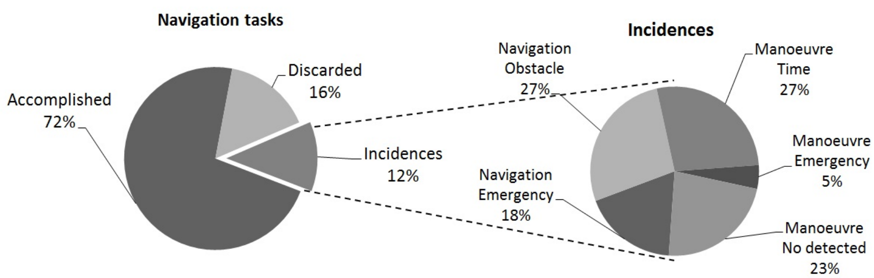

Figure 14: Stage II: Navigation tasks.

Concerning the navigation incidences (Figure 14), 55% (12 cases) were produced in the connect/disconnect maneuver at the battery charger and 45% (10 tasks) during path tracking. The most common problem was a connection timeout, most times caused by a mechanical problem in the pressure sensor located at the charging system.

The most frequent reason to cancel navigation (27%, 6 cases) was that an obstacle had been encountered in the robot's way for an exceeded time. Sometimes, a large localization inaccuracy led the robot too close to some obstacle for a long time (such as a wall or some furniture). In other cases, a guest blocked Sacarino's way for a long time. Sacarino gives out loud warnings to clear its path but, after a given time, it cancels navigation. It is not easy to distinguish whether a given navigation cancellation was actually required or not, but of course the priority of the robot is to avoid any dangerous situation, such as collisions with static or dynamic obstacles, thus ensuring safety. Another reason for the cancellations was that the emergency stops were pressed, in most cases without any particular reason (as in the case of navigation discards).

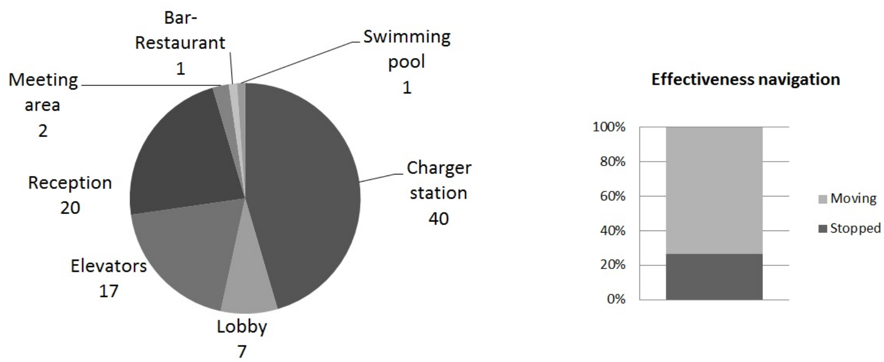

Figure 15: Navigation in Stage II: a) Destinations. b) Effectiveness navigation

The 88 navigation tasks completed covered a total distance of 1,043 meters. The distribution of these tasks, according to their destinations, is shown in Figure 15 (in %). It can be seen that the charger was the final destination of almost 45% of the trajectories. This was expected, because this destination is scheduled twice a day and is also generated asynchronously, at any battery drop under a given threshold.

The navigation effectiveness is shown in the said Figure 15.b. 73% of the navigation time Sacarino was moving, i.e. tracking its path without being stopped at any occasional obstacle.

5.3. Feedback and Improvement at Hardware Level

During stages I and II, the robot hardware was found to be fairly robust. Few malfunctions occurred in this level and, therefore, few modifications have been necessary.

Automatic battery charging system: Most navigation incidences were produced during connection/disconnection maneuvers due to the lack of robustness of a pressure sensor onboard the robot, as has been discussed. (Sacarino cancels the charging maneuver when it is not fully completed within a given timeout). Therefore, the docking system has been readjusted for an improved robustness.

Microphone: One of the outstanding issues found in Stage II was the low use of the voice recognition with respect to that of the touchscreen. There are several possible causes. In any case, we have replaced the existing microphone, installing a more directional one and changing its location from the head to the top of the touchscreen (Figure 16), in order to reduce the input noise coming from the head motors.

