Resume

The paradox of physiotherapists having the same amount of time for all patients, yet complex cases requiring more time and leaving less for documentation and treatment compared to simpler cases, motivates our research. We developed a prototype virtual assistant (VA), incorporating both speech recognition and speech synthesis technologies, to assess the viability of using such a system in physiotherapy practice for gathering patients' medical histories (anamnesis) in contrast to traditional interviews conducted by physiotherapists. We sought to alleviate the physiotherapists' responsibility of documentation and enable quicker initiation of necessary tests and treatment in practice. The two research questions for the project are:

How do physiotherapists perceive patient responses to anamnesis questions obtained through interaction with a virtual assistant compared to those obtained through conventional physiotherapy sessions?

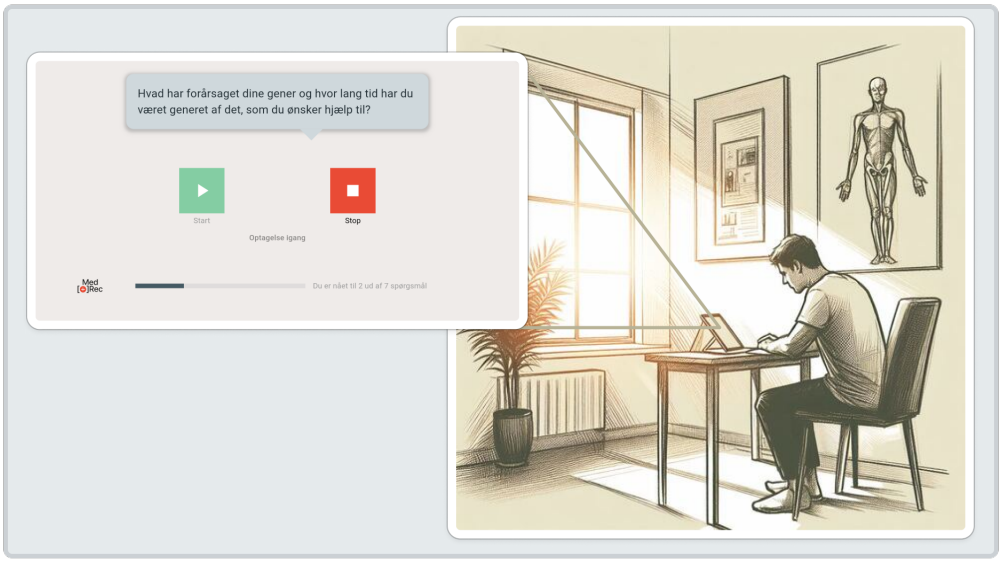
How do participants evaluate the communication when interacting with either a physiotherapist or the virtual assistant?

A literature review and an interview with a practicing physiotherapist were conducted in the preliminary stage of the project to understand the challenges in the field. During the initial phase, we also collected four anamnesis recordings from physiotherapists, utilized both as inspiration and evaluation. To gather potential patients' responses to the system, recruited participants were asked to interact with the prototype as well as provide feedback through questionnaires and semi-structured interviews, addressing the second research question. Subsequently, selected answers from the prototype test, along with responses from the collected conventional physiotherapy recordings, were presented to seven physiotherapists in a focus group interview. They were asked to distinguish between the two if possible.

Our findings reveal that physiotherapists could not distinguish between anamnesis data collected by the VA and that obtained through the conventional interviews. These results suggest that patients do not respond differently to questions related to their medical history when prompted by a system specifically designed for the purpose. From the patient perspective, the VA was well-received, particularly for its low-pressure environment that allowed patients to reflect before responding. Additionally, physiotherapists recognized its potential to streamline data collection. However, both groups noted some drawbacks that need to be addressed for the system to have a viable future. The system's ability to customize the questions based on a patient's specific situation and response was highlighted as vital for obtaining valuable data. Additionally, logistical challenges associated with possible implementation of the system was a concern among physiotherapists. Building on our achievements, it is crucial to address these challenges while acknowledging that additional obstacles may still need to be overcome to achieve successful implementation.

Self-anamnesis in Physiotherapy Practice

- Collecting patient data with a virtual assistant -



Master thesis

Digitalisation and Application Development
Cassiopeia - Department of Computer Science
Selma Lagerløfs Vej 300
DK-9220 Aalborg

AALBORG UNIVERSITY
STUDENT REPORT

Department of Computer Science
Aalborg University
http://www.aau.dk

AALBORG UNIVERSITY
STUDENT REPORT

Title:
Self-anamnesis in Physiotherapy Practice

Theme:
Master thesis

Project Period:
Spring Semester 2024

Project Group:
1

Participant(s):
Fie Louise Tengberg (20194329)
Alona Chaika (20221148)

Supervisor(s):
Niels van Berkel

Copies: 1

Page Numbers: 54

Date of Completion:
May 31, 2024

Abstract:

The paradox of complex patient cases requiring more time and leaving less for documentation and treatment compared to simpler cases in physiotherapy practice, motivates our research. The study aimed to investigate the use of a virtual assistant in physiotherapy practice to gather patients' medical histories (anamnesis). We sought to alleviate the physiotherapists' responsibility of documentation and enable quicker initiation of necessary tests and treatment in practice. For testing whether patients would respond differently to a system compared to a physiotherapist, a self-anamnesis prototype was developed, incorporating both speech recognition and speech synthesis technologies. The evaluation involved potential patients interacting with the system, followed by a survey and semi-structured interviews to gather their perspectives on the technology. A group of seven physiotherapists were then presented with a selection of answers from conventional anamnesis recordings and answers provided to the system. They were asked to distinguish between the two if possible. Overall, physiotherapists could not distinguish between anamnesis data collected by the VA and that obtained through conventional methods, supporting that patients do not respond differently when prompted by a system specifically designed for the purpose. Additionally, both physiotherapists and potential patients provided overall positive feedback on the system, while also highlighting some drawbacks. The primary concerns were the lack of customization and logistical challenges that could hinder successful implementation in practice.

Institut for Datalogi
Aalborg Universitet
http://www.aau.dk

AALBORG UNIVERSITET
STUDENTERRAPPORT

Titel:
Selv-anamnese i fysioterapeutisk praksis

Tema:
Speciale

Projektperiode:
Forårssemestret 2024

Projektgruppe:
1

Deltager(e):
Fie Louise Tengberg (20194329)
Alona Chaika (20221148)

Vejleder(e):
Niels van Berkel

Oplagstal: 1

Sidetal: 54

Afleveringsdato:
31. maj 2024

Abstract:

Paradokset at komplekse patienttilfælde kræver mere tid og efterlader mindre tid til dokumentation og behandling sammenlignet med enklere tilfælde i fysioterapeutisk praksis, motiverer vores forskning. Studiet havde til formål at undersøge anvendelsen af en virtuel assistent i til at indsamle patienternes sygehistorie (anamnese). Vi ønskede at lette fysioterapeuternes dokumentationsansvar og muliggøre hurtigere igangsættelse af nødvendige tests og behandling i praksis. For at teste, om patienter ville svare anderledes til et system sammenlignet med en fysioterapeut, blev der udviklet en selv-anamnese prototype, indeholdende både talegenkendelse og talesyntese teknologier. Evalueringen involverede deltagernes interaktion med systemet, efterfulgt af et spørgeskema og semi-struktureret interview for at indsamle deres perspektiver på teknologien. En gruppe på syv fysioterapeuter blev derefter præsenteret for et udvalg af svar fra konventionelle anamnesisoptagelser og svar givet til systemet. De blev bedt om at skelne mellem de to, hvis muligt. Generelt kunne fysioterapeuterne ikke skelne mellem anamnese data indsamlet af den virtuelle assistent og dem, der blev indsamlet gennem konventionelle metoder, hvilket understøtter, at patienter ikke svarer anderledes, når de bliver spurgt af et system, der er specifikt designet til formålet. Derudover gav både fysioterapeuter og potentielle brugere generelt positiv feedback på systemet, mens de også fremhævede nogle udfordringer. De primære bekymringer var manglen på personliggørelse af dialogen og logistiske udfordringer, der kunne hindre en vellykket implementering i praksis.

Contents

Preface

Aalborg University, May 31, 2024

Fie Louise Tengberg
<ftengb19@student.aau.dk>

Alona Chaika
<achaik22@student.aau.dk>

Chapter 1

Introduction

With our study, we aimed to investigate the feasibility of using a virtual assistant (VA) in physiotherapy practice to gather patients' medical histories without compromising the physiotherapists' ability to make meaning of the data for further treatment planning.

Striving to enhance patient outcomes, reduce costs, or streamline operations, VAs are one out of various types of artificial intelligence (AI) technology that have been tested and implemented within the healthcare sector. The outcomes from efforts made to digitally assist healthcare professionals bring awareness towards one challenge (Rowe, Nicholls, and Shaw 2022; O'Donovan et al. 2023; Davids, Lidströmer, and Ashrafian 2022). Amidst the integration of digital tools into healthcare practice, concerns arise about maintaining the human element of care, such as empathy and intuition. Healthcare providers are obligated to meet legal and ethical standards, and they argue that their human capabilities empower them to do so (O'Donovan et al. 2023).

Nevertheless, The use of AI has brought increased success across various healthcare professions, and physiotherapy is no exception, even though formal discussions have often overlooked this profession, according to Rowe, Nicholls, and Shaw 2022. Physiotherapist are an integral part of the healthcare system and the type of treatment that the profession offers is commonly used, in some cases as the most efficient or even sole curing ingredient (Davids, Lidströmer, and Ashrafian 2022).

In the contribution written by Davids, Lidströmer, and Ashrafian 2022, AI is believed to have the potential to provide a plethora of new supportive applications for physiotherapy. AI-based technology has also shown to be successfully used to support physiotherapists in their decisions regarding testing and diagnosing of patients, which most treatments begin with (Mirsky et al. 2020).

Receiving information on patients' medical history during the first appointment, in the following referred to as an *anamnesis*, is an essential part of healthcare practices involving diagnosis and therapy processes (Wienrich, Reitelbach, and Carolus 2021). A study from 2005 documented the accuracy of patients' anamnesis gathered through a

self-administered 92-item paper questionnaire in physiotherapy practice (Boissonnault and Badke 2005). The patients included in the study were attending a pre-operative visit at a physiotherapy sports clinic prior to a scheduled surgery. When comparing patients' responses to responses gathered by practitioners and existing medical records, results supported the accuracy of the self-administered reporting of anamnesis information (Boissonnault and Badke 2005).

Several digital alternatives for obtaining patient anamnesis have early on demonstrated enhanced accuracy and completeness of data compared to traditional paper-based questionnaires (Denecke et al. 2018; Dale and Hagen 2007). While electronic questionnaires enable patients to complete the questionnaire prior to their meeting with the physiotherapists, challenges identified with their use reveal potential issues with compliance. Patients may opt not to respond or only partially respond to the electronic questionnaire, affecting the reliability and completeness of the data gathered (Nayak and Narayan 2019).

Keeping in mind that parts of the aforementioned research into physiotherapy anamnesis' is carried out in 2005 and 2007 respectively, it is essential to recognize physiotherapy as an evolving practice, as it also applies to AI. Advancements in physiotherapy practices may influence the relevance and applicability of past findings, why we find it significant to reassess and retest the effectiveness of both conventional and digital methods for obtaining patient anamnesis.

While the initial study by Boissonnault and Badke 2005 suggests that self-administered anamnesis matches the quality of information obtained through physiotherapist interviews, the shift to digital platforms may introduce new factors that could impact our data. According to Rowe, Nicholls, and Shaw 2022, the public is seen to express concern and mistrust in relation to AI. Although, we aim to address the importance of the anamnesis with the implementation of an on-set solution, preventing patients from opting not to provide the necessary information, factors such as technical familiarity, privacy and security concerns remain critical to consider.

Our main research objective was to investigate if there were any impacts on the told patient anamnesis when interacting with a prototype of a virtual assistant, compared to the conventional conversation with a physiotherapist. Additionally, we also explored the evaluation of the communication in both scenarios from a patient perspective. This secondary objective complements our main research objective by providing insights into factors which, based on our research, could influence patient interactions with the system.

The concept of a self-anamnesis is by Denecke et al. 2018 described as *"... a procedure in which the patient answers questions about the personal medical history without direct interaction with a doctor or medical assistant"* (Denecke et al. 2018, p. 85). This overarching definition aligns closely with our product vision. However, it is essential to acknowledge that our

understanding of conventional practices is crucial for determining the primary purpose of the self-anamnesis prototype we are developing. Additionally, besides covering the existing landscape of supportive AI technology implemented within the healthcare sector, we will utilize our initial work to delve into the specified parts of physiotherapy practice, where various methods are employed for gathering anamnesis. Such methods could be either computerized, paper-based, or in-person (Wienrich, Reitelbach, and Carolus 2021; Denecke et al. 2018). Our preliminary investigations will serve as the foundation for this exploration, providing insights that will guide our main study in answering the following research questions:

How do physiotherapists perceive patient responses to anamnesis questions obtained through interaction with a virtual assistant compared to those obtained through conventional physiotherapy sessions?

We added a second research question for addressing the patient perspective:

How do participants evaluate the communication when interacting with either a physiotherapist or the virtual assistant?

The contributions provided by this work are: (1) advancements of AI technology into healthcare, (2) insights into benefits and considerations for implementing a digital self-anamnesis in physiotherapy practice and (3) informed future directions for taking the features of the self-anamnesis to the next level of effectiveness and usability.

The remainder of the paper is structured as follows: We start by covering related work within AI utilized in the healthcare sector and perceptions related hereto. Then, we present insights into physiotherapy practice and patient anamnesis' gathered through a pre-study interview, leading into the chosen main study methods for data collection and analysis. Following, we present the self-anamnesis prototype and describe our process of implementation before revealing the findings of our work. For the last chapters of this paper, the presented findings will be discussed in relation to related work along with a description of limitations and possibilities for future work.

