Article:

Malavasi, M., Turri, E., Atria, J.J. et al. (2017) An innovative speech-based user interface for smarthomes and IoT solutions to help people with speech and motor disabilities. Studies in Health Technology and Informatics, 242. pp. 306-313. ISSN: 0926-9630

https://doi.org/10.3233/978-1-61499-798-6-306

An innovative speech-based user interface for smarthomes and IoT solutions to help people with speech and motor disabilities

Massimiliano MALAVASI [a,1], Enrico TURRI [b], Joaquin ATRIA [c], Heidi CHRISTENSEN [d], Ricard MARXER [e], Lorenzo DESIDERI [f], Andre COY [g], Fabio TAMBURINI [h] and Phil GREEN [i]

[a] *AIAS Bologna, Italy*
[b] *University of Bologna, Italy*
[c,d,e] *University of Sheffield, UK*
[f] *CRA Centro Regionale Ausili, Italy*
[g] *The University of the West Indies, Jamaica*
[h] *FICLIT, University of Bologna, Italy*
[i] *University of Sheffield, UK*

Abstract. A better use of the increasing functional capabilities of home automation systems and Internet of Things (IoT) devices to support the needs of users with disability, is the subject of a research project currently conducted by Area Ausili (Assistive Technology Area), a department of Polo Tecnologico Regionale Corte Roncati of the Local Health Trust of Bologna (Italy), in collaboration with AIAS Ausilioteca Assistive Technology (AT) Team. The main aim of the project is to develop experimental low cost systems for environmental control through simplified and accessible user interfaces. Many of the activities are focused on automatic speech recognition and are developed in the framework of the CloudCAST project. In this paper we report on the first technical achievements of the project and discuss future possible developments and applications within and outside CloudCAST.

Keywords. Speech recognition, Smart home, Disability, Internet of things (IoT), Dysarthria, Assistive technology (AT)

1. Introduction

The Department of Computer Science at the University of Sheffield (UK), in collaboration with the universities of Toronto (Canada) and West Indies (Jamaica) and AIAS Bologna Onlus, is funded by the Leverhulme Trust to develop a cloud based computing resource, for clinical and educational applications related to technologies for speech recognition (CloudCAST: clinical applications of speech technology) [1,2,3]. The project aims to provide a route to bring automatic learning and speech recognition technology to professionals who deal with speech problems, such as therapists,

[1] Corresponding Author. AIAS Bologna and CRA Centro Regionale Ausili, Via Sant'Isaia 90, Bologna, Italy. Email: mmalavasi@ausilioteca.org

pathologists, AT experts and teachers, creating a self-sufficient community that continues to grow the resource after the three-year funding period (from January 2015 until December 2017). The project aims to achieve this by creating an internet-based and free resource, which will provide a set of software tools for personalized speech recognition and speech therapy. Moreover, it provides a personalized interactive dialogue; the automatic speech recognition system of CloudCAST is able to adapt to dysarthric speech, or to more general problems of language by adaptation processes which require a brief training. This opens up the possibility of providing users with an efficient voice control system, for example in the case where they are unable to use other user interfaces for Information and Communication Technologies (ICT) and electronic home devices like traditional mice, keyboards and remote controls.

Home automation systems, often used in combination with other assistive technologies, can greatly increase the autonomy and safety of persons with disabilities [4]. This sector can be considered an evolving reality, that can be further advanced by the increasing availability and dispersal of new low cost mobile ICT solutions (smartphone and tablets) and IoT solutions. It is important to understand how these two factors may change the home automation sector for people with disability and the possible synergies among them. Currently, many home automation systems require substantial changes of the domestic electronics to be installed and integrated with specific user control interfaces from the AT sector. This type of user interfaces is often quite expensive and with a poor choice of smart functions, if compared with those present in actual smartphones or tablets. In the next paragraphs, we will describe the development of the CloudCAST platform to efficiently interact both with traditional home automation systems and IoT solutions, using mobile ICT devices.

