Accepted version (postprint) of paper with reference: Følstad, A., Larsen, A. G., & Bjerkreim-Hanssen, N. (2023). The human likeness of government chatbots–An empirical study from Norwegian municipalities. In Proceedings of the International Conference on Electronic Government – EGOV2023 (pp. 111-127). Cham, Switzerland: Springer. The version of record is available at https://link.springer.com/chapter/10.1007/978-3-031-41138-0_8

The Human Likeness of Government Chatbots – an Empirical Study from Norwegian Municipalities

Asbjørn Følstad[1][0000-0003-2763-0996], Anna Grøndahl Larsen[1][0000-0002-6600-7606], and Nina Bjerkreim-Hanssen[2][0009-0009-7888-2393]

[1] SINTEF, Forskningsveien 1, 0373 Oslo, Norway
[2] Prokom, Karl Johans gate 37b, 0162 Oslo, Norway
asf@sintef.no

Abstract. While chatbots represent a potentially useful supplement to government information and service provision, transparency requirements imply the need to make sure that this technology is not confused with human support. However, there is a knowledge gap concerning whether and how government chatbots indeed represent a risk of such confusion, in spite of their resemblance with human conversation. To address this gap, we have conducted a study of a Norwegian municipality chatbot including interviews with 16 chatbot users and 18 municipality representatives, as well as analysis of >2600 citizen dialogues. Interviews with citizen and municipality representatives suggested that citizens typically understood well the chatbot capabilities and limitations, though municipality representatives reported on some examples of humanizing the chatbot in its early phases of deployment. Dialogue analyses indicated that citizens have a markedly utilitarian style in their communication with the chatbot, suggesting limited anthropomorphizing of the chatbot.

Keywords: Chatbot, government, transparency, human-likeness.

1 Introduction

Government service provision increasingly make use of chatbots to facilitate service delivery for efficiency gains and improved availability [25]. Chatbots are software agents which provide users with access to information and services through natural language interaction conducted in the form of dialogue [11]. Chatbots are considered an intuitive way of interacting with computer systems, due to the resemblance of the chatbot dialogue with that of a conversation with a fellow human being [17], which potentially lowers barriers to interaction and engagement. Furthermore, the human likeness resulting from a chatbot's use of natural language interaction and resemblance of human conversation, has been suggested as conductive to improved user experience [19] and to reflect positively on users' perceptions of a service provider [4].

However, humanlike chatbots may also entail negative implications for users and service providers. Human likeness may induce erroneous expectations concerning chatbot capabilities and limitations and lead to unwanted interaction patterns or strategic user behavior [29]. Chatbot human likeness may also lead to uncertainty or deception,

where users become uncertain with regard to the chatbot's status as an automated agent or even erroneously believe they are indeed interacting with a human [24]. Curbing such undesirable uncertainty and deception concerning the machine status of a chatbot is important with regards to transparency requirements for trustworthy AI systems [8, 18]. Furthermore, the proposed European Commission AI Act, specifically details such transparency requirements on chatbot providers [10, 33].

In this context, knowledge is needed on how citizens perceive government chatbots. Specifically concerning their perceptions of chatbot human likeness and how this may impact chatbot interaction and, by extension, government service provision. However, while current research has investigated determinants and implications of chatbot human likeness [e.g., 4, 15, 19], there is a lack of knowledge on whether and how such human likeness perceptions impact user interactions with government chatbots.

In response to this gap in knowledge, we have conducted a study to explore whether and how users interact with and perceive government chatbots as humanlike service providers. The study addressed a chatbot for Norwegian municipalities and included three method components. First, we interviewed 16 citizens on their experiences with the chatbot. Second, we interviewed 18 government representatives with a role in maintaining the chatbot. Third, we analyzed >2600 chatbot dialogues between citizens and a government chatbot.

The study contributes needed knowledge on human likeness and transparency in government chatbots. The interviews with municipality representatives and users contribute knowledge on how the government chatbot is perceived and the implication of such perceptions on behavior and service outcomes. The dialogue analysis contributes insight into users' communication style, indicative of markedly utilitarian goal-orientation.

2 Background

2.1 Government Chatbots

Chatbots are about to become a commonplace channel of government provision of information and services. Already in 2020, a survey identified a substantial appearance of chatbots as part of European government service provision [9]. Currently, chatbots are among the most frequently deployed AI applications in the public sector [36].

Within government service provision, chatbots have been taken up for a broad range of service sectors or areas such as health [31] and social services [37], and by broader service providers such as cities [35] and municipalities [1]. During the pandemic, the uptake of government chatbots saw a marked boost as part of pandemic response [3].

