

8-7-2018

Classification And Analysis Of Mobile Health Evaluation Through Taxonomy and Method Development

Alan T. Yang
Georgia State University

Follow this and additional works at: https://scholarworks.gsu.edu/cis_diss

Recommended Citation

Yang, Alan T., "Classification And Analysis Of Mobile Health Evaluation Through Taxonomy and Method Development."
Dissertation, Georgia State University, 2018.
https://scholarworks.gsu.edu/cis_diss/67

This Dissertation is brought to you for free and open access by the Department of Computer Information Systems at ScholarWorks @ Georgia State University. It has been accepted for inclusion in Computer Information Systems Dissertations by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact scholarworks@gsu.edu.

PERMISSION TO BORROW

In presenting this dissertation as a partial fulfillment of the requirements for an advanced degree from Georgia State University, I agree that the Library of the University shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to quote from, to copy from, or publish this dissertation may be granted by the author or, in his/her absence, the professor under whose direction it was written or, in his absence, by the Dean of the Robinson College of Business. Such quoting, copying, or publishing must be solely for the scholarly purposes and does not involve potential financial gain. It is understood that any copying from or publication of this dissertation which involves potential gain will not be allowed without written permission of the author.

Alan Yang

NOTICE TO BORROWERS

All dissertations deposited in the Georgia State University Library must be used only in accordance with the stipulations prescribed by the author in the preceding statement.

The author of this dissertation is:

Alan Yang

35 Broad Street

Atlanta, GA 30303

The director of this dissertation is:

Dr. Upkar Varshney

Department of Computer Information Systems

J. Mack Robinson College of Business

Georgia State University

35, Broad Street

Atlanta, GA 30303

**CLASSIFICATION AND ANALYSIS OF MOBILE HEALTH EVALUATION THROUGH
TAXONOMY AND METHOD DEVELOPMENT**

By
ALAN YANG

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy
In the Robinson College of Business of
Georgia State University

GEORGIA STATE UNIVERSITY
J. MACK ROBINSON COLLEGE OF BUSINESS
2018

**Copyright by
Alan Yang
2018**

ACCEPTANCE

This dissertation was prepared under the direction of the Alan Yang's Dissertation Committee. It has been approved and accepted by all members of that committee, and it has been accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Business Administration in the J. Mack Robinson College of Business of Georgia State University.

Richard Phillips, Dean

DISSERTATION COMMITTEE

Dr. Upkar Varshney,	Chair
Dr. Aaron Baird,	Member
Dr. Gregory Gimpel,	Member
Dr. Jeong-ha (Cath) Oh,	Member
Dr. Karen Minyard,	External Member

Acknowledgements

I give thanks to all those that supported me during my time in graduate school. Without you this dissertation would not have been possible. Thanks to my advisor and dissertation chair Dr. Upkar Varshney, who was active and helpful throughout every stage of this PhD student's life. I could not have asked for a better mentor. Thanks to those who served on my committee: Dr. Aaron Baird, Dr. Gregory Gimpel, Dr. Jeong-ha (Cath) Oh, and Dr. Karen Minyard. The feedback all of you provided made this document possible and has helped me to grow as a scholar. Thanks to those that have written in support of me during my job search: Dr. Upkar Varshney, Dr. Aaron Baird, and Dr. Subhashish Samaddar.

Thanks to all the teachers that have helped and inspired me throughout my life. Dr. Rory McDonald, thank you for providing the spark which ignited the twin flames of research and teaching necessary for pursuing an academic career. Mrs. Anne Woolweaver, thank you for getting me interested in programming and logical thought. Mr. Ken Fontenot, thank you for showing me the value of critical thinking and the importance of being an active reader.

Thank you to all my fellow graduate students, the experiences and interactions we shared made these last few years not only productive, but enjoyable as well. Finally, thanks to my parents, my sister, and my extended family living in Taiwan and elsewhere in the world for their unwavering support.

Table of Contents

Abstract

Chapter 1. Introduction	1
Section 1.1 Dissertation Outline	1
Section 1.2 Problem Significance	1
Section 1.3 Research Question	2
Chapter 2. Literature Review	3
Section 2.1 Project Management	3
Subsection 2.1.1 Evaluation and Project Management	4
Subsection 2.1.2 Evaluation and Information Technology	5
Subsection 2.1.3 Evaluation and Healthcare	7
Section 2.2 Mobile Health	8
Section 2.3 Design Science	10
Section 2.4 Taxonomy	11
Section 2.5 Method	13
Section 2.6 Research Opportunities	14
Chapter 3. Methodology	16
Section 3.1 Justification as Design Science	16
Section 3.2 Contributions	24
Chapter 4. Taxonomy Design and Evaluations	25
Section 4.1 Taxonomy Design	25
Section 4.2 Taxonomy Evaluation	41
Subsection 4.2.1 Taxonomy Survey	41
Subsection 4.2.2 Cluster Analysis	45
Chapter 5. Method Design and Evaluations	60
Section 5.1 Constructs and Relationships	60
Section 5.2 Guidelines	68
Section 5.3 Method Evaluation	71
Subsection 5.3.1 Attribute analysis	71
Subsection 5.3.2 Method Survey	82
Section 5.4 Theorizing the Artifact	90

Chapter 6. Discussion	97
Section 6.1 Limitations	97
Section 6.2 Contributions	100
Section 6.3 Future work	102
Section 6.4 Conclusion	104
References	106
Appendices	119

List of Figures

Figure 1. Three Cycle View of Design Science Research	20
Figure 2. Taxonomy Development Method (Adapted from Nickerson et al. 2009)	26
Figure 3. 2016 UN Human Development Index Report	40
Figure 4. Optimal number of clusters analysis with $n = 80$	49
Figure 5. Cluster plot with $n = 80$, Cycle 1	51
Figure 6. Cluster plot with $n = 78$ (outliers 54 and 9 removed), Cycle 2	52
Figure 7. Cluster plot with $n = 104$, Cycle 3	54
Figure 8. Cluster plot with $n = 104$ and 3 clusters, Cycle 4	55
Figure 9. Activity Framework for Design Science Research (Adapted from Venable 2006)	92
Figure 10. Analytical Model for m-health Evaluation Selection	93
Figure 11. Summary of Research Process	101

List of Tables

Table 1. Design Science Research Framework	17
Table 2. Taxonomy Ending Conditions	27
Table 3. First Iteration of Taxonomy	28
Table 4. Second Iteration of Taxonomy	30
Table 5. Third Iteration of Taxonomy	31
Table 6. Fourth Iteration of Taxonomy	32
Table 7. Fifth Iteration of Taxonomy	33
Table 8. Sixth Iteration of Taxonomy	35
Table 9. Seventh Iteration of Taxonomy	37
Table 10. Attribute values for taxonomy	38
Table 11. Summarized Responses for Taxonomy Survey	42
Table 12. Taxonomy attributes of groups of interest ($n = 104$)	58
Table 13. Taxonomy attribute occurrences ($n = 104$)	60
Table 14. Counts Table for Cross-Attribute Observations in Taxonomy ($n = 104$)	73
Table 15. Frequency Table for Taxonomy Column Dimensions ($n = 104$)	74
Table 16. Frequency Table for Taxonomy Row Dimensions ($n = 104$)	75
Table 17. Observations of Interest for Column Dimension Frequencies	78
Table 18. Observations of Interest for Row Dimension Frequencies	78

Table 19. Means and Standard Deviations for Q1 – Q11 (n = 19)	85
Table 20. Summarized results for Q12-Q17 (n = 19)	88

List of Equations

Equation 1. Individual word frequency in a text	47
Equation 2. Conditional statement for Patient and Long Duration Relationship	63
Equation 3. Conditional Statement for Provider and Simulation Relationship	64
Equation 4 Conditional Statement for Simulation and Short Duration Relationship	64
Equation 5 Conditional Statement for Simulation and Short Duration Relationship	65
Equation 6. Conditional Statement for Technology-Based and Short Duration Relationship	65
Equation 7. Conditional Statement for Developed Country and RCT Relationship	66
Equation 8. Conditional Statement for Developing Country and Field Study Relationship	66
Equation 9. Conditional Statement for Patient and Long Duration Relationship.....	67
Equation 10. Conditional Statement for RCT and Patients Relationship	67
Equation 11. Conditional Function for Cross-Attribute Table Cell	72
Equation 12. Conditional Function for Frequency Table Across Column Dimension	72
Equation 13. Conditional Function for Frequency Table Across Column Dimension (Simplified)	76
Equation 14. Conditional Function for Frequency Table Across Row Dimension.....	77

Abstract

CLASSIFICATION AND ANALYSIS OF MOBILE HEALTH EVALUATION THROUGH TAXONOMY AND METHOD DEVELOPMENT

By

ALAN YANG

2018

Committee Chair: Dr. Upkar Varshney

Major Academic Unit: Computer Information Systems

This manuscript documents the creation and evaluation of a taxonomy for mobile health (m-health) evaluation and a method for m-health evaluation. M-health as a field within IS has seen significant amounts of growth in recent years due to improvements in technology leading to more affordable and portable computing power. The application of these technologies to the healthcare domain has created many new opportunities and benefits for patients and providers alike. This research seeks to study how these m-health projects are being evaluated and to determine what the characteristics of these evaluations are.

To accomplish this goal, the research process is conducted as design science and the research outputs of taxonomy and method are presented as design science artifacts.

The two artifacts are evaluated during their creation and once more afterwards to determine their utility. The taxonomy is created by collecting and analyzing documentation on m-health evaluation and using that information to generate descriptive categories by following a series of guidelines for creating a classification system. After evaluation of the artifact, a method is created for conducting m-health

evaluation. This method is a series of guidelines built upon constructs and relationships derived from the taxonomy.

Evaluation of the artifacts consists of expert surveys, cluster analysis, and attribute analysis. After evaluation of both artifacts, a descriptive theory explaining the selection of m-health evaluation types is created and presented. Theory development is based on the idea of kernel theories and their transferability to the information systems (IS) and design science domains. Contributions of this research are as follows: a classification system for m-health evaluation, a series of guidelines for individuals working on evaluations in the field of m-health, and a descriptive theory on the selection of evaluation type in an m-health context.

Keywords: Mobile healthcare, design science, project evaluation, health information technology

Chapter 1. Introduction

Section 1.1 Dissertation Outline

The dissertation is organized into 6 chapters. Chapter 1 introduces the problem being addressed by the work as well as the research question. Chapter 2 is the literature review which establishes the domain this work is set in. Chapter 3 is a description of the chosen methodology of design science and covers how the research process relates to established techniques for conducting design science research. Chapter 3 also covers the scope of the research and provides a list of potential contributions. Chapter 4 describes the creation and evaluation of the taxonomy artifact. Chapter 5 describes the creation, evaluation, and theorization of the method artifact. Chapter 6 is a discussion on the insights and contributions generated throughout the research process. The chapter also discusses limitations as well as avenues for future work in this domain. The final section of chapter 6 concludes the dissertation by circling the conversation back to the original problem described in chapter 1 and how the work has addressed it.

Section 1.2 Problem Significance

The field of mobile health (m-health) has grown significantly in the last ten years to become a multi-billion-dollar industry. Recent market reports have appraised the global m-health market at over \$10 billion USD (Ben-Zeev et al. 2015, Silva et al. 2015). The m-health Economics 2016 report predicts that the m-health application market alone will grow by 15% to reach \$31 billion by the year 2020, with the entire m-health market projected to be worth \$59.15 billion (Jahns et al. 2016). Despite the growth of industry, effective evaluation of m-health projects has remained an area of limited focus in the

research field. This research project seeks to create a meaningful classification of existing mobile health evaluation and guidelines for future evaluation of m-health projects.

Research has shown that lack of effective project management techniques leads to a higher rate of project failure (Dasgupta et al. 1972; Harberger 1972; Frechtling 2002). Many projects tend to emphasize the implementation stage at the expense of evaluation (Alexander and Faludi 1989; Michie et al. 2011; Steckler et al. 2002). Studies on the topic of evaluation in m-health either cover the topic on a case-specific basis or address the issue within a very specific medical context (Whittaker 2012; Kumar et al. 2013; Mookherji et al. 2015). As the field of m-health continues to grow, effective evaluation will continue to play a role both in determining and increasing the value of the field.

Section 1.3 Research Question

The primary research question is as follows: “What are the characteristics of m-health evaluations?” The main research contributions of this study are the organization of m-health project evaluation into a taxonomy and method for guiding evaluation of m-health projects and research.

The goals of this research project are to create a categorization system of m-health evaluation and to create actionable suggestions for researchers and practitioners in the m-health domain. The research is based in the literature of m-health, taxonomy development, and project management. The main contributions of the research are presented as design science artifacts.

Chapter 2. Literature Review

The review covers five areas of research: project evaluation, m-health, design science, taxonomies, and methods. From this survey, opportunities are identified and form the basis for the problem definition and proposed contribution of the research project.

Section 2.1 Project Management

Project management as a field of study in the behavioral sciences has existed for over fifty years. The project management discipline seeks to achieve an optimum balance between the competing constraints of cost, time, and scope (Gray 2008, Burke 2013). Projects that are managed effectively confer many benefits to stakeholders, including, but not limited to: better control of financial and human resources, shorter development times, lower costs, higher quality, improved productivity, higher internal coordination, and increased worker morale (Meridith and Mantel 2011, Kerzner and Kerzner 2017, Heagney 2016).

The Project Management Institute, established in 1969, created the project management body of knowledge (or PMBOK) guide as an educational standard for project managers (Duncan 1996). The bulk of the project management literature is based on the PMBOK guide and its division of a project into the stages of initiating, planning, executing, controlling, and closing (Duncan 1993). Most frameworks describing a project life cycle have adhered to these stages and commonly define a project as having a planning stage, an implementation stage, and an evaluation stage; a beginning, middle, and an end. Within the planning stage, the objectives and constraints are presented and analyzed, typically by a team planning the project such as a steering

committee. The actualization of the planning takes place in the implementation phase with evaluation usually occurring after implementation to measure the effects of the project and determine if the goals established during the planning stage were achieved (Schwalbe 2015).

Subsection 2.1.1 Evaluation and Project Management

The main goal of evaluation is to determine the success of a project (Baccarini 1999, Munns et al. 1996). Evaluation provides information to stakeholders on whether established goals are being met and generates new insights from project outputs (Frechtling 2002). Project managers and other stakeholders will often create metrics of success during the planning phases that will later be used to measure outcomes and performance during evaluation (Liu and Walker 1998, Cooke-Davies 2002). Individuals working in projects often have difficulty determining metrics that are both agreed upon by a majority of stakeholders and are an accurate representation of the stated goals of a project (Bryde 2003). This problem is compounded by the fact that metrics of success and stakeholder motivations have a propensity to change throughout the duration of a project and rarely remain static (Kinsella 2002).

In response to these issues, writings on evaluation during the last 20 years have focused on the connection between the evaluation and implementation stages of a project lifecycle. New techniques promote more evaluation taking place throughout the implementation phase. This is contrary to the more traditional role of evaluation only taking place after planning and implementation which often causes the changes that happened during those stages of the project lifecycle to be missed by the evaluation. Techniques such as developmental evaluation have emerged from the idea that the

three cycles of implementation, planning, and evaluation are inherently linked and should be conducted simultaneously and repeatedly throughout a project (Patton 1994, Bryde 2003).

Another crucial role the evaluation stage plays is communicating information to a variety of stakeholders. Projects need to provide some form of information about their impacts to effectively measure success. This data is then fed back into the planning loop to provide insights for future projects. The information being passed through the cycle needs to be standardized and made understandable to the relevant stakeholders involved for it to be meaningfully applied. The volume and quality of the information generated during this stage has been changing at an accelerated pace in the past 30 years (Blomquist et al. 2010). Because of this deluge of new information, more evaluation techniques are emerging for determining what metrics are important and calculating the appropriate level of granularity of project information (Cicmil et al. 2006). These trends are unlikely to reverse, growth in information availability has led to increased complexity in designing effective project management approaches (Svejvig and Anderson 2015).

Subsection 2.1.2 Evaluation and Information Technology

Information technology (IT) has changed the way many activities are conducted in business and project management is no exception. IT has introduced new systems and overlapping layers of complexity which demand new tools for analysis and measurement. Seminal papers on measuring success in IT identify the quality of the IT system, the quality of the information passing through the system, the usage of the technology, and the satisfaction of users as key factors for predicting the individual and

organizational impacts of an IT implementation (DeLone and McLean 1992). As the field of information systems developed in line with the growing technological capabilities of early IT systems, ideas such as the technology acceptance model and the role of the user became factors in measuring success (Delone and McLean 2003). Updated models include context and the intentions of users to utilize the technology as additional indicators of a system's effectiveness (Venkatesh and Bala 2008; Petter et al. 2008).

IT projects also have unique characteristics that differentiate them from projects that were conducted before the emergence of the information systems discipline. IT enables a myriad of different stakeholders to oversee and participate in a project with diverse technologies on a global scale (Schwalbe 2015). The increased capabilities brought about by inclusion of IT adds additional complexity and exacerbates the existing challenges of performing effective evaluation in a project setting (Pich et al. 2002; Cooke-Davies et al. 2007; Thomas and Mengel 2008). Analysis of projects utilizing IT have uncovered many challenges common to these projects such as user inexperience with new technologies, the absence of technical specifications, and rapidly changing scope and objectives (Tesch et al. 2007, Schmidt et al. 2001). To address these issues, project planning techniques including agile, rapid application development, extreme programming, and more emerged from the software development field (Beck and Gamma 2000, Beck et al. 2001, Wysocki 2011, Fernandez 2008). These methodologies espoused relatively fast iterations with built in evaluation cycles and frequent testing and communication between stakeholders as an answer to the changing landscape of evaluation. Recent studies have explored the idea of applying these project management techniques beyond the realm of software engineering (Conforto et al.

2014, Hobbs et al. 2017). As IT becomes increasingly ubiquitous, the challenges of managing technological complexity and evaluating effective usage of technology by users have become intertwined with the extant difficulties of the project management discipline.

Subsection 2.1.3 Evaluation and Healthcare

Healthcare as a discipline has existed for much longer than IT, but it also presents its own set of unique challenges to the evaluation domain. Technical challenges and complexities are exacerbated by the fact that for many stakeholders, the outcome of success is measured in impacts to quality of life and, depending on context, decisions made during the project can mean the difference between life and death for individuals (Viney et al 2002, Heeks 2006). When combined with IT, healthcare can be conceptualized as healthcare information technology (HIT). HIT carries all the difficulties inherent to IT including the complexity of systems, the importance of technology usage, and the necessity for information quality and relevance (Ammenwerth et al. 2003).

Research in the field of HIT has revealed that project risks and failures can be attributed to themes that are commonly observed in IT and project management. Such themes are as follows: failure to properly address context, gaps between requirements and their metrics, lack of communication between stakeholders, and underestimating the complexity of systems (Heeks 2006, Littlejohns et al. 2003, Ranmuthugala 2011).

Many healthcare projects overlap with the public health and policy fields. The implementation of these projects has caused an increased focus on the economic component of evaluations. Evaluation studies in the field of healthcare have examined the interaction between stakeholders in the health environment, assessed the variability

of cost effectiveness within the context of healthcare implementations, and determined the economic value of patient data (Sculpher et al. 2004, Curtis and Netten 2012, Drummond et al. 2005). Economic aspects of projects are often underexamined and are one of the reasons why many healthcare IT projects fail (Anderson 2010).

Section 2.2 Mobile Health

m-health is defined as any system enabled through a wireless infrastructure that provides healthcare to individuals in a manner that decreases spatial and temporal constraints (Varshney 2014; Varshney 2007). Other names for this phenomenon include E-health, wireless-health, and pervasive healthcare (Eysenbach and Group 2011). Common consumer applications of m-health include the use of smartphone applications to improve healthy behaviors or simple message service (SMS) to shorten communication times and coordinate tasks between individuals.

Ideal m-health environments fulfil certain characteristics. The well-being of the users within the system, particularly the patients, is generally agreed to be the top priority (Barton et al. 2012; Demiris et al. 2008). The next major issue is that of data security and privacy protection (Kotz et al. 2009; Doukas et al. 2010). The third qualification of an ideal m-health environment is that the various components of the system, particularly the devices operated by the users, are functional, reliable, and usable (Asangansi et al. 2010). Finally, the wireless infrastructure that the system is built upon needs to fulfil the same requirements of reliability and functionality to provide a consistent quality of service to users (Varshney 2014).

Definitions for the ideal m-health environment begin to diverge when expanding upon the four characteristics mentioned above. Some studies have argued for closer examination of the development cycle of m-health systems and that different attributes of the design team, such as transparency, can provide strong hints towards eventual outcomes (Mandl et al. 2009). Other studies have looked at the application-level and argue that characteristics of software such as context-awareness and data visualization are the means to interpret the success of a m-health system (Chang et al. 2011). Some researchers have taken an outcome-based stance and argue that the likelihood of positive modification of patient behaviors is the best representation of system quality (Blaya et al. 2010; Cameron et al. 2015).

Beyond the established criteria for what makes a good m-health system, another emerging phenomenon is the gap between m-health usage in developing countries and developed countries. Healthcare in high-income countries has seen an increase in both efficiency and patient well-being through applications of m-health technology (Bastawrous and Armstrong 2013). Most m-health applications fall underneath one of the following categories: Health promotion and behavior, sensors and peripherals, medical education and training, remote diagnosis, health monitoring, and provider-side communication. The potential benefits of successful m-health implementation are high, wireless communication provides a tool to quickly and informally transmit information between individuals, allows for remote locations to benefit from the functional and structural properties of a mobile network, opens the opportunity for low start-up costs and flexible payment plans to the public, and supports real time feedback for burgeoning decentralized health systems (Mechael 2009).

Section 2.3 Design Science

Design science research contributes to a knowledge base through the creation and evaluation of artifacts. Artifacts are described in the design science literature as conceptualizations of IT systems and their interactions with individuals and organizations (Orlikowski and Iacono 2001). Artifacts themselves can manifest in many ways, despite a common misconception that the research output of design science must take the form of either a hardware or software product. For instance, a research project that creates and tests a framework describing user reactions to an information system and another project that creates and tests software to make a work process more efficient would both be considered design science (Von Alan et al. 2004).

The common links joining together design science research are a foundation in the design literature and the shared purpose to address problems that have emerged from the complexity of human interaction with IT (Cross 2001; March and Storey 2008). These problems are often described as wicked problems, a concept that originated in the field of social policy that describes problems that differentiate themselves from others in their complexity, dynamism, novelty, and tenacity (Rittel and Webber 1973). Wicked problems do not present a single, clearly defined problem and likewise do not offer themselves up to a singularly applied solution. The design field seeks to address the issue of wicked problems by incorporating the study of the different relationships and concepts that make up the problem context into the process for crafting a solution (Buchanan 1992; Coyne et al. 2015).

The idea of the wicked problem precedes the design science school of thought. However, many comparisons have been made between wicked problems and problems that have arisen because of the growth of IT. IT systems often have many interconnected parts and problems which rarely have a one-size-fits-all solution that can be readily applied (Davenport et al. 1990; Orlikowski and Baroudi 1991; Brynjolfsson, 1993). Design science expands on this idea and offers artifacts as the building blocks to solutions for problems arising from systems with human and technological interactions. These solutions are meant to be address problems within those systems that have no single, universally applicable remedy at any given time (Pries-Heje and Baskerville 2008; Hevner and Chatterjee 2010).

Artifacts are design science's response to the growing complexity of problems appearing in the IT domain. They are created for the sole purpose of addressing a specific issue (De Leoz et al. 2018). These artifacts are then designed and evaluated which progresses the collective understanding of problematic phenomena and improves the possibility of discovering applicable solutions.

Section 2.4 Taxonomy

Taxonomy development originated from the field of biology to classify organisms based on a predefined structure of characteristics. There are two forms of analysis related to taxonomy building, phonetics and cladistics. Phenetics is the practice of clustering together organisms which are deemed to be similar though patterns in their attributes determined through observation. Cladistics looks at the evolution of organisms and creates groups based on a shared heritage (Sneath 1995).

A combination of these two modes of grouping create the basis for sorting entities into a taxonomic categorization scheme. Taxonomies were created to organize different types of organisms, but soon saw use in other disciplines, including the social sciences. As usage spread, terminology began to branch and a new form of classification system called a typology began to emerge. A major guideline for the application of taxonomies is that they must be distinguished from typologies (Bailey 1994). The key difference between the two tools is that taxonomies create their classifications based on patterns observed in empirical data while typologies form classifications first through a conceptual foundation, then observe data to see if those classifications fit. The two categorization strategies run the risk of being confused whenever theory application is involved in research (Follette and Houts 1996). In many areas, including project management and information systems, the terms typologies and taxonomy are frequently used interchangeably as ways to classify information (Marradi 1990, Rich 1992, Mitchel and Shortell 2000, Park et al. 2012).

In the field of project management, typologies have been used to classify the different types of projects and the management styles that are best suited for each (Shenhar and Dvir 1996, Krishnan and Ulrich 2001, Winter et al. 2006). Examples are numerous, and include typologies that examine topics such as technical complexity, scope, managerial control, worker morale, project setting, and more (Ross and Staw 1993, Shenhar and Dvir 1996, Evaristo and Van Fenema 1999, Hobbs and Aubry 2008).

In the fields of m-health and HIT, application of taxonomy has been used as a means of organizing literature or grouping together medical interventions into like categories (Waterlander et al. 2014, Plachkinova et al. 2015). Taxonomies have recently begun to see more usage in IS research. Articles in the past few years have seen its application in fields such as crowdfunding, consensus systems, and even design science itself (Haas et al. 2014; Nakatsu et al. 2014, Glaser 2015, Prat 2015). HIT and m-health has also seen recent usage of taxonomies as a way to classify technological devices or user interventions (Alrige and Chatterjee 2015; Sobrino and Bertrand 2017).

The field of health information systems contains many research studies claiming to utilize taxonomy as the guiding structure for analysis of information. However, most of these taxonomies are constructed using either ad-hoc or intuitive reasoning and lack a conceptual, theoretical, or empirical foundation informing their construction (Nickerson et al. 2013). In recent years, some studies in the field have taken an evidence-based approach to taxonomy construction in topics such as mobile security and economic evaluations (Abdullah et al. 2015, Brennan et al. 2006).

Section 2.5 Method

The term method has many connotations. In this research, methods are defined as a series of guidelines to aid individuals working within a specific domain. These methods can also be called suggestions, lists, or best practices depending on the field they are being used in. In project management, HIT, and m-health, methods have been created to better guide researchers and practitioners in their work.

Examples of methods include lists of best practices for generalized project management, guidelines for evaluating an information system, tips for managing a stakeholder circle, and suggestions for determining the optimal route for implementation of a healthcare intervention (Loo 2003, Cao and Hoffman 2011, Bourne and Walkter 2008, Agarwal et al. 2016). In research, methods exist for the development of theoretical models, application of qualitative methods, and the study of IT usability in healthcare contexts (Anderson 2005, Pope et al. 2002, Beuscart-Zephir et al. 1997). The guidelines posited by these methods are all meant to be used by a specific group of individuals and are meant to be both descriptive of the phenomenon common to their contexts and prescriptive for readers to take an informed course of action.

Section 2.6 Research opportunities

Evaluation as a concept in project management is well understood and the field of project management is well established. The intersection of general IT and project management has also been heavily studied, particularly in conjunction with the technological acceptance model and the impact that it has on the traditional project management process (Straub, Keil, and Brenner 1997, Hsu and Lin 2008, Davis and Venkatesh 2004). Despite the abundance of research in the project management and IT fields on evaluation, the HIT discipline, particularly the field of m-health, is still relatively new. The specific context of HIT projects creates research opportunities for the study of evaluation. Repeated calls have been made throughout the literature for more research focusing on the topic of evaluation to maximize the value from the growing field of m-health projects (Istepanian et al. 2006, Bourdreaux et al. 2014, Kumar et al. 2013, Eysenbach et al. 2011).

Within the field of m-Health research there is a lack of evidence proving the actual utility resulting from individual projects, particularly in low-income countries (Bastawrous and Armstrong 2013). This lack of evidence-based research creates both an opportunity along with a sense of urgency for researchers in the field. As mobile technologies become more widely adopted worldwide, individuals and organizations may attempt to implement mobile health-related projects utilizing technology without a set of guidelines and standards that were created from objective analysis of existing m-health implementations.

Preliminary attempts have been made at creating a taxonomy for making sense of the rapidly changing m-health environments, but the studies do not incorporate a theoretical component or go beyond the planning phase to the implementation and evaluation phases of m-health project management (Olla et al. 2015, Plachkinova et al. 2015). An opportunity exists within the m-health field for the creation of a classification system that can both relate to theory and test existing projects. Creation of such a taxonomy can lead to practical contributions in the development of new m-health project management techniques and can contribute to the IS literature by clarifying the information that currently exists. Motivated by these issues uncovered during the literature survey, the taxonomy and method designed in this work are the vehicles by which the contributions to the literature are delivered.

Chapter 3 Methodology

Section 3.1 Justification as Design Science

Design science is how the research will seek to address these questions uncovered from the survey of the literature. Beyond providing the definitions for a solution, design science can help to identify and address difficult problems directly. The challenges associated with evaluation of m-health projects meets the description of a wicked problem as defined by Weber 1973. This research attempts to address the problem of effective m-health project evaluation which involves both human and technological factors within a complex environment with no one solution that is easily applicable at any point in time (Ammenwerth et al. 2003).

The focus of design science is on the creation and evaluation of an artifact. Table 1 contains a representation of a design science research framework proposed by March and Smith. The table has two axes which describe research outputs and research activities. Research outputs are the artifacts in design science research and activities are conducted with individual artifacts created during a research process. Artifacts themselves can take many different forms, which can lead to confusion when determining what constitutes one. Classifications exist for simplifying the identification of individual artifacts by studying their characteristics. March and Smith, in a 1995 paper, proposed a categorization of artifacts in design science as constructs, models, methods, or implementations (March and Smith 1995). This categorization has been widely accepted and most artifacts are presented as one of these four forms (Gregor 2006; Peffers et al. 2007; Vaishnavi and Kuechler 2015; Baskerville, Kaul, and Storey 2015).

		Research Activities			
		Build	Evaluate	Theorize	Justify
Research Output (Artifacts)	Constructs				
	Model				
	Method				
	Instantiation				

Table 1. Design Science Research Framework (Adapted from March and Smith 1995)

Constructs are the most basic type of artifact and are definitions of concepts in a domain or system. These conceptualizations can have many names, such as entities, observations, attributes, or occurrences. They are meant to be representations of phenomena. By describing phenomena, constructs add to the vocabulary of their field and increase the clarity of concepts. Models are a combination of constructs and are created to divine the relationships between them. A model is meant to be a representation of reality which combines both descriptive and explanatory elements to elucidate the problems in a domain. Models are framed in a problem-solution perspective which adds to their utility as tools meant to explain and solve complex questions. Methods are a series of steps that serve as guidelines for performing a task. Methods are informed by the underlying constructs and models of the domain in which they are to be applied. Method artifacts serve as actionable solutions to the problems identified in the development of constructs and models. The final artifact instance is the instantiation, which is the application of an artifact in its problem environment. The development and evaluation of instantiations often leads to the formalization of its underlying components.

The four artifact types provide a common template for framing design science research contributions. March and Smith also define four activities associated with research: build, evaluate, theorize, and justify. Build and evaluate are specific to design science artifacts and are mandatory for determining the utility of the finished product of research. Theorize and justify are grouped as activities which belong to the natural sciences but which can be performed on artifacts after they are completed. The task of building is the creation of an artifact to address a problem or to perform a specific task. Prerequisites for building an artifact are an acknowledgement that a problem exists within a domain and that a solution is required to advance knowledge in that field. Evaluation of artifacts determines whether progress has been made in the research and whether the artifact has achieved the goals it was created to accomplish.

Once the artifact has been created, researchers can then theorize the artifact by attempting to explain phenomena observed during its creation and evaluation. Like the artifacts themselves, theorizing can take different forms. For instance, a method could have its underlying relationships and constructs examined and theory could develop by exploring the interactions and the reasons for them; An instantiation of an artifact could be theorized by applying it to a specific setting which leads to outcomes that can become the basis of a theory. Justification of theory is the testing of the viability of theories based on the artifact. Justification is meant to determine the validity and generalizability of theories through testing of its component parts.

The research activities comprising this project and its outputs are based upon this framework. The main outputs are a model and a method. Both will be built and evaluated and the method artifact will enter the theorization stage.

