Designing Chatbots for Crises: A Case Study Contrasting Potential and Reality

Lara S. G. Piccolo
The Open University
Milton Keynes, UK
lara.piccolo@open.ac.uk

Shadrock Roberts
Ushahidi
Nairobi, Kenya
shadrock@usahidi.com

Anna Iosif
Ushahidi
Nairobi, Kenya
anna@usahidi.com

Harith Alani
The Open University
Milton Keynes, UK
h.alani@open.ac.uk

Chatbots are becoming ubiquitous technologies, and their popularity and adoption are rapidly spreading. The potential of chatbots in engaging people with digital services is fully recognised. However, the reputation of this technology with regards to usefulness and real impact remains rather questionable. Studies that evaluate how people perceive and utilise chatbots are generally lacking. During the last Kenyan elections, we deployed a chatbot on Facebook Messenger to help people submit reports of violence and misconduct experienced in the polling stations. Even though the chatbot was visited by more than 3,000 times, there was a clear mismatch between the users' perception of the technology and its design. In this paper, we analyse the user interactions and content generated through this application, and discuss the challenges and directions for designing more effective chatbots.

Chatbot, Community Resilience, Crises, Citizen Engagement, First-time Users, Human-computer interaction.

1. INTRODUCTION

Free and fair elections are the cornerstone of democracy. Yet, in some parts of the world, elections are often accompanied by much conflict and civil unrest. For example, the tension and violence during Kenya's presidential elections in 2007 resulted in more than 1,000 fatalities, and between 300 and 600 thousand people displaced (UN 2008). This scenario triggered the creation of a digital platform called Ushahidi[1], to support democracy and to empower citizens, by enabling them to report violent incidents by email or text messages (Hyman 2014).

Since then, the Ushahidi crowd-sourced and map-based technology has been deployed over 150 thousand times in around 160 countries, thus gathering more than 30 million reports on different crises, including natural disasters, man-made emergencies or accidents, or social and human-rights issues.

In 2017, the Kenyan elections remained to be a complex event, and much attention was given to limiting human rights violations and corruption. A socio-technical initiative called Uchaguzi[2] ("Election" in Kiswahili), based on the Ushahidi platform, was deployed to ensure safety and transparency by collecting and mapping citizens reports of witnessed incidents.

It is estimated that more than 6 million people in Kenya are Facebook users, which constitutes around 13% of the population[3]. As part of the Uchaguzi initiative, Facebook users were invited to report elections-related incidents via a dedicated chat on a Facebook page (Collins 2017). The reports were checked by volunteers, and then added to the freely-accessible Ushahidi platform.

Although the chatbot attracted a high number of curious and probably well-intentioned users, the data generated by these users reveal many challenges concerning the design of this new technology and appropriation by the users.

In this paper, we report our Kenyan pilot experience. The interaction design and logged dialogues between the chatbot and the users are analysed, searching for insights on how people made sense of the technology. The results led to discussions around the design and evaluation of chatbots created with an inclusive purpose, contrasting some potentials with the reality met in the field. These findings offer guidance to other socio-technical initiatives that also aim at exploring

---

[1] https://www.ushahidi.com

[2] https://uchaguzi.or.ke/

[3] http://www.internetworldstats.com/africa.htm

the potential of chatbots to generate a social impact.

The paper is organised as follows: in the next section, the potential of the technology is highlighted with a state of the art report, and some design challenges found in the literature are pointed out. Section 3 describes the Uchaguzi experience, detailing the functioning of the chatbot. After that, the interaction design is analysed by applying an inspection method, and an overview of the conversations between users and the chatbot is presented in Section 4. Section 5 discusses some reflections towards designing and evaluating a chatbot pursuing a social impact. Section 6 presents a reflection on chatbots for crises. Paper is concluded in Section 7.

2. BACKGROUND

Chatbots, also known as digital assistants or conversational interfaces (Dale 2016), are natural-language processing algorithms empowered with intelligence to simulate a human-like conversation. Based on the input of the user, they try to generate responses for engaging users in a dialogue for providing information, executing tasks, or offering services (Rashid and Das 2018; Dale 2016; Folstad and Brandtzaeg 2017).

The popularity of this technology has continuously grown in different platforms (Zamora 2017), both in the commercial domain and "for the good" (Zhou 2017), attracting investments and the attention of technology enthusiasts interested especially in the potential of Artificial Intelligence (AI) mechanisms. Chatbots have been recently created, for instance, for engaging with public health issues, e.g., assisting people to quit smoking[4], or for signing up as volunteers at the United Nations International Children's Emergency Fund (UNICEF)[5] (Zhou 2017).

Textual and dialogue-based user interfaces are not new, since they go back to the early days of computing in the 60's. However, exploring messaging applications to offer services as chatbots is a recent endeavour (Radziwill and Benton 2017). How to make the best of this type of interaction to reach real impact on people's lives is yet to be investigated (Folstad and Brandtzaeg 2017). Studies with this purpose are still rare in the literature.

