# Free software virtual assistants for designing pervasive gaming experiences to promote active aging: State of the art

Nacimiento-García, Eduardo[1][0000−0002−4075−0944], González-González. Carina S.[1][0000−0001−5939−9544], and Gutiérrez-Vela, Francisco L.[2][0000−0001−6629−7597]

[1] Universidad de La Laguna, La Laguna, Spain
{enacimie,cjgonza}@ull.es
[2] Universidad de Granada, Granada, Spain
fgutierr@ugr.es

**Abstract.** The design of pervasive gaming experiences to promote active aging through virtual voice assistants makes us raise two fundamental requirements, such as respect for user privacy and the possibility of developing proactive applications. These two requirements make the use of the main commercial assistants, unfeasible. In general, these devices are always listening and processing conversations without the need for explicit consent each time. These virtual assistants usually restrict some functionalities such as proactivity in Applications. Thus, in this paper, we present different free and open-source solutions to fulfill these requirements in implementing pervasive games.
Using free software-based tools, we can develop virtual assistants that meet our requirements and implement our pervasive gaming experiences to promote active aging. After analyzing the main free tools available, we highlight Mozilla DeepSpeech and Rasa, on the one hand, and on the other hand, the tools offered by the Stanford Open Virtual Assistant Lab (OVAL). It is important to note that these tools are modular to be mixed depending on our tastes and needs.

**Keywords:** pervasive game · active aging · virtual assistant · voice assistant · deep speech · rasa · free software.

## 1 Introduction

Nowadays, homes with a multitude of interconnected digital devices are increasingly common and these, along with mobiles, are already part of our daily lives.

In the same way, transversally, technology is linked to the daily lives of many people. We also have that the percentage of the world population over 60 years old does not stop increasing[1], and it is estimated that between 2000 and 2050, this percentage will more than double, reaching 11% to 22%[2].

It is essential to highlight the importance of active aging for people. For this, we must consider the specificities of this sector of the population and investigate

how to promote this active aging, which must include, for example, from performing physical activities to avoiding the problems derived from loneliness[2], because many times these types of people suffer from it.

Games, and especially pervasive games, together with the proliferation of interconnected digital devices, offer us a clear opportunity to deal with the issue of active aging from this perspective.

There are currently various virtual voice assistants on the market[3] and free tools to create virtual voice assistants specially adapted to our requirements[4]. We propose the use of virtual voice assistants together with pervasive games to promote active aging. There are two fundamental requirements when choosing which platform or technology to use. The first requirement is the possibility of the assistant's proactivity, and the second is respect for the user's privacy.

When it comes to using virtual assistants, we run into problems from the functionalities perspective and the privacy point of view. In the first case, we find the limitation of proactivity by the main existing virtual assistants, but this is a functionality that we consider very important. In the second case, we find how companies use and process the data of users who use their devices, which causes a severe loss of privacy. Thus, we will focus mainly on the analysis of free software tools that allow us to build virtual assistants that adapt to our needs and respect privacy.

This study aims to find tools that allow us to design experiences based on pervasive games through virtual assistants and that comply with the requirements that we have raised regarding proactivity and privacy.

This paper is organized as follows: first, we describe the main characteristics and needs in the design of pervasive gaming experiences to promote active aging; second, we introduce the use of virtual assistants considering the impossibility of adapting non-free virtual assistants to all our requirements, from among the free ones, we will see the most appropriate, and finally, the conclusions comment on the main findings of this work.

## 2 Use of pervasive gaming experiences to promote active aging

The percentage of people over 60 continues increasing, and there seems to be a clear trend for this percentage to continue to grow. Due to the large number of transformations produced in societies with the incorporation of ICT in daily life and especially in the field of digital entertainment, it is necessary to redesign and adapt these transformations to the needs and preferences of older people in order to incorporate this segment of the population into the digital culture and at the same time promote active aging through the use of these technologies.

It is necessary to adapt to these circumstances in all areas, including virtual assistants and games, among other things, to try to achieve active aging of the population. It is essential to analyze this type of technology, primarily through using the voice and avoiding graphic interfaces that could hinder older adults[5].

