HEALTH DATA SHARING FOR CLINICAL RESEARCH

Designing a patient-centric approach

ZAHRA GHASIA

MASTER THESIS

ACKNOWLEDGEMENTS

NEGOTIATE YOUR PARTICIPATION IN CLINICAL RESEARCH

Your data preferences :

Blood test reports

Monthly
Quarterly
Yearly

PROMOTE TRANSPARENT RESEARCH PRACTICES

PATIENT-CENTRICITY
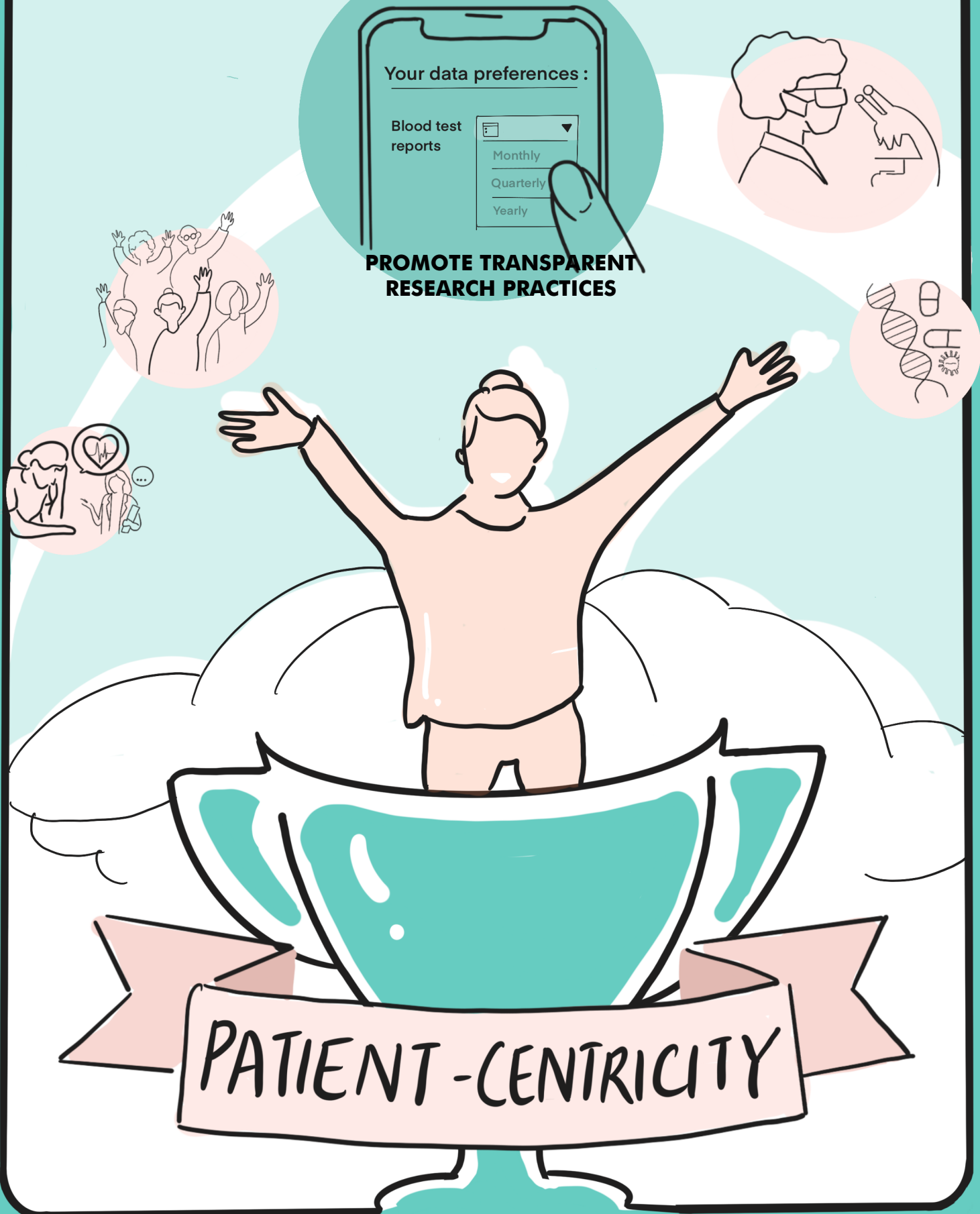
EXECUTIVE SUMMARY

The use of Electronic Health Records (EHRs) has seen a breakthrough in clinical research for personalized treatments (Hamburg & Collins, 2010.) Despite the potential advantages of vast EHR data available, constraints of privacy and legislation hinder its use (Rieke et al., 2020.) Health data exists in an interconnected healthcare system (Harris et al., 2009), comprising of stakeholders responsible for health data management within the constraints of GDPR. Data sharing platforms, through the use of secure data sharing practices and encrypted technology, can potentially change the landscape of health data in clinical research (Obermeyer & Emanuel, 2016.)

This project examines the privacy and stakeholder landscape of health data sharing through an evaluation of interaction with patients. It uses a Value-Sensitive design approach (Friedman et al., 2002) to contextualize the privacy values in clinical research for patients, doctors, patient organizations, clinical researchers, pharmaceutical industries and data sharing platforms. Through this examination, it identifies lack of transparent data use and research practices as a hindrance to the use of health data on a data sharing platform. Transparency is examined through a patient-centric lens, wherein information and control over preferences of participation are found to be crucial. This evaluation further identifies roles of researchers and data sharing platforms for a transparent approach.

The project results in a concept termed 'Negotiated Consent', which examines the use of data sharing platforms in offering individuals transparency. This is contextualized within patient consent for recruitment in clinical research, wherein data sharing platforms are responsible for data processing through the use of Federated Learning and Natural Language Processing. The stakeholder landscape for the same is defined in 3 functions for the platform- Access to Health Data, Patient Consent and Recruitment, and Data-Driven Insights. The results are tested through a study of the prototyped user experience of 'Negotiated Consent', wherein participants emphasized 'feeling more informed and in control' in comparison to the current informed consent.

Whilst the study examines a patient-centric approach towards health data sharing for clinical research, it has limitations in addressing the multi-faceted reality of patient's lives that informs their choices to enrol in clinical research. The further development of patient-centricity in the domain requires examination of motivations through a lens of disease-specific patient groups, demographics, and personal history with diseases (Hong et al., 2020) will lead to richer insights. Furthermore, Negotiated Consent is a proof-of-concept for dynamic consent (Mascalzoni et al., 2022), that demonstrates revision over participation choices within one moment of patient consent.

This project contributes to an understanding of the privacy and stakeholder dynamics for health data sharing. It illuminates a transparent approach to recruitment for clinical research through a collaborative effort that emphasizes on patient-centric approaches. Using a mutually responsible approach, data sharing platforms can lead the way for use of data in clinical research that simultaneously empowers patients to control their data through a negotiated lens.

READING GUIDE

This guide aids readers in comprehending the structure and content of the thesis. It provides visual cues, highlights the main topics in the project by chapter, and offers abbreviation explanations used in the report. Its aim is to facilitate swift and efficient access to desired content for both casual and focused readers.

VISUAL GUIDE

PROJECT RESEARCH
Within this box, you'll find text indicating research-oriented questions and summaries integrated throughout the project.

*RESEARCH QUESTIONS
AND SUMMARIES*
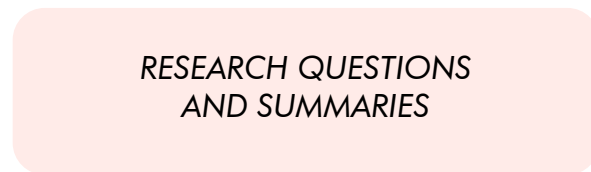
PROJECT HIGHLIGHTS
This box encompasses noteworthy points that occur throughout the project's duration.

DATA SEGMENTS
The results are presented systematically within different sections of data for clarity and coherence.

QUOTE SUMMARIES
This section contains text that highlights and summarizes key quotes.
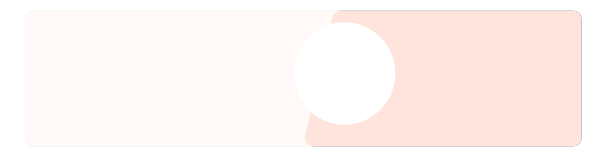
ICONS
This report relies on a significant use of icons. They represent the values and stakeholders in the thesis, and will be introduced systematically within the project.
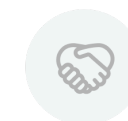
VALUES　　STAKEHOLDERS

GLOSSARY

EHRs	Electronic Health Records
AI/ML	Artificial Intelligence/Machine Learning
FL	Federated Learning
NLP	Natural Language Processing
GDPR	General Data Protection Regulation
DGA	Data Governance Act
EHDS	European Health Data Space

TABLE OF CONTENTS

1. INTRODUCTION

Understanding the assignment

This chapter introduces the initial assignment of this graduation project. In addition, it elaborates on the project approach. First, it describes the context and introduces the client in collaboration. Subsequently, it defines the problem and assignment of the graduation project by formulating overall research questions that serve as starting points for a literature review. Second, the project approach is described and visualised. It addresses the methods used throughout the research and design phases, and the design perspectives undertaken.

1.1 CONTEXT

Health data, comprising essential information about an individual's well-being, has evolved significantly with the transition from paper-based to electronic records. This transformation, driven by technological advancements, has not only made medical services more efficient but has also paved the way for breakthroughs in medical research (Daniels et al., 2021; Lorkowski & Pokorski, 2022).

EVOLUTION OF HEALTH DATA
The shift from traditional patient records to Electronic Health Records (EHRs) has enhanced the management, flexibility, and cost-effectiveness of medical services, ultimately benefiting patients (Lorkowski & Pokorski, 2022). EHRs, offering real-time, patient-centered information, extend beyond standard clinical data and include a comprehensive view of a patient's care (What Is an Electronic Health Record (EHR)? | HealthIT.gov, n.d.).

HEALTH DATA FOR CLINICAL RESEARCH
The availability of health data through EHRs, coupled with technological advancements in health information exchange and artificial intelligence, has catalyzed a paradigm shift in medical research (Daniels et al., 2021). The integration of EHR data with genomics has significantly contributed to understanding and addressing the root causes of diseases (What Is an Electronic Health Record (EHR)? | HealthIT.gov, n.d.). Notable breakthroughs in personalized medicine, reliant on vast datasets, have reshaped the landscape of medical inquiry (Hamburg & Collins, 2010). For instance, a diagnostic-focused study revealed insights into breast tumours, aiding in identifying patients benefiting from targeted therapies like 'trastuzumab' (Scholl et al., 2001).

INTERCONNECTED HEALTHCARE SYSTEM
The progress in technology has not only transformed health data utilization but has also interconnected the healthcare ecosystem. Data sharing technologies facilitate seamless exchange among doctors, researchers, and patients, aligning with responsible and ethical practices promoted by regulatory bodies (Davenport & Kalakota, 2019; Harris et al., 2009).

The synergy between technological progress and health data sharing is revolutionizing healthcare and medical research. As we delve deeper into this interconnected era, embracing responsible data-sharing practices becomes paramount for harnessing the full potential of health data in advancing medical knowledge and patient care.

1.2 ABOUT THE CLIENT

OASYS NOW is dedicated to transforming the landscape of clinical research by establishing a health data sharing platform that directly connects individuals to ongoing research studies. Recognizing the challenges in patient recruitment faced by current clinical research practices, the platform adopts a patient-facing approach, aiming to alleviate concerns related to drop-outs and low recruitment rates, ultimately minimizing the undesirable outcomes that contribute to delayed time-to-market for clinical solutions (mdGroup, 2023; Chaudhari et al., 2020).

ENHANCING PATIENT RECRUITMENT
The primary goal of OASYS NOW is to address the persistent issue of patient recruitment by introducing a privacy-preserving, patient-centred approach to data sharing. By linking Electronic Health Records (EHRs) generated by doctors to clinical researchers, the platform ensures that individuals, through their consent, become vital contributors to relevant research studies (see Figure 1.) Simultaneously, individuals gain access to their own health data via the platform, creating a two-fold benefit. Furthermore, OASYS NOW strives to bridge the gap between doctors seeking diverse treatment options for their patients and patient organizations advocating for increased participation in clinical research.

IMPACTING CLINICAL RESEARCH
In addition to its patient-centric approach, OASYS NOW seeks to provide substantial value to clinical researchers. The platform employs privacy-preserving technologies, including Artificial Intelligence, for data analysis. This facilitates predictive analysis by extracting meaningful insights from Electronic Health Records (EHRs), enhancing the efficiency and effectiveness of clinical research practices (Yang, 2022).

OASYS NOW's commitment to patient-centric health data sharing not only addresses the challenges in patient recruitment but also contributes to the advancement of clinical research. By leveraging innovative technologies, the platform ensures that individuals, doctors, and researchers collaborate seamlessly, fostering a transformative paradigm in the realm of healthcare and medical research.
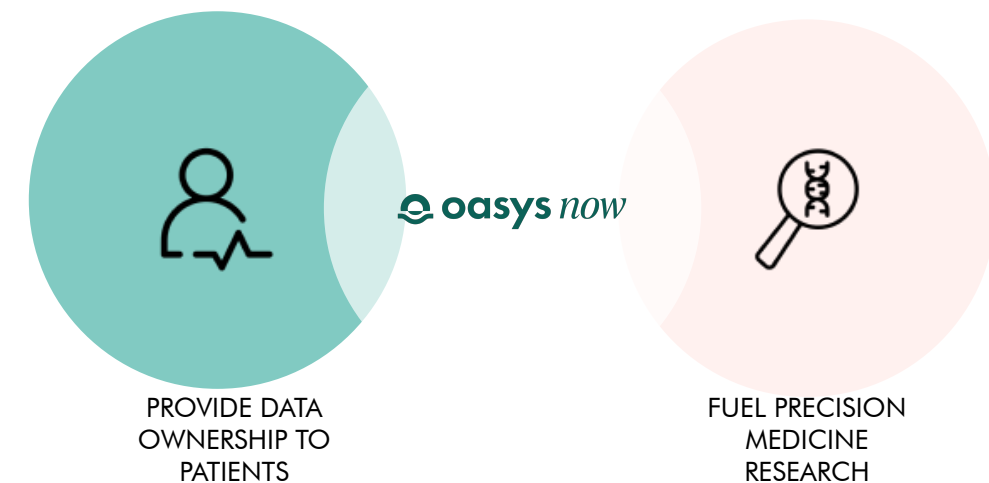

PROVIDE DATA OWNERSHIP TO PATIENTS

FUEL PRECISION MEDICINE RESEARCH

Figure 1 : The aim of company OASYS NOW B.V.

1.3 PROBLEM

The potential advantages of health data sharing are vast, yet it is accompanied by significant privacy concerns that demand careful consideration. This includes risks such as unauthorized access, data breaches, and the potential re-identification of individuals even after attempts at anonymization (Sweeney, 2002). Managing health data, especially in the era of big data, poses unique challenges, with encryption limitations and the associated risk of identifying individuals and their communities (Bonomi et al., 2020). This intricate landscape becomes more complex with the rise of sophisticated cyber threats.

STAKEHOLDERS IN HEALTH DATA SHARING
In the context of health data sharing for research, stakeholders, including doctors, researchers, and patients, grapple with a myriad of responsibilities. These include ensuring secure data transmission, obtaining informed consent, and protecting patient confidentiality. Doctors, for example, play a crucial role in obtaining and validating patient consent, while researchers must prioritize data integrity and transparency throughout the research process (COREON Foundation, 2023).

INSIGHT GENERATION FROM EHRs
The comprehensive nature of Electronic Health Records (EHRs) provides a holistic view of patients' health journeys. Leveraging EHRs enriches genomic findings with vital contextual information, effectively bridging the gap between research and clinical practices.

TECHNOLOGY AND REGULATION
The dynamic interplay between technological advancements and regulatory frameworks significantly shapes the landscape of health data sharing for clinical research. Privacy-preserving Artificial Intelligence methods hold promise for secure and transparent data sharing. However, this necessitates an evolving legal framework to align with European data protection laws such as the General Data Protection Regulation (GDPR) (Agbo et al., 2019). While current clinical research practices involve consent for data involvement, GDPR places decision responsibility on data controllers, leading to varied awareness levels among controllers (Van Ooijen & Vrabec, 2018). In this evolving scenario, data sharing platforms emerge as intermediaries with the potential to empower individuals in unique ways, providing informed control over their participation in clinical research.

The landscape of health data sharing is marked by intricate challenges and promising opportunities. Navigating privacy concerns requires a collaborative effort from all stakeholders, supported by evolving technologies and a robust regulatory framework. As data sharing platforms play a pivotal role, the focus shifts to empowering individuals with informed control over their data, fostering a responsible and transparent approach to health data sharing in clinical research.

1.4 ASSIGNMENT

This thesis centres on the intricate interplay of factors that shape the landscape of health data sharing, with EHR data that consists of genomic health data. This study endeavours to discern the dynamic roles of diverse stakeholders in this arena, including patients, doctors, patient organizations, data sharing platforms, clinical researchers, and the pharmaceutical industry. By illuminating their respective interests, concerns, and contributions, we aim to create a comprehensive understanding of the values underpinning collaborative endeavours in health data sharing for medical research.

Moreover, this thesis places patient centricity at its core. Recognizing the pivotal role patients play as the primary source of health data and the ultimate beneficiaries of advancements in healthcare, the study seeks to advocate for approaches that empower and prioritize patients in the data-sharing ecosystem. By championing patient rights, informed consent, and data ownership, we aim to cultivate a culture of trust and transparency among all stakeholders involved.

Furthermore, the research delves into the realm of technology and regulation implementation possibilities. In an era where technological innovation is reshaping healthcare practices, we endeavour to explore the potential of solutions for secure data sharing, and AI-driven analytics for deriving meaningful insights from genomic data for the purposes of clinical research (Obermeyer & Emanuel, 2016.) Simultaneously, the thesis examines the regulatory frameworks that govern health data sharing, considering their capacity to balance innovation with the imperative of safeguarding privacy and ethical integrity (COREON Foundation, 2023.)

To further understand the context of this thesis and identify potential problem areas for design, the following overall research questions are defined:

What are the privacy dynamics and challenges within the European health data sharing ecosystem for clinical research, considering the diverse stakeholders (patients, doctors, clinical researchers in the industry, and data sharing platforms), technological advancements (ML and encryption), and regulatory frameworks (GDPR, DGA and EHDS)?

How can a patient-centric approach be implemented within a collaboration among stakeholders in the health data-sharing ecosystem, leveraging technological capabilities to meet privacy needs and transparency whilst upholding diverse collaboration needs?

1.5 PROJECT APPROACH

DESIGN METHODOLOGY

The overall project was supported through the use of the Double Diamond model (Design Council, 2019), which illustrates the design process as a process of exploration and definition. The use of design tools and methods are herein defined at each stage of the Double Diamond, and illustrated in Figure 2.

DISCOVER
During the Discover phase, the aim was to better understand the context. After having defined the project assignment in the Project Brief, a divergent exploration of the data sharing platforms, stakeholders involved and context of use of health data sharing was explored (covering Chapter 2 and 3.) Moreover, presentations regarding the context were attended, both virtual and in-person.

DEFINE
During the Define phase, insight generation of information collected was the main highlight. The insights were generated at multiple stages of Literature Review and Stakeholder Interview using the research data collected. Final insights from Chapter 2 and 3 lead to discovering the scope. The same is done under Chapter 4, which led to a conclusion of Problem Definition.

DEVELOP
The problem was framed under three verticals of patient-centricity, collaboration and technology for transparency, which were developed further under Chapter 5. This lead to the Solution designed in Chapter 6, offered in three levels of systemic, technical and user-facing solutions.
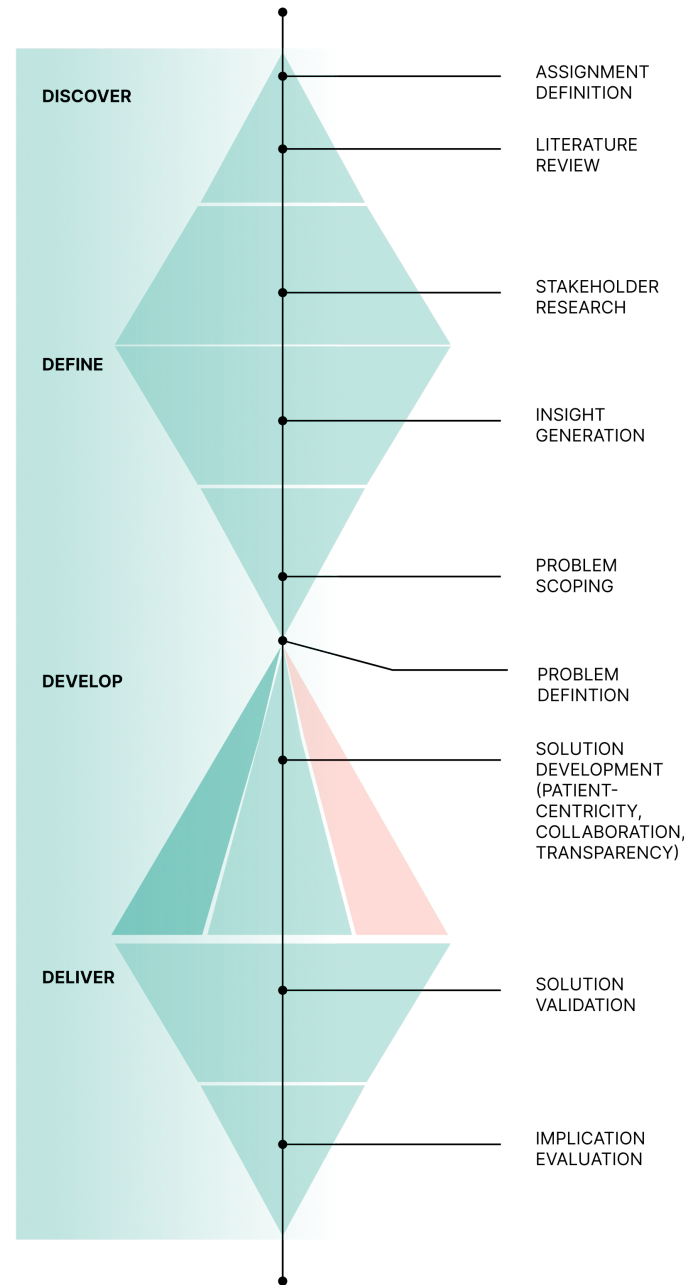
Figure 2 : Implementation of Double Diamond in project

DELIVER
The Solution presented in Chapter 6 was developed in prototypes for validation of concepts. The Health System Framework and Data Journey were iterated and validated at several stages throughout development through validation with the client. However, validation with stakeholders and patients for the implementation of Negotiated Concept was carried out in Chapter 7. As a final result, Chapter 8 explores the implementation of the Solution and provides a conclusion for the overall design process and an answer to the research questions defined.

PERSPECTIVES IN THE PROJECT

The following approaches were taken during the course of the project to ensure a human-centered and value-centered approach towards designing for health data sharing-

HUMAN-CENTERED DESIGN
Human-Centred Design is based on the use of design research techniques to obtain understanding of human needs, desires and experiences. Furthermore, it focuses on the questions,insights and activities from the people for whom the design is intended (Giacomin, 2015). This thesis showcases HCD through its integration of the various stakeholders and end-users in the research and design phases, along with an ongoing focus on understanding their experiences, tasks, interactions, and overall context.

PARTICIPATORY DESIGN THROUGH CO-DESIGN
Embracing participatory thinking is gradual, and is necessary to develop collaborative ideas that facilitate change in attitudes (Sanders & Stappers, 2008.) For the same, a co-design approach was incorporated during the project to challenges conventional beliefs, particularly in business, where a prevailing expert mindset resists relinquishing control.

VALUE-SENSITIVE DESIGN
Value Sensitive Design is an approach that originates from the field of information technology and human-computer interaction. It focuses on the design of socio-technical systems from a foundation of human values from the stakeholders involved, that are then considered throughout the design process (Friedman et al., 2002.) VSD is reflected in this thesis through the identified values of different stakeholders which becomes the foundation of definition of problem statement.

2. LITERATURE RESEARCH

Understanding the ecosystem of health data sharing

This chapter initiates the literature review for the project's focus. Firstly, it introduces sub-research questions based on the overall research questions. Subsequently it introduces two perspectives of examination of health data sharing. As part of the literature review, it explores the stakeholders in collaboration for health data sharing in the context of patient journey, and examines patient decision to share their data. Subsequently, technology of data sharing platforms is reviewed through examples and exploration of use and ethics of Machine Learning. Furthermore, it examines the context of health data sharing in the European environment through the lens of GDPR and other regulatory initiatives. Lastly, it concludes with key takeaways in light of the project focus from the literature research conducted.

Figure 3 : Health RI conference, 2023

2.1 INTRODUCTION

This section of the literature review is dedicated to a comprehensive exploration of the ecosystem of health data sharing. It delves into the an understanding of the stakeholders involved in the ecosystem, along with their association to patients and their health data. Furthermore, it assesses the use of machine learning technology used in health data sharing whilst defining the associated regulatory norms in place. The significance of an individual's outlook on health data sharing is also discussed, shedding light on the decision-making behind the motivation to share health data. The literature review to be carried out attempts to answer the following research question :

What are the privacy dynamics and challenges within the European health data sharing ecosystem for clinical research, considering the diverse stakeholders, technological advancements, and regulatory frameworks?

For the same, following sub-research questions were developed, that assist in examination of the scope through the different aspects of the research question.

What role does collaboration among stakeholders (patients, doctors, researchers, institutions, data sharing platforms) play in health data sharing for clinical research?

What are the key factors influencing individual's decisions to share their data ?

What technological advancements and infrastructure are needed to support health data sharing, whilst ensuring compliance with evolving regulatory standards (such as GDPR)?

These sub-research questions are included in the literature review to be followed, and will be highlighted at relevant moments.

TRANSPARENCY   TRUST

MUTUAL
RESPONSIBILITY
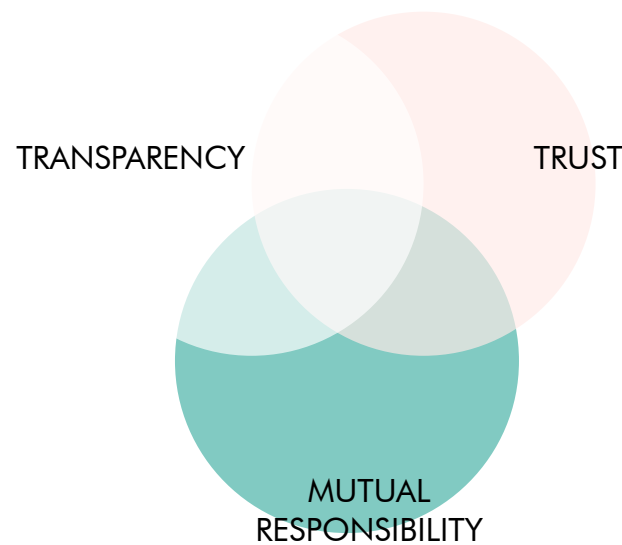
Figure 4 : Perspective on Health Data Sharing

2.2 PERSPECTIVES ON HEALTH DATA SHARING

Health data sharing can have many associated meanings, however within the clinical research context, they are associated with data shared amongst many stakeholders and data shared to make value out of it. For the same, two perspectives are herein identified- data sharing as collaboration and transaction.

