

Anfang Liu

**A CHATBOT TO SUPPORT INTERNA-
TIONAL STUDENTS' ENTRY TO
WORKLIFE**

Prototype Development and a User Study

Faculty of Information Technology and Communication Sciences

M. Sc. Thesis

December 2023

ABSTRACT

Anfang Liu: A Chatbot to Support International Students' Entry to Worklife: Prototype Development and a User Study

M.Sc. Thesis

Tampere University

Master's Degree Programme in Human-Technology Interaction

December 2023

Chatbots have been employed in various fields, incorporating artificial intelligence (AI) technology to enhance realism. Despite extensive research on recruitment chatbots catering to recruiters, this study investigates the utilization patterns of job seekers interacting with the Generative Pre-trained Transformer (GPT)-based chatbot.

This study addresses the employment challenges of international degree students in Tampere, Finland, aiming to facilitate their adaptation to the worklife through a GPT-based chatbot prototype and user study. Preceding prototype development, conversation flows are designed, and a Tampere job search Knowledge Base (KB) is constructed via cross-disciplinary research and web content analysis. The Voiceflow platform is then employed for prototype creation. I conducted 3 evaluation workshops with 7 international degree students seeking employment in Tampere. The workshops utilize the chatbot prototype, the User Experience Questionnaire (UEQ) and group interviews to explore challenges associated with job applications, users' expectations for the chatbot, and interaction behaviours. Subsequently, data is systematically organised into formative themes employing Grounded Theory and affinity diagrams.

The qualitative results highlight job seekers' challenges, expectations, perceptions of the chatbot, and interaction behaviour. Typically, they contend with real-world challenges and external barriers in the job market. Accordingly, they anticipate the chatbot's role in enhancing job search skills and understanding the local employment landscape. They favourably perceive the chatbot and prefer free-form questions to predefined conversation flows, though they face challenges in formulating coherent questions. Additionally, these results bear implications for prospective research on GPT-based chatbots within the context of job seekers, leading to concentrating on improving conversation flows, and user experience, and evaluating chatbots based on local job market criteria and professional skills.

In conclusion, the study indicates that the GPT-based chatbot was effective, encouraging, significant, and valuable in supporting international students with their job-seeking process. Further pertinent research in this domain is deemed valuable.

Key words and terms: Chatbot, GPT-based chatbots, International Degree Students, Job-seeking Assistant, Conversational Design, Knowledge Base, Evaluation Workshop

The originality of this thesis has been checked using the Turnitin Originality Check service.

PREFACE

My thesis project was an interesting process with various challenges, and I am grateful that it happened. Starting with the initial motivation for chatbots, narrowing down the research gap, and confirming the thesis topic were the most time-consuming stages and they forced me to learn and understand how research happened at the beginning. I would like to express my appreciation to my two outstanding supervisors: Thomas Olson and Sami Koivunen. With their guidance, support, encouragement, and patience, the thesis process was interesting, challenging, and rhythmic, and I really enjoyed this process. I learnt from them that not to limit myself to what I am capable of, but to challenge myself to go further. “I would like to challenge you to consider ...” was what they generally said to me. Furthermore, Thomas provided the opportunity to collaborate with researchers in the Trust-M project which assisted in the development of the Knowledge Base and improvement of the creative evaluation workshop.

I would like to thank my friends for their help. Particularly, Wang Yurui encouraged me to accomplish tiny things throughout the process, shared relevant information with me, and offered his spare laptop when mine failed to function properly. Lauri Parkkamäki advised me to employ the Pomodoro technique in my thesis writing. Zhang Wenbo always told me to concentrate on my studies and everything would be alright. Additionally, I would like to express my heartfelt gratitude to all the participants for their contribution.

I am truly grateful to my family, including my parents, sisters, and brothers, who have always been behind me and trusted me. Finally, be thankful for my persistent and breakthrough self, it is incredible that I accomplished this significant, meaningful event.

Tampere, 7 December 2023

Anfang Liu

Contents

1	Introduction	1
2	Related work/Literature review	3
2.1	Job-seeking in Finland	3
2.2	Chatbots of Job Seekers' Perspective	3
2.3	AI-based Chatbots.....	5
2.4	Conversation Design.....	6
3	Research Methodology	7
3.1	Overview of the Research Process and Methods	7
3.1.1	Research Process	7
3.1.2	Research Methods.....	7
3.2	Methodology of the Evaluation Workshop	8
3.2.1	Data Collection Methods	9
3.2.2	Design of the Evaluation Workshop.....	10
3.2.3	Procedure of Evaluation Workshop.....	11
3.3	Participants and their Recruitment	13
3.4	Data Analysis Methods	16
3.5	Research Platform	16
4	Chatbot prototype	19
4.1	Recruitment Chatbot in Finland	19
4.2	Chatbot Prototyping Tools	19
4.3	Knowledge Base Design	22
4.4	Design of Conversation Flow	26
4.5	Technical Design.....	30
5	Evaluation results	34
5.1	International Students' Perceived Challenges Related to Applying for Jobs 34	
5.1.1	Practical Challenges Faced by International Students.....	34
5.1.2	External Challenges Faced by International Students	35
5.2	International Students' Expectations Towards the Chatbot	36
5.2.1	Support for Understanding Jobs, Writing Job Applications, and Preparing for Interviews.....	36
5.2.2	Insights into the Local Job Market	37
5.3	Overall impression of this chatbot	38
5.4	Interaction between participants and the chatbot	39
5.4.1	Initial impressions have an impact on how participants interact with the chatbot	39
5.4.2	Methods used by participants posed questions	40

5.1.3	The potential connection between the chatbot’s usage and the participants’ back-ground	41
6	Discussion	42
6.1	Implications of the Findings for GPT-based chatbots	42
6.1.1	Improve the conversation design according to guidelines.....	42
6.1.2	Design for improving chatbot user experience	43
6.1.3	Support for evaluating users’ professional skills.....	43
6.2	Limitations	45
6.3	Future Work	45
7	Conclusions	46
	References.....	47
	Appendices	52

1 Introduction

Chatbots have become more popular in various fields recently: for example, customer service chatbots are used on many websites to answer users' questions. In the dictionary, a chatbot refers to "A computer program designed to simulate conversation with human users (Adamopoulou & Moussiades, 2020).

Various script-based chatbots are already vast in number, a disadvantage of them is that they only follow the pre-defined conversation flow. The emergence of Artificial Intelligence (AI) has promoted the naturalness and intelligence of conversation. Currently, ChatGPT, Google Bard, and Bing Chat are good examples of generative AI. Aishwarya et al. identified three categories of AI chatbots, namely menu/button-based chatbots, keyword recognition-based chatbots, and contextual chatbots (Aishwarya Gupta & Dayananda Sagar College of Engineering, 2020). Menu-based chatbots have buttons and top-down menus, the user can follow the instructions and make decisions by selecting options and deeper towards appropriate response from the AI. Nevertheless, in terms of performance, this type of chatbot is not as dependable as others and cannot always produce the intended response. Keyword recognition-based chatbots can identify keywords from user input and generate a desired response. However, these chatbots will fail when there are different keywords between several related questions. For example, if the user asked, "I would like to find a UX designer job in Tampere, how can I write my resume?" the bot would probably use keywords like 'job' and 'resume' to decide which answer is the best to give. Contextual chatbots make use of Machine Learning and AI technologies, such as voice and speech recognition to understand users' situations and present an intelligent answer. To build an AI-based chatbot, Large Language Models (LLMs) are applied, LLMs are trained on massive amounts of data, including articles, books, and other resources to produce natural responses. A Generative Pre-Trained Transformer (GPT) is one of the LLMs, which is an open source and applied in various generative AI.

This research aims to build a GPT-based chatbot and evaluate it with international degree students to find out their challenges, expectations, and interactive behaviour. International students are a vast job-seeking group defined as the target group in this study. This target group was having difficulty getting their first job after graduation because they lacked work experience, had limited awareness of the Finnish market, lacked networking, and lacked local language abilities. Tampere was chosen as the site of the research because it is the second-largest city in Finland and is popular among both Finns and international students. Furthermore, Tampere is attempting to internationalise through the integration programme (*City of Tampere Integration Programme 2021–2025*, 2021), which is a two-way street, with international students being guided to find jobs in Tampere, and Tampere benefiting from international talents. Building on the above, this

research concentrated on prototyping and evaluating a meta-level chatbot that assists international students in orienting towards career life, helping them understand how the job market in Tampere operates, and where to get more information or support. In other words, this study encompassed two critical stages: chatbot prototype development and chatbot evaluation. To provide natural responses and go beyond traditional script-based chatbots, GPT was used to train collected data from web content analysis related to job-seeking in Tampere to deliver natural responses and go beyond standard script-based chatbots. In addition, an initial nonlinear conversation flow was built to engage users to continue the conversation and ask recruitment-related questions. The author then conducted three qualitative self-designed evaluation workshops with 7 international students who have recently sought a job in Tampere. Participants in the workshops got the opportunity to test the GPT-based chatbot prior to group interviews.

The findings highlight several significant themes that demonstrate the critical need for international students to develop recruitment skills to enter the recruitment process. In this sample, the target group preferred to type free-form questions to ask how to strengthen their recruitment abilities and enter the local employment market.

2 Related work/Literature review

This chapter first describes the job-seeking situation in Finland which indicates the need for this research, and then reviews chatbots from the standpoint of job seekers. The following subsection covers AI-based chatbots and why AI technology was used in this research. The final subsection provides a brief introduction to the conversation design process and design guidelines for conversation flows.

2.1 Job-seeking in Finland

The job-seeking process consists of a series of activities to find a (new) job. These activities can include identifying one's career goals, preparing application materials (e.g., resume and portfolio), accessing job opportunities through companies and job searching platforms, applying to open positions, and preparing for interviews with interested companies (Wanberg et al., 2020).

According to the research from Viitasalo, the key issues with the employment of international students in Finland are the lack of language, social networking, and awareness of Finnish companies (Viitasalo, 2018). Language is a tool for communication, and in this case, Finnish is the specific language. As an international student, it is advantageous to be motivated and to have learned the language. Social networking is one of the sources of job information and employment opportunities for job seekers in the Finnish job market even though the well-established formal employment platform (Ahmad, 2005). Understanding the knowledge about Finnish companies is important when seeking a job in Finland, it helps jobseekers to prepare for applications, interviews, navigate the job market and build networking with potential employers (Finland, n.d.).

The positive perspective for jobseekers from Tampere city is that an integration program starts in 2021 and runs until 2025. This program offers various integration strategies for different fields, with the integration of employment and growth services and international talent services being the most important for international students (*City of Tampere Integration Programme 2021–2025*, 2021). To be specific, job seekers can get help from Multilingual info Mainio, International Skills Centre for Immigrants OSKE, and International House Tampere.

2.2 Chatbots of Job Seekers' Perspective

The lack of connection between students' academics and their career goals, as a result, a web application contains a chatbot was built to assist students in career planning, development, and real-time guidance from the counselling chatbot ('WEB BASED CAREER

COUNSELLING SYSTEM WITH CHATBOT', 2021). The Intelligent Career Counselling Chatbot (ICCC) provides outputs of suggestions to users based on their queries, acting as an assistant to help them find out career goals through career suggestions. While the ICCC is mainly for high school students (Goyal et al., 2023).

Recently, career counselling chatbots for university students were investigated. Summy (Lee et al., 2019) was created to assist college students with their career development, and the researchers discovered that chatbots can be used to provide information and recommendations, provide career development intervention, augment the work of career counsellors, and support career counselling. Typically, the research built an AI chatbot to assist colleges with university admissions counselling and career guidance (Truong et al., 2020). Furthermore, Zaidi et al. also developed an AI-based chatbot to help students resolve career-related requests with appropriate AI responses, and users are not forced to follow any certain pattern (Zaidi et al., 2021). Suresh et al. observed that newly graduating university students struggle to find employment and mentioned one of the possible reasons was the lack of meaningful career guidance (Suresh et al., 2021). Based on their findings, they constructed a career counselling chatbot built with AI to offer career counselling guidance and support students to discover possible career choices. Finally, CASExploer (Lee et al., 2022) was created to help college students majoring in science, technology, engineering and mathematics discover their career goals through self-exploration, academic exploration, and career exploration. The intelligent career counselling bot was designed which help users select career paths by giving relevant responses to their questions (Parab et al., n.d.).

An orientation-based chatbot system was developed to support university students and graduate students who wish to enter the workforce in choosing a job that matches their preferences. (Hourrane et al., 2020)

Furthermore, a CV (Curriculum Vitae) chatbot (Zhu & Han, 2021) based on the "STAR" technique was developed to assist undergraduates in producing "projects and working experiences" in their CV, so that they have more possibilities to present the beginning of the recruiting process.

According to the research about the future employment applications for underserved job seekers (Dillahunt et al., 2018), three concepts report job seekers' practical requirements: providing resume feedback, assisting job seekers in defining work experience, and a specific strategy to achieve their career goals.

2.3 AI-based Chatbots

Diederich's findings suggest that the use of preset answer options in conversational agents might be detrimental to service satisfaction, as they can impair the natural sense of human-conversational agent interaction (Diederich et al., 2019). To improve the natural performance of a chatbot, generative AI gives an example because the underlying principle is learning from existing data to generate new and realistic responses (*Generative AI*, n.d.). As a result, this section will proceed with the application of AI in the chatbot field with a brief introduction to Large Language Models (LLMs) and Generative Pre-trained Transformer (GPT).

ChatGPT is an AI chatbot created by GPT that has been employed in a variety of areas including creative writing, essay writing, prompt writing, and answering questions (Taecharunroj, 2023). GPT refers to a Generative Pre-trained Transformer, a sort of Large Language Models (LLMs) that generates human-like, conversational text using deep learning. LLMs represent AI tools based on multilayer recurrent neural networks that are trained on vast amounts of data to generate human-like text (Alberts et al., 2023). LLMs have grown in popularity because of their extensive application for a range of Natural language processing (NLP) tasks, including translation, classification and categorization, as well as conversational AI and chatbots (*What Are Large Language Models? | Definition from TechTarget*, 2023). These capabilities provide functional support for chatbots, especially conversational AI and translation which support multi-language questions. LLMs have numerous benefits, including extensibility and adaptability, flexibility, performance, accuracy, and ease of training (*What Are Large Language Models? | Definition from TechTarget*, 2023).

GPT can handle repetitive tasks, enabling humans to focus on more complicated tasks that require a higher degree of critical thinking. Furthermore, GPT can accept text input in multi-languages. Therefore, GPT could be utilized as a customer service centre to answer questions from customers or to support chatbots. GPT-3, a third-generation, autoregressive language model that uses deep learning to produce human-like texts and uses the previous distinction to analyse them (Floridi & Chiriatti, 2020). GPT-3.5 was widely used due to its versatility, intelligence, and ability to engage in human-like conversations, ChatGPT is a good example of this. GPT-4 which is multimodal, and it supports image inputs and text inputs, was released in 2023 and is more advanced than GPT-3.5.

Furthermore, Yang and Evans also identified a research gap in AI chatbots in higher education and proposed three separate AI chatbots to assist learning and teaching, educational application training, and helpdesk support (Yang & Evans, 2019). AI provided two benefits for the chatbot relevant to my research. To begin, the chatbots can understand input messages with grammatical and spelling mistakes, to be tolerant of users' language.

