Tampere University

Susanna Eskelinen

HOW CAN I HELP YOU?

The usability of answers provided by the customer service chatbots of student housing associations

Faculty of Information Technology and Communication Sciences
Master's thesis
April 2020

ABSTRACT

Susanna Eskelinen: How can I help you? The usability of answers provided by the customer service chatbots of student housing associations
Master's thesis
Tampere University
Master's Programme in Multilingual Communication and Translation Studies
April 2020

Chatbots are becoming more common in customer service. Because chatbots have a greater importance, the usability of their answers is also more important. This study evaluates the usability of the answers provided by customer service chatbots. The customer service chatbots examined in this study are from Finnish student housing associations. There are six associations from different parts of Finland.

The theoretical framework of this study comprises usability and the chatbots in the framework of technical communication. This study sees the answers of the customer service chatbots as part of technical communication and as instructional texts. Therefore, usability heuristics for documentation are included to the theoretical framework.

Cognitive walkthrough is used to analyse the usability of each chatbot's answers. Due to the evaluation method, the focus in the usability evaluation is on the learnability of the answers. For the cognitive walkthrough, the user profile is defined as a new user of the services which student housing associations provide.

The study material was gathered with a set of questions to the chatbots. These questions were developed from the frequently asked questions on the websites of the associations. The questions were placed on a simplified customer journey map which was created for this study. The simplified customer journey map was created to reflect the customer experience a user might have with any of the associations in this study.

The individual evaluations revealed several places for improving the usability of the answers. For example, the answers could use more approachable terminology and address the user directly while instructing them. This study could be used as a basis for creating comprehensive usability testing for the answers. This usability testing could include entire user tasks to perform with a chatbot.

Keywords: chatbot, customer service, technical communication, usability, cognitive walkthrough

TIIVISTELMÄ

Susanna Eskelinen: How can I help you? The usability of answers provided by the customer service chatbots of student housing associations
Pro gradu tutkielma
Tampereen Yliopisto
Monikielisen viestinnän ja Käännöstieteen maisteriohjelma
Huhtikuu 2020

Chatboteista on tulossa koko ajan tärkeämpi osa asiakaspalvelua, joten niiden vastausten käytettävyydelläkin on yhä suurempi merkitys. Tässä tutkimuksessa tarkastellaan asiakaspalveluun käytettävien chatbottien vastausten käytettävyyttä. Tutkimuksessa käytettävät chatbotit ovat kuuden suomalaisen opiskelija-asuntosäätiön chatbotteja. Säätiöt ovat eri puolilta Suomea.

Tämän tutkimuksen tieteellinen viitekehys pohjautuu käytettävyyteen ja chatbottien tarkasteluun teknisen viestinnän näkökulmasta. Tässä tutkimuksessa asiakaspalvelun chatbottien vastaukset nähdään teknisen viestinnän tuotteina ja ohjeteksteinä. Siten tutkimuksen teoreettisessa kehyksessä hyödynnetään myös dokumentaation käytettävyysheuristiikoita.

Yksittäisten chatbottien vastausten käytettävyyden arviointiin käytetään kognitiivista läpikäyntiä. Arvioinnissa korostuu opittavuus, koska metodina kognitiivinen läpikäynti korostaa sitä. Läpikäyntiä varten tutkimuksessa määritellään käyttäjäksi uusi opiskelija, joka ei ole aikaisemmin käyttänyt opiskelija-asuntosäätiön palveluita. Tutkimusmateriaali on kerätty kysymyslistalla chatboteilta. Nämä kysymykset on kehitetty säätiöiden nettisivuilta löytyvistä yleisimmin kysytyistä kysymyksistä. Sen jälkeen ne on asetettu tutkimusta varten kehitetylle yksinkertaistetulle asiakkaan palvelupolulle. Yksinkertaistettu asiakkaan palvelupolku on kehitetty heijastelemaan asiakaspolkua minkä tahansa tutkimukseen kuuluvan säätiön kohdalla.

Yksittäisissä arvioinneissa havaittiin, että vastausten käytettävyyttä voitaisiin parantaa monilla tavoilla. Vastauksissa voitaisiin esimerkiksi käyttää lähestyttävämpää terminologiaa ja puhutella käyttäjää suoraan ohjeistamisen aikana. Tätä tutkimusta voitaisiin kehittää vastausten laajemman käytettävyystutkimuksen suuntaan. Tällaisessa käytettävyystutkimuksessa voitaisiin huomioida vastausten luonne tarkemmin ja sisällyttää kokonaisia käyttäjätehtäviä suoritettavaksi chatbotin kanssa.

Avainsanat: chatbot, asiakaspalvelu, tekninen viestintä, käytettävyys, kognitiivinen läpikäynti

TABLE OF CONTENTS

1 INTRODUCTION

Chatbots, which are also called conversational agents, have a relatively long history (Mathur & Lopez 2018, 1). Traditionally, a chatbot is defined as a technological agent that communicates with a user through natural language (Khan & Das 2017, n.p.). This often means written language (Dale 2016, 3). They have become more common in everyday use. There are chatbots for highly specialized tasks. For instance, some chatbots give therapy (Sharma, Puri & Rawat 2018, n.p.), others provide the probability of a diagnosis during a pandemic situation (Terveyskylä 2020) and some are for personal banking assistance (Fintech 2017, n.p.). Chatbots are even delivering the news, for example *Duunibotti* (Björksten, Kanerva & Tuominen 2020, n.p.), which was published by the Finnish national public broadcasting company, Yle. *Duunibotti* can tell a user how their profession is doing, in which part of Finland it is easiest to be employed in that profession and even give tips for applying for a job (Björksten et al. 2020, n.p.). The website listing different types of chatbots has thousands of entries in multiple categories (Botlist 2020).

One industry where chatbots have become common is customer service (O'Brien 2019, 4). One reason for this could be the benefits that the chatbot provides businesswise. For example, chatbots are in general quicker in answering the users' questions than a human (Khan et al. 2017, n.p.). Chatbots are also a cheaper option than customer service representatives (Lester, Branting & Mott 2004, 3). For users, chatbots enable continuous service (Dal Porto 2017, 6). And as AI solutions are becoming more advanced, the number of customer service chatbots will likely become greater (O'Brien 2019, 4). For instance, according to research and advisory company Gartner's (Panetta 2017, n.p.) estimates by year 2021, more than 50% of enterprises will spend more on developing bots and chatbots rather than mobile applications. They also listed chatbots and virtual customer assistants as one of the top technological trends in the customer service industry that will garner more investments in the future (Blum 2020, n.p.). However, chatbots will most likely not replace customer service representatives, instead they will change the role of service representatives and help with repetitive tasks (Dal Porto 2017, 9). As chatbots have become more common as a first customer service experience which the user has, it is important – and in the interest of the user and the company – that this experience is good. Improving customer service is important for a business, because failed customer service can lead to both loss of revenue and customers (Lester et al. 2004, 3).

In previous studies, the focus has been on the communication between the users and the chatbots (Liu & Sundar 2018; Skjuve, Haugstveit, Følstad & Brandtzaeg 2019). Liu et al. (2018) examined whether a chatbot should express sympathy to the user whereas Skjuve et al. (2019) examined user experience with chatbots. There has also been research on the usability of the technological solution for a specific task (Saenz, Burgess, Gustitis, Mena & Sasangohar 2017). Some previous studies have specifically focused on the customer service chatbots (Følstad, Nordheim & Bjørkli 2018; Følstad & Skjuve 2019). These studies have focused on the user's trust in a customer service chatbot (Følstad et al. 2018) as well as user experience and motivation with a customer service chatbot (Følstad et al. 2019).

However, there are not many usability studies on the customer service chatbots and even less on the answers they provide. This study focuses on the answers to bring a new perspective into the usability studies on the chatbots. The user should understand the answers for the chatbot itself to be usable. Thus, it is relevant to examine these answers especially in a field in which they are becoming a common communication format.

These answers are often defined in conversation diagrams which function as directions for chatbots as how the conversation should happen (Williams 2018, n.p.). Depending on the technological solution of a chatbot, someone has to write these conversation diagrams (ibid.). Also, chatbot replies need to be designed, because the personality of the chatbot affects the way users respond to it (ibid.). This can be done by a technical writer, who is already controlling the help material in the company. For instance, at Danfoss the technical communications department has been developing their chatbot, as they manage the documentation in the company (Savola 2018). Thus, their chatbot can utilize already existing instructional material, and everything does not need to be written just for the chatbot (ibid.). Also, technical communicators in the company often have knowledge regarding the end-users. Knowledge regarding the users is important in developing an efficient chatbot that answers to the users' needs (Williams 2018, n.p.).

Thus, the research question is: How usable are the answers that customer service chatbots provide? This study researches the usability of customer service chatbots of student housing associations. These chatbots are designed to solve problems and provide answers to the users – therefore their usability is directly linked to the answers. Most of the user experience is created through these answers, not the technical solution that brings them to the user. If the provided answers are unusable to the user, the chatbot's fundamental purpose is not fulfilled.

Thus, this research focuses on the usability of the answers and whether those answers provide the value to the user they are supposed to.

This study's research field is in technical communication. Even though the focus is on the answers produced in interaction with the chatbot, this study will not focus on the interaction itself. The approach begins from my argument that the chatbot answers can be examined as instructional texts. As I define the term chatbot, I will also discuss the different types of chatbots and their characteristics. During this I will also explore the types of tasks chatbots are used for within customer service. Then I compare these tasks to the different types of technical communication products such as software user documentation. Based on the similarities with chatbot tasks and types of documentation, I will argue that chatbot answers can be examined as instructional texts.

In addition to exploring the nature of chatbots, the theoretical framework in this study consists of Nielsen's (1994) usability theory. In parallel I will also benefit from the usability aims defined by Shneiderman, Plaisant, Cohen, Jacobs, Elmqvist & Diakopoulos (2017). While discussing the different user types I will also include some aspects of cognitive psychology. Because I focus on usability of chatbots as instructional texts, I will also cover a few guidelines for effective and usable instructional texts.

The chatbots in this study are from six different student housing associations in Finland. The actual study material comprises the answers collected from them. The answers were collected with a set of questions, which was developed from the frequently asked questions on the associations' websites. In developing the questions, I also used customer journey mapping which is a management tool that visualizes the customer experience throughout the purchase process (Rosenbaum, Otalora & Ramirez 2017, n.p.). The questions are placed on a simplified customer journey map, which I will define for this study.

As usability evaluation method I will use cognitive walkthrough. With the walkthrough I will evaluate the usability of each chatbot's answers separately. In the individual walkthroughs I will examine the answers and their characteristics. After these walkthroughs, I will compare the answers to each other.

The definition and characteristics of a chatbot will be explored in chapter 2. In this chapter, I will also discuss the reasons why chatbot answers resemble instructional texts. In chapter 3, I will discuss usability, its aspects and its evaluation. The collection of the study material and

the evaluation method, cognitive walkthrough, will be presented in chapter 4. The chatbot answers will be evaluated individually and compared between each other in chapter 5. In chapter 6, I will discuss the conclusions of this study.

2 DEFINITION OF A CHATBOT

In this chapter, I will define a chatbot, which is the focal term in this study, as well as explore different types of chatbots and how they function. I will also discuss the different requirements for a functional chatbot from a technological and business standpoint, because these requirements can highlight the usage and the target users for chatbots. As I explore these, I use the following terms to keep the terminology consistent. I use *user input* to describe the written or spoken natural language message through which the user communicates with the chatbot. *UI* is used as an abbreviation for *user interface,* which means the method of communication between the users and a machine, a program or a system (Griffin & Baston 2014, n.p.).

Chatbot is the main term used in this study for describing different types of bots. However, as a term chatbot can be confused with terms *conversational agent* and *dialog system*, and sometimes creating distinctions between them is complicated. For instance, sometimes these terms are used to describe different *conversational UI*s and the key difference is the implementation method (Janarthanam 2017, n.p.). Sometimes, *conversational UI* is even used as a definition for a chatbot (Mayo 2017, n.p; Shevat 2017, n.p.). Thus, the way to distinguish between these systems is to consider how they are integrated, what modalities they have, and through which channels they are deployed (Janarthanam 2017, n.p.). *Dialog system* is a system which can have a conversation with another party, which usually is a user (Klüwer 2011, 3), whereas *conversational agent* is a type of dialog system (Radziwill & Benton 2017, n.p.). Essentially, *chatbot* is a type of *conversational agent* but there are other types of *conversational agents* such as *embodied conversational agents* (ibid.). *Embodied conversational agents* are programs embodied as animals, avatars or humans, unlike chatbots (ibid.). Figure 1 on the left describes the relationship between these systems. As Figure 1 shows, *chatbots* are a type of *conversational agent*, and *conversational agents* are a type of *dialog system*.

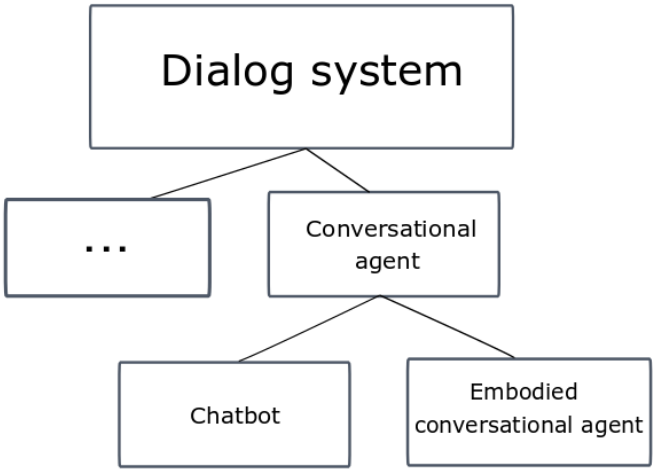


Figure 1. Relationships between software-based dialog systems. It is my modification from Radziwill & Benton's (2017, n.p.) Figure 2: Relationships between classes of software-based dialog systems.

A bot can be defined as a program, which based on its guidelines, can independently perform repetitive or routine tasks (TSK 2020). As I earlier mentioned, chatbots can be defined as computer programs that process user input (Khan et al. 2017, n.p.). This input is in natural language and in written format (Dale 2016, 3). As a basis for the idea of chatbots, the Turing Test from the 1950s is usually mentioned (Deshpande, Shahane, Gadre, Deshpande & Prachi 2017, n.p.; Shevat 2017, n.p.; Mathur et al. 2018, 1). Turing (1950, 445–460) developed the idea of a learning machine that can be taught similarly to a child through punishments and rewards. Furthermore, Turing (1950, 433-435) developed the Turing Test that he called the imitation game to evaluate the intelligence of a program. To pass the Turing Test, a program should be able to converse with a human so that the human cannot recognize it to be a program. Later, the Loebner Price Competition for chatbots was established to test whether chatbots can pass the Turing Test (Bradeško & Mladenić 2012, n.p). The competition is won by a chatbot that appears most human from the other competing chatbots (ibid.).

One of the earliest chatbots is ELIZA, which was developed by Weizenbaum in 1966 (Mathur et al. 2018, 1). ELIZA was a program that could answer to user input in natural language and could be taught better responses (Weizenbaum 1966, 36-37). ELIZA was programmed to talk similarly to a Rogerian psychotherapist (ibid., 42). It used pattern matching to determine key words from the user input and basic context in which the key word was stated (ibid., 37). Naturally, ELIZA had some limitations in its interactions. For instance, the user could not use the question mark, because that was interpreted as a delete character in the system on which ELIZA operated. (Weizenbaum 1966, 36.) A modernized version of ELIZA is still available online (Eclectic Energies 2020). ELIZA is a great example for how long the technology for chatbots has been around.

However, chatbots have changed over the years with technological development (Khan et al. 2017, n.p.). Therefore, modern chatbots can differ from the previous definition. For instance, modern chatbots can process verbal user input and respond verbally (Bruner & Barlow 2016, n.p.; Deshpande et al. 2017, n.p.; Shevat 2017, n.p.). Some modern chatbots are also able to perform tasks instead of just answering questions (Deshpande et al. 2017, n.p.). Therefore, modern chatbots, such as Alexa or Siri, are so advanced that the old definition does not cover all the tasks they can perform. In fact, sometimes Alexa, Siri and other voice activated bots are classified as voice-activated conversational agents (Radziwill & Benton 2017, n.p.) to distinguish them from chatbots.

The chatbots in this study reflect the more traditional description of chatbots as they are used through text and do not perform tasks for the user. Even though they are more advanced than ELIZA was, they still resemble it in their basic functionality. Therefore, the bots in this study can be classified as chatbots. I will also use the term chatbot in my examples and descriptions to maintain focus on them, but I do acknowledge that it is not self-explanatory whether a program is a chatbot or not.

2.1 Chatbot characteristics

In this subchapter, I will explore different characteristics of chatbots. With this exploration, my intention is to go beyond the definition of chatbots and provide more insight into them. I will discuss different elements of a chatbot's implementation and their UI. I will also briefly discuss social characteristics of chatbots. The focus in this subchapter is on modern chatbots, because the chatbots in this study are modern as well.

Many chatbots can be built on top of existing platforms such as Slack and Facebook Messenger (Bruner et al. 2016, n.p.). Thus, almost anyone can build their own chatbot and benefit from the existing resources for the chatbot (ibid). This practice lowers the costs of implementing a chatbot as it can be used through the platform (ibid.). This also makes it possible that the chatbot does not need to be downloaded by the user before it can be used (ibid.). Instead, the user can, for instance, call the chatbot by the name in the application and use it immediately (ibid.). For instance, IBM Watson and Microsoft Bot Framework are popular platforms for building chatbots (Davydova 2017, n.p.). These types of platforms provide the framework for a chatbot and tools to develop it further (ibid.). They are the way through which a user interacts with a chatbot. According to GoodFirms research (Sebastian 2019, n.p), websites are still the most preferred platform for using chatbots. However, among younger generations, messenger and mobile applications are more popular.

One thing in common with different chatbots is that they use a conversational UI (Batish 2018, n.p.). Conversational UI is modelled after a text interaction in which a bot answers the user's questions (ibid.). According to Batish (2018, n.p.), conversational UI can be defined by four characteristics. First, the UI is in written or oral format. Second, it is between two participants and the other one is a form of a computer. Third, it enables natural conversation between the participants, even though conversational ideas might not be exchanged between them. Fourth, it learns and is taught by enabling artificial intelligence (AI), machine learning (ML), deep learning (DL) and natural language understanding (NLU). Artificial intelligence

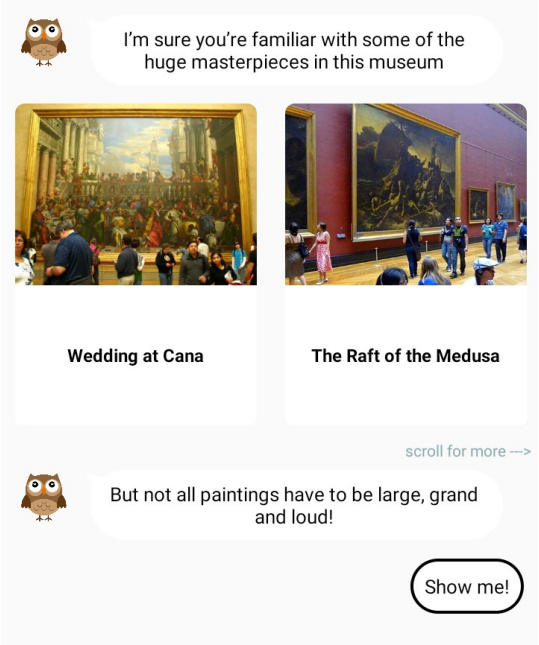
is a field which focuses on developing systems that can express characteristics associated with human intelligence (Tecuci 2012, 168). Machine learning is a branch of artificial intelligence that uses intelligent software to enable machines to perform tasks skilfully (Mohammed, Bashier & Khan 2016, n.p.). The intelligent software is taught through statistical learning methods (ibid.). Deep learning is a method of machine learning which can teach the intelligent systems more and more complex functions (LeCun, Bengio & Hinton 2015, 436). AI, ML and DL are not discussed in depth later as the technical solutions within chatbots are not the focus of this study.

Natural language understanding, NLU, is needed for the chatbot to comprehend the users' intent (Batish 2018, n.p.). Related, chatbots also need at least some level of natural language processing (NLP) (Lester et al. 2004, 4). This means that the chatbot has to connect the user input to an action the chatbot can accomplish (ibid.). Pattern matching is the most common NLP used in chatbots (Bradeško et al. 2012, n.p.). As I mentioned earlier, the first chatbot Eliza also used pattern matching. Another common approach is parsing. According to Bradeško et al. (2012, n.p), in parsing the chatbot breaks the user input into a set of words with features. Early parsing methods were simple and only searched for certain words in a certain order. For example, the sentences “please take the gold” and “can you get the gold” would both be parsed into a sentence “take gold”. Modern, more complex chatbots can grammatically parse entire sentences. Based on the commonality of pattern matching in chatbots, it is likely that the chatbots in this study also benefit from it.

The conversational UI of a chatbot is usually based on a chat interface (Khan & Das 2017, n.p.). Interacting with a chatbot resembles an online chat. Chat or online chat is real time conversation through exchanging messages (Oxford 2016). Common messages in the chat are a welcome message and a default message (Janarthanam 2017, n.p.). Welcome message is the message which a chatbot sends to a user as they open a chat with a chatbot (ibid.). Default message is the message which a chatbot sends to a user when it does not know how to answer to the user (ibid.).

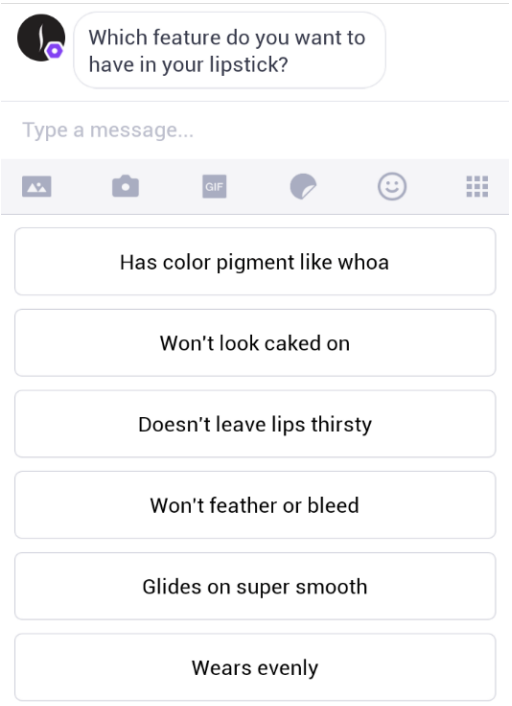
Besides the textual elements, chatbots can have several other UI elements to help the user perceive data or quickly answer the questions (Khan et al. 2017, n.p.). These elements include carousels, quick replies, buttons and web views (ibid.). These elements differ from the conversational UI and bring elements of graphical UI into a chatbot (Batish 2018, n.p.). Graphical UI includes more action, such as clicking on something on the screen, instead of

writing (ibid.). I will now introduce these UI elements in more detail with example pictures. Some of the introduced UI elements are also featured in the chatbots of this study.



Picture 1. Screenshot of my chat with Louvre chatbot on 3 February 2020.

Carousels are compilations of items that the user can browse horizontally (Khan et al. 2017, n.p.). Carousels resemble cards that can include an image, a title and buttons (ibid.). Usually, the user can click on a card in a carousel to be directed to a website or to start a conversation regarding the subject of the card (ibid.). In Picture 1, whilst chatting with Louvre chatbot, the sizable paintings of Louvre are displayed in a carousel. The paintings are in their own cards which include an image and the names of the paintings below them. The user can scroll through the paintings and tap on them to learn more about them. Picture 1 shows that Louvre chatbot also directs the user to scroll for more paintings with the instructions to *scroll for more* and an arrow to the direction of scrolling.



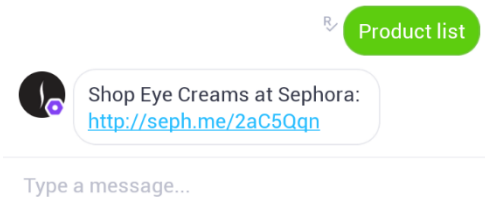
Picture 2. Screenshot of my chat with Sephora's chatbot on 3 February 2020.

Quick replies are buttons that appear on the screen as possible answers to the question, so that the user does not have to necessarily type the answer (Khan et al. 2017, n.p.). Quick replies disappear after the user chooses one of them or responds by typing a message (Janarthanam 2017, n.p.). Quick replies are useful in situations with multiple choices or with context sensitive options (ibid.). In Picture 2, there are multiple quick replies for a user who is searching for a lipstick with Sephora's chatbot. Once the user clicks on one them, the quick reply text goes to the chat similarly as if the user had written it themselves. Other options disappear.

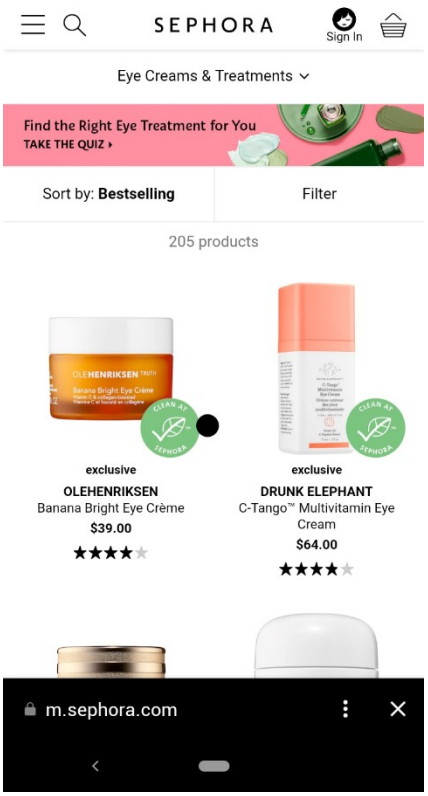
In comparison to quick replies, buttons are used to choose between options, and they do not disappear from the screen after the user taps or clicks on one (Khan et al. 2017, n.p.). Buttons can contain longer text content than quick replies can (ibid.). The button tells the user what action will happen if they click on it (Android 2020a). In picture 3, the bottom of the chat with Louvre chatbot has an exit button for the user to leave the conversation.

Picture 3. Screenshot of my chat with Louvre chatbot on 3 February 2020.