DATA SHARING AS COLLABORATION

As examined, health data sharing is within a network of several stakeholders. Collaboration herein implies balancing the diverse interests among stakeholders.

SOCIAL CONTRACT
The OECD (Organization for Economic Co-operation and Development) underscores collaboration's role in sustaining data sharing platforms for research (OECD, 2021) and defines a term 'social contract' for the same. Under this term, a 'contract' is a means of stakeholders acknowledging the mutual need for reciprocal obligations whilst seek individual gains. The 'contract' herein identifies a mutual responsibility wherein transparency and trust go hand-in-hand. It was further identified by Hoepman, 2018 that communication, as part of ensuring clarity and precision in terminologies, process and information about the clinical research is fundamental to fostering collaboration.

An examination of health data sharing as collaborative 'social contracts' helps to examine the necessity of mutual relationships and transparency in communication for building trust in the same.

DATA SHARING AS TRANSACTION
In the domain of FinTech, value is an outcome of data shared under transaction. To bring in the perspectives offered by the same, Innopay's 'Everything Transaction' (Innopay, 2022) discusses the nature of agreements among stakeholders in intentional data exchange transactions. A brief summary of its application for data sharing platforms is derived herein.

TRANSACTIONAL ENVIRONMENT
Considering the context of health data sharing, these transactions involve data exchange, where one stakeholder provides data in exchange for a service or insight with another, facilitated by the data sharing platform. The platform serves as the intermediary, aligning with conventional transactional value exchange. Each stakeholder bears responsibilities within this environment (including patients, doctors, clinical researchers, and data sharing platforms), ensuring the transaction's desired outcome through the intermediary's guidance.

The role of the data sharing platform as an intermediary is to facilitate agreements between different stakeholders. By facilitating agreements, ensuring adherence to regulations and mitigating the risks related to data quality, the platform becomes crucial to sustaining trust and credibility within this interaction (Adler-Milstein et al., 2017.)

While there are other perspectives of examination of health data sharing (such as data donation, crowd-sourcing, etc.), examining it through the given lens helps us identify and frame health data sharing as a mutual responsibility, supported by transparency and trust as illustrated in Figure 4.

2.3 STAKEHOLDER COLLABORATION

Health data sharing is an act of collaboration between patients, doctors, clinical researchers in the industry and data sharing platform. This section explores the same by using the following sub-question as guide:

What role does collaboration among stakeholders (patients, doctors, researchers, institutions, data sharing platforms) play in health data sharing for clinical research?

PATIENT'S JOURNEY

Individuals who suffer from rare genetic diseases undergo a series of interventions during their patient journey. At present the average time from disease onset to accurate diagnosis for a rare disease is 4.8 years (RARE Disease Facts - Global Genes, 2024; Blöß et al., 2017). Part of their interventions have much to do with the many interactions they make with stakeholders in the healthcare sector, such as with doctors, specialists of conditions, patient organizations and testing

centres (illustrated in Figure 6), during which health data is generated and shared across institutions (Business.gov.nl, 2023b.)

Acknowledging the need to examine the role of health data in a patient's journey, comes with the acceptance of how patient journeys may look different for each individual. It is then necessary to examine the relationships with other stakeholders, whilst staying open-minded.

Figure 6 : Patient Journey individuals with rare genetic diseases
(Solve-RD Infographic on the Patient Journey to Diagnosis -
EURORDIS, 2023b)

AT THE DOCTOR'S CLINIC

The time spent at a doctor's visit is crucial in fostering trust in a patient with the healthcare sector (Adeleke et al., 2011.) Demirsoy et al. (2021) highlight privacy as safeguarding personal boundaries, which allows patients to share openly during care-seeking, emphasizing the significance of confidentiality in patient-doctor interactions. This includes being informed of the patient's authority over their own data (Ballantyne, 2020.) In the Netherlands, where patient data is shared among doctors, educating patients about data exchange details and their rights to modify or erase data signifies the patient's central role in governing their clinical information (Business.gov.nl, 2023b). As part of

the informed consent process, the doctor plays an active role in decision-making with the patient and by informing a patient about their data rights, risks and benefits involved (Shah, 2023.)

Patient control over health information is pivotal in establishing trust, despite healthcare institutions holding health records. Viewing clinical data as co-created information shaped by both patients and doctors (Ballantyne, 2020) underlines patients' autonomy in determining data use and sharing practices. Through this lens, patient data shared among doctors can become an opportunity to include patients in the process.

PATIENT   **PATIENT ORGANIZATION**   **DOCTOR**

DATA SHARING PLATFORM   **CLINICAL RESEARCHER**   **PHARMACEUTICAL INDUSTRY**

Figure 5 : Stakeholders in Health Data Sharing

2.4 DATA SHARING PLATFORMS

PATIENT ORGANIZATIONS

Patient organizations contribute significantly to patient well-being. Organizations such as Patientenfederatie and PatientsLikeMe have platforms that offer insights on health conditions, treatment options, and create virtual communities for knowledge exchange (Patiëntenfederatie Nederland, n.d.; PatientsLikeMe, n.d.) Their virtual spaces enable patients to be informed and engaged, empowering them with control over their patient journeys.

Organizations such as these are also pioneers in ongoing clinical research, and they contribute by influencing research priorities and advocating for transparent data sharing (Koay & Sharp, 2013.) Their active participation in establishing research agendas and engaging patients in trials often leads to a commitment to patient-centric research practices through transparent and collaborative data sharing practices.

RESPONSIBILITY OF CLINICAL RESEARCHERS

Clinical researchers' interest in collaboration with patients lies in building disease-specific treatments for rare genetic diseases. Research of rare diseases and clinical trials are necessary to develop a foundation for discovering targeted medicines (Dharssi et al., 2017). Moreover, the relationship between patients and clinical researchers pivots on trust and patient engagement within research endeavours (Lemke et al., 2010) such as the process of informed consent.

Researchers are responsible to provide participants any new information that becomes known during the course of the study, whilst offering the participant the 'right to withdraw' their participation (Gupta, 2013.) Herein, an emphasis for tailoring privacy measures to individual preferences becomes crucial in empowering patients with control over their participation in clinical research studies.

In discussing health data sharing platforms, it's crucial to prioritize patients and their centrality in this landscape. Amidst technical and regulatory challenges in managing health data across institutions, it's essential not to overlook individuals providing this data. This can be done by examining the factors that influence a patient's decision to share their data in the patient journey. Within this section, we'll be covering the exploration of the sub-question :

> *What are the key factors influencing individual's decisions to share their data ?*

VARIATION IN PRIVACY CHOICES

It was found by Hong et al. (2020) that factors like race, medical history, and prior participation in clinical research greatly impacts an individual's need for data control. Everyone's choices for privacy are different, and Gabrielli et al. (2022) emphasizes the significance of a positive user experience and clear privacy practices in maintaining individual trust and interest in a Data Sharing Platform. Individuals using such a platform are found to need transparency, understandability, control and convenience during sharing and other services provided by such platforms (MyData Global, 2017.) Moreover, under a study carried out by Bussone et al. (2020), participants were found to want to make fine-grained choices over what data was released to whom, for how long, and for what purpose.

Moreover, data sharing platforms may offer patients ongoing control over individual data sharing commitments, under the concept of dynamic consent (Mascalzoni et al., 2022.) Under this model, participants can revise choices over time, ensuring continuous communication and aligning with transparency, understanding, and control principles in health data sharing practice.

INFORMATION PROVISION

The platform's ability to disclose information regarding the clinical research plays a crucial role in determining and influencing decisions of individuals. As discussed earlier, doctors play an active role in the patient's decision-making through ensuring patients about their data rights. Similar to the doctor, a data sharing platform can play an active role in informing patients of their involvement in clinical research.

For a data sharing platform, MyData Global (2017) advocates for principles aligned with a patient-centric data management approach. Privacy maintaining strategies such as the 'Inform' strategy ensures individuals are adequately informed when their personal data is processed. (Hoepman, 2018) Serving as intermediaries, the platform can empower individuals to access, control, and share their data with informed consent and transparent practices.

This focus on awareness through usability, including user-friendly interfaces that allow individuals to control their participation may lead to empowering individuals with data rights, aiming at patient empowerment.

2.5 USE OF TECHNOLOGY

Technology of data sharing platforms is examined through the lens of value creation for health data sharing. In the context, health data to be shared is examined for the use of clinical research outcomes, wherein size and diversity of data available are relevant factors (Oh & Nadkarni, 2023.) Within this section, we will be examining the infrastructure necessary to generate value of health data using a data sharing platform, and covering the following sub-question -

> *What technological advancements and infrastructure are needed to support health data sharing, whilst ensuring compliance with evolving regulatory standards (such as GDPR)?*

EXAMPLES OF DATA SHARING INITIATIVES

Some examples of current use of technologies for sharing of health data in clinical research are examined to scope concerns with use of EHR data and privacy preserving technology.

1. USE OF EHR DATA
Example of Geisinger MyCode stands as a useful evaluation for EHR Utilization in Precision Medicine Research (Carey et al., 2016.) Under this initiative, individual EHRs are contributed through a community of willing participants, which includes genomic and health data. The collaborative effort forms a vast dataset, enabling researchers insights into the correlation between genetics and health outcomes.

Through their informed consent process, they outline data collection purposes and privacy measures, however it doesn't explicitly address the re-usability of individual data. This poses a noteworthy point for consideration, especially concerning differences in data protection laws between regions like the US and Europe. (European Parliament's Policy Department for Citizens' Rights and Constitutional Affairs, 2015), and with the European data sharing initiatives.

2. USE OF PRIVACY ENHANCING TECHNOLOGIES (PETs)
Example of Nebula Genomics helps to evaluate the use of Privacy-Ensuring Technologies for protection of genomic data (Nebula Genomics, 2018.) This platform uses blockchain technology to ensures secure, decentralized storage of health data, enhancing participants' trust and control over their data. Through their use of 'smart contracts,' individuals can offer precise definitions for data access and sharing permissions, empowering individuals to manage who can utilize their genetic information.

However, the example further states challenges in data sharing across different uses and users. Furthermore, the example demonstrates the necessity of maintaining data accuracy and quality across diverse sources.

USE OF FEDERATED LEARNING (FL)

The use of ML in clinical research can support statistical analysis performed by clinical researchers (Gates et al., 2021) and to match patients for participant recruitment (Beck et al., 2020.) Through the aforementioned examples, we understand that the use of EHR data proves useful in clinical research requiring vast datasets. In this section, we will further examine the use of EHR data with use of ML for clinical research.

Health data currently exists in silos, bound by privacy concerns and data governance laws(Rieke et al., 2020), and is the reason for lack of implementation of ML to its full potential. Moreover, individual EHR data is generated and stored at different locations during an individual's lifetime (Vos et al., 2020.) Hence, relying solely on data from a single institute may not suffice, and may require multi-centre studies for pooling data. Herein, machine learning (ML) technology is examined to analyse how data can be made use of within the concerns of privacy and dynamic nature of health data amongst different stakeholders.

Federated Learning (FL), as a form of ML, is a privacy-preserving approach to cater to the challenges of multi-centre studies for health data, wherein FL models do not require centralized data repositories (McMahan, 2017.) This way, challenges of data governance across different institutes is addressed (Rieke et al., 2020.) Using FL, each institute will be responsible for managing its computing and data warehouse and will share only ML models (Oh & Nadkarni, 2023b), which can be used by the data sharing platform to build global models.

PRIVACY-ENHANCING TECHNOLOGY
Using health data in clinical research leaves individual at a risk of re-identification. A study done by Prayitno et al. (2021), introduces the use of Differential Privacy and Homomorphic Encryption algorithms to maintain individual privacy in the FL environment and will be herein examined in context of the data sharing platform.

Differential privacy, as discussed by Prayitno et al. (2021) introduces controlled random noise in FL models to reduce risk of identification in case of leak. However, data sharing platforms that pool multi-centre data sets may be at a risk of errors in the data used in such a case (Gianfrancesco et al., 2018), leading to incomparable datasets across institutions.

Meanwhile, the same study examines the use of homomorphic encryption algorithm in upholding privacy. This is done by enabling the execution of analytical functions directly on encrypted data, hence yielding encrypted outputs identical to plain-text executions, without revealing underlying data (Prayitno et al., 2021; Munjal & Bhatia, 2022.)

Through an examination of use of FL in processing health data, it is understood that privacy-preserving technology such as Homomorphic encryption support the use of health data for clinical research. Moreover, in a FL environment, health data across diverse healthcare institutions can be further made use of whilst following the regulatory standards of each institutions.

2.6 REGULATIONS

USE OF NATURAL LANGUAGE PROCESSING (NLP)

Whilst we have examined health data through a privacy-preserving lens, it is important to understand the nature of health data available in EHRs. Current EHRs comprise of unstructured data sets (Li et al., 2022), which require processing in a ML environment. As discussed in S. M. Shah and Khan (2020), EHR data consists of structured and unstructured data sets, wherein Natural language processing (NLP) methods are useful for extraction of structured clinical data from unstructured data sets (Kreimeyer et al., 2017.)

Using NLP, patients can be matched based on the specific characteristics discovered through processing of their unstructured EHR data (Weissler et al., 2021b) Moreover, the same can be conducted across healthcare institutions using an analysis of enrolment criteria (text) of research with patient records (Zhang et al., 2020.) Due to a vast dataset hence available, this may lead to reducing bias in participation in research (Vassy et al., 2018; Weber et al., 2017.)

However, bias in the EHR data and ML algorithms could also affect the performance leading to questions regarding their reliable use, and need further examination for the same.

ETHICS IN USE OF AI

In the pursuit of leveraging advanced technologies like Federated Learning (FL) and Natural Language Processing (NLP) for health data sharing, it's crucial to acknowledge the ethical implications that may accompany their implementation.

ML algorithms that lack interpretability (Weissler et al., 2021), may lead the researchers to be unable to distinguish an accurate result of the algorithm from an inaccurate one. This may lead to a lack of trust in machine learning in processing of individual data for clinical research.

Based on SUM Values, the societal and ethical impacts of the use of AI in the context are examined (Leslie, 2019). With the use of AI in a health data sharing environment, we can imagine a societal push towards strengthening the relationship between stakeholders, whilst ensuring that individuals are able to safeguard their autonomy and agency over their data. Furthermore, by minimizing the risk of misuse, use of AI may lead towards advancing fairer practices of participant recruitment in research.

Naik et al., (2022) further examines the necessity of responsibility towards upholding transparency of process, integrity of model, and accountability of outcomes of machine learning in the use of these tools. While doing so may help mitigate potential biases, ensure data quality, and preserve patient privacy, the responsible and ethical application of technology in health data sharing will not only drive innovation but also safeguard patient interests.

Regulatory standards such as General Data Protection Regulation (GDPR) shape the nature of health data sharing. Under the same, sharing data with third parties, including other medical professionals, requires patient consent (Consent - General Data Protection Regulation (GDPR), 2021.) To further examine the implementation of technology for use of health data sharing, regulatory standards that support the same are examined. This section further supports the follows sub-research question-

What technological advancements and infrastructure are needed to support health data sharing, whilst ensuring compliance with evolving regulatory standards (such as GDPR)?

IMPLEMENTATION OF DATA GOVERNANCE ACT

GDPR remains a crucial enabler in healthcare, granting individuals greater control over their personal data (Wolford, 2023.) Within the European Union, the implementation of the Data Governance Act (DGA) has reshaped the health data sharing landscape, emphasizing principles of transparency, accountability, and consent (European Data Governance Act, 2023; Vuković et al., 2022). This Act facilitates responsible data sharing, allowing researchers to access and use health information while upholding regulatory standards and ensuring secure cross-border data exchange (Spalding et al., 2021). While the Dutch government ensures secure medical data handling, aligning with EU standards, the implication of DGA implementation requires re-evaluating the aspect of 'control' that can be offered through the lens of GDPR (Business.gov.nl, 2023; Filkins, 2016.) In a data sharing landscape,

GDPR is an enabler in ensuring user rights over their data, pre-existing laws regarding data privacy and data sharing, sharing anonymized statistics, developing new data analysis approaches, patients' trust towards dealing with their health data and transparency (Vuković et al., 2022.)

INTEGRATED HEALTH DATA SPACE

Moving towards an integrated data sharing environment requires collaboration. The European Health Data Space initiative promotes collaborative, standardized data sharing to advance public health outcomes (The European Health Data Space (EHDS), n.d.). Embracing Findable, Accessible, Interoperable, Reusable (FAIR) principles aids in standardizing shared data, ensuring its value and accessibility (Kush et al., 2020). The European Health Data Space initiative necessitates further demonstration of health data sharing practices that address concerns about individual data protection and privacy (De European Health Data Space Wankelt Op Onderdeel Bescherming - Zorgvisie, 2022; Vuković et al., 2022.)

However, GDPR's stringent regulations regarding identifiable and individual-level data, time needed to complete the process, workload increase, differences with local legal legislations, different (and stricter) interpretations and access to data pose barriers to its secondary use in Europe.

2.7 KEY TAKEAWAYS

LEVERAGING MUTUAL RESPONSIBILITY TOWARDS HEALTH DATA SHARING

The patient's interaction with stakeholders such as doctors, patient organizations, data sharing platforms and clinical researchers illuminate the opportunities for developing mutual responsibility. All stakeholders were found to be connected to patients through privacy for data sharing, which impacts the experience of trust in individuals.

It is found that discussions over privacy and control of health information in patient-doctor interactions are crucial in building trust, and supports sharing of data. Furthermore, patient organizations are found to be advocates for clinical research, given a transparent practice of use of data.

Clinical researcher's interaction with patient through the informed consent can impact trust in patient participation for research. Data sharing platforms offer convenience of use, and individuals were found to need a clear understanding of the services offered.

IMPACT OF SERVICE OFFERED BY DATA SHARING PLATFORMS ON INDIVIDUAL TRUST

As intermediaries of service, the experience of use of data sharing platforms was found to impact individual decisions. Information provided and ease of use of platforms was found to impact individual decisions.

Moreover, individuals have a varied need for privacy on platforms. As service providers, the ability of data sharing platforms to influence decision making through experience of use offer value in their role as intermediaries in health data sharing.

NEED FOR STANDARDIZED DATA IN A PRIVACY-PRESERVING ENVIRONMENT

It was found that health data use in clinical research is impacted by the difficulty of use through GDPR and unstructured nature of data. For the same, technology such as Federated Learning and Homomorphic Encryption, that supports use of health data in a privacy-preserving way is found to support the use of vast data sets across multiple institutions for statistical analysis in clinical research.

Moreover, through the use of NLP, health data can be operated in a standardized manner during participant recruitment in clinical research. Together, the use of FL and NLP supports the use of health data in initiatives such as EHDS.

Figure 7 : Interview with stakeholders

3. STAKEHOLDER RESEARCH

Examination of the people involved

This chapter addresses the context of health data sharing within stakeholders described in Chapter 2. It further recognizes a need for a more nuanced understanding of the responsibilities of stakeholders within the context, and describes the stakeholder research that further supports the project focus. It describes the research methodology, including a new methodology for data analysis. The results are discussed for the same, which lead to a generation of personas, stakeholder map and value-based conclusions. Lastly, it examines the insights from the project perspective and discusses key takeaways.

3.1 INTRODUCTION

Examination of the context under literature research led to an understanding of the impact of stakeholders in the health data sharing environment. However, it revealed gaps in understanding of the responsibilities of individual stakeholders in relation to patient, patient data and clinical research. This is further examined in the following section, wherein a qualitative research methodology is employed to support an understanding within the following over-all research question:

What are the privacy dynamics and challenges within the European health data sharing ecosystem for clinical research, considering the diverse stakeholders, technological advancements, and regulatory frameworks?

VALUE-SENSITIVE DESIGN

Given the technological focus of this research, acknowledging the human values associated with health data sharing services became crucial. A value-sensitive design approach, as outlined by Friedman et al. (2013), was adopted to understand the underlying values of each stakeholder.

The qualitative research herein carried out acknowledges the diverse experiences of stakeholders and examines their roles in patient interaction, challenges associated with current responsibilities, existing privacy practices concerning medical data, and collaborations with other stakeholders.

To address these aspects comprehensively, sub-research questions were formulated in alignment with the over-all research question:

What are the needs, constraints, and motivations of each individual stakeholder?

How do these factors influence their roles in relation to one another?

3.2 METHODOLOGY

For this study, I selected a qualitative research approach to explore the nuanced aspects of stakeholders' goals, needs, motivations, and behaviours. This choice was grounded in the necessity to understand the qualitative dimensions of the research questions (Marshall et al., 2006.) For the same, I conducted interviews with stakeholders.

PARTICIPANT RECRUITMENT
Purposeful sampling was utilized for participant recruitment, focusing on individuals with expertise within stakeholder categories. Selection criteria were based on the assumption that participants possessed knowledge and experience relevant to the study objectives, ensuring detailed and generalizable information (Palinkas et al., 2013, see Figure 8.) This approach allowed for results that provide novel insights into the context without making overly generalized statements about participant experiences (Patton, 2002).

The primary demographic criterion considered was the location of expertise practice, with most experts based in the Netherlands to maintain relevance to the European context in research questions.

Table 2: Interview timeline and overview

Participant ID	Stakeholder category	Understanding of context
P1	Patient	Individual tasks within patient journey through interaction with HCP and PO
P2	Patient Organization	Individual responsibility and current collaboration with HCPs and research
P3	Clinical researcher (NL)	Individual responsibility and current collaboration with PO, and industry
P4	Industry expert	Individual responsibility and current collaboration with researchers and PO
P5	Data sharing platform	Individual responsibility and (desired) collaboration with HCPs, POs, clinical researchers, and industry

Figure 9 : Interview Timeline and overview

I contacted the participants via mail or LinkedIn, and consent forms were exchanged and signed before the interviews. While participants had no prior personal contact with the researcher, P1, P2, and P3 had previously known each other. excluding P4. No incentives were offered for participation, but participants were given the option to receive research findings.

ETHICAL CONSIDERATIONS
Ethical considerations were acknowledged within the research process. Participants signed a consent form indicating their approval for participation, with a choice to be re-contacted for future research. The consent form was examined and approved by the Human Research Ethics Committee (HREC) at TU Delft.

PROCEDURE
I conducted interviews on platforms like Microsoft Teams, and transcripts were generated using the built-in transcript generator. The interview schedule spanned three weeks, strategically designed to gradually build an understanding of the context (see Figure 9) and accommodate expert availability.

Participant ID	Stakeholder categories	Expertise	Location	Duration	Type of interview
P1	Patient	Member of PO and PhD in genomics	NL	60 min	In-person
P2	Patient organization member	Member of PO and PhD in genomics	NL	60 min	Online
P3	Clinical researcher	Epidemiologist in genomic research and professor for governance practices in research	NL	60 min	Online
P4	Industry expert	Associate Director of Clinical Operations, Clinical Research Assistant	NL	60 min	Online
P5	Data sharing platform	CEO of data sharing platform, Computer Science Engineer	NL	30 min	In-person

Figure 8 : Participant Recruitment for Stakeholder Resaearch

EMERGENT SCOPE METHODOLOGY

An evolutionary approach within Grounded Theory Method

The "Emergent Scope Method," operated within the framework of the Grounded Theory Method, offered a new approach for addressing systemic research questions. This method was discovered and developed by the researcher during the course of the study. In this methodology, the gradual comprehension of context informed and broadened the scope of research, and hence offered a wider research scope towards its end.

SIGNIFICANCE

The Emergent Scope Methodology can be used in scenarios where stakeholders have diverse understandings of the context.
Example: Within the research, understanding the Patient Organization's perspective offered insights into the nature of collaboration between them and the Pharmaceutical Industry. Leveraging this understanding in formulating interview guide for Pharmaceutical Industry expert led to richer insights regarding participant recruitment in collaboration with Patient Organizations.

KEY COMPONENTS

INITIATING RESEARCH WITH CORE STAKEHOLDERS

The methodology advocates commencing research by delving into the patient journey and experiences. This strategic initiation assists in identifying stakeholder associations with the patient, forming a foundational understanding into collaborations with patients for all stakeholders.

PROGRESSIVE DATA ANALYSIS

The Emergent Scope Methodology conducted data analysis in two distinct phases. The first phase was the data analysed during the research, and involved pre-emptive generation of codes from interview transcripts before the research's conclusion. These were used to further guide the interview process during the research, fostering an iterative and evolving analytical process.

TAILORED INTERVIEW APPROACH

Interview guides were structured based on differing stakeholder perspectives owing to their varying roles and evolving contextual comprehension throughout interviews. Additionally, the interview guide construction was influenced by the principles of Value Sensitive Design (Friedman et al., 2013)

Figure 10 : Emergent Scope Method description

DATA ANALYSIS

The transcripts collected were analysed during the research. Grounded Theory Method was used to analyse the research results and construct or build theory from data (Glaser & Strauss, 1967.) Moreover, a gradual understanding approach of the context was carried out during the research, discovered by the researcher. Based within the research methodology of Grounded Theory Method, this methodology (herein termed as "Emergent Scope Method") is described in Figure 10.

The transcripts generated were coded in two phases-
Phase 1, termed 'Needs', wherein needs, pain points, motivations, goals and tasks of each stakeholder were identified.
Phase 2, termed 'Values', wherein values of each stakeholder were generated.

CODING IN PHASE 1 & 2

Phase 1: Needs
Carried out during the research, this phase laid the foundation for Phase 2. I generated insights using raw transcript data by identifying stakeholder pain points, needs, motivations, goals, tasks, and responsibilities. Figure 11 details the number of codes generated for each stakeholder.

Partici pant ID	Stakeholder categories	Pain points identified	Needs identified	Motivations identified	Goals identified	Tasks identified
P1	Patient	6	3	4	4	5
P2	Patient organization	4	3	5	5	9
P3	Clinical researcher	8	5	5	4	9
P4	Industry expert	3	5	6	6	4

Figure 11 : Participant Recruitment for Stakeholder Resaearch

Phase 2: Values
This phase involved a Value-Sensitive Design approach to generate values for each stakeholder. The code NEEDS developed for first iterative analysis were used as the basis of creating value-based synopsis. Values defined by Ethical Stack (The Ethical Stack, n.d.) were used as a reference for the same. However, the original values mentioned under Ethical Stack were modified to fit within the context of health data sharing for clinical research and stakeholders involved. Terms such as 'patient,' 'clinical research' were added to define the context within the values.

Eg: Well-being (as defined by Ethical Stack)
"Paying attention to the physical and mental welfare of the users and developers, designers and testers of the product."
Well-being (redefined within the context)
Paying attention to factors that influence the physical and mental welfare of the users.

The differences between the values used can be found in Appendix A. Statements made by participants were compiled under each value, and a synopsis was generated that described the effect of the statements on the identified values. Due to the lack of availability of doctors for the interview, values of doctors were defined through the literature research conducted earlier in the project.

The values defined during Phase 2 were further used to examine stakeholder collaborations through a stakeholder map, and value-based analysis of value dams, flows and tensions between stakeholders, and are discussed in the following section.