The recent surge in government chatbots seems motivated by beneficial aspects of the technology both for government agencies and for citizens [35]. On the side of the government agencies, chatbots can enable reduced employee workload, and lowered service delivery cost, while improving users' service experiences through providing a more personalized and efficient service delivery [25]. Chatbots have also been explored as a means to strengthen citizen participation and engagement in government [32]

Government chatbots typically are implemented as intent-based solutions, where machine learning is applied to predict user intents from users' free text requests and

then provide needed information and services on the basis of predefined content [23]. Most government chatbots are set up as what Makasi et al. [25] refer to as chatbots for service triage, that is, they provide generic information and access to services without adaptation to a user profile.

There is a growing body of knowledge on how users experience government chatbots. Makasi et al. [26], in an interview study with users and designers of government chatbots, found that such chatbots were perceived to enable increased effectiveness and efficiency in service provision, while potentially also strengthening accessibility and ease-of-use for government services. Abbas et al. [1], in an interview study with municipality chatbot users, found users to appreciate the navigation support provided by the chatbot and its potential for simplifying access to government information.

2.2 Chatbot Human Likeness and Its Implications

While users typically have been found to have utilitarian motivations for chatbot use, such as efficiency and convenience [5], chatbots may also potentially improve user experience due to their human-likeness in appearance and communication style [19]. In consequence, there has been substantial industry and research interest in the benefits and limitations of chatbot human likeness [27], and the factors that may determine users' tendencies to *anthropomorphize* chatbots, that is, a tendency to imbue the behavior of an agent with motivations, intention, or emotions reflecting human likeness [7].

While chatbots arguably resemble human communication through their natural language processing capabilities and dialogical interaction [17], chatbot human likeness may be manipulated through the inclusion of humanlike *cues* in the chatbot design [4], that is, design features intended to strengthen users' anthropomorphizing the chatbot. Such design features may concern the visual appearance of the chatbot, such as providing a humanlike avatar, the presentation of the chatbot, such as having it present itself with a human name, the communication style of the chatbot, such as presenting the information in an informal tone of voice, and the communication intelligence of the chatbot, such as its capabilities to mimic a skilled human conversationalist.

Strengthening human likeness in chatbots has been shown to entail a range of potentially beneficial effects for service provision. For example, Go and Sundar [15], in an experimental study of chatbots in the e-commerce domain, found manipulation of human likeness to be associated with changes in user satisfaction and perceived chatbot expertise. Furthermore, Jain et al. [20], in a study of chatbot interaction design, found users to desire chatbot interactions that resemble conversations with humans.

Chatbot human likeness may also hold implications for user behavior and chatbot interaction outcomes. Adam et al. [2], in an experimental study in the e-commerce domain, found increased human likeness in the chatbot to be associated with increased user compliance during interaction. Park et al. [28], in an experimental study, found chatbot human likeness to impact willingness to donate to a fundraising initiative. Hence, chatbot providers may utilize humanlike design cues in chatbots to impact user behavior in a direction considered desirable from the point of view of the provider.

2.3 Chatbot Transparency Requirements

While human likeness may be desired by chatbot users [20], the potential for humanlike design cues in chatbots to unduly impact user perceptions and behavior have caused concern [27], e.g., regarding users' potential confusion of whether they interact with a chatbot or a human [30]. Such confusion could bias user decision-making during chatbot use or induce erroneous user beliefs concerning chatbot capabilities.

For AI-systems, transparency is considered a key requirement in ethics guidelines [18], including that of the EC high level expert group on trustworthy AI [8]. In the latter, transparency is defined so as to concern data, systems, and AI business models, and it is particularly noted that users should "be aware that they are interacting with an AI system, and must be informed of the system's capabilities and limitations".

For chatbots, this requirement implies a requirement on the part of the service provider to ensure that users are properly informed that they are interacting with a chatbot. This requirement has been formalized in the proposed European AI Act [10]. Here, AI service providers are obliged to ensure that users are aware that they are interacting with an AI-system and not a human service person. Chatbots are explicitly mentioned with regards to this transparency obligation.

3 Research Questions

In consequence of the potential implications of chatbot human likeness to user perceptions and behavior, as well as the transparency requirements for chatbots, it is important to know how users perceive and interact with government chatbots and, specifically, whether and how these perceptions and interactions suggest that users anthropomorphize such chatbots. In response to this knowledge need, we explicated the following research question:

> **RQ1:** How do users and service providers consider chatbots as humanlike interfaces to government information and service?

Furthermore, since chatbot human-likeness may impact user behavior, it is also important to explore whether and how users' interactions with government chatbots suggest a tendency to anthropomorphize such chatbots. In response to this, we asked:

> **RQ2:** How do users interact with government chatbots? And does such interaction suggest a tendency to anthropomorphize such chatbots?

4 Method

4.1 Research Design

In response to the research questions, we set up a three-component research design consisting of two qualitative interview series and an analysis of chatbot dialogues. In the first interview series, we interviewed users of the chatbot. In the second series, we

interviewed government representatives with responsibilities for chatbot implementations. In the dialogue analysis, we reviewed >2600 citizen chatbot dialogues.

