AN END-USER PERSPECTIVE ON THE CAMI AMBIENT AND ASSISTED LIVING PROJECT

I.A. Awada[1], O. Cramariuc[2], I. Mocanu[1], C. Seceleanu[3], A. Kunnappilly[3], A.M. Florea[1]

[1] *University Politehnica of Bucharest, Bucharest (ROMANIA)*
[2] *Centrul IT pentru Stiinta si Tehnologie, Bucharest (ROMANIA)*
[3] *Mälardalen University, Vasteras (SWEDEN)*

Abstract

In this paper, we present the outcomes and conclusions obtained by involving seniors from three countries (Denmark, Poland and Romania) in an innovative project funded under the European Ambient Assisted Living (ALL) program. CAMI stands for "Companion with Autonomously Mobile Interface" in "Artificially intelligent ecosystem for self-management and sustainable quality of life in AAL". The CAMI solution enables flexible, scalable and individualised services that support elderly to self-manage their daily life and prolong their involvement in the society (sharing knowledge, continue working, etc). This also allows their informal caregivers (family and friends) to continue working and participating in society while caring for their loved ones. The solution is designed as an innovative architecture that allows for individualized, intelligent self-management which can be tailored to an individual's preferences and needs. A user-centred approach has ranked health monitoring, computer supervised physical exercises and voice based interaction among the top favoured CAMI functionalities. Respondents from three countries (Poland, Romania and Denmark) participated in a multinational survey and a conjoint analysis study.

Keywords: user-centred design, artificial intelligence, fall detection, exergames, vocal interface

1 INTRODUCTION

In the context of unprecedented worldwide demographic changes, information and communication technologies (ICT) are increasingly sought for their potential to help aging adults and seniors to live independently in their home environment. While innovation in this field is rapidly picking up, the full impact of such technologies can be attained only through a wide spread adoption of such technologies by the elderly population. Consequently, several initiatives at both national and cross-national level are actively supporting the development of AAL ICT through a user-centered approach that is expected to increase acceptance and reduce learning barriers.

In this paper, we present the outcomes and conclusions obtained by involving seniors from three countries (Denmark, Poland and Romania) in a project funded under the European Ambient Assisted Living (ALL) program. CAMI stands for "Companion with Autonomously Mobile Interface" in "Artificially intelligent ecosystem for self-management and sustainable quality of life in AAL" [1]. The project consortium comprises eight SME's and universities from five European countries which are developing a fully integrated AAL solution at the overlap of tele- care and health, smart homes and robotics (see Figure 1).

The CAMI solution enables flexible, scalable and individualized services that support elderly to self-manage their daily life and prolong their involvement in the society (sharing knowledge, continue working, etc) while allowing their informal caregivers (family and friends) to continue working and participating in society whilst caring for their loved ones. The solution is designed as an innovative architecture that allows for individualized, intelligent self-management which can be tailored to an individual's preferences, culture, level of comprehension, skill, educational needs and learning style. A combination of advanced reasoning algorithms with a voice-based interface is easing the self-management and decision-making process. The decision support module is an expert system that acts as the central integration point for CAMI. It is a dockerized container, deployed in the cloud.

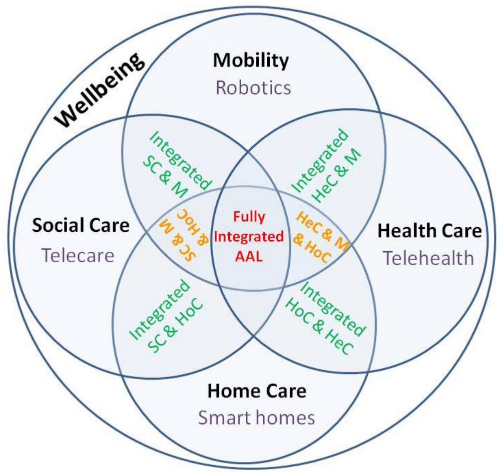

Figure 1. Venn diagram of a fully integrated AAL solution.

A user-centred approach has ranked health monitoring, computer supervised physical exercises and voice based interaction among the top favoured CAMI functionalities. Respondents from three countries (Poland, Romania and Denmark) participated in a multinational survey and a conjoint analysis study. E-health solutions were perceived as very useful and 59% of the respondents considered the graphic display of various health measurements (e.g. blood pressure, heart rate, oxygen levels) as an interesting feature. The ability to share health measurements with various doctors was perceived as useful by 60% of the respondents. Computer supervised physical exercises ranked third in the conjoint analysis. All the voice control functionalities received more than 50% votes which classifies them as promising features.

