

Designing AI chat-bots and User Assisted Systems in Business Applications

Creating Guidelines for a Graphical Interface

Master's thesis in Interaction Design & Technologies

LINNEA OLSSON & JENNY LU

Department of Computer Science and Engineering

CHALMERS UNIVERSITY OF TECHNOLOGY

Gothenburg, Sweden 2024

www.chalmers.se

MASTER'S THESIS 2024

Designing AI chat-bots and User Assisted Systems in Business Applications

Creating Guidelines for a Graphical Interface

JENNY LU

LINNEA OISSON



CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Computer Science and Engineering

Division of Interaction Design

CHALMERS UNIVERSITY OF TECHNOLOGY

Gothenburg, Sweden 2024

Designing AI chat-bots and User Assisted Systems in Business Applications
Creating Guidelines for a Graphical Interface
Linnea Olsson and Jenny Lu

© Jenny Lu and Linnea Olsson, 2024.

Supervisor: Pauline Belford, Department of Computer Science and Engineering
Industrial Advisors: Alexander Simola Bergsten, CANEA Partner Group AB
Examiner: Palle Dahlstedt, Department of Computer Science and Engineering

Master's Thesis 2024
Department of Computer Science and Engineering
Division of Interaction Design
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: CANEA Chat.

Typeset in L^AT_EX
Printed by Chalmers Reproservice
Gothenburg, Sweden 2024

Designing AI chat-bots and User Assisted Systems in Business Applications
Creating guidelines for a graphical interface
Jenny Lu, Linnea Olsson
Department of Computer Science and Engineering
Chalmers University of Technology and University of Gothenburg

Abstract

This thesis investigates the impact of graphical interfaces on AI chatbots and User Assisted Systems (UAS) within business applications, with a particular focus on designing interfaces for the CANEA ONE platform. The primary research questions revolve around optimizing customer acceptance, usage, and satisfaction of AI chatbots in business contexts, while also addressing key considerations and challenges in their design. By adopting a user-centered design approach, which includes design sprints and stakeholder engagement, the study integrates literature review, qualitative and quantitative data collection, and prototyping phases. The resulting high-fidelity prototype underwent rigorous user testing, serving as the foundation for the development of comprehensive guidelines. These guidelines successfully outline best practices for designing chatbots and UAS tailored specifically to business applications.

Keywords: AI Chatbot, Design Guidelines, User-Centered Design, Chatbot interface, User Assisted System.

Acknowledgements

In this thesis, we have benefited greatly from the knowledge, guidance, and insights provided by several helpful people.

Firstly, a big thank you to Pauline Belford, who supervised the thesis and offered unparalleled input and support throughout the entire project.

We also express our appreciation to Alexander Simola Bergsten, our supervisor from CANEA, for his essential role in the creation of this thesis and support during the project.

We are also grateful to everyone at CANEA who contributed to the project, especially those who shared their indispensable insights during our research and design phases. Thank you for your participation in interviews and workshops, which enriched the project with valuable reflections and knowledge.

We are also very grateful to our opponents Victor Cousin and Philip Lindström Rabia, who have provided us with invaluable insights and feedback on the project. You have our eternal thanks.

Lastly, we wish to express our gratitude to Chalmers University of Technology for providing us with an environment promoting learning and growth throughout these remarkable years.

Jenny Lu, Linnea Olsson, May 2024

List of Acronyms

Below is the list of acronyms that have been used throughout this thesis listed in alphabetical order:

| | |
|-----|---------------------------|
| AI | Artificial Intelligence |
| CA | Conversational Agent |
| CMS | Content Management System |
| UAS | User Assisted Systems |
| UCD | User Centered Design |

Contents

| | |
|--|-----------|
| List of Acronyms | ix |
| List of Figures | xv |
| 1 Introduction | 1 |
| 1.1 Aim & Research Question | 1 |
| 1.1.1 Research Question | 2 |
| 1.2 Stakeholders | 2 |
| 1.3 Limitations | 3 |
| 2 Background & Related Work | 4 |
| 2.1 CANEA & CANEA ONE | 4 |
| 2.1.1 Existing support | 5 |
| 2.1.2 CANEA Design System | 6 |
| 2.2 Types of Chatbots | 6 |
| 2.2.1 Social vs Task Oriented Interaction Style | 7 |
| 2.2.2 Why People Use Chatbots | 7 |
| 2.3 Similar Studies | 7 |
| 2.3.1 Survey on Social Characteristics in Chatbots | 8 |
| 2.3.2 A Critical Review of Chatbot Designs and Applications | 8 |
| 2.3.3 AI Chatbot for Customer Service | 8 |
| 2.3.4 Summary | 9 |
| 3 Theory | 10 |
| 3.1 Accessible Design Guidelines | 10 |
| 3.2 Graphical User Interface Guidelines | 10 |
| 3.2.1 5 Principles of Visual Design in UX | 11 |
| 3.2.2 Principles of User Interface Design | 12 |
| 3.2.3 Accessible Graphical User Interfaces | 14 |
| 3.3 Customer-Service Chat Design Guidelines | 14 |
| 3.4 Gender-bias in Chatbots | 16 |
| 3.5 Ethics of Artificial Intelligence | 17 |
| 3.6 Advanced User Assistance Systems | 18 |
| 3.6.1 User Assisted System in Business Information Visualization | 19 |
| 3.6.2 Notification Systems | 19 |
| 3.7 Autocompletion in Generative AI Interaction | 19 |
| 3.8 Design for AI | 20 |

| | | |
|----------|---|-----------|
| 3.9 | Trust In AI Chatbots | 21 |
| 3.10 | User-Centered Design | 22 |
| 4 | Methods | 23 |
| 4.1 | Design Sprint | 23 |
| 4.2 | Literature Review | 23 |
| 4.3 | Questionnaire | 24 |
| 4.4 | Interviews | 24 |
| 4.5 | Personas and Storyboards | 25 |
| 4.6 | Ideation | 25 |
| 4.6.1 | Crazy-eights | 25 |
| 4.6.2 | Brain-writing | 25 |
| 4.6.3 | Co-design Workshop | 26 |
| 4.7 | Prototype | 26 |
| 4.7.1 | Low & High Fidelity Prototype | 26 |
| 4.7.2 | Figma | 27 |
| 4.7.3 | Wireframing | 27 |
| 4.8 | Evaluation & Analysis | 27 |
| 4.8.1 | Heuristic Evaluation | 27 |
| 4.8.2 | User Testing | 28 |
| 4.8.3 | Affinity Diagram | 28 |
| 5 | Process and Execution | 29 |
| 5.1 | Literature Review | 29 |
| 5.2 | Questionnaires | 29 |
| 5.3 | Interviews | 30 |
| 5.4 | Affinity Diagram Interviews | 30 |
| 5.4.1 | Interview Findings | 31 |
| 5.5 | Persona and Storyboard | 33 |
| 5.5.1 | Persona: Per | 33 |
| 5.5.2 | Persona: Alice | 33 |
| 5.5.3 | Storyboard: Configuration | 33 |
| 5.6 | Crazy-eights | 34 |
| 5.7 | Brain-writing | 34 |
| 5.8 | Low Fidelity Prototypes | 35 |
| 5.9 | Co-design Workshop | 35 |
| 5.9.1 | Co-Design Workshop Findings | 35 |
| 5.10 | High Fidelity Prototypes | 36 |
| 5.10.1 | Heuristic Evaluation | 38 |
| 5.11 | Functional Prototype | 39 |
| 5.11.1 | Prototype Interface | 39 |
| 5.11.2 | UAS Interface | 44 |
| 5.12 | User Testing | 45 |
| 5.13 | Affinity Diagram User Tests | 47 |
| 5.13.1 | User Tests Findings | 47 |
| 6 | Results | 51 |

| | | |
|----------|---|------------|
| 6.1 | Problem Statement | 51 |
| 6.2 | Guidelines for AI Chatbots | 51 |
| 6.2.1 | Tool-tips | 51 |
| 6.2.2 | AI Personality | 52 |
| 6.2.3 | Usage of Icons | 52 |
| 6.2.4 | Minimalist Design | 52 |
| 6.2.5 | Size | 52 |
| 6.2.6 | Small-Interface | 53 |
| 6.2.7 | Perception of Control | 53 |
| 6.2.8 | Consistency | 53 |
| 6.2.9 | Complement Existing Solutions | 53 |
| 6.2.10 | Location of Chatbot | 53 |
| 6.2.11 | Chatbot Response | 54 |
| 6.2.12 | Anticipation | 54 |
| 6.2.13 | Accessible Design | 54 |
| 6.2.14 | Differentiating Chatbot and User Messages | 54 |
| 6.3 | Guidelines for User Assisted System | 54 |
| 6.3.1 | Accessible Design | 54 |
| 6.3.2 | Do not Disturb Workflow | 55 |
| 6.3.3 | Design for Easy Deactivation | 55 |
| 6.3.4 | Accepting and Canceling | 55 |
| 6.4 | Prototype Improvements | 55 |
| 7 | Discussion | 57 |
| 7.1 | Results | 57 |
| 7.1.1 | AI Chatbot Trust | 57 |
| 7.1.2 | System Complexity | 58 |
| 7.1.3 | Chatbot Humanness & Gender | 58 |
| 7.2 | Methods | 59 |
| 7.2.1 | Questionnaires | 59 |
| 7.2.2 | Interviews | 59 |
| 7.2.3 | User Group | 59 |
| 7.2.4 | User testing | 60 |
| 7.3 | Future Development | 61 |
| 8 | Conclusion | 64 |
| | Bibliography | 67 |
| A | Timeline | I |
| A.1 | Time plan | I |
| B | CANEA ONE | IV |
| C | Questionnaire | VII |
| D | Interview Questions | X |

Contents

| | |
|-----------------------------|------------|
| E User Testing | XII |
| F Ideation Crazy 8's | XIV |
| G Figma Wire-frames | XIX |

List of Figures

| | | |
|------|--|----|
| 2.1 | Example of a Dashboard in CANEA ONE | 5 |
| 4.1 | Design Sprint | 23 |
| 5.1 | Example of the cluster AI in CANEA ONE | 31 |
| 5.2 | Storyboard of Per | 34 |
| 5.3 | Examples of low-fi wireframes | 35 |
| 5.4 | Chatbot 1 | 36 |
| 5.5 | Chatbot 2 | 37 |
| 5.6 | UAS 1 | 37 |
| 5.7 | UAS 2 | 38 |
| 5.8 | Chatbot Menu Open | 40 |
| 5.9 | New Chat | 40 |
| 5.10 | Chatbot Typing | 41 |
| 5.11 | Chatbot Messages | 41 |
| 5.12 | Copy Message | 42 |
| 5.13 | Edit Message | 42 |
| 5.16 | Minimized Chatbot Messages | 42 |
| 5.14 | Minimized Chatbot | 43 |
| 5.15 | Minimized Chatbot History | 43 |
| 5.17 | Chatbot Closed | 44 |
| 5.18 | UAS | 44 |
| 5.19 | UAS Off | 45 |
| 5.20 | UAS Message | 45 |
| 5.21 | Tasks on a whiteboard | 46 |
| 5.22 | Example of the cluster Bot or Not in Miro | 48 |
| 6.1 | New Icon for New Chat | 56 |
| 6.2 | Tool-tips | 56 |
| B.1 | Strategy | IV |
| B.2 | Process | V |
| B.3 | Projects | V |
| B.4 | Workflow | VI |
| B.5 | Documents | VI |
| F.1 | Chatbot interface wire-frame sketches developed during crazy 8's part 1XIV | |
| F.2 | Chatbot interface wire-frame sketches developed during crazy 8's part 2XV | |

| | | |
|-----|---|-------|
| F.3 | Chatbot interface wire-frame sketches developed during crazy 8's part 3 | XVI |
| F.4 | Chatbot interface wire-frame sketches developed during crazy 8's part 4 | XVII |
| F.5 | Chatbot interface wire-frame sketches developed during crazy 8's part 5 | XVIII |
| G.1 | Chatbot wire-frame created in Figma where the chatbot is integrated in the stream functionality (comments) of CANEA ONE | XIX |
| G.2 | Chatbot wire-frame created in Figma where the chatbot appears like a modal and is re-sizeable and movable | XX |
| G.3 | UAS wire-frame created in Figma where the notification appears like a modal | XX |
| G.4 | UAS wire-frame created in Figma where the user drags a box over the functionality they want explained | XXI |

1

Introduction

The adoption of artificial intelligence (AI) in organizations is rapidly increasing. A recent report indicates that 85% of surveyed organizations are either contemplating or already employing AI-based applications [1]. Concurrently, the use of AI has greatly increased in recent years, particularly with the emergence of generative AI in the form of chatGPT and other large language models. An AI chatbot is an artificial intelligence application utilizing Natural Language Processing (NLP) to comprehend and facilitate communication between a human user and a machine. This technology enables the chatbot to interpret and respond to natural language inputs, allowing users to engage in interactive conversations with the AI system. The primary objective of an AI chatbot is to simulate human-like interactions, providing users with a seamless and responsive conversational experience. AI chatbots can be used in different areas such as customer support, e-commerce and education. Chatbots can for example be used for answering frequently asked questions and automating mundane tasks such as gathering information for orders, inquiries, and appointments, or executing various other functions through seamless integration with business software and platforms [2]. They can also consider intricate user feedback and give appropriate answers and product recommendations [3]. According to Okonkwo et al. chatbot technology can improve teaching and learning in education [4]. With these new tools it is crucial to consider how users interact with them and how to craft the desired user experience with AI chatbots.

1.1 Aim & Research Question

The project aims to develop an interface for an AI chatbot designed to assist users of the CANEA ONE application. CANEA ONE is a business application that digitizes work within an organisation. CANEA ONE has five different modules; Strategies, Processes, Projects, Cases and Documents which make up the application. While many chatbots exist, CANEA envisions a more sophisticated and advanced chatbot that surpasses existing applications in complexity and richness.

Most existing chatbots fall into simplistic categories, offering predefined responses or limited functionalities, further discussed in chapter 2. In contrast, the CANEA chatbot requires a higher level of sophistication. It needs to serve as an interactive assistant capable of engaging in meaningful conversations, providing tailored feedback, and offering nuanced recommendations based on users' ongoing tasks within the complex CANEA ONE program. This complexity makes designing an effective

interface more challenging compared to typical chatbots.

The AI encompasses two requested functions: a chatbot functionality for answering user questions about program usage and locating specific functions, and an advanced user assistance system capable of recognizing the user's current tasks within the program and providing relevant suggestions. This project's chatbot is more intricate, requiring an understanding of the context in which the user is working. The AI assistant also needs to adapt to various user types, including beginners, intermediate, and advanced users, as each category has distinct needs, and the AI should be able to cater to their specific requirements.

The expected outcome of the project is a set of guidelines for the graphical interface and desired behaviour of an AI chatbot. Additionally, a prototype developed based on the guidelines of the AI chatbot visualising the interface will be presented.

1.1.1 Research Question

The aim of the project is to create the interface of an AI chat-bot that will assist users of the program (CANEA ONE). The two research questions the thesis aims to investigate are as follows:

- How can we design the user interface of AI Chat-bots and user assisted systems in the field of Business Applications to improve customer acceptance, use, and satisfaction?
- What are the key considerations and challenges in designing AI Chat-bots and user assisted systems in the field of Business Applications?

1.2 Stakeholders

The thesis identifies a range of stakeholders, listed as follows:

- **Thesis Authors**
Linnea Olsson & Jenny Lu.
- **Chalmers University**
The University overseeing the thesis.
- **CANEA**
The industrial partners of this project who are providing guidance. CANEA may be interested in implementing the guidelines for an AI chatbot in future projects.
- **Customers of CANEA**
The end-users, being the primary consumers of the final product, play a pivotal role in its utilization.

1.3 Limitations

In the project, a functioning AI will not be developed since the focus will be on creating an interface that is intuitive and user-friendly. Instead, an API of a chatbot, such as Chat-GPT, will be utilized to develop a functional prototype that can be used for user tests. Consequently, controlling and fine-tuning the chatbot's responses will have limitations and won't be a focal point in this project.

2

Background & Related Work

This thesis revolves around developing the guidelines for an AI chatbot for the application CANEA ONE at CANEA. In this chapter CANEA and its application CANEA ONE will be described followed by related works within the area of AI chatbots.

2.1 CANEA & CANEA ONE

CANEA [5] is a company that offers management consulting services, IT solutions, and training while CANEA ONE is their business application. The application aims to digitise work within an organisation, enabling the control of work (i.e., ensuring that work is carried out correctly, structured, and goal-oriented) at various levels of an organisation. CANEA ONE has a total of 5 modules: Strategy, Project, Document, Workflow and Process, see appendix B. An example of how the start page can appear in CANEA ONE is depicted in Figure 2.1.

- **Strategy**

The strategy module makes it possible for organisations to establish a unique common thread, seamlessly connecting strategic endeavors to daily operations [6]. Companies can gain a clear vision by setting goals and planning how to achieve them. The goals can then be broken down into different perspectives, such as customers and finance. Users can decide the limits of the goals themselves, and these goals may vary from month to month. With this framework in place, users can easily track and monitor their goals on a regular basis.

- **Project**

The project module is a tool that supports all types of projects and parties [7]. This may include management, resource owners, clients, project managers and project participants. The module includes features that lets the user visualise data to analyse aspects such as budget vs proposed budget. In the project module users are also able to connect workflows, create tasks, prioritise tasks and manage risks.

- **Document**

The document module simplifies document management by eliminating complex folder structures, making it user-friendly for easy navigation and information retrieval [8]. It efficiently handles version control, allowing users to view documents without downloading them. Users can easily select a document

type to initiate a straightforward workflow.

- **Workflow**

The workflow module is designed to automate, ensure quality, and expedite administrative workflows [9]. The customers can define and set up cases according to their own needs. This results in a real-time overview of workflow processes and diagrams which supports decision making.

- **Process**

The process module supports companies with communicating business processes through an accessible and user-friendly graphical web interface [10]. Processes can be broken down through tree diagrams which helps companies both visualise the process but also to see the complete picture.

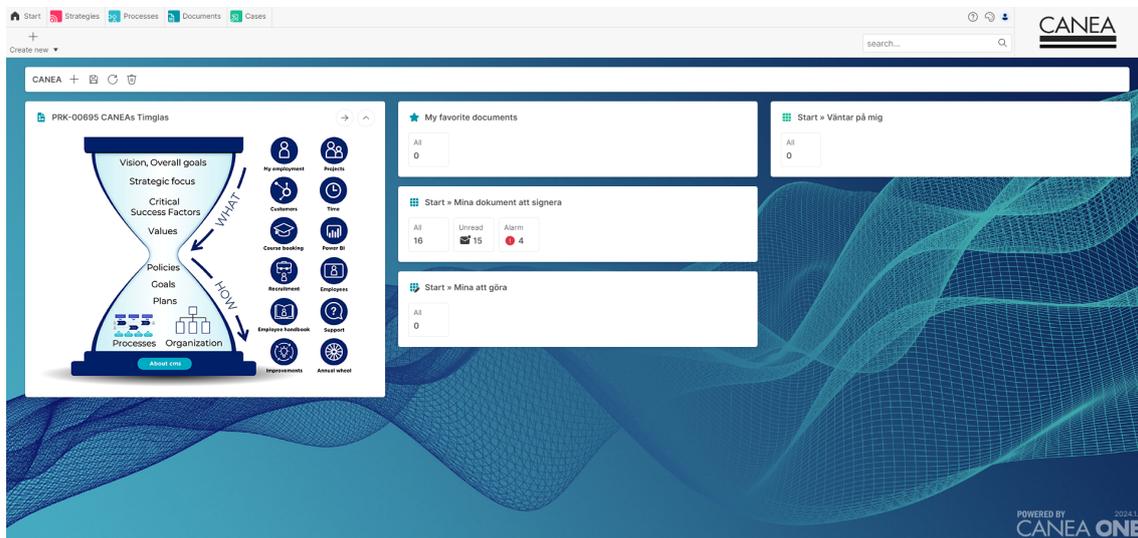


Figure 2.1: Example of a Dashboard in CANEA ONE

Together, these modules create a system designed for effective business management. Some features that are common in many modules, but not all, include the display of a workflow, members, work roles, and a dashboard containing widgets.

Additionally, senior management can have an overview of the entire organisation and monitor activities where services/products are developed and delivered. The system also eliminates the issue of different individuals working with various document versions since they are version-controlled and synchronised among all users. With the system, customers (other companies/organisations) can gather and organise their documents, cases, projects, processes, and strategic goals in one place.

2.1.1 Existing support

All CANEA ONE users currently have access to a manual within the system, which contains documentation on functions and system information. This documentation

is important, as the application involves numerous functions that users may find difficult to memorise or may not have learned yet. In addition to the manual, superusers responsible for configuring the system can seek assistance from CANEA support (a department in CANEA) if needed. However, it is expected that users consult the manual before reaching out to CANEA for assistance.

2.1.2 CANEA Design System

According to Fessenden [11], a design system can be defined as a comprehensive collection of standards intended to manage design on a large scale through the utilisation of reusable components and patterns. One of the benefits of using a design system is minimising repetition while establishing a common language and ensuring visual coherence across various pages and channels.

CANEA has developed its own design system, which includes multiple color schemes for various occasions and guidelines for the use of typography, icons, logos, and patterns. For example all icons used in CANEA ONE are fluent UI by Microsoft. Furthermore, all products intended to represent the company must adhere to these guidelines. Therefore, any design solutions created during the project will strictly follow the CANEA design system.

2.2 Types of Chatbots

According to Gupta et al. [12] there are three different types of chatbots; Menu/Button-Based Chatbots, Keyword Recognition-Based Chatbots, and Contextual Chatbots.

The most widely utilised and straightforward chatbots in today's market are menu-based chatbots [12], presented in the form of buttons and top-down menus. These chatbots operate on decision tree principles, where users navigate through decisions to reach the final answers. Users are guided to make selections, delving deeper to find the suitable response from the AI. However, these menu-based chatbots exhibit comparatively slower performance and may not be entirely reliable in delivering the desired answers.

Keyword recognition-based chatbots identify specific keywords to generate the desired outcome [12]. They analyze user inputs and respond accordingly. Leveraging AI technology and a tailored list of keywords, the bot utilises algorithms to determine an appropriate response. However, these chatbots may encounter challenges when there are keyword redundancies across multiple related questions. For instance, if a user inquires about setting up auto-login authentication on their phone, the bot may rely on keywords such as 'auto' and 'login' to determine the most fitting response.

Contextual chatbots represent the most advanced bots currently available in the market [12]. They leverage technologies such as Machine Learning and Artificial Intelligence, including voice recognition and speech-to-text conversion algorithms,

to comprehend user sentiments. The fundamental principle behind these bots is to discern the user's intentions and provide thoughtful responses by analyzing patterns within the database. These bots continually learn and evolve through various experiences over time. An example can be found in a food delivery application, where the bot considers factors such as previous order history, user payment options, and delivery address stored in the database. By analyzing the user's perspective, these chatbots offer recommendations based on past orders and user preferences. This project aims to design a contextual chatbot capable of discerning the user's current activity within the application to provide thoughtful responses.

2.2.1 Social vs Task Oriented Interaction Style

When interacting with humans chatbots can have either a social or task oriented interaction style [13]. Social chatbots have an informal tone [13] and focus on chatting with the user and making them engaged [14]. Task oriented chatbots however keep a more formal tone and engage on only on-task conversations to achieve useful goals [13]. The chatbot to be designed in this thesis will adopt a primarily task-oriented interaction style, as its primary objective is to aid users with tasks within the application. Moreover, the chatbot should refrain from responding to inquiries unrelated to the application, limiting users' ability to engage in social interactions with it.

2.2.2 Why People Use Chatbots

There are various reasons why individuals choose to interact with chatbots. Melián et al. [15] suggest that users primarily engage with chatbots based on their anticipation of effective performance. Similarly, Brandtzaeg and Følstad [16] found that users emphasize the ease of use, speed, and convenience as key factors contributing to effectiveness. Additionally, people use chatbots for purposes such as entertainment, social and relational interactions, as well as novel experiences.

However, in a different study Brandtzaeg and Følstad argues that chatbot may not necessarily be looking for social experiences with family and friends which they can find elsewhere [17]. They are instead looking for a place where they can share information without the fear of judgement. There are many reasons why users of CANEA ONE would use an AI chatbot within the application. For instance, the chatbot could enhance work efficiency by providing solutions quickly, and it could also optimize resource allocation, serving as the initial point of contact for issue resolution; if unsuccessful, users could then reach out to CANEA support.

2.3 Similar Studies

There have been several use cases of AI chatbots investigated in various studies and a common reason for implementing these chatbots is the convenience, effectiveness, accessibility and ability to lower costs. In this section diverse papers related to the topic of AI chatbots are presented.

2.3.1 Survey on Social Characteristics in Chatbots

Chaves and Gerosa identify the characteristics of chatbots in different fields [18]. In the realm of business chatbots, two key characteristics were identified: proactivity and personalisation. Proactivity, in this context, refers to the system's capacity to autonomously take action, thereby reducing the user's effort in completing tasks. This proactive behavior allows the chatbot to initiate conversations, propose new topics, offer additional information, and formulate follow-up questions. Proactivity not only sustains the conversation but also enhances engagement. On the other hand, personalisation involves the chatbot's ability to tailor its functionality, interface, and content to heighten personal relevance for the user. By adapting to individual preferences, personalisation fosters a sense of rapport and cooperation, ultimately leading to increased engagement with chatbots.

2.3.2 A Critical Review of Chatbot Designs and Applications

Although there have been numerous papers discussing the success of AI chatbots in various fields, Luo et al. [19] argue that there is a lack of an up-to-date critical review that considers both the various computational approaches used to develop AI chatbots and their usability across various business sectors. There are three main issues with existing literature reviews on chatbots. There is a lack of classification on the chatbots that naturally connects chatbot technologies with real life applications. Recent studies frequently categorise chatbots into either retrieval-based or generation-based systems based on their mechanism of response generation.

When considering chatbots, the acceptance of chatbot systems is influenced by various factors beyond task performance [19]. These factors include perceived ease of use, perceived usefulness, aesthetics, perceived enjoyment, and anthropomorphism, all of which are intricately connected to interface design. However, many existing chatbots do not offer adequate information about their interactive user interfaces, and there is a lack of documentation in research literature regarding these aspects of chatbot development. The author concludes by recommending further studies that focus on interface design and its evaluation.

2.3.3 AI Chatbot for Customer Service

In the paper by Ngai et al. [20] an AI was utilised for customer service in a clothing company to give advice and suggestions and also give information about price and stock. If the chatbot was unable to handle the questions it received, the customer service query was handed over to human staff. The results suggested that the implementation of AI led to improvements in customer service efficiency and satisfaction. The instantaneous responses provided by the AI chatbot minimized complaints and elevated customer satisfaction with the company's customer service. There was a noteworthy decrease in human effort, with the chatbot's accuracy remaining at 100% when compared to human staff. As a result, the evaluation findings suggested that the implemented AI chatbot can efficiently improve the handling of customer

queries, thereby contributing to effective customer relationship management.

The significance of identifying a clear use case and objective was emphasized in the paper [20]. Concentrating on a well-defined use case with specific objectives aids in expectation management and expedites returns. Involving users in the early stage was underscored as a crucial aspect. Engaging users in the initial phases enhances their comprehension and acceptance of emerging technologies such as AI. Involving humans in the loop was also highlighted, emphasizing the importance of maintaining a balanced division of labor between humans and machines to optimise results.

