Interfacing participation in citizen science projects with conversational agents

MANUEL PORTELA, Universitat Pompeu Fabra (Barcelona, Spain)

ABSTRACT

This paper assesses the use of conversational agents (chatbots) as an interface to enhance communication with participants in citizen science projects. We based our analysis on the current needs exposed in citizen science literature to assess the opportunities of using them to promote engagement and participation. We found that chatbots are great communication platforms that can help to engage participants as an all-in-one interface. Chatbots can benefit projects in reducing the need to develop an exclusive app while deployed on several platforms. Finally, we establish design suggestions to help citizen science practitioners to incorporate such platforms into new projects. We encourage the development of more advanced interfaces through the incorporation of Machine Learning into several processes.

1. INTRODUCTION

In its simple definition, citizen science (CS) involves laypeople and experts in the development of scientific projects or in the use of scientific knowledge to tackle societal problems (Eitzel et al., 2017; Irwin, 1995; Kullenberg & Kasperowski, 2016). The use of information technologies led to a boost of citizen science projects so that apps and websites encourage remote participation and fast dissemination (Brenton, von Gavel, Vogel, & Lecoq, 2019).

Nowadays, CS is well documented and reported along with multiple disciplines that took advantage of it to extend the scientific practices worldwide. However, CS presents challenges related to motivating, involving and keeping citizens engaged within the projects. CS literature reports that communication failures with the participants, lack of clear goals or ambiguous instructions can lead to a decrease in participation (Schröter et al., 2017).

Strategies and methods for reducing such failures include gamification strategies (Preece, 2016), co-design and co-creation processes (Shaw, Draux, García Martín, Martin, & Bieling, 2017). To contribute to creating new strategies for engagement, we propose to explore the potentialities of

using chatbots' artificial intelligence capabilities (European Commission, 2019; Independent High-Level Expert Group on Artificial Intelligence, 2019) as new means of interfacing between experts and citizens. In this article, we evaluate the requirements and needs of CS that can be addressed using conversational agents for engaging participants. Consequently, we articulate the following research question: **How can chatbots help enhance communication strategies for engaging participants in CS projects?**

1.1 The quest for useful conversational agents

Opportunities for using artificial intelligence (AI) go beyond the frequently implemented supervised machine learning and automation algorithms. Instead, through adaptive behaviour, AI can also improve experiences enhancing the collaboration and co-production between humans and non-humans (Portela & Granell-canut, 2017). Bots can be described as AI systems with automation capabilities for assisting humans in specific tasks efficiently (Brandtzaeg & Følstad, 2017). Moreover, conversational agents can be understood as bot systems using text-to-speech or chat interfaces (chatbots). In this universe, we can find different types of bots based on their architecture, where the most common follow a series of rules or patterns to provide the correct answer to users' requests (Surmenok, 2016).

A bibliometric study about conversational agents (Io & Lee, 2018) shows its increasing trend as a research topic that calls various disciplines' attention. Since 2015, the interest grew exponentially with a manifest presence in computer science and engineering. However, the authors highlight the fragmentation of use cases and real applications, which need more specialisation.

Experiments using virtual assistants are recurrent in the AI and HCI research fields, proving different opportunities to create affection (Partala, Surakka, & Lahti, 2004) and user engagement (Goebel, 2012). Recently, there have been various attempts to survey approaches and perspectives to consider chatbots as companions. Paikari & Van Der Hoek (2018) set up a six-category classification regarding its software development. In that sense, the challenges of ML algorithms development become a bottleneck to generate natural conversations.

To avoid mistaking with real-life conversational capabilities in chatbots, Klopfenstein et al. (2017) suggested level-down the expectations of users by using the term 'botplications' instead of 'chatbots' and focus on enhancing the actual capacities of providing guidance to specific tasks and taking advantage of using parallel conversations, notifications and contextual updates that are currently supported by chatbot apps. However, Piccolo, Mensi & Alani (2019) agreed that beyond technical challenges to provide capabilities to chatbots to behave as expected, there is much to be done to understand users' perceptions, expectations and contexts of use. Recent publications have shown novel findings in usability and user apprehension to these new interfaces and technologies (Luger & Sellen, 2016). In that direction, the book 'Studies on Conversational UX Design' (Moore, 2018) demonstrates the standardisation of chatbots as an interface and schematises the first guidelines to be followed. For that reason, we think that chatbots are one of many other tools that should be considered for deploying CS projects.

1.2 Citizen participation and e-government chatbots

A prominent field to explore the use of conversational agents is the public sector and governmental services oriented to citizens. Similarly, laypeople and non-experts are the typical types of users in CS. On the one hand, chatbots were tested in enhancing public services and communication. On the other hand, we account for the use of chatbots for public participation processes. Both sides are attempts to close the gap between the administration and civil society.

Public services can be easily derived or facilitated by chatbots. 'Bot-based transaction services may include payment of local taxes, issuance and electronic delivery of certificates and permits, as well as proactive notifications on all services (e.g. notification on the expiration of permits)' (Androutsopoulou, Karacapilidis, Loukis, & Charalabidis, 2019, p. 7).

A step beyond citizen service is establishing two-way communication, letting the bots facilitate citizens' participation in public surveys, crowdsourcing or decision-making processes. The advantages of using chatbots as facilitators are yet to be proven worthy. Tavanapour, Poser, & Bittner (2019) found deficiencies with chatbots that achieved to discuss complex ideas. For example, in one of such experiments, 'the participants, who received a wrong answer, submitted ideas on a lower level of completion and elaboration compared to other users' (Tavanapour et al., 2019, p. 12). In Lieven et al. (2021), a chatbot helped a participatory urban project were the chatbot was considered valued for online co-creation. However, the complexity or lack of congruence in participants' submitted annotations complicated the evaluation.

