SAATHI: An Urdu Virtual Assistant for Elderly Aging in Place

Anand Kumar[✉], Ghani Haider[✉], Maheen Khan[✉], Rida Zahid Khan[✉], and Syeda Saleha Raza[✉]

Dhanani School of Science and Engineering, Habib University, Karachi, Pakistan
{ak05173,gh05177,mk04389,rk04364}@st.habib.edu.pk,
saleha.raza@sse.habib.edu.pk

Abstract. With the rise of the digital age, life has become a lot easier for the vast majority of the population. However, the ever-increasing elderly population has suffered, especially in countries like Pakistan, where limited accessibility to technology, often due to language barriers, hinders elderly from reaping technological benefits. In this paper, an Urdu virtual assistant application is proposed which provides an intuitive and empathetic platform for the elderly in Pakistan that helps them perform essential tasks such as reminding them of their medications, organising their work, getting daily news highlights, and connecting them with their loved ones. It also provides entertainment in the form of user-specified video playlists or by positively engaging them in conversations on various topics.

Keywords: Elderly care · Virtual assistant · Conversational agent · Urdu natural language understanding (NLU)

1 Introduction

According to the World Health Organization (WHO), life expectancy is increasing due to advancement in science, strong economies, and healthy behaviours [10]. Due to this increased life expectancy, most people are expected to live beyond their sixties, allowing them to contribute to society as mentors, innovators, and in a variety of other ways. The latest census of Pakistan states that the current population of people over the age of 60 (old-age) is estimated to be 5.54% [2] and this number is set to increase. However, due to the increasing normalisation of a nuclear family structure [20] in Pakistani society and the hectic lifestyles of modern-day individuals, oftentimes there are not enough people to take care of older family members. As a result, the elderly might end up isolated and confined to their homes. This can negatively impact their emotional and physical well-being. Consequently, there is a growing demand for technology that supports the elderly at home [11] and can assist them in their daily tasks helping them stay connected to their loved ones.

To address the aforementioned problems, a solution called "Saathi" is proposed. Saathi intends to offer the elderly a supportive environment by allowing them to use an Urdu mobile virtual assistant to perform everyday tasks such as setting medicine and appointment reminders, calling their loved ones, navigating through different apps in their phones using visual icons or voice commands in Urdu to improve their everyday life. In addition to that, it provides them with entertainment, prompts them to take their meals on time, and generates empathetic conversation or actions accordingly. While an abundance of such applications exists in languages like English, there are currently no assistive applications for the elderly that support the Urdu language. Saathi acts as a bridge that allows the elderly to use their smartphone by making it more accessible through a voice-enabled Urdu virtual assistant as well as a consolidated and user-friendly user interface.

The rest of the paper is organized as follows. In Sect. 2, the existing work is reviewed in the domain of Socially Assistive Robots (SAR) and virtual assistants (VA) for elderly. Section 3 explains the proposed solution which includes the overall architecture and features of the application. Section 4 presents the application prototype and the results related to testing performed on the conversational agent. Section 5 discusses the obtained results and mentions the future scope of the project and Sect. 6 concludes the paper.

2   Related Work

Much research and development has been done in the domain of Socially Assistive Robotics (SAR) to develop companion robots to assist older people in a variety of tasks such as monitoring and promoting physical health in the elderly at home [3,11], helping them by autonomously navigating the user's home, reminding them of their medications, and providing entertainment [1]. Other than that, technologies such as Internet of Things (IoT) and wearable technologies could offer promising solutions for elderly care. Wearable devices are a part of IoT systems which can be worn by the people for monitoring the physical activities and physiological data [19]. These devices are embedded with sensors and algorithms that allow them to track, analyse and guide users' behaviour [17], vital signs or movement. Wearable tech has the potential to assist with several scenarios in elderly healthcare and can improve the quality of life for the elderly and allow them to maintain their independent lifestyle [19]. However, one of the major issues with SAR systems and wearable technology is the lack of social component and interaction, which hinders their acceptance among older people. Another problem with these assistive robots is their significant cost, which is not within everyone's budget, especially in low income countries like Pakistan.

Like other technological devices, virtual assistants have also become an element of great help to integrate and keep people active, in addition to facilitating the different everyday tasks that need to be performed [13]. A virtual assistant is an application that recognises voice instructions and does tasks on the user's behalf.

