
AUI Project 16-17

ABI for SACRE

Giovanni Brena, Andrea Canale, Edoardo Longo Politecnico of Milan

January 26th, 2017

Loneliness for elderly has always been an ethical issue of central importance. Aging can generate feelings of marginalization, especially if it is not possible for the family to provide ongoing support. Difficulties for old people to communicate with other people is the basis of the problem. This is why we often end up dispersing a huge amount of experience and stories about people's lives, images and memories that only those who lived that moments can remember. ABI is an application that was created with this intention: provide a narrative and conversational channel for the elderly who are alone. In this paper we will describe the initial stage of project development. During this phase it has been designed a Wizard of Oz version to be tested on a small sample of users to get a ground-truth for future developments.

Team

ABI GitHub Repository

Giovanni Brena
giovanni1.brena@mail.polimi.it
878773



Andrea Canale
andrea.canale@mail.polimi.it
858638



Edoardo Longo
edoardo.longo@mail.polimi.it
841677



Contents

Glossary & Acronyms

- **Wizard of Oz:** research experiment in which subjects interact with a computer system that subjects believe to be autonomous, but which is actually being operated or partially operated by an unseen human being [?].
- **Stable mental health:** we mean elderly that have not particular mental disease (e.g. Alzheimer)
- **Operator:** the operator is the person that type in the server application substituting the autonomous robot conversation.
- **Desktop application:** desktop application means any software that can be installed on a single computer (laptop or a desktop) and used to perform specific tasks.
- **Server (application):** we intend the application used by the operator.
- **Sentiment analysis:** this is the operation performed by IBM Alchemy that extract entity, relationship and sentiment with a score.
- **Porting:** it is the process of adapting software so that an executable program can be created for a computing environment that is different from the one for which it was originally designed.
- **FLAC:** it is an audio coding format for lossless compression of digital audio.
- **FTP:** The File Transfer Protocol (FTP) is a standard network protocol used to transfer computer files between a client and server on a computer network.
- **API:** Application Programming Interface (API) is a set of subroutine definitions, protocols, and tools and for building application software

1 INTRODUCTION

In many developed countries, including Italy, Japan and many other, the ratio of senior citizens continues to increase. It has been estimate that by 2050 there will be more than 2 billion elderly people in the world, and this number will for the first time in history exceed the number of teenagers[?]. For these reason in the last decade, there has been an increased interest in the use of robots in eldercare from researcher and scientific community[?]. I would just remind you the great success of *Paro*, *Pearl*, *Huggable* and many other. It's also proved that often the elderly lives alone and distant from their family and relatives, which would listen stories and memory of their dear.

Our project grow in this context. On the one hand, it wants to ensure to the elderly a company against loneliness and in the meantime letting sons and grandchildren to listen senior's conversations and stories. For these reasons we have developed ABI, it wants to solve the problem of loneliness in the third age.

In this initial phase or research, ABI is developed as desktop application, but it can be easily converted in a smartphone/tablet application.

The elderly, above the loneliness, want to tell about their past. In fact, they have a lot of experience, from both social point and an historical point. Very often these stories are lost, losing a big number of interesting tales that can create a sort of collective memory, with a great potential from an historical point of view. On the other hand the relatives, regardless of the fact that they have a very little time, want to live again the memories of their parents and grandparents in a simple and immediate way. With ABI this is possible thanks a web page with a research tab where the family can insert phrases or keywords regarding their beloved.

During the first phase ABI is designed as a *Wizard of Oz* version. In this step the senior believes that the dialogue is with an automatic system, but in reality the conversation is controlled by an operator that digits remotely the response. In this way we can test our application on real users and acquire a big number of conversation of elderly people for a research phase.

In this documentation we are explaining the fundamental phases of our research and the more technical details regarding the development of the software.

In the following section we clarify the goals and the needs of our target group. Section 3 talks about the state of the art presenting researches and projects on which we have based for the basic idea of ABI. In section 4 and 5 we explain our solution, respectively from the UX point of view and from the software development point of view. An empiric evaluation about our work is done in the section 6, while the critical aspects, potentiality and challenge are described in section 7. Finally, in the last part (8), we explain the future work. We append to this documentation the instruction manual for the desktop application.

2 TARGET GROUP, USER NEEDS AND GOAL

2.1 Main target group

ABI is used both from the elderly engaged in the conversation and both from the family which relive the stories of their darlings. So we identify two main target: the **elderly** and the **family**.