Second, chatbots can direct users to helpdesk-related websites to save time, which corresponds to the function of providing connections to job search platforms in my research.

2.4 Conversation Design

Conversation design is a language-based design in human conversation, it is an architect to map what users can accomplish in space while considering the technological constraints (Muhammad et al., 2020). The process of conversation design is iterative, including designing the conversation flow, testing with users, reviewing their data, and enhancing the design (Choi et al., 2021). Conversational design guidelines are critical for the chatbot design because they provide a set of principles and recommendations that help to ensure consistency, clarity, and usability across all areas of the design.

According to the Chatbot Conversation design guidelines, the conversation should have three attributes: transparency, naturalness, and emotionality (Silva & Canedo, 2022). Transparency ensures that users can understand the capabilities and limitations of the chatbot, and the chatbot can provide suggestions and even seek clarification from users to give better responses. Naturalness is related to the chatbot's ability to self-introduce, address users by name, make small talk, and echo responses with casual language. Emotionality implies that the chatbot uses graphical media and provides emphatic messages. McTear also proposed three types of dialogues in conversation modelling: one-shot queries, slot-filling and open-ended dialogues (McTear, n.d.). One-shot queries are simple input-output that do not require follow-up questions from the user. Slot-filling dialogues can be controlled by the system and users, this type of interaction can collect information from users. The open-ended dialogue has more natural conversation for example the Google Assistant.

3 Research Methodology

The Research Methodology chapter explains the research phases and describes different research methods utilized during the research procedure. In addition, the design platform used to build the prototype was named Voiceflow¹.

3.1 Overview of the Research Process and Methods

3.1.1 Research Process

This thesis research is divided into seven main phases. The primary duties and outcomes for each phase are listed below.

The first phase, the literature review aimed to discover related work conducted by other academics, identify the research gap, and format research questions. Furthermore, different research methods and design guidelines were investigated before deciding on the research methodology.

The second phase, cross-disciplinary research concentrates on attending diverse events to reach a broad audience and identify a potential research direction. Career workshops, research meetings, and visiting job-seeking organizations are examples of events.

The third phase built a Tampere job-seeking related Knowledge Base to serve as the database for the GPT-based chatbot by the web-content analysis method.

The fourth phase, the construction of a prototype, was responsible for creating the GPT-based chatbot prototype, which included conversation flows and Knowledge Base.

The fifth phase, which designed and conducted the evaluation workshop, was a validation of initial assumptions, insights, and learning gathered from the literature study and cross-disciplinary collaboration. Furthermore, evaluate the design to determine users' expectations and interactions.

The sixth phase, design evaluation, intended to determine how international students would prefer to utilise such a system, and what to expect from it. This phase is also beneficial to discover the weaknesses of the design and potential improvements for a better user experience. Additionally, identify potential research directions for future GPT-based chatbot research.

The discussion and conclusion phase follows, and it contains qualitative data analysis and findings from the evaluation phase, limitations of this research, and future work.

3.1.2 Research Methods

To understand the research gap and conduct the research, the following research approaches were employed in this study.

¹ <https://www.voiceflow.com/>

Literature review as a research approach can support researchers to address research questions, recognise current knowledge in a certain field, and discover research gaps (Snyder, 2019). A literature review was conducted in this study to better understand the job-seeking situation for international students in Tampere, identify the study of chatbots in the field of recruitment from the perspective of job seekers, comprehend the application of AI in chatbots, and define conversation design guidelines. The most crucial part was finding the research gap in international students' readiness to enter work life.

Cross-disciplinary research is an approach that aims to integrate knowledge and collaborate with other researchers from different disciplines to create new insights and solutions (Heitzmann et al., 2021). This strategy was used in the Knowledge Base development phase where the author collaborated with other researchers and organizations to collect reliable information.

Web content analysis is a systematic approach to collecting and analysing user-generated data from the World Wide Web (Wed) (Kim & Kuljis, 2010). During the Knowledge Base design, this method was employed to identify the job-seeking related themes and contributed to the design of conversation flows.

Construction of a prototype: during the design solution phase, designers create varying degrees of fidelity to fulfil users' requirements and capture design concepts. Prototyping is an experimental process where design teams transform ideas into tangible forms. To create the solution in this research, low-fidelity and high-fidelity prototyping were used. Paper prototypes are low-fidelity prototypes, which can contribute to the basic layout and initial insights, and they are simple to modify when new iterations are introduced. The disadvantage of paper prototypes is that they are too basic to reflect the user experience of the finished product, therefore high-fidelity will bridge this gap. Digital prototypes usually were used as high-fidelity prototypes. The most essential part of this research is the conversation flow, which causes a problem of what kind of software can support the conversation type of prototype, which is discussed in the research platform in Chapter 3.5 and the prototype development process in Chapter 4.

3.2 Methodology of the Evaluation Workshop

In collaborative workshops, the researcher acts as the organiser to prioritise participant requirements, and the participants along with their ideas and materials produced in this workshop become part of the research design (Ørngreen & Levinsen, 2017). I designed and conducted the collaborative workshop to explore the challenges for target users in job-seeking, expectations for the GPT-based chatbot, and how they will use such a system. Continue to investigate potential enhancements and research based on the results.

3.2.1 Data Collection Methods

Diverse data collection methods were used during this study, some of them were qualitative methods such as audio-recording, screen-recording, focus group interviews, note-taking, and Log data, while the questionnaires filled by participants collected both qualitative and quantitative data.

Questionnaires: Participants were asked to complete two questionnaires during the evaluation session, the first questionnaire was connected to background information to demonstrate their age group, study field, experience in generative AI, and job-seeking status in Tampere. The second questionnaire named User Experience Questionnaire (UEQ) (Schrepp et al., 2014) was completed by participants after using the GPT-based chatbot prototype. The scales of the questionnaire cover a comprehensive impression of user experience and could measure both classical usability aspects (efficiency, perspicuity, dependability) and user experience aspects (originality, stimulation). The purpose of UEQ is to explore the user experience and identify potential improvement aspects.

Audio-recording of the evaluation workshop: The entire evaluation workshop was audio-recorded in order to allow for verbatim transcripts of the discussion and data analysis to discover significant findings.

Screen-recording of the chatbot prototype: During the evaluation session, participants were instructed to utilise the chatbot prototype independently. This section was screen-recorded to allow for video observation to analyse participants' behaviours, what topics they asked the chatbot and how they formulated questions.

Focus group interview: A 15-minute focus group semi-structured interview was conducted and audio-recorded in which they discussed their experience related to the chatbot prototype, with themes including the benefits of using the chatbot, the disadvantages of using the chatbot, and improvement suggestions for this chatbot prototype. As well as whether would they recommend this chatbot to their friends.

Observation and note-taking: Except for the group focus interview, direct observation (Baker, 2006) was executed during the evaluation workshop session. To prevent making participants feel uncomfortable, the notes were written while they were undertaking individual work, group discussions, and chatbot prototype usage. Furthermore, the limitations and problems that occurred during this session have been noted.

Log data: During the use of the chatbot prototype, the log data (Dumais et al., 2014) will be stored in the research platform. Log data gave support to the analysis of what questions were asked by participants, how participants interacted with the chatbot prototype, and how to improve the chatbot prototype.

3.2.2 Design of the Evaluation Workshop

The purpose of the evaluation workshop is to offer an organised and participative way to assess the success and effect of the chatbot. The evaluation workshop can help to validate the assumption that international students face various challenges in the job-seeking process, test how they would like to use this chatbot system, understand what they expect from it, collect feedback from participants, and explore future improvements or research on this GPT-based chatbot.

Workshop Goals: The workshop goals included exploring how international students at Tampere would utilise this GPT-based chatbot, what they expected from the chatbot, and what challenges they encountered in job applications.

The cost of the workshop: It is important to consider the cost of the workshop including the number of participants, the duration, the venue, and the rewards. Generally, 2 to 4 participants as a group are typical, however, with only one researcher as facilitator in this case, 2 or 3 participants per group was the option. Three workshops were conducted during this research, each workshop had 2, 2, and 3 participants respectively. The workshop duration was set for one hour because it is a reasonable amount of time to cover the primary objectives while also keeping the participants engaged and motivated. The workshop was conducted in person at Linna 3016 at Tampere University because face-to-face communication is beneficial to group discussion and makes it easier for the facilitator to handle technical circumstances. Furthermore, rewards can benefit both participants and researchers. Rewards can make participants feel valued, respected, and appreciated for their contributions. Because participants are essential for the research, researchers can utilise rewards to attract more participants to the study.

Procedure design: Building on the workshop goals, three themes' activities were designed in the workshop. To begin, it validates the assumption that international students face challenges during job hunting in Tampere. At the same time, participants also shared their expectations for the GPT-based chatbot. The first stage consists of warm-up activities, individual work, and group discussion. Warm-ups could assist participants in becoming acquainted with one another and establish a welcoming environment. Following that, participants were instructed to put down the challenges, difficulties, and problems that they wished to question the GPT-based chatbot. Moving to the group discussion, participants had the opportunity to exchange their ideas and job-seeking situations. If participants did not have free-form questions to ask, all actions in this stage contributed to the next phase. Second, participants were requested to use the chatbot prototype on their laptops with screen recording so that the author could analyse data afterward. The goal of this stage was to explore how they would like to interact with this chatbot and what they expect from it. Finally, a semi-structured interview was conducted to collect feedback about the chatbot and their perspectives on the study.

Background information questionnaire design: A brief questionnaire was designed to collect background information from participants. Age group, original country, professional background, knowledge of generative AI, and work opportunities in Finland were all collected to investigate probable relationships between background and chatbot use. The list of questions can be seen in Appendix 1.

Interview questions design: The design of the interview questions was based on the goals of the workshop. The first two questions were aimed to find out the impression and experience of the GPT-based chatbot after actual use. The third question asked participants to reflect on the difference between their thoughts and the chatbot's capabilities. The fourth question asked participants to indicate the extent to which they will improve this chatbot. The final two questions were designed to indicate whether they will continue to utilise and recommend this chatbot to their social circle. The list of questions can be seen in Appendix 2.

Technical guidance for using the chatbot stage: Participants were asked to bring their laptops to the evaluation session, and during the using chatbot stage, they were instructed to access the chatbot prototype webpage through an incognito window on their laptops. To ensure smooth data collection by screen recording, the technical details should be documented down and checked. As a result, the author prepared technical guidance for using the chatbot, which helps participants understand what they can do while also allowing the facilitator to verify everyone's device. Firstly, the author used Teams to create distinct meetings for each participant so that participants could attend their meetings. Secondly, abbreviate the link to the chatbot prototype so that participants can quickly type it into the browser. Thirdly, for the screen recording, participants were requested to share the chatbot screen through the Teams meeting. Fourthly, the facilitator checked the first participant's device setting and started to record the meeting, verifying that the participant's side was recording and that there was a "recording started" in the meeting chat. Lastly, the facilitator left this meeting and repeated the process for the remaining participants. Appendix 3 contains technical instructions for utilising the chatbot.

3.2.3 Procedure of Evaluation Workshop

The evaluation workshops were conducted in person at TAU Linna 3016, and a pilot test as well as two subsequent workshops were conducted, each workshop lasted for an hour. Participant guidance with user ID and distinct links of the prototype, seen in Appendix 3, was created to guide the prototype utilising stage, the agenda involved the following subjects:

Instructing the participants: The main goal of the thesis project, and the purpose of the workshop were presented first. Participants signed the consent form to participate and filled out the background information. They were reminded that they were free to quit

this workshop at any time and without explanation. In addition, the facilitator explained the entire session was audio-recorded, how the collected data would be stored, and anonymity when reporting the results.

Background introduction and warm-up activities: The motivation and background of this thesis project were presented initially. Participants were asked to answer three job-related questions which depicted the job-seeking situation to a better understanding of the topic and each other.

Individual work: In this task, participants were instructed to use the sticky notes and pen on the table to write down what they wanted to ask the chatbot about job-seeking, what challenges they had when seeking a job in Tampere, and anything else connected to their job hunting.

Group discussion: The group discussion was based on the previous task, each participant explained their individual work, categorised notes, and highlighted repetitive issues. New insights were written down as well if they came up with something.

Chatbot using individually: Each participant followed the chatbot use in-person guidance (Figure 1), the researcher checked the technical settings and started recording the participant's screen, and participants had 10 minutes to use the chatbot prototype.

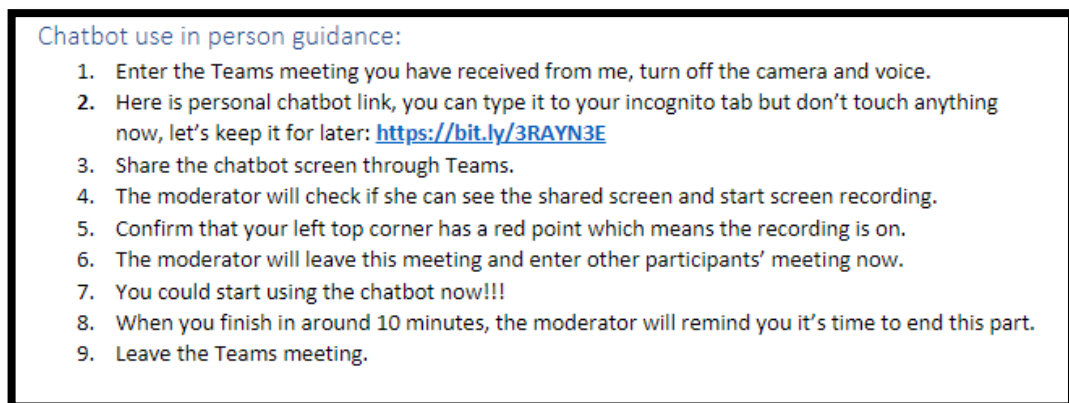


Figure 1 Chatbot use in-person guidance

User experience questionnaire (UEQ): After the chatbot prototype was used, participants were requested to fill out the short UEQ.

Group interview: A semi-structured group interview was conducted, and participants answered interview questions one by one.

Closing remarks: At the end of the session, participants were thanked for their time and effort, and the reward was emailed to them.

3.3 Participants and their Recruitment

After identifying the goals of the evaluation workshop, the target participants should possess the following criteria be an international master's degree student in their second or third year of studies, want to find professionally appropriate employment in Tampere, and actively seek a job. A LinkedIn advertisement was created to attract target participants. The advertisement included a sign-up link, where potential participants could learn more about the study and choose the best time slot for their schedule.

All the participants were Tampere University students with prior experience with generative AI (i.e., ChatGPT). In addition, three of 7 participants were research assistants (RA) in their respective fields (P1, P2, and P3), however, none of them had company work experience or work opportunities in Tampere so far. The participants' professional roles and other background information are presented in Table 1.

