AALBORG UNIVERSITY
STUDENT REPORT

Title: *Envisioning the Future of AI Chatbots in a General Practice:*
a Techno-Anthropological Perspective

Semester:	**10. Semester**
Semester theme:	**Master Thesis**
Project period:	**January to June 2023**
ECTS:	**30**
Supervisor:	**Mette Ebbesen**
Project group:	**IxD601f19**

Nikolaj Arildsen Jakobsen 20163851

Zuzana Hruba 20175664

Silas Jonatan Skjønnemand 20182443

Pages: 79.14

This thesis investigates how a general practice in Northern Jutland envisions the future in regards to the emergence of AI chatbots such as ChatGPT. This study was prompted by the recent adoption of AI chatbots in healthcare, driven by the COVID-19 pandemic, and the subsequent release of ChatGPT. This emergence calls for new knowledge and ethical considerations. By exploring the social and cultural aspects that influence the construction and interpretation of such technological tools, the study aims to understand their alignment with the needs, values, and preferences of the involved social group(s). The exploratory case study employed in this research utilises both the theoretical approach of Social Construction of Technology (SCOT) and the methodology of ethnographic co-envisioning. This combined approach allows for an examination of how the social groups at the general practice interpret and envision their future with AI chatbots, and what implications the technology might present. The findings of the study seek to inform the sustainable and responsible implementation of AI chatbots in healthcare settings, enabling related actors to make informed decisions about their possible integration. The thesis offers a techno-anthropological perspective on the ongoing development of AI chatbots in healthcare, emphasising the importance of context-specific knowledge and social influence on development of the technology.
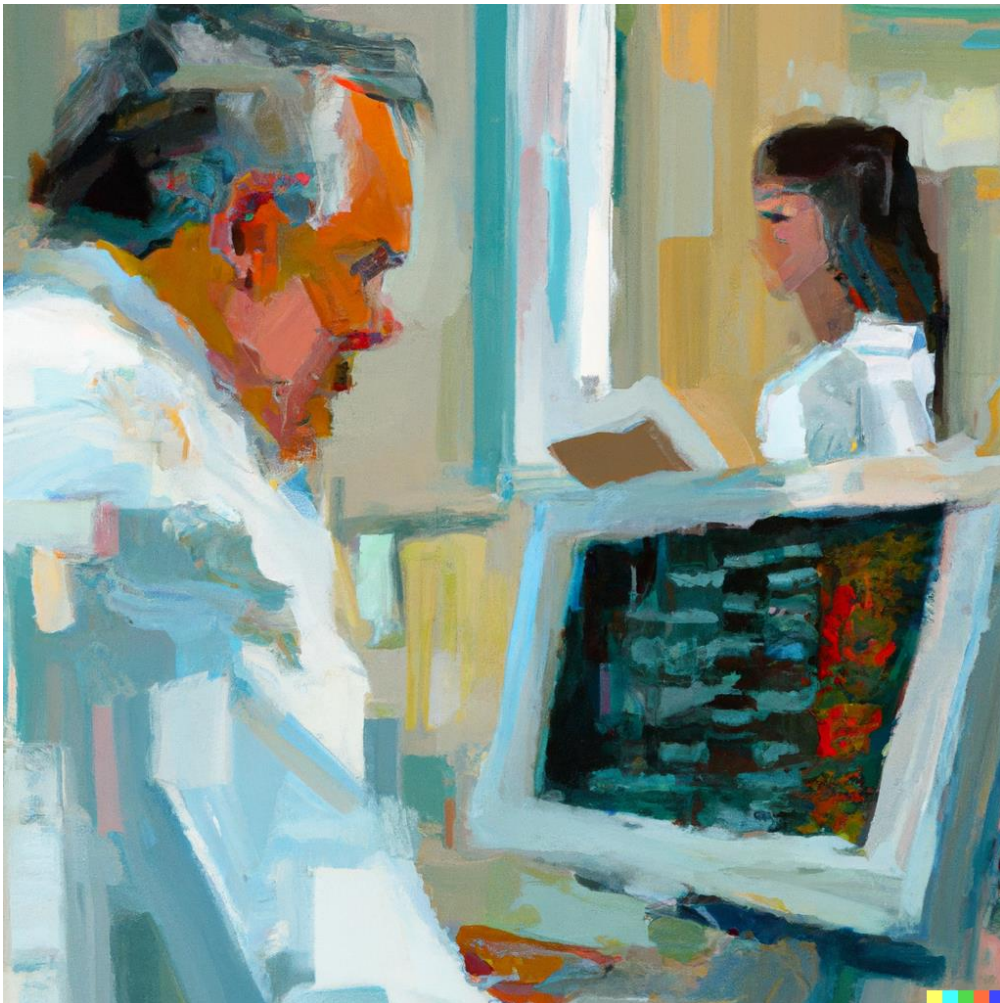
AALBORG UNIVERSITET

ENVISIONING THE FUTURE OF AI CHATBOTS IN A GENERAL PRACTICE: A TECHNO-ANTHROPOLOGICAL PERSPECTIVE

A CASE STUDY

Nikolaj Arildsen Jakobsen, Silas Jonatan Skjønnemand, Zuzana Hruba
Department of Planning, Techno-Anthropology

02-06-2023

INDHOLDSFORTEGNELSE

INTODUCTION

The recent release of ChatGPT on November 30, 2022 sparked a public debate across various fields, as this tool became widely used shortly after its release with the provided open access to a technology that is unlike any other previously accessible to the public (De Angelis et al. 2023). A field of particular importance is healthcare, where human knowledge and skills are essential for saving lives. Furthermore, the COVID-19 pandemic has led to a rapid digital leap in healthcare (Parviainen et al. 2022). This digital leap, which has brought about significant changes in the healthcare industry, has accelerated the adoption and implementation of AI chatbots. These chatbots were deployed to provide patients with up-to-date information, guidelines, and advice on the virus, among other things (ibid.). The pandemic created an urgent need for accessible health information and services, and chatbots were seen as a potential solution to address this demand while alleviating the pressure on healthcare professionals.

The release of ChatGPT has introduced a new dimension to the healthcare industry's adoption of AI chatbots. The capabilities of ChatGPT extend beyond traditional rule-based systems, enabling more dynamic and conversational interactions with users. The ability to understand and respond to natural language makes ChatGPT a promising tool for enhancing the accessibility and delivery of healthcare information and services.

The primary goal of this thesis is to explore the future of rapidly evolving AI chatbots through a case study by co-envisioning in the context of a general practice in Northern Jutland. By examining the social and cultural aspects that influence the construction of the technological artefact of AI chatbots, we aim to understand how these factors can shape the interpretation of AI chatbots in a way that aligns with the needs, values, and preferences of the social group(s) involved.

Additionally, this study seeks to provide insights that can inform the sustainable and responsible implementation of AI chatbots in general practice or healthcare settings in the future. By considering the findings of this thesis, stakeholders can make informed decisions regarding the integration of AI chatbots, ensuring that they are ethically and responsibly deployed to enhance patient care and respect the patient-practitioner relationship as well as the overall social and cultural contexts in which they are utilised.

The future of rapidly evolving AI chatbots in the context of a general practice in Northern Jutland presents both opportunities and challenges. Understanding how social and cultural aspects within the social group(s) can inform the construction of the technological artefact is crucial to ensure a user-centred and contextually appropriate implementation of AI chatbots. By investigating how AI chatbots are socially constructed, this study aims to contribute to the co-envisioning of the future of these technologies in general practice.

2. BACKGROUND

To study the influence of AI chatbots on a general practice, we find it important to understand the wider context of AI chatbots, as well as the status quo of general practices particularly workflows and relations. As such, we have reviewed relevant literature and this chapter will present its pertinent findings, for us to obtain a knowledge base, and frame our empirical study. The findings of this background chapter will ultimately inform our problem formulation, establishing the foundation of our empirical case study.

2.1 THE EMERGENCE OF ARTIFICIAL INTELLIGENCE CHATBOTS IN HEALTH CARE

Artificial intelligence (AI) chatbots, which are advanced computational programs that employ natural language processing and machine learning algorithms to simulate human-like dialogues, have shown promising results in the health care industry. By comprehending user inputs and generating contextually appropriate responses derived from pre-existing knowledge bases or experiential learning, AI chatbots can be seamlessly integrated into various platforms such as websites, messaging applications, and mobile devices. However, there are still limitations and possible dangers regarding the use of AI chatbots in healthcare that must be considered. These limitations and dangers will be presented later in this chapter.

In the article *Conversational Artificial Intelligence for Spinal Pain Questionnaire: Validation and User Satisfaction* (Nam et al. 2022), the authors discuss the development and validation of a Spoken Dialogue System (SDS), which is a type of conversational AI that employs natural language processing and machine learning algorithms to simulate human-like dialogues, for a spinal pain questionnaire. The SDS showed favourable results in terms of user satisfaction and performance accuracy, demonstrating the potential of conversational AI for diagnosing and remotely monitoring patients. However, improvements are needed in voice recognition technology and the collection of dialog sets between patients and medical staff to make the technology more effective (ibid.).

Despite their potential, discussions of limitations of AI chatbots in health care acknowledge that they will still fall short in terms of their ability to understand and respond to human emotions and expressions effectively (Powell 2019). This lack of emotional intelligence can impact trust, which is a crucial factor in health care interactions. Thus, AI should complement, not replace, medical professionals.

Parviainen, et al. (2022) conclude that a systemic perspective should be considered when implementing chatbots in health care. Ethico-political approaches are needed in professional ethics, as chatbots can become entangled with clinical practices in complex ways. Insufficient consideration of chatbot implementation can lead to poor professional practices and long-term side effects, harming both professionals and patients (Parviainen, et al. 2022). As an emerging phenomenon in health care, AI chatbots hold significant promise for improving patient care and reducing the burden on health care professionals. However, it is imperative to address the challenges and risks accompanying this technology, such as patient trust and ethical considerations, to warrant that AI chatbot adoption results in favourable outcomes for both patients and health care professionals (Nam et al. 2022; Powell 2019; Parviainen, et al. 2022).

In summary, the adoption of AI chatbots has been accelerated in the health care industry in recent years. While AI chatbots have shown promising results in providing accessible health information and reducing the burden on health care professionals, limitations such as the lack of emotional intelligence must be considered. To ensure that AI chatbot adoption results in

favourable outcomes for both patients and health care professionals, challenges, and risks such as patient trust and ethical considerations must be addressed (Nam et al. 2022; Powell 2019; Parviainen, et al. 2022).

2.2 STATE OF THE ART

The following section will introduce different cases of how chatbots have been implemented in health care context. This is done to develop insights into state-of-the-art of the technology, as illustrated in certain related literature, and by extension what opportunities for implementation, researchers have previously considered.

Since the initiation of this project concerns a rapid development of artificial intelligence (AI), and not all the chatbots found while investigating the literature, there is value in understanding how a distinction can be made between a non-AI chatbot and an AI chatbot. To explore this aspect, Kelley (2003) applied concepts stemming from biology and human cognition onto computational models, more specifically the project Adaptive Character of Thought-Rational architecture initiated by the US Army Research Laboratory (ARL) and Human Research and Engineering Directorate (HRED). They introduce a continuum of symbolic and sub-symbolic paradigms as a theoretical perspective that can be applied to computational architecture (ibid.).

The symbolic architectures can be likened to a process of manipulating symbols, with Kelley (2003) exemplifying this process through a sentence between humans; *Mary loves* Sam. In the example, humans, or symbolic computational models, can understand the different elements of the sentence as symbols as a series of symbols. *Mary* is understood as a symbol representing a person, who has an emotional attachment or connection, symbolically represented in *loves*, to the recipient of said affection, symbolised as *Sam* (ibid.).

The sub-symbolic architectures function in parallel, as opposed to the serial approach of symbolic architectures, likened to a collection of neurons in the brain working simultaneously to produce an output (Kelley 2003), hence also the nomenclature of neural networks. This means that sub-symbolic architectures are essentially able to autonomously learn by weighing an input on different internal connections and parameters to produce an output. These parameters and connections must be defined (or trained) beforehand (ibid.), requiring an exponential amount of training according to the complexity of a given input.

The dichotomy between the two architectures can be understood by expanding the example input given in *Mary loves Sam*, which would be relatively easy for a sub-symbolic architecture to process. However, due to the parallel process, the sentence *Mary loves Sam and Bob hates Sally* is considerably more difficult for a sub-symbolic architecture to handle, whereas the procedure of a symbolic architecture would process the sentence serially, relying on the sequence of the symbols (ibid.).

Although the concepts are relatively rudimentary, and based on rudimentary technology, considering the rapid evolution in recent years, the distinctions remain consistent with the contemporary chatbots found in health care. In this project, the distinction is used to identify the complexity of state-of-the-art chatbots and evaluate the strengths and weaknesses of the approaches. For the purpose of continuity, we will then equate chatbots using symbolic architecture to non-AI chatbots, and chatbots using sub-symbolic architecture to AI chatbots. Lastly, we classify these models from the information given in the specific studies, and as such, we do not account for possible changes post-publication or insufficient clarifications on the architecture of the specific chatbots.

2.2.1 NON-AI CHATBOTS

The classification of non-AI chatbots largely refer to chatbots that either deal with predetermined prompts, or can discern simple prompts, or questions, in order to provide a pre-specified answer relating to keywords. These are arguably the more efficient procedures, when it comes to tailored systems, however they lack the flexibility, and potential complex understanding of AI chatbots.

Alturaiki et al. (2022) present a rule-based expert system, and chatbot, to disseminate education to patients about β-thalassemia. In defining the chatbot as an expert system, it is meant that the knowledge base, and the interaction with a patient is designed by experts. This is to say, that the logic of the system is determined by an engineer, and the domain specific health care knowledge is gathered from human experts, books, and guidelines (ibid.). The system is rule-based, as a clear opposite to Natural Language Processing (NLP)-based, meaning it uses predefined categories, and rules in the interaction with users (ibid.), classifying it as a non-AI chatbot. In this case, the chatbot works as a question-and-answer system, where the chatbot asks a question, prompting the user to select an answer from a drop-down list of options. Since the rule-based expert system follows a rigid process, this factual knowledge is relatively easy to verify, given the proper experts. The chatbot was tested for user satisfaction, on 34 patients, where it was determined that the chatbot is user-friendly, accessible, and affordable, and makes managing β-thalassemia easier and more efficient (ibid.). However, the system seemed to lack in terms of patients learning the timetable for their follow-up assessments from the chatbot, learning the timetable for blood-tests, and learning their normal blood test results. Alturaiki et al. (2022) attribute these shortcomings to the short-term testing of their prototype. They underline that chatbots, as a technology in health care, has a large potential, particularly for long-term patients, since it can reduce the amount of hospital visits, which can be uncomfortable for certain patients (ibid.).

Schario et al. (2022) provide two case-based examples of the use of a chatbot, which they define as a software product, leading users in a conversation through text or audio-based methods. They focus on a chatbot technology used by University Hospitals of Cleveland, Ohio, in post-acute and chronic conditions. The chatbot includes a number of automated conversation pathways, specific to health conditions, that a care manager can provide to a patient (ibid.). The chatbot works through text-messaging, email, or a web-based platform, the choice of which is decided by the care manager. Within the given pathway, the chatbot will then automate the collection, analysis, and usage of patient data through a structured conversation that also assists in patient-education for their specific condition. The data is collected from the conversations, and processed through an algorithm that scores data, collected from the patient historically, as either, passing (green), low (rising, yellow) or high (red) risk. On completion of the interaction with the chatbot, the documentation of the conversation is entered into a care management platform and notifies the care manager of high-risk responses (ibid.). This enables the care manager to prioritise patients with most needs, and ultimately saves time spent on repetitive, low skill tasks. The two patient-case examples given by Schario et al. (2022) related to a man in his 50's with congestive heart failure, and a woman in her 70's chronic obstructive pulmonary disease, and congestive heart failure, both patients were assessed with a high risk of readmission.

Schario et al. (2022) conclude that a chatbot can help patients remotely, in terms of managing their conditions, and work through challenges, which can benefit patients who may want minimal interaction with the health care organisation. These patients are provided a safety net, a feeling of partnership, and a connectedness free from judgement, while also disseminating education and gathering clinical data, some of which would not be openly shared with health care personnel (ibid.).

2.2.2 AI CHATBOTS

What we classify as AI chatbots, relates to the underlying more complex neural network-based approach of sub-symbolic architecture, as opposed to the symbolic architecture. These approaches are more prevalent in the recent background, which may stem from development within computational power, and a focus on optimising the architecture related to training the models.

Within the field of AI, there has long been a focus of research in natural language processing, with the aim of building systems that can produce text in a human like manner. Recently, Large Language Models (LLM) such as Generative Pre-trained Transformer (GPT) have been developed, which are pre-trained on massive amounts of text data and fine-tuned for specific downstream tasks. These models have shown impressive performance on a range of natural language processing tasks, including text generation, language translation, and question answering (Radford et al. 2018; Brown et al. 2020).

De Angelis et al. (2023) provide an introduction into the technology behind ChatGPT, which initiated our curiosity regarding AI chatbots. ChatGPT is a language model (LLM) trained by the company OpenAI and was released in a free research preview on November 30th, 2022 (ibid.). The AI chatbot was optimised for human-like interactions and dialogue. This AI chatbot spread astonishingly fast, with over one million users in the first 5 days following its release. The model of the Generative Pretrained Transformer (GPT), on which ChatGPT is based, was one of the two *game-changing* NLP architectures introduced in 2018. These two models, GPT and the Bidirectional Encoder Representations from Transformers (BERT), were based on the self-attention network architecture, also known as Transformer, to bypass issues of NLP models previously being trained using supervised learning for specific tasks (ibid.). The new architecture presented the possibility for semi-supervised training, which expanded the capabilities, and opportunities for the technology. De Angelis et al. (2023) express a worry about medical dissemination that ChatGPT might present. With the precedent of misinformation on social media during the Covid-19 pandemic, ChatGPT might introduce new challenges to research and public health, in a novel concept that they call an *AI-driven infodemic*.

Kim et al. (2022) provide a general insight into the impact of artificial intelligence on the health care system in the Republic of Korea during the COVID-19 pandemic. The focus is on NLP solutions such as chatbots or clinical decision support systems, and the utility they can provide as a model for predicting medical specialty. Using a pre-trained language representation model, developed by Google, the Bidirectional Encoder Representations from Transformers (BERT), they propose an NLP model that can take input from a user, in the form of a question text, and tag the appropriate medical specialty for the user. Comparing two competitive deep-learning-based NLP models, they perform non-overlapping training and tests of the performance, concluding that their model generally outperforms those in terms of accurate and precise predictions (Kim et al. 2022, 4-5). Additionally, they apply the models on 27 different specialties, in the form of case-experiments, which gives convenient insights into the accuracy of both the competing models, and their proposed models. Comparing the accuracy of these models can provide some insights into the strengths and weaknesses of chatbots for clinical decision support between certain specialties. All models performed well for specialties such as dentistry, plastic surgery, Urology, and psychiatry, while general surgery, family medicine, nephrology and emergency medicine exemplified some of the worst accuracies for the models (ibid.). A worry concerning the low performance within emergency medicine is emphasised, as the potential impact and importance of accuracy are particularly salient in that specialty (ibid.). They attribute the low performance within general surgery and family medicine to the medical specificity, since those specialties deal with a wide range of organs and systems, compared to the rest of the specialties (ibid.).

Plastic surgery was one of the more accurate specialties, and Avila et al. (2023) illustrates how that might be the case, in a more specialist-centric evaluation of a telephone-accessed artificial intelligent virtual assistant (AIVA), simulating human conversation to answer preoperative frequently asked questions (FAQs). Time spent on patient education, involving instructions and counselling in relation to plastic surgery, can vary from 19% to 48% of the dedicated office time per patient, and some surgeons think that this process disrupts the patient-practitioner relationship (ibid.). Moreover, since patient education largely concerns repetitive dissemination of information, remaining largely the same between patients, FAQ data is easily generated, which can then be used as training for an AIVA, saving time for the surgeons, and providing a convenient entry point for patients (ibid.). In this case, the chatbot was developed to be available through text-messages or a telephone call, introducing the element of voice output and input, which was seemingly a successful element, with patients scoring a collective 98.5% accuracy and a general approval of the AIVA of 96.1%. However, the feedback from the patients deviated when it came to whether they would find the chatbot useful in assisting them in choosing a surgical procedure, with 42.3% disagreeing that it could serve this purpose.