2.3.4 Summary

To summarise this chapter the thesis is aimed at developing guidelines for an AI chatbot within the business application CANEA ONE. CANEA ONE consists of five modules each adding different functionality to streamline business management. The application currently relies on a manual as well as CANEA support.

There are primarily three types of chatbots in use: menu-based, keyword-based, and contextual-based, each with its own set of advantages and limitations. However, in terms of technical sophistication, contextual-based chatbots currently represent the most advanced type available in the market.

In addition to type, chatbots can also have an interaction style, either social or task-oriented. Social-oriented chatbots primarily focus on engaging users in conversation with an informal tone, while task-oriented chatbots maintain a more professional tone. The primary goal of task-oriented chatbots is to assist users in completing their tasks without addressing unrelated questions.

Finally, three cases were observed of the implementation of AI chatbots across various sectors, highlighting their convenience, effectiveness, accessibility, and cost-saving potential. According to Chaves and Gerosa [18], chatbots can proactivity and personalisation can be used reduce the users effort in completing tasks and to tailor the experience leading to increased engagement. Luo et al. [19] on the other hand highlights the fact that while there are numerous papers emphasising the success of AI chatbots there is a need to further study the interface design and its evaluation. Finally, Ngai et al. [20] discusses how AI was utilised in a clothing company. The AI chatbot handled simple questions and if it was unable to do so the query was send to a human. This resulted in a decrease in human effort while the chatbot had a accuracy at 100% in comparison to human staff. The study highlights the importance maintaining a balance in labour between human and machines to optimise results.

3

Theory

This chapter delves into the theoretical foundation of the project, aiming to provide a comprehensive exploration of the current research area.

3.1 Accessible Design Guidelines

The 7 Principles of Universal Design are guidelines for creating designs that are accessible to all users [21]. The principles include Equitable Use, Flexibility in Use, Simple and Intuitive Use, Perceptible Information, Tolerance for Error, Low Physical Effort, and Size and Space for Approach and Use. These design principles focus on creating designs that cater to individuals with diverse abilities. Achieving this involves allowing various modes of use, such as utilizing the mouse, keyboard, and voice commands to navigate the interface. By supporting different modes of use, the design enables users to choose their preferred method of navigation, thereby enhancing accessibility for a broader audience. The distinction between universal design and assistive design lies in their target users. Universal design aims to cater to as many users as possible, encompassing various abilities. In contrast, assistive design is tailored to meet the specific needs of users with disabilities. The advantage of following universal design guidelines is that they prioritize creating designs usable by all users, rather than focusing on a specific user group. An accessible design not only benefits users with disabilities but is also generally more user-friendly for those without disabilities.

3.2 Graphical User Interface Guidelines

A Graphical User Interface (GUI) provides an easy way for humans to interact and communicate with computers [22]. A well-designed GUI eliminates barriers to communication with the computer system, allowing users to focus directly on solving the task at hand [23]. Common GUI types include widget layouts, forms, hypertext, toolbars, windows, and menu systems [24]. The design of these GUIs significantly influences the system's usability, usefulness, learnability, and overall appeal, ultimately playing a pivotal role in the system's success and acceptance. What defines a good GUI? According to [25], the question can be answered by one simple sentence: "*Good design should be functionally effective and aesthetically pleasing*". There have been several attempts of creating guidelines for a good GUI and some are going to be presented below.

3.2.1 5 Principles of Visual Design in UX

When looking at designs users can usually tell if its appealing or flawed but they may find it difficult to explain with words why. To explain why designs may appear more visually appealing Nielsen Normal Group has defined 5 principles of visual design that impacts the UX [26]. Graphics that take advantage of these principle can help increase the usability of a design. These principles are useful when designing any interface and will also be used in this project to create a design that is easy to use.

1. Scale

The concept of scale involves utilizing relative size to convey significance and hierarchy within a composition. A design should not use more than three different sizes. Having a diversity of sizes creates variation within the layout and it is important to make sure that the most important elements are the largest.

2. Visual Hierarchy

The concept of visual hierarchy involves directing the viewer's gaze across the page, prioritizing attention to various design elements based on their importance. Visual hierarchy can be achieved by incorporating differences in scale, value, color, spacing, placement, and various other cues.

3. Balance

The principle of balance refers to achieving a harmonious arrangement or proportion of design elements. Balance is achieved when there is an evenly distributed (though not necessarily symmetrical) visual signal on both sides of an imaginary axis traversing the center of the screen. While this axis is typically vertical, it can also be horizontal.

4. Contrast

The concept of contrast involves placing visually distinct elements side by side to emphasize their differences, indicating distinctions such as belonging to different categories, serving different functions, or behaving differently. In other words contrast makes it visible to the eye that there is a notable difference(colour or size etc) between two objects.

5. Gestalt Principles

Gestalt principles were formulated in the early twentieth century by Gestalt psychologists and pertain to how humans comprehend images. Humans simplify and structure intricate images, consisting of numerous elements, by subconsciously arranging these parts into an organized system that forms a cohesive whole, rather than interpreting them as separate elements. The Gestalt principles consist of similarity, continuation, closure, proximity, common region, figure/ground, and symmetry and order. In the context of UX, proximity is especially significant. Proximity indicates that visually close items are perceived as belonging to the same group.

There are numerous benefits to adhering to design principles beyond achieving a

more aesthetic design. Some of these benefits include:

- **Increased usability**
Implementing visual-design principles, like the golden ratio, enhances both aesthetics and usability, contributing to improved readability and task success rates, especially when combined with strong interaction design.
- **Eliciting emotion and delight**
Beautiful designs not only evoke positive emotions but also contribute to the aesthetic–usability effect, suggesting users may overlook minor usability issues. Following good visual design principles is essential for creating UIs that look good and in turn enhance user experience.
- **Enhancing brand perception**
A strong design system attracts interest in the product, builds user trust and accurately represents and strengthens the brand.

3.2.2 Principles of User Interface Design

Over the years, specialists in information and communication technologies have outlined heuristics essential for enhancing the usability of user interfaces [27]. Noteworthy among them are Donald Norman, Ben Schneiderman, Jacob Nielsen, and Bruce Tognazzini, whose research and expertise form the foundation for numerous works. When examining the rule sets presented in each specialists findings, it is evident that they share commonalities, primarily emphasizing attributes such as appearance, consistency, feedback, efficiency, security, and error prevention. Utilizing these criteria, Nacheva [27] formulated a set of rules intended to facilitate the development of user-friendly interfaces, which are presented below. The outcome of the project is a prototype for the interface of an AI chatbot and UAS. Taking these principles into consideration during the design process will help create an interface that is pleasant to use.

- **Appearance**
 - **Sensory comfort**
Ensuring sensory comfort in software design involves addressing the visual and auditory perceptions of users, with considerations such as selecting suitable color combinations, ensuring label readability, and conducting specialized tests for individuals with specific needs like dyschromatopsia.
 - **Minimalism**
The beauty of design lies within simplicity. Excessive use of decorations and colours leads to discomfort and frustrations.
 - **Aesthetics**
Minimal design does not equal plain design, instead, it emphasizes the importance of selecting appropriate colors, fonts, and graphic elements,

particularly for visually-inclined users. Aesthetic design is crucial for fostering user satisfaction.

- **Intuitiveness**

Ensuring adherence to platform-specific terminology and visual concepts, as well as affordance, is crucial for intuitive design. Intuitive products reduce discoverability issues, although achieving this in complex systems remains a significant challenge.

- **Structuring interface objects**

Arranging interface elements appropriately can be guided by Fitt's Law, which states that the time it takes to reach a target is influenced by both the size and distance of the target [28].

- **Consistency**

- **Standardization**

Adhering to established standards such as ISO, IEC, and IEEE, along with compliance with platform-specific technological requirements. This adheres to commands, terminology, and icon actions

- **Conventions imposed in the real world**

Real-world conventions, linking physical objects and pre-existing terms predating technology, promote seamless knowledge continuity across diverse fields, aiding user adoption.

- **Mental models of users**

If the design aligns with users' preconceptions, it facilitates easier orientation for them.

- **Feedback**

- **Supporting information**

Add active windows in the shape of pointers or advice to new features.

- **Messages**

Present relevant messages that guide users during actions.

- **Progress of the work performed by the system**

Showing the status of the system.

- **Efficiency**

- **Ease of work**

The user needs to accomplish their specific goals quickly and easily.

- **Sense of freedom**

It's recommended to enable users to turn off or restart specified options. This results in a design that adapts to temporary needs.

- **Implement cognitive resources**
It is crucial to consider how much cognitive effort consumers need to invest in task execution. The technology should aim to minimize cognitive processes and reduce the load on working memory.
- **Multivariate approach to achieve the objectives**
Implementing tasks in various ways enables accessibility for diverse user groups, including beginners, intermediate users, advanced users, and those with special needs.
- **Security and prevention of errors**
Allowing users to perform a task in various ways helps identify errors and find solutions in undesired situations.

3.2.3 Accessible Graphical User Interfaces

Web Content Accessibility Guidelines (WCAG) involves a large set of recommendations aimed at enhancing the accessibility of web content [29]. Adhering to these guidelines expands content accessibility to a broader spectrum of individuals with disabilities, including those with blindness and low vision, deafness and hearing loss, limited movement, speech disabilities, photosensitivity, and various combinations of these conditions. While providing some accommodations for learning disabilities and cognitive limitations, these guidelines may not address every user need for individuals with these disabilities. Moreover, following these guidelines often results in improved usability for users in general.

While the Web Content Accessibility Guidelines (WCAG) provide a positive direction, researchers Sullivan and Matson [30] discovered that when evaluating content accessibility on the web as a continuous scale (not just as a yes/no or accessible/inaccessible), 29 out of the 50 most popular websites were still categorized as inaccessible. In other words, a significant portion of popular websites did not meet the criteria for full accessibility when assessed on a more nuanced and continuous scale. This indicates that even though the WCAG have existed since 1999 and there are laws that require their application there have been limited consequences or advancements in the field [31]. Since CANEA ONE is a web application the interfaces designed should adhere to the WCAG as much as possible to ensure that it is accessible to all of its users.

3.3 Customer-Service Chat Design Guidelines

Nielsen Norman Group has outlined 20 design guidelines for customer service chat [32]. Since we are creating a chatbot to assist users with CANEA ONE, we can take inspiration from some of these guidelines, which may also be applicable to an AI chatbot. They are listed below.

- **Keep users updated on how fast they will receive a response**
Users appreciate when there is an indication of when the answer is coming,

for example, if there is an "agent is typing" message.

- **Use a Separate Chat Window**
Having the chat in a separate window enables users to view other information simultaneously while reading the chat, making it easier to refer back to information and details within the application.
- **Visually Differentiate the Messages Coming from Different Participants in the Chat**
Assigning a unique color to each participant makes it easier to distinguish which message comes from whom. Making it easier for the user to navigate the chat and previous messages.
- **Prepare for Possible Interruptions**
Empower users with the ability to decide whether they want to continue the conversation or not. Additionally, make it easy for the user to pick up a paused chat session. The system should preserve both the context and the progress, eliminating the need for the user to re-enter questions or initiate the conversation anew.
- **Don't Make Users Type Their Question Multiple Times**
Making users type the same message several times frustrates them and is a waste of their time. The system should remember the previously inputted information.
- **Allow Users to Save a Transcript of the Session**
Saving the information in the chat can be helpful to users if they want to revisit it later.
- **Allow People to Upload Documents During the Chat Session**
Allowing users to submit documents can make it easier for them to describe their problem.
- **The Agent Should Not Start the Conversation with "How Are You"**
Although it is important for the chat agent to be polite, emphasizing efficiency is crucial. Polite but generic greetings such as "How are you today" or a simple "Hello" may lead to unnecessary delays for some users. A more effective way to start a chat is by saying, "Hello! How can I help you today?"
- **Don't Overuse Canned Responses**
While a small number of canned responses are acceptable, an excess of them can raise suspicion and make users think they are dealing with a bot. Users typically favor a conversational and informal tone that resembles communication with a real person.
- **Provide Users with Specific, Detailed Answers**
When users initiate a conversation, it suggests they likely need information that wasn't readily available to them elsewhere. Hence, it is essential to offer

precise and customized responses that cater to their specific needs.

3.4 Gender-bias in Chatbots

According to Zlotowski et al. [33] Anthropomorphism refers to the human tendency to perceive human-like shapes in the surrounding environment. However, anthropomorphism involves more than just making behavioral or dispositional assumptions about a nonhuman agent. It involves attributing human form or a human mind to the agent [34]. Since chatbots share many common traits with humans such as the utilization of natural language and conversational abilities this makes them susceptible to anthropomorphism from those they interact with. Moreover, the anthropomorphism of chatbots also tends to be accompanied by a bias of feminisation. In addition to their names, voices, or avatars, they complete tasks historically linked to women's labour[35].

Several studies have shown that using gendered cues in chatbot design can have both positive and negative outcomes [36]. A study by UNESCO highlights the negative consequences of having gendered design in specifically voice-assisted conversational assistants (CA) [37]. Most CAs are female by default which can contribute to the reinforcement of gender stereotypes. In addition to the study of voice assisted CAs, several studies have shown that the gender of the chatbot can be embodied without a spoken voice though for example an avatar.

Previous research has indicated that individuals tend to prefer female bots over male ones due to their perceived warmth and inclination to display emotions [38]. This preference is seen as advantageous, as emotions are considered inherently human qualities often lacking in machines, thus instilling AI chatbots with a sense of humanity [39]. Furthermore, studies suggest that social machines should bear resemblance to humans to enhance acceptance and facilitate smoother interactions between machines and humans [40].

Gender design can also have an impact on the degree of user trust in a chatbot. A study conducted by McDonnell et al. (2019) [41] indicates that the perceived trustworthiness, expertise, and intelligence of a chatbot are influenced by a chatbot's gender. This was demonstrated in the context of mechanics where users preferred a male version of the chatbot. This phenomenon may arise from individuals' tendency to categorize subjects based on gender, assuming that women are more knowledgeable about feminine topics and men about masculine topics [42].

Furthermore, assigning a gender to an chatbot could be harmful because users might tend to evaluate the success of a system according to their own biases and emotional attachment to the agent, rather than the actual performance of the system [43].

To not further reinforce the current gender bias towards female chatbots the design of the chatbot in this thesis should aim towards being as gender neutral as possible. This can be done by not giving any clues regarding the gender in the forms of a

name, avatar or any other possible factor.

3.5 Ethics of Artificial Intelligence

Leslie discusses the potential harms of AI [44]. These points are listed below. It is important to always consider ethics when creating a new design to ensure that the design is inclusive, particularly in the case of AI. Without careful consideration, individuals may receive incorrect information and be misled, which goes against the intended purpose of AI. Therefore these potential harms needs to be carefully considered in the process of creating the AI chatbot and UAS interface.

- **Bias and Discrimination**

AI is based on information from the existing structures in society. This leads to risks such as reproducing, reinforcing, or amplifying the marginalization, inequalities, and discrimination that already exist in society. Therefore, there needs to be careful consideration of which data samples are used to train the AI.

- **Denial of Individual Autonomy, Recourse, and Rights**

AI systems can automate cognitive functions that were previously made by humans. This complicates who is responsible if there are negative consequences of using the AI.

- **Non-transparent, Unexplainable, or Unjustifiable Outcomes**

Machine learning models derive their outcomes from processing high-dimensional correlations that surpass the interpretative capacities of human-scale reasoning. This becomes problematic if the AI results display discrimination, bias, inequity, or unfairness; the lack of transparency in the model can pose significant issues.

- **Invasions of Privacy**

AI systems present privacy threats in both their design and development stages, as well as during deployment. The construction of AI, reliant on data structuring and processing, often involves the utilization of personal data without proper consent, posing risks to individual privacy. In the deployment phase, AI systems engaging in targeting, profiling, or influencing individuals without their knowledge or consent may be viewed as violating their right to a private life. Such privacy infringements can hinder an individual's core right to pursue goals and life plans without undue influences, affecting their ability to navigate the transformative impacts of these technologies.

- **Isolation and Disintegration of Social Connection**

Excessive use of AI can lead to decreased social interactions, which can decrease human-to-human interaction and lead to polarized social relationships.

- **Unreliable, Unsafe, or Poor-Quality Outcomes**

Inadequate handling of data, careless design and production procedures, and

questionable deployment practices can independently contribute to the development and dissemination of AI systems that generate unreliable, unsafe, or subpar results. Such outcomes have the potential to directly harm the well-being of individuals and public welfare. Additionally, they may erode public trust in the responsible application of AI technologies for societal benefits and result in inefficient use of limited public resources devoted to ineffective or even harmful AI technologies.

Leslie also suggest the following goals to design for ethical AI systems [44]. Ensure the ethical permissibility of your AI project by assessing its potential impact on the wellbeing of stakeholders and communities. Prioritize fairness and non-discrimination by addressing potential biases, mitigating their influence on model outputs, and being conscious of fairness issues throughout the design and implementation phases. Build public trust in your AI project by prioritizing safety, accuracy, reliability, security, and robustness. Justify your AI project by emphasizing transparency in both the design and implementation process, as well as in the decisions and behaviors of the model.

3.6 Advanced User Assistance Systems

According to Maedche et al. [45] an advanced User Assistance system (UAS) is a system that can enhance information systems (IS) to optimize user task performance. These advanced UAS are intricately linked or fully integrated with IS. They go beyond offering guidance, actively sensing user activities and the environment to provide context-aware features. Additionally, they can suggest recommendations based on the user's context, fostering interaction within the user community and with the information system. Maedche et al. [45] lists these characteristics that identifies a UAS: allow users to decide whether to follow the given assistance, provide context-aware assistance, provide proactive assistance, contain adaptation capabilities regarding assistance behavior, detect users' needs while receiving assistance.

Moreover Maedche et al. [45] identifies three different categories of UAS: interactive, intelligent, and anticipating UAS. An interactive UAS is based on interaction between the user and the system, typically relying on behavioral patterns defined by the system engineers. Intelligent UAS can beyond interacting with the user adapt to the specific user and context. Anticipating UAS combines intelligence and interaction to predict future situations. They are able to self-learn and develop their assistance and behaviors to the user's needs and context over time and are not rigidly defined from the start.

Early attempts at comprehensive integrated assistance systems were unsuccessful [45]. A notable example is "Clippy, the paperclip," a cartoon character developed by Microsoft that automatically appeared to assist users of Microsoft Office. Despite its intended purpose of providing clear and precise guidance, studies show that users perceived Clippy as annoying, impolite, and disruptive. However, more recent efforts in user support systems, especially in personal contexts such as mobile devices (Siri,

Google Now, etc.), have proven successful, becoming more advanced and receiving a more positive response.

3.6.1 User Assisted System in Business Information Visualization

Schelkle & Grund investigated the use of a user-assisted system (UAS) in business information visualization to aid users in analyzing information [46]. According to the study, participants mentioned that a good design should visualize the changes made by the UAS and also noted that confirmations can be annoying. The results of the study indicated that the UAS contributed to minimizing errors, providing clear explanations, and ensuring transparency in chart-related modifications (i.e., the ability to trace the actions performed by the UAS) was important. From a perceived ease of use standpoint, participants appreciated easily understandable explanations and indicated the advantages of incorporating multiple languages in the UAS, possibly seen as a form of customization.

3.6.2 Notification Systems

For User Assistance Systems (UAS) to communicate with users, they need to notify them in some way. According to McCrickard and Chewar (2003), users desire notifications from various sources while engaging in multiple tasks. Notification systems can be defined as interfaces designed to aid users in accessing supplementary information that is less critical than their current tasks. However, many of these interfaces are perceived as distracting. This perception may be attributed to information being presented at inappropriate times or in an unsuitable format. In this project, the design of a User Assistance System interface and methods to notify users without disrupting their workflow will be explored.

3.7 Autocompletion in Generative AI Interaction

Lehmann and Buschek examines the interaction with generative AI as auto-completion in different fields, such as code completion, GUI sketching, and layouting [47]. They found five aspects of interface and interaction patterns in AI auto-completion across different fields, these are listed below. These aspects will be useful when designing a UAS that will give AI generated recommendations for the user within the interface of CANEA ONE.

- **User Interface**

The interface should be minimalistic with a field for user input and a separate suggestion area. The generated objects should be placed close to the input but stay separated so that it is clear to the user what is their input and what is generated by the AI.

- **Workflow**

The interface should allow for a continuous interaction between the system

and the user. The generated suggestions from the AI can also become part of the workflow. The interaction between the system and the user is continuous until the output is finally confirmed.

- **User Decision**

The user should be able to freely decide to accept a suggestion or not. The user should be able to ignore suggestions to ensure that the user is in control of their work and the AI has a more passive role. If the recommendations are ignored the system should continue to generate new suggestions.

- **Editing**

The user should be able to freely edit the suggestions to fit their vision. The recommendations can be extended, shortened or deleted. The system should detect errors and inform the user of them and suggest corrections.

- **Information**

Input from the user should never be considered fully complete only as partial input to the system. Therefore the AI should predict a complete version of the input.

3.8 Design for AI

Google has made six design guidelines for best practises when designing for AI [48]. The guidelines are based on data from experts and academic research. This project revolves around the interaction between AI and humans and therefore these guidelines will provide usefully when designing the interface for the AI chatbot and UAS.

- **User Needs + Defining Success**

To apply AI successfully, the users' needs and the capabilities of the AI should be considered first. AI should only be employed to address user problems if it can offer solutions that add unique value. Tasks that are unpleasant, challenging, and have a defined correct solution can be automated with AI. Conversely, tasks that are enjoyable, possess social capital, and lack a clear correct solution can be enhanced with AI.

- **Mental Models**

AI systems can adapt over time. Help users understand this capability by explaining the system's abilities, how it can evolve, and ways to enhance its performance. Additionally, communicate the limitations of the system to ensure users maintain realistic expectations.

- **Feedback + Control**

Collect feedback from users to enhance the AI model. Provide users with the option to choose not to give feedback.

- **Data Collection + Evaluation**

Gather high-quality data. Decide which data is needed to train the model based on user needs. Evaluate and ensure that the data is appropriate for the project.

- **Explainability + Trust**

Explaining how the AI works to the user can help them understand when they can trust the system and when they should use their own judgment.

- **Errors + Graceful Failure**

Understanding the terms "errors" and "failure" in the context of AI is subjective, tied to user expectations. Perceptions vary, as what may be considered a failure or success depends on individual users and the system's purpose. Managing these interactions shapes mental models and influences user trust calibration.

Identifying error sources in AI systems is challenging; errors may originate from various places and manifest in non-intuitive ways for both users and system creators.

To enhance the user experience, it is crucial to provide pathways for users to navigate errors. Given that AI capabilities evolve, enabling users to take action in response to encountered errors fosters patience, sustains the user-AI relationship, and contributes to an overall improved experience.

3.9 Trust In AI Chatbots

Trust has been characterized in numerous ways by various researchers from different disciplines [49]. In the context of AI, trust can be defined as a belief that an AI-enabled agent's service and reported results are reliable and trustworthy[50]. Følstad et al. highlights that the importance of user trust has been proven to play a pivotal role in the acceptance of innovative interactive solutions [51]. However, our understanding of how user trust in chatbots and the factors influencing it remains significantly constrained. This knowledge gap needs to be addressed if we wish to improve customer service via chatbots. Additionally, trust holds an outstanding importance in online interactions, impacting a customer's receptiveness to machine-provided information, willingness to follow their recommendations, and ability to leverage the inherent benefits of chatbot systems [52].

A study by Przegalinska et al.[53] suggests three dimensions of trust when developing chatbots in order to enhance their refinement, nuance, context-awareness, and transparency, thereby evolving into more sophisticated collaboration assistants. The dimensions consists of: transparency and honesty, predictability and control and benevolence. Similarly, Følstad et al.[51] argues that there are 5 main factors that affect the user trust and can be divided into factors that concern the chatbot and factors that concern the environment of the chatbot. The factors interpretation

and advice, Self-presentation, Human-likeness and professional appearance fall into the first category while the factors brand, security and privacy and risk fall into the latter category. In this project the user interface and design of the chatbot will be designed to gain user trust and will be evaluated with user-tests.

3.10 User-Centered Design

What is User-centered Design (UCD)? The definition varies depending on the source. According to Abras et al. [54] UCD, also known as human-centered design and customer-centered design [55], can be described as a broad term in design processes where the influence of end-users plays a significant role in shaping the design [54]. Miaskiewicz & Kozar [55] share a similar view, stating that UCD presents a broad idea to design that involves integrating users or consumers into the design process. By using this approach users can be treated as a source for innovation, enabling companies to uncover valuable insights by probing about their needs or, even more effectively, by observing them while utilizing existing products and monitoring their behavior throughout the use of the product [56].

In UCD design solutions can be evaluated by users [54]. By listening to users discuss alternative designs, designers' comprehension of the intended purpose can deepen. As the design process progresses iterations of prototypes of various fidelity can be produced and user tested or evaluated. User-testing should be highly considered because they aid in identifying measurable usability criteria encompassing effectiveness, efficiency, safety, utility, learnability, memorability (how quickly users remember common tasks), and subjective satisfaction. It is challenging for designers to anticipate all crucial usability criteria without user input. Only through feedback gathered in an interactive iterative process involving users can products be refined. It is also important to note that UCD is not the same as usability [57] which refers to the study of how humans relate to products. However, applying UCD will assure that your product has both great usability and user experience.

4

Methods

This chapter serves as a comprehensive guide to the methodological framework and scientific approaches employed in the course of this research project.

4.1 Design Sprint

Design sprint is a methodology developed by Google, used to solve problems through designing, prototyping, and testing ideas with users [58] and was the main framework used during the project. A design sprint consists of six phases: Understand, Define, Sketch, Decide, Prototype, and Validate, see fig 4.1 [59].

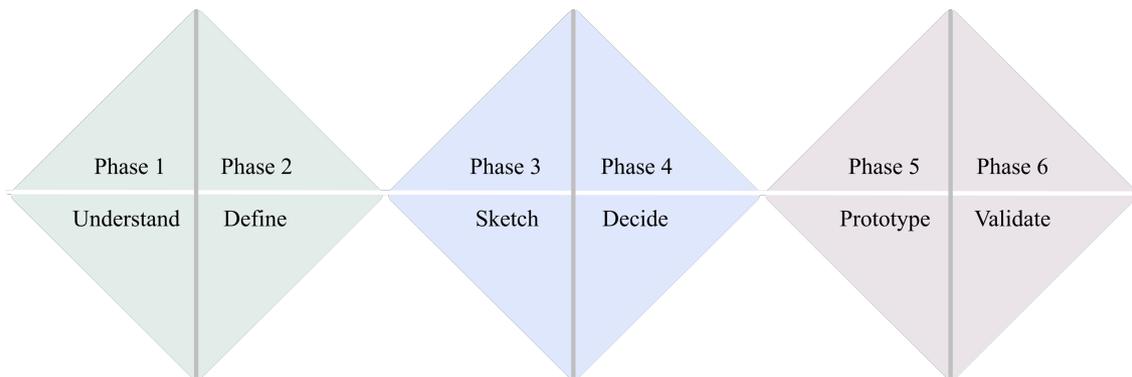


Figure 4.1: Design Sprint

This project began with research aimed at understanding the problem and formulating it effectively. To gain insights into the problem, the authors conducted a literature review analyzing previous solutions. This was followed by questionnaires and interviews with users to understand their perspectives. Subsequently, the process involved brainstorming and ideation to generate various potential solutions. Following this, prototypes of the ideas were developed and later assessed with users. For a more detailed timeline, please refer to Appendix A.