Table 1 Participant information

ID	Role(s)	Studying programme	Frequency of using GPT	Original country	Age group	Job chances in company?
P1	Student, RA	Factory Automation and Robotics	Between daily and weekly	Vietnam	25~30	No
P2	Student, RA	Factory Automation and Robotics	Daily	Sri Lanka	25~30	No
P3	Student, RA	Human Technology Interaction	Use it sometimes when need	India	25~30	No
P4	Student	Information Security	Daily	Bangladesh	25~30	No
P5	Student	Statistical Data Analytics	Use it sometimes	India	18~24	No
P6	Student	Human Technology Interaction	2 or 3 times per week	China	25~30	No
P7	Student	Human Technology Interaction	Occasionally	India	18~24	No

P1 was studying Factory Automation and robotics Engineering, as well as a research assistant in this field. The participant used generative AI on a daily or weekly basis. The participant will consider whether the job description matches the professional background when it comes to job-seeking. At this early job-seeking stage, the participant was gaining enough experience and skills to find a solid position in Tampere. To achieve those goals, the participant will attend job fairs, network with people and conduct background research on companies regarding job opportunities and the type of application they are looking for.

P2 was studying the same programme as P1 and was also a research assistant in this field. Because GPT was part of the RA's work, the participant was familiar with it. The first thing was to be fluent in the language when discussing job-seeking. The participant started applying to every potential company that leads to the same professional background, even though P1 did not receive any interviews, the participant maintained a positive attitude. As for contributions to job-seeking progress, the participant was working on building the network, for example, obtaining references from existing employers and colleagues. Another contribution was building a professional profile through getting experience from the research assistant work.

P3 was studying Human-Technology Interaction (HTI) and working as a research assistant in this field, but P3 would like to be a software developer. LinkedIn came to P3's mind when mentioning job-seeking because it is a place where people can access job vacancies and apply for them. Currently, the participant is not actively seeking a job since the job market has limited open positions. While the participant made some contributions to the job-seeking process, including asking for feedback on improving the curriculum vitae (CV) and building social networking with people who were senior-level in specific positions. P3 would pose questions like "What kinds of skills are popular right now?", based on the response from networking, the participant was polishing CV and building professional skills.

P4 was studying Information Security and was familiar with generative AI because of daily use. The participant felt useless about job-seeking for the string of rejections. In the last 16 months, P4 applied for roughly 200 jobs but received no constructive feedback or interview opportunities. The worst-case scenario was that 100% rejection would happen if showing face on the CV, as P4 explained: "It is not racism but more like circumstance biases because they do not prefer brown people and Asian people, Finland has this sort of issue." To receive accurate feedback, the participant revised the CV multiple times based on the opinions of others. In addition, P4 discussed with a few recruiters and discovered the evaluation standards for Finnish people and international people are different, they have higher requirements for international talents, and they prefer Finnish people. Building on this negative experience, the participant stopped applying for jobs and was considering applying for positions outside of Finland because of the better responses.

P5 was studying Statistical Data Analysis and would occasionally employ generative AI. The participant was considering what exactly the recruiters are looking for from the CVs of international students because some people have impressive CVs but without interview chances, while others without work experience got interviews. P5 applied for roughly 50 jobs with rejected, and only got one technical interview but did not advance to the next phase. Therefore, the participant was confused about the true requirements for a job in this field, but P5 positively worked on some projects relevant to this field and published them to GitHub.

P6 was studying HTI and using generative AI at least twice a week. The participant was considering the wage range in Finland because the country has a competitive salary compared to other countries. At the early stage of the job-seeking, P6 found it difficult to find a field-related job rather than a temporary job in Finland. The participant concentrated on networking with locals and met a senior Human Resource (HR) professional who could assist.

P7 was studying HTI and occasionally used generative AI. The only thing on P7's mind was whether the company requires Finnish as a mandatory requirement when talking about job-seeking. The participant was preparing a portfolio and CV because recruiters expect HTI candidates to have a good portfolio.

3.4 Data Analysis Methods

All the workshops were audio recorded and transcribed by Microsoft Office Word first, and then the author proofread the audio and transcriptions. The average length of a workshop was 50 min (min. 49 min and max.54 min). The author used the Grounded Theory (Jonathan Lazar, 2017) which is an inductive research method and useful for a new topic that has limited literature to build on. During the coding process, the author analysed the text and identified any interesting phenomena with a distinctive name or code. The author then iterative reviewed the data to identify new codes and categories that might lead to the most interesting codes. In practice, an affinity diagram (Plain, 2007) was employed in the *Miro* platform where the author made categories, themes, and codes. To create the affinity diagram, all the collected data was divided into separate sticky notes that covered only one aspect per note, and all notes were organized in a hierarchy form visualizing common phenomena, themes, codes, and relations among the data.

3.5 Research Platform

Voiceflow, seen in Figure 2, is the collaboration platform where conversational AI teams design, prototype, and launch amazing chat assistants. Below are the four main reasons why this tool was selected as the research platform.

Firstly, Voiceflow currently uses GPT for its AI features. For the free account, it supports the following AI models: GPT-3.0 (DaVinci) and GPT-3.5 Turbo (ChatGPT). While for upgrade accounts, it also supports GPT-4.0.

Secondly, Voiceflow provides the Knowledge Base (KB) which allows users to leverage their Probability Density Function (PDF), Uniform Resource Locator (URL), and Text documents data to power answers and define variables on personal Voiceflow Chatbot, for instance, add the specific company's webpage as a URL to provide company-related information. Voiceflow also can generate tasks, generate No Match, and generate AI steps. Generating tasks refers to creating content for your assistant to speed up the creation and optimize the content of conversations. Generating No Match responds to questions which assistant cannot answer, static Global No Match was recommended because designers can identify the content. Generative No Match is still an experimental feature for this reason and is not recommended to be used for serious production use cases. Generative AI Steps provide a prompt and have an AI assistant generate either a contextual response for users or a value for a variable. Based on this technique, designing a chatbot that responds to various situations is possible. At the same time, users can select an AI model to generate responses from the Knowledge Base, in this research, GPT-3.5 was utilized because of its versatility, intelligence, and ability to engage in human-like conversations.

Thirdly, Voiceflow has a friendly and intuitive graphical user interface (Figure 2), also its documentation is constantly updated to help designers study this platform.

Lastly, Voiceflow includes Conversation Memory which can automatically include the previous 10 turns in a single conversation into the prompt to provide context on what the user has been asking, specific entities or topics they have mentioned, and the response the chat assistant has provided. In this research, Memory and Prompt were applied since it is best to enable the most contextually appropriate response. When the user asks a question, it will be augmented with the previous 10 turns in this conversation, and the LLMs will generate an answer from both bits of data.

Furthermore, Voiceflow supports deploying to the webpage after complement which is beyond the scope of this research.

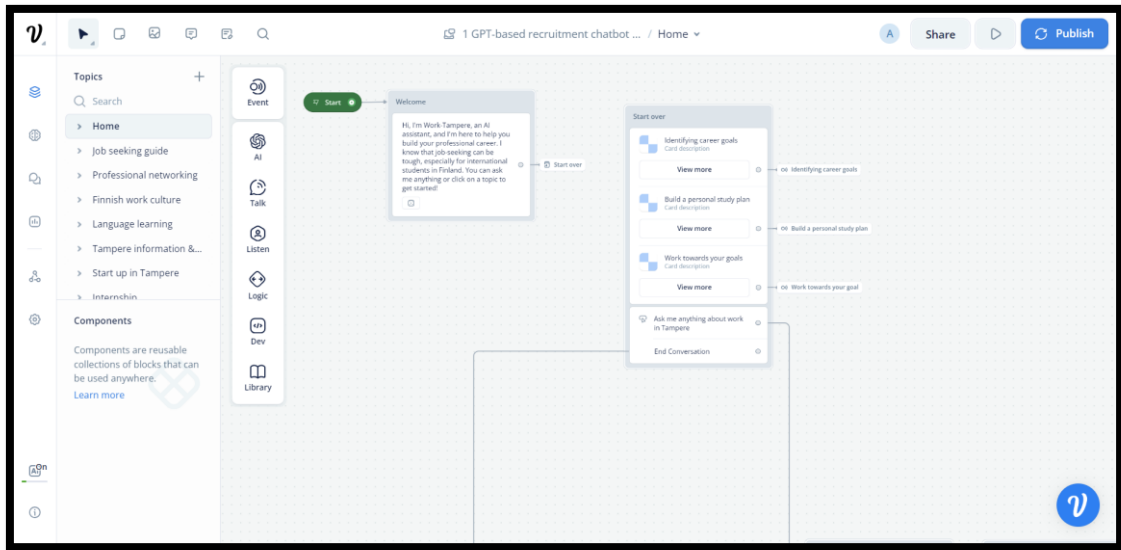


Figure 2 Graphical user interface of Voiceflow

4 Chatbot prototype

This chapter first explains the present recruitment chatbots applied in Finland to demonstrate the significance of this research topic, followed by the chatbot prototyping tools to choose a specific tool for this research. The last three subsection defines the details of the three stages of designing the GPT-based chatbot.

4.1 Recruitment Chatbot in Finland

Chatbots and GPT have grown in popularity in recent years, to figure out how GPT-based chatbots assist job seekers during the application process, I evaluated related products. Which companies used chatbots in Finland? The significant finding was that some companies deployed chatbots on their websites, but the majority did not. There are three significant findings. Firstly, current chatbots have limited capabilities. For example, Nokia's chatbot named Tom, will recommend blogs on why people choose this company, direct people to their career page, and help them through their career progression. Users can only select the options offered, it absent other interaction techniques like typing questions. Secondly, chatbots appear on the career page or job page on the company website, which reduces accessibility because users are required to exert some effort to discover a specific page. Thirdly, chatbots only support a limited number of languages. The chatbots of Verkko kauppa and Viking Line only support Finnish, Tom only supports English, and Aino from *encourage.fi* provides four languages while it is not a recruitment chatbot. Therefore, there is plenty of scope for research into GPT-based chatbots from the job seeker's perspective.

4.2 Chatbot Prototyping Tools

During this research, evaluating the chatbot is the main task which means that a chatbot prototype with AI features is required. Based on this, the author compared 14 potential design tools to determine the best option and Table 2 displays the four most relevant prototyping tools.

Table 2 Prototyping Tools

Name	Support interactive chatbot?	Support AI feature	Have Knowledge Base?	With pre-defined questions or open-ended questions?	Programming skills	Price	Conversation memory
Voiceflow	Yes	Yes, GPT3, Claude V1, GPT4 needs to be paid	Yes	Both	Coding understanding would be a plus	Free limitations, 20000 tokens(words)	10 turns of single conversation
Botpress	Yes	Yes, Based on ChatGPT (GPT 3.5)	Yes	-	Coding understanding would be a plus	1000 messages/month free, after that 0.005\$ per conversation	Yes, up to approximately 512MB of conversation
Figma	No	Yes	No	No	No	Free for students	No
Mocking Bot	Kind of	No	-	No	No	Free for limited use	-

The following criteria were examined throughout the selection process:

1. Is it available for free?
2. Could this tool create an interactive chatbot?
3. Could this prototype give a response after the user's input?
4. Could this tool build a prototype quickly? Or is it simple to use?
5. If this tool supports the AI model?
6. Could this prototype store users' input?
7. Could designers build their database for the prototype?

Those 14 tools include Voiceflow, Figma, Botframe, Mocking Bot, TARS, Botpress, Chatfuel, IBM Watson Assistant, Dialogflow by Google, Power Virtual Agents by Microsoft, GTP-4 API, meta-lama, mPLUS-Owl, and Visual-LLamA. Because of the tight research schedule, those tools that need programming skills were not selected at this stage, for example, IBM Watson Assistant, Dialogflow by Google, GTP-4 API, meta-lama, mPLUS-Owl, and Visual-LlamA. The problem of Power Virtual Agents by Microsoft is that generative AI is not available in Finland. Chatfuel is a chatbot tool that relies on Facebook and Instagram, and it needs to set all the conversation details. Even though it supports ChatGPT, it does not support its build database. The rest of the tools are Voiceflow, Figma, Botframe, Mocking Bot, and Botpress, all of them are available for limited usage, if the usage is out of range, then you need to pay for it. While the Botframe is a web-based product that supports the creation of chatbot conversation mock-ups easily and then displays the conversation on screen, it cannot create an interactive prototype.

Voiceflow is a no-code platform that allows users to design, prototype, and deploy a chat assistant or chatbot. It contains a free version and an upgraded version; I only discussed the free version in this research. The chatbot was created by designers that could output text, images, cards, and carousel messages, and it accepts buttons, text input, and voice input while only two modes could be used in the prototype: buttons with text input, and buttons with voice input. Buttons with text input were used for this study as the speech-to-text algorithm does not recognize "Tampere" which is the keyword in this study. Considering the user-friendly Voiceflow, it has a friendly user interface, maintains documents and video tutorials supports building a prototype quickly. Moreover, it provides AI models like GPT-3.0 (DaVinci) and GPT-3.5 Turbo (ChatGPT), where users have access to choose a specific AI model to generate output messages. Another advantage of Voiceflow is conversation memory, it refers to the chatbot prototype automatically include the previous 10 Turns (user inputs and system responses) in the conversation into the prompt as context, so the response from the chatbot can be more powerful. In addition, Knowledge Base (KB) allows users to build their data to power responses, and the data includes PDF, URL, and Text documents.

Overall, Voiceflow was selected as the prototyping tool because it fulfils the selection criteria.

4.3 Knowledge Base Design

A Knowledge Base is a collection of information that can be used to answer users' questions. During this study, building the Knowledge Base is essential to enhance the quality of the chatbot's response. A six-step procedure was used in Knowledge Base design, including: 1. Define the scope of Knowledge Base, 2. Gather information through web-content research, 3. Group information from the step 2, 4. Concentrate on one most relevant theme, 5. Organize information to a specific format to build the Knowledge Base, and 6. Test the Knowledge Base and improve.

Step 1: Define the scope of the Knowledge Base

International students face many challenges during their job-seeking process in Tampere even though Tampere plans to attract more international talents (*City of Tampere Integration Programme 2021–2025*, 2021). Generally, international students face a language barrier, missing local networking (Viitasalo, 2018), ignoring the local job-seeking culture, and information overload. Furthermore, the requirements for international students from Viitasalo (Viitasalo, 2018) included suitable personality, good English language skills, motivation, programming skills, communication skills and suitability for the job. Therefore, the scope of the Knowledge Base is job-seeking related information at Tampere, job-seeking culture at Tampere, and job-seeking relevant services at Tampere.

Step 2: Gather information through web-content research

At this step, the gathered information came from different sources, a total of 6 ways contributed to this stage and all collected information sources were saved (Figure 3). The important and most conventional way was keyword searching on the Internet, which contributed to general-level information. However, this method lacks reliability and trust, so comparing and validating collected information with other sources was necessary. The second way is collecting information through job-related research papers in Tampere. For example, the head of Immigrant Affairs in the City of Helsinki Annika Forsander mentioned approximately 60% of jobs in Finland are not advertised, employees found jobs through their social networks (Viitasalo, 2018). As well as, Tandem is considered a work-based language learning to support international talents and increase the hired possibility, and International House Tampere and Tampere universities can provide career life assistance (Abou Merhi, 2022). The third way is user manual learning from Management and Business faculty, this user manual was given to their students to better understand the

Finnish job market and culture. One key insight from this manual is to treat yourself as a professional instead of a student so that you will think professionally. Fourthly, I got support from the Trust-M project by reviewing the job-related information for immigrants and co-conducted a meeting which focused on immigrant service with an expert from International House Tampere (IHT). Building on this way, some web sources collected by the first way got validation. Fifthly, learned from the interactive *Tampere international talent service ecosystem* created by HIWE, in this ecosystem, work and economy, and professional networks were related to my study. The last way to gain information is by discussing with locals since they are familiar with the Finnish work culture and various local services.

