A new friend in our Smartphone? Observing Interactions with Chatbots in the search of emotional engagement

Full Paper

Manuel Portela
Universitat Jaume I
Av. Sos Baynat, S/N
Castellón de la Plana, Spain 12007
portela@uji.es

Carlos Granell-Canut
Universitat Jaume I
Av. Sos Baynat, S/N
Castellón de la Plana, Spain 12007
carlos.granell@uji.es

ABSTRACT

We present the findings of a quantitative and qualitative empirical research to understand the possibilities of engagement and affection in the use of conversational agents (chatbots). Based on an experiment with 13 participants, we explored on one hand the correlation between the user expectation, user experience and intended use and, on the other, whether users feel keen and engaged in having a personal, empathic relation with an intelligent system like chatbots. We used psychological questionnaires to semi-structured interviews for disentangle the meaning of the interaction. In particular, the personal psychological background of participants was found critical while the experience itself allowed them to imagine new possible relations with chatbots. Our results show some insights on how people understand and empathize with future interactions with conversational agents and other non-visual interfaces.

CCS CONCEPTS

• **Human-centered computing → Natural language interfaces**; *Collaborative and social computing design and evaluation methods*; *Empirical studies in collaborative and social computing*;

KEYWORDS

Mixed-method analysis, Conversational agents, Empathic relations, Emotional Engagement

ACM Reference format:
Manuel Portela and Carlos Granell-Canut. 2017. A new friend in our Smartphone? Observing Interactions with Chatbots in the search of emotional engagement. In *Proceedings of 18th edition of the International Conference promoted by the Spanish Human Computer Interaction Association, Quintana roo MEXICO, September 2017 (Interaccion'17),* 7 pages.
https://doi.org/

1 INTRODUCTION

Artificial Intelligence (AI) is a wide topic and involves different algorithmic-driven software for knowledge representation, reasoning, and social intelligence, among others. During the last years, virtual assistants and conversational agents have been exponents of advanced Machine Learning processes and new technological capabilities. Nowadays, the use of this type of assistants, based on Natural Language Processing (NLP) and Speech Recognition is becoming more than ever natural since these systems are embedded in almost all smartphones and web interfaces. Moreover, the development of conversational agents, such chatbots, are widely spreading thanks to open APIs in messaging platforms (i.e. Telegram, WeChat, Facebook Messenger, etc).

The opportunity of conversational agents for marketing campaigns and costumer care services brought a strong excitement in the industry reflected in multiple online publications and news websites. Yet this phenomenon is less studied in the academic field, specially in Human-Computer Interaction. Opportunities and consequences can be found not only in new utilitarian appropriations but also in semiotic and affective relations with these socio-technical assemblages [29]. The importance of science fiction in ubiquitous computing [25], and the use of fiction as a research method [6], allow technologists to understand the opportunities of algorithmic interfaces [14, 20], while preventing unwanted consequences [11].

Embodied interactions with technology have been analysed by a sociological and ethnomethodological point of view by Paul Dourish [16], understanding that symbolic relations are not only important but also accountable. The use of smartphones and algorithmic applications can also modify our emotions, aesthetics and taste in automation processes [3]. The importance of the topic is approached by Uday Gajendar who called for *"HCI professionals to ensure these advancements are truly improving the human condition, enabling real relationships and supporting our daily activities."* [19].

For that reason, we developed an experiment based on quantitative and qualitative methods to understand the opportunities and effects on human behaviour with this type of interfaces in real-life situations. We present our first results of a series of interviews with 13 people, who interacted with two self-designed conversational agents, by evaluating their conversational interactions by using well-defined surveys from psychology and ethnographic methods.

2 RELATED WORK

Starting by a raw definition, *bots* can be described as AI systems with automation capabilities for assisting humans in specific tasks. Moreover, Conversational Agents can be understood as bots systems using text-to-speech and/or chat interfaces (chatbots). In this universe we can find different types of bots depending on their architecture [43].

Interaccion'17, September 2017, Quintana roo MEXICO

ACM ISBN 978-1-4503-5229-1…$15.00
https://doi.org/

Experiments on virtual assistants have been recurrently made in the AI and Human-Robot Interaction research fields proving different opportunities to create affection [37] and engagement with users [21]; even home-assistant robots can be considered social entities [23]. *Generative bots* (those aimed to generate new and creative content by using NLP algorithms [44, 48]) and *social bots* (those to interact in social networks, even with good [42, 45] or bad [18] intentions) have been studied as well [12].

Regarding conversational agents, recent publications have shown results on its usability and user apprehension to these new interfaces and technologies [31]. In [4] it is remarked the importance of *small talk* as a way to generate trust with the user, generating a social relation thanks to characteristics such as familiarity, solidarity and affect. The authors acknowledge that only extrovert people increase trust on social talks during embodied conversations, remaining unclear what happens with chat- or phone-based conversational systems. In [1] it is considered the difficulty of these complex systems to recreate natural turn-taking, change the dialogue domain, recognize intention and follow an incremental generation model of content in contextual conversations mixing dialogue with task management.

