Hey Google, Have We Met Before?

Milica Pavlovic*, Stefanie Geukens, Kim Wery

Proximus

*Corresponding author e-mail: milica.pavlovich@gmail.com

Abstract:
Artificially intelligent (AI) personal assistants today are embedded in diverse devices and used in a truly ubiquitous manner, and we have already reached a stage in which major tech companies' assistants have an established identity towards global users. In this paper we discuss the concept of product familiarity in the light of designing interactions with existing AI agents. We present a design case study on embedding an existent AI assistant into an existent digital platform of a TV set-top box Pickx, and we investigate the implications it brings towards users' experiences. From testing the very initial prototype, we observed that there is a strong risk of perceiving discrepancy between the two interaction environments as well as non-seamless interactions. We believe that familiarity and user memory are prominent research areas when it comes to human-AI interactions, as the relations with existing, as well as new, AI assistants will become more established and their identities ever more entrenched.

Keywords: UX design; AI assistants; AI identity; product familiarity

1. Introduction

Human conversation is an essential component of social behaviour, considering it is among the main tools though which we express our intentions, emotions, attitudes as well as personality traits. We have seen this means of human interaction being leveraged within the human-computer interaction field, since it is considered to be very natural and intuitive for the users (Mariani et al., 2013). Even though conversational interfaces have been investigated in academia for quite a while, only in recent years we are witnessing their true expansion within consumer products, including chatbots, messaging apps and virtual personal assistants (McTear et al., 2016; McTear, 2016). Today these intelligent personal assistants are embedded in many different devices and used in a truly ubiquitous manner expanding the context of interaction on an ambient level (Santos et al., 2016; De Barcelos Silva et al., 2020). Major tech companies have been investing much into artificial intelligence (AI), such as deep learning and natural language processing, with the aim of creating commercially

available personal assistants that provide information and access to services through efficient natural conversational interactions (e.g. Google Assistant[1], Amazon Alexa[2], Apple Siri[3], Microsoft Cortana[4]).

As conversations are a natural and known means of interaction to humans, the expectations users put into conversational interfaces with personal agents are very high, and for the product to be successful it should be matching the seamless communications humans are used to. There are diverse approaches when designing human-AI interactions and recently we have seen a rise of design guidelines to this regard (see for e.g. Amershi et al., 2019). While accuracy of the received user input and provided response are a first parameter for a successful conversational interface, there are many other design aspects to take into consideration for enabling a seamless and natural communication. Among these aspects are the choice of words when responding to the user, comprehensive and engaging conversational flow, but also the affect that personal assistants bring as well as their personalities (McTear et al., 2016). Affect is a key factor in human conversation and thus modelling empathy in embodied personal assistants is seen as an important research area (McQuiggan & Lester, 2007). For achieving affect and empathy it is on one side needed to understand well the user (Mairesse & Walker, 2010; Ortigosa et al., 2014), and on the other to provide suitable and expected reactions, which could be embodied in a personal agent via, for e.g., emotional facial and tone of voice expressions (Moridis & Economides, 2012) or even eye contact (Hardjasa & Nakazawa, 2020).

Additional design aspect, which importance seems to be much on the rise in recent years, is the ethics of AI systems design that employ very personal and sensitive data of the users (Jobin et al., 2019; Jacovi et al., 2021). It has shown that the inherent risks that come with giving up sensitive data, as well as human control, are becoming a limiting factor for a broader adoption of AI agents. To this end we have seen design guidelines that reflect on ethics, where the aspects of transparency as well as explainability of the AI processes hold a significant position (Google PAIR, 2019; Samek et al., 2019). As transparency and explainability of AI are becoming more and more significant, we do see them rising as an inevitable new perceived user value. Users are expecting to have more knowledge and insights into an AI system they are interacting with, particularly in terms of which personal data are being used and for what scope, how are the data being processed and who has access to it (Pavlovic et al., 2019). It appears indeed that the focus in the design of AI systems have shifted from getting to know well the user, i.e. ensuring tailored personalized services, towards enabling the user to get to know well the AI assistant he/she is interacting with, i.e. ensuring transparency.