In (Zamora 2017), the author involved 54 people in the US and India to qualitatively study their perception of chatbots applications using voice and text. Among other results, participants expressed an interest in using chatbots to execute menial tasks, fulfil emotional needs, such as talking to someone to provide motivation or to listen without judging. Some participants found chatbots adequate for sensitive content that is too embarrassing to ask another human about. Participants expected an ideal chatbot to be high performing (fast, efficient, reliable), smart (knowledgeable, accurate and foreseeing), seamless (easy, smooth), and personable (understands me, likeable).

For being based on simple and usual interactive mechanisms, either voice or text, chatbots hold great potential as an inclusive technology, opening doors of digital services to people still unfamiliar with them (Folstad and Brandtzaeg 2017).

2.1. Design challenges

As stated in Folstad and Brandtzaeg (2017), designing chatbots redefines the notion of usability, moving from graphics and interactive mechanisms to a conversation, which is much more dependent on the users' input to lead users to attain their goals.

Looking for success criteria to evaluate chatbots, in Zamora (2017), the author goes back to the virtual agents' literature from the 90's, which includes being efficient in responsiveness and having human-like behaviour (Louwerse et al. 2005). To this end, a good chatbot relies not only on the user-interface design but also on the development of a robust AI-based process to support the conversation. The design, though, has to properly address eventual conversations breakdowns (Folstad and Brandtzaeg 2017).

According to Dale (2016), as the technology evolves, in the near future it should become hard to distinguish a bot from a human in a conversation. In most of cases, this clear distinction is important to avoid frustration, but for other menial tasks, this should not even be a concern for users. This, and other aspects of bots behaviour, and the eventual impact on people's lives, are still an open discussion around *Bot Ethics* (Berente and Salge 2017).

To this end, trust is an important aspect to be pursued in a chatbot design. Participants of Zamora's study (Zamora 2017) referred to this need of developing a trust relationship prior to engaging with potentially risky tasks that may fail due to mishandling data. Social media and financial-related are examples of such tasks.

Results from Zamora (2017) suggest that designing the best chatbot experience is always a contextual challenge. For instance, typing in a chatbot is adequate when the task requires confirmation or some logic thinking. Displaying menu options or also adding voice input can reduce error or recovery time. Facebook suggests a balance

---

between conversation and graphical user interfaces as part of their general chatbot guidelines (Facebook 2018). The guidelines include tips for establishing the bot identity, for sounding conversational, and an import remark for designers to not expect perfection. Regardless of how much the interaction is planned or anticipated, people constantly behave in an unexpected way (Facebook 2018).

Radziwill and Benton (2017) surveyed the literature and compiled a set of quality attributes expected for chatbots classified into 6 categories: 1) Performance, reflecting the ability to deal with unexpected input and the appropriate escalation to humans; 2) Functionality, which is related to linguistic accuracy; 3) Humanity, referring to humanised interactions; 4) Affect that encompasses enjoyability, politeness and personality traits; 5) Ethics and Behaviour, referring to respecting users' privacy, to sensitivity to social concerns and trustworthiness; and 6) Accessibility, by detecting users' meaning or intent and responding to social cues.

Going back to 2008, in their report that pointed future directions for Human-Computer Interaction research, Harper et al. (2008) (p. 42) raised 3 questions related to interaction and design in a context where we would be living with 'increasingly clever computers': *1) What will be an appropriate style of interaction with clever computers?*; *2) What kinds of tasks will be appropriate for computers; and when should humans be in charge?* and *3) How can clever computers be designed to be trustworthy, reliable and acting in the interests of their owners?*.

What the above literature overview evidences is that, 10 years later after the report by Harper et al. (2008), the bots are indeed becoming ubiquitous, but from a research perspective, these questions are still open, requiring systematic answers from HCI and AI together. The chatbots are still striving for robustness and, to provide valuable experiences for users, the design and evaluation require dealing with complex human and subjective components, some of them not yet fully explored from the HCI perspective.

3. THE UCHAGUZI EXPERIENCE

Uchaguzi is a customized deployment of the Ushahidi platform to monitor the Kenyan general elections, held on the 8th of August and 26th of October of 2017. The initiative taps into the collective intelligence of the crowd, where people are invited to send any information on the election/electoral process that they feel is significant.