Even though there have been several attempts in the literature to develop effective conversational agents, consumer oriented products still offer basic dialogue capacity. Since our intention is to analyse user expectation and experience, we worked with the current state of chatbots rather than looking at more capable systems.

2.1 Hedonomics, Anthropomorphism and the Uncanny Valley

Despite our study does not approach embodied interactions with robotics or physical objects, it is important to remark the place of aesthetic on creating affection. *Hedonomics* [22] is defined as a Ergonomics discipline that intends to evaluate how the user measures its emotions. In the case of Anthropomorphism, Lemaignan [30] presented a model to evaluate its Dynamics. Anthropomorphism is defined as a social phenomena and includes "emotional states, motivations, intentions ascribed by the user to the robot" [30, p. 226]. This topic has been discussed many years ago resulting in several controversies [13]. In the Lemaignan's model, *familiarity* is a key element for user engagement, but disruption and cognition can modify the long-term perception. These concepts are present in our findings and will be later on compared.

The idea behind Anthropomorphism is that human-like robots and embodied systems can rise positive emotions. But it was also stated that excessive similarities can produce the opposite effect, known as the *Uncanny Valley* in robotics [34]. In this regard, Misselhorn [33] suggested that leading something to the users' imagination and perception, can help to elude repulsive reactions. In words of Brahnam [8], conversational agent's designers can rely on the *ethos* (the Aristotle's definition of moral character and goodwill) to generate credibility and trust by adapting agent behaviour to the user ethos.

2.2 Computers as Social Actors (CASA)

As with human-like characteristics, there have been intentions to analyse computers as social actors. That is, to understand how users perceive them not as machines, but as actors that follow social rules. In [35] it is presented a method to determine which *social cues* lead to users to act as if they were interacting with social actors instead of machines. This perspective provided some insights on how users can recognize different types of actors and provide social attributions. It gave us a basic understanding to drive our experiment.

Beyond the efforts in HCI for understanding this paradigm, Science and Technology Studies have a long trajectory considering technological artifact as actants, including notably Latour's Actor-Network Theory (ANT) and further discussions [28, 29]. In [41], it is proposed a framework to analyse interactions with everyday context-aware objects, relating ANT to virtual agents. In ANT, it is not only under consideration that computers are social actors, but it is established that non-human actants (including computers) can have agency over others actors.

2.3 Affective Computer and Emotional Intelligence

MIT researcher Rosalind Picard developed a conceptual implication of affection in machine development [40]. She based her "affective computing" theory on psychological tests and neural responses, linking decision making with the role of emotions. She argued that machines should recognize these emotions to effectively assist people. As such, Picard offered a set of useful guidelines to design affective systems. Later, Picard and colleagues [39] argued that machines should provide some affective feedback to the user's emotional state. This state would be measured by specific signals and context. But authors disclaimed of it because even humans have difficulties to determine correctly other' emotions. The proposed solution was to match gesture patterns to psychological signals, admitting some limitations. First, emotional response changes from day to day, making it difficult to determine continuity of measurement. And second, only spme specific features can be actually matched with real emotional states.

Despite the efforts made to progress in this topic, there still remain unknown factors to study about the implications of AI in our daily lives. For example, Cassell [9] stated that language ability does not make the agent more social, but non-verbal actions in cooperation. Beyond of primitive responses and facial expressions, it remains unclear how conversational situations impact on our emotions.

Moreover, social and cultural effects on algorithmic powered technologies have recently raised great interest in the HCI, Computer Science, and Sociology communities [7, 14]. Beyond technicalities on affective computing, it is a current concern how this feedback can modify the social condition and personal relations.

3 RESEARCH METHODOLOGY

HCI is an *interdisciplinary* field [5] as it provides a place for achieving the necessity of observing new social phenomena through a variety of methodologies [46]. In this line, our focus is on understanding **the nature of emotional engagement between the individual psychological mindset and a chatbot during a conversation**. To respond to this question, we devised an experiment that consisted of testing and evaluating the use of conversational agents and

their effect on people. The design, settings, and evaluation methods of the experiment are described in the following.

3.1 Experimental Design and Task

The strategy was to create two agents to test different behaviours, in order to compare the affection or the possible empathy with the participants. For doing so, we designed two platforms that enabled one-to-one conversations using a mobile phone. For the main interface at the client side (mobile phone), we used the Telegram Bot API in both agents, because it allowed us to have a common interface with a simple implementation. As noted in [37], conversations are preferred in personal proximity level, influencing an affective dominance in participants. We designed the behaviour of agents to reach enough intimacy by using informal language.

For the server side, we found that beside the raise of platforms for creating conversations through Machine Learning algorithms, few of them supported other languages than English. Since the local spoken language was Spanish, we developed the platform based on what best suited to our needs.