Elderly The elderly to which our project is focused live alone or in a nursing home. They are in stable mental health.

Family The family to which is focused ABI lives physically far away from their darlings and the visits of their darlings occurs occasionally. In the specific, in the family, the main target are the **grandchildren** of the elderly. This for the special *grandparents-grandchildren relationship*, in fact the last ones are often desirous to listen stories

2.2 Context and needs addressed

Elderly The primary need of an elder who live alone is just talking and be heard to alleviate the sense of loneliness. This is linked to tell about their self, often about the past remembering event of their youth.

Family The needs of the family are related to those of the elderly. In fact they desires their darling are well and does not feel alone in spite of the distance between them.

2.3 Goal of the project

The goals of ABI are mainly two: create **conversation** and keep **memory** of the past.

Conversation The conversation is the best way to make elderly feel less alone. The aim of ABI is to listen the elder, answer him and keep company during everyday life.

To satisfy these goals we want to:

- Engage the elderly into simple conversations
- Use basic social interactions and provide proper feedback
- Give satisfaction to the interlocutor

Memory With our project we do not want that the memories of elderly will be lost in time. This has a double meaning; on one side the elder is pleased to tell about his past, on the other side these memories can be heard by the family and grandchildren.

To satisfy these goals we want to:

- Collect stories and experiences as the elderly has reported
- Let the family recover the memories
- Have a real experience when the elderly conversation are listened by the family

3 STATE OF THE ART

As we said previously a growing number of projects address the development of conversational robots for experiments in eldercare. Most of the study are focused on the *eldercare companion robot*, ignoring the relationship that could be established with the elderly and the possibility for the family to live the relationship with a darling despite the distance.

In Japan, Wada and Shibata performed experiments[?] with a seal shaped robot (*Paro*). These experiments showed that a robot could have the same beneficial effect on elders that a pet can have, making them feel happier and healthier. In the US, was developed *Pearl* [?], a robot that could actually provide some assistance to elders, reminding people about routine activities such as eating, drinking, taking medicine and guiding them through their environments.

Research related on acceptance of a conversational robot, and about retrieving useful data from conversation is described by De Ruyter et al [?]. It concerned a robotic interface (the *iCat* made by *Philips*), which was tested in a *Wizard of Oz* experiment where the robot was controlled remotely by an experimenter while it was suggested that the robot was autonomous. This experiment was done in a laboratory setting, with adult, but not elderly participants. The participants were asked to participate in an online auction, by using the *iCat* interface. To measure acceptance was used a questionnaire asking enjoyment, ease of use, utility, acceptance and other topics. The aim of the study was to find out to what extent participants would use the *iCat* at home after having experienced it.

A similar example is *Robovie*[?]. It is a conversational childlike robot developed in Japan and tested for more than three months in an elderly care where elderly with stable mental health live. Also *Robovie* is remote controlled by a person in order to obtain data about conversational habit of the elderly and to grant the most realistic experience possible. The test has been very successful: the elderly accepted *Robovie* because it made them feel important with just little attentions. The other strong point of the research was the childish physiognomy of *Robovie* that helped the elderly interact with the robot in a way that was familiar to them, even though the robot itself was not familiar.

Building on the great achievement made from studies mentioned above, our work focus on keeping company and alleviating the sense of loneliness through an initial solution of the *Wizard of Oz* for initial phase of research. The surplus lie on the storage of elderly conversation,

with which we create a sort of diary about the person easily accessible from the family.

4 SOLUTION - UX DESIGN

4.1 General Approach

4.1.1 Interaction paradigm

Aiming to a target group like elderly put several issues about the interaction that need to be resolved. Specifically, the relationship of the elderly with new technologies is often difficult: when the distance between the *machine* and the interlocutor is excessively marked it could generate anxiety and restlessness in the elderly. For this reason we think that the correct approach to these limitations is to reduce at most physical interaction (interaction with UI). We consider interaction through audio and video channels appropriate for this scope, using just simple vocal interfaces to solve possible hearing and/or vision problem.

4.1.2 Relationship elderly-bot

We focus now on the interactions challenges of the real conversation. Watching elderly people telling stories it's observable that exists some specific and expressive peculiarity that can be traced back to these categories: response time, expressions and lexical. Response time changes from case to case, a fast communication could be lead to a unpleasant experience for the user like as well as a long response time could make the conversation unrealistic.

