Adoption of Social Media Search Systems: An IS Success Model Perspective

Mandy Yan Dang
Northern Arizona University
yan.dang@nau.edu

Gavin Yulei Zhang
Northern Arizona University
yulei.zhang@nau.edu

Hsinchun Chen
University of Arizona
hchen@eller.arizona.edu

Abstract

The social media search system aims at providing an organized and integrated access and search support to a massive amount of unstructured, multilingual, user-generated content in an effective and efficient manner. Previous research on social media analytics mainly focuses on developing and applying advanced analysis methods and/or tools to make sense of the large amount of user-generated data over the Internet. Relatively little effort has been put to specifically examine the social media search system. In this study, we utilize and apply the DeLone and McLean IS Success Model to examine this type of systems. To do it, a lab experiment was conducted, and the results showed that all causal relationships, except for satisfaction to social benefit, specified in the DeLone and McLean IS Success Model hold in the context of the large-scale, social media search system. Specifically, we found that information quality and system quality associated with the system could significantly influence both users' intention to use and satisfaction toward it, both of which, in turn, had significant impacts on users' perceived individual benefit and social benefit. In addition, satisfaction could significantly influence intention to use the system.

Keywords: Social media search system, DeLone and McLean IS Success Model, lab experiment

Citation: Dang, M. Y., Zhang, G. Y. and Chen, H. (2018). " Adoption of Social Media Search Systems: An IS Success Model Perspective," *Pacific Asia Journal of the Association for Information Systems, 10*(2), pp. 55-78.
ISSN 1943-7536(print) / ISSN 1943-7544(online)

DOI: 10.17705/1PAIS.10203

Introduction

We are now in the big data era, with a huge amount of user-generated content being created and made publicly available to the Internet every day. It has been reported that 90% of the data over the Internet have been created in the past two years alone, and as to 2017, the daily generation rate of Internet data was about 2.5 quintillion bytes (one quintillion is 10^{18}) (IFL Science, 2017). In addition, the majority of these vast amount of data were generated by general Internet users on various social media sites, and there were 2.79 billion active social media users in the world as to 2017 (Stevens, 2017). In every minute of 2017, social media users generated over 456 thousand tweets on Twitter, posted over 46 thousand photos on Instagram, asked over 18 million questions about the weather forecast on The Weather Channel, shared over 527 thousand photos on Snapchat, made 600 new page edits on Wikipedia, and created over 120 pages of new professionals on LinkedIn (IFL Science, 2017).

Such user-generated content and the social media sites have changed our lives dramatically. Now, we are able to know the updates from our friends and connect with them by making comments on their posts over social networking sites, without calling or seeing them directly. The online product reviews posted by others could influence our purchase choices and decisions. Not just influencing people's personal lives, social media data also have made significant impacts on companies and the society as a whole. For example, social media marketing has become important to businesses (Akar & Topçu, 2011; Ashley & Tuten, 2015; Kaur, 2016), and a lot of companies start to use the social media channel to promote their products and services, make connections with existing customers, and attract new potential customers. In addition, some influential, social and political movement activities were started from and/or made possible via the use of social media, such as the Bank Transfer Day (https://www.facebook.com /Nov.Fifth) and the Spring Awakening Revolution in Egypt (Vargas, 2012).

To make sense of the vast amount of user-generated content, previous research has put a lot of effort in creating and/or assessing new methods and tools for conducting social media analytics (Batrinca & Treleaven, 2015; Fan & Gordon, 2014). Considering the vast amount of user-generated content, an integrated and organized way to provide an effective and efficient access to this type of data is of importance, and could potentially be used as a technical infrastructure to better support more advanced analyses on social media data. Along this line, social media search systems are created with the purpose of providing efficient and effective information search support across the huge amount of user-generated content collected from various types of social media sites. A social media search system should be able to deal with certain challenges particularly associated with user-generated content over the Web, which include the vast volume, the unstructured nature, and the multilingual issue.

Once a new type of systems is developed, it is important to investigate the adoption and acceptance from the user's point of view. In addition to focusing on the technical advancement of a system as most of the technical research does, it is also important to assess and make sure the system can be of help for the end users, since the ultimate goal of a new system is to better support users' needs. Thus, examining the success of a new system from the user's perspective is of great importance. In our case, such a system is the social media search system. Relatively little existing research has been done to assess the performance of it. To address the gap, this study aims to empirically examine the large-scale, social media search system. To do it, we leverage the well-known theory of the DeLone and McLean IS Success Model (DeLone & McLean, 1992, 2003) as our theoretical foundation. A lab experiment is conducted, and significant results are found on all causal relationships, expect for the one from satisfaction to social benefit, specified in the DeLone and McLean IS Success Model.

