The Importance of a Proper Introduction

A master thesis within the field of artificial intelligence on user expectations and evaluations of chatbot introductions.

Sammie Smaak
Utrecht University
Master Artificial Intelligence
5680344

Dr. Charlotte van Hooijdonk
Utrecht University
First supervisor

Frank Smit
OBI4wan B.V.
Daily supervisor

Dr. Ruud Hortensius
Utrecht University
Second supervisor

July 22, 2021

1 Abstract

Technological advances have made it possible for chatbots to be used in a variety of settings, like customer service. Despite this rise in use, users remain skeptical towards chatbots. One way to reduce this skepticism is to properly introduce the chatbot. A chatbot introduction is used to familiarize both the chatbot and its functionality and features towards the user. It often does so, using a header (the bar at the top of the chat), several first messages, and a way for the user to respond to this. This study focuses on the impact this introduction has on the expectations of chatbot users and the evaluation of the chatbot after users communicated with it. First, a content analysis has been performed on 48 Dutch customer service chatbots to find what features are present in their introduction. The content analysis showed a lot of variety in the chatbots' anthropomorphism (the presence of human traits), ranging from human to nonhuman. The most common introduction features in the header were a picture, subtext below the main title, and a button. The first messages on average contained twoa messages with a small picture of a human. Also, a question and an introduction were present in most chatbots. To respond to these questions, the average Dutch customer service chatbot used text, without a limit. Then a participant study with a human and a nonhuman chatbot was conducted to show how the amount of anthropomorphism in the introduction influences users' expectations and evaluation. The participants first saw a screenshot of either a human or a nonhuman chatbot introduction. Then they had to answer questions regarding the competence, perceived ease of use, and warmth of the introduction. After answering these questions, the participants had to use the same type of chatbot to order a bouquet of flowers for a friend. Subsequently, questions regarding the competence, perceived ease of use, and warmth of the chatbot had to be answered. The users were also asked about their intention to use the chatbot again in the future. The user study showed that anthropomorphism significantly influences the expectation and evaluation of chatbots when it comes to the competence of the chatbot and its perceived ease of use. Meaning that the chatbot with a lower amount of anthropomorphism was expected and evaluated as more competent and perceived as easier to use when compared to the chatbot with higher anthropomorphism.

Contents

1 Abstract **2**

2 Executive summary **4**
- 2.1 Introduction 4
- 2.2 Literature background 4
- 2.3 Content analysis 4
- 2.4 Participant study 5
- 2.5 Conclusion 5

3 Introduction **6**

4 Literature background **7**
- 4.1 Chatbots in customer service . . 7
- 4.2 Factors that influence chatbot use 7
- 4.3 Features in chatbot introduction 8
- 4.4 Using frames and anthropomorphism to communicate chatbots' competence 9
- 4.5 Using metaphors to communicate chatbots' warmth and competence 9
- 4.6 The impact of design features to communicate chatbots' warmth and competence 10
- 4.7 Research questions and hypotheses 11

5 Content analysis **11**
- 5.1 Data collection 11
- 5.2 Annotation 11
- 5.3 Inter-annotator agreement 12
- 5.4 Data analysis 13
- 5.5 Results 13
  - 5.5.1 Header 13
  - 5.5.2 First messages 13
  - 5.5.3 Response options 14
- 5.6 Qualitative analysis 14
- 5.7 Conclusion 15

6 Participant study **16**
- 6.1 Design 16
- 6.2 Chatbots 16
- 6.3 Participants 16
- 6.4 Materials and procedure 16
- 6.5 Data analysis 18
- 6.6 Results 19
  - 6.6.1 Participant expectations . 19
  - 6.6.2 Participant evaluation . . 19
  - 6.6.3 Expectations versus evaluations 19
- 6.7 Conclusion 20

7 Discussion **21**
- 7.1 Implication of the results 21
- 7.2 Limitations and future research . 21
  - 7.2.1 Limitations of the content analysis 21
  - 7.2.2 Limitations of the participant study 21
  - 7.2.3 Future research 22
- 7.3 Recommendation 22
- 7.4 Relevance for the field of AI . . . 22
- 7.5 Closing words 22

Appendices **27**

A Code book **27**

B Annotation table **33**

C Questionnaire **39**

D Website **41**

E Chatbot conversation blueprint **42**

2 Executive summary

2.1 Introduction

Customer services are increasingly using chatbased services. Companies use chatbots for several reasons, such as providing a 24/7 service and responding to frequently asked questions of customers. However, users of chatbots are often not satisfied with the communication with a chatbot. To manage users' expectations about the capabilities of the chatbot, the chatbot must be properly introduced. This study aims to find out how expectations and evaluations are influenced by design features in the introduction of the chatbot.

2.2 Literature background

According to the "computers are social actors" (CASA) principle, chatbots are seen as social entities that users respond socially to (Nass et al., 1994). That is why there is a lot of research on the amount of anthropomorphism (the presence of human traits in inanimate objects) in chatbots and the way users respond to this. This research shows that adding human elements to chatbots, like a human way of talking, seems to have a positive effect on the evaluations of the chatbot user (Gnewuch et al., 2018b; Liebrecht and Van Hooijdonk, 2019; Liebrecht and Van der Weegen, 2019; Ciechanowski et al., 2019).

A topic within chatbot communication that has been barely studied is the introduction of the chatbot. Within this introduction, the chatbot can introduce itself to the user and clarify how it can assist the user. This is the first interface users see when using the chatbot for the first time and often consists of a header at the top of the conversation, first messages from the chatbot, and a way the user can respond to these messages. Based on this introduction, the user will form expectations regarding the competence and warmth of the chatbot (Khadpe et al., 2020; Araujo, 2018).

The first goal of this study was to analyze the features in the introduction currently present in Dutch customer service chatbots.

The second goal of this study was to research the influence of the amount of anthropomorphism in chatbot introductions on the expectation and evaluation of the chatbot user.

2.3 Content analysis

For the content analysis, the introductions of 48 Dutch customer service chatbots were collected. For the analysis, the chatbot introductions were annotated per section. The chatbot introduction was divided into three sections for this analysis: a header, the first messages of a chatbot, and the way the user could respond. The analysis showed that the header often consisted of a picture (either human, nonhuman, or a logo), subtext (smaller text box underneath the main text section), and a button.

The results for the first messages showed that on average companies use two messages to introduce the chatbot. These messages are accompanied by a picture. Also, the first messages consist of utterances in which the chatbots introduce themselves. Moreover, these messages contain a question.

In the last section, the response, the most common response option was text without a text limit. With these findings, I built a human and a nonhuman introduction for a Dutch customer service chatbot to use in the participant study.

2.4 Participant study

The nonhuman chatbot used in this study was as following: a nonhuman picture and just the word "chatbot" in the header. The first messages contained an introduction, which was followed by an explanation of the chatbots' purpose. Also, it contained a message explaining that humans will be contacted if needed, after which the first question was asked to help the user. The user then had to respond to this using buttons.

The human chatbot used in this study had a human picture, the word "chatbot", and a human name in the header. Then the first messages contained a greeting, and an introduction, which was followed by an implicit explanation of the chatbots' purpose. Then the chatbot ended with a question to start the conversation. The user then had to respond to this with text.

The participant study showed the participants a screenshot of the chatbot introduction (either human or nonhuman), after which they had to fill in a questionnaire asking them about the competence, perceived ease of use, and warmth of the chatbot. Then the participants had to use the same version of the chatbot (human or nonhuman) to order a bouquet of flowers for a friend. Only the introduction of the chatbots was manipulated, the rest of the conversation was the same. When the task was completed, the participants got another questionnaire that asked them about the chatbots' competence, perceived ease of use, warmth, and their intention to use the chatbot again.

The results showed that after seeing the screenshot of the chatbot introduction, participants expected the nonhuman chatbot to be easier to use. After using the chatbot they evaluated the nonhuman chatbot to be more competent, easier to use and had a higher intention to use it in the future, compared to the human chatbot. Participants' evaluations about competence, perceived ease of use, and warmth were higher than their expectations. Overall, the nonhuman chatbot was seen as more competent and easier to use.

2.5 Conclusion

The first goal was to find the features in Dutch customer service chatbot introductions. The most common features found for the header were a picture, subtext, and a button. In the first messages, these features were found in two messages on average. The features found were an introduction, a picture, and a question. To which the user could most often respond with text without restrictions.

The second goal was to study the influence of anthropomorphism on the expectations and evaluations of chatbot introductions. From the study, it follows that less anthropomorphism influences the expectations and evaluations of chatbots and causes higher perceived competence and ease of use. Also, participants had lower expectations than their evaluations of the chatbot. This means that the introduction influences the user when using a chatbot.

So when designing chatbot introductions, it is important to keep in mind that if you want your chatbot to come across as more competent and to be perceived as easier to use, a nonhuman chatbot will deliver better results. The nonhuman design used in this participant study would result in a proper way to introduce your chatbot.

3 Introduction

Technological advances have made it possible for chatbots to be used in a variety of settings, like customer service. By implementing chatbots, the customer service of companies is available 24 hours a day. Another advantage of implementing chatbots is that they can address customers' frequently asked questions, whilst leaving the difficult questions and the complaints to the human service employees (Charlton, 2013; Cui et al., 2017).

The scope of research on chatbots has mainly been on their technological performance, like their architecture and the way they use natural language processing (Paul et al., 2019; Androutsopoulou et al., 2019). However, attention is lacking on the field of user experience of chatbots (Khadpe et al., 2020). Studying the user experience of chatbots is of importance because users of chatbots are often not satisfied with their use (Spierenburg, 2021). This study will focus on one specific part of the user experience, the introduction of a chatbot and its influence on user expectations and evaluations. Based on the introduction, users could form expectations with regards to the conversation (Khadpe et al., 2020).

Previous research on chatbot introductions focused on the explanation beforehand. The participants received a textual explanation on the workings of the chatbot they were about to use, e.g. *You are about to talk with a virtual agent.* (Araujo, 2018). However, there are no studies on the design features in the introduction of chatbots for customer service purposes. It is plausible that these features will evoke different expectations from chatbot users.

This study focuses on the features present in the introduction of Dutch customer service chatbots. This introduction consists of three parts: the header, the first messages, and the response as shown in Figure 1.

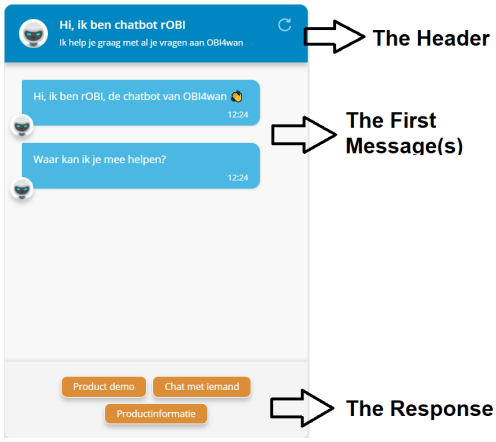


Figure 1: Chatbot Introduction.

