

Towards Vocally-Composed Personalization Rules in the IoT

Original

Towards Vocally-Composed Personalization Rules in the IoT / De Russis, Luigi; Monge Roffarello, Alberto; Borsarelli, Carlo. - ELETTRONICO. - 3053:(2021), pp. 1-5. (Intervento presentato al convegno 2nd International Workshop on Empowering People in Dealing with Internet of Things Ecosystems (EMPATHY 2021) tenutosi a Bari (Italy) nel August 30, 2021).

Availability:

This version is available at: 11583/2910012 since: 2021-12-22T14:49:21Z

Publisher:

CEUR-WS

Published

DOI:

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

Towards Vocally-Composed Personalization Rules in the IoT

Luigi De Russis¹, Alberto Monge Roffarello¹ and Carlo Borsarelli¹

¹Politecnico di Torino, Dipartimento di Automatica e Informatica, Corso Duca degli Abruzzi 24, 10129 Torino, Italy

Abstract

This paper presents a study aimed at understanding whether and how end users would converse with a conversational assistant to personalize their domestic Internet-of-Things ecosystem. The underlying hypothesis is that users are willing to create personalization rules vocally and that conversational assistants could facilitate the composition process, given their knowledge of the IoT ecosystem. The preliminary study was conducted as a semi-structured interview with 7 non-programmers and provided some evidence in support of this hypothesis.

Keywords

Internet of Things, End-User Development, Conversational Agents, Vocal Interaction

1. Introduction and Goal

Smart speakers like the Amazon Echo are entering our homes and enriching the Internet of Things (IoT) ecosystem already present in them. The Intelligent Personal Assistants (IPAs) they provide allow people to easily connect to their online searches, music, IoT devices, alarms, and wakes [1]. Such IPAs have some end-user personalization capabilities present in the system, but they are segregated in a mobile app and take no advantages from their Natural Language Processing capabilities, nor from their knowledge of the IoT ecosystem in which the smart speaker is inserted. In our previous work [2], we reflected on the more prominent roles of these devices in personalizing domestic IoT ecosystems, exploring different steps of the process (creation, explanation and debugging, etc.) where they can be particularly useful. Recently, other researchers investigated the role of IPAs for end-user development in the IoT, either proposing an evolution of the IoT ecosystem definition to include smart speakers and their conversational agents [3], or exploring how the voice-based support offered by Amazon Alexa could be integrated in a platform to support the creation of trigger-action rules [4].

Stemming from those reflections and motivated by the reported related works, we started to explore novel approaches for **creating** personalization rules through conversation between a user and an IPA, when the IPA is embedded in a smart speaker. In particular, we were interested in understanding *a*) whether and how end users are willing to create personalization rules,

EMPATHY: 2nd International Workshop on Empowering People in Dealing with Internet of Things Ecosystems.
Workshop co-located with INTERACT 2021, Bari, Italy

✉ luigi.derussis@polito.it (L. De Russis); alberto.monge@polito.it (A. Monge Roffarello);
carlo.borsarelli@studenti.polito.it (C. Borsarelli)

ORCID 0000-0001-7647-6652 (L. De Russis); 0000-0002-9746-2476 (A. Monge Roffarello)



© 2021 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

 CEUR Workshop Proceedings (CEUR-WS.org)

Table 1
Study participants

#	Age Range	Job	Smart Speaker's Owner
P1	23-27	Speech therapist	Yes
P2	18-22	Student	Yes
P3	48-52	Housewife	Yes
P4	23-27	Health technician	No
P5	23-27	Student	No
P6	43-47	Health worker	No
P7	23-27	Math teacher	Yes

vocally, *b*) by using which rule format, and *c*) which kind of support they are expecting from the IPAs during the creation process. To take a step forward this goal, we conducted 7 interviews with non-programmers with different occupations and backgrounds, described in the remainder of the paper.

2. User Study

We recruited participants through convenience and snowball sampling, by sending private messages to our social circles. To minimize self-selection bias, we selected participants to enroll end users that have a medium-high interest in home automation, have a mix of participants using/not using a smart speaker with an IPA, and balance our population in terms of occupation and educational background. Our final sample included 3 participants who self-identify as male and 4 who self-identify as female, with age ranging from 18 to 52 years old. Table 1 summarizes the relevant information. All participants currently live in Italy and the study was conducted in Italian.

All participants completed a two-part study session composed of a background interview and an imagination exercise. Due to the Covid-19 pandemic, those one-to-one study sessions were conducted partially online (with Zoom) and partially in-person during the month of April 2021. Study sessions lasted from 25 to 40 minutes.

Background interview. We first conducted a background, semi-structured interview to understand which are the relationships that users have with smart speakers (if they have one) or with IPA's in general. We also asked about the experience that participants have with home automation and IoT devices, providing examples when possible. Questions included: "*Which are the main issues you experienced with a smart speaker?*" and "*In an IoT-powered home, which activities would you like to automate?*"

Imagination exercise. After the background interview, we conducted an imagination exercise to elicit, directly from the interviewed participants, how they will create personalization rules in different scenarios by using a smart speaker. Since not all the participants might have knowledge of end-user personalization in the IoT, we briefly introduced them to the topic. The collected information allows us to explore the possibilities and approaches that a non-programmer would use to freely create custom rules via conversation. Participants received a

description of a home (e.g., a fully IoT-powered home with a smart speaker for each room) and two personalization goals: 1) “*you would like to turn on the central kitchen light every time you enter that room*” and 2) “*you want that, when you go to bed, the shutters of the house are closed and all the lights go out*”. Participants have to express, freely but vocally, an instruction (i.e., a rule) for realizing each goal and then answer some questions about it.