Pharmacology is an area that was not introduced by Kim et al. (2022), however in the same context of the COVID-19 pandemic. Daniel et al. (2022) scope in on assisting the decision making of caregivers and pharmacy technicians using an AI chatbot to avoid medication errors in French hospitals. The point of the chatbot is to enable dissemination of information about medication 24 hours of the day. Caregivers risk errors, particularly during late shifts at the hospital, since pharmacy technicians may not be available for pharmaceutical advice, leading to accessing information on poorly documented websites due to time pressures (ibid.). This became more critical during the COVID-19 pandemic, since pharmaceutical misinformation, particularly surrounding vaccines, had become abundant on websites and forums. Daniel et al. (2022) design a proof-of-concept chatbot utilising AI to disseminate pharmaceutical information, testing the design with 20 caregivers from four different services. The design was generally rated highly, as a beneficial tool to reduce time spent on calling pharmacy technicians, however primary issues related to ergonomics in using the chatbot and a limited database, which, according to Daniel et al. (2022) was mended. Miner et al. (2019) introduce different facets of how conversational AI can be introduced in psychotherapy, bringing up the interesting aspect of collaboration, or relationships, between AI systems, patients, general practitioners, and medical secretaries, which will be elaborated on later (2.3 Relations). Their study introduces conversational AI to help address an increasing insufficiency of general practitioners by undertaking or easing some of the tasks surrounding the practice of psychotherapy. Specifically, conversational AI has the potential to replace some of the visits that do not include talk therapy and might even be able to provide a better quality in terms of complete evaluation and dissemination of psychotherapy, while avoiding the feeling of stigmatisation (ibid.). However, they highlight safety, trust, and oversight as the first crucial steps in implementing conversational AI in psychotherapy, particularly since it relates to vulnerable people.

Easton et al. (2019) elaborate on the complexity of certain health care practices, introducing the concepts of comorbid mental health and physical long-term conditions (LTC), pointing to increasingly insufficient availability of mental health treatment. They underline how mental health can have a significant clinical impact on the general health of people with physical LTC, leading to a greater consumption of costly health care (ibid.). Furthermore, the task of health care is increasingly being moved to the community, emphasising taking control of one's own health as a focal theme, which is something that requires techniques on the part of the patients to be effective. In contrast to other cases of state-of-the-art chatbots, their research revolves around co-design and the process of meeting the users' requirements when

developing and proposing a solution. Their proposed solution consists of a virtual agent for managing these comorbidities and conditions, Avachat (close homophone to British have a chat) (ibid.). Through the process of co-design, they run into some important considerations concerning state-of-the-art technologies relating to health care; reinforcing the points made by Miner et al. (2019), that safety, and trust are particularly important when developing technologies affecting a large population of vulnerable people.

To give some final insights into the state-of-the-art of AI chatbots, Danieli et al. (2022) present a controlled trial of the mobile Health (mHealth) conversational agent *Therapy Empowerment Opportunity* (TEO). TEO is a cognitive behavioural therapy (CBT) agent, designed to converse with users through text-based dialogue, recognizing emotional states, beliefs, and personal events, and suggests professionally designed strategies on those grounds (ibid.). TEO is then a tool that is largely independent from treatment, and can be used both in parallel with treatment and during the absence of treatment. Danieli et al. (2022) study TEO in an experiment with adults above the age of 55, in four groups at different points of treatment and involvement of the agent during the COVID-19 pandemic. The participants were assessed in terms of symptoms related to stress, and general anxiety disorders prior to treatment, at mid-term, at the end of treatment, and three months after treatment, and relevant groups filled in a satisfaction questionnaire at the end of intervention (ibid.). The results of the study show that there is no significant statistical difference, when it comes to the stress-related symptoms or general anxiety disorders, between those who used the agent alone, and those who used it in conjunction with traditional treatment. However, TEO presents some significant improvements for subjects who are in treatment, in conjunction with greater levels of satisfaction with TEO amongst the subjects undergoing treatment (ibid.). This example contributes in several ways in terms of understanding the use of AI chatbots. First, it presents that introducing the chatbot in the treatment did improve said treatment, and that it is potentially competitive with traditional treatment. Additionally, it seems that there is a correlation, in terms of satisfaction, between using the chatbot alongside traditional treatment, as opposed to using it alone. These factors ultimately highlight the potential of mHealth treatment options, but also emphasises the importance of involving traditional treatment, for users to be satisfied with the chatbot.

2.3 RELATIONS IN GENERAL PRACTICE AND THEIR POSSIBLE DISRUPTION

According to Mittelstadt (2021, 7), it is essential for medical practice to nurture a good pratient-practitioner relationship. With the emergence of AI chatbots however, practitioner-patient-AI relationship has begun forming in health care, disrupting the status quo of relationships.

According to a report published by the Council of Europe, there are six potential areas of concerning the patient-practitioner relationship where AI could have an impact:

"(1) Inequality in access to high quality health care; (2) Transparency to health professionals and patients; (3) Risk of social bias in AI systems; (4) Dilution of the patient's account of well-being; (5) Risk of automation bias, de-skilling, and displaced liability; and (6) Impact on the right to privacy." (Mittelstadt 2021)

The report describes the patient-practitioner relationship as healing and AI as an actor which could potentially support this relationship. Furthermore, the report states that the roles in the relationship remain unchanged with the introduction of AI (Mittelstadt 2021). The report views AI as having a mediating role or even as having the role of a provider of medical care. It describes a change or a disruption to duties and expertise that comes with the introduction of the technology. A potential worry that the report raises is a lack of face-to-face interactions as well as a general uncertainty on the impact on the relationship (ibid.).

AI in general, as well as AI chatbots, have the ability to utilise complex analytics and vast amounts of data, far beyond what human clinical expertise can achieve. This may suggest that its effect on the patient-practitioner relationship could be unique and without precedent (Mittelstadt 2021). Therefore, the emergence of AI chatbots calls for a public debate regarding its possible future impact on this relationship.

After the COVID-19 pandemic there has been a boom in chatbots within health care and this caused professionals within the field to worry about their impacts on clinical practices (Parviainen and Rantala 2022). This was however even before the release and boom of open-access conversational agents such as ChatGPT.

Chatbots are being integrated into practices in various ways, to supplement, survey or completely replace human work and abilities. Although it is not expected that chatbots will replace the work of medical doctors they may however completely redefine the relationships within medical practices (Parviainen and Rantala 2022). In this section we will first introduce the important characteristics of the status-quo within patient-practitioner relations and then move to what the current research says about how these might be disrupted and transformed by the introduction of AI chatbots. We will dive into the perspectives of practitioners as well as patients and summarise how these technologies could potentially transform the status quo or pose dangers.

2.3.1 PATIENT-PRACTITIONER RELATIONS

In order to imagine futures where the relationship between the practitioner and the patient might be transformed due to an introduction of a new technology, we must first understand the status quo of these relations. Furthermore, with the practitioner being the one with power to help their patients as well as in possession of knowledge and skills that the patient has a need for, they are put in a position that requires ethical considerations and often difficult decision making. The patient sometimes needs to exhibit vulnerability and open about often intimate issues or give access to their body, pushing their boundaries for their own wellbeing. Therefore, to understand the knowledge and skills that lie behind the decision making of clinical practitioners and that make up their relations with patients, it is crucial to talk about ethical guidelines as well as rules and principles that practitioners follow in their daily practice.

Ways of Knowing in General Practice

In our thesis we are examining the field of general practice in Denmark and thus we reviewed the ethical principles and guidelines that exist for general practice here.

According to the Professional Ethics for General Practitioners in Denmark, general medicine requires a holistic approach while dealing with issues of human character, which are characterised by their complexity, pertaining to various areas of life while still having to be seen in a medical context (Dansk Selskab for Almen Medicin 2022). Furthermore, it is also crucial to earn the trust, not only of patients but of society as a whole.

With the help of Aristotle, we investigated how general practitioners utilise three forms of knowledge or abilities (in Danish "kunnen") in their practice: theory (*episteme*), technique or craft (*techné*) and judgement (*phronesis*) (Dansk Selskab for Almen Medicin 2022).

To begin with, episteme is the theoretical knowledge that the practitioners acquire through their lifelong education, beginning at university and continuing throughout life as they attend conferences and acquire knowledge about new medical findings from publications. This scientific knowledge includes understanding of universal principles and coherence of causes.

Another key type of wisdom is technique or craft, also called techné. This is the wisdom that regards the skill and craft of the procedures that the general practitioner conducts in their

praxis. In its essence it is a practical way of knowing, using a skill or a method to reach a result.

The final way of knowing is phronesis, also called judgement. This is the kind of knowledge required for good decision making and the skill of applying the latter ways of knowing to individual cases, thus always being contextual (Dansk Selskab for Almen Medicin 2022).
It is also the ability of a physician to make informed and appropriate decisions when the knowledge available is dependent on the specific circumstances at hand. This means that decision-making in medicine requires a flexible and interpretive approach that considers the specific context of each situation, all the while following specific standards and procedures when dealing with unique problems. In essence, it is a highly refined mode of reasoning that requires careful consideration and attention to detail (Parviainen and Rantala 2022).
General practitioners in their daily practice switch between these ways of knowing. For the episteme and techne there are textbooks and instructions, however, the use of phronesis is more nuanced (Dansk Selskab for Almen Medicin 2022).
Therefore, we will mainly focus on decision making and phronesis, as due to their more nuanced characteristics, they could become more challenged in the presence of new technologies. Especially technologies that can act as a support in decision making when meeting all the particular situations characteristic of general practice in all their uniqueness and complexity. We are interested in knowing how, if any, will the emergence of the field of open-access AI chatbots influence phronesis or the decision making of general practitioners and consequently their relationships with patients. Moreover, also the other way around, we want to examine how the understanding of this emerging technology becomes constructed with the understanding of the importance of phronesis and thus how it gets constructed in the context of phronesis.
The importance of being able to assess each individual case is essential when it comes to the relationships between GPs and their patients, since each one is unique with different problems and thus universal principles of episteme or the craftsmanship of techne are not enough without being able to look at the patient as a particular individual.
Phronesis, however, also serves as an alternative to following the abundance of guidelines and recommendations available, with "ethical decision-making based on an application of accumulated wisdom gained through previous practice" (Conroy et al. 2021, 1).

Models of the Patient-Practitioner Relationship
Along with the advancements in science and medicine, the relationship between patient and practitioner also gets transformed and has historically been altered. Emanuel and Emanuel (1992) dived into the question of what an ideal of this relationship should be in and proposed four interaction archetypes into which the patient-practitioner relationships can be categorised (Emanuel and Emanuel 1992).
The first archetypical relation, *the paternalistic model*, describes a type of relation that, as the name suggests, resembles the one of a parental or fatherly figure and a child. This is because in this archetype the practitioner uses their best abilities to determine what is best for the patient, without involving them in the decision making to any large extent. The practitioner might present information to the patient about why their decision is right to support their consent or in some cases only inform them on what will be done and when. Thus, the participation of the patient is very reduced, and the practitioner assumes what is best for them. The practitioner is then also committed, in the same way a parent is to their child, to put the interest of the patients above their own (Emanuel and Emanuel 1992).
The objective of the practitioner's role in the second archetype, *the informative model*, is to inform the patient as much as possible on which options they have available, with the patient

having full control and agency over their choice. The practitioner then executes this choice. (Emanuel and Emanuel 1992). An important element of this relationship is the interplay between the values that the patient holds and the knowledge of the practitioner. The practitioner brings in the facts, while the patient considers them and makes a choice based on their values. The values of the practitioner bear no significance, their sole role is to mediate their knowledge and as experts execute the decision of the patient (ibid.).

The third archetype, *the interpretative model*, presents a relationship where the practitioner's role is to clarify the values, needs and wants of the patient and guide them in choosing the next course of action so that it is in accordance with these. It is similar to the second model; however, the practitioner is more present in the decision making through helping them comprehend their values, wants and needs (Emanuel and Emanuel 1992).

The fourth archetype, *the deliberative model*, portrays a relationship where the practitioner has a role resembling a friend or a teacher, working with the patient to determine the best course of action. The practitioner not only proposes possible options, but also advises the patient on what would be best, based on their knowledge of them. Patient autonomy in this model revolves around moral self-improvement; the patient is encouraged to evaluate alternative health-related values, their merits, and their effects on treatment through discussion, rather than simply following unquestioned preferences or assessed values (Emanuel and Emanuel 1992).

The fifth, and final archetype, *the instrumental model*, presents a relationship, which can be classified as unethical, although it does take place. In this kind of relationship, the values of the patient are completely disregarded. The only subject of importance are the motivations of the practitioner. This could for example be some greater good such as contribution to society or science (Emanuel and Emanuel 1992).

Each model might have its place in different situations; however, the deliberative model is described as the ideal one.

2.3.2 PATIENT- PRACTITIONER RELATIONS AND INTRODUCTION OF CHATBOTS

In our literature search we also focused on what the current research says about how the relations between practitioners and patients transform or react when chatbots are added into the equation, since this is the primary focus of our thesis. We experienced that there is a knowledge gap in the existing research in this area.

When it comes to the patient-practitioner relationship with the complicated vulnerability and power dynamics it encompasses, trust is imperative. With the introduction of AI as a mediating factor, this trust could be negatively impacted due to the quantifiable nature of machine understanding and a possible failure to understand and consider the individual patient (Mittelstadt 2021).

Workflows

There are various degrees, or rather a broad spectrum of how AI can be implemented in the workflows of general practitioners or in care in general. AI can be used as a supportive element, complementing the work of practitioners and in such a case the disruption to the patient-practitioner relationships does not necessarily have to be significant (Mittelstadt 2021). In other cases, when AI is implemented on a larger scale and replaces the medical encounters instead, its influence on the relationship is much more complicated to anticipate (ibid.).

However, broader implementation of AI could also enable the practitioners to spend time with the patients that would have previously been taken up by time-consuming, low-skill tasks and lay ground for new norms and standards (Mittelstadt 2021).

According to the report by Mittelstadt (2021), it is highly improbable that the work of practitioners is going to be substituted by AI. However, the landscape has changed since the report was published. In our report we focus on similar issues, however in the face of the emergence of widely accessible and ungoverned AI chatbots, which both the practitioners and patients have access to.

Mental health

When it comes to conversational AI, a very relevant area appeared several times throughout our literature search, namely mental health. There is a global as well as local problem with increasing demand for psychological and psychiatric services, as well as a public debate about loneliness amongst patients. This situation also creates demand for solutions (Miner et al. 2019). Chatbots do however have the capability to relieve some of the stress and increasing workload.

Even though practitioners appreciate conversing with their patients, they do not receive any financial compensation for engaging in lengthy but meaningful conversations with patients where for example loneliness comes into play. Conversational AI can help address this issue by providing an alternative solution (Miner et al. 2019).

Practitioner-Chatbot Relations: Workflow Adaptations and Their Impacts

When it comes to the perspective of practitioners, and their attitude towards chatbots, mixed results have been reported. An area that was seen as having the biggest potential was administration, for example tasks related to scheduling.

In contrast, practitioners were most sceptical about the abilities of chatbots to: effectively care for all the patients' needs, abilities to display human emotion, and ability to provide detailed diagnosis and treatment because of not knowing all the personal factors associated with the patient (Palanica et al. 2019, 1).

Since chatbots lack "intelligence to accurately assess patients", which could also be called expertise or phronesis, they pose a risk to patients and might not accurately meet their needs (Parviainen and Rantala 2022, 6). Here it is however also important to note, that Paravien and Rantala refer to the chatbot breakthrough that happened consequently the COVID-19 pandemic, while our perspective is on the open access chatbots which boomed with the release of ChatGPT in November 2022 (De Angelis et al. 2023). These chatbots have been the subject of public debate ever since, mainly due to their persuasiveness and ability to appear human-like, thus some of the issues could either become outdated in the light of new technologies, or even more profound when it comes to for example human-like communication, persuasiveness, and self-diagnosis.

2.3.3 DANGERS

As previously mentioned, the introduction of AI chatbots into the field of health care, and in our case, a general practice, can not only disrupt and transform, but also pose threats to parts of the relational status quo that are valuable for clinical outcomes. In this section we will present dangers we identified through our literature search,

Phronesis and Expertise

Artificial intelligence and chatbots also come with risks and pose potential dangers to the relations within general practices. The systems are built by humans and thus reflect both human imperfections and attainments, also when it comes to clinical practices (Parviainen and Rantala 2022). Furthermore, according to Parvianen and Rantala, chatbots currently lack what could be called phronesis, which is problematic. Moreover, self-diagnosis could become a more profound problem due to their persuasiveness and patients could be more difficult to

convince when given a misdiagnosis by a chatbot, thus increasing the workload of practitioners, and increasing tensions in the relationships.

Standards

Generally, when people develop technology, it is usually designed to uphold certain standards or act as a solution to concrete problems. Just as human tools, human professions are also held up to certain standards, which is especially the case with general practitioners and the standards and virtues we described in previously (2.3.1 Patient-practitioner relations). There is however not much known about what goes on behind the closed door in companies such as OpenAI which stands behind ChatGPT that catalysed the boom of open-access AI chatbots. Transparency is lacking when it comes to these technologies, creating a 'black box' problem (Mittelstadt 2021). This means for example, that the logic that is behind the outputs of the chatbots is not accessible or observable or even not intelligible. However,

"Transparency and comprehensibility are generally desired because algorithms that are poorly predictable or interpretable are difficult to control, monitor and correct." (Mittelstadt 2021, 16)

The lack of transparency is however often deliberate. Although democratisation of AI has been somewhat of a trend, it does not come without dangers, since making large language models open access, opens for powerful technologies getting into the hands of individuals who do not mind using them unethically and losing control over their further development.

Furthermore, these AI solutions are capable of passing the same exams as the doctors certified to provide care, however, which standards were they taught to uphold? There exists an uncertainty regarding whether this form of intelligence will be able to abide by the same level of professional standard as a human would while taking care for another human (Mittelstadt 2021).

Technological solutions that are implemented in health care are often highly intricate and the end users, although being experts themselves, do not necessarily understand the logic behind inputs and outputs and the system functionalities in depth. Therefore, it becomes complex to foresee the implications of these systems and thus also to form standards that could ensure ethical and socially beneficial use (Parviainen and Rantala 2022).

Consequently, this situation could allow for various risks that could negatively impact the relationships within clinical practices as well as various factors that can change the relationships as we know them today. These could for example be: loss of quality in diagnostic practice, introducing market principles such as profit maximisation into health care, poor assessment of patients, increasing burden on patients through making them more self-reliant, loss of trust and increase of conflict, increase of workload and lower well-being of employees (Parviainen and Rantala 2022).

Social Division

Lastly, adoption of chatbots could possibly create a divide in society, with human connection being inaccessible to certain groups who might be entrusted to AI and miss out on connections with real humans when needing help (Parviainen and Rantala 2022).

3. PROBLEM FORMULATION

The following chapter will revisit parts of our background chapter that have played a large role in defining our problem formulation. The point is not to reiterate the reviewed literature, but rather to address the important impacts that have initiated, and informed how we will conduct our study into AI chatbots within general practice.

The area of artificial intelligence is a rapidly expanding one, not only in terms of developments of their technical capabilities, but also in terms of the fields to which it is applicable. Particularly within the area of chatbots, the field has seen a boom in terms of desirability, as well as accessibility, within multiple fields. The boom could largely be attributed to the necessity of remote care within health care, brought on by the Covid-19 pandemic ((Parviainen and Rantala 2022, 61-71); (Daniel et al. 2022, e39102); (Danieli et al. 2022, e38067)), and was further accelerated with the introduction of ChatGPT, making the cutting-edge technology widely accessible. Particularly within health care, AI chatbots have then become an area worth investigating, due to the intersection of machine and human intelligence, and the implications this might bring.