The physical exercise module is based on exergames with two avatars: the training avatar and the avatar of the user. The training avatar is performing different physical exercises which have to be reproduced by the user. The user's movements are compared with the movements of the avatar based on algorithms for comparing two nonlinear series. At the end of the exercise, the user receives a score that reflects the correctness of the performed exercises. The results are saved in a database, such that they can be analysed by a caregiver or by a medical specialist.

The voice interface is composed of five main parts: Automatic Speech Recognition, Natural Language Understanding, Dialog Management module, Natural Language Generation, and Text to Speech synthesis. Since, the user should always be able to interact with the CAMI platform, regardless of the status of the internet connection, the vocal interface should work also offline ensuring basic system functionalities. Therefore, the vocal interface will have two working modes: (1) an online mode that depends on internet connectivity and (2) an offline mode (a limited version) that doesn't depend on internet connection.

2 AN OVERVIEW OF THE CAMI PLATFORM ARCHITECTURE

The architecture of the CAMI system is developed aiming at the seamless integration of various assisted-living functionalities, like health monitoring, fall detection, home monitoring, robotic platform support etc., ensuring modularity and re-use. A detailed description of CAMI architecture was presented in our previous work [2], [3]. In this paper, we present a smaller implemented version of the system architecture as shown in Figure 2. The integration of various functionalities is achieved by three main modules: a) CAMI Gateway (running on the SNG-Gateway developed by the Eclexys Sagl partner, b) CAMI Cloud, and c) CAMI multi-modal user interface (e.g. the 3rd party health platforms (Linkwatch [4] and OpenTele [5] and vocal interaction with the CAMI Cloud).

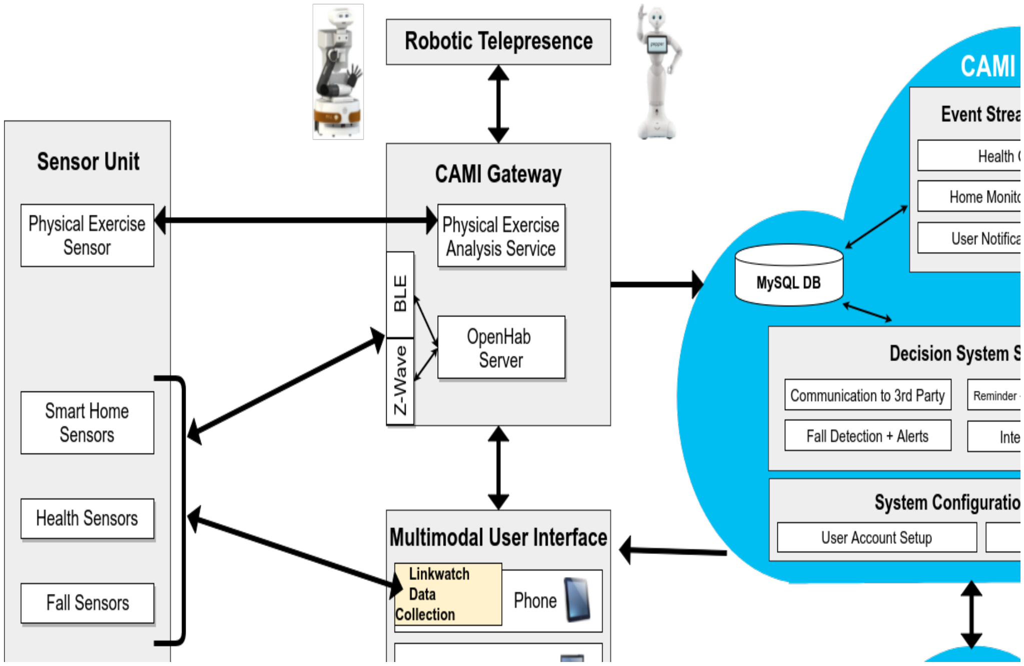

Figure 2. CAMI architecture block diagram.

The **CAMI Gateway** connects with a multitude of Bluetooth and Z-Wave compatible health measurement and home monitoring sensors (e.g. the A&D UA-651 BLE blood pressure meter, the Onyx II Model 9560 oxymeter, Fibaro Temperature and Motion Sensor FGMS-001, Z-Wave 3 in 1 Sensor (temperature, illumination, door) PHI_PSM01), Vibby fall detection sensor etc.