For this reason we think that a system that can adapt in a progressive way to the specific response mode of the single user.

Expression and lexical modes are related to the content of the conversations: which tone is correct to use with an elderly person? How to stimulate the user to talk more about himself and past experience to a bot? Which are the lexical terms suitable and easily understandable for the target? All these questions constitute the critical point of the conventional conversational system, or said in another way they are the main difficulties to overcome to reach comfortable dialogue for the target users.

We think that the way to overcome these difficulties is in the field trials, through tests with real users in order to study the interaction mode deeply and deploy voices, questions and answers as possible familiar.

4.1.3 Childish Style

Going on with the study we analyzed the behavior of the elderly and of the children. The relationship between the seniors and the little boy is very peculiar and it is different from elderly and man. The first one is more direct and not driven by rules, fear or embarrassment. Furthermore the narrative characteristic of our project address mainly to the dynamic between grandparent and grandchildren. For this reason we

provide a *childish-styled interface* (UI) that can infuse self-confidence and intimacy in the user, and can also stimulate him to have a confrontation with the system in a similar way to the one with his grandchild. The validation of these choices will be proved through the user-testing.

4.1.4 Contents, childhood and memories

In the first phase of experimentation, conversations are driven by an operator who listens to the elderly and manually develops the conversation. In order to reach the goal of a future perspective of creating automatic conversations, it is necessary to think about which are the most exciting themes to stimulate elderly talking. A test session could certainly provide precious data about the frequent memories, the events or the historical period mostly to refer to. As a preliminary analysis we consider that it could be a good idea to lead the user to remember memories of his childhood; the reason is that of a tendency of elderly people to generally idealize in their memory, that kind of moments (in most of the cases), as happy moments. Other possible peculiar themes could be connected to significant historical events (wars, political revolutions, scientific progress...) which remarked important moments. At last we think that another interesting scope is that of the family relationships, parents, sons, their own marriage.

4.2 Scenarios

In the following section we present two scenarios. The first one is a video scenario, the second one is textual.

4.2.1 Antonio and the mountain

Persona

- Name: Antonio
- Age: 84
- Job: pensioner
- Personality: pleasant and mournful
- Attitude: open-minded

Context Antonio, lives out of town in a terraced house. He has two sons and three grandchildren. The first one live in England with the family; the second son is very busy with the job and he come to find him usually on Sundays. Before retiring he was a journalist of a popular newspaper. So he loves to tell stories and to travel around the world.

Scenario The video scenario is available here: <https://www.youtube.com/watch?v=Jaluliachgc>

4.2.2 Maria Luisa with ABI for the first time

Persona

- Name: Maria Luisa

- Age: 79
- Job: pensioner
- Personality: sweet and polite
- Attitude: curious and meddler

Context Maria Luisa lives in a home in Venice. She lives alone from a while because all her sons went away for work. She stays most of the time alone in the house unless when her family or some acquaintance come to visit her in the weekend. The need of someone to talk is constant, but she cannot stay for long time outside because she is sickly. Also she wants to tell so much things to tell to her grandchildren.

Scenario Maria Luisa was talking with her daughter about his desire to have somebody to talk with and concerning about how many things wanted to tell to her grandchildren, when her daughter suggest her to use the new released ABI software. She accepted to use it, so the next weekend her daughter came to install it on her PC and explained how to use it and how her grandchildren can listen everything she told to ABI. She realized that she can finally delivered all her stories to her grandchildren and started using ABI more times in a day.

One day she remembered one event that happens when she was a child but, as most of the time, no one was there to listen her, so she sit in front on her PC, click on the ABI icon and push the START button and conversation start. ABI starts with greeting and a smiley face, things that cheer up and relax Maria Luisa, then it asks to Maria of what she want to talking about. Maria tells him that she want to tell a stories about her past and ABI says "*Sure, come on!*" with a surprise face. So Maria start to tell it what happened to her: in most of the conversation is just Maria Luisa talking and ABI nods, but when she says something particularly happy, ABI shows an happy and rotating smiley face or also when she tells melancholy tales ABI ask why that thing make her so sad. She finish telling the story and she feels quite tired because it has spoken a lot, so she says "*Goodbye ABI*", clicked on STOP button and close the application. That day, before going to bed, she called her daughter to tell that a new story was recorded and make her children listen it.