For *Chatbot 1* (CH1), we took the classical AIML (stands for Artificial Intelligence Mark-up Language, an XML-based format for specifying natural language software agents) processor based on the original ALICE project [47] using the open source software Program-O (PHP). The AIML-based processor is a widely used system to easily develop and configure conversational agents. Its simplicity though comes at the expense of some limitations for specifying sophisticated conversations to fully exploit the capabilities of natural language processing. Our choice was based on successful early experiments and its multi-lingual support, which fit the needs of the current experiment. We designed the conversations to make them apparently real and interesting by configuring the AIML-based processor with pre-defined conversational itineraries by means of matching short sentences or keywords during the conversation to respond accordingly. We also considered the possibilities of failures and errors, giving the chatbot the opportunity to behave and react accordingly to prevent it from getting stuck. For example, when a participant answered ambiguously, the chatbot either remained in the previous conversation thread or changed the topic. In case of confusion, the chatbot could avoid the answer or related it to a different topic. The lack of context, though, is the major problem of this type of processors, but also one of the main issues in AI in general.

For *Chatbot 2* (CH2), we used a Wizard of Oz method [36], where a person is pretending to be the bot. In order for a person (operator) to really operate the bot, we designed a web interface in NodeJS to interact with the Telegram API Bot to enabling a chat platform through which an operator could interact with the participants. Obviously, participants were not told about the real implementation behind CH2, and they believed the bot acted autonomous. To reduce bias in answers, the operator received guidance with behavioral rules similar to those followed by CH1. We decided to use this mechanism because only humans can behave freely and we wanted to improve the bot corpus. We expected that by having more contextual answers, participants would get more engaged in the conversation.

The experiment consisted of an individual session where each participant was given a mobile phone with Telegram installed and ready to talk with both chatbots. The expected time to spend for each chatbot was 10 minutes, but some of the conversations ended earlier on the request of participants.

3.2 Evaluation methods

The evaluation was conducted at three different stages of the experiment:

(1) Before the interaction with the bots, we gave the participants a set of surveys. As noted in [4], subjects might respond differently based on their predisposition. For measuring their emotional state minimising this issue, we used a Visual Analogue Scale (VAS) to account for 7 attitudes or variables (Happiness, Sadness, Anger, Surprise, Anxiety, Tranquility and Vigor). Next, we used the Spanish adapted version of the Interpersonal Reactivity Index (IRI) [32] to understand their disposition to empathy. Finally, we also developed a semi-structured interview to understand the previous knowledge and usage of the participants with respect to AI, Conversational Agents and chatbots (including Telegram and Facebook bots), Virtual Assistants (such as Siri, Alexa, and Google Now), and inquired them about their disposition to use these interfaces in the near future.
(2) After the interaction with both systems (CH1 and CH2), participants were asked to fill in a survey based on the Multidimensional Integrative Model (MIM), adapted from the proposed schema by Fernandez, Lopez & Marquez [17]. It was aimed to assess each of the interactions within a pro-social behaviour and sense of affiliation based on a cognitive and affective empathy. The MIM survey uses 5 variables (Discomfort, Frustration/anger/impotence, Interest, Satisfaction, Wellness) that correlate with different levels of Emotional Contagion and affective-cognitive empathy. It was followed by an interview where users explained their experience and highlighted whether their preconception changed about conversational agents.
(3) Once the session with participants ended, our analysis continued with a Thematic Network Analysis [2] of the resulting interviews. Similar to Luger & Sellen [31], we used this method to extract and build conceptual and sensitive insights based on recorded textual data. Additionally, we analysed response times from the timestamped logs of the conversations, giving us contextual bits of information on how the chatbots and participants have behaved.

Different ways of measuring usability, engagement and affection in interactions with Conversational Agents can be found in the literature [24, 37], but only providing partial results. It is clear that much remains to be discovered yet. As much of the communication process occurs in the participant interpretation, our methodology is oriented to consider previous states of the participants in order to clarify the affection process through an ethnographic process.

3.3 Participants

We conducted the experiment with 13 participants who can be classified in 3 groups by age: 3 participants in the 18 to 25 years

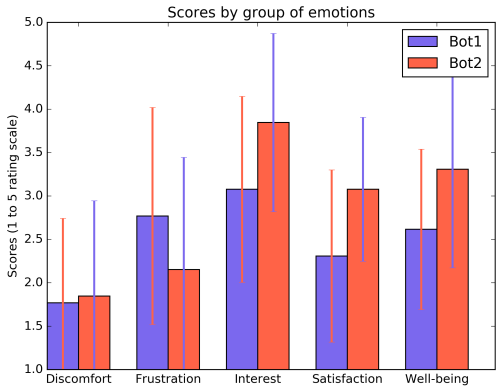

Figure 1: MIM scoring Mean and Standard Deviation for CH1 and CH2

group, 8 in the range 26-36 years, and 2 between 37 and 50. All of them have an undergraduate degree and mostly postgraduate studies (10 of 13, with background in Computer Science, Humanities and Psychology). In terms of gender, 9 of them were women.