Additionally, we have chosen to commit to a case study of a general practice in Northern Jutland. While this case study has benefits in relation to accessibility to the field, we also presented the findings, made by Kim et al. (2022), in terms of how family medicine appears to be one of the medical fields where chatbots exhibit low accuracy. This finding drew our attention towards general practice since it is equitable with the practice of family medicine and would then be an area where the potential for supplementing knowledge would be evident. Additionally, the field of general medicine is most often the first point of contact with the Danish health care sector, and an area that impacts most of the Danish population (DSAM). Caregiving, within general practice, has conventionally been undertaken by human professionals, with specialised knowledge, but the emergence of intricately developed sub-symbolic (Kelley 2003, 847-860) AI chatbots pose a potential disruption to this practice. What will the consequences of this intersection of machine and human intelligence be for the general practice? This is an important question that requires further research and exploration to fully understand the potential benefits and risks presented by this emerging technology.

Finally, we chose to delimit our case study to the perspective of the general practice, as opposed to involving patients. This was partially done, since most of the literature that we reviewed, was primarily engaged with the patient perspective. However, the phronesis of a general practitioner is comprised of unique experiential knowledge, and the ability to apply this knowledge on specific contexts (DSAM). Particularly regarding the human relations with their patients, this becomes a valuable aspect of their specialisation. Large Language Models, such as GPT, have been shown to lack in emotional intelligence, which can impact the direct applicability of the technology, and as such alternative applications, or the exclusion of the technology, should be explored. Subsequently, it is also important to ensure that the patient-practitioner relationship is supported, rather than undermined by the wide accessibility of AI chatbots and that the care received with support of these technologies is held up to the same standards as when provided solely by humans (Mittelstadt 2021).

Within the rapid expansion that we presented, we perceived the need to be able to study technologies that have insecure temporalities, meaning the state of the technology is changing so fast, that studies risk becoming outdated as soon as they are published. Easton et al. (2019) provides some insights into how the user can be effectively involved in the design process of AI chatbots through co-design, however, as techno-anthropologists, we possess unique competencies when it comes to investigating technology through more holistic, ethnographically grounded lenses, yet the uncertain temporality still poses peculiarities, that we wish to address through methodological and theoretical means.

With these facets of AI chatbots within general practice in mind, we seek to investigate the field through the following problem formulation:

"How might the future of rapidly evolving AI chatbots, in the context of a general practice in Northern Jutland, be co-envisioned?"

To operationalise the problem formulation into actionable activities, we will use the following research questions:
  ➢ How can social and cultural aspects within the social group(s) inform the construction of the technological artefact?
  ➢ How might the findings within this thesis inform the sustainable and responsible implementation of AI chatbots in general practice or health care?

4. METHODS

The following chapter will introduce our methodological and methodical approaches to our case. We will present how the case study is conducted and deliberate on reasons why our particular case of AI chatbots in a general practice in Northern Jutland is relevant to investigate. Additionally, we will present our methodology, encapsulated in the concept of ethnographic co-envisioning. Finally, we will introduce the methods used during empirical data generation, and how this data is subsequently coded and analysed.

4.1 CASE STUDY AND FIELD INTRODUCTION

The choice of the case study is also very relevant when talking about futures. Flyvbjerg (2006) describes how the narrative nature of case studies and the qualitative approach have an important role in informing futures, through understanding the object of study through the nuanced perspectives of subjects:

" . . . narratives not only give meaningful form to experiences we have already lived through but also provide us a forward glance, helping us to anticipate situations even before we encounter them, allowing us to envision alternative futures. . . . Narrative inquiries then develop descriptions and interpretations of the phenomenon from the perspective of participants, researchers, and others." (240)

Case studies are however often perceived as unreliable sources of information when it comes to general phenomena that they seek to enlighten through particular examples. In the conventional knowledge about case studies there exist many misconceptions and over-simplifications. For example, scepticism towards the ability to draw generalisations from a single case, or that case studies open for too much subjectivity, undermining the validity of studies (Flyvbjerg 2006).
Furthermore, based on the status quo, case studies always need to be connected to hypotheses, otherwise they are of no value. Flyvbjerg however mentions how several authors of texts that were sceptical towards case studies changed their perspectives and later became inclined to see their value, writing for example that the importance of case studies lies not in "proving anything", but in "learning something" (Flyvbjerg 2006, 221). He presents the value of case studies through the counterarguments against five common misunderstandings concerning case studies.
The first misunderstanding addressed is that knowledge that is general and theoretical in nature and this universal and "context-independent" is more significant than that which is concrete and context dependent (Flyvbjerg 2006, 224).
Studies of human affairs where human behaviour is in focus is however an *eternal beginning*, where the way to knowledge is determined by context and specific cases (ibid). Therefore, seeking proof and universality in social studies is addressed as *vain* by Flyvbjerg (2006), while learning is undoubtedly viable (ibid).
Flyvbjerg also addresses the well-known example of *all swans are white* and suggested that the observation of a single black swan would disprove this statement, prompting further exploration. Therefore, case studies are ideal for uncovering unexpected findings, as a closer examination of what appears to be *white* may reveal something *black* (Flyvbjerg 2006). This essentially envelopes the value of case studies, in that they provide insights into the particular, rather than the general.

The second misconception about case studies concerns generalisation, viewing it to be impossible to generalise from a single case and so unattainable to contribute to scientific development. Flyvbjerg (2006) however argues that it is possible to draw generalisations from case studies. However, he also argues that the notion of generalisation being so central to scientific development is overestimated, making us forget the power that lies in single examples.

Moreover, the third misconception is the belief that case studies are only useful for constructing hypotheses in the early days of the research. Other methods should then be used to work further with the hypothesis and either prove it or disprove it. Flyvbjerg (2006) however argues that generalisation through case study research is possible and therefore working further with hypotheses through this method is also useful.

The fourth misunderstanding regards that the researcher gets their preconceived beliefs verified. The opposite however tends to prove true, and the researcher actually learns through investigating their preconceptions and often disproving them in the process.

Finally, the last misunderstanding that Flyvbjerg (2006, 241) addresses is that drawing "general propositions and theories" from case studies can often be challenging. Although Flyvbjerg (2006) states that this indeed can be challenging, it does not concern case outcomes. Although, this challenge stems from the nature of what is studied and not the methodology. In our case, we have in this way looked at the phenomenon of the emergence of AI chatbots in general practice, interpreted and envisioned by the relevant social groups. The fact that we chose to conduct a case study does however also come with certain limitations that we cover here for transparency. Our case is indeed particular, covering a small general practice in a rural area. It is important to acknowledge that there are different kinds of general practices in Denmark, varying in size, location, regional governance and to certain extent also mode of operation.

There were 1675 general practices in Denmark in 2014, while practices with several practitioners are the most common. As of 2022 there were 3284 general practitioners in Denmark (Praktiserende Lægers Organisation 2022). The amount of consultations per day per practitioner has been increasing and the average lies at 52 daily consultations per practitioner. The majority of consultations are still physical, with an average of 21,8 physical consultations per day, followed by 11,7 email consultations and 11,5 telephone consultations. There is also a small number of video consultations (0,4) and other services. The amount of physical consultations per day has been decreasing, with email consultations increasing, suggesting a move towards streamlining processes surrounding consultations through digitalisation, making AI chatbots an interesting aspect to investigate Northern Jutland is also one of two regions where the practitioners receive the highest payments. The payments consist of a basic amount and an amount for the amounts of services the practitioners perform. (Praktiserende Lægers Organisation 2022)

The general practices in Denmark are financed from public resources, they are however driven as independent businesses. Besides general practitioners, it is common for a practice to employ medical secretaries, nurses, social and health care assistants, bioanalysts, pharmacologists, physical therapists or midwives.The number of this kind of personnel had been steadily increasing since 2008.

The average age of a general practitioner is around 50 years old. Over half of the general practitioners are women. The number of female practitioners have been steadily increasing since 1997, from 10% to around 59% in 2022 (Praktiserende Lægers Organisation 2022).

With regard to future developments, general practices in Denmark are expected to be challenged due to a population that lives longer and thus with more chronic diseases all the while the amount of general practitioners has been showing a decreasing trend. Therefore, the practitioners are getting under more pressure and their workload is getting increased.

To put this into perspective, the group of patients between 70 and 79 years has increased by 40% in the past 10 years and the group above 80 has increased by 25%, while this group is expected to increase even more, by a staggering 47% from 2022 to 2030 (Praktiserende Lægers Organisation 2022, 27).

Furthermore, as practices close, there are often no substitutes in many areas of Denmark and the patients are forced to seek care further away from their homes or at regional clinics where they cannot always see the same practitioner. From 2012 to 2016 the amount of practitioners who felt their work had become more demanding doubled, from 25% to 58% (Praktiserende Lægers Organisation 2022, 24).

Our case study is looking at a solo practice, with one practitioner, a secretary, and a nurse.

Based on the statistics, a practice in Northern Jutland makes for an interesting case, since practitioners in this region are experiencing the biggest workload (Praktiserende Lægers Organisation 2022) and thus the situation might call for streamlining. Additionally, we had unique access to the field through an internal gatekeeper, the secretary, at the general practice. In regard to the problem at hand, AI technologies might be tempting in such situations, however, the articles we reviewed in the background section warn about integrating AI chatbots for this very purpose. Therefore, we see value in looking into what tendencies might appear in a general practice when such a technology is introduced.

We also chose a case study and an expert approach, since accessing informants in the medical field is challenging due to the workload of the potential informants.

Finally, regarding the type of technology we have chosen to focus on, its emergence and recent boom can bring new challenges in the near future as well as solve the ones mentioned in our case study.

4.2 METHODOLOGY

As presented above, in this case study, we seek to explore the potential future of AI chatbots within a general practice in Northern Jutland. For this purpose, we seek a methodology that not only allows for exploring the status quo, but also for looking into imagined futures. Additionally, we are working within the theoretical framework of the Social Construction of Technology (SCOT), which involves understanding the importance of social groups in shaping technologies (5. SCOT Theory). Methodologically, we draw heavy inspiration from *An anthropology of Futures and Technologies* (Lanzeni et al. 2022), more specifically *Ethnographic co-envisioning*.

To think critically, we engage in proposing alternative ways, grounded in our ethnography to engage with the emergence of AI chatbots. We are not engaging in dominant narratives, particularly technologically determinist views (Lanzeni et al. 2022), yet we still wish for our thesis to provide interventionist insights that might inform responsible development of technologies within our case.

We view the marriage of SCOT and ethnographic co-envisioning to be fruitful, since it enables us, through a 'snapshot' of the inherently social, material, and political technology, to imagine how everyday practices might play a role in determining the use (or exclusion) of said technologies. Inspired by the sentiments of An anthropology of Futures and Technologies, we do not seek to find concrete solutions, or requirements, but rather we involve the central stakeholders of our case in finding new problems and solutions (Lanzeni et al. 2022). However, we acknowledge that futures, similar to those which we will envision, in cooperation with our informants, tend to radically differ from how they eventually unfold. With this in mind, we want to present a work that might intervene, or inform, the rapid development of AI chatbot models. As such, we draw on Heidegger's notion of 'enframing', and attempt to apply the

notion proactively, rather than retroactively (ibid.), in order to find alternatives in processes of development and innovation.

Since the technology of AI chatbots is undergoing rapid development, and the impacts, attitudes, and implications cannot necessarily be predicted through conventional testing, and user involvement (Lanzeni et al. 2022), we are ultimately limited by a temporal context, which is difficult to make tangible through conventional lenses, such as that of SCOT. Our approach to this, is to involve a snapshot of the technology, in GPT-4, which, at the time of writing, is the state of the art available AI chatbots. Anthropology has previously been criticised for dealing with the past, and present, and has historically dealt with foreign , and far away contexts, in order to explore why things are as they are, and how things could have been different, while struggling to study futures (Lanzeni et al. 2022). We seek to explore how anthropology, and ethnography can be utilised as tools to explore the future, in an interventionist approach.

4.2.1 ETHNOGRAPHIC CO-ENVISIONING

We aim to ground this future minded approach specifically in ethnographic co-envisioning, and interventionist anthropology, in order to disrupt the invention, innovation, and development processes of AI chatbots, in order to contribute to make those processes sustainable and responsible (Lanzeni et al. 2022). This process does not aim for the same reconstructions of futures, as exact natural sciences might, but rather aims to open up for debate and study into the everyday practices, and how the technology of AI chatbots may play a role in the future (ibid.). We aim to frame future visions more broadly, beyond the conventional approaches of technology design and creation, capital and power relation. Instead, we aim to investigate human activity, to understand what visions of the future might present themselves, in relation to AI chatbots (ibid.).

Ethnographic co-envisioning entails that our informants are not restricted to being respondents to our inquiry, but rather as active, and perhaps even dominant, participants in envisioning the technological futures within their general practice. In our case study, we co-envision through everyday life scenarios, and tangible, situated perspectives. More specifically, we utilise specific methods, such as explorative interviews to understand the status quo, as well as the PLEX scenario technique, and Wizard of Oz to envision futures involving AI chatbots (4.4 Empirical Data).

4.3 LITERATURE SEARCH

The process of searching for literature to form a foundation for this project can largely be separated in two phases; searching for preliminary literature, and acquiring supplementary literature.

Preliminary Literature

The preliminary research was conducted as an exploratory process of reviewing literature on three different databases; PubMed, CINAHL, and Embase through three different exclusion phases. Articles were excluded in terms of the relevance to the motivation for the project (1. Introduction), and articles were excluded in parallel between project group members, based first on titles, then on abstracts, and lastly a collective process of reading remaining articles, and categorising them in terms of how useful they could be for the project.

The literature was found, using the following Boolean search string: *"("Artificial Intell*" OR AI) AND (relate* OR relation* OR turing*) AND Chatbot AND (patient OR clinician OR "general pract*" OR nurs*)"*, which was constructed as a foundation for the areas of interesting

elements to investigate for the project. The resulting search, within the databases, were then sorted by peer-reviewed articles, and the exclusion phases were initiated.

The literature search on CINAHL resulted in 15 results, with 10 being included based on reading titles, and 5 being included based on reading abstracts.

The literature search on Embase resulted in 36 results, with 30 being included based on reading titles, and 17 being included based on reading abstracts.

The literature search on PubMed resulted in 36 results, with 28 being included based on reading titles, and 23 being included based on reading abstracts.

The last, collective, exclusion phase was structured by reading all articles once, inserting them into a table, noting themes, summary/main arguments, and relevance to the project. The articles were additionally colour-coded as either green, relating to very relevant articles, yellow, relating to potentially relevant articles, or red, relating to irrelevant articles. The articles, categorised as very relevant, were read by all project group members, followed by a collective discussion, and inductively creating themes that would become the primary influences on different segments of the background chapter; Introduction to background, State-of-the-art, and Patient-expert-chatbot relationships. 13 articles were ultimately included as the primary inspirations for the background chapter, which was supplemented by additional material.

Supplementary Literature

The supplementary literature was found through checking the references of the included articles and finding material that helped in exploring the subject matters of the background chapter. Additionally, material was found, and included, by consulting peers, and supervisors. This supplementary process was more pragmatic, and less rigid than the first process of searching for preliminary literature, and as a result, the background chapter is structured to provide the reader with appropriate insights into the research field, and not as a chronological process of knowledge acquisition for the project.

To ensure a coherent, and transparent exploration of empirical, theoretical, and analytical elements, this thesis will illustrate these elements, first based on background literature, and subsequently supplementing literature, if the empirical data generation presents new findings. These supplementary literary resources will be integrated into the Background or Theory chapters retroactively, in order to provide a coherent reading of the thesis.

4.4 EMPIRICAL DATA

The following section will present the methods that we applied to generate empirical data. These include interviews, that we used to explore the context of the general practice and our workshop elements, the PLEX scenario technique, and Wizard of Oz, respectively. Our empirical data was gathered at a general practice in Northern Jutland, in two meetings, involving two internal informants, a medical secretary and a general practitioner.

4.4.1 INTERVIEWS

This section outlines our methodical approach to conducting interviews as an empirical research tool, adhering to Kvale and Brinkmann's perspective on qualitative research interviews from their book, *Interview* (Kvale and Brinkmann 2015).

Qualitative research interviews provide insights into the emic world of the subject through dialogue between the researcher and the participant (interviewer and interviewee). To depict the subject's lifeworld or worldview, pertinent informants are interviewed to obtain specific knowledge about how experiences and actions are conveyed or articulated. Kvale and

Brinkmann (2015, 20) emphasise the necessity of basic conversational skills for this method, such as the ability to ask questions and engage in conversation.

"The research interview is based on everyday conversations and is a professional discussion; it is an inter-view where knowledge is constructed in the interplay or interaction between the interviewer and the interviewee."

This quote highlights a key element of the interview process: the exchange of perspectives between two individuals. Knowledge creation through this interaction is visually depicted by Kvale and Brinkmann by a white vase-like object between two faces, symbolising knowledge constructed during the verbal interplay between researcher and informant (Kvale and Brinkmann 2015, 21). Therefore, knowledge is constructed through the interrelationship between subject and researcher.

Kvale and Brinkmann (2015, 21) describe the interview as a conversation with structure and purpose. The researcher's essential task is to guide and direct the conversation toward a specific topic or question, presenting it to the subject, and then critically following up on their response.

Interview Guide

For the qualitative research interview, we employed an interview guide as our starting point, a script intended to structure the interview itself (Kvale and Brinkmann 2015, 185). Utilising a topic-oriented approach, we formulated relevant thematic research questions for the semi-structured research interview. These questions, which we seek answers to, are not directly presented to the subject during the interview. The primary role of research questions is to generate knowledge for subsequent analysis of empirical data (ibid., 185-186). Research questions are mediated through interview questions that consider and translate the research questions, typically composed in an academic, theoretical language, into the subject's assumed everyday language, contributing to a natural conversation flow (ibid., 187). The translation into the subject's everyday language aims to elicit spontaneous and rich responses or descriptions to the specified interview question.

Transcription

After completing the interview, it was transcribed to facilitate its use in our coding process, as described in section 4.6 Coding and Results. The transcript was prepared as uniformly as possible. Ideally, one group member would transcribe the interview in its entirety to maintain consistency. However, due to the time-consuming and complex nature of transcription, prone to technical issues and potential misinterpretations, we opted to distribute the interview recording among all group members. We established shared guidelines to ensure homogeneity, combining verbatim transcription with notation of relevant non-verbatim behaviour, such as laughter or sarcasm, which could convey specific meaning in context. Alongside the verbatim transcription format and notes of pertinent nonverbal behaviour, timestamps were included at the beginning of each speech segment to better structure the conversation. The interviews were transcribed in Danish, and subsequently translated to English, as they were used for analytical points.

4.5 WORKSHOP

For the second meeting, with the general practice of the case, a workshop was planned, with the purpose of exploring their views on AI chatbots, and how the future of the practice might be influenced by the technology. The workshop was structured in two rounds, involving two workshop elements; the *Playful Experiences* (PLEX) *scenario technique* (Lucero and Arrasvuori 2010), and a process of *Wizard of Oz* (Preece et al. 2019), respectively.

4.5.1 WORKSHOP SETUP AND STRUCTURE

The workshop was set up in the staff room of the general practice, around a round table. Before the actual meeting, the two present facilitators prepared workshop elements by preparing a laptop with ChatGPT, selecting to use GPT-4, using an initiative prompt to define the role of the chatbot. An A3 version of the PLEX scenario game board was placed in the centre of the table, with seating for the two participating informants, side by side in front of the game board. A tablet was placed behind the participants, to allow the third facilitator to participate, and to audio/video record the workshop session for analysis. An agenda was written on a whiteboard, already present at the general practice.