Chapter 2

Related work

In the following, we firstly present research in the field of natural language systems, as a sub-field of AI, introduced in healthcare. In continuation hereof we outline relevant research on perceptions of implementation and use of AI to either enable, empower or partially replace healthcare professionals.

2.1 Integration of natural language systems in healthcare

The rapid and ongoing evolution of AI has led to its widespread use across various fields of work. Van Hartskamp et al. 2019 assert healthcare as *"the next domain to be revolutionized by artificial intelligence"*. Although AI has not yet reached its full potential within this sector, its adoption is progressing significantly (He et al. 2019). The progress includes AI systems with integration of Natural Language Processing (NLP), such as speech recognition (SR), conversational agents (CA) and conversational user interfaces (CUI). Although the aforementioned systems are related concepts, they represent different applications within the field of NLP. The commonality of the systems are thought to improve the quality of healthcare by improving efficiency and the satisfaction level of both patients and professionals (Latif et al. 2020; Milne-Ives et al. 2020; Davids, Lidströmer, and Ashrafian 2022). Primarily, it entails relieving the workload for healthcare professionals and elevating the quality of work through error reduction and increased precision (Aung, Wong, and Ting 2021; Latif et al. 2020). According to Davids, Lidströmer, and Ashrafian 2022, AI systems for physical rehabilitation can also be cost-effective alternatives to more resource-intensive methods and expensive equipment. Each type of system will be elaborated on separately in this chapter, commencing with the broader term of SR and progressively narrowing down to voice assistance and CUI, both viewed as subsets of CA.

Speech recognition in healthcare applications

Speech recognition, the process of transcribing speech to text (Alharbi et al. 2021), primarily serves as an alternative way for healthcare professionals to document medical information, thereby enhancing the quality, efficiency, and speed of such documentation (Latif et al. 2020). Despite being a relatively outdated study in the context of AI, Levin and Levin 2006 recognized assistive SR as an extensive and practical tool to facilitate efficient and convenient patient communication. Also, the use of SR was then perceived as particularly useful in patient data collection. Within the study, a dialogue system was implemented and tested for collecting and monitoring data about patients pain descriptions through phone sessions. The researchers reported the system's effective functionality, with an error rate of just two percent (Levin and Levin 2006). To emphasize the progress made throughout the last decade, it was argued in a study carried out by Durling and Lumsden 2008, that the use of SR on mobile devices in healthcare contexts was limited. Nevertheless the technology has significantly improved in the past years and many applications incorporating SR are being developed for smartphones and intelligent personal assistants (Park et al. 2018).

Conversational agents in healthcare

To address existing challenges of healthcare services, such as insufficient numbers of healthcare providers, CAs, which are described as *"systems that mimic human conversation using text or spoken language"* (Laranjo et al. 2018, p.1248), are being introduced to the health sector (Milne-Ives et al. 2020). Laranjo et al. 2018 and Milne-Ives et al. 2020 both acknowledge the emergence and utilization of CAs in healthcare. Laranjo et al. 2018 describes the use of CA in healthcare as an emerging field and Milne-Ives et al. 2020 emphasizes that improvements in CAs are driven by the considerable advancements in NLP. This progress has led to the integration of more advanced AI technologies across various parts of the healthcare sector, including health condition screening, at-home health management support, and professional training. The systematic review studies (Milne-Ives et al. 2020; Laranjo et al. 2018) have investigated the effectiveness of CAs in healthcare. While concluding that there is *"a moderate amount of evidence supporting the effectiveness, usability, and positive user perceptions of the agents"* (Milne-Ives et al. 2020, p.12), the researchers of both review studies also highlight key challenges, such as poor dialogue management and difficulties with language understanding.

Two examples of CAs are chatbots and voice assistants. Chatbots facilitate communication with the system through natural language via text interface, while voice assistants enable communication through a voice interface (Sezgin et al. 2020). Given the project's primary focus on the latter aspect, only VAs will be addressed in the following section.

Voice assistants in healthcare

The evolution of voice assistants technology (VAT) has been noteworthy since the introduction of Siri in 2011, with continuous improvements in technology (Sezgin et al. 2020). Just like the systems mentioned above, VATs have also made way into healthcare. However, development hereof is still in its early stages (Wienrich, Reitelbach, and Carolus 2021). VAT, described as *"a game changer for the future healthcare"* (Chen et al. 2021), is used in different areas. For instance, Wienrich, Reitelbach, and Carolus 2021 highlights promising contributions in the area of medical diagnosis and therapy, as well as promising features in the area of anamnesis. The nature of VATs to facilitate hands-free interaction with a system in a natural way, resembling human-human interaction, is believed to improve the effectiveness of delivering health information and communication (Sezgin et al. 2020). Given the human-like characteristics of VATs, such as emotions, humor, gender, politeness, memory and the ability to perform multi-turn conversations (Kim, Jung, and Lim 2022), there is a necessity to comprehend the psychology of interaction with these systems (Wienrich, Reitelbach, and Carolus 2021). One result of including user perspectives in the investigation of VAT used by older adults, highlights feelings of frustration during interaction. The insights to users experience gave rise to a proposal for a voice-first interface, supplemented by alternative input-output modalities such as touch screens (Chen et al. 2021). By putting voice interaction at the forefront of the user experience, and including touch-based interfaces instead of using voice commands, the type of technology would then be closely aligned with a CUI.

Conversational user interface in healthcare

CUI differs from SR and CAs by providing an interface through which users can engage in conversation with systems. Given that CUIs are predominantly deployed on mobile devices (Kocielnik et al. 2021), their utilization in healthcare contexts offers promising benefits. For example, Jaber and McMillan 2020 suggest that integrating CUIs in healthcare could lower care costs and improve therapy accessibility and convenience by facilitating patient self-management and monitoring. Furthermore, Kocielnik et al. 2021 highlights the potential for increased patient engagement and support for clinical workflows.

The paper "Talking to Ana: a mobile self-anamnesis application with a conversational user interface" (Denecke et al. 2018) introduced the concept of a self-anamnesis. Self-anamnesis involves patients responding to questions about their personal medical history without engaging directly with a doctor or medical assistant. The aim of the concept is to improve the collection of data at the beginning of the treatment process. The implementation of a mobile application with a CUI was tested in the context of music therapy, yielding positive results. However, the researchers left unanswered questions regarding patients' feelings during interactions with the system, particularly whether patients feel more open and comfortable.

AI continues to emerge in the form of promising tools in healthcare for facilitating patient engagement and streamlining clinical workflows, but equally important as the design is the deployment, which often gets neglected (Norman and Stappers 2015). Huang, Blaschke, and Lucas 2017, along with other researchers, refers to the issue by the name "pilotitis" and emphasizes the necessity for longer deployment periods and larger-scale implementations to observe long-term effects in complex systems (Norman and Stappers 2015; Blandford 2019).

The aforementioned studies lead the shift of attention towards understanding the factors influencing the successful adoption and utilization of AI in healthcare settings. The following focus on perceptions surrounding the integration of AI in healthcare settings will transition from researchers perspective into factors navigating the reception of AI systems among healthcare professionals and patients.

2.2 Perceptions of AI in healthcare

Care settings are by O'Donovan et al. 2023 described as complex, unstructured, dynamic and sometimes unpredictable. Research carried out by the aforementioned focused on identifying and validating human capabilities relevant to care settings. Workshops were held with the aim of informing design and development of robotic platforms for addressing frailty in healthcare, and participants included both stakeholders, healthcare professionals and clinical researchers. The professionals discussed tasks and practices valued in their professions, and as an outcome, it was acknowledged that healthcare and care professionals bring forth a diverse array of capabilities in their roles. The research puts emphasis towards the importance of understanding what capabilities that matter in a specific care setting when designing and implementing autonomous systems leveraging AI technologies (O'Donovan et al. 2023).

The acknowledgement of care professionals and the quality they bring to complex settings is also evident in older research carried out years before the rapid development of AI. For instance, Vollenbroek-Hutten et al. 2015 conducted a study where Information and Communication Technology (ICT) for rehabilitation services are designed and pilot tested across twelve healthcare institutions and four different diagnosis groups. The results hereof show that it is important to distinguish between patient groups and tailor to their unique needs and characteristics, as is required of care professionals who treat across various groups.

Vollenbroek-Hutten et al. 2015 evaluated the designed ICT-supported rehabilitation services with both patients and care professionals. While the services where accepted and highly used by the patients, the same parameters were described as 'dramatically low' among care professionals. Research done by Kosterink 2014 was based on a curiosity towards why only very few attempts to implement ICT services in healthcare get

pass the project phase. She advocated that the value of telemedicine services is highly dependent on both the clinical purpose of what it is used for but also how the technology intervenes with healthcare work (Kosterink 2014). Kosterink 2014 refers to the definition provided by the World Health Organization defining telemedicine services as 'healthcare services where ICT is used by healthcare professionals to exchange information for the treatment of a patient' (WHO 1998). In relation to Kosterink 2014 and her research, the professionals included in the more recent study carried out by O'Donovan et al. 2023, expressed that they for the future wish to continue to do what they signed up for. Statements as such accentuate the importance of considering how any technical solution might intervene with the work carried out by care professionals and whether there might be reasons for any resistance towards technology interventions. Blandford 2019 describes it as 'the design of technology has to be both socially and culturally appropriate besides addressing the unique concerns related to particular health regions'. By fulfilling these criteria, human computer interaction (HCI) is expected to have significant impact on health and well-being (Blandford 2019).

According to Rowe, Nicholls, and Shaw 2022, there is no way to avoid the evolving technology and the fundamental changes it will bring to the future of work, even though technology is unconscious and incapable of empathy. While questioning if empathy and compassion is what physiotherapy patients really are looking for, Rowe, Nicholls, and Shaw 2022 refer to the likeliness of AI-based systems outperforming professionals with highly reliable answers given at a very low cost.

Physiotherapy, as many other healthcare professions, places much emphasis upon person-to-person interaction where human values can be applied, and Harari 2019 believes that they can keep doing so for a long time to come. Even though routine tasks carried out in specialized domains will be automated, the human care industry is likely to remain due to the simultaneously use of a wide range of skills (Harari 2019). However, while there is no reason to fear for technology to entirely eliminate the physiotherapy industry, it is critical for physiotherapist to know of the issues related to implementations of technology in order for the profession to maintain its integrity (Rowe, Nicholls, and Shaw 2022; Harari 2019).

According to Rowe, Nicholls, and Shaw 2022 the public is seen to express concern and mistrust in relation to AI. While physiotherapists might be concerned about "how's" and "why's" for any health predictions made by AI technologies, patients are likely more concerned about their privacy and data security (Rowe, Nicholls, and Shaw 2022). Regardless of the case and person, suggestions are to implement strategies for education, as well as ongoing verification and validation (Rowe, Nicholls, and Shaw 2022; O'Donovan et al. 2023). According to Holzinger et al. 2017, it is also a question of addressing the challenges posed by black box models and enabling systems to generate the underlying explanatory structures leading to decisions. Blandford 2019 emphasizes,

that HCI will play a crucial role in the process of making future AI systems "explainable", but there is a need for new design- and deployment strategies costumed for the complexity of these systems.

To summarize, our research both draws for inspiration and differentiation from the existing literature presented in this chapter. The research carried out by Denecke et al. 2018, introducing the concept of a self-anamnesis in the context of music therapy, has much in common with the chosen setup for investigating the effect of a similar solution in physiotherapy practice. While the aforementioned study implements a CUI, enabling the patient to get answers to any clarifying questions, the appearance of our prototype resembles to a greater extent with a traditional questionnaire. In chapter 5, decisions concerning the prototype design will be elaborated on.

Based on our preliminary research findings, there does not appear to be previous attempts of introducing the idea of replacing the physically present physiotherapist in the initial patient interview with technology. While the ways in which AI proves to contribute in other areas of physiotherapy practice have been explained by Davids, Lidströmer, and Ashrafian 2022, the way in which a self-anamnesis can add value to the conventional physiotherapy treatment is not considered clarified to this point. Wienrich, Reitelbach, and Carolus 2021 recognise the effectiveness of voice assistance for anamnesis which aligns with our decision to utilize the technology for this specific purpose in a distinct context. Given that our prototype will be designed to enable both verbal and physical interaction with an interface, we will refer to it as a virtual assistant (VA), which is a CUI utilizing VAT.

Other studies provide points of awareness concerning deployed methods relevant for our methodological approach presented in the two following chapters: first a pre-study interview with a experienced physiotherapist followed by the methods deployed during tests.

Chapter 3

Pre-study

We conducted an interview with a clinical practicing physiotherapist to achieve a better understanding of what a anamnesis is and how the concept contributes to aspects of the physiotherapists work. The interview was carried out as semi-structured with guidance from predefined themes and questions (appendix A). For the purpose of this initial interview it was not found necessary to invest time and resources in recruiting multiple physiotherapist. The participating physiotherapist had up until the interview gained 9 years of clinical experience after taking her master thesis in Musculoskeletal Physiotherapy. Alongside her clinical practice, she worked for the Association for Musculoskeletal Physiotherapy as an alternate at the board. Together with existing literature, the physiotherapist was thought to have the necessary expertise to provide a knowledge-based starting point for generating design ideas as well as finding a clear and relevant analytical purpose.

Questions were asked towards the time typically allocated to gather patient anamnesis' and to whether the questions asked during this session are standardized or not. The physiotherapist highlighted that her work is based on what is pronounced as the biopsychosocial model, aiming to understand the factors that have the most influence on the patient's condition or injury. She mentioned that there is always a certain structure to her approach, guided by suggested questions, but emphasized that it is not a rigid conversation, allowing flexibility in questions as well as time spent:

> *"Typically, in our clinic, we allocate 45 minutes for an initial consultation, which includes documentation. So, I usually plan for about 40 minutes of actual interaction with the patient. However, this can vary from 15 minutes to the entire allotted time, depending on the complexity of the case."*

We were also curious to whether the physiotherapist had identified any aspects of the anamnesis process that she found less effective or problematic. The physiotherapist explained that a thorough anamnesis is crucial as it lays the groundwork for selecting the

subsequent examinations, and with some patients 40 minutes feels insufficient.