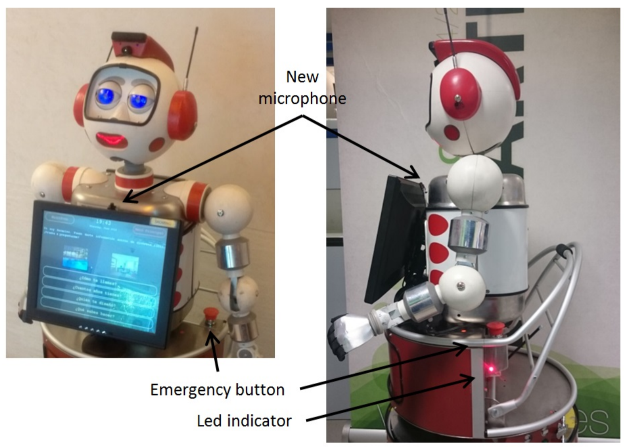

Figure 16: New hardware added.

Emergency buttons: Another issue observed during this Stage is that noticing an unexpected status of the emergency stop was not easy. Therefore, a pair of red LEDs has been added near the emergency buttons, which light up when the buttons are pressed (Figure 6.13). In this way, the hotel staff can readily notice that the buttons have been pressed.

5.4. Feedback and improvement at architecture level

Speech recognition: Speech recognition is an important issue given that speech is, indeed, a prominent communication mechanism among humans. Apart from the fact that users do not seem comfortable when talking loudly to a robot in the presence of other people, the speech recognition ratio is affected by the following evidence:

- Hotel environment is very noisy, in contrast with the lab environment.
- Users tend to maintain a personal distance from the robot.
- Sacarino was continuously analyzing the received sound, looking for any voice request.

In order to deal with these facts, we have added a touch-to-listening mechanism so the user must start recognition by touching a button on the interface. This method, although less natural than direct speech, is effective and well accepted by users who often use it to interact with their mobile devices.

Navigation: Concerning incidences during the navigation tasks, the vast majority of cases were not caused by localization inaccuracy, poor planning or charging maneuver failure, but by the said misuse of the emergency buttons or the presence of static obstacles in the robot's way. The waiting timeouts were tuned higher in order to solve some of these circumstances. Clear screen and voice messages have been added too, in order to inform users about an imminent navigation cancellation if they stay in the robot's path.

5.5. Feedback and improvement at application level

Behavior (Emergency buttons): It is remarkable that the emergency stops were activated during 12% of the total running time of Sacarino (most times without any actual reason). Therefore, a question that arises is how to act when an emergency button is pressed:

- Sacarino can continue with its social skills activated, interacting with people and discarding only navigation tasks when they are requested. With this behavior, Sacarino is available to interact and provide services for longer, but maintenance staff hardly notices that the robot needs assistance.

- On the contrary, if any operational error occurs, all Sacarino's capabilities can become blocked, thus not allowing any interaction. Sacarino can provide only screen and voice messages reporting the need for assistance so that the hotel staff can quickly notice the need (directly or through a guest's warning). However, the robot may continue to provide no service for an extended time unnecessarily.

In the present work, we have aimed at maximizing the robot's functionality, so we have chosen the first option.

6. Stage III: Refinement

This stage extended from May to August 2014. During this period, Sacarino operated in the hotel in the same conditions as in Stage II. The improvement of the robot's performance was derived from the modifications introduced in feedback II, measured using the same quantitative metrics, throughout the same period of time (60 calendar days). The hotel maintenance staff was responsible for turning the robot on/off and, when necessary, recovering it from operational errors (by bringing it back to the charger station).

Sacarino was turned on 51 days (according to the workload of the hotel maintenance-service manager), while in Stage II it was turned on only 23 days. This increase comes from the improvements introduced in the previous stages, which have resulted in enhanced service and degree of dependability (decreasing maintenance personnel requirements). There was also a slight decrease in running time (about 1 hour a day over 8 hours), but the total running time was about twice that of the previous stage (435 hours versus 220).

The overall operation data in this Stage III is summarized in Table 2. Moreover, we have discriminated between working days and weekends following the hotel manager's suggestion. It is worth noting that the guest profile is quite different on working days (generally business travelers) and weekends (families and tourism). Therefore, a separate analysis is useful for adapting the robot services and schedule to the day type.