4.2 Literature Review

Literature review is a method for collecting and analysing existing data. As outlined by Knopf, a literature review offers several advantages, including providing a comprehensive overview of unfamiliar research, highlighting well-executed prior work,

generating new ideas applicable to your research, and aiding in the identification of issues or flaws in existing research [60]. As mentioned in the previous section, literature review was mainly used during the research phase of the project where a deeper understanding for the problem at hand was critical.

4.3 Questionnaire

Questionnaires, which serve as a valuable tool for engaging a broad audience and collecting essential data on users' opinions, habits, and behaviors [61, 62], are subject to the crucial influence of question formulation. Studies, as highlighted by Lietz [63], have demonstrated that the careful curation of question length, wording, and order plays a pivotal role in ensuring the success of the communication process. This strategic approach helps avoid negative impacts on sample quality caused by non-responses, contributing to the effectiveness of questionnaires as a means to gather quantitative primary data. Despite their advantages, it's important to acknowledge some challenges associated with questionnaires, such as low response rates, the inability to address misunderstandings in real-time, difficulties in motivating respondents, and a preference for verbal responses over written ones. .

4.4 Interviews

Interviews are a qualitative research method used to gain a deeper understanding of users and their needs [64]. Unlike questionnaires, interviews excel at collecting narrative data, enabling researchers to delve more profoundly into individuals' perspectives [65]. Through user interviews, you can discern who the users are, understand their experiences, and identify their values and desires. Interviews help build empathy towards the users, making it easier for researchers to put themselves in the users' shoes.

There are mainly three types of interviews: structured, unstructured, and semi-structured interviews [66]. Structured interviews are somewhat similar to questionnaires, as the researcher presents predetermined questions, which tends to improve the response rate [67]. Unstructured interviews are the complete opposite and the interview is instead based on a set amount of issues or topics. In this kind of interview the researcher may need to adjust their questions and their sequence based on the interviewee's responses. The semi-structured interview is the most commonly used type. Semi-structured interviews manifest in various forms, with differing quantities of questions and varying degrees of question adaptation and order adjustment to suit the interviewee. There are usually a span of six to twelve well-chosen and well-phrased questions. There can be two to four sub-questions or prompts for each main question, and the interviewer employs them if needed to ensure a comprehensive exploration of the main question by the interviewee.

In a comprehensive exploration of interview dynamics, Guest et al. [68] examined how varying interview quantities impact data saturation and variability. Interviews

were analysed into codes and the research showed that 73% of codes had been identified following transcription of the initial six interviews. The author also concluded that data saturation had occurred after transcribing twelve interviews, where 92% of the codes had been developed. Additionally, researcher Morse suggests at least six participants for phenomenological studies [69] (observational studies).

4.5 Personas and Storyboards

Personas play a crucial role in design processes, enhancing user-centered design by providing an accurate representation of customer requirements, facilitating communication, and pinpointing key advantages [55]. Moreover, personas represent targeted users who share common traits, achieved through the creation of a unique narrative supplemented with a name, visual depiction, and identity [70].

Similarly, storyboards can aid designers in understanding the targeted user and the context of use [71]. Storyboards are a tool used by designers to create a common language that people from different backgrounds can understand and can be used to communicate with all parties involved such as stakeholders and targeted users [72].

4.6 Ideation

Utilizing ideation techniques enhances the processes of generating ideas and improves the resulting outcomes, as suggested by Daly et al. [73]. Concept generation also benefits from adhering to several recommended best practices, including producing a substantial number of concepts, creating a diverse range of ideas, and refraining from making immediate judgments about the worth of these concepts.

Different ideation methods were used in the project to come up with several design suggestions for the AI chatbot interface. Utilizing methods such as brainstorming, brain-writing, and crazy-eights, diverse ideas were generated.

4.6.1 Crazy-eights

Crazy-eights consists of each participant dividing a paper into 8 parts. A timer is set to 1 minute and every participant fills one of the rectangles of the paper with an idea until the timer runs out. The process is repeated until the entire paper is filled [74].

4.6.2 Brain-writing

Brain-writing is a method where the participants have a paper each and a timer is set for 1-2 minutes. The participants get to write or draw their ideas on the paper until the time runs out. Then the participants switch papers clockwise and the timer is set again. This time the participants expand on each other's ideas until the timer runs out. Repeating until every participant has contributed to every paper [75].

4.6.3 Co-design Workshop

Co-design can be defined as collective creativity within the design process[76]. The method involves a collaborative effort among diverse experts, including researchers, designers, developers, and potential users—who are recognized as "experts of their experiences". It is important to note that in co-design the experiences of users and customers is central. Utilizing co-design offers several advantages, including enhanced quality of system requirements, a more precise alignment between the system and users' needs, increased user satisfaction, and a reduction in development time.

4.7 Prototype

According to Lauff et al., a prototype is an embodiment, whether physical or digital, of the critical elements of the intended design. It is also an interactive tool that enhances communication, enables learning, and informs decision-making in the design process [77]. Prototypes are created to visualize ideas and test them with users. The prototyping phase began with the utilization of tools such as Figma to create wireframes [78], followed by the development of a functional prototype.

4.7.1 Low & High Fidelity Prototype

Prototypes of various fidelities and media are employed to test interface designs [79]. Low-fidelity prototypes are typically used in the early stages of a project and can be utilized to propose a design for the user interface [80]. This design can then be presented to users for feedback, which can be used to further iterate on low-fidelity prototypes or serve as requirements for higher-fidelity prototypes. The benefits of low-fidelity prototypes include their low cost, portability, and ability to address layout issues. However, they also have several disadvantages: they are not well-suited for usability testing, have limited flow and navigation, and are facilitator-driven.

High-fidelity prototypes, unlike their low-fidelity counterparts, are highly functional and interactive, making them excellent for addressing navigation and flow issues and ideal for user tests and evaluations. However, high-fidelity prototypes are often more time-consuming and costly to produce. Given the choice, developers may opt not to create any prototypes and instead test interactivity once the product has been fully developed.

In the early stages of the prototyping phase, Figma was used to create low-fidelity prototypes with minimal interactivity and functionality. These were presented to stakeholders for feedback, resulting in the development of a high-fidelity prototype using React for the frontend design and the Chat-GPT API [81] for chatbot functionality.

4.7.2 Figma

Figma is a powerful real time collaborative design tool for creating graphical interfaces [82]. The tool includes nearly everything required to design intricate interfaces, brainstorming and creating wireframes and prototypes. In Figma animations can be used to make interactive mock ups for applications.

Low fidelity prototypes were created in Figma to create the designs created in the ideation phase.

4.7.3 Wireframing

A wireframe can be described as a rough visual representation of a proposed application [83]. It is a specific type of prototype created within the user-centered design process to collect input and feedback from target users before finalizing designs. Wireframes are recognized for contributing to the gradual improvement of both the utility (usefulness) and usability (ease of use) of applications. Typically generated early in the digital design process, wireframes possess a limited amount of interactivity but serve as a valuable tool for providing a visual representation of the application's functionality and features. They act as a guide for understanding how users will navigate different parts of the application and interact with its various elements. By visualizing the layout and elements early in the design process, designers can proactively address any confusion or usability issues that users might face.

4.8 Evaluation & Analysis

To evaluate prototypes data triangulation, which refers to using several methods to get a broader understanding of a phenomenon, [84] was used. Furthermore, the data analysis process involved the utilization of affinity diagrams to systematically organize and manage the gathered information [85].

4.8.1 Heuristic Evaluation

Heuristic evaluation is a method for identifying usability issues in interface design [86]. This method involves a small set of evaluators examining and judging the interface based on how well it complies with usability principles, also known as "heuristics." According to Nielsen and Mack [87], the level of expertise of the evaluators plays a significant role, as there are consistent and systematic differences in performance when looking at groups of evaluators. The authors recommend using 3-5 evaluators if "single expert" usability specialists were employed and 2-3 evaluators if "double expert" usability specialists were used. A "single expert" evaluator refers to a person with general usability experience, while a "double expert" evaluator is defined as someone with a usability background and expertise in a specific application area.

4.8.2 User Testing

User testing is an observational method for identifying usability problems [88], where researchers observe real users and their interaction with the design [89]. In this project, qualitative usability testing will be conducted. Qualitative usability testing aims to collect insights, findings, and anecdotes about how users utilize the product or service. It is particularly effective in identifying issues within the user experience. Regardless of the type of user test, common elements include the facilitator, the task, and the participant. The facilitator assigns tasks to the participant, closely observing their behavior and attentively listening for feedback as they complete these tasks. Additionally, the facilitator may ask follow-up questions to gather more detailed information.

4.8.3 Affinity Diagram

The Affinity Diagram, also known as the KJ method [85], is a technique employed to organize a large quantity of data. This method is versatile and finds application in diverse tasks such as data analysis, idea generation and planning, as well as consensus-based decision-making [90]. The process of the method can generally be broken down into six steps[85]:

1. Transfer collected data onto sticky notes.
2. Place sticky notes onto a large surface such as a table or wall.
3. If needed divide the group into smaller groups of 2-3 members.
4. Have one group arrange the cards in silence, grouping them into related issues, ideas, or topics.
5. Rotate the groups of participants every 2-3 minutes until there are no sticky notes left.
6. Create suitable titles for each group of sticky notes and, if logical, combine smaller groups of sticky notes into larger supergroups.

Its recommended that the group takes the time to discuss the results of the affinity diagram and to review any surprises or disagreements. Some benefits of the methods include: speed, acceptance and new thinking.

5

Process and Execution

In this chapter the process and execution of the methods are presented.

5.1 Literature Review

The project started with a comprehensive literature review delving into existing materials on AI chatbot design and human/AI interaction. By examining solutions implemented by others, including conventional assistance chatbots, we aimed to gain insights and inform our approach. Search engines such as Google scholar, Zotero and Chalmers library were used to find high quality papers.

To begin with keywords such as "AI chatbot", "AI chatbot in business application", "AI chatbot interface", and "ethics in AI" were used. To gain a broader foundation of design theory we also search for "interface design guidelines" and "accessibility design guidelines".

5.2 Questionnaires

The questionnaire was designed using Google Forms to gather information from users of the CANEA ONE application, see the questionnaire in appendix C. It initially covered general demographics, including age, gender, and work role. Subsequent questions focused on users challenges with the program, the solutions they employ, the frequency of such challenges, their experiences with AI chatbots, and their expectations for a chatbot integrated into the application, see Appendix C. The questionnaire was sent out to people within the company who have experience using the CANEA ONE program as well as customers of CANEA. The questionnaire was open for responses for one week.

The questionnaire published to CANEA employees garnered 51 responses and featured a series of questions designed to explore whether users of CANEA ONE encounter difficulties with specific tasks, and if so, which ones. The findings indicate that a majority (84%) have encountered some form of challenges. The frequency of these difficulties varies among users, with the most common response being monthly. Participant proficiency spans across beginners (14%), intermediate (36%), and advanced (50%) levels. There are five main tasks with which a majority of users struggle, including: navigation, locating information, using the process-drawer, configuring different cases, and configuring documents.

Additionally, the questionnaire delved into users' perceptions of AI usage. A significant number expressed either a neutral stance or a positive/very positive outlook, particularly regarding AI chatbots, attributing their enthusiasm to the perceived efficiency boost. Respondents highlighted several potential roles for an AI chatbot within CANEA ONE, including assisting with information retrieval, answering general queries, and providing recommendations.

For the questionnaire that was shared to the customers of CANEA there were 23 replies and among these there were 14 (61%) that had experienced difficulties using the application. A difference between the customers and CANEA employees is that most customers have experience using the document, cases and process modules while most of the employees have experience with all modules. The customers experienced most difficulties with navigation, configuration of cases and using the process-drawer which was also the largest problems among the employees.

5.3 Interviews

For this project the interview type semi-structured interview was selected. The authors carefully prepared and formulated questions and sub-questions in a google document, see Appendix D.

A total of six participants were selected from the questionnaires considering that they need to have preceding experience with CANEA ONE and three participants of intermediate and three participants of advanced skill level were selected because no beginner users were available. The interviews were done with a mix of in person interviews and online interviews through Microsoft teams. All interviews were recorded using a smartphone with the consent of the participants to be used for transcription.

For the observation part of the interview the participant was asked to perform a task on CANEA's internal deployment of CANEA ONE called CMS, either on their own computer (if performed online) or on the computer of the authors. Interviews were conducted to gather valuable insights into users' attitudes and expectations concerning AI chatbots. By interviewing CANEA ONE users, we aimed to gain further insights into their preferences for the chatbot's functionality. The interviews consisted of two parts: a question segment and an observation segment where users completed a task in the application while explaining their thoughts.

After each interview the key findings were transcribed into post it notes on Miro. The interviews were then analysed using affinity diagrams to get a grasp of topics and issues.

5.4 Affinity Diagram Interviews

Affinity diagrams were used for analysing the interviews. When making the affinity diagram Miro was used as a virtual board with sticky notes.

Important quotes from the interviews were written down on post-it notes and later grouped based on theme. The themes were not predefined but were discussed between the authors during the making of the affinity diagram. The categories created were problems, the manual/existing solutions, help, chatbot, AI functionality, AI response, improving CANEA, Customers of CANEA, AI trust, AI advantages/disadvantages, AI area of use, use of CANEA ONE and AI in CANEA ONE. The largest topics were problems, use of CANEA ONE and AI in CANEA ONE. See figure 5.1 for reference of the cluster AI in CANEA ONE.

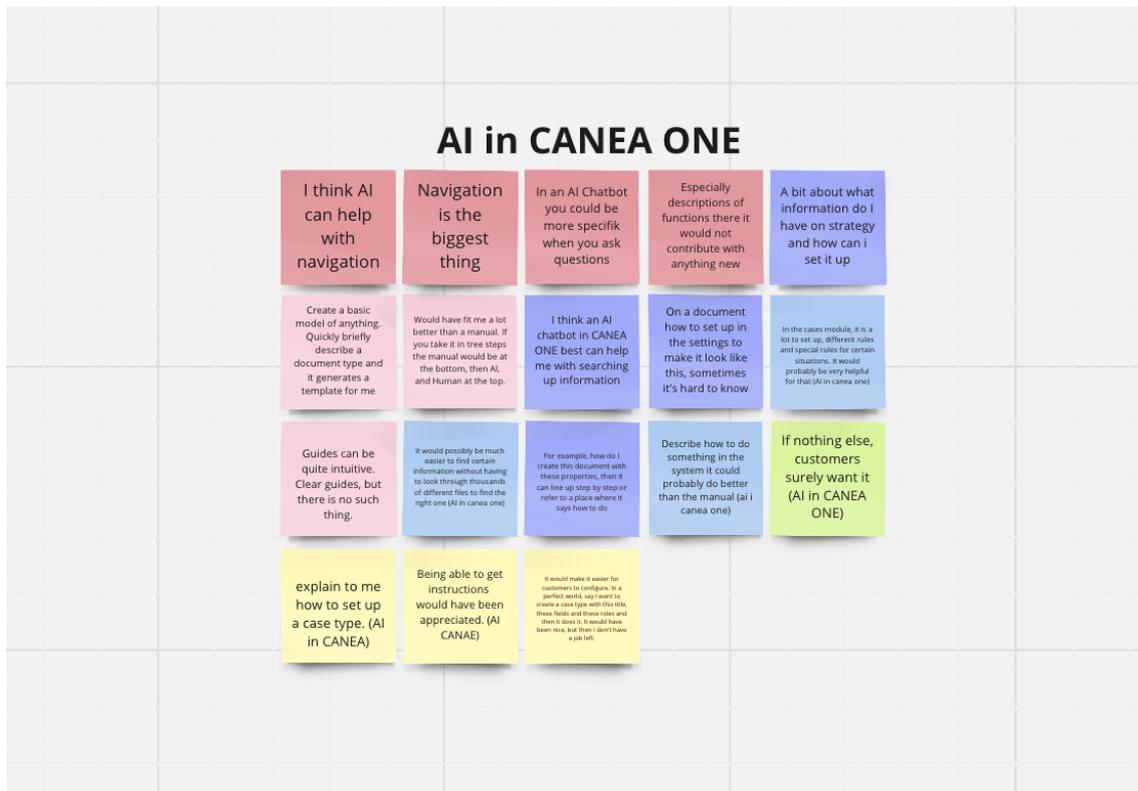


Figure 5.1: Example of the cluster AI in CANEA ONE

5.4.1 Interview Findings

One common reoccurring opinion amongst the participants of the interviews is that it can be challenging navigating and some functions can be difficult to find especially during the configuration of for example a project. This is due to the fact that settings for the configuration are scattered across the application.

"In general, I find it quite non-intuitive, messy, and complicated. It's a large program" - Participant 2

One participant also noted that the absence of feedback during the configuration design process causes significant issues. This lack of feedback permits users to assign

any values, even if they might not be functional, potentially leading to crashes when attempting to create something.

"When designing, it would have been beneficial to receive more input if one makes mistakes. One can set it up however they want, but it's not guaranteed to work. There is a lot of validation missing for users." - Participant 5

Furthermore, if a user notices that something is wrong in the middle of the designing process, they may need to redo the tasks they have already completed because they navigated away from the page, resulting in the loss of all previous information.

"Depending on where you are, some pages may be such that you have to start over if you go somewhere and then back." - Participant 5

When questioned about their preferred source of assistance while using CANEA ONE, the majority of participants indicated that they rely on their colleagues rather than consulting the manual. Many voiced concerns about the manual's inadequacy in explaining the functions comprehensively, noting inconsistencies in detail across its pages. Participants highlighted the manual's utility primarily for looking up tags, which can be challenging to memorize.

"I do try to look into the manual sometimes, and it does tell me what things are, but not how to use them." - Participant 1

The respondents think that an AI chatbot in the CANEA ONE application would be helpful for navigation, providing specific information, generating document templates, searching for information, setting up configurations, and guiding users through system processes. They mention that a chatbot would be more useful than a manual and express a desire for clear instructions and examples of how to use different functions.

"Being able to receive instructions would have been appreciated." - Participant 6

Users also mentions the potential benefits of having an AI chatbot in assisting with setting up different configurations and templates.

"Create a basic model of anything. Quickly describe a document type, and it generates a template for me." - Participant 2

It was also mentioned that AI could provide assistance faster than reaching out to

a colleague or calling support, enabling users to solve problems more easily on their own.

"It becomes quicker and easier, before reaching out to a real person, one can try on their own first." - Participant 3

After the interviews, we realized that there aren't many issues with basic functions, such as viewing documents, which are the functions most commonly used by regular users. The majority of problems arise for administrative users who handle more complex functions, such as configuring documents and managing cases. If the administrator sets up the application effectively, it will become easier to use for end users. Therefore, this is the area where the most assistance is needed from the AI assistant/chatbot.

5.5 Persona and Storyboard

A persona and storyboard was created using the data gathered from the qualitative and quantitative studies to further define the context where users encounter difficulties.

5.5.1 Persona: Per

Per is 35 years old and works at a company that utilizes CANEA ONE to manage its business operations. He is a daily user of all the modules within the application, making him an advanced user responsible for handling configurations. Per frequently communicates with CANEA employees if any issues arise.

5.5.2 Persona: Alice

Alice is 27 years old and uses the document and process modules of CANEA ONE weekly. Alice is an intermediate user and does not handle configurations of any kind and only views information in the program. This means that Alice mostly experiences difficulties finding the right documents and navigating to certain functions.

5.5.3 Storyboard: Configuration

The story board presents Per who is 35 years old, see fig 5.2. Per works at a company that uses CANEA ONE to manage their business. He is an advanced user and is responsible for taking care of configurations. In this story Per is trying to configure a case. However, he is unsure about some information that is required in order to create the configuration but he somehow manages to create the configuration anyway.

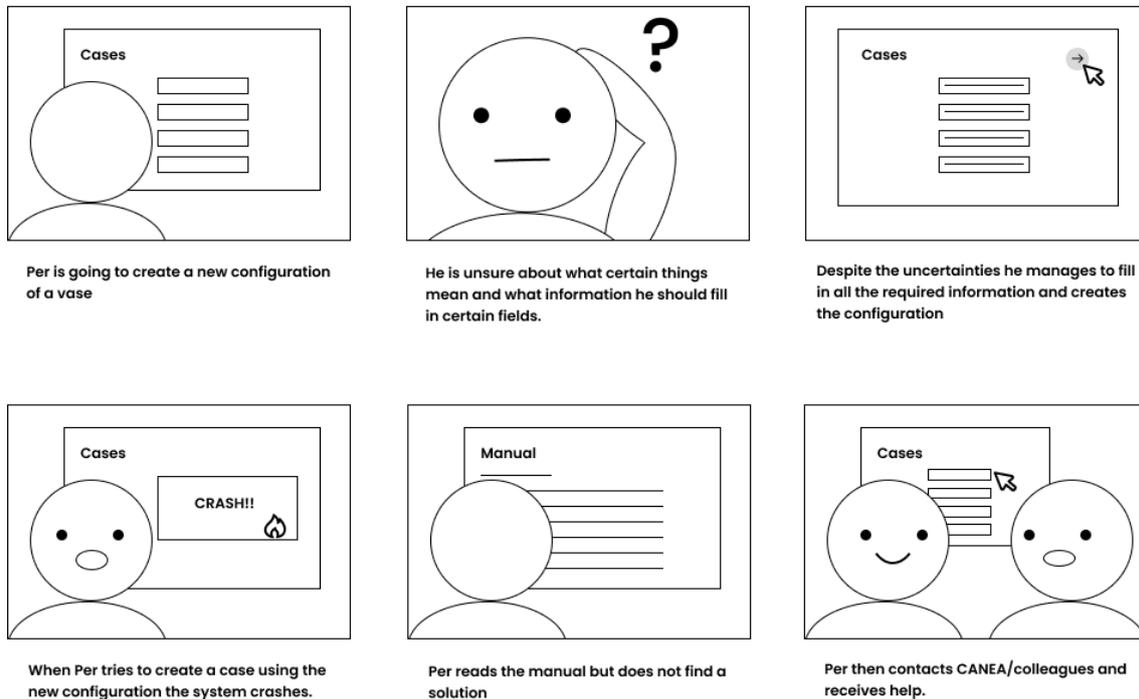


Figure 5.2: Storyboard of Per

When he tries to create a case using the configuration he just created the system crashes. He turns to the manual but does not find a solution. Per then contact CANEA/colleagues and receives help.

5.6 Crazy-eights

The crazy-eights method began by folding two A4 papers in half both horizontally and vertically, creating 8 rectangles. Instead of setting a timer for one minute, we opted for a two-minute duration, allowing us more time to contemplate various ideas. Once the timer elapsed and we had generated a total of 16 ideas, we proceeded to discuss and compare them to identify similarities and differences.

5.7 Brain-writing

The brain-writing session started with each author sketching an idea on a piece of paper. These ideas were more detailed and took inspiration from the previous generated concepts. No timer was used to to let the authors ideate freely without the pressure of a timer. When the ideas were finished the authors switched papers and continued to expand on each others concepts. This resulted in two more detailed concepts.

5.8 Low Fidelity Prototypes

To further develop and visualise the ideas generated from the brainstorming wireframes were created of each concept in figma. There were a total of 16 concepts generated, see fig 5.3 for examples.

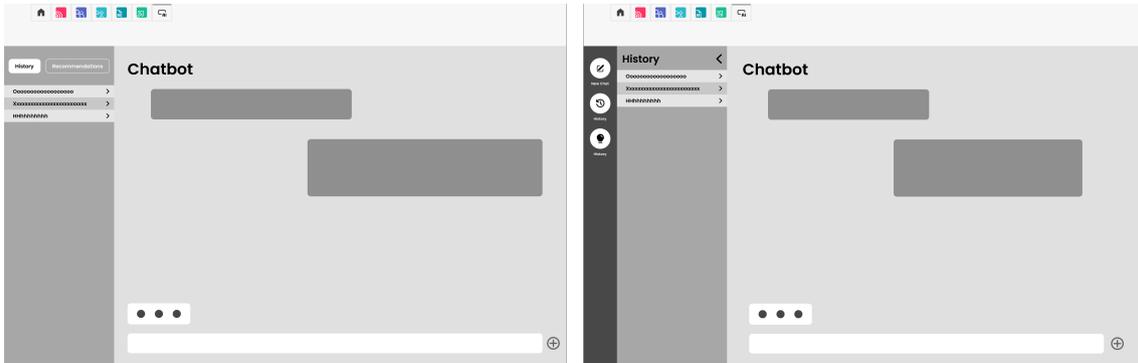


Figure 5.3: Examples of low-fi wireframes

5.9 Co-design Workshop

A co-design workshop was conducted with active participation from CANEA employees who shared their insights and ideas for the project. There were 6 employees who participated in the session. The workshop began with a presentation of findings from the literature review, surveys, and interviews. This was followed by a storyboard session where various scenarios depicting the users' different problems were created, aimed at increasing empathy with users and understanding of their needs. Lastly, a group brainstorming session took place during which stakeholders were encouraged to join and contribute new ideas. After the brainstorming the ideas were grouped and analysed with the affinity diagram method. Lastly a voting session was held where everyone could vote for their preferred ideas. Dot voting was used, by giving each participant 3 stickers, allowing participants to anonymously vote.

5.9.1 Co-Design Workshop Findings

During the workshop, participants were asked to create scenarios featuring a persona illustrating either how a problem could arise when using CANEA ONE or how a solution could resolve the problem. In many of the scenarios, emotions such as frustration arose because the user struggled to find the correct document, etc.

Furthermore, a scenario was portrayed, which was also described by one of the interviewees. In this scenario, within a company, there exists a dedicated individual responsible solely for creating configurations. However, when this individual quits, another user must take over their responsibilities. Unfortunately, the new user

struggles because the workload is overwhelming and there is insufficient support available.

5.10 High Fidelity Prototypes

High fidelity prototypes were created after the ideation phase which included both internal brainstorming with the authors but also a co-design workshop with CANEA, one of the stakeholders. The high fidelity prototypes were created in Figma and were meant to be interactive enough to be used for heuristic evaluations. Flows were created to illustrate how the interface would be used and what would happen when clicking different elements in the interface. The authors decided to proceed with two chatbot ideas, see Figure 5.4 and 5.5, and two UAS ideas, see Figure 5.6 and 5.7.

