University of Bergen
Department of Informatics

Western Norway University of Applied Sciences
Department of Computer Science, Electrical Engineering
and Mathematical Sciences

Integrating Wearable Technology for Enhanced Self-Assessment in Mental Health

---

Author: Petter Johansen

Supervisors: Svein-Ivar Lillehaug and Yngve Lamo

Western Norway University of Applied Sciences

October 2023

Abstract

Mental health is a critical aspect of overall well-being and affects a large proportion of the world's population. The first step of identifying possible issues in an individual is often using assessment method such as surveys, forms, or questionnaires. However, traditional assessment methods such as self-assessment questionnaires yield challenges and limitations like social desirability and response bias. These traditional assessment methods rely heavily on the patient's recollection of events, feelings, and current psychological state to work as intended.

The emergence of ubiquitous wearable technology shows promise that it can help mitigate the mentioned issue. These devices promise to collect data reliably and can be used to get an objective representation of the physiological state of a patient. In addition, it might lead to better response rates and a more enhanced patient experience. However, thorough testing and evaluation is needed when integrating these emerging technologies.

This thesis research the pressing need to improve these traditional assessment methods used in mental health by leveraging the potential of wearable technology. Moreover, the aim is to demonstrate how wearable technology can be integrated into self-assessment questionnaires through the development of an artifact that promotes reuse and interoperability. It consists of three general components: the questionnaire, the corresponding response, and the wearable data collection process for specific domains through digital biomarkers. The evaluation process involved a semi-structured interview, object-based evaluation experiment, and a user acceptance survey of the artifact. Based on this, our artifact poses as a viable solution and can be used as a starting point for future research in the problem domain.

Keywords: Mental Health Assessments - Wearable Technology - Design Science

Acknowledgments

I want to thank my supervisors, Yngve Lamo and Svein-Ivar Lillehaug, for the help and advice given throughout this process. I would also like to thank my family and friends for their support and understanding during a challenging time. Finally, I need to thank my partner, Hedda for her continuous love, encouragement, and patience.

Petter Johansen
October 2023

Contents

1 Introduction ***1***

1.1 Motivation 1

1.2 Problem Description 2

1.3 Research Questions 2

1.4 Research Methods 3

1.5 Terminology 3

1.6 Thesis Overview 4

2 Background ***6***

2.1 Internet-Delivered Psychological Treatment (IDPT) Systems 6

2.2 Open-Source Adaptive IDPT 6

2.3 Introduction to Mental Health Assessments 7

2.4 Role of Self-Assessment Questionnaires 8

2.5 Challenges and Limitations of Self-Assessment Questionnaires 9

2.6 Emergence of Wearable Technology 9

2.6.1 Ethical Considerations 10

2.6.2 Challenges and limitations 11

2.7 Physiological Data in Mental Health Assessments 12

2.8 Digital Biomarkers 13

2.8.1 Stress 13

2.8.2 Sleep 14

2.8.3 Physical Activity 15

2.9 Existing Solutions 15

2.9.1 CrossCheck 16

2.9.2 ExpiWell 17

3 Method and Design ***20***

3.1 Research Method 20

3.1.1 Using the Design Science Guidelines 20

3.2 Design Process 23

4 Implementation ***25***

4.1 Technology 25

4.2 System Architecture and Design Patterns 26

4.3 General components 27

4.3.1 Extending the IDPT framework. 28

4.3.2 Questionnaire 28

4.3.3 Questionnaire Response 31

4.3.4 Health Connect 35

4.3.3 Wearable Data Collection 39

4.4 Demonstrative component 44
4.4.1 Additions to the Existing Framework 44
4.4.2 Creation of the mHealth Application 47
4.4.3 User Interface and Functionalities 47
4.4.5 Data Processing and Storage 52

5 Research Evaluation **54**

5.1 Interviews 54
5.1.2 Interview Results 55

5.2 Outcome-Based Evaluation 56
5.2.1 Objectives 56
5.2.2 Participants 56
5.2.3 Data Collection and Analysis 57
5.2.4 Results 58
5.2.4 Conclusion 61

5.3 User Acceptance 62
5.3.1 Survey Results 62

6 Discussion **64**

6.1 Answering the Research Questions 64

6.2 Contributions 66

6.3 Reflections 66
6.3.1 Research Methodology and Evaluation Methods 66
6.3.2 Reflecting on the Design Process 67
6.3.3 Reflecting on the developed artifact 68

6.4 Project Limitations 68

7 Conclusion **70**

7.1 Summary 70

7.2 Future Work 71

List of Acronyms and Abbreviations **73**

Bibliography **74**

Appendix A **81**

Appendix B **86**

Appendix C **87**

Appendix D **90**

List of Figures

2.1: The adaptive IDPT system 7
2.2: Benefits of wearable technology in healthcare, derived from. 10
2.3: This figure depicts the physiological domains and their corresponding digital biomarkers used in this study..15
2.4: Crosscheck Data collection 17
2.5: Questionnaire response UI for the ExpoWell application...18
2.6: Dashboard for creating surveys and questionnaires in the EchoWIll platform. 19

4.1: Depicts the logical flow of the entities added in the IDPT framework. It follows the same flow as the other entities. 26
4.2: High-level architecture of the developed artifact...27
4.3: A specific questionnaire as represented in the database. 30
4.4: A specific questionnaireResponse resource as represented in the database. 32
4.5: A specific observation resource as represented in the database. 33
4.6: A specific input answer structure as represented in the database. 33
4.7: Platform architecture of the Health Connect platform (Android Developers 2023). 36
4.8: Representation of how a user can allow and deny measurements in Health Connect....38
4.9: Representation of the values collected for the stress data measures derived from the database. 42
4.10: Representation of the values collected for the sleep data measures derived from the database. 43
4.11: Representation of the values collected for the physical activity measures derived from the database. 43
4.12: A screenshot showing the questionnaire creation view in the IDPT framework. 45
4.13: A screenshot showing the question card in the IDPT framework. 46
4.14: Dashboard where the clinician or therapist can see all the responses for the questionnaires..46
4.15: Retrieved from Android Developers. It shows how a normal application works with Jeptack Compose. 48
4.16: Representation of the answer UI for different question types.......................................50
4.17: UI for the visualization of the collected data for each type...51
4.18: A representation of how permission for the collection of data is handled by the UI. ...53

5.1: Represents answers given by the participants in each round for question 2. 59
5.2: Line graphs depicting the difference between the two methods for question 2. 59
5.3: Represents answers given by the participants in each round for Question 5. 60
5.4: Line graphs depicting the difference between the two methods for question 5. 61

List of Listings

4.1: CraphQL schemas used for the questionnaire resource. 29
4.2: Representation of the questionnaire response class in the mHealth application 34
4.3: Representation of the questionnaire response item and meta class in the mHealth application adhering to HL7 FHIR standards. 34
4.4: Permission set up for data collection types. 37
4.5: Example of how data can be read for a specific measure with Health Connect. 39
4.6: Example of a HeartRateVariabilityRmssdRecord derived from Health Connect. 39
4.7: Represents sleep data, raw, aggregated and sleep stages for a given [SleepSessionRecord] 40
4.8: Represents data for a given Sleep Stage inside a [SleepSessionRecord]. Holds both raw and aggregated data about the individual sleep stage 40
4.9: Represents data both aggregated and raw, associated with a single exercise session. ...41
4.10: Represents Stress data, collected from [HeartRateRecord] and [HeartRateVariabilitySmmdRecord]. 41

Chapter 1

Introduction

1.1 Motivation

Mental health is a crucial aspect of overall well-being and is essential for individuals to lead fulfilling and productive lives. According to the World Health Organization (WHO 2022), approximately one in four people globally will experience a mental health disorder in their lifetime. However, mental health has a growing stigma, and individuals often have difficulty opening up about their symptoms and issues (Henderson et al. 2013). The healthcare sector commonly uses established self-assessments, like questionnaires, to monitor and address these issues. However, in today's day and age, these traditional assessment methods have their limitations (Colombo 2019). Among many, a significant challenge is response bias, which can significantly impact the results' validity (Kwak et al. 2021; Larson 2019).

Furthermore, another potential problem with these questionnaires is response burdens. A response burden can be defined as an individual's effort to fill out a questionnaire. Among the factors that affect response burden is questionnaire length and complexity, which is manifested, for example, in response rate (Rolstad et al. 2011).

According to Statista, two hundred and nineteen million people own a smartwatch in 2023 (Statista 2023). In recent times, there has been a notable increase in studies exploring the potential utilization of ubiquitous commercialized digital technologies, within the healthcare sector, including wearables. These devices can potentially gather large amounts of data about a person's physiological state. More interestingly, the research shows promise that it can fill specific gaps in the healthcare sector (Abernethy et al. 2022; Huhn et al. 2022). This emerging technology can already be used in a variety of ways, such as stress management apps that use HRV data collection through smartwatches (Jerath et al. 2023). Suppose healthcare professionals successfully integrate and use emerging technologies like wearables with tedious traditional self-assessment tools. In that case, we can reduce length and time while getting more accurate and objective data (Weiler 2023).

Naturally, with the widespread use of wearable technology worldwide, has brought about challenges and limitations that needs to be addressed. However, research suggests that utilizing wearable technology can provide a more holistic view of a person's physiological state (Aljehaili & Alomainy 2023). wearable technology has the potential to read objective digital biomarkers and thus should be leveraged for the greater good.

1.2 Problem Description

Evaluating and treating mental health issues is a tedious and complex task. Despite notable advancements in recent years, challenges persist, primarily due to the inherently subjective nature of mental health. Unlike many physical health conditions, mental health concerns are not always observable and hinge on an individual's unique psychosocial state, encompassing emotions, thoughts, and behaviours—all of which can vary significantly from person to person. The planning and implementation of treatment frequently depend on conventional self-reported assessments, typically administered through questionnaires.

Emerging, ubiquitous wearable devices have brought about exciting possibilities for improving the accuracy and effectiveness of traditional assessment tools. Utilizing these mainstream technologies shows great potential in filling gaps posed in the mental health sector (Huhn 2022). While self-assessment questionnaires have their merits, they also have limitations such as bias in responses, burden for the patient, limited data details, and dependence on patients' recollections of events. On the one hand, wearables like smartwatches and fitness trackers can offer objective real-time data on various aspects of an individual's well-being, including physical activity, sleep patterns, and other physiological indicators correlating with mental health issues (Guk 2019). On the other hand, it is tough to implement validly and reliably.

Despite the possible advantages of integrating wearable technology into traditional standardized health assessments, there is still a need to bridge the gap between these devices and conventional evaluation methods. This involves identifying which types of sensor data accurately correlates with mental health issues, understanding the limitations of self-assessment questionnaires, and developing effective strategies for integrating ubiquitous wearable technology into the assessment process. Research addresses that data gathered in a passive way through the likes of wearables can be a possible solution to the limitations posed on traditional assessment methods (Hart 2022).

Moreover, the main objective of this thesis is to create and deploy an mHealth application that combines wearable data collection with self-assessment questionnaires and utilizes wearable technology in the assessment process. Standardized data formats and an open-source IDPT framework are used to facilitate the creation and view of questionnaires and responses. The aim is to develop a digital tool for mental health assessment that can show proof of mitigating the mentioned challenges and limitations through an artifact.

1.3 Research Questions

This thesis will explore research on integrating wearable technology into self-assessment questionnaires in an adaptive IDPT and creating an mHealth application through the development of an artifact. The primary purpose of the thesis focuses on the development, design, and evaluation of an artifact that integrates wearable technology into self-assessment questionnaires through an IDPT and an mHealth application. Specifically, the following research questions will be addressed:

RQ1 How can questionnaires be implemented to facilitate for the use of wearable data collection in the assessment process?

RQ2 How can the integration of wearable technology with self-assessment questionnaires be implemented to enhance the self-assessment process?

RQ3 Building on the result from the last two research questions, how can a system be designed to simplify and combat common challenges of traditional self-assessments for therapists and patients?

1.4 Research Methods

For this thesis, design science is chosen as the preferred research method. Hevner et al. (2004) state that this methodology's aim is to contribute to the knowledge base for the specified problem domain by developing and designing an artifact. The design science methodology is described in Section 3.1.

In the project, an artifact is developed as a possible solution or a starting point for integrating wearable technology with traditional self-assessment questionnaires. The artifact consists of two distinct parts; first, an extension of the IDPT framework discussed in Section 2.1, where questionnaires can be created, and the corresponding responses can be viewed and stored. Furthermore, a mHealth application is developed as a minimum viable product that poses as our demonstrative component. Furthermore, we intend to address how wearable technology and passive data collection can be integrated in an innovative way. Both qualitative and quantitative methods are used to evaluate the artifact. See Chapter 5 for more information.

1.5 Terminology

This section address terminology used throughout the thesis.

Self-Assessment Questionnaire A *self-assessment questionnaire* is commonly used in psychology to assess a person's psychological state. It consists of the combined word *self-assessment,* which can gather information about individuals' feelings, thoughts, behaviors, and experiences without any outside noise. A *questionnaire* is often used in epidemiological surveys and mental health assessments to assess knowledge and information about a particular topic of interest. It usually consists of a set number of predefined questions based on the aim of the research or study (Sharma 2022).

IDPT stands for Internet-Delivered Psychological Treatments and can be described as any online psychological treatment (Andersson 2016). A specific IDPT framework is used in this thesis to facilitate the use of questionnaires and responses. See Section 2.1 for more information.

Wearable can be defined as “devices that can be worn or mated with human skin to continuously and closely monitor an individual's activities, without interrupting or limiting the user's motions” (Haghi et al. 2017). It is a device that can measure and collect data through different sensors attached to an individual’s body. It ranges from phones to smartwatches.

mHealth Application consists of the two words *mHealth* and *application. mHealth is an* acronym for mobile health and describes using mobile devices in public health and medicine to improve care and minimize costs (El-Sherif 2022*).* So, the word *mHealth application* refers to all the applications used to do what is described above. In our thesis, a *mHealth* ap*plication* is developed to show how wearables integrated with self-report questionnaires can be facilitated.

1.6 Thesis Overview

In this section, we present a brief, structural overview of the content of this thesis.

Chapter 1 introduces the research topic. It states the research problem's objectives and offers a rationale for the study. Furthermore, it outlines the scope and significance of the research with research questions, chosen method, and terminology.

Chapter 2 presents background information about the thesis and establishes the theoretical foundation of the study.

Chapter 3 discusses the research methodology and approach. This includes describing the chosen research design, methods, and tools. Additionally, ethical considerations are addressed here.

Chapter 4 details the development process of the integrated mental health assessment tool. It escribes the technical aspects of creating the artifact, including integrating wearable technology into self-assessment questionnaires. It also discusses any challenges or decisions made during the development phase.

Chapter 5 evaluates the implementation of the artifact. It summarizes key findings and insights from the evaluation and provides insights into the tool’s effectiveness in assessing mental health issues.

Chapter 6 analyzes the research findings in the context of the research questions and objectives. Furthermore, it addresses any limitations encountered during the study, answers to the research questions, contributions, and reflections. Lastly, a discussion on the project's constraints follows.

Chapter 7 summarizes the key findings of the research conducted and contributions made. In addition, future work and research is discussed.

Chapter 2

Background

This chapter presents and explains the theoretical background of essential terms regarding the thesis. We introduce the system we intend to extend, followed by a description of self-assessment questionnaires and their importance in health care and the mental health domain. Furthermore, an introduction to wearables and digital biomarkers follows. Lastly, a look at existing solutions is done.

2.1 Internet-Delivered Psychological Treatment (IDPT) Systems

Internet-Delivered Psychological Treatment (IDPT) represents a therapy delivered online (Andersson 2016). Other terms can be used in similar contexts, including web-based, internet-delivered cognitive-behavioral treatment, and e-therapy (Morgan et al. 2017). Guided IDPT involves active engagement from a trained clinician who essentially guides the patient through the therapeutic process via email and phone (Morgan et al. 2017). On the other hand, unguided IDPT systems put their trust in the patient, and they are solely responsible for engaging in the therapeutic process. (Morgan et al. 2017). A developed software platform is needed to facilitate these processes in guided form, including presenting treatment materials and health assessments like exercises and questionnaires (Andersson et al. 2016). These systems are what we coin as IDPT systems and include but are not limited to web and mobile applications or augmented or virtual reality. (Andersson et al. 2016). According to Mukhiya, most current IDPT systems are inflexible, tunnel-based, and non-interoperable (Mukhiya et al. 2020). Furthermore, scholars state that enhancing user adaptability improves treatment adherence and patient outcomes (Lamo et al. 2022).

2.2 Open-Source Adaptive IDPT

The framework we intend to extend as part of our artifact is the open-source adaptive IDPT framework developed by Mukhiya (2022). The framework builds upon the Figure 2.1, where the main idea is a digital system that delivers personalized care through self-adaptation. The main difference between the system discussed in Section 2.1 and the open-source adaptive variant is that it tailors the treatment plan or process based upon user input, feedback, and performance measures through the use of modules, cases, tasks, and assignments. The actor in the Figure is what triggers the adaptation based on analyzation of data, or answers to cases or assignments.

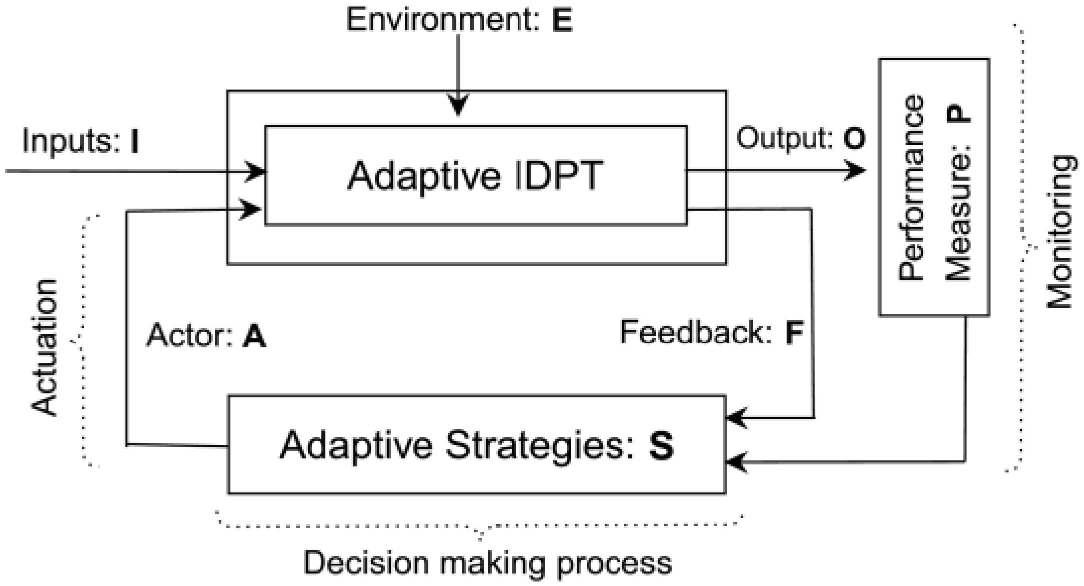


Figure 2.1: The adaptive IDPT system, derived from Mukhiya et al. (2020).

The way it works is that each module can have several cases and each case can have several tasks which are either informative or interactive. Interactive tasks can be questionnaires which requires participation and engagement from the user. Utilizing wearable technology shows proof that it can be valuable when providing assistive solutions and strategies in treatment process (Alhejaili & Alomainy 2023).

The thesis aims to extend the IDPT with general components which facilitates the possibility of integrating wearable technology into self-assessment questionnaires through our demonstrative component. We extend the IDPT framework with questionnaires and their corresponding responses while also providing a view where therapists can create the questionnaires and facilitate passive data collection, an in-depth discussion on the implementation phase can be found in Chapter 4.

2.3 Introduction to Mental Health Assessments

Mental health assessments represent a critical aspect of healthcare, offering invaluable insights into an individual's emotional and psychological well-being. These assessments evaluate a person's mental health status, identify potential disorders or disturbances, and provide a foundation for tailored treatment and support (Agency for Healthcare Research and Quality). Assessments are often used as a way of collecting information that can be used to plan interventions. In this section, we delve into the fundamental concepts and significance of mental health assessment tools, laying the groundwork for a deeper exploration of their role in mental healthcare and this thesis.

Mental health assessments serve a multifaceted purpose within the healthcare landscape:

Early Detection and Diagnosis: Assessments are often the first step in the diagnosis process and is used for early detection.

Treatment Planning and Monitoring: Mental health and individuals' thoughts, feelings, and emotions are complex. Assessment tools like questionnaires and surveys can pose a valuable way of monitoring the psychological state of an individual. It collects objective and real-time data that can be analyzed and leveraged to help therapists and patients through the treatment process.

Preventive Measures: Assessments used in mental health monitoring can, in many cases, help identify individuals early before they develop the disorders. The World Health Organization (WHO) states the mentioned (Preventing mental disorders: A research Perspective – WHO 2004).

Research and Data Collection: Assessments are a valuable tool for the collection of data used in research. Additionally, the bigger the data collected, the easier it is to locate trends and patterns, which can help in the early detection of mental health issues and symptoms. The National Institute of Mental Health provides resources and funding for creating and validating assessment tools, which promotes new solutions and ideas. (National Institute of Mental Health 2021)

Personalized Care: Assessments can be used as a valuable tool for creating customized care for a patient. Thus, it can tailor the treatment plan based on needs and symptoms. The IDPT framework discussed in Section 2.1 makes personalized patient care through interventions, and assessments can be used as a valuable tool to personalize the maintenance after the patients' needs.

Figure 2.1 shows the adaptive IDPT model for the IDPT framework developed by Mukhiya (2020) and the one to which we want to add self-assessment questionnaires. The self-assessment questionnaires are a performance measure that can be used to adapt and personalize the intervention in the decision-making process based on the answer.

2.4 Role of Self-Assessment Questionnaires

Historically, assessments used in mental health treatment and prevention have heavily relied on traditional self-assessment questionnaires. A questionnaire can be defined as "a series of questions asked to individuals to obtain statistically useful information about a given topic" (Roopa et al. 2012).

In psychology and mental health prevention, questionnaires are often used to collect self-reported data from a patient. Therapists often have limited time, and these assessments are an asset for collecting passive data about a patient that can be used to examine their general physiological state and potential issues (Pekrun 2020). Some examples are the Patient Health Questionnaire (PHQ-9 used for measuring depression) and the General Health Questionnaire (GHQ used for measuring general health). These standardized, pen-to-paper questionnaires identify present mental disturbances and disorders within a primary care setting (Anjara et al. 2020).

These well-known self-assessment questionnaires are a valid and reliable tool to measure mental health indicators like stress, depression, and lack of sleep. However, there is a growing consensus that these tools have limitations, and their validity can be questioned (Rolstad et al. 2011; Colombo et al. 2019). In other words, there is a growing need for more objective and reliable tools in the future.

2.5 Challenges and Limitations of Self-Assessment Questionnaires

Self-assessment questionnaires can provide valuable insight into a patient's physiological and mental state (Grassini & Laumann 2020) Furthermore, researchers opting for questionnaires rely on accurate and truthful responses from the participants. As mentioned, the main driver of this tool is that the patient, often in retrospect, provides information based on their current mental state and characteristics, leading to resistance to possible answers, which can increase the false alarm rate. (Rolstad 2011). Common drawbacks and limitations to self-assessment questionnaires are:

Response Bias, Larson (2019) describes respondents' potential to provide inaccurate or skewed answers due to various factors. In questionnaires, response bias is often referred to as social desirability bias, a type of response bias. Specifically, social desirability bias is when a participant portrays a positive image of themselves rather than the truth. (Johnson & Friedrich 2005). This is a common issue with self-assessment questionnaires, where participants tend to misrepresent or alter their true feelings, meanings, and thoughts. This results in a subjective answer that is not representable and could lead to misinterpretation of the response and possible mistreatment or diagnosis (Graeff 2005).