The **CAMI Cloud** allows the system to perform two other essential integrations:

- From the input perspective: the CAMI Cloud allows access to information from sensors that publish their data directly to the cloud service (e.g. Fitbit bracelet, WiFi weight scale, smartwatches etc.)
- From an output/sharing perspective: the CAMI Cloud allows dissemination/replication of data to other health monitoring platforms (e.g. Linkwatch, OpenTele). This type of integration allows end-users to monitor the health parameters collected through the CAMI system via web-accessible graphical interfaces.

The **CAMI Multi-Modal Interface** allows for user interaction with the system. This is achieved by smartphone application (implemented on iOS) and the CAMI Linkwatch web interface, and vocal interaction.

For each integration aspect, a dedicated micro service running in its own Docker container or using its own RESTful API is implemented.

3 RESULTS

3.1 The CAMI end-user perspective

A total of 105 respondents (55 to 75 years old) from Romania (42), Poland (37) and Denmark (26) have participated in the multinational survey. These are CAMI's primary end-users, that is, older adults and elderly who have benefited from the digital revolution and are therefore expected to have an increased acceptance of innovative ICT solutions. Out of the respondents 49 were males and 56 females living in urban and sub-urban areas (87%). Most of the respondents (46%) had a master degree or higher while 28% held a post-secondary school qualification. Most of the respondents were married or in partnership (64%), and almost 30% of them were living alone being widowed or separated. More than a half of the respondents were retired (55%), while the rest were still active (employed or running their own business). All respondents were informed about the anonymity of the survey and have provided an informed consent for their participation.

Several of the CAMI technologies were already used by the respondents as independent devices. For example, 85% of the respondents have 4 or less mobile devices (smartphone, tablet, etc). The median value of number of possessed devices for Denmark is 4, while for Poland and Romania is 3. The number of possessed devices does not depend on gender (η^2 - 0.02) and age (η^2 - 0.026). Small influence has the level of education (η^2 - 0.279) and number of people the respondents live with (η^2 - 0.206).

In addition, respondents expressed their interest in CAMI's technologies. E-health solutions were well perceived as very useful and 59% of the respondents considered the graphic display of various health measurements (e.g. blood pressure, heart rate, oxygen levels) as an interesting feature. The ability to share health measurements with various doctors was perceived as useful by 60% of the respondents. Physical and mental exercises were of interest for 41% of the respondents. All the voice control functionalities received more than 50% votes which classifies them as promising features. Nevertheless, more than half (60%) considered that a mobile device is an acceptable interface with the system. Robotic platforms were perceived as being useful for a number of tasks such as house supervision (51 %), telepresence (54%), manipulation of objects (50%), etc. A total of 87% of the respondents were worried about the price of the CAMI solution. Consequently, a renting scheme was more appealing to the participants. They were interested to own the health monitoring devices and rather rent other less personal devices and sensors such as those for home monitoring, serious games, robotic platforms, etc.

With the multinational survey outlined above establishing the interest of the CAMI primary users for the targeted technologies, we proceeded to rank these technologies according to the users' preference. For this purpose, we employed the best-worst scaling (BWS) method for a total of 57 elderly participants: 25 Romanians (11 female, 14 male), 20 Polish (13 female, 7 male), and 12 Danish (6 female, 6 male) respondents. The BWS belongs to the so-called 'stated preferences' family of method and was first proposed by Louviere in a series of articles, as an alternative to classical discrete choice experiments [6]. In BWS, respondents were not asked how much they prefer certain alternatives of the CAMI solutions compared with each other, but only to choose which options they prefer and which they don't. A list of 22 functionalities was defined as bases for the survey's combinations pool (see Figure 3).