5 SOLUTION - IMPLEMENTATION

During the testing phase ABI is developed as a Java desktop application. It is combined with the web site for the search of the conversation and a database to store every single phrase. For the development was fundamental *Google* and *IBM Watson* APIs.

Below we list the main elements of hardware and software architecture.

5.1 Hardware architecture

We do not need a specific hardware architecture, in fact ABI is simply a desktop application. We only need a microphone for the voice of the elderly and the speakers to reproduce the voice of ABI.

5.2 Software architecture

Our project is composed by:

- Desktop Application
- Website
- Database

Very important is the use of *Application Programming Interface* (API).

5.2.2 Desktop application

In Figure 1, we display the main component of the application that we are going to explain:

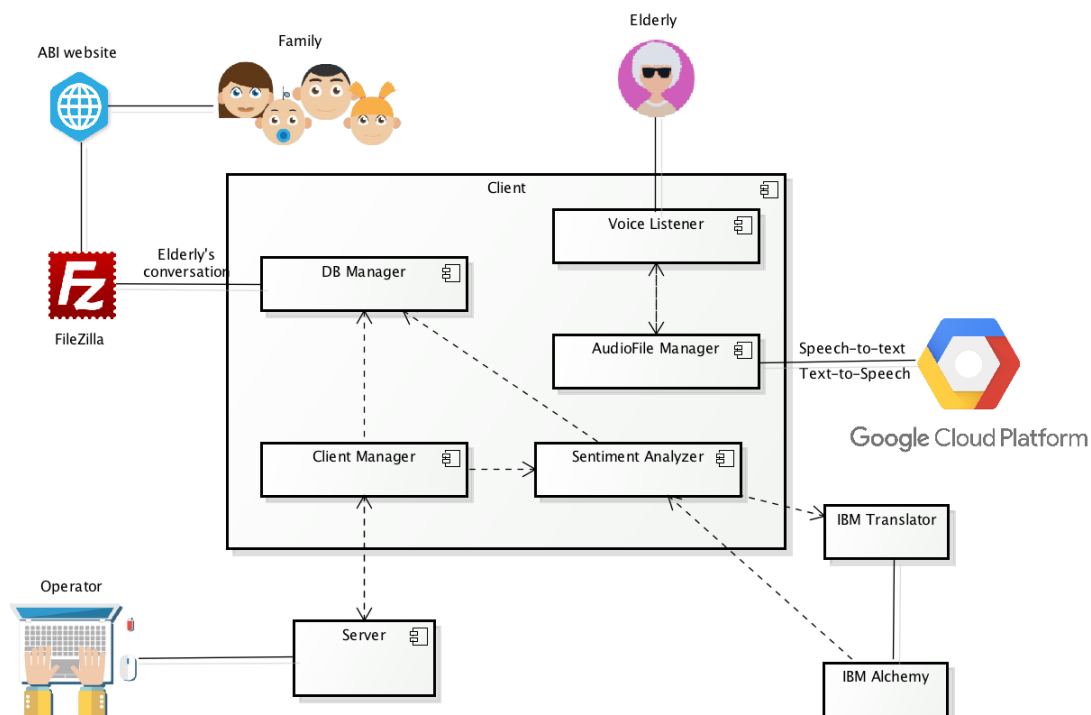


Figure 1: Component Diagram

5.2.1 API

For the deployment of our application we used the following APIs. We remember that we used all free version.

- **Google Cloud Speech:** to manage speech to text and text to speech.
- **IBM AlchemyLanguage:** it is the sentiment analysis. For each phrase Alchemy returns:
 1. Overall Sentiment: a score from -1 to +1. -1 shows a negative sentiment, 0 neutral one, +1 positive one.
 2. Anger, Disgust, Fear, Joy, Sadness. For each of these fields show how much that single emotion was found in the phrase. 0 show the absence, 1 is the maximum value.
 3. Entity: returns entity people or city.

AlchemyLanguage is available only in English, for this reason we have to use another API to translate from Italian to English.

- **IBM Language Translator:** used to translate the phrases said by the elderly to English for the analysis of AlchemyLanguage.
- **IBM Conversation:** to make a automatic conversation.

Elderly client Elderly clients is the core ABI. After launching the application, the GUI allows user to choose two different modes: *Wizard of Oz* and *Robot Conversation*. In both modes, ABI records voice and save it saved as a *FLAC* file. At this stage we used the *library J.A.R.V.I.S.* available on *GitHub* [?].

