

A Living Labs Approach for Usability Testing of Ambient Assisted Living Technologies

Miguel Sales Dias^{1,4}, Elisângela Vilar^{2(✉)}, Filipe Sousa³,
Ana Vasconcelos³, Fernando Miguel Pinto¹, Nuno Saldanha¹,
and Sara Eloy⁴

¹ Microsoft Language Development Center, Lisbon, Portugal
{Miguel.Dias, t-fecos, t-nunosa}@microsoft.com

² Centre for Architecture, Urban Planning and Design (CIAUD),
Rua Sá Nogueira, Pólo Universitário, Alto da Ajuda, 1349-055 Lisbon, Portugal
elipessoa@gmail.com

³ Fraunhofer Portugal Research, Porto, Portugal
{filipe.sousa, ana.vasconcelos}@fraunhofer.pt

⁴ ISCTE Instituto Universitário de Lisboa, Lisbon, Portugal
sara.elay@iscte-iul.pt

Abstract. This paper presents usability tests results with real users during the prototype development phase of two applications for seniors care, AALFred and SmartCompanion. To this aim, usability testing was performed considering a Living Lab approach. Seniors were invited to use the applications in an environment that simulates the one they would use the tested technology during their everyday life. Observation methods, thinking aloud and questionnaires were used to collect data related to the systems' effectiveness and users' satisfaction, namely their expectations, frustrations and difficulties. Evaluations were performed during the initial phases of product development and results were used to improve the applications, considering the development cycle of User-Centered Design methodology.

Keywords: Usability tests · User-centered design · Ambient assisted living · Living labs

1 Introduction

Global projections [1] point to a significant increase in the aging rate. These projections estimate, for instance, that the proportion of people with 60 or more years old will duplicate between 2009 and 2050, reaching two billion in 2050. This trend is also reflected in the increase of chronic diseases incidence such as Diabetes, which should reach more than 340 million people in 2030 [2].

In this context, it is clear that an active intervention to promote the several levels of social integration is needed. Ambient Assisted Living (AAL) is a new approach to the needs of population aging, with the main goal of applying the technologies of ambient intelligence in helping people with specific demands, and in building safe environments for the maintenance of independent living [3]. Considering this, two applications developed and integrated by partners in the scope of the QREN AAL4ALL – Ambient

Assisted Living for ALL interoperable and standardized platform [10], are the main focus of analysis in this paper, AALFred, from Microsoft and the PaeLife consortium partners [4], and the Smart Companion [5], from Fraunhofer Portugal AICOS (FhP). To engage real users in the various stages of development, both applications were developed using a user-centered design methodology [6].

The main objective of this paper is therefore, presenting usability testing with real users during the prototype development phase. For this, a Living Usability Labs methodology [7] was used. According to Schumacher [8], the concept of Living Labs is based on a systematic co-creation, directed to the user, which integrates the research and innovation processes. These processes are integrated throughout the development, exploitation, experimentation and evaluation of ideas, scenarios, concepts, products/services in real-life utilization.

In the following sections, this paper will present the usability testing performed with both AALFred and Smart Companion, describing the adopted methodology, main findings and improvements achieved with the applications' re-design.

2 The Applications

Two applications directed for senior's use, were the focus of usability tests, namely, AALFred and Smart Companion.

2.1 AALFred

AALFred is a personal life assistant that helps and guides users in the access to ICT's, developed by Microsoft and the PaeLife consortium partners [3] and improved and integrated on AAL4ALL ecosystem in the scope of AAL4ALL project. Older users interact with AALFred, via speech (in European Portuguese) and touch, using a Windows 8.1 Tablet. With AALFred, messaging with friends and relatives exchanged in social media (Facebook, Twitter), email, agenda (Outlook) and audio-videoconferencing calls (Skype) can be easily accessed and used to make seniors more active, engaged in social and community life and therefore less isolated. Additionally, interesting information such as news and access to nearby services, such as informal and formal healthcare, pharmacies and authorities, is delivered in an integrated way.