Throughout the interview we also gained insight into the importance of clarifying whether there might be any symptoms indicating serious pathology requiring further examination by other health professionals. Additionally, the physiotherapist described how, in her clinic, patients are sent a questionnaire prior to their first consultation. This questionnaire includes initial questions about the patient's condition, such as where they are experiencing pain, for how long, and the intensity of the pain. There is also a body diagram where patients can indicate the location of their pain. The questionnaire has the potential to provide valuable information for the initial assessment, if the patients provide comprehensive responses. However, according to the physiotherapist, it is not uncommon for patients to provide incomplete answers or even choose not to respond to the questionnaire at all.

Lastly, during the interview we sought the physiotherapist's opinion on the potential value of digital tools for conducting anamnesis and whether she knew of clinics utilizing such solutions. In her response, she acknowledged the potential time-saving benefits of pre-gathering medical history data using technology. While she saw value in streamlining the workflow, she emphasized the importance of personal interaction for building trust with patients, reasoning why the physiotherapist should not be entirely replaced.

The aforementioned insights have influenced our main study in several ways. Firstly, while some aspects of an anamnesis may require customization for complex cases, there is potential for standardization in certain areas, such as in identifying serious pathology. The interview has also brought awareness to the potential for technology to streamline certain aspects of the anamnesis process.

In our main study, we aim to explore the potential role of technology in standardized anamnesis processes. We are interested in whether users will respond differently to anamnesis questions when prompted by technology, challenging the notion that technology cannot replace physiotherapists. While customization of anamnesis processes is important, for the scope of our study, we will focus primarily on a standardized approach. However, we recognize the importance of addressing customization as a limitation in our study. Subsequent chapters will provide insight into the potential benefits of customization and avenues for future research in this area.

Chapter 4

Method

In the sections below, we describe the process of gathering different types of necessary data throughout the scope of our project, along with the recruitment of participants who have contributed with the collected data.

4.1 Data for analytical purpose

4.1.1 Scenario 1: Conventional physiotherapy

To investigate the perceptions of virtually-assisted self-anamnesis, we collected four anamnesis-recordings in conventional physiotherapy settings to establish a baseline for comparison and analysis. Two physiotherapists contributed to the data collection and recorded the anamnesis using mobile devices. The recordings served two purposes. Firstly, the questions asked during the recorded conventional scenarios were compared in the process of selecting questions for our prototype to ask. By comparing recordings from different physiotherapists and with different patients, we gained a broader perspective to what questions and formulations would reflect a general physiotherapy practice and related standards. The second purpose served by the conventional recordings will be described in a subsequent section.

Prior to recording, all participants signed a declaration of consent, thereby also committing to completing a questionnaire after the physiotherapy session. To mitigate potential bias, participants were instructed to seal their completed questionnaires in envelopes provided by the physiotherapist, ensuring confidentiality and anonymity, as they might have otherwise believed the physiotherapist would review their responses.

Eight questions were included in our questionnaire, all inspired by the Health Care Communication Questionnaire (HCCQ), an established questionnaire concerning one persons experience when encountering a health care professional (Gremigni, Sommaruga, and Peltenburg 2008). We excluded in total five questions from the HCCQ due to reasons

concerning their relevance in the two distinct scenarios investigated in our project. To apply the questionnaire in both the conventional and technical scenario, two questions pertaining to the presence of healthcare personnel were not applicable. The remaining three questions were omitted for reasons regarding their relevance to the anamnesis concept. Specifically, the questions focused on the patients response receipt and the physiotherapists problem resolution. These aspects do not align with the primary objective of the initial interview, wherein the physiotherapist is responsible for gathering answers and insights into the medical history of the patient. Lastly, modifications were made to the wording of the retained questions to ensure the suitability for use across both scenarios. This included modifying phrases such as 'The healthcare provider' to eliminate the use of third-person language in the final questionnaire. Figure 4.1 shows the final version of the questionnaire used for this projects purpose.

	Not at all	A little	Some what	Very much	Completely
I felt my needs were being respected					
I was asked questions in a clear manner					
I was asked questions in an aggressive manner					
I received clear and precise information					
I have been treated with kindness					
I have been treated in a rude and hasty manner					
The consultation was well managed					
My privacy was respected during the interaction					

Table 4.1: Final questionnaire translated from danish to English

To ensure that our assumptions about participants evaluation of the communication in each scenario were based on a sufficient amount of information, we collected extra responses to our questionnaire. Besides the four patients appearing in the received anamnesis recordings, another three physiotherapists were contacted and asked to hand out the questionnaire to any new patients of theirs, willing to contribute, after the initial interview.

4.1.2 Scenario 2: Self-anamnesis

While we had the option of utilizing our contacts in physiotherapy practice for recruiting participants to test our prototype, we decided to recruit participants outside of these en-

vironments due to risks of bias. Attempts of recruiting participants were made through social media, the university and our own network, based on two criteria. For one, any participant had to experience one or multiple types of physically pain or injury at the time of recruitment. The second criteria was that the pain or injury gave reason to seek physiotherapy treatment; however, participants were only selected if they had not yet received treatment. These two criteria were chosen to accommodate any bias related to differences in patients and participants starting point for providing the anamnesis information. Choosing to recruit our own participants was to avoid patients having to repeat their medical history twice, either risking omitting details due to boredom or inadvertently introducing new information or recalling additional details through repetition.

All of the activities related to the second scenario were conducted in a closed room, but at different locations in favor of those recruited. In some cases, tests were carried out in the participants' own homes, where it was more difficult for us to control the presence of any interfering elements. However, the self-anamnesis process in the intended usage scenario was simulate by allowing the user to choose a place in their home, where they felt physically private and separate from our presence. 10 participants were recruited to test the concept of self-anamnesis by interacting with the developed prototype on a tablet device. Prior to the test, all participants were briefed on the system's operation and usage. Furthermore, guidance on seeking technical assistance in case of any issues was provided, as we were not present in the room during the test. This approach aimed to closely simulate the self-anamnesis process in the intended usage scenario, preventing participants from interacting with or relying on any human during the interaction.

After testing the prototype for a self-anamnesis, the participants were asked to fill out the same questionnaire as in the first scenario, which was distributed to them beforehand. Furthermore, brief semi-structured interviews (appendix B) were conducted with each participant afterwards to gather qualitative data aimed at enhancing our understanding of the user experience with communicating and interacting with the self-anamnesis system.

The decision to exclusively conduct interviews following the tests of the self-anamnesis system, as opposed to the conventional physiotherapy setting, stemmed from several considerations. Integrating interviews within the conventional physiotherapy practice could introduce potential conflicts with established clinical protocols, potentially disrupting the therapeutic process. Coordinating interviews within the constraints of routine physiotherapy sessions necessitates proactive patient engagement, potentially burdening both patients and physiotherapists alike with additional administrative tasks. Besides demanding additional resources, reaching out to patients beforehand, might also risk biasing the patient-therapist interaction by foregrounding the research agenda over therapeutic objectives.

4.2 Focus group interview

As previously described, the recordings of conventional anamnesis sessions had an additional purpose besides supporting the selection of questions for our prototype. Parts of the anamnesis recordings from both scenarios were also compared to validate the participants answers. To do so, we carried out an online focus group interview with seven physiotherapists participating.

The recruited group comprised individuals with varying years of experience ranging from three to ten and a wide range of specializations, such as neurology, orthopedic surgery, geriatrics, cardiac rehabilitation, and sports injuries. The participating physiotherapists worked in various practice settings such as private clinics, hospitals, and municipal facilities, and brought perspectives from various regions and communities across the country. Our aim was to recruit for a rich diversity of perspectives and experiences within the field of physiotherapy, hoping it would lead to insightful discussions.

Patients' responses to certain anamnesis questions from the first scenario were compared with responses to corresponding questions from the second scenario. For each anamnesis question within the prototype, four to six answers were presented together, without indicating which answers originated from which scenario (appendix C). This approach aimed to facilitate discussions regarding differences in the presented data, enabling the identification of deviations and determining superiority (if applicable) along with the underlying reasons. Furthermore, we also facilitated a discussion on the overall idea and concept of the self-anamnesis with the participants of the focus group interview to assess their perspectives on its feasibility and practicality within clinical settings.

4.2.1 Thematic analysis

We proceeded to analyze the data gathered through our focus group interview using a thematic analysis approach. This is a method that entails the identification, analysis, and reporting of themes within data (Braun and Clarke 2006). Braun and Clarke 2006 explain the term 'theme' as *"something important about the data in relation to the research question"* (Braun and Clarke 2006, p. 82), which indicates a certain level of patterned response or meaning within the dataset. The researchers have developed a guide outlining a 6-step process (Doing reflexive TA), which we approached as follows. While listening to the recording of the focus group session, notes containing information relevant to the research question were taken, along with corresponding citations. Subsequently, this data was analyzed and organized into themes based on response patterns.

The data collected through the presented activities and methods provided insights, which will be elaborated on in chapter 6. Prior to this the developed prototype of a self-anamnesis system will be presented and described in the following chapter.

Chapter 5

Implementation

In this chapter we will present the chosen structure of the self-anamnesis prototype, providing insights into the reasoning behind the choices for its components and functionalities. The structure is depicted through three simplified models which highlights only the key components necessary for implementing the intended functionality and concepts of the prototype.

5.1 System structure and integration

Our project repository consisted of four main folders:

- A "lib" folder encompassing all UI screens and the file from which the execution of the app begins.
- An "assets" folder containing pre-generated audio files and text strings utilized during interaction.
- A "python" folder with the code responsible for establishing connections to both TTS and STT APIs.
- A "data" folder containing data from tests, both audio recordings of the participants' answers and the text data generated from these files.

The intended functionality of our VA is to offer an intuitive and easy-to-use interface, which enables patients of physiotherapists to provide comprehensive medical histories independently. When interacting with the system, standard anamnesis questions should be presented in a clear and structured manner for patients to respond to and the patient responses to anamnesis questions should be gathered accurately.

To achieve the described functionality, we developed three separate systems. One of these is responsible for establishing a connection to a text-to-speech(TTS) API (figure 5.1). To integrate the TTS functionality, we chose to rely on the TTS API in the

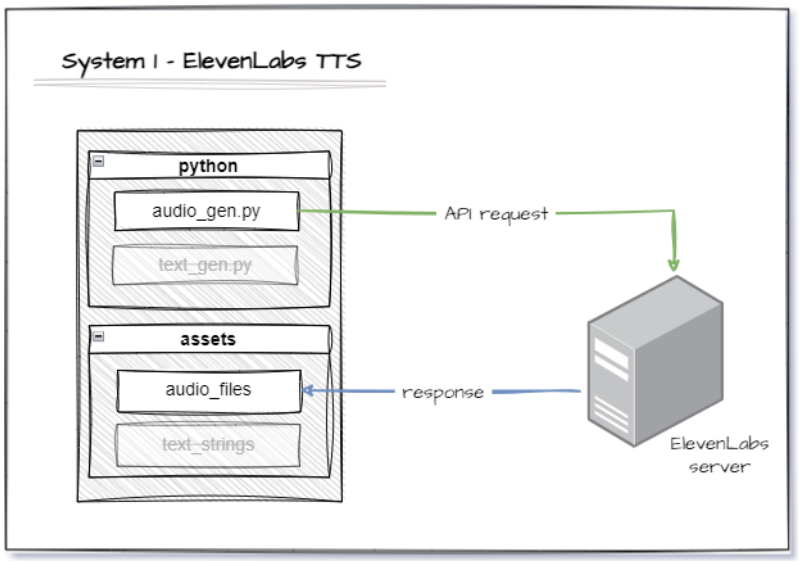


Figure 5.1: System 1, TTS

preliminary implementation phase, as the system's anamnesis questions are predefined. To implement TTS functionality, we established a connection to ElevenLabs API, which aligned well with the project's scope and requirements for seamless integration of a Danish language model. The company has earned widespread recognition for providing high-quality TTS software, enabling creation of audio content for various platforms and devices (Eden AI).

To provide insight into how the API requests are handled, we refer to 'Source code 5.1' as we explain the programming we implemented. The process involves sending a POST request to the ElevenLabs API endpoint, hosted by a server, and specifically designed for TTS conversion. The request variable contains the necessary information for handling the request - the specified URL and two variables containing the text to be converted, the desired language model, voice settings, API key required for authentication and more. The system utilizes a Python library to handle the HTTP request through the entire communication process with the API endpoint. Upon receiving the request, the ElevenLabs API processes the text and generates a response in return, as depicted in figure 5.1 - an audio file in the specified format.

Using a context manager, which ensures that files are properly closed after completion, the file 'output.mp3' gets opened in binary write mode ('wb'). The file will be created if it doesn't exist, or will overwrite an existing file. The response is then iterated over in chunks of the specified size (CHUNK_SIZE), using the iter_content method. After ensuring that the current chunk is not empty, the data gets written to the 'output.mp3' file. Each chunk of data received from the server is written to the

file sequentially, downloading the audio content chunk by chunk. We generated each anamnesis question separately located in the "audio_files" folder. These files were then utilized within the code to enable the VA to read the questions aloud.

```
28 response = requests.post(url, json=data, headers=headers)
29
30 with open('output.mp3', 'wb') as f:
1 31
32     for chunk in response.iter_content(chunk_size=CHUNK_SIZE):
33         if chunk:
34             f.write(chunk)
```

Source code 5.1: API TTS request

We primarily chose this approach because the questions remain consistent across sessions, obviating the need for repeated API requests during development and testing phases. Minimizing such requests not only conserves resources but also mitigates potential latency issues, thereby enhancing the user experience. However, it is important to acknowledge the limitations of this approach, which will be elaborated on in chapter 7.