STAKEHOLDER VALUES

The following list of values were defined and identified for individual stakeholders with regards to their responsibilities and concerns to health data sharing for clinical research.
Values mentioned herein are classified as 'care-focused' for those that focus on the relational value of privacy, and 'use-focused' for those that focus on purpose value of privacy.

USE−FOCUSED

TRANSPARENCY
Striving towards achieving clarity throughout the process about the use of health data and how the implementation of the clinical research would look like.

DATA PROTECTION
Control over, access to and use of private health data. Making sure that patients (or individuals) are not adversely affected by the health data that is collected, processed, or analysed about them for research- both as individuals and as groups. Giving them the control to erase or alter their data, should they wish to do so as a form of consent.

SUSTAINABILITY
Considering and accounting for the long-lasting impacts of approach towards clinical research.

ACCOUNTABILITY
Assuming responsibility to explain why certain decisions (in the clinical context) are taken the way they have been, if or when potential risks are identified or when adverse consequences of a decision take place.

CARE−FOCUSED

PARTICIPATION
Encouraging collaboration with other stakeholders throughout the clinical research and ensuring that their involvement is deemed important.

DIGNITY
The feeling of control (or offering the other control) over one's own (or their) destiny that entails relationships of respect. Patients (or individuals) being able to have a say in how their own experiences should be, and exerting control within the environment and within relationships with other stakeholders that can adversely affect their own experiences.

WELL-BEING
Paying attention to factors that influence the physical and mental welfare of the patients.

RESPONSIBILITY
Assuming duty to take care, overseeing their own individual tasks undertaken to advance the decisions taken in the clinical context, and may affect the other stakeholders involved.

3.3 RESEARCH FINDINGS

PERSONAS

The identified needs, pains, responsibilities, motivations of stakeholders are represented through personas herein described. Furthermore, it illustrates the values of each stakeholder examined under Phase 2 of data analysis.

The following personas are representations of stakeholders in the context of health data sharing. For purposes of further examination, these are henceforth used to identify the stakeholders with.

PATIENT
Values- Well-being, Dignity, Transparency, Data Protection

"I value receiving appropriate care for my condition"

BIO
Sam is an individual with a rare genetic disease, who is highly pro-active about matters related to his health. He is a strong supporter of clinical research and believes in his ability to impact people around him.

GOALS
Being connected to organizations and research that matters to him.

LIKES
Transparent information regarding use of data

DISLIKES
Not being valued and respected

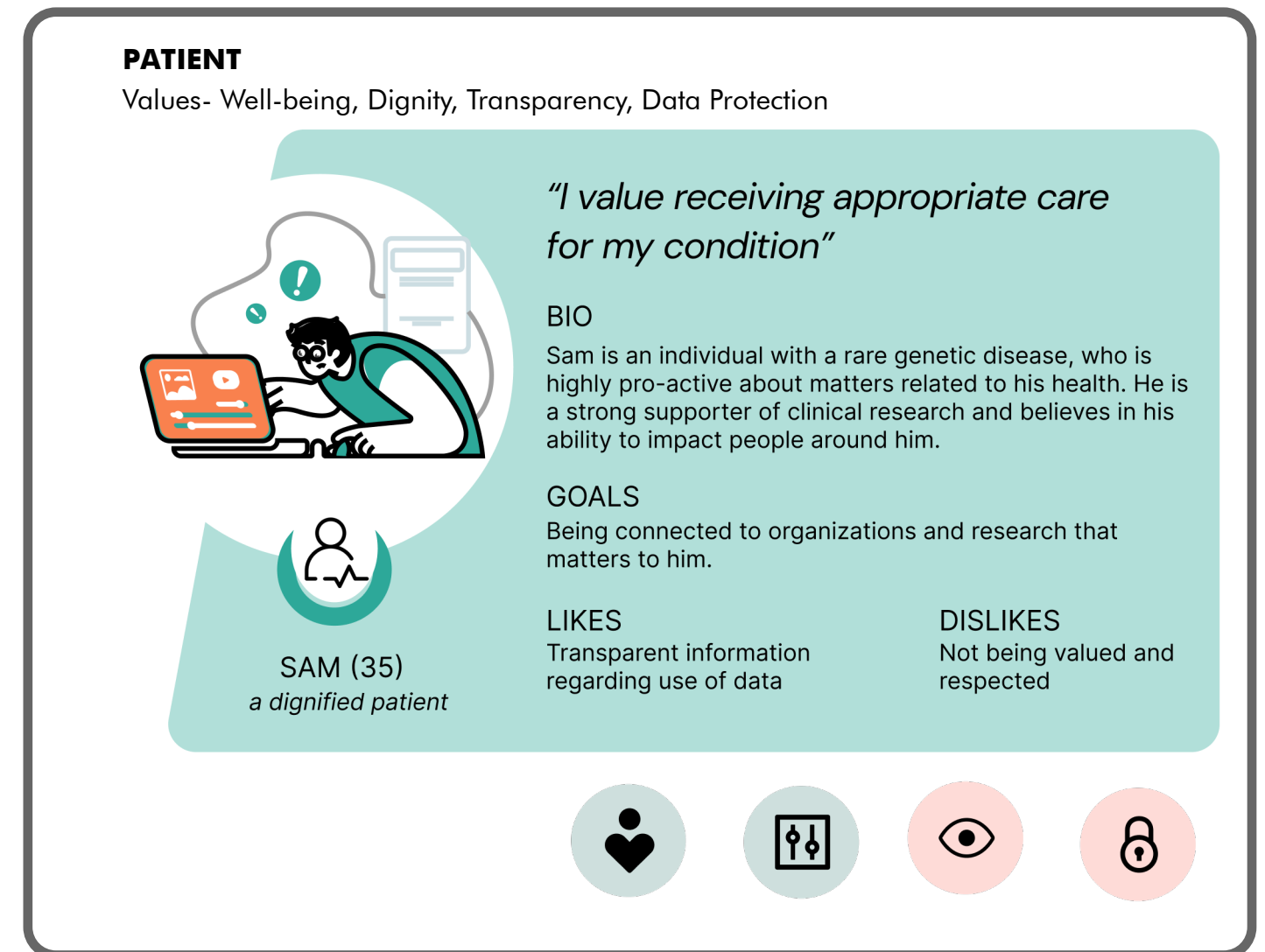**SAM (35)**
a dignified patient

Figure 12 : Personas of stakeholders

PATIENT ORGANIZATION
Values- Participation, Well-being, Accountability

"I want to advance research by involving patients in meaningful ways"

BIO
Ava is a retired patient organization member who studied and worked in the domain of her rare disease, and is now a volunteer at the patient organization that offers support and guidance to others.

GOALS
Using her professional experience to bridge the gap between research and patients

LIKES
Patient-focused care and support

DISLIKES
Non-transparent practices in research

AVA (57)
a friendly patient organization member

DATA SHARING PLATFORM
Values- Transparency, Data Protection, Accountability, Dignity, Participation

"I want to grant individuals control over their health data"

BIO
Yannes is the owner of the data sharing platform that wants to connect participants to clinical research. She believes in offering patients the ability to have control over their health data and hopes for her platform to offer the same.

GOALS
To connect patient data with clinical researchers

LIKES
Transparent practices in use of data

DISLIKES
Lack of agency in control over health data
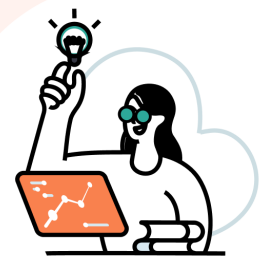
YANNES (25)
a responsible business owner

CLINICAL RESEARCHER
Values- Data Protection, Accountability, Sustainability, Participation, Dignity

"I want to ensure ethical practice of research"

BIO
Sergio is a clinical researcher who works mostly with anonymized patient data on observational studies. He ensures ethical practice by informing patients of implications of their participation, and using anonymized patient data.

GOALS
To have research alignment that is a priority for relevant patient groups

LIKES
Ethical consent practices

DISLIKES
Processing of data for use in research

SERGIO (30)
an accountable researcher

PHARMACEUTICAL INDUSTRY
Values- Transparency, Data Protection, Accountability, Sustainability, Participation, Well-being

"We work to develop solutions for patients that are effective"

BIO
Josie manages clinical operations in a pharmaceutical company and through her experience, believes that the public has a wrong image of the clinical practices of pharmaceuticals.

GOALS
Creating clinical outcomes that are reliable and without side effects

LIKES
Future use of patient data for other studies

DISLIKES
Offering patients agency over pharmaceutical practices

JOSIE (65)
a thorough pharmaceutical industrialist
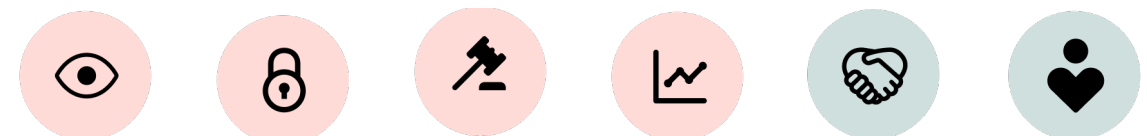
STAKEHOLDER VALUE MAP

Based on the identification of values of stakeholders, a stakeholder map was generated, which illustrates a relationship between different stakeholders through common values.

The value relationships herein identified were used to define and describe Value Dams, Value Flows and Value Tensions (Friedman et al, 2021), that help explore the relationship between different stakeholders based on their values.
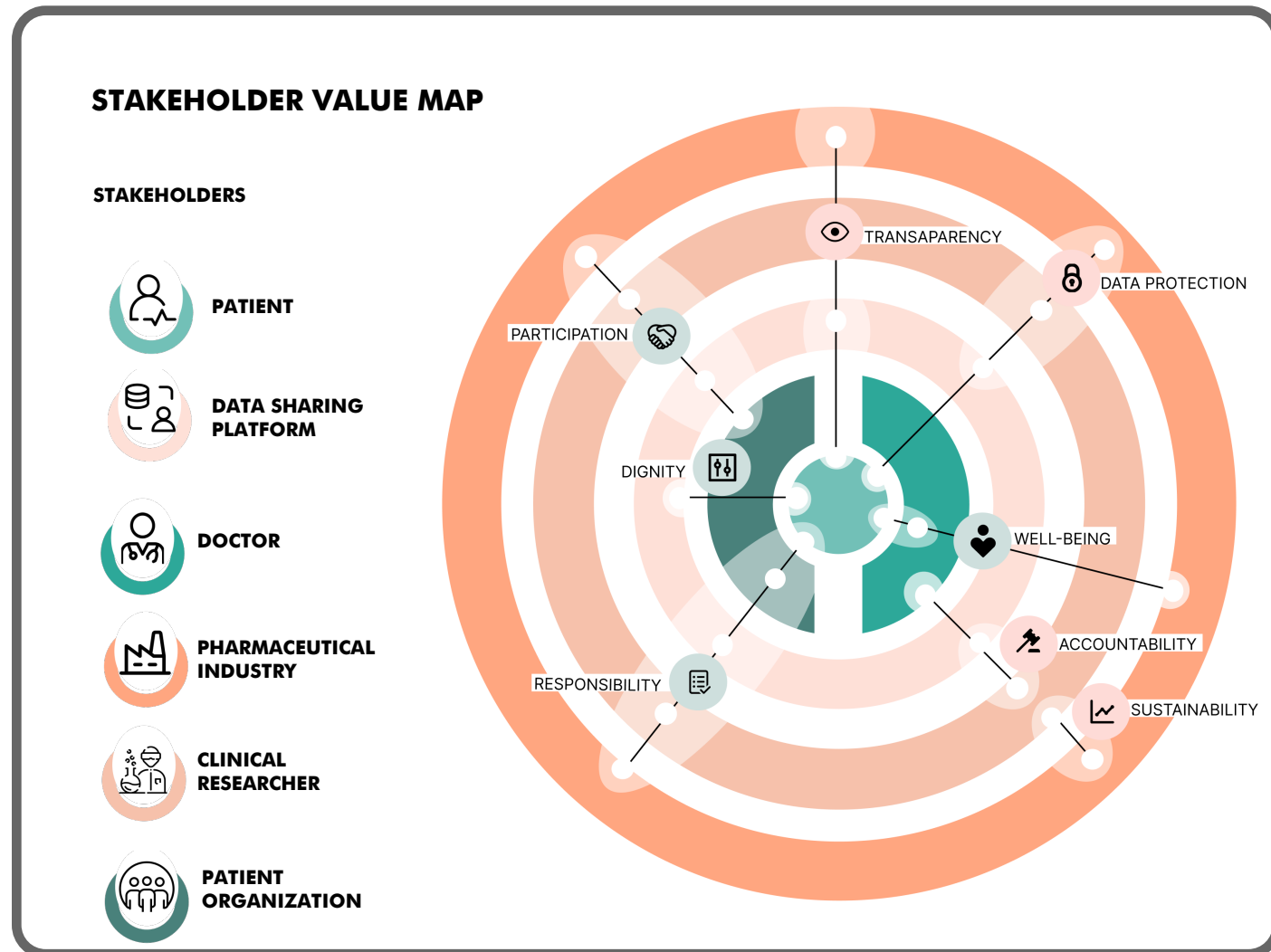
Figure 13 : Stakeholder value map

VALUE DAMS

Value dams, herein refer to absolutes that are strongly opposed by even a small set of stakeholders (Friedman et al, 2021.) These are considered as absolute no-go's as they may end up opposing the current context of health data sharing, if included. The following Value Dams were identified and described to fit the context:

"We do these data protection impact assessments before we start studies to understand what data might be at risk of reaching privacy concerns and how we can mitigate that"

DATA LEAK

Health data leaks arise from hacking platforms or technological challenges, constituting the majority of current breaches (Definitive Healthcare, n.d.). Stakeholders handling health data are concerned about unauthorized access and improper disposal.

Interviews revealed each stakeholder oversees data stewardship in health data management. Currently, they are not held accountable for the repercussions of data mismanagement by other stakeholders.

RESPONSIBILITY

"As middlemen, our job is to bring stakeholders in communication with each other, however it is the pharma's job to maintain their responsibilities towards the patients.."

"of course, you need to do everything according to the correct protocols.. there is so much regulation for clinical research. There is nothing there that clinical researchers can do. unethically. "

REGULATIONS

Stakeholders must adhere to ongoing EU regulatory practices within the GDPR framework, with their actions and responsibilities guided by legal practices governing health data sharing.

Data quality is assessed by its inter-operability, re-usability, and coherence. Inter-operability ensures seamless data integration from diverse sources, re-usability allows data for multiple purposes, and coherency facilitates research practices with minimal processing needs.

DATA QUALITY

"Because if you have the same image and you can use another algorithm or for it, then you can reuse the data, makes perfect sense"

Figure 14 : Value dams

VALUE FLOWS

Value flows (Friedman et al, 2021.), herein refer to value relations amongst stakeholders that are found to be in support of one another. The following Value Flows were identified and described to within the context:

COLLABORATION

Patient Organization, Clinical Researcher, Data Sharing Platform and Industry indicate an interest towards collaboration with other parties.

"So we can then give patients more the opportunity to be involved in research. Of course they will recruit from their own center but maybe people are wanting to be more involved"

Stakeholders express participation-based interests in interviews, fostering collaboration in clinical research. Collaboration terms vary; for example, the Patient Organization advocates for patient involvement, Clinical Research emphasizes understanding patient-relevant research, Industry values patient engagement in implementation but deems research a task for experts. The Data Sharing Platform supports these needs, connecting Patient Organizations with researchers. Collaboration is desired, but new definitions are needed for success.

AWARENESS

"what data are they going to use? Is it nice to know or need to know? also what's, where is the data stored? Who can access the data? is my data safe?"
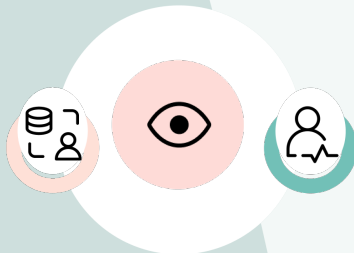
Patient and Data Sharing Platform denotes similar goals towards attaining transparency.

Patients express concerns about data use details, seeking more information. The Data Sharing Platform intends to address this by providing insights into data usage during clinical trial enrolment.

SUPPORT

Patient and Patient Organization interactions are fostered with care and support.

"So if you want to help with something, you can reach out to the info, email address, and we're going to see, okay, what can I do to help you? you can always email us with every question you have."

In the patient journey, recognition of medical conditions is essential. Patient Organizations serve as data hubs and advocates, fostering an environment where support is readily available.

AGENCY

"our goal is to start providing autonomy and agency over your health data for anybody who wants."
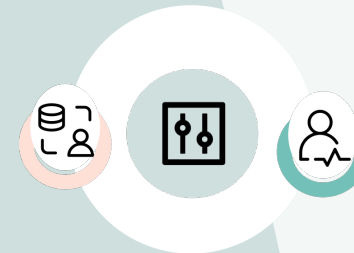
Patient and Data Sharing Platform exist in a dynamic of empowerment.

Patients express a desire for control in their journey, and the Data Sharing Platform supports this by providing control over health data through access and share controls.

BURDEN

Industry expert supports the cause of looking out for Patient's well-being.

"you can say, hey, instead of bringing on additional patients, we don't have to give them that burden and therefore we can reuse the data"

Patients seek research practices considering the strain of participation, while the Industry aims to reuse participant data for various studies, asserting awareness of participant strain.

QUALITY

"So as long as we don't expect clinical data from the patient, we can still locate it."

Data Sharing Platform agrees with the non-negotiable factor of data quality on the platform.

Recognizing the distinction between self-generated and clinically generated data, they aim to provide users with both. Emphasizing the value of clinically generated data, they advocate using it directly from healthcare providers for sharing in clinical research.

LEGALITY

Patient Organization, Data Sharing Platform, Clinical Researcher and Industry expert emphasizes on following the legal norms.

"also it's making sure that you are explicitly clear about everything (in the research) that will be done and will not be done and how the data will be used and not be used .."

Patient Organization ensures digital service recommendations avoid member risks, Clinical Research identifies and mitigates risks through informed consent, Industry adheres to clinical research laws, and the Data Sharing Platform strives to ensure user safety.
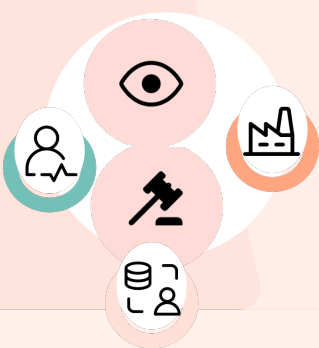
Figure 15 : Value flows

VALUE TENSIONS

Value tensions (Friedman et al, 2021.), herein refer to value relations that may concern the other stakeholder and cause disharmony. The following Value Tensions were identified and described within the context:

TRANSPARENCY

"We've done campaigns, had to be transparent and explain. They can never do it right. So in the beginning, everybody was really happy, and then you could see that the public started to be critical"

Patient are unable to receive clarity from the Industry as needed, and Data Sharing Platform further propogates the same through lack of research awareness.

Patients seek more awareness on data use in clinical research, including understanding research purposes, posing a challenge for Industry in a competitive environment. Industry acknowledges mistrust and feels powerless to change public perception. The Data Sharing Platform recognizes the need for awareness but maintains a neutral stance on responsibility.

PATIENT ALIGNMENT OF RESEARCH

"You often end up with a very different set of facts and assumptions after you start talking to participants.

Clinical Researcher and Patient want research to be more meaningful, whereas Industry expert believes it already is.

Clinical researchers emphasize patient engagement for relevant research goals aligned with patient concerns. Patients seek research with personal benefits. The Industry expert asserts responsibility for ensuring reliable and safe research outcomes, considering this as a benefit offered to individuals providing data.

DATA REPURPOSE

"If you don't want to re-purpose, then either you need to ask upfront. But if you do it upfront, it's very vague, and you need to be very specific."
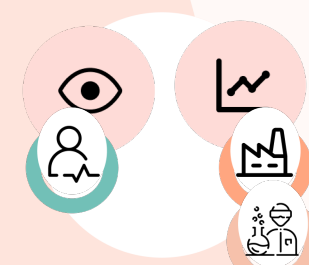
Clinical Researcher and Industry expert see the opportunities for data re-purposing, However they are bound by current regulations towards patient privacy.

Patients emphasize knowing data usage, but regulations limit data re-purposing in clinical research for privacy and consent. Both Clinical Research and Industry find this challenging, as they see potential for re-purposed data to provide deeper insights into future research results.

Figure 16 : Value tensions

3.4 DISCUSSION

ANSWERING RESEARCH QUESTIONS
The exploration of stakeholder collaboration helped frame the current practices and responsibilities such that opportunities can be further developed. An examination of their values further supports the same.

Stakeholders defined as 'use-focused' (Clinical researcher, Pharmaceutical Industrialist and Data Sharing Platform) were found to have diverse needs in relation to their responsibilities. It was found that their focus towards 'care' had primary associations with their requirement from the patient. Despite having distinct motivations, they were found to be pro-active towards their collaborations with one another (often too with the patient.)

Meanwhile, stakeholders defined as 'care-focused' (Patient, Doctors and Patient Organization) had a unanimous focus towards 'care' despite their diverse needs. Stakeholders such as Patient Organization that are not only focused on 'care' but also 'use' were found to be collaborative in their approach towards other stakeholders.

This nuanced understanding of stakeholder dynamics informs the broader examination of privacy challenges within the European health data sharing ecosystem, considering technological advancements and regulatory frameworks.

LIMITATIONS
Results (including personas, stakeholder map and values) defined during the study are based on interview with singular stakeholder representatives, and may mean that they acknowledge only a singular perspective of a diverse environment. The novel approach of Emerging Scope Methodology developed by the researcher towards data analysis has not previously been tried and tested, and may lead to unknown bias in the study.

3.5 KEY TAKEAWAYS

ON COLLABORATION
Stakeholders (Patient Organization, Clinical Researcher, Data Sharing Platform and Pharmaceutical Industry) were found to express interest in collaboration with each other for the purposes of fulfilment of their responsibilities.

ON TRANSPARENCY
Challenges were found in obtaining detailed information about data usage in clinical research, leading to a concern in Patients regarding transparent practices followed by the pharmaceutical industry.

ON PATIENT-CENTRICITY
Stakeholders such as Patient Organization and Data Sharing Platform were found to have the highest care-focused values towards the Patient. Furthermore, it was found that a patient's direct involvement with the research industry is minimal.

ON RESEARCH PRACTICES
Both Clinical Researcher and Pharmaceutical Industry emphasize on the quality of compliance followed by research institutions, wherein GDPR compliance is followed.

Data processing in clinical research is found to be the most cumbersome task followed by Clinical Researchers, wherein they examine the quality and nature of data to remove inadequacies within the same before use for research purposes.

The Pharmaceutical Industry further illuminates a need for re-use of data for sustaining the research practices, with a suggestion towards ease of burden towards Patients for the same.

4. PROBLEM DEFINITION

Framing the problem to be solved

The chapter re-frames the findings from Chapter 2 and 3 for development of a problem definition. First, it describes brief key takeaways and defines the important factors to be examined for development of a solution. Subsequently, it uses research findings from Chapter 2 to frame a problem, which is further developed. To finalize, it provides a Problem Statement and a Program of Criteria with necessary parameters to be included in the solution.

4.1 INTRODUCTION

OVERVIEW

Based on the challenges discovered in the stakeholder environment through Stakeholder research, an examination of the scope is considered for the problem definition. For the same, a network map was crafted (Refer Figure 17.) Within this framework, key components of the health data sharing ecosystem were defined as follows: Stakeholders and their values (includes Patients, Doctors, Patient Organizations, Clinical Researchers, Pharmaceutical Industry and Data Sharing Platform), health data exchanges (EHRs), Technology used (Federated Learning, Encryption Technology) and regulations (GDPR.) A detailed version of the same is provided in Appendix B.

SCOPE

The following criterion were considered pivotal for further development of the problem statement.

UTILIZATION OF HEALTH DATA

The concerns regarding privacy regulations, data quality and re-identification of individuals requires technological solutions capable of addressing privacy issues while maintaining the integrity of the data. Exploring methods like Machine Learning (ML) and encryption holds promise in addressing these concerns.

PATIENT-CENTRIC FOCUS

Patients are both sources of data and ultimate beneficiaries of healthcare advancements. The demand for transparency, and acknowledgment of patient rights aligns with the necessity of transparency required for patient recruitment through data sharing platforms for clinica research.

COMPLEX COLLABORATION DYNAMICS

Examination of current collaboration indicates an interest for collaboration, however this is a challenge given the variety of needs and concerns of each stakeholder associated in the data sharing environment. Moreover, this needs to further be aligned with technological capabilities of data sharing platforms.

BALANCING TECHNOLOGY AND REGULATIONS

Integrating technological innovations such as Federated Learning (FL) and encryption methods within regulatory frameworks, such as GDPR, requires an innovative apprach to individual privacy within the data sharing environment. Moreover, through the use of technology, data sharing platforms can further support the initiatives such as DGA and EHDS. Ethical challenges however need to be additionally acknowledged for application of the same.

Figure 17 : Network map

4.2 PROBLEM DEFINITION

The following values identified during earlier stages were used to further complement a problem statement with an opportunity.

PROBLEM OF TRANSPARENCY (VALUE TENSION)
The recognition of transparency as a value tension serves as the foundation of the problem to be solved. Challenges in provision of detailed information about data usage in clinical research, highlights the necessity for addressing challenges related to detailed information disclosure, ensuring that patients are well-informed about how their data is used in research.

OPPORTUNITY OF COLLABORATION (VALUE FLOW)
Numerous stakeholders were observed engaging in collaborative efforts to reinforce and improve their respective responsibilities. Their expressed interest in collaboration can contribute to redefining collaboration terms in terms of their engagement with patients.

Based on the elaboration of key elements from the scope and findings of opportunities within value relationships, the following problem definition is stated-

PROBLEM STATEMENT EXPLAINED
The problem statement addresses the imperative need for a transparent approach to health data sharing in clinical research, emphasizing collaboration to navigate privacy concerns and stakeholder dynamics.

Patients face a lack of clarity regarding the implications of their involvement during recruitment, primarily within the consent practice. The proposal of use of Value Flow identified, involves leveraging existing collaboration practices to establish familial relationships through a data sharing platform.

The stakeholders to be directly affected by the same include includes patients, researchers, and the data sharing platforms, while patient organizations, doctors, and pharmaceutical industries are indirectly affected.

The problem occurs during patient recruitment, driven by the increasing challenges in disclosure faced by the pharmaceutical industry. Transparency is crucial in accommodating these shifts and addressing recruitment challenges.

Create a transparent approach to health data sharing through a collaborative environment that supports the privacy and stakeholder dynamics in clinical research.

Figure 18 : Problem Statement

4.3 PROGRAM OF CRITERIA

Development of the problem statement helps to identify what needs to be further developed. Moreover, given the scope of research carried out before, criterion were defined that will help shape the solution and support in validation of the results. The following criteria were established-

PERFORMANCE

COLLABORATION
It should support the collaboration between different stakeholders, such that individual values for privacy of each stakeholder are maintained.

