5-9-2016

Essays on Health Information Technology: Insights from Analyses of Big Datasets

Langtao Chen
Georgia State University

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.

Langtao Chen

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:

Langtao Chen
Department of Computer Information Systems
J Mack Robinson College of Business
Georgia State University

The directors of this dissertation are:

Dr. Detmar W. Straub
Dr. Aaron M. Baird
J. Mack Robinson College of Business
Georgia State University
35 Broad Street
Atlanta, GA 30303

ESSAYS ON HEALTH INFORMATION TECHNOLOGY: INSIGHTS FROM ANALYSES OF BIG DATASETS

BY

Langtao Chen

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
ROBINSON COLLEGE OF BUSINESS
2016

ACCEPTANCE

This dissertation was prepared under the direction of the *Langtao Chen'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. Detmar W. Straub

Dr. Aaron M. Baird

Dr. Lisa Schurer Lambert

Dr. Jeong-ha (Cath) Oh

ACKNOWLEDGEMENTS

I would like to express my appreciation and thanks to many people whose help makes this dissertation possible.

I owe many thanks to my advisor Dr. Detmar Straub for encouraging my research and guiding me to grow as a scholar. I'm grateful to have been working with him. His advice has been and will be priceless to my career. Thanks Detmar for everything! Thanks to my advisor Dr. Aaron Baird for all the guidance and encouragement. Working with him is a treasure to me. His patience, kindness, and hardworking style influence me. I also own many thanks to my wonderful committee members, Dr. Lisa Lambert and Dr. Jeong-ha (Cath) Oh, for challenging and guiding me through the dissertation process. Their valuable comments and advice will continue to benefit my future career.

Thanks to the entire faculty and staff at the Department of Computer Information Systems and the Center for Health Information Technology for making my Ph.D. study there a wonderful experience. I thank Drs. Arun Rai, Bala Ramesh, Dan Robey, Lars Mathiassen, Mark Keil, Michael Gallivan, Veda Storey, and Vijay K. Vaishnavi for teaching me how to do research. I also want to thank my fellow doctoral students, especially Alan Yang, Arun Aryal, Jessica Pye, Liwei Chen, Neetu Singh, Sayed Almohri, Shane Givens, Vitali Mindel, Youyou Tao, and Zhitao Yin. I always learn from them.

I would like to express my special thanks to my beloved wife Phoenix and my mother-in-law for all of the sacrifices they have made on my behalf. I also thank my baby son Jinlin, a gift to my life. Thanks to my parents, my grandma, my younger sister, and my younger brother for their unconditional support and love.

ABSTRACT

ESSAYS ON HEALTH INFORMATION TECHNOLOGY: INSIGHTS FROM ANALYSES OF BIG DATASETS

BY

Langtao Chen

4/5/2016

Committee Chair: *Dr. Detmar W. Straub, Dr. Aaron M. Baird*

Major Academic Unit: *Computer Information Systems*

The current dissertation provides an examination of health information technology (HIT) by analyzing big datasets. It contains two separate essays focused on: (1) the evolving intellectual structure of the healthcare informatics (HI) and healthcare IT (HIT) scholarly communities, and (2) the impact of social support exchange embedded in social interactions on health promotion outcomes associated with online health community use. Overall, this dissertation extends current theories by applying a unique combination of methods (natural language processing, machine learning, social network analysis, and structural equation modeling etc.) to the analyses of primary datasets.

The goal of the first study is to obtain a full understanding of the underlying dynamics of the intellectual structures of HI and its sub-discipline HIT. Using multiple statistical methods including citation and co-citation analysis, social network analysis (SNA), and latent semantic analysis (LSA), this essay shows how HIT research has emerged in IS journals and distinguished itself from the larger HI context. The research themes, intellectual leadership, cohesion of these themes and networks of researchers, and journal presence revealed in our longitudinal intellectual structure analyses foretell how, in particular, these HI and HIT fields have evolved to date and also how they could evolve in the future. Our findings identify which research streams are central (versus peripheral) and which are cohesive (as opposed to disparate). Suggestions for vibrant areas of future research emerge from our analysis.

The second part of the dissertation focuses on comprehensively understanding the effect of social support exchange in online health communities on individual members' health promotion outcomes. This study examines the effectiveness of online consumer-to-consumer social support exchange on health promotion outcomes via analyses of big health data. Based on previous research, we propose a conceptual framework which integrates social capital theory and social support theory in the context of online health communities and test it through a quantitative field study and multiple analyses of a big online health community dataset. Specifically, natural language processing and machine learning techniques are utilized to automate content analysis of digital trace data. This research not only extends current theories of social support exchange in online health communities, but also sheds light on the design and management of such communities.

Table of Contents

List of Tables

List of Figures

CHAPTER 1

INTRODUCTION

1.1 Motivation

Understanding the intellectual structure of health informatics is crucial to the whole health informatics community. In general, the intellectual structure of a discipline bespeaks the topics and paradigms selected by a field, the research themes that emerge over time, the thought leaders who direct the efforts of its various research programs, and the relationships between various structural components. Gaining deep insights into the intellectual structure of a discipline can lead to defining moments for a community of scholars (Kuhn 1962). Whereas this structure often reifies what is already known in the knowledge base or else increments (Kuhn 1962), it can also shape the epistemologies that frame knowledge development work and alter the philosophical basis of these efforts (Crane 1972). Structural knowledge can help scholars set their future research directions by seeing patterns of work that have existed in the past and noting trend lines into the future (Platt 1964).

Although in-depth intellectual structure analyses have been conducted for the entire field of information systems (IS) in journals such as *MIS Quarterly* and *Management Science* (Culnan 1986; Culnan 1987), IS research intellectual structure analyses are notably lacking in the growing discipline of health informatics (HI) and its sub-discipline health information technology (HIT). Given that HI literature reviews and citation analyses have been conducted in HI journals and the HIT literature has been reviewed in information systems (IS) journals (Chiasson and Davidson 2004; Gallivan and Tao 2014; Raghupathi and Nerur 2010; Romanow et al. 2012), such articles are either becoming dated (especially in the case of many HI analyses) and/or use only one primary method (e.g., citation analysis, social network analysis, or latent semantic analysis). We contend that future progress is dependent on: (1) a more complete understanding of how the HI and HIT disciplines have grown and evolved in the context of IS research over the past two decades, (2) multi-method analyses of the structural relationships between and cohesion of research themes and thought leaders (we use citation and co-citation analysis, social network analysis, and latent semantic analysis), and (3) leveraging these intellectual structure analyses to guide future research.

The first essay of the current dissertation represents such an effort of more recent, more complete, and more thorough analyses of HI and, particularly, HIT intellectual structures. Deeper understanding of the evolving intellectual structures of HI and HIT provides a means by which to further expand, consolidate, and renew the discipline in a systemic and informed manner while also theoretically contributing back to coordinate and reference disciplines. Given that an in-depth intellectual structural analysis of HIT focused on research in top IS journals had not appeared before our study, we fill an important research gap in this essay. Using the multiple statistical methods including citation and co-citation analysis, social network analysis (SNA), and latent semantic analysis (LSA), we show how HIT research has emerged in IS journals and distinguished itself from the larger HI context.

The second essay of the current dissertation zooms in one specific emerging HIT research theme, online health communities, which are defined as social networks where people with common health interests can share experiences, request questions, seek or provide emotional support (Eysenbach et al. 2004). A 2011 national survey conducted in the U.S. by the Pew Research Center's Internet & American Life Project found that 80% of U.S. Internet users have looked for health information online, 34% of Internet users have read others' commentary or experience about health issues online, and 18% have sought online to find others with similar health concerns (Fox 2011). A more recent national survey by the same project found that 72% of U.S. Internet users have looked online for health information within the past year (Fox and Duggan 2013). Another survey showed that social media sites are emerging as a potential source of online health information, with 42% Internet users consulting online rankings or reviews and 32% using social networking sites for health (Thackeray et al. 2013). These statistics suggest that online health communities, or the Internet in general, are becoming a common source for health information seeking. As an inseparable part of the personalized preventative medicine (Swan 2012), online health communities are changing the way patients treat and/or manage their health.

Two major purposes of participants joining online health communities are to seek health information regarding self-management options and to receive emotional support by knowing that their peers care (Hajli et al. 2014). People can discuss conditions, symptoms, and treatments as well as seek and provide health-related advice and emotional support from each other. Moreover, advanced services such as posing questions to physicians, quantified self-tracking of health conditions, and clinical trials access can also be provided to consumers (Swan 2009). When individuals are sharing their personal health information with other online community peers, they are "crowdsourcing" the collective wisdom of a huge number of community members (Eysenbach 2008). This can significantly lower the cost of health care and alleviate burdens on the health care system. Ultimately, online health communities open up new opportunities for the health care industry to obtain the "triple aim" (Berwick et al. 2008, p. 760) including: (1) cutting costs, (2) enhancing the individual's experience of care, and (3) improving the health of populations. The wide use of online health communities leads naturally to the need to better understand the social relations in this context.

The rise of health social networks such as PatientsLikeMe, DailyStrength, and MedHelp provides unique opportunities for research focusing on healthcare decision support and patient empowerment (Miller 2012). User-generated content on these online communities is accessible not only to the patients and caregivers but also researchers. Specifically, digital trace data on the online communities are available for scholars to better address more complicated research questions proposed. Digital trace data are records of activities that are undertaken through an online information systems (Howison et al. 2011). Here, a trace represents an event occurred in the past. Following proper and rigorous ways, digital trace data can be used to measure theoretically interesting constructs (Howison et al. 2011). With the abundant big digital trace data being generated by online health communities, scholars are able to obtain insights into highly detailed, contextualized, and rich contexts, thereby obtaining insights that address the heterogeneous needs of individual patients. However, there is a lack of research in IS field that empirically addresses this phenomenon and its underlying theoretical relationships via analyses of big health data.

The second essay of the dissertations tends to fill such knowledge gap by probing the impact of social support provided and consumed in online health communities on individual health promotion outcomes through the analyses of big online health digital trace data. Contributions of this research not only extend current understanding of micro-mechanisms of social support exchange in online health communities as well as the catalytic role of social support in health promoting, but also shed light on the design and management of such online health communities.

1.2 Scope of Inquiry

This dissertation follows the multi-paper model and is comprised of two separate essays that respectively investigate: (1) the intellectual structure of the discipline health informatics (HI) and its sub-discipline health information technology (HIT), and (2) an emerging and interesting area of HIT research that explores the impact of social support on health promotion outcomes in online health communities. Table 1.1 summarizes the key characteristics of the two essays.

Table 1.1 Summary of Two Essays

Research Design	Essay 1: Intellectual Structure of Health Informatics	Essay 2: Online Health Communities
Research Topic	Intellectual structure of health informatics discipline	The effect of social support on health promotion outcomes
Data Source	Archival data	Digital trace data
Raw Data Volume	• 24,897 health informatics papers • 324 health information technology articles	• 2,305,288 online discussion posts • 238,617 threads • 32,405 members
Analytical Method	• Citation analysis • Co-citation analysis • Social network analysis (SNA) • Latent semantic analysis (LSA) • Cluster analysis	• Natural language processing (NLP) • Latent Dirichlet allocation (LDA) • Support vector machine (SVM) • Unified medical language system (UMLS) • Social network analysis (SNA)

CHAPTER 2

THE EVOLVING INTELLECTUAL STRUCTURE OF THE HEALTH INFORMATICS DISCIPLINE: A MULTI-METHOD INVESTIGATION OF A RAPIDLY-GROWING SCIENTIFIC FIELD[1]

Abstract

Scientific disciplines are self-defined and self-evolving to a large extent, but acknowledging that disciplines develop organically does not diminish the continuing need to more fully understand the underlying dynamics of their intellectual structures. Intellectual structures bespeak the topics (including paradigms) that a discipline selects, the sub-disciplines and sub-communities that emerge, the thought leaders who direct the efforts of its various research programs, and the relationships between these various structural components. One such discipline, the discipline of health informatics (HI), is not only a vitally important discipline for societies worldwide, but is also an enormous field that manifests itself in the natural and social sciences as well as in the information systems (IS) and applied disciplines including professionals such as physicians, nurses, paramedics, and so forth.

A subset of the HI field especially important to IS scholars is identified here as health information technology (HIT). The current study analyzes the intellectual underpinnings of the field of HI and, in particular, focuses on its sub-discipline HIT. Using the multiple statistical methods including citation and co-citation analysis, social network analysis (SNA), and latent semantic analysis (LSA), we show how HIT research has emerged in IS journals and

[1] Chen, L., Baird, A., and Straub, D. 2015. "The Evolving Intellectual Structure of the Health Informatics Discipline: A Multi-Method Investigation of a Rapidly-Growing Scientific Field," *Working Paper, Georgia State University*.

distinguished itself from the larger HI context. The research themes, intellectual leadership, cohesion of these themes and networks of researchers, and journal presence revealed in our longitudinal intellectual structure analyses foretell how, in particular, these HI and HIT fields have evolved to date and also how they could evolve in the future. Our findings identify which research streams are central (versus peripheral) and which are cohesive (as opposed to disparate). Suggestions for vibrant areas of future research emerge from our analyses.

Keywords: health informatics (HI); health information technology (HIT); intellectual structure; social network analysis (SNA); citation analysis; co-citation analysis; latent semantic analysis (LSA)

2.1 Introduction

A discipline or field of study is a community of scholars and teachers who develop expertise in a self-defined domain of knowledge (Abbott 1988). A discipline is distinguished, in part, by the power that this group exercises over expert matter, the more abstract term for such a community being a "profession" (Abbott 1988). Combining the terms leads us to the concept of an academic professional discipline which lays claim to knowledge in particular intellectual domains. Intellectual knowledge within domains grows and evolves over time, often in an organic manner, as geographically and temporally dispersed research is conducted by researchers who may or may not be familiar with the published, forthcoming, and/or ongoing works of others. Therefore, an "intellectual structure" underlying a discipline develops over time, as research topics, themes, and thought leaders emerge (and cohere and/or fragment), but the underlying structure between these elements is often difficult to identify without comprehensive analyses.

While in-depth intellectual structure analyses have been conducted for the entire field of information systems (IS) in journals such as *MIS Quarterly* and *Management Science* (Culnan 1986; Culnan 1987), IS research intellectual structure analyses are notably lacking in the growing discipline of health informatics (HI) and its sub-discipline health information technology (HIT). Granted, HI literature reviews and citation analyses have been conducted in HI journals and the HIT literature has been reviewed in IS journals (see Table 2.1 for a summary), but such articles are either becoming dated (especially in the case of many HI analyses) and/or use only one primary method (e.g., citation analysis, social network analysis, or latent semantic analysis). We contend that future progress is dependent on: (1) a more complete understanding of how the HI and HIT disciplines have grown and evolved in the context of IS research over the past two decades (our data span January 1992 to April of 2013), (2) multi-method analyses of the structural relationships between and cohesion of research themes and thought leaders (we use citation and co-citation analysis, social network analysis, and latent semantic analysis), and (3) leveraging these intellectual structure analyses to guide future research. Therefore, we contribute a more recent, more complete, and more thorough analysis of HI and, particularly, HIT intellectual structures.

The intellectual structure of a discipline bespeaks the topics (including paradigms) selected by a field, the themes that emerge, the thought leaders who direct the efforts of its various research programs, and the relationships between various structural components. Gaining deep insights into the intellectual structure of a discipline can lead to defining moments for a community of scholars (Kuhn 1962). Whereas this structure often reifies what is already known in the knowledge base or else increments (Kuhn 1962), it can also shape the epistemologies that frame knowledge development work and alter the philosophical basis of these efforts (Crane 1972). Structural knowledge can help scholars set their future research directions by seeing patterns of work that have existed in the past and noting trend lines into the future (Platt 1964).

Many authors see intellectual structures as a critical aspect of the history of a field, specifically, in this case, an intellectual history (Abbott 1999; Grafton 2006). Understanding the intellectual development of a discipline is of great importance for researchers in that it allows them to more effectively conduct studies based on prior research (Culnan 1986; Platt 1964). It can also aid in identifying gaps in the literature and subsequently forging research projects or programs that address these gaps (Platt 1964).

Studies of intellectual structures are likely a sub-dimension of a larger set of studies of how professional disciplines evolve. Some might even frame this as the sociology of a scientific discipline since intellectual structure studies examine how groups establish their identity and the social activities through which they establish their legitimacy (DeSanctis 2003). When they focus on knowledge creation and dissemination, they ask and answer questions about the "who" and "why" of the main research themes of the discipline. But, they can be broader in their vision, such as the current IS history initiative taken on by the Association for Information Systems (AIS) professional society, i.e., to create a record of historical artifacts about the discipline and how it has developed (see, for example, Abbott 1999). Intellectual scholarly activities are an important part of this overall story, but they are not the entire substance. The goal in the case of IS is, as articulated by Hirschheim et al. (2012):

> *We believe that a study of the history of the IS discipline can foster understanding of where the discipline of IS has come from, what has happened in the discipline, and how the discipline has evolved to the position it is in today (page ii).*

Clearly, the choice of discipline that is the focus of a structural study can be of equally great pertinence. Most people would place a premium on the history of nuclear physics over the history of basket-weaving even though the latter likely says a lot about changing cultural values and economics. For this reason, we are focusing the current study on the *information systems* that are heavily impacting *health* and *healthcare* in contemporary societies.

One hardly needs to argue for the criticality of healthcare (and thus healthcare studies) today. Healthcare budgets are soaring worldwide (Moses et al. 2013) and there appears to be no end in sight. Moreover, sizeable percentages of GDPs internationally are being absorbed by the delivery and consumption of healthcare products and services. Globally in 2013, healthcare was estimated at a rate in the range of 7-18% of GDP in nearly all developed economies, a rate that, in general, is climbing every year (Martin et al. 2014; OECD 2013). What is particularly disturbing about such trends is that even though the use of HIT seems to lead to better health outcomes (Garg et al. 2005; Jones et al. 2014) and may be able to lower the soaring costs of healthcare (Hillestad et al. 2005), HIT implementation barriers can be high (Jha et al. 2009).

Given the importance of this profession and discipline, and the need for a better understanding of the intellectual structures of HI and HIT in the context of IS research, we focus our efforts generally on the intellectual structure of HI and more specifically on the field of HIT. In the field of HIT in particular, research methods and citation trends have been reviewed (Chiasson and Davidson 2004; Gallivan and Tao 2014; Romanow et al. 2012), but comprehensive research on authorial and thematic leadership has not been fully addressed, leaving a research gap for both understanding the whole view of the HIT community and evaluating scholars and topics in this sub-discipline. Therefore, our main research questions are:

RQ1*: What is the intellectual structure of the entire field of HI?*

RQ2*: What is the emerging intellectual structure of the HIT sub-discipline?*

…including, in RQ2: (a) which HIT themes have been popular over time and what thematic shifts been observed over time; (b) which themes are the most prestigious, the most cohesive, and the most mature, both from the standpoint of content and networks of thought leaders; and (c) who are the intellectual leaders of the entire domain and the sub-domains?

The organization of this paper follows the standard format. First, we review the extant literature regarding intellectual structures and hone in on the HI and HIT literatures. This review will show the gaps in our current knowledge base about the intellectual leaders and the abiding topics in these fields. Our sampling and multi-methodological techniques, which include: (1) social network analysis (SNA), (2) variant forms of citation and co-citation analysis, and (3) latent semantic analysis (LSA), are then described, followed by data analysis. The paper concludes with observations about the state of the HIT field and areas that appear to be most fruitful for future work. Our multi-method approach to uncovering the nature of the HI discipline and its sub-discipline HIT yields vital information for academic research and theory development.

2.2 Distinctions, Definitions, and Background

2.2.1 Disciplinary Distinctions

What is a discipline? While we might wish to conceptualize a discipline such as HI as a well-defined, bounded body of knowledge, distinct from other disciplines, reality is of course much more complex. Abbott (2001) uses a fractal distinctions model of disciplinary development to show that the boundaries between academic disciplines are amorphous and ephemeral; this notwithstanding, many disciplines have an "axis of cohesion" (p. 144). When fields attempt to shift and up-scope their domain of interest, he argues that they inevitably move beyond their traditional boundaries and seek out interdisciplinary intellectual spaces. A novel interdisciplinary focus can share interests and paradigms from originating disciplines, but the point from which individual scholars start (i.e., their originating disciplines) dramatically affects how they ultimately position their interdisciplinary work. Rather than clarifying themselves through refinements, disciplines are continually fragmenting across thought and method. Equilibrium and stability are not possible because of fractation. Additionally, scientific disciplines are self-defined and self-evolving to a large extent, making full understanding of intra- and inter-disciplinary relationships a challenge. Therefore, there is a continuing need to more fully understand the underlying dynamics of their intellectual structures.

HI is one such discipline with complex structural properties, as it draws theoretical perspectives from many disciplines in the natural and social sciences as well as from IS. Given this interdisciplinary nature in which the discipline of health informatics has been approached and defined, we next show how HI is both distinct and related to research in the IS and health services sciences. We also show how the sub-discipline of HIT has emerged in the shared space between three more macro-level fields, namely: HI, health administration and management, and health services research (see Figure 2.1).

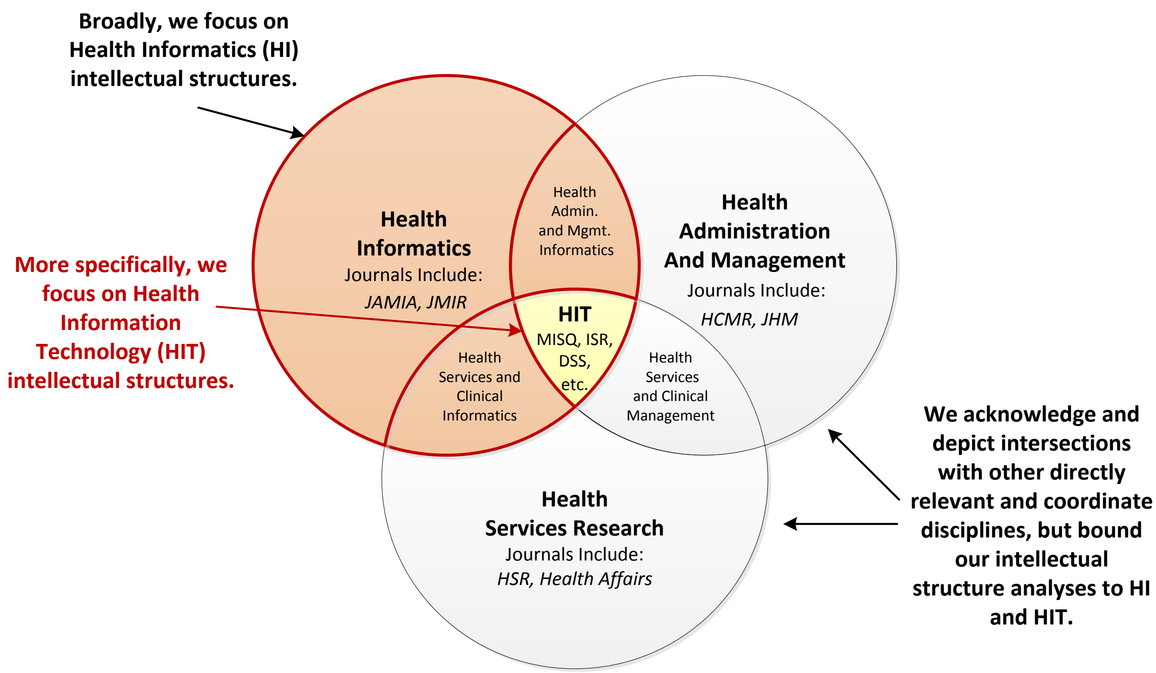


Figure 2.1 Distinguishing between HI, HIT, and Relevant Coordinate Disciplines

To differentiate HI-related research disciplines and sub-disciplines and to identify the centrality of the HIT sub-discipline for IS scholars, we utilize a preferred academic journal perspective. Journals are often used both to distinguish disciplines (Adler and Bartholomew 1992) and to identify overlap. For instance, Baumgartner and Pieters (2003) analyze the influence of journals in (and related to) marketing and demonstrate distinct differences and overlap in how these journals contribute to marketing sub-areas such as: core marketing, consumer behavior, managerial marketing, and marketing applications. In the HI domain, Morris and McCain (1998) demonstrate how clusters of citations in specific health informatics journals contribute to sub-areas such as: use of information for core medical informatics, medical decision making, and biomedical computing and engineering. While we focus most of our intellectual structure analyses on citations and journals specifically within the HIT field in this paper, we leverage favored journals to identify similarities and differences between research disciplines within the broadly considered field of HI.

Using this approach, we specifically identify three broad categories of health care journals (related to HIT research in management) that, at the intersection, are described by the journals concentrating on HIT research in the IS discipline (see Figure 2.1). These disciplines are: (1) *health informatics*, (2) *health administration and management*, and (3) *health services research*.[2]

HI is defined as: "The interdisciplinary study of the design, development, adoption and application of IT-based innovations in healthcare services delivery, management and planning" (Procter 2009).[3] HI includes applied clinical and public health informatics research. The broader field of HI, i.e., medical or health informatics, has been defined as a discipline that "draws on, and contributes to, multiple disciplines in the health sciences and information sciences" (Morris and McCain 1998, p. 448). Morris and McCain (1998) go on to note that

---

[2] We acknowledge that other domains, such as biology, also contribute to specific fields such as biomedical informatics. Based on our focus on HIT in the IS discipline, however, we focus our systematic analyses on the domains most relevant to researchers in business schools.
[3] More details on definitions and variations of definitions for HI and HIT are available in Appendix 2I.

"…while many definitions of the field can be found, most share two characteristics: reference to health sciences, biomedicine, and the healing arts; and reference to the use of information management techniques and technologies in support of those pursuits (p. 448)." The HI discipline includes journals such as the *Journal of the American Medical Informatics Association (JAMIA)* and the *Journal of Medical Internet Research (JMIR).*

Health administration is defined as "the decision making of program leaders and the supervision, controls, and other actions to ensure satisfactory performance and attain certain goals" (Roemer 1993). *Health management* is defined as "the profession that provides leadership and direction to organizations that deliver personal health services, and to divisions, departments, units, or services within those organizations" (Buchbinder and Shanks 2011, p.2). The field of health administration and management includes such journals as *Health Care Management Review (HCMR)* and the *Journal of Healthcare Management (JHM).*

Health services research is defined as "the multidisciplinary field of scientific investigation that studies how social factors, financing systems, organizational structures and processes, health technologies, and personal behaviors affect access to health care, the quality and cost of health care, and ultimately our health and well-being" (Lohr and Steinwachs 2002, p.15). This domain includes journals such as *Health Services Research (HSR)* and *Health Affairs.*

HIT is defined by the Office of the National Coordinator for HIT (ONC) as: "The application of information processing involving both computer hardware and software that deals with the storage, retrieval, sharing, and use of health care information, data, and knowledge for communication and decision making" (ONC 2014). We suggest that HIT research published in journals such as *MIS Quarterly (MISQ), Information Systems Research (ISR),* and *Management*

Science (MS), etc., sits at the intersection of the HI, health administration and management, and health services research domains.[4] As such, HIT research holds significant potential to contribute to the IS discipline as well as coordinate disciplines. We suggest that comprehensive analysis of the intellectual structures and research streams associated with HIT presents a unique opportunity to formalize our existing thinking in this important area of interdisciplinary research and provide a systematic foundation from which to build future HIT research in the IS domain.

2.2.2 Intellectual Structure of a Discipline

What is an intellectual structure and how does it apply to the analysis of a discipline? Intellectual structure bespeaks the topics that a field migrates to and selects,[5] the development of thematic sub-communities, the emergence of thought leaders who direct the efforts of its various research programs, and relationships between these components. Although the usage of the term "intellectual structure" may vary somewhat from one discipline or sub-discipline to another, it fundamentally has to do with the ideas that form the basis for impactful research. In this sense, an intellectual structure is a historical approach to knowledge creation and advancement in the sense that historians speak and write about the intellectual history of an era or a people.

"Intellectual" refers to ideas, but what does structure mean? While the concept of "structures" likely differs between the natural sciences and the social sciences as well as the arts and humanities,[6] under all circumstances, it would seem to be ways of thinking, old and new, that lie at the heart of a scholarly community of practice. Structure refers to the organization of

---

[4] As also mentioned in Appendix 2I, Health Information Systems (HIS) is likely a more appropriate term than HIT, as HIT indicates a focus on technology rather than a more comprehensive view of people, processes, technology, and information. However, the field most frequently uses the term "HIT" to refer both to the technology as well as to the more IS-comprehensive view. We adopt this more comprehensive view, but use the term HIT in accordance with the more frequent occurrence of this term.

[5] We take the term "topics" to be synonymous with the terms streams, themes, areas, or domains.

[6] In the natural sciences, for example, there appears to be greater stress on the value of linked research programs (Platt 1964).

the ideas themselves and also to relationships and distinctions between ideas among thematic sub-communities and contributors. The structure of a field depends not only on the ideas and knowledge being generated, but also on the thought leaders[7] who create networks of dependencies, most often revealed as patterns of citations and co-citations in studies. As these patterns develop and cohere and/or fragment, knowledge builds on knowledge and theories and paradigms compete until the community senses the need for a change and the paradigm shifts (Culnan 1987; Kuhn 1962).

Intellectual structure (and dynamics) emerges as a result of those who advance a discipline through thought leadership. Thought leadership is an important concept in the study of the intellectual structures of disciplines as well as innovations more generically (Rogers 1962). The central place of thought leaders in intellectual structures can be traced back to Crane's sociology of science work (1972) on invisible colleges. Building on de Solla Price's stress on the importance of citation networks (1963; 1965), Crane argues that scientists communicate their ideas through both formal and informal communication channels, which result in ideas that change over time. These form the so-called "invisible college" of a discipline. She also asserts that citation networks are a reasonable approximation of how these influences manifest themselves. Crane's views have been largely substantiated by Mulkay et al. (1975). Wagner (2008) has further updated the concept and has contextualized it within the Internet.

Why examine the intellectual structure of a discipline? The development and evolution of leaders, ideas, and concepts within and between disciplines provides a roadmap of the progression and current state of a scientific field and its relationships to coordinate disciplines

[7] In the diffusion of innovation literature (Rogers 1996), thought leaders are referred to as "opinion leaders" and they are deemed to be instrumental in the dissemination of new ideas.

(see Table 2.1 for examples). Examining past and current clusters of research activity also offers insights into which authors and ideas have become the most influential, what shifts have occurred over time, and which research streams are central (versus peripheral) or cohesive (as opposed to disparate). Knowledge gleaned from such analyses can be used to infer which research streams are still in their infancy, which research streams are mature and perhaps moving toward paradigmatic status, and which are ripe for disruption and revolution (Kuhn 1962). As our ultimate goal in research is to contribute to such theoretical understanding, it is vital to identify areas where future contributions can further extend our knowledge.

Table 2.1 Selected Works on Intellectual Structures of Various Disciplines (Ordered by Discipline)

Relevant Literature	**Research Domain**	**Research Method**	**Unit of Analysis**
Morris and McCain (1998)	Health Informatics	Citation analysis	Journal
Chiasson and Davidson (2004)	Health IT	Citation analysis	Author
Agarwal et al. (2010)	Health IT	Literature review	Unspecified
Romanow et al. (2012)	Health IT	Literature review	Article
Gallivan and Tao (2014)	Health IT	Co-citation analysis	Article
Raghupathi and Nerur (2008)	Health IT	Co-citation analysis	Author
Jones et al. (2014)	Health Services Research	Systematic review	Article
Culnan (1986), Culnan (1987)	IS	Co-citation analysis	Author

Polites and Watson (2009)	IS	Citation analysis & social network analysis	Journal
Sidorova et al. (2008)	IS	Latent semantic analysis	Article
Taylor et al. (2010)	IS	Co-citation analysis	Author
Li and Joshi (2012)	IS	Latent semantic analysis	Article
Euske et al. (2011)	Management Accounting	Citation analysis & social network analysis	Author
Baumgartner and Pieters (2003)	Marketing	Citation analysis	Journal
Pilkington and Meredith (2009)	Operations Management	Co-citation analysis	Author (and knowledge groups)
Nerur et al. (2008)	Strategic Management	Co-citation analysis & pathfinder analysis	Author

2.2.3 Intellectual Structure of HIT and Coordinate Disciplines

While the extant HIT literature provides a strong foundation from which to understand this growing sub-discipline, we suggest that little has yet to be done to: (1) compare and contrast HIT research with coordinate research in other disciplines; (2) comprehensively identify the intellectual structures of HIT research; and (3) highlight important HIT research streams (and shifts) within the IS discipline. Nor have the thought leaders of the discipline been exhaustively enumerated. We begin here by assessing the first point—how HIT research compares to coordinate research in other disciplines.

How have intellectual structures been previously analyzed in the HIT discipline? As can be seen in Table 2.1, literature reviews, systematic reviews (a term used by the medical community to indicate a rigorous literature search and review of a specific topic), and commentaries have been published, but analyses of HIT intellectual structure are wanting, especially from the IS scholar's point-of-view. Up to this point, systematic analyses of the HIT field have focused primarily on: how the healthcare context contributes to IS theory building and validation (e.g., Chiasson and Davidson 2004); reviews of research trends in the HIT literature (e.g., Romanow et al. 2012); and informed opinions regarding where the HIT discipline may be headed (e.g., Agarwal et al. 2010). The substantial quantity of empirical research work carried out on the impact of HIT on performance outcomes (such as cost, quality, and efficiency) has been systematically reviewed numerous times, typically drawing from the literature of many disciplines coordinate to HIT, including health management and health services research (e.g., Buntin et al. 2011; Jamal et al. 2009; Lau et al. 2010; Poissant et al. 2005; Wu et al. 2006). Findings related to the use of HITs [and "meaningful use" incentives in the USA (Blumenthal and Tavenner 2010)] have also been systematically reviewed. Such reviews typically synthesize the relevant literature from coordinate disciplines such as HI, health management, health services research, and health policy journals (e.g., Jones et al. 2014). Additionally, the growing body of HIT consumer acceptance work has also been systematically reviewed (Or and Karsh 2009).

What is glaringly missing is an analysis of the intellectual structure of the HIT literature. Granted, while intellectual structures have been assessed for the overall IS field (Culnan 1986; Culnan 1987; Pratt et al. 2012) and HI disciplines (as discussed in the next few paragraphs), these methods and analyses have yet to be rigorously applied to the HIT discipline.

In the HI or medical informatics discipline, several intellectual structure analyses have been conducted, with the bulk of this work focusing on intellectual structures emerging in the mid-1990s. Andrews (2003) assesses the relationships between authors and author influence using a co-citation analysis of medical informatics articles published between 1994 and 1998. Vishwanatham (1998) examines the most frequently cited journals in the medical informatics discipline between 1994 and 1996 using citation analysis. Morris and McCain (1998) conduct a co-citation analysis of medical informatics articles published between 1993 and 1995 and find that biomedical, decision support, and education were primary areas of focus. Eggers et al. (2005) use content maps and citations networks of medical informatics research published between 1994 and 1997 and find top and emerging content areas of that time to include: medical informatics, electronic medical records, information technology, decision support, medical students, protein sequencing, and neural networks. More recent analyses of medical informatics and HI intellectual structures have been conducted by Raghupathi and Nerur (2010) and Schuemie et al. (2009). Raghupathi and Nerur (2010) draw on HI and medical informatics literature published between 1998 and 2006 and, through an author co-citation analysis, demonstrate that distinct subfields are beginning to emerge including: artificial intelligence, user-interface design, and bioinformatics. Schuemie et al. (2009) conduct a similar analysis of the medical informatics literature published between 1993 and 2008, identifying three key clusters: (1) health information systems, (2) medical knowledge representation such as clinical guidelines and ontologies, and (3) data analysis and classification techniques and evaluation.

Whereas intellectual structure analyses in all of these coordinate disciplines are very informative, what is still needed is a comprehensive analysis of the HIT intellectual structures related to the IS discipline. Bounded by a set of core journals in the IS field, we next indicate the methods to be used in studying HIT intellectual structures.

2.3 Methods

How can the intellectual structure of a discipline be analyzed? And, what past approaches have been the most effective or informative? To better understand the intellectual structure of HI and its sub-discipline HIT, this paper employs as its major methods: (1) citation and co-citation analysis, (2) SNA, and (3) LSA. We also use other analytical tools, as appropriate. Table 2.2 shows the constructs being explored as well as the statistical toolsets employed.

Table 2.2 Constructs, Sub-Constructs, and Study Metrics

Constructs	**Sub-constructs**	**Definition**	**Measures Used**	**Analytical Method**
1. *Disciplinary structure*	—	Differentiation between disciplines by virtue of citation/co-citation patterns	Node in-degree; strength of tie	Citation and co-citation analyses; SNA
2. *Cohesion* (of HIT streams of research)	*Content cohesion*	The extent to which the semantics of a field or a sub-field cohere, that is, are common across article descriptors	Average intra-thematic sub-community factor loadings; changes in these average loadings over time	LSA; descriptive statistics
	Network cohesion or *maturity*	The extent to which a field or a sub-field is connected or integrated; intra-community citation cohesion	Network density	SNA; XY axes plot of maturity by prestige
	Prestige	The extent to which a field or a sub-field is cited by other fields or sub-fields	Node in-degree centrality and information centrality	SNA; XY axes plot of maturity by prestige

3. *Thought leadership*	*Overall HIT thought leaders*		Node in-degree; strength of tie	SNA; cluster analysis
	Sub-domain thought leaders		Raw citation counts by sub-theme	Citation analysis

Legend: SNA stands for Social Network Analysis; Node in-degree, strength-of-tie, and information centrality are centrality metrics in SNA; LSA is Latent Semantic Analysis.

2.3.1 Constructs and Measures

2.3.1.1 Disciplinary Structure

The relationship of disciplines to each other (and distinctions among them) is termed *disciplinary structure*. Table 2.2 indicates that this structure will be revealed by the citation pattern within and between disciplines, which, as noted earlier, are delimited by the journals that individual fields favor. We will examine this structure through both citation and co-citation patterns.

We further sub-divide the cohesion construct into two sub-constructs: *content cohesion* (related to semantic analysis of the usage of terms within articles) and *network cohesion* (related to citation patterns within and between articles). Research themes do not occur in a vacuum; they are created and nurtured by scholarly communities. Therefore we would argue that ideas are not separable from the people who create these ideas and tie their work to other individuals through publication citations. For this reason, we analyze intra-thematic citation patterns to uncover how tightly or loosely a community adopts the same linguistic terms in their work (i.e., article descriptors) and how tightly or loosely a community cites itself. In this way research themes also characterize the communities of scholars who study them. As Table 2.2 shows, the use of common semantics (i.e., common terminology) differentiates groups by means of our sub-construct *content cohesion,* while the sub-construct *network cohesion* relates to citation patterns.