2.2 SmartCompanion

Smart Companion is an Android launcher developed by Fraunhofer Portugal AICOS (FhP-AICOS). It is an Android customization that was specially designed to address seniors and caregivers' goals and needs. Its main objective is to facilitate the use of a smartphone by reducing its complexity. In this way, Smart Companion aims to promote the use of smartphones by seniors during their everyday activities, considering several tools, from messages to medication reminders, activity monitoring, fall risk analysis and fall detection.

3 Methodology

For the usability evaluation of the two selected applications (i.e., AALFred and Smart Companion), personas and scenarios were identified and tasks were defined, considering the User-product interaction. The Living Lab methodology was considered in order to make the simulated use closer to the real experience seniors would have, if they were performing the same tasks in their everyday life.

3.1 Data Collection

All data was collected through direct observation and questionnaires. For direct observation, a set of tasks was pre-defined during brainstorming meetings with the development team, and experts in Psychology, Ergonomics and Software Engineering. From this, an observation form was specifically developed to collect metrics such as task execution time, task completion rate (and how easily the participant completed the task), assistances during task completion, and the participant's visible emotional state. All tasks were decomposed into activities that were evaluated separately. Demographic data was collected at the beginning of the usability test through a questionnaire. A satisfaction questionnaire was also developed and applied after the completion of each pre-defined task. These questionnaires were developed by the project team and were pre-tested during the initial phases of the AAL4ALL Project. A Usability Scale, based on the International Classification of Functioning (ICF) was developed by the AAL4ALL partners and was also applied at the end of each task completion.

3.2 Protocol

Usability testing was performed considering a task-oriented analysis in which participants were asked to perform predefined tasks. Before the test, all participants were asked to sign a Consent Form. It was mandatory, in a way that if a participant did not sign the consent form, he/she was not allowed to continue the test.

After signing the Consent Form, participants were asked to fill out the demographic questionnaire together with technology-related and quality of life-related questions.

Then, a facilitator presented to the participants all AAL4ALL products involved in the usability test. In this phase, the facilitator explained to the participants the main functionalities of the products.

After the presentation phase, the facilitator started the task-related phase. At this moment, the facilitator explained to the participants that he/she should accomplish tasks with the presented products without time restrictions. Tasks were given in a sequential manner, in random order, in a way that only after completing one task and fulfilling the satisfaction questionnaire and the usability scale related to that task, the facilitator started with a new task with the participant. Each task was read by the facilitator and also delivered to the participant written on paper. The usability test finished when all tasks were given and completed by the participant.

3.3 Usability Testing with AALFred

The evaluation took place at Microsoft's Living Lab in Lisbon, Portugal. It simulates a regular house living room, in which participants will interact with AALFred through the use of a tablet and a TV, simulating real situations that seniors may face.

Scenarios and Personas. For usability testing of ALLFred, two main scenarios were considered: Monitoring seniors at home and social integration, entertainment and communication of senior users. Personas were defined based on a previous study of the AAL4ALL consortium considering the target market, and participant selection was done based on these persona definitions, which can be seen on Table 1.

Table 1. Personas, characteristics and criteria for participant's selection for usability testing with AALFred.

Personas	Characteristics	Criteria for participants selection
Teresa	She is no longer able to regulate body temperature as before. It occurs because her response to body temperature changes are not normal, consequently she is not able to normally react to hypothermia situations. Some parts of the body such as hands and toes are more critical and need special attention regarding temperature regulation	<ul style="list-style-type: none"> • Senior (≥ 70 years old) • Shoe size (42) • No speech impairments • No hearing impairments • With limited mobility
Formal caregiver	Health care professional responsible for seniors' monitoring	<ul style="list-style-type: none"> • Health care professional • No motor impairments • No visual impairments • Good technological literacy

Participants. For the usability testing four senior participants were recruited considering the selection criteria. The average participant age was 79.50 (SD = 5.74; min. 76, max. 88 years old), and all of them were female. All participants were retired. Half of the participants had an elementary school level education, and others had a high school level. All participants had used a cellphone, but none of them had ever used a smartphone. Half of the participants reported a moderate use of the cellphone while 33 % reported a frequent use. Regarding the usage of a computer, most of the participants (83 %) did not use it.