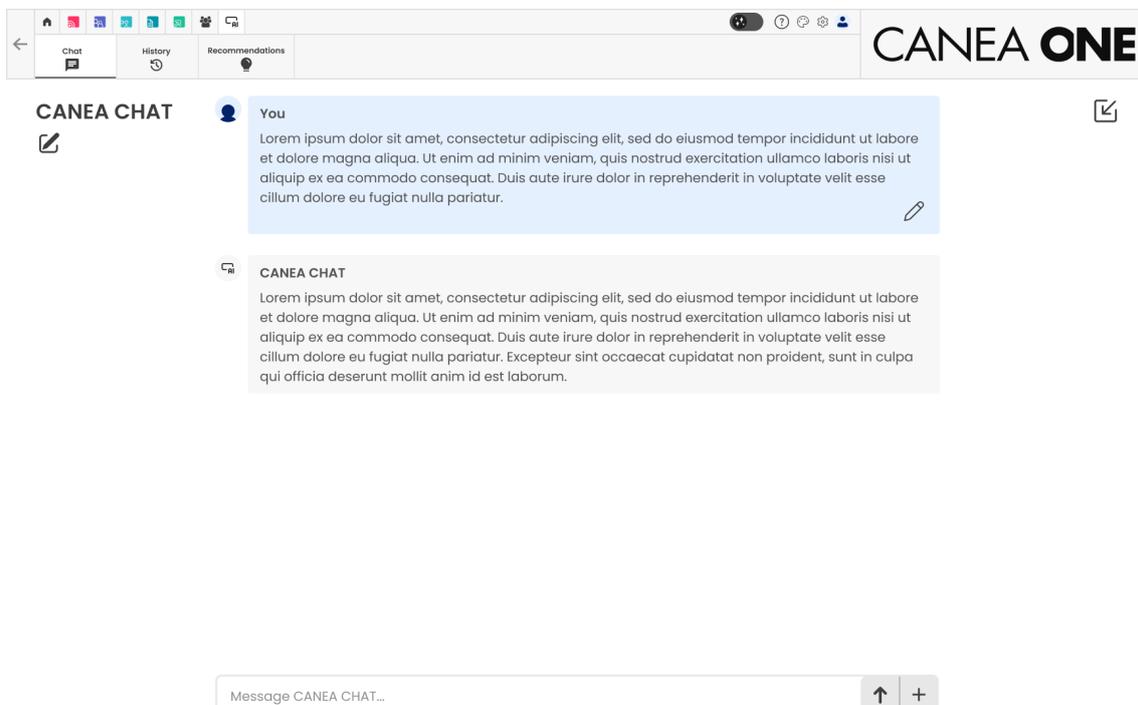


Figure 5.4: Chatbot 1

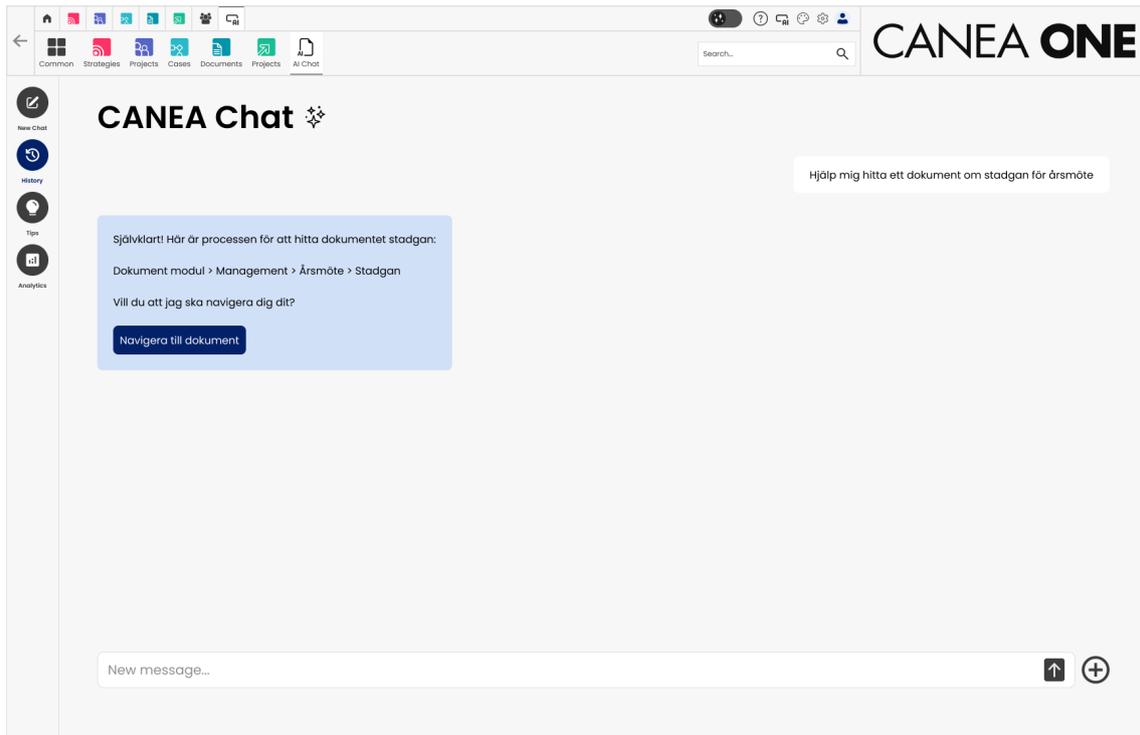


Figure 5.5: Chatbot 2

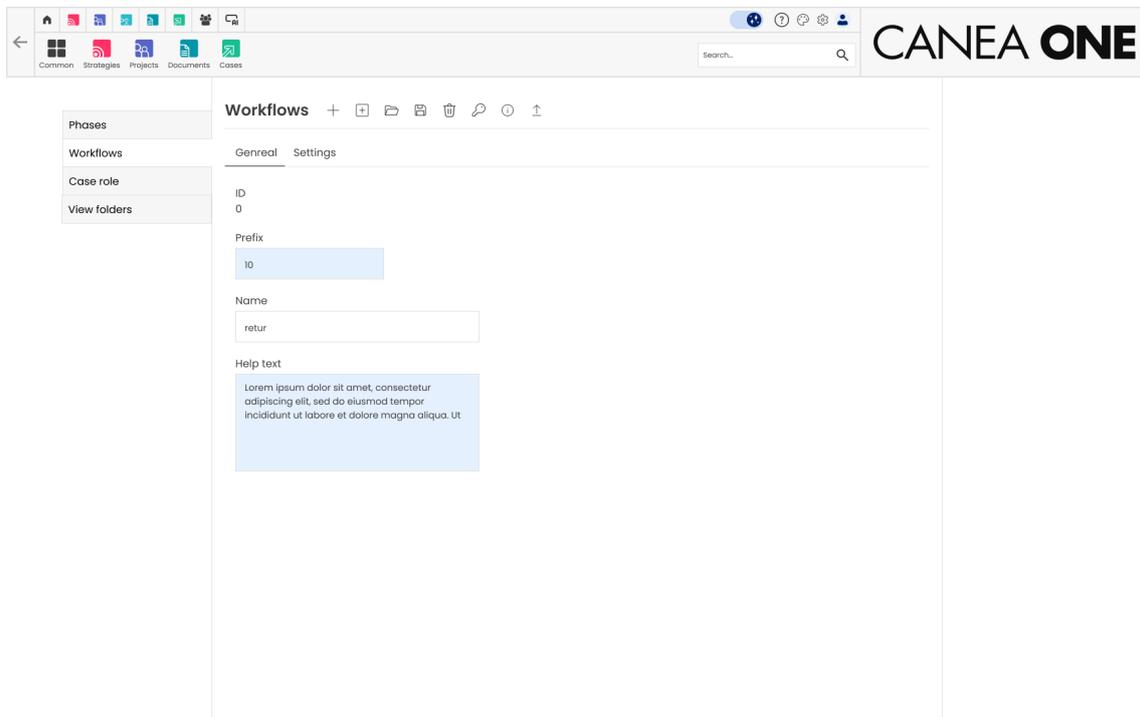


Figure 5.6: UAS 1

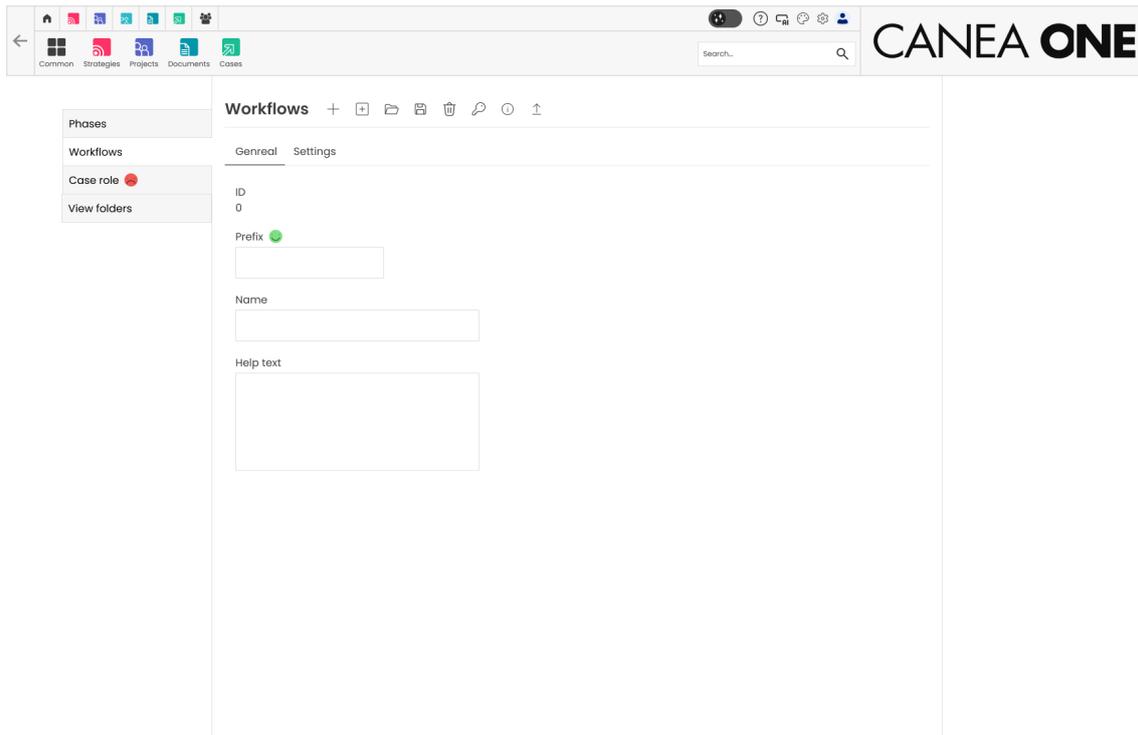


Figure 5.7: UAS 2

5.10.1 Heuristic Evaluation

To evaluate the high-fidelity prototypes and to detect potential usability problems Heuristic evaluations were conducted with a total of 3 usability experts. Two of the experts were recruited from the Interaction design program at Chalmers University of Technology and the third expert was a UX designer working at CANEA.

The results of the heuristic evaluation indicate that, to enhance usability, it is recommended to incorporate copy buttons for the answers provided by the chatbot. Additionally, a "scroll to bottom" button should be added in case conversations become lengthy. It was also suggested to include a feature allowing users to toggle between different chats in the minimized version of the chat. Inconsistencies were identified, particularly between the UAS toggle symbol (a glitter symbol) and the recommendations (which utilized smiley symbols). It is advised to review the icons used and ensure there are no inconsistencies. For example, symbols in the expandable window should align with the rest of the system by using the fluent system icons. To improve access to the function, it is recommended to add access to the recommendations tab in the minimized window. Furthermore, readability was found to be enhanced with a centered chat layout compared to having answers on each side of the screen.

5.11 Functional Prototype

A functional prototype was developed using React with TypeScript and the OpenAI API. The OpenAI API was utilized to retrieve responses from the ChatGPT-3 model, enabling the chatbot functionality. Additional instructions were provided to the chatbot to tailor its responses specifically for the CANEA ONE application, including information sourced from the manual and educational material on how to use CANEA ONE. The prompt given to ChatGPT started with:

"You are Canea Chat who help users with questions regarding the application CANEA ONE. Remember to be kind and polite. If you do not have the information refer to the Canea website or recommend contacting Canea support."

Flask and MongoDB were employed as the backend to facilitate message storage and chat history tracking. Git and GitHub were utilized for version control.

The authors and stakeholders at CANEA decided that the functional prototype would primarily consist of three pages: the homepage, case configuration, case viewing, and a page featuring the chatbot.

CANEA provided a design for a new menu system intended for implementation in CANEA ONE. Consequently, this new menu system was incorporated into the prototype instead of the existing one. This decision aimed to assist CANEA in visualizing how the chatbot would be integrated into future versions of their application. The prototype was based on the designs created in Figma, for reference see section 5.10. However, some adjustments were made due to the new menu system.

In addition to the chatbot a partially functioning prototype of the UAS interface was developed. A replication of the pages required to create a new configuration of a phase in a case was created. On this page, the UAS was integrated into the interface with the ability to click the buttons, but it was not fully functional in providing recommendations and auto-filling information.

5.11.1 Prototype Interface

The AI chatbot exists in CANEA ONE in two instances. One instance as a full-screen window and one as a smaller window. The full-screen chatbot is navigated to by clicking the chatbot icon in the menu to the left. When clicking the icon a modal will open which contains the different functions of the chatbot, new chat, history and recommendations, see Figure 5.8.

When starting a new chat the user is presented with four recommendation prompts. These prompts should assist users who are unsure of how to start a conversation and serve as inspiration, see Figure 5.9.

When a user sends a message a typing animation will appear indicating that the chatbot is thinking. This is to communicate the system status to the user, see Figure 5.10.

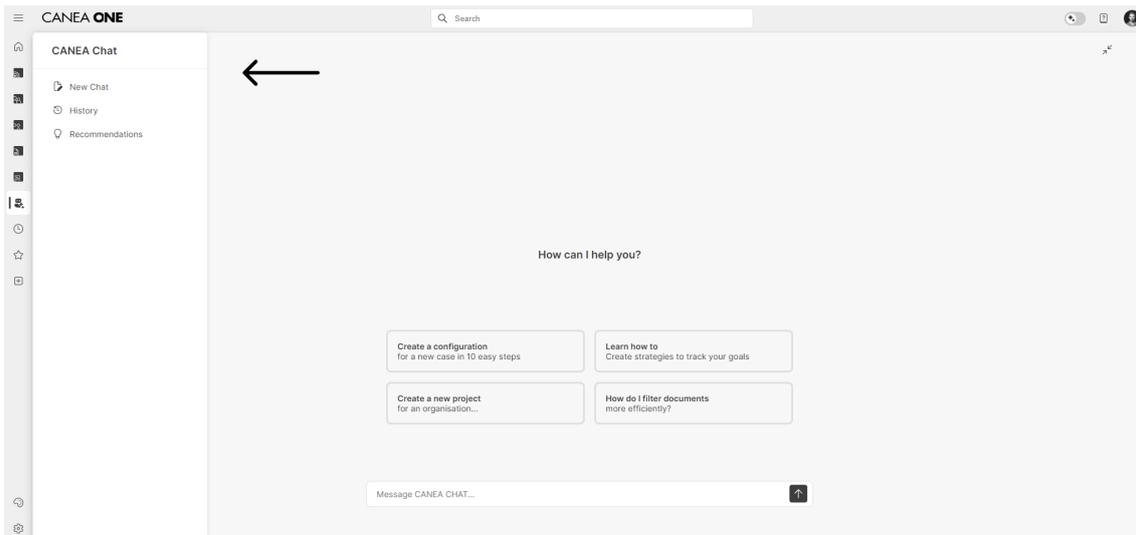


Figure 5.8: Chatbot Menu Open

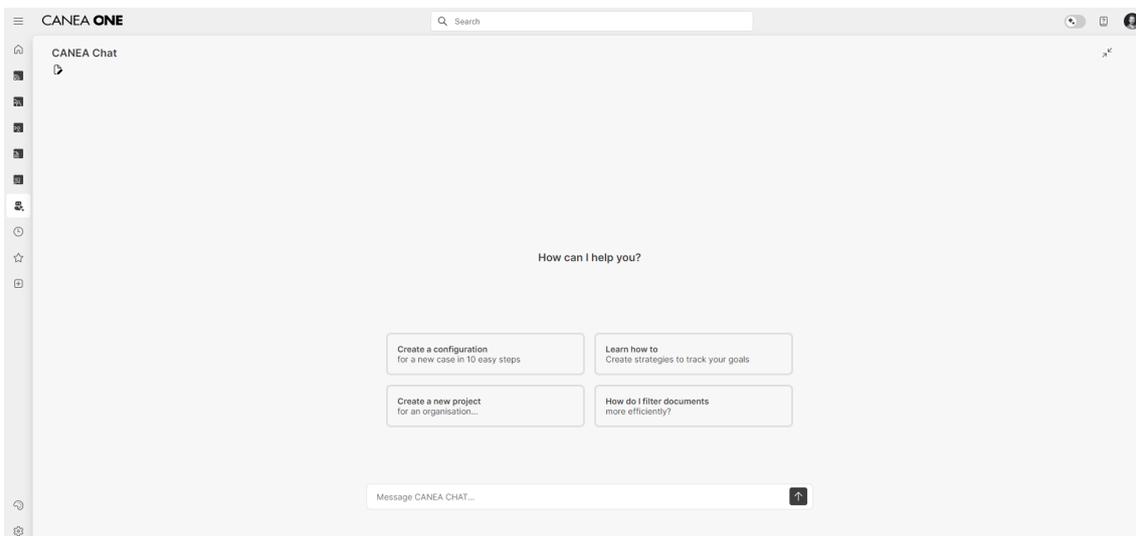


Figure 5.9: New Chat

When chatting the users messages are blue and the messages from the chatbot are white or gray, depending on if you are in the small or full screen mode, see Figure 5.11. This helps differentiate between your own and the chatbots messages. The messages are also accompanied with an icon of the chatbot or user depending on who sent the message, as well as a title that says either "You" or "Canea Chat". This makes sure that users can differentiate between the messages without relying solely on color.

When scrolling up a scroll to the bottom button appears that helps the user navigate to the bottom in a longer chat. The centered messages layout was chosen to increase readability for longer messages. While hovering on the messages a copy icon will appear on the messages from the chatbot and an edit icon will appear on the messages from the user. This helps the user copy information and adjusting

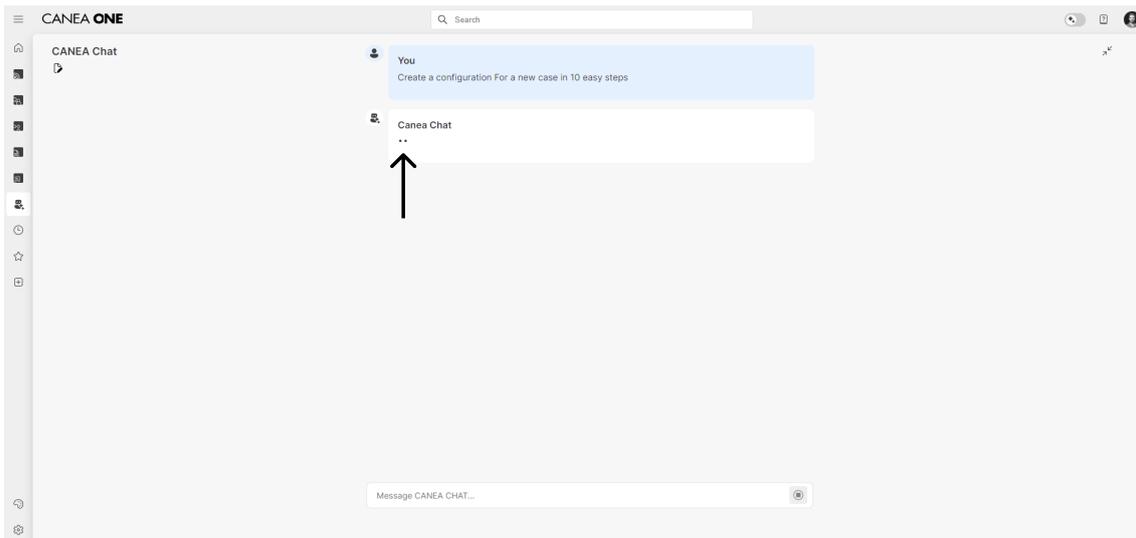


Figure 5.10: Chatbot Typing

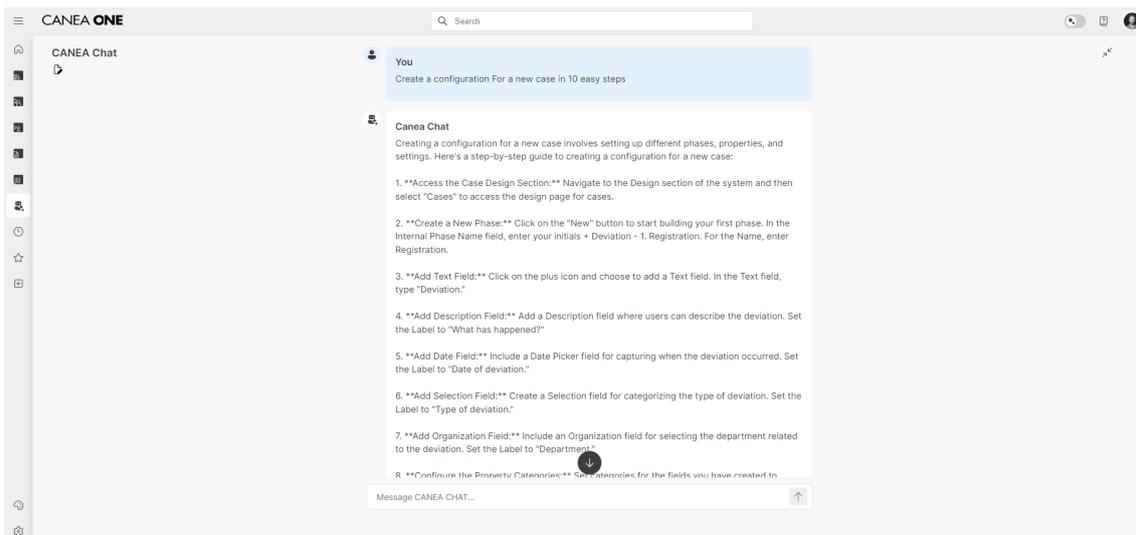


Figure 5.11: Chatbot Messages

previous messages, see Figure 5.12 and 5.13. The icons used are from Fluent icons because these are the icons used by CANEA. The colors are also the same as used in CANEA ONE to keep the chatbot cohesive with the rest of the application.

The minimized chatbot has mostly the same functionality as the full screen chatbot but in a more compact design to enable users to chat while they are using other parts of the application, see Figure 5.14. History and new chat are located in the top menu of the minimized chatbot, see Figure 5.15.

