

An Innovative Speech-Based Interface to Control AAL and IoT Solutions to Help People with Speech and Motor Disability

Massimiliano Malavasi, Enrico Turri, Maria Rosaria Motolese, Ricard Marxer, Jochen Farwer, Heidi Christensen, Lorenzo Desideri, Fabio Tamburini and Phil Green

Abstract The main aim of the project described in this paper is to develop an experimental low cost system for environmental control through simplified user interfaces and voice control, to better respond to the needs of users with motor speech impairments (dysarthria). The project is actually being conducted by Area Ausili, a department of Polo Tecnologico Regionale Corte Roncati in Italy, in collaboration with AIAS AT Team and the CloudCAST project. The prototype, that has been implemented in an experimental smart home in Italy, integrates a completely hands free commands recognition function based on the cloud-based voice recognition system developed within the CloudCAST project. The target of the project is to create a tool that is able to overcome the limits of some actual assistive technologies, and to take advantage from the evolution of home automation and internet of things technologies. Future developments of the projects and expected results are discussed.

Keywords Voice recognition · Smarthome · Disability · Internet of things (IoT) · Environmental control · Speech disability

Massimiliano Malavasi: **Area of Interest**—Disability and rehabilitation.

M. Malavasi (✉) · L. Desideri
CRA Centro Regionale Ausili, Area Ausili di Corte Roncati,
Via Sant’Isaia 90, 40123 Bologna, Italy
e-mail: mmalavasi@ausilioteca.org

L. Desideri
e-mail: ldesideri@ausilioteca.org

E. Turri
Università di Bologna, Facoltà di Ingegneria, Bologna, Italy
e-mail: enrico.torre91@yahoo.it

M.R. Motolese
CAAD Bologna, Bologna, Italy
e-mail: mmotolese@ausilioteca.org

R. Marxer · J. Farwer · H. Christensen · P. Green
University of Sheffield, Sheffield, UK
e-mail: r.marxer@sheffield.ac.uk

1 Introduction

The expansion of the functional control capabilities of home automation systems and Internet of Things (IoT) devices for the needs of users with disability is the subject of a research project currently being conducted by Area Ausili (Assistive Technology Area), a department of Polo Tecnologico Regionale Corte Roncati of the public health system authority of Bologna (Italy). The main aim of this project is to develop experimental low cost systems for environmental control through simplified user interfaces and voice control. Some of the activities have been set within the CloudCAST [1–3] project. Here we report on the first technical achievements of the project and discuss future possible applications within CloudCAST.

1.1 *The CloudCAST Project and Environmental Control*

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 computing resource based in the cloud, for clinical and educational applications related to technologies for speech recognition (CloudCAST: clinical applications of speech technology). The project aims to provide a way for automatic learning and speech recognition technology developments to be put in the hands of professionals who deal with speech problems, such as therapists, pathologists, 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, 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 voice recognition system of CloudCAST is able to adapt to dysarthric speech, or to more general problems of language by adaptation processes which require only a few minutes of training data. 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 interfaces like traditional mouse or keyboards.

J. Farwer
e-mail: j.farwer@sheffield.ac.uk

H. Christensen
e-mail: heidi.christensen@sheffield.ac.uk

P. Green
e-mail: p.green@sheffield.ac.uk

F. Tamburini
FICLIT, University of Bologna, Bologna, Italy
e-mail: fabio.tamburini@unibo.it

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, but can it be advanced by the increasing availability and dispersal of new low cost mobile ICT solutions? Another significant element is the growth, in recent years, of the availability of IoT solutions. It is important to understand how these two factors may change the home automation sector for people with disability. Currently, many home automation systems require substantial changes of the domestic electronics to be integrated with user control interfaces from the assistive technology (AT) field. This type of interfaces is often quite expensive and with a poor choice of smart functions, if compared with those present in smartphones or tablets. In the present contribution, we describe the development of the CloudCAST platform to efficiently control both traditional home automation systems and IoT solutions, using mobile ICT devices.

2 Materials and Methods

2.1 A Single Multi-standard Access Point

The first problem to solve was to create a single multi-standard access point to handle home automation systems and IoT devices which could also be connected to different types of human machine interfaces (HMI). The technical goal was to create a web-server system, possibly based on an open source 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 Touchscreen user interface
DomotiGa	Yes	92	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 control systems, not bound to a specific hardware device 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 actually supporting the project.

OpenHAB is fully based on OSGI (Open Service Gateway Initiative). It 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 are optional and 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.