There has been significant work done to develop user-centered virtual assistants that feature different functionalities to assist elderly people. Researchers have found that spoken language seems to be the most preferred mode of interaction for elderly people, they prefer to interact with virtual assistants in a simple, hassle free way [18]. Beskow et al. [6] designed a user-centric virtual assistant that could be used to set medicine reminders after surgery and had a virtual character with whom users could communicate and set reminders, among other things. Lesin [16] developed a virtual assistant solely for mobile platforms, which aimed to assist diabetic patients with their daily routines. In addition to providing reminders, it also possessed the option to reach out to doctors/medical staff. Bickmore et al. [7] developed a virtual assistant to foster fitness and exercises in elderly people. Mival et al. [12] developed a virtual assistant that resembled a dog so that people may have virtual pets and communicate with them. Klein [15] created a virtual assistant that could aid a user in managing stress and anxiety by offering a variety of solutions for reducing stress and anxiety. Kasap et al. [14] created a virtual assistant that could change its facial expressions depending on how the user was feeling. Yasuda [21] created a virtual assistant with a human-like aspect to encourage user attention in a purely conversational medium.

The aforementioned works are in English and hence would not be helpful in the local context of Pakistan as majority of the population does not converse in English. Little work has been done in the field of virtual assistants in Urdu. In 2019, C-Square collaborated with Genesys to launch Pakistan's first AI enabled Urdu voice recognition bot named RUBA (Real Urdu Bot Automation) [4] at the Smart CX conference. RUBA stands for Real Urdu Bot Automation and is a Siri-styled virtual voice-enabled personal assistant. The bot can converse with the user in Urdu and react in Urdu while performing simple activities like as checking the amount of the user's bank account, sending text messages, and so on [4]. However, the domain of this app is very limited and does not focus on solving problems specific to any particular group in the Pakistani society.

There are a few other mobile applications available such as Carer, Elderly Care and BaldPhone which are designed for elderly people, however all of these apps are solely in English and not voice activated hence they would be of little use to the majority of Pakistani elderly individuals who do not communicate in English.

3   Proposed Solution

Saathi is an Urdu virtual companion application for elderly people in Pakistan. It facilitates them in their daily tasks and provides them with a supportive environment to keep them engaged and entertained through technology. Saathi is built using Flutter for the frontend, Firebase for the backend and the conversational agent has been designed using the RASA Framework. Figure 1 represents a high-level architecture of Saathi.
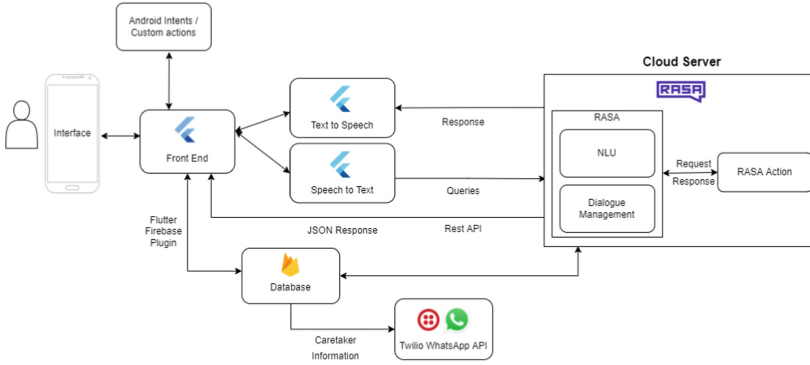
Fig. 1. System architecture diagram

3.1   Conversational Agent

The Urdu conversational agent is based on the RASA framework. RASA is a Python based, machine learning framework for conversational AI [8]. Since RASA is an open-source conversational AI framework, its code is readily available to use for free. It is an accessible, flexible, robust, and transparent framework. A benefit that RASA provides is that it follows a modular, extensible, micro-services architecture that fits well in any typical software development scenario. It is easy to integrate and customize to fit the needs of our application. Most importantly, RASA allows to develop conversational agents that leverage the power of NLP to determine user intents and mimic human-like conversation, as compared to chatbots or conversational agents with hardcoded logic and no capacity of learning.

Saathi uses RASA to engage the elderly in chit-chat about subjects that are most common among Pakistan's senior citizens. These topics of interest could include food-related discussions, asking them to describe their past experiences, or conversing with them about their favourite activities. Along with daily conversations, the conversational agent can be used to respond to queries about the weather, news, and time among other things.

RASA framework breaks down the processing of user queries and returning of responses into two main components, namely, Natural Language Understanding (NLU) and Core Dialogue Management [8]. Figure 2 shows the complete data-pipeline for the conversational agent module.

