Bringing digital scribes into orthopedic consultations

Towards AI-assisted clinical documentation

Master thesis
Integrated Product Design
by Reka Magyari

TU Delft
attendi

Bringing digital scribes into orthopedic consultations

Towards AI-assisted clinical documentation

Author

Reka Magyari

Master thesis

MSc. Integrated Product Design | Medisign
Faculty of Industrial Design Engineering
Delft University of Technology

Graduation committee

Chair | Fernando Secomandi
Assistant Professor of Service Design for Emerging Technologies;
Department of Design, Organisation & Strategy

Mentor | Hosana Morales
PhD Candidate in Evidence-based design for healthcare;
Department of Sustainable Design Engineering

Company mentor | Diederik de Rave
Co-Founder
Attendi

Master thesis
by Reka Magyari

April, 2022

Merging data science and design fascinates me, especially in a healthcare context. This project involved everything I was looking for in a Master thesis: UX, AI and startup momentum. Time to say thank you to the many people who helped me to deliver this work.

To my graduation committee

I would like to thank my supervisory team for guiding me with your valuable expertise. To Fernando, for chairing my supervision. I was so excited to bring service design into such a complex context and learnt so much from our discussions. Hosana, I am really glad you mentored me throughout this project. You were always there for thinking along, helping to clarify my own thoughts and sometimes nerding out together. Gracias!

To Attendi

Diederik and Berend, bedankt! I am lucky you trusted me with this awesome challenge to graduate on. I still remember our first call thinking: *"Yess, I like this energy*!" You gave me a playground and freedom to grow. Dear Diederik, thank you so much. All the weekly talks, smart inputs and making the whole experience super fun at the same time. For always making time for me in your packed agenda. And Berend, thank you for answering my never-ending tech questions to finalize the blueprint. You guys are so ambitious yet modest, I find it impressive.

To the whole team, JW, Tirza, Omar, Arjan and Kacper. Thanks to your team spirit and Slack emojis, I never felt lonely during this thesis time. You have so much value in your hands, I really hope (y)our speech reporting magic will reach more and more users soon.

To all the dear surgeons who spared precious time to talk to me. Especially Thony, thank you for all your inputs in developing the Assistant. I was lucky to have your expert inputs every single week. And Joost, dank je wel! You were such an asset of the process and the results. I cherish your enthusiasm about AI.

To my friends

To my great friends in Delft, Budapest, Eindhoven and Munich. Thank you for being in my life regardless of the the smaller or bigger distances. Special thanks to Eef, who shared every single moment of this online masters journey. I could not have made it without you.

My lovely roommates and ex-quarantine family, thanks for creating a home away from home.
To my VolleyBolley community: you guys are amazing, keep it up! Thanks for all the time off, good games and fun times together.

To my family

Last but not least, to my entire family, my main supporters and critics. My brother, my Mom and Dad who taught me that anything is possible if I set my mind to it. You never pressured me to get into medicine but somehow one project came after the other. You made it possible to follow my dreams, köszönöm. Mindig hálás leszek.

Executive summary

Clinical documentation takes up 40% of clinicians' time. To ease the administrative burden of clinicians, digital scribes offer the potential to automate clinical note taking. Digital scribes are intelligent documentation softwares that combine automated speech recognition (ASR) and natural language processing (NLP). Digital scribes transcribes clinician-patient consultations and convert the conversation into structured clinical notes. Attendi, the client is developing a digital scribe for orthopedics: the Assistant. Hip arthrosis consultations have a well-structured anamnesis, therefore a promising use case for the digital scribe.

This Master thesis investigates the clinician perspective of implementing the Assistant. Literature was reviewed to understand the problem space and the design implications of the enabling technologies. Core concepts in human-AI collaboration such as system transparency and human control were identified to design for hybrid documentation. Also, the perspectives of recording consultations were translated into values for hospitals, clinicians and patients. The findings lead to both mutual benefits and tensions for the clinician-patient relationship and obstacles for implementation.

To contribute to developing the Assistant, user research was carried out in context by shadowing orthopedic surgeons to see their day-to-day workflow and understand the current cycle of clinical documentation. Several surgeons were interviewed to gain more in-depth views about the digital scribe. As synthesis, personas and journey maps were created both for a typical consultation cycle and a daily workflow. From the research phase, a list of user requirements were gathered in order to aid the design phase and future development.

The vision of the Assistant is to structure clinical notes through organised templates in order to enable partial automation. The product roadmap was considered to define the design goal: design for partial automation to speed up clinical documentation with automated suggestions and provide flexibility to fit individual documentation styles.

To envision the future journey with the Assistant, a service blueprint was made in combination with a storyboard. The service blueprint shows the user steps as well as the software processes throughout a consultation cycle. Three user flows were extracted and translated into the interface: recording a consultation, filling out the anamnesis template and query a specific patient file from the database. The design phase focused on developing the interface for anamnesis template completion where the clinician can accept, reject and edit the automated suggestions. To improve trust towards the system, transparency is implemented in the design by providing the option to check where the suggestions come from the transcript.

The designed wireframes were validated with orthopedic surgeons to clarify the value proposition of the solution. From the findings, future opportunities are formulated and limitations of the study are discussed. Potential research areas for developing digital scribes are proposed as future work in order to transition towards AI-assisted clinical documentation.

I hope you will enjoy reading this thesis, sparking many discussions and future innovations!

Table of contents

Reading guide

This reading guides aims to support you as a reader to go through this report. At the beginning of each chapter, a small summary of the chapter's content is presented. At the end of the chapters, the main takeaways are listed from the chapter. Reading these takeaways can provide an understanding of the entire design process to understand the key project insights.

All illustrations were produced throughout the project to make sense of the complexity. I hope they will also help you to read this report.

Key takeaways and reflections per chapter.

This is my own interpretation.

Paragraph title This is an important paragraph.

Quotes are formatted this way.

AI - Artificial Intelligence
ASR - Automated Speech Recognition
EHR - Electronic Health Records
HCI - Human-Computer Interaction
NLP - Natural Language Processing
UX - User Experience
STT - Speech-to-text

Chapter 1

Motivation for the project

In this chapter the rationale for the project is articulated. Attendi, the client company of the graduation project is introduced with their mission. Also, a high-level view on the product roadmap is illustrated to understand how this Master thesis contributes to developing the Assistant: a digital scribe for orthopedics.

Image source: Pexels

Introduction

As of today, clinical documentation is completely human-led which leads to a high administrative burden for clinicians (figure 1). To ease this burden, digital scribes offer the potential to automate clinical note taking and enable clinicians to focus on what they love doing: taking care of the patient. With digital scribes, taking notes will become computer-led and contribute to the transition towards true patient-centred care.

Digital scribes can help clinicians to spend more time on talking to the patient and less on typing. The content of the consultations needs to be administered by the clinician and saved as part of the patient file in the Electronic Health Records (EHR). Hospital EHR systems store all medical information in one place and it is the duty of the clinicians to fill out the necessary documents.

Figure 1: Human-led documentation

Attendi is developing a digital scribe for orthopedics: the Assistant. The Assistant listens to the conversations between doctors and patients and converts their speech into usable notes to save time. To transition towards computer-led documentation, the first step is to design for mixed-initiative documentation where the computer and the user both have relative control over the documentation process (figure 2). This Master thesis contributes to developing the Assistant through researching the context with a focus on the clinician perspective and developing the interface of the software.

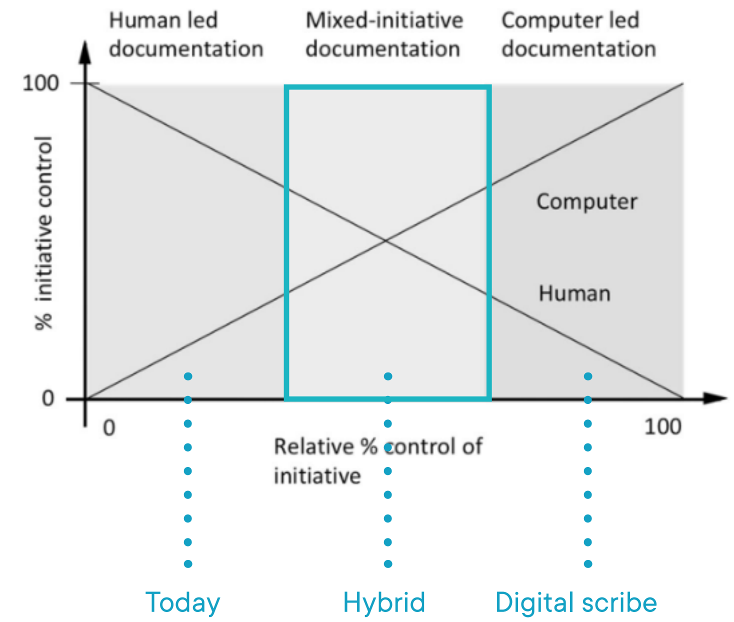


Figure 2: From human-led to computer-led documentation (Coiera et al., 2018)

1.1 Collaboration with Attendi

The client of the assignment is Attendi, a healthcare startup based in Amsterdam. The company's mission is to ease the administrative burden for healthcare professionals by building intelligent voice technologies that are optimised for care domains. If administrative tasks are automated, healthcare professionals regain valuable time, essential for delivering personalised patient care.

Voice-powered solutions allow healthcare professionals to interact with computers without shifting their focus from a patient to a computer. Attendi's speech technology accurately and instantly converts spoken words into written text (figure 3). Since the speech and language models are specifically trained on spoken Dutch healthcare data, the technology has a high accuracy for applications in healthcare.

The Attendi Assistant extracts medically relevant information and immediately converts this information into usable documentation. In collaboration with several hospitals, the company is introducing the first all-round digital scribe for orthopedics in the Netherlands.

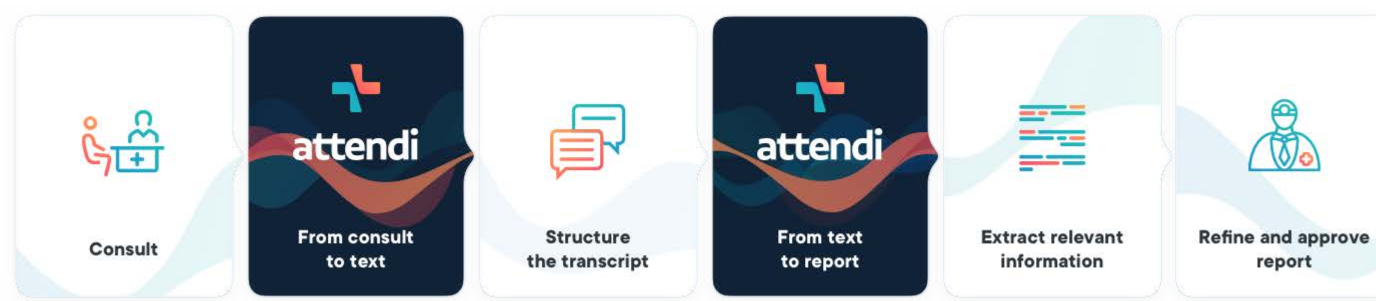


Figure 3: Product pipeline provided by the company

1.1.1 Project context

The context of the project is an orthopedic clinic where patient-doctor consultations take place concerning hip arthrosis. Hip arthrosis is a disease of the hip joint, caused by increasing age or injury and often treated by surgery. In a clinical consultation, typically a patient explains his or her health status and the clinician asks focused questions to formulate a diagnosis. In medicine, the information gained by the clinician is called the anamnesis* by asking specific questions relevant to the specialty. Hip arthrosis consultations are carried out by orthopedic surgeons, who also perform the surgery and welcome their patients for appointments in an outpatient clinic.

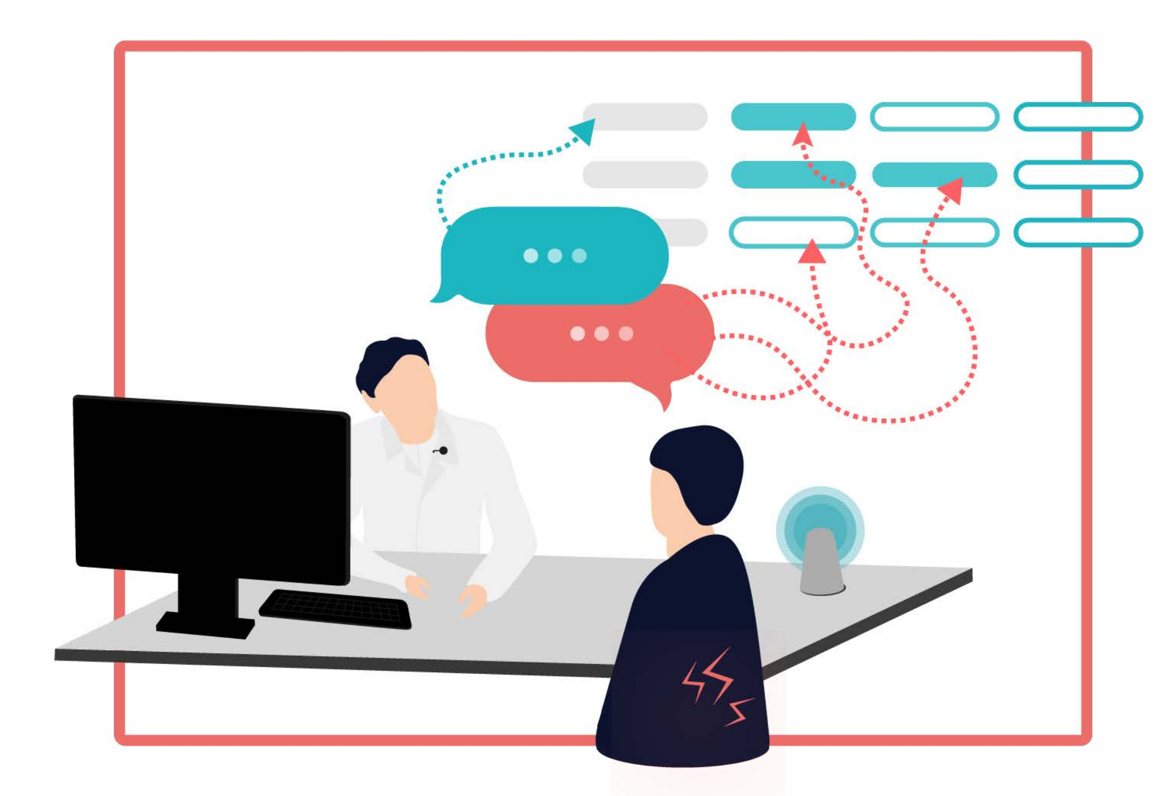


Figure 4: Hip arthrosis consultation with anamnesis

Why orthopedics?

Hip arthrosis consults are short, occur in high volumes and have a well-structured anamnesis (figure 4). Since the flow of the conversation influences the ability to compute the speech, these homogeneous consultations are a promising use case for the digital scribe. For the hip arthrosis use case, Attendi has curated an anamnesis standard jointly with orthopedic surgeons.

*the term **anamnesis** will be used frequently throughout the report*

1.1.2 Expected contribution

A roadmap (figure 5) for implementing the Assistant was provided by the company. The product is planned to be implemented in hospitals in 3 steps:

1) establishing a recording infrastructure
2) providing a template completion application
3) scaling the template options and enabling data sharing with the EHR system.

In order to bring digital scribes into hospitals, researching the consultation context and the end user is necessary. The cycle of clinical documentation has to be investigated to decide which tasks could the technology automate and support the clinicians in performing their jobs. Furthermore, the solution has to be integrated into the current and future clinical workflow of the medical professionals for better EHR development (Healthcare IT News, 2021). To provide a valuable report of the transcribed consultation to the clinician, in-depth user research is required. All in all, an interface needs to be designed to facilitate the interaction between the user and the software.

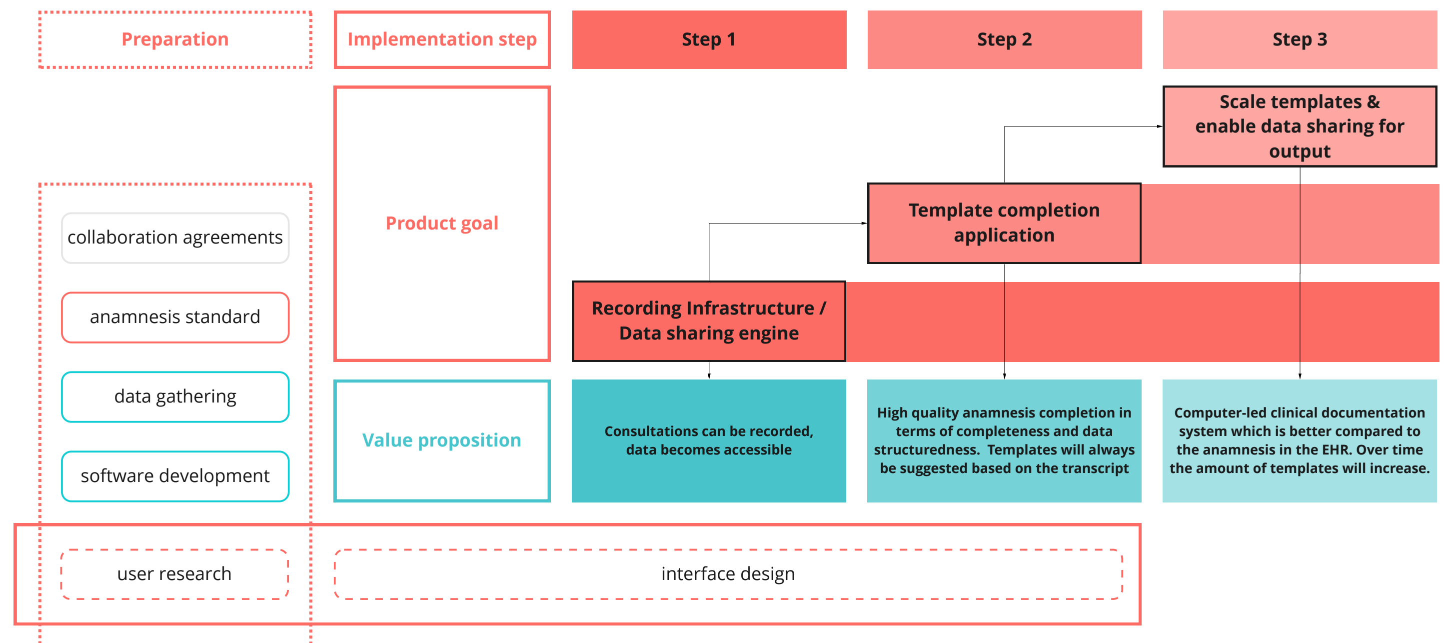


Figure 5: product roadmap with my contribution

To position my contribution in relation to the product roadmap, it is illustrated in figure 5. My main contributions are carrying out user research and design the interface. The research questions were formulated accordingly (figure 7).

1.1.2.1 Design approach & methods

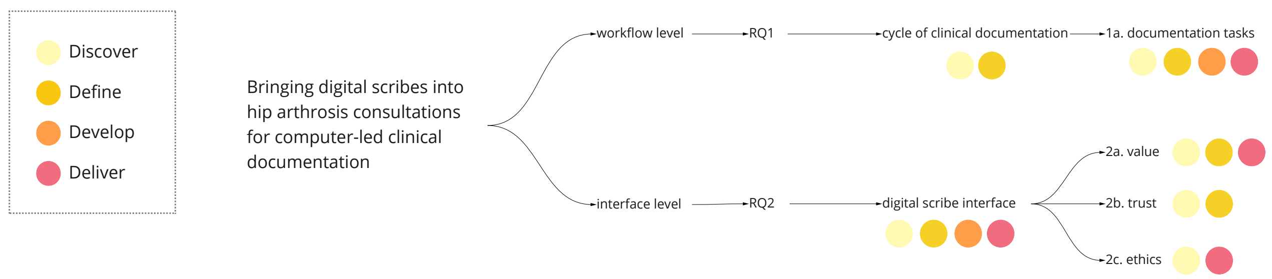


Figure 6: Research questions per design phase

RQ1. How can the full **cycle of clinical documentation** in hip arthrosis consultations be automated using NLP technologies?

1a. Which **documentation tasks** in hip arthrosis are time consuming/frustrating/sensitive to errors and technically feasible to automate?

RQ2. How can we design a clinically valuable, trusted and ethical **digital scribe interface** to aid computer-led documentation from a clinician's perspective?

2a. Valuable: What value can be delivered to hospitals, clinicians and patients by recording and transcribing consultations?

2b. Trusted: What are the key fears and obstacles for physicians and patients if consultations will be recorded and how can we manage those fears?

2c. Ethical: How can the digital scribe interface ensure that the clinician remains critical about the automated suggestions for entry in the EHR?

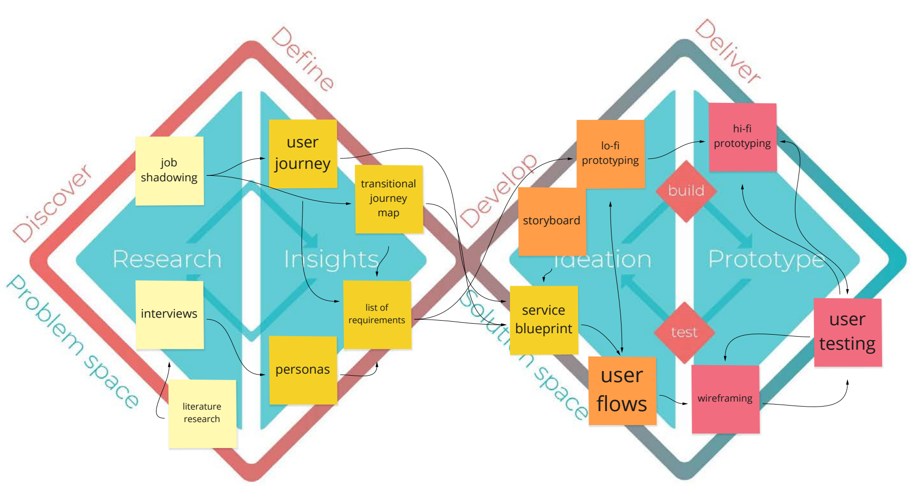


Figure 7: Planned methods during the double diamond process

The expected outcomes of this Master thesis were formulated to:

1. collect a **list of requirements** primarily from a clinician perspective
2. design **high fidelity wireframes** for the software interface based on the requirements
3. **validate** the envisioned user journey and interface
4. create a **roadmap** to suggest future features from a UX perspective

Chapter 1 takeaways

- The administrative burden is very high for clinicians since it takes up almost 40% of their time. It is the main rationale to invest efforts into the project.
- To develop the digital scribe, first a hybrid scenario (mixed-initiative) should be considered where both the computer and the clinician carries out parts of the tasks.
- Attendi's product pipeline was introduced to understand the steps how the data goes through from the consult to an approved report.
- To scope the project, hip arthrosis consultations were chosen as a use case due to the homogeneous nature of the conversation. The anamnesis standard is a central component of the design process.
- The Assistant will be implemented in 3 steps, step-by-step. To prepare for the implementation, the contribution of this Master thesis serves as a preparation in the form of user research and designing the interface.
- The research questions are divided on a workflow and an interface level. First, a high level understanding of the clinical documentation cycle is required then it is time to zoom in to the interface. Key aspects during the design phase will be value, trust and ethics.
- To design the interface of the Assistant, close collaboration is required with the company's Product manager, the Medical lead and the Tech lead.

Chapter 2

Potential of digital scribes

In this chapter, the potential of digital scribes is outlined. In order to understand intelligent documentation softwares, both Automated Speech Recognition and Natural Language Processing technologies are explained. Also, the clinician and the patient perspectives of recording consultations are discussed by reviewing secondary research. The goal of the literature research phase was to find answers to key questions such as:

What do digital scribes promise?
How do the involved technologies work?
What are the perspectives on recording consultations?

Image source: Pexels

The potential

In recent years, the healthcare sector has seen a rising number in clinician burnout due to the administrative burden that they face on a daily basis. On average, Dutch clinicians spend 40% of their time on administrative work ('Statement initiatiefnemers ORDZ', 2021) which comes at the direct expense of devoting time to the patient (Dugdale et al., 1999). For every face-to-face hour physicians spend directly with patients, two more hours are spent on desk work documentation in ambulatory care (Sinsky et al., 2016). Improving the work life of clinicians is one of the four pillars that drive innovation in healthcare according to the Quadruple Aim (Sinsky & Bodenheimer, 2014).

Clinical documentation refers to the process of typing a text record that summarises the interaction between patients and healthcare providers during clinical encounters (Rosenbloom et al., 2011). Implementing technology in a clinical context brings many advantages, for instance adopting Electronic Health Records (EHRs) initially offered the prospect of improved patient safety, better health care quality and clinician performance (King, Patel, Jamoom, & Furukawa, 2014). However, the regulated use of EHR systems has decreased physician satisfaction over time, increased documentation times (Ehrenfeld and Wanderer, 2018) and also negatively affected the clinician–patient relationship (Coiera et al., 2018). EHR systems suffer usability problems

Figure 8: Hybrid documentation

that make them inefficient, not usable or well aligned with clinical workflow (Gardner et al., 2018) which aspects add to decreased clinician satisfaction. To ease the administrative burden of clinicians, medical scribes were introduced who are trained individuals hired solely for administrative tasks, such as real-time documentation of clinician-patient encounters (Gellert, Ramirez, & Webster, 2015). Medical scribes proved to increase overall clinician satisfaction, chart quality and efficiency, without reducing patient satisfaction (Gidwani et al., 2017). However, medical scribes require extensive training before reaching their full potential, which is estimated to cost $6,317 (Walker et al., 2016) on average. Additionally, they are often medical students who tend to move on to attend medical school full time (Walker et al., 2016). These aspects do not make medical scribes an effective solution in the long term. Overall, the use of medical scribes does not reduce the clinical documentation burden, but simply shifts it from clinicians to others (Dusek et al., 2021).

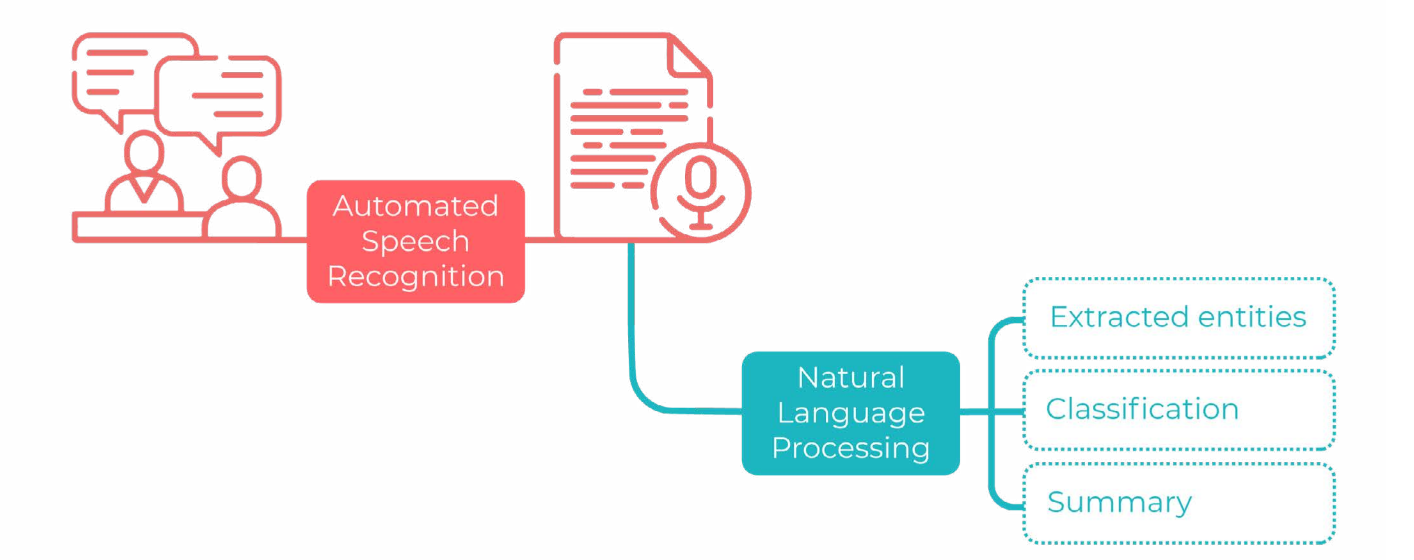


Figure 9: Digital scribe per definition involves ASR and NLP

2.1 Enabling technologies

The potential of digital scribes is to reduce the clinical documentation burden conducted by humans (figure 8). Digital scribes are intelligent documentation softwares that combine Automated Speech Recognition (ASR) and Natural Language Processing (NLP) technologies according to definition (Coiera et al., 2018) in figure 9. The software automatically transforms the recording of a clinical conversation into usable notes, by generating transcripts for encounter documentation (Chiu et al., 2018).