- **Wizard of Oz mode:** After recording user sentence, it is used *Google Cloud Speech* to text conversion. The resulting text is sent to the operator, saved in the database and analyzed by *IBM Alchemy* (*sentiment analysis*). At this point, the application waits for the operator response that is obtained as a string. By using the *Google text-to-speech* is converted into audio, reproduced and saved in the database. In addition, the operator can change the ABI expression displayed on the client using the commands in your console.

- **Robot conversation mode:** This function has been implemented for the purpose of testing the automated conversation, therefore without the use of an operator. Responses are provided by *IBM Conversation*: user *speech-to-text* input is analyzed by *IBM Alchemy*, which returns feeling, emotions and entities. Such data is loaded into the context variable *IBM Conversation* and sent to the latter in order to have a personalized response.

Server (application) The use of ABI Server is critical at this early stage of collection of conversations to get a solid base of ground-truth. ABI Server application is handled by an external operator who connects via IP address to the elderly client. It works as a simple chat: once obtained an elder phrase in the text string, the operator types a response that is sent to the client as text.

The following diagram (Figure 2) presents the interactions between clients, servers and the APIs.

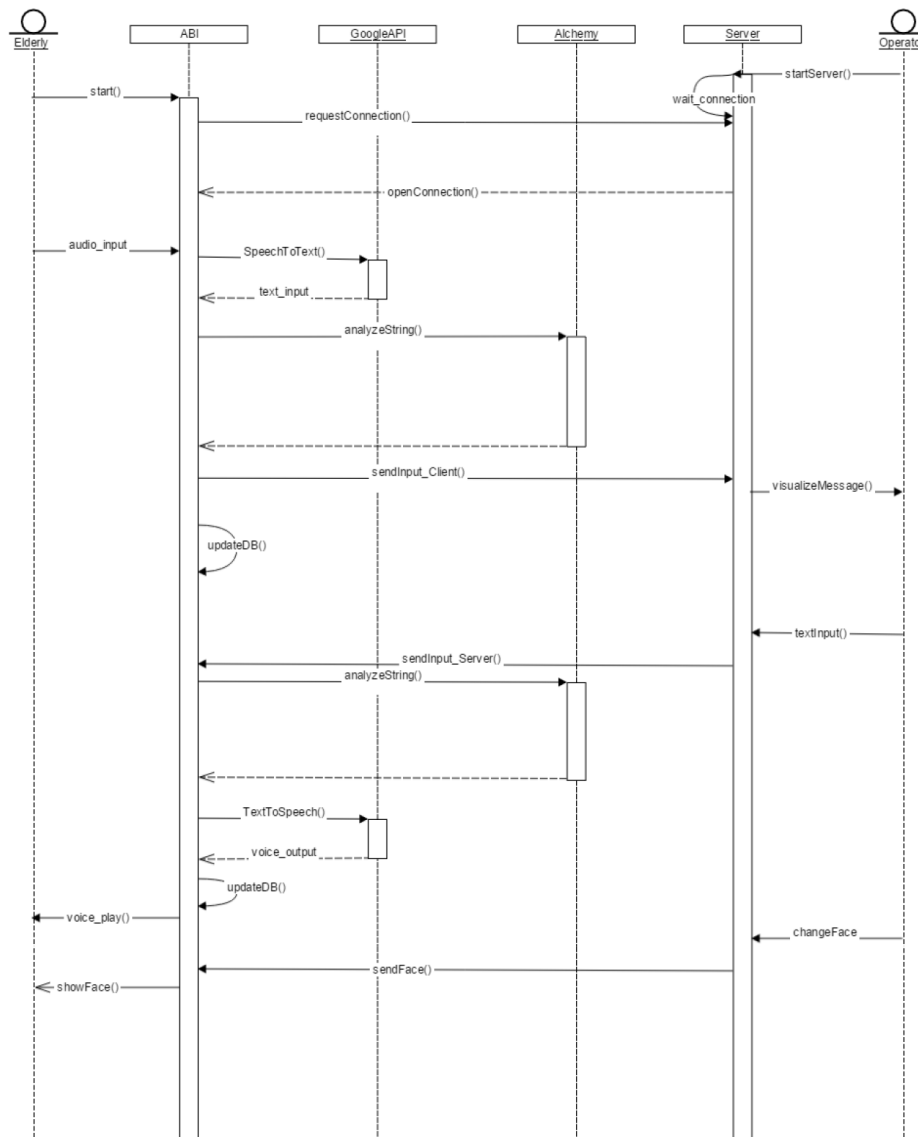


Figure 2: Sequence diagram

5.3 Database