The second system is the self-anamnesis system, to which users provide information related to their medical history (figure 5.2). We opted to develop for Android and carry out the tests on a physical device to best simulate the user experience as envisioned in physiotherapy practice. As a result of this decision, there were other matters to consider for the concept of the self-anamnesis to function during tests. For instance, how the participants' answers were to be recorded and stored, enabling us to access and analyze the data afterwards.

We chose to implement the necessary technology for the application to record each answer through the utilized tablet. Besides aligning with the intended functionality of the self-anamnesis, the decision to convert individual answers into separate recordings also made the collected data more manageable. The audio files were saved to the tablets external storage, as shown in figure 5.2, with a generic name determined by the number of the question the user is currently answering, to avoid one recording overwriting another. This allowed us to access each audio file after a test was completed and manually transfer them to our project repository. If the user records an answer but chooses to re-record a new answer to the same question, the new recording overwrites the previous one.

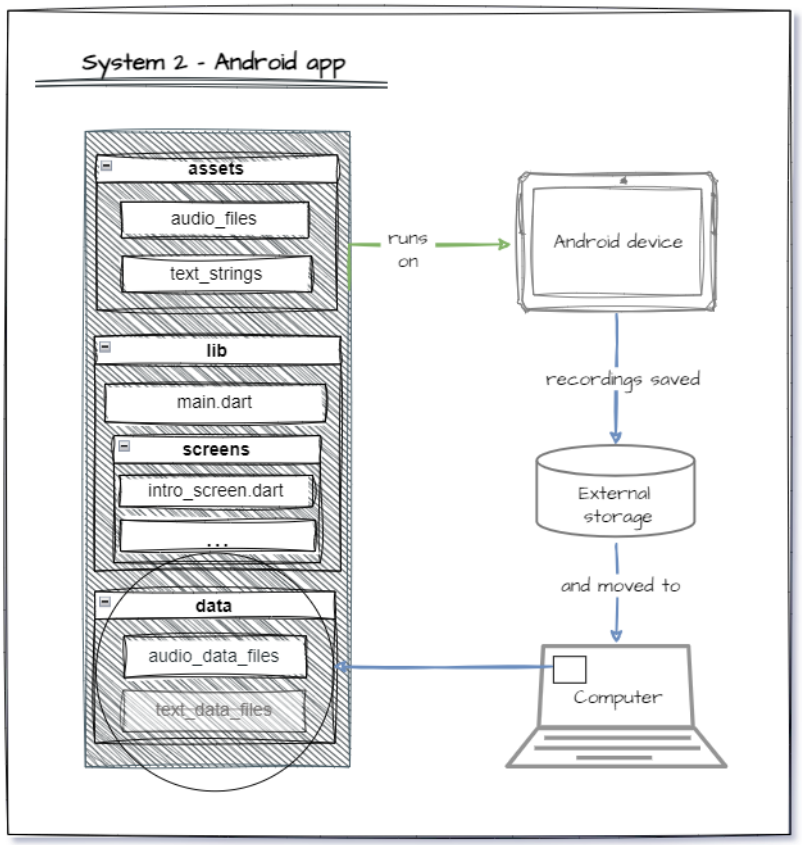


Figure 5.2: System 2, Android app

To fulfill the VAs purpose of assisting physiotherapists in the process of gathering the anamnesis information, crucial for determining the next treatment steps, the recorded answers need to be converted into text. Other related and necessary considerations involve protecting the data and making it immediately accessible for the physiotherapist to read prior to meeting the patient. This functionality is not within the scope of the project, hence the prototype does not include it. However, we have utilized STT technology within the third system we developed, to avoid manually transcribing the participants answers to questions during tests. As depicted in figure 5.3, the recorded audio files, saved in a separate folder in the 'data' folder of our repository, where utilized in the process of sending an API request with the code implemented in the 'text_gen.py' file.

To integrate STT functionality for processing each audio file from our test data, we opted to establish connection to the Whisper large-v3 API, developed by OpenAI (OpenAI - Whisper). To generate text files in Python, we utilised Hugging Face Transformers library, which offers an interface for working with NLP, including STT models.

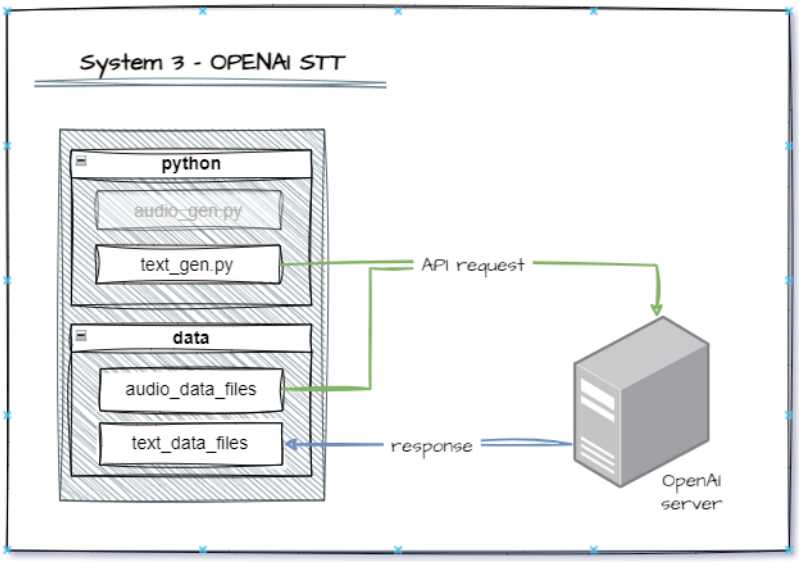


Figure 5.3: System 3, STT

For converting participants' audio recorded answers through the chosen API, there are two key components - a model and a processor. The model is configured with a specified ID, chosen among a number of available pre-trained models from OpenAI, along with a few other specifications. The processor ensures that the input is properly formatted for the chosen model and is initialized with tokenization and feature extraction, which refers to the initial steps taken to prepare the input data for processing. The two key components are then utilized within what is called a 'pipeline' depicted in Source code 5.2. The initialized pipeline, named 'pipe', serves as an interface for interacting with the chosen model and performing the requested automatic speech recognition (ASR) task. Besides including the model and processor specifications, the pipeline is also configured specifically for automatic speech recognition with additional parameters (line 27-33).

```
22 pipe = pipeline(
23     "automatic-speech-recognition",
24     model=model,
25     tokenizer=processor.tokenizer,
26     feature_extractor=processor.feature_extractor,
27     max_new_tokens=128,
28     chunk_length_s=30,
29     batch_size=16,
30     return_timestamps=True,
31
32     torch_dtype=datatype,
33     device=device,
34 )
```

Source code 5.2: Pipeline initialization

The initialized pipeline is utilized within the code after opening the requested audio file for conversion (Source code 5.3). Using the pipeline, ASR is performed on the audio sample (line 41), extracting the transcribed text from the result, and then writes the transcribed text to a text file named "transcribed_text".

```
37 audio_file_path = os.path.join("..", "assets", "audio_files", "
       name_of_file")
38 with open(audio_file_path, "rb") as audio_file:
39     sample = audio_file.read()
40
41 result = pipe(sample)
42 transcribed_text = result["text"]
43
44 with open("transcribed_text.txt", "w", encoding="utf-8") as file:
45     file.write(transcribed_text)
```

Source code 5.3: Handling STT request

Our choice of OpenAI model was primarily influenced by its support for the target language. At the time of implementation, the word error rate (WER), an accepted metric for assessing the accuracy of SR, was measured at 12% for the Danish language (Ali and Renals 2018). For the calculated WER it is applied that the lower the value, the better the precision of speech recognition. To understand the impact of 12% WER and if it poten-

tially could affect our data, we tested the technology prior to the implementation using a playground on Hugging Face's platform (Hugging Face), feeding the model with some of our TTS generated audio files. The provided audio files were successfully converted into text and saved in a separate folder in the "data" folder of our project repository (figure 5.3). In cases of wrong spelling, we chose to adjust the test data, ensuring that the recruited physiotherapists would not be able to recognize these answers as STT generated.

In the following we will showcase the final outcome of our prototype development, illustrating the design and interaction flow as presented to the participants recruited for testing.

5.2 System interface

The developed prototype is our take on a modern healthcare system within physiotherapy practice. The self-anamnesis is designed for streamlining and effectuating the gathering of crucial patient information. When arriving at the clinic upon the first physiotherapy session, the patient is thought to be directed to a private area or room, where (s)he can comfortably sit and complete the self-anamnesis. A receptionist or other personnel at the clinic briefly explains the purpose and how the system functions before leaving the patient.

The full navigational journey is depicted in figure 5.4, starting with the login screen in the upper left corner, marked with number 1. Before proceeding to the anamnesis questions, the patient has to login. The prototype prompts the user for its name, but in practice the system should require a social security number, ensuring that the collected patient data is documented in the correct patient journal. When logged in, the user is, on the screen marked with number 2, introduced to the VA, reading aloud a short introduction and emphasizing the importance of the information the user will be asked to provide.

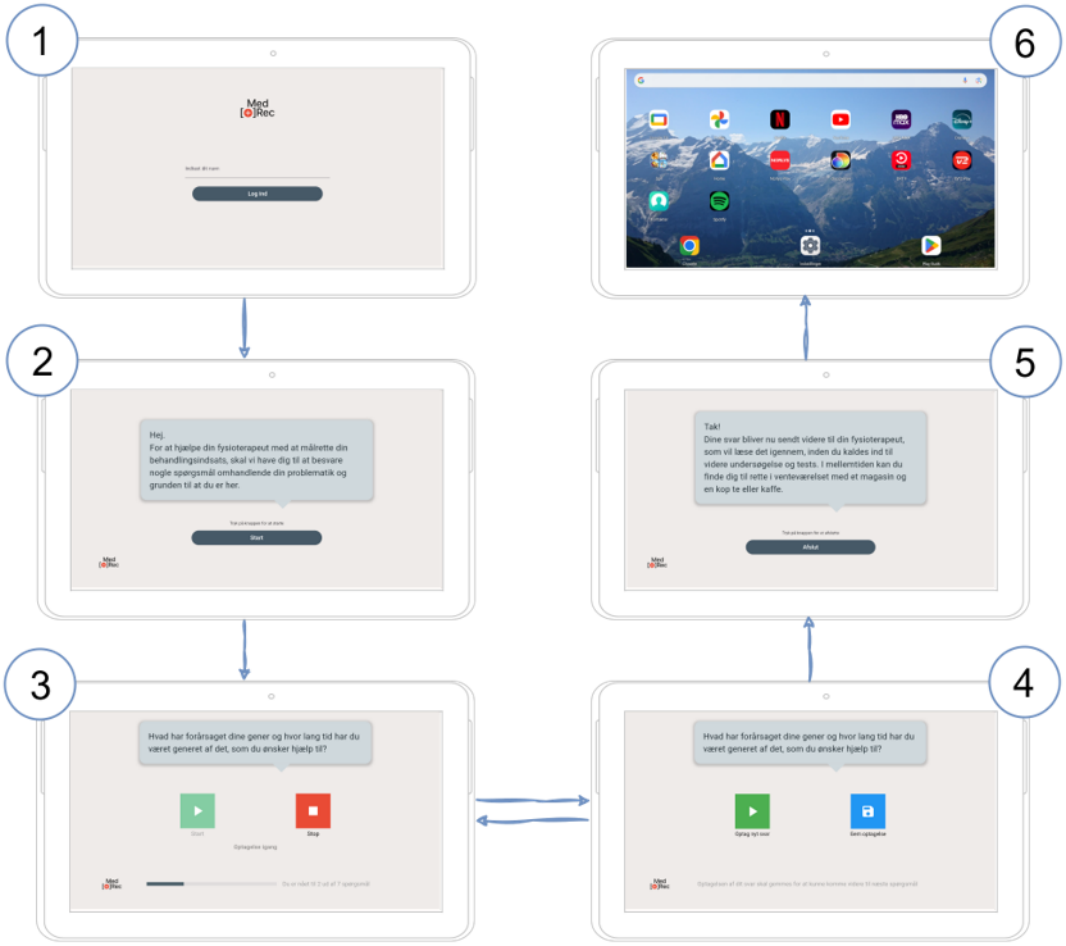


Figure 5.4: Navigation diagram

Once the user presses start on the second screen, a anamnesis question, as depicted on the screen marked with number 3, gets read aloud by the VA and displayed in text format on the screen as well. There is no option for requesting the VA to repeat a question, but the question remains displayed on the screen continuously until proceeding to the next question, helping the user stay focused on what to answer. Below the question, two buttons are displayed: "Start" and "Stop" a recording, but only the start button is active when the screen loads. The state switches when the user starts a recording. While recording, animated text with a fade transition displays just below the buttons, reassuring the user that the recorder is active. The last element on the recording screen is a progress bar which dynamically gets updated with feedback to the user on how many questions have been answered.

Figure 5.5 is a snippet of the navigational journey, indicating with two opposite arrows an unbreakable iterative cycle between the two screens. Once a recording is stopped,

the system loads the screen to the right, displaying the options of either recording a new answer to the current question or saving the recording, which was just stopped. With the option of undoing a recording, the total number of iterations will differ from one patient to another.

