Empirical Studies Aimed at Understanding Conversational Recommender Systems and Accessibility Aspects

Lucas Padilha Modesto de Araujo[0000−0002−4464−118X],
Cynthya Letícia Teles de Oliveira[0000−0001−9542−8189],
Kamila Rios da Hora Rodrigues[0000−0001−9176−8709], and
Marcelo Garcia Manzato[0000−0003−3215−6918]

University of São Paulo, São Carlos, Brazil
padilha.lucas@usp.br, cynthya@usp.br, kamila.rios@icmc.usp.br,
mmanzato@icmc.usp.br

Abstract. Conversational systems allow communication by voice, touch, and text between system and user through interactive interfaces. They also allow users to carry out tasks such as shopping, answering emails, and with few interactions, the user reaches the goal. The industry usually disregards the limitations of the elderly public when designing conversational systems. Such users may have problems with dexterity, vision, and hearing, impacting their interactions with the system and, consequently, requiring applications accessible to their cognitive and physical characteristics. The goal of this work was to propose a set of functional and non-functional accessibility requirements for the development of conversational systems for elderly. A first set of requirements were gathered through the Wizard of Oz technique, applied with eleven participants. For this, a conversational system, named Bob, was designed to recommend YouTube videos to the user. Recommendations were made according to the user's needs and preferences, adapting visual and auditory resources through the interface. After the trials, the results showed that, especially for the elderly population, conversational systems can also play the role of a friend or virtual company, with adaptable and adaptive accessibility and interaction settings.

Keywords: Human-Centered Computing · Empirical Studies in HCI · Conversational Recommender Systems · Accessibility · Elderly.

1 Introduction

The expressive amount of applications available has brought challenges to the design of accessible applications. The audience that uses the applications can be diverse, and therefore each one has their perceptions, needs, and interests. Elderly people, for example, who were not born immersed in technological languages need to migrate to a different reality from which they are used to. In addition, they have needs which are different from those usually faced by people

born with the use of technologies, such as young people and non-elderly people. This fact, consequently, brings the necessity of research on accessibility and usability in the technological context.

In the case of the elderly people, often, efforts to create solutions and tools that help are limited to visual aspects such as customizing contrast, adjusting font size on print screens, and easy pressing application buttons. But, there are other problems that are associated with usability, such as the inability to understand the procedures that make up the use of the proper tools for accessibility [3]. In other words, there is a problem to facilitate access to the different parameters offered by a solution regarding its usability, otherwise users may give up on using an application because they do not know how to configure it according to their needs and preferences. Providing accessibility means removing barriers that prevent people with disabilities from participating in substantial activities, including the use of services, products and information [24]. Conversational systems must meet this premise, in addition to considering usability requirements and being intuitive for users, especially for the elderly with low literacy or with difficulties and disabilities.

Conversational systems are designed to carry out a variety of tasks in order to simulate dialogues and interactions with humans. There are architectures capable of bringing relevant information in a conversational environment, such as Frequently Asked Questions Systems (known as FAQ's), which encompass frequently asked questions of a system based on a probabilistic model [1]. However, this strategy is not very relevant in the context of Natural Language Processing (NLP), as the user is not able to express their request or ask about other issues that are not frequent and related to the FAQ proposal.

Proposed in 1964, ELIZA was the first conversational system that simulated dialogues based on natural language [32]. The system proposed a conversation between the user and a psychologist, in which ELIZA was the psychologist and the user was the patient. Its success was determined by characteristics similar to human feelings. With the flow of conversation, the system could not maintain a coherent dialogue, but it was still a historic landmark for the study of conversational systems.

Currently, virtual assistants have become popular with the development of mobile devices such as smartphones and tablets. They consist of software that interacts with the user through audio and text interactions. Popular agents such as *Siri (Apple)*, *Google Assistant (Google)* and *Alexa (Amazon)*, support a variety of solutions, such as turning on lamps, closing curtains, turning on the television and other electronics in order to facilitate everyday tasks. They also offer other functionalities, such as providing weather information, opening applications, sending messages, among others.

In particular, such applications also offer content recommendations according to users' requirements, such as suggestions for restaurants near their location and tourist attractions. Even though agents have become increasingly efficient and useful for performing essential and non-essential tasks, the benefit generated by them cannot be restricted to the average audience only. Rather, there is

untapped potential for using these agents in the support needed for different groups of users with disabilities or difficulties, such as the elderly audience.

In this sense, this work aims to analyze a sample of the target audience and collect interaction and accessibility requirements to build an adaptable (by the user) and adaptive (by the system) conversational system that recommends content according to the user's interests. For this, studies were carried out with users using the Wizard of Oz technique[1], and the data were analyzed, resulting in a set of requirements to guide the development of conversational systems.

This paper is structured as follows: Section 2 presents the related works, Section 3 provides information on the preliminary case study conducted, Section 4 presents the methodology used in this work, in Section 5 the results obtained with the research are highlighted, Section 7 contains the discussion, and Section 8 contains the final considerations and future work.