The database is a fundamental component of ABI. On it all the conversations and the sentiment analysis are store. It was used *clearDB*, a relational database, provided to us in a free version thanks to *IBM Bluemix*. The database is composed of two tables: *Frase* (Phrase) and *Conversazione* (Conversation) In *Frase* are stored every phrase said by the elderly or by the operator with its timestamp, the value of sentiments returned by *Alchemy* and the id of the audio file. The table *Conversazione* identify every single conversation. There are the fields *data*, *ora inizio*, *ora fine*.

Client FTP For uploading audio files on the website at this stage we use *FileZilla*, a free cross-platform software that allows the network to transfer files via FTP (*File Transfer Protocol*)[?]. We opted for this solution for not going to burden the application logic.

5.4 Website

Search page of the website is available on: <https://abiproject.000webhostapp.com/search.html>.

Search Page In this page it is possible to search for one or more words said by the elderly or by the operator in the conversations' database.

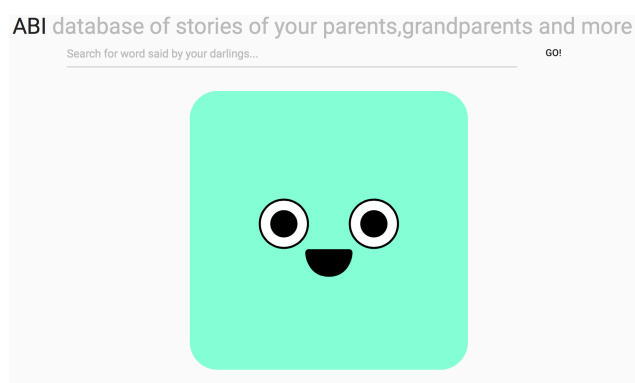


Figure 3: Search Page

Result Page On the next page are returned all the phrases matching the word you typed previously, including additional information such as the date and time.

Clicking on the phrase (or *Go to the dialog* button) will redirect to the full conversation between Elder and ABI.

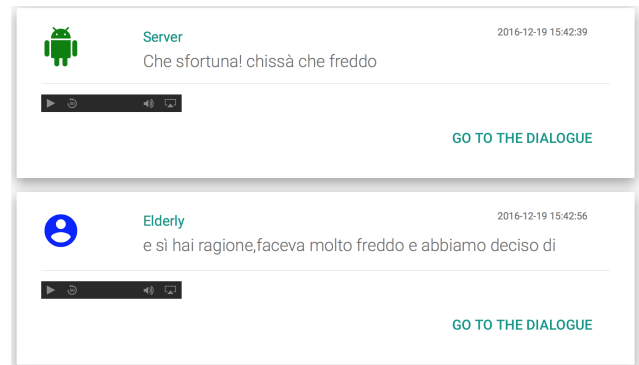


Figure 4: Result Page

This is the full dialog. Clicking on play button allows user to listen to recordings.

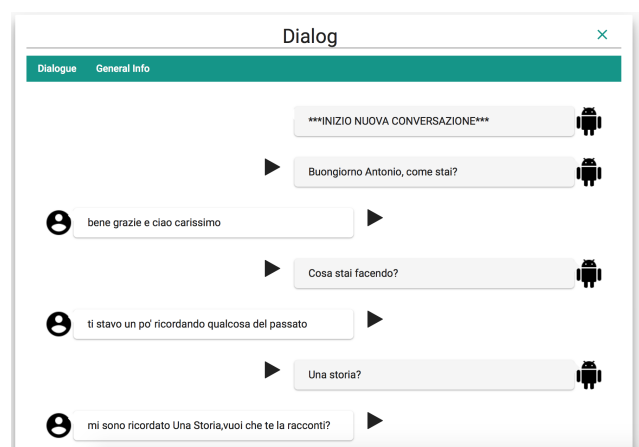


Figure 5: Full conversation

General Info tab shows the progress of the general feeling of the conversation, the date, and the number of sentences.