As some AI assistants exist on the global market for a few years already, we can observe that more and more people are introduced to them, meaning they already shaped an idea about what they can do with each of them and how the agents can be of assistance. We could thus deduct that some of the globally know agents (such as those of major tech companies previously mentioned) for some of the existing global users are already established entities, with which they already have developed certain degrees of familiarity.

In this paper we would like to present and discus a design case from the telecommunications company where we are currently engaged, Proximus, which refers to embedding an existing and established AI assistant within our current existing and established product. More precisely, it is a case of embedding the Google Assistant (GA) into a TV set-top box (i.e. TV decoder), through which our product platform Pickx offers both linear national TV channels as well as over-the-top content (via Netflix and Disney+, to name a few). Our aim was to improve the current product by offering to our users a possibility to interact via voice as well, ensuring more seamless interactions, and we opted for leveraging existent voice assistants rather than building one from scratch internally. In this context, the GA would be a means for our national customers to easily find the content they want to watch. In parallel with the technical analysis for implementing GA into our product, we needed to

---

understand the implications it brings from the side of users' experience (UX) as well. Besides having many guides on how to design an AI conversational assistant, we haven't come across any references that reflects on how to design interactions with an existing AI assistant which will be embodied in a new device and embedded within an existing product. This led us into making our own inquiry and motivated us in reporting it in this paper, in order to provide insights and support for design and research practitioners who will be facing similar issue in their practice.

In this paper we would like to discuss the concept of product familiarity in the light of designing interactions with existing AI agents. The concept of product familiarity is not new and was already discussed much in the field of marketing and branding as a potent aspect for successful strategies (Marks & Olson, 1981). Marks and Olson (1981) define product familiarity as cognitive structures of knowledge that are stored in memory and affect the behaviour, which could have derived from direct or indirect experiences with the product. While we do see some investigations related to interaction design and user familiarity (Viswanathan et al., 2016), as well as the usage of metaphors for designing interactive products for seniors (Turner & Van De Walle, 2006), we did not find any investigations related to familiarity with an established AI assistant, and in particular when it comes it its embodiment within an existing interactive product.

2. Research Questions and Methodology

Our main objective was to investigate UX implications when it comes to embedding an existing AI assistant, i.e. Google Assistant, into our existing product which represents a video content curation platform. More precisely, our research questions were the following:

- What are the perceived values and pain points of using a voice assistant within the TV context?

- How to seamlessly embed GA into our existing product Pickx from the UX point of view?

- What are the expected use cases from the users' point of view, and what is the perceived identity of the voice assistant in this context?

For responding to these questions, we conducted a qualitative study through semi-structured in-depth interviews and usability tasks. During the study, the GA was already integrated into Pickx as a very rough first version, which enabled us to already conduct some usability tasks to comprehend the reactions when communicating with the TV via the voice assistant. We firstly asked the participants about their knowledge on existing voice assistants, then we conducted the usability tasks, and finally we proceeded with gaining more detailed insights on their attitude and needs towards the feature. Interviews took place via online video calls, with a length of approximately 45 minutes per interview. In total we had 19 participants, which are our registered product testers (they have the product already installed at their homes), distributed according to the parameters indicted in Figure 1.