The level of digital literacy in the group was high. Seven used computers on a daily basis up to 7 hours, and the remaining six stated that they spent more than 7 hours a day using it. All of them use smartphones on a daily basis, and 12 answered they use them very frequently. Regarding the use of virtual assistants, participants were asked if they knew or have used some of the publicly available platforms. Ten of them have used Siri or Google Now at least once, but only 3 of them - males, between 30 and 34 with experience in programming and highly educated - have had a conversation with a chatbot using either Facebook or Telegram.

4 FINDINGS

Main findings are presented according to the two groups of evaluation methods described earlier.

4.1 Evaluating the emotional and psychological state

In a first instance we found that participants already had a previous idea of a possible relation with virtual and conversational agents, and described fears, hope, excitement and curiosity in different emotional insights:

- Some participants were skeptic about the opportunity and remained so after the experiment, suggesting that the expected functionality was more oriented to a rational aspect of its use. For example, before the experiment one participant (P04) stated,*"it offers a solution that you already have, it's like I feel more comfortable doing it by myself"*, but at the same time he said that he felt old or antique since he is not embracing this new technology and prefers to do it by himself. Looking at the IRI Survey results, P04 scored highly in Empathic Concern (EC) and Personal Discomfort (PD) sub-scales and comparatively low at Fantasy. Nevertheless, after the experiment, the same participant was open to use it for regular tasks and entertainment activities. Other participants, though, were already excited to embrace a relation and felt positive after trying it.
- Other participants changed their emotions after having a good experience, leading to imagination for creating new relations with the system. Before and after the experiment, all participants described new functionalities for specific and routine tasks, or new uses of these technologies: *"I think about it for elderly people who have more rejection from technology or displays. By the fact that it is a conversation it can be easy for them."* or *"I think it could be smarter, if people can access to something different, an answer that is not the same for everyone"*.
- The results of VAS survey indicated that the attitude of each individual was a critical factor both in the development of the experiment and for the possibility of creating more intense conversations. *Intensity* can be described here as more time for chatting, more lines of conversation, more user's *domination* over the conversation, and more and varied utterances and *social cues* used during the conversation.

4.2 Evaluating the conversation experience

We found that the experience with the second chatbot (CH2) was better than the first (CH1), which is reflected in the MIM survey. Participants scored higher for interest (probability of affiliation but with emotional detachment), satisfaction (high probability of affiliation) and well-being (medium/low probability of affiliation) in relation to the second chatbot, while more frustration (low probability of pro-social behaviour) with regards to the first chatbot (see Figure 1). In terms of the number of conversation lines and the duration of the conversation, the difference between both experiences is also maintained (see Figure 2).

In spite of the absence of clear insights with the IRI survey, Spearman's correlation [49] showed that Fantasy (FS) sub-scale is a linear relationship with lower scores of discomfort (coefficient: 0.412, p-value: 0.1609) and frustration (coefficient: 0.530, p-value: 0.0619).

After analysing and examining the interviews with the Thematic Network Analysis [2, 31] method, we highlight here the following findings that will be later on analysed and described in detail in next section:

- Random, empty or duplicated answers and *topic turn-taking* by the chatbots were perceived as bad behaviour or logic.
- Keeping the context, humour and timing were essential for humanizing the process.
- Positive effects were caused by letting the conversation be open, using *social cues* and empathic signs, and touching personal topics from the bot side.

5 DISCUSSION AND CONTRIBUTIONS

In this section we describe our finding and interpretations of the experiment with the aim to understanding the nature of emotional engagement between the individual psychological mindset and a chatbot during a conversation.

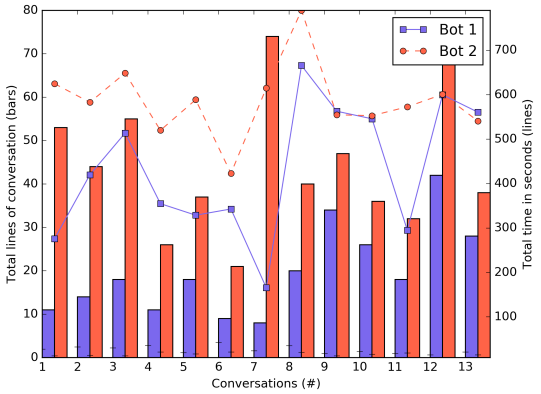

Figure 2: Time and Length Conversation Comparison between Bots

5.1 The preconception in the use of virtual assistants

We found that some participants already knew about AI systems because their personal or professional activities. Another distinctive group showed some reference to commonly used systems like Siri or Alexa. And some even showed an interest or pointed to Science Fiction series or films (i.e. *"After watching HER, the movie, I don't think so..."*). Those who brought the name of these fictions showed some concerns about privacy and also dystopic future ideas, while at the same time showed more excitement to do the test. Indeed, it is well known that familiarity with fiction is linked to embodied empathic behaviours [27], being also used as a HCI approach for designing new devices [15]. This suggests the importance of preconception about what these systems are able to do, and correlates with familiarity, as Lemaignan and colleagues aptly mentioned [30], when users expect pre-configured types of interactions. Some of these pre-configured behaviours during interviews include: answering while the participant was still writing (leaving the idea that the system can predict what is next), finding relation between topics, having a sense of humour or giving nimbly responses.