2.1.1 Automated Speech Recognition (ASR)

ASR is increasingly researched and used in healthcare (Jamal et al., 2017), where a microphone records a conversation and an ASR system transcribes the speech into text. The captured analogue speech signals get converted into digital forms which can be read by computers. Especially in medical conversations every word counts. To process the conversation, we need to understand 'what is being said' during a clinician-patient consultation and detect 'who spoke when'. In addition, punctuation marks such as periods, commas and question marks are automatically added to the text to generate readable and useful transcripts. The speech processing step is done by speaker diarization models in the software that assign speaker labels to dialogue segments of the conversation (Kanda et al., 2019). To rate the success of the speaker segmentation, diarization error rate (DER) and word error rate (WER) are calculated by the software. Both DER and WER accuracy indicators impact the consequent processing steps to extract relevant information.

2.1.2 Natural Language Processing (NLP)

The ASR system transcribes the conversation, and a set of NLP models are deployed to extract and summarise relevant information that is to be presented to the physician (van Buchem et al., 2021). NLP is at the intersection (figure 10) of linguistics and artificial intelligence (AI) that allows computers to process and analyse natural language data (Brownlee, 2019). In this context the NLP tasks could be split into extracting entities, classification categories and summarization (Du et al., 2019). In order to train the NLP algorithms, ground truth is established by manually annotating a medical transcript. While NLP comes with numerous technical challenges such as structuring and summarising clinician-patient conversations (Quiroz et al., 2019) in healthcare, literature suggests that the success of digital scribes will be dependent on the way interactions are designed between the clinician, computer and patient (Coiera et al., 2018).

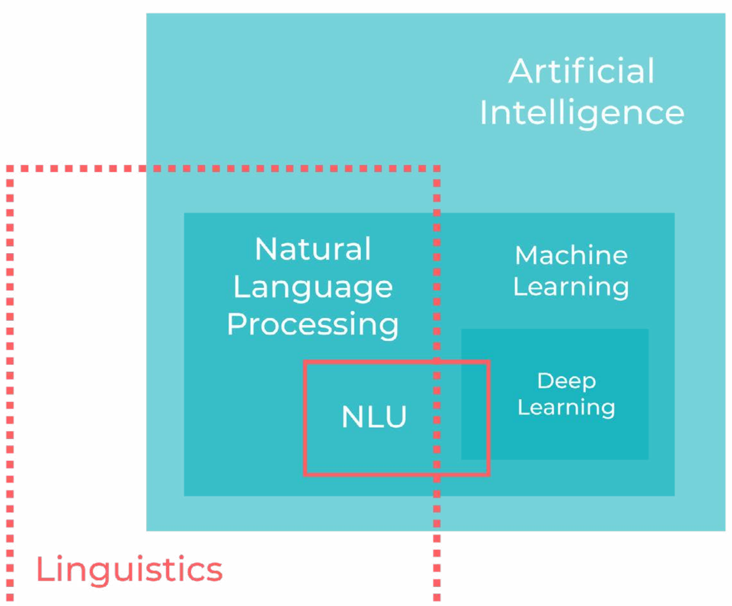


Figure 10: NLP is in the intersection of linguistics and machine learning

2.2 Perspectives on recording consultations

In order to implement digital scribes in hospitals, the challenge is to understand user requirements and tailor a desired experience between patient, computer and doctor. Traditionally clinician-patient encounters take place in hospitals, therefore the institution is also a stakeholder in the process of implementation. There are 'two sides' of the system, and both patients and clinicians have their own perspective on recording consultations. The key research questions to understand were the following:

2a. Valuable: What value can be delivered to hospitals, clinicians and patients by recording and transcribing consultations?

2b. Trusted: What are the key fears and obstacles for physicians and patients if consultations will be recorded and how can we manage those fears?

2.2.1 Value for hospitals

If consultations are recorded, the data offers numerous benefits (figure 11) for hospitals such as motivating clinician performance. Increased performance can also enable hospitals to conduct more transparent quality assurance (Turley and Metcalfe., 2020). Improving the quality of care can potentially lead to better health outcomes. The quality of care and safety depends on high quality data. The digital scribe structures data collection to enable research, access to higher quality and safety of care.

As of today, if patient consent and confidentiality is adhered to, recordings can be made. Hospitals face a rising number of patient-initiated recordings that clinicians are not always aware of. There is little evidence that recording influences consultation style (Turley and Metcalfe, 2020). A large-scale study found that patients would encourage normalising the process and expect the clinic to facilitate organising the recordings. In the Netherlands, an increasing number of healthcare professionals have experience with audio-recording consultations (van Bruinessen et al., 2017) and confirm the evidence-based benefits.

Figure 11: Hospital benefits

2.2.1.1 Obstacles for implementation

Moreover, from a hospital perspective there are further obstacles in the way of implementing digital scribes. The technical aspects of handling the recorded consultation data raise both privacy and legal concerns that should be managed on an institution level. In the Netherlands, a nation-wide survey (Voskens et al., 2022) also reported on these concerns where Dutch surgeons show a positive and open attitude towards AI. Furthermore, the recording has implications for the confidentiality of the conversation, therefore safely storing the data is an important factor and ensuring that the patient always remains the owner of the information. As a consequence, the use of a protected database is an important part of the system and in the process of certifying the technology (Ghatnekar et al., 2021) before its introduction in clinical use. In the future, integration with EHR systems will be essential and require further technical considerations.

These obstacles (figure 12) are input for business requirements during the implementation.

Figure 12: Obstacles for implementation

2.2.2. Clinician perspective

The primary user of the system will be the clinician so their perspective is central. Prior research from the US and UK discusses clinicians' attitudes towards audio recordings. According to a mixed methods study (Elwyn et al., 2015), it was reported that patients asked for approval to record their clinical conversation or did without asking. Joshi et al (2020) investigated physicians' attitude by an online survey among general medicine practitioners in the US. The participants considered the recording to be of benefit to the patient or their family, however 50% did not wish to allow recordings and 22% were neutral. The most common concerns were that patients might have undeclared motives, such as a wish to use the recording for litigation, and concerns that the audio files could be used inappropriately, for example, placed online (Joshi et al., 2020). Physicians stated that they might agree to record only part of the visit, provided they could control the process. It was also concluded that generally physicians had no knowledge of the laws governing patient-initiated recordings. Clinicians are open to allow recording if long writing in EHR is replaced (Barr et al., 2018).

2.2.2.1 Clinician benefits

There are clear benefits (figure 13) of recording consults, such as allowing the clinician to spend less time on documentation, and devote that time to the patient to communicate more effectively (Dommershuijsen et al., 2021). Enabling clinicians to provide personal attention and empathy makes a consultation successful and more valuable (Bensing et al., 2011). Decreased documentation time could increase physician satisfaction as well as reduce burnout rates in healthcare.

Moreover, recording the consultation could serve as an incentive to perform better when delivering care (Tsulukidze et al., 2015). Motivating performance could potentially lead to increased interest in maintaining medical knowledge (Elwyn and Buckman, 2015). The recordings could also benefit research purposes, and serve as input for teaching and learning opportunities (Turley and Metcalfe, 2020).

Figure 13: Clinician benefits

2.2.3 Patient perspective

From a patient perspective, there has been prior research executed to understand their views and motivations (figure 14) regarding recording consultations. A study (Elwyn et al., 2015) reported that in total 69% of the participants indicated their wish to record clinical consultations either in secret or with permission. Interviewed patients reported fears that recording would violate the expected etiquette in clinical encounters. Patients who recorded consultations without permission were mostly male and have a below university level education (Elwyn et al., 2015). Others advocated for asking for permission mainly to prioritise the relationship with the clinician and protect trust, make accountability public, and have an improved quality of recording.

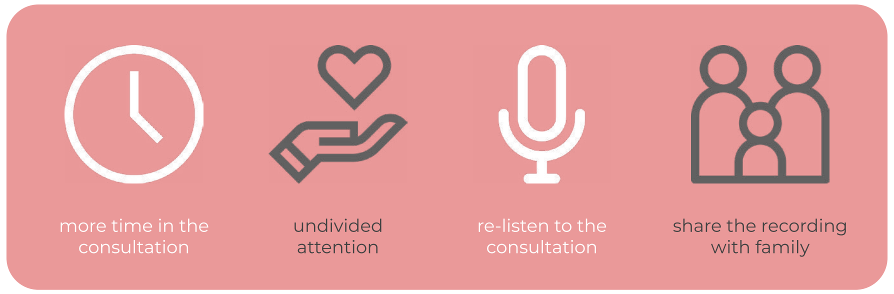


Figure 14: Patient benefits

2.2.2.1 Patient benefits

There are several patient benefits of recording consultations. There is evidence that 60% is forgotten as patients tend to block and have difficulty storing information during consults (Joshi et al., 2020). Therefore, the recording could allow patients to re-listen to the exchanged information for better understanding. Also, the option to share the recording with family and friends to make sense of the expert advice is reassuring to patients. An interesting aspect was raised that the recording has therapeutic purposes as well such that it can enable the patient to feel supported by the 'tone of voice' of the clinician (Elwyn et al., 2015). The main patient profiles to benefit from recordings are older patients, patients from lower socio-economic status and patients with impaired abilities (Dommershuijsen et al., 2021).

2.2.4 The clinician-patient relationship

Figure 15: The clinician-patient relationship

2.2.4.1 Mutual benefits

There are numerous mutual benefits of recording patient-doctor consultations (figure 15). Elwyn et al. (2015) suggests that recording has the potential to make clinicians more attentive from a patient perspective. In a study among patients aged 50 years or older (Dommershuijsen et al., 2021), more attention leads to improved confidence in the relationship for both sides and higher satisfaction with the treatment.

There is little evidence that recording influences consultation style (Turley and Metcalfe, 2020). Where complex and emotional subjects are discussed, such as oncology or pediatrics, clinicians already offer recordings (Elwyn et al., 2015). Being able to re-listen to the conversation allows the patient to better prepare for the next consultation which leads to more informed decision-making. Ultimately, improved communication (figure 16a) can even lead to better health outcomes (Dommershuijsen et al 2021).

Figure 16a: Mutual benefits

2.2.4.2 Tensions

When synthesising the perspectives from clinicians and patients, some tensions (figure 16b) occur regarding recording consultations such as consent, influence on conversation and the reason behind recording. Regarding consent, it is important to note that patient consent is strictly needed by law while clinician consent is not (Tsulukidze et al., 2015). On the other hand, if both parties do not consent to the recording it could damage the patient-doctor relationships (Elwyn et al., 2015).

Recording does not suit some doctor's own values (Tsulukidze et al., 2015) while patients favour the idea of having proof of received care. There are increasing statistics of recording in secret (Elwyn et al., 2015) that clinicians can not influence nor like. The patient-initiated recording behaviour challenges the established etiquette of clinician-patient interactions (Elwyn et al., 2015). For this reason many patients advocate recording consultations with permission to maintain the relationship and trust between the clinician and the patient. Furthermore, the recording can also influence what doctors say or not and patients might be less likely to admit to problems (Elwyn et al., 2015) during the consultation.

Figure 16b: Tensions

Chapter 2 takeaways

- Typing in the Electronic Health Records results in a number of problems: unstructured notes and also decreased physician satisfaction. Clinical documentation affects both the clinician and the patient and their relationship negatively.
- Medical scribes were proposed (mostly in the US) to ease the administrative burden, but their training is very expensive and also time-consuming. Therefore medical scribes are not a solution to the problem.
- Medical scribes raised the idea of digital scribes. There is a lot of potential in developing digital scribes, but from a technical perspective it is quite complex. Digital scribes combine ASR and NLP technologies and both are crucial to the success of the software, For this reason it is important to understand both technologies from a design perspective.
- To bring digital scribes into the consultation context, the value of recording and transcribing consultations (RQ1) for all involved stakeholders was important to investigate during the research.

2a. Valuable: What value can be delivered to hospitals, clinicians and patients by recording and transcribing consultations?

- Hospitals: less time spent on documentation, increased clinician performance, quality assurance, improved safety of care, better health outcomes, research purposes
- Clinicians: less time spent on documentation, increased physician satisfaction, reduced burnout rates, more time for empathising with the patient, teaching and learning opportunities
- Patients: more time in the consultation, ability to re-listen the exchanged information and better prepare for the next consultation, share the recording with family

2b. Trusted: What are the key fears and obstacles for physicians and patients if consultations will be recorded and how can we manage those fears?

- two-sided consent -> normalise recording on an institution level
- reason for recording -> introduce the benefits in order to get their consent to the recording
- confidentiality -> ensure that the audio recording and transcription is not used inappropriately
- data safety -> use a protected database and certify the technology

Image: Consultation room

Chapter 3

Translating the technology

In this chapter the enabling technologies are interpreted from a design perspective. Both ASR and NLP come with design implications that can inform the project. In terms of the interaction, the collaboration between the user and the AI algorithms need to be considered. Theory from human-computer interaction is summarised for the user experience as well as service design.

Image source: Pexels

Design implications of the enabling technologies

To facilitate the interaction between the digital scribe and the clinician, design implications of the enabling technologies should be understood in terms of human-computer interaction (figure 17). According to Jeblee (2019 et al.), automated note-taking has the potential to save clinicians valuable time and allow them to focus on interacting with their patients. Replacing typing directly into the EHR with intrusive commands provides a challenge to design an interface that facilitates interacting with the system and remains non-intrusive to the clinician-patient relationship.

Figure 17: Technology aspect of the future scenario

3.1 ASR

In the context of the project, hip arthrosis consultations typically involve an orthopedic surgeon and a patient. However, it could also happen that a family member escorts the patient to the consultation. The number of people present in the consultation informs the number of the microphones necessary to perform the speaker diarization step. Taking into account each and every speaker's voice characteristics is important to correctly assign labels to the speech segments with ASR.

There are a number of factors that can influence speech recognition such as natural variations of speech, speaking style, speaking rate, emotional state, and emphasis on certain words (Benzeghiba et al., 2006). In a difficult consultation, emotional state can become an influencing factor to expect. In terms of terminology, clinicians and patients might describe the same problem with different words therefore training the ASR on clinical conversations is essential. Furthermore, speech characteristics such as accents, dialects or foreign words influence the accuracy of speech recognition (Bent & Frush Holt, 2013). Environmental and background noise can also interfere with the success of speech recognition, but typically in a closed consultation room setting this aspect can be ignored.

Speech is the most natural form of human communication. Interacting with systems can also be enabled by using voice only with voice user interfaces (VUI). Users can control such interfaces by using a 'wake word' or voice commands. Voice commands can free up hands and eyes for other tasks and coupled to software features if the speech is processed real-time.

3.2 NLP

The transcribed speech is further processed by using Natural Language Processing models which is a subdomain of machine learning. ML models advanced application of conversational user interfaces which provides an opportunity to switch to a speech-based interface. In human-computer interaction, NLP has the potential to effectively complement other available modalities, such as windows, icons, menus, and pointing (Manaris, 2008) where modality is defined as a communication channel used to convey or acquire information. (Coutaz and Caelen, 1991). Cohen et al. (2013) defines speech as input modality for 'command and control' of systems or devices. After the user input is recognised by ASR, natural language understanding (NLU) components can identify user intents or commands (Clark et al., 2019).

VUIs need to know when a user starts speaking as well as when the speech has ended. Knowing when the user stops speaking is known as a timeout (Pearl, 2016). Timeouts allow the system to pause for some time before continuing recording the conversation which can be configured by ASR. For the interface, visibility of system status is important according to Nielsen's usability heuristics (1994). According to Clark et al. (2019) there is lack of design related research on the topic, meaning that there are currently no clear considerations or robust heuristics for developing user centred speech interactions. As Corbett & Weber (2016) highlight, HCI theories may not automatically translate from one context of interaction to another therefore qualitative work is encouraged to reveal UX issues within the context to use speech interfaces.

3.3 Interface level

3.3.1 Human-AI collaboration

It is argued that human-centred design is invaluable to AI system design, however it presents designers challenges to move from algorithms to systems (Beardow, Lomas & van der Maden, 2020). To facilitate the interaction between humans and the system, the collaborative aspect in-between needs to be carefully looked at to design the user experience (UX). Connecting AI to UX design is of increasing interest in the human-computer interaction (HCI) research community. Designing with AI is proven to be difficult since the dynamic nature of the technology can give unpredictable errors that impact the intended user experience or even lead to undesired societal impact (Yang et al., 2020). To aid designers in the process, Amershi et al. (2018) proposes guidelines to take into account when designing AI systems. The relevant guidelines are summarised in figure 18.

In future doctor-AI collaborative consultations, Kocaballi et al (2020) identified the core tasks of doctors namely clinical reasoning, human communication, embodied experience of interacting with a doctor and empathy. For this reason, patients will always require human interaction and empathy according to the results of a nationwide survey conducted among members of the Dutch Association of Surgery (Voskens et al., 2022). Designing for a collaborative interaction leads to the concept of partial automation where clinicians have control over the digital scribe.

To facilitate effective human interaction with fallible AI models, transparency (figure 19) is key as well as showing the confidence of each prediction to the user. Transparency refers to clearly communicating the level of certainty of the model which helps users in decision-making (Nguyen et al., 2018). Applications of Human-AI collaboration in complex domains, such as healthcare, require the user's ability to understand and predict agent behaviour (Miller, 2018). In the case of mixed-initiative documentation, the interface should be designed with the assumption that users may often wish to complete or refine an analysis provided by an agent for less bias (Malhi et al., 2020).

G7	**Support efficient invocation.** Make it easy to invoke or request the AI system's services when needed.
G8	**Support efficient dismissal.** Make it easy to dismiss or ignore undesired AI system services.
G9	**Support efficient correction.** Make it easy to edit, refine, or recover when the AI system is wrong.

Figure 18: Relevant design guidelines for Human-AI collaboration (Amershi et al. 2018)

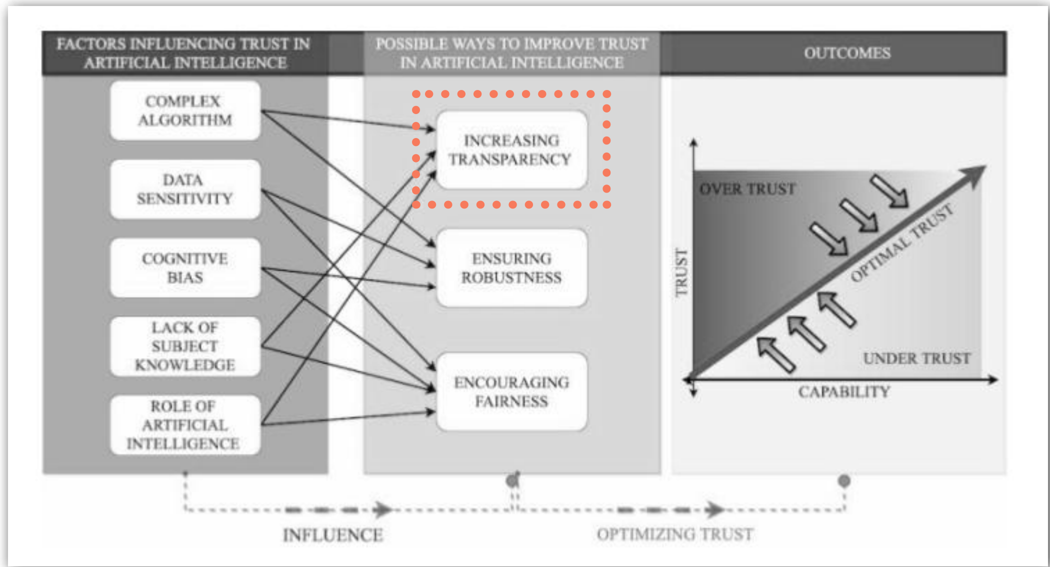


Figure 19: Transparency as a possible way to improve trust in AI (Asan et al. 2020)

3.3.1.1 Automation bias

Automation bias happens when clinicians incorrectly follow instructions from technology, and is an important new cause of clinical error (Goddard et al., 2012). It holds similar risks that clinicians will automatically accept scribe suggestions or complete documents without checking (Coeira et al., 2018). Trust is possibly the strongest driving factor in over-reliance, when trust is incorrectly calibrated against system reliability (Goddard et al., 2012). Number of factors influence trust such as the user's own experience with the system or the extrinsic source such as the reputation of the system technology in the user's social circle (Han, 2013). To mitigate automation bias, the notion of optimal trust is introduced by Asan et al. (2020) by advocating for fairness, transparency, and robustness in the system.

The European Commission set up a framework for Trustworthy AI, in which exposures are defined as ethical, technical, and legal considerations related to the use of the AI system." (Lemonne, 2018) The framework proposes seven requirements to self-assess trustworthy AI systems in terms of 1) Human agency and oversight, 2) Technical robustness and safety, 3) Privacy and data governance, 4) Transparency, 5) Diversity, non-discrimination and fairness, 6) Societal and environmental wellbeing, 7) Accountability. However, Zicari et al. (2021) argues that the framework offers a static checklist and does not provide specific guidelines during design phases.

With the advent of Europe's General Data Protection Regulation (GDPR), other issues also come into play. For instance, Article 22 states that "the data subject shall have the right not to be subject to a decision based solely on automated processing". A digital scribe by definition would thus always require a clinician to sign off on the final document, and patients might need to explicitly consent to have their record created in such a way (Coeira et al., 2018).

3.3.2 Intelligent documentation systems

Coiera et al. (2018) also states that there is little consensus on the core features that should constitute a digital scribe, since there has been little exploration on the topic. In a hospital, clinicians spend almost half of their day interacting with EHR system interfaces while carrying out documentation activities. It is evident that EHRs and decision support systems provide value to clinicians to carry out documentation tasks (Schnipper et al., 2008). However the main challenge is to simplify these tasks with intelligent documentation systems. Currently, much of a patient's story is captured in non-standardized narrative format which leads to highly unstructured clinical notes (Rosenbloom et al., 2011). An enabler is provided by Schnipper stating that requesting structured, and coded data entry could help. Structured entry systems may allow users to create templates to maximise data structure and completeness (Rosenbloom et al., 2011). Structured clinical data can also be linked to knowledge of evidence-based medicine and queried at the point of care (Schnipper et al., 2008).

Template-based systems can facilitate clinical documentation and the ordering of procedures (Henry et al., 1998). To configure the templates, Gardner and Pearce (2013) suggest sections to include the patient's medical history, family history, physical examination, the diagnosis, planning, and prescriptions to be listed in a tabular format on the computer screen. Several studies conclude that clinicians prefer narrative charting because it fits individual documentation styles (Gardner and Pearce, 2013) and flexible note taking (Rosenbloom et al., 2011).

Patient identifiers (i.e. name, medical record number, date of birth, physician name, and date of service) are needed for correctly processing inbound referrals, physician orders, clinical notes, and medical records requests (Mathioudakis et al., 2016). Extracting this information allows for intelligent indexing of incoming documents so they can be linked to the appropriate patient record in the EHR. Machine learning has the potential to securely and privately read the document, and analyse the content to speed up clinical and administrative processes. (Larkin, 2020).

Document automation also offers the possibility to provide automated suggestions*. The suggestions can be potentially extracted from the processed conversation as text input in the corresponding template sections. However, the challenge is to design the interface in a way that the clinician stays critical about the suggestions while displaying system transparency.

*the term **automated suggestions** will be used often later in the process

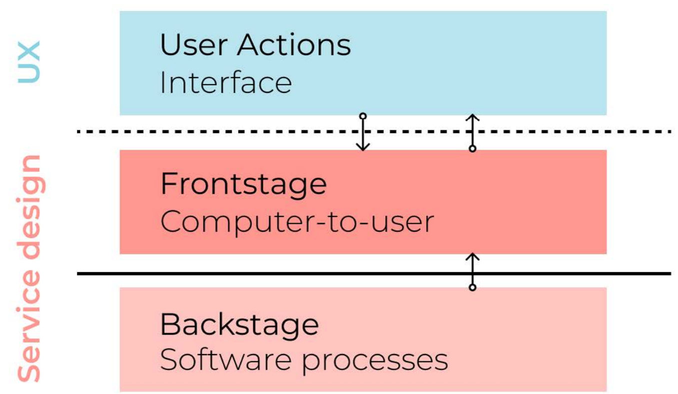


Figure 20: Experience levels

3.4 Workflow level

Literature mentions concerns related to a digital scribe's clinical utility, such as the effect on a physician's workflow (van Buchem et al., 2021). They are willing to consider IT solutions to support clinical practice but require that it fits into their workflow (Schnipper et al., 2008). To figure out the needs, qualitative research into clinician's daily workflow is encouraged as part of product development. Designing for the workflow is a holistic level of the entire user experience throughout clinical documentation. The term user experience (UX) is used for designing the experience on the interface, while the interface could also connect service touchpoints over time (figure 20). State of the art service design often embeds technologies throughout digital channels in service solutions (Jylkäs et al., 2019).

3.4.1 AI-enabled service design

In the case of the digital scribe, the core value is in the advanced technology behind the software. The recorded conversation can be referred to as a data source which feeds the natural language processing models in order to speed up clinical documentation. Processing power is increasingly utilised for automating mundane tasks and processes (Jylkäs et al., 2019). As products are becoming more and more data-driven, there is a trend to employ AI to personalise service content (Reavie, 2018). AI has the potential to become an enabler for value creation in digital service channels (Vargo & Akaka, 2012) connecting multiple touchpoints. Digital service solutions are often offered in the form of software as a service (SaaS) that allows delivering service experiences via the Internet. For ease of access and manageability, SaaS products make use of cloud-based integrations therefore often offered as a web-application.

Chapter 3 takeaways

- To design for ASR, two microphones are needed to listen to the conversation.
- The picked up words influence the ability to automate with NLP. The processed speech can potentially be coupled to commands in the software to control the interface.
- Since there is no checklist of what features a digital scribe should have, qualitative work is encouraged in the context.
- A collaborative scenario should be designed, where we automate what is possible (partial automation) but the clinician has ultimate control over the interface. Guidelines should be taken into account to design for human-AI collaboration.
- Trust towards the system will build up over time, and can not be quantified. An enabler for trust is system transparency.
- With a template-based system, the notes can become structured and lead to improved data quality.
- Many papers question the clinical utility of the digital scribe if it does not fit into the workflow of clinicians, therefore their daily operations need to be understood.

Chapter 4

Understanding the context

To research the context, user research was planned and separated into two activities. First, semi-structured interviews took place to understand the clinician perspective and build upon learnings from literature. Second, job shadowing was performed in two hospitals to witness clinical documentation and see the day-to-day operations of orthopedic surgeons. To synthesise the findings, user personas and journey maps were created from these activities. Two driving questions set the goal for this activity:

What are the current documentation habits of orthopedic surgeons?
How can the digital scribe aid surgeons during note taking?