TRANSPARENCY
It should promote transparency, such that individual differences in data sharing for individuals are incorporated.

DATA PROCESSING
It should facilitate the data processing during the data sharing process, such that it may promote use of data for stakeholders.

ETHICS
It should incorporate the ethical considerations of artificial intelligence, such that implementation of technology is examined.

RESPONSIBILITY
It may incorporate different roles and responsibilities (than current) for each stakeholder, such that patient-centricity in the process is defined and maintained.

STANDARDS

REGULATIONS
It should be designed within the regulatory frameworks of GDPR, whilst considering initiatives such as EHDS and DGA.

CONSENT PRACTICE
It may incorporate new consent practices for clinical research, that lead to new pathways for future regulatory practices.

5. IDEATION

Developing the solution

The chapter develops three verticals of Patient Centricity, Collaboration and Transparency for further development and implementation for the solution. Firstly, it explores the meaning of Patient-Centricity through a participatory approach of co-design. Subsequently, it develops interactions for stakeholders for a transparent approach. Finally, it examines the use of technology in implementation of transparency while addressing ethical challenges in the use of Machine Learning. To finalize, it provides Key Takeaways to establish a thorough understanding of the factors of three verticals to be included in the Solution.

5.1 INTRODUCTION

Through the problem definition, it was identified that there is a need to create a transparent approach to health data sharing that addresses the needs of stakeholders. Herein, transparency is defined through the lens of the patients that offer their health data for clinical research, hence bringing in a patient-centric approach towards health data sharing. Moreover, the necessity of examination of stakeholder collaboration and technology use, leads to development of three verticals herein: Patient-Centricity, Collaboration, and Transparency. In the upcoming sections, we will delve into these areas to devise a

comprehensive solution. Throughout this chapter, our exploration is guided by the overarching research question:

How can a patient-centric approach be implemented within a collaboration among stakeholders in the health data-sharing ecosystem, leveraging technological capabilities to meet privacy needs and transparency whilst upholding diverse collaboration needs?

5.2 PATIENT CENTRICITY

SCOPE

In exploring the essence of a patient-centric approach in health data sharing, a critical question emerges - what defines patient-centricity? This inquiry forms the basis of a participatory examination, which will further help leverage insights gained in previous chapters. A creative facilitation herein termed co-creation (Sanders & Stappers, 2008) was implemented to help reveal insights into the patient-centric nature of health data sharing.

CREATIVE FACILITATION

AIM
The goal of the co-creation session was to build a collaborative insight into patient-centricity for a health data sharing platform, such that an understanding of the relevant meaning of the terminology is generated within the context.

PARTICIPANT RECRUITMENT
The team's recruitment process utilized Convenience Sampling (Miles, Huberman, & Saldaña, 2014) by reaching out to potential participants through Whatsapp or email. The choice of Convenience Sampling was driven by the researcher's limited access to a broader

pool of interested individuals. However, I selected participants based on their relevant work backgrounds or past experiences, aiming for diversity to ensure an unbiased approach during the co-creation session. Additionally, members from a data sharing platform company were specifically invited to foster a collaborative atmosphere for idea generation. A total of 6 participants, each bringing a distinct background, willingly joined the co-creation session.

ETHICAL CONSIDERATIONS
Before the session started, participants were asked for approval for photography for later analysis. Furthermore, the co-creation session proposal was reviewed and approved by the HREC committee at TU Delft.

PROCEDURE
The co-creation session was carried out in a conference room at YES!Delft, under the facilitation of the researcher. The session

comprised four distinct sessions (see Figure 20, 21, 22, 23), with insights from the initial three sessions helping participants in conceptualizing themes for the final session.

The first session, Introduction to Context (30 minutes), involved participants acquainting themselves with session particulars, encompassing guidelines for fostering a secure and collaborative idea generation space. Participants introduced themselves, elucidating their motives for joining, and engaged in the 'Purge' activity to initiate preliminary idea generation.

In the second session, Problem Understanding (30 minutes), participants explored stakeholder values displayed on the wall. These values, derived from insights of a prior user research study, were presented in simple language, prompting participants to scrutinize and discuss the stakeholders involved in the health data sharing

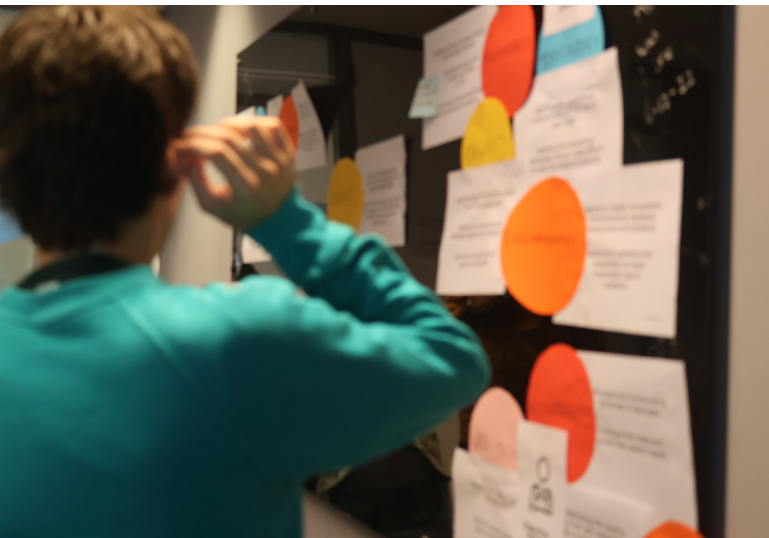Figure 19 : Co-creation session with participants

environment, enhancing their comprehension.

The third session, Exploration through Brainwriting (30 minutes), required all six participants to ideate on five probes related to values such as well-being, dignity, transparency, data protection, and accountability—identified as crucial during user research for patients. Using post-its for concise thought expression, participants spent five minutes on each question while standing. Before the subsequent session, participants reviewed answers to grasp the range of ideas generated.

In the fourth session, Theme Generation and Reflection (45 minutes), participants named and crafted three themes important to them from the accumulated ideas. The two company-affiliated participants stepped aside to observe, minimizing bias. The other four participants collated post-its from the five probes, forming three themes with mixed post-its. This stage proved challenging, with facilitator assistance in fostering creative thinking. Once themes were generated, they

were summarized and presented to the company. Subsequently, all six participants selected their top five ideas across the three themes.

DATA COLLECTION
All ideas generated were captured on post-its. The ideas generated during Brainwriting were then used for the theme generation. No new post-its were generated during the last session.

DATA ANALYSIS
The analytical process focused solely on the three themes crafted by participants. This deliberate choice emerged because participants had already placed all pertinent post-its during the Brainwriting session, indicating the belief that all relevant ideas were encapsulated there. The three identified themes (see Figure 24) formed the basis for cluster analysis. Post-its within each theme were systematically grouped to extract insights on topics deemed relevant under each thematic category.

Figure 20-23 : Session activities

HOW DO WE INFORM PATIENTS?
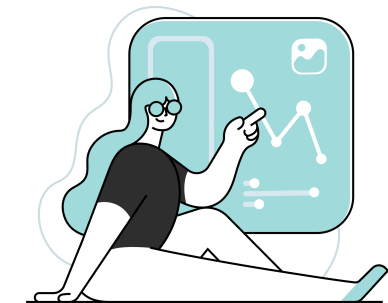
HOW DO WE ENSURE CONSIDERATION?
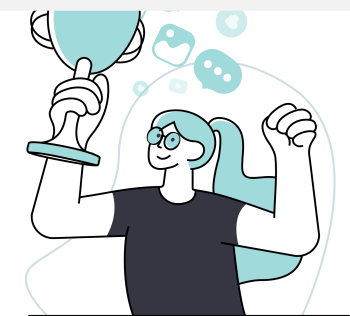
HOW DO WE CONTROL A CRISIS?

Figure 24 : Co-creation themes

5.3 COLLABORATION

RESEARCH FINDINGS

In order to approach patient-centricity that incorporates all three themes generated during the co-creation, the topics identified in data analysis were clustered to create an overview of how to meet the needs of individuals to ensure patient centricity in the health data sharing process.

Two clusters- information and action were defined.

INFORMATION

Information regarding the data and how it is being shared is essential for patients to feel that the data sharing platform is reliable. Moreover, statements that Illustrate the need for research benefits that can help inform patients, based on the nature of the study were identified.

| This demonstrates a need for a clear understanding of not only the data sharing process, but also of the research study to whom the individual data is donated. | INFORMATION | *"How do they share the data with researchers?"* *"Who gets access?"* |

PREFERENCES

Individuals acknowledge the need of having control over their participation. Consideration of their needs further necessitates action towards being offered benefits that are relevant for individuals. More statements were made that suggest a need for individualized approach towards receiving benefits from enrolment.

| This demonstrates a need for the health data sharing space to include aspects during enrolment that allow for individuals to incorporate their preferences in participation. | PREFERENCES | *"Will they pay or offer other benefits?"* |

DISCUSSION

A patient-centric approach in the health data-sharing ecosystem can be implemented by prioritizing Information and Preferences, as covered through the research results. This will be further discussed under key-takeaways at the end of the chapter.

SCOPE

Due to the diverse stakeholder needs and values, it was important to explore a collaborative approach to interaction for the purposes of health data sharing. Based on the Network Map (see Fig X.) of stakeholders performed under Problem Definition, crucial interactions amongst stakeholders were identified to be included in framing the collaboration.

LEVELS OF AGENCY FOR INTERACTION BETWEEN STAKEHOLDERS

In the collaborative framework of health data sharing, 11 individual-based actions (see Figure 25) were identified as crucial for effective collaboration among patients, patient organizations, doctors, clinical researchers, and data sharing platforms. To define a mutual responsibility towards health data sharing,

three levels of agency offered to the individuals were herein examined- 'Traditional,' 'individual agency,' and 'mutual responsibility.' In the 'traditional' agency, interactions reflect current stakeholder responsibilities. In 'individual agency,' individuals have complete autonomy in interactions, while stakeholders maintain their existing responsibilities. In 'mutual responsibility,' both individuals and stakeholders share responsibility in each interaction, modifying their responsibility as needed. Three levels of interactions are further illustrated in Appendix C.

1. individual has ___ access to health data on their platform.
2. individual gives ___ access to healthcare to generate insights.
3. individual wants to ___ delete access to health data on the platform.
4. individual ___ shares data with different healthcare practitioners.
5. individual ___ receives suggestions (for research) from all involved.
6. individual ___ inputs conditions for connection with other patients.
7. individual ___ inputs medical conditions for research study.
8. individual ___ chooses relevant data to offer for all research studies.
9. individual must ___ make informed choices for enrollment in study.
10. individual ___ approves of the data use in study.
11. individual seeks ___ benefit from offering data.

Figure 25 : Individual based actions

TRADITIONAL ROLES

INDIVIDUAL AGENCY

MUTUAL RESPONSIBILITY

Figure 26 : Levels of agency explored

EXAMPLE OF INTERACTION FOR PATIENT CONSENT

From the 11 individual-based actions identified, the interaction for informing patients for consent, between patients, data sharing platforms, clinical researchers and pharmaceutical industry was found to require establishment of different stakeholder responsibilities. The same is described herein-

The example (see Figure 27) illustrates the action of making informed choices for enrolment in the study. The difference between the three iterations demonstrates the change in context-

TRADITIONAL ROLES
Based on the current practice, individuals are informed of details that are offered by the researcher of the study. This is of a "FIXED" nature, in the sense that it is not modified throughout the process. Current problems associated with transparency are identified within this iteration.

INDIVIDUAL AGENCY
Under individual agency however, individuals can choose to be informed of details that they make preferences for under "CHOICE-BASED"

options provided by the data sharing platform. Within this iteration however, the current practice of details offered is followed. While this practice allows autonomy to individuals on the platform, the challenges of information clarity remain due to the "FIXED" nature of responsibility by the Clinical Researcher and Pharmaceutical Industry.

MUTUAL RESPONSIBILITY
The interaction under Individual Agency is found to lack change in responsibilities of clinical researcher and pharmaceutical industry. Herein, "CHOICE-BASED" information provided by the researcher involved in enrolment of participants offers data sharing platforms to provide information on clarification required by the individuals.

DISCUSSION
This example indicates the role played by researchers in offering information. It examines the change in responsibility required by the researcher to ensure transparent practices in information provided. While details are further elaborated in the "Transparency" section below, it implies that researchers and industry may need to assume additional responsibilities for collaboration on the platform.
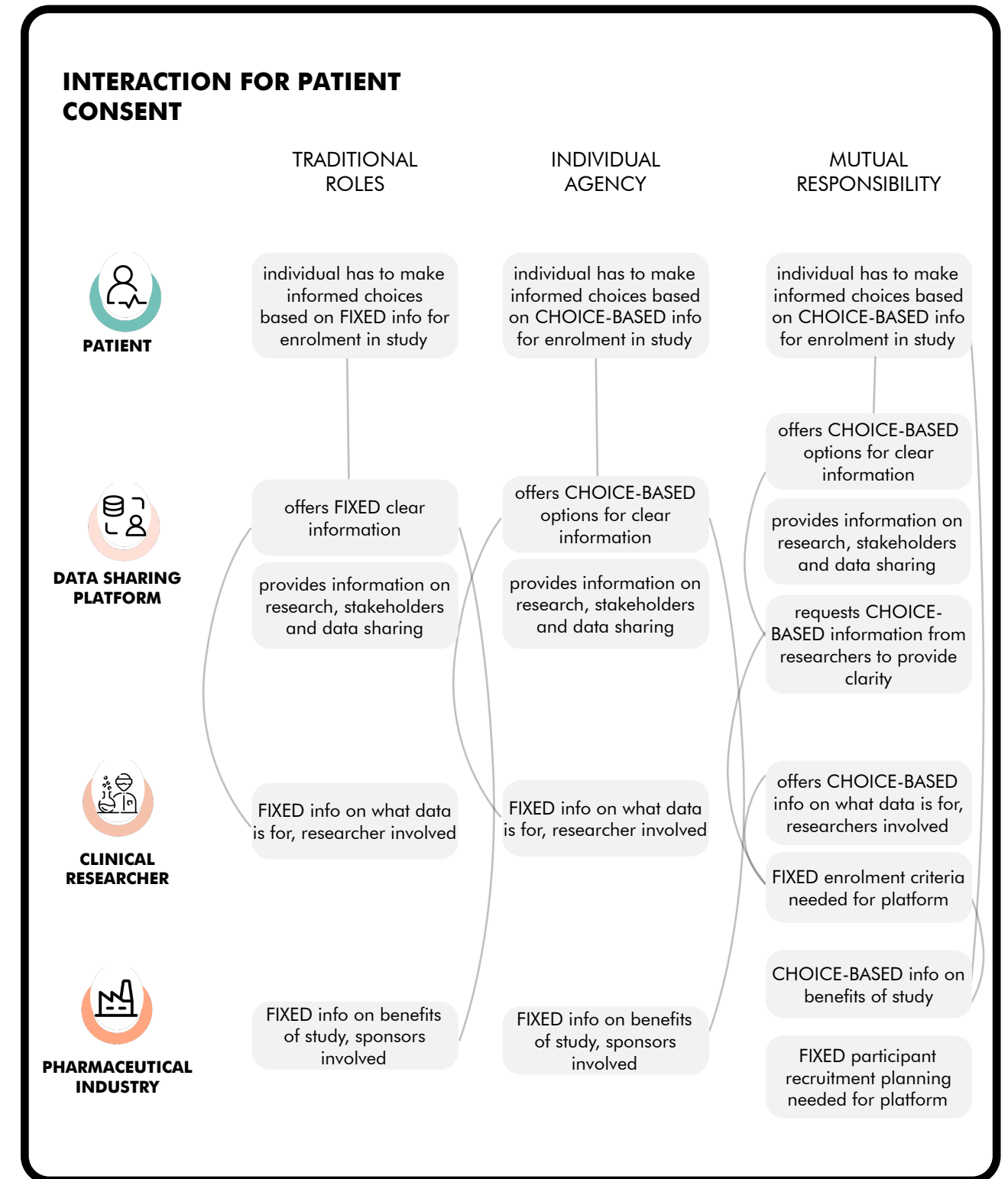

INTERACTION FOR PATIENT CONSENT

	TRADITIONAL ROLES	INDIVIDUAL AGENCY	MUTUAL RESPONSIBILITY
PATIENT	individual has to make informed choices based on FIXED info for enrolment in study	individual has to make informed choices based on CHOICE-BASED info for enrolment in study	individual has to make informed choices based on CHOICE-BASED info for enrolment in study
DATA SHARING PLATFORM	offers FIXED clear information / provides information on research, stakeholders and data sharing	offers CHOICE-BASED options for clear information / provides information on research, stakeholders and data sharing	offers CHOICE-BASED options for clear information / provides information on research, stakeholders and data sharing / requests CHOICE-BASED information from researchers to provide clarity
CLINICAL RESEARCHER	FIXED info on what data is for, researcher involved	FIXED info on what data is for, researcher involved	offers CHOICE-BASED info on what data is for, researchers involved / FIXED enrolment criteria needed for platform
PHARMACEUTICAL INDUSTRY	FIXED info on benefits of study, sponsors involved	FIXED info on benefits of study, sponsors involved	CHOICE-BASED info on benefits of study / FIXED participant recruitment planning needed for platform

Figure 27 : Example of levels of agency explored

5.4 TRANSPARENCY

SCOPE

In exploring transparency within the realm of health data sharing using artificial intelligence (AI) for data processing, it is necessary to examine the ethical challenges associated with its implementation. This examination discusses issues such as explainability of decisions, accountability of outcomes, and the presence of data bias. These ethical considerations play a pivotal role in ensuring the responsible and transparent use of AI in the context of clinical research and health data sharing.

ETHICAL CHALLENGES

Some ethical challenges arise in the deployment of AI in health data sharing as discussed below-

EXPLAINABILITY OF DECISIONS
Explainability, defined as "the goal of an explanation is to make it possible for a human being (designer, user, affected person, etc.) to understand (a specific outcome or the whole system) the logic of algorithmic decision systems" (Henin & Métayer, 2021b)

IMPLICATION
The use of ML algorithms by the data sharing platforms in processing data, lacks interpretability of outcomes generated, and may undermine trust in the processing of individual data for clinical research.

ACCOUNTABILITY OF OUTCOMES
Accountability defined as "a party A is accountable to a party B with respect to its conduct C, if A has an obligation to provide B with some justification for C and may face some form of sanction if B finds A's justification to be inadequate." (Henin & Métayer, 2021b)

IMPLICATION
Moreover, the use of ML algorithms require providing justifications on decisions regarding parameter use in the algorithm. Data sharing platforms require to take accountability for their use of the same to generate results for clinical research purposes. The lack of the same may hinder collaborative accountability.

DATA BIAS at source of data
Data bias, defined as "when the data samples used to train and test algorithmic systems are insufficiently representative of the populations from which they are drawing inferences, leading to discriminatory outcomes because the data being fed into the systems is flawed from the start " (Leslie, 2019)

IMPLICATION
Data sharing platforms are further accountable to transparently indicate the nature of training data for generation of ML models. Improper use of datasets may lead to incorrect outcomes.

PROPOSED SOLUTIONS

To address the potential ethical challenges of AI in health data sharing, the following approaches are recommended:

EXPLAINABILITY AND ACCOUNTABILITY towards researchers

Validation of algorithms and the creation of model cards can ensure ML models are appropriate for clinical research (Weissler et al., 2021.) In the health data sharing ecosystem, platforms can enlist model parameters and share details on the population gathered, fostering trust in ML models for research use.

Moreover, through sharing details on population gathered, these can be reviewed by the researcher, further build ing trust in machine learning models for the use of research. However, this further implies the need for specialist involvement for understanding and correlating machine learning characteristics with research purposes. Current researchers may not be capable of such responsibilities.

DATA BIAS for health data
Issues of data bias relate to data quality and access, impacting interoperability (Weissler et al., 2021.) To ensure system interoperability, ML tools should read and understand inputs across different data sources for intended use (Askin et al., 2023.)

Structured data with insights generated from NLP of patient reports will supplement the necessary information required for eligibility screening. Moreover, the same can be used to make categorizations of data set that can be used to match patients to clinical research for patient recruitment.

DISCUSSION
Mutual collaboration for establishing trust and transparency of use of ML requires additional responsibilities carried out by data sharing platforms and researchers. This is further examined in the Key Takeaways section.
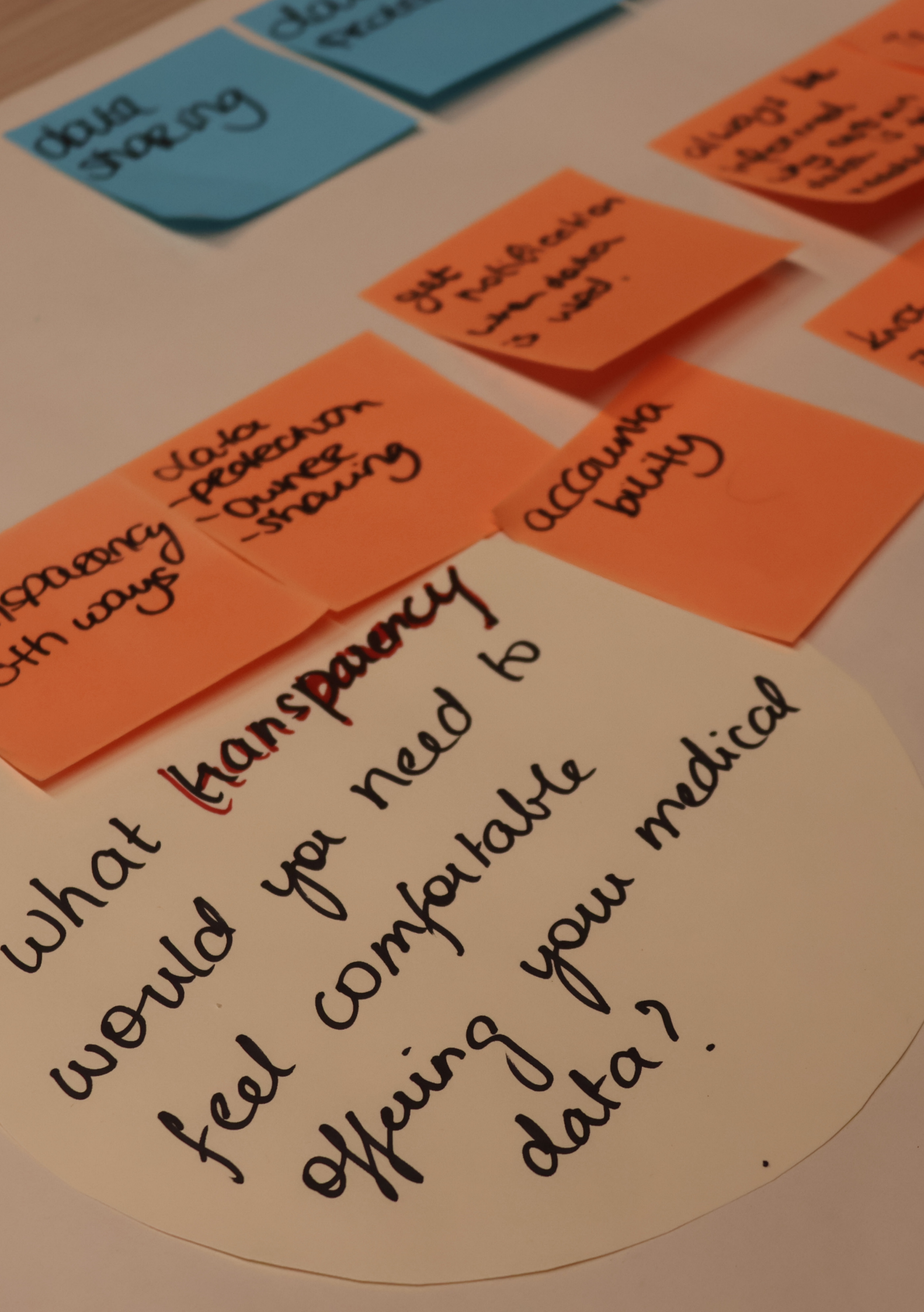
5.5 KEY TAKEAWAYS

PATIENT CENTRIC APPROACH

INFORMATION and PREFERENCES were identified as fundamental requirements of a patient-centric approach towards transparency. Individuals were found to prioritize understanding the nature of use of their data. This was further described as information regarding the research study affiliations and data processing risks. Furthermore, individuals define the necessity of preferences of research benefits for the purposes of their interest.

COLLABORATION FOR MUTUAL RESPONSIBILITY

It was identified that responsibilities of clinical researchers and data sharing platforms needed to change for a mutual collaboration approach towards transparent practices.

RESPONSIBILITY OF CLINICAL RESEARCHER

To promote the Information requirement of individuals for transparency, researcher and pharmaceutical industry were found to have a change in their current roles. They were required to offer Information regarding the use of data and their affiliations and provide choices for research benefits for Preferences.

Moreover, due to the use of ML, researchers require to closely monitor the parameters of use in ML models to be used by data sharing platforms. This requires review of FL models based on the parameter and training data used.

RESPONSIBILITY OF DATA SHARING PLATFORM

As intermediaries between individuals and researchers, data sharing platforms were found to require offering Information regarding data processing to be carried out.

Moreover, in collaboration with the researcher, they require sharing details of parameters, training data for a mutual collaboration based on transparency and trust, that addresses the ethical challenges of Explainability, Accountability and Data Bias in clinical research.

This section summarizes the nature of Collaboration, Transparency, Data Processing , Ethics and Responsibility as stated under Performance criterion for Program of Criteria.
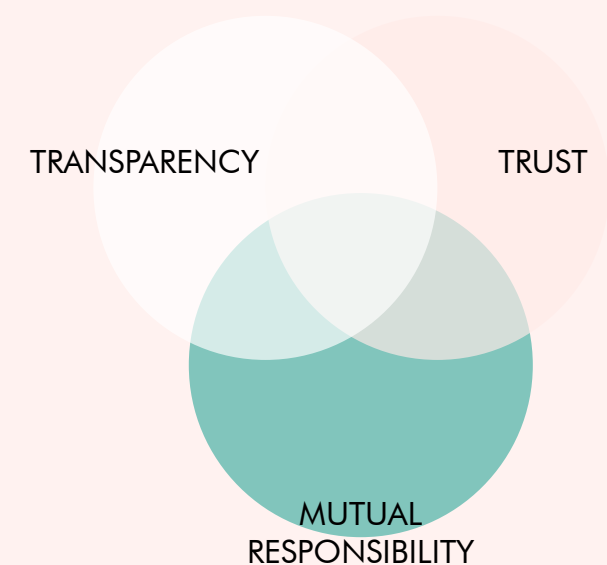
TRANSPARENCY   TRUST

MUTUAL RESPONSIBILITY

Figure 27 : Validation of Health System Framework

6. SOLUTION

Negotiated consent in the health system

This chapter describes the solution proposed based on the context and problem identified. Firstly, it introduces the concept of Negotiated Consent through three stages, and subsequently demonstrates it's implementation through them. Herein, three solutions are explored- Health System Framework, which indicates change in responsibility for collaboration, Data Processing, which demonstrates the use of technology for a transparent approach, and Patient Consent, wherein the user experience of Negotiated Consent for transparency are demonstrated.