The two artifacts being created are a taxonomy of m-health evaluation and a series of guidelines for evaluation meant to aid individuals working in the domain of m-health. Taxonomy as a design science artifact is a relatively new concept. The idea was introduced by Nickerson, Varshney, and Muntermann in their 2013 paper establishing the guidelines for taxonomy development following design science principles (Nickerson et al. 2013). Summarized, the taxonomy development cycle states that the creation of a taxonomy must be started by a guiding meta-characteristic followed by multiple cycles of entity observations to create dimensions and attributes for classification. Mapped to the design science context, the meta-characteristic establishes the domain of the artifact and the characteristics of taxonomy entities leads to the creation of dimensions, which are comparable to construct artifacts. Grouping these dimensions together into a classification system of the taxonomy is equivalent to forming a model from constructs in the design science context (Nickerson et al. 2013).

After the taxonomy is created and evaluated, the method artifact will be built. As the method is comprised of a series of guidelines, it satisfies the definition of a method in March and Smith as a “a set of steps used to perform a task.” Guidelines are meant to be acted upon by individuals working in a domain related to the method artifact (March and Smith 1995). Method examples range from a series of guidelines for a practitioner to enact while performing a task to a research methodology for researchers to follow while studying a new phenomenon (Gregor and Hevner 2013).

The research follows the three stages proposed in Hevner 2007 of relevance, design, and rigor (Figure 1). Relevance comes from the literature review and the use of existing design science methods. The design cycle generates the proposed artifacts of

taxonomy and method. The rigor cycle draws upon existing work in the literature and creates a contribution to the literature in the form of the artifacts.

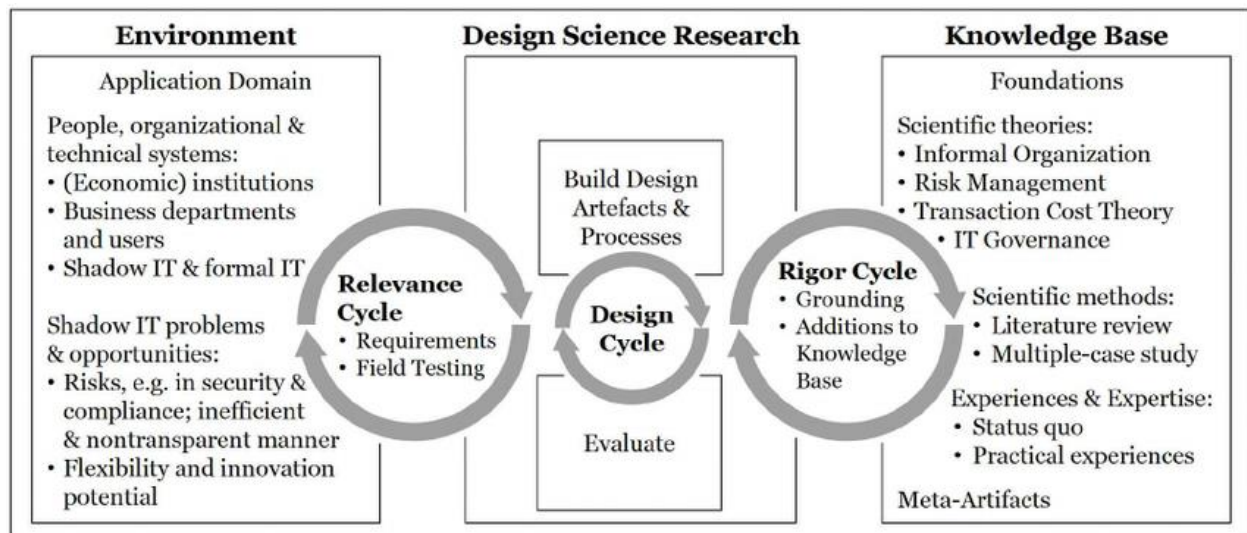


Figure 1. Three Cycle View of Design Science Research (Hevner 2007)

Design science principles stipulate that an artifact generated from research needs to be tested to ascertain its utility and subsequently generate a meaningful contribution from the research project. The evaluation process for both the taxonomy and method is informed by the definitions proposed in Venable et al. 2012 of three separate dimensions in evaluation: Artificial/naturalistic evaluation, formative/summative evaluation, and ex-ante/ex-post evaluation. The distinction between artificial and naturalistic evaluation is where the evaluation is to take place, in a controlled environment with little external influence (artificial), or less controlled environment with the possibility of larger external influences impacting the evaluation (naturalistic). Formative evaluation involves the testing of the artifact while it is in the process of being created. Summative evaluation is testing of the artifact after its creation. Ex-ante evaluation occurs before the artifact in question is implemented into study and practice,

the evaluation relies on incorporating existing data or information into a model. Ex-post evaluation examines the aftermath of an artifact implementation and looks at the resulting impact to determine utility based on the influence of the artifact on the outcome (Venable et al. 2016; Pries-Heje and Baskerville 2008). The evaluation strategy for the two artifacts is an artificial, ex-ante process that uses both formative and summative evaluation techniques. The evaluation techniques for both artifacts is as follows:

- Formative evaluation of taxonomy: Expert survey
- Summative evaluation of taxonomy: Cluster analysis of entity characteristics
- Formative evaluation of method: Attribute analysis of constructs and relationships
- Summative evaluation of Method: Expert survey

The conclusion of evaluation of both artifacts leads to the final research stage of theorizing. The method is the artifact of interest at this stage and the underlying assumptions informing the guidelines will be translated to constructs and relationships of a descriptive theory.

Adherence to Hevner 7-step process

Alan Hevner proposed in a 2004 paper seven guidelines for design science research. Mapping each guideline to a related task defines the methodology and clarifies the way in which the research is adherent to design science principles. The exercise of connecting parts of the research to the guidelines has an added benefit of creating an overview of each major milestone in the project. This research is undertaken with the understanding that simply following the 7-step process does not guarantee a worthwhile contribution. Hevner's steps further justify the research as design science but the

artifacts, evaluations, and contributions should be analyzed based on their individual merits.

Guideline 1: Design as an Artifact

The goal of this research is to create two artifacts as defined by March and Smith 1995: A taxonomy for m-health evaluation and a method for conducting m-health evaluation and research.

Guideline 2: Problem Relevance

Through the literature review of the proposal, m-health is demonstrated to be a timely problem with complex characteristics that make effective evaluation difficult (Tachakra 2003; Kumar et al.; Ammenwerth et al. 2003). The field of HIT has grown tremendously in the past twenty years and effective evaluation techniques and standards have been identified as key areas that need to be addressed if the discipline is to expand (Lazar et al. 2013, Agarwal et al. 2010). Current research shows that the field of m-health will not diminish soon. Existing problems will persist and most likely grow in the coming years.

Guideline 3: Design Evaluation

The Taxonomy will be evaluated formatively with an expert survey comparing iterations of the taxonomy and summatively by cluster analysis of the entities comprising it. The method will be evaluated formatively through attribute analysis of its underlying relationships and summatively with an expert survey to determine the aptness of the guidelines and relationships proposed in the method.

Guideline 4: Research Contributions

This research examines the field of m-health evaluation through the creation and evaluation of the taxonomy and method artifacts. Development and evaluation of the artifact draws upon the rigor and relevance cycles defined in Hevner 2007 and produces outputs to both the environment and knowledge base. Theorization of the artifact creates an analytical model for interpreting m-health evaluation phenomena. Practical contributions of this research are a better understanding of the role of evaluation in the information systems field and the possible development of new assessment practices and principles.

Guideline 5: Research Rigor

The primary motivation for the research is to examine the field of m-health evaluation through taxonomy and method development. The artifacts are created from data obtained through m-health project documentation. Rigor is demonstrated through adherence to procedures of taxonomy development and grounding in the fields of project management, HIT, and m-health for the design and evaluation of the artifacts.

Guideline 6: Design as a Search Process

The search process requires that design research follows a pattern seeking constant feedback and exploration of phenomena to generate insights. The three-step process of taxonomy development and the proposed research process are consistent with the definition of iterative design and evaluation. Development of artifacts will lead to evaluation which then feeds back into further development as revisions are made.

Guideline 7: Communication of Research

Communication of the research is in the form of a completed dissertation, followed by attempts to publish the findings from the design and evaluation process of the artifacts.

Section 3.2 Contributions

The goals of this research are as follows:

- Applying the field of taxonomy construction to m-health evaluation
- Examining and defining common phenomenon in m-health evaluation
- Suggesting and explaining relationships between common phenomena in m-health evaluation
- Creating guidelines for individuals creating evaluations in the domain of m-health
- Contributing to the field of design science through the design and evaluation of a taxonomy and a method for m-health evaluation
- Theorizing a method artifact and creating a descriptive theory on an aspect of m-health evaluation
- Furthering the IS discipline by applying ideas from project management and evaluation to the field of m-health

Boundaries are a necessity for defining the scope of any type of project and this research is no different. The work does not seek to accomplish the following:

- Expand the project management literature by introducing new ways to conceptualize or conduct evaluation
- Push the boundaries of research techniques such as cluster analysis, primary components analysis, or expert surveys
- Generate predictive theories on HIT or project management

Chapter 4. Taxonomy Design and Evaluation

The initial taxonomy was created from a survey of papers within the m-health field that were related to the field of project management. The creation of this taxonomy went through 4 iterations and some preliminary patterns were observed. After formative evaluation, the taxonomy underwent 3 more iterations and concluded on the seventh. The steps that were taken along with the conceptual backing guiding the research are presented below.

Section 4.1 Taxonomy Design

The creation of the taxonomy followed the three-level, seven-step indicator model described in Bailey 1984 and adapted in Nickerson et al. 2013 (Figure 2). The justification for use of this model is that it adheres the taxonomy development literature closely and is a practical means of classifying information. The final taxonomy was created through seven iterations of the development cycle. The explanation of the process for taxonomy creation is presented as a series of steps adhering to the development method.

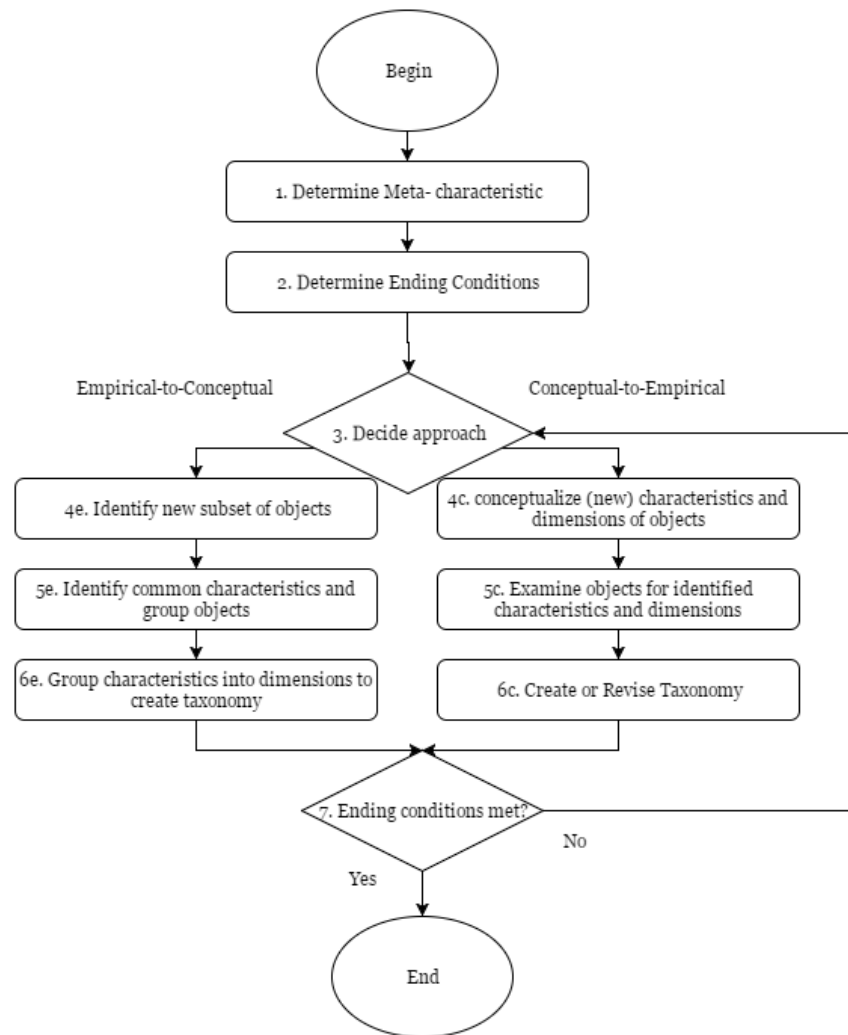


Figure 2. Taxonomy Development Method (Adapted from Nickerson et al. 2009)

Iteration 1, Step 1

The meta-characteristic of the taxonomy was determined to be: “project implementation and evaluation” within the field of mobile health. Establishing this aspect of the taxonomy prior to the analysis of the data serves a two-fold purpose. One, it prevents an ad-hoc search through a broad field of literature with the hope of stumbling upon a pattern. Two, it helps to guide the creation of the taxonomy past its genesis by creating a standard to which dimensions can be evaluated.

Iteration 1, Step 2

Ending conditions for the taxonomy were established before the papers from the sample were analyzed. Establishing the termination criteria prior to creation of the taxonomy prevents the situation where papers are added or removed because of reasons other than their relevance to the classification system. The criteria for ending conditions were evaluated based on the guidelines proposed in Nickerson 2013 of both subjective and objective ending conditions (Table 2).

Taxonomy ending conditions	
Objective criteria	Subjective criteria
1) All objects of a representative sample of objects have been examined 2) No object was merged or split in the last iteration 3) At least one object is classified under every characteristic of every dimension 4) No new dimensions or characteristics were added in the last iteration 5) No dimensions or characteristics were merged or split in the last iteration 6) Every dimension is unique 7) Each cell is unique	1) Taxonomy is concise 2) Taxonomy is robust 3) Taxonomy is comprehensive 4) Taxonomy is extendible 5) Taxonomy is explanatory

Table 2. Taxonomy Ending Conditions (Adapted from Nickerson et al. 2013)

Iteration 1, Step 3, 4e, 5e, and 6e

Step three involves a decision point indicating what type of approach is taken in creating new dimensions and characteristics for the taxonomy. We chose to follow the empirical-to-conceptual track for this first iteration to obtain a representative sample and to begin analyzing units to obtain initial dimensions.

At stage 4e, empirical data was obtained through a review of the literature. A search was conducted with the terms “Mobile health”, “Implementation”, and “Evaluation” on the academic databases Web of Science and IEEE explore from the year 2000 until November 2015. The search terms were decided based on the meta-characteristic determined in step 1. Combined, various permutations of the search terms yielded over one-thousand results. Papers were then further narrowed through identification of shared characteristics. Any papers that did not specifically deal with some form of mobile health implementation in a project setting were not considered for the taxonomy. Filtering at this level resulted in a vast decrease in the number of papers for consideration. 64 papers were identified as relevant and were the basis for our sample.

Focus		Biometric Monitoring		Implementation Obstacles		Duration	
Physician	Patient	Invasive	Non-Invasive	Hardware	People	0-6 Weeks	6-18 Weeks

Table 3. First Iteration of Taxonomy

At stage 5e, the papers were analyzed and themes and patterns of the various mobile health projects within the papers began to emerge. The emergence of these common characteristics leads to stage 6e, where the first dimensions and characteristics of the initial taxonomy (Table 3) were formed. Four dimensions were identified with two attributes each: Focus of study (Physician or patient), type of biometric monitoring (invasive or non-invasive), implementation obstacles (hardware or people), and duration of study (0–6 weeks to 6-18 weeks). An initial grouping of 10 papers were categorized by these dimensions.

Iteration 1, Step 7

A review of the initial taxonomy quickly revealed that it did not meet multiple ending conditions. Multiple dimensions had unpopulated characteristics and the dimensions themselves were too broad to reasonably provide any value in subsequent analysis. The cycle then progressed back to stage 3 and lead to the start of iteration 2.

Iteration 2, Step 3, 4c, 5c, and 6c

We decided on the conceptual-to-empirical approach for the second iteration. We believed our existing sample was representative of m-Health evaluation, but that our dimensions needed modification to properly portray the differences between entities. Step 4c involved multiple changes to the existing dimensions and addition of two new dimensions (Table 4).

“Focus” was seen as too broad and changed to “Determinant of success”, with the characteristics modified with the word “outcomes” rather than simply “provider” or “patient”. “Implementation Obstacles” was modified and the characteristics made more specific. The dimension of “Biometric monitoring” was determined to be an inadequate dimension and subsequently deleted from the taxonomy as M-health applications with a body-invasive hardware component were not represented in the sample. “Study type” was a newly added dimension as the initial analysis of the papers revealed that projects fell into one of two categories, “simulation” or “field study”. Simulations focused on the more theoretical application of m-health while field studies recorded the influence of implementation of a m-health technology in a live setting. Finally, a new dimension was added titled “Implementation type” with characteristics of “Developing” and “Improving”. The dimension determines whether m-health projects developed a new information system or were expanding upon an existing one. Finally, the dimension of duration was

altered from “weeks” to “months”, as nearly all papers described a timeline of research and evaluation lasting for over a 6-week period.

Determinant of Success		Implementation Type		Primary Obstacle to Implementation		Duration		Study Type	
Provider – outcomes	Patient – outcomes	Develop	Improve	Tech – centric	People – centric	0-6 Months	6-18 Months	Simulation	Field Study

Table 4. Second Iteration of Taxonomy

Iteration 2, Step 7

The taxonomy did not meet all ending conditions as new dimensions were added and other dimensions modified during the iteration. This led to another cycle and a return to stage 3.

Iteration 3, Step 3, 4c, 5c, and 6c

The conceptual-to-empirical approach was used again to further clarify the existing dimensions for iteration 3 (Table 5). Stages 4c and 5c saw the addition of two new dimensions, “Theory based” and “Project setting”. The theory dimension tracked if m-health projects with certain objectives utilized theory in the application of their solutions. Project setting had the characteristics of “Developed country” or “Developing country” and was meant to serve as a guideline for where projects tended to take place depending upon their goals. The final change was a clarification of the “implementation type” dimension into the “Impact on existing healthcare system” dimension. A minor change in the duration column resulted in “6-18 months” becoming the broader category of “over 6 months”. This change was made as continued analysis of sample documents

revealed evaluations that lasted over a year and a half. The final change was the removal of the “Primary obstacle to implementation” dimension. Projects were initially classified as either impeded by “hardware” or “people”. However, many projects did not neatly fit into either category, as both issues with technology and people arose and became hindrances to successful project completion.

Determinant of Success		Theory Based		Impact on Existing Healthcare System		Duration		Study Type		Project Setting	
Provider – outcomes	Patient – outcomes	Yes	No	Incremental Improvement	New System	0-6 months	6+ Months	Simulation	Field Study	Developed Country	Developing Country

Table 5. Third Iteration of Taxonomy

Iteration 3, Step 7

The taxonomy did not meet all ending conditions yet as new dimensions emerged from a review of the sampled literature. This led to another cycle and a return to stage 3.

Iteration 4, Step 3, 4c, 5c, and 6c

The fourth iteration saw the addition of two new dimensions, “Type of network” and “Application usage”. Type of network was divided into “3G or older” or “4G or newer”. Application usage was a simple “yes” or “no”. The addition of these dimensions added a new layer of description for the projects, the dimensions can be seen on Table 6.

Type of Network	3G or older
	4G or newer
Application usage?	Yes
	No
Determinant of Success	Provider – outcomes
	Patient – outcomes
Theory Based	Yes
	No
Impact on Existing Healthcare System	Incremental Improvement
	Introduction of New System
Duration	0-6 weeks
	6-18 weeks
Study Type	Developed Country
	Developing Country

Table 6. Fourth Iteration of Taxonomy

Iteration 4, Step 7

The taxonomy saw two new dimensions during this iteration. The original research cycle saw the end of iterations at this stage. Formative evaluation took place during iterations 4, 5, and six. New issues were identified and additional iterations were conducted.

Iteration 5, Step 3, 4c, 5c, and 6c

The type of network was determined to be an irrelevant description of a m-health evaluation. The change here marked a shift away from the taxonomy describing m-health projects and towards describing m-health evaluations themselves. In line with this new direction, “Application usage” was also dropped, as it was more of a

description of the nature of an m-health project rather than the evaluation of one. A new attribute of “Randomized Control Trial” was added to the study type dimension and the duration dimension was made more robust by having no upper limit, the attributes were adjusted to “0 to 6 months” and “Over 6 months”. Finally, a new dimension of “Evaluation Stages” was added to provide more nuance to the temporal aspect of evaluations. The fifth iteration can be seen on table 7.

Determinant of Success	Provider - outcomes
	Patient - outcomes
Theory Based	Yes
	No
Impact on Existing Healthcare System	Incremental Improvement
	Introduction of New System
Duration	0 to 6 months
	6+ months
Study Type	Simulation
	Randomized Control Trial
	Field Study
Project Setting	Developed Country
	Developing Country
Number of Stages	1 - stage
	2- stages
	3- stages or more

Table 7. Fifth Iteration of Taxonomy

Iteration 5, Step 7

The taxonomy saw the removal of a dimension, the addition of an attribute, and the inclusion of a new dimension. The changes led to another cycle and a return to stage 3.

Iteration 6, Step 3, 4c, 5c, and 6c

In iteration 6, we sought to describe evaluations with more detail and to remove some of the broader dimensions. The dimension of “Impact on existing healthcare system” was removed as it was determined to be too broad. Two new dimensions were added, “Primary evaluation metrics” with the attributes of “Cost-based”, “Technology-based”, and “User-based”, and “Theory application” with the attributes of “Exists” or “Does not exist”. To align the taxonomy more with evaluation concepts, the dimension of “Determinant of Success” was changed to “Primary Stakeholders” and the attributes changed to “Patients” and “Providers”. The sixth iteration can be seen on table 8.

Primary Stakeholders	Providers
	Patients
Theory Based	Yes
	No
Duration	0 to 6 months
	6+ months
Study Type	Simulation
	Randomized Control Trial
	Field Study
Project Setting	Developed Country
	Developing Country
Number of Stages	1 - stage
	2- stages
	3- stages or more
Primary Evaluation Metric	Cost-based
	Technology - based
	User-based
Theory application	Exists
	Does not exist

Table 8. Sixth Iteration of Taxonomy

Iteration 6, Step 7

A dimension was removed and two new dimensions were added. The changes led to another cycle, iteration 7, and a return to step 3.

Iteration 7, Step 3, 4c, 5c, and 6c

Iteration 7 was originally meant to be the concluding iteration confirming the changes made in iteration 6. However, formative evaluation of the taxonomy between iteration 6 and 7 led to the removal of the “Number of stages” evaluation. While understanding the number of stages throughout an evaluation was important, it was closely linked to the dimension of evaluation duration. An “Evaluation conclusion” dimension was also added to provide more information on the results of the evaluation. Iteration seven of the taxonomy can be seen on table 9. Two dimensions were also renamed for clarity. “Primary Stakeholders” became “User Stakeholders” and “Study Type” became “Evaluation Type”. An additional step was taken here to update the taxonomic entities. The search process for m-health evaluation documents was repeated with the end date expanded to 2017. This resulted in 40 more papers being added for a total of 104 entities.

Iteration 7, Step 7

The removal and addition of a dimension precipitated one more iteration. Iteration 8 saw no changes made to the taxonomy and after a check of the ending conditions, a determination was made that the taxonomy at the end of iteration 7 would be the final taxonomy used for summative evaluation.

User Stakeholders	Providers
	Patients
Theory Based	Yes
	No
Duration	0 to 6 months
	6+ months
Evaluation Type	Simulation
	Randomized Control Trial
	Field Study
Project Setting	Developed Country
	Developing Country
Primary Evaluation Metric	Cost-based
	Technology - based
	User-based
Theory application	Exists
	Does not exist
Evaluation Conclusion	Positive
	Negative
	Neutral

Table 9. Seventh Iteration of Taxonomy

The final taxonomy contained 7 dimensions and 17 attributes informed by 104 observations of m-health implementation. The attribute values for the taxonomy can be seen on table 10. Appendix item A1 contains citations for all documents included in the taxonomy.

User Stakeholders	Patients	65
	Providers	39
Evaluation Type	Field Study	42
	RCT	17
	Simulation	45
Evaluation Setting	Developing Country	30
	Developed Country	74
Main Metric	Cost-based	2
	Technology - based	23
	User-based	79
Evaluation Duration	0 to 6 months	83
	6+ months	21
Evaluation Conclusion	Positive	68
	Negative	2
	Neutral	34
Theory-application	Exists	4
	Does not exist	100

Table 10. Attribute values for taxonomy

Throughout the iterations, new dimensions and attributes were created as ways to describe the entities. What follows is a description of the dimensions and attributes in the final taxonomy.

User Stakeholders: Patient or Provider

Stakeholder management is an integral component of project management. The mobile health projects all had some component of end-user stakeholder identification within the studies. These end-users would be the individuals who would utilize the finished m-health product being developed by the project.

Type of Study: Simulation, Randomized Control Trial, Field Study

Study type was the technique that the m-health project was being evaluated with during the project. Field studies encompassed any open deployment of the finished m-health product into an environment where individuals utilized the technology. Randomized control trials were more controlled releases of the m-health component being developed to at least two separate groups. The efficacy of the m-health being evaluated was then determined through a comparison between a group of individuals utilizing the m-health to a control group. Finally, simulations modelled the potential effectiveness of the m-health item and provided analyses concluding whether an actual deployment of the product would be beneficial to stakeholders.

Evaluation Setting: Developed Country, Developing Country

Every m-health project included in the taxonomy identified where the m-health was intended to be deployed for usage. Depending on characteristics of the deployment setting, projects would show marked differences in the way that they were designed and evaluated. The most striking difference was whether the country was a developing or developed country. To determine where countries fall between these two categories, we used the metric of the world bank's calculation of human development index, or HDI. (World Bank 2017). Countries that were ranked 'high' or 'very high' were considered developed countries. Countries ranked 'medium' or 'low' were considered developing countries. None of the m-health evaluations took place in a country with data unavailable. A world map of the HDI index this category was based on is available on figure 3.

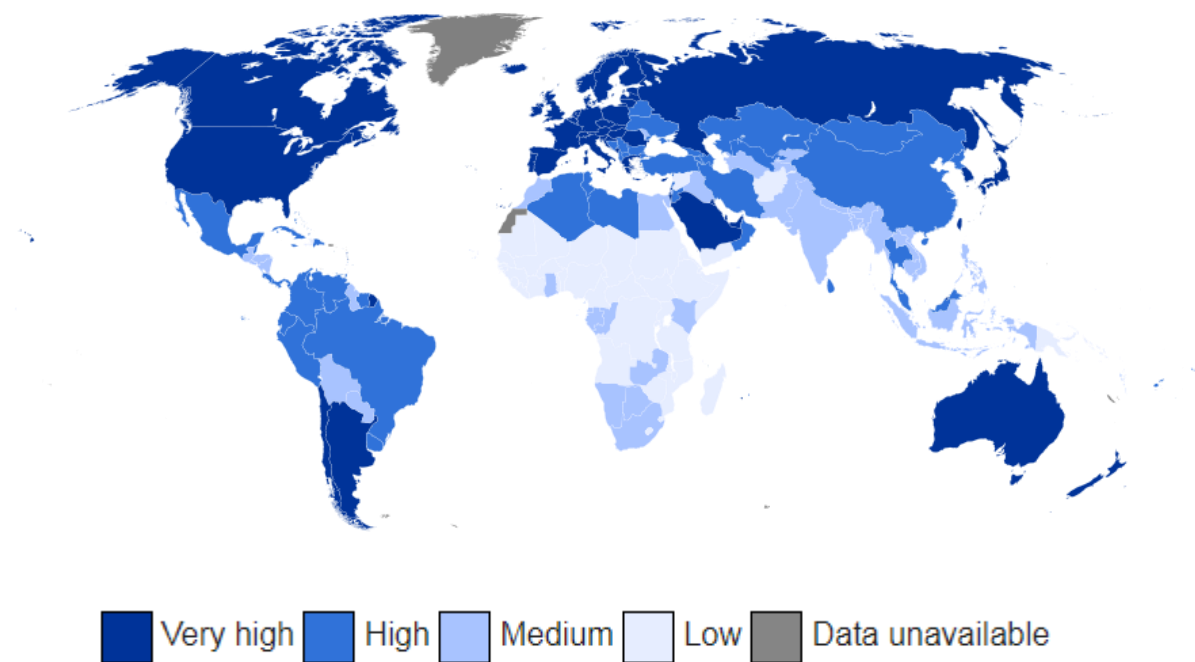


Figure 3. 2016 UN Human Development Index Report (Source: Wikimedia Commons)

Primary Evaluation Metric: Cost-centric, User-centric, Technology- centric

Traditional views of project management adhere to the idea of scope, quality, and cost. For m-health projects, quality is represented by the functionality and usability of the information technology being developed from the project. For instance, we found that not all projects in the sample mentioned scope management as a defining metric for project success, but many focused on stakeholder satisfaction and usage of the developed m-health intervention. Based on the sample, we defined the three main evaluation metrics as being based on either user, cost, or technology.

Evaluation Duration: 0 – 6 Months, 6+ Months

Documented evaluation durations ranged from as short as a few weeks to over a year. A median point of 6 months divided evaluations into short and long-term.

Evaluation Conclusion: Positive, Negative, Neutral

We looked at the overall conclusion of the evaluation and whether the m-health project being evaluated met the criteria for success established by the evaluating team. The characteristics we chose to represent this dimension were a positive result, negative result, or a neutral result.

Theory Application: Exists, Does Not Exist

This dimension addressed whether any type of theory was used in the project. The theories did not have to be in the information systems field, any incorporation of existing theory into the evaluation, or development of new theory resulting from the evaluation, would warrant a project's inclusion in the "Yes" characteristic of this dimension.

Section 4.2 Taxonomy Evaluation

The taxonomy evaluation was conducted in two stages. The first was a formative evaluation done during the creation of the taxonomy. This evaluation was administered as an expert survey. The second evaluation was a summative cluster analysis of the taxonomic entities after the taxonomy was finalized.

Subsection 4.2.1 Taxonomy Survey

Formative evaluation took place during the taxonomy design processes and informed the transition from iteration six to iteration seven of the taxonomy. An expert survey was used as the instrument to conduct the evaluation. Expert surveys are a documented source of qualitative evaluation, and are commonly deployed in design science contexts as a means of ex-ante, formative evaluation prior to the completion of the artifact (Gregor and Hevner 2013). Surveys can also reveal the sentiment of individuals

towards the artifact and can generate feedback that informs future design cycles (Parsons and Wand 2008, Cleven et al. 2009).

The survey for the taxonomy can be seen at appendix item A2. A critique of earlier taxonomy iterations was that they were too focused on m-health rather than m-health evaluation. One of the objectives of the survey was to determine whether the dimensions and attributes of later iterations of the taxonomy were more representative of m-health evaluation rather than general m-health. Adjustments were made between taxonomy iterations five and six, an interim taxonomy (Taxonomy A) was presented along with the fourth iteration of the taxonomy (Taxonomy B) for comparison. Taxonomy A differs compared to iteration six of the taxonomy in that it includes a new dimension titled “Evaluation Conclusion”. In the taxonomy design section, this dimension was not formally added until iteration seven. The survey was sent to 25 individuals working in the m-health or information systems field. There were 20 responses in total. Table 11 shows the questions and the number of responses for each sentiment.

	Response	Occurences	Notes
Q1	A	11	Which taxonomy is preferable, A or B?*
	B	5	
Q2	Yes	10	Do the taxonomies cover all aspects of mHealth evaluation?
	No	10	
Q3	Yes	7	Are there any unncessary dimensions?
	No	13	
Q4	Yes	11	Are there any additional dimensions you can think of?
	No	9	

*4 respondents had no preference

Table 11. Summarized Reponses for Taxonomy Survey

Question one is meant to determine whether Taxonomy A was more representative of mobile health projects compared to taxonomy B. While five respondents preferred B, eleven preferred A for classifying mobile health projects. Most responses preferring A mentioned how it referenced evaluation directly in the dimensions and differentiated between different evaluation concepts such as stakeholders and metrics rather than utilize broad definitions such as “impact on existing healthcare system”. Responses that preferred B mentioned that it was more parsimonious compared to A and that it dealt more closely with information technology and the technical side of mobile health. Another response mentioned that it was easier to understand at a glance compared to taxonomy A. Four responses showed no preference for either A or B. The unifying trend of these four responses was that insufficient context or information was provided in the survey to make an informed comparison. One response suggested that more dimensions would be better.

