Project presentation at CONVERSATIONS 2023 – the 7th International Workshop on Chatbot Research and Design, hosted by the University of Oslo, Norway, November 22-23, 2023.

User Perceptions, Experiences and Interactions with Municipalities' Chatbots Differing in Human Likeness and Interaction Design

Charlotte van Hooijdonk[1](✉), Christine Liebrecht[2], Naomi Kamoen[2], and Gabriëlla Martijn[1]

[1] Utrecht University, The Netherlands
<c.m.j.vanhooijdonk@uu.nl, g.n.martijn@uu.nl>
[2] Tilburg University, The Netherlands
<c.c.liebrecht@tilburguniversity.edu, n.kamoen@tilburguniversity.edu>

Abstract. Municipalities are increasingly implementing chatbots as a part of their digital service provision. The extent to which users embrace the chatbot plays a role in determining the success of this implementation. Multiple factors play a role in users' perceptions, experiences, and interactions with chatbots, such as the human likeness (e.g., avatar, name, and communication style) and interaction design (e.g., free text versus buttons). This project examines how users perceive and interact with Dutch municipality chatbots. A unique feature of the project is that users interact with multiple Dutch municipal chatbots that differ in terms of humanlikeness and interaction designs. A mixed-methods approach is adopted encompassing both a qualitative interview study and a content analysis. The project is expected to have key implications for theory and practice on municipality chatbots.

Keywords: Digital Government · Chatbot, · Human-likeness · Interaction Design

1 Introduction

1.1 Background

Recent advances in AI have led to the wide adoption of chatbots in public sectors, such as in municipal administrations [1,3,5]. Chatbots, or conversational agents, are software agents designed to communicate with users in everyday language [6]. According to Makasi et al. [11], municipality chatbots commonly serve the purpose of service triaging, which involves assisting users in accessing information and services. For example, the Dutch omnichannel chatbot 'Gem' was created by a municipal collective and provides generic information about the municipalities and related services. This chatbot is currently used in 20 municipalities.

There is a growing body of knowledge on citizens' experiences with municipality chatbots. For example, Abbas et al. [1] conducted an interview study

examining citizens' intentions to use a municipality chatbot. Their findings indicated that citizens found the chatbot useful for navigating available information and services. Følstad et al. [3] adopted a mixed methods approach to investigate considerations about the design and interactions with a human-like municipality chatbot. Citizens and municipality representatives were interviewed to explore their considerations regarding human-like features (e.g., name and avatar) of the chatbot. Results showed citizens expected that the chatbot could give them swift responses to simple requests. In addition, 2,663 conversations were analyzed on citizens' adopted communication style. The results of this dialogue analysis showed citizens had short conversations with the chatbot and formulated their requests briefly and without politeness markers. It was therefore concluded that citizens are highly goal-oriented when interacting with a municipality chatbot and its human-like features had a limited impact on their expectations and behavior [3].

The previous studies provide relevant insights into the citizens' perceptions of the humanness of municipality chatbots. The human-like features of chatbots, such as avatar, name, and communication style, have the potential to boost the perception of chatbots as more human, thereby encouraging users to interact with them in a socially-oriented manner. However, the interaction design of chatbot may also have an impact on these perceptions. Users can communicate with a text-based chatbot by means of free text input and/or buttons with predefined answers [13]. On the one hand, research has demonstrated buttons increase the chatbot's usability, prevent miscommunication, and decrease humanness perceptions [8,9]. On the other hand, free text input mimics computer-mediated human-to-human interaction, but may also lead to miscommunication [8,12]. Therefore, it is relevant to investigate the chatbots' human-like features (i.e., identity, visual and conversational cues) as well as the interaction design, as on a meta-level there is an interesting paradox in the results of studies investigating these factors in isolation [7,8,12].

In this project, we methodologically replicate the research of [1] and [3], by investigating participants' humanness perceptions, usability experiences, and interactions with chatbots. Three Dutch municipality chatbots are used that differ not only in the extent to which they incorporate human-like features, but also in interaction design: free-text interaction, button interaction, or a combination of both. This way, the interplay between human-like features and the interaction design of municipality chatbots can be investigated in an ecologically valid context.