To build an accurate and citizen-driven picture of the event, all reports are reviewed by a team of trained volunteers in terms of credibility and the eventual need of action. As Figure 1 illustrates, the involvement of volunteers happens at different levels.
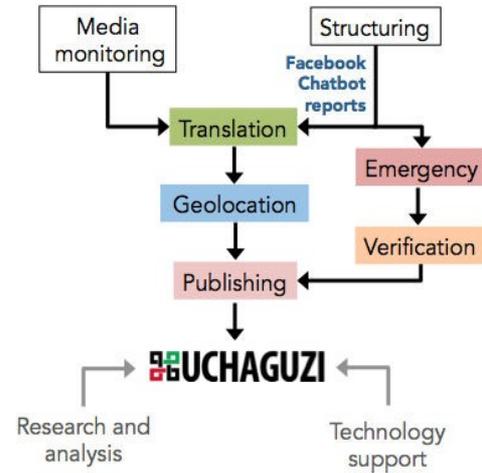

Figure 1: Uchaguzi workflow

In the first level, teams of volunteers check citizens reports from social media streams (media monitoring), while other groups work on adapting reports posted on Twitter, Facebook chatbot, and received via SMS into the structure of the platform (structuring). At the next level, while some people work on translating posts from the local language, mainly Kiswahili, to English, another team is verifying the posts in terms of geolocation. Urgent reports that require actions are flagged by an emergency team leader that escalates them to response partners. In the third level, the volunteers finally check the veracity and categorisation of the posts to be published. Additional to this work flow to populate the platform, there is an Analysis and Research team that cognitively process the data collected and provide reports to stakeholders, and a technical team responsible for the infrastructure and support.

To accomplish this, Ushahidi built partnerships with local civil society organisations, official NGOs, governmental and religious organisations, and local election observation groups. As a result, the initiative involved 700 monitoring individuals located at polling stations around the country, and 93 trained volunteers to help triaging the data.

The Facebook chatbot was one possible user interface for people to make a new report. While the platform remained active throughout the elections, the chatbot was piloted exclusively during the first round of the elections on the 8 of August.

3.1. Facebook chatbot

The chatbot design transposed the interaction on Ushahidi platform to the Facebook messenger, following the design principles by Facebook of keeping the interaction short, mixing conversation with Graphical user interface (GUI), using buttons, quick replies and menu to structure user input, using typing indicator to let people when the bot is in-progress, among others described in (Facebook 2018).

Figure 2 illustrates the initial message on Facebook Messenger to activate the chatbot; and Figure 3 shows the welcoming messages and the 'Send a report' button, which kickstarts the reporting process.
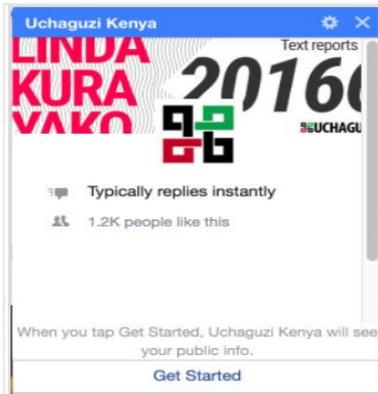
Figure 2: *Facebook messenger screen to start the chatbot*
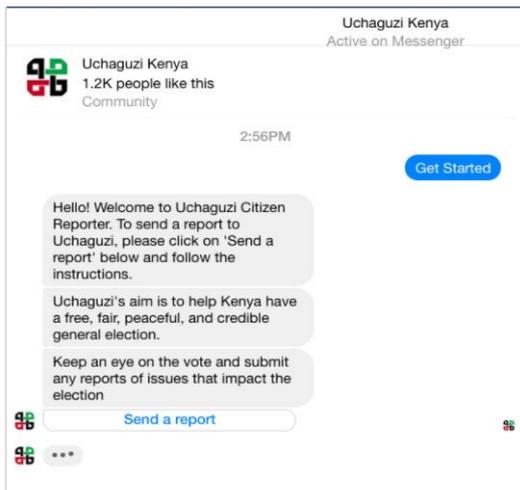
Figure 3: *Screenshot of the welcome messages*

After tapping 'Send a report', the users should select from the menu the type of the issue being reported, then type the report itself, as illustrated below in Figures 4 and 5 respectively.
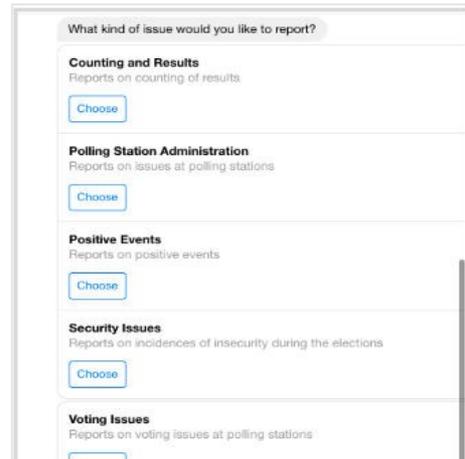
Figure 4: *Part of the menu with the types of issues for selection*
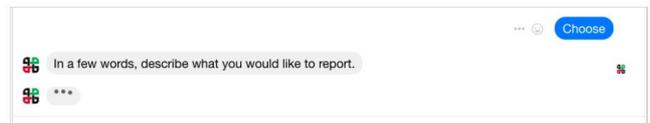
Figure 5: *Instruction for the user to type the report*