Table 2: Overall operation time data summary in Stage III

	Days robot working	Total working time	Total social interaction time	Total time in motion tasks	Total number social actions	Total number users/ interactions	Average interaction duration
Working days (mon-fri)	43	350.22 h (8.14 h/day)	35.7 h (10.2%)	6.35 h (1.8%)	9232 (214.69/day)	677 (15.74/day)	2.44 min
Weekends (sat-sun)	8	84.76 h (10.5 h/day)	12.39 h (14.6%)	0.31 h (0.3%)	3616 (452/day)	209 (26.1/day)	2.61 min
Total Stage III	51	434.98 h (8.52 h/day)	48.09 h (11%)	6.66 h (1.5%)	12848 (251.9/day)	866 (17.3/day)	2.58 min

6.1. Social interaction

There was a significant increase in the robot use in Stage III, as can be seen in table 2. A total of 12848 social actions from 886 different users took place, which represents more than twice those corresponding to the previous stage (table 1). Sacarino was interacting with people for 11% of the total time (4% in Stage II) and was running navigation tasks over 1.5% of the total time (1.11% in Stage II). Thus, the time during which Sacarino was busy increased from 5% to 12%. The other 88% of the time, Sacarino was awaiting user demands.

The robot operated on 43 working days (Monday to Friday) and 8 weekend days (Saturday and Sunday). On average, Sacarino's running time (per day) was higher on weekends than on working days. Moreover, more users interacted with the robot on weekends and for more time, as was expected. Concerning navigation, the robot spent more time at the charger on weekends because, in general, there were no navigation tasks scheduled on these days. The improvements introduced in feedback II also resulted in a decrease in the number of robot reboots to less than once a day.

The average number of social actions through the day time is reported in Figure 17. Sacarino was busiest at about 10am and 10pm. The afternoon maximum happened two hours later than in the previous stage due to the different time of the year (spring and summer in the current stage). The average user interaction duration increased from 1.54 to 2.58 minutes, thanks to the increased

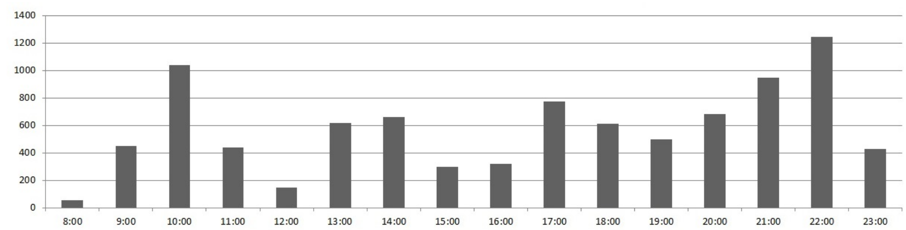

Figure 17: Stage III: Number of social actions versus time of the day.

performance and smoothness of the interaction.

The number of social actions, considering the day type, is shown in Figure 18. The distribution of these actions throughout the day is also different depending on the day type. Noon, after lunch and mid-afternoon are preferred on weekends.

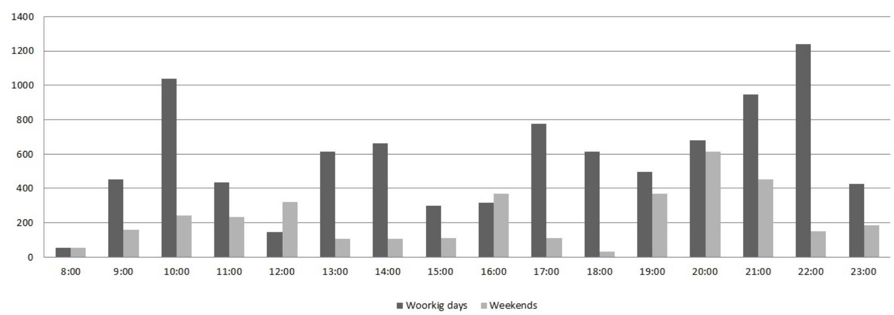

Figure 18: Stage III: Number of social actions versus time of the day (working days versus weekends)

The distribution of the services requested by the users (over 51 days) is reported in Figure 19.a. The chart is similar to that of the previous stage (Figure 13). Information about Sacarino itself is still the most requested topic (43%). However, *"Dialog"* has increased significantly. This topic concerns information that is only accessible through voice (not through the touchscreen). Certainly, this increase owes much to the improvement of the speech recognition system addressed in Stage II. In fact, voice interaction increased from 0.05% in the previous stage to 30% in the current one, as can be seen in Figure 19.b. This increase represents a significant advance, and future efforts should focus on making further improvements, because voice is considered the most natural way of interaction for humans.