2 Related Works

This work considered three distinct themes for the identification of related works. The first is related to the conversational interface that considers topics of accessibility and usability, the second considers conversational recommender systems and, finally, the third considers participatory design.

Ryu et al. [27] consider human psychological aspects in his work, and uses a chatbot to relate and contribute to the mental health of elderly people. In this sense, the research proposes a conversational system based on buttons to make interactions. Button-based systems make it easy for the chatbot to successfully identify what the user's interaction intentions are, but limit interactions as users may have other preferences related to the use of buttons, audio, text, or other forms of input. Considering that works´s Ryu et al. [27] is a conversational system for the elderly, the results obtained in the case study we did with the chatbot Bob, suggests more types of interactions, which were not attributed to Ryu et al. [27]. Because it is limited to this type of interaction, content recommendations are not entirely accessible either.

Valtolina and Hu [31], in turn, using strategies to improve the quality of life of the elderly, considers a chatbot that enables a connection between family members and doctors. It is also able to remind users of their appointments and medications, which for the elderly is promising. In addition, a proposal contributes to factors related to loneliness and contextualizes the experiences lived during a COVID-19 pandemic. Therefore, even though this is a preliminary work, it considers the current aspects of the elderly's daily life, but it was not considered that the elderly may need and/or prefer the means of interaction

Georgieva [12] raises questions about digital inclusion and how it affects the daily lives of the elderly population. This work considers a sample of first generation and second generation elderly people, to make a comparison between issues

[1] Term was invented by John Falk ("Jeff") Kelley, around 1980, during the development of his dissertation at Johns Hopkins University. More details: http://www.musicman.net/oz.html.

of digital inclusion, thus it is concluded that first generation elderly people believe that technologies fail when they need it most, besides also showing that others elderly people believe that technology makes things better. The author brings these data to a discussion involving the behavior of chatbots. Additionally, natural language processing would be an important factor for the tasks of a chatbot that supports the dynamics and behaviors of elderly users, as they can assist in complex tasks involving financial risk issues, such as banking services per application. Although natural language processing helps the elderly to communicate with the system, there is still a gap on how chatbots should behave and be built for this specific audience, as there are language, physical and cognitive limitations that are characteristic of the decline of human functions in relation to the natural aging process.

The work of Furini et al. [11] considers a problem still open in relation to accessibility, which refers to the profile of each user, in a way that a child interacts differently from an adult, who, in turn, interacts differently from elderly. With the results obtained in this work, accent personalization strategies and regionalism are shown to be a promising strategy regarding the user's experience in relation to their profile.

Pradhan, Mehta and Findlater [23], on the other hand, consider the approach through voice commands that serves many people with disabilities. Thus, people with reduced mobility can carry out tasks such as closing doors, turning on the TV, while the visually impaired also benefit from questions such as "How is the weather today?". So, with just one accessibility feature, the user can carry out everyday functions. Even though conversational agents have evolved to be more accessible, researchers are still trying to understand the behavior of average users in this type of system. The interactions of users with specific needs also need to be researched, so, considering conversational solutions, this work shows some results that address the need to go beyond the voice resources available on the market, or even its evolution in different contexts of user.

Yu, Shen and Jin [33] bring a conversational recommendation system, in which the user can evaluate content recommendations by visual resources. Grasch, Felfernig and Reinfrank [13], in turn, feature a conversational recommendation system, which supports voice interactions, but its recommendations are limited to textual mode. Although they are two systems that enable different types of interaction, they are similar in using hybrid dialogue strategies. Hybrid strategies enable greater accessibility in terms of interaction, so Ikemoto et al. [18] combined natural language models with structured elements based on forms, using visual resources for their responses to the user, although they did not propose voice or text interactions.

The domain of a conversational application is a widely discussed topic in the literature, and it should be task-oriented, considering precise and short interactions, in order to allow people to communicate more easily and find what they want with few messages [15]. From the consulted literature, it was evidenced that some applications use hybrid interaction approaches for content recommendation. However, they are still restricted to a few accessibility features. Therefore,

this paper explores conversational systems from another aspect, which is to include accessibility features in applications for users with multiple difficulties, such as the elderly.

In isolation, each system considers issues that evolve conversational systems, but considering the elderly audience and their needs, more information and resources are needed for the use of a conversational system.

The importance of participatory design is due to the objectives it seeks, which are: to give voice to "invisible" people in the context in which they are inserted; perform actual tasks in places familiar to the user in their real context; mutual learning with the participants, in search of an intersection; use of tools that effectively, in practical, concrete and specific situations, help the participant; alternative visions about technology that can generate equality in any environment, and; democratic practices learned in the form of equality practices or models [20]. However, Schlomann et al. [29] point out in their study that elderly people, especially with cognitive deficits, are neglected by Voice Assistant developers, thus showing the need for participatory design. The authors recommend these users' high level of involvement in research on trade assistants, in order to identify the best ways to introduce the use of these assistants, their benefits, limitations, and possible risks to the elderly.