5. Process and Execution

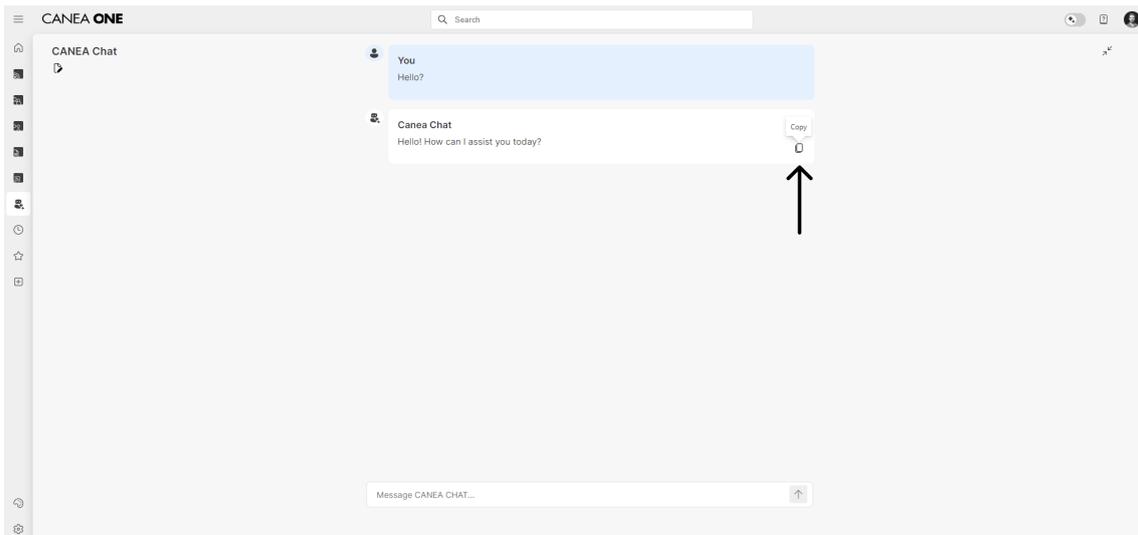


Figure 5.12: Copy Message

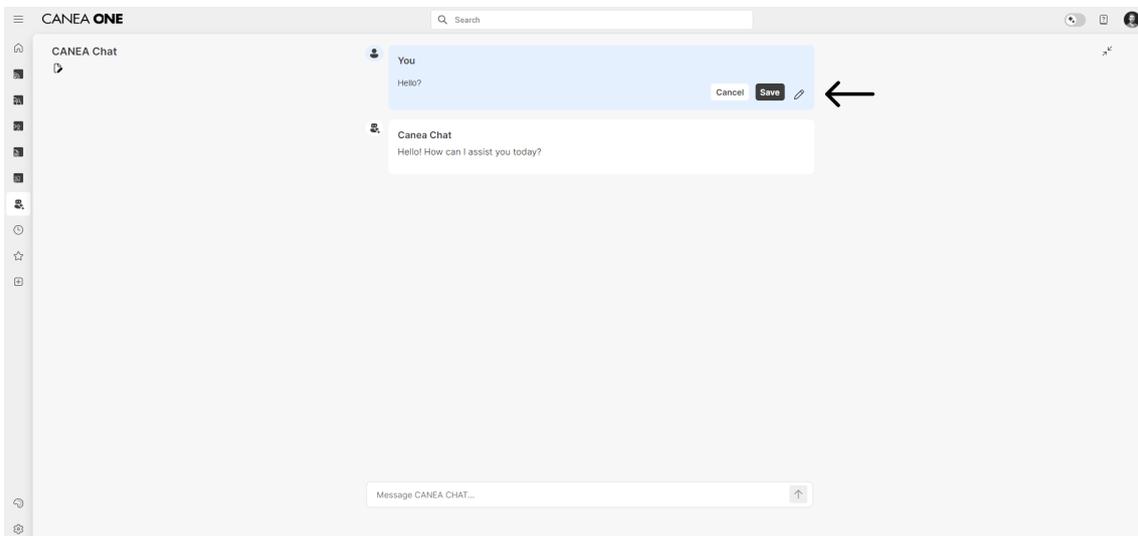


Figure 5.13: Edit Message

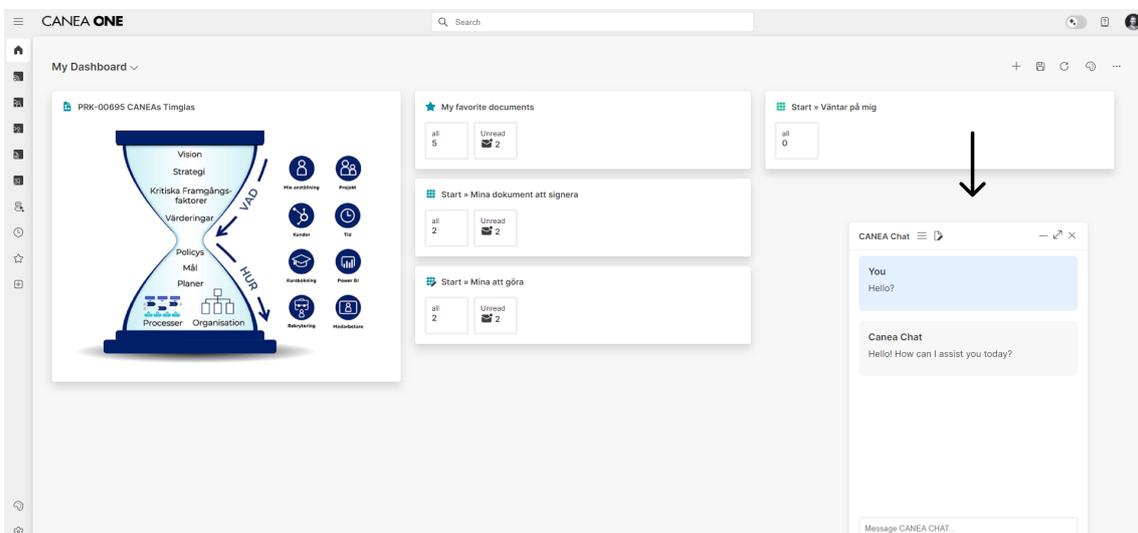


Figure 5.16: Minimized Chatbot Messages

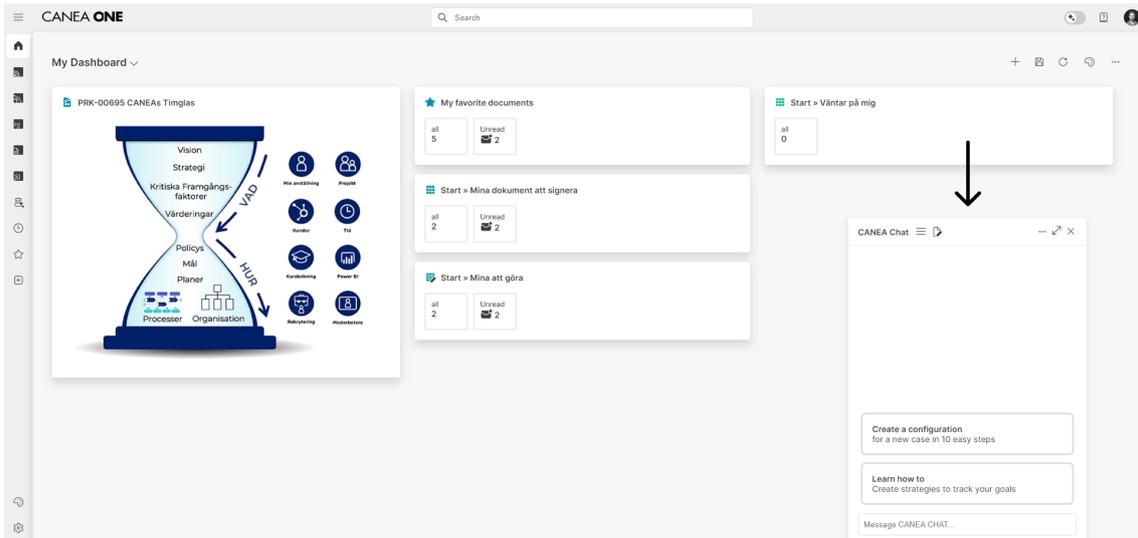


Figure 5.14: Minimized Chatbot

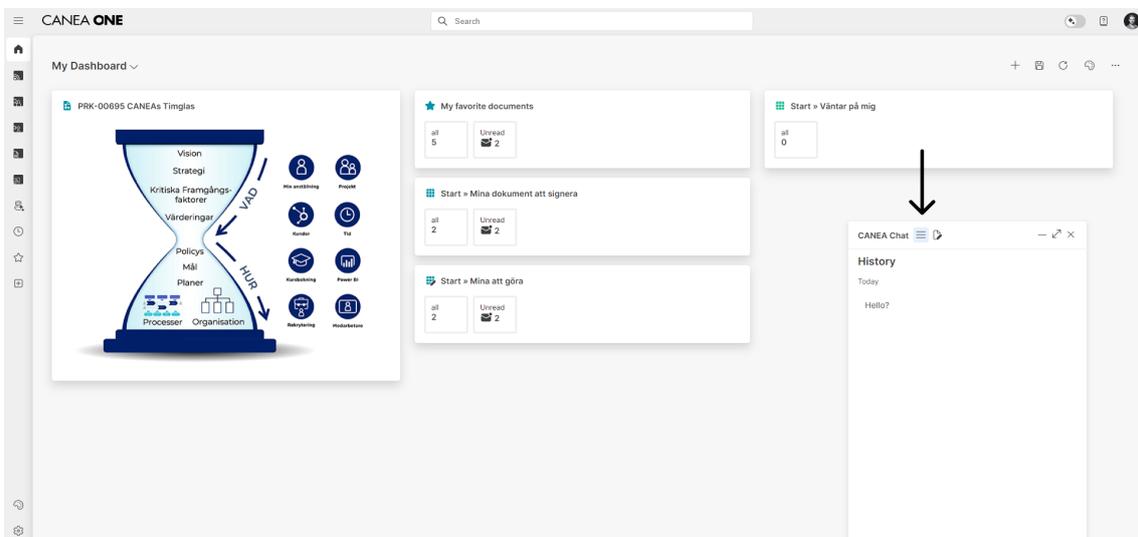


Figure 5.15: Minimized Chatbot History

The small chatbot can be minimized further or closed, then a blue button will appear in right corner to enable users to easily open up the chatbot again, see Figure 5.17.

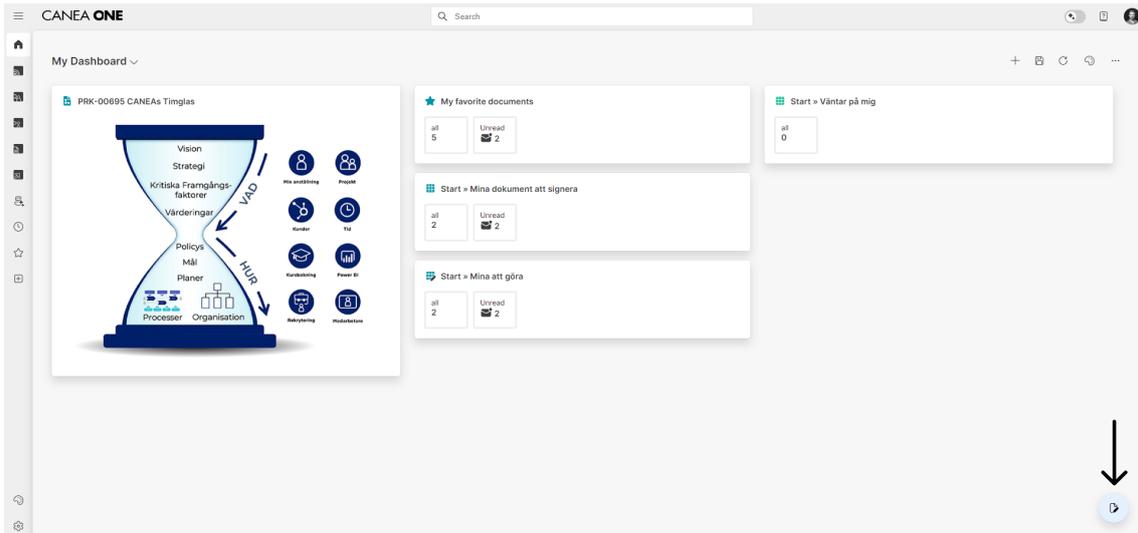


Figure 5.17: Chatbot Closed

5.11.2 UAS Interface

The UAS function can be toggled on or off using the switch with the sparkle icon in the upper right corner. This enables users to easily turn on or off the functionality when needed, see Figure 5.18. It was decided to keep this function easily available for users in case they find the recommendations annoying they can quickly turn them off and on again when they are needed. The icon turns blue when switching the UAS on and gray when the UAS is off, see Figure 5.19.

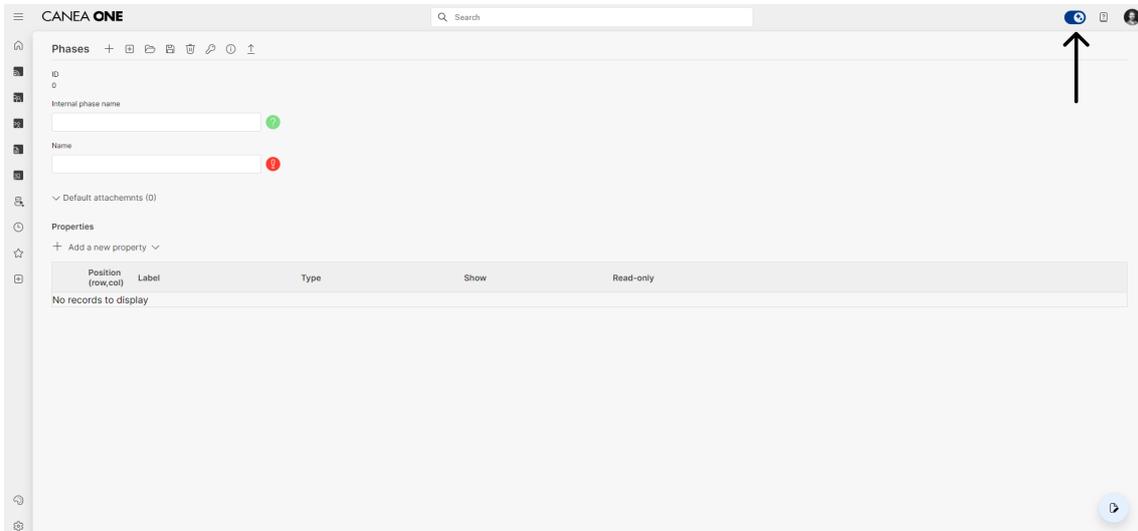


Figure 5.18: UAS

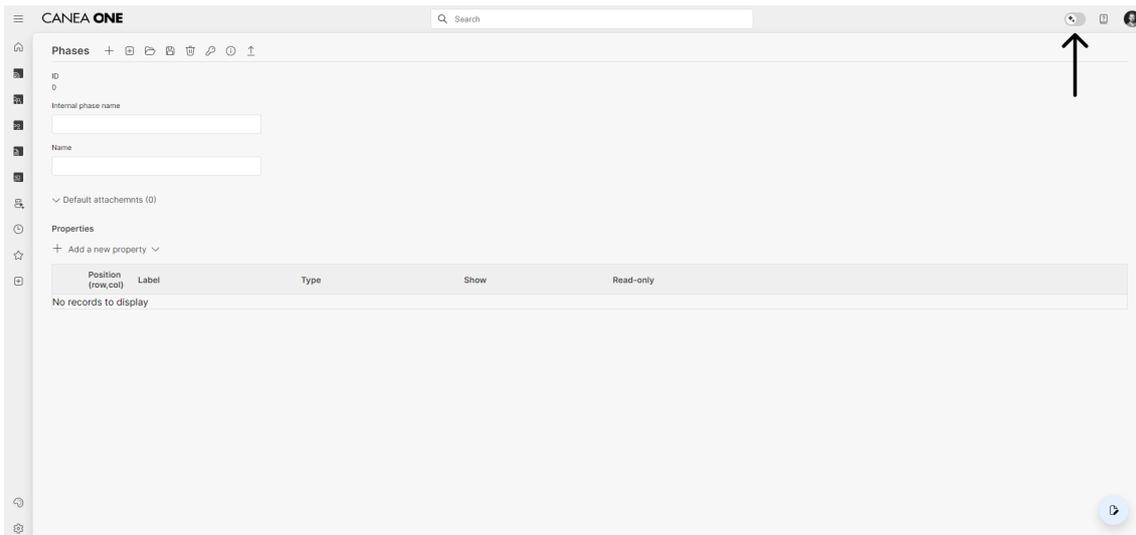


Figure 5.19: UAS Off

The UAS gives recommendations or warnings to the user in the interface visualised by a green question mark icon for the recommendations and a red exclamation icon for the warnings. To use the recommended change the user can click the Accept button on the message otherwise the user can close down the recommendation by clicking Cancel or the X icon to the top right of the message, see Figure 5.20.

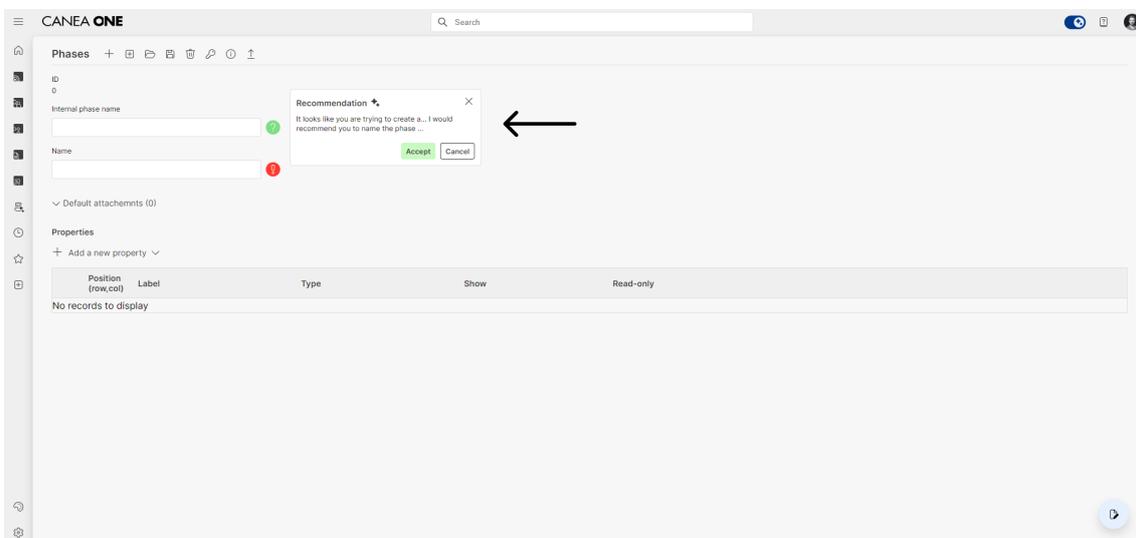


Figure 5.20: UAS Message

5.12 User Testing

To evaluate the functional prototype, user tests were conducted with a total of 6 participants. The same participants who were previously interviewed were invited to participate in the user tests, as they were already familiar with our work and interested in taking part in the study. However only four of those previously interviewed

5. Process and Execution

were available to do the user tests therefore two new participants were recruited as well. Since the chatbot needs to be used in conjunction with a functional version of CANEA ONE, the user testing was conducted on one computer with the functional prototype of the CANEA Chatbot and another computer with the real CANEA ONE software. The goal of the tests was to understand if users found the chatbot useful when completing a task and to assess how easy the interface was to use.

The participants were asked to carry out 3 tasks with help from the chatbot:

1. Make an absence report between the dates 9-19th of April.
2. Turn on subscriptions to email and subscribe to documents view "instructions".
3. Create a configuration of a phase, a case type and a case with the following instructions:
 - There was an accident outside the CANEA office yesterday, a car collided with a bicycle and we want you to create a new case with the case type "Accident" which contains a phase called accident registration. The phase must contain a description of the accident and the date.

The researchers sat next to the participant and observed in silence unless a question was asked. The tasks were also written on a whiteboard in front of the participants to make it more accessible, see Figure 5.21

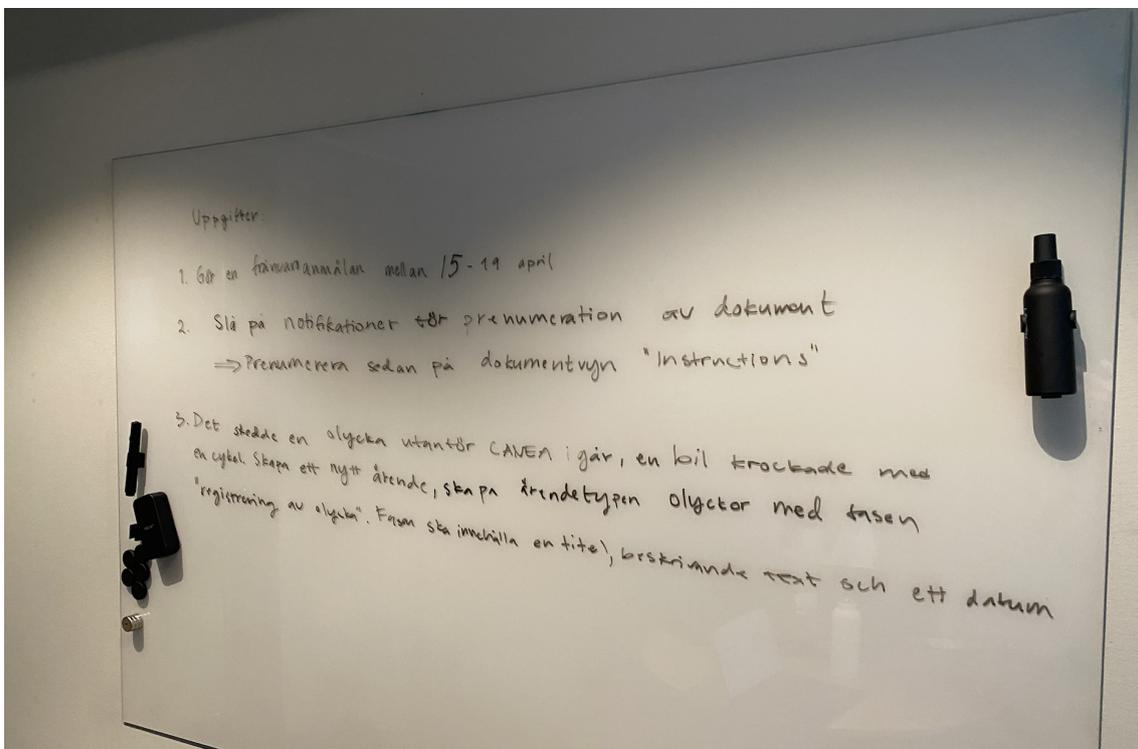


Figure 5.21: Tasks on a whiteboard

The user tests were performed on two computers: one computer with the functional

prototype and one computer with the existing CANEA ONE. Moreover, to emulate the real action of using the chatbot in the application (on the same screen) the participants were informed that when using the large chatbot interface they could not perform any tasks on the computer with the existing CANEA ONE, instead they could use the smaller chatbot window which is designed to always be accessible.

For half of the user tests the participants got a thorough walk-through of the functional AI chatbot and the rest were asked to perform the tests without any further explaining. This was done in order to test if the perceived ease of use/intuitiveness of the chatbot would change based on if the user got a "tutorial" beforehand and if so how big is the difference.

After performing the tasks the participants were asked followup questions that had been prepared in advance, see Appendix E.

5.13 Affinity Diagram User Tests

Miro was used to create an affinity diagram for the user tests. All user tests were voice recorded to be transcribed where important quotes from the user tests were written on post it notes. These post it notes could then be grouped into clusters forming different themes. The themes identified were: Functionality, Interface, UX, Trust, Bot or Not, Usability, UAS and Intuitiveness. The themes with the most post it notes were UX and user interface. See Figure 5.22 for reference of the cluster Bot or Not. The colors of the post it notes represents the different participants in the study. For example participant 1 has the color yellow so every quote on a yellow post it note is from participant 1.

5.13.1 User Tests Findings

The users were mostly positive toward the chatbot prototype. The interface was considered minimalist and well integrated into the application.

Several participants noted that the icon chosen for initiating a new chat was not clearly indicating its function. One participant mentioned that the icon is similar to the document edit icon used elsewhere in the application and could therefore be misunderstood as being connected to documents. Choosing a different icon that better aligns with users' expectations would enhance usability. Making sure that the icon is not used elsewhere in the application is important so that the chatbot feels seamlessly integrated with the application.

"I think it's a bit unclear that the icon for the chatbot is a document icon; I believe it's used elsewhere..." - Participant 2

Since users had to use the smaller version of the chatbot while completing tasks, some commented that it was too small or expressed a desire to adjust the size of the

5. Process and Execution

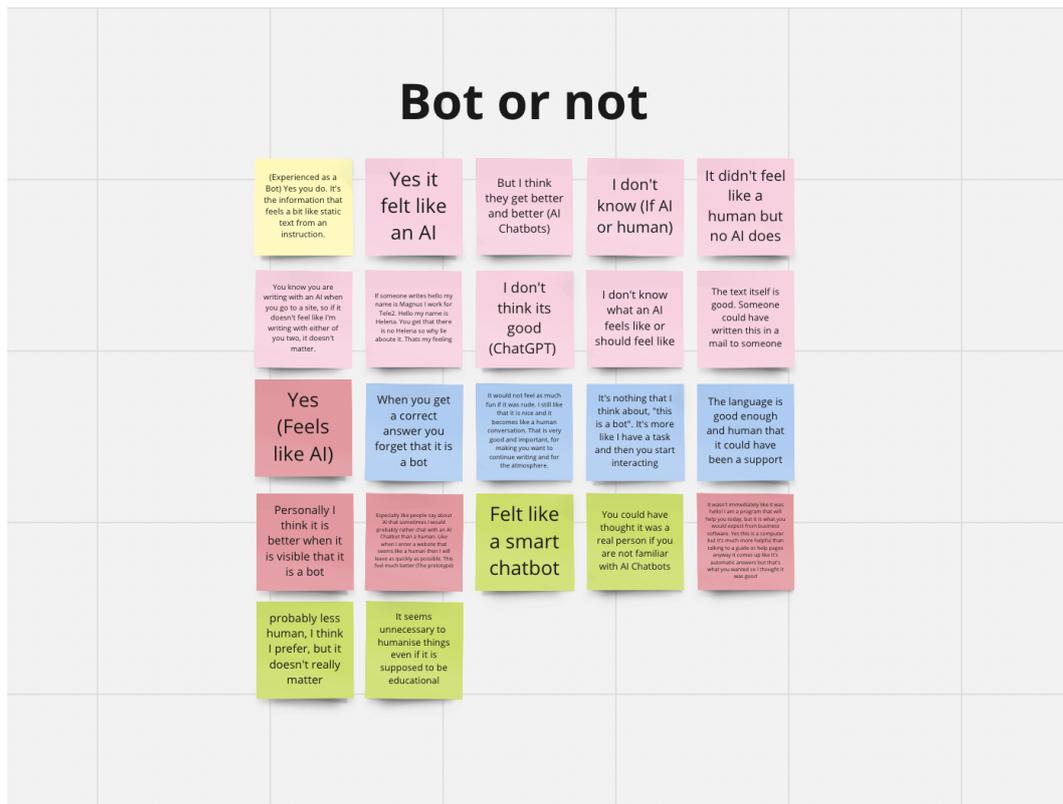


Figure 5.22: Example of the cluster Bot or Not in Miro

chatbot window. While some thought the smaller chatbot was a good size. Allowing users to customize the window size to their preference would be preferable, ensuring it suits a wider range of users.

"I think the small one works fine too. It's mostly the UI thing that you can adjust how big it is yourself." - Participant 1

Otherwise the minimized version of the chat was considered an important addition to enable users to view instructions at the same time as completing tasks in the interface. One of the participants noted that it would have been very annoying to move back and forth between the chat and the application window if there had not been a minimized version.

"It would have been really annoying if one had to go back and forth all the time." - Participant 6

Participant 6 described how current users who need assistance use the manual while performing tasks in CANEA ONE by having the manual in a separate tab on a web-browser.

"First, you check the manual, then you backtrack, and then you refer back to the manual." - Participant 6

One participant compared the chatbot to those commonly used on e-commerce websites. Many e-commerce platforms feature chatbots that automatically pop up when users visit the webpage, initiating a chat without their consent. Such intrusive behavior can lead to irritation and frustration in the user experience. The participant expressed a preference for a more discreet interface that only appears when the user chooses to engage with it.

"I think that's more annoying than when it's discreet and pops up when you want it to." - Participant 2

According to the respondents the interface was considered trustworthy, however the main factor affecting trust was the quality of the answers. When the participants found the answers to be helpful and clear, trust increased, while if some of the information was misleading, trust decreased. One of the participants mentioned that it is important for the chatbot to provide correct answers at the beginning of its use to build trust. Then, if the chatbot is wrong at some point, users will be more forgiving. The crucial aspect is that the chatbot is correct most of the time. The prototype in its current state has varying correct answers, which led to some users having a very good experience, while others received more false information and felt like the chatbot could not be trusted at all.

"The interface feels robust and reliable." - Participant 2

"If it responds with something that's wrong, then I trust it less." - Participant 3

"It's important in the beginning that you feel like you're doing the right things with what you've got, because then you feel like 'I trust this.' If you were to receive wrong answers, you become a bit skeptical and think, 'Do I really trust this?' But you still do, and then you overlook it when it's wrong." - Participant 5

When asked whether the prototype felt like an AI chatbot, most participants agreed that it did indeed feel like one, though they didn't necessarily view this as either positive or negative, it was merely an observation. One participant in particular noted that when the chatbot consistently provided accurate responses, they momentarily forgot it was a bot. Moreover, this participant valued conversational interactions that resembled human interaction, as it encouraged prolonged engagement and enhanced the overall atmosphere.

Several participants expressed that it was more disturbing when AI chatbots try to

convince users into thinking it is a human behind the answers by typing "Hi, my name is ..." as many chatbots that are used as webpage assistants use. Another participant appreciated that the chatbot was polite, saying "Hope you have a nice day" and similar phrases, which made the chatbot feel more human and friendly. This made the person feel like the conversation was more human-like and enjoyable.

Hello, my name is Helena. You understand that it's not really Helena, so why lie? That's just how I feel. - Participant 2

"It wouldn't have felt as fun if it was rude. I think it's polite, and it becomes like a human conversation." - Participant 5

When it comes to the design of the answers most of the respondents thought that bullet points made the steps easy to follow and appreciated this format.

"I like that it lists it in bullet points. When seeking help, you want some point guidance." - Participant 4

When it comes to the UAS, all participants had a positive attitude towards the interface and the concept. However, one participant expressed concern over the "accept" button on warnings, as it is unclear what accepting implies. Instead, they suggested only providing necessary information to allow users to make manual changes. Having an "accept" button on recommendations, however, was deemed acceptable since it isn't as critical as with warnings.

"The warning might be unfamiliar. Then one doesn't know how to handle it. One doesn't know what happens when clicking (accept)." - Participant 3

The respondents also appreciated the functionality of being able to easily turn on and off the UAS. This ensures that if the recommendations are annoying to a person they will not have to use them.

"The most important thing is that it's not there all the time. You have the option to toggle it here." - Participant 5

Taking the users' responses into consideration, minor changes were made to the interface. Since users were mostly satisfied with the interface, we did not find it necessary to make significant alterations to the interface that was already intuitive to use. The updated version of the interface can be viewed in section 6.4.

6

Results

In this chapter the results of the design process will be presented. The guidelines for designing a chatbot and UAS for business applications will be described, followed by the final updates of the prototype.

6.1 Problem Statement

CANEA ONE is a large and complex application with users of varying levels and purposes. Issues may arise in different sections of the program. Users with infrequent usage and no administrative roles typically focus on viewing documents and statuses within the document modules. These users commonly encounter navigational challenges and difficulties locating specific documents and information.

Administrators, tasked with configuring workflows or documents, face more intricate problems, such as designing processes using the "process designer" feature. Configuring can be challenging because each setup is unique, and there are many ways to do it. Simply relying on the manual isn't always enough, and asking colleagues for help might result in different answers. Users can refer to a manual or to others when faced with challenges however many experience that the information in the manual only covers what tools do but not how they should be used.

6.2 Guidelines for AI Chatbots

Based on the data collected and the findings of the evaluation, several guidelines have been established to serve as guidance when designing an AI chatbot. In this section the guidelines established will be presented together with a motivation. The motivation behind each guideline can be traced back to reflections and revelations found during literature reviews, questionnaires, interviews and user testing of the prototype.