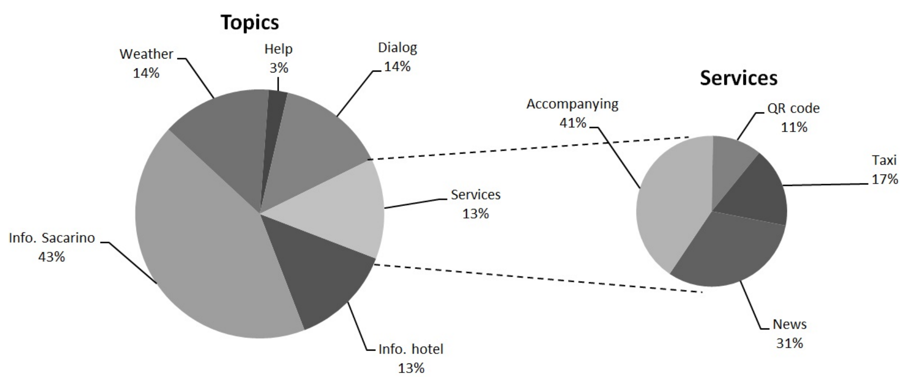

Figure 19: Stage III: a) Topics distribution. b) Interaction channel

6.2. Navigation

The navigation tasks represent about 1% of the total running time of the robot (as in the previous stage). However, a total of 240 navigation tasks were launched, which represents twice the number of navigation tasks launched in Stage II. Their distribution is shown in Figure 20. 5.65% (13 tasks) were discarded before starting, which compares favorably to the 16% found in Stage II. Moreover, the tasks were not completed in 15 cases (6.52%) due to incidences produced during execution. In total, 212 navigation tasks were successfully accomplished, i.e. 88%, thus resulting in an increase of 16% with respect to Stage II.

It is worth remembering that discards do not actually correspond to a navigation system malfunction, but to a good management of the robot's priorities and its state machine. In fact, just 227 of the navigation tasks required movement, 94% of which were completed successfully (while only 6 % finished due to an incidence).

Discarded tasks dropped from 16% in Stage II to 6% in the current stage (see Figure 20), which suggests that the emergency stops remained pressed for

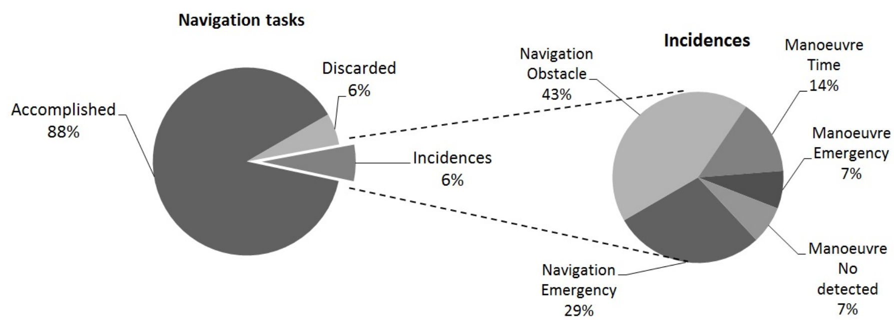

Figure 20: Stage III: Navigation tasks

less time (thanks to the added LED indicators).

The effect of the battery charging system modifications (at hardware and architecture levels) is also noticeable. The maneuver failures decreased from 57% in the previous stage to 28% in the current one. The connection/disconnection maneuvers were not accomplished only 4 times.

Concerning the path tracking, on-the-way task cancellations are similar (in %) to those of Stage II. The most frequent reason to cancel navigation (43%, 3 times) was an obstacle staying in the robot's way for an extended time.

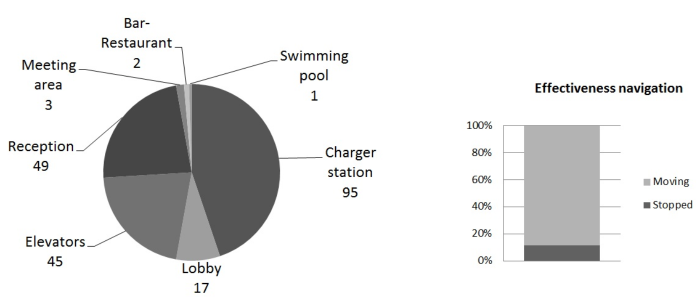

Figure 21: a) Destinations. b) Effectiveness navigation

The 212 navigation tasks completed in this stage covered a total distance of 2,490 meters (more than twice the distance covered in the previous stage). The distribution of the navigation tasks, according to their destinations, remains similar to that of the previous stage, as can be seen in Figure 21.a. The charger continues to be the most common destination.

Finally, the navigation effectiveness increased from 73% to 88%, as can be seen in Figures 15.b. and 21.b. This means that the navigation system improved significantly thanks to the modifications made in feedback II: Sacarino cancels its navigation due to the presence of obstacles less than before.