Response Burden is defined by Rolstad et al. (2011) as "the effort required by the patient to answer a questionnaire. Traditional self-assessment questionnaires can be quite long and requires substantial effort from the patient. Yan et al. (2019) proposes that the response burden in questionnaires occurs through many different factors: motivation, effort needed, difficulty, time, and more. Furthermore, the response burden in patient-reported outcomes should be addressed to avoid high rates of missing or "false" data. In the end, this may lead to poor data quality, and the decision-making process for a clinician or therapist becomes complex (Aiyegbusi et al. 2022). A study by Edwards et al. (2004) indicates that responses to a questionnaire can potentially be increased with a shorter questionnaire. Additionally, they state that making changes to a shorter questionnaire has a better effect than making minor changes to a longer questionnaire.

2.6 Emergence of Wearable Technology

In recent years, there has been a growing prevalence around the use of wearable technology, which includes smartwatches and fitness trackers. These devices show promising potential to become a staple in patient diagnostics and patient monitoring thanks to the passive continuous collection of data (Henriksen et al. 2018). These devices enable individuals to monitor and track various facets of their health and well-being, including physical activity, heart rate, and sleep patterns (Xie et al. 2021). The continuous growth in popularity of wearable technology has led to a steady rise in the constant collection and storage of individual health data. This can significantly improve healthcare outcomes and aid in developing new treatment methods and tools (Chen et al. 2021). However, scholars and scientists argue that commercialized-grade wearable technology differs in how data is measured compared to gold-standard equipment, such as medical-grade sensors. One study state that fitness trackers continuously underestimate the number of steps taken (Tudor-Locke et al. 2015). As technology advances, ubiquitous wearable devices can play a vital role in the future of healthcare.

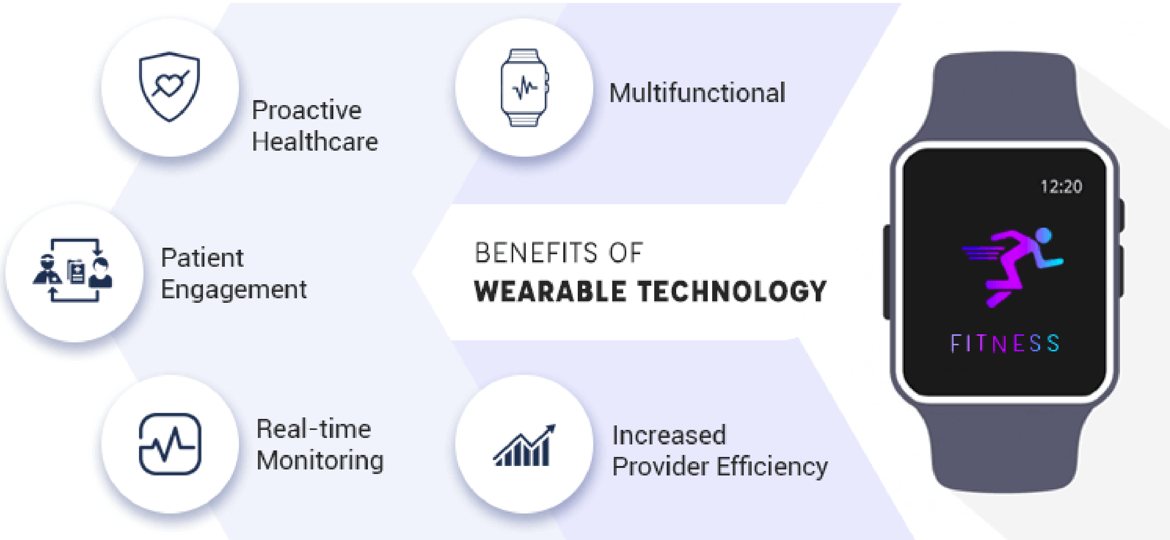


Figure 2.2: Benefits of wearable technology in healthcare, derived from (Sinhasane 2018).

2.6.1 Ethical Considerations

Commercialized wearable technology shows great promise. With these promises, ethical considerations arise that are important to address and understand when working with sensitive health data.

Privacy and Data Security is a crucial aspect that must be addressed. Since we will work with individual patient health data and the sharing of this, it is essential to understand the audience and concept. Wearables can continuously monitor and record data that only a handful of people should have access to. This includes but is not limited to health data and geographical location (Quach, et al. 2022). With this technology, one often must share the data between services, and they can essentially be leaked, stolen, or tampered with. Wearable data collection is still an emerging market, and it is essential to note and ensure the privacy and security of data to keep and gain public trust. A study by Nicholas et al. (2019) concludes that the threshold of sharing sensitive health data through assessments depends on the data type collected and the recipient of the collected data. Nevertheless, this gives us an important reminder that systems implementing this should conform with sensed data sharing, explicitly tell what is measured, and state the validity and role of the recipient.

Informed Consent is the way of telling the user how their data is used. These can often be lengthy and time-consuming, leading to users not paying attention to it. It is essential to state what the data is used for, how it is collected, and which third-party systems have access to their data. (Jiabong 2021). This can have unintended consequences; Fitbit users might unintentionally agree to share their data with third-party apps when they create an account (Fitbit 2023)

Data Ownership is crucial. If an individual owns a watch and synchronizes their health data with the company's companion app, it is stored in one of their databases. Still, an individual should have complete control of any sensitive data. Without correctly stating how the data is stored, potential issues can arise.

Accuracy and Reliability is something that needs to be addressed. Commercialized consumer-grade wearables often use proprietary software with algorithms that are not visible to the public (Godfrey et al. 2018). This creates a blackbox problem where users do not know how their sleep score, or stress level is tracked and scored. Different wearables have different biomarkers and sensor that collects sensor data. How can we rely on these wearables if they use different methods or means to track the same digital biomarkers?

These are important ethical considerations that we need to consider when collecting sensitive health data.

2.6.2 Challenges and limitations

One must consider possible challenges and limitations with new and widespread technology that wants to revolutionize the health sector. When developing a new way of answering self-assessment questionnaires, it is essential to have this in mind. The challenges and limitations listed are crucial; however, more could be argued.

Interoperability The International Organization for Standardization/International Electrotechnical Commission (ISO/IEC) defines interoperability as "the capability to communicate, execute programs, or transfer data among various functional units in a manner that requires the user to have little or no knowledge of the unique characteristics of those units". We often speak about interoperability between different information systems in informatics and the health sector. Since there is no uniform set of data structures for the exchange of wearable data, it might lead to potential pitfalls when trying to implement it into existing systems. See the data standards underneath for more information.

Data quality grasps the concept of how valuable the collected data is for the context of where it is used. One must learn and use data quality and the collected data to facilitate the collection of meaningful, reliable data for use in the healthcare sector. This also encompasses the collected data's integrity, accuracy, and validity. Bottcher et al. (2022) conducted a study that shows data quality in wearable monitoring. They state that the types measured are suitable for accuracy and validity; however, since it is an emerging technology, one must create a common consensus on how each data type should be collected to enhance data quality in wearables.

Data Standards are a heavily discussed topic when incorporating new technologies into the old manual way of data collection. Data standards can be defined as "a type of standard, which is an agreed-upon approach to allow for consistent measurement, qualification or exchange of an object, process, or unit of information." (National Library of Medicine 2023) The use of medical data hinges on its transformation into meaningful information. This transformation necessitates the availability of high-quality data, efficient communication, and the service of understandable standardized data; without it, collected data is nearly useless (Lehne et al. 2019). This is also applicable to wearable data collection. Countless wearable devices are floating around in every continent. They all have different manufacturers, and they collect and store data in different ways. Naturally, there is no common consensus and widely used data structures/standards for the storing, sending, and retrieving these data. This is because each manufacturer processes and stores the data in the format that best suits their system.

However, there have multiple attempts to accommodate this problem. Standardized data structures for interoperability are one of the key issues that health informatics continuously tries to solve. One of these attempts is Fast Healthcare Interoperability Resources (FHIR) made by Health Level 7 International (HL7). FHIR seeks to categorize the fundamental entities in healthcare information exchange as resources. Where each resource represents a unique and identifiable entity (Bender et al. 2013). According to research, mobile applications, research tools, and SMART on FHIR applications are the primary beneficiaries of this standard, followed by web-based applications (Ayaz et al. 2021)

2.7 Physiological Data in Mental Health Assessments

In our pursuit of creating a more reliable and accurate self-assessment tool, we focus on what type of domains are linked to specific mental health issues. To facilitate wearable data collection in self-assessment questionnaires, it is crucial to understand what data types can be harnessed and how reliable the data is for a specific domain. Stress, sleep, and physical exercise are the domains in focus due to their strong correlations with mental health (Stead et al. 2010; Zou et al. 2020; Mikkelsen et al. 2017) and therefore, it is researched further to understand how we can measure and utilize these domains in our research.

Stress can be defined as "any influence of internal and surrounding environment on living being which disrupt its homeostasis (Shahsavarani et al. 2015). In phycology it is defined as "a state of worry or mental tension caused by a difficult situation" (World Health Organization 2023) Stress is correlated to many mental health issues such as anxiety and depression.

Sleep promotes healthy brain function and an excellent physiological state (NHLBI 2022). Scholars suggest that sleep-related problems correlate to poor mental health outcomes, including anxiety, aggression, and depression. (Ramsawh et al. 2009; Zou et al. 2020). Indeed, sleep is a complex concept to grasp and a subject that scientists have struggled with for a long time. A study conducted by Scott et al. (2021) concludes that a well-rounded night of sleep repeatedly has a significant impact on mental health. On the other hand, sleep deprivation can lead to increased anxiety and distress levels in healthy individuals (Columbia University Department of Psychiatry 2023)

Physical Exercise or physical activity is vital for a good mental and physical health (HelseNorge 2023). Exercise is "a subset of physical activity that is planned, structured, and repetitive and has a final or an intermediate objective, the improvement or maintenance of physical fitness" (Caspersen et al. 1985). Research continuously suggests that physical exercise repeatedly done has a significant effect on an individual's mental health (Mikkelsen et al. 2017). According to multiple studies, regular physical exercise lifts the mood and creates better self-esteem while stress levels go down (Mahindru et al. 2023)

2.8 Digital Biomarkers

To facilitate the substitution of self-report questions with passive data collection, it is essential to understand what digital biomarkers can be collected that correlate with one of the domains discussed in Section 2.6. *A digital biomarker* is an "objective, quantifiable, physiological, and behavioral measure that is collected utilizing digital devices that are portable, wearable, implantable, or digestible" (Babrak et al. 2019). To monitor and passively collect data about individuals, the devices used must be accurate, validated, and reliable (Ahmed & Sundas 2022).

In the scope of this thesis, digital biomarkers are the raw or aggregated data collected from the patient's wearable device. Furthermore, by harnessing this technology and using these devices, healthcare professionals can access a combination of objective and patient-reported health data. This new access can lead to better treatment outcomes, enhanced decision-making, faster treatment, and better quality (Liao et al. 2019; Cox et al. 2018). Outlined in the rest of the section are reliable digital biomarkers that can be found in consumer-wearable devices for each domain. The biomarkers discussed are chosen based on their positive correlation to one of the domains, and because they can be found in ubiquitous commercialized wearable devices.

2.8.1 Stress

Heart Rate: While HRV is considered a more reliable way of measuring stress, heart rate metrics such as resting, average, and maximum heart rate can be a good indicator of stressful events or prolonged stress over a given period (Lee et al. 2014). A study conducted by Lee et al. (2014) reveals that, on average, participants who did not engage in stressful live events or had little to no engagement generally had a lower resting heart rate.

HRV Heart rate variability, referred to as HRV, is "changes in time intervals between adjacent intervals" (McCraty 2015). McCraty also states that having slight variation between adjacent heartbeats leads to inadequate function in various levels of the body or chronic stress (2015). This biomarker is commonly used to detect stress in a person, and research sees it as an accurate and reliable way of measuring stress in an individual (Lee et al. 2022; Kim et al. 2018; Makovac et al. 2022)

Both skin temperature and electrodermal activity shows proof of the ability to indicate certain stressors, however, they are left out since the technology in ubiquitous wearable does not support it.

2.8.2 Sleep

When measuring sleep in individuals through wearable technology, research is divided. Sleep is indeed hard to measure as it is intricate due to the complex way sleep works. Different manufacturers tend to have different ways of scoring sleep through algorithms. However, there is a consensus on what data types are measured and how they should be calculated. However, sleep data is often aggregated, and the potential to get hold of the raw data is challenging. IBM defines aggregated data as the data expressed and gathered from raw data in a summary form for statistical analysis (IBM YEAR). However, the data is aggregated by hidden algorithms provided by the manufacturer, which can cause issues since we need to know how it is measured and translated. Aggregated data is still objective and can be seen as a better representation of an individual's physiological state than patient recollection of events. Listed underneath are what is measured in a sleep response from wearable devices.

HRV is a versatile digital biomarker commonly found in various wearables. In addition to the positive correlation with stress, it provides a good measurement for sleep analysis. Research suggests that HRV gathers essential data on autonomic changes in an individual's sleep pattern (Tobaldini et al. 2013)

Resting Heart Rate states the heart rate of an individual during sleep. It is often lower than when in a waking state. A study conducted underlines that heart rate is associated with sleep; lower resting heart rate correlates with a good night of sleep and vice versa. (Sajjadieh et al. 2020)

Time Awake in bed indicates a patient's struggles of falling asleep and waking up, perhaps.

Sleep Stages can help to depict the quality of sleep. Specific mental health issues can alter time in different sleep stages (Patel et al. 2022). Thus, measuring time spent in each sleep stage can provide valuable information about an onset or change in psychological state.

Total time slept can give a good indication of the quality of a night's sleep. An adult should average about 7 to 8 hours of sleep a night (HelseNorge 2023). The use of a wearable device can effectively measure this, and it is measured as the total time in bed minus the time spent awake.

2.8.3 Physical Activity

When measuring physical activity in individuals with wearable technology, research and consumer-grade wearable devices use activity metrics that measure aggregated data such as steps taken, calories burned, heart rate metrics, distance (Watanabe & Tsutsumi 2022).

These three areas were chosen because of their positive correlation to identifying mental health issues. Stress is usually measured with HRV, heart rate metrics and skin temperature in commercialized wearables. While sleep and physical exercise use accelerometers and activity metrics. Based on the literature search, these areas are a valid and reliable way to recognize patterns that can help clinicians and therapists. In other words, collecting this data passively instead of asking the user to self-report on specific questions can help to create a new and enhanced version of self-assessment questionnaires.

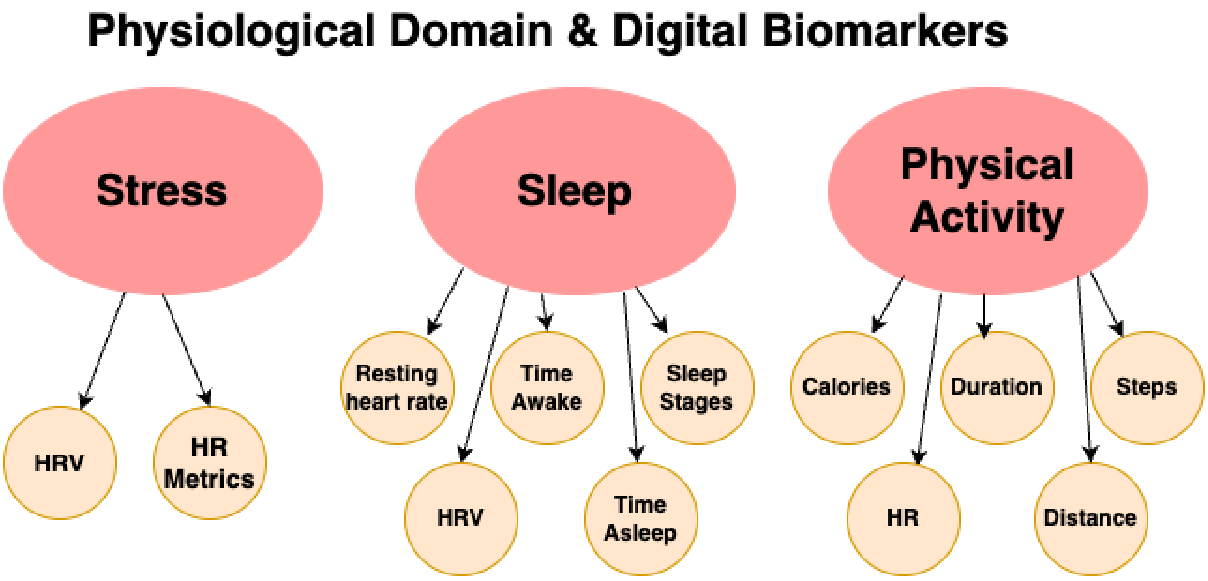


Figure 2.3: This figure depicts the physiological domains and their corresponding digital biomarkers used in this study.

2.9 Existing Solutions

From what is provided by research and found in the literature search, integrating wearables into self-assessment questionnaires is a novel approach; by other means, there are no publicly available solution according to the literature search conducted. However, studies and projects with a slightly different approach have been conducted. Revilla et al. (2016) did a research study where they installed a tracking application on participants' devices to see if answers would differ between the group that used both the tracking application and handed in self-report data versus those who only reported data from their own recollection of events. The study showed that answers from participants who provided passive and active data differed significantly from answers provided by the same respondents who only filled out the web-based survey. Additionally, they state that based on these factors, combining several data collection sources in the future is necessary to create better and more accurate tools. Applications or research projects that have tried to do something similar are discussed in the following sections.

2.9.1 CrossCheck

CrossCheck resulted from a research study by the UW Medicine Department of Psychiatry and Behavioral Sciences at the University of Washington. They made a mHealth application that integrated ecological momentary assessment (EMA) with self-report questionnaires administered through a smartphone (Ben-Zeev et al. 2017). The goal was to integrate EMA, self-report, and a smartphone to identify better digital indicators correlating with psychotic relapse. (Ben-Zeev et al. 2017). The study concludes that implementing mHealth applications with enhancing technology is possible to help clinicians and therapists identify critical patterns that lead to relapse earlier than what can be done the traditional way (Ben-Zeev et al. 2017). Additionally, they state that using, understanding, and applying new technology to existing health assessments is essential to facilitate novel applications that can help improve individual treatment outcomes. (Ben-Zeev et al. 2017)

There is no place where we could find and investigate the demonstrational component or the source code the platform was built upon. However, they show a figure that indicates how the data in question are captured, sent, and analyzed for remote monitoring.

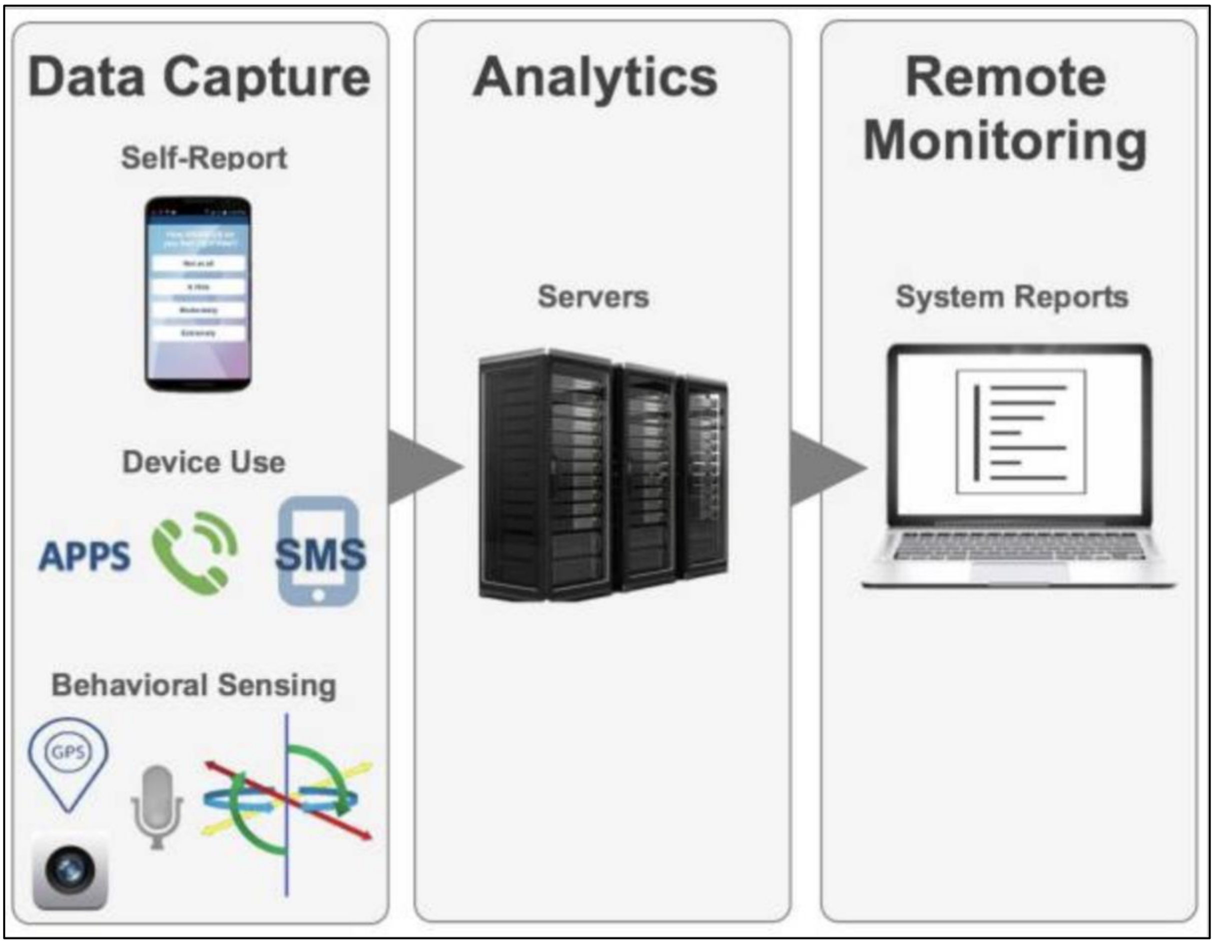


Figure 2.4*:* Crosscheck Data collection. (Been-Zeev et al. 2017)

2.9.2 ExpiWell

ExpiWell is a platform developed by researchers and can be downloaded through the Google Play or Apple store. It is a tool for "human experience, innovation, and practice." (ExpiWell 2022). Scientists make it from the field of EMA and experience sampling methodology. It provides a platform where subscribers can create user-centered surveys consisting of passive EMA data collection and participant self-reporting. In this app, one can survey everything, including logic for questions and what type of EMA should be measured. In Google Play, the app is described as "the premium experience sampling method (ESM) and ecological momentary assessment (EMA) app for surveys, mood tracking, daily diary, just-in-time adaptive interventions (JITAI), and digital ethnography.