Web views are elements that can display the information from a web page that would not fit the chat (Khan et al. 2017, n.p.). A web view can open a website for the user (ibid.). A web view is an embedded browser which enables the user to view web information without opening a separate browser (Android 2020b). Pictures 4 and 5 display a link for eye creams when the user is shopping with Sephora's chatbot. When the user clicks the link, a web view opens to the Sephora's eye cream selection as is shown in Picture 5. Picture 5 shows a black bar at the bottom with an X button. When the user clicks it, they are back in the chat with the chatbot. Thus, the user does not have to return to the chatbot by opening it again.



Picture 4. Screenshot of my chat with Sephora's chatbot on 3 February 2020.



Picture 5. Screenshot of my chat with Sephora's chatbot on 3 February 2020.

Lastly, I will briefly discuss social characteristics of chatbots. Social characteristics are evident in their interaction with a user, so they are also represented in a chatbot's messages to a user. Unconsciously, users react to media in a similar manner as they react to people; this is called the media equation (Reeves & Nass 1996, 251-253). This means that users attribute characteristics and personality to the media they are interacting with (ibid., 253). Media in this case means computers, TV and other media which can be used for communication (ibid., 5).

Chaves and Gerosa (2019) did a survey of different social characteristics that have been found useful for chatbots in multiple other studies. They divided these characteristics into three groups: conversational intelligence, social intelligence and personification. Conversational intelligence represented characteristics that facilitate the conversation management (ibid., 2–3). A chatbot with conversational intelligence can actively participate in the conversation with the user (ibid., 3). It can also demonstrate awareness about the topic in discussion, the context of the conversation and the conversation flow (ibid). Social intelligence is the chatbots ability to behave in accordance with socially acceptable protocols for a conversation (ibid., 12)

Personification represented how the personality and identity of chatbots was perceived by users (ibid., 2–3).

A chatbot should have a personality that defines the way it interacts with the users (Batish 2018, n.p.). Besides interaction style, personality should describe the character of the chatbot and enable the user to understand its general behaviour (De Angeli, Lynch & Johnson 2001, n.p.). A chatbot should display a consistent and stable personality (ibid.). If the personality is unpredictable and unexpected, the user is uncomfortable with the chatbot (ibid.). Personality is also important for building trust between the user and the bot (Batish 2018, n.p.). Personality affects the tone and style of the chatbot's speech, for example word choices and whether it uses abbreviations or emojis (ibid.). According to Moran (2016, n.p.), tone of voice can be compared in four dimensions. First, whether the tone is funny or serious. Second, whether the tone is formal or casual. Third, whether the tone is respectful or irreverent. And fourth, whether the tone is enthusiastic or matter of fact. The tone can be somewhere between dimensions or at the other end of the spectrum (ibid.). The differences in the tone of voice is represented well in Pictures 1 and 2. Picture 1 shows that Louvre chatbot is more formal even though it is enthusiastic. Picture 2 shows that Sephora's chatbot is more informal and purposefully humorous.

In this study, the chatbot's platform, its UI elements or its personality are not directly evaluated. However, these characteristics affect the interaction with a chatbot, so it is possible that some notions related to them might surface. Especially the social characteristics or the personality of the chatbots might be present in their answers.

2.2 Types of chatbots

Chatbots can be divided into different types depending on the approach. I do not intent to explore every possible type of a chatbot, but I will examine a few methods of categorizing them. I will categorize them based on technological solutions, the services they provide, intended use and the businesses in which they are used. My expertise is not in the technological approach, nor is it the focus in this study, therefore the technological categorizations are only on a general level. More important for this study are the categorization in businesses. This categorization reveals what capabilities the chatbots are expected to have to be useful to the user and the company implementing it. Because the chatbots in this study are customer service chatbots, I will especially focus on them.

Based on the technological solution, chatbots can be separated into two categories: rule-orientated and AI-powered chatbots (Hassan 2019, n.p.; Hupli 2018, n.p.). Rule-orientated chatbots have been programmed to follow a set of rules (ibid.). They follow a dialog which is written by a human and cannot answer questions from the user that are not in the prewritten dialog (ibid.). Therefore, this type of chatbot is limited in its functions to respond only to programmed commands (Schlicht 2016, n.p.). In comparison, AI-powered chatbots benefit from AI-solutions and/or machine learning (Hassan 2019, n.p.). They are more advanced in understanding users' inputs and can even predict the users' needs (ibid.). They do not just follow commands but can analyze the input from the user (Schlicht 2016, n.p.). However, the answers that the chatbot gives to users are usually written by a human and not generated by the chatbot itself (Hupli 2018, n.p.).

Besides technological categorizations, chatbots can be grouped based on the services they provide. As an overall categorization, chatbots can be divided into domain specific chatbots and super bots (Shevat 2017, n.p.). A domain specific bot focuses on providing a certain service to the user and it performs tasks related to that service (ibid.). In comparison, super bots are not focused on single service (ibid.). For instance, Google assistant is a super bot, which can provide multiple services, including calling someone and finding a travel route from the map application (ibid.). In some cases, Google assistant can even provide subservices under its services (ibid.). Chatbots in this study are domain specific as they do not provide multiple services other than the basic customer service.

In addition, chatbots can be divided by their intended usage. Business chatbots are used for work by a company (Shevat 2017, n.p.). Business chatbots can facilitate complicated work processes and improve communication between employees (ibid.). They can also be used to automate repeated tasks such as clearing expenses within the company (ibid.). In contrast, consumer chatbots are used by customers of a company (Shevat 2017, n.p.). Consumer chatbots have a wider range of possible uses than the business chatbots (ibid.). They can be used for entertainment, shopping, information retrieval or even to improve productivity (ibid.). Consumer chatbots can have more personality than business chatbots, because they need to be more focused on the customer experience whereas business chatbots must function more precisely (ibid.). Consumer chatbots could also be called enterprise assistants. Enterprise assistants are designed after customer service representatives and store assistants (Janarthanam 2017, n.p.). Their purpose is to serve customers (ibid.).

In contrast to enterprise assistants, personal assistants are chatbots which focus on helping with user's personal needs (Janarthanam 2017, n.p.) They can help in personal tasks, for example managing user's calendar, listening to music and answering calls (ibid.). For instance, Alexa is a personal assistant (ibid). Another difference between personal assistants and enterprise assistants is that personal assistants can be extended to do more (ibid.). For instance, PizzaHut and Starbucks have developed features related to their own business that can be taken into use with a personal assistant Alexa (ibid.). The chatbots in this study are customer service chatbots, therefore they are also consumer or enterprise chatbots. They are used by the customers of the company whom I will refer to as users.

Lastly, chatbots can be categorized based on the business fields in which they are used (Lester et al. 2–3). According to Lester et al. (2004, 2–3), these can be the following five categories: customer service, help desk, website navigation, guided selling, and technical support. Three of the businesses focus on answering the users' questions: customer service, help desk and technical support. Website navigation and guided selling are more focused on general guidance and providing support for users' tasks (navigation and buying), but they also need to respond to possible questions along the process. However, all these businesses need to have a dialog with their customers. Because the chatbots in this study are customer service chatbots, I will now focus on the aspects of customer service and chatbots role in it.

As customer service has begun to include more and more conversational solutions, especially messaging applications, Messina (2015, n.p.) created the term conversational commerce to describe this trend. Conversational commerce means using chat, messaging or other natural language interfaces to interact with brands, services or bots which previously have not been part of the bidirectional messaging context (Messina 2016, n.p.). It focuses on extreme personalization to every user (ibid.). The rise of conversational commerce is largely due to the communication trend and the developments in AI (Gentsch 2018, 92). Messaging services are established and popular, so companies want to offer their services through them (ibid.). The developments in AI enable that and the future development of conversational commerce as well (ibid.).

Chat-based interactions in conversational commerce include among other things: sending the customer order confirmation and shipping information, thanking them for their purchase, recommending products based on their purchase history and the conversation with them, and troubleshooting (Schlicht 2018, n.p.). Troubleshooting in this context means helping the

customer with problems they might have with the purchased product or service (ibid.). While the customer is still browsing the service online, a chatbot can act similar to a customer service representative in a physical store (ibid.). It can answer questions about products, provide offers and explain more about the services (ibid.). Thus, chatbots can perform multiple different type of things for the user. Some of them are more task focused, for instance the troubleshooting, while others are more informative, for instance answering the customer's questions. As chatbots in this study are customer service chatbots, they could be expected to perform these types of tasks.

2.3 Chatbot answers as instructional texts

In this subchapter, I will argue that chatbot answers can be seen as a form of technical communication, and even further, can be examined as instructional texts. To present this argument, I will introduce definitions and characteristics of both technical communication and software user documentation, comparing them to chatbot answers. My focus is on customer service chatbots and their answers, because all the chatbots in this study are customer service chatbots.

As I mentioned earlier in the introduction, technical communicators have begun writing content for chatbots. More specifically, they are altering the existing content for chatbots to use as answers. Because technical communicators are producing content for chatbots, it could be interpreted that chatbots fall under the tasks of technical communication. However, defining technical communication is not easy, even inside the discipline itself (Allen 1996, 9). Technical communication includes multiple actions, for instance writing, designing and technical illustration (Jones 1996, v). Usually its subjects are science and technology, but it can include other subjects as well (ibid.). Technical communication has also been defined as writing which someone does in their profession or discipline and is meant to elicit a behavioural response from the reader (Stratton 1996, 39). Another possible definition is that it manages technical information to allow the readers to act (Priest 2010, 865). An even more modern definition of technical communication describes it as delivering clear, consistent and factual information to the users (TCBOK 2020). This definition in the Technical Communication Body of Knowledge by the Society for Technical Communication (TCBOK 2020) states that "technical communication is a user-centered approach for providing the right information, in the right way, at the right time so that the user's life is more productive". The

products which technical communicators produce are multiple, including how-to guides, online helps and user interface texts (ibid.).

Compared to these definitions, customer service chatbots and their answers can be classified under technical communication. They can provide answers which are meant to elicit an action from the user. As I mentioned in chapter 2.2, they can help the user with troubleshooting or shopping. The customer service chatbots can also provide factual information which enables productivity in the user's life. Overall, chatbots' answers can be characterized as instructional texts which the technical communication products usually are; instructional texts help the user reach their goals and perform an action. If we further consider what form of technical communication the chatbots answers could be, it seems sensible to turn to online helps or software user documentation. A chatbot, after all, is a software communicating with a user.

According to IEEE 1063-2001 (2001, 3), software user documentation is "electronic or printed body of material that provides information to users of software". Based on this definition, chatbot answers could be defined as software user documentation: the answer is electronic material, which provides information to the users of chatbot. However, this is not a straightforward comparison. According to ISO/IEC/IEEE 26511 (2012, 6), user documentation is defined as: "information to describe, explain, or instruct how to use software". As chatbots provide information related to the issues other than how to use the software, they would not qualify as software user documentation according to this international standard. According to Simpson and Casey (1988, 11), all software user documentation has the same purpose, which is to provide information of the software features and help the user gain proficiency in using it.

One form of software user documentation is embedded user documentation. Embedded user documentation or embedded help is accessed through the user interface and does not entirely cover the task which the user is performing at the moment (Ames 2001, 114). Embedded help can open in another window or a pane to display the instructions (ibid.). The ISO/IEC/IEEE 24765 (2017, 122) defines embedded documentation as documentation which is accessed as an integral part of the software. For example, a pop-up help and a help text on the screen are embedded documentation (ibid.). Online help systems are to assist the user to accomplish tasks with a software and to help users solve problems with user interface, process, options or other elements (Ray & Ray 2001, 105).

Chatbots in this study offer information on processes, such as the application process, to help the user gain proficiency in going through those processes. Therefore, it does have similarities with software user documentation. However, it obviously is not the same thing as it does not provide information on the use of the software. Of course, if a chatbot is designed to answer software related questions, then it could be described as software user documentation. Also, the presentation of chatbot answers is similar to embedded user documentation. If the chatbot is used on a website, it has its own chat interface, which does not cover the entire screen and the task at hand. If we further consider the nature of chatbots' answers, the concept of information type from technical communication becomes useful. Actually, the previous direct quote from ISO/IEC/IEEE 26511 (2012, 6) alludes to information types, as it mentions information that can explain or instruct.

Although technical communication products are typically referred to as instructional texts or instructions, they can include many information types (Karreman, Ummelen & Steehouder 2005, 328). The most important information types are procedural and declarative (ibid.). Procedural information concerns the user's actions and while using the system, it is the most important information type (ibid.). Procedural information is often written in a step-by-step fashion (Estrin & Elliot 1990, 50; Simpson et al. 1988, 10). The user reads procedural information because they want to perform a task (Ummelen 1994, 117).

In comparison, declarative information is explanatory (Karreman et al. 2005, 328). It includes the necessary facts to know about the system and is the foundation for learning the system (Simpson et al. 1988, 10). The user reads declarative information because they want to learn about the system and be able to use it without instructions (Ummelen 1994, 117).

However, the distinction between procedural and declarative information is not entirely clear (Karreman et al. 2005, 330; Ummelen 1994, 123). They are broad terms which encompass many subtypes of information (Ummelen 1994, 124). Besides the content of the text, also its form can be either declarative or procedural (ibid.). Procedural form could mean instructive form and declarative form could mean narrative, argumentative or descriptive form (ibid., 123).

Customer service chatbots can provide answers including both procedural and declarative information. They can provide the necessary facts about the service or product, which the user needs. They can also provide procedural information by describing how to perform an action.

For instance, in shopping or troubleshooting situations mentioned in chapter 2.2, the chatbot can describe the necessary steps to the action. Answers which state the necessary facts are declarative, whereas answers which include the steps to performing a task are procedural. However, distinguishing every answer only into procedural or declarative information types might be difficult and even impossible. The answers can include both types of information, so making a clear distinction can be difficult. Thus, when we look at the chatbot answers from the viewpoint of information types, they resemble other technical communication products.

3 USABILITY

In this chapter, I will examine usability. As usability is closely related to user experience, I will begin by describing their relationship and the importance that user experience has. This is not the focus of the theoretical framework of this study, but it is to provide context. Next, I will define usability and its central aspects out of which the learnability will be central in this study. Then I discuss how users can be characterized in usability research and how usability testing is performed. Lastly, I will examine usability heuristics developed for documentation, as in this study I approach the chatbot answers as a form of technical communication.

According to the ISO/IEC/IEEE 24765 (2017, 495–496), user experience is the perceptions and responses of the user that result from the use of a system. User experience depends on multiple aspects, for example the system functionalities, brand and user's attributes. The user experience includes all the touchpoints through which the user interacts with the product brand (Rosenzweig 2015, 8). This includes the store, website, online help and other possible touchpoints (ibid., 8). The user experience is an abstract concept that is divided into smaller parts as it is difficult to describe otherwise (Sinkkonen, Kuoppala, Parkkinen & Vastamäki 2009, 225). The idea behind user experience is that the user and their experience with the system should be considered in the design of the system (Rosenzweig 2015, 10–11). Thus, the design should also consider the limitations the users have, for example that stressed users are more likely to make mistakes (ibid., 11). Usability is considered part of user experience (ibid., 7). Usability as a field benefits from cognitive psychology and human-computer interaction research (Sinkkonen et al. 2009, 12).

In this study, the main theoretical framework is Nielsen's usability theory, which I will introduce next. Nielsen (1994, 24) uses the term system acceptability to describe whether the system can satisfy the demands of the users and the potential stakeholders. This system acceptability is a combination of social and practical acceptability (ibid.). Practical acceptability is further divided into usefulness and other attributes, such as cost and reliability (ibid., 24–25). Usefulness describes whether the system can be used to achieve a desired aim and can be divided into utility and usability (ibid., 24). Nielsen (1994, 25) defines them in them in the following manner: *"*--utility is the question of whether the functionality of the system in principle can do what is needed, and usability is the question of how well users can use that functionality.*"* Thus, usability and utility together describe how useful a system is (Nielsen 2012, n.p.). ISO/IEC/IEEE 24765 (2017, 492) defines usability as "extent to which a

system, product or service can be used by specified users to achieve specified goals with the effectiveness, efficiency and satisfaction in a specified context of use".

Nielsen (1994, 26–27) highlights that usability should be systematically approached, evaluated and improved in the system. The idea is to go beyond notions like "intuitive" and "natural" in the design of the system (Shneiderman et al. 2017, 33). As usability affects the entire system and the processes user performs with it (Nielsen 1994, 25), it is an important attribute to improve. For instance, usability is paramount for a website as users leave the website immediately if the website fails to provide what they need (Nielsen 2012, n.p.). Because websites are common and plenty, users can just switch to another if there is a problem with usability (ibid.).

3.1 Aspects of usability

In this subchapter, I will discuss Nielsen's (1994, 26) usability theory in which usability is divided into aspects. In parallel I will introduce the usability measures by Shneiderman et al. (2017, 33–34). The purpose is to further discuss usability and what it entails.

Usability is composed of five different aspects: learnability, efficiency, memorability, errors, and satisfaction (Nielsen 1994, 26). These attributes are studied with different usability methods to show whether they have been accounted for in the design (ibid., 27). Similar aspects are included in Shneiderman et al.'s (2017, 33–34) definition of usability measures: time to learn, speed of performance, rate of errors by users, retention over time, and subjective satisfaction.

Learnability refers to the ease with which a user can learn to use the program (Nielsen 1994, 27–28). Learnability does not mean users have to learn everything about the UI and the program, but that they can use it in a needed level (ibid., 28–29). Learnability is easily measured with novice users, as they need to learn everything about the program for the first time (ibid.). Time to learn refers to the same principle as learnability: it measures how quickly a typical user from the user community learns to perform relevant actions (Shneiderman et al. 2017, 33). Users can learn to use even difficult systems with training, but it is not necessarily an effective solution as it takes time and money (Sinkkonen et al. 2009, 194).

Efficiency means the time it takes a user of certain expertise level to perform tasks (Nielsen 1994, 30–31). Thus, efficiency can be measured with expert users (ibid.). Usually, this means

just deciding what constitutes as an expert user (ibid., 31). Then, it is measured how quickly they can perform tasks with the system (ibid., 30–31). Shneiderman et al.'s (2017, 34) speed of performance focuses more on the time it takes a user to perform basic tasks with the program. An efficient system is consistent in workflows and terminology, and its structure is explicit (Sinkkonen et al. 2009, 194).

Memorability means how well users can remember the workflows of the program (Nielsen 1994, 31–32). This is often tested with casual users, who do not use the program all the time, and therefore can use the program, but are not experts (ibid.). However, many modern programs inform the user well while in use, so the user does not necessarily have to remember how it is used (ibid., 32). Memorability and retention over time have very similar definitions. Shneiderman et al. (2017, 34) also note that time to learn and frequency of use are closely linked to the retention over time. Repetition is important for remembering things but is not the only thing determining how memorable something is (Sinkkonen et al. 2009, 150). Users remember information that is relevant and easily connected to previous information better (ibid.).

Users should not make a significant amount of errors while using a program (Nielsen 1994, 32). Some errors are more fatal to the work the user is trying to accomplish, for example, errors that destroy or ruin the task that the user is trying to do (ibid., 32–33). These types of errors should be avoided as they greatly hinder usability (ibid., 33). The rate of errors by users also refers to the amount of errors users perform during basic tasks (Shneiderman et al. 2017, 34). The recovery time from errors could be accounted in the speed of performance, but because error handling is important, it should be studied on its own (ibid.). Errors that users make can usually be divided into two groups: *intentional errors* and *lapses* (Sinkkonen et al. 2009, 43). *Intentional errors* occur when the user chooses to perform a task taking a route that is not the fastest for performing it (ibid.). *Intentional errors* are caused by misunderstandings, wrong information, wrong inferences or wrong generalizations (ibid.). *Lapses* occur when the user has understood the aim correctly and their intentions are correct, but the task is performed in the wrong way (ibid.). *Lapses* are usually small mistakes, which can be easily fixed by the user (ibid.).

Satisfaction describes whether users enjoy using the program (Nielsen 1994, 33). For programs such as games, satisfaction is one of the most important aspects of usability (ibid.). Nielsen (1994, 33) notes that how users feel about the entire operating system should be

accounted in the social acceptance. Therefore, satisfaction is more related to the user experience with the program (ibid.). Subjective satisfaction informs whether users enjoy using the different aspects of the user interface and it can be measured with questionnaires or interviews (Shneiderman et al. 2017, 34). Identifying positive feelings, however, can be difficult even to the users themselves (Sinkkonen et al. 2009, 223). Therefore, usability research often focuses on what the users do not like as that is easier to identify (ibid., 225).

According to Shneiderman et al. (2017, 34–35), these aspects serve a function as a measurement for usability, but it might not be possible to attain all of them in the same program. For instance, if a speed of performance is high, then it might be necessary to have a longer time to learn. Therefore, it is important to understand the users' needs and which aspects they would want the most from the program (ibid.). It might also be a good idea to make it visible to the users which aspects have been sacrificed for another (ibid.). A usable interface is a necessary for the program to survive businesswise, as competition and demand for usability by users have risen (ibid.). Sometimes other requirements hinder the overall usability or an aspect of it, for instance, security concerns can cause a situation in which usability has to be partly sacrificed (Nielsen 1994, 42–43).

3.2 Users, usability testing and evaluation

I will discuss user profiling and briefly touch upon methods that can be used to assess usability. I will also include aspects of cognitive psychology in the discussion of users and their characteristics. In chapter 4.3, I will discuss the users in this study and the chosen method for usability evaluation.

Users are an important aspect of usability (Nielsen 1994, 43). Users can be grouped based on different characteristics. Based on previous experience, users can be grouped to novices, casual users and experts (ibid., 31, 43). Novices are users who do not really know anything about the system or the task (Shneiderman et al. 2017, 89). Novices are separate from first time users, who know about the task well but are unfamiliar with the system (ibid.). Casual users use the system intermittently (Nielsen 1994, 31). This type of users can also be called knowledgeable intermittent users (Shneiderman et al. 2017, 89). They are familiar with the tasks and have used the system before, but they might not remember the entire workflow in the system (ibid.). Experts are users who are experienced in using the system (Nielsen 1994, 30; Shneiderman et al. 2017, 90). Experience can be defined by the users themselves, or more formally by setting a number of hours the user has to have spent using the system (Nielsen

1994, 30). However, the distinction is not always clear, as a user might be an expert in using certain parts of the system but a novice in other parts of the system (ibid., 45). Also, users who have experience with multiple similar systems usually have better expertise with a new system compared to users who have experience with just one system (ibid., 45–46).

These different types of users have different needs and wants for the system (Shneiderman et al. 2017, 89-90). Novices need support in learning the system, such as tutorials and video demonstrations (ibid., 89). They benefit from positive feedback as a task is successfully accomplished (ibid.). When the user uses a system for the first time, they try to remember a similar system for analogy (Sinkkonen et al. 2009, 186). This might be a similar system or a system with similar user interface, buttons or even just a shape (ibid., 186). Casual, or knowledgeable intermittent users benefit from a system that allows relaxed exploration and emphasizes recognition rather than recall in the user interface (Shneiderman et al. 2017, 89). Expert users want to complete their tasks quickly and prefer short response times, shortcuts and non-distracting feedback after completing a task (ibid., 90). The system can be designed for users of multiple knowledge levels, but it is naturally more difficult than designing just for users of one knowledge level (ibid.).

Users naturally have other characteristics, such as cultural background and physical capabilities, out of which some are more personal and others more universal within their user group (Sinkkonen et al. 2009, 17–19). Language and cultural habits can be similar between users from the same culture whereas the place and surroundings during use may differ greatly (ibid.). For instance, a user's age affects the way they learn new things (ibid., 205). Thus, for an adult or elderly user, previous experiences and previously learned things are in an important part of the learning process (ibid., 205–206). However, aging is also a personal experience so it cannot be said users of the same age learn in the same way (ibid., 205). All users are, after all, humans that have different characteristics (Nielsen 1994, 43). There are also several psychological things which affect the use situation, such as the user's motivation, beliefs and emotions (Sinkkonen et al. 2009, 222–230). For instance, a positive state of mind can affect the situation so much that the user tolerates small usability issues (ibid., 222). Vice versa, having to constantly use unusable systems and products causes frustration and stress in users (ibid., 234–235).

As users, everyone brings to the situation their culture, background and personality. Thus, usability is not a definitive attribute of a system but something that can always be developed

further to make it more usable to some users or usable to a new user group. Usability can be measured in relation to certain users and certain tasks (Nielsen 1994, 27). Thus, the results of usability testing might be different when the system is used by different type of users or the performed tasks are changed (ibid.). In usability testing, a representative set of tasks is used to test usability (ibid.). To evaluate the overall usability of the system, the mean value of each aspect of usability can be calculated and compared to some agreed minimum (ibid.).

Usability testing is a fundamental usability method, as it is done with real users of the system (Nielsen 1994, 165). Usability testing reveals how users actually use the system and what issues they have with it (ibid.). The testing can be done for academic purposes to support theories and it can also be done to reveal usability issues on a singular system and develop it further (Shneiderman et al. 2017, 175). Usability testing is usually done in usability laboratories in which techniques, such as recording and user activity logging, can be utilized (ibid.,175–179). Proper usability testing takes multiple weeks including planning, piloting, testing and reporting (ibid.). The testing can be done with a prototype or with an existing product (ibid., 168).

Besides usability testing, there are several other methods for usability evaluation (Nielsen 1994, 207). These methods include methods such as observation, interviews and logging the actual usage of the system (ibid., 207–217). These methods can be used to gather additional information besides usability testing (ibid., 207). They include information from the real users, but they are lighter to implement compared to a proper usability testing (ibid., 207–217).

Besides usability testing, there are different expert reviews which have proven effective in evaluating usability (Shneiderman et al. 2017, 171). Expert review methods include heuristic evaluation, guidelines review, consistency inspection, cognitive walkthrough and formal usability inspection (ibid., 171–175). All these methods rely on an expert or experts, who can come from different fields (ibid. 171). These methods can be executed in a faster pace than a usability testing with real user participants (ibid). In this study, I will use cognitive walkthrough in usability assessment, but that is introduced in more detail in chapter 4.3.

Usability research has produced guidelines, principles and theories (Shneiderman et al. 2017, 82). The guidelines are advice about good practices (ibid.). The principles are rules to analyze and compare whilst choosing design alternatives (ibid.). The theories are frameworks to

utilize in design and evaluation of the system (ibid.). In the following chapter 3.4, I will examine some of the guidelines and principles with the aim of producing usable instructional texts.

3.4 Usability in instructional texts

In this subchapter, I will focus on usability in instructional texts by examining the usability heuristics for documentation (Purho 2000). The purpose of this subchapter is to have more information on the probable usability issues affecting documentation before evaluating the usability of the answers in chapter 5. As I previously argued in chapter 2.3, the chatbot answers can be classified as instructional texts and therefore, I want to examine the usability issues affecting them more closely.