This study is guided by the research questions RQ1 and RQ2. RQ 1: *Which features do Dutch customer service chatbots use in their introduction?* and RQ2: *How does the anthropomorphism in a chatbot's introduction influence the expectation and evaluation of users?*. These questions will be answered by performing two different studies. The first study is a content analysis in which verbal and non-verbal features are analyzed in the header, the first message(s), and the response of 48 Dutch customer service chatbots. The second study entails a participant study in which the verbal and non-verbal features in the introduction of a chatbot are systematically manipulated. For this, two chatbots are built. In an experiment, users see the introduction of the chatbot after which they answer questions about their expectations. Subsequently, they perform a task by conversing with the chatbot after which they evaluate the chatbot. In these chatbots, features are altered in

the introduction. With these chatbots, I measure the difference in the warmth and competence that users experience.

For the building of chatbots for the participant study, I use a chatbot building platform developed by OBI4wan. A company that helps businesses with their web care, reputation management, and data analysis. This thesis is written in cooperation with OBI4wan. In exchange for the use of their tool, their knowledge, and their contacts, the answer to this research question will provide OBI4wan with a recommendation on the introduction of chatbots.

4 Literature background

4.1 Chatbots in customer service

Chatbots originally refer to software that simulates human natural language with a text-based dialogue system, to interact with users (Zumstein and Hundertmark, 2017). At first, chatbots' main goal was to mimic human conversation. However, after a lot of improvement, they are now able to show distinct and diverse characteristics (Chaves and Gerosa, 2019).

Within the customer service domain, chatbots can be used to answer frequently asked questions, which makes them perfect for handling the first questions consumers have when looking for service. Chatbots also are able to be on 24/7. Which means that they are always approachable. Even though chatbots have a lot of potential, Dutch research shows that 78% of their participants($N = 10.000$) did not receive a valid or clear answer to their question when they communicated with a chatbot (Spierenburg, 2021). This shows that a lot of improvement is still required within the field of chatbot communication.

Previous research has focused on improving chatbots' technical capabilities. Research that focuses on the programming of a chatbot, dives into whether chatbots should be a rule-based system (have a specific response for a situation) or a corpus-based system (selecting a response from a large dataset or using a neural network to find a suitable answer). Research of Serban et al. (2016) focused on a neural network approach to build a chatbot for realistic and flexible interactions. They showed that their neural network model approach could outperform other models, from which they concluded that a bigger dialogue dataset is needed to improve chatbots' performance. Research on understanding user input with named entity recognition (the extraction and annotation of real-world objects from text) could help chatbots better understand what a user is saying to them and then respond better to this (Reshmi and Balakrishnan, 2018). Next to this, research on intent recognition, also known as intent classification, is important to improve the chatbots' understanding of the users' intent in conversations (Setyawan et al., 2018). Intent recognition focuses on identifying what the users want to achieve in the conversation, which results in a better understanding of textual content and would allow for a more appropriate response from a chatbot (Kamphaug et al., 2017).

4.2 Factors that influence chatbot use

Despite these technical advances, customers can be hesitant in interacting with chatbots, as chatbots can fail to live up to their expectations (Araujo, 2018). This hesitant behavior in customers could be explained by the research of Davis. He came forward with research on the design and implementation of information systems such as chatbots and developed and tested a model to help explain how people perceive these

systems. This model was called the Technology Acceptance Model (TAM) and provides a theoretical basis for system designers to evaluate new systems prior to their implementation (Davis, 1985). The model helps to explain which variables have an impact on the users' motivation to use a system and how causally related variables are related to system characteristics and user behavior (i.e., the way users behave when using systems). The TAM shows, that before people use a system, they need enough motivation (Davis, 1985).

The motivation to use a system is driven by perceived usefulness and perceived ease of use of a system, and these two factors influence the attitude toward using a system. Perceived usefulness is the measure of how useful the user thinks the chatbot was or will be, does the chatbot help the user achieve their goal? Perceived ease of use is the measure for how difficult it was or will be to use the chatbot. Design features of a chatbot, such as the appearance of a system, influence the perceived usefulness and the perceived ease of use of a system (Davis, 1985; Zarouali et al., 2018).

Perceived usefulness and perceived ease of use are part of the cognitive aspects of the chatbot. However, these are not the only aspects that play a role within the TAM, affective features also play a role (Davis, 1985). These affective features influence the attitude of the user towards the chatbot, such as the warmth of the chatbot.

For the design features, it is essential to take into account that users see chatbots as a social actor (Nass et al., 1994). A social actor can engage in social actions. Various studies show that users appreciate it if a chatbot shows qualities or characteristics as seen in humans, also known as human attributes, like a human way of communicating (Gnewuch et al., 2018a; Liebrecht and Van Hooijdonk, 2019; Liebrecht and Van der Weegen, 2019; Ciechanowski et al., 2019). So if these design features, as explained in the TAM, entail human attributes such as a human way of communicating, the chatbot could be more acceptable to people (Davis, 1985).

The research of Appel et al. (2012) focuses on the importance of social cues during human interaction and found that chatbots that use a virtual avatar caused more social presence (which refers to a sense of being together (Biocca et al., 2003)) than when the chatbot uses a profile picture. This means that the way a chatbot is designed influences the way the user perceives the chatbot.

However, if this design incorporates too many human features, the user could have trouble finding out if the chatbot is a chatbot or a human. This tipping point is called the uncanny valley principle and is an eerie and uncomfortable feeling towards the chatbot (Ciechanowski et al., 2019).

4.3 Features in chatbot introduction

These design features are also present in the introduction of a chatbot. This can be done using verbal (e.g. words used in the introduction) and nonverbal cues (such as pictures) (Hendriks et al., 2020). Okuda and Shoda (2018) explain that the introduction can be used to explain to the user what the chatbot's capabilities are. Research that focuses on manipulating the features in the introduction of chatbots are the studies of Khadpe et al. (2020) and Araujo (2018).

4.4 Using frames and anthropomorphism to communicate chatbots' competence

Araujo (2018) focuses on the framing of the chatbot, but also on the amount of anthropomorphism (i.e., the attribution of human traits). The frame of the chatbot was how a chatbot was introduced. The frame was either an intelligent frame or a neutral frame. The intelligent frame was described as "*a virtual agent powered by artificial intelligence (AI)*" that uses machine learning and AI technology, while the neutral frame was described as a "*virtual agent*". The anthropomorphism was manipulated in the following ways: the human chatbot used informal language, used a human name, and the participant had to start and end the conversation using verbal cues usually associated with human-to-human communication (such as "hello" and "goodbye"). The nonhuman chatbot used formal or computer-like language, had a robotic name, and the participant started the conversation with "start" and "quit".

After the participants interacted with the chatbot, they filled in a questionnaire measuring mindful anthropomorphism (Araujo, 2018). Mindless anthropomorphism was measured by asking the participants to what extent they evaluated the agent as likable, sociable, friendly, and personal on a 10-point scale.

Their results show that an intelligent frame reduces the perception of mindless anthropomorphism for machine-like agents. Araujo (2018) tries to explain this by saying that human-like cues could have been more important for participants when they were evaluating the agent. If these cues were absent, the participants were more aware that their interaction was with a machine, due to the framing. Their findings also include, that the usage of a human name or human-like language resulted in people increasingly seeing the chatbot as human-like.

4.5 Using metaphors to communicate chatbots' warmth and competence

Araujo (2018) introduced their chatbots with a description to see how this would influence the expectation of users. Khadpe et al. (2020) used metaphors in the introduction of the chatbot and studies the evaluation of chatbots that use these metaphors. It is important to know that evaluating people is done on two fundamental dimensions using measures drawn from the stereotype content model (Fiske et al., 1999). The fundamental dimensions are warmth and competence (Fiske et al., 2002). So if chatbots are seen as social actors by their users, it would be possible that they have expectations of these chatbots which can be measured using the same fundamental dimensions warmth and competence. According to the research of Khadpe et al. (2020), this is the case. Khadpe et al. (2020) investigated user expectations concerning the competence (i.e. capability, skillfulness, and efficacy) and warmth (i.e. friendliness, helpfulness, and sociability) of a chatbot based on how the chatbot was introduced.

When users interact with a chatbot, they create a mental model of that chatbot. Their experience with the chatbot is influenced by this model (Norman, 1988; Khadpe et al., 2020). Understanding these mental models will help understand the system behind the chatbot. However, the mental model does not explain why two chatbots with the same technology can evoke different reactions. So, that is why Khadpe et al. (2020) look at how a metaphor that emphasizes warmth and competence influences the chatbots'

evaluation. These metaphors are short descriptions for the chatbot that suggest its functionality and intentions, like Tay's description from Microsoft *AI that's got no chill* which indicates less warmth (Khadpe et al., 2020).

Khadpe et al. (2020) performed an experiment using a Wizard of Oz paradigm which means that when participants think they are communicating with a chatbot, they are actually communicating with a human. For this, they hired customer-support professionals to be the conversational AI (a travel assistant bot) and left them blind to the treatment condition of the participants. These customer-support professionals were instructed on how to interact with the participants. For the metaphors that were mentioned in the introduction, Khadpe et al. (2020) used warmth and competence as axes that differed in the amount of warmth and competence they communicated, such as a toddler, teenager, professional, and executive. A toddler would show a lot of warmth but a low amount of competence, while an executive would show a lot of competence but a low amount of warmth.

In the first study, the introduction of the chatbot was manipulated. Four different metaphors were used which differed in warmth and competence. Also, a control condition was included in which the participants did not see a metaphor in the chatbot's introduction. The participants first got to see the introduction of the chatbot, after which they had to answer questions regarding their expectations for warmth and competence. After these questions, they started their conversation with the chatbot. After the participant had communicated with the travel assistant bot, they were asked to evaluate their experience in a questionnaire. The measures for the warmth of the system were drawn from the stereotype content model (Fiske et al., 1999). These questions were: "*This AI system was (will be) good-natured*" and "*This AI system was (will be) warm*". An example of questions to measure competence are: "*The AI system was easy to use*" and came from Kujala et al. (2017).

Khadpe et al. (2020) found that the introductions without a metaphor yielded higher expectations regarding warmth and competence. Chatbots that used a high competence metaphor, were evaluated as less warm, but competent. Chatbots that were introduced with a low competence metaphor were evaluated higher than chatbots that were introduced with a high competence metaphor (Khadpe et al., 2020).

4.6 The impact of design features to communicate chatbots' warmth and competence

From the papers of Araujo (2018) and Khadpe et al. (2020) it follows that the introduction of customer service chatbots has an influence on the expectations of users in terms of warmth and competence. Khadpe et al. (2020) and Araujo (2018) focus on affective and cognitive aspects of the chatbot in their studies. These aspects are warmth and competence and are part of the design features and user motivation sections in the TAM from Davis (1985). So combining the TAM from Davis (1985) with the studies from Khadpe et al. (2020) and Araujo (2018) would suggest that if a chatbot shows human attributes in the introduction, the perceived ease of use and the perceived usefulness would be higher compared to a chatbot without those human attributes.

