Multimodal Interaction for Accessible Smart Homes*

Nuno Almeida
DETI/IEETA, Universidade de Aveiro
Aveiro, Portugal
nunoalmeida@ua.pt

Samuel Silva
DETI/IEETA, Universidade de Aveiro
Aveiro, Portugal
sss@ua.pt

António Teixeira
DETI/IEETA, Universidade de Aveiro
Aveiro, Portugal
ajst@ua.pt

Maksym Ketsmur
DETI/IEETA, Universidade de Aveiro
Aveiro, Portugal
mvk@ua.pt

Diogo Guimarães
Bosch Termotecnologia S.A.
Aveiro, Portugal
Diogo.Guimaraes@pt.bosch.com

Emanuel Fonseca
Bosch Termotecnologia S.A.
Aveiro, Portugal
Emanuel.Fonseca@pt.bosch.com

ABSTRACT

Nowadays, houses are being equipped with new smart products and smart sensors, from multiple manufacturers, each offering their own options for interaction with providing varying degrees of usability and user experience. This diverse nature of smart homes and buildings, in general, poses new challenges to interaction design and accessibility. To tackle them, human-building interaction needs to move from articulating different interactive artifacts towards an holistic view of the house as an interactive ecosystem. In a joint effort with Bosch Termotecnologia, S.A., and profiting from recent contributions including an architecture and framework supporting multimodal Interaction, the authors aim to explore novel ways of approaching interaction design with a smart house and proposing smart home applications for all. This paper presents the status of this ongoing work, in the scope of project Smart Green Homes, proposing how multimodal interaction can be supported in the scenario of a smart home and showing first results of tackling human-building interaction through a home assistant serving a family in their daily interactions with the house.

CCS CONCEPTS

• **Human-centered computing → User interface management systems**; **Accessibility systems and tools**; *Ubiquitous and mobile computing systems and tools*; *Accessibility technologies*;

KEYWORDS

Smart Homes, Accessibility, Interfaces and interaction, Multimodal Interaction, User Centred Design, Personas, Design for All , Multi-platform design, Speech interfaces, Context–aware interfaces.

ACM Reference Format:
Nuno Almeida, Samuel Silva, António Teixeira, Maksym Ketsmur, Diogo Guimarães, and Emanuel Fonseca. 2018. Multimodal Interaction for Accessible Smart Homes. In *DSAI 2018: 8th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion, June 20–22, 2018, Thessaloniki, Greece.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3218585.3218595

*In the scope of project Smart Green Homes (SGH).

DSAI 2018, June 20–22, 2018, Thessaloniki, Greece

ACM ISBN 978-1-4503-6467-6/18/06…$15.00
https://doi.org/10.1145/3218585.3218595

1 INTRODUCTION

Smart homes have gained popularity, in the last few years, boosted by technological evolutions supported, for instance, on the evolutions stemming from the Internet of Things and lower technology costs. New products, such as smart lights, smart plugs, smart sensors and even smart home appliances are widely spread. Each manufacturer offers an application capable of handling and control their own devices, but if users use more than one brand they have to use multiple applications. Connecting all these devices to use them seamlessly is complex and poses several challenges.

Currently, there are some solutions to tackle and aggregate the diversity of devices into a more unified interaction and user experience, but, in general, do not originate from an effort of designing for a diverse set of users. Taking, for instance, Amazon Echo, it can aggregate and control some smart devices supporting voice control by programming Alexa's skills. However, and even though speech is our most natural way of communication, such a solution is not suitable for all, e.g., people with hearing problems, nor for all situations. To fully harness the smart devices' capabilities in service of people's expectations and needs the approach adopted to support interaction design needs to deliver easy communication with the house considering the potential diversity of abilities and needs of its occupants. Also, the amount, complexity and richness of information available from the smart home demands diverse ways of communicating it to the users, in context, such as graphics with data regarding smart sensors or the bits and pieces of home appliances such as a water heater.

1.1 Smart Green Home (SGH) project

Our work is motivated by a set of challenging scenarios of Human-Building Interaction [2], from smart homes to smart campi, and a notable example is the ongoing project Smart Green Homes (SGH), a partnership between University of Aveiro and Bosch Termotecnologia, S. A. where the challenges go even beyond the integration of the different technologies and applications. In the household, we need to ensure an increased awareness and optimization of how energy resources are being spent (e.g., in water heaters or lighting) and its associated environmental impact without loosing sight of comfort and well being [3]. In this regard, the literature shows that fully automated methods for managing these aspects are not the best option, making it paramount that users can take control and provide feedback [18]. This entails that the smart home needs to be accessible, highly adaptable, and reach users in their own context, enabling a certain level of mutual “understanding” between the house, as an entity, and each occupant [21].

1.2 Contributions and Overview

In this article, we address the complexity and needs of a diverse set of occupants, in a smart home, by proposing a vision and implementation for multimodal interaction with homes and a multimodal home assistant that serves as an accessible and adaptable way of communicating with the house, controlling its features and accessing a wide range of data and information.