Name	URL	Notes	Done	Tags
Ecosystem map. Online Whiteboard fo	https://miro.com/app/board/uXjVPiyh7a8=	Trust-M	<input checked="" type="checkbox"/>	
Womento mentoring - Västölitto.fi	https://www.vaestolitto.fi/en/immigrants/w	Women mentoring in Finnish	<input checked="" type="checkbox"/>	
Working and internships during studie	https://migri.fi/en/working-and-internships-		<input checked="" type="checkbox"/>	Student work hours
Working in Finland for employers - Edu	https://www.edunation.co/working-finland/		<input checked="" type="checkbox"/>	Information student
Sign up » FinTandem	https://www.fintandem.fi/registering-en-us		<input checked="" type="checkbox"/>	Language
Frontpage - Talent CV Gallery	https://talent-cv-gallery.com/		<input checked="" type="checkbox"/>	Tips
Finnish and Swedish language in Tamp	https://www.infofinland.fi/en/tampere/finnis		<input checked="" type="checkbox"/>	Language Tampere
Information about Tampere	https://www.infofinland.fi/en/tampere/infor		<input checked="" type="checkbox"/>	Information Tampere
Work and Enterprise in Tampere	https://www.infofinland.fi/en/tampere/work		<input checked="" type="checkbox"/>	Business start-up Tampere
Finnish working culture	https://www.infofinland.fi/en/work-and-ent		<input checked="" type="checkbox"/>	work culture
Taxation	https://www.infofinland.fi/en/work-and-ent		<input checked="" type="checkbox"/>	Taxation
Starting a business in Finland	https://www.infofinland.fi/en/work-and-ent		<input checked="" type="checkbox"/>	Business start-up Guideline
Conditions of employment and salary i	https://www.infofinland.fi/en/work-and-ent		<input checked="" type="checkbox"/>	Salary
Employee's rights and obligations	https://www.infofinland.fi/en/work-and-ent		<input checked="" type="checkbox"/>	Information
Foreign diploma or degree in Finland	https://www.infofinland.fi/en/work-and-ent	Related to degree	<input checked="" type="checkbox"/>	Information
Job application and CV	https://www.infofinland.fi/en/work-and-ent	Step by step navigation	<input checked="" type="checkbox"/>	Guidelines
Find a job in Finland	https://www.infofinland.fi/en/work-and-ent		<input checked="" type="checkbox"/>	Guidelines
Skills intelligence CEDEFOP	https://www.cedefop.europa.eu/en/tools/sk	Data for different filed about sk	<input checked="" type="checkbox"/>	
Work culture - Work in Finland	https://www.workinfinland.com/en/why-finl		<input checked="" type="checkbox"/>	work culture
Job Market Finland	https://tyomarkkinatori.fi/en		<input checked="" type="checkbox"/>	Information services
Employee's rights and obligations	https://www.infofinland.fi/en/work-and-ent		<input checked="" type="checkbox"/>	Information
Job market - Work in Finland	https://www.workinfinland.com/en/why-finl		<input checked="" type="checkbox"/>	Information services
Working in Tampere www.tampere.fi	https://www.tampere.fi/en/relocation-guide		<input checked="" type="checkbox"/>	work culture Information
SIMHE - Supporting Immigrants in Hig	https://www.metropolia.fi/en/rdi-rdi-project		<input checked="" type="checkbox"/>	services
Etusivu - Suomi - International House	https://internationalhousetampere.fi/fi/etus		<input checked="" type="checkbox"/>	Information services

Figure 3 Information sources

Step 3: Group information from the step 2

To categorize information from step 2, all information pieces were transferred onto sticky notes on a Miro board. Next, find sticky notes that are similar in some way, and group them on a new frame. Finally, labelled the group name with a theme label and selected the most relevant group for the next phase. There were 6 themes in this stage (Figure 4), including Finnish Work Culture, Language Learning, Local Information and Services, Work Requirements, Startup, and During Recruitment. The first 5 themes were constructive information, while the last one was the concentration of this study. Therefore,

6 sub-themes were used at this stage: job-seeking suggestions, networking in Tampere, access to job vacancies, research companies, job application, and employer's expectations.

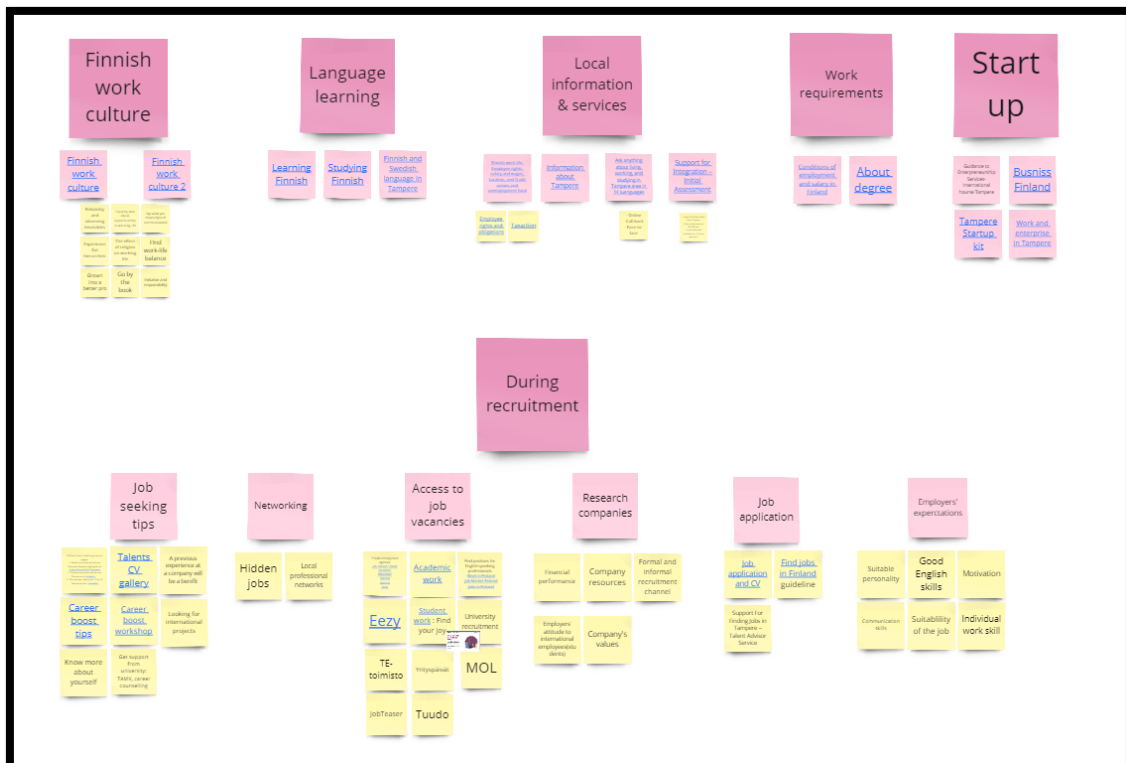


Figure 4 Six themes of information sources

Step 4: Concentrate on one most relevant theme

Compared with other information-based themes, During Recruitment was selected as the focus because it is a practical theme that international students can build themselves. At this point, I reviewed all the collected information and continued to gather relevant information that could contribute to this theme. Following that, content analysis and coding subthemes for the conversation flow design are performed, as stated in the following section. Hence, this chatbot focuses on the preparation of job-seeking, consisting of a job-seeking guide, building professional networking, growing as a professional and finding possible internship opportunities as shown in Figure 5.

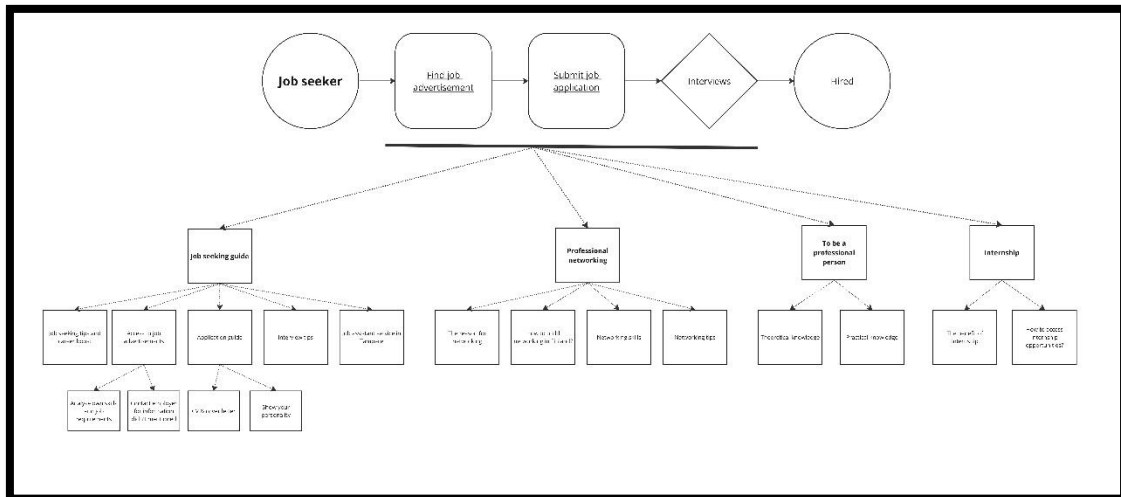


Figure 5 The focus of this chatbot

Step 5: Organize information to a specific format to build the Knowledge Base

URL(s), Sitemap, Text, PDF, and Doc are accepted by Voiceflow, therefore organizing information to those formats was the key task in this step. The collected information was organized into three sections, six themes, others, and an update (Figure 6). The first section was a document based on themes in step 3, while recruitment themes had sub-documents since overload information should be divided into sub-themes. The second section included all collection of sources, questions with answers related to job-seeking in Tampere, and the self-introduction of the chatbot. Lastly, with future research in mind, the updated information was separated into a new section. To be more detailed and intuitive, Figure 7 displayed all documents at the Knowledge Base in the Voiceflow platform, even though the order of documents was messed up, it is clear to notice them by names.

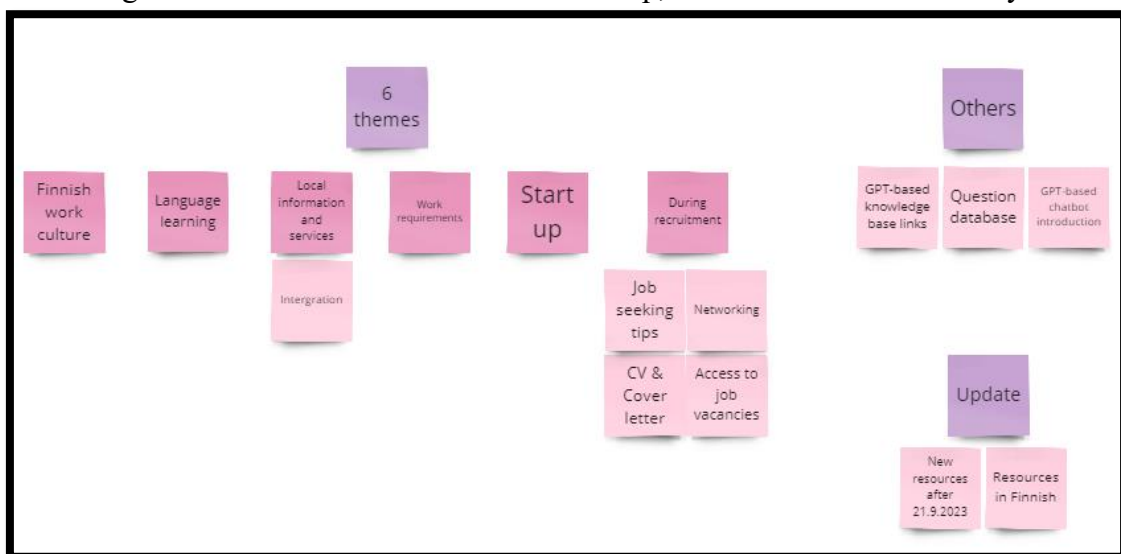


Figure 6 Structure of Knowledge Base

Knowledge Base		Search 15 documents		
	Name	Type	Status	
<input type="checkbox"/>	Access to job vacancies.docx	DOCX	✓	
<input type="checkbox"/>	Question database.docx	DOCX	✓	
<input type="checkbox"/>	CV & Cover letter.docx	DOCX	✓	
<input type="checkbox"/>	GPT-based knowledge base links.pdf	PDF	✓	
<input type="checkbox"/>	GPT-based chatbot introduction.txt	TEXT	✓	
<input type="checkbox"/>	New resources after 21.9.2023.docx	DOCX	✓	
<input type="checkbox"/>	Resources in Finnish.docx	DOCX	✓	
<input type="checkbox"/>	Networking.docx	DOCX	✓	
<input type="checkbox"/>	Work requirements.docx	DOCX	✓	
<input type="checkbox"/>	Job seeking tips.docx	DOCX	✓	
<input type="checkbox"/>	Language learning.docx	DOCX	✓	
<input type="checkbox"/>	Start-up.docx	DOCX	✓	
<input type="checkbox"/>	Integration.docx	DOCX	✓	
<input type="checkbox"/>	Local information and services.docx	DOCX	✓	
<input type="checkbox"/>	Finnish work culture.docx	DOCX	✓	

Figure 7 Knowledge Base in Voiceflow platform

Step 6: Test the Knowledge Base and make Improvements

The final stage in creating a Knowledge Base was to upload all material to the Knowledge Base in the Voiceflow. In addition, testing and improvement were the following phases, my supervisors examined all information sources and tested it through the chatbot prototype. Building on their contribution, several Finnish sources were added to the Knowledge Base, for example, *duunitori*, *Oikotie*, and *te-office*. This was because LLMs have no technical limitations and the AI model will power response from those sources.

4.4 Design of Conversation Flow

During the iterative design process for chatbots, the most significant phase is designing a conversation flow, following includes testing with target users, analysing data from tests, and improving design (Choi et al., 2021).

According to the characteristics of Voiceflow, it is possible to build multiple conversation flows based on various themes in a chatbot. The main conversation flow is how the chatbot displays the major capabilities and guides users through the conversation. In this study, the chatbot will greet users with a brief self-introduction, then present three themes

by cards based on the general job-seeking preparation, a button for free-form questions, and also offer the option to end the conversation (Figure 8). These three topics are identifying career goals, building a personal study plan, and working towards your goal separately, which provides a blueprint for how to begin the preparation of job-seeking (Wanberg et al., 2020). Under the branch of the main conversation flow, each topic has its independent conversation flow and provides a button for initiating the conversation from the beginning.

To begin, Figure 9 indicates that identifying career goals starts to guide users to think about their goals building on the understanding evaluation process in the recruitment, and understanding Finnish work culture can help users to better choose a position or company that matches their personality. The chatbot also provides recruitment information and services in Tampere, so that users can seek assistance if they cannot handle their problems.