The remainder of this article is organized as follows. Next section presents the related literature and theoretical background, followed by a summary of the main functionality of the large-scale, social media search system. Then, details about the research model and hypotheses are provided. After that, the research method is discussed, followed by the description on data analyses and results. The article concludes with the discussion on research contributions and future research directions.

Related Literature and Theoretical Background

Social Media Analytics

Social media analytics is defined as the development and evaluation of informatics tools and frameworks that can be used to collect, monitor, analyze, summarize, and visualize social media data with the purpose of extracting and identifying useful patterns and intelligence (Fan & Gordon, 2014). To conduct social media analytics, existing research has leveraged various types of techniques, including sentiment analysis (or referred to as opinion mining), topic modeling, social network analysis, trend analysis, and visualization (or referred to as visual analytics) (Fan & Gordon, 2014).

Sentiment analysis, or sometimes called opinion mining, is the automatic detection and extraction of user sentiments or opinions from text sources by utilizing computational linguistics methods (Li & Wu, 2010; Pang & Lee, 2008). Research in sentiment analysis typically focuses on either determining whether a textual document is objective or subjective (Pang, Lee, & Vaithyanathain, 2002), or whether it contains positive or negative sentiments (Wiebe, Wilson, Bruce, Bell, & Martin, 2004). In terms of the granularity level of sentiment analysis, it can be either for the whole textual documents, or for sentences, or even for phrases (Pang & Lee, 2008). To enable sentiment analysis, two types of methods are typically used, including the machine learning approach and the semantic orientation approach (Liu, 2007; Pang & Lee, 2008). The machine learning approach treats sentiment analysis as a text classification problem to automatically put textual pieces into different groups with different sentiment orientations, in most cases, positive and negative. Any text classification algorithm can be employed, such as Naïve Bayes and Support Vector Machines (SVM) (Liu, 2007; Pang & Lee, 2008). The semantic orientation approach performs classification based on positive and negative sentiment words and phrases contained in the data collection, and often relies on external knowledge resources (Liu, 2007; Pang & Lee, 2008). Previous research has conducted sentiment analysis on social media data. For example, Fang and Zhan (2015) performed sentiment analysis on product reviews collected from Amazon.com. They proposed a process for doing it based on the machine learning approach, and conducted both the sentence-level and review-level analyses. In another study, Saif, He, Fernandez, and Alani (2016) proposed a new method for sentiment analysis on Twitter, based on the semantic orientation approach. Specifically, they improved the way of creating the lexicon used for the classification. By comparing against three existing lexicons on three Twitter datasets, their method achieved the best performance for both entity-level and tweet-level sentiment analysis in the detection of subjectivity (i.e., neutral vs. polar) and polarity (i.e., positive vs. negative).

Topic modeling aims at detecting dominant themes or topics from a large collection of textual documents (Fan & Gordon, 2014). To conduct it, advanced statistical and/or machine learning techniques need to be developed and utilized. Topic model has been applied on various types of social media data. For example, Samtani, Chinn, and Chen (2015) conducted topic modeling on hacker forums. They proposed a research framework for hacker asset analysis, and demonstrated its performance on five hacker forums. By leveraging the Latent Dirichlet Allocation (LDA) method, which is a statistical technique used to model latent topics for textual documents in a hierarchical manner, topics were automatically extracted from messages posted on those five forums. In another study, Resnik et al. (2015) utilized three advanced, supervised topic modeling methods to detect topics related to depression messages from Twitter. The three techniques they utilized were supervised LDA, supervised anchor topic modeling, and supervised nested LDA. In a more recent study, Pu, Wu, and Yuan (2017) developed a new method, called User Item Sentiment Topic (UIST), for conducting topic modeling on online reviews. Different from existing methods that only modeled the text, their method also considered the user (who expressed the opinions) and the item (which the opinion was expressed on). The new method incorporated users and items for topic modeling and could produce topic–word, user–topic, and item–topic distributions simultaneously.