Along with aging, people are prone to present some problems derived from age, such as cognitive deterioration, loneliness, depression, etc. The use of virtual assistants can serve to try to avoid or improve the situation of older people regarding these problems[6]. Voice-controlled assistants have great potential in monitoring older people's health conditions, mainly due to the variety of sensors available that can be used. For example, sensors can be used to detect movement or pressure on a bed and depth sensors, among others, that can help us control the parameters of health and activity that these people carry out, respiratory or heart rates, behavior patterns, etc[7].

We want to place particular emphasis on proactivity due to the target audience for the resulting product. The devices must be capable of generating interactions with older people without waiting to be activated previously. This feature is essential for the correct use of virtual assistants' possibilities because we may be dealing with people unfamiliar with the use of technology, but also with people who suffer depression as a consequence of loneliness, among other things.

Technological evolution has allowed games to play an essential role in mobile devices today[8], and increasingly their redesign to be used with virtual assistants will become more common. A challenge arises when adapting games to the characteristics that these devices offer us and discover the new possibilities that arise. The field of pervasive games is wide and diverse in the approaches and technologies used. These games are used to create new experiences that can combine elements of real and virtual games[9]. With the advancement of pervasive games, we can achieve until now incredible or unthinkable forms of entertainment[10].

It is crucial to design experiences based on older adults' pervasive games since evidence indicates that older people who use games regularly tend to have a more positive attitude towards aging than other people at that age. From the socio-emotional point of view, applications must adapt to the requirements and preferences of the elderly be considered and developed[11].

The development of pervasive games focused mainly on older adults with virtual assistant technologies with voice and preserving users' privacy can serve this field's innovations. These innovations could be extrapolated to other areas of society with requirements or conditions other than the hegemonic ones in games and virtual assistants, such as childhood or people with functional or intellectual diversity.

## 3 Virtual assistants

One of Artificial Intelligence (AI) 's main objectives has been to achieve dialogues between people and computers. In recent years, the branch of AI focused on conversation systems has grown the most. Apart from voice recognition, other topics such as gesture recognition, image recognition, or videos are investigated[3].

In this section, we will see the main problems related to using the most used virtual assistants, especially considering the stated objectives related to the use

of these in the development of experiences based on pervasive games to promote active aging. Later we will focus on analyzing the main alternatives in the world of free software to get a virtual assistant that meets our requirements.

### 3.1 Top virtual assistants on the market

Many large companies implement their virtual assistant systems, such as Amazon Alexa, Google Home, Microsoft's Cortana, Apple's Siri, or Facebook's M. A critical issue to keep in mind regarding the most popular commercial voice assistants is that some of its functionalities are limited. So, features such as proactivity are not available[12], which could seriously limit this line of research's possible applications.

Serious security problems have been detected due to the way these virtual voice assistants work. In part due to the lack of a reliable authentication method, since it is possible that identity can be easily supplanted and, for example, perform purchases not authorized by the person who owns the assistant[13],[14]. Given the mistrust of these devices' users' vulnerabilities, a clear tendency has been detected to use them in private environments, mainly and avoid using them in public spaces[15].

Around the main virtual assistants, special attention must be paid to the privacy and data management that corporations do, because their commercial interests make them seek to maximize profits and take advantage of the value of user information. We have that while the people who use it express their concern that their rights are violated and want to control their personal information[16].

This concern regarding privacy and data use without user permissions seems to stand out, especially in Amazon Alexa, based on the number of articles that refer to it. However, this is probably mainly because Amazon controls 70% of the virtual voice assistant market. However, probably none of the different commercial systems is free from this concern. The use made of the collected data derived from a non-voluntary interaction with the devices is especially delicate. Analysis of the traffic produced by Alexa shows that this is part of regular operation[17].

The fact that these virtual assistants are continuously "listening" represents a serious privacy problem in which many people are either unaware of its possible implications or finally end up resigning themselves in exchange for using the technological advances that this type of platform offers[18].