Through this multi-method approach, we were able to gain rich insight into the research questions, combining the perspectives of users and government representatives with data from actual chatbot interactions.

4.2 The Case: A Municipality Chatbot

The study was conducted in the context of a specific chatbot: the municipality chatbot 'Kommune-Kari'. This chatbot is provided for service triage rather than service negotiation [25], and is available to citizens in about 100 Norwegian municipalities; about one third of the Norwegian population. The chatbot has been operational since 2017 and engaged in several hundred thousand dialogues in 2022.

The chatbot provides access to municipality information and services through a text-based chat user interface. It is provided by Prokom and based on the boost.ai conversational platform. The chatbot is implemented as an intent-based solution [23] leveraging a machine learning model to predict users' intents based on their textual input. Following intent prediction, the chatbot provides a predefined response through a rule-based approach. This response typically includes options for follow-up or refinement of the answer through button or free text interaction. The chatbot is set up for anonymous use. It does not have access to a user profile, but provides the users with general information about the municipality and its services. Information is provided either directly through the chatbot or by links to relevant sources on the municipality website or elsewhere.

Of particular interest to this study, the chatbot is presented through a human-like cartoon avatar resembling a female face (Fig 1). When activated, the chatbot presents itself in a welcome message greeting the user, states its own name (Kari – a common Norwegian female name), explains that it is a chatbot, encourages the users to phrase their requests in a concise manner, and reminds the user that the service is for anonymous use only. The chatbot avatar, name and presentation was decided by the chatbot provider following an analysis of current practice.

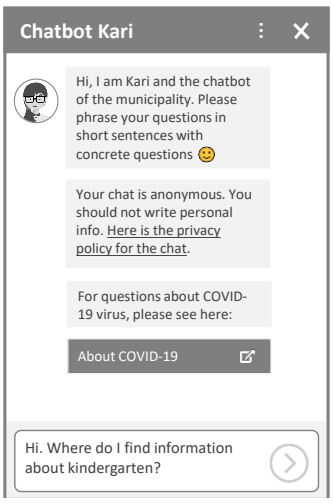

Fig. 1: Municipality chatbot welcome message. Translated from Norwegian.

The chatbot is arguably a good case for exploring the research questions. First, it has been operational for several years and is a much-used chatbot. Furthermore, it is used for citizen interaction with a relatively large set of government actors and broad range of information seeking. The large number of municipalities in which it is in use allows for variation in how it is implemented and perceived. The broad range of information and services provided through a municipality, spanning, e.g., healthcare, education, renovation, planning and construction, as well as sport and leisure, implies that the chatbot is used for a broad range of citizen requests – which is valuable when exploring user perceptions and interactions.

4.3 Interviews with Citizens and Municipality Representatives

Participants and Recruitment. The study included 16 citizen participants and 18 municipality representatives.

Citizen participants were recruited from four municipalities where the chatbot is implemented, four from each municipality. The municipalities varied in size and regional location to ensure breadth in the data collected. The citizen participants were recruited by a panel provider, Norstat; eight females and eight males in the age range of 21-68 years (median: 47,5). To ensure recent experience with the chatbot, all citizen participants were asked to use the chatbot at least twice prior to the interview. Most participants (11) reported also to have used the chatbot prior to this preparatory use.

The municipality representatives were recruited through the network of the chatbot service provider. The municipalities represented 18 different municipalities or municipality constellations of various sizes and regional locations. All participants held roles in the municipality with responsibility for the chatbot. They were mainly organized as part of the communication team (12) or the citizen service centre (5). Most municipalities had been involved in the implementation of the municipality chatbot, hence holding substantial experience on its use in the municipality.

Interview Procedure. The interviews were semi-structured and conducted individually with the participants, in total 34 interview sessions. The interview guide was set up to address ethical and societal aspects of chatbot implementation in government, and focused on perceptions and implications of chatbot human-likeness as one of several topics. In this paper we focus specifically on the participants' reflections on the human-likeness of the chatbot. Analyses and findings focusing on other aspects of the interview datasets, will be presented elsewhere [16]. The interviews were conducted by video link, through Microsoft Teams. The interviews with the citizen participants lasted 20-40 minutes. The interviews with the municipality representatives each lasted about 1 hour.

Analysis process. The interviews were audio recorded and transcribed. The transcripts were analyzed through thematic analysis following the guidelines of Braun and Clarke [6]. To maintain quality in analysis, the citizen interviews were analyzed separately from the interviews with the municipality representatives. The analysis of the interviews with the citizens was led by the second author, the analysis of the interviews with the municipality representatives was led by the first author.

4.4 Dialogue Analysis

An analysis of citizen dialogues with the chatbot was conducted to understand how users interact with government chatbots. The analysis was conducted as part of a larger research effort to understand user interaction with government chatbots through the lens of an analysis framework for customer service. This larger research effort will be published elsewhere [12]. In the study presented here, we combine the dialogue analysis with findings from interviews data to shed light on whether users' interactions with a government chatbot suggest a tendency to anthropomorphize such chatbots.

