Co-funded by the
Erasmus+ Programme
of the European Union

Digital Twins in Healthcare

Conceptualisation and Privacy Aspects

MASTERTHESIS
to obtain the Erasmus Mundus Joint Master Degree

in Digital Communication Leadership (DCLead)

of

Faculty of Cultural and Social Sciences
Paris Lodron University of Salzburg

Technical Faculty of IT and Design
Aalborg University in Copenhagen

Submitted by
Sarah Julia Jeske
11836861
sarah@tecsi.de
Haager Str. 52b, 85435 Erding, GERMANY

Reza Tadayoni
Ursula Maier-Rabler
Daniel Angus

Department of Communication Studies

Salzburg, 28.07.2020

Contents

Table of Figures

Table of Tables

List of Acronyms

AI	Artificial Intelligence
API	Application Programming Interface
GDPR	General Data Protection Regulation
IoT	Internet of Things
PbD	Privacy by Design
RQ	Research question

Executive Summary

The digital twin technology is experiencing increasing popularity and is being applied in fields such as manufacturing, product lifecycle management, prognostics and health management, waste management, and smart cities. A digital twin is a virtual representation of a physical product, consisting of three main components: the physical object, its virtual equivalent, and the data connection between the physical and the virtual entity. Digital twins support the physical system regarding maintenance, prediction, and simulation. Healthcare is an up-and-coming, yet relatively uncharted and complex field of application for digital twins as the physical object is the human. This master thesis aims to explore how digital twins can be implemented in healthcare in a privacy friendly manner. Firstly, specific privacy challenges associated with digital twins in healthcare are identified. Secondly, the influence of the identified privacy risks on the implementation of digital twins is of interest. For this purpose, an adapted step-by-step process based on Requirements Engineering serves as the underlying methodological approach to identify essential prerequisites from stakeholders while uncovering the challenges that digital twins in healthcare pose. Within these steps, semi-structured expert interviews with six participants knowledgeable in the field of digital twins, digital twins in healthcare, and privacy studies are conducted. The findings gathered in the interviews are analysed using Mayring's approach for qualitative text analysis and are interpreted in relation to relevant literature and theories in the field of digital health and digital twins. Finally, the insights are discussed critically and assembled into a model. The outcome is a visualisation of the digital twin infrastructure that includes the main components in the physical and virtual space, as well as their relationship with another. Additionally, the challenges and privacy friendly design strategies to counteract the issues are included. This conceptual framework of the implementation of digital twins in the field of healthcare acts as a guide on what needs to be taken into consideration when developing such technology and serves as a source for further research in the field of digital twins in the health domain.

1 Introduction

In late 2019, cases of pneumonia were detected in Wuhan, Hubei province in China. A couple of weeks later, Chinese authorities revealed that the pneumonia was caused by a novel coronavirus (CNN, 2020). In mid-January, the first case outside of China was confirmed in Thailand (CNN, 2020) and the virus then rapidly continued to spread globally, reaching over 216 countries, areas, or territories with a total of 10 million infected cases worldwide (as of June 29, 2020) (World Health Organization, 2020b). On March 11, the World Health Organisation declared the virus outbreak a pandemic (World Health Organization, 2020a). Following the rapid emergence of the virus and the associated respiratory disease COVID-19, the healthcare systems of several countries, such as Italy (Horowitz, 2020), Spain (Jones, 2020), and others, were pushed to their limits. In response to this and to avoid being hit as severely as some European countries, Armidale Hospital in New South Wales, Australia, planned to equip non-hospitalised low-risk COVID-19 patients with home monitors. While the patient sleeps, the device tracks their vital signs such as heart rate, body temperature, and blood pressure that could indicate their overall health condition. An artificial intelligence program, together with doctors, monitor this status in a 'virtual hospital' and could order the patient to admit themselves to the 'physical hospital' if necessary (Mannix, 2020). Australia's first virtual hospital is only one example of how the COVID-19 pandemic has led to rethinking the overall health system and how technologies could facilitate healthcare in the future. Saracco (2020a) takes the idea of the virtual hospital a step further by employing "personal digital twins" in epidemics control. The idea is to collect data from a person through their smartphone and wearable sensors. The data might consist of their movement, their interaction with other people and their general health condition, indicated by their vital signs. As fever is a common symptom for COVID-19, healthcare institutions could use such data to analyse the people's movement and compare this to other digital twins that show signs of high body temperature. Moreover, it could help to find out which areas are affected most. "This would create an awareness of an incipient epidemics and make accurate forecast of possible contagion based on red flagged people movement. It would provide a most timely and most accurate picture of the situation, worldwide" (Saracco, 2020a).

The hype surrounding digital twins has increased recently. Gartner's *Hype Cycle* explores how certain technologies and innovations might evolve and support businesses when adopted over

the next years (Gartner, n.d.). In the *Hype Cycle for Emerging Technologies* in 2018, the term 'digital twin' was on the peak of the curve and was said to reach mainstream adoption in five to ten years (Panetta, 2018a). In their *Top 10 Strategic Technology Trends for 2019* (Panetta, 2018b), the research and advisory company Gartner once again mentioned digital twins and named this technology the number four trend that businesses need to explore. The term "digital twin" is herein defined as "a digital representation of a real-world entity or system. […] They are linked to their real-world counterparts and are used to understand the state of the thing or system, respond to changes, improve operations and add value" (Cearley & Burke, 2018, p. 16). Therefore, the general set-up around a digital twin includes the physical object, its digital equivalent, and the data connection between the physical and the virtual entity. The latter is what distinguishes a digital twin from a mere digital model. Digital twins further support the physical system regarding maintenance, prediction, and simulation (Grieves, 2014). Possible future behaviour, both desired and undesired, can be detected, and measures to elicit or prevent these behaviours can be undertaken (Grieves & Vickers, 2017).

Saracco's vision of a personal digital twin is still just that. However, leading healthcare companies such as Siemens (Siemens Healthcare GmbH, 2019), Philips (van Houten, 2018), or General Electric (GE Healthcare Partners, 2018) have recognized the technology's potential of revolutionizing the health sector and have invested in researching digital twins in healthcare. It is important to note that digital twins in healthcare can be viewed from different perspectives. On the one hand, digital twins can be used to manage expensive health equipment (Tao, Zhang, Liu, & Nee, 2018). On the other hand, digital twins can be used for modelling the patient, as explained in the article by Saracco. The crucial difference is that the latter is based on humans. Despite the promising benefits of using patient digital twins in healthcare, this application reveals several risks related to privacy and surveillance, amongst others. Saracco refers to the totalitarian surveillance society depicted in George Orwell's *1984* and that digital twins might pose a more significant threat of uncertainty.

1.1 Relevance and Objectives

Gartner's report predicts that half of the large industrial companies will employ digital twins by 2021 and increase their effectiveness by 10% (Cearley & Burke, 2018). More and more fields of application are becoming familiar with the concept of digital twins and are starting to make use of the benefits of this technology. However, digital twins in healthcare is a field that still lacks research and practical application. This master thesis aims to explore how digital

twins can be implemented in healthcare in a privacy friendly manner, taking the benefits and the challenges as well as how they affect the implementation process into account. A strong focus is set on the privacy risks associated with the application of digital twins in the health domain. The goal is to create a conceptual framework of the implementation of digital twins in the field of healthcare that can act as a guide on what needs to be taken into consideration when developing such technology. The findings are expected to contribute to further research on the impact of digital twins in healthcare regarding the implementation process and the privacy debate.

To reach this goal, the thesis aims to answer the following research questions:

Main research question: How can digital twins in healthcare be implemented in a privacy friendly manner?

- **Sub research question 1**: What are the privacy challenges of the implementation of digital twins in healthcare?
- **Sub research question 2**: How do privacy issues affect the implementation of digital twins in healthcare?

The main research question is concerned with implementing digital twins in the healthcare sector while taking privacy protection measures into account. This question involves two aspects. Therefore, two sub research questions were established to specifically address these two sides. On the one side, specific privacy challenges associated with digital twins in healthcare are identified. On the other side, the influence of the identified privacy risks on the implementation of digital twins is of interest. The sub-questions pinpoint the research objectives in more detail. When put together in the end, they will help answer the leading question of how digital twins can be implemented in the healthcare sector by including privacy protection measures.

1.2 Structure of the Report

The present thesis is structured as follows. At first, the subsequent chapter to this introduction includes a literature review that examines existing research relevant to the research goal. This includes an overview of digital healthcare and how digital twins fit into digital health. Moreover, the concept of digital twins, together with its definition and existing application fields beyond healthcare, will be explored before turning to potential challenges connected to digital healthcare. Chapter 3 describes the theoretical framework that serves as the foundation

of this research. It contains frameworks for the implementation of digital twins, privacy design strategies, and theoretical models concerned with surveillance, datafication, and privacy. The conceptual framework that was developed based on these theories and the literature review is presented in this chapter as a result. The methodology is illustrated in chapter 4, which explains the approach that was chosen to answer the research questions mentioned above and clarifies the methods used to collect empirical data. A more detailed explanation of the data collection can be found in chapter 5, together with the findings of the empirical data acquisition. The findings are then related to the theories and literature that serve as the foundation for this thesis, and the research questions are answered (chapter 6). Finally, the last chapter contains a summary and reviews the key insights that can be derived from this research, as well as the limitations of this study. The chapter concludes with an outlook on the next steps of digital twins in healthcare.

2 Literature Review

The purpose of this chapter is to define the necessary terms to create a common understanding and investigate the current state of research in digital healthcare and digital twins. Such a literature review is the base for further research. It reveals findings from previous studies related to the research topic and helps identify areas that require additional contribution (Creswell, 2009). Firstly, a glimpse into digital healthcare research will be given in chapter 2.1, including the paradigm shift that led to the emergence of the application of digital twins in healthcare. Secondly, the concept of digital twins will be defined with its characteristics, underlying technologies, and application fields (chapter 2.2). Chapter 2.3 concludes this section with a review of the challenges that digital healthcare faces.

2.1 Digital Healthcare

Due to the rapid emergence and innovation of information and communication technologies, digital health has gained importance in recent years. Digital health refers to a range of technology innovations related to health and medicine, such as telemedicine and telehealth, health informatics (e.g. electronic patient records), and patient self-care and monitoring (Lupton, 2014b). This includes “e-health” (electronic health), which is concerned with “health services and information delivered or enhanced through the Internet and related technologies” (Eysenbach, 2001), or “m-health” (mobile health), which focuses on healthcare applications offered on mobile devices (Lupton, 2012). Concepts such as the Internet of Things (IoT), big data, and cloud computing play a significant role in the realm of digital health and are immense facilitators of its dissemination (Lupton, 2014b). Therefore, these three concepts will be explained to highlight how these technologies contribute to digital health. This sets the basis for further investigation of the current state of digital health technologies and their benefits to healthcare.

Internet of Things

In a nutshell, the Internet of Things refers to the interconnection of physical objects (“things”) over public or private Internet Protocol networks. These physical objects are not limited to static devices. IoT can be an interaction through the Internet of people to people, people to machine/things, or machine/things to other machine/things (Patel & Patel, 2016). Swan (2012b) defines four functional layers, which act as a technical guide to understand the interrelation between the different components in the IoT ecosystem. The layers relate to (1) data acquisition,

(2) information creation, (3) meaning-making, and finally (4) action-taking. Firstly, to collect data, the real-world objects are connected to the Internet by equipping them with sensors (Swan, 2012b). This allows for the communication between the physical entity and the digital world, transforming the physical objects into so-called "smart" objects (Patel & Patel, 2016). Sensors can collect all sorts of data, depending on the desired parameters necessary to reach the goal of a specific use case. Examples for possible parameters are temperature, location, movement, but also heart rate or blood flow volumes (Swan, 2012b). In the second layer, information creation, the collected data are processed and transmitted using technology standards like Wi-Fi, Bluetooth, or 4G. The acquired data are then transformed in the third layer, making it comprehensible to the human to facilitate meaning-making. Swan (2012b) states that visualisation is a powerful tool to represent data intuitively. The final layer – Swan alternatively calls this the "So What?" layer – is concerned with action-taking. "The 'So What' layer makes sense of the data, and allows action items to be derived based on real information" (Swan, 2012b, p. 235). The resulting action step, however, may not always be immediately evident from the data as the data flows can be totally new, which is why the layer has this alternative name.

The introduction and vast adoption of IoT in the healthcare sectors have led to the establishment of new digital health technologies regarding medical consultation, diagnosis and treatment, medical equipment management, and health promotion, amongst others, and are supposed to improve conventional healthcare measures. There are many digital health applications designed to promote self-tracking to self-monitor one's health. An accessible technology that people integrate into their everyday lives are wearables – sensor-equipped devices that patients wear on their bodies and that collect different physiological parameters. The sensors can then be connected to an application on a person's smartphone, for instance, where the collected data are visualised and analysed to detect any changes (Sharon, 2016). These self-tracking devices prove to be beneficial for the owner's health. Vogel et al. (2017) examined the use of smart wearables in treating cardiovascular disease by employing physical activity. Their results show that wearing a smart bracelet that tracks physical activity (including the number of steps, calories burned, and time spent sitting) can increase physical activity and improve the overall performance of the cardiovascular system. Johns Hopkins Medicine together with Apple developed a health digital platform consisting of "an Apple CareKit App with a collaborative Apple Watch application, Bluetooth blood pressure cuff, and backend data monitoring" (Spaulding et al., 2019, 2) to help patients recover from heart attacks and prevent readmissions ("Corrie Health," 2019). The platform is concerned with monitoring medication and vital signs

and emphasises educating the users about cardiovascular diseases. The preliminary study suggests that patients engaged in the platform have a lower risk of being readmitted in comparison to those that undertook conventional treatment (Dobkowski, 2019). Next to wearables like bracelets or smartwatches carried by the patient, sensors fixed straight onto the skin promise additional and more precise measurements for healthcare. In 2011, Kim et al. introduced a tiny skin patch that monitors the electrical activity of the heart, brain, and skeletal muscles. They name this approach "epidermal electronics" and describe it as the following:

> *"[T]he electrodes, electronics, sensors, power supply, and communication components are configured together into ultrathin, low-modulus, lightweight, stretchable 'skin-like' membranes that conformally laminate onto the surface of the skin by soft contact, in a manner that is mechanically invisible to the user, much like a temporary transfer tattoo."* (Kim et al., 2011, p. 838)

S. K. Ameri et al. (2017) propose an even less obtrusive version of an electronic tattoo made of graphene, a very thin electrically conductive material. They demonstrate that this type of tattoo adapts better to the skin's reaction of stretching during movement without being damaged as easily. According to the authors, electronic tattoos allow for long-term, high-fidelity biometric sensing and have the advantage of sitting extremely close to the body as opposed to mechanic wearable sensors mounted in a bracelet.

In the area of hospital care, Hu, Xie, and Shen (2013) define the term "medical Internet of Things" as "a kind of technology that embeds wireless sensors in medical equipment, combines with the internet and integrates with hospitals, patients and medical equipment to promote the new development of modern medical model" (p. 2054). They further state that the application of medical IoT can benefit the fields of medical equipment and medication control, medical information management, telemedicine, and mobile medical care, for example the facilitation of quick medical information sharing and transmission between medical institutions, and remote diagnosis and consultation. Moreover, medical IoT can be applied in health management, which has the goal to prevent, detect and contain diseases, decrease medical cost, and detect factors that impede the general health status to increase the quality of life. It includes the collection of data from the physical object through sensors that send the medical data to a household device (e.g. phone, computer) which can then be forwarded to management facilities that process it and define action steps (Hu et al., 2013) (see Figure 1). This also underlines the general functionality of the Internet of Things on an abstract level, as introduced by Swan (2012b).

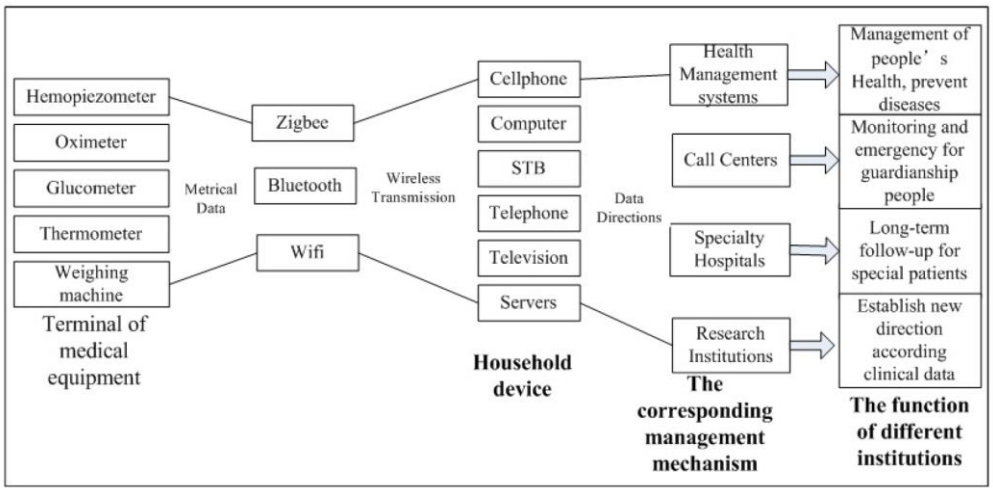

Figure 1. The relationship between the Internet of Things and Health Management (Hu et al., 2013, p. 2058).

Big Data and Datafication

Due to the technical possibility and financial feasibility of collecting a lot of data through smartphones and wearables, as mentioned above, the concept of "big data" emerged. Most scholars agree that big data is defined by three Vs – volume, variety, and velocity (Johnson, Friend, & Lee, 2017; Roski, Bo-Linn, & Andrews, 2014) – while others argue that its characteristics are composed of five Vs, adding veracity and value (Demchenko, Ngo, Laat, Membrey, & Gordijenko, 2013; Ishwarappa & Anuradha, 2015; Kaur & Sood, 2017). An overview of the five Vs of big data can be found in Figure 2. Volume is concerned with the extensive amount of data, whereas variety means that the data can be of various types and formats, both structured and unstructured. With big data infrastructure, the collection, processing, and storage of data are made easier, which allows for more flexibility and rapidity concerning data management. This is big data's third characteristic: velocity (Roski et al., 2014). Veracity of data refers to the quality and accuracy of the data and whether it is correct or "dirty". The resulting analysis thereby greatly depends on the credibility of the data source. Finally, big data's fifth and most important feature is value. To exploit big data's full potential, it is necessary to draw value from the combination of data volume and variety, the IT infrastructure, the emerging velocity of data analysis, and its veracity (Ishwarappa & Anuradha, 2015).

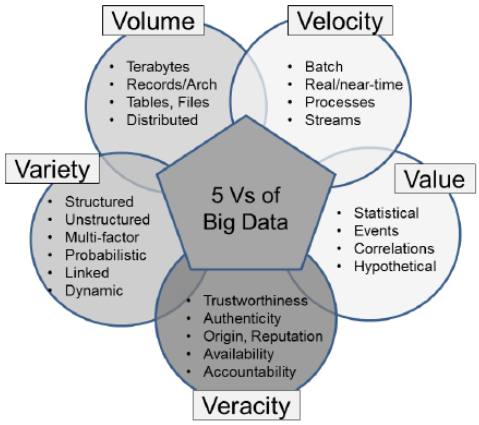

Figure 2. 5Vs of big data (Demchenko et al., 2013, p. 79).

Regarding value, Roski et al. (2014) draw from multiple sources and conclude that one of big data's potential value lies in improved clinical value driven by clinical decision support systems, personalised diagnostic and treatment decisions based on patient-generated data or a patient's detailed risk profile, patient behaviour modifications as well as big data-driven population health analyses. Big data can support decision-making by creating decision models based on individuals' data. The collection of large amounts of individual-level health data such as medical history, diagnoses, and treatments from electronic health records and mobile applications, hospital admission, and recharge, but also health insurance enrolment information together with advancing computing power can help to build prediction models. These models are then translated into actions (e.g. treatment choices) and, in turn, support decision-making by illustrating outcome probabilities and possible drivers that influence the different scenarios (Bjarnadóttir et al., 2014). "Aggregated health care data can help build a picture of the overall functioning of the health care system and has the potential to support health care decision for a wide range of stakeholders" (Bjarnadóttir et al., 2014, p. 287), such as entrepreneurs, policymakers, and researchers (Bjarnadóttir et al., 2014). A concept that emerged from the facilitated access to and storage and processing of massive amounts of big data is datafication. "To datafy a phenomenon is to put it in a quantified format so it can be tabulated and analyzed" (Mayer-Schönberger & Cukier, 2013, p. 78). Datafication offers a new lens through which people's behaviour and society as a whole can be viewed and understood by looking at "metadata" collected through social media platforms, for instance (van Dijck, 2014). Public health surveillance is a field that can benefit from big data (Richterich, 2018) and datafication by utilising different sources of information on the Internet (e.g. social media, blogs, search engine queries) for health promotion, public health initiatives (Lupton, 2017), or digital disease detection (Brownstein, Freifeld, & Madoff, 2009).

Cloud Computing

The short and concise definition by Mell and Grance (2010) describes cloud computing quite well. “Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (for example, networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service-provider interaction” (Mell & Grance, 2010, p. 50). Next to being very cost-effective (Kuo, 2011), the big benefit of using cloud computing in healthcare is that collected information from sensors can be sent to the “cloud” automatically via a wireless network and is stored there for further processing and analysing. This enables automated real-time data collection and reduces human transcription errors. Moreover, it can be accessed easily by other people or systems regardless of their location, facilitating knowledge distribution (Rolim et al., 2010). These aspects are crucial for IoT, as described earlier, which demonstrates overlapping borders between the Internet of Things, big data, and cloud computing.

The Shift Towards Personalised Preventive Care

As shown above, IoT, big data, and cloud computing are strong facilitators to the digital health environment. Due to these advancements, the purpose exceeds that of monitoring the current health status and shifts towards preventive medical care (Lupton, 2012). The patient actively participates to better understand and improve their health condition, recognise early signs of deviations, and initiate countermeasures, facilitated by technologies like self-tracking (Swan, 2012a). Predictive, Preventive and Personalised Medicine (PPPM) is an innovative concept that accentuates exactly what its name promises and is believed to be “the medicine of the future” (Chaari, 2019, vi). PPPM aims to predict if an individual is susceptible to a disease before its outbreak. That way, preventive measures and personalised treatment algorithms that are both directly aimed at the respective person can be provided (Chaari, 2019). The European Association for Predictive, Preventive and Personalised Medicine (EPMA) describes their vision as “to promote the paradigm change from delayed reactive medical services to evidence-based Predictive, Preventive & Personalised Medicine (PPPM) as an integrated science and healthcare practice” (European Association for Predictive, Preventive and Personalised Medicine, n.d.). Meskó, Drobni, Bényei, Gergely, and Győrffy (2017) go as far as saying that the shift that digital health is experiencing is not limited to be technological, but also comes with a cultural transformation due to the changing role of the individual in the healthcare domain.

Digital health today can be exercised by anyone owning a smartphone. Parameters such as heart rate, step counts, calorie intake, sleep quality, and mental well-being can be collected and used to not only monitor and improve the individual's health condition but also prevent illnesses from happening to a certain degree. At the same time, risks such as privacy implications and surveillance must not be overlooked. The negative side effects of digital health are delineated in more detail in chapter 2.3.

2.2 Digital Twins

The concept of digital twins facilitates the paradigm shift of digital health towards being more preventive due to the possibility of simulation and prediction based on (big) data analysis, artificial intelligence, and machine learning. The following subchapter will define what a digital twin is and explore its application fields, as well as the role digital twins currently and could potentially play in healthcare.

2.2.1 Definition and Characteristics of Digital Twins

Michael Grieves first explored the concept of digital twins in 2003 at the University of Michigan in the field of product lifecycle management (Grieves, 2014). He defines a digital twin as "a virtual, digital equivalent to a physical product" (Grieves, 2014, p. 1) consisting of three main components that are "a) physical products in Real Space, b) virtual products in Virtual Space, and c) the connections of data and information that ties the virtual and real products together" (Grieves, 2014, p. 1). Many scholars follow Grieves' approach and highlight three main elements of a digital twin: the physical component, virtual presentation, and their interconnection (Cearley & Burke, 2018; Glaessgen & Stargel, 2012; Tao, Cheng et al., 2018, 2018). Nevertheless, other scholars have evolved what Grieves sketched out by adding new perspectives and incorporating technological change. Grieves' definition focuses on machines, non-living objects. El Saddik (2018) underlines that the physical entity should be extended to include living ones as well. Research today reveals that this is indeed possible. The specific application fields will be discussed more detailed in chapter 2.2.3. Gartner's trend report supports this notion as well by stating that digital twins are a "digital representation of a real-world entity or system" (Cearley & Burke, 2018, p. 16), thereby broadening the focus of Grieves' "physical products" to include all things in real life. Tao, Zhang et al. (2018) built upon Grieves' three-component structure and proposed a five-dimensional model. In their paper on managing healthcare equipment, they added digital twin data and services. The digital twin

data model has data from multiple entities of the whole digital twin ecosystem. It includes data from the physical entity, data from the virtual representation, data from the services domain, as well as their fusion. The services aspect focuses on the services for both the physical and the virtual entity. The services model is composed of the function, input, output, as well as the quality and state of the services.