Users are then invited to add their geographical location and an image to the report before sending it to be checked and added to the platform, as Figure 6 depicts.
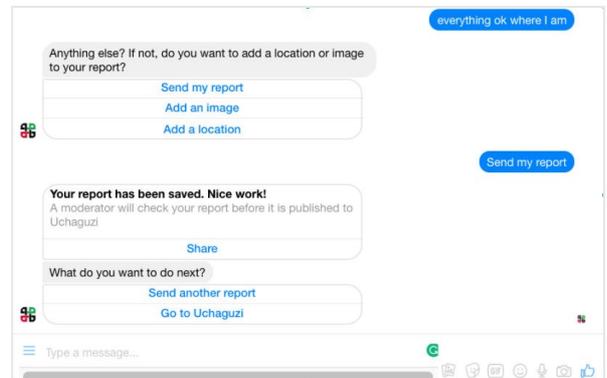
Figure 6: *Chatbot screenshot presenting options to send the report, add an image or the location*

Figures 7 and 8 refer to the instructions for the user do share their location and to add an image, respectively. And Figure 9 illustrates the confirmation that the report has been sent to be checked and eventually published on Ushahidi.
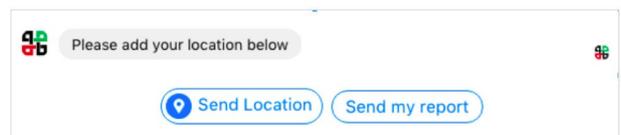
Figure 7: *Instructions to add a location*

4
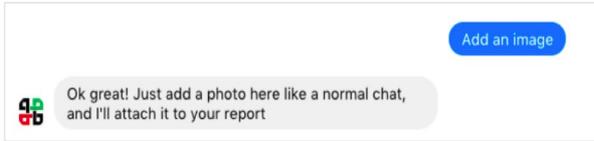
Figure 8: Instructions to add an image to the report
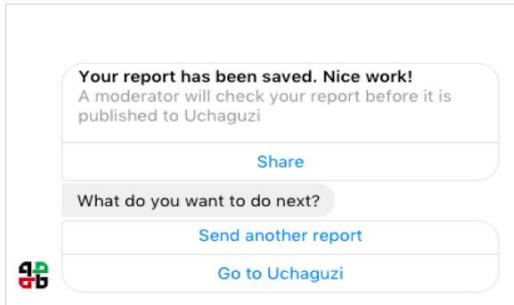
Figure 9: Confirmation that the report has been sent for analysis

3.2. Impact

Uchaguzi and the chatbot were advertised on Facebook targeting Kenyan residents during the first round of the elections. As a result, more than half of the visitors to the Uchaguzi website came from Facebook.

During this time, from the 8th to the 10th of August, the Uchaguzi Facebook chatbot received 3,034 visits, as described in Table 1.

In almost 50% of the visits, the users did not proceed with the interaction beyond the welcoming message. Presumably, the visitors were either curious people with no intention to create a report, or they did not make sense of the bot functioning.

In 38% of the visits, the reports were actually created, but the users ignored the navigation flow suggested by the chatbot GUI elements. In 222 visits (7%) the users started following the navigation, but failed at some point before sending the report. In 106 visits (3.5%) the users wrote the report, but faced a technical fault when submitting it to the platform. This technical issue lasted for a few hours due to communication overload between the platform and the bot. Only in 55 interactions, less than 2% of the visits, had the task completed and the reports were successfully sent to the platform.

Table 1: Interactions with the Facebook chabot

Interaction with the chatbot	Nr of reports
"Just browsing"	1,501
Did not follow the user flow	1,150
Tried to follow the user flow but failed	222
Technical issue to send the report	106
Sent to Uchaguzi	55
Total of reports originated	**912**

Total of visits	3,034

The data from the uncompleted interactions were manually checked, resulting in 912 reports originated from Facebook that were analysed and eventually added to the platform.

In total, Uchaguzi received 6,875 reports from all the channels (Twitter, SMS, etc.). The great majority of the reports were either commentary pieces or they did not fit into the classification of data relevant to Uchaguzi. Only 10% of this total (687) were actually structured, geolocated, verified, escalated if necessary, and published.

4. ANALYSIS

Although the number of visitors exceeded our expectations, the low number of finalised reports that were actually sent to Uchaguzi via the chatbot (2%) evidences an issue with the user interaction, and the causes can be many.

This analysis aims at diagnosing the issues experienced by the Facebook chatbot users in Kenya that led to this low success rate. To this end, first the interaction is going to be inspected, then the content of the conversation is examined to seek insights into the users' expectations and perceptions.