Social network analysis is the process of identifying and examining the social structures through the use of networks and graph theories (Otte & Rousseau, 2002). A network or graph is depicted as a set of nodes connected by links. The nodes could be people or any other things/roles, and the links represent the relationships or interactions among them (Fan & Gordon, 2014). A social network is a powerful way to show the topological structure of data being analyzed, as well as to identify important and influential nodes in such a structure. Several important measures have been developed to characterize individuals' roles in a network, such as degree, betweenness, and closeness (Wasserman & Faust, 1994). For a given node in the network, its degree means the number of direct links it has; its betweenness is the number of the shortest paths between any two nodes, which passes through it (called geodesics); and its closeness refers to the number of all the geodesics between it and every other node. Previous research has applied social network analysis on different contexts in order to study different phenomenon, on both traditional textual data and social media data, such as co-authorships on literature and patent data (Chen & Roco, 2008), consumer interactions in online blogs (Chau & Xu, 2012), and communication patterns of travel-related electronic word-of-mouth on social networking sites (Luo & Zhong, 2015). For example, when studying blogger interactions in online communities, Chau and Xu (2012) conducted two case studies on blogs related to iPad and Starbucks, respectively. For each case study, they reported the detailed results on content analysis, interaction networks, central bloggers, and implicit communities by performing social network analysis.

Trend analysis aims to use the existing data on hand to predict future outcomes and behaviors (Fan & Gordon, 2014). It has been applied on different types of social media data to examine various scenarios. For example, Mathioudakis and Koudas (2010) developed an automatic system, called TwitterMonitor, for trend detection over Twitter stream data. The proposed system performed trend analysis in three steps, including the identification of bursty keywords that were words suddenly appeared in tweets at an unusually high rate, the grouping of bursty keywords into trends based on their co-occurrence patterns, and the extraction of additional information from tweets associated with each identified trend with the purpose of further discovering some interesting patterns. In another study, Liu, Omar, Liou, Chi, and Hsu (2015) developed a novel event-based recommendation approach for online blogs, by combining trend analysis and personal preference. The purpose was to help blog readers to effectively discover the most useful information that was beneficial or of interest to them, from emerging and popular events mentioned in a large number of blog articles. By comparing with three widely used traditional approaches, including content-based filtering, item-based collaboration filtering, and user-based collaboration filtering, their proposed approach achieved the highest quality in terms of recommendation.

Visualization can help better make sense of a huge amount of data points (Fan & Gordon, 2014). It typically is used to present information search or analysis results to users in a cognitively efficient, intuitive, easily comprehensible, and transparent manner (Börner, Chen, & Boyack, 2003; Zhu & Chen, 2005). As classified by Shneiderman (1996), in general, there are seven types of information visualization representations, including one-dimensional (1D), two-dimensional (2D), three-dimensional (3D), multi-dimensional, tree, network, and temporal representations. The 1D representation is used to show abstract information using one-dimensional visual objects, and display them on the screen in a linear or a circular manner. With a 2D representation, information is represented as two-dimensional visual objects, with x and y axes. A 3D representation reveals information as three-dimensional visual objects, with x, y and z axes. A multi-dimensional representation leverages a three-dimensional or a two-dimensional space, often projecting document clusters or themes into that space via the use of some dimensionality reduction algorithms. A tree representation is usually employed to show hierarchical relationships among objects. A network representation is often used when a simple tree structure is insufficient for depicting complex relationships. A temporal visualization can organize information according to the temporal sequence, with location and animation being commonly used. Previous research has utilized visualization techniques to better make sense of social media data. For example, Cao et al. (2015) developed an advanced system framework for spatiotemporal analysis of location-based social media data. The visualization component of the system framework provided interactive flow mapping interfaces through which users could explore the geographical dynamics of movement (organized as spatiotemporal data cubes) retrieved from location-based social media data. To demonstrate the performance and effectiveness of the proposed framework, they applied it on Twitter steam data for both single-source (i.e., based on one region) and multiple-source (i.e., for multiple regions) analyses. In another study, Scharl et al. (2016) developed Westeros Sentinel, a Web-based intelligence system, that could capture and analyze both news and social media data about Game of Thrones, an American drama television series. The system extracted data from the Web site of Anglo-American news media, as well as from four social media sites, including Twitter, Facebook, Google+ and YouTube. An advanced visualization component was embedded in the system, presenting to the user as an interactive dashboard. The dashboard display enabled users to keep track of how often characters were mentioned by journalists and reviewers, as well as to identify concepts associated with the unfolding storyline and each new episode.