This thesis will investigate to what extent users' expectations and evaluations are the result of the design features in the chatbot's introduction. This introduction is the screen that pops up when the user initiates a conversation

with the chatbot. The introduction consists of the header, the first message(s) of the chatbot, and the response option(s). Within these subsections, both verbal and nonverbal design features are used. The chatbot could introduce itself, ask questions, show a profile picture, et cetera. The features in each of these could evoke different user expectations regarding the competence and warmth of the chatbot. More anthropomorphism would result in more competence and warmth, while less anthropomorphism would result in less competence and warmth (Khadpe et al., 2020).

To find out what kind of expectations chatbot users have and to what extent users' expectations are influenced by the verbal and nonverbal features in the chatbot's introduction, this study consists of two separate studies. Study one is a content analysis in which the verbal and nonverbal features in the header, first message(s), and the response options are analyzed for 48 Dutch customer service chatbots. The second study is the participant study in which I systematically manipulate a set of (non-)verbal features in the chatbot's introduction. I will test what expectations those features will produce beforehand and how users evaluate the chatbot afterward.

4.7 Research questions and hypotheses

This results in two different research questions. The first research question (RQ1) being: Which features do Dutch customer service chatbots use in their introduction? The second (RQ2) is: How does the anthropomorphism in a chatbot's introduction influence the expectation and evaluation of users?

For the second research question, three hypotheses arise from the literature. H1: Participants' expectations regarding the chatbot's competence, perceived ease of use, and warmth will be higher when they see a human-like chatbot introduction with anthropomorphism compared to a machine-like chatbot introduction.

H2: Participants' evaluations regarding the chatbot's competence, perceived ease of use, and warmth will be higher when they see a human-like chatbot introduction with anthropomorphism compared to a machine-like chatbot introduction.

H3: Participants' intentions to use the chatbot will be higher when they see a human-like chatbot introduction with anthropomorphism compared to a machine-like chatbot introduction.

5 Content analysis

5.1 Data collection

To answer the question which features are being used in the introduction of Dutch customer service chatbots, a content analysis was conducted. For this content analysis, a representative sample of the Dutch chatbots of different branches was collected within the customer service domain. In total 48 chatbots were included from 8 different branches: nonprofit, electronics, furniture, insurance, logistics and postal service, retail, telecom, and utility. The chatbots that were collected mainly helped customers with their questions. For more information about the data set, see Appendix B.

5.2 Annotation

First, a screenshot was taken from these introductions. Then the introductions from these chatbots were annotated on the presence of ver-

bal and nonverbal cues. This introduction consists of three parts: the header, the first messages, and the response as shown in Figure 1. Within the header of the chatbot, the cues that were annotated were: the presence of a picture, the picture type, subtext in the header, the chatbot's name, chatbot description, what the chatbot was called (like virtual assistant or chatbot), the purpose of the chatbot, a greeting, and a button. Figure 2 shows an example of these elements in the header. For this header, the header picture is present and its type is nonhuman. Furthermore, it is called a "chatbot" and a greeting, a name, and a button are present.

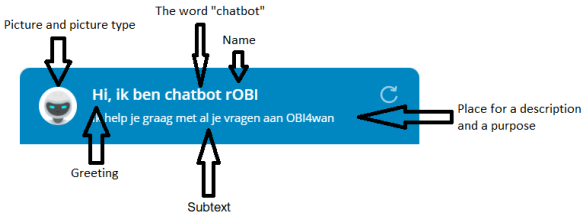


Figure 2: Header annotation.

In the first message(s) section the chatbots were annotated several features. The first features were the number of messages, if there was a picture and if there was a picture, its type. Then I looked for the presence of a question and the type of a question (open or closed). Then, if the chatbot introduced itself by explaining what it is to the user and if the chatbot explained what it would be able to do for the user (the purpose). Finally, I looked at the description used to describe the chatbot (virtual agent, chatbot, or something else) and if the chatbot stated that the conversation could be handed over to a human service employee. Figure 3 shows what this could look like. In this figure, you see that a chatbot introduces itself with two messages, both with a nonhuman picture and an open question. In these messages, there is also an introduction from the chatbot and a chatbot description. In this picture, there is no chatbot purpose or handover present.

Figure 3: First message(s) annotation.

The response was the final part of the annotation process. This first feature was how the user could respond to the first messages from the chatbot and if this was either with text, buttons, or both. The second feature was if there was a text limit when the user could respond with text. The third feature was the number of buttons and checked if the user was able to respond with buttons. Figure 4 shows an example. In this example, you can see that the user was only able to respond with buttons and that there are three buttons to choose from. For more in-depth information on these annotations, see Appendix A.

5.3 Inter-annotator agreement

To assure that the annotation was trustworthy, 15 chatbot introductions were annotated by a second annotator in order to calculate the inter-annotator agreement. Subsequently, the Cohen's Kappa and the Krippendorff's Alpha were cal-

Figure 4: Response annotation.

culated. The scores showed that the annotation was adequate to perfect. For the chatbot purpose and the button in the introduction, no score could be calculated. The values of the Kappa and the Alpha can be found in Table 1, 2, and 3. I proceeded with the annotation of the other chatbot introduction screenshots, after ensuring the annotation approach was reliable.

5.4 Data analysis

After all the data was annotated, the statistical data analysis could start. For this, I mainly looked at frequencies, and in some cases at averages. A χ^2 test was run in R to determine if the presence of features was significant or not.

5.5 Results

5.5.1 Header

In 30 chatbot headers a picture was present ($\chi^2(2) = 20.38$, $p < .001$). However, the picture type did not differ: logo, nonhuman and human pictures equally occurred ($\chi^2(3) = 4.17$, $p = 0.24$). There was significantly more often a subsection present in the header($\chi^2(2) = 11.63$, $p < .001$), but the chatbot's name was often absent in the header ($\chi^2(2) = 13.88$, $p < .001$). Also, the purpose of the chatbot ($\chi^2(2) = 69.13$, $p < .001$), greeting ($\chi^2(2) = 49.63$, $p < .001$), a chatbot description ($\chi^2(2) = 34.13$, $p < .001$), and the word "chatbot" ($\chi^2(2) = 20.38$, $p < .001$) were often absent in the header. There often was a button in the header of the chatbot ($\chi^2(2) = 27.88$, $p < .001$). In 40% of the annotated chatbots, the word to describe the chatbot was "*chatbot*". Both "*digitale assistent*" and "*virtueel assistent*" were each used in 20% of the cases, while "*chatbot in opleiding*" in 10% of the cases. "*Virtuele agent*" and "*virtuele chatbot*" were each used in 5% of the cases.

5.5.2 First messages

The average number of first messages was 1.94 and these were often accompanied by a picture ($\chi^2(2) = 34.13$, $p < .001$). The picture type was significantly more often human than nonhuman or a logo ($\chi^2(3) = 1.67$, $p < .001$). Often a question was present in the first messages ($\chi^2(2) = 32.00$, $p < .001$). Although both open and closed questions occurred equally. ($\chi^2(2) = 0.50$, $p = 0.78$). The chatbot often introduced itself in the firs messages ($\chi^2(2) = 34.13$, $p < .001$), like the chatbot from Kruidvat that introduces

Table 1: Cohen's Kappa scores for the annotation of the header variables

	Picture	Picture type	Subtext	Name	Description	Greeting	"chatbot"
Cohen's Kappa	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 2: Cohen's Kappa and Krippendorff's alpha scores for the annotation of the first message(s) variables

	Number of messages	Picture	Picture type	Question	Question type	Introduction	Purpose	Handover
Cohen's Kappa		1.000	1.000	1.000	1.000	1.000	1.000	0.765
Krippendorff's Alpha	1.000							

Table 3: Cohen's Kappa and Krippendorff's alpha scores for the annotation of the response variables

	Text or button	Text limit	Button amount
Cohen's Kappa	0.868	1.000	
Krippendorff's Alpha			1.000

itself by saying *"Hoi! Ik ben Kiki, de chatbot van Kruidvat."*. The purpose of the chatbots was often absent in the first messages ($\chi^2(2) = 30.13$, $p < .001$). Moreover, the announcement that the customer could be handed-over to a human service employee was often absent ($\chi^2(2) = 60.13$, $p < .001$). In the first messages different labels were used for the chatbot, i.e., "*chatbot*" (35.4%), "*virtuele assistent*" (12.5%), "*digitale assistent*" (6.3%), "*digitale hulp*", "*digitale vraagbaak*", "*virtuele agent*", and "*virtuele klant expert*" (each 2.1%). In 35.4% of the chatbots' first messages section, there was no such label on the chatbot.

5.5.3 Response options

The customer could most often respond to the chatbot by means of text ($\chi^2(2) = 13.63$, $p = 0.0011$), and the other options (buttons or buttons and text) were present approximately the same amount. There was often no text limit for the customer's response ($\chi^2(2) = 16.13$, $p < .001$), and the average amount of buttons to respond with was 4.80.

5.6 Qualitative analysis

There was a lot of difference between the chatbots and how their introductions were designed. Within their introductions, there was a lot of difference in anthropomorphism. They ranged from very human to very nonhuman. This range can be found in the differences between Figure 5 and Figure 6. Figure 5 is the chatbot from the municipality of Hollands Kroon and makes it hard to distinguish whether or not the user is communicating with a human or a chatbot. With features such as a human avatar, a human name, the possibility to respond with text, and

that it never mentions that it is a chatbot make it hard to see if it is a chatbot or a human. Figure 6 on the other hand makes it very clear that you are communicating with a chatbot. It does so with features such as its logo as a picture, calling itself a chatbot and a digital assistant, and the buttons to respond with.

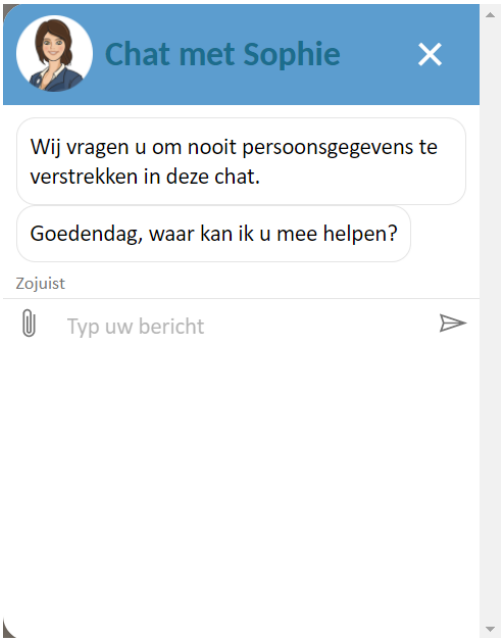


Figure 5: Chatbot with human features.

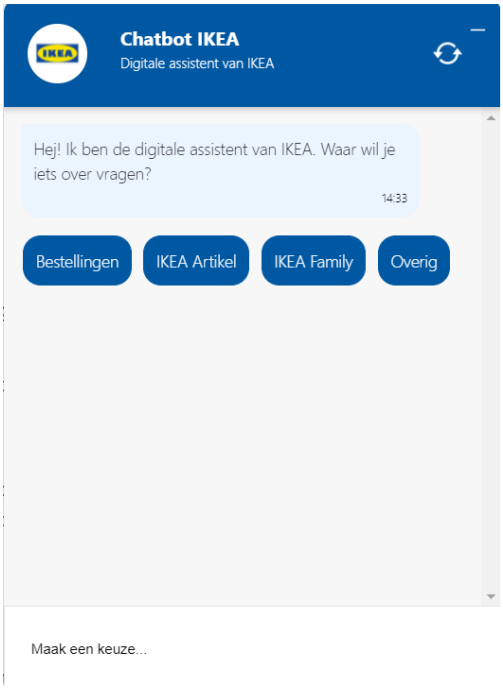


Figure 6: Chatbot with nonhuman features.