4.1. Inspection method

As a transient service during the elections time, meeting the actual chatbot users to collect their perceptions afterwards was not a feasible option. Considering this limitation, an inspection method was more appropriate for this research purpose.

To the best of our knowledge, there is not yet a consolidated inspection method targeting chatbots. We then relied on the traditional Cognitive Walkthrough (Wharton et al. 1994), a classical method applied by a specialist that simulates, or predicts, what the users will be thinking at every stage of the interaction. For every action towards completing the interaction (sending the report to Uchaguzi), the specialist answers these four questions (Wharton et al. 1994):

(Q1) Will the correct action be sufficiently evident to the user? (in other words, will the user know what to do to achieve the task?(Rogers et al. 2011))

(Q2) Will the user notice that the correct action is available? (Can users see the button or menu item that they should use for the next action? (Rogers et al. 2011))

(Q3) Will the user associate and interpret the response from the action correctly? (will users know from the feedback that they have made a correct or incorrect choice of action? (Rogers et al. 2011))

(Q4) If the correct action is performed, will the user see that progress is being made towards the completion of the task?

The actions towards completing the interaction are:

1. Get the chatbot started
2. Choose to send a report
3. Choose the kind of the issue
4. Type the report
5. Add the location (optional)
6. Add an image (optional)
7. Send the report

The analysis with the Cognitive Walkthrough simulated a real interaction with the Uchaguzi chatbot on Facebook messenger by a human-computer interaction specialist not involved with the chatbot design.

Additionally to the classic method application, the log file generated during the elections period was consulted for confirming some hypothesis on the users' perception. The log files describe every step of the interaction between the chatbot and the users, including the conversations. It is worth mentioning that for privacy reasons the Facebook users' id were not revealed for the analysis.

4.2. Inspection results

In this section, the most relevant findings of the inspection are described, shedding light mainly on the negative results obtained by answering the questions for every action previously outlined.

To activate the chatbot, all the 3,034 logged visitors had to tap 'Get Started' (Figure 2). Thus, the inspection did not reveal any issues related to this action.

Answers to the questions Q1, Q2 and Q3 for the Action (2) 'Choose to send a report' (illustrated on Figure 3) revealed a conflict between the bot interaction design and the users' perception. Although the button 'Send a Report' is evident (Q1), the log files evidenced that more than 70% of the users started typing the report directly in the chat without clicking this button (Q2). Also, the response the users get is unclear (Q3), since the typing indicator had been displayed for many seconds, potentially suggesting to the users that there was still more information to be presented from the bot side, as illustrated on Figure 3.

Results to Action (3) 'Choose the kind of issue' reinforced the understanding that the users were not expecting a menu of options. The majority ignored the selection and typed the report directly. Less than 10% of the users that made a report selected the type of the issue. For them, the

dialogue structure did not help (Q4). Independently of selection, the feedback was the word 'Choose' instead of the option chosen, as Figure 4 illustrates.

After typing the report text, action (4), the bot says 'Anything else? if not, do you want to add a location or image to your report?' (Figure 6), which is not a clear instruction towards the next step, raising a problem related to question Q4.

Adding the location (action 5) and an image (action 6), were optional steps taken in only 5% of the interactions each. Strategically, the users should be convinced to add both the images and geolocation by being informed on how valuable these information would be for the report, as revealed by answering Q1 to action (5) 'Add the location'. Question Q3 pointed out that the instruction provided by the bot; 'Add your location below' (Figure 7), which did not match the button label 'Send Location', may have suggested to one-third of the users that reached this point to actually type the location manually instead of using the button 'Send Location', as the log analysis revealed. The pointer to the next steps (Q4) could also generate misunderstandings when giving the user the option to "restart" at this point of the interaction.

Similarly, Q3 for action (6) 'Add an image' suggesting to 'add a photo as a normal chat' could be confusing, since for the users the distinction between the bot and a "normal chat" may not be evident.

Less than 5% of the interactions reached the last action (7) 'Send my report', available from the actions (4), (5) and (6). Question Q1 brings to light whether the users were expecting to press a Send button on a conversational, not transactional, interaction. The feedback message (Q4) 'Your report has been saved', illustrated in Figure 9 may also not make sense for users situated in a dialogue context.

4.3. Overview of the conversations

Complementing the interaction analysis described above, the content of the conversations between the bot and the users has been evaluated in order to find indications on the users' perception and expectations. This evaluation does not take into account the technical aspects of natural language processing, nor analyses the relevance of the content generated for the social issue. Instead, the analysis is supported by desirable characteristics of chatbot found in the literature, especially concerning robustness and humanised elements (Zamora 2017).

Following the design guidelines by Facebook (2018), Uchaguzi chatbot is strongly supported by the graphical and interactive elements, practically limiting the expected textual input to the report content. Even though, a total of 2,390 text entries

were created for the 912 reports. Beyond the report itself, the entries include greetings, questions to the bot, and personal opinions.