Dialogue Sampling. Dialogues between citizens and the chatbot were sampled from six different municipalities. The municipalities varied in size and regional location, as well as how they had chosen to implement the chatbot, so as to enable substantial variation in chatbot users and interactions.

Data were sampled over a two-month period. For sufficient breadth in request topics and communication styles, we sampled between 4-500 dialogues from each municipality, 2663 in total. The sampling was conducted by the chatbot service provider in line with the chatbot terms of use and following confidentiality agreements with the researchers. To ensure dialogue anonymity, all sampled dialogues were checked for person data by personnel at the chatbot service provider prior to analysis.

Analysis Process. The dialogues were first analyzed from perspective of dialogue descriptives, including user request characteristics, e.g., message brevity and predicted user intent, and dialogue characteristics, such as the number of dialogue turns and indications of understandability issues.

Furthermore, and key to this study, the dialogues were analyzed to identify the *users' communication style* in the interactions, that is, their tendency to engage in utilitarian or socially oriented interactions [14]. Dialogues with a socially oriented communication style more closely resemble dialogues to be expected between human conversationalists, with use of politeness markers and first- or second person pronouns. Dialogues with a utilitarian communication style lack one or both of these characteristics.

4.5 Research Ethics

The presented research has been conducted in line with ethical guidelines for research involving human subjects. Interviews were voluntary and only conducted following informed consent, and upon approval of the research organization data protection officer. Chatbot dialogue analyses were conducted in line with terms of use and only following manual checks of user anonymity.

5 Results

In the results section, we first present findings from the interviews with citizens, followed by findings from the interviews with municipality representatives, and the findings from the dialogue analysis.

5.1 Results from the Interviews with Citizens

Chatbot Considered Efficient Complementary Channel. In the citizen interviews, the participants described themselves as highly efficiency-seeking and goal-oriented in their interactions with the chatbot. Such efficiency concerned fast response (16) and navigational support (11). The participants typically also noted that they expect the chatbot to be able to help them with simple, general requests rather than complex and personal questions.

> *I think of it as a more advanced search engine. That can help you to sort out what you need to know. You may ask general questions that everyone would ask […]. And then you get answers without having to search the website forever* (C1).

In line with this, the participants typically described the chatbot as and additional self-service channel and noted that they appreciated the opportunity for self-service provided by the municipality's digital channels (14). The participants, furthermore, reported on not being worried that digitalization will remove their opportunity to get in touch with humans in the municipality. Rather than a substitute for human communication, the participants considered the chatbot as an additional public service channel, allowing them to get swift responses to general requests around the clock. Participants indeed underlined the importance of available human resources when needed, but they reported to prefer self-service for general requests (14). The following quote illustrates this perspective:

> *Preferably a chat service, because using the phone you have to wait for them to answer the phone, and they don't always respond, […]. So, I find that the chat service is better. You get a response much faster* (C5).

Humanlike Chatbot Features May Be Pleasant but Not Important. While the chatbot is presented through a humanlike avatar image, a humanlike name, and communicate in an informal style, all participants noted that such humanlike features of the chatbot have little or no implications for their use of it. Some noted that the humanlike features did *not* matter to them whatsoever whereas others said they appreciated such features. The latter participants noted that this made talking to the chatbot more personal and that they found it somewhat "fun" that the chatbot has humanlike features. The following quote illustrates this latter perspective:

> *I wouldn't say it is important to me […]. But I thought some of the features were fun. That* Kommune-Kari *is a character, and when I did a search on football and sports clubs it said, "football is fun!" with a football emoji. There are some fun features like that […] it has some personality to it* (C6).

Yet, although some reported to appreciate humanlike features in the chatbot, all participants argued that these had little or no bearing on their assessment or use of the chatbot.

Machine Nature of Chatbot Clear but Interaction May Improve on Experience. All participants expressed that they found it clear that they were interacting with a chatbot and not a human service provider, and they were aware that the chatbot provides general rather than personalised responses to inquiries. The participants further found it relatively easy to understand how to interact with the chatbot. Yet, participants also pointed out that the chatbots' usefulness had increased over time, as they had become more experienced chatbot users, and better understood how to pose questions (*i.e.*, shorter sentences and/or single words). To illustrate, one participant noted that:

> *I tried some longer sentences, and then I realised that it didn't work, and then I started using very short sentences, and sometimes just one word* (C6).

This may suggest that the human likeness and the communication style of the chatbot initially may lead users to ask longer and more complicated questions. Participants did not express this as a drawback, though, noting that they found it easy to figure out how to use the chatbot.