Following that building a personal study plan (Figure 10) on this topic, the chatbot provides three factors that affect the study plan. The first factor is the timeline of the study plan, typically, international students will spend two years for their master's degree study. Under the job market situation, demonstrating professional competence in conjunction with the study plan is an essential object for a student. The second factor is language, the chatbot explains the reason why learning Finnish is important, and provides learning methods, and courses to study Finnish such as Tandem approach learning. The last factor is the employer's expectations which assist users in planning how to develop their skills for their future career life.

Finally, implementing the plan is the only way to make progress towards the goal (Figure 11). The practical things include following the job-seeking guide to prepare relevant objects and starting to build professional networking for a better understanding job market or getting references. In addition, the chatbot would like to transmit that being a professional person is important. Experts are more demanding than students so that the user can challenge themselves and tap into their strengths. Furthermore, finding an internship opportunity is a method to gain work experience to present on the CV.

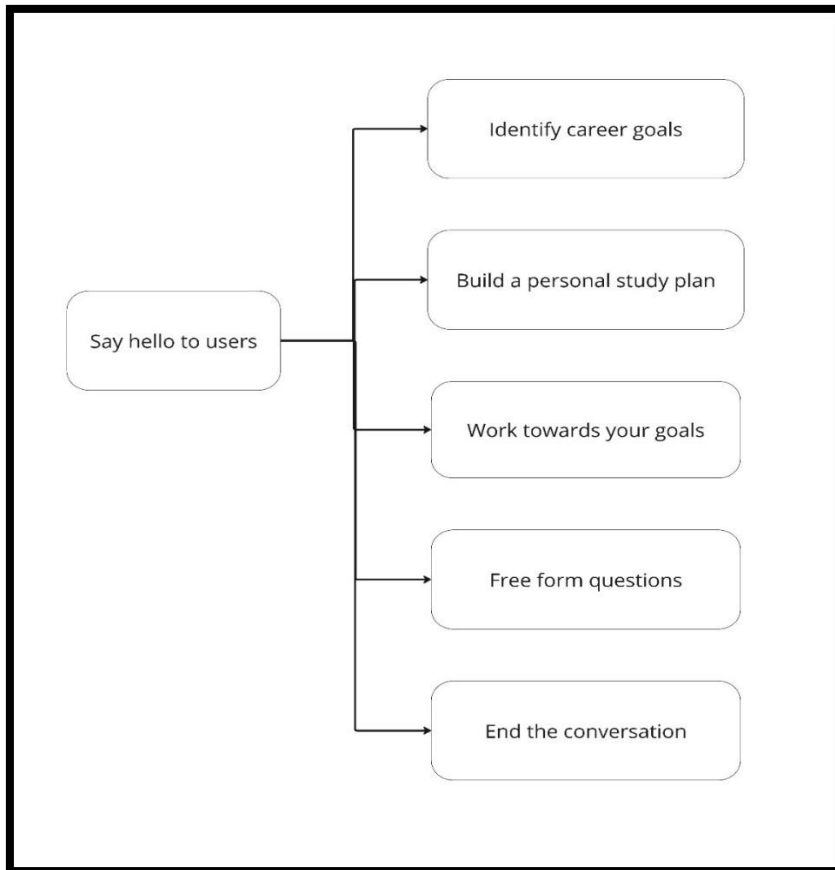


Figure 8 The main conversation flow for the chatbot

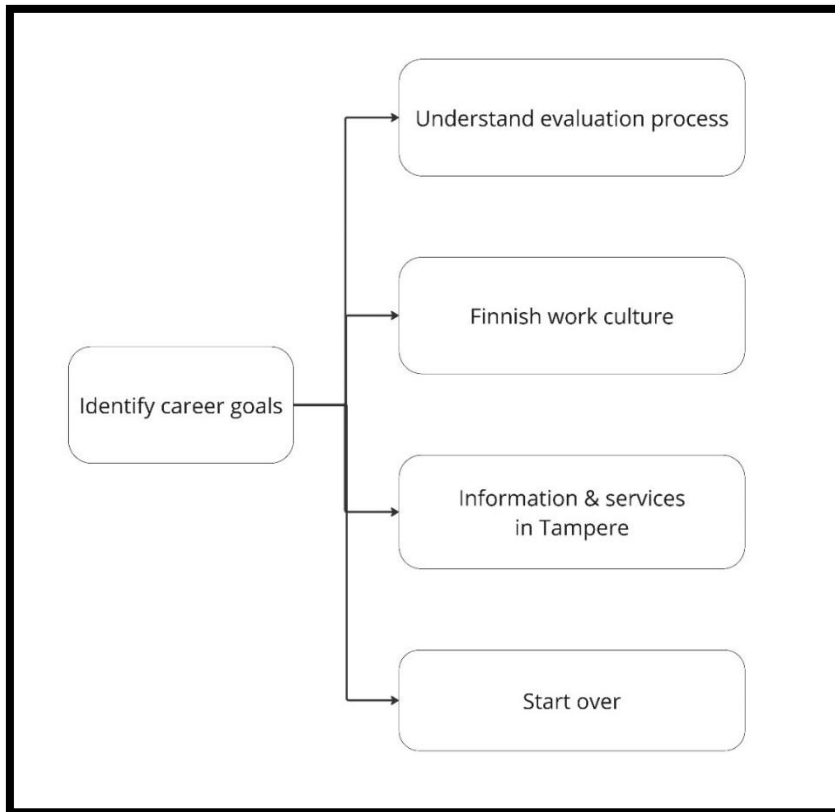


Figure 9 Conversation flow: Identify career goals

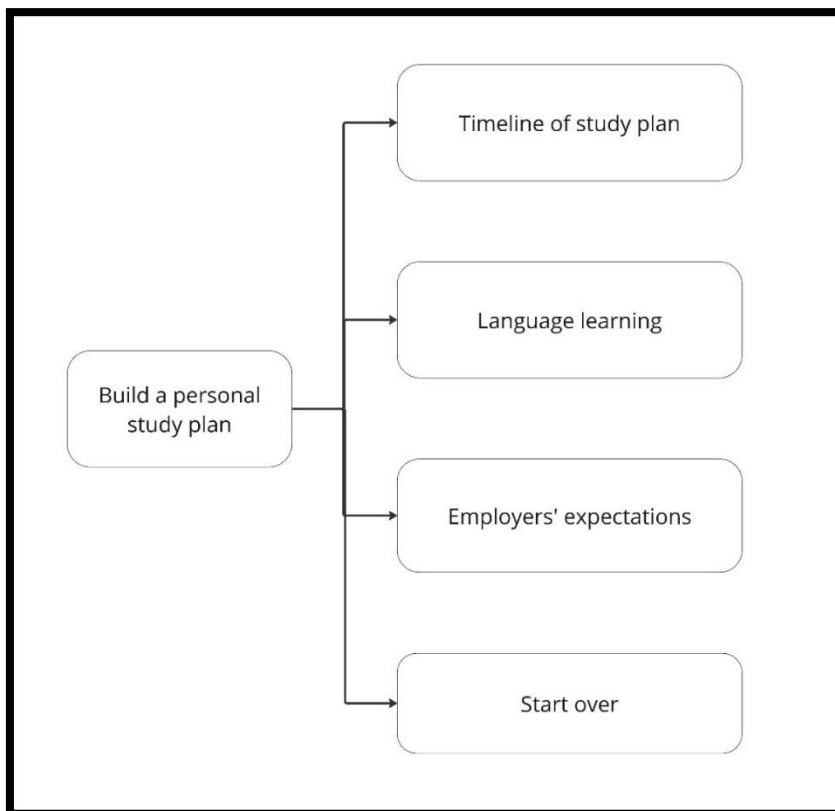


Figure 10 Conversation flow: Build a personal study plan

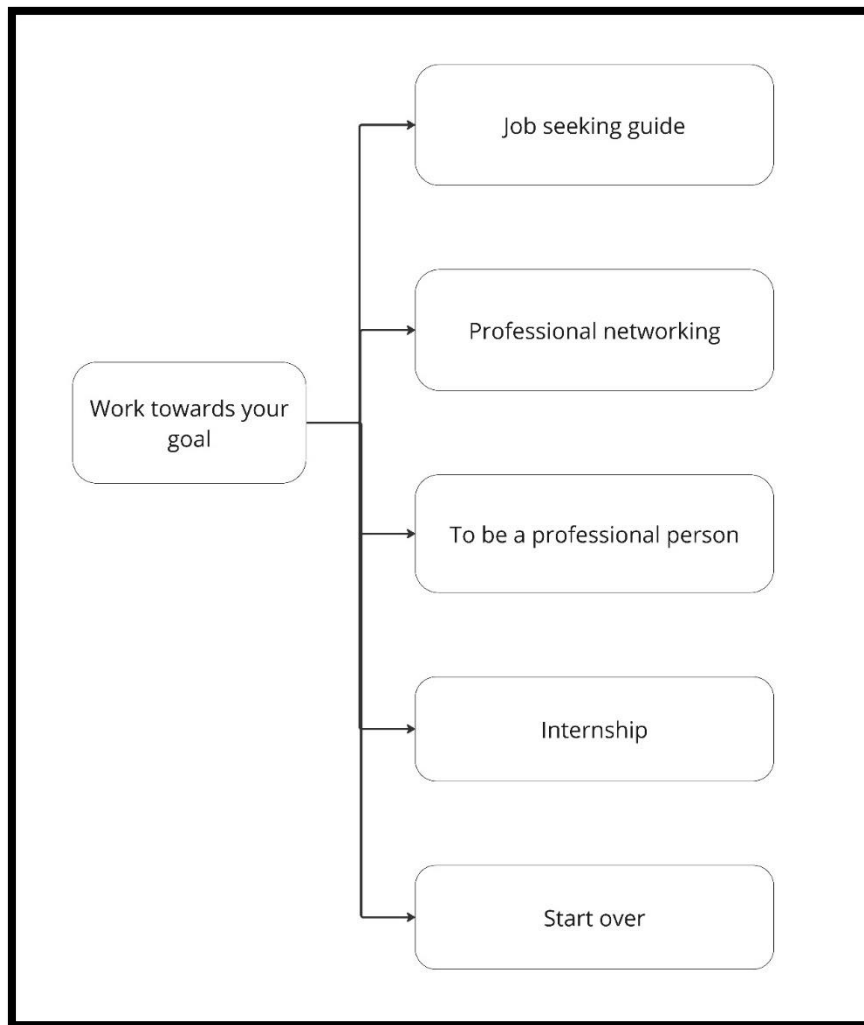


Figure 11 Conversation flow: Work towards your goals

4.5 Technical Design

The GPT-based chatbot designed for international students in Tampere is named Work-Tampere. According to the guidelines for chatbot conversation design (Silva & Canedo, 2022), assigning a name to a chatbot can enhance its naturalness and foster engagement in conversations. Drawing on the insights from conversational interface design, considerations such as branding, interaction methods, and the choice and use of various media for input and output will contribute to the system more naturally and intuitively (McTear, n.d.). In this study, the chatbot's branding was achieved by incorporating a picture representing Tampere, which indicates this chatbot is specific to the city. For the interaction method, button clicking and typing were selected based on the capabilities of the design platform. Buttons work for initial conversation flow and typing for addressing free-form questions. The conversation user interface displayed in Figure 12, the chatbot first starts

with a self-introduction and provides the options for users to either continue with the initial conversation or pose free-form questions.

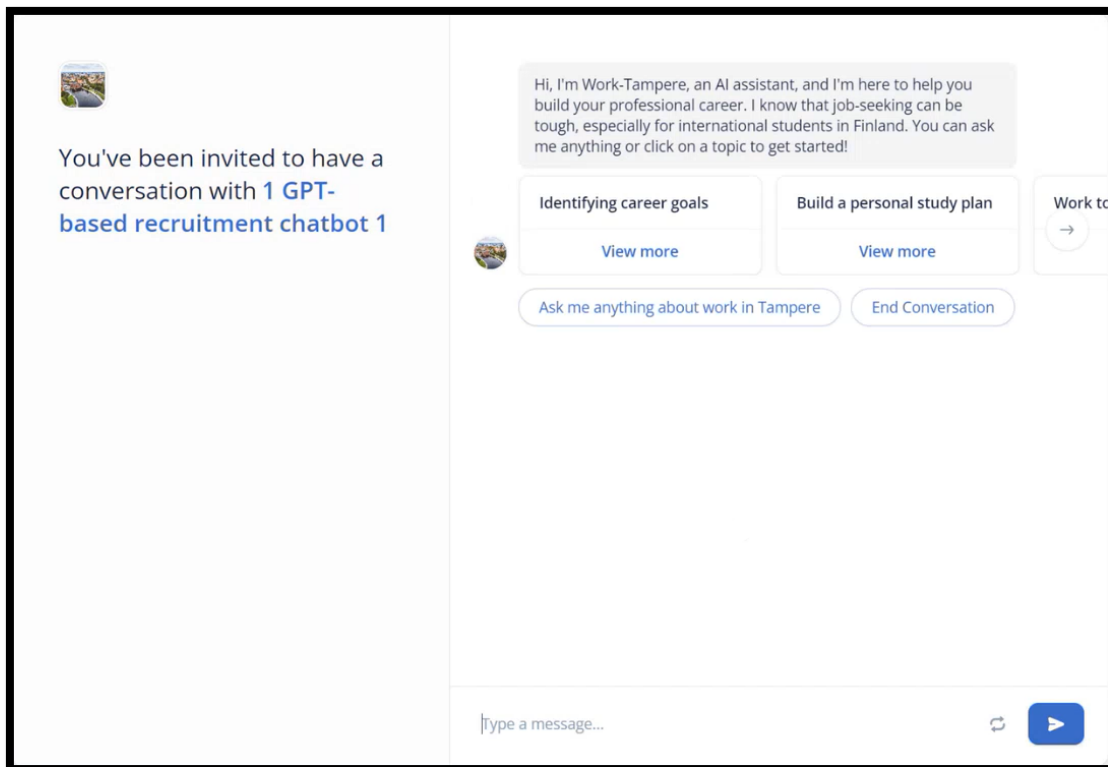


Figure 12 The user interface of GPT-based chatbot

In the architecture of the chatbot, topics, intents, entities, utterances, and prompts were strategically employed to construct a flexible and dynamic conversation. Topics are subsections of the chatbot that correspond to a specific function. During this study, topics link to themes within the Knowledge Base, where each theme corresponds to a topic. Intents refer to the purpose of the user in the conversation flow, sub-themes in Knowledge Base which were seen as intents. Figure 13 presents a comprehensive overview of all topics and intents embedded in this chatbot. Entity means a data point or value which can be extracted from a conversation. Utterance means the inputs from end-users. To give an example, the question “How to design my CV and cover letter” serves as an utterance, with the entities extracted from it are “CV” and “cover letter”, and then signifying the user’s intent to comprehend the requirements of the application materials. Hence, the relevant topic in the chatbot is identified as “Job-seeking guide”. Additionally, the concept of a prompt is a pre-defined message to guide GPT in generating specific responses to questions. In this context, the prompt for users’ free-form questions is specified as “Based on the user’s question, respond with two brief sentences in English” (Figure 14), and the chatbot will generate responses based on GPT-3.5 from the Knowledge Base (Figure 15).