The workshop session was structured as follows:

1. A group member introduced the agenda for the workshop, and reiterating some significant themes and elements from the first meeting.
2. The two present group members conducted a short, two minute, pre-planned game of the PLEX scenario technique, using the initiating case of virtual reality in a coffee shop.
3. The two rounds of the workshop followed the same structure, first relating to introducing AI chatbots during consultations at the general practice, and secondly relating to introducing AI in activities surrounding the consultation.
   a. The PLEX scenario technique was played, using the game board.
   b. A Wizard of Oz was performed, using GPT-4 on a laptop.

During the workshop, the participants requested that the facilitators wrote down points made during the workshop, for them to focus on the conversations and workshop elements, to which the facilitators obliged.

4.5.2 PLEX SCENARIO TECHNIQUE

For the first activity, the PLEX framework was used to engage the participants in some creative, imaginary processes (Lucero and Arrasvuori 2010) to explore the area of AI chatbots in their general practice, which could be an abstract idea, depending on their knowledge and insights into the technology itself. The framework includes templates for playing cards, and a scenario game board (Figure 4.1), which can be used for two proposed techniques; brainstorming, and scenario (ibid.). The brainstorming technique assists in exploring a concept broadly, encouraging idea generation of multiple ideas, without necessarily diving deeper into the peculiarities of the ideas. *The* PLEX scenario technique, on the other hand, focuses on proposing a case, or a concept, which is explored through the gameboard, containing a start, a continuation, and an end of a scenario to be built around the case. These three stages of the scenario are presented on the game board, in three frames, similarly to a short cartoon strip (ibid.), that are sized appropriately for the cards to fit.

Figure 1: An example of the PLEX card, Fellowship (on the left), as well as the card containing rules for the PLEX scenario technique and a model of the PLEX game board (on the left).

The PLEX cards were developed through an iterative process of developing 22 categories, and imagery providing concepts that allow players to engage creatively, and playfully in conversations (Lucero and Arrasvuori 2010). Within both techniques, the players should then be encouraged to think aloud, which is also explained in the rules of the two techniques (Figure 4.1). Additionally, players can write down their thoughts, feelings, and ideas on sticky notes, or draw quick sketches, depending on their preferences.

As we did know of some specific situations, that were central to the general practice, we decided that the scenario technique was the most appropriate to explore their thoughts and feelings about AI chatbots. Additionally, the scenario technique could enable a continuity between the workshop activities, as we could ask them to first explore the two initiating situations, using AI chatbots during a consultation, and in activities surrounding the consultation, respectively, followed by the *Wizard of Oz*, where we allowed them to act out the scenario.

In preparation for the *PLEX scenario technique*, we printed out the *PLEX cards*, and *game board*. The cards were cut out, and laminated, and as such served as a relatively low-cost, and time-efficient workshop element (Lucero and Arrasvuori 2010).

4.5.3 WIZARD OF OZ

For the second activity of the workshop, we have been inspired by the low-fidelity prototyping method called Wizard of Oz (Preece et al. 2019). The method is conventionally used by creating some rudimentary software, with primary elements of a user interface, and some fundamental functionality, as it is traditionally a method used to test low-fidelity prototypes (ibid.). This software is operated by a facilitator, working behind the scenes, to present the

intended functionality of a system. As such, it is a method that is successfully used for various applications, particularly when analysing dialogue between a user and virtual agents or within human-robot interaction studies, since the interaction can be distilled down to a request-response interaction (ibid.). The method allows for playing out situations with testing participants, which aligns well with how we used the activity as a continuation of the PLEX scenario. Additionally, Preece et al. (2019) introduces how the operator can be replaced by artificial intelligence, as the technology matures, which is the approach that we had during the workshop.

We used this method by setting up an initiating prompt on a laptop with ChatGPT, which we formulated relatively abstractly, to attempt to have the AI work for us. The initiating prompt was additionally formulated in Danish, in order to test the system in as close a context as could be expected in a Danish general practice and explore how GPT-4 could handle Danish interaction. The prompt, which was written in Danish, was as follows: "You are now an artificial intelligent assistant, in a general practice in Denmark, who has the purpose of advising me, as a general practitioner in my work day".

Proceeding the *PLEX scenario technique*, the participants were asked to try and use the chatbot how they would, if their scenario was implemented in the general practice. During this process, we had conversations regarding their thoughts, and feelings surrounding the interaction and potential implementation in the general practice. In order to avoid a technological determinist approach, we encouraged both positive and negative critique of the technology.

4.5.4 WORKSHOP DATA GENERATION

who was participating remotely and recorded the meeting. We used this footage to analyse the workshop.

Additionally, we took pictures of the game board after the PLEX scenario technique, to see the result of the activity. Since the facilitators obliged to jot down the thoughts of the participants on sticky notes, these pictures were not necessarily beneficial as a tool to review the insights and elements emphasised by the participants. The pictures were then used in order to review the order that the participants placed the cards on the game board, more as a supplement to the audio and video, than as empirical data by itself.

4.6 CODING AND RESULTS

As delineated earlier, our data collection involved methods including conducting interviews and facilitating a workshop in a general practice setting. To analyse this diverse set of data, we drew inspiration from Braun and Clark's (2006) thematic analysis. According to Braun and Clark (2006, 5), thematic analysis is a technique used to identify, analyse, and report patterns or themes within data. This method was used to compare and contrast the generated empirical data in order to discover similarities and differences. Our initial step involved transcribing the interview, thereby converting the auditory content into a textual format, enabling us to conveniently sift through the data and recognise recurring patterns or themes. According to Braun and Clarke (ibid., 10) "A theme captures something important about the data in relation to the research question, and represents some level of patterned response or meaning within the data set". The analysis of the workshop that was informed by the identified themes from the workshop, followed a similar procedure to that of the interview. In contrast to the interview data, the workshop data was not transcribed but rather analysed through a group approach, wherein patterns and themes were recognised by collectively viewing the workshop recording, and including relevant quotes regarding these patterns and themes.

Through a systematic examination of each individual data item, followed by a comparative analysis of the identified themes, we ascertained as overarching themes that were found to be recurrent in two or more data items.

Given that our informants contributed a diverse range of perspectives and considerations to each of the overarching themes, we incorporated subcategories within each theme to account for this diversity. To illustrate, under the broader theme of *Workflow* we included the following subcategories: *Referrals, Knowledge Seeking, Documentation, Preparation, Diagnosis* (Figure 4.2).

Figure 2: Miro board illustrating the overarching themes and its sub-categories derived from the interview's coding process.

Subsequent to coding our data using this approach, we compared our identified themes with our problem statement to effectively determine which themes were pertinent for inclusion in our analysis. As Braun and Clark (ibid., 10) pose, the significance of a theme does not solely hinge on quantifiable measures, but rather on its relevance to the research question at hand.

After the identification of overarching themes and subcategories within the themes, we further disassembled/subdivided statements into specific element in accordance to the technological frame, which will be elaborated in section 5.3 Technological Frames, clarifying relevant claim and thus their perspectives on the technology. The empirical data's subdivision was performed by attaching relevant pictograms to differentiate between certain elements from the technological frame (Figure 4.3).

Figure 3: Miro board illustrating the workshop's coding process in conjunction with pictograms to incorporate the technological frames.

5. SCOT THEORY

In the following chapter we will present the Social Construction of Technology (SCOT) framework, examining its key concepts including the multidirectional model, technological frames, closure and stability, and interpretative flexibility. Drawing heavily on the works of Bijker and Pinch (2012), this chapter serves as a comprehensive discourse on how technological artefacts are socially constructed. This chapter centres the role of social dynamics and interpretation in shaping technology.

5.1 FRAMEWORK AND TERMINOLOGY

The multidirectional model is developed by conceptualising technological artefacts as a fluctuation of variation and selection (Bijker et al. 1993, 22), which serves as a stark alternative to linear innovation paradigms. However, Bijker and Pinch (2012, 22) emphasise that when looking retrospectively at a technology's evolutionary trajectory, it is feasible to simplify the multidirectional model into a more straightforward, linear version. Figure 5.1 provides a multidirectional perspective of the historical development of the Penny Farthing bicycle. The artefact, the Penny Farthing bicycle, is situated centrally and depicted by a hexagon. Squares with rounded edges represent the relevant social groups associated with the artefact. The term *relevant social groups* encompass institutions, organisations, and both organised and unorganised groups of individuals relevant to the artefact. As suggested by Bijker (2012, 23), it is essential to question the artefact's importance to the social groups under scrutiny. Relevance is not exclusive to consumers and users of artefacts; groups of less apparent relevance may also be important, as demonstrated by the *anti-cyclists* in the bicycle case, who are deemed a relevant social group due to their apparent engagement with the bicycle. Moreover, the SCOT approach necessitates structuring the relevant social groups as homogeneously as possible, in terms of the artefact's relevance to the group (Bijker et al. 1993, 27). It may be beneficial to segregate heterogeneous user groups, as demonstrated by the distinction between men and women in the bicycle case, due to variations in the artefact's significance within the specific case (Bijker et al. 2012, 28). A comprehensive description of the relevant social groups will make it possible to explain the artefact's developmental process, to be discussed further in section 6.1 Relevant Social Groups.
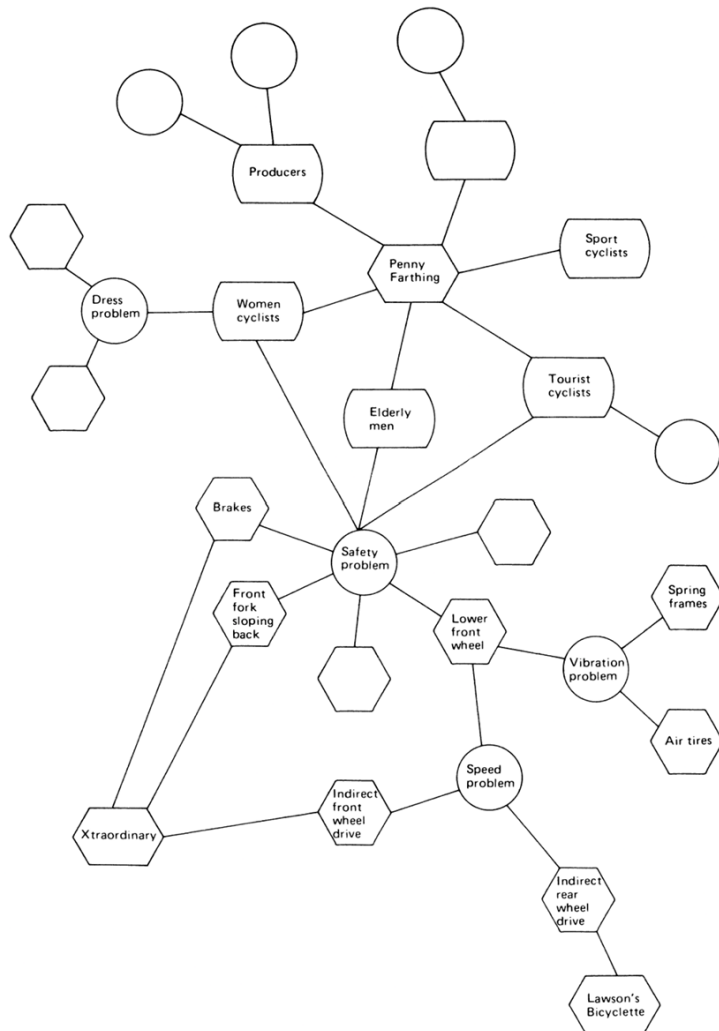
Figure 4: The multidirectional SCOT model portraying examples of social groups, problems, and solutions.

Upon identification of the relevant social groups associated with the artefact, subsequent identification focuses on the problems each group discerns in relation to the artefact (Bijker et al. 2012, 28). Numerous solution variants can be identified around each problem. These solutions are not necessarily technological but may be legal or moral in nature. The model embodies the interpretive flexibility of technological artefacts amongst relevant social groups and, additionally, the model can display the stabilisation degree of the artefact, which will vary across different relevant social groups (Bijker et al. 2012, 30).

5.2 TECHNOLOGICAL FRAMES

Bijker (1997, 123) provides a definition for the term technological frame:

"A technological frame comprises all elements that influence the interaction within relevant social groups and lead to the attribution of meanings to technical artefacts– and thus to constituting technology." (Bijker et al. 1997, 123).

The technological frame thus makes it possible to clarify why social groups construct a technological artefact in a particular way. Bijker (1997, 125) suggests a flexible list of elements that a technological frame could potentially comprise: *Goals, Key problems, Problem-solving strategies, Current theories, Tacit knowledge, Test procedures, Design methods and criteria, Requirements to be met by problem solutions, User's practice, and Exemplary artefacts*. This list is subject to change; certain elements may carry more weight than others in the context of the subject being examined, as well as within different relevant social groups. Hence, some of the suggested elements may be eliminated or incorporated in the designated study.

Technological frames traditionally serve as a concept in understanding how relevant social groups comprehend and utilise technological artefacts, particularly in terms of how these groups attribute different meanings to an artefact (Bijker et al. 2012, 102). Each relevant social group adheres to a specific technological frame. An individual actor may be a part of multiple relevant social groups, and consequently, multiple technological frames, while simultaneously navigating through these groups.

Besides understanding both how relevant social groups understand and potentially use the technology, AI chatbots, as well as the meanings they attribute to it, we have employed technological frames to code and categorise data (4.6 Coding and Results). This method will later facilitate the interpretation of interaction within relevant social groups. The technological frame thereby enables the understanding of problems and potentials that different relevant social groups associate with the technological artefact. Bijker (1997, 124) emphasises the efficacy of technological frames when the researcher's focus lies on situations marked by instability, controversy, and change.

5.3 CLOSURE AND STABILITY

The second phase of the Social Construction of Technology (SCOT) framework pertains to the stabilisation of the artefact. This stabilisation manifests through closure mechanisms that seemingly make issues disappear (Bijker et al. 2012, 37). Notably, it is not required to resolve a technological controversy entirely. Rather, it is critical that the relevant social group perceives the problem as solved. One method of achieving this is through rhetorical closure, exemplified by the attempt to quell the safety controversy around the high-wheeled bicycle, Facile. By stating that the bicycle is nearly entirely safe, the problem appears resolved:

"Bicyclists! Why risk your limbs and lives on high Machines when for road work a 40 inch or 42 inch 'Facile' provides all the advantages of the other, coupled with almost absolute safety" (Illustrated London News 1880; cited in Woodforde 1970, 60; cited in Bijker et al. 2012, 38).

Another mechanism fostering stability and closure involves the redefinition of the problem. As the concept suggests, this mechanism redefines the issue, thereby creating a new interpretation:

"One could argue, we believe, that the meaning of the airtire was reframed to represent a solution to an entirely different problem" (Bijker et al. 2012, 39).

5.4 INTERPRETATIVE FLEXIBLITY

Interpretative flexibility refers to the capacity of technological artefacts to be understood and used in diverse ways, depending on individual or group perspectives (Bijker and Pinch 1992, 40). In other words, there is not a single determinant interpretation or use of a technological artefact; instead, there exists a multitude of plausible interpretations influenced by cultural, social, and individual contexts. Demonstrating the social construction and interpretation of an artefact necessitates the display of its interpretive flexibility. This concept signifies not just the flexibility in how individuals perceive and interpret the artefact, but also the adaptability in

its design (Bijker et al. 2012, 24). Methodologically, interviews with actors from relevant social groups can be utilised to ascertain their interpretations of the artefact (Bijker & Pinch 1992, 40). However, demonstrating interpretative flexibility is not confined to interviews and historical sources. The study of technology offers other methods, such as revealing how different social groups have radically different interpretations of a single technological artefact (Bijker & Pinch 1992, 41). The interpretative flexibility includes the different interpretations of relevant social groups of the artefact's content leading to distinct developments. This shows the influence of interpretative flexibility in shaping technology's trajectory.

Thus, SCOT's interpretative flexibility demonstrates how technology is culturally constructed and interpreted, challenging deterministic views of technological development. This perspective underscores the influence of social and cultural factors in shaping the design, interpretation, and use of technological artefacts.

6. ANALYSIS

The following chapter contains our analytical findings within our empirical case study at the general practice. Since the theoretical framework of the analysis is the social construction of technology (SCOT), we will initiate the analysis by identifying relevant social groups. Additionally, the chapter will be sectioned into the significant themes, found during coding, and analytical points will be related to the theoretical frameworks, presented in 5. SCOT Theory.

6.1 RELEVANT SOCIAL GROUPS

In the following section, we aim to delineate the significant social groups that intersect with the use and function of the technological artefact in question. This critical step draws inspiration from the perspective shared by Bijker and Pinch (1993, 34), who argue that "We need to have a detailed description of the relevant social groups to define better the function of the artifact with respect to each group. Without this, one could not hope to be able to give any explanation of the developmental process". By identifying and elaborating on these relevant social groups, we can better understand not only how the artefact operates within different social contexts, but also how these interactions may shape its ongoing evolution and development. This understanding will enable a richer interpretation of the artefact's role and significance within its broader socio-technological milieu.

To classify pertinent problems associated with the artefact for analysis, it is first necessary to designate relevant social groups, as a problem can only be defined in context to the social group within which the problem is constituted (Bijker et al. 2012, 22). Relevant social groups include institutions and organisations, as well as both organised and unorganised groups of individuals, where the key requirement is that all members of a particular social group share the same kind of meanings attached to the artefact. As Bijker and Pinch (1993, 34) suggest, there is no predefined methodology available, and it is ultimately at the discretion of the researcher to determine the course of identifying relevant social groups, the ensuing segment, anchored by the question: does the artefact bear any importance to the members of the examined social groups?, will identify and select the social groups that are relevant to the utilisation of AI chatbots in general practice. This focus will be particularly directed towards practitioners from the case study, while consciously not elaborating on other social groups such as patients and developers.
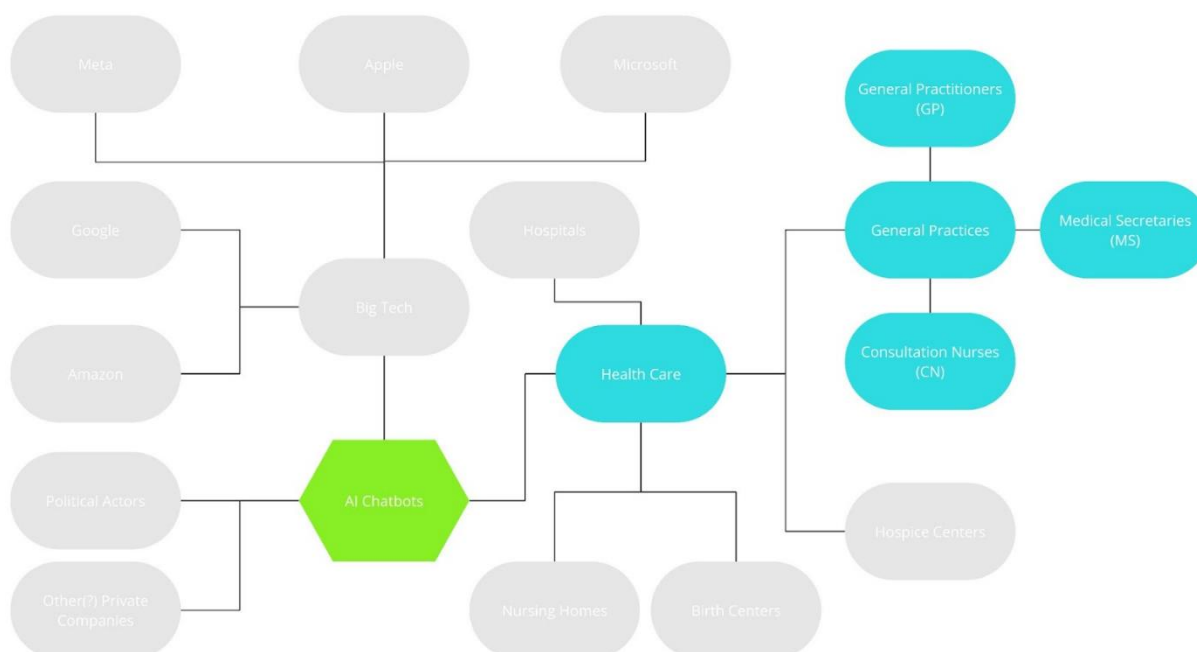
Figure 5: A visualisation inspired by SCOT's multidirectional model, delimiting the relevant social groups in general practice surrounding the technological artefact, AI chatbots.