The implications on the preconception in the use of chatbots can still be threshed out as follows.

5.1.1 The influence of the professional background. Participants who work in computer science, or have programming skills, showed skeptic expectations about the possibilities of the chatbots. Most of them evaluated the system regarding its functionality, and were less inclined to be open-minded to have a different experience. In the previous section we mentioned a participant with background in Computer Science (P04) who scored high in PD sub-scale. Other two participants, who also were skeptic before starting the experiment, obtained similar results. On the contrary, people akin to humanities or in activities related to cognition abilities got surprised about the possibility of engagement with this sort of systems.

5.1.2 The expected learning process. The assumption that the system *did* actually learn (expected learning process) was a clear preconception highlighted by some participants. They believed that if they kept talking, the system would get better answers, because *"they will know better"*. One of the participants said, *"I suppose that the more it (the chatbot) deal with me, the more it can respond, right? And the more useful it will be, the more agile it will be. I suppose if that would give you more opportunities..."*. This belief though was not clearly showed up in the conversation logs, and it can be explained by considering the preconception effect. This shows that the modernist vision on the future of AI systems facilitates the engagement process.

5.2 The cross-empathy behaviour

Participants were strongly engaged when the chatbot remembered something that they have said before, even when it was just the participant's name. This effect of empathy from the system provoked on them the same empathic feeling. In another example, when the chatbot asked for details about some anecdote expressed by the user earlier, the rise of affection was immediately highlighted by the participant. The participants also mentioned an affection when the chatbot asked something personal. For example, one participant mentioned when the bot *"asked me a special experience for my life. In my childhood, he told me to tell him a little more of that. And I felt it as real."*

5.3 The timing bias

In several answers, we found that the response time of the system was critical. But it was unclear the effect of it over the participants' affection. Some of them found that the delay in the response was related to the intelligence of the system (assuming that it takes time to give a smart answer) while others expressed their disappointment in this lag effect. There is no receipt for this matter, since it depends on the situation and the user expectation which can be variable depending on the context and the intention.

5.4 The effect of unexpected

Unexpected behaviours in the system side brought some controversial insights. Most people found annoying that the system changed the topic of the conversation, but only some of them understood this behaviour as a way to keep the conversation alive. As noted in [30], the unexpected, depends on the perceived agent intention to be judged as a failure. This need of turn-taking behaviour in embodied conversational agents was already pointed by Cassell [9], where non-verbal actions can manage this kind of unexpected situations. But with the unexpected, we also found the presence of *serendipity connections* (the act of finding new and unexpected situations) that could bring different choreographs on the use of the system, supporting the value of previous findings in Kefalidou & Sharples [26].

6 CONCLUSIONS

The whole intention of this research was aimed to understand the possibilities to embrace some kind of personal relationship with chatbots. In the current stage of AI technology, most systems are oriented to specific functions instead of real or contextual intelligence. For that reason, we were aware of the low possibility to fulfill

the expected learning preconceptions, although, preconceptions were found critical for encouraging engagement with the chatbot.

We asked our participants if they would have future relations with these systems. Most of them were skeptic to have personal relations or more confidence as they perceived the current state of the art of conversational agents was not so advance to the point to enabling this sort of conversations and interactions. But, some of them talked about their experience with Siri which was more empathic because it makes jokes and gives funny answers. Also two participants expressed their intention to sharing this situation with friends, as a way of socializing their findings. Socialization is an important advantage because it shows a breaking point of the previous "functional" conception of systems. Regarding the CASA and ANT paradigms mentioned earlier, some participants agreed on being increasingly confident with trustworthy systems, even letting them make decisions on their behalf, which is undoubtedly a confirmation of both theories. Future designs should consider not only the opportunity to build new agents as a social actors but also consider ethical consequences of it.

It is worth exploring the combination of utterances, "social cues", discourse postures [10] and the animism [38] effect of AI, which may contribute to greater levels of engagement. We agree with Misselhorn who suggested that "an object which does not display a great number of human-like features, but very typical and salient ones, might do better in terms of perception-based empathy..." [33, p. 356].

We conclude that combining different behaviour strategies in the design of the agents would be helpful to positively engage with users, but the level of human-likeness should be relative and well balanced to the purpose. In other words, promoting anthropomorphism is helpful, as it also provides machine-like signals, preventing the user from being confused.