1.2 Objectives

This project has two objectives which are addressed in two studies:

1. To investigate how human-like features and interaction designs of Dutch municipality chatbots affect users' perceptions of humanness and their usability experiences.
2. To investigate how human-like features and interaction designs of Dutch municipality chatbots affect users' interactions.

2 Study Methods and Planning

To reach these objectives two studies will be conducted: an interview study and a content analysis of chatbot conversations.

2.1 Interview study

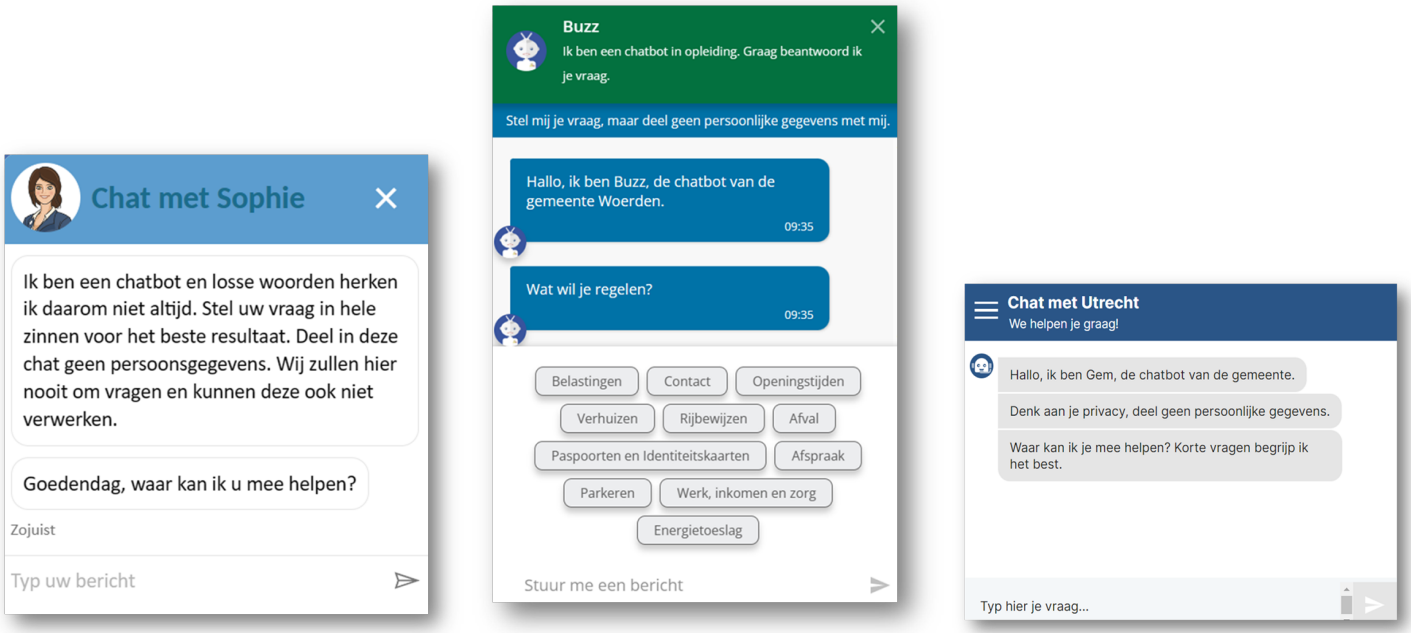
We conducted a qualitative interview study to address the first objective. To investigate the impact of human-like features and interaction design on users' humanness perceptions and usability experiences, participants performed a set of tasks with three Dutch chatbots on the website of three different municipalities.

Participants. Fifty participants were recruited through convenience sampling. Among the participants, 64% were women (N = 32), and 36% were men (N = 18). The participants' ages ranged from 19 to 81 (M = 33.98, SD = 18.82). Regarding educational backgrounds, 20% of the participants had completed or were enrolled in a master's degree program, 18% had a bachelor's degree, 40% had completed higher professional education, 10% had completed intermediate vocational education, and 12% had a high school education or lower. All participants reported prior experience with (customer service) chatbots, with varying frequencies of interaction (83% had contact several times per year). Finally, 90% (N = 45) of the participants had prior experience with contacting their municipality, while 10% (N = 5) had no prior experience.