For evaluating the usability of documentation, Purho (2000, n.p.) has developed ten usability heuristics. These heuristics describe what usable documentation is. They can be used as a checklist while designing the documentation (ibid.). I will also include mentions from the IEC/IEEE 82079-1 International standard for Preparation of information of use (instructions for use) of products (2019) as this standard includes many similar aspects to Purho's heuristics.

First, the documentation needs to match with the real world (Purho 2000, n.p.). The documentation should be in a language understandable to the user, including words, phrases and concepts familiar to the user. Information should also be presented in a logical order (ibid.). The information about a task should be in the order in which the actions are performed (IEC/IEEE 82079-1 2019, 44–45). In the terminology, any atypical variations of names and product features should be avoided (ibid., 31).

Second, the documentation should match the product for which it is provided (Purho 2000, n.p.). The same terminology should be used in both the documentation and the product (ibid.). This might contradict the first heuristic, if the product includes odd terminology (ibid.). The provided information about the product should enable the users to safely, efficiently and effectively use the product (IEC/IEEE 82079-1 2019, 20).

Third, the documentation should be purposeful (Purho 2000, n.p.). The purpose of each document and their intended use ought to be clear to the user (ibid.). This also includes the media in which information is presented to the user (ibid.). The media in which the

documentation is provided should accommodate the needs of the users (IEC/IEEE 82079-1 2019, 24–25). It should also be considered whether the documentation is offered embedded in the product or separately (ibid.).

Fourth, the documentation should support different types of users (Purho 2000, n.p.). This might be users of different knowledge levels, but also users who need to achieve different tasks in that domain (ibid.). If the users are not experts, terminology should be explained (IEC/IEEE 82079-1 2019, 31). This can be done, for example, by including definitions, by adding links or with glossaries (ibid.). If abbreviations, acronyms or technical terms cannot be avoided, they should be explained as well (ibid.).

Fifth, the information design should be effective (Purho 2000, n.p.). This means that information is presented in a way that it is easily found and understood (ibid.). Different ways of presenting information, for example graphics and tables, are used to support the user's information needs (ibid.). The text itself is written in a way that supports the user and their needs (ibid.). This includes using short lines and paragraphs, using imperative in instructions and addressing the user directly (ibid.). Each instructional step should include only a single action (IEC/IEEE 82079-1 2019, 45).

Sixth, the documentation should support different ways for searching information (Purho 2000, n.p.). The users have different search strategies and that should be accounted in documentation (ibid.). The layout of a document should also support browsing, so that the important information can be identified while browsing (ibid.). If possible, only the relevant information to the task should be presented to the user (IEC/IEEE 82079-1 2019, 46). Links to other relevant information can be provided, but they should not distract the user from the instructions (ibid., 47).

Seventh, the documentation should be task orientated (Purho 2000, n.p.). It should not focus on the tools which the user uses to achieve their goal, but the tasks they are performing for achieving that goal (ibid.). The provided information should be usable and relevant to the users with respect to their tasks and goals (IEC/IEEE 82079-1 2019, 20).

Eighth, there should be troubleshooting information available to the user (Purho 2000, n.p.). Troubleshooting provides guidance to the common issues and the information how to analyse rarer issues (ibid.). If there are troubleshooting procedures which require a skilled user, it is preferable to separate that information from troubleshooting procedures all users can perform

(IEC/IEEE 82079-1 2019, 37). For troubleshooting all users can perform, there should be instructions for the actual troubleshooting procedure and any testing which is done after the procedure (ibid.).

Ninth, the documentation should be consistent and adhere to standards (Purho 2000, n.p.). The documentation should be consistent, so that the terminology and actions mean the same thing (ibid.). There also should not be unnecessary overlap of the same information in different documents (ibid.). Consistent content is unambiguous and correct (IEC/IEEE 82079-1 2019, 22). The information about the use of the product should also be consistent with other information, which is provided, for instance, in training or promotional materials (ibid.).

Tenth, there should be help provided for using the documentation (Purho 2000, n.p.). If the document set is large, there should be instructions on how they are supposed to be used (ibid.). The documentation updating information can also be useful (ibid.).

Even though this study does not use heuristic evaluation as the evaluation method, these heuristics provide valuable insight to the usability in instructional texts. They show how many facets affect the usability of instructional texts. Besides knowing the information which the user needs, it has to be considered how this information is conveyed understandably to the user. I will return to some of these heuristics when comparing the chatbots' answers in chapter 5.7.

4 STUDY MATERIALS AND METHODS

In this chapter, I will first introduce the student housing associations and their chatbots, which are part of this study. Second, I will explain why and how I chose the questions that I asked from the chatbots. I benefitted from frequently asked questions and customer journey mapping in compiling the questions. Lastly, I will introduce the cognitive walkthrough as the main usability method that I will apply in analysing the answers of the chatbots. As part of the cognitive walkthrough, I will construct the action sequence for the tasks and profile the type of user in this study.

As I was searching for suitable subjects for this study, I considered different types of customer service chatbots. To make the comparing easier, I have chosen chatbots from the same industry and similar companies, so that the chatbots would ideally have as similar functions as possible. All chatbots are also from Finland and operate in Finnish. Thus, collecting study material from them and comparing them is less complicated because they have similar target audiences and their answers are in the same language.

To gather the study material from the chatbots, I had to ask them questions. These questions needed to reflect questions that a real user could ask the chatbot and not to be engineered for receiving certain type of answers, such as the correct way of solving the issue. The questions also cannot be guaranteed to produce answers from the chatbot, because a possible situation is that a chatbot does not have an answer to the user. Therefore, part of the reason for choosing these exact companies and chatbots was that they have frequently asked questions (FAQ) available on their websites alongside the chatbot. These frequently asked questions can give direction as to what real users could ask and what subjects they are interested in.

4.1 Customer service chatbots of the student housing associations

I have chosen six companies and their chatbots for this study. These companies are student housing associations located in different parts of Finland. Their chatbots are publicly available on their respective websites. Also, a reason for choosing these associations is that they have frequently asked questions or information pages on their websites. These questions or information pages include information on usual problems and questions that users have with the services of the associations. All the information pages resemble frequently asked questions in their content, but the formatting is different from the question and answer format in that it is descriptive text divided by titles. If the housing association had both frequently

asked questions and information pages on their website, I only used the frequently asked questions and not the information pages. Because frequently asked questions should be the most common questions that the users have (Christensson 2014, n.p.; Nielsen 2002, n.p.), they should reflect the needs and problems of the users. Table 1 below lists all the student housing associations, their chatbots and their websites.

Table 1. The student housing associations and their chatbots chosen for this study.

Name of the chatbot	Name of the student housing association	Website on which the chatbot is hosted
Helmi	HOAS - Helsingin seudun opiskelija-asuntosäätiö sr	hoas.fi
Kaija	KOAS - Keski-Suomen opiskelija-asuntosäätiö sr	koas.fi/fi
Toivo	TYS - Turun Ylioppilaskyläsäätiö	tys.fi
DAS bot	DAS - Domus Arctica - Säätiö	das.fi/fi
Elli	Joensuun Elli - Opiskelija-asunnot Oy	joensuunelli.fi/fi
Tane	TOAS - Tampereen opiskelija-asuntosäätiö sr	toas.fi

As can be seen from Table 1, one of the chatbots has a name which reveals it is a bot, *DAS bot*. Others have Finnish first names: *Elli, Helmi* and *Kaija* are common Finnish female names (Digi 2020a; Digi 2020b; Digi 2020c) whereas *Toivo* is a common male name (Digi 2020d). *Tane* is a slightly more common name among men than women, but it is a rare first name in Finland (Digi 2020e). Later in this study, I will refer to the chatbots by their names and using the third person *it* pronoun to highlight that I am discussing the chatbot.

These student housing associations provide apartments to students studying in post-comprehensive schools or in upper secondary and vocational schools (Tys, Toas). Tys also provides youth housing, which is meant for young adults, 18–29-year-olds, who do not study anything. A few associations provide summer housing to young adults who need temporary housing in the city due to summer studies or work (Toas, Hoas, Tys). The biggest associations are from southern Finland and they have 4,000–10,000 apartments (Toas, Hoas, Tys). The smaller associations have around 1,500–2,500 apartments (Das, Koas, Elli). Five of the student housing associations are foundations (Toas, Hoas, Koas, Das, Tys), and one is a limited company (Elli). The foundations are directed by delegations and boards (Toas, Hoas).

The foundations have different organizations behind them, such as the city and different student coalitions (ibid).

All the chatbots in this study appear to be simple customer service chatbots, which do not perform any user tasks. For example, they do not collect user information for the application process. They only answer questions from the user and all their introductory messages include a mention that they assist customer service. They have a simple chat window which can be opened and collapsed by the user. None of the chatbots are developed by the housing associations, instead they have an outside provider for the platform. All of the chatbots in this study are hosted on the *Giosg* platform, which is apparent because the chatbots have the *Giosg* logo at the bottom of their chat window. *Giosg* is a live chat platform for customer services which includes real time reporting and other services (Giosg 2020). Most of the chatbots appear to be developed by *Get Jenny,* which designs customer service chatbots (GetJenny 2020a). *Get Jenny* does deploy some AI assisted NLP on their chatbots (GetJenny 2020b), but it is unclear whether the chatbots of this study utilize this technology.

The housing associations' customer service chatbots are very suitable for the purposes of this study, because they are domain-specific, consumer chatbots. Their purpose is limited to answering questions to the possible customers of the associations. Because providing student housing is a clearly defined industry rather than one providing multi-services, profiling the possible customer, in other words, the user, is easier.

As I am also a student and have lived in apartments provided by student housing associations, I am part of their target audience. Thus, I have insight into the user experience and the customer journey one might have with a student housing association. This means that my role as an evaluator is varied. I am both the expert evaluating the usability and the possible user using the product. My personal experiences also increase my expectations regarding the services. For instance, I have applied for student housing in multiple different cities at different times and therefore know how quickly the associations usually respond with an apartment offer or when the deposit is paid back. Of course, the focus of this study is not in the quality of services or their functioning, but it is fair to acknowledge that I can have expectations towards these services, which their users might also have.

4.2 Collection of the study material from the chatbots

The study material is obtained with a set of questions to the chatbots. In this subchapter, I will explain the process of first gathering the base questions and then developing these questions to the final set of the questions. The final set of questions ought to resemble real questions from the users. The questions should also be compiled in a logical order from the user's point of view, so that questions regarding the apartment are not after questions about moving out of the apartment. Therefore, the questions are compiled and organized with the help of a customer journey. Customer journey is a tool used to visualize the customer experience throughout the purchase of products or services (Rosenbaum et al. 2017, n.p.). All the questions are in Finnish because that is the language used by the chatbots. The questions in English are my own free translations.

As mentioned, the base questions have been selected from the frequently asked questions (FAQ) and the information pages on the websites of the housing associations. From the FAQs I collected the questions into an Excel table. These were the base questions. The Excel table had two columns, one for the questions and the other for marking in how many FAQs the question was included. This way, I could choose the most common questions. If two questions had the same subject, but were formulated in different ways, for example if one was in passive voice and the other in active voice, I treated them as the same question. Because the information pages did not include questions as the FAQs did, I compared the topics on the information pages to the base questions I had collected from the FAQs. If the information pages included the same subjects, I marked them down as the same questions as the FAQs. After going through all the websites, I selected the questions from the table that were on two or more websites.

According to Hill, Ford & Farreras (2015, 248), users send shorter messages to chatbots than to people in a chat conversation. Users also send more messages to the chatbot (Hill et al. 2015, 248-249). In a conversation with a chatbot, users often start with a greeting, but do not necessarily say goodbye at the end of the conversation (Kopp, Gesellensetter, Krämer & Wachsmuth 2005, 11-12). Based on this information of the interactions between users and chatbots, I edited the base questions into the final questions.

Because the base questions had been collected from FAQs, their style was at times very formal. For instance, the question *What are the terms of returning the deposit* includes formal

terminology and is formulated in a way that applies to every user but is not the way a user might ask about their own deposit being returned. Many of the base questions were in passive voice, for instance *How to terminate the tenancy*. These types of base questions I edited to the active voice, as in *How can I terminate the tenancy*. Some of the base questions included references to housing association specific information, such as *myToas portal,* which refers to the electronic portal of the Toas housing association. These references I changed to a more generic term, such as *electronic portal*. A few of the base questions were more of a statement instead of a question or had an implicit answer in them, such as *I received a bill for cleaning, what I need to do*. These I edited to questions without the answer explicitly in them, as *Why have I received a bill for cleaning?* My aim in editing the questions was to produce neutral standard language. I did include the idiom *jonon hännilla* [in the last place of the queue] into a question. I decided to include it in the final question as it was already in the base questions. After finishing these edits, I created a customer journey map for this study.

As mentioned above, customer journey is a management tool which visualizes the customer experience throughout the purchase process (Rosenbaum et al. 2017, n.p.). Customer journey is a visualization of the events during which a customer might interact with a service provider (ibid.). During these events, customers can interact with the business through different touchpoints (Lemon & Verhoef 2016, 69). On the map, the journey is divided into three phases: before services, during services and after services (Rosenbaum et al. 2017, n.p.). The journey map lists all the touchpoints through which the customer can interact with the service provider at different phases (ibid.). Thus, the customer journey map is usually company specific.

Because complex services might mean several different touchpoints, a traditional customer journey map can be difficult to follow (Rosenbaum et al. 2017, n.p.). One way of diminishing the confusion is to create a customer journey with market research in which it is evaluated how important different touchpoints are to the customers (ibid). To use customer journey as a strategic tool, the company can combine other strategic information such as marketing or customer emotions along the journey (ibid). The risk is that it becomes too complex and does not reflect the actual customer experience anymore (ibid).

The customer journey I created for this study has a narrower scope than a regular customer journey map, as for the purposes of this study other touchpoints excluding the chatbot are not relevant. Therefore, the customer journey map in this study focuses on the phases during

which the user can interact with the chatbot. Also, this customer journey map is designed to suit all of the companies and their chatbots, instead of being company specific. Creating company specific customer journey maps in this study would mean six different maps. Managing and following these maps at the same time could become confusing, especially when the resulting chatbot answers are compared. The maps would also need more company specific information, such as marketing information, which I am not able attain for this study. Therefore, I have focused on the common phases between the student housing associations. Thus, the focus of this journey are the commonalities in the services and the interaction with the user. As the journey map in this study is in these regards different from a usual customer journey map, I will refer to it as simplified customer journey.

While I collected the base questions, I noticed that the questions in the FAQs and the information pages were already divided into sections based on different criteria. The common divisions were based on the user groups – such as the applying user and the user who already lives in the apartment – and on the services, such as keys and appliances in the apartment and the web services which the housing association provides. The divisions based on the user groups was the foundation for the customer journey map of this study. These user groups were the applying user, the user who lives in the apartment and the user who is moving out. These user groups are similar to the three phases in the customer journey if they are compared to the states during which the user interacts with the housing association. Therefore, the phases in the simplified customer journey are the applying for an apartment (before services), living in the apartment (during services) and moving out (after services).

After determining the phases of the customer journey map, I divided the questions between these phases. I used the divisions which the housing associations had used on their websites. These phases could have some overlap, which is evident in the questions. For instance, there is a transition from the applying user to a user living in the apartment during which the user might have questions related to this transition, such as signing a contract on the apartment or how to receive the keys for the apartment. These types of questions were placed into the living in the apartment phase as that was the division already on the websites of the housing associations. It is also logical in a sense that the difference between these phases is whether the user is actively using the services of the association or applying for these services. However, the questions are in an order that user could have them, meaning that the transitional questions between the phases are first in the phase. Figure 2 visualizes the

simplified customer journey. The green circles represent the phases in the journey, to be read from left to right.

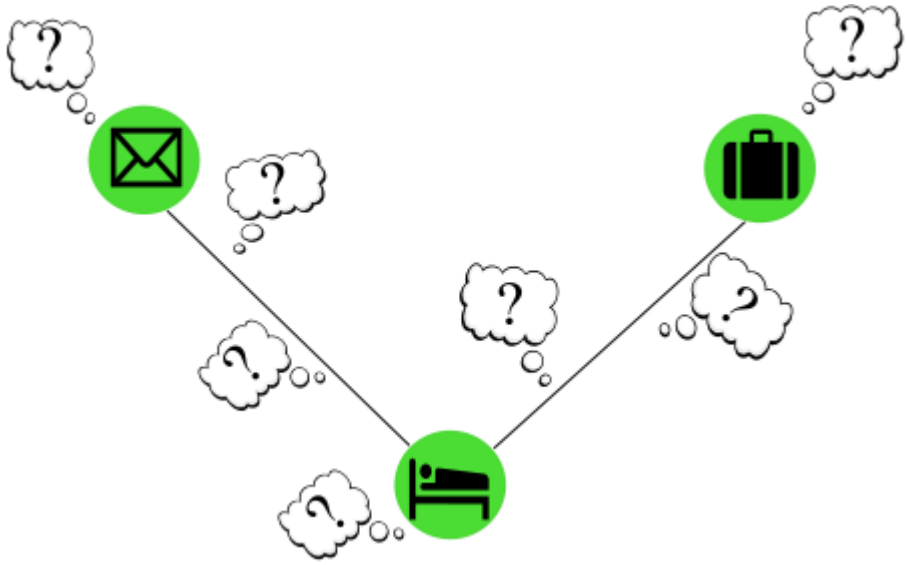


Figure 2. The simplified customer journey in this study.

In Figure 2, the question marks represent the problems and the questions based on these problems that the user has before, during and after the phases in the journey. In the first phase, the questions are related to the application process and generally to the apartment. In the second phase, the questions are more specific and most of them concern the apartment. During the third phase, the questions relate to the process of moving out and different expenses. All in all, there are 14 questions in the journey. The first phase has four questions and the other two phases have five questions each. The order of these questions responds to the possible needs of the user. Therefore, questions such as changing the batteries of the smoke alarm are after the question how to receive keys to the apartment, as the user needs the keys first to live in the apartment and the question of the smoke alarm probably arises after living in the apartment for a while.

After all of this, I piloted the questions with a similar chatbot as the ones in this study. The chatbot is provided by a housing association POAS, which provides housing for students and working people below the age of 30 (Poas). The UI of their chatbot *Onni* is shown in Picture 6. *Onni* uses the same platform *Giosg* as other chatbots of the study. Picture 6 shows how the *Giosg* logo is visible at the bottom of the chat window. The pilot testing was done on 21st of January 2020. During the pilot I discovered three things: first, the chatbot worked properly only with the Chrome browser; second, some questions were overly formal for a chat format,

and third, the chatbot could not answer several questions noticeably because these questions were complex. Due to these findings, I made changes to the test situation.

Picture 6. A screenshot of Onni *chatbot used in the pilot testing.*

First, I decided to only use Chrome browser to minimalize any technological issues. Second, I edited a few questions to more concise structures. For instance, instead of *I paid the rent a few days late* --, I edited the phrasing to *I paid the rent late* --. Third, I prepared a second version of the questions that the chatbot could not answer due to their complexity. For instance, the question *Can I install a washing machine and a dishwasher to my apartment?* is long and includes two different machines, therefore it most likely was a difficult for the chatbot. At first, I had planned not asking the same question in another form, because the non-response could be evaluated as well. However, the pilot showed I would not necessarily receive many responses to evaluate if I did not rephase any of the questions. Therefore, I decided to add a second version to four questions that I could ask if the chatbot could not answer the question. These were the most difficult questions due to their phrasing, not content. They include either multiple questions in the one question, for instance in the washing machine example, or the chosen terms were misleading to the chatbot like the term *muuttoilmoitus* [notification of change of address] instead of *irtisanomisilmoitus* [notice of termination]. For instance, to the previous example question I prepared a second version *Can I install a washing machine to my apartment?* I did not include a second option for the other questions, as in their case the complexity was in the subject of the question more than the phrasing of the question. Table 2 below lists the final set of questions after these edits.

Table 2. The final set of questions for the chatbots.

Phase in the customer journey	**Question in Finnish**
Applying	Voinko hakea asuntoa?
Applying	Miten pitkään yksiön saamisessa menee?
Applying	Milloin saan asunnon?
Applying	Jos lisään hakemukseeni kohteen, joudunko jonon hännille? / Jos lisään hakemukseeni kohteen, joudunko jonon viimeiseksi?
Living in the apartment	Mistä saan asunnon avaimet?

Living in the apartment	Voinko asentaa asuntooni pyykinpesukoneen ja astianpesukoneen? / Voinko asentaa asuntooni pyykinpesukoneen?
Living in the apartment	Pitääkö minun vaihtaa palohälyttimen patterit itse?
Living in the apartment	Asunnossani on tosi kylmä, mitä teen?
Living in the apartment	Maksoin vuokran vähän myöhässä, mitä nyt tapahtuu?
Moving out	Mikä on irtisanomisaika?
Moving out	Miten teen muuttoilmoituksen? / Miten teen irtisanomisilmoituksen?
Moving out	Koska saan vakuuteni takaisin?
Moving out	Miksi en saanut koko vakuutta takaisin?
Moving out	Miksi sain laskun siivouksesta? / Miksi olen saanut laskun asunnon siivoamisesta?

As I interviewed all the chatbots, I screen recorded the process. The initial plan was to interview all the chatbots on the same day, on 25th of January 2020. However, while I started these interviews, the chatbot from Toas, *Tane*, was not in operation. Instead of the chatbot there was a contact request form. Because other chatbots were operational, I interviewed them, and then interviewed *Tane* on 4th of February 2020 when it was operational.

With the screen recordings, I created a simple spreadsheet that showed whether the chatbot could answer the question or not. This also included whether the chatbot could answer the second phrasing of the question. This way, I could easily compare the rate of the success in answering the questions.

4.3 Cognitive walkthrough

For analysing the results, I will use cognitive walkthrough. Cognitive walkthrough focuses on how easy it is to learn to use the product (Wilson 2014, 66). Cognitive walkthrough does not require a real user, because the method is based on the hypothetical user (ibid.). In the walkthrough, the evaluator is supposed to mimic the actions of the user during a task (Ranne 2005, 125). The evaluator answers four questions during each action in the task to evaluate usability (ibid.). The four questions are: does the user have the right aim regarding the UI, does the user notice that the right functionality is available, does the user recognize the right functionality as part of their aim and if the user chooses the right functionality, do they receive enough feedback that their aim is now closer (ibid. 132). These questions are not

directly asked in this study, because the walkthrough is slightly modified to suit the purposes of this study. These questions would better suit a study that includes a larger task to perform. The tasks in this study is asking the questions, therefore the above questions are too broad for this study.

There are several different versions of cognitive walkthrough, for instance streamlined cognitive walkthrough, informal cognitive walkthrough (Wilson 2014, 66) and cognitive jogthrough (Rowley & Rhoades 1992, 389). Cognitive jogthrough is less formal in structure and faster to perform than a cognitive walkthrough (Rowley et al. 1992, 389). It is often used with web services and applications, and the process is recorded (ibid). Cognitive walkthrough can also be adjusted depending on the aim of the evaluation. For example, this was done by Gray, Yilmaz, Daly, Seifert & Gonzalez (2015, n.p) when they adapted the walkthrough into empathetic walkthrough to inspire their students to design systems better suited for the real user. Empathetic walkthrough was developed to consider the use context in evaluating how the user interacts with a system (ibid.).

Cognitive walkthrough is especially suitable for assessing learnability (Wilson 2014, 68). It is also suitable for evaluating smaller programs (Ranne 2005, 125) and systems that are intended to be used without documentation or other help (ibid. 133). However, it has been criticized for focusing on too few tasks, as it is laborious to do more (Wilson 2014, 68). Another point of criticism has been that many tasks could be resolved in several ways, so choosing one method for the walkthrough can be difficult (ibid., 68–69). The evaluators must also consider how the user would react if the system gave feedback to the user that they have chosen the wrong choice (Ranne 2005, 133). Thus, cognitive walkthrough is not meant to replace usability testing (ibid., 125).

In this study I will use cognitive walkthrough that resembles cognitive jogthrough as there is only one evaluator and the evaluated systems are smaller applications. The evaluated tasks in this study are also shorter and less complicated, which makes the full cognitive walkthrough difficult to implement. Therefore, the less formally structured and lighter evaluation suits the aim of this study. The questions asked from the chatbots in this study are related to real issues or situations that the user might have at home. For example, they might receive a bill they did not expect. It would be ideal to include consideration for the use context into this study. However, due to resource limitations this is not possible, and I will only allude to the use context during the evaluations.

Wilson (2014, 70) provides a comprehensive list of documents to be used in cognitive walkthrough. These documents include the representation of UI, the problem reporting form, the user profile, the task list and the action sequence for each task in the list. I used these as a basis for my own analysis but have adapted parts of it to suit my study. The materials called representation of the UI I have left out, because it is for walkthroughs that are done on prototypes, not finished products. As these chatbots are on the web, they are themselves representations of the UI. Chapter 5, which includes the analysis of the chatbot answers, can be interpreted as the problem reporting form.

The following is the user profile in this study:

- A new, full-time higher education student
- 20–30-years-old
- Has not lived in a student apartment before
- Has intermediate skills in using everyday technology.

As I mentioned in chapter 3.1, novice users usually participate in measuring the learnability of a system. As cognitive walkthrough as a method focuses on learnability, a novice user appears as a logical choice. In this case, the user is novice in a sense that they have not used the services of a student housing association before. The style of the questions in the simplified customer journey imply that the user does not have experience with an association. For example, the journey includes questions about whether they are qualified to apply for a student apartment. This implies they have not done the applying procedure before and are unsure about the qualifications. A new student who has not lived in a student apartment before most likely does not have extensive knowledge regarding the services and apartments of an association. Someone with more experience might already know whether they are qualified or not.

The estimated age of the user is 20–30 as that is the most common age range of the new students in the bachelor's degree level students in Finland (SVT 2012). Gender is not considered relevant because gender does not influence housing need. Furthermore, none of the questions are related to gendered needs, such as a shared apartment only for one gender. As Finland has a high level of access to internet and technologies, and media and digital literacy is taught in schools (Kupiainen 2011, n.p), it can be assumed that the user has intermediate skills in using websites and other technologies such as chatbots.

The task list correlates strongly with the question list. In this study, there is not one bigger procedure such as applying for an apartment that is tested. Therefore, the question list is regarded as the task list. The questions of the simplified customer journey are considered separate questions even though the user could ask some of them at the same time if they wanted. Thus, I have created the following action sequence for asking a question that can be applied to all the individual questions in the customer journey:

1. Open a browser.
2. Go to the website of the association.
3. Open the chat window.
4. Ask the question from the chatbot.
5. Follow the instructions given by the chatbot.
6. Close the chat.