6.1 INTRODUCTION

The envisioned solution is the proposal of a concept termed 'Negotiated Consent' defined within the context of the stakeholder environment and technology needed for a transparent approach towards health data sharing in clinical research.

The concept is described herein through three stages-
1. The systemic stage, using a Health System Framework that addresses the responsibilities of stakeholders for a collaborative approach.

2. The technical stage, using a Data Journey Map that illustrates the use of technology for data processing whilst following the regulatory standards and,
3. The user experience lens, using a Customer Journey and User Interface that demonstrates the implementation of the concept for patients.
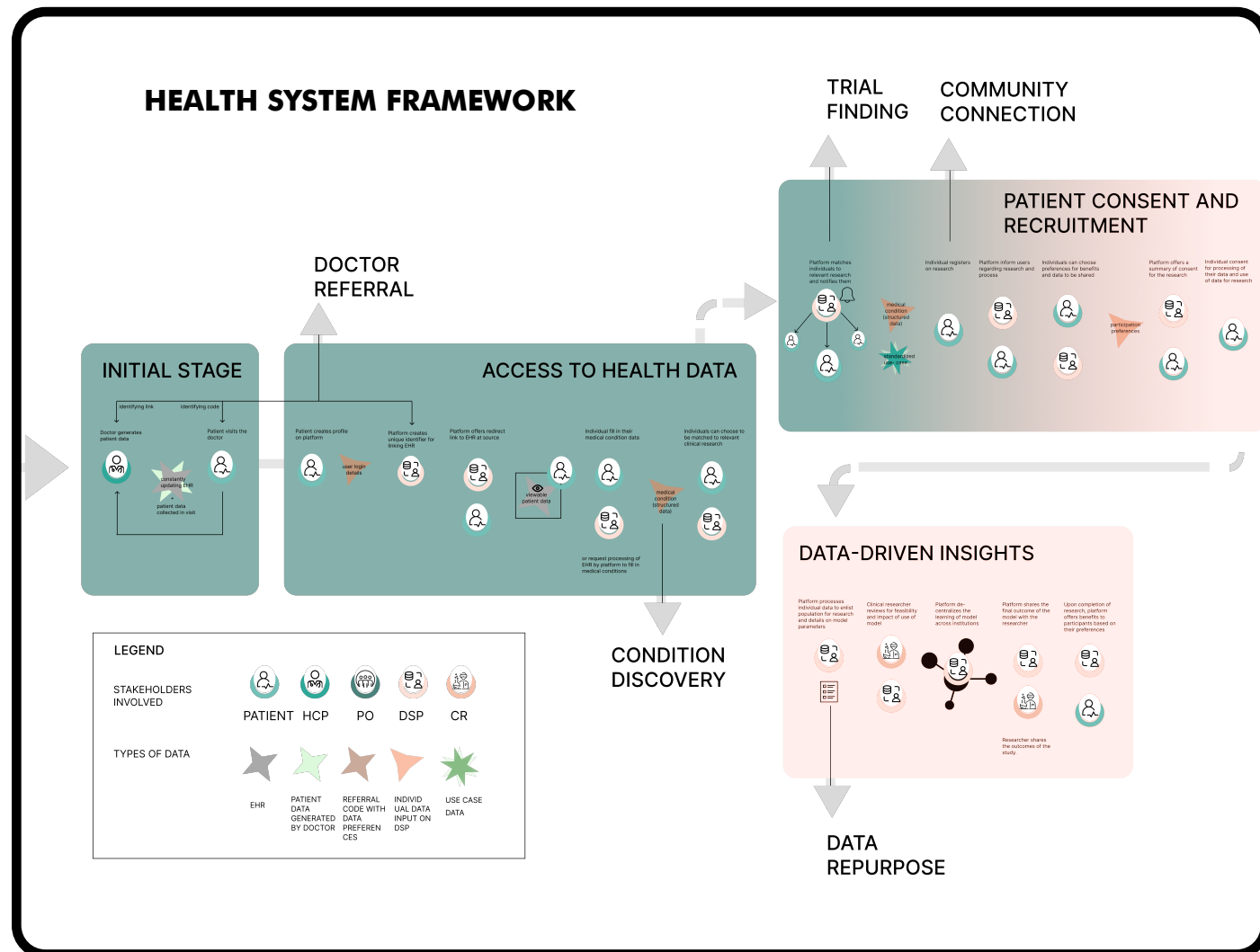
6.2 HEALTH SYSTEM FRAMEWORK

INTRODUCTION

The envisioned framework incorporates the changed responsibilities of multiple stakeholders during their interactions with the patient in the patient journey. It is built around the idea that a data sharing platform can collaborate with each stakeholder during an individual's patient journey, providing an ease of access to shared data through a transparent, patient-centric approach (see Figure 28.)

The framework addresses patient data generated at the source of doctors and evaluates the pathways for use of this data towards meeting needs and responsibilities for each stakeholder, whilst focusing on connecting individuals to clinical research. The patient-centric approach included in the framework is brought about through choices for preferences in interactions, and information for transparency. The patient data is shared through the consent of the individual on the data sharing platform, which connects individuals and stakeholders together.

STRUCTURE

The final structure of the framework includes 8 components based on the Performance criteria Collaboration. Three components including - Access to Health Data, Patient Consent and Recruitment, Data-Driven Insights cover essential aspects of connecting individuals to clinical research through a data sharing platform, and will be the main focus of discussion.

The details of other 5 components, such as- Doctor Referrals, Community Connection, Condition

Figure 28 : Health system framework

Discovery, Trial Finding, Data Repurpose, which illustrate the various uses of the same platform to ensure meeting the requirements of the stakeholders involved, and can be found in Appendix D.

ASSUMPTIONS IN THE HEALTH DATA FRAMEWORK

To clarify the framework components, the following assumptions are outlined:

1. Healthcare institutions are identified as patient data sources. Patient visit to the doctor is the considered initial stage.
2. Web APIs facilitate data sharing to the platform from healthcare institutions.
3. Individuals access health data through a redirect link to institutional infrastructures.
4. Monetary benefits vary based on study financiers.
5. Researchers detail research data for participant enrolment.
6. Data connectivity to the platform depends on prior collaborations with healthcare institutions.
7. The platform has access to studies' research data.
8. Collaboration with patient organizations may differ, primarily noted in their presence on the data-sharing platform.
9. The motivations of Pharmaceutical Industry is represented through the actions of Clinical Researcher, and remains inconsequential as an independent body herein.

While these aspects have been addressed in the framework, they are aspects outside the scope of the problem definition and further work needs to be done to develop an understanding of how to incorporate them into the system.

INITIAL STAGE

The initial step mandates a pre-established Data Processing Agreement (DPA) between the data sharing platform and healthcare data controllers, outlining data processing procedures (Exams, 2023.) An illustrative DPA example is available (Brancheorganisaties Zorg, 2022.)

In this phase, doctors inform patients about data access via the platform, seeking consent. Upon consent, a unique patient enrolment link is generated and shared for registration using an email address.

For patients with existing platform profiles, data linking can be requested by providing a unique identifier code generated on the platform.

The individual components are herein described, with the interactions of stakeholders involved for each. Each component has been described through explanation of their individual aim, relation to the framework, and functions within each. Whilst examining the same, patients using the system are herein termed as individuals within the functions.
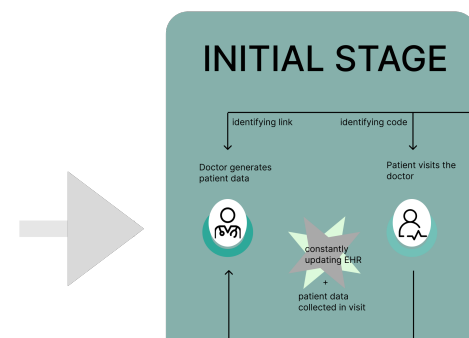

Figure 29 : Initial Stage (Health system framework)

ACCESS TO HEALTH DATA

AIM
Empowering individuals with access to medical data by enabling viewing of personal medical records across different institutions (see Figure 30).

RELATION TO FRAMEWORK
By having access to medical data, individuals can view and make informed decisions over control of their data shared amongst different stakeholders.
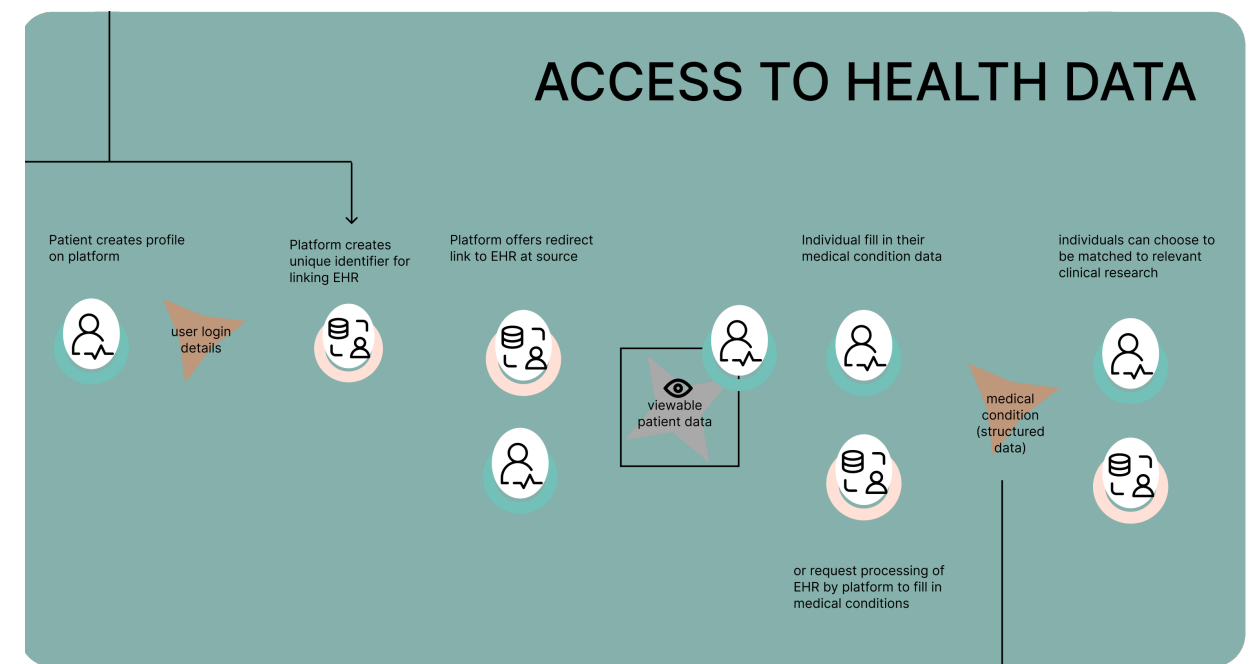

Figure 30 : Access to Health Data (Health system framework)

FUNCTIONS

F1 HEALTH DATA DISPLAY
Upon registration and linkage with doctors, the platform establishes a redirect link to the individual's database, displaying their health data without processing. No data is processed, ensuring the individual's health information remains inaccessible for processing on the platform till individual consent.

F2 INDIVIDUAL DATA INPUTS
Users can personalise their profiles by adding personal details and medical conditions based on medical history.

F3 PARTICIPATION IN CLINICAL RESEARCH
Individuals, based on their medical conditions, may opt to be matched with ongoing clinical research. Approval allows the platform to process Electronic Health Records (EHRs), accurately categorizing medical conditions and incorporating them into the user profile.

VALUES FOR STAKEHOLDERS

The platform empowers individuals to control their experiences by facilitating access to their health data at healthcare institutions. This reinforcement of self-assertion contributes to the value of **dignity**.

Doctors inform individuals about privacy rights, emphasizing choices in data sharing with the platform and seeking consent. Doctors, central to individual **well−being**, become collaborators in the enrolment decision-making process.

The platform, in **collaboration** with doctors, caters to individuals' access needs, providing them with control over their medical records.

PATIENT CONSENT AND RECRUITMENT

AIM
Users can define data-sharing terms upon matching with a clinical research study.

RELATION TO FRAMEWORK
The platform facilitates collaboration between stakeholders and individuals, enabling personalized contracts for clinical research enrolment.

FUNCTIONS

F1 PATIENT MATCHING
Patients are matched to ongoing research studies, incorporating internet-posted and platform-requested studies. Standardized by the platform, this research data ensures accurate matching based on medical conditions from F2 and F3 of Access to Health Data.

F2 INFORMING ABOUT CLINICAL RESEARCH
Registered individuals undergo the consent process, with the platform detailing study implications, data usage, and ownership outside the consent process. A thorough explanation is provided on the platform.

F3 NEGOTIATION OF PREFERENCES
Participants make choices and negotiate data set terms. The platform creates a "negotiation contract" adding preferences to individual's profil data.

F4 REVIEWING NEGOTIATION
The platform summarizes the negotiated contract for individuals to review, enabling approval for data processing and research use after patient consent.
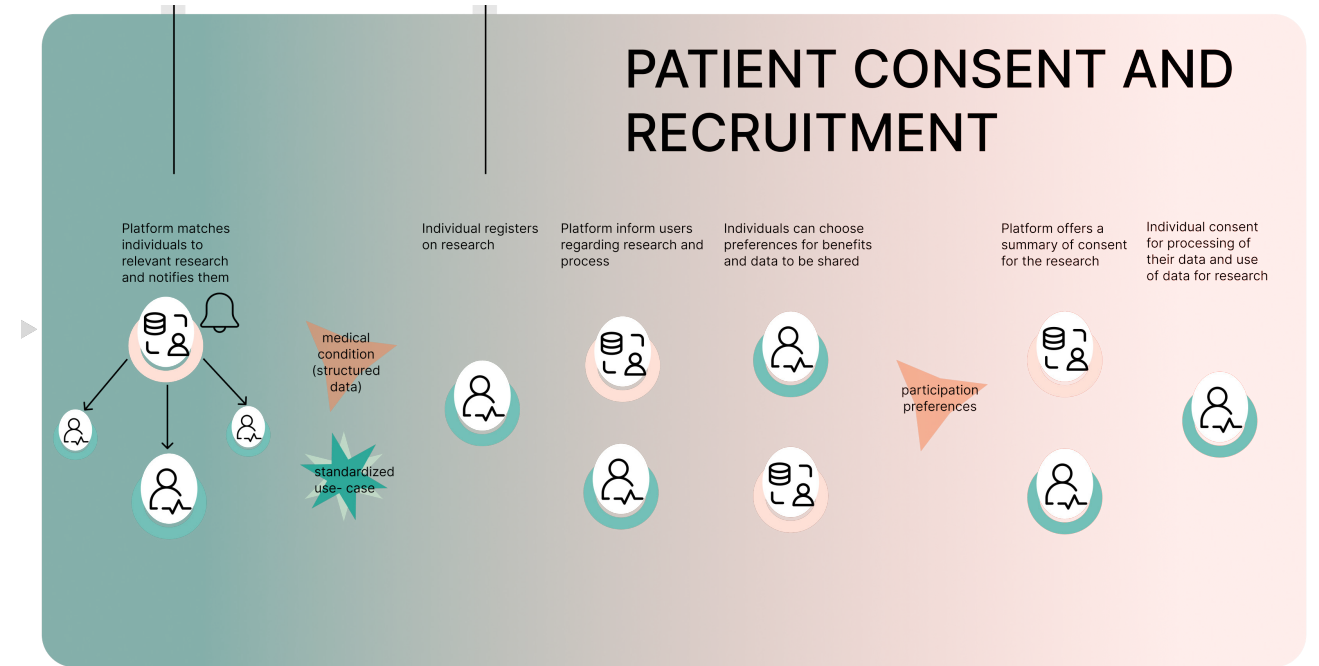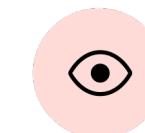
PATIENT CONSENT AND RECRUITMENT

Figure 31: Patient Consent and Recruitment (Health system framework)

VALUES FOR STAKEHOLDERS

The platform promotes **transparency** in clinical research, ensuring individuals are informed about the research process and data sharing. Researchers clarify study implications and data use.

The platform ensures individuals receive information on data usage within ethical standards for **data protection**. Clear understanding of data processing activities simplifies consent for study participants.

By empowering individuals to negotiate **participation** and shared data value, the platform and researchers collaborate, giving individuals a voice in their clinical research involvement.

DATA-DRIVEN INSIGHTS

AIM
Facilitate collaboration between individuals and researchers using the data sharing platform for successful research outcomes.

RELATION TO FRAMEWORK
The data sharing platform's technological expertise provides diverse outcomes for researchers utilizing the service.

FUNCTIONS:

F1 REVIEWING PARTICIPANT DATA AND MODEL FEATURES

Post-negotiated contract approval, the platform processes participant data, generating a feasibility study. Researchers assess population relevance and ML model uncertainties, approving or requesting modifications. The platform proceeds to other functions in Data Driven Insights or Data Repurpose based on researcher feedback.

F2 RECEIVING OUTCOMES OF MODEL
The platform updates the model from local gradients, sharing it across healthcare data providers. The final model is shared with the researcher, including encrypted patient data for various study types. Researchers share outcomes with the platform for participant notification.

F3 OFFERING RESEARCH BENEFITS
Based on preferences in the negotiated contract, the platform provides research benefits, such as study insights tailored to participant preferences and a portal for monetary benefits from researchers.
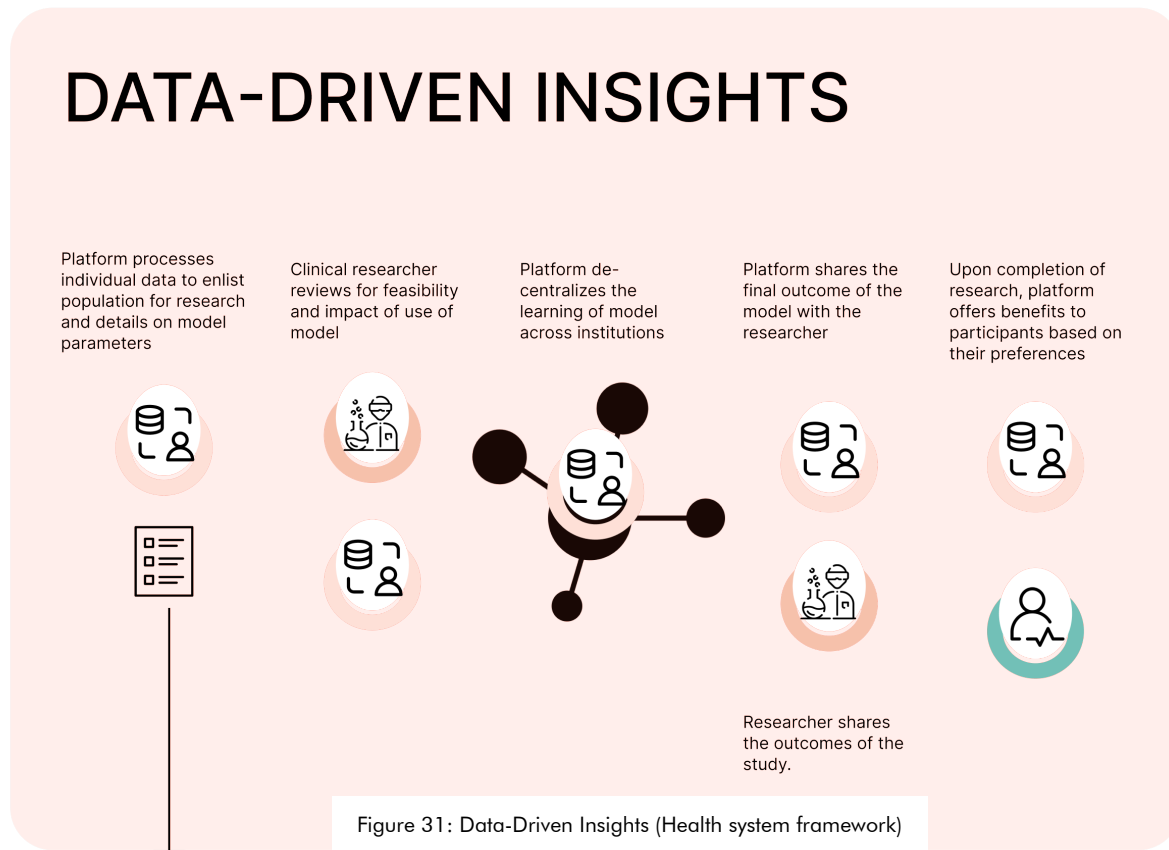
SUMMARY

The above framework demonstrates the stakeholder roles and responsibilities for a mutual collaboration under Negotiated Consent. Herein, collaboration through the data sharing platform creates new responsibilities, wherein the connection to health data and platform are demonstrated for each stakeholder. It further indicates the fulfilment of values of stakeholders under each component.

VALUES FOR STAKEHOLDERS

The data sharing platform minimizes **data protection** concerns of health data through anonymization efforts. Population-based insights protect individual identification, ensuring study relevance for enrolled participants.

Collaborating with clinical researchers, the data sharing platform ensures beneficial outcomes for researchers and participants enrolling through the platform.

DATA-DRIVEN INSIGHTS

Platform processes individual data to enlist population for research and details on model parameters

Clinical researcher reviews for feasibility and impact of use of model

Platform de-centralizes the learning of model across institutions

Platform shares the final outcome of the model with the researcher

Upon completion of research, platform offers benefits to participants based on their preferences

Researcher shares the outcomes of the study.

Figure 31: Data-Driven Insights (Health system framework)

6.3 DATA PROCESSING

INTRODUCTION

The data journey to be described illustrates the processing of data carried out by the platform for the use of clinical research. It includes the process of consent and offers and explains the method of processing to be carried out by the data sharing platform.

ASSUMPTIONS

The data processing to occur on the platform is examined through parameters of data sharing platforms discussed in PHC Catalyst Alliance & Roche Nederland B.V., (2020) and further supports the data journey. Based on the parameters discussed, the following assumptions have been made -

DATA STORAGE
Patient data (EHRs) is not stored outside healthcare institutions and doctor registries. Only local AI models developed by the Data Sharing Platform are transferred across healthcare institutions. Doctors connect patient data across different healthcare institutions through the use of Data Processing Agreements.

DATA PROCESSING
Patient data from doctors' sources undergoes standardization before feeding into ML models. This process, termed data standardization, categorizes both structured and unstructured data, achieving FAIR principles (Findable, Accessible, Interoperable).

DATA GOVERNANCE
Connection permissions to EHR occur during a doctor's visit through a unique enrolment link (doctor-initiated) or a unique identifying code (patient-initiated.)

DATA PERMISSION
Data use permission on the Data Sharing Platform follows the concept of 'Negotiated Consent.'

DATA DISTRIBUTION
The platform accesses EHR data via APIs for display and processing, sharing outcomes through secured servers. Individuals access their EHR records directly, provided with a link by the Data Sharing Platform pre-consent without access to individual data.

RESULTING CONTRACT
The Data Sharing Platform generates and stores a negotiated contract, connecting relevant parties involved based on the agreement.

DATA JOURNEY MAP

The data journey map in Figure 32 illustrates the data processing carried out by data sharing platforms under the Health System Framework discussed above. It indicates three stages of the process- Access to Health Data, Negotiated Consent and Data-Driven Insights for Researchers. It was further validated with cybersecurity and software engineers for its feasibility. In the following section, the legend and component based process will be described.
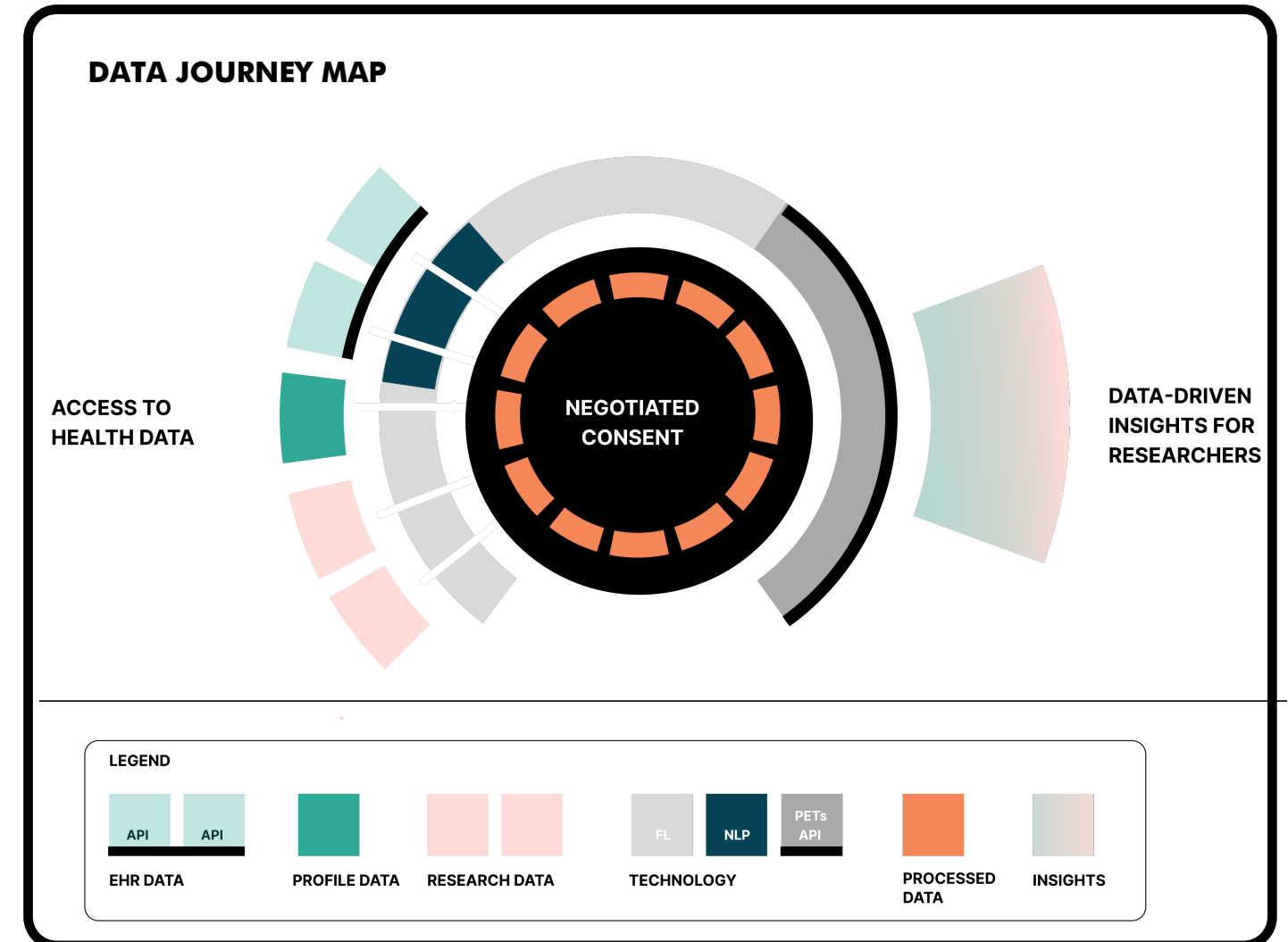

Figure 32 : Data Journey Map

LEGEND
The data journey map shows the use of three types of data- EHR, Profile and Research data. Herein, EHR data is of two kinds, structured EHR data and unstructured, which is transferred to the data sharing platform using API.

