A Review of the Current Intelligent Personal Agents

Sean Li[1(✉)] and Xiaojun (Jenny) Yuan[2]

[1] Rosa International Middle School, Cherry Hill, NJ 08003, USA
seanhaoranli@gmail.com
[2] College of Emergency Preparedness, Homeland Security and Cybersecurity
University at Albany, State University of New York, Albany, NY 12222, USA
xyuan@albany.edu

Abstract. In artificial intelligence, an intelligent personal agent (IPA) is a software designed to use online resources to help people complete simple tasks and respond to questions with natural language. In the market, people wonder if there is a best IPA and have done preliminary testing on them [4]. In this study, we aim at investigating the use of IPAs in people's daily lives. We are also interested in identifying factors that may have an impact on the usability of IPAs. The IPAs we select are Apple Siri, Amazon Alexa, Google Assistant and Microsoft Cortana. We first compare the natural language user interfaces of the above mentioned four IPAs. Next, we examine how researchers have evaluated the use of the four IPAs with the focus on identifying factors affecting the usability of the IPAs. At the end, future research directions are discussed.

Keywords: Intelligent personal agents · Usability · User interface

1 Introduction

According to [1], an intelligent personal agent (IPA) is "software that has been designed to assist people with basic tasks, usually providing information using natural language". In [7], an IPA is "an application that uses inputs such as the user's voice, vision (images), and contextual information to provide assistance by answering questions in natural language, making recommendations, and performing actions". With the rapid technical development of natural language processing and artificial intelligence, IPAs are becoming an important assistant in people's professional work and daily lives [2]. IPAs respond to questions from users through using online resources. Users' questions vary depending on what they need at that specific moment, for example, weather, restaurant or driving directions. Most importantly, an IPA employs voice commands via a natural language user interface to assist users by answering their voice queries and carrying out tasks.

In a study comparing the use of four popularly used IPAs, including Apple Siri, Amazon Alexa, Google Assistant and Microsoft Cortana, Dunn [4] took an initiative to find out if there is a best IPA in the market. In this paper, we carried out a review on studies investigating the use of IPAs. In the following, we first introduce the user

interfaces of the four IPAs, and then review existing research on the use of IPAs with the focus on factors affecting the usability of IPAs. At the end, future research directions are discussed.

2   Four Intelligent Personal Agents (IPAs)

In the market, Apple Siri, Amazon Alexa, Google Assistant and Microsoft Cortana have been widely used by the public. When people talk to IPAs, a user-friendly interface is critical. The interfaces of the above-mentioned four IPAs are shown as below in Fig. 1. As shown in Fig. 1, Alexa is the only one that has no textual interaction with the users. The other three IPAs look similar at the interface level by providing a textual dialogue and recommendations to the users. All the four IPAs have voice input.
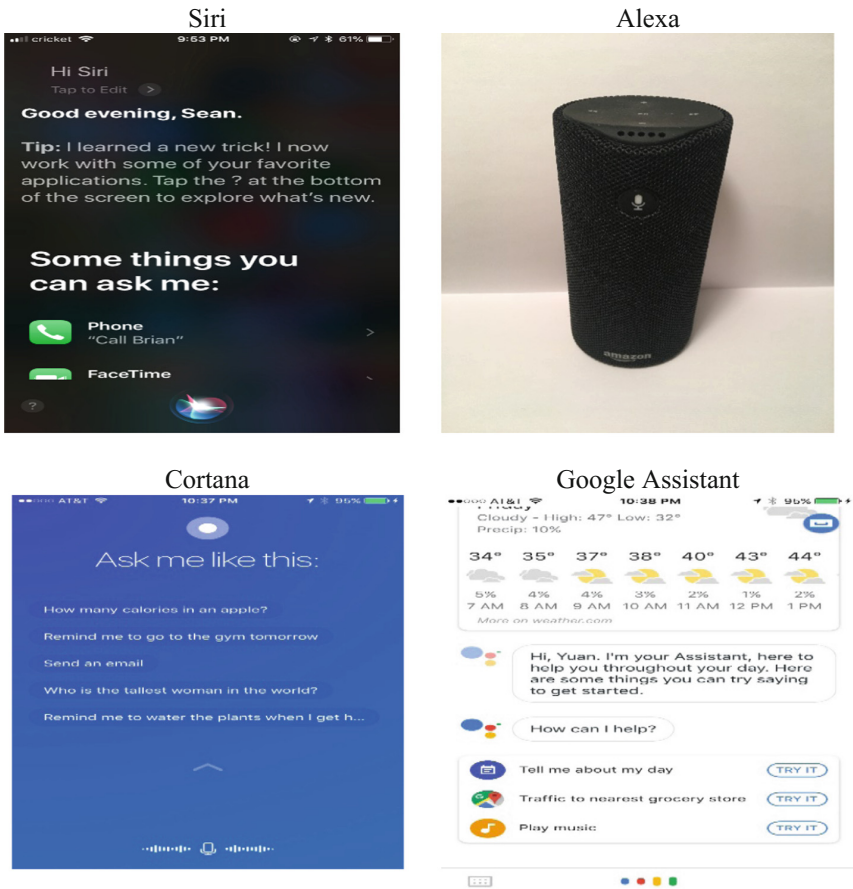