Natural Language Understanding (NLU) Component: The NLU component of the RASA framework is further divided into two components:

1. **Natural Language Processing (NLP) Component:** The raw Urdu text is first passed through the Urdu Natural Language Processing pipeline built using the SpaCy library. The Urdu NLP pipeline consists of a tokenizer, a parts-of-speech tagger, a parser and a named entity recognition module.
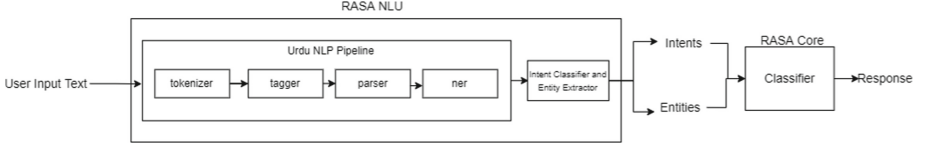
Fig. 2. Data pipeline for RASA framework

The tokenizer breaks the input sentence into smaller pieces called tokens. These tokens are tagged and categorized in correspondence with a particular Urdu part of speech by the tagger. The parser then extracts the syntactic structure of the input text by analyzing the words based on the grammar of the language. Lastly, the named entity recognition module identifies and classifies the named entities in the text. These named entities refer to the proper nouns found in the text.

2. **Intent Classifier and Entity Extractor:** The processed query from the NLP component is passed on to the intent classifier and the entity extractor module. Entities in RASA are structured pieces of information inside a user message [8]. The module maps the user query to a pre-defined intent and extracts important entities from it. RASA uses DIET (Dual Intent and Entity Transformer) architecture as part of its NLU component. DIET is a multi-task transformer architecture that handles both intent classification and entity recognition together [9]. The DIET architecture comes with the ability to plug-and-play various pre-trained embeddings and support for custom components and pipelines to use any other ML model.

Core Dialogue Management Component: The extracted intents and entities from the NLU component are then passed on to the second component of the framework called RASA core dialogue management. RASA core uses a classifier as a response selector. The response selector finds and outputs the best response to the user input query. In the case where the RASA module cannot map a query to a pre-defined intent, a fallback policy is activated by the RASA module which sends a fallback response to user and asks the user to rephrase their query.

Custom Actions: Sometimes, the queries by users might require the execution of custom code to process the response required by the user. Such queries are handled by the custom actions module in RASA. RASA's custom actions module allows running custom code which can be used to perform database queries, make API calls, etc. Saathi uses custom actions to add events to reminders, open mobile applications, dial calls, send messages, etc. Following are a few custom actions that the user can perform using Saathi conversational agent:

1. **Feeling connected:** Saathi allows the elderly to open any application that might be installed in their smartphones through Urdu voice commands. It

also helps them to dial a call or send a message to their loved ones by using either their name or phone number. These actions are performed by using an android intents plugin supported by Flutter. Such features help keep the elderly feel connected and give them a sense of independence.

2. **Reminders:** While the elderly are physically capable of performing most everyday tasks, they frequently forget to do so. Therefore Saathi can remind them of their medicines, meals, and any events they choose to be reminded of. They can do this through voice commands which triggers a custom action that takes the date, time and reminder description as entities and adds a reminder to the user's schedule. Other than user-defined reminders, there are reminders like reminding the elderly to charge their phone when the battery is running out.

3. **Entertainment:** According to research, the elderly's entertainment needs are equally vital for their well-being and joyous life [5]. Saathi several features for providing daily entertainment to the elderly. At the time of sign up, they can specify a playlist they prefer to listen to and whenever they want to play it, they can just ask the conversational agent to do so. Moreover, they can also get the latest news, ask about the weather of a particular city or inquire current time.

3.2   Empathetic Responses

It is important to emphasize that all intents and responses in the conversational agent have been pre-defined in Urdu. This means that the scope of the application has been set up during the development phase. The responses and interactions are tailored in such a way that they are as empathetic and human-like as possible. Firstly there are intents, which are used to categorize user messages. For example, for an intent called "greet" all the different kinds of greetings in Urdu are defined. Whereas, responses are what the chatbot sends to the user. After defining all the anticipated messages from the user and the replies from the chatbot, multiple intents and responses are tied together to form stories, which act as templates for possible conversations. These templates are provided as training data for the chatbot's dialogue management model [8] which in turn helps generalize the model to deal with unseen conversation paths. In Fig. 3, the story shows a happy path where the user might say hello and the chatbot greets them in return. It then asks them about their mood and depending on the mood, different story paths are chosen. In this particular example, the user's mood is great and so the chatbot replies with an appropriate response.