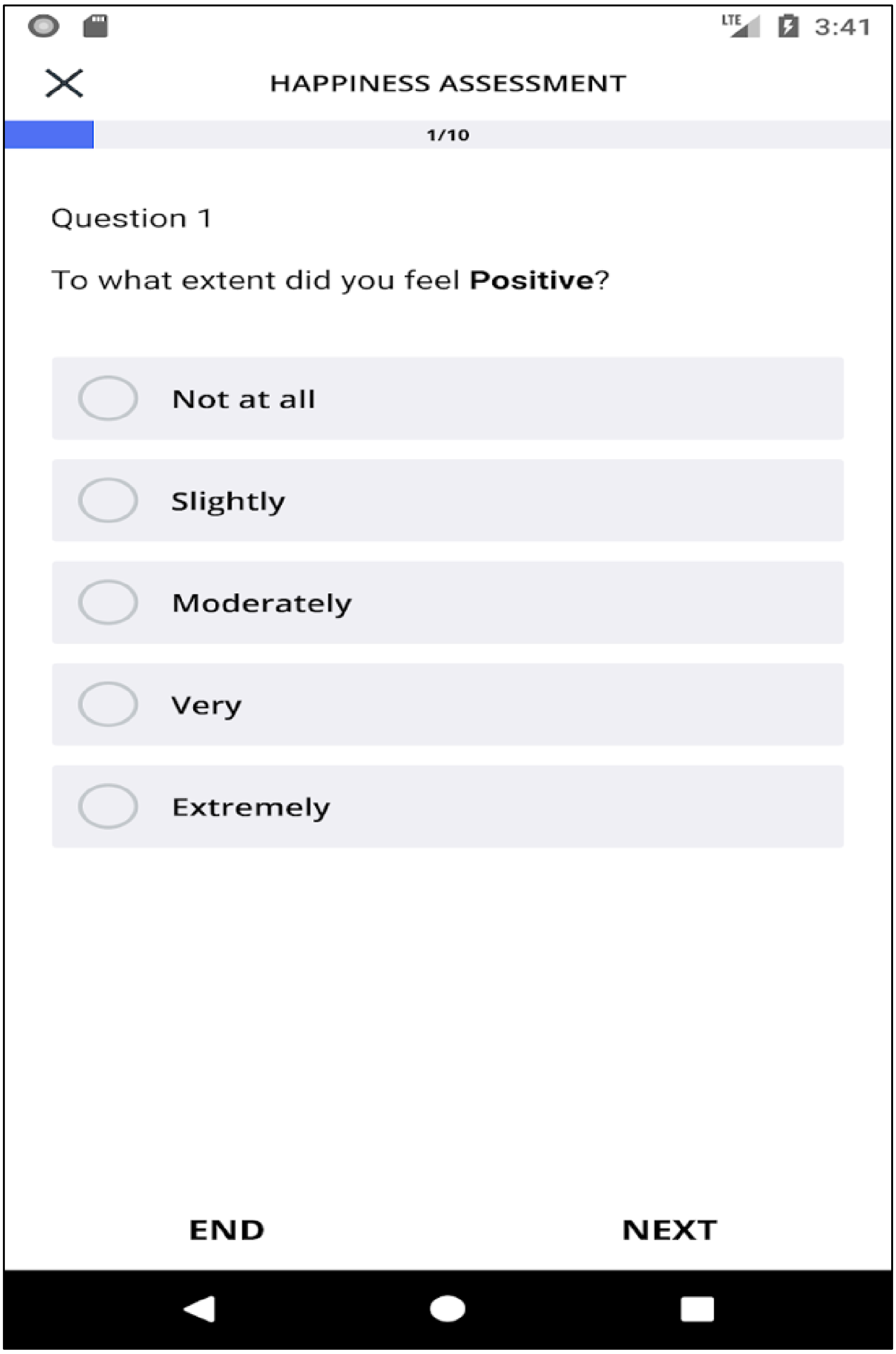


Figure 2.5: Questionnaire response UI for the ExpoWell application (Google Play Store 2023).

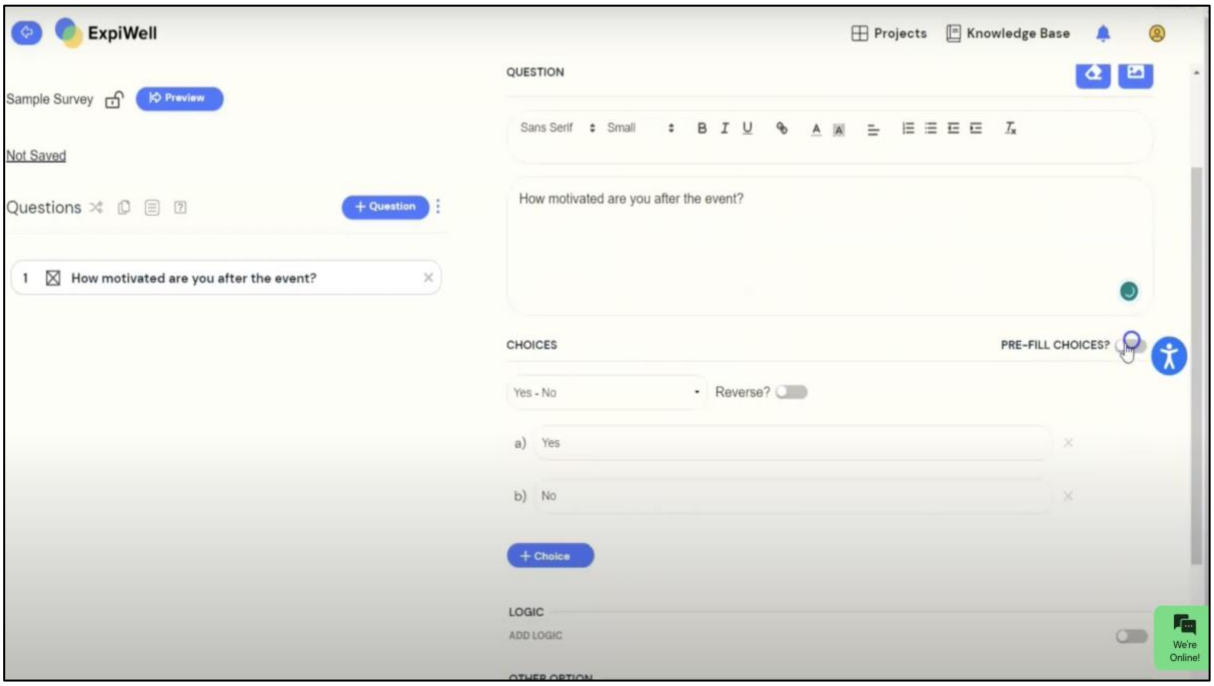


Figure 2.6: Dashboard for creating surveys and questionnaires in the EchoWIll platform. (EchoWill 2023)

It supports the creation of surveys and assessments with EMA data collection in the background, given the type. There needs to be more information from the system's website, research, or GitHub on how it is implemented and works. The surveys or questionnaires can be set to be answered multiple times a day, daily, weekly, or monthly, to gather the appropriate amount of data. (EchoWill 2023). The app itself looks like how we want to implement our demonstrative component; however, since it is surveyed and can be for anything and not particularly mental health-related problems, it makes it hard to compare them. In addition, no open-source code exists or documentation on how thet retrieve the wearable data and the formats they use to store them in. We tried to reach out to them to get valuable information, however, we were left on read.

Both existing solutions mentioned try to make use of EMA, however EMA is usually something measured when a certain event unfolds and therefore is not something that we want to do. Moreover, extensive documentation and open-source code is needed to understand how they standardize the data in for the values measured and the response created. Neither of the platforms states how the application works, and no system architecture was found. Our general components focus on implementing questionnaires, questionnaire responses, and passive data collection. More on this can be found in Chapter 4.

Chapter 3

Method and Design

In this chapter, we intend to explain our chosen research methodology and touch base on how we want to utilize the paradigm guidelines to design and create an artifact that facilitates wearable integration with self-assessment questionnaires. This section includes information on the research method chosen and a walkthrough of the different iterations of the development process.

3.1 Research Method

As mentioned in Section 1.4, design science is the research methodology chosen for this thesis. *Design science* is a research method coined by Hevner et al. in a paper from 2004 named Design Science in Information Systems. Hevner et al. (2004) state that design science is a research method that swerves around the concept of contributing to a set of problems/specific problem domains in the shape of an artifact.

3.1.1 Using the Design Science Guidelines

Hevner et al. (2004) describe seven guidelines that should be used to apply the design science paradigm in their research effectively. The described guidelines serve as a base for the use of design science, dwells around the core principles of the paradigm, and are used to address what design science strives to achieve. In this section, we address the different guidelines and touch base on how we intend to use them.

Guideline 1: Design as an artifact
Hevner et al. (2004) state that the first guideline in design science focuses on the development of an artifact. The artifact itself is the main goal or focus of design science research and must apply to the described problem domain. Furthermore, the artifact can be developed in one of the forms below:

A construct serves as a vocabulary of symbols for communicating specific problems in the domain questioned.

A model or an abstraction that gives a well-rounded representation of the problem domain. These models are often constructed with the use of various constructs.

A method that aims to capture specific processes within the scope of the problem domain and then provide guidance on how a solution can be found. A typical example is an algorithm.

An instantiation or practical system solution that illustrates how constructs, models, or methods can be implemented effectively in real-world practice.

We intend to create an instantiation, which is an artifact that exist of an extension of the IDPT framework and an mHealth application to demonstrate how wearable data collection can be facilitated for the use in questionnaire responses. The problem domain discussed consist of both methods and constructs that we can use and leverage to create an instantiation in the form of an artifact.

Guideline 2: Problem Relevance
Guideline 2 addresses the importance of the relevance of the created artifact. Moreover, the developed artifact should be created to solve a relevant problem in the domain or bridge the gap between the current state and the ideal state; in other words, the artifact should be original and resourceful. The artifact developed aims to contribute to the problem domain described in Section 1.2

Guideline 3: Design Evaluation
It is essential to thoroughly evaluate and understand the developed artifact in design science research. With the help of evaluation methods, the resulting artifact has a high probability of contributing to a solution or being a step in the right direction for the specific problem domain. Additionally, having good design evaluation fosters the crucial points discussed in Guideline 2 by making the developed artifact relevant to the problem domain it seeks to answer.

Since this is a novel approach for the specified problem domain, we evaluated our artifact using qualitative and quantitative evaluation methods. This includes the likes of domain experts and potential users. More information about the evaluation of our artifact can be found in Chapter 5.

Guideline 4: Research Contributions
The following guideline proposed by Hevner et al. (2004) addresses how the research conducted in design science should lead to contributions in the problem domain that can help expand the knowledge base. The contributions for the problem domain must be in at least one of the formats listed below:

A Design Artifact can be novel and innovative, extending the knowledge base for the proposed problem domain. In addition, it can also be an artifact that solves a gap in the knowledge base through it being visionary and ground-breaking. This is the most common research contribution in the design science paradigm.

Foundations encompass constructs, models, methods, or instantiations that contribute to the knowledge base by extending the knowledge for the problem domain.

Methodologies are the third and last contribution form in design science., Hevner et al. (2004) state that methodologies can be seen as evaluation methods (see Guideline 3) that can lead to an extension of the knowledge base. In other words, this is when evaluation metrics are applied to better understand the artifact or proposed solution in the context of design science research.

In the thesis, research contributions of the designed artifact can be found in Section 6.2. The main contribution to the chosen problem domain is the design and development of an artifact.

Guideline 5: Research Rigor
The fifth guideline encompasses using rigor in design science research. In other words, Hevner et al. (2004) state that rigor in design science revolves around using the existing knowledge base efficiently.

Due to the novel approach of the project done in this thesis, it was essential to use existing literature through extensive literature searches to understand the data types that has a positive correlation with mental health issues and what type of data is reliable and accurate enough to measure for the domains. Additionally, the involvement of domain experts in mental health and wearable technology is an efficient use of the current knowledge base. Lastly, it is crucial to understand what the potential user of the artifact thinks about the solution. We believe this was a bright and effective way to give us the experience and knowledge needed to develop and evaluate the created artifact. The development of the artifact and how it was done is described in Sections 3.2 and 4.1. We used methods to evaluate our work, which can be found in Chapter 5.

Guideline 6: Design as a Search Process
The sixth guideline swirls around, searching for the ideal solution to the problem addressed in the domain. It is often done using an iterative method where solutions are tested against the domain. Naturally, it is an adequate approach if the problem domain and knowledge base are well documented. However, when designing and developing new solutions, especially when incorporating new or emerging technologies like consumer-grade wearable devices, comparing it to existing solutions might not be that easy. The problem is that the artifact is novel, and there are few or no solutions to compare it with. It is important to try to design a satisfactory solution.

In our thesis, we intend to create a novel solution with the help of wearable technology. The problem domain consists of emerging and new technologies and thus a comparison to already developed solutions is challenging. However, there is evidence in the knowledge base that encompasses a part of our solution, like the passive collection of data through wearables and the issues with traditional self-assessment questionnaires. Naturally, this is a complex problem, and for the readiness of this thesis, we focus on a possible solution to the challenges of these questionnaires. Moreover, qualitative, and quantitative evaluations was posed on the artifact to get an understand of the usefulness.

The design process used in the development of our artifact can be found in Section 3.2

Guideline 7: Communication of Research
The seventh and last guideline states that communication is vital. Hevner et al. (2004) state that research through design science must be communicated precisely to make it easy for all

audiences to understand. This includes developers and individuals who use and understand the proposed solution and further develop the solution made.

The communication of the research and the result of this thesis can be found throughout the paper. In addition, both the IDPT framework and the mHealth application boosts open-source code and development documentation. This can promote future work and help one get a better understanding of the artifact developed. Furthermore, this thesis provides an in-depth explanation of the created artifact and thus can be referenced and used as a documentation document for the

3.2 Design Process

The following section describes the iterative process of developing and designing our artifact. It includes a literature search, semi-structured interviews of domain experts, and surveys to understand what potential users think about our solution. This gave us valuable feedback throughout the implementation and design that was used in different ways to enhance the result.

The implementation process was divided into three primary iterations to highlight the different steps to develop the artifact. The three primary iterations are described below.

Iteration 1: Extending the IDPT framework.
With a good understanding of the problem domain, the first iteration focused on extending the IDPT framework. We added two entities to the IDPT framework, namely questionnaire and questionnaire response. Moreover, it creates a way where therapists can create questionnaires with input on which questions should be self-reported by the user and which questions should be answered using passive data collection. In this iteration, we also added the questionnaire response entity to the IDPT framework to facilitate the patient responses that should be sent for analyzing and displaying. Change the last sentence.

Iteration 2: Wearable Integration and mHealth application
With the addition of questionnaires and questionnaire responses in the IDPT framework discussed in Chapter 2.1, this iteration focused on the creation of the mHealth application so that we could integrate wearable technology to facilitate passively collecting data for specific questions. The iteration's focus was to successfully integrate the wearable data collection for specific domains and biomarkers and adding this as an observation of the questionnaire response. In addition, the skeleton for the mHealth application developed in Android Studio was made. It focused on functionality rather than user-friendly UI development in this iteration.

Iteration 3: System design and development
In the third and last iteration, we focused on making all the different parts work together as a unit. This involved getting the mHealth application to work with the IDPT system discussed in Section 2.1. In addition, we created views for therapists in the IDPT framework and UI for the mobile application. Moreover, the questionnaire response generated in the mHealth application is sent to the IDPT framework for storing and visualization. A semi-structured

interview, an object-oriented experiment, and a user acceptance survey were conducted with domain experts and users to acquire a qualitative and quantitative evaluation of the artifact following the end of this iteration—more on this in Chapter 5.

Chapter 4

Implementation

In this chapter, we will address the process of developing the artifact discussed throughout this paper. For our thesis, as described in Section 2.3, we wanted to look at the possibility of switching out questions in a traditional self-assessment questionnaire where it was applicable to create a new and innovative tool that would potentially lead to less response burden, response bias, and more accurate and reliable answers. First and foremost, we had to determine what type of questions could potentially use passive data collection instead of active self-reporting and determine what data types are reliably used to measure these indicators. For this we concluded extensive literature searches to find domains and digital biomarkers that can be used to measure values for the domain. This can be seen in Section 2.7 and 2.8. Wearable technology showed promise because of the potential ability to measure signals that have direct relations with mental health issues. (Lee et al. 2014; McCraty 2015; Zamkah et al. 2020). Integrating ubiquitous wearable technology into standardized traditional self-assessment questionnaires could potentially give therapists and clinicians more reliable answers to help early understand and prevent mental health issues.

First, we mention the different tools used, followed by an extension of the IDPT framework to facilitate the use of questionnaires in our demonstrative contribution, followed by an in-depth description of how the demonstrative contribution was designed and developed.

4.1 Technology

This section addresses the tools and frameworks used to develop our artifact. The artifact extends an existing framework and creates a new system for using questionnaires and their corresponding responses created by the extension developed. The new system serves as a demonstrative contribution to the general contribution; in other words, it showcases a potential solution to the problem domain. This section will address the technology used in creating the new system and touch base on technology from the existing framework. The IDPT framework has already been discussed in earlier research projects, so there is no need to add more information to what already exists. The following technologies and framework were utilized in the making of the mHealth application, see Appendix D for open-source code access.

MongoDB is a no relational database that gives the user a way of storing flexible documents in separate collections through queries and mutations (MongoDB 2023)

Kotlin is a statically typed programming language running on the Java Virtual Machine. It is commonly used in developing Android applications. (Kotlin 2023)

Room Database is a database library for Android applications. It provides an additional abstraction layer over SQLite, commonly used in relational database management systems. It provides an easy way to run local offline databases in Android applications. (Room Database 2023).

Jetpack Compose is an Android UI framework developed by Google that makes it easy for developers to create native UIs when developing an Android application. It provides a declarative UI, reactive updates, state management, and many other beneficial features for creating interactive user interfaces (Jetpack Compose 2023).

GraphQL is a query language for APIs available as an open-source project. It boosts the creation of flexible data fetching by making the client request only the data they need through queries and mutations instead of traditional RESTful APIs. In addition, this means that one only needs to create a single endpoint for all queries and mutations (GraphQL 2023).

4.2 System Architecture and Design Patterns

This section addresses the architecture of the mHealth application to facilitate the use of answering the self-assessment questionnaires made in IDPT and sending the questionnaire response back. Before we look at the new system, let us understand how the IDPT framework works. It consists of a backend and a frontend. One finds endpoints, repositories, services, and databases in the backend. The repository in the backend performs tasks such as database management for a specific entity, like assignments, users, questionnaires, and questionnaire responses. The services work on logic related to the entities mentioned by facilitating the endpoints made in the backend. Any external system can initiate a call to one of these endpoints.

Queries and mutations are sent to the API endpoint for the backend of the IDPT. These graphQL operations uses a service for the specified entity, which again utilizes a repository that performs the database operations. Below is a Figure showcasing how our components work in the IDPT framework.

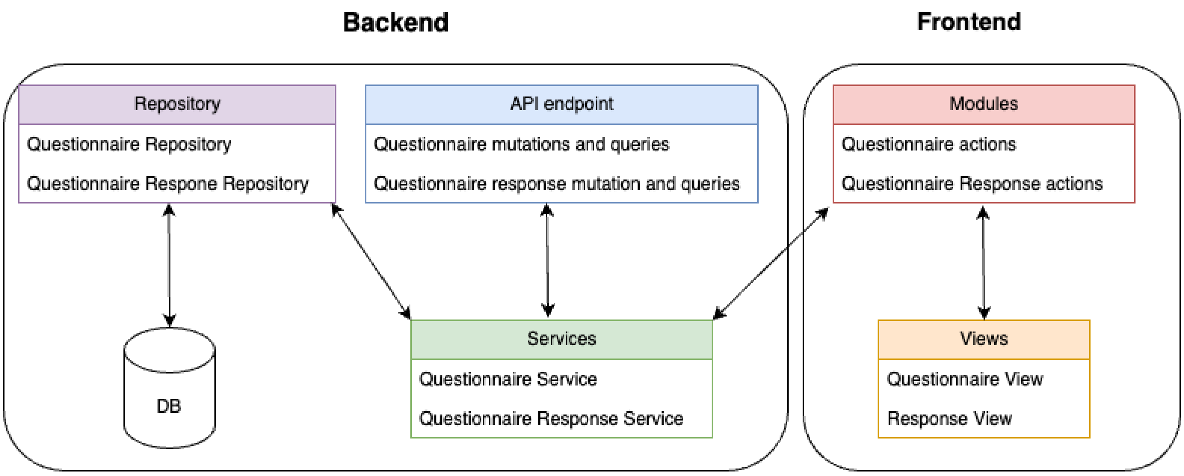


Figure 4.1: Depicts the logical flow of the entities added in the IDPT framework. It follows the same flow as the other entities.

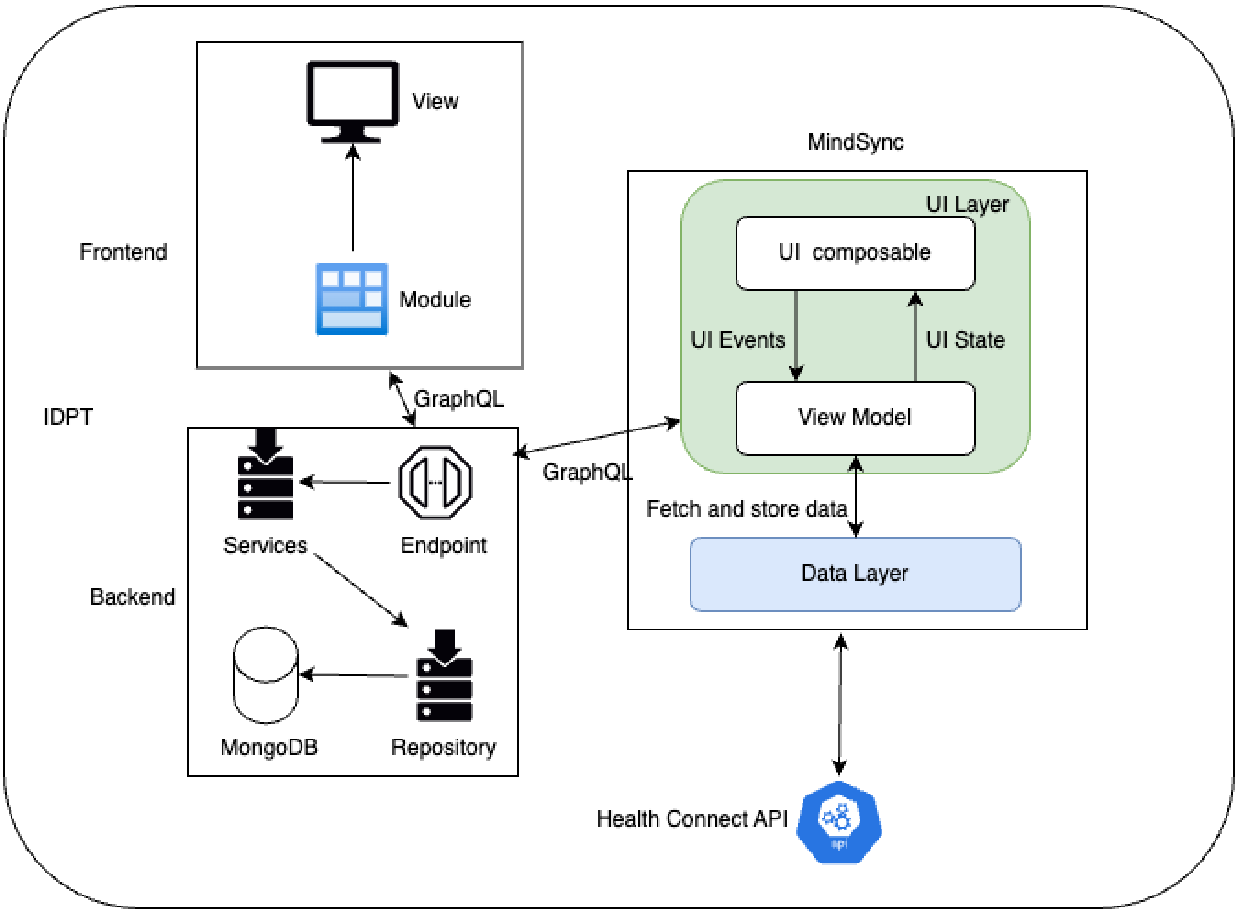


Figure 4.2: High-level architecture of the developed artifact.