The next section describes more aspects of the elderly audience and their interaction with technologies.

3 Elderly and Technology

A Brazilian Institute of Geography and Statistics (IBGE, in Portuguese) study [8] on the use of Internet, television and cell phones in Brazil made it possible to verify the increase in Internet use by 8 age groups. The study showed that two groups had the same growth rate (6.7%) and were the ones that increased the most between the years 2018 and 2019, one of them being the elderly population. This shows that the use of the Internet by the elderly population is growing, and the software that meets their needs must support the demand. In fact, the digital inclusion of the elderly is something recent, and has benefits such as reducing isolation, increasing autonomy and independence, as well as improving health [16]. Making content, such as videos, accessible is part of the task of digitally including elderly people [25, 26].

Youtube® is the most accessed video platform in the world [7]. Users access *Youtube*® in search of entertainment and information on various subjects. It is a platform for easy dissemination of opinion-forming content. Videos made available often have subtitles and other accessibility features available for use. To enable them, it is necessary to interact with the layout to identify what the buttons mean and how to customize their content in order to improve the user experience. In a few steps it is possible to configure the player of the video and watch the personalized content. However, the availability of resources does not mean that the user will be able to reach the goal of using them, because when it comes to users with low literacy, or elderly people who are not used to

technologies, they have difficulty accessing, identifying and activate the features [30].

Elderly people find it difficult to see, understand texts and symbols and remember instructions, which is an issue related to sensory, motor and cognitive [10] problems, inherent to age, which make it difficult to use smartphones. The ability to handle smartphones is also related to how accessible are the systems that comprise it. In fact, older people tend to write more slowly and can also make mistakes when using buttons by pressing more than one button at a time. Therefore, larger buttons can contribute to increase the chance of the user to interact according to their expectations [17]. Customizing buttons can be useful and makes the system accessible for some people, but when considering the elderly audience, the interface design needs to be done in a personalized way, for each individual [19], considering not only an interaction mechanism, but also assigning others that support the needs of each user.

Considering these limitations with the use of buttons, conversational systems exempt the user from dealing with customization only through buttons, which is the solution employed by *Youtube*®. It is possible to customize the use according to the needs of each individual, whether by text, voice, buttons, forms or some visual aid. This way, the interface design is adjusted in advance so that the content personalization interactions are done in a satisfactory way from the beginning of the conversation.

Based on the preliminary case study of the *Youtube*® platform reported in this paper, the goal of this search group is to develop a conversational video recommendation system, which uses different types of inputs and considers the customization of the parameters of the recommended video. Initially, the system aims to integrate with the *Youtube*® platform, but it can be adapted to any other streaming service that has accessibility features and a programming interface.

4 Preliminary Case Study

In this paper, the results of the first phase of the development of the accessible conversational system will be described. This first phase corresponds to the collection of requirements related to interaction and accessibility. For this, a study was carried out with users in order to understand what needs and particularities would be necessary for the system. In the next subsections, the study methodology is presented in detail.

4.1 Participants

The case study had eleven voluntary participants, five men and six women, aged between 62 and 73 years, attending a digital literacy course. Of these, two have visual difficulties and therefore need, in addition to eyeglasses and a device configured with large fonts, the aid of a magnifying glass. One of them was color

blind. Some participants reported having hearing problems, one of them using a hearing aid.

Participants were selected using a non-probabilistic sampling technique known as convenience sampling. In this type of sampling, participants are selected according to availability, location and interest in participating in the research [9]. For the selection of participants, the profile of the application's target audience and the group for which there was approval by the Research Ethics Committee, CAAE number: 57875016.3.0000.5390, was also taken into account.

4.2 Procedures and Methods

The study aimed to collect information that would allow collecting requirements for the construction of a chatbot capable of making recommendations for *YouTube*® videos to users. The procedure of the conducted study consisted of simulating a dialogue between the chatbot and the user, applying the Wizard of Oz technique [6]. In this technique, the user's interaction with the system is simulated, when, in fact, there is a person performing the activities that simulate the interaction. Sometimes, as in this case, acting in such a way that the user believes that the system is fully functional. That said, the participation of two researchers was necessary in the application of the test, one being responsible for observing the interaction between the user and the bot represented by the other researcher. It is noteworthy that at all times the participants believed they were interacting with a functional chatbot.

To carry out the study, the two applicators of the study and one participant at a time met, through the *Google Meet*® platform, through which the participant shared the screen of the device he was using, either smartphone or computer. For the simulation of interaction with the chatbot, called Bob, an account was used in the *Whatsapp*® messaging application, in the Web version, in which one of the applicators, simulating the bot, started the conversation with the user. The chatbot introduced itself saying: *"Hi! I'm Bob and I'm here to help you find YouTube videos you might like. What's your name? Tell me a little about yourself."*. After the participant's response, the following question was sent: *"Nice name! ¡participant's name¿, I know you like movies. What movie genre do you like best?"*. Then, the researcher that simulated the bot brought to the conversation a video related to the participant's response, and asked if the participant liked the recommendation. The next two questions were about the participant's favorite singer and travel destination. Similarly to the first question, the researcher representing Bob searched for videos in *Youtube*® referring to the participant's preferences, so that the participant believed that he/she was receiving personalized recommendations. A simulation with a user of this study is illustrated in Fig. 1 and it contains screenshots of the real interaction between the user and the simulator. Some details, such as name, have been hidden to ensure user privacy.