2.1.1 Basic Automation Functions Available in the Prototype

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

Table 2 Basic automation functions available in the prototype

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

2.2 *Implementing the Prototype*

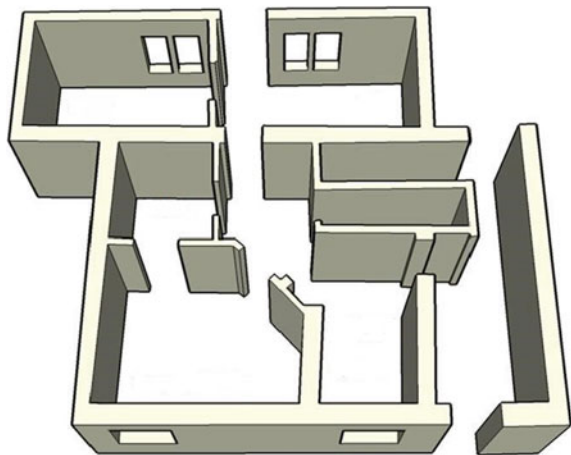
The prototype has been developed and implemented in a context of real application (Fig. 1): the two Experimental Domotic Apartments (ADS) of the Polo Tecnologico Regionale Corte Roncati of the public health system authority of Bologna (Italy) [7]. They are used as an exhibition site for assistive and AAL solutions, where end users and professional users can test them. They are also used for temporary residential experiences of independent living within the project “Weekends of Autonomy” dedicated to young people with disability. The goal of this project is to give the opportunity to users with disability 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. One of the main activities of the project was to integrate this system with the prototype.

Specifically, many function have been integrated such as lighting controls, automation of doors, windows, blinds and shutters, heating/air conditioning systems and environmental sensors (such as temperature, presence, flooding, light and smoke sensors).

2.3 *Integrating IoT Devices and Low Cost ICT Solution*

In a logic of creating a low cost control system, the webservice has been implemented and tested on a Rasperry Pi 1 system [9]. 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:

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



1. RGB Philips Hue bulbs (a lighting system)
2. Netamo Weather Station
3. Logitech Harmony HUB (an infrared universal remote control system).

The third device is really important for daily use as it can allow target users to control many audio/video home devices.

3 Results and Discussion

3.1 A Completely Hands-Free Home Control Interface

Some smart functions, particularly useful for the needs of target users, were developed and tested:

1. Non-invasive monitoring based on environmental sensors
2. Functions for deaf and blind users
3. Cloud based and remotely controlled functions.

At present, the voice control of the devices of the two ADS is made through the use of a 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 voice recognition module requires a brief user training session. This kind of product belongs to the AT commercial sector. A key feature of the project was therefore to integrate voice control also in the low cost prototype based on mobile devices. The official app of the OpenHAB community allows users to operate the mobile device's microphone 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 apartments. 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. In the present phase, it uses Google voice recognition services. The major limitation of this system for voice control is the need for an internet connection. Other research projects have involved the use of a computer with voice recognition systems that run locally. On the other hand, the designed solution has an important economic advantage: it only requires a device with Android OS which can be purchased for a few hundred euros.

3.2 The CloudCAST Based Solution

To extend the performance and the functions available, a second prototype, based on CloudCAST technology, was developed. Clinical applications of speech technology 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 since it is difficult to collect large speech corpus, the only way to address this problem is to collect material which is produced by systems already in use. The second is customization: this field requires individual solutions, and a technology that adapts to its user rather than demanding that the user adapts to it. CloudCAST addresses these two problems making the adaptive technology available at a distance to professionals who work with language. The CloudCAST resources also facilitate the collection of voice data needed to improve the machine learning techniques that are the basis of this technology: it will be able to automatically collect data from systems that are already in use, in addition to providing a system for the collection of databases. From the beginning, it was planned 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 inferior in terms of the choice of recognizer output, the flexibility of the recognition process, the personalization of the speech models and modes 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 decodings 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 (Fig. 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 also has the task of applying the changes to the parameters of the models concerned for the recognition process; both application and server have access to a common storage database for models, recordings and authentication data. The CloudCAST site will be visible to users who can manage recordings, developers who wish to obtain API keys, professionals who want to create models, and so on.

In the initial stages, in order to facilitate the creation of services that use CloudCAST, a speech recognition client was developed in JavaScript on the basis of dictate.js existing library. The final client is planned to extend CloudCast dictate.js 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 environmental control prototype (Fig. 3), it was possible to view the same initial interface of OpenHAB in a different format with the possibility to provide feedback of the actions performed by the program in a dialog box. Compared to the initial prototype, in this version, the items and commands to execute were single words rather than sentences. In this other interface, the user navigates the command tree using single commands because the microphone is always active and a word is recognized as soon as it matches with one of those displayed. The system then either proceeds to a submenu or executing the