4.3 General components

This section provides an overview of the general components of our artifact. This section will serve as a high-level introduction to the core components and functionalities developed and later applied in the demonstrative component. The demonstrative component makes use of the general components, and they overlapped in the design process, however, for the readiness of the thesis we decided to divide them and address them separately. To demonstrate our research contributions, these where the general components developed in the implementation phase.

- **(i)** **Extending the IDPT framework.**
- **(ii)** **Integrating wearable data collection**
- **(iii)** **Fetching, storing, and sending of wearable data**

4.3.1 Extending the IDPT framework.

The first thing was to extend the IDPT framework. The framework consists of a backend and a frontend that communicates with GraphQL queries and mutation through an endpoint. The first thing we had to do was to extend the framework with a questionnaire entity that could be used for the creation of questionnaires in the frontend by therapists. At first, we tried to add questionnaire as an option in the assignment entity, however that changed after we found out that the structure of the questionnaire schema had to be different than the set up for the assignment schema. In addition, we added a questionnaire response entity to facilitate for the answers to the questionnaires.

4.3.2 Questionnaire

The questionnaire entity holds the data that should be used to At first, the questionnaire entity followed the same structure as the other entities in the framework. After thoughtful considerations we opted out of this structure to facilitate the use of a standardized format that promotes interoperability and easy integration for future development. Specifically, we decided to apply the HL7 FHIR specifications that makes use of standardized data formats commonly used in the healthcare sector for our entities. The questionnaire entity and its corresponding questionnaire response entity have their own models for visualization in the frontend, more on that in Section 4.4.

According to the HL7 FHIR's Guide to Resources (FHIR 2023), a resource or a model is seen as a component. FHIR lists 157 resources on their homepage, and they can consist of different subcomponents that is a part or can be referenced to the main component. Each resource has their own pre-defined structure that can be utilized to enhance interoperability. Below we have listed each HL7 FHIR resource we opted on using in our artifact.

Questionnaire is a resource type utilized for creating questionnaires. The questionnaire resource consists of a set of predefined fields that can be used when creating a questionnaire resource. In our project, it was utilized to create the questionnaire resource by therapists in the IDPT framework.

Questionnaire Response is a resource type for the list of answers to a specific questionnaire resource. It holds predefined fields for the structure of the answers and other data fields needed. It is used as a way of creating response objects to the questionnaires mentioned in our project.

Observation is a resource type used for measurements taken about a patient, device, or other subject. It consists of predefined required and optional fields. In our project it is used when we add wearable data to the answers in the questionnaire response.

The questionnaire makes use of the FHIR standard for creating questionnaire components found on the FHIR website and has the following structure.

```
const schema = `
 type Questionnaire {
   id: String!
   resourceType: String!
   status: String!
   title: String!
   type: QuestionnaireTypeEnum!
   item: [QuestionnaireSchema]
   subjectType: String!
   publisher: User
   date: DateTime
   repeats: FrequencyEnum!
 }

 type QuestionnaireSchema {
   linkId: String!
   type: QuestionTypeEnum
   text: String!
   required: Boolean!
   answerOption: [Options]
 }

 type Options{
   value: String!
   label: String!
 }
```

Listings 4.1: CraphQL schemas used for the questionnaire resource.

As seen in Listings 4.1, the questionnaire resource consists of the following fields.:

resourceType describes the resource type, for our entity it is questionnaire.

status is used to indicate the current state of the questionnaire. Ranges between draft, active, retired, or unknown.

title states the title of the questionnaire.

type specifically states what type of questionnaire we are talking about. For our project it ranges between anxiety, depression, general_health and physical_activity.

item holds the questions for the questionnaire. It references to the QuestionnaireSchema seen in Figure 4.2 and each item has linkId, type, text, required and answerOption as their fields.

subjectType states what type of subject the questionnaire applies to. For our project it states patients. Patients is an entity in the IDPT but does not conform to the FHIR standard itself. However, to showcase the possibilities we added patients as the subjectType.

publisher states the name of the person or organization that is responsible for the questionnaire. In our project it is a reference to a user object in another database collection.

date states the last date the questionnaire was modified or last updated. In the case of none of the mentioned it states the date in which it was created.

repeats addresses if the questionnaire can be repeatedly answered or not. It can either have null, daily, weekly, monthly as a value.


```
_id: ObjectId('652970469e94a138e831d527')
resourceType: "questionnaire"
status: "active"
type: "depression"
subjectType: "Patients"
repeats: "weekly"
title: "General Health Questionnaire"
description: "Questionnaire for finding out the general health of a patient"
▾ item: Array
  ▾ 0: Object
      type: "wearable"
      _id: ObjectId('652970469e94a138e831d529')
      linkId: "1"
      text: "In the last week, how stressed have you felt?"
      required: true
    ▾ answerOption: Array
      ▾ 0: Object
          _id: ObjectId('652970469e94a138e831d52b')
          value: "Stress"
          label: "dataOption"
      ▾ 1: Object
          _id: ObjectId('652970469e94a138e831d52a')
          value: "week"
          label: "timeIntervalOption"
  ▸ 1: Object
publisher: ObjectId('63e1205324f7d531b6aadd4f')
date: 2023-10-13T16:28:54.659+00:00
__v: 0
```


Figure 4.3: A specific questionnaire as represented in the database.

The questionnaire entity implies the data input we want for each question. Every type facilitates the use of user self-report data except the type regarding wearable, which implies that the answer to this question should be passively collected in the background when a patient answers the questionnaire. The questionnaire entity facilitates the use of creating and fetching questionnaires with the use of GraphQL. The calls must be made to the backend to create a questionnaire resource or to be able to fetch the questionnaires. Other entities in the IDPT framework adhere to the use of createdAt, updatedAt, and id. We initially used that in our entity; however, we wanted to see the possibility of using FHIR standards that facilitate the use of date instead in our GraphQL schemas that relate to the endpoints addressed. It is essential to address that not all fields listed in the documentation of the FHIR standard for a questionnaire component are used; however, some fields were omitted since they create complexity and do not apply to the entity at the time of writing

4.3.3 Questionnaire Response

To store and create responses to the questionnaire entity discussed in the previous section we added another entity to the IDPT framework named questionnaireResponse. This entity holds the answers for each questionnaire question alongside other meaningful and related data. Each answer links to a question in the questionnaire and can hold input data or wearable data. Input data is data collected directly from the user, which can be either through text input, multiple choice, or checkboxes. Wearable data is collected when the input type for that question is set to the wearable value. A questionnaire response is closely linked to the questionnaire. The structure of a questionnaire response has certain familiarities with the questionnaire component. As mentioned this entity also conforms to the FHIR resource named QuestionnaireResponse and consists of the following fields.

resourceType describes the resource type, for this entity is is set to questionnaireResponse.

questionnaire is a reference to the answered questionnaire. In our project it holds an ObjectId which is a reference to the questionnaire in the database collection.

status is used to indicate the current state of response.

authored holds the date for when the response resource was sent.

item holds the answer for each question. Reference to another graphQL schema that holds the linkId of the question in the questionnaire, the type which is either an observation resource holding the wearable data or input from user as a string.

subject reference to the patient that this response belongs to.

meta holds information about the device the response comes from

date states the last date the questionnaire was modified or last updated. In the case of none of the mentioned it states the date in which it was created.

Since passive data collection measures vital signs, sleep patterns, or other objective patient data, we add an observation resource to the answer that holds the observation measure taken. More on the observation resource in the next section. Below we have figures showing how the questionnaire response looks like in our MongoDB database collection. And the different schemas the answer can hold.


```
_id: ObjectId('6529721ae8b0d3392e887d14')
resourceType: "QuestionnaireResponse"
questionnaire: ObjectId('652970469e94a138e831d527')
status: "completed"
subject: ObjectId('65294eddd8e5fb2c632c3f05')
authored: "2023-08-13T18:36:43.036+02:00[Europe/Oslo]"
▾ item: Array
  ▾ 0: Object
      _id: ObjectId('6529721ae8b0d3392e887d16')
      linkId: "1"
      text: "In the last week, how stressed have you felt?"
      type: "wearable"
    ▸ answer: Object
  ▾ 1: Object
      _id: ObjectId('6529721ae8b0d3392e887d15')
      linkId: "2"
      text: "In the last week, have you felt lonely?"
      type: "input"
    ▸ answer: Object
▾ meta: Object
    versionId: "1"
    lastUpdated: "2023-08-13T18:36:43.036+02:00[Europe/Oslo]"
    source: "Android"
__v: 0
```


Figure 4.4: A specific questionnaireResponse resource as represented in the database.

```
▾ answer: Object
    type: "FhirObservation"
  ▾ value: Object
      resourceType: "Observation"
      status: "final"
    ▾ code: Object
      ▸ coding: Array
        text: "Stress Data"
      subject: "65294eddd8e5fb2c632c3f05"
      effectiveDateTime: "2023-07-13T18:36:37.708+02:00[Europe/Oslo]"
    ▾ value: Object
      ▾ data: Object
          type: "Stress"
        ▸ value: Object
        origin: "Android"
```


Figure 4.5: A specific observation resource as represented in the database.


```
▾ 1: Object
    _id: ObjectId('6529721ae8b0d3392e887d15')
    linkId: "2"
    text: "In the last week, have you felt lonely?"
    type: "input"
  ▾ answer: Object
      type: "input"
    ▾ value: Array
        0: "yes"
      valueType: "string"
```


Figure 4.6: A specific input answer structure as represented in the database.

The questionnaire response is simple and complies to the FHIR standards for the given resource. However, when we created the answer schema, we had a few issues. The data collected through wearables can hold a lot of observations. To keep it simple and adhere to the standards we opted at adding all the measurements in a single object. This is not the ideal way to do it, each measurement derived can be seen as an observation and for future development one should investigate the possibility of doing that. However, due to time-constraints and possible complexity this would add to our project we opted on the solution discussed. This promotes interoperability and makes it possible for future development.

```
@Serializable
data class FHIRQuestionnaireResponse(
    val resourceType: String,
    val questionnaireId: String,
    val status: String,
    val subject: String,
    val authored: String,
    val item: List<FHIRQuestionnaireResponseItem>,
    val meta: FHIRMeta
```

Listings 4.2: Representation of the questionnaire response class in the mHealth application

```
@Serializable
data class FHIRQuestionnaireResponseItem(
    val linkId: String,
    val text: String,
    val type: String,
    val answer: AnswerData
)

@Serializable
data class FHIRMeta(
    val versionId: String? = null,
    val lastUpdated: String? = null,
    val source: String? = null
)
```

Listings 4.3: Representation of the questionnaire response item and meta class in the mHealth application adhering to HL7 FHIR standards.

Listings 4,2 and 4.3 shows how the questionnaire response is made in the demonstrative component discussed in Section 4.4 before it is sent to the backend with the use of a GraphQL mutation where it is retrieved and stored in the database. It must be created in this way to make calls to the endpoints set up for the questionnaire response entity. In addition, the answer data holds the type and value of the answer. However, as mentioned we have two different ways of adding answers, so the type is set to either “input” or “fhirObservation” so that the value either holds the self-reported data or the an instance of the FHIR observation resource as seen in Figure 4.5.

4.3.4 Health Connect

In this section, we discuss how we facilitated and set up the wearable data collection in our application. However, wearable data integration is a complex topic and one of the main issues we had to solve. At the start of our project, we wanted to directly integrate a smartwatch into our application using a Bluetooth connection with a Samsung Galaxy 4 watch. However, it became clear that better solutions existed. Even though it is possible to retrieve some raw data from APIs provided by Samsung, the documentation and development process for this was scarce.

Additionally, if we were to opt in on adding devices directly in the artifact through Bluetooth, we would have to integrate data fetching from different APIs. This means that if we or others want to develop the application further, there needs to be device compatibility for numerous devices, and this will make our artifact complex and not user-friendly because most manufacturers structure their data differently and provide multiple APIs that distinguish themselves from each other. This provides no standardized data structures and complexity in future development. Furthermore, all privacy components must be developed from scratch, and we did not have time for that since we programmed an Android application that none of the project participants had before; this is where Health Connect comes in.

Health Connect is a proprietary platform developed by Samsung and Google to facilitate the secure access of data from a set of apps and devices using APIs and a standardized data schema (Google 2023). The Beta version is available in the Google Play library and can be downloaded for use in developing applications. With Health Connect, we can both read raw and aggregated data and synchronize data from multiple platforms like Google Fit, Samsung Health, and Fitbit (Android Developers 2023). The best part is that Health Connect uses a standardized data format that simplifies health data acquisition. In other words, while data coming into the Health Connect platform has different data standards, the app conforms them into one standard for each data type.

Another vital aspect is that the data available is raw data or aggregated data, so we can get hold of the digital biomarkers we need through the platform. Health Connect can be an excellent way to streamline the data collection process in the future when it becomes widely accessible. The domain experts interviewed also states that a common platform for data retrieval could enhance the artifact and make it easy to further develop. Therefore, it is chosen as the preferred data collection platform for this study. In the future the platform can also be used to write data to that other health applications can read the data and use it in their applications.

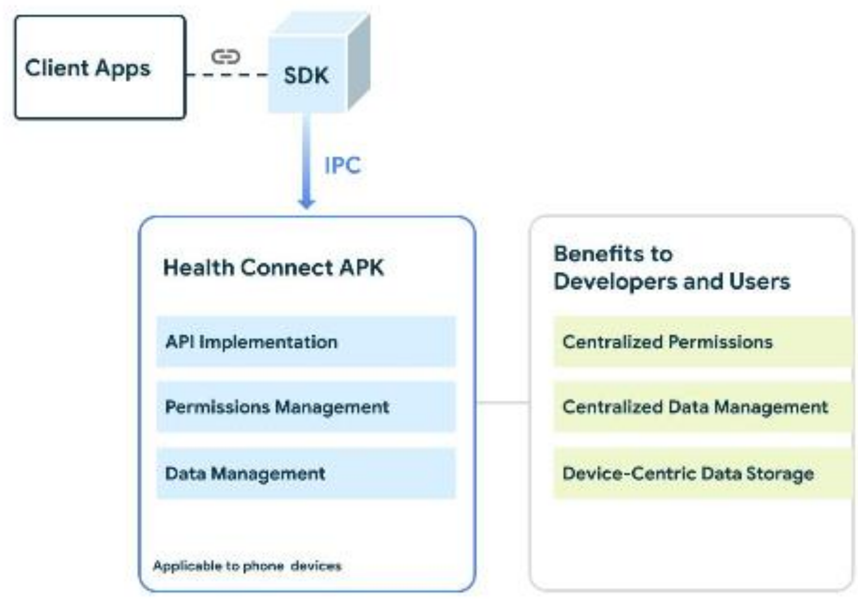


Figure 4.7: Platform architecture of the Health Connect platform (Android Developers 2023).

Health Connect architecture can be seen in Figure 4.7 and depicts the integration of a client app through the Health Connect API and the SDK layer. To use the Health Connect platform in our application, we had to add the Health Connect SDK to our dependency list, and then one can make calls to the API through a HealthConnectManager class. This client gives us easy access to permissions, data management, and the availability to read and potentially write the health data we need and ensure we can interact with the Health Connect API. Health Connect reads and writes data from the following well known applications: Fitbit, Samsung Health, Google Fit, Oura, MyFitnesPal, and Withings Health Mate.

Each application is well known and has one or multiple devices that is commonly used worldwide. This indicates that Health Connect could be a great tool to use and implement because market leaders and strong brands have chosen to integrate it with their applications.

Health Connect is still a beta app but will become a member of Android 14 as an Android system app (Android Developers 2023), and thus can be used as a tool to showcase the possibility of integrating it into novel approaches, which again can be developed further due to its mainstream access in Android 14. The Health Connect API is used in our demonstrative component to read raw and aggregated data when the question is to be answered with passive data collection. To implement Health Connect, one must add it to the dependency list as first mentioned and initialize a client that holds the different functions. It is essential to write out the permission for the data one wants to collect, as seen in Listing 4.2

```kotlin
// The minimum android level that can use Health Connect
const val MIN_SUPPORTED_SDK = Build.VERSION_CODES.O_MR1

class HealthConnectManager(val context: Context) {
    private val healthConnectClient by lazy { HealthConnectClient.getOrCreate(context) }

    val permissions = setOf(
        HealthPermission.getReadPermission(HeartRateRecord::class),
        HealthPermission.getReadPermission(HeartRateVariabilityRmssdRecord::class),
        HealthPermission.getReadPermission(SleepSessionRecord::class),
        HealthPermission.getReadPermission(StepsRecord::class),
        HealthPermission.getReadPermission(DistanceRecord::class),
        HealthPermission.getReadPermission(SpeedRecord::class),
        HealthPermission.getReadPermission(TotalCaloriesBurnedRecord::class),
        HealthPermission.getReadPermission(ExerciseSessionRecord::class),
    )

    val _isPermissionGranted = MutableStateFlow(false)
```

Listings 4.4: Permission set up for data collection types.

By adding these permissions and asking the user to permit the reading of these values from Health Connect, we get the data needed to create the answer for the passive data collection. The permissions can be turned on and off by the user in the Health Connect app or through the app we have created. Data stored and collected can also be seen in the app, and it allows the user to delete certain data types if they want. In addition, the user can go into the Health Connect app and delete or change which apps should have access to certain data types.

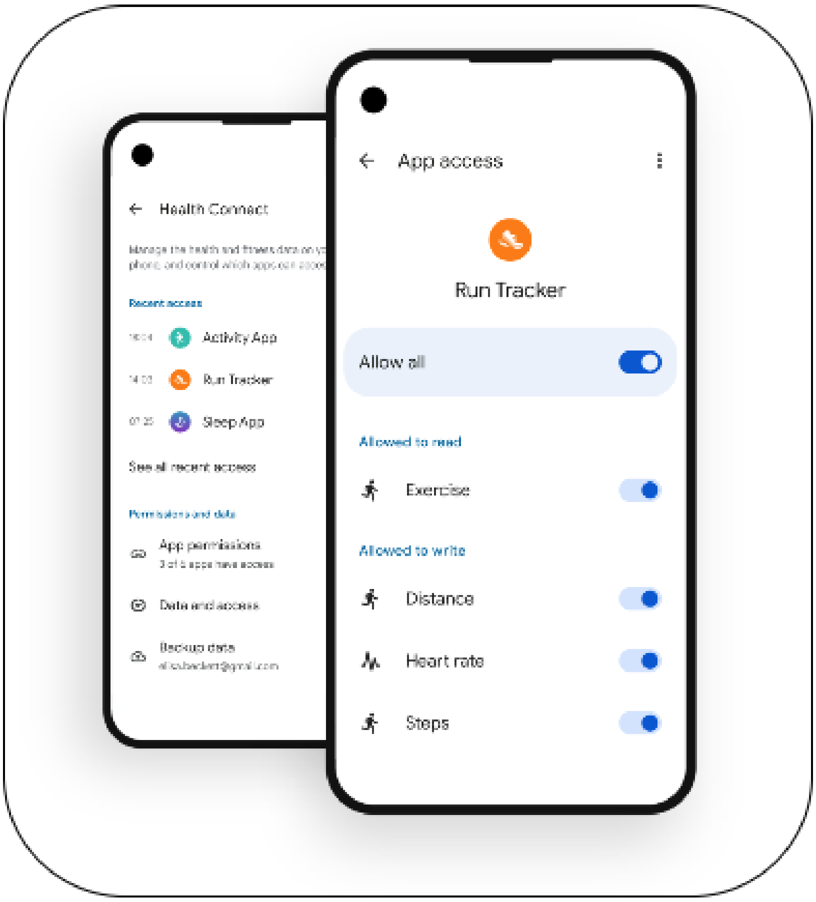


Figure 4.8: Representation of how a user can allow and deny measurements in Health Connect

To get the data, one must make a call to the Health Connect API with the use of the SDK dependency. Underneath is an example of how we read data from Health Connect. It is essential to state the record type, the data type one wants to retrieve, and the time range for the data that should be returned.

```
suspend fun readHeartRateVariabilityRecord(start: Instant, end: Instant): List<HeartRateVariabilityData> {
        val request = ReadRecordsRequest(
            recordType = HeartRateVariabilityRmssdRecord::class,
            timeRangeFilter = TimeRangeFilter.between(start, end)
        )
        val response = healthConnectClient.readRecords(request)

       return response.records.map { record ->
            HeartRateVariabilityData(
                heartRateVariability = record.heartRateVariabilityMillis,
                id = record.metadata.id,
                time = record.time,
            )
    }
```
Listings 4.5: Example of how data can be read for a specific measure with Health Connect.

HeartRateVariabilityRmssdRecord (
time: Instant,
zoneOffset: ZoneOffset?
heartRateVariabilityMillis: Double,
metadata: Metadata
)
Listings 4.6: Example of a HeartRateVariabilityRmssdRecord derived from Health Connect.

As seen in the listing over and Figure 9, a request returns a list of records or one record dependent on what one wants. The readHeartRateVariabilityRecord is a function from our application and showcases how data is retrieved. Additionally, each record is stored in an object containing the data for each record. We implemented this for each record listed in Figure 8, which provides a convenient way of quickly integrating wearable data collection into applications—more on how we store and send wearable data in Section 4.3.4.

4.3.3 Wearable Data Collection

The last general component regards how we structure the data we collect through Health Connect. Since the wearable data was added to the questionnaire response, we had to ensure that the data was easy to read. There is no standardized format for the wearable data component; the main goal was to see if the proposed solution was possible. However, with the addition of the observation resource we were able standardize the measurement but not the values retrieved from the observations. The data collected is structured into three different types of dependent on the information provided in the questionnaire. These specific types are sleep, stress, and physical activity. We collect measures as mentioned in Section 2.7 for each type.

Below one can see a representation of each data type. It holds measurements that can be used to find out the amount of stress, quality of sleep, and physical activity respectively. The data is retrieved from Health Connect and needs to be further analyzed to hold any value. Due to time constraints and effort needed this is outside the scope of this thesis, however, future development can be done to create or use a valid algorithm for the scoring of the data collected.

```kotlin
// Represents sleep data, raw, aggregated and sleep stages, for a
    given //[SleepSessionRecord].

@Serializable
data class SleepSessionData(
    val uid: String,
    @Serializable(with = InstantSerializer::class) val startTime:
    Instant,
    @Serializable(with = ZoneOffsetSerializer::class) val
    startZoneOffset: ZoneOffset?,
    @Serializable(with = InstantSerializer::class) val endTime:
    Instant,
    @Serializable(with = ZoneOffsetSerializer::class) val
    endZoneOffset: ZoneOffset?,
    @Serializable(with = DurationSerializer::class) val duration:
    Duration?,
    val stages: List<SleepStageData> = listOf(),
    val heartRateMetrics: HeartRateMetrics
)
```

Listings 4.7: Represents sleep data, raw, aggregated and sleep stages for a given [SleepSessionRecord]

```kotlin
@Serializable
data class SleepStageData(
    val stage: String,
    @Serializable(with = InstantSerializer::class) val startTime:
    Instant,
    @Serializable(with = InstantSerializer::class) val endTime:
    Instant,
    @Serializable(with = DurationSerializer::class) val duration:
    Duration,
    val metadata: String
)
```

Listings 4.8: Represents data for a given Sleep Stage inside a [SleepSessionRecord]. Holds both raw and aggregated data about the individual sleep stage

```
// Represents data, both aggregated and raw, associated with a single    exercise session. Used to collate results from aggregate and raw reads from Health Connect in one object.

@Serializable
data class ExerciseSessionData(
    val uid: String,
    @Serializable(with = DurationSerializer::class) val totalActiveTime: Duration? = null,
    val totalSteps: Long? = null,
    @Serializable(with = LengthSerializer::class) val totalDistanceKm: Length? = null,
    @Serializable(with = EnergySerializer::class) val totalCaloriesBurned: Energy? = null,
    val minHeartRate: Long? = null,
    val maxHeartRate: Long? = null,
    val avgHeartRate: Long? = null,
    @Serializable(with = CustomVelocitySerializer::class) val minSpeed: Velocity? = null,
    @Serializable(with = CustomVelocitySerializer::class) val maxSpeed: Velocity? = null,
    @Serializable(with = CustomVelocitySerializer::class) val avgSpeed: Velocity? = null,
)
```

Listings 4.9: Represents data both aggregated and raw, associated with a single exercise session.

```
@Serializable
data class StressData(
    val hrv: List<HeartRateVariabilityData>,
    val hr: HeartRateMetrics
)
```

Listings 4.10: Represents Stress data, collected from [HeartRateRecord] and [HeartRateVariabilitySmmdRecord].