Initially, we will sketch the complex social network related to the technological artefact, although the depiction may lack precision due to the wide-ranging application of AI chatbots across professional activities. The purpose of Figure 6.1 is to provide a visual guide for the reader, drawing attention to the social groups that hold relevance to our study. This is illustrated by highlighting the relevant social groups with colour, while the less relevant ones are depicted in a subdued grey tone. Further delimitation narrows the focus of social groups to those relevant to health care, proposing groupings relevant to the technological artefact. This results in an emphasis on general practice and medical clinics, with the internal social groups of this context becoming the central point of investigation being *General Practitioners, Medical Secretaries, and Consultation Nurses*.

6.1.1 DESCRIPTION OF RELEVANT SOCIAL GROUPS

General Practitioners

General practitioners (GPs) constitute a crucial social group within the Danish health care sector being the first point of contact. Trained specialists in their field, they engage in patient care in its entirety, accommodating various factors such as age, gender, and the nature of health complications. This comprehensive engagement demands a personal approach to medical care, wherein an individual's specific familial, societal, and cultural contexts are considered respecting patient autonomy (Ministeriet for Sundhed og Forebyggelse 2008). GPs undertake a commitment to their communities, exercising their professional responsibility in a manner that incorporates a spectrum of physical, psychological, social, cultural, and existential factors. The integration of these factors into their treatment planning is based on the trust and knowledge derived from regular patient interactions (ibid.). Moreover, GPs are dedicated to promoting health, preventing disease, and delivering appropriate treatment, care, or palliation. This role may be enacted directly or indirectly through coordination with other health care personnel. They strive to meet the specific health requirements of their community given the available resources and assist patients in accessing these necessary health services (ibid.). A critical aspect of the role of GPs is their commitment to ongoing professional

development, maintaining their skillset, and upholding personal integrity and core professional values (ibid.). These attributes are essential for providing effective and safe patient care.

As a relevant social group in health care, GPs demonstrate a distinct, complex relationship with the technological artefact of AI chatbots, characterised by a combination of scepticism and prospective optimism. This nuanced interaction is rooted in their comprehensive medical expertise and professional experience. The optimism towards AI chatbots arises from the vision of these tools as silent assistants that could potentially streamline administrative and lower-skilled tasks. As highlighted by the GP in our case study, such tasks may include ad hoc documentation, prescription generation, note-taking for future reference, email correspondence with nurses, scheduling of follow-up consultations and appointments, among others. This positions AI chatbots as promising tools capable of enhancing the efficiency of GPs in their practices. Conversely,according to the GP, this group also exhibits scepticism, primarily concerning the reliability, legitimacy, and accuracy of medical information generated by the AI chatbot. GPs, vested with the responsibility of delivering high-quality patient care, express apprehension about the potential risks associated with incorrect or misleading information dispensed by these AI tools. While GPs see potential in AI chatbots to enhance the operational efficiency of their practices, they also underscore the importance of ensuring the reliability and accuracy of these tools.

Consultation Nurses

In a typical Danish general practice, consultation nurses (CNs) often serve as the primary point of contact for patients, handling inquiries, providing advice, and triaging cases to ensure appropriate care levels. They hold a critical role in interpreting and conveying medical information accurately. Our informant and general practitioner perceive that consultation nurses are aware of their medical limitations, acknowledging that certain areas of expertise might be beyond their scope. The GP assesses the potential engagement of consultation nurses with technological artefacts like AI chatbots by reflecting on their existing practices, such as their adherence to stringent methodological workflows involving general and ethical guidelines, as well as decision-making. AI chatbots could be an excellent fit to complement the competencies of this social group, given their established practice of following rigid guidelines and checklists to ensure patient care and safety, as highlighted by the GP. However, it is essential to note that we have not had the opportunity for direct dialogue with the consultation nurses. Our empirical data concerning this social group is solely derived from secondary knowledge provided by key informants who are able to represent the consultation nurses due to their close collaboration with the group's members. This lack of primary representation ultimately eliminates the possibility of determining the social group's interpretative flexibility of the technological artefact. Although this approach is not ideal, we deemed the insights gathered through the informants and sources too significant to exclude from an analysis conducted within the framework of SCOT theory.

Medical Secretaries

Medical secretaries (MSs), as a relevant social group within general practice, exhibit a unique, pragmatic relationship with the technological artefact of AI chatbots, influenced by their roles both centred around administrative efficiency and seamless communication as well as lighter clinical tasks (Allen et al. 2011). This is due to the artefact's already developed administrative capabilities to automate routine tasks, such as scheduling appointments, issuing reminders, handling general inquiries, and facilitating communication aligning well with the primary functions of the social group. Our informant in the social group anticipates the potential of the artefact to increase administrative efficiency and reduce workload, and further envisions the possibility of allowing more focus on tasks requiring human judgement and interpersonal skill.

6.1.2 INTERPRETIONS OF AI CHATBOTS

In this section we will distinguish the social groups' interpretative flexibility toward the technological artefact. Our empirical data generally points towards a consensus in technological interpretation between the social groups. However, this concurrence might stem from a variety of social factors including, but not limited to, inherent power structures (6.2.3 Perceived Expectations and Performativity) within the general practice and its social groups as well as our methodical compromises.

During the PLEX scenario technique, the GP and MS envisioned the potential role AI chatbots might play as a technological actor within their particular clinic. Both the informants had explicit interpretations of how the artefact might be utilised in the early stages of the workshop. The GP articulated a desire for, and foresaw possibilities of, how the artefact might promote objectivity and professionalism. Additionally, they anticipated that the AI could function as a warning system to prevent errors surrounding patient consultations. Conversely, the MS identified areas of administrative and practical spheres. These included the capability to control scheduling, remind users of tasks, assist with low-skill tasks, and the potential to guide patients in scheduling consultations. While both groups perceive the artefact's utility, the distinct emphasis indicates interpretative flexibility shaped by their roles and needs within the general practice.

However, as the workshop progressed, an unexpected dynamic emerged. The informants engaged in collaborative brainstorming, bouncing ideas off each other, which ultimately led them towards a shared understanding of the technological artefact. This convergence muddied the waters of interpretative flexibility, making the differentiation of interpretations between social groups increasingly challenging. It appeared that through the course of this interactive discourse, individual interpretations began to blur and overlap, lending the artefact a more unified, albeit less distinct, social interpretation.

6.1.3 ACTORS' PERFORMANCE WITHIN THEIR RELEVANT SOCIAL GROUP

We will, in the following section, explain how we utilise the concepts of performance and performativity (Butler 2006) to understand how the social groups act at the general practice, providing a context to the analysis, regarding why the user's practice might be unique.

Incorporating Judith Butler's ideas of performance and performativity, we can understand how the practitioner's actions reflect the performative nature of their role in the medical context.

They perceive the rest of the consultation just as a medically unnecessary service, and in this sense, they perform it, as if it was of medical importance.

Butler pondered whether humans equally operate based on expectations when it comes to gender, with these expectations then constructing "the very phenomenon that they anticipate" (Butler 2006), XV). Incorporating Butler's ideas of performance and performativity, we can understand how the practitioner's actions reflect the performative nature of their role in the medical context.

They perceive the rest of the consultation just as a medically unnecessary service, and in this sense, they perform it, as if it was of medical importance

Although these ideas were primarily written about gender, Butler (2006) themselves uses a literary example portraying law. In our case, we see parallels with this line of constructivist thinking, as the practitioner explains how her work contains various performative elements, constructing her role as the general practitioner as well as the course of the consultation based on the expectations of patients. In this way, the practitioner's actions can be seen as performative, echoing Butler's notion that performances can be instrumental in shaping and maintaining certain social norms and expectations such as in this situation, the notions of how a practitioner should behave and when a consultation is necessary. The practitioner continues:

"I could have said already in the waiting room, 'You with the elbow, just go home again, you won't get a referral for an X-ray,' but I don't."

By carrying out the consultation, even when they may perceive it as medically unnecessary, the practitioner reinforces the idea of the consultation's significance through performativity and adheres to both their own as well as the patient's perceived standards. Furthermore, the practitioner then elaborates on the consultation as a service, introducing more performative elements through communication:

"So, we talk, have a casual chat, engage in small talk as we walk down the hallway, then we enter the room, and say "it's nice" 'Oh, you like green, don't you?" "Yes, we do." "You take a seat," "I'll sit here." "So, it was your elbow we were supposed to talk about today." "Yes, that's right!" "Because it has been hurting for so long." "Oh, how long has it been?" "At least for 2 days." "Oh, yes, that is a long time. I can understand it's frustrating."

However, one could argue that the social and psychological aspects of the consultation in fact gain the medical importance through acts of performativity - the repeated acts that reinforce the construction of what it means to be a GP and what it means to conduct and undergo a medical examination. Understanding these perceived expectations in regard to how the interactions with the practitioner should be performed becomes central knowledge when it comes to the emergence of AI chatbots, since they have the potential to change the interactions which we call performativity. A good example of this is the use of touch, tactile communication, and placebo to which our informants assigned great importance.

6.2 RELATIONS

After identifying the social groups, we move to the second step in SCOT analysis, namely identifying problems that the groups might have with the functionality of the technology. These include the technical requirements and moral conflicts (Bijker et al. 1993). The SCOT framework analyses these retrospectively, to understand how technologies came to exist in their current forms and so it tells a story of iterations. Since we are looking at the technology prospectively, we will look at which areas explored in our case with our informants paint a picture of potential problematic areas. We do this both in regard to the functionality of the technology that we mainly explored through PLEX scenarios and Wizard of Oz, or possible moral or ethical problems that emerge throughout our empirical material. The first pivotal areas are the relations, which include our findings about the use of knowledge and judgement, patient expectations, the GP's role, and patient typologies within the general practice.

6.2.1 PERCEIVED PATIENT EXPECTATIONS IN REGARD TO USE OF TECHNOLOGY

There exist various dynamics when it comes to patient practitioner relations as well as different expectations. This pertains to the use and presence of technologies in the interactions, such as a computer present and how much a practitioner looks on the screen. According to the practitioner we interviewed, some patients feel more comfortable with the GP interacting with a computer than others. The presence of a computer was, however, described rather as a barrier in the relationship. During the workshop, the practitioner mentioned that the computer has already partially destroyed a good deal of the relationality between them and their patients. Similar concerns were raised when we were envisioning the future with our informants and talking about the possible presence of AI in their interactions with patients:

"It should not affect relationships. That's what it could have the most impact on, that may be my prejudice. It should not interfere, and I could imagine that happening.

How could it ruin the relational aspect?

It takes the focus away from the patient."

This worry aligns with an issue raised by Mittelstadt (2021), that AI could potentially be contributing to a lack of face-to-face interactions and thus when envisioned in the light of the issues raised, it could perpetuate the damage to the relations.
Furthermore, the GP also raised the issues of trust when digital technologies such as the computer or AI are present in the interactions with patients, which is a known threat, in relation to AI as described by Parviainen and Rantala (2022). According to the practitioner, the patients lose confidence in their expertise if they spend too much time looking into the screen:

"It ruins the consultation; it erodes the trust. They won't trust me if I keep looking things up all the time."

Moreover, the presence of technologies in the relationship was also described negatively with respect to online consultations. According to the practitioner, the online exchanges with patients contributed to a loss in the quality of the relations in real life. It made the patients get used to communicating independently of the relations and as a consequence, they also acted with more distance in real life. Another way the relations can lose quality due to presence of technologies can be clearly seen in the real life versus video consultations, where the practitioner experiences patients crying on a regular basis at the physical consultations, but never over video.

6.2.2 PHRONESIS AND EXPERTISE

As mentioned in section 2.3.1 Patient-practitioner relations, general practitioners possess different types of knowledge, where contextual knowledge is most nuanced as well as most important to dive into when it comes to a technology that can serve as decision making support. In this section we will present how our informants perceive and use knowledge under current circumstances, as well as envision its importance or possible challenges in the presence of AI chatbots. During the interview the practitioner described their knowledge in the following way:

"…it's all I have, it's all my experience and all my knowledge. And also what the patient tells and expresses, and all the relational signals that come, the objective examinations and reactions. So, it's all a comprehensive package. There's not one thing that weighs more than the other."

A good example of the use of contextual knowledge was given when the practitioner explained how the process of diagnosing begins already in the waiting room at their initial encounter with the patient:

"Then they come into the waiting room, and they already start there saying: 'I just have so much pain in my elbow, I would like to get an X-ray'... but then, without any difficulty, they take their jacket off and hang it over there, right?' So you think, 'Okay, I already know how this is going to end.' The rest is just service."

Furthermore, the MS and the GP explained that this first impression consists of various sensory inputs, such as looking at the patient's behaviour, movements, noticing their smell or whether they sweat amongst other things. Moreover, they also described how they can sometimes already over telephone distinguish who really needs to come and be seen at the clinic and who just needs to be heard. They say that this is their professional judgement which does not build upon the knowledge that we refer to as techné, but it builds upon the knowledge about the particular individuals, and so it becomes contextual as phronesis. The MS defined it as "human knowledge". Additionally, the informants express how an assessment made by telephone, might be superior to one made in a video-consultation. The video-consultation may have additional features, in being able to see the patient, however, the GP expresses that it is not a sufficient sensory input to make an assessment. Speaking with patients on telephone, however, was viewed as better, since the lack of sensory inputs would be more obvious, leading to more thorough explanations on the side of the patient. The practitioner also explains how they utilise their expertise. The knowledge they had acquired through personal encounters with patients makes it possible to evaluate the patients from short interactions, just by sight.

6.2.3 PERCEIVED EXPECTATIONS AND PERFORMATIVITY

To the patient the touch of the practitioner serves as the healing touch and "laying on of the hands" in some way is anticipated, while from the practitioner's point of view it can be called the diagnostic touch and its importance in the face of technologies being present in health care was by Bruhn (1978). Furthermore, Bruhn (1978) illustrates how essential touch is to the formation of patient-practitioner relationship historically and overall, as a sign of empathy and bonding tool. Also, he highlights the importance of "performing the expected activities of a clinical examination" as we also described above as performance or what the practitioner calls a service and placebo. The practitioner compared the effect of placebo medicine to the effect of placebo in interpersonal communication with the patient, which according to her can be in the right situations just as effective as medical treatment, which then on the other hand might become ineffective when the communicational part is left out:

"Just the fact that someone has listened, touched them, made them feel seen, examined, thoroughly examined, and understood what has happened and why they feel the way they do. That in itself can have a healing effect on many things. And you can provide a completely effective treatment that has no effect at all, if they don't receive all of that as well."

This statement also aligns with the thinking of Mittelstadt (2021), who in his report also addressed the patient-practitioner relationship to be something that is healing.
The MS could also relate to the expertise of the practitioner in regard to being able to evaluate patients already in the waiting room. This can be illustrated through an example, where children enter the medical centre, accompanied by their parents. The MS notices that they are fine, playing around and that they do not behave as if they have a fever, although

the patient expressed having had stomach pain for the past 3 days. They also mention the importance of touch and performing the consultation as might be expected by the parents:

"If you don't poke it (the child) with your finger and actually hurt the children... And you listen to them, and you look into their ears, they should scream and they should not sit still. If you do all this almost against what you personally think is necessary...But then the parents leave and feel safe, and that's really nice. If you hadn't done it, then it's like...

Practitioner: Then they come back the next day. Or they call for acute help [the Danish Lægevagten]."

This example portrays how physical contact during the consultation is a reassuring element and creates a "feeling that something is being done" and that in a literal manner "matters are being taken in hand" (Bruhn 1978, 1470), such as poking the children where it hurts.
This important expectation can be defined as a rhetorical closure mechanism, particularly in relation to text-based technologies, as they are not designed to perform physical tasks. Since the technology is not capable of performing the desired task of fulfilling the expectations of touch, the subsequent closure is inhabited in a dismissal of the artefact, in favour of the conventional human interaction.
Similar to patients reacting positively to their physical expectations being met, they can also react when the opposite takes place. The practitioner talked about how they received many frustrated patients from other clinics, but in their journals, they could see they received the right treatment. These patients did not feel properly examined. The practitioner's guess was that perhaps in the other clinics they did what was necessary but did not have the resources to properly explain their actions to the patient. She would then often give the same course of treatment and hear the patients say ". . . 'Oh, thank you for listening to me'".
Here, just as touching, the act of a human practitioner listening to the human patient becomes a tool of reassurance which is uniquely human and not imagined as something that could be supplied or substituted by the AI chatbots.
Furthermore, The practitioner and the MS described how the procedure sometimes can feel almost in opposition to their professional judgement. Bruhn described how this can feel as a rather religious than a medical act, drawing parallels to examples of laying of hands and healing touch as described in biblical stories, describing this healing touch, especially of persons in a religious or professional position of power such as the practitioner as something engraved in culture (Bruhn 1978).
Should an AI chatbot be used in these situations of informing the consultative process, its lack of contextual and emotional knowledge (Powell 2019; Mittelstadt 2021), might not be able to accommodate the counterintuitive use of phronesis of the practitioner and the MS. This defines a key problem of the artefact, since despite the superfluousness of this medical inspection, it becomes a vital element for a successful consultation to take place.
Furthermore, the touch has a diagnostic value for the practitioner (Bruhn 1978). Our informants described how they can see the effect of the touch in their patients and how the touch also helps them to diagnose when the relation to the patients is maintained:

"But it's touch, you know, something happens with people when you touch them. Whether it's just saying, 'Oh, I can see that you're having a tough day,' or when you give them a hand as they get up from the chair, or something else, something happens. You can see it in them, you

also develop a different connection with people. It's easier to spot things if you have a relationship with them."

Our informants explained their process of diagnosing, how they utilise their contextual knowledge and the essential importance of touch. This is a particularly important aspect, when it comes to dealing with people expressing vulnerability, who are particularly prevalent in health care contexts (Miner et al. 2019). In the face of technological progress and instrumentalisation, we want to highlight the importance of this knowledge, especially in regard to the importance of and maintenance of the patient-practitioner relationship as it can influence quality of care and medical outcomes (Nam, Kyoung Hyup et al. 2022; Powell 2019; Parviainen et al. 2022). Moreover, if chatbots should be used as a decision-making support in general practices, awareness of the holistic nature of diagnosing is needed.

Moreover, the practitioner also mentioned that many of the tests they prescribe are solely to make the patients feel they were properly examined and not because it was medically required. The whole consultation in general is largely about meeting the expectations of the patient through the ways we described above. However, it goes into even smaller details such as the practitioner deliberately using communication techniques such as mirroring, which has become their tacit skill:

"…when I start mirroring them, the first 5 minutes of the consultation, they start following me slowly, because then I lean back a bit and then they also lean back. And then I sit a bit relaxed, like this, and then they start to relax too."

Additionally, the practitioner also consciously uses pauses in speaking, pauses for thought or leaves the room to use Google search. Such breaks can uplift the consultation as well as the relationship in general. When technologies are used, in this case compared to video consultation, such pauses cannot be utilised.

Returning to Butler's notions of performance and performativity, expectations on various subjects play an important role in how humans construct them. Another example of these expectations is how grey hair archetypically gets associated with wisdom and authority. The practitioner mentioned that it helped tremendously to get grey hair, as people began listening to her much more. What is performed gets constructed through repeated acts of performativity. These acts of performativity then naturalise the construction that is anticipated and perceived as something "outside itself" (Butler 2006, XV).