Tasks. Each task had a theme introduced by a scenario in which the context and relevant details were explained, such as requesting a new passport. The tasks differed per chatbot, but were comparable in difficulty among the three chatbots. Participants engaged in the same series of tasks whereas the allocation of chatbots was randomized.

Chatbots. The tasks were performed with three chatbots of medium to large Dutch municipalities: Hollands Kroon (chatbot Sophie), Utrecht (chatbot Gem), and Woerden (chatbot Buzz). All chatbots provided generic information on a broad range of topics, such as garbage, taxes, and passports, which were either presented within the dialogue or by presenting a link to the municipality website. Thus, all chatbots provided service triaging [10]. However, the chatbots differed in type and number of human-like features present both in the welcome messages (e.g., avatar and name) and during the interaction (e.g., communication style). Moreover, the chatbots differed in their interaction design: free text, buttons, or a combination. Figure 1 shows the three chatbots that were used by all participants in the study.

Procedure. The semi-structured interviews were audio-recorded and participants' interactions with the chatbots were screenrecorded. The interview consisted of three parts. In the first part, participants were interviewed regarding their prior



(a) Hollandse Kroon (b) Woerden (c) Utrecht

Fig. 1: The three municipality chatbots

experiences with chatbots and contacting their municipality. In the second part, the chatbots were introduced individually. The chatbot's welcome messages were shown and participants reflected on its humanness and their expectations about the usability of the chatbot. Next, they performed a task with the chatbot, after which they reflected on the task performance and it's communication style. In the third part, participants reflected on the humanness of three chatbots as well as their usability experiences with the chatbots and ranked them accordingly.

Analysis. First, we analyzed the screen recordings on whether participants were able to complete the tasks with the chatbots. Second, the audio recordings were transcribed verbatim and the transcripts are now being analyzed using a thematic analysis following the guidelines of Braun and Clark [2]. The objectives of the analysis were data-driven themes that are associated with:

1. participants' humanness perceptions and expectations about the usability of the chatbot based on the welcome messages;
2. participants' usability experiences regarding the task performance and their interaction with the chatbot.

2.2 Content analysis.

The next step in this project will be to investigate the second objective. A corpus of dialogues will be collected from the three municipalities. We will manually analyze the dialogues using the framework of Følstad et al.[6]. Also, building on another study of Følstad et al. [5], we will analyze users' input on message brevity and communication style.

3 Preliminary findings and expected contributions

3.1 Preliminary findings

Participants' humanness perceptions and usability expectations. Our preliminary findings show that some participants based their usability expectations on the human-like features present in the chatbots' welcome message. For example, participants expressed that they could perform the task efficiently with the chatbot of Hollands Kroon which was based on presence of the human avatar and the-human name. This is exemplified in this quote by one of the participants below:

> *"Sophie... I expect it will go well because it is a common name. I have the idea that I am chatting with someone"* (P.49)

Participants' usability experiences. After completing the tasks, participants reflected on the chatbots' interaction design. As exemplified in the following quote participants expressed positive experiences with the combination of free text and button interaction of the chatbot of Hollands Kroon:

> *"I prefer to write a text because, with buttons, you have to click a lot before you eventually get somewhere. From my own experience, if I just have a sentence with, for example, the word 'passport,' I get to the content more quickly. But I did find that with Sophie, there was a combination of both, which was ideal. So first just a sentence and then the buttons"* (P.32)

Participants had mixed experiences with the chatbot of Woerden. Although the buttons allowed an easy way of interacting with the chatbot, participants thought the chatbot could not assist them if the button labels did not correspond with their request.