Profile data comprises of individual inputs and preferences collected on the platform, and is collected as structured datasets. Moreover, research data includes enrolment criterion offered by clinical researchers and study data found on the internet.

The platform uses FL and NLP to process EHR data, whilst simultaneously processing individual data and research data for successful patient matching. Furthermore, the use of PETs is implemented for results to be shared with clinical researchers.

The results of data processing shared with researcher can vary between encrypted patient data, feasibility studies, retrospective study results, observational research data and statistically analysed data.

COMPONENT BASED SUMMARY

1. ACCESS TO HEALTH DATA

EHR- Health data sharing across providers is facilitated through API. Patients view data via a redirect on the platform, ensuring that individual data remains private from the platform.

INDIVIDUAL DATA- The platform generates structured individual data, including personally identifiable and user behaviour data, stored under one node.

2. PATIENT CONSENT AND RECRUITMENT

Research data - Research study data, created by researchers, are processed as structured datasets, labelled under one node for patient matching by the platform.

MEDICAL CONDITION FILLING - For EHR data consent, patient data is treated as structured + unstructured datasets, standardized using machine learning. The platform identifies and adds patients' medical conditions to user profiles.

PATIENT MATCHING - Collected individual data on the platform is a structured dataset used for patient matching. Matching involves relevant categories of patients across user profiles and research data processed by the platform.

3. DATA-DRIVEN INSIGHTS

NEGOTIATED PREFERENCES - During patient enrolment, negotiated preferences are stored by the platform under the user profile to facilitate EHR data processing for research outcomes.

MODEL PARAMETERS - Before processing patient data, the platform develops a local model with parameters and uncertainties. Reviewed by the researcher, a local model is generated and sent to healthcare institutions, following a federated learning model.

OUTCOMES OF DATA PROCESSING - The platform uses machine learning to offer researchers various outcomes, including anonymized and encrypted patient datasets, and statistical analyses for feasibility, retrospective, and observational studies. Anonymized patient records may use differential privacy learning, while genomic data is encrypted through homomorphic encryption.

NOTE ON REGULATORY STANDARDS

GDPR
By placing patient consent at the centre of all data processing processed by the platform, 'Negotiated Consent' places the control of processing of data at the forefront of enrolment of participation in clinical research.

Data Governance Act
By connecting patients to their health data, the data sharing platform is central to offering individuals the platform to further progress the impact of voluntary data sharing practices under the Data Governance Act.

European Health Data Space
The use of de-centralized data addresses the variation in nature of regulatory standards of healthcare across institutions and borders. This is a push towards the European initiative of health data usability across EU states. The acknowledgement also indicates the necessity of data sharing platforms to develop partnerships with institutions that allow this kind of connectivity.
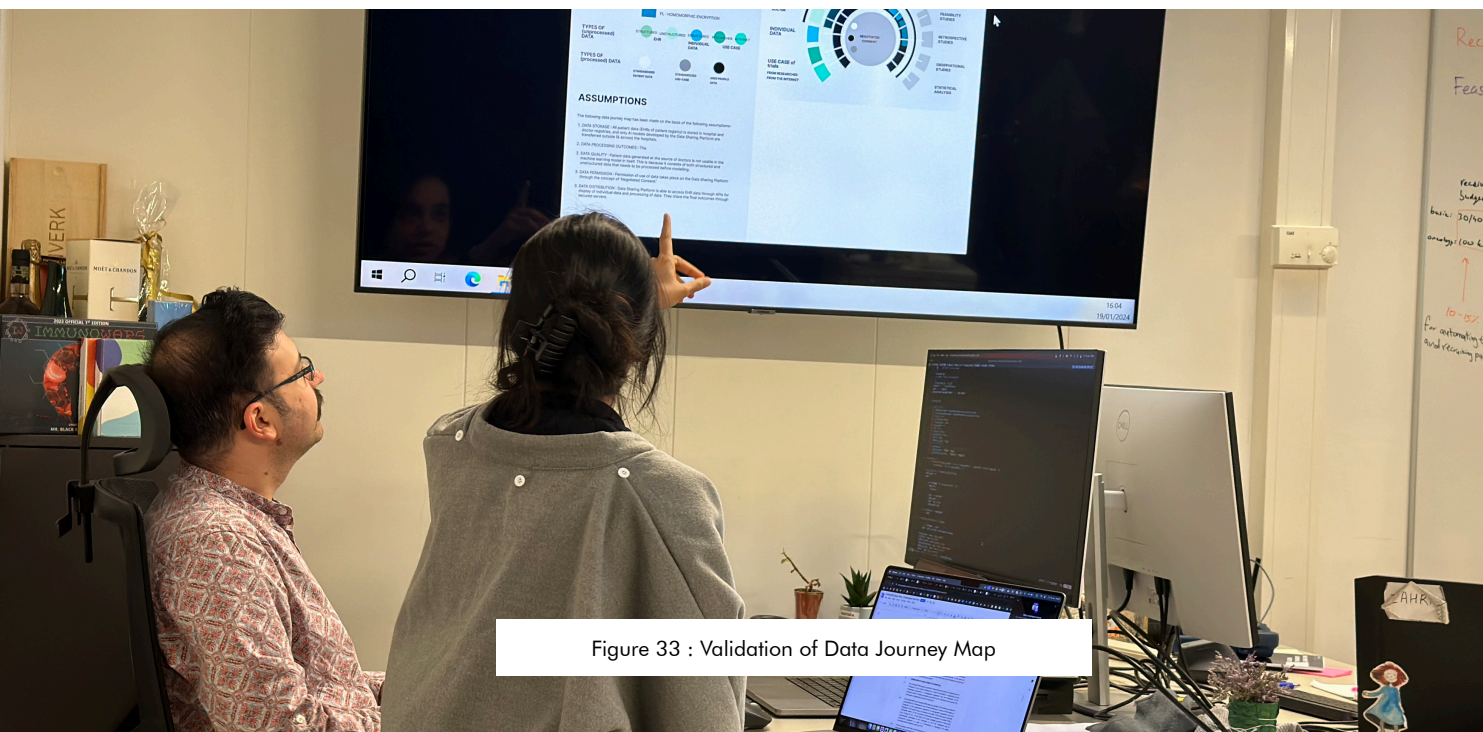

Figure 33 : Validation of Data Journey Map

6.4 PATIENT CONSENT

INTRODUCTION

In the context of participant recruitment through data sharing platforms, the translation of informed consent onto the interface is a critical consideration. In the process of offering data for clinical research, individuals undergo informed consent, where they learn about the study, evaluate options, and understand associated risks. The signing of informed consent signifies an individual's agreement to the study details. However, with the Data Governance Act emphasizing data donation as a part of Data Altruism, the implementation and significance of consent require careful attention.

DATA NEGOTIATION
A recent study by Ortega et al. (2023) explores the importance of data literacy in the context of meaningful data donation and informed consent. The challenge lies in ensuring that individuals are 'adequately informed' and possess a clear understanding of how their data will be used. Data literacy empowers individuals to define their 'privacy boundaries' more effectively, fostering positive associations with contributing data to research.

"Negotiation is not just about what value the individuals can receive, but it should also be about how much of their data individuals want to give for research."
- MSc. Gomez Ortega, A.
PhD candidate at the TU Delft, working on Designerly Data Donation

Within the framework of health systems, particularly under Patient Consent and Recruitment, the concept of "negotiation" has been introduced. This entails offering individuals information on the research and enabling proactive decision-making on data sharing platforms regarding their participation in clinical research. The process of Patient Consent and Recruitment explored within Health System Framework requires further conceptualization.

In conclusion, the incorporation of negotiation into the patient consent and recruitment process represents a significant advancement. By providing individuals with information and the ability to actively shape their participation, the negotiation component adds depth and autonomy to the traditional informed consent process. As data donation becomes a pivotal aspect of Data Altruism, emphasizing data literacy and negotiation in the informed consent process is crucial for fostering positive attitudes toward contributing data for research.

CUSTOMER JOURNEY

The conceptualization of negotiation was initiated with examination of the Customer Journey (see Fig X.) Herein, the key activities defined within the Health System Framework are defined in 4 stages-

INFORM, wherein individual learns about their participation in the clinical research study;
NEGOTIATE, wherein individual makes their 'personal boundaries' for participation;
REVIEW, wherein individual reviews their terms of participation and;
CONSENT, wherein individual finalizes their consent to participation in the research study.

This is further examined for each stage using the User Interface framework developed.

STAGES OF PATIENT CONSENT

INFORM

This stage follows the individual's matching to relevant study, notification, and registration on the study. Herein, the platform shares crucial information to establish transparency and empower individuals in their decision-making process.

The platform provides transparency through two types of information, ensuring a clear understanding for participants regarding the research and data processing activities to be performed on individual data. The researcher's responsibility herein is providing the necessary information for the platform's display.

STAGE	INFORM	NEGOTIATE	REVIEW	CONSENT
KEY ACTIVITIES	1. PATIENTS REGISTER ON STUDY 2. PLATFORM INFORMS PATIENTS ABOUT PROCESS 3. PATIENT APPROVES/CHANGES DETAILS/EXITS STUDY	4. PATIENTS CHOOSE RESEARCH BENEFITS 5. PATIENTS CHOOSE DATA PREFERENCES 6. PATIENT APPROVES/CHANGES DETAILS/EXITS STUDY	7. PATIENT REVIEWS THEIR PREFERENCES AND INFORMATION OFFERED 8. PATIENT APPROVES/CHANGES DETAILS/EXITS STUDY	9. PATIENT CONSENTS/ CHANGES DETAILS/EXITS STUDY
OBJECTIVES	PROVIDE TRANSPARENCY ENHANCE PERSONAL DATA LITERACY	ENABLE INDIVIDUALS TO DEFINE THEIR 'PRIVACY BOUNDARIES'	PROVIDE INDIVIDUALS WITH A REVIEW OF STUDY DETAILS	ALLOW A FINAL CONFIRMATION
STAKEHOLDERS INVOLVED				

Figure 34 : Customer Journey in Participant Recruitment

The two types of transparency displayed are -

1. Based on Researcher's study data
These are to include research goals, specifications on data to-be-collected, researchers and institutions involved and potential risks associated with the research.

2. Based on Platform's data processing
This includes data processing methods, data privacy details and measures taken in case of a data breach or leak.

This is further illuminated in Fig X. The presentation of details related to the platform's responsibility may vary, but following this disclosure, the interface seeks individual approval before proceeding. Individuals retain the autonomy to review details, proceed to the next step, or exit the study.

This stage serves as an opportunity for the platform to provide transparency about the research study to individuals, enhance personal data literacy among them regarding the data sharing process and assist individuals in understanding the use of their data.

NEGOTIATE

This stage follows the individual's initial approval of reviewing the details under INFORM. Herein, the platform empowers individuals to negotiate various aspects of their involvement, fostering a personalized and consent-driven approach.

The platform provides negotiation on two aspects, hereby allowing individuals to choose their benefits for the dataset they offer for the research. The researcher's responsibility herein is providing a clear understanding on the nature

INFORM

NEGOTIATE

Figure 35-36 : User experience framework

of data sets required for their research. This may vary in terms of frequency, level of anonymity or duration of participation.

The two types of negotiation offered are -

1. Negotiation on Research Benefits
For studies offering monetary benefits, individuals are presented with varied payout options, enabling them to choose based on personal preferences. In studies providing research results, individuals have the freedom to select their preferred payout option.

2. Negotiation on Data Offered
Individuals can tailor the granularity of data shared, choosing from a range of options relevant to the examined study. For demonstration purposes of the same, a sample study was used which requires haemoglobin test results, vaginal ultrasound, and questionnaires. (Sonographic Features of Fibroids Before and During Non-surgical Therapy and/or Expectant Management, n.d.)

Upon selection, the platform stores preferences under user profiles, seeking patient approval before advancing. Individuals retain the autonomy to review details, proceed to the next step, or exit the study.

This stage serves as an opportunity to provide individuals with a platform to articulate personalized preferences for study participation by enabling them to define their 'privacy boundaries' concerning their data. This leads to a consent of enrolment which is stored as a negotiated contract developed on the platform.
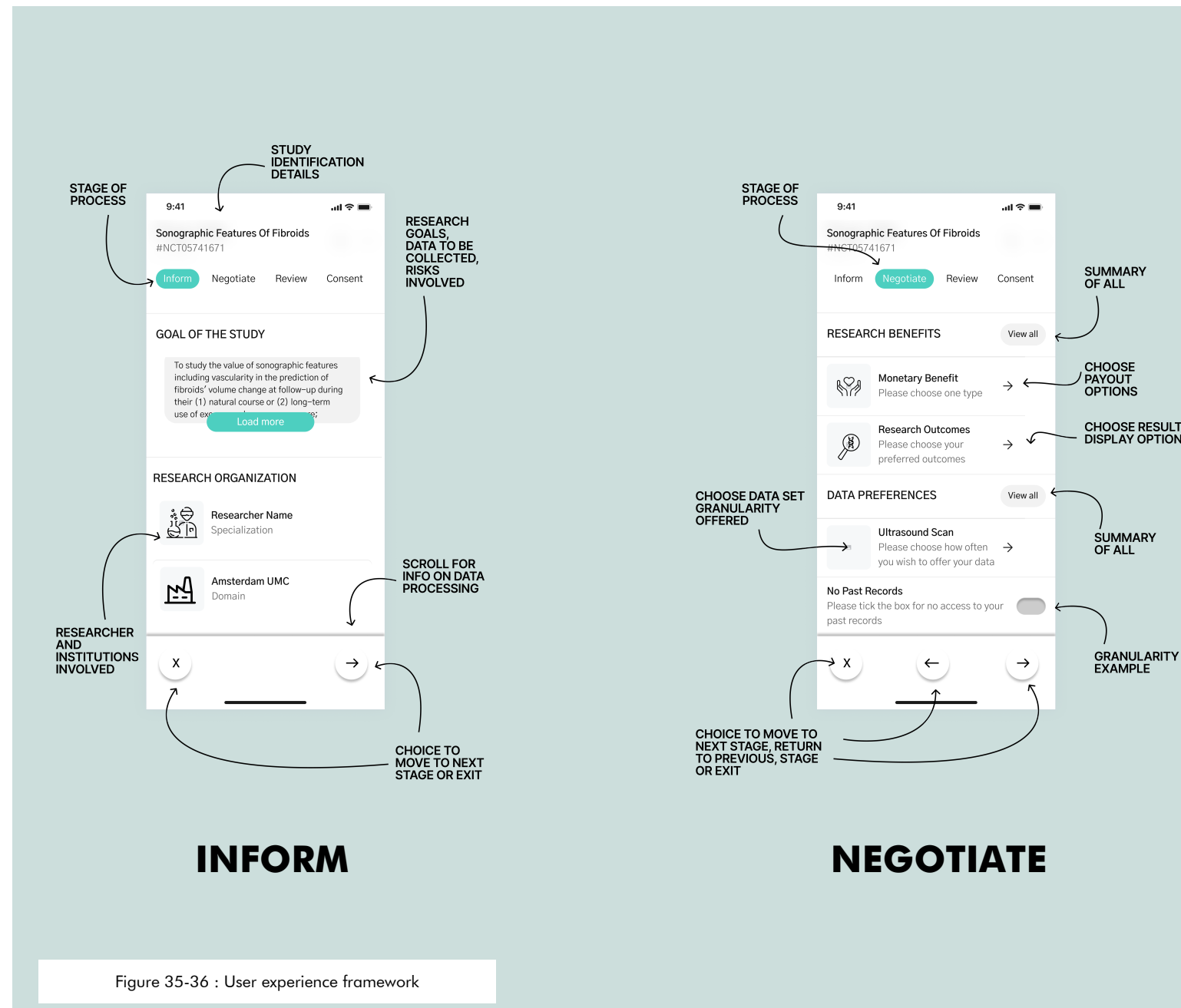
REVIEW

This stage follows the individual's negotiation of participation under NEGOTIATE. Herein, the platform empowers individuals to review their preferences and study details before providing explicit consent.

Herein, individuals can review the details of the study, data processing details, risks associated and their negotiated preferences for participation. At this stage, individuals retain the autonomy to review details, and proceed to the next step, or exit the study.

This stage serves as an opportunity for the platform to provide individuals with a transparent review of study details, to enable them to make an informed decision to consent on their negotiation in the next stage.

CONSENT

This stage follows the individual's review of details and preferences under REVIEW. Herein, the platform allows individuals to formally acknowledge their commitment to the study.

This stage serves as an opportunity for the platform to ensure that individuals enter into the study with a clear understanding, and the agency to make informed decisions for the consent.

By offering comprehensive details about the research and the platform's responsibilities, individuals are empowered to make informed decisions about their participation. The implementation of Negotiated Consent on a user experience aligns with the overarching goal of fostering patient-centric practices within the realm of health data sharing for clinical research.
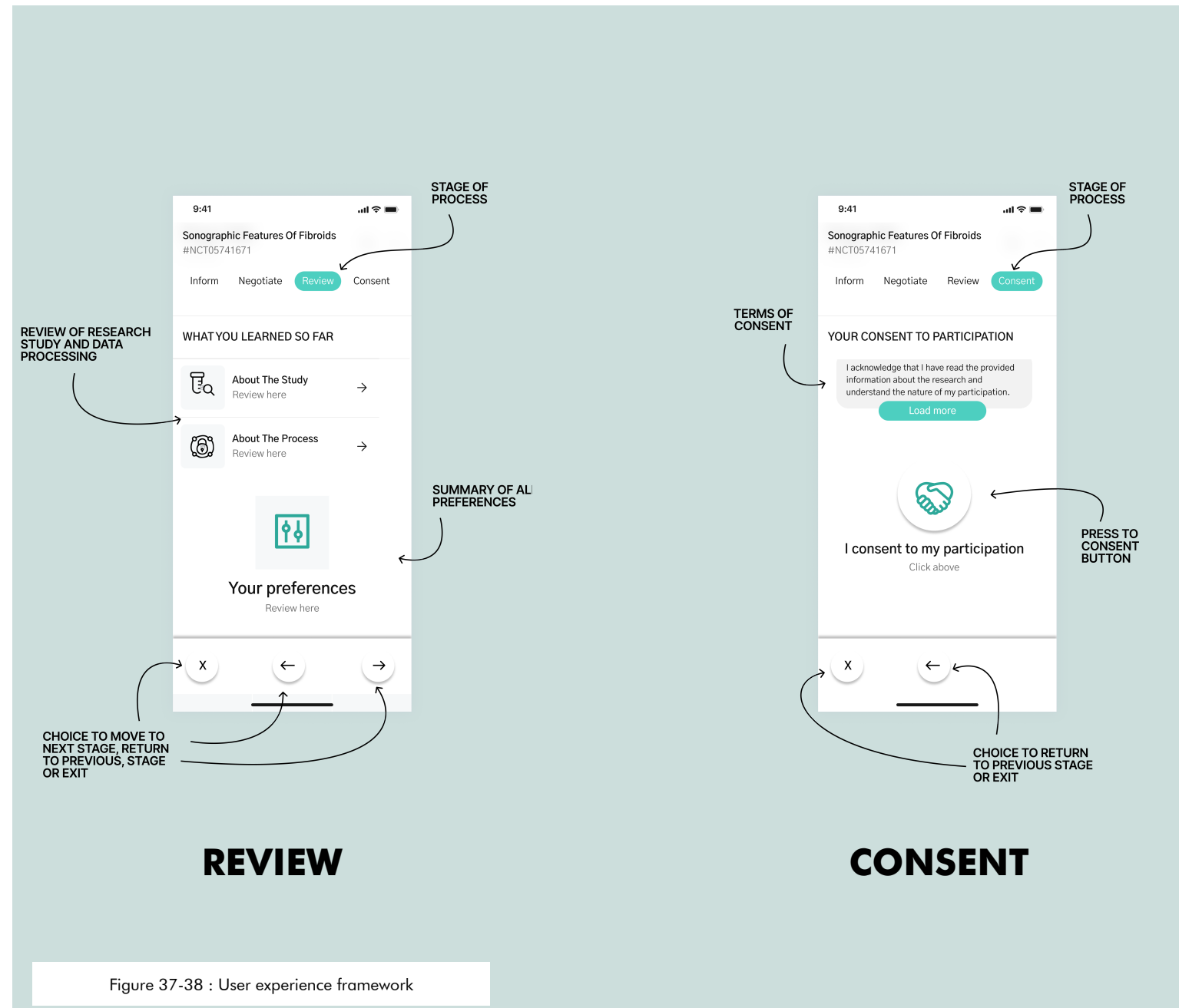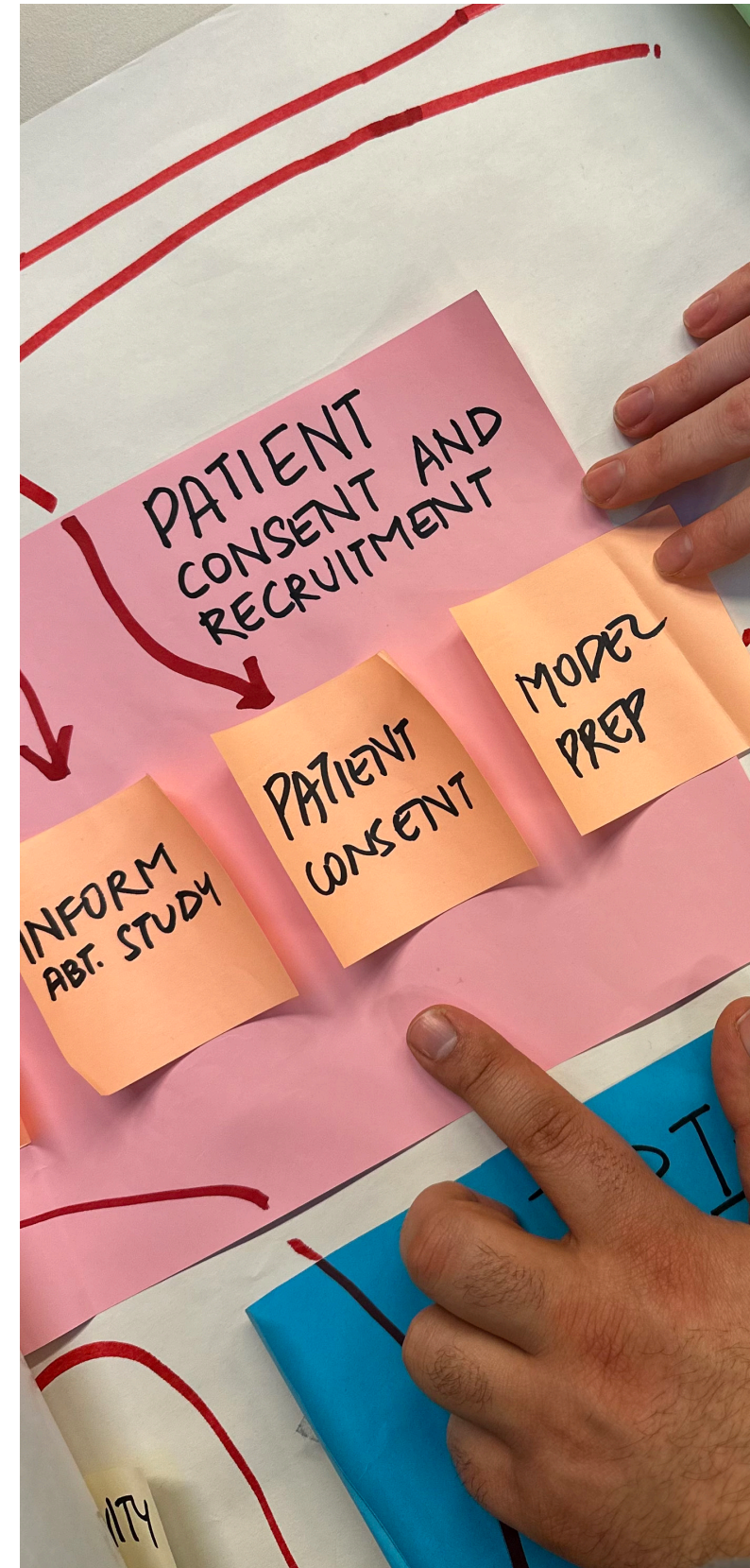
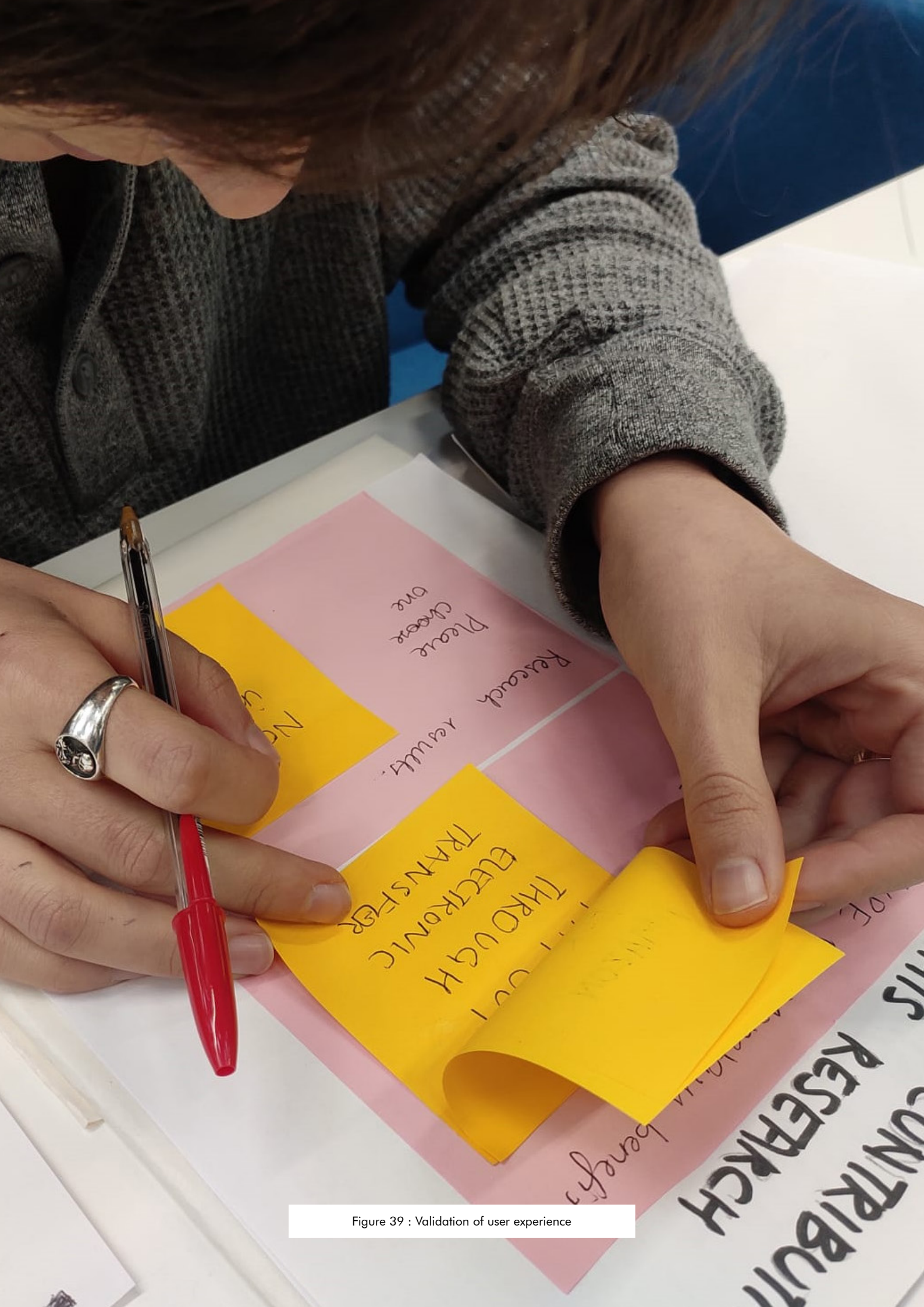

Figure 37-38 : User experience framework

REVIEW

CONSENT

7. VALIDATION

Testing the solution with users

This chapter validates the concept for its feasibility. Firstly, it examines the user experience of Negotiated Consent by development of a prototype of the same. Subsequently, it addresses the transparency criterion developed in solution under Chapter 6. Secondly, it examines the role of Negotiated Consent in the stakeholder dynamic of patient-clinical researcher-data sharing platform to study the effect of the same on the responsibilities of stakeholders.