For this research, the following characteristics remain within the bounds of a digital twin setting:

1. Physical entity (physical twin)

While this is not the digital twin per se, the physical entity plays a crucial part as it is the object (or subject) the digital twin mirrors. Haag and Anderl (2018) call this component the "physical twin". Examples for such a physical entity will be given when exploring the application fields of digital twins in 2.2.3.

2. Virtual representation of the physical entity (digital twin)

The digital representation of the physical entity is known as the actual digital twin, the virtual doppelgänger, so to speak, and can mimic and simulate the physical twin's behaviour. It evolves in accordance with its physical counterpart (Haag & Anderl, 2018). Nevertheless, this does not necessarily mean that the digital twin internalises every detail of its physical equivalent. Instead, it depicts the key features that are of importance to meet a particular purpose (Batty, 2018). The digital twin is built based on behavioural, as well as technical characteristics and features of the physical part.

3. Connection between the physical entity and its virtual representation

Data and information are part of the digital twin set-up as they link the physical and the digital aspects (Grieves, 2014; Tao, Qi, Wang, & Nee, 2019). Tao, Cheng et al. (2018) attempt to make a more specific distinction between the different sources of data and their relationship with each other. Firstly, information is collected from the physical entity through sensors, or other collection means embedded in the physical twin. These data synchronise with the virtual entity and are updated continuously to ensure that the digital twin is always a "real-time reflection" (Tao, Cheng et al., 2018, p. 3566) of the physical entity. Secondly, the data interact and converge on different levels: 1) on the physical side only, 2) between historical data (or expert knowledge) and real-time data, and 3) between physical and virtual space (Tao, Cheng et al., 2018).

4. Added value of the connection between the physical and virtual component

This characteristic is what distinguishes a digital twin from a simple digital model. The link between the physical product and the virtual representation creates value by supporting the physical system regarding visualization, analysis, maintenance, prediction, simulation, and optimisation of operations (Cearley & Burke, 2018; Dohrmann, Gesing, & Ward, 2019; Grieves, 2014). This concept can save time and resources during the design stage as simulations and prototypes can be made virtually. Possible future behaviour, both desired and undesired, can be detected, and measures to elicit or prevent these behaviours can be undertaken (Grieves & Vickers, 2017). Thus, the main goal of a digital twin is "to learn from the past, understand the present, and predict the future to achieve improved business outcomes" (Bohm, 2018, p. 43).

To sum up, digital twins consist of a physical and a virtual component, as well as their interconnection. Necessary data and information are collected, exchanged, and updated continuously at regular intervals. The connection between the physical and the virtual object can help enhance the physical entity's performance (e.g. maintenance, prediction, simulation, optimization).

Underlying Technologies

Dohrmann et al. (2019) give a brief overview of the fundamental technologies that digital twins use. These include the Internet of Things, cloud computing, APIs (application programming interfaces) and open standards, artificial intelligence, and augmented, mixed, and virtual reality (Dohrmann et al., 2019). IoT and cloud computing were explained in the previous chapter; therefore, the remaining technologies will be explained briefly. More and more APIs and open standards are becoming publicly available and facilitate data sharing and exchange, leading to the possibility of interlinking data from various sources. Using real-time and historical data, artificial intelligence and machine learning, allow for systems to make autonomous decisions and predictions. Finally, augmented, mixed and virtual reality enable the visualisation of digital models in a three-dimensional environment, the interaction with such digital models in a physical environment, and the development of new virtual environments. Figure 3 gives an overview of these technologies and how they are used in the context of digital twins.

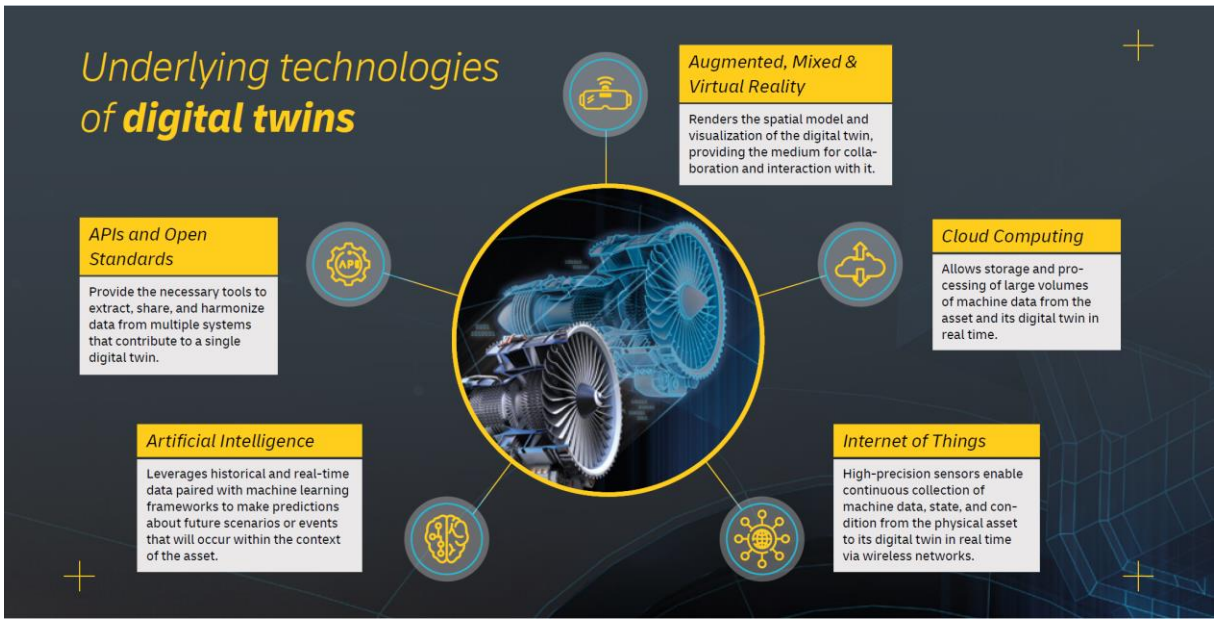

Figure 3. Technologies behind digital twins (Dohrmann et al., 2019, p. 7).

2.2.2 Application Fields of Digital Twins

Gartner's report (Cearley & Burke, 2018) mentions the following five business opportunities for digital twins: improved maintenance and reliability, business process and asset optimisation, monetization of data and models, research and development, and the establishment of new business models. These opportunities can take place in different settings as digital twins have been applied in several fields, such as manufacturing, aerospace, city planning, and healthcare. The following section will give an overview of what digital twins can accomplish in these fields.

Product Lifecycle Management and Manufacturing

In the field of product lifecycle management and manufacturing, the potential benefits through digital twins are recognized, especially when it comes to maintenance, prediction and sustainability (Schleich, Anwer, Mathieu, & Wartzack, 2017; Tao, Qi et al., 2019; Tao, Sui et al., 2019). Grieves (2014) clusters the areas in which digital twins can help humans assess situations in the scope of manufacturing into conceptualization, comparison, and collaboration. By conceptualization, Grieves means that "the digital twin lets us directly see the situation and eliminate the inefficient and counterproductive mental steps of decreasing the information and translating it from visual information to symbolic information and back to visually conceptual information" (Grieves, 2014, p. 4). Simulations are of major help here as they visualise possible scenarios instead of having to manually conceptualise them based on performance reports. Digital twins further support decision-making when it comes to adjusting future operations by

comparing different values and scenarios with each other. Finally, digital twins simplify sharing digital models, concepts, and real-time production with other individuals, despite their location, allowing for improved collaboration. Tao, Cheng et al. (2018) identify similar benefits enabled by digital twins for manufacturing. The researchers highlight the lack of convergence between physical and virtual space and emphasize how accessing, controlling, and maintaining a product on-site at a customer can be achieved with the help of a digital twin. They propose digital twin-driven solutions for product lifecycle management in the areas of *product design*, *manufacturing,* and *service*, leading to more efficiency and sustainability.

- **Digital twin-driven product design:** Product design can benefit from the digital twin technology in all its stages, which the authors categorise into conceptual design, detailed design, and virtual verification. In the first stage, the conceptual design, digital twins assist by integrating a large amount of data related to the product (for instance customer satisfaction or product sales) into one point of information. In addition, the product can be improved faster and more precisely as the communication between designers and clients is facilitated due to real-time feedback transmission. As for the step of detailed design, prototypes are built and assessed for their performance with simulation tests. With the help of real-time data from the product and its environment, these simulations reach more reliable results. In the last stage of virtual verification, a small amount of the product is produced and checked for its validity and feasibility. “[U]sing digital twin technology, designers can create vivid simulation scenarios to effectively apply simulation tests on prototypes and accurately predict the actual performance of the physical products as far as possible” (Tao, Cheng et al., 2018, p. 3568), all without having to physically produce. Real-time data, together with data from previous simulations, enables the identification of defects and their origin as well as finding solutions to the error.

- **Digital twin-driven manufacturing:** Regarding manufacturing, Tao, Cheng et al. (2018) refer to a digital twin shop-floor. Tao and Zhang (2017) utilised the digital twin technology to improve a shop-floor system not only during but also post and prior to production. Their digital twin solution consists of the physical (PS) and virtual shop-floor (VS), a shop-floor service system (SSS) as well as the shop-floor digital twin data (SDTD). The latter is the heart of the whole system. It generates two types of data from the other three components: 1) simple, raw data collected through sensors in the PS, and 2) "fused" data, meaning data that has been processed with other physical or virtual data (e.g. data comparison or clustering). With this, the PS can receive optimal orders based on evaluations and predictions priorly made by the VS. Additionally, optimizations strategies for the SSS help to manage and control the PS and improve the VS (Tao & Zhang, 2017).
- **Digital twin-driven service:** In the service domain, Tao, Cheng et al. (2018) list nine categories of services that can be supported by digital twins, of which service of real-time state monitoring, service of user management and behaviour analysis, and service of intelligent optimization and update are examples for such services. Here, the focus lies on the post-sale tasks such as product usage and maintenance. The latter, especially, is of utmost importance as defects can have severe consequences regarding system failures and general safety (Tao, Cheng et al., 2018). Digital twins allow for the detection of different kinds of behaviours that might inhibit a system's unobstructed use. Grieves and Vickers (2017) describe four categories: Predicted Desirable (PD), Predicted Undesirable (PU), Unpredicted Desirable (UD), and Unpredicted Undesirable (UU) (see Figure 4). Taking these behaviours into account, digital twins seek to reach PD, get rid of PU, and reduce UU system behaviour.

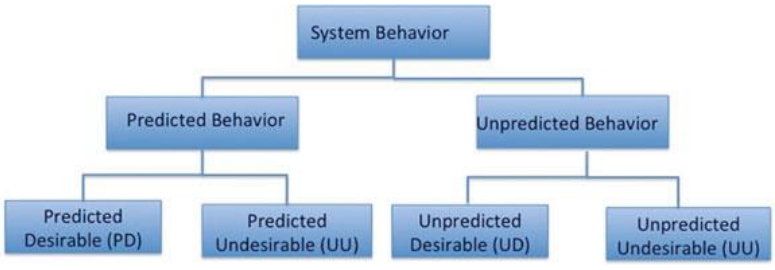

Figure 4. Four categories of system behaviour (Grieves & Vickers, 2017, p. 90).

F. Ameri and Sabbagh (2016) introduced the "digital factory", a digital twin of a real factory, that represents the technological capabilities of a supplier and thereby supports the decision-making in supply chain partnerships and facilitates automated supply chain formation.

Wang, Ye, Gao, Li, and Zhang (2019) used the concept of digital twins for monitoring the state of rotating machinery in manufacturing processes. By developing a prototype based on a digital twin reference model for rotating machinery fault diagnosis, the diagnosis error was reduced to less than 5% (Wang et al., 2019).

The examples show that digital twins have been applied to a smaller extent, for instance when maintaining machinery as well as in larger settings where this technology is used throughout the complete lifecycle, from product design to production and, finally, to its usage.

Smart Buildings and Smart Cities

Integrated Environmental Solutions (IES) offers a range of application examples for digital twins for the built environment to create solutions for sustainable buildings. Together with the University of Nottingham, the software and consultancy company developed an interactive online platform that visualizes both historical and real-time energy data from buildings and utilities of the Trent Basin neighbourhood in Nottingham. The data are collected using smart home and IoT technologies and are supposed to help predict energy consumption and behaviour (IES, 2018). What is more, the platform allows its residents to access and interact with the data to educate them about their energy use and make energy-efficient decisions (Woods & Freas, 2019). In 2014, IES and the Nanyang Technological University in Singapore launched the EcoCampus initiative intending to reduce their energy, water, and waste footprints by 35% by 2020 and be the world's most eco-friendly campus. The digital twin solution used energy and build stock data from campus buildings to give an overview of the current energy consumption and identify areas where energy could be saved. It also allowed for simulation and comparison of energy-saving technologies (e.g. high resistance envelope for walls and roofs or windows, or lighting occupancy sensors) to determine which one worked best (Woods & Freas, 2019). The simulations revealed a 31% energy savings potential and approximately 4.7 million Singapore Dollars financial savings (IES, n.d.).

From one building to multiple buildings to a whole city – digital twins have reached the domain of city planning by helping to create so-called "smart cities". Smart cities are cities that use information and communication technology (ICT), such as networks, sensors, and intelligent management systems, to connect things, organisations, and people. The goal is to improve the city's state in areas like infrastructure, sustainability and energy efficiency, and citizen participation (Clarke, 2013). The digital twin technology can foster the transformation to smart cities, as Mohammadi and Taylor (2017) demonstrate in their smart city digital twin paradigm

based on the city of Atlanta. A 3D model of the city, consisting of infrastructure data such as transport, power, or water supply, and human/social networks, monitors the dynamics and performance of the city and supports in terms of operation, resource allocation, and consumption. Another example would be *Virtual Singapore*, a project of the National Research Foundation of Singapore. It is a 3D model of the city, containing 3D maps and city components such as water bodies, vegetation, transportation infrastructure, and buildings. Other real-time data and information like demographics, movement, and climate will be fed into the city's digital twin as well. It is a collaborative data platform aimed to be used by the public, private, people, and research sectors once completed. Its goal is to "enable users from different sectors to develop sophisticated tools and applications for test-bedding concepts and services, planning, and decision-making, and research on technologies to solve emerging and complex challenges for Singapore" (National Research Foundation Singapore, 2018). The government highlights that Virtual Singapore can be used for virtual experimentation, virtual test-bedding to test the provision of services, planning and decision-making, as well as research and development in terms of the innovation and development of new technologies.

Other Application Fields

The application of digital twins is certainly not limited to the fields mentioned above. There are many other areas digital twins see use. According to Grieves and Vickers (2017), NASA can address some of their major issues with the use of digital twins. The biggest problem, in their opinion, is that NASA produces few yet costly systems that have not been created before. Instead of producing prototypes that cost a lot of money and time, digital twins could address these issues in a more efficient and effective way. Glaessgen and Stargel (2012) took a look at the applicability of digital twins for NASA and US Air Force Vehicles fleet management. By incorporating information from the physical model, sensor updates, or fleet history, for instance, the "flying twin", as they call it, can mimic its physical counterpart. With data from the integrated vehicle health management system as well as maintenance and other historical data, this information can support the maintenance of these vehicles in the long run. Moreover, it leads to detecting atypical, formerly unknown behaviour, and making more informed decisions regarding the consequences of alterations to a vehicle's mission. This, in turn, increases the probability of mission success. Tuegel, Ingraffea, Eason, and Spottswood (2011) propose a digital twin for predicting the structural life of aircrafts. Similar to what Glaessgen and Stargel (2012) described, the vehicle's – in this case, the aircraft's –, condition can be

monitored faster, and its maintenance can be better managed with the help of this technology (Tuegel et al., 2011).

This brief review of the application fields of digital twins showed how this technology is applied and the benefits that come with it. There are mutual benefits across different application areas that were repeatedly mentioned in the literature, namely real-time monitoring, simulation, and prediction. To sum up, the digital twin is updated regularly with real-time data, enabling the monitoring of the current condition of the physical object and detecting faulty areas faster by looking at the digital model. The physical object's digital model allows for simulations to be carried out and eliminates the necessity of building physical prototypes. Simulations also help to test out different behaviours and scenarios as well as new technological solutions to see how the physical object might react and make better decisions regarding improvements. The benefits of being able to predict behaviour and changes of the physical object are a big part of the benefits of digital twins. With the help of different data sources and artificial intelligence, the digital twin can forecast the consequences certain variables have on the product. This allows for better maintenance as well as performance optimisation of the physical twin.

Within the scope of this research, the focal point is on the application of digital twins in healthcare. While the previous section had the purpose of highlighting the broad application field of digital twins and how these fields benefit from this technology, the next section will go into depth about digital twins in the health sector.

2.2.3 Digital Twins in Healthcare

Digital twins in healthcare fall under the umbrella term of digital health technology. What makes research in digital twins in (digital) healthcare an especially interesting matter is that the physical component can be a living one (patient digital twin). Lauzeral et al. (2019) explored digital twins in the context of biomechanical modelling. The researchers focused on real-time human liver models and their deformation resulting from breathing by using patient-specific medical data. Liu et al. (2019) developed a cloud healthcare system based on digital twins called CloudDTH. The research offers valuable insights into the overall design of a scenario that involves the patient, the digital twin, and the doctor, as well as the technical implementation of a digital twin prototype. CloudDTH should support personal health management, especially for the elderly, in monitoring, diagnosing, and predicting their health status. The core elements in the framework are the physical object, the virtual object, the cloud healthcare service platform,

and the healthcare digital twin data. The data are a combination of personal data, which can be derived from medical exams, for instance, third-party data from insurances, but also simulation data. Individuals can further collect data outside the clinic through wearable devices that measure heart rate, blood pressure, body fat, or blood glucose (Liu et al., 2019). These devices are a key aspect as digital twins in healthcare as well as in any other field depend on constantly updated data (Grieves, 2014). Mayer-Schönberger and Cukier (2013) call this "datafication". "To datafy a phenomenon is to put it in a quantified format so it can be tabulated and analyzed" (Mayer-Schönberger & Cukier, 2013, p. 78). The authors further refer to "self-trackers" that "measure every element of their bodies and lives in order to live better – or at least, to learn new things they couldn't have known in an enumerated way before" (Mayer-Schönberger & Cukier, 2013, pp. 94–95). Therefore, the concepts of big data and datafication are indispensable when discussing digital twins in healthcare. Employing digital twins in healthcare might be a step beyond self-tracking and a possible solution to detect diseases in a timely manner and employ treatment accordingly. One example of a concept that comes close to that of digital twins is "The Living Heart Project" (Dassault Systèmes). The collaborative platform offers a default model of the cardiovascular system, which can be customised to a patient using their medical data. The model is able to simulate the behaviour of the patient's cardiovascular system and allows clinicians to run tests on the model to find the best treatment option for the patient (Shugalo, 2019). Björnsson et al. (2019) support the usage of digital twins for personalised medicine and treatment. They investigate how the digital twin technology can help identify the drug that has the best effect on the patient by using digital twins of the patient and testing treatment scenarios with these twins. Figure 5 shows their illustration of this process.

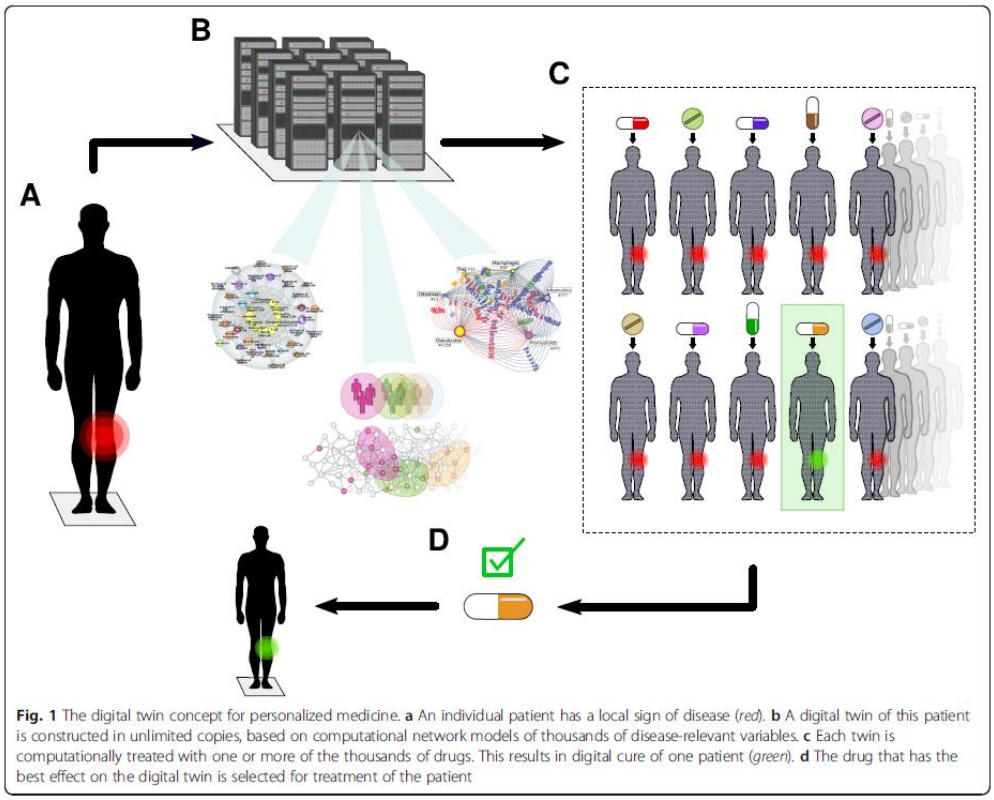

Fig. 1 The digital twin concept for personalized medicine. **a** An individual patient has a local sign of disease (*red*). **b** A digital twin of this patient is constructed in unlimited copies, based on computational network models of thousands of disease-relevant variables. **c** Each twin is computationally treated with one or more of the thousands of drugs. This results in digital cure of one patient (*green*). **d** The drug that has the best effect on the digital twin is selected for treatment of the patient

Figure 5. The digital twin concept for personalised drug treatment (Björnsson et al., 2019, p. 2).

Utilising digital twins in healthcare is gaining importance in research as well as in the industry. Interdisciplinary organisations such as DigiTwins (https://www.digitwins.org), or the Swedish Digital Twin Consortium (https://www.sdtc.se) have emerged to drive the application of digital twins in personalised medicine and companies such as Siemens (Siemens Healthcare GmbH, 2019) or Philips (van Houten, 2018) have recently started to invest in research in digital twins in healthcare as well.

This subchapter focused on the definition and characteristics of digital twins, their application fields, and the benefits of this technology. Applying digital twins in the health sector, as it is the focus of this research, promises to revolutionise healthcare by offering real-time monitoring of a patient's health status, early disease detection, and improved personalised treatment, amongst others. As digital twins are applied very close to the individual, it is just as important to look at the issues that digital twins potentially pose.

2.3 Challenges of Digital Healthcare

The application fields prove that digital twins are a versatile technology and bring many benefits to different areas of business. While it is important to look at the benefits of digital twins and

the reasons why this technology is so revolutionary, it is just as significant to investigate the risks and challenges that can be associated with digital twins. While this project focuses on the healthcare domain, there are not many practical implementations and studies regarding digital twins in the field of health. As digital twins in healthcare belong under the umbrella term of digital healthcare, this chapter will focus on some of the challenges that digital healthcare in general poses.