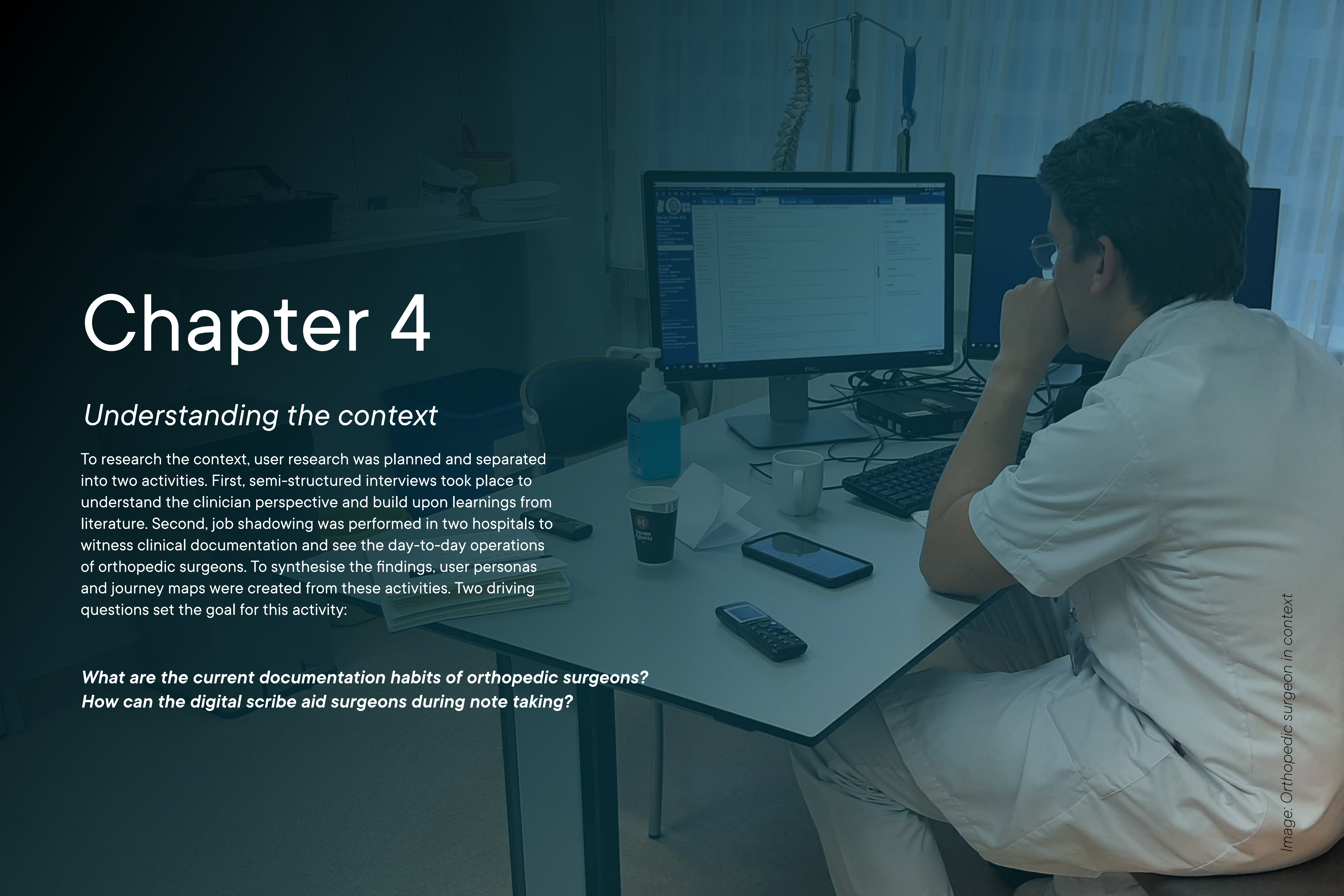


Image: Orthopedic surgeon in context

User research

4.1 User interviews

In order to gather first-hand perspectives on the implementation of the digital scribe, a set of user interviews were carried out with target end users. For the interviews a semi-structured interview guide was prepared including open-ended questions and topics informed by prior literature study (Chapter 2). These topics are namely the influence on consultation style, controlling the recording, and the utility of audio compared to transcripts. Literature also informed some assumptions as well as possible biases that could influence the inputs. For example, the level of computer literacy is an influencing factor (Sutton et al. 2020). The understanding of the involved technologies also influences their attitude towards new solutions. Also, experience level is indirectly related to age (Goddard et al. 2012).

In total, 6 orthopedic surgeons were interviewed and two non-orthopedic clinicians for practical reasons. The interview participants were recruited via the network of the company, which could also lead to a generally positive attitude towards the solution. All interview participants were informed about their voluntary participation in the study beforehand and presented with a consent form so that inputs can be quoted in this thesis report.

The interviews were conducted in a 1-on-1 setting, either in person or remotely. The recruitment was possible via the collaborating surgeons of the company. The interviews were audio recorded upon approval and transcribed. The transcriptions were analysed using Atlas.ti using qualitative labelling. In the software, it is possible to add co-occurring codes which also show relationships between clusters. During the labelling, 10 clusters were identified as a result of the analysis. In figure 21, an overview of the clusters are displayed. For each cluster, the findings are first summarised as an explanation of the cluster and followed by some representative quotes from the interviews.

Cluster
Frustrations
Focused attention
Documentation style
Utility of the transcription over the audio recording
Behaviour and interaction during consultation
Implications for intelligence
Concerns about language use
Fear of control
Indifferences
Trust towards the system

Figure 21: Clusters from interview analysis

4.1.1 Insights

Frustrations

Clinicians have numerous frustrations in their current workflow and interaction with the current EHR system. Due to the time pressure, they typically need to start typing the documentation during the consultation already which leads to having to typing blindly that could lead to typing errors.

"I type blindly [during the consultation] not 10 fingers, but I think five or something."

"Sometimes I notice after a couple of seconds that my hands are a little bit moved off to the right or left and then it's completely jibberish what I'm typing."

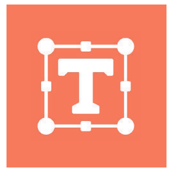
Focused attention

The mentioned frustrations also influence the behaviour of the clinician as well as their interaction with the patient during consultations. It is clear that they value interacting carefully with the patient, but the computer stands between the clinician and the patient. Many questions of the conversation are repeated in each consultation which adds to the value of standardising the anamnesis. Having to ask the same set of questions does not leave room for the clinician sometimes to ask deeper questions to the patient.

"Most of the time with a certain complaint, I ask the same questions. And so that's not really beneficial for the consultation, [...] if you get more space inside the consultation, you can move focus on what's really the question behind the question."

"I think it would be very helpful because I am able to type and talk and I think my focus will be greater if I don't have to type. So I only have to correct it or check it afterwards."

Documentation style



Every clinician has their own style of documentation and writing up their notes. Due to the repetitiveness and time pressure, many clinicians use abbreviations in their notes for example PT (physiotherapist), LI (links =left), RE (rechts =right), A/ (anamnesis), O/ (onderzoek = research). All clinicians are aware of the commonly used abbreviations. Often they have to paraphrase their medical advice in simple terms to patients, but they have to do documentation in medical terms.

Utility of the transcription over the audio recording

If we start recording and transcribing clinical conversations, there will be both an audio recording and a transcription of the conversation. The two contain the same information, but they might provide different value to both the clinicians and patients. It was highlighted by multiple clinicians that the fact that you can search in the transcription file would be valuable. However, they would not have time to go through the entire transcription, and definitely not for listening back to the audio.

"And again, I can imagine that you use the transcription then you can just search in the transcription file, whether you already discussed it with the patients."

Behaviour and interaction during consultation

The need to carry out documentation during the consultation also hinders the behaviour as well as the interaction with the patient. The computer is between the clinician and the patient, requiring the clinician to multitask between talking and typing which divides their attention.

"I think that's how the computer really stands literally between me and my patients. And I'm looking at the screen. ... and that's why I find it interesting that I think I can add value to my consultations."

"... during the conversation I'm typing, and I'm looking at the screen instead of the patient during the consultation."

Implications for intelligence

Clinicians wish to design the system such that it is self-learning. Also, the system could notice differences between behaviours, if for example the same task is done slower compared to other times, the system could sort of 'check on' the clinician to avoid error-prone notes. The preference was also raised to let the algorithm adapt to individual changes so that it can accommodate individual documentation styles.

"My ultimate goal would be that every clinician has a personal algorithm just during the whole process but when you're doing an outpatient clinic, that algorithm you use is trained by adaptations you made.. because your if you have a different doctors, you will also get subtle changes in the way they talk or what they find more important."

Image: UMCG

Concerns about language use

Next to opportunities, naturally there are also concerns towards the system. Multiple clinicians have concerns to what extent the system can work for different clinicians if they use different words for the same topic. Also, the fact that it is a computerised system makes clinicians worried that not all scenarios are pre-programmed in the system.

"So that's a concern of mine that it's not optimal for things that are just outside of what is known to the computer. You know, if I say something slightly different or use another word then it doesn't fill in anything"

Fear of control

In addition to concerns, there are also fears. Not having control over what goes into the EHR was mentioned, which should be carefully considered in a hybrid scenario. Also, spreading the data inappropriately i.e. on the internet was a potential fear. Also, as any human being, clinicians also have better and worse days which might influence their consultations. In case a consultation is recorded that they are less confident about, it might make them feel uncomfortable.

Indifferences

While some clinicians have fears, others feel indifferent towards recording their consultations. Being indifferent was true for multiple interviewees, which could be influenced by personal characteristics or level of experience. Clinicians are motivated to deliver the best care, and are professional to focus on their work. If they focus on their work, they are confident that the recorded data will also reflect the quality of care they deliver.

"It's an interesting question. I don't know. I think that if you do what you do well and do it conscientiously then I would let you do it [recording the consultation] without any problem"

Trust towards the system

Trust towards the system will be built up over time. If the product has clear benefits for the clinicians, they are open to it and could get used to it. They will trust the system if it proves itself over time, or a statistical overview of success metrics would encourage them to implement it.

"If It proves itself over and over again. Okay, I have to have successful experiences a couple of times, and then I would trust it."

4.2 Job shadowing

To understand the context, job shadowing was performed both in an academic and a non-academic hospital in the Netherlands. Shadowing is a qualitative research method where the researcher follows the end user throughout a day to see and learn from their day-to-day operations. Observing reality can lead to a better understanding of existing behaviours in order to design a product experience. The research question was the following:

RQ1. What is the full cycle of clinical documentation in hip arthrosis consultations and how can it be automated using NLP technologies?

1a. Which documentation tasks in hip arthrosis are time consuming/ frustrating/sensitive to errors and technically feasible to automate?

In terms of procedure, the shadowing was arranged taking into account the surgeon's availability and schedule. Upon agreement, the plan was to follow one surgeon throughout his entire shift and just observe without interfering. At the end of the shift, questions were raised in a follow-up interview. From a relationship perspective, it was also interesting to pay attention to the interaction between the clinician and the patient and note how documentation affects it. Mainly the clinician's emotions are relevant for the journeys in this project. The detailed plan for job shadowing can be found in Appendix C.

As a preparation, the observation was planned in advance to make sure it will go smoothly. The main objective was to identify the documentation tasks that the clinician does during a typical shift. Also, the phases of a consultation were important to recognize as well as measure how much time is spent on each phase. From the shadowing, the plan was to develop a better understanding of the end user and also to map out his observed experience during the full shift and from a typical consultation. From

Job shadowing was performed in two hospitals: at the University Medical Center Groningen (UMCG) and at the Elisabeth-TweeSteden Ziekenhuis (ETZ). At each hospital, one day was spent following the same shadowing plan. From the shadowing activity, differences were identified between academic and a non-academic hospital. The next table (figure 22) highlights these differences.

umcg etz

	Academic hospital	*Non-academic hospital*
Length of consultation	10-30 minutes	5-10 minutes
Number of patients in a shift	around 10	around 30
Patient cases	complex cases	simple cases
Homogeneity	not true for complex cases	true
Flow of conversation	depending on case	typical order
Doc. after consult	there could be time	no time for it
Pressure on time schedule	low	high
Pressure on time schedule	low	high
User persona	*The balancer*	*The multitasker*

Figure 22: Differences between academic and non-academic hospitals

4.2.1 Academic hospital

In an academic hospital, there are about 10-15 consultations scheduled during a half day shift. Here typically more complex patient cases are presented therefore the consultations can take longer varying between 10-30 minutes. The complexity of the case also influences the flow of the conversation. Consultations in an academic hospital are less homogeneous therefore more challenging to automate with NLP. Here an orthopedic surgeon might mostly do documentation tasks after the consultation (figure 23) which could take almost the same time as the consultation itself (46%). Since more time passes from the conversation until documenting the exact content, it is possible to forget what to document in the notes which is a challenge for the clinician.

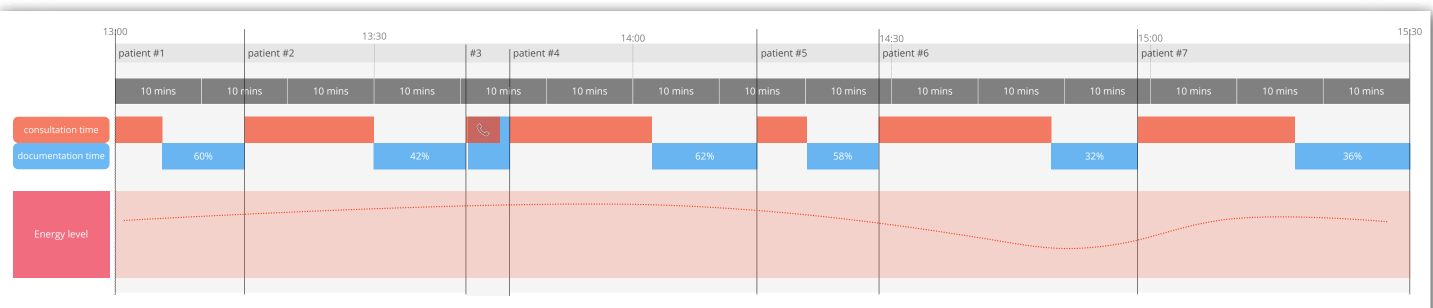


Figure 23: Length of consultations in an academic hospital

4.2.2 Non-academic hospital

In a non-academic hospital, it is usual to have around 30 consultations scheduled in a day at an outpatient clinic. These patient cases are more simple, therefore the consultations are shorter (5-10 minutes). The flow of the conversation follows a typical order of questioning which makes these consultations much more homogeneous. In this context, it is not possible to not start on documentation tasks during the consultation time. There is a very high pressure on time and a strict schedule to follow. If for any reason there is a delay in time, the surgeon needs to finish the documentation tasks after the entire shift.

In both contexts, the documentation tasks are the same for orthopedic surgeons. Since less complex cases are typically presented in a non-academic hospital, the conversations are more structured in those consultations. For this reason the more structured conversations are a better use case for the digital scribe to automate with NLP. Last but not least, since there are many more non-academic hospitals in the Netherlands, there is also bigger market potential for the Assistant.

4.3 Synthesis

4.3.1 Personas

Based on the user research activities, two types of clinicians were identified that were translated into user personas. Personas are a representation of the target user profile which summarises the key characteristics, motivations and frustrations of the end user of the product. The two types of user personas were named as The multitasker and The balancer (figure 24).

The multitasker	Non-academic hospital	does documentation mostly <u>during</u> consultation
The balancer	Academic hospital	does documentation mostly <u>after</u> the consultation

Figure 24: User personas

The names were inspired by the nature of the repetitive documentation cycle that requires orthopedic surgeons to juggle between a variety of tasks. The key difference is whether the clinician does the documentation during or after the consultation. The balancer mostly does documentation after the consultation that typically takes almost as much time as the consultation itself thus trying to balance the time between caring for the patient and documentation. In comparison, The multitasker mostly does documentation during the consultation while constantly multitasking between talking to the patient and note taking.

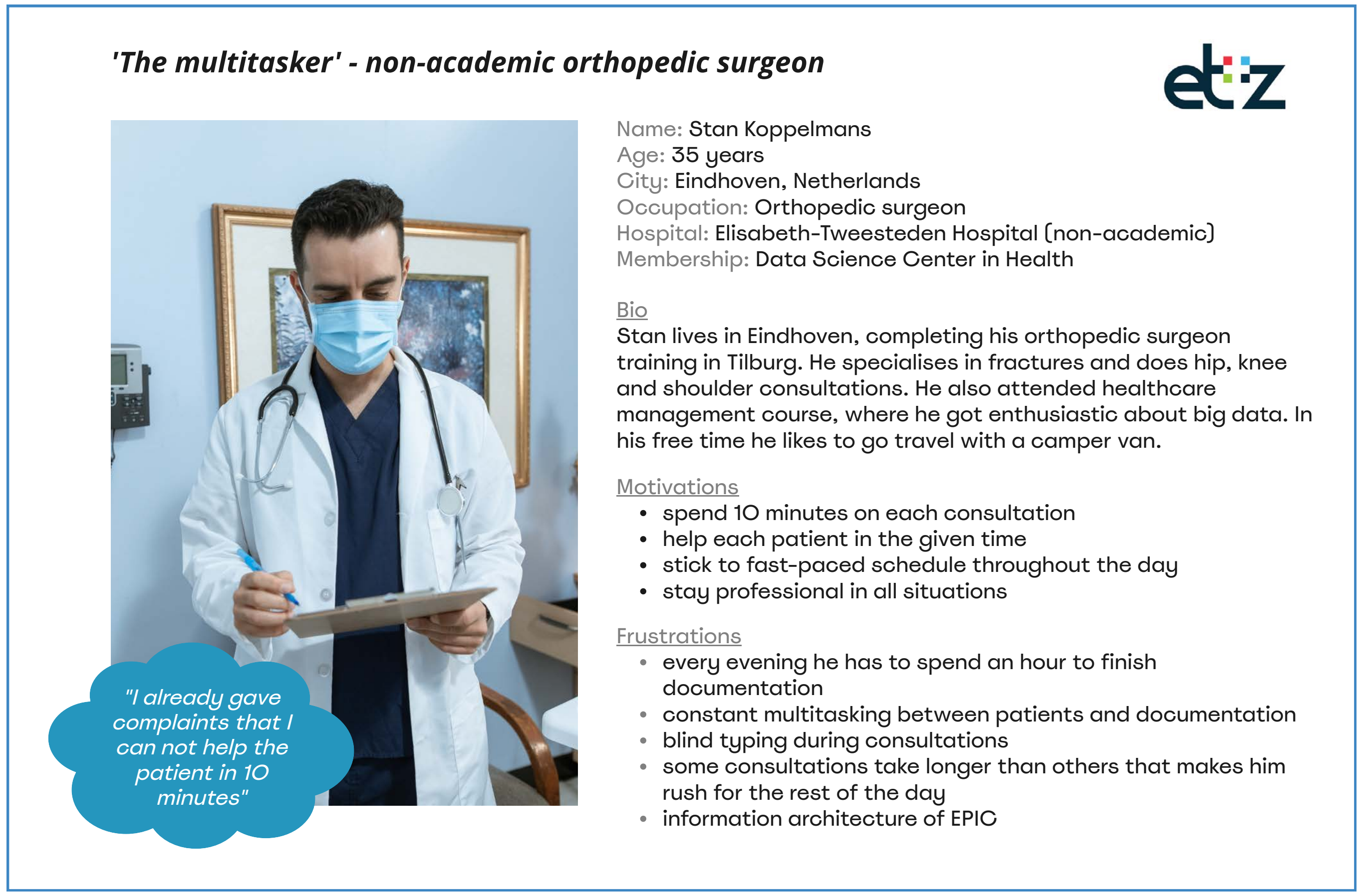


Figure 25: The multitasker

The multitasker (figure 25) typically works in a non-academic hospital responsible for both orthopedic surgery and seeing those patients in an outpatient clinic. He gets 10 minutes scheduled for every single consultation according to schedule. He is motivated to help each patient in the given time but already gave complaints to the hospital that it is almost impossible to help the patient in such a short time. He does his best to stick to a fast-paced schedule throughout the day and stay professional in all situations.

Patients are human beings therefore all are different in a way. Since some patients require more time than others, it is likely that there is not enough time to finish documentation for each. Therefore every evening he has to spend an hour to finish documentation, which is one of his main frustrations. Due to the time pressure sometimes he has to type blindly. He is really frustrated with the information architecture of EPIC.

The balancer is an orthopedic surgeon who typically works at an academic hospital. The balancer persona can be found in Appendix F. In an academic hospital there is less stress on the schedule, he can take the time to answer all individual questions of each patient. That means that he usually does documentation tasks in-between consultations. He is motivated by perfecting every 'Conclusion' of the notes. In an academic hospital a diverse range of patients occur from simple to complex cases, which means it is different to empathise with all of them.

The balancer is frustrated by the constant switching between the notes and the Electronic Patient Dossier. Also, he is irritated by the fact that he has to search for the MRI scan of each patient while it is relevant to show in every single consultation. He needs to send each referral separately in the EHR for which he is always copy-pasting text from the notes.

The Assistant could be of value to both types of clinicians, but the implications are different (figure 26). For The Multitasker, the speed for documentation needs to be much higher. For him, the probability of making errors is high due to time pressure and blind typing. For The Balancer, there is a chance that he forgets content to document by the time he starts doing documentation since more time passes. For both, it is very challenging to improve documentation at the end of the shift.

	The multitasker	*The balancer*
speed for documentation	very fast	mid
probability of making errors	high	mid
reason for making errors	multitasking (talking + typing)	more time passed since conversation
probability of forgetting documentation content	mid/high	mid
chance to improve documentation at the end of the shift	impossible	possible but error-prone

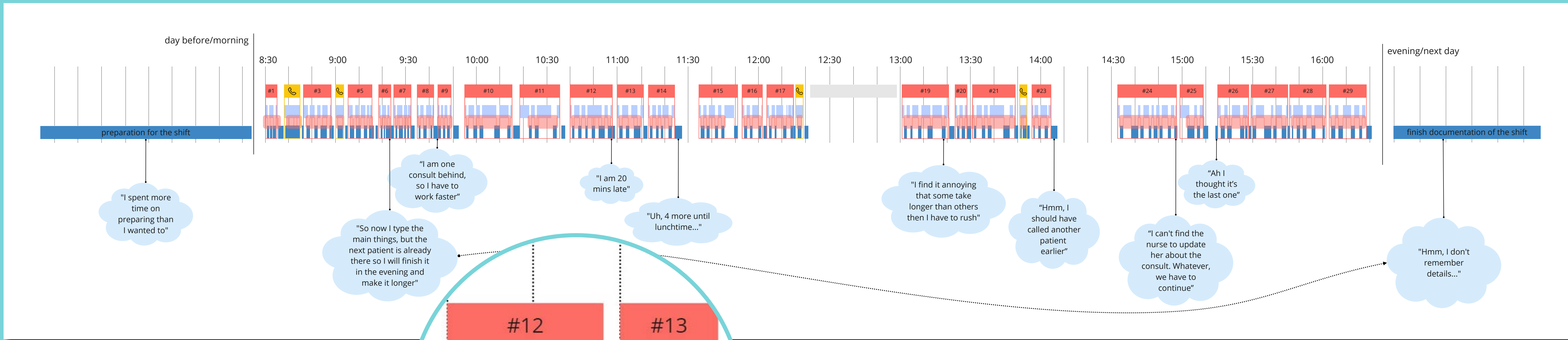
Figure 26: Implications of the solution for both user personas

4.3.2 Journey maps

From the job shadowing activities, two journey maps were made: one for the entire day and one for a typical consultation cycle (figure 27). The shadowing activity made it possible to understand existing behaviours of the end user and journey maps visualise the observed user experience.

User experience level	**Journey method**
Consultation cycle	Consultation journey
Workflow of a day	Transitional journey map

Figure 27: Journey mapping methods



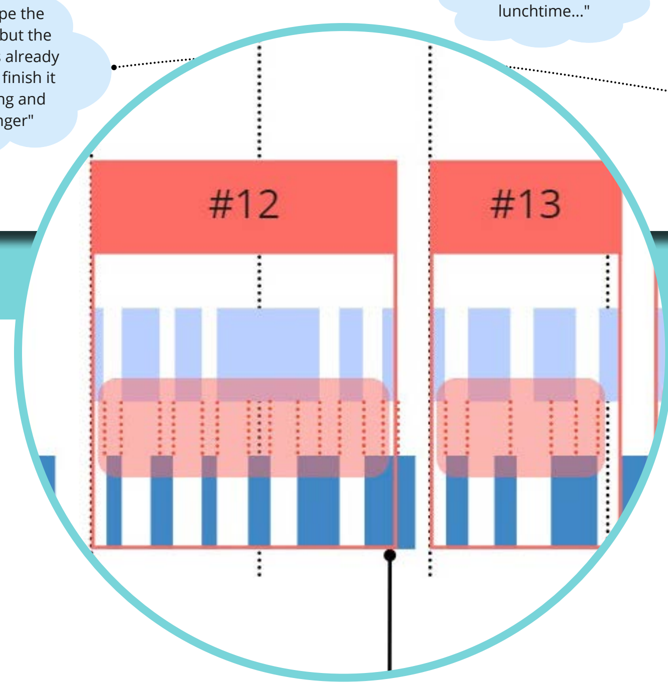


Figure 28: Transitional journey map (non-academic hospital)

4.3.2.1. Workflow of a day

From the job shadowing activities, two journey maps were made: one for the entire day (figure 28). and one for a typical consultation cycle The shadowing activity made it possible to understand existing behaviours of the end user and journey maps visualise the observed user experience.

Figure 29: Legend to interpret the journey

The transitional journey map is a visual representation of a typical day of an orthopedic surgeon in an outpatient clinic scheduled with ~30 consultations. In figure 29, a legend for interpreting the map is shown.

Before a working day, there is already time spent on preparing a day at the outpatient clinic. On a typical working day, consultations are scheduled between 8 and 5. The consultations are mapped out throughout the time, of which some of them were phone consultations. For each consultation, the time snippets spent on talking to the patient and spent on documentation is visualised. There are constant transitions from one activity to another. The transitions and constant switching lead to divided attention that could also lead to error-prone note typing. The multitasking both pressures and tires the clinician.

Due to the time pressure, starting on documentation is inevitable during the conversation already. The documentation is focused on the main points and formed by the clinician's preferences and documentation style. Typically, physical examination takes the most time during consultations which is more relevant for new patients (patient intakes) or if the patient state worsened. Overall, the documentation cycle is very much repetitive throughout each consultation.

Between the consultations some time is always spent on finalising notes of the consult, while already preparing for the upcoming consultation. If one consultation takes longer due to the patient being more difficult, that requires more attention and all the next consultations are delayed, which stresses the clinician to speed up.

> *"So now I type the main things, but the next patient is already there so I will finish it in the evening and make it longer..."*

If making up for the delay is not possible, the surgeon needs to finish doing the documentation in the evening which adds to burnout rates and reduces physician satisfaction. The quotes give contextual understanding of the workflow of an orthopedic surgeon, that was collected via asking questions in the meantime.

4.3.2.2. The consultation

Figure 30: Legend to interpret the journey

After analysing the workflow of an orthopedic surgeon, it is time to zoom into the consultation room. A journey map for a consultation cycle was created (figure 30 and 31). The cycle of documentation starts before the patient enters with a quick preparation of the patient file. In this step, the surgeon looks back on the conclusion of the previous consultation and checks on any update.

Typically each consultation starts with a general check to find out the current state of the patient and the reason for visiting. Next, during the evaluation phase the surgeon checks if the treatment is working (in case of recurring patients) with more focused questions. The patient explains the complaints and pain (s)he is experiencing and gives an update about physiotherapy if applicable. In case of pain, then comes the physical examination phase where both the surgeon and the patient move to the examination table. Here the surgeon assesses the range of motion and the patient's capacity to move and find the source of pain. The physical examination takes the most time of a consultation which needs to be accurately documented, but in the meantime it is impossible for the clinician to take notes. In the last phase, the surgeon formulates the diagnosis and responds to the patient's individual questions.

"I always struggle with this because I really want the conclusion to be perfect."

After the patient leaves the consultation room, the cycle of documentation continues since the notes need to be formulated and finalised. The 'Conclusie' is the most important part of the notes and can be reused for follow up documentation tasks such as sending a letter to the GP or placing an order for a lab.