Fig. 1. The user interfaces of the four IPAs.

3   Comparison of the Four IPAs

Research has been done to compare the four IPAs. A study conducted by Dunn [4] highlighted the advantages and disadvantages each IPA had over the other in terms of categories including travel, email, messaging, sports, music, weather, calendar, social, translation, basic tasks, general knowledge and personality. Dunn reported that Google Assistant did very well in assisting with direction and sending emails. Siri came out on top in the areas related to phone call, text, and checking emails. Cortana was also quite good when sending texts, therefore tying with Google Assistant and Siri in this area. Alexa was efficient at reading tweets.

Each IPA has its pros and cons. Schultz [11] found that Cortana can analyze data at an efficient speed with accuracy, and correct pronunciation. However, at times, Cortana pulls up Bing when the answer is very simple, and it is unnecessary to input the query into Bing. As described by [10], Siri is useful when accessing settings, finding emails, doing mathematics, and transforming measurements. The few problems that Siri has are, at times, Siri has trouble comprehending what one is trying to say even though the level of speech is very basic. Also, if Wi-Fi dies then Siri goes along with it. Brandon [3] reported that Alexa understands complex speech and can come up with a reasonable if not accurate response. The main problem that Alexa has is that it crashes. In his article, Moore [9] indicated that Google Home, which runs Google Assistant, provides factual responses to questions asked, with efficiency also. Google Assistant can also understand follow up questions. The problem with Google Assistant is that it lags behind compared to other assistants when it comes to third party support.

With the growing demand of IPAs in the world, how to design a usable IPA that can satisfy users' professional and daily needs is becoming an important research topic. In the following, we discuss some factors that may have an impact on the usability of IPAs.

4   Factors Affecting the Use of IPAs

Research has shown that voice can affect the usability of interactive voice response systems that provide ubiquitous user interfaces to enable customers to collect information and perform tasks [5]. Specifically, [5] found out that voice personality and speaker gender have an impact on the perceived usability of the system. For example, male voices can lead to higher usability metrics than female voices.

The tasks an IPA performs include professional and personal tasks, and an IPA was designed to help users in doing his professional tasks while taking care of the personal tasks [2]. Users' personal tasks may change frequently based on their immediate needs. For different tasks, their complexity can vary. Dunn [4] indicated the potential impact of task type and task complexity on the usability of IPAs. In completing tasks, voice-based interactions may cause increased cognitive workload for users. Stayer et al. [12] tested the effect of voice-based interactions using 3 different IPAs (e.g., Apple's Siri, Google's Google Now for Android phones, and Microsoft's Cortana) on the cognitive workload of the driver. It seems that systematic differences exist between the smartphones. The Google system placed lower cognitive workload on the driver than

the other two systems. Further analysis demonstrated that such differences were associated with the number of system errors, the time to complete tasks, and the complexity and intuitiveness of the devices.

Miangah and Nezarat [8] claimed that "the speech aspect of mobile learning is as significant as textual aspect of it, since it enables learners to comfortably speak with a system recording their voice and allowing them to listen back to themselves (p. 314)".

Based on this claim, the way users interact or communicate with the IPAs could motivate users' learning [6]. For example, Goksel-Canbek and Mutlu [6] found out that the speech/language dialogues between users and Siri may help users improve language skills on speaking (pronunciation) and listening. Therefore, the design of the dialogue structure of the user interface is one of the factors that may influence the usability of IPAs.