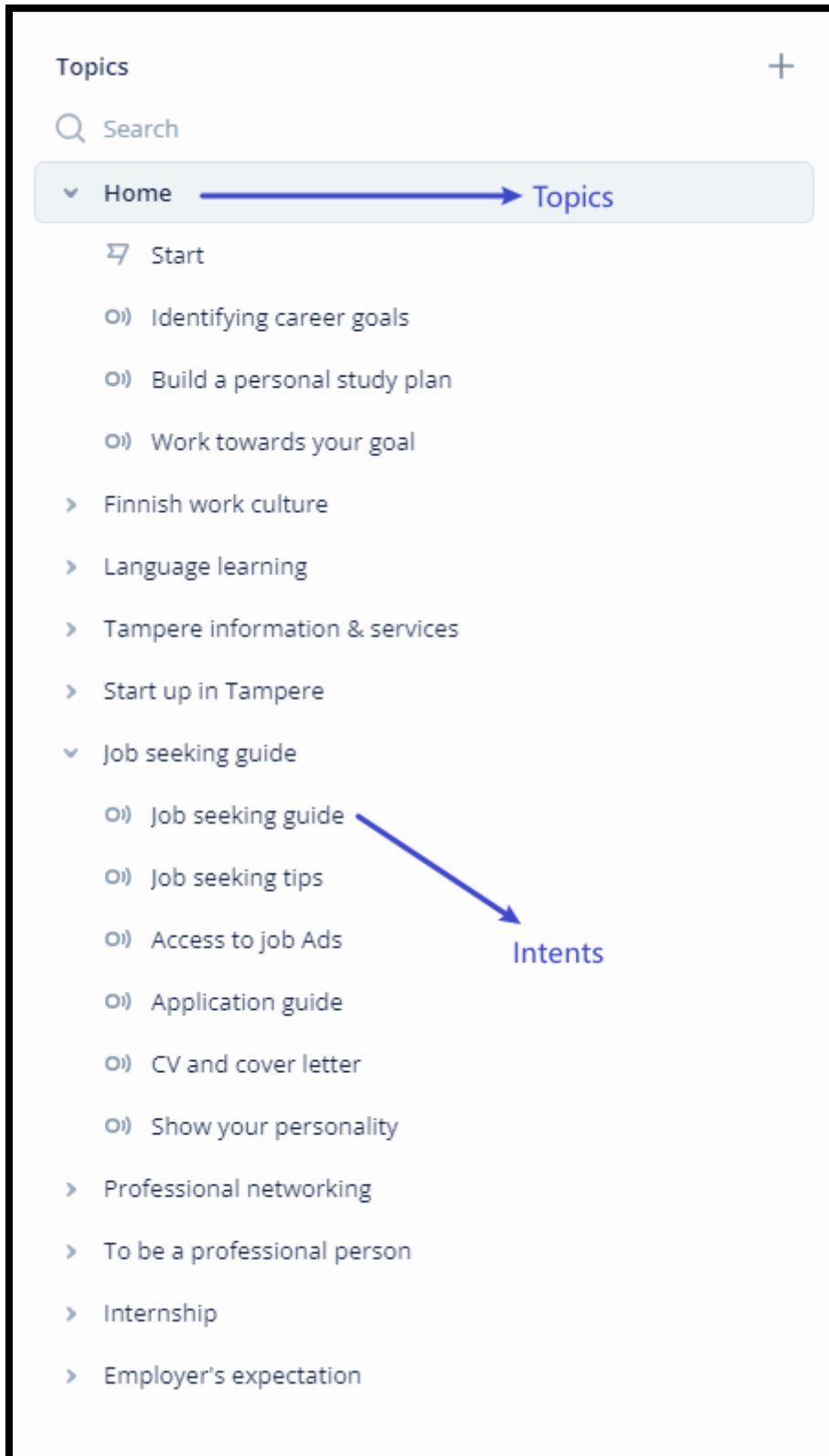


Figure 13 Topics and Intents in this chatbot

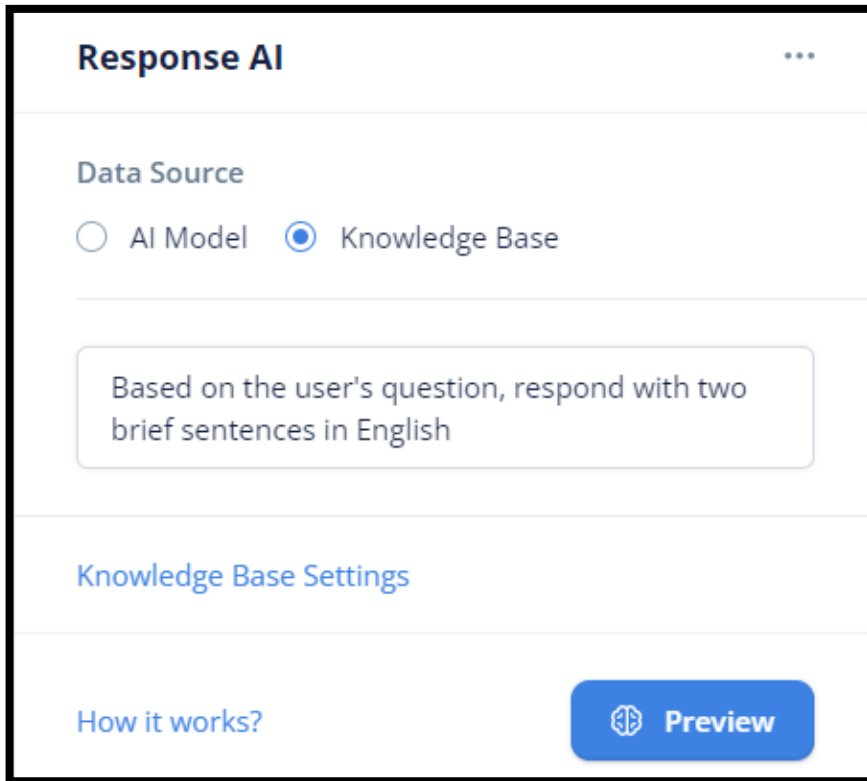


Figure 14 The prompt for free-form questions in the chatbot

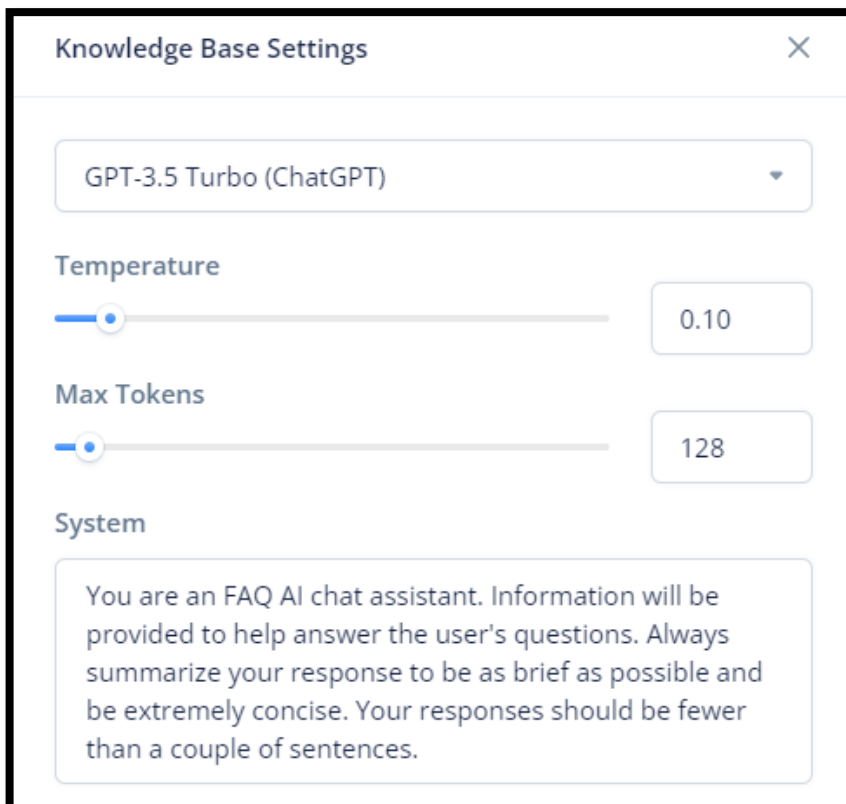


Figure 15 Knowledge Base setting

5 Evaluation results

I first investigated the challenges behind the job-seeking process from international students' perspective, then analysed the participants' expectations for the GPT-based chatbot. Finally, I analysed the ways to pose questions and participants' interaction behaviours with the chatbot.

5.1 International Students' Perceived Challenges Related to Applying for Jobs

The participants stressed practical challenges and external challenges during the early stage of the job-seeking process.

5.1.1 Practical Challenges Faced by International Students

Language is an important and the first challenge for international students, and it is unclear what exact Finnish level is needed in specific positions. The majority of participants mentioned the language barrier during the workshop, for example, P3 worried that people without Finnish skills have less competitiveness compared to those who speak Finnish. Participants who were willing to learn Finnish anticipated knowing the certain requirements of Finnish so that they could work on the explicit target as the preparation for job hunting.

“In Finland, Finnish is mandatory. I got rejected from at least 50 jobs. They need Finnish even though no one was Finnish people.” (P4)

Second, Finnish work experience or experience in Finnish culture seems to affect the job opportunities for international students. Some participants had work experience in their original country in their current field or other fields, but from their application experience, that work experience did not advance them to obtain job opportunities. While participants had no work experience, their challenge was how to demonstrate personal work attributes to recruiters and how to get the first job opportunity.

“Asking the university students for three years of experience for this job.” (P2)

“There is another thing I learned from networking with different people, if you are experienced, it does not matter if you are experiencing outside of Finland. They require some kind of experience in Finland to be considered for the job. I do not have experience in Finland, but I have experience back in my home country. So that is another obstacle which does not allow me to find jobs.” (P3)

“I also have this seen: Finnish work culture. For example, there are two companies outside of Finland, one is Amazon, the other one is Nordea. As a person, I have two years of experience in Amazon and this guy same thing on all the profiles and has like 2 years of experience in Nordea. So, they will consider the person from Nordea more than the person from Amazon because Nordea is a Finnish company, and they know that they have a different work culture going over there.” (P3)

Third, the noteworthy phenomenon was that P1 and P2 were concerned that they would be hard to select compared to the fake appealing applications, and all participants were struggling in preparing application materials including CVs, cover letters or motivation letters, and portfolios. The example, participants asked questions related to building and refining their applications. Hidden behind this phenomenon is the ability to demonstrate competence. At the same time, focusing on the improvement of capability is more important than excessive worry, because it is impossible to avoid faking applications and difficult to discriminate from recruiters' perspective. From a practical point of view, learning to tap into the abilities, storytelling skills, and language skills (English) is necessary to improve. P4 and P6 indicated that they do not understand how to display personality or incorporate interests and experience in a job application.

“What should be in my CV for a software engineer position? How long the CV should be? Give me the list of topics that should be present.” (P4)

“As a beginner without any work experience, how can I make my CV appeal?” (P5)

“... I want to know how many projects you think I should have done so far, to be eligible to apply for an intermediate or senior level designer position?” (P7)

5.1.2 External Challenges Faced by International Students

In this study, participants encountered three external challenges related to applying for a job.

To begin, participants as job seekers found it difficult to keep up with companies for the information gap. This was demonstrated by the understanding of job descriptions, to give an example, participants had their interpretation of whether they have the skills needed for the position and what specific skills need to be enhanced, while the evaluation process is a black box for them. Participants with experience of rejection did not receive any constructive feedback on their job application (P4 and P5), which causes the information gap to persist and is difficult to narrow. In other words, poor information affects participants' self-improvement and reduces their motivation in the job-seeking. Additionally, this divergent understanding can also lead participants to deviate from their goal position. However, this type of challenge is less relevant because the chatbot cannot address these structural challenges.

“It (job-seeking) is a useless activity for me right now. Because I have applied to nearly 200 jobs so far in the last 16 months. But I did not get any feedback or interview chance.” (P4)

“I did not get proper feedback from the company regarding the rejection and if they gave any proper reason, it would be great for applying for the next job. I can improve what I am missing.” (P5)

Second, intense competition is a common occurrence which is the result of the abundant supply of labour. Numerous candidates competing for one position as P3 said: “I find there are not many jobs right now and a lot of competition for a single job. Like 70 or 80 people for one position. I am pretty sure that is for one position.”. In this case, students who are job seekers are prone to receive rejection letters, especially when the job market requests graduates with 3 years of work experience as P2 mentioned. Furthermore, intense competition results in rejection letters, which leads to a negative job search for those who do not receive positive feedback. P3 stated that the big reason for not actively looking for a job currently is this, and P4 was hesitant to give up looking for a job in Finland.

“How to match skills with the position? For example, 3 years of experience is missing for many students.” (P2)

Third, from the participants’ perspective, they face rejection because of nationality, which is related to the company’s attitude toward international talents. P4 had rejection experiences because of showing the face on the CV, and P2 mentioned some companies do not prefer internationals.

“There are some companies that do not prefer internationals. So, we have to go through to show that even though we are not from your nationality, we are as good as them. So that is a challenge that comes up.” (P2)

5.2 International Students’ Expectations Towards the Chatbot

One of the goals of the evaluation workshop was to explore the participant’s expectations for the GPT-based chatbot, which will be covered in what follows.

5.2.1 Support for Understanding Jobs, Writing Job Applications, and Preparing for Interviews

All the participants were hopeful that chatbots could provide them with local, reliable data that would help them improve their employment skills and increase the hiring rate. Employment skills include job skills, job application materials skills, and interview skills.

Job skills refer to individual competitiveness which can help participants build themselves, four aspects were highlighted in this study. First, participants expected to get a better understanding of the professional skills for a specific job title. For instance, P1 and P2 underlined the technical skills for a robotics vacancy that would help them to improve and meet the requirements of the company. Second, soft skills are important to match the company culture and team environment. Participants struggled to show their personalities and were uncertain how to match their soft skills with career life. P2 emphasized that having a good LinkedIn account is beneficial because it is a way to display what have you done and what is in progress. Third, all participants stated language skills, they hoped

to know the necessity of Finnish for a position, as well as the specific Finnish level in a certain position for the preparation. Last, participants noticed networking is a key point to access hidden positions, therefore they desired to expand their social networking.

Typically, job application materials consist of a resume, cover letter or motivation letter, and portfolio. Application materials are necessary for recruiters to assess candidate's competence, participants hoped to display their skills based on local job-seeking rules. There are two types of questions from participants which depend on their situation. Participants who had a ready resume, desired to get improvement suggestions or evaluation feedback from the chatbot. However, for those who were struggling with the first step, asked the checklist of a resume.

“How to refine my resume?” (P2)

“As a beginner without any work experience, how can I make my CV appeal?” (P5)

As for the interview skills, participants hoped to get practical and interactive support from chatbots. For example, giving a list of typical interview questions or suggestions so that they can prepare for an interview. It is more advanced if the chatbot can act as an interviewer and users can practice interview questions with it.