Based on the responses, the researcher responsible for the dialog would look for videos on *YouTube*® and send them to the user, who would reply whether he/she liked the recommendation made by the bot or not. After this step, the

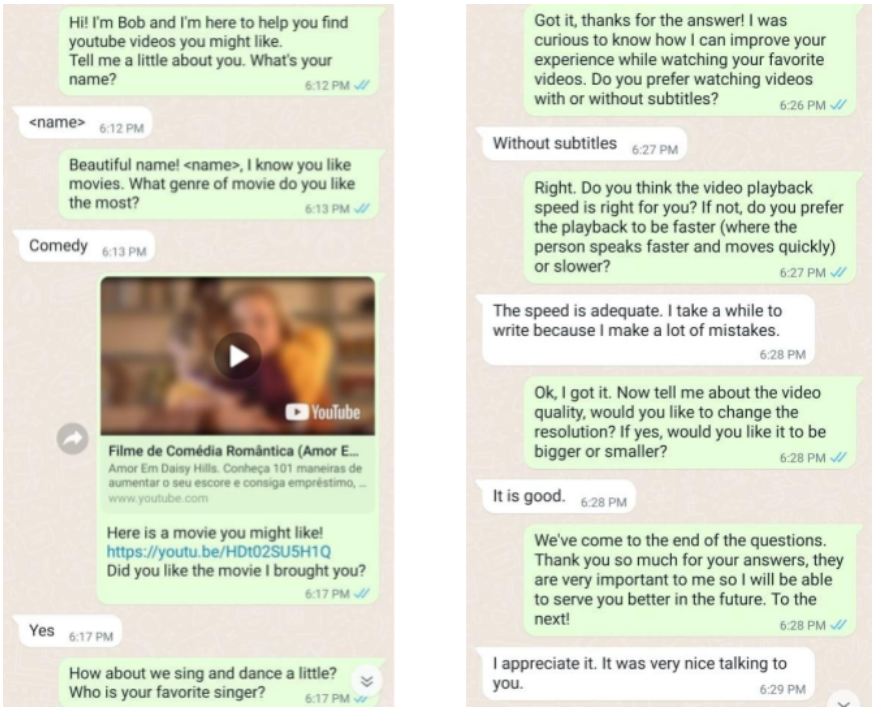

Fig. 1. Example of a snippet of dialogue between participant and Bob.

bot asked the user some questions, related to some accessibility issues of *Whatsapp*®, such as preference for chatting via text or audio, and *YouTube*®, such as the preference by the use of subtitles, playback speed and video resolution.

After the study was carried out by *Whatsapp*®, a unstructured interview was started with each participant, in order to reinforce some statements that they had made previously or so that they could go deeper into any answer that they might have answered very briefly, leaving applicants with doubts.

5 Results

This section presents the results of our preliminary case study. They are based on a qualitative assessment carried out from the analysis of the participants' discourse on the studies conducted, from which categories of meaning emerged that highlight important points about the use and interaction with Bob and are presented below.

5.1 Audio and/or Text Interaction

During the study, the participant was free to choose the way they wanted to interact with Bob (audio or text), but all of Bob's answers were sent by text.

One participant with vision problems, even though he chose to interact via text, commented: "*[...], even with the people I talk to, it's usually always via audio, because then it's quick. Typing takes a long time, I make a mistake, because of my vision problem.*". In line with this, another participant stated: *"I think it's perfect that he could answer by audio, it would look like we're doing here now."*. This participant interacted via text, but tried to interact via audio. He had difficulties with the *Whatsapp*® interface, and accidentally erased the audios without being able to send them.

Some participants suggested that the audio interaction was optional and allowed to be carried out by only one of the parties. One participant mentioned that *"It would be nice if Bob answered me with audio, but I don't like to send audio."* while another participant said: *"I don't think it's interesting for Bob to answer audio, I think more easy to send audio, but I prefer it answers with text."*. It was also suggested by some participants that the bot should adapt to the user's context or ask their preference for interaction. One of the attendees said: *"As for interacting via audio, I think it depends on context, if I'm 'cooking' I won't type. Likewise, if I'm 'on the notebook' I won' t send audio."* , which was reinforced by the speech of another participant who stated: *"It depends a lot, it depends on where the person is, whether he/she can hear at that moment. Maybe if it asked before sending it, it would be interesting."*.