Furthermore, the practitioner was also critical of how the expectations of patients can influence her role. She explained that in the past, the general practitioner was perceived as the all-knowing authority, who could just by sight evaluate whether a patient needed antibiotics. The expectations of patients in regard to tests brought disruption to this authoritative role, since patients demand testing, they themselves deem necessary and thus undermine the authority of the practitioner, or their own perception of this authority. These descriptions resemble the paternalistic model of patient-practitioner relationship, where the practitioner has the role of an authority which assumes what course of treatment is best for the patient to their best ability, without the involvement of the patient in the decision making. Here the decision that the patient can take is whether to accept or decline. The practitioner described her own, and the clinic's, role as follows:

"But we are never equal as health care providers in relation to the patient. Never. Never. It is simply a condition that one must accept - it is us who determine the format and framework of the consultation, and we offer ourselves to the patient and present a hypothesis. It is us who

decide the treatment or investigation options available, and they can then choose whether to accept or decline. But it is us who define the boundaries and determine the format of the consultation."

Here the practitioner mentions another attribute of the paternalistic style, stating that is them "offering themselves to the patient". Emanuel and Emanuel (1992), also mention that in this type of relation the practitioner is then also committed to put the interest of the patients above their own.

The paternalistic role and authority of the practitioner also became evident during the Wizard of Oz exercise regarding both the technology and the MS based on our observations. The practitioner provided a verbal answer describing the course of treatment for the symptomatology we provided as an example. Then we suggested testing what treatment the chatbot would suggest and the practitioner prompted the following: "peanut in the ear. What do I do". The MS was positively surprised with the medical knowledge of the chatbot as it was presented with symptoms and gave a detailed guide of how to handle them (Appendix, prompt 1). The MS then stated that the answer was just as good as the one the practitioner provided, to which the practitioner reacted in a way that could be interpreted negatively, raising their eyebrows at the MS. This reaction suggests that they might have felt their authority could be undermined by the technology. Additionally, it might suggest that the paternalistic role and the authority of the practitioner creates power dynamics between the two social groups.

The paternalistic role then seems to be undergoing a transformation into a role where the patient is more involved, such as the informative model, the interpretative model or the deliberative model as described in section 2.3.1 Patient-practitioner relations. This is however not because the practitioner deems it necessary, but because the patients demand the treatment that they imagine to be necessary or else they would not feel properly treated. The "all knowing" general practitioner of the past was allowed to prescribe treatments just by sight, however, medical technologies such as testing available nowadays undermined this authority and created more demands from the patients.

In the end, it is in everyone's best interest that the involvement of these demands is fulfilled, and the consultation is performed as the patient expects, so that the patient feels satisfied, and the practitioner does not have to see them again the next day.

In the view of the GP, the undermining of authority can also be due to the technologies entering the relationship, such as a chat function, which was implemented at the general practice. The chat function in particular made patients more demanding, perceiving the general practitioner as being more accessible and thus more equal according to the practitioner.

The patients however still perceive the practitioner's authority very strongly according to the MS, since she described that often patients do not have full confidence in her or the nurses' evaluations and need the practitioner to say the words, sometimes even the exact same, to have their expectations met.

However, the practitioner sees authority as an important tool. She stated that hers, the medical MS', and the nurse's authority makes the patients feel safe and that it is rather essential. Moreover, a lack of authority can result in patients losing confidence in the abilities of practitioners, which can be seen with residents who are still receiving their training when they become "too equal" in their patient relationships.

Along with the importance of authority, the informants also emphasised the value of empathy and being human.

Finally, the practitioner also explained how some of her colleagues can use a highly authoritative approach or even arrogance, which paradoxically can make some types of

patients feel safe. While these behaviours may seem counterintuitive to creating a safe and inclusive environment, they can be understood as performative gestures that align with some patietient's preconceived norms and expectations regarding authority and expertise:

"When a practitioner behaves as if they are an asshole, well, then they attract the patients who feel most comfortable with assholes because 'he is simply unbearable, but damn, he is talented. So they feel safe with it.'"

The Importance of Patient-Practitioner Relations
The patient-practitioner relations are also essential and closely intertwined with the subject of professional judgement and phronesis. The practitioner described how the long-term relations they form with their patients support the diagnostic process. Although in these conversations chatbots were not mentioned, it is valuable knowledge as the nature of this relation must be understood, protected, and supported in the face of emerging technologies, since they directly influence medical outcomes. The knowledge of the individual patients becomes tacit, and thus the practitioner explained that they can recognize patterns through tacit knowledge, or deviations from patterns for that matter:

". . . or maybe it's Karen who is a bit paler, a bit thinner than she usually is. Then I really should [do something]."

It is also important to note that this continuity, that supports the practitioner's phronesis, is something unique to general practice, where the patients are followed for many years, sometimes large parts of their lives, by the same practitioner.
Furthermore, as we mention in the background section 2.3.1 Patient-practitioner Relations, for a good patient-practitioner relationship to exist, trust must be maintained. This trust is described as something that must be earned and protected, since the practitioners are dependent on it, both on individual level with patients and with the society as a whole (Dansk Selskab for Almen Medicin 2022).
According to Mittelstadt (2021) trust is one of the areas that could come under threat with the introduction of AI, especially since phronesis in understanding the individual patient is needed. This threat could however be imagined in various ways, for example starting with the existing technologies such as the computer, that already now disrupts this relationship as we already established. Furthermore, phronesis also manifests as emotional intelligence, which is important for the maintenance of trust. According to Powell (2019), chatbots lack the abilities to respond with emotional intelligence and this factor could further undermine trust if they should be included in interactions,
Although the practitioner undoubtedly has vast knowledge in all the three forms of knowledge we presented in the section 2.3.1 Patient-Practitioner Relations; episteme, techne and phronesis. The theoretical knowledge, also called episteme, cannot cover for all existing diagnoses. Especially in the field of general practice, which is the first point of entry for the patients, a helping tool such as Google search engine becomes an important tool. The practitioner sometimes however deliberately hides this computer interaction, as she is worried it might negatively influence the trust and the patient's confidence in her abilities. On the topic of AI possibly supplementing knowledge, the practitioner expressed that it would be beneficial to have a machine that could be quicker than a Google Search to recognize rare diseases and give the practitioner the ability to quickly explain them to the patients.

Another aspect that the practitioner mentioned after the value of continuity, is the general ability of technology to exhibit the humanity that they can in their clinic, doubting that a machine could display empathy:

"Where is the continuity, humanity, and the relational aspect? Because we communicate in a way that enhances the expectations and ability of both the sick and healthy individuals to manage their own health. We involve the patient's thoughts, concerns, wishes, and so on, taking into account the patient's resources, particularly family and network. A robot simply cannot do that. We follow the patient and their relatives throughout their life trajectory, illnesses, and diseases. This is the heart; this is the human. A robot has no heart."

This description also aligns well with the ethical values described in Professional Ethics for General Practitioners in Denmark as presented in section 2.3.1 Patient-practitioner Relations, stating that in a general practice holistic approach is needed, since the practitioner must tackle human issues which are in themselves complex and often of psychological, social, or existential nature (Dansk Selskab for Almen Medicin 2022, 6).
Furthermore, she also questioned whether a machine could utilise contextual knowledge and professional judgement, defining what can be seen as empathy as in reality being her professionalism, which is built upon the vast knowledge she accumulated:

"Can a machine translate science and facts and adapt them to each individual patient as well as I can? I have a heart, which the computer does not. However, my argument is that it is not our heart that we rely on, but what I call our professionalism, which allows us to possess extensive knowledge. Not just medical knowledge, but also the knowledge beyond."

This presents the key problem of the GP distrusting the results, which appears if the AI chatbot is applied to an area, where the contextual knowledge is important.

Typology of Patients, Cultural Understanding and Communication
Following on the topic of professional judgement, the practitioner elaborated that when it comes to the relationality, it means to take into consideration subjects such as the person's personality type as well as their fears and worries and so the context of each consultation individually.
The phronesis and professional judgement of the practitioner also come into play when communicating with patients from different backgrounds. Being able to distinguish how different cultures interpret symptoms requires a skill set of its own. The practitioner has extensive knowledge in this area due to her experience of working in an asylum centre for many years. She explained how there are many nuances in general practice, not just across continents but also within Europe's borders:

"You can't just be a general practitioner in Denmark and think that you can work in Norway under the same conditions because it's a different culture, a different way of presenting symptoms. It's completely different in Poland, and something completely different in Romania, and something entirely different in Asia … and you need to have knowledge about each culture in order to read, understand, listen, and treat correctly with the person sitting in front of you."

She continued to explain that one can even encounter cultural differences within Denmark's borders. The practitioner also provided en example which could be especially relevant in the AI debate, portraying how essential it is to understand the cultural background of the patient to provide the right treatment:

"In the Middle East, there is no [diagnosis of] anxiety; there are no mental illnesses, and such. But if one experiences what corresponds to anxiety, worries or restlessness, they come and say: 'I'm sure my blood pressure is too high.'

And I have measured many blood pressures in Middle Eastern patients, all of which were normal. You just have to say: 'Hm, you keep saying your blood pressure is too high, and I keep measuring it objectively to be completely normal. So, what does it mean when you say your blood pressure is too high? 'Ahh, well, it means that you are afraid of this and that.' "

The practitioner's phronesis, or practical wisdom, is required to navigate these differences and interpret the patient's concerns accurately. By objectively measuring the patient's blood pressure and engaging in meaningful dialogue, the practitioner can uncover the underlying fears and anxieties that the patient is expressing through their belief about high blood pressure. This demonstrates the importance of the practitioner's knowledge, experience, and ability to approach the situation with sensitivity and cultural competence. All of which are skills that need to be taken into consideration when technologies such as AI could be brought into the diagnostic process in the future.

A lack of this contextual understanding could be dangerous and lead to worse medical outcomes, since many types of patients that deviate from what could be understood as standard, go undiagnosed and untreated. This is a problem that also exists in the present and not only in regard to patients from other ethnic and cultural backgrounds, but also with patients who are not neurotypical or are socially marginalised. While textbooks and instructions provide a foundation for episteme and techne, the application of phronesis requires a more nuanced approach, as described in the ethical guidelines we presented in background section 2.3.1 Patient-practitioner Relations. This is also an issue that the practitioner mentioned:

"We overlook a significant number of common conditions among individuals of different ethnic backgrounds and socially marginalised individuals. They present their symptomatology in a completely different manner. Asperger's, autism diagnoses, and even ADHD, for that matter, present their symptom profiles differently from the standard described in textbooks."

Previously we have mentioned the problems regarding the presence of technologies such as the computer in the consultation and their damage to the relations between practitioners in patients. When it comes to this group of patients however, the practitioner can see benefits in utilising the computer and looking into the screen instead of overwhelming the non-neurotypical patients such as the ones on the autism spectrum with too much eye contact, which they often are uncomfortable with.

Algorithmic Knowledge and Techné versus Phronesis
This knowledge of particular patient types must be accounted for when emerging technologies such as AI chatbots enter the field of general practice, in order to ensure equitable outcomes for all patients. Moreover, the theoretical knowledge that we refer to as techne, is

standardised. If chatbots operate primarily based on textbook information, their assistance in health care can bring the risk of exacerbating health disparities. Furthermore, according to Mittelstadt (2021), the quantifiable nature of the machine's understanding can fail in understanding the particular contexts and particular patients and their needs. Deviations from standards is something that the practitioners must cope with daily, therefore the skills that come with contextual knowledge are essential for good outcomes:

"One can perform exceptionally well as a medical professional, adhering to all guidelines and directions whatsoever, and still overlook, lose, and mismanage a significant number of patients simply because they do not fit the standard. Our guidelines and our expertise - they are standardised. But there isn't much in this world that is standardised."

Another area where the context and demographic of patients requires phronesis is prescription of medications. The practitioner explained during the interview how the patient populations that medicine is normally tested on are usually young healthy males. Therefore, she stated that it is fundamental to understand that science cannot be applied directly, since there is a lot of bias. The practitioner compared it to taking something from an aquarium and using it in the sea, with her task then being the recognition and evaluation of which parts of the knowledge acquired in aquarium conditions can be applied in the sea, as well as when, how or whether they can be applied at all to the situations at hand.

Moreover, contextual knowledge is also important when it comes to communicating difficult information to the patients. The practitioner explained how they sometimes evaluate that a certain patient should not receive details or their diagnosis due to their age and psychological frailty. Therefore, this area also needs to be accounted for if technologies such as AI chatbots should be interacted with by the patients in diagnostic processes.

During one of the Wizard of Oz interactions, the practitioner tested the response on the chatbot to this prompt: "What can be wrong if one has a pain in the shoulder?"

She was excited to see whether the chatbot would suggest cancer, as this has been a frequent concern with using the internet. Relating to the existing technology of the Google search engine, our informants expressed the key problem, that if the technology is too quick to suggest the worst scenarios while supporting the decision making of the practitioner, it poses a risk for overtreatment or overmedication. However, it may also fail at picking up on signs of diseases and so make the practitioner overlook diagnoses.

At the same time, the practitioner also voiced concerns about the contextual knowledge of the chatbot, as it comes from experience and focuses on the context of a particular situation to make the right decision. According to the practitioner, one should be careful when challenging the chatbot when it comes to medical proficiency since the ability to distinguish between different articles is something that even themselves may not have. Furthermore, they must assess what is most relevant for each individual patient and the validity of sources is important.

Similarly, Parvianen and Rantala (2022), raised the same issue of chatbots showing a problematic deficiency in contextual knowledge as well as the aforementioned loss of quality in diagnostic practice. This pertains to the danger of reliability, which can be further enhanced by the persuasive and human-like communication of the chatbot. The practitioner stated that AI should strive to be objective instead of human-like and that the chatbot could be dangerous if the person using it for medical advice does not know much about what they are doing, with the chatbot response being just as "safe" as the question being asked.

A potential area where our informants could envision an AI chatbot as an enhancement of their expertise could be patients with concurrent conditions, helping them when in doubt. The MS

could see a potential in using less clicks than when using Google, however, the practitioner raised again the issue of expertise, which she applies when using Google search. She knows what she is looking for and which sites can be considered reliable. The reliability of chatbots such as ChatGPT is highly questionable due to the black box problem we describe in the background. There is not enough transparency from the big tech field to properly account for the logic behind the outputs of the AI chatbots and cannot figure the logic out just by observing and using the technology. Thus, according to Mittelstadt (2021) the logic is neither accessible nor intelligible.

6.2.4 PROFESIONALITY

Another important aspect of the practitioner's knowledge is professional judgement, which was a recurring theme throughout the interview and the workshop. This is also a form of phronesis as it is an executive virtue and comes into play when the knowledge acquired through textbooks or guidelines is no longer sufficient. Phronesis is built from the experiences of the practitioner and as we describe in the background section, it is an "accumulated wisdom" (Conroy et al. 2021, 1).

For the practitioner, to be professional, it means to be able to distance themselves from any initial emotional impressions they would get from the patients. Understanding, and responding properly to the emotions of their patients, is also an important element of health care (Powell 2019), that establishes trust in the care that is provided. The accumulated wisdom can also include personal aversions or prejudices. The practitioner described a theoretical example of a situation that could arise, with a patient having poor hygiene, coming in the clinic, and shouting. The way the practitioner's professional judgement then comes into play is to separate the psychological irritation the patient is causing and focus as an expert, on what might medically cause the patient to act in such a manner. They describe that in this way the clinic can hold up to treating all patients equally. Although it is not always possible to adhere to these values completely, the practitioner always does their best to find what the underlying reason might be for the patients seeking medical help.

6.3 WORKFLOW

During both the interview, and the workshop, workflow was a repeated theme. The informants often related to how their current workflow was, and by extension, it was relatively easy for them to envision how the introduction of AI chatbots might influence this in the general practice. We define workflow as activities, or procedures, that relate to the everyday function of the general practice. These can include, but are not limited to, scheduled events (Alturaiki et al. 2022), such as phone hours, or breaks, as well as tasks that are often made in the same way in the user's practice.

6.3.1 TAILORED CARE

The informants repeatedly expressed how a general practice is unique and complex. Particularly, when comparing to the workflow at hospitals, it is exemplified in how consultations are not exclusively handled by specialists or nurses; administrative personnel also have a large involvement, both in the fact that they also carry out consultations, and that they play a large part in finding the appropriate care giver, and time frame for the patient in question:

"Then we are sort of started up, where they [the patients] do not want to research any information yet. Then [the MS] assesses the most appropriate place for them, when during the day, and who of us are going to see them. And if they have booked themselves, then move around a little. It is very very complex."

Here, the general sentiment, that the general practice is more personalised towards the patients, is also expressed. Providing care, and having a degree of compassion and consideration, is something that extends beyond the allotted time window of the consultation. In our background chapter, we introduce two cases, presented by Schario et al. (2022), where patients are apprehensive in their engagement with the health care sector, and that providing care, specifically around the consultation, can be utilised for the benefit of the patient. While our informants express that it is a complex undertaking, Schario et al. (2022) explain how taking steps towards not restricting caregiving to an appointed consultation can result in long-term benefits. This point is also expanded on, when relating to concrete consultations. As described previously (6.2.2 Phronesis Expertise and 6.2.3; Perceived Expectations and Performativity), our informants express how the assessment of patients starts in the waiting room.

These kinds of extracurricular activities illustrate that the workflow in a general practice can be very hectic, and that one day in the practice is never like the last. This may be due to the requisite elements of being a general practice; having to deal with, as the informants mention; "5000 different ways" of communicating with patients, or it may be based on the personalities of the clinicians. The GP mentions how they ". . . cannot follow a flow chart. I skip past all the time.", relating to differences between their way of working, as compared to their perception of nurses and specialists at hospitals. Additionally, they point out that the universities do not educate for the real-life circumstances of a general practice, particularly when it comes to activities that are not based in their medical professionalism. In our background we illustrate how chatbots show potential when it comes to administrative tasks, and self-learning aspects of AI chatbots might be able to take the humane elements of administrative tasks into account. This is also one of the areas where the informants envision the implementation; at the workshop they requested that the chatbot, as an exemplary artefact, should preferably not interfere directly with their patient interaction, but that it could be a quiet listener, that gets to know their preferences, guidelines, and behaviours:

"Then I might want to have one [a chatbot] that listened to my consultation, and everything I said during, it would list and make the note, and then I could just press accept on the referrals, I wanted it to make, and then it would make it for me. . . . It should be in the consultation for around half a year to get to know me, so I can trust that it is doing what it should."

Interestingly, this illustrates a degree of insights into the neural, or sub-symbolic, potential of AI chatbots, not only that it could be able to listen and take notes, which a symbolic model with speech-to-text capabilities could carry out, but that it should be able to essentially understand the sentiments expressed by the user.

6.3.2 PLANNING

During the workshop, the informants expanded on aspects of planning, and workflow, engaging in a degree of envisioning how AI chatbots might be able to benefit them in the general practice. Drawing the card of Excitement in the PLEX scenario technique, they comment:

"You should avoid that, so you do not get into that [excitement]. It should make our weekday predictable."

The informants further envisioned the idea of using the chatbot as a scheduling tool, that could assist them in planning their days:

"[It should] ensure that we do not overbook ourselves. Now we are going to book two of [the same types] of assessments in a row. Are you sure that you have the time for that? You also need your lunch. Or, it is the third day, this week, that you do not have your lunch. . . . It should dispose of our time. So it will throw us home, and shut off the entire system. . . . a checklist, to keep track of what we need to do."

This envisioning relates to how the chatbot should have control of their system, in order to decide, and enact, the plan for the day. Much like smart devices, or the internet of things, they envision a system that, not only, gives them advice or writes out a plan, but has more practical uses, such as turning off the lights, nudging them towards certain desired behaviours. This type of interaction has similarities to what the informants expressed during the interview, where they mentioned how they had placed their computers in consultation rooms, so that it would be hard to use it, while also conducting a consultation.
Using the available technology of the workshop, ChatGPT, they narrowed down their envisioning to a text-based plan of their day. Specifically in relation to phone times, they envisioned how the chatbot could assist in planning, providing suggestions in a systematic way, that they might not have thought of:

"It might be, that it could tell me, that it is not such a good idea that [I] talk on the phone from eight to eight thirty. . . . Or this many patients have not had contact with you today but might need to. It makes you think differently, or specifically, more systematically, it could provide suggestions, you had not thought of yourself."