3. Results

All the participants have some knowledge and experience with smart speakers; as expected, smart speaker’s owners have a more extended knowledge of the possibilities and limitations of such devices, while the others used them in a more sporadic way, e.g., while at a friend’s home. They demonstrated, however, to know at least the basic features of smart speakers, especially of the Amazon Echo. For what concerns home automation and personalization rules, instead, only 2 participants (P2 and P3) have smart devices at home but they are not in charge of configuring devices and creating personalization rules. However, P1, P3, and P7 knew about the rule composition capability included in the mobile app of smart speakers, the Amazon Echo in this case. P1 knew about such a capability, but never tried to create a rule. P3 came up with interesting rules, but he was not able to actually set all of them. P7 was the only one who actually created a rule through the Amazon Alexa’s app, namely a “goodnight” scenario to be activated on a vocal command.

In the imagination exercise, all participants created rules with a structure similar to the **trigger-action** formalism, even if they were not instructed or primed to do it. As mentioned in previous work about trigger-action programming, also in this case participants used triggers *one level of abstraction higher* than direct sensors [5].

We noticed, however, a clear difference in the rules composed by participants who owned a smart speaker and those who did not. For instance, while speaking with the kitchen’s smart speaker to realize the “you would like to turn on the central kitchen light every time you enter that room” personalization goal, the former group composed the following rule, with very minor differences among participants: “*Alexa, every time I walk into the kitchen, turn on the central light.*” These participants specified the room where the rule should happen, even if they knew that rule was set in the kitchen and that the smart speaker was in the kitchen. Conversely, with the same instructions, the latter group composed different rules, such as “*Alexa, turn on the light every time I pass by*” (P3) or “*Alexa, whenever you see someone walk through the door, turn on the main light*” (P6). None of those participants mentioned the kitchen, given that they were speaking with the smart speaker located in that same room. When asked to fulfill the same personalization goal but while speaking with a smart speaker in a different room, all the participants generated rules very similar to the one elicited from smart speaker’s owners. This is possibly due to the participants’ experience with smart speakers: they could have been primed by the current possibilities of the devices, which require them to be very precise in their requests. When the participants do not have an extended experience with smart speakers, instead, they seem to consider that the smart speaker may possess some **implicit knowledge**, e.g., where it is located.

When asked about the possible answers of the speaker after the rule creation, participants

preferred to have an **explicit acknowledgment** that the rule was correctly understood, i.e., by having the smart speaker repeat the entire rule in its own terms, with a confirmation at the end.

Finally, participants commented about what should happen if the IPA did not fully understand the rule. They recognized two main options: *i*) the composed rule has a missing/unclear info (e.g., which lamp to turn on) or *ii*) there is one or more mistakes in the rule. In the former case, participants would accept either an **auto-complete feature**, if possible (e.g., if there is only one lamp that can be turned on), or an **explicit request** from the speaker (e.g., “which lamp do you want to turn on among these?”). In the latter case, participants would use a **trial-and-error approach**, where they rephrase the rule until the IPA understands it correctly.

4. Conclusions

This paper presents a pilot study, in the form of semi-structured interviews, aimed at understanding whether and how end users would converse with a smart speaker to personalize their IoT ecosystem at home. Preliminary results are promising and seem to indicate that end users would be inclined to compose personalization rules by speaking with a smart speaker. Also, they naturally used the trigger-action format to compose such rules, thus confirming the versatility of that formalism. Interesting, non-programmers would create rules with more or less details according to their own experience with smart speakers. The most experienced user would specify several details, even if the IPA could possess that implicit knowledge (e.g., the room where the rule is composed), while less experienced users would “trust” more the smart speaker’s knowledge. In any case, they appreciated the possibility to perform such a relatively complex task with a conversational approach and they would all have an explicit acknowledgement that the IPA correctly understood the rule they composed with different strategies for error recovering or missing/not precise information. They would accept a certain degree of automation in deriving which are the devices they would involve in the rule, while in case of more significant errors they would like to adopt a trial-and-error approach. Future work will expand the pilot study to confirm these results and will consider other approaches for the rule composition, complementary to the vocal interaction, such as the possibility to physically act on the home devices to show the IPA what to do. In addition, a prototype of a conversational agent in the form of a smart speaker will be designed, developed, and evaluated with users in a real IoT ecosystem at home.

References

- [1] T. Ammari, J. Kaye, J. Y. Tsai, F. Bentley, Music, search, and iot: How people (really) use voice assistants, *ACM Transactions on Computer-Human Interaction* 26 (2019). doi:10.1145/3311956.
- [2] L. De Russis, A. Monge Roffarello, Personalizing iot ecosystems via voice, in: *Proceedings of the 1st International Workshop on Empowering People in Dealing with Internet of Things Ecosystems*, volume 2702 of *EMPATHY 2020*, CEUR-WS, 2020, pp. 37–40.
- [3] B. R. Barricelli, E. Casiraghi, S. Valtolina, Virtual assistants for end-user development in the

internet of things, in: End-User Development, Springer International Publishing, Cham, 2019, pp. 209–216. doi:10.1007/978-3-030-24781-2_17.

- [4] M. Manca, P. Parvin, F. Paternò, C. Santoro, Integrating alexa in a rule-based personalization platform, in: Proceedings of the 6th EAI International Conference on Smart Objects and Technologies for Social Good, GoodTechs '20, Association for Computing Machinery, New York, NY, USA, 2020, p. 108–113. doi:10.1145/3411170.3411228.
- [5] B. Ur, E. McManus, M. Pak Yong Ho, M. L. Littman, Practical trigger-action programming in the smart home, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '14, Association for Computing Machinery, New York, NY, USA, 2014, p. 803–812. doi:10.1145/2556288.2557420.