However, some participants explained their reasons why they are not in favor of the appeal. Two participants commented on the need to read or listen to the statement more than once to be able to understand. The first one said: *"I think answering by audio wouldn't be very nice, because the voice will deliver that it's a robot, I prefer it just by text. I also read some questions three times, in the audio it wouldn't be much nice to keep repeating."*, which was in line with the opinion of another participant: *"I already had an experience with audio (with another application), I didn't clearly understand what it was asking, you know? At my age it is common to have hearing problems, so depending on the phoneme it can be bad to understand, until you understand it has already said some other things. So I think I might have the choice 'do you want audio or written?'* In particular, this case can be solved by customizing the way of interacting with Bob.

5.2 Subtitles, Playback Speed and Video Resolution

Participants were unanimous regarding the subtitles of the *YouTube*® videos. Everyone said they only use it when the video's audio is not in Portuguese, as one participant mentioned: *"I only use subtitles if I don't speak the language."*. Only one participant mentioned that he sometimes prefers to slow down the video playback speed a little, and that he does this when the pronunciation is fast or the speaker has an accent, and there are no subtitles available.

As for the resolution, everyone said it was good, but that it could start with a better resolution whenever available and compatible with the user's Internet speed.

5.3 Use of Emojis and Stickers

The use of emojis and stickers was associated with reinforcing the message or feeling the user wants to convey, as one participant mentioned: *"I love talking using stickers and gifs, I laugh. I would love for it to talk to me like that."*. In this study, Bob only responded with text, without emojis.

However, a negative point was noted regarding the lack of consensus for the interpretation of emojis. *"I use a lot of emoji and stickers, I think it's really cool, but I don't think it's interesting for Bob to answer me like that, because like... I know what I mean, but I don't know what the other person means."* . When asked if he used emojis when chatting through *Whatsapp*, one participant said: *"Sometimes there are things I don't use because I don't know what it means, it can lead to misinterpretations."*, showing the need for *chatbot* customization options.

5.4 Friend or Virtual Assistant

Among the elderly participants, the idea of a virtual friend was cited a lot, as in *"Oh, I'd rather talk to Bob than stay there* [on YouTube®]... *Because I'm going to talk to bot, it'll answer me ... Even if it's virtual, even if I can't see it, answering what I ask is already something very important. You have no idea at this moment (of the pandemic) what it's like."*, and in *"Bob is working with a very wide range of options it seems, that's why it looks like a friend, because it seems to talk about everything. Bob talks about movies, singing, traveling and such, it seems like a friend who can talk about any subject."*.

5.5 Use of Buttons

Three participants showed interest in using the buttons with quick answers, for example, "yes", "no", "like it" and "did not like it", as we can see in: *" I think it would be interesting to use buttons, the more it facilitates interaction, I think it's interesting."*, and in *"I would definitely use buttons, you look at the options, identify what you want, then it goes faster, sometimes it even expresses better because sometimes we want to say something and can't express it, right?"*.

5.6 Information Verification

Currently, a lot of information reaches users through messaging apps and social networks. There is, therefore, a difficulty for the elderly public to determine if the information is true or false, especially if it comes from a source they believe to be reliable. In this sense, there is also the problem of authenticity verification, for example, when purchasing a product or service. One participant reports: *"Another thing would be for Bob to help me know what is true, for example, now I want to take a course but I don't know who will teach, Bob could help me know if he is trustworthy."*.

5.7 Return on Response Time

In cases where Bob was slow to respond, there was a certain feeling of frustration in the user. One of the participants suggested: *"Would you be able to put a 'Wait' message, so that it informs you that it is processing the information? The person might wonder what happened due to the bot's response time."*. This issue can be resolved with a message to the user that Bob is still verifying the order and asking the user to wait.

5.8 Accents and Regionalisms

Conducting the study online allowed for participants from all regions of Brazil. Thus, it could be noted that the participant could feel more comfortable in the conversation if Bob used regional terms, as well as being able to understand these terms when used by the user. This can be seen in the speeches of the participants, as in *"Bob writes very well, but these (regional) expressions would make Bob more natural."*, and in *"Bob needs to understand accents (in this case of voice interaction) and regionalisms, it has to know who it is 'speaking with. From Pará, to Rio Grande do Sul, to the northeast, it changes a lot."*. To resolve this issue, a solution will be developed based on informal text mining approaches and heterogeneous data classification (considering the informal text and the geographic information where the text was written), so the sentences can be validated by the classifier and later, used in synthetic voice tools. Thus, each user could instantiate a textual set according to the linguistic variations of their region.

5.9 Classification of Requirements

The study allowed classifying the aforementioned categories into functional and non-functional requirements of the proposed conversational system. For classification and understanding of non-functional requirements, the set of software quality metrics ISO/IEC [22] was used. Tables 1 and 2 illustrate the requirements collected with the study.

Table 1. Functional Requirements.

Functional Requirements	**Description**
Using buttons and forms	The system must provide use of buttons and forms
Audio interaction	The system must offer voice interactions
Text interaction	The system must offer text interactions
Use of emojis and stickers	The system must offer interactions by emojis and stickers
Accents and Regionalism	The system must be customized for each person with regard to accents and local slang

Table 2. Non-Functional Requirements.