2. Methods

Since the very beginning of the project, the first problem to solve appeared to be the creation of a single multi-standard access point to handle both home automation systems and IoT devices, which could also be connected to different types of customizable human machine interfaces (HMI). After an analysis of the possible technical architecture of such e system, the main promising technical choice appeared to develop and implement a web-server system, possibly based on an open architecture. Different solutions have been evaluated (see table 1) to define the architecture underlying the system.

Table 1. Web-server systems for home Automation and IoT

Name	Open Source	Technologies supported	HMIs available
OpenHAB	yes	121	iOS and Android Apps Web Interface (Classic UI) Web-app (GreenT) XML-based (Comet Visu)
Open Remote	yes	46	iOS and Android Apps Web interface
IKON	no	29	App iOS e Android Web interface
Calaos	yes	14	iOS and Android Apps Web application

DomotiGa	yes	92	Touchscreen user interface iOS and Android Apps Web interface
Thinknx	no	44	App Android Web interface
Jeedom	yes	101	iOS and Android Apps Web interface

The choice fell on OpenHAB (Home Automation Bus) framework [5,6] which is specifically aimed at creating integrated smart home control systems, not bound to a specific hardware device or system and capable of using a single communication protocol. The main reasons for this choice was the huge number of different technologies and standards supported, the full open source architecture and the active community of developers and users that is supporting the project.

OpenHAB is fully based on OSGI (Open Service Gateway Initiative) and makes it possible to build modular application components (bindings). Any technology, device, social network or integrated cloud platform is supported by a specific binding. These packages can be added or removed to expand or limit the functionalities of a specific installation. OpenHAB is designed to run independently and brings together different types of field-bus systems, hardware devices and interface protocols for dedicated applications. These allow an application to send and receive commands and status updates, enabling the design of personalized user interfaces with a unique appearance and keeping open the possibility of operating with multiple devices and services. These allow an application to send and receive commands and status updates on the bus enabling the design of personalized user interfaces with a unique appearance while keeping open the possibility of operating with multiple devices and services. It also allows the development of automation logics between the different sections of the system. This type of architecture allows the system to be ready also for the new technologies that will be placed on the market. Among the most promising of these, are surely the new cloud based personal assistants dedicated specifically to environmental control and daily home activities, such as Google Home or Amazon Alexa. In this way the system can also assure a great choice of different technologies not only for the home environment adaptation, but also for creating adequate and innovative user interfaces.

2.1 The development of the first prototype

The prototype was developed and implemented in a real context (Fig.1), one of the Experimental Domotic Apartments (ADS-1) of the Polo Tecnologico Regionale Corte Roncati [7]. These smart homes are currently used as an exhibition site for assistive and AAL (Active Assisted Living) solutions, where end users and professional users can test them. They are also used for temporary residential experiences of independent living within the project "Independent living weekends" targeting young people with disability. The goal of this project is to provide young people with disabilities the opportunity to live a few days in a smart home, to test the benefits they would get by installing similar systems in their homes [8]. The two ADS are equipped with a full automation system based on KNX standard [9]. As one of the main activities of the project was to integrate this system with the prototype, many of the home automation functions implemented have been integrated in the webserver, such as lighting controls, doors, windows, blinds and shutters automation, heating, air conditioning and environmental sensors (such as temperature, presence, flooding, light and smoke sensors).

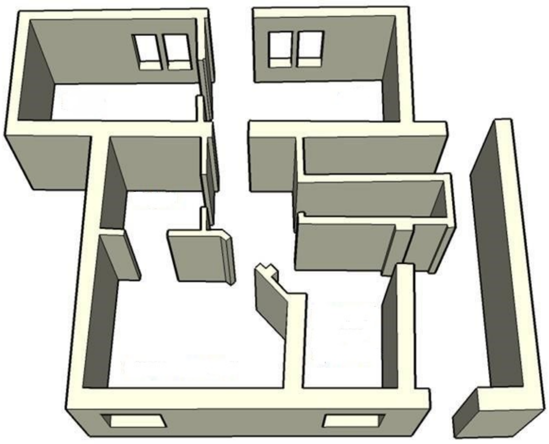

Figure 1. The ADS 1 Smart Home, Corte Roncati, Bologna.