In addition to the demographic and technology-related data, participants were also asked to self-evaluate their quality-of-life and memory abilities. Most of participants were completely satisfied (33 %) or very satisfied (50 %) with their ability to carry out

their daily activities and evaluated their quality of life as good (50 %) or very good (33 %). Considering memory abilities, all participants reported it as good. Most of the participants reported being apprehensive (67 %) before the test. The others stated being motivated to perform the tests (33 %).

Tasks. For AALFred usability tests, three external devices were also used: the Smartshoe, a chestband, and a body temperature and pulse sensor.

The Smartshoe is a device that aims to contribute with feet temperature regulation. It has sensors integrated on the shoes that reads room temperature and humidity (outside the shoes) and regulates feet temperature (temperature inside the shoes) according to the external data. The values can be observed through AALFred’s user interface, since it is linked to AALFred via wireless means (Bluetooth).

The Chestband by Plux¹, is a device that enables the remote monitoring of a senior’s respiratory rate, electrocardiogram (ECG) and home location (via inertia unit sensing) by a caregiver. It is connected with AALFred via wireless (Bluetooth) and can be accessed also by the caregiver during an AALFred-to-AALFred Skype call over the Internet.

Another device considered for the test was a set of sensors (EXA ALL-in-One²) that collect body temperature and heart rate for personal and remote monitoring.

All devices (Smartshoe, Chestband, EXA ALL-in-One) and AALFred communicate over the internet with peer remote ALLFred apps (used by formal or informal caretakers), via the AAL4ALL interoperable and standard service-oriented communication architecture.

Three main tasks were defined and were presented to the participants considering a hypothetical scenario:

Task 1 – “Imagine you arrive home after buying your Smartshoes that regulate the temperature of your feet. Your task is to turn them on, tell me what is the internal temperature of the shoes and the humidity of the environment, and finally charge their battery”.

Task 2 – “You have installed at home a system called Chestband that allows your doctor, or a family member to observe some of your vital signs and talk to you through a Skype video call. Your task is to answer a video call on the tablet and put the Chestband so that your data is sent to your doctor/family member”.

Task 3 – You need to frequently monitor your vital signs, so you use the “EXA ALL-in-One “ device. Your task is to check and inform me about your temperature and your pulse using the “EXA All in One”.

3.4 Usability Testing with SmartCompanion

The evaluation took place at Fraunhofer’s Living Lab, more specifically in its living room and all participants were volunteers from Colaborar³ network.

¹ <http://www.plux.info>.

² <http://www.exatronic.pt/en/home/>.

³ <http://colaborar.fraunhofer.pt/en/>.

Scenarios and Personas. For the usability testing of Smart Companion five main scenarios were considered: Time management, Time Management with TV, Mobility - Navigation, Mobility – Activity Monitoring and Medication Reminders.

The Personas defined for usability testing with the Smart Companion can be seen on Table 2.

Table 2. Personas, characteristics and criteria for participants' selection for usability testing with the Smart Companion.

Personas	Characteristics	Criteria for participants selection
Ana	Active and healthy, with an autonomous life. She practices exercise everyday	<ul style="list-style-type: none"> • Senior (≥ 65 years old) • Active and autonomous life • No speech impairments • No hearing impairments • No mobility limitations
Maria	She lives in a nursing home. She has a chest angina, which requires daily medication. She also has memory loss, so she is afraid of forgetting to take her medicines	<ul style="list-style-type: none"> • Senior (≥ 65 years old) • Can read and write • Without severe vision limitations • Without severe speech limitations • Without severe hearing limitations • No experience in dealing with cellphones, computers and internet use

Participants. For the evaluation 12 senior participants were recruited as volunteers. The average participant age was 68.42 (SD = 3.40; min. 64, max. 76 years old), eight male and four female. Almost all participants were retired, with the exception of one participant who was self-employed. Most participants had high-school level education. All participants had a cellphone, 67 % of which were smartphones. 83 % of the participants reported a frequent and moderate use of the cellphone, while 17 % reported little use. Regarding computer use, most of the participants (83 %) were users and only 17 % did not have a computer.