It is to note that, among 19, 8 participants are already using GA via Google Home and/or a smartphone as a hub for their IoT devices (e.g. light bulbs, sound system, cameras, vacuum cleaner, oven).
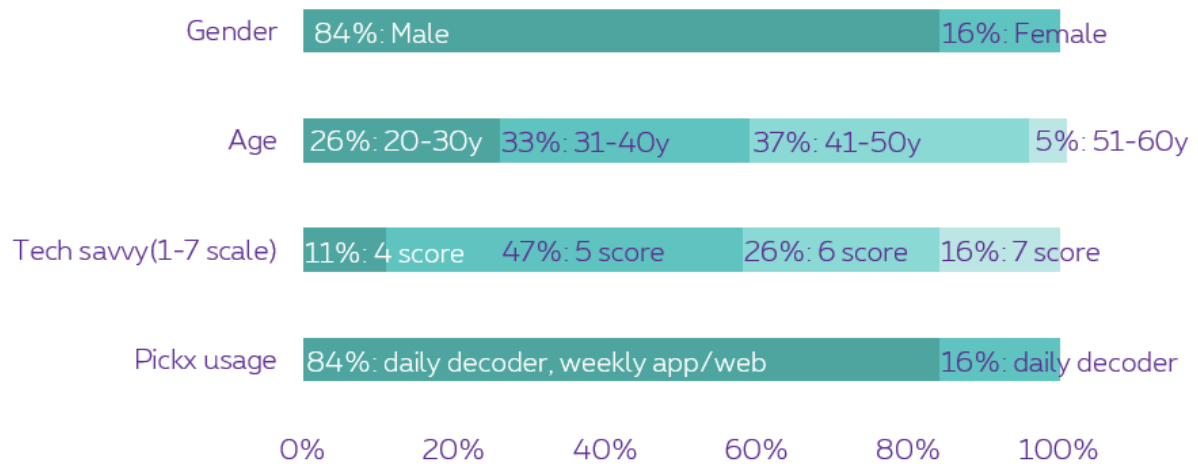
*Figure 1. Characteristics of the study participants.*

3. Results and Discussion

We will layout the results and provide a discussion related to UX and design implications according to the following points:

- Perceived value of the GA voice assistant in the TV context,
- GA embeddedness within Pickx,
- Assistant's perceived identity.

3.1 Perceived Value of the GA Voice Assistant in the TV Context

When asking the participants whether they find the usage of a voice assistant useful for the interaction with the TV, the opinions were quite divided (see Figure 2). It appears that the value of using a voice assistant is very much related to the context of usage; namely, if the context requires a hands-free interaction the value is higher (e.g. when cooking).
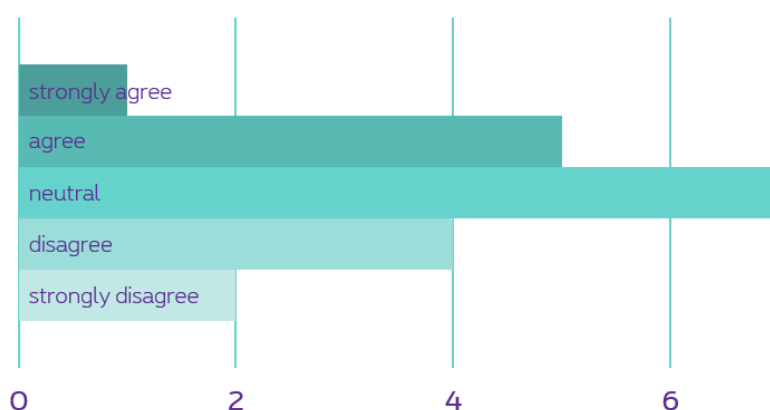
I find the voice assistant useful for the TV

Figure 2. Opinions about perceived usefulness of a voice assistant for the TV are divided.

In overall, the perceived advantages of the voice assistant for the TV context are seen in the light of it being (i) useful for simple commands, (ii) useful for navigation shortcuts, and (iii) providing effortless

rapid interaction (for e.g. there is no need for typing). On the other side, perceived disadvantages are that it (i) appears not useful for complex actions (currently it does not enable more complex conversation flows), and it (ii) provides an impression that the user has to adapt to the "language" of the assistant in order to interact with it successfully.