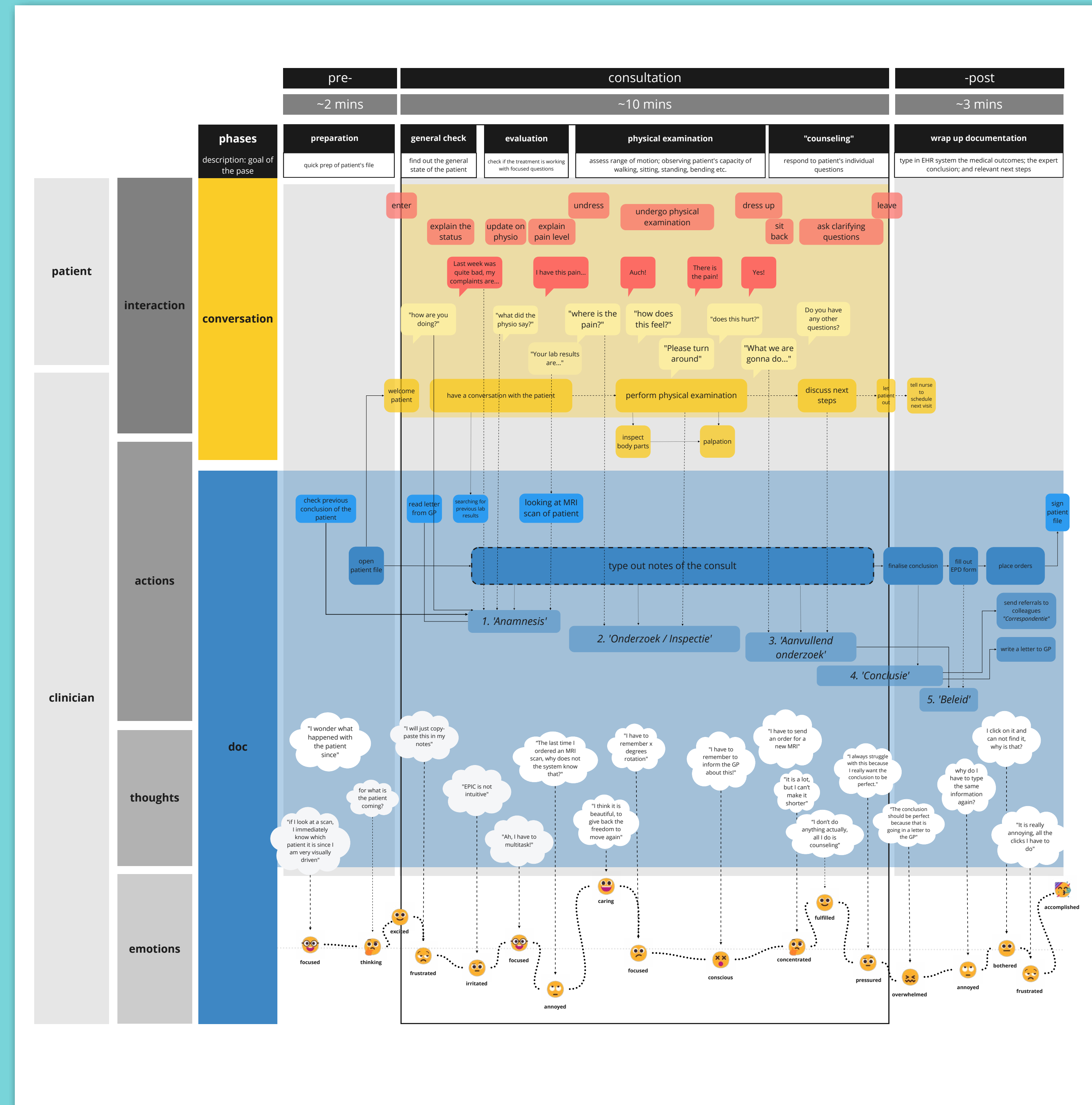


Figure 31: Consultation journey (non-academic hospital)

Opportunities

RQ1. How can the full cycle of clinical documentation in hip arthrosis consultations be automated using NLP technologies?

1a. Which documentation tasks in hip arthrosis are time consuming/ frustrating/sensitive to errors and technically feasible to automate?

The full cycle is illustrated in the journey map. From the experience line, the negative emotions can imply opportunity areas to improve with the Assistant. For example, reading a letter from the GP causes frustration. Some part of the content is often reused in the notes of the clinician. Since that is text, that can be also utilised as a data source to help the clinician. Also, during physical examination there are a lot of details to remember which leads to frustration. Physical examination is the longest part of the consultation during which it is impossible to type notes. The intuitiveness of the EHR leads to annoyance which provides an opportunity to improve the user experience. Follow up documentation tasks can also be automated by picking out certain words from the conversation. However to automate that, an integration with the EHR will be required.

Image: Context

Chapter 4 takeaways

- Both interviewing and shadowing helped to empathise with the end user.
- Qualitative insights made me understand that each clinician has their own documentation style which they are also proud of. It will be a challenge to design an interface that fits all the needs and standardise data entry.
- Journey mapping helped to identify the phases of the consultation and see how they are connected to the sections of the notes.
- It was an unexpected finding that there are differences between academic and non-academic hospitals. Due to the homogeneity of the conversation, together with the company we decided to focus on the non-academic context in the project.
- Shadowing in a non-academic hospital showed the very repetitive routine to help by developing the Assistant.
- The multitasker persona has a very high-paced schedule and many frustrations that could potentially be improved.

Chapter 5

From research to design

This chapter bridges the research phase with the design phase. The vision of the product is introduced to scope the design space. The product roadmap for the Assistant is revealed in detail with the implementation steps and corresponding value propositions. In light of that, the design goal is defined and a concluded list of user requirements is presented.

Image source: Pexels

Vision of the Assistant

To win time with the Assistant, we should try to automate what is possible but also provide the flexibility to edit the notes. In order to automate note taking, the text entry needs to be structured through the anamnesis standard (figure 32). Structured data enables partial automation and can be organised into templates.

By introducing the Assistant, the value proposition has multiple pillars. On one hand, the Assistant aims to provide a much better user experience than how typing is currently done in the EHR. As of today, typing free text in the EHR is limited to that person and time and is not stored in a transferable data format. Without structuring the data entry, documentation tasks can not be automated. Through the Assistant, we also standardise structuring data entry through design. That is why the design phase focuses on designing for template completion where the data input is as structured as possible but free text input always remains possible. The Assistant will be a web-application since in the beginning of the implementation it is risky for EHR vendors to integrate third party products.

The Attendi Assistant will offer a part of the total data collection: the consultation anamnesis. For the anamnesis, templates are built by clinicians in their specific domain and will be reused by all users. In the future, the Assistant will expand its offering, and gather data throughout the entire patient journey pre- and post-consultation.

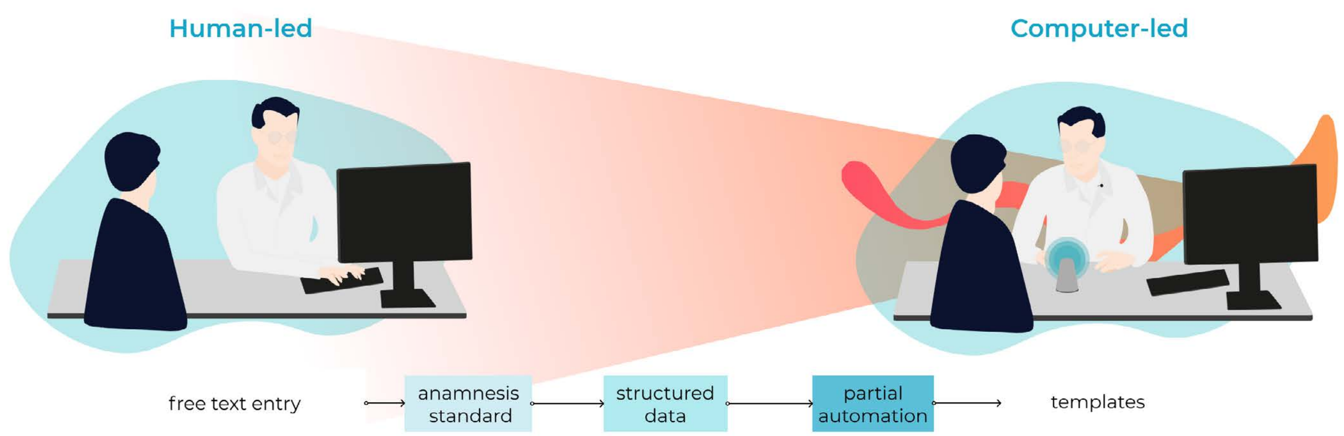


Figure 32: Data structuring through the Assistant

5.1 Product roadmap

A roadmap for implementing the digital scribe was provided by the company (figure 33). The Assistant is planned to be implemented in hospitals in 3 steps: 1) establishing a recording infrastructure 2) providing an application for template completion 3) scaling the template options and finalising integration with the EHR system. Each step leads to a value proposition and product features. To support the value proposition, the design phase was informed by extracting UX features from Step 1 and Step 2 (Figure 33). so that they are considered.

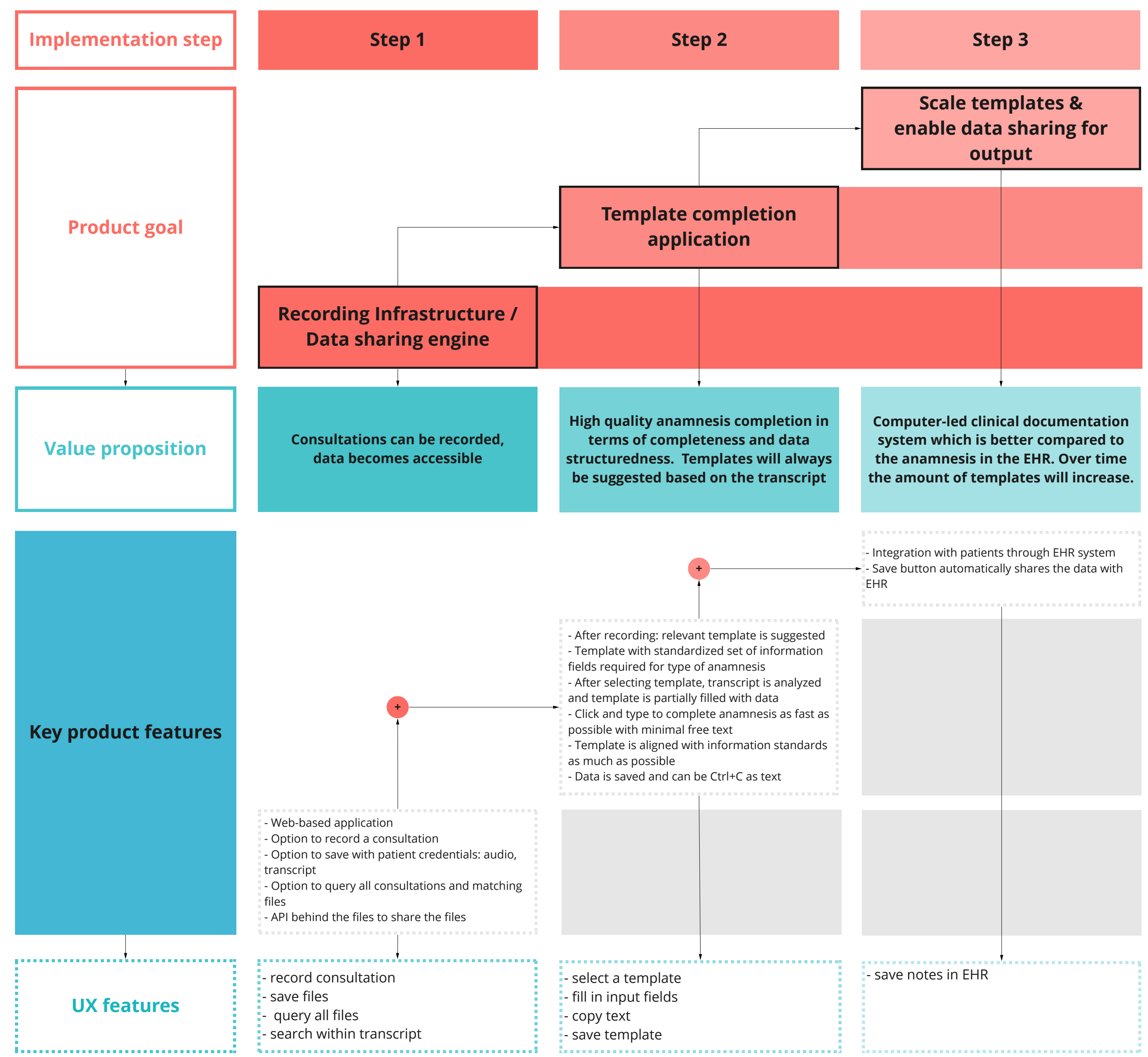


Figure 33: Product roadmap based on company input

5.1.1 Design goal

Partial automation also implies that the clinician has full control over the text entry. That is how the theory of the collaborative aspect between the user and the algorithms translates into designing the interface. To provide value to clinicians, automated suggestions could potentially speed up clinical documentation. Using the Assistant should allow for flexibility where the clinician can always edit the notes to fit individual documentation styles. The interface should support system transparency to increase trust towards the system. By displaying where the suggestions come from, the clinicians can also judge whether to accept, reject or edit the suggestions and add their own notes.

User requirements						
UX level	Task type	Group		#	*Requirement*	Source
UX	Note taking	General		1	*Using the assistant can be initiated from the EHR system of the hospital*	Shadowing (Participant #4, Participant #6)
				2	*Using the assistant has a minimal learning curve*	Shadowing (Participant #4, Participant #6)
				3	*Display patient credentials for each transcript: Name or patient ID [Open text field], Date & Time*	AI Design Guidelines G2: show contextually relevant information
		State of recording		4	*Visual feedback of the system listening*	Voice-user interfaces (VUIs) need to know when a user starts speaking as well as when the speech has ended (Pearl, 2016)
				5	*Visibility of system status: recording started/listening/ended*	Nielsen's usability heuristics (Nielsen, 1994)
				6	*Be able to manually start the recording*	physicians stated that they might agree to record only part of the visit, provided they could control the process (Joshi et al., 2020)
					Be able to manually pause the recording	Roadmap, user control
				7	*Be able to manually stop the recording*	Roadmap, user control
				8	*Receive warning if recording goes into a technical error*	Nielsen's usability heuristics
		STT	Transcript	9	*Be able to search in the transcript*	Quote 11: And again, I can imagine that you use the transcription then you can just search in the transcription file, whether you already discussed it with the patients. (Participant #4)
				10	*Display separately who said what according to the dialogue*	Quote 12: It's very important to see what the patient said, whether it was a day or two weeks or two years before, to see what they said then, even if it's again and repeat it and focus on where they are now. (Participant #3)
			Notes	11	*Option to export/copy/print the notes*	Roadmap
				12	*Summary to be structured: 'Anamnesis', 'Onderzoek', 'Aanvullend Onderzoek', 'Conclusie', 'Beleid'*	Shadowing (Participant #4, Participant #6)
				13	*Option to copy 'Conclusie' text*	Shadowing (Participant #4, Participant #6)
			Audio	14	*Option to listen back to the audio*	Quote 12: It's very important to see what the patient said, whether it was a day or two weeks or two years before, to see what they said then, even if it's again and repeat it and focus on where they are now. (Participant #3)
				15	*Match audio with transcript as you play the audio*	Quote 12: It's very important to see what the patient said, whether it was a day or two weeks or two years before, to see what they said then, even if it's again and repeat it and focus on where they are now. (Participant #3)
		Template	During interaction	16	*Enable the user to accept suggestions for template field completion*	AI Design Guidelines G8: Support efficient dismissal: Make it easy to edit, refine or recover when the AI system is wrong
				17	*Enable the user to accept suggestions for template field completion*	AI Design Guidelines G8: Support efficient dismissal: Make it easy to edit, refine or recover when the AI system is wrong
				18	*Enable the user to edit suggestions for template field completion*	Quote 31: So if it then goes into error, I have to do it myself anyway. So that's that's a concern of mine that it's not optimal for things that are just outside of what is known to the computer, you know, yeah. I say slightly different or use another word, that it doesn't fill in anything. (Participant #5)
				19	*Option to sign off patient file before exporting to EPD*	A digital scribe by definition would thus always require a clinician to sign off on the final document, and patients might need to explicitly consent to have their record created in such a way (Coeira et al., 2018).
			Over time	20	*Enable the system to learn from the changes the user make*	Quote 12: My ultimate goal would be that every clinician has a personal algorithm just during the whole process but when you're doing an outpatient clinic, that algorithm you use is trained by adaptations you made.
				21	*Personalize automated suggestions to each clinician*	Quote 13. because your if you have a different doctors, you will also get subtle changes in the way they talk or what they find more important.
SX	Follow-up documentation tasks			22	*Suggest actions for follow-up documentation tasks i.e. place a lab order*	Shadowing
				23	*Suggest actions for follow-up documentation tasks i.e. send a letter to GP*	Shadowing
				24	*Suggest actions for follow-up documentation tasks i.e. send a letter to colleague(s)*	Shadowing

Figure 34: List of user requirements

5.2 User requirements

To guide both design and development, a list of requirements were collected from a clinician perspective (Figure 34) as one of the deliverables. The requirements are extracted from the findings from literature research, user research and product roadmap and separated on UX and service design (SX) level. The requirements were prioritised using the MoSCoW method and first discussed with the medical lead. MoSCoW is a prioritisation technique used to reach a common understanding with stakeholders on the importance of each requirement.

Prioritisation

In the process, many requirements were influenced by technical considerations and also product strategy. For this reason, the final list was consolidated together with the product manager and the CTO of the company (figure 35). For example, the decision had to be made whether we process the speech real-time during the consultation or only at the end of the consultation. We made the decision that during the first build we will not process the conversation real-time, so voice commands from the requirements were eliminated. Also, advanced features (i.e. personalise suggestions) were deprioritized.

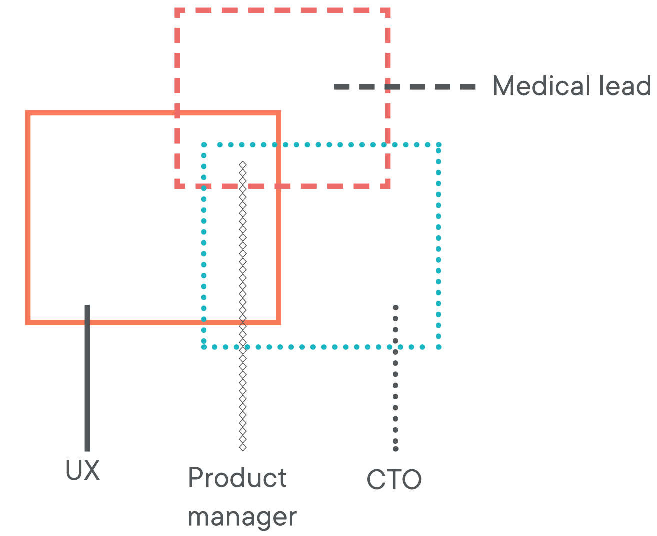


Figure 35 Collaboration

Lastly, the features that depend on EHR integration became less important in terms of priority.

5.2.1 Scope

At this stage it was important to define the scope for the prototype (figure 36)that will also shape the design process. In light of that, the prioritisation of the MoSCoW method was also influenced. To provide the most value to the company, the decision was made to design for the first build and leave out features that will only be technically feasible in some years.

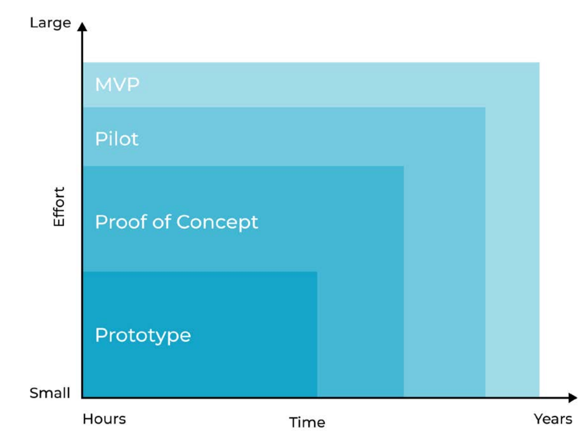


Figure 36: Product development approaches

5.3 Trends

To inspire the design phase, trends from both NLP and healthcare were considered. The trends provide knowledge for applying the technologies in the product.

5.3.1 NLP

Natural language processing is increasingly employed in the form of chatbots, text assistants and information extraction (Websfarm, 2021). There is a trend for NLP-enabled functions such as auto-complete, spell check and auto-correct. Also, NLP allows for speeding up search results so that users can get faster results. The main power of NLP lies in its ability to extract and combine information from text sources. The algorithm can use the information it gathers and also make decisions if applicable as the technology advances.

5.3.2 Healthcare

The healthcare sector is driven by four pillars to innovate: improving patient and clinician experiences, reducing the costs of care and achieving better health outcomes (Sinsky & Bodenheimer, 2014). More and more innovations pitch value propositions around patient centred care that requires an understanding of what patients value. To measure patient experiences, PREMs and PROMs (Patient Reported Experience/Outcome Measures) are increasingly utilised (Weldring & Smith, 2013). These measurements can also serve as a data source to complement the recorded consultation.

Chapter 5 takeaways

- Partial automation and flexible editing will be key aspects to pay attention to during the design phase.
- Understanding the product vision helps to translate the value in a bigger picture. The Assistant also contributes to data standardisation through structuring the notes.
- The roadmap visualises the steps that will be taken to implement the digital scribe. From each step, UX features could be extracted to support design decisions.
- The design goal was formulated in light of the user research and key concepts from theory: design the interface for a hybrid scenario where partial automation and human control is combined.
- By going through the user requirements collaboratively with the company, the strategic decisions and technical limitations were understood. To aid the company, the decision was made to design for the first steps in the roadmap.
- Reviewed trends from NLP and healthcare provided know-hows to apply the technologies into the product. For example, NLP can support auto-complete options that could aid clinicians to search in the transcript. As the product will evolve, PROMs and PREMs can be added to complement the anamnesis data.

Chapter 6

Envisioning the Assistant

To realise the interface of the Assistant, a service blueprint was made to map out both the user processes as well as the software processes based on the gathered requirements. The service blueprint serves as the envisioned user journey with the Assistant to guide both design and product development.

Image source: Pexels

Service blueprint

A service blueprint was made to visualise the user steps as well as the software processes throughout a consultation cycle. The service blueprint represents the envisioned user journey with the Assistant. With this method the relationships between service components are visualised to support communication with diverse stakeholders and understand how the software processes are related to the user steps. To highlight the experiential aspect of the service blueprint, it was combined with a storyboard for both the clinician and the patient. To read the blueprint, an illustration is provided in figure 37.

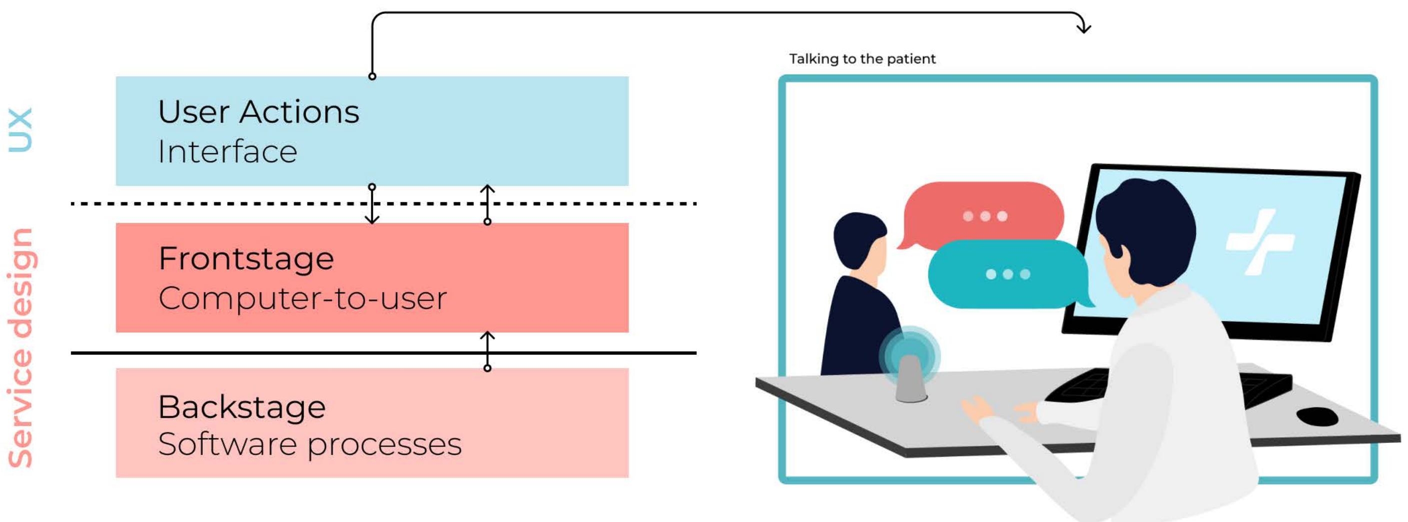


Figure 37: Simplified depiction of the service blueprint combined with storyboard

6.1 User actions

On the top, the envisioned patient journey is visualised. All efforts are for the patient therefore the key moments are highlighted. It is logical that the patient is only included in the journey throughout the consultation time. Under it a detailed clinician journey is highlighted as (s)he is the primary user of the digital scribe. The moments in the clinician journey are connected to corresponding user steps. In this service blueprint the touchpoints will take place on the interface of the Assistant, but some of them have implications (under Line of interaction) for the hardware set up. During recording, the user interactions will control the hardware set up so turning the microphone on and off. The hardware component only gathers the audio data and then the software handles it further (under Line of visibility).

6.2 Software processes

The software processes (Backstage) are mapped out from processing the audio data through automated speech recognition to NLP processes. As a result of the software processes, some actions have implications for the user actions (Frontstage). Throughout processing the speech (figure 38), with Attendi's technology multiple APIs (Application Programming Interface) are combined. The ASR results in a transcript with timestamps which is further processed with speaker segments. Based on the transcript, a relevant template is suggested for the anamnesis. The segmented transcript (figure 39) goes into the NLP module, where the information is classified into categories and keywords. These words are mapped to the anamnesis standard and a score of confidence (%) indicates the certainty. Above a set threshold, the classified words become automated suggestions that are presented on the interface. The clinician can accept, reject and edit these suggestions during template completion. The NLP processes and the clinician input is part of a training loop that allows the software to learn over time.

Figure 38: Processing the speech

Upon completing the template, it is saved and the text can be copied into the EHR or used for follow documentation tasks. The Assistant also offers the option to search within patient files in the database that stores the audio recordings, the transcripts and the anamnesis templates.