7. Maintenance

In this section, the dependability of Sacarino is analyzed, thus providing interesting guidelines aimed at the marketing of this technology. The robot must perform a real service with acceptable robustness and dependability without requiring costly maintenance or developers' assistance.

The assessment presented in this section covers stages II and III, the robot operating in normal conditions in the hotel without the attendance of the developers. The setup period involving the definition of services and the construction of the localization and navigation map has not been included.

The introduction of the automatic battery charging system greatly increases Sacarino's autonomy. The hotel maintenance staff is only responsible for turning the robot on and off. Moreover, when a malfunction happens, generally during navigation, a hotel employee has to check whether the emergency stops are pressed or Sacarino is blocked at a stationary obstacle. Then, he/she must unlock the emergency stop (in the former case) or move the robot to an open area (in the latter) and, eventually, to the charging station if Sacarino is unable to relocalize (the navigation system is restarted at the charging station).

Some hardware problems could not be solved by the hotel staff and required our assistance. During Stage II, the staff reported that Sacarino's body had been damaged and lost some of its movements. The neck tilting joint (up/down movements), the right arm servomotor and an eyelid servomotor had been damaged due to external forces (applied generally by children). Springs and compliant joints had been included during feedback II so the problems no longer happened in Stage III.

Furthermore, the hotel staff twice reported that the robot could not be turned on. The reason was that the battery had got fully discharged due to a malfunction of the pressure sensor onboard the robot. The problem was solved by improving the docking system hardware and related software during feedback II. The total cost of man-hours employed on maintenance is around 22 hours.

At the end of stage III, the hotel staff has not reported any other need for assistance, thanks to the described improvements.

8. Lessons Learned

As a result of the field experimentation approached in this work, some important lessons have been learned that may guide researchers and developers of service robots. These lessons are summarized below.

- The robot must be fully autonomous. In a first experimentation stage, Sacarino was placed at the charger station when its batteries were low, but manual connection to the charger station was required. This task, although simple and quick, was sometimes overlooked by the hotel maintenance staff (due to their other duties), thus limiting the running time of the robot. The development of a fully automatic charging station has allowed this problem to be overcome. Moreover, Sacarino's autonomy has increased significantly, given that the robot can charge during idle periods, even while interacting with users. A manual on/off routine is still used because it allows the robot to be disconnected during nights (thus reducing power consumption) and can be assumed by non-technical persons (e.g. the receptionists). Of course, if desired, this operation could be automated by using a simple relay-based circuit permanently supplied with a low power.

- Navigation must be robust, thus avoiding localization losses that may prevent the robot from reaching its destination. A detailed in-field analysis is recommended to address this issue. Many cases can be solved by changing the robot's routes to avoid undesirable situations (routes passing through large featureless areas or near windows, mirrors, one-legged tables and other furniture, crowded areas, etc.). Some landmarks can also be added to the environment to reduce localization errors. In our case, we added a data-matrix and some infra-red LEDs to the charging station to guide the charging maneuver. Moreover, the robot localization is restarted at the charging station (uniquely identified with the said data-matrix) to avoid cumulative errors.

- Voice interaction is especially challenging because of the current limitations of speech recognition and the noisy hotel environment. Multi-modal interaction systems allowing both speech and touch interaction is recommended. In this case, it is also necessary to reconcile the close interaction distance required by the touchscreen with the personal distance used in face-to-face voice interaction. The required compromise can be alleviated by including large touchscreens and fonts. In addition, the interaction should include help and feedback mechanisms so that the user can know about the robot's capabilities and its scope of understanding.

- Another relevant aspect is how to initiate the interaction. It is recommended that the robot greets and incites interaction when a person is detected nearby. However, our observations have shown that speaking too loud may cause reluctance in some guests, given that they become the center of attention for the surrounding people, which may not be pleasant. Starting interaction with a greeting volume adapted to the guest's distance (which can be measured by the laser or other distance sensor) favors interaction.

- The most requested information from Sacarino has been about the robot itself (its capabilities and functionality), followed by requests about the hotel facilities and services. The technological novelty of the robot arouses guests' curiosity, which eases engagement and can be exploited. However, added value services must be incorporated to the robot to prevent users from losing interest over time. Regarding the requested services, the most popular one is accompanying guests through the hotel facilities. This owes much to the expectation generated by robot movement. However, it is assumed that other services, such as taxi calls, check-out or item delivery, will gain relevance in the future, provided that dependable behavior is attained.