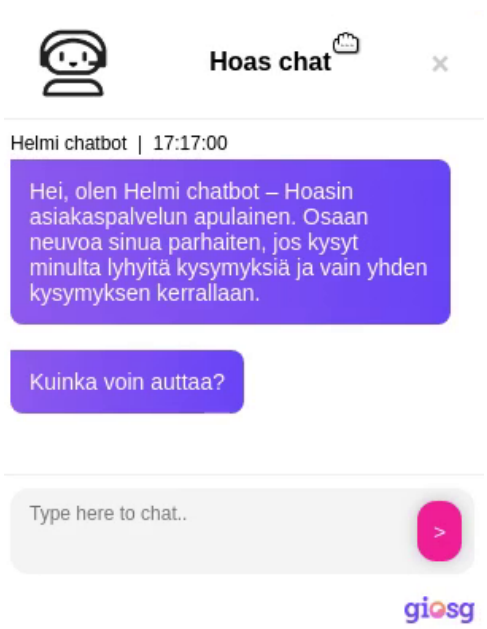
The fifth and sixth step could be in a different order, if the instructions by the chatbot require to contact the customer service or going to the office of the student housing association. This is the order in which the steps are carried out if the user can execute the action immediately or if the question has a more descriptive answer which does not require an action from the user. The fifth step does not include what to do whether the chatbot does not provide instructions. The failure of the chatbot is not part of a successful action sequence or the correct way of solving the issue, which is the idea behind the action sequence. The ideal situation is that the user receives instructions and can follow them.

5 EVALUATION OF THE ANSWERS

In this chapter I will analyse the usability of the collected answers. At first, I will walkthrough each chatbot's answers separately. I use the action sequence presented in the previous chapter. The first three actions in the action sequence are common between all the questions that represent the task list. Therefore, I describe them at the beginning, but not between every question. After that, I continue to compare the chatbots to each other.

5.1. Hoas

I opened the browser and searched for Hoas. From the right side of the screen I opened the chat window which was labelled as chat. Picture 7 shows the chat window as it opened. The *Helmi* chatbot tells the user that it is a chatbot and the assistant of customer service. The welcome message also includes directions that *Helmi* can best answer short questions which are asked one at a time. After the welcome message, *Helmi* asks the user how it can help.



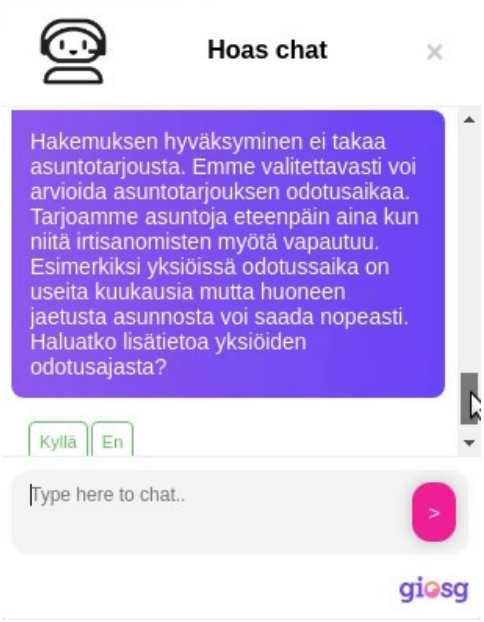
Picture 7. Screenshot of the welcome message sent by the Helmi *chatbot.*

1. Phase: Applying

The first phase in the customer journey is the applying phase. The first question in this phase was about who can apply for an apartment. As I asked the first question, *Helmi* answered with a long message which did not fit the chat window. The answering time was not long, as the answer appeared almost immediately. After the first answer, *Helmi* provided three quick replies. The chat window could not include both the answer and the quick replies at the same time. Therefore, the user has to scroll to be able to read the answer. *Helmi* answered to the question with an explanation that the applicant should study full-time, their studies should aim for a degree, it should qualify them for governmental support and be in the capital region or the surrounding cities. *Helmi* also provided a link where the user can read more about who can apply for an apartment. The three quick replies included questions about when the apartment can be applied, if you are an underage applicant and if you are a postgraduate student. The information about who can apply is in the long reply, so I did not choose any of the quick replies. I did, however, click on the link which *Helmi* provided.

The link opened in the same browser window as the start page where I opened the chat. It took a few seconds, but also the chat window with the ongoing chat opened in the same window. On the webpage which opened, it is mentioned that students outside of capital region can apply for an apartment if they are doing an internship or research required for their degree in the capital region. It also stated that during spring and summer anyone with a temporary job or summer studies in the capital region can apply for an apartment. This information is interesting, but not necessarily something relevant to the user who is applying because they are about to start their primary studies.

The second question in the applying phase was about waiting for a studio apartment. *Helmi* answered with an even longer text to this second question. In the answer, *Helmi* explained that single apartments are offered based on the waiting time and they are primarily offered from the same city where the studies are. Then it lists reasons which do not expedite the process, such as health reasons. Then *Helmi* provides estimates of the waiting times in each city which are at minimum six to nine months. It reminds the user to renew their application every three months. Presumably, it means that the application is valid for three months only. However, that it not explicitly stated in the answer.



Picture 8. Screenshot of Helmi's *answer to the question time about the waiting time for an apartment.*

The third question in the phase is when the user can expect an offer of an apartment. Interestingly, in the answer, *Helmi* states that acceptance of the application does not guarantee an apartment offer. Picture 8 shows the entire answer which includes information that the association cannot estimate the waiting time for the apartment offer. Even though the answer is not as long as the previous ones, the user still has to scroll up to see the entirety of the answer. Lastly, *Helmi* asked whether the user wants more information on the waiting time for a studio apartment. The bottom of Picture 8 displays the quick replies *Yes* and *No.* As I had already asked about that, I chose *No. Helmi* answered it by asking whether I wanted to ask something else.

The fourth question regarded whether the user ends up last in the waiting queue if they add an apartment to their application. This question has the Finnish idiom *jonon hännille* [last in the

queue] in it. *Helmi* did not directly state whether the user does end up last in the queue. However, it stated that the user can renew their application or update the information in the application to maintain it as an active application. This could be interpreted as an answer that the user does not end up last in the waiting queue, but instead it is renewed and active for longer if new apartments are added. This is speculative, because *Helmi's* answer does not directly address the question. *Helmi* again provided a link for more information. When I clicked on the link, it redirected me to a page on which I could update, renew or withdraw my application. The page does not include more information about the renewal process or the placement in the queue after renewing the application. This page would have required logging into my application with personal information.

2. Phase: Living in the apartment

First in this phase is the question about where the user can acquire their keys after receiving an apartment from the association. *Helmi* gave a straightforward answer. It is the first answer that fits entirely on the chat window even with the quick replies. In the answer, *Helmi* tells the user that the keys can be acquired from their offices during the opening hours. It also provides the address to the offices and reminds the user to bring valid identification documentation with them.

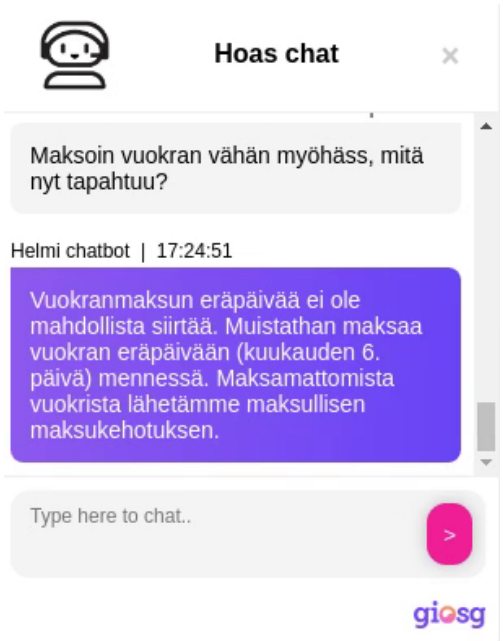
The second question in the phase is about installing a washing machine and a dish washer to the user's apartment. *Helmi's* answer though did not provide relevant information to the question. *Helmi's* answer regarded the common laundry room rules and obeying them. It does not deal with a washing machine in the user's own apartment. The answer explains that housing association gives directions in using the common laundry room and they are sorry if someone has used the time which the user has booked. This probably means the time a user can book for a washing machine. *Helmi* encouraged to reach out to the association if the user thinks the rules are not comprehensive enough. After the answer, there was a quick reply for the contact information of the housing services.

Using the second version of the question, which only addresses installing the washing machine and not the dish washer, did not change the answer. As there is no feedback possibility, the user cannot indicate that the answer is entirely incorrect. Neither is there any link to a page related to washing machines or appliances, thus leaving the user to search for the information on their own.

The third question was whether the user has to change the batteries to the fire alarm by themselves. *Helmi* did not know how to answer the question. It instead instructed the user to rephrase the question and pay attention to the spelling or contact the customer service. It provided in quick replies the options to ask again, send a message or close the chat. I had not prepared another version of this question and thus I did not ask it again and instead moved onto another question.

The fourth question was about the room temperature in the user's apartment being too low. To this question *Helmi* gave a long, in-depth answer. The answer explained when the apartments of the housing association are heated and what the temperature in the apartment should be. *Helmi* mentioned a "heating season" during which the apartments are heated. This season begins as temperatures outside are below 15–17 Celsius. It also explained that the threshold for room temperature during this heating season should be 18–26 Celsius. All in all, the answer included a lot of information, but it was complex. The concepts of a heating season and threshold temperatures are confusing. Especially in a chat format, processing such formal terminology takes time. Besides, the initial answer did not offer any action to be taken to heat the apartment or who to contact about this issue. At the end of the answer, there was a link which I clicked.

The link directed me to an information page explaining that during the autumn users often contact the association regarding the room temperature. According to the information page, sometimes it would be enough to make sure that the radiator is not covered by curtains or something else, but sometimes it would require a professional to check whether it is working properly. Even though this is more information, it still did not reveal who exactly to contact for the professional to check the radiator. Of course, both already include information which can be helpful in a case that the temperature is around threshold temperatures, in other words no action is required. And this of course explains why the apartment might feel cold during the time that it is not heated, outside the mentioned heating season.



Picture 9. Screenshot of Helmi's answer to the question about late rent payment.

The last question in the living in the apartment phase was the question about late rent payment. *Helmi* answered the question correctly, even though I had made a small spelling mistake in the word *myöhässä* [late] as Picture 9 shows. The tone of the answer, however, was strict and does not appear to reflect the situation. *Helmi* stated that the date of payment cannot be rescheduled, the payment should be done on time and in case of unpaid rent the association sends a demand to pay. Picture 9 shows that the answer did not include any information regarding the user's situation – the rent is already paid, but it was paid late. Based on this answer it appears that the association would still send a demand to pay and the user has to wait for it.

3. Phase: Moving out

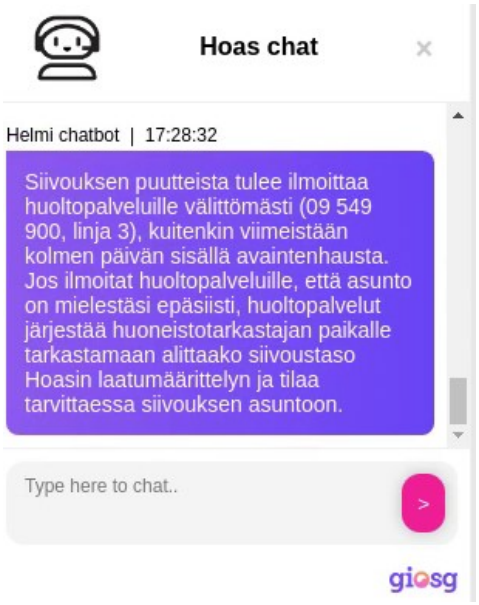
The first question in this phase was about the time of notice before moving out. *Helmi* answered that the time of notice for the resident is one calendar month. *Helmi* gave an example: if the user had given the notice that day on which they asked the question, the contract would have ended at the end of the following month. The wording in the example was slightly complex as it said that *päättyy sopimus ensi kuun loppuun* [the contract ends to the end of the month]. *Helmi* also mentioned the summertime apartments and furnished apartments. These apartments have a fixed term in their contract and the user cannot leave them earlier. This is not really relevant information for the user in this study.

The second question in this phase was about how to notify the housing association about terminating the tenancy. This question had two versions, as was shown in Table 2 in the chapter 4.2. *Helmi* could not answer the question and instead instructed the user to rephrase the question and pay attention to the spelling or contact the customer service. The answer was same to the second version. After the answer there were quick replies which provided options to ask again, send a message to the customer service or close the chat.

The third question in the phase was about when the user receives their deposit back. *Helmi* answered with relevant information, stating that deposits are paid back the following month after moving out. However, the first sentence of the answer was complex in structure and difficult to understand. It explained the date on which the association begins paying back the deposits. The next sentences were clearer and less complicated. *Helmi* also provided two links at the end of its answer. The first one was for the user to leave their bank account information and the second one was for more information about deposits. I clicked on the link for more information. It directed me to an information page about deposits, which had a figure of the schedule for the deposit returns. This made the information easier to process than the explanation by *Helmi.* The figure included the same information, but it was structured more clearly than *Helmi's* answer.

Fourth question was about the reason why the user has not received their entire deposit back. *Helmi* gave a long answer. *Helmi* stated that the deposit is a guarantee to the association in case the user has not paid their rent or other expenses. It also said the association can without consultation keep the deposit as payment for unpaid expenses. The style of the answer is very formal, and I had to read it a few times before understanding it entirely. The explanation was clouded by the complex sentence structure and format compared to the small chat where it is presented.

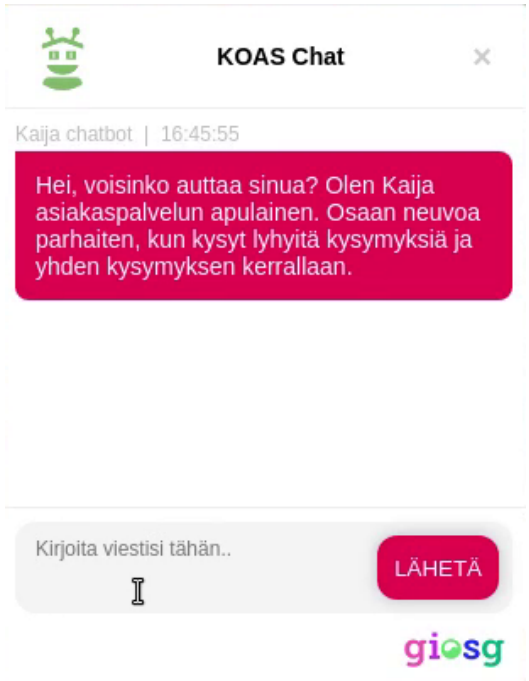
At the end of the answers there was a link and a quick reply. The link was for more information about returning deposits. The quick reply was also called *returning the deposit.* It was unclear whether the information behind the link and quick reply would differ somehow as they appeared to be the same. I clicked on the link, but it just reloaded the page where I already was. At the bottom of the page I noticed information about reasons the deposit is not returned to the user. The page stated that deposit is paid back if the apartment is in proper shape, the keys are returned on time to the right place and there are no unpaid expenses. The page also mentioned that the user has to ensure they have given the correct account information for returning the deposit. In my opinion, this explanation was more comprehensive than *Helmi's* answer. It also does not use the expression of deposit being a guarantee to the association. Instead the phrasing was more direct, thus easier to comprehend in one reading. Also, *Helmi* did not give contact information in case the user wanted to talk about the possible unpaid expenses to someone.



Picture 10. Screenshot of Helmi's answer to the question about receiving a cleaning bill.

The final question was why the user has received a bill for cleaning. This question had two versions. *Helmi* did not respond to either versions correctly. Picture 10 shows how *Helmi* talked about the cleanliness of a new apartment in which the user would have just moved. It encouraged to contact the housing association if the cleanliness of the apartment was not satisfactory. As Picture 10 shows, there is no quick reply or other way to indicate that the answer is incorrect.

Overall, *Helmi* could answer most of the questions, 9 out of 14. However, *Helmi* could not answer 5 out of 14 questions. One of these wrong answers was related to the subject but did not answer the question properly. The tone of the answers was rather formal. At times, *Helmi* directly addressed the user and gave directions. However, in descriptive answers it usually did not address the user directly. I had to learn to construe the formal style of the answers. I also noticed that the linked webpages often included the same information as the answers in a more understandable format. These links were therefore quite useful.



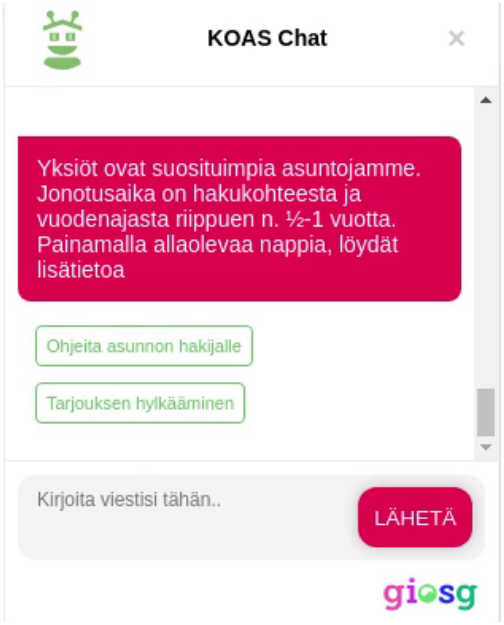
Picture 11. Screenshot of the welcome message sent by Kaija.

5.2 Koas

I opened the browser and searched for Koas. From the right side of the screen I opened the chat window which was labeled as chat. In the Picture 11 is the chat window as it opened. The *Kaija* chatbot asks the user if it can help. It also introduces itself as an assistant to the customer service and includes directions that it can best answer short questions which are asked one at a time.

1. Phase: Applying

To the first question, *Kaija* gave a long answer, which was divided into two parts. In the first part



Picture 12. Screenshot of Kaija's answer to the quick reply.

of the answer, *Kaija* explained that anyone who studies after primary school in the city of Jyväskylä or nearby can apply for an apartment. User can apply as soon as they have been admitted to school. *Kaija* also mentioned that for family apartments only one of the applicants has to be a student. The second half of the answer popped up to the chat with a short delay so that I had a chance to read some of the first half. In the second half *Kaija* explained the association rents apartments with shorter leases for course study students, interns and summer workers.

Kaija answered the second question in multiple parts. These parts also popped up with a small delay to the chat. In the first part *Kaija* explained that demand for apartments fluctuates between seasons and apartments. In the following parts, it explained that the user can apply as soon as they have been admitted to school and that residents have a moth for time of notice. It continued that due to this time of notice, user would receive an apartment offer about a month before the apartment is available for moving in. Then it stated that apartments without any applicants are posted on their website. *Kaija* offered three quick replies with options *available apartments, instructions for applicants* and *applying for a studio.* There was no time estimate or mention that they cannot estimate the time for waiting on a studio offer. Because this information was not mentioned, I clicked on the quick reply *applying for a studio.* I expected general information and maybe a link to their website, but *Kaija* actually provided a time estimate for an offer. This answer is in Picture 12. This answer is more concise and includes the information I had asked about. Picture 12 shows that *Kaija* said the studios are their most popular apartments and the waiting time for one is from half a year to a year. Below the answer are the quick replies for instructions and turning down an offer. It is strange that *Kaija* did not just immediately offer this time estimate because it was what I had asked about.

To the third question, *Kaija* again provided a long answer. It said the apartments are offered as previous tenants move out. *Kaija* also reminded to check the junk mail folder of the email in case the apartment offer has been directed there. It stated the association cannot estimate how long the user has to wait for the offer, and that it cannot give personalized information

regarding the user's application. At the end, *Kaija* provided two quick replies, *progress of application* and *studio situation.* As *Kaija* had already said it cannot give a time estimate, I did not choose any of the quick replies.

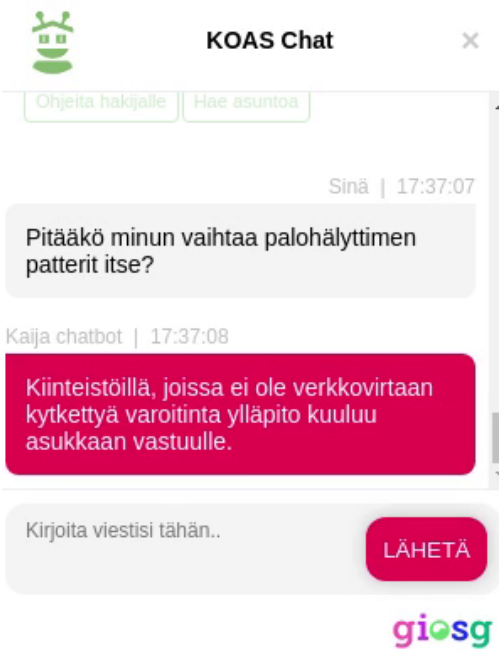
To the fourth question *Kaija* did not know the answer. It answered with the default message. The message started with with *Kaija* saying it cannot unfortunately answer the question. After that *Kaija* said the customer service is closed but the user can leave a contact request via email. It also provided the email address for the customer service. It could not answer the second version of the question either and gave the same default message. After the default message, there were quick replies for continuing with *Kaija,* contacting the customer service or getting back to the question on a later date. I chose continuing with *Kaija.* It answered that it is happy to help and urged to ask another question.

2. Phase: Living in the apartment

Unsurprisingly, *Kaija* had a long, two-part answer to the first question. In the first part, *Kaija* described when the keys can be acquired from the office of the association. However, the answer did not include the address of the office. *Kaija* also said the user can write a commission for someone else to acquire the keys for them. The last sentence in its answer was rather odd. It stated the user can agree with the association for the keys to be delivered to a paid pick-up point. There is no other mention of this pick-up point or where it is located. Thus, this information is an odd addition, because there is no clear reason why this pick-up point would be preferable over the office. In the second part, *Kaija* gave direct instructions to contact the association in case the user wants to move into the apartment earlier. It provided quick replies *instructions for the user moving in* and *the contact information of the association*. The quick replies did not mention the paid pick-up point again. Neither did it include the address of the office.

To the second question, *Kaija* gave the exact same answer as to the first question in applying phase. To the second version of the question, *Kaija* again provided the same answer. There was no option to indicate the answer is not correct or to give feedback on the answer. It was also unclear why *Kaija* answered with an explanation of the applying process to the question about appliances in the apartment.

For a change, the answer to the third question was short, only one sentence. Picture 13 shows the entire answer. *Kaija* states that in a property without a fire alarm connected to the current,



Picture 13. Screenshot of Kaija's answer to the third question.

the maintenance of the fire alarm is the responsibility of the tenant. However, the sentence structure is unnecessarily complicated. It also does not directly address the user, which would make the answer more understandable. As can be seen from Picture 13, there are no quick replies or links for more information. Thus, the user should somehow know whether they have a fire alarm which is connected to the current. Maybe the user can easily know this by looking at the fire alarm in their apartment. However, it could be explained briefly how the user can identify the current connected fire alarm.

Kaija answered in two parts to the fourth question. In the first part it instructed to ensure there are no curtains or furniture in front of the radiator. It also instructed to measure the room temperature because the warmth of the radiator does not necessarily indicate whether it functions properly. In the second part, *Kaija* said the room temperature should be +21 Celsius. If the temperature is constantly below +20, *Kaija* instructed to contact the association. Then it provided a quick reply for contacting the association.

Kaija's answer to the fifth question was calming and friendly. It stated that the user does not have to worry if the rent is late by a few days. *Kaija* told the user can pay their rent in a service portal without fees even if they have received a reminder to pay. If the user does not react to this reminder then they would receive a demand to pay. At the end of the answer, *Kaija* also instructed to contact the controller of the association if needed. I assume this is for cases in which the user has bigger problems with paying rent. However, this is not really explained in the answer clearly. *Kaija* also included quick replies called *paying rent* and *contact information of Koas*.

3. Phase: Moving out

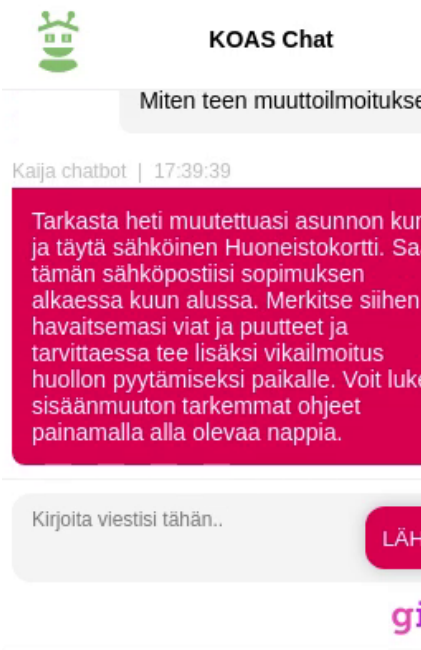
The longest answer thus far was *Kaija's* answer to the first question. It was in four parts. In the first part *Kaija* stated the time of notice is one calendar month. It also said the notice of termination can be done earlier if the user wants. The temporary leases end on the agreed date

and there is no need for a notice. This first part included the most relevant information regarding the question and answered the question quite well.

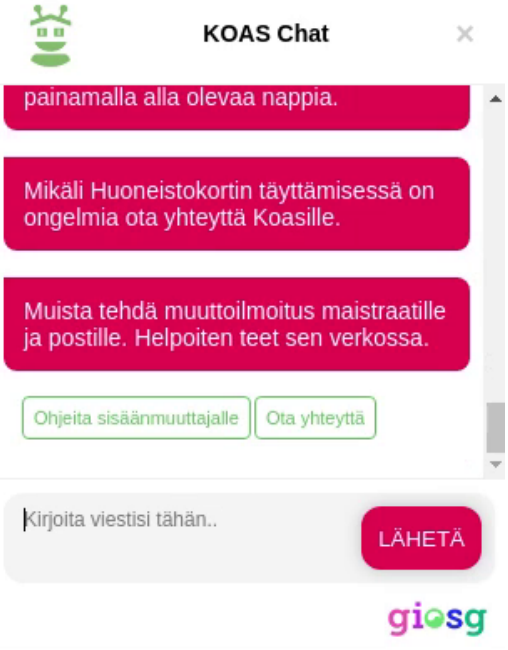
In the second part, *Kaija* said the notice of termination must always be in written form and done by all principle tenants of the apartment. It said the easiest way of leaving the notice is through their website but did not offer a link to the page or the form. Lastly, *Kaija* wished the user's phone number could be given to the person moving into the apartment so that they could agree on seeing the apartment before this person accepts the offer on the apartment. However, *Kaija* did not mention whether this is an option to fill on the notice on their website or if it has to be stated otherwise. This second part was relevant, but it would be more relevant with a link to the notice form. Also, the information about the phone numbers was not that relevant as the question was only what the time of notice is.