Non-Functional Requirements	**Description**
Usability (button size)	The system must contain larger buttons compared to the popular patterns [5]
Adaptability (customizable contents)	The system must contain custom interface and configurable
Performance (Response Time)	The system should reply the user with a time less than 5 seconds [2]
Performance against memory	The system must provide the conversation history whenever requested, without prior storage on the user's device
Maintainability	The system must be implemented by modules [14]
Interoperability	The system must work on the most used platforms by the target audience
Usability	The system must be self-explanatory [28]
Availability	The system must persist the modules independently
Accessibility	The system must be responsive
Accessibility	The system must be adaptive (by user) and adaptive (by system)

To deal with non-functional requirements, software quality metrics that referred to the results obtained during the user testing phase were considered. Thus, for system usability, two requirements were raised: to have a self-explanatory system, allowing users to understand how to use the system and if the user is confused, the system must clearly guide him/her on how to proceed; and consider button sizes, as it is one of the factors found in the case study that will impact not only functionality but also implementation issues.

Regarding adaptability, as this paper presents questions of user preference and/or needs, the system must have a customizable and configurable interface, that is, for each user with different preferences, the system must support the demands related to interaction factors.

Considering the performance of conversational systems for the elderly, two factors were found: the first with regard to memory, which should not store user content and information on the device the user is using; another regarding the response time, something that was also found through the case study carried out in this research.

As the functionalities are diverse, the systems must be implemented by modules and this is a maintainability requirement, since in case of failure of some interaction functionality, developers will be able to deal with it without impeding other functionalities, thus, it also stands out as a requirement availability, since the user will be able to take advantage of others and will not be frustrated by not being able to reach their final objective.

This system should work on conversational platforms that are most used in the world, taking advantage of the fact that users already use it to talk to other people and facilitating access without having to install new applications, which is a requirement for system interoperability.

Regarding accessibility, the system must be configurable to support the elderly audience and their interaction preferences, in addition to being responsive.

To classify the degree of relevance of each requirement, the definitions high, medium and low [21] were used. Regarding the functional requirements, all are of high relevance, except the use of 'emojis and stickers', 'accents and regionalism', which were classified as low relevance. Non-functional requirements were rated as high relevance, except 'performance against memory' which was rated as medium relevance.

6 Pilot Evaluation

Some features were implemented in the Bob chatbot and then a 1-user test was performed. The pilot test was intended to find possible coding problems and new proposals, in addition to considering a restaurant recommendation domain. The participant's profile is configured as non-elderly and their specialties are related to psychology. The evaluation was carried out in a controlled environment and had the participation of two applicators, one of them was in person mode, accompanying the user and coordinating the test, the other was in remote mode, following the test and also observing the system logs. After the test, the user answered the SAM (Self-Assessment Manikin) questionnaire, an instrument that uses images for personal affective assessment in the domains of pleasure, arousal and dominance, after exposure to a stimulus [4], in this case, referring to the use of the chatbot. The results allowed verifying problems in the implementation, such as: increasing the flow of conversations, giving users more options for paths, recommending more restaurants, especially if the user doesn't like the recommendation, and providing filtering options by types of food.

New tests are been carried out with a larger sample of users and also with new application domains, such as recommendation of establishments and news.

7 Discussion

The analysis of the requirements presented in Section 5.9 allowed the research team to verify important aspects in the construction of the chatbot. One of the points observed was the existence of greater agreement on preferences or needs for interaction and accessibility among the elderly. For this reason, the points discussed take into account this population and their suggestions.

It was observed that users would like there to be voice interaction by at least one of the parties, but it was a consensus that the audios sent by Bob should be short. However, in the participants' speeches, it was noted that the user's context needs to be analyzed, either with Bob asking, at the beginning

of the conversation, which type of interaction is preferred, either allowing personalization with a kind of user profile, or doing this adaptation automatically. Customizing according to user preferences and needs can also include Bob's use of emojis and stickers, as not all users would like to use them.

The interest of elderly users in a virtual friend was also noted. The statements of some elderly participants showed that they sometimes experience moments of solitude. Having a companionship, even in a virtual form, that can maintain a dialogue, or even help you with some tasks, such as searching online. Another functionality in this case would be for Bob to make a call or send a message via *Whatsapp*® itself to an emergency user contact, simply by means of a voice command or by touching a button.

The idea of having Bob as a facilitator also extends to the verification of news, which usually arrives by *Whatsapp*® and in large volumes, that the elderly, in particular, cannot or do not know whether it came from a reliable source and if it is true. Thus, Bob would check the news and certify the users, preventing them from propagating fake news.

After conducting the studies and analyzing the results, it was possible to conclude that several features can be added to Bob, making it more than a conversational recommendation system. Customizing Bob to meet the user's needs proves to be a necessary viable solution, allowing the user to manually configure the system preferences (adaptable), or this custom configuration to be done automatically by the system (adaptative).