The Focus of This Study in Social Media Analytics

As argued by Fan and Gordon (2014), the entire process or spectrum for conducting social media analytics has three stages, including "capture," "understand," and "present," with each of the latter ones being built upon the success of the former one(s). A great body of existing research on social media analytics has focused on developing, applying, and/or assessing advanced methods and/or frameworks based on the techniques discussed in the above sub-section, with the purpose of better making sense of the user-generated content over the Internet. The majority of those studies belong to the "understand" and "present" stages of the overall process for social media analytics.

However, relatively less effort has been made to the "capture" stage, which is the foundation of the other two stages. The success of the "capture" stage could play an important role in influencing the performance of the "understand" and "present" stages, thus cannot be overlooked. Although the main focus of this foundation stage is not to directly analyze social media data by using the advanced techniques discussed in the above sub-section, it is still an essential stage that can provide a better organization and access of the massive amount of "noisy" data generated by the general public over the Web, which can later be utilized by the two higher-level stages for further data processing and analysis.

As mentioned by Fan and Gordon (2014), two of the most important tasks in the "capture" stage of social media analytics are: the gathering of data from various sources, and the preprocessing of the data being collected. To perform these two tasks, one way is to adopt the information systems approach, by creating a Web-based system that can effectively and efficiently integrate user-generated content from different social media sources and present it to the user in an organized way. Along this line, the social media search system is developed. In our previous research (Dang et al., 2014), we have created a large-scale, social media search system which provides an integrated and consistent access to user-generated content collected from various social media sites, and enables information retrieval and presentation to the end user via its search functions (including searching information within one particular data source, and across all data sources embedded in the system). In this study, we aim to empirically investigate the large-scale, social media search system by using the DeLone and McLean IS Success Model, which is one of most well-known theories for assessing the success in the adoption of and the usability about information systems, as our theoretical foundation. To the best of our knowledge, little research has been done to empirically examine this relatively new type of systems in order to understand their adoption or usability.

Theoretical Lens: the DeLone and McLean IS Success Model

Through an extensive review of definitions on IS success and their corresponding measures, DeLone and McLean classified them into six major categories, including system quality, information quality, use, user satisfaction, individual impact, and organizational impact, and further specified their causal relationships to develop the IS Success Model (DeLone & McLean, 1992). Ten years later, they refined the model based on the results of many validation attempts, by adding the service quality dimension and combining individual impact and organizational impact into one construct which is net benefits (DeLone & McLean, 2003). The updated model states that information quality, system quality, and service quality directly impact usage (including usage intentions and system use) and user satisfaction, which, in turn, influence net benefits. As a feedback loop, net benefits can further influence both usage intentions and user satisfaction.

Over the past decades, the DeLone and McLean IS Success Model has been applied to assess various types of information systems, and has been tested and validated in different contexts. In some recent studies, it has been adopted to examine systems such as the e-commerce website (Hsu, Chang, Chu, & Lee, 2014), the e-government system (Rana, Dwivedi, Williams, & Weerakkody, 2015), the m-banking system (Tam & Oliveira, 2016), the digital library (Alzahrani, Mahmud, Ramayah, Alfarraj, & Alalwan, 2017), the nursing information system (Lin, 2017), and the e-learning system (Aparicio, Bacao, & Oliveira, 2017).

For example, when investigating the online group-buying behavior via the e-commerce website, Hsu et al. (2014) found that perceived quality of website, perceived quality of sellers, satisfaction with website, and satisfaction with sellers could significantly influence users' repurchasing intention. In addition, perceived quality of website and perceived quality of sellers had significant impacts on users' satisfaction with website and satisfaction with sellers, respectively. Rana et al. (2015) adopted the DeLone and McLean IS Success Model to examine the e-government system, particularly the online public grievance redressal system. They found that all three types of quality (i.e., system quality, information quality, and service quality) had significant impacts on users' perceived satisfaction toward this type of systems. Tam and Oliveira (2016) used the DeLone and McLean IS Success Model as a theoretical lens to investigate users' m-banking behavior, and found that all three types of quality could significantly influence user satisfaction toward m-banking. However, no significantly enough impacts were found about the three types of quality on the use of m-banking.