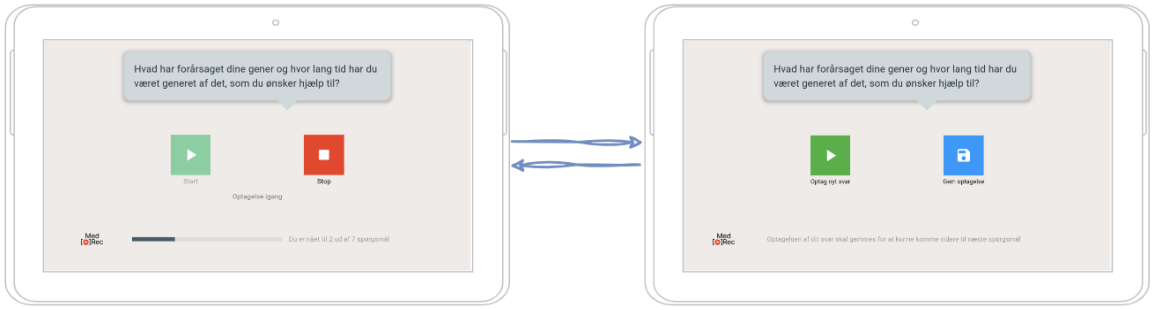


Figure 5.5: Recording an answer

Once the user has recorded and saved answers to all questions, the system navigates to the last screen, marked with number 5 on figure 5.4. Here the VA directs a patient to sit in the waiting area of the clinic, while the physiotherapist uses the provided information to plan steps for tests and treatment. If the user presses "finish", the application closes itself, returning to the home screen of the device.

During tests of the prototype, users went through the journey as described above, but the full context of being in a physiotherapy clinic, with waiting area and receptionist were too circumstantially to simulate. However, while carrying out tests at different locations, we endeavored to establish a private setting with an enclosed room. In the following chapter we will present and explain the results of our research.

Chapter 6

Findings

Besides providing material for our focus group interview, the outcome of the prototype test also provided answers to our questionnaire and additional insights from individual interviews, encompassing the patient perspective. The following chapter is divided in three sections, first revealing the insights gained into patients' perspectives, leading into the main study findings focusing on how the recruited group of physiotherapists perceived the test answers to the chosen anamnesis questions. To ensure clarity, quotes will be referenced as follows: "PA" will denote a participant from the prototype test interview, and "PH" stands for a physiotherapist from the focus group interview.

6.1 Findings of questionnaire study

Due to the nature of our study, wherein participants were exclusively exposed to one of the two scenarios, it was not feasible to draw direct comparison between participants evaluation of the communication in both scenarios. Therefore, when presenting the data, we opt to analyze each scenario independently and focus on identifying patterns and discrepancies within each scenarios responses. In the following analysis, a clear distinction will be made between responses from '**patients**' in the conventional scenario and those from '**participants**' testing our prototype.

A total of 12 patients answered the questionnaire handed out by physiotherapists in the conventional physiotherapy scenario, and 10 participants where recruited to test the prototype. The group of patients who contributed to the study ranged in age from 22 to 84 years, with a mean of 53.5, while the participants ranged from 25 to 85, with a mean of 48.7. Among the patients were four men and eight women, and for the participants, five men and five women.

Among the 12 patients, there was a consensus when answering to whether they felt that questions were asked in an aggressive manner or they had been treated in a rude and hasty manner (figure 6.1). Only one patient chose the option 'A little' for feeling

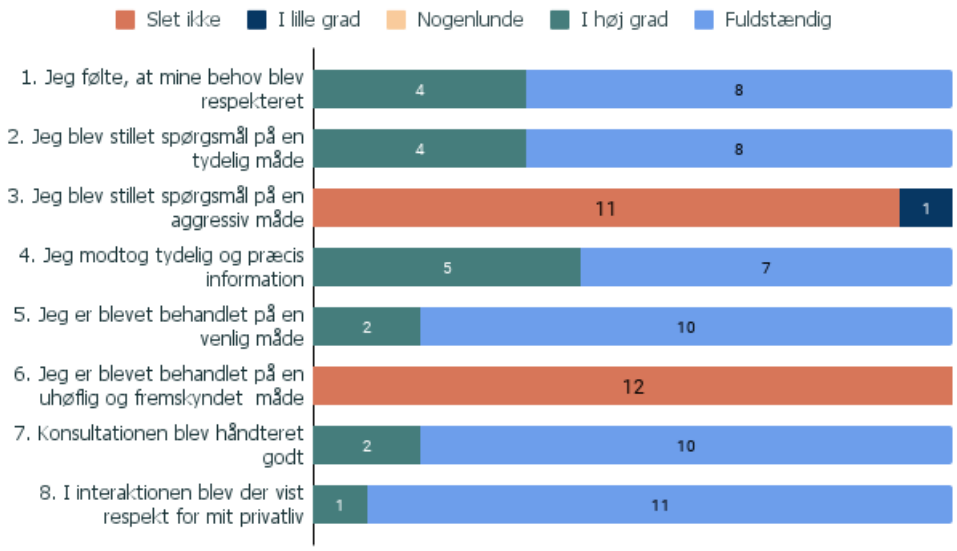


Figure 6.1: Chart for patients answers to questionnaire

that questions were asked in an aggressive manner, while all other answers to these two questions where 'Not at all'.

For the remaining questions, answers varied between 'Completely' and 'Very much', with highest consensus among the patients when answering whether they felt that their privacy was respected during the interaction. 11 out of the 12 patients, felt that their privacy was completely respected. The majority of patients, specifically 10 out of 12, also expressed agreement that they felt completely treated with kindness. Similarly, an equal number felt that the consultation was well managed. When asked if their needs were respected and if questions were clearly asked, eight answered 'Completely' leaving four responding 'Very much'. An almost similar distribution of answers was seen for the question of receiving clear and precise information, with one response changing from 'Completely' to 'Very much'.

Among the 10 participants recruited to test the prototype, there also appeared to be somewhat consensus on how they evaluated the communication when interacting with the VA (figure 6.2). In general, the participants agreed that they did not feel treated in a rude and hasty manner by the VA. When answering whether they felt that they were asked questions in an aggressive manner, two participants opted to answer 'A little', while the remaining answered 'Not at all'.

There was not as broad a consensus when answering the other six questions as there appeared to be between the group of patients in the conventional physiotherapy scenario. The options for answering varied between 'Somewhat', 'Very much' and 'Completely'.

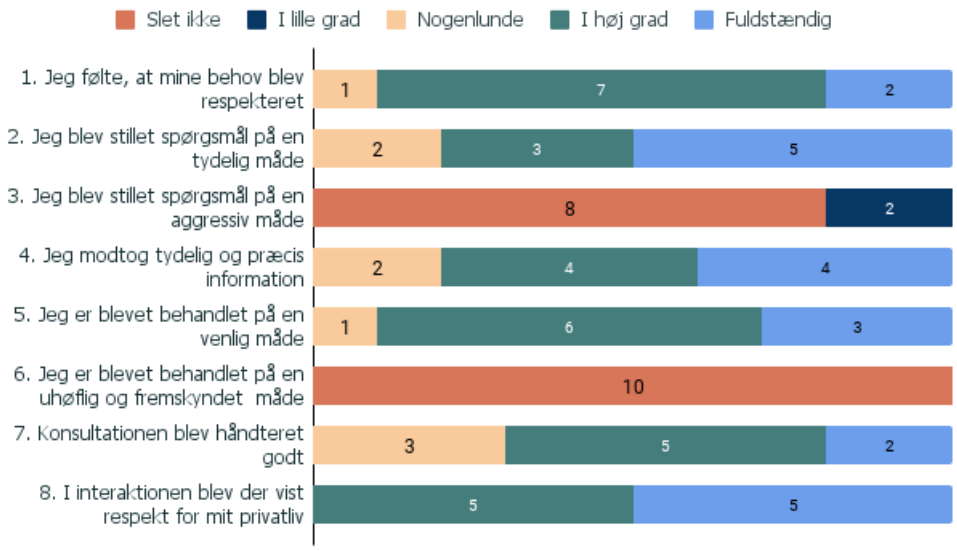


Figure 6.2: Chart for participants answers to questionnaire

Highest consensus appeared among the group when answering whether they felt that their needs were being respected with seven participants answering 'Very much' and two answering 'Completely' and a single 'Somewhat'. When asked if they felt that their privacy was respected during the interaction, five responded 'Very much' while the other half responded 'Completely'.

Half of the participants expressed agreement on that questions were completely asked in a clear manner. The last responses were distributed as follows: three participants answered 'Very much' and two 'Somewhat'. An almost similar distribution of answers was seen for the question of feeling that they received clear and precise information, with one response changing from 'Completely' to 'Very much'. For the last two questions, which asked if participants felt they were treated with kindness and whether the consultation was well managed, most chose 'Very much': six participants for kindness and five for management. Fewer participants felt they were 'Somewhat' treated with kindness compared to those who thought the consultation was 'Somewhat' well managed. Conversely, more participants felt they were 'Completely' treated with kindness compared to those who felt the consultation was 'Completely' well managed.

To summarize, there was a consensus among the patients that they weren't asked questions aggressively or treated rudely. Most felt their privacy was respected, and they were treated kindly during the interaction. Among the prototype testers, there wasn't as strong a consensus, but generally, they didn't feel rudely treated. The highest agreement was on whether they were treated kindly and their privacy and needs were respected, while opinions on consultation management and the clarity of both information and

questions were more varied. Overall, both groups generally had positive experiences, but opinions were more mixed among those who tested our prototype.

6.2 Findings of post-test-interview

For this part of our research, we were mainly focused on two aspects - the participants' perception of answering questions aloud when there is no person physically present to speak to and whether the UI was intuitive or not. We refer to appendix B to see the full interview guide including both interview questions and general information provided to participants before beginning the test. For analysing the collected data we utilized a thematic analysis method as for the data collected through our focus group interview (section 4.2.1). Below we have listed some general insights organized in themes based on response patterns of the participants' feedback, evolving around our second research question.

- **Preference for interaction style**: Participants showed a range of responses regarding their interaction preference. While some found it comfortable and efficient to interact with the system without a human presence, others expressed a familiarity with and appreciation for the human element in consultations. One participant said: "*Perhaps it's a need to have one's concerns acknowledged as genuine problems. Not meaning that someone needs to say 'I feel sorry for you'*"(PA04). Another participant also values human presence but emphasizes, "*Well, you do like to see who the person is, but honestly, I can live without it in this case*"(PA09). Similarly, one remarked, "*I actually don't mind it. Um, also because [unclear speech] you could definitely hear that it was a real person maybe, I don't know if it was, but it sounded like a real person, and that made the experience feel safer*"(PA01).

  The observations underscore the importance of authenticity in communication, even in digital interactions. However, participants widely accepted the use of technology as an initial interaction because they were aware of the eventual human communication in subsequent stages. One participant said, "*I actually thought it was okay for an initial conversation*"(PA07), suggesting a pragmatic approach to integrating technology into the consultation process. Lastly the oldest participant mentioned, "*I have experienced before at a doctor's office that the person is more focused on typing on the computer than making eye contact with me*"(PA08). The statement was reflection upon the impact a self-administered anamnesis might have on the first consultation with a physiotherapist, potentially reducing their need to type during the consultation.

- **Favoring of recording answers**: Participants generally viewed the system favorably compared to traditional questionnaires. They highlighted advantages such as ease of use, clarity of questions, and the ability to express nuances in responses

through recording. One participant said: *"These are my words, my own words. In questionnaires, it is often yes/no, and there can be a bit of putting words in your mouth, but if you ask these open-ended questions, it becomes a bit more nuanced"*(PA05). Another participant said: *"In terms of writing a text versus speaking, I would prefer to speak, but it is always easier to just tick some boxes, though I imagine that is not particularly useful in a situation like this, where you need to explain things"*(PA03). This comment underscores the limitations of checkbox-style responses in capturing the complexity of medical histories, highlighting the importance of allowing participants to articulate their thoughts freely with a self-administered anamnesis. Additionally, one participant remarked, *"I actually think that is better. All things considered, spoken language is easier, I mean you get the opportunity to be more impulsive than if you have to write your answers"*(PA10). The last observation aligns with the sentiment expressed by other participants regarding the ease and spontaneity of verbal responses compared to written ones.

Several participants appreciated the opportunity for reflection that the system provided. They noted that having time to contemplate their responses allowed for more thoughtful answers compared to immediate face-to-face interactions. One participant explained: *"I was sitting there reading the questions, and then probably took a minute before answering, just reflecting. If you're sitting face-to-face, you tend to give a more immediate response"*(PA05). The opportunity for reflection was also appreciated by another participant, who said: *"I thought it was fine because there was time to reflect, take in the question, and then respond"*(PA07).

- **Desire for clarification and additional information**: Some participants expressed a need for clarification on certain questions or desired the option to add more information beyond the provided prompts. One said: *"What challenged the interaction was that if there was something I was unsure about, I did not have the opportunity to ask 'what do you mean by that?' or 'what are you referring to?'"*(PA07). This sentiment aligns with another participant's observation: *"There was a question, how would you describe your activity level, or something like that, and I thought, what do they mean, is it in relation to my peers or what should I base it on?"*(PA10).

  Even though some participants did not have difficulty understanding the anamnesis questions, they still preferred having the option to ask questions to the system. For instance, one participant said: *"I had no trouble understanding the questions, but if someone did, I think it would be good"*(PA03). Moreover, another participant expressed: *"Now I can't recall, but was there an opportunity for providing additional information at the end, other comments or something? I could have used that. When you finish the questions, there might be something else you would like to say, but have not had the chance to"*(PA01). Offering a feature for seeking clarification or providing supplementary information could enhance the accuracy patients' responses.

Overall, participants appreciated the system's simplicity, clarity of questions, and the

opportunity for reflection. Some suggested improvements such as the option to add additional information, seek clarification, or edit previous responses. The absence of a human face was generally acceptable, and the interactive nature of the system was well-received.

6.3 Findings of focus group interview

Parts of the acquired data, gathered from the tests, constituted the material presented to the recruited group of physiotherapists. While the outcome of the prototype test also provided answers to our questionnaire and additional insights from individual interviews, the primary results of the study were derived from the qualitative analyses and discussions during the focus group session. The insights from the focus group interview will be presented in this section organized according to the identified themes.