The strength of connections within and between thematic communities can be described by the terms *maturity* and *prestige*, concepts which are further defined in Table 2.2. We will compare the HIT scholarly sub-communities on these constructs in order to posit which research sub-domains can more fully evolve.

2.3.1.2 Thought Leadership

Our third major construct is *thought leadership* (see Table 2.2). As noted earlier, groups of scientists form invisible colleges (Crane 1972) as they engage in their thematic pursuits. Both citation patterns and networks can portray which individuals lead these communities of practice (Crane 1972; de Solla Price 1963; de Solla Price 1965). We use these citation counts (in SNA these are known as in-degree or centrality measures) to determine which scholars are heading up the intellectual discourse in the overall network of HIT research. We also subdivide the HIT dataset into sub-communities and examine the HIT intellectual leadership through this lens.

2.3.2 Data Collection and Sampling Procedures

Regardless of analytical methods, the first issue in a scientometric, intellectual structure study such as this is to determine which data and which samples are to be used. Many structural studies focus on a highly limited set of representative journals (e.g., Euske et al. 2011; Ramos-Rodríguez and Ruíz-Navarro 2004). Our view is that this is too tenuous, given the interdisciplinary and emerging nature of HI research. Therefore, we used keywords to search bibliographic databases and did not limit our initial search to a predefined set of journals, with the purpose of investigating the entire spectrum of the HI, in general, and the HIT sub-discipline in particular. Since the foundation of the present study is both citation analysis and co-citation analysis, article information was retrieved from the Web of Science (formerly ISI Science Index and Social Science Index), which contains source article information and a comprehensive reference list (Bernroider et al. 2013), thus facilitating the citation and co-citation analyses.

Data collection followed terms used in previous systematic reviews (Higgins and Green 2008). Multiple healthcare-related keywords (such as "health-care," "healthcare," "health care," "health," "medical," "clinical," "hospital," "physician," "doctor," "patient," "nurse," and "medicine" etc.) were combined with IT-related keywords (such as "information technology," "information system," "computer" etc.) to retrieve articles potentially related to HI. Also, keywords such as "healthcare information technology," "healthcare information system," "health information technology," "health information system," "health informatics," "medical informatics," "healthcare IT," "health care IT," "health IT" etc. were directly used to retrieve relevant articles. Articles under Web of Science Category "Medical Informatics" were further checked and added into the dataset if they were not explicitly included in the search result. We limited our search to academic articles in English language. As a result, 62,249 papers formed the initial dataset, as shown in Figure 2.2.

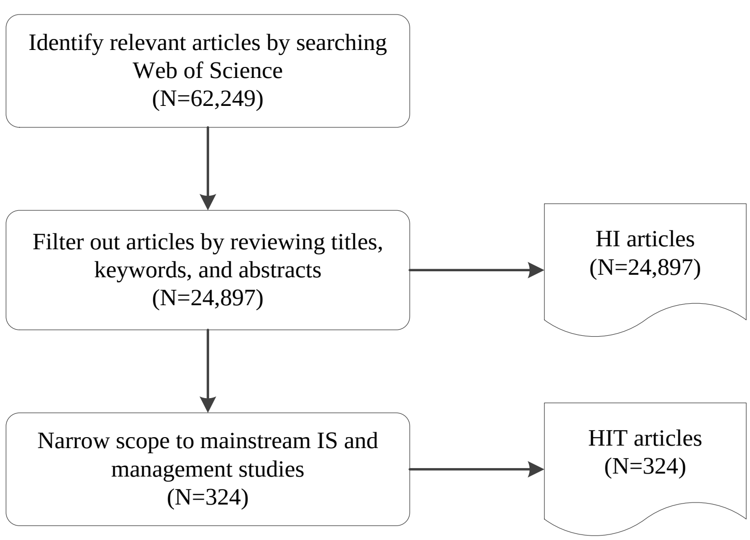


Figure 2.2 Sampling Frames and Filtering Procedures

To refine the dataset, we examined the title, keywords, and abstract of each paper in order to exclude articles that were included in the search result but not actually related to HI. By doing so, 24,897 HI papers published in an approximately 30-year period from 1983 to April 2013 qualified as the HI sampling frame. Most of the HI articles in the sampling frame were published in medical informatics journals. This dataset was used to explore the overall intellectual structure of HI research.

Finally, to uncover the intellectual structure of the sub-discipline of HIT, the sampling frame for HI research was narrowed to articles published in mainstream IS and management journals such as IS Senior Scholars' Basket of Journals,[8] Decision Support System, and Communications of the ACM (refer to Appendix 2B for a complete list of HIT journals). At this stage, 324 HIT articles were identified within the approximately 21-year period from 1992 to April 2013. Figure 2.2 shows the sampling frames and the filtering procedures employed. Summaries of exemplar HI and HIT publications are attached in Appendix 2A and Appendix 2B respectively.

2.3.3 Multi-Method Selection Procedure

Two major bibliometric techniques, citation and co-citation analyses, have been widely deployed to explore the intellectual structure of a variety of disciplines. These techniques form the foundation of our multi-method approach which, overall, includes (Figure 2.3): (1) data collection and sampling (described above), (2) creation of citation and co-citation matrices, (3)

---

[8] These eight journals include the following and are further described at http://aisnet.org/general/custom.asp?page=SeniorScholarBasket: *MIS Quarterly, Information Systems Research, Journal of MIS, Journal of AIS, European Journal of Information Systems, Information Systems Journal, Journal of Information Technology,* and *Journal of Strategic Information Sy*stems.

extraction of research themes via LSA, and (4) conducting SNA on the final matrices for the purposes of understanding networks of themes and thought leaders.

Figure 2.3 summarizes the overall design of this multi-method data analysis approach and the order in which the analyses were conducted for the investigation of HI and HIT intellectual structures. References exported from Web of Science contain bibliographic information which can be used to construct the citation relationship among articles. For each article, authors, year, journal, title, abstract, and all articles cited by it were imported into a database. Then a computer program parsed the bibliographic information to build article citation and co-citation matrices for the HI and HIT datasets, respectively. An LSA procedure was used to extract research themes from HIT article abstracts. Based on the article citation and co-citation matrices, citation and co-citation matrices at discipline, author, and HIT research theme levels were calculated. The detail of the multi-method data analysis approach is explained in the following.

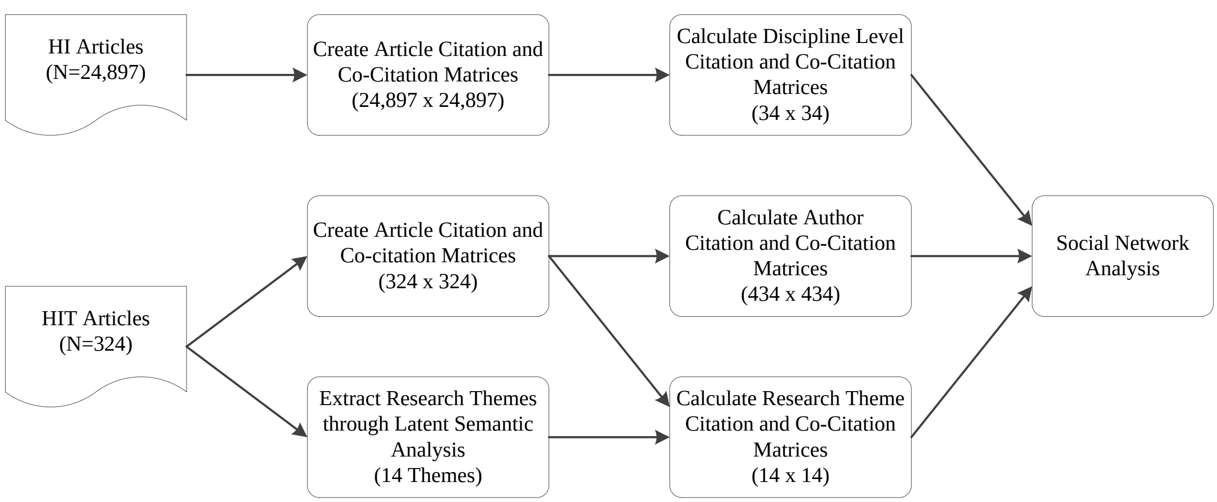


Figure 2.3 Flowchart of the Multi-Method Approach Utilized in This Investigation

Citation analysis is based on the assumption that the bibliographic references cited in a research paper are a valid indicator of their influence on the citing paper (Cole and Cole 1972; Ramos-Rodríguez and Ruíz-Navarro 2004). Thus, repeatedly cited references are thought to be more influential on the intellectual structure of a discipline than less frequently cited articles (Culnan 1986). A complementary perspective, co-citation analysis, takes the number of articles citing two particular documents to be a surrogate for the intellectual association between any two documents (Small 1973; White and Griffith 1981). Co-citation analysis is a powerful tool to identify clusters of authors, research themes, or paradigms. It particularly helps in understanding how such clusters interrelate (Pilkington and Liston-Heyes 1999).

To analyze the intellectual structure of the overall HI across multiple disciplines, we aggregated article-level citation and co-citation matrices to the discipline level based on the Web of Science categories of journals. Thomson-Reuters Journal Citation Report (JCR) contains information on influence, impact, and subject relationships for leading journals. Subject categories of each journal in our dataset were retrieved from both the JCR for the Social Science Citation Index (SSCI) 2012 and the JCR for the Science Citation Index (SCI) 2012 and treated as academic disciplines for the citation and co-citation analysis. In total, 34 disciplines were identified as publishing HI research. As a result, 34 x 34 matrices for discipline citation and co-citation relationships were created.[9]

For the dataset of the 324 HIT articles, two levels of analysis, including author and research theme, were addressed. Since the analysis of authors for HIT articles identified 434 HIT scholars, there were 434 x 434 resulting matrices for examining author citation and co-

[9] We created a 34 x 34 *citation* matrix and a 33 x 33 *co-citation* matrix, because one discipline (Biochemistry & Molecular Biology) does not co-cite with any other disciplines.

citation relationships. These were calculated from article level citation and co-citation relationships by checking the authors for each article. Next, to extract the research themes in the extant HIT literature for the purposes of creating theme level citation and co-citation matrices, we employed the same LSA procedure used by Sidorova et al. (2008) (please refer to Appendix 2C for details of the LSA procedure). Traditional literature reviews that are manually coded and analyzed by researchers are subject to two substantive limitations: (1) the huge amount of time and effort to analyze large datasets and (2) the researcher bias in coding and analyzing textual data (Larsen et al. 2008). LSA is a text mining technique that provides another way to unveil hidden concepts from textual data, thus discovering core research themes within whole bodies of literature (Sidorova et al. 2008). The underlying logic of LSA is that the aggregate of all the word contexts in which a given word does or does not appear provides a set of mutual constraints that largely determine the similarity of meaning of words and sets of words to each other (Landauer et al. 1998). HIT research theme level citation relationships were also calculated, with 14 x 14 citation and co-citation matrices being created. Appendix 2D shows the detailed procedure for constructing theses citation matrices at different levels.

We then used SNA to assess both the citation and co-citation patterns in the HI and HIT disciplines, as applied to the discipline-level (HI), author-level (HIT), and theme-level (HIT) citation and co-citation matrices developed through the procedures explained above. We selected SNA for its ability to make inferences about our key constructs as revealed in the citation and co-citation matrices. SNA can analyze network structures rather than patterns of individual (i.e., node) attributes. Thus, the results of SNA can complement general statistical methods which generally ignore network structures and topologies. Metrics in SNA such as centrality (e.g., degree centrality, closeness centrality, Bonacich power, and information centrality) are methodologically mature and hold the potential of analyzing a variety of citation and co-citation relations (Scott and Carrington 2011). SNA has been employed in prior studies to assess the relationships between inter-journal citation patterns in academic literatures. To rank IS journals, Polites and Watson (2009) rely on SNA's ability to disclose the underlying structure of the entire IS discipline. Euske et al. (2011) investigate the tribalism of management and accounting scholars by analyzing networks of literature citation. Benckendorff (2009) conduct network analysis to reveal themes and trends in tourism research in Australia and New Zealand. In this study, directed graphs unveiled the structure of citation relationships while co-citation relationships were represented by undirected graphs. In our case, the software package NetDraw (Borgatti 2002) was used to investigate citation and co-citation relationships.

2.4 Results

2.4.1 Disciplinary Structure of HI

A primary goal of this research is to investigate how HIT has emerged from the larger HI setting (RQ2). Thus, citation analysis and co-citation analysis first reveal where HIT fits in the larger HI context. The citation network of HI research disciplines is shown in Figure 2.4 where the size of each node is proportional to the in-degree of the node (that is, citations coming to a sub-discipline), with the thickness of the arrows and lines representing the relative strengths of the citation relationship between two nodes. Clearly, Medical Informatics dominates the HI intellectual structure as the central node. But, the major contributing sub-disciplines are Health Care Sciences & Services, General and Internal Medicine, Information Systems, and Computer Science, in that order. This suggests that IS and its closely related technical field, computer science, are key drivers of knowledge creation in this space.

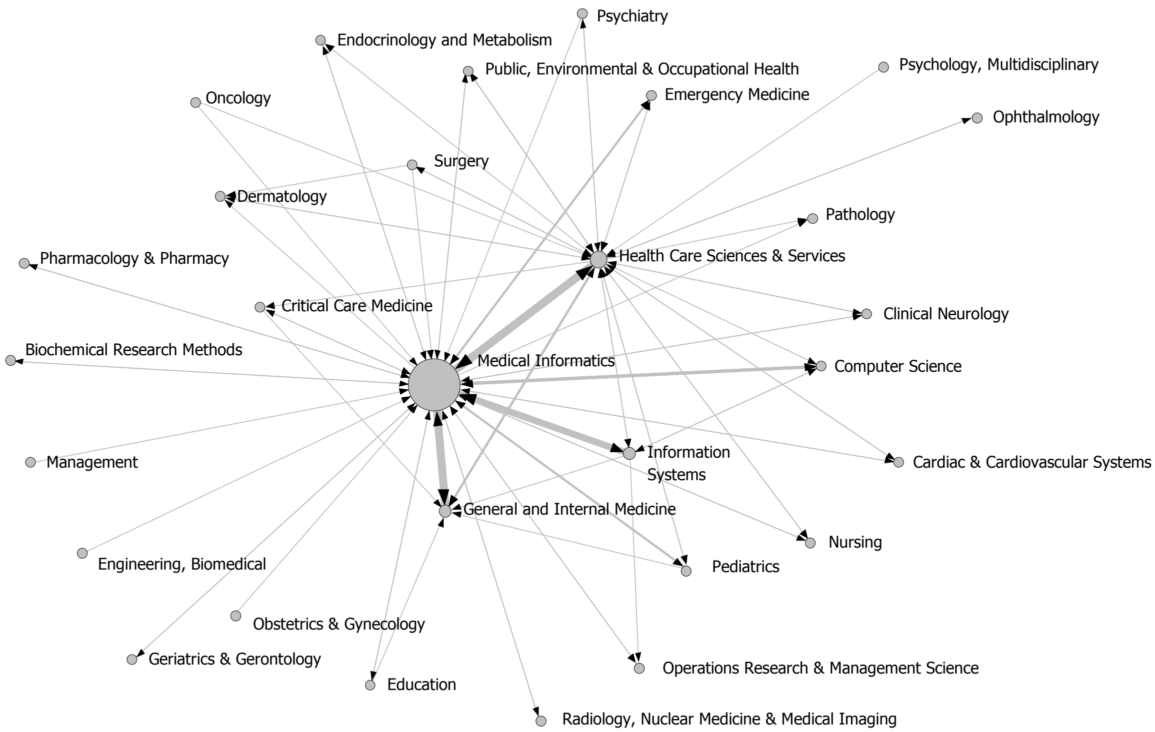


Figure 2.4 Citations among Sub-Disciplines of HI Research (Strength-of-ties ≥ 25[10])

In a similar manner, the intellectual structure of HI can also be inferred from the analysis of the co-citation network of the HI research disciplines, as shown in Figure 2.5. What the graphic shows is that, with the exception of Health Care Sciences & Services and General and Internal Medicine, Information Systems is most often co-cited among the sub-disciplines (including Operations Research & Management Science and Computer Science). This is consistent with recent studies on the intellectual structure of IS that find that management,

[10] Showing all ties in the diagram would lead to insuperable difficulties in interpreting the network structure. To simplify the diagram, only relationships with strength-of-ties equal to or larger than a specific threshold are displayed. In this we are consistent with the approach used by Euske et al. (2011) iteratively increasing the cutoff point to the point where the network structure becomes visually apparent. The interpretability of the network structure at a particular cutoff point strongly suggests the threshold to be used to reveal the social network structure. The same method is used to display other subsequent networks.

operations research, and management science are major contributors to the IS discipline (e.g., Polites and Watson 2009). For this reason, we next narrow our analysis to the sub-discipline of HIT.

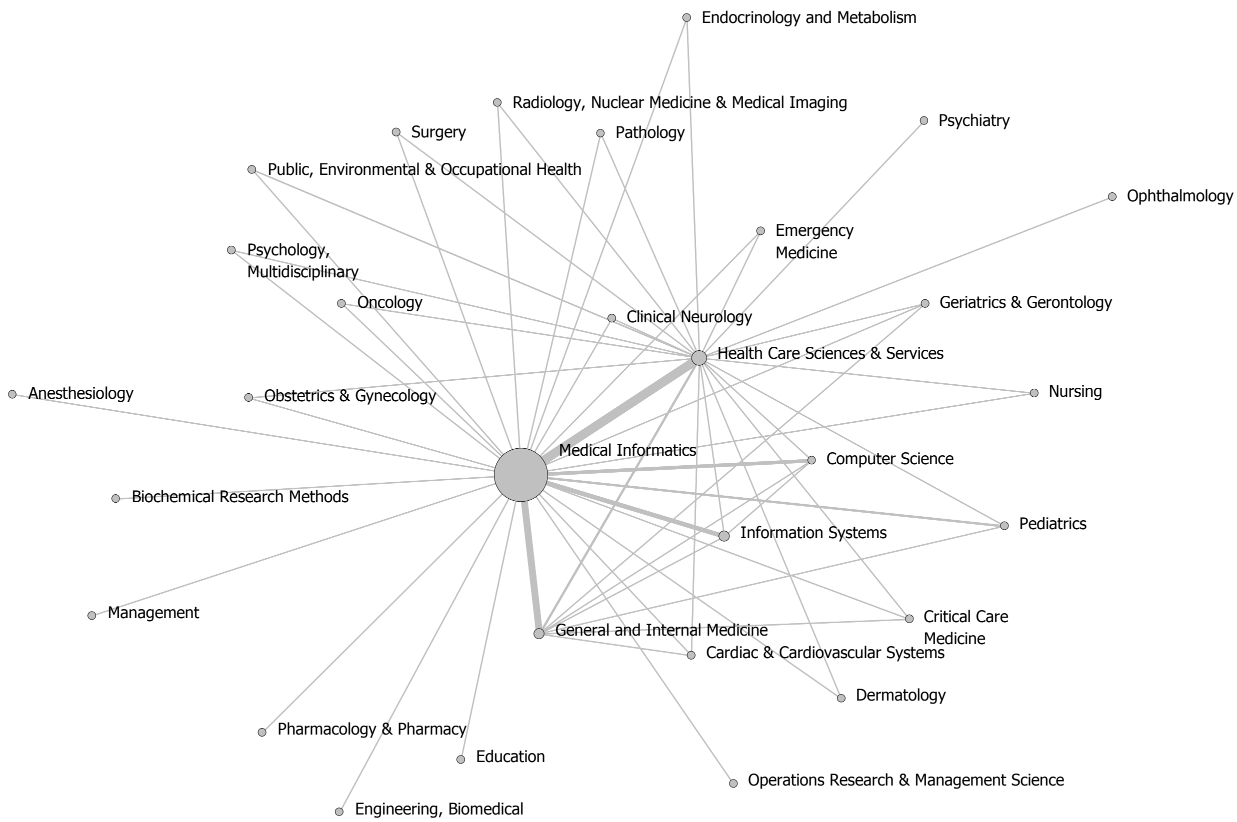


Figure 2.5 Co-Citations among the Sub-Disciplines Making up HI Research (Strength-of-ties ≥ 90)

2.4.2 Thematic Structure of HIT

An LSA of the term-document matrix (using a Varimax rotation) was best resolved with a 14-factor solution of HIT research themes. Each of these identified 14 factors represents a collection of articles that contain semantically similar groups of terms. For instance, the top loading factor, which we labeled Security of HIT, contains articles that similarly use joint terms (in their root forms) such as: secur, hipaa, comput, polici, and issu. The detailed high-loading terms and documents for the 14-factor solution can be found in Appendices 2E and 2F. The final set of extracted core HIT research themes (factors) includes (in order of the average loading): (1) Security of HIT; (2) Implications of HIT; (3) Medical Information Retrieval; (4) Medical Image Processing and Management; (5) Trust in HIT; (6) EMR and EHR; (7) Knowledge Management in Healthcare; (8) TAM of HIT; (9) National HIT Programs; (10) General HIT Application; (11) HIT Innovation; (12) HIT and Organizations; (13) Clinical Decision Support; and (14) Telemedicine.

2.4.2.1 Centrification of Most Content-Cohesive Core Themes

Table 2.3 shows the *content cohesion* of these 14 HIT research themes. We distinguish this form of cohesion from network cohesion, which will be examined later. Content cohesion of a research theme is defined as the average loading of papers belonging to this research theme.[11] A higher level of content cohesion of a specific theme means that the thematic sub-community is mature in using certain language and terminology in their articles, that is, they share common semantics in describing their research topic. Among the 14 HIT research themes, Security of HIT, Implications of HIT, Medical Information Retrieval, and Medical Image Processing and Management have the highest average factor-document loadings (i.e., ≥ 0.50). This suggests that these four research themes are the most content cohesive and thereby the most tightly-connected sub-communities with respect to semantic maturity. Research themes including HIT Innovation, HIT and Organizations, Clinical Decision Support, and Telemedicine have the lowest average factor-document loadings (i.e., < 0.30). This indicates that these four sub-communities are, at

[11] In this analysis of HIT research themes, we counted articles with document-factor loading coefficients ≥ 0.178, which is a threshold used to distinguish significant document-factor loadings from insignificant ones (Sidorova et al. 2008). The purpose of such cutoff point decisions is to retain 1/k of the loadings for a k-factor solution such that each term and document will just load on one factor, on average.

the present time, the least semantically consistent and are, therefore, exhibit low levels of content cohesion.

Table 2.3 Content Cohesion of Core HIT Research Themes from 1992 to April 2013

Factor	Label	Avg. Loading of Sig. Papers	% of Papers
1	Security of HIT	0.59	2.16%
2	Implications of HIT	0.58	2.47%
3	Medical Information Retrieval	0.54	1.85%
4	Medical Image Processing and Management	0.50	2.78%
5	Trust in HIT	0.41	4.32%
6	EMR and EHR	0.38	4.63%
7	Knowledge Management in Healthcare	0.37	5.25%
8	TAM of HIT	0.35	8.33%
9	National HIT Programs	0.33	6.17%
10	General HIT Applications	0.31	10.19%
11	HIT Innovation	0.26	14.51%
12	HIT and Organizations	0.25	11.73%
13	Clinical Decision Support	0.22	12.35%
14	Telemedicine	0.14	6.48%

2.4.2.2 Thematic Dynamics

Dynamic Year-to-Year Thematic Charts: Given that our sample of HIT articles spans an approximately 21-year period in which the discipline evolved considerably, HIT research themes are likely to shift over time. Therefore, we analyzed the temporal dynamics of above listed HIT research themes extracted via LSA. Figure 2.6 shows the dynamics of publication counts amongst the core HIT research themes [aggregated by counting articles with significant document-factor loadings (i.e., loading coefficients ≥ 0.178)]. The 14 research themes identified had sporadic publications before year 1996, while from the year 1997 to year 2003 we see quite a few fluctuations. Since year 2004, publications of core HIT research themes have steadily increased, with the exception of year 2007 which saw a spike in publication within a single year. The waxing and waning of HIT publication across the years speaks of the extreme volatility of yearly dynamics. Thus, to make more sense of the resulting counts in the subsequent section, we divided the overall range into 2 periods and conducted further analysis (next section).

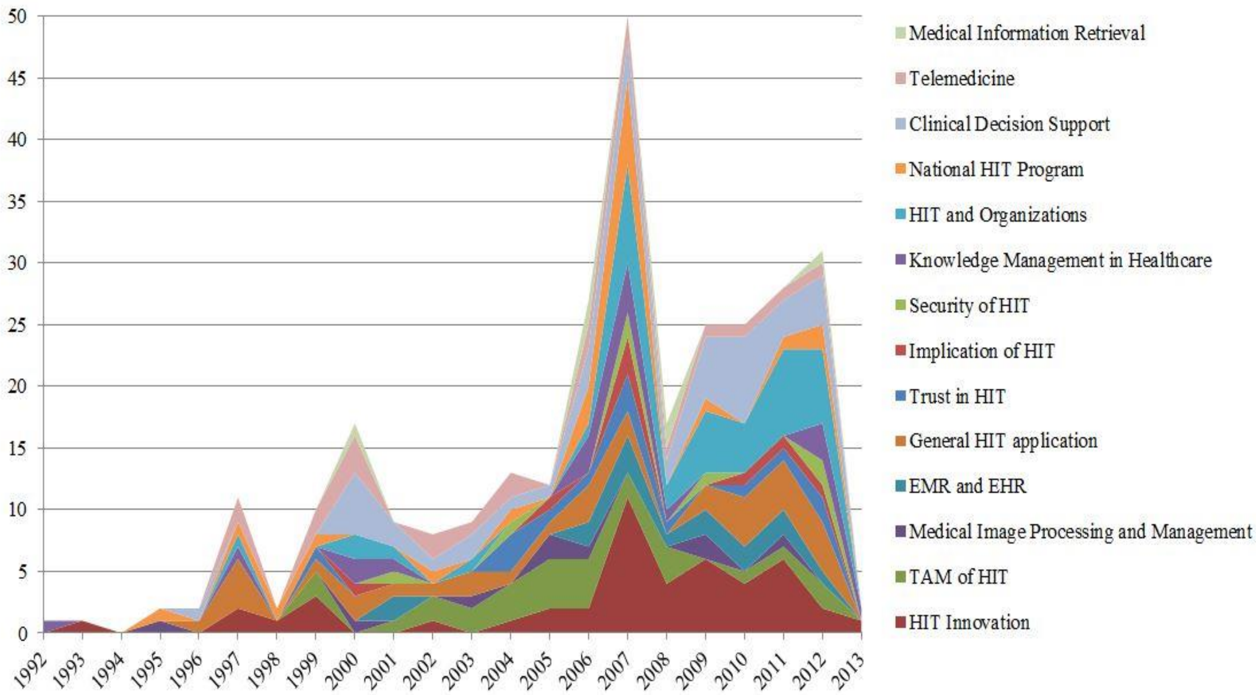


Figure 2.6 Waxing and Waning of Core HIT Research Themes[12]

Visualization of Trends by Using Era Analysis: We compared HIT research trends across two separate periods: (1) 1992 – 2002, and (2) 2003 – 2013. Figure 2.7 highlights the change of publication count percentages for all core HIT research themes across the two study periods. In the second period, HIT and Organizations, Trust in HIT, and HIT Innovation

[12] Please note that data collection was finalized in April of 2013, thus including fewer publications from 2013 in our sample.

changed most dramatically in popularity (downward trends) while research themes such as EMR and EHR, Implications of HIT, TAM of HIT, Security of HIT, Medical Information Retrieval, Medical Image Processing and Management, Clinical Decision Support, National HIT Programs, and Knowledge Management in Healthcare had modest percentage deltas, meaning that publication counts were more consistent between the two periods for these themes. Interestingly, the field also seemed to lose interest in two research themes, General HIT Applications and Telemedicine from one time period to the next. These themes were drastically downplayed in period 2 as compared to period 1.

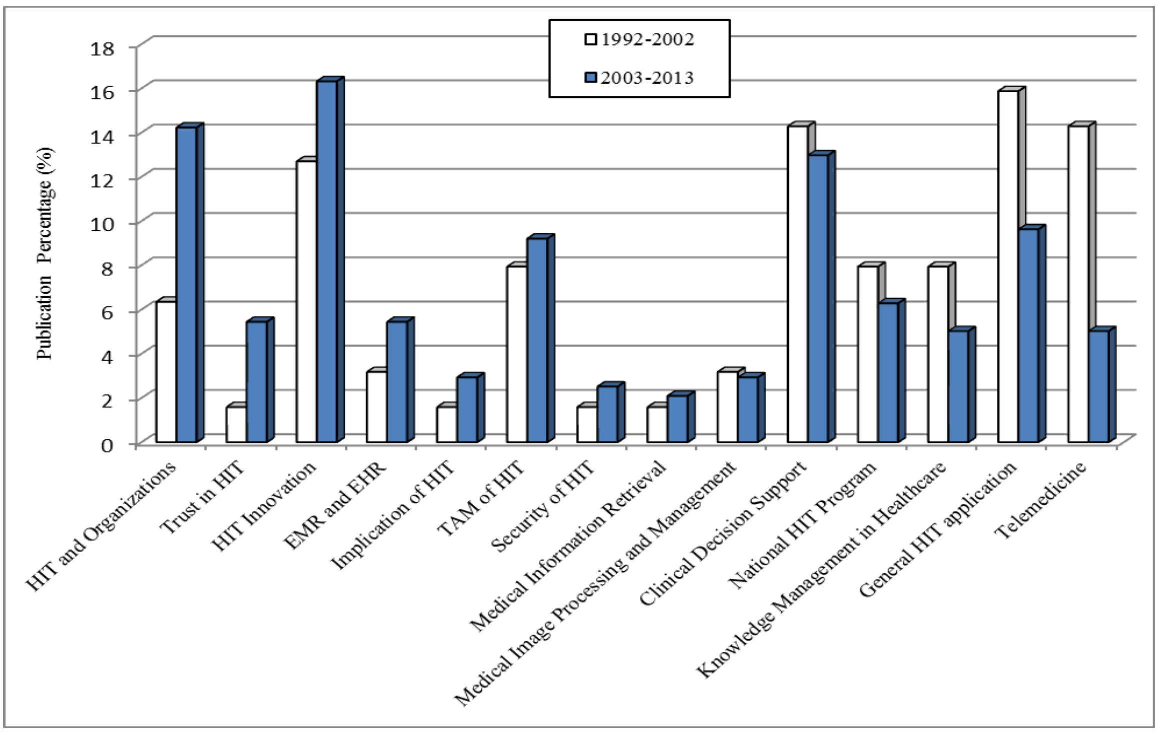


Figure 2.7 Changes in Paper Percentages
(Sorted in Descending Order from Period 1 to Period 2)

Research in the first period focused more on General HIT Applications, Clinical Decision Support, Telemedicine, and HIT Innovation while in the second period the themes of General HIT Applications and Telemedicine fell in interest levels as more research began to address the organizational context of HIT as well as trust in HIT settings. In terms of raw publication counts, HIT Innovation saw the largest number of publications in the most recent period, followed by HIT and Organizations and Clinical Decision Support. The areas least studied (based on raw publication counts) were, in descending order, Medical Image Processing and Management, Security of HIT, and Medical Information Retrieval.

What is clear is that there have been dramatic shifts toward and away from certain topics. One partial explanation for this shift could be that telemedicine issues have been solved, at least from a technical standpoint, and thus interest has declined.

Changes in Semantic Association of Core Themes: We analyzed the publication trends of these HIT themes across two periods. Table 2.4 reveals the change of average loading coefficients for all core HIT research themes across the two study periods. Although research themes such as Medical Information Retrieval, Implications of HIT, and Medical Image Processing and Management are still not strong foci of the HIT sub-discipline in terms of percentage of overall production, the linguistic connections are becoming stronger within these sub-communities. In contrast, previous cohesive HIT themes including General HIT Applications and TAM of HIT are becoming less cohesive. We can conclude that these newly addressed HIT research areas are still in the process of maturing, providing the potential for future research to fully address related research topics.

Dependencies among HIT Thematic Domains: SNA on the citation relationships among HIT research themes helps reveal those themes that are contributing most to the overall scholarly discourse, thus having more influence on the intellectual structure of the HIT community. The SNA citation relationships among the 14 core HIT research themes are shown in Figure 2.8. As before, the size of each node is proportional to the in-degree of the node, while thickness of the arrows and lines represents the relative strength of the citation relationship between any two nodes. We can classify 14 research themes into four categories, ordered according to degree centrality.

Group 1. Highly central themes

1. TAM of HIT
2. General HIT Applications

Group 2. Marginally central themes

1. HIT and Organizations
2. Telemedicine
3. HIT Innovation
4. Implications of HIT

Group 3. Specialized themes

1. Trust in HIT
2. Security of HIT
3. EMR and EHR
4. National HIT Programs
5. Clinical Decision Support
6. Knowledge Management in Healthcare

Group 4. Isolated themes[13]

1. Medical Information Retrieval
2. Medical Image Processing and Management

[13] The research themes Medical Information Retrieval and Medical Image Processing and Management are especially independent from the other themes, that is, having no citation relationship to any of the other HIT research themes (after applying the threshold criteria). Thus these two themes are not displayed in Figure 2.8. This makes sense, given the fact that Medical Information Retrieval and Medical Image Processing and Management are traditional focus areas of the HI discipline rather than the HIT sub-discipline.

Table 2.4 Trend of Core HIT Research Themes (Cutoff of Paper Loading ≥ 0.178)

1992- 2002			**2003- 2013**		
Rank	Theme	Avg. Loading (Percent)	Rank	Theme	Avg. Loading (Percent)
1	Security of HIT	0.66 (1.6%)	1	Implications of HIT	0.60 (2.9%)
2	TAM of HIT	0.49 (7.9%)	2	Medical Information Retrieval	0.58 (2.1%)
3	Implications of HIT	0.46 (1.6%)	3	Security of HIT	0.57 (2.5%)
4	Medical Image Processing and Management	0.40 (3.2%)	4	Medical Image Processing and Management	0.52 (2.9%)
5	EMR and EHR	0.38 (3.2%)	5	Trust in HIT	0.42 (5.4%)
6	General HIT Applications	0.37 (15.9%)	6	EMR and EHR	0.38 (5.4%)
7	Knowledge Management in Healthcare	0.35 (7.9%)	6	Knowledge Management in Healthcare	0.38 (5.0%)
8	Medical Information Retrieval	0.34 (1.6%)	7	National HIT Programs	0.35 (6.3%)
9	HIT Innovation	0.26 (12.7%)	8	TAM of HIT	0.32 (9.2%)
10	National HIT Programs	0.25 (7.9%)	9	General HIT Applications	0.28 (9.6%)
10	HIT and Organizations	0.25 (6.4%)	10	HIT Innovation	0.25 (16.3%)
11	Telemedicine	0.23 (14.3%)	10	HIT and Organizations	0.25 (14.2%)
12	Trust in HIT	0.22 (1.6%)	11	Telemedicine	0.24 (5.0%)
12	Clinical Decision Support	0.22 (14.3%)	12	Clinical Decision Support	0.22 (13.0%)

What does the intellectual structure of the HIT sub-discipline as shown in Figure 2.8 suggest? Except for Group 4, which shows no citations of the other HIT themes, a high percentage of works cite the TAM of HIT literature and General HIT Applications literature. What appears to be the case is that these citations by scholars are used, in many cases, to motivate their own work. To lesser extent, they also cite the HIT and Organizations, HIT

Innovation, Implications of HIT, and Telemedicine literatures. Group 3 are specialized areas that are themselves not as central in the citation patterns, no doubt due to their tighter focus on a particular aspect of HIT. Security of HIT is a good example of this kind of niche research.

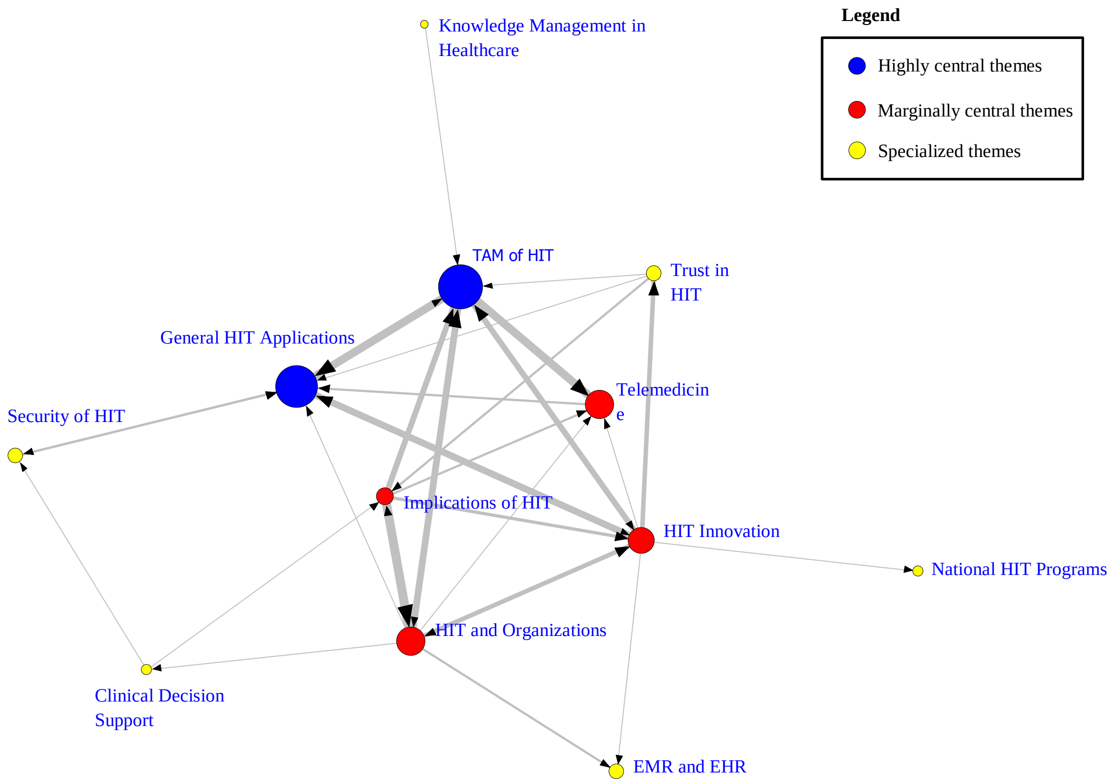


Figure 2.8 Citation Relationships among Core HIT Themes
(1992 – April 2013, Strength ≥ 0.355)

To compare the citation patterns and growth of all thematic groups, we next assessed the centrality and maturity of each HIT sub-community. Centrality refers to the extent to which a node connects to a social network. In this study, we used in-degree centrality and information centrality, metrics widely used to evaluate the *prestige* of network nodes. In-degree centrality is a localized metric or the number of direct relationships a research theme has with other themes. Information centrality takes into account all paths between HIT research themes, thus providing a measure for the relative drop in network efficiency if a particular theme is removed from the network (Polites and Watson 2009).