The participants also pointed out that they know someone, or assumed that there may be someone, who may be challenged to use a chatbot (10). One concern among these participants was that other users might fail to understand the chatbot interaction format, where the chatbot input should be presented in a concise manner. That is, they foresaw that some users might use the same interaction strategies that they would when interacting with a human, leading to a suboptimal outcome. Related to this, one participant made specific note that the chatbot did not understand them when asking a difficult question and suggested that this may indicate difficulties for users struggling with reading or writing. Another noted as follows:

> *For chatbots you must be as short and concise as possible […]. But I helped my mother-in-law, and she had written [a very long sentence]* (C13).

5.2 Results from the Interviews with Municipality Representatives

In the interviews with the municipality representatives, the participants reported on their perceptions of citizen interactions with the municipality chatbot. Their reports were based on their experience from reviewing interaction logs with the chatbot and on feedback from citizens – for example through the citizen service centre.

Chatbot for Efficient Interactions with the Municipality. The participants described the chatbot as a useful navigation support for users (13), that may simplify access to the municipality information and services (11). The participants noted that they considered the chatbot to be particularly useful for general information requests (13), while chatbot responses for precise or personal questions could be insufficient (9).

> *We find the chatbot to answer quite well, with some important limitations. Specific questions and the like do not work that well […] But very good at general questions. Short, general questions work well.* (M3)

Machine Nature of Chatbot Clear to Most Users. When asked about implications of the chatbot human likeness, the participants considered this to be limited. Several of the participants pointed out that users typically understand what a chatbot is and which opportunities and limitations it entails (7).

> *Most understand that it is a chatbot. This is actually very clearly explained. That it is not a human. Even though it may appear like one.* (M9)

The participants also noted that the chatbot was clearly different from a human service provider also in terms of its relative lack of flexibility (8). While a human service provider may show high levels of flexibility in adapting to the requests and messages of a user, the chatbot does not have such conversational intelligence. Hence, the risk of confusing the chatbot for a human was seen as limited.

> *I believe it is pretty clear that it is not a person, that is, the dialogue is very structured and it includes new links, new buttons, as you ask new questions.* (M15)

Some of the participants also made note of the chatbot being clear on its limitations (4). For example, by asking for questions to be provided in short sentences, or by clarifying to the user in cases of insufficient prediction confidence regarding user intents.

Some may not Understand Chatbot Interaction or Capabilities. While the participants argued that most users understand that the chatbot is a machine and that they should interact with it differently than with a human service person, they also typically noted that some users might misunderstand.

Most participants reported to have observed that some users interact with the chatbot in a manner suggesting that they do not understand how to ask questions to a chatbot in a productive manner (13). This could, for example, be that users were observed to not ask direct questions, but instead presented their inquiries over multiple questions; something that is challenging for the chatbot to interpret correctly.

> *There are some who ask the chatbot as if they believe they are about to chat with a real human. This can be seen in the way people ask questions and follow up, can be seen sometimes.* (M8)

Some participants suggested that chatbot human likeness may lead users to get false expectations regarding chatbot capabilities. For example, they may believe that the chatbot has higher conversational flexibility than it actually does, or they may ask questions at a level of complexity that the chatbot cannot answer.

> *I find that sometimes they think of her as a human […] they write long sentences […] then the chatbot does not work and it becomes a source of irritation* (M14)

Signs of Increasing Maturity in Users. While not all users may understand the chatbot interaction or capabilities, some participants noted what they saw as an increased maturity in user interaction with the chatbot over time (4). As an example of this, some noted a reduction in playful or exploratory chatbot interactions and a general tendency to fewer users engaging with the chatbot as if it were a human.

> *Before we experienced perhaps someone believed it was a real person behind it. This we could see in the questions coming in. But I do not see this much anymore.* (M4)

Further Reflections on Implications of Human Likeness. Finally, reflecting further on implications of human likeness in the chatbot, some participants also noted that chatbot human likeness could potentially lead to changes in how citizens engaged with the municipality information and services. Some participants noted that a humanlike chatbot may entail that users see the service interaction as more personal and – thereby – more attractive (4).

> *The dialogue interaction I believe is beneficial. To feel that you talk to someone, this is a human need basically.* (M12)

Others noted that chatbot human likeness may reduce citizens threshold for getting in touch with the municipality, because it is easier to ask questions in an interaction format resembling that of human conversation (3).

> *I believe that the threshold for getting in touch is a little lower […] We see that it is used for very much now* (M9)

Some participants also noted that the humanlike character of a chatbot may motivate users to ask more personal or specific questions than they would e.g. in a search interface. This may be beneficial as responses may be more relevant to the user, but it may also entail a challenge in cases where the user asks questions at a level of specificity to which the chatbot cannot provide an answer without knowledge of the user context.

> *We see that there may also be very personal questions in the chatbot, and hard to answer [these] in a general way* (M2)

5.3 Results from the Dialogue Analysis

In total, 2663 dialogues were included in the sample from the six municipalities. The dialogues were about equally distributed across the six municipalities involved in the analysis, ranging between 430 and 475 dialogues for each municipality.