The remainder of this document, after a brief overview of related work, presents our overall vision for multimodal interaction with homes, proposing how a multimodal architecture and framework can support interaction in this complex ecosystem. Next, we look into the concrete scenario of project SGH and present a family of Personas, characterizing the envisaged household, followed by a presentation of a multimodal multidevice home assistant and several examples of its use. Finally, we draw some conclusions and set the routes for future developments.

2 BACKGROUND AND RELATED WORK

This sections provides a brief overview of the topics more closely related with the context of our work: smart homes, how interaction with them is supported, and to what extent of accessibility.

2.1 Smart homes

The term "smart homes" can be described as a house capable of enhanced monitoring and control functionalities. These houses equipped with advanced devices allows to collect house information but also to control appliances [19, 26].

Nowadays, environments are populated by a myriad of devices including smartphones, smartwatches, smart home appliances or sensors harnessed with technologies enabling richer and more natural ways of interaction (e.g. speech, gestures), moving towards the way humans communicate with each other. Such diversity of interaction options (modalities) potentially increases usability rates and acceptance and these coexisting devices create novel opportunities for an extended interaction ecosystem that could blend the characteristics of all devices into a seamless interaction experience.

Improving usability is particularly relevant for scenarios such as Pervasive Assistance or Ambient Assisted Living (AAL), allowing technology to reach everyone, including, but not limited, people with special needs and elderly.

2.2 Interaction with Smart homes

Devices such as Amazon Echo with Alexa [7] or Google home with the Google assistant [20] enabled a more natural way to interact, using speech, with smart devices and working as a hub to control all devices. These devices are endowed with speech recognition capabilities and are connected to the network, providing the users with the control of smart home devices. However they are very focused in speech interaction and control.

Smart devices manufacturers usually provides an application to allow the control of one device but also to collect data and present that data to the user.

2.3 Accessible Interaction with Smart home

Finding smart home applications to interact with the different household devices (e.g., lighting, heating, appliances), which are easy to use by all is not easy. For instance, Amazon Alexa or Google Home may enable a blind person to interact with the system in many situations it may not be suitable for a deaf person or even a person with special needs in a noisy environment.

The literature proposes some systems addressing interaction with the smart home by groups with a certain level of diversity in their set of abilities. Two notable examples are Bempong [9], which presents a system to enable deaf and hard-of-hearing to access and control a smart home, and the work by Pawale and Chidrawar [22] presenting an application to create an accessible display, targeting people of disabilities.

Other solutions tackling accessibility for people with disabilities based on mobile applications can be found in literature [11, 13, 24]. Oliveira [17] in their paper the authors present a evaluation of mobile application for smart homes using a screen reader.

Accessible solutions for smart home interaction are mostly limited to one device controlling everything or are oriented to one kind of disability. Given the variety of devices available and technologies, we argue that they can (and should) be used together to accomplish a ubiquitous and transparent interaction between humans and their houses.

3 A CONCEPTUAL VIEW FOR MULTIMODAL INTERACTION FOR ACCESSIBLE SMART HOMES

The scenario of a smart home is rendered highly complex by the number of technologies involved and by the transient nature of how these technologies move among rooms or enter and leave the household. Providing high levels of accessibility and adaptability needs to rely on a versatile infrastructure that ensures a wide pathway between the user, at one end, and the equipments (e.g., a water heater) at the other. In this section we present our vision for multimodal interactive smart homes based on two core options: the adoption of a decoupled multimodal interaction architecture, side-by-side with a supporting framework, and the consideration of an assistant, fostering communication and a certain level of understanding between the user and the home.

3.1 Multimodal Interaction with the Home

Our vision for the multimodal interactive smart home stems from our previous work on multimodal interaction [4, 5, 25]. We argue that the solution adopted for supporting the interaction with the home would benefit from a versatile, decoupled approach that can dynamically encompass multiple diverse devices and interaction technologies. Aligned with the W3C Multimodal Interaction Architecture [15], we propose an architecture to endow a smart home with multimodal interaction capabilities serving adaptable natural interaction for all, but also potentially improving accessibility for users with disabilities.

A simplified version of the adopted architecture supporting multimodal interaction is shown in Figure 1 and depicts its main components:

- The *Interaction manager*, the central component of the architecture, receives the events created by the input modalities and creates new messages to the output modalities. Several interaction managers can co-exist, in the architecture.
- The *Modalities* are the interface between the user and the system. Modalities can be for: (a) input, receiving the user's commands; (b) output, transmitting information to the user; (c) mixed, dealing with both input an output; and (d) passive, gathering information, for instance, regarding the environment.
- The *Fusion* module is a component that has some similarities with the interaction manager. It receives events from the modalities and generates new ones, with the particularity that this component can combine (fuse) the information of two events into one, an indispensable feature to enable truly multimodal interaction.
- The *Fission* module is the component responsible for selecting the output modality based on the nature of the content that needs to be transmitted, on the available output modalities, and on the user/task context, ensuring that the message is presented to the user.
- The *Context and User model* services store relevant information of the user context and preferences. Applications can update the information and/or access it to address adaptation to the user preferences.
- The *Evaluation Services* provide different levels of support for context-aware evaluation following the concepts and proposal described in [23].
- The *Other Services* broadly represent additional services for a wide variety of purposes, also for controlling different aspects of the smart home such as temperature and window blinds, with legacy equipment being integrated through portals [16].
- Finally, the *Application* can work as single or multi-device [5] and provides the application logic for the system.