Many issues brought up in this discussion are directly related to Natural Language Processing techniques, another idea that emerged during the study, is to extract behavioral patterns and identify possible depressive users through these techniques. Thus, Bob could check if the users are having depressive episodes and recommend content that helps them in this regard.

8 Final Remarks

This work presented a study carried out to raise interaction and accessibility requirements for the construction of a conversational recommendation system aimed at the elderly audience.

With the studies carried out, it was possible to verify the preferences and needs for interaction and accessibility for the elderly. The use of resources, such as buttons with quick responses and voice interaction, proved to be important, but especially when the user is allowed to use it according to their will. Therefore, it is interesting to build a chatbot with an adaptive and adaptable interface, to improve user interaction with the system interface, as well as use the accessibility features according to their needs.

At first there were expectations regarding user interactions and what would be possible to collect to understand and build an adaptive and adaptable conversational system, but only after the study was it possible to sustain the ideas and obtain others that are significant in the user experience. A future work is to expand the language training of the conversational system, from the point

of view of natural language processing and also of regional aspects, considering that users are more attracted by the accents and customs of their regions.

Another important future work will be the recommendation of videos through a recommendation system that filters out fake content. Note that the elderly audience is more likely to believe in untrue content, unconsciously disseminating this information on social networks. Thus, the support of a fake news verification system will be of great relevance.

Adapting the system to different domains can also be a future work, so, in addition to recommending personalized content, users will be able to interact with the bot for other entertainment activities, such as sustaining a dialogue to talk about everyday life.

It is also intended to carry out a new stage of studies using Whatsapp, with 36 volunteers, arranged in 3 groups. One group is interacting with the conversational agent by voice, and the second group is doing the interactions by text. The third group could use both types of interactions, by audio or text. The choice of which users would go to a given group was made randomly, except for users who had some disability that made it impossible to interact with a given group, so it was transferred to another. At the end of each test, two researchers are interviewing the volunteer, and they are answering a questionnaire. It has a similar methodology to the first, however, with Bob working and being adaptable, dispensing with the role of the Wizard of Oz. At this stage, we want to check if the accessibility issues implemented in Bob are practical and easy to configure.

Finally, regarding the current state of the area of accessible conversational recommendation systems and, considering the lack of conversational corpora for applications, an innovative contribution of this work is the construction of an annotated *corpus* that can be used by others researchers, since this was not found in the literature.

9 Acknowledgments

The authors would like to thank the financial support from Brazilian agencies: CNPq, CAPES and FAPESP. We would also like to thank the participants in the preliminary case study and the pilot evaluation.

References

1. Abo Khamis, M., Ngo, H.Q., Rudra, A.: Faq: questions asked frequently. In: Proceedings of the 35th ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems. pp. 13–28 (2016)
2. Ahmadvand, A., Choi, I., Sahijwani, H., Schmidt, J., Sun, M., Volokhin, S., Wang, Z., Agichtein, E.: Emory irisbot: An open-domain conversational bot for personalized information access. Alexa Prize Proceedings (2018)
3. Akatsu, H., Miki, H.: Usability research for the elderly people. Oki Technical Review (Special Issue on Human Friendly Technologies) **71**(3), 54–57 (2004)

4. Bradley, M.M., Lang, P.J.: Measuring emotion: the self-assessment manikin and the semantic differential. Journal of behavior therapy and experimental psychiatry **25**(1), 49–59 (1994)
5. Cunha, B.C., Rodrigues, K.R., Pimentel, M.d.G.C.: Synthesizing guidelines for facilitating elderly-smartphone interaction. In: Proceedings of the 25th Brazillian Symposium on Multimedia and the Web. pp. 37–44 (2019)
6. Dahlbäck, N., Jönsson, A., Ahrenberg, L.: Wizard of oz studies: Why and how. In: Proceedings of the 1st International Conference on Intelligent User Interfaces. pp. 193–200. IUI '93, ACM, New York, NY, USA (1993). https://doi.org/10.1145/169891.169968, http://doi.acm.org/10.1145/169891.169968
7. Demir, M., Taken, K., Eryilmaz, R., Aslan, R., Ertas, K.: Youtube as a health information source: Covid-19 and andrology. Medical Bulletin of Haseki/Haseki Tip Bulteni **59**(2) (2021)
8. Educa, I.: Uso de internet, televisão e celular no brasil (2019)
9. Etikan, I., Musa, S.A., Alkassim, R.S.: Comparison of convenience sampling and purposive sampling. American journal of theoretical and applied statistics **5**(1), 1–4 (2016)
10. Fisk, A.D., Czaja, S.J., Rogers, W.A., Charness, N., Sharit, J.: Designing for older adults: Principles and creative human factors approaches. CRC press (2020)
11. Furini, M., Mirri, S., Montangero, M., Prandi, C.: Do conversational interfaces kill web accessibility? In: 2020 IEEE 17th Annual Consumer Communications & Networking Conference (CCNC). pp. 1–6. IEEE (2020)
12. Georgieva, L.: Digital inclusion and the elderly: The case of online banking. In: Proceedings of the LREC 2018 Workshop "Improving Social Inclusion using NLP: Tools, Methods and Resources"(ISI-NLP 2). vol. 7, pp. 8–12 (2018)
13. Grasch, P., Felfernig, A., Reinfrank, F.: Recomment: Towards critiquing-based recommendation with speech interaction. In: Proceedings of the 7th ACM Conference on Recommender Systems. pp. 157–164 (2013)
14. Hayes, J.H., Zhao, L.: Maintainability prediction: A regression analysis of measures of evolving systems. In: 21st IEEE International Conference on Software Maintenance (ICSM'05). pp. 601–604. IEEE (2005)
15. Hill, J., Ford, W.R., Farreras, I.G.: Real conversations with artificial intelligence: A comparison between human–human online conversations and human–chatbot conversations. Computers in human behavior **49**, 245–250 (2015)
16. Holgersson, J., Söderström, E., Rose, J.: Digital inclusion for elderly citizens for a sustainable society (2019)
17. Iancu, I., Iancu, B.: Designing mobile technology for elderly. a theoretical overview. Technological Forecasting and Social Change **155**, 119977 (2020)
18. Ikemoto, Y., Asawavetvutt, V., Kuwabara, K., Huang, H.H.: Tuning a conversation strategy for interactive recommendations in a chatbot setting. Journal of Information and Telecommunication **3**(2), 180–195 (2019)
19. Lim, C.K.: The implementation of 3d printing in customized interactive design for elderly welfare technology. In: International Conference on Human-Computer Interaction. pp. 299–303. Springer (2014)
20. Luck, R.: What is it that makes participation in design participatory design? Design Studies **59**, 1–8 (2018)
21. Marçal, E.K., et al.: Auditoria da qualidade de um software de contabilidade. Gestão & Regionalidade **23**(66) (2007)

