Spring 2022

Privacy and Security Concerns Associated with MHealth Technologies: A Social Media Mining Perspective

Damion Mitchell
Dakota State University

DAKOTA STATE UNIVERSITY

PRIVACY AND SECURITY CONCERNS ASSOCIATED WITH MHEALTH TECHNOLOGIES: A SOCIAL MEDIA MINING PERSPECTIVE

A doctoral dissertation submitted to Dakota State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Information Systems

October 31, 2022

By
Damion R. Mitchell

Dissertation Committee:
Dr. Omar El-Gayar, Chair
Dr. Jun Liu
Dr. Insu Park
Dr. Renae Spohn

DISSERTATION APPROVAL FORM

This dissertation is approved as a credible and independent investigation by a candidate for the Doctor of Philosophy degree and is acceptable for meeting the dissertation requirements for this degree. Acceptance of this dissertation does not imply that the conclusions reached by the candidate are necessarily the conclusions of the major department or university.

Student Name: Damion R. Mitchell

Dissertation Title: PRIVACY AND SECURITY CONCERNS ASSOCIATED WITH MHEALTH TECHNOLOGIES: A SOCIAL MEDIA MINING PERSPECTIVE

Dissertation Chair/Co-Chair: DocuSigned by: Omar El-Gayar BE671FC34D9845C... Date: October 31, 2022

Name: Omar El-Gayar

Dissertation Chair/Co-Chair: _____ Date: _____

Name:

Committee member: DocuSigned by: Jun Liu E0CE171C3CDE419... Date: October 31, 2022

Name: Jun Liu

Committee member: DocuSigned by: Insu Park AD181583719C41C... Date: October 31, 2022

Name: Insu Park

Committee member: DocuSigned by: Renae Spohn 05B32D18A30E436... Date: October 31, 2022

Name: Renae Spohn

ACKNOWLEDGMENT

I would like to express my gratitude to my Dissertation Chair, Dr. Omar El-Gayar, for his support and guidance. He has been enduring, encouraging and a great source of motivation throughout my tenure as a Ph.D. student. Dr. El-Gayar has inspired me from his calm and peaceful demeanor even during unprecedented times. His technical and editorial advice were essential to the completion of this dissertation and has taught me invaluable lessons and insights. I would like to sincerely thank my committee members Dr. Jun Liu, Dr. Insu Park, and Dr. Renae Spohn for their valuable comments, suggestions and input to the dissertation. Thanks a lot for your patience and time. I express my gratitude and love to my dear wife, Syrie for her immense care and support throughout this journey; along with my three children Zoie, Zayne, and Zahra for providing me with the necessary motivation. Thanks to the valuable support and encouragement of my Mi Familia Group and especially my praying friend, Dr. Louis James. Thanks to my accountability partner, Dr. Gloria Gregory for ensuring that I completed my tasks in a timely manner. Last, but by no way the least, I want to thank God for His sustaining power.

ABSTRACT

mHealth technologies seek to improve personal wellness; however, there are still significant privacy and security challenges. With social networking sites serving as lens through which public sentiments and perspectives can be easily accessed, little has been done to investigate the privacy and security concerns of users, associated with mHealth technologies, through social media mining. Therefore, this study investigated various privacy and security concerns conveyed by social media users, in relation to the use of mHealth wearable technologies, using text mining and grounded theory. In addition, the study examined the general sentiments toward mHealth privacy and security related issues, while unearthing how the various issues have evolved over time. Our target social media platform for data collection was the microblogging platform Twitter, which was accessed through Brandwatch providing access to the "Twitter firehose" to extract English tweets. Triangulation was conducted on a representative sample to confirm the results of the Latent Dirichlet Allocation (LDA) Topic Modeling using manual coding through ATLAS.ti.

By using the grounded theory analysis methodology, we developed the D-MIT Emergent Theoretical Model which explains that the concerns of users can be categorized as relating to data management, data invasion, or technical safety issues. This model claims that issues affecting data management of mHealth users through the misuse of their data by entities such as wearable companies and other third-party applications, negatively impact their adoption of these devices. Also, concerns of data invasion via real-time data, security breaches, and data surveillance inhibit the adoption of mHealth wearables, which is further impacted by technical safety issues. Further, when users perceived that they do not have full control over their wearables or patient applications, then their acceptance of these mHealth technologies is

diminished. While a lack of data and privacy protection policies contribute negatively to users' adoption of these devices, it also plays a pivotal role in the data management issues presented in this emergent model. Therefore, the importance of having robust legal and policy frameworks that can support mHealth users is desired. Theoretically, the results support the literature on user acceptance of mHealth wearables. These findings were compared with extant literature, and confirmations found across several studies.

Further, the results show that over time, mHealth users are still concerned about areas such as security breaches, real-time data invasion, surveillance, and how companies use the data collected from these devices. The findings reveal that more than 75% of the posts analyzed were categorized as depicting anger, fear, or demonstrating levels of disgust. Additionally, 70% of the posts exhibited negative sentiments, whereas 26% were positive, which indicates that users are ambivalent concerning privacy and security, notwithstanding mentions of privacy or security issues in their posts.

DECLARATION

I hereby certify that this dissertation constitutes my own product, that where the language of others is set forth, quotation marks so indicate, and that appropriate credit is given where I have used the language, ideas, expressions or writings of another.

I declare that the dissertation describes original work that has not previously been presented for the award of any other degree of any institution.

Signed,

Damion R. Mitchell

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

This chapter presents a detailed discussion on the background to the research problem, the statement of the problem and the objectives of the research. It commences with an comprehensive review of the background of the research problem and then examines key factors that were significant to the formation of the research objectives and then concludes with an overview of the structure and flow of this document.

Background of the Problem

Social media platforms have seen unprecedented growth worldwide, with these platforms being more favored that the more traditional media sources for obtaining and sharing information in real time. Social media are online, often mobile, platforms that support the creation and exchange of user-generated content (Kaplan & Haenlein, 2010). According to Pew Research Center, about 72% of U.S. adults use some type of social media sites[1]. These sites have served various purposes and functions, with diversities in information dissemination, personal activities posting, product reviews, advertisements, and sentiments. Consequently, researchers have tapped into this large data source that pointedly increases the range of what can easily be measured, and thus enables computational knowledge discovery. People's online activity in social media is increasingly being used as a source of data for research (Wilson et al., 2012). Studies are done seeking to understand users' behaviors, demographics, interaction and networks, or users' responses or sentiments towards particular topics, products, or policies (Anstead & O'Loughlin, 2015).

[1] http://www.pewinternet.org/fact-sheets/social-networking-fact-sheet/

In recent years, mobile technologies have been evolving at an unprecedented rate, and mobile health (mHealth), the use of mobile technologies in medicine, has also surged analogous to these technological developments. The World Health Organization (2011) defines mHealth as "medical and public health practice supported by mobile devices, such as mobile phones, patient monitoring devices, personal digital assistants, and other wireless devices." This definition is supported by Kotz et al. (2016) who defined mHealth as the use of mobile technologies—wearable, implantable, environmental, or portable—by individuals who monitor or manage their own health, perhaps with the assistance of individual caregivers or provider organizations. It was further opined that mHealth technologies can be utilized for wellness goals such as losing weight, eating a healthy diet, quitting smoking, or becoming physically fit. mHealth technologies have transformed the means by which individuals seek and receive health care, manage chronic conditions, and access medical records (Acquisti et al., 2015; Filkins et al., 2016; Gallagher et al., 2017; Goldberg et al., 2016).

Through different types of technologies and platforms, mHealth has enabled individuals to become more engaged in their health care experience (Yardley et al., 2016). In addition, telemedicine services provide instantaneous access to urgent care and mental health providers through the patient's smartphone or tablet (Miller et al., 2019; Stowell et al., 2018). Healthcare workers are able to remotely monitor patients' blood pressure or weight by using Bluetooth-enabled devices (Ganapathy et al., 2016; Vegesna et al., 2017). mHealth wearables with self-monitoring and tracking functions have dramatically increased the users' ability to self-manage and monitor their health and well-being (Haghi et al., 2017; Metcalf et al., 2016). There are many types of mHealth devices; two examples include an armband, BodyMedia Fit, that tracks daily activities and a wristwatch, Glucowatch, that monitors blood glucose level.

Statement of the Problem

Although there is enormous potential for mHealth technologies to increase healthcare quality, expand access to services, and improve personal wellness, there are still significant privacy and security challenges (Al Ameen et al., 2012; Giannetsos et al., 2011). Literature commonly cited privacy problems as the primary barrier to the persistent adoption of mHealth technologies such as wearables (K. Kang et al., 2013; L. Lee et al., 2015). Previous research has shown that privacy concerns and perceptions of security risks can hinder the usage of e-commerce systems (Eastlick et al., 2006; Malhotra et al., 2004), online health information systems (Bansal et al., 2010) and in particular of location-based services (LBS) of mHealth technologies (Zhou, 2012). The concept of privacy is not new, and it has generally been defined as an individual's ability to control the terms by which their personal information is acquired and used (Westin, 1968). Privacy is also described as protecting personal information from being misused by malicious entities and allowing certain authorized entities to access that personal information by making it visible to them (Bunnig & Cap, 2009).

Mobile health sensing devices can help individuals work towards a healthier lifestyle or allow them to share the collected information with their doctor to diagnose health issues or manage a chronic disease (Prasad et al., 2012). Although mHealth technologies may indeed improve quality of healthcare and quality of life, they also generate security and privacy issues. Past research has focused on privacy and security concerns in the context of mHealth technologies (Arora et al., 2014; Avancha et al., 2012; Iwaya et al., 2020; Zhao et al., 2020). Other research has examined different health related issues through the use of social media mining (R. Correia et al., 2020; Domalewska, 2021; Sarker et al., 2016). Other studies have used social media analytics for knowledge discovery in domains such as public health

surveillance (Fung et al., 2015; Y. Kang et al., 2017; Paul et al., 2016); discovering adverse drug events (R. B. Correia et al., 2016; Nguyen et al., 2017; Wu et al., 2013); discovery of health related information (Lu et al., 2013; Tuarob et al., 2014); disease trend prediction (McGough et al., 2017; Nagar et al., 2014; Santos & Matos, 2014); and disease intervention (Robinson et al., 2015; Tanner et al., 2016). However, with social networking sites serving as lens through which public sentiments and perspectives can be easily accessed, to the best of our knowledge, little has been done to investigate the privacy and security concerns of users, associated with mHealth technologies, through social media mining. Accordingly, this research seeks to fill the gap by examining mHealth security and privacy related topics, compare and contrast the findings with extant literature, and propose an emergent theoretical framework that explains these user expressed concerns.

Objectives of the Dissertation

This research aims to systematically analyze social media users' privacy and security concerns with mHealth devices and to contribute an emergent theoretical framework to the body of knowledge that can be used to explain these concerns. Based on the huge volume of data available online, and the need for efficient data analysis, the research employed text mining and grounded theory (Al-Ramahi et al., 2016) with an interpretivist worldview. Grounded theory from an interpretivist worldview enables researchers to inductively develop a theory or pattern of meaning rather than start with a theory, as in a post positivist worldview (Creswell, 2003). The study seeks to answer the following research questions:

RQ1: What are the expressed privacy and security concerns of social media users in the context of mHealth technologies?

RQ2: What are the general sentiments toward mHealth privacy and security related issues?

RQ3: How has the perception of various mHealth related issues evolved over time?

From a *theoretical perspective*, the findings of this study contribute to the literature of users' acceptance of health consumer technology, by unearthing the privacy and security concerns that may inhibit their adoption. Further, the findings provide evidence through the D-MIT Emergent Theoretical model, that users of mHealth wearables are concerned about data management, data invasion or technical safety issues, which finds support from extant literature dealing with privacy and security concerns in the wearables domain. Finally, the study reveals which of the privacy and security concerns mHealth users are most concerned about.

From a *methodological perspective*, the capability of text mining within the grounded theory context was utilized. We used the LDA algorithm for topic modeling, to automatically extract concepts from large amounts of text data, instead of manually analyzing and coding the tweets, which is time-consuming and subjective. As far as we know, this is the first work that leverages social media mining to understand the privacy and security concerns of mHealth users. Automatically evaluating social media users' posts with the utilization of machine learning tools, can assist in understanding the themes and topics that exist in the tweets shared by online users.

From a *practical perspective*, the research further contributes to the growing mHealth industry, particularly with the proliferation of wearable devices. The findings of the study can help policy makers with developing comprehensive guidelines to govern data collection, dissemination, and processing on these devices. Additionally, the findings can guide

companies which develop and distribute mHealth wearable devices in better understanding the expressed concerns of users, especially in the area of having greater control over the wearables and also apps used for patient care. Doctors and other health practitioners who use these mHealth devices can understand reasons which may inhibit the adoption of these wearables by their patients, and develop strategies to mitigate these concerns.

Structure of the Dissertation

The remainder of the dissertation is structured as follows. In chapter 2, a theoretical background and a comprehensive literature review of related work are represented. Chapter 3 examines the research methodology adopted in this dissertation. Chapter 4 presents the findings from the grounded theory analyses, emotion and sentiment analyses, and the general evolution of privacy and security concerns. Chapter 5 presents the emergent theoretical model. Finally, Chapter 6 concludes the report by presenting an overview of the contributions of this project, the limitations of the study as well as the future directions of this research project.

CHAPTER 2

LITERATURE REVIEW

It is of import that any research conducted should seek to investigate past efforts in order to appreciate what is known about a specific area and discover and contribute to the ongoing discussion and evolution towards viable solutions to a related but new research problem (Machi & McEvoy, 2016). This section explores the current state of knowledge in relation to mHealth technologies, social media usage, privacy and security with the main aim of understanding the privacy and security concerns of these users, through social media mining.

mHealth Technologies

Mobile health technologies have revolutionized the way individuals seek and receive health care, manage chronic conditions, and access medical records. Therefore, with the advent of miniaturized sensors, low-power body-area wireless networks, and pervasive smartphones, the burgeoning field of mobile health (mHealth) technologies have attracted tremendous commercial activity, consumer interest, and adoption by major healthcare providers (Kotz et al., 2016a). mHealth has emerged over the past 20 years as an integrative discipline, focusing on developing and implementing wireless, portable, or implantable technology for improving human health (Andreu-Perez et al., 2015; Kumar et al., 2013). The global mHealth market is projected to grow at a rate of 36.5% between 2016 and 2022, and would ultimately reach a size of US$ 22.31 billion by 2022 (Market Research Focus, 2020). Through mHealth technology (e.g., mobile apps and wearable devices), health care-related

information, knowledge, and consultations can be delivered to patients at any time, which is helpful for disease prediction and self-management behaviors (Kumar et al., 2013).

McCallum et al. (2018) identify different categories of mHealth technologies such as wearables for encouraging physical activity and mobile applications for medication adherence which have become ubiquitous in modern society as tools for health promotion. The proliferation of mHealth technologies has also seen infrastructural updates across the world, such as investments in the expansion of both cellular and internet facilities (Adibi et al., 2013). In addition, engineers and physicians are collaborating to develop innovative mobile intervention approaches, with greater focus on the integration of mobile technology into clinical treatment approaches with health care organizations (Charani et al., 2017; Krohn, 2015). Considering the popularity of mHealth technologies, privacy and security play a pivotal role in safeguarding the users' data.

Theoretical Foundation

In order to ground further understanding of the expressed privacy and security concerns of mHealth technologies, we will explore a theoretical foundation that combines the privacy calculus theory, protection motivation theory (PMT), and the theory of planned behavior (TPB) which extends the theory of reasoned action (TRA) (Ajzen, 1991; Ajzen & Fishbein, 2000; Maddux & Rogers, 1983; Rogers, 1975). Since mHealth technologies collect users' personal health information on an ongoing basis, concern about data privacy risk increases. An individual's decision to adopt these technologies would involve an obvious privacy calculus, in which users may consider the trade-off between perceived benefit and perceived privacy risk (Li et al., 2016). The theory assumes that individuals evaluate anticipated benefits and perceived risks in order to make a rational decision regarding the

disclosure of their personal data (Culnan & Armstrong, 1999; Dinev & Hart, 2006). The results of Kim et al. (2019) demonstrated that both perceived benefits and perceived privacy risks have an effect on the willingness to provide personal information when using different IoT services.

Perceived privacy refers to "an individual's self-assessed state in which external agents have limited access to information about him or her" (Dinev et al., 2013, p. 299). Most research treats privacy as a state whether it is implicit or explicit (Dinev et al., 2013). For example, Westin (1967) discussed 'states of privacy' while both Altman (1976) and Westin (1967) refer to 'state of control' and 'state of limited access'. Privacy concern and trust are two known proxies of perceived privacy (Dinev et al., 2013; Flavián et al., 2006). Both privacy concern and trust are attitudinal factors indicating people's current mental state toward certain objectives (Vaske & Donnelly, 1999). Privacy concern is the negative mental state and trust is the positive mental state that influence the overall self-assessed state of perceived privacy (Dinev et al., 2013).

A fundamental aspect of privacy is the control over personal data. It was Altman (1976) who posited that privacy is "selective control of access to the self or one's group" (p. 24). This suggests that users' control over personal data and its utilization must be guaranteed in order to ensure privacy. Perceived control is also one of the foundations of the theory of planned behavior (Ajzen, 1991). This theory states that behavioral intentions and subsequently actual behavior of individuals reflect the interplay of their attitudes, perceived social norms and perceived control over an action. Wang and Nepali (2015) state that control over the purpose of information collection determines users' provision of personal data. Brandimarte et al. (2013) further found that individuals disclose even personal identifiable

information when they perceive to be in control over its release and access. Therefore, control provides a sense of security and is a crucial factor in assessing the associated risks. In the context of TRA, behavioral intentions determine individual behavior, i.e., an individual's attitude toward the behavior and subjective norms about the behavior (Ajzen & Fishbein, 1980).

The security concerns of users characterize a grave issue that can influence the trust levels of mHealth technologies and may impede the adoption of these devices or applications (AlHogail, 2018; Falcone & Sapienza, 2018; Gao & Bai, 2014). Therefore, these devices or applications must certainly gain the users' confidence and provide assurance that they are safe to use. The security risks associated with mHealth technologies are much higher compared to other ICT technologies, since these mobile devices typically have limited computational power, battery life, and run a weaker encryption system in general. Based on the fact that mHealth technologies collect a variety of information from users via their sensors, it is obvious that the way security is implemented, and the effect it has on the users' perception is directly related to the demonstrated trust in these technologies (AlHogail, 2018). Among all the theories that explain health behavior, PMT is regarded as a better theory than others (Prentice-Dunn & Rogers, 1986; Weinstein, 1993) to investigate individual's behaviors toward health information technology. Perceived vulnerability refers to the possibility that one will experience health threat, while perceived severity represents the extant of threat from unhealthy behaviors (Rogers, 1975). Based on PMT, Guo et al. (2015) found that age plays a major role with threat appraisal and coping appraisal factors in mHealth acceptance.

Another frequently utilized framework that focuses on individuals' (and groups') decision-making processes regarding privacy is the Concerns and Privacy Management (CPM) theory (Petronio, 2002). CPM theory claims that privacy should not be measured as establishing a maximum boundary for keeping others out, but rather as a conciliation between

accessibility and retreat (Taddicken, 2014; Trepte et al., 2015). Consequently, privacy management entails a dynamic process within which individuals use stratagems—called privacy rules— to control these boundaries. Thus, a key reason for the incongruity between privacy concerns and behavior may be that users perceive the risk to privacy to be lower than the benefits of sharing.

Security and Privacy in mHealth

Privacy and security issues impede the adoption and diffusion of technology in the IT domain (Cho et al., 2009; C. Lee et al., 2011). The first reference in the literature to an instrument for measuring privacy concern about personal information is the scale called "Concern for Information Privacy" (Smith et al., 1996). In this research, the authors presented a theoretical framework that conceptualized privacy concerns about personal information in five key dimensions, namely: collection, unauthorized secondary internal use, unauthorized secondary external use, improper access, and errors in personal data. Privacy interests can be affected by various activities, i.e. (1) information collection, (2) information processing, (3), information dissemination, and (4) invasion (Solove, 2006). Clarke (1999) defined four categories of privacy, including privacy of the person, privacy of personal data, privacy of personal behavior and privacy of personal communication. However, Finn et al. (2013) expanded Clarke's categorization to seven types of privacy: privacy of the person, privacy of behaviour and action, privacy of personal communication, privacy of data and image, privacy of thoughts and feelings, privacy of location and space, and privacy of association.

Faudree and Ford (2013) postulated that the use of mHealth technologies among healthcare providers and consumers may bring significant issues, such as security and privacy challenges. Owing to the high data sensitivity and the mobility of the devices, privacy

concerns have proved to be more important in the context of health wearables than other technological devices (Miltgen et al., 2013). Most of the previous works on user privacy have focused on mobile devices and their applications (Shklovski et al., 2014), social networks (Gurses & Diaz, 2013), or web applications (Reidenberg et al., 2014).