In order to understand what is perceived as the most valued use case when it comes to interacting with the voice assistant within the TV context, we gave the participants a task and we observed their behaviour. Namely, we asked them to think about their usual activities they do when inside Pickx, and then to try to do the same but relying on the voice assistant, we indicated that they can feel free to try out interactions that they find most useful in this context. Initially we assumed that the most expected and highest valued use case would be the search for content, however, the interviews showed us that it was not the case. It is to note that 79% of the participants stated that they do prefer saying out loud a content title via voice instead of typing it in via the remote, so the Search use case is valued high. However, the highest value and expectations were attributed to the navigation of the graphical user interface (GUI) via the voice (table 1). When navigating the GUI, majority of the participants see the remote controller and voice assistant co-existing in best case. Big difference is the number of clicks needed for arriving to a certain content, i.e. if it is 1-2 clicks via the remote then this the preferred interaction, otherwise voice is more preferred.

Table 1. Results from the exploration task.

UI Navigation						Content Search					Media Controls		
Open an app	Zap a channel	Go to Home	Go to TV guide	Go to personal recordings	Open settings	Content per title	Top content (e.g. movies)	Content per genre	Weather and news	What is on TV tonight	Volume control	Start recording	Pause and rewind
10	10	3	3	3	1	9	3	2	2	1	2	2	1

Most of the interviewees (89%, 17 persons) used the structure "verb + target" to communicate with GA, e.g. *"Go to my videos", "Start Pano".* Minority of them (11%, 2 persons) used complete sentences to give commands, e.g. *"What's on TV tonight?", "I would like to watch the recordings of my husband from September 30th at 8:15 p.m.".* By adding more words to the command, the participants expected the GA search to be more accurate. Although most participants used the structure "verb + target", (e.g. "open Netflix"), even in this small group, there was a big diversity in the wordings that were used. For example, channel zapping was communicated as *"Go to + channel", "Open + channel", "I want to see + channel", "Go to the next channel", "Go to channel + number", "Switch channel up".*

Some participants expected a conversational flow, meaning that by asking for more information the GA would guide them towards more refined results. In the large majority of observed commands users did not get the expected results, at least 47% of users rephrased their communications spontaneously, which they found very frustrating. Enabling a conversational flow with the GA would imply asking and receiving questions back in loops, thus not relying just on one-command interaction. An interviewee observed *"adding a "did you mean" response in order to correct your error would be good… get feedback until you get where you wanted to go."*

According to derived insights, we generated the following recommendations to support the expected values GA brings to the TV context, as well as an overall better UX:

- Supporting use cases beyond search for content, by enabling expected UI navigation and leveraging the value of simple rapid interactions via voice (starting off with the elements not easily reachable via the remote controller);

- Contributing semantic enrichment for navigation and Search use cases, while using as a base structure "verb + target" and expanding towards and enabling a more natural conversational flow.