The dialogues reflected the breadth of the information and services offered by the municipalities, including general healthcare (18%), COVID-19 (12%), general municipality information (8%), leisure (6%), applications and case processing (4%), contact information (4%), water and sewage (4%), education (3%), and renovation (3).

The dialogues provided insight into the characteristics of the user requests, the length of dialogues, and the users' communication style. We detail these in the following.

User Requests. The user requests to the chatbot were typically brief and concise. In the analyzed dialogues, the initial requests had a median length of 19 characters (25th percentile = 12; 75th percentile = 36). The vast majority of such initial requests (92%) were 60 characters or less, despite the maximum message length in the chatbot was set to 110 characters.

This implies that the user requests to the chatbot typically were highly pointed, with little detail or contextual explanations. Examples of such pointed request formulations include the following: "Where is covid test", "Registration for vaccination", "Dirty water in the tap", "When is the boating license course", and "Summer school".

Dialogue Characteristics. The citizen dialogues with the chatbot were typically brief. Most chatbot dialogues (77%) included only one user message, 15% included two messages, 8% included three or more. Moreover, the dialogue analysis identified that the users typically received useful help (65%), either by information included in the chatbot message or through information or services linked to by the chatbot. Furthermore, only 3% of the dialogues were found to indicate understandability issues such as failure to formulate requests interpretable to the chatbot or failure to make use of interaction mechanisms.

In the following, a typical conversation with only one user message is presented:

- *User:* "Status for vaccination in [municipality name]"
- *Chatbot:* "We provide the latest updates on vaccination status here: [link]"

Users' Communication Style. Dialogues were categorized as having a socially oriented communication style if including social markers such as greetings and use of first- or second person pronouns. Otherwise, the user dialogues were categorized as having a utilitarian communication style. In the analysis, we found an overwhelming proportion of the dialogues to be in a utilitarian style (95%), whereas only a small minority (5%) were in a socially oriented style.

To illustrate the two communication styles of the users, we include below examples of user requests in utilitarian and socially oriented styles respectively:

- *Utilitarian style user request:* "Status for vaccination in [municipality]"
- *Socially oriented style user request:* " Hi. Where in the municipality may I take a rapid test?"

While the low proportion of dialogues with a socially oriented communication style was consistent with the brevity or user requests and short dialogues, it was a surprise as previous research on customer service chatbots has found higher prevalence of socially oriented dialogues [14].

6 Discussion

In the following, we first discuss citizen and municipality perspectives on chatbots with regard to human likeness. Second, we discuss citizen behavior during chatbot interactions and how this may shed light on any anthropomorphizing of the chatbot. Finally, we address implications for theory and practice and reflect on limitations and future research.

6.1 Citizen and Municipality Representative Considerations of Human Likeness

Our findings suggest that citizens typically have an adequate understanding of chatbot capabilities and limitations. Specifically, it is interesting that the citizens were found to hold adequate expectations on chatbot capabilities, since previous research has suggested that chatbot human likeness may induce inflated capability expectations [27]. While the chatbot in this study had marked humanlike characteristics [4, 15] in visual appearance and communication style, with a humanlike avatar, human name, and informal tone of voice, the users did not find this confusing or problematic. Rather, the users argued that the chatbot human likeness was not important for their use of it, and a low rate of understandability issues was found in the chatbot dialogues.

Chatbot human likeness was, however, suggested by some of the citizen participants to have some pleasurable aspects. This is in line with previous findings in the context of customer service, where users have been found to appreciate chatbot human likeness even though humanlike characteristics are not considered key [13]. The relative lack of perceived importance of human likeness in government chatbots may be due to the

highly goal-oriented user of such chatbots [1], which is in line also with the participants accentuation of potential efficiency benefits in the chatbot. Hence, while human likeness may be highly important in other chatbots [20], such as for example companion chatbots [34], this characteristic do not seem to be important to government chatbots.

However, the study participants noted that some users may be confused by the chatbot human likeness, shown for example in terms of inadequate strategies for chatbot interaction. Such confusion is reminiscent of what has been found in research on voice-based agents where inexperienced users fail to understand how to interact with a conversational user interface [22]. Possibly, confusion due to chatbot human likeness may be caused by lack of experience. This assumption is supported by our participants noting increased maturity over time for chatbot interactions.

6.2 Reflections of Anthropomorphizing in Citizens' Chatbot Dialogues?

Citizen and municipality representative reports on chatbot human likeness typically not confusing its users, are corroborated by our findings from the dialogue analysis. In these analyses, the concise requests of users and the typically short dialogues suggest that users are highly goal oriented and that their mode of interaction is one of efficiency. This use of chatbots aligns with previous findings, where utilitarian motivation has been identified as users' main motivation to engage with chatbots [5]. Furthermore, efficient interactions are in accordance with the aim of chatbots for service triage [25], where the chatbot is used to identify and access needed information and services.