6.3.1 Length and quality of responses

To begin with, the physiotherapists generally agreed on the expectations regarding responses given to a VA. It was assumed that longer responses were more likely to be given in conventional anamnesis, while shorter, more concise responses might be given when interacting with a VA. The prevailing descriptors for responses in the technical scenario were *short, precise* and *well-considered*. In explaining why a response was perceived as directed at a VA one participant explained: *"it is very short and precise. It sounds like a kind of monologue where you don't have the opportunity to have a dialogue by asking follow-up questions*"(PH02). This sentiment resonated with another participant's observation, "*the long [answers] with several variations are where they [participants] have spoken to a physiotherapist and are able to elaborate on what they have said*"(PH05). However, a contradictory viewpoint emerged later during the interview when another participant, while acknowledging the response to VA as "*being concise*" expressed uncertainty, suggesting, "*if I were to explain something to a computer, I would think that I should make sure to include all the details because I do not know what the next question is and I can't return back if I feel that I didn't explain it fully*"(PH06). The physiotherapists did not correlate the length of responses with their quality. As one participant stated, "*there are no good and bad answers, i.e. you cannot predict whether it was a person or a machine*"(PH07), which others agreed on. While concise answers were deemed efficient and well-considered by participants, the value in more detailed responses was also acknowledged for providing deeper insights for health professionals into the patients condition.

Based on the focus groups feedback, there was no consensus saying the length of a response correlates to one specific scenario. One participant emphasized this point claiming that patients communication styles differentiate. This observation reflects the inherent diversity in individuals' communication styles during interactions, a phenomenon

also observed in responses to the VA. Indeed, some participants tended to provide longer and more detailed answers compared to others, mirroring the diversity seen in physiotherapy practice.

6.3.2 Need for clarification and challenges in standardization

As there appears to be acknowledgement of different communication styles among patients, the group of physiotherapists expressed concern of how that difference would affect the collected anamnesis data. Therefore, the need for clarifying some questions was discussed during the focus group interview. One participant emphasized: "*You also need to prompt them [patients] and influence them a little to give us the right information*"(PH02). In relation to asking for more detailed responses, especially when patients offer ambiguous answers or seem to be 'talking off track', as one participant described it, the topic of formulating the good questions was discussed. It was agreed upon that formulating a clear question is crucial for extracting useful answers from patients. As a participant explained, "*It requires that you ask a very precise question, then I think you will also get a very precise answer*"(PH02). However, participants found it challenging to formulate questions that fit to all patients' needs. One participant speculated: "*What are the right questions for that patient? It is not always the same question that gets the patient to understand what you ask them. So you have to reformulate them. I think it's really difficult to find the right formulated questions that can suit everyone*"(PH05). The participants unanimously agreed on this point, deeming it advantageous for the VA to have the ability to personalize, which will be referred to as 'customization' in subsequent sections of the report.

6.3.3 Role of active listening

During the interview, the significance of active listening in human-human interaction was highlighted. Active listening was described as something that "*makes the patients tell more, elaborate a little more*"(PH04) and as something that "*causes some people to redefine or prompt them to think differently*"(PH02). Moreover, participants emphasized the importance of confirmation during communication. When attempting to identify the answers given during conventional scenarios one participant noticed: "*you can feel when you read it that the patient really needs some confirmation. When the patient says "I think it is about getting the core and legs strengthened", the patient is likely trying to talk to the therapist and get the therapist to nod appreciatively*"(PH01). Another participant added: "*You will, even if it is not on purpose, influence them [patients] to think of some other examples with the response you give when you are actively listening. They can sense that you want to hear more of one or the other*"(PH02).

Despite the focus group's acknowledgement of active listening as an important element of the conventional scenario, the physiotherapists recognized, that the lack of this

feature in our VA may not necessarily be a drawback, assuming that the absence of active listening might cause patients to provide more precise and thorough responses. As one expressed, "*There is something in the human interaction, a lot of non-verbal communication that might distract people from actually answering what you are being asked about*"(PH02) and later adding, "*What will be given to a virtual assistant, if I had to guess, would be very precise and well-considered*"(PH02).

6.3.4 Expectations and patient engagement

One other theme which appeared in the dialogue between the physiotherapists was centered around an assumption about why the answers provided by those who tested the prototype were hard to distinguish from a conventional scenario. The belief was that patients' expectations of the interaction, whether with a physiotherapist or a VA, may influence the depth and clarity of their responses. Some of the participating physiotherapists assumed that providing the participants with clear expectations regarding follow-up on their VA responses could improve the quality of their input.

The assumption above aligns with some of the feedback given by those who tested the prototype. As one explained, "*I think it was okay because I know I have to go in and talk to someone afterwards. If I wasn't to have a consultation at all, then I would find it strange*"(PA05). Another of the participants who tested the prototype expressed a recognized concern when also asked towards the experience of not having a face to talk to: "*It didn't bother me. I didn't think about it. It's more that I feel like there could be some information lost if my answers aren't followed up on*"(PA04).

The expressed concern is also shared among the physiotherapists during the focus group interview. One pointed towards a specific answer displayed and remarked: "*It seems like someone who wanted to say a lot. I could also imagine that it could be given to a virtual assistant because there are many things you want to have said, but the information doesn't necessarily make sense as a whole*"(PH05). The answer which was referred to was indeed recorded in the digital scenario, but also provided by the person who expressed the mentioned concern about potential loss of coherence if responses are not followed up on. Two comments by other physiotherapists insinuate the importance of being able to make meaning of patients responses: "*With some questions, I expect to receive an answer that has a sort of conclusion. If it's not concrete enough, then I'm not satisfied*"(PH02) and one other saying "*If it's unspecific, then it's difficult to measure*"(PH06).

The physiotherapists agreed on that patients need to be informed about the conventional session that follows the digital anamnesis recording. The information was expressed as crucial in practice for eliciting the appropriate type of responses. One physiotherapist said: "*I believe that if they know in advance that their responses to these machine questions will be followed up on, then I don't think they will go off on a tangent, because they know there will be*

follow-up. I think this will lead them to give more precise answers"(PH02). Although all participants recruited to test the prototype were briefed on the context for engaging with the self-anamnesis (appendix B), there still appeared to be single unsatisfying answers identified by the physiotherapists during the focus group interview.

6.3.5 Summary and tangential insights

To summarize, the focus group interview revealed somewhat consensus among the participating physiotherapists regarding the absence of a discernible distinction between responses given in a digital scenario versus a conventional one. Their varying perspectives and hypotheses regarding what characterizes a response to a VA indicated a level of uncertainty and inconsistency in their understanding and perception. However, their feedback predominantly revolved around the prototypes limitations posed by the lack of physiotherapists' capabilities, particularly in terms of customizing to each patient's needs, asking follow-up questions, and listening actively while physically present. While referring to patients needs, their feedback also implies that the self-anamnesis prototype cannot entirely supplant their need to seek clarification and obtain in-depth answers in certain cases.

One last identified, but tangential, theme occurs within our data as relating to more practical implications of implementing the system in practice. As mentioned in chapter 4, we engaged in a conversation with participants of the focus group interview discussing the concept of the self-anamnesis to assess their perspectives. One participant expressed: *"I think it sounds like a good idea. The only thing I'm thinking about at my clinic, where we are 7-8 physiotherapists, is that we constantly have patients coming in at the same time. It would be difficult logistically to have space and opportunity, and there isn't a full-time secretary available during the clinic's opening hours to assist the patients*(PH05)". Another participant followed up on that comment and said: *"I have the same concerns. We would also have space problems. We are three at work and have three rooms, so we don't have the option to send them into a room where it can be recorded. The idea is good if it can be solved*"(PH07).

The identification of these practical concerns underscores the need for further exploration and refinement of the self-anamnesis tool to ensure its seamless integration into clinical practice. When we towards the end of the following chapter outline avenues for future work that could build upon our achievements, we will elaborate on two ways to approach the logistical challenges.

Chapter 7

Discussion

The discussion chapter reflects on our work, synthesizing our results and comparing them with prior research. At the outset, our research endeavors aimed to streamline and enhance the crucial process of collecting anamnesis information and enable quicker initiation of necessary tests and treatments in physiotherapy practice. By transferring some responsibilities from the physiotherapist to a self-anamnesis system, we sought to alleviate the responsibility of documentation and record-keeping, allowing physiotherapists to focus more on diagnosis and treatment. Our efforts focused on elucidating whether patient data is at risk of deterioration if collected digitally without the physiotherapist's presence.

As our findings show, there appears to be no difference in how patients respond to anamnesis questions when interacting with a prototype of a VA, because physiotherapists can not distinguish their answers from those in conventional anamnesis conversations. Additionally, we were curious about how patients and physiotherapists perceive the self-anamnesis tool for clinical practice. In accordance with our second research question, the additional parts of our findings highlight areas for improvement, which will be elaborated on in the two following sections, 7.1 and 7.2.

A similar research objective to ours was found in two of the included studies in this report, though there were slight differences in their and our study parameters. The research by Boissonnault and Badke 2005 focused on documenting the accuracy of self-administered patient anamnesis in physiotherapy practice, however the data was gathered using a 92-item paper questionnaire. Even though parameters such as publication date and the applied anamnesis collection tool affect the applicability of their findings to ours, both results support the utility of self-administered anamnesis reporting.

The other study included is a systematic review by Dale and Hagen 2007 exploring how personal digital assistants perform as collection tools of patient-reported outcomes in clinical research, compared to pen and paper diaries. As specified metrics, the research focused on system acceptance and data accuracy, among other, somewhat

aligning with our own research objectives. Despite clear differences in clinical scope and specified research field, our findings support the use of digital tools for collecting patient data. This indicates that even with advancements in both IT and physiotherapy since 2007, the recognition of digital tools' effectiveness in health care remains consistent.

Through empirical investigation and analysis, we have tried to illuminate whether AI, in the form of a VA, can be utilized to collect patient data in physiotherapy practice. One of the key contributions of our work lies in providing insights into some of the practical implications of implementing a digital self-anamnesis tool in physiotherapy practice. By examining the perspectives of patients and experiences of healthcare professionals, we have been able to identify opportunities for optimizing the effectiveness and usability of such a tool. Our study stands out due to its focus on the concept of self-anamnesis within VA, a concept which to our knowledge was overlooked by researchers thus far. However, a related study by Denecke et al. 2018 explored self-anamnesis using CUI utilizing a chatbot.

As considerations for future work, Denecke et al. 2018 proposed expanding the concept towards a VA for gathering patient information - mirroring our research idea. However, the focus of our study was different. While Denecke et al. 2018 focused on the usability of their application, our primary research objective was to assess whether the data collected through our system would be different from data collected through conventional methods. Nevertheless, we also investigated how users perceived their communication with the system. Our findings align to some extent with those of the aforementioned study. For instance, the *low-pressure environment* defined by Denecke et al. 2018, allowing patients to take their time in answering anamnesis questions, was also a principle of our intended concept with the self-anamnesis. This element was positively referred to by those who tested our prototype.

Additionally, one of the key benefits highlighted in their application was the users' ability to ask clarifying questions if something was unclear. While this feature was not present in our prototype, data from both prototype testing and focus group interview underscored its importance. This aspect will be further addressed in the subsequent sections.

Our research has illuminated the nuanced dynamics of patient-provider interactions within the realm of AI-driven healthcare, thereby informing future directions for research and development in this field. Simultaneously, it provides insights with significant implications for practice. Practitioners across various healthcare domains, such as medical practice and occupational therapy, can leverage the findings from our study to explore the integration of a self-anamnesis system into their respective workflows. This expansion of the concept opens up new opportunities for interdisciplinary collaboration and innovation in healthcare delivery.

As digital healthcare technologies continue to proliferate, ensuring data privacy and security becomes paramount. In continuation of our research, there is a need for politicians to establish clear guidelines and standards for the development and implementation of a self-anamnesis system. As prior research indicates, the specific focus might be on ensuring the patient's confidentiality and trust in digital health solutions.

Besides learning that more actively engaging patients in the data collection process with a self-anamnesis system yields no discernible difference in patients' answers, it also does not deter patients from accepting the implementation of such a system, especially when assured of a follow-up consultation with a physiotherapist. As an implication on society, embracing patient-centered approaches that merge the efficiency of digital solutions with human care can foster a healthcare environment that prioritizes both clinical effectiveness and patient experience. This represents a shift towards patients becoming more invested in their own care journey and encourages more collaboration between patients and healthcare providers.

In the following, we will address the shortcomings of our work and delve into specific technical limitations of the prototype. We also outline future work that could build upon our achievements and address identified conceptual challenges inherent in the self-anamnesis process.

7.1 Limitation - Lack of customized self-anamnesis

Initializing research with a standardized self-anamnesis model offered advantages, such as consistency in data collection and ease of implementation. The pre-defined questions helped ensure consistency in the information collected from patients across different tests but also facilitated easier comparison to patients' answers in the conventional physiotherapy scenario, as the prototype's questions were identified in at least one of the collected anamnesis recordings. While implementing a standardized self-anamnesis with fixed questions is simpler and requires fewer resources, compared to developing a customized VA, we can not ignore the complexity of health care encountered both through our pre-study, related work and findings.

Even though, it was possible to generate themes of questions for the standardized self-anamnesis, the collected conventional recordings were highly diverse, and the questions asked were inevitably tailored to each patient, their issues, and the context of being physically present with each other - physiotherapist and patient. To exemplify the latter, during one of the recordings the physiotherapist asked: ″Can you show me where the pain is?", requiring his physical presence in order to see where the patient pointed to. When listening to the conventional anamnesis recordings, we also recognized the use of follow-up questions. For example, one patient mentioned: ″I have crooked legs", prompting the physiotherapist to ask ″Why do you feel that you have crooked legs?". For one, the follow-up question allowed the physiotherapist to show interest and gain a

further understanding but also re-using the patient's terminology, 'crooked legs', could enhance the patient's sense of being heard and understood.