Another measure employed is the density of directed citation network within each thematic group, which is the ratio of all present relationships to all possible ties (Hanneman and Riddle 2005). A higher value of network density indicates a higher connectedness within the thematic group, and thus the thematic sub-community is "more integrated and interdisciplinary" (Biehl et al. 2006, p. 363). As a measure at the whole thematic subnetwork level, density indicates the *network cohesion* or *maturity* of each thematic HIT sub-community. Appendix 2G shows the citation network measures of HIT research themes with their rankings.

Before drawing inferences about these measures, how do in-degree centrality, information centrality, and network density relate to each other? Spearman correlation coefficients among rankings of three network measures appear in Table 2.5.[14] The Spearman R between degree centrality and information centrality is 0.853 ($p < 0.01$). This means that rankings obtained by the two measures move together to a large extent. This makes perfect sense in that a theme with strong direct connections with other themes will have an impact on the information flow of the overall network if it is removed. What is instructive, however, is that network density is neither significantly correlated with degree centrality nor information centrality (and the correlation coefficients have much lower explained variances of 0.39 and 0.21, respectively). This suggests that a thematic group which contributes the most to other thematic groups is not necessarily mature within its own group.

[14] "Medical Information Retrieval" and "Medical Image Processing and Management" were excluded from these and the subsequent analysis since these themes are isolated from the others. Thus, we were left with 12 themes for further analysis.

Table 2.5 Spearman correlation coefficients among network measures

	In-degree Centrality	Information Centrality	Subnetwork Density
Degree Centrality	1.000		
Information Centrality	0.853**	1.000	
Subnetwork Density	0.573	0.463	1.000

**: $p < 0.01$

To differentiate the maturity and prestige of HIT themes, we compared in-degree centrality with network density for each thematic group, since these two centrality measures are highly correlated. Figure 2.9 compares the prestige and maturity of the remaining 12 thematic groups. Trust in HIT and Security of HIT had high network density, but their centrality values were relatively low. This means that these two themes are cohesive (or mature) within their own group, but they do not receive high levels of citation from the other thematic groups. Contrariwise, although not cohesive within its own thematic group, General HIT Applications received numerous citations from other themes. It is also evident that current HIT research has stressed work on TAM in terms of both prestige and network cohesion while other HIT research themes are closer to the point of origin in Figure 2.9, including Knowledge Management in Healthcare, EMR and EHR, Clinical Decision Support, National HIT Programs, and Implications of HIT. These latter themes are thus emerging thematic domains. We later argue that these areas need more directive leadership so that future research can better support these less mature and less prestigious topics.

2.4.3 Thought Leadership in HIT

Up to this point, we have primarily discussed key HIT research themes and relationships between the identified themes. We now turn our attention to *thought leadership*, with a particular emphasis on authors of HIT papers in IS journals. We begin with some general and informative descriptive statistics, as descriptive statistics tell us a great deal about the makeup of the thought leadership in this domain. Our dataset of HIT papers contains 700 authors in total, with most authors publishing fewer than 2 articles; specifically, with 85.0% authors publishing only one HIT study and 8.7% authors two papers. The most prolific authors represent 6.3% of the author pool.[15] This finding is consistent with those conducted in other disciplines such as management control (e.g., Euske et al. 2011). It is also quite consistent with the power distributions uncovered by Chua et al. (2002) across baskets of 4 to 58 IS journals. What it also means in this context is that a small and elite group of authors constitute the thought leaders of the field and the burden of further developing the field falls heavily on their shoulders.

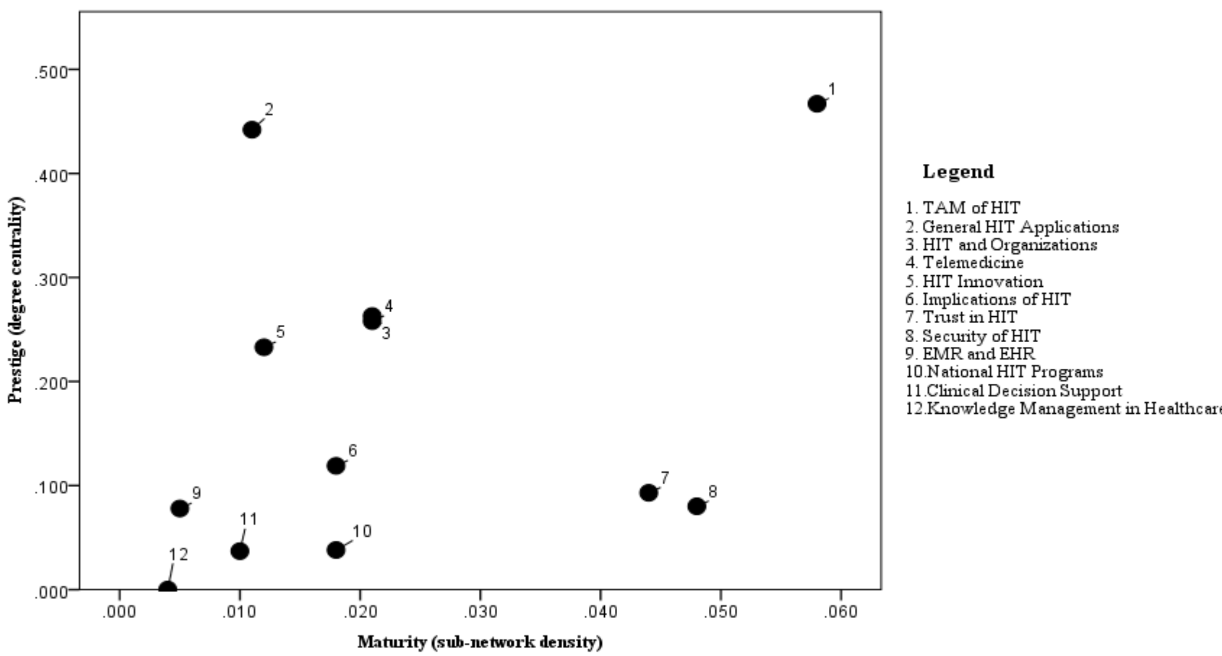


Figure 2.9 Comparison between the Prestige and Maturity of Thematic Groups

[15] A summary of author productivity can be found in Appendix 2H.

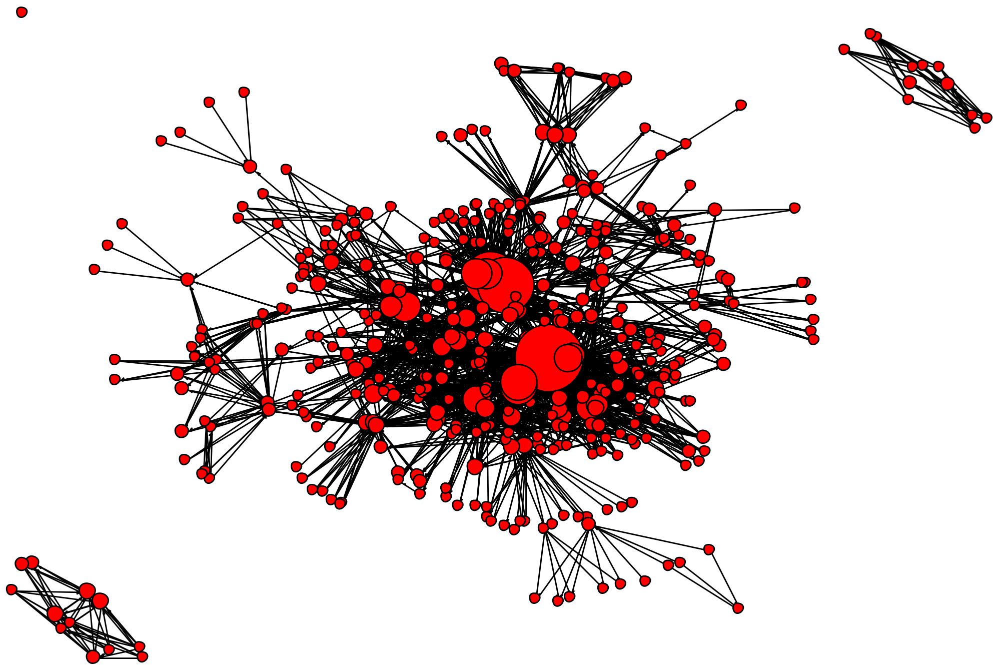


Figure 2.10 Overall HIT Author Citation

Figure 2.10 visualizes the overall author citation relationships in the HIT discipline, in which several scholars dominate the citation structure with two small outlying clusters of citation relationships among small, isolated cliques. To make better sense of this important element of the intellectual structure of HIT, Figure 2.10 displays HIT scholars who have been cited by other HIT scholars at least once. The figure is unlabeled to demonstrate how complex a network structure appears when filtered at this most elementary level. Ironically, and like most real world networks, the HIT thought leadership network is actually a very sparse network. Because network density is a factorial, most real world networks are exactly like this. As soon as several dozen nodes are defined in a network, the likelihood that they would all be connected to each other drops exponentially. The result of this filtering is patterns among 263 HIT scholars. Similar to the prior analyses, the size of each node is proportional to the in-degree of the node.[16]

Categorizing all the HIT scholars by their in-degrees, we obtained a 4-cluster solution (with more specific information on these thought leaders discussed next):

Cluster 1: Kohli, R.

Cluster 2: Hu, P.J.H. and Chau, P.Y.K.

Cluster 3: 14 HIT scholars including Devaraj, S., Davidson, E.J., Rivard, S., Lapointe, L. and 10 other authors (see Table 2.6 for a complete listing)

Cluster 4: 246 remaining scholars

To further explore the citation relationships among HIT thought leaders and scholars, we zoomed in one end of the distribution by showing only scholars with an in-degree ≥ 12 and citation strength-of-tie ≥ 2, as shown in Figure 2.11. This simplified network contains 45 HIT highly-cited scholars. The top 20 most highly cited HIT scholars are listed in Table 2.6 with their rankings.

These scholars (Table 2.6) represent the intellectual thought leaders of the HIT field. Given the *network centrality* demonstrated by the in-degree citations, these scholars have been setting the direction for research for the last several decades. However, thought leadership is often focused on particular themes and, in recognition of this observation, we also analyzed thought leadership by HIT research theme (Table 2.7). This analysis provides more granular insights into the primary contributors and influencers per research theme, which hopefully gives

[16] As an exception, we found one scholar with no citation relationship with other HIT scholars, all of whom were cited at least once by the entire HIT community.

current and future researchers a better idea of which authors to search for when seeking seminal and influential articles to cite and build upon in their own work.

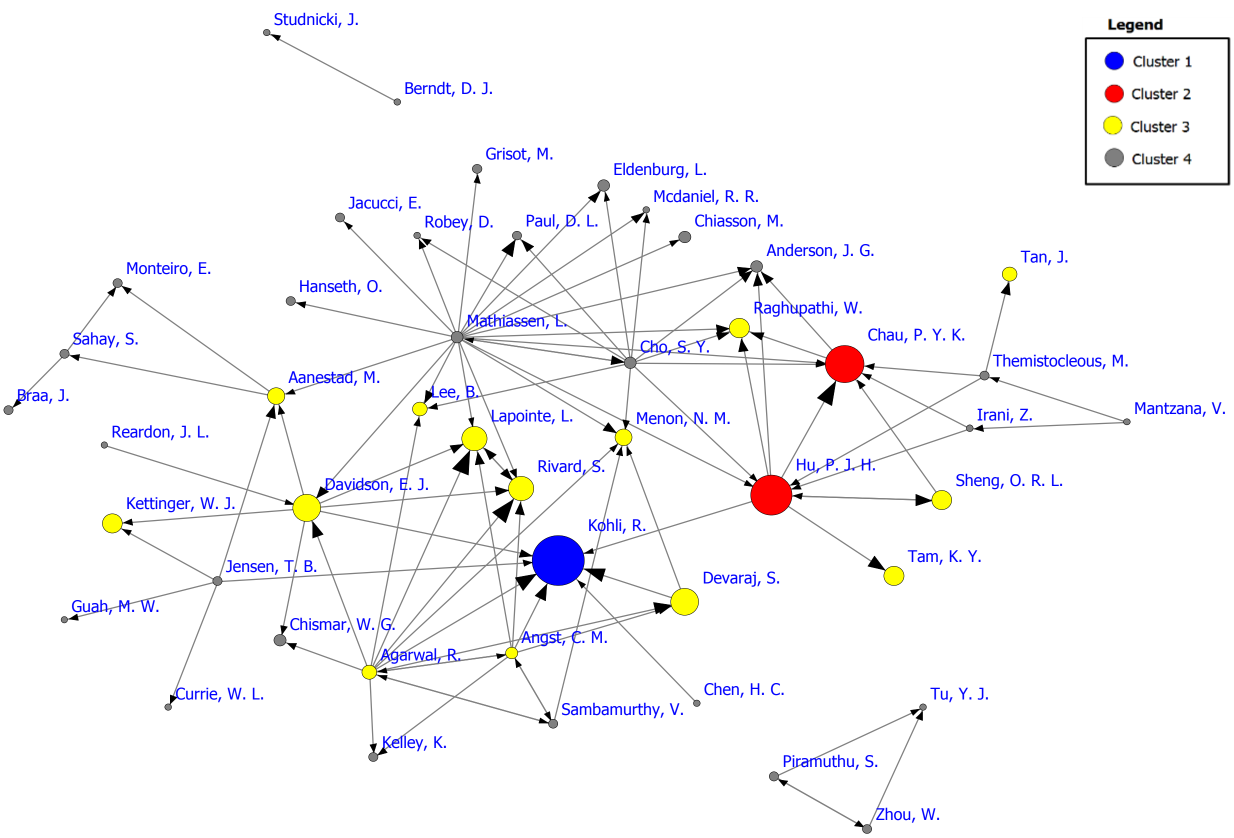


Figure 2.11 Most Highly-Cited HIT Authors
(Top 45 Scholars, In-degree ≥ 12, Strength-of-ties ≥ 2)

Table 2.6 Top HIT Scholars according to In-Degree Citation Counts

Rank	Author	In-Degree	Rank	Author	In-Degree
1	Kohli, R.	153	14	Angst, C. M.	32
2	Hu, P. J. H.	119	15	Chismar, W. G.	28
3	Chau, P. Y. K.	115	16	Anderson, J. G.	27
4	Devaraj, S.	82	16	Eldenburg, L.	27
5	Davidson, E. J.	79	16	Chiasson, M.	27
6	Rivard, S.	72	17	Mathiassen, L.	25
6	Lapointe, L.	72	17	Cho, S. Y.	25

7	Sheng, O. R. L.	56	18	Jensen, T. B.	24
8	Tam, K. Y.	53	19	Hikmet, N.	22
8	Raghupathi, W.	53	19	Bhattacherjee, A.	22
9	Kettinger, W. J.	50	19	Paul, D. L.	22
10	Menon, N. M.	47	20	Sambamurthy, V.	20
11	Aanestad, M.	45	20	Sahay, S.	20
12	Agarwal, R.	40	20	Monteiro, E.	20
12	Lee, B.	40	20	Kelley, K.	20
13	Tan, J.	35			

Table 2.7 Leadership within HIT Core Themes (Top 3 Authors)

Theme	**Author**	**Citation**
TAM of HIT	Chau, P. Y. K. Hu, P. J. H. Lapointe, L.	45 45 24
General HIT Applications	Raghupathi, W. Tan, J. Mercuri, R. T.	18 10 5
HIT and Organizations	Kohli, R. Devaraj, S. Agarwal, R. Angst, C. M.	29 25 11 11
Telemedicine	Chau, P. Y. K. Hu, P. J. H. Devaraj, S. Kohli, R.	32 32 25 25
HIT Innovation	Davidson, E. J. Aanestad, M. Chismar, W. G.	17 10 10
Implications of HIT	Bhattacherjee, A. Hikmet, N. Brooks, R. G. Kayhan, V. O. Menachemi, N.	7 7 4 4 4

Trust in HIT	Paul, D. L. Mcdaniel, R. R. Zahedi, F. M.	10 7 3
Security of HIT	Mercuri, R. T. Huston, T. L.	5 4
EMR and EHR	Aanestad, M. Jensen, T. B. Huston, T. L.	5 5 4
National HIT Programs	Currie, W. L. Guah, M. W. Eason, K.	6 6 2
Clinical Decision Support	Walczak, S. Lee, B. Menon, N. M.	5 4 4
Knowledge Management in Healthcare	Davidson, E. J. Heslinga, D. Paul, D. L.	3 3 3
Medical Image Processing and Management	Aboulafia, A. Blum, J. M.	1 1
Medical Information Retrieval	Chen, H. C. Qin, J. L. Zhou, Y. L.	1 1 1

2.5 Discussion and Conclusion

We build upon and extend prior work by contributing multi-method analyses that span two decades of HI and HIT research and provide insights into cohesion of content and networks, thematic dynamics, and thought leadership. Our findings indicate that whereas the raw bulk of research in the HI field is currently taking place outside of the IS discipline, the field of IS is the "second among equals" of those disciplines that are key contributing disciplines (Lee 2003, p. 319). The most powerful forms of scientific influence are found in the citation numbers that are naturally generated by a preponderance of non-IS journals. Medical informatics and healthcare sciences and services currently dominate this space, but information systems and computer science share the next most cited position among the others.

We find that the sub-disciplines of HI are co-citing, however, and there is a developing mutual influence, as shown by the citation and co-citation networks. What is also evident from these visualizations is that IS is somewhat better positioned in the networks (more co-citations and in-degree citations) than other key fields like computer science and much better positioned than operations research, management science, and general management. HIT leaders can increase this influence in obvious ways such as interacting more frequently with the larger HI communities. Greater attention will be paid to the value and impact of HIT research through specialized publication outlets such as those suggested by Lucas et al. (2013).

2.5.1 Research Themes of HIT

We have demonstrated that 14 themes characterize the overall production of HIT research, but two of these themes (Medical Information Retrieval and Medical Image Processing and Management) are tangential and isolated from the others, garnering the lowest levels of citations from the remainder of the HIT field. What this means, essentially, is that these themes are more closely connected to the HI community than to the HIT community. Whereas tying them more closely to the HIT field is feasible, it might be preferable to expend the scarce organizational energies of HIT scholars on the other 12 emergent themes. With this positioning in mind, we discuss findings related to the remaining HIT themes.

Over the two decades of HIT activity studied, these themes have shifted in frequency of publications, content cohesion, and network density. One theme appears to be highly citation-central to the other themes in motivating work (TAM of HIT), but over time decreasing drastically in consistent semantics to refer to the theme (content cohesion). If the content cohesion of TAM continues to decrease this much, HIT TAM studies could well become less frequently cited as other means of motivation emerge.

Themes with increasing interest levels (i.e., deltas in frequency of articles) include HIT and Organizations, Trust in HIT, HIT Innovation, and EMR and EHR. Stable themes include TAM of HIT, Security of HIT, and Clinical Decision Support. Interest levels in Telemedicine, General HIT Applications, Knowledge Management in Healthcare, and National HIT Programs have been dropping off over the decades. The HIT community as a whole will decide whether to rejuvenate these themes or not.

One desideratum for determining whether to develop these themes further is the currently low levels of both prestige and maturity of Knowledge Management in Healthcare, EMR and EHR, Clinical Decision Support, National HIT Programs, and Implications of HIT. As demonstrated in the XY plotting of maturity and prestige, Security of HIT and Trust in HIT are mature in their use of consistent semantics, whereas (likely because they are niche areas) they are low on network density of citations from other HIT sub-fields. These themes thus appear to be maintaining their positions in the overall HIT scholarly community. The other named themes can be much further developed along the lines of both maturity and prestige. Whether this occurs is also a function of whether the thought leaders identified by this study will step forward and advance the work of the thematic community. We discuss this possibility next.

2.5.2 Leading Scholars in HIT

With leadership comes responsibility. We have identified the HIT leaders both in the overall metrics and in analyses of its sub-communities. The thematic sub-communities of the HIT field, no doubt, have high expectations of their leaders and our analysis helps the community by identifying those authors whose work has been most influential to date.

Intellectual leaders can likely be even more proactive in advocating for and heading up special issues in our top journals. They can also take on more organizational roles that should come naturally with idea leadership. Is it not time for an HIT Special Interest Group (SIG) in AIS? Would not the HIT community be well served with pre-conference workshops and an online knowledge forum for sharing working papers and completed work?

Given that our analysis shows which sub-themes need to be more concerned with cohesion across semantics and intra-citation patterns, the identified thought leaders can serve as role models for remedial actions. There is a sense that thought leaders, more than other members of the sub-communities, can and should lead by being aware of all of the relevant work in the sub-community and making full use of it. As exemplified in their citation patterns, their journal and conference papers can highlight the important knowledge creation taking place in the sub-community and encourage others in the research stream to be cognizant of critical prior work. Consistent use of language by leading scholars in describing intellectual themes will help greatly in the cohering of sub-themes. Intra-theme citation of important work will help the field to mature and lead to greater prestige.

It would also seem to be the natural outcome of identifying those who are leading the idea generation in HIT that these leaders would also forge ahead with "blue ocean" ideas in their own work (Straub 2009). It is devoutly wished that they also encourage the work of others in innovating beyond the topics that have dominated the field for the last twenty years. We offer suggestions for what these novel areas might look like in the following sections.

2.5.3 Limitations

Our research is limited by: (1) limitations of methods, (2) limitations of data collection (e.g., time frame and reliance on Web of Science), and (3) limitations in inference and generalization. However, even with such limitations, we believe our analyses, findings, and interpretations offer interesting insights into the development and evolution of this growing discipline.

2.5.4 Opportunities for Future Research

Future research on publications and research within the HI and HIT academic disciplines could: (1) expand the time frame of analysis as time progresses and as research trends evolve, (2) delve deeper into the sub-communities identified in our analyses (e.g., Security of HIT) for further and more fine-grained insights, and (3) apply new and novel methods to the content of published articles and relationships between articles. Potential future research within the HIT academic discipline, as motivated by the findings and interpretations in this paper, is discussed in the following sections.

2.5.4.1 Maintaining or Increasing Cohesion and Research Theme Life Cycles

We suggest that research themes undergo life cycles, similar to products and services in a marketing context metaphorically represented by growth stages ranging from infancy to maturity, and eventually are disrupted or renewed. Significant research opportunities are available in all such stages in the HI and HIT disciplines as the discipline itself is relatively young. Therefore, many new research themes and topics are emerging (as discussed below), many themes discussed in this paper are moving to adolescence and maturity and could benefit from application of mature methods and theories toward the goal of increasing content and network cohesiveness, and many opportunities for renewal will continue to become available as a dynamic environment impacts the context of HI and HIT use. To the last point, regulation and policy are currently in a dynamic state, especially in the areas of healthcare payment reform and HIT meaningful use, both of which are having an enormous impact on patterns of HI and HIT design, adoption, use, and evaluation. Therefore, even as research themes mature, the environment is changing and, as such, provides new and interesting opportunities to further validate and/or update our understandings. In particular, we suggest that research themes identified in this paper, such as Security of HIT, Implications of HIT, HIT and Organizations, and EMR and EHR which are regularly impacted by changes and updates to meaningful use policies (Blumenthal and Tavenner 2010), may become more cohesive research themes as policy, research, and use interact to further validate and incrementally refine existing findings. However, it is also likely that many of the interactions between policy, research, and use will result in disruptive findings. Therefore, as briefly discussed in the next sections, we are likely to witness much iteration of research theme life cycles over the next several years.

2.5.4.2 Spanning Boundaries (Where Appropriate)

HI and HIT research could benefit enormously from boundary spanning research that seeks to develop insights beyond insular patterns that often impact maturing research streams. For instance, TAM of HIT is identified in this paper as a mature and cohesive theme, but consumer acceptance of HIT is likely to be impacted by a complex mix of economic and behavioral constraints and incentives. This research stream is likely to benefit from research that incorporates theories and constructs from other academic disciplines, such as marketing and consumer behavior, that leverage the unique and dynamic context of HIT to both validate and update existing theoretical notions of correlation and causation. Existing research on services (especially in complementary contexts where physical interactions are difficult to substitute with technology, as is often the case in health care), consumer choice and decision making patterns, and supply-side challenges with addressing demand heterogeneity while retaining revenue and market share could be applied, tested, and refined to and within the HIT context. Such research would be especially beneficial to the HIT discipline as well as coordinate and reference disciplines as multi-theoretic impacts are likely to be the norm. Consumer choice and decision making will not occur in a vacuum. Economic and policy considerations are likely to impact this process, offering an opportunity for researchers to expand current theories through the use of a complex (rather than "reductionist") context.

2.5.4.3 Novel Areas That Will Further Enrich the Intra-Community Knowledge Base

As acknowledged in our limitations, our data collection went through April of 2013, but we are already witnessing significant new contributions to the HIT academic discipline. For instance, while themes such as General HIT Applications and Implications of HI are maturing and cohering, new HIT artifacts and ways of using (or updating) existing artifacts are emerging, creating new opportunities to further explore existing constructs and to develop new constructs [or update existing theories in new contexts, as suggested by Johns (2006)]. We are now beginning to observe the expansion of existing research themes into new sub-communities of thematic interest. We acknowledge that many potentially impactful future research streams are discussed (Agarwal et al. 2010; Baird 2014; Jones et al. 2014; Kellermann and Jones 2013; Romanow et al. 2012) and we seek to further contribute to this growing list by considering how the themes in this study are providing the foundation for recently emerging themes:

Consumer HIT and Consumer Informatics: Many new technologies have emerged that intermediate the "supply-side" of health care (providers, payers, suppliers) with the "demand-side" (consumers). As found in this paper, much prior research has explored the TAM of HIT, the Implications of HIT, and the Security of HIT, to name a few related themes, but only recently are researchers applying (and expanding) these themes into the context of consumer-facing and patient-facing HITs. The consumer context is uniquely heterogeneous where choice and usage decisions are individual, rather than firm-centric. Technologies have even emerged that allow consumers to manage their own health without the need of a provider (substitute) or as a significant complement to in-person health care services. Recognized technologies include: personal health records (PHRs), patient portals, social media, online health communities, health tracking devices and services, telehealth, and mHealth. Future technologies and informatics research challenges in this emerging context will likely expand upon existing research and renew many current HIT research themes. Specialized topics that may afford significant and interesting opportunities as well as direct connections to existing theory include: personalized medicine (e.g., genomics and pharmacogenomics) and personalized health services, business model challenges (e.g., reimbursement for telehealth across U.S. state lines), and the challenges associated with meeting individualized (heterogeneous) needs in a resource-constrained (and dynamic) environment.

Advanced "User-Centric" Artifact Designs: Existing research themes such as Medical Information Retrieval, Medical Image Processing and Management, Knowledge Management in Healthcare, and Clinical Decision Support often assume significant limitations associated with expert and decision systems, especially given the complex and difficult to predict nature of provider-patient interactions, diagnoses, and treatment. Therefore, research in these themes often focuses on research issues such as overcoming usage resistance or effectively dealing with search and retrieval challenges. We suggest that these assumptions are beginning to be challenged with the ever increasing capabilities of expert and decision systems, especially now that such systems are becoming more accurate even when the logic required is "fuzzy" or based more on patterns, connections, and correlations than hard-and-fast rules. Concurrently, health care professionals are realizing the benefit of unstructured data and are seeking novel ways to leverage HITs to balance the need for structure (as is often needed for billing) with the need for variation (as is often needed in clinical settings). Thus, the limitations of standard “rule-based” algorithms are becoming apparent, especially in health care. Applying standardized user-interfaces or clinical decision support algorithms to entire population segments can result in significant clinical and administrative errors, especially as observed with the challenges associated with standardized EHR UIs and CDS engines. While reducing variation can improve quality, heterogeneity must also be addressed. Therefore, the potential benefits of artificially intelligent and context aware technologies that leverage the benefits of machine learning are significant, but also require significant HIT research contributions if effectiveness is to be realized.

Optimal Decision Making: Health care is replete with trade-offs and optimizations, especially at firm and individual levels. While some research themes consider (or infer) trade-offs and the need for optimizing between multiple (and often competing) attributes, as is often the case with Trust in HIT, HIT Innovation, and EMR and EHR research themes, explicit consideration of the complexity of trade-offs is only now beginning to emerge. This is primarily due to policy efforts focused on reforming many aspects of health care simultaneously. At the highest level of policy making, the question of effectively lowering costs, improving health outcomes, and improving health care (referred to as the “triple aim” (Berwick et al. 2008)) remains open, as achieving all three simultaneously has proven to be an enormous challenge. Going forward, many theories could contribute to our understanding of any one of these items, such as how to lower costs by increasing information transparency, for instance. Achieving all three simultaneously, however, will likely require research that evaluates how various theories and models interact. For instance, how might challenges associated with economic notions of switching costs and lock-in interact with consumer behavior constructs such as involvement, diagnosticity, and decision (or choice) models? Or how might the need for efficiency and quality in service delivery be optimized (Rust and Huang 2012; Rust and Huang 2014)? Given that multiple stakeholders at multiple units of analysis must jointly and interdependently make choices that result in optimal balances between attributes for all parties (e.g., policy makers, providers, payers, suppliers/producers, and consumers), it is highly likely that interdisciplinary research will be essential to furthering our theoretical understandings.

Population Health (and Analytics): While the Implications of HIT is a maturing area of research, new challenges are emerging at multi-level units of analysis that will require new research and new points-of-view. This is especially true as policy making efforts seek to improve overall health care of entire populations. One of the biggest challenges in health care is balancing the needs of the individual with the needs of the population and considering the implications of various approaches to balancing sometimes conflicting goals. This challenge has never been more apparent as new models of health care delivery are emerging (e.g., patient centered medical homes, PCMHs, and accountable care organizations, ACOs), but have not been fully researched. We could discuss this area at length, but, in short, researchers need to ask how large datasets ("big data") and associated technologies and informatics approaches can be leveraged to generate population-level insights that trickle down to the heterogeneous needs of individuals with three overall goals in mind (as mentioned earlier): lowering costs, improving health, and improving health care (Berwick et al. 2008).

2.6 Conclusion

We began this paper by discussing the importance of understanding the intellectual structure of an academic discipline. As academic disciplines grow, expand, and even fracture, so do the research themes and sub-communities within them. Over time, knowledge can fragment, especially in multi-disciplinary fields such as HI and HIT. Deeper understanding of the evolving intellectual structures of innovative and contextually interesting disciplines provides a means by which to further expand, consolidate, and renew the discipline in a systemic and informed manner while also theoretically contributing back to coordinate and reference disciplines. Given that an in-depth intellectual structural analysis of HIT focused on research in top information systems journals had not appeared before our study, we fill an important research gap in this paper. We used multiple, rigorous methods, including citation and co-citation analyses, LSA, and SNA, to probe the intellectual structures of HIT. Our results clearly show that the field of HIT has evolved by shifting its research stream foci, through the changes in content cohesion, prestige and maturity of its sub-communities, and the emergence of its thought leaders. This is an exciting time in the HIT discipline and we are optimistic about the plethora of research projects that have already been carried out and those that will be conducted in years to come. We take a natural step to instantiate this optimism by providing insights into potential future directions of HIT research that should continue to enhance the depth and breadth of HIT intellectual structures. In conclusion, we encourage current and future HIT researchers alike to recognize how they are contributing to the intellectual structures that will systematically consolidate, expand, and renew the HIT knowledge base.

Appendix

Appendix 2A: HI Article Selection

Table 2.8 shows the number of articles identified for major HI journals.

Table 2.8 Major HI Journals (HI Articles in Our Dataset > 50)

Journal	***Articles***
Journal of the American Medical Informatics Association	4100
Computer Methods and Programs in Biomedicine	2056
International Journal of Medical Informatics	1556
Methods of Information in Medicine	1195
Artificial Intelligence in Medicine	1078
IEEE Transactions on Information Technology in Biomedicine	1059
Journal of Telemedicine and Telecare	997
Journal of Biomedical Informatics	909
Journal of Medical Internet Research	755
Journal of Medical Systems	725
CIN-Computers Informatics Nursing	510
Telemedicine Journal and E-Health	505
BMC Medical Informatics and Decision Making	502
M D Computing	264
Medical Informatics and the Internet in Medicine	194
Computers in Biology and Medicine	191
International Journal of Bio-Medical Computing	186
Telemedicine and E-Health	182
Medical Informatics	178
Artificial Intelligence in Medicine, Proceedings	143
IEICE Transactions on Information and Systems	112
Medical Decision Making	108
Informatics for Health & Social Care	102
Health Information Management Journal	100
Health Informatics Journal	94
Journal of Health Communication	86

International Journal of Clinical Monitoring and Computing	85
Health Information and Libraries Journal	85
Journal of the Medical Library Association	82
Decision Support Systems	81
Journal of Evaluation in Clinical Practice	66
Journal of the American Society for Information Science and Technology	65
Computers and Biomedical Research	62
Biomedical Engineering-Applications Basis Communications	61
Journal of General Internal Medicine	60
Journal of Digital Imaging	60
Biomedizinische Technik	59
Telemedicine Journal	57
Pediatrics	56
Aslib Proceedings	55
Mathematical and Computer Modelling	54
Bulletin of the Medical Library Association	54
Health Affairs	51

The yearly publication counts of HI research are depicted in Figure 2.12.

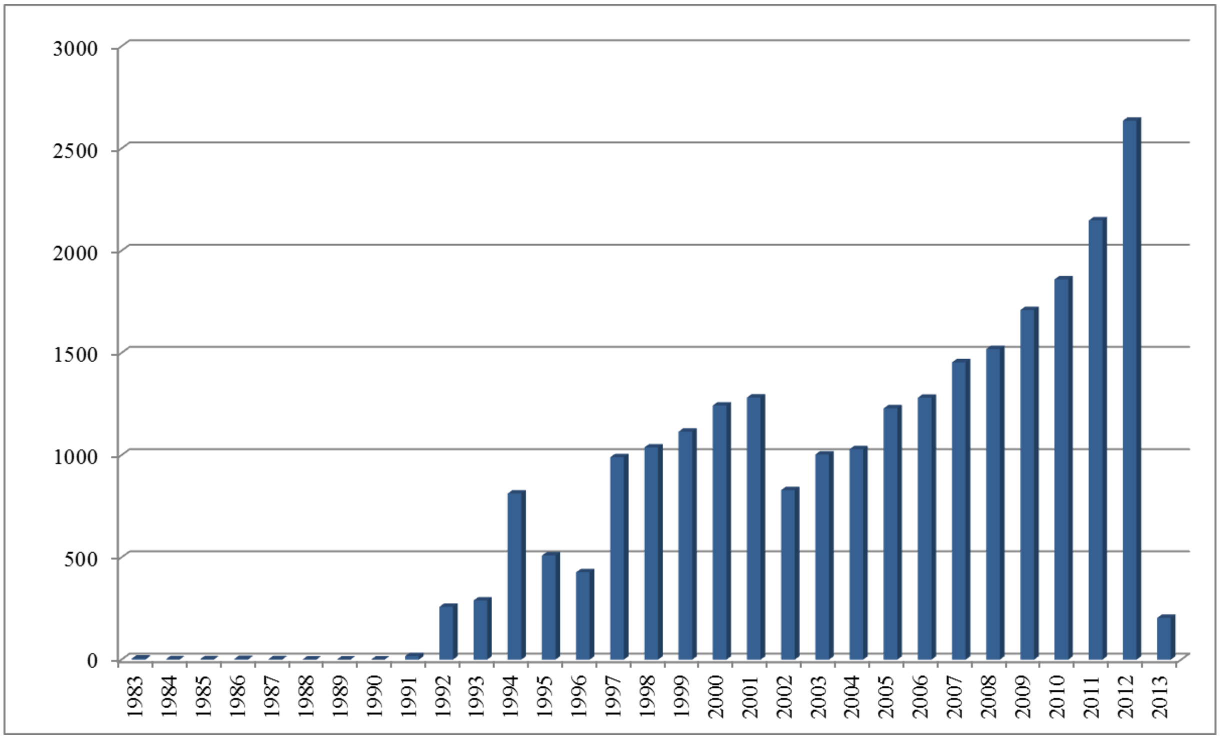


Figure 2.12 HI Yearly Publication Counts

Appendix 2B: HIT Article Selection

Table 2.9 shows the number of articles identified for mainstream management and IS journals.

Table 2.9 Journal Selection

Mainstream Management and IS Journals	***# of Retrieved Articles***
Decision Support Systems	78
Communications of the ACM	42
European Journal of Information Systems	28
Information Systems Frontiers	27
Information & Management	24
Journal of Information Technology	21
Journal of Management Information Systems	16
MIS Quarterly	15
Information Systems Research	13
Journal of the Association for Information Systems	11
Journal of Computer Information Systems	11
Information Systems Management	11
Journal of Strategic Information Systems	8
Information Systems Journal	7
Management Science	5
Organization Science	4
Human Relations	3
Total	324

The yearly publication counts of HIT research are depicted in Figure 2.13.