- Children get closer to the robot than adults, as expected, and tend to play severely with it. They grab the robot by its arms trying to move them and manipulate the head and related elements (eyelids, mouth), in most cases, without paying any attention to the screen. Therefore, special care should be taken to build a robust robotic platform to prevent damage. Springs and compliant joints are recommended to protect the robot against external forces and, therefore, minimize maintenance requirements. The most critical points are the arms and the neck.

- The robot must be accepted not only by users but also by the hotel staff, looking for their complicity and involvement in the integration of the robot in the hotel's everyday activity. Of course, the staff should not perceive the robot as an inconvenience or a threat to their jobs. Moreover, an easy-to-use interface should be provided so that the staff can remotely plan the robot's activity and modify information contents (menu of the day, events, meetings and conferences, etc).

- The robot must provide added-value services. It is able to attract great attention from users, but they lose interest over time if the robot cannot offer anything else. Quantifiable, added-value services that can be amortized in a short period of time must be provided to introduce this technology in the market.

- A final issue of cardinal relevance concerning the marketing of this kind of robots is that they must comply with all safety standards for human robot interaction, in particular with the already mentioned standard ISO/NP ISO13482. It is important to include the requirements from the initial stages, covering all design levels: hardware, architecture and application. Attention must be paid to battery and electrical isolation mechanisms, power, force control, speed and stopping functions. In some cases, the inclusion of these safety measures penalizes the autonomy of the robot. For example, emergency stop buttons are often pressed by some users (especially children) when there is no collision risk or malfunction. This is a severe problem given that Sacarino cannot be operational until the stop button is manually released. (Obviously, the robot should not, for safety reasons, unlock itself, although it could). Moreover, emergency stops must be clearly visible and accessible, therefore easing misuse. A mechanism that remotely informs the staff about the incident is probably a good solution.

9. Conclusions

In this paper, Sacarino, a service robot, has been evaluated in a hotel environment over long time periods. A three-stage evaluation methodology has been used to evaluate both qualitative and quantitative aspects.

In the first stage, the analysis and evaluation was addressed by directly observing the behavior of the users interacting with the robot and asking for services. This evaluation resulted in a number of guidelines that have been undertaken to redesign the robot at hardware, architecture and application levels.

In the second stage, a number of performance metrics were designed and the appropriate mechanisms for the automated acquisition of these metrics were implemented. Then the robot operated in the hotel in normal conditions, without the presence of the developers. This has allowed the continuous evaluation of the robot over long time periods to be addressed, in real operating conditions, without external influences that could affect the way the user behaves. Hotel staff was only in charge of switching the robot on/off and intervening in special situations (when the emergency-stop button was pressed, or the robot could not reach a given destination due to external disturbances). Metrics concerning social interaction and navigation were evaluated, resulting in guidelines for further hardware, architecture and application improvements.

Finally, in the third stage, the effects of previous improvements have been measured upon the defined metrics, in order to effectively quantify the progress. In total, the robot has been operating in the hotel for 74 days, 9 hours a day, during the second and third stages and has interacted with 17 users a day, on average. Figure 3 is shows a summary of the main results and improvements by stages.

To sum up, the presented *iterative data collection* ⇒ *analysis* ⇒ *redesign methodology* has allowed us to improve the robot's performance through successive refinements, upon continuous experimentation in real conditions that were suspended only during the improvement works in the laboratory. Moreover, a number of lessons have been learned during this extensive experimentation, which can serve as guidelines for robot researchers and manufacturers towards reaching the definitive integration of social service robots into daily life, thus meeting society's demands.

We are currently developing a new robot, so we will be able to switch between it and the existing one during the improvement periods. Thus, continuous robot service at the hotel will be assured. To do so, we will start from the accumulated experience and the already developed functional prototype and we will force the evolution of the prototype to higher levels of technology readiness. We will seek the usage (and definition when necessary) of a wider range of standardized user-centered metrics and evaluation procedures that will serve as research methodology and performance indicators. Future experiments will analyze the long-term functionalities and operation of the robot in a real-world environment and under realistic operational conditions, trying to add complexity to the offered services and improvements on performance effectiveness. Based on all the outcomes, our final goal will be to obtain a ready-for-the-market platform.

Acknowledgments

This work was supported by the Ministry of Science and Innovation, fundamental research project ref. DPI2014-56500-R and Junta de Castilla y León ref. VA001A10-2 and VA036U14.

References

[1] A.L.I.C.E. (Artificial Linguistic Internet Computer Entity) (2015). http://www.alicebot.org/. Accessed: 2015-05-27.