The RASA conversational agent requires datasets in the form of intents and responses. For our application, we have defined 28 intents in total for now, covering a range of queries, concerns and actions that the elderly might require. However, this set of intents can further be extended depending on the future needs of our application. Each of these intents have 10–15 Urdu examples on which our model has been trained. These example sentences have been crowd-sourced from native Urdu speakers to train the model to best identify each requisite intent. Each intent is associated with a response or custom action that

(a) nlu.yml

(b) domain.yml

(c) stories.yml

Fig. 3. RASA file format

is sent as output to a triggered intent to help the elderly. RASA allows making the conversational agent's replies more interesting if multiple response variations are provided to choose from for a given response name. Where possible, we have added variability in responses by providing multiple possible response utterances mapping to specific intents. RASA then selects one of the provided utterances at random and sends it to the user.

In contrast to question-answer bots, Saathi is categorized as a companion or a conversational agent. Therefore, special care is taken in appropriately responding to different scenarios. One such example is when the elderly user mentions that they are feeling sad or anxious, an empathetic response is sent, which asks them if anything can be done to help improve their mood. In some cases, a YouTube video of narration of an Urdu short story is also sent to cheer them up.

3.3 User Interface

To use Saathi, the user first needs to sign up for the application and fill in the required information such as their email, password, medicine timings, and preferences. These preferences can be related to their meal, medicine or appointment reminders, specifying a Youtube playlist they like to watch, etc. The registration process also requires the user to provide information of at least one caretaker, so that they can be alerted in cases of emergency such as when the elderly might be feeling unwell. After signing up, the elderly can interact with the application and use features such as the visual assistant, which allows them to navigate through their phones easily, the scheduler, where they can see their reminders and events, and the virtual assistant which can be used by their voice in the Urdu language.

To use the virtual assistant, the user can give their queries in the form of speech or voice commands in Urdu. The Urdu voice data is first converted to text using the speech to text feature of the underlying mobile platform and then passed on to the conversational agent module in the form of a text query. On receiving the text response from the conversational agent, it is converted to voice using the text to speech feature and sent to the user.

Saathi relieves the caretakers - people in charge of looking after the elderly - the burden of constantly being worried about the well-being of their elderly loved ones. This is done by keeping them up to date on the elderly person's medicine

and meal intake, the status of their appointments and other such events. To update the caretaker, a daily log of above-mentioned activities is maintained and shared with them on WhatsApp using the Twilio API.

4 Prototype and Results

This section presents the application prototype to validate the proposed solution and the results obtained while testing the conversational agent.

4.1 Application Prototype

The mobile application is installed on the elderly's smartphone. It is used to communicate with the conversational agent which is deployed on the cloud. Through the application, the elderly can navigate to different screens and choose the activities that they are interested in performing. The application allows them to do the following tasks:

1. **Registration and Login Screens:** Used to register their information, preferences and caretaker information as shown in Fig. 4.
2. **Visual Screen:** A screen which displays all the features consolidated in a single place for easy navigation as shown in Fig. 5a.
3. **Conversational Agent Screen:** Used to chat with the conversational agent in order to perform various tasks or engage in chit chat as shown in Fig. 5b.
4. **Schedule Screen:** Used to set, update and delete reminders related to meal, medicine, medical appointments and other such events as shown in Fig. 5c.

4.2 Prototype Validation

In order to test the application, unit and integration tests were written for different components and modules of the application. Usability testing was also performed to evaluate the accessibility and ease of use of the application from the user's perspective. The testing was carried out on 10 users over the age of 60. One of the feedbacks received from the users was that they faced difficulty in navigating to different screens of the application. Another issue pointed out was the limitation in the type of conversations with the conversational agent.

After getting the feedback, the user interface was made more intuitive and easy to use for the elderly by increasing the font, icon and image sizes. To provide ease in navigation, we provided a visual screen where all the basic features are consolidated on a single screen. For the conversational agent, more data was added around different topics such as checking elderly mood, discussing their interests, likes and dislike to keep the elderly engaged.