For the communication among components and their configuration, we consider the standards adopted by the W3C in their multimodal architecture recommendation. Events are encoded using markup languages, such as EMMA (Extensible MultiModal Annotation markup language), and the interaction manager and fusion modules are configured using SCXML (State-chart XML) [8, 16].

Adopting this architecture, we propose that the different smart devices in a home (e.g., smart lighting, smart plugs) are integrated as modalities (with legacy devices integrated through portals [16], connected to the interaction manager, and that additional interaction modalities are distributed across the house to support, for instance, speech input and output, gestures and touch.

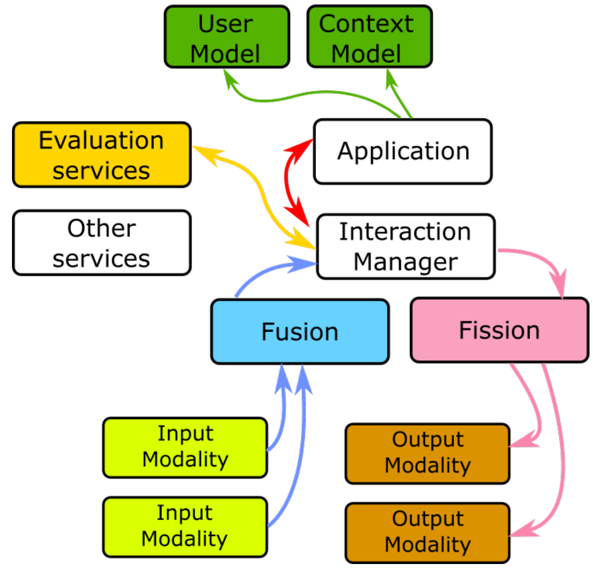

Figure 1: The different modules composing the multimodal interaction architecture considered as base for the proposed multimodal interactive home.

The Multimodal Framework

The proposed architecture and framework will offer developers the possibility to easily integrate new modalities and new smart devices to an existing smart home. The current proposal combines previous work, such as multidevice interaction [5], fusion of events from input modalities [6] and the use of the multimodal interaction framework in AAL [25].

The key improvements to our previous work on the multimodal interaction framework are the new and improved modalities that better suit the home environment usage scenarios, the portability of modalities to multi-platform and the user and context models.

A diversity of modalities using different transmission mediums, such as sound or visual, is important in the design for all.

- *Generic Speech Modality* included in the framework is multilingual cloud based modality, using web technologies such as HTML5 the audio is recorded and processed in the cloud. It runs in any platform that owns a browser. The modality is capable of speech recognition and speech synthesis.
- *User detection and identification (passive modality)* is also based in the cloud, when a user is detected and identified generates an event, which is used to update the current context. Therefore, the information can be used to adapt the system to the current user.
- *Graphical Output Modality* is responsible to deal with the content to be presented to the user, it can be text, graphical. The modality may adapt the content according to the user needs or preferences and to the screen used to display the information.
- *Gestures Modality* is a simple modality that detect body motion gestures using the kinetic technology. Since the Kinect SDK is a closed solution for a defined platform this particular modality can only run on windows. But as the architecture is decoupled the modality can be changed by other to do the same job, other technologies are available to do the same work.

3.2 Home Assistant

The multimodal interactive home, full of diverse interactive technologies, supported on the architecture described above, still remains a complex setting for users and our work is also driven by the need to turn the house into an interaction ecosystem rather than a set of interactive artifacts.

Our vision is that, on top of the proposed architecture, a home assistant is an essential component, managing the interactions with users, providing a sense of a single interface for everything, in the house. This assistant should inherit the multimodal nature of the architecture, with a particular emphasis (but not limited to) on speech interaction, our most natural way of communication, and contribute to an accessible home for all.

4 REQUIREMENTS FOR A MULTIMODAL MULTIDEVICE HOME ASSISTANT

A user-centered design [12] approach was considered to improve our understanding of the target users, their needs, and the envisaged context scenarios. The adopted methodology, aligned with Cooper et al. [14], was an important tool to support the dialog with our partners from industry and settle on a common ground for the system. In this context, we have created Personas to understand the needs of the targeted users, defined a set of usage scenarios (for the sake of simplicity, a subset of the devised scenarios is shown ahead, in the article, as scripts for the proof-of-concept system), created an AMITUDE (Application type, Modalities, Interaction, Tasks, Users, Devices, Environment) table and defined the requirements.

4.1 Personas

The created Personas [1] are the members of a typical Portuguese family composed of five members: the parents, two children and a grandmother.