3.2 GA Embeddedness within Pickx

All of the participants stated that having guides and suggestions on how to use the GA and how/which commands to form is highly needed, and it is currently missing. One participant noted: *"I still don't know the exact commands you can use with it, I have no clue, it would help to have a guide."* Surprisingly enough, currently there are suggestions provided by the GA on what a user can ask, but these are not seen as much helpful as they are not strictly related to the Pickx environment (Figure 3). These are suggestions generated by the GA by default, where the first suggestion is not a real and common use case, and the others are recognized as strictly Google-related.
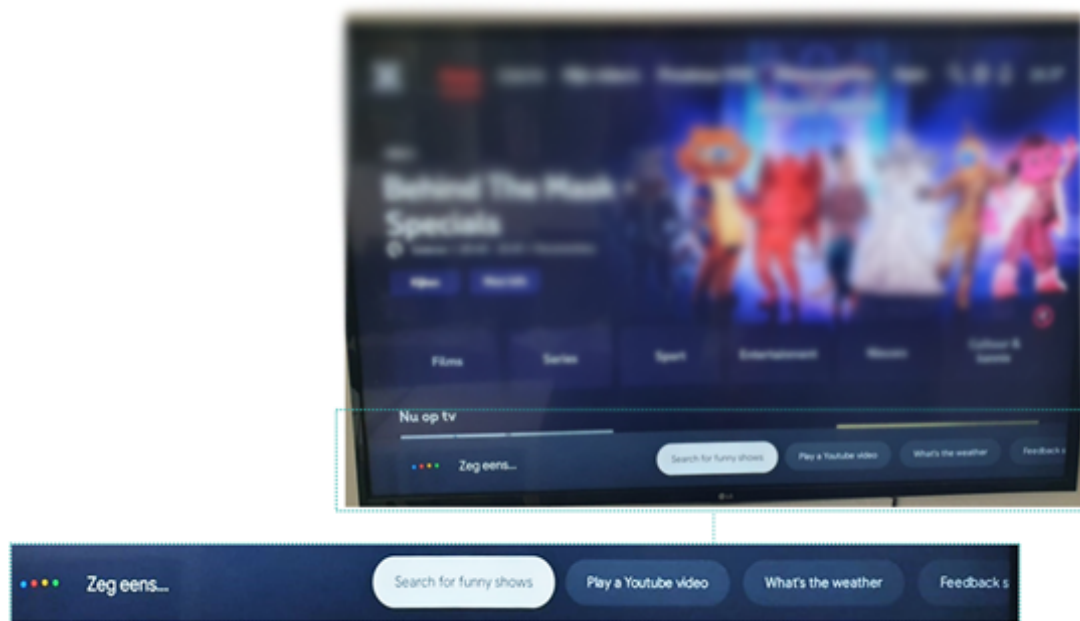

Figure 3. GA suggestions to the user on what to ask the assistant, embedded within the Pickx GUI.

Suggested improvements indicate that the current recommendations should be more Pickx related, and in terms of the GUI they should be presented with higher visual hierarchy. A participant stated that *"it looks more like a generic help from Google, not so linked to* Pickx*."*

Furthermore, 84% (16) of the participants expect the assistant to know their personal taste and provide a personalized overview of the content search. One interviewee suggested *"…you can add also a sentence like "surprise me", I would expect it to learn what I watch…"*.

According to derived insights, we generated the following recommendations to support better and seamless embeddedness of GA within Pickx, as well as an overall better UX:

- Improving accuracy and tightening the link with Pickx-related elements, features and contents;

- Providing guidelines and improving command suggestions for interaction with the GA inside of Pickx;

- Linking GA search results with Pickx personalized content recommendations.

3.3 Assistant's Perceived Identity

All 19 participants recognized the GA button on the remote controller as a function for voice recognition and recording; however, only 11 named it as the Google Assistant. Furthermore, 84% (16) of the participants discussed that they do not see the GA as a Pickx assistant, rather there is a distinguish between a Google and a Pickx interactive environment which are not synced. Supporting this observation of discrepancy, a participant noted *"…Pickx is the basic TV navigator, but GA is on top, it is not a tool for it because GA does not show the basic Pickx stuff…"*.

We observed that acceptance of GA and expectations related to it are very much dependent on the habits already developed with similar products. Some participants perceive that the voice interaction is "an inevitable future" they should adapt to, while others believe that they would stick to the remote controller because of the habit. An older representative of the interviewees noted *"…yes, why not via voice… but I am already old and have my habits, getting rid of the remote will be difficult…".* On the other side, representatives of the younger age group seem to be more experimental and more likely to multitask across devices, i.e. browse the web while watching the video, in order to gather additional information on the video content they are watching. Type of information they would be looking for are for e.g. video ratings, but also actors, places, historical events, etc. They underlined that using the GA for similar multitasking could be desirable, but it has to be implemented in a very seamless manner.

Furthermore, some participants stated that they do see GA as a Pickx assistant because it is actually limited to Pickx and the set-top box and cannot do what other GAs across other devices can. However, this could be influenced by the fact that these participants are already users of the Google IoT ecosystem. Additionality, it is to note that the ones who had interacted with a GA from before were already confident about which commands to ask and in which way, while the others were quite insecure in their construction of commands.