“I will use the chatbot as an interviewer and let it ask me some interview questions, for example, how many years of work experience do you have? What are your advantages? Do you have to qualify for this job?” (P6)

“How to refine my interview for a Robotics vacancy?” (P2)

5.2.2 Insights into the Local Job Market

While recruitment is a two-way process, and for job seekers, conducting the appropriate background research can help them to be better informed and apply. Hence, participants expected chatbots could provide insights into the local job market, which consists of job market research, job opportunities, job descriptions, and job factors.

In the early stage of job-seeking, participants need support relevant to job market information since it might influence their decisions. For example, the job searching channels in Tampere, specific companies in a certain field, and job market situations. Participants anticipated getting specific job vacancies directly by chatbot because of time saving, P5 typed this type of question in the chatbot “Job opening for data analyst in Tampere?” After access to job opportunities, participants had higher requirements for the chatbot that explained the job description related to talents' responsibilities with reasonable salary and background information. An understanding of explanation can provide a blueprint which guides job seekers' applications. Furthermore, participants would conduct background research about the company, including salary range, work environment, inclusivity, work hours, and benefits if they were hired. Those job factors will influence the decision of

candidates, they expect a position where the company culture matches with them. P2 emphasized the importance of the company’s attitude toward international talents so that the work environment is inclusivity.

“Do I get a coach or leader to help me when I am at a lost in a project or something?” (P5)

5.3 Overall impression of this chatbot

During the evaluation workshop, a short version User Experience Questionnaire (UEQ) was applied to collect the initial impression of the GPT-based chatbot from participants. Then the author analysed collected data by the Short UEQ Data Analysis Tool, the result showed that this chatbot is a constructive design with both pragmatic quality and hedonic quality (Figure 16).

Item	Mean	Variance	Std. Dev.	No.	Negative	Positive	Scale	Short UEQ Scales	
1	1.3	0.9	1.0	7	obstructive	supportive	Pragmatic Quality		
2	1.4	3.0	1.7	7	complicated	easy	Pragmatic Quality		
3	0.9	1.8	1.3	7	inefficient	efficient	Pragmatic Quality		
4	0.7	2.9	1.7	7	confusing	clear	Pragmatic Quality		
5	1.1	2.5	1.6	7	boring	exciting	Hedonic Quality		
6	1.7	1.2	1.1	7	not interesting	interesting	Hedonic Quality		
7	0.3	1.9	1.4	7	conventional	inventive	Hedonic Quality		
8	1.0	2.0	1.4	7	usual	leading edge	Hedonic Quality		
								Pragmatic Quality	1.071
								Hedonic Quality	1.036
								Overall	1.054

Figure 16 The result of a short UEQ

In addition to the quantitative data analysis, the data included qualitative group interviews. Participants highlighted that the domain-specific chatbot had an intuitive conversation flow to assist them in posing more questions. The chatbot was able to provide clickable links based on participants’ input messages which left a deep positive impression. It was perceived to be a significant advantage compared to current ChatGPT. Lastly, participants were impressed with the chatbot’s problem-solving skills that build on their situation. Like P4 said, the chatbot can provide comprehensive information about what should need to be present on a CV.

“I asked the chatbot what should be presented in the CV, and it provided all the needed information.” (P4)

Additionally, the quality assurance of the chatbot is necessary in this case, because P1 found several infinite loops and unexpected conversation endings in the conversation flow. Hence, participants would like to recommend this chatbot to their friends if it is a complete product with good quality assurance.

“I would like to introduce my friend to this complete because I see that there are still some major bugs. I would love to tell you that you still have bugs inside this. I would like to introduce my friends because it is really good that it provides a lot of information. Even I am impressed.” (P1)

“I think the development, yes, the links are a hugely important thing that even GPT cannot solve at the moment because whatever the links that charity gives, it never works,

and Google bar is known for giving wrong links. So, if you can develop this, it is going to be very good.” (P2)

“I think everyone should try, because everyone’s way of asking is different, everyone’s mentality is different, so we will get a lot of inputs and outputs from them.” (P5)

To summarize, this GPT-based chatbot was effective, supportive, meaningful, and useful in assisting international students’ job-seeking process and it is valuable to conduct deeper research.

5.4 Interaction between participants and the chatbot

This chatbot has an initial conversation flow if the participants are at a loss regarding a specific inquiry to pose, and an input area for free typing questions. Based on the log data from workshops, the majority of participants preferred to interact with the chatbot by typing instead of clicking preconfigured buttons. At the same time, whether it is a conversation flow or a free question, the job-seeking guide and how to increase professionalism are popular themes. Therefore, this chapter explains how participants interact with the chatbot, their methods to pose questions and potential connections between the utilise of the chatbot and their background.

5.4.1 Initial impressions have an impact on how participants interact with the chatbot

In general, participants checked the conversation flow first and changed to typing interaction when they found the chatbot only provided general-level information which was not eligible for their real situation. During the typing interaction, participants were interested in the reliability of the information provided by the Knowledge Base, and the capability of this chatbot. The reliability of information would assist participants in trusting this chatbot and seeking for help from it while exploring the capability is a method to get better assistance from the chatbot. After several rounds of conversation with the chatbot, participants understood the capabilities and started to decide on the main interaction. P1 explored the different conversation branches to understand the whole structure and did not pose any free-form questions. While other participants had an initial impression of the chatbot and started to ask questions based on their circumstances from different angles. P2 underlined it is possible to post content questions and receive informative responses, however, from the instruction of the chatbot and the conversation flow where did not display those abilities of the chatbot. Furthermore, P2 declared it would be difficult for users to format more questions without displaying the functionalities.

5.4.2 Methods used by participants posed questions

The ways of asking are critical because the received responses are based on how the GPT model understands the input message. Based on the analysis, participants had various ways to ask questions which can be divided into well-defined questions and opinion-based questions. The chatbot seemed to perform better when the questions were well-defined.

Well-defined questions refer to questions with short context and clear descriptions. Participants who had positive perspectives for the chatbot typically asked context questions and follow-up questions.

Typically, participants posed questions based on their background and their career goals, for example, “What qualification will they look for in an electrical engineer?” (P2) This question provided the background was electrical engineer position, and the focus was the qualifications. In addition, participants found the chatbot’s response was lack of readability, then typed “Can you give it in point form, so it is easy for me to understand” (P2) to increase the readability of responses.

“As a data analyst job seeker what is necessary?” (P5)

Another well-defined question was the follow-up question because participants recognized the chatbot has the conversation memory. P2 typed in “Robotics interview questions” first, and after reading the response, “How to answer them?” was the next question. P2 was statistical with the response to those questions.

Opinion-based questions will reduce the performance of the GPT-based chatbot, especially questions that are contextless, out of range, misleading, self-awareness, and have long complex descriptions.

Firstly, the chatbot would match participants’ questions with the Knowledge Base, then it would provide the most related general suggestion if they asked contextless questions. A question like “Do you have some recommended job titles?” (P6) is a challenge to answer even for face-to-face communication because it is contextless. Secondly, the chatbot only support questions related to job-seeking in Tampere, out-of-range question will not be answered properly. For example, log data shows this question “Do you know any job opportunities in Helsinki or Rovaniemi?” (P6) The chatbot navigated the participant to a job-seeking guide instead of answering it. Thirdly, misleading questions were confusing for the chatbot, to give an example, one question confused the concepts of interviewer and interviewee, then it was impossible to get the correct answer. Fourthly, participants with negative perspectives on job-seeking tried to ask self-awareness questions. To give an example, “If I am studying information security, what would my hobbies be?” (P4) There seems to be a need for guiding participants to know more about themselves through the chatbot at the early stage of recruitment. Lastly, long complex questions were asked by participants, which got unstable responses from the chatbot. Literally,

the participants' insight was that providing enough context to expect a better response, while the question was too long with several questions which would influence the chatbot's performance. For instance, "I am a master's student doing Human-Technology Interactions in Tampere, Finland and I am looking for jobs in the design field and to be more specifically UI/UX designer jobs. Do you think Finnish is mandatory for a designer and if yes, what level of Finnish is exactly needed for each level of position?" (P7)

Additionally, one phenomenon was that formulating questions was the biggest challenge for participants. Even though P7 had a context-based mind model during the interaction with the chatbot, the participant provided a too-long and complex question which overflowed the input box. Then the participant copied the first question to the input box and modified several expressions, so P7 spent too much time on this which caused the participant to be frustrated.

5.1.3 The potential connection between the chatbot's usage and the participants' background

The noticeable fact is the performance of interaction with the chatbot, majority of participants acquired proper responses from it.

There are two commonalities related to receiving informative responses from the chatbot, first, those participants had a high frequency of using generative AI (P1, P2, and P4) which means they were familiar with how to interact with generative-based products. Second, they have a technical-related background which helps them quickly find the method to interact. For example, P5 found the chatbot could not answer ambitious questions and then changed the strategy to get the desired response. On the contrary, participants with negative feedback about this chatbot do not have a technical background, they mentioned the chatbot did not answer questions directly and only provided general suggestions. Furthermore, they preferred to type opinion-based questions as mentioned in the previous subsection. For instance, "Do you have some recommended job titles" (P6) is a contextless and opinion-based question. In this situation, P6 changed the interaction way to click buttons while found the chatbot did not answer the questions directly.

6 Discussion

This GPT-based chatbot was effective, supportive, meaningful, and beneficial in assisting international students' job-seeking process and it is valuable for conducting more deeper research. While this study still had available space for improvements, and faced technical challenges and challenges related to the use of the chatbot. Therefore, in this chapter, three potential improvement perspectives are provided first, followed by the challenges and limitations of this research. Furthermore, future work is presented.

6.1 Implications of the Findings for GPT-based chatbots

6.1.1 Improve the conversation design according to guidelines

Based on the conversation design guidelines of the chatbot from Silva and Canedo (Silva & Canedo, 2022), transparency, naturalness, and emotionality should be considered during the chatbot conversation design process.

Transparency refers to the chatbot's ability to present capabilities, acknowledge limitations, make suggestions, and ask for clarification. From this research, participants were interested in the capabilities of the chatbot and emphasised the importance of displaying the abilities in a way which can assist them in interacting with the chatbot. Especially the feature that users can ask questions at any time, but the initial conversation flow will misdirect this. Which caused users cannot recognize the typing function. Acknowledging limitations is an essential improvement since the participants asked several out-of-range questions. If the chatbot can acknowledge limitations properly that can increase the user experience and save their time when it does not meet their situations. In addition, the majority of participants suggested the chatbot provides more personality and concrete suggestions to them because general-level suggestions cannot help too much. For example, collecting necessary background information from users first, and considering users' backgrounds before giving responses is a method to provide a personalised experience. As for the clarification from the chatbot side, it could be thought of as a future improvement which was not touched on in this study. For instance, the chatbot is allowed to ask follow-up questions if it needs more context information from users.

Naturalness means the chatbot supports self-introduction, addresses the user by name, handles small talk, provides echoing responses, and uses casual language during the interaction. In other words, a chatbot with naturalness can engage users to enjoy and continue the conversation. Based on participants' feedback, they expected the chatbot to be more personalized, answer their questions directly, and have a soft response if it cannot answer their questions.

Emotionality describes that a chatbot can offer exclamatory feedback, graphical media, emphatic messages, and humour responses. The output of this chatbot contains graph-

ical media and simple emphatic messages. Improving the output messages to more emphatic and humorous is essential. Job-seeking is tough work for most job seekers, if the chatbot can understand their situations and provide proper comfort messages, it would be an advantage compared with the current chatbot.

6.1.2 Design for improving chatbot user experience

The key to promoting the use of this chatbot is increasing the user experience and being recognized by users. The findings imply that the conversation flow, the Knowledge Base, and the response messages could be improved.

In the conversation flow, consistency and accuracy of the description are essential to users' understanding and continuing with the conversation. At the same time, this chatbot provides free-from questions interaction, a method to assist users in getting better responses is incorporating question guidance into the conversation flow. Furthermore, powerful personalized services can help users locate problems and develop themselves. For example, a slot-filling conversation of mixed-initiative (McTear, n.d.) in which both system and user can take control of the interaction, in this way, the chatbot collects necessary information from users and gives a personalized response.

A Knowledge Base with a certain scale, reliable source, and updated in real-time can enhance users' trust in the chatbot and have references for its assistants. From the findings, the Knowledge Base provided clickable links and valuable local information or services, while participants suggested boarding it to a certain scale by cooperating with local organizations like International House Tampere and conducting more tests on the chatbot. On the other side, checking the source of the information can assist users in determining whether the information is valid, this is a method to create trust between the chatbot and users. The update of the Knowledge Base is critical because the information is specific and real-time, outdated information may mislead users and lost their trust. Hence, maintaining the Knowledge Base is an urgent aspect of the next-generation GPT-based chatbot.

Another practical prospect in the study is that the chatbot could help job seekers effortlessly follow the conversation state and review the conversation history. If this feature was employed, the chatbot could better support users to study how to interact with it.

6.1.3 Support for evaluating users' professional skills

The lack of constructive feedback during the job-hunting process is a general challenge for job seekers from this study. One potential feature of the chatbot is training a skills evaluation system to address candidates' professional skills and offer practical suggestions.

The author mentioned that the prerequisites for international students were the same as those for Finnish students, based on Viitasalo's research (Viitasalo, 2018). Those requirements include a suitable personality, good English language skills, motivation, programming skills, communication skills, and suitability for the job. Building on the above, below is the description of the evaluation chatbot from three perspectives. However, suitable personality and suitability for the job depend on the recruiter's perspective which cannot be addressed in this chatbot. Regardless of the perspective, users are required to provide a specific job description before using the evaluation feature of the chatbot.

First, **CV evaluation of the chatbot**, users can upload their CV through the typing area, and then the chatbot will match and compare their CV with the local requirements for a CV. After that, users will receive feedback from the chatbot, which includes general level attractions, limitations, and personality of their CV, even what their professional skills match with the specific job description. However, a CV typically contains personal information such as email address and phone number, which indicates that privacy and security should be a concern during the design and research process.

Second, a **dialogue-based evaluation chatbot** provides interview-type experiences. Based on a certain job description, the chatbot extracts descriptions of key skills from it and then compares them with the existing Knowledge Base to generate relevant interview questions. The chatbot will ask interview questions one by one, and the user through text or voice input to answer corresponding questions. After all questions are asked and answered, the chatbot provides evaluation feedback based on the interactive history. Typically, users can receive an evaluation of their professional skills that match the job position, their English and communication skills from dialogue evaluation, and their motivation for this job.

Last, an **assignment-driven evaluation chatbot** that helps users better understand their practical knowledge. After collecting the job advertisement from the user, the chatbot will check whether the hiring process for this position includes an assignment component, and if not the chatbot will suggest ending the conversation or using other features. If it is included, the chatbot can generate an assignment with the description, requirements, and recommended time for completion. At this point, the user needs to complete this assignment in the allotted time and submit it to the chatbot. Furthermore, the chatbot starts evaluating the result from the user and provides comments or feedback about the professional skills in practice, and the matching situation with the job position.

To summarize, the chatbot supports evaluating users' professional skills in order to provide users that the main inspection points, room for improvement, and areas worthy of encouragement. In this case, the main purpose is to assist international students in altering their minds from students to professionals, at the same time, the research scope is possible to extend to university students.