Question two relates to the taxonomy ending condition of exhaustiveness. Responses were split, ten believed the taxonomy covered all relevant dimensions and characteristics while ten had suggestions or were unsure. Those that did not respond positively made suggestions supporting their response. Common suggestions were to include more context or to include dimensions pertinent to mobile technology such as location based metrics. Multiple responses asked for additional nuance in the dimensions, such as more descriptive theory application, an expansion of the stakeholder dimension to include types of patients and types of providers, and a more granular representation of project setting such as differentiating between rural and urban settings in addition to developed and developing countries.

Question three was meant to identify any extraneous dimensions and to help streamline the taxonomy. Seven respondents thought both taxonomies were fine, 13 respondents gave suggestions. Two respondents identified that the two dimensions referencing temporal aspects of the evaluation in taxonomy A, number of evaluation stages and duration of evaluation, were redundant. Three commented on the wording of “Theory based” and “Theory application”, but none opposed the inclusion of the dimension itself. “Type of network” and “Impact on existing healthcare system” in taxonomy B were both identified as either unnecessary dimensions or non-generalizable descriptions of m-health evaluations.

Question four served as an extension of question two and is meant to generate possible dimensions that did not exist before. Nine respondents did not make suggestions and 11 did. The suggested dimensions tended towards the technical side of m-health, requesting more dimensions differentiating between mobile technologies and user interaction with devices. Some responses also suggested more dimensions to add additional context for better classification. Suggestions included the following: more descriptions of the projects, additional dimensions reflecting project team composition, and an expansion of the taxonomy to include additional user context such as culture.

Question five is meant to determine general user sentiment towards the taxonomies and the survey itself. Two responses mentioned was that the presentation of the taxonomies was confusing and that the research goals were not clear. Three other responses offered additional suggestions and elaborated on the responses to questions two through four.

Analysis of the responses revealed that the dimensions listed in taxonomy A were generally more apt for m-health evaluation compared to those in taxonomy B. However, many revisions were suggested for the dimensions that existed and comments were made to include additional dimensions to achieve a finer view of the projects being classified. Multiple respondents also expressed confusion regarding the presentation of terms and concepts in the survey. The feedback received during this stage led to iteration seven of the taxonomy which saw the elimination of one temporal dimension and the formal inclusion of the dimension of “Evaluation conclusion”. Multiple dimensions pertaining to the technical aspects of m-health were considered based on suggestions, but ultimately were not included because they deviated from the taxonomy meta-characteristic of m-health evaluation.

The feedback on the survey itself also informed the research process. The subsequent survey for the method evaluation was heavily influenced by feedback received during this process.

Subsection 4.2.2 Cluster Analysis

Summative evaluation of the taxonomy took the form of cluster analysis. Clustering is an empirical method of data analytics that began to appear in the social sciences literature starting in the late 1960s (Wilmink and Uytterschaut 1984; Lorr 1983). During that time, work was emerging on applying classification systems to the social sciences in the form of taxonomies, typologies, and other classification systems (Sneath and Sokal 1973; Johnson 1967; Blasfield 1976). The then-newly established technique of data clustering was considered a logical step towards classification and was further refined in applications involving categorization (Orloci 1967; Wallace and Boulton 1968;

Clifford and Stephenson 1975). As computing power expanded, cluster analysis broadened from classification and became a tool that could be combined with ever-increasing computing power to analyze large sets of data (Michalaski et al. 1983; Gauch 1980; Rose et al. 1990).

Clustering remains a technique that is relied upon to address the prevalence of large datasets in research by grouping together similar data points by following a set of rules, also known as a clustering algorithm (Hair et al. 1998; Jain et al. 1999; Webb 2003). While the technique first gained prominence in the computer science field, clustering is now commonly employed in the information systems field as a data analytics method (Berkhin 2006; Kaufman and Rousseeuw 2009; Prat et al. 2015). Clustering was chosen as the evaluation technique because of its history as a method both for classifying information and generating observations from a dataset. In the context of taxonomy, clustering appears to be an appropriate method as it seeks to create like groups that exist naturally within data.

Data from taxonomic entities was used as the basis for cluster development. Each entity within the taxonomy corresponds to a document explaining the m-health project and its subsequent evaluation. These documents took the form of either research papers or project reports. A word count was performed on each individual paper to determine the number of occurrences of words in a text. All documents contained a different number of total words, so occurrences alone were not a sufficient representation for a dimension, for instance, two documents could both contain 200 occurrences of the word “mobile”, but the first document may have 5000 words while the second has 10,000. To address this, the total number of words in a document was also recorded to calculate

individual word frequency. Finally, certain words that had high occurrence rates across multiple papers but were deemed irrelevant were removed. Examples of such words include: “One”, “Two”, “et”, “al”, “however”, and “because”. A list of all words removed during the clustering iterations is available in appendix item A3. Equation 1 is was used for calculating individual word frequency in a document:

$$\nu = o / (\tau - \iota)$$

Equation 1. Individual word frequency in a text

The following is a key for each symbol in the equation:

- ν - word frequency
- o – occurrences of word in text
- τ - total words counted in a text
- ι – all occurrences of irrelevant words in text

Using this equation, the frequency of all words was calculated. The next step was determining which words to include in the clustering. For the first analysis, the 172 most common words across 80 documents were used. The reasoning for this is that a word needs to appear at least in two texts to form clusters of any meaning. By taking the most common words across all documents, there is a higher likelihood of observances forming interpretable clusters (Steinbach et al. 2000; Berry and Castellanos 2004; Hotho et al. 2005).

The clustering itself was conducted in the statistical tool R. At this stage, a decision was made to use k-means clustering for analysis of the data. K-means clustering organizes

data around some number of clusters provided to the algorithm. The algorithm then determines where each data point is in relation to the different means and assigns the individual data point to the cluster of the mean it is closest to (Huang 1998; Xu and Wunsch 2005; Jain et al. 2010). The strengths of this technique are that it can determine clusters more naturally through the iterative assignment by distance and results from this method lend themselves to more direct interpretation because of the technique's relative simplicity compared to other clustering methods (Webb 2003, Berkhin 2006, Gan et al. 2007). Weaknesses of k-means include a higher propensity to overvalue outliers and noise because of its assignment algorithm and the requirement of selecting the number of clusters before the analysis is conducted (Wagstaff et al. 2001; Kanungo et al. 2002; Likas et al. 2003).

Compared directly to other clustering algorithms, K-means offers some direct benefits. Three other popular clustering techniques are: mean shift clustering, density based spatial clustering (DBSCAN), and hierarchical agglomerative clustering. Mean shift clustering requires the selection of a window-size or radius for selection space prior to the analysis. Given the relatively small size of the dataset, choosing an appropriate sizing would have been difficult compared to choosing many clusters through k-means. DBSCAN is affected when cluster sizes are of greater varying density. Given the nature of the data, small size with high dimensionality, this weakness with the DBSCAN algorithm would have been exacerbated. Finally, hierarchical clustering was not selected because we do not assume that a natural hierarchical structure exists within the data itself. The dimensions of the clusters throughout the hierarchical aggregate process would not be readily interpretable. Ultimately, K-means was selected over other

clustering techniques based on its perceived applicability to the word frequency dataset and potential for more direct interpretation of results.

Before k-means clustering could begin, two inputs were required, the number of clusters to form and the number of iterations to run. The number of iterations was capped at 50. This number was determined by the literature on k-means clustering applied to the context of the taxonomy dataset. The K-means technique tends to arrive at data convergence after a short number of iterations, argued anywhere between 10 to 25 (Fraley 1998). A consensus on the methodology is that any results after the first 20 iterations do not display much variation and any changes are relatively small compared to earlier iterations (Alsabti et al. 1997, Pelleg and Moore 2000). Combined with the nature of the taxonomy datasets used for clustering which each had a relatively small number of observations, 80 and 104, a decision was made to cap the number of iterations at 50 for all runs.

The average silhouette method was used to choose the optimum number of clusters. Silhouette is a cluster validation technique that measures the similarity of a data point to the cluster it belongs to against the difference that point has compared other clusters (Rousseeuw 1987). The range of the silhouette value ranges from -1 to 1. Values closer to one represent clusters where data points are closer to other data points within their cluster and farther from points outside of their cluster (Kaufman and Rousseeuw 2009). Results of the average silhouette calculation is available on **figure 4**. The code used in R to determine the optimal number of clusters, cluster data, and plot results is available in appendix item A4.

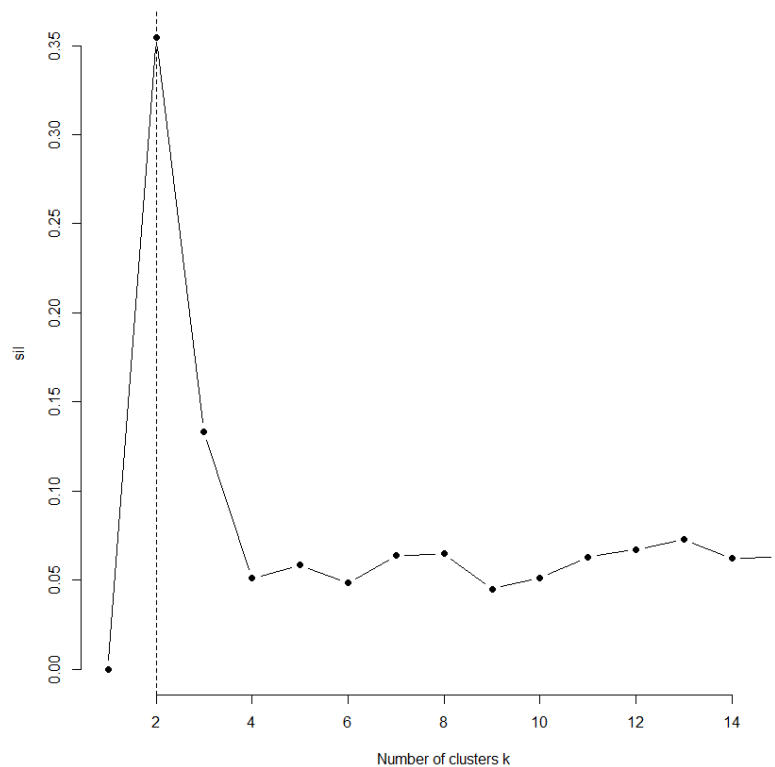


Figure 4. Optimal number of clusters analysis with $n = 80$

The optimum number of clusters for the first run was determined to be 2 with an average silhouette value of .3561. K-means clustering with 50 iterations and 2 clusters produced the results in appendix item a5. At this point in the process, principal component analysis (PCA) was selected to visually represent the clusters.

PCA is a statistical method for simplifying a dataset by reduction of dimensions through comparison of eigenvalues until only a few elements remain (Duntelman 1989). The remaining elements in PCA are responsible for the most variance in the dataset (Wold et al. 1987). PCA was chosen as the visualization method because of its value in reducing the total number of dimensions and ability to present a more concise

visualization of data, it is also commonly used in cluster analysis to reduce data dimensionality and visualize results (Scholkopf et al. 1997, Ding and He 2004).

The potential benefits of PCA were valuable given the context of the taxonomy dataset and the 172 separate dimensions used in clustering. Figure 5. contains the cluster plot generated through k-means and subsequent PCA of cycle 1.

The two dimensions represent the primary components generated from PCA of the clustering results. The percentages beside them represent the amount of variability explained by that principal component (Abdi and Williams 2010). Visual inspection of the plot shows that the two clusters are nearly completely overlapping. Additionally, two outliers, 54 and 9, seem to be strongly influencing the clustering results.

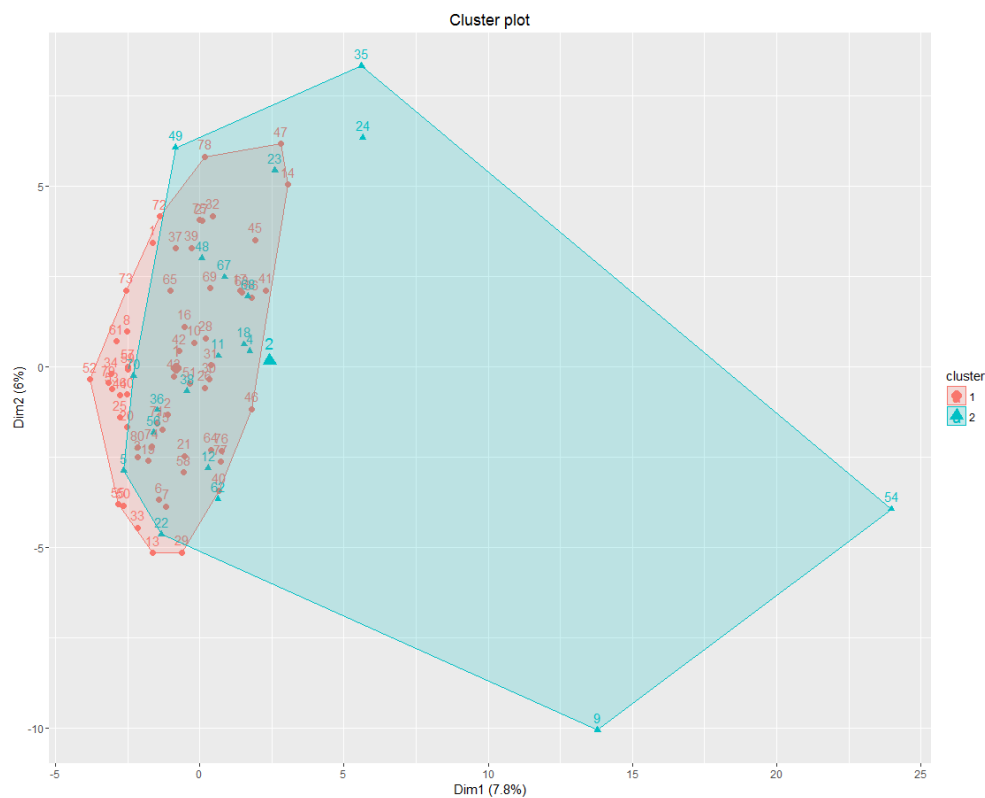


Figure 5. Cluster plot with n = 80, Cycle 1

The next cycle of analysis removed the outliers to achieve more distinct clusters. The number of dimensions remained the same at 172 and the number of observations was reduced to 78. Average silhouette analysis was performed again to determine the optimum number of clusters. The results of the silhouette analysis and k-means clustering are available in appendix items A6 and A7. Figure 6 contains the cluster plot generated through k-means and subsequent PCA of cycle 2.



Figure 6. Cluster plot with n = 78 (outliers 54 and 9 removed), Cycle 2

The results from the second cycle revealed that removal of the outliers made the clusters somewhat more distinct, but that many of observations were still overlapping between clusters. The amount of variance explained by the dimensions, and the clustering, was low. Two characteristics of this stage of the summative evaluation were identified as the source of these problems: the low observation size and the high number of dimensions for clustering. Cycle three saw two changes to the dataset to address these issues: the number of observations was increased by 24 to include all 104 taxonomic entities and the number of dimensions was reduced from 172 to 51. To achieve this reduction in dimensions, the taxonomy was used to create a list of words that related to attributes within the taxonomy. A total of 51 words were identified as matching or representative of taxonomy attributes. These words and their corresponding taxonomy attributes are listed at appendix item A8.

Cycle 3 of the cluster analysis then began with the new dataset of 104 values and 51 dimensions. Average silhouette analysis put optimum number of clusters at 2 and k-means analysis was run, results at appendix items a9 and a10. Figure 7 contains the cluster plot generated through k-means and subsequent PCA of cycle 3.

More variance is explained by the two primary components compared to previous cycles, which is consistent with the two changes we made, increasing observations and decreasing total dimensions. Visual inspection of the plot shows that while there is still overlap, cluster 2 has many more unique observations. Observations 54, 35, 24, appeared to be outliers in the data. However, instead of removing them outright, another k-means clustering was conducted with 3 clusters as opposed to 2, this became

cycle 4. The results of this 3-cluster analysis are in appendix item A11. Figure 89 contains the cluster plot generated through k-means and subsequent PCA of cycle 4.

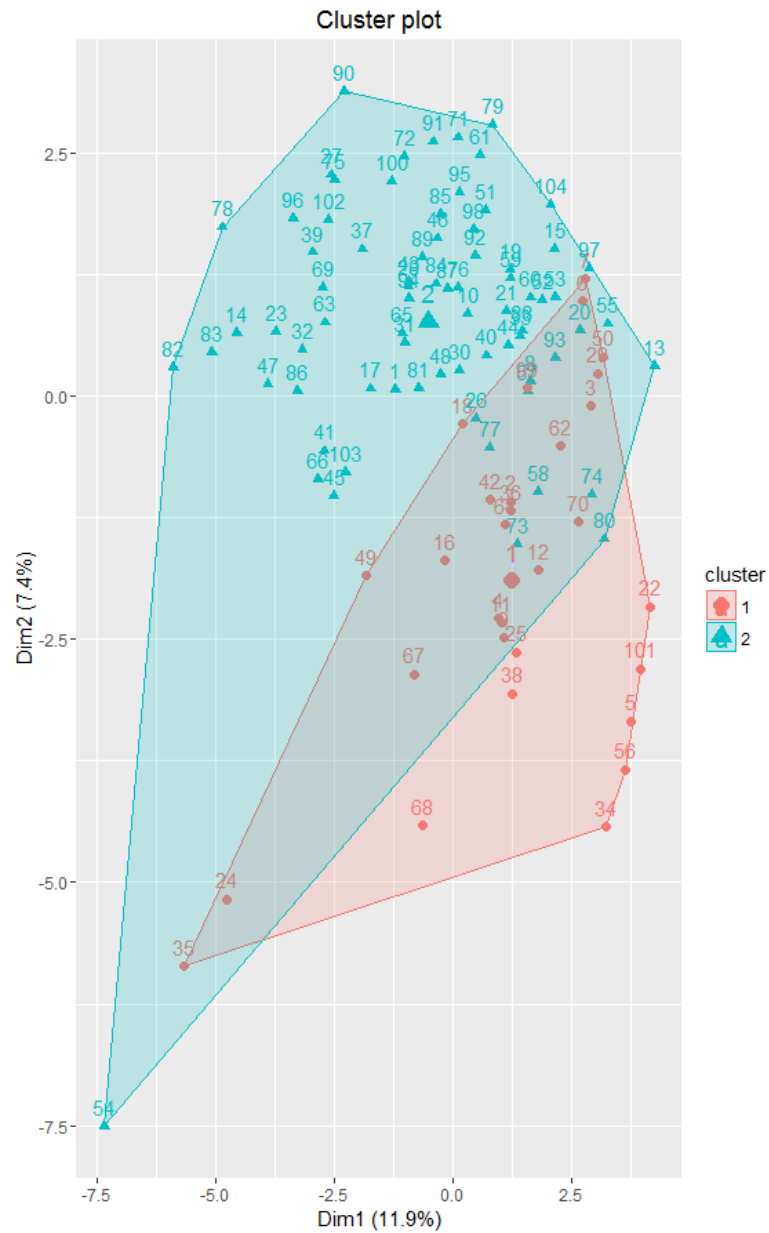


Figure 7. Cluster plot with n = 104, Cycle 3

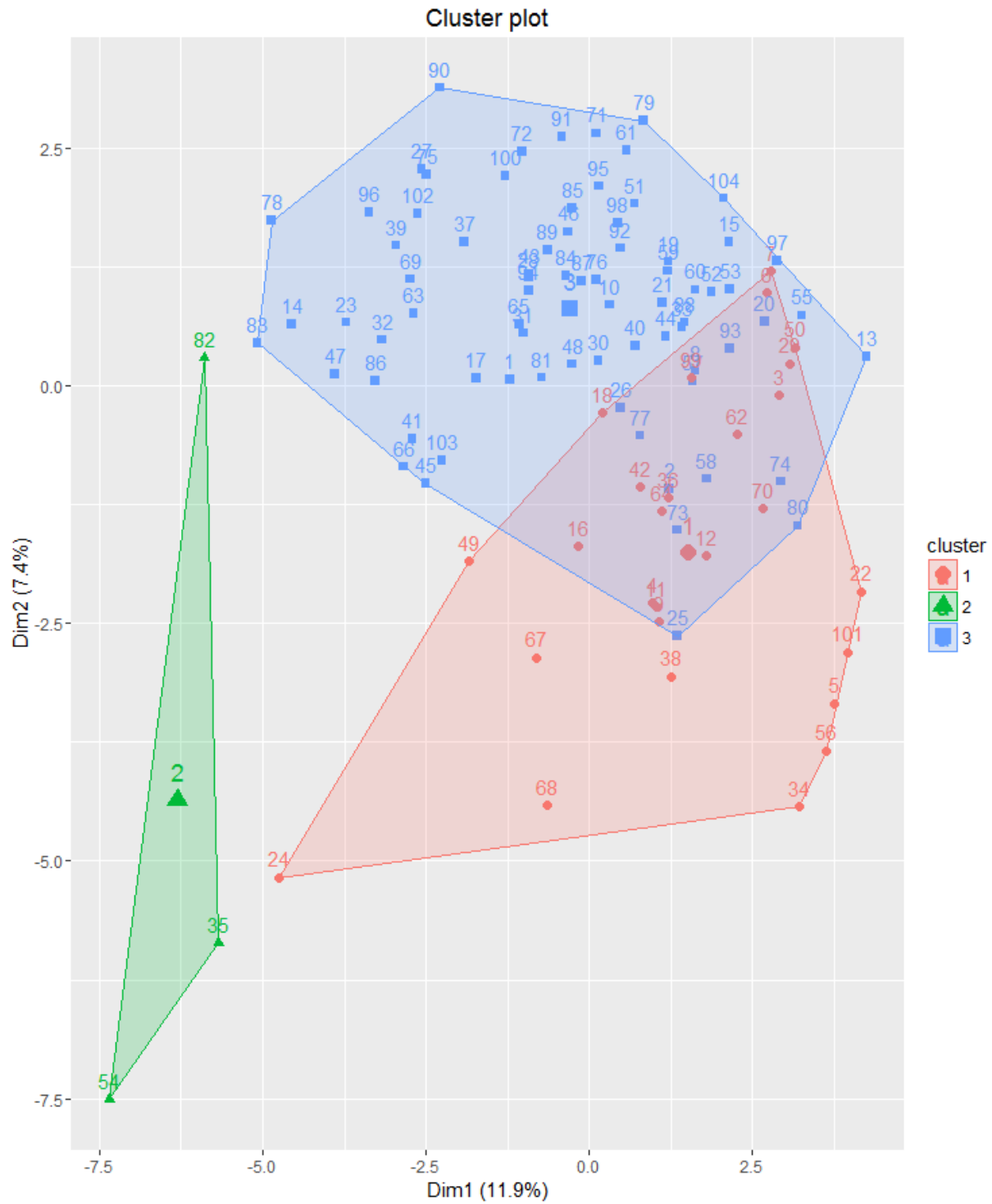


Figure 8. Cluster plot with n = 104 and 3 clusters, Cycle 4

Visual inspection of the cluster shows that observances 82, 35, and 54 were separated into their own cluster. Clusters one and three are the main groups that were formed from the previous cycle. The mean of cluster 1 still overlaps with cluster 3, indicating that there is still overlap between the two clusters.

Analysis of the results across cycles focuses on the “within cluster sum of squares by cluster” statistic, or “between sum squares” / “total sum squares”. The statistic of “total sum squares” represents the sum of squared distances between the mean of all observations to each data point. “Between sum squares” is the sum of squared distances between the mean of all observations to the mean of each cluster. Higher values indicate that observations are more spread and clusters are more distinctly separated, while lower values for this variable indicate that clusters are more compact and observations are closer together (Scholkopf et al. 1997). In general, the statistic measures the amount of variance in the data that can be explained by the clustering. Higher numbers of clusters and observations will also increase this statistic. The likelihood of the values of “between sum squares” and “total sum squares” being the same, 100%, increases as the number of clusters increases (Abdi and Williams 2010).

Cycles 3 and 4 had a “within cluster sum of squares by cluster” value of 11.4% and 19.1% compared to a value of 8.3% for cycles 1 and 2. This increase is not surprising given the changes to the dataset, an increase to the number of observations will generally increase the value of the statistic assuming the clustering method remains the same (Ding and He 2004). However, the overall values for these statistical measures is still comparatively low. These values could have been increased by prompting the algorithm with more clusters. However, the purpose of the evaluation is to determine the

aptness of the taxonomy in relation to a representation of the natural state of the field. Higher clusters would have increased the amount of variance explained by the clustering, but would not have produced any further insights. Additional application of alternative inputs or even other algorithms would not have eliminated the weaknesses inherent in the dataset of low sample size and high dimensionality. As a result, cycle 4 was the final clustering cycle and was used for interpretation of results.

Visual analysis of the graphs shows that there is overlap between clusters and that most of the observations are grouped tightly together. This phenomenon was likely a combined result of the data and the analytical method. In the context of cluster analysis, the datasets were relatively small (80 and 104 observations) with a high number of dimensions (172 and 51) compared to datasets typically studied with clustering. The tendency for clustering techniques to decrease in efficacy as dimensionality increases is an established phenomenon and has been dubbed the 'curse of dimensionality' (Indyk and Motwani 1998, Hinneburg and Keim 1999; Houle et al. 2010; Har-Peled et al. 2012). This issue is also common whenever word frequencies are used as dimensions and observational points (Delen et al. 2008; Tuffery 2011; Assent 2012). Another interpretation of the results is that the content and wording of the dataset documents are very similar to each other. This would reaffirm the need for approaches that delve into the nuance of differences between projects to determine the differences between m-health evaluation techniques.

Despite these setbacks, further observations were made when data points within the 3-cluster plot were observed. Four groups of data points were analyzed, those exclusive to cluster 1, those exclusive to cluster 2, those exclusive to cluster 3, and those that

belonged to both clusters 1 and 2. The attribute totals for each of these groups was calculated and can be seen on Table 12.

The results on table 12 show that some of the different groupings display unique characteristics. For the grouping exclusive to cluster 1, 10 out of 10 observations were provider-centric, simulation evaluations taking place in a developed country. 9 out of 10 evaluations had user-based metrics as the main determinant of success and were of short duration, lasting from 0 to 6 months. Keyword means with higher values in cluster one that support this finding were: “service”, “care”, “university”, “clinic”, “system”, “application”, and “project”.

		Exclusive to cluster 1	Exclusive to cluster 3	Shared clusters 1 and 3	Cluster 2
Main Stakeholders	Patients	0	44	20	1
	Providers	10	18	9	2
Evaluation Type	Field Study	0	25	16	1
	RCT	0	15	2	0
	Simulation	10	22	11	2
Evaluation Setting	Developing Country	0	18	12	0
	Developed Country	10	44	17	3
Main Metric	Cost-based	0	2	0	0
	Technology - based	1	12	9	1
	User-based	9	48	20	2
Duration	0 to 6 months	9	47	24	3
	6+ months	1	15	5	0
Evaluation Conclusion	Positive	7	43	17	1
	Negative	1	0	1	0
	Neutral	2	19	11	2
Theory-application	Exists	0	3	1	0
	Does not exist	10	59	28	3

Table 12. Taxonomy attributes of groups of interest (n = 104)

The data points in the group exclusive to cluster 3 were more varied. 15 of the 17 total RCT evaluations in the taxonomy are represented in this group along with all instances of cost-based evaluation and 3 out of the 4 evaluations with theory. The entity attributes for the group inside clusters 1 and 3 were understandably more mixed. The three

entities in cluster 2 had characteristics like those of the group exclusive to cluster one. Visual inspection of the plot shows points 35 and 54 would have been included into cluster 1, and the clustering separated them by grouping the two points with point 82. Distinct groupings are visible for short-duration, provider-centric simulation evaluations in developed countries as well as patient—centric RCT evaluations. Keyword means supporting this conclusion were as follows: “change”, “intervention”, “positive”, “study”, “monitor”, “assess”, “measure”, “feedback”, and “trial”. Because of the small size of cluster 2, meaningful interpretation of means in relation to clusters 1 and two is not reasonable.

From a statistical standpoint, the formative evaluation did not uncover strong empirical evidence for the existence of natural groupings in the dataset. However, observations made from the evaluation results shows that there are dimensions within the taxonomy that are useful categorizations for m-health evaluation. Overall, the results show that the m-health field is homogenous with some slight derivations that can be partially explained by the taxonomy. The next step is to study the dimensions and attributes within the taxonomy to generate additional observations for designing the method artifact.

Chapter 5 Method Design and Evaluations

Section 5.1 Constructs and Relationships

The goal of the method design and evaluation is to create a series of guidelines to help individuals to create better m-Health evaluations that generate value for project stakeholders. Researchers in the field of m-Health evaluation could also utilize the method to determine the state of evaluation in the field and what types of evaluation are overrepresented or underrepresented.

Design science methods are based on the underlying artifacts of models and constructs. The dimensions and attributes of the taxonomy were used as the basis for the constructs and relationships that informed the method. For convenience, the attribute occurrences for the 104 taxonomy entities is repeated on Table 13.

Main Stakeholders	Patients	65
	Providers	39
Evaluation Type	Field Study	42
	RCT	17
	Simulation	45
Evaluation Setting	Developing Country	30
	Developed Country	74
Main Metric	Cost-based	2
	Technology - based	23
	User-based	79
Duration	0 to 6 months	83
	6+ months	21
Evaluation Conclusion	Positive	68
	Negative	2
	Neutral	34
Theory-application	Exists	4
	Does not exist	100

Table 13. Taxonomy attribute occurrences (n = 104)

Informed by the evaluation and design cycles of the taxonomy, five observations and eight relationships based on the dimensions and their interactions were created.

Observations

1) Theory-based evaluations are uncommon in m-health evaluation

Out of the 104 evaluations in the taxonomy, only 4 explicitly referenced a theory as part of the design of the evaluation. Despite the importance of theory development and application in the IT literature, theory applications in the field of m-health evaluation remain sparse.

2) Patient and provider metrics are different

While this observation is implicit given the existence of the “User Stakeholders” dimension, the comparison between patient and provider metrics bears repeating as it predicates two of the eight relationships derived from the taxonomy.

3) Few evaluations report negative results

Out of the 104 evaluations in the taxonomy, only 2 reported a definitively negative conclusion. This finding is surprising given that most of the m-health literature reports that healthcare implementations are prone to failure in nearly half of all cases (Agarwal et al. 2010). A point worth noting is that most documents included in the taxonomy were academic articles. Although the taxonomy is specific to the field of m-health evaluation, this observation may be indicative of a wider trend in scientific reporting to shy away from research that reports negative results. This tendency towards positive reporting has been reported on and is not exclusive to m-health, IT, or even the social sciences (Fanelli et al. 2012).

4) User-based metrics are common in m-health evaluation

79 out of 104 evaluations focused on user-metrics as the primary determinant of project success. The prevalence of these types of metrics is consistent with discussions in the HIT field on the added burden of healthcare projects caused by the direct impact of decisions on the quality of life and well-being of individuals (Devaraj et al. 2013; Free et al. 2013).

5) Cost-based metrics are uncommon in m-health evaluation

Only 2 out of 104 evaluations focused on cost-metrics as the primary determinant of project success. This finding is surprising relative to the amount of technology-based metrics, which comprised 26 out of 115 evaluations. Many authors in the literature have called for closer examinations of the financial implications of m-health implementations and evaluations (Bardhan 2013). M-health and mobile technology have a proven capability to generate cost-savings by decreasing communication barriers and granting users access to information in new settings (Amadi-Obi et al. 2014). The fact that cost-specific evaluations are so underrepresented in the taxonomy may be an indicator that the focus of the field is still on the benefits m-health can bring to users through technological advancement rather than on economic feasibility of interventions.

Relationships

The relationships posited below are derived from a combination of the insights gained during the design and evaluation of the taxonomy and formative evaluation of the

method. A caveat is that these are general trends observed from the taxonomy data. The statements made are not meant to be predictive of all m-health evaluations. The relationships propose that the existence of a certain attribute in an evaluation can be an indicator that another attribute will likely be in the evaluation as well. These propositions are supported by data from the taxonomy and are accompanied with a brief, possible explanation for their occurrence. Relationships that mention the formative evaluation will be revisited during that section (5.3.1).