Mother — Joana Sousa was born the 10th of November 1973 in Moita, Anadia county, Aveiro district. She is married with Afonso Sousa and has two children, a girl with 13 years old, Sofia Sousa, and a boy with 5 years old, Pedro Sousa. Joana has an MSc and works as an architect. Recently, she started working as a freelancer, and spends some of the days of the week working from her office, at home. As hobbies, she likes reading, playing the piano, and traveling, and is very interested in environmental causes and on sustainability. She uses electronic devices throughout her day to manage email and her agenda and have online meetings with colleagues. She uses a tablet to produce first drafts of ideas to send to clients.

She sees her home has a good scenario to understand energy consumption, a subject she also finds useful to improve her work as an architect.

Goal: Joana would like to have a better understanding of how resources (water, gas, electricity) are spent, at home, and the impact it has on the environment.

Father — Afonso Sousa was born in Porto, Portugal, on the 9th of January, 1972, and currently lives in Moita, Anadia county, district of Aveiro. He is married to Joana Sousa and has two children, a girl aged 13, named Sofia, and a boy aged 5, named Pedro.

He has a BSc in Management and he works as an external market manager in a multinational company. His work demands that he is constantly doing business travels. During weekends, one of his favorite activities is to ride a bicycle with his daughter Sofia and he enjoys watching a good football match.

Goal: Afonso would like to improve house management and have easier control of the devices.

Teenager daughter — Maria Sousa was born on the 30th of May, 2004 and lives with her parents and 13 years old brother in Moita, Anadia county, Aveiro district. She has a little brother aged 5.

Currently, she is in the 8th grade and is a very good student. She is proficient with technology, especially mobile devices, using them mainly to access different social networks and to perform content searches. Outside school she enjoys ballet and plays violin. At home, she likes watching television, reading and helping her mother with the domestic chores. She often likes a warmer temperature for her bath and bedroom. However, her parents explained to her that she needs to be responsible and be aware that exaggerating on the temperature costs money that could be spent in other important things. When questioned about her future and professional plans she mentions she would like to attend the university and graduate in Veterinary so she can treat animals.

Goal: Maria would like to control everything without leaving her room.

Youngest son — António was born on the 5th of September 2013 and lives with his parents and sister.

Currently, he goes to preschool during most of the week days. Sometimes, his mother keeps him with her at home. He still does not know how to read, but is very fond of playing with his mother's tablet and smartphone.

He also likes to play with legos and has a large collection in a big box, in his bedroom. Sometimes, his mother finds him, completely concentrated playing, late in the afternoon, without even remembering of turning the lights on.

Goal: António would like to be able control more things in the house, particularly those that are out of reach or demand reading capabilities.

Grandmother — Maria was born in 3 of December of 1948 in Moita, Anadia in the Aveiro district. Her husband passed away 5 years ago, and she currently lives with her daughter and family. When her husband was alive they in a small family house and have a small orchard and a few animals.

She was a nurse at a local clinic, but she is now retired. She passes her days at home or going out with a few friends she has. She enjoys being with her grandsons, and she normally picks him from the school.

Maria has some normal senior health problems, she wears glasses but that is not a big issue for her. Since a few years ago, with age, she became sensitive with cold.

Maria doesn't have much experience with computers, but she loves to surf the web, reading the last news and being updated in social networks.

Goal: Maria would like to use all the capabilities of the house assistant and have ways of automatically controlling the house.

4.2 AMITUDE

Adopting AMITUDE [10], Table 1 provides an overall description of the main features, overall context of use and required supporting technology for the home assistant. The main goal is to enable different users to interact in a multimodal way (using multiple modalities, e.g. speech, touch) to control and retrieve information regarding the home appliances from any room in the house.

Table 1: AMITUDE description of the Multimodal Multidevice Home Assistant Developed

AMITUDE	Description
Application Type	application, multi-user, multi-platform, with multiple forms of interaction and multi-device. possibly collaborating with other devices with the same application
Modalities	Input: • Spoken language commands (in Portuguese, English, German) • Touch and Multi-touch • User identification (face/speaker) • User location (passive/ implicit modality) • 3D gestures with mobile device Output : • 2D graphics (text, images, graphics, movies) • Spoken output (using Text-to-Speech) • Sound • (eventually) Emoticons
Interaction	Two-way communication (deliberate). Voice User Interfaces (command and control)
Tasks	In the next section are presented some scenarios, describing users interaction and tasks.
User	All family
Devices	Input: Kinect (depth + RGB), Microphone, Sensors (humidity, temperature, air quality). Output: Display (in tablet), Speaker, projector, Lights
Environment	at home, low/limited noise, no extreme lighting, internet access available, use when socially acceptable.