Commodification of Personal (Health) Data

Back in 2010, Lyon identified the challenge of collecting data for reasons other than well-being and preventive care and drew first conclusions to personal data being treated as commodities. An example is online patient-support networks, consisting of blogs, Facebook groups, Twitter hashtags, and YouTube videos that allow people to not only inform themselves about medical information, diagnoses or treatment options, but to also share their own experiences (Lupton, 2014a). Lupton (2014a) discusses the emergence of a "digital patient experience economy, in which patients' online accounts and details of their medical conditions and their ratings and opinions of healthcare providers and institutions have become valued not only for the support and information they offer to other patients but also for the increasing commercial or research value they have for others" (p. 858). The patient-generated content is exploited and sold by for-profit companies without compensating the creator. The notion of commodifying social media data, in general, is well supported by researchers (for example Floridi, 2015; Keen, 2015; Mai, 2016; Taplin, 2017; van Dijck, 2014). In the healthcare sector, social media (big) data can also be used for improving public health. In the context of digital disease detection, Vayena, Salathé, Madoff, and Brownstein (2015) point out that public health is a common good, which raises the question of the trade-offs between maintaining an individual's rights and serving the common good. Big data especially shines a new light on this debate. For instance, social networking data that is initially collected for public health purposes, e.g. digital disease detection (Vayena et al., 2015) or mental health (Coppersmith, Dredze, & Harman, 2014), can also be used for making corporate profit, such as advertising (Vayena et al., 2015). The example of using the same data for both health and profit purposes shows that the vast collection of digital personal health information in general (not only through social media) reveals ethical concerns of commodification without compensation.

Surveillance and Behaviour Manipulation

Third parties (especially advertising companies) that buy personal data acquired through social media platforms often employ "behavioural targeting". Here, the aim lies in tailoring ads that fit a consumer (Brown, 2016). Exploiting data for corporate venture has become a business model, and terms such as "surveillance marketing" (Taplin, 2017), "dataveillance" (Raley, 2013), and "surveillance capitalism" have emerged. Shoshana Zuboff (2019) coined the latter. The concept is related to a system that uses personal data to generate behavioural predictions and sell these predictions in so-called "behavioural futures markets". Beyond knowing what an individual will be doing in the future, the goal is to shape their behaviour as well: "[…] it is no longer enough to automate information flows *about us*; the goal now is to *automate us*" (Zuboff, 2019, p. 8). This concept is closely connected to the next challenge of digital health – surveillance. Lupton (2012) recognises the use of mobile devices for medical care as part of the "surveillance society". Surveillance society is a term coined by David Lyon. First of all, he defines surveillance as the following:

> *"[I]t [surveillance] is any collection and processing of personal data, whether identifiable or not, for the purposes of influencing or managing those whose data have been garnered. […] Today, the most important means of surveillance reside in computer power, which allows collected data to be stored, matched, retrieved, processed, marketed and circulated" (Lyon, 2002, p. 2).*

As a next step, surveillance society means the inclusion of surveillance technologies into everyday life (Lyon, 2010). Surveillance especially threatens an individual's equality, freedom, and privacy (Floridi, 2015). Equality can be threatened by ideas such as "social sorting", an associated risk of surveillance identified by Lyon (2010). With automated classification based on all sorts of data, big corporations can make judgements that directly impact people's lives (Lyon, 2010). His concern brought to paper in 2010 is about to manifest itself ten years later in China (Campbell, 2019). This can foster exclusion (Lupton, 2012), inequality, and discrimination (Montgomery, Chester, & Kopp, 2018). O'Neil (2016) gives the example of wellness programs in companies and how employees are penalised if they do not reach certain health metrics. To monitor the employee's metrics, wearables can be a means for verification. This inhibits the employees' freedom of choice and forces them to monitor their lives for the sake of the wellness program, in case they cannot pay the penalty (Christl & Spiekermann, 2016). Limiting an individual's choice distorts their sense of freedom and is connected to Zuboff's idea of using big amounts of data for not only monitoring but moulding an individual's

behaviour. Together with the power of big data-driven nudging, which aims at personalised selection optimisation, behaviour manipulation can be achieved and abused just as easily (Yeung, 2017).

Privacy Paradox and the Greater Public (Health) Good

In 1967, Alan Westin defined "privacy" as the "claim of individuals, groups, or institutions to determine for themselves when, how and to what extent information about them is communicated to others" (Brown, 2016). At the base of the aforementioned challenges lies the violation of what Westin defined – personal health data are passed on to third parties and are used for other purposes, out of the control of the data subject. Therefore, maintaining the protection of one's privacy is a very big, if not the most challenging issue faced by digital health. The protection is further hindered by what is called the privacy paradox. This paradox is concerned with the contradiction of highlighting the importance of privacy concerns while openly sharing personal information online (Barnes, 2006; Hargittai & Marwick, 2016). The privacy calculus theory can offer an explanation for the privacy paradox (Chen, 2018; Pentina, Zhang, Bata, & Chen, 2016). Li, Wu, Gao, and Shi (2016) researched the adoption of healthcare wearable devices and concluded that this decision is determined by a risk-benefit analysis (i.e. the privacy calculus theory), meaning that if the benefits outweigh the privacy risk, the chance of adopting the device is higher. "This calculus governs the decision-making process of individuals to decide whether to disclose personal information" (Majumdar & Bose, 2016, p. 192). Perceived risk – the fear of one's information being misused wrongly – and perceived benefit – the potential positive outcomes – are the key components linked to privacy calculus (Majumdar & Bose, 2016). In healthcare, health-related emotion was revealed to be an additional strong factor in the calculus, weakening the relationship between perceived risks and intention to use technology for healthcare. This implies that people are less likely to worry about the risks and aim their attention more on the benefits when they have high condition emotions (Rahman, 2019). This perception, in turn, fosters the privacy paradox. More recently, Adorjan and Ricciardelli (2019) identified a "new privacy paradox". They found out that the attitude towards privacy has shifted amongst young adults to a "nothing to hide" mindset, i.e. privacy is not of major relevance to them. Solove (2007) discussed this argument in the context of surveillance by the US National Security Administration (NSA), stating that the problem lies in the understanding of privacy: "there is no threat to privacy unless the government uncovers unlawful activity, in which case a person has no legitimate justification to claim that it remain private" (p. 746). This shift in people's mindset today, however, must not be seen as an excuse

to disregard the importance of maintaining an individual's privacy. Together with privacy, confidentiality and security are main issues of digital health technologies because of the big amount of data flow between devices and systems, according to Jumelle and Ispas (2015). Data breaches of health information systems that hold personal health information about an individual are at the forefront of the threats. A health data breach occurs when personal health records are lost or revealed to any other party that is not involved in the confidentiality agreement (Khan & Hoque, 2016). A data breach happens due to cyber-attacks, hacking of databases, or data kidnapping (data held by hackers who ask for blackmail money for return) (Vayena, Haeusermann, Adjekum, & Blasimme, 2018).

The debate about privacy and (mobile) applications that are supposed to have beneficial effects for an individual's but also for the general population's health has been rolled up due to recent events regarding the pandemic. In light of COVID-19, the idea of developing a mobile tracing application unfolded. This application is supposed to track down and inform people who were recently in contact with an infected person faster, thereby helping to contain the spread of the virus. Critics warn about "hastily written software" that might pose serious threats to privacy and has the ability to offer easy access for hackers and government surveillance (Valentino-DeVries, Singer, & Krolik). Cho, Ippolito, and Yu (2020) point out that the weak points of mobile apps for contact-tracing for COVID-19 lie in guaranteeing privacy from snoopers (linkage attacks, in which an unauthorised actor can potentially link data to a user), privacy from contacts (i.e. people that the user was near to and the tracing app has exchanged some information with), and privacy from authorities. Public identification of diagnosed patients and access of personal information (e.g. location data) by third parties utilising it for mass surveillance purposes are potential privacy violations that were identified to arise from contact-tracing technology (Raskar et al., 2020). The questions concerning the digital divide and discrimination against people that do not own or have access to this technology are additional issues worth addressing during this discussion (Zintl & Melia, 2020).

Despite the uproar of critics, many countries have already implemented mobile tracing applications and experience a relatively widespread adoption by their citizens, e.g. Norway, Singapore, or China (Woodhams, 2020). This is where the discussion returns to the adopter's privacy paradox and the question about how big the risk trade-offs should really be to manage public health.

Digital Health Policy

One of the biggest challenges for policymakers is the rapid revolution and constant enhancement of technologies and, more importantly, the co-existence of these in people's everyday life (Castro & Atkinson, 2009). The emergence of wearables, applications, and devices that are advertised to improve one's health but are neither officially licensed medical devices nor monitored by health professionals puts enormous pressure on the importance of digital health policymakers. Accountability and responsibility in the use of automated decision-making facilitated by artificial intelligence (AI) is an example of questions that policymakers need to address in digital health. In addition, the massive collection of data used for population health benefits reveals the issue of international data governance, which is additionally complicated by the diverse (or the blunt lack of) privacy protection policies across the globe (Vayena et al., 2018). The enforcement of the General Data Protection Regulation (GDPR) in the European Union in 2018 tries to tackle this issue. In general, the GDPR is a legislation that defines rules for the processing of individuals' data, their rights to the protection of their personal data as well as the penalties in case of rule-breaking, beyond the European Union's borders (The European Parliament and the Council of the European Union, 2016b). While the GDPR focuses on companies obtaining the data subject's explicit consent, it does not exactly state how. A report by Sanchez-Rola et al. (2019) showed that websites could make opting-out very difficult for the user (which could explain the privacy paradox to an extent). Next to this, the GDPR can override rights when it comes to the health domain. Article 17 (right to erasure or "right to be forgotten") can be annulled if it is "for reasons of public interest in the area of public health" (Art. 17, paragraph 3(c)). Recital 52 allows for "[d]erogating from the prohibition on processing special categories of personal data […] where it is in the public interest to do so" (The European Parliament and the Council of the European Union, 2016a), listing public health, management of healthcare services as well as scientific or historical research as example purposes where data processing is allowed without prior consent from the data subject. Apart from the issue of data governance, the connected lack of one collective global policy, and the weaknesses of existing ones, policymakers also face the challenge of a concept called "net neutrality". Network neutrality, "an Internet that does not favor one application" (Wu, 2003, 145), is a term coined by Tim Wu in the field of broadband communication. "Paid prioritisation" – the counterpart of network neutrality – is argued to foster information gatekeeping and censorship. In the area of healthcare, the discussion of net neutrality takes a course towards impacts on health literacy and health equity, caused by access restrictions to information (Early & Bustillos, 2018).

This concludes the review of relevant literature to this thesis' purpose. The previous chapter shone a light on digital healthcare with its benefits and challenges. A special focus was set on digital twins, their characteristics, and application fields, as well as the current state of digital twins in healthcare. The following chapter will continue with creating the necessary theoretical foundation to build a framework that is used throughout the rest of the research process, including the empirical research as well as the analysis and discussion of the findings.

3 Theoretical Framework

The theoretical framework introduces theories and frameworks that are used as the basis for establishing a conceptual framework as well as the analysis for further research work. The goal of this research is to investigate how digital twins in healthcare can be implemented in a privacy friendly manner. Firstly, it is important to explain how digital twins are implemented, from the components to the set-up, as well as the interaction between these elements. For this purpose, a framework for building a functional digital twin will be introduced (chapter 3.1). To identify challenges and privacy aspects that stakeholders need to give special attention to, it is necessary to understand how applications can be designed privacy friendly. The privacy design strategies will be introduced in chapter 3.2 to explain how systems can be realized in a privacy positive setting. Next to that, chapter 3.3 reviews models of privacy addressing surveillance and datafication to further underline privacy challenges that come along with this technology. Finally, the resulting conceptual framework that combines said theories will be presented (chapter 3.4).

3.1 Conceptual Design of Digital Twins

In the bounds of the research by Tao, Sui et al. (2019), who developed a framework of digital twin-driven product design, the researchers outlined the general process of building a functional digital twin for a product. According to the authors, there are six general steps to create a fully functional digital twin for an existing physical entity. They further highlight the possibility of carrying out these steps parallel to each other in practice instead of following a strict order. The steps are as follows:

1. **Build the virtual representation of the physical product**: With technologies such as computer-aided design and 3D modelling, the virtual product is built, and is made up of the following three features:
    - **Elements:** The virtual product model consists of the geometric and physical model of the product as well as other environmental variables (e.g. user).
    - **Behaviours:** Here, the behaviour of both the products and the users are analysed as well as the interaction between the two that was created through the behaviour and modelling.

- **Rules**: These include models that are concerned with evaluating, optimising, and forecasting based on certain laws (in this application field it was the law of product operation).

2. **Process data to facilitate (design) decision-making:** Data collection happens mainly at the physical entity but can also be collected from other sources. The data are analysed, integrated, and visualised, as explained in the following:
    - Data analytics, where data are transformed into concrete information
    - Data integration, where "hidden patterns" are revealed with the help of multiple data sources
    - Data visualisation, which helps to present data in a straightforward way
    - Advanced artificial intelligence, which helps to increase the digital twin's "cognitive ability (e.g. reasoning, problem solving and knowledge representation), so that certain relatively simple recommendations can be made automatically" (p. 3941)
3. **Simulate product behaviours in the virtual environment:** In a virtual setting, simulations are used to replicate essential functions and behaviours of the physical product. Technologies such as virtual reality can help as they allow direct interaction with the virtual entity.
4. **Command the physical product to perform recommended behaviours:** Actuators, being part of the physical product, are used to adapt the physical entity based on what was recommended by the digital twin, whereas sensors simply take note of the current situation.
5. **Establish real-time, two-way, and secure connections between physical and virtual product:** Connections between the physical and the virtual product are established using network technologies such as Bluetooth, QR (quick response) code, and Wi-Fi. Using these technologies, data are sent to the cloud where the virtual twin is created and maintained. Storing the twin in the cloud has the advantage that it can be accessed by different parties regardless of their location. Lastly, the security of the connection is mentioned briefly.
6. **Collect all kinds of product-related data from different sources:** Different types of data such as product data, environment data, customer data, and interactive data are collected using sensor and IoT technology in real-time, fed into step one and, therefore, create a loop that can enable the development of further virtual products.

Figure 6 presents an overview of these six steps.

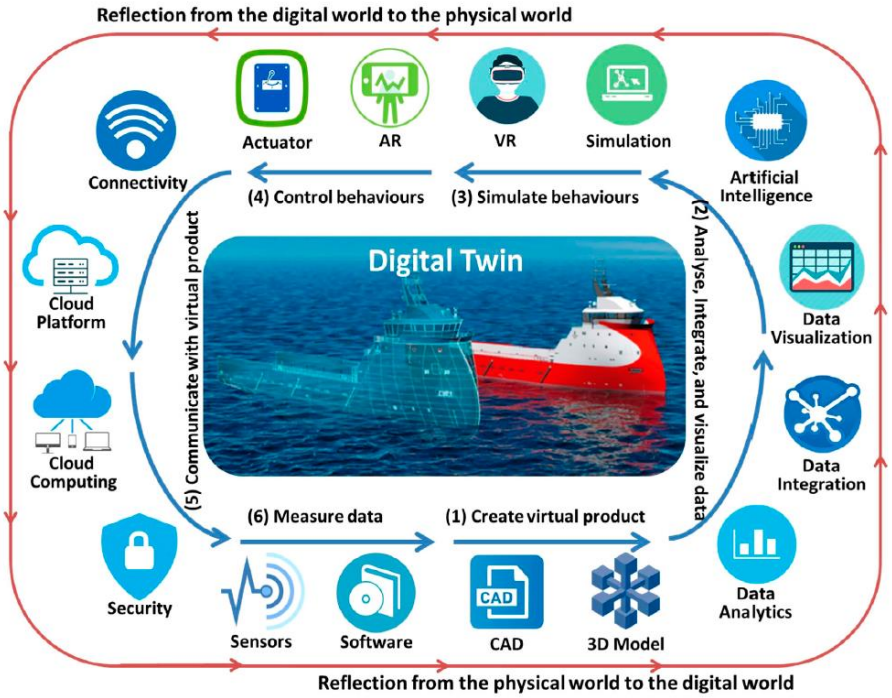

Figure 6. Overview of the six steps to create a fully functional digital twin (Tao, Sui et al., 2019, p. 3940).

In this case, the product was related to a static entity, not specifically a product in the health domain. Nevertheless, the process outlined in their paper will be used to set the basis for creating a digital twin in healthcare. Further research in this report will build upon this framework and concretise it regarding special components or processes necessary in the health domain.

When taking a look at a healthcare digital twin-like solution, Liu et al. (2019) reveal useful insights. Their paper focused on a cloud-based solution for healthcare services for the elderly using the digital twin idea (CloudDTH). They delineated a reference framework supported by cloud computing, health IoT, digital twin, and big data, similar to the technologies mentioned by Tao, Sui et al. (2019). Therefore, the CloudDTH reference framework will be explained shortly in the following. It consists of eight layers, which are visualised in Figure 7.

- **Resource layer:** The resource layer receives data healthcare resources (healthcare equipment and healthcare professional software), healthcare capability (expert knowledge, diagnostic ability, or other intellectual resources), and patient information (health record, wearable devices).
- **Perception layer:** On this layer, resources such as doctors and patients as well as healthcare devices and medicine are determined in order to manage them better.

- **Virtual resource layer:** Here, digital models of the physical objects from the resource layer are stored and form a pool of virtual models (resource, capability, and patient digital twin).
- **Middleware layer:** As the name suggests, this layer holds the middleware of services, such as service, data, knowledge, simulation, and user management middleware. The service management middleware, for instance, is responsible for maintaining the quality of services of the healthcare services, whereas the data management takes care of the storage, analysis, and transport of medical data.
- **Service layer:** Services that give the involved users functional support during the usage are located at this layer. This includes services for medical institutions, ensuring that they receive patient information and are able to send advice to their patients. Services for patients are related to real-time monitoring, crisis warning, and medical guidance. Finally, third parties (insurances, government) find their place at this layer.
- **User interface layer:** The interfaces on this layer support service provision, service request, and platform operation.
- **Application and user layers:** Users can interact with the CloudDTH platforms through their mobile phones, computers, or special medical devices. Application services include remote diagnosis and treatment, health consultation, simulation, and decision-making as well as real-time monitoring and crisis warning, to name a few.
- **Security system**: To prevent unwanted access to the health data and ensure the user's privacy, security is a top priority.
- **Standard system and specification:** A range of standards and system specifications regarding the electronic exchange of health records and information, for instance, are of importance to "guarantee the standardization of healthcare data collection, data sharing and exchange, and the application service management" (Liu et al., 2019, p. 49096).

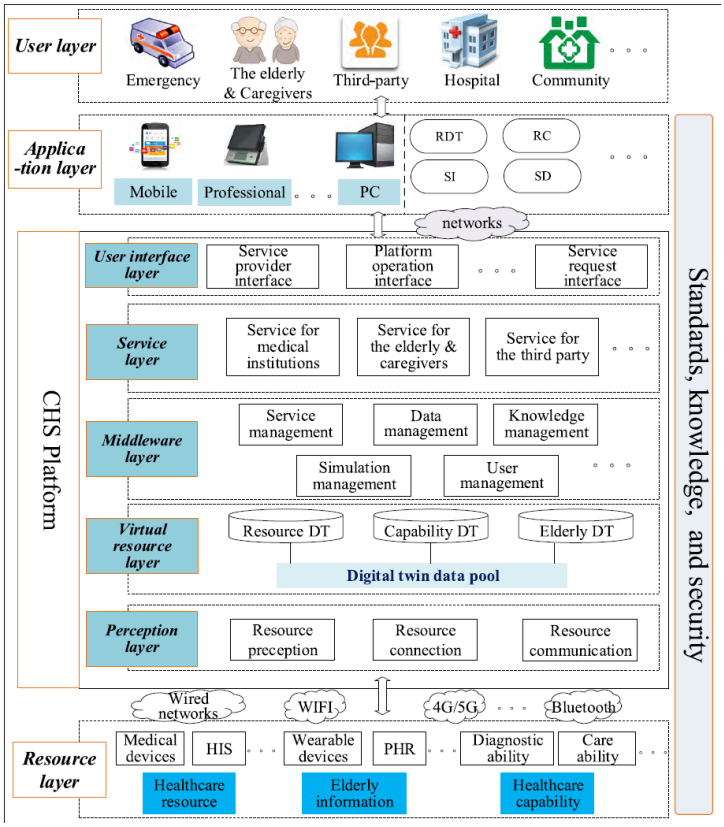

Figure 7. Reference framework of CloudDTH (Liu et al., 2019, p. 49094).

3.2 Privacy Design Strategies

As identified in the previous chapter, the collection of personal data in the healthcare spectrum reveals risks of invading the individual's privacy as these types of data are intimate and sensitive. Therefore, it is necessary to approach the development of healthcare applications by keeping the retention of people's privacy in mind.

"Privacy by design" (PbD) was developed by Cavoukian (2012) to emphasise that privacy assurance needs to be included in organisations' procedures from the beginning on, especially when they deal with personal information, and even more so when highly sensitive data are involved. As the data that are dealt with in the healthcare sector are mostly very intimate and personal, maintaining privacy in the design of health applications is of utmost importance. According to the author, the goal of privacy by design is "ensuring privacy and gaining personal control over one's information and, for organizations, gaining a sustainable competitive advantage" (Cavoukian, 2011).

To accomplish privacy by design, the information and privacy commissioner recommends seven principles that each include a range of actions. These actions are carried out by different

roles involved in the overall implementation process, such as regulators, senior management, application, and product owners, as well as software engineers and developers.

1. **Proactive not reactive; preventative not remedial:** Privacy invasion should be predicted and prevented *before* it takes place.
2. **Privacy as the default setting:** A user should not have to put a lot of effort into protecting their privacy when using a system or business practice. It is the system's responsibility to incorporate privacy measurements to keep people's data secure from the start on.
3. **Privacy embedded into design:** Privacy has to be an integral part of the system, without compromising its functionality.
4. **Full functionality - positive-sum, not zero-sum:** Privacy by design includes the interests and objectives of all parties involved without compromises, such as privacy vs. security, but choosing both.
5. **End-to-end security - full lifecycle protection:** Throughout the whole lifecycle management of information, data must be kept safe and secure – from the collection to its deletion.
6. **Visibility and transparency - keep it open:** All parties involved should be assured that the business practices or technologies are functioning in the way that was previously agreed on and are complying with the established goals. The system, including all its elements and operations, must be clear and transparent to all stakeholders involved.
7. **Respect for user privacy - keep it user-centric:** Systems should be developed in the people's interest by including strong privacy defaults, appropriate notice, and empowering user-friendly options.

These principles have been the basis for several other researchers and regulatory bodies attempting to solve the issue of incorporating privacy into health applications (Ataei, Degbelo, & Kray, 2018; Cavoukian, Fisher, Killen, & Hoffman, 2010; Colesky, Hoepman, & Hillen, 2016; Helm & Georgatos, 2014; Kotz, Avancha, & Baxi, 2009; Nordgren, 2015; O'Connor, Rowan, Lynch, & Heavin, 2017; Safavi & Shukur, 2014).

The European Commission developed a code of conduct on privacy for mobile health applications that "aims to promote trust among users of mHealth apps, and will provide a competitive advantage for those who sign up to it in the future" (European Commission, 2018). It reveals evident parallels to Cavoukian's privacy by design principles, while focusing strongly

on the mHealth environment. It was created in cooperation with industry stakeholders and addresses requirements posed by the GDPR legislation. The main guidelines for developers of mobile health applications are summarised in the following:

- **User's consent and transparent information:** Users must give free, specific, and informed consent when installing health apps before any of their data can be processed (this also applies to advertising within the app). The purpose – what personal data and why it is collected – must be easily understandable to the user. The consent should be written in concise and clear language. After having given consent, users must still have the option to opt-out and withdraw their consent (e.g. when uninstalling the app).
- **Purpose limitation and data minimisation:** Collection and processing of data are strictly limited to the data and time frame necessary for the defined purpose. If the data should be used for other purposes, it must be made anonymous or another consent must be claimed from the user. Any change of the purpose must be communicated to the user.
- **Privacy by design and privacy by default:** At every step during the implementation process, the implications that the app might have on the user's privacy must be considered and minimised as much as possible to support the user's privacy (privacy by design). Additionally, the user's privacy should be protected regarding the processing of personal data by selecting the least privacy intrusive option (privacy by default).
- **Data subject rights:** Users are required to access their stored data, including its rectification and deletion.
- **Data storage:** The period of storing data may not exceed the time frame necessary for the purpose (except if required by law). The deletion of the data must follow explicit rules which the user has to be informed about. Data may also be anonymised instead of deleted as long as they cannot be traced back to its owner.
- **Security measures:** Personal data must be protected from "accidental or unlawful destruction, loss, alteration, disclosure, access and other unlawful forms of processing" (p. 11). Therefore, the design of the app should follow secure smartphone app development and software development guidelines. To identify the specific technical and organisational measures required for the application, developers must perform a risk assessment and a privacy impact assessment, of which a template can be found in the code of conduct.