Each of the relationships are first described by a general statement which references a taxonomy dimension and attribute. Each relationship also has an associated conditional statement linked to it. For instance, the first relationship is that “Patient-focused evaluation often have longer durations.” The associated conditional statement is written out in Equation 2:

$$(Ea = \text{Patients FOR } D = \text{User Stakeholders} \rightarrow Ea = 6+ \text{ Months FOR } D = \text{Duration})$$

Equation 2. Conditional Statement for Patient and Long Duration Relationship

Where D is the dimension and Ea is the entity attribute. The components preceding the arrow (\rightarrow) are the determinants of the components after the arrow. Equation 2 can be read as follows: If an evaluation has attribute ‘patients’ for dimension ‘user stakeholders’, then the evaluation will likely have attribute ‘6+ months’ for dimension ‘duration’.

- 1) Provider-focused evaluations often use simulation

This relationship was derived from observations made in the cluster analysis. The group of 10 data points exclusive to cluster 1 all contained these two attributes. Possible

explanations for this relationship are the practice of using archival hospital data to facilitate simulation design, the design of prototype pilot studies that focus on improving hospital worker efficiency as a contribution, and the modelling of m-health efficacy as an argument for more widespread adoption (Aranda-Jan et al. 2014). Equation 3 shows the conditional statement for this relationship.

(Ea = Providers FOR D = User Stakeholders → Ea = Simulation FOR D = Evaluation Type)

Equation 3. Conditional Statement for Provider and Simulation relationship

2) Simulation evaluations often have shorter durations

This relationship was also drawn from the clustering results. The points exclusive to cluster 1 contained all simulations with only one evaluation lasting over 6 months. An explanation for this relationship is that simulation evaluations do not involve the coordination of as many elements as RCT or field studies and can be completed more quickly as a result (Denkinger et al. 2013). Equation 4 shows the conditional statement for this relationship.

(Ea = Simulation FOR D = Evaluation Type → Ea = 0 to 6 Months FOR D = Duration)

Equation 4 Conditional Statement for Simulation and Short Duration Relationship

3) Technology-based evaluations often use simulation

This relationship was uncovered during formative evaluation where it was found that simulations make up most of the evaluations focusing on technological metrics. An explanation for this relationship is that technical metrics lend themselves more easily to this type of evaluation as simulated machine-performance tends to be more consistent

with real-world performance compared to simulations of human behaviors (Conroy et al. 2014). Equation 5 shows the conditional statement for this relationship.

(Ea = Technology-Based FOR D = Primary Evaluation Metric → Ea = Simulation FOR D = Evaluation Type)

Equation 5 Conditional Statement for Simulation and Short Duration Relationship

4) Technology-based evaluations have shorter durations

This relationship was uncovered during formative evaluation where it was discovered that all evaluations focused on technical metrics took less than 6 months. The explanation for this relationship is likely related to relationship 2. The study of technical-metrics can be conducted faster compared to user or cost-metrics as it does not involve the coordination of as many elements (Afshin et al. 2016). Equation 6 shows the conditional statement for this relationship.

(Ea = Technology-Based FOR D = Primary Evaluation Metric → Ea = 0 to 6 months FOR D = Duration)

Equation 6. Conditional Statement for Technology-Based and Short Duration Relationship

5) Evaluations in developing countries do not tend to have RCT evaluations

This relationship was uncovered during formative evaluation which revealed a very low number of RCT evaluations set in developing countries. A possible explanation of this relationship is that, on average, there is less established infrastructure in developing countries compared to developed countries (Chib et al. 2015; Anderson et al. 2016).

This may cause difficulty for evaluators who want to conduct an RCT and may lead them to perform another type of evaluation. Equation 7 shows the conditional statement for this relationship.

$$(Ea = \text{Developing Country FOR } D = \text{Setting} \rightarrow (Ea = \text{Field Study OR } Ea = \text{Simulation}) \\ \text{FOR } D = \text{Evaluation Type})$$

Equation 7. Conditional Statement for Developed Country and RCT Relationship

6) Evaluations in developing countries tend to have field study evaluations

This relationship is related to relationship 5 and was also uncovered during the formative evaluation. Most of evaluations in developing countries were field studies. An explanation for this is that the infrastructure may not be present for consistent RCT evaluations and there is not enough prior data to conduct simulations (Braun et al. 2013; Dwivedi et al. 2016). These factors may cause evaluators to rely more upon field studies out of necessity rather than choice. Equation 8 shows the conditional statement for this relationship.

$$(Ea = \text{Developing Country FOR } D = \text{Setting} \rightarrow Ea = \text{Field Study FOR } D = \\ \text{Evaluation Type})$$

Equation 8. Conditional Statement for Developing Country and Field Study Relationship

7) Patient-focused evaluations often have longer durations.

While this relationship is particularly context-dependent, a trend was discovered during formative evaluation that most of the longer-duration evaluations had patients as user stakeholders. The reasoning for this could be that studying patient metrics can take longer than studying provider metrics for comparable m-health projects (Bert et al. 2014). Another likelihood is that m-health interventions involving patients often deal with patient health as a metric, which can take longer to properly evaluate compared to metrics focused on provider-side improvements (Arsand et al. 2012). Equation 9 shows the conditional statement for this relationship.

(Ea = Patients FOR D = User Stakeholders \rightarrow Ea = 6+ Months FOR D = Duration)

Equation 9. Conditional Statement for Patient and Long Duration Relationship

8) RCT evaluations often have patient stakeholders

This relationship was uncovered during formative evaluation. Most of the RCT evaluations in the taxonomy were focused on patients rather than providers. The history of RCT may explain this relationship, as the methodology was developed to test the medical effectiveness of a new treatment (Blackwood et al. 2010). The concept has been applied to different types of interventions, but is still commonly used, and understood, as a technique for measuring patient- metrics in the healthcare domain (Bowling 2014). Equation 10 shows the conditional statement for this relationship.

(Ea = RCT FOR D = Evaluation Type \rightarrow Ea = Patients FOR D = User Stakeholders)

Equation 10. Conditional Statement for RCT and Patients Relationship

Section 5.2 Guidelines

Based on the observations and relationships, a series of guidelines was developed to aid researchers and practitioners working in the field of m-health implementation. Each of these guidelines offers a suggestion for more effective and robust evaluation, but the best case will often vary depending on the context of individual projects. For instance, guideline 2 suggests not to overemphasize user and technological metrics at the expense of cost metrics. However, if a key project stakeholder only wants evaluation on user metrics and wishes for analysis to focus exclusively on them, then the guideline will have to be applied at the discretion of the evaluating team.

The guidelines were developed from a combination of the summative evaluation of the taxonomy and the formative evaluation of the method itself. Guidelines one and two were based on the clustering results, while guidelines 3 through 6 were based on results from the attribute analysis evaluation of the method.

Guideline 1: Basing your evaluation on theory can help distinguish it from others

Observations from the taxonomy reveal that theory-based evaluations are uncommon in the field of m-health. This guideline suggests the inclusion of theory to make individual evaluations stand out and better transition from field documentation towards an academic manuscript.

Guideline 2: Do not emphasize user and technological metrics at the expense of cost-based metrics

Very few evaluations from the taxonomy mentioned cost metrics, much less considered them a determinant of project success. User metrics being paramount in a healthcare

project makes logical sense, but the near absence of cost-considerations for most projects indicates that for many projects, longevity and scalability of results is not a key consideration. For the beneficial user impacts to continue after the cessation of the project, the feasibility for others to replicate the project is information that should be captured by the evaluation.

Guideline 3: Be wary of the following combinations of evaluation characteristics if your goal is to generate an academic contribution: Provider-centric evaluations in developed countries; technology-based evaluations using simulation; and short duration evaluations using simulation

A few areas of research in m-health evaluation are currently inundated with evaluation examples. To avoid going over covered ground, the three combinations of m-health evaluation attributes should be conducted with an understanding that there is a large amount of extant work in the field. Of note is that simulation evaluations appear in two of the three overrepresented groups. Simulations are a relatively easy means to complete an evaluation without a large upfront investment, the consequence is that many works in the field rely upon simulation as the evaluation methodology.

Guideline 4: Do not be afraid to report negative results

The evaluation conclusions in the taxonomy were largely positive or neutral, very few reported any sort of negative outcome. Multiple surveys of the literature would suggest that the success rates of m-health projects are not as high as reflected in the taxonomy (Chiasson et al. 2007). If an evaluation is not representative of reality, it loses value and only serves to confuse stakeholders (Cicmil et al. 2006). Negative outcomes can still form the basis for research, this idea has been demonstrated in the fields of project

management, IT, and HIT (Fichman et al. 2011). The reasoning for the lack of negative results may be the fledgling status of the m-health evaluation discipline. Academically, for the field to grow, negative outcomes should be reported and studied so that past mistakes are not repeated. Reporting negative results also gives managers a more transparent picture of the project and allows them to make better informed decisions.

Guideline 5: Allow enough time for evaluation

Many evaluations in the taxonomy were of shorter duration (between 0 to 6 months). The findings in these studies were frequently positive or neutral, but often did not examine the impacts of implementation fully. Another problem brought about by short duration evaluations is that the long-term effects of a healthcare intervention are often left unexamined. M-health evaluations can have immediate impacts, but documented cases exist where the full consequences, both positive and negative, of an m-health project are not made apparent until the system has been in place and used for an extended period of time (at least 6 months). Longer duration evaluations allow the different components of a project to be studied with more nuance and provides a better temporal perspective to stakeholders and any others reading the evaluation.

Guideline 6: Allocate more time if an evaluation involves users or is set in a developing country

Evaluations set in developing countries tend to require a longer duration evaluation. Those that involve the study of user metrics also tend to take longer compared to technical or cost-based metrics (Iribarren et al. 2017). These observations also reflect a relationship between study type and duration. Specifically, the fact that simulations are of generally shorter duration compared to RCT and field studies. As many m-health

evaluations place an emphasis on user metrics, allowing adequate time to study the impacts of an m-health project is often just as important as ensuring the successful planning and implementation of the project itself.

Section 5.3 Method Evaluation

The method evaluation occurred in two parts, a formative and summative stage. The formative evaluation examined the taxonomy attribute data and was meant to test the constructs and relationships underlying the method guidelines as well as to discover new relationships and add to the method artifact. Summative evaluation took the form of an expert survey which presented a series of statements for respondents to evaluate. The survey also presented respondents with a hypothetical situation where they could design and describe their ideal m-health evaluation within a simple context.

Subsection 5.3.1 Attribute Analysis

Formative evaluation of the method was an analysis of cross-attribute taxonomy data. Three tables were created from this evaluation (Table 14, Table 15, and Table 16). Each of the tables contains the full list of taxonomy attributes as the axes. The difference between tables are the contents of the cells. The cells in the first table (Table 14) contain a whole number representing the number of entities within the taxonomy containing both the Row Attribute (AR) and the Column Attribute (AC) relative to the current cell c . The square brackets represent an Iverson bracket which evaluates to 1 if the statement in the brackets is true and 0 if it is false. Equation 11 represents the cell contents as a conditional function where C represents Cell. It can be read as follows:

The value of any given cell is the number of all (104) taxonomy values that contain the row attribute and column attribute relative to the cell.

$$C = \sum_{n=1}^{104} n[ARc \text{ AND } ACc]$$

Equation 11. Conditional Function for Cross-Attribute Table Cell

The cells in table 15 contain a percentage. This value represents the prevalence of attribute combination AR and AC relative to other combinations containing AR for the Column Dimension (DC). Equation 12 represents the cell contents as a conditional function. For instance, to determine the value of the cell where AR is “Patients” and AC is “Field Study” in table 15, we find the corresponding cell value in table 14, 26. As the AC of the cell-of-interest is “field study”, the DC is the dimension of “Evaluation Type” which also includes the ACs of “RCT” and “Simulation”. The value of C in table 15 is determined by dividing the corresponding value of table one by the sum of all row attributes for the given column dimension, or $26/(26+15+24)$ which equals .4000 and is represented as a rounded percentage to the 2nd decimal of 40.00%. This means that 40% of all evaluations with patient user stakeholders used field studies to conduct evaluation.

$$C = \frac{\sum_{n=1}^{104} n[ARc \text{ AND } ACc]}{\sum_{n=1}^{104} n[ARc \text{ AND } (ACc \text{ FOR } DCc)]}$$

Equation 12. Conditional Function for Frequency Table Across Column Dimension

	Patients	Providers	Field Study	Randomized Control Trial	Simulation	Developing Country	Developed Country	Cost-based	Technology-based	User-based	0 to 6 months	6+ months	Positive	Negative	Neutral	Exists	Does not exist
Patients	x	x	26	15	24	17	48	1	9	55	50	15	43	1	21	1	64
Providers	x	x	16	2	21	13	26	1	14	24	33	6	25	1	13	3	36
Field Study			x	x	x	21	21	0	6	36	29	13	29	1	12	1	41
Randomized Control Trial			x	x	x	1	16	0	0	17	12	5	10	0	7	0	17
Simulation			x	x	x	8	37	2	17	26	42	3	29	1	15	3	42
Developing Country						x	x	0	4	26	21	9	21	1	8	0	30
Developed Country						x	x	2	19	53	62	12	47	1	26	4	70
Cost-based								x	x	x	1	1	1	0	1	0	2
Technology - based								x	x	x	23	0	13	0	10	1	22
User-based								x	x	x	59	20	54	2	23	3	76
0 to 6 months											x	x	52	2	29	3	80
6+ months											x	x	16	0	5	1	20
Positive													x	x	x	4	64
Negative													x	x	x	0	2
Neutral																0	34

Table 14. Counts Table for Cross-Attribute Observations in Taxonomy (n = 104)

	Patients	Providers	Field Study	Randomized Control Trial	Simulation	Developing Country	Developed Country	Cost-based	Technology-based	User-based	0 to 6 months	6+ months	Positive	Negative	Neutral	Exists	Does not exist
Patients			40.00%	23.08%	36.92%	26.15%	73.85%	1.54%	13.85%	84.62%	76.92%	23.08%	66.15%	1.54%	32.31%	1.54%	98.46%
Providers			41.03%	5.13%	53.85%	33.33%	66.67%	2.56%	35.90%	61.54%	84.62%	15.38%	64.10%	2.56%	33.33%	7.69%	92.31%
Field Study						50.00%	50.00%	0.00%	14.29%	85.71%	69.05%	30.95%	69.05%	2.38%	28.57%	2.38%	97.62%
Randomized Control Trial						5.88%	94.12%	0.00%	0.00%	100.00%	70.59%	29.41%	58.82%	0.00%	41.18%	0.00%	100.00%
Simulation						17.78%	82.22%	4.44%	37.78%	57.78%	93.33%	6.67%	64.44%	2.22%	33.33%	6.67%	93.33%
Developing Country								0.00%	13.33%	86.67%	70.00%	30.00%	70.00%	3.33%	26.67%	0.00%	100.00%
Developed Country								2.70%	25.68%	71.62%	83.78%	16.22%	63.51%	1.35%	35.14%	5.41%	94.59%
Cost-based											50.00%	50.00%	50.00%	0.00%	50.00%	0.00%	100.00%
Technology-based											100.00%	0.00%	56.52%	0.00%	43.48%	4.35%	95.65%
User-based											74.68%	25.32%	68.35%	2.53%	29.11%	3.80%	96.20%
0 to 6 months													62.65%	2.41%	34.94%	3.61%	96.39%
6+ months													76.19%	0.00%	23.81%	4.76%	95.24%
Positive																5.88%	94.12%
Negative																0.00%	100.00%
Neutral																0.00%	100.00%

Table 15. Frequency Table for Taxonomy Column Dimensions (n = 104)

	Patients	Providers	Field Study	Randomized Control Trial	Simulation	Developing Country	Developed Country	Cost-based	Technology-based	User-based	0 to 6 months	6+ months	Positive	Negative	Neutral	Exists	Does not exist
Patients			61.90%	88.24%	53.33%	56.67%	64.86%	50.00%	39.13%	69.62%	60.24%	71.43%	63.24%	50.00%	61.76%	25.00%	64.00%
Providers			38.10%	11.76%	46.67%	43.33%	35.14%	50.00%	60.87%	30.38%	39.76%	28.57%	36.76%	50.00%	38.24%	75.00%	36.00%
Field Study						70.00%	28.38%	0.00%	26.09%	45.57%	34.94%	61.90%	42.65%	50.00%	35.29%	25.00%	41.00%
Randomized Control Trial						3.33%	21.62%	0.00%	0.00%	21.52%	14.46%	23.81%	14.71%	0.00%	20.59%	0.00%	17.00%
Simulation						26.67%	50.00%	100.00%	73.91%	32.91%	50.60%	14.29%	42.65%	50.00%	44.12%	75.00%	42.00%
Developing Country								0.00%	17.39%	32.91%	25.30%	42.86%	30.88%	50.00%	23.53%	0.00%	30.00%
Developed Country								100.00%	82.61%	67.09%	74.70%	57.14%	69.12%	50.00%	76.47%	100.00%	70.00%
Cost-based											1.20%	4.76%	1.47%	0.00%	2.94%	0.00%	2.00%
Technology-based											27.71%	0.00%	19.12%	0.00%	29.41%	25.00%	22.00%
User-based											71.08%	95.24%	79.41%	100.00%	67.65%	75.00%	76.00%
0 to 6 months													76.47%	100.00%	85.29%	75.00%	80.00%
6+ months													23.53%	0.00%	14.71%	25.00%	20.00%
Positive																100.00%	64.00%
Negative																0.00%	2.00%
Neutral																0.00%	34.00%

Table 16. Frequency Table for Taxonomy Row Dimensions (n = 104)

Because of the layout of the table, the denominator for equation 12 can be simplified. The sum of all attributes containing AC for DC and AR is the same as the sum of all attributes containing AR. For the previous example, the value derived for the denominator in the equation is 75. This value corresponds to the total number of entities in the taxonomy with a patient attribute for the main stakeholder dimension (Table 13). A simplified equation is provided in equation 13.

$$C = \frac{\sum_{n=1}^{104} n [ARc \text{ AND } ACc]}{\sum_{n=1}^{104} n [ARc]}$$

**Equation 13. Conditional Function for Frequency Table Across Column
Dimension (Simplified)**

The cells in the third table (Table 16) also contain a percentage. This value represents the prevalence of attribute combination AR and AC relative to other combinations containing AC for the Row Dimension (DR). Equation 14 represents the cell contents as a conditional function. This equation is similar to equation 12, with a minor change to the denominator.

As an example, to determine the value of the cell where AR is “Patients” and AC is “Field Study” in table 16, we start by finding the corresponding cell value in table 14, 26. Because the AR of the cell-of-interest is “Patients”, the DC is the dimension of “Main Stakeholder” which also includes the AR of “Providers”. We then come to the value of C by dividing the corresponding value in table one by the sum of all values sharing AC and have an AR corresponding to the DR. This results in the calculation of $26/(26+16)$ which equals .61904 and is represented as a rounded percentage to the 2nd decimal of

61.90%. This percentage means that among all Field Studies, 61.90% had patients as user stakeholders.

$$C = \frac{\sum_{n=1}^{104} n [ARc \text{ AND } ACc]}{\sum_{n=1}^{104} n [ACc \text{ AND } (ARc \text{ FOR } DRc)]}$$

Equation 14. Conditional Function for Frequency Table Across Row Dimension

Observations of interest are grouped in tables 17 and 18. Table 17 corresponds to the values in the frequency table across the column dimension. Table 18 corresponds to the values in the frequency table across row dimension. The five columns in the tables are the same. Each row contains an observation which is continuous across tables to ease discussion of results. AR and AC are row attribute and column attribute. Frequency is the value in the cell of the corresponding table. The group column puts together observations with similar frequencies and indicates that the attribute combination of the AR and AC are either underrepresented or overrepresented relative to all other occurrences of the row attribute or the column attribute.

For instance, on Table 17, observation 3 shows providers and RCT at 5.13%. This means that out of all evaluations with main stakeholder of providers, only 5.13% used RCT. The remaining 94.88% of evaluations were split between using simulation (53.85%) and using field study (41.03%). On Table 18, observation 24 has RCT and Developing Country at a frequency of 3.33%. For this table, that result means that for all evaluations in developing countries, only 3.33% used RCT. The remaining 96.67% of evaluations were split between using field study (70%) or using simulation (26.67%). All observations are in groups except for observation 36 which had a frequency value that was not within the ranges but was included as a relationship of interest.

Group	Observation	AR	AC	Frequency
0-10%	1	Technology-based	6+ Months	0%
	2	RCT	Tech-Based	0%
	3	Providers	RCT	5.13%
	4	RCT	Developing Country	5.88%
	5	Simulation	6+ Months	6.67%
10-20%	6	Developing Country	Tech-Based	13.33%
	7	Patients	Tech-Based	13.85%
	8	Field Study	Tech-Based	14.29%
	9	Providers	6+ Months	15.38%
	10	Developed Country	6+ Months	16.22%
80-90%	11	Simulation	Developing Country	17.78%
	12	Simulation	Developed Country	82.22%
	13	Developed Country	0 - 6 Months	83.78%
	14	Patients	User-Based	84.62%
	15	Providers	0 - 6 Months	84.62%
	16	Field Study	User-Based	85.71%
	17	Developing Country	User-Based	86.67%
90-100%	18	Simulation	0 - 6 Months	93.33%
	19	RCT	Developed Country	94.12%
	20	RCT	User-Based	100.00%
	21	Tech-based	0 - 6 Months	100.00%

Table 17. Observations of Interest for Column Dimension Frequencies

Group	Observation	AR	AC	Frequency
0-10%	22	Technology-Based	6+ Months	0%
	23	RCT	Technology-Based	0.00%
	24	RCT	Developing Country	3.33%
10-20%	25	Providers	RCT	11.76%
	26	Simulation	6+ months	14.29%
	27	RCT	0 to 6 months	14.46%
	28	RCT	Positive Outcome	14.71%
	29	6+ months	Neutral outcome	14.71%
	30	Developing Country	Technology-Based	17.39%
	31	Technology-based	Positive Outcome	19.12%
80-90%	32	Developed Country	Technology-Based	82.61%
	33	0 to 6 months	Neutral outcome	85.29%
	34	Patients	RCT	88.24%
90% - 100%	35	User - Based	6+ Months	95.24%
	36	Simulation	Technology-Based	73.91%

Table 18. Observations of Interest for Row Dimension Frequencies

Some combinations are not represented on the tables, this was due to the characteristics of the taxonomy. The attributes that contained less than 5 occurrences across the taxonomy were not considered for determining relationships between attributes. These attributes are as follows: “Negative” for the dimension “Evaluation Conclusion” with 4 occurrences, “Cost-based” for the dimension “Main Metric” with 2 occurrences, and “Exists” for “Theory Application” with 4 occurrences.

The attributes of “RCT”, “6+ months”, and “technology-based” also have relatively lower occurrences compared to other dimensions (respectively, 17, 23, and 21), but were still considered for analysis.

The results showed partial support for the first two relationships and generated evidence that informed the other six. The first four method guidelines were already in place prior to the evaluation. Analysis of the results led to the creation of guidelines 5 and 6. Each relationship and guideline will be briefly discussed and the corresponding observations either supporting or refuting them will be referenced.

Relationship 1 posits that provider-focused evaluations often use simulation. This relationship is not totally supported and was likely just a unique characteristic of the group exclusive to cluster 1 in the cluster analysis. Observation 3 shows that provider-focused evaluations using RCTs are underrepresented, however the remaining evaluations that are provider-focused utilize field study and simulations without an overrepresentation of either group.

Relationship 2 states that simulation evaluations often have shorter durations. This tendency is demonstrated in observation 26, where out of all long-duration evaluations, those utilizing simulations are underrepresented at a frequency of 14.29%.

Relationship 3 posits that technology based evaluations often use simulations, this relationship is slightly supported by observation 36, where out of all technology-based evaluations, simulations make up 73.91%. Of note, observation 36 is not in any official grouping as the frequency value is higher than 20% and below 80%. However, we can still interpret that simulations make up the bulk of technology-based evaluations.

Relationship 4 states that technology-based evaluations have shorter durations. This relationship is supported by observations 1 and 22 and the fact that no evaluation in the taxonomy which utilized simulation took over 6+ months (Table 14).

Relationship 5 states that evaluations set in developing countries do not tend to have RCT evaluations. This relationship is supported by observations 4 and 19 which show that 94.12% of RCT evaluations occur in developed countries.

Relationship 6 states that evaluations in developing countries tend to have field study evaluations. This relationship is only partially supported, and can be observed on table 15 where out of all evaluations in developing countries, 70% are field studies.

Relationship 7 states that Patient-focused evaluations tend to have longer durations. This is supported by observations 9 and 15, which shows that most provider evaluations are of shorter duration, and observations 14 and 25 which show that most patient-focused evaluations are user based and that user-based evaluations make up most of the evaluations lasting over 6+ months.

Relationship 8 states that RCT evaluations tend to have patient stakeholders.

Observations 2, 20, and 23 support this relationship. A point worth noting is that the methodology of RCT inherently involves the testing of user groups, so evaluations utilizing RCT testing users is not a surprising finding.

Guidelines 1 through 4 were established from analyzing the taxonomy attributes along with the cross-attribute occurrences on table 14. Few evaluations are based on theory, report negative results, or focus on cost-based metrics. Most evaluations focus on user-based evaluations, tend to be of shorter duration, and are set in developed countries.

Guidelines 5 and 6 each mention the duration of evaluations. Observation 35 shows that user-based evaluations make up most of the long duration evaluations, which is consistent with other observations which show that technology-based evaluations along with those utilizing simulations tend to take shorter amounts of time to evaluate (observation 18 and observation 21). Observations 10, 17, and 32 also indicate that most technological evaluations take place in developed countries and that most evaluations in developing countries involve user-metrics. These characteristics lead to the guidelines that sufficient time should be allowed for evaluation and to prepare for a longer duration if users are directly involved.

The formative evaluation took some of the early relationships and method guidelines and tested them by examining the characteristics of taxonomy entities more closely.

From this analysis, more guidelines and relationships were added to make the 8 relationships and 6 guidelines that constitute the method artifact. The next stage in the research process was to perform a summative evaluation on the method.

Subsection 5.3.2 Method Survey

Summative evaluation was an expert survey to test the constructs, relationships, and guidelines of the method artifact. The survey questions can be seen at appendix item A12. There were 17 questions, the first 11 were answered on a 5-point Likert scale and the remaining 6 questions were a combination of free response and multiple choice.

The first part of the survey with 11 questions presented a statement related to m-health evaluation and asked for the respondent's level of agreement. The last 6 questions posed a hypothetical situation where the respondent oversaw an m-health project and its evaluation and asked questions related to the design of the evaluation.

Each of the survey questions will be briefly described along with an explanation of which part of the method artifact the question is referring to. Question 1 asks whether there is a difference between patient and provider metrics, this is to establish that the dimension of "main stakeholders" has merit and that relationships including the attribute have conceptual support.

Question 2 directly asks whether theory is important in evaluation. Guideline 1 suggests the application of theory to evaluations; The question is meant to determine how open potential evaluators would be to adding a theoretical component to evaluations.

Questions three, four, and five compares whether user, technical, or cost metrics are more important than the others when conducting m-health evaluations. The three questions are meant to test guideline 2 which suggests not to over-emphasize user and technological metrics at the expense of cost metrics. Results from the question can also

provide insight as to which group of metrics are generally considered to be the most important in the domain.

Question six asks whether studying user metrics takes longer compared to cost and technological metrics. This question is meant to test guidelines 5 and 6 which deal with allocating enough time to perform effective evaluations, particularly when evaluations involve testing user metrics or involve users in a non-simulation environment.

Questions 7 through 11 are on the impact of evaluation setting on evaluation effects. Survey takers were presented with a map of worldwide Human Development Index (HDI) of countries to provide context for the question (Appendix A13). Question 7 asks whether higher HDI countries have more complex network infrastructure. This question is meant to test parts of guideline 3, which states that provider-centric evaluations in developed countries, technology-based evaluations using simulation, and short-duration evaluations using simulation tend to be overrepresented in the literature. The underlying relationships for this guideline were uncovered during formative evaluation (Observations 5, 7, 8, and 10, Table 17; Observations 30 and 32, Table 18).

Question 7 tests whether a more complex network infrastructure exists in developed countries. If developed countries do have more complex networks, evaluations are more likely to be based on technical metrics or simulation to study the technological effects. Question 8 asks whether m-health implementations are costlier to implement in developed countries. This question is meant to determine whether there is a propensity for cost-based evaluations in developed countries or whether simulations are more popular in developed countries for financial reasons. Questions 9, 10, and 11 ask about the connection between setting and evaluation type. Each question asks whether it is

easier to conduct RCT, field studies, or simulations in developed countries. Question 9 checks relationship 5, that evaluations in developing countries do not tend to have RCT evaluations. Question 10 checks relationship 6, that evaluations in developing countries tend to have field study evaluations. Question 11 checks relationships 1, 3, and 4, which all relate to simulation evaluations and their tendencies.

Questions 12 through 17 were prefaced by a hypothetical situation presented to the survey taker. Respondents were told that they were leading a research study to develop and evaluate an m-health implementation in a country with either high or low HDI. They were also told that they had unlimited funding to pursue whatever research questions they want. Question 12 is a free response question that asks what methodology they would choose to conduct the research. Question 13 asks whether they would use RCT, field study, or simulation to evaluate their research. Question 14 references the research setting directly and asks whether they would focus on user-based metrics, technology-based metrics, or cost-based metrics. Question 15 asks respondents to provide more detail and give some specific metrics they would study. Questions 16 and 17 ask how long users would expect the research process to take and how long they would expect the evaluation process to take.

The second half of the survey meant to determine whether having an evaluation set in a developing or developed country will change the way an evaluation is designed. The questions are also meant to determine what an ideal m-health evaluation context would look like, as respondents were told that they would have unlimited funding to work on any research questions they wanted.

The questions and flow of the survey was designed with feedback from the taxonomy survey in mind. Many respondents to the earlier survey mentioned that presenting the taxonomies without much context made it difficult for them to effectively make a comparison. They also stated that it was hard to come up with specific ideas when every question was a freeform response. In response to these comments, over half of the method survey questions were multiple-selection. Additionally, descriptions and examples were provided for each new concept that was introduced.

The method guidelines and relationships were not directly presented to the user. If the respondents knew that a specific relationship or guideline was what was being evaluated, like they did for the taxonomy survey, there was a tendency to state or agree that they were satisfied with what they saw without any further information. While these responses were appreciated, they did not provide additional insights

The survey was distributed to 25 individuals working in either the m-health or information systems field. There were 19 respondents, the means and standard deviations for the first 11 questions are summarized on table 19.

	Average (Mean)	Standard Deviation	Note
Question 1	4.00	0.75	Patient and Provider metrics are different
Question 2	3.53	1.26	Theory is important
Question 3	3.79	1.13	User metrics > Tech metrics
Question 4	3.84	1.01	User metrics > Cost metrics
Question 5	2.47	1.22	Cost metrics > Tech metrics
Question 6	3.63	1.38	User evaluations take longer
Question 7	4.16	0.83	Higher HDI means more developed infrastructure
Question 8	2.79	1.40	Higher HDI means more costlier implementations
Question 9	3.16	1.07	Higher HDI means easier to conduct RCT
Question 10	2.53	1.31	Higher HDI means easier to conduct Field Studies
Question 11	3.68	1.16	Higher HDI means easier to conduct Simulations

Table 19. Means and Standard Deviations for Q1 – Q11 (n = 19)

To aid in interpretation, Likert scale values are as follows: 1 – Strongly disagree, 2 – Disagree, 3 – Neither agree nor disagree, 4 – Agree, 5 – Strongly agree. The mean average is displayed on the second column and the standard deviation is on the third.

Question one averaged a score of 4, so respondents generally agreed that patient and provider metrics are different. Responses 2 had a score of 3.53, indicating that most respondents felt that inclusion of theory was important but not vital to an evaluation.

This result is interesting considering the low number of existing theory-based evaluations in the field. Questions 3 and 4 average 3.79 and 3.84, indicating that most respondents felt that user metrics were more important to measure compared to technical metrics and cost metrics. This is consistent with what was seen in the taxonomy and subsequent evaluations.