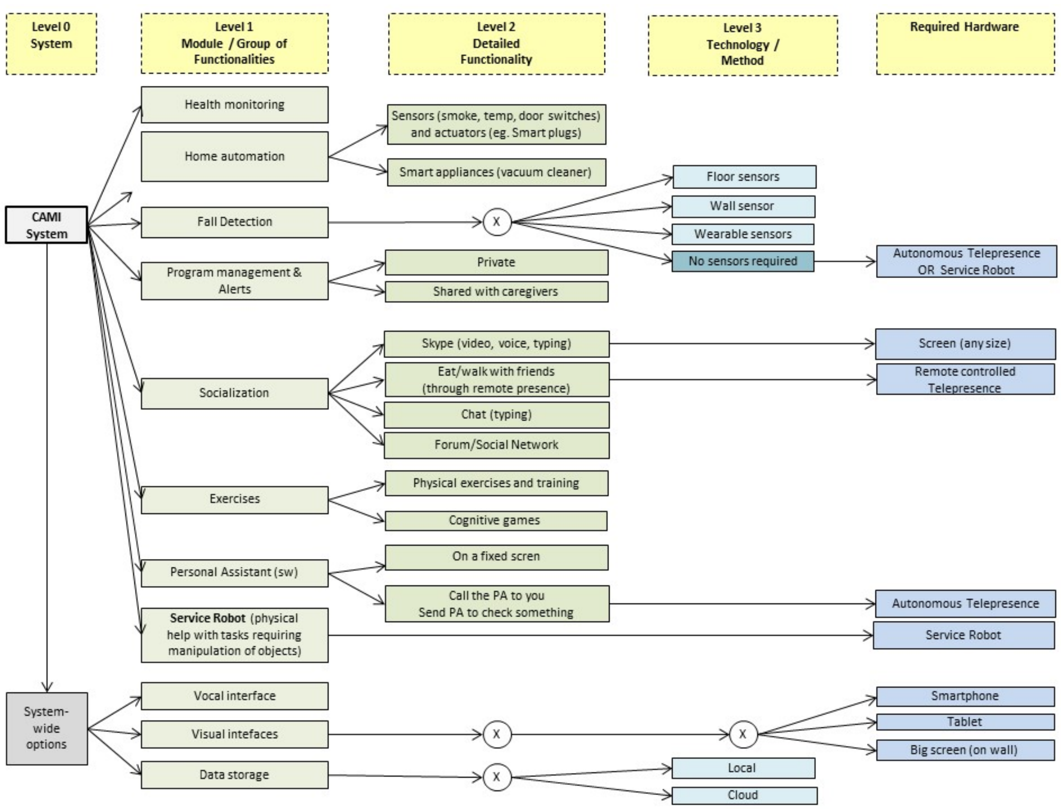

Figure 3. Summary sketch of CAMI's envisioned functionalities distributed hierarchically.

All relevant functionalities were subject to a randomization algorithm utilized by the BWS method, whose purpose was to maximize the probability of appearance of each functionality in a certain choice set. The BWS calculated scores are directly comparable in terms of strength or magnitude of preferences. More specifically, a functionality with the computed preference weight of 8 is twice as preferred as a functionality with the computed preference weight of 4. In our research, the two most preferred functionalities (1 and 2) were twice as preferred as more than half of all preferences (for instance, 4, 18, 12, 10, 17 and so on).

As depicted in Figure 4, functionalities 1 (basic health parameters monitoring), 2 (smart house with various sensors, such as smoke, temperature, open doors, etc.) and 14 (computer supervised physical exercise and training program) were ranked the three most preferred, whereas functionalities 4 (fall detection alert-able floor), 13 (socialization via forums), and 22 (have the system-acquired data stored in the cloud) were ranked least preferred.