- **Data transfer to third parties:** Users have to be informed about the data transfer to a third party. In addition, a legal agreement between the developer and the third party must be established that clearly outlines the purpose for which the third party may use the data, coordinated with the information that was communicated to the user.
- **Data breach:** In case of a data breach, affected users are notified depending on the local law.
- **Apps aimed at children:** When data from minors are involved, it is necessary to design the data processing as restricted as possible and strive to obtain parental consent.

The code of conduct is currently awaiting approval from the European Data Protection Board.

While the code of conduct never actually stated to have parallels to Cavoukian's principles (but the similarities are, nevertheless, quite distinct), Colesky et al. (2016) accentuate that the PbD principles acted as a base for their research. The researchers attempted to concretise the privacy by design propositions by translating the legal requirements posed by the GDPR into privacy friendly design strategies. The aim was to "bridge the gap between data protection requirements set out in law, and system development practice" (p. 33). The authors propose eight privacy design strategies (minimize, hide, separate, abstract, inform, control, enforce, demonstrate) with respective tactics. Table 1 gives an overview of these strategies and the extent to which they are novel or related to Cavoukian's approach.

Table 1. Overview of privacy design strategies by Colesky et al. (2016) and their relation to PbD.

Privacy design strategy (Colesky et al., 2016)	**PbD counterpart** (Cavoukian, 2012)	**Concretisation and supplement**
Minimise tactics: exclude, select, strip, destroy	2. Privacy as the default setting: data minimisation (minimising the collection of personal data) and collection limitation (purpose specification) as two separate steps	The *minimise* strategy stresses the reduction of collecting personal data simultaneously with fair use of the collected data. In short, the authors "encourage the non-collection of purposeless data" (Colesky et al., 2016, p. 35).

Hide tactics: restrict, mix, obfuscate, dissociate	2. Privacy as the default setting: keep data as anonymous as possible, de-identification	*Hide* concretises Cavoukian's notion by highlighting the importance of access control strategies, reduction, or removal of correlation between pieces of data as well as obfuscating data so that it becomes unreadable to laymen
Separate tactics: distribute, isolate	-	The goal of the *separate* strategy is isolating and distributing the collection, storage, and processing of data so that correlation is more difficult.
Abstract tactics: summarise, group	-	The *abstract* strategy focuses on data abstraction and including as little detail as possible before processing data.
Inform tactics: supply, notify, explain	6. Visibility and transparency – keep it open: transparency and openness regarding policies, procedures, and controls as well as audit trails that communicate how personal data are stored, utilised, and secured	In addition to providing transparent explanations regarding the storage, collection, retention, and operations on personal data, *inform* also includes notifications regarding any changes and breaches.
Control tactics: consent, choose, update, retract	-	Data subjects should be in *control* of how their data are stored, used, and shared with other parties. Moreover, the option of updating data or having them removed must be given.

Enforce tactics: create, maintain, uphold	5. End-to-end security – full lifecycle protection: security throughout the whole lifecycle management of information	Given the scope of the paper is to design strategies that comply with the GDPR, the *enforce* strategy focuses on meeting and upholding technical as well as legal policy obligations at all times – before, during, as well as after the development stage.
Demonstrate tactics: log, audit, report	6. Visibility and transparency – keep it open: audit trails on how data are stored, secured, and obtained	By using logging, auditing, and reporting, compliance to policies and technical controls is *demonstrated.*

Like Colesky et al., O'Connor et al. (2017) also took the introduction of the GDPR as an opportunity to associate the regulation's demands with PbD requirements but focused specifically on informed consent in the health domain. Their paper suggests practices for what to consider when designing and developing IoT applications for collecting and sharing data in the field of health. The proposed practical approach combines privacy by design principles with universal usability challenges by Shneiderman (2000) (technology variety, user diversity, gaps in user knowledge) to provide electronic consent (eConsent). This should cover the GDPR's legal requirement of making the user aware of how their data are processed and used before asking for the user's consent. The practical solutions for the respective privacy by design principle as well as the usability challenges can be found in Figures 8 and 9.

Privacy by Design Principle	Description[14,31]	Proposed Solution
Proactive not Reactive	Seeks to anticipate and prevent privacy-invasive events before they happen.	Constantly updated anti-virus, anti-malware, anti-ransom ware in place. Protection of data guaranteed to user.
Privacy as the Default Setting	Seeks to deliver the maximum degree of privacy by ensuring that personal data are automatically protected.	All user data is kept private. Access by third parties is requested from user and agreement sought prior to disclosure.
Privacy Embedded into Design	Embedded into the design and architecture of the system.	The eConsent process must be clearly articulated from the start. Terms and conditions and privacy policy statements must be clearly located by the user.
Full Functionality	Seeks to accommodate all legitimate interests and objectives in a positive-sum, win-win manner, not through a dated, zero-sum approach where unnecessary trade-offs are made.	Transparency and honesty in platform design by provider. The elimination of bias in design and the promotion of trustworthiness to the end user.
Privacy as the Default Setting	The digital health citizen gives a general agreement but some restrictions in terms of the person, data and purpose are defined.	This coincides with the different levels of consent that digital health citizens can decide from. Digital health citizens should be able to set the level of privacy which best suits their needs.
End-to-End Security	Must detail the first element of information being collected, extends securely throughout the entire lifecycle of the data involved — strong security	The encryption of identifiable user details and personal health information i.e. full name, DOB, email, health condition,

Figure 8. Proposed practical approach by O'Connor et al. (2017, p. 656).

Privacy by Design Principle	Description	Proposed Solution
Visibility and Transparency	(1) To inform users about privacy risks and their implications; and (2) to be as open and transparent as possible.	eConsent must move away from text-heavy and jargon-based documentation to a more visual approach, with voice-over capabilities, to easily inform the user.
Respect for User Privacy	Requires architects and operators to keep the interests of the individual uppermost by offering such measures as strong privacy defaults, appropriate notice, and empowering user-friendly options.	This coincides with the different levels of consent that digital health citizens can decide from.
Universal Usability Principle	**Description**	**Proposed Solution**
Technology Variety	Supporting a broad range of hardware, software, and network access.	The eConsent process must adapt to the user's devices/network. This should be an automatic process which is not disruptive to the user.
User Diversity	Accommodating users with different skills, knowledge, age, gender, disabilities, disabling conditions (mobility, sunlight and noise), literacy, culture, income, and so forth.	By moving to a more visual approach with voice over capabilities the eConsent process is more accessible to a broader range of users. eConsent could be delivered across multiple modalities and multiple languages.
Gaps in User Knowledge	Bridging the gap between what users know and what they need to know.	The eConsent process must highlight in simple language (and across a variety of languages) what the terms and conditions/privacy policy document imply. Introducing a quiz on the statements could be a beneficial step in identifying what users know and understand about the terms and conditions.

Figure 9. Proposed practical approach by O'Connor et al. (2017, p. 657) (cont.).

As the frameworks by Colesky et al. and O'Connor et al. were developed with the GDPR in mind, it can be argued that they are inflexible to other settings and contexts that do not have to

comply with this regulation (e.g. outside of the EU). However, the basic principles of the GDPR regulate the collection, usage, and processing of an individual's personal data (European Commission, n.d.b) and, therefore, attempt to retain an individual's privacy. This is an important demand when dealing with especially sensitive data, which is the case in the field of healthcare. With this in mind, several points can be derived from this framework and mapped to the healthcare environment. This, together with the guidelines from the privacy by design framework as well as its extensions mentioned above, will be combined into one framework in chapter 3.4.

3.3 Surveillance, Datafication and Privacy Models

As the frameworks in 3.1 highlight, digital twins rely on data. New technologies enable the easy collection of a vast amount of health data – whether it is through telemedicine and telecare technologies that support remote examination and diagnosis, through wearable devices worn by the patients themselves, or through the patient's mobile device. As discussed in 2.3, data collection for medical care can be considered part of the surveillance society concept. Data are very important for the creation and upkeep of digital twins. A surveillance society, therefore, is inevitable for making use of its functions and benefits. To recap, the term surveillance society refers to integrating surveillance technologies into people's day-to-day life (Lyon, 2010). In a surveillance society, a vast amount of data can be collected. Throughout the lifecycle of the collection process, there are several privacy harms that individuals can be exposed to. Privacy law expert Daniel J. Solove (2006, 2007, 2008) created a taxonomy of privacy that sheds light on the potential privacy invasions arising from the collection of personal data.

Solove's Taxonomy of Privacy

There are four main activities within the collection of personal data from a person. Each activity poses a number of privacy issues. The activities, together with their possible privacy threats, are explained in the following.

1. **Information collection** is concerned with how an individual's data are gathered. Privacy harms can arise due to surveillance and interrogation.
2. **Information processing** means the storing, analysis, and manipulation of data. Information processing can lead to privacy problems arising from aggregation, identification, insecurity, secondary use, and exclusion.

3. **Information dissemination** focuses on the transfer of data to others. Privacy issues related to information dissemination are breach of confidentiality, disclosure, exposure, increased accessibility, blackmail, appropriation, and distortion.
4. **Invasion** relates to direct interferences with individuals. Here, intrusion and decisional interference are possible privacy harms.

An overview of the interconnection between the four activities is shown in Figure 10. It depicts the start of the collection process where information about the data subject is gathered by businesses, the government, or other people. It continues with the processing, storage, manipulation, and dissemination of the subject's data and ends in privacy invasion of the individual. The potential privacy risks that can happen are listed under each activity.

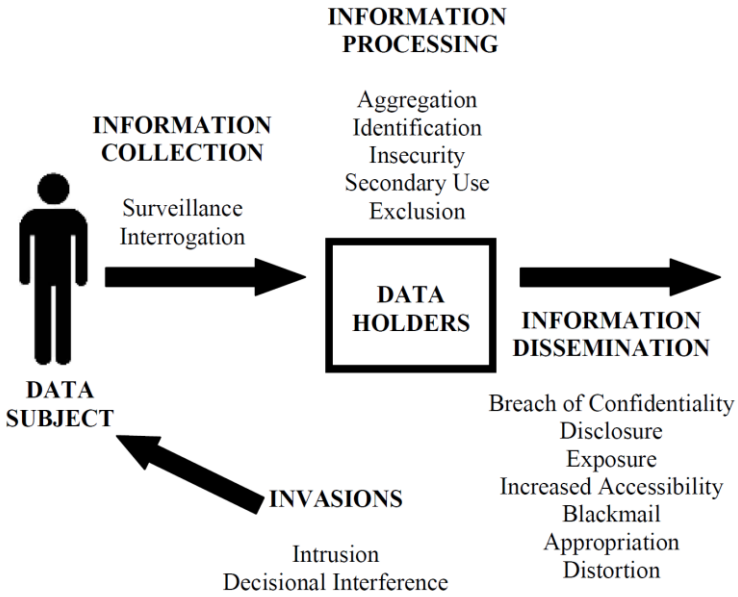

Figure 10. Solove's taxonomy of privacy (Solove, 2006, p. 490).

Solove's taxonomy accentuates potential privacy harms that can occur during the process of collecting personal data. As previously mentioned, many daily activities today are collected, transformed into data, stored, and manipulated for different purposes. This concept is called "datafication".

Datafication

Regarding datafication of personal information, Mai (2016) believes it is necessary to shift from focusing on privacy and its characteristics to privacy models that identify how privacy works. For this purpose, he introduced two privacy models – the surveillance and the capture model

by Agre (1994) (as cited in Mai, 2016) – and proposed a new model, the datafication model of informational privacy, which allows for the analysis of privacy concerns and tackles privacy challenges in the era of big data.

- **The surveillance model** focuses on the relationship between "the watchers" and "the watched". Here, the activity of being watched is non-disruptive and done secretly by intruding into an individual's private space. Watching is related to centralized orchestration and rather done by the state in a political sense. Collected data are seen as nothing more than a representation of reality. The model, however, ignores technological developments, which is why Agre defined the capture model to include the advancements in computer technology.
- **The capture model** is more concerned with the acquisition of data (epistemological notion) and the interpretation of what the data reflect (ontological notion). The model has a philosophical stand as opposed to the political view in the surveillance model. In a nutshell, "[t]he capture model focuses on the sociotechnical nature of computer technology, and on the unclear purposes of data collection" (Mai, 2016, p. 198).
- **The datafication model** concentrates on the processing and analysis of data. Where the goal of data collection is to represent simple facts and the relationship between different data, data processing and analysis are about the realities that can be constructed from the data. Therefore, the model focuses on "anonymous creation of new personal information, the reinterpretation and statistical analysis of data, and the commoditized nature of personal information" (Mai, 2016, p. 198) based on already collected data. According to Mai (2016), both the surveillance and the capture model address the protection of privacy in the collection process. In contrast, the datafication model aims its attention at privacy violations that happen due to data processing and analysis. What is more, the datafication model addresses challenges relevant to the datafication of personal information, which are delineated in the following:

    1. The production of new personal information based on big data analytics is out of the individual's control. Ownership of and rights to it are unclear in privacy frameworks.
    2. Big data analysis tends to classify and sort a big group of people, which is out of reach for individuals in terms of access and control, as the information technically never belonged to them.

3. Understanding and controlling the flow of data, the creation of big data, and statistical calculations in the digital sphere is not within the capacity of an individual. Within the proposed model, these types of actions fall under the processing and analysis of data, which is why they are related to privacy.
4. The line between private and public spheres in the digital world is blurred. Organisations use personal information to create new data, which people have not agreed to. Because Mai's model deals with the data itself and not with the time and place of its collection, this is within the interest of privacy.
5. Finally, the datafication model recognizes privacy risks when organisations process and analyse digital traces of personal information.

Each of the models has their own viewpoint and concentrate on different aspects. Together, they act as a holistic approach to reconceptualise the understanding of privacy and privacy concerns related to the collection, processing, and analysis of personal information.

Surveillance, the surveillance society, and datafication all play a big role in the scope of implementing digital twins in healthcare. Data are a crucial part of this technology. The concepts above will help shape this research process further to identify the privacy challenges that come with digital twins in healthcare. Now that the general design of digital twins as well as relevant theories focusing on surveillance, datafication, and privacy have been discussed, the next subchapter will combine the theories and ideas outlined in this chapter into one coherent conceptual framework. This will be used for the following empirical research and analysis of this research.

3.4 Conceptual Framework – Privacy Friendly Digital Twins in Healthcare

This research is guided by the frameworks, theories, and concepts above, which are integrated and linked with each other in one coherent conceptual framework. Ravitch and Riggan (2017) define a conceptual framework as "both guide and ballast for empirical research, situating specific question and strategies for exploring them within the wider universe of what is already known about a given topic or question" (p. xv). It is a helpful tool to create new knowledge and perspectives by using the experience and expertise of other researchers to match the research question at hand. Conceptual frameworks further support the choice of research methods,

evolve during the research process, and assist the researcher in reflecting critically on their own study (Ravitch & Riggan, 2017).

The conceptual framework of this thesis is depicted in Figure 11. The main block holds the main research question. To help answer it, two blocks were added that each focus on a sub research question. The following section will delineate the blocks with further detail, explain the research focus, and the outcomes.

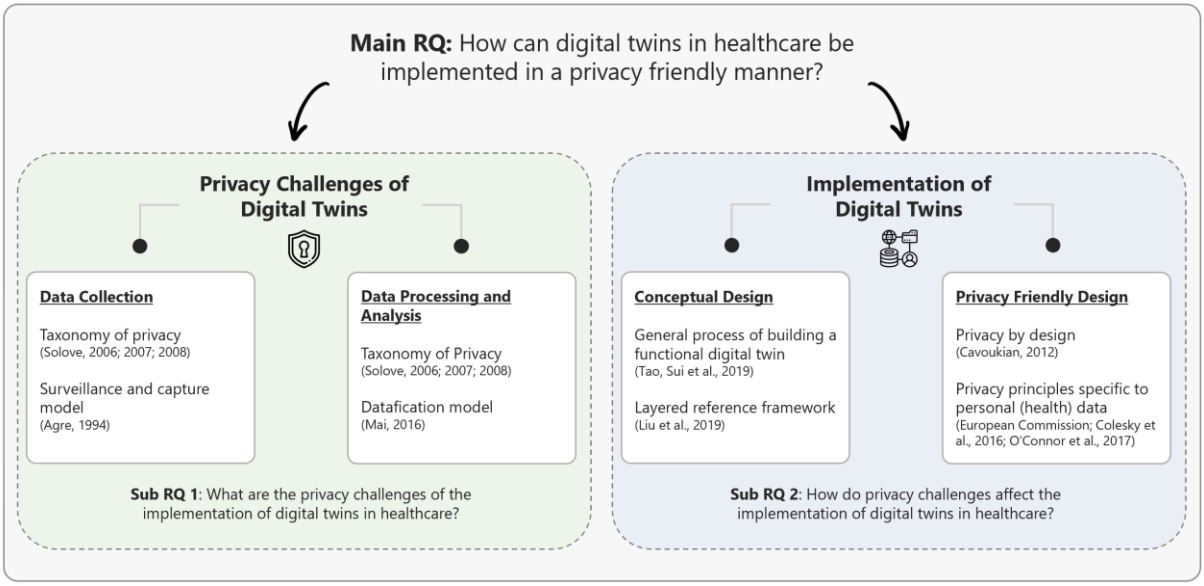

Figure 11. Conceptual framework.

3.4.1 Privacy Challenges of Digital Twins

The first block, coloured in green in Figure 11, focuses on answering the first sub research questions to identify privacy challenges of the implementation of digital twins in healthcare. This block is divided into two sections and is composed of Solove's taxonomy of privacy and the privacy models discussed in chapter 3.3. One side is concerned with data collection and privacy harms, taking Solove's taxonomy of privacy (2006, 2007, 2008) and the surveillance and capture model by Agre (1994) into account. The other side aims its attention on the privacy risks occurring when processing and analysing data by utilising the datafication model by Mai (2016). Solove's taxonomy is also part of this section, but the focus will be solely on the privacy problems of the "information processing" activity.

This block will be used as the base for the empirical research to help identify additional privacy challenges specific to digital twins in healthcare. Possible solutions to tackle these problems prior to and during the implementation process will be an outcome of this part of the conceptual framework.

3.4.2 Implementation of Digital Twins

The second block, highlighted in blue in Figure 11, concentrates on the general implementation of digital twins and how privacy affects this process. Firstly, the frameworks from chapter 3.1 that were concerned with the general conceptual and technical set-up of digital twin scenarios are used here. This included the six-step process for building a functional digital twin by Tao, Sui et al. (2019), and the eight-layer reference framework for the cloud digital twin healthcare platform by Liu et al. (2019). Secondly, privacy friendly design principles are utilised to set the base for what to consider when implementing digital twins in healthcare. The guidelines from the theories and concepts mentioned in 3.2 will be combined and used for the scope of this research. The following table (Table 2) shows an overview of the principles and the respective frameworks.

Table 2. Overview of privacy design principles and the respective framework or model.

Principle/Theory	**Privacy by Design** (Cavoukian, 2011)	**Code of Conduct** (European Commission)	**PbD and GDPR** (Colesky et al., 2016)	**PbD and GDPR in health** (O'Connor et al., 2017)
Foresee and prevent privacy invasion before it occurs	✓			✓
Privacy as the default setting, including data minimisation and purpose limitation	✓	✓	✓	✓
Privacy by design	✓	✓		✓
End-to-end security throughout the whole lifecycle (from its collection to its storage and transfer until its deletion)	✓	✓	✓	✓

Visibility and transparency regarding stakeholder communication as well as informed user consent	✓	✓	✓	✓
Designing a system in favour of the user's privacy by ensuring user control and data subject rights	✓	✓	✓	✓
Notification about any changes as well as data breaches		✓	✓	
Keeping data as anonymous as possible by de-identification, obfuscating and separating data where possible	✓		✓	

The concepts will be used to analyse the empirical data of this research regarding guidelines and privacy requirements towards digital twins. Research outcomes at this point are solutions to better ensure privacy and securely protect personal health information in digital twin healthcare applications. This part of the conceptual framework also helps to accentuate the developer's perspective to better guide them through the privacy requirements.

Together, these blocks aim to build the foundation for this research to answer the main research question: How can digital twins in healthcare be implemented in a privacy friendly manner? This conceptual framework will be extended with the results of the empirical research, guided by the theories and literature above. The next chapter will continue with a detailed explanation of the methodology of this research and outline how the empirical research and the conceptual framework are combined to reach the study's goal.

4 Methodology

The previous section discussed the theoretical foundation necessary for this research, which will be further built upon. It included a literature review of digital health, the status quo of research and applications of digital twins, as well as challenges that come along with such technology. Next to the theoretical background, a conceptual framework was introduced. This section describes the study's methodology that will help to answer the research questions. The methodical procedure, as well as the approach of the empirical investigation, will be explained. Explanations for the chosen methods are given throughout this chapter.

4.1 Research Approach

The main research question concerns the role privacy plays in the implementation of digital twins in healthcare. To answer it, the goal is to build a conceptual design/framework based on literature and primary data that emphasises the technical perspective and the privacy challenges that need to be taken into account. The philosophical worldview that underlies the research at hand is constructivism, meaning that views are constructed based on the participants' views of the studied phenomena (Creswell, 2009). There are multiple meanings to it, of which the research presents one of them (Bryman, 2012). The study follows a combination of a deductive and an inductive approach. Conclusions are drawn from pre-established assumptions from existing theory and literature (deductive), and new knowledge is added emerging from the empirical data (inductive). Therefore, a thorough review of relevant literature and research in the field of digital twins and digital healthcare, as well as the associated challenges, is conducted to create the necessary foundation. Primary data are gathered to confirm as well as build upon the existing body of literature.

The step-by-step approach of Requirements Engineering (RE) is the basis for the methodology and adapted to this research purpose. This will help identify important demands from the stakeholders involved in this process while uncovering the challenges digital twins in healthcare pose, specifically related to privacy. The theories and frameworks in the conceptual framework act as guiding theories within the steps of RE.

RE is used to develop software systems (Pohl, 1996). Dick, Hull, and Jackson (2017) define Requirements Engineering as "the subset of systems engineering concerned with discovering, developing, tracing, analyzing, qualifying, communicating and managing requirements that define the system at successive levels of abstraction" (Dick et al., 2017, p. 9). In such a system,

a group of elements, e.g. machines, software, and people, work together to reach a common goal (Dick et al., 2017). Requirements Engineering involves the stakeholders at all times and emphasises the establishment of user satisfaction by meeting the needs of the user (Pohl, 1996). A stakeholder is defined as "[a]n individual, group of people, organisation or other entity that has a direct or indirect interest (or stake) in a system" (Dick et al., 2017, p. 8). The stakeholders' requirements are so important as neglecting them might lead to system failure, or, if it is generally functional, to non-use (Dick et al., 2017).

Fricker (2015) dedicated a whole book to Requirements Engineering for digital health. Due to its suitability to this research topic, the book will be used as the basis for the methodology.

4.2 Requirements Engineering for Digital Health

Brost and Hoffmann (2015) designed a security engineering process that helps identify security requirements and privacy concerns in digital healthcare systems. As this report focuses primarily on privacy, the original four-step security engineering process was modified to address privacy instead of security requirements (see Figure 12). The main research question addresses the implementation of digital twins in healthcare that are as privacy friendly as possible. The privacy engineering process can help to find answers to this question by guiding the research process according to the individual steps. Within each step, suitable research methods are used to generate scientific results that, when put together in the end, create a valid answer to the research question. The series of stages and the actions carried out within them will be of focus in this chapter.

Figure 12. Privacy engineering process based on the security engineering process by Brost and Hoffmann (2015, p. 138).

4.2.1 Step 1: Decompose System and Determine Assets

To be able to detect privacy threats and protection mechanisms therefor, it is necessary to understand how the system works, inside and out, including the assets and data flow patterns. Assets can be tangible, e.g. computers or servers, and intangible, e.g. patient data. Architecture diagrams or use case descriptions can help reach that understanding. The literature review

section on digital healthcare, digital twins that states previous work regarding digital twins' application fields together with the frameworks by Tao, Sui et al. (2019) and Liu et al. (2019) help to comprehend the general set-up of digital twins, the components that are necessary and how these elements interact with each other.