The focus group also advocated the use of follow-up questions, especially when patients offer ambiguous answers, veer off-topic or use a descriptive terminology unfamiliar to physiotherapists. It is clear, that follow-up questions have a multi-dimensional purpose. Follow-up questions not only allow therapists to gather detailed information but also provide patients the opportunity to clarify their thoughts, express concerns, and add additional context. Moreover, they demonstrate the genuine interest of the physiotherapist in the patient's well-being. This need for follow-up questions was also echoed in interviews held with those who tested the prototype, expressing a desire for the ability to clarify any misunderstandings, revisit previous questions to add details they remembered later, or record additional information towards the end of the session that wasn't initially addressed.

Recognizing the control that physiotherapists inherently bring to structure and guide the anamnesis conversation, the responsibility of ensuring that pertinent information is elicited and documented effectively shifts to the software's capability. With features enabling customized interactions and the discernment, the system would mirror the decision-making and information processing found in patient-physiotherapist interaction.

7.2 Future work

Recognizing the strengths and benefits of the conventional physiotherapy setting for anamnesis gathering underscores the need to explore avenues for adapting the prototype to individual patient needs and preferences, a topic we will further elaborate on.

7.2.1 Customizing the self-anamnesis

Because customization is a limitation in our study, future research could contribute to the development of more effective and patient-centered self-anamnesis solutions in physiotherapy practice. Such contributions could be for the investigation of methods for dynamically adjusting the questioning strategy based on patient responses. This approach would allow for a more tailored and responsive interaction, ensuring that the self-anamnesis tool adapts in real-time to the specific concerns and circumstances of each patient. Another avenue could be to integrate and fully utilize the potential of NLP technology into the self-anamnesis model, enabling a more natural and conversational interaction between patients and the technology. The tools leveraged by NLP techniques allow more nuanced and contextually relevant responses from the technology, which might not only enhance the user experience but potentially also lead to more precise

assessments.

One last suggestion, related to customization, could be for researchers to explore the use of AI-driven algorithms with the ability to modify and individualize based on additional factors such as demographic information or previous medical history. To exemplify, the latter would particularly be relevant in cases where a patient has undergone a recent operation related to the current symptoms or has been referred specifically to physiotherapy treatment from another healthcare provider.

Although our findings indicate a lack of conversational flexibility in the prototype, it is important to acknowledge the potential mitigations that increased customization may offer. One main concern is how customization might affect the physiotherapists' workflow. For example, in a complex patient case involving multiple diagnoses or a long history of operations and treatment related to the current pain, the number of questions asked by the VA might be significantly higher compared to a patient with a one-time sprained ankle. This concern arose from the pre-study interview, where the physiotherapist explained, that the first session for anamnesis gathering can vary from 15 minutes to the entire allotted time, depending on the complexity of the patient case. This creates a paradox, when physiotherapists are left with less time to uncover and document the more complex patient cases and more time for simpler ones. The question then might be how to design a balanced system that accommodates complex patients while still time-effectively gathering the necessary information to avoid disrupting the physiotherapist's work schedule. One suggestion could be to implement a time limit and maintain some level of standardized questioning, with the system then providing information to the physiotherapist about what it did not have time to uncover.

7.2.2 Adapting for implementation in practice

Despite improving the prototype with the explored suggestions for features, one concern remains in regards of implementing the self-anamnesis in physiotherapy practice. The previously mentioned logistical challenges, involving aspects such as lack of space, simultaneous treatment of multiple patients and (partly) absence of administrative assistance in the reception area, are crucial to address.

One approach is to explore possible adjustments in practice to align with the intended concept. If redesigning existing clinical spaces is not feasible, future work could focus on investigating options for flexible scheduling within the clinic to accommodate the self-anamnesis process. This could involve identifying periods of lower patient flow, staggering the first appointments with new patients or allocating specific time slots for self-anamnesis sessions.

If the potential adjustments in practice fail to align with the principles of the concept, it is worth considering alternative strategies, such as adjusting the concept itself to better

suit physiotherapy practice. To accommodate the absence of administrative assistance, future work could focus on creating structured guidelines or prompts to assist patients in completing the self-anamnesis independently. Finally, we can not ignore a rather obvious solution enabling remote completion prior to an appointment. While there is a concern for how we ensure patient compliance with remote completion, some innovative and unexplored initiatives might have potential. One suggestion could be to conduct studies assessing the effectiveness of initiatives aimed to improve patients understanding and motivation for completing the self-anamnesis remotely. Additionally, research could explore the impact of providing clear instructions and encouraging incentives, or investigate the efficacy of systems developed to prompt and remind patients to complete the self-anamnesis.

7.3 Summary

While the introduction of a virtual assisted self-anamnesis in physiotherapy is a novel approach to our knowledge, we could still draw parallels between our work and several related studies, revealing coherence in our findings. The primary objective of our work was to determine whether the anamnesis data would be affected if the patient were to record answers to anamnesis question using a VA when comparing with answers from conventional anamnesis recordings. Our analysis suggests that it is not the case.

Among our findings we identified several limitations of the developed prototype, evolving around the absence of customization feature such as the inability to ask follow-up questions. While our analysis did not reveal any definitive evidence of distinguishable responses compared to the conventional scenario, the concept of customization deserves further investigation. Physiotherapists underscored the importance of specific, concrete, and measurable data, while also acknowledging the value of more elaborated responses in certain contexts. By integrating features that foster customization, there's a potential for improving patient's responses to the extent that the physiotherapist may experience less need for follow-up questions during patient consultations after the completion of self-administered anamnesis.

Furthermore, limitations of the concept itself were identified, particularly logistical challenges associated with its implementation. To address these different limitations, ideas for future work were discussed, such as fully utilizing the technology within NLP and discovering the possibilities of adjusting either the environment or the concept to align.

Chapter 8

Conclusion

In this chapter, we will summarize the project and draw conclusions based on how effectively it addresses the initially defined research questions. We will clarify the contributions of the work, highlight key points and look towards the future of AI in healthcare as informed by our work and findings. The research questions for this project were formulated:

How do physiotherapists perceive patient responses to anamnesis questions obtained through interaction with a virtual assistant compared to those obtained through conventional physiotherapy sessions?

How do participants evaluate the communication when interacting with either a physiotherapist or the virtual assistant?

In our work, we set out to explore the impact of a virtual assisted prototype on patient-reported anamnesis compared to conventional physiotherapist interviews, and to evaluate the communication from a patient perspective. Our findings affirm that we have delivered on these objectives. The empirical investigation reveals that physiotherapists were unable to distinguish answers provided through interaction with the self-anamnesis system from those provided during conventional interviews. These results suggest that patients do not respond remarkably different to questions related to their medical history when prompted by a system specifically designed for the purpose.

Furthermore, from the patient perspective, the evaluation of communication revealed that while traditional face-to-face interactions were valued for their human touch and active listening, the virtual assistant was also perceived favorably. Participants appreciated the opportunity for reflection and design for verbal responses, but requested for enhanced customization features enabling them to pose clarifying questions to the system. These communication preferences are important to consider when integrating new tools and technologies into physiotherapy practice.

We have contributed to the understanding of adoption and advancement of AI in healthcare by showcasing the application and integration of AI technologies with NLP, and the ways in which they so far improve clinical settings. Although our findings highlight the positive reception and effectiveness of collecting patient data with a VA and voice recording, we must also consider the shortcomings identified in our research. Building upon our achievements, it becomes crucial to investigate how to solve logistical challenges while acknowledging that additional obstacles may still need to be overcome.

As emphasized by Norman and Stappers 2015, deployment is as important as design, yet it often gets neglected. This brings awareness to areas of caution, which are also applicable to our work. Related research advocates that systems need to be observed over longer deployment periods, acknowledging the intricate process of effecting change in complex settings. Given that our findings are based on a limited number of participants in a simulated physiotherapy practice, future work must focus on ensuring long-term observation in real-world settings to fully realize and uncover the potential of our system.

As we continue to refine the technology and address challenges related to physiotherapy practice, we potentially move towards more efficient and patient-centered treatment. The implications of our work highlight the self-anamnesis system's potential for integration across various healthcare domains, enabling the responsibility of data collection and initial assessments to be shared among the system, providers of healthcare and patients themselves. While the technology cannot fully replace healthcare professionals, it has potential to augment their capabilities by freeing up more time for patient engagement and providing the necessary data to initiate further treatment.

Bibliography

Alharbi, Sadeen et al. (2021). "Automatic speech recognition: Systematic literature review". In: *IEEE Access* 9, pp. 131858–131876.

Ali, Ahmed and Steve Renals (2018). "Word error rate estimation for speech recognition: e-WER". In: *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pp. 20–24.

Aung, Yuri YM, David CS Wong, and Daniel SW Ting (2021). "The promise of artificial intelligence: a review of the opportunities and challenges of artificial intelligence in healthcare". In: *British medical bulletin* 139.1, pp. 4–15.

Blandford, Ann (Nov. 2019). "HCI for health and wellbeing: Challenges and opportunities". en. In: *International Journal of Human-Computer Studies* 131, pp. 41–51. ISSN: 10715819. DOI: 10.1016/j.ijhcs.2019.06.007. URL: https://linkinghub.elsevier.com/retrieve/pii/S1071581919300771 (visited on 03/06/2024).

Boissonnault, William G and Mary Beth Badke (June 2005). "Collecting Health History Information: The Accuracy of a Patient Self-administered Questionnaire in an Orthopedic Outpatient Setting". en. In: *Physical Therapy* 85.6, pp. 531–543. ISSN: 0031-9023, 1538-6724. DOI: 10.1093/ptj/85.6.531. URL: https://academic.oup.com/ptj/article/85/6/531/2805028 (visited on 03/18/2024).

Braun, Virginia and Victoria Clarke (2006). "Using thematic analysis in psychology". In: *Qualitative research in psychology* 3.2, pp. 77–101.

Chen, Chen et al. (2021). "Understanding barriers and design opportunities to improve healthcare and QOL for older adults through voice assistants". In: *Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility*, pp. 1–16.

Dale, Oystein and Kaare Birger Hagen (Jan. 2007). "Despite technical problems personal digital assistants outperform pen and paper when collecting patient diary data". en. In: *Journal of Clinical Epidemiology* 60.1, pp. 8–17. ISSN: 08954356. DOI: 10.1016/j.jclinepi.2006.04.005. URL: https://linkinghub.elsevier.com/retrieve/pii/S0895435606001934 (visited on 03/18/2024).

Davids, Joseph, Niklas Lidströmer, and Hutan Ashrafian (2022). "Artificial Intelligence for Physiotherapy and Rehabilitation". en. In: *Artificial Intelligence in Medicine*. Ed. by Niklas Lidströmer and Hutan Ashrafian. Cham: Springer International Publishing, pp. 1789–1807. ISBN: 978-3-030-64572-4 978-3-030-64573-1. DOI: 10.1007/978-3-030-64573-1_339. URL: https://link.springer.com/10.1007/978-3-030-64573-1_339 (visited on 02/24/2024).

Denecke, Kerstin et al. (2018). "Talking to ana: A mobile self-anamnesis application with conversational user interface". In: *Proceedings of the 2018 international conference on digital health*, pp. 85–89.

Durling, Scott and Jo Lumsden (2008). "Speech recognition use in healthcare applications". In: *Proceedings of the 6th international conference on advances in mobile computing and multimedia*, pp. 473–478.

Gremigni, Paola, Marinella Sommaruga, and Michael Peltenburg (2008). "Validation of the Health Care Communication Questionnaire (HCCQ) to measure outpatients' experience of communication with hospital staff". In: *Patient education and counseling* 71.1, pp. 57–64.

Harari, Yuval N. (2019). *21 lessons for the 21st century*. eng. Spiegel & Grau trade paperback edition. OCLC: 1114327591. New York: Spiegel & Grau. ISBN: 978-0-525-51219-6.

He, Jianxing et al. (2019). "The practical implementation of artificial intelligence technologies in medicine". In: *Nature medicine* 25.1, pp. 30–36.

Holzinger, Andreas et al. (2017). "What do we need to build explainable AI systems for the medical domain?" In: Publisher: [object Object] Version Number: 1. DOI: 10.48550/ARXIV.1712.09923. URL: https://arxiv.org/abs/1712.09923 (visited on 03/06/2024).

Huang, Fei, Sean Blaschke, and Henry Lucas (Dec. 2017). "Beyond pilotitis: taking digital health interventions to the national level in China and Uganda". en. In: *Global Health* 13.1, p. 49. ISSN: 1744-8603. DOI: 10.1186/s12992-017-0275-z. URL: http://globalizationandhealth.biomedcentral.com/articles/10.1186/s12992-017-0275-z (visited on 03/06/2024).

Jaber, Razan and Donald McMillan (2020). "Conversational user interfaces on mobile devices: Survey". In: *Proceedings of the 2nd Conference on Conversational User Interfaces*, pp. 1–11.

Kim, Hyeji, Inchan Jung, and Youn-kyung Lim (2022). "Understanding the negative aspects of user experience in human-likeness of voice-based conversational agents". In: *Proceedings of the 2022 ACM Designing Interactive Systems Conference*, pp. 1418–1427.

Kocielnik, Rafal et al. (2021). "Can I Talk to You about Your Social Needs? Understanding Preference for Conversational User Interface in Health". In: *Proceedings of the 3rd Conference on Conversational User Interfaces*, pp. 1–10.

Kosterink, Stephanie Maria (2014). *The added value of telemedicine services for physical rehabilitation*. eng. OCLC: 899237348. Enschede: University of Twente. ISBN: 978-90-823196-0-6.

Laranjo, Liliana et al. (2018). "Conversational agents in healthcare: a systematic review". In: *Journal of the American Medical Informatics Association* 25.9, pp. 1248–1258.