Furthermore, the utilitarian orientation reflected in the chatbot dialogues are indicative of chatbot human likeness not imposing on users a tendency towards anthropomorphism. On the contrary, the prevalence of socially oriented interactions – found in only 5% of the analyzed dialogues – was lower than in a similar analysis of chatbots for customer service [14]. Also, the social orientation in users' communication style was surprisingly low contrasted with observations in other domains, such as conversational search [21].

In conclusion, the brevity in user requests, efficiency in dialogues, and prevalence of a utilitarian communication style all point in the same direction as the findings from the citizen and municipality representative interviews. In spite of humanlike design cues in the studied chatbot, users perceive and engage with this in a way that suggests a concern for efficiency and effectiveness, rather than one of anthropomorphizing.

6.3 Implications for Theory and Practice

The study findings hold several implications for theory and practice. We note the following implications to be of particular interest for theory:

- **Chatbot objective may determine user perceptions and behavior:** Much previous work has addressed how chatbot design may impact user perceptions and behavior [e.g., 19]. Our findings complement this, by suggesting that also chatbot objective – e.g., to provide government service triage [25] – may potentially determine user perceptions and communication style. This is in particular seen when contrasting our findings on communication style to previous work [e.g. 14, 21].

- **Humanlike chatbot design cues may have limited impact on user behavior:** Previous work has shown that user perceptions and behavior may be determined by humanlike design cues in the chatbot [e.g., 4, 15]. Our findings indicate that while humanlike design cues concerning chatbot presentation and appearance may hold implications for user perceptions, they may have limited impact on user behavior, e.g., in a government chatbot for service triage.

For practice, we see the following implications to be of particular interest:

- **A chatbot for government service triage may comply with transparency requirements:** Chatbots are expected to comply with transparency requirements, as per ethics guidelines [18] and regulatory frameworks [33], that is, it should be evident to users that they are interacting with a chatbot. Our findings suggest that complying with such requirements is indeed feasible for chatbots for service triage.
- **Some users may nevertheless fail to understand the chatbot:** In spite of the machine nature of a chatbot typically being clear to users, the conversational interaction may lead to some confusion, in particular for inexperienced users. Design of government chatbots should take into consideration how to also support users who are inexperienced with chatbots so as to avoid confusion.

6.4 Limitations and Future Research

While the presented study provides needed knowledge on citizen perceptions of human likeness in chatbots, and their limited anthropomorphizing of such chatbots, the study has important limitations. These limitations suggest paths for future research.

First, while the study employs different methods for data collection and analysis, it only involves one chatbot, the municipality chatbot 'Kommune-Kari'. While this chatbot is implemented in a large number of municipalities, it far from covers chatbots at the level of all government services and organizations. Hence, the findings from this study would benefit from being complemented with findings from other chatbots.

Second, the study is limited to a single country, Norway, which may limit findings in terms of the characteristics of citizen population and structure of government. We foresee future studies replicating our approach in other countries or regions.

Third, the study only addresses user perceptions and interactions with a government chatbot for service triage, following the typology of Makasi et al. [25]. While this chatbot type arguably is most commonly deployed in current government service provision, this is likely to change in the future. Both in terms of the increasing availability of chatbots for personalized support, and also the emerging availability of chatbots based on large language models with improved capabilities for humanlike interaction. There will arguably be a need for continued research into the implications of government chatbot human likeness as more advanced technology is taken up in government chatbots and chatbot capabilities for humanlike and personalized interactions change. We hope our study is a useful initial contribution to this important area of research.

Acknowledgement. The study was supported by EC H2020 grant no. 101004594, ETAPAS.

References

1. Abbas, N., Følstad, A., Bjørkli, C.A.: Chatbots as part of digital government service provision–a user perspective. In: CONVERSATIONS 2022 (pp. 66-82). Springer, Cham (2022).
2. Adam, M., Wessel, M., Benlian, A.: AI-based chatbots in customer service and their effects on user compliance. Electronic Markets, 31(2), 427-445 (2021).
3. Amiri, P., Karahanna, E.: Chatbot use cases in the Covid-19 public health response. Journal of the American Medical Informatics Association, 29(5), 1000-1010 (2022).
4. Araujo, T.: Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. Computers in Human Behavior, 85, 183-189 (2018).
5. Brandtzaeg, P.B., Følstad, A.: Why people use chatbots. In: Proceedings of Internet Science: - INSCI 2017 (pp. 377-392). Springer, Cham (2017).
6. Braun, V., Clarke, V.: Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77-101 (2006).
7. Epley, N., Waytz, A., Cacioppo, J. T.: On seeing human: A three-factor theory of anthropomorphism. Psychological Review, 114(4), 864–886 (2007).
8. European Commission: Ethics guidelines for trustworthy AI. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai (2019)
9. European Commission: AI watch - artificial intelligence in public services. Technical report. https://publications.jrc.ec.europa.eu/repository/handle/JRC120399 (2020)
10. European Commission: Proposal for a regulation on a European approach for artificial intelligence (No. COM(2021) 206 Final). https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence (2021)
11. Følstad, A., Araujo, T., Law, E.L.C., Brandtzaeg, P.B., Papadopoulos, S., Reis, L., ... Luger, E.: Future directions for chatbot research: an interdisciplinary research agenda. Computing, 103(12), 2915-2942 (2021).
12. Følstad, A., Bjerkreim-Hanssen, N.: User interactions with a municipality chatbot – lessons learnt from dialogue analysis. Research paper submitted for review (2023).
13. Følstad, A., Skjuve, M.: Chatbots for customer service: user experience and motivation. In: Proceedings of Conversational User Interfaces - CUI 2019 (pp. 1-9). ACM, New York (2019)
14. 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).
15. 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).
16. Grøndahl Larsen, A., Følstad, A.: The impact of chatbots on public service provision – a qualitative interview study with citizens and public service providers. Research paper submitted for review (2023).
17. Hall, E.: Conversation Design. A Book Apart, New York (2018).
18. Hagendorff, T.: The ethics of AI ethics: An evaluation of guidelines. Minds and Machines, 30(1), 99-120 (2020).
19. 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).
20. Jain, M., Kumar, P., Kota, R., Patel, S.N.: Evaluating and informing the design of chatbots. In: Proceedings Designing Interactive Systems – DIS '18 (pp. 895-906). ACM, New York (2018).