The third and fourth part included irrelevant information regarding the question. *Kaija* told the user receives a confirmation email after the notice of termination. If the user does not receive confirmation, the notice of termination has not been received. Then *Kaija* gave the contact email for the association if there are problems with the notice. In the last part it only stated the user should contact the association if the other tenants in the apartment want to stay. These parts could be an answer to any further questions about the notice of termination, but in this case, they appeared excessive. The answer was already long with only two parts. Besides, the information feels disjointed as the timeline in leaving the notice is not clear. Should the user contact the association about other tenants before leaving the notice? Or afterwards?

Most of the information mentioned in the first answer was the answer to the second question. However, *Kaija* did not answer the second question correctly. Its answer was about filing an apartment card, which is filed when a tenant moves into a new apartment. To the second version, *Kaija* gave the same answer. This answer is entirely in Pictures 14 and 15. The answer is again long and in multiple parts. There was no way of indicating that this is not the correct answer.



Picture 14. Screenshot of Kaija's *answer to the second question.*



Picture 15. Screenshot of Kaija's *answer to the second question.*

Kaija answered the third question briefly. *Kaija* only stated that the deposits are usually paid within a month after the lease ends. It also said the bank account for returns has been given in the notice. This answers the question clearly and concisely.

Kaija could not answer the fourth question and instead gave the same default message as in phase one. After the default message the user can choose from quick replies whether to continue chatting with *Kaija* or not. I naturally continued.

In the fifth answer, *Kaija* said the apartment is inspected during moving out and the inspector leaves notes about cleanliness and possible damages. If these notes are not handled, and the inspector has to order cleaners or maintenance workers to the apartment, the tenants are billed. *Kaija* specified that the tenants in the shared apartments are all responsible for the common areas. Also, *Kaija* reminded the user that they are responsible for compensating any damages they have caused by purpose, neglect or carelessness. *Kaija* gave an email contact to ask for more details about the bill. Lastly, it provided a quick reply for compensation price list.

Overall, *Kaija* could answer most of the questions, 10 out of 14. However, *Kaija* could not answer 4 out of 14 questions. The tone of answers was formal at times, but rather friendly. However, many of the answers were long. They did include information related to the

situation, but it made reading the answers cumbersome. Some of the related information was also situational, so that its relevance depends on the user's personal situation. This type of information could be added by offering the user a choice via the quick replies instead of offering it immediately. I learned that for the most part, the relevant information was at the beginning of the answers. Sometimes *Kaija* also did not offer all the relevant information, for instance as it did not include the address of the office.

Picture 16. Screenshot of the welcome message sent by Toivo.

5.3 Tys

I opened the browser and searched for Tys. From the right side of the screen I opened the chat window which was labelled as TYS chat. Picture 16 shows the chat window as it opened. The *Toivo* chatbot introduced itself as a virtual customer service chatbot. It included directions that it can best answer short, clear questions which are asked one at a time. *Toivo* also sais it cannot help with matters related to the user's own application or contract. Lastly, *Toivo* asked how it can help the user.

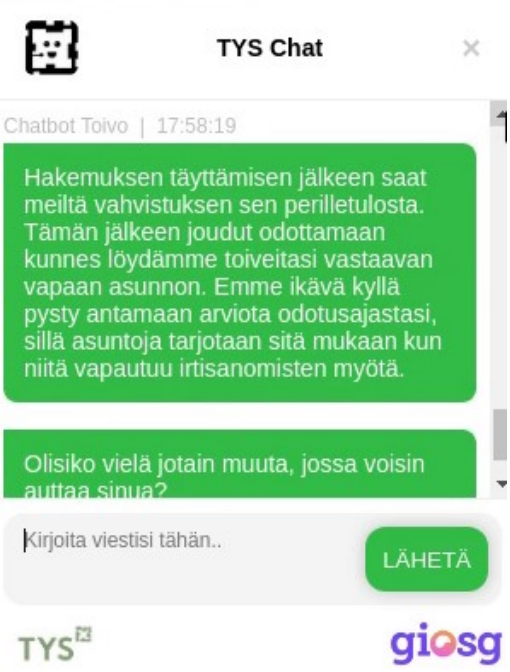
1. Phase: Applying

Toivo could not answer the first question and gave a default message. In it *Toivo* stated it unfortunately does not understand the question. Then *Toivo* personifies itself by telling it has just begun working in the customer service and its orientation is still ongoing. Therefore, it does not know the answer. It explained the customer service is closed, but the user can send an email. *Toivo* provided the email address at the end of the message. *Toivo* could not answer the second question either. It answered with the same default message.

Toivo's answer to the third question was brief and related to the question. According to *Toivo*, the apartments are offered as they become available. Thserefore, the offered apartments are usually apartments which become available in a month. This does not directly answer the question, as *Toivo* did not provide any timeline for waiting an apartment offer or clearly state

it cannot provide one. Thus, the answer to the question remained unclear, even though *Toivo* provided relevant information.

After the answer *Toivo* asked whether it can help with something else. It provided quick replies for *yes* and *no*. I chose *yes*. To the quick reply, *Toivo* answered with encouragement to ask another question and telling it is happy to help.



Picture 17. Screenshot of Toivo's *answer to the fourth question.*

Picture 17 shows *Toivo's* answer to the fourth question in its entirety. *Toivo* said that after sending the application, user receives an email confirmation. Then the user has to wait until an apartment which they want is available. *Toivo* said it cannot provide an estimate how long this takes. This does not exactly answer the question though, as it does not provide any information on whether the application can be modified. As Picture 17 shows, after the answer *Toivo* asks whether it can help with something else.

2. Phase: Living in the apartment

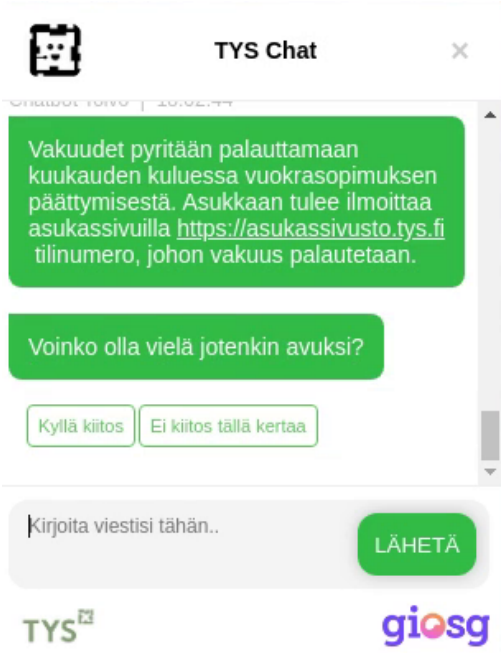
Toivo answered the first question thoroughly in two parts. In the first part, it told the keys can be acquired from the office of the association and provided the address of the office. In the second part, *Toivo* explained that keys can be also acquired from a kiosk outside the office hours. It instructed to contact the association the previous day if the user wanted the keys from the kiosk. The instructions for this were brief and explained clearly.

Toivo could not answer either version of the second question. Instead *Toivo* answered with the same default message as previously in the applying phase. *Toivo's* answer to the third question is in Picture 18. It is very brief and matter of fact. *Toivo* only stated that changing the batteries to the fire alarm is the tenant's responsibility. *Toivo* could have addressed the user directly as that is the clear answer to the question. Picture 18 also shows the quick reply options for either continuing the conversation or ending it.

Picture 18. Screenshot of Toivo's *answer to the third question.*

Toivo's answer to the fourth question was its longest one thus far. *Toivo* instructed the user to first check that nothing is covering the radiator. It also stated that the radiator can feel cold to the touch even if it is working. Thus, measuring the temperature is the best way to determine whether the radiator is working. *Toivo* explained the temperature should be around +22 Celsius. In case the temperature is consistently below +20 degrees, it instructed to notify the association. Then it provided a link to the page where the user can make a fault notification to the association. I clicked on the link, but it was a page which required authentication for login. The chat with *Toivo* loaded and opened on this site after a little while.

To the fifth question, *Toivo* gave a reassuring answer. It said it is no problem if the rent is somewhat late. If the rent is late over 14 days, the user must pay a fee. *Toivo* told the fee is 5 euros, which I think is good information to include so the user knows it beforehand. It also said the penalty interests are billed biannually.



Picture 19. Screenshot of Toivo's *answer to the first question.*

3. Phase: Moving out

Toivo answered the first question briefly. Picture 19 shows this answer. It stated that the time of notice for tenant is one calendar month. This means that the lease ends on the last day of the following month. *Toivo* also provided a link to the page where the notice of termination has to be done. This answer concisely explained everything relevant to the question.

Toivo could not answer either version of the second question. Instead *Toivo* answered with the same

default message as the one mentioned previously in the applying phase.

To the third question, *Toivo* said that the deposits are usually paid within a month from moving out. It also said the tenant has to provide their bank account information for returning the deposit. It provided the link to the page in which this can be done. These instructions are clear. However, *Toivo* does not directly address the user, but instead only refers to a tenant.

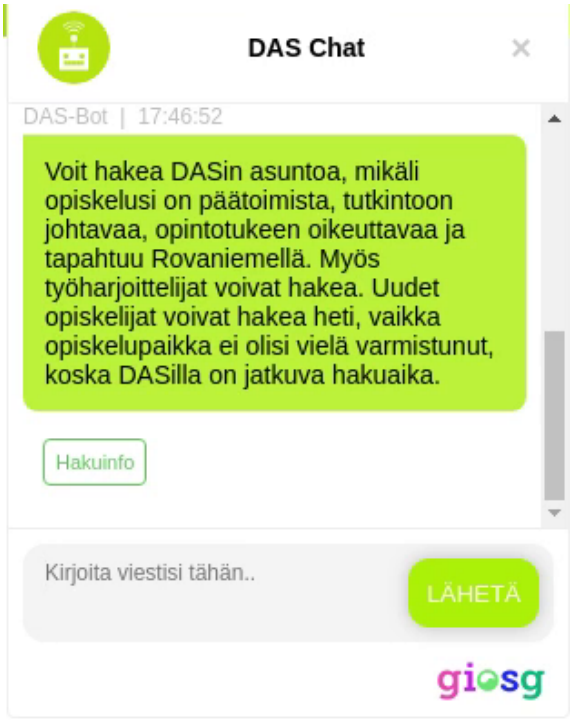
In its answer to the fourth question, *Toivo* explained the deposit is a guarantee to the association in case the tenants do not pay rent, cause damages to the apartment or do not clean it properly while moving out. If the tenant has unpaid payments or has damaged the apartment, the association can withhold the deposit or parts of it for repayment without consulting the tenant. *Toivo's* tone in this answer is very formal. It does not provide any contact information to ask for further details. *Toivo* could not answer either version of the fifth question. Instead *Toivo* answered with the same default message as the one mentioned in the applying phase.

Overall, *Toivo* could answer half of the questions, 7 out of 14. It answered with short and concise answers. Also, it did not offer much additional information with links or quick replies. The links it did provide were directly linked to the task which the user wanted or should do, for instance the link to the notice of termination. *Toivo's* answers were for the most part quick to read and comprehend. They did not really require learning or familiarizing with its style. Sometimes, it could have addressed the user directly so that the instructions would be as clear as possible. *Toivo* asked after each question whether it could help with something else. The wording changed a little bit in these messages in between, but they were quite similar, and a few were repeated during the discussion. In a longer chat like the one we had, it begun to feel a bit unnecessary question to answer all the time. However, this is the type of question which customer service representatives often ask, so it also vaguely created a feeling of conversing with a customer service representative.

5.4 Das

I opened the browser and searched for Das. From the right side of the screen I opened the chat window which was labeled as DAS chat. Picture 20 shows the chat window as it opened. The *DAS bot* first greeted the user with a hello. It introduces itself as the user's virtual assistant. *Das bot* also included directions that it can best answer short questions which are asked one at a time. Lastly, it asked the user how it can help.

Picture 20. Screenshot of the welcome message sent by DAS bot.



Picture 21. Screenshot of Das bot's *answer to the first question.*

1. Phase: Applying

To the first question *Das bot* answered that the user can apply for an apartment if they study for a degree, they are eligible for the governmental support and their studies are in the city of Rovaniemi. It also told the user can apply for an apartment before they are admitted to school. This is shown in Picture 21. Below the answer was a quick reply for applying information which I clicked on. It opened a new window in which the chat with *Das bot* loaded after a while. The page listed the different schools whose students can apply for an apartment. This list did not include further information compared to the *Das bot's* answer.

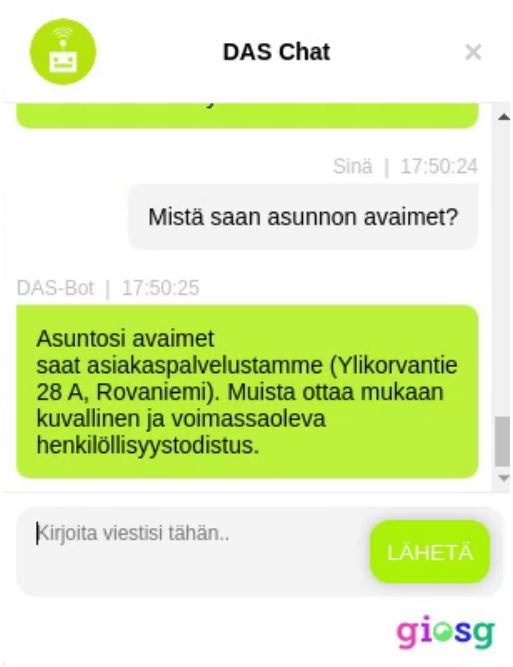
Das bot could not answer the second question correctly. *Das bot* told the user that they can apply for an apartment by filling out the application and provided a quick reply to the application. This was not relevant to the question of how long the waiting time for a studio is. There was no way of indicating to the chatbot that this was not the correct answer.

Das bot's answer to the third question is relatively long. *Das bot* said the apartments are offered as they become available. Thus, the apartment is offered about a month before it is available for moving in. It explained it cannot estimate when the user receives an apartment offer. It said that immediately available apartments can be checked from their website or by contacting the housing association. Below the answer was a quick reply for immediately free

apartments. This was good additional information, which had not been offered by the other chatbots before.

Das bot could not answer either version of the fourth question. Both times it gave the same answer. It explained how user receives a confirmation email after the application is sent and after that the user just must wait. It also mentioned the association cannot estimate the time it takes to receive an apartment offer. There is, however, no mention of altering the application. Nor is there a way to indicate to the chatbot this is not a relevant answer to the question.

2. Phase: Living in the apartment



Picture 22. Screenshot of Das bot's *answer to the first question.*

Das bot answered to the first question briefly and clearly. This answer is shown in Picture 22. It said the keys can be acquired from the customer service of the association and provided the address. Then it reminded the user to bring valid identification documents with them. This includes all relevant information except the opening hours of the office.

Das bot did not know the answer to either version of the second question. Instead, it answered with the default message both times. First *Das bot* said it unfortunately does not know the answer but next time it will. It also suggested checking the message for any spelling mistakes. Then *Das bot* said the user can search for the answer on their website or contact the customer service via email. It offered the contact information as a quick reply at the bottom of the chat.

Das bot's answer to the third question is brief. *Das bot* stated the batteries in the fire alarms are fixed and cannot be changed by the user. If the fire alarm is beeping, the user should notify the association. It provided a link to the webpage in which this can be done.

In an answer to the fourth question, *Das bot* instructed to check nothing is covering the radiator and it is on. It explained the radiator might feel cold to the touch even though it is functional. Therefore, it instructed it is best to measure the room temperature. If the room

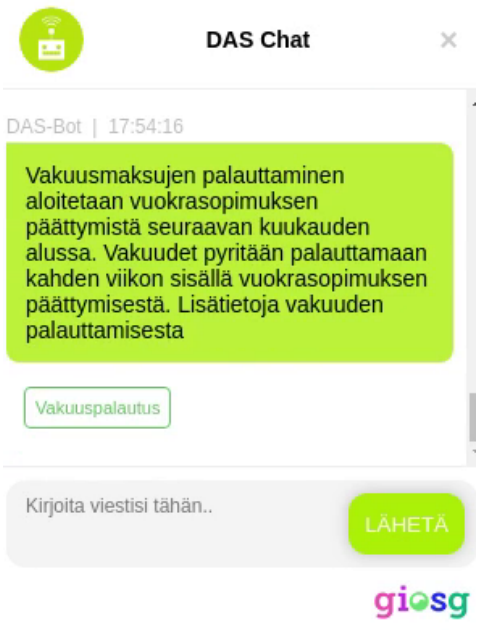
temperature is below +20 Celsius repeatedly, the user should notify the association. It instructed how to measure the temperature correctly. There was also a quick reply for notifying the association.

Das bot's answer to the fifth question is brief and friendly. It said there are no repercussions if the rent was paid a few days late. However, if the rent is not paid by the 20th of the month, the user might receive a fee. Interestingly, the wording in *Das bot's* answer gives the impression it is not certain whether the user receives a fee and has to pay it.

3. Phase: Moving out

To the first question, *Das bot* briefly stated that the time of notice is one calendar month for tenants. It clarified this means the lease will end on the last day of the following month of leaving the notice. This brief answer included everything relevant to the question.

Das bot could not answer the first version of the second question and gave the default message. It did however answer the second version. First, *Das bot* asked whether the user wanted more information about the notice of termination or cancelling the lease, which ends the lease immediately and requires lawful reasons. It provided quick replies for both options. I clicked on the quick reply for the notice of termination. In an answer *Das bot* explained the notice must always be done in writing. If there are issues with the electronic form, the user should contact the association. It also stated that the time of notice is one calendar month. It provided a quick reply for the notice form. I clicked on the quick reply and it opened a page to another window. This page would have required authentication. Eventually *Das bot* did then answer the question. The mention of cancelling the lease was surprising and I did not immediately realize the difference. However, the explanation was good and well placed as this choice was rare and user might not be familiar with the difference.



Picture 23. Screenshot of Das bot's *answer to the third question.*

Das bot's answer to the third question is in Picture 23. It said the association begins paying back the deposits at the beginning of the month. Usually, the deposits are paid within two

weeks from the end of the lease. After the answer, it provided a link for more information about returning the deposits. I clicked on the link, which again opened another window. The page would have required authentication.

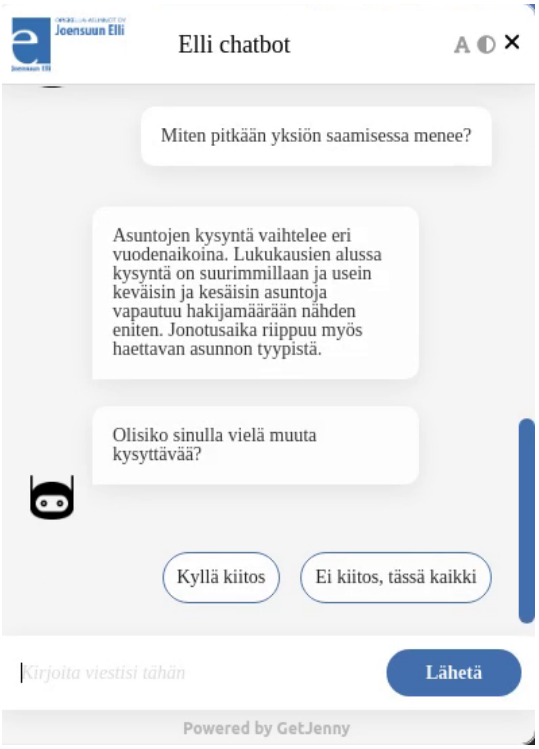
Das bot could not answer either the fourth or the fifth question. Instead it answered with the default message which was mentioned earlier in the living in the apartment phase.

Overall, *Das bot* answered 9 out of 14 questions and could not answer five questions. The style of the answers was instructive and relatively concise, similar to *Toivo's* style. It often addressed the user directly, thus, the answers were easier to comprehend. It also provided more information in the form of quick replies. The quick replies always opened another window, which was frustrating as there were then multiple tabs open with the *Das bot. Das bot* asked after each answer whether it could help with something else. There was some variation to the phrasing of this question, but some forms were already repeated during our discussion.

5.5 Joensuun Elli

I opened the browser and searched for Joensuun Elli. At the bottom of the screen I opened the chat window which was labeled as Elli chatbot. Picture 24 shows the chat window as it opened. The chat window is slightly wider than the chat window of any other chatbots in this study. *Elli* introduced itself as a virtual Ellibot and told it has started as a trainee in the customer service. *Elli* included the directions that it can best answer short questions which are asked one at a time. Lastly, it asked how it can help. From Picture 24 can be seen there is a banner at the top of the chat window. In the banner is a reminder not to write any personal information to the chat. It also includes a link to the privacy statement of the company.

Picture 24. Screenshot of the welcome message sent by Elli.



Picture 25. Screenshot of Elli's *answer to the second question.*

1. Phase: Applying

While *Elli's* answer was loading, the screen showed three grey dots as if the bot was actively writing. *Elli's* answer to the first question was concise. *Elli* said the user can apply for an apartment if they are studying for a degree, they are eligible for the governmental support and the studies are near the city of Joensuu. It notified the user that the students of the open university cannot apply for the student apartments.

Elli's answer to the second question is in Picture 25. *Elli* said the demand for apartments fluctuates between the seasons and the demand is greatest during the fall. During the spring and the summer more apartments are released compared to the number of applicants. As Picture 25 shows, it also said the waiting time for an offer depends on the type of the apartment. Even though the answer is related to the waiting for an apartment, it does not really answer to the waiting time for a studio.

To the third question, *Elli* gave a more relevant answer. *Elli* said the apartment offers are done as apartments are released. It also said the association cannot estimate the time it takes to receive an apartment offer or the waiting time in the apartment queue.

Elli did not know the answer to either versions of the fourth question. To the first version, *Elli* answered that the association will ask for a reason if the user declines an apartment offer so that they can serve the use better in the future. This does not relate to the question at all. To the second version, *Elli* answered that the user can

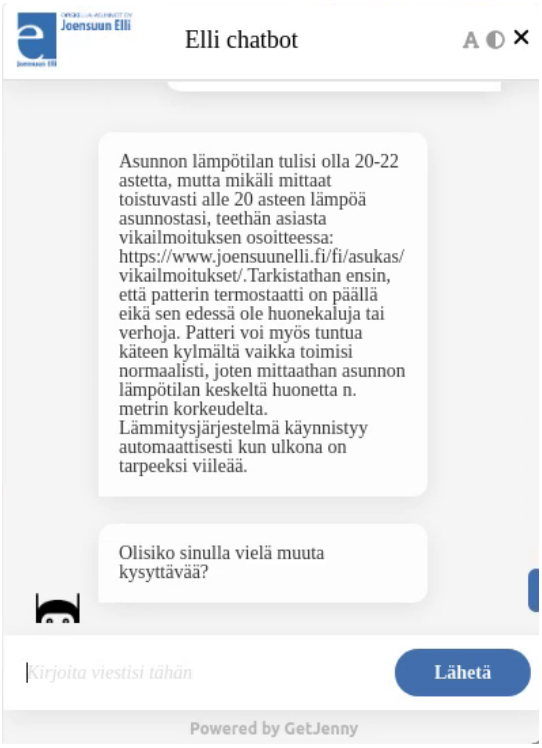
apply for an apartment as soon as they have been admitted to school. Neither of the answers were related to altering the application. There was no way of indicating to the chatbot that the answer is not correct.

2. Phase: Living in the apartment

Elli's answer to the first question was brief. *Elli* told the keys can be acquired from the office of the association and provided the address of the office. It also reminded the user to bring valid identification documentation with them. The answer did not include the office hours but otherwise included everything relevant to the user.

Elli was the first chatbot to answer the first version of the second question. *Elli* said that in case the apartment has the needed connections for the installations, the user can install the dish washer and the washing machine to the apartment. For installing the dishwasher, *Elli* recommended using a professional. It also told that these installations can be bought from the association as a service. It reminded the user about having a home insurance and that the machines need to be removed whilst moving out of the apartment.

To the third question, *Elli* said that the user must change the batteries if the fire alarm is installed by a tenant. If the user is not sure whether the fire alarm is installed by a tenant or not, *Elli* instructed to contact the association. These instructions are clear, although it seems odd that the user has to know whether the fire alarm is installed by a tenant or not.



Picture 26. Screenshot of Elli's *answer to the fourth question.*

Elli's answer to the fourth question is in Picture 26. *Elli* said the room temperature should be 20–22 Celsius, but if it is constantly below 20 degrees, the user should notify the association. Then it provided a link to the website where the user can notify the association. However, the link was not clickable and instead it had to be manually copied from the answer. That is rather difficult from the chat with relatively small text. Picture 26 shows that after the link, *Elli* told the user to first check nothing is covering the radiator and to measure the room temperature. Thus, the answer does not follow the order in which the user should perform these actions. They should first check the radiator and measure the temperature, and only after that notify the association.

Elli's answer to the fifth question was friendly. *Elli* said there are no consequences if the rent is paid a few days late. If the rent is paid after the reminder, the fee is 5 Euros. It is convenient for the user that *Elli* told the sum of the fee beforehand, so the user knows what to expect. The sentence structure in this second sentence was odd as it stated -- *jos maksusi tulee vasta maksumuistutuskirjeiden tulostamista joudut maksamaan* -- [-- if your payment arrives the printed reminder you have to pay--]. The sentence is still understandable: the user has to pay the fee, if the reminder has already arrived. It is confusing though.

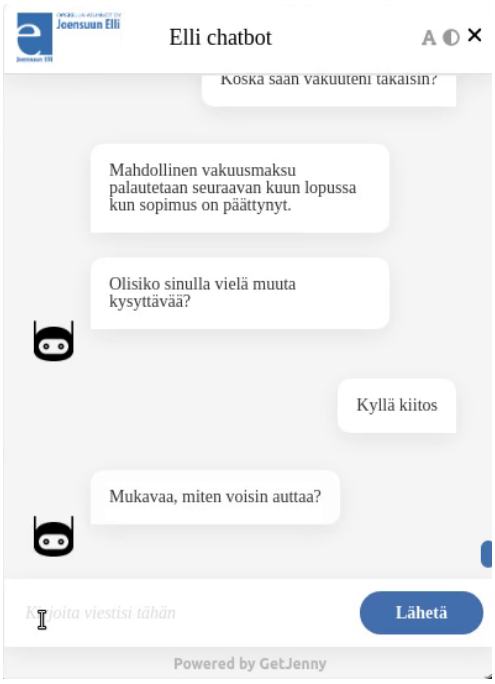
3. Phase: Moving out

To the first question, *Elli* explained that the time of notice is one calendar month for a tenant and the association should always be notified in writing. It mentioned the notice form is available on their website but did not provide a link or a quick reply to this form.