4.2.2 Step 2: Determine Threats

For every use case, Brost and Hoffmann (2015) suggest finding a potential "misuse" case, meaning the way the respective use case can be abused. This is followed by a threat analysis "to cover as many attack angles as possible so that a complete set of security requirements can be derived" (Brost & Hoffmann, 2015, p. 139). The authors continue to introduce "attacker models", which characterise the attacker's capabilities regarding their tools and the channels they could get access to the system through, and "attack trees". Both help to get a better sense of possible scenarios and then evaluate these threats. In the bounds of this research, the evaluation will not play a big role as the focal point is identifying the privacy challenges rather than measuring the degree of the threat's severity. For this purpose, the literature review is, once again, the method to address this step. Instead of focusing on digital twins and their functionality and application fields, the focus here is set on the challenges that come along with digital twins. This extends the scope of digital twins by including the problems that arise from digital health applications in general, using chapter 2.3 as a reference.

4.2.3 Step 3: Identify Privacy Requirements

Depending on the goal, a range of privacy requirements suitable for achieving that goal should be identified. This is where the general procedure of Requirements Engineering comes into play to identify these specific requirements. The traditional outcome is a ranked order of functional as well as non-functional requirements. It is important to note that the research process of this report leans on the steps of RE. The steps are adapted and simplified where necessary, resulting in a different outcome – a critical qualitative review of the requirements and threats that allows space for interpretation.

Pohl (2010) proposes a framework for structuring the Requirements Engineering process that is divided into four blocks: system context, core requirements engineering activities, cross-sectional activities, and requirements artefacts. The core and cross-sectional activities play a big role in the development facet, i.e. the actual development process of a system (Pohl, 2010). These are said to be the common phases within the requirements engineering process (Pohl,

1996). Therefore, the focus is set on applying these activities for identifying the privacy requirements. The former refers to the documentation, elicitation, and negotiation activities that are performed to ensure that the requirements are understandably documented and that all stakeholders agree upon these. The latter include validation and management and are employed to support and secure the core activities. The following section delineates these phases more thoroughly, explains how they will be integrated into the research process and what methods will be used.

Data Collection – Elicitation

The *elicitation* stage serves to help understand and clarify the requirements of all involved stakeholders and other requirement domains (Pohl, 2010). Elicitation happens continuously throughout the process (Pohl, 1996). One of the elicitation methods is interviewing relevant interest groups (Fricker, Grau, & Zwingli, 2015; Pohl, 2010). Primary data are collected through semi-structured interviews with digital twin experts, people working in the 'medtech' sector, and scholars with a research focus on privacy issues in digital health. This opportunity is also used to add to the threats identified in step 2. The interview guide's design, the recruitment of the participants, and an introduction of the interviewees can be found in chapter 5.

Data Analysis – Negotiation, Specification and Documentation

Negotiation aims to reach an agreement among all stakeholders (Pohl, 1996). The elicited requirements are organised, refined, and, finally, approved by all parties involved (Marcelino-Jesus, Sarraipa, Agostinho, & Jardim-Goncalves, 2014). For this research, these requirements will not be reviewed and evaluated in follow up questions with the participants from the previous phase. It is of interest to gather as many requirements as possible and compare the statements from each participant with another. The insights gathered in the interviews will be analysed using Mayring's (2016) approach for qualitative text analysis. More details regarding the analysis of the interviews can be found in the findings.

In the *specification & documentation* stage, all requirements towards the system are noted down in a requirements specification (Pohl, 1996). Techniques are informal modelling, object-oriented analysis using UML (unified modelling language) diagrams, or prototyping (Fricker, Grau, & Zwingli, 2015). The negotiation results will be arranged in a model that aims to give an overview of the digital twin healthcare environment and its privacy challenges, understandable for developers, healthcare personnel, policymakers, and any other stakeholder.

In a nutshell, the goal is to balance the digital twin components and the relationships, and the privacy issues related to those.

Data Validation – Verification and Validation

"…[T]he purpose of the verification task is to check the specification according to formally defined constraints, whereas the purpose of the validation task is to certify that the specified requirements are consistent with the user/customer intentions" (Pohl, 1996, p. 14). The goal is to check if the requirements specification is valid amongst the stakeholders (Fricker, Grau, & Zwingli, 2015). The results from the expert interviews are compared and combined with the literature review to reach a consensus on the privacy challenges, and how these influence the implementation process, drawing from existing knowledge from the literature and new insights gathered through the interviews. The model from the previous step is then revised and modified so that it includes this comparison.

4.2.4 Step 4: Mitigate Threats

Once the threats are identified, measures to counteract these need to be agreed upon and installed into the system. The decision is based on a comparison between financial costs, impact on usability as well as performance and threat severity (Brost & Hoffmann, 2015). The goal of this research is to investigate how digital twins can be applied in healthcare by looking at the role of privacy and its challenges. Possible solutions to counteract privacy threats will be touched upon during the interviews; however, they do not play a central part in this research report. The outlook and concluding chapter will loosely discuss countermeasures for the identified privacy challenges and acts as a source of inspiration for further research.

4.3 Summary

In summary, this research project's methodology follows the steps within Requirements Engineering, with a special focus on digital health applications. The research questions relate to a privacy friendly implementation of digital twins. To find answers to the research questions, the steps of the security engineering process by Brost and Hoffmann (2015) are followed and adapted so that the focus is set on privacy (instead of security). Within these steps, different research methods are applied to reach the goal of the respective step. An overview of the methodology can be found in Table 3.

Table 3. Overview of the research methodology.

	Method	Focus
1: Decompose system & determine assets	Literature review	Digital healthcare, digital twins
2: Determine threats	Literature review	Challenges of digital healthcare
3: Identify privacy requirements	Requirements Engineering Modification 1. Elicitation: qualitative expert interviews 2. Negotiation, Specification & Documentation: qualitative text analysis based on Mayring; informal modelling 3. Verification & Validation: comparison of expert interview findings with theory and literature	Literature, theories from conceptual framework
4: Mitigate threats	---	Loosely discussed in outlook

5 Empirical Research

This chapter illustrates the chosen method for collecting empirical data in more detail. It explains why expert interviews were chosen for collecting empirical data and should help to understand how the theory served as the foundation for the data collection process (chapter 5.1). The acquisition of the sample and the sample itself is further delineated in chapter 5.2 before the analytical procedure and the resulting findings are covered (chapter 5.3).

5.1 Design of the Expert Interviews

The expert interviews followed a semi-structured interview approach. Semi-structured interviews give the interviewee the opportunity to answer according to their knowledge and expertise. The interview guide acts as such – it leads the conversation but leaves room for the interviewer to ask questions about the participant's response. As for the type of question, open questions were used during the interview. Open questions allow the interviewee to answer in their own terms as opposed to closed questions, where the answer is chosen from a set of fixed options. Open questions also have the advantage of generating unbiased and unexpected responses (Bryman, 2012). Two different interview guides were developed. Both can be found in Appendix B. An excerpt from the interview guide is presented in Figure 13, which shows how the research's purpose and the topic were introduced during every interview. In the beginning, each participant was asked to introduce themselves briefly.

Interview Guide

Digital Twins in Healthcare

DCLead Master in Digital Communication Leadership

Co-funded by the Erasmus+ Programme of the European Union

Group 1: Digital Twin Experts

INTRODUCTION

Hello, thank you for taking the time today to talk to me about digital twins in healthcare.

As you know, this interview is part of the empirical research for my master thesis, which aims at identifying the challenges of digital twins in healthcare, with a strong focus on privacy. The goal is to create a conceptual framework of the implementation of digital twins in the field of healthcare that can act as a guide on what needs to be taken into consideration when developing such technology.

I appreciate your effort and time. I believe that this interview will give me valuable insights for the research.

Our discussion will be recorded and archived to ensure the accuracy and validity of the research. It will not be made public, but my notes and some citations from it may be included in the master thesis. Please let me know if you would like me to anonymise the citations.

Do you have any questions before we start?

If a question is not clear, please feel free to ask me and I will clarify it.

WARM-UP, POSITION AND EXPERIENCE WITH DIGITAL TWINS

DT-Q0. Could you shortly introduce yourself and the experience you have with digital twins?

Figure 13. Interview guide excerpt.

The first guide addresses digital twin experts and focuses on the second sub research question regarding the implementation of digital twins and how it is affected by privacy issues. To design the questions, the theories from the blue block in the conceptual framework were used. A total of eleven questions are included in the first interview guide. Table 4 shows an overview of the first interview guide with its four sections, the theoretical focus of each section, and the number of questions. The interview phase "warm-up" helped generate a common understanding of the concept of a digital twin, and the expert's experience in this field is discussed. It also included a question regarding the benefits of digital twins in that respective field. After talking about the benefits, the questions in the next section deal with potential risks as well as the role that data subjects and the protection of their information play in this technology. The phase labelled "privacy friendly design" focused on identifying guidelines, principles, and technical solutions that can be incorporated to protect an individual's privacy in digital twin applications. To

conclude the interview, the participants were asked to assess the future of digital twins in healthcare.

Table 4. Theoretical foundation for the digital twins interview guide.

Interview phase	Theoretical focus	Number of questions
Warm-up	Characteristics of a digital twin (Cearley & Burke, 2018; Grieves, 2014; Tao, Cheng et al., 2018); general process of building a functional digital twin (Tao, Sui et al., 2019)	2
Privacy risks	Purpose specification (Cavoukian, 2011; Colesky et al., 2016; European Commission, n.d.a; O'Connor et al., 2017); user privacy (Cavoukian, 2011); data subject rights (European Commission, n.d.a); control strategy (Colesky et al., 2016)	4
Privacy friendly design	Privacy by default and design (Cavoukian, 2011; European Commission, n.d.a), hide and inform strategy (Colesky et al., 2016), visibility and transparency (Cavoukian, 2011; European Commission, n.d.a), security system (Liu et al., 2019)	4
Cool-down	Future of digital twins in healthcare	1

The second interview guide is tailored around the first sub research question that is concerned with the privacy challenges that come along with the implementation of digital twins in healthcare. Here, privacy experts were interviewed. A total of ten questions, divided into five sections, were designed by taking the theories from the green block in the theoretical framework as guiding concepts, as shown in Table 5. The questions in the "warm-up" phase served to find out how familiar the interview partner was in the field of digital twins. Depending on this, the subsequent interview questions were phrased in relation to digital twins or – if the interviewee possessed no prior knowledge – digital health applications in general. The interview guide then proceeded to the topic of personal health data collection. The next phase, "data processing and analysis", focused on the individual's role in big data analysis. After focusing on the individual's role in big data analysis and processing, the participants were asked to express their evaluation between the risks and the benefits by taking surveillance and concepts like the "transparent citizen" into account. Final questions revolved around the responsibility of privacy protected systems and the future of digital health applications.

Table 5. Theoretical foundation for the digital health and privacy interview guide.

Interview phase	Theoretical focus	Number of questions
Warm-up	Characteristics of a digital twin (Cearley & Burke, 2018; Grieves, 2014; Tao, Cheng et al., 2018); general process of building a functional digital twin (Tao, Sui et al., 2019)	2
Data collection	Taxonomy of privacy (Solove, 2006); capture model (Agre, 1994)	1
Data processing and analysis	Public health surveillance (Richterich, 2018); datafication model (Mai, 2016); surveillance society (Lyon, 2010); surveillance capitalism (Zuboff, 2019); privacy calculus (Majumdar & Bose, 2016)	4
Surveillance	Surveillance capitalism (Zuboff, 2019); surveillance model (Agre, 1994)	1
Cool-down	Future of digital twins in healthcare	2

5.2 Qualitative Expert Interviews

To create a framework for the implementation of digital twins in healthcare that informs about the set-up requirements and the privacy requirements arising from the potential issues, it is necessary to look at two sides. These sides were delineated in more detail in chapter 3. To address both sides of the framework, it was decided to use two sample groups of experts that have specific knowledge related to the theories and concepts within the sub-blocks of the conceptual framework. Group 1 is comprised of experts working in the conception and implementation of digital twins or doing research in this field. As the application of digital twins in healthcare is quite novel, it was decided to also include digital twin experts of other application fields. Group 2 consists of experts in the field of digital health and (data) privacy. A fixed sample size was not decided upon before the research. Firstly, purposive sampling was followed to determine as many experts as possible that are relevant and fit the requirements (Bryman, 2012). Experts that had talked about digital twins publicly were identified through Google and LinkedIn and contacted via said platform or email. Secondly, snowball sampling was employed where more experts were addressed using previously contacted experts and the researcher's network of contacts (Bryman, 2012) (e.g. university network). Suitable experts were contacted directly or through a middle person that established the connection between the

defined expert and the researcher. Every potential expert received a predefined cover letter (see Appendix A), which informed them about the researcher's academic background, the purpose of the project, and why they were thought to be a suitable candidate to contribute to the research. Additionally, information about the overall process of the interview and its approximate duration were given. Ultimately, six experts make up the final sample size of this research, of which four belong to group 1 and two to group 2. Table 6 gives an overview of the experts that took part in the interviews. It states their job description, the industry they work or research in, the date, and interview duration.

Table 6. Overview of the acquired experts.

Expert	Position	Industry	Date of Interview	Duration (min)
E1.1	Physician and university professor	Healthcare, Academia (School of Medicine)	10.04.2020	31
E1.2	Program Manager Patient Digital Twin	Healthcare	14.04.2020	32
E1.3	Engineer, Privacy Foundation Chair Health Committee	Healthcare	17.04.2020	60
E1.4	Solutions Centre of Excellence Lead	Healthcare	13.05.2020	27
E2.1	Professor	Academia (Faculty of Law)	22.04.2020	39
E2.2	Professor	Academia (Department of media and communication studies)	06.05.2020	32

After defining the sample and recruiting suitable experts, the next step was the execution of the interviews. Prior to interviewing the experts, the interview guide was tested on March 30, 2020, with a person outside of the recruited sample. This pilot study served to detect questions phrased in a way that is unclear or difficult to understand for the participant. Moreover, pre-testing helps to check if the questions are in the right order that generates a nice conversational flow. It can also improve the interviewer's ability to conduct a confident and smooth interview by gaining experience in a real setting (Bryman, 2012, p. 263–264). As a result of this pre-test, the wording

of three questions was rephrased and refined slightly in both interview guides. The overall structure and order of the questions did not change. Finally, interviews with the six experts were conducted between April 10 and May 13, 2020. The shortest interview was 27; the longest was 60 minutes. The average duration is 36,8 minutes and the median is 32.

5.3 Findings

All interviews were conducted virtually with a teleconference tool chosen by the expert to keep their efforts as low as possible. The tools used throughout the interviews were Microsoft Teams, Zoom, and Skype for Business. Each interview was recorded using the in-application recording feature, when possible, or with recording software on the researcher's device that was used during the interview. Transcriptions of the interviews based on these recordings were generated with the help of the transcription software Otter.ai (https://otter.ai/). This software was suggested by a colleague from academia and has a solid privacy policy. It does not sell or share the transcribed data with third parties and only uses data for training purposes with explicit consent. The resulting six transcripts are the basis for the summary of the findings and the analysis of the conducted empirical data. The interviews were analysed following Mayring's (2016) structured content analysis. The analysis is based on a category system with nine main categories, of which two have three subcategories and one has two subcategories. The categories were created deductively prior to the analysis based on the theoretical framework. More categories were added inductively during the coding process based on the findings. Each category has a definition, a prime example, and coding rules. Figure 14 shows an excerpt of how the categories are prepared in the codebook. The codebook with the categories and their properties can be found in Appendix C.

Codebook

Digital Twins in Healthcare

Co-funded by the Erasmus+ Programme of the European Union

Category 1: Position	
Definition	The current occupation of the interviewee.
Coding Rules	Statements are coded that relate to the person's job title and their responsibility in that position.
Prime Example	"I am a physician. I'm trained in paediatrics, anaesthesia, critical care. I have special training in something called clinical informatics. That's a very new field in the US where physicians are trained to do data driven research."

Figure 14. Codebook excerpt.

The following findings are structured based on the order that the topics occurred in the semi-structured interviews.

Knowledge about and Experience with Digital Twins

As the interviews focused on digital twins, all experts had prior knowledge of the concept. Four of them have worked or are working in healthcare and have encountered the digital twin principle in their job. E1.1 (line 69-73) is a physician specialising in data-driven research and has hands-on experience with a digital twin-like application in a hospital surgery recovery setting. Two of the experts are engineers, of whom one used to do mathematical modelling (of the human intestine, for instance) (E1.3, l. 79-82) and the other has worked in healthcare their whole career (E1.4, l. 52-56). The latter, as well as E1.2, work for a healthcare provider and have leading roles in driving the company's digital twin research (E1.2, l. 2-4; E1.4, l. 46-51). E2.1 (l. 64-65) and E2.2 (l. 122-128) had not heard of the application in the healthcare sector before the interview; however, they were familiar with digital twins in the area of smart cities. Moreover, they are academics with research experience in privacy, data protection, the impact of technology on society (E2.2, l. 73-76, 95-103), and digital data collection and information privacy law (E2.1, l. 361-362).

Definition of Digital Twins in the Healthcare Domain

In general, three of the experts that knew digital twins agreed that the definition of digital twins from chapter 2.2.1 also applies to the healthcare sector (E1.1, l. 88-96; E1.2, l. 18;). However, differences in terms of application progress (E1.1, l. 96-97) and challenges, such as acceptance (E1.1, l. 97-106; E2.2, l. 226-227), privacy (E1.1, l. 97-106), ethical and legal (E1.3, l. 67-75) issues were mentioned as discriminators to other fields. These will be summarised in more detail at a later point in this chapter. E1.3 distinguished between three fields of digital twins within healthcare, those being healthcare engineering (l. 137-140), research (l. 141), and clinical medicine (l. 165) and that each of those areas has a different approach, set-up and, challenges. When focusing on digital twins in clinical medicine, E1.3 highlighted that, at this moment, the data that is fed into a digital twin is not updated continuously but is less automatic. In their opinion, "a digital twin is a digital representation of a specific individual informed by data acquired *at the time of care*" (l. 540-541). Where it needs to be is that the digital twin has up-to-date data for it to be reliable and beneficial (E1.3, l. 307-309). The concept of a negative feedback loop is also an enabler for creating benefit: "you measure the status of the patient, you apply a treatment, a drug or something and you have the feedback loop, such that is always adapting to the patient being treated" (E1.3, l. 549-551). In the context of digital twins, this means that the data collected are fed into models. These analytical models are derived from medical science (E1.3., l. 517) and are based on the knowledge of how a certain system behaves and reacts to certain stimuli. This output then informs the diagnostic process (E1.3, l. 505-509). Knowledge is crucial to the application of digital twins in healthcare (E1.2, l. 223-226). This is why the focus is currently set on digital twins for certain organs (E1.4, l. 255-257), especially because there is much more knowledge in some areas such as cardiology. To expand on this knowledge and detect patterns for certain conditions, for instance, "AI is a huge enabler for the digital twin concept" (E1.4, l. 240-241).

E1.3 took a critical viewpoint and argued that digital twins (including artificial intelligence) are not the "magic bullet" (l. 332), meaning that there are other tools that can accomplish similar results (l. 330-332). It is also a question of whether the outcomes actually justify the effort of the means (E1.3, l. 159-164).

Benefits of Digital Twins in Healthcare

In medical research, digital twins are useful for drug development, for instance (E1.3, l. 150-155). For patient digital twins, E1.2 (l. 107) pointed out that the predictive part is essentially

what distinguishes this technology from other health applications. With the help of predictions, better personalised treatment is possible. Different treatment scenarios can be investigated in the digital twin (E1.2, l. 96-98), or the digital twin can be compared with "others that have the same characteristics along many dimensions" and how they reacted to certain treatments (E1.4, l. 120-125). Another positive effect resulting from predictions that was mentioned by the experts is that of preventive care. With a digital twin, the monitoring of a patient is facilitated, and predictions regarding their health status can be made to detect if the current course of treatment might lead to a decline in health and preventive countermeasures can be undertaken to maintain the patient's well-being (E1.1, l. 124-127; E1.2, l. 103-104). In summary, digital twins benefit the clinician by assisting them with informed decision-making: "from the data that you put into the digital twin with the intelligence that you have in the digital twin, it will reveal things hidden in the data that a clinician cannot see without this extra intelligence. So, it provides extra insights into the current status of a patient" (E1.2, l. 91-95).

Risks of Digital Twins in Healthcare

The majority of the experts identified privacy as one of the challenges connected to digital twins in healthcare (E1.2, l. 124-126; E1.3, l. 388; E1.4, l. 83; E2.1, l. 180-181; E2.2, l. 162). Here, the difficulty of defining sensitive data was referred to by the experts. It depends on the context (E1.3, l. 381-383) and is subjective (E1.4, l. 90-98). "It's the conclusions, the diagnosis and the treatment that has privacy aspects, not the details, […] those things that, to use a loose term, can be used against you" (E1.3, l. 369-373). This, in turn, can lead to inhibiting other areas such as legislation (E1.3, l. 382-383) but also innovation (E1.4, l. 98-101). Another hurdle for innovation and thereby the faster application of digital twins in healthcare, is acceptance (E1.1, l. 200-201; E1.2, l. 226-227). However, E1.1 (l.145-146) believed that this would get better over time due to the novelty of the concept. The number of benefits as opposed to the risks was highlighted as another factor that might reduce acceptance issues (E1.2, l. 233-233; E1.3, l. 390-292; E1.4, l. 124-126), together with the notion of offering incentives to consent, such as preferential provision of certain services (E1.4, l. 194-205) or financial benefits (E1.2, l. 236-239).

Data Collection

Another threat mentioned by the experts was that of inconsistent purpose limitation. First of all, the experts pointed out that it is difficult to make sure that the collected data are used for the right purposes (E1.4, l. 221-223). Either the rationale is not comprehensible (E2.1, l. 209-215;

E2.2, l. 213-220) or the networks of data collection are big, meaning that if the data are collected through a commercial wearable developed by a non-medical provider, this company's intentions become of importance as well (E2.1, l. 287-292). In terms of the devices used, E2.1 accentuated the importance of the quality of the sensor and whether the collected data are suitable for the desired goal (E2.1, l. 189-193).

Data Processing, Analysis and Surveillance

The emergence of business models that use the data for other reasons than well-being (E2.1, l. 228-230) and aim to model and shape behaviours by nudging towards a specific behaviour (E2.1, l. 230-236, E2.2, l. 246-252) was considered another potential risk for the application of patient digital twins. E2.1 (l. 160-174) is of the opinion that companies might not stop at monitoring the health condition but will extend the surveillance to tracking an individual's lifestyle since a lot of health conditions arise from the individual's lifestyle choices, such as diet and exercise habits. This would then lead to "physical intrusion into privacy that is done through a digital context" (E2.1, l. 180-181). E2.2 (l. 251-252) stated that "predicting human behaviour can be very beneficial profit-wise, but the best prediction is when you decide yourself how the human will behave". This gives rise to the question of whether companies shape their customer's lives for the individual's benefit or the company's own benefit (E2.1, l. 235-328), and how this can be a threat to an individual's autonomy (E2.1, l. 236).

Finally, the risk of the unknown effect of digital twins on the societal dynamics between patients and doctors and, in the long run, the medical evaluation was mentioned by E2.1 (l. 251-255).

Mitigating Risks – Privacy Friendly Design

When discussing how to reduce the risks, two experts pointed out that the digital twin itself and the environment it is functioning in need to be secure to prevent unauthorised access (E1.2, l. 126-128; E1.4, l. 138-140). E2.2 stressed on building privacy protection measures directly into the system in a transparent way. Not only should this be included from the start, but the compliance with these measures should be checked. On the one hand, this can take place directly in the technology, for example with the help of auditing algorithms that ensure the privacy rules are followed (E2.2, l. 185-186). On the other hand, there should be an independent organisation that has insight into the workings of the systems and check how these affect and benefit society (E2.2, l. 337-349). Involving the individual of the digital twin was also considered as an important topic amongst the experts. People should have control over their digital twin, its data, and who has access (E1.2, l. 151-155; E1.4, l. 187-189). "I think the main

guiding principle is that in the end, the digital twin belongs to the person" (E1.2, l. 151). This goes hand-in-hand with the concept of informed consent, where the individual grants access and use of the data to a certain purpose (E1.4, l. 193-194). A prerequisite to this, however, is a clear purpose statement regarding the collection and analysis of the data by the digital twin providers (E2.2, l. 242-245) as well as better education of people about data literacy and data agency (E2.2, l. 350-355).

To be prepared for possible data breaches, anonymisation, and encryption of the data are of utmost importance (E1.4, l. 140-145). "Data breach is going to happen. It's not the question of *whether* it's going to happen, it's *when* it's going to happen. But as long as you have adequate de-identification of the data, there's no problem" (E1.1, l. 190-192). Using blockchain to "de-identify" the data was mentioned as an idea to counteract this issue (E1.1, l. 193-196; E1.2, l. 168-169). E1.3 (l. 397-405), however, criticised this approach since full anonymisation is very difficult without losing any value of the data. When handling data breaches, it was suggested that the system should have the possibility to trace the intrusion and reveal the data's audit trail (E1.4, l. 145-148) along with informing people about the incident and taking responsibility (E1.4, l. 155-158).