### 3.2 Virtual assistants with Free Software

The main virtual assistants do not meet our demands for proactivity and privacy. After searching for alternatives, it is necessary to focus our efforts on finding free alternatives that allow us to implement virtual assistants who are proactive and respect the privacy of people who use them. We also need these tools to offer us the possibility of adapting it as much as possible to the people's demands and tastes for whom they are intended. In this case, mainly for older adults, without forgetting the other sectors of the population.

After analyzing the main free software tools available, it is observed that there are some tools with the capacity to be used to implement a virtual assistant with our demands. It is important to note that we usually do not have a single software that does all the work necessary to implement our virtual assistant. However, there are usually programs that do part of the work, and then each of these tools interconnects with each other, thus giving rise to the final product. We must emphasize that this modular approach often allows us to interconnect these "modules" with each other to adapt it even more to our needs. Even when the same project offers us several of these "modules," we can often use some of them and interconnect them with others without using all the same project elements.

Therefore, below we pay special attention to free tools that can allow us to build systems that respect the user's privacy and personal environment.

Also, we have that these systems can be adapted to the needs of our research and requirements, contrary to what can happen when the possibilities of adapting a proprietary system are usually limited.

• **Deep Speech:** The Baidu research team proposed a voice recognition system that uses deep learning and simplifies voice systems[19]. Other essential features of Deep Speech are that they do not use a phoneme dictionary, and that "you do not need hand-designed components to model background noise, reverberation, or speaker variation, but rather learn it directly"[19].

There is an implementation of Deep Speech made by Mozilla and released under the MPL (Mozilla Public License)[20]. It uses a neural network to convert the captured voice's spectrogram into the transcribed text; it is beneficial for us to use on the Internet of Things (IoT) devices. It can be used with or without a continuous internet connection. It includes pre-trained data sets. However, we can also add new sets[21].

Mozilla Deep Speech uses Google's Tensorflow to facilitate the speech's implementation to the text engine[22]. Such a project requires a large amount of heterogeneous data to function correctly. There is the Common Voice project, also from Mozilla, to collect this data using crowdsourcing through more than 50,000 people. The information has been collected in 38 languages, making it the largest body of free audio for voice recognition[23].

In non-ideal and rather noisy environments, the Mozilla implementation has a much lower error rate than other commercial systems from Google, Apple, or Microsoft. It also has a lower error rate in noise-free and combined environments, although the difference, especially with the Google system, is not wide[4].

We can find success stories where Mozilla tools can be integrated with other free tools such as Rasa. A virtual assistant can be built in a local environment, careful with privacy since it avoids sending packets over the internet. Mozilla DeepSpeech and Mozilla TTS are used to convert speech to text and vice versa. In this case, Rasa would be in charge of understanding natural language and managing dialogue[24].

• **OVAL:** Also notable is the "Stanford Open Virtual Assistant Lab (OVAL)," which aims to create a free virtual assistant ecosystem that respects privacy[25]. This project is divided into other projects, such as Thingpedia, LUInet, Genie, Almond, a communication protocol for a virtual assistant, and Brassau.

Almond is a free virtual assistant through crowdsourcing, and that preserves privacy; Almond is programmable and is designed to offer online and internet of things services[26]. Thingpedia offers us a public knowledge base with an open API and natural language interfaces[26]. Brassau has a graphical virtual assistant that converts natural language commands into a graphical interface[27]. We also have Genie, a natural language semantic parser generator for virtual assistant commands[28]. Soundr allows us to use an array of microphones like those included in most intelligent voice systems and infer the speaker's spatial locations and the head[29].

• **Rasa:** When we analyze Deep Speech, we talked about Rasa as part of a joint-use proposal between both tools. Rasa is a machine learning framework that allows us to automate texts and voice conversations. Rasa is free software under Apache 2 license. With Rasa, we can understand messages from conversations, but it also offers us the possibility of connecting to a wide variety of messaging channels and APIs[30].

Basically, Rasa's operation consists of receiving a message and converting it into a dictionary to be processed. Rasa has two main components, Rasa Core and Rasa NLU. Rasa Core uses dialogue management based on machine learning, and with this framework, we can create, for example, chatbots. Rasa NLU offers us a module of classification of intentions and extraction of entities[31]. Rasa NLU handles natural language processing.