Responses from the bot to unexpected input (i.e., 'Oh my, I'm not programmed to understand what you're saying. Sorry!' or 'Sorry, I'm just a bot built for reporting, so my vocabulary is pretty limited') represented 10% of the bot entries in the dialogue.

The word cloud in Figure 10 brings to light the 30 most frequent English words in the users' entries, grouping similar words like voting, vote, etc., and ignoring the most common words such as articles (a, the, etc.) and conjunctions (but, or, etc.), among others.

Figure 10: the 30 most frequent words from the users' entries

The highlighted words 'peaceful', 'good' and 'well' suggest that users used the Facebook channel to report mostly positive events. The main issues reported were related to 'queue', 'line', 'slow', 'wait'.

Even though the users were aware that they were communicating with a bot, many people used greetings and emoticons; 43 entries started with 'hello', 'good morning' or 'good afternoon'; and 40 'thanks', 'thank you', and 'sorry' were found in the users' text.

5. DISCUSSIONS

As a short-lived, "in-the-wild" (Chamberlain et al. 2012), and concerned with participants' privacy, this research had to consider alternatives to overcome the lack of involvement of actual users.

Additional to that, there is still a gap in the literature related to methods and criteria to evaluate chatbots. Although the inspection method Cognitive Walkthrough (Wharton et al. 1994) does not address particularities of conversational interfaces, the application was viable due to the interactional nature of the Uchaguzi chatbot design. Nevertheless, supported by the log files, the method was effective in suggesting the users' most evident understandings that could not have been predicted by the designer.

As the literature review points out, defining what are the desirable characteristics of a chatbot still demands contextual investigations. So far, it has been acknowledged that evaluating a chatbot involves a multitude of subjective aspects (Morrissey and Kirakowski 2013), many of them are

too complex to measure and almost impossible to be estimated without the direct involvement of the users, a limitation of this study.

In this scenario, reflections on the results of the analysis are presented, taking into account the potential and constraints faced during the pilot in Kenya, and some lessons learned are discussed from the perspective of desirable characteristics of a chatbot as found in the literature Radziwill and Benton (2017).

5.1. Potential as an inclusive technology

Chatbots have been recognised as a promising way to engage people with any sort of technology-related service, due to them being easy to access, to interact with, and simplicity in deploying as a service. The high number of visitors to the Uchaguzi chatbot during the elections (3,034) surpassed the expectations, therefore confirming its potential in reaching new users.

Most users that accessed the chatbot were likely to be unaware of this crowd-sourced platform, or were unwilling to learn how to operate it to submit a report. While the Facebook messenger clearly brought the Uchaguzi initiative closer to a number of first-time users, it revealed several design implications, as further described.

5.2. The pilot design

Problems with the interaction design, as detected in our inspection, could be adjusted to enhance the users' experience for further events. Naming a few, the modifications include splitting and re-phrasing the instructions to add the geolocation and to attach an image (Figure 6), therefore guiding the users to complete both actions; and removing the typing feedback (as visible, for instance, in Figure 3), which suggests that the bot was still preparing some information to be presented, whereas it was actually waiting for the users' input.

Although these and other minor modifications would improve user-chatbot interaction, they may not solve the mismatch between the users' perception of a conversational interface and the graphical-centred interaction designed, which was a barrier that presumably prevented many users from completing their reports.

5.3. Design considering first-time users

Although users were probably familiar with Facebook Messenger on their mobile devices, it is likely that many of them were encountering a chatbot for the very first time. As they were not expecting having a graphical user interface and interactive mechanisms on the chat, the majority ignored or got lost in the navigation flow planned in the design.

Simply switching to a fully conversational-based interaction is not yet an option, mainly when including inexperienced users. It is important to keep in mind the immaturity of the technology in general when dealing with the unpredictability of user-generated texts as input. People not yet aware of the technology limitations and its potential failures may feel frustrated and demotivated, thus counteracting the dissemination purpose of the initiative.

This current state of the technology justifies the recommendation in (Facebook 2018) to pursue a balance between conversational and graphical interaction styles. More than that, as pointed out by Zamora (2017), the best interaction mechanism should be chosen according to the context, for instance replacing text-based input, which is more susceptible to errors, with making selections whenever it is possible.

To create a truly inclusive experience, the design should guide the users throughout the navigation, step-by-step, telling them exactly what is expected in every step until fulfilling the final objective. They also need to be informed of the value of the information they are about to provide in their report to the social issue.

However, more frequent users must also be catered for in the design. Again, the design should aim for a balance between being practical and explanatory enough, or considering tailoring the instructions and dialogue for first-time or recurring users.