All these data types can help to indicate a specific mental health issue, as discussed in Section 2.1.4. These data types are collected through API requests through the Health Connect SDK, making it possible to capture passive data retrospectively. Since the FHIR documentation does not provide a specific type for the smartwatch data, we needed a better way of implementing a standardized format for the wearable data answer option. In the future, this can be done; however, due to time constraints and the time-consuming part of creating the demonstrative component, we still need to implement a standardized way of adding the wearable data to the answer portion of the questionnaire response.

This works because when a user answers the questionnaire from the demonstrative component and reaches the question with the wearable type, the application collects the data from the Health Connect platform and continues the questionnaire. There are some pitfalls to this method; one is that the user has not synched their sensor data with the companion app, which means that Health Connect does not have the newest data, which could lead to no data collected. Additionally, if the user turns off permissions for one of the data types, it can lead to the same outcome. However, solutions are in place in the demonstrative component to solve these issues. The data collected are temporarily stored and added to the questionnaire response when submitted. Data is then added as JSON data to the answer parameter.

Since we collect raw and aggregated data about vital signs and other objective measurements, we opted to add the FHIR HL 7 resource type observation when wearable data is collected and added to the response. Each observation holds all the data for one of the domains under. In the documentation HL 7 states that each measurement should have their own observation resource, however, since the data collected can be hundreds of observations, we decided to add the data in a single observation each time data is collected. This might not be the preferred approach; however, it makes it possible for integration and sharing of data in the future, so we see it as a suitable starting point. Underneath is a representation of each answer for the specified domain.

Stress Data


```
▾ value: Object
  ▾ data: Object
      type: "Stress"
    ▾ value: Object
      ▸ hrv: Array
      ▾ hr: Object
          startTime: "2023-08-06T16:36:35.775Z"
          endTime: "2023-08-13T16:36:35.775Z"
          bpmMax: 133
          bpmMin: 41
          bpmAvg: 65
          measurementCount: 9883
        ▸ sourceAppInfo: Object
    origin: "Android"
```


Figure 4.9: Representation of the values collected for the stress data measures derived from the database.

Sleep Data


```
▾ value: Object
  ▾ data: Object
      type: "Sleep"
    ▾ value: Array
      ▾ 0: Object
          uid: "9466b92a-53ad-3ca3-ae3c-15c7d584f39e"
          startTime: "2023-08-12T17:09:35.567Z"
          startZoneOffset: "+02:00"
          endTime: "2023-08-13T17:09:35.568Z"
          endZoneOffset: "+02:00"
          duration: "PT8H32M"
        ▸ stages: Array
        ▾ heartRateMetrics: Object
            startTime: "2023-08-12T23:56:00Z"
            endTime: "2023-08-13T08:28:00Z"
            bpmMax: 68
            bpmMin: 42
            bpmAvg: 49
            measurementCount: 512
          ▸ sourceAppInfo: Object
    origin: "Android"
```


Figure 4.10: Representation of the values collected for the sleep data measures derived from the database.

Exercise Data


```
▾ value: Object
  ▾ data: Object
      type: "Exercise"
    ▾ value: Array
      ▾ 0: Object
          startTime: "2023-08-07T18:04:17+02:00"
          endTime: "2023-08-07T18:58:54+02:00"
          id: "7176dd5a-8b16-30ad-a4e6-a97dd2b9568c"
          typeOfExercise: "EXERCISE_TYPE_WALKING"
        ▾ sourceAppInfo: Object
            packageName: "com.fitbit.FitbitMobile"
            appLabel: "Fitbit"
        ▾ sessionData: Object
            uid: "7176dd5a-8b16-30ad-a4e6-a97dd2b9568c"
            totalActiveTime: "PT54M37S"
            totalSteps: 5900
            totalDistanceKm: 4.3054749879675605
            totalCaloriesBurned: 469.30788305494525
            minHeartRate: 80
            maxHeartRate: 126
            avgHeartRate: 109
      ▸ 1: Object
```


Figure 4.11: Representation of the values collected for the physical activity measures derived from the database.

4.4 Demonstrative component

In this subsection, we delve into the demonstrative component that serves as a demonstration of our work. It showcases how the mHealth application developed collects passive data through a wearable device, core functionalities, user interface, and data storage. Additionally, it addresses the additions made to the existing IDPT framework to showcase how the two systems work together. The general components discussed in Section 4.3 serves as building blocks for the demonstrative component, and the demonstrative component shows how they are applied in practice.

We start by showing an overview of the additions to the IDPT that serve our demonstrative component. It is followed by an introduction to the mHealth application made for answering the questionnaires, where we will address how and why it is developed as it is. Furthermore, a discussion around the user interface is done, and we highlight a study conducted to test the usability and effectiveness of the artifact.

4.4.1 Additions to the Existing Framework

In Section 4.3, we discussed the general components of our artifact. That involved the addition of two entities and logic related to these to facilitate data storage and management. That is indeed an addition to the existing framework; however, to facilitate the visualization of these, we had to create separate views modules for the questionnaire and the response. As mentioned earlier in the thesis, the questionnaire could be a sub-group in the assignment entity; however, with the addition of FHIR resources for the questionnaire entity, we created a separate view for this. Below is a screenshot of the questionnaire view.

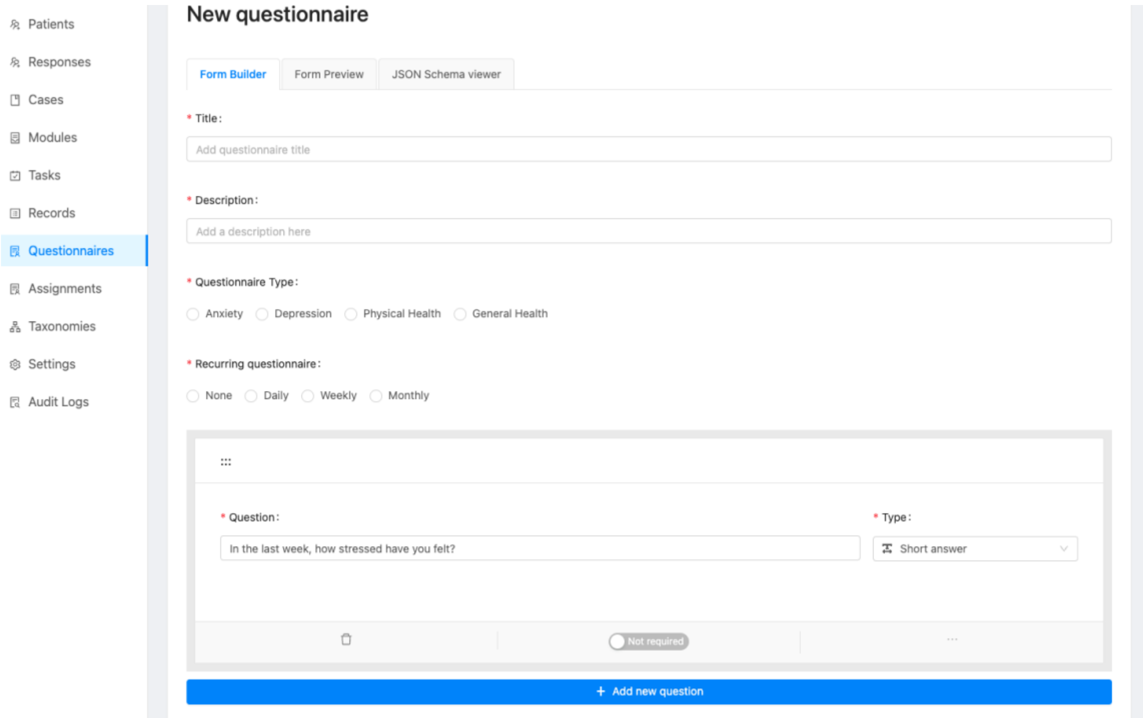


Figure 4.12: A screenshot showing the questionnaire creation view in the IDPT framework.

In this view, therapists can create questionnaires. The type holds input, multiple choice, checkboxes, paragraphs, and wearable dat. The wearable type should be used for passive data collection instead of self-reporting for the question. The view itself looks like the assignment view, apart from the wearable functionality and the fact that it creates an FHIR resource type. When selecting the wearable option, it will require the therapist to enter both indicators, either sleep, stress, or physical exercise, and the timeframe in which the data should be collected. It is important to note that data is collected in retrospect so if one chooses day as value it collects data from the last day. Figure 16 shows an example of how it looks like when smartwatch data is selected.

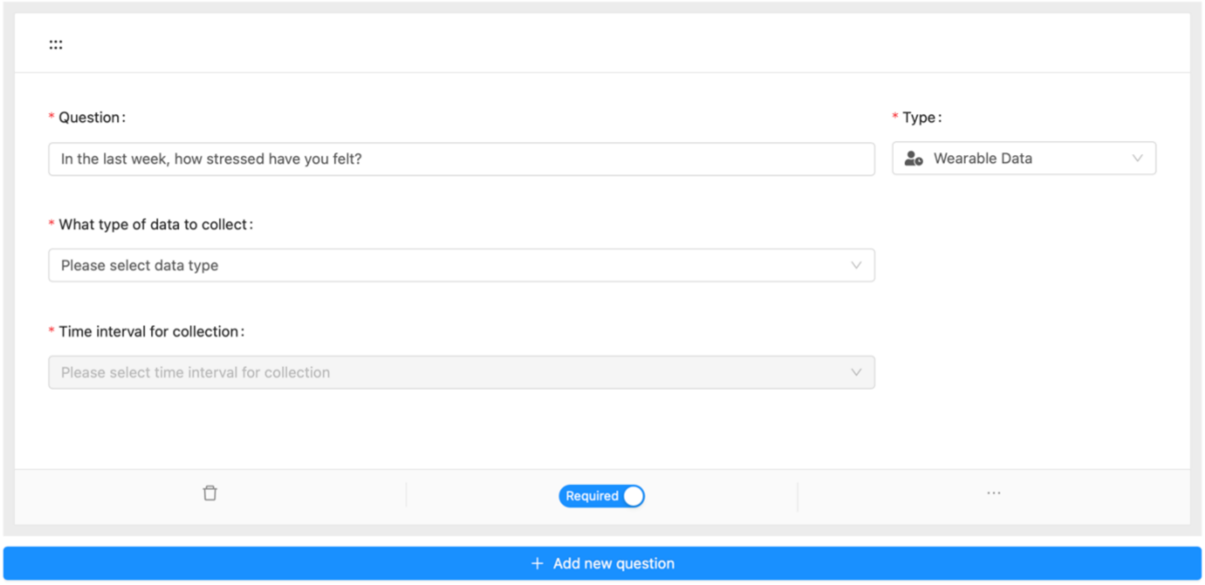


Figure 4.13: A screenshot showing the question card in the IDPT framework.

A view is implemented to show the questionnaire response to the therapist/clinician. This shows all the data related to the response and data is not manipulated or analyzed. For future work, scoring of the data collected is crucial. There are no changes to the architecture of the IDPT when adding the views, they behave the same way as the other entities. Figure 4.12 shows how the response view looks like and the answer can be viewed when clicking on it.

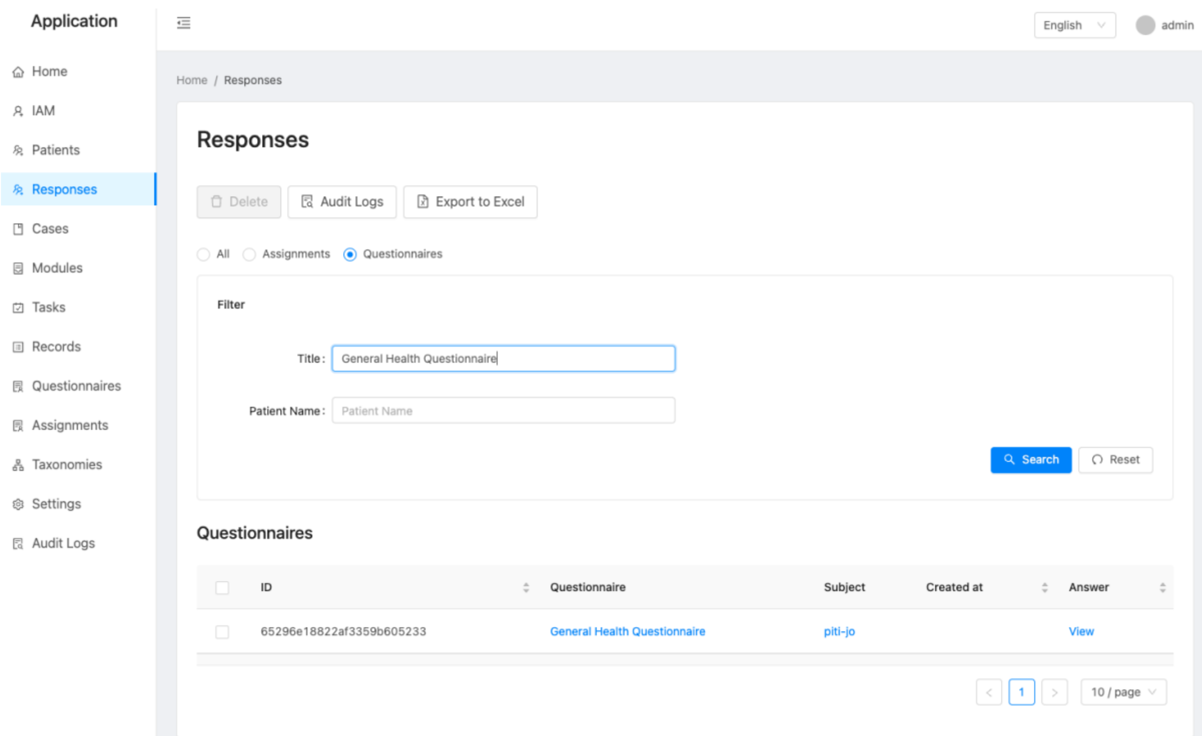


Figure 4.14: Dashboard where the clinician or therapist can see all the responses for the questionnaires.

4.4.2 Creation of the mHealth Application

In this section, we address and show how the mHealth application that serves as the main demonstrative component for our artifact works. This application retrieves and sends data to the IDPT's backend for fetching and storing questionnaires and questionnaire responses. It is an Android application created in Android Studio with Kotlin. The application was developed after implementing the two new entities in the IDPT framework. It is a minimum viable product (MVP) to showcase a possible way of implementing wearable technology with self-assessment questionnaires to provide a better tool for mental health assessments. The first thing we did was to make sure that we could retrieve the questionnaires from the backend, which required the possibility of signing up and logging in to the application. For the sake of the MVP, if a user registered, a profile was created in the IDPT framework with the patient role. This means that they can only access data and other tasks that belong or are assigned to them through a clinician/therapist. How the passive data collection logic works is previously discussed in Sections 4.3.3 and 4.3.4, and for the rest of this section, we will not address what has already been stated in the thesis.

Creating the Android application was tedious and challenging since we needed to gain experience with mobile application development within the Android ecosystem. At first, we looked at the possibility of creating a React Native app; however, since we wanted to take advantage of the Health Connect platform, we needed to utilize Android Studio and develop the application in Kotlin. After thoughtful consideration, we ended up making an Android application with the use of these technologies and frameworks mentioned in Section 4.1.

4.4.3 User Interface and Functionalities

For this section, we want to discuss and show how the mobile application's user interface has been developed. Since the functionality and proof of concept were more critical than the design of the user interface, we have yet to spend substantial time on this. However, we have created a fully working application with a user interface. To get help, we asked a professional UX designer for input and feedback after the first implementation of the user interface. We showed a quick demo of our application and sent over screenshots of each screen and functionality. The UX designer provided some feedback to help us in our design implementation.

Additionally, after we had created a functional minimum viable product (MVP), we sought to create a pilot study to check if the created artifact could have potential benefits in the problem domain discussed in Section 1.2. The conducted study is discussed in Section 5.2. We had three main objectives for the design of the application:

(i) Ease of use
(ii) Dynamic and visualized questionnaire answer process
(iii) Visualization of results to make users want to come back.

Let us first describe the screens in the application before we discuss and address the answer process and the visualization of the questionnaire response in detail. Navigation using Jetpack Compose works in a way that the MainActivity initializes a navigation host that creates a composable for each screen. Each screen has a composable, a view model, and a view model factory. The latter is responsible for injecting necessary instances of a class into the view model. The view model factory and nav host also help us with the separation of concern and make it possible to inject instances of a class. This uses the singleton pattern, which is excellent for the application as that gives it less complexity.

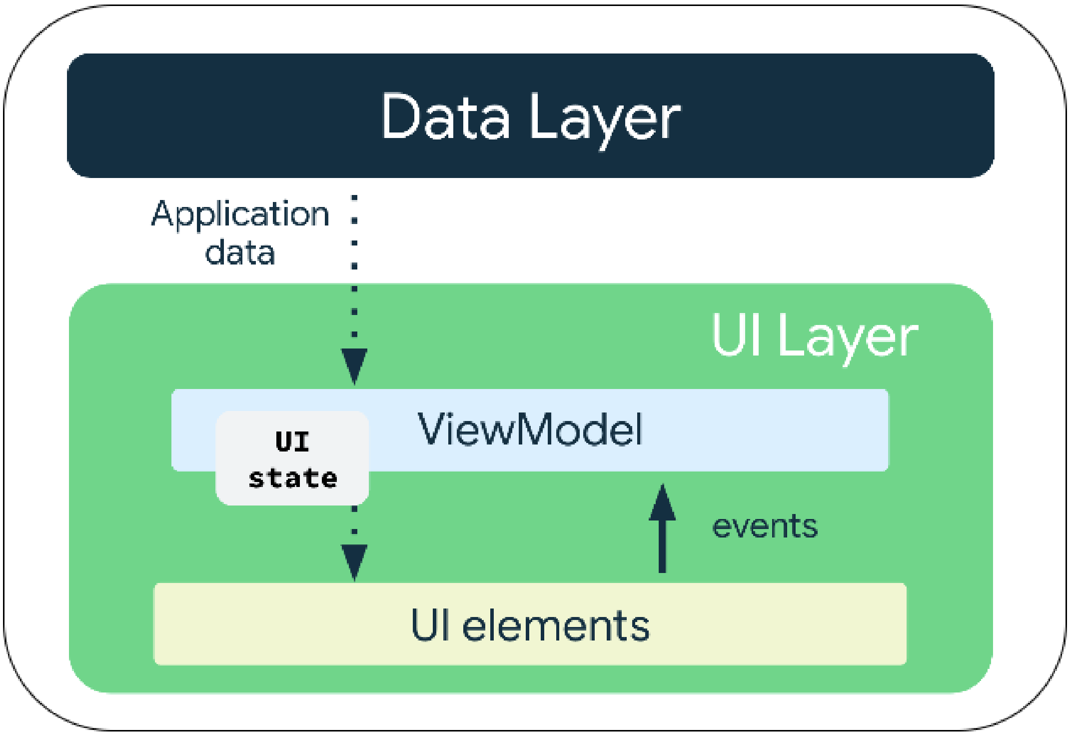


Figure 4.15: Retrieved from Android Developers. It shows how a normal application works with Jeptack Compose.

As seen in Figure 18, this is a generic example of how a Jetpack Compose app works. Our application works the same way: it has a view model for each composable, and UI elements are sent to the view model, which updates the UI state based on the events. The data layer is responsible for sending instances to the view model so that it can be used to update and change the UI composable. Let us address each screen in the application.

Login screen – The login screen is responsible for the login part of the application. If the user has no profile, they can access the register screen.

Register Screen – The register screen is responsible for the registration of new users. It performs a GraphQL mutation to the IDPT's backend to create a new user.

Setup Screen – The setup screen gives information about data processing, storage, collection, and notifications. It provides a helper for setting up Health Connect and permission for collecting user health data—more on that in Section 4.4.5.

Home Screen – This is the main screen, showing two buttons that take the user to the questionnaires or the result.

Privacy Policy Screen – States the privacy policy for the application concisely with different interactable cards for the user.

About Screen: Tell the user about the application and its intended purpose.

Settings Screen: A screen that allows the possibility of deleting the account and changing permission for the app and other apps in Health Connect. In addition, users can toggle notifications here.

Questionnaire Screen: A screen that shows all available questionnaires. If a questionnaire has been answered, it is placed in the completed tab. We have a scheduled task that runs every minute to check if the questionnaire can be answered again based on the frequency given by the therapist. If it is available again, it is displayed in the ready tab, and a notification is sent to the user to prompt them to answer.

Questionnaire Answer Screen – The core composable of the application. It serves as a dynamic screen where users can answer the provided questionnaires through self-report and passive data collection.

Result Screen – This screen shows all the questionnaire responses for the user. It expands once clicked and shows detailed information about the passive data collected and the answer provided by the user.

One of the crucial things to consider was the implementation of the functionality for answering questionnaires. When first developed and designed, it did not adhere to user-friendly design principles. Feedback from the UX designer stated that it should be fixed to get users to come back to the application since a good first impression is crucial. To show the process, let us use an example questionnaire.