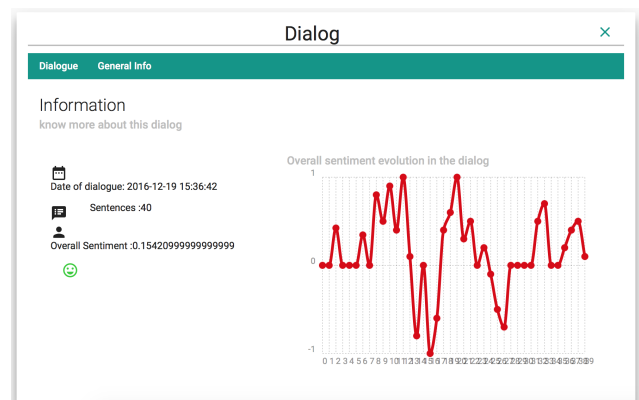


Figure 6: General Info Tab

6 EMPIRICAL EVALUATION

The work developed so far does not intend to produce a complete solution, but act as a base for exploration of the problem. We believe that the study undertaken must necessarily be followed by an empirical testing in order to validate the basics. The *ABI Wizard of Oz modes (Client-Server)* has been specifically developed to produce results through a User-Research.

6.1 Evaluation Goals

We intend to validate the idea that underlies ABI is realistic. Can an automated bot create a pseudo-human relationship with an elder in order to collect stories and memories?

We also want to perform usability and interface effectiveness tests on the specific target groups, with particular reference to the above mentioned *childish approach*.

6.2 Research Questions

The open-ended questions to which aims to answer the *user-research* concern the full conversation. Through the study of real conversations, manually controlled by an operator, we want to collect technical data on the average response time of the elderly. In that way we could better understand the correct timing to be assigned to the bot sentences.

Also we need to care about other details such as tone, voice type, vocabulary and language in order to put at ease the user and facilitate the relationship with the bot.

Another broad area of interest is the content of the conversations. Through field-testing we can gather insight on the which arguments are more relevant for users. Specific periods of life, events or persons of which the elderly tend to talk more will allow to design more engaging conversations automatic.

6.3 Test Execution Process

Test execution contexts may be different since it is not required a particular hardware equipment. We believe that short individual sessions can be performed both at home and in elderly care centers.

Two operators are necessary. The first operator is responsible for physically follow the elderly during the test, introducing the system and providing technical and emotional support.

The second operator acts remotely (from another room or elsewhere) by sending text or voice responses to sentences pronounced by the elder.

7 VALUE PROPOSITION

7.1 Challenges

Elderly Target

The use of the technology by the elderly is a crucial aspect in recent years. More and more people are us-

ing computers and smartphones, but still difficulties remains for old people to easily understand technology improvements. In our work we target people that already has some technological experience, though not experts.

The key issue is the use of an interface suited to the elderly, with clearly visible buttons and easy to understand. The volume of ABI entry must be adapted to the needs of the individual.

In this way we want to encourage the elderly in the use of our application, focusing on the immediacy of the GUI. In particular ABI interface shows only the start button to begin a dialogue with ABI. Once the dialogue is available only the STOP button to end the call.

7.2 Critical Aspects

7.2.1 Create a real conversation

Physical interaction In a normal two-person spoken interaction, there is a set of implicit rules that each party spontaneously follows in order to keep the conversation smooth and make the other side comfortable. Hands gesture, eyes and face movement and body orientation play an important role during social interaction[?]. In this first initial phase, ABI, is a desktop application so there are not physical interaction that can give a feedback to the elderly. These feedback is useful to get the elderly more involved in the conversation.

Response timing Another critical aspect of the dialogue is the speed of response on the part of the system, both in the version *Wizard of Oz*, both in the automated version. Audio must inevitably use the recognition services. Then the operator must type the answer and be again converted into audio. This process causes a slowdown in the speed of response that makes less fluid dialogue. A partial solution is to use the premium version of the various APIs that go to offer higher speed.

ABI voice ABI uses a synthesized voice and pronunciation of some words unrealistic. Often this problem is due to the still weak development of the Italian language in the *text-to-speech* APIs. Also we can not ensure that the elderly uses speakers with an optimal configuration. These problems can cause the elderly disaffection in the use of ABI.

7.2.2 Privacy

ABI must ensure privacy respect. The conversations of elders concern personal topics, kind of information that hardly want to share with strangers. It is important that conversations stored in the database are accessible only by family or, in the case study, only by authorized personnel. In addition ABI listening phase should be triggered by specific commands and in any case automatic.