6.2 Limitations

The methodological choice to conduct evaluation workshops in a specific context has limitations on generalizability. First, the design of the evaluation workshop was a creative and non-traditional interview, which was a challenge for the author. The author considered the workshop process, technical details while using the chatbot prototype, and the manuscript for facilitating. However, it is impossible to plan all and avoid any exposure, there was something out of the plan which relied on the author's facilitator skills. For example, technical problems occurred in participants' devices, and the technical guidance will affect participants' time spent with the chatbot. Second, it is unavoidable that voluntary-based participation is probable to attract participants with a positive viewpoint on the topic, and the sample size of the evaluation study was small, hence, conducting more evaluation sessions is needed when continuing deeper research on this topic. Should the filter conditions improve to meet the target group, it would be advantageous to conduct more qualitatively focused follow-up research. Last, participants had various understandings about this chatbot which caused them to have different interactions with it. To be specific, participants were struggling with formatting questions and addressing their situations when they utilised the chatbot. This indicates that the chatbot has room for improvement, especially in engaging and assisting users to consider their situation and accurately format questions.

The author notices in the findings that participants' expectations and interactions mainly focus on free-form questions, while the initial conversation flow is also included in the prototype. Hence, there are not too many specific suggestions and expectations related to the conversation flow from them. Adding the outlook to gain treasured understandings to find out improvements at this topic more entirely for the design of GPT-based chatbot.

6.3 Future Work

The approach in this study was explorative and it provides numerous perspectives for the chatbot focus on job seekers. The author encourages other scholars to continue exploring this area from different viewpoints, for example by focusing on the skills evaluation chatbot which is probable from both the job seekers' viewpoint and recruiter's perspective and conducting the user study to find out their requirements. Significantly, as recruitment chatbots attract more attention, job seekers' observations would have more treasured research, ideally by directing on the user study for the GPT-based chatbot design.

7 Conclusions

In this thesis research, a GPT-based chatbot prototype for international students at Tampere was created, and the author conducted evaluation workshops to understand the challenges that the target users face, how they would interact with the chatbot, and what they expect from such a system.

This study has identified three main findings following the constructive research process. To begin, international students as job seekers face various practical and external challenges, practical challenges like the language barrier, graduate students being required to have work experience and the ability to demonstrate personal competence. External challenges came from companies' attitudes towards hiring international talents, the information gap between the company and job seekers, and intense competition between more candidates than available positions. Following that, external challenges cannot be addressed by the chatbot, but realistic methods to handle practical challenges are building recruitment skills and recognizing local job market situations to find potential opportunities. Finally, the findings show that participants preferred to interact with the chatbot by typing free-form questions rather than selecting options, even the initial conversation flow can provide insights to them. In addition, users need assistance to formulate questions from the two types of questions asked by users: well-defined questions and opinion-based questions. Well-defined questions will receive better responses compared to opinion-based questions, and the participants who asked the former type of questions had more optimistic feedback for this study.

Additionally, a workshop as an evaluation method is creative and valuable in this study, which exercises researchers' workshop planning, execution, and adaptability abilities. In future, to evaluate the workshop plan and improve the researcher's ability, following ten simple rules for measuring the impact of workshops by Sufi (Sufi et al., 2018) is crucial.

As a result of this study, it is possible to conduct future work on this topic: how GPT-based chatbot helps job seekers to develop required skills in the recruitment process. Overall, the results of this thesis can be applied when developing a chatbot for job seekers, particularly those in the early stages of the recruitment process.

References

- Abou Merhi, A. (2022). *Influencing factors for hiring international students: A case study of MSMEs in the Tampere region* [fi=AMK-opinnäytetyö|sv=YH-examensarbete|en=Bachelor's thesis]. <http://www.theseus.fi/handle/10024/751152>
- Adamopoulou, E., & Moussiades, L. (2020). An Overview of Chatbot Technology. In I. Maglogiannis, L. Iliadis, & E. Pimenidis (Eds.), *Artificial Intelligence Applications and Innovations* (pp. 373–383). Springer International Publishing. https://doi.org/10.1007/978-3-030-49186-4_31
- Ahmad, A. (2005). *Getting a Job in Finland. The Social Networks of Immigrants from the Indian Subcontinent in the Helsinki Metropolitan Labour Market*.
- Aishwarya Gupta & Dayananda Sagar College of Engineering. (2020). Introduction to AI Chatbots. *International Journal of Engineering Research And*, V9(07), IJERTV9IS070143. <https://doi.org/10.17577/IJERTV9IS070143>
- Alberts, I. L., Mercolli, L., Pyka, T., Prenosil, G., Shi, K., Rominger, A., & Afshar-Oromieh, A. (2023). Large language models (LLM) and ChatGPT: What will the impact on nuclear medicine be? *European Journal of Nuclear Medicine and Molecular Imaging*, 50(6), 1549–1552. <https://doi.org/10.1007/s00259-023-06172-w>
- Baker, L. (2006). Observation: A Complex Research Method. *Library Trends*, 55. <https://doi.org/10.1353/lib.2006.0045>
- Choi, Y., Monserrat, T.-J. K. P., Park, J., Shin, H., Lee, N., & Kim, J. (2021). Proto-Chat: Supporting the Conversation Design Process with Crowd Feedback. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW3), 225:1-225:27. <https://doi.org/10.1145/3432924>
- City of Tampere integration programme 2021–2025*. (2021).

Diederich, S., Brendel, A., Lichtenberg, S., & Kolbe, L. (2019, June 8). *Design for Fast Request Fulfillment or Natural Interaction? Insights from an Experiment with a Conversational Agent*.

Dillahunt, T. R., Lam, J., Lu, A., & Wheeler, E. (2018). Designing Future Employment Applications for Underserved Job Seekers: A Speed Dating Study. *Proceedings of the 2018 Designing Interactive Systems Conference*, 33–44.

<https://doi.org/10.1145/3196709.3196770>

Dumais, S., Jeffries, R., Russell, D. M., Tang, D., & Teevan, J. (2014). Understanding User Behavior Through Log Data and Analysis. In J. S. Olson & W. A. Kellogg (Eds.), *Ways of Knowing in HCI* (pp. 349–372). Springer.

https://doi.org/10.1007/978-1-4939-0378-8_14

Finland, S. A. W. D., Turku. (n.d.). *Finding Work in Finland & Securing a job*. Retrieved 21 November 2023, from https://www.expats-finland.com/employment/finding_work.html

Generative AI: What Is It, Tools, Models, Applications and Use Cases. (n.d.). Gartner. Retrieved 4 December 2023, from <https://www.gartner.com/en/topics/generative-ai>

Goyal, R., Chaudhary, N., & Singh, M. (2023). Machine Learning based Intelligent Career Counselling Chatbot (ICCC). *2023 International Conference on Computer Communication and Informatics (ICCCI)*, 1–8.

<https://doi.org/10.1109/ICCCI56745.2023.10128305>

Heitzmann, N., Opitz, A., Stadler, M., Sommerhoff, D., Fink, M. C., Obersteiner, A., Schmidmaier, R., Neuhaus, B. J., Ufer, S., Seidel, T., Fischer, M. R., & Fischer, F. (2021). Cross-Disciplinary Research on Learning and Instruction – Coming to Terms. *Frontiers in Psychology*, 12. <https://www.frontiersin.org/articles/10.3389/fpsyg.2021.562658>

- Hourrane, O., Ouchra, H., Eddaoui, A., Benlahmar, E. H., & Zahour, O. (2020). Towards a Chatbot for educational and vocational guidance in Morocco: Chatbot E-Orientation. *International Journal of Advanced Trends in Computer Science and Engineering*, 9, 2479–2487.
<https://doi.org/10.30534/ijatcse/2020/237922020>
- Jonathan Lazar, H. H., Jinjuan Heidi Feng. (2017). *Research Methods in Human-Computer Interaction*. Elsevier Science.
- Kim, I., & Kuljis, J. (2010). Applying Content Analysis to Web-based Content. *Journal of Computing and Information Technology*, 18(4), 369–375.
<https://doi.org/10.2498/cit.1001924>
- Lee, T., Jagannath, K., Aggarwal, N., Sridar, R., Wilde, S., Hill, T., & Chen, Y. (2019). *Intelligent Career Advisers in Your Pocket? A Need Assessment Study of Chatbots for Student Career Advising*.
- Lee, T., Zhu, T., Liu, S., Trac, L., Huang, Z., & Chen, Y. (2022). CASExplorer: A Conversational Academic and Career Advisor for College Students. *The Ninth International Symposium of Chinese CHI*, 112–116.
<https://doi.org/10.1145/3490355.3490368>
- McTear, M. (n.d.). *CONVERSATIONAL MODELLING FOR CHATBOTS: CURRENT APPROACHES AND FUTURE DIRECTIONS*.
- Muhammad, A. F., Susanto, D., Alimudin, A., Adila, F., Assidiqi, Moh. H., & Nabhan, S. (2020). Developing English Conversation Chatbot Using Dialogflow. *2020 International Electronics Symposium (IES)*, 468–475.
<https://doi.org/10.1109/IES50839.2020.9231659>
- Ørngreen, R., & Levinsen, K. (2017). *Workshops as a Research Methodology*. 15(1).
- Parab, A., Palkar, S., Maurya, S., & Balpande, S. (n.d.). *An Intelligent Career Counseling Bot*. 04(03).

- Plain, C. (2007). Build an affinity for KJ method. *Quality Progress*, 40(3), 88.
- Schrepp, M., Hinderks, A., & Thomaschewski, J. (2014). *Applying the User Experience Questionnaire (UEQ) in Different Evaluation Scenarios*. 383–392.
https://doi.org/10.1007/978-3-319-07668-3_37
- Silva, G. R. S., & Canedo, E. D. (2022). Towards User-Centric Guidelines for Chatbot Conversational Design. *International Journal of Human–Computer Interaction*, 1–23. <https://doi.org/10.1080/10447318.2022.2118244>
- Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104, 333–339.
<https://doi.org/10.1016/j.jbusres.2019.07.039>
- Sufi, S., Nenadic, A., Silva, R., Duckles, B., Simera, I., Beyer, J. A. de, Struthers, C., Nurmikko-Fuller, T., Bellis, L., Miah, W., Wilde, A., Emsley, I., Philippe, O., Balzano, M., Coelho, S., Ford, H., Jones, C., & Higgins, V. (2018). Ten simple rules for measuring the impact of workshops. *PLOS Computational Biology*, 14(8), e1006191. <https://doi.org/10.1371/journal.pcbi.1006191>
- Suresh, N., Mukabe, N., Hashiyana, V., Limbo, A., & Hauwanga, A. (2021). Career Counseling Chatbot on Facebook Messenger using AI. *Proceedings of the International Conference on Data Science, Machine Learning and Artificial Intelligence*, 65–73. <https://doi.org/10.1145/3484824.3484875>
- Taecharunroj, V. (2023). “What Can ChatGPT Do?” Analyzing Early Reactions to the Innovative AI Chatbot on Twitter. *Big Data and Cognitive Computing*, 7(1), Article 1. <https://doi.org/10.3390/bdcc7010035>
- Truong, H., Thi-Yen-Linh, T., Thi-Duyen-Ngoc, N., Bao-Tuyen, L., Ha-Phuong-Truc, N., & Su, L. (2020). Development of an AI Chatbot to Support Admissions and Career Guidance for Universities. *International Journal of Emerging Multidisciplinary Fluid Sciences*, 13–20. <https://doi.org/10.22662/IJEMR.2020.4.2.013>

- Viitasalo, A. (2018). *Enhancing Employment of International Degree Students in Tampere University of Applied Sciences* [fi=AMK-opinnäytetyö|sv=YH-examensarbete|en=Bachelor's thesis]. Tampereen ammattikorkeakoulu. <http://www.theseus.fi/handle/10024/145796>
- Wanberg, C. R., Ali, A. A., & Csillag, B. (2020). Job Seeking: The Process and Experience of Looking for a Job. *Annual Review of Organizational Psychology and Organizational Behavior*, 7(1), 315–337. <https://doi.org/10.1146/annurev-orgpsych-012119-044939>
- WEB BASED CAREER COUNSELLING SYSTEM WITH CHATBOT. (2021). *International Research Journal of Modernization in Engineering Technology and Science*.
- What are Large Language Models? | Definition from TechTarget.* (2023). <https://www.techtarget.com/whatis/definition/large-language-model-LLM>
- Yang, S., & Evans, C. (2019). Opportunities and Challenges in Using AI Chatbots in Higher Education. *Proceedings of the 2019 3rd International Conference on Education and E-Learning*, 79–83. <https://doi.org/10.1145/3371647.3371659>
- Zaidi, D., Raza, S., & Sharma, L. (2021). Artificial Intelligence Based Career Counseling Chatbot A System for counselling. *Annals of the Romanian Society for Cell Biology*, 25(6), Article 6.
- Zhu, C., & Han, L. (2021). *CV chatbot based on “STAR” method*.

Appendices

Appendix 1

Background Questionnaire

Date:

User ID:

1. What's your age group?
 - 18-24
 - 25-34
 - 35-44
 - 45+

2. What's your original country?

3. What are you studying at university now? (IT, HTI, etc.)

4. How familiar with chatbot/generative products? (First time to use, used every day, or other-please specify)

5. If you have any work opportunities in Finland? Or work in Finland?

Appendix 2

Interview questions

- 1) What impressed you most about the process of using this chatbot? / What aspects of the product work well for you? (For example, advantages, helpful, anything you like?)
- 2) What are the difficult parts when you use this chatbot? Or do you feel frustrated? Could you explain the reason?
- 3) What didn't you ask compared to your notes in individual work?
- 4) Do you have any improvement recommendations or expectations for this chatbot?
- 5) Would you recommend this chatbot to your friends? Why or why not?
- 6) Would you use this in the future? Why or why not?

Appendix 3

Guidance for participants – User 1

Hi, my name is Anfang, welcome to my evaluation session! Here is a guide for you! It is nice to see you here.

User ID: User 1

Date: 27.09.2023

Here are some steps you can take now.

1. Please select a seat and seat down.
2. There is a plug-in board here to connect to your computer, we will use the chatbot in your laptop later, so keeping enough power is very important.
3. You could enter your Teams meeting now if you are comfortable, if not, we can do it later.
4. Your user ID is “**User 1**” which you need to fill out in questionnaires, so that you don’t need to use your real name during the test.

The moderator will remind you when to start to use the chatbot system, below are the steps you need to do.

Chatbot use in person guidance:

1. Enter the Teams meeting you have received from me, turn off the camera and voice.
2. Here is personal chatbot link, you can type it to your incognito tab but don’t touch anything now, let’s keep it for later: <https://bit.ly/3RAYN3E>
3. Share the chatbot screen through Teams.
4. The moderator will check if she can see the shared screen and start screen recording.
5. Confirm that your left top corner has a red point which means the recording is on.
6. The moderator will leave this meeting and enter other participants’ meeting now.
7. You could start using the chatbot now!!!
8. When you finish in around 10 minutes, the moderator will remind you it’s time to end this part.
9. Leave the Teams meeting.

Thank you for participating in my thesis evaluation session! The movie ticket will be sent to your email soon! 😊