• **Others tools:** To finish this section of virtual assistants, we will see other free tools that we could also use to construct a virtual assistant.

Tacotron 2 is a neural network architecture for speech synthesis from the text, developed by Google, which allows us to synthesize waveforms from spectrograms[32]. Tacotron allows an end-to-end voice system, which achieves very similar to that of a real human voice[33].

Studies have been carried out that mix the use of Deep Speech and Tacotron in the same project. Tacotron is used to recognize speech and convert it to text automatically, and Deep Speech is used to convert that text to speech[34].

Another system, designed by Facebook and under MIT license, is wav2letter++, a light, and straightforward tool to build speech recognition systems, which uses the ArrayFire library and is written in C++[35].

As an alternative to Rasa NLU, we also have Snips NLU, a machine learning-based voice platform connecting with embedded systems[36]. Snips are designed to work without an internet connection, so it allows you to perform processing without using the cloud, although you can also connect to the internet if it is wanted[12]. Snips allow us to guarantee the privacy of users[37].

## 4 Defense of privacy

The privacy of virtual assistant users is an important point to keep in mind. We have that the main commercial services do not comply with this requirement because confidential information is sent through the internet to be later processed in the cloud without the person's explicit consent who uses the device consciously or unconsciously. It is important to establish protocols that assure us that this type of "cheating" practice is not carried out.

The need arises to process the voice input locally, even in a "more primitive" way, until a request or an affirmative response to the assistant's proactive action is detected. After obtaining the user's permission, it would be possible to analyze the content obtained below and during that conversation. These protocols must allow us to enjoy the benefits of virtual assistants without violating privacy.

Many of the privacy problems that we have using the main virtual voice assistants are not unique in this type of device. However, problems derived from the Service as a Software Substitute (SaaSS)[38] are added to the voice that may be continuously being recorded and processed.

The problem related to privacy is not only solved with the drafting of terms and conditions of use. Your data will be collected and processed at all times. Also, many people will accept these abusive conditions even if they do not agree[15], only to enjoy these devices' technical advantages, so there should even be legal directives that regulate this type of practice by companies.

**Table 1.** Relationship between the tools and the required objectives.

| Tool | Proactivity | Privacy | Easy to use |
|---|---|---|---|
| Google Home | Low | Low | High |
| Amazon Alexa | Low | Low | High |
| Apple Siri | Low | Low | High |
| Mozilla & Rasa | High | High | Medium |
| OVAL | High | High | Low |

## 5 Conclusions

In this paper, we analyzed virtual assistants' use to design pervasive gaming experiences to promote active aging, and together we propose two essential requirements that virtual assistants must meet. These requirements are that the user's privacy is respected and that our application can be proactive because it is mainly intended for older people.

The main virtual assistants in the market are not respectful of user privacy, and also because these assistants usually have limited functionality, such as proactivity. For this reason, we decided to investigate whether there are free tools that allow us to develop virtual assistants that meet our requirements and allow us to create pervasive gaming experiences to promote active aging.

After analyzing the main free tools available, we have mainly chosen Mozilla and Rasa's tools, on the one hand, and on the other hand, the tools offered by the Stanford Open Virtual Assitant Lab (OVAL). It is important to remember that these tools are modular to be mixed depending on our tastes and needs.

Simplifying, we can divide the analyzed tools based on their proactivity, respect for privacy, and ease of use in our experiences based on pervasive games. In *Table 1*, we show the classification of the different tools based on the analysis carried out.

## Acknowledgement

Supported by "*Predoctoral training program for research staff in the Canary Islands of the Ministry of Economy, Knowledge, and Employment co-financed by the European Social Fund (ESF), with a co-financing rate of 85%*" and the project "*Design of pervasive gaming experiences using virtual assistants to promote active aging in older people within the digital home environment (RTI2018-096986-B-C32)*".