(a) Step 1   (b) Step 2   (c) Step 3

Fig. 4. Registration screens

4.3 Results for Conversational Agent

Since the conversational agent will be directly interacting with elderly people it must be as accurate and well-performing as possible. To do this, the model was tested through the RASA test module. The cross-validation method was used to test the chatbot, where the dataset is split into k number of groups, each of equal size, in this case, $k = 5$. One of these k splits is then chosen for testing and the rest of the splits are used for training the chatbot. This process is repeated k number of times until each split is used for testing. Finally, a weighted average of all k iterations is taken to validate the model's performance. Figure 6 shows the confusion matrix of intents with the cross-validation approach. It shows that even though the diagonal values are high, there are several misclassifications of the intents. The ratio of misclassifications to correct classifications is 0.247. While the model performs well on a majority of the intents even with cross-validation, there is still room for improvement. Due to these misclassifications, a threshold of 0.7 is set as the confidence level of an intent classification. This implies that if the confidence score of any intent goes below the defined threshold value, a fallback policy is triggered and the user is asked to rephrase the sentence. The overall goal is to ensure that no inconvenience or discomfort is caused to the user by any incorrect or hurtful responses.

5 Discussion

From the results, it is apparent that the conversations are not strictly domain-specific and can instead be more open-ended. As the conversations are open-

(a) Visual Screen   (b) Chat Screen   (c) Schedule Screen

Fig. 5. Application prototype

ended, it is difficult to account for all possible conversational intents and their relevant replies. This is where a need for an Urdu conversation generation model arises.

However, to our knowledge, no such model exists to perform this task in the Urdu language. Hence, this area can potentially be explored further.

Urdu, being Pakistan's national language is understood by almost everyone irrespective of their ethnic identity. For a language so widely spoken, there need to be technological advancements so applications like Saathi can be developed with ease. The features that makeup Saathi are usually only provided by applications that do not support the Urdu language and most elderly people in Pakistan are not fluent English speakers. This huge gap makes technology inaccessible for around 15 million [2] elderly people of Pakistan. Additionally, Urdu being a resource-poor language with respect to the NLP resources currently available means that it is very difficult for researchers and developers to effectively solve Urdu NLP related problems like Saathi.
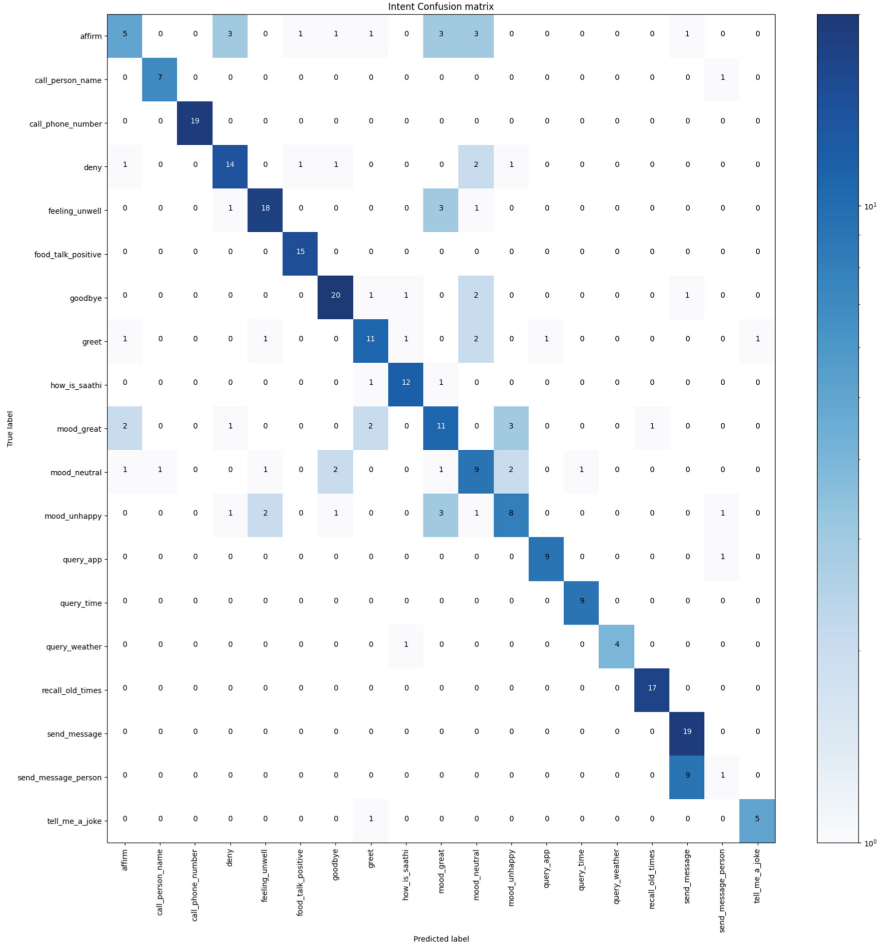
Fig. 6. Confusion matrix of intents with cross-validation

6   Conclusion

There are many virtual assistants that exist in English and other commonly used languages specifically designed for the elderly population. However, no such application currently exists in the Urdu language as per our knowledge. This paper presented a one of its kind Urdu language based virtual assistant named "Saathi", in the form of a mobile application. It can help the Urdu-speaking elderly individuals navigate through their smartphones with ease and provide them with a supportive environment to stay entertained and engaged in daily activities. Saathi serves as a companion to the elderly and helps them stay connected with their loved ones.