In some way, the use of chatbots looks promising to include citizens in-the-loop of the administration channels and public life. For example, the project Talkin'Piazza takes advantage of a chatbot to use public objects to 'raise urban issues to the citizens, by requiring their contributions in the form of local memories or suggestions and ideas; […] and, bring information to the citizens in a more proactive way' (Celino, Calegari, & Fiano, 2016, p. 4). Another example is the Uchaguzi chatbot, used in Kenya Messenger to help people submit reports of violence and misconduct experienced in the polling stations (Piccolo, Roberts, Iosif, & Alani, 2018). In the field of health service, chatbots are helping patients with depressions by getting help in an anonymous way (Skjuve & Brandtzæg, 2018). Chatbots also provided practical guidance in promoting collaborative work and enhance the debate in the context of sensitive areas such as assessing recidivism risk (Van Berkel et al., 2019).

It is not our intention to expand the list of examples but to reflect the increasing interest in using this technology to enhance communication and engagement with citizens. In these fields, bots have been proved valuable in different tasks and scenarios. However, because many of the actual examples are short-length projects and not stable implementations, more extended tests are required to prove their utility.

1.3 Citizen science and engagement

Citizen science (CS) is a prominent and multidisciplinary field (Kullenberg & Kasperowski, 2016). The context of this Special Issue, the European Union's COST Action CA15212 on

Citizen Science[1], demonstrates its importance for the scientific and social contexts. Several research projects (i.e. Making Sense[2], D-Noses[3], CitieSHealth[4], Growth Observatory[5]) are conceived as CS practices because take advantage of engaging citizens to take part in the research process, to co-create and co-design the process of raising matters of concern and, ultimately, of acquiring scientific knowledge (Vallabh, Lotz-Sisitka, O'Donoghue, & Schudel, 2016).

In this regard, increasing the engagement level promotes more extended participation and a more substantial contribution from participants. Through the CS literature, we can find different typologies of CS projects concerning the level of engagement of their participants. Therefore, projects can be classified as volunteer, crowdsourced, contributive, collaborative, and co-created (Eitzel et al., 2017; Haklay, 2013; Shaw et al., 2017; Shirk et al., 2012). However, many studies on CS projects agree that fluid communication with participants is one of the main issues that face project organisers in order to keep them engaged along with the projects and to reinforce their motivations (Peters, Zhan, Schwartz, Godoy, & Ballard, 2017; Skarlatidou, Hamilton, Vitos, & Haklay, 2019). Other needs for ensuring a successful engagement and participation are related to generating trust, having clear goals and objectives, promoting knowledge exchange among participants, understanding the importance of their contribution, among others (Cvitanovic et al., 2018; Schröter et al., 2017).

Understanding engagement is a complex goal in CS. For Haklay (2019), engagement is considered, together with knowledge, key concepts to reduce the ambiguity of the term 'participation' for evaluating projects, where both concepts are tangled in different ways. Therefore, a higher level of engagement is not always desired and should be evaluated in each project's needs and goals, whereas the projects' design should be aligned with who will be engaged (Spiers et al., 2019).

In that sense, chatbots are a promising user interface for enhancing collaboration, reducing complexity in the guidance of knowledge, and communicating with participants, among other activities. Not all chatbots and AI behaviours are desired or needed in every situation. That is because the interaction between citizens, researchers and other stakeholders in a citizen science project should not be just a matter of data collection. The experience of getting involved in such projects is an embodiment and affective encounter with the object of research - i.e. natural environment, animal species, urban pollution (Sharma et al., 2019). Moreover, we argue that it is in that aspect where chatbots can stand out compared to other communication tools.

Nevertheless, a few examples of chatbot use exist in the field of citizen science in recent years. Ethnobot was developed to help volunteers make sense of their contribution and knowledge improvement (Tallyn, Fried, Gianni, Isard, & Speed, 2018). In another example, a chatbot helps

---

[1] Citizen Science Cost Action - CA15212. Available at: https://www.cs-eu.net (accessed: 31/12/2020)
[2] Making Sense - H2020-ICT-2015. Available at: http://making-sense.eu (accessed: 31/12/2020)
[3] D-NOSES - H2020-SwafS-2017-1. Available at: https://dnoses.eu (accessed: 31/12/2020)
[4] CitieS-Health - H2020-SwafS-2018-1. Available at: https://citieshealth.eu (accessed: 31/12/2020)
[5] Grow Observatory - H2020-SC5-2015. Available at: https://growobservatory.org (accessed: 31/12/2020)

to augment the participants' collaboration on data contribution. A chatbot can be a strategic part in collaboration with other media or social networks. Thus, in the FloodBook project in Italy, the strategy was a multi-source approach is to achieve a considerable number of structured data, involving as many users as possible (Isacco et al., 2018). In linguistics, another Italian project called DialettiBot helped crowdsource dialectic data from different parts of the country (Sangati, Abramova, & Monti, 2018). Beyond how efficient and straightforward other UIs can be, there is much to explore within the chatbot development to achieve this embodied and emotional bonding that could support citizen science projects (Zamora, 2017).

In the following sections, we describe our methodology to address this question. In the first place, we explain the methodology used for an exemplary case where we deployed a chatbot platform. Later, we compare the results of the study case with the requirements for successful CS projects presented by Kieslinger et al. (2018). Lastly, we describe a set of rules and suggestions to use such conversational agents for enhancing engagement in citizen science projects.

2. **METHODOLOGY**

During 2016, we developed a project within our university to understand how people are engaged with the use of chatbots and what are the factors that keep people motivated to continue such interaction.

We conducted the experiment with 13 participants who can be classified in 3 groups by age: three participants in the 18 to 25 years' group, eight in the range 26-36 years, and two between 37 and 50. All of them have an undergraduate degree and mostly postgraduate studies (10 of 13, with a background in Computer Science, Humanities and Psychology). In terms of gender, nine of them were women.