## References

1. Lutz, W., Sanderson, W., Scherbov, S.: The coming acceleration of global population ageing. Nature **451**(7179) (February 2008) 716–719. https://doi.org/10.1038/nature06516
2. Arslantas, H., et al.: Loneliness in Elderly People, Associated Factors and Its Correlation with Quality of Life: A Field Study from Western Turkey. Iranian Journal of Public Health **44**(1) (January 2015) 43–50 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4450013/
3. Këpuska, V., Bohouta, G.: Next-generation of virtual personal assistants (Microsoft Cortana, Apple Siri, Amazon Alexa and Google Home). In: 2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC). (January 2018) 99–103. https://doi.org/10.1109/CCWC.2018.8301638
4. Ordiales, H.: Comparativa de métodos de conversión de voz a texto Open Source. (2019) 6
5. De los Santos Cicutto, Santiago: Experiencia de uso de asistentes de voz sin GUI en personas mayores (2017)
6. Sunghoon, K., Parasuraman, G.M., Jaunbuccus, S.: Elderly Care Assistant: A Discreet Monitoring Tool. In Fleming, P., Lacquet, B.M., Sanei, S., Deb, K., Jakobsson, A., eds.: Smart and Sustainable Engineering for Next Generation Applications. Lecture Notes in Electrical Engineering, Cham, Springer International Publishing (2019) 287–301. https://doi.org/10.1007/978-3-030-18240-3_27
7. Shalini, S., et al.: Development and Comparison of Customized Voice-Assistant Systems for Independent Living Older Adults. In Zhou, J., Salvendy, G., eds.: Human Aspects of IT for the Aged Population. Social Media, Games and Assistive Environments. Lecture Notes in Computer Science, Cham, Springer International Publishing (2019) 464–479. https://doi.org/10.1007/978-3-030-22015-0_36

8. Arango-López, J., Gallardo, J., Gutiérrez, F.L., Cerezo, E., Amengual, E., Valera, R.: Pervasive games: giving a meaning based on the player experience. In: Proceedings of the XVIII International Conference on Human Computer Interaction. Interacción '17, Cancun, Mexico, Association for Computing Machinery (September 2017) 1–4. https://doi.org/10.1145/3123818.3123832
9. Magerkurth, C., Cheok, A.D., Mandryk, R.L., Nilsen, T.: Pervasive games: bringing computer entertainment back to the real world. Computers in Entertainment **3**(3) (July 2005) 4. https://doi.org/10.1145/1077246.1077257
10. Hinske, S., Lampe, M., Magerkurth, C., Röcker, C.: Classifying Pervasive Games: On Pervasive Computing and Mixed Reality. 21
11. Allairea, Jason C., et al.: Successful aging through digital games: Socioemotional differences between older adult gamers and Non-gamers. (2013). https://doi.org/10.1016/j.chb.2013.01.014
12. Jesús-Azabal, M., et al.: Voice Assistant to Remind Pharmacologic Treatment in Elders. (February 2020) 123–133. https://doi.org/10.1007/978-3-030-41494-8_12
13. Lei, X., et al.: The Insecurity of Home Digital Voice Assistants – Amazon Alexa as a Case Study. (November 2019) http://arxiv.org/abs/1712.03327
14. Zhang, N., et al.: Understanding and Mitigating the Security Risks of Voice-Controlled Third-Party Skills on Amazon Alexa and Google Home. (June 2018) http://arxiv.org/abs/1805.01525
15. Easwara Moorthy, A., Vu, K.P.L.: Voice Activated Personal Assistant: Acceptability of Use in the Public Space. In Yamamoto, S., ed.: Human Interface and the Management of Information. Information and Knowledge in Applications and Services. Lecture Notes in Computer Science, Cham, Springer International Publishing (2014) 324–334. https://doi.org/10.1007/978-3-319-07863-2_32
16. Norberg, P.A., Horne, D.R., Horne, D.A.: The Privacy Paradox: Personal Information Disclosure Intentions versus Behaviors. Journal of Consumer Affairs **41**(1) (2007) 100–126. https://doi.org/10.1111/j.1745-6606.2006.00070.x
17. Ford, M., Palmer, W.: Alexa, are you listening to me? An analysis of Alexa voice service network traffic. Personal and Ubiquitous Computing **23**(1) (February 2019) 67–79. https://doi.org/10.1007/s00779-018-1174-x
18. Lau, J., Zimmerman, B., Schaub, F.: Alexa, Are You Listening? Privacy Perceptions, Concerns and Privacy-seeking Behaviors with Smart Speakers. Proceedings of the ACM on Human-Computer Interaction **2**(CSCW) (November 2018) 102:1–102:31. https://doi.org/10.1145/3274371
19. Hannun, A., et al.: Deep Speech: Scaling up end-to-end speech recognition. (December 2014) http://arxiv.org/abs/1412.5567
20. Mozilla: mozilla/DeepSpeech (May 2020) original-date: 2016-06-02T15:04:53Z.
21. Firmansyah, M.H., Paul, A., Bhattacharya, D., Urfa, G.M.: A.I. based Embedded Speech to Text Using Deepspeech. (February 2020) http://arxiv.org/abs/2002.12830
22. : Welcome to DeepSpeech's documentation! — DeepSpeech 0.7.1 documentation
23. Ardila, R., et al.: Common Voice: A Massively-Multilingual Speech Corpus. (December 2019) https://arxiv.org/abs/1912.06670v2
24. Petraityte, J.: Build an AI voice assistant with Rasa Open Source and Mozilla tools | Rasa (August 2019)
25. OVAL: Stanford Open Virtual Assistant Lab (2020)
26. Campagna, G., Ramesh, R., Xu, S., Fischer, M., Lam, M.S.: Almond: The Architecture of an Open, Crowdsourced, Privacy-Preserving, Programmable Virtual Assistant. In: Proceedings of the 26th International Conference on World Wide