In addition, Berendt et al. (2005) conducted a study to better understand users' privacy concerns. Even though the focus of this work was on e-commerce applications, it showed a gap between reported concerns and actual users' behaviors, reinforcing that those users generally sacrifice their privacy in exchange of benefits. Previous studies primarily use the privacy calculus theory to investigate individuals' willingness to share personal health information (PHI) voluntarily if they expect that perceived benefits from data disclosure outweigh the perceived costs (Anderson & Agarwal, 2011; Li et al., 2016). This trade-off theory has been presented as "the most useful framework for analyzing contemporary consumer privacy concerns" (Culnan & Bies, 2003, p. 326). Plachkinova et al. (2015) presented a three-dimensional model for classifying mHealth applications in terms of security and privacy concerns. They posited that privacy-related threats can be classified as identity threats, where patients may lose their identity credentials, thus allowing access to their personal health information; and Access threats, where patients have ultimate control on the collection, use, and disclosure of PHI, but if they fail to express their consent broader-than-intended access may be granted.

Security refers to the safeguards, techniques, and tools used to protect against the inappropriate access or disclosure of information. As such it is one of the key factors in protecting the users from any type of uncertainties and risks. In the mHealth context specifically, security covers the three triads of confidentiality (ensuring that the collected data

is accessible only to the authorized entities), integrity (ensuring the correctness and trueness of the data being transmitted), and availability (survivability despite security attacks). Users' security concerns are a serious issue that can affect the trust levels and hinder the adoption rate (AlHogail, 2018; Al-Momani et al., 2016; Falcone & Sapienza, 2018; Gu et al., 2017). In mHealth, information is often transmitted at a high frequency and transferred over wireless networks, which can be more susceptible to monitoring and interception than broadband (Internet) networks, making security protocols the only barriers protecting data against a breach (Luxton et al., 2012). Therefore, mHealth technologies such as wearables must gain the users' confidence and provide assurance that they will be safe. Thus Arora et al. (2014) in their study shared several solutions to prevent privacy and security breaches in mHealth while maintaining the benefits (Table 1).

Table 1. Addressing confidentiality, privacy, and security challenges in mHealth

	Risk	**Solution**
	De-identification	Share data in aggregate Separate transmission of identifying information (name, location) from other data
	Consent	Use consent to educate participants about what data are being collected and what can be inferred from such data Include privacy and safety training for participants Consider allowing patients to choose which data to share and with whom
	Breaches from intended user	Enable password, pin, or passphrase on phones before distribution Enable remote wiping
	Encryption	Use WPA2 and 128-bit key encryption Add a tag or header to the encrypted message
	Data transmission	Use non-sensitive messages to contact participants Store data remotely, such as on a secure server or in a cloud
	Data accessibility	Store critical data in two locations to ensure availability
	Data integrity and quality	Have a second system to collect the same data, such as in-person visits or surveys, to verify mobile data integrity and quality

Location	Have adjustable security settings for trusted and untrusted locations
Authentication	Use two-factor authentication, such as with a pin/password and a token/smart card/dongle
Audits and risk assessment	Include audits in security protocols, potentially with the help of a "red team"; risk assessment should be done at each stage of implementation

Social Media Mining

Social networking sites are lenses through which public sentiment can be easily accessed. These sites have been used in research to predict stock prices (Bing et al., 2014), to gauge political opinions (Ahmad et al., 2019), and for the highlighting of areas of focus for public health (Seltzer et al., 2017). Social media mining is the process of extracting and analyzing patterns from user data available online (Han et al., 2011; Tan et al., 2005). Social media mining includes an array of analyses, from simple counting of the likes, retweets, and users' demographics to more sophisticated measuring of quantifiable information such as sentiment, popularity, or reach. Data mining techniques encompass social network analysis, Bayesian networks, decision trees, natural language processing, and other algorithms (Domalewska, 2021). However, text mining is an automated technique that uses computational algorithms to extract meaning and patterns from already existing text (Gemar & Jiménez-Quintero, 2015). Text mining discovers new knowledge by analyzing and identifying the relevant information from large amounts of currently existing unstructured data. In addition, text mining aims at recognizing associations between words in sentences rather than just discovering words, as done in popular search engines.

Research Gap

Public sentiments and perspectives can be easily accessed through the lens of social networking sites. However, to the best of our knowledge, little has been done to investigate

the privacy and security concerns of users, associated with mHealth technologies, through social media mining. The review of literature revealed a lack of comprehensive research studies conducted to provide sufficient findings in highlighting privacy and security concerns that may inhibit the acceptance of mHealth wearables. Additionally, the literature is clear that there is significant need for additional empirically based theories on the subject of privacy and security concerns in mHealth technologies and their impact on the adoption of these devices. This problem represents a gap in literature that this study addresses, by examining mHealth security and privacy related topics, and comparing and contrasting the findings with extant literature, and proposing an emergent theoretical framework that explains users expressed concerns. Further, to address this research gap in literature, a grounded theory (empirical) with text-mining is performed to identify varying privacy and security concerns of mHealth users, and then via the established techniques of the research methodology, we will develop substantive theory to explain the observed phenomena of mHealth users expressed privacy and security concerns.

CHAPTER 3

RESEARCH METHODOLOGY

This chapter discusses the technique used in this project for discovering privacy and security concerns associated with mHealth technologies through social media mining. First, it describes the data collection and preparation. It then explains the topic modeling technique used to extract privacy and security concerns of social media users. Figure 1 shows the research methodology adapted in this study. According to Al-Ramahi et al. (2016), text mining and grounded theory are seen as epistemologically compatible since text mining allows for the extraction of concepts and theories from the data. Therefore, we sought to automate the extraction and analysis of social media posts through text mining within the grounded theory framework (Culnan & Armstrong, 1999; Dinev & Hart, 2006). The first stage involved data collection, based on a specific time and keywords of interest. The collected tweets were pre-processed and open-coded using text mining. The Latent Dirichlet Allocation (LDA) algorithm was used for topic modeling to automatically extract concepts from the large corpus of text data. These findings were then confirmed using manual coding through ATLAS.ti on a representative sample. We performed axial coding and selective coding to extract relevant higher-level categories and propositions. Brandwatch (BW), a social media mining platform was used to analyze the data for aspects such as sentiment and trend analyses (Grimmer & Stewart, 2013; Hopkins & King, 2010).

.

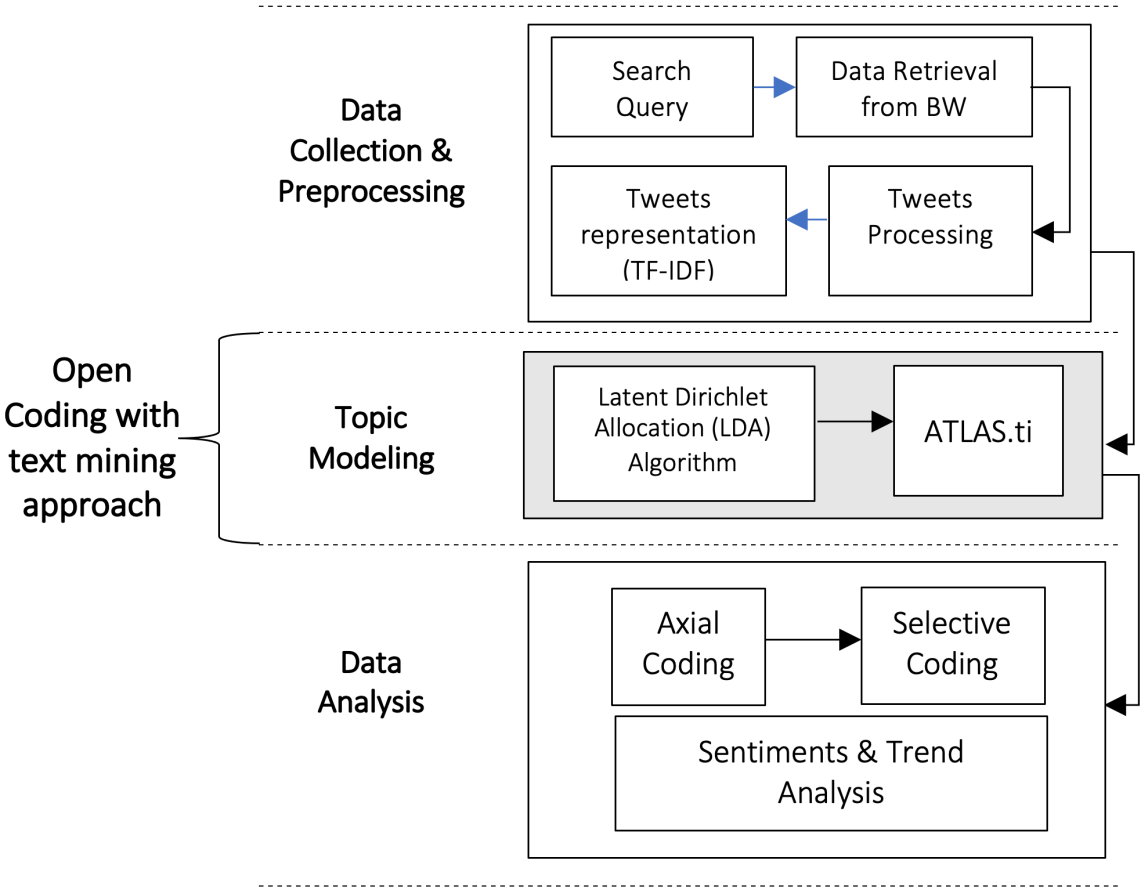

Figure 1. *Research Approach (Adapted from Al-Ramahi et al. 2016)*

Ground Theory Methodology

Grounded Theory (GT) approach has the capability of conceptual thinking and theory building rather than theory or hypothesis testing (Charmaz, 2011). GT differs from other qualitative methods, as it permits the "development of theories directly from raw data, data collection and analysis in a systematic manner, and maintains the data to be grounded, rather than forcing data to fit with current theories, thus fostering creativity" (Charmaz, 2014, p. 8). Charmaz (2006) suggests that the exploration experience commences with "finding data" (p. 14). He further postulated that data will uncover the unique situation and structure of the

respondent's lives notwithstanding the disclosure of their sentiments, perspectives, aim and activities (Charmaz, 2006).

The inductive approach espoused by GT will depend on the researcher methodically collecting, coding, categorizing, and analyzing the data (Charmaz, 2006), to determine the theory that clarifies the phenomenon in the instance of the privacy and security concerns of mHealth technologies, from the viewpoint of social media users. The rich data used was based on text mining techniques applied to social media posts using the Brandwatch platform. There are generally three analytic types of coding in grounded theory, namely: open coding, axial coding, and selective coding. Open coding comes up with concepts, while axial coding represents the process of developing main categories and their sub-categories. Lastly, selective coding deals with the integration of the categories that have been developed to build theoretical framework (Pandit, 1996).

Studies which use a grounded theory methodology can make three contributions to research: development of theory, development of a model, or a rich description of phenomena (Wiesche et al., 2017). According to Bacharach (1989) a theory is a statement of relations among concepts within a set of boundary assumptions and constraints providing detailed explanations. Further, a model is a visual illustration of abstract variables and their respective relationships amongst one another (Sutton & Staw, 1995). Additionally, rich descriptions are narratives based on observations with few generalizations or abstractions (Van Maanen, 1990). The goal of this dissertation was to create an emergent theoretical model.

Data Collection & Preprocessing

This stage involved data collection, based on a specific time period and keywords of interest. The collected tweets were preprocessed by removing stop words, retweets, addresses,

and certain words that are not context appropriate. We performed lemmatization and represented each document using the well-known Term Frequency Inverse Document Frequency (TF-IDF) weighting scheme (Haddi et al., 2013). Specifically, TF-IDF weight of a word *i* in a document *j* is given by

$$F_{i,j} * Log\ (N/DF)$$

Where $F_{i,j}$ is the frequency of the word *i* in the document *j*, N indicates the number of documents in the corpus, and DF is the number of documents that contains word *i*. Our target social media platform for data collection was the microblogging platform Twitter. We used Brandwatch which provides access to the "Twitter firehose" with the search query shown in Figure 2, where a total of 64,179 English tweets were extracted for the period January 1, 2010 to April 30, 2022. The keywords were identified by examining the literature (V. G. Motti & Caine, 2014; Solove, 2006) as well as through the use of online synonym generators.

```
((wearable* OR Fitbit* OR "Apple Watch" OR "Google Glass*" OR "Samsung Gear*" OR wristwatch
OR "acitivity track*" OR smartwatch* OR  "smart watch*" OR "fitness track*" OR "sport watch*" OR
implantable OR Garmin* OR "Samsung Galaxy watch*" OR "ECG Monitor*" OR "Blood Pressure
Monitor*" OR pedometer OR  glucometer OR "heart rate monitor*" OR mhealth OR AmazFit*)

AND (privacy OR intrude OR intrus* OR expose* OR spy* OR distrust OR consent* OR authentic* OR
insecur* OR unsecur* OR leak* OR anonym* OR malicious* OR hijack* OR compromise* OR
unauthoriz* OR harvest* OR hack* OR theft OR risk* OR breach* OR invad* OR captur* OR invasi*
OR secur* OR disclos* OR priva* OR (third AND part*) OR (3rd AND part*) OR sensitive* OR
unauthoriz* OR vulnerabl* OR violate* OR (privacy AND polic*) OR (data AND collection) OR
surveillance* OR (data AND capture) OR (personal AND data)))

AND - (RT OR http OR https OR author:(fitbit* OR apple* OR garmin*) OR beach* OR hill*)
```

Figure 2. Search query used for data collection

Open Coding using Text Mining

Data analysis is a fundamental component in grounded theory, since theories are developed from the data (Corbin & Strauss, 1990). In this phase of the analysis the labeling and categorization of the phenomena discovered in the posts was done (Charmaz, 2006). According to Myers (2009) the open coding phase forms the basic foundation for grounded theory construction. Several researches using the grounded theory methodology, normally utilize manual content analysis; however in this study, we utilized text mining to automatically extract concepts and associated theories from the corpuses of social media posts. Text mining is a process of obtaining useful information from document collections through the identification and exploration of interesting patterns (Feldman & Sanger, 2006). The text-mining process, like that of grounded theory, requires impartiality which will allow for categories to emerge from the data (Yu et al., 2011). This approach for data analysis was chosen, as text mining allowed us to code using automated algorithms, that is, we were able to extract meaningful information out of the corpuses, thus eliminating the subjectivity and many delimitating factors associated with manual coding (Yu et al., 2011).

Topic Modeling with Latent Dirichlet Allocation (LDA)

Topic models are statistical algorithms that can be used to discover the hidden thematic structure (i.e., topics) from large unstructured collections of documents by analyzing the words within the texts (Blei, 2012). Topic modeling algorithms do not necessitate any prior labeling or annotations of the documents and allow the topics to emerge from the examination of the original texts. In this study, Latent Dirichlet Allocation (LDA)-based topic Modeling (Blei, 2003a) was used, which is known to have the highest performance among several topic modeling algorithms when dealing with large-scale documents and interpreting

identified latent topics (Chiru et al., 2014). The model produces automatic summaries of topics in terms of a discrete probability distribution over words for each topic; additionally it deduces per-document discrete distributions over topics. The interface between the observed documents and hidden topic structure is revealed in the probabilistic generative process associated with LDA (Blei, 2012). LDA assumes the following generative process for a corpus D consisting of M documents which were extracted from Brandwatch, each of length N_i.

To demonstrate the results of LDA, Let M be the number of documents in a collection, K the number of topics, N the number of words in a document, *and* V the vocabulary size. The first result is the $M \times K$ matrix, where the weight $w_{m,k}$ is the relationship between a document d_m and a topic t_k. The second result is the $N \times K$ matrix, where the weight $w_{n,k}$ is the connection between a word w_n and a topic t_k. The notations *Dirichlet* $(\cdot)$ and *Multinomial* $(\cdot)$ represent Dirichlet and multinomial distribution with parameter $(\cdot)$, respectively. The graphical representation of LDA is shown in Figure 3, and the corresponding generative process is shown below:

(1) For each topic $t \in \{1, \ldots, K\}$,

(a) draw a distribution over vocabulary words

$\boldsymbol{\beta} t \sim Dirichlet(\boldsymbol{\eta})$.

(2) For each document d,

(a) draw a vector of topic proportions

$\boldsymbol{\theta}_d \sim Dirichlet(\boldsymbol{\alpha})$.

(b) For each word w_n in document d, where $n \in \{1, \ldots, N\}$,

(i) draw a topic assignment

$\mathbf{z}_n \sim Multinomial(\boldsymbol{\theta}_d)$;

(ii) draw a word $w_n \sim Multinomial(\boldsymbol{\beta}_{\mathbf{z}n})$.

The notations $\boldsymbol{\eta}$ and $\boldsymbol{\alpha}$ represent the hyperparameters of the corresponding Dirichlet distributions. The notation $\boldsymbol{\beta}_t$ is the V-dimensional word distribution for topic t, and $\boldsymbol{\theta}_d$ is the K-dimensional topic proportion for document d. After the topic modeling was done, a representative sample of the data was imported to the ATLAS.ti, a Qualitative Data Analysis and Research software, where manual coding was done to further confirm the LDA topic-modeling results.

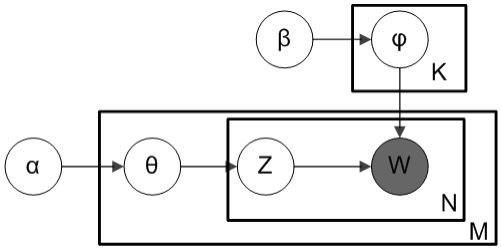

Figure 3. LDA-based topic modeling process

Predictive Power of Topic Models

To measure the predictive power of LDA models with varying number of topics, we utilized a metric called perplexity that is standard in language modeling (Azzopardi & Rijsbergen, 2003). Typically, the evaluation of topic models includes determining how well a model performs when unobserved documents are being predicted. Furthermore, when estimating the probability of unseen held-out documents, given a set of training documents, an ideal model should give rise to a higher probability of held-out documents. According to Blei (2003) a lower perplexity over a held-out document is equivalent to a higher log-likelihood,

which indicates better predictive and generalization performance. Formally, for a test set D_{test} of M documents, the per-word perplexity is defined as:

$$Perplexity(D_{test}) = \exp(-\sum_{d=1}^{M} \log p(w_d) / \sum_{d=1}^{M} N_d)$$

Where N_d is the number of words in document d (Blei, 2003).

In our research we evaluated different LDA models, by varying the number of topics (k) and evaluated them against the held-out test items. Therefore, the perplexity of a held-out test set was computed to evaluate the generated models. The dataset was divided in which 80% of the data was used to train the models, while the remaining 20% was used for the held-out test set. The predictive power of the models in terms of the held-out per-word perplexity was accomplished by changing the number of topics.

Topic Labeling

According to Chang et al. (2009), topics are typically manually labeled to ensure high labeling quality particularly when such classification requires domain knowledge. To guarantee that the labeling was not biased, two independent researchers reviewed and labeled the ten (10) topics. The level of agreement between the two researchers was measured using Cohen's Kappa. The calculated Kappa statistics = .80, which indicates substantial agreement between the labeling of the two researchers (Landis & Koch, 1977; McHugh, 2012). Since the agreement was high, the labeling was completed using one researcher.

Axial Coding

While the focus of open-coding is on generating categories and their properties and determining how the groupings vary dimensionally, the focus of axial-coding is on relating

categories to their subcategories at the level of properties and dimensions (Strauss & Corbin, 1998) and noting the dynamic interrelationships between categories to form the basis for theory construction (Goulding, 2002). The goal of axial coding is to initiate the formation of conceptual groupings. In the open coding with text-mining activities the data was broken into its most granular units. During the axial coding the researcher compared code against code and looked for emergent relationships that form conceptual groupings. These concepts form the constructs of the developing theory.

Selective Coding

Selective-coding is the process of integrating categories to build a theory and to refine the theory (Glaser & Strauss, 1967). Its task is to relate categories found in axial-coding to a core category which represents the main theme of research. To discover the central category and its relationship with the other categories, we used two techniques: (a) using diagrams (the final diagrams are shown in Figures 3 and 4) and (b) reviewing extant literature. We started this process after some categories had been discovered in axial-coding and continued with modification and refinement until we reached theoretical saturation. After each loop of coding (open-axial-selective), we could further develop the multiple layers of categorized theoretical statements. Theoretical codes are developed to explain and describe the relationships between the categories developed at the axial coding level (Charmaz, 2006).

The Six C's Approach

To organize emerged categories and their mutual relationships we applied the ''Six C's'' coding family (Glaser, 1978). According to Glaser this is the first general code to keep

in mind when analyzing the data. Table 2 describes the terms used in the Glaser's Six C's model. In our case, these code categories were a good fit for the phenomena that emerged.