Considering the participants' self-evaluation of their quality-of-life and memory abilities, all participants were completely or very satisfied with their ability to carry out their daily activities, however 50 % of the participants evaluated their quality of life and their memory only as reasonable. All participants but one stated being motivated to perform the tests and were confident on their abilities to work with the tested technologies.

Tasks. All tasks were selected as a representation of the main features of each application and the most frequently used as well. For each task the participants were given instructions on what they were expected to do, given a hypothetical scenario:

Task 1: "Imagine that you decide to invite your friends for a lunch at your house. Your task is to use your smartphone to choose the best day to schedule the lunch,

considering the week of the 19th to the 23rd of May, schedule the appointment and invite your friends Maria and Ana.”

Task 2: “[The participant was told to watch some TV and let the facilitator know if something out of the ordinary happened] Please read what appeared on the TV and dismiss it when you are ready”.

Task 3: “Imagine that you need to go to a new place and you don’t know how to get there. You will use your smartphone’s navigation system to help you get there. Your task is to introduce the address in the system and try to understand the directions”.

Task 4: “In your phone you have an application that monitors your daily activity. One of your doctor’s recommendations was that you should do a 30 min walk at least three times a week. To know when was the last time you accomplished this goal you can check your activity history. Your task is to verify if you did a 30 min walk today and how many times you did it last week”.

Task 5: “During your last medical appointment the doctor prescribed you a new medication that you are not used to, and therefore you are afraid that you will forget about it. You can use the medication reminder application in your smartphone to ensure that you do not forget to take the medicine. Your task is to create a new medication reminder on the phone”.

Task 6: “[The participant was told to watch some TV and let the facilitator know if something out of the ordinary happened] Please read what appeared on the TV and dismiss it when you are ready”.

Task 7: “You can’t remember if you took the medicine you were supposed to on May 9th. Therefore you decide to check you intake history. Your task is to check your intake history on the smartphone”.

4 Results

4.1 ALLFred

Performance Results. The overall performance results suggest that the apps did not present major usability issues, with less than half tasks ($n = 2$) requesting assistance during the test and a mean unassisted task completion effectiveness of 91.7 % ($SD = 13.9$; $min = 66.7$; $max = 100$). Table 3 summarizes the performance results.

Even though we did not have a comparison threshold for the time needed to perform the tests, the average time of approximately 6 min for all tasks ($SD = 2.04$; $min = 5$; $max = 10$) can be considered a reasonable value.

Each task was divided into activities (A1.. AN) that were scored according to their completion and how easily it was for the participant (easily completed, completed, completed with difficulty, not completed). In general, all activities were completed without a challenge. The only exception that can be noticed happened in task 3, with the activities A4 and A5 not being completed by half of the users.

Activity 4 of task 3 refers to the access, through AALFred’s interface, to the “Vital signs” item using touch and/or speech commands. Activity 5 refers to the verbalization of the values for the body temperature and pulse.

Table 3. Summary of performance results for AALFred

User #	Unassisted Task Effectiveness [(%)Complete]	Duration (min) ¹	Assists
P1 – task 2	100	5	0
P1 – task 3	100	5	0
P2 – task 1	83.3	5	1
P3 – task 3	66.7	5	2
P4 – task 1	100	5	0
P4 – task 2	100	10	0
Mean	91.7	5.8	0.5
Standard Deviation	13.9	2.04	0.8
Min	66.7	5	0
Max	100	10	2

The difficulty with task 3 (activities 4 and 5) however could be due to a lack of comprehension about the terms used. According to a user’s verbalization, “the terms and icons are difficult to understand”. As they were sequential activities, not completing activity 4 would lead to difficulties in completing activity 5.

Main verbalizations were related to failures on voice commands, font size, color scheme, and difficulties on understanding the meaning of terms and icons.