We designed two chatbot 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[6] for both agents because it allowed us to have a familiar interface with a simple implementation.
As noted in Partala (2004), conversations are preferred at a personal proximity level, influencing participants' experienced affective valence and intimacy with the agent. We designed the behaviour of agents to reach enough intimacy by using informal language. For the server-side, we found that besides the raising of platforms for creating conversations through Machine Learning algorithms[7], few of them supported other languages than English at the time of developing our

---

[6] Telegram Bot API. Available at: https://core.telegram.org/bots/api
[7] At the moment of the project some platforms were starting, i.e. PandoraBots (Available at: http://www.pandorabots.com), Recast (Available at: https://recast.ai), Aiva (Available at:http://kengz.me/aiva/), Howdy (Available at: http://howdy.ai), or Hubot (Available at: https://hubot.github.com). However, nowadays more complete and multilanguage platforms are available from big corporations such as IBM's Watson (https://www.ibm.com/watson/how-to-build-a-chatbot) and Microsoft Bot Framework (https://dev.botframework.com)

experiment. 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 processor using the open-source software Program-O (PHP). AIML stands for Artificial Intelligence Mark-up Language, an XML-based format for specifying natural language software agents, based on the original ALICE project (Wallace, 2009). It is a widely-used system to develop and configure conversational agents quickly. Its simplicity comes at the expense of some limitations for specifying sophisticated conversations to exploit natural language processing capabilities fully. We based our choice on successful early experiments and its multi-lingual support, which fit the needs of the current experiment. We designed the conversations to make them appear natural and attractive 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, allowing the chatbot 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 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 (Okamoto, Yang, & Ishida, 2001), consisting of a person pretending to be the bot. Wizard of Oz is a frequently used method for researching about conversational interfaces and dealing with human dialogues. This method allows to test complex utterances following more human-like conversations when Natural Language Processing developed capabilities are insufficient (Dahlbäck, Jönsson, & Ahrenberg, 1993). For a person to operate the bot from behind, we designed a web interface in NodeJS to interact with the Telegram Bot API to enabling a chat platform through which an operator could interact with the participants. For obvious reasons, we did not tell participants about the actual implementation behind CH2, and they believed the bot acted autonomously. To provide the same style of answers, the operator received guidance with conversation structures and behavioural rules similar to those followed by CH1. We decided to use this mechanism because the human operator behind the system can manage the conversation without altering the context, and we wanted to improve the bot's corpus of answers. That is, giving freedom of improvising more human-like conversations. 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 participants' request. Due to our limitation in developing this study longer, we do not pretend to study a long-term engagement but understand how an initial conversation can increase confidence and interest in this kind of tools.

2.1 Data collection and analysis

In literature, we found different ways of measuring usability, engagement and affection in interactions with conversational agents like Hubal et al. (2008) and Partala et al. (2004), but only provide partial results. We developed a set of interviews with our participants in an individual session with each of them. First, we developed an unstructured interview to assess the usage and experience with the participants. Second, we let them try each of the two chatbots for 10 minutes as maximum. Third, we developed a second interview to know their experience and know if their perception and willingness to use them changed after the test.

After the experiment, we recovered logs captured in the server from the conversations with the chatbots. We carefully studied them using Conversational Analysis (CA), a method approached by Sacks and Heritage (Heritage, 1991; Sacks, 1992; Sacks, Schegloff, & Jefferson, 1974) (Sacks, 1992; Sacks, Schegloff, & Jefferson, 1974) to understand the structure of conversations by systematically analysing each transcription. Additionally, we used the time-log of conversations to analyse the response time from both sides, giving us some insights and feedback on how the system and participants have behaved.

Our analysis continued by building a Thematic Network Analysis (TNA) (Attridge-Stirling, 2001) from the interviews' transcriptions. TNA is a tool that helps to build a 'web-like network as an organising principle and a representational means, and it makes explicit the procedures that may be employed in going from text to interpretation' (Attridge-Stirling, 2001, p. 388). For building conceptual and sensitive insights of the textual data, the data is organised into three themes that go from the basic premises to more global themes. Furthermore, this method was also used for similar purposes in Luger and Sellen (2016).

2.2 Further analysis in the context of Citizen Science

After our experiment, we explored the work of previous publications to understand the possibilities of chatbots to contribute to CS projects. For that, we took as a reference the needs in the CS technologies spotted by Skarlatidou et al. (2019) in an extensive literature review.

The authors found more than 20 key design features distributed in 6 categories, where technologies can be helpful in CS projects. We recovered these features to study where can chatbots can contribute and facilitate the implementation of such features. Last, we explore the next steps for implementing more advanced support with this and other machine learning technologies. Therefore, this article's main contribution is a consolidation of positive and negative aspects of using chatbots that could benefit or harm a communication strategy in any citizen science project.

3. **FINDINGS**

Since we carried out our experiment in 2016, other studies worked with chatbots to assess its performance and usefulness with similar results in terms of technical development (Amato et al., 2017). Moreover, the citizen science approach became an emergent trend, originated by the VGI and other participatory GIS communities (Eitzel et al., 2017). We wanted to understand how likely a chatbot can be useful for CS projects, but developing one was out of our scope. CS projects usually take place in urban or rural environments, observing natural or social phenomena. For that reason, we added some questions in our questionnaires regarding the platform's use in the context of urban spaces because we wanted to know if they would consider it as a possibility.