As the prototype is based on OpenHAB technology, the basic automation functions are mainly the ones available in a standard OpenHAB server. The principal automation functions and data types with related commands are described in Table 2.

Table 2. Basic automation functions available in the prototype

Data Type	Description	Automation commands
Color	RGB data for color lighting	On, Off, Increase, Decrease, Percent, HSB
Contact	State of binary sensors	Open, Closed
DateTime	Date and time data	
Dimmer	Lighting dimmers data	On, Off, Increase, Decrease, Percent
Group	Useful for creating groups of basic functions and scenarios	
Location	Location data	
Number	Numeric data in different formats	
Rollershutter	Data for motors, blinds, and shutters automations	Up, Down, Stop, Move, Percent
String	Text data	String
Switch	Data used for any kind of switching functions	On, Off

In a logic of creating a low cost control system, the webserver has been implemented and tested on a Rasperry Pi system [10]. In the same logic, several IoT devices were integrated in the system and tested, but a full functional integration was possible only with three systems:

- RGB Philips Hue bulbs (a lighting system)
- Netatmo Weather Station
- Logitech Harmony HUB (an infrared universal remote control system)

The third device is really important for daily use as it allows target users to control many home electronic devices, natively equipped with an infrared (IR) remote control. Some other smart functions, particularly useful for the needs of target users, were developed and tested: non-invasive monitoring based on environmental sensors, facilitated interfaces for deaf and blind users, cloud based remote control and monitoring.

3. First results

Before the development of the prototype, a voice control of the of the ADS-1 smart home was possible only through the use of a specific universal remote control that allows the management of infrared (IR) codes. The device has the capability to learn IR codes directly from other remote controls and the embedded voice recognition system requires a training session for each command. This kind of products belong to the AT commercial field and are in many cases relatively expensive. A key feature of the project was therefore to integrate voice control also in a more flexible and smart low cost prototype, based on a standard mobile device. The official app from OpenHAB developer community allows users to operate the mobile device's microphone for acquiring a command phrase and send the recognized text to a specific module of the web server. A specific script was developed to parse the command phrases and decode the single commands to activate the functions available in the ADS-1 apartment. In order to make this service accessible to people with upper limbs disability, it has been developed and integrated with another software unit to provide a complete hands-free voice activation (no button or touch command needed). This first implementation used Google voice recognition services. The major limitation of this system for voice control is the need for an internet connection, but on the other hand, it has an important economic advantage: it only requires a device with Android OS on board, which can be purchased for a few hundred euros or even less.

3.1 A CloudCAST based command and control user interface for smart homes

To extend the performance and the functions available, a second prototype, based on CloudCAST technology and able to support also dysarthric speakers, was developed.

The clinical applications of speech technology to the needs of people with disability face two challenges. The first is the lack of data: there are few corpora available to support the techniques that rely on machine learning and it is also difficult to collect large speech corpuses. The second is customization: this field requires individual personalized solutions, and a technology that adapts to its users rather than demanding that the users adapt to it. CloudCAST addresses both these two issues, developing an adaptive technology simply available, also remotely, to professionals who work with language and AT. The CloudCAST resources can also facilitate the collection of voice data needed to improve the machine learning techniques that are the basis of this technology: it's able to automatically collect data from systems that are already in use, in addition to providing tools for the collection of specific databases.

To develop the second user interface for the prototype, it was planned from the beginning to use a technology based on open-source services, like Kaldi for automatic speech recognition and OpenHAB for home automation control.