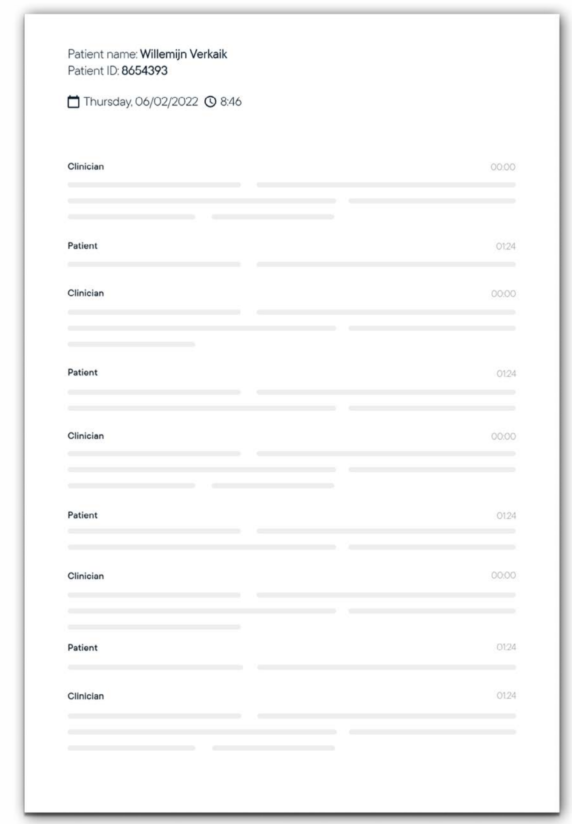


Figure 39: Segmented transcript

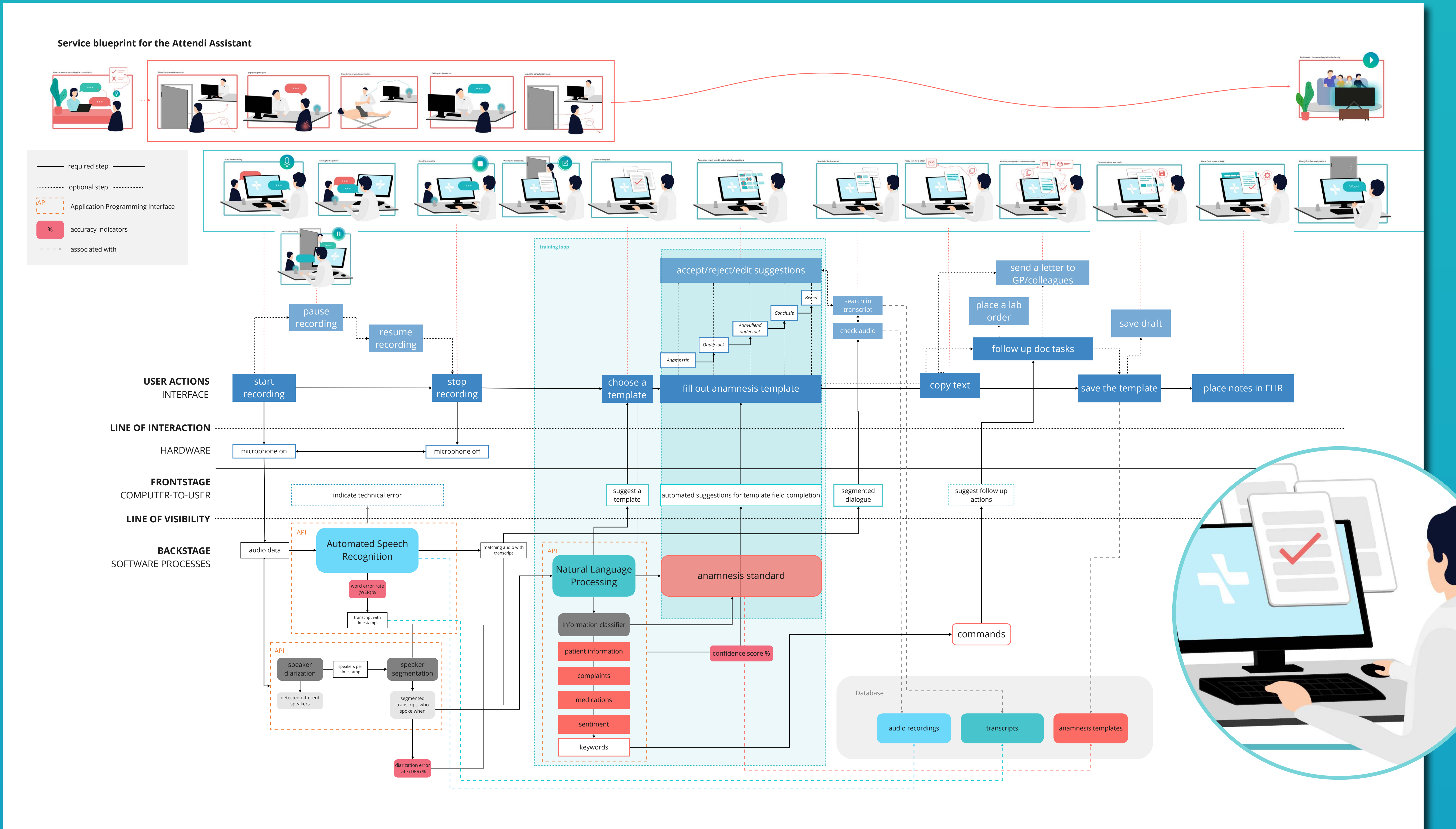


Figure 40: Service blueprint for the Attendi Assistant

Image source: Pexels

Chapter 6 takeaways

- Translating the technology into the service blueprint was a detailed process during which multiple discussions with the CTO served as important inputs.
- The combination of the service blueprint and the storyboard was a new method I learnt. The storyboard illustrates the steps that the user goes through, so anyone can understand it visually. On the other hand, the backstage part of the blueprint is flowchart scheme that also works well with technical people. Together, the user perspective and the technical considerations can be connected to see how one influences the other.
- During the process, the service blueprint was useful to have one overview at hand and facilitate discussions. The steps also informed the interfaces that should be implemented in the prototype.

Chapter 7

Ideation

In this chapter, the design approach is explained zooming into the interface development. For the concept, different approaches are considered with its implications. To design the anamnesis template, certain elements need to be defined to aid the design process and discussions. From the blueprint three core user flows are presented.

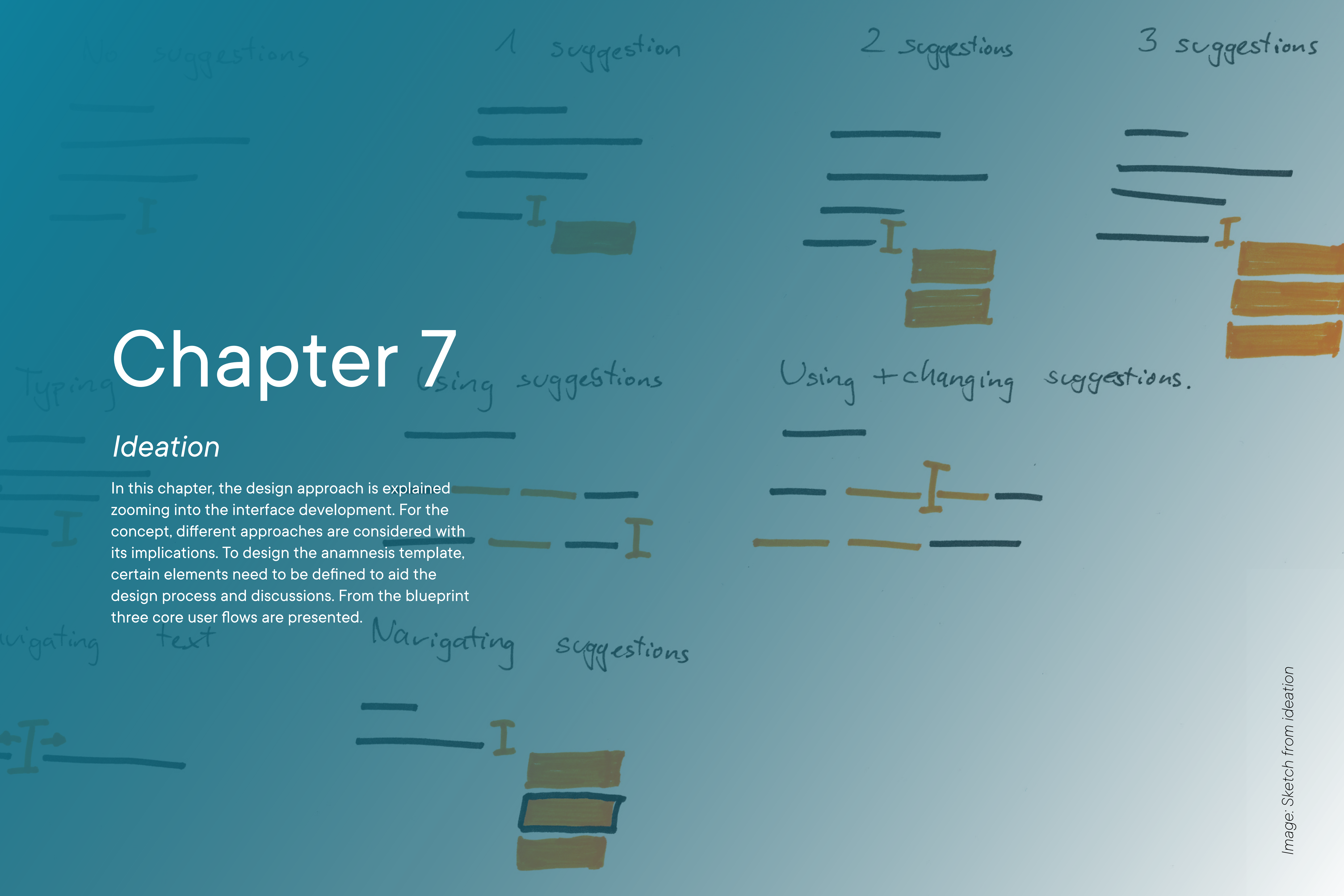


Image: Sketch from ideation

Design approach

An iterative design approach was used for ideation where both the product manager and the Medical lead of the company was involved during multiple discussions. The plan for prototyping was to start with low fidelity sketches and go higher fidelity. Furthermore since the defined scope guides the design phase during this thesis time, the decision was made to design for scale. To do so, a component library was considered to define interface elements that are reused throughout digital product development. This means the components that are used in the prototype can be reused in future development of the Assistant when adding new features.

To start low-fidelity, some sketches were made using pen and paper. Paper prototypes help designers to iterate fast and communicate ideas early in the process. During this phase many points were discussed internally that helped to establish a common terminology.

For example, it was important to agree on what to include in the navigation panel before going higher fidelity (figure 41). Also, templates were still up for interpretation so different layouts were considered.

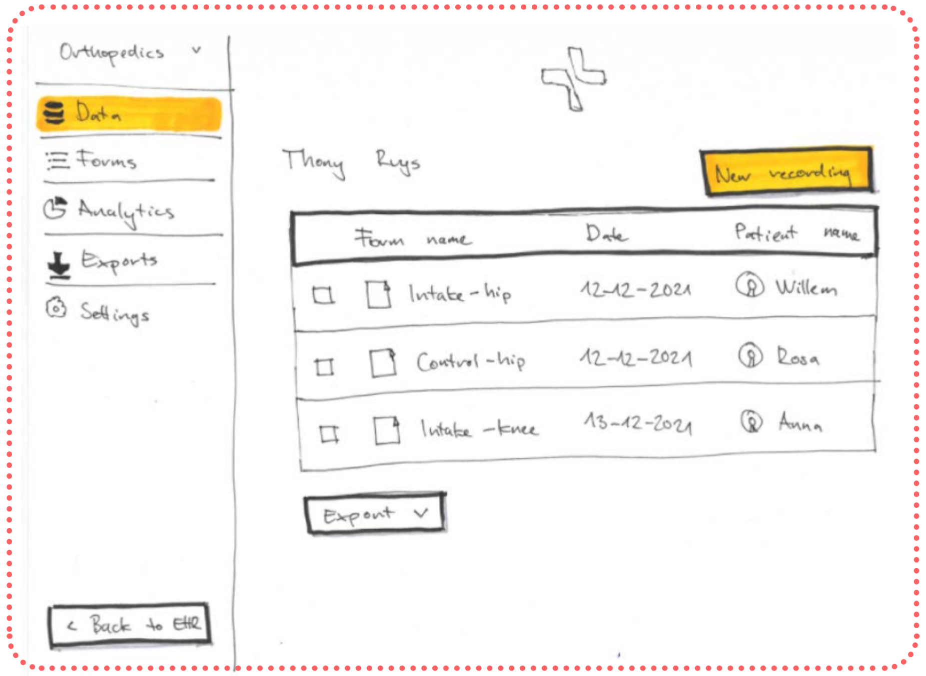


Figure 41: Early sketch for overview of templates

It was critical to discuss the utility of providing the transcript of the conversation in comparison with a summary (figure 42). Although summary sounds like a great promise, it is really hard for NLP to summarise the notes based on the transcript. The reason for that is that the computer is good at analysing text within a sentence such as the syntax, but to comprehend logic behind a full paragraph is very challenging. A computer can not summarise a conversation the same way as a human can.

Figure 42: Comparing transcript, summary and template fields

7.1 Concepts

To design a template-based system, several approaches to the concept were considered. After the conversation is recorded and transcribed, we have the entire consultation text to work with. That means that we can organise the conversation text into templates (figure 43) however that means clinicians would start editing text and we do not win time with that. Also, for the software, it is hard to segment the transcript and know which part belongs to Anamnesis, Onderzoek, Aanvullend onderzoek etc sections since certain sentences in the conversation could belong to multiple sections due to the non-linearity of doctor-patient conversations (Kocaballi et al., 2020). Next to that, the templates can be organised into a menu-like clickable format (figure 44) which also allows for structuring the data. Each approach has advantages and disadvantages which were discussed internally with the medical lead. Logically, it is also possible to combine the two approaches which could allow for both speed and adding detail (see Design goal in chapter 5.1.1).

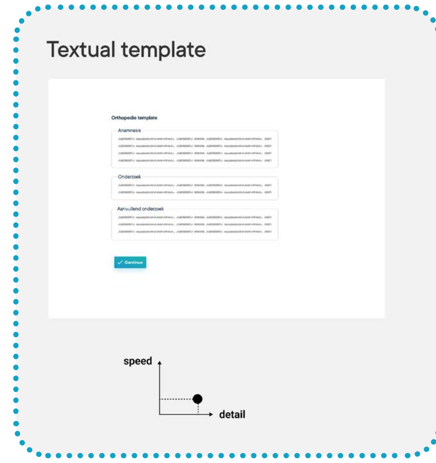


Figure 43: Textual template approach

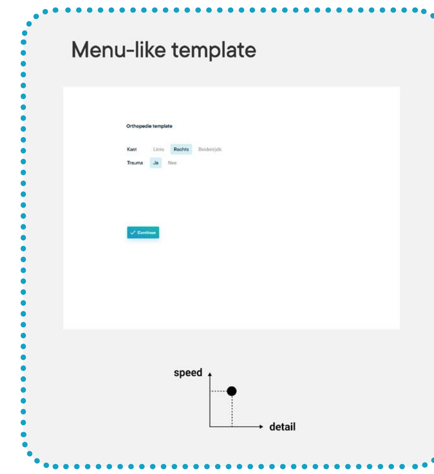


Figure 44: Menu-like template approach

To conclude, a combined approach (figure 45) is preferred by surgeons that requires the least amount of clicks to go through the anamnesis. Next to clicking, they should have the option to type as well but should not start editing the transcript. For system transparency, the transcript can be provided on a subtab of the interface.

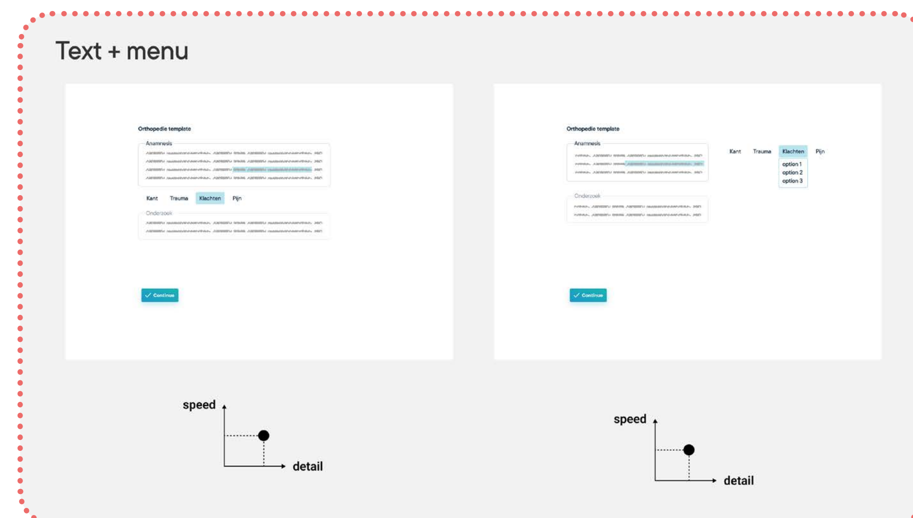


Figure 45: Combined approach

7.2 Anamnesis template

To structure the data, an anamnesis standard was provided by the company which was curated by collaborating orthopedic surgeons. In the anamnesis standard, questions relevant for hip arthrosis consultations are listed along with possible answer options (Appendix E). In the anamnesis standard, the questions follow the mentioned sections that clinicians use during their note-taking (Journey map in 31). Each section has a typical content that is summarised in figure 46.

Section name (Dutch)	***Section name (English)***	***Content of the section***
Anamnesis	Anamnesis	State of the patient
Onderzoek (or Lichamelijk onderzoek)	Research; physical examination	Results from labs (i.e. MRI) and from physical examination
Aanvullend onderzoek	Required research	Required more labs
Conclusie	Conclusion	Conclusion of the consultation
Beleid	Policy	Expert advice for next steps

Figure 46: Identified notes sections from the shadowing activity

To design for an anamnesis template, the questions in the anamnesis standard need to be looked at. Each question leads to a certain data input and also should have a consistent interaction type. There are 4 question types in the anamnesis standard: yes/no questions, multiple choice questions, checkboxes with multiple options and also open questions (figure 47).

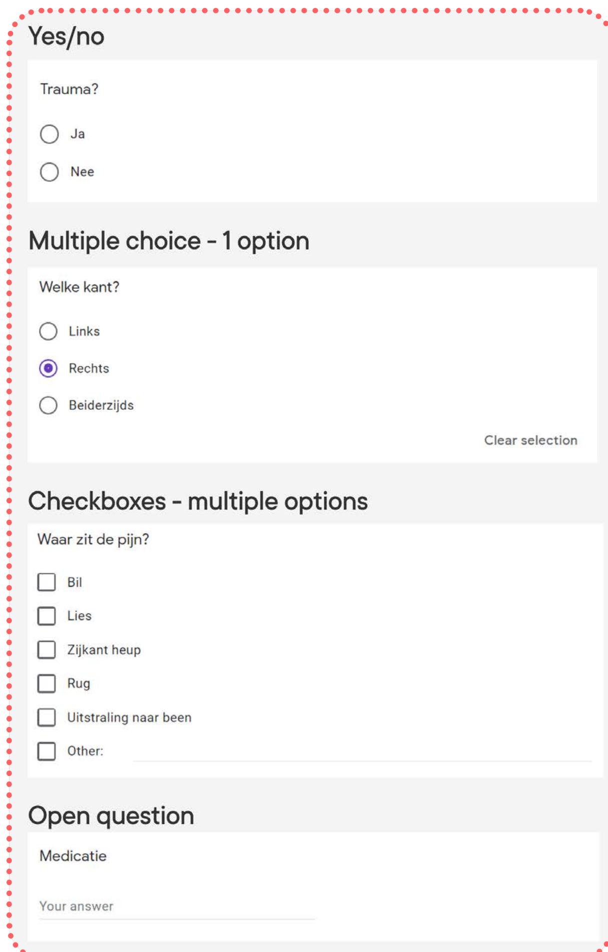


Figure 47: Question types from the anamnesis standard

7.2.1 Definitions

To further design and also guide design discussions, some definitions had to be clarified. Simplified visuals (figure 48) were used to have these discussions and find a common terminology with the product manager.

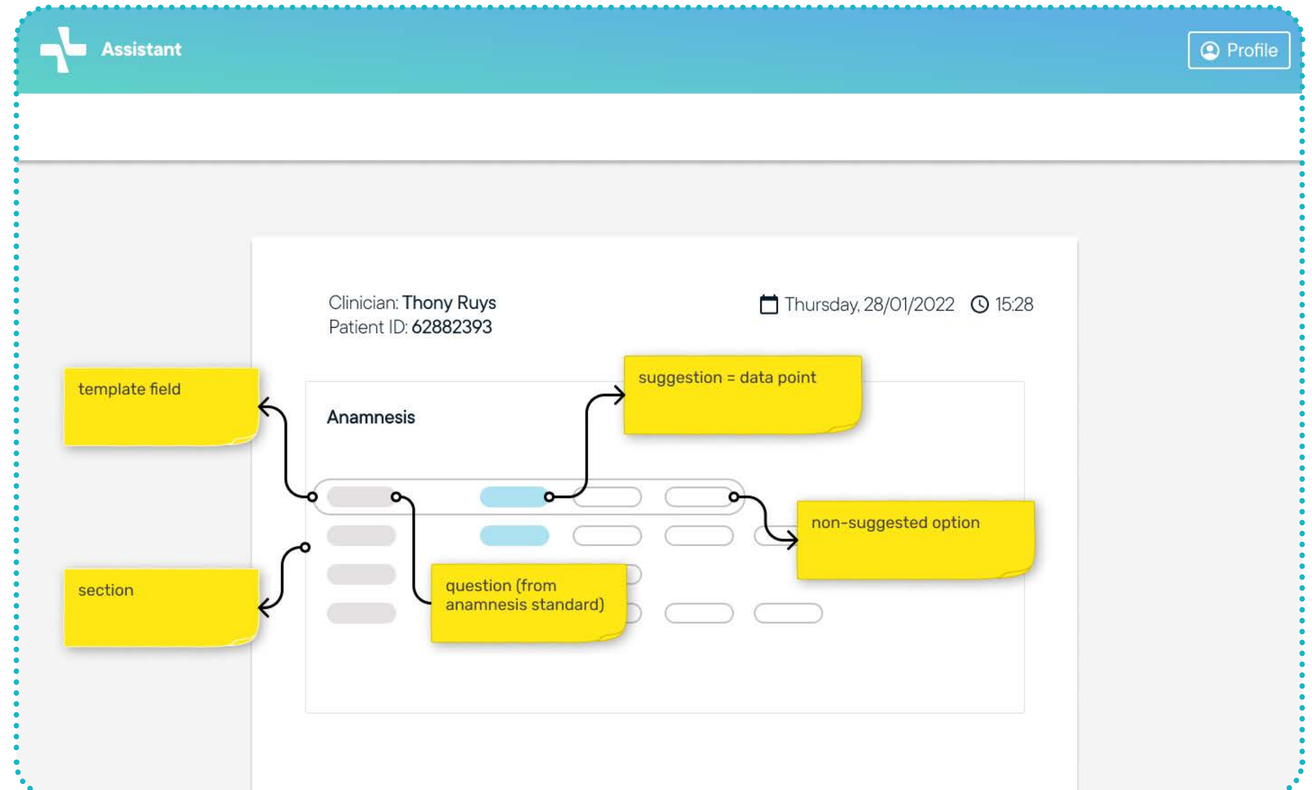


Figure 48: Simplified visual to discuss definitions: elements

The template fields are ordered by the questions from the anamnesis standard. From the transcribed transcript, words can be picked out and used as a data point (figure 49). The Assistant can provide automated suggestions to add value with NLP. To allow the clinician to have control over the automation, the non-suggested options are also clickable but not highlighted. Next to both, text input in each template field should be possible.

Figure 49: Simplified visual to discuss definitions: result

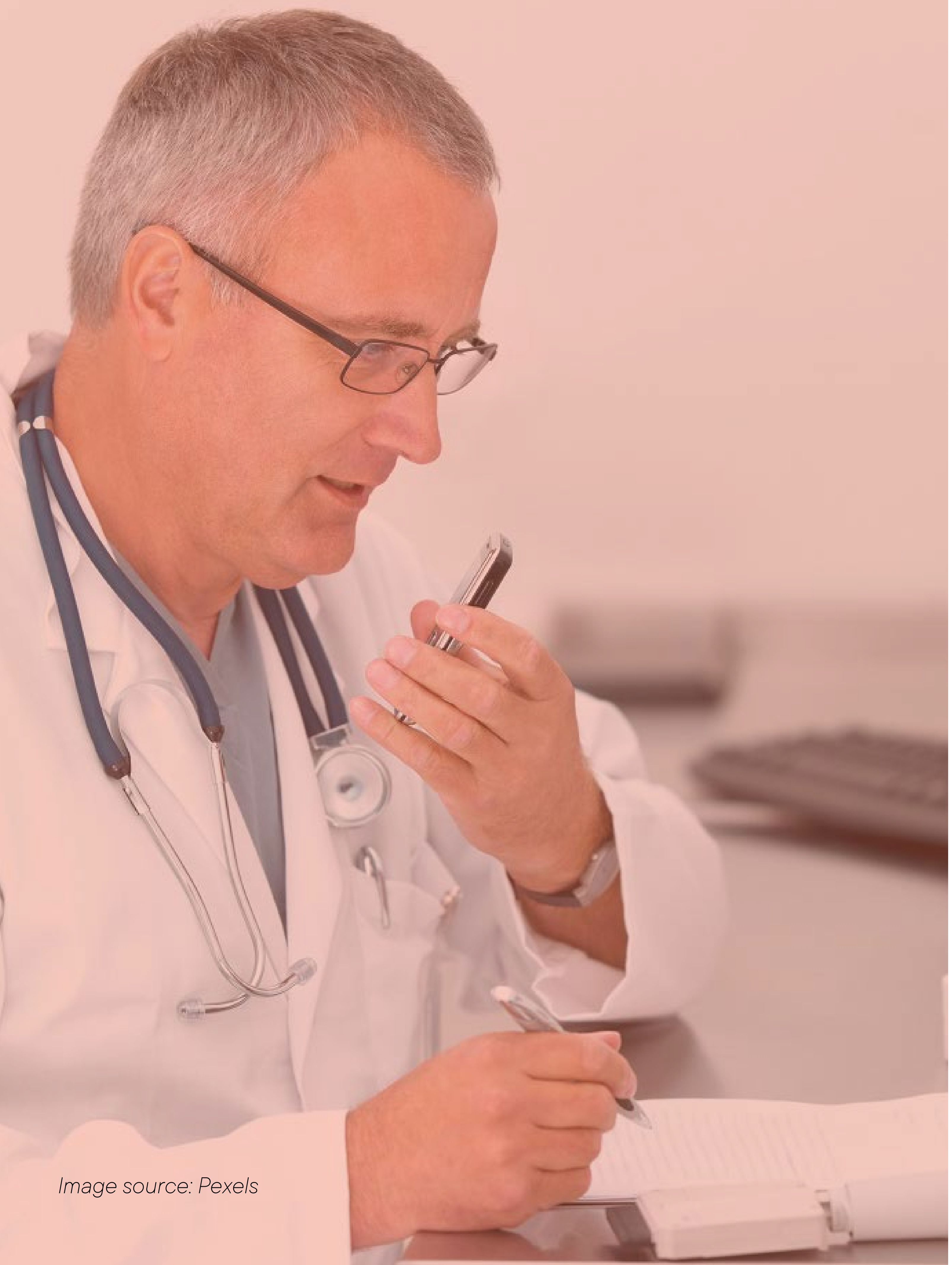


Image source: Pexels

7.3 User flows

Next, it was time to start mapping out user flows. A user flow is a series of steps a user takes to achieve a task on a digital interface (figure 50). In terms of structure, a user flow should have a purpose and go only in one direction. User flows help to communicate how the technology works and how it affects users while interacting with it. Since it is a flowchart, it is also helpful to communicate with non-designers. User flows also help to go higher in terms of prototype fidelity.

After clarifying the software processes and what the user steps are, three user flows were extracted from the service blueprint:

1) Record a consultation
2) Fill out the anamnesis template
3) Query specific patient file from the database.