Another critical aspect of privacy is the use of third party services. In our case the database, run by

ClearDB, and *Google APIs* and *IMB* to which we send each sentence pronounced by the user.

7.3 Potential

7.3.1 Collective Memory

The possibility to share stories to anyone in the world, could create a world collective memory which otherwise would be lost.

A social platform could create digital archives, for example about history, accessible by people all around the world.

7.3.2 ABI as assistant for healthcare

A possible extension of the project is to integrate health-related capabilities in order to meet the needs of the elderly health care.

Control the physical and mental health can be made in a newspaper not too invasive manner through simple questions in a conversation.

7.4 ABI is a good solution?

- **Elderly user:** ABI is designed specifically to provide to elderly users a simple and enjoyable tool to use.
- **Memory functionality:** The memory function, already amply explained in this document, it is something still not developed and ABI may be a precursor or even a standard.
- **Compatibility and Scalability:** ABI does not require any expensive hardware and can be integrated into any compatible device (smartphone, tablet, desktop...) without any great effort.

7.5 What is new in ABI?

Many studies have been done about elderly care and companion robots often with excellent results. Our project takes its place alongside these, but by developing a new concept. Unlike previous projects, the ABI target not only the elderly, but also the family and especially the grandchildren. Our application can also have a historical and social aspect. The use of ABI on a specific target population can collect large amounts of data relating to that period.

8 FUTURE WORK

Tablet / Smartphone client The smartphone and tablet solution enables rapid deployment and mobility.

Improved integration with text analysis algorithms

In the current version is an analysis of the general feeling and emotions for each individual elder sentence. The next step could be a deeper analysis of entities

(people, places), keywords and concepts in order to provide better conversations.

Video Conversations Video files could be included on elderly database, empowering the stories.

9 Used Tools

During the development of the project we use the following tools:

- **Eclipse Neon:** for the software development.
- **Sequel Pro:** for the management of the database.
- **GitHub Desktop:** for the software versioning.
- **Slack:** for the communication among the team.
- **Brackets:** for the web development.
- **FileZilla:** for the FTP transfer.
- **Adobe Premiere** and **Final Cut:** for the videos.
- **Google Drive** and **Dropbox:** for the files sharing.
- **L^AT_EX:** to compile and format the document.

References

- [1] Bella, M. and Hanington, B., *Universal Methods of Design*, 2012.
- [2] V. Logar, *The Role, Meaning and Dimension Health Service in the Homes for Elderly People*. (Slovenian: Vloga, pomen in obseg zdravstvene slube v Domovih za stareje obcane.) In: Pentek Metka (eds.): *Zdravo staranje*: 1. Nacionalna konferenca o promociji zdravja v Sloveniji Ljubljana. 29. In 30 marca 1995. First edition, Gerentoloko drutvo Slovenije, Intitut za varovanje zdravja RS, Ljubljana. (1995) 7078
- [3] M. Pollack, *Intelligent Technology for an Aging Population: The Use of AI to Assist Elders with Cognitive Impairment*, AI Magazine, vol. Summer, 2005.
- [4] T. Shibata, K. Wada, and K. Tanie, *Statistical Analysis and Comparison of Questionnaire Results of Subjective Evaluations of Seal Robot in Japan and U.K.*, 2003.
- [5] M. E. Pollack, S.Emberg and other, *Pearl: A Mobile Robotic Assistant for the Elderly*, 2003.
- [6] B. de Ruyter, P. Saini, P. Markopoulos, and A. J. N. van Breemen, "Assessing the Effects of Building Social Intelligence in a Robotic Interface for the Home", *Special Issue of IwC: social impact of emerging technologies*, 2005.

- [7] A. Sabelli, T. Kanda and N.Hagita, *A Conversational Robot in an Elderly Care Center: an Ethnographic Study*, 2011.
- [8] M. A. Sheili, F. Yang and L. Devillers, *Attention Detection in Elderly People-Robot Spoken Interaction*, 2014.
- [9] Marci Meingast, *Security and Privacy Issues with Health Care Information Technology*, 2006.
- [10] FileZilla, FTP protocol client. Website: <https://filezilla-project.org>.
- [11] J.A.R.V.I.S. Speech API. GitHub link <https://github.com/lkuza2/java-speech-api>.