Satisfaction Results. The overall results of the satisfaction questionnaire related to the users’ satisfaction with the app and perceived usefulness indicate that all participants understood the service as a potential benefit on their daily lives and something that they would be willing to acquire and learn how to use. All participants considered that the presented solution facilitates the health monitoring and 83 % of the participants agree that it promotes relaxation and would be willing to pay for the service.

Usability Scale Results. The average score for the Usability Scale was 22,5 (SD = 6,12; min = 15; max = 30), which, considering that the maximum score is 30, indicates a successful result. Figure 1 presents the scores attained for all tasks.

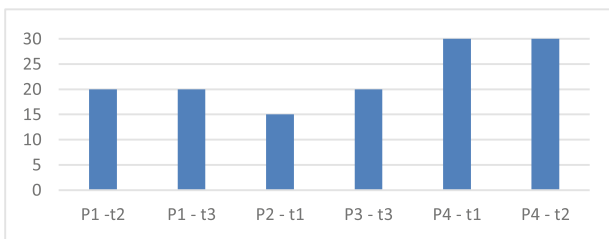


Fig. 1. ICF-Usability scale score by participant

4.2 SmartCompanion

Performance Results. The overall performance results suggest that the products did not present major usability issues, with less than half the users requesting assistance during the test and a mean unassisted task completion effectiveness of 88.1 % (SD = 15.92; min = 57.14; max = 100). Table 4 summarizes the performance results.

Table 4. Summary of performance results for Smart Companion

User #	Unassisted Task Effectiveness [(%)Complete]	Duration (min) ¹	Assists
P1	100	15	0
P2	71.43	17	2
P3	85.71	15	1
P4	100	13	0
P5	100	23	0
P6	100	12	0
P7	100	15	0
P8	100	12	0
P9	57.14	19	4
P10	100	21	0
P11	71.43	21	2
P12	71.43	21	2
Mean	88.10	17	0.92
Standard Deviation	15.92	3.88	1.31
Min	57.14	12	0
Max	100	23	4

The average time for all tasks completion was 17 min for all tasks (SD = 3,88; min = 12; max = 23) which can be considered also a reasonable value.

In general, all activities were completed without challenge. The only exception that can be noticed happened in task 1, with the activities A2 and A3 not being completed by half the users. This fact however was due to deviations in the activity flow, i.e., there were different flows that allowed the participants to achieve the same result. In this case, half of the users chose a flow which did not require the completion of these two specific activities but that produced result required to complete the task.

Satisfaction Results. The overall results of the satisfaction questionnaire related to the users satisfaction with the app and perceived usefulness indicate that all participants understood the service has a potential benefit on their daily lives and something that they would be willing to acquire and learn how to use. Participants also indicated the three most valued characteristics of this service as the immediate access to information, the easing of a problem's resolution and the ease of use.

Usability Scale Results. The average score for the Usability Scale was 24.6 (SD = 2,64; min = 19; max = 27), which, considering that the maximum score is 30, indicates a successful result. Figure 2 presents the scores attained for all tasks.

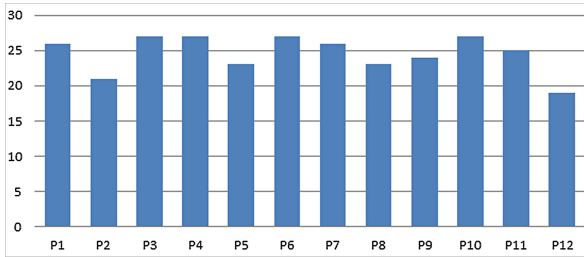


Fig. 2. ICF-Usability Scale score by participant



Fig. 3. Main changes in GUI of ALLFred: left – GUI used during the studies reported on this paper; right – new GUI (darker theme) that took into account the learning from this study.



Fig. 4. Current version of Smart Companion

5 Conclusions

The usability testing during the prototype phase of app development, allowed a better understanding of senior users' needs and expectations. From this analysis some changes were considered for the Graphical User Interface (GUI) and information architecture of both apps.