Table 2. Six C's Terminologies

Term	Description
Context	The setting where the category is at play
Condition	A factor that is a prerequisite for the category to emerge
Cause	A reason for the category to occur
Consequence	Outcomes or effects as a result of the occurrence of the category
Contingency	A moderating factor between categories and consequences
Covariance	Categories or parts thereof can co-vary with each other, meaning that a change in one category inflicts a change in the other.

The *Context* in our case is the mHealth privacy and security concerns expressed by social media users. In general each of the core categories will arise under certain *Conditions*, and has certain *Causes* and *Consequences*. These causes are mitigated by *Contingencies*. Finally, possible combination of categories or parts thereof as known as *Covariances*. In our study there was insufficient information to reliably show how causes and consequences varied across users, as such the covariance component was not discussed in the result section. Figure 4 shows our adaptation of Glaser's Six C's Model.

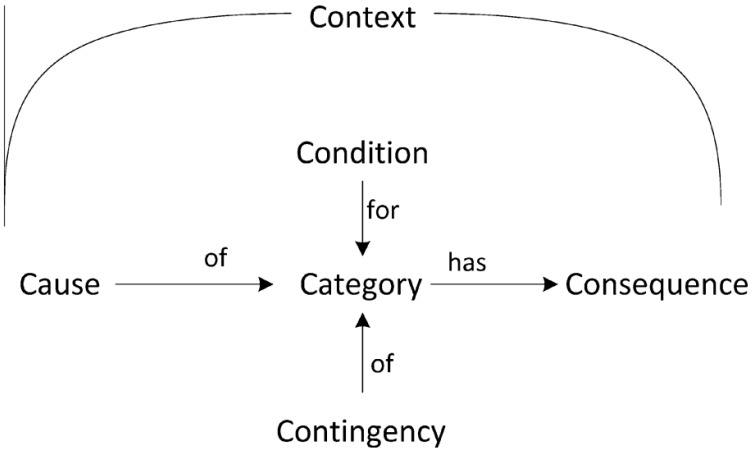

Figure 4. Adaptation of Glaser's Six C's Model

Sentiments & Trend Analysis

Textual data can be broadly categorized into facts and opinions; facts are objective expressions such as entities and events and their properties, while opinions are subjective expressions that describe people's sentiments, appraisals, or feelings (Liu, 2010). Sentiment analysis involves the task of automatically ascribing positive, negative, or neutral sentiment to portions of text that express opinions (Jeong et al., 2019). Furthermore, emotion analysis provides an additional layer of contextual analysis by the utilization of "Ekman 6" (Anger, Fear, Disgust, Joy, Surprise, and Sadness) basic human emotions (Ekman, 1993). The researchers used Brandwatch which employs BrightView, a supervised algorithm which is an updated version of the ReadMe algorithm developed by (Hopkins & King, 2010). The algorithm is based on aggregate analysis to allow flexibility and accuracy, which is primarily suited when the researcher wants to depict the volume of tweets that fit in to specific categories over time.

The algorithm requires the researcher to manually code a training set of documents into a set of predefined groups. In contrast to traditional classification methods that focus on maximizing the percent of documents correctly classified into a given set of categories, the ReadMe algorithm emphasize the broad categorization about the whole sets of documents (Hopkins & King, 2010). Accordingly individual-level classification is not a result of this method and traditional classification performance metrics based on the confusion matrix do not apply. Examples further illustrating the use of the algorithm and its supporting platform include Al-Ramahi et al. (2021), El-Gayar et al. (2021), Jamal et al. (2015), Kim et al. (2013), and Runge et al. (2013). In this study, the collected tweets represent the set of documents, and the predefined categories were obtained from the topic modeling stage. The researcher

assigned at least 20 tweets into each category, after which the BrightView algorithm was executed on past and future tweets returned by the search query. The tweets were examined based on the assigned categories, and further training was conducted where necessary.

CHAPTER 4

FINDINGS

This chapter presents the study's findings based on the three research questions under investigation. It will elaborate on the detailed codes and their relationships that emerge from the data and provide an appropriate diagram. A table overview of the findings is presented followed by a detailed discussion of each construct. The chapter concludes with a discussion of the theoretical model and the relationships between the constructs.

Tweets Information

Figure 5 represents a total of 64,179 unique English tweets which were returned for the period June 1, 2010, to April 30, 2022. These tweets according to Figure 6 were primarily from users residing in United State of America (54%), United Kingdom (17%), Canada (6%), India (3%), and Australia (3%).

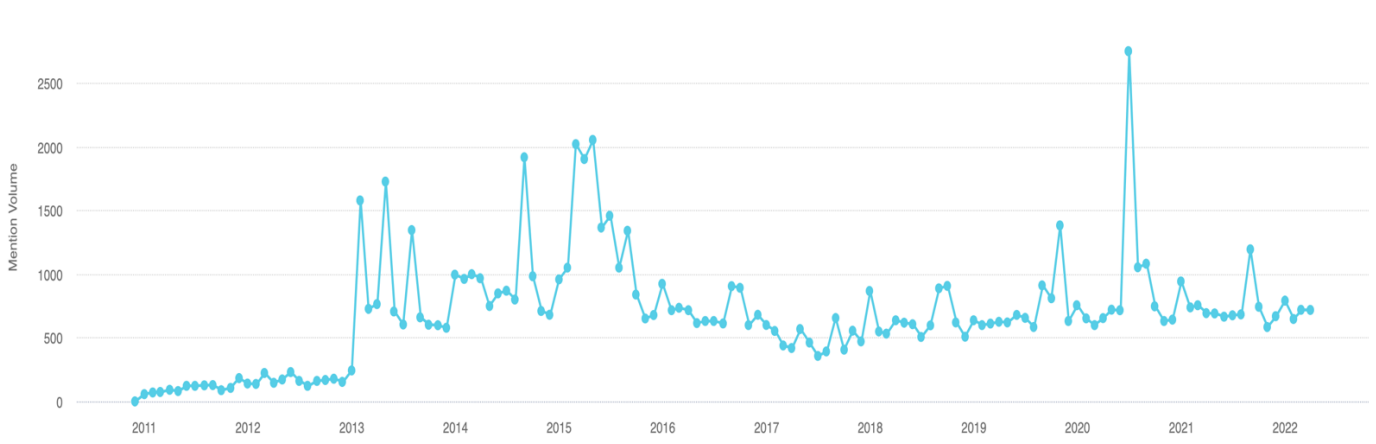

Figure 5. Volume of Tweets for search period

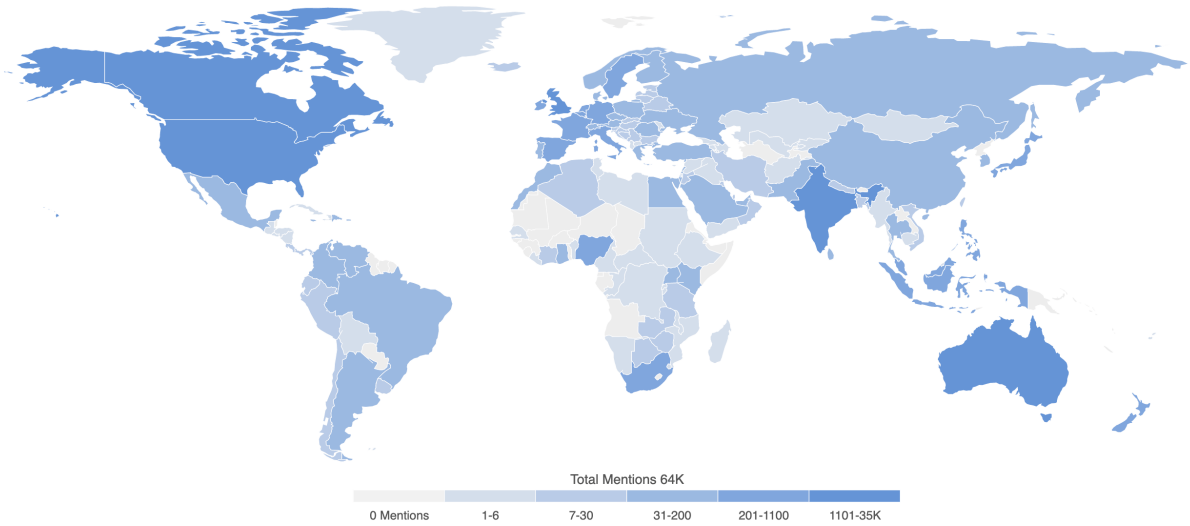

Figure 6. Volume of tweets by countries

The demographics by gender for the corpus of tweets as shown in Figure 7, highlights that 68% of tweets under were tweeted by males, and 32% by females.

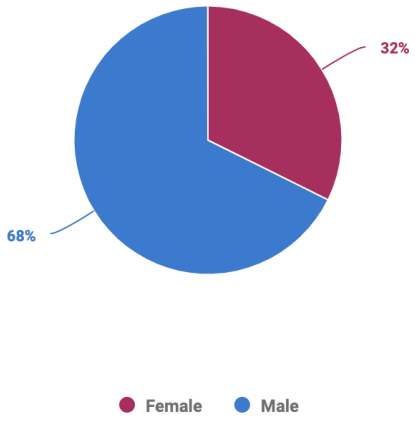

Figure 7. Distribution of tweets by gender

Open Coding Results

Table 3 illustrates the result of a 10-topic LDA model produced during the open coding phase, where each topic was represented by the top-15 weighted words in its

vocabulary distribution. A descriptive word or phrase was then ascribed to each topic to signify the main privacy and security concerns related to healthcare wearable technologies.

Table 3. Privacy and security concerns

Categories – abstractions based on the concepts	**Concepts** – abstractions from the open codes	**Open Codes** – top-15 weighted words generated by the LDA algorithm
Data Management Issues impact the acceptance of mHealth wearables.	Misuse of Data	datum, personal, health, information, use, device, collect, secure, record, harvest, know, share, wearable, patient, protect
	Capture of Personal Data	datum, capture, people, need, personal, good, think, activity, want, track, acquire, include, really, point, research
	Company Use of Data	Fitbit, datum, personal, user, health, company, information, buy, sell, google, wellness, acquisition, pay, say
	Third-Party Data Access	access, data, change, body, level, medical, year, increase, work, info, company, tech, phone, monitor, third-party
	Data and Privacy Protection	Protection, datum, privacy, data, issue, consumer, say, healthcare, security, concern, share, challenge, need, service, high
Data Invasion Issues impact the acceptance of mHealth wearables.	Real Time Data Invasion	time, private, make, datum, real, enable, invasion, store, use, thank, protect, hand, data, people
	Security Breach	wearable, security, breach, privacy, tech, wear, bring, human, connect, device, camera, create, data, capture
	Data Surveillance	surveillance, look, right, watch, day, Apple, business, product, disease, use, monitor, market, long, trust
Technical Safety Issues impact the acceptance of mHealth wearables.	Control over wearables	control, wearable, gesture, device, technology, come, let, use, thing, project, glass, home, remote, smartwatch, want
	Control over Patient Apps	control, patient, app, help, mobile, way, phone, allow, use, health, love, improve, device, monitoring, care

Throughout the open coding phase using text-mining, the generated weighted words were constantly compared to identify relationships and common concepts in the corpus of tweets. Through this process of comparison different categories emerged from the data. Figure 8, provides a summary of the main concepts and abstracted categories that materialized.

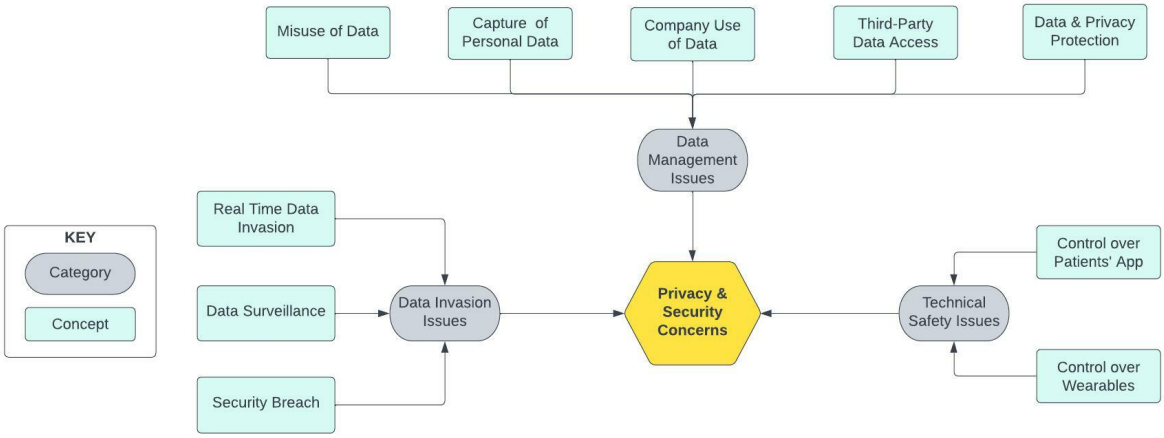

Figure 8. Summary of Concepts and Categories

Data Management Issues

Misuse of Data: There's nothing intrinsically wrong with the collection of data, but it's what happens to the information after it's been collected that should be a cause for concern. The inappropriate use of data through data harvesting was one of the themes that emerged from the data. One tweet highlighted the challenge of companies or their workers stealing and selling the data. It was mentioned that:

> *"Nothing stops Fitbit – or a crooked employee of Fitbit – from stealing that data and selling it. What would stop it is making the engineers who created Fitbit legally liable for misuse – would force them to design a Fitbit with inherent privacy security."*

Other users were concerned that if their data was being inappropriately harvested, they are generally not in the know, as to who the data may be sold to, and if that occurs how to opt out of using those mHealth wearables. It was shared that users need to know if the *"harvested*

data will pseudonymised for selling", or *"who will it be sold to and how to request your right to be removed"*. The concern was further amplified when data such as blood pressure and heart rate are captured by smartwatches, and the possible harm that can be caused from the *"misuse of this data or the inappropriate use of the data."* Another user exclaimed that *"other people do not deserve that much data about me."*

Capture of Personal Data: Although organizations are generally required to obtain meaningful consent from users for the collection, use, and disclosure of information, several users shared the apprehension when using these mHealth wearables and their personal data is being captured. It was intimated by one user that *"a reason to avoid wearing any tech wearables is because they collect and send data about your body all the time without your conscious consent."* In the domain of implantable medical monitoring devices, one user disclosed that *"they send info about you automatically without consent to your doctor."* From all indications users are generally seeking for *"people to be given control and consent over their wearable data."* Conversely, some mHealth devices like Fitbit may grant or provide some information on how the collected data is managed after consent;, at least one user expressed the following:

> *"Sure, I know that many users will just click through. Ironically, Fitbit gives Ok-ish info in how they manage data after consent has been given. But good luck finding info on what data they process and how long they store it. It's hard to tell."*

Furthermore, since *"many wearable devices seem to be connected to applications that capture data that people could use to monitor aspects of your daily life"*, the concern of granting users' proper consent for information gathering was expressed by many users.

Company Use of Data: mHealth users are primarily concerned about the ability of people or entities to both see and use the data being captured by these wearable devices. In one instance, it was shared that *"Fitbit is just another acquisition that will give Google access to hugely valuable sensitive data about us."* This concern was further amplified when it was shared that *"every footstep, heartbeat and location (every scrap of personal data that a #Fitbit device harvested from its owner's human body) is now the commercial property of Alphabet/#Google... and they do whatever they want with it..."* Additionally, while users are fearful about what these companies may choose to do with their data, there was a general concern about how these companies will use their data for financial gains at their expense, for example *"It's not Google ads that I'm worried about – it's what happens with our private health data, making (more) money for Google wasn't what I signed up for."* Further, it is *clearly "worrisome for personal data being used/sold" especially to third parties, with the trepidation that "they probably steal your data/location & sell to companies"* or *"sells all your personal health and movement data like a fitbit?"* Even though these companies promise not to use the personal data of the users for advertising, the concern of having *"unskippable YouTube ads"* was expressed by several users.

Third-Party Data Access: While there are benefits to be derived from using third-party applications that are integrated with mHealth wearables, users generally were concerned about how much of their data were being collected and inappropriately used by these entities. One user expressed fear that their data was being leaked via the third-party data access as gleaned from this post *"please help me understand how is my data with a third party and how should I trust your brand? This is a point of concern. So please look into the data leakage."* The apprehension of one user was evident, and it was basically voiced that *"it's not about*

wearable tech, it's about third-party data access." One recommendation was given, that unknown third-party apps should never be downloaded from the app store as these could further compromise the data storage and access. Another challenge is when data processors do not adequately de-identify the data, and third-party apps are able to easily re-identify the data for their purposes. It was quite worrisome for some when it was apparent that consumer data from wearable devices were being sent to various third-party companies. For example: *"we are worried about third party companies taking our healthy data,"* as it seems that *"most wearables have an open API that lets third part somewhat suck out our data."* It was vented that mHealth apps were seemingly sending unencrypted data and storing it on third-party servers.

Data & Privacy Protection: Users expressed concerns about their data and privacy protection, and about the associated policies and regulations to ensure the protection of their personal data. One user shared that *"data protection is definitely key to mHealth"* and that *"wearable tech is seen as the new data protection conundrum."* The major concern in this emerged theme, was that in the realm of data and privacy protection, regulations are still unclear, and all-encompassing policies should be created to enable public trust and confidence in these systems. One user stated that "*If they record any private data, image or audio, wearables should come with privacy and security policies and mechanisms."* Furthermore, since wearables generally collect a lot of data about the users, it was consistently presented that the devices and applications must come with security and privacy protections and policies. One user shared that *"data protection, personal privacy and security issues need to be addressed before potential benefits from integrating secondary data from health apps into healthcare can be realized."* The policies that govern how third party applications access and

use data, and governments and other agencies that readily access the data must be clearly delineated. It was posited that this concern can be seen as the *"the tip of the iceberg, with issues compounded by lack of privacy protection under 3rd party, where govt can access wearable data."*

Data Invasion Issues

Real Time Data Invasion: A major factor that inhibits the users from utilizing mHealth wearables stems from the presumed invasion of privacy of real time data, which involves the unjustifiable intrusion into the personal data collected without consent. It is apparent from the data that users are overly concerned about the technology invasion that takes place in their personal life for example: *"Wearable technology, how safe is it? Not sure. Only time can tell you. So much of technology invasion in my personal life."* There are organizations that choose to use the devices to track employees health, but a few posts indicated that it *"feels a bit like invasion of privacy and too much control to the employer."* Other users expressed the trepidation that some of these wearables require their location setting to be enabled, which for them constitute an invasion of privacy, for example: *"Why do yall need my location to sync my fitbit? Seems like an invasion of privacy".* Additionally, one user postulated *"No Fitbit, you are being returned. Forcing me to allow you to track my personal data and report it to the US government before I can even setup my watch means invasion of privacy. It's going back."*

Security Breach: Security breach generally represents any incident that results in unauthorized access to data, applications, networks, or devices. That is, the unauthorized access of information by intruders who are able to circumvent installed security measures. mHealth wearables data breaches and malicious attacks are obvious disquiets for users. In one

instance, one user shared the concern that *"tough times for #digital health tech – Fitbit plunges after letter, under Armour's MyFitnessPal suffers 150m ppl data breach"* and *"Cloudflare security breach exposes data from Fitbit"*. With all the personal data being captured and stored by these devices*, "Wearables like smartwatches and smart thermometers could increase the risk of data security breach. Healthcare providers need to be aware of these issues in order to comply with HIPAA"*.

Other security vulnerabilities expressed involved a user whose account was hacked on surge Fitbit, and the concern raised about the time it takes to have answers to these high priority security and data breaches by the affected companies. Some users were very scared about the prospects of a malicious attacks impacting their wearables, for example *"@Garmin's current ransomware outage brings back old thoughts. What happens when a malicious actor takes control of GPS? So many things rely on it nowadays, that is quite scary."* A similar sentiment about security vulnerabilities was presented by another user, who wrote: *"@Garmin Not certain that encryption categorically means that they cannot access sensitive information, like payment. I would like a clear and open statement on the reasons behind the "outage" and its impact on data security."*

Data Surveillance: The expressed concerns of surveillance or location tracking featured prominently in the data. Surveillance normally is a form of monitoring, which is performed to obtain specific data without the users' knowledge. The tracking mechanisms employed in these wearables or applications give rise to privacy concerns. A common sentiment outlined that *"#Fitbit is surveillance"* and that *"all wearables and connected clothing are hackable & surveillance vulnerable."* A major concern revolving around tracking mechanisms is when *"tech companies roll out wearable child tracking devices to*

'normalize' intrusive surveillance for new generation". Many users were concerned about wearables such as Apple Watch as it is seen as an *"obtrusive, stealthy piece of mass surveillance"*, which is probably a device used by different agencies. Even when users restrict certain activities on their wearables or applications, there is still the concern that some other tracking mechanisms may have been embedded in this system, for example: *"I have opted out of sharing my activity data, but perhaps there's a shared tracking cookie that could be leaking my location"*.