Compared to CloudCAST goals, existing solutions for the target user groups are in general inferior in terms of the choice of recognizer output, the flexibility of the recognition process, the personalization of the speech models and methods of interaction. CloudCAST services include the possibility to provide interactive voice recognition where the user is able to change the grammar, the acoustic models, and other essential parameters. The recognizer may also provide feedback on its own performance, for examples partial decoding and confidence measures. Subsequently, the interaction of different users with CloudCAST should provide data resources to improve the recognition process and the training of future models.

The general architecture of CloudCAST (Figure 2) can be divided into two sections: the application and the CloudCAST server. The server processes the audio data according to the models and provides the speech recognition results using a Kaldi library. The server has also the task managing the changes to the parameters of the recognition models; both application and server have access to a common storage database for models, recordings and authentication data. The CloudCAST site will be accessible to users who can manage recordings, developers who wish to obtain the API keys, and professionals who want to create new recognition models.

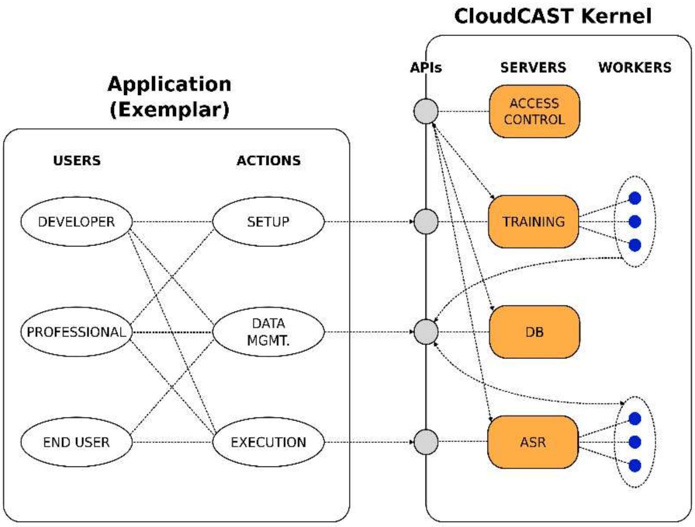

Figure 2. The CloudCAST architecture

In the initial stages, in order to facilitate the creation of services that use CloudCAST, a first speech recognition test client was developed in JavaScript on the basis of dictate.js existing library. The final Command & Control client is planned to extend the functionalities allowing for more types of interaction with the server, such as swapping grammars, models and other parameters, so as to interpret the results provided by different servers.

In the CloudCAST based prototype, it was possible to view the same initial interface of OpenHAB in a different format, with the possibility to provide feedbacks of the actions performed by the system. Compared to the first prototype, in this version, the items and commands to execute were single words rather than sentences, to better support dysarthric speakers. In this new interface the microphone is always active and the user navigates the command tree using single commands: a specific word is recognized as soon as it matches with one of those displayed. The system then either proceeds to a submenu or executes the command. For example, to open the door of the kitchen, the user will have to pronounce the words in a sequence to access the relevant page for the room and finally the object concerned, as for example: "ground floor", "kitchen" and "open door." This procedure can appear unnatural, but it’s often required for users with severe dysarthric problems, who usually have difficulties in articulating a whole sentence instead of a word at a time. This also helps the CloudCAST speech recognition system since for every access to an interface page, it only needs to consider the allowed words for that context. Keeping the grammar simple, with only a few commend options (low perplexity) at every stage of the control sequence, makes the system less prone to error detection.

Future developments include the possibility of activating the recognition after the delivery of a specific word (vocal password) and saving an adaptation session in order to make specific recognitions models more constant and performing for a certain user.

As the CloudCAST cloud server will be available by the end of the project, the first tests have been performed using a demo Kaldi version with a standard language model. When the service will be on-line, selecting few options it will be possible to specify a configuration based on a model to be used by dysarthric users.