Finally, we presented a set of mixed qualitative and quantitative methods for evaluating these systems. Even though more and deeper proofs of validity are necessary, cross-domain analytic methods for evaluating this kind of interfaces have been proven valuable. While the experience is accountable through a qualitative approach, quantitative methods allowed us to have a closer understanding when they are used to systematically capture data.

7 LIMITATIONS AND FURTHER WORK

The presented results are partial in the sense that encounters were not strong enough to make a safe statement. More and varied experiments should be made to deeply analyse these phenomena. The experiment was held in an Spanish-centered environment potentially causing social bias, due to language and cultural constraints. Gender specific differentiation might also provide new insights, specifically with methods that presented differences in their results, like the IRI survey. Moreover, the interface used (Telegram) was chosen because it was regularly used by all participants, but cognitive effects could have been different using other types of user interfaces.

While the sample of participants is not enough to having a valid quantitative result, the experiment was focused on how mixed methods can complement each other, focused on the output of subjective appreciations. In this regard, highlighted results should be considered starting points to new future studies.

Although the technology possibilities of creating more sophisticated AI systems that would enhance bots capabilities are dramatically raising, future commercial, connected and location-based products will surely bring new scenarios to create fictional and non-fictional opportunities to study conversational interfaces in many and diverse real-life situations which are ever unimaginable today.

8 ACKNOWLEDGMENTS

The authors gratefully acknowledge funding from the European Union through the GEO-C project (H2020-MSCA-ITN-2014, Grant Agreement Number 642332, http://www.geo-c.eu/). Carlos Granell is funded by the Ramón y Cajal Programme of the Spanish government (grant number RYC-2014-16913).

REFERENCES

[1] James F Allen, George Ferguson, and Amanda Stent. 2001. An Architecture for More Realistic Conversational Systems. *Proceedings of Intelligent User Interfaces (IUI-01)* (2001), 1–8.

[2] Jennifer Attridge-Stirling. 2001. Thematic networks: an analytic tool for qualitative research. *Qualitative Research* 1, 3 (2001), 385–405. https://doi.org/10.1177/1468794107085301

[3] Nello Barile and Satomi Sugiyama. 2015. The automation of taste: A theoretical exploration of mobile ICTs and social robots in the context of music consumption. *International Journal of Social Robotics* 7, 3 (2015), 407–416. https://doi.org/10.1007/s12369-015-0283-1

[4] Timothy Bickmore and Justine Cassell. 2005. Social Dialogue with Embodied Conversational Agents. In *Advances in Natural Multimodal Dialogue Systems*, Jan C J van Kuppevelt, Laila Dybkjær, and Niels Ole Bernsen (Eds.). Springer Netherlands, Dordrecht, 23–54. https://doi.org/10.1007/1-4020-3933-6{_}2

[5] Alan Blackwell. 2015. Filling the big hole in HCI research. *Interactions* 22, 6 (2015), 37–41. https://doi.org/10.1145/2830317

[6] Julian Bleecker. 2009. Design Fiction: A Short Essay on Design, Science, Fact and Fiction. *Near Future Laboratory* March (2009), 49. http://tinyurl.com/nfl-dfiction

[7] Mark Blythe and Elizabeth Buie. 2014. Chatbots of the Gods: Imaginary Abstracts for Techno-Spirituality Research. *Proc. NordiCHI 2014* (2014), 227–236. https://doi.org/10.1145/2639189.2641212

[8] Sheryl Brahnam. 2009. Building Character for Conversational Agents : Ethos, Credibility and Believability. *PsychNology Journal* 7, 1 (2009), 9–47.

[9] Justine Cassell. 2000. Embodied conversational interface agents. *Commun. ACM* 43, 4 (2000), 70–78. https://doi.org/10.1145/332051.332075

[10] Justine Cassell, Yukiko I Nakano, Timothy W Bickmore, Candace L Sidner, and Charles Rich. 2001. Annotating and Generating Posture from Discourse Structure in Embodied Conversational Agents. *Workshop on Representating, Annotating, and Evaluating Non-Verbal and Verbal Communicative Acts to Achieve Contextual Embodied Agents, Autonomous Agents 2001 Conference 29* (2001). http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.21.171

[11] Kate Crawford. 2016. Artificial Intelligence's White Guy Problem - The New York Times. (2016). http://www.nytimes.com/2016/06/26/opinion/sunday/artificial-intelligences-white-guy-problem.html?smid=tw-share&_r=0

[12] Clayton A. Davis, Onur Varol, Emilio Ferrara, Alessandro Flammini, and Filippo Menczer. 2016. BotOrNot: A System to Evaluate Social Bots. In *Proceedings of the 25th International Conference Companion on World Wide Web*. 273–274. https://doi.org/10.1145/2872518.2889302