4.3 Requirements

For scenarios such as distributed architectures capable of handling dynamic multiple input and output modalities in multiple devices needs to be created. This architecture will provide the basis for the creation of an Assistant. The main requirements for both are:

4.3.1 User interaction requirements .

- Allow interaction using a mobile device
- Support Portuguese and English spoken interaction
- Allow users to interact in different ways: touch, speech, gestures, user identification (implicit).
- Offer redundant output and input alternatives
- Interaction with the system should be possible in different rooms
- The language that is used by the application should be simple and informal since technical language can be difficult for the users
- Simple to use and avoid content overload, i.e., large amounts of content and information at once.
- Adapt the application content based on the users' characteristics and the context.
- Allow the configuration of the application to suit users' preferences.

4.3.2 Functional requirements . As the Assistant will be the interface with the Home and the support architecture/infrastructure, the functions relevant for the user will be its responsibility, and, in consequence, the relevant functional requirements for the Assistant are:

- Retrieve information from home appliances and take action, e.g., detect temperature changes and notify user
- Control home appliances, lighting and personalized temperature
- Can call assistant

4.4 Technical requirements

The technical requirements for the assistant are two-fold. On one hand they focus what the ICT infrastructure needs to provide and, on the other, the technological support where the assistant will run.

4.4.1 For the architecture/infrastructure.

- Support available in all rooms
- Controllable lights
- Services providing info on aggregated sensors data
- Direct integration of sensors as passive modalities
- Provide easy integration of input modalities available in house and devices (ex: users tablets)
- API to control appliances

4.4.2 For the assistant.

- Multiplatform
- Usable inside and outside the house

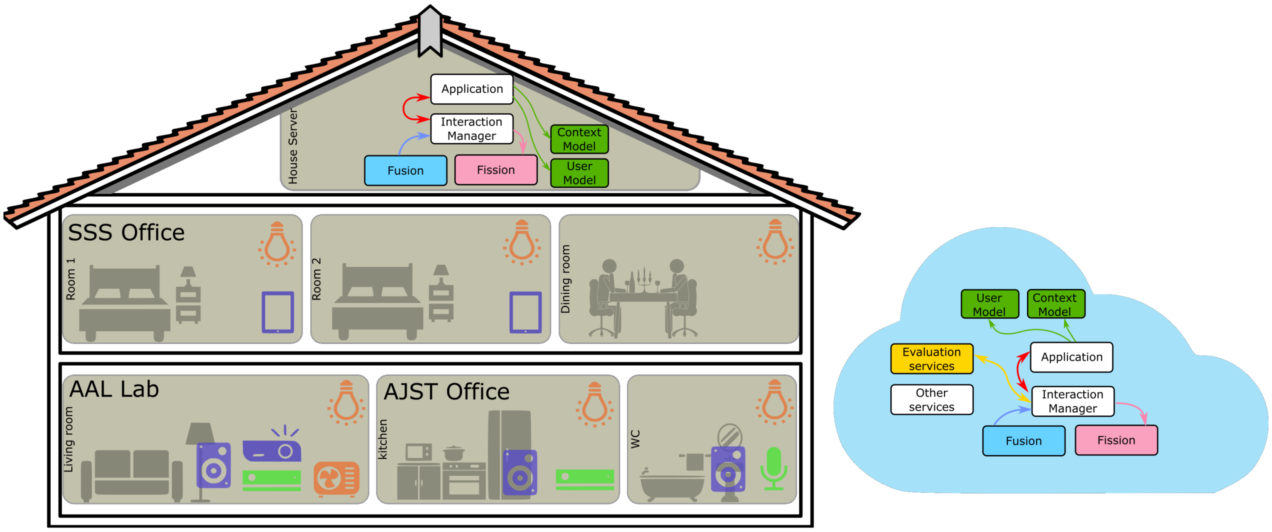

Figure 2: Smart home architecture, smart devices and modalities

5 THE MULTIMODAL INTERACTIVE SMART HOME

As proofs-of-concept for the proposed approach and assistant, we have defined a set of usage scenarios deemed relevant for the household context and implemented the required infrastructure and application logic to tackle them.

5.1 The household

In order to have a concrete scenario to support the implementation of our ideas, we designed a house and populated it with smart devices and devices supporting human-building interaction. Figure 2 shows a high level depiction of the components distributed across the house. First, the home server, providing the basic interaction infrastructure components, abiding to the proposed architecture. This infrastructure can also be supported (or complemented) by cloud services. Second, each house division shows a representation of the connected devices (e.g. light bulbs, heater, and multiple technologies to support modalities) that will always be installed. For the sake of simplicity, we do not show an explicit connection between the devices and the infrastructure (interaction manager). The supported modalities are as follows: (i) *Microphone:* speech input modality; (ii) *Kinect:* user identification modality; speech input modality; (iii) *Speaker:* speech output modality; (iv) *Projector:* graphical output modality; (v) *Tablet:* touch modality; speech input modality; graphical output modality; speech output modality;

For the implementation, we mapped each home division to a room in our lab. This allowed a more tangible setting to perform some testing and to better understand the scenarios.

User and Context model

The user and context models are services with databases capable of storing and handling useful information regarding users and usage contexts, respectively. In this first instantiation our User and Context models are document based (JSON like) providing enough flexibility to add new parameters, in the future.