3.2 An Italian speech recognition model for Italian dysarthric users

Due to the absence of a speech recognition model for Italian dysarthric users, another ongoing activity involves the acquisition of voice data by some volunteers who belong to the Ausilioteca Living Lab (Bologna, Italy). So far, 5 potential users have been involved in the recordings of two word lists. The first list includes 120 most frequent polysyllabic words; the second list comprises 75 words for command and control of ICT and home automation devices or systems (e.g., on, off, open, close). Participants recruited (mean age 40,2; standard min-max: 23-52; 2 females) have moderate to severe dysarthria due to cerebral palsy. All of them are expert users of assistive technologies and two live in a fully equipped smart home. They also have volunteered to test the future user interfaces developed within the project.

From these data it's under development the first Italian speech recognition model for users with dysarthria. This model will be tested for the first time in the CloudCAST Command and Control prototype connected to the ADS-1 smart home automation system.

4. Discussion

The initial goal of the project was to create a low cost and highly accessible home automation control system, to be tested in already existent smart homes, and which could be operated from off the shelf mobile ICT devices. The prototype has fully achieved the functional objectives, while the integration with a completely hands free voice recognition system and the integration of the CloudCAST technology have created a significant added value in relation to the target user needs.

The versatility of the system developed has allowed the creation of a solution that is able to replace or integrate specific products belonging to the AT market for environmental control, which currently have often relatively high costs, functional limits and a certain level of obsolescence due to the rapid evolution of home automation and IoT technologies.

Thus, the solution developed can be considered a starting point for the creation of low cost but high-performance custom home automation systems, useful for improving the levels of safety and autonomy in daily living activities, and particularly dedicated to people with full or partial inability in the use of the upper limbs, but not limited to this: the integration with the cloud based speech recognition systems, provided in CloudCAST, will also allow, the possibility of an efficient use by dysarthric users.

The development of an automatic speech recognition engine also for Italian dysarthric users can have a very positive impact for this users group. The method used for its development can also be easily generalized and extended to other languages, making the system developed easily adaptable to other international contexts.

Acknowledgements

CloudCAST is an International Network (IN-2014-003) funded by the Leverhulme Trust.

References

[1] Green, P., Mrxer, R., Cunningham, S., Christensen, H., Rudzicz, F., Yancheva, M., Coy, A., Malavasi, M., Desideri, L. Remote Speech Technology for Speech Professionals - the CloudCAST initiative. 6th Workshop on Speech and Language Processing for Assistive Technologies, Dresden, Germany, 11 September, 2015.

[2] Green, P., Mrxer, R., Cunningham, S., Christensen, H., Rudzicz, F., Yancheva, M., Coy, A., Malavasi, M., Desideri, L., Tamburini, F. (2016). CloudCAST - Remote Speech Technology for Speech Professionals. Interspeech 2016, 8-12 September, San Francisco, USA.

[3] http://cloudcast.rcweb.dcs.shef.ac.uk/

[4] Giovanni del Zanna, Massimiliano Malavasi, Giulio Vaccari "Manuale illustrato per la domotica a uso sociale-la casa flessibile al servizio dell'uomo", Tecniche Nuove (2009)

[5] https://community.openhab.org/

[6] https://github.com/openhab/openhab/wiki

[7] R. Agusto. F.M. Ioele, M. Malavasi, S. Martinuzzi, M.R. Motolese, M.Rimondini "Living in the Living Lab! Adapting Two Model Domotic Apartments for Experimentation in Autonomous Living in a Context of Residential Use (2014)" In S. Longhi, P. Siciliano, M. Germani, A. Monteriù (Eds.), Ambient Assisted Living - Italian Forum 2013, pp. 325-333

[8] Krantz O. "Assistive devices utilization in activities of everyday life. A proposed framework of understanding a user perspective." Disability and Rehabilitation: Assistive Technology, pp. 189-198 (2012)

[9] https://www.knx.org/knx-en/index.php

[10] https://it.wikipedia.org/wiki/Raspberry_Pi