In a previous publication, we reported findings related to the quantitative analysis of this experiment (Portela & Granell-canut, 2017). One main finding was related to the agency of bots in the conversation. Some indexical expressions (Lynch, 1994) that appeared during the conversations were not interpreted by the bot and generated confusions; for example, in the case of using words like *here* or *us*, these terms became confusing for participants since they did not have a point of reference, and because of the invisible nature of the system (Gajendar, 2016). Contextual understanding is essential if Citizen Science projects will make use of conversational aid, including geographical references. Designers of conversations should be aware of these blind spots in common utterances. For linguistics, nouns are less critical than a well-referenced context through indexical expressions. Our human capacity of interpretation in an ambiguous conversation can be underrated. Dealing with uncertainties is part of our habitual behaviour and is difficult to replicate in a machine.

However, in this article, we extend some additional analysis over our raw material from the study by following the needs of CS projects giving significant importance to engagement and facilitation. Social cues were more useful for this purpose. Expressions that generate more emotional recognition, recall and persuasiveness have been reported useful in a chatbot conversation (Ruijten, Midden, & Ham, 2016).

Our focus in this article is to understand the participants' perception of chatbots and identify key points to engage them and the negative aspects of their interaction. This goal was mainly addressed in a qualitative approach through the interviews. We analysed the interviews through a Thematic Network Analysis, highlighting the following findings (see Figure 2).

Figure 2. *Thematic Network (TNA) of themes touched by participants*

Following the methodology, we aggregated the findings until having only two global themes, one referred to the way they perceived the conversation during the test, and another based on the opportunities of use and the chatbot features. Both themes are correlated with how participants made sense of using the chatbots as they see them through the social imaginaries (i.e. science fiction movies and advertising). Furthermore, both themes are interrelated and embodied during any interaction.

For *use feedback and perception,* we found three primary themes. All three are related to how participants felt more or less engaged regarding the interaction, and more importantly, how they aligned themselves to specific behaviour. It was evident that errors and unexpected or random behaviour caused disappointment and distrust. Instead, social cues and small talk provoked more confidence and trust. Familiarity with other chatbots or conversational interfaces like Siri or Alexa was crucial for them to accept and understand the study's limits and opportunities.

Regarding the user expectation, we noticed that before and after the experiment, all participants described new desired functionalities for specific and routine tasks or new uses of these technologies. One participant said, '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'. In some way, participants imagine that a conversational interface is more appealing to diverse publics. Even more, it seems that they expected some kind of custom conversation that fits their needs and understand their desires. For example, 'think it could be smarter if people can access to something different, an answer that is not the same for everyone'. Of course, in the current state-of-the-art, such an idea is not feasible and is something that can disappoint them in case of failing their expectations. However, using social cues and other linguistic tricks, it could be easy to simulate some kind of personal treatment.

Another expectation was the value of using chatbots for specific tasks. One participant said, 'it offers a solution that you already have; it's like I feel more comfortable doing it by myself'. Nevertheless, 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. Our chatbot did not have a specific goal, and neither was developed to perform a specific task that could be evaluated. However, this is something to consider when deploying a chatbot for specific purposes. Designers should critically consider if the chatbot is facilitating or adding value to a process in a way that would be impossible by other means.

For the topic *system future opportunities,* we grouped data into three primary themes. These themes respond to how the participants see themselves engaged or not by using them in the future. For example, some participants imagined using chatbots only for simple tasks. Others came up with valuable ideas in the context of urban space (something that we asked explicitly). Ideas were like the opportunity to talk to urban furniture and public offices without dealing with people or knowing contextual information of the city and its heritage.

These answers were interrelated with the experience they had and the trust that was transmitted in such interaction. We found that participants already had a prefiguration on how to deal with virtual and conversational agents. Some participants were sceptic about the opportunity to use chatbots for various tasks and remained so after the experiment, suggesting that the expected functionality was more oriented to a rational aspect of its use. On the contrary, other users thought that chatbots had learning capabilities, have a more advanced adaptation that they actually had, and solve problems faster or efficiently.

Other answers were related to data and privacy questions and the learning capacity that the chatbot presents as an opportunity. 'I would not give any advanced task to do' spotted one participant who had low confidence in the actual capabilities of chatbots. 'But it would be nice to give them simple and repetitive tasks', she added. Engagement with chatbots was also studied in Zamora (2017), which was consistent with our findings. It is linked with trust, confidence and the fulfilment of user expectations.

Another finding is the socialisation of chatbots. As chatbots are accessed through personal computers or smartphones, they are embedded in the context of daily life activities. Some participants acknowledge that they shared their conversations or played with conversational agents like Siri o Alexa with other people. Sometimes, they said they used it just as a game to see the system's response or if it can make jokes. How a conversation is socially shared can be a benefit for two reasons. First, a chatbot can enhance collaboration to develop a task between different people. That is something desired in many CS projects. Second, a socialisation aspect would help communicate with the participants and their daily lives, encouraging them to participate in different contexts and let them promote their participation with their colleagues, friends, and families. Lastly, designers should consider not only the interaction with the chatbot in a social context but also interacting with the entire ecology of artefacts that they are familiar with (Sørensen & Kjeldskov, 2014).

4. SUGGESTIONS FOR ENGAGING CITIZENS

Our experiment's findings open the door for new prospects oriented to design specific relations between humans and virtual agents. We spotted opportunities to develop systems that engage people and consider them as companions, guides or facilitators in complex scenarios. Implementing chatbots can help participants reduce the need for specific skills and help augment participation (Haklay, 2019).

However, we understand that such a platform's development does not come without costs regarding time and resources. Even if we encourage adopting chatbot as an interface, developers and designers should consider the positive and negative aspects beforehand.

Additionally, some conditions might be considered in advance. First, we believe that easy access to people's smartphone and using notification systems to communicate with them is a simple way of being present during any project's lifecycle. Second, users may be familiar with chatbots and the idea of conversational agents in advance, or at least, having a popular reference to naturalise its presence, understand its limitations and the way it communicates. In other words, in the current state of the art, a conversation with a chatbot is considerable different from a person-to-person conversation. Third, technological and model building of chatbots must be assured and tested long before deploying them in a real scenario. Any project should develop enough user experience (UX) research to avoid failures and risking expectations and motivations in the first instance.