The user model stores information related to the users preferences regarding interaction or even custom parameters for certain home appliances. Additionally, it stores some personal information, including security information, for instance if the user identification modality is allowed to share the result with the the entire system.

The User Context stores information of the house and its rooms, such as the available devices and interaction modalities, the users in each division, and lighting status. Table 2 illustrates an example of the structure and data for the living room.

Table 2: Example structure and contents of the user context regarding the living room.

JSON

```
{    "divId" : "livingroom",
     "devices" : [
          "minipc",
          "speaker",
          "projector",
          "kinect",
          "lights",
          "heater" ],
     "modalities" : [
          "speech_in",
          "speech_out",
          "user_identification",
          "GUI"],
     "dynamicDevices" : [],
     "dynamicModalities" : [],
     "lightN" : 1,
     "current_context" : {
          "currentUsers" : [],
          "numberOfUsers" : 0,
          "lightState" : 0,
          "heaterState" : 0,
          "roomTemperature" : 20 } }
```

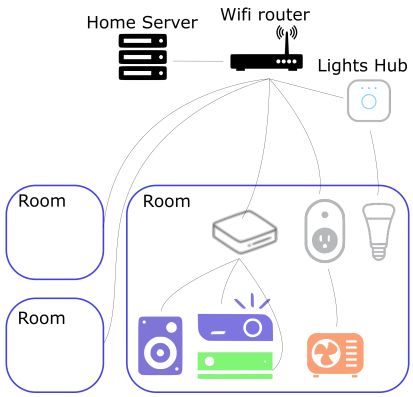

Figure 3: Technological infrastructure supporting the multimodal interactive home.

5.2 Technological Infrastructure

In the proposed architecture a home server (refer to Figure 3) will provide services and will be connected to a router, which connects to the home appliances and modalities locally supported by a mini computer. Additional modalities can be connected to the system such as those instantiated in a tablet.

5.3 Proof-of-concept Scenarios

The presented scenarios, selected from those considered to define the requirements, illustrate how the proposed architecture and assistant serve concrete purposes. With each scenario, we also present the architectural infrastructure set in place to support it. These proof-of-concept scenarios depict actual implementations, illustrating the current state of developments for the proposed multimodal interaction assistant.

Scenario 1 (Fig. 4): The father wakes up and goes to the living room. The system is configured to give information about the weather the first time it sees him, each morning. The system speaks 'Good morning, today it is a rainy day, I'm preparing your bath setting the water to your preferences'. The father woke a bit latter and he needs to bath quicker than usual. To know how much time he usually spends taking a shower he asks the system for the information, "How long did my last baths last?". The system shows a graphic with the time he took to take showers each day in the past week.

Scenario 2 (Fig. 5): The mother gets up, to turn on the lights she tells the system to turn on the lights by saying "dim the light to lower brightness". This way the intensity of the light will not be to high. She makes the bed, prepares herself, and then leaves to the kitchen to have breakfast. While she is preparing the breakfast, she remembers the light on the room is on, so she says to the system, "turn off the light" using her tablet to touch the room division. The fusion will combine the two interactions to turn off the lights in her room.

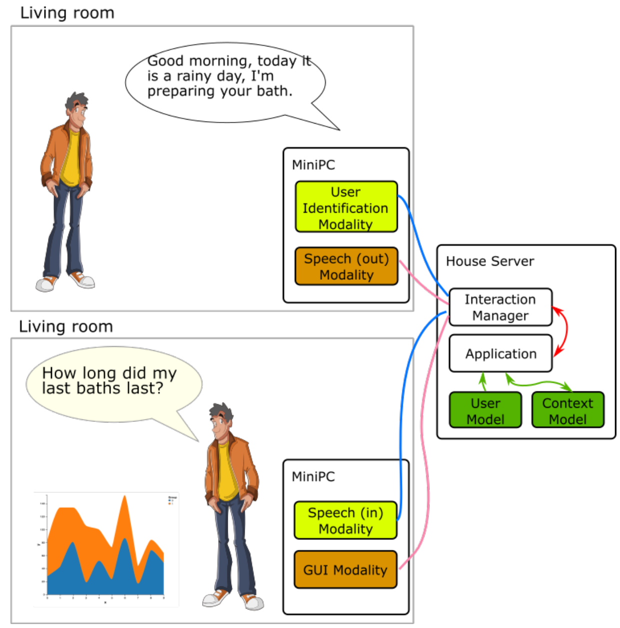

Figure 4: Diagram illustrating the scenes of Scenario 1 and modalities involved.