[2] Bainbridge, W., Hart, J., Kim, E., and Scassellati, B. (2008). The effect of presence on human-robot interaction. In *Robot and Human Interactive Communication, 2008. RO-MAN 2008. The 17th IEEE International Symposium on*, pages 701–706.

[3] Brachman, R. and Levesque, H. (2004). *Knowledge Representation and Reasoning*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

[4] Breazeal, C., Kidd, C., Thomaz, A., Hoffman, G., and Berlin, M. (2005). Effects of nonverbal communication on efficiency and robustness in human-robot teamwork. In *Intelligent Robots and Systems, 2005. (IROS 2005). 2005 IEEE/RSJ International Conference on*, pages 708–713.

[5] Choset, H., Burgard, W., Hutchinson, S., Kantor, G., Kavraki, L. E., Lynch, K., and Thrun, S. (2005). *Principles of Robot Motion: Theory, Algorithms, and Implementation*. MIT Press.

[6] Dautenhahn, K. and Saunders, J. (2011). *New Frontiers in Human-robot Interaction*. Advances in interaction studies. John Benjamins Pub.

[7] Dominguez, S., Zalama, E., and Gómez-García-Bermejo (2012). Corrección de la posición mediante marcas planas utilizando el sensor ir del wiimote. In *Actas de las XXXIII Jornadas de Automática*, pages 711–719.

[8] Durrant-Whyte, H. and Bailey, T. (2006). Simultaneous localisation and mapping (slam): Part i the essential algorithms. *IEEE Robotics and Automation Magazine*, 2:99–110.

[9] Eurobotics Aisbl (2015). http://www.eu-robotics.net/.

[10] Fong, T. (2003). A survey of socially interactive robots. In *Robotics and Autonomous Systems 42 2003*, pages 143-166.

[11] Forlizzi, J. (2006). Service robots in the domestic environment: A study of the roomba vacuum in the home. In *In Proc. HRI 2006*, pages 258–265. Press.

[12] Gat, E. (1998). On three-layer architectures. In *Artificial Intelligence and Mobile Robots*. MIT Press.

[13] Gockley, R., Bruce, A., Forlizzi, J., Michalowski, M., Mundell, A., Rosenthal, S., Sellner, B., Simmons, R., Snipes, K., Schultz, A., Wang, J. (2005). Designing robots for long-term social interaction. In *2005 IEEE/RSJ international conference on intelligent robots and systems.* pages 1338–1343.

[14] Hegel, F., Gieselmann, S., Peters, A., Holthaus, P., and Wrede, B. (2011). Towards a typology of meaningful signals and cues in social robotics. In Christensen, H. I., editor, *RO-MAN*, pages 72–78. IEEE.

[15] Holz, D., Iocchi, L., and van der Zant, T. (2013). Benchmarking intelligent service robots through scientific competitions: The robocup@home approach. In *AAAI Spring Symposium: Designing Intelligent Robots*.

[16] ISO13482: Robots and robotic devices – Safety requirements for personal care robots. (2014). Available: http://www.iso.org/.

[17] Kanda, T., Shiomi, M., Miyashita, Z., Ishiguro, H., Hagita, N. (2010). A communication robot in a shopping mall. In *IEEE Trans on Robotics* 26(5):897–913.

[18] Kirby, R., Forlizzi, J., Simmons, R. (2010). Affective social robots. *Robot Auton Syst* 58(3):322–332.

[19] lai Lee, S., Lau, I., Kiesler, S., and Chiu, C.-Y. (2005). Human mental models of humanoid robots. In *Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference on*, pages 2767–2772.

[20] Leite, I., Martinho, C., and Paiva, A. (2013). Social robots for long-term interaction: a survey. *International Journal of Social Robotics.*, 5(20):291–308.

[21] Lester, J., Branting, K., and Mott, B. (2004). Conversational agents. In Singh, M. P., editor, *The Practical Handbook of Internet Computing*. Chapman and Hall CRC, London.

[22] Maartje,M.A. de G. and Somaya, B. A. (2013). Exploring influencing variables for the acceptance of social robots. In *Robotics and Autonomous Systems. Volume 61, Issue 12, December 2013*, pages 1476-1486.

[23] Madhavan, R., Lakaemper, R., and Kalmar-Nagy, T. (2009). Benchmarking and standardization of intelligent robotic systems. In *Advanced Robotics, 2009. ICAR 2009. International Conference on*, pages 1–7.