5.7 Conclusion

This part of the study was meant to find out which features are used in the introduction of Dutch customer service chatbots. The content analysis showed that in its header, a picture (either human, nonhuman, or a logo) would be found, but also subtext, and a button. For the First Message(s) section, I would find almost two messages, a picture (either human, nonhuman, or a logo), a question, and an introduction in which the chatbot introduces itself. In the last section, the response, I would find a text box to respond with. Without a text limit.

However, from the content analysis, it does not become clear how these elements influence the expectations and evaluations of the introduction of a chatbot. Therefore, a participant study was conducted which investigates anthropomorphism in chatbot introductions and its influence on the expectations and evaluations of the participants. For this participant study, two chatbots are needed: a human and a nonhuman chatbot. The results from the content analysis informed the design of the chatbot introductions used in the participant study.

6 Participant study

6.1 Design

To study if a difference in anthropomorphism influence the expectations and evaluations of chatbot users an experiment was conducted using a 2X2 between-subjects study design. The independent variable was the type of chatbot: nonhuman or human. The dependent variables were participants' expectations concerning competence, perceived ease of use, and warmth and their evaluations concerning competence, perceived ease of use, warmth, and the intention to use the chatbot.

6.2 Chatbots

The chatbots used for this study are shown in a simple website, build for this study, see Appendix D. Within the website, participants used one of two different chatbots: either the human chatbot or the nonhuman chatbot. Both chatbots were designed using the Bots platform from OBI4wan. The goal of those chatbots was to help its user to order a bouquet of flowers from the website by only interacting with the chatbot. The chatbots are made to strictly respond to the script that is used in the study. The only differences in the chatbots were in the introduction. The first differences can be seen in the header (see Figures 7 and 8) and the entire chatbot script can be found in Appendix E. After the introduction of the chatbot, both of these chatbots were the same in what the chatbot said and in the way the user could respond.

In the header, the human chatbot includes both a human profile picture and name, while the nonhuman chatbot consists of an image of a robot and no name. Next, in the first messages, the human chatbot greets the user and introduces itself. Following with explaining its purpose implicitly. The nonhuman chatbot has no greeting, but starts with an introduction and explicitly explaining its purpose. After which a handover is present, which explains to the user that an employee will be contacted if extra help is needed. Both versions end by asking the user a question. In the response options, the human version of the chatbot requires an answer to this question in text, while the nonhuman chatbot includes multiple buttons of which the user should select one. The design differences are focused on how human interaction would work (a greeting, introduction, implicitly explaining its purpose, and the ability to respond in text).

6.3 Participants

In total, 71 participants participated in the participant study: 35 participating in the human version and 36 participating in the nonhuman version. The participants were asked questions regarding their demographics. 68 out of the 71 participants were familiar with chatbots. The participants in the two conditions were comparable concerning age ($t(66.36) = 1.24$, $p = .22$, cohen's $d = .29$) and educational level ($\chi^2(5) = 6.05$, Cramer's $V = .92$). The average age was 32 ($SD = 13.97$). The participants were also asked to select their gender, which resulted in 35 male, 33 female, one different, and two people that would rather not tell.

6.4 Materials and procedure

Next to the use of chatbots and a website, this study used a questionnaire to research the expectations and evaluations of users. There were two versions of the questionnaire which had the same questions but used either the human or the non-

human chatbot introduction. The questionnaire started with an introduction in which the experiment was explained, the user was informed on how their data was handled, and whom to contact in case of a complaint. All participants gave informed consent to participate in the research voluntarily.

On the next page, the user got to see the scenario and the image of the chatbot introduction (see Figures 7 and 8). The task was to order a bouquet of flowers for a friend with a birthday coming up. The participants then received questions about their expectations on the same page. These questions had to be answered on a 7-point Likert scale ranging from completely disagree to completely agree.

Figure 7: Chatbot with nonhuman features.

Figure 8: Chatbot with nonhuman features.

Competence was measured with three items. These items were: Based on the image of the chatbot, the chatbot will be competent, intelligent, and independent. The items were based on Fiske et al. (2002). The reliability of the scale was sufficient (Cronbach's alpha: .62).

Subsequently, participants answered five questions about perceived ease of use. The items were based on Ashfaq et al. (2020) and were: based on the image of the chatbot, this chatbot will be able to order a bouquet quickly and easily. Based on the image of the chatbot, the interaction with the chatbot will be clear and understandable. Based on the image of the chatbot, the chatbot will be easy to use. The reliability of the scale was good (Cronbach's alpha: .88).

To measure warmth, three items were used based on Fiske et al. (2002). The items were: based on the image of the chatbot, the chatbot will be warm, helpful, and sincere. The reliability of the scale was sufficient (Cronbach's alpha: .61).

After the questions regarding the users' expectations, the users got instructed to visit the website (see Appendix D) and use the chatbot to order a bouquet of flowers for the birthday of their friend. Within these instructions, information was given about the flower preference, size preference, cost preference, delivery address, and e-mail address. Before clicking on the link of the website, there was a warning that informed the participant that they were about to leave the page and that they needed to come back to this page to finish the questionnaire.

After participants used the chatbot to order the flowers, they answered questions about the competence of the chatbot based on their experience and answering the same questions as before. The reliability of the scale was good (Cronbach's alpha: .80). Participants also evaluated the perceived ease of use of the chatbot using the same questions as before. The reliability of the scale was excellent (Cronbach's alpha: .92). Moreover, participants evaluated the warmth of the chatbot using the same questions as before. The reliability of the scale was good (Cronbach's alpha: .75). Finally, participants answered two questions about their intention to use the chatbot. These questions were based on Oghuma et al. (2016). The items were: if they would use the chatbot again if they had to do the same task again and if they would recommend the chatbot to others. The reliability of the scale was excellent (Cronbach's alpha: .92). The questions can also be found in Appendix C.

Then the participants were asked a few questions regarding demographic data based on Liebrecht and Van der Weegen (2019). The questions asked the users about their sex, age, current or highest achieved education, and how often they online shopped. Also, if they knew what chatbots were, and if they did, how often they have contact with a chatbot and in which situations. After which a final set of questions was asked about their experience with chatbots in the past on a 7-point Likert scale. These questions asked to what extent the user experienced its previous use of customer service chatbots as useful, to what extent these customer service chatbots improved the communication with the organization, if the communication went well, and if they prefer contact with a chatbot when communicating with an organization. The participant then got an option to leave some remarks, after which they were sent to the end page of the questionnaire. On this page, the participants were thanked for their participation. They were also debriefed about the purpose of the research. They were told that the website and the chatbot were fictitious and were only made for research purposes.

6.5 Data analysis

When all the participants filled in the questionnaire it was important to get all the data neatly in one file for the analysis. So first, because the study consisted of two separate questionnaires, these needed to be added together and labeled whether it was an entry from the human or the nonhuman questionnaire. This resulted in 91 entries. Then the chatbot conversation was checked and all the participants that did not complete the conversation with the chatbot, were filtered out (20 participants). This resulted in 71 complete answers to the participant study.

Which was followed by checking the reliability of the constructs with the Cronbach's Alpha. If the constructs were reliable ($\alpha \geq .60$), the averages per construct were calculated.

To test the hypotheses, independent samples t-test were performed. Moreover, a repeated-measures ANOVA was conducted to investigate whether participants' expectations differed from their evaluations and whether the chatbot type moderated this difference. These statistical tests were performed in SPSS.

6.6 Results

6.6.1 Participant expectations

Before using the chatbot, so when the participants only saw the screenshot of the introduction, there was no significant difference for competence (t(69) = 1.82, p = 0.07, Cohen's d = .43) and warmth (t(69) = -.50, p = 0.6, Cohen's d = -.12). There was a significant difference in perceived ease of use between the human and nonhuman chatbot introduction (t(69) = 2.50, p = 0.02, Cohen's d = .59). With the nonhuman chatbot introduction having higher expectations about the ease of use than the human chatbot. The means can be found in Table 4.

Table 4: Difference in expectation between the human and nonhuman chatbot introduction

Before	Human		Nonhuman	
	Mean	SD	Mean	SD
Competence	3.82	1.15	4.23	.96
Perceived Ease of use	4.40	1.28	5.10	1.07
Warmth	3.92	1.28	3.79	1.00

6.6.2 Participant evaluation

After using the chatbot, there was no significant difference in warmth between the two conditions(t(53.72) = .88, p = .4, Cohen's d = .21), but there was a significant difference for experienced competence(t(69) = 3.35, p = .001, Cohen's d = .79), perceived ease of use(t(53.92) = 2.13, p = .04, Cohen's d = .51), and the intention to use this chatbot again in the future(t(63.15) = 2.79, p = .007, Cohen's d = .67). In all three cases, the nonhuman chatbot scored higher than the human chatbot. The mean values of the constructs can be found in Table 5.

Table 5: Difference in evaluation between the human and nonhuman chatbot introduction

After	Human		Nonhuman	
	Mean	SD	Mean	SD
Competence	3.86	1.43	4.92	1.23
Perceived Ease of use	5.36	1.40	5.94	.80
Warmth	3.94	1.52	4.20	.87
Intention	3.10	1.98	4.26	1.49

6.6.3 Expectations versus evaluations

For the first construct, competence, there was a significant higher evaluation compared to the expectation of the chatbot (Pillai's Trace = 0.81, F(1,69) = 6.08, p = .02, Partial Eta Squared = .081). Also, the nonhuman chatbot was seen as significantly more competent than the human chatbot (F(1,69) = 9.15, p = .003, Partial Eta Squared = .12). Finally, there is a significant difference in the expectation and the evaluation of competence between the human and nonhuman condition (Pillai's Trace = .065, F(1,69) = 4.79, p = .03, Partial Eta Squared = .07). After

talking to the chatbot, the nonhuman chatbot was evaluated significantly more competent ($p = .001$). This difference was not found for the human chatbot ($p = .85$).

Then for the second construct, perceived ease of use, there was also a significant difference in the expectation and evaluation of the chatbot (Pillai's Trace = 0.41, $F(1,69) = 48.06$, $p < .001$, Partial Eta Squared = .041). The evaluation was significantly higher than the expectations. Subsequently, the nonhuman chatbot was perceived as significantly easier to use ($F(1,69) = 6.97$, $p = .01$, Partial Eta Squared = .092). However, there was no significant difference in the expectations and evaluations between the human and nonhuman chatbot regarding this construct (Pillai's Trace = .003, $F < 1$).

Furthermore, for the construct warmth the evaluation was significantly higher than the expectation (Pillai's Trace = .059, $F(1,69) = 4.34$, $p = .04$, Partial Eta Squared = .059). However, no significant difference was found between the human and nonhuman chatbot for warmth ($F < 1$). Neither was a significant difference found in the difference in warmth for the expectation and evaluation between the human and the nonhuman chatbot (Pillai's Trace = .050, $F(1,69) = 3.61$, $p = .06$, Partial Eta Squared = .050).