For AALFred, main changes were made considering the: (i) color scheme and font size/type to improve contrast and legibility, (ii) icons design to a better information comprehension, and (iii) rearrangement of the layout elements to meet users' expectations and to be updated using Windows 8 look and feel and design guidelines. Figure 3 shows AALFred's evolution with re-designed icons and a new with a new layout, with a more comprehensive flow of information. Users were also given choice to customize their user experience by having access to GUI parameters like multiple application themes (dark/light), list orientation (vertical or horizontal scrolling), or speech interaction customization (multiple synthetic voices). This is easily configurable via a step-by-step wizard available when the app first starts and in the preferences screen.

New usability tests performed with most recent version of AALFred, yet to be published, have shown a clear choice of the darker theme when compared to the clearer one. The clearer theme was perceived as harder to read and understand by the seniors.

As for the Smart Companion, the results of this evaluation indicated that the proposed solution meet the criteria to be considered suitable for the target users. Participants' performance did not vary significantly, and the low number of assistances and high completion rate are considered as positive results. The only task including activities that were not completed, only suggests that the alternative flows allow users to also easily complete tasks. Since satisfaction results were also positive, there were no recommendations to change the application design at the time of the usability testing during the prototype phase of development (Fig. 4).

We believe that these results are a direct consequence of the use of a user-centered methodology adopted early in the apps design and development. Since requirements elicitation phase to the early prototype phases, both AALFred and Smart Companion have been built on users' input and according to design guidelines for this target audience [9]. Since the tested design had already gone through a series of iterations that included usability tests and redesign of the user interface, the above presented results validate the current design. Both AALFred and SmartCompanion are currently being used in extensive field trials in various regions of Portugal, in the context of QREN AAL4ALL, whose results will be subject to further publication. Future work for these apps includes the development of new features covering seniors' needs and iterations to improve new designs.

Acknowledgments. Portuguese Project AAL4ALL (<http://www.aal4all.org/>), co-funded by the European Community Fund FEDER through COMPETE – Programa Operacional Factores de Competitividade (POFC).

References

1. United Nations: World Population Ageing 2009. United Nations, New York (2009)
2. World Health Organization: Diabetes action now. In: W. H. Organization (ed.): World Health Organization and International Diabetes Federation (2004)
3. Kleinberger, T., Becker, M., Ras, E., Holzinger, A., Müller, P.: Ambient intelligence in assisted living: enable elderly people to handle future interfaces. In: Stephanidis, C. (ed.) UAHCI 2007 (Part II). LNCS, vol. 4555, pp. 103–112. Springer, Heidelberg (2007)
4. AAL PaLife. <http://www.aal-europe.eu/projects/paelfife/>. Accessed 6 Dec 2014
5. Smartcompanion. <http://smartcompanion.projects.fraunhofer.pt/>. Accessed 6 Dec 2014
6. International Organization for Standardization (ISO): Human-centred design processes for interactive systems. ISO 13407:1999. International Organization for Standardization, Geneva, Switzerland (1999)
7. Pacheco da Rocha, N., Queirós, A., Oliveira, C., Teixeira, A., Pacheco, O., Pereira, C., Martins, A. I.: Integrated development and evaluation of innovative AAL services; A Living Lab approach. Paper presented at the Information Systems and Technologies (CISTI), 2013 8th Iberian Conference on Information Systems and Technologies, Lisbon, Portugal, 19–22 June 2013 (2013)
8. Schumacher, J.: Alcotra Innovation project: Living Labs. Definition, Harmonization Cube Indicators & Good Practices (I. Alcotra, Trans.), Alcotra (2013)
9. Correia de Barros, A., Leitão, R., Ribeiro, J.: Design and evaluation of a mobile user interface for older adults: navigation, interaction and visual design recommendations. *Procedia Comput. Sci.* **27**, 369–378 (2014). Elsevier B.V.
10. QREN AAL4ALL, Ambient Assisted Living for ALL. <http://www.aal4all.org/>. Accessed 6 Dec 2014