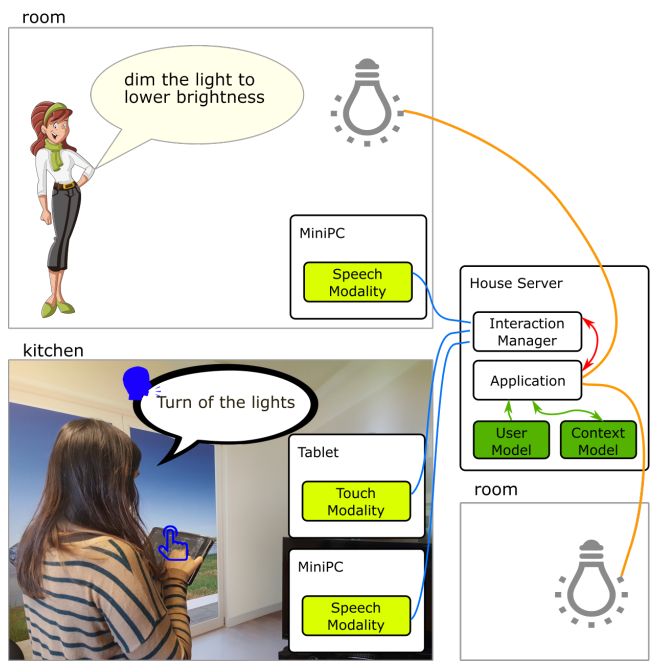

Figure 5: Diagram illustrating the scenes of Scenario 2 and modalities involved.

Scenario 3 (Fig. 6): The smart home system detects that the air quality is poor and tries to inform someone. The user identification modalities update the context model, and the system knows that the father is at home, in the kitchen, cooking. The system shows an alarm to the father by blinking the lights, in the kitchen. The father realizes the alarm and asks the system, 'what is the problem?'. The system responds to the father: 'the system detected bad air quality, you should open the window'.

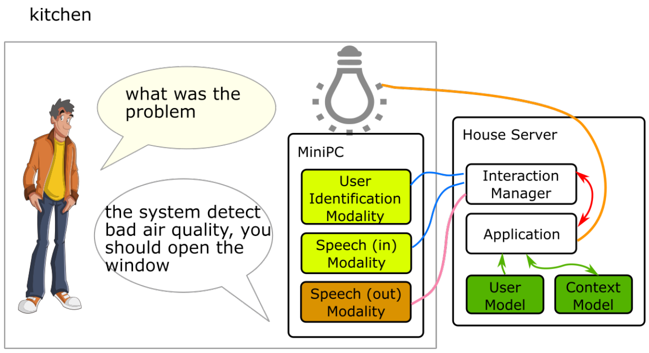

Figure 6: Diagram illustrating the scenes of Scenario 3 and modalities involved.

6 CONCLUSION

In this paper, we propose an architecture for an accessible smart home assistant, enabling multimodal interaction and addressing the needs of integration of different smart devices/components in this complex scenario. The proposed conceptualization and architecture have a modular and decoupled approach, with the different modules loosely distributed around the house and modified or changed independently. The diversity of modalities and ways of interaction enables an interaction for a diverse set of users including people with disabilities.

In order to instantiate a first proof-of-concept of the proposed assistant, a user-centered design approach was adopted and a set of Personas and scenarios used to define the requirements. The infrastructure needed to support a subset of the scenarios, providing a basic set of devices and modalities has been deployed and the capabilities of the assistant illustrated in real contexts proving the suitability of the proposed architectural approach.

At its current stage, the system has several modules, which allowed to create the proof-of-concept scenarios. To truly fulfill our entire vision of the assistant some improvements to the modalities have to be done, such as more adaptable modalities to user and context model, exploration of new modalities and devices. Also, in the future, we intend to do more iterations of our design methodology, finding new requirements by assessing the current prototypes.

ACKNOWLEDGMENTS

Research partially funded by project Smart Green Homes (POCI-01-0247-FEDER-007678), a co-promotion between Bosch Termotecnologia S.A. and the University of Aveiro, and IEETA Research Unit funding (UID/CEC/00127/2013.). Samuel Silva's work is funded by Portugal 2020 under the Competitiveness and Internationalization Operational Program, and by the European Regional Development Fund through project SOCA – Smart Open Campus (CENTRO-01-0145-FEDER-000010).

REFERENCES

[1] Tamara Adlin and John Pruitt. 2010. *The Essential Persona Lifecycle: Your Guide to Building and Using Personas.* Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

[2] Hamed S. Alavi, Denis Lalanne, Julien Nembrini, Elizabeth Churchill, David Kirk, and Wendy Moncur. 2016. Future of Human-Building Interaction. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16).* ACM, New York, NY, USA, 3408–3414. https://doi.org/10.1145/2851581.2856502

[3] Hamed S. Alavi, Himanshu Verma, Michael Papinutto, and Denis Lalanne. 2017. Comfort: A Coordinate of User Experience in Interactive Built Environments. In *IFIP Conference on Human-Computer Interaction.* Springer, 247–257.

[4] Nuno Almeida. 2017. *Multimodal Interaction - Contributions to Simplify Application Development.* Ph.D. Dissertation. Universidade de Aveiro.

[5] Nuno Almeida, Samuel Silva, António Teixeira, and Diogo Vieira. 2017. Multi-Device Applications Using the Multimodal Architecture. In *Multimodal Interaction with W3C Standards - Toward Natural User* (1 ed.), Deborah Dahl (Ed.). Springer International Publishing, 367–383.