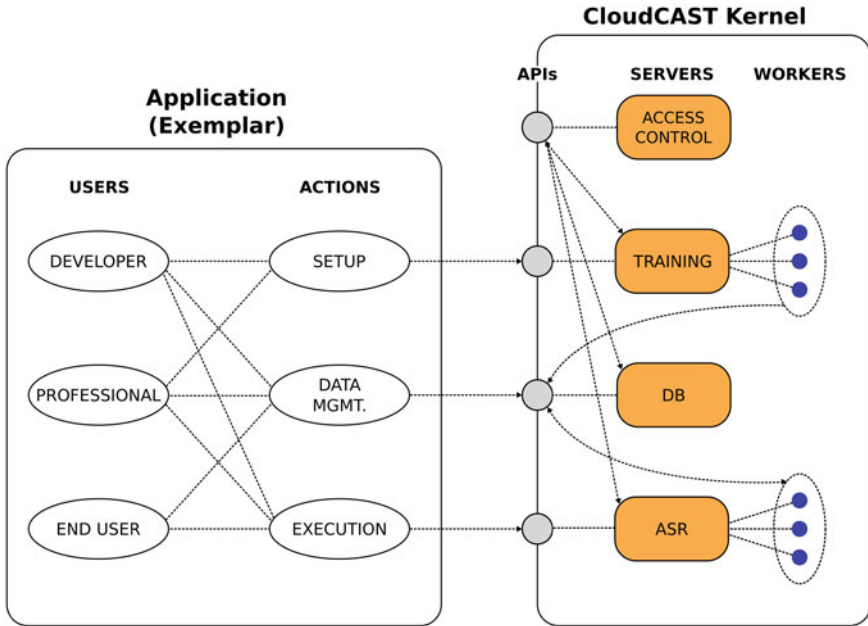


Fig. 2 The CloudCAST architecture

OpenHab voice controller powered by CloudCast

Home / piano terra / soggiorno

man/automatico [off]
ingresso [off]
apri porta [on]
chiudi porta [on]
apri scuri [off]
chiudi scuri [off]
apri finestra [off]
chiudi finestra [on]
veneziana [off]
modalità comfort [off]
modalità notte [off]
automatico/manuale [off]
presenza soggiorno [off]

vnt soggiorno [CLOSED]
diff setpoint soggiorno []
setpoint soggiorno []
Temperatura [22.0 °C]

```
msg: Partial command detected: null From transcript: porte
msg: Executing: "apri porta"
msg: Partial command detected: [{"type":"switch","parameters":{"label":"apri porta","widget":{"widgetId":"PT_soggiorno_2","type":"Switch","label":"apri porta","icon":"none","item":{"type":"SwitchItem","name":"porta_soggiorno_PT_apr","state":"ON","link":"http://10.9.8.8:8080/rest/items/porta_soggiorno_PT_apr"}}}]}]
From transcript: apri
msg: Executing: "apri porta"
msg: Partial command detected: [{"type":"switch","parameters":{"label":"apri porta","widget":{"widgetId":"PT_soggiorno_2","type":"Switch","label":"apri porta","icon":"none","item":{"type":"SwitchItem","name":"porta_soggiorno_PT_apr","state":"ON","link":"http://10.9.8.8:8080/rest/items/porta_soggiorno_PT_apr"}}}]}]
From transcript: apri
msg: Executing: "apri porta"
```

Fig. 3 An early prototype of the CloudCast voice recognition interface for environmental control

command. For example, to open the door of the kitchen, the user will have to pronounce the words in sequence to access the relevant page for the room and finally the object concerned: “ground floor”, “kitchen”, “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 voice recognition system since for every access to an interface page, it only needs to consider the possible words by the processing server. The interface is configured to accept the most similar words, before sending a signal declaring its readiness to receive the command. Keeping the grammar simple, with only a few command options (low perplexity) at every stage of the control sequence, makes the system less prone to error detection.

An improvement has been made to include the function of directly and automatically use the OpenHAB interface in JSON format, without the need to download it. Future developments include the possibility of activating the recognition after the delivery of a specific word and saving an adaptation session in order to make specific recognitions constant and regular for a certain user. The CloudCAST cloud server will be available within the next year, so the 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.

4 Conclusion

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

The versatility of the system developed has allowed the creation of a tool that is able to replace specific products belonging to the assistive technology sector for home automation control, which currently have high costs, functional limits and a certain level of obsolescence due to the evolution of home automation and IoT technologies.

Thus, the solution developed can be considered a starting point for the creation of low cost custom home automation systems, useful for improving the levels of independence 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, in the future, the possibility of an efficient use by dysarthric users.

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 (2015) Remote speech technology for speech professionals—the CloudCAST initiative. In: Paper accepted for 6th workshop on speech and language processing for assistive technologies, Dresden, Germany, 11 Sept 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 8–12 Sept 2016, San Francisco, USA
3. <http://cloudcast.rcweb.dcs.shef.ac.uk/>
4. del Zanna G, Malavasi M, Vaccari G (2009) Manuale illustrato per la domotica a uso sociale-la casa flessibile al servizio dell'uomo. Tecniche Nuove
5. <https://community.openhab.org/>
6. <https://github.com/openhab/openhab/wiki>
7. Agosto R, Ioele FM, Malavasi M, Martinuzzi S, Motolese MR, Rimondini M (2014) Living in the living lab! Adapting two model Domotic apartments for experimentation in autonomous living in a context of residential use. In: Longhi S, Siciliano P, Germani M, Monteriù A (eds) Ambient assisted living—Italian forum 2013, pp 325–333
8. Krantz O (2012) Assistive devices utilization in activities of everyday life. A proposed framework of understanding a user perspective. *Disabil Rehabil Assistive Technol* 7:189–198
9. https://it.wikipedia.org/wiki/Raspberry_Pi