To showcase how it works, we created a sample questionnaire with the following questions alongside a screenshot of the UI for each question:

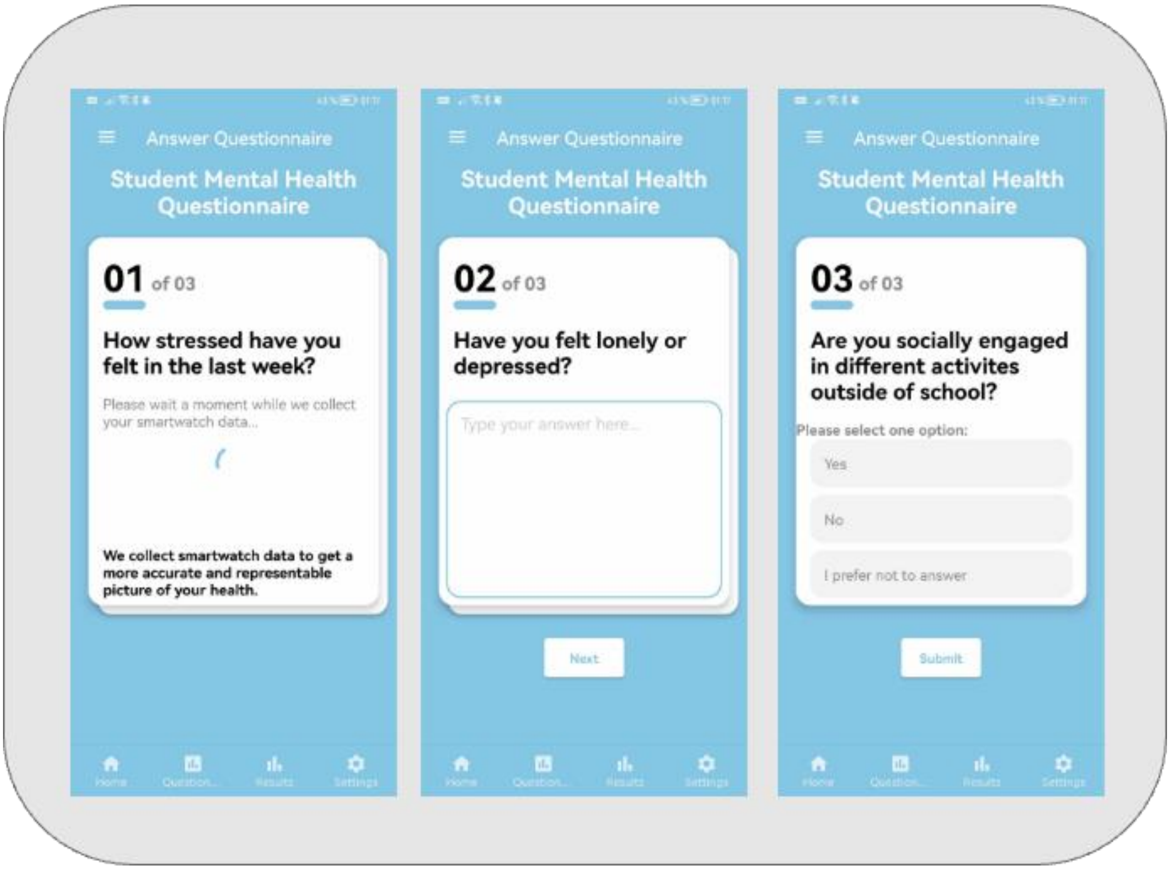


Figure 4.16: Representation of the answer UI for different question types.

Each question has a different type. The first one is set to wearable input to capture passive data from the user. The second is regular input from a text field, while the third question is multiple choice. This shows how our answer component works and the visualization of the answering process that may help to get more answers and combat the response burden in traditional self-assessment questionnaires. One can see that even though data is collected passively, we still want to show the question to the user and display information so that they understand what is going on. At first, we had the app to collect data passively without showing the question; however, after feedback from a UX design expert, it became clear that it is crucial to show and display information about what is done to the user, so they understand what is going on.

Another essential aspect we wanted to add to our UI was the visualization of answers. Each user has control and can visualize their answers. In this section, they can also delete the answer if they want. This is vital because we want to engage the user and showcase their data. This will likely lead to better engagement and less social desirability bias because it motivates them. Underneath, one can see screenshots of all the different results screens.

Figure 4.17: UI for the visualization of the collected data for each type.

This figure shows how the visualization of data works. Depending on the type of measured information about that data, it is displayed to the user once they click on the corresponding result card. This will hopefully engage the user and encourage them to answer the questionnaires again so that they also can compare and see their changes over time.

Notifications are another thing added to the functionality of our application. Our goal is that it may lead to a more engaged user. Since the questionnaire can be repeated after a specific time, we wanted to look at the possibility of sending a simple notification to the user. A notification channel sets up the notification itself, and the permission is set to true when the app is installed. However, in the start-up phase, the user is prompted with the information and is given the possibility of turning it off. A coroutine worker is created and schedules a check of available questionnaires at a specific time each day. The completion time in milliseconds and when the user can retake the questionnaire in milliseconds are stored locally in a Room database, and the scheduled worker checks if the questionnaire can be answered again. If yes, a notification is displayed to the user, prompting them to answer the questionnaire again. The notification can be clicked, taking one to the screen where all available questionnaires are listed.

4.4.5 Data Processing and Storage

In this section, we will look at the data processing and storage of the data, with a focus on the potentially sensitive data collected from the user's wearable. As discussed in Section 2.1.5, there exist challenges and limitations to the use of wearable technology in health. One of them is the processing of data and how it is stored. As mentioned, our data is collected either from the user itself or from the Health Connect platform. The Health Connect platform retrieves the data from the companion applications for the devices used. As mentioned in the previous Section, the response adheres to the FHIR standard for a questionnaire response resource; however, the wearable data in the answer is not standardized. For future work, this should be investigated. Moreover, future applications must promote standardization of answer data. However, wearable technology is still an emerging tool in the health sector; no commonly used standardized data format exists at the time of writing.

Data is retrieved from Health Connect and temporarily stored in a mutable state map while the patient answers the questionnaire. On submission, each answer is added to the corresponding questionnaire response item. Before sending it to the backend, the item part that consists of self-report data and passive data collected from the smartwatch is encrypted. The encryption algorithm itself could be better since the key used is the same and can be found both in our artifact's backend and front end. Moreover, this is to show how it should be done, and for the future, better encryption algorithms and transport layer security should be added to conform to privacy and GDPR. Each questionnaire response is decrypted when received and stored in the backend and can only be viewed by users with the administrator role in the IDPT framework. According to research, for individuals to share their sensitive health data, it is essential to understand how it is stored and enhance openness about the data process of sensitive health data (Genevieve et al. 2019).

The privacy policy describes how data is stored and what it is used for concisely. Permission to collect the passive data is asked when a user enters the application and can be changed later in settings. A complete demo of the artifact can be found by following the URL in Appendix D. Future development should address the privacy issues regarding GDPR and the secure sharing of data.

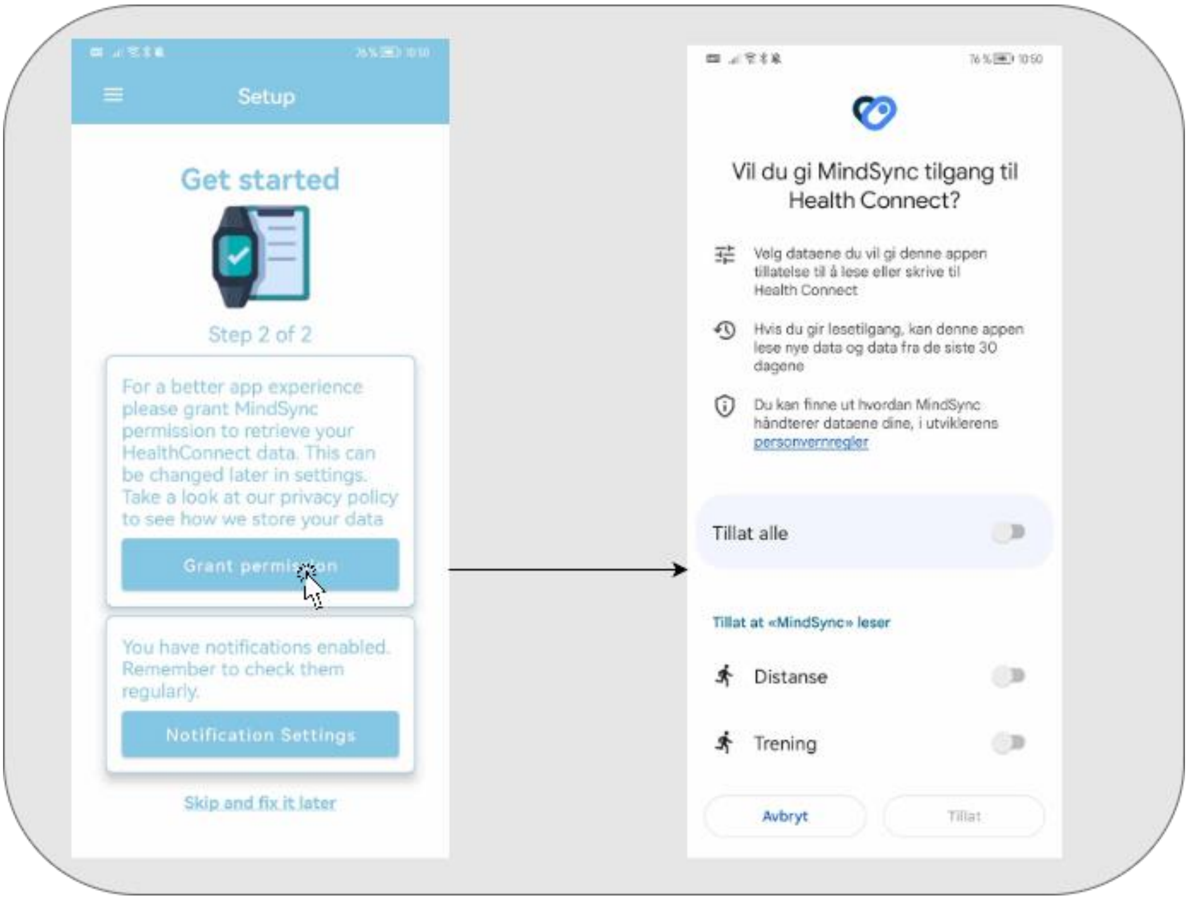


Figure 4.18: A representation of how permission for the collection of data is handled by the UI.

Chapter 5

Research Evaluation

This section describes the evaluation process of our artifact. An outcome-based evaluation study was conducted to see if there were any differences between the answers when a participant answered a set of questionnaires both traditionally and using our artifact. Additionally, a semi-structured interview with domain experts and a user acceptance survey were conducted.

5.1 Interviews

Two semi-structured interviews were conducted with two different domain experts after the completion of the artifact. The one interviewee was a domain expert in psychology and mental health and had prior experience with using traditional self-assessment methods in the field of mental health. The second interviewee was a domain expert in software engineering and had previous experience monitoring patients with wearables and other technological devices.

The semi-structured interviews conducted is a qualitative evaluation method. In software engineering, developed artifacts have certain qualitative features that facilitate qualitative research evaluation in the specified domain. The semi-structured interviews are a qualitative evaluation method where valuable feedback can be retrieved from domain experts in the problem domain. The interview is used to gather data from the experts; the benefit of this method is that while the domain expert answers set questions from the interviewer, it facilitates the possibility of the interviewee to come up with ideas. In contrast, the interview is conducted (Adeoye 2021).

The interview process started with a description of the artifact developed by the domain experts. This included a demo of the developed application and the extensions of the IDPT framework. It was followed by the interview process, where we asked the domain experts our questions (See Appendix A). Since it was a semi-structured interview, it included additional discussion and feedback outside of the question scope. However, since it was a semi-structured interview, it is valuable to get information that is derived from discussion outside of the question scope stated by the interviewer. The questions and summarized and answers can be found in Appendix A.

5.1.2 Interview Results

During and after the interview process, the domain experts stated that the artifact developed is a step in the right direction; however, as it is a novel approach, some questions of concern were raised.

The domain experts mentioned several positive aspects of our artifact that show a novel way of integrating wearable technology into traditional mental health assessments. The domain experts were asked to state the potential benefits of this integration, including the possibility of providing objective data that can mitigate biases often associated with self-assessment questionnaires. In addition, the experts stated the importance of a more accurate representation of a patient's current physiological state and the potential for faster diagnosis and assistance. However, they also mention that this is a novel solution or a hypothesis to solve a problem in the stated problem domain, and further research is needed to conclude on the validity and accuracy of this tool. An additional benefit emphasized in the interview process was the point of patient engagement. Research suggests that patient engagement can lead to a more motivated patient and can benefit the patient and the therapist (Marzban et al. 2022). Moreover, he domain experts address that this can lead to better patient involvement in the care process.

The domain experts addressed that the measured digital biomarkers are a good starting point since research state that they correlate with mental health issues. However, ubiquitous commercialized wearable devices used in patient monitoring are a novel approach that must be validated and tested extensively before widespread use. One domain expert state that it is also vital that the data should be raw and not some numbers that are the output of an algorithm before therapists and clinicians retrieve it.

Additionally, the domain experts raised some concerns about the artifact that could potentially hamper its use. Naturally, using wearable technology and emerging technologies in mental health is a relatively new concept, and more research needs to be conducted to see how patients embrace it. Challenges mentioned during the interview were data privacy, GDPR compliance, informed consent, the need for data quality and accuracy in the passively collected data, and other ethical concerns. Furthermore, the domain experts stated concerns about synchronization of data, standardization of the collected data, and interoperability. The last two go together because, as the domain experts state, it is easier to instantiate interoperability between systems with standardization. Trustworthiness of the artifact and how data is handled and stored where other concerns are raised.

The semi-structured interview gave us a good understanding of the need for future research in the problem domain. Moreover, the domain experts interviewed imply that this solution can contribute to the knowledge base. However, further research and evaluation is needed (see Section 7.2).

5.2 Outcome-Based Evaluation

To properly evaluate the artifact, we conducted an object-based effectiveness evaluation. Outcome-based evaluation is a systematic and structured process of assessing the effects, changes, or accomplishments of a program, project, or intervention (Schalock 2001). For our artifact, we want to check the effectiveness of the created system by trying to understand and see if it fulfills the objectives listed in Section 5.3.1.

5.2.1 Objectives

The pilot study was conducted to see if the answers given to the questionnaire differed when using the artifact versus the traditional pen-to-paper method. Moreover, we came up with two objectives for our study:

(i) To assess whether the artifact influences questionnaire responses.
(ii) Is there a significant difference between the two methods

These objectives, combined with the evaluation, can represent a culmination of our research effort, and provide insights into the real-world impact of our artifact.

5.2.2 Participants

The app and the IDPT framework are not commercially available. Therefore, participants in this effectiveness study are individuals directly related to us. This approach was chosen for several reasons:

Accessibility: The app and the IDPT framework were not accessible to the public due to their research-oriented nature. In addition, an Android smartphone and a compatible watch were needed. Therefore, individuals closely connected to the researcher allow us to obtain feedback and get the needed data within the constraints of the study.

Controlled Environment: Selecting participants who knew the research and the research topic could help us get the data we needed.

Ethical Considerations: Involving individuals with a direct connection to us, we could ensure that we maintained high transparency and ethical rigor in obtaining informed consent and protecting participant privacy.

While we were limited to individuals who knew us and the constraints that an Android phone and corresponding smartwatch were needed, it is essential to highlight that some degree of bias may be seen in the evaluation results because they know the content of the research and can alter their answers to get a favorable outcome. Future research should focus on and evaluate a broader audience to assess the artifact's generalizability and real-world impact.

5.2.3 Data Collection and Analysis

After finding 10 participants and downloading the application to their phones, we ensured they were guided through the setup process and that everything was in place. Additionally, we ensured that all necessary conditions for data collection were met, and participants were given clear instructions on using both the traditional pen-to-paper method and the artifact in question. The questionnaire provided can be seen in Appendix B. The questionnaire use was explicitly made for this experiment to see if there are any indications on whether the answers provided by the user and the passive-data collection differ, which can indicate that it has the potential to eliminate response biases and response burden.

Data Collection Process:
The data collection process spanned one week to allow the participants sufficient time to complete both the traditional and digital versions of the experimental questionnaire. The questionnaire is listed in Appendix B. It is not a commonly used questionnaire and is made for use in this experiment. The goal is to provide questions that can show if the answer provided significantly differs when applying the two different methods. These can give indications of the promise of our developed artifact. The process was structured as follows.

> **Traditional Method (Pen-to-Paper):** Participants were initially instructed to complete the questionnaire using the traditional pen-and-paper method. They were provided with hard copies of the questionnaire and a designated completion timeframe.
>
> **Digital Method (Artifact**): After completing the traditional method, participants were guided through using our application to answer the PSS questionnaire digitally. They were asked to follow the same timeframe for completion. Additionally, each time it was time to answer a questionnaire, a notification was sent to remind the participant through the application.
>
> **Randomized Order:** To mitigate order effects and potential biases, the order in which participants completed the traditional and digital methods was randomized. Half of the participants started with the traditional method, while the other half began with the digital method.

Data Analysis:
Data analysis was used to assess the differences and potential advantages of using our artifact compared to the traditional method. This includes both quantitative and qualitative measures. We compared the scores with emphasis on the question where we collect passive wearable data instead of user self-report and see if it differs significantly from the self-reported answer. No data analysis is implemented in the artifact. However, the questions asked can be analyzed on their own. When asking a stress question, lower HRV can indicate certain moments of stress (Kim 2018), and the other questions regarding physical exercise and sleep provide aggregated data such as time asleep and awake, which can directly showcase if it is any biases in the participant's answers. However, for this evaluation the primary goal was to see if the questions in general differ between the two methods discussed. For the calculations, we created a Python program that gave us the mean

difference rounded, the p-value rounded, and the effect size rounded using Cohens D after plotting the numbers as seen in the Tables in Section 5.2.4.

Mean difference: Is used in this evaluation to get the mean value of each method and then see the mean difference between the two.

P-value: We choose to use the p-value as a measure because it can help to indicate if one method is better than another method for an entire population (Singh 2013) If $p < 0.05$ means that the difference between the two hypothesis are significantly different and one can reject the null hypothesis (Singh 2013).

Cohens D: Measures effect size and is used to measure if it is a significant difference between the self-reported answer and the wearable data collected for a specific question.

5.2.4 Results

Results show that questions that utilize passive wearable data collection show a significant difference from the answers provided by the individual through self-report. Let us go through each question that tried to investigate this. Because of time constraints, no scoring of the data sent in is developed. Therefore, the questions discussed hold either raw or aggregated data that is easily measured with the response from the user through the traditional method.

2. On average, how much do you sleep each day (in hours and min)?

Participant	Round	Traditional Method	mHealth Application
1	1	7 hours and 20 min	6 hours and 45 min
1	2	7 hours and 45 min	7 hours and 25 min
2	1	6 hours and 10 min	6 hours and 0 min
2	2	8 hours and 5 min	7 hours and 43 min
3	1	7 hours and 30 min	6 hours and 50 min
3	2	8 hours and 0 min	7 hours and 45 min
4	1	6 hours and 0 min	5 hours and 50 min
4	2	6 hours and 30 min	6 hours and 18 min
5	1	7 hours and 30 min	7 hours and 10 min
5	2	7 hours and 30 min	7 hours and 7 min
6	1	8 hours and 0 min	7 hours and 38 min
6	2	8 hours and 30 min	8 hours and 2 min
7	1	6 hours and 45 min	6 hours and 10 min
7	2	7 hours and 0 min	6 hours and 49 min
8	1	5 hours and 30 min	4 hours and 54 min
8	2	7 hours and 0 min	6 hours and 42 min
9	1	8 hours and 30 min	8 hours and 25 min
9	2	8 hours and 0 min	7 hours and 45 min
10	1	7 hours and 20 min	6 hours and 52 min
10	2	7 hours and 20 min	6 hours and 38 min

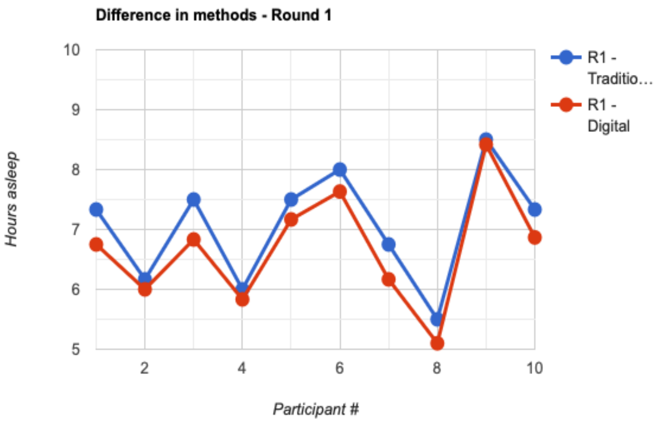
Figure 5.1: Represents answers given by the participants in each round for question 2.

Round 1

Mean Difference: 0.9
p-value: 0.001662
Effect Size: 2.846

Round 2

Mean Difference: 0.5
p-value: 0.049
Effect Size: 2.236



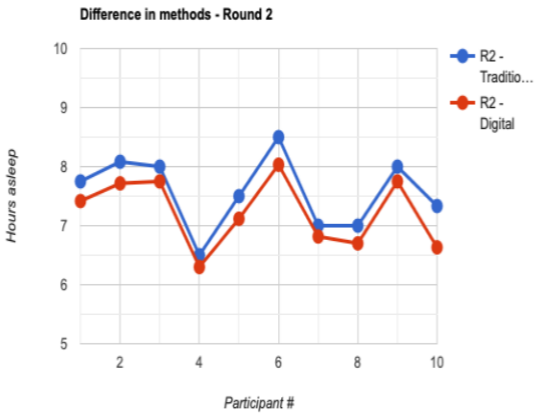


Figure 5.2: Line graphs depicting the difference between the two response methods for question 2.

Both rounds indicate that there is a significant difference in values between the traditional pen-to-paper questionnaire and the artifact developed (p-value less than 0.05). In addition, the effect size is large and that can indicate that the artifact can indeed help in creating a more objective and reliable health assessment tool.

5 In the last week, how many times have you engaged in physical activity?

Participant	Round	Traditional Method	mHealth Application
1	1	2	4
1	2	5	5
2	1	0	1
2	2	0	0
3	1	2	4
3	2	1	3
4	1	1	3
4	2	1	2
5	1	3	4
5	2	3	5
6	1	5	6
6	2	0	2
7	1	2	3
7	2	1	1
8	1	2	4
8	2	2	4
9	1	1	3
9	2	3	5
10	1	1	2
10	2	1	2

Figure 5.3: Represents answers given by the participants in each round for Question 5.

Round 1
Mean Difference: 1.4
p-value: 0.000894
Effect Size: 1.086

Round 2
Mean Difference: 1.3
p-value: 0.00037
Effect Size: 1.050

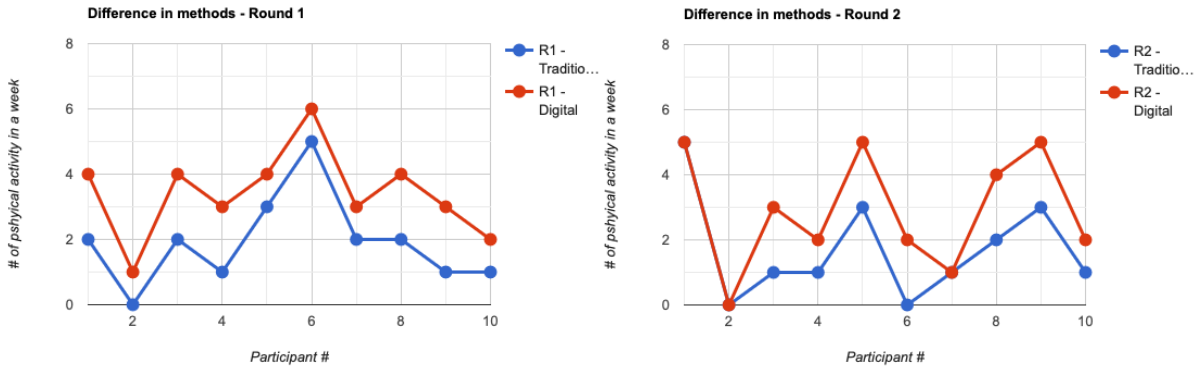


Figure 5.4: Line graphs depicting the difference between the two response methods for question 5.

Both rounds indicate that there is a significant difference in values between the traditional pen-to-paper questionnaire and the artifact developed (p-value less than 0.05). In addition, the effect size is large and that can indicate that the artifact can indeed help in creating a more objective and reliable health assessment tool.

5.2.4 Conclusion

The results show a strong indication that the artifact developed can be influential. Moreover, let us answer the objectives we wanted to investigate in this outcome-based evaluation.