Additionally, our findings suggest that any chatbot system should comply with several principles or rules that various authors already pointed out (Amato et al., 2017; Moore, 2018; Piccolo et al., 2019; Tavanapour et al., 2019). Chatbots should manage the unexpected by keeping the context, the coherence, and the conversation flow; have the capacity of summarisation but also to repair, repeat or paraphrase; recognise user utterances and preferences; provide affective, emotional and social cues; may be entertaining; must be transparent and provide trustfulness.

For any implementation, it should be decided which style of conversation is better for the project: system-centric, content-centric, visual-centric or conversation-centric (Moore, 2018). Each style delineates more efficient or long-explaining conversations, more natural or command-oriented queries, responses using visual resources or using turn-taking approaches in the conversation. As we found, users tend to expect a more efficient and direct answer. However, engagement is rarely achieved with simple queries or requests. Instead, the setup of a conversation that requests personal point of views, opinions, personal details or invitations to socialise augments the possibilities of long-term engagement. Therefore, we suggest designing chatbots that can keep simple dynamics and raise the level of explanation if it is requested.

Another way to explore the opportunities for CS projects is to approach the needs that were recognised in the literature. So we recovered the six categories defined by Skarlatidou et al. (2019) and adapted them to the use of chatbots.

Basic features and design recommendations

Skarlatidou et al. point to general categories on a website. In a chatbot interface, there are no sections like in a visual interface. Rather, the conversations are organised chronologically, allowing the user to browse previous conversations if needed. However, common chat commands for simple information and queries should be available (i.e. /start, /help, /info, /about). Here we suggest having commands that allow users to orientate themselves or to restart the conversation in the case of need. Also, commands related to data management (i.e. get, delete, upload, share) and the status of the projects would be useful.

Authors refer in this section to registration and login issues. Whether a chatbot is used within other apps, like Telegram or Facebook, it would request additional personal information to know participants better. We agree with the authors' recommendations about not relying on social network logins, even though we recommend keeping it simple, avoiding overlapping information requests.

Regarding tutorials, chatbots are essential to deploy tutorial or step-by-step features. We suggest the use of chatbots as guides or facilitators as a central feature. However, the ability to visualise or collect data can complement the tutorials, creating an all-in-one platform.

Lastly, regarding usage and interface by designing a chatbot. The cultural and environmental differences are reduced when chatbots are deployed in platforms that are already familiar to users. We encourage to avoid developing ad-hoc apps but using existing platforms to reduce misunderstandings and promote engagement. Even in remote places, chatbots servers can be deployed through SMS services, although the conversation should be reduced to the minimum because of SMS's transaction costs.

4.1 Design for communication functionality

Chatbots are communication platforms by definition, providing a unique advantage because of their real-time availability and data exchange. Possibilities extend beyond simple chat conversations when other smartphone capabilities (using the camera, the sensors or the touchscreen) or server-side processing (data aggregation or analysis) are used. Chatbots can also

be used as an interface between project coordinators and participants through regular chat interaction, as a notification system for project updates or reminders, as a facilitator for collaboration or, using it as a research diary or note-taking activities.

However, it is limited in its interface. For example, maps are less friendly than in an ad-hoc application and tend to be perceived as too small for input in the mobile phone (Lieven et al., 2021). Nevertheless, location can be shared through various mechanisms like the geolocation service what3words (https://what3words.com/) or messengers' location sharing features. There are always opportunities to generate flows of communication by combining apps. Thus, we encourage designers to think out-of-the-box and extend the capabilities of a simple interaction through a smartphone to a multi-modal interaction with multiple physical objects (i.e. in the city, in the office, or the countryside) with the chatbot as an interface or a medium.

4.2 Design for data collection

Chatbots can also help to define more precise data in real-time when results are not transparent or incomplete. In opposition to other types of interfaces, chatbots can be designed to insist or rephrase data request, encouraging participants to provide more accurate data.

Other types of data can be shared through the most common chat platforms, like photos, location, video or audio recordings. Also, more comprehensive qualitative data can be added in the form of comments and conversational logs. Another option is Telegram add-ons that allow generating in-app features like surveys and polls. All this data can be smoothly requested in a conversation, with more personalised feedback and sensitivity to user utterances.

A weakness in chatbot platforms based on Telegram or Messenger applications is the need for internet network to work. This limitation should be considered when deploying a project. Depending on the context, maybe it will be necessary to have an alternative way of data collection or to use offline based platforms like SMS or in-platform chatbots that can save information in the device itself.

4.3 Design for data processing and visualisation

The other side of data collection is data provision. Chatbots can also provide contextual information from servers. For example, a participant can request data gathered in a particular space, where the chatbot can answer summarised information or a cartographic image with data points. Sometimes chatbots cannot provide detailed visualisations of data because their interfaces are optimised for text or images rather than complex interactions. Even more, downsizing the data to be transferred is also desired to reduce costs and infrastructure. For that reason, the power of pre-processing the data on the server-side can be helpful for contextual, in-the-wild interactions.

4.4 Gamification features

Beyond affective, emotional and social cues and other conversational tricks, chatbots can help to gamify any project. Gamification consists of adding game-based elements (e.g. story, challenges,

feedback or rewards) and can help to increase spatial learning and to learn scientific knowledge faster (Devisch, Poplin, & Sofronie, 2016). The use of mobile notifications or participant-to-participant chat can be a form that can create scenarios for competition, team collaboration or individual contribution to a project.

At this point, we do not follow the authors' recommendations about avoiding gamification for a CS project when it includes a chatbot. Gamification apps indeed showed few improvements in data contribution. However, our findings showed that socialisation leads to better engagement with the chatbots and the purpose of using them. Even with simple game strategies, the focus should be oriented to encourage participants to exchange knowledge, feelings and experiences with each other. That would motivate them to create an ecosystem that can improve teamwork and collaboration.