In a more recent study, Alzahrani et al. (2017) applied the DeLone and McLean IS Success Model to investigate the success of the digital library. The data analysis results showed that all three types of quality had significant impacts on both users' satisfaction toward the digital library and their intention to use it, which, in turn, significantly influenced users' actual usage behavior. When leveraging the DeLone and McLean IS Success Model to understand the nursing information system adoption, Lin (2017) found that system quality (in the alignment with both usefulness and ease of use) and information quality (in the alignment with both usefulness and ease of use) had significant impacts on nurses' satisfaction with the system. However, service quality was found to be a significant indicator of satisfaction only when being in the alignment with ease of use, but not with usefulness. The DeLone and McLean IS Success Model also was recently applied in the context of the e-learning system, and it was found that all three types of quality significantly influenced users' satisfaction toward the system, and both information quality and service quality had significant impacts on use (no significant result was found from system quality to use though) (Aparicio et al., 2017). In addition, both use and satisfaction could significantly influence individual impact, which was defined as the individual student's perception of the impact from using the e-learning system in terms of his/her learning performance.

System Description

As pointed out in previous literature, different from data in traditional sources (such as news articles, scientific publications, and official reports), user-generated content is always "noisy" (Cao et al., 2015; Fan & Gordon, 2014). This type of data are noisy in two ways: the massive amount available over the Internet, and their unstructured nature (Cao et al., 2015). In addition, according to the most recent report provided by the Internet World Stats (http://www.internetworldstats.com/stats.htm), as of 2017, at least more than 80% of the Internet users were from non-English speaking countries, indicating that a vast amount of user-generated content was actually written in languages other than English. Therefore, an effective social media search system should be able to deal with the two noisy aspects of the user-generated content, as well as take into account the multilingual nature of it. As mentioned in the previous Section, in our previous research, we have developed a large-scale, social media search system that can deal with all three issues. In this study, we use it as our study site. The general process of the system framework is summarized in Figure 1.

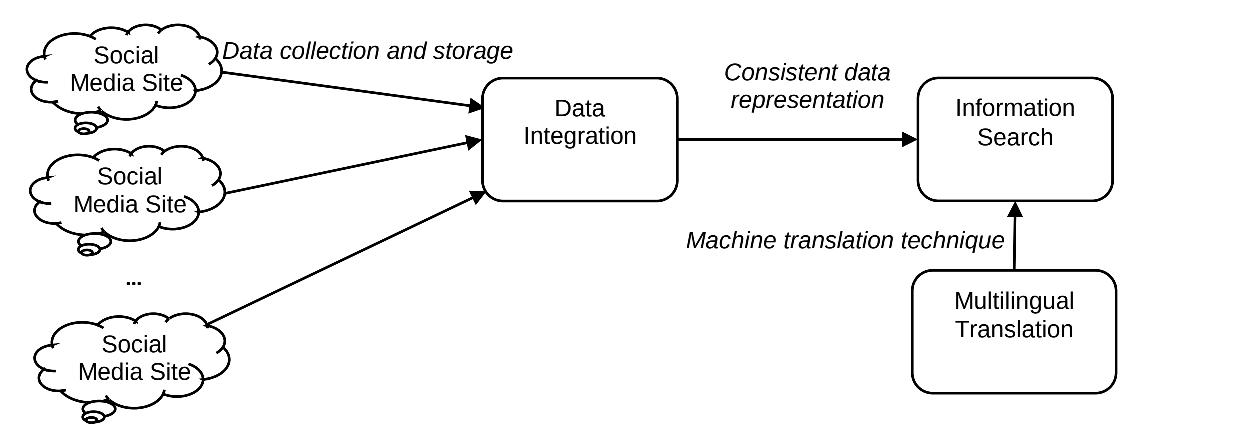

Figure 1 - The General Process of the System Framework of the Large-Scale, Social Media Search System

As shown in Figure 1, the system has three major functions: data integration, information search, and multilingual translation. To develop the data integration function, automatic data collection programs were created, which could collect user-generated content from various social media sites. Considering the massive amount of data typically associated with each social media site, to make the data collection process efficient, both complete and incremental data collection programs were utilized. The complete data collection program was run when a social media site was newly added to the system. After that, the incremental data collection program was used on a regular basis (such as weekly) to only collect user-generated content that was posted after the last update. Further, to better organize the unstructured data collected from various social media sites, a consistent data representation format was used, and the data were stored in a large-scale, relational database with a unified design schema.