To assess whether the artifact influences questionnaire responses.
The artifact influences questionnaire responses. Section 5.2.3 shows a significant difference between answers from the patient's recollection and the objective data collected through wearables. The answers differ significantly between the two methods, which can indicate that this approach can help solve the social desirability bias problem.

Is passive data collection for a specified question a good approach?
This is important because it lays the foundation for future research; our evaluation shows that passive data collection can provide more accurate and reliable information than traditional methods. Both the p-values can be seen as significant since it is less than 0.05 and the effect size is large, which indicates a substantial difference.

This indicates that the artifact can serve as a good contribution to the problem domain. However, more testing and evaluation must be done to conclude the matter. The study focused on the difference between the answers provided by the two methods, and future research should focus on more generalization of the study participants and measure its significance. A whole thesis can be concluded on the quantitative testing of this artifact, this evaluation was conducted to investigate if the artifact can pose as a viable solution in the problem domain.

5.3 User Acceptance

A user acceptance survey was given to the participants in the evaluation described in Section 5.2 after the completion of the study. It was provided to investigate what potential users think about the developed artifact and to see if we can potentially draw some conclusions on how the demonstrative artifact works. The user acceptance survey can serve as a good measurement of what potential users think about the developed artifact. The survey can be seen in Appendix C and uses both qualitative and quantitative methods. It is comprised of open-ended questions and questions with predefined answers so that it can give us valuable information about the potential of a digital platform for answering questionnaires.

5.3.1 Survey Results

The survey uses a combination of qualitative and quantitative evaluation methods. Let us investigate the overall significance and the questions that imply something valuable to the research conducted. The demographics are not really of the essence now since we have a homogenous group of participants due to the limitation that they all know the researcher.

Two-thirds of the participants state that they have used a mental health assessment tool before, with the majority stating the use of traditional pen-and-paper and online surveys or forms as the method used. On a scale from 1 to 5, the participants, on average, gave the application a 4 when it comes to user-friendliness. Some participants wanted to know where their data was sent and how to access the collected data. One person specifically stated that they had to reinstall the application, and the setup process of installing Health Connect and setting it up was a hassle. An Average score of 4.5, on a scale from 1 to 5, was given about the satisfaction and ease of the wearable data integration, which implies that this can be a good way of doing it in the future.

9 out of 10 participants preferred the mental health assessment tool instead of the traditional pen-and-paper questionnaires. One answered no preference. When explaining why they chose the preferred method, the participants stated things such as ease of use, mobile application, speed, and the fact that only a little effort is needed to answer it digitally. This gives us valuable feedback indicating that the solution does what is intended. All participants stated that the application was more convenient than the traditional method.

However, 5 out of 10 answered that they felt more comfortable sharing their data in traditional pen-and-paper questionnaires, 2 had no preference, and 3 picked the used application. This might imply that the public knows they cannot alter or see the passively collected data before it is submitted. Research and information about the data collection process need further research to make users comfortable with the passive data collection and sharing between individuals and platforms. 8 out of 10 said they would be willing to use the application. In addition, some participants stated the need to understand why passive data collection is needed and better understand who and where the data is sent and stored.

The survey gave us valuable feedback from potential users and showed promise for the artifact in question. However, this survey has certain limitations regarding participants, and future research should focus on generalization through testing, see Section 7.2.

Chapter 6

Discussion

Chapter 6 discusses the artifact developed, findings addressed throughout the process, and the research evaluation. We will start by addressing the research questions from Section 1.3. In other words, answers will be provided for the specified questions. This is followed by a section addressing our research contributions to the specified problem domain for the thesis. Additionally, a discussion on the use of our research method, techniques, and principles applied in developing our artifact is concluded in this chapter. Lastly, we will address and point out the limitations and challenges posed by our project and the implied result.

6.1 Answering the Research Questions

This section presents answers to the research questions addressed in Section 1.3.

RQ1 How can questionnaires be implemented to facilitate for the use of wearable data collection in the assessment process?

The first research questions surround the process of creating questionnaires that facilitates for the utilization of wearable data collection. This question has been in answered in Section 4.3.2 and 4.3.3 and mentioned in the semi-structured interview results in 5.1.2. We extended the IDPT framework with the possibility of creating questionnaires and receiving the corresponding response to the questionnaire.

Questionnaires and other traditional assessment methods are commonly used, however, as mentioned in Section 2.4 limitations are linked with these traditional assessments such as social desirability bias, response burden and motivation. With the possibility of passive data collection on questions we could mitigate these common challenges and collect objective data that gives a better representation of a patient physiological state. The extension provides a way for the therapist to choose questions for wearable data collection when asking question regarding physiological domains such as stress, sleep, and physical exercise.

The extension of both questionnaire and the response followed the same structure as the other components of the IDPT in the first iteration, however, to promote interoperability and reuse of the questionnaires in other systems we opted for a standardized data format. Both the questionnaire and the response utilize the HL7 standards for the specific resource and can thus be reused as discussed in Section 4.3. The questionnaire facilitates the wearable data collection while the response can hold the data collected for the specific question. We opted for a solution where the therapist can choose if the question should prompt patient self-report or passive data collection. In addition, both what type of data should be measured and the time interval for the collection must be provided to create a reliable integration. Based on this, we have successfully created a way where questionnaires facilitate for the use of wearable data collection in the assessment process.

RQ2 How can the integration of wearable technology with self-assessment questionnaires be implemented to enhance the self-assessment process?

This question prompts to answer how wearable technology can be integrated to combat the challenges and limitations regarding traditional self-assessment questionnaires discussed in Section 2.5. This question has been answered in Section 4.3 and discussed in Chapter 5 through evaluations and a user acceptance survey (see Appendix B). Through literature search and feedback from domain experts, the use of wearable technology to combat these challenges shows promise. As mentioned in Section 2.6, wearable technology can collect passive data about a patient's health; this data is objective and, if used correctly, shows promise to restrict response bias and response burden. As addressed in Chapter 5, the integration of wearable data collection into traditional has the potential to promote a better representation of the answers provided by the patient.

As mentioned in Section 4.3.3, we use a publicly available data collection platform that ensures the standardization and collection of data from multiple vendors. This promotes interoperability and reduces complexity, however, there exists different ways to do it. Section 2.8 mentions digital biomarkers that has positive correlations to the psychological domains discussed in Section 2.7. These biomarkers are collected at the exact time the user answers the questionnaire and ensures and temporarily stored in a data structure. When adding it to the response we created a HL7 FHIR observation resource to prompt interoperability of the collected data. However, the structure of the values collected are not in a standardized format, more on this in Section 4.3.3 and 4.3.4. Based on the semi-structured interview and the object-oriented evaluation discussed in Section 5.1 and 5.2, we successfully integrated wearable data collection into the self-assessment process. Thus, it can be used as a starting point for future development and research.

RQ3 Building on the result from the last two research questions, how can a system be designed to simplify and combat common challenges of traditional self-assessments for therapists and patients?

This research question builds upon the answers provided in RQ1 and RQ2. The answer to this research question is addressed in Section 4.4, 5.1 and 5.3. A demonstrative component was created to showcase how the proposed solution can be used and developed (See Section 4.4). The demonstrative component is a mHealth application that poses as a platform for users to answer the questionnaires created by the therapists in the IDPT framework. The demonstrative component shows how data can be efficiently collected while adhering to privacy and sensitive data challenges and ethical considerations, as described in Section 2.6. Through evaluations and a user acceptance survey, this component proves that it can assist patients and therapists.

The answer received shows potential to be more reliable than the ones from traditional methods, as discussed in Section 5.2.2. The demonstrative component serves as an example or a starting point on how wearable technology can be integrated into traditional self-assessment questionnaires for a more reliable assessment tool and how such a system can be developed to facilitate this. However, it is a novel solution, and additional testing and evaluation must be conducted to prove the proposed solution's potential further. Naturally, a way of scoring the questionnaire response should be developed in the future for the wearable data to hold any value. More on future work in Section 7.2.

6.2 Contributions

Sections 1.2 and 1.3 define and explain the problem domain and research questions for this thesis. These questions describe relevant problems using traditional self-assessment methods such as questionnaires. Hevner et al. (2004) state that to adhere to the methodology proposed by design science, the artifact developed should yield valuable contributions to the problem domain. See Section 3.1.1 and the fourth guideline in the design science principles for more information. A summary of the contributions from the artifact developed in this thesis follows.

The artifact is a starting point for future research in the described problem domain. It is a tool that can potentially make traditional self-assessment questionnaires and similar interventions more reliable. The artifact showcases how wearable technology is integrated into the mentioned traditional methods.

The main target for developing the artifact was to showcase a possible solution to the growing concern about the reliability of traditional self-assessment questionnaires. The artifact is publicly available, and the codebase is open-source, which can promote further research on the artifact and the problem domain in general. Based on the evaluation, our artifact and problem domain require more research due to the use of new and emerging technology. Our artifact provides valuable research and show a possible solution that can be used as a starting point for the future.

The general components and the demonstrative component discussed in Section 4.3 and 4.4 makes up the artifact. All the components developed and discussed can be integral to our contributions to the problem domain. The artifact in question contributes to the knowledge base for the problem domain. However, further research, development, and testing are needed to validate the artifact for use in real-life settings. As such, it is still a hypothesis in the form of an artifact.

6.3 Reflections

6.3.1 Research Methodology and Evaluation Methods

We aimed to develop an artifact that could provide valuable insights into a possible solution to the problem domain. Design science was the chosen methodology for the project conducted in this thesis. It was a good choice; it gave us tools and guidelines for conducting the research and for the development, design, and evaluation of the artifact that should contribute to the knowledge base.

In retrospect, it would have been beneficial for the implementation of our artifact to use domain experts throughout the process. Because a lot of time was spent on the understanding both Android development and the IDPT framework we lost valuable feedback since we did not include domain experts continuously in the implementation process.

A set of domain experts was interviewed, and a professional UX designer was asked for feedback on the design part of the demonstrative component. This gave us valuable feedback in the development process that helped us to evaluate the artifact. However, this was done in the latter part of the implementation phase, and we would have benefitted by including them earlier in the process so that they could see hot it unfolded. In addition, a small experiment was conducted to show that the artifact shows proof and is a step in the right direction. This included a user acceptance survey and comparing answers when self-reporting and collecting passive data for specified questions.

However, more testing of the artifact through quantitative methods would be beneficial to showcase the validity and reliability since our artifact is a hypothesis right now. In addition, taking advantage of qualitative methods throughout the development process would have been beneficial. This includes several interviews with domain experts and may have an evaluation of the artifact in each iteration of the process. For future reference, use domain experts throughout the implementation phase.

6.3.2 Reflecting on the Design Process

Several things hampered the process of designing the artifact. First, we created a mHealth application with React Native to be available on iOS and Android devices. However, after a while, it became clear that the React Native library did not have a way for us to incorporate the use of Health Connect. After opting into using that platform because of the possibility of accessing aggregated and raw data from multiple wearable companion apps, we had to change to Android development. The change provided many challenges; we needed to gain experience with Android development. Thus, much time was spent on understanding how it works. This led to valuable time spent reading documentation instead of improving and developing the artifact. The iterative process described by Hevner et al. (2004) worked well with the development of the artifact. The artifact comprises several general components, and the iterative process gave us a good way of dividing the necessary components needed for further development into separate iterations.

In the final iteration we conducted an evaluation of the artifact through an interview with domain experts and an outcome-based evaluation of the artifact. This gave us valuable information that we used in the latter part of the design and development process.

6.3.3 Reflecting on the developed artifact

The artifact developed provides a possible solution to an emerging problem domain. As mentioned in Section 2.7, no existing solutions could be compared to our artifact due to no open-source code and documentation of how it was developed. The solutions mentioned are different since they focus on ecological momentary assessment (EMA) with wearable integration. However, the artifact developed has contributed to an already existing open-source framework and created an open-source mHealth application available to the general population. Thus, our artifact and development can aid future work and research in the problem domain.

The artifact promotes the use of a novel platform that connects an individual's health data(Health Connect), which will be publicly available for all in Android 14. In addition, we have shown how to use the FHIR data standards and provided a starting point for how the passive data can be collected to combat the limitations and challenges described in Section 2.4.

Our goal was to provide a recognized data standard for the wearable data collected; however, no recognized standards were to be found, and creating such a standard was outside the scope of our project. Instead, we opted for the use of the FHIR observation resource that gave us a standardized format, however the values collected does not conform to a standardized format. In the future, a standardized wearable data format should be implemented in the IDPT to make the storing and retrieving of wearable data interoperable and standardized. The artifact developed promotes ease of use and a publicly available way of collecting data that can be used to create a more reliable health assessment tool. Based on this, this solution can be further developed or used as a starting point for other new solutions. See Section 7.2 for more information about future work in the problem domain.

We would have liked to implement additional features and functionalities in the artifact, like the scoring of the collected data for each domain. However, due to unforeseen delays mentioned in Section 6.3.2, this was not the case.

6.4 Project Limitations

In this subsection, we will address the limitations posed on our project. First, after thorough research, we could not find any existing solutions for the problem domain. Since this was a novel approach and no source code or architecture overview of similar applications was encountered, comparing our system and similar existing solutions could not be done.

Time constraints were a big challenge for us. At first, we started creating an application with ReactNative, after integrating a Samsung Galaxy watch through Bluetooth connecitivity we understood that it was hard to get hold of the raw data needed and for the future it would be a tedious task to add new devices. After the discovery we changed to Android development and had to start from scratch somewhat late in the development phase. Which resulted in less time spent on evaluation and the writing process. In hindsight, we understand that this is something we should of have researched before doing it.

Moreover, developing an Android application and integrating wearable technology into the mHealth application was time-consuming since this was new and untouched technology for us. This resulted in an MVP product, which is good, but we had the idea of adding this to the Google Play store to test it better and make it available to a larger audience. In addition, we chose to extend the IDPT framework which has not been in use for a year, so a considerable amount of time was spent on getting it to run and understand the architecture and logical flow.

Additionally, we chose to integrate and collect data with the Health Connect platform. Health Connect has some documentation. However, it is a new application, and the beta version can be found in the Google Play store. Very little source code with examples of how it can be used was found. Therefore, a reasonable amount of time was spent on understanding and implementing it as the data collection source in our artifact.

Furthermore, only a few domain experts and interviews were conducted during the development and evaluation. They gave us valuable information; however, to get a more in-depth understanding of how to implement our artifact best, it would be ideal to have more input from experts who know the domain. Moreover, we should have included them earlier in the process and continuously interviewed them throughout the implementation phase.

Lastly, our artifact was hampered during the testing and the object-oriented effectiveness evaluation because the mHealth application was not available in the Google Play Store, thus making it hard to test on individuals who did not know the creator. GDPR issues needs to be solved before it can be published to the Google Play Store.

In retrospect, it might have been clever to focus on one certain questionnaire and provide a way of scoring the data for that. Let us say that we took the PSS questionnaire and figured out which questions can be answered with wearable data and provide a stress measure algorithm based on HR and HRV measurements to get a better understanding of the capabilities and contributions of the artifact. However, the artifact developed showcases how wearable integration can be done and is a good starting point for future research.

Chapter 7

Conclusion

In this thesis, we have presented and evaluated the design and implementation of an artifact that integrates wearable data collection with traditional self-assessment questionnaires. The artifact serves as a mHealth application where patients can answer the questionnaires created by the therapist in the IDPT framework extension. The mHealth application serves as a digital tool that can be used to answer these questionnaires that can help therapists get a more reliable answer, which possibly scrutinizes the issues around social desirability bias and response burden.

This chapter serves as a conclusion of our research. It starts with a summary of the essential parts of the thesis. Lastly, it provides pathways for future research based on the thesis result.

7.1 Summary

We developed an artifact that incorporates wearable technology through passive data collection into traditional self-assessment questionnaires using the design science paradigm. This artifact consists of three general components: a questionnaire and a questionnaire response, an extension of the IDPT framework that facilitates the development of the demonstrative component. The focus has been creating a way for clinicians to create these questionnaires and a demonstrative component that uses the questionnaire and incorporates passive data collection for specific questions. The development of the novel artifact happened through an iterative process where domain experts, a UX designer, and potential users provided valuable feedback and data. In addition, efforts have been taken to standardize the data sent to the IDPT framework using HL7 FHIR standards, which promotes interoperability and reduce complexity.

The goal was to provide an artifact that could be a step in the right direction for future work in using ubiquitous wearable technology with self-assessment questionnaires. Through the development process, both qualitative and quantitative evaluation methods have been concluded, and the artifact developed serves as a good measure for future research in the specified problem domain. It can impact mental health care by reducing the risk of overlooking critical indicators when assessing patients with a questionnaire.

In retrospect, we would have liked to include the domain experts' and professionals' feedback in multiple iterations instead of just a single instance. It took significant time to develop the artifact because we needed to gain experience in Android development and the integration of wearable devices. The project would also have benefitted from setting aside time to implement an example on how the passively collected data could be scored and analyzed.

7.2 Future Work

The developed artifact is a novel solution and a minimum viable product as we conclude this thesis. There are several possibilities for future research and development. The artifact developed poses as a starting point that contributes to the knowledge base. Multiple functionalities can be added, and further research can be conducted on this emerging problem domain. The first thing that comes to mind is to thoroughly investigate the efficiency and accuracy of the answers given by the passive data collection. Underneath is a list of areas for improvement for future research:

Data Privacy and Security – Ensure that the data collected and sent for storing follows GDPR. We are working with sensitive health data, and ensuring the data does not fall into the wrong hands is essential. Adding a robust encryption algorithm and transport layer security to the application can solve that.

Create a separate platform for therapists – Even though the IDPT is a good framework, if someone is to develop the artifact or build a similar solution, we would encourage them to investigate the possibility of creating a stand-alone platform specifically for this application that can serve as a backend and workspace for therapists. This makes it easier to scale the platform and enhances the possibility of scoring questionnaires and analyzing the data collected from the wearable.

Longitudinal studies – Conducting longitudinal studies over an extended period can offer a greater insight into the developed artifact's reliability and effectiveness in the problem domain. The domain experts stated that the artifact poses as a hypothesis now and extensive testing needs to be concluded. Longitudinal studies are a great way of testing it. This includes how data is collected and how a patient's needs and preferences change over time.

Further investigate the reliability of the passive data collected – Incorporating wearable technology into existing solutions in the health sector is tedious. Future research can continue to check the reliability of the artifact against traditional methods. In other words, conduct studies to investigate if it does what it promises, creating a more reliable assessment tool. This can be done by studying the answers and seeing if it promotes less response burden and social responsibility bias for the user. A validation of the accuracy of the data collected can also prove to be suitable for future research.

Additional Wearable Data Sources – The artifact uses Health Connect, which promotes the use of multiple wearable devices. However, it is limited to a set amount, and future work should investigate how we can promote the use of additional wearable data sources to make the artifact available to a larger audience.

Implement a way of scoring the Questionnaires and the Wearable Data – A way of scoring the questionnaires in the IDPT framework is the scoring of questionnaires so that a score is automatically given to the therapist or user when submitting their answer. This could also include visualizing the data collected with graphs to the therapist.

Analyze the data – The data collected from Health Connect are raw or aggregated. To interpret and use this answer to guide the therapist, we need a way of scoring each data type. This can be done using an already studied algorithm or through a new algorithm and provides a good starting point for future research.

Standardization – Currently, the data collected from Health Connect and sent to the IDPT framework lacks standardization. The measurements are added as a FHIR observation resource, but the actual data collected does not conform to a standardized data format. This is something that should be researched and studied in the future.

List of Acronyms and Abbreviations

API	Application Programming Interface
EDA	Electrodermal Activity
EMA	Ecological Momentary Assessment
FHIR	Fast Healthcare Interoperability Resources
GHQ-9	General Health Questionnaire-9
HL7	Health Level 7
HR	Heart Rate
HRV	Heart Rate Variability
IDPT	Internet Driven Psychological Treatment
mHealth	Mobile Health
PSS	Perceived Stress Scale
UI	User Interface

Bibliography

Android Developers (2023) ‘’Health Connect: simplify connectivity between apps’’ URL: https://developer.android.com/health-and-fitness/guides/health-connect (Accessed 10.08.2023)

Abernethy, A. et al. (2022). ‘’The Promise of Digital Health: Then, Now, and the Future’’. *NAM perspectives*. doi: https://doi.org/10.31478/202206e

Adeoye, O.A. et al. (2021). "Research and scholarly methods: Semi-structured interviews" J. Am. Coll. Clin. Pharm – 4,1258-1367. doi: https://doi.org/1002/jac5.1441

Aiyegbusi, O. A. et al. (2022). "Key considerations to reduce or address respondent burden in patient-reported outcome (PRO) data collection." Nature communications vol. 13(1), 6026. doi: https://doi.org/10.1038/s41467-022-33826-4

Alhejaili R. & Alomainy A. ‘’The Use of Wearable Technology in Providing Assistive Solutions for Mental Well-Being’’. *Sensors*. 2023; 23(17):7378. doi: https://doi.org/10.3390/s23177378

Andersson G. (2016).” Internet-Delivered Psychological Treatments”. *Annual review of clinical psychology*, *12*, 157–179. doi: https://doi.org/10.1146/annurev-clinpsy-021815-093006

Anjara, S G. et al. "Using the GHQ-12 to screen for mental health problems among primary care patients: psychometrics and practical considerations." International journal of mental health systems vol. 14 62. 10 Aug. 2020, doi: https://doi.org/10.1186/s13033-020-00397-0

Ayaz, M. et al. (2021). "The Fast Health Interoperability Resources (FHIR) Standard:Systematic Literature Review of Implementations, Applications, Challenges and Opportunities." JMIR medical informatics vol. 9,7 e21929 doi: https://doi.org/10.2196/21929

Azza, Y. et al. (2020). ‘’How stress affects sleep and mental health: nocturnal heart rate increases during prolonged stress and interacts with childhood trauma exposure to predict anxiety.’’ *Sleep*, *43*(6), doi: https://doi.org/10.1093/sleep/zsz310

Babrak, L. M. et al (2019). ‘’Traditional and Digital Biomarkers: Two Worlds Apart?’’. *Digital biomarkers*, *3*(2), 92–102. doi: https://doi.org/10.1159/000502000

Bender, D. & Sartipi, K. (2013). "HL7 FHIR: An Agile and RESTful approach to healthcare information exchange," Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems, Porto, Portugal. pp. 326-331. doi: https://doi.org/10.1109/CBMS.2013.6627810

Ben-Zeev, D. et al. (2017). "CrossCheck: Integrating self-report, behavioral sensing, and smartphone use to identify digital indicators of psychotic relapse." Psychiatric Rehabilitation Journal vol. 40,3: 266-275. doi: https://doi.org/10.1037/prj0000243

Burke, L. E. et al. (2017). "Ecological Momentary Assessment in Behavioral Research: Addressing Technological and Human Participant Challenges." *Journal of Medical Internet Research vol. 19,3 e77*. 15 doi: http://doi.org/10.2196/jmir.7138

Böttcher, S. et al. (2022) "Data quality evaluation in wearable monitoring." Sci Rep 12, 21412 doi: https://doi.org/10.1038/s41598-022-25949-x

Chen, J. et al. (2019). ''Design of personnel big data management system based on blockchain.'' *Future Generation Computer Systems*, *101*, 1122-1129. doi: https://doi.org/10.1016/j.future.2019.07.037.

Chen, J. et al. (2021) "Pain and stress detection using wearable sensors and devices—a review," Sensors, 21(4), p. 1030. doi: https://doi.org/10.3390/s21041030.