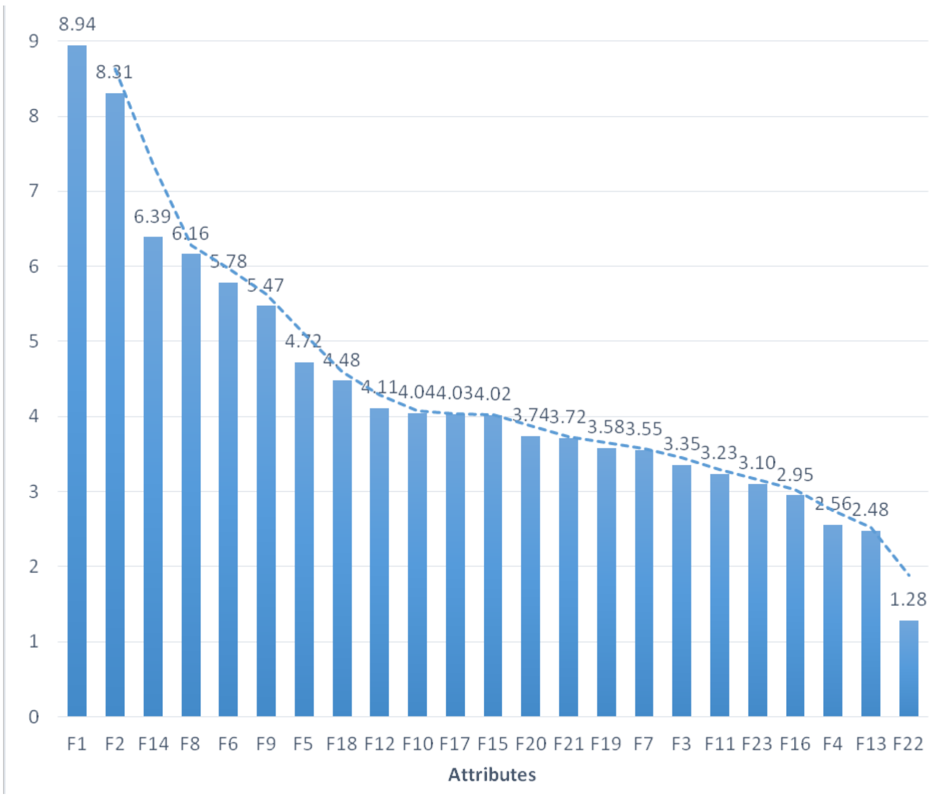

Figure 4. Preferences hierarchy of all functionalities for all respondents. Values represent MaxDiff scores.

Further discussions within focus groups organized in all three countries, i.e. Romania, Poland and Denmark, have revealed that fall detection was ranked low only in implementations requiring substantial changes of the home environment and high investments (alert-able floor). On the contrary, fall detection through wearable sensors was considered an essential feature, especially by the caregivers. The focus groups were conducted with 5 – 8 elderly and with a large variety of secondary stakeholders such as nurses, IT specialists, insurance companies, etc. The vocal interface feature, which was greatly overlooked during both the multinational survey and the BSW analysis was appreciated during the focus groups and subsequent demonstrations performed during the CAMI project.

In the next sections, we are presenting the implementation of some of the CAMI technologies which were ranked top by the users participating in the project.

3.2 Health monitoring and fall detection

The health monitoring functionality in CAMI is allowing users to perform scheduled (e.g. according events scheduled in the user's calendar) monitoring of important health parameters, *i.e.* blood pressure, heart rate, blood glucose, weight, blood oxygenation. The integrated medical devices are transferring the acquired data via Bluetooth to the CAMI Gateway and then further to the CAMI database. The acquired data is checked against the user's profile to identify important deviations from acceptable limits. Deviations trigger alert message to the user and the caregiver. The stored data can be used to plot the evolution in time of the physiological parameters. Normal or acceptable limits for each of the parameters will be plotted together with the recorded data.

Falls are serious threats to elderly people living alone and can even result in life threatening situations when the fall is critical and not addressed within a specific time. As a result, in CAMI, the fall detection functionality together with alarm generation on the occurrence of a serious fall, needed in order to automatically inform caregivers and family is given due importance. To detect a fall, we employ the Vibby Oak fall detection sensor, along with its IoT gateway, Vibby Leaf (both produced by Vitalbase). Vibby Oak is a wearable sensor that can be worn around the wrist like a watch, or as a pendant around the neck. They are designed to automatically detect dangerous (heavy) falls of the wearer who lies on the floor. The device also has a manual trigger to push the alarm off, if the user has recovered successfully. A dangerous fall is detected if: a) a body in standing position is followed by: b) a quick and sudden loss of gravity or verticality followed by: c) a sudden and strong impact on the floor, and d) a lying position on the floor with or without activity. If these four phases have occurred, an automatic alarm is then activated. The Vibby Oak algorithm reduces the false fall alarms significantly, although it does not completely remove them.

The Vibby Leaf gateway is an ISM 868Mhz transceiver designed to connect to the Vibby Oak fall detector, with data received by ISM-band radio communication. To integrate it with the CAMI gateway, we forward the fall event to the CAMI gateway, from where it gets pushed to the Decision Support System (DSS) placed in the CAMI cloud, which generates a notification for the caregiver regarding the fall.

3.3 Computer supervised physical exercises

This functionality aims to (1) improve the user's lifestyle by providing regular physical activity that is consistent with his / her medical condition; (2) monitor the user's activity and provide motivational feedback; provide a personalized exercise program that can be adapted by a medical specialist, depending on patient progress. The application is developed using the Unity 3D engine and is based on two avatars: the training avatar and the avatar of the user. The training avatar is performing different physical exercises which the user must reproduce. The user's movements are compared with the movements of the avatar based on algorithms for comparing two nonlinear series. The results are saved in a database, such that they can be analyzed also at later times. At the end of the exercise, the user is receiving a score reflecting the correctness of the performed exercises. A screen shot of the application, with both avatars, is given in Figure 5.

Figure 5. Screenshot of the application.

The Kinect v2 sensor is used to map the user's movements to its avatar. The sensor provides 25 joints for a user at approximately 30 frames per second. For the implemented exercises, only 20 joints are used. For each joint we compute a 3D rotation using quaternions relative to the parent bone. The application aims to personalize the exercises according to the medical or physical condition of each user. For this purpose, we associated weights with each joint in order to reflect the user's condition. Each weight has a value in between 0 and 1 which can be personalized for each user.