Web, Perth Australia, International World Wide Web Conferences Steering Committee (April 2017) 341–350. https://doi.org/10.1145/3038912.3052562
27. Fischer, M., Campagna, G., Xu, S., Lam, M.S.: Brassau: automatic generation of graphical user interfaces for virtual assistants. In: Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services, Barcelona Spain, ACM (September 2018) 1–12. https://doi.org/10.1145/3229434.3229481
28. Campagna, G., et al.: Genie: a generator of natural language semantic parsers for virtual assistant commands. In: Proceedings of the 40th ACM SIGPLAN Conference on Programming Language Design and Implementation - PLDI 2019, Phoenix, AZ, USA, ACM Press (2019) 394–410. https://doi.org/10.1145/3314221.3314594
29. Yang, J.J., et al.: Soundr: Head Position and Orientation Prediction Using a Microphone Array. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Honolulu HI USA, ACM (April 2020) 1–12. https://doi.org/10.1145/3313831.3376427
30. : Build contextual chatbots and AI assistants with Rasa
31. Singh, A., Ramasubramanian, K., Shivam, S.: Introduction to Microsoft Bot, RASA, and Google Dialogflow. In: Building an Enterprise Chatbot: Work with Protected Enterprise Data Using Open Source Frameworks. Apress, Berkeley, CA (2019) 281–302. https://doi.org/10.1007/978-1-4842-5034-1_7
32. Shen, J., et al.: Natural TTS Synthesis by Conditioning WaveNet on Mel Spectrogram Predictions. (February 2018) http://arxiv.org/abs/1712.05884
33. Wang, Y., et al.: Tacotron: Towards End-to-End Speech Synthesis. (April 2017) http://arxiv.org/abs/1703.10135
34. Chandran, S., Giri, S.: Voice Converter Using DeepSpeech and Tacotron. (2019) 6
35. Pratap, V., et al.: wav2letter++: The Fastest Open-source Speech Recognition System. ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (May 2019) 6460–6464. https://doi.org/10.1109/ICASSP.2019.8683535 arXiv: 1812.07625.
36. Staudigl, F.: Design and Implementation of an End-to-End Speech Assistant. (June 2019) https://dspace.cvut.cz/handle/10467/83444
37. Coucke, A., et al.: Snips Voice Platform: an embedded Spoken Language Understanding system for private-by-design voice interfaces. (December 2018) http://arxiv.org/abs/1805.10190
38. Richard Stallman: Who does that server really serve? (2010)