[6] Nuno Almeida, António Teixeira, Samuel Silva, and João Freitas. 2016. Towards Integration of Fusion in a W3C-based Multimodal Interaction Framework: Fusion of Events. In *Proc. IberSpeech.* Lisboa.

[7] Amazon. [n. d.]. Amazon Alexa. ([n. d.]). https://developer.amazon.com/alexa

[8] Paolo Baggia, Daniel C. Burnett, Jerry Carter, Deborah A. Dahl, Gerry McCobb, and Dave Raggett. 2009. EMMA: Extensible MultiModal Annotation markup language. (2009). http://www.w3.org/TR/emma/

[9] Joan Bempong, Kate Gleason, Joseph Stainslow, and Gary W. Behm. 2015. Accessible Smart Home System for the Deaf and Hard-of- Hearing.

[10] N Bernsen and L Dybkjaer. 2009. *Multimodal Usability.*

[11] Brody Bruns and Caleb Ogbonnaya. 2017. Expanding the Accessibility of Conventional Smart Home Systems. In *Proceedings of the SouthEast Conference.* ACM, 255–258.

[12] Stephanie Chamberlain, Helen Sharp, and Neil Maiden. 2006. Towards a Framework for Integrating Agile Development and User-Centred Design. In *Extreme Programming and Agile Processes in Software Engineering SE - 15*, Pekka Abrahamsson, Michele Marchesi, and Giancarlo Succi (Eds.). Lecture Notes in Computer Science, Vol. 4044. Springer Berlin Heidelberg, 143–153. https://doi.org/10.1007/11774129_15

[13] Paulo A Condado and Fernando G Lobo. 2015. A system for controlling assisted living environments using mobile devices. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility.* ACM, 33–38.

[14] A. Cooper, R. Reimann, and D. Cronin. 2007. *About Face 3: The Essentials of Interactive Design* (3^{rd} ed.). Wiley Publications.

[15] Deborah A. Dahl. 2013. The W3C multimodal architecture and interfaces standard. *Journal on Multimodal User Interfaces* 7, 3 (apr 2013), 171–182. https://doi.org/10.1007/s12193-013-0120-5

[16] Deborah A. Dahl. 2017. Standard Portals for Intelligent Services. In *Multimodal Interaction with W3C Standards.* Springer International Publishing, Cham, 257–269. https://doi.org/10.1007/978-3-319-42816-1_11

[17] Gabriela Amaral Araújo de Oliveira, Raphael Winckler de Bettio, and André Pimenta Freire. 2016. Accessibility of the smart home for users with visual disabilities: an evaluation of open source mobile applications for home automation. In *Proceedings of the 15th Brazilian Symposium on Human Factors in Computer Systems.* ACM, 29.

[18] Giuseppe Desolda, Carmelo Ardito, and Maristella Matera. 2017. Empowering End Users to Customize Their Smart Environments: Model, Composition Paradigms, and Domain-Specific Tools. *ACM Trans. Comput.-Hum. Interact.* 24, 2, Article 12 (April 2017), 52 pages. https://doi.org/10.1145/3057859

[19] Steven K Firth, Farid Fouchal, Tom Kane, Vanda Dimitriou, and Tarek M Hassan. 2013. Decision support systems for domestic retrofit provision using smart home data streams. CIB.

[20] Google. [n. d.]. Google Assistant. ([n. d.]). https://developers.google.com/assistant/sdk/overview

[21] Denis Lallanne et al. 2016. Human-Building Interaction in the Smart Living Lab. In *Future of Human-Building Interaction workshop, 34rd ACM Conference on Human Factors in Computing Systems (CHI 2016).*

[22] Satish R. Pawale and Sadhana K. Chidrawar. 2015. Real Time Implementation of Accessible Display Design to Control Home Area Network.

[23] Carlos Pereira. 2016. *Dynamic Evaluation for Reactive Scenarios.* Ph.D. Dissertation. Universidade de Aveiro (MAPi).

[24] Samuel Tang, Vineetha Kalavally, Kok Yew Ng, and Jussi Parkkinen. 2017. Development of a prototype smart home intelligent lighting control architecture using sensors onboard a mobile computing system. *Energy and Buildings* 138 (2017), 368–376.

[25] António Teixeira, Nuno Almeida, Carlos Pereira, Miguel Oliveira e Silva, Diogo Vieira, and Samuel Silva. 2017. Applications of the Multimodal Interaction Architecture in Ambient Assisted Living. In *Multimodal Interaction with W3C Standards - Toward Natural User* (1 ed.), Deborah Dahl (Ed.). Springer International Publishing, 271–291.

[26] Charlie Wilson, Tom Hargreaves, and Richard Hauxwell-Baldwin. 2015. Smart homes and their users: a systematic analysis and key challenges. *Personal and Ubiquitous Computing* 19, 2 (2015), 463–476.