Latif, Siddique et al. (2020). "Speech technology for healthcare: Opportunities, challenges, and state of the art". In: *IEEE Reviews in Biomedical Engineering* 14, pp. 342–356.

Levin, Esther and Alex Levin (2006). "Evaluation of spoken dialogue technology for real-time health data collection". In: *Journal of medical Internet research* 8.4, e590.

Milne-Ives, Madison et al. (2020). "The effectiveness of artificial intelligence conversational agents in health care: systematic review". In: *Journal of medical Internet research* 22.10, e20346.

Mirsky, Reuth et al. (Jan. 2020). ""PhysIt"—A Diagnosis and Troubleshooting Tool for Physiotherapists in Training". en. In: *Diagnostics* 10.2, p. 72. ISSN: 2075-4418. DOI: 10.3390/diagnostics10020072. URL: https://www.mdpi.com/2075-4418/10/2/72 (visited on 03/18/2024).

Nayak, MSDP and KA Narayan (2019). "Strengths and weaknesses of online surveys". In: *technology* 6.7, pp. 0837–2405053138.

Norman, Donald A. and Pieter Jan Stappers (2015). "DesignX: Complex Sociotechnical Systems". en. In: *She Ji: The Journal of Design, Economics, and Innovation* 1.2, pp. 83–106. ISSN: 24058726. DOI: 10.1016/j.sheji.2016.01.002. URL: https://linkinghub.elsevier.com/retrieve/pii/S240587261530037X (visited on 03/06/2024).

O'Donovan, Cian et al. (July 2023). "Empowering future care workforces: scoping human capabilities to leverage assistive robotics". en. In: *Proceedings of the First International Symposium on Trustworthy Autonomous Systems*. Edinburgh United Kingdom: ACM, pp. 1–5. ISBN: 9798400707346. DOI: 10.1145/3597512.3600210. URL: https://dl.acm.org/doi/10.1145/3597512.3600210 (visited on 02/23/2024).

Park, Jinhwan et al. (2018). "Fully neural network based speech recognition on mobile and embedded devices". In: *Advances in neural information processing systems* 31.

Rowe, Michael, David A. Nicholls, and James Shaw (Nov. 2022). "How to replace a physiotherapist: artificial intelligence and the redistribution of expertise". en. In: *Physiotherapy Theory and Practice* 38.13, pp. 2275–2283. ISSN: 0959-3985, 1532-5040.

DOI: 10.1080/09593985.2021.1934924. URL: https://www.tandfonline.com/doi/full/10.1080/09593985.2021.1934924 (visited on 02/24/2024).

Sezgin, Emre et al. (2020). "Readiness for voice assistants to support healthcare delivery during a health crisis and pandemic". In: *NPJ Digital Medicine* 3.1, p. 122.

Van Hartskamp, Michael et al. (Apr. 2019). "Artificial Intelligence in Clinical Health Care Applications: Viewpoint". en. In: *Interact J Med Res* 8.2, e12100. ISSN: 1929-073X. DOI: 10.2196/12100. URL: https://www.i-jmr.org/2019/2/e12100/ (visited on 02/24/2024).

Vollenbroek-Hutten, Miriam et al. (Oct. 2015). "From telemedicine technology to telemedicine services". en. In: *Proceedings of the 3rd 2015 Workshop on ICTs for improving Patients Rehabilitation Research Techniques*. Lisbon Portugal: ACM, pp. 113–117. ISBN: 978-1-4503-3898-1. DOI: 10.1145/2838944.2838972. URL: https://dl.acm.org/doi/10.1145/2838944.2838972 (visited on 02/23/2024).

WHO (1998). *A health telematics policy in support of WHO's Health-for-all strategy for global health development : report of the WHO Group Consultation on Health Telematics, 11-16 December, Geneva, 1997*.

Wienrich, Carolin, Clemens Reitelbach, and Astrid Carolus (2021). "The trustworthiness of voice assistants in the context of healthcare investigating the effect of perceived expertise on the trustworthiness of voice assistants, providers, data receivers, and automatic speech recognition". In: *Frontiers in Computer Science* 3, p. 685250.

Appendix A

Interview guide - pre-study

- Introduktion til os
- Introduktion til specialet
- Formålet med interviewet

– Introduktion af hende

Uddannelsesbaggrund, års arbejdserfaring, patienttyper m.m.

– Fokus 1: Hvordan foregår anamnesen

1. Kan du med dine egne ord forklare, hvad en anamnese er?
2. Hvor lang tid bruger du på anamnesen?
3. Udfører du anamnesen på samme måde med alle dine patienter?
4. Følger du en form for skabelon?
5. Er der dele af anamnesen som ikke fungerer godt for dig?

– Fokus 2: Subjektiv oplevelse af anamnesens/samtalens værdi

1. Bidrager anamnesen til din fysioterapeutiske praksis? Hvis ja, hvordan?
2. Bidrager anamnesen til din relation til patienter? Hvis ja, hvordan?

– Fokus 3: Umiddelbare tanker om en digital anamnese-model

1. Kender du til klinikker hvor anamnesen ikke bliver prioriteret eller den samtale foregår på en anden måde?
2. Kender du til klinikker hvor anamnesen udføres ved brug af digitale løsninger? (chatbot, video, talegenkendelse)?
3. Gør du brug af anden form for teknologi, som hjælper dig i din fysioterapeutiske praksis?
4. Tror du et digitalt værktøj til udførelse af anamnesen er noget der vil tilføje værdi til din arbejdsgang, således at du ikke skal afsætte tiden til det selv?

- Mulighed for data gennem et par nystartede patienter (+ spørgeskema)
- Eventuelle kolleger vi kan interviewe
- Spørgeskema som patienten udfylder, kan du sende det til os (det gør journalsystemet) $\rightarrow$ kunne det være denne i AI form??

Appendix B

Interview guide - prototype test

– Før test - Forklaring af formålet med testen

Det der skal siges til alle patienter:

– Introduktion til systemet og testforløbet

- Under testen af applikationen vil du blive præsenteret for vores virtuelle assistent, som vil stille dig nogle forskellige spørgsmål omhandlende din skade/smerter/-gener
- Formålet med testen, som du skal igennem, er at indsamle lydoptagelser af din fortalte skades/sygehistorie, som man med et sundhedsfagligt term kalder for en "anamnese"
- Den applikation som du skal interagere med vil stille dig nogle forskellige spørgsmål, som du skal besvare ved at starte en lydoptagelse og stoppe den, når du er færdig med at besvare spørgsmålet
- Når du stopper en optagelse bliver du præsenteret for to muligheder - enten kan du optage et nyt svar eller gemme det svar, som du lige har optaget
- Når svaret er gemt, stiller prototypen dig det næste spørgsmål. Der er syv spørgsmål i alt, og når alle er besvaret, kan du afslutte testen
- Efter du har besvaret alle prototypens spørgsmål, vil vi gerne bede dig om at udfylde spørgeskemaet med otte spørgsmål der omhandler kommunikationen som du oplever den mellem dig og den virtuelle assistent
- Når du har udfyldt spørgeskemaet, må du gerne komme ud af lokalet, så vi ved, at du er færdig. Derefter vil vi stille dig nogle enkelte uddybende spørgsmål til din oplevelse med systemet

– Introduktion til scenariet

- Du skal forestille dig, at du har bestilt en tid til at blive tilset af en fysioterapeut
- Du er mødt ind på fysioterapiklinikken, men før du møder fysioterapeuten, skal du besvare nogle standardspørgsmål, som du selv står for at administrere på en tablet
- Du bliver henvist til et lokale, hvor der ligger en tablet på et bord, som guider dig gennem de spørgsmål, som fysioterapeuten skal bruge dine svar til for at kunne tage beslutninger relateret til eventuelle tests og din videre behandling

– Interview efter test

– Spørgsmål omhandlende konceptet

1. Hvad synes du om at blive stillet spørgsmål og skulle besvare dem med den her digitale opsætning?
2. Hvordan var det ikke at have ansigt på den, som du taler til?
3. Hvordan fungerede det for dig som et alternativ til det traditionelle spørgeskema, hvor du enten skulle udfylde svarene ved at skrive det ned eller afmærke et felt?

– Omhandlende UI

1. Var systemet intuitivt?
2. Var der noget som forvirrede dig undervejs?
3. Var der noget du manglede mens du interagerede med systemet?

Appendix C

Interview guide - focus group interview

– Overblikket

- Introduktion til os og projektet (figure C.1)
- Forklar hvordan vi har testet (figure C.2)
- Formålet med interviewet → hvad vi skal igennem (figure C.3)
- Præsentation af indsamlet data (figure C.4 and C.5)
- Præsentation af koncept (figure C.6)

Figure C.1: Introduction

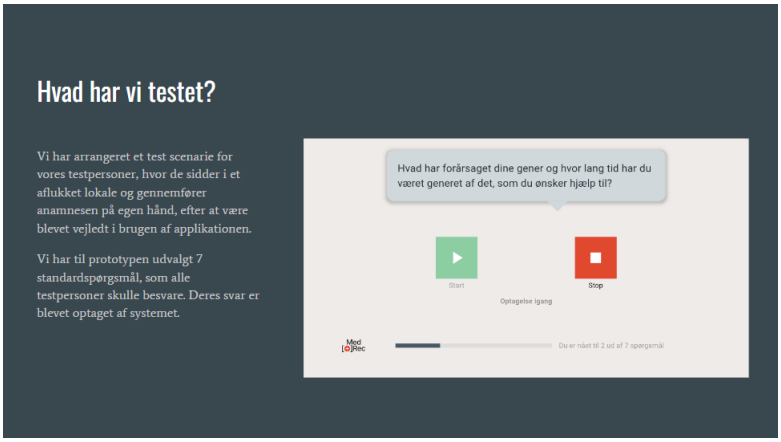


Figure C.2: Test explanation

– Spørgsmål under interview ved præsentation af indsamlet data

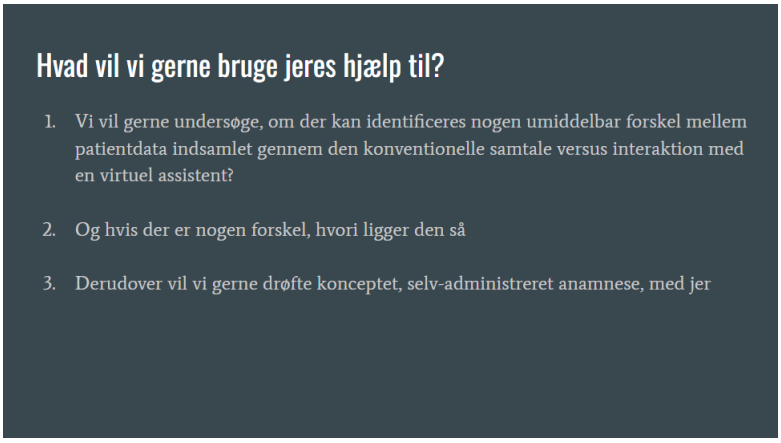


Figure C.3: Interview focus

1. Er der nogen af de her svar, som I vil udpege som givet under samtale med den virtuelle assistent?
2. Hvad er det, der gør, at du/I udpeger den/dem som digitale besvarelser?

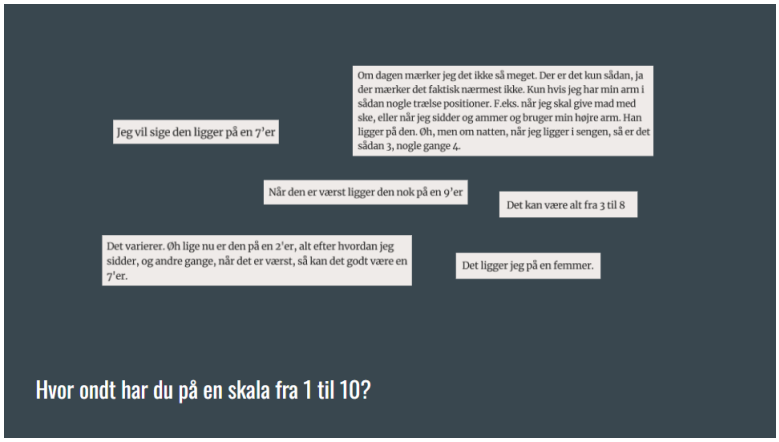


Figure C.4: Answers to anamnesis question, example 1

Figure C.5: Answers to anamnesis question, example 2

– Spørgsmål til konceptet

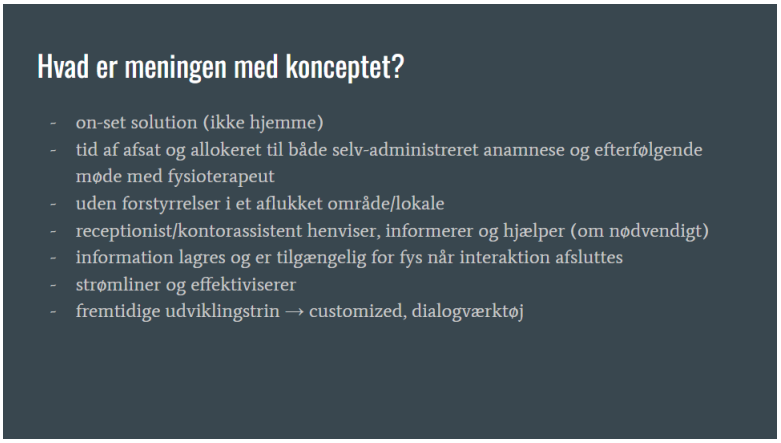


Figure C.6: Concept explanation and questioning

1. Hvad er jeres umiddelbare tanker, når I hører om konceptet?
2. Er der noget ved sådan en løsning som bekymrer jer?
3. Ser I nogle muligheder i at implementere en digital anamneseoptagelse?
4. Gør I brug af spørgeskemaer i praksis og hvad er jeres erfaring med det?