In order to compare the trainer and user movements, the similarity between the set of quaternion values for each joint, computed for each frame, is compared using the Dynamic Time Warping algorithm (DTW) [7]. DTW computes the similarity between two series by calculating the minimum distance between them.

The Kinect sensor is introducing input noise when collecting the user's movements. Hence, the result obtained by applying the DTW algorithm to two sets of frames must be processed in order to obtain a robust result that reflects the correctness of the user's movements. The score provided to the user, reflecting his results is computed using eq. (1):

$$\text{Similarity}(f_{ref}, f_{user}) = 1 - \frac{DTW}{ref_{MAX}} \quad (1)$$

where DTW is the distance between f_{ref} and f_{user} computed using DTW; f_{ref} and f_{user} are are the set of frames associated to the trainer and to the user; ref_{MAX} represents the highest value obtained by applying the DTW algorithm. The latter is obtained experimentally based on the user's performance.

The similarity between joints is computed with DTW using the Euclidean distance. Values obtained for joints similarities are normalized in the range [0, 1] by dividing to the maximum obtained value. If one value is greater than a fixed threshold (experimentally obtained) then the movement wasn't performed correctly based on that joint.

We experimented with different distance metrics for computing the DTW, *i.e.* inner product, euclidean distance, squared Euclidean distance and Manhattan distance. Four cases were considered: a) M1 – user movements are performed similar with the reference exercise; b) M2 – user movements are performed slower than the reference exercise; c) M3 – user movements are partially wrong than the reference one and d) M4 – user movents are slower and different than the refence one. The similarity computed with the squared Euclidean distance and the inner product distance can differentiate between these 4 type of cases. The other two considered methods are not able to differentiate between M2 and M3. The user can visualize the results obtained during a week, as shown in Figure 6. One day exercises can be also selected for visualization.

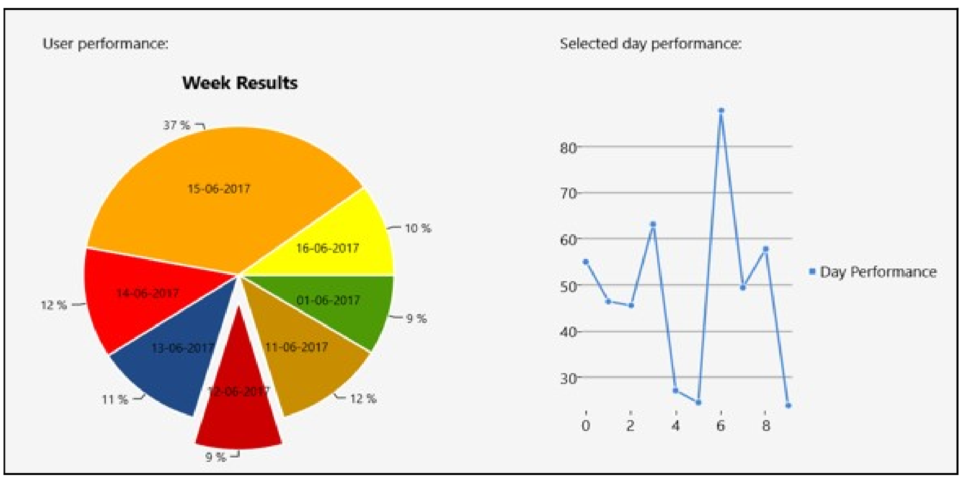

Figure 6. Scores obtained by one user.

3.4 CAMI vocal interface

The vocal interface integrated in CAMI allows oral (speech and not speech) interactions between the user and the system (see Figure 7). After analyzing existing solutions in the field, we developed a vocal interface that is composed of five main modules: Automatic Speech Recognition (ASR), Natural Language Understanding (NLU), Dialog Management (DM), Natural Language Generation (NLG), and Text to Speech synthesis (TTS). Since, the user should always be able to interact with the CAMI platform, regardless of the status of the internet connection, the vocal interface should work also offline ensuring basic system functionalities. Therefore, the vocal interface has two working mode: one that depends on the internet connectivity that will be used normally (online mode) and the other that doesn't depend on internet connection that will be used in case of connection lost (offline / limited mode).