By using the data processed by the data integration function, two types of information search functions were created and enabled, including information search within a particular data source (i.e., user-generated content collected from one given social media site), and information search across all data sources embedded in the system. A consistent, keyword-based, search interface design was used for both types of search functions. Logical operations (i.e., “AND” or “OR”) were enabled to allow the specification of multiple keywords in the conduction of searching.

The information search function also was supported by the multilingual translation function, which provided real-time processing of user-generated content written in different languages. The machine translation technique was used to implement this function, which could conduct translations across over 80 different languages. When searching information in a non-English data source, the system allowed the user to specify the search terms in either the original language or in English. If the search terms were specified in English, the multilingual translation function would convert them to the equivalent terms in the original language and then triggered the information search function. This process was automatically conducted without user awareness. Then, the search results were returned and displayed back to the user in both the original language and its English translation.

Based on the framework, a large-scale, social media search system was created, with a massive amount of user-generated content (over 13 million postings) in different languages (such as English, Arabic, French, German, and Russian) from 29 Web forums related to homeland security. We chose this domain because of its importance to the lives of all individuals as well as our society as a whole, and the emphasis and effort (in terms of money and personal) that our government has put onto this area. However, the proposed framework is generic, and can be applied to create social media search systems for any domains.

Research Model and Hypotheses

In this study, we aim to assess the large-scale, social media search system by leveraging the DeLone and McLean IS Success Model as the theoretical lens. To better fit the nature of the social media search system and the context of the study, our proposed research model (shown in Figure 2) is different from the updated DeLone and McLean IS Success Model in several ways, including: excluding "service quality," utilizing a simplified usage construct which only contained "intention to use" (without "use"), creating and incorporating five first-order constructs for system quality, and splitting the construct "net benefits" into two – "individual benefit" and "social benefit."

Information quality is defined as the quality of information that a system is able to store, deliver, and/or create, and system quality is defined as the overall quality of the system itself (DeLone & McLean, 1992, 2003). Information quality typically measures the semantic success of a system, while system quality generally measures the technical success of it. In our context, information quality is about the quality of the massive amount of multilingual, user-generated content that the large-scale, social media search system organizes, stores, and presents, while system quality is about the quality toward the three major functions that the system enables and provides to its users. Service quality is not included in our research model, since social media search systems are a relatively new type of systems, and not widely available or being commercialized with dedicated support teams that focus on providing related services. Therefore, in this study, we only focus on the information quality and system quality dimensions of the large-scale, social media search system. For the same reason, since it is still in the research stage, without a lot of commercialized development and/or adoption yet (thus lacking the actual usage experience from the general public), the construct of use (which is in the DeLone and McLean IS Success Model) is not included in our research model.