6.7 Conclusion

Before the participant study, three hypotheses were retrieved from the literature. The first hypothesis (H1) was: Participants' expectations regarding the chatbot's competence, perceived ease of use, and warmth will be higher when they see a human-like chatbot introduction with anthropomorphism compared to a machine-like chatbot introduction. The participant study results in sufficient evidence to reject this hypothesis. The nonhuman chatbot introduction was even expected to be easier to use, compared to the human chatbot introduction.

H2: Participants' evaluations regarding the chatbot's competence, perceived ease of use, and warmth will be higher when they see a human-like chatbot introduction with anthropomorphism compared to a machine-like chatbot introduction. This hypothesis is also rejected. From the participant study, it follows that the nonhuman chatbot introduction resulted in higher evaluations for competence and perceived ease of use when compared to the human chatbot introduction.

H3: Participants' intentions to use the chatbot will be higher when they see a human-like chatbot introduction with anthropomorphism compared to a machine-like chatbot introduction. The last hypothesis is also rejected, based on the results from the participant study. This study shows that participants have a higher intention to use the chatbot with the nonhuman introduction, than the chatbot with the human introduction.

From these results, it can be concluded that all the hypotheses are rejected. Chatbots that use more human traits in the introduction are not evaluated as warmer or more competent. For competence, the opposite is true. If a chatbot uses more nonhuman traits in the introduction it is evaluated as more competent. This means that anthropomorphism is important in the introduction, but not as hypothesized. The nonhuman chatbot led to significantly higher evaluations in terms of competence and perceived ease of use.

7 Discussion

7.1 Implication of the results

The purpose of this study was to gain a better understanding of chatbot introductions in Dutch customer service chatbots and the influence of anthropomorphism on the expectation and evaluation of chatbots. Based on this content analysis I see that chatbots use a wide variety of ways to introduce themselves. Generally, these chatbots use a header and a form of interaction with the user in which the chatbot used 2 messages to introduce themselves and ask a question.

The participant study was to find the influence of anthropomorphism in chatbot introductions on the expectations and evaluations of its users. The hypotheses are rejected. This can be caused by a couple of factors. First of all, the hypotheses were established from prior research. In the meantime, chatbots could have appeared more regularly and people might have gotten used to them. Second, in the prior research from which the hypotheses originate, they manipulated the entire chatbot. My manipulation solely happened in the introduction. The rest of the conversation was the same between the human and nonhuman chatbot. Third, the nonhuman chatbot explicitly states what it can do for its user and also tells the user that it can connect them to a human employee if more help is needed. Chavesa and Gerosaa (2020) and Araujo (2018) show that user expectations are very important when it comes to the evaluation of the chatbot. So another explanation could be that the nonhuman chatbot created better expectations, while the human chatbot did this with more ambiguity.

The findings of the participant study show that design features are important in the introduction of customer service chatbots. The same goes for the cognitive aspects, such as perceived ease of use. This is on par with the TAM from Davis (1985). This means that if these features in the introduction of the chatbot are nonhuman, the user could be more motivated to use the chatbot and more accepting of it.

7.2 Limitations and future research

7.2.1 Limitations of the content analysis

During this study and after gathering the results, I found some issues which limited this study. My contact at OBI4wan gave me a list of chatbots that they built to help me find more chatbots. Other chatbots were found by using google and searching for them. There are currently more Dutch customer service chatbots up and running. So, for a more complete overview of the Dutch customer service chatbot introductions, more chatbots should be annotated.

7.2.2 Limitations of the participant study

Within the participant study, I found three different noteworthy limitations. First, the manipulation that happened in the chatbot introduction consisted of manipulating different features. This causes uncertainty in finding out which manipulation(s) caused the nonhuman chatbot to be better evaluated. In future research features that are manipulated in this study should be systematically manipulated to find out the effect of each feature individually.

Second, the chatbots that were built were very simplistic. They were designed for one purpose only, ordering a bouquet of flowers. Also, if the task had been in a different branch than the

flower business, it could have resulted in different findings.

Third, a qualitative study design, instead of a quantitative study design, could have brought us more insight in why people have these expectations and evaluations of chatbot introductions. This could be done by interviewing the participants instead of using an online questionnaire, or in addition to.

7.2.3 Future research

Future research on this topic should keep these limitations in mind and use them to their advantage. Furthermore, it would be useful to extend the current findings by examining each of the features in the introduction separately in a controlled environment. This participant study only investigated the impact of a set of features, so the results can only be used for this set of features. Next to this, the questionnaire used in the participant study focused on competence, perceived ease of use, warmth, and the intention to keep using the chatbot. There are other constructs that could also be studied from Oghuma et al. (2016) such as perceived security. Work remains to be done before a full understanding of the extent to which anthropomorphism influences chatbot introductions is established.

7.3 Recommendation

When designing chatbots there are different sections of the chatbot to focus on. This study focused on the chatbot introduction and found that when users interacted with a nonhuman chatbot, they expected it to be easier to use and experienced the chatbot as more competent and easy to use. What is also important, is that compared to the human version of the chatbot, the nonhuman chatbot gave people more the intention to use the chatbot again in the future for similar activities. The nonhuman design that is used in this study adopts the following features: a nonhuman picture and just the word "chatbot" in the header. Then the first messages containing an introduction, which is followed by an explanation of the chatbots' purpose. But also a message explaining that humans will be contacted if needed, after which the first question can be asked to help the user. The user then has to respond to this with buttons. Resulting in a proper way to introduce your chatbot. So when designing chatbot introductions, it is important to keep in mind that if you want your chatbot to come across as more competent and to be perceived as easier to use, a nonhuman chatbot, as described before, will deliver better results.

7.4 Relevance for the field of AI

During the Master Artificial Intelligence at the Utrecht University, I came into contact with subjects such as informatics, logic, cognition, psychology, philosophy, and linguistics. For chatbots, a lot of these areas are of importance. The technique behind the chatbot, the human interaction with chatbots, but also the verbal cues from chatbots, are all research areas within the Master Artificial Intelligence. The topic discussed in this Master's thesis, therefore, fits into this Master's program, because it looks at a key tool (chatbots) in which Artificial Intelligence could be used and focuses on humanity within linguistic and psychological aspects for that tool.

7.5 Closing words

Despite the limitations, the present study has enhanced the understanding of the relationship be-

tween anthropomorphism and the introduction of chatbots. I hope that the current research will stimulate further investigation of this important area.

References

Androutsopoulou, A., Karacapilidis, N., Loukis, E., and Charalabidis, Y. (2019). Transforming the communication between citizens and government through ai-guided chatbots. *Government Information Quarterly*, 36(2):358–367.

Appel, J., von der Pütten, A., Krämer, N. C., and Gratch, J. (2012). Does humanity matter? analyzing the importance of social cues and perceived agency of a computer system for the emergence of social reactions during human-computer interaction. *Advances in Human-Computer Interaction*, 2012.

Araujo, T. (2018). 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.

Ashfaq, M., Yun, J., Yu, S., and Loureiro, S. M. C. (2020). I, chatbot: Modeling the determinants of users' satisfaction and continuance intention of ai-powered service agents. *Telematics and Informatics*, 54:101473.

Biocca, F., Harms, C., and Burgoon, J. K. (2003). Toward a more robust theory and measure of social presence: Review and suggested criteria. *Presence: Teleoperators & virtual environments*, 12(5):456–480.

Charlton, G. (2013). Consumers prefer live chat for customer service: stats.

Chaves, A. P. and Gerosa, M. A. (2019). How should my chatbot interact? a survey on human-chatbot interaction design. *arXiv preprint arXiv:1904.02743*.

Chavesa, A. P. and Gerosaa, M. A. (2020). How should my chatbot interact? a survey on social characteristics in human-chatbot interaction design.

Ciechanowski, L., Przegalinska, A., Magnuski, M., and Gloor, P. (2019). In the shades of the uncanny valley: An experimental study of human–chatbot interaction. *Future Generation Computer Systems*, 92:539–548.

Cui, L., Huang, S., Wei, F., Tan, C., Duan, C., and Zhou, M. (2017). Superagent: A customer service chatbot for e-commerce websites. In *Proceedings of ACL 2017, System Demonstrations*, pages 97–102.

Davis, F. D. (1985). *A technology acceptance model for empirically testing new end-user information systems: Theory and results*. PhD thesis, Massachusetts Institute of Technology.

Fiske, S. T., Cuddy, A. J., Glick, P., and Xu, J. (2002). A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. *Journal of personality and social psychology*, 82(6):878.

Fiske, S. T., Xu, J., Cuddy, A. C., and Glick, P. (1999). (dis) respecting versus (dis) liking: Status and interdependence predict ambivalent stereotypes of competence and warmth. *Journal of social issues*, 55(3):473–489.

Gnewuch, U., Morana, S., Adam, M., and Maedche, A. (2018a). Faster is not always better: understanding the effect of dynamic response delays in human-chatbot interaction.

Gnewuch, U., Morana, S., Adam, M. T., and Maedche, A. (2018b). "the chatbot is typing..."– the role of typing indicators in human-chatbot interaction. In *Proceedings of the 17th Annual Pre-ICIS Workshop on HCI Research in MIS*, pages 0–5.

Hendriks, F., Ou, C. X., Amiri, A. K., and Bockting, S. (2020). The power of computer-mediated communication theories in explaining the effect of chatbot introduction on user experience. *interaction*, 12:15.

Kamphaug, Å., Granmo, O.-C., Goodwin, M., and Zadorozhny, V. I. (2017). Towards open domain chatbots—a gru architecture for data driven conversations. In *International Conference on Internet Science*, pages 213–222. Springer.

Khadpe, P., Krishna, R., Fei-Fei, L., Hancock, J. T., and Bernstein, M. S. (2020). Conceptual metaphors impact perceptions of human-ai collaboration. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW2):1–26.

Kujala, S., Mugge, R., and Miron-Shatz, T. (2017). The role of expectations in service evaluation: A longitudinal study of a proximity mobile payment service. *International Journal of Human-Computer Studies*, 98:51–61.

Liebrecht, C. and Van der Weegen, E. (2019). Menselijke chatbots: een zegen voor online klantcontact?: Het effect van conversational human voice door chatbots op social presence en merkattitude. *Tijdschrift voor Communicatiewetenschap*, 47(3-4):217–238.

Liebrecht, C. and Van Hooijdonk, C. (2019). Creating humanlike chatbots: what chatbot developers could learn from webcare employees in adopting a conversational human voice. In *International Workshop on Chatbot Research and Design*, pages 51–64. Springer.

Nass, C., Steuer, J., and Tauber, E. R. (1994). Computers are social actors. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 72–78.

Norman, D. A. (1988). *The psychology of everyday things.* Basic books.

Oghuma, A. P., Libaque-Saenz, C. F., Wong, S. F., and Chang, Y. (2016). An expectation-confirmation model of continuance intention to use mobile instant messaging. *Telematics and Informatics*, 33(1):34–47.