Furthermore, evaluate whether every action is indeed necessary or could be simplified, making the interaction shorter and more direct, which is desirable in a mobile context. Further developments of Uchaguzi chatbot, for instance, will assign the action of choosing the type of issue (Figure 4) to the analysts of the reports instead of the users.

5.4. Towards design guidelines

Some findings are discussed from the perspective of desirable chatbot characteristics compiled by (Radziwill and Benton 2017). This set of characteristics does not intend to be exhaustive nor to contest the design guidelines provided by Facebook (2018). Instead, they are framing the lessons learned in this experience, pointing towards design aspects that should be considered in further studies or initiatives applying chatbots for crises, or in similar situations where first-time users are attracted.

- **Performance and robustness**: Considering the low success rate in sending the reports to Uchaguzi, the chatbot was not efficient in reaching its main objective. For ensuring a good performance, a chatbot should guide the inexperienced users throughout the navigation, convincing them to complete every action towards the end of the task.

Another aspect of performance and robustness, the impact of conversations breaking down, could not be measured in this analysis. But it is acknowledged that open dialogue-based interaction should be explored with care, avoiding exposing the users to unnecessary pitfalls and frustrations.

- **Functionality**: In (Radziwill and Benton 2017), functionality refers mainly to accuracy in interpreting input and linguistic precision of the out-put. In terms of natural language processing, the scenario revealed an extra complexity due to the combination of languages, since a few users mixed English and Swahili in some reports. In the Uchaguzi context it is not a problem considering that the reports will be analysed and translated by volunteers before the publication. But in other domains, the co-existence of languages may lead to far more frequent conversations breaking down.

- **Humanity and affect:** Humanity refers to the ability to be convincing, satisfying, and a natural interaction (Radziwill and Benton 2017). The fact that many users used greetings and emoticons reinforce this expectation from people. A certain degree of informality on the dialogue can bring some humanity to the interaction. Regarding affect, some chatbot-based services rely on fun and entertainment to engage users. In the context of crises and violence, this is not an advisable approach since some people may be facing critical situations. However, showing politeness, informality, and responding to users' moods in a friendly way are desirable aspects. Users' entries like 'Good afternoon first of all', 'Sorry for the typo' or ending their reports with 'thank you', reinforce the importance of these aspects for the users.

- **Ethics:** In a crises context, it is crucial that the chatbot reveals and reinforce its identity and do not pretend to be a human, since reports may address sensitive content or require immediate actions. Also, the users should be informed about the privacy around their identity and the reports generated. It is important that the users trust the initiative towards the social issue, the technical platform, and the user interface.

- **Accessibility:** In (Radziwill and Benton 2017) accessibility refers to responding to social cues, detecting intentions and meaning. In crises contexts, the idea of accessibility should be broader than that, comprising also bandwidth and network availability, timing in the interaction, compatibility with difference types of devices, dictation features, etc. More research

in the field is necessary to detect real-life constraints that affect accessibility.

6. REFLECTIONS ON CHATBOTS FOR CRISES

While the use of chatbots is relatively new for crisis, the idea of reaching out to affected populations or engaging them for collecting information is not. Roberts and Doyle (2017) highlight several case studies for this form of engagement during various crises.

Indeed, individuals and groups – such as spontaneous unaffiliated volunteer groups – within an affected population often possess local knowledge for specific aspects of a crisis and access to communication channels and informal networks that may be more effective than official systems during a crisis event.

As Roberts and Doyle (2017) point out, a single piece of technology alone is not the answer: crisis responding organisations are encouraged to build relationships with the local population around questions of data interoperability, data sharing, and understanding how their policies might inhibit or affect data sharing and collection. The creation of a chatbot itself must be considered in conjunction with these other factors to maximise its utility for crowd sourcing.

Community's practices and values should also be taken into account when featuring some human-like aspects of the chatbots. Rapport with the community and with the situation they are facing, which can be translated into interaction and dialogue elements, are examples. Although not explored in this pilot, and in line with the literature, this study recognises the human elements as important steps towards building trust in the chatbot application and in the socio-technical initiative as a whole.

As suggested by this study results, the Facebook chatbot was effective in attracting people, reaching citizens that probably would not engage with the initiative otherwise. Although this is a positive outcome, possible side effects should also be taken into account in the design. For example, a system overload (as recorded), and the need to deal with a high volume of non-informative reports by citizens who might be unaware of the platform's functioning and impact.

6.1. Limitations and further works

This study during the Kenyan election revealed insights regarding challenges and potentials of using chatbots to engage citizens as reporters in crisis situations. Therefore, the nature and impact of a crisis and emergency circumstances to people's lives can be so diverse that it is hard to predict that people's behaviour and the impact of the technology would be the same in another scenario with different variables in play.