Figure 39 : Validation of user experience

7.1 USER EXPERIENCE TEST

INTRODUCTION

The concept of 'Negotiated Consent,' was developed under a patient-centric approach to health data sharing. However, aspects of patient-centricity, such as 'Information' and 'Control,' defined for the solution, remained unexamined. Transparency as a means of patient-centricity is evaluated herein for performance validation of criteria stated under Chapter 4. This user experience validation study was designed to investigate the impact, usability, and concerns related to the aspects of patient-centricity within 'Negotiated Consent.'

GOAL
The goal of this study was to determine if the proposed 'Negotiated Consent' concept fulfils the requirements for transparency through information and control offered to individuals in health data sharing.

HYPOTHESIS
The primary assumption guiding the research was that individuals perceive the process of 'Negotiated Consent' as a more transparent approach compared to a fixed consent practice.

RESEARCH QUESTIONS
This research was conducted to examine the solution with individuals using the data sharing platform based on the following sub-research questions-

In what ways does the information offered by the data sharing platform support a transparent informed consent?

How do individuals experience a negotiation of data sharing?

METHOD

The study adopted a user experience validation approach to assess the impact, usability, and concerns associated with aspects of patient-centricity, specifically focusing on 'Information' and 'Control' within the 'Negotiated Consent' framework. The following sections provide an in-depth exploration of each stage of the research methodology-

USABILITY TEST
The evaluation of the concept served as a means to validate these features and provided insights on how to enhance the usability of the 'Negotiated Consent' concept (van Boeijen et al., 2021). Usability tests, including the Likert Scale (Likert et al., 1934) and NASA's Task Load Index (The NASA TLX Tool: Task Load Index, n.d.), were employed to gauge participants' attitudes toward the two types of consent forms and compare their workload between them.

PROTOTYPE
The prototype utilized for the user experience test was developed to incorporate features of interaction expected during the user experience. A paper-based, interactive prototype of 'Negotiated Consent' was developed, encompassing four stages: INFORM-NEGOTIATE-REVIEW-CONSENT. Each stage incorporated features based on the proposed solution, with all research information displayed from the internet (see Fig X).

Similarly, a prototype for Fixed Consent was developed, representing the current consent process, consisting of information about the research study, associated risks, and data to be used (Gupta, 2013). This prototype comprised

three stages: INFORM-REVIEW-CONSENT, mirroring the structure of 'Negotiated Consent.' The distinction between the two prototypes lay in the description of data usage. While Fixed Consent outlined data to be collected throughout the study, Negotiated Consent specified data to be collected within specific time frames over the study's duration, subsequently used to establish data preferences during the NEGOTIATE stage.

An observational research study (Sonographic Features of Fibroids Before and During Non-surgical Therapy and/or Expectant Management, n.d.) was chosen from the internet based on the researcher's personal experience with a medical condition. This decision was influenced by the researcher's understanding of the nature of the data used and the context in which the data is generated.

PARTICIPANT RECRUITMENT
Participants were recruited based on their relevance to the chosen research study. Due to the study's nature, only women were recruited for testing, totalling 10 participants. While prior experience with a consent form was not a prerequisite, this detail was recorded by the researcher for data analysis.

PROCEDURE

Participants were introduced to the context of utilizing the consent form within a data sharing platform as part of the research study. Each participant underwent two distinct tests, experiencing one consent practice at a time. During this process, participants engaged in a comprehensive journey through the various stages, encompassing INFORM, NEGOTIATE (if applicable), REVIEW, and finally, the formal request for CONSENT. Upon completion of the study, participants were provided with a testing

sheet containing usability questions (refer to Figure 40), administered by the researcher following each consent form.

Upon concluding the first test, participants proceeded to undertake the second consent practice. Subsequently, at the conclusion of the second test, participants were prompted to evaluate the usability of both experiences on the same testing sheet, using a distinct coloured pen. This deliberate approach aimed to delve into participants' cognitive processes, offering insights into their considerations while analyzing responses for both scenarios. This method facilitated the measurement of divergences in participants' experiences between the developed 'Negotiated Consent' concept and a conventional 'fixed' consent practice.

Throughout the procedure, participants were encouraged to vocalize their thoughts, enabling the researcher to correlate test results with participants' comments during subsequent

Figure 40 : User test sheet

analysis. Additionally, the researcher meticulously ensured an equal distribution of participants who followed the sequence of Negotiated-Fixed and vice versa. This careful consideration aimed to forestall any potential bias in results arising from participants' shared experiences when undertaking both tests in the same order.

DATA COLLECTION

Two types of data were collected: the testing sheet filled by participants and hand-written notes taken by the researcher during participant interactions. Data on the testing sheet included non-identifiable information, with participants marked as participant number P(n). Researcher notes identified each participant as P(n).

DATA ANALYSIS

Data generated on the testing sheets were processed quantitatively. Each participant's change in attitude over the two types of consent forms were analysed. The results of Likert Scale values were analysed on three levels that indicated change in attitude as LESS-EQUALLY-MORE. Two pie-charts titled INFORMED and CONTROL (see Figure 41) were made to represent the same. Through the use of pie charts, the researcher was able to quantitatively and visually represent the variation in attitudes amongst the participant group, such that initial insights were further derived.

The NASA's Task Load Index results were analysed on a scale (see Figure 42). Similar to Likert Scale analysis, the results were analysed on three levels that represents the change in experience over the type of consent forms, stated as 'Similar Experience,' 'Minimal change in experience' and 'Drastic change in experience.'
The visualised representations of the data were used to draft research findings. The notes taken by the researcher during the test were used to supplement the findings by contextualising the participant's choices.

LESS *"i like the precise information.. for health data you need that"*

EQUALLY

INFORMED?

MORE
"options gives me more information on what's happening to my data"

EQUALLY *"signing consent is enough"*

EXTREME DIFFERENCES
"not being able to choose.. ..feels pushed in my face"

CONTROL?

MORE

Figure 41 : Likert Scale Research Findings

SIMILAR EXPERIENCE   SLIGHT CHANGE IN EXPERIENCE   DRASTIC CHANGE IN EXPERIENCE

SAFE?
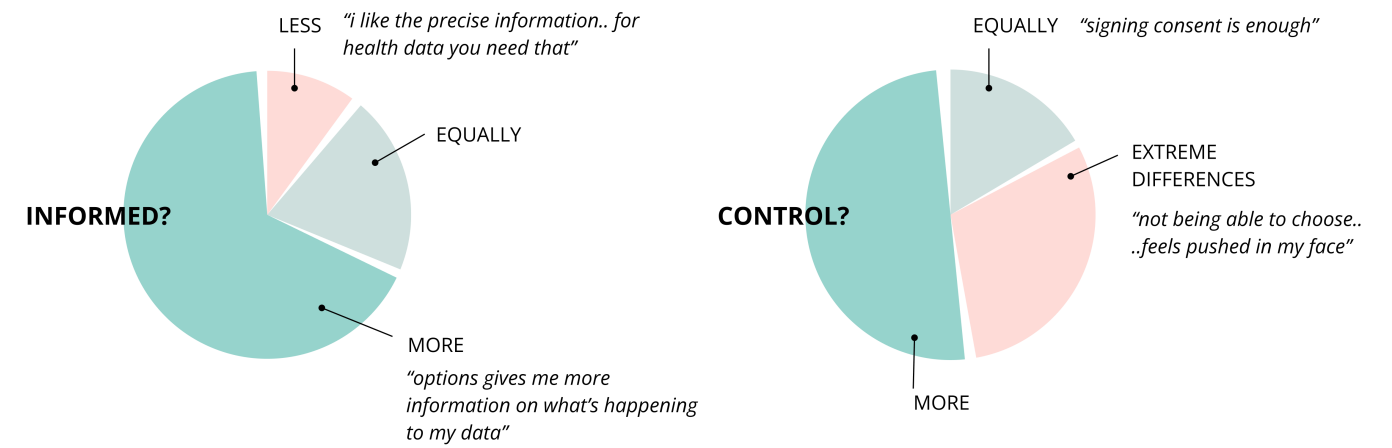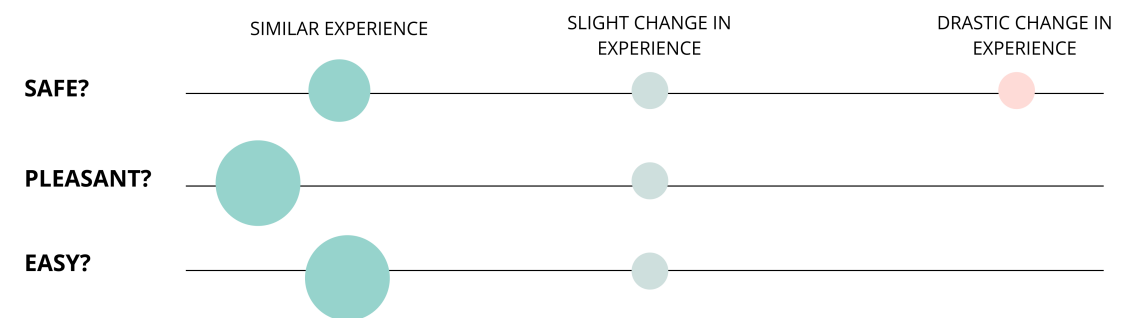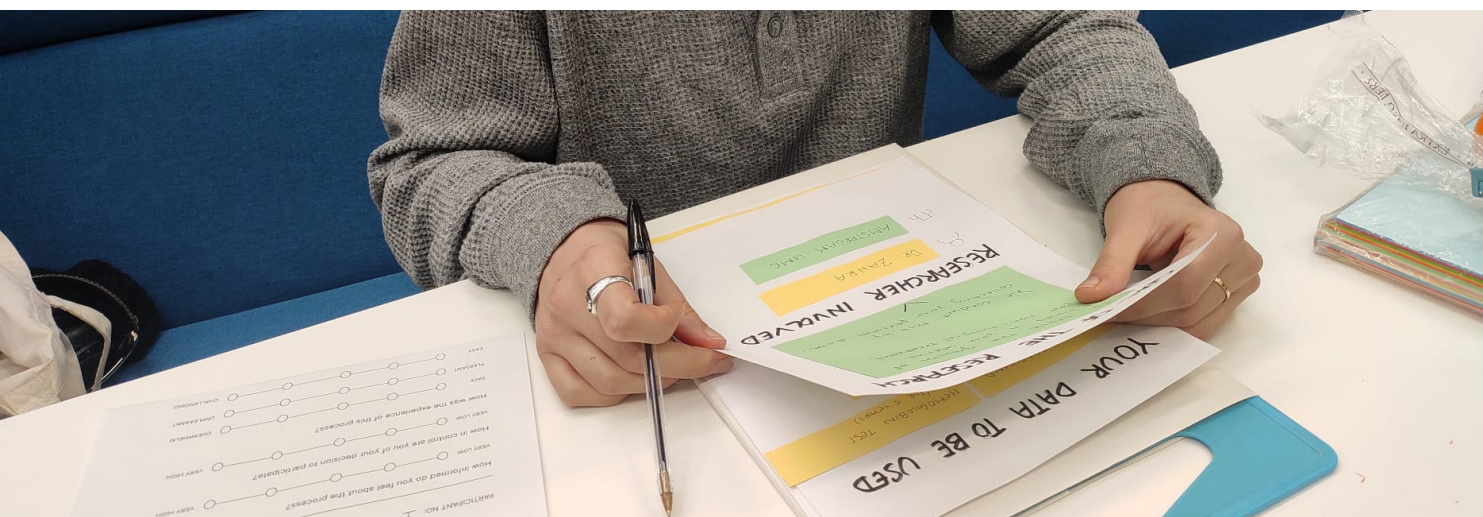
PLEASANT?

EASY?

Figure 42 : NASA TLX Scale Research Findings

RESEARCH FINDINGS

The research led to an understanding of individual variances in the use of 'Negotiated Consent' in comparison with that of 'Fixed Consent' regarding the nature of transparency proposed. They are stated herein, with quotes that illustrate the findings-

PREFERENCES LEAD TO MORE INFORMED CHOICES

"feel more in control... it's scarier because more it feel more real, but i'm more informed now..."

INFORMATION

9/10 participants were found to associate control over their participation as analogous to being more informed about their participation, and positively associated having control over their data in negotiated consent. Knowing more, being asked to think about their choices already felt like they had more agency in the process.

NUANCES IN INFORMATION

Participants were found to have a varied need for the depth, detail and nature of information provided during both processes.

INFORMATION

"not sure how they're making it secure.. as someone with a technical background, i would like to know some technical details.."

"do i need to know the risks in such detail? I already know that there's a risk of data leak.."

CONCERN OVER DATA PREFERENCES

"if i don't mind one time, why will i mind twice? it doesnt matter if i already have the data, because research findings may be affected"

PREFERENCES

3/10 participants were concerned about making choices for data preferences. While they liked the agency, they were concerned if offering less would affect the research outcomes.

IMPACT OF EXPERIENCE ON FAVOURABILITY

Most participants of the study were found to be consistent in their experience of the two forms even in comparison. 2/4 participants were found to have a change in experience, in an unfavourable manner towards either one of the consent forms. 1/10 participant was found to have an extreme difference in their experience. 9/10 participants preferred the negotiated consent towards the end, regardless of their experience over both.

"having more choices made me question my participation.. with the risks involved as mentioned, i prefer clarity in my participation"

PREFERENCES

It was found that preference for fixed consent was a suggestion for clear information, which they found necessary for health data.

BURDEN OF PARTICIPATION

All participants evaluated the data to be given based on the burden associated with it.

For the two datasets to be collected, they were found indifferent (4/10) to how much of their data they offered. However, for one data set, wherein they had to do more work, they were found to be unfavourable of offering more.

BURDEN

"if i had to go to give my data then maybe i would offer less"

7.2 STAKEHOLDER COLLABORATION TEST

DISCUSSION

The exploration of health system with 'Negotiated Consent' helped frame the benefits and challenges such that opportunities can be further developed.

Transparency, when defined as a means of being informed and having control over participation, further demonstrates the negotiation within the concept. It was found that the demonstration of Negotiated Consent under Solution offers individuals with more information and control over their participation preferences within the informed consent. Individuals when informed about the research, data processing involved, risks and data variations in participation, were able to make conscious choices regarding their consent to participate.

The experience of individuals in consenting through negotiation (when evaluated in comparison with fixed consent), revealed that most individuals experienced the process similarly. Herein, positive experiences remained positive for most individuals. However, upon having a negative association to the process led to further negative experiences through the Negotiated Consent.

This understanding demonstrates the importance of tailoring the implementation of a patient-centric approach, that can be further supported by leveraging the technological capabilities embedded in the 'Negotiated Consent' concept.

LIMITATIONS
Due to a lack of researcher resources, limited participants were evaluated during the study and the research findings may not be revealing of population. Moreover, due to the choice of study, the results could be biased and require a standard study that is gender inclusive.

INTRODUCTION

Literature research and stakeholder interviews showed the dynamics of health data sharing in the current landscape for clinical research. The intention of conducting this study is to validate criteria of Collaboration, Transparency and Responsibility under Performance through an understanding of impact on stakeholders, whilst sharing their possible concerns and finding out new insights for future development of this ecosystem.

GOAL
To determine if the interactions under the health data system designed fulfil the needs of the stakeholders involved, and how the developed solution of 'Negotiated Consent' impacts them.

HYPOTHESIS
The primary assumption of the research was that stakeholders are interested in using the service provided by the platform for participating in the eco-system of health data sharing for clinical research.

RESEARCH QUESTIONS
This research was conducted to examine the solution with stakeholders and answer the following sub-questions-

> *In what ways does the current framework meet the requirements of individual stakeholders?*
>
> *What are the challenges experienced in the stakeholder collaboration?*

METHOD

To validate the fulfilment of stakeholder requirements in health data sharing for clinical research, a role-play based test was designed. One of the major advantages of role-playing is that it involves a more real interaction with the outcome compared to using storyboards/scenarios (van Boeijen et al, 2021)

PARTICIPANT RECRUITMENT
The participants for the study were stakeholders previously contacted for interviews in earlier research stages, including a data sharing platform member, clinical researcher, and patient. Member of patient organization and doctor were not available during the course of this study. Patient representative was chosen based on personal experience of the individual with healthcare institutions and frustrations with their chronic conditions.

PROCEDURE

The virtual test was conducted through Microsoft Teams to accommodate remote participation. Participants were provided with a Character Sheet (see Figure 43), symbolizing their primary values and motivations within the given system. This approach, chosen due to the absence of a physical prototype in the developed health system, proved feasible within the constraints of time. The procedural steps taken by the researcher are described as follows:
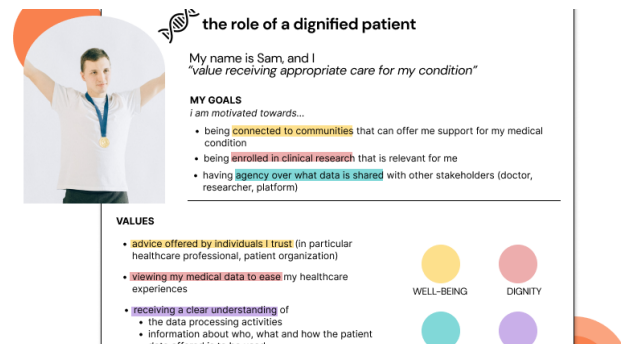
Figure 43 : Character Sheet

COMMUNITY FINDING

FUNCTIONS

Participation in POs   P   D

Patient matching   P   D

Informing about POs   P   D   O

DATA   PATIENT
SHARING   ORGANIZATION

Figure 44 : Example of demonstrated framework

RESEARCH FINDINGS

The research findings were drafted through the lens of stakeholders and are stated herein, with quotes that illustrate the findings-

PATIENT

"for your data to be processed so that you can be matched with the patient organization I think I would want to know like what data is necessarily being processed?"

DATA PROCESSED INFORMATION

At stages of requesting consent for participation during stages of the system, participant was found to require more data about what data that was going to be processed for their consent purposes.

"what options there are for why I can sign up for instead of you just signing me up for a list of stuff.. Or at least showing me that there is an option to do that eventually "

SEARCHABLE COMMUNITIES

Participant wishes for communities to be searchable rather than directly connected, such that they can choose to enroll in what they wish.

DATA SHARING PLATFORM

"how easy is it to sign up and is the process already clear and transparent from the memory that she gets offered the option...So how to get process as simple as possible and easy to understand"

ONBOARDING OF PARTICIPANTS

The member of a data sharing platform indicates the sign-up as the most crucial stage of intervention for information clarity and use, and anticipates ease of use as a path to successful onboarding.

INITIATION
The researcher commenced by outlining the study's objectives and detailing the tasks to be undertaken. Each participant received a personalized Character Sheet, crafted by the researcher to define their role in the test. The researcher clarified the utility of the Character Sheet and explained the sequential procedure for the test.

SENSITIZATION
Despite the participants' professional familiarity with their assigned roles, a sensitization process was incorporated to mitigate preconceived opinions about the system. This aimed to immerse participants in the perspective of the defined characters. The researcher initiated the session with a narrative grounded in the health system framework.

FRAMEWORK EXPLORATION
The researcher systematically presented each component of the framework. Stakeholders associated with each component were invoked to enact their roles and respond to queries (if requested), guided by their respective Character Sheets. This approach facilitated a comprehensive exploration of each framework component, with relevant stakeholders participating as required.
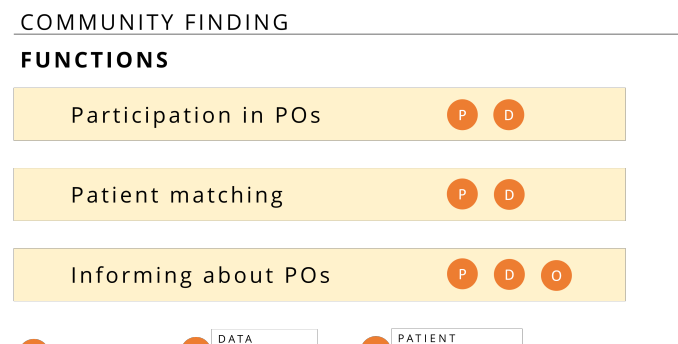
Upon concluding the session, the researcher facilitated a reflective segment, allowing participants to articulate their experiences. The researcher directed the conversation with targeted questions aligning with the research goals. The session reached its conclusion once every participant had shared their insights successfully.

DATA COLLECTION
Data was collected through a live transcript generated during the session. This transcript was anonymized and individuals were tagged based on the roles played during the session.

DATA ANALYSIS
The qualitative nature of the focus group study led to an analysis of the transcripts through themes within the transcript. The researcher used the transcript to identify the statements made by stakeholders in relation to the research aim and research question of the study. The statements made by the stakeholders were compared to the values and motivations of the Character Sheet used during the study, and research findings were drafted.

"we want to make sure that the researcher side of this platform also can get what they want ..the time-wise and all the change of processes etc. "

DATA FOR RESEARCH

Data sharing platform member finds it important to meet the current needs of researchers with regards to efficient processes and ease of workflows.

CLINICAL RESEARCHER

The researcher anticipates concerns with data inconsistencies with the solution of 'Negotiated Consent' implemented in the health system, and is eager to examine the effect of the variability in consent on participation.

DATA INCONSISTENCIES

"I can tell that it will be challenging if everyone is consenting to sharing different data points for sure that it's real-world data.. but that happens a lot in life already. "

The researcher addresses the need for data to be trackable across different studies in case of data repurposing, and suggests being informed on the population percentage of data being re-used.

DATA TRACKABILITY

"it's important to know if when your study we're also part of another study. So I would say that's also a challenge or something you think about of how do you somehow alert"

DISCUSSION

The exploration of health system with 'Negotiated Consent' helped frame the benefits and challenges such that opportunities can be further developed.

It was found that all participants in their functional roles were positively supportive of the impact of Negotiated Consent in the process of health data sharing for clinical research despite the challenges involved. This helped to understand the positioning of the same in the ecosystem, such that impact can be further investigated. The challenges analysed by the stakeholders further indicates a push for development of Negotiated Concept within the health data sharing ecosystem.

In answering the research question earlier proposed regarding the patient-centric approach through a negotiated consent, it is understood that implementation requires a collaborative understanding of information clarity, researcher responsibility and technological abilities of data sharing platforms.

This study further illustrates the importance of addressing the diverse privacy needs for stakeholders by leveraging technological capabilities for a successful implementation of transparency in the 'Negotiated Consent' concept.

LIMITATIONS
The absence of a doctor and patient organization in the study is a limitation of the study. Given the role of the doctor in offering access to health data for the patient, an evaluation of their perception of the process may lead to richer insights of the scope. Moreover, patient organization's advocacy towards patient care and research can further necessitate different functions that are found to be useful and concerning for them. The absence of both during this study can influence the results to be funnelled towards an understanding of the pathway of patient-data sharing platform-clinical research, rather than the eco-system it actually is.

8. CONCLUSION

Exploring implications and wrapping up

The final chapter concludes with an examination of the impact of the project in the society. Firstly, it discusses implications of use of Negotiated Consent in the Business Model. Subsequently it examines the use of technology in the Technological Landscape and the proposed concept's Impact on Research. Secondly, it describes the outcomes of research in light of the overall research questions. Herein, the scope of Negotiated Consent along with Limitations and Future Recommendations are provided. Lastly, the project concludes with a Personal Reflection wherein I reflect on my learnings in the past few months.

8.1 IMPLICATIONS

In this section, we will examine the implications of the proposed solution on business models, technological landscape and impact on research in society.

BUSINESS MODEL IN COLLABORATIVE RELATIONSHIPS

The implications of a data sharing platform's services are analysed by examining financial contributors such as the pharmaceutical industry, healthcare institutions, patient organizations, and individual users. It questions who pays for the platform and how it influences accessibility, ethical considerations, and societal outcomes.

PHARMACEUTICAL INDUSTRY
The pharmaceutical industry may fund the data sharing platform through participant recruitment budgets for clinical research. Collaboration however, could introduce bias of unethical practices in research.

HEALTHCARE INSTITUTIONS
Healthcare institutions may financially support the platform for its versatile data processing capabilities for single-centre studies carried out at the institution.

PATIENT ORGANIZATIONS
Patient organizations may financially collaborate with the platform for outreach purposes, leading to a boost in social engagement with patients.

INDIVIDUALS USING SERVICE
Individuals accessing the platform's health data services may face payment requirements, with payment nature varying.

While this thesis explores the systemic nature of health data sharing through a data sharing platform, a detailed study on a supporting business model is essential for further development of a patient-centric approach towards the same.

DE-CENTRALIZED DATA IN TECHNOLOGICAL LANDSCAPE

This exploration examines the current state and potential evolution of these platforms, emphasizing their value, computational strengths, and the imperative need for partnerships.

COMPUTATIONAL CAPABILITIES
Outsourcing certain features such as data standardization and addressing server downtime risks ensure a robust infrastructure. Amidst current non-interoperability, these platforms offer a glimpse into the future—tapping into existing datasets for anonymized, encrypted patient data. Redefining consent as dynamic adds a flexible dimension to its nature across diverse studies. Aligning with EHDS and FAIR practices, the data sharing platform can look forward to data standardization for increased re-usability and decreased computational requirements.

NEED FOR PARTNERSHIPS
Multi-center data processing necessitates collaborations with localized healthcare institutions, which the platform may outsource through partnerships. Moreover, given the collaborative nature of use of ML, researchers' input and feedback on uncertainties contribute to the platform's continual improvement.

Navigating the terrain of building computational strength requires strategic partnerships, innovative approaches, which may improve viability in offering meaningful changes in the domain of clinical research.