During the second round of Wizard of Oz, they made their ninth prompt as "Can you plan my day". To the output from this prompt, they had some grievances with the chatbot, saying that it "is hard on us", and that it "does not know us". The comments were specifically related to how the chatbot gave them long work hours, and that it did not include time to receive phone calls. While these factors could subsequently be adjusted, it ultimately illustrates how they envision a system that could take their particular general practice, and elements specific to it, and tailor a weekday to their wants and needs.

6.3.3 GUIDELINES

During the interview, the informants relayed what entities had influence on the workflow at the general practice. The Agency for Patient Safety (Styrelsen for Patientsikkerhed) was mentioned as an entity that involved some strict and rigid guidelines, which were important to consider at the general practice:

"Other than the Agency for patient safety, that is certainly based on rigid structures, then…"

The statement was expressed in a conversation about professional judgement, in relation to patient interaction, and was given as an example of something that would not require their own professional judgement, but rather could be an element to inform potential systems to either ensure that the guidelines were followed or inform personnel at the general practice on proper procedures. The point was later expanded on, in relation to Danish Society for General Medicine (Dansk selskab for almen medicin 2022), where the thought of robots, or chatbots, were introduced:

"I mean, DSAM's guides are gradually becoming extremely complex. . . . and if you can use a robot like this to help finding the proper level from some patient parameters, it would be very helpful."

This underlines the frustrations that were present at the general practice, about how certain outside entities had an impact on how the workflow was. These outside entities particularly came to the forefront, when they mentioned the concept of service registration (ydelsesregistrering), to which their frustrations were embedded in the practicality, the economy, and the public optics related to certain required procedures:

"The service registration takes up a whole lot of effort. And every time there is a story in the media, about how GPs have accounted for x amount of millions for faulty calculations, yes, but that is multiplied by 10, and, on average, we [GPs] have that much missing, because of some technicality or other. So, we are actually getting [paid] for much less than what we all actually work for."

Dismissing the technicalities, this quote clearly illustrates a frustration with certain faulty calculations, navigating within bureaucratic structures, that, with the help of some form of reliable automation, might be avoidable. Additionally, it underlines the continuous disapproval for external entities having too much influence on how they run their general practice. Towards the end of the workshop, they reiterate their statement, mentioning how the chatbot might not be viable for general practices:

"[the chatbot would be] better in more limited specialisations. Our specialisation is a lot about the exceptions, those who do not fit into the boxes, where you should not follow the guidelines. We are too lopsided."

This comment largely related to using chatbots during a consultation, as opposed to using it in surrounding activities, and illustrates a common theme of lacking professional capabilities, of the chatbot within health care. As presented, the envisioned artefact and its uses, largely concern administrative or planning, activities, which aligns with our findings relating to highest potential for these activities (Mittelstadt 2021).

6.3.4 SAFETY NET

When speaking of the workflow of the informants, and an idea of their envisioned exemplary artefact of chatbots, the theme of alarm bells, or a kind of safety net, was introduced:

"It could be nice, if it could put up alarm bells, and help finding solutions, schedules, and structures: knowledge seeking."

These alarm bells were envisioned to be safety nets that could catch mistakes, or avoid risks, and inform the general practice on how to combat them. It was expanded on, as a tool to ensure patient safety, as well as remind them of loose ends to be tied up:

"I would use it if it could give me better patient safety, better treatment, and more control; fewer loose ends, and fewer risk factors."

In the second round of Wizard of Oz, while not relating to surrounding activities, the practitioner and the MS further explore the idea of using the chatbot as a safety net. In relation to an output concerning medication, they ran into an issue of potential drug interaction, aligning well with how Daniel et al. (2022) presents the benefits of chatbots in relation to pharmacology. An idea of alarm bells for drug interaction was envisioned, and they commented that something similar was already implemented:

"Our patient systems actually already do that. So, if you prescribe simvastatin, and they already receive amlodipine, it [the system] will go 'woah [wait]'."

Continuing the theme of existing factors that could become features for a future chatbot, they mention a tool called the visitation guide (visitationsguiden), which is a tool for MS, used to commit to an informed decision on how to treat a call from a patient. This was brought up, when exploring how a chatbot might be used on their own website, as a mediator between patients and the general practice:

"It already exists in the form of the visitation guide, where professional personnel can take a call and ask questions to clarify what time to schedule or if it is necessary to call 112 [emergency services]."

The aspect of using AI chatbots for such tasks was also introduced in our background chapter by Avila et al. (2022), in relation to frequently asked questions, however our informants seemingly did not view this as having a large potential, and the closure mechanism of "it already exists" was reached.

6.4 EFFICIENCY AND EFFECTIVENESS

The workshop introduced noteworthy findings, when compared to the findings presented previously in section 2.1 The Emergence of Artificial Intelligence in Health Care. Particularly, when it comes to the area concerning efficiency and effectiveness. As presented in 4.1 Case Study and Field Introduction general practices in Northern Jutland are coming under pressure due to workload and these tendencies are expected to worsen (Praktiserende Lægers Organisation 2022). This sentiment was also something expressed by our informants interpreting the technology of the AI chatbot as a potential help with exactly the tasks related to making their work easier and more effective. However, this is especially an area where ethical considerations are relevant, since using AI chatbots for streamlining should not affect the patients and the care they receive.

In the background chapter, we also presented how AI chatbots lack emotional intelligence (Powell 2019), and GPs are generally also sceptical concerning the accuracy in terms of

assessing patients emotionally, and symptomatically (Palanica et al. 2019). The informants from the general practice address these points, during the interview, when speaking about the importance of the phronesis, or authority of the GP in terms of sorting patients as elaborated upon in section 6.2 Relations. The unique benefit of 'the old' archetypical practitioner was a much more paternalistic authority, as we have introduced in 2.3.1 Patient-practitioner Relations. Introducing new technologies can potentially threaten this authority, less efficient process of tasks such as assessing ailments or prescribing drugs. This sentiment is continuous throughout our work with the informants, particularly for the GP, who towards the conclusion of the workshop reiterates an apprehension for implementing technological capabilities at a general practice:

"You need to watch out for what you are able to do, because it can trigger a lot [of tasks]."

This presents a key problem concerning effectiveness and efficiency, and the implications that being a more effective institution might trigger additional work loads, and responsibilities. This worry is well founded, as technological solutions increase effectiveness, risk cynical approaches to employees and ultimately reduce their well-being (Parviainen and Rantala 2022).
In terms of phronesis, it is indicated that the implementation process should be carefully considered, to avoid poor professional practices, and long-term side effects, harming both professionals and patients (Parviainen and Rantala 2022). During the workshop, the informants expressed scepticism in relation to the capabilities, and prescription of AI chatbots directly:

"It is actually not a benefit to act like humans, because we need something else."

The strengths of general practitioners, and the scepticism towards capabilities of technology replacing the work tasks, and phronesis, illustrate that the sub-symbolic AI approach to chatbots (Kelley 2003), in terms of replacing GPs, causes significant apprehension amongst our informants. This apprehension is reinforced throughout the workshop, particularly during the Wizard of Oz exercises. After their fifth prompt; "what is the waiting period for a dermatologist in the northern Jutland region?", and receiving a response from the chatbot, the GPs explains:

"It [the chatbot] has sort of general knowledge, but it does not have specific knowledge of Northern Jutland. And if it should do anything, then it is precisely the specific changes, existing in northern Jutland, as compared to the general, and where it can match the specialised physician level. That stuff [chatbot output] is layman level."

The GP points to the bulk of inputs and outputs, concerning medical procedures, symptom checking, and patient referrals, indicating a frustration with the capabilities in these areas. There is a particular scepticism towards the professional capabilities, as well as the relations with patients (6.2.3 Perceived Expectations), which are essential during consultations, leading to our informants wishing to explore capabilities surrounding the consultation. These capabilities envelop tasks that are related to administrative work, particularly tasks that are considered low skill, and time consuming (Mittelstadt 2021; Schario et al. 2022). While there has been a theme of optimising the workflows, which might be considered increasing efficiency and effectiveness, this section will present some considerations that relate to problems, worries or quasi-technical requirements. As a result, this chapter will address some

discrepancies between the technology used in gathering empirical evidence, GPT-4, and the wishes expressed by our informants.

6.4.1 RESPONSE TIME

In terms of improving traditional treatment, the first prompt, made by the informants, initially concerned medical procedures, specifically the removal of a peanut in the ear of a patient. While our informants agreed with the general procedure, suggested by GPT-4, the answer was considered "layman", as it did not address important factors such as a peanut being a biological foreign object, as well as a frustration with "small talk" in the answer. This was reiterated for the second prompt in the workshop, requesting guidance for a person with high blood pressure, and reduced kidney function. For the output produced for this, the frustration with response time was addressed specifically:

" . . . And then it is too much to read if you have to find that ACE II."

We classify this point broadly as response time, since there is both a process of the chatbot writing an output, and the process of the user reading the output. The key problem with outputs being too wordy, particularly concerning liability disclaimers on the side of the chatbot, was a recurring one, and while it is possible to request generally shorter answers from GPT-4, this was not obvious to the informants, resulting in long respond times for the prompts. The informants later compare the technology to technological solutions, that they already utilise, more specifically the Google search engine, where they express that for the chatbot to be useful, it should be more efficient than Google, in relation to amount of clicks, and time used in the interaction. Additionally, the GP mentioned that they knew what sources they trusted, so at the time, a search on Google would require shorter time investments.

6.4.2 VIRTUAL ASSISTANT AND CONTEXTUAL KNOWLEDGE OF THE USER

During the debrief of the workshop, the practitioner expressed a central key problem, in terms of what they would want from a technological solution:

"I had some prejudice about how it was not very good at answering. I would like such an assistant that could perform some tasks. . . . It can become just another computer. [it can become] an element of irritation."

The GP expresses that introducing technology, just for the sake of introducing it, becomes an element of irritation, and if it does not benefit them, or if it is not integrated with how they do things at the general practice, it becomes an element of irritation, rather than a positive addition.

Expanding on the element of a personalised assistant, as an exemplary artefact, our informants repeatedly requested features relating to adaptability to the context and user. They mention that an AI assistant, or chatbot, should have knowledge of local systems and guidelines within the general practice, surrounding institutions, and the specialty of the users. The MS specifically mentions that the chatbot should know whether it is a GP, MS, CN, or resident using it, so it might adjust the expertise to said user. Additionally, the practitioner mentions significant differences between patients, and how they explain things:

"It [the chatbot] should be adapted and understand the user. They [the patients] have 5000 ways of expressing the exact same need."

This is taking the perspective of using the chatbot as a mediator between GP, and patient, yet ultimately underlines a general point, that the general practice is unique and a chatbot should be aware of this. They later elaborate that they believe the chatbot to be " . . . Better in more limited specialisations.", and that their specialisation " . . . is very much about the exceptions. Those who do not fit within the boxes. Where you do not follow the guidelines.". These comments related to their interpretative flexibility, in relation to how the specialisation of the GP is more complicated than other medical specialisations, leading to the closure mechanism of the possible use of the chatbot, in their context, since it did not meet the expectations, in terms of techné, episteme, and phronesis, desired by the GP. We previously presented the opportunities for administrative tasks, illustrating the interpretive flexibility, present at the general practice.

6.4.3 RADICAL INVENTIONS

In relation to SCOT, the concept of *radical inventions*, as compared to *conservative inventions*, is introduced (Bijker et al. 1993, pp. 51-82), where radical inventions are technologies that disrupt the status quo, and conservative inventions improve upon it. These two concepts of inventions can help in explaining why there are some discrepancies between the properties of the technological artefact, and the wishes from the general practice. GPT-4 may be considered a conservative invention, within the space of large language models (LLMs), in that the technology increments, in a way, on the previous models, and attempts to improve the on the language models, to better perform tasks. However, the introduction of AI chatbots within general practice, and specifically in the context of our case, is certainly very disruptive, primarily because there is no equivalent technology in place, but additionally, the technology has potential to solve a multitude of tasks that might not previously have involved technological artefacts. The informants did somewhat compare the technology to search engines, more specifically Google, remarking the closure mechanism, that the chatbot should require "fewer clicks than Google", which may imply that their use of the chatbot was based on a presumption that the chatbot would serve the same tasks. This presumption is certainly reasonable, considering how the LLM is trained on internet resources, and the implications of similar input/output user interfaces, yet the capabilities of the AI chatbot do exceed those of a search engine, when it comes to the ability to make sense of data, and formulate sentiments. Since there is no direct comparison to AI chatbots in their current practice, our informants were somewhat apprehensive in exploring the radical impacts that could be introduced by the technology, this was particularly prominent in relation to interpersonal connections, and 'human' qualities:

"It [the chatbot] should also not destroy the social feeling, but that is not its task."

An argument can be made for LLMs to be specifically designed with these social interactions in mind, attempting to imitate human language, however the informants express a definite prescription that it is not the chatbot's purpose to act in such manner. This can be seen as an issue of the radical invention being inaugurated as an entire new system, rather than developed to improve upon existing structures (Bijker et al. 1993, p. 57), and as a result, the existing structures are not predisposed to nurture the introduction of the invention.

7. DISCUSSION

Following the analysis of the empirical material, there are some discussion elements, regarding the concepts of radical and conservative inventions, as well as some unexpected findings, which we will unpack in the following chapter.

7.1 RADICAL OR CONVERSATIVE INVENTIONS

The analysis presents the concepts of radical and conservative inventions (6.4.3 Radical Inventions). We briefly introduced that the existing structures within the general practice were not predisposed to nurture the introduction of radical inventions. In this section of the discussion, we will unpack this point further.

We define AI chatbots as a radical invention, due to how it presents the possibility to solve tasks that were not previously possible to solve technologically. AI chatbots are, at the time of writing this thesis, the closest thing to imitating human interaction through text. As an LLM, It has the capability of accessing large amounts of information, and disseminate it concisely in response to a user input. However, our informants had apprehensions towards envisioning the implementation of the chatbot, which we attributed to the fact that there were no comparable technologies in place. This presents the social constructivist argument, for no technological intervention to take place, since the general practice is functioning well already. However, if we entertain a more technologically determinist approach, it presents the question of whether the technology is superfluous or redundant, requiring an entirely different technology to reach desired outcomes, or if the features of the technology enable it to cater to user needs. An argument could be made for the latter, since GPT-4, and other versions of GPT, are pre-trained on large amounts of data, enveloping many different fields of expertise, and with the technical know-how, a system, powered by GPT, could be tailored to the desires expressed by the general practice.

However, much of the reluctance, concerning our informants' use of the chatbot, related to how it was too general, and did not take into account the unique nuances of the general practice. As such, the countering argument to tailoring a ChatGPT variant would be that the sheer amount of training data, as well as a predisposition to not provide medical advice, due to a risk of misinformation, makes for a too broad foundation. Our empirical evidence points towards a system, which is trained on specific data, both relating to the context of the general practice in question, but also limited to the sources of medical knowledge that they trust.

EXPECTATIONS AND RESULTS

When we conducted our literature review, we prioritised and sorted through text materials based on expectations of what is relevant to our problem statement and what topics we expected to discuss with our informants. However, as mentioned in the Methods section, we chose an exploratory approach and let the social groups, our informants, show us how they interpret the technology, its problems, and the possible future. All the while, we also had our personal preconceived notions. As we write in 4.1 Case Study and Field Introduction, Flyvbjerg (2006, 237) addresses the subject of preconceptions through the fourth misconception regarding the bias toward verification of these. He then argues that in fact the opposite usually takes place and these notions become instead falsified and it is exactly the activity of casting off these preconceptions where the learning takes place and new insights are achieved (ibid., 236-237).

In hindsight, this methodological point is also something that describes our process very well. Although we did not form any explicit hypothesis, we did have various expectations about the

practitioner's Wizard of Oz interactions, evaluation of the technology as well as where they might identify the key problems or how they might imagine the exemplary artefact.

The first counterintuitive outcome was that the practitioner did not express any interest or worry in the possibility of patients self-diagnosing with AI chatbots such as ChatGPT. They acknowledged that self-diagnosing has always been an issue, even in an analogue way through conversations with one's family or neighbours, and with the technology through the Google search engine. However, considering how the informants described meeting of expectations as a crucial element in consultations and how nuanced and sometimes problematic it can be to tackle the situations when patients believe something is wrong and they need medical attention, when in fact they need reassurance and being heard, it is noteworthy how quickly they disregarded the technology's threat to perpetuation of these issues. They explained how the internet makes the patients tremendously worried as the worst scenarios pop up as the first search results.

In addition, the practitioner also mentioned internet of things and wearables giving patients worries: "They have access to all kinds of data that they don't know how to respond to."

The only remark they had regarding ChatGPT was saying how self-diagnosing already is a problem.

Some of our preconceived notions were thus that the practitioner would become much more worried about AI Chatbots being accessible to patients. As already mentioned in the Analysis, they acknowledged its humanlike manner of communication as well as persuasiveness and untrustworthiness could perpetuate problems related to self-diagnosis. Although there also exists a possibility that we failed to portray the capabilities of the technology properly that could explain the disinterest, we can still conclude that what we perceived as a potentially strong threat and expected catastrophic future scenarios from the practitioner's perspective seems to simply be a human tendency that is well known to the practitioner, taking place with and without technology although still made worse in its presence. Therefore, we want to highlight the value of the case study approach and acknowledge how it led to many specific and in this case unforeseen outcomes that aided our learning.

Furthermore, when it came to the Wizard of Oz interactions during the workshop, we observed that the practitioner often sought help with tasks that would require the chatbot to have access to local and practice-specific knowledge and looked for ways its answers could be relevant in making their job easier. Here we expected that the practitioner would attempt to challenge the chatbot more on its techné to get a sense for the depth of its knowledge, such as a follow-up on the interview where the need for Google search engine was mentioned regarding rare diseases. This was however again an important learning as it shows how the practitioner interprets and thus constructs the technology, showing authentically which areas are most relevant or problematic for their practice and particular context.

Additionally, the informants did not seem to perceive the particular technology of AI chatbots as a threat, rather they described the presence of technologies, both in regard to relationships as well as their workflow and general operation as problematic, presenting us with the key problems. The informants portrayed how the growing presence of technology creates a shift in the relationship where the patient gets access to large amounts of information, data, and previously unseen options which they can demand. This finding was also partially unanticipated, since we expected more emphasis on what the consequences will be for the practitioner when patients get access to such powerful technology. In its current form it was not perceived as a threat, but neither as a miracle solution for the problems of the social groups. In this sense there was more emphasis on the exemplary artefact and its potential applications within the practice, imagined outside of the current technological bounds.

This leads us to the next unforeseen finding, which is how the technology was imagined as an omnipresent tool in the future, a tool that is one step ahead, making the work of the practitioner easier through anticipation. The practitioner envisioned how the technology should not be in the way of the physical, face-to-face interactions and thus they do not desire to directly engage with it as much.

Finally, when first learning about and presenting phronesis, we expected to see how or whether it could be supplemented or potentially substituted by machine intelligence. Or rather, in what form will it exist in the presence of such intelligence. What we learned instead is that in the context of the particular practice, phronesis is still a human skill in a very deep and profound way and should be protected as such. It consists of various sensory inputs, accumulated knowledge, continuity and thus very specific contextuality. The skill was described as so nuanced that the practitioner did not see it as being under threat in any way or transforming any time soon.

8. CONCLUSION

In line with Flyvbjerg (2006), our case study does not attempt to prove specific claims but rather to gain deep insights and context specific knowledge, as inherent in case studies. Through the knowledge acquired in this thesis, we aim to contribute to the ongoing and dynamic development of the field of AI chatbots through a social-constructivist lens.