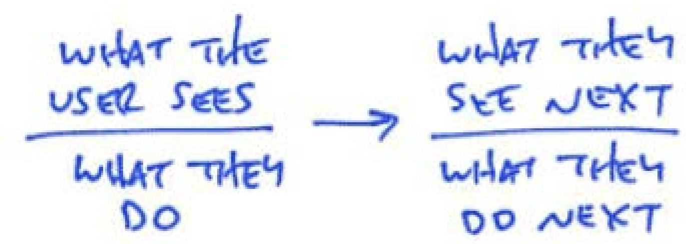


Figure 50: User flow in simple terms

1. Record a consultation

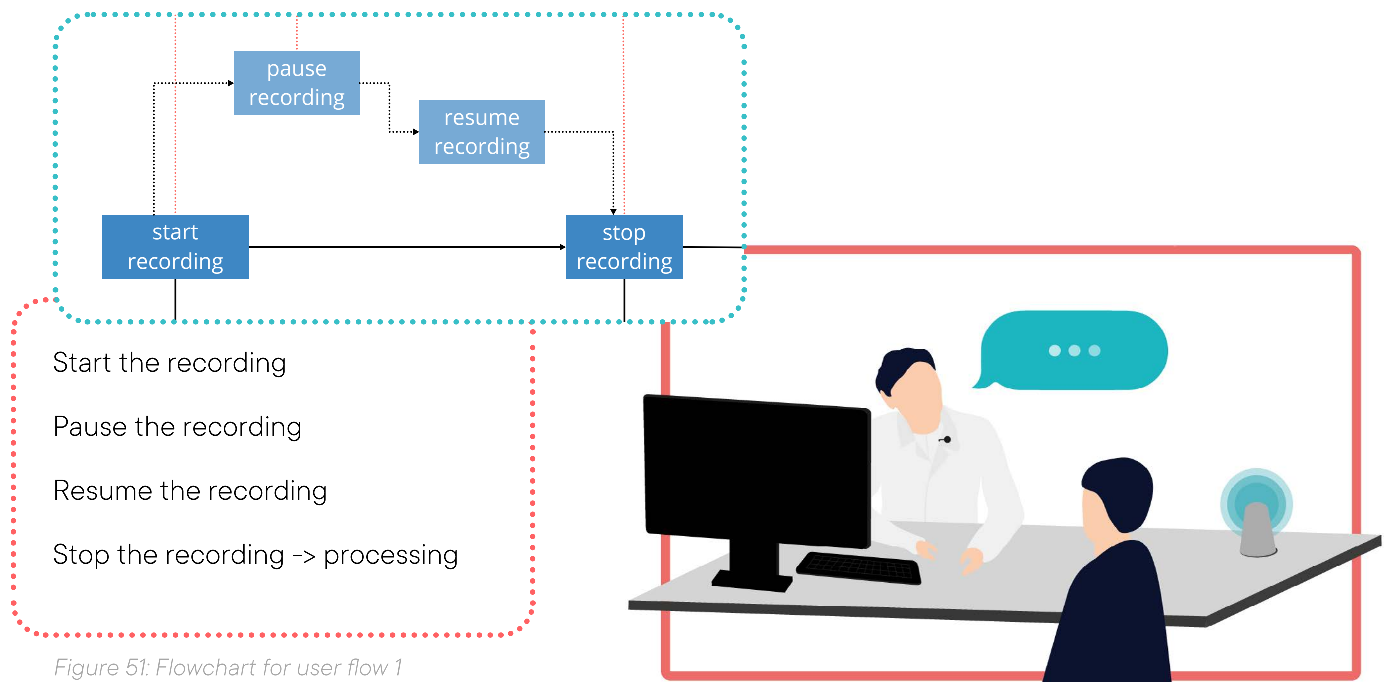


Figure 51: Flowchart for user flow 1

2. Fill out anamnesis template

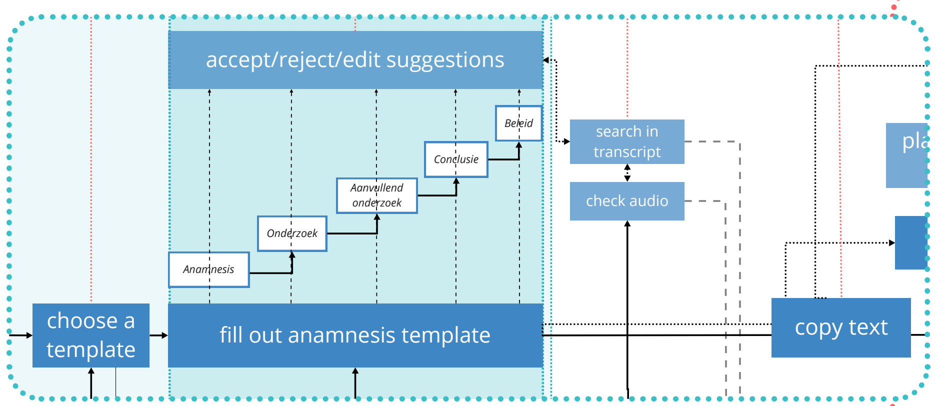


Figure 52: Flowchart for user flow 2

- Choose a suggested template
- Fill out the anamnesis template
  - Accept suggestions
  - Reject suggestions
  - Edit suggestions
    - Option to type
- Copy text

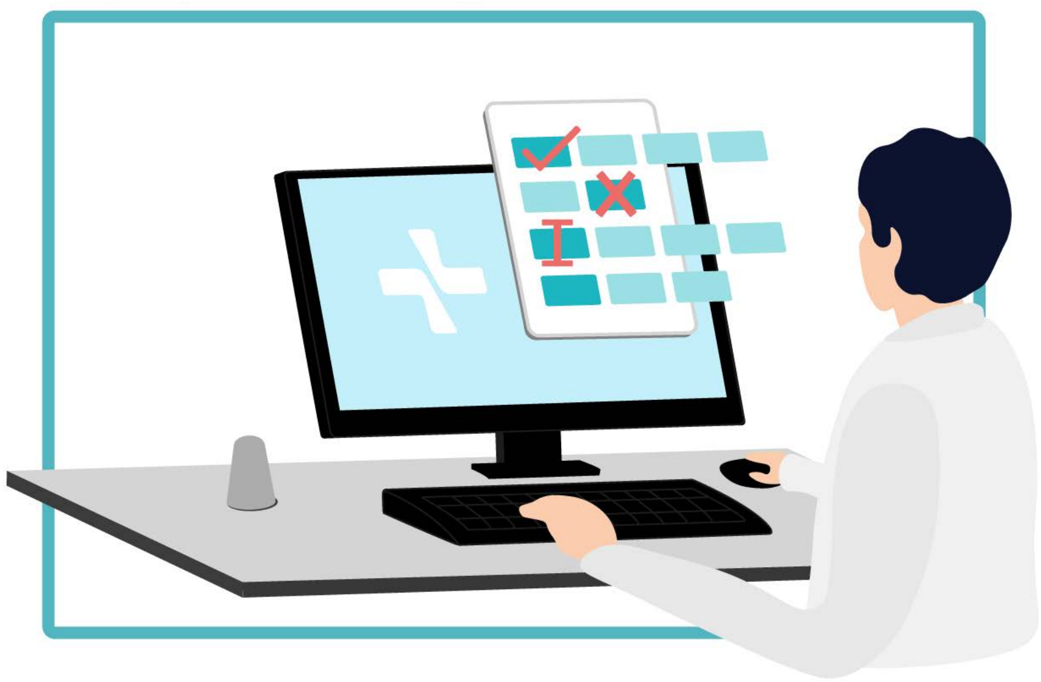


Figure 53: Storyboard - accept/reject/edit suggestions

3. Query specific patient file

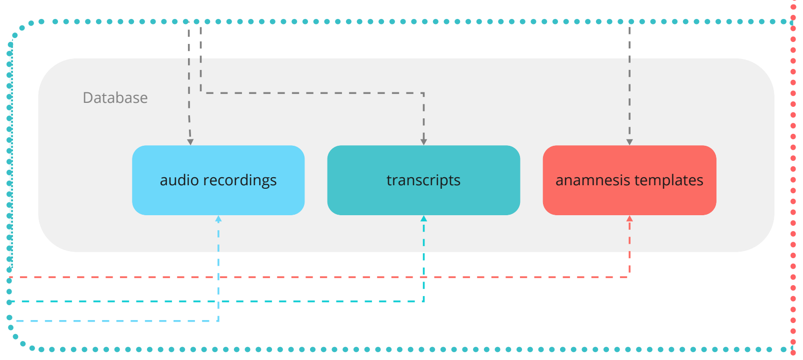


Figure 54: Flowchart for user flow 3

- Check overview of recorded consultations
  - Filter per draft templates
  - Filter per finished templated
- Search by patient name
- Export patient files: audio and transcript
- Search within transcript

Chapter 7 takeaways

- In many rounds, design decisions were made by reviewing ideas internally with the product manager.
- The most interesting discussions were around the summary. Clearly, getting a summary from the entire conversation with one click would be desired. However, from a technical perspective it had to be understood that it is not very feasible. As a designer, this was an important moment to balance between user wishes and technical considerations.
- It took time to clarify the definition of a template and its elements. Using visuals helped to manage these discussions.
- The design aim with the interface should be speed and the ability to add quality details as a clinician, therefore a combined approach was chosen for the template. On one hand, a menu-like clickable format and next to it providing the transcript for system transparency.
- Looking into the anamnesis template was a relevant step to understand the type of questions. Each question type should lead to a consistent interaction.
- From the blueprint, three user flows were extracted to aid the UX design process and develop the interface.

Chapter 8

Meet the Assistant

This chapter discusses the interface of the Assistant starting with the use case. The interface screens are explained which is illustrated by the storyboard from the service blueprint. The three user flows are realised: recording a consultation, completing the anamnesis template and looking up a patient file within the application.

The data in the demonstrator is fictional and does not claim any proved health outcomes.

The interface

8.1 The use case

Using the Assistant will be initiated from the EHR system that is in place in hospitals (Requirement #1). Upon opening the patient file, typically clinicians already have the 'Notes' panel open ready to type. In this panel, a button could be potentially integrated to initiate recording with the Assistant. The button opens the web-application in a new tab where the recording can be started.

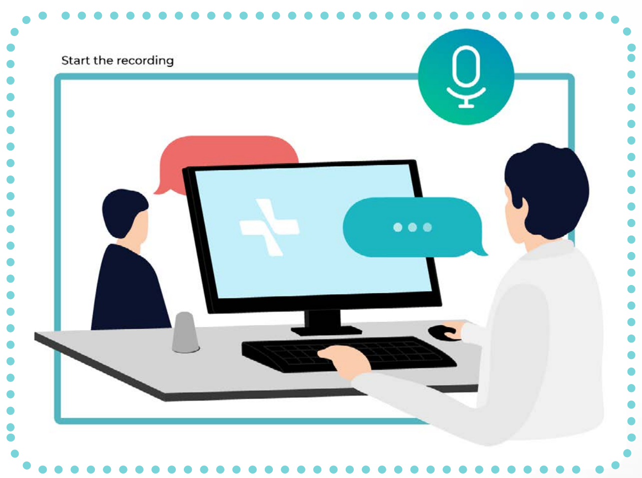
During the consultation the Assistant is listening to the conversation between the clinician and the patient. In case it is necessary, the recording can be paused or restarted. From the shadowing activity it was witnessed that clinicians still need to have the EHR open to search for lab results for example, therefore a widget is proposed (figure 55) that could still keep the user informed about the state of the recording. In the Assistant web-app, the same listening screen is available.

8.2 Storyboard

1. Record a consultation

Recording a consultation can also be started from the Assistant interface. Upon clicking the Start button (figure 56), the automated speech recognition is in progress.

The user has three options during recording: to restart, to pause or stop it. Providing the option to pause is necessary in case of any ad hoc event, when for example a nurse enters the consultation room. In that case, the system should not pick up the third person's voice and the clinician also wants to control that. The icons for the control actions are informed by common voice recording and music player interfaces.



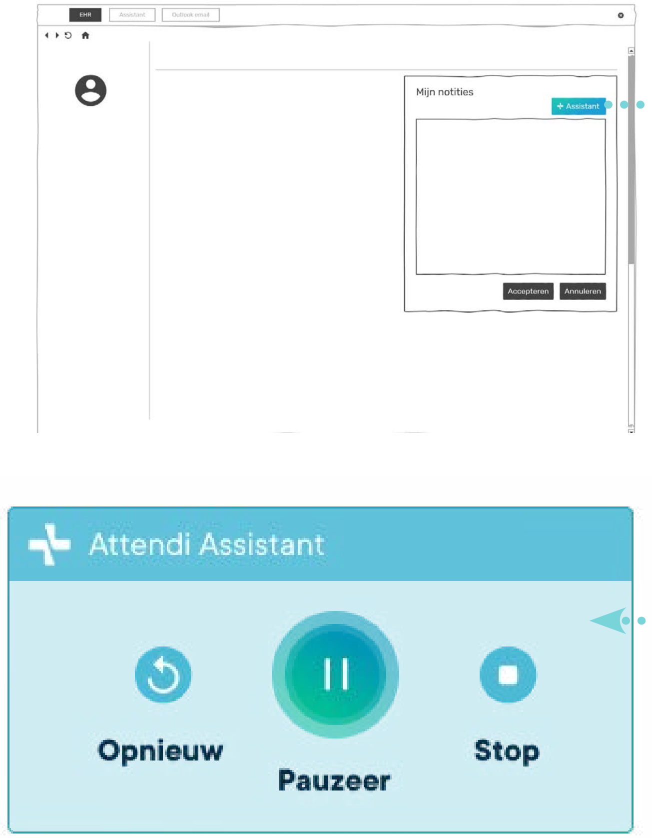


Figure 55: Widget

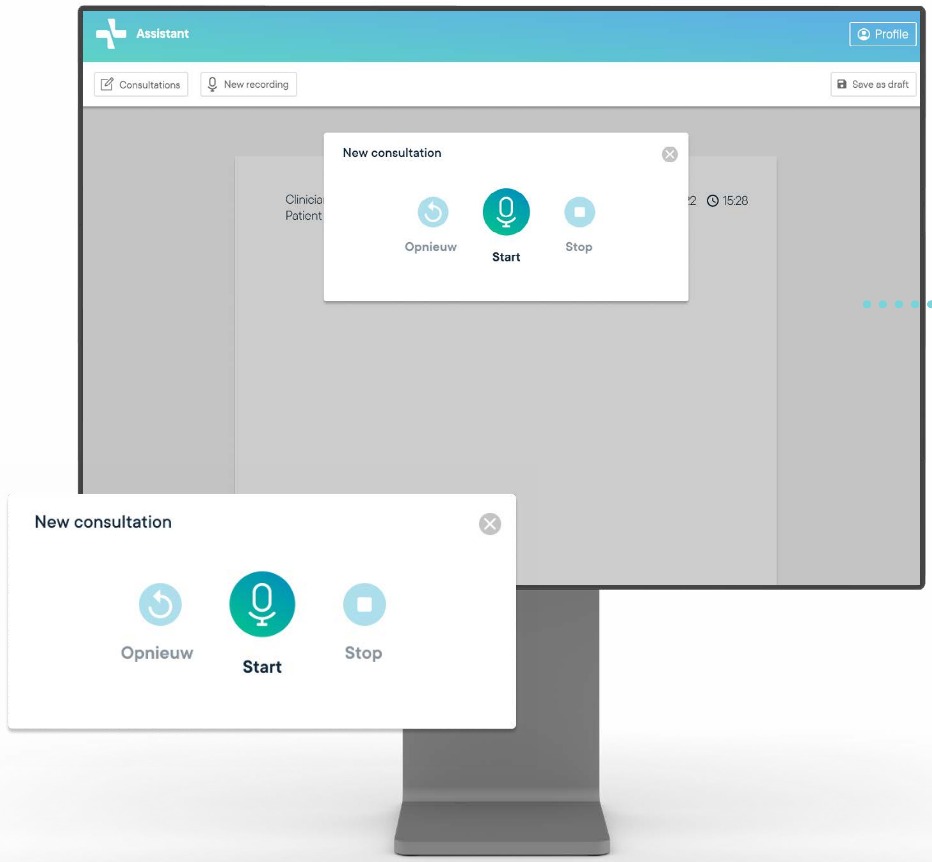
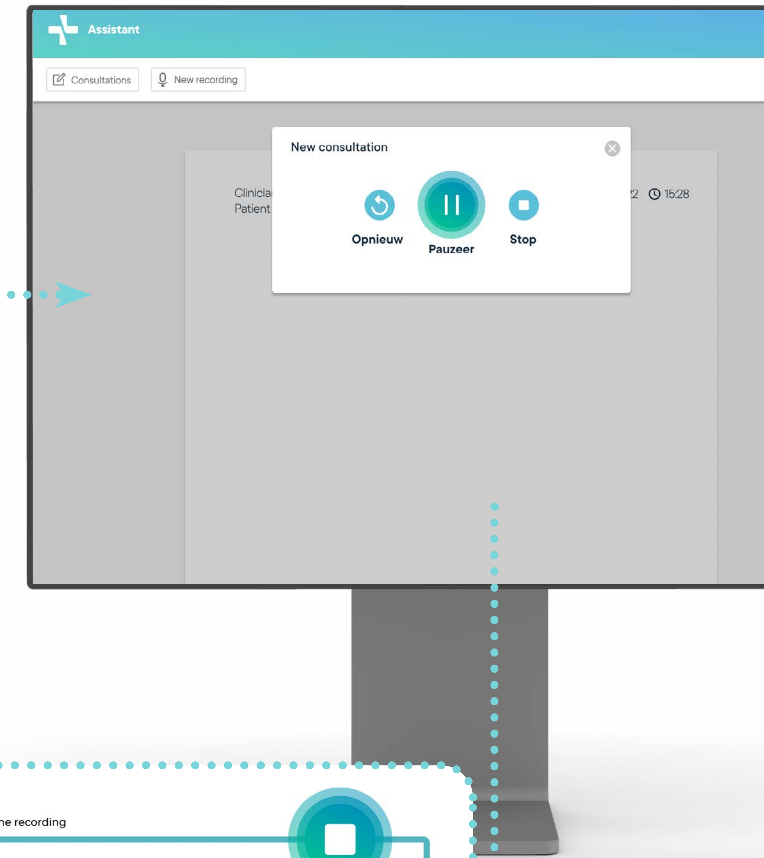
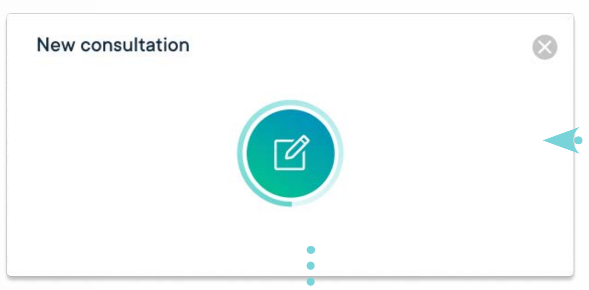


Figure 56: Start of the recording

During the recording there is an indication that the system is listening with the concentric circles around the pause button. The circles dynamically change as the voice changes to give visual feedback (figure 57) to the user. Own testing with several laypeople confirmed that visual indication is straightforward.

Figure 57: Recording in progress: pause or stop





Once the consultation is over, the recording should be stopped by the clinician. Stopping the recording immediately initiates processing the conversation.

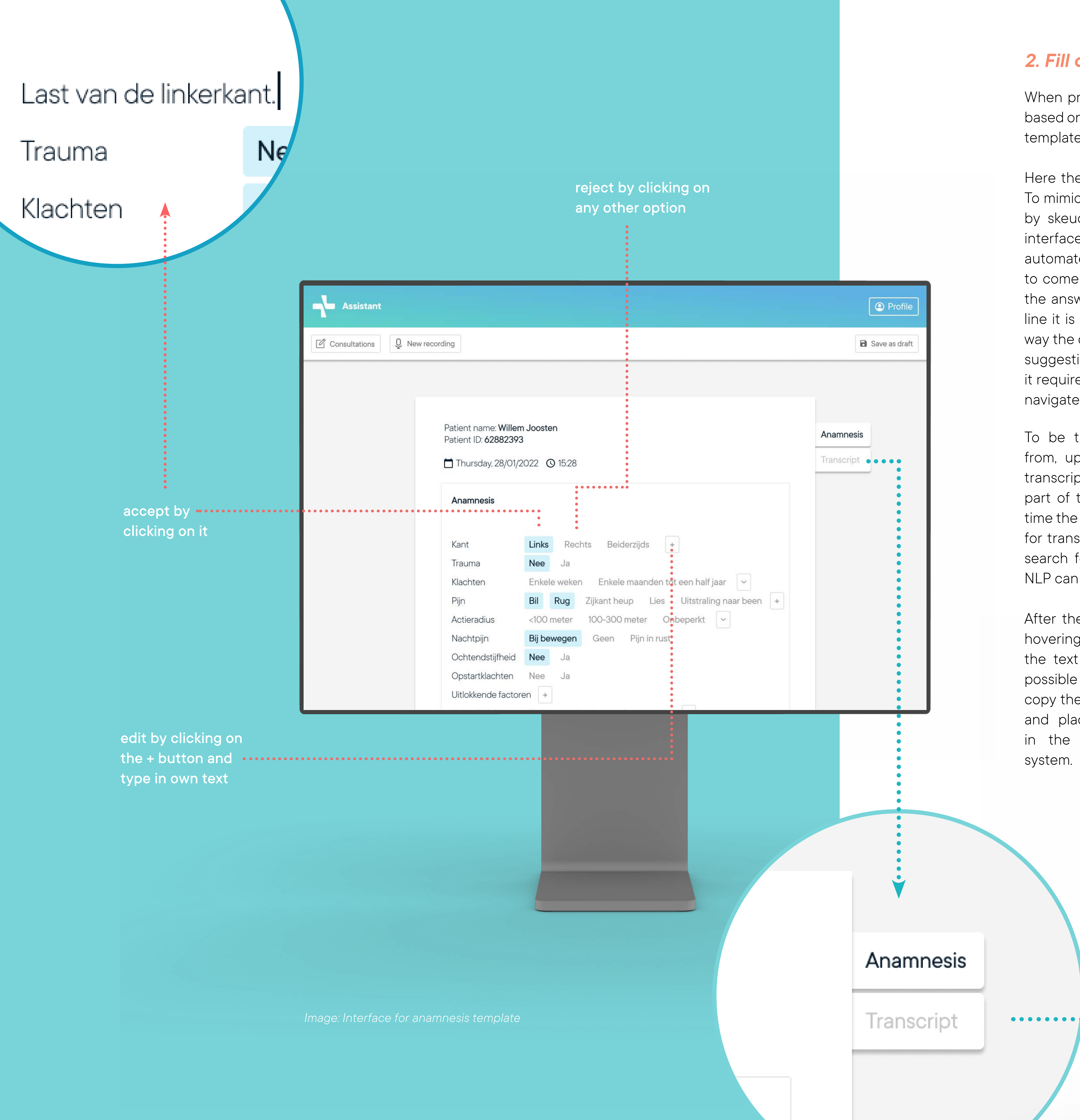


Image: Interface for anamnesis template

2. Fill out anamnesis template

When processing the conversation is done, the most relevant template option is suggested based on the transcript. In this case, a template for hip arthrosis is suggested. By confirming the template, the software knows which algorithms to use for the template type.

Here the hip arthrosis template is presented to the clinician, and partially filled in with data. To mimic real-world medical notes, the template is placed on a piece of 'smart paper' inspired by skeuomorphism. Skeuomorphism is a design concept in interaction design of making interface items mimic real-world objects. For each anamnesis question, a suggested option is automated using NLP from the conversation. Many recorded consultations allow the system to come up with accurate suggestions compared with the transcript of the consultation. All the answer options are clickable, but the automated suggestion is highlighted. Also, in each line it is possible to add own text as a clinician. By this way the clinicians are able to accept, reject and edit the suggestions (Requirements 16-18). For each user action, it requires one click from the clinician and they can also navigate with the keyboard.

To be transparent of where the suggestions come from, upon clicking on the anamnesis question, the transcript is displayed with a highlight of the relevant part of the text. While filling out the template, at any time the user also has the option to check the transcript for transparency (figure 58) where it is also possible to search for keywords. While typing in the search field, NLP can enable auto-complete to help the user.

After the several template fields are completed, upon hovering on the text it is possible to copy the text and place it in the EHR system.

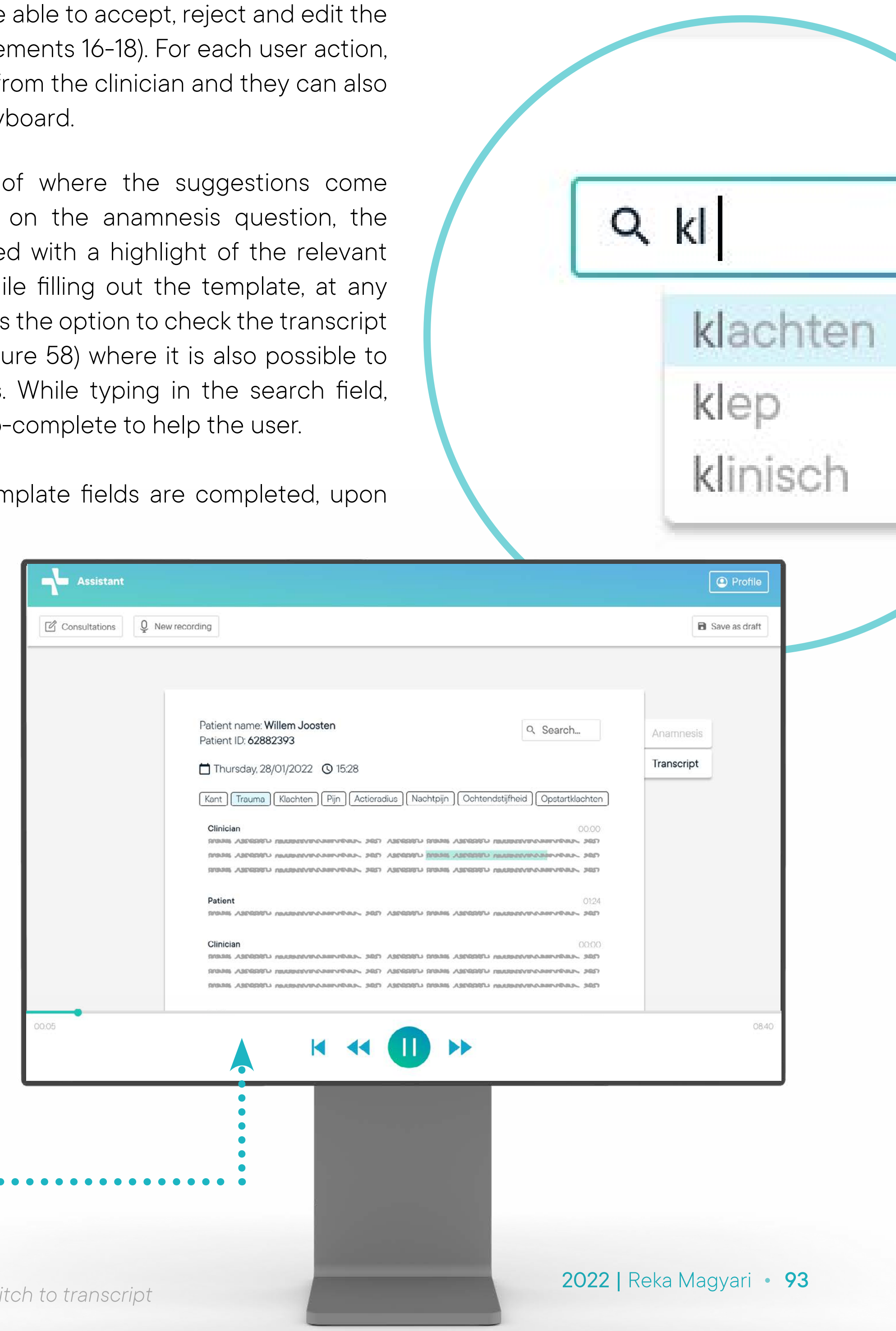


Figure 58: Switch to transcript

3. Query specific patient file

In the Assistant, an overview of the recorded consultations are provided per patient. For each patient file, the audio recording, the conversation transcript and the anamnesis template is provided. Since it is possible to save unfinished templates, the files can be filtered as per drafts and finished notes as well. If needed the files can be exported from the system or sent as a letter (figure 59).