[13] Abbe Don, Susan Brennan, Brenda Laurel, and Ben Shneiderman. 1992. Anthropomorphism: from Eliza to Terminator 2. *CHI '92: Proceedings of the SIGCHI conference on Human factors in computing systems* (1992), 67–70. https://doi.org/10.1145/142750.142760

[14] Paul Dourish. 2016. Algorithms and their Others: Algorithmic Culture in Context. *Big Data & Society* December (2016), 1–11. https://doi.org/10.1177/2053951716665128

[15] Paul Dourish and Genevieve Bell. 2014. "resistance is futile": Reading science fiction alongside ubiquitous computing. *Personal and Ubiquitous Computing* 18, 4 (2014), 769–778. https://doi.org/10.1007/s00779-013-0678-7

[16] Paul Dourish and Graham Button. 1998. On "Technomethodology": Foundational Relationships Between Ethnomethodology and System Design. *Human-Computer Interaction* 13, 4 (1998), 395–432. https://doi.org/10.1207/s15327051hci1304{_}2

[17] I. Fernández-pinto, B. López-pérez, and M. Márquez. 2008. Empatía: Medidas, teorías y aplicaciones en revisión. *Anales de psicología* 24, 2 (2008), 284–298.

http://www.redalyc.org/articulo.oa

[18] Emilio Ferrara, Onur Varol, Clayton Davis, Filippo Menczer, and Alessandro Flammini. 2016. The rise of social bots. *Commun. ACM* 59, 7 (6 2016), 96–104. https://doi.org/10.1145/2818717

[19] Uday Gajendar. 2016. Empathizing with the smart and invisible. *Interactions* 23, 4 (2016), 24–25. https://doi.org/10.1145/2935195

[20] Tom Geller. 2016. Can Chatbots Think Before They Talk? *Commun. ACM* (4 2016). http://cacm.acm.org/news/201579-can-chatbots-think-before-they-talk/fulltext

[21] Randy Goebel. 2012. Intelligent Virtual Agents. In *12th International Conference, IVA 2012, Santa Cruz, CA, USA, September, 12-14, 2012. Proceedings (Lecture Notes in Computer Science)*, Yukiko Nakano, Michael Neff, Ana Paiva, and Marilyn Walker (Eds.), Vol. 7502. Springer Berlin Heidelberg, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-33197-8

[22] Martin Helander and Ming Po Tham. 2003. Hedonomics–affective human factors design. *Ergonomics* 46, 13-14 (2003), 1269–1272. https://doi.org/10.1080/0014013031000161081O

[23] Matthias Hoenen, Katrin T. Lübke, and Bettina M. Pause. 2016. Non-anthropomorphic robots as social entities on a neurophysiological level. *Computers in Human Behavior* 57 (2016), 182–186. https://doi.org/10.1016/j.chb.2015.12.034

[24] Robert C. Hubal, Diana H. Fishbein, Monica S. Sheppard, Mallie J. Paschall, Diana L. Eldreth, and Christopher T. Hyde. 2008. How do varied populations interact with embodied conversational agents? Findings from inner-city adolescents and prisoners. *Computers in Human Behavior* 24, 3 (2008), 1104–1138. https://doi.org/10.1016/j.chb.2007.03.010

[25] Jofish Kaye and Paul Dourish. 2014. Special issue on science fiction and ubiquitous computing. *Personal and Ubiquitous Computing* 18, 4 (4 2014), 765–766. https://doi.org/10.1007/s00779-014-0773-4

[26] Genovefa Kefalidou and Sarah Sharples. 2016. Encouraging serendipity in research: Designing technologies to support connection-making. *International Journal of Human-Computer Studies* 89 (5 2016), 1–23. https://doi.org/10.1016/j.ijhcs.2016.01.003

[27] David Comer Kidd and Emanuele Castano. 2013. Reading literary fiction improves theory of mind. *Science (New York, N.Y.)* 342, 6156 (2013), 377–80. https://doi.org/10.1126/science.1239918

[28] Bruno Latour. 1996. On actor-network theory. A few clarificaitons plus more than a few complications. *Soziale Welt* 25, 3 (1996), 1–16. https://doi.org/10.2307/40878163

[29] Bruno Latour. 2006. *Reassembling the Social: An introduction to Actor-Network Theory*. Oxford University Press, USA, Oxford. 320 pages.

[30] Séverin Lemaignan, Julia Fink, and Pierre Dillenbourg. 2014. The dynamics of anthropomorphism in robotics. *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction - HRI '14* (2014), 226–227. https://doi.org/10.1145/2559636.2559814

[31] Ewa Luger and Abigail Sellen. 2016. "Like Having a Really Bad PA". In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*. ACM Press, New York, New York, USA, 5286–5297. https://doi.org/10.1145/2858036.2858288

[32] Vicenta Mestre Escrivá, Maria Dolores Navarro Frías, and Paula García Samper. 2004. La Medida De La Empatía: Análisis Del Interpersonal Reactivity Index. *Psicothema* 16, 2 (2004), 255–260.