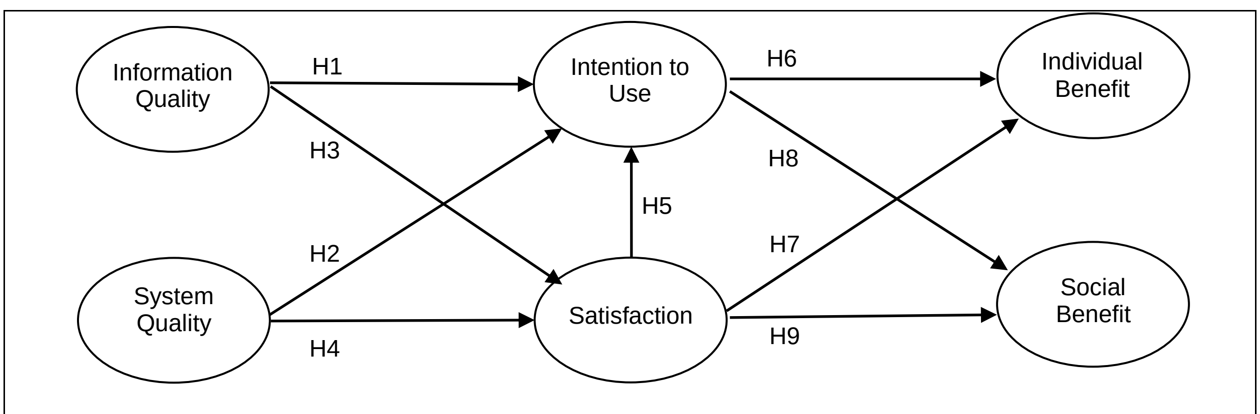

Figure 2 - The Research Model

Intention to use refers to an individual's willingness to use an information system in the future, and satisfaction means the positive idea or experience associated with the direct use of a system (DeLone & McLean, 2003; Venkatesh, Morris, Davis, & Davis, 2003). In our context, both constructs are about users' related perceptions toward the large-scale, social media search system. When examining different types of information systems, previous research found significant impacts of information quality and system quality on intention to use (Alzahrani et al., 2017; Hsu et al., 2014) and satisfaction (Alzahrani et al., 2017; Aparicio et al., 2017; Lin, 2017). In our context, we expect those relationships, as specified in the DeLone and McLean IS Success Model, will also hold for the large-scale, social media search system. Specifically, if a user perceives the quality of data stored, maintained, and presented by the system is of high quality, and the quality of the overall system functionality is high, he/she would be more willing to use the social media search system in the future, and would be more satisfied in using the system to conduct information searching. Thus, we hypothesize:

> *H1: Information quality can positively influence users' intention to use the large-scale, social media search system.*

> *H2: System quality can positively influence users' intention to use the large-scale, social media search system.*

> *H3: Information quality can positively influence users' satisfaction on the large-scale, social media search system.*

> *H4: System quality can positively influence users' satisfaction on the large-scale, social media search system.*

The DeLone and McLean IS Success Model also specifies a causal relationship from satisfaction to intention to use. Such a relationship has been validated in existing literature on various types of information systems (Alzahrani et al., 2017; Hsu et al., 2014). In our study, we posit that it is still true when applying to the context of the large-scale, social media search system. That is, if a user is more satisfied with using the social media search system to find information in user-generated content, he/she will be more willing to use it in the future to conduct this type of information searching. Therefore, we hypothesize:

> *H5: Satisfaction can positively influence users' intention to use the large-scale, social media search system.*

In this study, instead of using one construct, which is net benefits, as the dependent variable, we specifically look into two levels of potential benefits that the large-scale, social media search system could bring to the users, including individual benefit and social benefit. In our context, individual benefit is about a user's perception toward the level of value the system can provide to himself/herself, and social benefit means a user's perception on the overall value that the system can bring to our society and nation as a whole. The DeLone and McLean IS Success Model states that both intention to use and satisfaction can influence net benefits. In our case, we propose that intention to use and satisfaction have direct impacts on both individual benefit and social benefit. Previous research has validated the positive impacts of intention to use and satisfaction on individual benefit/impact (Aparicio et al., 2017). However, much fewer studies have been seen to specifically investigate social benefit in their proposed research models. In this study, we posit that both users' intention to use the large-scale, social media search system, and their satisfaction toward it, can influence their perceptions on the individual benefit and social benefit the system could potentially provide. Thus, we hypothesize:

> *H6: Users' intention to use the large-scale, social media search system can positively influence their perception on*

the individual benefit the system can provide.

H7: Users' satisfaction on the large-scale, social media search system can positively influence their perception on the individual benefit the system can provide.

H8: Users' intention to use the large-scale, social media search system can positively influence their perception on the social benefit the system can provide.

H9: Users' satisfaction on the large-scale, social media search system can positively influence their perception on the social benefit the system can provide.

Research Method

Research Process

The research method used in this study is the lab experiment. To empirically examine the large-scale, social media search system, as well as to test the research model, the system used in this study is the social media search system we created with data sources related to homeland security. Our subjects were senior-level students from a police university in Taiwan. After graduation, the majority of them would become either policemen/policewomen or security analysts. Therefore, we believe they were an appropriate group of subjects for this study, who could potentially need to use or have interest in using this type of systems in their future careers. In total, 78 subjects participated in the study, including 67 males and 11 females.