4.5 User privacy issues

This point is more than necessary in the use of chatbots. In our experiment, participants showed signs of distrust in what chatbots can do and how they manage the data they gather. Not all the participants understand how chatbots work, how data is managed, and its capabilities.
Therefore, it is crucial to set clear rules at the beginning of any participation, placing the information available all the times. Maybe, a specific command can show information or FAQs about privacy, security measures and data management. A link to a website with an extended explanation is also suggested.

5. CONCLUSIONS

We found that conversational agents promote the socialisation of activities and reinforce engagement, opening the opportunities to keep participants motivated while they are oriented and trained through the project. In short, we suggest that treating conversational agents as companions for exploration and discovery would contribute to personalised feedback about citizen science projects. In a few words, chatbots can help to put participants in-the-loop with less dependence on direct communication. Even more, chatbots can facilitate and encourage cooperation between participants.

The use of ICT for data management in CS is essential and has been extended to several projects to let participants develop observation tasks, monitor data collection and data classification. Data management is central for any CS project, but also data analysis and visualisation. In that sense, automation and collaboration with Machine Learning processes are being tested and implemented systematically (Starkey et al., 2017).

Connecting multiple databases as a knowledge base for chatbot and process it with ML capabilities can take us to a new generation of extended facilitators. Androutsopoulou et al. (2019) explored different ways that AI-guided chatbots can solve complex tasks and populate information in both ways. On the one hand, the multiple databases, on the other hand, the user.

This approach goes far beyond simple guidance chatbots based on commands and conversational rules. It can provide contextual synthesised data, data optimisation, image and metadata classification, spatial navigation, and smart guidance. Without going further with more examples, the exploration of ML can be enhanced in terms of UX.

Chatbots, when properly designed, can mediate the development to deliver contextual, real-time and straightforward feedback. Nevertheless, for accomplishing an engaging and successful conversational interface, our findings suggest that designing and developing a chatbot requires a participatory approach and several iterations where dialogues and features achieve users' expectations and provide sufficient guidance towards the purpose of the CS project in play. The experiment presented in this article is exemplary for being reproduced to test and evaluate a chatbot with few participants in the context of new development.

Beyond effectivity and engagement, chatbots can be considered simple platforms for continuous development, testing and deployment in long-term projects. That way, developers and researchers do not need to rely on complex visual interfaces that can delay research projects and require maintenance. There is a need for continuing the exploration and experimentation in more extended implementations, more usability tests and evaluations. Not without limitations, our findings and guidelines can be the background for this continuous research in conversational agents and citizen science.

6. LIMITATIONS TO OUR WORK

Because our initial case study was not in the context of any SC project, we cannot give empirical evidence of the real impact of chatbot use in a CS project. Nevertheless, carrying a CS project requires many resources and planning. Our work opens the door to such new experiments and research. A broader study about the value of chatbots in the urban context would be needed together with vast knowledge and experience with current technological state-of-the-art.
We acknowledge that our study with chatbots only 10 minutes for each participant cannot reflect a long-term relation, neither a real scenario. However, our findings contribute to a better understanding of its potential benefits. We encourage more "in the wild" research and more extended studies in future research. Also, we would like to see and measure the impact of a conversational platform being used inside a CS project. We acknowledge that current projects seek to include chatbots in their strategies, and their findings can enlighten our contribution.

7. ACKNOWLEDGEMENTS

8. REFERENCES

Amato, F., Marrone, S., Moscato, V., Piantadosi, G., Picariello, A., & Sansone, C. (2017). Chatbots meet ehealth: Automatizing healthcare. *CEUR Workshop Proceedings*, *1982*, 40–49.

Androutsopoulou, A., Karacapilidis, N., Loukis, E., & Charalabidis, Y. (2019). Transforming the communication between citizens and government through AI-guided chatbots. *Government Information Quarterly*, *36*(2), 358–367. https://doi.org/10.1016/j.giq.2018.10.001

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

Brandtzaeg, P. B., & Følstad, A. (2017). Why people use chatbots. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 10673 LNCS, pp. 377–392). https://doi.org/10.1007/978-3-319-70284-1_30

Brenton, P., von Gavel, S., Vogel, E., & Lecoq, M.-E. (, 2019). Technology infrastructure for citizen science. In *Citizen Science* (pp. 63–80). https://doi.org/10.2307/j.ctv550cf2.12

Celino, I., Calegari, G. R., & Fiano, A. (2016). Towards Talkin'Piazza: Engaging citizens through playful interaction with urban objects. In *2016 IEEE International Smart Cities Conference (ISC2)* (pp. 1–5). IEEE. https://doi.org/10.1109/ISC2.2016.7580809

Cvitanovic, C., van Putten, E. I., Hobday, A. J., Mackay, M., Kelly, R., McDonald, J., … Barnes, P. (2018). Building trust among marine protected area managers and community members through scientific research: Insights from the Ningaloo Marine Park, Australia. *Marine Policy*, *93*(May), 195–206. https://doi.org/10.1016/j.marpol.2018.04.010

Dahlbäck, N., Jönsson, A., & Ahrenberg, L. (1993). Wizard of oz studies-why and how. *International Conference on Intelligent User Interfaces, Proceedings IUI*, *Part F1275*, 193–200.