Elli answered the first version of the second question with the default message. It assured that the next time it will know the answer to the question. To the second version it answered with relevant information. *Elli* explained that the notice of termination should be done in writing. Again, it mentioned the notice can be done on their website but does not offer a link or other

directions to find it. It only mentioned that the notice can be printed from their website if there are problems with the electronic form. It also stated that it is the tenant's responsibility to ensure the association receives the notice of termination.

Elli's answer to the third question was brief. Picture 27 shows *Elli's* answer. *Elli* only stated that the possible deposits are paid back at the end of the month after the lease has ended. Picture 27 also shows how *Elli* asked after each question whether it can help with something else.



Picture 27. Screenshot of Elli's *answer to the third question.*

Elli answered with the default message to the fourth question. *Elli* told that the question is registered and the next time it will know the answer to the question.

Elli's answered to the fifth question that the user must pay for any damages caused to the apartment on purpose, due to neglect or irresponsibility. *Elli* said the user can ask for more information from the maintenance and provided an email address.

Overall, *Elli* could answer 11 out of 14 questions and could not answer 3 questions. The tone of the answers is relatively informal at least compared to some other chatbots in this study, like *Helmi*. Often, *Elli* provided clear instructions but did not provide additional resources. For example, it did not provide a link to the notice of termination even though it mentioned it can be done on their website. The only link it provided was not clickable and had to be copied from the answer. The same applied to the contact information. That is a rather unusable solution for providing links. It also felt like *Elli's* answers loaded a little slower than other chatbots' in the study, but I did not time them, so I cannot say for sure. The small spelling mistake in one of the answers did not render the answer incomprehensible. However, the mistake is distracting and could be clearly noticed from the answer.

5.6 Toas

I opened the browser and searched for Toas. From the right side of the screen I opened the chat window which was labeled as chat. Picture 28 is the chat window as it opened. *Tane* greeted the user with a colloquial hello [*moro*] and introduces itself as a chatbot from the city of Tampere. It said it does not know the user, thus, they should not ask personal questions from it. It told the user that they can look at the frequently asked questions or the news on their website. It also provided the quick replies for these options.

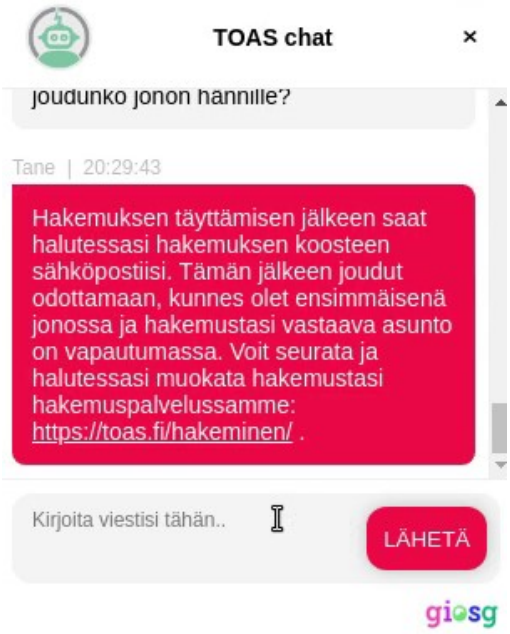
Picture 28. Screenshot of the welcome message sent by Tane.

1. Phase: Applying

Tane answered to the first question briefly that the user can apply for an apartment if they are studying for a degree, their studies entitle them to the governmental support and are in the city of Tampere or in the nearby towns. This answer included everything relevant to the question. *Tane* also directly addressed the user in its answer. After the question *Tane* asked if it can help with something else. The style of the question was informal and friendly. It also provided quick replies for *yes please* and *no thank you.*

Tane could not answer the second question. *Tane* provided a link for the different apartment types which the association has. There was no way to indicate to the chatbot that this is not the correct answer.

Tane's answer to the third question was long. *Tane* told that the apartments are offered as they become available. It stated it cannot estimate the waiting time for sure and mentioned that the



Picture 29. Screenshot of Toivo's *answer to the fourth question.*

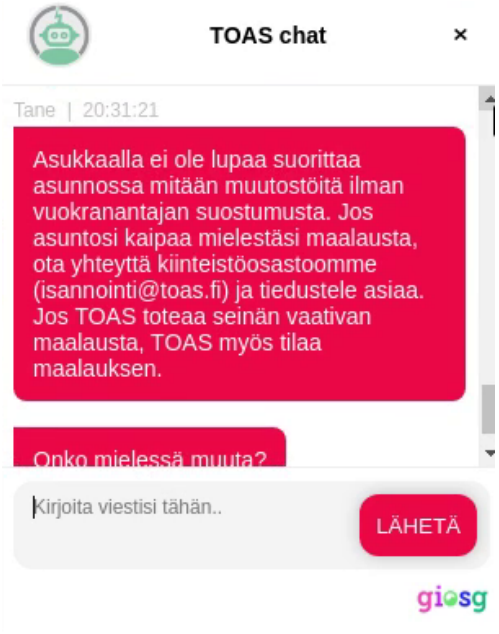
waiting times for the studios are between half a year and over a year. It explained the waiting time also depends on the season and popularity of the apartments. According to *Tane,* there are usually quite many shared apartments available and the family apartments might even be immediately available. This breakdown between apartment types was good information and something other chatbots had not done.

Tane could not answer either version of the fourth question. Picture 29 shows its answer to both versions. *Tane* told the user receives an email confirmation of their application. After that the user can just wait until they are offered an apartment. *Tane* mentioned the user can modify their application but it did not explain whether this affects the placement in the queue.

2. Phase: Living in the apartment

The first question *Tane* answered briefly and clearly. *Tane* told the keys can be acquired from the office of the association and provided the address of the office. It reminded the user to bring valid identification documentation with them.

Tane could not answer either version of the second question. Both times it gave the same answer which is in Picture 30. *Tane* said the tenant has no permission to change anything in the apartment without the permission of the association. *Tane* said the user should contact the association if they think the walls in the apartment need painting. This does not answer the question and there is no way to indicate that to the chatbot.



Picture 30. Screenshot of Tane's *answer to the second question.*

Tane answered to the third question that the tenant needs to ensure the fire alarm has working batteries. However, the user should contact the association if the fire alarm is connected to the current. Interestingly, *Tane* talked about a tenant in the first sentence and switched in the following sentences to addressing the user directly. It also did not explain how the user can know whether they have the current connected fire alarm.

Tane's answer to the fourth question was long. *Tane* instructed to first check the radiator is on and nothing is covering it. Then it instructed to measure the room temperature. If the room temperature is constantly under +20 degrees, the user should notify the association. These instructions are clear in their order of execution.

Tane's answer to the fifth question was friendly and reassuring. *Tane* explained there are no consequences if the rent is paid a few days late. If the rent is paid after the 15th day of the month, the user receives a fee. *Tane* also provided the contact information for a situation in which the collection agency is already handling the user's unpaid rents.

3. Phase: Moving out

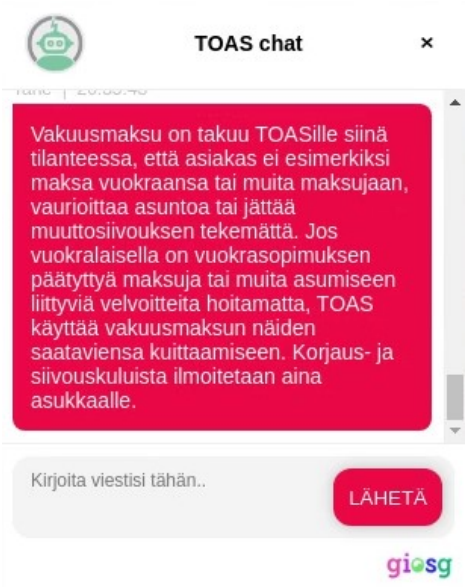
To the first question *Tane* explained that the time of notice for the tenant is one calendar month. It also explained the time of notice begins at the end of the same month the notice of termination is delivered to the association. *Tane* also provided an example and told the fixed term leases end as agreed. This was a comprehensive answer to the question.

Tane answered the first version of the second question with the default message. The default message was very brief, only stating that *Tane* unfortunately does not know the answer. It did not include any instructions for contacting the customer service. However, *Tane* answered the second version of the second question with a relevant answer. *Tane* said the notice of termination must always be done in writing and it is easiest to do in the electronic portal of the association. It also said the notice of termination can be printed from their website and

offered a link to the form. I clicked on the link and it opened the page in the same window. The chat with *Tane* loaded to the page quite quickly. The form was available on the page.

Tane answered the third question in detail. First *Tane* says the deposits are paid back within the following month after the lease has ended. Second it told that returning the deposit might take longer if a part of the deposit is withheld. The user is always informed of the redactions. Third it instructed to contact the association if the deposit has not been paid back, the lease ended over a month ago and no redactions have been informed to the user.

Tane answered in a long manner to the fourth question as shown by Picture 31. *Tane* explained the deposit is a guarantee to the association in case the tenant does not pay their rent or other expenses. Thus, the association can withhold the deposit or parts of it to cover the unpaid payments. *Tane* said the association always informs the tenant of any repair or cleaning costs. In this answer, *Tane* does not directly address the user, but instead only talks about a tenant.



Picture 31. Screenshot of Tane's *answer to the fourth question.*

To the first version of the fifth question, *Tane* gave the default message which was mentioned earlier. *Tane* answered the second version with an explanation that the user should apply for an apartment as soon as they have been admitted to school. This was not a relevant answer to the question. There was no way of indicating to the chatbot that it is not the correct answer.

Overall, *Tane* answered 10 out of 14 questions and did not answer four questions. *Tane's* answers shifted between informal and formal. It addressed the user directly at times, especially in the more instructional answers. However, in some answers it switched between addressing styles which is disorienting. It could have always directly addressed the user. The instructions it provided were relatively concise and logical in order. The more descriptive answers could have been more concise. *Tane* also asked after each answer whether it could help with something else.

5.7 Comparing the answers

In this subchapter I will compare the answers to each other. I will also compare them in each phase separately for easier comparison.

1. *Phase: Applying*

Table 3 lists whether the chatbots could answer (Y) or could not (N) answers the questions in the first phase. The one chatbot which stands out is *Toivo* because it could not answer any of the questions in the first phase.

Table 3. Chatbot answers in the first phase.

Question in Finnish	Helmi	Kaija	Toivo	Das bot	Elli	Tane
Voinko hakea asuntoa?	Y	Y	N	Y	Y	Y
Miten pitkään yksiön saamisessa menee?	Y	Y	N	N	N	N
Milloin saan asunnon?	Y	Y	N	Y	Y	Y
Jos lisään hakemukseeni kohteen, joudunko jonon hännille? / Jos lisään hakemukseeni kohteen, joudunko jonon viimeiseksi?	N	N	N	N	N	N

As Table 3 shows, rest of the chatbots answered consistently to 2–3 questions. The question which no chatbot could answer was about modifying the application and its affects to the placement in the waiting queue. The other difficult question was the second question about the waiting time for a studio. Many of the chatbots mentioned in their answers that the waiting times cannot be estimated. Thus, it is surprising they did not answer so to the second question.

In general, the answers in this phase included a lot of information. For example, the chatbots included information about the different waiting times for different apartment types and explained that the offered apartments are usually available a month from the offer. For a novice user this is helpful as they receive information that they might not realise to ask. The style of answers resembles declarative information, which was discussed in chapter 2.3, as they stated the facts and background information but did not often direct the user to perform actions. There were answers in which the chatbots directly addressed the user, but it could have been done more consistently.

2. *Phase: Living in the apartment*

Table 4 lists whether the chatbots could (Y) or could not (N) answer the questions in the second phase. In this phase the chatbots answered very similarly, most of them answered to four questions.

Table 4. Chatbot answers in the second phase.

Question in Finnish	Helmi	Kaija	Toivo	Das bot	Elli	Tane
Mistä saan asunnon avaimet?	Y	Y	Y	Y	Y	Y
Voinko asentaa asuntooni pyykinpesukoneen ja astianpesukoneen? / Voinko asentaa asuntooni pyykinpesukoneen?	N	N	N	N	Y	N
Pitääkö minun vaihtaa palohälyttimen patterit itse?	N	Y	Y	Y	Y	Y
Asunnossani on tosi kylmä, mitä teen?	Y	Y	Y	Y	Y	Y
Maksoin vuokran vähän myöhässä, mitä nyt tapahtuu?	Y	Y	Y	Y	Y	Y

As Table 4 shows the only chatbot which answered all of the questions was *Elli.* This was the phase in which the chatbots performed best in regards of providing answers, meaning they provided most correct answers in this phase. The difficult question in this phase was the second question about installing appliances in the apartment. No one, except *Elli,* could answer either version of this question.

Answers in the second phase included more procedural information than in the previous one. However, the style of the answers was always not ideal considering the information type. For example, all the answers did not follow the order in which the task should be executed. Some of the answers begun by addressing the user directly but switched to talking about a tenant or did not directly address the user. In chapter 3.4, I discussed the usability heuristics for documentation out of which the fifth one included that instructions should be in imperative and the user addressed directly. A good example are the answers to the question about changing the batteries to the fire alarm. A clear answer in accordance with the usability heuristics would state directly whether it is the user's responsibility or not. Most of the chatbots answered this question but referred to the responsibility of a third person, a tenant. Also, the answers to the question about cold apartment were relatively confusing. The answers switched between explaining the usual room temperature and how to measure the room temperature correctly. The expected actions from the user could have been clearly

separated from the description about the usual room temperature. That way those answers would have matched with the real world better, which was the first of the usability heuristics in chapter 3.4.

3. *Phase: Moving out*

Table 5 lists whether the chatbots could (Y) or could not (N) answer the questions in the third phase. If the chatbot could answer the second version of the question, it is marked with an asterisk (*). This phase had most dispersion between the answers. The chatbots answered 3–4 questions each.

Table 5. Chatbot answers in the third phase.

Question in Finnish	**Helmi**	**Kaija**	**Toivo**	**Das bot**	**Elli**	**Tane**
Mikä on irtisanomisaika?	Y	Y	Y	Y	Y	Y
Miten teen muuttoilmoituksen? / Miten teen irtisanomisilmoituksen?	N	N	N	Y*	Y*	Y*
Koska saan vakuuteni takaisin?	Y	Y	Y	Y	Y	Y
Miksi en saanut koko vakuutta takaisin?	Y	N	Y	N	N	Y
Miksi sain laskun siivouksesta? / Miksi olen saanut laskun asunnon siivoamisesta?	N	Y	N	N	Y	N
* The chatbot answered to the second version of the question.						

As Table 5 shows, there is dispersion between the difficult questions. The most difficult question was the last question about the cleaning bill. Other hard questions were about the notice of termination and not receiving the entire deposit back.

Answers to the three lasts questions were often most formal in their style. The subjects of these questions are serious; therefore, the formal style was somewhat understandable. However, the formal style often translated into complicated structures which made reading the answers difficult. The answers were also quite long. Even the chatbots which usually gave shorter answers, such as *Toivo*, in this phase gave longer answers. When the answers were both long and complicated in structure, it required time to read and comprehend them. It was also common that in this phase the chatbots did not directly address the user. That could have clarified the otherwise complicated answers. Effective information design, which was one of the usability heuristics discussed in chapter 3.4, could have prevented some of these issues.

Now the answers were not always effectively structured and designed, so the user had to search for the relevant pieces of information.

Throughout all the phases, the answers had a lot of similarities in their style, content and even the level of formality. All the chatbots required learning in a sense that they had a style of informing which was not the most direct. For instance, the chatbots in their instructions referred to a tenant and a tenant's responsibilities. In these situations, the user could have been directly addressed to make the instructions clearer. It would also have made the conversation seem more natural because all the questions were phrased in the first person instead of the third person. Now the answers resembled more declarative information type, which was discussed in chapter 2.3. Thus, they were not really directed at the user. Due to this style, I had to read some of the answers multiple times to ensure my understanding. After some time, the style appeared more natural and I did not have to exert as much attention to reading comprehension. However, a user might not notice these similarities in answers as they likely would only ask questions from one chatbot. This might also mean the user does not have time to learn the style of the answers.

All of the chatbots occasionally used terms related to housing industry which the user might not know without previous experiences or research. For instance, *Kaija* used the term *päävuokralainen* [principle tenant]. This is a legal term related to the tenant contract. Its meaning can be deduced to an extent so that the answer is understandable. In my user profile, I discussed that this user is a novice who has not lived in a student apartment before. Thus, they might not be familiar with the terminology related to this area. This is, however, a grey area. If the user has lived on their own before, they most likely have had a lease and are to some extent familiar with the terminology. Therefore, I cannot state for certain whether the user knows the terminology or not. A good thing was that all chatbots were consistent in their terminology, as was recommended in the ninth usability heuristic in chapter 3.4. The terms were also consistent between answers, so that a tenant was always a tenant and not something else.

Also, some of the answers were difficult to read in a chat window as they were so long or in multiple parts. This required scrolling back and forth. That is not the easiest thing to do as the scrolling bar is also small in a small chat window. Sometimes scrolling was necessary because the answers were split into smaller parts. These parts were usually easier to read and follow than one long answer. In these cases, the scrolling did not feel as bothersome.

The chatbots were divided in how they offered additional information. For instance, *Helmi* offered links to the relevant information on the association's website. Others, like *Kaija* and *Tane,* offered many quick replies which sometimes meant another message and sometimes there was a link to the association's website. *Kaija* also answered with the longest answers, thus providing a lot of additional information in the answer itself. Then were the ones like *Elli* which did not really offer any additional information in a form of links or quick replies. Sometimes the additional information was useful in processing the answer and understanding the circumstances properly. For example, *Helmi's* answer about returning the deposits was easier to understand with the help of the additional information. Other times this additional information was useful as it was directly related to the task, for instance many of the chatbots provided a link to the notice of termination which could be done on the association's website. This type of additional information helps the user as they do not need to search for it themselves. However, not all additional information that the chatbots provided was relevant to the type of a user in this study. Therefore, it is better to offer additional information in a form which allows the user to choose the information relevant to their situation.

Overall, the relevant answers could be improved in certain aspects. The style of the answers required learning as it was at times complex and at times even resembled statements from a contract. As mentioned in chapter 4.3, cognitive walkthrough focuses on the learnability of a system and it shows in the evaluations of this study as well. One of the main findings was that the answers actually required learning. The answers could be improved at least in the following aspects: the user could be addressed directly; the sequencing of actions could match the real life better; and the relevance of the offered information could be specified. Some of the answers provided more information than necessary in the actual answer when the chatbot could have asked whether the user wants more information on the subject. If these aspects were improved, the answers would be easier to understand from the start and would not require as much effort from the user.

6 CONCLUSIONS

I began this study with the research question: how usable are the answers that customer service chatbots provide to the users? I would say the answer is moderately. During the individual walkthroughs, there were multiple usability findings. There were usable aspects, for instance the chatbots often provided additional information, which is helpful to a novice user. All the correct answers to the questions included useful and relevant information. However, these answers often required consideration instead of being plainly understandable. I often had to read them twice to comprehend the instructions. For instance, the phrasing in some answers was difficult to read in a chat window. One answer from *Elli* chatbot even included a clear spelling mistake. These types of issues can be noticed with a usability evaluation similar to the one performed in this study. The answers could be improved with small corrections, for example ensuring the proper sequencing of actions. It can also be noted that none of the chatbots could answer all the questions in this study. The results of this study demonstrate how the existing answers can be improved. They also show what to consider while designing new answers, so that they would be usable.

Surprisingly, the answers did require the user to learn their style. I had expected that the chatbots as applications might require learning, but not necessarily the answers. It might be that this style has been logical to the associations which are, of course, knowledgeable about the subject. However, this style does not serve a novice user as well as it does a more experienced user. This could indicate that using customer service chatbots and understanding their answers is not always as simple as it might appear. Of course, the chatbots in this study are all from the same field. Thus, it is not a wide representation of customer service chatbots in general. Further studies in customer service chatbots from other fields could show similar or different results to this study.

As a method, the cognitive walkthrough enabled me to focus on the style of the answers and how quickly I could understand the answers. It also ensured I could include six chatbots into this study. With another method, the findings might be more dispersed to other usability aspects. For instance, heuristic evaluation of the answers could provide more information on all the aspects of usability. However, it might not be as flexible a method for six chatbots in one study. If the study material would be collected with 14 questions, as in this study, this might be especially true.

Chatbots are not exactly a new technology, but this is the time when they are becoming more and more common in everyday life. As they are the new common in interacting with businesses, their usability could be studied more. Because the answers are an important way through which we experience the interaction with chatbots, studying their usability even further would be relevant. Even with just one evaluator, the cognitive walkthrough revealed multiple possibilities for improvement in the answers. If this would be replicated with multiple evaluators from different disciplines, there could be even more usability findings. When the answers are examined as technical communication, the focus can be in the content of the answers. This approach could be combined to other approaches, so that social characteristics and personality of the chatbot can also be examined.

This study could be further developed by having larger tasks which the evaluator performs with the help of a chatbot. For instance, this type of task could be actually applying for an apartment and performing all the steps required for that with the help of a chatbot. Usability testing could also be helpful in determining what type of answers the users prefer. Using both usability evaluations and usability testing would be ideal in finding out the different usability issues. Usability evaluations are easier and cheaper to perform and, therefore, could be replicated a few times before a proper usability test.

The approach to the answers as products of technical communication provided a new and useful perspective to the answers. During the evaluation of answers in chapter 5.7, there were multiple instances in which usability heuristics of documentation could be linked to the answers. The usability of them could be improved by considering the usability heuristics. Thus, a heuristic evaluation of the answers could be fruitful. Also, this technical communication approach could be replicated in other studies. Thus, this study shows that technical communication research can have a role in studying the usability of chatbots. I would like to encourage more usability evaluation of the chatbot answers. As chatbots are becoming more and more common, the demands for their usability will likely also rise. The usability of answers might even become one of the reasons why some chatbots are preferred over others.

SOURCES

The study material

Hoas 2020. [Website and chatbot] Retrieved: https://www.hoas.fi/ [Accessed 6.2.2020.]

Toas 2020. [Website and chatbot] Retrieved: https://toas.fi/ [Accessed 6.2.2020.]

Koas 2020. [Website and chatbot] Retrieved: https://www.koas.fi/fi/ [Accessed 6.2.2020.]

Joensuun Elli 2020. [Website and chatbot] Retrieved: https://www.joensuunelli.fi/fi/ [Accessed 6.2.2020.]

Tys 2020. [Website and chatbot] Retrieved: https://tys.fi/ [Accessed 6.2.2020.]

Das 2020. [Website and chatbot] Retrieved: https://das.fi/fi [Accessed 6.2.2020.]

Poas 2020. POAS, Find a home. [Website and chatbot] Retrieved: https://poas.fi/en/find-a-home/ [Accessed 6.2.2020.]

Reference pictures

Louvre chatbot 2020. A mobile application. Two screenshots of the same chat between the chatbot and the writer. [Accessed 3.2.2020.]

Sephora 2020. A chatbot on the mobile application Kik. Three screenshots of the same chat between the Sephora's chatbot and the writer. [Accessed 3.2.2020.]

References

Allen, Jo 1996 [1990]. The Case Against Defining Technical Writing. In Dan Jones (ed.), *Defining Technical Communication.* Arlington, VA: Society for Technical Communication. 9–14.

Ames, Andrea L. 2001. Just what they need, just when they need it: an introduction to embedded assistance. In *Proceedings of the 19th annual international conference on Computer documentation* (*SIGDOC '01*). New York, USA: Association for Computing Machinery. 111–115.

Android 2020a. *Android Developers*: *Documentation, Guides, User Interface*. [Documentation] Retrieved: https://developer.android.com/guide/topics/ui/controls/button [Accessed 4.2.2020.]

Android 2020b. *Android Developers*: *Documentation, Reference, Web View*. [Documentation] Retrieved: https://developer.android.com/reference/android/webkit/WebView [Accessed 3.2.2020.]

Batish, Rachel 2018. *Voicebot and Chatbot Design: flexible conversational interfaces with Amazon Alexa, Google Home, and Facebook Messenger.* Birmingham, Mumbai: Packt Publishing. O'Reilly Safari Learning Platform: Academic Edition. Retrieved: https://learning.oreilly.com/library/view/voicebot-and-chatbot/9781789139624/ [Accessed 24.11.2019.]

Björksten, Tuomo, Joel Kanerva & Stina Tuominen 2020. *Mistä töitä?* Yle Uutiset. [Web article] Retrieved: https://yle.fi/uutiset/3-11090530 [Accessed 6.3.2020.]

Blum, Kelly 2020. *5 Customer Service Tech Trends for 2020.* Gartner. [Web article] Retrieved: https://www.gartner.com/smarterwithgartner/5-customer-service-tech-trends-for-2020/ [Accessed 5.3.2020.]

Botlist 2020. Botlist.co. [Website] Retrieved: https://botlist.co/ [Accessed 6.3.2020.]

Bradeško, Luka & Dunja Mladenić 2012. *A Survey of Chatbot Systems through a Loebner Prize Competition.* Ljubljana, Slovenia: Artificial Intelligence laboratory, Jozef Stefan Institute.

Bruner, Jon & Mike Barlow 2016. *What Are Conversational Bots?* Sebastopol, CA: O'Reilly Media Inc. O'Reilly Safari Learning Platform: Academic Edition. Retrieved: https://learning.oreilly.com/library/view/what-are-conversational/9781492042808/ [Accessed 1.2.2020.]

Chaves, Ana Paula & Marco Gerosa 2019. *How should my chatbot interact? A survey on human-chatbot interaction design.* Retrieved: https://arxiv.org/abs/1904.02743 [Accessed 29.2.2020.]

Christensson, Per 2014. *FAQ Definition. TechTerms*. Sharpened Productions. [Website] Retrieved: https://techterms.com/definition/faq [Accessed 20.2.2020.]