Okuda, T. and Shoda, S. (2018). Ai-based chatbot service for financial industry. *Fujitsu Scientific and Technical Journal*, 54(2):4–8.

Paul, A., Haque Latif, A., Amin Adnan, F., and Rahman, R. M. (2019). Focused domain contextual ai chatbot framework for resource poor languages. *Journal of Information and Telecommunication*, 3(2):248–269.

Reshmi, S. and Balakrishnan, K. (2018). Enhancing inquisitiveness of chatbots through ner integration. In *2018 International Conference on Data Science and Engineering (ICDSE)*, pages 1–5. IEEE.

Serban, I., Sordoni, A., Bengio, Y., Courville, A., and Pineau, J. (2016). Building end-to-end dialogue systems using generative hierarchical neural network models. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 30.

Setyawan, M. Y. H., Awangga, R. M., and Efendi, S. R. (2018). Comparison of multinomial naive bayes algorithm and logistic regression for intent classification in chatbot. In *2018 International Conference on Applied Engineering (ICAE)*, pages 1–5. IEEE.

Spierenburg, G. (2021). Consumenten niet tevreden over chatbots.

Zarouali, B., Van den Broeck, E., Walrave, M., and Poels, K. (2018). Predicting consumer responses to a chatbot on facebook. *Cyberpsychology, Behavior, and Social Networking*, 21(8):491–497.

Zumstein, D. and Hundertmark, S. (2017). Chatbots–an interactive technology for personalized communication, transactions and services. *IADIS International Journal*, 15(1).

Appendices

A Code book

Codebook for Chatbot Introduction by S.P.J. Smaak

Preface

This codebook described how certain aspects in the introduction of chatbots in the customer service domain are being coded. This chatbot introduction consists of 3 sections:

1. The header of the chatbot application

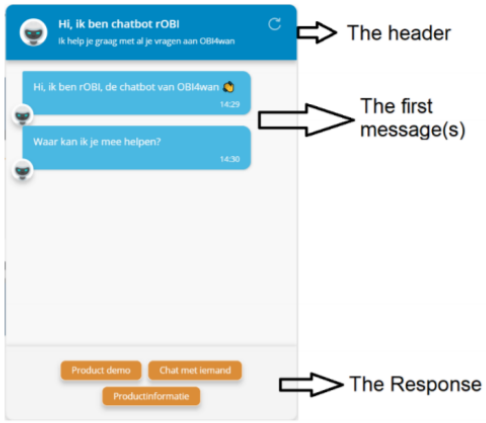
   The header of the chatbot interface is the space above the messages from the chatbot.

2. The first message(s) from the chatbot

   The first message(s) is or are characterized by their location under the header and by often being on the left side of the message window.

3. The options for a user to respond with

   The response is the bottom part of the chatbot interface and is the way the user can react to the chatbot its first message(s) and is either by text, buttons, or both.



Each section contains a set of variables for which code must be given.

Approach

Open the excel spreadsheet and the corresponding chatbot pictures next to this codebook. Look at the first chatbot and look at the division between the header, first message(s), and the response. Then start coding according to the steps mentioned in this document for that chatbot. After finishing this document, start again at the top for the next chatbot.

General remark

Sometimes, a chatbot can lack one of the three aforementioned sections. All the corresponding codes for the variables in this section then become an empty cell, and the coder can move on to the next section.

Section 1: The Header

The variables with their corresponding code options for the header are presented below and are coded in the corresponding yellow columns in the excel file. The choice for each code is explained per variable. The variables are coloured yellow in the excel file.

1. Header_picture

   Is there a profile picture or image present in the header? The profile picture is often placed in the top left corner and often round.

   If there is, the corresponding code is **yes**.

   If there is no profile picture or image in the header, the corresponding code is **no**.

2. Header_pictureType

   If the answer was no in the previous variable, leave the cell **empty**.

   If the answer was yes in the previous variable and the picture or image is the company's logo, the corresponding code is **logo**.

   If the answer was yes in the previous variable and the picture or image resembles a human avatar, the corresponding code is **human**. Resembling a human avatar means that in the picture an entity is present that looks like a human being, albeit animated or not. This is an example of a humanlike avatar:

   If the answer was yes in the previous variable and the picture or image resembles a non-human avatar, the corresponding code is **nonhuman**. Resembling a non-human avatar means that in the picture an avatar is present which does not resemble a human being. It often is the case that it then looks like a robot. This is an example of a nonhuman avatar:

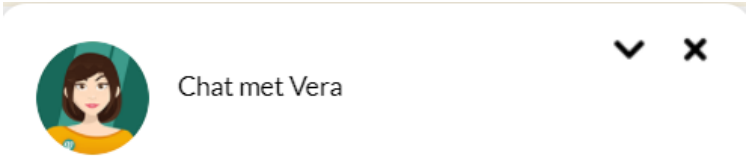
3. Subsection_present

   A subsection could be presented under the main title or section in the header and often has a smaller typeface. If a subsection is present in the header, the corresponding code is **yes**.

   If a subsection is not present in the header, the corresponding code is **no**.

4. Subsection_content_name

   If the content of the header contains the name for the chatbot, the corresponding code is **yes**. An example for this situation is as follows.



If this is not the case, the corresponding code is **no**.

5. Subsection_content_chatbotDescription

   If the content of the header is a description of the chatbot, the corresponding code is **yes**. An example of a chatbot description is "*This is a chatbot that is still learning!"* or "Digital Assistant".

   If this is not the case, the corresponding code is **no**

6. Header_ChatbotDescription

   This feature is not filled in by the second annotator, just by me, Sammie Smaak. The corresponding code for this feature is the word(s) used to describe the chatbot in the header, otherwise, the cell is left empty.

7. Subsection_content_chatbotPurpose

   If the content of the header is the purpose of the chatbot, the corresponding code is **yes**. This should explain what the chatbot will do for the user. An example of a chatbot purpose is: "*I am here to help you find an answer to your question!"*.

   If this is not the case, the corresponding code is **no**.

8. Subsection_content_greeting

   If the content of the header is a greeting, the corresponding code is **yes**. An example of a greeting is: *"Welcome to the chat"*.

   If this is not the case, the corresponding code is **no**.

9. Subsection_content_button

   If there is at least one button present in the header, the corresponding code is **yes**. Buttons in a header could look like the buttons on the right in the following example.

   If there is no button present in the header, the corresponding code is **no**.

10. Subsection_content_chatbot

    If the word "chatbot" occurs in the header, the corresponding code is **yes**.

    If the word "chatbot" does not occur in the header, the corresponding code is **no**.

Section 2: The First Message(s)

The variables with their corresponding code options for the first message(s) are presented below in and can coded in the corresponding orange columns in the excel file. The choice for each code is explained per variable..

1. FirstMessage_amount

   The corresponding code is an **integer** with a minimum value of 0 and resembles the number of messages the chatbot has sent.

2. FirstMessage_picture

   If there is at least one picture or image present next to the messages in this section, the corresponding code is **yes**. This could look like the example on the right, but the picture or image could also be located on the bottom left or top left from the message. In this picture you could find an avatar (either human or non-human), a company logo, or just a colour.

   If there is no picture or image present next to the messages in this section, the corresponding code is **no**.

3. FirstMessage_pictureType

   If the answer was no in the previous variable, leave the cell **empty**.

   If the answer was yes in the previous variable and the picture or image is the company's logo, the corresponding code is **logo**.

   If the answer was yes in the previous variable and the picture or image resembles a human avatar, the corresponding code is **human**. Resembling a human avatar means that in the picture an entity is present that looks like a human being, albeit animated or not. An example of a humanlike avatar:

   If the answer was yes in the previous variable and the picture or image resembles a non-human avatar, the corresponding code is **nonhuman**. Resembling a non-human avatar means that in the picture an avatar is present which does not resemble a human being. It often is the case that it then looks like a robot. An example of a nonhuman avatar:

4. FirstMessage_question

   If in the first message(s) the chatbot has asked a question, the corresponding code is **yes**.

   If in the first message(s) the chatbot has not asked a question, the corresponding code is **no**.

5. FirstMessage_questionType

   If the answer was no in the previous variable, leave the cell **empty**.

If the answer was yes in the previous variable and the question was an open-ended question, the corresponding code is **open**. An open-ended question is a question that cannot be answered with yes, no, or another clear set of possible answers.

If the answer was yes in the previous variable and the question was a closed ended question, the corresponding code is **closed**. A closed question is a question which has a clear set of possible answers (e.g. yes or no questions). A question with only buttons as a response option is coded as **closed**. If it also has an option to write a response, it can be **open**.

6. FirstMessage_introduction

If within the first message(s) the chatbot has introduced itself, the corresponding code is **yes**. A chatbot Introducing itself means for it to tell someone its name when they meet for the first time and perhaps a small description of itself. A way a chatbot can introduce itself is the following:

If within the first message(s) the chatbot has not introduced itself, the corresponding code is **no**.

7. FirstMessage_chatbotPurpose

If the purpose of the chatbot becomes clear from the first message(s), the corresponding code is **yes**. The purpose of a chatbot within the customer service domain is often helping a customer to find an answer to their question or a solution to their problem.

If the purpose of the chatbot does not become clear from the first message(s), the corresponding code is **no**.

8. FirstMessage_ChatbotDescription

This feature is not filled in by the second annotator, just by me, Sammie Smaak. The corresponding code for this feature is the word(s) used to describe the chatbot in the first message(s), if this is not present, the cell is left **empty**.

9. FirstMessage_Handover

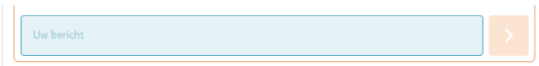
Does the chatbot show the ability to connect you to a human? Then the corresponding code is **yes.** It should be the chatbot that connects you to a human, so something like “you can call the customer service if I cannot help you” is not a handover. And if there is no handover, the corresponding code is **no.**

Section 3: The Response

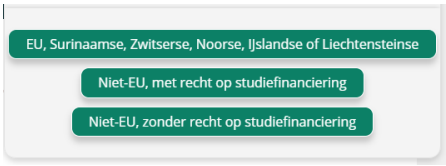
The variables with their corresponding code options for the response are presented below and can coded in the corresponding blue-grey columns in the excel file. The choice for each code is explained per variable.

1. Response_textButton

   How can the user react to the first message(s) from the chatbot? If this can only be done through typing an answer in a text box, the corresponding code is **text**. Such a textbox could look something like the following:

   If the user can solely react to the chatbot by choosing and clicking an answer from given buttons, the corresponding code is **button**. To give an example, the buttons the chatbot of the Erasmus University uses are displayed on the right.

   If the user can do both, the corresponding code is **both**.

2. Tekst_limit

   If the answer was button in the previous variable, leave the cell **empty**.

   If there is no limit visible for how much text the response can hold, the corresponding code is to leave the cell **no**.

   If there is a limit visible for the response, the corresponding code is **yes**. A limit is an instruction of the chatbot for the user how the user should formulate their answer. This could be a maximum number of words, letters, or something else and is often displayed within the text box. This could look something like this:

3. Button_amount

   If the answer was text for the first question of this section variable, leave the cell **empty**.