We concluded that, in our case, expectations and perception on the assistants' identity and what a user can do with is directly dependent on the user's previous experience with Google products (Figure 4).
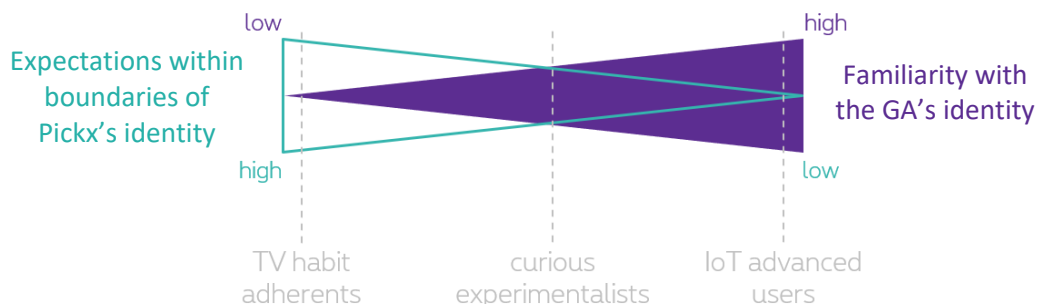

Figure 4. GA familiarity scale and design implications for the Pickx product.

In this line, we observed three diverse profiles of users:

1) <u>IoT advanced users with high expectations</u> - the ones that already use the Google IoT ecosystem, have quite high expectations on feature performance and consider using the GA from Pickx as part of the ecosystem enabling actions beyond TV (e.g. connecting their front-door camera);

2) <u>Curios experimentalists</u> - users with some previous experience with voice assistants who would be motivated to use GA more within the TV setting context, they see it as an advanced level of interaction that is going to be even more (inevitably) wide spread in the future;

3) <u>TV habit adherents with low expectations</u> - users not so familiar with voice assistants who lack stronger motivation for using GA, they do not have expectations of usage beyond the TV setting context and they strictly expect a Pickx-related assistant and not a GA.

According to derived insights, our recommendation for shaping a clearer identity of the voice assistant embodied within Pickx, and for avoiding perception of discrepancy, lies in taking a clear position on whether the use cases should expand towards connecting with the Google IoT ecosystem and steering the GUI design accordingly, or eliminating traces of GA identity and shaping a more Pickx-related unique identity for TV-only use cases.

4. Conclusion

We have presented a design case study on embedding an existent AI assistant (Google Assistant) into an existent digital platform of a TV set-top box Pickx, and we investigated the implications it brings towards users' experiences. From testing the very initial prototype, we observed that there is a strong risk of perceiving discrepancy between the two interaction environments as well as non-seamless interactions. Furthermore, we observed that assistant' identity within Pickx is very much dependent on user's familiarity with the Google Assistant, and the user's expectations around the interaction with the assistant are being built accordingly. To this end, in our design case we have identified three persona types that support us in further product definition and decision making, namely (1) IoT advanced users with high expectations, (2) Curious experimentalists, and (3) TV habit adherents with low expectations. Insights derived from this research can be leveraged in similar UX research and design cases when incorporating established AI agents into existing and established interactive products.

We believe that familiarity and user memory are prominent research areas when it comes to human-AI interactions, as the relations with existing as well as new AI assistants will become more established and their identities ever more entrenched. Therefore, for successful UX designs we should not overlook the expectations users are building in this context, as well as their need for transparency.

In the present work we investigated familiarity and AI assistant's identity mostly in the light of functionalities it provides towards the user. However, in the future work we would like to investigate the relations between affect and familiarity, as designing AI assistants will involve more affective and empathetic response towards the user, leaving him/her with the unique memory of how the assistant makes them feel.

References

Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., ... & Horvitz, E. (2019, May). Guidelines for human-AI interaction. In *Proceedings of the 2019 chi conference on human factors in computing systems* (pp. 1-13).