Colombo, D. et al. (2019). ''Current state and future directions of technology- based ecological momentary assessment and intervention for major depressive disorder A systematic review.''. *Journal of Clinical Medicine* 8(4). doi: https://doi.org/10.3390/jcm8040465

Columbia University Department of Psychiatry (2023). ''How Sleep Deprivation Impacts Mental Health''. URL: https://www.columbiapsychiatry.org/news/how-sleep-deprivation-affects-your-mental-health

Cox, M. et al. (2018). "Use of Wearable, Mobile, and Sensor Technology in Cancer Clinical Trials". JCO Clinical Cancer Informatics 2018:2, pp. 1–11. doi: https:////doi.org/10.1200/CCI.17.00147

Edwards P. et al. (2004). "Follow-up by mail in clinical trials: does questionnaire length matter?" Controlled Clinical Trials, Volume 25, Issue 1, pp. 31–52. ISSN: 0197–2456. doi: https://doi.org/10.1016/j.cct.2003.08.013

El-Sherif, D. M., & Abouzid, M. (2022). ''Analysis of mHealth research: mapping the relationship between mobile apps technology and healthcare during COVID-19 outbreak''. *Globalization and health*, *18*(1), 67. doi: https://doi.org/10.1186/s12992-022-00856-y

ExpiWell (2022). URL: https://www.expiwell.com (Accessed at 16.06.2023)

FHIR (2023). URL: https://www.hl7.org/fhir/resourceguide.html?fbclid=IwAR36AJ-8BuUmwUeettSbBvRMPwQlhbBCovuD1-Jwho11R0KcvHoJRRf0A1Q (Accessed at 12.08.2023)

Fitbit**.** (2023). Fitbit Privacy Policy. URL: https://www.fitbit.com/global/us/legal/privacy-policy (Accessed at 07.07.2023).

Geneviève, L. D. et al. (2019). ''Factors influencing harmonized health data collection, sharing and linkage in Denmark and Switzerland: A systematic review.'' *PloS one*, *14*(12). doi: https://doi.org/10.1371/journal.pone.0226015

Godfrey, A. et al (2018). ''From A to Z: Wearable technology explained''. *Maturitas*, *113*, 40-47. doi: https://doi.org/10.1016/j.maturitas.2018.04.012.

Graeff, T. R. (2005). Response bias. In K. Kempf-Leonard, Encyclopedia of Social Measurement (pp. 411–441). Academic Press.

GraphQL (2023) URL: https://graphql.org (Accessed at 29.08.2023)

Grassini, S., & Laumann, K. (2020). Questionnaire Measures and Physiological Correlates of Presence: A Systematic Review. *Frontiers in Psychology*, 11. doi: https://doi.org/10.3389/fpsyg.2020.00349

Guk, K. et al. (2019). ''Evolution of wearable devices with real-time disease monitoring for personalized healthcare''. In *Nanomaterials* 9(6). doi: https://doi.org/10.3390/nano9060813

Haghi, M. et al. (2017). ''Wearable devices in medical internet of things: scientific research and commercially available devices''. *Healthc Inform Res*. 23: 4–15. doi: https://doi.org/10.4258/hir.2017.23.1.4

Hart, A. et al. (2022). ''Using Smartphone Sensor Paradata and Personalized Machine Learning Models to Infer Participants' Well-being: Ecological Momentary Assessment.'' *Journal of medical Internet research*, *24*(4). doi: https://doi.org/10.2196/34015

HelseNorge (2023). ''Mental Health and Physical Activity''. URL: https://www.helsenorge.no/en/psykisk-helse/mental-health-and-physical-activity/ (Accessed at 23.07.2023)

Henderson C. et al. (2013). "Mental illness stigma, help seeking, and public health programs." Am J Public Health. 2013;103(5):777-780. doi: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3698814/

Henriksen, A. et al. (2018). ''Using Fitness Trackers and Smartwatches to Measure Physical Activity in Research: Analysis of Consumer Wrist-Worn Wearables''. *Journal of medical Internet research* 20(3), e110. doi: https://doi.org/10.2196/jmir.9157

Hevner, A. R. et al. (2004-03). "Design Science in Information Systems Research". *MIS Quarterly* 28(1), pp. 75–105

Huhn, S. et al. (2022). "The Impact of Wearable Technologies in Health Research: Scoping Review". *JMIR mHealth and uHealth*, *10*(1), doi: https://doi.org/10.2196/34384

Jetpack Compose (2023) URL: https://developer.android.com/jetpack/compose (Accessed at 11.09.2023)

Jerath, R. et al. (2023). "The Future of Stress Management: Integration of Smartwatches and HRV Technology." *Sensors* (Basel, Switzerland) vol23,17 7314. doi: https://doi.org/10.3390/s23177314

Johnson, T., & Fendrich, M. (2005). ''Modeling sources of self-report bias in a survey of drug use epidemiology''. *Annals of epidemiology*, *15*(5), 381–389. doi: https://doi.org/10.1016/j.annepidem.2004.09.004

Kwak, D. H. et al (2021) ''When does social desirability become a problem? Detection and reduction of social desirability bias in information systems''. *Information & Management. 58(7). doi: https://doi.org/10.1016/j.im/2021/103500*

Kim, H. G. et al. (2018). Stress and Heart Rate Variability: A Meta-Analysis and Review of the Literature. *Psychiatry investigation*, *15*(3), 235–245. doi: https://doi.org/10.30773/pi.2017.08.17

Kotlin (2023) URL: https://kotlinlang.org/spec/introduction.html (Accessed at 08.27.2023)

Lamo, Y. et al. (2022). ''Towards adaptive technology in routine mental health care''. *DIGITAL HEALTH*. 8. doi: https://doi.org/10.1177/20552076221128678

Larson, R. B. (2019). “Controlling social desirability bias.” *International Journal of Market Research*, *61*(5), 534-547. doi: https://doi.org/10.1177/1470785318805305

Lee, J. M. et al*.* (2014) ''Association between stressful life events and resting heart rate''. *BMC Psychol* 2(29). doi: https://doi.org/10.1186/s40359-014-0029-0

Lehne, M. et al. (2019). "Why digital medicine depends on interoperability." *NPJ digital medicine* vol. 2 79., doi: https://doi.org/10.1038/s41746-019-0158-1

Liao, Y. et al. (2019). "The Future of Wearable Technologies and Remote Monitoring in Health Care." *American Society of Clinical Oncology Educational Book* pp. 115-121 doi: https://doi.org/10.1200/EDBK/_238919

Mahindru, A. et al. (2023). ''Role of Physical Activity on Mental Health and Well-Being: A Review''. *Cureus*, *15*(1). doi: https://doi.org/10.7759/cureus.33475

Makovac, E. et al. (2022). ''Safe in my heart: resting heart rate variability longitudinally predicts emotion regulation, worry, and sense of safeness during COVID-19 lockdown''. *Stress (Amsterdam, Netherlands)*, *25*(1), 9–13. doi: https://doi.org/10.1080/10253890.2021.1999408

Marzban, S. et al. (2022). "Impact of Patient Engagement on Healthcare Quality: A Scoping Review." *Journal of patient experience* vol. 9 23743735221125439. 16 Sep. 2022, doi: https://doi.org/10.1177/23743735221125439

McCraty, R., & Shaffer, F. (2015). Heart Rate Variability: New Perspectives on Physiological Mechanisms, Assessment of Self-regulatory Capacity, and Health Risk. *Global advances in health and medicine*, *4*(1), 46–61. doi: https://doi.org/10.7453/gahmj.2014.073

Mikkelsen, K. et al. (2017) "Exercise and mental health" Maturitas, 106, 48-56. doi: https://doi.org/10.1016/j.maturitas.2017.09.003

MongoDB (2023) URL: https://www.mongodb.com/docs/ (Accessed at 08.24.2023)

Morgan, C. et al. (2017). ''The effectiveness of unguided internet cognitive behavioural therapy for mixed anxiety and depression.'' *Internet Interventions*, *10*. doi: https://doi.org/10.1016/j.invent.2017.10.003

National Library of Medicine. (2023) Data Standards URL: https://www.nnlm.gov/guides/data-glossary/data-standards (Accessed at 10.08.2023)

Nicholas J, et al. (2019) ''The Role of Data Type and Recipient in Individuals' Perspectives on Sharing Passively Collected Smartphone Data for Mental Health: Cross-Sectional Questionnaire Study'' *JMIR Mhealth Uhealth* 7(4) doi: https://doi.org/10.2196/12578

Patel, A. K. et al. (2022). ''Physiology, Sleep Stages''. *StatPearls Publishing*. URL: https://www.ncbi.nlm.nih.gov/books/NBK526132/

Pekrun, R. (2020). "Commentary: Self-Report is Indispensable to Assess Students' Learning." *Frontline Learning Research*, Vol. 8(3), 185–193. ISSN 2295-3159 doi: https://doi.org/10.14789/flr.v8i3.637

Quach, S. et al. (2022). ''Digital technologies: tensions in privacy and data''. *Journal of the Academy of Marketing Science*, 50, 1299–1323. doi: https://doi.org/10.1007/s11747-022-00845-y.

Ramsawh, H. J et al. (2009). "Relationship of anxiety disorders, sleep quality, and functional impairment in a community sample." *Journal of Psychiatric Research vol. 43,10: pp. 926–33*. doi: https://doi.org/10.1016/j.jpsychires.2009.01.009

Revilla, M. et al (2016). "Using Passive Data From a Meter to Complement Survey Data in Order to Study Online Behavior" Social Science Computer Review, 35(4), 521-536 doi: https://doi.org/10.1177/0894439316638457

Rolstad, S. et al. (2011). "Response burden and questionnaire length: is shorter better? A review and meta-analysis." *Value in health: The Journal of the International Society for Pharmacoeconomics and Outcomes Research*, 14(8), 1101–1108. doi: https://doi.org/10.1016/j.jval.2011.06.003

Room Database (2023). URL: https://developer.android.com/jetpack/androidx/releases/room (accessed at 07.08.2023)

Roopa, S & Rani, M. S. (2012). "Questionnaire Designing for a Survey." *The Journal of Indian Orthodontic Society*. 46. 37-41. doi: https://doi.org/10.5005/jp-journals-10021-1104

Sajjadieh, A. et al. (2020). ‘’The Association of Sleep Duration and Quality with Heart Rate Variability and Blood Pressure’’. *Tanaffos*, *19*(2). pp.135–143. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7680518/

Scott, A. J. et al. (2021). "Improving sleep quality leads to a better mental health: A meta-analysis of randomized controlled sleep trials." *Sleep Medicine Reviews*, Volume 60, 101556. ISSN 1087-0792. doi: https://doi.org/10.1016/j.smrv.2021.101556.

Shahsavarani A. M. et al. (2015). "Stress: Facts and Theories through Literature Review" International Journal of Medical Reviews, 2(2), 230-241. URL: https://www.ijmedrev.com/article_68654_37adc02e9432adfa017b8d6095cb6760.pdf

Sharma H. (2022). "How short or long should be a questionnaire for any research? Researchers' dilemma in deciding the appropriate questionnaire length." *Saudi journal of anesthesia*, *16*(1), 65–68. https://doi.org/10.4103/sja.sja_163_21

Sinhasane, S. (2018). Wearable Technology: The Coming Revolution in Digital Health. *Mobisoft Infotech Blog*. URL: https://mobisoftinfotech.com/resources/blog/wearable-technology-in-healthcare/

Singh, P. (2013). ‘’P Value, Statistical Significance and Clinical Significance’’. *J Clin Prev Cardiol*, 2(4) 202-204. URL: https://www.jcpcarchives.org/full/p-value-statistical-significance-and-clinical-%09significance-121.php

Statista. (2023). "Number of users of smartwatches worldwide from 2018 to 2027" *Statista* URL: https://www.statista.com/forecasts/1314339/worldwide-users-of-smartwatches(Last visited on 10.08.2023)

Tobaldini, E. et al. (2013). «Heart rate variability in normal and pathological sleep”. *Frontiers in Physiology, 4*(294). doi: https://doi.org/10.3389/fphys.2013.00294

Tudor-Locke, C. et al. (2015) “Improving wear time compliance with a 24-hour waist-worn accelerometer protocol in the International Study of childhood obesity, lifestyle and the environment (ISCOLE),” *International Journal of Behavioral Nutrition and Physical Activity,* 12(1). doi: https://doi.org/10.1186/s12966-015-0172-x.

Jiaobing, T. & Gao, W. (2021). "Ethical Considerations of Wearable Technologies in Human Research." *Advanced healthcare materials* 10(17) doi:10.1002/adhm.202100127 !!!

Watanabe, K. & Tsutsumi, A. ''The Passive Monitoring of Depression and Anxiety Among Workers Using Digital Biomarkers Based on Their Physical Activity and Working Conditions: 2-Week Longitudinal Study'' *JMIR Form Res. 6(11)* doi: https./doi.org/10.2196/40339

Weiler, A. (2016). "mHealth and big data will bring meaning and value to patient-reported outcomes." *mHealth* 2(2), doi: 10.3978/j.issn.2306-40.2016.01.02

World Health Organization. (2022) "Mental Health". URL: https://www.who.int/news-room/fact-sheets/detail/mental-health-strengthening-our-response (last visited 07.07.2023)

- (2023). *Stress*. URL: https://www.who.int/news-room/questions-and-answers/item/stress (Accessed at 15.05.2023)

Xie, Y. et al. (2021) "Integration of artificial intelligence, Blockchain, and wearable technology for Chronic Disease Management: A new paradigm in smart healthcare," *Current Medical Science,* 41(6), pp. 1123–1133. doi: https://doi.org/10.1007/s11596-021-2485-0.

Yan, T. et al. (2019). "Response Burden: What Is It and What Predicts It?" *Advances in Questionnaire Design, Development, Evaluation and Testing* pp 193-212. Wiley Online Books. doi: https://doi.org/10.1002/9781119263685.ch8

Zamkah, A. et al. (2020). ''Identification of suitable biomarkers for stress and emotion detection for future personal affective wearable sensors.'' *Biosensors* 10(4). doi: https://doi.org/10.3390/bios10040040

Zou, P. et al. (2020). "Poorer sleep quality correlated with mental health pests in college students: A longitudinal observational study among 686 males". *Journal of Psychosomatic Research*, 136(110177) doi: 10.116/j.jpsychores.2020.110177.

Appendix A

Semi-Structured Interview Answers

Answers to the semi structured interview conducted with two separate domain experts in the field.

1: What kind of experience do you have with wearable technology in the health care sector?

Domain Expert 1:
- Patient surveillance
- Patient monitoring

Domain Expert 2:
- Patient surveillance
- Patient monitoring
- EMA
- Used it under COVID-19 to assess patients and try to see themes and patterns through the sensor data that could help us depict symptoms.

2: What do you think of traditional standardized self-assessment questionnaires when it comes to mental health?

Domain Expert 1:
- Can be a good tool, however we tend to see that answers given are misleading and that the patient answers what we want to see rather than how they feel or the experiences they have been through.
- Patients with mental health issues like anxiety, depression and so on often don’t have the motivation to fill out these.

Domain Expert 2:
- Divided
- Can serve as an indication of mental but it has the potential of providing false information that can lead to wrong treatment or diagnosis.
- Biases because the patient often have troubles in their life, and they are not in the right state of mind to answer objectively.

3: Do you have any experience with the use of traditional assessment in mental health?

Domain Expert 1:
- Yes, follow ups after assessments and the PHQ-9 to indicate potential symptoms in patients.
- Qualitative Surveys conducted after new modules and methods to assess the potential value from users.

Domain Expert 2:

- Yes
- Used them for insight into mental health under the COVID-19 pandemic.
- Physical exercise questionnaires where the patient states how physical they have been which is often an indication into how they feel.

4: What do you think about the use of wearable technology in the field of mental health?

Domain Expert 1:
- Good idea
- Long way to go because the lack of standardization and the fact that it is hard to get hold of raw data from big companies like Fitbit, Garmin etc.
- Can be good, but it must be implemented in a way that serves the patient.

Domain Expert 2:
- Interesting
- Faster and more accurate recollection of a patient's mental health if done the right way.
- Challenges with interoperability, integration, and standardization

5: Would integrating wearable technology into these questionnaires be beneficial?

Domain Expert 1:
- If done in the right way yes
- We are collecting sensitive data so privacy, quality and storing of the data is essential.
- Can make the scoring and analysis process complex and time-consuming.

Domain Expert 2:
- Yes, the healthcare sector needs to embrace new emerging technology.
- However, it will take time and resources.
- Make sure that the data collected is accurate and reliable, if not there is no need.

6: [Follow up for Q5] Can you list any potential benefits by implementing this?

Domain Expert 1:
- Objective data, get rid of the bias surrounding answers where it is applicable.
- Faster diagnosis or help which can prevent huge costs for the health care sector and society.

Domain Expert 2:
- Yes, a more holistic representation of a patient's mental state early in the process.
- This can lead to faster help and save a lot of time.
-

7: Do you know of any similar existing solutions?

Domain Expert 1:
- No

Domain Expert 2:

- I know of some research done on the integration of wearable technology into EMAs, however that is a bit different since EMA is used when like a stress response occurs.
- As far as I have understood were talking about collecting passive data retrospectively and I like the idea of that.

8: Are there any potential challenges and limitations to this new method?

Domain Expert 1:

- We're talking about a fairly new technology which leads to a set of challenges:
    - Data Privacy
    - GDPR
    - Consent
    - Data quality
    - Accuracy
    - Ethics

Domain Expert 2:

- Yes
    - What about if no data is collected because data is not synced, or an issue occurs?
    - Privacy and storage
    - Standardization of the wearable data – use FHIR?
    - Effectiveness
    - Reliability
    - Trustfulness to the application, the manufacturer of the integrated wearable and the way in which collection happens. Who gets to see it? We are discussing the collection and use of sensitive data so one must be careful.

9: Do you think that integrating wearable technology into these questionnaires to collect passive data for some questions can be useful?

Domain Expert 1:

- Yes
- It is important to get a good understanding or selection method on which questions can be collected passively and which ones need user-input.

Domain Expert 2:

- Absolutely – if done the right way. There are a lot of pitfalls so one must be careful when developing this "tool".

10: What type of measurements would be valuable to collect instead of asking the patient to self-report?

Domain Expert 1:

- There are a few valid ways. For example, there exists way to measure stress with the use of heart rate variability, resting heart rate, skin temperature, ECG and so on. I would guess that a regular smartwatch does not have the ability to measure all of them accurately.

Domain Expert 2:

- As mentioned, I have used physical exercise data before in questionnaires. They can have an indication of certain mental health issues and by assessing raw or aggregated data instead of a patient's own recollection of an event we can get more accurate and valuable data.
- This also applies to sleep. Manufacturers tend to measure sleep in different ways with their devices, however, they are more accurate at stating statistics like time in each sleep stage, time awake, asleep and so on that could be valuable. Sleep is a universal measurement for most diagnoses.
- Heart rate variability is also a common way which indicates stress, so that would be something to try out.

11: Do you think the proposed application can be useful and adaptable for future use?

Domain Expert 1:

- It could provide valuable information on how it can be done.
- Yes, if it shows promise it can yield future research.

Domain Expert 2:

- If it shows proof that it is doing something right, then yes.
- Faster diagnoses and help

12: How would you think that clinicians/therapists would benefit from the implementation and use of this proposed tool?

Domain Expert 1:

- As mentioned earlier:
    - Better representation of the current physiological state
    - Help earlier in the process.
    - Resources – could help them in allocating their time elsewhere.
    - However, only if there is a way to analyze the data measured.

Domain Expert 2:

- Patient engagement helps the therapist by actively including the patient since they can see their scores and it is presented in a more modern way.
- Faster help

13: What do you think about the use of a third-party app for collection that gathers data from different vendors into a standardized format through an API?

Domain Expert 1:

- Good idea
- If there is a possibility to collect data from different wearables from one platform in a standardized way, it is great.

Domain Expert 2:
- Great
- As long as privacy and sharing are fine
- Untapped possibilities if one can gather all the data in one place.

14: Have you heard about Health Connect?

Domain Expert 1:
- No

Domain Expert 2:
- Yes, I have seen something about it but it is fairly new so don't know how it works and how efficient it is.

Appendix B

Questionnaire provided to the participants of the object-oriented experiment described in Section 5.3 Questions annotated with “*” shows which is substituted or is answered by passive data collection in the application tested. The questionnaire is made for this specific test and serves no purpose outside of the experiment conducted.

General Well-Being Questionnaire

1. Over the course of the last week, have you felt anxious or nervous? (On a scale from 1 to 5 where 1 = nothing and 5 = a lot)

[ ] 1
[ ] 2
[ ] 3
[ ] 4
[ ] 5

2. On average, how much do you sleep each day? *

__ h ___m

3. Do you spend a lot of time awake in bed?(Choose one) *

[ ] Yes
[ ] No
[ ] Maybe
[ ] Prefer not to answer

4. In the last week, how many times have you engaged in social activity?

[ ] 1
[ ] 2
[ ] 3
[ ] 4
[ ] 5 or more

5. In the last week, how many times have you engaged in physical activity? *

[ ] 1
[ ] 2
[ ] 3
[ ] 4
[ ] 5 or more

Appendix C

User Acceptance Survey

This was given to the participants of the object-oriented effectiveness evaluation and the result is discussed in Section 5.2. It was given through Survey Monkey, but a written example of the Survey can be found below

Thank you for participating in our study. We would love to get some feedback from you that would help to improve our research and the developed artifact.

Part 1: General Information

1 Gender:

Male [ ] Female [ ] Other [ ] Prefer not to say [ ]

2 Age:

Under 18 [ ] 18-24 [ ] 25-34 [ ] 35-44 [ ] 45-54 [ ] 55-64 [ ] 65 or older [ ]

3 Have you ever used a mental health assessment tool before (questionnaires, forms, surveys)?

Yes [ ] No [ ]

4 If yes, which methods have you used? (Select all that apply)

[ ] Traditional pen-and-paper questionnaires
[ ] Online surveys or forms
[ ] Mobile application
[ ] Others (please specify): ____

Part 2: Experience with the application

1 On a scale of 1 to 5, how user-friendly did you find the used application? (1=Not user-friendly, 5=Very user-friendly

[ ] 1
[ ] 2
[ ] 3
[ ] 4
[ ] 5

2 Please describe any difficulties or challenges you encountered while using the application (if any):

3 How satisfied were you with the integration of wearable data into the mental health assessments? (1 = Very dissatisfied, 5 = Very satisfied)

[ ] 1
[ ] 2
[ ] 3
[ ] 4
[ ] 5

Part 3: Comparison with Traditional Methods

1 Which method do you prefer for mental health assessments: the application used or the traditional pen-and-paper questionnaires?

[ ] Used application
[ ] Traditional pen-and-paper questionnaires
[ ] No preference

2 Please explain why you prefer one method over the other

3 Did you find the used application more convenient to use than the traditional methods?

Yes [ ] No [ ]

4 Did one method make you feel more comfortable sharing your data?

[ ] Used application
[ ] Traditional pen-and-paper questionnaires
[ ] No preference

Part 4: Additional Feedback

1 Do you have any suggestions for improving the digital mental health assessment tool or the integration of wearable data?

3 Do you have any prior experience with wearable technology?

Yes [ ] No [ ]

Part 6: Additional Comments (Optional)

1 Is there anything else you would like to share about your experience with the application use or this study?

Appendix D

Source Code, Demo and Documentation

(i) The source code for the mHealth application is available at this URL: https://github.com/P1T1B0Y98/MindSync

(ii) The source code for the IDPT framework extended in this thesis can be found at this URL: https://github.com/P1T1B0Y98/idpt/tree/feat/mental-health-assessment

(iii) Documentation and development information can be found here: https://mindsync.gitbook.io/mindsync/

(iv) The demo for the IDPT extension and the mHealth application can be found in the readme.md file in the repository URL in (i).