In general, the protection of the digital twin data should be handled no differently than the protection of data in medical records, according to E1.1 (l. 154-161), for instance through rules set by the HIPAA (Health Insurance Portability and Accountability Act) in the US context (E1.1, l. 177-178).

Future of Digital Twins in Healthcare

All experts apart from E2.1, with whom this topic was not discussed due to time constraints, had a positive attitude regarding the future of digital twins and believed that this technology would play an essential role in the long run in healthcare: "Yes, digital twins, in my definition, are critical for the future of clinical medicine" (E1.3, l. 581-584). Digital twins and their data are expected to become part of the electronic medical record (E1.1, l. 245-246) as "the new norm" (E1.1, l. 108). Such progress is believed to be inevitable and necessary, which underlines the importance of recognising the concerns connected to this technology (E2.2, l. 328-332). In the current development, the focus is and will continue to be set on creating digital twins of particular organs (E1.2, l. 203-206; E1.4, 254-257), but the vision for the usage of digital twins in healthcare, in the long run, is to have one for the whole body (E1.2, l. 180-181).

6 Analysis and Discussion

After presenting the findings of the qualitative expert interviews in the previous chapter, the following section will analyse these in relation to the delineated literature and theories in this thesis, complemented with a critical discussion. Based on this, a framework for implementing digital twins in healthcare in a privacy friendly manner will be illustrated. Throughout this chapter, the goal is to provide answers to the main research question and the supporting ones.

Firstly, the focus is set on providing answers to the two sub research questions by comparing the findings from chapter 5.3 with the literature from chapter 2 and the conceptual framework from chapter 3.4. To recap, the first sub research question is concerned with the privacy challenges based on Solove's taxonomy, the surveillance and capture model by Agre (1994) as well as the datafication model by Mai (2016). The second sub research question aims at how these privacy challenges affect the implementation of digital twins based on literature by Tao, Sui et al. (2019) and Liu et al. (2019) as well as privacy design principles.

6.1 Sub RQ 1: What are the Privacy Challenges of the Implementation of Digital twins in Healthcare?

Digital twins facilitate the justification of privacy infringement.

When discussing the challenges with the digital twin experts, it was revealed that they were aware of the privacy challenges that come along with the digital twin technology in healthcare. These challenges were said to possibly complicate things such as innovation and adoption but could be overcome by highlighting the benefits of the digital twin application. The literature review and the findings of the expert interviews certainly indicate that digital twins come with many beneficial aspects to patient-centred healthcare. What the experts touched upon is rooted in the risk-benefit-analysis notion. Because the benefits are related to health, the emphasis on the risks is loosened (as discussed by Rahman, 2019), which could potentially lead to people turning a blind eye to the possible privacy risks. This would override the "traditional" phenomenon of the privacy paradox and trigger a shift towards the "new privacy paradox". What was described by Adorjan and Ricciardelli (2019) in the social media context would thereby manifest itself outside the bounds of data from social networking sites and extend to the health context as well. As some of the experts mentioned, the purpose of data collection is often ambiguous and unclear. In the context of digital twins, together with a privacy neglecting mindset, health benefits such as public health surveillance can be easily used as an excuse to

intrude into a person's privacy and act as a doorway for governments to employ mass surveillance. Ultimately, this is a danger to the right to privacy.

Digital twins have the potential to threaten an individual's autonomy.

The issue of jeopardising people's autonomy was briefly brought to attention during the expert interviews. This threat sits at several steps during the personal data collection process based on Solove. In the first step, the information collection, an individual's freedom of choice can be limited if incentives, as suggested by E1.4, are utilised to lure people into using digital twins and disclosing their data. Offering benefits that lead to a disadvantage for those who are not willing to use digital twins, such as preferential or exclusive treatment, might lead to discrimination and exclusion. This is already an acknowledged privacy challenge in the field of digital health applications as outlined in the debate about the corona tracking apps (Zintl & Melia, 2020), for instance. As for information processing, the second step in Solove's taxonomy, the analysis and manipulation of data can be utilised to not only predict but shape an individual's behaviour. What was discussed by Zuboff (2019) aligns with what the interviewed experts said. In the context of digital twins, the further inclusion of surveillance technologies into everyday life is facilitated, pushing towards an establishment of a surveillance society whose behaviour can be moulded based on big data analysis. In Solove's invasion step in the taxonomy, intrusion is classified as one of the privacy harms. As soon as the digital twin concept goes beyond the mere monitoring of an individual's health and is extended to other areas of the person's lifestyle that are not necessarily related to health, the possibilities to shape behaviour increase. In combination with the smart home concept, digital twins can take on several different roles outside the purpose of health, as accentuated by E2.1, when employing nudging frameworks as shown in Yeung's (2017) work. This is where physical intrusion as potential privacy harm can occur.

Digital twins can be an enabler for new business models that support the commodification of health data.

As mentioned in the literature review section, the commodification of personal data has been a widely discussed topic and is certainly not limited to health information. Digital twins, however, could amplify this since personal data are continuously collected from the individual at potentially any place. Real-time data open new possibilities for businesses to make a profit from this information. Advertising is a very popular example of commodifying personal information. Here, ads can be tailored even better and shown to the individual at the right moment when

information about the individual's mood or health status is given in real-time. E2.1 picked up on the commodification aspect and raised a question regarding the company's intentions that is merchandising digital twins for health improvement and whether they see their goal in creating benefit primarily for themselves or their customers. Moreover, moulding of behaviours, as already mentioned in the previous statement, does not only threaten an individual's autonomy. It can lead to new business models that specifically see their goal in making revenue from it by giving false health-related pretences. According to Solove, information dissemination, which is the transfer of data to others, can lead to privacy risks such as distortion and appropriation. In the digital twin scenario, the components and transfer of data in its network are enormous and might not be fully comprehensible to the individual, as underlined by E2.1. This insight can also be found in Mai's (2016) datafication model regarding the data flow that is taking place to create big data analysis and calculations. It also states that new data are produced based on previously collected personal information, which is closely connected to Solove's threat of distortion. This, as well, is a possibility for the emergence of business models that make it their primary goal to make revenue with the excuse of providing health-improving applications.

It is uncertain what effect digital twins might have on the relationship between patient and healthcare provider.

The surveillance model by Agre (1994) views the relationship between those being surveilled and those employing the surveillance as one-sided. The watched person is not aware that someone is watching them. As for digital twins, the patient would (and should) be aware that clinicians and other health personnel view their twin data. The possible change in the relationship between the patient and the healthcare provider might not seem to be a privacy challenge at first. It is important to note, however, that the information in the digital twin could be much more personal than that of an electronic medical record when it includes intimate variables such as diet and exercise, for instance, as discussed by the experts. This changes the meaning and accentuates the importance of confidentiality and privacy protection from the healthcare provider's side.

6.2 Sub RQ 2: How do Privacy Challenges affect the Implementation of Digital Twins in Healthcare?

The new privacy paradox emphasises the digital twin developer's responsibility for embedding privacy strategies directly into the system from the beginning on.

Because of the threat of people not caring about their privacy (new privacy paradox) and thereby facilitating the justification of surveillance, including built-in privacy design strategies, as discussed in 3.2, gain importance. The principles highlight that the system is responsible for protecting the user's privacy from the start on and not relying on the user to act privacy proactively (privacy by default, privacy by design, user-centric). This was supported by the experts for the field of digital twins as well. E2.2 went a step further and appealed to extend this notion by including technical auditing solutions that check if the user's privacy is preserved. The threat of society's privacy neglect and its abuse for mass surveillance affect the implementation so that privacy protection must be embedded into the system early on. In a nutshell, it is the developer's responsibility to implement a secure system, maintaining the user's privacy as best as possible from the start on and constantly review if the system complies with privacy protecting measures.

Questions of ownership influence the implementation of digital twins in healthcare.

The threats of unnecessary invasion and mass surveillance are not the only challenges influencing the implementation process in pushing towards privacy friendly design strategies. The potential manifestation of a surveillance society in which behaviours are shaped by for-profit companies that threatens an individual's autonomy calls for a reflection of possession of the digital twin during the implementation process. Ownership of the digital twin was understood as an important topic during the expert interviews. Access permission and informed consent were topics that were mentioned in this context. This has similarities with what is stated in the principles in chapter 3.2 about giving users control as well as data subject rights when designing a system that maintains an individual's privacy. When combining the experts' statements with the propositions from the theories, the implementation of digital twins in healthcare must deal with the question of who the owner of the digital twin is and how to handle access control. The changing relationship dynamic between the patient and the healthcare provider also plays a role in this. It forces developers to think about the topic of digital twin ownership and rights.

Privacy challenges force digital twin developers to take notice of the set-up and network around a digital twin in healthcare.

Even though the experts mostly agreed that a digital twin set-up is characterised by the basic elements of the physical entity, the virtual entity, their connection, and value, the importance of the network around a digital twin was briefly touched upon as well. Here, end-to-end security

is an aspect that is referred to in theory (PbD) as well as in the expert interview findings. The experts highlighted that the digital twin itself and the components surrounding it should be secure to prevent unauthorised parties from accessing the data. The previously discussed privacy challenge related to the establishment of new business models briefly introduced the difficulty of understanding and retracing the data flow within the digital twin setting. This is not only a challenge for users but for those wanting to implement a part within this environment as well, especially when taking the end-to-end security argument into account. To implement digital twins in a privacy friendly manner, developers would have to be aware of the individual surrounding components and what their intentions are. Therefore, they cannot solely focus on their own system but must consider the whole setting. This, in turn, highlights the importance of visibility and transparency from each stakeholder.

Digital twins in healthcare pose similar privacy challenges as digital health applications – the difference is the characteristic of scalability.

The potential challenges of digital twins in healthcare discussed by the experts coincide with the challenges that common digital health applications pose, as delineated in the literature and theories used within this research. The difference that can be seen between common digital health applications and the digital twin scenario, is scalability. Digital twins promise a much more personal and faster delivery of care, including real-time data of not only factual data like vital signs but also diagnoses and more accurate treatment options. The possibilities for digital twins to be connected with each other in a huge network of data transfers, connections, and trails with fields in as well as outside of the healthcare sector are enormous. The bigger this network of digital twins and its connection to other areas, the bigger the opportunities for privacy intrusion and abuse. Digital twins do not reduce the privacy challenges, but retain, if not complicate the existing challenges. This accentuates the importance of a reflection on how to implement digital twins to reduce the risks and make abuse as difficult as possible. This leads to answering the main research question of this research project.

6.3 Main RQ: How can Digital Twins in Healthcare Be Implemented in a Privacy Friendly Manner?

The previous section discussed the privacy challenges and how those influence the implementation of digital twins in healthcare. When developing digital twins, the guidelines and frameworks by Tao, Sui et al. (2019) and Liu et al. (2019) that were introduced in chapter

3.1 are helpful tools to understand the technical set-up of digital twins. The following section will discuss how affected steps and processes can be altered to ensure a privacy friendly digital twin implementation in a healthcare context, taking the identified privacy challenges as well as privacy design strategies and solutions into account. The outcome of this is an overview of a digital twins scenario in healthcare that includes the privacy challenges and a brief explanation of what needs to be taken into account for a privacy friendly implementation.

A patient digital twin should belong to its physical counterpart.

The privacy friendliest answer to the question of ownership in the context of digital twins in healthcare is the patient or, more precisely, the person whom the digital twin is modelled after and whose data are fed into that virtual representation. After all, it is their personal and intimate information that is stored in the digital twin. This answer may seem simple in terms of protecting an individual's privacy but reveals many new questions and difficulties when it comes to the technical implementation. Here, ownership does not mean that the digital twin and its environment is developed and in possession of the patient. Firstly, it is impossible that everyone has the expertise and knowledge to create and maintain their twin, let alone the storage capacity to store the data around the digital twin in their own server space. This would also complicate communication with healthcare providers and other twins. What is important is that the handling, once personal data are fed into it, is in control of the patient. This includes the collection of data, the transfer to any other entity, and the processing of the data once transferred. It is the developers' task to keep any of these steps to the absolute minimum without harming the individual's privacy. This is where clear purpose limitation and transparency play a big role. The individual must understand what their data are being used for and be given a free choice to decide whether they want to disclose their information.

Involved parties in the digital twin scenario should collectively establish cooperative responsibility.

One of the experts mentioned that the company they work for follows specific guidelines, but the company itself designed these. While some companies set up their own privacy guidelines, this is not enough in the digital twin domain. As discussed, the network around a patient digital twin is highly scalable and includes many different stakeholders. A possible solution to ensuring to involve everyone is that of cooperative responsibility, an idea shaped by E2.2. First, all involved parties need to be identified carefully. These include obvious parties such as health equipment engineers, health service providers, and health institutions, but other important

players such as patients' organisations or informal caregivers must not be forgotten. By including each stakeholder's needs and point of view, values and issues that need to be tackled can be determined. By employing this course of action, different interests are represented democratically, and relevant and beneficial topics to society are of focus. The next step is to decide upon different levels of responsibility among the stakeholders. The concept of cooperative responsibility ensures that all parties have a say in the values and challenges that need to be addressed to develop effective principles and solutions to maintain the benefits and manage the challenges. Together with visible and transparent communication between the stakeholders as well as a clear purpose statement as already recognised by previous privacy design strategies, this can be a solution to implement digital twins in a privacy friendly manner and audit any privacy wrongdoings on many levels.

Universal privacy regulations need to be set in place to ensure privacy protection of the patient digital twin on a global scale.

There have been attempts to develop comprehensive principles and guidelines, for instance Privacy by Design but also ones that apply to a territorial scope and specific fields such as the EU guidelines for mobile health applications or the GDPR. A combination of cooperative responsibility and privacy design strategies is a good way to implement privacy friendly applications and systems. The set-up around digital twins in healthcare does not involve only one provider but is potentially made up of several (wearable provider, infrastructure, healthcare providers etc.) that are based across the globe. What must follow a concept like cooperative responsibility is legal consequences. A challenge here, however, are the differences in or even lack of digital health policies. Because digital twins have the potential to function in a big world-wide network of other digital twins and stakeholders, the necessity for basic universal stringent regulations for the use of commercial patient digital twins is of utmost importance prior to their adoption.

The right to privacy must not lose importance.

Finally, what can be taken from the theoretical foundation, as well as the expert findings, is that the understanding of privacy is subjective and fluid. It changes depending on the context, environment, a person's needs and situation, as well as their attitude and prior experience with digital health. With the issues of fast technological change and rapidly growing surveillance societies, it is difficult for health legislators to define what privacy should mean in the context of health and how to best protect it. Technology should not inhibit a person's privacy – it should

empower them to make use of its benefits and live a better life. E2.2 called this empowerment by design. When implementing digital twins in healthcare, the right to privacy must not be undermined, despite possible lack or inconsistency of regulatory measures. After all, it is the shield against loss of control and privacy invasions from governmental institutions, state surveillance, and big corporations. E2.1 made the analogy of privacy being a protector against power. Privacy friendly patient digital twins should not attack this shield but work with it.

Visualisation of the Digital Twin Healthcare Universe (and How to Keep it Privacy Friendly)

After discussing the findings in relation to the theories and literature critically, the existing and the acquired empirical data in this thesis are put together and visualised. The frameworks used for the implementation of digital twins were adjusted to fit the healthcare sector and extended by combining the literature on digital health and digital twins with the key findings of the empirical data. An overview of the digital twin universe in healthcare can be found in Figure 15. It shows the main components in the physical and virtual space on an abstract level and which layer of the framework by Liu et al. (2019) they refer to. Additionally, the challenges and privacy friendly design strategies to counteract the issues were added to in Figure 16.

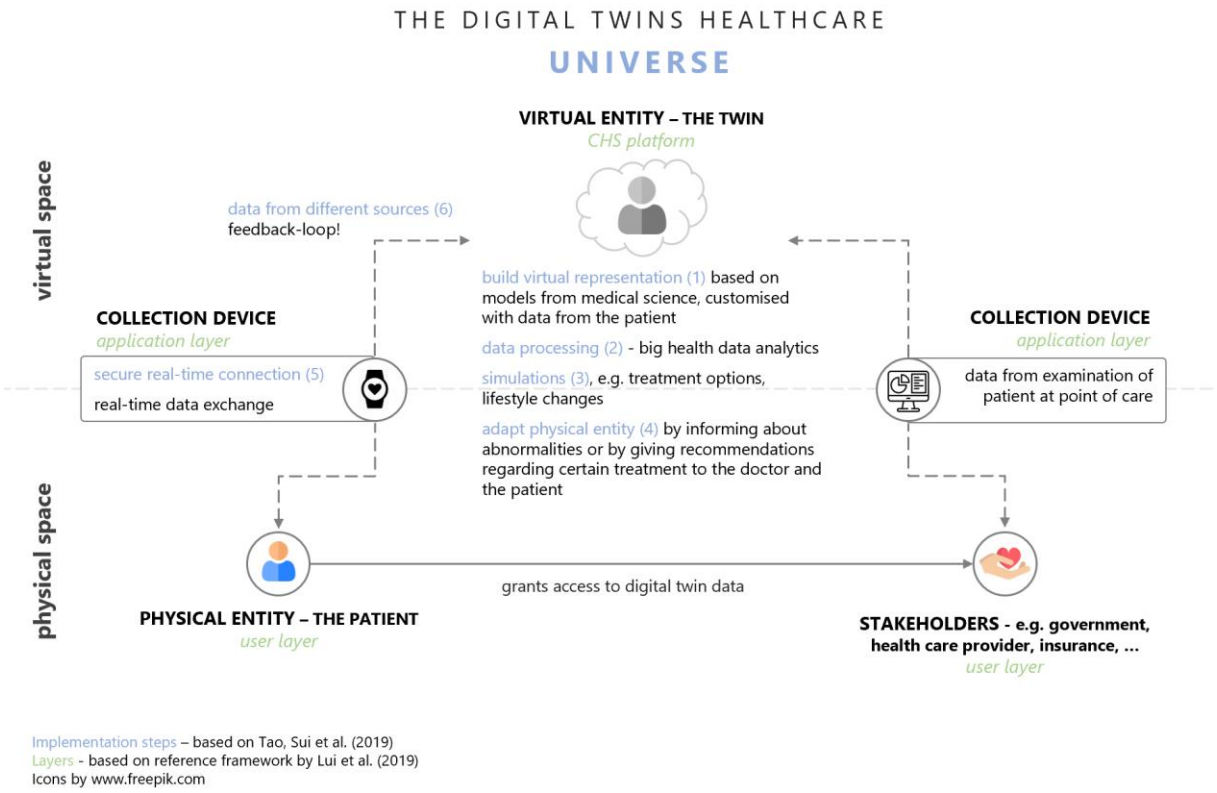

Figure 15. Overview of the digital twins healthcare universe.

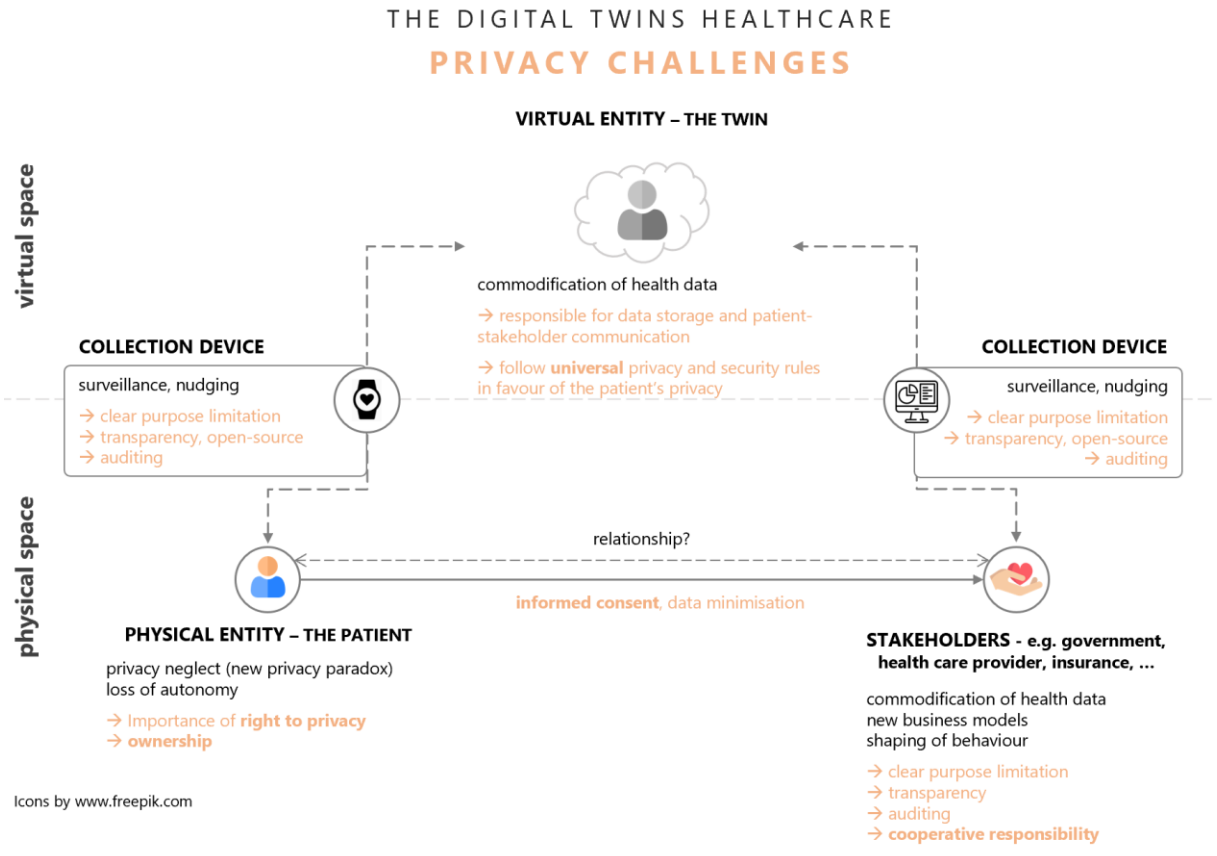

Figure 16. Overview of the digital twins healthcare privacy challenges and mitigation approaches.

7 Conclusion

The last chapter will summarise the research process, the results, and how they contribute to the existing body of knowledge. Next, the limitations of this study will be assessed to critically review the validity of this research and offer improvement points for future studies. The chapter concludes with a brief outlook of what can be expected of the future of digital twins in healthcare.

7.1 Summary

This master thesis aimed to investigate the implementation of digital twins in healthcare in a privacy friendly manner. The main research question was supported by two sub research questions that, on the one side, addressed the privacy challenges that come along with digital twins in healthcare and, on the other side, dealt with how these issues affect the implementation. The methodology of Requirements Engineering for digital health was followed, which consists of four steps. The first two steps covered the understanding of the concept, the benefits, and the challenges of digital twins, which was accomplished through a thorough literature review. The third step was concerned with identifying requirements for a digital twins set-up in healthcare. Here, semi-structured interviews were utilised to build upon existing knowledge and help answer the research questions. The interview guide was designed using digital twins frameworks by Tao, Sui et al. (2019) and Liu et al. (2019) and various privacy design strategies based on Cavoukian's (2011) Privacy by Design principles. Six expert interviews were conducted with experts from the field of digital twins in healthcare and academics with expertise in privacy studies. The last and fourth step of requirements engineering for digital health concerns the mitigation of the challenges that pose a threat to the system. These were not a focal point in the research project but will be discussed in an outlook on future steps for digital twins in healthcare.

Resulting from the analysis and discussion, the key learning points that can be drawn from this research project are summarised in the following to provide answers to the research questions that were of focus in this thesis:

Sub RQ 1: What are the privacy challenges of the implementation of digital twins in healthcare?

Because of the many advantages that digital twins can bring to healthcare, a shift in people's mindsets towards privacy indifference is possible, as seen in other areas such as social media.

This, together with the beneficial aspects of this technology, fosters the justification of privacy infringement. Surveillance and constant real-time monitoring of the daily life are not only an enabler for new business models that have the intent to monetise a person's health but also threaten the individual's autonomy. Together with clear purpose limitation and transparency on the side of the businesses, consumers need to acquire a certain degree of digital literacy to understand what is happening to their data. Moreover, the involved parties (healthcare providers, insurances, relatives etc.) must be aware of the transformation that digital twins can forward concerning the patient-stakeholder relationship.