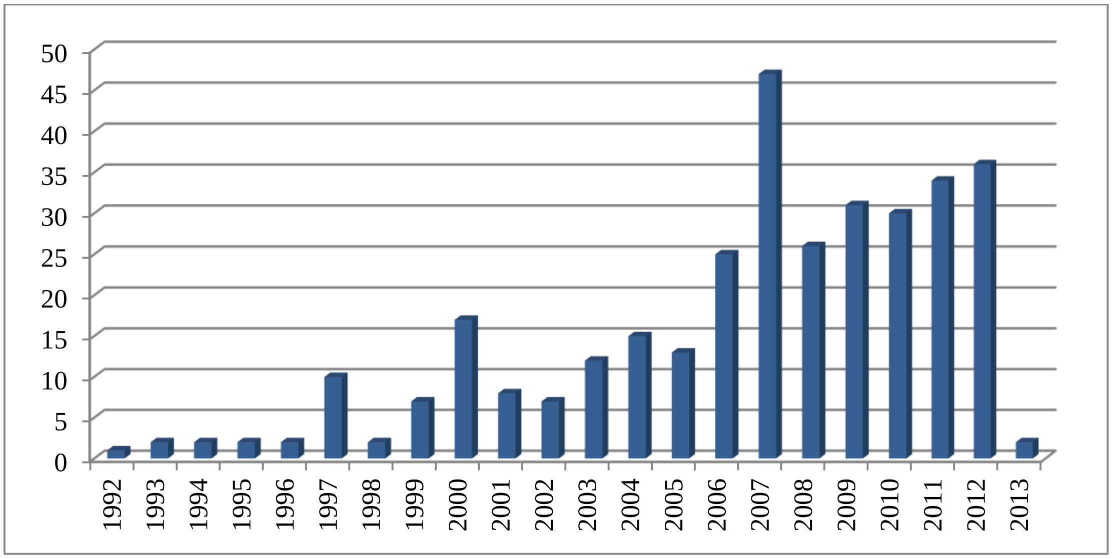


Figure 2.13 HIT Yearly Publication Counts

Appendix 2C: Latent Semantic Analysis Procedure

Latent semantic analysis (LSA) was initially proposed as an information indexing and retrieval approach based on conceptual content rather than exact match of inquiry words (Deerwester et al. 1990). Following the similar LSA procedure used by Sidorova et al. (2008), we systematically analyzed the research themes of HIT via the following procedure:

***Step 1.* Text Preprocessing and Term Reduction**

Abstracts were extracted from all exiting papers. Then the abstracts were tokenized by filtering out non-letter characters. Stop words such as "the", "this", and "a" etc. were filtered out since they only have trivial meaning in English. All tokens with just one letter (such as "c", "d", and "e" etc.) were also removed. After transferring all tokens into lower case, the Porter stemming algorithm (Porter 1980) was used to remove term suffices. For example, tokens such as "collaborate", "collaborating", "collaboration", and "collaborative" were replaced by their common stem "collabor". Finally, terms with only one occurrence were also filtered out since they did not load to more than two documents and were trivial to LSA. As a result, we obtained 1,879 terms.

***Step 2.* Generating TF-IDF Matrix**

LSA analyzes the relationships between a set of documents and terms contained in these documents by generating a set of concepts that are related to both the documents and the terms. LSA starts with a term-document matrix which describes the occurrence of terms in corresponding documents. In this study, a TF-IDF (term frequency–inverse document frequency) term-document matrix with 1,879 rows (terms) and 324 columns (documents) was created, which represented the relevant importance of terms to a corpus of documents (Wu et al. 2008).

***Step 3.* Applying SVD on the TF-IDF Matrix**

Central to LSA is singular value decomposition (SVD), which reduces the dimensionality of the term-document matrix to derive a particular latent semantic structure model. The latent semantic structure model is comprised of a set of orthogonal factors from which the original matrix can be approximated by linear combination (Deerwester et al. 1990). The SVD was applied to the TF-IDF matrix to reduce dimensionality. As a result, three matrices were obtained: 1) a term-by-factor matrix describing the term loadings to latent factors; 2) a document-by-factor matrix showing the document loadings to latent factors; and 3) a diagonal matrix containing scaling values in descending orders. We explored several solutions with different number of factors.

***Step 4.* Factor Rotations and Interpretation**

After dimension reduction, a factor analysis is typically applied for interpretive purposes. In this research, an orthogonal rotation method, Varimax, was applied to rotate the term-factor loading matrix and document-factor loading matrix to give more interpretable factor loadings on the solution. Finally a 14 factor solution appears most appropriate to capture most important factors of HIT research themes.

Appendix 2D: Citation and Co-Citation Matrix

To analyze the intellectual structure of the overall HI discipline across multiple disciplines, we aggregated the document-level citation and co-citation information to the discipline level. The Thomson Reuters Journal Citation Report (JCR) contains information of influence, impact, and subject relationships for leading journals. Subject categories of each journal in our dataset were retrieved from both the JCR for Social Science Citation Index (SSCI) 2012 and JCR for Science Citation Index (SCI) 2012 and treated as academic disciplines for the citation and co-citation analysis. In total, 34 disciplines were identified which had published HI research. Table 2.10 shows a subset of the raw discipline citation matrix. A subset of the lower-half raw discipline co-citation matrix is depicted in Table 2.11.

Table 2.10 Raw Discipline Citation Matrix (7 x 7 Subset)

	1	*2*	*3*	*4*	*5*	*6*	*7*
1. Computer Science	138	11	24	48	0	634	1
2. General and Internal Medicine	2	310	246	3	0	938	2
3. Health Care Sciences & Services	34	364	3,449	69	0	1,451	89
4. Information Systems	115	47	80	606	10	1,286	0
5. Management	4	1	7	11	4	29	0
6. Medical Informatics	619	1,355	1,408	727	4	39,419	20
7. Surgery	0	21	90	0	0	44	39

Table 2.11 Raw Discipline Co-Citation Matrix (7 x 7 Subset)

	1	*2*	*3*	*4*	*5*	*6*	*7*
1. Computer Science	402						
2. General and Internal Medicine	108	1,882					
3. Health Care Sciences & Services	370	2159	25,254				
4. Information Systems	347	224	621	2,038			
5. Management	2	11	30	26	10		
6. Medical Informatics	4,496	11,484	15,217	6,616	92	31,8758	
7. Surgery	3	64	620	3	0	193	134

The document-level citation and co-citation information can also be easily aggregated into author and research theme levels, thereby providing a more accurate measure for citation and co-citation analysis at higher levels than document-level analysis. This information aggregation provides more flexible and valid measures than traditional methods which rely on the first authors without the consideration of co-authorship (e.g., Culnan 1986; Culnan 1987; Ding et al. 1999; Pilkington and Meredith 2009). For HIT articles, we aggregated the document-level citation and co-citation matrix to an authorial level to examine thought leadership in the HIT sub-discipline. Table 2.12 shows a subset of the raw HIT author citation matrix. We noticed that some author names have multiple initials. For example, "Anderson, C." and "Anderson, C. L." represent the same author, and "Hu, P. J. H." sometime displays as "Hu, P. J.". For such case, we analyzed the data at a more detailed level and kept an identical scholar name if multiple initials represented the same scholar.

Table 2.12 Raw HIT Author Citation Matrix (7 x 7 Subset)

	1	*2*	*3*	*4*	*5*	*6*	*7*
1. Chau, P. Y. K.	1	0	0	1	0	0	0
2. Davidson, E. J.	0	4	1	0	3	3	3
3. Devaraj, S.	0	0	0	0	5	0	3
4. Hu, P. J. H.	6	1	6	2	0	0	0
5. Kohli, R.	0	0	1	0	4	0	0
6. Lapointe, L.	0	0	0	0	1	3	3
7. Rivard, S.	0	0	0	0	1	3	3

To investigate research themes in the HIT sub-discipline, we also aggregated the document-level citation/co-citation matrix into research theme levels for all HIT articles. Since the document-factor loadings represent the strength of the association between particular documents and factors, the weight for the research theme citation/co-citation matrix is defined according to formula (1):

$$w_{i,j} = \sum_{m,n} l_{m,i} \times l_{n,j} \tag{1}$$

where *i* and *j* are research themes, $l_{m,i}$ is the loading of document *m* on research theme i, $l_{n,j}$ is the loading of document *n* on research theme *j*, document *m* cites document *n* in document-level citation matrix or documents *m* and *n* are co-cited in document-level co-citation matrix. Table 2.13 and Table 2.14 show the HIT research theme level citation and co-citation matrices, respectively.

Table 2.13 HIT Research Theme Citation Matrix

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*
1. HIT and Organizations	1.934	0.950	0.445	1.366	0.431	0.409	0.476	0.425	0.324	0.102	0.133	–	–	–
2. HIT Innovation	0.554	2.039	1.267	1.255	0.207	0.116	0.387	0.427	0.125	0.415	0.232	1.025	–	–
3. General HIT Applications	0.274	0.396	1.655	0.512	–	0.050	0.234	0.100	0.513	–	–	0.201	–	–
4. TAM of HIT	0.525	0.395	1.617	8.996	0.160	0.055	–	1.510	–	–	–	0.267	–	–
5. Implications of HIT	1.815	0.827	–	1.125	0.437	0.171	0.121	0.476	–	–	–	–	–	–
6. Clinical Decision Support	0.111	0.050	0.287	0.298	0.402	0.735	0.210	0.093	0.362	–	–	–	–	–
7. EMR and EHR	0.130	0.101	0.345	0.145	0.211	–	0.090	0.033	0.166	0.103	–	–	–	–
8. Telemedicine	0.049	0.076	0.594	0.276	0.160	0.012	–	0.522	–	–	–	–	–	–
9. Security of HIT	0.140	–	0.582	–	–	–	0.281	–	1.252	–	–	–	–	–
10. National HIT Programs	0.206	0.272	–	–	–	–	0.159	0.139	–	0.833	–	0.169	–	–
11. Knowledge Management in Healthcare	0.132	0.206	0.122	0.462	–	0.045	0.073	0.074	–	0.228	0.204	0.290	–	–
12. Trust in HIT	0.093	0.183	0.357	0.419	0.477	–	–	0.138	–	–	–	2.378	–	–
13. Medical Information Retrieval	–	–	–	–	–	–	–	–	–	–	–	–	0.528	–
14. Medical Image Processing and Management	–	–	–	0.351	–	–	–	0.107	–	–	–	–	–	–

Table 2.14 HIT Research Theme Co-Citation Matrix

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
1. HIT and Organizations	7.027											
2. HIT Innovation	3.840	6.237										
3. EMR and EHR	0.855	0.468	0.387									
4. Knowledge Management in Healthcare	0.236	0.396	0.324	0.000								
5. Telemedicine	0.146	0.127	0.372	0.627	1.692							
6. TAM of HIT	7.254	3.872	1.466	1.065	2.826	21.822						
7. National HIT Program	1.321	3.652	0.123	0.000	0.546	1.434	4.979					
8. Clinical Decision Support	1.034	0.621	0.216	0.133	0.042	1.048	0.000	0.972				
9. Implication of HIT	2.194	2.379	0.778	0.000	0.383	3.357	0.166	0.729	1.127			
10. Trust in HIT	1.141	3.356	0.267	0.603	1.380	5.250	1.312	0.164	2.236	7.413		
11. General HIT application	1.875	2.556	1.168	1.299	1.194	2.735	0.376	0.507	1.112	3.418	4.424	
12. Security of HIT	0.140	0.125	0.678	0.234	0.312	0.698	0.000	0.135	0.984	0.417	1.910	0.801

Note: Research themes “Medical Information Retrieval” and “Medical Image Processing and Management” are not co-cited with any other theme, so that they are not listed in the co-citation matrix.

Appendix 2E: 14 Factors of HIT Research

Table 2.15 HIT Factors

Factor	Label	Top 30 Terms (Stemmed)
1	Security of HIT	secur, hipaa, comput, polici, issu, mobil, collabor, social, perceiv, perspect, privaci, behavior, medic, record, data, iso, healthcar, implement, represent, work, inform, efficaci, care, critic, commun, variou, control, self, protect, threat
2	Implications of HIT	hit, implic, technolog, inform, health, medic, strateg, usag, issu, healthcar, system, perform, resist, solv, routin, davidson, chiasson, data, adopt, co, cite, problem, agenc, clinic, hospit, cost, invest, research, measur, huge
3	Medical Information Retrieval	search, languag, engin, chines, web, non, cmedport, user, portal, retriev, modul, develop, session, approach, brows, project, domain, tool, build, multilingu, mesh, techniqu, speak, categor, benchmark, thesauri, issu, consum, research, medicin
4	Medical Image Processing and Management	imag, retriev, medic, tool, visual, learn, applic, softwar, radiologist, sourc, model, pac, radiolog, autom, featur, read, deform, fetch, process, practic, method, registr, evalu, implement, rank, regist, data, pre, compartment, transform
5	Trust in HIT	trust, collabor, infomediari, person, interperson, virtual, disposit, belief, health, consum, portal, trait, onlin, vc, type, commun, review, role, vcr, build, posit, model, measur, compet, opportun, perform, disclos, affect, web, individu
6	EMR and EHR	record, electron, medic, vista, ignor, implement, strategi, ehr, health, nation, data, respons, secur, issu, hidden, index, except, analyz, system, care, incent, emr, risk, patient, adopt, phr, match, share, articl, physician
7	Knowledge Management in Healthcare	knowledg, transfer, project, manag, clinic, virtual, medicin, learn, npd, collabor, flow, dkm, activ, integr, share, pathwai, process, barrier, nurs, hipp, develop, tacit, support, km, parti, internet, requir, case, articl, medic
8	TAM of HIT	accept, tam, technolog, model, physician, user, perceiv, usag, telemedicin, profession, us, individu, test, context, intent, eas, behavior, decis, research, resist, fit, explanatori, mobil, factor, construct, attitud, examin, support, evalu, explain
9	National HIT Programs	nation, servic, project, programm, nh, chang, govern, organis, year, institut, health, npfit, critic, trust, technolog, mobil, billion, uk, implement, invest, complex, manag, local, time, strategi, introduct, web, sector, reluct, period

10	General HIT Applications	care, health, inform, system, privaci, strateg, deliveri, technolog, advanc, commun, adapt, manag, vr, servic, enabl, design, centuri, asynchron, polici, challeng, interoper, warehous, person, patient, framework, provid, data, complex, onlin, develop
11	HIT Innovation	innov, implement, process, project, theori, adopt, context, actor, practic, organ, system, organiz, structur, integr, work, action, level, healthcar, group, research, analysi, develop, strategi, case, standard, design, conting, collabor, studi, institute
12	HIT and Organizations	hospit, adopt, privaci, physician, emr, patient, assimil, healthcar, ehr, cost, innov, learn, perform, influenc, practic, complianc, satisfact, organiz, person, technolog, effect, electron, invest, crm, inform, manag, usag, factor, impact, exchang
13	Clinical Decision Support	decis, data, medic, patient, cost, healthcar, treatment, problem, support, hospit, predict, comput, analyz, model, qualiti, optim, neural, diagnosi, clinic, provid, network, perform, error, accuraci, dss, servic, make, databas, mine, evalu
14	Telemedicine	telemedicin, practic, realiti, medicin, privaci, comput, learn, collabor, health, polici, healthcar, context, program, patient, telehealth, technolog, medic, physician, diagnosi, treatment, tele, virtual, person, countri, saharan, sub, human, profession, complianc, theori

Appendix 2F: High-Loading Papers for 14-Factor Solution

Table 2.16 High-Loading Papers for 14-Factor Solution

Factor	Label	High Loading Papers	Journal	Loading
1	Security of HIT	Ng et al., 2009	*Decision Support Systems*	0.722
		Stahl et al., 2012	*Information Systems Journal*	0.700
		Vaast, 2007	*Journal of Strategic Information Systems*	0.667
		Huston, 2001	*Communications of the ACM*	0.657
		Mercuri, 2004	*Communications of the ACM*	0.610
		Thomas & Botha, 2007	*Information Systems Management*	0.511
		He et al., 2012	*Information Systems Frontiers*	0.232
2	Implications of HIT	Goldschmidt, 2005	*Communications of the ACM*	0.777
		Agarwal et al., 2010	*Information Systems Research*	0.681
		Goh et al., 2011	*Information Systems Research*	0.678
		Bhattacherjee et al., 2007	*Information Systems Management*	0.642
		Zhang et al., 2009b	*European Journal of Information Systems*	0.555
		Romanow et al., 2012	*MIS Quarterly*	0.511
		Sheng, 2000	*Decision Support Systems*	0.460
		Bhattacherjee & Hikmet, 2007	*European Journal of Information Systems*	0.323
3	Medical Information Retrieval	Zhou et al., 2006	*Decision Support Systems*	0.742
		Chau et al., 2008	*Decision Support Systems*	0.711
		Chung et al., 2006	*Decision Support Systems*	0.702
		Lu et al., 2008	*Decision Support Systems*	0.442
		Houston et al., 2000	*Decision Support Systems*	0.344
		Wang et al., 2012	*Decision Support Systems*	0.311

4	Medical Image Processing and Management	Metaxas, 2005	*Communications of the ACM*	0.681
		Sheng et al., 2000	*Decision Support Systems*	0.593
		Yoo & Ackerman, 2005	*Communications of the ACM*	0.578
		Tang & Ip, 2009	*Information Systems Frontiers*	0.565
		Wong et al., 2009	*Information Systems Frontiers*	0.522
		Hu et al., 2006	*Decision Support Systems*	0.495
		Da Silva et al., 2011	*Decision Support Systems*	0.494
		Blum & Aboulafia, 2003	*Information Systems Frontiers*	0.315
		Law et al., 1995	*Information & Management*	0.216
5	Trust in HIT	Brown et al., 2004	*Journal of Management Information Systems*	0.683
		Zahedi & Song, 2008	*Journal of Management Information Systems*	0.673
		Paul & Mcdaniel, 2004	*MIS Quarterly*	0.635
		Song & Zahedi, 2007	*Decision Support Systems*	0.614
		Leimeister et al., 2005	*Journal of Management Information Systems*	0.516
		Luo & Najdawi, 2004	*Communications of the ACM*	0.440
		Bansal et al., 2010	*Decision Support Systems*	0.435
		Eason, 2007	*Journal of Information Technology*	0.317
		Randell, 2007	*Journal of Information Technology*	0.284
		He et al., 2012	*Information Systems Frontiers*	0.234
6	EMR and EHR	Hoffmann, 2009	*Communications of the ACM*	0.791
		Venkatraman et al., 2008	*Communications of the ACM*	0.699

		Bhaskar, 2010	*Communications of the ACM*	0.603
		Cantrill, 2010b	*Communications of the ACM*	0.590
		Ozdemir et al., 2011	*Information Systems Research*	0.427
		Huston, 2001	*Communications of the ACM*	0.421
		Charette, 2006	*Communications of the ACM*	0.394
		Bell & Sethi, 2001	*Communications of the ACM*	0.332
		Poston et al., 2007	*Information Systems Management*	0.272
7	Knowledge Management in Healthcare	Lin et al., 2008	*Information & Management*	0.573
		Leiter et al., 2007	*Human Relations*	0.521
		Pedersen & Larsen, 2001	*Decision Support Systems*	0.501
		Mohan et al., 2007	*Decision Support Systems*	0.495
		AlKaraghouli et al., 2013	*Information Systems Management*	0.458
		Paul, 2006	*Journal of Management Information Systems*	0.458
		RubensteinMontano et al., 2000	*Journal of Computer Information Systems*	0.440
		Ghosh & Scott, 2007	*Information Systems Management*	0.391
		Detmer & Shortliffe, 1997	*Communications of the ACM*	0.355
		Yang et al., 2012	*Information Systems Frontiers*	0.353
		Mitchell, 2006	*MIS Quarterly*	0.323
		Sheng et al., 2000	*Decision Support Systems*	0.290
		KamsuFoguem et al., 2012	*Decision Support Systems*	0.280
8	TAM of HIT	Hu et al., 1999	*Journal of Management Information Systems*	0.709
		Chau & Hu, 2002a	*Information & Management*	0.688

		Cantrill, 2010a	*Journal of Management Information Systems*	0.614
		Yi et al., 2006	*Information & Management*	0.581
		Moores, 2012	*Decision Support Systems*	0.496
		Hu et al., 2003	*Information & Management*	0.441
		Wang et al., 2006	*Information Systems Journal*	0.433
		Walter & Lopez, 2008	*Decision Support Systems*	0.421
		BurtonJones & Hubona, 2006	*Information & Management*	0.418
		Lai & Li, 2005	*Information & Management*	0.405
		Bhattacherjee & Hikmet, 2007	*European Journal of Information Systems*	0.395
		Bhattacherjee & Hikmet, 2008	*Journal of Computer Information Systems*	0.353
		Shih, 2004	*Information & Management*	0.353
		Deng et al., 2005	*Information & Management*	0.325
		Liu & Ma, 2005	*Information & Management*	0.280
		Wu et al., 2011	*Decision Support Systems*	0.267
		Pendharkar et al., 2001	*Journal of Computer Information Systems*	0.264
		Barki et al., 2008	*Journal of Information Technology*	0.261
9	National HIT Programs	Currie & Guah, 2006	*Information Systems Management*	0.583
		Currie & Guah, 2007	*Journal of Information Technology*	0.512
		Brennan, 2007	*Journal of Information Technology*	0.497
		Clegg & Shepherd, 2007	*Journal of Information Technology*	0.491
		Fernando et al., 2012	*Information Systems Frontiers*	0.465
		Mark, 2007	*Journal of Information Technology*	0.404
		Tan et al., 2009	*Journal of Computer Information Systems*	0.332

		Eason, 2007	*Journal of Information Technology*	0.307
		Mcgrath, 2002	*European Journal of Information Systems*	0.293
		Currie, 2012	*Journal of Information Technology*	0.281
		Aanestad & Jensen, 2011	*Journal of Strategic Information Systems*	0.280
		Gillies, 1995	*Journal of Information Technology*	0.273
		Wiredu & Sorensen, 2006	*European Journal of Information Systems*	0.270
10	General HIT Applications	Raghupathi, 1997	*Communications of the ACM*	0.804
		Thompson & Dean, 2009	*Communications of the ACM*	0.525
		Rindfleisch, 1997	*Communications of the ACM*	0.494
		Raghupathi & Tan, 2002	*Communications of the ACM*	0.486
		Tan et al., 2005	*Communications of the ACM*	0.405
		Berndt et al., 2003	*Decision Support Systems*	0.392
		Meiller et al., 2011	*Decision Support Systems*	0.369
		Smith & Bullers, 1999	*Journal of Computer Information Systems*	0.367
		Dutta & Heda, 2000	*Decision Support Systems*	0.332
		Agrawal et al., 2007	*Communications of the ACM*	0.322
		Johnson & Ambrose, 2006	*Communications of the ACM*	0.322
		Singh et al., 2011	*Journal of the Association for Information Systems*	0.317
		Wilson, 2003	*Communications of the ACM*	0.311
		Strickland, 1997	*Communications of the ACM*	0.299

		Mouttham et al., 2012	*Information Systems Frontiers*	0.294
		Zhou & Piramuthu, 2010	*Decision Support Systems*	0.285
		Pendharkar et al., 2001	*Journal of Computer Information Systems*	0.282
		Balka et al., 2012	*Information Systems Frontiers*	0.281
		Gianchandani, 2011	*Journal of Information Technology*	0.277
11	HIT Innovation	Igira, 2008	*Journal of Information Technology*	0.445
		Yetton et al., 1999	*Journal of Information Technology*	0.442
		Cho & Mathiassen, 2007	*European Journal of Information Systems*	0.410
		Mitchell & Zmud, 1999	*Organization Science*	0.374
		Kaganer et al., 2010	*Journal of the Association for Information Systems*	0.370
		Cho et al., 2007	*Journal of Information Technology*	0.355
		Jensen et al., 2009	*Journal of Information Technology*	0.335
		Leidner et al., 2010	*Journal of Strategic Information Systems*	0.322
		Cho et al., 2008	*European Journal of Information Systems*	0.315
		Braa et al., 2007	*MIS Quarterly*	0.314
		Lapointe & Rivard, 2007	*Organization Science*	0.313
		Hanseth et al., 2006	*MIS Quarterly*	0.299
		Sahay et al., 2009	*Journal of the Association for Information Systems*	0.289
		Hussain & Cornelius, 2009	*Information Systems Journal*	0.288
		Wainwright & Waring, 2007	*Journal of Information Technology*	0.284
		Fedorowicz & Gogan, 2010	*Information Systems Frontiers*	0.267

12	HIT and Organizations	Miller & Tucker, 2009	*Management Science*	0.416
		Reardon & Davidson, 2007	*European Journal of Information Systems*	0.376
		Hung et al., 2010	*Decision Support Systems*	0.333
		Chang et al., 2009	*Information & Management*	0.312
		Kohli et al., 2001	*Decision Support Systems*	0.312
		Angst et al., 2010	*Management Science*	0.294
		Mishra et al., 2012	*Information Systems Research*	0.293
		Leidner et al., 2010	*Journal of Strategic Information Systems*	0.289
		Angst & Agarwal, 2009	*MIS Quarterly*	0.283
		Angst et al., 2012	*Journal of Management Information Systems*	0.279
		Anderson & Agarwal, 2011	*Information Systems Research*	0.272
		Klein, 2007	*European Journal of Information Systems*	0.270
		Lee & Shim, 2007	*European Journal of Information Systems*	0.270
		Menon & Lee, 2000	*Decision Support Systems*	0.266
		Davidson & Heslinga, 2007	*Information Systems Management*	0.265
		Warkentin et al., 2011	*European Journal of Information Systems*	0.265
13	Clinical Decision Support	Poston et al., 2007	*Information Systems Management*	0.371
		Delen et al., 2012	*Decision Support Systems*	0.323
		Hu et al., 2007	*Decision Support Systems*	0.285
		Yeh et al., 2011	*Decision Support Systems*	0.261
		Forgionne & Kohli, 1996	*Decision Support Systems*	0.256
		Menon & Lee, 2000	*Decision Support Systems*	0.251
		Cao et al., 2012	*Decision Support Systems*	0.250
		Bielza et al., 2008	*Decision Support Systems*	0.248
		Mangiameli et al., 2004	*Decision Support Systems*	0.246

		Churilov et al., 2005	*Journal of Management Information Systems*	0.244
14	Telemedicine	Huston & Huston, 2000	*Communications of the ACM*	0.394
		Miscione, 2007	*MIS Quarterly*	0.338
		Chau & Hu, 2004	*Communications of the ACM*	0.300
		Tarakci et al., 2009	*Decision Support Systems*	0.299
		Tan et al., 2002	*Journal of Computer Information Systems*	0.258
		Mbarika, 2004	*Communications of the ACM*	0.239
		Fichman et al., 2011	*Information Systems Research*	0.235
		KlecunDabrowska & Cornford, 2000	*Information Systems Journal*	0.229
		Nicolini, 2007	*Human Relations*	0.226
		Kifle et al., 2006	*Information Systems Frontiers*	0.225

Appendix 2G: Citation Network Measures of Core HIT Research Themes

Table 2.17 Degree Centrality

Rank	Theme	Normalized Score
1	TAM of HIT	0.467
2	General HIT Applications	0.442
3	HIT and Organizations	0.263
4	Telemedicine	0.258
5	HIT Innovation	0.233
6	Implications of HIT	0.119
7	Trust in HIT	0.093
8	Security of HIT	0.080
9	EMR and EHR	0.078
10	National HIT Programs	0.038
11	Clinical Decision Support	0.037
12	Knowledge Management in Healthcare	0

Table 2.18 Information Centrality

Rank	Theme	Raw Score
1	TAM of HIT	1.115
2	HIT Innovation	1.104
3	HIT and Organizations	1.072
4	Implications of HIT	1.040
5	General HIT Applications	1.038
6	Telemedicine	0.948
7	Trust in HIT	0.848
8	Clinical Decision Support	0.632
9	Security of HIT	0.554
10	EMR and EHR	0.552
11	Knowledge Management in Healthcare	0.370
12	National HIT Programs	0.343

Table 2.19 Subnetwork Density

Rank	Theme	Normalized Score
1	TAM of HIT	0.058
2	Security of HIT	0.048
3	Trust in HIT	0.044
4	Telemedicine	0.021
4	HIT and Organizations	0.021
5	National HIT Programs	0.018
5	Implications of HIT	0.018
6	HIT Innovation	0.012
7	General HIT Applications	0.011
8	Clinical Decision Support	0.010
9	EMR and EHR	0.005
10	Knowledge Management in Healthcare	0.004

Appendix 2H: Summary of Author Productivity

Table 2.20 Summary of Author Productivity

Article	Frequency	Percent	Cumulative Percent
1	595	85.0	85.0
2	61	8.7	93.7
3	28	4.0	97.7
4	9	1.3	99.0
5	3	0.4	99.4
6	1	0.1	99.6
7	1	0.1	99.7
8	1	0.1	99.9
10	1	0.1	100.0
Total	700	100	

Appendix 2I: Definitions for Health Informatics and HIT

Informatics is defined as a "the discipline focused on the acquisition, storage, and use of information in a specific setting or domain" and is focused on "using technology to help people do cognitive tasks better" (Hersh 2009). When applied to the context of health (i.e., "health informatics"), many definitions abound. Table 2.21 summarizes the key definitions (in chronological order, newest first). To define health informatics in this paper, we adopt the definition from the National Library of Medicine (NLM) and National Institutes of Health (NIH) (the first definition in the table). We also mention other available definitions for completeness. It is valuable to note from these definitions that health informatics (and the comparable definitions for biomedical and medical informatics) spans multiple disciplines and knowledge areas.

Table 2.21 Health Informatics (and Closely Related) Definitions

Domain	**Definitions**	**Source**
Health Informatics	*"The interdisciplinary study of the design, development, adoption and application of IT-based innovations in healthcare services delivery, management and planning."*	*National Library of Medicine (NLM) and National Institutes of Health (NIH) (Procter 2009)*
Medical Informatics	"The field of information science concerned with the analysis, use and dissemination of medical data and information through the application of computers to various aspects of health care and medicine"	National Library of Medicine (NLM 2014)

Health Informatics	"The application of multidisciplinary sciences to transform (not just automate) the structure and behavior of health-related systems, organizations, and individuals (including patients, professionals, and support personnel) who interact to provide personalized care."	Brown et al. (2012, p. 2)
Biomedical Informatics	"The effective uses of biomedical data, information, and knowledge for scientific inquiry, problem solving, and decision making, driven by efforts to improve human health."	Kulikowski et al. (2012, p. 933)
Biomedical and Health Informatics	"Optimal use of information, often aided by the use of technology, to improve individual health, health care, public health, and biomedical research"	Hersh (2009)
Medical Informatics	"While many definitions of the field can be found, most share two characteristics: reference to health sciences, biomedicine, and the healing arts; and reference to the use of information management techniques and technologies in support of those pursuits."	Morris and McCain (1998, p. 448)
Medical Informatics	"Medical informatics is the field concerned with the cognitive, information processing, and communication tasks of medical practice, education, and research, including the information science and technology to support these tasks."	Greenes and Shortliffe (1990, p. 1115)
Medical Informatics	"[T]he hybrid child of medicine and those logical sciences that are suggested by computer technology."	Lincoln and Korpman (1980, p. 262)

We also briefly examined definitions for "Health Information Technology." We first acknowledge that Health Information Systems (HIS) is likely a more appropriate term than HIT in that HIT indicates a focus on technology rather than a more comprehensive view of people, processes, technology, and information. However, the field most frequently uses the term "HIT" to refer to both to the technology as well as to the more comprehensive view. We take the more comprehensive view, but use the term HIT in conformance with the more common use of this term. To define HIT in this paper, we adopt the definition put forth by the Office of the National Coordinator (ONC) (Table 2.22).

Table 2.22 Health Information Technology (HIT) Definitions

Domain	Definitions	Source
HIT	*"The application of information processing involving both computer hardware and software that deals with the storage, retrieval, sharing, and use of health care information, data, and knowledge for communication and decision making."*	*Office of the National Coordinator (ONC) for HIT (ONC 2014)*
HIT	"Health information technology (IT) encompasses a wide range of products and services—including software, hardware and infrastructure—designed to collect, store and exchange patient data throughout the clinical practice of medicine."	American Medical Association (AMA 2014)
HIT	"Term used to describe the application of computers and technology in health care settings."	Hersh (2009)
Clinical Information Systems	"Clinical information systems support patient care and provide information for use in strategic planning and management. Applications include computerized patient records systems; clinical department systems such as pharmacy, laboratory, and radiology; automated medical instrumentation; clinical decision support systems (computer-aided diagnosis and treatment planning); and information systems that support clinical research and education."	Glandon et al. (2008, p. 20)
HIT	"HIT consists of an enormously diverse set of technologies for transmitting and managing health information for use by consumers, providers, payers, insurers, and all other groups with an interest in health	Blumenthal and Glaser (2007, p. 2527)

	and health care."	
Health Information Systems	"The health information system provides the underpinnings for decision-making and has four key functions: data generation, compilation, analysis and synthesis, and communication and use. The health information system collects data from the health sector and other relevant sectors, analyses the data and ensures their overall quality, relevance and timeliness, and converts data into information for health-related decision-making."	World Health Organization (WHO 2008)
HIT	"The application of information processing involving both computer hardware and software that deals with the storage, retrieval, sharing, and use of health care information, data, and knowledge for communication and decision making."	Thompson and Brailer (2004, p. 38)

CHAPTER 3

HEALTH PROMOTION IN ONLINE HEALTH COMMUNITIES: EXPLAINING THE EFFECT OF SOCIAL SUPPORT ON HEALTH PROMOTION OUTCOMES

Abstract

Health consumers are increasingly using online health communities to exchange health-related social support between each other. As a result of these exchanges, health care consumers may be socially influenced by such virtual interactions in ways that affect individual health promotion outcomes. However, questions remain as to the effectiveness of online consumer-to-consumer social health support, particularly when such support is in the form of user-generated content and unstructured data. Thus, an emerging and interesting area of research is to comprehensively understand the relationship between social support provided and received in online health communities and individual members' health promotion outcomes. To further explain this relationship, the current study leverages a theoretically derived conceptual framework which integrates social capital theory and social support theory in the context of online health communities. This framework is applied in a quantitative field study and multiple analyses of a big online health community dataset. Methodologically, a computational multi-method approach, which combines natural language processing and machine learning techniques, is utilized to automate content analysis of big health digital data. Contributions of this research include: (1) confirming the advantages of being positioned at a high level of structural social capital for social support exchange in online health communities; (2) extending current understanding of the reciprocity mechanism of social support interaction in online health communities by unpacking the social interactions down to specific informational and emotional

support exchange; (3) presenting evidence on the mixed role of social support exchange in health promoting; and (4) shedding light on the design and management of online health communities.

Keywords: online health communities (OHCs), social support, social capital, health promotion, big data, automatic content analysis, natural language processing, machine learning, social network analysis

3.1 Introduction

Online health communities, which are social networks where people with common health interests can share experiences, post questions, and seek or provide emotional support (Eysenbach et al. 2004), are becoming a common source for health information seeking by health care consumers. A 2011 national survey conducted in the U.S. by the Pew Research Center's Internet & American Life Project found that 80% of U.S. Internet users have searched for health information online, 34% of Internet users have read others' commentary or experience about health issues online, and 18% have sought others with similar health concerns online (Fox 2011). A more recent national survey by the same project found that 72% of U.S. Internet users have looked online for health information within the past year (Fox and Duggan 2013). Another survey showed that social media sites are emerging as a potential source of online health information, with 42% Internet users consulting online rankings or reviews and 32% using social networking sites for health (Thackeray et al. 2013).

As an inseparable part of the move toward the so-called personalized preventative medicine (Swan 2012), online health communities are significantly changing the way patients treat and/or manage their own health. The core principle of personalized preventative medicine involves the empowerment of individuals to self-monitor and self-manage their health and wellness (Swan 2012). Online health communities offer various kinds of participation possibilities for individuals to self-manage their health with no limit of time and space. Specifically, participants can discuss conditions, symptoms, and treatments as well as seek and provide health-related advice and emotional support. As patients and consumers are beginning to use online health communities to exchange health-related social support, they may be socially influenced in ways that may impact their health.

When individuals are sharing their personal health information with other online community peers in this manner, they are "crowdsourcing" the collective wisdom of a huge number of community members (Eysenbach 2008). This can significantly lower the cost of health care and alleviate burdens on the health care system. Ultimately, online health communities open up new opportunities for the health care industry to obtain the "triple aim" (Berwick et al. 2008, p. 760) including: (1) cutting costs, (2) enhancing the individual experience of care, and (3) improving the health of entire populations. While previous research has investigated the impact of social support on health outcomes, such research has not fully explored the underlying nuanced mechanisms of such influence (Swan 2009; Thoits 2011) and often assumes a simple mechanism which explicates the influence of social interactions on individual's health (Zhu et al. 2013). Thus, questions remain as to the effectiveness of such communities and little empirical work has examined in detail the impact of the social support exchanged in these communities on individual health promotion, particularly given that much of this support is provided in the form of user-generated content and unstructured data. Therefore, motivated by this gap and need to revisit such assumptions, this paper seeks to provide a more comprehensive understanding of the dynamics of self-managed care enabled by online health communities as well as unpack the complicated micro-mechanisms embedded in the pathways from social interactions to health promotion. Given such purpose, the current research intends to tackle the following research question:

> ***RQ****: What is the effect of social support provisioning and consumption on individual health promotion outcomes in online health communities?*

The structure of this paper is as follows. First, we review the extant literature and set forth the theoretical background of the study. Then, the proposed research model and hypotheses are presented. We then test the proposed model empirically using a computational multi-method framework which combines various natural language processing and machine learning techniques applied toward empirically a big dataset (i.e., where "big" is defined as many observations as well as many potential variables) collected from nine online health communities. Lastly, we discuss how our study contributes to theory development and seeks to improve our understanding of social support exchange in online health communities and its impact on health promotion.

3.2 Literature Review

We first discuss the theoretical background that provides the basis for the proposed research model and key constructs within the proposed model. Then in section 3.3, we elaborate the underlying relationships between the variables, which lead to hypothesis development.

3.2.1. Social Support in Online Health Communities

The phenomenon of general social support has been extensively investigated for decades. Social support refers to the extent to which an individual's basic social needs, such as affection, esteem or approval, belonging, identity, and security are met through interaction with others (Kaplan et al. 1977; Thoits 1982). Social support, by its definition, is a multidimensional concept. Barrera (1986) suggests three perspectives of social support: (1) the social integration or embeddedness, which focuses on the social connections that an individual has to significant others in the social settings; (2) the perspective of perceived social support as the subjective cognitive appraisal of social support provided by others (Cohen and Wills 1985); and (3) the enacted or received support perspective, which characterizes social support as actions rendered by others to a focal individual to protect against the health consequences of stress, focusing on the objective aspects of social support (Cobb 1976). Compared with other two views, social embeddedness perspective provides a very indirect index of the social support functions and usually fails to illuminate the mechanism of the hypothesized influence of social support on stressful life events (Barrera 1986; Cohen and Wills 1985). Thus, in this research, we take the enacted support view, as the availability of large digital trace dataset in online health communities allows a more accurate account of received social support within a given time period than perceived social support (through self-reports) which relies on participants' retrospective evaluations (Barrera 1981; Barrera 1986; Scholz et al. 2013).

Although extant literature posits strong and consistently beneficial effects of perceived social support on physical and mental health, findings on received social support often find weak or contradictory effects (Haber et al. 2007; Nurullah 2012; Thoits 2011). Therefore, the methodological distinction between different perspectives on social support is important. Specifically, the current study is motivated by the curiosity about the exact effect of received social support in promoting health wellbeing.

Social support is now being studied empirically in the context of online services. Specifically, with the advent of Web 2.0, social media technologies such as social networking sites, wikis, forums and message boards, blogs, consumer reviews and opinions sites, and online support groups have emerged to support virtual social interactions for patients and caregivers. Consequently, research on online health communities is becoming one of the most interesting and vibrant research areas. Various studies have been conducted to address different research themes. Current efforts in social support under the setting of online health communities can be categorized into four research streams (see Table 3.1 for a summary). Specific findings are discussed in the following subsections.