There is also the fear of invasive surveillance projects in which *"wearable tracking devices which are termed as a patient tracking tool but will be used for national security purposes."* In one country, smartwatches were banned for children because of privacy reasons, as the *"watches can be hacked to change tracking location."* Additionally, wearables such as pedometers may have some hidden tracking mechanisms, and one tweet highlighted that *"some pedometer devices also use location tracking feeding your running route to website."* Clearly, even with the collected data being anonymized, it has been demonstrated that such data can be compromised.

Technical Safety Issues

Control Over Wearables: Users are concerned that they may not have the requisite control over their wearables. In one instance it was clearly expressed that certain permissions were inactive, and there were seemingly no settings or options to *"address the selective shutting of access and control of my device."* One tweet outlined that "*people with implants or wearable are vulnerable to code acting against their interests. They need security and control*." The issue of not having exclusive dynamic control over wearables have caused users to experience apprehension. For example: *"If you go the Fitbit/Fitness tracker route you will*

have some privacy but not exclusive control." This claim was supported by another tweet, which stated that *"you may be the person wearing a given piece of wearable technology, but that doesn't mean you control where the data goes or how it's used.*"

Control over Patients' Apps: The ultimate question arising from a closer examination of the data is *"who has the nexus of control over the patient's data, the patient or the provider?"* One user expressed that *"only patients should own their data and control who get to see, what and how long. #mhealthapp."* Therefore, companies should not exploit the privacy and control needed by users, as suggested by one user: *"I think if companies don't respect data collected, then wearables and health apps will face a backlash soon, privacy and control are important."* Consequently, as more individuals use health apps, there are questions about what happens to that sensitive data and how much control users have over these.

Emotion and Sentiment Analyses

The emotion analysis presented in figure 9 is showing that 56% of the posts were depicting anger, 15% portraying fear, and 6% and 2% showing levels of disgust and fear respectively, which emphasizes the attitude of the users of these healthcare wearable technologies. For example: *"All of your location and fitness data just got acquired by the world's largest surveillance company and there's nothing you can do about it. How do you feel about breaking up some of these companies now?"*. Additionally, 21% of the posts were expressing joyful emotions, which suggests that users were uncertain about privacy and security concerns, albeit mentions of privacy or security concerns being made in their posts. For example: *"Great data! I always saw a future for people sharing their biometric data via wearables socially. Just be mindful that it can be interpreted as private health info too. Data privacy aside, awesome stuff!"*

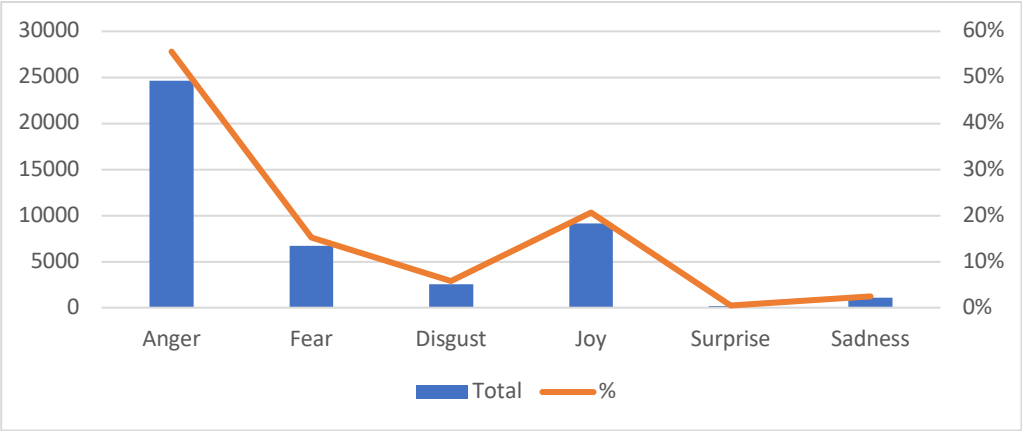

Figure 9. Emotion Analysis

Furthermore, in Figure 10 the sentiment analysis demonstrates that 74% of the posts were categorized as a negative sentiment, whereas 26% were positive, which indicates that users are ambivalent concerning privacy and security, notwithstanding mentions of privacy or security issues in their posts, there was a general positive tone. This behaviour is in line with the Privacy Calculus theory, where individuals always rationally weigh the potential benefits and potential risks of data disclosure decisions (Culnan & Armstrong, 1999). For example:

> *I have great hopes for IoT. A shirt you wear that monitors your BP, heart rate, temp, heart rhythm, exertion, pulse, etc and feeds that to your smart watch to keep you safe. And if you have an AFIB moment while driving, it can alert EMS. Hard to secure - but the value is high.*

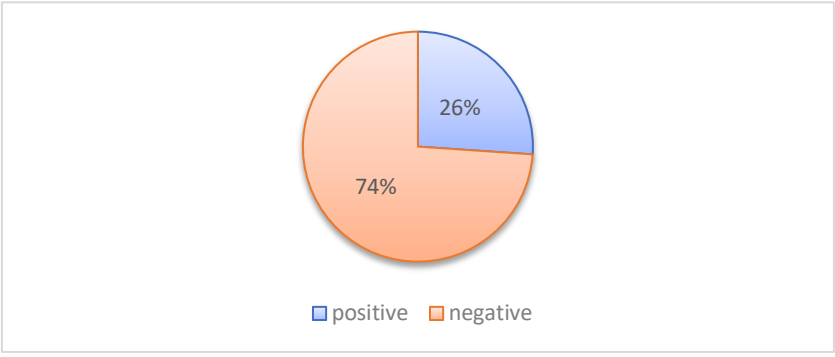

Figure 10. Sentiments Analysis

Evolution of issues relating to mHealth Wearables

In answering the research question on how has the perception of various mHealth related issues evolved over time, the Brightview supervised algorithm based on aggregate analysis was used to manually code a training set of documents into a predefined set of groups generated from the LDA model. We assigned at least 20 tweets into each category (see Codebook in Appendix A), after which the algorithm was executed on past and future tweets returned by the search query. The tweets were examined based on the assigned categories, and further training was conducted where necessary.

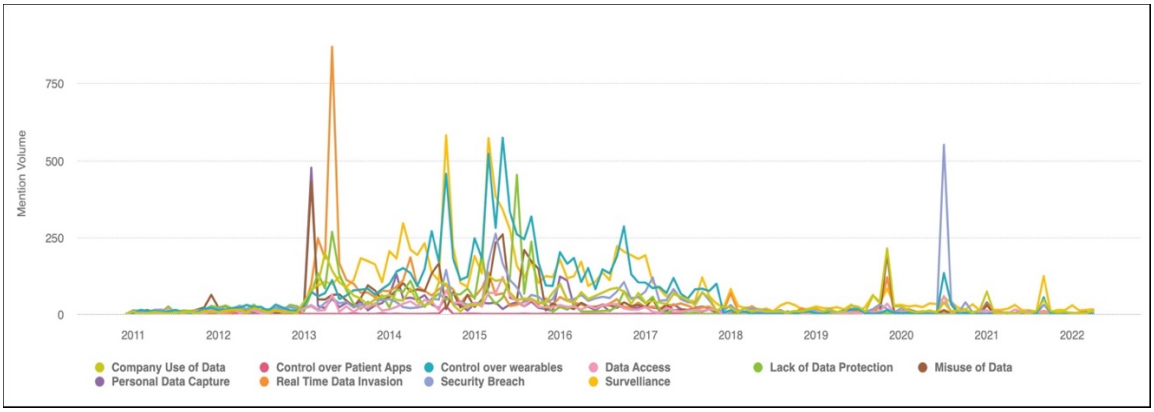

Figure 11. Posts made over time based on different categories

Figure 11 shows the volume of tweets over time by category. The period 2013 to 2017 had several posts being made especially in concerned areas of surveillance, real time data invasion, control over wearables, security breach, misuse of data, and lack of data protection. The evolution of healthcare wearables issues being discussed on Twitter is not as prolific in terms of number of posts from 2018 to 2022, but shows that mHealth users are still concerned about areas such as security breach, real time data invasion, surveillance and how these companies are actually using the data collected from these devices.

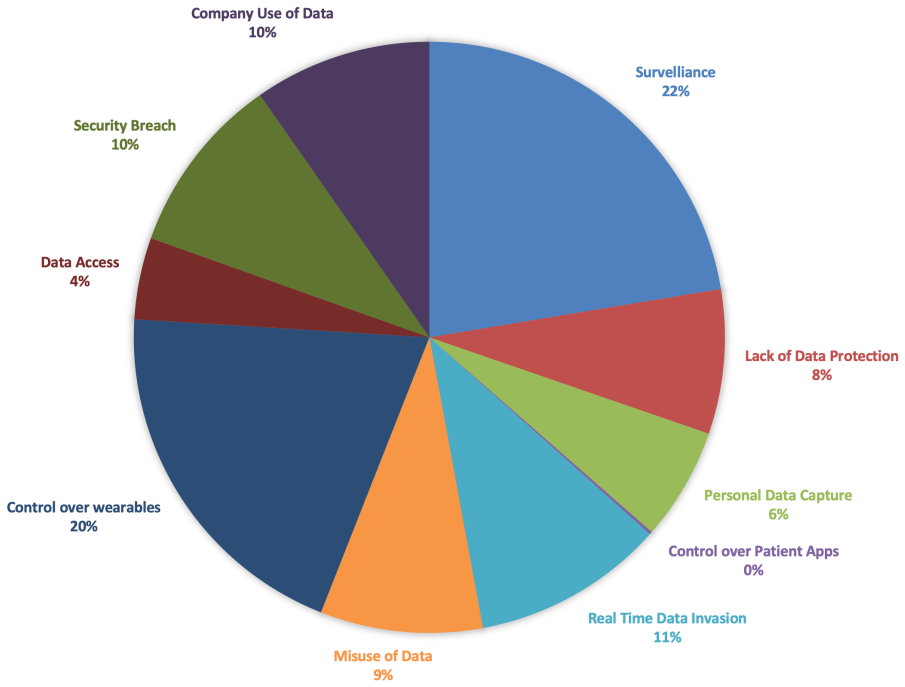

Figure 12. Evolution of mHealth security and privacy issues

In figure 12, it shows the distribution of mHealth security and privacy issues being discussed over the period under examination. The results show that 22% of the posts were related to surveillance issues being expressed by the users, while 20% were concerned about how much control they truly have over these wearable devices. Further, the concern about the invasion of real time data accounted for 11% of the posts, while security breach and company use of the data were represented by 10% each. Other issues included the misuse of data, 9%, lack of data protection, 8%, personal data capture, 6%, and data access, 4%. The results also show that users seemingly were not having major issues with the control over patients apps, as less than 1% of the posts were related to that concern.

CHAPTER 5

EMERGENT THEORETICAL MODEL

This chapter provides support for the different categories and concepts which emerged during the analysis of the data. Each proposition in the model is discussed and confirmed by sample data, which demonstrates their emergent nature, which is followed by a presentation of emergent theoretical model. This section ends with a discussion, which compares the findings of the model with extant literature.

Overview

The Data Management, Data Invasion, and Technical Safety (D-MIT) Emergent Theoretical Model represent several mHealth user concerns which may inhibit the adoption of these wearables. The primary concepts that describe the user expressed concerns are listed as, data management (P1), data invasion (P2), technical safety (P3). Within each primary concept are the supporting elements alphabetically labelled with each conceptual category. All the propositions are theorized to inhibit the adoption or use of mHealth wearables. We will now discuss each proposition in the model and relate it to the data demonstrating their emergent nature.

Data management issues impact the acceptance of mHealth devices (P1)

The underlying concepts developed in the Findings section contribute to the theoretical category which highlights how the data management issues emerged from the core concepts as shown in Figure 13.

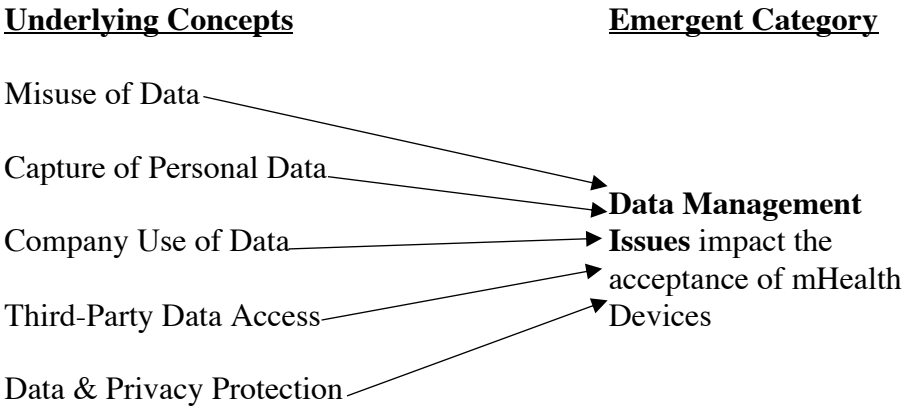

Figure 13. Emergence of data management category

Figure 14 shows our representation of the category *Data Management Issues* and its constituents as it emerges from the data. The *cause* section of the diagram represents constructs in the emerging theoretical model.

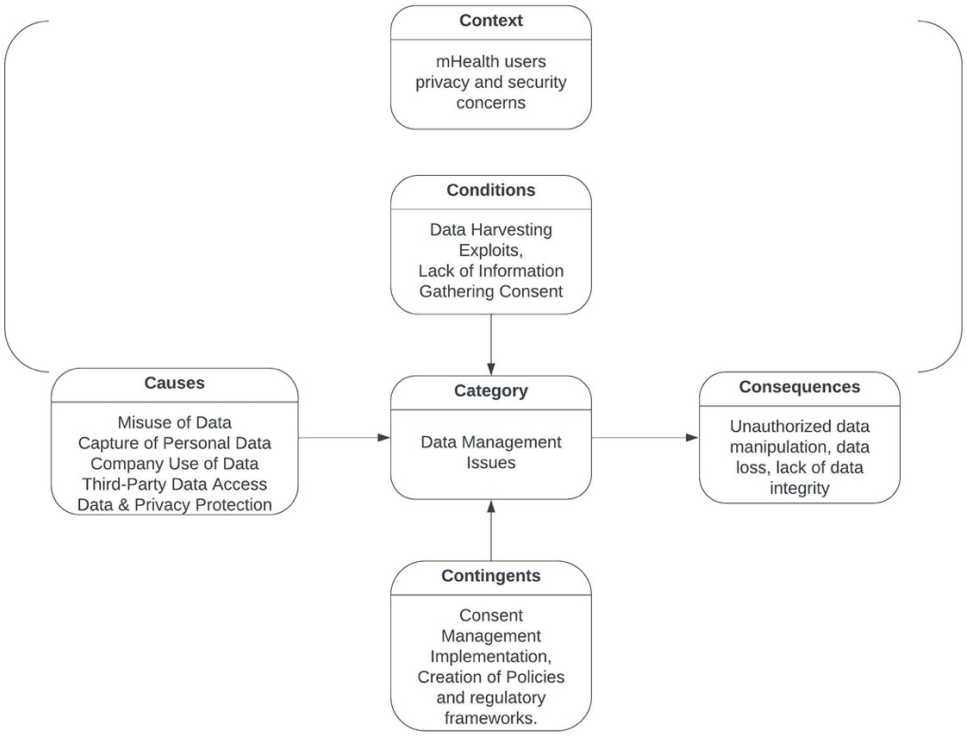

Figure 14. The category Data Management Issues

From the analysis of the data, two *conditions* that contribute to the data management issues emerged, namely, *data harvesting exploits and lack of information gathering consent*. It was apparent that some users felt that their data were being slyly harvested by companies, sometimes under the guise of conducting a survey or the *"unauthorized mining of personal health data"*. With the notion of having mHealth data kept private, some were concerned that *"the dark market in brokerage of personal biometric & behavioral data were being harvested from #Fitbit devices."* Furthermore, it was opined that companies may also be harvesting the intimate personal health data of individuals, and referencing them with all other data they have collected, for exploits unknown to the users. For example, *"I'm super bummed that they will be getting their grubby little hands on my personal wellness data. Are they going to share it with @Facebook & any data-mining service that pays them a pretty penny too? Ugh. Sadness."*

The *lack of information gathering consent* served as another condition for this issue of data management to occur. This is where users felt that the current notice and consent models used on these devices are flawed to the point of being meaningless. For example, *"Personal health information can be used, shared, or sold, without consent. Consumers have no control over who can access their health data."* While mHealth wearables are seen as advantageous, when neither permission nor consent is granted for the sharing of private information, that leads to data management issues. In other words, users are desirous of granting *"informed and explicit"* consent for the collection and use of the personal data generated by mHealth wearables; otherwise, this lack of consent can be seen as *"unethical practices to store and use my data."*

Five *causes* that lead to issues relating to data management were *misuse of data, capture of personal data, company use of data, third-party data access, and data & privacy protection* which represent the constructs in the emerging theoretical model. In the Data Management issues group of propositions, the *misuse of data inhibits the adoption of mHealth wearables P1(a)* was mentioned in several tweets. It was stated in one instance that based on the wealth of data collected by these mHealth devices, it is pertinent that the *"data is kept private and safe"* so as to minimize the misuse of the data collected. Users were willing to discontinue the use of certain devices as they were perturbed by the mandatory harvesting and use of contacts when all that was needed was a simple step counter.

The discussions also highlight the point that many users *"don't think they should sell your data"* particularly those related to their personal data. Furthermore, comments indicate that users were agitated about what may happen if the data acquired via these devices were *"shared with Facebook and any data-mining services"*. Therefore, the result of the misuse of data inhibits the adoption of mHealth wearables shows that users generally prefer to grant appropriate consent for the data that is to be shared, rather than companies or other third-parties simply accessing and utilizing at will. Figure 15 shows an example tweet of how the theoretical construct emerged through the different GT phases (open, axial, and selective).

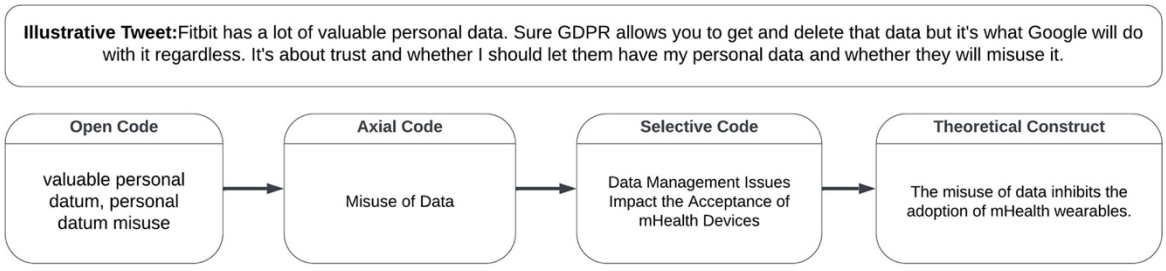

Figure 15. Emergence of "misuse of data" construct

Another aspect in the data management issues group was the sense that users were concerned that the personal data being captured about them were being done without their direct approval, as such proposition P1(b), *the capture of personal data must be consensual*. The most common tweet intimated that these devices and by extension the companies were tracking and capturing a plethora of personal data without the users' consent, which was seen as a major threat to privacy. Obviously, users want to be able to grant appropriate consent for data capture, especially when personal data such as heart rate, sleep patterns, workouts, etc. are being collected and used by companies such as Google.

With data privacy being foremost in the minds of many of the users, some were concerned that wearables were capturing data on a permanent basis, and not knowing what else the sensors can detect and collect, supports the argument about users being able to grant consent so as to allay their fears. Interestingly, many deemed wearable data as *"intensely personal and intimate"* and suggested that *"good privacy practices are the gateway to trust."* Further, users wanted reassurance from organizations such as Fitbit that the captured data would not be combined with other data sources without informed consent. Therefore, it is imperative that opportunities be provided for new standards of user-centered informed consent to be available in the age of mHealth. Additionally, some users have suggested that new legislations should be enacted to ensure that *"data collected through fitness trackers, smartwatches, health apps cannot be sold or shared without consumer consent."* An illustration of how the theoretical construct emerged from an example tweet is shown in Figure 16.