   The corresponding code is an **integer** with a minimum value of 0 and resembles the number of buttons messages the chatbot has shown as available replies.

B Annotation table

Company Name	Branche	Header_picture	Header_pictureType	Subsection_present	Subsection_content_name
Gemeente Goes	Non Profit	yes	nonhuman	yes	yes
Gemeente Woerden	Non Profit	yes	nonhuman	yes	yes
Gemeente Veenendaal	Non Profit	yes	human	yes	yes
Gemeente Hollands Kroor	Non Profit	yes	human	no	yes
Kvk	Non Profit	yes	nonhuman	yes	no
AMC	Non Profit				
BKR	Non Profit	yes	human	yes	yes
Albert Heijn	Retail	no		no	no
Hello Fresh	Retail				
Kruidvat	Retail	yes	human	no	no
Ici paris	Retail	yes	human	no	no
Swapfiets	Retail	no		no	no
Zalando	Retail	yes	logo	yes	no
Bol.com	Retail	no		yes	yes
postnl	Logistics and Postal Service	no		no	no
DHL	Logistics and Postal Service	yes	nonhuman	yes	yes
Logistiekconcurrent	Logistics and Postal Service	yes	human	yes	yes
Mediamarkt	Electronics	yes	human	yes	no
Phoenix contact	Electronics	yes	human	yes	yes
Samsung	Electronics	yes	logo	yes	yes
HP	Electronics	yes	logo	no	no
TOMTOM	Electronics	no		yes	no
Kantoorinrichtingkopen	Furniture	yes	nonhuman	yes	no
Karwei	Furniture	yes	nonhuman	yes	no
BeterBed	Furniture	yes	logo	no	no
IKEA	Furniture	yes	logo	yes	no
Lab21	Furniture	yes	logo	yes	yes
NUON/vattenfall	Utility				
OXXIO	Utility	yes	nonhuman	no	no
Green choice	Utility	yes	nonhuman	yes	yes
Energie direct	Utility	yes	human	no	no
Eneco	Utility	no		no	no
Symio	Telecom	yes	human	no	yes
Tele2	Telecom	yes	logo	no	no
Ben.nl	Telecom	yes	logo	no	no
t-mobile	Telecom	yes	logo	no	no
vodafone	Telecom	no		yes	yes
ziggo	Telecom	no		yes	yes
ohra	Insurance	no		no	no
VGZ	Insurance	no		no	no
Aon	Insurance	yes	nonhuman	yes	no
FBTO	Insurance	no		no	no
Centraal Beheer	Insurance	no		no	yes
InShared	Insurance				
ASR	Insurance				
Nationale Nederlanden	Insurance	yes	nonhuman	no	no
Scildon	Insurance	yes	nonhuman	yes	no
Menzis	Insurance	no		yes	yes

Company Name	Branche	Subsection_content_chatbotDescription	Header_ChatbotDescription	Subsection_content_chatbotPurpose
Gemeente Goes	Non Profit	yes	chatbot in opleiding	no
Gemeente Woerden	Non Profit	yes	chatbot in opleiding	no
Gemeente Veenendaal	Non Profit	no	chatbot	no
Gemeente Hollands Kroor	Non Profit	no		no
Kvk	Non Profit	yes	chatbot	no
AMC	Non Profit			
BKR	Non Profit	no	chatbot	no
Albert Heijn	Retail	no		no
Hello Fresh	Retail			
Kruidvat	Retail	no		no
Ici paris	Retail	no		no
Swapfiets	Retail	no		no
Zalando	Retail	no		no
Bol.com	Retail	no	chatbot	no
postnl	Logistics and Postal Service	no		no
DHL	Logistics and Postal Service	no	chatbot	no
Logistiekconcurrent	Logistics and Postal Service	no	chatbot	no
Mediamarkt	Electronics	no	Virtueel assistent	no
Phoenix contact	Electronics	no	chatbot	no
Samsung	Electronics	yes	virtuele assistant	no
HP	Electronics	yes	virtuele agent	no
TOMTOM	Electronics	no		no
Kantoorinrichtingkopen	Furniture	no		no
Karwei	Furniture	no	chatbot	no
BeterBed	Furniture	no		no
IKEA	Furniture	yes	Digitale assistent	no
Lab21	Furniture	yes	virtuele chatbot	no
NUON/vattenfall	Utility			
OXXIO	Utility	no		no
Green choice	Utility	Yes	Virtuele assistent	no
Energie direct	Utility	no		no
Eneco	Utility	no	digitale assistent	no
Symio	Telecom	no		no
Tele2	Telecom	no		no
Ben.nl	Telecom	no	digitale assistent	no
t-mobile	Telecom	no		no
vodafone	Telecom	no		no
ziggo	Telecom	no		no
ohra	Insurance	no	Virtuele assistent	no
VGZ	Insurance	no		no
Aon	Insurance	no		no
FBTO	Insurance	no		no
Centraal Beheer	Insurance	no		no
InShared	Insurance			
ASR	Insurance			
Nationale Nederlanden	Insurance	no		no
Scildon	Insurance	no		no
Menzis	Insurance	no		no

Company Name	Branche	Subsection_content_greeting	Subsection_content_button	Subsection_content_chatbot
Gemeente Goes	Non Profit	no	yes	yes
Gemeente Woerden	Non Profit	no	no	yes
Gemeente Veenendaal	Non Profit	yes	yes	yes
Gemeente Hollands Kroor	Non Profit	no	yes	no
Kvk	Non Profit	no	yes	yes
AMC	Non Profit			
BKR	Non Profit	no	yes	yes
Albert Heijn	Retail	no	yes	no
Hello Fresh	Retail			
Kruidvat	Retail	no	yes	no
Ici paris	Retail	no	yes	no
Swapfiets	Retail	no	yes	no
Zalando	Retail	no	yes	no
Bol.com	Retail	no	yes	yes
postnl	Logistics and Postal Service	no	yes	no
DHL	Logistics and Postal Service	no	yes	yes
Logistiekconcurrent	Logistics and Postal Service	no	yes	yes
Mediamarkt	Electronics	no	yes	no
Phoenix contact	Electronics	no	yes	yes
Samsung	Electronics	yes	yes	yes
HP	Electronics	no	no	no
TOMTOM	Electronics	yes	yes	no
Kantoorinrichtingkopen	Furniture	no	yes	no
Karwei	Furniture	no	yes	yes
BeterBed	Furniture	no	yes	no
IKEA	Furniture	no	yes	yes
Lab21	Furniture	no	yes	yes
NUON/vattenfall	Utility			
OXXIO	Utility	no	no	no
Green choice	Utility	no	yes	no
Energie direct	Utility	no	no	no
Eneco	Utility	no	no	no
Symio	Telecom	no	yes	no
Tele2	Telecom	no	no	no
Ben.nl	Telecom	no	yes	no
t -mobile	Telecom	no	no	no
vodafone	Telecom	no	yes	no
ziggo	Telecom	no	yes	no
ohra	Insurance	no	yes	no
VGZ	Insurance	no	yes	no
Aon	Insurance	yes	no	no
FBTO	Insurance	no	yes	no
Centraal Beheer	Insurance	no	yes	no
InShared	Insurance			
ASR	Insurance			
Nationale Nederlanden	Insurance	no	yes	no
Scildon	Insurance	no	no	no
Menzis	Insurance	no	no	no

Company Name	Branche	FirstMessage_amount	FirstMessage_picture	FirstMessage_pictureType	FirstMessage_question
Gemeente Goes	Non Profit	2	yes	nonhuman	yes
Gemeente Woerden	Non Profit	2	yes	nonhuman	yes
Gemeente Veenendaal	Non Profit	2	yes	human	yes
Gemeente Hollands Kroon	Non Profit	2	no		yes
Kvk	Non Profit	3	yes	nonhuman	no
AMC	Non Profit	2	yes	nonhuman	yes
BKR	Non Profit	2	yes	human	yes
Albert Heijn	Retail	1	no		yes
Hello Fresh	Retail	2	no		yes
Kruidvat	Retail	4	yes	human	no
Ici paris	Retail	3	yes	human	no
Swapfiets	Retail	2	yes	logo	no
Zalando	Retail	2	no		no
Bol.com	Retail	1	yes	human	yes
postnl	Logistics and Postal Serv	3	yes	nonhuman	yes
DHL	Logistics and Postal Serv	2	yes	logo	yes
Logistiekconcurrent	Logistics and Postal Serv	3	yes	human	yes
Mediamarkt	Electronics	1	yes	human	no
Phoenix contact	Electronics	3	no		yes
Samsung	Electronics	1	yes	logo	yes
HP	Electronics	2	yes	human	no
TOMTOM	Electronics	2	yes	nonhuman	no
Kantoorinrichtingkopen	Furniture	1	no		yes
Karwei	Furniture	3	yes	logo	yes
BeterBed	Furniture	3	yes	logo	yes
IKEA	Furniture	1	no		yes
Lab21	Furniture	2	yes	logo	yes
NUONvattenfall	Utility	1	yes	human	yes
OXXIO	Utility	1	yes	nonhuman	no
Green choice	Utility	3	yes	nonhuman	yes
Energie direct	Utility	3	yes	human	no
Eneco	Utility	2	yes	logo	no
Symio	Telecom	2	yes	human	no
Tele2	Telecom	2	no		yes
Ben.nl	Telecom	2	no		no
t-mobile	Telecom	2	no		yes
vodafone	Telecom	2	no		yes
ziggo	Telecom	2	no		yes
ohra	Insurance	1	yes	nonhuman	yes
VGZ	Insurance	2	no		yes
Aon	Insurance	1	yes	human	yes
FBTO	Insurance	1	no		yes
Centraal Beheer	Insurance	2	yes	nonhuman	no
InShared	Insurance	1	yes	human	yes
ASR	Insurance	2	yes	logo	no
Nationale Nederlanden	Insurance	1	no		no
Scildon	Insurance	2	yes	logo	yes
Menzis	Insurance	1	yes	nonhuman	yes