Table 3.1 Research Streams of Social Support in Online Health Communities

RS#	Description of Research Stream	Unit of Analysis	Relevant Literature
1	Content analysis of social support exchange	Message/post	Blank et al. (2010); Chuang and Yang (2012); Coulson et al. (2007); Coursaris and Liu (2009); Huang et al. (2014); Loane and D'Alessandro (2013); Mo and Coulson (2008); Sillence (2013)
2	Social support reception and empowerment	Individual	Mo and Coulson (2012); Mo and Coulson (2014); Nambisan (2011); Yan and Tan (2014); Zhu et al. (2013)
3	Social support provisioning	Individual	Huang and Chengalur-Smith (2014); Huang et al. (2012)
4	Participation and commitment in online health communities	Individual or individual-period	Kordzadeh et al. (2014); McLaughlin et al. (2012); Wang et al. (2014); Wang et al. (2012)

3.2.1.1 Content Analysis of Social Support Exchange

The first research stream involves content analysis of social support exchanged online. This research stream has been extensively studied and there are mature content analysis methods for online user-generated content. For example, Loane and D'Alessandro (2013) investigated communication between participants with high levels of disability in an Amyotrophic Lateral Sclerosis (ALS) online community. Their results showed that high levels of social support evident in the ALS community include informational support, network support, and emotional support. Sillence (2013) analyzed messages in an online breast cancer support forum and found that major types of advice solicitation are through problem disclosure and requests for information and opinion.

Although there are different classification schemes of social support, the most widely accepted typology in the literature on online health communities was developed by Cutrona and Suhr (1992), a typology (refer to Appendix 3A for the detailed definition) that includes: (1) informational support (providing suggestion or advice on coping with the stress), (2) emotional support (communicating love, care, or empathy), (3) esteem support (communicating respect and confidence in abilities), (4) tangible support (providing or offering to provide goods or services), and (5) network support (affirming individuals' belonging to a group or persons with similar interests and concerns). Among the different types of social support, *informational support* and *emotional support* have been found to be the two most frequent types of social support exchanged online (Braithwaite et al. 1999; Coulson et al. 2007; Coursaris and Liu 2009; Gooden and Winefield 2007; Huang and Chengalur-Smith 2014; Mo and Coulson 2008). Tangible support is least frequently provided in the online community setting (Mo and Coulson 2008). In this research, we focus on informational support and emotional support exchanged within online health communities.

3.2.1.2 Social Support Reception and Empowerment

This second line of research concerns the effects of social support reception on health outcomes such as: (1) self-efficacy, psychological well-being, and functional well-being; (2) the benefits that online health community interactions can bring to the participants; and (3) how social support empowers patients and often leads to positive health outcomes. Berkman et al. (2000) suggest that the provisioning of social support is one of the pathways through which social relationships and affiliation can influence physical and mental health. Mo and Coulson (2012) propose that the use of online health communities was positively associated with occurrence of empowering processes for patients living with HIV/AIDS. A later study by Mo and Coulson (2014) identifies six empowering processes and six empowering outcomes including: increased optimism, emotional well-being, social well-being, being better informed, improved disease management, and feeling confident in relationships with physicians. Nambisan (2011) suggests that information seeking effectiveness rather than the social support affects patient's perceived empathy in online health communities which are run by healthcare organizations. An empirical study by Yan and Tan (2014) shows that informational support and emotional support given and received in online health communities have positive effects on patient's self-reported health functionality levels. Using structural equation modeling method, Zhu et al. (2013) suggest that perceived social support fully mediates the influence of social ties on subjective well-being. However, even though many studies have investigated the impact of social support on health outcomes, the underlying mechanisms of such influence has not yet been fully addressed (Swan 2009; Thoits 2011).

3.2.1.3 Social Support Provisioning

The third research stream addresses the provisioning of social support, particularly factors or antecedents that influence the provisioning of social support. Drawing from social capital theory, a seminal study by Huang and Chengalur-Smith (2014) explored the determinants of social support provisioning in healthcare virtual support communities. Their study demonstrated that an individual's provisioning of emotional support can be predicted by her/his extent of social interaction with other community members as well as her/his social identification within the online community, while the contribution of informational support can be determined by the provider's level of healthcare-related expertise. Although many extant studies on online health communities pay a great deal of attention to the health-promoting consequences of social support (as reviewed in the previous section 3.2.1.2), few research addresses the intricate micro-mechanisms of social support receipt and provisioning. In that this research stream has not been extensively addressed by extant literature, future (and current) research can delve deeper into this research theme by expanding the research scope and applying various theoretical perspectives and innovative methods.

3.2.1.4 User Participation and Commitment in Online Health Communities

The last research stream concentrates on the sustainability and effectiveness of online health communities, particularly in relation to the continued commitment of participants who are seeking health information and social support. This research theme is of significance, as attracting and maintaining user participation through voluntarily provided (and consumed) information and social support is one of the biggest challenges for the success of online health communities and, ultimately, patient engagement is a key component of improving health outcomes.

The effect of social support receipt on continued commitment to the community has been demonstrated by several empirical studies. For example, Wang et al. (2012) showed that emotional support receipt is negatively associated with the risk of participant dropout while informational support has a relatively weaker positive effect on commitment in online health communities. They argued that emotional support enhances member relationships with others or the online group as a whole, whereas informational support only gratifies an individual’s short-term information needs. Another empirical study by Wang et al. (2014) found the similar results. In addition, Wang et al. (2014) suggested that the level of user engagement in an online health community is related to not only social support but also companionship. McLaughlin et al. (2012) found that young adult cancer survivors participating in a social networking and video-sharing intervention program were more involved in the social networking intervention, particularly in situations characterized by weak social bonding with other cancer survivors and little social support from friends and family. Applying theories of group identity and theories of interpersonal bonds, Ren et al. (2012) argued that both identity-based and bond-based online community features enhance member attachment and participation. Additionally, Kordzadeh et al. (2014) suggested that short-term reciprocity exists in online health communities such that as more social support is received more active participation occurs, on average.

3.2.2. Social Capital Theory

Although there is no agreement on the definition of social capital in extant research, the concept of social capital generally refers to "resources embedded in a social structure which are accessed and/or mobilized in purposive actions (Lin 1999, p. 35)." Social capital is rooted in social relationships between individuals as well as individuals' connections with other peers in the community (Lin 1999; Putnam 1995). The principal proposition of social capital theory is that resources embedded in networks of relationships can facilitate collective action for mutual benefits (Woolcock 1998). According to Nahapiet and Ghoshal (1998), social capital can be conceptualized as three dimensions: (1) the structural dimension refers to the existence of social ties that facilitate social interaction; (2) the relational dimension is defined as social assets created and leveraged through relationships such as trust, norm of reciprocity, and identification; and (3) the cognitive dimension is manifested as shared vision and shared language, which represents resources providing shared representations, interpretations, and meanings among actors.

Social capital theory has been widely employed by information systems (IS) literature to explain knowledge sharing (Chiu et al. 2006; Wasko and Faraj 2005), social support contribution in online health communities (Huang and Chengalur-Smith 2014), open source project success

(Singh et al. 2011), and IS project control (Chua et al. 2012). In the setting of online communities, the role of social capital in current research is treated either as a dependent variable or an explanatory variable. For example, Ellison et al. (2007) took the first view to study the impact of online interactions in social network services on the formation and maintenance of social capital. On the other hand, Faraj et al. (2015) used social capital embedded in online interaction network to predict leadership in the online communities. As the purpose of this study is to explain social support exchanged online and its impact on individual health promotion, we take the second perspective to investigate how a participant's structural social capital affects his/her social support interaction with other members in online health communities. In short, we hypothesize that a high level of structural social capital in the online health community provides advantageous resources, thereby facilitating the receipt and provisioning of social support.

3.2.3. Health Promotion Outcomes

To assess potential improvements in the quality of healthcare, various quality and patient safety (QPS) metrics such as structure, outcome, process, and volume have been devised by the healthcare industry (Donabedian 1966; Lazar et al. 2013). As outcomes are the ultimate or acid test for effective healthcare (Lazar et al. 2013), health outcomes emerging from social interactions are of vital importance for meaningful research on online health communities. Typically high level outcome indicators include morbidity, recovery or restoration of function, and quality of life (Donabedian 2005; Lazar et al. 2013). However, there is an opportunity to examine intermediate level outcomes that may ultimately contribute to final outcomes such as morbidity and quality of life. With the definition of health and healthcare being extended to wellness maintenance and condition prevention rather than the single target of curing disease

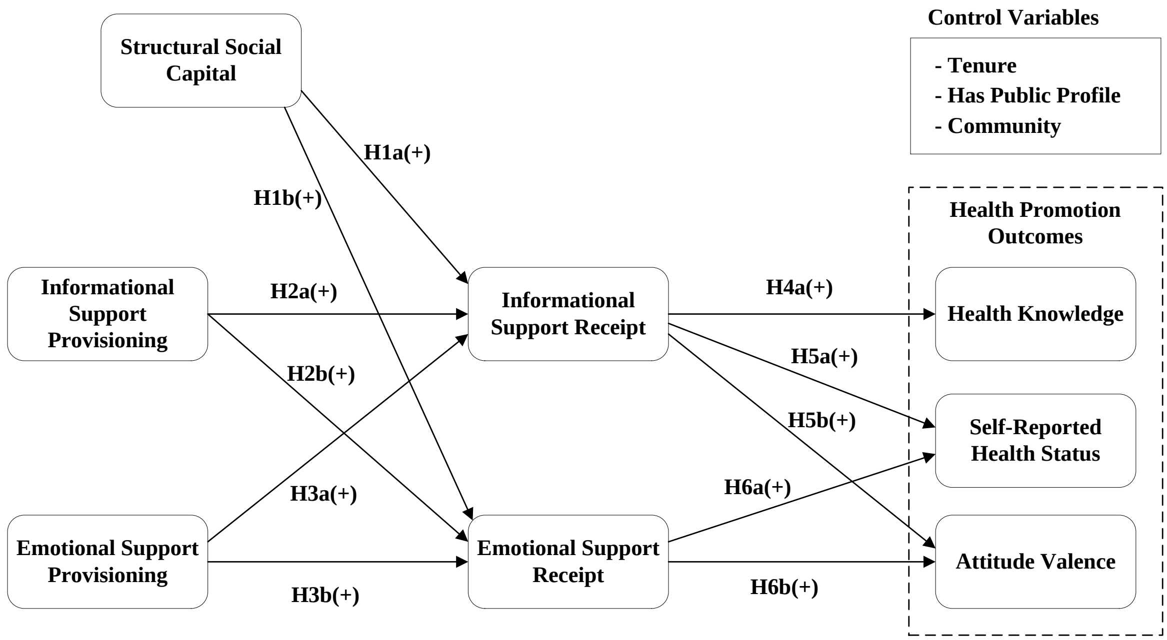
(Swan 2012), there are various kinds of supplementary outcome measures reported in extant literature (Eysenbach et al. 2004).

As defined in the Ottawa Charter for Health Promotion by World Health Organization, health promotion is the process of enhancing people's self-management and control over their health and thereby improve their health outcomes (World Health Organization 1986). In the setting of online health communities, information sharing as well as emotional support exchange facilitates participants to better engage in diagnosis, treatment, and self-management of diseases (Frost and Massagli 2008; Wicks et al. 2010). Thus, a comprehensive evaluation of the quality of health intervention through online health communities should also include health promotion outcomes such as changes in the individual's attitudes, knowledge, skills, confidence, and behaviors related to self-management of health (Fowles et al. 2009; Yoo and Bock 2014). Given the context of online health communities where people exchange social support to improve the self-management of health, it is appropriate to measure health promotion outcomes through attitudes towards health, health-related knowledge, and self-reported health status, especially when the bio-medical status of members is not directly accessible by analyzing online user-generated content. Thus, we analyze intermediate health promotion outcomes in this study, as a first step toward further understanding in this area.

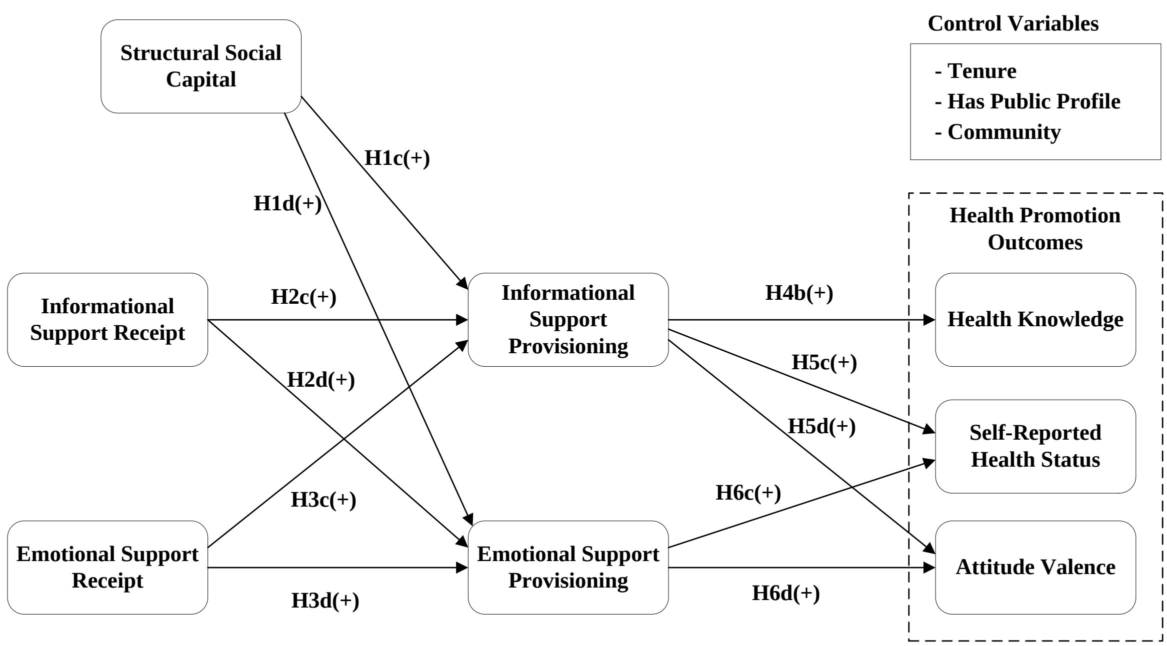
3.3 Research Model and Hypotheses

3.3.1 Research Model

This study seeks to explain variation in individual health promotion outcomes through the mechanism of social support receipt and provisioning in online health communities. To unpack the heterogeneity of social support interaction, we disaggregate social support into informational support and emotional support. As shown in Figure 3.1, the proposed research model integrates social capital and social support theories and provides a more comprehensive understanding of the dynamics of self-managed care enabled by online health communities. Essentially, structural social capital and the norm of reciprocity explain the degree of social support interchange while social support is used to explain health promotion outcomes. Specifically, in model specification A, structural social capital and the provisioning of informational and emotional support explain informational and emotional support receipt which further enhances health. In model specification B, structural social capital and informational and emotional support receipt explain the provisioning of informational and emotional support which further explains health promotion outcomes. The rational for the proposed research model is explained in the next section.



Model Specification A



Model Specification B

Figure 3.1 Research Model

3.3.2 Research Hypotheses

3.3.2.1 Relationships between Structural Social Capital and Social Support Exchange

Social capital theory is used in this study as the basis to explain social support exchange in online health communities. From the network perspective of social capital (Lin 1999), patterns of relationships define resources and social capital that are embedded in the network structure of social interaction. Structural social capital (SSC) refers to the potential resources embedded in the social interaction ties that individuals have access to by virtue of their network structural positions (Faraj et al. 2015; Thoits 2011). In the context of online health communities, structural social capital stands for the capability of participants to secure benefits by virtue of participation in the community, such as reading and posting messages in forums and locating people with similar interests or concerns.

As the building block of structural social capital, social interaction ties refer to the social connections that an individual has with others in the social setting through membership in groups (Thoits 2011). Social interaction ties are important for online health communities as they bond participants with common health interests together and provide access to resources (Nahapiet and Ghoshal 1998). The social interaction ties in online health communities provide an effective way for members to obtain and exchange health-related support resources. Granovetter (1973) distinguishes between two types of social ties, namely strong ties versus weak ties, where the strength of a dyadic tie depends on the amount of time spent interacting, the emotional intensity of the relation, the intimacy of the tie, and the reciprocal services provided to one another. Strong ties are formed by social relations with frequent contact, deep feelings of affection and obligation, and broad focus of domains; weak ties are relationships with infrequent contact, superficial and easily broken bonds, and narrow focus (Kraut et al. 1998).

This strong and weak ties distinction is similar to the difference between primary groups (e.g., family members, relatives, and friends) and secondary groups (e.g., work, voluntary, and religious organizations) (Thoits 2011). With different levels of social integration among members, strong ties and weak ties provide different types of supportive resources (Wellman and Wortley 1990). The power of weak ties is to provide innovative and non-redundant information and access to disparate networks (Granovetter 1973; Wellman et al. 2001). In contrast, the strength of strong ties lies in its capability to sustain commitment, friendship, and exchange of resources such as emotional aid and companionship (Kraut et al. 1998; Wellman et al. 2001; Wellman and Wortley 1990).

From the weak tie or brokerage view of social capital (Burt 1992), online health community members who bridge disconnected parts of interaction network have a competitive advantage in getting higher levels of returns directly toward themselves. From the strong tie or bonding view of social capital, trusting and cooperative relations between online health community participants account for the social support exchange. The higher degree of structural social capital obtained by a member in the online health community, the greater the intensity, frequency, and intimacy of the social relation there will be, thus granting the individual potential to obtain higher level of social support. Thus, we hypothesize that:

> ***H1a****: Participants' degree of structural social capital (SSC) will positively relate to their informational support receipt (ISR) in online health communities.*
>
> ***H1b****: Participants' degree of structural social capital (SSC) will positively relate to their emotional support receipt (ESR) in online health communities.*

Just as social interactions convey social support, structural social capital formed and sustained through social interactions between online health community peers should also explain why individuals provide various types of social support (Wellman and Wortley 1990). We term this “social support provisioning” in this thesis. The sustainability and effectiveness of online health communities depend on the continued commitment of participants who are seeking health information and social support. Social support exchange among peers in online health communities is a common objective for all participants. According to theories of collective action (Marwell and Oliver 1993; Olson 2009), participants of online health communities tend to contribute to collective benefits through voluntarily providing information and social support rather than free ride. Community members with high levels of structural social capital are more likely to initiate and sustain collective action through active collaboration such as knowledge contribution (Wasko and Faraj 2005). Such participants, due to their centrally embedded positions in the online interaction network and the resulting high demands from other members, are therefore more likely to contribute social support to other peers. Recent studies (e.g., Hwang et al. 2014; Hwang et al. 2010) empirically support the linkage between social capital and the provisioning of informational and emotional support. Given different levels of the structural social capital that participants hold in online health communities, the degree of social support contribution will vary. Consequently, we expect:

> ***H1c****: Participants’ degree of structural social capital (SSC) will positively relate to their informational support provisioning (ISP) in online health communities.*

> ***H1d****: Participants’ degree of structural social capital (SSC) will positively relate to their emotional support provisioning (ESP) in online health communities.*

3.3.2.2 Relationships between Social Support Receipt and Social Support Provisioning

A cornerstone of social relations is the norm of reciprocity, which refers to the universal social rule that forces us to repay others for what we have obtained from them to sustain ongoing exchange (Gouldner 1960). Different with the perspective of social dilemmas which posit that participants tend to get from the community rather than give to it, reciprocity concerns with individuals’ behaviors of both giving and rewarding in a community that is formed based on shared understandings, rules, as well as conventions on continuing social interactions (Preece 2001; Yang et al. 2009). From the perspective of social exchange theory (Blau 1964), individuals participate in social interactions based on the expectation that their efforts will be reciprocated with social rewards. As the major purpose of participants joining online health communities is to receive social support (Hajli et al. 2014), obtaining social support from others is what participants expect as a reward. Bowling et al. (2005) showed the existence of reciprocity in social support exchange. Their empirical study demonstrated the positive correlation between provisioning and receipt of social support in the workplace setting.

Thus, in the online health community setting, the norm of reciprocity works as a catalyst for both social support provisioning and receipt. Given a strong norm of reciprocity in online health communities, individuals trust that their social support provisioning efforts will be reciprocated, thus encouraging them to provide social support to others and stimulating more social support from others as a result. Given this study's focus on informational support and emotional support exchanged within online health communities, we hypothesize the following relationships among the provisioning and receipt of informational as well as emotional support:

> ***H2a****: Informational support provisioning (ISP) in online health communities will be positively associated with informational support receipt (ISR).*
>
> ***H2b****: Informational support provisioning (ISP) in online health communities will be positively associated with emotional support receipt (ESR).*
>
> ***H2c****: Informational support receipt (ISR) in online health communities will be positively associated with informational support provisioning (ISP).*
>
> ***H2d****: Informational support receipt (ISR) in online health communities will be positively associated with emotional support provisioning (ESP).*
>
> ***H3a****: Emotional support provisioning (ESP) in online health communities will be positively associated with informational support receipt (ISR).*
>
> ***H3b****: Emotional support provisioning (ESP) in online health communities will be positively associated with emotional support receipt (ESR).*

__H3c__: Emotional support receipt (ESR) in online health communities will be positively associated with informational support provisioning (ISP).

__H3d__: Emotional support receipt (ESR) in online health communities will be positively associated with emotional support provisioning (ESP).

3.3.2.3 The Relationship between Social Support and Health Promotion Outcomes

Various perspectives can be drawn on to explain the health-promoting function of social support. The perspective of supportive actions posits that received support enhances coping, which buffers the harmful impacts of stressors on health (Lakey and Cohen 2000). From the perspective of analogical behavioral processes, social support facilitates healthy behaviors such as exercising, eating right, quitting smoking, and actively engaging in medical regimens (Uchino 2006). Cohen (2004) suggests that stress buffering is the primary mechanism explicating the effect of social support in promoting health. According to the stress buffer theory, social support not only bolsters one's perceived ability to cope with stressful events, but also alleviates the impact of stress by provisioning of solutions to specific problems (Cohen 2004). In this study, we focus on three health promotion outcomes including: (1) health knowledge, (2) self-reported health status, and (3) attitude valence. Thus, it is hypothesized that social support exchanged in online health communities will positively influence each of these health promotion outcomes.

Sharing information about health conditions and treatments is one important aspect of the online health community discourse. Being better informed about health self-management, patients or consumers sharing information within online communities can clearly benefit from the process (Frost and Massagli 2008). The motivation of information support seekers is different with participants who want to obtain emotional support from online health communities in that information support is oriented to problem solving (Cutrona and Russell 1990). As a platform for health crowdsourcing, online health communities can aggregate distributed health-related information together, thus empowering patients with more knowledge and confidence in self-management of health and stress. Through the exchange of informational support in the online health community, individuals get more information and knowledge on their health conditions and available treatment options. Hence, we expect that:

> ***H4a**: Participants' informational support receipt (ISR) in online health communities will positively relate to their health knowledge (HK).*

> ***H4b**: Participants' informational support provisioning (ISP) in online health communities will positively relate to their health knowledge (HK).*

Besides the level of health knowledge that a participant learns from and exchanges with the online health community, self-reported health status (SHS) and attitude valence (AV) are two important health promotion outcomes. Various empirical studies provide evidence that self-reported health status is an important predictor of mortality (e.g., Idler and Benyamini 1997; Miilunpalo et al. 1997; Mossey and Shapiro 1982). Applying the theories of reasoned action and planned behavior (Madden et al. 1992), attitudes towards health should produce behavioral intentions that subsequently determine health behavior. In online health communities, distributed health-related information, experience, and emotional supportive resources are aggregated to effectively satisfy the needs of participants, thereby nurturing their self-reported health status and attitude towards self-management of health and stress. While informational support satisfies relatively short-term information needs of online community participants, emotional support meets their relatively long-term affective needs such as love, caring, sympathy, and encouragement (Thoits 2011). From the perspective of optimal matching theory (Cutrona and Russell 1990), the relative importance of informational and emotional support is moderated by the controllability of the stressors that the individual encounters. Optimal matching theory suggests that emotional support provides more effective health promotion under an uncontrollable stressor while informational support is more important in enhancing health outcomes if the individual has relatively more control on the stressor (Cutrona and Russell

1990). Based on above argument, we propose that both informational and emotional support obtained through online health community interactions benefit participants in terms of empowering their health self-management by promoting the level of their self-reported health status and attitude valence towards health. Specifically, we hypothesize:

> ***H5a**: Participants' informational support receipt (ISR) in online health communities will positively relate to their self-reported health status (SHS).*

> ***H5b**: Participants' informational support receipt (ISR) in online health communities will positively relate to their attitude valence (AV).*

> ***H6a**: Participants' emotional support receipt (ESR) in online health communities will positively relate to their self-reported health status (SHS).*

> ***H6b**: Participants' emotional support receipt (ESR) in online health communities will positively relate to their attitude valence (AV).*

In online health communities, all members are encouraged to participant in the peer-to-peer social interaction. Given the informational and emotional social support exchanged in this setting are both provided for and given by community peers, the effect of social support exchange on health promotion is not only through receipt of social support but also via the provisioning of such support. Although the receipt of social support from other peers promotes a participant's health, a higher level of involvement in providing social support to others makes it easier for this individual to assimilate and internalize social support received from others. Hence, we expect:

> ***H5c**: Participants' informational support provisioning (ISP) in online health communities will positively relate to their self-reported health status (SHS).*

> ***H5d**: Participants' informational support provisioning (ISP) in online health communities will positively relate to their attitude valence (AV).*

> ***H6c**: Participants' emotional support provisioning (ESP) in online health communities will positively relate to their self-reported health status (SHS).*

> ***H6d**: Participants' emotional support provisioning (ESP) in online health communities will positively relate to their attitude valence (AV).*

3.3.2.4 Control Variables

To more fully account for the unobserved heterogeneity, three control variables are included in the research model. These sets of control variables include: (1) tenure in the community (Huang and Chengalur-Smith 2014), (2) whether a member has a public profile (Wang et al. 2012), and (3) the community to which a member belongs.

3.4 Research Method

Given the explanatory nature of this study, we conducted a quantitative field study on online health communities to empirically test the proposed model. Previewing how we integrate text mining techniques with a general quantitative research approach, Figure 3.2 presents the overall research method. The detailed methods are explained in the following sections.

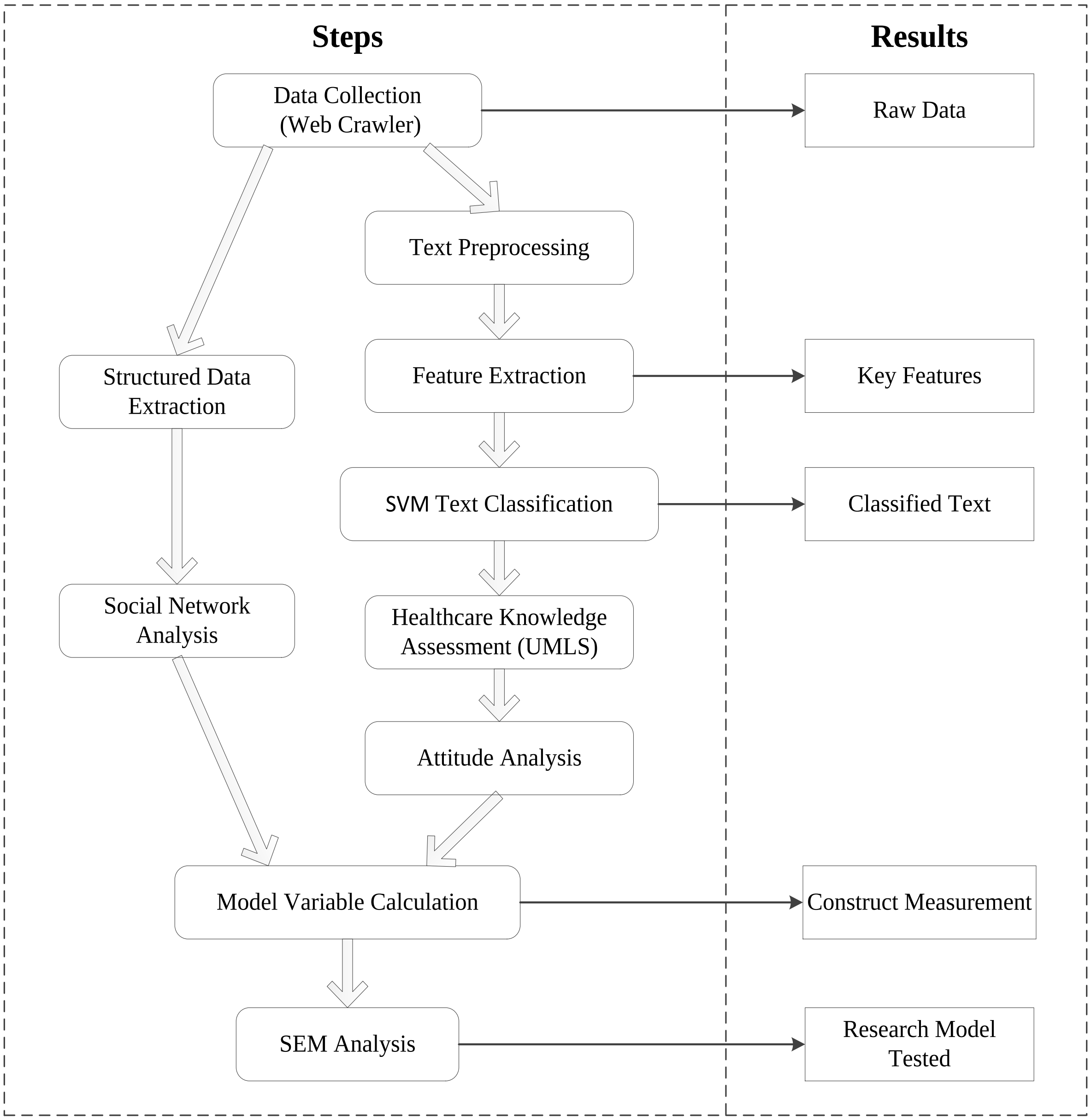


Figure 3.2 Research Method

3.4.1 Data Collection

3.4.1.1 Data Source

Data were collected from a large online health community. To obtain representative samples, we selected 9 forums hosted in the United States covering various kinds of health conditions including: (1) general conditions (chronic pain and obesity), (2) behavioral conditions (depression, anxiety, alcoholism, physical & emotional abuse, and insomnia), and (3) specific diseases (type 2 diabetes and HIV). An Internet crawler program was used to extract user-generated content from the online health community.

In total, we obtained 238,617 online discussion threads containing 2,305,288 posts generated by 32,405 members. A thread is a group of messages discussing a question or topic initiated by a member, while a post or response is a message by another member replying to the initial message. These messages were posted during the 8 years from July 2006 to November 2014. Appendix 3B presents some summary statistics of the data. About 87% of the responses were submitted within 24 hours after the thread initiation (refer to Appendix 3B Figure 3.8).

3.4.1.2 Ethical Considerations and IRB Review

Potential invasion of personal privacy in this research is expected to be minimal. Researchers do not have any direct interaction or intervention with users in the online community. The target online community is a public space and all the personal posts can be searched through search engines such as google.com. The object of our analysis is the communication patterns in the online community rather than how individual personalities interact. To ensure that no highly unlikely harm could come to subjects, we also “de-identified” the data collected by removing any names from the online user profile. Since the user-generated content in this online community was publically accessible, informed consent from members was not considered to be necessary (Flicker et al. 2004). Georgia State University institutional review board (IRB) approval for an exempt study was received.

3.4.2 Measurement

Two major types of construct development are reflective and formative measurement models. While reflective constructs assume that each measure is a reflection of the underlying latent construct (MacCallum and Browne 1993), formative constructs are conceptualized as composite of multiple indicators, with each item capturing a specific aspect of the construct. Literature suggests that formative constructs have been misspecified as reflective in research disciplines such as marketing (Jarvis et al. 2003) and information systems (Petter et al. 2007).

The distinction between formative and reflective constructs is critical for any empirical study in that misspecification of the measurement models may lead to Type I and Type II statistical errors (Jarvis et al. 2003; Petter et al. 2007). Criteria suggested to distinguish formative constructs from reflective constructs (Jarvis et al. 2003) include: (1) the direction of causality for formative constructs is from indicators to the construct; (2) indicators do not need to be interchangeable and co-vary with each other; (3) dropping an indicator may significantly change the conceptual domain of the construct. Table 3.2 presents the operationalized definition of constructs being explored as well as their measurement items and analytical methods used to extract them.

Table 3.2 Constructs and Measurements

Constructs/ Variables	**Definitions**	**Measures**	**Analytical Methods**
Structural Social Capital (SSC)	The potential resources embedded in the social interaction ties that individuals have access to by virtue of their network structural positions.	Five network measures are: ▪ SSC1: Betweenness ▪ SSC2: Closeness ▪ SSC3: In-degree ▪ SSC4: Out-degree	Social Network Analysis
Informational Support Receipt (ISR)	The amount of informational support *received* from other community members.	The total number of informational support messages provided by other members to the member.	SVM Text Classification
Emotional Support Receipt (ESR)	The amount of emotional support *received* from other community members.	The total number of emotional support messages provided by other members to the member.	
Informational Support Provisioning (ISP)	The amount of informational support *provided* to other community members.	The total number of informational support messages provided by the member to other members.	
Emotional Support Provisioning (ESP)	The amount of emotional support *provided* to other community members.	The total number of emotional support messages provided by the member to other members.	
Health Knowledge (HK)	The extent to which health professional knowledge is embedded in informational support provisioning.	The average number of UMLS terms used in the member's informational support posts.	UMLS Term Identification
Self-Reported Health Status (SHS)	The health status self-reported by the member.	Possible values include: horrible (1), bad (2), OK (3), good (4), and excellent (5).	Descriptive Statistics
Attitude Valence (AV)	The direction and strength of attitude expressed in the member's posts.	The average of attitude valence score expressed in the member's posts.	Sentiment Analysis

The degree of individuals' centrality in the interaction network is used to measure their structural social capital (Wasko and Faraj 2005). The higher degree of network centrality a member has in online interaction, the greater the intensity, frequency, and intimacy of the social relation there will be, thus providing different resources for members to obtain and exchange health-related social support. Specifically, social network measures for structural social capital include betweenness centrality, closeness centrality, in-degree centrality, and out-degree centrality. Betweenness indicates the extent to which a participant is in the middle of the communication between members in the community (Faraj et al. 2015; Shaw et al. 2005). From the perspective of bonding social capital or strong ties, closeness and degree centrality are used to measure a community member's capability to sustain commitment, friendship, and exchange of resources (Kraut et al. 1998; Wellman et al. 2001; Wellman and Wortley 1990).

These social network indicators are supposed to contribute to the structural social capital construct. As these indicators increase or decrease in magnitude, structural social capital also increases or decreases in magnitude. In contrast, an increase or decrease in the structural social capital does not necessarily lead to an increase or decrease of betweenness, closeness, in-degree, and out-degree simultaneously. Thus, the structural social capital construct is identified as a formative measurement model. All the formative indicators jointly determine the conceptual as well as empirical meaning of the structural social capital construct (Jarvis et al. 2003).

Other constructs are operationalized as single-indicator constructs. The detailed calculation procedures of all constructs are explained in the following sections.

3.4.3 Analysis of Digital Trace Data

In the current era of "Big Data," data generated from Web 2.0, social media, mobile devices, and ubiquitous sensors have been experiencing an exponential growth in terms of volume, velocity, and variety (Russom 2011). The rise of health social networks such as PatientsLikeMe, DailyStrength, and MedHelp provides unique opportunities for research focusing on healthcare decision support and patient empowerment (Miller 2012). User-generated content within these online communities are accessible not only to the patients and caregivers but also researchers. Specifically, digital trace data from the online communities are available for scholars to address more complex research questions than in the past.

Digital trace data have been suggested as a novel data source for IS scholarly efforts that address contemporary activities and behaviors (Hedman et al. 2013; Takeda et al. 2013). Howison et al. (2011) define digital trace data as "records of activity (trace data) undertaken through an online information system (thus, digital) (p. 769)." A trace represents an event occurring in the past that has been recorded by the information system, such as information a consumer posts about his/her prior health experiences. The rise of online health communities brings vast amount of digital trace data that can be used by researchers to address more complex research questions than in the past. Compared with traditional datasets collected through experiments, survey, or interviews, digital trace data hold three general characteristics: (1) the data is found rather than produced for research purposes; (2) the raw data is event-based with details at activity level; and (3) the data is longitudinal in nature (Howison et al. 2011). Following proper and rigorous procedures, digital trace data can be used to measure theoretically interesting constructs (Howison et al. 2011).

Given these characteristics, digital trace data are suitable for research on online communities (Johnson et al. 2014). With abundant digital trace big data being generated by online health communities, scholars are able to obtain insights into highly detailed, contextualized, and rich contexts, thereby obtaining insights that address the heterogeneous needs of individual patients. However, there is a lack of research in IS field that empirically addresses social relations within online health communities and its underlying theoretical relationships via analyses of big health data.

The current study represents a step toward obtaining insights into highly detailed, contextualized, and rich contexts from online health digital data. The task for this study is to map the digital trace data recorded in online health communities into measures of theoretically interesting constructs by following proper and rigorous procedures (Howison et al. 2011).

Some digital trace data in our target online health community are structured, such as the number of responses in a discussion thread. However, the messages posted in the online health community are textual and thus ill-structured. In this study, we apply a computational multi-method approach (Gaskin et al. 2014) which combines various natural language processing and machine learning techniques to process the digital trace data to extract measures for theoretical constructs represented in the proposed research model.

3.4.4 Social Network Analysis

To obtain social network measures, a directed network was constructed based on post-response relationships. The network also considers the strength of each tie between two community members. Figure 3.3 shows an example of the social network in online health community. As the example shows, Ted gets 6 replies from Ross and 12 replies from Mike, while Ross receives 4 responses from Ted, 7 from Daisy, and 5 from Anne. After the directed and weighted network is constructed, the focal social network metrics can be easily calculated via social network analysis software tools. The social network analysis package Pajek (De Nooy et al. 2011) was chosen with the consideration of its capability in analyzing large networks. In this study, the online health community request-response network contains 54,192 individual actors with 1,908,005 ties.