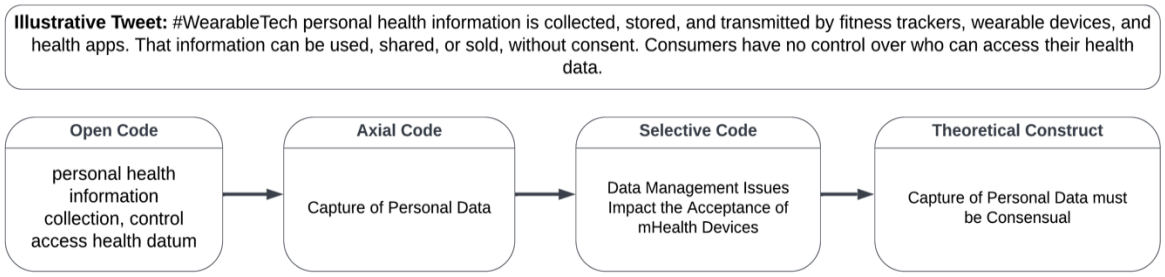

Figure 16. Emergence of "Capture of Personal Data must be consensual" construct

In the Data Management issues group of propositions, the concern about how mHealth device companies are using the data collected from users, gave rise to proposition P1(c), the inappropriate use of data by companies inhibits the adoption of mHealth wearables. When Fitbit was bought by Google, many users became apprehensive not knowing how the data would not be used, and many posted that they wanted a way to recover their data from Fitbit so that it becomes only available to them. It was apparent that many users viewed the merger as an avenue through which their health/wellness data would be monetized, and possibly accessed by third-party applications. Many users were reluctant and opted to dispose of their devices, as in some instances *"it wouldn't let them see the data on phone without uploading to their server."* This concern was further compounded as *"every scrap of personal data that a #Fitbit device harvested from its owner's body"* was now considered the commercial property of Google, and users were uneasy, knowing that they would do whatever they want with the data. The disquiet was also seen across several tweets, as the notion of companies selling *"personal health and movement data"* impacted the trust exhibited by the users. Figure 17 shows an example tweet of how the theoretical construct emerged through the different GT phases (open, axial, and selective).

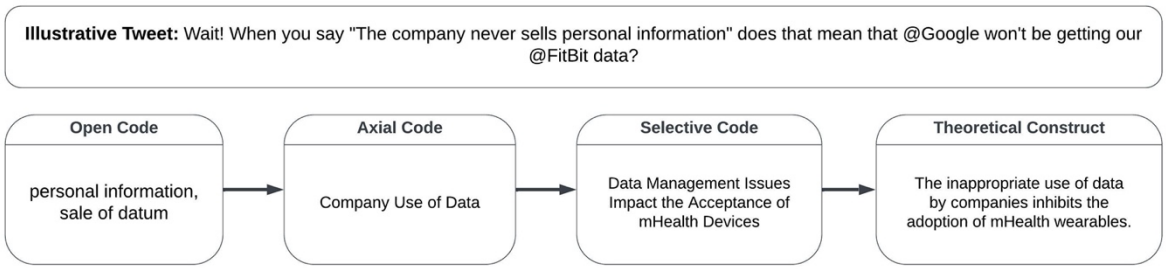

Figure 17. Emergence of "company use of data" construct

In the Data Management issues group of propositions, the concern about how third-parties access the personal data of users, even without the requisite permissions, gave rise to the proposition P1(d), *illegal third-party data access inhibits the adoption of mHealth wearables.* When users give permission to third-party apps they are often unaware of the amount data they give to these entities. On the other hand, these third-parties may act unknowingly and access things such as contacts, browsing history, and other personal information. For users it was necessary to know *"what are the third party data protection commitments."* Further, *"it's not about losing recorded data, it's more whether the data is in the hands of a third party,"* was a popular sentiment expressed. Some felt safer by simply discarding the wearables or uninstalling third-party apps, as the belief was that there would be *"no more third party spying from these companies."* So it is disconcerting for users to have a mHealth that collects and stores personal information, and then having this data access through third-party applications illegally. Figure 18 shows an example tweet of how the theoretical construct emerged through the different GT phases (open, axial, and selective).

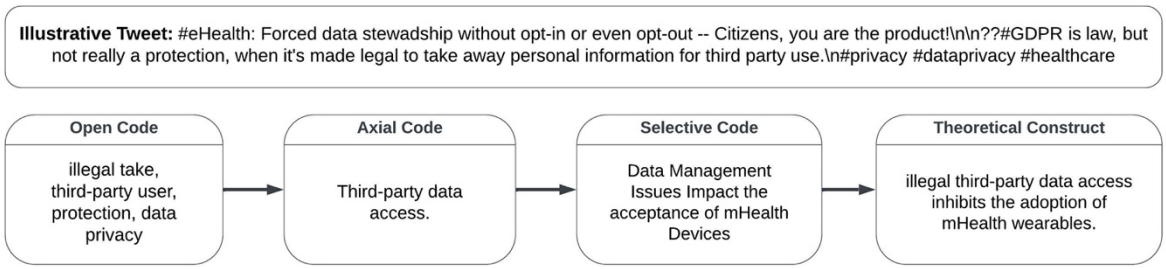

Figure 18. Emergence of "third-party data access" Construct

In the Data Management issues group of propositions, the proposition, the *lack of data & privacy protection inhibits the use of mHealth wearables P1(e)* was mentioned by multiple users, since wearable technologies of all categories raise a broad range of probable legal and policy issues. Users felt that the needed regulations were either not clear or non-existent and one person shared that *"personal data from #fitbit and other user public data like weather need regulatory frames."* Others were willing to seek other wearable options, as it was lamented that *"regulation of data protection"* was needed, since technology alone would fail. It was also felt that there was a "*lack of #eHealth legislation framework"* especially as there is an increase in cybercrimes, the lack of data protection policies was concerning. It was expressed that while some companies, like Apple has a track record on personal data protection, there was a fear factor with other companies like Amazon or Google. Users generally wanted to know that *"#mHealth tracking stays voluntary and apps policies are clear"* with the requisite mechanisms in place to cover data protection and security. Otherwise, it was felt that there are inconsistent privacy policies in mHealth apps, where it is perceived that companies can simply *"collect and sell your personal data for profits."* Surprisingly, the notion of having wearable technology to track data, felt for some an invasion of privacy, but it was believed that *"HIPPA already protects me and my data and I choose*

who sees it." Overall, users are more concerned when there are no existing policies to safeguard the collection, storage, dissemination, and transfer of data, especially when third-party applications become involved. Figure 19 shows an example of how the theoretical construct emerged from an example tweet through the open, axial, and selective coding phases resulting in the theoretical construct P1(e).

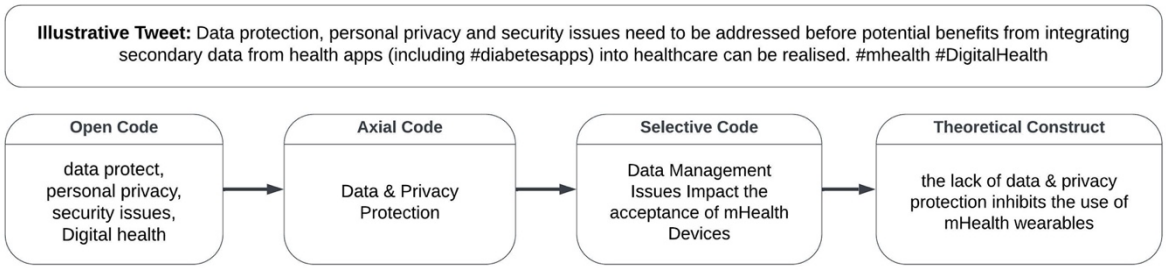

Figure 19. Emergence of "Data & Privacy Protection" Construct

Our research exposes several data management consequences, to include unauthorized data manipulation, data loss, and lack of data integrity. Several users felt that when their private health data lands in the possession of certain companies, *"they'll use this data to manipulate us"*. Data manipulation on *"health, nutrition, what we buy, where we go & what we do."* This resulted in some users feeling like *"a puppet being manipulated."* Further, users expressed that loss of personal data was a concern or outcome of the data management issues mentioned above. In one instance it was highlighted that not having access to wearables for a protracted period due to a systems challenge, resulted in a *"catastrophic data loss"* which caused a loss of privacy.

Integrity is an important security requirement for information systems especially for wearable systems where collected data is usually sensitive and private. It is important to make sure that data is not altered in transit and being received by authorised parties only. While

some companies seek to address the data integrity challenge, many users underscore the dilemma of wearable data integrity issues. For example *"there is a reason why I have always been wary of things like Fitbit and other tech tied to your body and information collection. I never trusted the integrity of the data."*

An interesting strategy gleaned from posts made by users to address the data management issues by mediating between the aforementioned causes and consequences, is the need *for proper consent management features*. It was clear that users wanted to be in control and grant appropriate consent for which data can be accessed and harvested. In one instance is was presented that *"every time a user interacts with the wearable app, the information is collected and stored & tied to a device ID & location."* This meant that the users wanted to provide the requisite consent to not only protect their data, but also determine which external entities the collected data are shared with. Furthermore, the development of *legal policies and regulatory frameworks is deemed important*. In one instance, after examining the security measures necessary, one user called for the "*creation of policies to minimize security risks."* Further, the point was made that *"some wearable apps don't even have privacy policies."* This therefore supports the strategy of users wanting clear and concise policies to guide their data collection, storage, and dissemination to alleviate their legal & policy issues.

Data Invasion Issues impact the acceptance of mHealth devices (P2)

The underlying concepts developed in the Findings section contribute to the theoretical category which highlights how the data management issues emerged from the core concepts as shown in Figure 20.

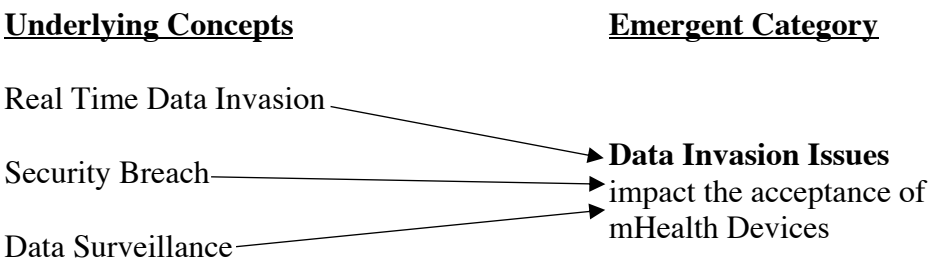

Figure 20. Emergence of Data Invasion Category

Figure 21 shows our representation of the category *Data Invasion Issues* and its constituents as it emerges from the data. The *cause* section of the diagram represents constructs in the emerging theoretical model.

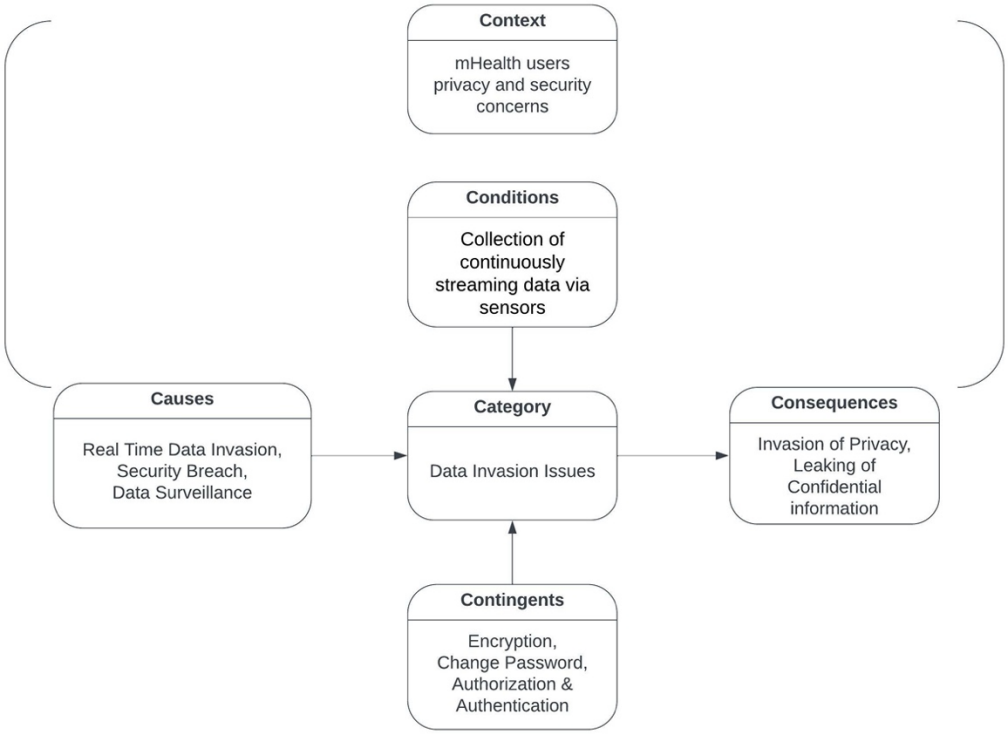

Figure 21. The category Data Invasion Issues

From the analysis of the data, one *condition* that contributes to the data invasion issues associated with mHealth wearables emerged, namely, the *collection of continuously streaming data via sensors*. mHealth wearable sensors are increasingly employed to monitor patient health, rapidly assist with disease diagnosis, and help predict and often improve patient

outcomes. However, the wide range of data collected through these sensors gives rise to real-time data invasion, security breach, and data surveillance issues. For instance, even though mHealth wearables and sensors open up an entirely new field of possibilities for data collection, it was posited that the major challenge with the data being collected via these many sensors is that of privacy and what is going to be done with the data.

Three *causes* that lead to issues relating to data invasions were *real-time data invasion, security breach, and data surveillance* which represent the constructs in the emerging theoretical model. In the Data Invasion issues group of propositions, *the invasion of real-time data inhibits the use of mHealth wearables P3(a)*. Most of these devices are continuously capturing the location and other real-time data, where users felt it was an invasion of their privacy to be demanding that the location-feature be continuously enabled in order to benefit from some of the features. Some felt that health insurance companies could invade real-time data such as the total number of hours slept, which may see them *"use the data against users to raise rates."* It was evident that the continuous tracking of location by Fitbit, was deemed a *"huge breach of privacy"*, especially when the users perceived that their personal data was being shared without permission. Further, some users would prefer to discontinue the use of these devices as they felt for example, that *"when you are wearing a Fitbit and you are breathing and sleeping data is on the internet"* which broadens the concern of the invasion of real-time data. Figure 22 shows an example tweet of how the theoretical construct emerged through the different GT phases (open, axial, and selective).

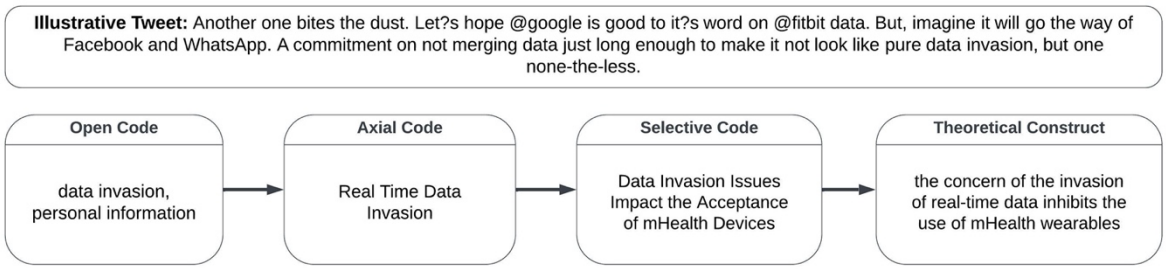

Figure 22. Emergence of "invasion of real-time data" construct

Furthermore, in the Data Invasion issues group of propositions, the concern associated with the unauthorized access to data, applications or devices, gave rise to proposition P3(b), *security breaches inhibits the adoption of mHealth wearables.* Many wearable devices store data in local storage without encryption or data protection. Accordingly, there could be a high risk of losing confidential and personal health data. The need for the mHealth industry to *"work overdrive to ensure data security as the applications grow"* is an opportunity for the confidence levels of users to increase. Due to the nature of these mHealth wearable devices they present "data breach concern with hackers gaining confidential information."

Some users cited a data breach which occurred in a particular company, and lamented the *"mental anguish over not protecting their personal information."* Several users were generally concerned that wearables are prone to security breaches, and also the ease with which, *"personal information could easily be made public due to hackers."* Further, based on the design of devices such as smartwatches, there are increased risks associated with data security breach. Users were not willing to take any chances when they were alerted about security breaches on things such as password, and many intimated they immediately changed devices or stopped their usage. This occurs because the users are unaware if their personal data has been stolen or lost. Additionally, some users felt that *"wearables are most likely to*

be the source of a security breach among Internet of Things devices." Figure 23 shows an example tweet of how the theoretical construct emerged through the different GT phases (open, axial, and selective).

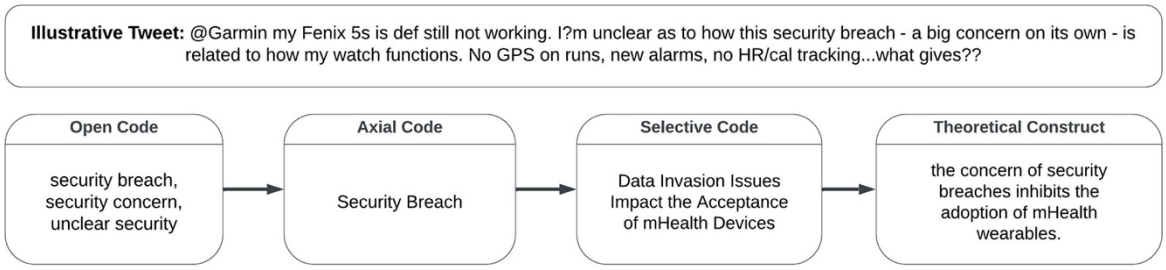

Figure 23. Emergence of "security breach" construct

Finally, in the Data Invasion issues group of propositions, the concern associated with the monitoring of data associated with mHealth wearables without the users' knowledge gave rise to proposition P3(c), *data surveillance inhibits the adoption of mHealth wearables.* Users were overly concerned about the fact that *"Google and Apple monitor you through your devices."* This concern was further highlighted as users believed that these companies were using security purposes as an excuse to be monitoring these devices, which is tantamount to some form of extensive tracking. Ironically, some understood the Apple Watch and Fitbit to be examples of luxury surveillance and celebrated the benefits of tracking and monitoring.

However, others expressed the views that, *"surveillance and exploitation of personal data are an unavoidable reality at this point but I want more people to know that you personally can resist quantifying yourself."* In another instance, it was stated that *"my fitbit as just another surveillance technology from the big tech oligarchs. I'm disgusted."* From all indications the user perception about being under constant surveillance, in which privacy is almost entirely eradicated, surely will have a negative impact on the acceptance of these devices. Since, it is *"inevitable that all trackers eventually default to surveillance*

capitalism." Figure 24 shows an example tweet of how the theoretical construct emerged through the different GT phases (open, axial, and selective).

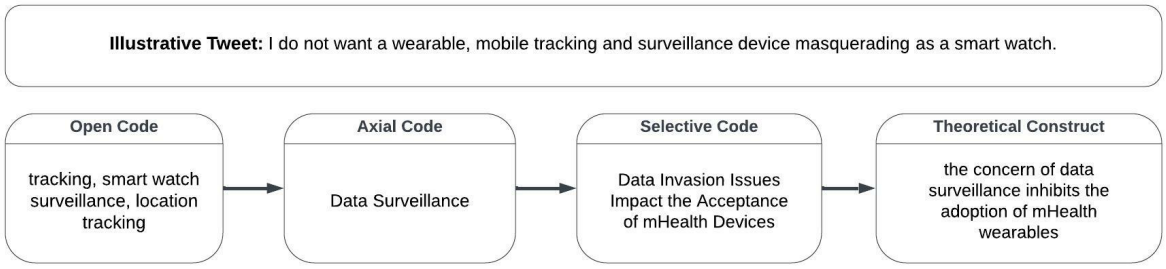

Figure 24. Emergence of "data surveillance" construct

Our research reveals two data invasion *consequences*, to include leaking of confidential information and invasion of privacy. The leaking of confidential data can cause emotional and financial loss. This was highlighted by one user who shared that "*you may want to think twice before buying a #smartwatch as these gadgets can leak your data to #hackers."* Several users felt that in the face of a security breach, their personal data could be leaked, and that was most disturbing to them. For example: *"a potential leak of our personal information including address, health, and location data in addition to an expensive device not working is definitely something to rage about."* The next outcome of data invasion issues, stems from the belief that the users' privacy is being invaded with the notion that these devices *"will track us literally all day and night."* For example: *"I've been interested in a Fitbit, but I do not want to subject myself to the invasion of privacy."*

The *contingents* to address these data invasion issues, include the use of encryption, password management, and authorization techniques. Encryption generally involves the conversion of data into a format that prevents unauthorized access. Users wanted to be sure for privacy and cybersecurity reasons that companies like Fitbit were *"encrypting both the storage and transmission of their personal data."* This is a desired feature to help in allaying

the concerns of users of mHealth wearables. In another instance, users felt that if a security breach exposes their personal data, then *"password changes are recommended".* Some of the issues were addressed when users did a reset of their passwords. It was apparent that frequent password changes and the use of strong passwords served as a means of making these wearables more robust. Due to the *"unauthorized mining of personal data"* users felt that proper authorization facilities should be available, so as to ensure that entities cannot simply access the different data resources.