To reaffirm the purpose of this study, we will restate the problem formulation and research questions, which are as follows:

"How might the future of rapidly evolving AI chatbots, in the context of a general practice in Northern Jutland, be co-envisioned?"

➢ How can social and cultural aspects within the social group(s) inform the construction of the technological artefact?
➢ How might the findings within this thesis inform the sustainable and responsible implementation of AI chatbots in general practice or health care?

In addressing the first research question, which focuses on the influence of social and cultural factors on the development of a technological artefact, we conducted an ethnographic case study at a general practice in Northern Jutland. Through an interview and a workshop session, we employed the methodological and theoretical perspectives of ethnographic co-envisioning and the Social Construction of Technology (SCOT) framework, respectively. Our objective was to examine how the social groups at the general practice envisioned the future of their practice in the presence of AI chatbots, specifically ChatGPT. We discussed the conventional retrospective application of SCOT and utilised its key concepts of interpretative flexibility, closure and stability, technological frames, and social groups to analyse the empirical data we collected. While SCOT provided the theoretical basis for framing our findings, the rapidly evolving nature and early-stage implementation of AI chatbots in medical contexts necessitated a more forward-looking approach. To address this, we employed the methodological framework of ethnographic co-envisioning, introduced in An Anthropology of Futures and Technology (4.2 Methodology). This approach allowed us to envision the uncertain future implications of AI chatbots as a technological artefact. We conducted interviews to gain contextual insights into the general practice, which informed the design of the workshop and provided analytical understanding of the current state. The subsequent workshop aimed to co-envision future scenarios with the participants using methods such as the PLEX scenario technique and Wizard of Oz, enabling exploration and enactment of potential scenarios.

Regarding the second research question, which pertains to the way the findings in this thesis can inform the sustainable and responsible implementation of AI chatbots in general practice or healthcare, we present our specific empirical and analytical findings. When addressing the aspect of sustainable implementation, we consider sustainability in terms of developing technology that will remain viable in the future, particularly within the context of the social groups at the general practice. On the other hand, responsible implementation refers to achieving an ethical, social, and culturally appropriate integration.

9. REFLECTIONS

As techno-anthropologists, we seek not only to focus on areas of *technical rationality*, but rather to operate within a framework of being *reflective practitioners* (Schön 1992). As reflective practitioners we engage in a practice of *reflection-on-practice*, where the purpose is, to recognise the tacit aspects of our professional practice (ibid.). Additionally, we also *reflect-in-action*, involving the construction of appropriate methodology, and theory, for the given case in question (ibid.). The following reflective chapter will make these aspects explicit, to illuminate how the project was constructed, what might have been done differently, and how this study might contribute to further creation of knowledge.

9.1 THE UNPREDICTABLE FUTURE

The unique value of a techno-anthropological approach is that we seek to bridge gaps between deterministic technological views, and more abstract ideas of technology, concerning empirically grounded observations. During this study, we adopted the methodological and theoretical approaches of ethnographic co-envisioning (4.2 Methodology), and the social construction of technology (SCOT) (5. SCOT Theory), to approach the rapidly developing technology of AI chatbots by grounding our findings in empirical material, while making it tangible through the theoretical framing of SCOT. In alignment with co-envisioning, we seek to challenge the approach of technological determinism, however, we do not suggest a radical social constructivist intervention,but rather to make these socially constructed co-envisionings tangible, and able to inform the development of technology.

Throughout the case study, we experienced struggles, particularly relating to finding existing theoretical frameworks that were inclined towards imagining futures in case-specific studies. As part of our reflection-on-practice, we eventually settled on the theoretical framework of SCOT, with the imperative marriage with our methodological approach of ethnographic co-envisioning. This was done to repurpose the conventionally retrospective approach of SCOT with an anthropological methodology that enabled us to investigate futures, as part of working with a rapidly evolving technology, with a temporal context that is difficult to predict.

Practically, we introduced the approach of engaging with the technology as a snapshot, where we used the cutting-edge state of our technology, GPT-4, to explore potential impacts of this state. This was possible as a direct result of inherent qualities, and features, of the technology in question, AI chatbots based on Large Language Models, as they seek to be widely applicable to different contexts. Subsequently, we were able to use the inherent features to apply the technology to our case study, by prompting the chatbot to act as an artificially intelligent assistant, in general practice in Denmark, who has the purpose of advising a general practitioner in their workday.

9.2 EMPIRICAL DATA GENERATION

As part of reflecting-in-action, we made certain decisions, during the case study, that impacted how the project developed. Particularly, when it came to the methods we used, and how we employed the methods, we engaged in certain techniques that proved to have an impact on the generated empirical data. For the workshop, we used the PLEX scenario technique (Lucero and Arrasvuori 2010), as a means to engage our informants in constructing real-life scenarios, as part of the process of co-envisioning (Lanzeni et al. 2022). During the workshop debrief, the informants expressed that they did not particularly see the value of the PLEX scenario technique, but rather enjoyed the Wizard of Oz, and the conversations surrounding the workshop. This resulted in empirical data, which was less tied to the resulting constructed scenarios, while more reliant on the conversations. This led to a coding process,

and subsequent thematic analysis, which did not involve their constructed scenarios. Whether the reasons for this reluctance were not understanding the rules, or simply not enjoying the technique, the conversations during the process turned out to be valuable. However, an alternative approach to envisioning these scenarios could have been beneficial, since we wanted a reciprocal relationship with our participants, not only in providing efficacy in terms of allowing them to express their opinions and envisioning, but also in relation to providing a pleasant experience.

Another reflection on the workshop was that the participants appeared to struggle with how to formulate prompts in ChatGPT, which was also a point brought up by (Parviainen and Rantala 2022). We, as facilitators of the workshop, had insights into how to formulate certain prompts, to receive more contextually appropriate and useful responses, as well as to remedy some of their issues with the chatbot. To allow our informants to be the primary influence on the empirical data, we decided not to intrude, and suggest solutions for their issues. Ultimately, we found this approach to be preferable, since the frustrations, on the part of our informants, provided insights into fundamental discrepancies between the features of the technology, and the desires of our informants. Also, it allowed us an authentic view of how the informants interpreted the technology. This also led to the finding, relating to radical and conservative inventions, that the informants approached the technology through their most relatable comparison, the Google search engine. However, for future studies, it might be valuable to briefly educate informants on how to create prompts, to be able to let them investigate the capabilities of the technology to a fuller extent.

9.3 FUTURE WORKS

This case study relied on investigating a specific context, to contribute to the scientific corpus of related research. We could not cover all aspects of all the circumstances relating to our research area and as a result, we delimited the study to the context of a case. With this in mind, we have some propositions for research that could expand on this study, to unpack the topic further.

As part of our theoretical framing with SCOT, we identified a multitude of different social groups, many of which we chose to exclude from the case study, as part of delimiting our case. Notably, we chose to not focus on the patient perspective at the general practice, which was primarily as a result of this perspective being very prominent focal points within the related works that we presented in our background chapter. Since relations were a recurring theme, and patients are an essential element within a general practice, committing to investigating how the technology impacts the perspective of the patients, could provide supplemental insights into these relations.

Additionally, combining the two perspectives, of the general practice, and patients, respectively, could contextualise the relations to a higher degree. Through observational studies of a consultation, at a general practice, the study of the user's practice could be unpacked with a higher degree of validity, as it would provide direct insights, as opposed to verbal exchange.

A direct approach to continuing the study into the general practice, would be to study all the different social groups, that were identified, as part of SCOT. Our ethnographic process was structured through an interview and a workshop, which included two of the social groups, the GP, and the MS, respectively. However, including these two social groups in the same conversations, led to consensus in interpretations of the technology, which limited the potential to investigate the flexibility of its interpretation within the general practice. Isolating the social groups, as part of generating empirical data, could contribute to a deeper understanding into the interpretative flexibility that is present amongst the social groups.

Another social group, that could be investigated, as a next step, would be developers of the technology. Traditionally, a more determinist social group, it could provide a completely different perspective on the area of AI chatbots in general practice, and it would provide a perspective from a group that inherently has deeper insights into the technology. One could additionally provide this thesis, and discuss the results, to engage with the findings.

Further, one could also structure workshops, including multiple social groups, to facilitate dialogue of potentially contradictory perspectives, which could inform aspects of social constructivist ethnographic co-envisioning. Potentially involving a different general practice, to investigate whether the findings, made in this thesis, are congruent or divergent with other, similar, contexts, could contribute to an understanding of how general practices relate to AI chatbots, in a more general sense.

Lastly, we have repeatedly referred to AI chatbots as a rapidly evolving technology, particularly in recent years. We engaged with GPT-4, which was the cutting-edge state of available AI chatbots, at the time of writing this thesis. Repeating, and reiterating case studies, similar to this project, as the technology evolves, could provide interesting findings into how perspectives might change, and the impacts brought on by the continuous evolution. Additionally, the introduction of the publicly available ChatGPT has accelerated discussions surrounding the concept of *Artificial General Intelligence* (AGI), which seeks a model for recreating, rather than imitating, human intelligence (Figure 9.1). AGI would present an even more radical invention that could bring on a paradigm shift, which would further reinforce the importance of being able to study futures, and rapidly evolving technologies.
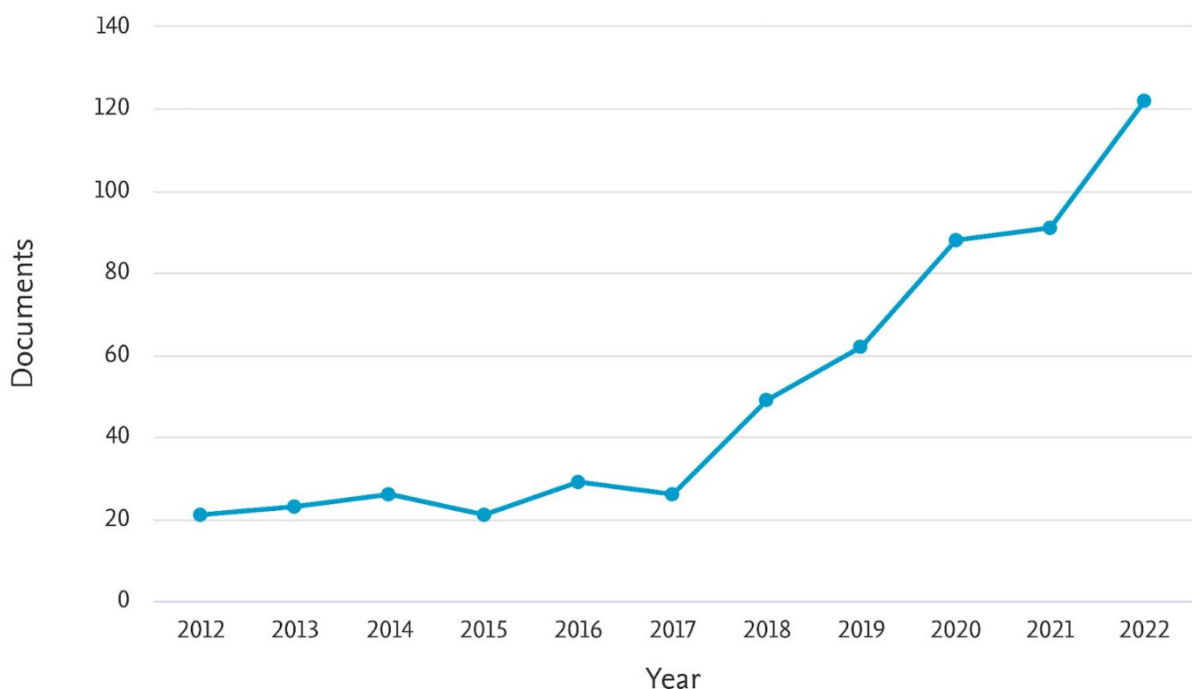

Documents by year

Figure 6:Document entries on Scopus, concerning the exact literature search of "Artificial General Intelligence", using the Analyse function to illustrate documents released by year, within the scope of 2012-2022.

BIBLIOGRAPHY

Allen, Justin, Bernard Gay, Harry Crebolder, Jan Heyrman, Igor Svab, Paul Ram, Philip Evans, et al. 2011. The European Definition of General Practice / Family Medicine: Network organisation within WONCA Region Europe.

Alturaiki, Alma Mohammed, Haneen Reda Banjar, Ahmed Salleh Barefah, Salwa Abdulrahman Alnajjar, and Salwa Hindawi. 2022. "A Smart Chatbot for Interactive Management in Beta Thalassemia Patients." International Journal of Telemedicine and Applications; Int J Telemed Appl 2022: 9734518-13. doi:10.1155/2022/9734518.

Avila, Francisco R., Daniel Boczar, Aaron C. Spaulding, Daniel J. Quest, Arindam Samanta, Ricardo Torres-Guzman, Karla C. Maita, John P. Garcia, Abdullah S. Eldaly, and Antonio J. Forte. 2022. "High Satisfaction with a Virtual Assistant for Plastic Surgery Frequently Asked Questions." Aesthetic Surgery Journal; Aesthet Surg J 43 (4): 494-503. doi:10.1093/asj/sjac290.

Bijker, Wiebe E., W. E. Bijker, and Wiebe E. Bijker. 1995. Of Bicycles, Bakelites, and Bulbs : Toward a Theory of Sociotechnical Change. 1995th ed. Cambridge, Mass: The MIT Press.

Bijker, Wiebe E., Thomas P. Hughes, and Trevor Pinch. 1993. The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology. 4th ed. The MIT Press.

Bijker, Wiebe E., Thomas Parke Hughes, Trevor Pinch, and Deborah G. Douglas. 2012. The Social Construction of Technological Systems, Anniversary Edition: New Directions in the Sociology and History of Technology. Cambridge: MIT Press.

Brown, Tom B., Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, et al. 2020. "Language Models are Few-Shot Learners." . doi:10.48550/arxiv.2005.14165.

Bruhn, J. G. 1978. The Doctor's Touch: Tactile Communication in the Doctor-Patient Relationship. Vol. 71. Birmingham, Ala.] : Southern Medical Association. doi:10.1097/00007611-197812000-00008.

Butler, Judith. 2006. Gender Trouble: Feminism and the Subversion of Identity. Florence: Routledge. doi:10.4324/9780203824979.

Conroy, Mervyn, Aisha Y. Malik, Catherine Hale, Catherine Weir, Alan Brockie, and Chris Turner. 2021. "Using Practical Wisdom to Facilitate Ethical Decision-Making: A Major Empirical Study of Phronesis in the Decision Narratives of Doctors." BMC Medical Ethics; BMC Med Ethics 22 (1): 16. doi:10.1186/s12910-021-00581-y.

Daniel, Thomas, Alix de Chevigny, Adeline Champrigaud, Julie Valette, Marine Sitbon, Meryam Jardin, Delphine Chevalier, and Sophie Renet. 2022. "Answering Hospital Caregivers' Questions at any Time: Proof-of-Concept Study of an Artificial Intelligence–Based Chatbot in a French Hospital." JMIR Human Factors 9 (4): e39102. doi:10.2196/39102.

Danieli, Morena, Tommaso Ciulli, Seyed Mahed Mousavi, Giorgia Silvestri, Simone Barbato, Lorenzo Di Natale, and Giuseppe Riccardi. 2022. "Assessing the Impact of Conversational Artificial Intelligence in the Treatment of Stress and Anxiety in Aging Adults: Randomized Controlled Trial." JMIR Mental Health 9 (9): e38067. doi:10.2196/38067.

Dansk Selskab for Almen Medicin. 2022. Professionsetik for Praktiserende Læger i Danmark.

De Angelis, Luigi, Francesco Baglivo, Guglielmo Arzilli, Gaetano Pierpaolo Privitera, Paolo Ferragina, Alberto Eugenio Tozzi, and Caterina Rizzo. 2023. ChatGPT and the Rise of Large Language Models: The New AI-Driven Infodemic Threat in Public Health. Vol. 11. Lausanne : Frontiers Editorial Office. doi:10.3389%2ffpubh.2023.1166120.

Easton, Katherine, Stephen Potter, Remi Bec, Matthew Bennion, Heidi Christensen, Cheryl Grindell, Bahman Mirheidari, et al. 2019. "A Virtual Agent to Support Individuals Living with Physical and Mental Comorbidities: Co-Design and Acceptability Testing." Journal of Medical Internet Research; J Med Internet Res 21 (5): e12996. doi:10.2196/12996.

Emanuel, E. J. and L. L. Emanuel. 1992. "Four Models of the Physician-Patient Relationship." Jama 267 (16): 2221-2226.

Flyvbjerg, Bent. 2006. "Five Misunderstandings about Case-Study Research." .

Kelley, Troy D. 2003. "Symbolic and Sub-Symbolic Representations in Computational Models of Human Cognition: What can be Learned from Biology?" Theory & Psychology 13 (6): 847-860. doi:10.1177/0959354303136005.

Kim, Yoojoong, Jong-Ho Kim, Young-Min Kim, Sanghoun Song, and Hyung Joon Joo. 2023. "Predicting Medical Specialty from Text Based on a Domain-Specific Pre-Trained BERT." International Journal of Medical Informatics (Shannon, Ireland); Int J Med Inform 170: 104956. doi:10.1016/j.ijmedinf.2022.104956.

Lanzeni, Débora, Karen Waltorp, Sarah Pink, and Rachel C. Smith. 2022. An Anthropology of Futures and Technologies. Milton: Taylor & Francis Group. doi:10.4324/9781003084471.

Miner, Adam S., Nigam Shah, Kim D. Bullock, Bruce A. Arnow, Jeremy Bailenson, and Jeff Hancock. 2019. "Key Considerations for Incorporating Conversational AI in Psychotherapy." Frontiers in Psychiatry; Front Psychiatry 10: 746. doi:10.3389/fpsyt.2019.00746.

Ministeriet for Sundhed og Forebyggelse. 2008. Almen Praksis' Rolle i Fremtidens Sundhedsvæsen. Slotsholmsgade 10-12, 1216 København K: Ministeriet for Sundhed og Forebyggelse.

Mittelstadt, Brent. 2021. The Impact of Artificial Intelligence on the Doctor-Patient Relationship: Council of Europe.

Nam, Kyoung Hyup, Da Young Kim, Dong Hwan Kim, Jung Hwan Lee, Jae Il Lee, Mi Jeong Kim, Joo Young Park, et al. 2022. "Conversational Artificial Intelligence for Spinal Pain Questionnaire: Validation and User Satisfaction." Neurospine 19 (2): 348-356. doi:10.14245/ns.2143080.540.

Palanica, Adam, Peter Flaschner, Anirudh Thommandram, Michael Li, and Yan Fossat. 2019. "Physicians' Perceptions of Chatbots in Health Care: Cross-Sectional Web-Based Survey." Journal of Medical Internet Research; J Med Internet Res 21 (4): e12887. doi:10.2196/12887.

Parviainen, Jaana and Juho Rantala. 2022. "Chatbot Breakthrough in the 2020s? an Ethical Reflection on the Trend of Automated Consultations in Health Care." Medicine, Health Care, and Philosophy; Med Health Care Philos 25 (1): 61-71. doi:10.1007/s11019-021-10049-w.

Powell, John. 2019. "Trust Me, i'M a Chatbot: How Artificial Intelligence in Health Care Fails the Turing Test." Journal of Medical Internet Research; J Med Internet Res 21 (10): e16222. doi:10.2196/16222.

Praktiserende Lægers Organisation. 2022. PLO Faktaark 2022 .

Preece, Jennifer, Yvonne Rogers, and Helen Sharp. 2019. Interaction Design, 5th Edition Wiley.

Schario, Mark E., Carol A. Bahner, Theresa V. Widenhofer, Joan I. Rajaballey, and Esther J. Thatcher. 2022. "Chatbot-Assisted Care Management." Professional Case Management 27 (1): 19-25. doi:10.1097/NCM.0000000000000504.

Schön, Donald A. 1992. "From Technical Rationality to Reflection-in- Action." In , 31-80: Routledge. doi:10.4324/9781315237473-8.