Sub RQ2: How do privacy issues affect the implementation of digital twins in healthcare?

Developers have long had more than just the responsibility of ensuring the functionality of a system. Embedding privacy protecting measures directly into the systems has been recognised as an important aspect of software development but gains even higher importance when it comes to sensitive data concerning one's health. It is not solely developers that need to think of embedding privacy design strategies early on into the system. Each stakeholder involved in this platform has to realise that they do not control the patient and their data. They are service providers that offer a certain benefit to the customer and use the data made available to them by the patient for the benefit of that patient.

Main RQ: How can digital twins be implemented (in healthcare) in a privacy friendly manner?

The complexity and the virtual infinite universe surrounding digital twins in healthcare, however, make it difficult to keep track of the components and interconnections, which is why universal privacy regulations need to be established collectively to ensure cooperative responsibility and global privacy protection equality. Digital twins in healthcare have the potential to expand fast and grow into a big network that goes beyond preventive and predictive healthcare. Due to this scalability, the role of the digital twin owner – the physical twin, so to speak – is a highly important one. The owner stands in the centre of this universe with their digital twin provider acting as a mediator between their needs and the surrounding stakeholders. Implementing the set-up in this way mitigates the threats and ensures that access control over their data remains with the patient. Finally, policymakers must not downplay the significance of privacy protection and enforce rules in favour of the digital twin owner.

The outcome of this master thesis is a conceptual framework of the implementation of digital twins in the field of healthcare that acts as a guide on what to look out for when developing

such technology. The findings can contribute to further research on the impact of privacy on the implementation process of digital twins in healthcare.

7.2 Limitations

Despite the depicted values that this thesis adds to research and the applicability of digital twins in healthcare, there are limitations that this research is confronted with, which will be evaluated critically in the following.

First, the empirical data exhibits some limitations. When looking at the empirical data collection, it is said that purposive sampling does not allow to generalise to a population (Bryman, 2012). The experts were recruited based on the researcher's interpretation of their fit to the research goal. This, however, was necessary as the interest in digital twins in healthcare is growing, but the number of companies and research institutions that are applying this concept practically is limited. Thus, the number of the potential sample is restricted due to the specificity of the subject. In addition, developers of such applications declined the interview invitations because of the company's policy. Another factor that played into the relatively low number of experts is that the time of the research for this thesis happened during the peak of the COVID-19 outbreak. Even though it can be argued that the number of experts is not sufficient to reveal significant insights, the recruited sample is a pool of knowledgeable experts that account for relevant experience appropriate to this research goal and the findings show many overlaps with the statements amongst the experts. To overcome shortcomings in the technical and medical area, an extensive literature review was included to provide the contextual environment. Another interesting viewpoint would have been beyond the business and expert side by including potential patients of digital twins and what they require as well as fear from such technology. However, the lack of applicability of commercial digital twins and time constraints did not allow for such investigation.

While it was attempted to remain as objective as possible with this information, the possibility that the author's own experience and views could have influenced the interview and the analysis of the findings cannot be ruled out completely. It is also important to note that all interviews were held online because the pandemic did not allow for face-to-face interviews, and technical issues such as inaudible audio sections occurred. Next to this, some questions were not asked specifically during the interview due to time issues or because they were answered while replying to another question.

Beyond that, the results of this research are limited to the point that they do not provide stakeholders of the digital twin universe such as researchers, developers, or policymakers with a detailed manual that holds instructions on building build a digital twin in healthcare. It is rather an informative research report with tools and recommendations on what to consider when building a digital twin in healthcare. It provides thoughts on a privacy friendly implementation and should serve as inspiration for further research and training in this field. This thesis scratched the surface of a fast-evolving technology that has the potential to manifest itself within our everyday life; hence, more detailed research is certainly necessary in the future.

7.3 Outlook

This research took a critical standpoint on the applicability of digital twins in healthcare. While digital twins have the potential to revolutionise healthcare by offering a big amount of benefits, the concept can certainly be abused if fallen into the wrong hands. The question of cost-effectiveness, not only economically but also in terms of privacy and acceptance, will become an important one in the realm of digital twins in healthcare. The application of digital twins in healthcare seems inevitable, but is it necessary to create twins for the whole body in the long run? The conceptualisation of the digital twin universe on a high-fidelity level with all its actuators, connections, policies, stakeholders, and use cases must be mapped out. Government bodies, health institutions, policymakers and developers must think not only one but several steps ahead and approach the issues in a timely manner before they are overtaken once again by technological advancements. Coming back to Saracco (2020b) and his thought experiment of using digital twins in epidemics control, he sees the main challenge as follows: “The challenge is to match societal benefits/needs with personal privacy, to be able to create an awareness that does not result in fear. All in all, it is a matter of creating trust.” The foundation of this trust must not be built with rash promises and false pretences, but with a stable and well-thought-through plan that follows approaches similar to cooperative responsibility and puts the patient, their well-being and their privacy first – always.

References

Adorjan, M., & Ricciardelli, R. (2019). A new privacy paradox? Youth agentic practices of privacy management despite "nothing to hide" online. *Canadian Review of Sociology*, *56*(1), 8–29. https://doi.org/10.1111/cars.12227

Agre, P. E. (1994). Surveillance and capture: Two models of privacy. *The Information Society*, *10*(2), 101–127. https://doi.org/10.1080/01972243.1994.9960162

Ameri, F., & Sabbagh, R. (2016). Digital factories for capability modeling and visualization. In I. Nääs, O. Vendrametto, J. Mendes Reis, R. F. Gonçalves, M. T. Silva, G. von Cieminski, & D. Kiritsis (Eds.), *IFIP Advances in Information and Communication Technology. Advances in Production Management Systems. Initiatives for a Sustainable World* (Vol. 488, pp. 69–78). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-51133-7_9

Ameri, S. K., Ho, R., Jang, H., Tao, L., Wang, Y., Wang, L. [Liu], . . . Lu, N. (2017). Graphene electronic tattoo sensors. *ACS Nano*, *11*(8), 7634–7641. https://doi.org/10.1021/acsnano.7b02182

Ataei, M., Degbelo, A., & Kray, C. (2018). Privacy theory in practice: Designing a user interface for managing location privacy on mobile devices. *Journal of Location Based Services*, *12*(3-4), 141–178. https://doi.org/10.1080/17489725.2018.1511839

Barnes, S. B. (2006). A privacy paradox: Social networking in the United States. *First Monday*, *11*(9). https://doi.org/10.5210/fm.v11i9.1394

Batty, M. (2018). Digital twins. *Environment and Planning B: Urban Analytics and City Science*, *45*(5), 817–820. https://doi.org/10.1177/2399808318796416

Bjarnadóttir, M. V., Agarwal, R., Crowley, K., Jin, Q., Barnes, S. [Sean], & Prasad Kislaya (2014). Improving decision-making using health data analytics. In K. Marconi & H. Lehmann (Eds.), *Big Data and Health Analytics* (pp. 285–307). New York, NY: Auerbach Publications.

Björnsson, B., Borrebaeck, C., Elander, N., Gasslander, T., Gawel, D. R., Gustafsson, M., . . . Benson, M. (2019). Digital twins to personalize medicine. *Genome Medicine*, *12*(4). https://doi.org/10.1186/s13073-019-0701-3

Bohm, R. (2018). *Industrial Internet of Things for developers*. Hoboken, NJ: John Wiley & Sons, Inc.

Brost, G. S., & Hoffmann, M. (2015). Identifying security requirements and privacy concerns in digital health applications. In S. A. Fricker, C. Thümmler, & A. Gavras (Eds.), *Requirements engineering for digital health* (Vol. 14, pp. 133–154). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-09798-5_7

Brown, I. (2016). The economics of privacy, data protection and surveillance. In J. Bauer & M. Latzer (Eds.), *Handbook on the economics of the Internet* (pp. 247–261). Cheltenham, UK: Edward Elgar.

Brownstein, J. S., Freifeld, C. C., & Madoff, L. C. (2009). Digital disease detection - Harnessing the Web for public health surveillance. *The New England Journal of Medicine, 360*(21), 2153-5, 2157. https://doi.org/10.1056/NEJMp0900702

Bryman, A. (2012). *Social research methods* (Fourth edition). Oxford, UK: Oxford University Press.

Campbell, C. (2019, January 16). How China is using "social credit scores" to reward and punish its citizens. Retrieved from https://time.com/collection-post/5502592/china-social-credit-score/

Castro, D., & Atkinson, R. (2009). Ten ideas for policymakers to drive digital progress. *IEEE Internet Computing, 13*(2), 69–73. https://doi.org/10.1109/MIC.2009.47

Cavoukian, A. (2011). Privacy by design: The 7 foundational principles. Retrieved from https://www.ipc.on.ca/wp-content/uploads/resources/7foundationalprinciples.pdf

Cavoukian, A. (2012). *Operationalizing privacy by design: A guide to implementing strong privacy practices*. Ontario, Canada: Information and Privacy Commissioner.

Cavoukian, A., Fisher, A., Killen, S., & Hoffman, D. A. (2010). Remote home health care technologies: How to ensure privacy? Build it in: Privacy by design. *Identity in the Information Society, 3*(2), 363–378. https://doi.org/10.1007/s12394-010-0054-y

Cearley, D., & Burke, B. (2018, October 15). *Top 10 strategic technology trends for 2019*. Retrieved from Gartner website: https://www.gartner.com/en/doc/383829-top-10-strategic-technology-trends-for-2019-a-gartner-trend-insight-report

Chaari, L. (2019). *Digital health approach for Predictive, Preventive, Personalised and Participatory Medicine* (Vol. 10). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-11800-6

Chen, H.-T. (2018). Revisiting the privacy paradox on social media with an extended privacy calculus model: The effect of privacy concerns, privacy self-efficacy, and social capital on privacy management. *American Behavioral Scientist, 62*(10), 1392–1412. https://doi.org/10.1177/0002764218792691

Cho, H., Ippolito, D., & Yu, Y. W. (2020). *Contact tracing mobile apps for COVID-19: Privacy considerations and related trade-offs*. Retrieved from https://arxiv.org/pdf/2003.11511

Christl, W., & Spiekermann, S. (2016). *Networks of control: A report on corporate surveillance, digital tracking, big data & privacy*. Wien, Austria: facultas.

Clarke, R. Y. (2013). Smart cities and the Internet of everything: The foundation for delivering next-generation citizen services [White paper]. Retrieved from https://www.cisco.com/c/dam/en_us/solutions/industries/docs/scc/ioe_citizen_svcs_white_paper_idc_2013.pdf

CNN (2020, April 16). Coronavirus outbreak timeline fast facts. Retrieved from https://edition.cnn.com/2020/02/06/health/wuhan-coronavirus-timeline-fast-facts/index.html

Colesky, M., Hoepman, J.-H., & Hillen, C. (2016). A critical analysis of privacy design strategies. In *2016 IEEE Security and Privacy Workshops (SPW),* San Jose, CA.

Coppersmith, G., Dredze, M., & Harman, C. (2014). Quantifying mental health signals in Twitter. In P. Resnik, R. Resnik, & M. Mitchell (Eds.), *Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality* (pp. 51–60). Stroudsburg, PA, USA: Association for Computational Linguistics. https://doi.org/10.3115/v1/W14-3207

Corrie Health: Re-engineering heart attack discharge and recovery (2019). Retrieved from https://corriehealth.com/

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Thousand Oaks, CA: SAGE Publications Inc.

Dassault Systèmes. The living heart project. Retrieved from https://www.3ds.com/products-services/simulia/solutions/life-sciences/the-living-heart-project/

Demchenko, Y., Ngo, C., Laat, C. de, Membrey, P., & Gordijenko, D. (2013). Big security for big data: Addressing security challenges for the big data infrastructure. In *10th VLDB Workshop on Secure Data Management*, Trento, Italy.

Dick, J., Hull, E., & Jackson, K. (2017). *Requirements engineering* (4th ed.). Cham, Switzerland: Springer.

Dobkowski, D. (2019, April 9). MiCORE: Smartphone app with Apple watch decreases readmissions after MI. Retrieved from https://www.healio.com/cardiology/vascular-medicine/news/online/%7B99a3f453-88db-4ce8-ac74-8be6a6fd2b9e%7D/micore-smartphone-app-with-apple-watch-decreases-readmissions-after-mi

Dohrmann, K., Gesing, B., & Ward, J. (2019). *Trend report: Digital twins in logistics: A DHL perspective on the impact of digital twins.* Retrieved from https://www.logistics.dhl/content/dam/dhl/global/core/documents/pdf/glo-core-digital-twins-in-logistics.pdf

Early, J., & Bustillos, D. (2018). An Internet for some threatens health for All: What effects could the repeal of net neutrality in the USA have on individual and population health? *Global Health Promotion*, *27*(2), 109-113. https://doi.org/10.1177/1757975918785354

El Saddik, A. (2018). Digital twins: The convergence of multimedia technologies. *IEEE MultiMedia*, *25*(2), 87–92. https://doi.org/10.1109/MMUL.2018.023121167

European Association for Predictive, Preventive and Personalised Medicine (n.d.). EPMA mission. Retrieved from http://www.epmanet.eu/mission/epma-mission

European Commission (n.d.a). Draft code of conduct on privacy for mobile health applications. Retrieved from http://ec.europa.eu/newsroom/dae/document.cfm?action=display&doc_id=16125

European Commission (n.d.b). What does the General Data Protection Regulation (GDPR) govern? Retrieved from https://ec.europa.eu/info/law/law-topic/data-protection/reform/what-does-general-data-protection-regulation-gdpr-govern_en

European Commission (2018, December 10). Privacy code of conduct on mobile health apps: Shaping Europe's digital future. Retrieved from https://ec.europa.eu/digital-single-market/en/privacy-code-conduct-mobile-health-apps

The European Parliament and the Council of the European Union (2016a). Document 32016R0679. Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1589281487422&uri=CELEX:32016R0679

The European Parliament and the Council of the European Union (2016b, April 27). Regulation (EU) 2016/679: GDPR. Retrieved from https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:02016R0679-20160504

Eysenbach, G. (2001). What is e-health? *Journal of Medical Internet Research, 3*(2). https://doi.org/10.2196/jmir.3.2.e20

Floridi, L. (2015). *The onlife manifesto: Being human in a hyperconnected era*. Cham: Springer Open. Retrieved from http://www.doabooks.org/doab?func=fulltext&rid=17320 https://doi.org/10.1007/978-3-319-04093-6

Fricker, S. A. (2015). *Requirements engineering for digital health*. Cham, Switzerland: Springer.

Fricker, S. A., Grau, R., & Zwingli, A. (2015). Requirements engineering: Best practice. In S. A. Fricker, C. Thümmler, & A. Gavras (Eds.), *Requirements engineering for digital health* (pp. 25–46). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-09798-5_2

Gartner (n.d.). Gartner hype cycle: Interpreting technology hype. Retrieved from https://www.gartner.com/en/research/methodologies/gartner-hype-cycle

GE Healthcare Partners (2018, March 14). Applying simulation modeling to the hospital environment [White paper]. Retrieved from https://uscan.gehealthcarepartners.com/insight-detail/applying-simulation-modeling-to-the-hospital-envir

Glaessgen, E. H., & Stargel, D. S. (2012). The digital twin paradigm for future NASA and U.S. Air Force vehicles. In *53rd Structures, Structural Dynamics, and Materials Conference: Special Session on the Digital Twin* (pp. 1–14). Honolulu, HI.

Grieves, M. (2014). Digital twin: Manufacturing excellence through virtual factory replication [White paper]. Retrieved from https://research.fit.edu/media/site-specific/researchfitedu/camid/documents/1411.0_Digital_Twin_White_Paper_Dr_Grieves.pdf

Grieves, M., & Vickers, J. (2017). Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems. In F.-J. Kahlen, S. Flumerfelt, & A. Alves (Eds.),

Transdisciplinary perspectives on complex systems (Vol. 89, pp. 85–113). Cham: Springer International Publishing.

Haag, S., & Anderl, R. (2018). Digital twin – Proof of concept. *Manufacturing Letters*, *15*, 64–66. https://doi.org/10.1016/j.mfglet.2018.02.006

Hargittai, E., & Marwick, A. (2016). "What can I really do?" Explaining the privacy paradox with online apathy. *International Journal of Communication*, *10*, 3737–3757. Retrieved from https://ijoc.org/index.php/ijoc/article/view/4655/1738

Helm, A., & Georgatos, D. (2014). Privacy and mhealth: How mobile health apps fit into privacy framework not limited to HIPAA. *Syracuse Law Review*, *64*(1), 131–170.

Horowitz, J. (2020, March 12). Italy's health care system groans under coronavirus - A warning to the world. *The New York Times*. Retrieved from https://www.nytimes.com/2020/03/12/world/europe/12italy-coronavirus-health-care.html

Hu, F., Xie, D., & Shen, S. (2013). On the application of the Internet of Things in the field of medical and health care. In *2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing* (pp. 2053–2058). Piscataway, NJ: IEEE. https://doi.org/10.1109/GreenCom-iThings-CPSCom.2013.384

IES (n.d.). NTU Singapore. Retrieved from https://www.iesve.com/ntu-singapore

IES (2018). University of Nottingham project SCENe: Trent Basin. Retrieved from https://www.iesve.com/icl/case-studies/2821/University-of-Nottingham-Project-SCENe:-Trent-Basin

Ishwarappa, & Anuradha, J. (2015). A brief introduction on big data 5Vs characteristics and Hadoop technology. *Procedia Computer Science*, *48*, 319–324. https://doi.org/10.1016/j.procs.2015.04.188

Johnson, J. S., Friend, S. B., & Lee, H. S. (2017). Big data facilitation, utilization, and monetization: Exploring the 3Vs in a new product development process. *Journal of Product Innovation Management*, *34*(5), 640–658. https://doi.org/10.1111/jpim.12397

Jones, S. (2020, March 24). Spain: Doctors struggle to cope as 514 die from coronavirus in a day. *The Guardian*. Retrieved from https://www.theguardian.com/world/2020/mar/24/spain-doctors-lack-protection-coronavirus-covid-19

Jumelle, A. K. L., & Ispas, I. (2015). Ethical issues in digital health. In S. A. Fricker, C. Thümmler, & A. Gavras (Eds.), *Requirements engineering for digital health* (Vol. 16, pp. 75–93). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-09798-5_4

Kaur, N., & Sood, S. K. (2017). Dynamic resource allocation for big data streams based on data characteristics (5Vs). *International Journal of Network Management*, *27*(4), e1978. https://doi.org/10.1002/nem.1978

Keen, A. (2015). *The Internet is not the answer*. London, UK: Atlantic Books.

Khan, S. I., & Hoque, A. S. (2016). Digital health data: A comprehensive review of privacy and security risks and some recommendations. *Computer Science Journal of Moldove*, *24*(2), 273–292.

Kim, D.-H., Lu, N., Ma, R., Kim, Y.-S., Kim, R.-H., Wang, S., . . . Rogers, J. A. (2011). Epidermal electronics. *Science*, *333*(6044), 838–843. https://doi.org/10.1126/science.1206157

Kotz, D., Avancha, S., & Baxi, A. (2009). A privacy framework for mobile health and home-care systems. In L. J. Camp & E. Ferrari (Eds.), *Proceedings of the First ACM Workshop on Security and Privacy in Medical and Home-Care Systems* (p. 1). New York, NY: Association for Computing Machinery. https://doi.org/10.1145/1655084.1655086

Kuo, A. M.-H. (2011). Opportunities and challenges of cloud computing to improve health care services. *Journal of Medical Internet Research*, *13*(3), e67. https://doi.org/10.2196/jmir.1867

Lauzeral, N., Borzacchiello, D., Kugler, M., George, D., Rémond, Y., Hostettler, A., & Chinesta, F. (2019). A model order reduction approach to create patient-specific mechanical models of human liver in computational medicine applications. *Computer Methods and Programs in Biomedicine*, *170*, 95–106. https://doi.org/10.1016/j.cmpb.2019.01.003

Li, H., Wu, J. [Jing], Gao, Y., & Shi, Y. (2016). Examining individuals' adoption of healthcare wearable devices: An empirical study from privacy calculus perspective. *International Journal of Medical Informatics*, *88*, 8–17. https://doi.org/10.1016/j.ijmedinf.2015.12.010

Liu, Y. [Ying], Zhang, L. [Lin], Yang, Y., Zhou, L., Ren, L., Wang, F., . . . Deen, M. J. (2019). A novel cloud-based framework for the elderly healthcare services using digital twin. *IEEE Access, 7*, 49088–49101. https://doi.org/10.1109/ACCESS.2019.2909828

Lupton, D. (2012). M-health and health promotion: The digital cyborg and surveillance society. *Social Theory & Health, 10*(3), 229–244. https://doi.org/10.1057/sth.2012.6

Lupton, D. (2014a). The commodification of patient opinion: The digital patient experience economy in the age of big data. *Sociology of Health & Illness, 36*(6), 856–869. https://doi.org/10.1111/1467-9566.12109

Lupton, D. (2014b). Critical perspectives on digital health technologies. *Sociology Compass*, *8*(12), 1344–1359. https://doi.org/10.1111/soc4.12226

Lupton, D. (2017). *Digital health: Critical and cross-disciplinary perspectives*. *Critical approaches to health*. London, UK: Routledge Taylor & Francis Group.

Lyon, D. (2002). *Surveillance society: Monitoring everyday life*. *Issues in society*. Buckingham, UK: Open University Press.

Lyon, D. (2010). Surveillance, power and everyday life. In P. Kalantzis-Cope & K. Gherab-Martín (Eds.), *Emerging digital spaces in contemporary society: Properties of technology* (Vol. 40, pp. 107–120). London, UK: Palgrave Macmillan. https://doi.org/10.1057/9780230299047_18

Mai, J.-E. (2016). Big data privacy: The datafication of personal information. *The Information Society, 32*(3), 192–199. https://doi.org/10.1080/01972243.2016.1153010

Majumdar, A., & Bose, I. (2016). Privacy calculus theory and its applicability for emerging technologies. In V. Sugumaran, V. Yoon, & M. J. Shaw (Eds.), *E-life: Web-enabled convergence of commerce, work, and social life* (Vol. 258, pp. 191–195). Cham, Switzerland: Springer International Publishing.

Mannix, L. (2020, March 29). Australia's first virtual hospital rolls out for COVID-19 patients. *The Sydney Morning Herald*. Retrieved from https://www.smh.com.au/national/australia-s-first-virtual-hospital-rolls-out-for-covid-19-patients-20200329-p54ezj.html

Marcelino-Jesus, E., Sarraipa, J., Agostinho, C., & Jardim-Goncalves, R. (2014, September). A requirements engineering methodology for technological innovations assessment. In J.

Cha (Ed.), *CE 2014 - The 21st ISPE - International Conference on Concurrent Engineering*. Amsterdam: IOS Press.

Mayer-Schönberger, V., & Cukier, K. (2013). *Big Data: A revolution that will transform how we live, work, and think*. New York, NY: Houghton Mifflin Harcourt Publishing Company.

Mayring, P. (2016). *Einführung in die qualitative Sozialforschung: Eine Anleitung zu qualitativem Denken* (6th ed.). Weinheim, Germany: Beltz.

Mell, P., & Grance, T. (2010). The NIST definition of cloud computing. *Communications of the ACM*, *53*(6), 50.

Meskó, B., Drobni, Z., Bényei, É., Gergely, B., & Győrffy, Z. (2017). Digital health is a cultural transformation of traditional healthcare. *MHealth*, *3*, 38. https://doi.org/10.21037/mhealth.2017.08.07

Mohammadi, N., & Taylor, J. E. (2017). Smart city digital twins. In *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, Honolulu, HI.

Montgomery, K., Chester, J., & Kopp, K. (2018). Health wearables: Ensuring fairness, preventing discrimination, and promoting equity in an emerging Internet-of-Things environment. *Journal of Information Policy*, *8*, 34. https://doi.org/10.5325/jinfopoli.8.2018.0034

National Research Foundation Singapore (2018). Virtual Singapore. Retrieved from https://www.nrf.gov.sg/programmes/virtual-singapore

Nordgren, A. (2015). Privacy by design in personal health monitoring. *Health Care Analysis*, *23*(2), 148–164. https://doi.org/10.1007/s10728-013-0262-3

O'Connor, Y., Rowan, W., Lynch, L., & Heavin, C. (2017). Privacy by design: Informed consent and internet of things for smart health. *Procedia Computer Science*, *113*, 653–658. https://doi.org/10.1016/j.procs.2017.08.329

O'Neil, C. (2016). No safe zone: Getting insurance. In C. O'Neil (Ed.), *Weapons of math destruction: How big data increases inequality and threatens democracy* (pp. 161–178). New York, NY: Crown Publishing Group.