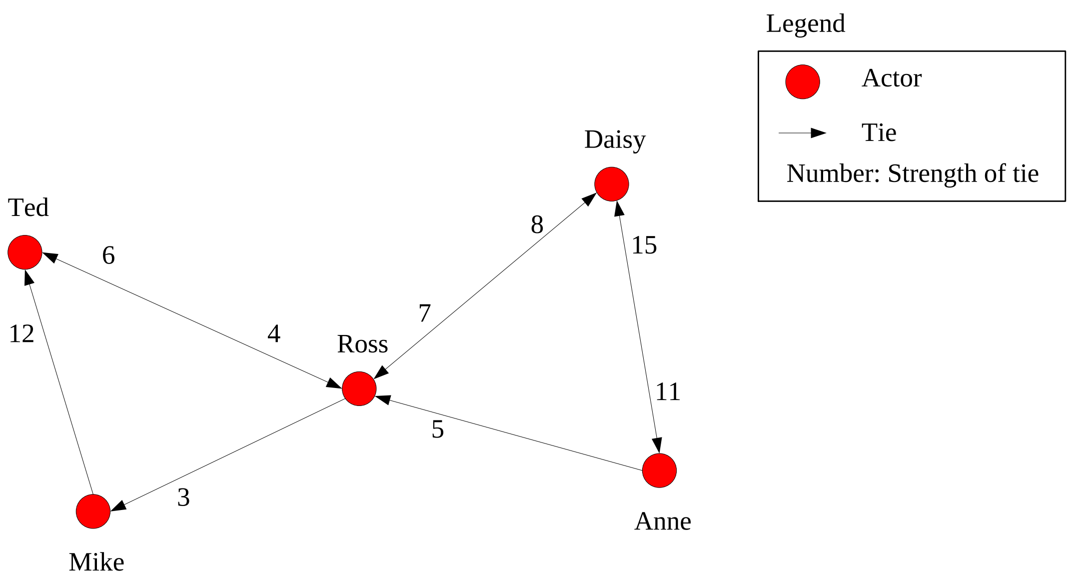


Figure 3.3 An Example of the Social Network in Online Health Community

Centrality refers to the extent to which an online health community participant connects to the interaction network. In this study, we used degree centrality, closeness centrality, and betweenness centrality to capture the network characteristics of online participants. Degree and closeness centrality measure the reachability of a participant with the network. In-degree refers to the number of incoming interactions for a participant. Out-degree is the number of outgoing interactions for a participant. Closeness centrality measures the extent to which a participant can reach other peers quickly. Closeness is generally calculated as the inverse of farness which is the sum of distances to other actors in the network (Freeman 1979). Newman (2001) extended the general calculation logic to handle a weighted network by transforming the weights of the network as costs and then applying Dijkstra's algorithm (Dijkstra 1959) to find the shortest path between two nodes. From another perspective, betweenness measures the centrality based on the idea that a person posits at a more central position if he or she is more important in intermediating communication for others. Betweenness is defined as the share of times a participant resides on the shortest path between other two individuals (Freeman 1979). The higher the betweenness for a participant, the more this participant can exploit the advantage of brokerage.

3.4.5 Content Analysis

Content analysis refers to "a research technique that makes replicable and valid inference from texts (or other meaningful matter) to the contexts of their use (Krippendorff 2004, p. 18)." Content analysis provides an unobtrusive way for researchers to gather information. Most previous research on online health communities employs a manual content analysis approach, whereby researchers read through the online messages and manually assign categories to them. Such manual approaches significantly reduce the scale of this type of research. Although some recent literature (e.g., Huang and Chengalur-Smith 2014; Huang et al. 2014; Wang et al. 2012) utilizes text mining algorithms to automate part of the content analysis work, thereby increasing scale, automatic content analysis in past scholarship has been severely limited in terms of both scope and depth. Thus, we seek to provide an analysis that is both of greater scale as well as more granularly scoped.

The unit of analysis for this study is at the individual level. Automatic content analyses of social support, attitudes, and the degree of healthcare knowledge expressed at the message level were aggregated to individual level to calculate the indicators for all focal constructs in the proposed research model.

3.4.5.1 Manual Coding of Social Support

To guide the content analysis of the online health messages, we used the Social Support Behavior Code (SSBC) developed by Cutrona and Suhr (1992) to code the social support for 3,083 replies randomly chosen from the dataset (refer to Appendix 3B Table 3.9 for the detailed definition of SSBC). This typology of social support is thought to be ideal for content analysis of online messages as it does not require the access to full range of nonverbal cues for the identification of social support (Braithwaite et al. 1999). Explanation with examples of social support provided by Mo and Coulson (2008) were also consulted. Noting that many messages indicate more than one type of social support, we followed the rule used by Loane and D'Alessandro (2013) to allow multiple social support types to be assigned to a single post.

To validate the applicability of the coding scheme, two coders independently assessed 1,000 replies for the types of social supported provided. The Cohen's Kappa was 0.87, indicating satisfactory inter-rater reliability (Straub et al. 2004). Then the first coder manually coded the left 2,083 messages. Precisely 1,387 replies among the 3,083 messages contain social support. Table 3.3 summarizes the frequency and percentage of different social support with examples. The finding shows that 91.1% social support exchanged in the online health communities are informational and emotional support.

Table 3.3 Summary and Examples of Social Support Coding

Social Support	Frequency (Percent)	Example
Informational Support	662 (44.1%)	"I have been on insulin for many, many years. I wish I was still on pills because I think if it's controlling it, I would stick with it. Insulin should be a last resort I think. But, of course, you should probably ask your doctor about it."
Emotional Support	706 (47.0%)	"Sorry to hear you are having a hard time with the meds. I would be apprehensive too. Anyway, I'm new to all of this, so I'm not sure what to say. I just wanted you to know that I heard you."
Network Support	85 (5.7%)	"Welcome, Joe! You'll find that people are really nice and supporting in this group. We're all here for each other, and now for you, too. On my worst pain days, I can always come here and feel better. People here really understand."
Esteem Support	46 (3.1%)	"You don't have anything to feel guilty about. I don't go out on weekends and if anyone asks i always say that i stayed home and kept busy. theres nothing wrong with staying home."
Tangible Support	2 (0.1%)	"...Have you ever written a gratitude list? Focusing on what you do have, the people who do care about you? It's a great way to lift your spirits a little bit. I sent you something in the mail the other day. You should get it today or tomorrow. :)"

3.4.5.2 Classification of Social Support

Given our aim to analyze the big data associated with online communities, automatic content analysis is the most tractable, efficient, and effective way to code our large dataset. This study applies text mining approaches to build classifiers for informational and emotional support respectively. The manually coded 3,086 replies were used as a training pool to train the automatic text classifiers which are based on support vector machine (SVM) model, a widely used text classification technique. A 10-fold cross-validation shows that the classification accuracy is 87.4% for the informational support classifier and 84.0% for the emotional support classifier. Then the training classifiers were used to automatically code the rest of the online community posts. The classification results were used to calculate the amount of social support that a participant provided to and received from other community members. The SVM-based automatic qualitative content analysis has been shown to provide results comparable to those concluded from traditional manual content analysis (Huang et al. 2010; Wang et al. 2012). The detailed procedure of social support classification is explained in Appendix 3C.

3.4.5.3 Health Knowledge Assessment

Following the method used by Huang and Chengalur-Smith (2014), we employed the count of the unified medical language system (UMLS) terms presented in informational support messages to assess the level of individual's health related knowledge. UMLS is a repository of biomedical and health-related terminologies developed by the US National Library of Medicine (NLM) (Bodenreider 2004). UMLS provides a representation of health-related knowledge in the UMLS semantic network. We used the Java API (application programming interface) of MetaMap[17], a software tool that maps text to concepts in the UMLS ontology, to identify UMLS terms from online health community posts. The mean number of UMLS terms used in a participant's informational support posts represents his/her health knowledge in the provisioning of informational support (Huang and Chengalur-Smith 2014).

3.4.5.4 Attitude Analysis

Opinion mining techniques were used to classify individuals' attitudes expressed in user-generated content in the online health communities. Opinion mining, a sub-discipline within

[17] MetaMap is available at https://metamap.nlm.nih.gov

data mining and computational linguistics, is the field of study that uses computational techniques to extract, classify, understand, and assess the opinions towards entities such as products, services, organizations, individuals, issues, events, and topics (Lim et al. 2013). With the explosion of text information written in natural languages, opinion mining has attracted the attention of many scholars in information systems (IS) and other disciplines such as computer science and linguistics. Sentiment analysis has been widely used in opinion mining to identify people's sentiments, evaluations, appraisals, attitudes, and emotions in settings such as social network sites (Agarwal et al. 2011), blogs (Melville et al. 2009), and online communities (Li and Wu 2010).

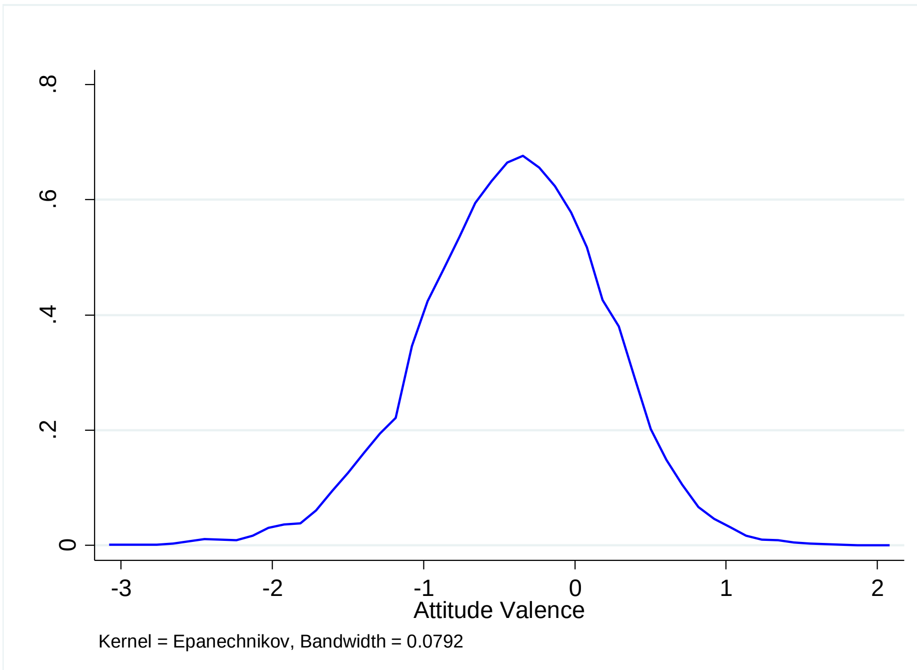
In this study, we used the tool SentiStrength[18](Thelwall et al. 2012; Thelwall et al. 2010) to measure the strength of positive and negative attitude expressed within online health community posts. The algorithm of SentiStrength has been demonstrated to provide better performance than a wide range of general machine learning approaches (Thelwall et al. 2010). SentiStrength allocate texts a positive attitude strength on a scale of 1 (no positive attitude) to 5 (very strong positive attitude) and a negative attitude strength on a scale of -1 (no negative attitude) to -5 (very strong negative attitude). Each message in our dataset is given both a positive and a negative attitude score. Then we applied the formula (3.1) by Stieglitz and Dang-Xuan (2013) to obtain the attitude valence for each message.

$$\textit{Attitude Valence} = \textit{Positive Attitude Score} + \textit{Negative Attitude Score} \tag{3.1}$$

Based on the scales of positive and negative attitude scores, the measure of attitude valence is in the range of -4 (very strong negative valence) to 4 (very strong positive valence). Then we aggregated the degree of attitude valence of messages to individual level by mean. Figure 3.4

[18] The tool SentiStrength is available at http://sentistrength.wlv.ac.uk

shows the distribution of attitude valence for all participants in the online health communities. In general, participants expressed a weak negative attitude in online health communities (mean = -0.41).



Mean = -0.41, SD = 0.60

Figure 3.4 Distribution of Participants' Attitude Valence

3.5 Results

Structural equation modeling (SEM) was used to test the hypotheses in the proposed research model. Compared with linear regression models, SEM has the capability of integrating the measurements (i.e., measurement model) and the hypothesized causal paths (i.e., structural model) and analyzing them simultaneously (Gefen et al. 2011). We can select one of the two most widely used SEM techniques in IS field, namely partial least squares SEM (PLS-SEM) and covariance-based SEM (CB-SEM).

PLS-SEM was selected for this study for four major reasons: (1) the research model contains formative items; (2) it includes both metric data as well as quasi-metric (ordinal) scaled data in dependent variables; (3) it contains non-normal data; and (4) its exploratory purpose to build novel theory (Chin et al. 2008; Gefen et al. 2011; Hair et al. 2013).

SmartPLS was used to test the research models. To assess the quality of results, the measurement model was first evaluated in terms of reliability and validity. Then the structural model was estimated to test the proposed hypotheses. Given our research objective of exploring the effect of social support on individual health promotion outcomes, we dropped missing values for self-reported health status and health knowledge. As a result, we obtained 24,506 observations of participants for structural equation modeling.

3.5.1 Measurement Model

Structural social capital is a formatively measured construct. As the indicators of the formative measurement model does not necessarily covary, criteria used to assess reflective measurement model such as composite reliability or average variance extracted (AVE) are not applicable to evaluate a formative measurement model (Hair et al. 2012). Recommendations in the literature (Cenfetelli and Bassellier 2009; MacKenzie et al. 2011; Petter et al. 2007) were applied to develop and validate the formative construct measurement. The results are summarized in Table 3.4. Correlations and descriptive statistics of the variables are shown in Appendix 3D. We noted that the weights of formative indicators are different under different model specifications, as “any measure (whether formative or reflective) is necessarily context-specific and, therefore, should not be considered in isolation of the context” (Diamantopoulos 2011, p. 341).

Table 3.4 Formative Measurement Collinearity, Weights, and Loadings

Item	VIF	Weight	T Value	Loading	T Value
Model Specification A					
SSC1	2.681	0.132*	2.520	0.720***	18.006
SSC2	1.058	0.065***	7.368	0.268***	29.339
SSC3	3.401	-0.073	1.889	0.754***	25.541
SSC4	2.378	0.948***	25.635	0.995***	285.086
Model Specification B					
SSC1	2.681	0.186***	3.655	0.831***	26.012
SSC2	1.058	0.021**	3.538	0.240***	26.534
SSC3	3.401	0.972***	19.644	0.989***	183.508
SSC4	2.378	-0.178***	3.293	0.678***	20.769

* $p<0.05$, ** $p<0.01$, *** $p<0.001$

3.5.1.1 Multicollinearity among Indicators

The formative measurement model evaluation began with an assessment of collinearity among the formative items. Evidence of substantial collinearity among formative indicators not only influences the estimation of their weights as well as statistical significance (Hair et al. 2013), but also may indicate that multiple indicators tap into the same aspect of the latent variable (Petter et al. 2007). Collinearity of an indicator is tested by regressing it on all other indicators of the structural social capital construct using ordinary least squares (OLS) method. A variance inflation factor (VIF) value of 5 implies that 80% of the indicator's variance is explained by the remaining formative indicators. All indicators except SSC3 (in-degree centrality) satisfy the recommended strict collinearity criterion, i.e., VIF < 3.33 (Diamantopoulos and Siguaw 2006). SSC3 has a moderate level of multicollinearity (VIF = 3.401) but still satisfy a less strict criterion, i.e., VIF < 5 (Hair et al. 2011). As the in-degree centrality does not have major conceptual overlap with other social network metrics, we do not need to remove any indicator at this point.

3.5.1.2 Significance and Relevance of Formative Indicators

An important aspect of formative measurement evaluation is to assess the contribution of each indicator through the calculation of its outer weight. The outer weight is the result of OLS by regressing the latent variable score on the formative indicators (Hair et al. 2013). The significance of outer weights was tested by bootstrapping procedure with 5,000 bootstrap samples (Hair et al. 2013). As Table 3.4 shows, all structural social capital indicators have significant outer weights. Thus, there is empirical support to retain all the formative indicators.

We also checked the co-occurrence of negative and positive indicator weights. In model specification A, the SSC3 (in-degree centrality) indicator has a negative outer weight (-0.073) significant at the 0.05 level. As suggested by Cenfetelli and Bassellier (2009), a suppressor effect might cause the negative weights. In this case, the bivariate correlation between SSC3 indicator and its construct is 0.754 (refer to the loading column in Table 3.4), which is less than the bivariate correlation between SSC3 and SSC1 (betweenness centrality), i.e., 0.781 (refer to Appendix 3D Table 3.13). That means SSC3 shares more variance with SSC1 than with the formatively measured construct. Thus, a suppressor effect of SSC1 on the correlation between SS3 and the construct explains the negative weight of the SSC3 indicator. The interpretation of the negative weight of SSC3 is that an increase of the betweenness will reduce the degree of structural social capital, holding other indicators constant. As SSC3 has a significant loading, there is empirical support to retain it in the formative measurement model (Cenfetelli and Bassellier 2009; Hair et al. 2013).

Similarly, in model specification B, the negative weight of item SSC4 can be explained by the suppressor effect given that SSC4 has a higher bivariate correlation (0.748) with indicator SSC3 than with the structural social capital construct (loading = 0.678). As the weight and loading of SSC4 are both significant, we chose to retain it in the formative measurement model.

3.5.1.3 Modified Multitrait-Multimethod (MTMM) Analysis

Convergent and discriminant validity of the structural social capital construct was evaluated by a modified multitrait-multimethod (MTMM) analysis (Campbell and Fiske 1959; Loch et al. 2003). Convergent validity requires that indicators of the same construct should correlate significantly with each other. To establish discriminant validity, each item should have a higher correlation with its construct than its correlations with other constructs. Table 3.5 summarizes the results. Convergent validity was achieved for the structural social capital construct in that its inter-indicator correlations are all significant at the 0.001 level. In terms of discriminant validity, there are some violations. In model A, indicators SSC1 and SS3 correlate slightly higher with ISP than with SSC, and SSC3 has a high level correlation with ESP. In model B, SSC2 correlates slightly higher with ISR than with SSC, meanwhile SSC4 has higher correlations with ESP and ESR than with its construct SSC. As suggested by Campbell and Fiske (1959), some violations to the basic MTMM principle in a large matrix are not necessarily meaningful. Since SSC is hypothesized to positively influence ISP, ISR, ESP, and ESR, high level correlations between the formative indicators and the later three constructs are expected due to the causal links. To conclude the modified MTMM analysis, we note a few exceptions but infer that the overall measurement validity is acceptable with regard to the overall discriminant validity of the structural social capital construct.

Table 3.5 Multitrait-multimethod (MTMM) Analysis

	SSC1	SSC2	SSC3	SSC4
SSC1	-			
SSC2	0.180	-		
SSC3	0.782	0.228	-	
SSC4	0.669	0.207	0.748	-
Model A **SSC**	0.720	0.268	0.754	0.995
Model B **SSC**	0.831	0.240	0.989	0.678
ISP	0.723	0.201	0.794	0.474
ESP	0.714	0.213	0.912	0.690
ISR	0.586	0.253	0.562	0.778
ESR	0.634	0.205	0.709	0.903
HK	-0.015	-0.100	-0.024	-0.025
SHS	0.004	0.002	0.017	0.000
AV	0.043	0.152	0.079	0.064

Note: latent variables are in bold.

3.5.2 Structural Model

3.5.2.1 Collinearity Assessment

PLS-SEM estimates the path coefficients of the structural model based on OLS of each endogenous latent variable on its predecessor variables (Hair et al. 2013). The path coefficient would be biased if there is significant level of multicollinearity among the predecessor variables. For model specification A, the VIFs are in the range of 1.066 to 2.896, with an average of 1.485. For model specification B, the VIFs are in the range of 1.066 to 2.657, with an average of 1.518. All the predecessor variables satisfy the recommended collinearity criterion VIF < 3.33 (Diamantopoulos and Siguaw 2006), showing no problem of multicollinearity.

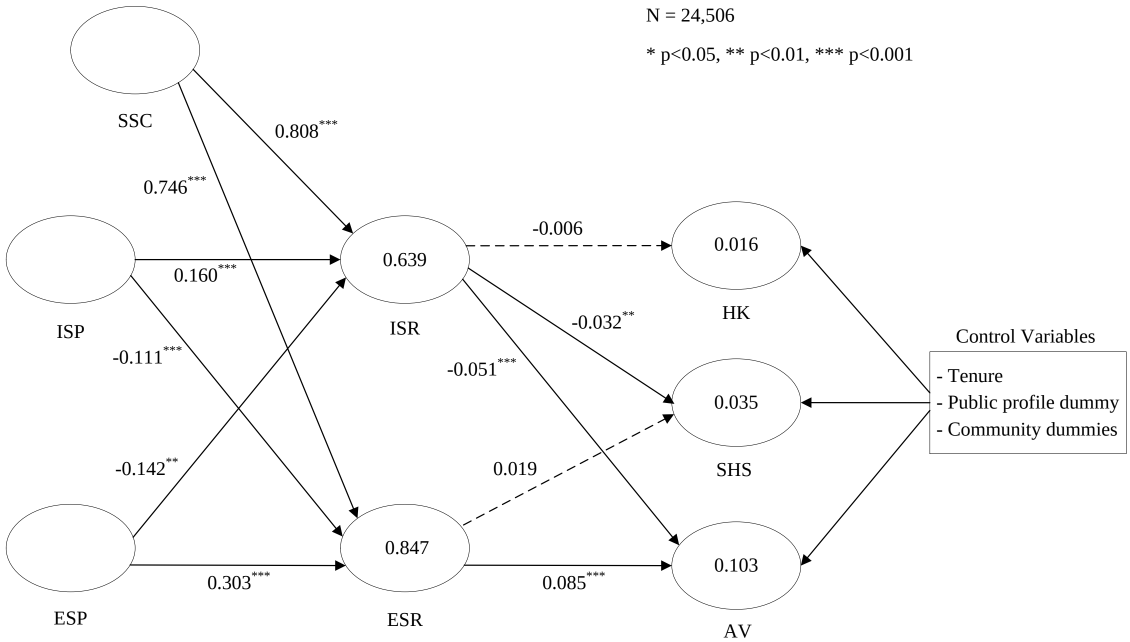
3.5.2.2 Overall Results

The structural model was assessed by standardized path coefficients, explained variance (R^2), and significance levels through bootstrapping with 5,000 bootstrap samples. Correlations among latent variables for each model specification are presented in

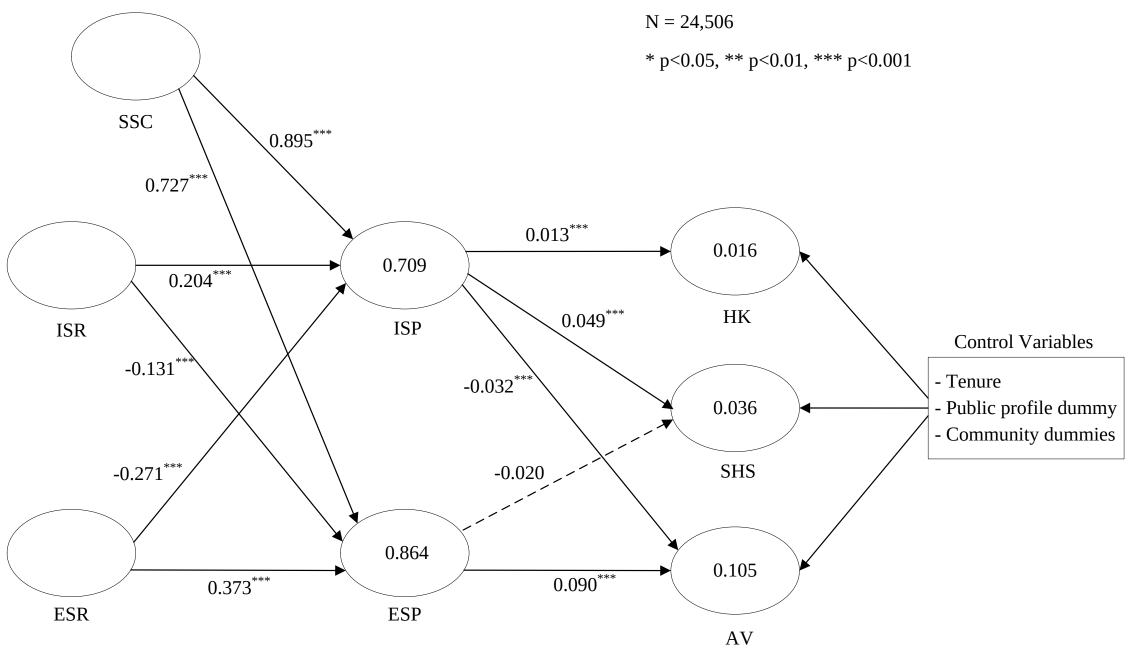
Table 3.6. Figure 3.5 summarizes the estimation results.

Table 3.6 Latent Variable Correlations

	1	2	3	4	5	6	7	8
Model Specification A								
1. Structural Social Capital	-							
2. Informational support provisioning	0.499	-						
3. Emotional support provisioning	0.696	0.671	-					
4. Informational support receipt	0.790	0.469	0.528	-				
5. Emotional support receipt	0.901	0.465	0.747	0.750	-			
6. Health knowledge	-0.031	0.004	-0.023	-0.017	-0.021	-		
7. Self-reported health status	-0.001	0.034	0.005	-0.016	-0.010	0.018	-	
8. Attitude valence	0.070	0.047	0.086	0.067	0.067	-0.087	0.105	-
Model Specification B								
1. Structural Social Capital	-							
2. Informational support provisioning	0.826	-						
3. Emotional support provisioning	0.901	0.671	-					
4. Informational support receipt	0.523	0.469	0.528	-				
5. Emotional support receipt	0.650	0.465	0.747	0.750	-			
6. Health knowledge	-0.023	0.004	-0.023	-0.017	-0.021	-		
7. Self-reported health status	0.017	0.034	0.005	-0.016	-0.010	0.017	-	
8. Attitude valence	0.077	0.047	0.086	0.039	0.067	-0.087	0.105	-



Model Specification A



Model Specification B

Figure 3.5 Structural Model Results

As shown in Figure 3.5, the basic structure of predictors of social support exchange in online health community as well as the impact of social support exchange on individual health promotion is confirmed. The results validate structural social capital and norm of reciprocity as important predictors of social support exchanged online. In model specification A, a community member's structural social capital is positively associated with his/her informational support receipt (path = 0.808, $p < 0.001$) and emotional support receipt (path = 0.746, $p < 0.001$). While informational support provisioning has positive effect on information support receipt (path = 0.160, $p < 0.001$), the provisioning of emotional support has negative effect on information support receipt (path = -0.142, $p < 0.01$), holding other factors constant. Similarly, the provisioning of emotional support has positive effect on emotional support receipt (path = 0.303, $p < 0.001$), but the effect of informational support provisioning on emotional support receipt is negative (path = -0.111, $p < 0.001$) after controlling for other factors. This result suggests that the provisioning of social support has a positive effect on the receipt of the same type of social support, but a negative effect on the receipt of different type of social support.

In model specification B, the level of structural social capital positively predicts the amount of informational support provisioning (path = 0.895, $p < 0.001$) and emotional support provisioning (path = 0.727, $p < 0.001$). While informational support receipt is positive related to information support provisioning (path = 0.204, $p < 0.001$), the receipt of emotional support has negative partial effect on information support provisioning (path = -0.271, $p < 0.001$). Similarly, receipt of emotional support has positive effect on emotional support receipt (path = 0.373, $p < 0.001$), while the effect of informational support receipt on emotional support provisioning is negative (path = -0.131, $p < 0.001$). The result reveals that the receipt of social support has positive effect on the provisioning of the same type of social support but negative effect on the provisioning of different type of social support.

In terms of the effects of social support exchange in promoting health, model A shows that informational support receipt has negative effects on self-reported health status (path = -0.032, $p < 0.01$) and attitude valence (path = -0.051, $p < 0.001$), while the receipt of emotional support has positive effect on attitude valence (path = 0.085, $p < 0.001$). In model B, informational support provisioning has mixed effects on health promotion outcomes: it contributes positively to health knowledge (path = 0.013, $p < 0.001$) and self-reported health status (path = 0.049, $p < 0.001$) but is negatively related to attitude valence (path = -0.032, $p < 0.001$). The provisioning of emotional support has positive effect on attitude valence (path = 0.090, $p < 0.001$)

The variance in informational and emotional support receipt explained by structural social capital and informational and emotional support provisioning is high at 63.9% and 84.7% respectively. Similarly, the variance in informational and emotional support provisioning explained by structural social capital and informational and emotional support receipt is high at 70.9% and 86.4% respectively. In contrast, the variance in health promotion outcomes explained by social support exchange is relatively low, ranging from 1.6% (health knowledge explained by informational and emotional support receipt in model A) to 10.5% (attitude valence explained by informational and emotional support provisioning in model B).

3.5.2.3 Hypothesis Testing

The results of hypothesis testing are summarized in Table 3.7. For the effect of structural social capital on social support exchange, the empirical results reveal significant positive impact of structural social capital on informational and emotional support receipt as well as provisioning. Thus, hypotheses H1a, H1b, H1c, and H1d are supported.

Table 3.7 Hypothesis Testing Results

#	Path	Hypothesis Description	Supported?
H1a	SSC → ISR	Participants' degree of structural social capital (SSC) will positively relate to their informational support receipt (ISR) in online health communities.	Yes
H1b	SSC → ESR	Participants' degree of structural social capital (SSC) will positively relate to their emotional support receipt (ESR) in online health communities.	Yes
H1c	SSC → ISP	Participants' degree of structural social capital (SSC) will positively relate to their informational support provisioning (ISP) in online health communities.	Yes
H1d	SSC → ESP	Participants' degree of structural social capital (SSC) will positively relate to their emotional support provisioning (ESP) in online health communities.	Yes
H2a	ISP → ISR	Informational support provisioning (ISP) in online health communities will be positively associated with informational support receipt (ISR).	Yes
H2b	ISP → ESR	Informational support provisioning (ISP) in online health communities will be positively associated with emotional support receipt (ESR).	No
H2c	ISR → ISP	Informational support receipt (ISR) in online health communities will be positively associated with informational support provisioning (ISP).	Yes
H2d	ISR → ESP	Informational support receipt (ISR) in online health communities will be positively associated with emotional support provisioning (ESP).	No
H3a	ESP → RIS	Emotional support provisioning (ESP) in online health communities will be positively associated with informational support receipt (ISR).	No
H3b	ESP → RES	Emotional support provisioning (ESP) in online health communities will be positively associated with emotional support receipt (ESR).	Yes

H3c	ESR → ISP	Emotional support receipt (ESR) in online health communities will be positively associated with informational support provisioning (ISP).	No
H3d	ESR → ESP	Emotional support receipt (ESR) in online health communities will be positively associated with emotional support provisioning (ESP).	Yes
H4a	ISR → HK	Participants' informational support receipt (ISR) in online health communities will positively relate to their health knowledge (HK).	No
H4b	ISP → HK	Participants' informational support provisioning (ISP) in online health communities will positively relate to their health knowledge (HK).	Yes
H5a	ISR → SHS	Participants' informational support receipt (ISR) in online health communities will positively relate to their self-reported health status (SHS).	No
H5b	ISR → AV	Participants' informational support receipt (ISR) in online health communities will positively relate to their attitude valence (AV).	No
H5c	ISP → SHS	Participants' informational support provisioning (ISP) in online health communities will positively relate to their self-reported health status (SHS).	Yes
H5d	ISP → AV	Participants' informational support provisioning (ISP) in online health communities will positively relate to their attitude valence (AV).	No
H6a	ESR → SHS	Participants' emotional support receipt (ESR) in online health communities will positively relate to their self-reported health status (SHS).	No
H6b	ESR → AV	Participants' emotional support receipt (ESR) in online health communities will positively relate to their attitude valence (AV).	Yes
H6c	ESP → SHS	Participants' emotional support provisioning (ESP) in online health communities will positively relate to their self-reported health status (SHS).	No
H6d	ESP → AV	Participants' emotional support provisioning (ESP) in online health communities will positively relate to their attitude valence (AV).	Yes

Test on the existence of norm of reciprocity in online health community shows mixed results: (1) social support provisioning has positive effect on the receipt of same type social support but negative effect on the receipt of different type social support; and (2) social support receipt has positive effect on the provisioning of same type social support but negative effect on the provisioning of different type social support. So, hypotheses H2a, H2c, H3b, and H3d are supported, while hypotheses H2b, H2d, H3a, and H3c are not supported.

The structural model analysis on the effect of social support exchange in promoting health supports hypotheses H4b, H5c, H6b, and H6d, with all other hypotheses not supported. As what we predicted, informational support provisioning positively influences the levels of health knowledge and self-reported health status. But the effect of informational support provisioning is negatively related to attitude valence. Contrary to our hypotheses, the effects of information support receipt on self-reported health status and attitude valence are significantly negative. Consistent with hypotheses, emotional support receipt and provisioning exerts positively effects on attitude valence.

3.5.2.4 Mediation Analysis

For each model specification, we conducted a formal mediation test at the structural model level by applying the linear regression using Zellner's seemingly unrelated regression (Zellner 1962). A major benefit of using seemingly unrelated regression is that it allows joint estimates by allowing errors associated with the dependent variables (i.e., health knowledge, self-reported health status, and attitude valence) to be correlated, thus leading to more efficient estimates than running multiple regressions separately. Compared with multivariate regression which regresses each dependent variable on the same set of independent variables, seemingly unrelated regression allows us to regress dependent variables on different sets of independent variables. Specifically, in our model health knowledge is hypothesized to be influenced only by informational support receipt and provisioning (H4a and H4b), while self-reported health status and attitude valence are supposed to be explained by the exchange of both informational and emotional support (hypotheses H5a through H6d). Thus, seemingly unrelated regression is preferred to specify the mediation model based on the PLS structural model.

The mediation effects were tested by the assumption-free bootstrapping procedure suggested by Preacher and Hayes (2008). A significant mediation effect requires three conditions to be satisfied: (1) significant path a (X→M); (2) significant path b (M→Y); and (3) significant indirect effect ab (X→M→Y). Mediation test results are summarized in Table 3.8. Only possible mediation effects with significant paths a (X→M) and b (M→Y) in the structural model are presented.

As shown in Table 3.8, in model A the effect of informational support provisioning on self-reported health status is fully mediated by information support receipt. Also, the effects of informational and emotional support provisioning on attitude valence are fully mediated by informational support receipt. In model B, the effects of structural social capital on self-reported health status and attitude valence are fully mediated by social support provisioning. There are partial mediation effects for social support provisioning in promoting health: (1) informational support provisioning partially mediates the effect of informational support receipt on self-reported health status; (2) informational and emotional support provisioning together mediate the effects of informational and emotional support receipt on attitude valence.

Table 3.8 Mediation Effect Testing Results

Mediation Path	Total Effect (c)	Direct Effect (c')	Indirect Effect (ab)	SE	Bias-C. 95% C. I.		Mediation
					Lower	Upper	
Model A							
SSC -> ISR -> SHS	-0.004	0.060	**-0.049**	0.013	-0.077	-0.025	none
SSC -> ISR -> AV	0.002	0.043	**-0.037**	0.009	-0.057	-0.023	none
SSC -> ESR -> AV			-0.005	0.011	-0.026	0.017	
ISP -> ISR -> SHS	0.049^{***}	0.055	**-0.010**	0.005	-0.021	-0.003	full mediation
ISP -> ISR -> AV	-0.032^{***}	-0.027	**-0.007**	0.003	-0.013	-0.003	full mediation via ISR
ISP -> ESR -> AV			0.001	0.002	-0.002	0.004	
ESP -> ISR -> SHS	-0.018	-0.020	**0.009**	0.004	0.002	0.019	none
ESP -> ISR -> AV	0.089^{***}	0.084	**0.007**	0.003	0.002	0.012	full mediation via ISR
ESP -> ESR -> AV			-0.002	0.004	-0.010	0.006	
Model B							
SSC -> ISP -> HK	-0.006	-0.113^{***}	**0.060**	0.012	0.041	0.082	none
SSC -> ISP -> SHS	0.040^{***}	0.019	**0.047**	0.012	0.024	0.072	full mediation
SSC -> ISP -> AV	0.041^{***}	0.019	**-0.023**	0.011	-0.049	-0.005	full mediation via both ISP and ESP
SSC -> ESP -> AV			**0.061**	0.011	0.044	0.086	
ISR -> ISP -> HK	0.006	0.003	**0.014**	0.004	0.007	0.022	none
ISR -> ISP -> SHS	-0.033^{**}	-0.049^{***}	**0.011**	0.005	0.004	0.021	partial mediation
ISR -> ISP -> AV	-0.052^{***}	-0.068^{***}	**-0.005**	0.002	-0.010	-0.002	partial mediation via both ISP and ESP
ISR -> ESP -> AV			**-0.011**	0.002	-0.016	-0.008	
ESR -> ISP -> HK	-0.013	0.002	**-0.018**	0.006	-0.032	-0.009	none
ESR -> ISP -> SHS	-0.005	0.023	**-0.014**	0.006	-0.028	-0.006	none
ESR -> ISP -> AV	0.060^{***}	0.047^{**}	**0.007**	0.004	0.002	0.016	partial mediation via both ISP and ESP
ESR -> ESP -> AV			**0.031**	0.006	0.021	0.044	

(1) * $p<0.05$, ** $p<0.01$, *** $p<0.001$; (2) indirect effects in bold are significant at 0.05 level;
(3) Bias-C. 95% C. I. refers to the bias-corrected 95% confidence interval of the indirect effects.

3.6 Discussions and Conclusions

3.6.1 Theoretical Implications

Drawing from the tenets of multiple theoretical bases selected, including social support theory and social capital theory, we have built a theoretical framework to identify the predictors of social support exchange in online health communities and explain the role of such social support exchange in promoting health. Rather than relying on a simple mechanism which explicates the influence of social relationships on individual's health [e.g., the full mediating role of social support provisioning suggested by Zhu et al. (2013)], we argue that the pathway from social interactions to health promotion is heterogeneous and nuanced, with multiple micro-mechanisms embedded in each other. We propose this more comprehensive framework to better understand the social interactions in online health communities and underlying theoretical relationships among them.

3.6.1.1 Do Structural Network Positions Matter?

Based on a big dataset collected from nine online health communities, the empirical results reveal that structural social capital has significant and positive effects on social support exchange including the provisioning and receipt of informational and emotional support. The resources embedded in the social interaction ties by virtue of individual's network structural positions do explain the amount of social support the individual receives from others as well as contributes to online health communities. Positioned at a high level of structural social capital in the online health community provides participants advantageous resources in facilitating social support exchange. This finding conforms to previous studies that structural social capital can predict the functions of social interactions such as knowledge contribution (Chiu et al. 2006; Wasko and Faraj 2005) and social support exchange (Huang and Chengalur-Smith 2014).