Technical Safety Issues impact the acceptance of mHealth devices (P3)

Figure 25 shows how the underlying concepts developed in the Findings section above contribute to the theoretical category which highlights the emergence of the technical safety issues category from the underlying concepts.

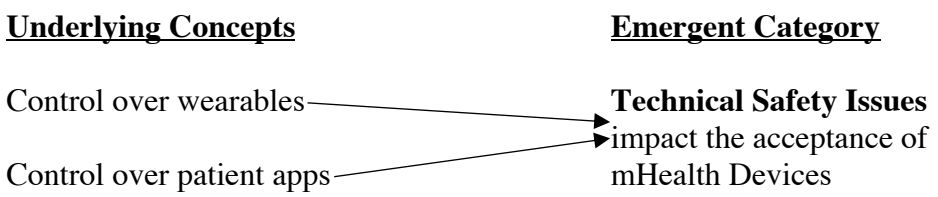

Figure 25. Emergence of Technical Safety Category

Figure 26 shows our representation of the category *Technical Safety Issues* and its constituents as it emerges from the data. The *cause* section of the diagram represents constructs in the emerging theoretical model.

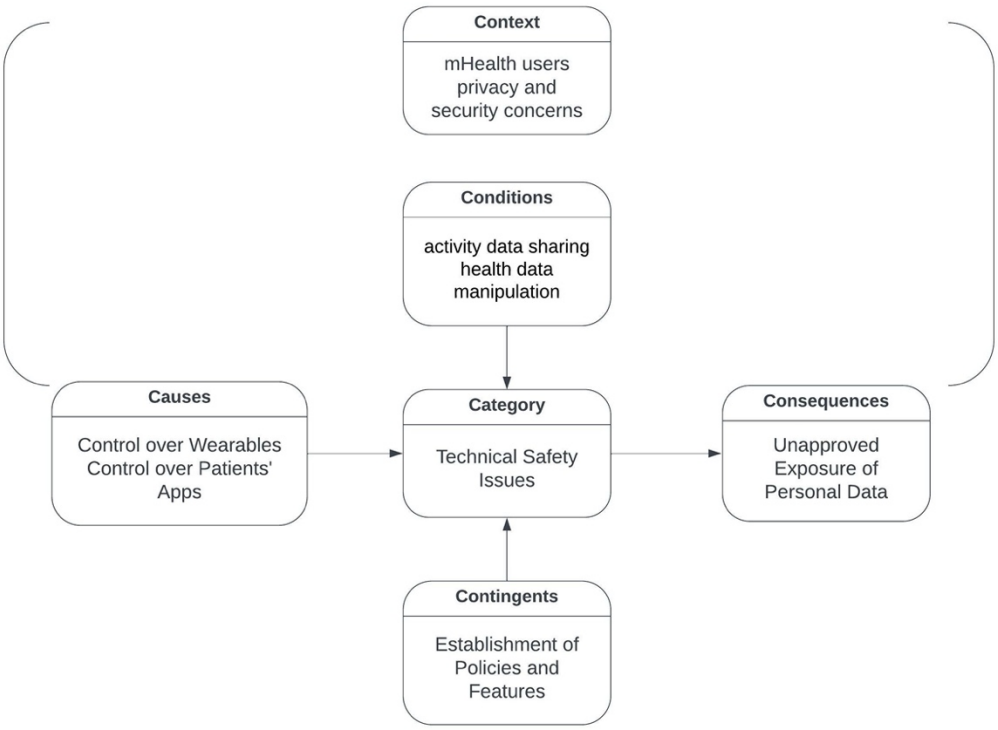

Figure 26. The category Technical Safety Issues

From the analysis of the data, two *conditions* that contribute to the technical safety issues emerged, namely, *activity data sharing* and *health data manipulation*. Based on the many sensors associated with mHealth wearables, there are a plethora of data sharing activities which are taking place, which require the user to have control over these devices. Additionally, in the context of mHealth applications, the condition that triggers this cause occurs during the different stages of the manipulation of health data.

In the technical safety issues users were concerned that they did not have control over their wearables, data permissions, and the flexibility to deactivate certain sensors; this gave rise to the proposition P2(a), *the lack of control over wearables inhibits their use.* In one instance, it was presented that even though the deactivation of activity data sharing was done, there was the concern that there might have been other features that would have made those

changes null and void. For example, *"I have opted out of sharing my activity data, but perhaps there's a shared tracking cookie that could be leaking my location."* It was intimated that *"putting control in the hands of the user is a basic human right. #wearables".* This suggests that users want to have more control over different facets of these devices, and others have decided to simply terminate their use based on this concern. It was also mentioned that users *"need to access their data & control who sees it #wearabletech".* Overall, users want to be able to absolutely control data access, activity sensors and other settings; otherwise it may deter their acceptance of mHealth wearables. Figure 27 shows an example tweet of how the theoretical construct emerged through the different GT phases (open, axial, and selective).

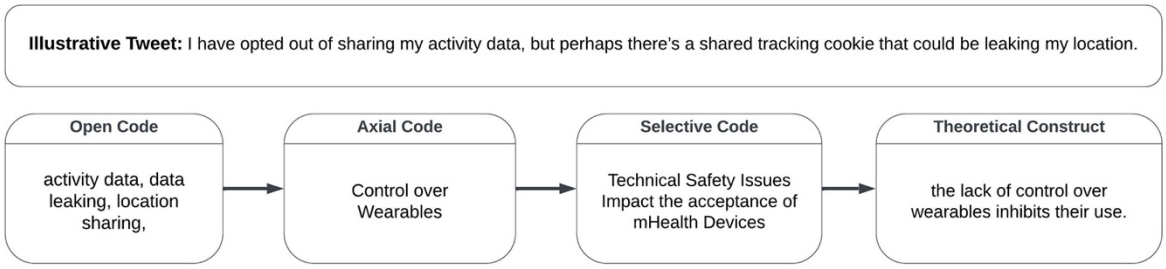

Figure 27. Emergence "Control over wearables" Construct

Furthermore, in the technical safety issues users were concerned that they did not have control over their patient mHealth applications, and the flexibility to cancel data collection and this gave rise to the proposition P2(b), *the lack of control over patient's mHealth application inhibits their acceptance.* Many users have resorted to the use of mHealth applications to manage their day-to-day medical experiences; however, it was obvious that users wanted to be able to enforce data restrictions as they feared that *"you Give Apps Sensitive Personal Information. Then They Tell Facebook." #healthdata #privacy."* Therefore, it was shared that *"apps must respect the user's permission settings and not attempt to manipulate, trick, or force people to consent to unnecessary data access."* Some

users have opted to uninstall these applications when they perceive they are not in control of especially data collection and use, and some were perturbed that *"when you install most apps, the EULA / contract you sign will allow them to pass all your personal data along to any company that buys them up. #Fitbit #Google #data #privacy."* Additionally, when users believe that their personal health data being captured on these mHealth applications are being exposed, and there is not much they can do about it, then their acceptance will be impeded. Figure 28 shows an example tweet of how the theoretical construct emerged through the different GT phases (open, axial, and selective).

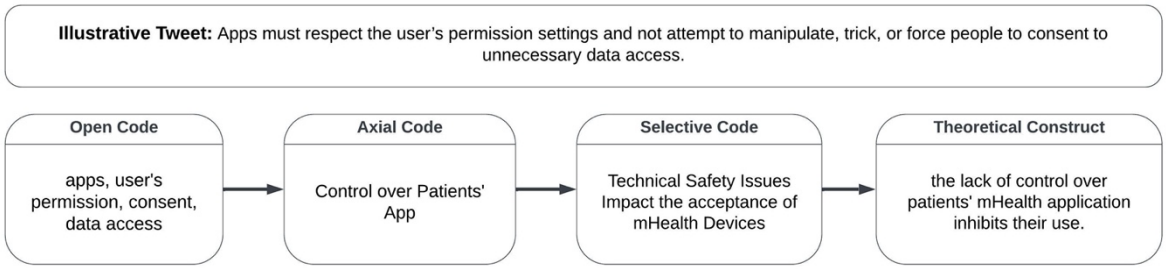

Figure 28. Emergence of "Control over Patients" App

With all the sensors available on mHealth wearables, and the plethora of data being transmitted via these mHealth applications, our research reveals one technical safety *consequence*, which involves the unapproved exposure of personal data. When a user does not have full control over determining what gets captured and what gets shared, they may end up with their data being exposed to even unauthorized third-party applications. For example, *"your Personal #HealthData Is Not #Safe: You go to the #doctor to get well, or check your #health. You don't expect the doctor's apps to expose your #privacy. But they do."* If companies do not respect data collected, *"then wearables/health apps will face a backlash soon"*. As such, *contingents* that can be employed to address this effect, involve the establishment of policies and features that will provide the users with greater control over

what gets activated and what gets turned off. This occurs because, *"some wearable apps don't even have privacy policies."*

The Theoretical Model

The D-MIT Emergent Theoretical model in Figure 29 communicates that mHealth users believe that there are several areas that inhibit their acceptance of mHealth wearables. That is, issues pertaining to their data management through the misuse of their data by entities such as wearable companies and other third-party applications negatively impacts their adoption of these devices. It also communicates the impact of data invasion and technical safety on the general use of mHealth wearables.

Discussion

The D-MIT Emergent Theoretical model shown in Figure 29, highlights three (3) abstracted categories, namely, Data Management, Data Invasion, and Technical Safety issues. Data Management issues encapsulates the misuse of data, the capture of personal data, the use of data by wearable companies, illegal third-party data access, and lack of data and privacy protection. A study conducted by de Arriba-Pérez et al. (2016) showed that users of healthcare wearables are worried that data that is harvested via the sensors available in these devices may be misused by different individuals. In addition, Abdolkhani et al. (2020) shared from their research, that users lament the lack of transparency on who owns and has access to the data, also, the lack of information gathering consent for continuous data collection and use. This concern is intensified since sensors in wearable devices allow the collection of a wide array of user data ubiquitously and unobtrusively on a continuum basis, and in most cases, without the explicit consent of the user.

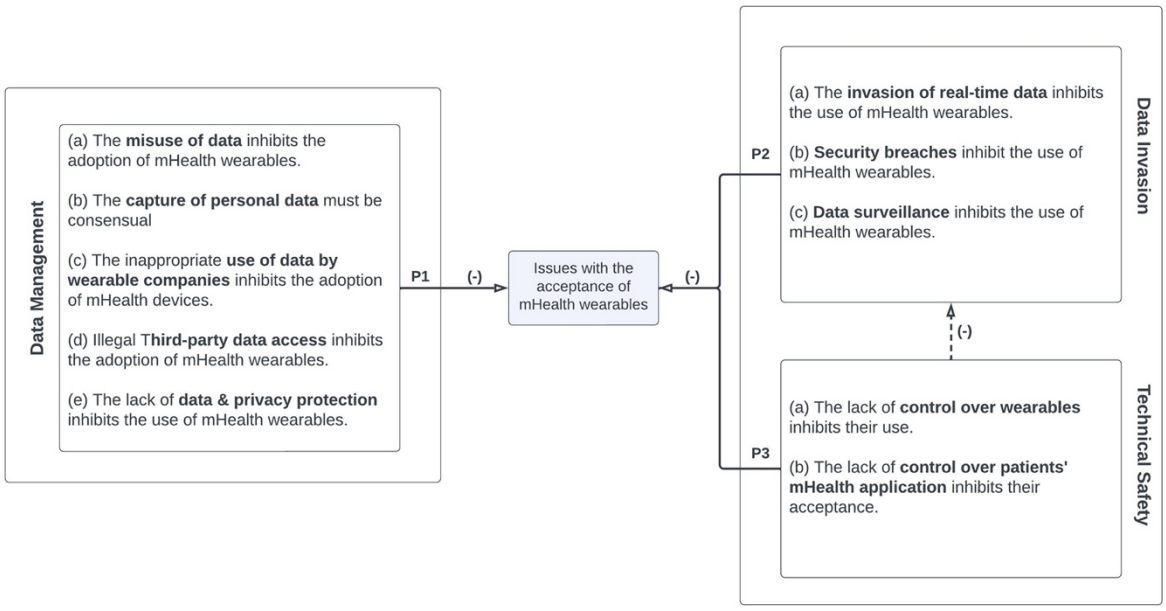

Figure 29. D-MIT Emergent Theoretical Model

The findings also demonstrated that surveillance through different tracking mechanisms results in Data Invasion issues where information is collected most times without the knowledge of the users (Datta et al., 2018). Surveillance can be seen as "any collection and processing of personal data, whether identifiable or not, for the purposes of influencing or managing those whose data have been garnered" (Lyon, 2001, p. 2). This is confirmed by Young (2018) where the top five wearable vendors were analysed to understand how they amass digital data on their users through surveillance assemblage, from which many concerns were discovered.

Our findings revealed that real-time data is affected by privacy invasion and security breach for healthcare wearables are caused by different security vulnerabilities, which all present data invasion concerns for users. This was confirmed by a study conducted by Ching & Singh (2016) outlining security and privacy vulnerabilities on wearable devices. It was

shown that there exists some security weakness that makes wearable devices vulnerable to attack. One of the critical attacks on wearable technology is authentication issues.

The Technical Safety concern is due to the lack of control over devices and data permissions, where users cannot choose to shut down a sensor individually or cancel data collection, making it difficult to authorize the viewing and use of data (Jiang & Shi, 2021). Therefore, users are concerned that they do not have dynamic control over wearables and patient apps which all have the ability to sense, collect, and store data which are often personal, confidential or sensitive; that is the user interaction with a wearable. On the other hand, users should have influence that will readily allow them to apply fine-grained control about what is collected and shared (Motti & Caine, 2015). Apparently, users believed that if they have more control over wearables and by extension patients' application, then they would be better able to minimize the data invasion issues such as security breaches, data surveillance etc. and thus increase their acceptance of these mHealth wearables.

Our study highlight issues that relate to the acceptance of mHealth wearables to include data management, data invasion, and technical safety. However, it is clearly shown on the D-MIT emergent theoretical model, that the concerns around data management as further compounded with the perceived lack of data and privacy protection, seeing that users were perturbed about the non-existence of legal and policy frameworks. Users are always concerned about their data and privacy protection, but it was apparent that there are legal and policy issues. Legal & Policy issues refer to a lack of policies and regulations on data security and privacy protection for wearable devices, especially healthcare wearables devices, once the manufacturers sell user data privately (Jiang & Shi, 2021). This concern was amplified by Lazzarotti (2015), in which it was suggested that Health Insurance Portability and

Accountability Act (HIPAA) does not apply directly to wearable devices, but may be applied to wearables and their collection of health-related data only when related to a group health plan. In other words, the number of heartbeats, steps, and sleep history tracked are not formally considered PHI unless they are shared with a doctor or third party vendors and are therefore not subject to HIPAA regulations.

Table 4. Comparison with Extant Literature

Study		**Propositions**								
	P1 (a)	**P1 (b)**	**P1 (c)**	**P1 (d)**	**P1 (e)**	**P2 (a)**	**P2 (b)**	**P2 (c)**	**P3 (a)**	**P3 (b)**
Vijayan et al. (2021)	*	*		*			*		*	
Arora et al. (2014)				*		*		*	*	
Kotz et al. (2016)		*		*					*	
Jusob et al. (2022)	*			*	*			*	*	
Datta et al. (2018)			*					*		
Ching & Singh (2016)						*	*		*	*
Zhang et al. (2020)	*	*			*					*
Kapoor et al. (2020)	*				*	*			*	
Sampat & Prabhakar (2017)	*				*	*				*
Habibipour et al. (2019)	*		*		*				*	

The comparisons in Table 4 further depict previous studies that examined the privacy and security concerns expressed by users of mHealth wearable devices. Vijayan et al. (2021) confirm our findings relating to concerns such as the misuse of data, capture of personal data, third-party data access, and security breaches. The authors showed that there was a high risk of losing confidential and personal health data, as many wearable devices were storing data in local storage without encryption or data protection. This study further confirms that the susceptibility of data hacking is increased based on how wearable sensors are always

synchronized with smartphones. Additionally, Arora et al. (2014) found that data surveillance, control over wearables, third-party data access and security breach constituted some of the main concerns in mHealth. These concerns were said to impede the full adoption of these devices and also showed that users were generally distrustful of how their data was being collected and manipulated. Further, based on a privacy framework developed by Kotz et al. (2009) three (3) concerns presented in our study were confirmed; these include data capture, data protection, and third-party data access. The study presented by Jusob et al. (2022), also confirms our findings with empirical support for five (5) of the concerns such as misuse of data, third-party data access etc. In a study examining the privacy concerns in wearable devices, Datta et al. (2018) demonstrated that issues surrounding surveillance and distrust of company continue to be major challenges for mHealth users.

In another study by Ching & Singh (2016), it was highlighted that security breaches, surveillance, and invasion of real-time data are considered major concerns, which must be tackled in order for the adoption of wearables to increase. A security and privacy analysis was done on popular wearables such as Fitbit, Google Glass, and Samsung Smartwatch, and it was revealed that users' location or places visited can be tracked, which also supports our findings. It was intimated that wearable sensors are capable of collecting a vast amount of data, including sensitive data such as health related data and credit card information with a corresponding increase in the danger of information leakage (J. Lee et al., 2016). It was also confirmed that data is collected in an obtrusive manner beyond end-user awareness (Bower & Sturman, 2015).

While Zhang et al. (2020) highlighted the strengths of popular wearables, their findings also support that data security is one of the major security vulnerabilities found in

many mobile health devices. They opined that while the European Union (EU) emphasizes data protection for tracking and monitoring patient's health information, it was presented in support of our findings that the future of using wearable devices and their applications can prove severely vulnerable. Kapoor et al. (2020) conducted an intrinsic review on privacy issues in wearable teachnology, and it was shown that concerns relating to misuse of data, surveillance, third-party data access, and lack of data and privacy protection continue to be major issues of mHealth users. They further confirmed that access to wearable data such as quality of sleep, heart rates, etc. from malicious access continues to be a threat to privacy.

Sampat & Prabhakar (2017) examined the privacy risks and security threats in mHealth apps and it was shown that while there is much convenience in using them, a lot of personal and sensitive data about users are collected, stored, and shared. Their findings also confirm the misuse of data, security breaches, control over the patients' application, and the lack of data and privacy protection as expressed concerns of users. Surprisely, while the studied assessed several health apps, it was found that some had some policies to safeguard the collection and use of personal data, but they were varied and inconsistent. This placed greater emphasis on the user to properly examine apps before they are downloaded to better understand the level of information the apps may be requesting. Further, Habibipour et al. (2019) from their study on the social, ethical, and ecological issues in wearable technologies confirmed that the misuse of data, distrust of wearable companies, control over the wearables, and lack of data and privacy protection are also expressed concerns. These issues challenge the adoption of wearable technologies.

CHAPTER 6

CONCLUSIONS

This chapter concludes the research project and presents an overview of the theoretical, methodological practical contributions of the study, as well as the limitations and future research directions.

Summary

The purpose of this study is to explore the various privacy and security concerns conveyed by social media users in relation to the use of mHealth wearable technologies, using text mining and grounded theory. In addition, the study examined the general sentiments toward mHealth privacy and security related issues, while unearthing how the various privacy and security issues have evolved over time. The results of the emerging theory explain that the concerns inhibiting the adoption of mHealth wearables can be categorized as relating to data management, data invasion, or technical safety issues. Additionally, the findings of our research reveal specific concerns to include the “misuse of data”, “data capture”, “distrust of companies”, “third-party data access”, “data and privacy protection”, “invasion of real-time data”, “data surveillance”, “security breach”, “control over wearables”, and “control over patients’ apps.” These findings were compared with extant literature, and found confirmation across several studies.

The findings reveal that more than 75% of the posts analyzed were categorized as depicting anger, fear, or demonstrating levels of disgust. Further, the study shows that 70% of the posts demonstrated negative sentiments, whereas 26% were positive, which indicates that

users are ambivalent concerning privacy and security, notwithstanding mentions of privacy or security issues in their posts, there was a general positive tone. Additionally, the findings show that overtime, users have been more concerned about issues relating to surveillance and how much control they truly have over these mHealth wearables, along with the invasion of real time data and security breaches. It also shows that users generally do not trust how companies such as Fitbit use the personal data collected from them.