Devisch, O., Poplin, A., & Sofronie, S. (2016). The Gamification of Civic Participation: Two Experiments in Improving the Skills of Citizens to Reflect Collectively on Spatial Issues. *Journal of Urban Technology*, *23*(2), 81–102. https://doi.org/10.1080/10630732.2015.1102419

Eitzel, M. V, Cappadonna, J. L., Santos-Lang, C., Duerr, R. E., Virapongse, A., West, S. E., … Jiang, Q. (2017). Citizen Science Terminology Matters: Exploring Key Terms. *Citizen Science: Theory and Practice*, *2*(1), 1. https://doi.org/10.5334/cstp.96

European Commission. (2019). Building Trust in Human-Centric Artificial Intelligence. Retrieved from https://ec.europa.eu/digital-single-market/en/news/communication-building-trust-human-centric-artificial-intelligence

Gajendar, U. (2016). Empathising with the smart and invisible. *Interactions*, *23*(4), 24–25. https://doi.org/10.1145/2935195

Goebel, R. (2012). Intelligent Virtual Agents. In Y. Nakano, M. Neff, A. Paiva, & M. Walker (Eds.), *12th International Conference, IVA 2012, Santa Cruz, CA, USA, September, 12-14, 2012. Proceedings* (Vol. 7502). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-33197-8

Haklay, M. (2013). Citizen science and volunteered geographic information: Overview and typology of participation. In *Crowdsourcing Geographic Knowledge: Volunteered Geographic Information (VGI) in Theory and Practice* (Vol. 9789400745, pp. 105–122). https://doi.org/10.1007/978-94-007-4587-2_7

Haklay, M. (2019). Participatory citizen science. In *Citizen Science* (pp. 52–62). https://doi.org/10.2307/j.ctv550cf2.11

Heritage, J. (1991). *Garfinkel and Ethnomethodology. Garfinkel and Ethnomethodology*. Wiley.

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

Independent High-Level Expert Group on Artificial Intelligence. (2019). *Definition of AI: Main Capabilities and Disciplines*. Retrieved from https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines

Io, H. N., & Lee, C. B. (2018). Chatbots and conversational agents: A bibliometric analysis. *IEEE International Conference on Industrial Engineering and Engineering Management*, *2017-Decem*, 215–219. https://doi.org/10.1109/IEEM.2017.8289883

Irwin, A. (1995). *Citizen science: A study of people, expertise and sustainable development*. *Routledge*. https://doi.org/10.1177/017084069701800109

Isacco, S., Claps, P., Grasso, S., Ferrari, E., Guercio, M. B., Musumeci, R. E., … Laio, F. (2018). Floodbook: a Social Platform for Flood Hydrology, *3*, 941–931. https://doi.org/10.29007/fb6d

Kieslinger, B., Schäfer, T., Heigl, F., Dörler, D., Richter, A., & Bonn, A. (2018). Evaluating Citizen Science: Towards an open framework. In *Citizen Science - Innovation in Open Science, Society and Policy* (p. in press).

Klopfenstein, L. C., Delpriori, S., Malatini, S., & Bogliolo, A. (2017). The rise of bots: A survey of conversational interfaces, patterns, and paradigms. *DIS 2017 - Proceedings of the 2017 ACM Conference on Designing Interactive Systems*, 555–565. https://doi.org/10.1145/3064663.3064672

Kullenberg, C., & Kasperowski, D. (2016). What is citizen science? - A scientometric meta-analysis. *PLoS ONE*, *11*(1), 1–16. https://doi.org/10.1371/journal.pone.0147152

Lieven, C., Lüders, B., Kulus, D., & Thoneick, R. (2021). Enabling Digital Co-creation in Urban Planning and Development. In *Smart Innovation, Systems and Technologies* (Vol. 189, pp. 415–430). https://doi.org/10.1007/978-981-15-5784-2_34

Luger, E., & Sellen, A. (2016). "Like Having a Really Bad PA." In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16* (pp. 5286–5297). New York, New York, USA: ACM Press. https://doi.org/10.1145/2858036.2858288

Lynch, M. (1994). *Scientific practice and ordinary action*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511625473

Moore, R. J. (2018). *Studies in Conversational UX Design*. https://doi.org/10.1007/978-3-319-95579-7

Okamoto, M., Yang, Y., & Ishida, T. (2001). Wizard of Oz Method for Learning Dialog Agents. In M. Klusch & F. Zambonelli (Eds.), *Cooperative Information Agents V: 5th*

InternationalWorkshop, CIA 2001 Modena, Italy, September 6--8, 2001 Proceedings (pp. 20–25). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/3-540-44799-7_3

Paikari, E., & Van Der Hoek, A. (2018). A framework for understanding chatbots and their future. *Proceedings - International Conference on Software Engineering*, 13–16. https://doi.org/10.1145/3195836.3195859

Partala, T., Surakka, V., & Lahti, J. (2004). Affective Effects of Agent Proximity in Conversational Systems. In *Proceedings of the Third Nordic Conference on Human-computer Interaction* (pp. 353–356). New York, NY, USA: ACM. https://doi.org/10.1145/1028014.1028070

Peters, C. B., Zhan, Y., Schwartz, M. W., Godoy, L., & Ballard, H. L. (2017). Trusting land to volunteers: How and why land trusts involve volunteers in ecological monitoring. *Biological Conservation*, *208*, 48–54. https://doi.org/10.1016/j.biocon.2016.08.029

Piccolo, L. S. G., Mensio, M., & Alani, H. (2019). Chasing the Chatbots. In *Studies in Computational Intelligence* (Vol. 342, pp. 157–169). Springer International Publishing. https://doi.org/10.1007/978-3-030-17705-8_14

Piccolo, L. S. G., Roberts, S., Iosif, A., & Alani, H. (2018). Designing chatbots for crises: A case study contrasting potential and reality. *Proceedings of the 32nd International BCS Human Computer Interaction Conference, HCI 2018*. https://doi.org/10.14236/ewic/HCI2018.56

Portela, M., & Granell-canut, C. (2017). A new friend in our smartphone ? Observing Interactions with Chatbots in the search of emotional engagement. In *Proceedings of Interacción '17*. https://doi.org/10.1145/3123818.3123826