[33] Catrin Misselhorn. 2009. Empathy with inanimate objects and the uncanny valley. *Minds and Machines* 19, 3 (2009), 345–359. https://doi.org/10.1007/s11023-009-9158-2

[34] Masahiro Mori, Karl F. MacDorman, and Norri Kageki. 2012. The uncanny valley. *IEEE Robotics and Automation Magazine* 19, 2 (2012), 98–100. https://doi.org/10.1109/MRA.2012.2192811

[35] Clifford Ivar Nass, Jonathan Steuer, and Ellen R. Tauber. 1994. Computers are social actors. *Computer-Human Interaction (CHI) Conference: Celebrating Interdependence 1994* JANUARY (1994), 72–78. https://doi.org/10.1145/259963.260288

[36] Masayuki Okamoto, Yeonsoo Yang, and Toru Ishida. 2001. Wizard of Oz Method for Learning Dialog Agents. In *Cooperative Information Agents V: 5th InternationalWorkshop, CIA 2001 Modena, Italy, September 6–8, 2001 Proceedings*, Matthias Klusch and Franco Zambonelli (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 20–25. https://doi.org/10.1007/3-540-44799-7{_}3

[37] Timo Partala, Veikko Surakka, and Jussi Lahti. 2004. Affective Effects of Agent Proximity in Conversational Systems. In *Proceedings of the Third Nordic Conference on Human-computer Interaction (NordiCHI '04)*. ACM, New York, NY, USA, 353–356. https://doi.org/10.1145/1028014.1028070

[38] Matteo Pasquinelli. 2016. Abnormal Encephalization in the Age of Machine Learning. (2016). http://www.e-flux.com/journal/abnormal-encephalization-in-the-age-of-machine-learning/

[39] R.W. Picard, E. Vyzas, and J. Healey. 2001. Toward machine emotional intelligence: analysis of affective\nphysiological state. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23, 10 (2001), 1175–1191. https://doi.org/10.1109/34.954607

[40] Rosalind W. Picard. 1997. Affective Computing. *Affective Computing* (1997), 304. https://doi.org/10.1007/BF01238028

[41] Manuel Portela and Carlos Granell-Canut. 2016. Methods to Observe and Evaluate Interactions with Everyday Context-Aware Objects. In *Ubiquitous Computing and Ambient Intelligence*. Number 1. Springer, 385–392. https://doi.org/10.1007/978-3-319-48746-5{_}39

[42] Saiph Savage, Andres Monroy-hernandez, and Tobias Hollerer. 2016. Botivist : Calling Volunteers to Action using Online Bots. *19th ACM conference on Computer-Supported Cooperative Work and Social Computing* (2016), 813–822. https://doi.org/10.1145/2818048.2819985

[43] Pavel Surmenok. 2016. Chatbot Architecture. (2016). https://medium.com/@surmenok/chatbot-architecture-496f5bf820ed#.29q6brn2u

[44] Nava Tintarev, Ehud Reiter, Rolf Black, Annalu Waller, and Joe Reddington. 2016. Personal storytelling: Using Natural Language Generation for children with complex communication needs, in the wild.... *International Journal of Human-Computer Studies* 92 (2016), 1–16. https://doi.org/10.1016/j.ijhcs.2016.04.005

[45] Carlos Toxtli, Claudia Flores-Saviaga, Flor Aguilar, Alejandra Monroy, Juan Flores, Jeerel Herrejon, Norma Chavez, William Dai, Needa Almani, Shloka Desai, and Saiph Savage. 2016. BotViz: Data Visualizations for Collaborations With Bots and Volunteers. *Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion* (2016), 421–424. https://doi.org/10.1145/2818052.2869132

[46] Tommaso Venturini and Bruno Latour. 2010. The Social Fabric: Digital Traces and Quali-quantitative Methods. *Proceedings of Future En Seine 2009* (2010), 87–101. http://www.tommasoventurini.it/web/uploads/tommaso_venturini/TheSocialFabric.pdf

[47] Richard S Wallace. 2009. The Anatomy of A.L.I.C.E. In *Parsing the Turing Test: Philosophical and Methodological Issues in the Quest for the Thinking Computer*, Robert Epstein, Gary Roberts, and Grace Beber (Eds.). Springer Netherlands, Dordrecht, 181–210. https://doi.org/10.1007/978-1-4020-6710-5{_}13

[48] Alex Wilkie, Mike Michael, and Matthew Plummer-Fernandez. 2015. Speculative method and Twitter: Bots, energy and three conceptual characters. *Sociological Review* 63, 1 (2015), 79–101. https://doi.org/10.1111/1467-954X.12168

[49] Daniel Zwillinger. 2002. *CRC Standard Mathematical Tables and Formulae, 31st Edition*. CRC Press. https://books.google.es/books?id=gE_MBQAAQBAJ