Myers (2009) recommended two vital conditions that must be met during the evaluation of grounded theory research: 1) rigor and validity; 2) generalization. In this study, the rigor and validity of the data analysis was realized through the use of a text-mining approach where concepts were extrapolated from a large corpus of tweets. This was also supported by the systematic approach in conducting the different grounded theory phases. Additionally, several tweets were identified which supported the privacy and security concerns deduced. Importantly, compared to manual coding with limited occurrences of the data, a higher degree of consistency and reliability can be realized through the mining of knowledge from a sizable volume of data (Yu et al., 2011). In terms of generalizability, we developed the D-MIT emergent theoretical framework by extracting knowledge from the large corpus of text data. The framework demonstrated three (3) overarching privacy and security concerns: data management, data invasion, and technical safety issues.

Contributions

Theoretically, the findings of this study contribute to the literature of users' acceptance of health consumer technology, by unearthing the privacy and security concerns that may inhibit their adoption. Further, the findings provide evidence through the D-MIT Emergent Theoretical model, that users of mHealth wearables are concerned about data management,

data invasion, and technical safety issues, which finds support from extant literature dealing with privacy and security concerns in the wearables domain. Finally, the study reveals which of the privacy and security concerns mHealth users are most concerned about.

Methodologically, the capability of text mining within the grounded theory context was utilized. We used the LDA algorithm for topic modeling, to automatically extract concepts from large amounts of text data, instead of manually analyzing and coding the tweets, which is time-consuming and subjective. As far as we know, this is the first work that leverages social media mining to understand the privacy and security concerns of mHealth users. Automatically evaluating social media users' posts with the utilization of machine learning tools, can assist in understanding the themes and topics that exist in the tweets shared by online users.

Practically, it can help policy makers with developing comprehensive guidelines to govern data collection, dissemination, and processing on these devices. Additionally, the findings can guide companies who develop and distribute wearable devices in better understanding the expressed concerns of users, especially in the area of having greater control over the wearables and also apps used for patient care. Doctors and other health practitioners who use these mHealth devices can understand reasons which may inhibit the adoption of these wearables by their patients, and develop strategies to mitigate these concerns. Furthermore, better indicators of the acceptance and use of mHealth devices can be established through available data on the web which provides opportunities for tracking and analyzing actual users' opinions about a phenomenon (Motiwalla et al., 2019).

Limitations & Future Research

A limitation of the study is the potential noise that accompanies social media posts and the impact of pulling data from only the Twitter social media platform. Further, another limitation is the difficulty in generalizing the findings emerging from the analyzed tweets. Future research may also investigate other factors relating to privacy and security concerns in healthcare wearables usage and adoption such as the role of age, gender, and culture. Further studies will examine the relationships that exist between expressed sentiments and each privacy and security concerns. Other studies can investigate the generalizability of the developed emergent theory. In addition, understanding the concerns from users on other popular social media platforms like Reddit and Facebook may be beneficial. Finally, quantitative research studies may aim to explore the propositions outlined in the D-MIT emergent theoretical model, as this will help to understand the relationships among the suggested constructs.

REFERENCES

Abdolkhani, R., Gray, K., Borda, A., & DeSouza, R. (2020). Quality Assurance of Health Wearables Data: Participatory Workshop on Barriers, Solutions, and Expectations. *JMIR MHealth and UHealth*, *8*(1). https://doi.org/10.2196/15329

Acquisti, A., Brandimarte, L., & Loewenstein, G. (2015). Privacy and human behavior in the age of information. *Science*, *347*(6221), 509–514. https://doi.org/10.1126/science.aaa1465

Adibi, S., Mobasher, A., & Tofigh, T. (2013). *LTE networking: Extending the reach for sensors in mHealth applications*. https://www.ezproxy.dsu.edu:2102/doi/abs/10.1002/ett.2598

Ahmad, T., Alvi, A., & Ittefaq, M. (2019). The Use of Social Media on Political Participation Among University Students: An Analysis of Survey Results From Rural Pakistan. *SAGE Open*, *9*(3), 2158244019864484. https://doi.org/10.1177/2158244019864484

Ajzen, I. (1991). The Theory of Planned Behavior. *Organizational Behavior and Human Decision Processes*, *50*, 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Prentice-Hall.

Ajzen, I., & Fishbein, M. (2000). Attitudes and the Attitude-Behavior Relation: Reasoned and Automatic Processes. *European Review of Social Psychology - EUR REV SOC PSYCHOL*, *11*, 1–33. https://doi.org/10.1080/14792779943000116

Al Ameen, M., Liu, J., & Kwak, K. (2012). Security and Privacy Issues in Wireless Sensor Networks for Healthcare Applications. *Journal of Medical Systems*, *36*, 93–101. https://doi.org/10.1007/s10916-010-9449-4

AlHogail, A. (2018). Improving IoT Technology Adoption through Improving Consumer Trust. *Technologies*, *6*(3), 64. https://doi.org/10.3390/technologies6030064

Al-Momani, A. M., Mahmoud, M. A., & Sharifuddin, M. (2016). *Modeling the adoption of internet of things services: A conceptual framework*. 7.

Al-Ramahi, M., El-Gayar, O., & Liu, J. (2016). Discovering Design Principles for Persuasive Systems: A Grounded Theory and Text Mining Approach. *2016 49th Hawaii*

International Conference on System Sciences (HICSS), 3074–3083. https://doi.org/10.1109/HICSS.2016.387

Altman, I. (1976). A Conceptual Analysis. *Environment and Behavior*, *8*(1), 7–29. https://doi.org/10.1177/001391657600800102

Anderson, C. L., & Agarwal, R. (2011). The Digitization of Healthcare: Boundary Risks, Emotion, and Consumer Willingness to Disclose Personal Health Information. *Information Systems Research*, *22*(3), 469–490. https://doi.org/10.1287/isre.1100.0335

Andreu-Perez, J., Leff, D., Ip, H., & Yang, G.-Z. (2015). From Wearable Sensors to Smart Implants-–Toward Pervasive and Personalized Healthcare. *IEEE Transactions on Bio-Medical Engineering*, *62*. https://doi.org/10.1109/TBME.2015.2422751

Anstead, N., & O'Loughlin, B. (2015). Social Media Analysis and Public Opinion: The 2010 UK General Election. *Journal of Computer-Mediated Communication*, *20*(2), 204–220. https://doi.org/10.1111/jcc4.12102

Arora, S., Yttri, J., & Nilsen, W. (2014). Privacy and Security in Mobile Health (mHealth) Research. *Alcohol Research : Current Reviews*, *36*(1), 143–151.

Avancha, S., Baxi, A., & Kotz, D. (2012). Privacy in mobile technology for personal healthcare. *CSUR*. https://doi.org/10.1145/2379776.2379779

Azzopardi, L., & Rijsbergen, V. (2003). Investigating the relationship between language model perplexity and IR precision-recall measures. *Annual ACM Conference on Research and Development in Information Retrieval*.

Bacharach, S. B. (1989). Organizational theories: Some criteria for evaluation. *The Academy of Management Review*, *14*(4), 496–515. https://doi.org/10.2307/258555

Bansal, G., Zahedi, F. "Mariam," & Gefen, D. (2010). The impact of personal dispositions on information sensitivity, privacy concern and trust in disclosing health information online. *Decision Support Systems*, *49*(2), 138–150. https://doi.org/10.1016/j.dss.2010.01.010

Berendt, B., Günther, O., & Spiekermann, S. (2005). Privacy in E-commerce: Stated Preferences vs. Actual Behavior. *Commun. ACM*, *48*, 101–106.

Bing, L., Chan, K. C. C., & Ou, C. (2014). Public Sentiment Analysis in Twitter Data for Prediction of a Company's Stock Price Movements. *2014 IEEE 11th International*

Conference on E-Business Engineering, 232–239. https://doi.org/10.1109/ICEBE.2014.47

Blei, D. M. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, *3*, 30.

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, *55*(4), 77–84. https://doi.org/10.1145/2133806.2133826

Bower, M., & Sturman, D. (2015). What are the educational affordances of wearable technologies? *Computers & Education*, *88*, 343–353. https://doi.org/10.1016/j.compedu.2015.07.013

Brandimarte, L., Acquisti, A., & Loewenstein, G. (2013). Misplaced Confidences: Privacy and the Control Paradox. *Social Psychological and Personality Science*, *4*(3), 340–347. https://doi.org/10.1177/1948550612455931

Bunnig, C., & Cap, C. H. (2009). Ad Hoc Privacy Management in Ubiquitous Computing Environments. *2009 Second International Conference on Advances in Human-Oriented and Personalized Mechanisms, Technologies, and Services*, 85–90. https://doi.org/10.1109/CENTRIC.2009.20

Chang, J., Boyd-Graber, J., Gerrish, S., Wang, C., & Blei, D. M. (2009). *Reading Tea Leaves: How Humans Interpret Topic Models*. 10.

Charani, E., Gharbi, M., Moore, L., Castro-Sánchez, E., Lawson, W., Gilchrist, M., & Holmes, A. (2017). Effect of adding a mobile health intervention to a multimodal antimicrobial stewardship programme across three teaching hospitals: An interrupted time series study. *The Journal of Antimicrobial Chemotherapy*, *72*. https://doi.org/10.1093/jac/dkx040

Charmaz, K. (2006). *Constructing grounded theory*. Sage Publications.

Charmaz, K. (2011). *Grounded Theory Methods in Social Justice Research* (pp. 359–380).

Charmaz, K. (2014). Grounded Theory in Global Perspective: Reviews by International Researchers. *Qualitative Inquiry*, *20*(9), 1074–1084. https://doi.org/10.1177/1077800414545235

Ching, K. W., & Singh, M. M. (2016). Wearable Technology Devices Security and Privacy Vulnerability Analysis. *International Journal of Network Security & Its Applications*, *8*(3), 19–30. https://doi.org/10.5121/ijnsa.2016.8302

Chiru, C., Rebedea, T., & Ciotec, S. (2014). Comparison between LSA-LDA-lexical chains. In *WEBIST 2014—Proceedings of the 10th International Conference on Web Information Systems and Technologies* (Vol. 2).

Cho, H., Rivera, M., & Lim, S. S. (2009). A multinational study on online privacy: Global concerns and local responses. *New Media & Society - NEW MEDIA SOC*, *11*, 395–416. https://doi.org/10.1177/1461444808101618

Clarke, R. (1999). Introduction to dataveillance and information privacy, and definitions of terms. *Roger Clarke's Dataveillance and Information Privacy Pages*.

Corbin, J., & Strauss, A. (1990). *Grounded theory research: Procedures, canons, and evaluative criteria*. 19.

Correia, R. B., Li, L., & Rocha, L. M. (2016). MONITORING POTENTIAL DRUG INTERACTIONS AND REACTIONS VIA NETWORK ANALYSIS OF INSTAGRAM USER TIMELINES. *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing*, *21*, 492–503.

Correia, R., Wood, I., Bollen, J., & Rocha, L. (2020). *Mining social media data for biomedical signals and health-related behavior*.

Creswell, J. W. (2003). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. SAGE Publications.

Culnan, M. J., & Armstrong, P. K. (1999). Information Privacy Concerns, Procedural Fairness, and Impersonal Trust: An Empirical Investigation. *Organization Science*, *10*(1), 104–115. https://doi.org/10.1287/orsc.10.1.104

Culnan, M. J., & Bies, R. J. (2003). Consumer Privacy: Balancing Economic and Justice Considerations: **Consumer Privacy**. *Journal of Social Issues*, *59*(2), 323–342. https://doi.org/10.1111/1540-4560.00067

Datta, P., Namin, A. S., & Chatterjee, M. (2018). A Survey of Privacy Concerns in Wearable Devices. *2018 IEEE International Conference on Big Data (Big Data)*, 4549–4553. https://doi.org/10.1109/BigData.2018.8622110

de Arriba-Pérez, F., Caeiro-Rodríguez, M., & Santos-Gago, J. M. (2016). Collection and Processing of Data from Wrist Wearable Devices in Heterogeneous and Multiple-User Scenarios. *Sensors (Basel, Switzerland)*, *16*(9), 1538. https://doi.org/10.3390/s16091538

Dinev, T., & Hart, P. (2006). An Extended Privacy Calculus Model for E-Commerce Transactions. *Information Systems Research*, *17*(1), 61–80.

Dinev, T., Xu, H., Smith, J. H., & Hart, P. (2013). Information privacy and correlates: An empirical attempt to bridge and distinguish privacy-related concepts. *European Journal of Information Systems*, *22*(3), 295–316. https://doi.org/10.1057/ejis.2012.23

Domalewska, D. (2021). An analysis of COVID-19 economic measures and attitudes: Evidence from social media mining. *Journal of Big Data*, *8*. https://doi.org/10.1186/s40537-021-00431-z

Eastlick, M. A., Lotz, S. L., & Warrington, P. (2006). Understanding online B-to-C relationships: An integrated model of privacy concerns, trust, and commitment. *Journal of Business Research*, *59*(8), 877–886. https://doi.org/10.1016/j.jbusres.2006.02.006

Ekman, P. (1993). Facial expression and emotion. *American Psychologist*, *48*(4), 384–392. https://doi.org/10.1037/0003-066X.48.4.384

Falcone, R., & Sapienza, A. (2018). On the Users' Acceptance of IoT Systems: A Theoretical Approach. *Information*, *9*(3), 53. https://doi.org/10.3390/info9030053

Faudree, B., & Ford, M. (2013). Security and Privacy in Mobile Health. *CIO Journal*.

Feldman, R., & Sanger, J. (2006). *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*. Cambridge University Press. https://doi.org/10.1017/CBO9780511546914

Filkins, B. L., Kim, J. Y., Roberts, B., Armstrong, W., Miller, M. A., Hultner, M. L., Castillo, A. P., Ducom, J.-C., Topol, E. J., & Steinhubl, S. R. (2016). Privacy and security in the era of digital health: What should translational researchers know and do about it? *American Journal of Translational Research*, *8*(3), 1560–1580.

Finn, R., Wright, D., & Friedewald, M. (2013). Seven types of privacy European data protection: Coming of Age. *Springer*, 3–32.

Flavián, C., Guinalíu, M., & Gurrea, R. (2006). The role played by perceived usability, satisfaction and consumer trust on website loyalty. *Information & Management*, *43*(1), 1–14. https://doi.org/10.1016/j.im.2005.01.002

Fung, I. C.-H., Tse, Z. T. H., & Fu, K.-W. (2015). The use of social media in public health surveillance. *Western Pacific Surveillance and Response Journal: WPSAR*, *6*(2), 3–6. https://doi.org/10.5365/WPSAR.2015.6.1.019

Gallagher, R., Roach, K., Sadler, L., Glinatsis, H., & Belshaw, J. (2017). Mobile Technology Use Across Age Groups in Patients Eligible for Cardiac Rehabilitation: Survey Study. *JMIR MHealth and UHealth*, *5*(10), e161. https://doi.org/10.2196/mhealth.8352

Ganapathy, R., Grewal, A., & Castleman, J. S. (2016). Remote monitoring of blood pressure to reduce the risk of preeclampsia related complications with an innovative use of mobile technology. *Pregnancy Hypertension*, *6*(4), 263–265. https://doi.org/10.1016/j.preghy.2016.04.005

Gao, L., & Bai, X. (2014). A unified perspective on the factors influencing consumer acceptance of internet of things technology. *Asia Pacific Journal of Marketing and Logistics*, *26*(2), 211–231. https://doi.org/10.1108/APJML-06-2013-0061

Gemar, G., & Jiménez-Quintero, J. (2015). Text mining social media for competitive analysis. *Tourism & Management Studies*, *11*, 84–90.

Giannetsos, T., Dimitriou, T., & Prasad, N. R. (2011). People-centric sensing in assistive healthcare: Privacy challenges and directions. *Security and Communication Networks*, *4*(11), 1295–1307. https://doi.org/10.1002/sec.313

Glaser, B. G. (1978). *Theoretical sensitivity* /. University of California,. https://eduq.info/xmlui/handle/11515/17665

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*.

Goldberg, R. M., McHenry, G., Zambrano Ramos, L., & Chen, C. (2016). *Trust in Internet Privacy and Security and Online Activity* (SSRN Scholarly Paper ID 2757369). Social Science Research Network. https://doi.org/10.2139/ssrn.2757369

Goulding, C. (2002). *Grounded Theory* (1st ed.). University of Birmingham, UK. https://uk.sagepub.com/en-gb/eur/grounded-theory/book210303

Grimmer, J., & Stewart, B. M. (2013). Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts. *Political Analysis*, *21*(3), 267–297. https://doi.org/10.1093/pan/mps028

Gu, J., Xu, Y. (Calvin), Xu, H., Zhang, C., & Ling, H. (2017). Privacy concerns for mobile app download: An elaboration likelihood model perspective. *Decision Support Systems*, *94*, 19–28. https://doi.org/10.1016/j.dss.2016.10.002

Guo, X., Han, X., Zhang, X., Dang, Y., & Chen, C. (2015). Investigating m-Health Acceptance from a Protection Motivation Theory Perspective: Gender and Age Differences. *Telemedicine Journal and E-Health : The Official Journal of the American Telemedicine Association*, *21*. https://doi.org/10.1089/tmj.2014.0166

Gurses, S., & Diaz, C. (2013). Two Tales of Privacy in Online Social Networks. *Security & Privacy, IEEE*, *11*, 29–37. https://doi.org/10.1109/MSP.2013.47

Habibipour, A., Padyab, A., & St, A. (2019). *Social, Ethical and Ecological Issues in Wearable Technologies*. 10.

Haddi, E., Liu, X., & Shi, Y. (2013). The Role of Text Pre-processing in Sentiment Analysis. *Procedia Computer Science*, *17*, 26–32. https://doi.org/10.1016/j.procs.2013.05.005

Haghi, M., Thurow, K., & Stoll, R. (2017). Wearable Devices in Medical Internet of Things: Scientific Research and Commercially Available Devices. *Healthcare Informatics Research*, *23*(1), 4–15. https://doi.org/10.4258/hir.2017.23.1.4

Han, J., Kamber, M., & Pei, J. (2011). *Data Mining: Concepts and Techniques* (3rd ed.). Morgan Kaufmann Publishers Inc.

Hopkins, D. J., & King, G. (2010). A Method of Automated Nonparametric Content Analysis for Social Science. *American Journal of Political Science*, *54*(1), 229–247. https://doi.org/10.1111/j.1540-5907.2009.00428.x

Iwaya, L. H., Ahmad, A., & Babar, M. A. (2020). Security and Privacy for mHealth and uHealth Systems: A Systematic Mapping Study. *IEEE Access*, *8*, 150081–150112. https://doi.org/10.1109/ACCESS.2020.3015962

Jamal, A. A., Keohane, R. O., Romney, D., & Tingley, D. (2015). Anti-Americanism and Anti-Interventionism in Arabic Twitter Discourses. *Perspectives on Politics*, *13*(1), 55–73. https://doi.org/10.1017/S1537592714003132

Jeong, B., Yoon, J., & Lee, J.-M. (2019). Social media mining for product planning: A product opportunity mining approach based on topic modeling and sentiment analysis. *International Journal of Information Management*, *48*, 280–290. https://doi.org/10.1016/j.ijinfomgt.2017.09.009

Jiang, D., & Shi, G. (2021). Research on Data Security and Privacy Protection of Wearable Equipment in Healthcare. *Journal of Healthcare Engineering*, *2021*. https://doi.org/10.1155/2021/6656204

Jusob, F. R., George, C., & Mapp, G. (2022). A new privacy framework for the management of chronic diseases via mHealth in a post-Covid-19 world. *Journal of Public Health*, *30*(1), 37–47. https://doi.org/10.1007/s10389-021-01608-9

Kang, K., Pang, Z., & Wang, C. (2013). Security and privacy mechanism for health internet of things. *The Journal of China Universities of Posts and Telecommunications*, *20*, 64–68. https://doi.org/10.1016/S1005-8885(13)60219-8

Kang, Y., Wang, Y., Zhang, D., & Zhou, L. (2017). The public's opinions on a new school meals policy for childhood obesity prevention in the U.S.: A social media analytics approach. *International Journal of Medical Informatics*, *103*, 83–88. https://doi.org/10.1016/j.ijmedinf.2017.04.013

Kaplan, A., & Haenlein, M. (2010). Users of the World, Unite! The Challenges and Opportunities of Social Media. *Business Horizons*, *53*, 59–68. https://doi.org/10.1016/j.bushor.2009.09.003

Kapoor, V., Singh, R., Reddy, R., & Churi, P. (2020). Privacy Issues in Wearable Technology: An Intrinsic Review. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3566918

Kim, A. E., Hansen, H. M., Murphy, J., Richards, A. K., Duke, J., & Allen, J. A. (2013). Methodological considerations in analyzing Twitter data. *Journal of the National Cancer Institute. Monographs*, *2013*(47), 140–146. https://doi.org/10.1093/jncimonographs/lgt026

Kim, D., Park, K., Park, Y., & Ahn, J.-H. (2019). Willingness to provide personal information: Perspective of privacy calculus in IoT services. *Computers in Human Behavior*, *92*, 273–281. https://doi.org/10.1016/j.chb.2018.11.022

Kotz, D., Avancha, S., & Baxi, A. (2009). A privacy framework for mobile health and home-care systems. *Proceedings of the First ACM Workshop on Security and Privacy in Medical and Home-Care Systems - SPIMACS '09*, 1. https://doi.org/10.1145/1655084.1655086

Kotz, D., Gunter, C. A., Kumar, S., & Weiner, J. P. (2016a). Privacy and Security in Mobile Health: A Research Agenda. *Computer*, *49*(6), 22–30. https://doi.org/10.1109/MC.2016.185

Kotz, D., Gunter, C. A., Kumar, S., & Weiner, J. P. (2016b). Privacy and Security in Mobile Health: A Research Agenda. *Computer*, *49*(6), 22–30. https://doi.org/10.1109/MC.2016.185

Krohn, R. (2015). mHealth: A pathway to the intelligent hospital. *MHealth*, *1*, 16. https://doi.org/10.3978/j.issn.2306-9740.2015.08.01

Kumar, S., Nilsen, W., Pavel, M., & Srivastava, M. (2013). Mobile Health: Revolutionizing Healthcare Through Transdisciplinary Research. *Computer*, *46*, 28–35. https://doi.org/10.1109/MC.2012.392

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174.