> *"These buttons were easy to use, but I do feel that if my option is not there, I will not be able to get further assistance."* (P.33)

Another experience about the interaction design relates to the large amounts of text in the chatbot's responses. As exemplified in the following quote, participants noted that they expected to receive short answers from the chatbot:

> *"Because he wants me to formulate my request briefly and concisely, I expected the chatbot to be brief as well"* (P.15)

In November 2023, during the conference, we can share the findings of the interview study.

3.2 Expected contributions

This project will augment the body of knowledge on the municipality chatbots. By methodologically replicating previous research [1,3,5], we make a theoretical contribution by extending and generalizing insights on European municipality chatbots. Additionally, chatbots are seen as low-threshold interfaces to access government information and services [4]. By examining how the interplay between the chatbot's human-like features and interaction design influences users' humanness perceptions, usability experiences, and interactions, we will be able to formulate practical guidelines for government chatbot designers to enhance the accessibility and user-friendliness of municipality chatbots.

Acknowledgements. The authors would like to thank Nienke Breugelmans, Rémy Chefiri, Nicole Hagendoorn, Sanne Hoogendoorn, and Anne Molegraaf, for conducting the interview study. This research is part of the NWO-funded project 'Smooth Operators: Development and Effects of personalized conversational AI', grant no: KIVI.2019.009.

References

1. Abbas, N., Følstad, A., Bjørkli, C.A.: Chatbots as part of digital government service provision–a user perspective. In: International Workshop on Chatbot Research and Design. pp. 66–82. Springer International Publishing, Cham (November 2022)

2. Braun, V., Clarke, V.: Using thematic analysis in psychology. Qualitative Research in Psychology **3**(2), 77–101 (2006)
3. Følstad, A., Larsen, A.G., Bjerkreim-Hanssen, N.: The human likeness of government chatbots–an empirical study from norwegian municipalities. In: International Conference on Electronic Government. pp. 111–127. Springer Nature Switzerland, Cham (August 2023)
4. Følstad, A., Araujo, T., Law, E.L.C., Brandtzaeg, P.B., Papadopoulos, S., Reis, L., Baez, M., Laban, G., McAllister, P., Ischen, C., et al.: Future directions for chatbot research: an interdisciplinary research agenda. Computing **103**(12), 2915–2942 (2021)
5. Følstad, A., Bjerkreim-Hanssen, N.: User interactions with a municipality chatbot—lessons learnt from dialogue analysis. International Journal of Human–Computer Interaction pp. 1–14 (2023)
6. Følstad, A., Taylor, C.: Investigating the user experience of customer service chatbot interaction: a framework for qualitative analysis of chatbot dialogues. Quality and User Experience **6**(1), 6 (2021)
7. Go, E., Sundar, S.S.: Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions. Computers in Human Behavior **97**, 304–316 (2019)
8. Haugeland, I.K.F., Følstad, A., Taylor, C., Bjørkli, C.A.: Understanding the user experience of customer service chatbots: An experimental study of chatbot interaction design. International Journal of Human-Computer Studies **161**, 102788 (2022)
9. Klopfenstein, L.C., Delpriori, S., Malatini, S., Bogliolo, A.: The rise of bots: A survey of conversational interfaces, patterns, and paradigms. In: Proceedings of the 2017 conference on designing interactive systems. pp. 555–565 (2017)
10. Makasi, T., Nili, A., Desouza, K., Tate, M.: A typology of chatbots in public service delivery. IEEE Software **39**(3), 58–66 (2021)
11. Makasi, T., Nili, A., Desouza, K., Tate, M.: Public service values and chatbots in the public sector: reconciling designer efforts and user expectations. In: Proceedings of the 55th Hawaii International Conference on System Sciences. pp. 2334–2343. University of Hawai'i at Manoa (2022)
12. Nguyen, Q.N., Sidorova, A., Torres, R.: User interactions with chatbot interfaces vs. menu-based interfaces: An empirical study. Computers in Human Behavior **128**, 107093 (2022)
13. Valério, F.A., Guimarães, T.G., Prates, R.O., Candello, H.: Comparing users' perception of different chatbot interaction paradigms: a case study. In: Proceedings of the 19th Brazilian Symposium on Human Factors in Computing Systems. pp. 1–10 (2020)