Question 5 had a result of 2.47, showing that most respondents did not think that cost metrics were superior to technical metrics. While not every possible combination was tested, a conclusion that can be drawn from this result is that user metrics are considered the most important, followed by technical metrics and finally cost metrics. Question 6 averaged 3.63, showing a slight agreement that user evaluations tend to take longer.

Questions 7 through 11 dealt with evaluation setting. Question 7 averaged 4.16, showing general agreement that developed countries tend to have more complex infrastructure. Question 8 averaged 2.79, showing that respondents generally disagreed with the contention that implementations are costlier in developed countries. This result is interesting in that it counters the argument that simulations are more popular in developed countries because of cost concerns. The general perception that there is no

major difference in costliness of implementation and evaluation between developed and developing countries means that there is probably another reason for the prevalence of simulations in developed countries.

Question 9 averaged 3.16 and asked whether it was easier to conduct RCT evaluations in developed countries. This is not a strong level of agreement and indicates that respondents did not feel country setting and prevalence of RCT evaluation are connected. However, the lack of RCT evaluations in developing countries is clear from the formative evaluation. The disconnect between the two evaluations is likely due to a combination of contextual factors in individual evaluations and the low number of RCTs evaluations in the taxonomy.

Question 10 saw even less support, averaging 2.53 when asking whether field studies were easier to conduct in a developed country. This shows slight disagreement with the statement, giving partial support to the assertion that field studies are more common in developing countries. The assertion that field studies are more common in developing countries because simulations and RCTs are more difficult to perform still holds some merit. Question 11 averaged 3.68 and asked whether it was easier to conduct simulations in developed countries. This shows a higher level of agreement compared to the previous two questions and is consistent with observations made during the formative evaluation.

The second half of the survey had four free response questions and two questions with 3 choices each. Respondents were randomly told that they were designing a study and evaluation for an m-health implementation in a developing or developed country. The summarized results for questions 12 through 17 are on table 20. Analysis of the results

show that there was not a noticeable distinction between the responses based on the provided setting.

Dimension	Attribute	High HDI	Low HDI
Type of evaluation	Field Study	6	8
	RCT	2	2
	Simulation	1*	0
Main Metric	User Metrics	8**	9
	Tech Metrics	0	1
	Cost Metrics	0	0
Project Length	0 - 6 months	0	1
	6+ months	9	9
Evaluation Length	0 - 6 months	4	6***
	6 + months	5	3

*Respondent made the caveat that they would select simulation is unlimited funding was not available

**One respondent elected not to answer, saying that metric of choice would depend on research question

***One respondent elected not to answer, said length of evaluation depends on research question

Table 20. Summarized results for Q12-Q17 (n = 19)

Most respondents chose field study as the evaluation technique of choice with four selecting RCT and one selecting simulation. The lack of choosing simulation does not match with what was uncovered during earlier evaluations which showed many evaluations utilizing simulation. This result may be because of the hypothetical context of the survey, which allowed respondents to determine their own temporal constraints and explicitly stated that funding was unlimited. All but one respondent had timelines of

at least one year for study implementation. However, only 8 out of 19 had evaluation periods over six months. 2 out of the 4 respondents that selected RCT as the evaluation type had evaluation periods of less than 3 months. One of the other two allocated 1 year for evaluation and one stated that the evaluation duration would depend on the research question.

Question 15 asked what specific metrics users would want to use. User metrics were the most popular, mortality rate, user engagement, turnover rate, and satisfaction were mentioned multiple times. Tech-centric metrics were mentioned the second-most frequently and included ease of use, attitudes towards technology, and Technology Acceptance Model (TAM) constructs. 2 respondents mentioned cost metrics in addition to user metrics, no specific cost metrics such as return on investment or reduction in healthcare expenditures were mentioned.

Summative evaluation revealed that most of the relationships and guidelines comprising the method were supported or generally acknowledged as having some truth. A few results suggested that some relationships may have a different underlying reason or that the relationship itself does not make apparent sense. The relationship between developed and developing countries and the impact on evaluation characteristics is not as clear cut as the formative evaluation indicated. Many of the discrepancies and groupings observed in previous analyses may be attributed to contextual factors independent to projects that are not immediately discernable from a reading of the project report. Supported relationships were the importance of user metrics relative to cost and technology metrics and an agreement that theoretical components are important in evaluation. An interesting finding from the summative evaluation is the

popularity of field studies and the near-complete neglect of simulation as an evaluation method when an ideal environment is presented to evaluation designers. The prevalence of simulation usage in the domain is an indicator that evaluations are likely taking place in environments where a field study or RCT may be preferable; But due to temporal or financial restraints, a simulation is used to evaluate instead.

Section 5.4 Theorizing the Artifact

The final stage in the research was to create an analytical framework, or theory, from the findings gathered during design and evaluation of the artifacts. The theory design process comes with a few caveats, the conceptual basis informing it are the two artifacts and there is no justification cycle to determine construct, internal, or external validity. The theory is meant to be analytical and is a representation of what was observed during the research process. Testing and potential extensions of the theory are covered in the discussion section.

The conceptual foundation for creating a theory from artifacts has been explored in the design science literature. Kuechler and Vaishnavi in their 2008 paper present a series of guidelines for the development of design science and information systems theories. They describe ideas from other disciplines as kernel theories, which can then be adapted into the domain of information systems through the incorporation of those ideas into design science research processes (Kuechler and Vaishnavi 2008). Recent research in the domain of design science has also explored the possibilities of theory development from taxonomies in the IS domain (Nickerson et al. 2017).

Gregor defined five types of theory in her seminal 2006 paper on the nature of theory in information systems: type I theory which analyzes phenomena, type II theory which explains phenomena and includes insights, type III theory which predicts and presents testable propositions, type IV theory which predicts and explains testable propositions, and Type V theory which is comprised of design and action, or the instantiation of the theory in an open, naturalistic environment (Gregor 2007).

Each type of theory tests and builds upon the previous type. The transfer of kernel theories from another literature to the design science and information systems contexts can lead to the development of Type I, Type II, Type III, or Type IV theories; Insights gathered from design science research cycles can also be directed into analytical and exploratory (Type I and Type II) theories that establish and describe relationships between constructs (Friedman 2003; Baskerville and Pries-Heje 2010; Kuechler and Vaishnavi 2012).

The development of the taxonomy and method draws upon the domains of project management and healthcare along with IS. The overarching kernel domain is project evaluation in a healthcare setting, the information systems component is the specific focus on m-health technologies. The evaluation of projects and the concepts of stakeholder management and metric design serve as kernel theories which inform the creation of the artifacts and their theorization. The results from these stages in the research process leads to the development of a type I analytical theory for m-health evaluation. This cycle follows the suggestion made by Venable for theory development in a design science framework, figure 9 shows the model adapted to this research (Venable 2006).

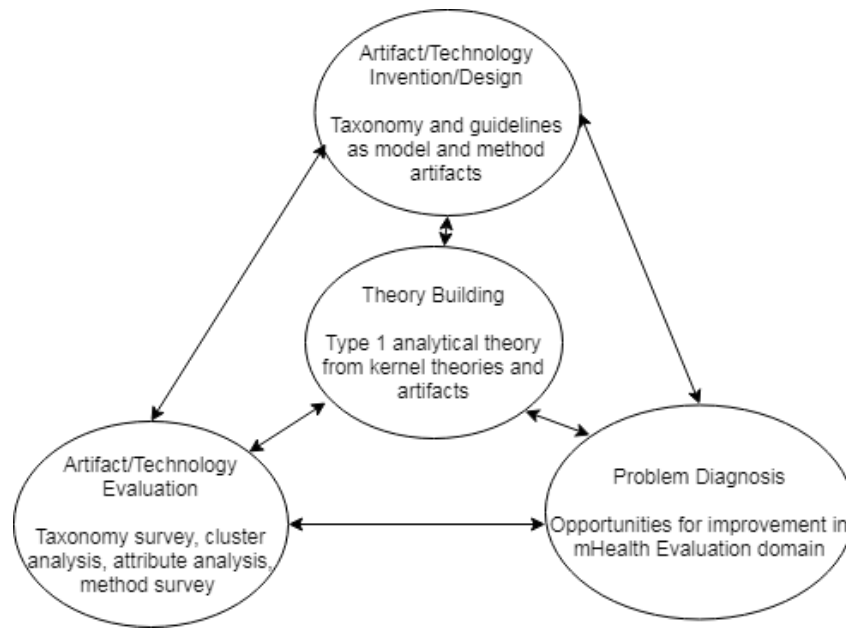


Figure 9. Activity Framework for Design Science Research (Adapted from Venable 2006)

The theory is meant to serve as an extension of the artifacts. The taxonomy artifact sought to answer the question of “what are the characteristics of m-health evaluations?” The method artifact expanded on this question and sought to answer the questions of “what are the relationships between m-health evaluation characteristics?” and “why do these relationships occur?” The theory is meant to serve a stepping stone towards creating an answer for these questions.

The theory is built upon ideas from the taxonomy and method artifacts. From the taxonomy, dimensions were reinterpreted as constructs. From the method, the relationships between the dimensions and attributes were expanded into interactions between the constructs. A diagram of the constructs and relationships is shown on figure 10.

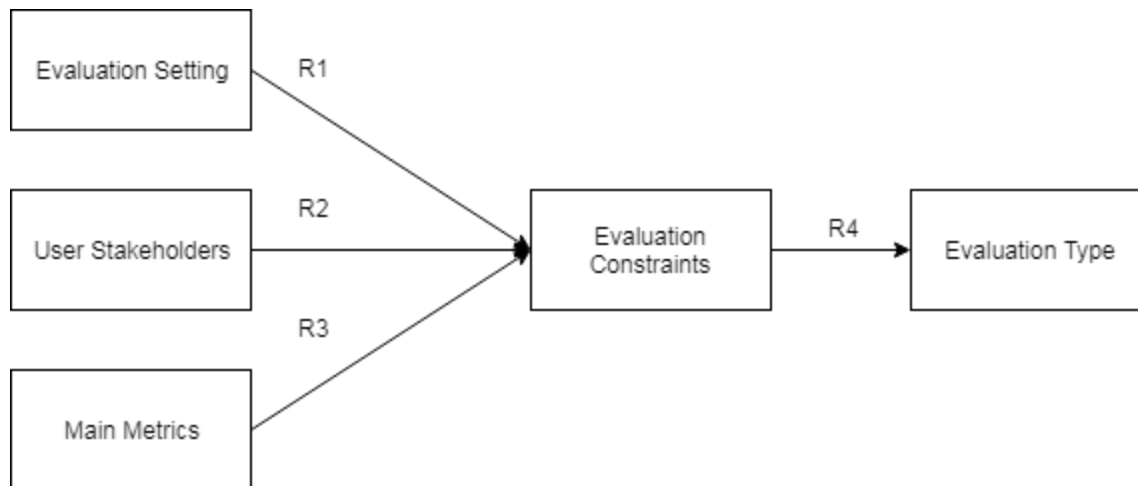


Figure 10. Analytical Model for m-health Evaluation Selection

The unit of observation and unit of analysis of this theory are both at the project level. The model is meant to examine the phenomenon of disparity between evaluation goals and the chosen methodology for conducting evaluations in an m-health context. The flow reads from left to right, evaluation setting, user stakeholders, and the main metrics serve to create a combination of evaluation constraints. These constraints are partially created by the three constructs, but also include contextual stipulations unique to each evaluation. These constraints then affect the choice of which evaluation type is used for the project. The links between constructs are labeled R1, R2, R3, and R4, short for relationships 1, 2, 3, and 4. Each of the constructs and relationships is described in further detail below.

Evaluation type is the dependent construct and represents the chosen methodology for evaluation of an m-health project. The construct is informed by the taxonomy dimension of the same name with the attributes of “Field study”, “RCT”, and “Simulation”.

Evaluation constraints are influenced directly by the three preceding constructs and represent the temporal, financial, and logistical obstacles that prevent an evaluation from achieving an ideal state.

Evaluation setting maps to the dimension of the same name in the taxonomy. The construct is meant to represent the environment which the evaluation is taking place in. R1 is the impact the evaluation setting has on creating or assuaging evaluation constraints. This can take the form of existing infrastructure, data availability, or other factors that serve as benefits or hindrances to an evaluation.

User Stakeholders represents the intended users of the m-health implementation being evaluated. This construct maps to the taxonomy dimension of “Main Stakeholders” with the attributes of “Patients” and “Providers”. R2 represents how user stakeholders can generate constraints in m-health evaluation. Different types of users can have different needs, necessitating different standards and metrics for evaluation. Certain users may require more time to effectively evaluate, leading to additional constraints.

Main metrics maps to the dimension of “Main metric” in the taxonomy, with the attributes “technology-based”, “user-based”, and “cost-based”. This construct represents the main indicators of success for the m-health project that are ideally meant to be the most clearly and objectively assessed metrics during the evaluation. R3 represents how selection of these metrics can create a constraint upon the evaluation. Measuring one set of metrics can take more time compared to others. Certain metrics also lend themselves more readily to an evaluation type and trying to implement another would take additional effort. R4 is the combination of these constraints and their effect on the selection of an evaluation type. This relationship was explored in the design of the

method artifact. Examples include the need to measure user data leading to more frequent field studies or RCT, simulations being prevalent in projects where constraints on evaluation duration were high, and field studies becoming the methodology of choice when user metrics were the most important.

While the constructs from the model were derived from the taxonomy and method artifacts, theoretical work related to the constructs developed in this study are not new. All the constructs have grounding in the IT, HIT, and evaluation literature. For the construct of evaluation setting, there is general agreement that developing countries contain different contexts compared to developed countries. These contexts can influence IT implementation and challenge traditional theories of IT adoption (Heeks 2002; Kamsu-Foguem et al. 2014). In the HIT context, adoption attempts of health interventions or technological innovations in developing countries requires different perspectives of user requirements and expectations (Kijisanayotin et al. 2009; Park et al. 2009).

In the HIT literature, the impact of the user stakeholder and the metric of success on evaluation constraints has been widely documented (Eccles 2005). On the provider side, ideas such as planned behavior, context, and information flow have been used to describe the levels of willingness provider-side users have had in adopting new technologies for improving efficiency (Godin 2008; Korttesisto et al. 2010; Thomas et al. 2014). On the patient side, many theories have been conceived suggesting that interventions and evaluations should be tailored to the context of patient needs (Koplan et al. 1999; Ciechanowski et al. 2002). The interpretation of patient metrics as well as individual intentions of patients on the adoption of new technologies are also important

factors that influence the design of HIT evaluations (Devaraj et al. 2013; Rai et al. 2013; Constand et al. 2014). Differences in user stakeholders or metrics can generate widely different implementation and evaluations for m-health interventions that may look similar on a technological level, but are evaluated completely differently (Zapata et al. 2015).

Constraints in evaluation have been documented ever since the foundation of the project management discipline. In the HIT context, these limitations have been documented to influence the process and timeline of evaluation designs (Linnan and Steckler 2002, Sridharan et al. 2006). The relationship of these constraints to the chosen evaluation technique is an exercise in balancing the granularity of metrics to the complexity of collected information (Sridharan et al. 2006).

Recent work in the evaluation literature has identified the value of theory-driven evaluations (Walshe 2007; Astbury et al. 2010, Coryn et al. 2011). The incorporation of theory and the practice of having theory serve as one of the conceptual foundations when creating an evaluation is becoming an increasingly accepted and encouraged practice (Donaldson 2012, Fishman et al. 2013). The type I theory presented in this research was created from observations and relationships generated from the analysis of m-health evaluations in a design science research process. The analytical model has the potential to provide utility to any individuals creating evaluations in the domain of m-health and can serve as a building block for future theories on m-Health evaluation.

Chapter 6. Discussion

Discussion will be done in four parts. The first section will cover limitations of this research from methodological and epistemological standpoints. The section also discusses the general scope of the research and what actions were taken to try and mitigate the effects of limitations. The next section is contributions, which will go over where this work fits in advancing knowledge and the role of the artifacts as intellectual products of the research process. The third section will cover future directions of this research and the final section contains the conclusion. All sections will comment on ideas and decisions spanning the timeline of the work, from the early literature review and taxonomy design to the concluding method summation and theory development.

Section 6.1 Limitations

The creation of the taxonomy and the method was done following the framework proposed in Nickerson et al. 2013. However, inside the twin cycles of empirical-to-conceptual and conceptual-to-empirical, the creation, reformation, and deletion of dimensions and attributes was left to the discretion of the taxonomy designer. The creation of constructs and relationships of the method was also done based on an interpretation of results from the taxonomy design and evaluation process. As no single individual can have perfect information concerning all aspects of a domain, this required subjective interpretation of information. Major portions of the design of both artifacts lean towards the interpretivist philosophy of research and bear its potential shortcomings. To lessen the impact from this design decision, more objective criteria

such as empirical analysis and feedback from experts was used in the evaluation of the artifacts.

The bulk of data gathering for the taxonomy itself occurred during early 2017 to mid-2017. As a result, most of the evaluations within the taxonomy are dated 2016 and earlier. The field of m-health is still growing, and more evaluations have occurred since mid-2017. A search conducted during early 2018 revealed that potential candidates for inclusion into the taxonomy numbered at least 50 and up to 100 documented evaluation cases could have been included. This issue was addressed throughout the research by having multiple stages of data collection. Once the taxonomy reached 104 entities, no further documents were added to maintain the integrity of the evaluations and the constructs and relationships formed in the method.

Formative evaluation of the taxonomy was conducted as an expert survey. As the lead researcher knew each potential respondent of the survey, there was a risk that respondents may have altered their answers knowing that they were opining on the work of someone they know. While some responses may have displayed this tendency, the overall results were not unanimously positive and constructive feedback was provided regarding both the evaluated artifact and the design and administration of the survey itself. These shortcomings were kept in mind and informed the method survey conducted later in the research.

The summative evaluation of the taxonomy was done through cluster analysis. The clustering vector was the frequency of word counts within the documents describing the evaluation. Direct word count is one of the simpler natural language processing techniques. More complex methods such as sentiment analysis, concept mining, and

terminology extraction exist. As the goal of the evaluation was to determine the aptness of the taxonomy dimensions and attributes, word frequencies were considered a sufficient clustering variable as the words themselves could be linked to the attributes and dimensions of the taxonomy. The results of the clustering also showed a high amount of overlap between different clusters. This is a limitation of the data that was used to generate the analysis. The data had the characteristics of high dimensionality with a relatively low sample size. To address these problems, primary components analysis was used to better visualize data and manual selection of keywords relevant to taxonomy attributes was also conducted to reduce total dimensionality.

The development of the constructs, relationships, and guidelines of the method artifact were based on the results from the taxonomy design and evaluation. Any shortcomings in the design and evaluation of the taxonomy could have influenced the method. To mitigate this problem, summative evaluation of the method was kept separate from the taxonomy through survey design.

Formative evaluation of the method involved attribute analysis of taxonomy values. Characteristics of the taxonomy data made interpretation of certain values difficult. To mitigate this, attributes that saw very low occurrences within the taxonomy such as cost-based metrics and theory-based evaluations were not considered in the development of constructs and relationships. The lack of occurrences in select attributes did wind up serving as general observations.

Summative evaluation of the method was another expert survey. Limitations revealed from the previous survey were addressed by changes to the survey design. One of these changes was not showing and asking respondents about method guidelines

directly. Instead, questions related to the constructs and the relationships underlying those constructs were presented. The survey also presented respondents with a hypothetical ideal m-health evaluation environment. The situation set in the survey (unlimited funding, implied unlimited time) was meant to gauge what respondents felt was an ideal evaluation, but is not a reflection of actual m-health implementations.

Based on Orlikowski's description, the artifacts developed from this study were presented as a model (Taxonomy) and a method (Guidelines), no instantiation artifact was created which involves the application and evaluation of the artifacts in a naturalistic environment. For these same reasons, an ex-post evaluation was not possible because of the extent of the research process. Creating an instantiation artifact based on this work is possible, but was beyond the scope of research project.

Theory development was inspired by the method artifact and informed by the concepts of transferring kernel theories and generating theory from design. The theorization of the artifact generated a model for m-health selection, but it has not been tested or justified in an m-health setting as that was also beyond the scope of this project. As a result, the validity of the model itself is not proven. However, the conceptual foundations of it are the sum of all research processes that have been described thus far.

Section 6.2 Contributions

Figure 11 contains a summary of the research process. Each of the major contributions in the project are represented by a circular node, apart from the start and end nodes. The numbers indicate the order research activities occurred. Dotted lines represent that one event influenced another.

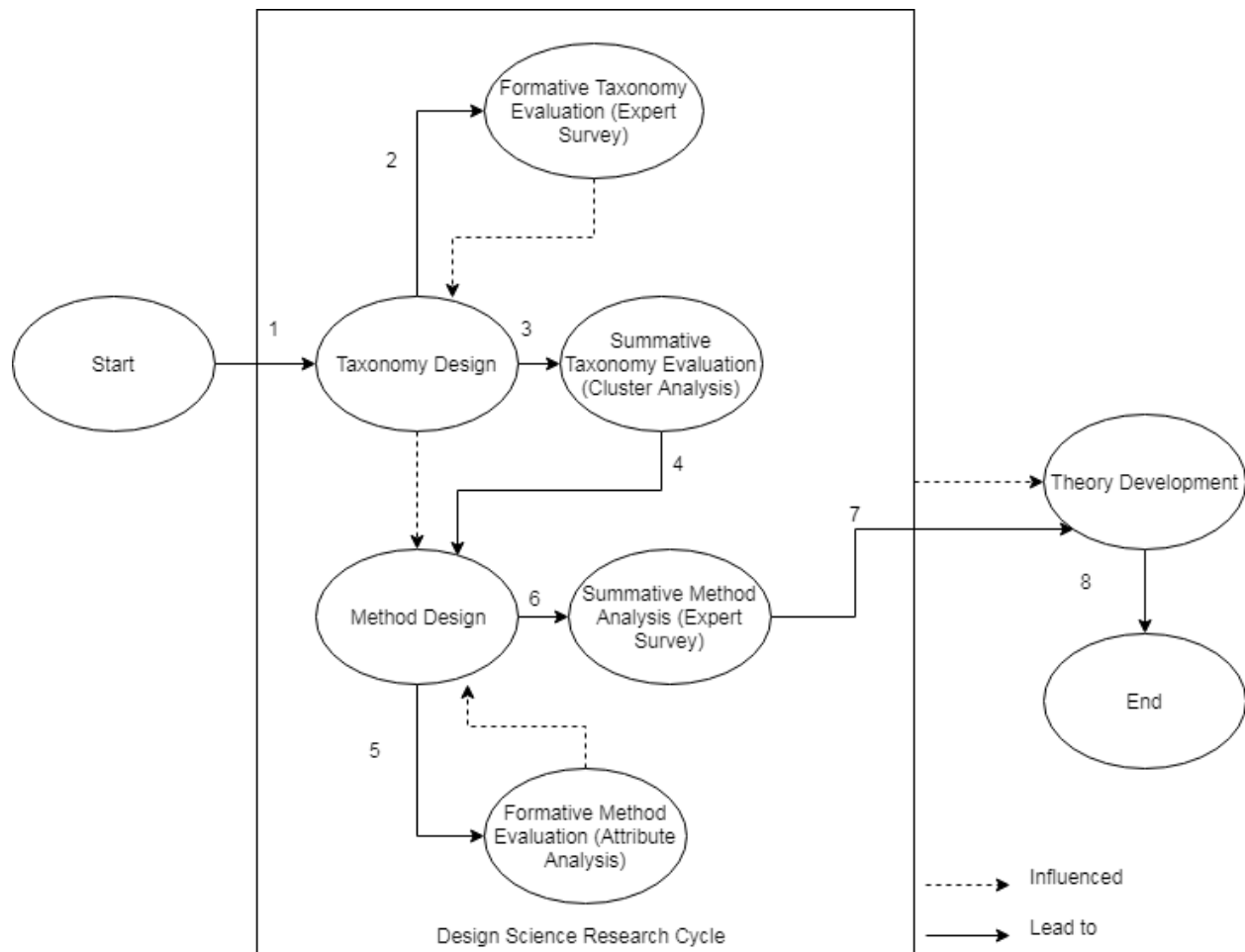


Figure 11. Summary of Research Process

This research is positioned between the intersection of the domains of m-health and evaluation. M-health is itself a subset of the fields of HIT and IT; Evaluation is a subset of the field of project management. The methodology of the research is design science and the two main contributions are created, evaluated, and presented as design science artifacts. The work adds to the methodological literature of information systems by applying the technique of taxonomy development to the sub-domain of m-health. The work also furthers the design science literature by presenting and evaluating the taxonomy as an artifact. A caveat to this contribution is that this work is not the first

instance of a classification system being presented as an artifact. However, the creation of the taxonomy adheres to a new technique from the design science literature and pioneers the application of that technique to the m-health evaluation domain.

The taxonomy serves as the foundation for the series of constructs and relationships underlying the second artifact. The design cycle saw the creation of guidelines for m-health evaluation. The underlying relationships uncovered during the evaluation stages led to the creation of an analytical theory for the selection of evaluation techniques in m-health projects.

Practical contributions of this research are the categorization of m-health evaluation attributes and the observations gleaned throughout the research process. Insights gained during clustering analysis revealed that the m-health evaluation field is very homogenous. Analysis of the individual characteristics of m-health projects revealed that theory-based and cost-based evaluations were underrepresented and that very few evaluations reported negative effects despite the high failure rate of m-health projects (Kaplan et al. 2009). These findings, along with the analytical theory from the method, can help guide future research and evaluation in the field of m-health. Optimistically speaking, the filling of conceptual gaps and the diversification of evaluation techniques are all steps towards a future where replication of this research reveals a heterogenous environment containing unique, innovative, and effective m-health evaluations.

Section 6.3 Future Work

Expansion of this work combines the activities of addressing the limitations and building on the contributions of the research. Data collection for the taxonomy stopped in late-

2017. Evaluations that have occurred between 2017 until now (mid 2018) have shown an increased emphasis on the security and privacy of user data (Idrish et al. 2018; Hussain et al. 2018; Vasudevan et al. 2018). Another recent development is the formal announcement of a new wi-fi protected access (WPA) protocol, WPA3, which is set to replace the existing standard of WPA2 in the coming years (Figuerola 2018). These changes in the field have already influenced the field of evaluation. A replication of the taxonomy design and evaluation can update the work to better reflect the current state of m-health evaluation and can determine if the dimensions and attributes created during the seven iterations of this research cycle are still applicable to more recent developments.

The taxonomic design process used in this research can be expanded. The determination of the meta-characteristic of the taxonomy and the ending conditions could be further explained. The subjective criteria serving as ending conditions in the taxonomy are difficult to objectively verify and findings from the formative and summative evaluation of the taxonomy show that other methods of testing a taxonomy exist. The cluster analysis can be expanded with new data and techniques. Sentiment analysis could be conducted on evaluation documents to generate data and further empirical analysis could be used to determine groupings between taxonomic entities.

Testing of the method could be expanded, the summative evaluation presented an ideal environment to survey takers and results from the formative evaluation were influenced by the low occurrences of certain attributes. The method itself could be applied to a naturalistic environment and evolve to become an instantiation artifact. This would most likely take the form of an action research project where m-health evaluators utilize the

findings from the taxonomy and guidelines from the method to inform the design and implementation of their evaluation. This would serve to further test the relevance of both the taxonomic attributes and the relationships between them.

A theory was developed from the method artifact, the next stage would be justification of the theory. This would involve additional construct definition and the creation of metrics for representing construct components. For instance, the construct of evaluation constraints may be expanded to include a characteristic such as data availability; the construct of evaluation setting may be expanded to include network complexity. The strength of the relationships between constructs could then be tested through additional surveys or included as part of a field study or action research project.

Expansion of the theory could occur after justification of the initial constructs. Testing of the analytical model may reveal additional constructs and relationships. A potential construct based on the method evaluation could be theory inclusion. The resultant relationship would explore how having a theoretical basis for evaluation influences the choice of evaluation type. Continued development of the theory could lead to explanatory or predictive models built upon the analytical model developed from this work.

Section 6.4 Conclusion

The field of information systems has ushered in continuous improvement and reshaped the way we look at interactions between humans and technology. benefits from the application of mobile technology to the field of healthcare are noteworthy not only because of their technological capabilities, which were impossible just a few decades

ago, but also because of their potential for the direct betterment of human lives. These opportunities have served as an impetus driving m-health projects and fueling the rise of the discipline. Whenever such an explosion in the popularity of a domain occurs, it is paramount to take a more sobering look at the state of the field. This research has held up a lens to the field of m-health implementation by examining what goes into the process of evaluating these m-health projects. As the field continues to progress, reflective work that pauses and asks the questions of “what, how, and why?” will become increasingly important as we collectively seek answers for why things work in the present and what comes next in the future.

References

- Abdi, H., and Williams, L. J. 2010. "Principal Component Analysis," Wiley interdisciplinary reviews: computational statistics (2:4), pp. 433-459.
- Afshin, A., Babalola, D., Mclean, M., Yu, Z., Ma, W., Chen, C. Y., Arabi, M., and Mozaffarian, D. 2016. "Information Technology and Lifestyle: A Systematic Evaluation of Internet and Mobile Interventions for Improving Diet, Physical Activity, Obesity, Tobacco, and Alcohol Use," Journal of the American Heart Association (5:9), p. e003058.
- Agarwal, R., Gao, G., DesRoches, C., and Jha, A. K. 2010. "Research Commentary—the Digital Transformation of Healthcare: Current Status and the Road Ahead," Information Systems Research (21:4), pp. 796-809.
- Alexander, E. R., and Faludi, A. 1989. "Planning and Plan Implementation: Notes on Evaluation Criteria," Environment and Planning B: Planning and Design (16:2), pp. 127-140.
- Alrige, M., and Chatterjee, S. 2015. "Toward a Taxonomy of Wearable Technologies in Healthcare," International Conference on Design Science Research in Information Systems: Springer, pp. 496-504.
- Alsabti, K., Ranka, S., and Singh, V. 1997. "An Efficient K-Means Clustering Algorithm,").
- Amadi-Obi, A., Gilligan, P., Owens, N., and O'Donnell, C. 2014. "Telemedicine in Pre-Hospital Care: A Review of Telemedicine Applications in the Pre-Hospital Environment," International journal of emergency medicine (7:1), pp. 1-11.
- Ammenwerth, E., Gräber, S., Herrmann, G., Bürkle, T., and König, J. 2003. "Evaluation of Health Information Systems—Problems and Challenges," International Journal of Medical Informatics (71:2), Pp 125-135.
- Anderson, J. G., and Aydin, C. E. 2005. "Overview: Theoretical Perspectives and Methodologies for the Evaluation of Healthcare Information Systems," in Evaluating the Organizational Impact of Healthcare Information Systems. Springer, pp. 5-29.
- Anderson, K., Burford, O., and Emmerton, L. 2016. "Mobile Health Apps to Facilitate Self-Care: A Qualitative Study of User Experiences," PLoS One (11:5), p. e0156164.
- Anderson, R. 2010. "Systematic Reviews of Economic Evaluations: Utility or Futility?," Health economics (19:3), pp. 350-364.
- Aranda-Jan, C. B., Mohutsiwa-Dibe, N., and Loukanova, S. 2014. "Systematic Review on What Works, What Does Not Work and Why of Implementation of Mobile Health (M-health) Projects in Africa," BMC public health (14:1), p. 1.
- Årsand, E., Frøisland, D. H., Skrøvseth, S. O., Chomutare, T., Tatara, N., Hartvigsen, G., and Tufano, J. T. 2012. "Mobile Health Applications to Assist Patients with Diabetes: Lessons Learned and Design Implications," Journal of diabetes science and technology (6:5), pp. 1197-1206.
- Asangansi, I., and Braa, K. 2010. "The Emergence of Mobile-Supported National Health Information Systems in Developing Countries," Stud Health Technology Inform (160:Pt 1), Pp 540-544.