NEGOTIATED CONSENT IN IMPACT ON RESEARCH

This exploration focuses on prolonging the use of data through refined consent processes, addressing research goals, preferences

ALIGNMENT WITH REGULATIONS
Within GDPR constraints, the solution promotes ethical data use, offering individuals agency in consenting to data use. Its patient-centric approach fosters awareness of data value, and negotiated research participation further motivates individuals in clinical research, as promoted by DGA.

LONG TERM AND RE-USE OF DATA
Through Negotiated Consent offered on a data sharing platform, changes in research and use of data can be incorporated in a seamless and transparent manner. Long term research studies may further lead to offering clarification in information regarding use of data, unlike current broad consent forms (Maloy & Bass, 2020.)

RESEARCHER RESPONSIBILITIES
To incorporate Negotiated Consent, the current clinical researcher would need to assume new responsibilities. Whilst they require being clear on data requirements for research, a new understanding of data quality for research needs to be defined, incorporating individual control over datasets.

In the pursuit of ethical, viable, and feasible long-term data use, the platform pioneers advancements in consent processes, aligning with evolving research goals and individual preferences. By mitigating information overload, it aims to enhance the overall experience, ensuring meaningful contributions to the research domain.

8.2 DISCUSSION

REFLECTION ON RESEARCH QUESTIONS

The potential advantages of health data sharing for clinical research are vast, yet it is accompanied by significant privacy and legislative concerns that requires consideration. This project examines the privacy and stakeholder landscape of health data sharing through an evaluation of interaction with patients. It solves for a lack of transparent data use and research practices for health data sharing on a platform. The study develops a patient-centric approach towards transparency, through an examination of change in responsibilities of stakeholders for a transparent approach. The examination of results is carried out through reflecting on the research questions herein.

What are the privacy dynamics and challenges within the European health data sharing ecosystem for clinical research, considering the diverse stakeholders (patients, doctors, clinical researchers in the industry, and data sharing platforms), technological advancements (ML and encryption), and regulatory frameworks (GDPR, DGA and EHDS)?

The study discovers the nature of patient engagement and interaction with stakeholders such as Doctors, Patient Organizations, Data Sharing Platforms, Clinical Researcher and Pharmaceutical Industry. It was found that interactions with care-providers such as Doctors and Patient Organizations were found to have an impact on trust in data sharing. Moreover,

it was found that these stakeholders were often advocates for ongoing research studies.

Furthermore, the impact of data sharing for clinical research was explored, wherein lack of information regarding research practices impacted the trust towards data sharing. Individual decision to share data was examined through their use of data sharing platforms, wherein it was discovered that individuals had varied need for privacy on such platforms.

Within the technological and legal landscape, it was found that use of health data is challenged by GDPR restrictions as well as individual identification risks. This is supported through the use of Federated Learning wherein privacy-enhancing technology such as Homomorphic Encryption can support multi-centre studies on encrypted data. Moreover, EHR data was identified as unstructured, for which the use of Natural Language Processing is found to offer insights into data for patient matching. This further supports the use of ML in clinical research for the purposes of data processing, which was found to be a cumbersome task for researchers. In light of EHDS, the processing capabilities of data sharing platforms were found to align with data re-use purposes for clinical research.

How can a patient-centric approach be implemented within a collaboration among stakeholders in the health data-sharing ecosystem, leveraging technological capabilities to meet privacy needs and transparency whilst upholding diverse collaboration needs?

The study results in a concept termed 'Negotiated Consent', wherein a transparent approach towards individual participation towards clinical research is developed.

It was found that transparency was defined as a need for information and preferences in enrolment. Moreover, a transparent approach is found to be supported by mutual responsibility carried out between data sharing platforms and researchers. Herein, researchers take responsibility for the information on data use and affiliation clarifications. By defining clearer terms on data to be used in research, individuals were found to be informed and in control of their participation in clinical research.

Furthermore, the responsibility of data sharing platforms in a transparent practice is found to be supported through taking accountability towards the use of ML algorithms for data processing.

Herein, it is found that examination of transparency and trust in health data sharing for clinical research is intertwined with the mutual responsibility taken by each stakeholder. Therefore, to offer a patient-centric approach, it is found that the examination of these three factors is not carried out separately.

SCOPE OF NEGOTIATED CONSENT

Negotiated Consent herein developed, builds on the study of dynamic consent in participant recruitment for clinical research (Mascalzoni et al., 2022), wherein participants can revise choices over time. Herein, the concept demonstrates the same nature of agency, and has been conceptualized within the moment of patient consent for research participation. This concept is further inspired by the depth of information provided within study-specific consent (Mikkelsen et al., 2019), whilst acknowledging the lack of full awareness regarding outcomes at the onset of clinical research studies. The ethical acceptability of information herein lies in a shared responsibility towards providing information regarding data use and control. Furthermore, it adds to the literature on attitudes towards data donation for research (Richter et al., 2021), wherein individuals were found to request control of use and protection of health data by independent bodies in case of commercial use. Through the offering of Negotiated Contract, the data sharing platform can become the independent body, whilst following regulation norms.

Moreover, in consideration of the motivation to donate, factors such as concern for others and self-benefit were identified under a study by Skatova and Goulding (2019.) Herein, Negotiated Consent is a means of motivating individuals to benefit from participation through their concerns for both-the other and themselves.

LIMITATIONS AND FUTURE RECOMMENDATIONS

While this study contributes to an understanding of the privacy and stakeholder dynamics for health data sharing, limitations of the study are herein acknowledged.

STAKEHOLDER PERSPECTIVES

The results are derived from a specific set of stakeholder perspectives, potentially limiting the inclusivity of diverse collaboration approaches within the health data sharing ecosystem. Recognizing the diversity in stakeholders' approaches could be instrumental in defining distinct partnerships within the industry. It is recommended that future research distinguishes between these approaches to shed light on unique collaborative models and potential early adopter strategies.

PATIENT-CENTRICITY

The definition of 'patient' employed in this study is somewhat singular, overlooking the dynamic and multifaceted reality of patients' lives. It is suggested that future research undertakes a more nuanced examination of patient personas, considering factors such as disease-specific patient groups, race, culture, and personal history with diseases. Moreover, a patient's motivation towards participation in clinical research may vary and needs a further examination. This nuanced understanding will enrich the design of patient-centric approaches, ensuring they are tailored to the diverse needs and capabilities of different patient populations.

USE OF FEDERATED LEARNING

Regarding the incorporation of Federated Learning, the proposed approach is gradual. As the industry moves toward connected data generation and sharing practices, data sharing platforms can unlock the potential of existing datasets. To accelerate this process, it is recommended that platforms leverage partnerships to establish the necessary technological infrastructure and data access. Platforms, by becoming sources of patient connectivity, can differentiate themselves in competitive markets while fostering technological competency growth.

In conclusion, acknowledging these limitations opens avenues for future research that can address these gaps, promoting inclusivity, personalized approaches, and strategic collaborations within the evolving landscape of health data sharing for clinical research.

8.3 PERSONAL REFLECTION

My personal goals for this project were simple, and yet my journey was not. Through this project, I was able to identify some of my beliefs, strengths and weaknesses. It is through the same lens that I will be carrying out this reflection.

I had a strong appeal towards the systemic nature of this project for its complexity and inter-disciplinarity, which has in the past often offered me opportunities for bringing unique perspectives. Within the field of healthcare, I recognize an imperative need for connectivity amongst many experts, and the same was met with my belief towards collaboration. I thrive and enjoy being in a connected environment, because it offers me a chance to learn plenty and find a way to make sense of it together. I learned that contrary to the glamour of systemic projects, I had to be highly pragmatic with my time. For example, I acknowledged learning about the regulatory practices at a 'workable knowledge level' that allowed me time to implement my learnings. As a result, I was able to put together some things that existed sparsely apart- a blender of sorts.

The project came with many learning opportunities. There were new skills and knowledge I had to gain for my vision to become achievable, for which I was comfortable and curious. Some skills were easier to learn and implement, such as creative facilitation. However, I was often challenged with domains outside of my comfort zone. An example is the implementation of machine learning in the project, which was hard to grasp in a short duration given the many concepts and jargons. Making the effort gave me great joy and pride in my capabilities. Through the challenge, I was able to implement an integrated product designer's approach towards this project. I realize, I re-learned a lot of my coursework through a new lens!

Given the nature of the graduation project, I was self-motivated towards defining the nature of research, design activities and design directions. I quite enjoyed taking ownership of my project! I had a strong support of my supervisor team for the same. Upon looking back however, some design moments in the project were crucial for supervisor intervention. Examples of moments such as sharp turning points of divergence to convergence were a struggle, because for me, a switch takes time to process. My supervisors intermittently played a role in helping me spot the next steps, which came as right reminders, at the right time.

Lastly, a short note on stress.
During the course of the project, I experienced a lot of stress. While part of it was natural due to the project, the effects of stress due to my unstable health were often a cause of concern in relation to the project. The sharp focus that I enjoyed bringing into the project was often affected due to the same. I realised that it impacted my time management the most, for which I requested more time and was always supported by my supervisor team. In complete honesty, I must say that it was hard to juggle my health and the project together. However, having seen the project through to its completion in a healthy and happy way, gives me hope in my ability to take it as it comes. Many thanks to my supervisor team for the same.

Man weet niet wat man mist.

THE END.

REFERENCES

As

Agbo, C. C., Mahmoud, Q. H., & Eklund, J. (2019). Blockchain Technology in Healthcare: A Systematic review. Healthcare, 7(2), 56. https://doi.org/10.3390/healthcare7020056

Askin, S., Burkhalter, D., Calado, G., & Dakrouni, S. E. (2023). Artificial Intelligence Applied to clinical trials: opportunities and challenges. Health and Technology, 13(2), 203–213. https://doi.org/10.1007/s12553-023-00738-2

Bs

Beck, J. T., Rammage, M., Jackson, G. P., Preininger, A. M., Dankwa☐Mullan, I., Roebuck, M. C., Torres, A., Holtzen, H., Coverdill, S., Williamson, M., Chau, Q., Rhee, K., & Vinegra, M. (2020). Artificial intelligence tool for optimizing eligibility screening for clinical trials in a large community cancer center. JCO Clinical Cancer Informatics, 4, 50–59. https://doi.org/10.1200/cci.19.00079

Bonomi, L., Huang, Y., & Ohno-Machado, L. (2020). Privacy challenges and research opportunities for genomic data sharing. Nature Genetics, 52(7), 646–654. https://doi.org/10.1038/s41588-020-0651-0

Business.gov.nl. (2023). Keeping and sharing medical records. business.gov.nl. https://business.gov.nl/regulation/medical-records/

Bussone, A., Kasadha, B., Stumpf, S., Durrant, A., Tariq, S., Gibbs, J., Lloyd, K., & Bird, J. (2020). Trust, identity, privacy, and security considerations for designing a peer data sharing platform between people living with HIV. Proceedings of the ACM on Human-computer Interaction, 4(CSCW2), 1–27. https://doi.org/10.1145/3415244

Cs

Carey, D. J., Fetterolf, S. N., Davis, F. D., Faucett, W. A., Kirchner, H. L., Mirshahi, U. L., Murray, M. F., Smelser, D. T., Gerhard, G. S., & Ledbetter, D. H. (2016). The Geisinger MyCode community health initiative: an electronic health record–linked biobank for precision medicine research. Genetics in Medicine, 18(9), 906–913. https://doi.org/10.1038/gim.2015.187

Chaudhari, N., Ravi, R., Gogtay, N. J., & Thatte, U. M. (2020). Recruitment and retention of the participants in clinical trials: Challenges and solutions. Perspectives in Clinical Research, 11(2), 64. https://doi.org/10.4103/picr.picr_206_19

Consent - General Data Protection Regulation (GDPR).

(2021, October 22). General Data Protection Regulation (GDPR). https://gdpr-info.eu/issues/consent/

COREON Foundation. (2023). Code of conduct for health research. NA.

Ds

Daniels, H., Jones, K., Heys, S., & Ford, D. (2021). Exploring the use of genomic and routinely collected data: narrative literature review and interview study. Journal of Medical Internet Research, 23(9), e15739. https://doi.org/10.2196/15739

Davenport, T. H., & Kalakota, R. (2019). The potential for artificial intelligence in healthcare. Future Healthcare Journal, 6(2), 94–98. https://doi.org/10.7861/futurehosp.6-2-94

De European Health Data Space wankelt op onderdeel bescherming - Zorgvisie. (2022, August 26). Zorgvisie. https://www.zorgvisie.nl/blog/de-european-health-data-space-wankelt-op-onderdeel-bescherming/

Definitive Healthcare. (n.d.). Most common types of healthcare data breaches. Definitive Healthcare. https://www.definitivehc.com/resources/healthcare-insights/most-common-healthcare-data-breaches

Dharssi, S., Wong☐Rieger, D., Harold, M., & Terry, S. F. (2017). Review of 11 national policies for rare diseases in the context of key patient needs. Orphanet Journal of Rare Diseases, 12(1). https://doi.org/10.1186/s13023-017-0618-0

Es

European Data Governance Act. (2023, October 3). Shaping Europe's Digital Future. https://digital-strategy.ec.europa.eu/en/policies/data-governance-act

European Parliament's Policy Department for Citizens' Rights and Constitutional Affairs. (2015). A comparison between US and EU data protection legislation for law enforcement purposes. n.a.

Exams, H. (2023, February 2). Data Processing Agreements: What is a DPA? HIPAA Exams. https://www.hipaaexams.com/blog/data-processing-agreements

Fs

Filkins, B. L. (2016). Privacy and security in the era of digital health: what should translational researchers know and do about it? PubMed Central (PMC). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4859641/

Friedman, B., Hendry, D. G., & Borning, A. (2017). A survey

of value sensitive design methods. Foundations and Trends in Human-computer Interaction, 11(2), 63–125. https://doi.org/10.1561/1100000015

Friedman, B., Kahn, P. H., Borning, A., & Huldtgren, A. (2013). Value sensitive design and information systems. In Philosophy of engineering and technology (pp. 55–95). https://doi.org/10.1007/978-94-007-7844-3_4

Gs

Gates, A., Gates, M., Sim, S., Elliott, S. A., Pillay, J., & Hartling, L. (2021). Creating efficiencies in the extraction of data from randomized trials: a prospective evaluation of a machine learning and text mining tool. BMC Medical Research Methodology, 21(1). https://doi.org/10.1186/s12874-021-01354-2

Gianfrancesco, M., Tamang, S., Yazdany, J., & Schmajuk, G. (2018). Potential biases in machine learning algorithms using electronic health record data. JAMA Internal Medicine, 178(11), 1544. https://doi.org/10.1001/jamainternmed.2018.3763

Gupta, U. C. (2013). Informed consent in clinical research: Revisiting few concepts and areas. Perspectives in Clinical Research, 4(1), 26. https://doi.org/10.4103/2229-3485.106373

Hs

Hamburg, M. A., & Collins, F. S. (2010). The path to personalized medicine. The New England Journal of Medicine, 363(4), 301–304. https://doi.org/10.1056/nejmp1006304

Harris, P. A., Taylor, R. J., Thielke, R. J., Payne, J. F., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. Journal of Biomedical Informatics, 42(2), 377–381. https://doi.org/10.1016/j.jbi.2008.08.010

Henin, C., & Métayer, D. L. (2021). Beyond explainability: justifiability and contestability of algorithmic decision systems. AI & SOCIETY, 37(4), 1397–1410. https://doi.org/10.1007/s00146-021-01251-8

Hoepman, J. (2018). Privacy Design Strategies (The Little Blue Book). n.a. https://repository.ubn.ru.nl/handle/2066/195397

Hong, S. J., Drake, B. F., Goodman, M. S., & Kaphingst, K. A. (2019). Race, trust in doctors, privacy concerns, and consent preferences for biobanks. Health Communication, 35(10), 1219–1228. https://doi.org/10.1080/10410236.2019.1623644

Is

Innopay. (2022). Everything Transaction. n.a.

Ks

Kreimeyer, K., Foster, M., Pandey, A., Arya, N., Halford, G., Jones, S. F., Forshee, R. A., Walderhaug, M., & Botsis, T. (2017). Natural language processing systems for capturing and standardizing unstructured clinical information: A systematic review. Journal of Biomedical Informatics, 73, 14–29. https://doi.org/10.1016/j.jbi.2017.07.012

Kulkarni, N. (2023). The language of reflection: A Linguistic Exploration of Reflection Written by Design Students. TU Delft Repository, 70. https://repository.tudelft.nl/islandora/object/uuid%3A45a3bd5f-6861-4ca4-970f-be20294e9088?collection=education

Kush, R., Warzel, D., Kush, M., Sherman, A., Navarro, E. A., Fitzmartin, R. D., Pétavy, F., Gálvez, J. L. V., Becnel, L. B., Zhou, F. L., Harmon, N., Jaúregui, B., Jackson, T., & Hudson, L. D. (2020). FAIR data sharing: The roles of common data elements and harmonization. Journal of Biomedical Informatics, 107, 103421. https://doi.org/10.1016/j.jbi.2020.103421

Ls

Leslie, D. (2019). Understanding artificial intelligence ethics and safety. arXiv (Cornell University). https://doi.org/10.5281/zenodo.3240529

Li, I., Pan, J. L., Goldwasser, J., Verma, N., Wong, W. P., Nuzumlali, M. Y., Rosand, B., Li, Y., Zhang, M., Chang, D. D., Taylor, R. A., Krumholz, H. M., & Radev, D. (2022). Neural Natural Language Processing for unstructured data in electronic health records: A review. Computer Science Review, 46, 100511. https://doi.org/10.1016/j.cosrev.2022.100511

Likert, R., Roslow, S., & Murphy, G. (1934). A simple and reliable method of scoring the Thurstone attitude scales. The Journal of Social Psychology, 5(2), 228–238. https://doi.org/10.1080/00224545.1934.9919450

Lorkowski, J., & Pokorski, M. (2022). Medical Records: a historical narrative. Biomedicines, 10(10), 2594. https://doi.org/10.3390/biomedicines10102594

Ms

Maloy, J. W., & Bass, P. F. (2020). Understanding broad consent. The Ochsner Journal, 20(1), 81–86. https://doi.org/10.31486/toj.19.0088

McMahan, B. (2017, April 10). Communication-Efficient Learning of Deep Networks from Decentralized Data. PMLR. https://proceedings.mlr.press/v54/mcmahan17a.html

mdGroup. (2023). Why Do Patients Drop Out Of Clinical Trials? Mdgroup. https://mdgroup.com/blog/why-do-patients-drop-out-of-clinical-trials/

Mikkelsen, R. B., Gjerris, M., Waldemar, G., & Sandøe, P. (2019). Broad consent for biobanks is best – provided it is also deep. BMC Medical Ethics, 20(1). https://doi.org/10.1186/s12910-019-0414-6

Munjal, K., & Bhatia, R. (2022). A systematic review of homomorphic encryption and its contributions in healthcare industry. Complex & Intelligent Systems, 9(4), 3759–3786. https://doi.org/10.1007/s40747-022-00756-z

MyData Global. (2017). Understanding MyData Operators. n.a.

Ns

Naik, N., Hameed, B. M. Z., Shetty, D. K., Swain, D., Shah, M., Paul, R., Aggarwal, K., Ibrahim, S., Patil, V., Smriti, K., Shetty, S., Prasad, B., Chłosta, P., & Somani, B. (2022). Legal and ethical consideration in artificial intelligence in healthcare: Who takes responsibility? Frontiers in Surgery, 9. https://doi.org/10.3389/fsurg.2022.862322

Nebula Genomics. (2018). Blockchain-enabled genomic data sharing and analysis platform. n.a.

Os

Obermeyer, Z., & Emanuel, E. J. (2016). Predicting the future — big data, machine learning, and clinical medicine. The New England Journal of Medicine, 375(13), 1216–1219. https://doi.org/10.1056/nejmp1606181

OECD. (2021). Building and sustaining collaborative platforms in genomics and biobanks for health innovation. n.a.

Oh, W., & Nadkarni, G. N. (2023). Federated Learning in Health care Using Structured Medical Data. Advances in Kidney Disease and Health, 30(1), 4–16. https://doi.org/10.1053/j.akdh.2022.11.007

Ortega, A. G., Bourgeois, J., Hutiri, W. T., & Kortuem, G. (2023). Beyond data transactions: a framework for meaningfully informed data donation. AI & SOCIETY. https://doi.org/10.1007/s00146-023-01755-5

Ps

Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P.,

Duan, N., & Hoagwood, K. (2013). Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. Administration and Policy in Mental Health and Mental Health Services Research, 42(5), 533–544. https://doi.org/10.1007/s10488-013-0528-y

Patiëntenfederatie Nederland. (n.d.). Patiëntenfederatie. https://www.patientenfederatie.nl/

PatientsLikeMe. (n.d.). PatientsLikeMe. https://www.patientslikeme.com/

PHC Catalyst Alliance & Roche Nederland B.V. (2020). Towards a Data Sharing Platform for Personalized Healthcare: Feasibility study. n.a.

Prayitno, P., Shyu, C., Putra, K. T., Chen, H., Tsai, Y., Hossain, K. S. M. T., Jiang, W., & Shae, Z. (2021). A Systematic Review of federated learning in the healthcare area: From the perspective of data properties and Applications. Applied Sciences, 11(23), 11191. https://doi.org/10.3390/app112311191

Rs

RARE Disease Facts - Global genes. (2024, January 3). Global Genes. https://globalgenes.org/rare-disease-facts/

Richter, G., Borzikowsky, C., Hoyer, B. F., Laudes, M., & Krawczak, M. (2021). Secondary research use of personal medical data: patient attitudes towards data donation. BMC Medical Ethics, 22(1). https://doi.org/10.1186/s12910-021-00728-x

Rieke, N., Hancox, J., Li, W., Milletarì, F., Roth, H. R., Albarqouni, S., Bakas, S., Galtier, M., Landman, B. A., Maier‐Hein, K. H., Ourselin, S., Sheller, M. J., Summers, R. M., Trask, A., Xu, D., Baust, M., & Cardoso, M. J. (2020). The future of digital health with federated learning. Npj Digital Medicine, 3(1). https://doi.org/10.1038/s41746-020-00323-1

Ss

Sanders, E., & Stappers, P. J. (2008). Co-creation and the new landscapes of design. CoDesign, 4(1), 5–18. https://doi.org/10.1080/15710880701875068

Scholl, S., Beuzeboc, P., & Pouillart, P. (2001). Targeting HER2 in other tumor types. Annals of Oncology, 12, S81–S87. https://doi.org/10.1093/annonc/12.suppl_1.s81

Shah, P. (2023, June 5). Informed consent. StatPearls - NCBI Bookshelf. https://www.ncbi.nlm.nih.gov/books/NBK430827/

Shah, S. M., & Khan, R. A. (2020). Secondary use of Electronic health Record: Opportunities and challenges.

IEEE Access, 8, 136947–136965. https://doi.org/10.1109/access.2020.3011099

Skatova, A., & Goulding, J. (2019). Psychology of personal data donation. PLOS ONE, 14(11), e0224240. https://doi.org/10.1371/journal.pone.0224240

Solve-RD infographic on the patient journey to diagnosis - EURORDIS. (2023, June 5). EURORDIS. https://www.eurordis.org/publications/solve-rd-infographic-on-the-patient-journey-to-diagnosis/

Sonographic features of fibroids before and during non-surgical therapy and/or expectant management. (n.d.). ctv.veeva.com. https://ctv.veeva.com/study/sonographic-features-of-fibroids-before-and-during-non-surgical-therapy-and-or-expectant-management

Sweeney, L. (2002). K-ANONYMITY: a MODEL FOR PROTECTING PRIVACY. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 10(05), 557–570. https://doi.org/10.1142/s0218488502001648

Ts

The ethical stack. (n.d.). https://ethicalstack.virteuproject.eu/ethics.html

The European Health Data Space (EHDS). (n.d.). https://www.european-health-data-space.com/

The NASA TLX Tool: Task Load Index. (n.d.). TLX @ NASA Ames - Home. https://humansystems.arc.nasa.gov/groups/tlx/

Tribune, V. (2022). ‘De vertrouwelijkheid van onze gezondheidsgegevens dreigt in gevaar te komen’ Knack. https://www.knack.be/nieuws/de-vertrouwelijkheid-van-onze-gezondheidsgegevens-dreigt-in-gevaar-te-komen/

Vs

Van Ooijen, I., & Vrabec, H. U. (2018). Does the GDPR Enhance Consumers' Control over Personal Data? An Analysis from a Behavioural Perspective. Journal of Consumer Policy, 42(1), 91–107. https://doi.org/10.1007/s10603-018-9399-7

Vassy, J. L., Ho, Y., Honerlaw, J., Cho, K., Gaziano, J. M., Wilson, P. W., & Gagnon, D. (2018). Yield and bias in defining a cohort study baseline from electronic health record data. Journal of Biomedical Informatics, 78, 54–59. https://doi.org/10.1016/j.jbi.2017.12.017

Vos, J. F., Boonstra, A., Kooistra, A., Seelen, M. A., & Van Offenbeek, M. (2020). The influence of electronic health record use on collaboration among medical specialties. BMC Health Services Research, 20(1). https://doi.

org/10.1186/s12913-020-05542-6

Vuković, J., Ivanković, D., Habl, C., & Dimnjaković, J. (2022). Enablers and barriers to the secondary use of health data in Europe: general data protection regulation perspective. Archives of Public Health, 80(1). https://doi.org/10.1186/s13690-022-00866-7

Ws

Weber, G. M., Adams, W., Bernstam, E. V., Bickel, J., Fox, K., Marsolo, K., Raghavan, V., Turchin, A., Zhou, X., Murphy, S. N., & Mandl, K. D. (2017). Biases introduced by filtering electronic health records for patients with "complete data." Journal of the American Medical Informatics Association, 24(6), 1134–1141. https://doi.org/10.1093/jamia/ocx071

Weissler, E. H., Naumann, T., Andersson, T., Ranganath, R., Elemento, O., Luo, Y., Freitag, D. F., Benoit, J., Hughes, M. C., Khan, F., Slater, P., Shameer, K., Roe, M. T., Hutchison, E., Kollins, S. H., Broedl, U. C., Zhao, M., Wong, J., Curtis, L. H., . . . Ghassemi, M. (2021). The role of machine learning in clinical research: transforming the future of evidence generation. Trials, 22(1). https://doi.org/10.1186/s13063-021-05489-x

What is an API-enabled EHR? (2022, June 21). https://www.elationhealth.com/resources/blogs/what-is-an-api-enabled-ehr

What is an electronic health record (EHR)? | HealthIT.gov. (n.d.). https://www.healthit.gov/faq/what-electronic-health-record-ehr

Wolford, B. (2023). What is GDPR, the EU's new data protection law? GDPR.eu. https://gdpr.eu/what-is-gdpr/

Ys

Yang, C. C. (2022). Explainable artificial intelligence for predictive modeling in healthcare. Journal of Healthcare Informatics Research, 6(2), 228–239. https://doi.org/10.1007/s41666-022-00114-1

Zs

Zhang, X., Xiao, C., Glass, L. M., & Sun, J. (2020). DeepEnroll: Patient-Trial Matching with Deep Embedding and Entailment Prediction. ACM Digital Library. https://doi.org/10.1145/3366423.3380181