21. Liao, Q.V., Geyer, W., Muller, M., Khazaen, Y.: Conversational interfaces for information search. Understanding and Improving Information Search: A Cognitive Approach (pp. 267-287). Springer, Cham (2020).
22. Luger, E., Sellen, A.: "Like having a really bad PA" The gulf between user expectation and experience of conversational agents. In Proceedings of CHI '16 (pp. 5286-5297). ACM, New York (2016).
23. Luo, B., Lau, R. Y., Li, C., Si, Y.W. A critical review of state-of-the-art chatbot designs and applications. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 12(1), e1434 (2022).
24. Maedche, A., Legner, C., Benlian, A., Berger, B., Gimpel, H., Hess, T., ... Söllner, M.: AI-based digital assistants: Opportunities, threats, and research perspectives. Business & Information Systems Engineering, 61, 535-544 (2019).
25. Makasi, T., Nili, A., Desouza, K.C., Tate, M.: A typology of chatbots in public service delivery. IEEE Software, 39(3), 58-66 (2021).
26. 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 Hawaii International Conference on System Sciences. University of Hawaii, Manoa (2022).
27. Murtarelli, G., Gregory, A., Romenti, S.: A conversation-based perspective for shaping ethical human–machine interactions: The particular challenge of chatbots. Journal of Business Research, 129, 927-935 (2021).
28. Park, G., Yim, M. C., Chung, J., Lee, S.: Effect of AI chatbot empathy and identity disclosure on willingness to donate: the mediation of humanness and social presence. Behaviour & Information Technology, https://doi.org/10.1080/0144929X.2022.2105746 (2022)
29. Rapp, A., Curti, L., Boldi, A.: The human side of human-chatbot interaction: A systematic literature review of ten years of research on text-based chatbots. International Journal of Human-Computer Studies, 151, 102630 (2021).
30. Ruane, E., Birhane, A., Ventresque, A.: Conversational AI: Social and ethical considerations. In: Proceedings of AICS 2019 (pp. 104-115). DBLP, Dagstuhl (2019).
31. Sagstad, M. H., Morken, N. H., Lund, A., Dingsør, L. J., Nilsen, A. B. V., & Sorbye, L.M. (2022). Quantitative user data from a chatbot developed for women with gestational diabetes mellitus: observational study. JMIR Formative Research, 6(4), e28091.
32. Segura-Tinoco, A., Holgado-Sánchez, A., Cantador, I., Cortés-Cediel, M.E., Bolívar, M.P.R.: A conversational agent for argument-driven e-participation. Proceedings of DG.O 2022 (pp. 191-205). ACM, New York (2022).
33. Schaake, M.: The European Commission's Artificial Intelligence Act. Policy Brief, Stanford HAI (2021).
34. Skjuve, M., Følstad, A., Fostervold, K.I., Brandtzaeg, P.B.: A longitudinal study of human–chatbot relationships. International Journal of Human-Computer Studies, 168, 102903 (2022).
35. Van Noordt, C., Misuraca, G.: New wine in old bottles: Chatbots in government. In: Proceedings of Electronic Participation 2019 (pp. 49-59). Springer, Cham (2019).
36. van Noordt, C., Misuraca, G.: Artificial intelligence for the public sector: results of landscaping the use of AI in government across the European Union. Government Information Quarterly, 39(3), 101714 (2022).
37. Verne, G.B., Steinstø, T., Simonsen, L., Bratteteig, T.: How can I help you? A chatbot's answers to citizens' information needs. Scandinavian Journal of Information Systems, 34(2), 7 (2022).