- Assent, I. 2012. "Clustering High Dimensional Data," Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery (2:4), pp. 340-350.
- Astbury, B., and Leeuw, F. L. 2010. "Unpacking Black Boxes: Mechanisms and Theory Building in Evaluation," American journal of evaluation (31:3), pp. 363-381.
- Baccarini, D. 1999. "The Logical Framework Method for Defining Project Success," Project management journal (30), pp. 25-32.
- Bailey, K. D. 1994. "Typologies and Taxonomies: An Introduction to Classification Techniques", Vol. 102 (Sage, 1994)
- Bardhan, I. R., and Thouin, M. F. 2013. "Health Information Technology and Its Impact on the Quality and Cost of Healthcare Delivery," Decision Support Systems (55:2), pp. 438-449.
- Barton, A. J. 2012. "The Regulation of Mobile Health Applications," BMC Medicine (10:1), P 46.
- Baskerville, R. L., Kaul, M., and Storey, V. C. 2015. "Genres of Inquiry in Design-Science Research: Justification and Evaluation of Knowledge Production," Mis Quarterly (39:3), pp. 541-564.
- Baskerville, R., and Pries-Heje, J. 2010. "Explanatory Design Theory," Business & Information Systems Engineering (2:5), pp. 271-282.
- Bastawrous, A., and Armstrong, M. J. 2013. "Mobile Health Use in Low-and High-Income Countries: An Overview of the Peer-Reviewed Literature," Journal of the royal society of medicine (106:4), pp. 130-142.
- Beck, K., and Gamma, E. 2000. Extreme Programming Explained: Embrace Change. addison-wesley professional.
- Beck, K., Beedle, M., Van Bennekum, A., Cockburn, A., Cunningham, W., Fowler, M., Grenning, J., Highsmith, J., Hunt, A., and Jeffries, R. 2001. "Manifesto for Agile Software Development,").
- Ben-Zeev, D., Schueller, S. M., Begale, M., Duffecy, J., Kane, J. M., and Mohr, D. C. 2015. "Strategies for M-health Research: Lessons from 3 Mobile Intervention Studies," Administration and Policy in Mental Health and Mental Health Services Research (42:2), pp. 157-167.
- Berkhin, P. 2006. "A Survey of Clustering Data Mining Techniques," in Grouping Multidimensional Data. Springer, pp. 25-71.
- Berry, M. W., and Castellanos, M. 2004. "Survey of Text Mining," Computing Reviews (45:9), p. 548.
- Bert, F., Giacometti, M., Gualano, M. R., and Siliquini, R. 2014. "Smartphones and Health Promotion: A Review of the Evidence," Journal of medical systems (38:1), pp. 1-11.
- Blackwood, B., O'halloran, P., and Porter, S. 2010. "On the Problems of Mixing Rcts with Qualitative Research: The Case of the Mrc Framework for the Evaluation of Complex Healthcare Interventions," Journal of Research in Nursing (15:6), pp. 511-521.
- Blashfield, R. K. 1976. "Mixture Model Tests of Cluster Analysis: Accuracy of Four Agglomerative Hierarchical Methods," Psychological Bulletin (83:3), p. 377.
- Blaya, J. A., Fraser, H. S., and Holt, B. 2010. "E-Health Technologies Show Promise in Developing Countries," Health Affairs (29:2), Pp. 244-251.

- Blomquist, T., Hällgren, M., Nilsson, A., and Söderholm, A. 2010. "Project-as-Practice: In Search of Project Management Research That Matters," *Project Management Journal* (41:1), pp. 5-16.
- Bowling, A. 2014. *Research Methods in Health: Investigating Health and Health Services*. McGraw-Hill Education (UK).
- Braun, R., Catalani, C., Wimbush, J., and Israelski, D. 2013. "Community Health Workers and Mobile Technology: A Systematic Review of the Literature," *PloS one* (8:6), p. e65772.
- Brennan, A., Chick, S. E., and Davies, R. 2006. "A Taxonomy of Model Structures for Economic Evaluation of Health Technologies," *Health economics* (15:12), pp. 1295-1310.
- Bryde, D. J. 2003. "Modelling Project Management Performance," *International Journal of Quality & Reliability Management* (20:2), pp. 229-254.
- Brynolfsson, E. 1993. "The Productivity Paradox of Information Technology," *Communications of the ACM* (36:12), pp. 66-77.
- Buchanan, R. 1992. "Wicked Problems in Design Thinking," *Design issues* (8:2), pp. 5-21.
- Charrad, M., Ghazzali, N., Boiteau, V., Niknafs, A., and Charrad, M. M. 2014. "Package 'Nbclust'," *Journal of Statistical Software* (61), pp. 1-36.
- Cameron, J., Ramaprasad, A., and Syn, T. 2015. "An ontology of m-health," *Proceedings of 21st Americas Conference on Information Systems, AMCIS 2015*
- Chang, L. W., Kagaayi, J., Arem, H., Nakigozi, G., Ssempijja, V., Serwadda, D., Quinn, T. C., Gray, R. H., Bollinger, R. C., and Reynolds, S. J. 2011. "Impact of A m-health Intervention for Peer Health Workers on Aids Care in Rural Uganda: A Mixed Methods Evaluation of a Cluster-Randomized Trial," *AIDS and Behavior* (15:8), Pp. 1776-1784.
- Chiasson, M., Reddy, M., Kaplan, B., and Davidson, E. 2007. "Expanding Multi-Disciplinary Approaches to Healthcare Information Technologies: What Does Information Systems Offer Medical Informatics?," *International journal of medical informatics* (76), pp. S89-S97.
- Chib, A., van Velthoven, M. H., and Car, J. 2015. "M-health Adoption in Low-Resource Environments: A Review of the Use of Mobile Healthcare in Developing Countries," *Journal of health communication* (20:1), pp. 4-34.
- Cicmil, S., Williams, T., Thomas, J., and Hodgson, D. 2006. "Rethinking Project Management: Researching the Actuality of Projects," *International journal of project management* (24:8), pp. 675-686.
- Ciechanowski, P. S., Walker, E. A., Katon, W. J., and Russo, J. E. 2002. "Attachment Theory: A Model for Health Care Utilization and Somatization," *Psychosomatic medicine* (64:4), pp. 660-667.
- Cleven, A., Gubler, P., and Hüner, K. M. 2009. "Design Alternatives for the Evaluation of Design Science Research Artifacts," *Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology: ACM*, p. 19.
- Clifford, H. T., and Stephenson, W. 1975. *An Introduction to Numerical Classification*. Academic Press New York.

- Conforto, E. C., Salum, F., Amaral, D. C., da Silva, S. L., and de Almeida, L. F. M. 2014. "Can Agile Project Management Be Adopted by Industries Other Than Software Development?," *Project Management Journal* (45:3), pp. 21-34.
- Conroy, D. E., Yang, C.-H., and Maher, J. P. 2014. "Behavior Change Techniques in Top-Ranked Mobile Apps for Physical Activity," *American journal of preventive medicine* (46:6), pp. 649-652.
- Constand, M. K., MacDermid, J. C., Dal Bello-Haas, V., and Law, M. 2014. "Scoping Review of Patient-Centered Care Approaches in Healthcare," *BMC health services research* (14:1), p. 271.
- Cooke-Davies, T. 2002. "The "Real" Success Factors on Projects," *International journal of project management* (20:3), pp. 185-190.
- Cooke-Davies, T., Cicmil, S., Crawford, L., and Richardson, K. 2007. "We're Not in Kansas Anymore, Toto: Mapping the Strange Landscape of Complexity Theory, and Its Relationship to Project Management," *Project Management Journal* (38:2), p. 50.
- Coryn, C. L., Noakes, L. A., Westine, C. D., and Schröter, D. C. 2011. "A Systematic Review of Theory-Driven Evaluation Practice from 1990 to 2009," *American journal of Evaluation* (32:2), pp. 199-226.
- Coyne, R. 2005. "Wicked Problems Revisited," *Design studies* (26:1), pp. 5-17.
- Cross, N. 2001. "Designerly Ways of Knowing: Design Discipline Versus Design Science," *Design issues* (17:3), pp. 49-55.
- Curtis, L., and Netten, A. 2012. *Unit Costs of Health and Social Care*. University of Kent At Can.
- Dasgupta, P., Sen, A., and Marglin, S. 1972. "Guidelines for Project Evaluation," *Guidelines for Project Evaluation*.
- Davenport, T. H., and Short, J. E. 1990. "The New Industrial Engineering: Information Technology and Business Process Redesign,").
- Davis, F. D., and Venkatesh, V. 2004. "Toward Preprototype User Acceptance Testing of New Information Systems: Implications for Software Project Management," *IEEE Transactions on Engineering management* (51:1), pp. 31-46.
- De Leoz, G., and Petter, S. 2018. "Considering the Social Impacts of Artefacts in Information Systems Design Science Research," *European Journal of Information Systems* (27:2), pp. 154-170.
- Delen, D., and Crossland, M. D. 2008. "Seeding the Survey and Analysis of Research Literature with Text Mining," *Expert Systems with Applications* (34:3), pp. 1707-1720.
- DeLone, W. H., and McLean, E. R. 1992. "Information Systems Success: The Quest for the Dependent Variable," *Information systems research* (3:1), pp. 60-95.
- Delone, W. H., and McLean, E. R. 2003. "The Delone and Mclean Model of Information Systems Success: A Ten-Year Update," *Journal of management information systems* (19:4), pp. 9-30.
- Demiris, G., Afrin, L. B., Speedie, S., Courtney, K. L., Sondhi, M., Vimarlund, V., Lovis, C., Goossen, W., and Lynch, C. 2008. "Patient-Centered Applications: Use of Information Technology to Promote Disease Management and Wellness. A White Paper by the AMIA Knowledge in Motion Working Group," *Journal of the American Medical Informatics Association* (15:1), Pp 8-13.

- Denkinger, C., Grenier, J., Stratis, A., Akkihal, A., Pant-Pai, N., and Pai, M. 2013. "Mobile Health to Improve Tuberculosis Care and Control: A Call Worth Making [Review Article]," *The International Journal of Tuberculosis and Lung Disease* (17:6), pp. 719-727.
- Devaraj, S., Ow, T. T., and Kohli, R. 2013. "Examining the Impact of Information Technology and Patient Flow on Healthcare Performance: A Theory of Swift and Even Flow (Tsef) Perspective," *Journal of Operations Management* (31:4), pp. 181-192.
- Ding, C., and He, X. 2004. "K-Means Clustering Via Principal Component Analysis," *Proceedings of the twenty-first international conference on Machine learning: ACM*, p. 29.
- Donaldson, S. I. 2012. *Program Theory-Driven Evaluation Science: Strategies and Applications*. Routledge.
- Doukas, C., Pliakas, T., and Maglogiannis, I. Year. "Mobile Healthcare Information Management Utilizing Cloud Computing and android OS," *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, Pp. 1037-1040.
- Drummond, M., Manca, A., and Sculpher, M. 2005. "Increasing the Generalizability of Economic Evaluations: Recommendations for the Design, Analysis, and Reporting of Studies," *International journal of technology assessment in health care* (21:2), pp. 165-171.
- Duncan, W. R. 1993. "The Process of Project Management," *Project Management Institute*.
- Duncan, W. R. 1996. "A Guide to the Project Management Body of Knowledge,").
- Dunteman, G. H. 1989. *Principal Components Analysis*. Sage.
- Dwivedi, Y. K., Shareef, M. A., Simintiras, A. C., Lal, B., and Weerakkody, V. 2016. "A Generalised Adoption Model for Services: A Cross-Country Comparison of Mobile Health (M-Health)," *Government Information Quarterly* (33:1), pp. 174-187.
- Eccles, M., Grimshaw, J., Walker, A., Johnston, M., and Pitts, N. 2005. "Changing the Behavior of Healthcare Professionals: The Use of Theory in Promoting the Uptake of Research Findings," *Journal of clinical epidemiology* (58:2), pp. 107-112.
- Evaristo, R., and Van Fenema, P. C. 1999. "A Typology of Project Management: Emergence and Evolution of New Forms," *International Journal of Project Management* (17:5), pp. 275-281.
- Eysenbach, G., and Group, C.-E. 2011. "Consort-Ehealth: Improving and Standardizing Evaluation Reports of Web-Based and Mobile Health Interventions," *Journal of medical Internet research* (13:4).
- Fanelli, D. 2012. "Negative Results Are Disappearing from Most Disciplines and Countries," *Scientometrics* (90:3), pp. 891-904.
- Fernandez, D. J., and Fernandez, J. D. 2008. "Agile Project Management—Agilism Versus Traditional Approaches," *Journal of Computer Information Systems* (49:2), pp. 10-17.

- Fichman, R. G., Kohli, R., and Krishnan, R. 2011. "Editorial Overview—the Role of Information Systems in Healthcare: Current Research and Future Trends," *Information Systems Research* (22:3), pp. 419-428.
- Figuerola, Edgar 2018 "Discover Wi-Fi: Security" Wi-Fi Alliance, <https://www.wi-fi.org/discover-wi-fi/security> Accessed: June 29, 2018
- Fishman, B. J., Penuel, W. R., Allen, A.-R., Cheng, B. H., and Sabelli, N. 2013. "Design-Based Implementation Research: An Emerging Model for Transforming the Relationship of Research and Practice," *National society for the study of education* (112:2), pp. 136-156.
- Follette, W. C., and Houts, A. C. 1996. "Models of Scientific Progress and the Role of theory in Taxonomy Development: A Case Study of the DSM," *Journal of Consulting and Clinical Psychology* (64:6), P. 1120.
- Frechtling, J. 2002. "The 2002 User-Friendly Handbook for Project Evaluation," *National Science Foundation* (2012).
- Free, C., Phillips, G., Galli, L., Watson, L., Felix, L., Edwards, P., Patel, V., and Haines, A. 2013. "The Effectiveness of Mobile-Health Technology-Based Health Behaviour Change or Disease Management Interventions for Health Care Consumers: A Systematic Review," *PLoS medicine* (10:1), p. e1001362.
- Friedman, K. 2003. "Theory Construction in Design Research: Criteria: Approaches, and Methods," *Design studies* (24:6), pp. 507-522.
- Gan, G., Ma, C., and Wu, J. 2007. *Data Clustering: Theory, Algorithms, and Applications*. Siam.
- Gauch, H. G. 1980. "Rapid Initial Clustering of Large Data Sets," in *Classification and Ordination*. Springer, pp. 103-111.
- Glaser, F., and Bezenberger, L. 2015. "Beyond Cryptocurrencies-a Taxonomy of Decentralized Consensus Systems,").
- Godin, G., Bélanger-Gravel, A., Eccles, M., and Grimshaw, J. 2008. "Healthcare Professionals' Intentions and Behaviours: A Systematic Review of Studies Based on Social Cognitive Theories," *Implementation Science* (3:1), p. 36.
- Gray, R. M., Cook, M. B., Natera, M. T., Inglis, M. M., and Dodge, M. L. 2008. "Project Management: The," *Managerial Process*, McGraw-Hill: Citeseer.
- Gregor, S. 2006. "The Nature of Theory in Information Systems," *MIS quarterly*), pp. 611-642.
- Gregor, S., and David Jones. 2007. "The Anatomy of a Design Theory." *Journal of the Association for Information Systems* 8, no. 5 (2007): 312.
- Hevner, A. R. "A Three Cycle View of Design Science Research," *Scandinavian Journal of Information Systems* (19:2), P 4.
- Gregor, S., and Hevner, A. R. 2013. "Positioning and Presenting Design Science Research for Maximum Impact," *MIS quarterly* (37:2).
- Gregor, S., and Jones, D. 2007. "The Anatomy of a Design Theory," *Journal of the Association for Information systems* (8:5), p. 312.
- Haas, P., Blohm, I., and Leimeister, J. M. 2014. "An Empirical Taxonomy of Crowdfunding Intermediaries,").
- Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., and Tatham, R. L. 1998. *Multivariate Data Analysis*. Prentice hall Upper Saddle River, NJ.
- Harberger, A. C. 1972. *Project Evaluation: Collected Papers*. Springer.

- Har-Peled, S., Indyk, P., and Motwani, R. 2012. "Approximate Nearest Neighbor: Towards Removing the Curse of Dimensionality," *Theory of computing* (8:1), pp. 321-350.
- Heagney, J. 2016. *Fundamentals of Project Management*. AMACOM Div American Mgmt Assn.
- Heeks, R. 2002. "Information Systems and Developing Countries: Failure, Success, and Local Improvisations," *The information society* (18:2), pp. 101-112.
- Heeks, R. 2006. "Health Information Systems: Failure, Success and Improvisation," *International journal of medical informatics* (75:2), pp. 125-137.
- Hevner, A., and Chatterjee, S. 2010. "Design Science Research in Information Systems," in *Design Research in Information Systems*. Springer, pp. 9-22.
- Hevner, Alan R. "A Three Cycle View of Design Science Research." *Scandinavian journal of information systems* 19.2 (2007): 4.
- Hevner, R., March, S. Park, J., Ram., S. "Design Science in Information Systems Research." *MIS Quarterly* 28, no. 1 (2004): 75-105.
- Hinneburg, A., and Keim, D. A. 1999. "Optimal Grid-Clustering: Towards Breaking the Curse of Dimensionality in High-Dimensional Clustering."
- Hobbs, B., and Petit, Y. 2017. "Agile Methods on Large Projects in Large Organizations," *Project Management Journal* (48:3), pp. 3-19.
- Hotho, A., Nürnberger, A., and Paaß, G. 2005. "A Brief Survey of Text Mining," *Ldv Forum: Citeseer*, pp. 19-62.
- Houle, M. E., Kriegel, H.-P., Kröger, P., Schubert, E., and Zimek, A. 2010. "Can Shared-Neighbor Distances Defeat the Curse of Dimensionality?," *International Conference on Scientific and Statistical Database Management: Springer*, pp. 482-500.
- Hsu, C.-L., and Lin, J. C.-C. 2008. "Acceptance of Blog Usage: The Roles of Technology Acceptance, Social Influence and Knowledge Sharing Motivation," *Information & management* (45:1), pp. 65-74.
- Huang, Z. 1998. "Extensions to the K-Means Algorithm for Clustering Large Data Sets with Categorical Values," *Data mining and knowledge discovery* (2:3), pp. 283-304.
- Hussain, M., Zaidan, A., Zidan, B., Iqbal, S., Ahmed, M., Albahri, O., and Albahri, A. 2018. "Conceptual Framework for the Security of Mobile Health Applications on Android Platform," *Telematics and Informatics* (35:5), pp. 1335-1354.
- Idrish, S., Rifat, A., Iqbal, M., and Nisha, N. 2018. "Mobile Health Technology Evaluation: Innovativeness and Efficacy Vs. Cost Effectiveness," in *Health Economics and Healthcare Reform: Breakthroughs in Research and Practice*. IGI Global, pp. 20-41.
- Indyk, P., and Motwani, R. 1998. "Approximate Nearest Neighbors: Towards Removing the Curse of Dimensionality," *Proceedings of the thirtieth annual ACM symposium on Theory of computing: ACM*, pp. 604-613.
- Iribarren, S. J., Cato, K., Falzon, L., and Stone, P. W. 2017. "What Is the Economic Evidence for M-health? A Systematic Review of Economic Evaluations of M-health Solutions," *PloS one* (12:2), p. e0170581.
- Istepanian, R. S., and Lacal, J. C. 2003. "Emerging Mobile Communication Technologies for Health: Some Imperative Notes on M-Health," *Engineering in*

- Medicine and Biology Society. Proceedings of the 25th Annual International Conference of the IEEE, IEEE2003, Pp. 1414-1416.
- Istepanian, R., Laxminarayan, S., and Pattichis, C. S. 2006. M-health. Springer.
- Jahns, Ralf-Gordon, Pohl, M. 2016 "M-health App Developer Economics 2016 The Current Status and Trends of the M-health App Market." Berlin: M-health Economics Research Program, 1 Oct. 2016. PDF.
- Jain, A. K. 2010. "Data Clustering: 50 Years Beyond K-Means," Pattern recognition letters (31:8), pp. 651-666.
- Jain, A. K., Murty, M. N., and Flynn, P. J. 1999. "Data Clustering: A Review," ACM computing surveys (CSUR) (31:3), pp. 264-323.
- Johnson, S. C. 1967. "Hierarchical Clustering Schemes," Psychometrika (32:3), pp. 241-254.
- Kamsu-Foguem, B., and Foguem, C. 2014. "Telemedicine and Mobile Health with Integrative Medicine in Developing Countries," Health Policy and Technology (3:4), pp. 264-271.
- Kanungo, T., Mount, D. M., Netanyahu, N. S., Piatko, C. D., Silverman, R., and Wu, A. Y. 2002. "An Efficient K-Means Clustering Algorithm: Analysis and Implementation," IEEE transactions on pattern analysis and machine intelligence (24:7), pp. 881-892.
- Kaplan, B., and Harris-Salamone, K. D. 2009. "Health IT Success and Failure: Recommendations from Literature and an AMIA Workshop," Journal of the American Medical Informatics Association (16:3), pp. 291-299.
- Kaufman, L., and Rousseeuw, P. J. 2009. Finding Groups in Data: An Introduction to Cluster Analysis. John Wiley & Sons.
- Kerzner, H., and Kerzner, H. R. 2017. Project Management: A Systems Approach to Planning, Scheduling, and Controlling. John Wiley & Sons.
- Kijsanayotin, B., Pannarunothai, S., and Speedie, S. M. 2009. "Factors Influencing Health Information Technology Adoption in Thailand's Community Health Centers: Applying the Utaut Model," International journal of medical informatics (78:6), pp. 404-416.
- Kinsella, S. M. 2002. "Activity-Based Costing: Does It Warrant Inclusion in a Guide to the Project Management Body of Knowledge (Pmbok Guide)?," Project Management Journal (33:2), pp. 49-49.
- Koplan, Jeffrey P., Robert L. Milstein, and S. Wetterhall. "Framework for program evaluation in public health." *Atlanta, Georgia.: US. Department of Health & Human Services* (1999).
- Kortteisto, T., Kaila, M., Komulainen, J., Mäntyranta, T., and Rissanen, P. 2010. "Healthcare Professionals' Intentions to Use Clinical Guidelines: A Survey Using the Theory of Planned Behaviour," Implementation Science (5:1), p. 51.
- Kotz, D., Avancha, S., and Baxi, A. Year. "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, Journal of the ACM (2009), Pp. 1-12.
- Krishnan, V., and Ulrich, K. T. 2001. "Product Development Decisions: A Review of the Literature," Management science (47:1), pp. 1-21.

- Kuechler, B., and Vaishnavi, V. 2008. "On Theory Development in Design Science Research: Anatomy of a Research Project," *European Journal of Information Systems* (17:5), pp. 489-504.
- Kuechler, W., and Vaishnavi, V. 2012. "A Framework for Theory Development in Design Science Research: Multiple Perspectives," *Journal of the Association for Information systems* (13:6), p. 395.
- Kumar, S., Nilsen, W. J., Abernethy, A., Atienza, A., Patrick, K., Pavel, M., Riley, W. T., Shar, A., Spring, B., and Spruijt-Metz, D. 2013. "Mobile Health Technology Evaluation: The M-health Evidence Workshop," *American journal of preventive medicine* (45:2), pp. 228-236.
- Likas, A., Vlassis, N., and Verbeek, J. J. 2003. "The Global K-Means Clustering Algorithm," *Pattern recognition* (36:2), pp. 451-461.
- Linnan, L., and Steckler, A. 2002. *Process Evaluation for Public Health Interventions and Research*. Jossey-Bass San Francisco.
- Littlejohns, P., Wyatt, J. C., and Garvican, L. 2003. "Evaluating Computerised Health Information Systems: Hard Lessons Still to Be Learnt," *Bmj* (326:7394), pp. 860-863.
- Liu, A. M., and Walker, A. 1998. "Evaluation of Project Outcomes," *Construction Management & Economics* (16:2), pp. 209-219.
- Lorr, M. 1983. *Cluster Analysis for Social Scientists*. Jossey-Bass Inc Pub.
- March, S. T., and Storey, V. C. 2008. "Design Science in the Information Systems Discipline: An Introduction to the Special Issue on Design Science Research," *MIS quarterly*, pp. 725-730.
- March, Salvatore T., and Gerald F. Smith. "Design and Natural Science Research on Information Technology." *Decision support systems* 15.4 (1995): 251-266.
- Marradi, A. 1990. "Classification, Typology, Taxonomy," *Quality and Quantity* (24:2), pp. 129-157.
- Meachael, P. N. 2009. "The Case for m-health in Developing Countries," *innovations* (4:1), pp. 103-118.
- Michalski, R. S., and Stepp, R. E. 1983. "Learning from Observation: Conceptual Clustering," in *Machine Learning, Volume I*. Elsevier, pp. 331-363.
- Michie, S., Abraham, C., Eccles, M. P., Francis, J. J., Hardeman, W., and Johnston, M. 2011. "Strengthening Evaluation and Implementation by Specifying Components of Behaviour Change Interventions: A Study Protocol," *Implementation Science* (6:1), p. 10.
- Mookherji, S., Mehl, G., Kaonga, N., and Meachael, P. 2015. "Unmet Need: Improving M-health Evaluation Rigor to Build the Evidence Base," *Journal of health communication* (20:10), pp. 1224-1229.
- Munns, A., and Bjeirmi, B. F. 1996. "The Role of Project Management in Achieving Project Success," *International journal of project management* (14:2), pp. 81-87.
- Nickerson, R. C., Varshney, U., and Muntermann, J. 2013. "A Method for Taxonomy Development and Its Application in Information Systems," *European Journal of Information Systems* (22:3), Pp. 336-359.
- Olla, P., and Caley Shimskey. "m-health Taxonomy: A Literature Survey of Mobile Health Applications." *Health and Technology* 4, no. 4 (2015): 299-308.
- Patton, M. Q. 1994. "Developmental Evaluation," *Evaluation Practice* (15:3), Pp 311-319.

- Orlikowski, W. J. 1992. "The Duality of Technology: Rethinking the Concept of Technology in Organizations," *Organization science* (3:3), pp. 398-427.
- Orlikowski, W. J., and Baroudi, J. J. 1991. "Studying Information Technology in Organizations: Research Approaches and Assumptions," *Information systems research* (2:1), pp. 1-28.
- Orlikowski, W. J., and Iacono, C. S. 2001. "Research Commentary: Desperately Seeking the "It" in It Research—a Call to Theorizing the It Artifact," *Information systems research* (12:2), pp. 121-134.
- Orloci, L. 1967. "An Agglomerative Method for Classification of Plant Communities," *The Journal of Ecology*, pp. 193-206.
- Park, N., Roman, R., Lee, S., and Chung, J. E. 2009. "User Acceptance of a Digital Library System in Developing Countries: An Application of the Technology Acceptance Model," *International journal of information management* (29:3), pp. 196-209.
- Parsons, J., and Wand, Y. 2008. "Using Cognitive Principles to Guide Classification in Information Systems Modeling," *MIS Quarterly*, pp. 839-868.
- Peppers, K., Tuunanen, T., Rothenberger, M. A., and Chatterjee, S. 2007. "A Design Science Research Methodology for Information Systems Research," *Journal of management information systems* (24:3), pp. 45-77.
- Pelleg, D., and Moore, A. W. 2000. "X-Means: Extending K-Means with Efficient Estimation of the Number of Clusters," *icml*, pp. 727-734.
- Petter, S., DeLone, W., and McLean, E. 2008. "Measuring Information Systems Success: Models, Dimensions, Measures, and Interrelationships," *European journal of information systems* (17:3), pp. 236-263.
- Pich, M. T., Loch, C. H., and Meyer, A. d. 2002. "On Uncertainty, Ambiguity, and Complexity in Project Management," *Management science* (48:8), pp. 1008-1023.
- Plachkinova, M., Steven Andrés, and Samir Chatterjee. 2015. "A Taxonomy of m-health Apps--Security and Privacy Concerns." In *System Sciences (HICSS), 2015 48th Hawaii International Conference on*, pp. 3187-3196. IEEE, 2015.
- Rittel, H. W., and Webber, M. M. 1973. "Dilemmas in a General theory of Planning," *Policy Sciences* (4:2), Pp. 155-169.
- Prat, N., Comyn-Wattiau, I., and Akoka, J. 2015. "A Taxonomy of Evaluation Methods for Information Systems Artifacts," *Journal of Management Information Systems* (32:3), pp. 229-267.
- Pries-Heje, J., and Baskerville, R. 2008. "The Design Theory Nexus," *MIS quarterly*, pp. 731-755.
- Rai, A., Chen, L., Pye, J., and Baird, A. 2013. "Understanding Determinants of Consumer Mobile Health Usage Intentions, Assimilation, and Channel Preferences," *Journal of medical Internet research* (15:8).
- Ranmuthugala, G., Plumb, J. J., Cunningham, F. C., Georgiou, A., Westbrook, J. I., and Braithwaite, J. 2011. "How and Why Are Communities of Practice Established in the Healthcare Sector? A Systematic Review of the Literature," *BMC health services research* (11:1), p. 273.
- Rich, P. 1992. "The Organizational Taxonomy: Definition and Design," *Academy of Management Review* (17:4), pp. 758-781.

- Rittel, H. W., and Webber, M. M. 1973. "Dilemmas in a General Theory of Planning," *Policy sciences* (4:2), pp. 155-169.
- Rose, K., Gurewitz, E., and Fox, G. 1990. "A Deterministic Annealing Approach to Clustering," *Pattern Recognition Letters* (11:9), pp. 589-594.
- Ross, J., and Staw, B. M. 1993. "Organizational Escalation and Exit: Lessons from the Shoreham Nuclear Power Plant," *Academy of Management journal* (36:4), pp. 701-732.
- Rousseeuw, P. J. 1987. "Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis," *Journal of computational and applied mathematics* (20), pp. 53-65.
- Schölkopf, B., Smola, A., and Müller, K.-R. 1997. "Kernel Principal Component Analysis," *International Conference on Artificial Neural Networks*: Springer, pp. 583-588.
- Schwalbe, K. 2015. *Information Technology Project Management*. Cengage Learning.
- Sculpher, M., Pang, F., Manca, A., Drummond, M., Golder, S., Urdahl, H., Davies, L., and Eastwood, A. 2004. "Generalisability in Economic Evaluation Studies in Healthcare: A Review and Case Studies,").
- Shenhar, A. J., and Dvir, D. 1996. "Toward a Typological Theory of Project Management," *Research policy* (25:4), pp. 607-632.
- Silva, Bruno MC, et al. "Mobile-Health: A Review of Current State in 2015." *Journal of biomedical Informatics* 56 (2015): 265-272.
- Sneath, P. H. 1995. "Thirty Years of Numerical Taxonomy," *Systematic Biology* (44:3), Pp. 281-298.
- Sneath, P. H., and Sokal, R. R. 1973. *Numerical Taxonomy. The Principles and Practice of Numerical Classification*.
- Sokal, R. R. 1963. "The Principles and Practice of Numerical Taxonomy," *Taxon*), pp. 190-199.
- Sridharan, S., Campbell, B., and Zinzow, H. 2006. "Developing a Stakeholder-Driven Anticipated Timeline of Impact for Evaluation of Social Programs," *American Journal of Evaluation* (27:2), pp. 148-162.
- Steckler, A. B., Linnan, L., and Israel, B. 2002. *Process Evaluation for Public Health Interventions and Research*. Jossey-Bass San Francisco, CA.
- Steinbach, M., Karypis, G., and Kumar, V. 2000. "A Comparison of Document Clustering Techniques," *KDD workshop on text mining*: Boston, pp. 525-526.
- Straub, D., Keil, M., and Brenner, W. 1997. "Testing the Technology Acceptance Model across Cultures: A Three Country Study," *Information & management* (33:1), pp. 1-11.
- Svejvig, P., and Andersen, P. 2015. "Rethinking Project Management: A Structured Literature Review with a Critical Look at the Brave New World," *International Journal of Project Management* (33:2), pp. 278-290.
- Tesch, D., Kloppenborg, T. J., and Frolick, M. N. 2007. "It Project Risk Factors: The Project Management Professionals Perspective," *Journal of computer information systems* (47:4), pp. 61-69.
- Thomas, A., Menon, A., Boruff, J., Rodriguez, A. M., and Ahmed, S. 2014. "Applications of Social Constructivist Learning Theories in Knowledge Translation for

- Healthcare Professionals: A Scoping Review," *Implementation Science* (9:1), p. 54.
- Thomas, J., and Mengel, T. 2008. "Preparing Project Managers to Deal with Complexity—Advanced Project Management Education," *International journal of project management* (26:3), pp. 304-315.
- Tufféry, S. 2011. *Data Mining and Statistics for Decision Making*. Wiley Chichester.
- Vaishnavi, V. K., and Kuechler, W. 2015. *Design Science Research Methods and Patterns: Innovating Information and Communication Technology*. Crc Press.
- Varshney, U. 2007. "Pervasive Healthcare and Wireless Health Monitoring," *Mobile Networks and Applications* (12:2-3), Pp 113-127.
- Varshney, U. 2014. "Mobile Health: Four Emerging Themes of Research," *Decision Support Systems* (66), Pp 20-35.
- Varshney, U. 2014. "Mobile Health: Four Emerging Themes of Research," *Decision Support Systems* (66), pp. 20-35.
- Vasudevan, L., Zeller, K., and Labrique, A. 2018. "Mobile Health," in *Digital Health*. Springer, pp. 15-25.
- Venable, J. 2006. "The Role of Theory and Theorising in Design Science Research," *Proceedings of the 1st International Conference on Design Science in Information Systems and Technology (DESRIST 2006)*, pp. 1-18.
- Venable, J., Pries-Heje, J., and Baskerville, R. 2012. "A Comprehensive Framework for Evaluation in Design Science Research," *International Conference on Design Science Research in Information Systems: Springer*, pp. 423-438.
- Venable, J., Pries-Heje, J., and Baskerville, R. 2016. "Feds: A Framework for Evaluation in Design Science Research," *European Journal of Information Systems* (25:1), pp. 77-89.
- Venkatesh, V., and Bala, H. 2008. "Technology Acceptance Model 3 and a Research Agenda on Interventions," *Decision sciences* (39:2), pp. 273-315.
- Viney, R., Lancsar, E., and Louviere, J. 2002. "Discrete Choice Experiments to Measure Consumer Preferences for Health and Healthcare," *Expert review of pharmacoeconomics & outcomes research* (2:4), pp. 319-326.
- Von Alan, R. H., March, S. T., Park, J., and Ram, S. 2004. "Design Science in Information Systems Research," *MIS quarterly* (28:1), pp. 75-105.
- Wagstaff, K., Cardie, C., Rogers, S., and Schrödl, S. 2001. "Constrained K-Means Clustering with Background Knowledge," *ICML*, pp. 577-584.
- Wallace, C. S., and Boulton, D. M. 1968. "An Information Measure for Classification," *The Computer Journal* (11:2), pp. 185-194.
- Walshe, K. 2007. "Understanding What Works—and Why—in Quality Improvement: The Need for Theory-Driven Evaluation," *International Journal for Quality in Health Care* (19:2), pp. 57-59.
- Waterlander, W., Whittaker, R., Mcrobbie, H., Dorey, E., Ball, K., Maddison, R., Smith, K. M., Crawford, D., Jiang, Y., and Gu, Y. 2014. "Development of an Evidence-Based m-health Weight Management Program Using a Formative Research Process," *JMIR m-health and uHealth* (2:3).
- Webb, A. R. 2003. *Statistical Pattern Recognition*. John Wiley & Sons.
- Whittaker, R., Merry, S., Dorey, E., and Maddison, R. 2012. "A Development and Evaluation Process for m-health Interventions: Examples from New Zealand,"

- Journal of Health Communication (17:Sup1), Pp 11-21.
- Wikimedia Commons, 2017. User Happenstance: "2016 UN Human Development Report". Created April 8, 2017. accessed December 1, 2017.
- Wilminck, F., and Uytterschaut, H. 1984. "Cluster Analysis, History, Theory and Applications," in Multivariate Statistical Methods in Physical Anthropology. Springer, pp. 135-175.
- Winter, M., Smith, C., Morris, P., and Cicmil, S. 2006. "Directions for Future Research in Project Management: The Main Findings of a Uk Government-Funded Research Network," International journal of project management (24:8), pp. 638-649.
- Wold, S., Esbensen, K., and Geladi, P. 1987. "Principal Component Analysis," Chemometrics and intelligent laboratory systems (2:1-3), pp. 37-52.
- Wysocki, R. K. 2011. Effective Project Management: Traditional, Agile, Extreme. John Wiley & Sons.
- Xu, R., and Wunsch, D. 2005. "Survey of Clustering Algorithms," IEEE Transactions on neural networks (16:3), pp. 645-678.
- Zapata, B. C., Fernández-Alemán, J. L., Idri, A., and Toval, A. 2015. "Empirical Studies on Usability of M-health Apps: A Systematic Literature Review," Journal of medical systems (39:2), p. 1.