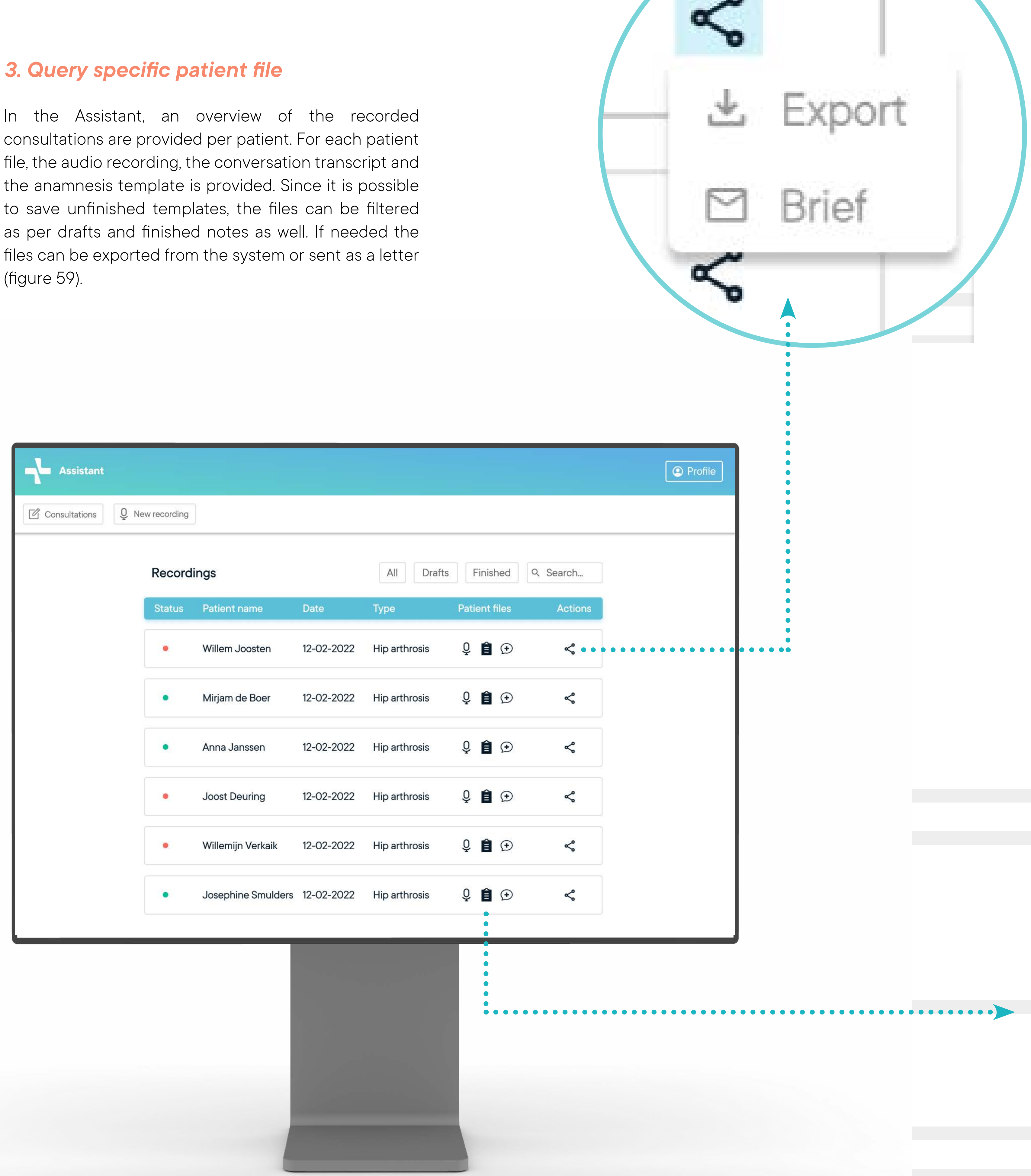


Figure 59: Interface for recorded consultations

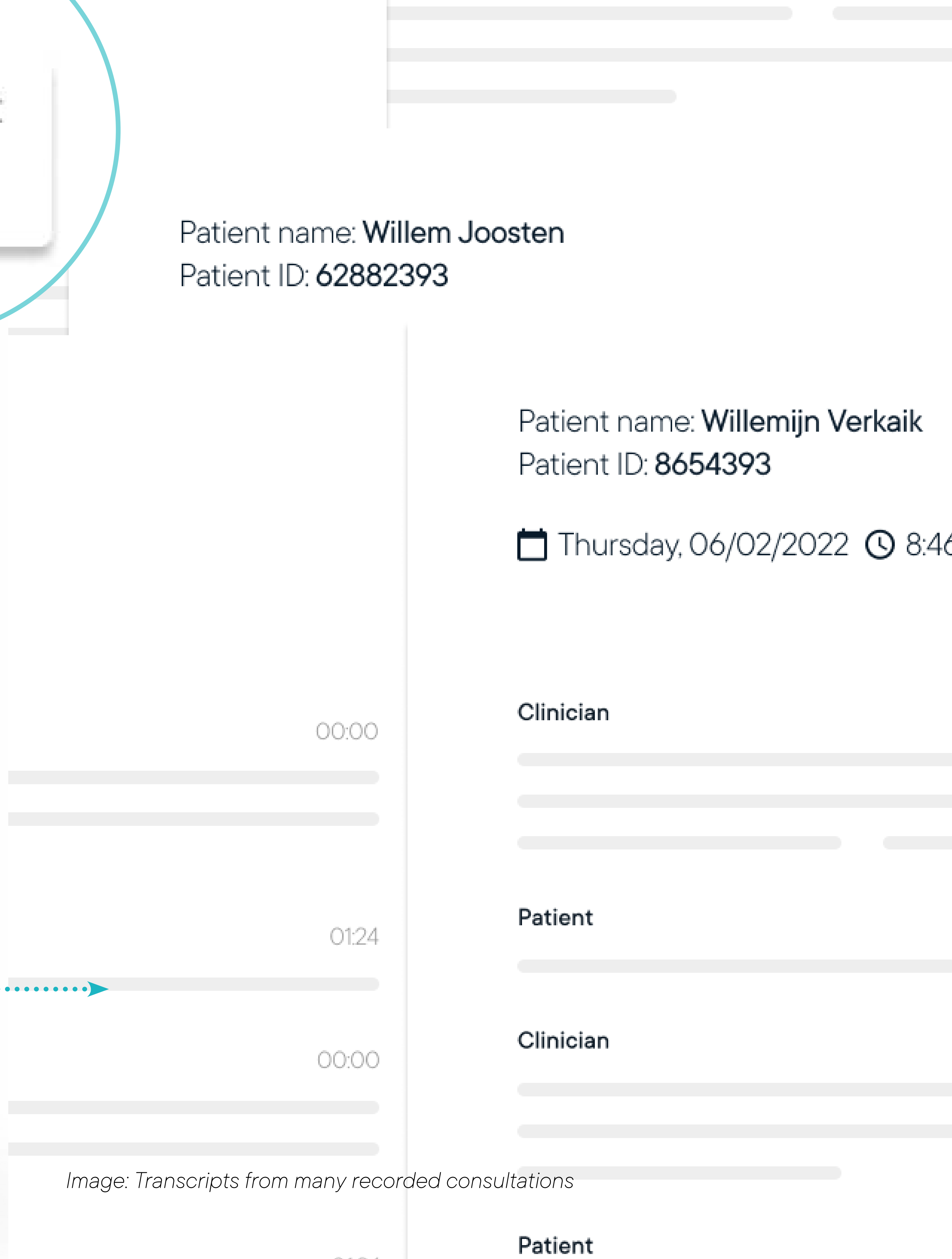


Image: Transcripts from many recorded consultations

RQ2. How can we design a clinically valuable, trusted and ethical digital scribe interface to aid computer-led documentation from a clinician's perspective?

Value - The value lies in speeding up note taking with NLP along with providing the option to add nuanced details. This way partial automation and human control is combined in a hybrid documentation scenario.

Trust - To improve trust towards the system, increasing transparency is implemented. In the transcript, it is highlighted from where the automated suggestions come from. Trust will build up over time when using the product.

Ethics - For maximum human control, the clinician can accept, reject, and edit the suggestions for each question. Ultimately it is the clinician's decision to sign off the final text and place it in the EHR. Further ethical implications are suggested to consider.

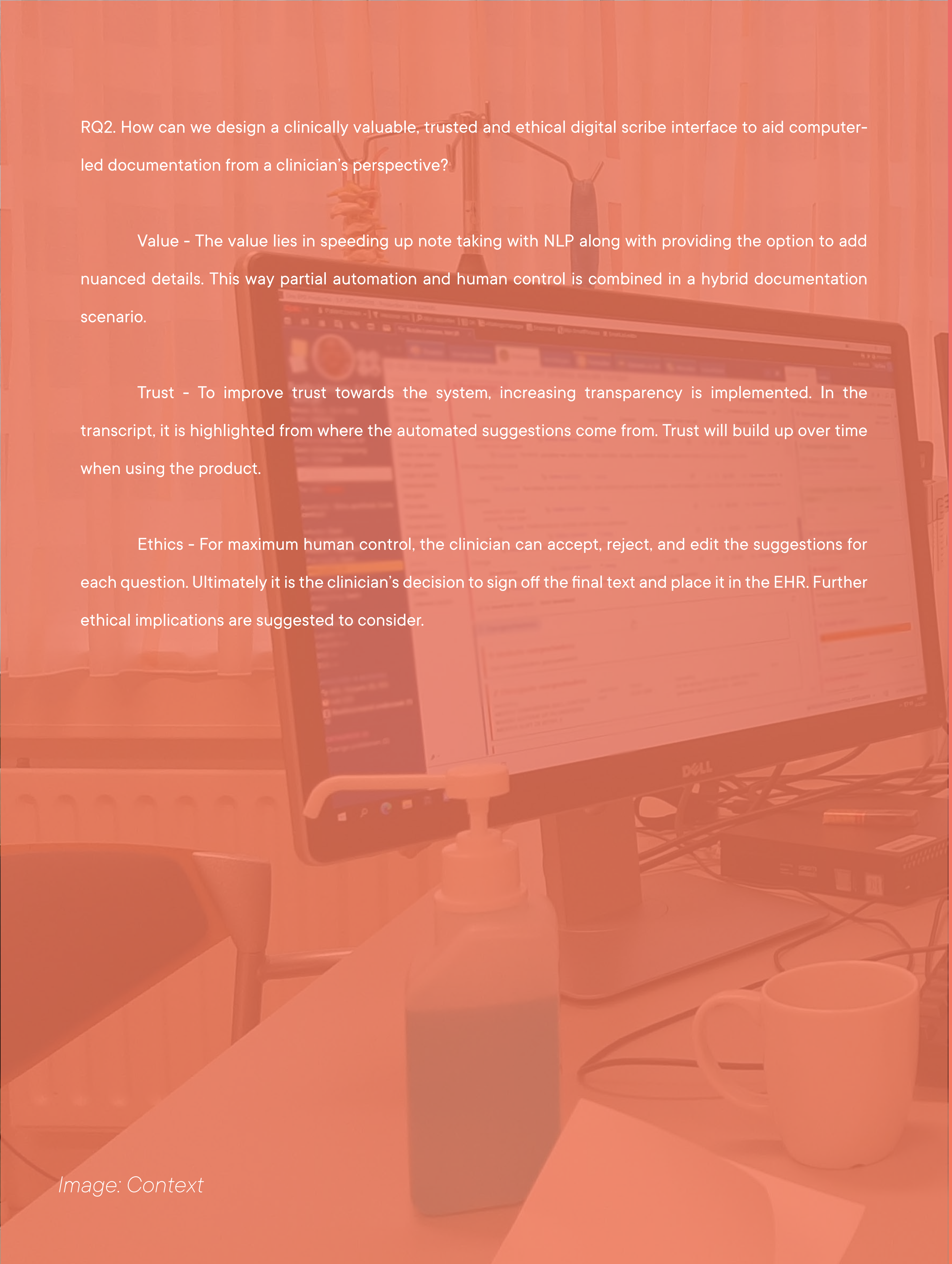


Image: Context

Chapter 8 takeaways

- Designing the interface was an iterative process and a result of multiple discussions, internal reviews and testing.
- A widget is proposed to be overlaid on the EHR during the recording. It makes sense for the use case as the Assistant needs no interaction during the consultation, however implementing it technically will require an EHR integration.
- Processing the speech is an important step of the journey, but without real data it is hard to claim its duration and impact on the user experience.
- The list of questions are informed by the anamnesis standard to integrate the medical expertise in the project.
- Use of icons are informed by existing interfaces that the users are already familiar with so they don't have to learn their meanings.
- During template completion the buttons indicate that they are clickable, but it is not clear that you can also navigate through with the keyboard.
- Transcript is provided in case the clinician wants to check it but they don't have time to do that during documentation.
- Providing an overview of the recorded consultations is in line with the roadmap (Step 1), but organising safely storing the data is not considered in the scope of this thesis.

Chapter 9

Validation

In this chapter inputs from end users are discussed. Throughout the project a surgeon gave weekly inputs, a focus group session was held and a final user testing round was organised with orthopedic surgeons. The qualitative results are discussed in order to understand the potential value that the Assistant can provide. The project is evaluated from a desirability, feasibility and viability aspects.

Image source: Pexels

Validating the design

9.1 Involving users

9.1.1 Weekly input

Throughout the project, it was extremely useful to have direct contact with a surgeon. With his expertise the anamnesis standard could be translated into the interface. The weekly medical input helped to make smaller decisions and faster until showing the work to more orthopedic surgeons. The main learnings from this collaboration was to implement interaction types that surgeons already use for i.e. clicking through buttons (figure 60) or formatting text. Also it was very important to design the template completion with the goal to finish it with the least amount of clicks possible.

Figure 60: Clickable buttons

9.1.2 Focus group session

Furthermore, an inspiration session was organised at the ETZ hospital where the shadowing was performed. In a presentation, the findings from research were presented (figure 61) to an audience of about 10 orthopedic surgeons. The members were very diverse, from young to older surgeons. As mentioned before, age and experience is relevant towards the attitude of the solution. Each surgeon was specialised in an area, not only hip arthrosis consultations.

The presentation was used to facilitate a table discussion around the topic. The audience was triggered by seeing the activities visualised and raised many questions for example if using the Assistant requires to follow a specific list of questions in order to get it working. Also, the learning curve is important for clinicians in order to adapt to the new way of working.

"Do you have to follow a specific list of questions in order to use the Assistant?"

"Did you do any investigation into the learning curve of using this stuff? Because we as doctors are very conservative, so if we adapt a way of working we stick to it. So your line of emotions would go down very much, I think you have a challenge there."

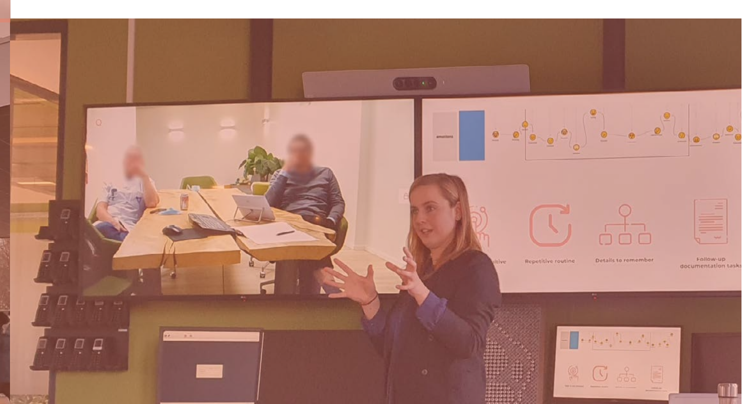


Figure 61: Presentation at ETZ, February 11

Also, concerns were raised that they will be busy editing and lose time. This aspect had to be clarified explaining the utility of the template over transcript concept. Clinicians care about time and speed, less about data structure.

"You want to be really complete, so you write down everything and the risk is that you are busy editing. But I think this is a system that can evolve and solve that."

"My hunch is that it might be very useful when doing history taking only [Anamnesis]... because my history taking is always different, my physical exam is sort of smart text, and my conclusion and way of treatment [policy / beleid] is always the same."

Furthermore, they were asking if it is possible to couple the Assistant to features in the EHR to allow them to decide which one to use for each section of the consultation. During physical examination (figure 62) smart text is favoured.

"I don't use smart phrases at all."

"I am very fond of smart phrases, but the thing is, if we use them a lot we do not recognize our patients [in the notes] anymore. That's why I want some flexibility in my history taking, but my decision taking for treatment is always the same."

Figure 62 Physical examination bed

9.1.3 Final user testing

Lastly, a final user testing was organised. The prototype was tested with 4 orthopedic surgeons who were also interviewed earlier in the project. The goal of the testing was: Evaluating the flexible ease of use of anamnesis template completion

As preparation, the users were informed about their voluntary participation. The testing was organised either remotely or in person based on the participant's preference and availability. Upon consent, the prototype was presented. In the prototype, no data input is logged and was told to the participants.
In the test the participants were asked to interact with the prototype (figure 63) and fill out the presented template. They were asked to talk out loud during testing to understand their perceived experience better. Regarding qualitative insights, the expectations of getting a summary of the conversation in one click were high. It implies that explaining the technology to end users is advised.

Surgeon 1 is from a hospital, where recently auto-anamnesis was introduced in which patients fill out PREMs before the consultation. Depending on the hospital, this aspect also influences the value. What is general, that during physical examination logically it is impossible to type notes and there are many details to remember.

"Can we skip the template completion step and get a summary?" - Surgeon 1

"To be honest, I don't see much value in filling out this template than the patient. But during physical examination, it would be very useful because my hands are occupied" - Surgeon 1

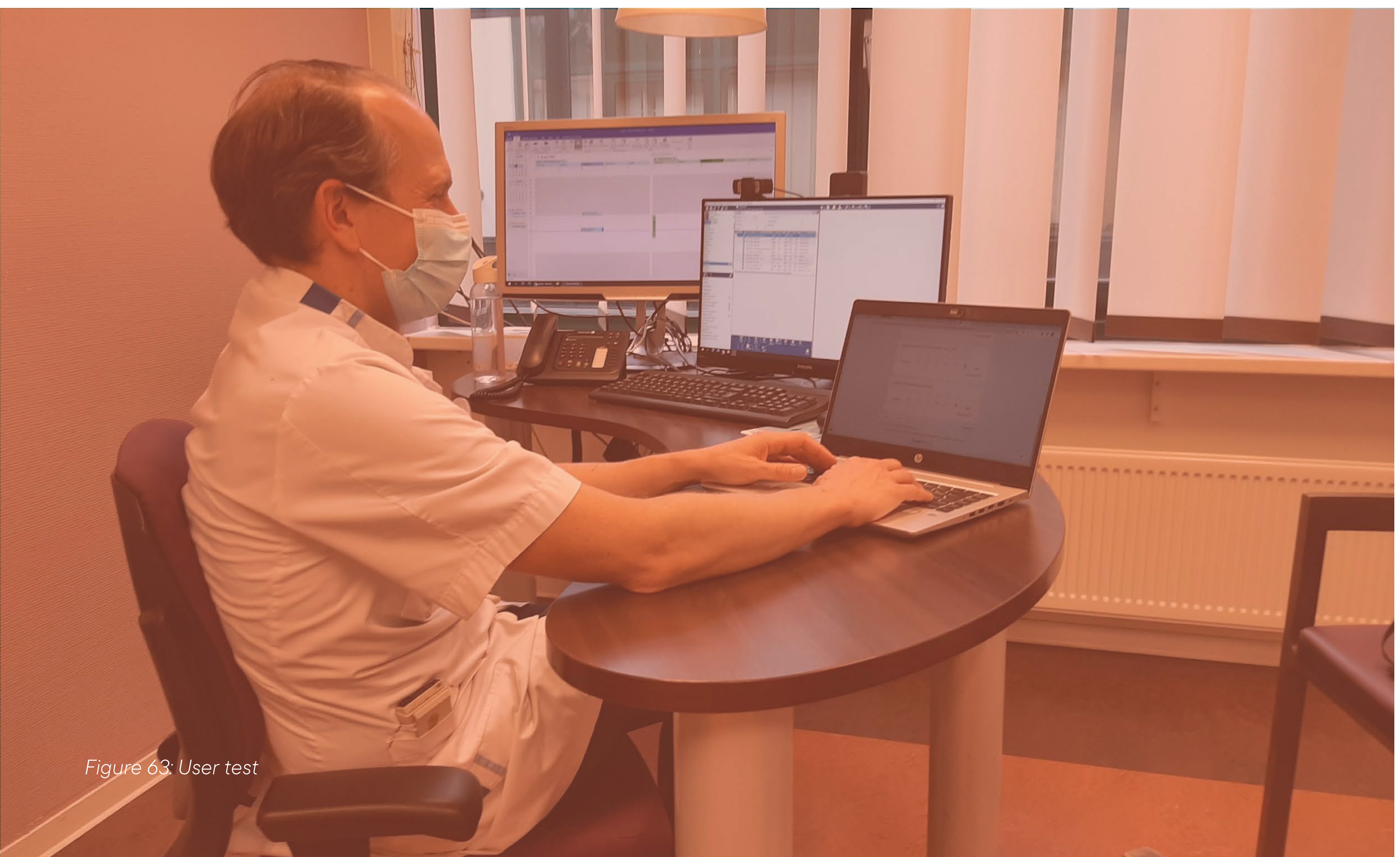


Figure 63: User test

Regarding flexibility, the buttons are easy to click. Throughout completing the template, questions should be possible to skip. When clicking on a suggestion, a commonly used phrase is displayed. According to even this small sample size, the preferences are different.

"What if I want to skip this question?" - Surgeon 3

"I would expect the system to create a nice note with one click" - Surgeon 3

"For me a list is fine, I don't care if it is a sentence or a list. But maybe my colleagues will not agree" - Surgeon 4

Last van de linkerkant.
Er is geen trauma aan vooraf gegaan.
Patiënt had al 2 weken klachten.

Kant: links
Trauma: nee
Klachten: 2 weken

Furthermore the user test was combined with a questionnaire informed by the Technology Acceptance Model (TAM). TAM scale is a validated questionnaire from the healthcare domain that focuses on measuring the perceived Usefulness and Ease of Use. To determine the perceived usefulness of a product, users are asked to rate the product on six points using a Likert scale. The averaged results are as follows:

"It is very important to be able to click on the other options as well. Oh, these are clickable." - Surgeon 2

Perceived Usefulness

- 5.75 — 1. Using this product at work would help me complete tasks faster.
- 5.75 — 2. Using this product would improve my job performance.
- 5.75 — 3. Using this product would increase my productivity.
- 5.75 — 4. Using it would increase my effectiveness at work.
- 6.0 — 5. Using this product would make it easier to do my job.
- 5.75 — 6. I would find this product useful at work.

Perceived Use of use

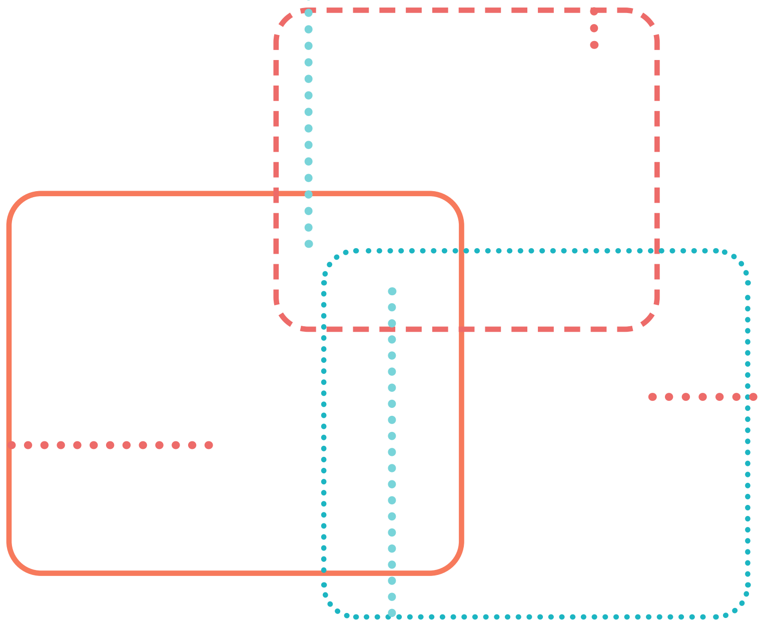
- 5.0 — 7. Learning how to handle the product would be easy for me.
- 4.25 — 8. I would find it easy to let the product do what I want it to.
- 4.25 — 9. My interaction with this product would be clear and smooth.
- 5.75 — 10. I would find this product flexible to work with.
- 5.0 — 11. It would be easy for me to become faster at my work with the product.
- 5.25 — 12. I would find it easy to use.

9.2 Evaluation

9.2.1 Desirability

From a human-centred design perspective, desirability translates into user value. The desirability aspect of the product was in the core of the project throughout the entire graduation project. User research was carried out to understand the current workflow of orthopedic surgeons as well as their attitude and wishes towards the system. Findings were translated into requirements that informed features in the prototype. The value that the Assistant provides is speeding up the clinical documentation with automated suggestions through NLP. Also, the option to add nuanced details in an easy way is valuable for the clinician. Ultimately, partial automation and human control is combined. Apart from testing with end users, an expert in Human-AI collaboration was also consulted for feedback.

"This is a very practical example of human control."
- PhD Candidate in Human-AI collaboration, Philips Design



9.2.2 Feasibility

In terms of technology, internal collaboration was carried out with the CTO of the company. With his inputs, the service blueprint could be finalised. An internal validation was organised with all the developers in the company to discuss the technology aspect. The integrated features for recording the consultation (ASR) is already feasible. The designed features for the Assistant will only be feasible after many consultations are recorded. Organising the recordings on an institution level takes time, therefore can not be concluded yet.

The service blueprint was praised by the tech team to combine the user steps and the technology. From the service blueprint, the extracted user flows are also input for feasibility to show which wireframe leads to another aiding developers to build the product. Usually user flows are handed over for developers in order to make sense of connections between wireframes as well as interactions.

"It is really interesting to see how the software processes influence the user" - CTO

9.2.3 Viability

For viability, the financial aspect was not relevant from the perspective of this thesis. To aid the company, the design phase was scoped with the first implementation steps. The initial plan was to implement a design system (component library) to design for scale. In the end it was more realistic to learn from existing design systems and start building up the company's own. A design system is a set of components and guidelines that aid digital product design and coordinating development. For the template completion interface, a set of components are designed with each interaction state so that buttons only need to be coded once and can be replicated throughout development time. To implement the Assistant, first awareness should be raised about its value in order to start setting up collaboration agreements with hospitals. To contribute, a whitepaper was written to support company communications.

"The storyboard is such a great fit with the healthcare industry. It is immediately clear how the solution works" - Product manager

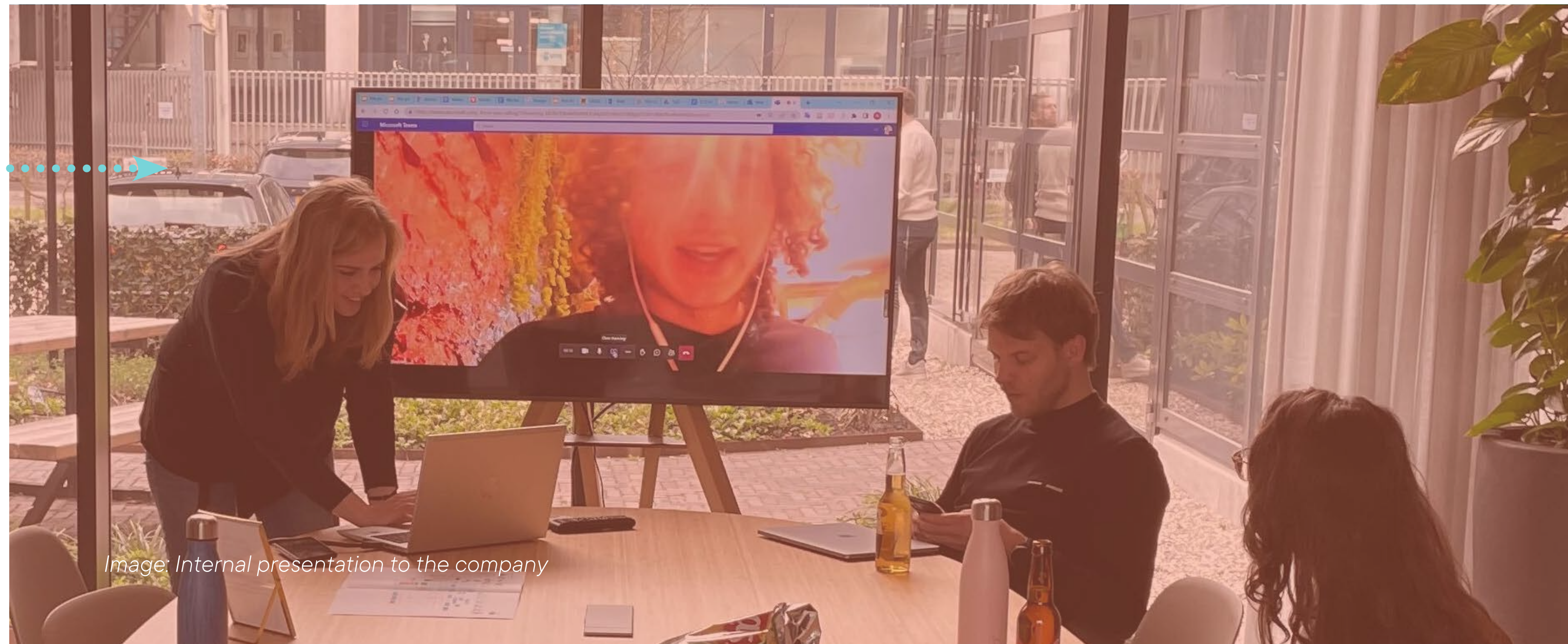


Image: Internal presentation to the company

Chapter 9 takeaways

- User inputs gathered throughout the project influenced many design decisions.
- Presenting to multiple surgeons was an interesting experience to get immediate reactions, hear diverse opinions and facilitate the discussion. Opposing views from users are always hard to deal with as a designer.
- Final user testing was performed with the surgeons I interviewed. Since it is a small sample size, it is advised to take the results as qualitative insights and not draw plain conclusions from the averages.
- For evaluation, all three aspects were considered. Desirability was in the core throughout the project therefore is integrated. The feasibility aspect was validated through an internal presentation and discussions with the technical team. The viability aspect was not central in this project.
- Wrapping up, all the takeaways lead to the final chapter.