De Barcelos Silva, A., Gomes, M. M., da Costa, C. A., da Rosa Righi, R., Barbosa, J. L. V., Pessin, G., ... & Federizzi, G. (2020). Intelligent personal assistants: A systematic literature review. *Expert Systems with Applications*, *147*, 113193.

Google PAIR. (May 8, 2019). *People + AI Guidebook*. Retrieved from <u>pair.withgoogle.com/guidebook</u>

Hardjasa, L., & Nakazawa, A. (2020, July). An Examination of Gaze During Conversation for Designing Culture-Based Robot Behavior. In *International Conference on Human-Computer Interaction* (pp. 475-488). Springer, Cham.

Jacovi, A., Marasović, A., Miller, T., & Goldberg, Y. (2021, March). Formalizing trust in artificial intelligence: Prerequisites, causes and goals of human trust in ai. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 624-635).

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, *1*(9), 389-399.

Mairesse, F., & Walker, M. A. (2010). Towards personality-based user adaptation: psychologically informed stylistic language generation. *User Modeling and User-Adapted Interaction*, *20*(3), 227-278.

Mariani, J., Rosset, S., Garnier-Rizet, M., & Devillers, L. (2013). *Natural Interaction with Robots, Knowbots and Smartphones*. Springer.

Marks, L. J., & Olson, J. C. (1981). Toward a cognitive structure conceptualization of product familiarity. *ACR North American Advances*.

McQuiggan, S. W., & Lester, J. C. (2007). Modeling and evaluating empathy in embodied companion agents. *International Journal of Human-Computer Studies*, *65*(4), 348-360.

McTear, M. F., Callejas, Z., & Griol, D. (2016). *The conversational interface* (Vol. 6, No. 94, p. 102). Springer, Cham.

McTear, M. F. (2016, November). The rise of the conversational interface: A new kid on the block?. In *international workshop on future and emerging trends in language technology* (pp. 38-49). Springer, Cham.

Moridis, C. N., & Economides, A. A. (2012). Affective learning: Empathetic agents with emotional facial and tone of voice expressions. *IEEE Transactions on Affective Computing*, *3*(3), 260-272.

Ortigosa, A., Carro, R. M., & Quiroga, J. I. (2014). Predicting user personality by mining social interactions in Facebook. *Journal of Computer and System Sciences*, *80*(1), 57-71.

Pavlovic, M., Botto, F., Pillan, M., Criminisi,C., Valla, M. (2019). Social Consensus: Contribution to Design Methods for AI Agents That Employ Personal Data. In *Proceedings of the International Conference on Intelligent Human Systems Integration*, (pp. 877-883). Springer, Cham.

Samek, W., Montavon, G., Vedaldi, A., Hansen, L. K., & Müller, K. R. (Eds.). (2019). *Explainable AI: interpreting, explaining and visualizing deep learning* (Vol. 11700). Springer Nature.

Santos, J., Rodrigues, J. J., Casal, J., Saleem, K., & Denisov, V. (2016). Intelligent personal assistants based on internet of things approaches. *IEEE Systems Journal*, *12*(2), 1793-1802.

Turner, P., & Van De Walle, G. (2006). Familiarity as a basis of universal design. *Journal of Gerontechnology*, *5*(3), 150-159.

Viswanathan, V., Tomko, M., & Linsey, J. (2016). A study on the effects of example familiarity and modality on design fixation. *AI EDAM*, *30*(2), 171-184.

Author Bios:

Milica Pavlovic, design strategist and researcher, holds a PhD from Politecnico di Milano, within Proximus she was engaged as product and solutions specialist.

Stefanie Geukens, holds a Master's degree in Business Engineering from Universiteit Antwerpen, at Proximus she has held different positions in consumer products' development and management of data analytics projects.

Kim Wery, UX team lead for Proximus entertainment products, focuses on ensuring a coherent experience of products across the different platforms such as the TV decoder, Apple TV, Android TV, mobile, tablet and web.