Appendices

A1. References for 104 documents included in final taxonomy

- Ahmed, T., Lucas, H., Khan, A. S., Islam, R., Bhuiya, A., and Iqbal, M. 2014. "Ehealth and M-health Initiatives in Bangladesh: A Scoping Study," *Bmc Health Services Research* (14), p. 9.
- Aikens, J. E., Trivedi, R., Heapy, A., Pfeiffer, P. N., and Piette, J. D. 2015. "Potential Impact of Incorporating a Patient-Selected Support Person into M-health for Depression," *Journal of General Internal Medicine* (30:6), pp. 797-803.
- Allman-Farinelli, M., Partridge, S. R., McGeechan, K., Balestracci, K., Hebden, L., Wong, A., Phongsavan, P., Denney-Wilson, E., Harris, M. F., and Bauman, A. 2016. "A Mobile Health Lifestyle Program for Prevention of Weight Gain in Young Adults (Txt2bfit): Nine-Month Outcomes of a Randomized Controlled Trial," *JMIR m-health and uHealth* (4:2).
- Amaro, P. I. N. 2017. "Designing Evaluation Plans for M-health Interventions: A Case Study on the Weal Life App."
- Anantraman, V., Mikkelsen, T., Khilnani, R., Kumar, V. S., Pentland, A., and Ohno-Machado, L. 2002. "Open Source Handheld-Based Emr for Paramedics Working in Rural Areas," *Proceedings of the AMIA Symposium: American Medical Informatics Association*, p. 12.
- Andersson, B., Wu, T.-S. J., ODonoghue, J., Mastellos, N., Thompson, M., Hardy, V., Ide, N., Heavin, C., O'Connor, Y., and Carlsson, S. 2017. "Enhancing Information Quality as Part of the Disease Surveillance System in Malawi, Africa: Reflections on a M-health Intervention,").
- Andry, F., Wan, L., and Nicholson, D. 2011. "A Mobile Application Accessing Patientshealth Records through a Rest Api," *HEALTHINF 2011 INTERNATIONAL CONFERENCE ON HEALTH INFORMATICS*, pp. 27-32.
- Ashcroft, N., Shelus, V., Garg, H., McLarnon-Silk, C., and Jennings, V. H. 2017. "Implementation of Cycletel Family Advice: An Sms-Based Service to Provide Family Planning and Fertility Awareness Information in India," *M-health* (3).
- Asiimwe, C., Gelvin, D., Lee, E., Amor, Y. B., Quinto, E., Katureebe, C., Sundaram, L., Bell, D., and Berg, M. 2011. "Use of an Innovative, Affordable, and Open-Source Short Message Service–Based Tool to Monitor Malaria in Remote Areas of Uganda," *The American journal of tropical medicine and hygiene* (85:1), pp. 26-33.
- Ben-Zeev, D., Scherer, E. A., Wang, R., Xie, H., and Campbell, A. T. 2015. "Next-Generation Psychiatric Assessment: Using Smartphone Sensors to Monitor Behavior and Mental Health," *Psychiatric rehabilitation journal* (38:3), p. 218.
- Brazionis, L., Jenkins, A., Keech, A., Ryan, C., and Bursell, S.-E. 2017. "An Evaluation of the Telehealth Facilitation of Diabetes and Cardiovascular Care in Remote Australian Indigenous Communities:-Protocol for the Telehealth Eye and Associated Medical Services Network [Teamsnet] Project, a Pre-Post Study Design," *BMC health services research* (17:1), p. 13.
- Brown, W., III, Ibitoye, M., Bakken, S., Schnall, R., Balan, I., Frasca, T., and Carballo-Diequez, A. 2015. "Cartographic Analysis of Antennas and Towers: A Novel

- Approach to Improving the Implementation and Data Transmission of M-health Tools on Mobile Networks," *Jmir M-health and Uhealth* (3:2).
- Burgess, L., Cooper, J., and Sargent, J. P. 2008. "Optimising User Acceptance of Mandated Mobile Health Systems (Mhs): The Epoc (Electronic Point-of-Care) Project Experience,").
- Burgess, L., and Sargent, J. 2007. "Enhancing User Acceptance of Mandated Mobile Health Information Systems: The Epoc (Electronic Point-of-Care Project) Experience,").
- Burn, E., Nghiem, S., Jan, S., Redfern, J., Rodgers, A., Thiagalingam, A., Graves, N., and Chow, C. K. 2017. "Cost-Effectiveness of a Text Message Programme for the Prevention of Recurrent Cardiovascular Events," *Heart*), pp. heartjnl-2016-310195.
- Bursell, S. E., Brazionis, L., and Jenkins, A. 2012. "Telemedicine and Ocular Health in Diabetes Mellitus," *Clinical and Experimental Optometry* (95:3), pp. 311-327.
- Capela, N., Lemaire, E., Baddour, N., Rudolf, M., Goljar, N., and Burger, H. 2016. "Evaluation of a Smartphone Human Activity Recognition Application with Able-Bodied and Stroke Participants," *Journal of neuroengineering and rehabilitation* (13:1), p. 5.
- Chang, S., Plotkin, D., Mulligan, R., Polido, J., Mah, J., and Meara, J. 2003. "Teledentistry in Rural California: A Usc Initiative," *J Calif Dent Assoc* (31).
- Cherrington, A. L., Agne, A. A., Lampkin, Y., Birl, A., Shelton, T. C., Guzman, A., and Willig, J. H. 2015. "Diabetes Connect: Developing a Mobile Health Intervention to Link Diabetes Community Health Workers with Primary Care," *The Journal of ambulatory care management* (38:4), pp. 333-345.
- De Tolly, K., Alexander, H., and Cell-Life, C. T. 2009. Innovative Use of Cellphone Technology for Hiv/Aids Behaviour Change Communications: 3 Pilot Projects. World wide web consortium (W3C).
- DeRenzi, B., Borriello, G., Jackson, J., Kumar, V. S., Parikh, T. S., Virk, P., and Lesh, N. 2011. "Mobile Phone Tools for Field-Based Health Care Workers in Low-Income Countries," *Mount Sinai Journal of Medicine: A Journal of Translational and Personalized Medicine* (78:3), pp. 406-418.
- DeVito Dabbs, A., Song, M., Myers, B., Li, R., Hawkins, R., Pilewski, J., Bermudez, C., Aubrecht, J., Begey, A., and Connolly, M. 2016. "A Randomized Controlled Trial of a Mobile Health Intervention to Promote Self-Management after Lung Transplantation," *American Journal of Transplantation* (16:7), pp. 2172-2180.
- Eysenbach, G., and Group, C.-E. 2011. "Consort-Ehealth: Improving and Standardizing Evaluation Reports of Web-Based and Mobile Health Interventions," *Journal of medical Internet research* (13:4).
- Fang, X., Wang, Q., Zhu, G., and Li, S. 2015. "An Intelligent Doctor-Patient Interaction System for Chronic Disease Telemonitoring," in *Proceedings of the 2015 2nd International Conference on Electrical, Computer Engineering and Electronics*, T. Chen and L. Xu (eds.). pp. 1206-1211.
- Ford, J. H., II, Alagoz, E., Dinauer, S., Johnson, K. A., Pe-Romashko, K., and Gustafson, D. H. 2015. "Successful Organizational Strategies to Sustain Use of a-Chess: A Mobile Intervention for Individuals with Alcohol Use Disorders," *Journal of Medical Internet Research* (17:8).

- Furber, G., Jones, G. M., Healey, D., and Bidargaddi, N. 2014. "A Comparison between Phone-Based Psychotherapy with and without Text Messaging Support in between Sessions for Crisis Patients," *Journal of Medical Internet Research* (16:10), pp. 43-54.
- García, I. P., Sánchez, C. M., Espilez, Á. S., García-Magariño, I., Guillen, G. A., and García-Campayo, J. 2017. "Development and Initial Evaluation of a Mobile Application to Help with Mindfulness Training and Practice," *International Journal of Medical Informatics* (105), pp. 59-67.
- Genz, A., Kirk, G., Piggott, D., Mehta, S. H., Linas, B. S., and Westergaard, R. P. 2015. "Uptake and Acceptability of Information and Communication Technology in a Community-Based Cohort of People Who Inject Drugs: Implications for Mobile Health Interventions," *Jmir M-health and Uhealth* (3:2).
- Geoffroy, E., Harries, A. D., Bissell, K., Schell, E., Bvumbwe, A., Tayler-Smith, K., and Kizito, W. 2014. "Bringing Care to the Community: Expanding Access to Health Care in Rural Malawi through Mobile Health Clinics," *Public health action* (4:4), pp. 252-258.
- Georgsson, M., Staggers, N., and Weir, C. 2016. "A Modified User-Oriented Heuristic Evaluation of a Mobile Health System for Diabetes Self-Management Support," *Computers, Informatics, Nursing* (34:2), p. 77.
- Germanakos, P., Mourlas, C., and Samaras, G. 2005. "A Mobile Agent Approach for Ubiquitous and Personalized Ehealth Information Systems," *Proceedings of the Workshop on 'Personalization for e-Health' of the 10th International Conference on User Modeling (UM 2005)*, pp. 67-70.
- Ginsburg, A. S., Delarosa, J., Brunette, W., Levvari, S., Sundt, M., Larson, C., Agyemang, C. T., Newton, S., Borriello, G., and Anderson, R. 2015. "Mpneumonia: Development of an Innovative M-health Application for Diagnosing and Treating Childhood Pneumonia and Other Childhood Illnesses in Low-Resource Settings," *Plos One* (10:10).
- Gordon, J. S., Armin, J., Hingle, M. D., Giacobbi Jr, P., Cunningham, J. K., Johnson, T., Abbate, K., Howe, C. L., and Roe, D. J. 2017. "Development and Evaluation of the See Me Smoke-Free Multi-Behavioral M-health App for Women Smokers," *Translational behavioral medicine* (7:2), pp. 172-184.
- Goyal, S., Morita, P., Lewis, G. F., Yu, C., Seto, E., and Cafazzo, J. A. 2016. "The Systematic Design of a Behavioural Mobile Health Application for the Self-Management of Type 2 Diabetes," *Canadian journal of diabetes* (40:1), pp. 95-104.
- Herbert, L. J., Mehta, P., Monaghan, M., Cogen, F., and Streisand, R. 2014. "Feasibility of the Smart Project: A Text Message Program for Adolescents with Type 1 Diabetes," *Diabetes Spectrum* (27:4), pp. 265-269.
- Heslop, L., Weeding, S., Dawson, L., Fisher, J., and Howard, A. 2010. "Implementation Issues for Mobile-Wireless Infrastructure and Mobile Health Care Computing Devices for a Hospital Ward Setting," *Journal of medical systems* (34:4), pp. 509-518.
- Holmen, H., Wahl, A., Torbjørnsen, A., Jenum, A. K., Småstuen, M. C., and Ribu, L. 2016. "Stages of Change for Physical Activity and Dietary Habits in Persons with

- Type 2 Diabetes Included in a Mobile Health Intervention: The Norwegian Study in Renewing Health," *BMJ Open Diabetes Research and Care* (4:1), p. e000193.
- Huang, A., and Xie, L. 2015. "Smart for Mobile Health: A Study of Scheduling Algorithms in Full-Ip Mobile Networks," *Communications Magazine, IEEE* (53:2), pp. 214-222.
- Huang, A., Xu, W., Li, Z., Xie, L., Sarrafzadeh, M., Li, X., and Cong, J. 2014. "System Light-Loading Technology for M-health: Manifold-Learning-Based Medical Data Cleansing and Clinical Trials in We-Care Project," *Biomedical and Health Informatics, IEEE Journal of* (18:5), pp. 1581-1589.
- Iribarren, S. J., Sward, K. A., Beck, S. L., Pearce, P. F., Thurston, D., and Chirico, C. 2015. "Qualitative Evaluation of a Text Messaging Intervention to Support Patients with Active Tuberculosis: Implementation Considerations," *Jmir M-health and Uhealth* (3:1).
- Jones, V., Rensink, A., and Brinksma, E. 2005. "Modelling Mobile Health Systems: An Application of Augmented Mda for the Extended Healthcare Enterprise," *Ninth IEEE International EDOC Enterprise Computing Conference (EDOC'05): IEEE*, pp. 58-69.
- Kallander, K., Strachan, D., Soremekun, S., Hill, Z., Lingam, R., and Tibenderana, J. 2015. "Evaluating the Effect of Innovative Motivation and Supervision Approaches on Community Health Worker Performance and Retention in Uganda and Mozambique: Study Protocol for a Randomised Controlled Trial," *Trials* (16:1), p. 157.
- Kaonga, N. N., Labrique, A., Mechael, P., Akosah, E., Ohemeng-Dapaah, S., Baah, J. S., Kodie, R., Kanter, A. S., and Levine, O. 2013. "Using Social Networking to Understand Social Networks: Analysis of a Mobile Phone Closed User Group Used by a Ghanaian Health Team," *Journal of medical Internet research* (15:4).
- Kapoor, L., Mishra, S., and Singh, K. 2005. "Telemedicine: Experience at Sgpgims, Lucknow," *Journal of postgraduate medicine* (51:4), p. 312.
- Kimaro, H. C., and Nhampossa, J. L. 2005. "Analyzing the Problem of Unsustainable Health Information Systems in Less-Developed Economies: Case Studies from Tanzania and Mozambique," *Information Technology for Development* (11:3), pp. 273-298.
- Kleinman, N. J., Shah, A., Shah, S., Phatak, S., and Viswanathan, V. 2017. "Improved Medication Adherence and Frequency of Blood Glucose Self-Testing Using an M-Health Platform Versus Usual Care in a Multisite Randomized Clinical Trial among People with Type 2 Diabetes in India," *Telemedicine and e-Health* (23:9), pp. 733-740.
- Lang, A. R., Martin, A., Condon, L., McKinstry, B., Atkinson, S., Gómez, S., Puigdomènech, E., Adorni, F., and Rashid, R. 2017. "Promoting Healthy Teenage Behaviour through the Use of a Novel M-health Technology Platform: Pegaso, Developed with and for Teenagers," *Journal of Adolescent Health* (60:2), pp. S65-S66.
- Lanzola, G., Ginardi, M., Mazzanti, A., and Quaglini, S. 2014. "Gquest: Modeling Patient Questionnaires and Administering Them through a Mobile Platform Application," *Computer methods and programs in biomedicine* (117:2), pp. 277-291.

- Lee, Y., and Moon, M. 2016. "Utilization and Content Evaluation of Mobile Applications for Pregnancy, Birth, and Child Care," *Healthcare informatics research* (22:2), pp. 73-80.
- Lv, Z., Chirivella, J., and Gagliardo, P. 2016. "Bigdata Oriented Multimedia Mobile Health Applications," *Journal of medical systems* (40:5), p. 120.
- Lwin, M. O., Vijaykumar, S., Lim, G., Fernando, O. N. N., Rathnayake, V. S., and Foo, S. 2016. "Baseline Evaluation of a Participatory Mobile Health Intervention for Dengue Prevention in Sri Lanka," *Health Education & Behavior* (43:4), pp. 471-479.
- Mann, D. M., Quintiliani, L. M., Reddy, S., Kitos, N. R., and Weng, M. 2014. "Dietary Approaches to Stop Hypertension: Lessons Learned from a Case Study on the Development of an M-health Behavior Change System," *JMIR m-health and uHealth* (2:4), p. e41.
- Massey, T., Marfia, G., Potkonjak, M., and Sarrafzadeh, M. 2010. "Experimental Analysis of a Mobile Health System for Mood Disorders," *Information Technology in Biomedicine, IEEE Transactions on* (14:2), pp. 241-247.
- Masters, K. 2014. "Health Professionals as Mobile Content Creators: Teaching Medical Students to Develop M-health Applications," *Medical teacher* (36:10), pp. 883-889.
- Mburu, S. 2017. "A Predictive Model for Optimizing Acceptance and Use of M-health Interventions in Low-Resource Settings: A Case of Mamacare Prototype," *AFRICON, 2017 IEEE: IEEE*, pp. 518-523.
- Mehralizade, A., Schor, S., Coleman, C. M., Oppenheim, C. E., Denckla, C. A., Borba, C. P., Henderson, D. C., Wolff, J., Crane, S., and Nettles-Gomez, P. 2017. "Mobile Health Apps in Ob-Gyn-Embedded Psychiatric Care: Commentary," *JMIR m-health and uHealth* (5:10).
- Menshaw, M. E., Benharref, A., and Serhani, M. 2015. "An Automatic Mobile-Health Based Approach for Eeg Epileptic Seizures Detection," *Expert Systems with Applications*.
- Mercer, K., Giangregorio, L., Schneider, E., Chilana, P., Li, M., and Grindrod, K. 2016. "Acceptance of Commercially Available Wearable Activity Trackers among Adults Aged over 50 and with Chronic Illness: A Mixed-Methods Evaluation," *JMIR m-health and uHealth* (4:1).
- Mohammadzadeh, N., Safdari, R., and Rahimi, A. 2013. "Cancer Care Management through a Mobile Phone Health Approach: Key Considerations," *Asian Pacific Journal of Cancer Prevention* (14:9), pp. 4961-4964.
- Mukherjee, A., Purkayastha, S., and Sahay, S. 2010. "Exploring the Potential and Challenges of Using Mobile Based Technology in Strengthening Health Information Systems: Experiences from a Pilot Study," *AMCIS*, p. 263.
- Mushamiri, I., Luo, C., Iiams-Hauser, C., and Amor, Y. B. 2015. "Evaluation of the Impact of a Mobile Health System on Adherence to Antenatal and Postnatal Care and Prevention of Mother-to-Child Transmission of Hiv Programs in Kenya," *BMC public health* (15:1), p. 1.
- Nhavoto, J. A., Gronlund, A., and Chaquilla, W. P. 2015. "Smsaude: Design, Development, and Implementation of a Remote/Mobile Patient Management

- System to Improve Retention in Care for Hiv/Aids and Tuberculosis Patients," *Jmir M-health and Uhealth* (3:1).
- Nundy, S., Dick, J. J., Chou, C.-H., Nocon, R. S., Chin, M. H., and Peek, M. E. 2014. "Mobile Phone Diabetes Project Led to Improved Glycemic Control and Net Savings for Chicago Plan Participants," *Health Affairs* (33:2), pp. 265-272.
- O'Connor, Y., Hardy, V., Heavin, C., Gallagher, J., and O'Donoghue, J. 2015. "Supporting Life: Mobile Health Application for Classifying, Treating and Monitoring Disease Outbreaks of Sick Children in Developing Countries," in *New Horizons in Design Science: Broadening the Research Agenda*. Springer, pp. 366-370.
- Oren, E., Bell, M. L., Garcia, F., Perez-Velez, C., and Gerald, L. B. 2017. "Promoting Adherence to Treatment for Latent Tb Infection through Mobile Phone Text Messaging: Study Protocol for a Pilot Randomized Controlled Trial," *Pilot and feasibility studies* (3:1), p. 15.
- Oren, E., Bell, M. L., Garcia, F., Perez-Velez, C., and Gerald, L. B. 2017. "Promoting Adherence to Treatment for Latent Tb Infection through Mobile Phone Text Messaging: Study Protocol for a Pilot Randomized Controlled Trial," *Pilot and feasibility studies* (3:1), p. 15.
- Pande, T. 2017. "Developing and Evaluating a Smartphone Application for Tuberculosis Amongst Private Sector Academic Clinicians in India." Université Laval.
- Parmanto, B., Pramana, G., Yu, D. X., Fairman, A. D., Dicianno, B. E., and McCue, M. P. 2013. "Imhere: A Novel M-health System for Supporting Self-Care in Management of Complex and Chronic Conditions," *JMIR m-health and uHealth* (1:2), p. e10.
- Poropatich, R. K., Pavlisacsak, H. H., Tong, J. C., Little, J. R., and McVeigh, F. L. 2014. "Mcare: Using Secure Mobile Technology to Support Soldier Reintegration and Rehabilitation," *Telemedicine and e-Health* (20:6), pp. 563-569.
- Quanbeck, A. R., Gustafson, D. H., Marsch, L. A., McTavish, F., Brown, R. T., Mares, M. L., Johnson, R., Glass, J. E., Atwood, A. K., and McDowell, H. 2014. "Integrating Addiction Treatment into Primary Care Using Mobile Health Technology: Protocol for an Implementation Research Study," *Implementation Science* (9), p. 11.
- Quintiliani, L. M., Mann, D. M., Puputti, M., Quinn, E., and Bowen, D. J. 2016. "Pilot and Feasibility Test of a Mobile Health-Supported Behavioral Counseling Intervention for Weight Management among Breast Cancer Survivors," *JMIR cancer* (2:1).
- Santos, J., Rodrigues, J. J., Silva, B. M., Casal, J., Saleem, K., and Denisov, V. 2016. "An Iot-Based Mobile Gateway for Intelligent Personal Assistants on Mobile Health Environments," *Journal of Network and Computer Applications* (71), pp. 194-204.
- Schnall, R., Rojas, M., Bakken, S., Brown III, W., Carballo-Diequez, A., Carry, M., Gelaude, D., Mosley, J. P., and Travers, J. 2016. "A User-Centered Model for Designing Consumer Mobile Health (M-health) Applications (Apps)," *Journal of biomedical informatics* (60), pp. 243-251.
- Schuttner, L., Sindano, N., Theis, M., Zue, C., Joseph, J., Chilengi, R., Chi, B. H., Stringer, J. S., and Chintu, N. 2014. "A Mobile Phone-Based, Community Health

- Worker Program for Referral, Follow-up, and Service Outreach in Rural Zambia: Outcomes and Overview," *Telemedicine and e-Health* (20:8), pp. 721-728.
- Schwitters, A., Lederer, P., Zilversmit, L., Gudo, P. S., Ramiro, I., Cumba, L., Mahagaja, E., and Jobarteh, K. 2015. "Barriers to Health Care in Rural Mozambique: A Rapid Ethnographic Assessment of Planned Mobile Health Clinics for Art," *Global Health-Science and Practice* (3:1), pp. 109-116.
- Serrano, K. J., Yu, M., Riley, W. T., Patel, V., Hughes, P., Marchesini, K., and Atienza, A. A. 2016. "Willingness to Exchange Health Information Via Mobile Devices: Findings from a Population-Based Survey," *The Annals of Family Medicine* (14:1), pp. 34-40.
- Shieshia, M., Noel, M., Andersson, S., Felling, B., Alva, S., Agarwal, S., Lefevre, A., Misomali, A., Chimphanga, B., and Nsona, H. 2014. "Strengthening Community Health Supply Chain Performance through an Integrated Approach: Using M-health Technology and Multilevel Teams in Malawi," *Journal of global health* (4:2), pp. 020406-020406.
- Sibeko, G., Temmingh, H., Mall, S., Williams-Ashman, P., Thornicroft, G., Susser, E. S., Lund, C., Stein, D. J., and Milligan, P. D. 2017. "Improving Adherence in Mental Health Service Users with Severe Mental Illness in South Africa: A Pilot Randomized Controlled Trial of a Treatment Partner and Text Message Intervention Vs. Treatment as Usual," *BMC research notes* (10:1), p. 584.
- Singh, K., Drouin, K., Newmark, L. P., Rozenblum, R., Lee, J., Landman, A., Pabo, E., Klinger, E. V., and Bates, D. W. 2016. "Developing a Framework for Evaluating the Patient Engagement, Quality, and Safety of Mobile Health Applications," *Issue Brief (Commonw Fund)* (5:1), p. 11.
- Smordal, O., Moen, A., Kristiansen, T., Refseth, Y., Mortensen, B., and Osnes, T. 2004. "Participatory Design and Infrastructures for Coordination and Communication in Interdisciplinary Mobile Health Teams," *Stud Health Technol Inform* (107), pp. 1443-1447.
- Somers, T. J., Abernethy, A. P., Edmond, S. N., Kelleher, S. A., Wren, A. A., Samsa, G. P., and Keefe, F. J. 2015. "A Pilot Study of a Mobile Health Pain Coping Skills Training Protocol for Patients with Persistent Cancer Pain," *Journal of Pain and Symptom Management*.
- Spyrou, S., Vartzopoulos, D., Bamidis, P., and Maglaveras, N. 2007. "Cross-Border Collaboration between Greece and Fyrom: Mobile Healthcare Provision," *Studies in health technology and informatics* (136), pp. 653-658.
- Stoyanov, S. R., Hides, L., Kavanagh, D. J., and Wilson, H. 2016. "Development and Validation of the User Version of the Mobile Application Rating Scale (Umars)," *JMIR m-health and uHealth* (4:2).
- Strachan, D. L., Källander, K., ten Asbroek, A. H., Kirkwood, B., Meek, S. R., Benton, L., Conteh, L., Tibenderana, J., and Hill, Z. 2012. "Interventions to Improve Motivation and Retention of Community Health Workers Delivering Integrated Community Case Management (Iccm): Stakeholder Perceptions and Priorities," *The American journal of tropical medicine and hygiene* (87:5 Suppl), pp. 111-119.

- Strong, D., Tulu, B., Agu, E., and He, S. Q. 2014. "Design of the Feedback Engine for a Diabetes Self-Care Smartphone App,"). In Twentieth Americas Conference on Information Systems.
- Szarvas, I., Nitsch, M., Lambacher, O., Howe, J., Reichwaldt, N., von Bargaen, T., and Haux, R. 2012. "Possibilities and Acceptance of Mobile Information Technologies for the Elderly-First Results of an Empirical Study," *Studies in health technology and informatics* (192), pp. 948-948.
- Takizawa, M., Sone, S., Hanamura, K., and Asakura, K. 2001. "Telemedicine System Using Computed Tomography Van of High-Speed Telecommunication Vehicle," *Information Technology in Biomedicine, IEEE Transactions on* (5:1), pp. 2-9.
- Tewari, A., Kallakuri, S., Devarapalli, S., Jha, V., Patel, A., and Maulik, P. K. 2017. "Process Evaluation of the Systematic Medical Appraisal, Referral and Treatment (Smart) Mental Health Project in Rural India," *BMC psychiatry* (17:1), p. 385.
- Vallely, A., Shagi, C., Kasindi, S., Desmond, N., Lees, S., Chiduo, B., Hayes, R., Allen, C., and Ross, D. 2007. "The Benefits of Participatory Methodologies to Develop Effective Community Dialogue in the Context of a Microbicide Trial Feasibility Study in Mwanza, Tanzania," *Bmc Public Health* (7).
- Van Halteren, A., Bults, R., Wac, K., Dokovsky, N., Koprinkov, G., Widya, I., Konstantas, D., Jones, V., and Herzog, R. 2004. "Wireless Body Area Networks for Healthcare: The Mobihealth Project," *Stud Health Technol Inform* (108), pp. 181-193.
- Van Olmen, J., Van Pelt, M., Malombo, B., Ku, G. M., Kanda, D., Heang, H., Darras, C., Kegels, G., and Schellevis, F. 2017. "Process Evaluation of a Mobile Health Intervention for People with Diabetes in Low Income Countries—the Implementation of the Text4dsm Study," *Journal of telemedicine and telecare* (23:1), pp. 96-105.
- Van Reijen, M., Vriend, I. I., Zuidema, V., van Mechelen, W., and Verhagen, E. A. 2014. "The Implementation Effectiveness of the 'Strengthen Your Ankle'smartphone Application for the Prevention of Ankle Sprains: Design of a Randomized Controlled Trial," *BMC musculoskeletal Disorders* (15:1), p. 1.
- Vedanthan, R., Blank, E., Tuikong, N., Kamano, J., Misoi, L., Tuliengo, D., Hutchinson, C., Ascheim, D. D., Kimaiyo, S., Fuster, V., and Were, M. C. 2015. "Usability and Feasibility of a Tablet-Based Decision-Support and Integrated Record-Keeping (Desire) Tool in the Nurse Management of Hypertension in Rural Western Kenya," *International Journal of Medical Informatics* (84:3), pp. 207-219.
- Vélez, O., Okyere, P. B., Kanter, A. S., and Bakken, S. 2014. "A Usability Study of a Mobile Health Application for Rural Ghanaian Midwives," *Journal of Midwifery & Women's Health* (59:2), pp. 184-191.
- Vidrine, D. J., Fletcher, F. E., Danysh, H. E., Marani, S., Vidrine, J. I., Cantor, S. B., and Prokhorov, A. V. 2012. "A Randomized Controlled Trial to Assess the Efficacy of an Interactive Mobile Messaging Intervention for Underserved Smokers: Project Action," *BMC Public Health* (12:1), p. 696.
- Wakholi, P., Chen, W., and Klungsøyr, J. 2011. "Workflow Support for Mobile Data Collection," in *Enterprise, Business-Process and Information Systems Modeling*. Springer, pp. 299-313.