Further studies are necessary to assess solutions in different contexts, considering technical, social and human aspects. Examples are: evaluating the technical effort to quickly build a functional and robust chatbot in a crisis situation both from the AI perspective and interoperability with other platforms; evaluating users' perception and appropriation according to distinct levels of familiarity with technology; investigate people's perception (and tolerance) when coping with failures or conversation breaking down in stressful conditions, to name a few.

Future studies will also consider possibilities for the participation of actual users to collect their feedback in terms of perception and expectation, advancing into a more systematic approach to design chatbots for crises considering the current stage of this technology.

7. CONCLUSIONS

The potential of chatbots in engaging users has recently attracted much attention both in the technical and commercial domains, and new applications are continuously emerging in many different aspects of life.

Adding to the literature on how to achieve a real social impact with this technology, this paper reported a pilot study, in a real life context, using Facebook chatbot as a way to engage citizens in Kenya with a socio-technical initiative towards a fair presidential election in 2017. The chatbot opened the doors of a crowd-sourced based platform to more than 900 people that reported incidents or their own experiences with the elections process. This number, though, could be far bigger if more people out of the 3,000 visitors could have made better sense of the functioning of the chatbot and properly followed the navigation design. Even though the design was based on a platform as popular as Facebook, the users of the chatbot tried to engage in a pure conversational style, instead of using the graphical elements that are supposes to facilitate the interaction. As a result, less than 2% of the users managed to complete their reports. An inspection of the interaction design revealed various issues that are likely to have hindered the general user experience, although the mismatch between the chatbot design and the perception of the users could not have been predicted by the designer without adequate user evaluations.

Results from this research confirm the potential of chatbots as an inclusive technology and point to directions for the design of other socio-technical

initiatives using chatbots that need to ensure a good experience, especially for first-time users.

REFERENCES

Berente, N. and C. Salge (2017, 02). Is that social bot behaving unethically? *Communications of the ACM 60*(9), 29–31.

Chamberlain, A., A. Crabtree, T. Rodden, M. Jones, and Y. Rogers (2012). Research in the wild: Understanding 'in the wild' approaches to design and development. In *Proceedings of the Designing Interactive Systems Conference*, DIS '12, New York, NY, USA, pp. 795–796. ACM.

Collins, K. (Aug, 2017). Facebook Messenger bots help monitor Kenya election violence. https://www.cnet.com/news/facebook-messenger-bots-help-monitor-kenya-election-violence/

Dale, R. (2016). The return of the chatbots. *Natural Language Engineering 22*(5), 811817.

Facebook (2018, Jan). Messenger platform design best practices. https://developers.facebook.com/docs/messenger-platform/introduction/general-best-practices.

Folstad, A. and P. B. Brandtzaeg (2017, June). Chatbots and the new world of HCI. *interactions 24*(4), 38–42.

Harper, R. et al. (2008) Being Human: Human-Computer Interaction in the Year 2020. Microsoft Research, Cambridge, U.K., 2008.

Heidelberg. Rashid, K. and A. Das (2018). *Build Better Chatbots: A Complete Guide to Getting Started with Chatbots*. Apress.

Hyman, P. (2014, January). 'Peace technologies' enable eyewitness reporting when disasters strike. *Communications of the ACM 57* (1), 27–29.

Louwerse, M., A. Graesser, S. Lu, and H. Mitchell (2005). Social cues in animated conversational agents. *Applied Cognitive Psychology 19*(6), 693–704.

Morrissey, K. and J. Kirakowski (2013). 'Realness' in chatbots: Establishing quantifiable criteria. In M. Kurosu (Ed.), *Human-Computer Interaction. Interaction Modalities and Techniques*, pp. 87–96. Springer Berlin

Radziwill, N. M. and M. C. Benton (2017). Evaluating quality of chatbots and intelligent conversational agents. *CoRR abs/1704.04579*.

Roberts, S. and T. Doyle (2017). *Understanding Crowdsourcing and Volunteer Engagement*, pp. 121–134. John Wiley and Sons, Inc.

Rogers, Y., H. Sharp, and J. Preece (2011). *Interaction Design: Beyond Human - Computer Interaction* (3rd ed.). Wiley Publishing.

UN (2008, 02). United nations high commissioner for human rights -report from ohchr fact-finding mission to kenya 6-28 february 2008.

Wharton, C., J. Rieman, C. Lewis, and P. Polson (1994). Usability inspection methods. Chapter The Cognitive Walkthrough Method: A Practitioner's Guide, pp. 105–140. New York, NY, USA: John Wiley & Sons, Inc.

Zamora, J. (2017). I'm sorry, dave, i'm afraid i can't do that: Chatbot perception and expectations. In *Proceedings of the 5th International Conference on Human Agent Interaction*, HAI '17, New York, NY, USA, pp. 253–260. ACM.

Zhou, A. (2017, Mar).6 ways bots are positively changing the world. www. forbes.com/sites/adelynzhou/2017/02/16/see-six-ways-bots-are-positively-changing\-the-world/.