Davydova, Olga 2017. *25 Chatbot Platforms: A Comparative Table.* Chatbots Journal. [Web article] Retrieved: https://chatbotsjournal.com/25-chatbot-platforms-a-comparative-table-aeefc932eaff [Accessed 11.3.2020.]

Dale, Robert 2016. The return of the chatbots. *Natural Language Engineering* 22:5. 811–817.

Dal Porto, Lindsay 2017. Chatbot Chatter. *Quality Progress* 50:7. 6–7, 9. Retrieved: https://search.proquest.com/openview/b78e79d6dbb67de2a9a8b1ee688773ba [Accessed 11.3.2020.]

De Angeli, Antonella, Paula Lynch & Graham Johnson 2001. Personifying the e-market: A framework for social agents. In Michitaka Hirose (ed.), *Human-Computer Interaction, INTERACT. International Conference on Human-Computer Interaction*. Amsterdam, Netherlands: IOS Press. 198–205.

Deshpande, Aditya, Alisha Shahane, Darshana Gadre, Mrunmayi Deshpande & Prof. Dr. Joshi M. Prachi 2017. A survey of various chatbot implementation techniques. *International Journal of Computer Engineering and Applications* IX:Special issue.

Digi 2020a = Digi- ja väestötietovirasto 2020. *Etunimihaku*. Search term: Elli. Retrieved: https://verkkopalvelu.vrk.fi/nimipalvelu/nimipalvelu_navigaatio.asp [Accessed 21.2.2020.]

Digi 2020b = Digi- ja väestötietovirasto 2020. *Etunimihaku*. Search term: Helmi. Retrieved: https://verkkopalvelu.vrk.fi/nimipalvelu/nimipalvelu_navigaatio.asp [Accessed 21.2.2020.]

Digi 2020c = Digi- ja väestövirasto 2020. *Etunimihaku.* Search term: Kaija. Retrieved: https://verkkopalvelu.vrk.fi/nimipalvelu/nimipalvelu_navigaatio.asp [Accessed 21.2.2020.]

Digi 2020d = Digi- ja väestövirasto 2020. *Etunimihaku.* Search term: Toivo. Retrieved: https://verkkopalvelu.vrk.fi/nimipalvelu/nimipalvelu_navigaatio.asp [Accessed 21.2.2020.]

Digi 2020e = Digi- ja väestövirasto 2020. *Etunimihaku.* Search term: Tane. Retrieved: https://verkkopalvelu.vrk.fi/nimipalvelu/nimipalvelu_navigaatio.asp [Accessed 21.2.2020.]

Eclectic Energies. Landsteiner, Norbert 2020. *Ego, Eliza.* [Chatbot] Retrieved: https://www.eclecticenergies.com/ego/eliza [Accessed 5.3.2020.]

Estrin, Herman A. & Norbert Elliot 1990. *Technical Writing in the Corporate World. Understanding, Developing, and Editing Technical Documents.* Los Altos: Crisp Publications Inc.

Fintech 2017. *Citi debuts its first facebook messenger banking chatbot in Singapore.* Fintech Innovation. Retrieved: https://libproxy.tuni.fi/login?url=https://search-proquest-com.libproxy.tuni.fi/docview/1938296353?accountid=14242 [Accessed 20.3.2020.]

Følstad, Asbjørn, Cecilie Bertinussen Nordheim & Cato Alexander Bjørkli 2018. What Makes Users Trust a Chatbot for Customer service? An Explanatory Interview Study. In Svetlana Bodrunova (ed.), *Lecture Notes in Computer Studies, 11193*. International Conference on Internet Science, INSCI 2018. Cham: Springer International Publishing. 194–208. Retrieved: https://doi.org/10.1007/978-3-030-01437-7_16 [Accessed 12.4.2020.]

Følstad, Asbjørn & Marita Skjuve 2019. Chatbots for Customer Service: User Experience and Motivation. In *Proceedings of the 1st International Conference on Conversational User Interfaces*. New York, USA: Association for Computing Machinery. 1–9. Retrieved: https://doi.org/10.1145/3342775.3342784 [Accessed 12.4.2020.]

Gentsch, Peter 2018. *AI In Marketing, Sales and Service: How Marketers Without a Data Science Degree Can Use AI, Big Data and Bots.* Cham, Switzerland: Palgrave Macmillan.

GetJenny 2020a. *How chatbot Helmi is helping to improve the overall customer satisfaction score in cooperation with the customer service team.* GetJenny. [Blog] Retrieved: https://www.getjenny.com/housing-chatbot-improves-the-overall-customer-satisfaction [Accessed 6.2.2020.]

GetJenny 2020b. GetJenny. *Products, Chatbot*. [Website] Retrieved: https://www.getjenny.com/chatbot [Accessed 6.2.2020.]

Giosg 2020. *Products, Live chat*. Giosg. [Website] Retrieved: https://www.giosg.com/en/products/live-chat [Accessed 5.2.2020.]

Gray, Colin M., Seda Yilmaz, Shanna R. Daly, Colleen M. Seifert & Richard Gonzalez 2015. Idea Generation Through Empathy: Reimagining the 'Cognitive Walkthrough'. In *Association for Engineering Education – Engineering Library Division Papers.* Atlanta, USA: American Society for Engineering Education-ASEE. Retrieved: https://lib.dr.iastate.edu/industrialdesign_conf/10/ [Accessed 24.2.2020.]

Griffin, Ben & Laurel Baston 2014. *Interfaces.* Presentation. Presentation material retrieved: http://peace.saumag.edu/faculty/kardas/Courses/CS/Interfaces2007_files/Interfaces2007.ppt [Accessed 2.2.2020.]

Hassan, Zubair 2019. *Which type of chatbot will have a positive impact on your business?* ReadWrite. [Blog] Retrieved: https://readwrite.com/2019/02/14/which-type-of-

chatbots-will-have-a-positive-impact-on-your-businesses/ [Accessed 10.11.2019.]

Hill, Jennifer, W. Randolph Ford & Ingrid G. Farreras 2015. Real conversations with artificial intelligence: A comparison between human-human online conversations and human-chatbot conversations. *Computers in Human Behavior* 49:8. 245–250.

Hupli, Mikko 2018. *Chatbot FAQ – kaikki mitä chatboteista on syytä tietää juuri nyt.* Salesforce blog. [Blog] Retrieved: https://www.salesforce.com/fi/blog/2018/chatbot-usein-kysytyt-kysymykset.html [Accessed 11.11.2019.]

IEC/IEEE 82079-1 2019 = IEC/IEEE International Standard for Preparation of information for use (instructions for use) of products - Part 1: Principles and general requirements 2019. In *IEC/IEEE 82079-1:2019*. Geneva: IEC; New York: IEEE.

IEEE 1063-2001 2001 = IEEE Standard for Software User Documentation 2001. In *IEEE Standard 1063-2001*. New York: IEEE.

ISO/IEC/IEEE 26511 2012 = ISO/IEC/IEEE International Standard – Systems and software engineering -- Requirements for managers of user documentation 2012. In *ISO/IEC/IEEE 26511*. The corrected version. Geneva: ISO; Geneva: IEC; New York: IEEE.

ISO/IEC/IEEE 24765 2017 = ISO/IEC/IEEE International Standard – Systems and software engineering – Vocabulary. In *ISO/IEC/IEEE 24765:2017.* Second edition. Geneva: ISO; New York: IEEE.

Janarthanam, Srini 2017. *Hands-On Chatbot and Conversational UI Development.* Birmingham, Mumbai: Packt Publishing. O'Reilly Safari Learning Platform: Academic Edition. Retrieved: https://learning.oreilly.com/library/view/hands-on-chatbots-and/9781788294669/ [Accessed 5.3.2020.]

Jones, Dan 1996. Introduction. In Dan Jones (ed.), *Defining Technical Communication.* Arlington, VA: Society for Technical Communication. v.

Karreman, Joyce, Nicole Ummelen & Michaël Steehouder 2005. Procedural and Declarative Information in User Instructions: What We Do and Don't Know About These Information Types. In *IPCC 2005: IEEE International Professional Communication Conference Proceedings.* New York: IEEE. 328–333.

Khan, Rashid & Anik Das 2017. *Build Better Chatbots: A Complete Guide to Getting Started with Chatbots.* Cambridge, MA: AP Professional. O'Reilly Safari Learning Platform: Academic Edition. Retrieved: https://learning.oreilly.com/library/view/build-better-chatbots/9781484231111/ [Accessed 17.10.2019.]

Kopp, Stefan, Lars Gesellensetter, Nicole C. Krämer & Ipke Wachsmuth 2005. A Conversational Agent as Museum Guide – Design and Evaluation of a Real-World Application. In Themis Panayitopolous, Jonathan Gratch, Ruth Aylett, Daniel Ballin, Patrick Olivier & Thomas Rist (eds.), *Intelligent Virtual Agents: 5th International Working Conference*, *IVA 2005 Kos, Greece*. Berlin, Heidelberg: Springer-Verlag. 329–343.

Kupiainen, Reijo 2011. Finnish media literacy policies and research tendencies within a European Union context. *International journal of media and cultural politics* 6:3. 335–341.

Klüwer, Tina 2011. From chatbots to dialog systems. In Diana Perez-Marin & Ismael Pascual-Nieto (eds.), *Conversational agents and natural language interaction: techniques and effective practices.* Hershey: Information Science Reference (IGI Global). 1–22.

LeCun, Yann, Yoshua Bengio & Geoffrey Hinton 2015. Deep learning. *Nature* 521:7553. 436–444.

Lemon, Katherine N. & Peter C Verhoef 2016. Understanding Customer Experience Throughout the Customer Journey. *Journal of Marketing* 80:6. 69–96.

Lester, James, Karl Branting & Bradford Mott 2004. Conversational agents. In M.P. Singh (ed.), *Practical Handbook of Internet Computing.* New York: Chapman and Hall, CRC. 220–240.

Liu, Bingjie & S. Shyam Sundar 2018. Should machines express sympathy and empathy? Experiments with a health advice chatbot. *Cyberpsychology, Behavior and Social Networking* 21:10. Retrieved: https://www.liebertpub.com/doi/10.1089/cyber.2018.0110 [Accessed 16.10.2019.]

Mathur, Saurabh & Daphne Lopez 2018. A scaled-down neural conversational model for chatbots. *Concurrency and Computation: Practice and Experience* 31:10. Retrieved: https://doi.org/10.1002/cpe.4761 [Accessed 28.11.2019.]

Mayo, Joe 2017. *Programming the Microsoft Bot Framework: A Multiplatform Approach to Building Chatbots.* Pearson Education Inc., Microsoft Corporation. O'Reilly Safari Learning Platform: Academic Edition. Retrieved: https://learning.oreilly.com/library/view/programming-the-microsoft/9781509305001/ [Accessed 30.11.2019.]

Messina, Chris 2015. *Conversational commerce: Messaging apps bring the point of sale to you.* Medium. [Web article] Retrieved: https://medium.com/chris-messina/conversational-commerce-92e0bccfc3ff [Accessed 3.4.2020.]

Messina, Chris 2016. *2016 will be the year of conversational commerce.* Medium. [Web article] Retrieved: https://medium.com/chris-messina/2016-will-be-the-year-of-conversational-commerce-1586e85e3991#.f50xaqevj [Accessed 3.4.2020.]

Mohammed, Mohssen, Eihab Mohammed Bashier & Muhammad Badruddin Khan 2016. *Machine Learning: Algorithms and Applications.* Boca Raton; London; New York: CRC Press, Taylor & Francis Group, LLC. O'Reilly Safari Learning Platform: Academic Edition. Retrieved: https://learning.oreilly.com/library/view/machine-learning/9781315354415/ [Accessed 4.3.2020.]

Moran, Kate 2016. *The Four Dimensions of Tone of Voice.* Nielsen Norman Group. [Web article] Retrieved: https://www.nngroup.com/articles/tone-of-voice-dimensions/ [Accessed 5.3.2020.]

Nielsen, Jakob 1994. *Usability Engineering.* Boston: AP Professional.

Nielsen, Jakob 2002. *Top 10 Web Design Mistakes of 2002.* Nielsen Norman Group. [Web article] Retrieved: https://www.nngroup.com/articles/top-ten-web-design-mistakes-of-2002/ [Accessed 6.3.2020.]

Nielsen, Jakob 2012. *Usability 101: Introduction to Usability.* Nielsen Norman Group. [Web article] Retrieved: https://www.nngroup.com/articles/usability-101-introduction-to-usability/ [Accessed 6.3.2020.]

O'Brien, Jason 2019. 5 technology trends that are changing business communication. *Communication world* 4. 4.

Oxford 2016. *A Dictionary of Social Media.* Online version. Search term*:* online chat*.* Daniel Chandler & Rod Munday (eds.). Oxford University Press. Retrieved: https://www.oxfordreference.com/view/10.1093/acref/9780191803093.001.0001/acref-9780191803093-e-168 [Accessed 11.3.2020.]

Panetta, Kasey 2017. *Gartner Top Strategic Predictions for 2018 and Beyond*. Gartner. [Web article] Retrieved: https://www.gartner.com/smarterwithgartner/gartner-top-strategic-predictions-for-2018-and-beyond/ [Accessed 5.3.2020.]

Priest, Susanna Hornig 2010. *Encyclopedia of Science and Technology Communication Volume 1*. Los Angeles; London: Sage Publications, Inc.

Purho, Vesa 2000. Heuristic Inspections for Documentation – 10 Recommended Documentation Heuristics*. Usability Interface, STC Usability SIG Newsletter* 6:4.

Radziwill, Nicole & Morgan Benton 2017. *Evaluating Quality of Chatbots and Intelligent Conversational Agents.* Retrieved: https://arxiv.org/abs/1704.04579 [Accessed 4.3.2020.]

Ranne, Sanna 2005. Kognitiivinen läpikäynti. In Saila Ovaska, Anne Aula & Päivi Majaranta (eds.), *Käytettävyystutkimuksen menetelmät*. Tampere: Tampereen yliopisto. 125–140. Retrieved: https://trepo.tuni.fi/handle/10024/96627 [Accessed 25.3.2020.]

Ray, Deborah. S., & Eric. J Ray 2001. Embedded help: Background and applications for technical communicators. *Technical Communication* 48:1. 105–115.

Reeves, Byron & Clifford Nass 1996. *The Media Equation: How People Treat Computers, Television and New Media Like Real People and Places.* Méridien; Frutiger: CSLI Publications.

Rosenbaum, Mark S., Mauricio Losada Otalora & Germán Contreras Ramirez 2017. How to create a realistic customer journey map. *Business Horizons* 60:1. 143–150.

Rosenzweig, Elizabeth 2015. *Successful User Experience: Strategy and Roadmaps.* Amsterdam, Netherlands: Morgan Kaufmann, Elsevier Science and Technology.

Rowley, David & David Rhoades 1992. The cognitive jogthrough: a fast-paced user interface evaluation procedure. In *Proceedings of the SIGHI Conference on Human Factors in Computing Systems (CHI '92).* New York, USA: Association for Computing Machinery. 389–395.

Saenz, Julia, Walker Burgess, Elizabeth Gustitis, Andres Mena & Farzan Sasangohar 2017. The usability analysis of chatbot technologies for internal personnel communications. In *67th Annual Conference and Expo of the Institute of*

Industrial Engineers 2017. Pittsburgh, USA: Institute of Industrial Engineers. 1357–1362.

Savola, Pasi 2018. *Danfoss Drives – how to utilize modular documentation with a chatbot*. Seminar lecture 1.11.2018. STVY Fall Seminar 2018. Lecture material retrieved from: https://www.stvy.fi/wp-content/uploads/2018/10/Savola_Pasi.pdf [Accessed 30.11.2019.]

Schlicht, Matt 2016. *The Complete Beginners Guide To Chatbots: Everything you need to know.* Chatbots magazine. [Web article] Retrieved: https://chatbotsmagazine.com/the-complete-beginner-s-guide-to-chatbots-8280b7b906ca [Accessed 3.4.2020.]

Schlicht, Matt 2018. *The Complete Guide to Conversational Commerce: Everything you need to know*. Chatbots magazine. [Web article] Retrieved: https://chatbotsmagazine.com/the-complete-guide-to-conversational-commerce-e47059293efa [Accessed 3.4.2020.]

Sebastian, Nathan 2019. *Usage & Statistics for an Engaging Online Chatbot Design: GoodFirms Research*. GoodFirms. [Web article] Retrieved: https://www.goodfirms.co/resources/online-chatbot-technology [Accessed 5.2.2020.]

Sharma, Bhuvan, Harshita Puri & Deepika Rawat 2018. Digital Psychiatry - Curbing Depression using Therapy Chatbot and Depression Analysis. In *2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT).* Coimbatore, India: IEEE. 627–631.

Shevat, Amir 2017. *What Are Bots? - Designing Bots.* Sebastopol, CA: O'Reilly Media Inc. O'Reilly Safari Learning Platform: Academic Edition. Retrieved: https://learning.oreilly.com/library/view/designing-bots/9781491974810/ch01.html [Accessed 2.2.2020.]

Shneiderman, Ben, Catherine Plaisant, Maxine Cohen, Steven Jacobs, Niklas Elmqvist & Nicholas Diakopoulos 2017 [1988]. *Designing the User Interface: Strategies for Effective Human-Computer Interaction.* Global Edition. 6th Edition. England: Pearson Education Limited.

Simpson, Henry & Steven M. Casey 1988. *Developing Effective User Documentation: A Human-Factors Approach.* New York, USA: McGraw-Hill Inc.

Sinkkonen, Irmeli, Hannu Kuoppala, Jarmo Parkkinen & Jarno Vastamäki 2009 [2006]. *Käytettävyyden psykologia.* Electronic version. Adage Oy.

Skjuve, Marita, Ida Maria Haugstveit, Asbjørn Følstad & Petter Bae Brandtzaeg 2019. Help! Is my chatbot falling into uncanny valley? An empirical study of user experience in human-chatbot interaction. *Human Technology* 15:1. 30–54.

Stratton, Charles R 1996 [1979]. Technical Writing: What It Is and What It Isn't. In Dan Jones (ed.), *Defining Technical Communication.* Arlington, VA: Society for Technical Communication. 39–42.

SVT 2012 = Suomen virallinen tilasto 2012. *Yliopistotutkinnon suorittaneet 2012, Liitetaulukko 4. Yliopistojen alemman korkeakoulututkinnon uudet opiskelijat ja opiskelijat iän mukaan 2012.* Helsinki: Tilastokeskus. [Web publication] Retrieved: http://www.stat.fi/til/yop/2012/01/yop_2012_01_2013-04-23_tau_004_fi.html [Accessed 6.2.2020.]

TCBOK 2020 = Technical Communication Body of Knowledge 2020. *Definition of Technical Communication*. Society for Technical Communication. [Wiki] Retrieved: https://www.tcbok.org/wiki/about-tc/definition-technical-communication/ [Accessed 2.4.2020.]

Tecuci, Gheorghe 2012. Artificial Intelligence. *WIREs Computational Statistics* 4:2. 168–180.

Terveyskylä 2020. Koronabotti. Terveyskylä. [Chatbot] Retrieved: https://www.terveyskyla.fi/koronabotti/ [Accessed 16.3.2020.]

Turing, A.M 1950. I. - Computing Machinery and Intelligence. *Mind: A Quarterly Review of Psychology and Philosophy* LXI:236. 433–460.

TSK 2020 = Sanastokeskus TSK ry 2020. *TEPA-termipankki: Erikoisalojen sanastojen ja sanakirjojen kokoelma*. Search term: bot. Sanastokeskus TSK ry. Retrieved: http://www.tsk.fi/tepa/fi/haku/bot [Accessed 2.3.2020.]

Ummelen, Nicole 1994. Procedural and declarative information: a closer examination of the distinction. In Michaël Steehouder, Carel Jansen, Pieter van der Poort & Ron Verheijen (eds.), *Quality of technical documentation*. Amsterdam: Rodopi. 115–130.

Weizenbaum, Joseph 1966. Computational Linguistics. ELIZA – A Computer Program For The Study of Natural Language Communication Between Man And Machine. *Communications of the ACM* 9:1. 36–45.

Williams, Sam 2018. *Hands-On Chatbot Development with Alexa Skills and Amazon Lex*. Birmingham, Mumbai: Packt Publishing. O'Reilly Safari Learning Platform: Academic Edition. Retrieved: https://learning.oreilly.com/library/view/hands-on-chatbot-development/9781788993487/ [Accessed 17.10.2019.]

Wilson, Chauncey 2014. *User Interface Inspection Methods: A User-Centered Design Method.* Waltham, MA: Morgan Kaufmann, Elsevier. Retrieved: https://doi.org/10.1016/C2012-0-06519-2 [Accessed 16.3.2020.]

SUOMENKIELINEN LYHENNELMÄ

1 Johdanto

Chatbotit ovat yleistyneet monilla aloilla helpottamaan ihmisten työtehtäviä (O'Brien 2019, 4). Perinteisesti chatbot määritellään teknologisena agenttina, joka kommunikoi käyttäjän kanssa luonnollisen kielen välityksellä (Khan & Das 2017). Monesti tällä tarkoitetaan kirjoitettua kieltä (Dale 2016). Nykyisin chatbotteja käytetään muun muassa terapiassa (Sharma, Puri & Rawat 2018) ja henkilökohtaisena pankkiirina (Fintech 2017). Chatbotteja käytetään myös osana uutisia, esimerkiksi Ylellä on ollut *Duunibotti* (Björksten, Kanerva & Tuominen 2020).

Ala, jolla chatbotit ovat todella yleistyneet, on asiakaspalvelu (O'Brien 2019, 4). Yrityksen kannalta ne ovat hyödyllinen ratkaisu. Chatbotit ovat esimerkiksi nopeampia vastaamaan kysymyksiin kuin ihmiset (Khan ym. 2017). Ne ovat myös halvempaa työvoimaa (Lester, Branting & Mott 2004, 3). Käyttäjän kannalta ne ovat hyödyllisiä, koska ne tarjoavat jatkuvaa palvelua (Dal Porto 2017, 6). Koska tekoälylliset ratkaisut kehittyvät koko ajan, on todennäköistä, että chatbottien määrä asiakaspalvelussa vain kasvaa (O'Brien 2019, 4). On siis tärkeää, että chatbottien käytettävyyttä parannetaan niiden yleistyessä ensimmäisenä asiakaspalvelukokemuksena. Hyvä asiakaspalvelu on tärkeää yritykselle, koska huono palvelu voi johtaa tulojen ja asiakkaan menetykseen (Lester ym. 2004, 3).

Chatbottien käytettävyyden arviointi on siis hyvin ajankohtaista. Aikaisemmissa tutkimuksissa keskiössä on ollut ihmisen ja chatbotin välinen kanssakäyminen (Liu & Sundar 2018; Skjuve, Haugstveit, Følstad & Brandtzaeg 2019). Näissä tutkimuksissa on tarkasteltu chatbotin ja käyttäjän välistä kommunikaatiota sekä käyttäjien toivomaa kommunikaatiotyyliä chatboteilta. On tutkittu myös teknologisen ratkaisun käytettävyyttä tiettyyn tehtävään (Saenz, Burgess, Gustitis, Mena & Sasangohar 2017). Aikaisemmassa käytettävyystutkimuksessa ei ole laajalti huomioitu pelkästään chatbottien käyttäjille tarjoamia vastauksia. Vastaukset ovat kuitenkin iso osa käyttäjäkokemusta, sillä ne ovat chatbotin tapa kommunikoida käyttäjän kanssa. Tässä tutkimuksessa keskitytään nimenomaan vastausten käytettävyyteen. Tutkimuksen tutkimuskysymys onkin, kuinka käytettäviä asiakaspalvelun chatbottien vastaukset ovat.

2 Chatbottien määritelmä ja luokittelu

Kuten jo mainitsin, chatbot voidaan määritellä teknologisena agenttina, joka kommunikoi käyttäjän kanssa luonnollisen kielen välityksellä (Khan ym. 2017). Chatbotit ovat kuitenkin muuttuneet

paljon teknologian kehityksen myötä (mt.). Sen vuoksi modernit chatbotit voivat erota perinteisestä määritelmästä. Esimerkiksi jotkut modernit chatbotit pystyvät käsittelemään käyttäjän verbaalisia komentoja ja vastavuoroisesti vastaamaan verbaalisesti (Bruner & Barlow 2016; Deshpande, Shahane, Gadre, Deshpande & Prachi 2017; Shevat 2017). Jotkut modernit chatbotit pystyvät myös suorittamaan tehtäviä pelkän kysymysten vastaamisen sijaan (Deshpande ym. 2017). Sen vuoksi chatbotit kuten Alexa tai Siri eivät vastaa perinteistä chatbotin määritelmää. Joskus ne luokitellaan ääniohjatuiksi keskusteluagenteiksi (*voiceactivated conversational agent*)[1] (Radziwill & Benton 2017), jotta ne eroteltaisiin chatboteista.

Teknologisen toteutuksen mukaan chatbotit voidaan jakaa kahteen kategoriaan: käsikirjotettuihin ja älykkäisiin (Hassan 2019; Hupli 2018). Käsikirjoitetut chatbotit on ohjelmoitu seuraamaan tiettyjä sääntöjä (Hassan 2019). Ne seuraavat ihmisen kirjoittamaan dialogia eivätkä pysty vastaamaan kysymyksiin, jotka eivät ole osa tätä dialogia (mt.). Sen vuoksi ne pystyvät vastaamaan vain niihin ohjelmoituihin käskyihin (Schilicht 2016). Sen sijaan älykkäät chatbotit hyödyntävät tekoälyä ja/tai koneoppimista (Hassan 2019). Ne pystyvät hienostuneemmin käsittelemään käyttäjän viestejä ja jopa ennustamaan käyttäjän tarpeita (mt.). Ne eivät vain seuraa komentoja vaan pystyvät analysoimaan käyttäjän viestejä (Schlicht 2016). Niiden vastaukset käyttäjälle ovat kuitenkin yleensä valmiiksi kirjoitettuja eivätkä sen itsensä generoimia (Hupli 2018).