[24] Martinoia, D. Calandriello, D. and Bonarini, A. (2013). Physically Interactive Robogames: Definition and design guidelines. In *Robotics and Autonomous Systems Volume 61, Issue 8, August 2013*, pages 739-748.

[25] Rautaray, S. and Agrawal, A. (2015). Vision based hand gesture recognition for human computer interaction: a survey. *Artificial Intelligence Review*, 43(1):1–54.

[26] Robotic Operating System. ROS. (2015). [online] Available: http://www.ros.org. Accessed: 2015-05-26.

[27] Rodriguez-Lizundia, E., Marcos, S., Zalama, E., Bermejo, J. G.-G., and Gordaliza, A. (2015). A bellboy robot: Studyof the effects of robot be-

haviouron user engagement and comfort. *International Journal of Social Robotics.*, In press.

[28] Ros, R., Baronib, I. and Demirisa, Y.(2014). Adaptive humanrobot interaction in sensorimotor task instruction: From human to robot dance tutors. In *Robotics and Autonomous Systems, Volume 62, Issue 6, June 2014*, pages 707-720.

[29] Sabanovic, S. (2010). Robots in society, society in robots - mutual shaping of society and technology as a framework for social robot design. *I. J. Social Robotics*, 2(4):439–450.

[30] Severinson-Eklundh, K., Green, A., Httenrauch, H. (2003). Social and collaborative aspects of interaction with a service robot. *Robot Auton Syst*, 42(34):223234

[31] Sim, D. Y. Y. and Loo, C. K. (2009). Extensive assessment and evaluation methodologies on assistive social robots for modelling humanrobot interaction a review. *Information Sciences*, 301(20):305–344.

[32] Speech Recognition System (ASR) Loquendo (2015). http://research.nuance.com/. Accessed: 2015-06-22.

[33] Steinfeld, A., Fong, T., Kaber, D., Lewis, M., Scholtz, J., Schultz, A., and Goodrich, M. (2006). Common metrics for human-robot interaction. In *HRI '06: Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*, pages 33–40, New York, NY, USA. ACM.

[34] Tec-Shs (2008). *Technology Readiness Levels Handbook for Space Applications.* European Space Agency.

[35] Wickens, C. D. and Hollands, J. G. (2000). *Engineering psychology and human performance (fourth edition).* Prentice-Hal., Upper Saddle River, NJ, USA.

[36] Wisspeintner, T., van der Zant, T., Iocchi, L., and Schiffer, S. (2009). RoboCup@Home: Scientific competition and benchmarking for domestic service robots. *Interaction Studies*, 10(3):392–426.

[37] Zalama, E., García-Bermejo, J. G., Marcos, S., Dominguez, S., Alonso, R. F., Pinillos, R., and López, J. (2013). Sacarino, a service robot in a hotel environment. In *ROBOT 2013: First Iberian Robotics Conference - Advances in Robotics, Vol. 2, Madrid, Spain, 28-29 November 2013*, pages 3–14.

Table 3: Overall assessment summary

Stage I	**Setup and Study on Proxemics**	
Results		
Users: 95	Average interaction duration: 0,59 minutes	
Feedback		
Level	**Weaknesses**	**Improvements**
Hardware	Servomotors damage s due to misuse.	Add compliant mechanism.
	Users interaction distance is not closer.	Increase touchscreen size (17-inch).
	Lack of robot autonomy .	Develop automatic battery charging system.
	Odometry drifts.	Add Gyroscope.
Architecture	Navigation and localization is not good enough .	Use topological map to planning trajectories .
		Sacarino stops at obstacles and waits until they are removed.
Application	Behavior.	The robot takes the initiative for interaction.
	Lack of simplicity and clarity of the interface.	Design menus and buttons more intuitive.
Stage II	**Metrics Assessments**	
Results		
Users: 349 (15,17/day)	Average interaction duration: 1,54 minutes	
Feedback		
Level	**Weaknesses**	**Improvements**
Hardware	Microphone: input noise from the head motors.	Microphone is relocated.
	The hotel staff can readily notice that the Emergency buttons have been pressed.	Add LEDs status .
Architecture	The speech recognition ratio is low.	Include touch-to-listening button on the interface to start recognition.
	The majority of incidences during the navigation tasks due to presence of obstacles.	Increase timeouts before cancellation.
		Add clear screen and voice messages to prevent user before cancellation.
Application	Behavior.	Sacarino can continue with its social skills even though the emergency buttons are pressed
Stage III	**Refinement**	
Results		
Users: 349 (15,17/day)	Average interaction duration: 1,54 minutes	