6.2.1 Tool-tips

Not all users are familiar with using AI chatbots. To assist them in understanding parts of the interface that are new to them, functions should incorporate tool-tips to provide clarification. When icons lack accompanying text, users may feel confused or hesitant to interact with them, as they're unsure of the function's purpose. For

instance, the 'new chat' function proved to be such a case during user testing. The guidelines tool-tips was derived from the user tests 5.13 and literature review 3.2.

6.2.2 AI Personality

The preferred personality of the AI chatbot is very subjective. While most do not mind that the chatbot seems like an AI some users prefer when the AI incorporates human phrases and makes the chat feel more like a human conversation. Giving the user control of the personality would be preferred to enhance the personal experience of each user.

Additionally, it is crucial not to impose a sense of human-like identity onto the chatbot agent such as the chatbot saying "hello, it's [insert human name]", as this could result in an uncanny valley experience. Instead, the bot's identity should remain distinct and transparent. The guideline AI Personality was derived from user tests 5.13.

6.2.3 Usage of Icons

When using icons to describe functions it is important to consider the application since an icon elsewhere may have another meaning in the system. Do not take for granted the meaning of an icon and always test on several users before implementing. Additionally, icons should have pointers to indicate that they are pressable. The guideline Usage of Icons was derived from user tests 5.13.

6.2.4 Minimalist Design

A minimalist design highlights the important functions and helps the users focus on what is important without being distracted by the visuals. This is important to make sure that the interface dose not become cluttered when adding the chatbot into an existing interface. The quick access button should also be subtle so that users can ignore the chatbot if they do not want to use it. The guideline Minimalist Design was derived from literature studies 3.2

6.2.5 Size

When considering the size of the chatbot, it's essential to offer users the flexibility to keep it open while engaging in other tasks. Therefore, the chatbot should be compact enough to seamlessly fit within existing screens. The dimensions chosen for the prototype were 350px * 400px. User testing determined that the dimensions should not be reduced beyond this size. Another factor to consider when deciding on the size of the chatbot interface is the amount of scrolling the user will have to do. The guideline Size was derived from user tests 5.13

6.2.6 Small-Interface

The chatbot should have the capability to be minimized into a smaller version that remains accessible while users perform tasks. This feature streamlines navigation, allowing users to follow instructions more easily by keeping all relevant information visible on a single screen. The guideline Small-Interface was derived from user tests 5.13.

6.2.7 Perception of Control

The chatbot should never initiate contact with the user; instead, it should only appear when the user chooses to open it. If the chatbot appears unexpectedly, the user may feel irritated and surprised in a negative way. The user should also be able close down the chatbot in case they do not want it to be visible at all times. The guideline Perception of Control derived from literature studies 3.2.

6.2.8 Consistency

The design needs to have consistency throughout the entire application. This is true both for colors, fonts and icons to seamlessly integrate the chatbot into the application. An interface without consistency will throw the user off and will decrease the perceived seriousness and professionalism. Additionally, the placement of the chatbot should remain the same in all pages to make it predictable and accessible. The guideline Perception of Control derived from literature studies 3.2.

6.2.9 Complement Existing Solutions

An AI chatbot should complement existing solutions such as the manual and CANEA support and not replace them. Customers will still need to initially be introduced to the application and get educated on how it works. The AI chatbot will instead help maintain the system. Only administrators within a company have access to CANEA support which means that beginners can only rely on the manual or colleagues for support. The AI should serve as an additional alternative to offer users assistance that is quick and always available. The guideline Perception of Control derived from literature studies 3.2 and interviews 5.4.

6.2.10 Location of Chatbot

The location of the chatbot is highly crucial since it determines how accessible the chatbot is. The chatbot should ideally be accessible via the main menu, which may differ across systems. This ensures its visibility throughout the entire application. Moreover, if a smaller version of the chatbot is to be integrated, placing it at the bottom right of the page is advisable, as this aligns with the standard practice in many other applications. The guideline Location of Chatbot derived from literature studies 3.2 and from user tests 5.13.

6.2.11 Chatbot Response

Chatbot responses are more clear and easy to follow when steps are provided in bullet points. When providing steps this format should be used to enhance the readability of answers. The guidelines Chatbot Response derived from literature studies 3.2 and user tests 5.13.

6.2.12 Anticipation

The chatbot should always show the system status. When it is typing there should be a typing animation (for example three dots in a row) communicating to the user. This will increase user patience. The guideline Anticipation was derived from the literature study 3.3.

6.2.13 Accessible Design

The chatbot should be designed following the WCAG (Web Content Accessibility Guidelines). Some of these guidelines include contrast in color and font size. However, to ensure that the design is accessible, the product should be evaluated with these criteria in mind. The guideline Accessible Design was derived from literature studies 3.1.

6.2.14 Differentiating Chatbot and User Messages

To enhance readability and distinguish between user and chatbot messages, it's important to clearly differentiate them. This can be achieved by using icons or images for both the user and the chatbot, using different colors for each conversation, and incorporating clear titles. Having several of these factors ensures that the design is accessible and clear. The guideline Differentiating Chatbot and User Messages was derived from the literature studies 3.3 and confirmed during the user tests 5.13.

6.3 Guidelines for User Assisted System

Based on the data collected and the findings of the evaluation, several guidelines have been established to serve as guidance when designing an user assisted system. In this section the guidelines established will be presented together with a motivation.

6.3.1 Accessible Design

The status of the recommendation should not be indicated by only color since it is not accessible to the colorblind. Instead, they can be accompanied by an icon that further indicates the meaning behind the notification. The guideline Accessible Design was derived from literature studies 3.1.

6.3.2 Do not Disturb Workflow

The recommendations should not stop the users work, only act as additional help if the user wants it. If the user do not want to use it it should not affect the users ability to complete tasks as usual. The guideline Do not Disturb Workflow was derived from literature studies 3.6.2.

6.3.3 Design for Easy Deactivation

If recommendations are disruptive to users' experience, for example, an advanced user might not require recommendations to complete tasks quickly. Therefore, the recommendations should be easily deactivated for times when they are not needed. The same goes for activation that should be as accessible as deactivation. The guideline Design for Easy Deactivation was derived from the user tests 5.13.

6.3.4 Accepting and Canceling

The option to accept and cancel should always be available on recommendations. However, when it comes to warnings, this should be more carefully considered since the stakes are higher. Users may feel uncomfortable and reluctant with pressing accept if they are unsure about what will happen once they press the button. In this case it might be better to simply inform the user of the changes that needs to be made. The guideline Accepting and Canceling was derived from user tests 5.13.

6.4 Prototype Improvements

Based on the guidelines established in section 6.2 a final iteration of the prototype was developed. However, much of the design already adhered to the guidelines resulting in only minor adjustments. According to the guidelines the usage of icons had to be considered and therefore the icon for "New Chat" used was replaced. The previous icon was associated with documents and caused confusion and was replaced with the "compose" icon from fluent icons to not confuse the function with the document modal, see figure 6.1.

Additionally, tool-tips were added to all functions to clarify what each icon stands for since it is not always obvious. The tool-tips will also aid in making new users feel more confident in exploring the new functions, see figure 6.2.

Finally, the font-size for all bodytext was increased from 12px to 16px since it is more accessible and readable.

6. Results

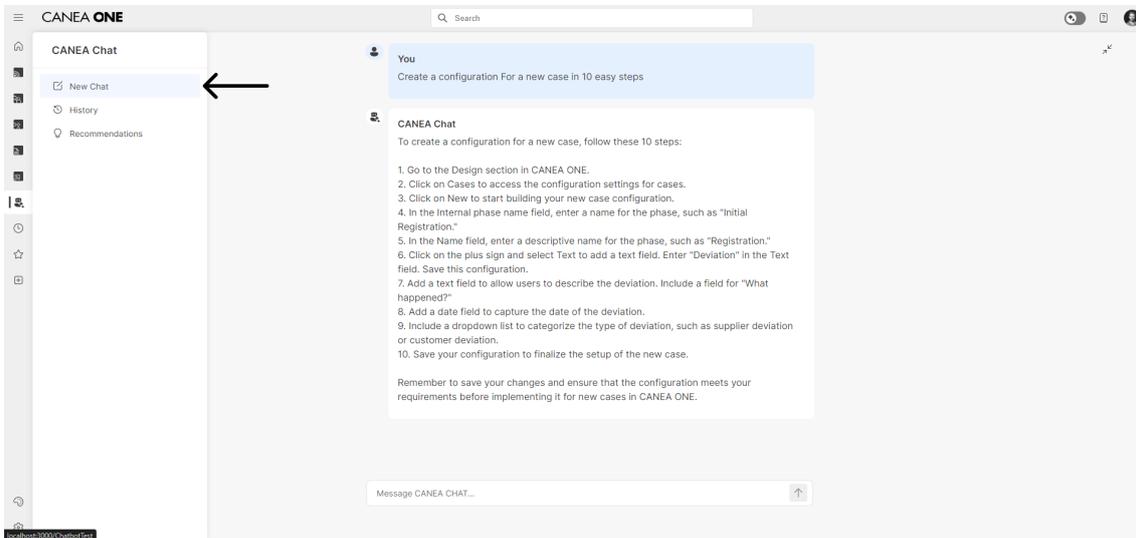


Figure 6.1: New Icon for New Chat

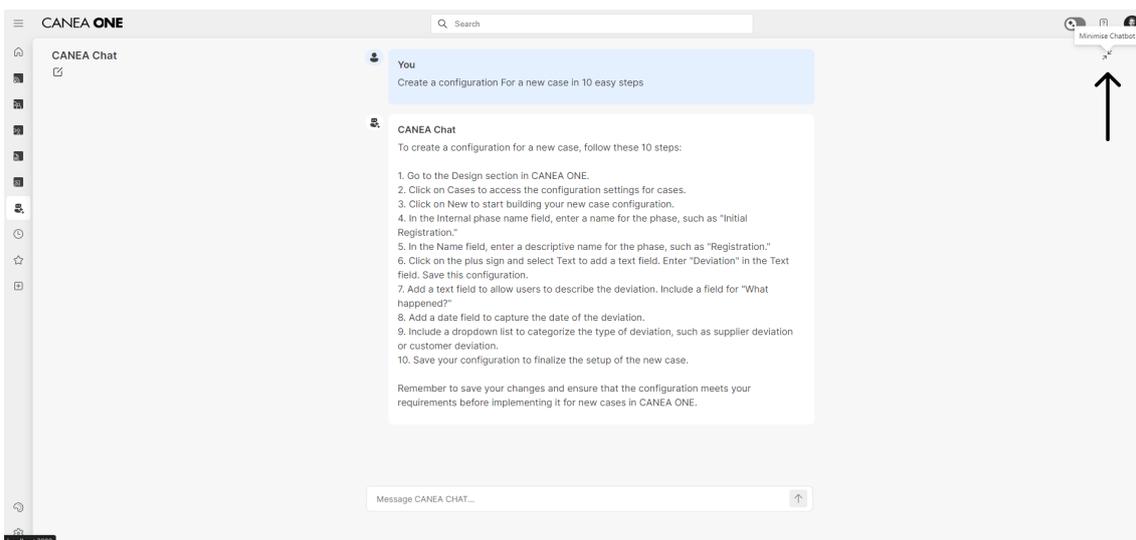


Figure 6.2: Tool-tips

7

Discussion

The discussion chapter of the report presents reflections of the final result as well as interesting aspects that could be considered for future works that were out of scope for this thesis.

7.1 Results

From the literature review it could be noticed that there was little to no existing research that mainly focused on the graphical interface of AI chatbots. Instead many research papers left out the interface with no discussion regarding any design process. The guidelines developed during this study serve as a step in the right direction for the design of AI chatbot interfaces especially for business applications. Moreover, the guidelines were made to be generic so that they could be applied for any chatbot within a system.

The UAS interface was more challenging to evaluate because it was not as functional as the chatbot. Therefore, it was not possible to test the UAS in conjunction with completing tasks as we did with the chatbot. The feedback on the UAS was solely based on the interface and users' thoughts on the design. This hindered the evaluation of how useful the UAS would be to users, and the guidelines are more based on users' perceived usefulness of the UAS.

7.1.1 AI Chatbot Trust

The results of the user tests validated that the interface of the chatbot gains user trust. However, this could become problematic if the chatbot provides incorrect information. This occurred during our user tests, highlighting the risks of relying on a chatbot that may not always provide accurate answers. If users are unaware that the chatbot can provide incorrect information, they may unknowingly follow misleading instructions from the chatbot. This presents a challenging issue because the objective is to ensure that the chatbot only provides answers based on the information it has been given and not give false information. In our prototype, it was difficult to control the chatbot and prevent it from generating false information, resulting in instances of misinformation. This is particularly significant for novice users of the system, as they may lack sufficient prior knowledge to discern whether the chatbot's responses are accurate, rendering them more susceptible to misinformation. Therefore it is important to inform users that the chatbot could be wrong sometimes to give them the right expectations and not blindly trust whatever

the chatbot says. Explainability and trust was mentioned in section 3.8 where the importance of explaining how AI works to help users know when they can trust the system and when they should use their own judgement is highlighted.

Additionally, it's important to inform users that the chatbot is not intended to replace the assistance of CANEA employees or colleagues; instead, it should serve as the first step when encountering an issue in the application.

7.1.2 System Complexity

It's important to recognize that CANEA ONE is a vast application, so throughout this project, it was not possible to cover every aspect of it. The application's usage varies significantly based on users' needs, making usability testing challenging. For instance, in our testing, we assigned tasks for participants to complete, one of which involved creating configurations for a fictional accident scenario. The necessity for new configurations depends heavily on the context and scenario, as users must add properties like text fields to elements. Often, specific terminology or knowledge, such as ISO standards, is required, which is difficult to simulate because the authors lack such expertise.

Configurations are intricate because they're interconnected and rely on previous steps. For example, to configure a case type, users must first configure phases, and configuring phases might require setting up case roles beforehand. These functions are spread across different pages and must be done separately. In summary, configuration is a complex process that can span hours or even days.

7.1.3 Chatbot Humanness & Gender

When designing the chatbot, it was unclear whether to make it appear more human or more like a machine. Various research had conflicting conclusions; some studies presented evidence that human-like traits are appreciated, while others argued that human-likeness is deceptive and that bots should remain robotic. The prototype that underwent user testing was a compromise, using human-like conversation bubbles while displaying machine elements through cues such as a robotic name, CANEA Chat, and a robotic avatar. Like the literature review, the user testing yielded divergent results. Preferences are highly subjective, so neither a fully bot-like nor a fully human-like approach is one hundred percent correct. Instead, there are positive and negative aspects to both, and the choice depends on the context.

The chatbot was intentionally designed to be gender-neutral to counteract the prevailing practice of defaulting chatbots to a female gender, a tendency that can have detrimental consequences. This was done by giving the chatbot a gender neutral name CANEA Chat, a robot avatar and a conversational background colour of white. Although all these measurements were taken it is unclear whether the chatbot is perceived as gender neutral or not since it may very much be affected by the context of the business application or company branding.

Furthermore, it's crucial to recognise the ethical complexities associated with gender design and anthropomorphism in AI. Assigning a female gender to an AI can perpetuate harmful gender stereotypes, portraying women in traditional roles such as that of an assistant. Moreover, it may lead to objectification of women, reducing them to mere tools for fulfilling customer needs. While some studies [38] suggest that developing chatbots with feminine traits like warmth can enhance their human-like qualities, we must question whether it's appropriate to humanise bots at the expense of dehumanising females.

7.2 Methods

In the following sections, some of the methods used and the results of utilizing these methods will be discussed. Additionally, areas that might be lacking and could be improved upon will be addressed.

7.2.1 Questionnaires

There were a total of two questionnaires sent out: one to the employees of CANEA and one to a customer group of CANEA. The sample size of the questionnaire was 51 and the company has roughly 100 employees which means we had quite a high response rate. When it comes to replies from customers on the other hand, we only had responses from 23 participants and the population size is unknown which makes it hard to draw accurate conclusions with reliable data. The results of this project are based on the responses from employees of CANEA who use the application CANEA ONE. However it is not certain that the customers have a similar experience to the employees.

7.2.2 Interviews

During the interviews, participants were frequently prompted to elucidate the reasons behind their difficulties with certain tasks. Many users identified the program's interface as a significant factor contributing to their challenges. For instance, within CANEA ONE, one tool in particular is known by many employees to suffer from bugs and is perceived as outdated in terms of interface design.

Additionally, many navigational issues stem from the layout of the program and its placement of functions. These problems are inherent to the application and need to be fixed to solve the main issues. While the integration of AI could potentially alleviate navigational problems, having an intuitive design would undoubtedly provide significant assistance.

7.2.3 User Group

The application has many different types of users, varying not only in skill level but also in the modules they use. CANEA ONE offers five different modules, and

companies can choose which ones to use. Since some explicitly use only two of the modules, it is difficult to pinpoint the needs of a typical user.

The primary user group for CANEA ONE comprises the company's customers, making them the key beneficiaries of any proposed design solutions. Regrettably, during the project, our interactions with customers were limited, hindering both data collection in the initial phase and user testing in the final phase. Consequently, our primary target group throughout the project development has been CANEA employees—encompassing both developers and end-users. The insights gathered from these employees have proven invaluable. However, it is crucial to recognize that customer challenges might differ, and a more diverse set of perspectives, obtained through interviews with customers, could have provided additional valuable insights.

Moreover, the age range of CANEA ONE users spans a wide spectrum. Questionnaire results indicate that users within the company fall between the ages of 21 to 60 years and older. Recognizing the diversity in age groups is essential, as different preferences regarding the design and functionality of the AI chatbot may arise across various age brackets.

7.2.4 User testing

One recognised reason why users may turn to interacting with chatbots is the perceived absence of judgment compared to interactions with humans. Chatbots provide a non-threatening environment where users feel comfortable asking questions without the fear of being judged or feeling incompetent. This anonymity and lack of social pressure can be particularly appealing, especially when seeking information or assistance on sensitive or personal topics.

However, during the user testing phases of the project, the researchers closely observed participants' behavior as they engaged with the chatbot interface. It became evident that the mere presence of researchers monitoring the interactions could influence how participants interacted with the chatbot. Some participants may have altered their behavior, consciously or subconsciously, to avoid appearing ignorant or making mistakes in front of the observers.

This may have led to skewed or distorted results in the user testing. The data collected may not accurately reflect how users would naturally engage with the chatbot in a real-world setting where there are no observers.

Another factor that contributed to the findings of the user test was the time span of the testing. Most of the user tests were conducted over the span of 1 hour, in which the user did not always interact with the chatbot. This means that they did not get the chance to completely familiarise themselves with it. For example one function that users did not find intuitive or did not discover very quickly was the function where you had to hover on the chat messages in order to copy or edit a message. If the tests were done during a longer period of time such as letting the users use the chatbot in their own natural setting while documenting in a diary they

may have had a different opinion/experience. Additionally, these functions may not be of interest to beginner users and then a more scaled down interface with less functions is a positive thing.

Additionally, user tests and interviews were conducted with users of CANEA ONE that considered themselves at least intermediate users of the application. Consequently, our conclusions are applicable solely to the chatbot's performance for intermediate and advanced users. However, during the user tests, intermediate participants were tasked with completing two beginner-level tasks while utilizing the AI chatbot at all times, which they accomplished without any difficulties. This observation supports the assertion that the chatbot is suitable for beginners as well. Nevertheless, it's crucial to acknowledge that beginners may encounter other challenges that were not discovered. Since the chatbot was helpful to intermediate users, it's reasonable to assume it could also assist beginners, who may be tasked with more basic assignments that don't involve complex configurations, which are among the system's more challenging aspects.

7.3 Future Development

This project mainly focused on the interface of the chatbot. However there are other factors that affect the interaction between the chatbot and user such as the quality of the replies of the chatbot. For more accurate usability testing on the interface the quality of response from the chatbot has to be higher since it distracts the user from focusing on the interface part of the user experience. Additionally, the chatbot occasionally started to hallucinate, which means that the chatbot generated fabricated information but presented it as if it was correct which also caused disturbance to the user test.

Another factor that affected the results of the studies is that the AI was not contextually aware, instead it could only answer simple questions if the information has been provided in the system message. However, many users expressed wanting the chatbot to be able to answer general questions, helping them find information and helping them find objects. But because the AI was not contextually aware it was not possible to create an interface where this was possible. Therefore, further development could put more of a focus on the design of the response of the chatbot. For example "How could navigational information be presented?" and "how would the chatbot help users navigate to their desired destination other than in text format". One frequent response from the surveys when asked "How can an AI chatbot in CANEA ONE assist you?" was "Give me recommendations" which strengthens the importance of presenting recommendations in a way that both captures the attention of the user but also is a of relevance. In other words, the recommendations should not be static but should be tailored to the users specific needs in order to provide a better user experience.

This thesis explored a first iteration of how the user interface for an AI chatbot integrated into CANEA ONE could look like. However, for further development

iterations based on user feedback, development of these functions and user testing would further improve the user experience. For example during the first round of user tests many expressed the wish to be able to move the smaller chatbot around. This featured might want to be explored since the smaller chatbot was never really tested when using the existing CANEA ONE and could potentially be in the way of the underlying interface and cause disturbance. Being able to move the window of the chatbot could solve this problem.

When designing the interface the authors wanted to consider accessibility and therefore tried to adhere to the web accessibility guidelines (WAG). One example of this is that the containers for the messages inside of the chatbot are distinguishable by colour, icon and text. Our toggle icon for the UAS has an acceptable contrast according to WAG and the notification system for the UAS uses both icon and colour to convey the status.

Additionally, the user experience of the UAS system relied heavily on the functionality which had not been developed. Thus, the user testing only led to surface level interactions such as toggling on and off and opening/closing the notification modals.

In the project, the prototype was only tested for accessibility by using online contrast checkers to ensure the contrast between the background and text, as well as the main visual elements, was accessible. We have also ensured that there are several channels to convey information by utilising both text, symbols and color and not solely relying on color which could be problematic for people who are colorblind. However, there have not been tests to ensure the chatbot is accessible, for example, for visually impaired people. Further user testing and development would be crucial to continue developing the chatbot and ensure that it is accessible. For example, implementing text-to-speech and similar features would improve accessibility further.

Furthermore, adding a function such as text-to-speech would require more thought as to what kind of voice the chatbot would adopt since it would affect the gender design of the chatbot, as previously discussed in section 3.4. This would in turn require more user testing in order to study how it affects the user trust in the chatbot. Since the preferred gender of a chatbot is linked to the context of use it would be interesting to evaluate if users perceive CANEA ONE to be more masculine or feminine and if this has any impact on the information that is presented.

This project is a first step towards establishing guidelines for designing AI chatbot and UAS interfaces. However, there is still more research needed to refine and expand the guidelines created. We encourage further exploration to deepen our understanding of designing AI chatbots and assistants in businesses applications.

Additionally, we recommend that future studies place a strong emphasis on enhancing the accessibility of chatbots, which has the potential to make AI chatbots significantly more user-friendly and inclusive to a broader audience. By investing resources in accessibility, organizations can not only better serve users with specific

needs but also improve the overall experience and usability for all users, thereby maximizing the impact and reach of AI chatbot technologies.

8

Conclusion

This thesis aimed to study how the the graphical interface influences AI chatbots in business applications. The purpose was to design a graphical interface of an interface for the business application CANEA ONE. The research questions guiding the research are as follows:

- How can we design the user interface of AI Chat-bots and user assisted systems in the field of Business Applications to improve customer acceptance, use, and satisfaction?
- What are the key considerations and challenges in designing AI Chat-bots and user assisted systems in the field of Business Applications?

To answer the research question a user-centered design process using design sprint methodology was followed whilst adapting it to the business application CANEA ONE. The process was initialized with a literature review collecting relevant information of design guidelines and previous and similar solutions. This was followed with quantitative and qualitative data collection through surveys and interviews. Later low fidelity prototypes were created in Figma. Stakeholders were included in the design process through a co-design workshop where new low fidelity prototypes and storyboards were created. In order to evaluate the prototype it was developed into a high fidelity prototype using React Typescript, Flask and OpenAI API. The prototpe was evaluated through user testing where users used the chatbot to complete tasks in the CANEA ONE application. With the feedback from the users guidelines were created for how to design chatbots and UAS for business applications. The following section outlines the guidelines that have been developed. For a more detailed description of the guidelines, see section 6.2.

These are the guidelines created for chatbots:

- **Tool-tips** - Give explanations to functions that may be new to users.
- **AI Personality** - Make it clear to the user that the chatbot is an AI and do not pretend that it is human.
- **Usage of Icons** - Make sure Icons are clear to users and not used elsewhere

in the application.

- **Minimalist Design** - Keep the design as minimal as possible to decrease clutter in the application.
- **Size** - Do not make the chatbot smaller than 350px * 400px to ensure readability.
- **Small-interface** - There should be a small version of the chatbot that is accessible in all pages of the application.
- **Perception of Control** - The chatbot should only write to the user after the user has contacted it first.
- **Consistency** - The chatbot design should be consistent with the existing application in colors, fonts and icons.
- **Complement Existing Solutions**, - The aim of the chatbot is to compliment existing solutions such as customer support and manual not to replace these.
- **Location of Chatbot**, - The chatbot should be easily accessible trough either the main menu or at the bottom right of the screen.
- **Chatbot Response** - Responses should be offered in bullet points when steps are discribed.
- **Anticipation** - The status of the chatbot shuld be visualised to indicate it's status to the user.
- **Accessible Design** - The design should follow the WCAG to ensure accessibility.
- **Differentiating Chatbot and User Messages** - The design should clearly indicate wich messages are from the user and which are from the Chatbot.

These are the guidelines created for UAS:

- **Accessible Design** - Utilise both color and icons to visualise recommendations.
- **Do not Disturb Workflow** - Recommendations should not affect the users ability to complete tasks without utilising the recommendations.
- **Design for Easy Deactivation** - The recommendations should be easy to deactivate and activate.

The guidelines uncovered in this thesis significantly contribute to the topic of crafting user-friendly chatbots and UAS for business systems. Applying these guidelines during the development of chatbots and UAS within business applications contribute

8. Conclusion

to enhancing their usability and user experience. These guidelines are also intended to inspire and support further research and development of designing chatbots and UAS. Consequently, these findings effectively address the research questions posed in this study.

Bibliography

- [1] Lorentsa Gkinko and Amany Elbanna. The appropriation of conversational ai in the workplace: A taxonomy of ai chatbot users. *International Journal of Information Management*, 69:102568, 2023.
- [2] Elitza Stoilova. Ai chatbots as a customer service and support tool. *ROBONOMICS: The Journal of the Automated Economy*, 2:21–21, 2021.
- [3] E Nichifor, A Trifan, and EM Nechifor. Artificial intelligence in electronic commerce: Basic chatbots and the consumer journey. *amfite-atru economic*, 23 (56), 87–101, 2021.
- [4] Chinedu Wilfred Okonkwo and Abejide Ade-Ibijola. Chatbots applications in education: A systematic review. *Computers and Education: Artificial Intelligence*, 2:100033, 2021.
- [5] Canea Consulting Group. <https://www.canea.com/>. Accessed on: 2024-05-13.
- [6] CANEA Strategy Activation. Canea strategy, Access date: 2024-01-24.
- [7] CANEA Project Management, Access date: 2024-01-24.
- [8] CANEA Document Management, Access date: 2024-01-24.
- [9] CANEA Workflow Case Management, Access date: 2024-01-24.
- [10] CANEA Process. Canea process modeling, Access date: 2024-01-24.
- [11] Therese Fessenden. Design systems 101, April 2021. Published on April 11, 2021.
- [12] Aishwarya Gupta, Divya Hathwar, and Anupama Vijayakumar. Introduction to ai chatbots. *International Journal of Engineering Research and Technology*, 9(7):255–258, 2020.
- [13] Veena Chattaraman, Wi-Suk Kwon, Juan E Gilbert, and Kassandra Ross. Should ai-based, conversational digital assistants employ social-or task-oriented interaction style? a task-competency and reciprocity perspective for older adults. *Computers in Human Behavior*, 90:315–330, 2019.
- [14] Ssu Chiu, Maolin Li, Yen-Ting Lin, and Yun-Nung Chen. Salesbot: Transitioning from chit-chat to task-oriented dialogues. *arXiv preprint arXiv:2204.10591*, 2022.