3.6.1.2 Presence of Reciprocity in Online Health Community

This study has investigated the existence of reciprocity in the setting of online health communities. Our analysis shows that the norm of reciprocity exists between informational support provisioning and receipt as well as between emotional support provisioning and receipt. The universal social rule that people repay others for benefits obtained from them operates in informational or emotional social support exchange, but not across different types of social support. Social support exchange in online health community is a continuous and interacting process. Obtaining a high level of informational or emotional support from other community peers implies a significant level of contributing the same type of social support to the community.

However, there may be suppression effects between the provisioning and receipt of different types of social support. On one hand, receiving higher level informational or emotional support from the online community seems to inhibit one's motivation to contribute different types of social support to others. On the other hand, provisioning of informational and emotional support could negatively influence the receipt of the same type of social support from other peers. Thus, this study extends findings by Bowling et al. (2005) that the reciprocity rule in that it empirically tests the reciprocity rule down to detailed types of social support exchanged rather than at the aggregate level.

Our findings support the view that reciprocity works as a catalyst for both social support provisioning and receipt. With a strong norm of reciprocity in the peer-to-peer online health communities, participants are assured that their social support provisioning efforts will be rewarded, thus motivating them to contribute more social support to others. As a result, social support exchange at the community level is heavily stimulated, thus ensuring the sustainability and prosperity of the online health communities.

Moreover, by unpacking the reciprocity mechanism of social support exchange from the general social support level down to specific social support types, our findings uncover a more comprehensive view on the underlying theoretical relationships in the complicated social interaction process. The findings have significant theoretical implications for understanding the intricate social support exchange in online as well as offline settings.

3.6.1.3 Role of Social Support Exchange in Health Promotion

This study presents evidence on the health promoting role of social support exchange. We found that informational support provisioning does influence one's level of health knowledge and self-reported health status. Given the setting of a peer-to-peer online social support exchange, participants who are involved in the online interactions seek and provide information and advice regarding the treatment of diseases as well as encouragement and emotional support on health self-management. The level of involvement in online social interaction is largely determined by social support provisioning rather than social support receipt. By participating in informational support provisioning, a member needs to absorb and assimilate external knowledge from online social interaction as well as self-learning of other materials. As a result, individuals appear to be accumulating and enhancing their health-related knowledge. Similarly, we argue that active participation in informational support provisioning enhances one's capability for self-managing health, thus increasing the level of self-reported health.

In contrast, our results do not show a profound effect of informational support receipt on health knowledge and, even more counter to prediction, informational support receipt has a significantly negative effect on self-reported status. This implies that just receiving informational support does not automatically increase one's level of health knowledge, but the provisioning of informational support does. Even more curious is that, the more informational support received, the lower the level of self-reported status. Active involvement in sharing and providing information rather than passive receipt of informational support tends to better inform participants about health self-management, thereby obtaining benefits from the social interaction process in terms of accumulating health-related knowledge and bolstering one's self-reported health status.

With respect to the effects of social support exchange on attitude valence, our results show mixed effects: (1) the provisioning and receipt of emotional support have significantly positive effects on attitude valence and (2) the provisioning and receipt of informational support have significantly negative influences on attitude valence. As emotional support conveys love, care, sympathy, encouragement, or empathy, both the provider and the recipient benefit from the emotional support exchange in terms of expressing more positive attitudes in their online posts. Emotional support exchanged in online health communities boosts participants' perceived ability to cope with stressful events as well as alleviating the negative impact of stressors.

Interestingly, our empirical analysis reveals negative effects of informational support exchange on attitude valence. Future studies, especially those using qualitative methods or field study or experimental designs to investigate the behavioral and psychological aspects of social support exchange are needed to further explain such mixed results.

3.6.2 Practical Implications

Online health communities have become and are increasingly regarded as an inseparable part of today's personalized preventative medicine. The flexible peer-to-peer interaction mechanism and the advantages of no limit of time and space allow participants to be maximally involved in online social support exchange, through which individuals are empowered to better self-monitor and self-manage their health and wellness. Empirical findings of this study via analyses of big datasets of user-generated content have implications for online health community management as well as decision making regarding health intervention and promotion.

This research suggests a broader view of how structural social capital explains the levels of social support exchange and how such social support exchange improve health outcomes for those actively engaged in managing their behaviors. This general view provides valuable insights on the design and management of online health communities. As our findings confirm the positive effects of structural social capital on social support exchange, online health community managers and health policy makers should provide website features and guidance to encourage the social interaction that helps building structural social capital resources for participants. For example, social network features such as chatting, following specific users, friendship building, and mentioning/ referencing users in posts may facilitate maintaining and enhancing interactions among community members, thereby building social capital for online participants.

Our findings confirm the catalytic role of reciprocity in social support exchange. Specifically, by unpacking the reciprocity mechanism down to specific social support types, our study reveals the presence of reciprocity between informational support provisioning and receipt as well as between emotional support provisioning and receipt; it also finds that the provisioning and receipt of different types of social support have suppression effects. Thus, online health communities could provide corrective instructions to guide participants in seeking and exchanging different types of social support. Participants whose main purpose is to obtain one single type of social support (either informational or emotional support) could be encouraged to involve in online interactions of the same type of social support. Focusing on one type of social support exchange ensures the effectiveness of such social support exchange, thus satisfying the needs of participants seeking specific type of social support. Such online health intervention guidance can help community managers in maintaining the continued commitment of current members.

Lastly, the present study employs and validates various text mining techniques for automatic content analysis of digital trace data. Our analytical approaches can be applied by online health community managers and health policy makers to similar settings to evaluate the social interaction efficacy, health promotion effect, and leadership of social support exchange within specific online health communities. The natural language processing techniques and machine learning approaches used in classifying social support expressed in short messages can be used in real time to evaluate the effectiveness and efficiency of social support exchange in online health communities. Social support requests that have not been effectively satisfied can be routed to community moderators or leaders who are experienced in promoting social support exchange (Wang et al. 2012). Moreover, our analytical methods such as social network analysis and health-related knowledge assessment can be used to identify leaders in social support exchange, such that online health community managers can collaborate closely with these leaders to better serve all participants in the online communities.

3.6.3 Limitations and Future Research

There are several limitations that need to be noted. The first limitation pertains to the generalizability of the findings. Our data were collected from nine online health communities hosted in the United States, which has a very high rate of Internet users who look online for health information. The findings may not apply to other cultures that do not actively participant in such online discussions. Future research can extend the current study by applying it to other countries and cultural backgrounds.

The second limitation concerns the cross-sectional design of this study. Although our results reveal associations between structural social capital and social support exchange as well as between social support exchange and individual health promotion, it is not conclusive on the direction of such associations. Future research can apply more advanced techniques such as latent growth model (LGM) to empirically analyze longitudinal data of the online health communities to obtain more confirmative results on the direction of these effects, thus confirming the causality of the underlying relationships.

The third potential limitation is related to the PLS-SEM method used in our data analysis. In our study, the construct of structural social capital is modeled formatively. Though we note that the issue of formative measurement model has been debated for decades and there is no a single solution as to best analyze formative constructs, the error free assumption of PLS-SEM for formative constructs may lead to inflated estimation of weights (Cenfetelli and Bassellier 2009). Such potential issues need to be acknowledged in interpreting the results.

In addition, as discussed in section 3.6.1.3, an opportunity for future study is to use field studies or experimental designs to investigate the behavioral and psychological aspects of social support exchange that can triangulate our findings or better explain the mixed roles of various kinds of social support exchange in promoting individual health. Qualitative studies of online participants, such as netnographic studies, could also be useful.

This study focuses on individual characteristics of social interactions in online health communities to build the general framework of the predictors of social support exchange as well as its role in health promotion. As an extension of the current research, future study can take into account the collective social capital at the community level (Yang et al. 2009) to further explore the social support exchange in online communities.

Appendix

Appendix 3A: Social Support Behavior Code

Table 3.9 Definition of Social Support Behavior Code, Adapted from Cutrona and Suhr (1992)

Support Type	Definition
Informational Support	
Suggestion/advice	Offers ideas and suggests actions
Referral	Refers the recipient to some other source of help
Situation Appraisal	Reassesses or redefines the situation
Teaching	Provides detailed information, facts, or news about the situation or about skills needed to deal with the situation
Tangible Support	
Loan	Offers to lend the recipient something
Direct task	Offers to perform a task directly related to the stress
Indirect task	Offers to take over one or more of the recipient's other responsibilities while the recipient is under stress
Active	Offers to join the recipient in action that reduces the stress
Willingness	Expresses willingness to help
Emotional Support	
Relationship	Stresses the importance of closeness and love in relationship with the recipient
Physical affection	Offers physical contact, including hugs, kisses, hand-holding, shoulder patting
Confidentiality	Promises to keep the recipient's problem in confidence
Sympathy	Expresses sorrow or regret for the recipient's situation or distress
Listening	Attentive comments as the recipient speaks
Understanding/ Empathy	Expresses understanding of the situation or discloses a personal situation that communicates understanding
Encouragement	Provides the recipient with hope and confidence
Prayer	Prays with the recipient
Esteem Support	
Compliment	Says positive things about the recipient or emphasizes the recipient's abilities
Validation	Expresses agreement with the recipient's perspective on the situation
Relief of blame	Tries to alleviate the recipient's feelings of guilt about the situation
Network Support	
Access	Offers to provide the recipient with access to new companions
Presence	Offers to spend time with the person, to be there
Companions	Reminds the person of availability of supportive companions, of others who are similar in interests or experience

Appendix 3B: Summary Statistics of Online Health Communities

Figure 3.6 Daily Message Count per Forum

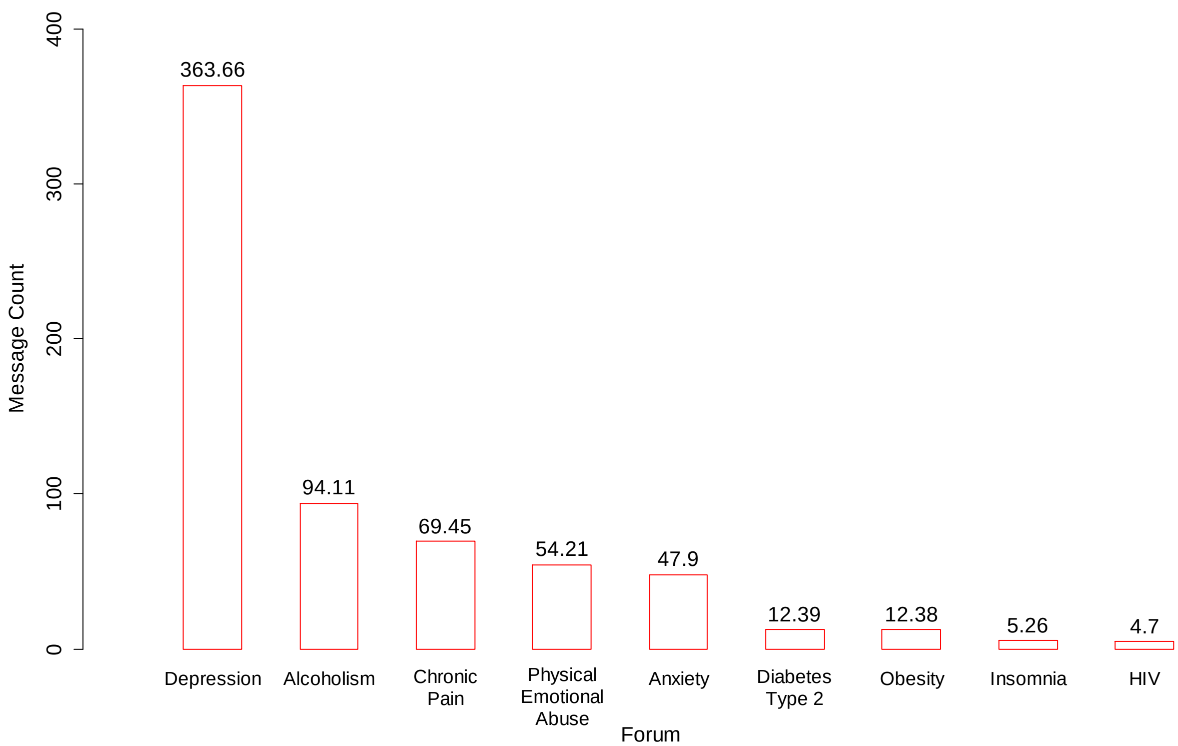


Figure 3.7 Distribution of Number of Posts per Thread

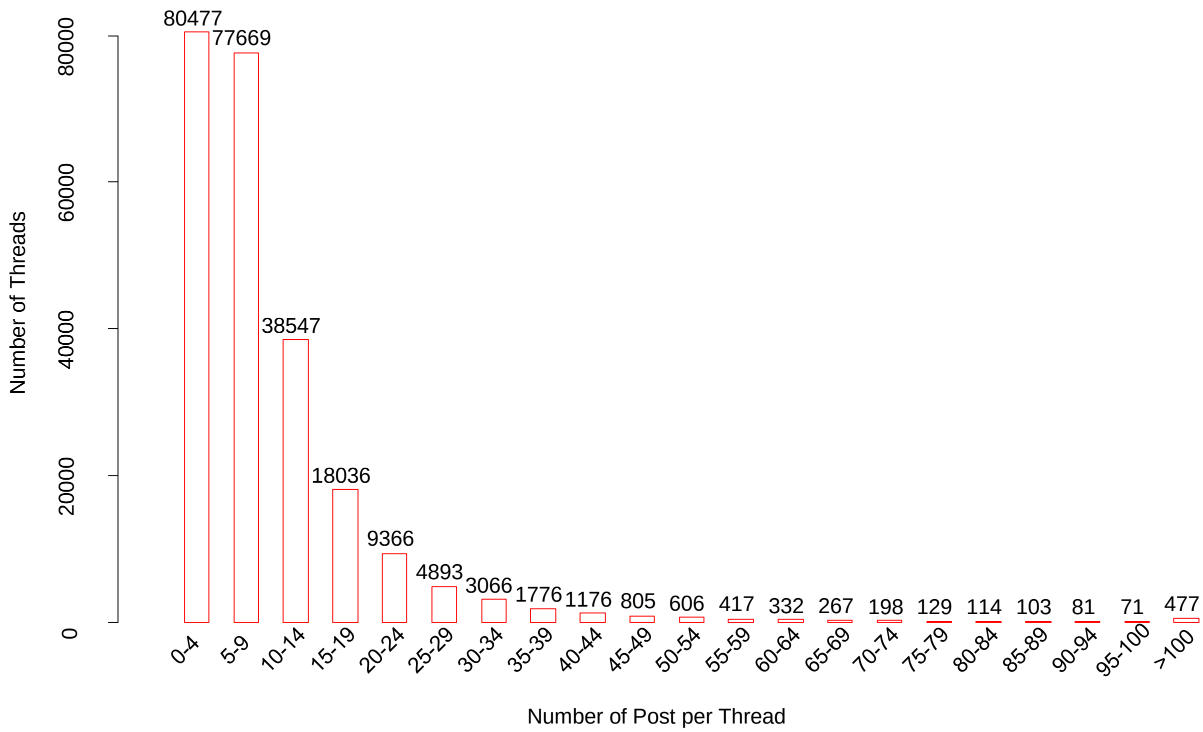
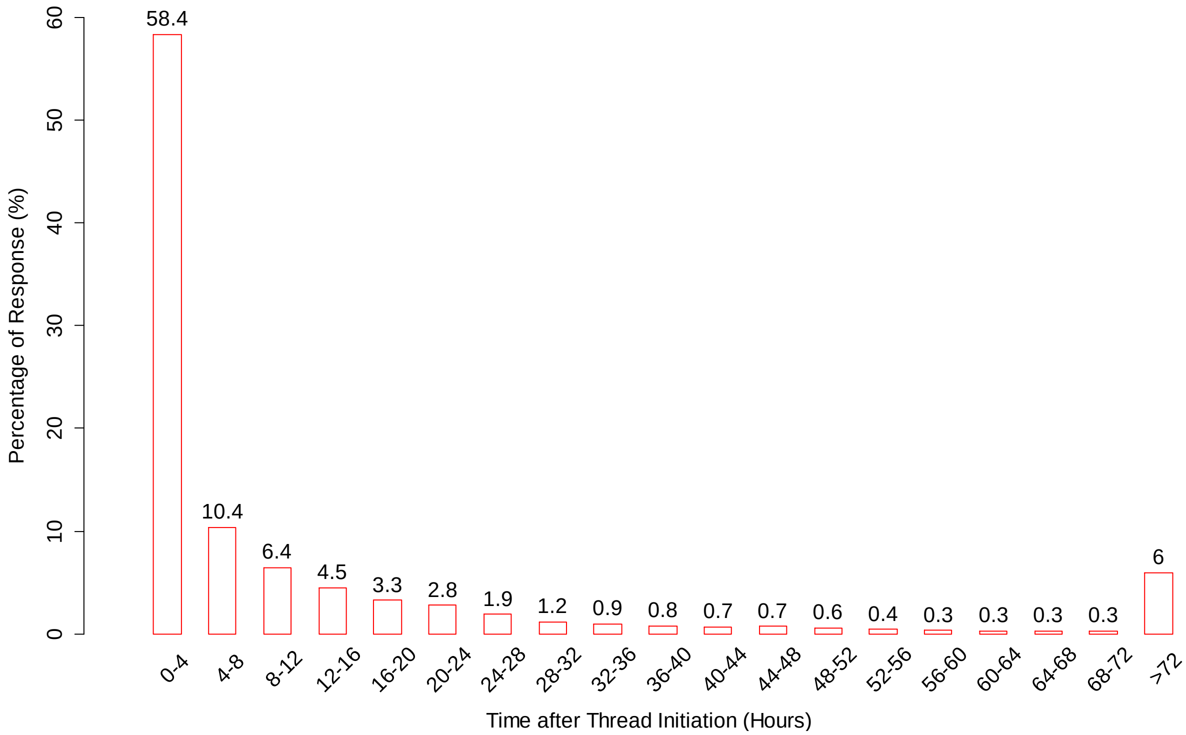


Figure 3.8 Distribution of Response after Thread Initiation



Appendix 3C: SVM-Based Social Support Classification

The technical details of the social support classifiers are discussed in this appendix. Generally, a text classification process begins with the preparation of features that are extracted from the text. Then these features are used to train a classifier [e.g., support vector machine (SVM), Naïve Bayes, decision tree, or artificial neural network model etc.]. With satisfactory performance, the classifier can be used to actually assess new textual contents. In this study, informational and emotional support was systematically analyzed through the following procedure.

Step 1. Extract Textual Features

The participants tend to apply different writing styles and elements in expressing different type of social support in their online communications (Wang et al. 2012). To capture these characteristics of social support expression, we extracted four major types of features in the text classification of social support. Table 3.10 provides a summary of these features. The basic linguistic and part of speech (POS) features were extracted by using natural language processing techniques. The sentiment features of messages were analyzed by using the MPQA corpus[19] (Wiebe et al. 2005). Latent Dirichlet allocation (LDA) topic modeling approach (Blei et al. 2003) was used to extract topic features from the online discussion messages. In LDA, each online post is modeled as a mixture over an underlying set of latent topics, while each topic is characterized by a distribution over various words (Blei et al. 2003). Following the same rule as used by Wang et al. (2012) and Wang et al. (2014), we set the LDA model to generate 20 latent topics. Table 3.11 presents the topics extracted from LDA and their highly associated terms.

---

[19] The MPQA corpus is available at http://mpqa.cs.pitt.edu

For each message, the probability of belonging to each topic (i.e., the topic distribution) was used as topic features of this message for social support classification.

Table 3.10 Summary of Features for Social Support Classification

Feature Sets	**Features**
Basic Linguistic Features	Count of sentences in the message
	Average term count in each sentence
	Count of sentences that contains negation terms (e.g., “not”, “never”, “n't”, and “no”)
	Count of sentences that contains a question mark, i.e., “?”
	Count of sentences that follow a format of <you + MODAL> (e.g., “you can”, “you could”, “you may”, “you might”, “you must”, “you shall”, “you should”, “you’d”, and “you had better” etc.)
	Count of advice terms (e.g., “advise”, “advocate”, “ask”, “desire”, “expect”, “necessitate”, “propose”, “recommend”, “request”, and “require” etc.) in the message
	Count of “if you” in the message
	Count of emoticons [e.g., “(-:”, “(:”, “:-D”, and “:-)” etc.] in the message
	Count of URLs (uniform resource locators) in the message
	Count of the Internet slang words (e.g., “alol”, “cid”, “cyo”, and “idk” etc.) in the message
Part-of-Speech (POS) Features	Count of numerical numbers in the message
	Count of proper nouns in the message
	Count of adjectives in the message
Sentiment Features	Count of terms with positive sentiment
	Count of terms with negative sentiment
	Count of terms with strong subjectivity
	Count of terms with weak subjectivity
Topic Features	20 topic distributions extracted from LDA

Table 3.11 20 Topics Extracted from Latent Dirichlet Allocation

Topic#	Top 30 Terms (Stemmed)
1	lol, love, thank, funni, laugh, yeah, lmao, look, gui, ye, hei, dont, ass, fuck, that, bump, sorri, mean, post, girl, shit, name, gonna, nice, damn, hell, littl, fun, head, omg
2	anxieti, feel, help, attack, time, breath, try, panic, fear, start, mind, worri, stress, relax, heart, bodi, happen, symptom, anxiou, sometim, dai, calm, yourself, caus, bad, think, lot, deep, control, head
3	eat, food, drink, chocol, water, cook, coffe, cream, sugar, lol, ic, tea, chicken, chees, cup, cake, love, tast, milk, hot, dinner, bread, egg, potato, fruit, bake, juic, sweet, cooki, butter
4	feel, help, talk, dont, time, peopl, try, sorri, hope, understand, tell, care, pleas, yourself, friend, hard, bad, depress, mayb, hug, hurt, sometim, life, dai, call, cant, support, happen, that, lot
5	hug, love, hope, thank, sorri, glad, happi, post, welcom, feel, prayer, wish, friend, dai, send, hear, wonder, help, share, pleas, care, god, support, bless, soon, hugs, prai, hun, peac, lot
6	job, call, monei, pai, help, peopl, time, insur, live, care, phone, school, servic, health, home, compani, hous, look, free, local, check, polic, bill, abl, law, medic, disabl, month, offic, legal
7	kid, time, famili, love, mom, life, mother, son, friend, parent, live, children, daughter, child, dai, husband, dad, home, father, tell, told, sister, school, care, babi, talk, feel, brother, ago, own
8	pain, doctor, med, help, doc, medic, hope, surgeri, time, chronic, dai, caus, patient, nerv, care, relief, try, test, drug, take, sorri, luck, bad, manag, treat, told, tell, month, call, feel
9	http, comwatch, wwwyoutub, song, love, music, listen, sing, favorit, video, movi, youtub, john, lyric, plai, band, danc, rock, beauti, blue, live, lol, version, link, watch, heard, classic, michael, jame, nice
10	feel, life, time, yourself, peopl, love, try, help, live, chang, person, dai, happi, hard, look, learn, real, own, sometim, friend, care, posit, lot, depress, hope, start, do, pain, take, understand
11	dai, sleep, night, time, morn, feel, hope, bed, hour, try, week, pain, wake, start, dream, help, get, bad, home, tomorrow, rest, tire, stai, do, littl, sometim, mayb, lol, fall, watch

12	peopl, post, person, read, friend, help, time, support, feel, try, site, agre, board, talk, mean, thread, thank, lot, repli, ignor, look, comment, sorri, mayb, opinion, word, tell, messag, sometim, understand
13	http, link, site, wwwdailystrength, help, click, page, dailystrength, check, www, read, websit, dont, comput, happy, yes, found, book, com, name, look, free, googl, people, chat, org, day, info, search, onlin
14	look, dog, love, lol, walk, time, cat, littl, hous, dai, plai, watch, car, live, home, hair, wear, hand, sit, run, head, door, drive, clean, ey, water, nice, kid, fun, light
15	abus, feel, relationship, time, yourself, love, leav, person, chang, tell, control, hurt, try, life, care, emot, women, pleas, real, help, stai, own, victim, behavior, believ, wrong, physic, husband, do, happen
16	drink, alcohol, sober, meet, time, step, life, dai, help, peopl, stop, sobrieti, recoveri, stai, sponsor, drunk, start, program, addict, live, god, quit, try, chang, do, real, book, diseas, lot, found
17	depress, help, medic, peopl, therapi, mental, anxieti, therapist, doctor, ill, issu, med, disord, person, treatment, caus, health, feel, suffer, life, deal, understand, lot, support, time, talk, brain, physic, symptom, experi
18	god, believ, peopl, life, world, power, live, person, faith, book, own, religion, human, read, word, church, christian, belief, love, spiritu, mean, mind, bibl, religi, jesu, true, question, understand, creat, real
19	weight, eat, food, diabet, lose, exercis, diet, sugar, dai, start, gain, blood, time, help, bodi, fat, lost, try, healthi, pound, week, carb, doctor, walk, lot, control, lb, test, meal, low
20	med, take, effect, help, doctor, anxieti, drug, medic, dai, start, time, week, dose, sleep, depress, feel, month, pill, doc, try, stop, xanax, luck, caus, tri, bodi, prescrib, addict, vitamin, lexapro

***Step 2.* Train Support Vector Machine Classifier**

In this study, support vector machine (SVM) model was used to classify informational and emotional support expressed in messages replied in the online health communities. Since each message may contain both informational support and emotional support, we build a classifier for each kind of social support. The manually coded 3,086 reply messages were used to train the SVM-based classifiers. Among the 3,086 messages, 662 posts contain informational support and 706 posts contain emotional support. This is an unbalanced dataset, with approximately not equal classification categories. The result of the unbalanced dataset is the bad accuracy performance of standard classifiers (Japkowicz 2000). To solve the unbalanced dataset issue, SMOTE algorithm (Chawla et al. 2002) was used to generates synthetic minority cases to over-sample the minority categories.

The LIBSVM library[20] (Chang and Lin 2011) was used to build the SVM classifiers. We chose the C-Support Vector Classification (C-SVC) with RBF kernel to train the social support classifiers. A grid-search strategy with 10-fold cross-validation was utilized to determine the best parameters *c* and *gamma* of the RBF kernel. As illustrated in Figure 3.9, the best parameters for information support classifier are $c = 8$ and $gamma = 0.5$, resulting in accuracy performance at 87.41% level. Figure 3.10 shows that the emotional support classifier with parameters $c = 2$ and $gamma = 2$ provides best accuracy performance at 84.01% level.

[20] The LIBSVM library is available at https://www.csie.ntu.edu.tw/~cjlin/libsvm/

Figure 3.9 Optimization of Informational Support Classifier

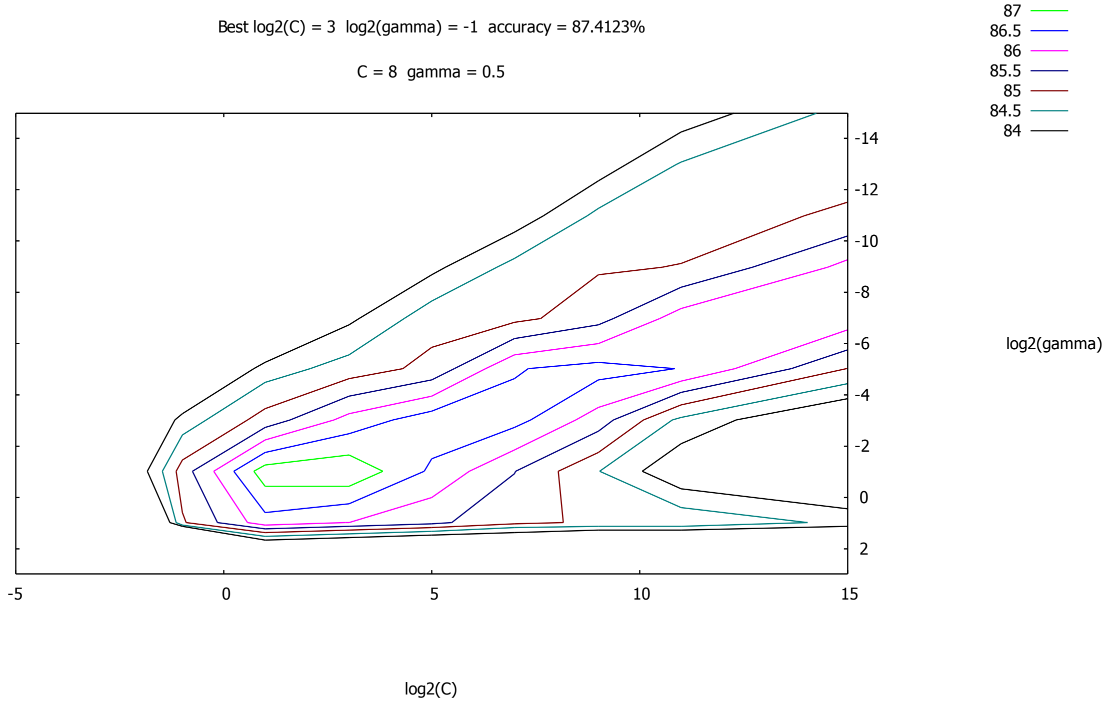
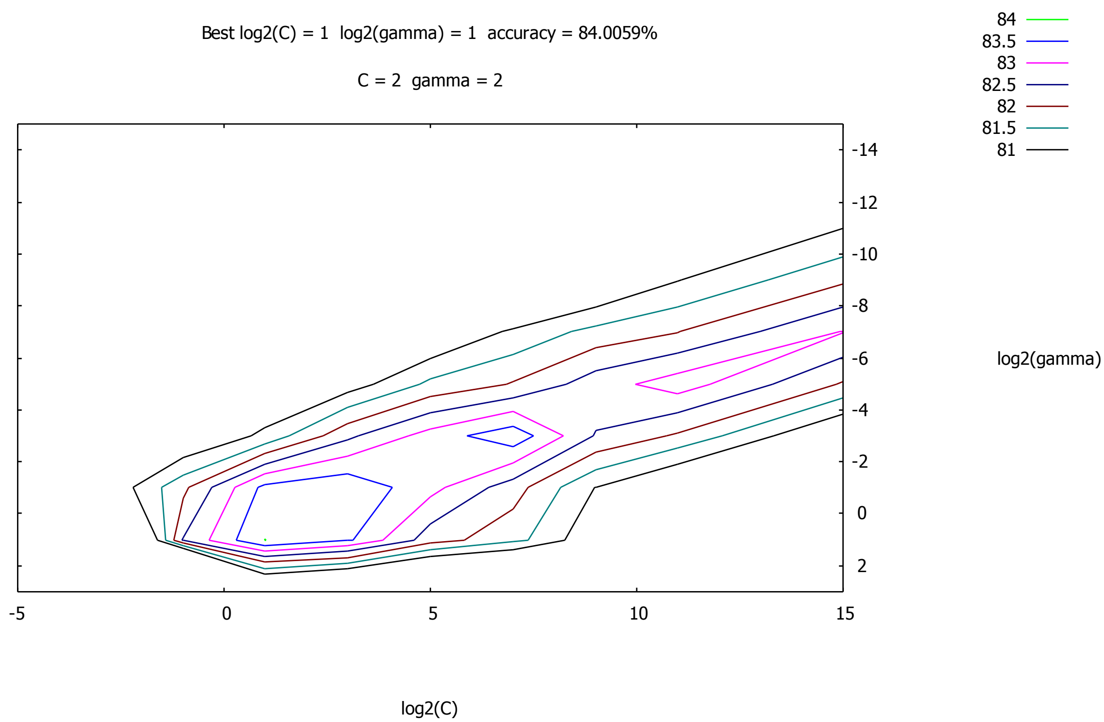


Figure 3.10 Optimization of Emotional Support Classifier



We also compared the performance of the SVM-based classifiers with other classification algorithms, summarized as in Table 3.12. For both information support classification and emotional support classification, the SVM-based classifiers outperform other commonly used algorithms including Naïve Bayes, Logisitc, C4.5 decision tree, and AdaBoost. Given the comparison, we were more convinced to choose the SVM-based methods.

Table 3.12 Comparison of Accuracy Performance for Different Classifiers

	SVM	**Naïve Bayes**	**Logistic**	**C4.5**	**AdaBoost**
Informational Support	87.41%	79.32%	84.87%	85.43%	82.73%
Emotional Support	84.01%	66.90%	80.85%	82.37%	80.05%

Step 3. **Classify Social Support**

After the SVM-based social support classifiers trained and evaluated, the classification algorithms classifiers were used to automatically code the rest of the online community posts. The results of social support classification were used to calculate the social support measures (refer to section 3.4.2).

Appendix 3D: Correlations and Descriptive Statistics of Variables

Table 3.13 Correlations and Descriptive Statistics (N = 24,506)

	1	2	3	4	5	6	7	8	9	10	11
1. Betweenness	-										
2. Closeness	0.180	-									
3. In-degree	0.782	0.228	-								
4. Out-degree	0.669	0.207	0.748	-							
5. Informational support provisioning	0.723	0.201	0.794	0.474	-						
6. Emotional support provisioning	0.714	0.213	0.912	0.690	0.671	-					
7. Received informational support	0.586	0.253	0.562	0.778	0.469	0.528	-				
8. Received emotional support	0.634	0.205	0.709	0.903	0.465	0.747	0.750	-			
9. Health knowledge	-0.015	-0.100	-0.024	-0.025	0.004	-0.023	-0.017	-0.021	-		
10. Self-reported status	0.004	0.002	0.017	0.000	0.034	0.005	-0.016	-0.010	0.018	-	
11. Attitude valence	0.043	0.152	0.079	0.064	0.047	0.086	0.039	0.067	-0.087	0.105	-
Mean	0.000	0.152	54.373	47.604	14.153	13.036	11.469	11.156	37.769	3.053	-0.453
Standard Deviation	0.000	0.037	367.931	303.279	92.777	96.485	43.212	72.469	30.556	1.121	0.835

REFERENCES

Abbott, A. 1988. *The System of Professions: An Essay on the Division of Expert Labor*. Chicago, IL USA: University of Chicago Press.

Abbott, A. 1999. *Department and Discipline*. Chicago, IL USA: University of Chicago Press.

Abbott, A. 2001. *Chaos of Disciplines*. Chicago, IL USA: University of Chicago Press.

Adler, N.J., and Bartholomew, S. 1992. "Academic and Professional Communities of Discourse: Generating Knowledge on Transnational Human Resource Management," *Journal of International Business Studies* (23:3), pp. 551-569.

Agarwal, A., Xie, B., Vovsha, I., Rambow, O., and Passonneau, R. 2011. "Sentiment Analysis of Twitter Data," in: *Proceedings of the Workshop on Languages in Social Media*. Portland, Oregon: Association for Computational Linguistics, pp. 30-38.

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.

AMA. 2014. "Health Information Technology." Retrieved June 1, 2014, from http://www.ama-assn.org/ama/pub/advocacy/topics/health-information-technology.page

Andrews, J.E. 2003. "An Author Co-Citation Analysis of Medical Informatics," *Journal of the Medical Library Association* (91:1), pp. 47-56.

Baird, A. 2014. "Making the Case for Leveraging the Patient-Centered E-Health (PCEH) Context to Expand Information Systems Theory," *Communications of the Association for Information Systems* (34), pp. 365-380.

Barrera, M. 1981. "Social Support in the Adjustment of Pregnant Adolescents: Assessment Issues," in *Social Networks and Social Support*. Beverly Hills: Sage, pp. 69-96.

Barrera, M., Jr. 1986. "Distinctions between Social Support Concepts, Measures, and Models," *American Journal of Community Psychology* (14:4), pp. 413-445.

Baumgartner, H., and Pieters, R. 2003. "The Structural Influence of Marketing Journals: A Citation Analysis of the Discipline and Its Subareas over Time," *Journal of Marketing* (67:2), pp. 123-139.

Benckendorff, P. 2009. "Themes and Trends in Australian and New Zealand Tourism Research: A Social Network Analysis of Citations in Two Leading Journals (1994–2007)," *Journal of Hospitality and Tourism Management* (16:1), pp. 1-15.

Berkman, L.F., Glass, T., Brissette, I., and Seeman, T.E. 2000. "From Social Integration to Health: Durkheim in the New Millennium," *Social Science & Medicine* (51:6), pp. 843-857.

Bernroider, E.W., Pilkington, A., and Córdoba, J.-R. 2013. "Research in Information Systems: A Study of Diversity and Inter-Disciplinary Discourse in the AIS Basket Journals between 1995 and 2011," *Journal of Information Technology* (28:1), pp. 74-89.

Berwick, D.M., Nolan, T.W., and Whittington, J. 2008. "The Triple Aim: Care, Health, and Cost," *Health Affairs* (27:3), pp. 759-769.

Biehl, M., Kim, H., and Wade, M. 2006. "Relationships among the Academic Business Disciplines: A Multi-Method Citation Analysis," *Omega* (34:4), pp. 359-371.

Blank, T.O., Schmidt, S.D., Vangsness, S.A., Monteiro, A.K., and Santagata, P.V. 2010. "Differences among Breast and Prostate Cancer Online Support Groups," *Computers in Human Behavior* (26:6), pp. 1400-1404.

Blau, P.M. 1964. *Exchange and Power in Social Life*. New York: Wiley.