Chapter 10

Discussion & the future

The final chapter discusses the project and the presented results. Furthermore limitations are addressed for the interpretation of the project results. Based on the results, recommendations are provided to the company from a UX perspective. Directions for future work are given to build upon the findings. Lastly, the thesis ends with personal reflections.

Image source: Pexels

Concluding

The digital scribe has great potential in order to shift from human-led documentation towards computer-led documentation. The administrative burden is indeed high according to both literature as well as carried out user research. There is a lot of interest in academia about bringing in the technologies into the context, but little exploration on developing the interface of digital scribes. The answeres for the research questions can be found at the end of the relevant chapters.

This Master thesis concludes that combining partial automation with human control is a promising step towards digital scribes. The interface is a proposed way for anamnesis template completion that aids ease of use as well as usefulness. However, fitting individual documentation styles and convincing hospitals for implementation will be a very challenging process. Part of the value proposition is establishing a recording infrastructure which indeed results in many benefits for hospitals, clinicians and patients. Following a step-by-step implementation, through recording consultations value can already be provided as well as it serves as data gathering purposes to further develop the software. The designed interface can already support this step.

For the Assistant, the most value could lie in scoping the digital scribe capabilities into the physical examination part of a consultation. If it is technically feasible, integrating with the EHR is favoured. Providing a portal to the patient to complement the data collection could be valuable to save more time in the consultation, and allow clinicians to spend less time on documentation.

10.1 Discussion

The concept of the digital scribe has great potential but still very futuristic for the healthcare industry. Also, dealing with sensitive data in healthcare always leads to challenges. Organising that aspect and arranging safety aspects of the solution was outside of this thesis scope. It was not anticipated that shadowing in an academic and non-academic hospital will lead to different results. From many discussions, I learnt that the medical community is interested in seeing the differences. The qualitative insights of the consultations could also serve as input for future collaborations.

From literature it was clear that part of the problem is that clinical notes today are unstructured. In the product vision, the Assistant also aims to structure the notes in order to enable partial automation. However, clinicians are more concerned about speed and less about data structure. They are also a bit stubborn to adjust their way of working, even if it is "claimed" to be fitting into their workflow. Fitting individual documentation styles with standardised templates is a difficult aspect of the solution. This leads to a tension to expect during future implementation.

Throughout the project, a small number of users were involved during user research as well as user testing. Increasing the sample size with a diverse group of surgeons could lead to different results. From many impressions it was understood that age and experience level also influences the attitude towards the solution. Furthermore, most of the involved surgeons were familiar with the anamnesis standard. If a surgeon is not familiar, he or she might be more reluctant towards the template completion in general. For this reason, consider the presented averages from the measured ease of use and usefulness with caution.

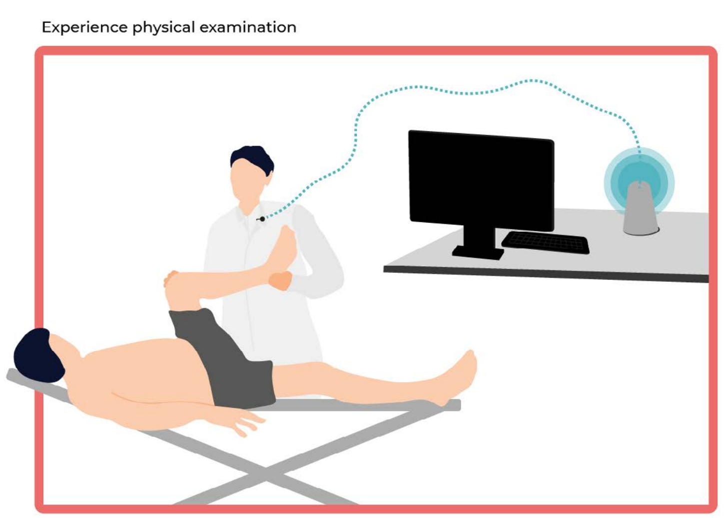


Figure 64: Physical examination

10.2 Recommendations

10.2.1 Future features

The digital scribe is a futuristic concept for end users. Through many discussions I got the impression that part of the implementation will be to communicate the value of the solution, translating it into clear benefits for both the clinician and the patient. Also, since it is a complex technology, it needs careful explanations of what the technology is capable of doing and how it could help during consultations.

As the Assistant evolves, personalising suggestions as well as adding adaptive features are advised. By making it personalised, ideally it could fit individual documentation styles. Once real-time speech processing is possible, the picked up conversations can enable voice commands or voice interactions. During physical examination (figure 64) that could be especially valuable. In the future, keywords should be gathered to aid the development.

With an EHR integration, follow-up documentation tasks could also be addressed. This thesis also presents inputs into what those tasks are. The recorded conversation could serve as a data source to design for an improved workflow and improve the repetitive follow-up tasks (figure 65).

These features are not yet realistic to map out on a roadmap, since it is challenging to estimate how long it will take to get there both from an organisational as well as a technical perspective.

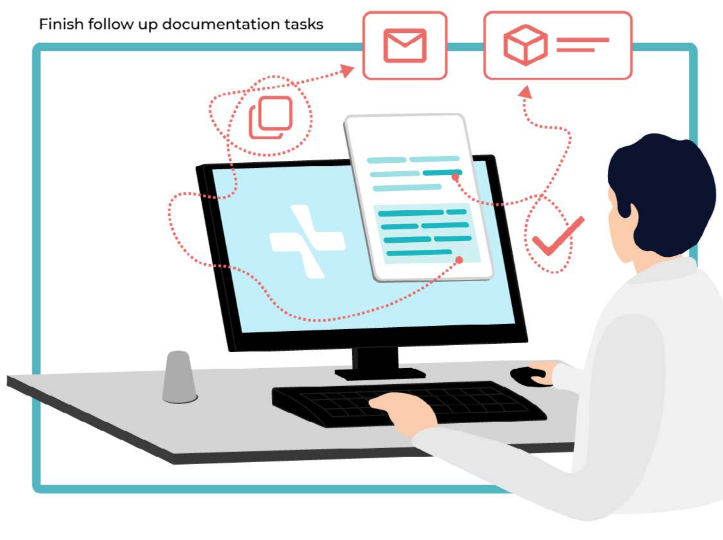


Figure 65: Follow-up documentation tasks

10.2.2 Future research

This thesis focused on the clinician perspective of the digital scribe. From a design research perspective, it could be interesting to investigate and understand the patient's perspectives. By providing the transcript and audio, the data is already of value to the patient. Possible more values could be researched to complement this work.

More research is advised to dive deeper into the trust and ethical aspects of the digital scribe. Trust will be built over time, therefore should be carried out in a longer study. The ethical aspect has a vast theoretical foundation in literature to build upon, but little applied research on the topic.

Above a certain confidence score (figure 66), the Assistant can produce automated suggestions. Future research is necessary to set the threshold for the confidence score. It is also advised to research how the technology influences the user experience with high-fidelity prototypes.

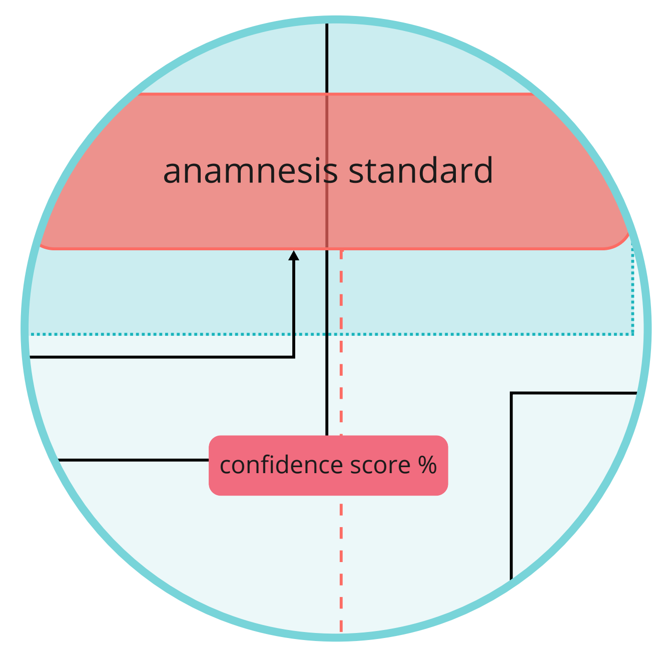


Figure 66: Confidence score from service blueprint

10.3 Reflections

I entered the Masters with a vision to combine user-centred design and data science with a focus on medical design. I truly think this project is a manifestation of my design studies and old and new learnings. Therefore, I am grateful for all the coaches who shaped my path during the past 5 years.

Growth

During this project I wanted to grow skills in service design. Job shadowing was a chosen method I wanted to experiment with. It was quite a challenge to figure out what to look for in order to visualise "useful "observed behaviours. Second time I did it, I could be more strategic with note taking and what to pay attention to.

I am proud of the service blueprint and the fact that it is complex. In the beginning I thought it would just become a flowchart. The tip to combine it with a storyboard made a big difference in the work I delivered and final feedback I got. It is a good lesson to be creative with design methods as well and make it fit with any project I might work in in the future.

This project intensively combined both service design as well as UX design. While both are useful to be skilled at, I feel like my heart is closer to hands-on interface development than journey mapping which clarified throguhout the process.

I wanted to combine UX with AI, and I had to make sense of the complexity. AI is such a dynamic technology to design with and for that excites me. However, innovating in the healthcare context is incredibly hard.

When I started my design studies, one of my goals was to become a designer who is confident in juggling between various design tools. This project required that intensively.

Process

This was the first ever project where I could get medical input every single week, which was a huge help. Initially I planned two rounds of user testing, but I had to be realistic. Clinicians are super busy people so whenever I got the opportunity to meet any, I had to be very straightforward.

During the process sometimes I made decisions without documenting them. I could have saved time at the end if I pay attention to it throughout the process.

I had to give many presentations to different audiences. I was not happy with how my midterm presentation went. It was a good lesson to learn how to filter per audience and tailor the message accordingly.

The role of design in this product is crucial and I think there was a lot of responsibility on me. People thought I will jump into designing the interface right away which was not the case. The challenging aspect is that people want to see things but the biggest part of the job is synthesis.

I truly think good design lies in the details. However, as a designer it is really challenging to know when to stop diving deeper in an activity and switch to another. To be specific, I really wanted to do the service blueprint 'correctly' which got me lost in details. Some feedback and reactions made me realise it probably was not necessary and I lost time with it.

I could have gone deeper into the ethical aspect of the project. I find the angle very interesting and crucial for the topic. However I had to prioritise my to-do list.

Collaboration

There was no other designer in the company which challenged me on a new level and my role in many aspects. I actually enjoy interdisciplinary settings where I learnt that terminology is key to collaboration.

I learnt a lot from working with technical people. It is both super challenging and rewarding to show the value of design (researc). I often realised that I was explaining design methods while other times I was using a lot of design jargons without even noticing.

In my opinion a good UX designer does more than making products look pretty. Yes, it has to look pretty. Personally, I want to produce work where reactions go beyond "It looks good", to a point where clients admit "... and it actually makes sense."

Talking about design is hard. Talking about design with non-designers is even harder. At some point I decided to read the book "Articulating design decisions" which was a game changer during the process. By showing my thinking process and coming up with supporting examples for design decisions, meetings became much smoother. I learnt how to steer conversations from subjective aesthetics to objective functionality.

Over time I built resistance to not get attached to anything I design because design is so subjective. It can always be better and is never really done. You have to embrace that people will critique your work and somehow learn how to enjoy it. I think it is a core skill to develop, to be able to objectively judge your own design and decide on improvement points.

Looking back it gave me a lot of energy to work in a small company and see the immediate client reactions to my work. During this thesis two things motivated me: to provide value to the company and be proud of what I did at the very end.

Personal

To me design means translating: between user, technology, and business - and that is what I did here and aim to keep doing.

I sincerely hope that this work will inspire both the design and the medical community, lead further development of digital scribes and spark discussions around the topic. There is soo much potential, yet still so much to do! Anyway, I am curious to hear any opinions on this project once it is not only me reading these words.

Reka

References

Image source: Pexels

Barr PJ, Bonasia K, Verma K, et al. Audio-/Videorecording Clinic Visits for Patient's Personal Use in the United States: Cross-Sectional Survey. J Med Internet Res. 2018;20(9): 11308.

Beardow, Caiseal & Lomas, Derek & van der Maden, Willem. (2020). Designing Smart Systems: Reframing Artificial Intelligence for Human-centered Designers.

Bensing JM, Deveugele M, Moretti F, Fletcher I, van Vliet L, Van Bogaert M, Rimondini M. How to make the medical consultation more successful from a patient's perspective? Tips for doctors and patients from lay people in the United Kingdom, Italy, Belgium and the Netherlands. Patient Educ Couns. 2011 Sep;84(3):287-93. doi: 10.1016/j.pec.2011.06.008. Epub 2011 Jul 26. PMID: 21795007.

Bodenheimer T, Sinsky C. From triple to quadruple aim: care of the patient requires care of the provider. Ann Fam Med.

Blijleven, V., Koelemeijer, K., & Jaspers, M. (2017). Identifying and eliminating inefficiencies in information system usage: A lean perspective. International Journal of Medical Informatics, 107, 40–47. https://doi.org/10.1016/j.ijmedinf.2017.08.005

Brownlee, J. (2019, August 7). What is natural language processing? Machine Learning Mastery. Retrieved October 8, 2021, from https://machinelearningmastery.com/natural-language-processing/.
van Bruinessen, I., Leegwater, B., & Dulmen, A. M. (2017). When patients take the initiative to audio-record a clinical consultation. Patient Education and Counseling, 100. https://doi.org/10.1016/j.pec.2017.03.001

van Buchem, M. M., Boosman, H., Bauer, M. P., Kant, I. M. J., Cammel, S. A., & Steyerberg, E. W. (2021). The digital scribe in clinical practice: A scoping review and research agenda. Npj Digital Medicine, 4(1), 57. https://doi.org/10.1038/s41746-021-00432-5

Calisto, F. M., Santiago, C., Nunes, N., & Nascimento, J. C. (2021). Introduction of human-centric AI assistant to aid radiologists for multimodal breast image classification. International Journal of Human-Computer Studies, 150, 102607. https://doi.org/10.1016/j.ijhcs.2021.102607

Chiu, C.-C., Tripathi, A., Chou, K., Co, C., Jaitly, N., Jaunzeikare, D., … Zhang, X. (2018). Speech recognition for medical conversations. ArXiv:1711.07274 [Cs, Eess, Stat]. Retrieved from http://arxiv.org/abs/1711.07274

Clark, L., Doyle, P., Garaialde, D., Gilmartin, E., Schlögl, S., Edlund, J., Aylett, M., Cabral, J., Munteanu, C., & Cowan, B. (2018, October 16). The state of speech in HCI: Trends, themes and challenges. arXiv.org. Retrieved December 15, 2021, from https://arxiv.org/abs/1810.06828

Coiera, E., Kocaballi, B., Halamka, J., & Laranjo, L. (2018). The digital scribe. NPJ digital medicine, 1, 58. https://doi.org/10.1038/s41746-018-0066-9

Dommershuijsen, L. J., Dedding, C. W. M., & Van Bruchem-Visser, R. L. (2021). Consultation Recording: What Is the Added Value for Patients Aged 50 Years and Over? A Systematic Review. Health Communication, 36(2), 168–178. https://doi.org/10.1080/10410236.2019.1669270

Du, N., Wang, M., Tran, L., Lee, G., & Shafran, I. (2019). Learning to Infer Entities, Properties and their Relations from Clinical Conversations. Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), 4979–4990. Hong Kong, China: Association for Computational Linguistics. https://doi.org/10.18653/v1/D19-1503

Dugdale, D. C., Epstein, R., & Pantilat, S. Z. (1999). Time and the patient-physician relationship. Journal of general internal medicine, 14 Suppl 1(Suppl 1), S34–S40. https://doi.org/10.1046/j.1525-1497.1999.00263.x

Dusek, H. L., Goldstein, I. H., Rule, A., Chiang, M. F., & Hribar, M. R. (2021). Clinical documentation during scribed and non-scribed ophthalmology office visits. Ophthalmology Science, 100088. https://doi.org/10.1016/j.xops.2021.100088

Elwyn, G., Barr, P. J., & Grande, S. W. (2015, August 1). Patients recording clinical encounters: A path to empowerment? assessment by mixed methods. BMJ Open. Retrieved October 24, 2021, from https://bmjopen.bmj.com/content/5/8/e008566.

Elwyn, G. (2014, March 11). "patientgate"-digital recordings change everything. The BMJ. Retrieved October 24, 2021, from https://www.bmj.com/content/348/bmj.g2078.
Enarvi, S., Amoia, M., Teba, M. D.-A., Delaney, B., Diehl, F., Hahn, S., … Ramamurthy, R. (2020, July). Generating Medical Reports from Patient-Doctor Conversations Using Sequence-to-Sequence Models. 22–30. https://doi.org/10.18653/v1/2020.nlpmc-1.4

Gellert, G. A., Ramirez, R., & Webster, S. L. (2015). The Rise of the Medical Scribe Industry: Implications for the Advancement of Electronic Health Records. JAMA, 313(13), 1315–1316. https://doi.org/10.1001/jama.2014.17128

Ghatnekar, S., Faletsky, A., & Nambudiri, V. E. (2021). Digital scribe utility and barriers to implementation in clinical practice: A scoping review. Health and Technology, 11(4), 803–809. https://doi.org/10.1007/s12553-021-00568-0

Gidwani, R., Nguyen, C., Kofoed, A., Carragee, C., Rydel, T., Nelligan, I., … Lin, S. (2017). Impact of Scribes on Physician Satisfaction, Patient Satisfaction, and Charting Efficiency: A Randomized Controlled Trial. Annals of Family Medicine, 15(5), 427–433. https://doi.org/10.1370/afm.2122

Goddard, K., Roudsari, A., & Wyatt, J. C. (2012). Automation bias: A systematic review of frequency, effect mediators, and mitigators. Journal of the American Medical Informatics Association : JAMIA. Retrieved October 8, 2021, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3240751/.

Jamal, N., Shanta, S., Mahmud, F., & Sha'abani, M. N. A. H. (2017, September 14). Automatic speech recognition (ASR) based approach for speech therapy of aphasic patients: A Review. AIP Publishing. Retrieved October 8, 2021, from https://aip.scitation.org/doi/10.1063/1.5002046.

Jeblee, S., Khattak, F. K., Crampton, N., Mamdani, M. & Rudzicz, F. Extracting relevant information from physician-patient dialogues for automated clinical note taking. In Proc. of the Tenth International Workshop on Health Text Mining and Information Analysis (LOUHI), 65–74 (Association for Computational Linguistics, 2019).

Jeblee, S., Khattak, F. K., Crampton, N., Mamdani, M., & Rudzicz, F. (n.d.). Extracting relevant information from physician-patient dialogues for automated clinical note taking. ACL Anthology. Retrieved October 8, 2021, from https://aclanthology.org/D19-6209/.

Joshi, A., Farberov, M., Demissie, S. et al. Attitudes of Physicians to Recording Clinical Encounters: Responses to an Online Survey. J GEN INTERN MED 35, 942–943 (2020). https://doi.org/10.1007/s11606-019-05127-y

Jylkäs, T., Augsten, A., & Miettinen, S. (2019, June 24). From Hype to Practice: Revealing the Effects of AI in Service Design.

JP;, E. J. M. W. (n.d.). Technology as friend or foe? Do electronic health records increase burnout? Current opinion in anaesthesiology. Retrieved October 8, 2021, from https://pubmed.ncbi.nlm.nih.gov/29474217/.

Kanda, N., Horiguchi, S., Fujita, Y., Xue, Y., Nagamatsu, K., & Watanabe, S. (2019). Simultaneous Speech Recognition and Speaker Diarization for Monaural Dialogue Recordings with Target-Speaker Acoustic Models. ArXiv:1909.08103 [Cs, Eess]. http://arxiv.org/abs/1909.08103

King, J., Patel, V., Jamoom, E. W., & Furukawa, M. F. (2014). Clinical Benefits of Electronic Health Record Use: National Findings. Health Services Research, 49(1pt2), 392–404. https://doi.org/10.1111/1475-6773.12135

Kocaballi, A. B., Ijaz, K., Laranjo, L., Quiroz, J. C., Rezazadegan, D., Tong, H. L., Willcock, S., Berkovsky, S., & Coiera, E. (2020). Envisioning an artificial intelligence documentation assistant for future primary care consultations: A co-design study with general practitioners. Journal of the American Medical Informatics Association, 27(11), 1695–1704. https://doi.org/10.1093/jamia/ocaa131

Lacson, R. C., Barzilay, R., & Long, W. J. (2006). Automatic analysis of medical dialogue in the home hemodialysis domain: Structure induction and summarization. Journal of Biomedical Informatics, 39(5), 541–555. https://doi.org/10.1016/j.jbi.2005.12.009

Malhi, A., Knapic, S., & Främling, K. (2020). Explainable Agents for Less Bias in Human-Agent Decision Making. Explainable, Transparent Autonomous Agents and Multi-Agent Systems: Second International Workshop, EXTRAAMAS 2020, Auckland, New Zealand, May 9–13, 2020, Revised Selected Papers, 12175, 129–146. https://doi.org/10.1007/978-3-030-51924-7_8

Mani, A., Palaskar, S., & Konam, S. (2020, July). Towards Understanding ASR Error Correction for Medical Conversations. 7–11. https://doi.org/10.18653/v1/2020.nlpmc-1.2

Onzinnige regels drukken nog steeds zwaar op zorgverleners. (2021, March 11). Retrieved 22 June 2021, from De stem en steun van zorgverleners website: https://www.vvaa.nl/voor-leden/nieuws/ontregelmonitor2021

Rajkomar, A., Kannan, A., Chen, K., Vardoulakis, L., Chou, K., Cui, C., & Dean, J. (2019). Automatically Charting Symptoms From Patient-Physician Conversations Using Machine Learning. JAMA Internal Medicine, 179(6), 836. https://doi.org/10.1001/jamainternmed.2018.8558

Rosenbloom, S., Denny, J., Qi, W., Lorenzi, N., Stead, W., & Johnson, K. (2011). Data from clinical notes: A perspective on the tension between structure and flexible documentation. Journal of the American Medical Informatics Association : JAMIA, 18, 181–186. https://doi.org/10.1136/jamia.2010.007237

Shachak, A., & Reis, S. (2009). The impact of electronic medical records on patient-doctor communication during consultation: A narrative literature review. Journal of Evaluation in Clinical Practice, 15(4), 641–649. https://doi.org/10.1111/j.1365-2753.2008.01065.x

Shafey, L. E., Soltau, H., & Shafran, I. (2019). Joint Speech Recognition and Speaker Diarization via Sequence Transduction. Interspeech 2019, 396–400. ISCA. https://doi.org/10.21437/Interspeech.2019-1943

Sinsky C, Colligan L, Li L, Prgomet M, Reynolds S, Goeders L, Westbrook J, Tutty M, Blike G. Allocation of Physician Time in Ambulatory Practice: A Time and Motion Study in 4 Specialties. Ann Intern Med. 2016 Dec 6;165(11):753-760. doi: 10.7326/M16-0961. Epub 2016 Sep 6. PMID: 27595430.

Statement initiatiefnemers ORDZ. (2021, May 23). Retrieved 23 June 2021, from De stem en steun van zorgverleners website: https://www.vvaa.nl/voor-leden/nieuws/statement-initiatiefnemers-ontregel-de-zorg

Turley, D. P., & Metcalfe, N. H. (2020). Patients recording their clinical consultations: A new challenge for medical ethics. InnovAiT: Education and Inspiration for General Practice, 13(5), 306–310. https://doi.org/10.1177/1755738020907358

Tai-Seale, M. et al. Electronic health record logs indicate that physicians split time evenly between seeing patients and desktop medicine. Health Aff. 36, 655–662 (2017).

Tajirian, T., Stergiopoulos, V., Strudwick, G., Sequeira, L., Sanches, M., Kemp, J., … Jankowicz, D. (2020). The Influence of Electronic Health Record Use on Physician Burnout: Cross-Sectional Survey. Journal of Medical Internet Research, 22(7), e19274. https://doi.org/10.2196/19274

Pearl, C. (2016). Designing Voice User Interfaces: Principles of Conversational Experiences (1st ed.). O'Reilly Media, Inc.

Pros and cons of letting patients record doctor visits. (n.d.). Retrieved December 15, 2021, from https://amednews.com/article/20121105/profession/311059956/5/

Quiroz, J. C., Laranjo, L., Kocaballi, A. B., Berkovsky, S., Rezazadegan, D., & Coiera, E. (2019, November 22). Challenges of developing a digital scribe to reduce clinical documentation burden. Nature News. Retrieved October 8, 2021, from https://www.nature.com/articles/s41746-019-0190-1/.

Tsulukidze, M., Grande, S. W., Thompson, R., Rudd, K., & Elwyn, G. (2015). Patients Covertly Recording Clinical Encounters: Threat or Opportunity? A Qualitative Analysis of Online Texts. PLOS ONE, 10(5), e0125824. https://doi.org/10.1371/journal.pone.0125824

Tsulukidze, M., Durand, M.-A., Barr, P. J., Mead, T., & Elwyn, G. (2014). Providing recording of clinical consultation to patients - a highly valued but underutilized intervention: A scoping review. Patient Education and Counseling, 95(3), 297–304. https://doi.org/10.1016/j.pec.2014.02.007

Walker KJ, Dunlop W, Liew D, Staples MP, Johnson M, Ben-Meir M, Rodda HG, Turner I, Phillips D. An economic evaluation of the costs of training a medical scribe to work in Emergency Medicine. Emerg Med J. 2016 Dec;33(12):865-869. doi: 10.1136/emermed-2016-205934. Epub 2016 Jun 28. PMID: 27352788.

Weldring, T., & Smith, S. M. (2013). Patient-Reported Outcomes (PROs) and Patient-Reported Outcome Measures (PROMs). Health services insights, 6, 61–68. https://doi.org/10.4137/HSI.S11093

Websfarm_ltd. (2021, April). 5 natural language processing trends in 2021. Retrieved from https://www.websfarm.net/5-natural-language-processing-trends-in-2021/

Yu, Han & Shen, Zhiqi & Leung, Cyril & Miao, Chunyan & Lesser, Victor. (2013). A Survey of Multi-Agent Trust Management Systems. IEEE Access. 1. 35-50. 10.1109/ACCESS.2013.2259892.

Image sources

https://www.istockphoto.com/nl/search/2/image?phrase=doctor+walking+profile
https://www.pexels.com/photo/a-doctor-reading-a-medical-chart-6129041/
https://www.health3-0.com/patient-outcomes/what-are-patient-outcomes/
https://www.buildingbetterhealthcare.com/news/article_page/Interview_How_artificial_intelligence_is_transforming_the_doctor-patient_relationship/153948
https://www.pexels.com/photo/people-woman-sitting-technology-7089401/
https://www.pexels.com/photo/person-s-hands-on-top-of-macbook-air-1181307/
https://www.pexels.com/photo/pile-of-folders-357514/
https://www.pexels.com/photo/close-up-photo-of-programming-of-codes-546819/
https://www.pexels.com/photo/cheerful-senior-mother-and-adult-daughter-using-smartphone-together-3791664/

Thank you