- [15] Santiago Melián-González, Desiderio Gutiérrez-Taño, and Jacques Bulchand-Gidumal. Predicting the intentions to use chatbots for travel and tourism. *Current Issues in Tourism*, 24(2):192–210, 2021.
- [16] Petter Bae Brandtzaeg and Asbjørn Følstad. Why people use chatbots. In *Internet Science: 4th International Conference, INSCI 2017, Thessaloniki, Greece, November 22-24, 2017, Proceedings 4*, pages 377–392. Springer, 2017.
- [17] Petter Bae Brandtzaeg and Asbjørn Følstad. Chatbots: changing user needs and motivations. *interactions*, 25(5):38–43, 2018.
- [18] Ana Paula Chaves and Marco Aurelio Gerosa. How should my chatbot interact? a survey on social characteristics in human–chatbot interaction design. *International Journal of Human–Computer Interaction*, 37(8):729–758, 2021.
- [19] Bei Luo, Raymond YK Lau, Chunping Li, and Yain-Whar Si. A critical review of state-of-the-art chatbot designs and applications. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 12(1):e1434, 2022.
- [20] Eric WT Ngai, Maggie CM Lee, Mei Luo, Patrick SL Chan, and Tenglu Liang. An intelligent knowledge-based chatbot for customer service. *Electronic Commerce Research and Applications*, 50:101098, 2021.
- [21] Molly Follette Story. Maximizing usability: the principles of universal design. *Assistive technology*, 10(1):4–12, 1998.
- [22] Wendy L Martinez. Graphical user interfaces. *Wiley Interdisciplinary Reviews: Computational Statistics*, 3(2):119–133, 2011.
- [23] Bernard J Jansen. The graphical user interface. *ACM SIGCHI Bulletin*, 30(2):22–26, 1998.
- [24] Antti Oulasvirta, Niraj Ramesh Dayama, Morteza Shiripour, Maximilian John, and Andreas Karrenbauer. Combinatorial optimization of graphical user interface designs. *Proceedings of the IEEE*, 108(3):434–464, 2020.
- [25] Klaus Hinum and Silvia Miksch. Human centred design for graphical user interfaces. *Master’s thesis, Technical University of Vienna, Austria*, 2004.
- [26] Kelley Gordon. 5 principles of visual design in ux, March 1 2020.
- [27] Radka Nacheva. Principles of user interface design: important rules that every designer should follow. *Izvestia Journal of the Union of Scientists-Varna. Economic Sciences Series*, 1:140–149, 2015.
- [28] Raluca Budiu. Fitts’s law and its applications in ux, July 2022. Published on July 31, 2022.
- [29] Web content accessibility guidelines (wcag) 2.1, June 2018.
- [30] Terry Sullivan and Rebecca Matson. Barriers to use: usability and content accessibility on the web’s most popular sites. In *Proceedings on the 2000 conference on Universal Usability*, pages 139–144, 2000.

-
- [31] Stefan Leuthold, Javier A Bargas-Avila, and Klaus Opwis. Beyond web content accessibility guidelines: Design of enhanced text user interfaces for blind internet users. *International Journal of Human-Computer Studies*, 66(4):257–270, 2008.
- [32] Raluca Budiu. The user experience of customer-service chat: 20 guidelines, January 2019.
- [33] Jakub Złotowski, Diane Proudfoot, Kumar Yogeeswaran, and Christoph Bartneck. Anthropomorphism: opportunities and challenges in human–robot interaction. *International journal of social robotics*, 7:347–360, 2015.
- [34] Adam Waytz, John Cacioppo, and Nicholas Epley. Who sees human? the stability and importance of individual differences in anthropomorphism. *Perspectives on Psychological Science*, 5(3):219–232, 2010.
- [35] Pedro Costa. Conversing with personal digital assistants: On gender and artificial intelligence. *Journal of Science and Technology of the Arts*, 10(3):59–72, 2018.
- [36] Johannes Feine, Ulf Gnewuch, Stefan Morana, and Alexander Maedche. Gender bias in chatbot design. In *Chatbot Research and Design: Third International Workshop, CONVERSATIONS 2019, Amsterdam, The Netherlands, November 19–20, 2019, Revised Selected Papers*, volume 3, pages 79–93. Springer, 2020.
- [37] Mark West, Rebecca Kraut, and Han Ei Chew. I’d blush if i could: closing gender divides in digital skills through education. 2019.
- [38] Sylvie Borau, Tobias Otterbring, Sandra Laporte, and Samuel Fosso Wamba. The most human bot: Female gendering increases humanness perceptions of bots and acceptance of ai. *Psychology & Marketing*, 38(7):1052–1068, 2021.
- [39] Kurt Gray and Daniel M Wegner. Feeling robots and human zombies: Mind perception and the uncanny valley. *Cognition*, 125(1):125–130, 2012.
- [40] Peter AM Ruijten, Antal Haans, Jaap Ham, and Cees JH Midden. Perceived human-likeness of social robots: testing the rasch model as a method for measuring anthropomorphism. *International Journal of Social Robotics*, 11:477–494, 2019.
- [41] Marian McDonnell and David Baxter. Chatbots and gender stereotyping. *Interacting with Computers*, 31(2):116–121, 2019.
- [42] Clifford Nass, Youngme Moon, and Nancy Green. Are machines gender neutral? gender-stereotypic responses to computers with voices. *Journal of applied social psychology*, 27(10):864–876, 1997.
- [43] Kimberly E Culley and Poornima Madhavan. A note of caution regarding anthropomorphism in hci agents. *Computers in Human Behavior*, 29(3):577–579, 2013.
- [44] David Leslie. Understanding artificial intelligence ethics and safety. *arXiv preprint arXiv:1906.05684*, 2019.

- [45] Alexander Maedche, Stefan Morana, Silvia Schacht, Dirk Werth, and Julian Krumeich. Advanced user assistance systems. *Business & Information Systems Engineering*, 58:367–370, 2016.
- [46] Michael Schelkle and Christian Karl Grund. Identifying design features to increase the acceptance of user assistance systems: Findings from a business information visualization context. In *13th International Conference on Design Science Research in Information Systems and Technology (DESRIST 2018)*, 2018.
- [47] Florian Lehmann and Daniel Buschek. Examining autocompletion as a basic concept for interaction with generative ai. *i-com*, 19(3):251–264, 2021.
- [48] Google PAIR. People ai guidebook, 2019.
- [49] D Harrison McKnight and Norman L Chervany. What is trust? a conceptual analysis and an interdisciplinary model. 2000.
- [50] Xusen Cheng, Xiaoping Zhang, Jason Cohen, and Jian Mou. Human vs. ai: Understanding the impact of anthropomorphism on consumer response to chatbots from the perspective of trust and relationship norms. *Information Processing & Management*, 59(3):102940, 2022.
- [51] Asbjørn Følstad, Cecilie Bertinussen Nordheim, and Cato Alexander Bjørkli. What makes users trust a chatbot for customer service? an exploratory interview study. In *Internet Science: 5th International Conference, INSCI 2018, St. Petersburg, Russia, October 24–26, 2018, Proceedings 5*, pages 194–208. Springer, 2018.
- [52] Peter A Hancock, Deborah R Billings, Kristin E Schaefer, Jessie YC Chen, Ewart J De Visser, and Raja Parasuraman. A meta-analysis of factors affecting trust in human-robot interaction. *Human factors*, 53(5):517–527, 2011.
- [53] Aleksandra Przegalinska, Leon Ciechanowski, Anna Stroz, Peter Gloor, and Grzegorz Mazurek. In bot we trust: A new methodology of chatbot performance measures. *Business Horizons*, 62(6):785–797, 2019.
- [54] Chadia Abras, Diane Maloney-Krichmar, Jenny Preece, et al. User-centered design. *Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications*, 37(4):445–456, 2004.
- [55] Tomasz Miaskiewicz and Kenneth A Kozar. Personas and user-centered design: How can personas benefit product design processes? *Design studies*, 32(5):417–430, 2011.
- [56] Claudio Dell’Era and Paolo Landoni. Living lab: A methodology between user-centred design and participatory design. *Creativity and Innovation Management*, 23(2):137–154, 2014.
- [57] Travis Lowdermilk. *User-centered design: a developer’s guide to building user-friendly applications*. " O’Reilly Media, Inc.", 2013.

-
- [58] Richard Banfield, C Todd Lombardo, and Trace Wax. *Design sprint: A practical guidebook for building great digital products.* " O'Reilly Media, Inc.", 2015.
- [59] Google. Design sprint kit. Accessed on 17-01-2024.
- [60] Jeffrey W Knopf. Doing a literature review. *PS: Political Science & Politics*, 39(1):127–132, 2006.
- [61] Bill Gillham. Developing a questionnaire. 2008.
- [62] Siddegowda Roopa and Menta Satya Rani. Questionnaire designing for a survey. *Journal of Indian Orthodontic Society*, 46(4_suppl1):273–277, 2012.
- [63] Petra Lietz. Research into questionnaire design: A summary of the literature. *International journal of market research*, 52(2):249–272, 2010.
- [64] M. Rosala and K. Pernice. User interviews 101, 2023. Accessed: 15/1-2024.
- [65] Hamza Alshenqeeti. Interviewing as a data collection method: A critical review. *English linguistics research*, 3(1):39–45, 2014.
- [66] Nick Fox. Using interviews in a research project. *The NIHR RDS for the East Midlands/Yorkshire & the Humber*, 26, 2009.
- [67] Jennifer Rowley. Conducting research interviews. *Management research review*, 35(3/4):260–271, 2012.
- [68] Greg Guest, Arwen Bunce, and Laura Johnson. How many interviews are enough? an experiment with data saturation and variability. *Field methods*, 18(1):59–82, 2006.
- [69] Janice M Morse. Designing funded qualitative research. 1994.
- [70] Tomasz Miaskiewicz, Tamara Sumner, and Kenneth A Kozar. A latent semantic analysis methodology for the identification and creation of personas. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 1501–1510, 2008.
- [71] Khai N Truong, Gillian R Hayes, and Gregory D Abowd. Storyboarding: an empirical determination of best practices and effective guidelines. In *Proceedings of the 6th conference on Designing Interactive systems*, pages 12–21, 2006.
- [72] Corrie Van der Lelie. The value of storyboards in the product design process. *Personal and ubiquitous computing*, 10:159–162, 2006.
- [73] Shanna R Daly, Colleen M Seifert, Seda Yilmaz, and Richard Gonzalez. Comparing ideation techniques for beginning designers. *Journal of Mechanical Design*, 138(10):101108, 2016.
- [74] Google. Crazy 8's. Accessed on 2024-01-15.
- [75] Mind Tools. Brainwriting. Accessed on 2024-01-15.
- [76] Marc Steen, Menno Manschot, and Nicole De Koning. Benefits of co-design in service design projects. *International journal of design*, 5(2), 2011.

- [77] Carlye A Lauff, Daria Kotys-Schwartz, and Mark E Rentschler. What is a prototype? what are the roles of prototypes in companies? *Journal of Mechanical Design*, 140(6):061102, 2018.
- [78] Figma. How you design, align, and build matters. Do it together with Figma. Accessed on 2024-01-15.
- [79] Miriam Walker, Leila Takayama, and James A Landay. High-fidelity or low-fidelity, paper or computer? choosing attributes when testing web prototypes. In *Proceedings of the human factors and ergonomics society annual meeting*, volume 46, pages 661–665. Sage Publications Sage CA: Los Angeles, CA, 2002.
- [80] Jim Rudd, Ken Stern, and Scott Isensee. Low vs. high-fidelity prototyping debate. *interactions*, 3(1):76–85, 1996.
- [81] G. Brockman, A. Eleti, E. Georges, J. Jang, L. Kilpatrick, R. Lim, L. Miller, and M. Pokrass. Introducing chatgpt and whisper apis. Accessed on 2024-01-15.
- [82] Fabio Staiano. *Designing and Prototyping Interfaces with Figma: Learn essential UX/UI design principles by creating interactive prototypes for mobile, tablet, and desktop*. Packt Publishing Ltd, 2022.
- [83] Robert E Roth, David Hart, Rashauna Mead, and Chloë Quinn. Wireframing for interactive & web-based geographic visualization: designing the noaa lake level viewer. *Cartography and Geographic Information Science*, 44(4):338–357, 2017.
- [84] Data Source Triangulation. The use of triangulation in qualitative research. In *Oncol nurs forum*, volume 41, pages 545–7, 2014.
- [85] Craig Plain. Build an affinity for kj method. *Quality Progress*, 40(3):88, 2007.
- [86] Jakob Nielsen. How to conduct a heuristic evaluation. *Nielsen Norman Group*, 1(1):8, 1995.
- [87] Jakob Nielsen. Usability inspection methods. In *Conference companion on Human factors in computing systems*, pages 413–414, 1994.
- [88] Limin Fu, Gavriel Salvendy, and Lori Turley. Effectiveness of user testing and heuristic evaluation as a function of performance classification. *Behaviour & information technology*, 21(2):137–143, 2002.
- [89] Nielsen Norman Group. Usability testing 101, 2019. Accessed: 15/1-2024.
- [90] Gunnar Harboe and Elaine M Huang. Real-world affinity diagramming practices: Bridging the paper-digital gap. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*, pages 95–104, 2015.
- [91] CANEA. Your toolbox for more efficient operations management, Access date: 2024-01-24.

A

Timeline

The timeline below describes the working process that was followed during the thesis work and can mainly be divided into the phases of the design sprint: *Understand & Define, Sketch & Decide, Prototype & Validate*. The last phase *report writing & presentation* was not a part of the design sprint but an important part of the work.

A.1 Time plan

| Understand & Define | | |
|---------------------|--|--|
| Week | Task | Methods |
| 3 | Literature Study | Google scholar |
| 4 | Literature Study | Google scholar |
| 5 | Literature Study prepare questionnaire Send out questionnaire Prepare interview questions | Google forms Google drive |
| 6 | Literature Study | Google scholar |
| 7 | Conduct interviews Transcribing interviews Formulate problem | Semi-structured interviews Miro Google drive |

A. Timeline

| Sketch & Decide | | |
|-----------------|---------------------------------|--|
| Week | Task | Methods |
| 8 | Ideation | Crazy-eights Brainwriting Co-design workshop |
| 9 | Prototyping | Figma |
| 10 | Prototyping | Figma |
| 11 | Evaluation Affinity diagrams | User testing Miro |
| 12 | Evaluation Affinity diagrams | Heuristic evaluation Miro |

| Prototype & Validate | | |
|----------------------|--|---------|
| Week | Task | Methods |
| 13 | Functional prototype | |
| 14 | Functional prototype | |
| 15 | Functional prototype | |
| 16 | Report writing | |
| 17 | User testing Affinity diagrams | Miro |
| 18 | Functional prototype User testing Affinity diagrams | Miro |

| Report writing & presentation | | |
|-------------------------------|-----------------------|---------|
| Week | Task | Methods |
| 19 | Report writing | |
| 20 | Report writing | |
| 21 | Make presentation | |
| 22 | Present Master Thesis | |

B

CANEA ONE

The Figures below showcases the different modules of CANEA ONE, Strategy, Process, Projects, Workflow, and Documents [91].

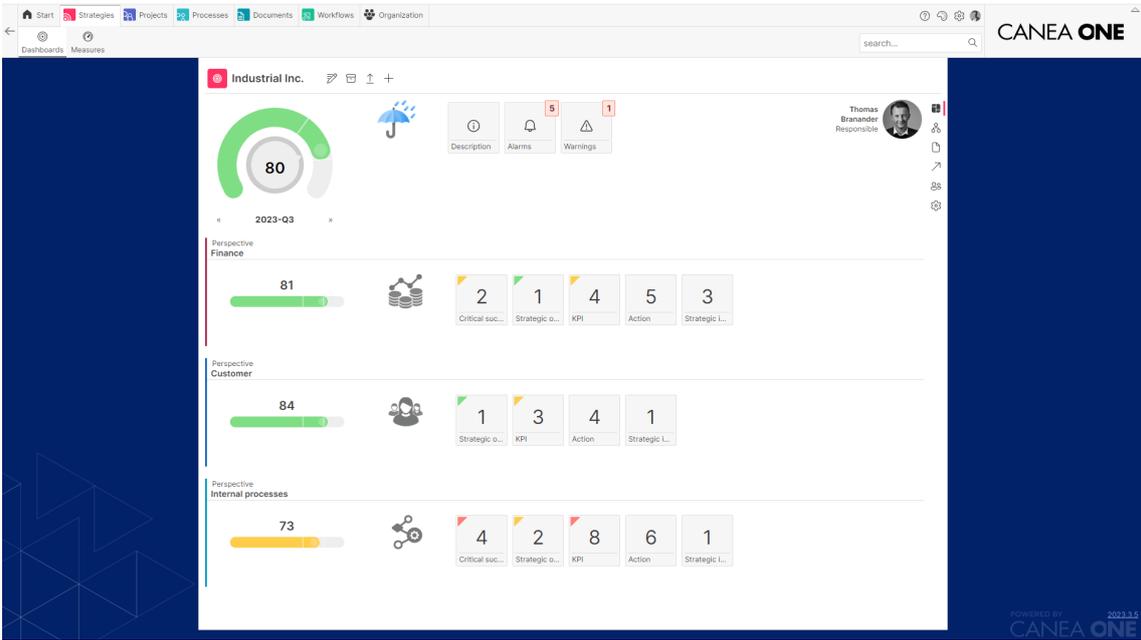


Figure B.1: Strategy

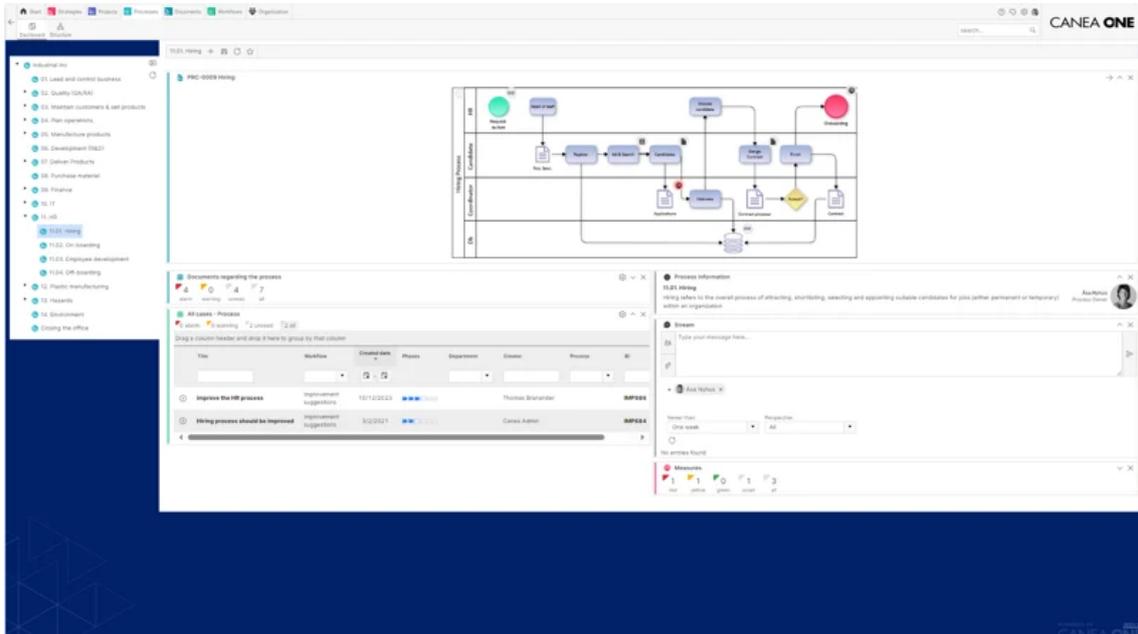


Figure B.2: Process

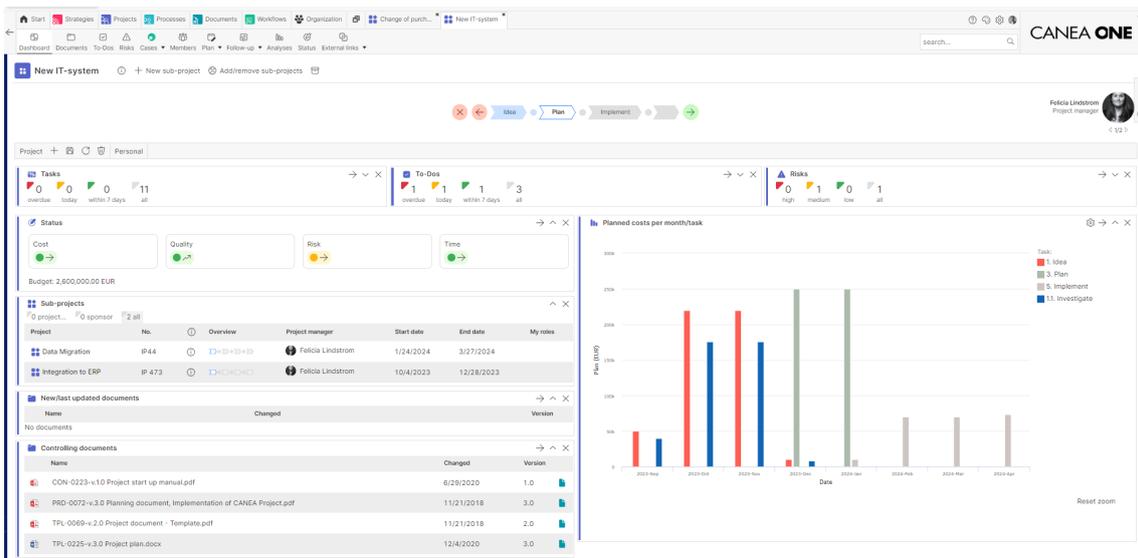


Figure B.3: Projects

B. CANEA ONE

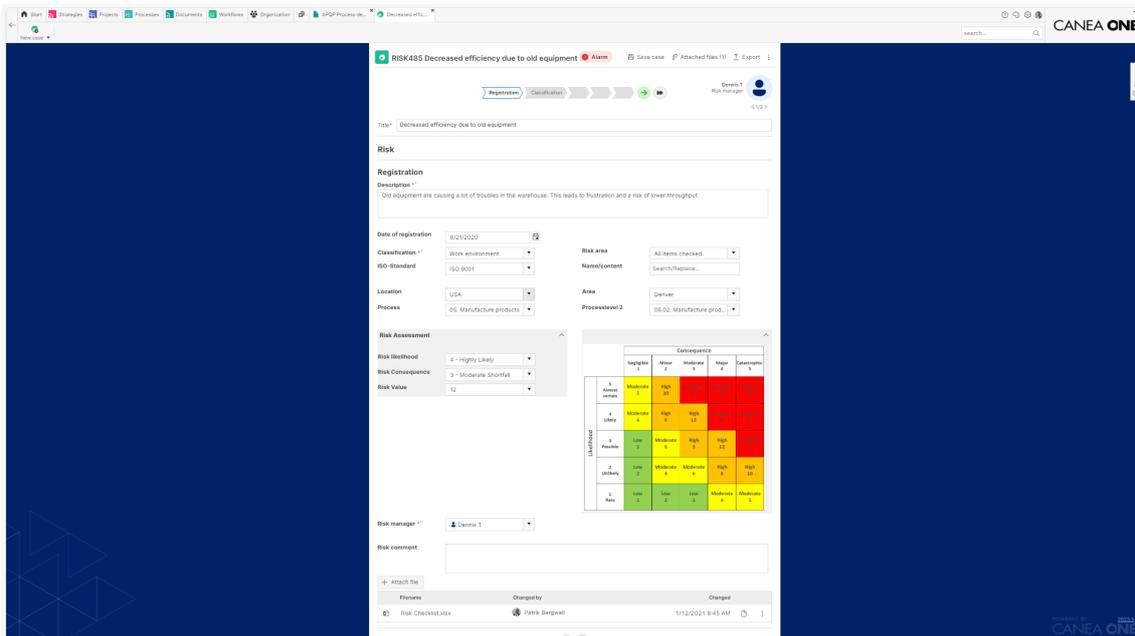


Figure B.4: Workflow

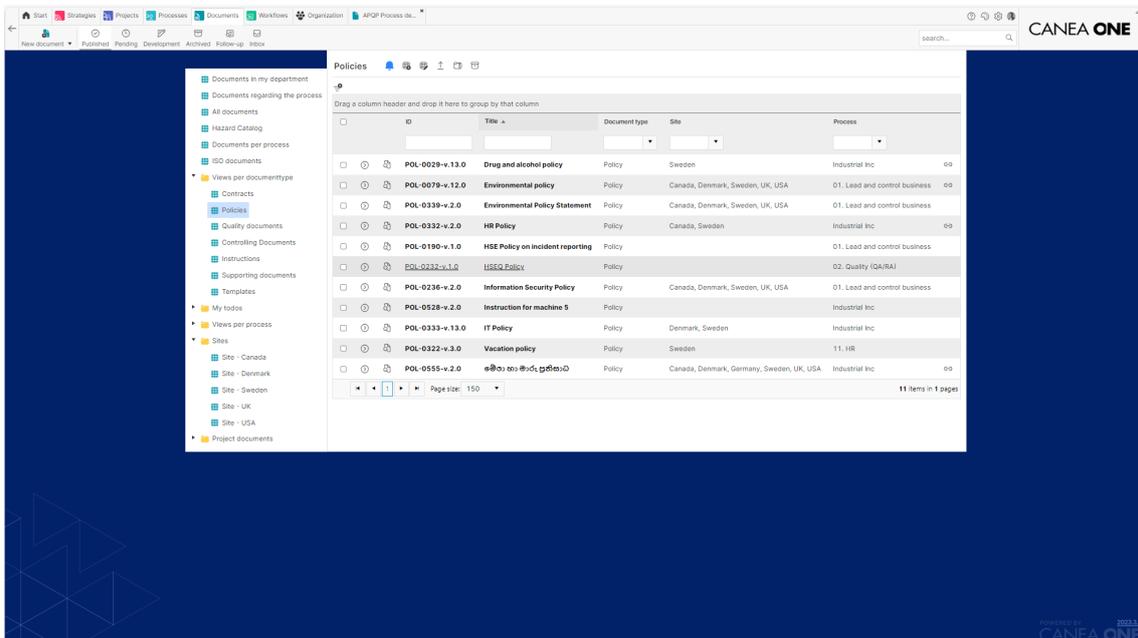


Figure B.5: Documents

C

Questionnaire

AI Assistant

We're excited to gather your insights on the implementation of an AI chatbot in CANEA ONE. Your feedback is invaluable as we explore ways to enhance the functionality and user experience within the application. Thank you for contributing.