Preece, J. (2016). Citizen Science: New Research Challenges for Human–Computer Interaction. *International Journal of Human-Computer Interaction*, *7318*(June), 1–28. https://doi.org/10.1080/10447318.2016.1194153

Ruijten, P. A. M., Midden, C. J. H., & Ham, J. (2016). Ambiguous Agents: The Influence of Consistency of an Artificial Agent's Social Cues on Emotion Recognition, Recall, and Persuasiveness. *International Journal of Human-Computer Interaction*, *32*(9), 734–744. https://doi.org/10.1080/10447318.2016.1193350

Sacks, H. (1992). *Lectures on Conversation. Victoria* (Vol. 1). Wiley. https://doi.org/10.1002/9781444328301

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. In *Language* (Vol. 50, pp. 696–735). Linguistic Society of America. Retrieved from http://www.jstor.org/stable/412243

Sangati, F., Abramova, E., & Monti, J. (2018). DialettiBot: A telegram bot for crowdsourcing recordings of Italian dialects. *CEUR Workshop Proceedings*, *2253*.

Schröter, M., Kraemer, R., Mantel, M., Kabisch, N., Hecker, S., Richter, A., … Bonn, A. (2017). Citizen science for assessing ecosystem services: Status, challenges and opportunities. *Ecosystem Services*, *28*, 80–94. https://doi.org/10.1016/j.ecoser.2017.09.017

Sharma, N., Greaves, S., Siddharthan, A., Anderson, H. B., Robinson, A., Colucci-Gray, L., … van der Wal, R. (2019). From citizen science to citizen action: Analysing the potential for a digital platform to cultivate attachments to nature. *Journal of Science Communication*, *18*(1), 1–35. https://doi.org/10.22323/2.18010207

Shaw, B. J., Draux, H., García Martín, M., Martin, J., & Bieling, C. (2017). Contributions of citizen science to landscape democracy: potentials and challenges of current approaches. *Landscape Research*, *42*(8), 831–844. https://doi.org/10.1080/01426397.2017.1385750

Shirk, J., Ballard, H. H. L., Wilderman, C. C., Phillips, T., Wiggins, A., Jordan, R., … Bonney, R. (2012). Public participation in scientific research: a framework for intentional design. *Ecology and Society*, *17*(2), 29. https://doi.org/10.5751/ES-04705-170229

Skarlatidou, A., Hamilton, A., Vitos, M., & Haklay, M. (2019). What do volunteers want from citizen science technologies? A systematic literature review and best practice guidelines. *Journal of Science Communication*, *18*(1), 1–23. https://doi.org/10.22323/2.18010202

Skjuve, M. B., & Brandtzæg, P. B. (2018). Chatbots as a new user interface for providing health information to young people. In Y. Andersson, U. Dalquist, & J. Ohlsson (Eds.), *Youth and news in a digital media - Nordic-Baltic perspectives* (pp. 59–66).

Sørensen, H., & Kjeldskov, J. (2014). Concepts of Multi-artifact Systems in Artifact Ecologies. In *7th International Conference on Advances in Computer-Human Interaction* (pp. 141–146). IARIA. Retrieved from http://people.cs.aau.dk/~jesper/pdf/conferences/Kjeldskov-C74.pdf

Spiers, H., Swanson, A., Fortson, L., Simmons, B. D., Trouille, L., Blickhan, S., & Lintott, C. (2019). Everyone counts? Design considerations in online citizen science. *Journal of Science Communication*, *18*(1), 1–32. https://doi.org/10.22323/2.18010204

Starkey, E., Parkin, G., Birkinshaw, S., Large, A., Quinn, P., & Gibson, C. (2017). Demonstrating the value of community-based ('citizen science') observations for catchment modelling and characterisation. *Journal of Hydrology*, *548*, 801–817. https://doi.org/10.1016/j.jhydrol.2017.03.019

Surmenok, P. (2016). Chatbot Architecture. Retrieved September 9, 2016, from https://medium.com/@surmenok/chatbot-architecture-496f5bf820ed#.29q6brn2u

Tallyn, E., Fried, H., Gianni, R., Isard, A., & Speed, C. (2018). The Ethnobot. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18* (pp. 1–13). New York, New York, USA: ACM Press. https://doi.org/10.1145/3173574.3174178

Tavanapour, N., Poser, M., & Bittner, E. A. C. (2019). Supporting the Idea Generation Process in Citizen Participation - Toward an Interactive System With a Conversational Agent As Citizen Participation - Toward an Interactive. *Proceedings of the European Conference on Information Systems (ECIS)*, 1–17.

Vallabh, P., Lotz-Sisitka, H., O'Donoghue, R., & Schudel, I. (2016). Mapping epistemic cultures and learning potential of participants in citizen science projects. *Conservation Biology*, *30*(3), 540–549. https://doi.org/10.1111/cobi.12701

Van Berkel, N., Goncalves, J., Hettiachchi, D., Wijenayake, S., Kelly, R. M., & Kostakos, V. (2019). Crowdsourcing perceptions of fair predictors for machine learning: A recidivism case study. *Proceedings of the ACM on Human-Computer Interaction*, *3*(CSCW). https://doi.org/10.1145/3359130

Wallace, R. S. (2009). The Anatomy of A.L.I.C.E. In R. Epstein, G. Roberts, & G. Beber (Eds.), *Parsing the Turing Test: Philosophical and Methodological Issues in the Quest for the Thinking Computer* (pp. 181–210). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-1-4020-6710-5_13

Zamora, J. (2017). I'm Sorry, Dave, I'm Afraid I Can't Do That. In *Proceedings of the 5th*

International Conference on Human Agent Interaction - HAI '17 (Vol. 41, pp. 253–260). New York, New York, USA: ACM Press. https://doi.org/10.1145/3125739.3125766