References

1. HOBBIT Project FP7. http://hobbit.acin.tuwien.ac.at/
2. Population by 5 year age groups - Pakistan. https://www.pbs.gov.pk/sites/defa ult/files//tables/POPULATION%20BY%205%20YEAR%20AGE%20GROUPS% 20-%20PAKISTAN.pdf
3. RAMCIP—Robotic Assistant for MCI Patients at home. https://ramcip-project. eu/
4. World's first Urdu based 'Siri-styled' voice-recognition AI bot launched. https://profit.pakistantoday.com.pk/2019/05/02/smart-cx19-focuses-customer- satisfaction-through-technology/
5. Alm, N., et al.: Engaging multimedia leisure for people with dementia. Gerontech-nology **8**, 236–246 (2009)
6. Beskow, J., Edlund, J., Granström, B., Gustafson, J., Skantze, G., Tobiasson, H.: The MonAMI reminder: a spoken dialogue system for face-to-face interaction. In: Tenth Annual Conference of the International Speech Communication Association. Citeseer (2009)
7. Bickmore, T.W., et al.: A randomized controlled trial of an automated exercise coach for older adults. J. Am. Geriatr. Soc. **61**(10), 1676–1683 (2013). https:// agsjournals.onlinelibrary.wiley.com/doi/abs/10.1111/jgs.12449
8. Bocklisch, T., Faulkner, J., Pawlowski, N., Nichol, A.: Rasa: open source language understanding and dialogue management. arXiv preprint arXiv:1712.05181 (2017)
9. Bunk, T., Varshneya, D., Vlasov, V., Nichol, A.: DIET: lightweight language under-standing for dialogue systems. arXiv preprint arXiv:2004.09936 (2020)
10. Com, I.: Global strategy and action plan on ageing and health (2017). http://apps. who.int/bookorders
11. Costa, A., Martinez-Martin, E., Cazorla, M., Julian, V.: PHAROS-PHysical assis-tant RObot system. Sens. (Switz.) **18**(8), 1–19 (2018)
12. Eisma, R., et al.: Mutual inspiration in the development of new technology for older people, January 2003
13. Gutierrez, F.J., Muñoz, D., Ochoa, S.F., Tapia, J.M.: Assembling mass-market technology for the sake of wellbeing: a case study on the adoption of ambient intelligent systems by older adults living at home. J. Ambient. Intell. Humaniz. Comput. **10**(6), 2213–2233 (2017). https://doi.org/10.1007/s12652-017-0591-4
14. Kasap, Z., Moussa, M.B., Chaudhuri, P., Magnenat-Thalmann, N.: Making them remember-emotional virtual characters with memory. IEEE Comput. Graph. Appl. **29**(2), 20–29 (2009)
15. Klein, J., Moon, Y., Picard, R.: This computer responds to user frustration: theory, design, and results. Interact. Comput. **14**(2), 119–140 (2002). https://doi.org/10. 1016/S0953-5438(01)00053-4
16. Lesin, S.: Frederick' medical virtual assistant. Final project in Software Engineering Department (2016)
17. Schüll, N.D.: Data for life: wearable technology and the design of self-care. BioSo-cieties **11**(3), 317–333 (2016)
18. Thakur, N., Han, C.Y.: An approach to analyze the social acceptance of virtual assistants by elderly people. In: Proceedings of the 8th International Conference on the Internet of Things, pp. 1–6 (2018)

19. Tun, S.Y.Y., Madanian, S., Mirza, F.: Internet of things (IoT) applications for elderly care: a reflective review. Aging Clin. Exp. Res. **33**(4), 855–867 (2020)
20. Uzma, A., Amir Zada, A.: The rising old age problem in Pakistan. J. Res. Soc. Pakistan **54** (2017)
21. Yasuda, K., Aoe, J.i., Fuketa, M.: Development of an agent system for conversing with individuals with dementia. In: vol. 27, pp. 3C1IOS1b2–3C1IOS1b2 (2013)