In sum, factors such as voice personality, speaker gender, task types, the complexity and intuitiveness of the devices, as well as the design of the dialogue structure of the user interface could have an impact on the usability of IPAs.

5   Conclusion and Future Work

We conducted a review on studies investigating the use of IPAs. Four IPAs, that is, Apple Siri, Amazon Alexa, Google Assistant and Microsoft Cortana, were selected and compared. Factors that may have an impact on the usability of IPAs were identified.

As [2] claimed, "IPA will play a very important role in the near future". It is time for researchers and developers to catch the opportunities and also confront the challenges of the development of IPA. With the advance of artificial intelligence and natural language processing technologies, the future application of IPAs would be able to take into account users' emotion, personal characteristics, and their personal needs, and to deal with complex tasks in various settings, including education, health, and entertainment, etc.

In the field of human computer interaction, how to make IPAs more usable and make users feel pleasing and satisfactory is becoming a very important topic. As can be seen from Fig. 1, there exist differences between the interfaces of the four IPAs. We need to answer the question if and how the various interface features affect the usability of IPAs. In addition, it would be interesting to explore the relationship between different tasks (including task complexity and task types) and the usability of IPAs. Besides gender, other user characteristics such as domain knowledge may be considered. For example, we can examine if users' domain knowledge have an impact on the interaction between users and IPAs, in particular when dealing with complicated tasks in the medical field.

In the near future, we plan to do a crowdsourcing study to explore if and how the level of task complexity, the types of tasks, and users' domain knowledge can have an impact on the usability and user experience of IPAs.

References

1. Beal, V.: IPA - intelligent personal assistant. https://www.webopedia.com/TERM/I/intelligent-personal-assistant.html. Accessed 2 Jan 2018
2. Bhinderwala, A., Shukla, N., Cherarajan, V.: Intelligent personal agent. Int. J. Comput. Appl. 9–12 (2014). (0975 – 8887) National Conference on Role of Engineers in Nation Building 2014 (NCRENB-14)
3. Brandon, J.: The pros and cons of using Amazon Alexa as a model for chatbots. https://venturebeat.com/2016/07/30/the-pros-and-cons-of-using-amazon-alexa-as-a-model-for-chatbots/. Accessed 3 Dec 2018
4. Dunn, J.: We put Siri, Alexa, Google Assistant, and Cortana through a marathon of tests to see who's winning the virtual assistant race—here's what we found. Business Insider. http://www.businessinsider.com/siri-vs-google-assistant-cortana-alexa-2016-11. Accessed 2 Jan 2018
5. Edwards, R., Kortum, P.: He says, she says: does voice affect usability? In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 56, no. 1, pp. 1486–1490. SAGE Publications, Los Angeles (2012)
6. Goksel-Canbek, N., Mutlu, M.E.: On the track of artificial intelligence: learning with intelligent personal assistants. Int. J. Hum. Sci. 13(1), 592–601 (2016)
7. Hauswald, J., Laurenzano, M.A., Zhang, Y., Li, C., Rovinski, A., Khurana, A., Dreslinski, R.G., Mudge, T., Petrucci, V., Tang, L., Mars, J.: Sirius: an open end-to-end voice and vision personal assistant and its implications for future warehouse scale computers. In: Proceedings of the Twentieth International Conference on Architectural Support for Programming Languages and Operating Systems, pp. 223–238. ACM (2015)
8. Miangah, T.M., Nezarat, A.: Mobile-assisted language learning. Int. J. Distrib. Parallel Syst. 3(1), 309–319 (2012)
9. Moore, Q.: Google home pros and cons review: advantages and disadvantages. https://techranker.net/google-home-pros-cons-review-advantages-disadvantages/. Accessed 3 Dec 2018
10. Norton team: hey Siri: the pros and cons of voice commands. https://uk.norton.com/norton-blog/2016/02/hey_siri_the_prosa.html. Accessed 3 Dec 2018
11. Shultz, G.: The pros and cons of Cortana in build 9926 of Windows 10, 30 Jan 2015. https://www.techrepublic.com/article/the-pros-and-cons-of-cortana-in-build-9926-of-windows-10/. Accessed 3 Dec 2018
12. Strayer, D.L., et al.: The smartphone and the driver's cognitive workload: a comparison of Apple, Google, and Microsoft's intelligent personal assistants. Can. J. Exp. Psychol. 71(2), 93–110 (2017)