Jenny & Linnea

What is your gender identity?

- Male
- Female
- Non-binary
- Prefer not to say

How old are you?

- 18-20
- 21-29
- 30-39
- 40-49
- 50-59
- 60 or older

Have you used the CANEA ONE program?

- Yes
- No

Use of CANEA ONE

What is your primary work role?

- Marketing
- Developer
- Management

How often do you use the CANEA ONE program?

- On a daily basis
- Weekly
- Monthly
- Less than once per month

C. Questionnaire

Which CANEA ONE modules have you used?

- CANEA Strategy
- CANEA Process
- CANEA Project
- CANEA Workflow
- CANEA Document

Choose one option that best describes your proficiency level in the modules selected above

- Beginner
- Intermediate
- Advanced

Have you ever experienced any difficulties using any of the modules in CANEA ONE?

- Yes
- No

Problem Areas

Have you ever experienced any difficulties performing any of the following tasks?

- Navigation (locating functions in the program)
- Information location
- Configuration of flowcharts
- Configuration of different types of projects
- Other...

How often do you experience the mentioned difficulties?

- On a daily basis
- Weekly
- Monthly
- Less than once per month

AI Chatbots

Do you have any experience using the following types of chatbots?

- Customer service chatbots
- Virtual assistants (Alexa, Siri, Google Assistant)
- AI chatbots (Chat GPT)
- No, I don't have any experience using any of the mentioned chatbots
- Other...

If you answered yes above, how would you rate your experience with such chatbots?

Poor 1 2 3 4 5 Excellent

How can an AI chatbot within CANEA ONE assist you?

- Automatically fill in information
- Help me locate different functions
- Help configure cases
- Help me find information
- Give me recommendations
- Answer general questions
- Other...

Interview

Thank you for taking the time to participate in this questionnaire. Your input is greatly appreciated.

Would you be interested in participating in an interview to further discuss the possibility of an AI chatbot in CANEA ONE?

- Yes
- No

D

Interview Questions

Introduction

Welcome!

About us

- Thank you for participating in this study.
- We are from the masters programme Interaction design and technologies at Chalmers university and are currently working on our master thesis here at CANEA.

Session

- This session will last max 60 minutes. First you will answer some questions (15-30 min), then you will complete a task in the program while explaining your process (10-20 min).
- You are free to withdraw from the study at any time.
- The session will be recorded for documentation purposes however after transcription will be disposed of according to GDPR.
 - Your answers will be anonymous.
 - Do we have your consent to do this?
 - Do you have any questions before we start?

Interview

About the participant

- In the questionnaire you answered . . . Could you please elaborate on how you use CANEA ONE in your work?

Current usage

- Is there any part of the program you avoid using/ have trouble using?
 - Can this be due to lack of knowledge?
- If you need help when using CANEA ONE where do you go to/who do you turn to?
 - Manual?
 - CANEA HR?

- Have you ever experienced any difficulties using any of the modules in CANEA ONE?
 - Please describe the difficulties you have encountered while using CANEA ONE.

Questions about AI

- Have you previously used an AI chatbot?
 - What was your experience using the AI chatbot?
 - What did you like/dislike?
- Would you trust an AI chatbot?
 - Why do you trust/not trust AI?
- Do you see any benefits in using an AI over the existing solutions?
- How can an AI chatbot within CANEA ONE assist you?
- Which types of responses would you prefer from the AI? Text, images, sound, suggestions, combination.

Observation

- We will now ask you to complete a task in one of the modules of CANEA ONE.
- The task will shortly be described
- Please think aloud as you proceed with the task.

Questions

- How difficult was the task 1 very easy 10 very hard
- How was the experience, annoying? Irritating? Satisfying?

Debriefing

- Thank you for participating, your input has been very valuable for our project.
- Do you have any questions?

E

User Testing

Introduction

Welcome!

- Thank you for participating in this study.
- The session today is to conduct user test on CANEA Chat, an AI chatbot integrated into CANEA ONE.

Session

- This session will last max 60 minutes. We will ask you to perform 3 tasks, more info about them will come later.
- You are free to withdraw from the study at any time.
- We want to remind you that there is no right or wrong answer so don't worry about making any mistakes. We also want to encourage you to be honest, don't be afraid to hurt our feelings when you give criticism.
- The session will be recorded for documentation purposes however after transcription will be disposed of according to GDPR.
 - Your answers will be anonymous.
 - Do we have your consent to do this?
 - Do you have any questions before we start?
- Before we start with the tasks, we would like to remind you that we want you to use the AI chatbot as much as possible.

Tasks

With the help of the AI chatbot, solve the following tasks:

1. Make an absence report between the dates 9-19th of April.
2. Turn on subscriptions to email and subscribe to documents view "instructions".
3. Create a configuration of a phase, a case type and a case with the following instructions:
 - There was an accident outside the CANEA office yesterday, a car collided with a bicycle and we want you to create a new case with the case type "Accident" which contains a phase called accident registration. The phase must contain a description of the accident and the date.

Followup Questions

- How would you describe your experience with the chatbot?

- What did you like most about the chatbot?
- What didn't you like about the chatbot?

- Did anything surprise you?
- Was there anything that frustrated you?

- How satisfied are you with the interface of the chatbot?
- Is the chatbot naturally integrated into the application?
- Is there any part of the interface that is not intuitive?
- Is there anything you would like to change about the interface?

- Is there anything you don't like about the chatbot?
- How do you feel about the visual design of the chatbot?

- What do you think of the responses from the chatbot?
- Feel like you're chatting with a bot? Why?
- Do you trust the answers from the chatbot?
- Is there anything that would have made you trust the chatbot more or less?
- On a scale of 1-5 where 1 is very unlikely and 5 is very likely, how likely is it that you would recommend the chatbot to a colleague.

Debriefing

- Thank you for participating, your input has been very valuable for our project.
- Do you have any questions?

F

Ideation Crazy 8's

The Figures below showcases the sketches developed during the ideation phase of the project using the method crazy 8's.

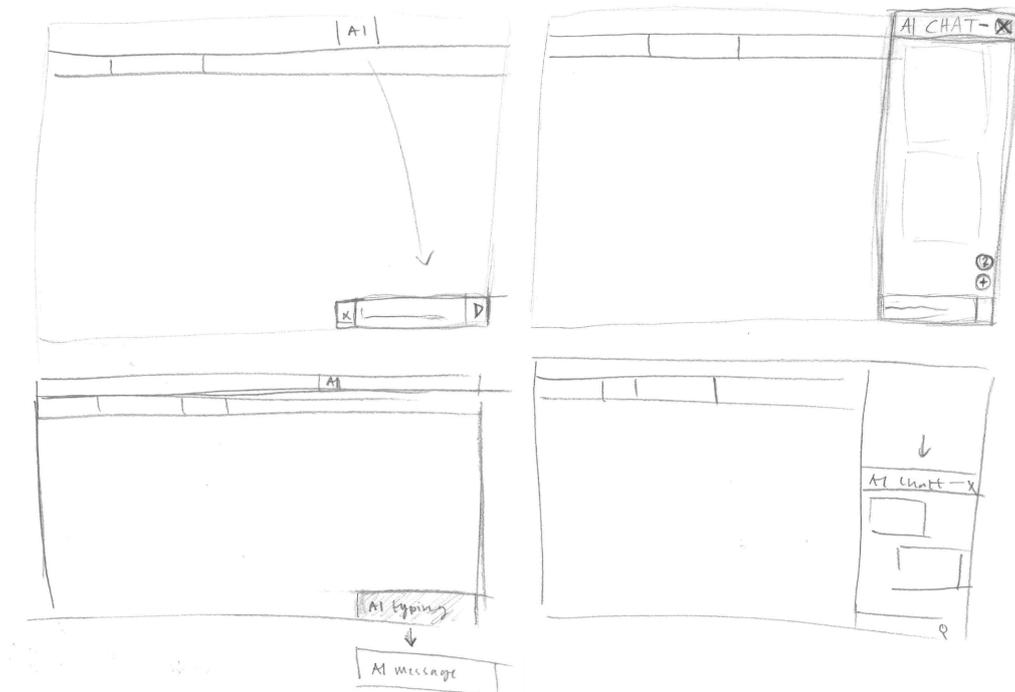


Figure F.1: Chatbot interface wire-frame sketches developed during crazy 8's part 1

Crazy 8's Chat-bot interface

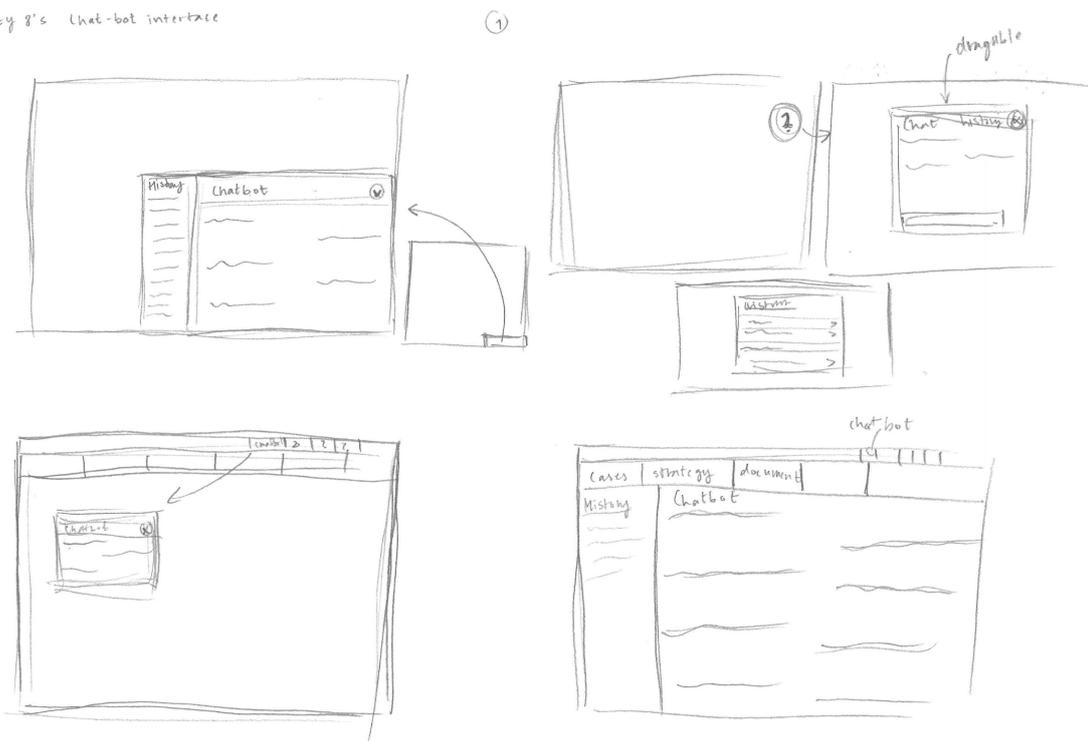


Figure F.2: Chatbot interface wire-frame sketches developed during crazy 8's part 2

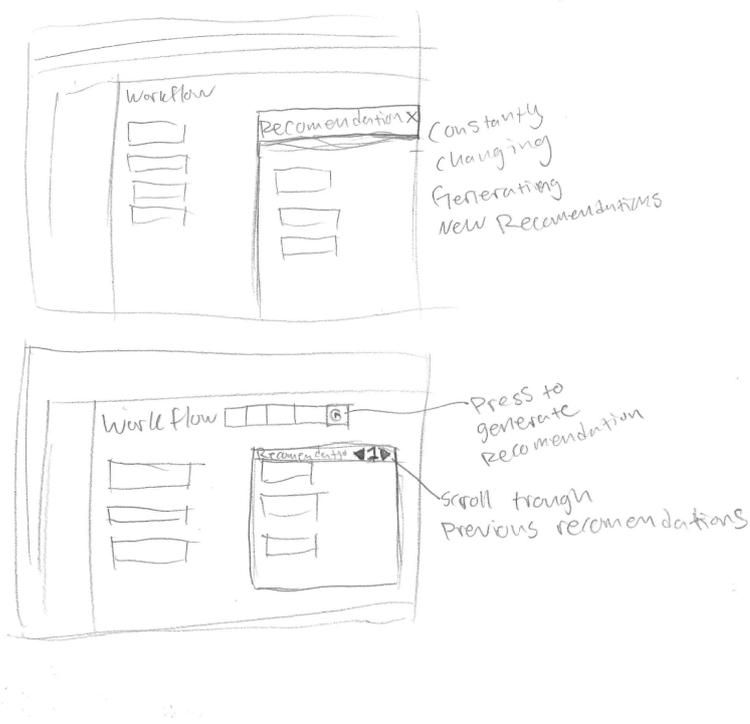


Figure F.3: Chatbot interface wire-frame sketches developed during crazy 8's part 3

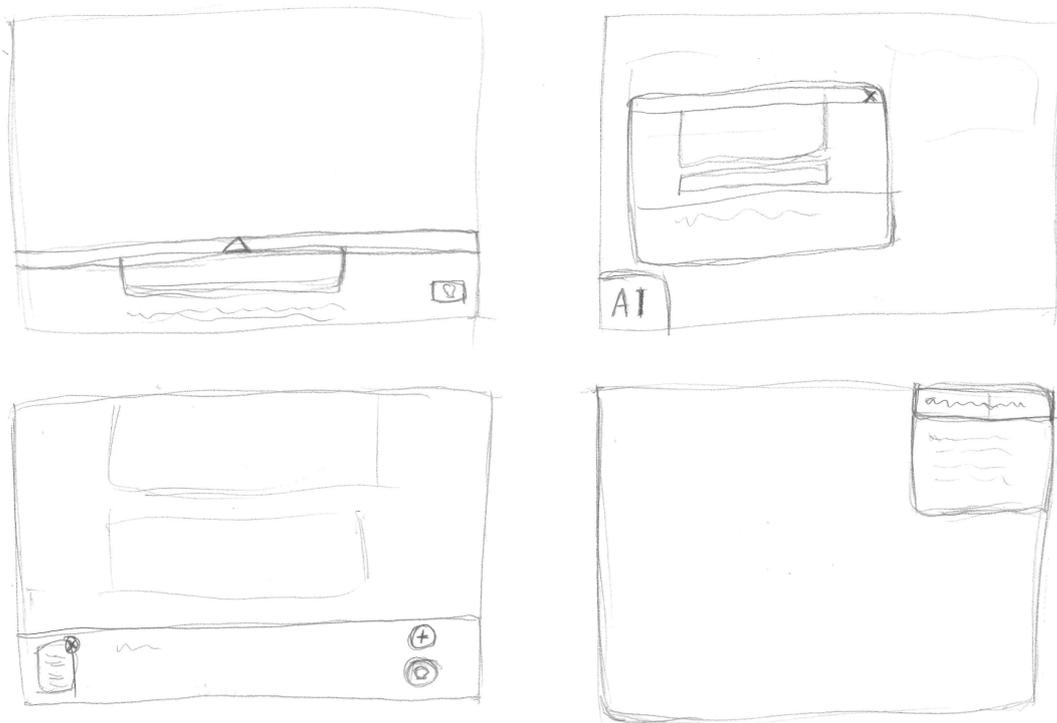


Figure F.4: Chatbot interface wire-frame sketches developed during crazy 8's part 4

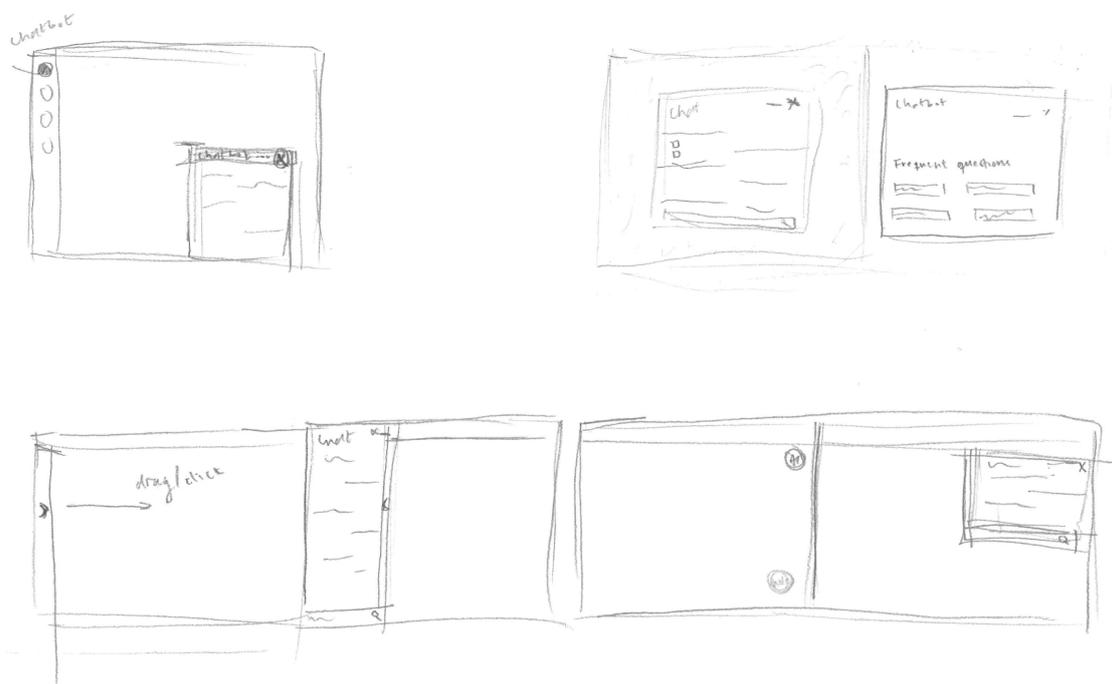


Figure F.5: Chatbot interface wire-frame sketches developed during crazy 8's part 5

G

Figma Wire-frames

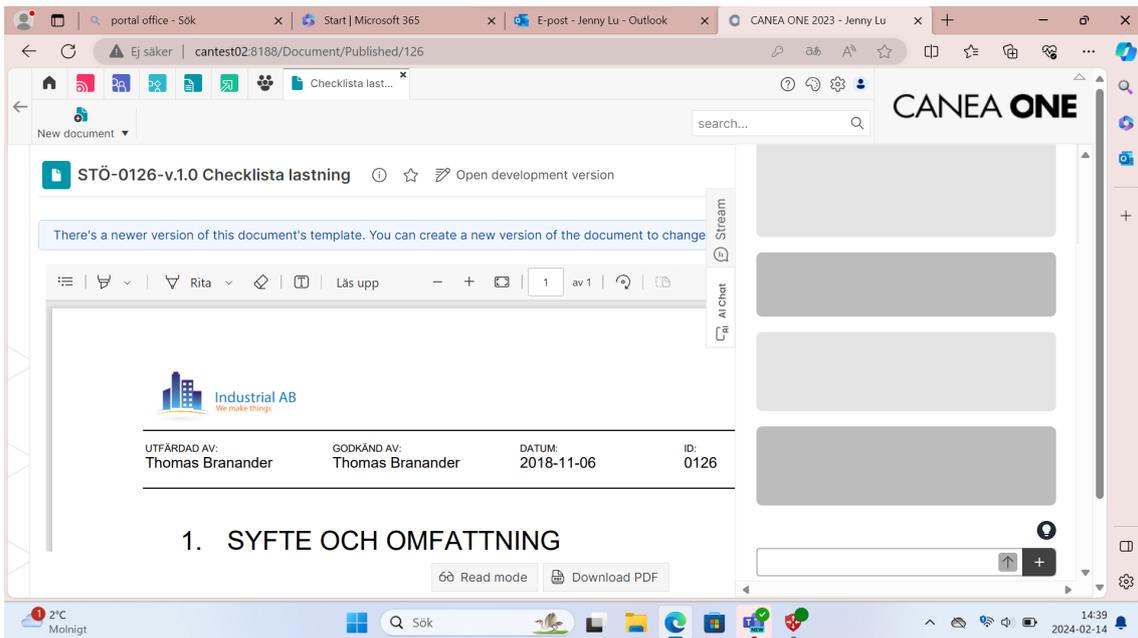


Figure G.1: Chatbot wire-frame created in Figma where the chatbot is integrated in the stream functionality (comments) of CANEA ONE

G. Figma Wire-frames

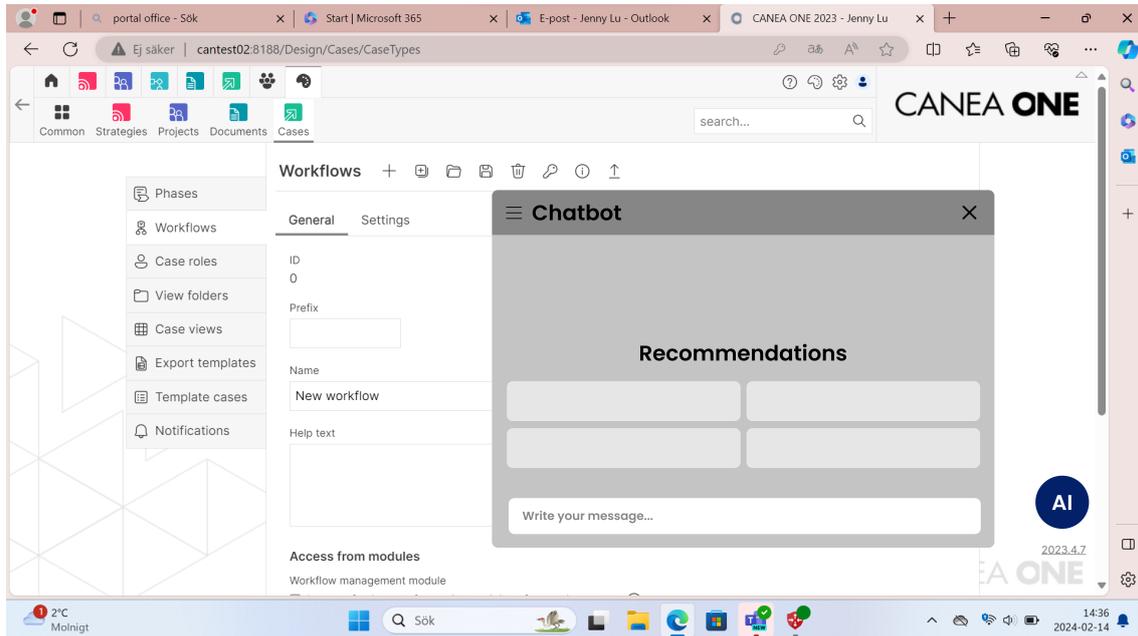


Figure G.2: Chatbot wire-frame created in Figma where the chatbot appears like a modal and is re-sizeable and movable

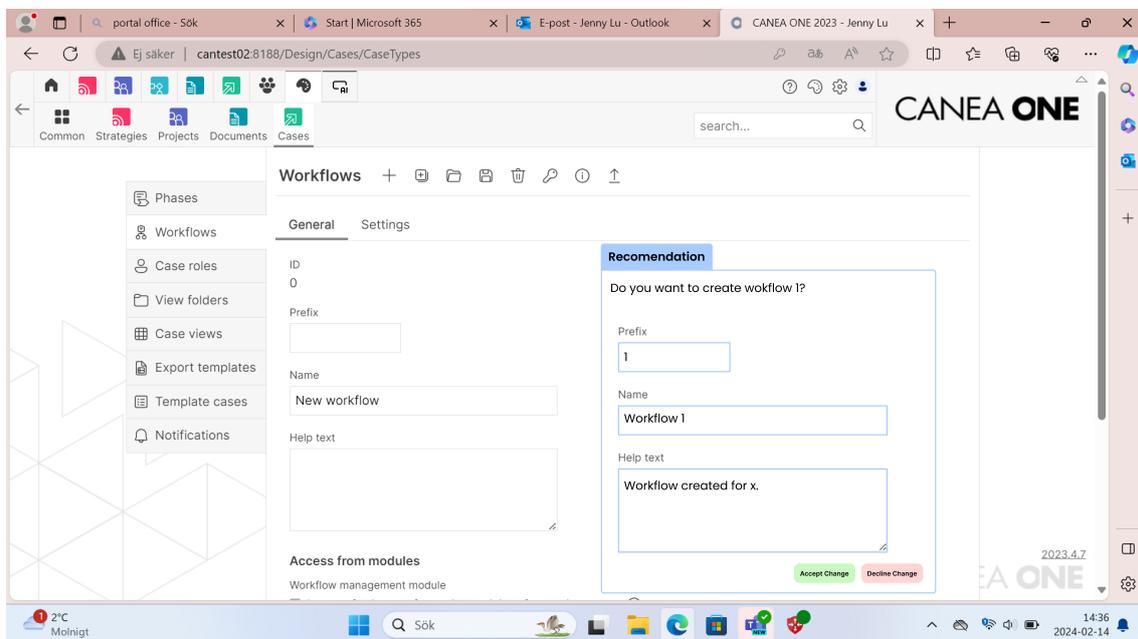


Figure G.3: UAS wire-frame created in Figma where the notification appears like a modal

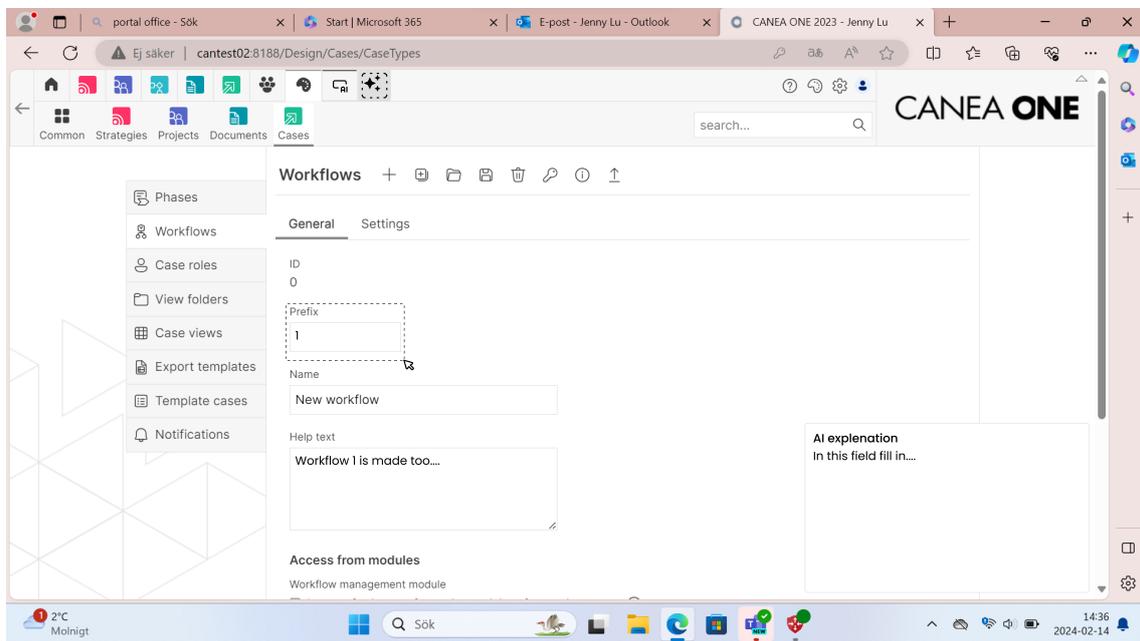


Figure G.4: UAS wire-frame created in Figma where the user drags a box over the functionality they want explained