Company Name	Branche	FirstMessage_questionType	FirstMessage_introduction	FirstMessage_chatbotPurpose	FirstMessage_ChatbotDescription	FirstMessage_Handover
Gemeente Goes	Non Profit	open	yes	no	digitale vraagbaak	no
Gemeente Woerden	Non Profit	open	yes	no	virtuele medewerker	no
Gemeente Veenendaal	Non Profit	closed	no	no		no
Gemeente Hollands Kroon	Non Profit	open	no	no		no
Kvk	Non Profit		yes	yes	virtuele assistent	no
AMC	Non Profit	open	no	no		no
BKR	Non Profit	closed	yes	no	chatbot	no
Albert Heijn	Retail	open	no	no	Digitale Assistent	No
Hello Fresh	Retail	open	no	no		no
Kruidvat	Retail		yes	no	Chatbot	Yes
Ici paris	Retail		yes	no	digitale hulp	no
Swapfiets	Retail		no	yes		no
Zalando	Retail		yes	yes	chatbot	yes
Bol.com	Retail	open	yes	no	chatbot	no
postnl	Logistics and Postal Serv	closed	no	yes		no
DHL	Logistics and Postal Serv	closed	yes	no	chatbot	no
Logistiekconcurrent	Logistics and Postal Serv	closed	no	no		no
Mediamarkt	Electronics		yes	no	virtuele assistent	no
Phoenix contact	Electronics	closed	yes	yes		no
Samsung	Electronics	closed	yes	no	virtuele assistent	no
HP	Electronics		yes	yes	virtuele agent	no
TOMTOM	Electronics		no	yes		no
Kantoorinrichtingkopen	Furniture	open	no	no		no
Karwei	Furniture	closed	yes	no	chatbot	no
BeterBed	Furniture	closed	yes	no	chatbot	no
IKEA	Furniture	closed	yes	no	digitale assistent	no
Lab21	Furniture	closed	no	no		no
NUON/vattenfall	Utility	open	yes	no		no
OXXIO	Utility		no	no		no
Green choice	Utility	open	yes	yes	chatbot	no
Energie direct	Utility		yes	yes	chatbot	no
Eneco	Utility		no	no		no
Symio	Telecom		yes	no	virtuele klant expert	no
Tele2	Telecom	open	yes	no	virtuele assistent	no
Ben.nl	Telecom		yes	yes	digitale assistent	no
t-mobile	Telecom	open	yes	yes	chatbot	yes
vodafone	Telecom	closed	yes	yes	chatbot	yes
ziggo	Telecom	closed	yes	yes	chatbot	yes
ohra	Insurance	open	yes	no		no
VG2	Insurance	closed	yes	yes	chatbot	yes
Aon	Insurance	open	yes	no	chatbot	no
FBTO	Insurance	open	yes	yes	virtuele assistent	yes
Centraal Beheer	Insurance		yes	yes	chatbot	no
InShared	Insurance	open	yes	no	chatbot	no
ASR	Insurance		no	no		no
Nationale Nederlanden	Insurance		yes	yes	chatbot	no
Scildon	Insurance	open	no	no		no
Menzis	Insurance	open	yes	no	virtuele assistent	no

Company Name	Branche	Response_textButton	Tekst_limit	Button_amount
Gemeente Goes	Non Profit	both	no	11
Gemeente Woerden	Non Profit	both	no	12
Gemeente Veenendaal	Non Profit	button		8
Gemeente Hollands Kroon	Non Profit	text	no	
Kvk	Non Profit	text	yes	
AMC	Non Profit	text	no	
BKR	Non Profit	button		6
Albert Heijn	Retail	text	no	
Hello Fresh	Retail	text	no	
Kruidvat	Retail	text	yes	
Ici paris	Retail	both	yes	6
Swapfiets	Retail	text	no	
Zalando	Retail	text	no	
Bol.com	Retail	text	no	
postnl	Logistics and Postal Serv	button		3
DHL	Logistics and Postal Serv	text	yes	
Logistiekconcurrent	Logistics and Postal Serv	text	no	
Mediamarkt	Electronics	both	no	2
Phoenix contact	Electronics	button		4
Samsung	Electronics	both	no	2
HP	Electronics	both	no	5
TOMTOM	Electronics	both	no	3
Kantoorinrichtingkopen	Furniture	text	no	
Karwei	Furniture	button		2
BeterBed	Furniture	button		6
IKEA	Furniture	button		4
Lab21	Furniture	button		6
NUON/vattenfall	Utility	text	no	
OXXIO	Utility	text	yes	
Green choice	Utility	text	no	
Energie direct	Utility	both	no	4
Eneco	Utility	text	yes	
Symio	Telecom	text	yes	
Tele2	Telecom	text	no	
Ben.nl	Telecom	text	no	
t -mobile	Telecom	text	no	
vodafone	Telecom	button		2
ziggo	Telecom	button		2
ohra	Insurance	text	no	
VGZ	Insurance	button		3
Aon	Insurance	text	no	
FBTO	Insurance	text	no	
Centraal Beheer	Insurance	text	no	
InShared	Insurance	text	yes	
ASR	Insurance	both	no	5
Nationale Nederlanden	Insurance	text	no	
Scildon	Insurance	text	no	
Menzis	Insurance	text	no	

C Questionnaire

Before using the chatbot, after seeing the screenshot of the chatbot introduction, competence was measured with the following items:

- Based on the image of the chat-bot, the chatbot will be competent.
- Based on the image of the chat-bot, the chatbot will be intelligent.
- Based on the image of the chat-bot, the chatbot will be independent.

Before using the chatbot, after seeing the screenshot of the chatbot introduction, perceived ease of use was measured with the following items:

- Based on the image of the chat-bot, I will be able to order a bouquet of flowers quickly with this chatbot.
- Based on the image of the chat-bot, I will be able to order a bouquet of flowers easily with this chatbot.
- Based on the image of the chat-bot, the interaction with the chatbot will be clear.
- Based on the image of the chat-bot, the interaction with the chatbot will be understandable.
- Based on the image of the chat-bot, the chatbot will be easy to use.

Before using the chatbot, after seeing the screenshot of the chatbot introduction, warmth was measured with the following items:

- Based on the image of the chat-bot, the chatbot will be warm.
- Based on the image of the chat-bot, the chatbot will be helpful.
- Based on the image of the chat-bot, the chatbot will be sincere.

After using the chatbot, competence was measured with the following items:

- Based on the conversation with the chatbot, the chatbot was competent.
- Based on the conversation with the chatbot, the chatbot was intelligent.
- Based on the conversation with the chatbot, the chatbot was independent.

After using the chatbot, perceived ease of use was measured with the following items:

- Based on the conversation with the chatbot, I was able to order a bouquet of flowers quickly with this chatbot.

- Based on the conversation with the chatbot, I was able to order a bouquet of flowers easily with this chatbot.
- Based on the conversation with the chatbot, the interaction with the chatbot was clear.
- Based on the conversation with the chatbot, the interaction with the chatbot was understandable.
- Based on the conversation with the chatbot, the chatbot was easy to use.

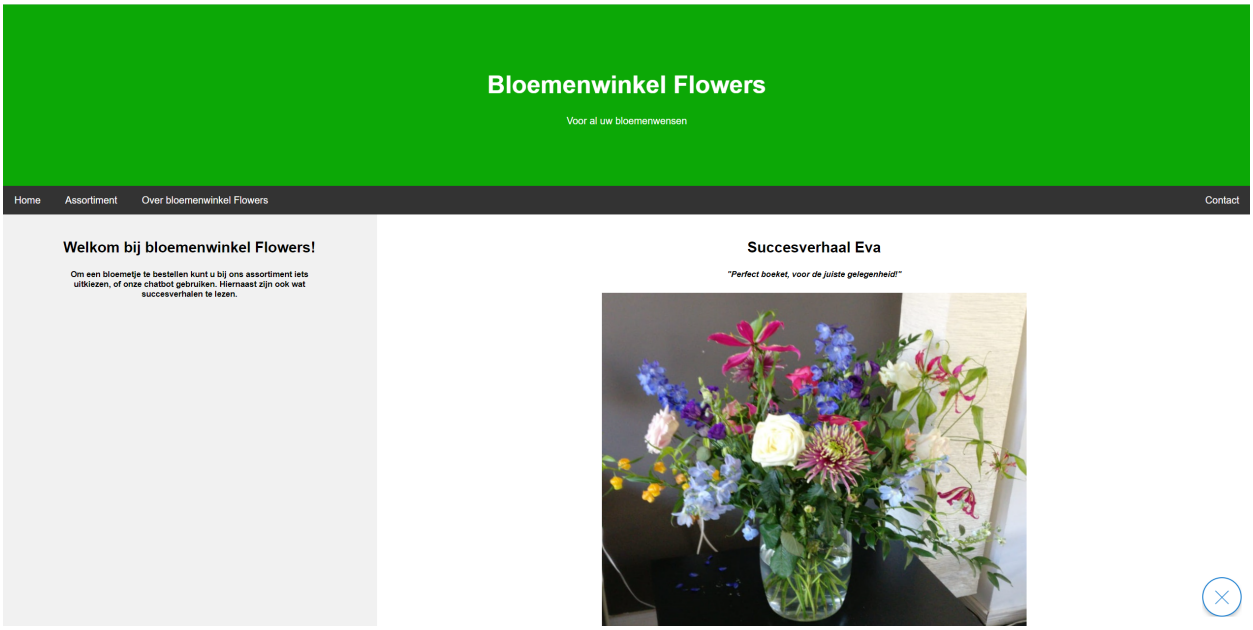
After using the chatbot, warmth was measured with the following items:

- Based on the conversation with the chatbot, the chatbot was warm.
- Based on the conversation with the chatbot, the chatbot was helpful.
- Based on the conversation with the chatbot, the chatbot was sincere.

After using the chatbot, the intention to use the chatbot was measured with the following items:

- If I would order a bouquet of flowers online again, I would use this chatbot.
- I would recommend this chatbot to others for ordering a bouquet of flowers online.

D Website


Bloemenwinkel Flowers
Voor al uw bloemenwensen
Home
Assortiment
Over bloemenwinkel Flowers
Contact
Welkom bij bloemenwinkel Flowers!
Om een bloemetje te bestellen kunt u bij ons assortiment iets uitkiezen, of onze chatbot gebruiken. Hiernaast zijn ook wat succesverhalen te lezen.
Succesverhaal Eva
"Perfect boeket, voor de juiste gelegenheid!"

E Chatbot conversation blueprint

Human chatbot	User
Hallo, ik ben Fleur de chatbot van bloemenwinkel Flowers. Ik kan je helpen en hiervoor ga ik je een aantal vragen stellen. Voor welke gelegenheid is het boeket?	
	Response
Wat leuk! Wat mag het verjaardagsboeket kosten?	
	Response
En hoe groot mag het boeket zijn?	
	Response
Op basis van je antwoorden heb ik drie boeketten voor je uitgekozen. [3 afbeeldingen] Welk boeket wil je bestellen?	
	Response
Op welke datum wil je dit boeket laten bezorgen?	
	Response
Wat zijn de adresgegevens?	
	Response
Wat is je e-mailadres, zodat wij de factuur kunnen sturen?	
	Response
Bedankt voor je bestelling! Ik ga er mee aan de slag! Sluit nu de chat en ga terug naar de vragenlijst.	

Nonhuman chatbot	User
Ik ben de chatbot van bloemenwinkel Flowers. Ik kan je helpen met het bestellen van een boeket bloemen. Mocht ik er helemaal niet uitkomen, dan verbind ik je door met een menselijke collega. Voor welke gelegenheid wil je een boeket bestellen?	
	Response
Wat leuk! Wat mag het verjaardagsboeket kosten?	
	Response
En hoe groot mag het boeket zijn?	
	Response
Op basis van je antwoorden heb ik drie boeketten voor je uitgekozen. [3 afbeeldingen] Welk boeket wil je bestellen?	
	Response
Op welke datum wil je dit boeket laten bezorgen?	
	Response
Wat zijn de adresgegevens?	
	Response
Wat is je e-mailadres, zodat wij de factuur kunnen sturen?	
	Response
Bedankt voor je bestelling! Ik ga er mee aan de slag! Sluit nu de chat en ga terug naar de vragenlijst.	