Lazzarotti, J. (2015). *Wearables, Wellness and Privacy*. The National Law Review. https://www.natlawreview.com/article/wearables-wellness-and-privacy

Lee, C., Eze, U., & Ndubisi, N. (2011). Analyzing key determinants of online repurchase intentions. *Asia Pacific Journal of Marketing and Logistics*, *23*, 200–221. https://doi.org/10.1108/13555851111120498

Lee, J., Kim, D., Ryoo, H.-Y., & Shin, B.-S. (2016). Sustainable Wearables: Wearable Technology for Enhancing the Quality of Human Life. *Sustainability*, *8*(5), 466. https://doi.org/10.3390/su8050466

Lee, L., Egelman, S., Lee, J. H., & Wagner, D. (2015). Risk Perceptions for Wearable Devices. *ArXiv:1504.05694 [Cs]*. http://arxiv.org/abs/1504.05694

Li, H., Wu, J., 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, B. (2010). *Sentiment Analysis and Subjectivity*. 38.

Lu, Y., Zhang, P., Liu, J., Li, J., & Deng, S. (2013). Health-related hot topic detection in online communities using text clustering. *PloS One*, *8*(2), e56221. https://doi.org/10.1371/journal.pone.0056221

Luxton, D. D., Kayl, R. A., & Mishkind, M. C. (2012). mHealth data security: The need for HIPAA-compliant standardization. *Telemedicine Journal and E-Health: The Official Journal of the American Telemedicine Association*, *18*(4), 284–288. https://doi.org/10.1089/tmj.2011.0180

Lyon, D. (2001). *Surveillance society: Monitoring Everyday Life* (1st edition). Open University Press.

Machi, L. A., & McEvoy, B. T. (2016). *The literature review: Six steps to success*.

Maddux, J., & Rogers, R. (1983). Protection Motivation and Self-Efficacy: A Revised Theory of Fear Appeals and Attitude Change. *Journal of Experimental Social Psychology*, *19*, 469–479. https://doi.org/10.1016/0022-1031(83)90023-9

Malhotra, N., Kim, S., & Agarwal, J. (2004). Internet Users' Information Privacy Concerns (IUIPC): The Construct, the Scale, and a Causal Model. *Information Systems Research*, *15*, 336–355. https://doi.org/10.1287/isre.1040.0032

Market Research Focus. (2020). *Mhealth Market Size, Share,Trends,Growth, Analysis -2027*. https://www.marketresearchfuture.com/reports/mobile-health-market-1816

McCallum, C., Rooksby, J., & Gray, C. M. (2018). Evaluating the Impact of Physical Activity Apps and Wearables: Interdisciplinary Review. *JMIR MHealth and UHealth*, *6*(3), e58. https://doi.org/10.2196/mhealth.9054

McGough, S. F., Brownstein, J. S., Hawkins, J. B., & Santillana, M. (2017). Forecasting Zika Incidence in the 2016 Latin America Outbreak Combining Traditional Disease Surveillance with Search, Social Media, and News Report Data. *PLoS Neglected Tropical Diseases*, *11*(1), e0005295. https://doi.org/10.1371/journal.pntd.0005295

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, *22*(3), 276–282.

Metcalf, D., Milliard, S., Gomez, M., & Schwartz, M. (2016). Wearables and the Internet of Things for Health: Wearable, Interconnected Devices Promise More Efficient and Comprehensive Health Care. *IEEE Pulse*, *7*, 35–39. https://doi.org/10.1109/MPUL.2016.2592260

Miller, S., Ainsworth, B., Yardley, L., Milton, A., Weal, M., Smith, P., & Morrison, L. (2019). A Framework for Analyzing and Measuring Usage and Engagement Data

(AMUsED) in Digital Interventions: Viewpoint. *Journal of Medical Internet Research*, *21*(2), e10966. https://doi.org/10.2196/10966

Miltgen, C., Popovič, A., & Oliveira, T. (2013). Determinants of end-user acceptance of biometrics: Integrating the "Big 3" of technology acceptance with privacy context. *Decision Support Systems*, *56*, 103–114. https://doi.org/10.1016/j.dss.2013.05.010

Motiwalla, L., Deokar, A. V., Sarnikar, S., & Dimoka, A. (2019). Leveraging Data Analytics for Behavioral Research. *Information Systems Frontiers*, *21*(4), 735–742. https://doi.org/10.1007/s10796-019-09928-8

Motti, V., & Caine, K. (2015). *Users' Privacy Concerns About Wearables: Impact of form factor, sensors and type of data collected* (Vol. 8976). https://doi.org/10.1007/978-3-662-48051-9_16

Motti, V. G., & Caine, K. (2014). Human Factors Considerations in the Design of Wearable Devices. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *58*(1), 1820–1824. https://doi.org/10.1177/1541931214581381

Myers, M. D. (2009). *Qualitative research in business & management* (pp. xii, 284). Sage Publications Ltd.

Nagar, R., Yuan, Q., Freifeld, C. C., Santillana, M., Nojima, A., Chunara, R., & Brownstein, J. S. (2014). A case study of the New York City 2012-2013 influenza season with daily geocoded Twitter data from temporal and spatiotemporal perspectives. *Journal of Medical Internet Research*, *16*(10), e236. https://doi.org/10.2196/jmir.3416

Nguyen, T., Larsen, M. E., O'Dea, B., Phung, D., Venkatesh, S., & Christensen, H. (2017). Estimation of the prevalence of adverse drug reactions from social media. *International Journal of Medical Informatics*, *102*, 130–137. https://doi.org/10.1016/j.ijmedinf.2017.03.013

Pandit, N. (1996). The Creation of Theory: A Recent Application of the Grounded Theory Method. *The Qualitative Report*, *2*(4), 1–15. https://doi.org/10.46743/2160-3715/1996.2054

Paul, M. J., Sarker, A., Brownstein, J. S., Nikfarjam, A., Scotch, M., Smith, K. L., & Gonzalez, G. (2016). SOCIAL MEDIA MINING FOR PUBLIC HEALTH MONITORING AND SURVEILLANCE. *Biocomputing 2016*, 468–479. https://doi.org/10.1142/9789814749411_0043

Petronio, S. (2002). *Boundaries of privacy: Dialectics of disclosure* (pp. xix, 268). State University of New York Press.

Plachkinova, M., Andrés, S., & Chatterjee, S. (2015). A Taxonomy of mHealth Apps – Security and Privacy Concerns. *2015 48th Hawaii International Conference on System Sciences*, 3187–3196. https://doi.org/10.1109/HICSS.2015.385

Prasad, A., Sorber, J., Stablein, T., Anthony, D., & Kotz, D. (2012). Understanding sharing preferences and behavior for mHealth devices. *Proceedings of the 2012 ACM Workshop on Privacy in the Electronic Society*, 117–128. https://doi.org/10.1145/2381966.2381983

Prentice-Dunn, S., & Rogers, R. W. (1986). Protection Motivation Theory and preventive health: Beyond the Health Belief Model. *Health Education Research*, *1*(3), 153–161. https://doi.org/10.1093/her/1.3.153

Reidenberg, J., Breaux, T., Cranor, L., French, B., Grannis, A., Graves, J., Liu, F., McDonald, A., Norton, T., Ramanath, R., Russell, N., Sadeh, N., & Schaub, F. (2014). Disagreeable Privacy Policies: Mismatches between Meaning and Userss Understanding. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2418297

Robinson, J., Rodrigues, M., Fisher, S., Bailey, E., & Herrman, H. (2015). Social media and suicide prevention: Findings from a stakeholder survey. *Shanghai Archives of Psychiatry*, *27*(1), 27–35. https://doi.org/10.11919/j.issn.1002-0829.214133

Rogers, R. W. (1975). A Protection Motivation Theory of Fear Appeals and Attitude Change1. *The Journal of Psychology*, *91*(1), 93–114. https://doi.org/10.1080/00223980.1975.9915803

Runge, K. K., Yeo, S. K., Cacciatore, M., Scheufele, D. A., Brossard, D., Xenos, M., Anderson, A., Choi, D., Kim, J., Li, N., Liang, X., Stubbings, M., & Su, L. Y.-F. (2013). Tweeting nano: How public discourses about nanotechnology develop in social media environments. *Journal of Nanoparticle Research*, *15*(1). https://doi.org/10.1007/s11051-012-1381-8

Sampat, B., & Prabhakar, B. (2017). *Privacy Risks and Security Threats in mHealth apps*. *26*.

Santos, J. C., & Matos, S. (2014). Analysing Twitter and web queries for flu trend prediction. *Theoretical Biology & Medical Modelling*, *11 Suppl 1*, S6. https://doi.org/10.1186/1742-4682-11-S1-S6

Sarker, A., O'Connor, K., Ginn, R., Scotch, M., Smith, K., Malone, D., & Gonzalez, G. (2016). Social Media Mining for Toxicovigilance: Automatic Monitoring of Prescription Medication Abuse from Twitter. *Drug Safety*, *39*(3), 231–240. https://doi.org/10.1007/s40264-015-0379-4

Seltzer, E. K., Horst-Martz, E., Lu, M., & Merchant, R. M. (2017). Public sentiment and discourse about Zika virus on Instagram. *Public Health*, *150*, 170–175. https://doi.org/10.1016/j.puhe.2017.07.015

Shklovski, I., Mainwaring, S., Skúladóttir, H., & Borgthorsson, H. (2014). Leakiness and Creepiness in App Space: Perceptions of Privacy and Mobile App Use Mobile. In *Conference on Human Factors in Computing Systems—Proceedings*. https://doi.org/10.1145/2556288.2557421

Smith, H. J., Milberg, S. J., & Burke, S. J. (1996). Information privacy: Measuring individuals' concerns about organizational practices. *MIS Quarterly*, *20*(2), 167–196. https://doi.org/10.2307/249477

Solove, D. J. (2006). A Taxonomy of Privacy. *University of Pennsylvania Law Review*, *154*(3), 477. https://doi.org/10.2307/40041279

Stowell, E., Lyson, M. C., Saksono, H., Wurth, R. C., Jimison, H., Pavel, M., & Parker, A. G. (2018). Designing and Evaluating mHealth Interventions for Vulnerable Populations: A Systematic Review. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–17. https://doi.org/10.1145/3173574.3173589

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory, 2nd ed* (pp. xiii, 312). Sage Publications, Inc.

Sutton, R. I., & Staw, B. M. (1995). What Theory is Not. *Administrative Science Quarterly*, *40*(3), 371–384. https://doi.org/10.2307/2393788

Taddicken, M. (2014). The 'Privacy Paradox' in the Social Web: The Impact of Privacy Concerns, Individual Characteristics, and the Perceived Social Relevance on Different Forms of Self-Disclosure*. *Journal of Computer-Mediated Communication*, *19*(2), 248–273. https://doi.org/10.1111/jcc4.12052

Tan, P., Steinbach, M., & Kumar, V. (2005). *Introduction to Data Mining*. Pearson.

Tanner, A. E., Mann, L., Song, E., Alonzo, J., Schafer, K., Arellano, E., Garcia, J. M., & Rhodes, S. D. (2016). weCARE: A Social Media-Based Intervention Designed to

Increase HIV Care Linkage, Retention, and Health Outcomes for Racially and Ethnically Diverse Young MSM. *AIDS Education and Prevention: Official Publication of the International Society for AIDS Education*, *28*(3), 216–230. https://doi.org/10.1521/aeap.2016.28.3.216

Trepte, S., Teutsch, D., Masur, P. K., Eicher, C., Fischer, M., Hennhöfer, A., & Lind, F. (2015). Do People Know About Privacy and Data Protection Strategies? Towards the "Online Privacy Literacy Scale" (OPLIS). In S. Gutwirth, R. Leenes, & P. de Hert (Eds.), *Reforming European Data Protection Law* (pp. 333–365). Springer Netherlands. https://doi.org/10.1007/978-94-017-9385-8_14

Tuarob, S., Tucker, C. S., Salathe, M., & Ram, N. (2014). An ensemble heterogeneous classification methodology for discovering health-related knowledge in social media messages. *Journal of Biomedical Informatics*, *49*, 255–268. https://doi.org/10.1016/j.jbi.2014.03.005

U.S. Dept. of Health & Human Services. (2020). *Summary of the HIPAA Security Rule*. https://www.hhs.gov/guidance/document/summary-hipaa-security-rule-1

Van Maanen, J. (1990). GREAT MOMENTS IN ETHNOGRAPHY: An Editor's Introduction. *Journal of Contemporary Ethnography*, *19*(1), 3–7. https://doi.org/10.1177/089124190019001001

Vaske, J., & Donnelly, M. (1999). A Value-Attitude-Behavior Model Predicting Wildland Preservation Voting Intentions. *Society and Natural Resources*, *12*. https://doi.org/10.1080/089419299279425

Vegesna, A., Tran, M., Angelaccio, M., & Arcona, S. (2017). Remote Patient Monitoring via Non-Invasive Digital Technologies: A Systematic Review. *Telemedicine Journal and E-Health: The Official Journal of the American Telemedicine Association*, *23*(1), 3–17. https://doi.org/10.1089/tmj.2016.0051

Vijayan, V., Connolly, J. P., Condell, J., McKelvey, N., & Gardiner, P. (2021). Review of Wearable Devices and Data Collection Considerations for Connected Health. *Sensors (Basel, Switzerland)*, *21*(16), 5589. https://doi.org/10.3390/s21165589

Wang, Y., & Nepali, R. K. (2015). Privacy impact assessment for online social networks. *2015 International Conference on Collaboration Technologies and Systems (CTS)*, 370–375. https://doi.org/10.1109/CTS.2015.7210451

Weinstein, N. (1993). Testing four competing theories of health-protective behavior. *Health Psychology : Official Journal of the Division of Health Psychology, American Psychological Association*, *12*(4), 324–333. https://doi.org/10.1037//0278-6133.12.4.324

Westin, A. (1968). Privacy And Freedom. *Washington and Lee Law Review*, *25*(1), 166.

Westin, A. F. (1967). *Privacy and freedom* ([1st ed.]). Atheneum.

Wiesche, M., Jurisch, M., Yetton, P., & Krcmar, H. (2017). Grounded Theory Methodology in Information Systems Research. *MIS Quarterly*, *41*, 685–701. https://doi.org/10.25300/MISQ/2017/41.3.02

Wilson, R. E., Gosling, S. D., & Graham, L. T. (2012). A Review of Facebook Research in the Social Sciences. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *7*(3), 203–220. https://doi.org/10.1177/1745691612442904

World Health Organization. (2011). *mHealth: New horizons for health through mobile technologies: second global survey on eHealth*. World Health Organization. https://www.cabdirect.org/globalhealth/abstract/20113217175

Wu, H., Fang, H., & Stanhope, S. J. (2013). Exploiting online discussions to discover unrecognized drug side effects. *Methods of Information in Medicine*, *52*(2), 152–159. https://doi.org/10.3414/ME12-02-0004

Yardley, L., Spring, B. J., Riper, H., Morrison, L. G., Crane, D. H., Curtis, K., Merchant, G. C., Naughton, F., & Blandford, A. (2016). Understanding and Promoting Effective Engagement With Digital Behavior Change Interventions. *American Journal of Preventive Medicine*, *51*(5), 833–842. https://doi.org/10.1016/j.amepre.2016.06.015

Young, S. (2018). Agency and the Digital Alter Ego: Surveillance Data and Wearable Technologies. *International Journal of Sociotechnology and Knowledge Development*, *10*(3), 41–53. https://doi.org/10.4018/IJSKD.2018070103

Yu, C., Jannasch-Pennell, A., & DiGangi, S. (2011). Compatibility between Text Mining and Qualitative Research in the Perspectives of Grounded Theory, Content Analysis, and Reliability. *The Qualitative Report*, *16*(3), 730–744. https://doi.org/10.46743/2160-3715/2011.1085

Zhang, C., Shahriar, H., & Riad, A. B. M. K. (2020). Security and Privacy Analysis of Wearable Health Device. *2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC)*, 1767–1772. https://doi.org/10.1109/COMPSAC48688.2020.00044

Zhao, W., Shahriar, H., Clincy, V., & Bhuiyan, Z. A. (2020). *Security and Privacy Analysis of Mhealth Application: A Case Study*. 1882–1887. https://doi.org/10.1109/TrustCom50675.2020.00257

Zhou, T. (2012). Examining Location-Based Services Usage from the Perspectives of Unified Theory of Acceptance and Use of Technology and Privacy Risk. *Undefined*. https://www.semanticscholar.org/paper/Examining-Location-Based-Services-Usage-from-the-of-Zhou/d90f803b9496ee3aca7aee1c6c382862c2ebe9ea

APPENDICES

APPENDIX A: CODEBOOK FOR LABELING CATEGORIES

Category	Description	Keywords	Examples
Misuse of Data	mHealth users are concerned about the misuse of their collected data.	Misuse, inappropriate use, abuse of data, exploit	Another example could be the collection of data from smartwatches include blood pressure and heart rate. The misuse of this data or the inappropriate use of this data has the potential for harm.
Surveillance	Captures the tweets for mHealth users who are worried about being watched via their wearable devices.	Tracking, surveillance, location tracking, GPS	All wearables and connected clothing are hackable & surveillance vulnerable.
Personal Data Capture	Relates to whom and how the personal data is captured.	Capture, personal data, sensing, data transfer	Fitbit represents just another set of personal data (heart rate, sleep patterns, workouts etc.) that Google can capture from its users.
Control over Patient Apps	Captures tweets that show how concerned mHealth users are with their overall control over the	Data control, app settings control, access control	You may be the person wearing a given piece of "wearable technology" but that doesn't mean you control where the data goes or how it's used.

	health applications used.		
Control over Wearables	Relates to the degree of control a user has over their wearable devices.	Security control, data control, sensors, personal control, user control	Thanks to the glaring obvious security flaw in the futuristic Google glass wearable computer, a hacker could within minutes take control
Real Time Data Invasion	Captures the tweets for users who are concerned about the invasion of their privacy through the use of the real time data.	Invasion, real time data, invasion of privacy	Why do yall need location in order to sync Fitbit? Seems like an invasion of privacy.
Lack of Data Protection	Relates to the lack of data protection experience by mHealth users.	Data protection, privacy protection, protect privacy, unclear policies	Challenges for #wearables & hardware in healthcare: validity/reliability of measures, data protection, still unclear regulations.
Security Breach	Captures the different security breach and concern from the tweets.	Hack, data breach, vulnerable, data security	Wearables present several opportunities for a data breach. Most are relatively east to a hack a wearable with password-fingerprint ID security.
Third-party Data Access	Captures the tweets which highlight the concern of third-party data access.	Access, unauthorized, third-party access	Consider privacy in #wearable technology, who has access to the #data and for what purposes

Company Use of Data	mHealth users are concerned about how the different companies use the data acquired.	Company, Fitbit, Garmin, Apple, Google, trust, sensitive data, third-party.	I haven't allowed it and don't want to. Honestly, I don't trust Fitbit with this information and I don't feel like you need it.
Irrelevant (off-topic)	The posts that have no related content for any of the above categories.		Next in my buying list …. A wireless surveillance camera

APPENDIX B: WORD CLOUDS

Topic	Word Cloud	Topic	Word Cloud
T1: Collection of Personal Health Data		T2: Control over wearables	
T3: Control over apps		T4: Company Use of Data	
T5: Invasion of Real Time Private Data		T6: Capture of Personal Data	
T7: Third-party Access to Data		T8: Data and Privacy Protection	
T9: Security Breaches		T10: Data Surveillance