Panetta, K. (2018a, August 16). 5 trends emerge in the Gartner hype cycle for emerging technologies, 2018. Retrieved from https://www.gartner.com/smarterwithgartner/5-trends-emerge-in-gartner-hype-cycle-for-emerging-technologies-2018/

Panetta, K. (2018b, October 15). Gartner top 10 strategic technology trends for 2019. Retrieved from https://www.gartner.com/smarterwithgartner/gartner-top-10-strategic-technology-trends-for-2019/

Patel, K. K., & Patel, S. M. (2016). Internet of Things-IOT: Definition, characteristics, architecture, enabling technologies, application & future challenges. *International Journal of Engineering Science and Computing*, *6*(5), 6122–6131.

Pentina, I., Zhang, L. [Lixuan], Bata, H., & Chen, Y. (2016). Exploring privacy paradox in information-sensitive mobile app adoption: A cross-cultural comparison. *Computers in Human Behavior*, *65*, 409–419. https://doi.org/10.1016/j.chb.2016.09.005

Pohl, K. (1996). Requirements engineering: An overview. *Aachener-Informatik Berichte*. (05).

Pohl, K. (2010). *Requirements engineering: Fundamentals, principles, and techniques*. Berlin, Germany: Springer.

Rahman, M. S. (2019). Does privacy matters when we are sick? An extended privacy calculus model for healthcare technology adoption behavior. In *2019 10th International Conference on Information and Communication Systems (ICICS)* (pp. 41–46). Piscataway, NJ: IEEE. https://doi.org/10.1109/IACS.2019.8809175

Raley, R. (2013). Dataveillance and countervailance. In L. Gitelman (Ed.), *"Raw data" is an oxymoron* (pp. 121–145). Cambridge, MA: MIT Press.

Raskar, R., Schunemann, I., Barbar, R., Vilcans, K., Gray, J., Vepakomma, P., . . . Werner, J. [John] (2020). Apps gone rogue: Maintaining personal privacy in an epidemic [White paper]. Retrieved from https://arxiv.org/pdf/2003.08567

Ravitch, S. M., & Riggan, M. (2017). *Reason & rigor: How conceptual frameworks guide research* (2nd ed.). Thousand Oaks, CA: SAGE Publications Inc.

Richterich, A. (2018). Big data-driven health surveillance. In A. Richterich (Ed.), *The big data agenda: Data ethics and critical data studies* (pp. 71–90). London, UK: University of Westminster Press. https://doi.org/10.16997/book14.e

Rolim, C. O., Koch, F. L., Westphall, C. B., Werner, J. [Jorge], Fracalossi, A., & Salvador, G. S. (2010). A cloud computing solution for patient's data collection in health care institutions. In J. Finkelstein (Ed.), *2010 Second International Conference on eHealth,*

Telemedicine and Social Medicine (pp. 95–99). Piscataway, NJ: IEEE. https://doi.org/10.1109/eTELEMED.2010.19

Roski, J., Bo-Linn, G. W., & Andrews, T. A. (2014). Creating value in health care through big data: Opportunities and policy implications. *Health Affairs*, *33*(7), 1115–1122. https://doi.org/10.1377/hlthaff.2014.0147

Safavi, S., & Shukur, Z. (2014). Conceptual privacy framework for health information on wearable device. *PloS One*, *9*(12), e114306. https://doi.org/10.1371/journal.pone.0114306

Sanchez-Rola, I., Dell'Amico, M., Kotzias, P., Balzarotti, D., Bilge, L., Vervier, P.-A., & Santos, I. (2019). Can I opt out yet? GDPR and the global illusion of cookie control. In S. Galbraith, G. Russello, W. Susilo, D. Gollmann, E. Kirda, Z. Liang, & S. D. Galbraith (Eds.), *Proceedings of the 2019 ACM Asia Conference on Computer and Communications Security* (pp. 340–351). New York, NY: The Association for Computing Machinery. https://doi.org/10.1145/3321705.3329806

Saracco, R. (2020a, March 17). Personal digital twins role in epidemics control - I. Retrieved from https://cmte.ieee.org/futuredirections/2020/03/17/personal-digital-twins-role-in-epidemics-control/

Saracco, R. (2020b, March 18). Personal digital twins role in epidemics control – II. Retrieved from https://cmte.ieee.org/futuredirections/2020/03/18/personal-digital-twins-role-in-epidemics-control-ii/

Schleich, B., Anwer, N., Mathieu, L., & Wartzack, S. (2017). Shaping the digital twin for design and production engineering. *CIRP Annals*, *66*(1), 141–144. https://doi.org/10.1016/j.cirp.2017.04.040

Sharon, T. (2016). Self-tracking for health and the quantified self: Re-articulating autonomy, solidarity, and authenticity in an age of personalized healthcare. *Philosophy & Technology*, *30*(1), 93–121. https://doi.org/10.1007/s13347-016-0215-5

Shneiderman, B. (2000). Universal usability. *Communications of the ACM*, *43*(5), 84–91. https://doi.org/10.1145/332833.332843

Shugalo, I. (2019, April 29). Digital twin technology: Should healthcare jump on the bandwagon? Retrieved from https://hitconsultant.net/2019/04/29/digital-twin-technology-should-healthcare-jump-on-the-bandwagon/#.XThL8ugzbZt

Siemens Healthcare GmbH (2019). The value of digital twin technology [White paper]. Retrieved from https://www.siemens-healthineers.com/at/services/value-partnerships/asset-center/white-papers-articles/value-of-digital-twin-technology

Solove, D. J. (2006). A taxonomy of privacy. *University of Pennsylvania Law Review*, *154*(3), 477–560.

Solove, D. J. (2007). "I've got nothing to hide," and other misunderstandings of privacy. *San Diego Law Review*, *44*, 745–772.

Solove, D. J. (2008). *Understanding privacy*. Cambridge, MA: Harvard University Press.

Spaulding, E. M., Marvel, F. A., Lee, M. A., Yang, W. E., Demo, R., Wang, J. [Jane], . . . Martin, S. S. (2019). Corrie health digital platform for self-management in secondary prevention after acute myocardial infarction. *Circulation Cardiovascular Quality and Outcomes*, *12*(5). https://doi.org/10.1161/CIRCOUTCOMES.119.005509

Swan, M. (2012a). Health 2050: The realization of personalized medicine through crowdsourcing, the quantified self, and the participatory biocitizen. *Journal of Personalized Medicine*, *2*(3), 93–118. https://doi.org/10.3390/jpm2030093

Swan, M. (2012b). Sensor mania! The Internet of Things, wearable computing, objective metrics, and the quantified self 2.0. *Journal of Sensor and Actuator Networks*, *1*(3), 217–253. https://doi.org/10.3390/jsan1030217

Tao, F., Cheng, J., Qi, Q., Zhang, M., Zhang, H. [He], & Sui, F. (2018). Digital twin-driven product design, manufacturing and service with big data. *The International Journal of Advanced Manufacturing Technology*, *94*(9-12), 3563–3576. https://doi.org/10.1007/s00170-017-0233-1

Tao, F., Qi, Q., Wang, L. [Lihui], & Nee, A.Y.C. [A.Y.C.] (2019). Digital twins and cyber–physical systems toward smart manufacturing and industry 4.0: Correlation and comparison. *Engineering*, *5*(4), 653–661. https://doi.org/10.1016/j.eng.2019.01.014

Tao, F., Sui, F., Liu, A., Qi, Q., Zhang, M., Song, B., . . . Nee, A. Y. C. (2019). Digital twin-driven product design framework. *International Journal of Production Research*, *57*(12), 3935–3953. https://doi.org/10.1080/00207543.2018.1443229

Tao, F., & Zhang, M. (2017). Digital twin shop-floor: A new shop-floor paradigm towards smart manufacturing. *IEEE Access*, *5*, 20418–20427. https://doi.org/10.1109/ACCESS.2017.2756069

Tao, F., Zhang, M., Liu, Y. [Yushan], & Nee, A.Y.C. [A.Y.C.] (2018). Digital twin driven prognostics and health management for complex equipment. *CIRP Annals*, *67*(1), 169–172. https://doi.org/10.1016/j.cirp.2018.04.055

Taplin, J. (2017). *Move fast and break things: How Facebook, Google, and Amazon cornered culture and undermined democracy*. London, UK: Pan Macmillan.

Tuegel, E. J., Ingraffea, A. R., Eason, T. G., & Spottswood, S. M. (2011). Reengineering aircraft structural life prediction using a digital twin. *International Journal of Aerospace Engineering*, *2011*(3), 1–14. https://doi.org/10.1155/2011/154798

Valentino-DeVries, J., Singer, N., & Krolik, A. A Scramble for virus apps that do no harm. *The New York Times*. Retrieved from https://www.nytimes.com/2020/04/29/business/coronavirus-cellphone-apps-contact-tracing.html

Van Dijck, J. (2014). Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. *Surveillance & Society*, *12*(2), 197–208. https://doi.org/10.24908/ss.v12i2.4776

Van Houten, H. (2018, November 12). How a virtual heart could save your real one. Retrieved from https://www.philips.com/a-w/about/news/archive/blogs/innovation-matters/20181112-how-a-virtual-heart-could-save-your-real-one.html

Vayena, E., Haeusermann, T., Adjekum, A., & Blasimme, A. (2018). Digital health: Meeting the ethical and policy challenges. *Swiss Medical Weekly*, *148*, w14571. https://doi.org/10.4414/smw.2018.14571

Vayena, E., Salathé, M., Madoff, L. C., & Brownstein, J. S. (2015). Ethical challenges of big data in public health. *PLoS Computational Biology*, *11*(2), e1003904. https://doi.org/10.1371/journal.pcbi.1003904

Vogel, J., Auinger, A., Riedl, R., Kindermann, H., Helfert, M., & Ocenasek, H. (2017). Digitally enhanced recovery: Investigating the use of digital self-tracking for monitoring leisure time physical activity of cardiovascular disease (CVD) patients undergoing cardiac rehabilitation. *PloS One*, *12*(10). https://doi.org/10.1371/journal.pone.0186261

Wang, J. [Jinjiang], Ye, L., Gao, R. X., Li, C., & Zhang, L. [Laibin] (2019). Digital Twin for rotating machinery fault diagnosis in smart manufacturing. *International Journal of Production Research*, *57*(12), 3920–3934. https://doi.org/10.1080/00207543.2018.1552032

Woodhams, S. (2020, March 20). COVID-19 digital rights tracker. Retrieved from https://www.top10vpn.com/research/investigations/covid-19-digital-rights-tracker/

Woods, E., & Freas, B. (2019). Creating zero carbon communities: The role of digital twins [White paper]. Retrieved from https://learn.iesve.com/digital-twin-white-paper/?utm_source=blog&utm_campaign=navigant

World Health Organization (2020a, April 8). WHO Timeline - COVID-19. Retrieved from https://www.who.int/news-room/detail/08-04-2020-who-timeline---covid-19

World Health Organization (2020b, June 29). Coronavirus disease (COVID-19) pandemic. Retrieved from https://www.who.int/emergencies/diseases/novel-coronavirus-2019/

Wu, T. (2003). Network neutrality, broadband discrimination. *Journal of Telecommunications and High Technology Law*, *2*, 141-179. https://doi.org/10.2139/ssrn.388863

Yeung, K. (2017). 'Hypernudge': Big data as a mode of regulation by design. *Information, Communication & Society*, *20*(1), 118–136. https://doi.org/10.1080/1369118X.2016.1186713

Zintl, T., & Melia, E. (2020, April 22). Is the pandemic deepening the digital divide? Retrieved from https://www.die-gdi.de/en/the-current-column/article/is-the-pandemic-deepening-the-digital-divide/

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power* (1st ed.). New York, NY: PublicAffairs.

Appendix

Appendix A: Cover Letter

Subject: Digital Twins in Healthcare – Master Thesis Research

Dear [Name of Expert],

my name is Sarah Jeske, I am a second-year master student in the DCLead programme (http://dclead.eu). Currently, I am conducting my master thesis on digital twins in healthcare at the Digital Media Research Centre (https://research.qut.edu.au/dmrc) at Queensland University of Technology in Brisbane, Australia, as a visiting student.

I am reaching out to you as [give reason, e.g. you are a research scientist working on the Digital Twin project at company XY]. I believe that your insights and expertise would be a great contribution to my research.

Let me introduce my thesis briefly. The project aims to identify the benefits and challenges of digital twins in healthcare, with a strong focus on privacy. The goal is to create a conceptual framework of the implementation of digital twins in the field of healthcare that can act as a guide on what needs to be taken into consideration when developing such technology. This includes the technical requirements specific for digital twins as well as the privacy issues that come along with it.

Would you be willing to spare 30 minutes of your time to talk to me about this topic? The interview would be conducted virtually (e.g. Skype, Zoom etc.) any time after April 8. In case

you are not available to take part in an interview, I would very much appreciate if you could refer me or forward this email to a colleague for yours.

If you have any questions, please do not hesitate to contact me at sjeske19@student.aau.dk or on LinkedIn (https://www.linkedin.com/in/sarah-jeske-b06923170).

Thank you for considering this request. I am looking forward to hearing from you!

Kind regards,

Sarah Jeske

Appendix B: Interview Guides

Group 1: Digital Twin Experts

INTRODUCTION
Hello, thank you for taking the time today to talk to me about digital twins in healthcare. *As you know, this interview is part of the empirical research for my master thesis, which aims at identifying the challenges of digital twins in healthcare, with a strong focus on privacy. The goal is to create a conceptual framework of the implementation of digital twins in the field of healthcare that can act as a guide on what needs to be taken into consideration when developing such technology.* *I appreciate your effort and time. I believe that this interview will give me valuable insights for the research.* *Our discussion will be recorded and archived to ensure the accuracy and validity of the research. It will not be made public, but my notes and some citations from it may be included in the master thesis. Please let me know if you would like me to anonymise the citations.* *Do you have any questions before we start?* *If a question is not clear, please feel free to ask me and I will clarify it.*

WARM-UP, POSITION AND EXPERIENCE WITH DIGITAL TWINS	
DT-Q0. Could you shortly introduce yourself and the experience you have with digital twins?	
DEFINING DIGITAL TWINS	
DT-Q1. Are you familiar with the concept of digital twins?	
[yes] *In general, a digital twin setting consists of the physical entity, the digital entity, the connection between the two, and the value derived from that connection. Is there anything that you would like to add, specifically related to the health domain?*	[no; give explanation of digital twins and continue interview with focus on digital healthcare systems and applications]
DT-Q2. In your opinion, what are the biggest benefits of digital twins in healthcare/digital healthcare applications?	
PRIVACY RISKS	
DT-Q3. What risks do you associate with such technology?	
DT-Q4. How do you think these risks stack up against the benefits?	
DT-Q5. What role does the protection of personal health information play in such technology?	
DT-Q6. What role does the individual, whose data are being collected, play in all of it? How involved are they, how much control do they have?	
PRIVACY FRIENDLY DESIGN	
DT-Q7. How can an individual's privacy be protected in such systems?	
DT-Q8. What principles or guidelines should be followed to maintain privacy when implementing healthcare systems?	
DT-Q9. What (technical) solutions can be incorporated into such technology to ensure privacy?	
DT-Q10. How would a data breach be handled?	
FUTURE, COOL-DOWN	
DT-Q11. How do you perceive the future of digital twins/digital healthcare regarding privacy protection?	

Group 2: Privacy Experts

INTRODUCTION	
Hello, thank you for taking the time today to talk to me about the challenges of digital health. *As you know, this interview is part of the empirical research for my master thesis, which aims at identifying the challenges of digital twins in healthcare, with a strong focus on privacy. The goal is to create a conceptual framework of the implementation of digital twins in the field of healthcare that can act as a guide on what needs to be taken into consideration when developing such technology.* *I appreciate your effort and time. I believe that this interview will give me valuable insights for the research.* *Our discussion will be recorded and archived to ensure the accuracy and validity of the research. It will not be made public, but my notes and some citations from it may be included in the master thesis. Please let me know if you would like me to anonymise the citations.* *Do you have any questions before we start?* *If a question is not clear, please feel free to ask me and I will clarify it.*	
WARM-UP, POSITION (IF APPLICABLE: EXPERIENCE WITH DIGITAL TWINS)	
P-Q0. Could you shortly introduce yourself (if applicable: and the experience you have with digital twins)?	
DEFINING DIGITAL TWINS	
P-Q1. Are you familiar with the concept of digital twins?	
[yes] *In general, a digital twin setting consists of the physical entity, the digital entity, the connection between the two, and the value derived from that connection. Is there anything that you would like to add, specifically related to the health domain?*	[no; give explanation of digital twins and continue interview with focus on digital healthcare systems and applications]
P-Q2. In your opinion, what are the biggest benefits of digital twins in healthcare/digital healthcare applications?	

DATA COLLECTION
P-Q3. With legislations such as the GDPR in Europe, the protection of an individual's privacy when collecting data has become a top priority. How is the collection of health information affected by this paradigm shift?
DATA PROCESSING AND ANALYSIS
P-Q4. How is the individual affected if their data are used for big data analysis such as public health surveillance?
P-Q5. How much control should individuals have over what happens to their data after its collection (e.g. big data analysis)?
P-Q6. What do you think about the threat of health information being treated as a commodity and used for ulterior financial motives?
P-Q7. To what extend do you believe the benefits generally outweigh and justify the risks?
SURVEILLANCE
P-Q8. In your opinion, is digital healthcare rather a burden or a blessing to society when taking surveillance and concepts such as "the transparent citizen" into account?
FUTURE, COOL-DOWN
P-Q9. To what extent is it the developer's responsibility to design a system in a privacy friendly manner? How could an individual's privacy be protected in such systems?
DT-Q10. How do you perceive the future of digital twins/digital healthcare regarding privacy protection?

Appendix C: Codebook

Category 1: Position	
Definition	The current occupation of the interviewee.
Coding Rules	Statements are coded that relate to the person's job title and their responsibility in that position.
Prime Example	"I am a physician. I'm trained in paediatrics, anaesthesia, critical care. I have special training in something called clinical informatics. That's a very new field in the US where physicians are trained to do data driven research."

Category 2: Experience	
Definition	The extent to which the interviewee has prior knowledge of and/or experience with digital twins (in healthcare and in general).
Coding Rules	Statements are coded that relate to the person's professional background and qualifications as well as their familiarity and the contact they have had with digital twins previously.
Prime Example	"I'm an engineer by training. I have a master's in industrial engineering, bachelor's in mechanical engineering, and spent my entire career in healthcare, both on the consulting side, I worked for a couple of large consultancies in the US, I worked on the provider side as a hospital administrator at a large academic medical centre for a few years." "My responsibility related to digital twin is anywhere from R&D to operational and clinical space."

Category 3: Digital Twins	
Category 3a: Definition	
Definition	In general, a digital twin setting consists of the physical entity, the digital entity, the connection between the two, and the value derived from that connection. The purpose of this code is to find an agreement upon the general definition of digital twins in the health domain.
Coding Rules	Statements are coded that indicate agreement as well as any deviations and/or special features regarding digital twins in healthcare.
Prime Example	"I think the set-up is the same in, I would say, every industry. There will be differences if you look at the different components of the seller."

Category 3b: Conceptualisation	
Definition	Any information about the set-up and elements in the digital twin scenario in healthcare.
Coding Rules	Statements are coded that explain the set-up and the (technical) functionality of digital twins (in healthcare).
Prime Example	"So, really, AI is a huge enabler for the digital twin concept because without AI, you need someone to be able to create that digital representation, and be able to quantify that and that's not something you can do with a radiologist or pathologist."
Category 3c: Benefits	
Definition	The advantages that digital twins can bring for healthcare.
Coding Rules	Statements are coded that relate to the positive outcomes from using digital twins in healthcare.
Prime Example	"The benefit certainly is that it can enable more better personalized treatment."

Category 4: Risks	
Category 4a: Privacy Risks	
Definition	Disadvantages that digital twins can bring for healthcare which are related directly to privacy.
Coding Rules	Statements are coded that relate to threats and negative outcomes from using digital twins in healthcare that affect one's privacy directly.
Prime Example	"What tends to be privacy issues are those things that, to use a loose term, can be used against you."
Category 4b: Other Risks	
Definition	Disadvantages that digital twins can bring for healthcare which are not related directly to privacy.

Coding Rules	Statements are coded that relate to any other threats and negative outcomes from using digital twins in healthcare.
Prime Example	"A lot of the information in the health record should only be taken as a guide, because there's all sorts of errors that could happen in the testing or the patient may change."
Category 4c: Risk-Benefit-Analysis	
Definition	The relationship between the benefits and the risks.
Coding Rules	Statements are coded that contrast the advantages of digital twins with the disadvantages and threats.
Prime Example	"Privacy is overwritten or balanced out by the value of getting that data. So, the benefits of people seeing your data outweigh the risk."

Category 5: Privacy Friendly Design	
Category 5a: Privacy Protection	
Definition	Privacy protection measures to maintain an individual's privacy.
Coding Rules	Statements are coded that indicate the role that the protection of personal health information plays in a technology like digital twins in healthcare and how the privacy is maintained.
Prime Example	"Which is no different than how do I protect your blood pressure, your heart rate. When you go to a doctor, right? How is your data protected? The same rules, because what you're doing is, you're now doing a, you're just increasing the electronic medical record to a different level. The grounds of privacy are the same. The only difference is you're making the electronic medical record more dynamic now."
Category 5b: Principles and Guidelines	
Definition	Any principles or guidelines that can be or are being followed during the implementation process.

Coding Rules	Statements are coded that indicate principles, guidelines, protocols, or rules that can be applied during the implementation process to maintain an individual's privacy.
Prime Example	"I think the main guiding principle is that in the end, the digital twin belongs to the person."

Category 6: Data Collection	
Definition	The process of collecting personal (health) data.
Coding Rules	Statements are coded that regard the data collection of personal health information and to what extent this might be different from other personal data.
Prime Example	"What are actually the sensorized devices that are being used to collect data for this model? But also, what's the accuracy quality levels of these sensors in these devices that are potentially being designed for something else?"

Category 7: Data Processing and Analysis	
Definition	The handling and manipulation of collected personal (health) data.
Coding Rules	Statements are coded that explain how the manipulation of the data after its collection affects the individual and the risks that are associated with it.
Prime Example	"And I guess what's interesting here is that the ability to collect and model this type of data means that you can use not only the data, but the service outcome or product for a whole range of different activities that go beyond say something like an X ray. Or a more accurate medical imaging of the heart. You can push it once you connect it to sensorized devices, you can push and nudge these types of behaviours. That can have a significant impact on the autonomy of the individual."

Category 8: Surveillance	
Definition	Surveillance in digital (twin) health technologies and the calculus between them being a burden or a blessing.
Coding Rules	Statements are coded that include the interviewee's perspective on surveillance in digital health technologies and whether the benefits justify the risks.
Prime Example	"I don't think you can clearly delineate health considerations from social aspects of how we live our lives. So, once you get to that sort of situation, you're moving from health surveillance through monitoring how the heart is going to inevitably looking at the lifestyles of individuals. So, you're looking at using ostensibly sensor eyes technologies to not only track and monitor the health and the health of the heart, you're actually surveilling individuals for their lifestyles as well."

Category 9: Future	
Definition	The interviewee's assessment of the future for digital twins in healthcare.
Coding Rules	Statements are coded that indicate next steps and what the future holds for digital twins in healthcare.
Prime Example	"I think in the short run, it's certain parts of the anatomy and that's also how medicine has been working for many, many years. So, you have specialties on one particular organ or one particular system. In the short run, that's also where we're going with that. We're focused on certain areas"

Appendix D: Ethical Clearance

Co-funded by the Erasmus+ Programme of the European Union

Copenhagen, 02.04.2020

Ethical Clearance

For the study entitled: Conceptualizing Digital Twins in Healthcare

To be carried out by: Sarah Jeske

Between: 01.03.2020 – 31.07.2020

The author of this study has demonstrated to the satisfaction of the supervisors a suitable level of knowledge of applicable ethical principles, following the meeting held on 12.03.2020

Name of the supervisors of the thesis: Reza Tadayoni, Ursula Maier-Rabler

1

On behalf of both supervisors,

Name: Reza Tadayoni

AALBORG UNIVERSITET

UNIVERSITÄT SALZBURG

VUB VRIJE UNIVERSITEIT BRUSSEL