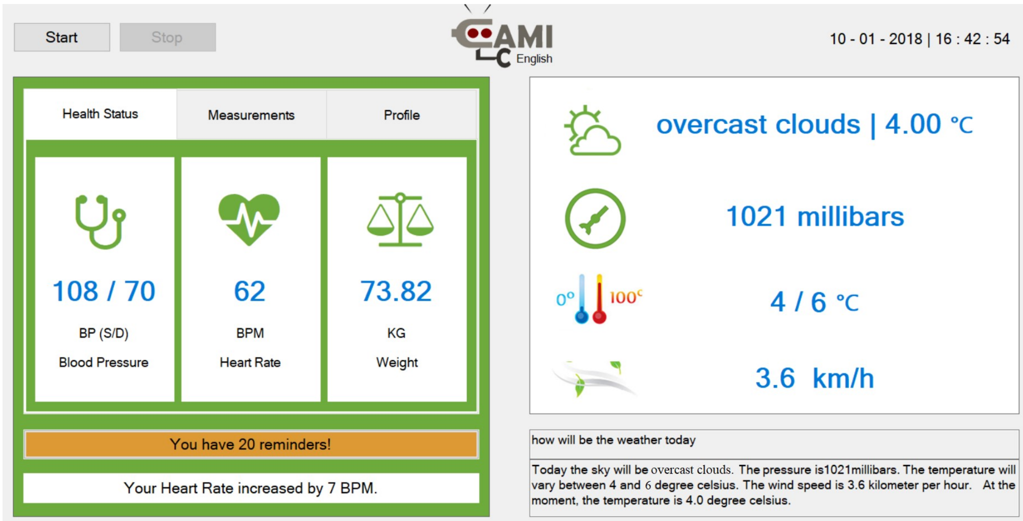

Figure 7. How will be the weather today command. The vocal output of the interface can be seen in the lower right panel.

Briefly, ASR is the technology that allows a computer to identify the words spoken by a person into a microphone and to convert them to writing text. We use the Windows Speech Technology as our ASR solution. It offers a basic speech recognition infrastructure that digitizes the acoustic signals and recovers words and speech elements from the audio input. To access and extend this basic speech recognition technology, we use the "System.Speech.Recognition namespace" and define algorithms for identifying and acting on specific phrases or word patterns. We also manage the runtime behavior of this infrastructure and we created a grammar that consist of a set of rules and constraints to define words and phrases that will be recognized as meaningful input.

The NLU module aims to extract the semantic meaning from the text received from the ASR module by converting the received words to a machine-reading representation. This module is also responsible to correct any errors made by the ASR module. We use the language Understanding Intelligent Service (LUIS) from Microsoft as our NLU solution. By sending the resulted text from the ASR module to LUIS, the system will receive back relevant, detailed information regarding the user's request. For each domain of interaction, we designed its proper domain-specific language model and tailored it to the need of the system.

The DM module is responsible for the state and flow of the conversation. It receives as input some relevant information regarding the user's request that were generated by the NLU. The output of this module is a semantic representation of a list of instructions of the dialog system that determines which should be the system's answer in response to the user's processed input.

The NLG module is the natural language processing task of generating natural language from a machine representation system (such as a knowledge base or a logical form). It may be viewed as the opposite of the NLU. We use the Microsoft Bot Framework from Microsoft as a solution that combines the DM and NLG modules, it is a platform that allows to build, connect, test, and deploy powerful and intelligent bots. We built the bots from scratch, using the Bot Builder SDK for .NET and Node.js provided by the Microsoft Bot Framework.

The TTS module is responsible to artificially produce any generated output normal language text into human speech that will be heard over the speakers of the system. The quality of a speech synthesizer is judged by its similarity to the human voice and by its ability to be understood clearly. We use the SpeechSynthesizer Class included in the Windows Speech Technology.

The implementation outlined above is the way in which the modules of the voice interface work in the online mode. Regarding the offline mode of the voice interface there is no change at the levels of ASR and TTS modules with the exception of the number of the recognized commands which are reduced to those that ensure the basic system functionalities. Regarding the other three modules, we developed an algorithm in which the NLU compares its generated output with an array that contains all the commands that should be recognized. If no match is found, the system will tell the user that the command is not recognized, and will ask for a new one. If a match is found, the system will execute the predefined content and will generate the outputs that correspond to the match (local variable involved only).

We tested our voice interface in the laboratory. The tests have been done on a set of 200 interactions between 5 different users and the machine. Each user repeated 20 interactions with the system 2 times using a Plantronics Voyager 5200 UC Microphone integrated into the system in a quiet then noisy environment. We tested the solution on an HP ZBook 15 G3 (Core i7 2.60 Ghz, 8 GB RAM, integrated stereo speakers), having Windows 10 - 64 bits operating system. For the ASR module we used the Levenshtein Distance to calculate the differences between the speech recognition results and the original texts [8]. Some of the results are listed in Table 1.