Blei, D.M., Ng, A.Y., and Jordan, M.I. 2003. "Latent Dirichlet Allocation," *Journal of Machine Learning Research* (3), pp. 993-1022.
Blumenthal, D., and Glaser, J.P. 2007. "Information Technology Comes to Medicine," *New England Journal of Medicine* (356:24), pp. 2527-2534.
Blumenthal, D., and Tavenner, M. 2010. "The 'Meaningful Use' Regulation for Electronic Health Records," *New England Journal of Medicine* (363:6), pp. 501-504.
Bodenreider, O. 2004. "The Unified Medical Language System (UMLS): Integrating Biomedical Terminology," *Nucleic Acids Research* (32:suppl 1), pp. D267-D270.
Borgatti, S.P. 2002. *Netdraw Software for Network Visualization*. Lexington, KY: Analytic Technologies.
Bowling, N.A., Beehr, T.A., and Swader, W.M. 2005. "Giving and Receiving Social Support at Work: The Roles of Personality and Reciprocity," *Journal of Vocational Behavior* (67:3), pp. 476-489.
Braithwaite, D.O., Waldron, V.R., and Finn, J. 1999. "Communication of Social Support in Computer-Mediated Groups for People with Disabilities," *Health Communication* (11:2), pp. 123-151.
Brown, G.D., Patrick, T.B., and Pasupathy, K. 2012. *Health Informatics: A Systems Perspective*. Chicago, IL: Health Administration Press.
Buchbinder, S.B., and Shanks, N.H. 2011. *Introduction to Health Care Management*. Jones & Bartlett Publishers.
Buntin, M.B., Burke, M.F., Hoaglin, M.C., and Blumenthal, D. 2011. "The Benefits of Health Information Technology: A Review of the Recent Literature Shows Predominantly Positive Results," *Health Affairs* (30:3), pp. 464-471.
Burt, R.S. 1992. "The Social Structure of Competition," in *Networks and Organizations*, N. Nohria and R.G. Eccles (eds.). Boston: Harvard Business School Press, pp. 57-91.
Campbell, D.T., and Fiske, D.W. 1959. "Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix," *Psychological Bulletin* (56:2), pp. 81-105.
Cenfetelli, R.T., and Bassellier, G. 2009. "Interpretation of Formative Measurement in Information Systems Research," *MIS Quarterly* (33:4), pp. 689-707.
Chang, C.-C., and Lin, C.-J. 2011. "LIBSVM: A Library for Support Vector Machines," *ACM Transactions on Intelligent Systems and Technology (TIST)* (2:3), pp. 1-39.
Chawla, N.V., Bowyer, K.W., Hall, L.O., and Kegelmeyer, W.P. 2002. "SMOTE: Synthetic Minority over-Sampling Technique," *Journal of Artificial Intelligence Research*, pp. 321-357.
Chiasson, M.W., and Davidson, E. 2004. "Pushing the Contextual Envelope: Developing and Diffusing IS Theory for Health Information Systems Research," *Information and Organization* (14:3), pp. 155-188.
Chin, W., Peterson, R., and Brown, S. 2008. "Structural Equation Modeling in Marketing: Some Practical Reminders," *The Journal of Marketing Theory and Practice* (16:4), pp. 287-298.
Chiu, C.-M., Hsu, M.-H., and Wang, E.T.G. 2006. "Understanding Knowledge Sharing in Virtual Communities: An Integration of Social Capital and Social Cognitive Theories," *Decision Support Systems* (42:3), pp. 1872-1888.
Chua, C., Cao, L., Cousins, K., and Straub, D.W. 2002. "Measuring Researcher-Production in Information Systems," *Journal of the Association for Information Systems* (3), pp. 145-215.

Chua, C.E.H., Lim, W.-K., Soh, C., and Sia, S.K. 2012. "Enacting Clan Control in Complex IT Projects: A Social Capital Perspective," *MIS Quarterly* (36:2), pp. 577-600.
Chuang, K.Y., and Yang, C.C. 2012. "Interaction Patterns of Nurturant Support Exchanged in Online Health Social Networking," *Journal of Medical Internet Research* (14:3), p. e54.
Cobb, S. 1976. "Social Support as a Moderator of Life Stress," *Psychosomatic Medicine* (38:5), pp. 300-314.
Cohen, S. 2004. "Social Relationships and Health," *American Psychologist* (59:8), pp. 676-684.
Cohen, S., and Wills, T.A. 1985. "Stress, Social Support, and the Buffering Hypothesis," *Psychological Bulletin* (98:2), pp. 310-357.
Cole, J.R., and Cole, S. 1972. "The Ortega Hypothesis," *Science* (178:4059), pp. 368-375.
Coulson, N.S., Buchanan, H., and Aubeeluck, A. 2007. "Social Support in Cyberspace: A Content Analysis of Communication within a Huntington's Disease Online Support Group," *Patient Education and Counseling* (68:2), pp. 173-178.
Coursaris, C.K., and Liu, M. 2009. "An Analysis of Social Support Exchanges in Online HIV/AIDS Self-Help Groups," *Computers in Human Behavior* (25:4), pp. 911-918.
Crane, D. 1972. *Invisible Colleges: Diffusion of Knowledge in Scientific Communities*. Chicago and London: The University of Chicago Press.
Culnan, M.J. 1986. "The Intellectual Development of Management Information Systems, 1972-1982: A Co-Citation Analysis," *Management Science* (32:2), pp. 156-172.
Culnan, M.J. 1987. "Mapping the Intellectual Structure of MIS, 1980-1985: A Co-Citation Analysis," *MIS Quarterly* (11:3), pp. 341-353.
Cutrona, C.E., and Russell, D.W. 1990. "Type of Social Support and Specific Stress: Toward a Theory of Optimal Matching," in *Social Support: An Interactional View,* B.R. Sarason, I.G. Sarason and G.R. Pierce (eds.). Oxford, England: John Wiley & Sons, pp. 319-366.
Cutrona, C.E., and Suhr, J.A. 1992. "Controllability of Stressful Events and Satisfaction with Spouse Support Behaviors," *Communication Research* (19:2), pp. 154-174.
De Nooy, W., Mrvar, A., and Batagelj, V. 2011. *Exploratory Social Network Analysis with Pajek*. New York: Cambridge University Press.
de Solla Price, D.J. 1963. *Little Science, Big Science*. New York: Columbia University Press.
de Solla Price, D.J. 1965. "Networks of Scientific Papers," *Science* (149:3683), pp. 510-515.
Deerwester, S., Dumais, S.T., Furnas, G.W., Landauer, T.K., and Harshman, R. 1990. "Indexing by Latent Semantic Analysis," *Journal of the American Society for Information Science* (41:6), pp. 391-407.
DeSanctis, G. 2003. "The Social Life of Information Systems Research: A Response to Benbasat and Zmud's Call for Returning to the IT Artifact," *Journal of the Association for Information Systems* (4:7), pp. 360-376.
Diamantopoulos, A. 2011. "Incorporating Formative Measures into Covariance-Based Structural Equation Models," *MIS Quarterly* (35:2), pp. 335-358.
Diamantopoulos, A., and Siguaw, J.A. 2006. "Formative Versus Reflective Indicators in Organizational Measure Development: A Comparison and Empirical Illustration," *British Journal of Management* (17:4), pp. 263-282.
Dijkstra, E.W. 1959. "A Note on Two Problems in Connexion with Graphs," *Numerische mathematik* (1:1), pp. 269-271.
Ding, Y., Chowdhury, G., and Foo, S. 1999. "Mapping the Intellectual Structure of Information Retrieval Studies: An Author Co-Citation Analysis, 1987-1997," *Journal of Information Science* (25:1), pp. 67-78.

Donabedian, A. 1966. "Evaluating the Quality of Medical Care," *The Milbank Memorial Fund Quarterly* (44:3), pp. 166-206.
Donabedian, A. 2005. "Evaluating the Quality of Medical Care," *Milbank Quarterly* (83:4), pp. 691-729.
Eggers, S., Huang, Z., Chen, H., Yan, L., Larson, C., Rashid, A., Chau, M., and Lin, C. 2005. "Mapping Medical Informatics Research," in *Medical Informatics*, H. Chen, S. Fuller, C. Friedman and W. Hersh (eds.). Springer US, pp. 35-62.
Ellison, N.B., Steinfield, C., and Lampe, C. 2007. "The Benefits of Facebook 'Friends:' Social Capital and College Students' Use of Online Social Network Sites," *Journal of Computer-Mediated Communication* (12:4), pp. 1143-1168.
Euske, K.J., Hesford, J.W., and Malina, M.A. 2011. "A Social Network Analysis of the Literature on Management Control," *Journal of Management Accounting Research* (23:1), pp. 259-283.
Eysenbach, G. 2008. "Medicine 2.0: Social Networking, Collaboration, Participation, Apomediation, and Openness," *Journal of Medical Internet Research* (10:3), p. e22.
Eysenbach, G., Powell, J., Englesakis, M., Rizo, C., and Stern, A. 2004. "Health Related Virtual Communities and Electronic Support Groups: Systematic Review of the Effects of Online Peer to Peer Interactions," *BMJ* (328:7449), pp. 1-6.
Faraj, S., Kudaravalli, S., and Wasko, M. 2015. "Leading Collaboration in Online Communities," *MIS Quarterly* (39:2), pp. 393-412.
Flicker, S., Haans, D., and Skinner, H. 2004. "Ethical Dilemmas in Research on Internet Communities," *Qualitative Health Research* (14:1), pp. 124-134.
Fowles, J.B., Terry, P., Xi, M., Hibbard, J., Bloom, C.T., and Harvey, L. 2009. "Measuring Self-Management of Patients' and Employees' Health: Further Validation of the Patient Activation Measure (Pam) Based on Its Relation to Employee Characteristics," *Patient Education and Counseling* (77:1), pp. 116-122.
Fox, S. 2011. "The Social Life of Health Information, 2011." Retrieved 2/1/2015, from http://www.pewinternet.org/2011/05/12/the-social-life-of-health-information-2011/
Fox, S., and Duggan, M. 2013. "Health Online 2013." Retrieved 2/1/2015, from http://www.pewinternet.org/2013/01/15/health-online-2013/
Freeman, L.C. 1979. "Centrality in Social Networks Conceptual Clarification," *Social Networks* (1:3), pp. 215-239.
Frost, J.H., and Massagli, M.P. 2008. "Social Uses of Personal Health Information within Patientslikeme, an Online Patient Community: What Can Happen When Patients Have Access to One Another's Data," *Journal of Medical Internet Research* (10:3), p. e15.
Gallivan, M., and Tao, Y. 2014. "The Value of Co-Citation Analysis for Understanding a Field's Intellectual Structure: An Application to Healthcare Information Technology (HIT) Research," in: *Proceedings of the 20th Americas Conference on Information Systems*. Savannah, Georgia, USA.
Garg, A.X., Adhikari, N.K., McDonald, H., Rosas-Arellano, M.P., Devereaux, P., Beyene, J., Sam, J., and Haynes, R.B. 2005. "Effects of Computerized Clinical Decision Support Systems on Practitioner Performance and Patient Outcomes: A Systematic Review," *JAMA* (293:10), pp. 1223-1238.
Gaskin, J., Berente, N., Lyytinen, K., and Yoo, Y. 2014. "Toward Generalizable Sociomaterial Inquiry: A Computational Approach for Zooming in and out of Sociomaterial Routines," *MIS Quarterly* (38:3), pp. 849-872.

Gefen, D., Rigdon, E.E., and Straub, D. 2011. "An Update and Extension to SEM Guidelines for Administrative and Social Science Research," *MIS Quarterly* (35:2), pp. iii-xiv.
Glandon, G.L., Smaltz, D.H., and Slovensky, D.J. 2008. *Austin and Boxerman's Information Systems for Healthcare Management*. Health Administration Press.
Gooden, R.J., and Winefield, H.R. 2007. "Breast and Prostate Cancer Online Discussion Boards: A Thematic Analysis of Gender Differences and Similarities," *Journal of Health Psychology* (12:1), pp. 103-114.
Gouldner, A.W. 1960. "The Norm of Reciprocity: A Preliminary Statement," *American Sociological Review* (25:2), pp. 161-178.
Grafton, A. 2006. "The History of Ideas: Precept and Practice, 1950-2000 and Beyond," *Journal of the History of Ideas* (67:1), pp. 1-32.
Granovetter, M.S. 1973. "The Strength of Weak Ties," *American Journal of Sociology* (78:6), pp. 1360-1380.
Greenes, R.A., and Shortliffe, E.H. 1990. "Medical Informatics: An Emerging Academic Discipline and Institutional Priority," *JAMA* (263:8), pp. 1114-1120.
Haber, M., Cohen, J., Lucas, T., and Baltes, B. 2007. "The Relationship between Self-Reported Received and Perceived Social Support: A Meta-Analytic Review," *American Journal of Community Psychology* (39:1-2), pp. 133-144.
Hair, J., Sarstedt, M., Ringle, C., and Mena, J. 2012. "An Assessment of the Use of Partial Least Squares Structural Equation Modeling in Marketing Research," *Journal of the Academy of Marketing Science* (40:3), pp. 414-433.
Hair, J.F., Hult, G.T.M., Ringle, C., and Sarstedt, M. 2013. *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*. SAGE Publications, Incorporated.
Hair, J.F., Ringle, C.M., and Sarstedt, M. 2011. "PLS-SEM: Indeed a Silver Bullet," *Journal of Marketing Theory and Practice* (19:2), pp. 139-152.
Hajli, M.N., Shanmugam, M., Hajli, A., Khani, A.H., and Wang, Y. 2014. "Health Care Development: Integrating Transaction Cost Theory with Social Support Theory," *Informatics for Health and Social Care*, pp. 1-11.
Hanneman, R.A., and Riddle, M. 2005. "Introduction to Social Network Methods." Riverside, CA: University of California, Riverside.
Hedman, J., Srinivasan, N., and Lindgren, R. 2013. "Digital Traces of Information Systems: Sociomateriality Made Researchable," in: *Proceedings of the 34th International Conference on Information Systems*. Milan, Italy.
Hersh, W. 2009. "A Stimulus to Define Informatics and Health Information Technology," *BMC Medical Informatics And Decision Making* (9:24).
Higgins, J.P., and Green, S. 2008. *Cochrane Handbook for Systematic Reviews of Interventions*. Wiley Online Library.
Hillestad, R., Bigelow, J., Bower, A., Girosi, F., Meili, R., Scoville, R., and Taylor, R. 2005. "Can Electronic Medical Record Systems Transform Health Care? Potential Health Benefits, Savings, and Costs," *Health Affairs* (24:5), pp. 1103-1117.
Hirschheim, R., Saunders, C., and Straub, D. 2012. "Historical Interpretations of the IS Discipline: An Introduction to the Special Issue," *Journal of the Association for Information Systems* (13:4), pp. i-viii.
Howison, J., Wiggins, A., and Crowston, K. 2011. "Validity Issues in the Use of Social Network Analysis with Digital Trace Data," *Journal of the Association for Information Systems* (12:12), pp. 768-797.

Huang, K.-Y., and Chengalur-Smith, I. 2014. "A Social Capital Perspective to Understand Individual Contribution of Social Support in Healthcare Virtual Support Communities," in: *Proceedings of the 47th Hawaii International Conference on System Sciences*. pp. 3489-3498.

Huang, K.-Y., Chengalur-Smith, I., and Ran, W. 2014. "Not Just for Support: Companionship Activities in Healthcare Virtual Support Communities," *Communications of the Association for Information Systems* (34:29), pp. 561-594.

Huang, K.-Y., Chengalur-Smith, S., Uzuner, Ö., Nambisan, P., and Choi, N. 2012. "Supporters in Deed–Studying Online Support Provision from the Perspective of Social Capital," *ICIS 2012*.

Huang, K.-Y., Nambisan, P., and Uzuner, Ö. 2010. "Informational Support or Emotional Support: Preliminary Study of an Automated Approach to Analyze Online Support Community Contents," in: *Proceedings of the 31st International Conference on Information Systems*. St. Louis.

Hwang, K.O., Etchegaray, J.M., Sciamanna, C.N., Bernstam, E.V., and Thomas, E.J. 2014. "Structural Social Support Predicts Functional Social Support in an Online Weight Loss Programme," *Health Expectations* (17:3), pp. 345-352.

Hwang, K.O., Ottenbacher, A.J., Lucke, J.F., Etchegaray, J.M., Graham, A.L., Thomas, E.J., and Bernstam, E.V. 2010. "Measuring Social Support for Weight Loss in an Internet Weight Loss Community," *Journal of Health Communication* (16:2), pp. 198-211.

Idler, E.L., and Benyamini, Y. 1997. "Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies," *Journal of Health and Social Behavior* (38:1), pp. 21-37.

Jamal, A., McKenzie, K., and Clark, M.J. 2009. "The Impact of Health Information Technology on the Quality of Medical and Health Care: A Systematic Review," *Health Information Management Journal* (38:3), pp. 26-37.

Japkowicz, N. 2000. "The Class Imbalance Problem: Significance and Strategies," in: *Proceedings of the 2000 International Conference on Artificial Intelligence (IC-AI'2000)*. Las Vegas, Nevada.

Jarvis, C.B., MacKenzie, S.B., and Podsakoff, P.M. 2003. "A Critical Review of Construct Indicators and Measurement Model Misspecification in Marketing and Consumer Research," *Journal of Consumer Research* (30:2), pp. 199-218.

Jha, A.K., DesRoches, C.M., Campbell, E.G., Donelan, K., Rao, S.R., Ferris, T.G., Shields, A., Rosenbaum, S., and Blumenthal, D. 2009. "Use of Electronic Health Records in Us Hospitals," *New England Journal of Medicine* (360:16), pp. 1628-1638.

Johnson, S.L., Faraj, S., and Kudaravalli, S. 2014. "Emergence of Power Laws in Online Communities: The Role of Social Mechanisms and Preferential Attachment," *MIS Quarterly* (38:3), pp. 795-808.

Jones, S.S., Rudin, R.S., Perry, T., and Shekelle, P.G. 2014. "Health Information Technology: An Updated Systematic Review with a Focus on Meaningful Use," *Annals of Internal Medicine* (160:1), pp. 48-54.

Kaplan, B.H., Cassel, J.C., and Gore, S. 1977. "Social Support and Health," *Medical Care* (15:5), pp. 47-58.

Kellermann, A.L., and Jones, S.S. 2013. "What It Will Take to Achieve the As-Yet-Unfulfilled Promises of Health Information Technology," *Health Affairs* (32:1), pp. 63-68.

Kordzadeh, N., Liu, C.Z., Au, Y., and Clark, J. 2014. "A Multilevel Investigation of Participation within Virtual Health Communities," *Communications of the Association for Information Systems* (34:26), pp. 493-512.

Kraut, R., Patterson, M., Lundmark, V., Kiesler, S., Mukophadhyay, T., and Scherlis, W. 1998. "Internet Paradox: A Social Technology That Reduces Social Involvement and Psychological Well-Being?," *American Psychologist* (53:9), pp. 1017-1031.

Krippendorff, K. 2004. *Content Analysis: An Introduction to Its Methodology*. Thousand Oaks, California: Sage Pulications, Inc.

Kuhn, T.S. 1962. *The Structure of Scientific Revolutions*. Chicago, IL USA.: University of Chicago Press.

Kulikowski, C.A., Shortliffe, E.H., Currie, L.M., Elkin, P.L., Hunter, L.E., Johnson, T.R., Kalet, I.J., Lenert, L.A., Musen, M.A., and Ozbolt, J.G. 2012. "Amia Board White Paper: Definition of Biomedical Informatics and Specification of Core Competencies for Graduate Education in the Discipline," *Journal of the American Medical Informatics Association* (19:6), pp. 931-938.

Lakey, B., and Cohen, S. 2000. "Social Support Theory and Measurement," in *Social Support Measurement and Intervention: A Guide for Health and Social Scientists*, S. Cohen, L.G. Underwood and B. Gottlieb (eds.). New York: Oxford University Press, pp. 29-52.

Landauer, T.K., Foltz, P.W., and Laham, D. 1998. "An Introduction to Latent Semantic Analysis," *Discourse Processes* (25:2-3), pp. 259-284.

Larsen, K.R., Monarchi, D.E., Hovorka, D.S., and Bailey, C.N. 2008. "Analyzing Unstructured Text Data: Using Latent Categorization to Identify Intellectual Communities in Information Systems," *Decision Support Systems* (45:4), pp. 884-896.

Lau, F., Kuziemsky, C., Price, M., and Gardner, J. 2010. "A Review on Systematic Reviews of Health Information System Studies," *Journal of the American Medical Informatics Association* (17:6), pp. 637-645.

Lazar, E.J., Fleischut, P., and Regan, B.K. 2013. "Quality Measurement in Healthcare," *Annual Review of Medicine* (64), pp. 485-496.

Lee, A.S. 2003. "Building and Testing Theory on New Organizational Forms Enabled by Information Technology," in *Exploring Patterns in Information Management: Concepts and Perspectives for Understanding IT-Related Change*, B. Sundgren, P. Mårtensson, M. Mähring and K. Nilsson (eds.). Stockholm, Sweden: The Economic Research Institute, Stockholm School of Economics, pp. 305-320.

Li, N., and Wu, D.D. 2010. "Using Text Mining and Sentiment Analysis for Online Forums Hotspot Detection and Forecast," *Decision Support Systems* (48:2), pp. 354-368.

Li, Y., and Joshi, K.D. 2012. "The State of Social Computing Research: A Literature Review and Synthesis Using the Latent Semantic Analysis Approach," in: *Proceedings of the 18th Americas Conference on Information Systems*. Seattle, Washington.

Lim, E.-P., Chen, H., and Chen, G. 2013. "Business Intelligence and Analytics: Research Directions," *ACM Transactions on Management Information Systems* (3:4), pp. 1-10.

Lin, N. 1999. "Building a Network Theory of Social Capital," *Connections* (22:1), pp. 28-51.

Lincoln, T., and Korpman, R. 1980. "Computers, Health Care, and Medical Information Science," *Science* (210:4467), pp. 257-263.

Loane, S.S., and D'Alessandro, S. 2013. "Communication That Changes Lives: Social Support within an Online Health Community for ALS," *Communication Quarterly* (61:2), pp. 236-251.

Loch, K.D., Straub, D.W., and Kamel, S. 2003. "Diffusing the Internet in the Arab World: The Role of Social Norms and Technological Culturation," *IEEE Transactions on Engineering Management* (50:1), pp. 45-63.

Lohr, K.N., and Steinwachs, D.M. 2002. "Health Services Research: An Evolving Definition of the Field," *Health Services Research* (37:1), pp. 7-9.

Lucas, H.C., Agarwal, R., Clemons, E.K., Sawy, O.A.E., and Weber, B. 2013. "Impactful Research on Transformational Information Technology: An Opportunity to Inform New Audiences," *MIS Quarterly* (37:2), pp. 371-382.

MacCallum, R.C., and Browne, M.W. 1993. "The Use of Causal Indicators in Covariance Structure Models: Some Practical Issues," *Psychological Bulletin* (114:3), pp. 533-541.

MacKenzie, S.B., Podsakoff, P.M., and Podsakoff, N.P. 2011. "Construct Measurement and Validation Procedures in MIS and Behavioral Research: Integrating New and Existing Techniques," *MIS Quarterly* (35:2), pp. 293-334.

Madden, T.J., Ellen, P.S., and Ajzen, I. 1992. "A Comparison of the Theory of Planned Behavior and the Theory of Reasoned Action," *Personality and social psychology Bulletin* (18:1), pp. 3-9.

Martin, A.B., Hartman, M., Whittle, L., and Catlin, A. 2014. "National Health Spending in 2012: Rate of Health Spending Growth Remained Low for the Fourth Consecutive Year," *Health Affairs* (33:1), pp. 67-77.

Marwell, G., and Oliver, P. 1993. *The Critical Mass in Collective Action: A Micro-Social Theory*. Cambridge University Press.

McLaughlin, M., Nam, Y., Gould, J., Pade, C., Meeske, K.A., Ruccione, K.S., and Fulk, J. 2012. "A Videosharing Social Networking Intervention for Young Adult Cancer Survivors," *Computers in Human Behavior* (28:2), pp. 631-641.

Melville, P., Gryc, W., and Lawrence, R.D. 2009. "Sentiment Analysis of Blogs by Combining Lexical Knowledge with Text Classification," in: *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Paris, France: ACM, pp. 1275-1284.

Miilunpalo, S., Vuori, I., Oja, P., Pasanen, M., and Urponen, H. 1997. "Self-Rated Health Status as a Health Measure: The Predictive Value of Self-Reported Health Status on the Use of Physician Services and on Mortality in the Working-Age Population," *Journal of Clinical Epidemiology* (50:5), pp. 517-528.

Miller, K. 2012. "Big Data Analytics in Biomedical Research," *Biomedical Computation Review*, pp. 14-21.

Mo, P.K.H., and Coulson, N.S. 2008. "Exploring the Communication of Social Support within Virtual Communities: A Content Analysis of Messages Posted to an Online HIV/AIDS Support Group," *CyberPsychology & Behavior* (11:3), pp. 371-374.

Mo, P.K.H., and Coulson, N.S. 2012. "Developing a Model for Online Support Group Use, Empowering Processes and Psychosocial Outcomes for Individuals Living with HIV/AIDS," *Psychology & Health* (27:4), pp. 445-459.

Mo, P.K.H., and Coulson, N.S. 2014. "Are Online Support Groups Always Beneficial? A Qualitative Exploration of the Empowering and Disempowering Processes of Participation within HIV/AIDS-Related Online Support Groups," *International Journal of Nursing Studies* (51:7), pp. 983-993.

Morris, T.A., and McCain, K.W. 1998. "The Structure of Medical Informatics Journal Literature," *Journal of the American Medical Informatics Association* (5:5), pp. 448-466.

Moses, H., Matheson, D.H., Dorsey, E.R., George, B.P., Sadoff, D., and Yoshimura, S. 2013. "The Anatomy of Health Care in the United States," *JAMA* (310:18), pp. 1947-1964.
Mossey, J.M., and Shapiro, E. 1982. "Self-Rated Health: A Predictor of Mortality among the Elderly," *American Journal of Public Health* (72:8), pp. 800-808.
Mulkay, M.J., Gilbert, G.N., and Woolgar, S. 1975. "Problem Areas and Research Networks in Science," *Sociology* (9:2), pp. 187-203.
Nahapiet, J., and Ghoshal, S. 1998. "Social Capital, Intellectual Capital, and the Organizational Advantage," *The Academy of Management Review* (23:2), pp. 242-266.
Nambisan, P. 2011. "Information Seeking and Social Support in Online Health Communities: Impact on Patients' Perceived Empathy," *Journal of the American Medical Informatics Association* (18:3), pp. 298-304.
Nerur, S.P., Rasheed, A.A., and Natarajan, V. 2008. "The Intellectual Structure of the Strategic Management Field: An Author Co-Citation Analysis," *Strategic Management Journal* (29:3), pp. 319-336.
Newman, M.E.J. 2001. "Scientific Collaboration Networks. II. Shortest Paths, Weighted Networks, and Centrality," *Physical Review E* (64:016132), pp. 1-7.
NLM. 2014. "Medical Informatics." Retrieved August 1, 2014, from http://www.nlm.nih.gov/tsd/acquisitions/cdm/subjects58.html
Nurullah, A.S. 2012. "Received and Provided Social Support: A Review of Current Evidence and Future Directions," *American Journal of Health Studies* (27:3), pp. 173-188.
OECD. 2013. "OECD Health Data 2013 - Frequently Requested Data." Paris: OECD Health Division.
Olson, M. 2009. *The Logic of Collective Action*. Cambridge, MA: Harvard University Press.
ONC. 2014. "What Is Health IT?" Retrieved June 1, 2014, from http://www.hrsa.gov/healthit/toolbox/oralhealthittoolbox/introduction/whatishealthit.html
Or, C.K., and Karsh, B.-T. 2009. "A Systematic Review of Patient Acceptance of Consumer Health Information Technology," *Journal of the American Medical Informatics Association* (16:4), pp. 550-560.
Petter, S., Straub, D., and Rai, A. 2007. "Specifying Formative Constructs in Information Systems Research," *MIS Quarterly* (31:4), pp. 623-656.
Pilkington, A., and Liston-Heyes, C. 1999. "Is Production and Operations Management a Discipline? A Citation/Co-Citation Study," *International Journal of Operations & Production Management* (19:1), pp. 7-20.
Pilkington, A., and Meredith, J. 2009. "The Evolution of the Intellectual Structure of Operations Management—1980–2006: A Citation/Co-Citation Analysis," *Journal of Operations Management* (27:3), pp. 185-202.
Platt, J.R. 1964. "Strong Inference," *Science* (146:3642), pp. 347-353.
Poissant, L., Pereira, J., Tamblyn, R., and Kawasumi, Y. 2005. "The Impact of Electronic Health Records on Time Efficiency of Physicians and Nurses: A Systematic Review," *Journal of the American Medical Informatics Association* (12:5), pp. 505-516.
Polites, G.L., and Watson, R.T. 2009. "Using Social Network Analysis to Analyze Relationships among IS Journals," *Journal of the Association for Information Systems* (10:8), pp. 595-636.
Porter, M.F. 1980. "An Algorithm for Suffix Stripping," *Program* (14:3), pp. 130-137.

Pratt, J., Hauser, K., and Sugimoto, C. 2012. "Defining the Intellectual Structure of Information Systems and Related College of Business Disciplines: A Bibliometric Analysis," *Scientometrics* (93:2), pp. 279-304.

Preacher, K.J., and Hayes, A.F. 2008. "Asymptotic and Resampling Strategies for Assessing and Comparing Indirect Effects in Multiple Mediator Models," *Behavior research methods* (40:3), pp. 879-891.

Preece, J. 2001. "Sociability and Usability in Online Communities: Determining and Measuring Success," *Behaviour & Information Technology* (20:5), pp. 347-356.

Procter, R. 2009. "Health Informatics." Retrieved June 1, 2014, from http://www.nlm.nih.gov/hsrinfo/informatics.html

Putnam, R.D. 1995. "Bowling Alone: America's Declining Social Capital," *Journal of Democracy* (6:1), pp. 65-78.

Raghupathi, W., and Nerur, S. 2008. "Research Themes and Trends in Health Information Systems," *Methods of Information in Medicine* (47:5), pp. 435-442.

Raghupathi, W., and Nerur, S. 2010. "The Intellectual Structure of Health and Medical Informatics," *International Journal of Healthcare Information Systems and Informatics (IJHISI)* (5:4), pp. 20-34.

Ramos-Rodríguez, A.-R., and Ruíz-Navarro, J. 2004. "Changes in the Intellectual Structure of Strategic Management Research: A Bibliometric Study of the Strategic Management Journal, 1980–2000," *Strategic Management Journal* (25:10), pp. 981-1004.

Ren, Y., Harper, F.M., Drenner, S., Terveen, L.G., Kiesler, S.B., Riedl, J., and Kraut, R.E. 2012. "Building Member Attachment in Online Communities: Applying Theories of Group Identity and Interpersonal Bonds," *MIS Quarterly* (36:3), pp. 841-864.

Roemer, M.I. 1993. *National Health Systems of the World*. Oxford University Press Oxford, UK.

Rogers, E.M. 1962. *Diffusion of Innovations*, (1st ed. ed.). New York: The Free Press.

Rogers, E.M. 1996. *Diffusion of Innovations*, (4th ed. ed.). New York: The Free Press.

Romanow, D., Sunyoung, C., and Straub, D. 2012. "Riding the Wave: Past Trends and Future Directions for Health IT Research," *MIS Quarterly* (36:3), pp. iii-A18.

Russom, P. 2011. "Big Data Analytics," in: *TDWI Best Practices Report, Fourth Quarter*.

Rust, R.T., and Huang, M.-H. 2012. "Optimizing Service Productivity," *Journal of Marketing* (76:2), pp. 47-66.

Rust, R.T., and Huang, M.-H. 2014. "The Service Revolution and the Transformation of Marketing Science," *Marketing Science* (33:2), pp. 206-221.

Scholz, U., Ochsner, S., Hornung, R., and Knoll, N. 2013. "Does Social Support Really Help to Eat a Low-Fat Diet? Main Effects and Gender Differences of Received Social Support within the Health Action Process Approach," *Applied Psychology: Health and Well-Being* (5:2), pp. 270-290.

Schuemie, M., Talmon, J., Moorman, P., and Kors, J. 2009. "Mapping the Domain of Medical Informatics," *Methods of information in medicine* (48:1), pp. 76-83.

Scott, J., and Carrington, P.J. 2011. *The SAGE Handbook of Social Network Analysis*. SAGE publications.

Shaw, J.D., Duffy, M.K., Johnson, J.L., and Lockhart, D.E. 2005. "Turnover, Social Capital Losses, and Performance," *Academy of Management Journal* (48:4), pp. 594-606.

Sidorova, A., Evangelopoulos, N., Valacich, J.S., and Ramakrishnan, T. 2008. "Uncovering the Intellectual Core of the Information Systems Discipline," *MIS Quarterly* (32:3), pp. 467-A420.

Sillence, E. 2013. "Giving and Receiving Peer Advice in an Online Breast Cancer Support Group," *Cyberpsychology, Behavior, and Social Networking* (16:6), pp. 480-485.

Singh, P.V., Tan, Y., and Mookerjee, V. 2011. "Network Effects: The Influence of Structural Capital on Open Source Project Success," *MIS Quarterly* (35:4), pp. 813-830.

Small, H. 1973. "Co-Citation in the Scientific Literature: A New Measure of the Relationship between Two Documents," *Journal of the American Society for Information Science* (24:4), pp. 265-269.

Stieglitz, S., and Dang-Xuan, L. 2013. "Emotions and Information Diffusion in Social Media—Sentiment of Microblogs and Sharing Behavior," *Journal of Management Information Systems* (29:4), pp. 217-248.

Straub, D.W. 2009. "Creating Blue Oceans of Thought Via Highly Citable Articles," *MIS Quarterly* (33:4), pp. iii-vii.

Straub, D.W., Boudreau, M.-C., and Gefen, D. 2004. "Validation Guidelines for IS Positivist Research," *The Communications of the Association for Information Systems* (13:1), pp. 380-427.

Swan, M. 2009. "Emerging Patient-Driven Health Care Models: An Examination of Health Social Networks, Consumer Personalized Medicine and Quantified Self-Tracking," *International Journal of Environmental Research and Public Health* (6:2), pp. 492-525.

Swan, M. 2012. "Health 2050: The Realization of Personalized Medicine through Crowdsourcing, the Quantified Self, and the Participatory Biocitizen," *Journal of Personalized Medicine* (2:3), pp. 93-118.

Takeda, H., Cuellar, M., Young, B., and Sainsbury, R. 2013. "Online Community Influence: A Study Using the Hirsch Metric and Social Network Analysis," in: *Proceedings of the 19th Americas Conference on Information Systems*. Chicago, Illinois: pp. 1-8.

Taylor, H., Dillon, S., and Van Wingen, M. 2010. "Focus and Diversity in Information Systems Research: Meeting the Dual Demands of a Healthy Applied Discipline," *MIS Quarterly* (34:4), pp. 647-A621.

Thackeray, R., Crookston, B.T., and West, J.H. 2013. "Correlates of Health-Related Social Media Use among Adults," *Journal of Medical Internet Research* (15:1), p. e21.

Thelwall, M., Buckley, K., and Paltoglou, G. 2012. "Sentiment Strength Detection for the Social Web," *Journal of the American Society for Information Science and Technology* (63:1), pp. 163-173.

Thelwall, M., Buckley, K., Paltoglou, G., Cai, D., and Kappas, A. 2010. "Sentiment Strength Detection in Short Informal Text," *Journal of the American Society for Information Science and Technology* (61:12), pp. 2544-2558.

Thoits, P.A. 1982. "Conceptual, Methodological, and Theoretical Problems in Studying Social Support as a Buffer against Life Stress," *Journal of Health and Social Behavior* (23:2), pp. 145-159.

Thoits, P.A. 2011. "Mechanisms Linking Social Ties and Support to Physical and Mental Health," *Journal of Health and Social Behavior* (52:2), pp. 145-161.

Thompson, T.G., and Brailer, D.J. 2004. *The Decade of Health Information Technology: Delivering Consumer-Centric and Information-Rich Health Care*. Washington, D.C.: U.S. Dept. of Health and Human Services, Office for the National Coordinator for Health Information Technology.

Uchino, B.N. 2006. "Social Support and Health: A Review of Physiological Processes Potentially Underlying Links to Disease Outcomes," *Journal of Behavioral Medicine* (29:4), pp. 377-387.

Vishwanatham, R. 1998. "Citation Analysis in Journal Rankings: Medical Informatics in the Library and Information Science Literature," *Bulletin of the Medical Library Association* (86:4), pp. 518-522.

Wagner, C.S. 2008. *The New Invisible College: Science for Development*. Washington DC: Brookings Institution Press.

Wang, X., Zhao, K., and Street, N. 2014. "Social Support and User Engagement in Online Health Communities," in *Smart Health*. Springer, pp. 97-110.

Wang, Y.-C., Kraut, R., and Levine, J.M. 2012. "To Stay or Leave?: The Relationship of Emotional and Informational Support to Commitment in Online Health Support Groups," in: *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*. Seattle, Washington, USA: ACM, pp. 833-842.

Wasko, M.M., and Faraj, S. 2005. "Why Should I Share? Examining Social Capital and Knowledge Contribution in Electronic Networks of Practice," *MIS Quarterly* (29:1), pp. 35-57.

Wellman, B., Haase, A.Q., Witte, J., and Hampton, K. 2001. "Does the Internet Increase, Decrease, or Supplement Social Capital?: Social Networks, Participation, and Community Commitment," *American Behavioral Scientist* (45:3), pp. 436-455.

Wellman, B., and Wortley, S. 1990. "Different Strokes from Different Folks: Community Ties and Social Support," *American Journal of Sociology* (96:3), pp. 558-588.

White, H.D., and Griffith, B.C. 1981. "Author Cocitation: A Literature Measure of Intellectual Structure," *Journal of the American Society for information Science* (32:3), pp. 163-171.

WHO. 2008. *Toolkit on Monitoring Health Systems Strengthening: Health Information Systems*. Geneve: WHO.

Wicks, P., Massagli, M., Frost, J., Brownstein, C., Okun, S., Vaughan, T., Bradley, R., and Heywood, J. 2010. "Sharing Health Data for Better Outcomes on Patientslikeme," *Journal of Medical Internet Research* (12:2), pp. 1-17.

Wiebe, J., Wilson, T., and Cardie, C. 2005. "Annotating Expressions of Opinions and Emotions in Language," *Language Resources and Evaluation* (39:2-3), pp. 165-210.

Woolcock, M. 1998. "Social Capital and Economic Development: Toward a Theoretical Synthesis and Policy Framework," *Theory and Society* (27:2), pp. 151-208.

World Health Organization. 1986. "The Ottawa Charter for Health Promotion," in: *First International Conference on Health Promotion*. Ottawa.

Wu, H.C., Luk, R.W.P., Wong, K.F., and Kwok, K.L. 2008. "Interpreting TF-IDF Term Weights as Making Relevance Decisions," *ACM Transactions on Information Systems (TOIS)* (26:3), pp. 1-37.

Wu, S., Chaudhry, B., Wang, J., Maglione, M., Mojica, W., Roth, E., Morton, S.C., and Shekelle, P.G. 2006. "Systematic Review: Impact of Health Information Technology on Quality, Efficiency, and Costs of Medical Care," *Annals of Internal Medicine* (144:10), pp. 742-752.

Yan, L., and Tan, Y. 2014. "Feeling Blue? Go Online: An Empirical Study of Social Support among Patients," *Information Systems Research* (25:4), pp. 690-709.

Yang, S., Lee, H., and Kurnia, S. 2009. "Social Capital in Information and Communications Technology Research: Past, Present, and Future," *Communications of the Association for Information Systems* (25:1), p. 23.
Yoo, H.-S., and Bock, G.-W. 2014. "The Impact of Social Media on Healthcare," in: *Proceedings of the 35th International Conference on Information Systems*. Auckland, New Zealand.
Zellner, A. 1962. "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias," *Journal of the American Statistical Association* (57:298), pp. 348-368.
Zhu, X., Woo, S.E., Porter, C., and Brzezinski, M. 2013. "Pathways to Happiness: From Personality to Social Networks and Perceived Support," *Social Networks* (35:3), pp. 382-393.