Teknologisen toteutuksen lisäksi chatbotit voidaan jakaa muun muassa sillä perusteella, minkälaisia palveluja ne tarjoavat, miten niitä on tarkoitus käyttää ja millä aloilla niitä käytetään. Näitä aloja on ainakin viisi: asiakaspalvelu, neuvonta, nettisivujen navigointi, ohjeistettu ostaminen ja tekninen tuki (Lester ym. 2004, 2–3). Näistä viidestä nimenomaan asiakaspalvelu on tämän tutkimuksen kohteena, koska kaikki tutkimuksen chatbotit ovat asiakaspalveluun käytettäviä chatbotteja. Chatbottien ja muiden vuorovaikutukseen perustuvien teknologisten ratkaisujen yleistyessä asiakaspalvelussa Messina (2015) kehitti termin keskusteleva kaupanteko (*conversational commerce)* kuvailemaan tätä ilmiötä. Keskusteleva kaupanteko tarkoittaa chatin, viestien tai muiden luonnollisen kielen käyttöliittymien käyttämistä vuorovaikutuksessa brändien, palveluiden tai bottien kanssa (Messina 2016). Perinteisesti nämä osapuolet eivät ole keskenään osallistuneet kaksisuuntaiseen viestittelykontekstiin (mt.). Tämä kaupanteko keskittyy äärimmäiseen yksilöllistämiseen jokaiselle käyttäjälle (mt.). Viestipohjaiseen vuorovaikutukseen keskustelevassa kaupanteossa kuuluvat muun muassa tilausvahvistuksen tai lähetystietojen lähettäminen asiakkaalle, ostoksesta kiittäminen, tuotteiden suositteleminen keskustelujen tai tilaushistorian avulla sekä

[1] Kaikki tässä lyhennelmässä olevat termikäännökset ovat kirjoittajan omia käännöksiä.

vianselvitys (Schlicht 2018). Vianselvitys tässä kontekstissa tarkoittaa asiakkaan ongelmien selvittämistä (mt.). Asiakkaan vielä selatessa palveluita netissä chatbot voi toimia asiakaspalvelijan roolissa (mt.). Tämä tarkoittaisi muun muassa kysymyksiin vastailua, tarjousten tekemistä ja palveluiden selittämistä (mt.). Chatbotit asiakaspalvelijana voivat siis tehdä monia asioita käyttäjälle. Jotkut näistä tehtävistä keskittyvät tehtävien suorittamiseen, kuten vianselvitys. Toiset ovat enemmän informatiivisia, kuten käyttäjän kysymyksiin vastaaminen.

Tässä tutkimuksessa esitän, että chatbottien vastauksia voi verrata teknisen viestinnän tuottamiin teksteihin eli tarkemmin ohjeistaviin teksteihin. Teknisen viestinnän määritteleminen ei ole helppoa (Allen 1996, 9). Yhden määritelmän mukaisesti tekninen viestintä toimittaa selkeää ja totuudenmukaista tietoa käyttäjille (TCBOK 2020). Teknisessä viestinnässä tuotetaan monenlaisia asioita, kuten online-ohjeita, oppaita ja käyttöliittymätekstejä (mt.).

Teknisen viestinnän tuotoksissa voi olla useita informaatiotyyppejä (Karreman, Ummelen & Steehouder 2005, 328). Keskeisimmät informaatiotyypit ovat ohjeistava ja kuvaileva (mp.). Ohjeistava tieto kuvailee toimintaa ja onkin kaikista tärkein informaatiotyyppi järjestelmän käytön aikana (mp.). Menetelmä informaatio kirjoitetaan yleensä askel-askeleelta muodossa (Estrin & Elliot 1990, 50; Simpson & Casey 1988, 10). Sen sijaan kuvaileva tieto on selittävää (Karreman ym. 2005, 328). Kuvailevaa tietoa ovat tiedot järjestelmästä, jotka käyttäjä tarvitsee oppiakseen käyttämään sitä (Simpson ym. 1988, 10).

Jako ohjeistavaan ja kuvailevaan tietoon ei ole aina selkeä (Karreman ym. 2005, 330; Ummelen 1994, 123). Ne ovat laajoja termejä, jotka käsittävät muita informaatiotyyppejä (Ummelen 1994, 124). Myös tekstin muotoilu voi olla ohjeistava tai kuvaileva (mp.).

Asiakaspalvelussa käytettävät chatbotit voivat tarjota sekä ohjeistavia vastauksia että kuvailevia vastauksia. Ne voivat esimerkiksi tarjota tarpeellisia tietoja palvelusta tai tuotteesta. Ne voivat myös ohjeistaa toiminnassa, esimerkiksi aikaisemmin mainituissa vianselvitystapauksissa. Tämän takia chatbottien vastauksia voidaan tarkastella myös ohjeteksteinä.

3 Käytettävyys

Tässä tutkimuksessa pääasiallinen teoreettinen viitekehys on Nielsenin (1994) teoria käytettävyydestä. Nielsen (1994, 25) kuvailee käytettävyyttä järjestelmän ominaisuutena, joka kuvailee, kuinka hyvin käyttäjä voi käyttää systeemin toimintoja. ISO/IEC/IEEE 24765 (2017, 492) määrittelee, että käytettävyys kuvaa, kuinka hyvin käyttäjä voi saavuttaa tietyt tavoitteet systeemin,

tuotteen tai palvelun avulla. Tämän pitäisi onnistua tehokkaasti ja miellyttävästi (mp.). Nielsen (1994, 26–27) korostaa, että käytettävyyttä pitää lähestyä systemaattisesti. Sitä pitäisi myös arvioida ja parantaa ohjelmassa (mp.). Suunniteltaessa käytettävyyden kuuluisi mennä pidemmälle kuin käsitteisiin "intuitiivinen" ja "luonnollinen" (Shneiderman, Plaisant, Cohen, Jacobs, Elmqvist & Diakopoulos 2017, 33). Käytettävyys vaikuttaa koko systeemiin (Nielsen 1994, 25), joten sen kehittäminen on tärkeää. Esimerkiksi nettisivuilla käytettävyys on todella tärkeää, koska käyttäjät voivat heti lähteä sivuilta, joilla on käytettävyysongelmia (Nielsen 2012).

Käytettävyys jakautuu viiteen ominaisuuteen: opittavuus (*learnability*), tehokkuus (*efficiency*), muistettavuus (*memorability*), virheellisyys (*errors*) ja tyytyväisyys (*satisfaction*) (Nielsen 1994, 26). Näitä ominaisuuksia voidaan tarkastella erilaisilla käytettävyysmetodeilla (mts. 27). Samankaltaiset ominaisuudet ovat Shneidermanin ym. (2017, 33–34) määritelmät käytettävyyden mittareista: oppimisaika (*time to learn*), suorittamisen nopeus (*speed of performance*), käyttäjän tekemien virheiden määrä (*rate of errors by users*), tiedon säilyvyys ajan kuluessa (*retention over time*) ja subjektiivinen tyytyväisyys (*subjective satisfaction*).

Opittavuus kuvaa sitä, kuinka helposti käyttäjä oppii käyttämään ohjelmaa (Nielsen 1994, 27–28). Se ei tarkoita, että käyttäjä olisi oppinut jokaisen yksityiskohdan ohjelmasta, vaan hän osaa käyttää sitä tarvittavalla tasolla (mts. 28–29). Oppimisaika tarkoittaa lähes samaa, koska siinä mitataan, kuinka nopeasti tyypillinen käyttäjä käyttäjäryhmästä oppii suorittamaan olennaisia tehtäviä (Shneiderman ym. 2017, 33). Tehokkuudella tarkoitetaan aikaa, joka tietyn taitotason käyttäjällä kuluu ohjelmistolla toimintojen suorittamiseen (Nielsen 1994, 30–31). Suorittamisen nopeudella mitataan perustoimintojen suorittamista ohjelmalla (Shneiderman ym. 2017, 34). Tehokas ohjelma on yhtenäinen työnkulussaan ja terminologiassaan (Sinkkonen, Kuoppala, Parkkinen & Vastamäki 2009, 194).

Muistettavuus kuvaa, kuinka hyvin käyttäjä muistaa ohjelman työnkulun (Nielsen 1994, 31–32). Monet nykyajan ohjelmistoista kuitenkin muistuttavat käyttäjää työnkulun etenemisestä käytön aikana, joten käyttäjän ei tarvitse välttämättä muistaa niin paljon (mts. 32). Tiedon säilyvyys ajan kuluessa tarkoittaa käytännössä samaa. Toisto on myös tärkeä asia muistamisessa (Shneiderman ym. 2017, 34). Käyttäjän ei myöskään pitäisi tehdä useita virheitä ohjelmaa käyttäessään (Nielsen 1994, 32). Isompia virheitä, jotka estävät halutun toiminnan suorittamisen, pitäisi erityisesti välttää (mts. 32–33). Käyttäjän tekemien virheiden määrässä mitataan, kuinka paljon käyttäjä tekee virheitä perustoimintoja käyttäessään (Shneiderman ym. 2017, 34). Tyytyväisyys puolestaan kuvaa nauttiiko käyttäjä ohjelman käyttämisestä (Nielsen 1994, 33). Esimerkiksi peleissä tyytyväisyys on yksi

tärkeimmistä käytettävyyden ominaisuuksista (mp.). Subjektiivista tyytyväisyyttä voidaan mitata muun muassa kyselyillä ja haastatteluilla (Shneiderman ym. 2017, 34). Kaikkia näitä ominaisuuksia ei välttämättä voi saavuttaa yhdessä ohjelmassa samaan aikaan (mts. 34–35).

Käytettävyyttä tutkitaan joko käytettävyystutkimuksella tai erilaisilla arviointimenetelmillä (Nielsen 1994, 165; Shneiderman ym. 2017, 171). Käytettävyystestaus tehdään ohjelmiston oikeilla käyttäjillä (Nielsen 1994, 165). Arviointimenetelmissä arvioija on jonkin alan ammattilainen, joka arvioi ohjelmiston käytettävyyttä (Shneiderman ym. 2017, 171). Näihin menetelmiin kuuluvat muun muassa heuristinen arviointi ja kognitiivinen läpikäynti (mp.). Kognitiivinen läpikäynti on tässä tutkimuksessa käytetty arviointimenetelmä.

Käytettävyystutkimuksen seurauksena on kehitetty ohjeistuksia, toimintaperiaatteita ja teorioita (Shneiderman ym. 2017, 82). Koska tässä tutkimuksessa tarkastellaan chatbottien vastauksia ohjeteksteinä, myös niiden käytettävyyttä on tarkasteltu. Purho (2000) on kehittänyt dokumentaation käytettävyydestä heuristiikoita, joiden avulla käytettävyyttä voidaan tarkastella. Näihin heuristiikkoihin lukeutuu muun muassa dokumentin ja tosielämän vastaavuus, erilaisten käyttäjien tukeminen sekä yhtenäisyys ja standardit (mt.).

4 Käytettävyyden arviointimenetelmä ja tutkimusmateriaalin kerääminen

Tässä tutkimuksessa käytän arviointimenetelmänä kognitiivista läpikäyntiä. Kognitiivinen läpikäynti keskittyy ohjelman oppimisen helppouteen (Wilson 2014, 66). Läpikäynnin aikana arvioija matkii oikean käyttäjän toimia (Ranne 2005, 125). Kognitiivisesta läpikäynnistä on monia eri muotoja, kuten epävirallinen kognitiivinen läpikäynti (Wilson 2014, 66) ja kognitiivinen läpijuoksu (*cognitive jogthrough*) (Rowley & Rhoades 1992, 389). Kognitiivinen läpijuoksu ei ole yhtä virallinen muodoltaan ja se on nopeampi toteuttaa kuin läpikäynti (mp.). Sitä käytetään usein nettipalveluiden ja sovellusten arviointiin (mp.). Tässä tutkimuksessa tehty kognitiivinen läpikäynti muistuttaa läpijuoksua, koska se ei ole yhtä virallinen rakenteeltaan, arvioijia on vain yksi ja arvioinnin kohteena on pienempiä sovelluksia. Tämän tutkimuksen tehtävät ovat myös pienempiä, eivätkä rakenteeltaan kovin monimutkaisia, joten kokonaisen virallisen läpikäynnin toteuttaminen voisi olla vaikeaa.

Kognitiiviseen läpikäyntiin tarvittavat materiaalit ja dokumentit ovat seuraavat: esitelmä käyttöliittymästä, raportti ongelmista, käyttäjäprofiili, tehtävälista ja toimintakuvaus jokaisesta tehtävästä (Wilson 2014, 70). Tässä tutkimuksessa olevat chatbotit ovat luonnollisesti käyttöliittymä

itsessään. Tämän tutkimuksen analyysiosiot voidaan tulkita raportiksi ongelmista. Käyttäjäprofiili on seuraava:

- uusi, päätoiminen opiskelija
- 20–30-vuotias
- ei ole aikaisemmin asunut opiskelija-asunnossa
- osaa keskinkertaisesti käyttää jokapäiväistä teknologiaa.

Perustuen tutkimusmateriaalin keräykseen käytettyihin kysymyksiin kyseinen käyttäjä ei olisi kokenut opiskelija-asuntosäätiöiden palvelujen käyttäjä. Koska arviointimenetelmä myös painottaa opittavuutta, uusi opiskelija ilman aikaisempaa kokemusta soveltuu tähän tarkoitukseen. Suomessa uusien kolmannen asteen opiskelijoiden ikä on 20–30 vuotta (SVT 2012), joten se soveltuu ikähaarukaksi. Sukupuolta ei ole huomioitu, koska tässä tutkimuksessa ei ole kysymyksiä sukupuolesta riippuvista asioista, kuten sukupuolen mukaan jaetuista soluasunnoista. Käyttäjällä voidaan olettaa olevan vähintään keskinkertaiset teknologiset taidot, koska Suomessa internet ja teknologia ovat hyvin saatavilla ja media- sekä digitaalista lukutaitoa opetetaan kouluissa (Kupiainen 2011).

Tehtävälista koostuu kysymyksistä, joihin palaan myöhemmin tässä luvussa. Koska tehtävälistan kaikki tehtävät muistuttavat toisiaan pitkälti, niille on laadittu sama toimintakuvaus:

1. Avaa selain.
2. Mene asuntosäätiön nettisivuille.
3. Avaa keskustelu chatbotin kanssa.
4. Kysy kysymys.
5. Seuraa chatbotin ohjeita.
6. Sulje keskustelu.

Tämä toimintakuvaus kuvastaa toivottua tilannetta, jossa chatbot osaa vastata kysymykseen. Jos se ei osaa vastata, askel 5 voidaan jättää väliin ja mennä suoraan askeleeseen 6.

Tähän tutkimukseen on valittu kuusi eri opiskelija-asuntosäätiötä. Niiden valintaan vaikutti, että ne kaikki ovat samalla alalla, joten niiden tarjoamat palvelut ja chatbotit ovat tietyissä määrin samanlaisia. Täten niiden vertailu keskenään on helpompaa. Lisäksi jokaisen chatbot oli julkisesti saatavilla heidän nettisivuillaan samoin kuin usein kysytyt kysymykset, joita käytettiin pohjana haastattelukysymyksille. Kaikkien säätiöiden chatbotit myös operoivat suomeksi, joten kaikki

kysymykset esitettiin suomeksi. Valitut asuntosäätiöt olivat: HOAS – Helsingin seudun opiskelija-asuntosäätiö sr, KOAS – Keski-Suomen opiskelija-asuntosäätiö sr, TYS – Turun Ylioppilaskyläsäätiö, DAS – Domus Arctica -Säätiö, Joensuun Elli -Opiskelija-asunnot Oy ja TOAS – Tampereen opiskelija-asuntosäätiö sr.

Viisi näistä on varsinaisia opiskelija-asuntosäätiöitä: TOAS, HOAS, KOAS, DAS ja TYS. Joensuun Elli on osakeyhtiö. Yleisesti kuitenkin puhun kaikista kuudesta, kun käytän termiä säätiö. Nämä säätiöt tarjoavat opiskelija-asuntoja toisen ja kolmannen asteen opiskelijoille (Tys, Toas). Isoimmat säätiöt ovat Etelä-Suomesta ja niillä on 4 000–10 000 asuntoa (Toas, Hoas, Tys). Pienemmillä säätiöillä on 1 500–2 500 asuntoa (Das, Koas, Elli).

Läpijuoksussa arvioitava tutkimusmateriaali on kerätty haastattelemalla valittujen opiskelija-asuntosäätiöiden chatbotteja. Tätä varten tarvittiin kysymyskokoelma. Tämä kysymyskokoelma on kehitelty opiskelija-asuntosäätiöiden nettisivuilta löytyvistä usein kysytyistä kysymyksistä. Usein kysyttyjen kysymysten pitäisi koostua käyttäjien tavallisimmista kysymyksistä (Christensson 2014; Nielsen 2002). Tämän vuoksi ne voivat tarjota suuntaa siitä, mitä oikeat käyttäjät kysyisivät ja mistä aiheista he ovat kiinnostuneet. Ensin keräsin usein kysytyt kysymykset taulukkoon ja merkkasin ne, jotka olivat useamman säätiön sivuilla. Valitsin kysymykset, jotka mainittiin useamman kuin kahden säätiön sivuilla. Kysymykset olivat kuitenkin muotoilultaan kohtalaisen virallisia. Osa oli esimerkiksi passiivissa tai kolmannessa persoonassa. Muokkasin tällaiset kysymykset ensimmäiseen persoonaan, kuten oikea käyttäjä saattaisi ne kysyä. Poistin kysymyksistä myös asuntosäätiökohtaiset termit tai tiedot. Tämän jälkeen asettelin kysymykset yksinkertaistetulle asiakkaan palvelupolulle.

Asiakkaan palvelupolku on hallintatyökalu, jolla kuvataan asiakaskokemusta palvelun tai tuotteen hankinnan aikana (Rosenbaum, Otalora & Ramirez 2017). Se sisältää kaikki kontaktipisteet, joiden kautta asiakas on vuorovaikutuksessa yrityksen kanssa (Lemon & Verhoef 2016, 69). Palvelupolku jaetaan kolmeen osaan: ennen palvelua, palvelun aikana ja palvelun jälkeen (Rosenbaum ym. 2017). Täten palvelupolku on monesti yrityskohtainen. Tässä tutkimuksessa on kuitenkin kuusi eri asuntosäätiötä, joten jokaiselle oman palvelupolun määritteleminen olisi ollut hankalaa. Myös tulosten vertailu keskenään erilaisten polkujen välillä olisi voinut olla ongelmallista. Sen takia olen tässä tutkimuksessa toteuttanut kapea-alaisemman palvelupolun, johon ovat valikoituneet todennäköisesti yhteiset osat jokaisen säätiön käyttäjäkokemuksesta. Siten tässä tutkimuksessa käytetty palvelupolku on yksinkertaistettu.

Yksinkertaistetun asiakkaan palvelupolun vaiheiden jako tulee jaeottelusta, joka oli asuntosäätiöiden useiden kysytyissä kysymyksissä. Nämä vaiheet ovat samankaltaiset palvelupolun yleisessä jaossa. Siten polun kolme eri vaihetta ovat: hakeminen (ennen palveluja), asunnossa asuminen (palvelujen aikana) ja pois muuttaminen(palvelujen jälkeen). Chatboteille esitettävät kysymykset on jaettu näille kolmelle eri vaiheelle. Ensimmäisessä vaiheessa on neljä kysymystä ja seuraavassa kahdessa vaiheessa on kummassakin viisi kysymystä. Kysymyksiä on siis yhteensä 14.

Pilotoin kysymykset asuntosäätiö POASsin *Onni*-chatbotilla. Pilotoinnin tulosten seurauksena tein pieniä muokkauksia kysymysten muotoiluun. Pilotoinnissa huomasin myös, että osa kysymyksistä oli muotoilunsa vuoksi hankalia chatbotille. Esimerkiksi kysymys *Voinko asentaa asuntooni pyykinpesukoneen ja astianpesukoneen?* oli vaikea todennäköisesti siksi, että siinä on kaksi eri kysymystä yhdessä. Sen vuoksi tein neljästä kysymyksestä toiset muotoilut, kuten *Voinko asentaa asuntooni pyykinpesukoneen?* Tällä tavoin halusin varmistaa, että saan jokaisen chatbotin kohdalla tarpeeksi arvioitavaa tutkimusmateriaalia.

5 Kognitiivinen läpikäynti chatbottien vastauksista

Käyn seuraavaksi läpi löydöksiä jokaisen chatbotin läpijuoksusta. Ensin vuorossa oli HOASin *Helmi* chatbot. *Helmi* vastasi suurimpaan osaan kysymyksistä, mutta ei osannut vastata viiteen kysymykseen. Yhteen näistä vastaamattomista kysymyksistä se antoi kysymyksen aiheeseen liittyvän vastauksen, mutta ei suoraa vastausta. Sen vastausten tyyli oli kohtalaisen virallinen. Se myös usein antoi vastauksessaan linkin säätiön informaatiosivuille. Monesti nämä sivut sisälsivät vastauksen kysymykseen ymmärrettävämmässä muodossa.

KOASin *Kaija* chatbot vastasi myös suurimpaan osaan kysymyksistä, mutta se ei osannut vastata neljään kysymykseen. Sen vastausten sävy oli myös paikoittain virallinen, mutta yleensä ystävällinen. *Kaijan* vastaukset olivat todella pitkiä ja tärkein sisältö oli yleensä aivan vastauksen alussa.

TYSsin *Toivo* chatbot osasi vastata puoleen kysymyksistä eli seitsemään. Sen vastaukset olivat melko tiiviitä ja lyhyitä. Jokaisen vastauksen jälkeen se kysyi voisiko se auttaa jotenkin muuten.

DASsin *Das bot* ei osannut vastata viiteen kysymykseen. Senkin vastaustyyli oli kohtalaisen tiivis, mutta ei yhtä tiivis kuin *Toivon.* Usein sen vastausten alla oli pikavastauksia (*quick replies*), joita käyttäjä voi klikata saadakseen lisätietoa.

Joensuun Ellin *Elli* chatbot ei osannut vastata kolmeen kysymykseen. Yhteen näistä sen vastaus liittyi kysymyksen aiheeseen, mutta ei varsinaisesti vastannut kysymykseen. Kun sen vastauksessa oli linkki, sitä ei pystynyt klikkaamaan, vaan linkkiteksti olisi pitänyt manuaalisesti kopioida vastauksesta, mikä on hankalaa.

TOASsin *Tane* chatbot ei osannut vastata neljään kysymykseen. *Tanen* vastausten tyyli oli kohtalaisen epävirallinen ja se puhutteli suoraan käyttäjää ehkä useimmiten. Joissakin vastauksissa se kuitenkin vaihtoi puhuttelutyyliä kesken vastauksen, mikä oli hämmentävää.

Yleisesti vastauksissa oli paljon samankaltaisuuksia. Erityisesti samankaltaisuutta oli vastausten tyylissä ja sisällössä. Oikea käyttäjä ei välttämättä huomaisi näitä samankaltaisuuksia, koska hän keskustelisi todennäköisesti vain yhden chatbotin kanssa. Vastausten tyyli vaati opettelua, jotta vastauksen pystyi tulkitsemaan kunnolla. Monissa vastauksissa esimerkiksi viitattiin asukkaaseen ja asukkaan velvollisuuksiin, eikä puhuteltu käyttäjää suoraan. Sen takia oli ajoittain vaikea tulkita, kenelle tieto oli suunnattu. Tämä oli myös keskustelun etenemisen kannalta erikoista, koska kaikki kysymykset oli kuitenkin muotoiltu ensimmäiseen eikä kolmanteen persoonaan. Puhuttelutyylin takia vastauksia joutui joskus lukemaan useamman kerran. Ajan kanssa tyyli alkoi tuntua normaalimmalta, eikä siihen kiinnittänyt yhtä paljon huomiota. Kaikki chatbotit myös käyttivät paikoittain hyvin virallisia termejä. Esimerkiksi *Kaija* käytti termiä *päävuokralainen.* Vaikka tämän termin tarkoitus on kohtalaisen helposti pääteltävissä, voi se olla noviisikäyttäjälle hetkellisesti hämmentävä.

Chatbotit erosivat sen suhteen, kuinka paljon ne tarjosivat lisätietoja. Esimerkiksi *Helmin* vastauksissa oli monesti linkkejä säätiön nettisivuille. Sen sijaan *Kaija* ja *Tane* useammin tarjosivat pikavastauksia vastauksensa jälkeen. Joskus nämä pikavastaukset johdattivat säätiön nettisivuille, mutta välistä ne tarjosivat toisen vastauksen niiden seurauksena. *Kaija* myöskin vastasi paljon pidemmillä vastauksilla ja samalla tarjosi lisätietoja jo itse vastauksessa. Lisätietojen tarjoaminen helpottaa käyttäjää, koska silloin hänen ei tarvitse etsiä tai kysyä näistä tiedoista erikseen. On kuitenkin tärkeää rajata lisätiedon määrää, ettei käyttäjälle tarjota vain kaikkea aiheeseen liittyvää tietoa vaikkei sillä olisi hänelle varsinaisesti merkitystä.

6 Päätelmät

Tämän tutkimuksen tutkimuskysymyksenä oli kuinka käytettäviä asiakaspalvelun chatbottien vastaukset ovat. Vastaus on kohtalaisen käytettäviä. Oikeat vastaukset kysymyksiin sisälsivät

tarvittavat tiedot ja olivat aina ymmärrettäviä. Niissä myös ajoittain tarjottiin lisätietoa, mikä on erityisen hyödyllistä noviisikäyttäjälle. Samaan aikaan ne eivät olleet niin selkeitä, että ne olisi aina ymmärtänyt heti, vaan niitä piti tulkita. Esimerkiksi ohjeistavista osioista ei aina ymmärtänyt, missä järjestyksessä asiat olisi pitänyt toteuttaa. Yleisesti vastausten tyyliä voisi siis kehittää käytettävämmäksi. Vastausten tyyli oli monimutkainen, minkä vuoksi niiden tulkitseminen vaati opettelua. Tyyliä voisi parantaa muun muassa puhuttelemalla käyttäjää suoraan, tarkemmalla toiminnan askeleiden jaksotuksella ja rajaamalla tarjottua tietoa. On myös huomionarvoista, ettei yksikään chatboteista osannut vastata kaikkiin niille esitettyihin kysymyksiin.

Chatbotit eivät varsinaisesti ole uusi teknologia, mutta nykyään ne ovat yleistymässä jokapäiväisessä elämässä. Niiden käytettävyyttä voitaisiin tutkia enemmän, koska ne ovat uusi tapamme olla vuorovaikutuksessa yritysten kanssa. Vaikka tässä tutkimuksessa oli vain yksi arvioija, tuloksena oli monia käytettävyyshuomioita. Useammalla eri alan ammattilaisarvioijalla tehtynä voisi tällainen tutkimus tuottaa vielä enemmän huomioita. Tutkimukseen voisi myös sisällyttää laajempia tehtäväkokonaisuuksia, kuten asunnon hakeminen kokonaisuudessaan. Arvioija voisi suorittaa tällaisen tehtävän kokonaan chatbotin avulla, jolloin voisi paljastua vielä enemmän niiden käytettävyydestä.