Table 1. Results of the ASR module performance tests.

Users Input	Av. Lev. Score	
	Quiet Env.	Noisy Env.
Who are you	98.51	94.63
How will be the weather today	98.74	95.72
What I have scheduled for today	98.05	94.62
What is my health status	99.73	95.70
Call Bogdan	99.89	96.12
How much have I walked	97.15	94.21
Show my calendar	98.56	94.81
Take a new health measurement	98.16	94.31
I forgot how to use the device	99.15	94.29
Display my sugar blood measurements	98.89	94.58

For the TTS module we used a questionnaire to find out if the user understood the spoken output and its degree of the clearness. For each question, the user answered with a score from 0 to 10 where 10 represents represent full understanding or excellent clearness. Results are listed in Table 2.

Table 2. Results of TSR performance testing.

System Spoken Output	User Satisfaction (Average Score)			
	Understood		Clearness	
I am CAMI, your smart personal assistant. How can I help you?	10.00	9.75	9.50	9.50
Here is the home control main page.	10.00	9.50	9.50	9.50
What health measurement do you want to take?	9.75	8.75	9.00	9.00
Here is your weekly calendar. Your time is getting filled.	10.00	9.75	9.50	9.50
Here are your sugar blood variations during the past month.	9.75	8.75	9.25	9.25

The obtained results are satisfying and are considerably improved when using an advanced microphone. Furthermore, we expect also the TTS module to exhibit improved results when using high-quality speakers.

4 CONCLUSIONS

A user centered design involving participants from Denmark, Poland and Romania was employed for the development of the CAMI platform in order to best fit the requirements and needs of elderly people. In addition to a multinational survey, the BWS method was used to obtain not only a general opinion of the elderly people on the CAMI solution but also a ranking of the CAMI functionalities. Functionalities related to health monitoring and prevention, including fall detection and computer supervised exercises, are highly ranked by the elderly users. Further discussions within focus groups and demonstration sessions organized in three countries also reveal vocal interaction with the CAMI platform is considered important when elderly people are involved.

The obtained results helped us propose guidelines for decreasing acceptance barriers of ICT solutions among the aging population. Moreover, we implemented these in the design and development of CAMI's functionalities. Through the integration of e-health application for self-monitoring of health parameters (e.g. blood pressure, glucose, heart rate, weight, etc.) with computer supervised physical exercise the CAMI end-users can learn how to for maintaining a healthy and independent lifestyle. In addition, fall detection and alerting is increasing their level of security and confidence. Extensive field trials are planned to gather users' assessment of the implementation of the CAMI functionalities described in this paper.

ACKNOWLEDGEMENTS

This work was supported by the Active and Assistive Living (AAL) program through a grant of the Romanian National Authority for Scientific Research and Innovation, CCCDI – UEFISCDI, CAMI - "The Artificially intelligent ecosystem for self-management and sustainable quality of life in AAL", project number AAL-2014-1-087".

REFERENCES

[1] Active and Assisted Living Program: CAMI (AAL-2014-1-087) Official Website: www.camiproject.eu – accessed on: 2018-01-15.

[2] A. Kunnappilly, A. Sorici, I.A. Awada, I. Mocanu, C. Seceleanu, A.M. Florea, "A Novel Integrated Architecture for Ambient Assisted Living Systems" *Computer Software and Applications Conference (COMPSAC), 2017 IEEE 41st Annual*, vol. 1, 2017.

[3] A. Sorici, I. A. Awada, A. Kunnappilly, I. Mocanu, O. Cramariuc, L. Malicki, C. Seceleanu, A. Florea, "CAMI - An Integrated Architecture Solution for Improving Quality of Life of the Elderly", *International Conference on IoT Technologies for HealthCare*, pp. 141-144, 2016.

[4] Official Website: www.linkwatch.se – accessed on: 2018-01-15.

[5] Official Website: www.opentelehealth.com – accessed on: 2018-01-15.

[6] J.J. Louviere, "Conjoint analysis," 223-259, Blackwell Publishers, 1994.

[7] H. Sakoe, C. Seibi, "Dynamic Programming Algorithm Optimization for Spoken Word Recognition" *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. ASSP-26, no. 1, pp. 43-49, 1978.

[8] V.I. Levenshtein, "Binary codes capable of correcting deletions, insertions, and reversals", *Soviet Physics Doklady*, vol.10, no. 8, pp. 707–710, 1966.