22. Morais, R., Costa, A.L., Góes, W.M., Somera, S.C.: Aplicação de métricas de software em um modelo de avaliação de qualidade para sistemas de informação de saúde: um estudo de factibilidade em um hospital público universitário (2013)
23. Pradhan, A., Mehta, K., Findlater, L.: " accessibility came by accident" use of voice-controlled intelligent personal assistants by people with disabilities. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. pp. 1–13 (2018)
24. Rocha, J.A.P., Duarte, A.B.S.: Diretrizes de acessibilidade web: um estudo comparativo entre as wcag 2.0 e o e-mag 3.0. Inclusão Social **5**(2) (2012)
25. Rodrigues, K., Zaine, I., Orlandi, B., da Graça Pimentel, M.: Ensinando configurações do smartphone e aplicações sociais para o público 60+ por meio de aulas semanais e intervenções remotas. In: Anais do XII Workshop sobre Aspectos da Interação Humano-Computador para a Web Social. pp. 25–32. SBC, Porto Alegre, RS, Brasil (2021). https://doi.org/10.5753/waihcws.2021.17541, https://sol.sbc.org.br/index.php/waihcws/article/view/17541
26. Rodrigues, K.R., Onuki, L.A., Assunção, D.M., Junior, S.G., Pimentel, M.G.: Possibilities for the digital literacy of the older people in times of social distancing. In: Proceedings of the 19th Brazilian Symposium on Human Factors in Computing Systems. pp. 1–6 (2020)
27. Ryu, H., Kim, S., Kim, D., Han, S., Lee, K., Kang, Y.: Simple and steady interactions win the healthy mentality: Designing a chatbot service for the elderly. Proceedings of the ACM on Human-Computer Interaction **4**(CSCW2), 1–25 (2020)
28. dos Santos, V.C.: Agentes inteligentes na educação a distância: uso de sistemas tutores inteligentes como auxiliares no estabelecimento da comunicação dialógica. LínguaTec **3**(2) (2018)
29. Schlomann, A., Wahl, H.W., Zentel, P., Heyl, V., Knapp, L., Opfermann, C., Krämer, T., Rietz, C.: Potential and pitfalls of digital voice assistants in older adults with and without intellectual disabilities: relevance of participatory design elements and ecologically valid field studies. Frontiers in Psychology **12** (2021)
30. Silveira, M.M.d., Portuguez, M.W., Pasqualotti, A., Colussi, E.L., et al.: Envelhecimento e inclusão digital: Significado, sentimentos e conflitos. Geriatrics, Gerontology and Aging **8**(3), 178–184 (2014)
31. Valtolina, S., Hu, L.: Charlie: A chatbot to improve the elderly quality of life and to make them more active to fight their sense of loneliness. In: CHItaly 2021: 14th Biannual Conference of the Italian SIGCHI Chapter. pp. 1–5 (2021)
32. Weizenbaum, J.: Eliza—a computer program for the study of natural language communication between man and machine. Commun. ACM **9**(1), 36–45 (Jan 1966). https://doi.org/10.1145/365153.365168, https://doi.org/10.1145/365153.365168
33. Yu, T., Shen, Y., Jin, H.: A visual dialog augmented interactive recommender system. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. pp. 157–165 (2019)