- Wang, J. B., Cadmus-Bertram, L. A., Natarajan, L., White, M. M., Madanat, H., Nichols, J. F., Ayala, G. X., and Pierce, J. P. 2015. "Wearable Sensor/Device (Fitbit One) and Sms Text-Messaging Prompts to Increase Physical Activity in Overweight and Obese Adults: A Randomized Controlled Trial," *Telemedicine and e-Health* (21:10), pp. 782-792.
- Wang, S.-L., Chen, Y. L., Kuo, A. M.-H., Chen, H.-M., and Shiu, Y. S. 2016. "Design and Evaluation of a Cloud-Based Mobile Health Information Recommendation System on Wireless Sensor Networks," *Computers & Electrical Engineering* (49), pp. 221-235.
- Wannenburg, J., and Malekian, R. 2015. "Body Sensor Network for Mobile Health Monitoring, a Diagnosis and Anticipating System,").
- Wickramasinghe, N., Goldberg, S., and Bali, R. K. 2008. "Enabling Superior M-Health Project Success: A Tricountry Validation," *International Journal of Services and Standards* (4:1), pp. 97-117.
- Worthen-Chaudhari, L., McGonigal, J., Logan, K., Bockbrader, M. A., Yeates, K. O., and Mysiw, W. J. 2017. "Reducing Concussion Symptoms among Teenage Youth: Evaluation of a Mobile Health App," *Brain injury* (31:10), pp. 1279-1286.
- Xu, B., Xu, L., Cai, H., Jiang, L., Luo, Y., and Gu, Y. 2017. "The Design of an M-Health Monitoring System Based on a Cloud Computing Platform," *Enterprise Information Systems* (11:1), pp. 17-36.
- Yang, C., Yang, J., Luo, X., and Gong, P. 2009. "Use of Mobile Phones in an Emergency Reporting System for Infectious Disease Surveillance after the Sichuan Earthquake in China," *Bulletin of the World Health Organization* (87:8), pp. 619-623.
- Zurovac, D., Sudoi, R. K., Akhwale, W. S., Ndiritu, M., Hamer, D. H., Rowe, A. K., and Snow, R. W. 2011. "The Effect of Mobile Phone Text-Message Reminders on Kenyan Health Workers' Adherence to Malaria Treatment Guidelines: A Cluster Randomised Trial," *The Lancet* (378:9793), pp. 795-803.

A2. Taxonomy survey

Please compare the following two taxonomies for mobile health projects:

Taxonomy A

Primary Stakeholders	Patients
	Providers
Type of Study	Field Study
	Randomized Control Trial
	Simulation
Evaluation Setting	Developing Country
	Developed Country
Primary Evaluation Metric	Cost-based
	Technology - based
	User-based
Duration of evaluation	0 to 6 months
	6+ months
(So what) Evaluation Conclusion	Positive
	Negative
	Neutral
Theory application	Exists
	Does not exist
Number of evaluation stages	1-Stage
	2-Stages
	3-Stages or More

Taxonomy B

Type of Network	3G or older
	4G or newer
Application usage?	Yes
	No
Determinant of Success	Provider outcomes
	Patient outcomes
Theory Based	Yes
	No
Impact on Existing Healthcare System	Incremental Improvement
	Introduction of New System
Duration	0-6 Weeks
	6-18 Weeks
Study Type	Simulation
	Field Study
Project Setting	Developed Country
	Developing Country

- 1) Which taxonomy do you think is more effective for evaluating mobile health projects: A or B? Why?

- 2) Do the taxonomies cover all the relevant dimensions and characteristics* of mobile health projects, to your knowledge?

*Dimensions are on the left side of the taxonomies (for example, duration, study type), characteristics are on the right side (for example, 0-6 weeks or 6-18 week, simulation or field study).

- 3) Are there any unnecessary dimensions you feel could be removed from either taxonomy?

- 4) Do you have any suggestions for additional dimensions not listed in either taxonomy?

- 5) Do you have any comments regarding either taxonomy not covered by the questions above?

A3. Words removed from clustering

- techn, ne, one, however, dure, include, al, et, how, eg, figure, require, issu, because, facle, improv, due, phon, ex, via, environ, imple, found, l, per, sett, compar, integrat, table

A4. R code for clustering

Note: lines preceded by # represent comments

Libraries

Cluster, Stats, Factoextra

Silhouette code

#based on analysis performed in Charrad et al. 2014), comments by Alan Yang

```

#determining variables for average silhouette method
abline(v = which.max(sil), lty = 2)
maxClusters <- 15
sil <- rep(0, maxClusters)
#loop to determine optimal number of clusters based on silhouette width
for(i in 2:k.max){
  km.res <- kmeans(data, centers = i, nstart = 25)
#data variable contains the file with word information
#note: keep different cluster analysis separate
#km.res$cluster variable contains km.res info from previous line
  ss <- silhouette(km.res$cluster, dist(data))
  sil[i] <- mean(ss[, 3])
}
#display graph with maxClusters, set to 15
plot(1:k.max, sil, type = "b", pch = 19,
  frame = FALSE, xlab = "Number of clusters k")
abline(v = which.max(sil), lty = 2)
#largest y-axis value is the optimal clustering amount

```

K-means code

```

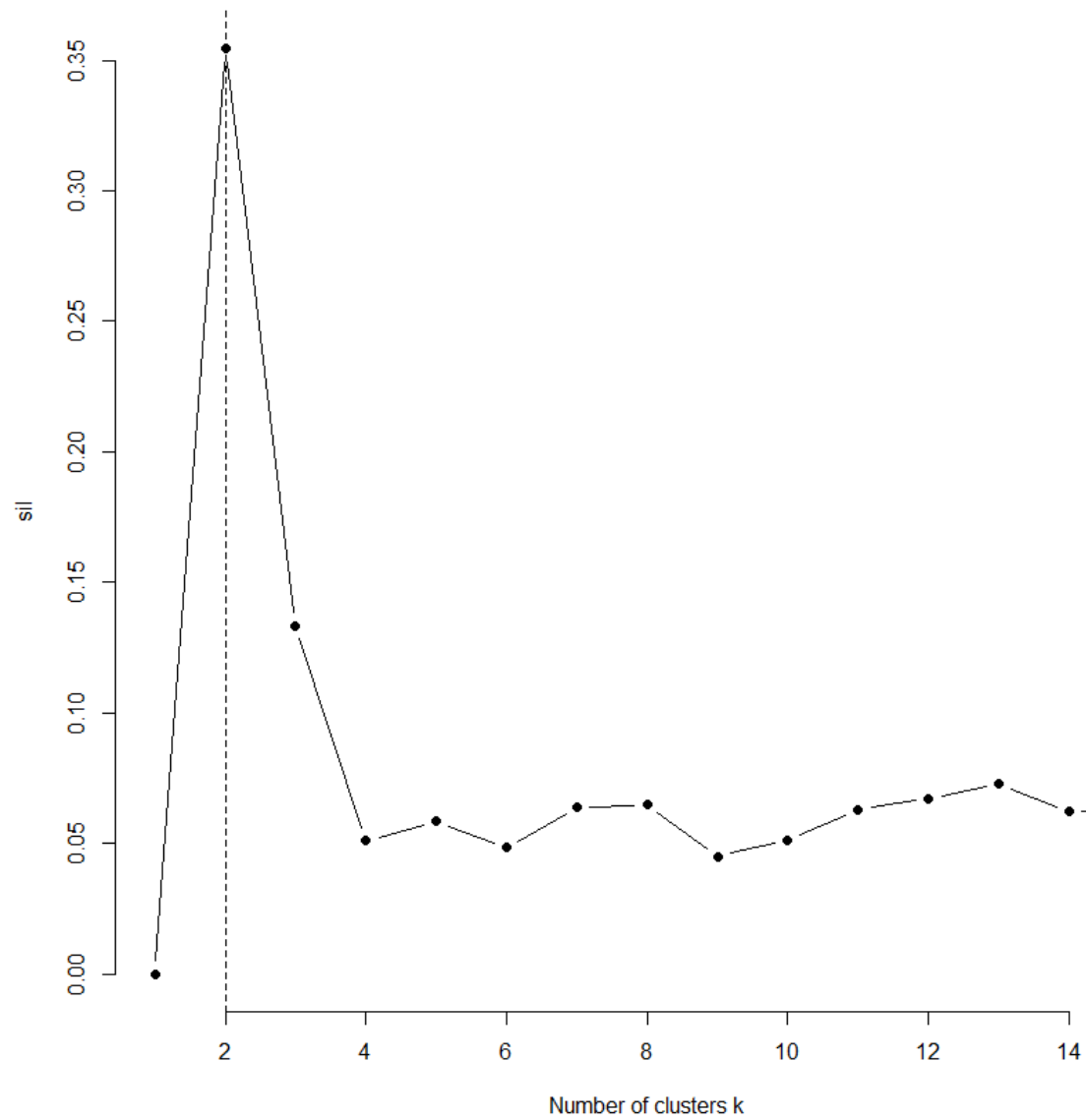
#data represents the csv file containing processed word data of taxonomic entities
#numClusters represents the number of clusters as a parameter to the kmeans function
Km.results <- kmeans(data, numClusters)
#PCA and visualization of results
Fviz_cluster(km.results, data)

```

K-means clustering with 2 clusters of sizes 60, 20

131

A6. Average silhouette graph for $n = 78$



A7. Cluster n = 78 results

K-means clustering with 2 clusters of sizes 59, 19

```

Cluster means:
1. .health use mobile time data information provide system study develop result design through patient research care service base
1 0.00470622 0.004710167 0.003640239 0.002241619 0.003820249 0.002451828 0.002349250 0.004703608 0.003423598 0.003204565 0.001703627 0.001938307 0.001241840 0.004579638 0.001678272 0.002422929 0.002447065 0.001184990
2 0.01687000 0.005374293 0.009146859 0.001934301 0.005251747 0.003293980 0.002146771 0.006852806 0.003609061 0.001959280 0.001889607 0.001077733 0.001523711 0.003785664 0.001157697 0.005246876 0.003388546 0.001385439
work support report include number medical management application project available test level user clinical improve group first two
1 0.0013051986 0.001787395 0.001436212 0.0009384776 0.0009911423 0.002017068 0.001173304 0.001746209 0.002255547 0.0009406925 0.001122874 0.001214648 0.0024346092 0.0016089463 0.0006696441 0.001797768 0.0007514166 0.0008885199
2 0.0009017405 0.001840390 0.004280605 0.0011236696 0.0021713729 0.001359529 0.002273419 0.002184514 0.002767437 0.0009410914 0.001362606 0.001789429 0.0008937598 0.0009292484 0.0009477964 0.001764539 0.0007620297 0.0008794201
treat access monitor over implementation follow allow communication method new intervention program increase effective device table approach
1 0.001492177 0.0009831093 0.001130362 0.0006866525 0.001274124 0.0006699922 0.0007683709 0.001256636 0.0008924117 0.0010636912 0.002280107 0.001282074 0.0008626275 0.0006376896 0.0015944078 0.0006067186 0.0010315554
2 0.002366270 0.0016863579 0.001199892 0.0007514188 0.001179543 0.0007452013 0.0007069813 0.001683804 0.0004878424 0.0005517927 0.002284648 0.002086842 0.0007190843 0.0009770360 0.0009358097 0.0008389616 0.0006081587
able assess process phone high network model problem tool analysis cost measure mhealth north significant potential different
1 0.0005226285 0.0007887441 0.0008766707 0.0005952482 0.0005053928 0.001202098 0.0011143814 0.0008132176 0.0006130491 0.0006873243 0.001016301 0.0009164065 0.001289079 0.0005205707 0.0007684809 0.0004687111 0.0007608152
2 0.0003978845 0.0005507212 0.0002661633 0.0036379792 0.0009644704 0.002312763 0.0004272738 0.0002145373 0.0010011981 0.0011035541 0.001806068 0.0003414820 0.002683134 0.0018842856 0.0007346999 0.0013053752 0.0005664274
disease show well current conduct change help example important year many case outcome train quality provider record
1 0.0006342828 0.0006351706 0.0005873161 0.0006875724 0.0005911973 0.0010078292 0.0004412799 0.0006492430 0.0004388981 0.0004616994 0.0005692031 0.0005492979 0.0007682213 0.0006754242 0.0005070769 0.0004883938 0.0005689175
2 0.0020774310 0.0005209975 0.0003586374 0.0004418605 0.0006059422 0.0003136018 0.0006443161 0.0002980335 0.0007709231 0.0007915988 0.0005257256 0.0014486736 0.0009110721 0.0009037695 0.0009298663 0.0006944595 0.0003970548
participant delivery population community activity shown limit three require resource receive country message center paper unit possible
1 0.001512915 0.0006202365 0.0003960251 0.001047373 0.0006657136 0.0003419598 0.0002807342 0.0005872071 0.0001172585 0.0004464228 0.0004625573 0.0003833267 0.001262785 0.0004362631 0.0004923470 0.0005423105 0.0002915791
2 0.001266730 0.0004253129 0.0013074209 0.002316491 0.0003254997 0.0003189917 0.0006899006 0.0002581379 0.0001105042 0.0004752344 0.0005777841 0.0012866682 0.001656127 0.0007768365 0.0003548625 0.0005782105 0.0003950476
review individual make remote evaluation mean healthcare practice week university further people relat electronic component present team
1 0.0005466234 0.0004956813 0.0003060643 0.0004045823 0.0005471139 0.0005104226 0.0007317629 0.0004771880 0.0009026912 0.0004353695 0.0003136407 0.0005369881 0.0004680766 0.0004888465 0.0005420798 0.000375049 0.0006242705
2 0.0003166066 0.0004555070 0.0005009533 0.0009904426 0.0003056622 0.0001490316 0.0007099356 0.0005086937 0.0006421856 0.0002475267 0.0003723405 0.0006048614 0.0000000000 0.0002758978 0.0001292216 0.0004685075 0.0005132095
range consider day impact control primary part found focus several address collect set worker form collection select
1 0.0003490906 0.0002937465 0.0005785779 0.0003424683 0.000532867 0.0004194665 0.0003890734 0.0003443858 0.0004716046 0.0003313845 0.0002668074 0.0003233366 0.0005350262 0.000315923 0.0002903341 0.0004246702 0.0003788041
2 0.0002146014 0.0003710692 0.0008876875 0.0003229319 0.001051002 0.0009720089 0.0001691659 0.0003926041 0.0001603689 0.0002045134 0.0003817829 0.0003438979 0.0001603328 0.002448278 0.0005587407 0.0006543477 0.0001964688
areas successful period screen local experience understand staff trial feedback computer offer technical question public text future
1 0.0003036166 0.0002507267 0.0003684722 4.720234e-04 0.0002616988 0.0003508825 0.0003700927 0.0008233726 0.0008556689 0.0007723357 0.0003834512 0.0002559677 0.0004162567 0.0004491069 0.0003131215 0.001190989 3.198165e-04
2 0.0010055909 0.0004116513 0.0006739021 9.667253e-05 0.0008221878 0.0002120083 0.0002326016 0.0006638113 0.0003439307 0.0001988243 0.0001663080 0.0003561338 0.0002343772 0.0002574250 0.0010287364 0.001073184 3.470216e-05
clinic function call order pilot condition national software member positive visit hospital page did strategy protocol term
1 0.0005374698 0.0003418939 0.0002691914 0.0003743569 0.0003347924 0.0004424799 0.0001849686 0.0004184503 0.0002234480 0.0002706372 0.0002643992 0.0005928702 0.0003515210 0.0002830799 0.0003498768 3.892820e-04 0.0002400626
2 0.0029894061 0.0002717499 0.0006285166 0.0002306449 0.0009636963 0.0002741287 0.0014970731 0.0004746585 0.0005341102 0.0004480026 0.0028753806 0.0005844207 0.0002526111 0.0005381956 0.0003826383 5.162489e-05 0.0002329176
standard
1 0.0003325090
2 0.0002525005

Clustering vector:
[1] 1 1 1 2 2 1 1 1 1 2 2 1 1 1 1 1 2 1 1 1 1 1 2 2 2 1 1 1 1 1 1 1 1 2 2 2 1 2 1 1 1 1 1 1 1 1 1 1 2 2 1 1 1 1 1 2 1 1 1 1 1 1 1 2 1 1 1 1 2 2 1 2 1 1 1 1 1 1 1 1 1

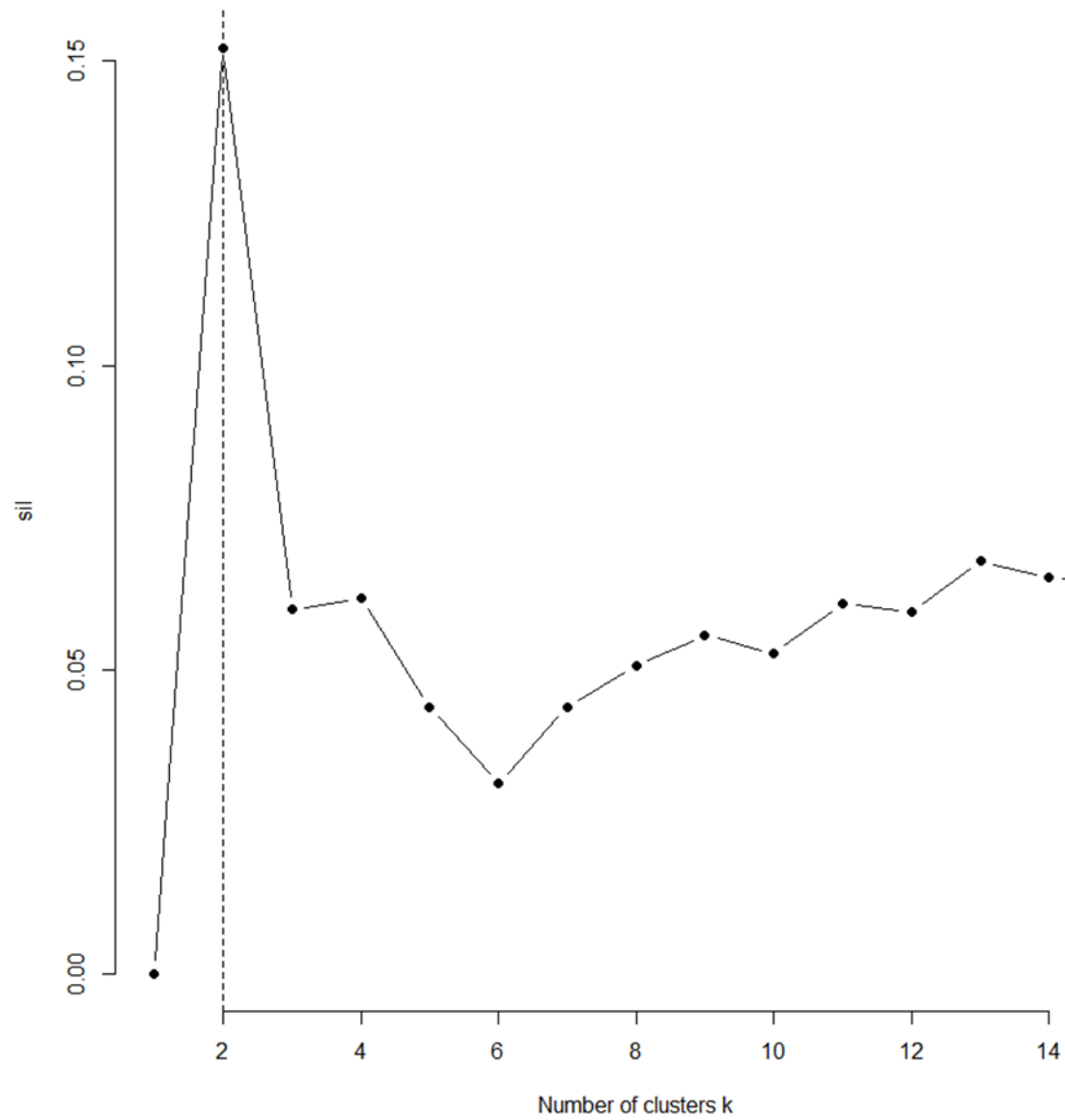
within cluster sum of squares by cluster:
[1] 0.02922690 0.01355924
(between_SS / total_SS = 8.3 %)

```

A8. Taxonomy keyword mapping (7th iteration)

Dimension	Attribute	Matching words
Primary Stakeholders	Patients	Patient, health, delivery
	Providers	Worker, clinic, record
Type of Study	Field Study	study, problem, case
	Randomized Control Trial	participant, trial, significant, mean, monitor, intervention, control
	Simulation	assess, university
Evaluation Setting	Developing Country	Country, local, national
	Developed Country	Country, national
Primary Evaluation Metric	Cost-based	Cost
	Technology - based	mobile, data, develop, information, application, device, network
	User-based	patient, health, care, user, disease, healthcare, project, communication, message, delivery
Duration of evaluation	0 to 6 months	Day, month
	6+ months	Year, period, change
(So what) Evaluation Conclusion	Positive	positive
	Negative	outcome, record, measure, analysis
	Neutral	feedback
Theory application	Exists	Theory
	Does not exist	

A9. Average silhouette graph for $n = 104$



A10. Cluster n = 104 results, 2 clusters, targeted words

K-means clustering with 2 clusters of sizes 30, 74

cluster means:

	1..Health	mobile	data	information	care	service	user	intervention	communication	network	problem	analysis	change	conduct
1	0.015311280	0.008032879	0.004283976	0.005037614	0.006236152	0.003238244	0.002683054	0.001661694	0.0017533806	0.001672355	0.0005989132	0.0005563690	0.0002584432	0.0004175639
2	0.004209207	0.003560869	0.003654623	0.002369793	0.001648241	0.002233828	0.002325187	0.003353689	0.000965346	0.001032544	0.0008411709	0.0008222794	0.0010538024	0.0007121418
	record	participant	delivery	message	country	week	university	control	local	period	clinic	positive	System	Study
1	0.0008463287	0.001113370	0.0007494562	0.001597821	0.001148920	0.0004885800	0.0004540654	0.0007370833	0.0005561235	0.0004622092	0.0019281027	0.0002158077	0.007821240	0.002771343
2	0.0003171742	0.002801082	0.0004021533	0.001729974	0.000331512	0.0009876489	0.0003421884	0.0007256866	0.0002668527	0.000421384	0.0004758912	0.0003821515	0.003516731	0.004362874
	develop	patient	application	project	monitor	device	assess	disease	significant	month	year	measure	cost	case
1	0.004158630	0.005945618	0.002928012	0.003227995	0.0008783456	0.001364334	0.0004882578	0.0019657998	0.0005941861	0.001202410	0.0006734982	0.0004101568	0.0014857055	0.0012247020
2	0.002080518	0.003667194	0.001398301	0.001402036	0.0011900085	0.001377889	0.0007719475	0.0004907335	0.0009220641	0.000932465	0.0005814527	0.0009569060	0.0008677965	0.0004073732
	outcome	healthcare	mean	remote	day	worker	trial	feedback	national					
1	0.0007505181	0.0008793896	0.0002503772	0.0008059398	0.0006758089	0.0020184844	0.0003176610	0.0003966891	0.0010488585					
2	0.0008414951	0.0006145968	0.0005418952	0.0002908195	0.0006079376	0.0001024451	0.0009083732	0.0007734077	0.0001695226					

clustering vector:

```
[1] 2 1 1 1 1 1 1 2 1 2 1 1 2 2 2 1 2 1 2 2 2 1 2 1 1 2 2 2 1 1 1 2 1 2 2 2 1 2 2 2 2 1 1 2 1 2 2 2 1 2 2 2 2 1 2 2 2 2 1 2 1 2 2 1 1 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2  
[90] 2 2 2 2 2 2 2 2 1 2 1 2 2 2
```

within cluster sum of squares by cluster:

```
[1] 0.01642693 0.02178099
```

(between_55 / total_55 = 11.4 %)

A11. Cluster n = 104 results, 3 clusters, targeted words

K-means clustering with 3 clusters of sizes 27, 3, 74

Cluster means:

	i..Health	mobile	data	information	care	service	user
1	0.014990415	0.008556354	0.004522519	0.005488745	0.006693856	0.003598049	0.002941833
2	0.013416056	0.008302887	0.004123015	0.003141317	0.000978994	0.002534212	0.000000000
3	0.004403113	0.003358925	0.003574113	0.002282069	0.001694369	0.002131090	0.002339540
	intervention	communication	network	problem	analysis	change	conduct
1	0.0008508734	0.0018708236	0.001802744	0.0005729448	0.0004629891	0.0002206343	0.0004639599
2	0.0233849937	0.0024180247	0.001681485	0.0030410543	0.0024104720	0.0022552373	0.0013086667
3	0.0027688543	0.0008959287	0.000984599	0.0007516401	0.0007811841	0.0014366424	0.0006590877
	record	participant	delivery	message	country	week	university
1	0.0007564239	0.001237077	0.0007933907	0.001343018	0.0008389195	0.0005428666	0.0004247532
2	0.0000000000	0.005885326	0.0010136847	0.017880829	0.0011205440	0.0014497797	0.0002899560
3	0.0003842879	0.002562488	0.0003754112	0.001162820	0.0004457706	0.0009288740	0.0003595365
	control	local	period	clinic	positive	System	Study
1	0.0004997508	0.0005785766	0.0004742274	0.0021029980	0.0002397864	0.008257535	0.002493985
2	0.0033022470	0.0015205270	0.0016894747	0.0000000000	0.0000000000	0.003739448	0.020893517
3	0.0007082878	0.0002195629	0.0003679993	0.0004902443	0.0003821515	0.003523021	0.003729390
	develop	patient	application	project	monitor	device	assess
1	0.003945425	0.006404536	0.0030406429	0.003361207	0.0008164118	0.001338904	0.0002970479
2	0.0000000000	0.004476425	0.0004639293	0.0000000000	0.0008447373	0.0000000000	0.0016518200
3	0.002326902	0.003559313	0.0014571016	0.001484296	0.0012139685	0.001442479	0.0007945418
	disease	significant	month	year	measure	cost	case
1	0.0018928416	0.0006208684	0.0009774424	0.0006558170	0.0004025538	0.0015310943	0.0012485966
2	0.0004059383	0.0021456740	0.0053168607	0.0011598237	0.0000000000	0.0022880980	0.0000000000
3	0.0005805911	0.0008494305	0.0008477458	0.0005681881	0.0009763080	0.0008187064	0.0004483049
	outcome	healthcare	mean	remote	day	worker	trial
1	0.0007075505	0.0009239236	0.0002781969	0.0008561504	0.0005502803	0.0012380713	0.0001866724
2	0.0027184700	0.0000000000	0.0002899560	0.0000000000	0.0031248450	0.0055871497	0.0036070493
3	0.0007773906	0.0006339988	0.0005301402	0.0003051726	0.0005544533	0.0002425148	0.0008228128
	feedback	national					
1	0.0004407657	0.0010434566					
2	0.0000000000	0.0013294403					
3	0.0007734077	0.0001601187					

Clustering vector:

```
[1] 3 3 1 1 1 1 1 3 1 3 1 1 3 3 3 1 3 1 3 3 3 1 3 1 3 3 3 1 3 3 3 1 2 1 3 1 3 3 3 1 3 3 3
[46] 3 3 3 1 1 3 3 3 2 3 1 3 3 3 3 3 1 3 1 3 3 1 1 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
[91] 3 3 3 3 3 3 3 3 1 3 1 3 3 3
```

within cluster sum of squares by cluster:

```
[1] 0.014180103 0.002160401 0.018546089
(between_SS / total_SS = 19.1 %)
```

A12. Method survey

The following 11 statements pertain to information systems research projects and their evaluations. On a scale of 1 to 5, 1 being lowest and 5 being highest, please indicate how much you agree with each of the following statements.

1) Technologies designed for healthcare patients and technologies designed for healthcare providers are fundamentally different

- 1 - Strongly Disagree
- 2 - Disagree
- 3 - Neither agree nor disagree
- 4 - Agree
- 5 - Strongly Agree

2) An underlying theoretical component, such as the technology acceptance model (TAM), is important for evaluating information technology.

- 1 - Strongly Disagree
- 2 - Disagree
- 3 - Neither agree nor disagree
- 4 - Agree
- 5 - Strongly Agree

The next four statements refer to user metrics, technical metrics, and cost metrics. The following link contains a table with examples of each: [metricExamples](#)

User Metrics	Cost Metrics	Technical Metrics
Mortality rate Hospital readmission rate Satisfaction	Return on investment Technology costs Personnel costs	System downtime Network stability Information transfer failure rate

Table 1. Examples of three different metric types

3) User metrics are more important than technical metrics when evaluating healthcare technologies

- 1 - Strongly Disagree
- 2 - Disagree
- 3 - Neither agree nor disagree
- 4 - Agree
- 5 - Strongly Agree

4) User metrics are more important than cost metrics when evaluating healthcare technologies

- 1 - Strongly Disagree
- 2 - Disagree

- 3 - Neither agree nor disagree
- 4 - Agree
- 5 - Strongly Agree

5) Cost metrics are more important than technical metrics when evaluating healthcare technologies

- 1 - Strongly Disagree
- 2 - Disagree
- 3 - Neither agree nor disagree
- 4 - Agree
- 5 - Strongly Agree

6) On average, studying user metrics takes longer than studying either cost metrics or technical metrics

- 1 - Strongly Disagree
- 2 - Disagree
- 3 - Neither agree nor disagree
- 4 - Agree
- 5 - Strongly Agree

The next five statements refer to the human development index (HDI) of countries. The following link contains a map of global HDI reports: [2016un hdi](https://data.unhcr.org/en/situations/human-development-index/hdi)

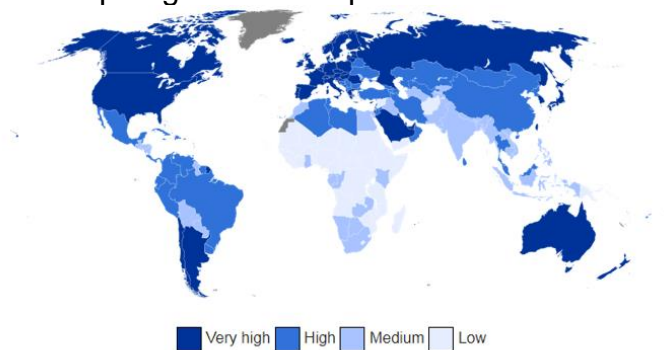


Figure 1. 2016 UN Human Development Index Report (Source: Wikimedia Commons)

7) Countries with higher HDI tend to have more complex networking infrastructure

- 1 - Strongly Disagree
- 2 - Disagree
- 3 - Neither agree nor disagree
- 4 - Agree
- 5 - Strongly Agree

8) On average, the same healthcare implementation would be costlier to implement in a higher HDI country

- 1 - Strongly Disagree
- 2 - Disagree
- 3 - Neither agree nor disagree
- 4 - Agree
- 5 - Strongly Agree

9) Randomized Control Trials (RCT) are easier to conduct in a higher HDI country

- 1 - Strongly Disagree
- 2 - Disagree
- 3 - Neither agree nor disagree
- 4 - Agree
- 5 - Strongly Agree

10) Field studies are easier to conduct in a higher HDI country

- 1 - Strongly Disagree
- 2 - Disagree
- 3 - Neither agree nor disagree
- 4 - Agree
- 5 - Strongly Agree

11) Simulations are easier to conduct in a higher HDI country

- 1 - Strongly Disagree
- 2 - Disagree
- 3 - Neither agree nor disagree
- 4 - Agree
- 5 - Strongly Agree

You are leading a research study to develop and evaluate a new mobile health implementation in a country with high/low* HDI. The sponsor of your research wants you to answer a few questions before releasing funding. Link to global HDI map: [2016un hdi](https://data.un.org/en/data/indicators/ny.hdi)

*Respondents were randomly shown either high or low

12) What methodology will you use to conduct your research? Assume you have unlimited funding to pursue any research question(s) you may have.

13) Which of the following three techniques would you use to evaluate your research?

Field Study
Randomized control trial

Simulation

14) Given your research setting, will you focus on user-based metrics, technology-based

metrics, or cost-based metrics?

User-based metrics

Technology-based metrics

Cost-based metrics

15) Related to the previous question, what are the main metrics you plan to use while evaluating the efficacy of your research?

16) How long do you expect the research process to take?

17) How long do you expect the evaluation process to take?