The process of the lab experiment included several steps. First, the subjects were informed that they were participating in a research study, and their job was to use a new type of system – a social media search system – to complete several information search tasks. Then, a brief introduction of the system was given to the subjects to show them where to find different functions to use. After that, each participant got a package. The first part of the package included four tasks they needed to perform and the second part of the package listed the questionnaire items, shown in 7-Likert scale, for them to provide their ratings based on their system usage experience. To ensure the success of their participation as well as the quality of data they provided based on their perceptions associated with using the system to perform information search tasks, all four tasks were presented to the subjects in Chinese which is their mother language. The tasks asked them to search for information across both English and Arabic data sources embedded in the system, through the use of the search function and multilingual translation function provided in the system. The measurement items in the questionnaire also were translated to Chinese, in order to make sure the subjects could fully understand them and provided choices/answers based on their best knowledge. Back translation was then conducted to ensure the quality of the translation.

Measures

To measure information quality, we adopted the set of seven items proposed and developed by Rai, Lang, and Welker (2002), with minor wording changes to fit our context. To measure intention to use and satisfaction, we adopted the related items from Bhattacherjee (2001). Measures on individual befit were adapted from Wixom and Watson (2001), and measures on social benefit were adapted from Seddon (1997), with minor modifications to fit our context.

As to measuring system quality, previous research argued that ease of use was an important component of it, and had leveraged the measurement items on ease of use to measure system quality (Doll & Torkzadeh, 1998; Rai et al., 2002). In this study, to obtain a more comprehensive assessment of different aspects related to the quality of the system itself, we adopted the Questionnaire for User Interaction

Satisfaction (QUIS 7.0, http://lap.umd.edu/quis/) (Chin, Diehl, & Norman, 1988; Kaipio et al., 2017; Naeini & Mostowfi, 2015), which is a multi-faceted scale consisting of items specific to users' assessments of their interactions with a system in terms of overall reactions (QUISA), screen factors (QUISB), terminology and system information (QUISC), learning (QUISD), and system capabilities (QUISE). Specifically, we modelled system quality as a second-order construct, with the five dimensions of QUIS as its formative, first-order constructs. All measurement items used in this study are listed in Appendix A.

Data Analyses and Results

Based on the concepts and following the original measurement sources, indicators of all constructs in the research model, except for system quality, are modeled as reflective measures. For system quality, the five first-order constructs are its formative indicators, since each of them contributes to one certain aspect of the second-order, latent construct of system quality. Well, the measures for each of the first-order constructs are reflective ones.

To conduct detailed data analyses, structural equation modeling (SEM) techniques are used. Specifically, Smart PLS 2.0 (M3) beta (Ringle, Wende, & Will 2005) is utilized. We present the detailed data analyses and results for both the measurement model and structural model in the following two sub-sections.

Measurement Model Assessment

Table 1 shows the reliability test results of all first-order reflective constructs. All loadings (except for IQ4) are above the minimum value of 0.4 (Hair, Anderson, Tatham, & Black, 1998), and are statistically significant at the 0.01 level. IQ4 is dropped for later analysis. In addition, since QUISC5 doesn't pass the more stringent threshold of 0.7 (Au, Ngai, & Cheng, 2008; Hair et al., 1998), it is also dropped for later analysis. The Cronbach's Alpha values for all constructs are greater than the 0.7 guideline (Au et al., 2008; Hair et al., 1998).

Table 2 shows the descriptive statistics, and Table 3 shows the composite reliability, average variance extracted (AVE), square root of AVE, and correlations among constructs. The composite reliability values of all constructs are above the recommended level of 0.70, indicating adequate internal consistency between items (Au et al., 2008; Hair et al., 1998). Convergent validity is demonstrated as the AVE values for all constructs are higher than the suggested threshold value of 0.50, which is the same as the requirement of the square root of AVE to be at least 0.707 (Gefen, Straub, & Boudreau, 2000). Comparing the square root of AVE with the correlations among the constructs indicates that each construct is more closely related to its own measures than to those of other constructs, and discriminant validity is therefore supported (Chin, 1998).

Table 1 - Reliability Test Results

Construct	Cronbach's Alpha	Item	Loading	T-stats
QUISA	0.960	QUISA1	0.909	46.217 *
		QUISA2	0.942	86.831 *
		QUISA3	0.932	82.031 *
		QUISA4	0.931	78.625 *
		QUISA5	0.875	31.233 *
		QUISA6	0.885	29.444 *
QUISB	0.916	QUISB1	0.858	26.485 *
		QUISB2	0.891	40.237 *
		QUISB3	0.908	38.640 *
		QUISB4	0.916	50.735 *
QUISC	0.901 (0.907**)	QUISC1	0.814	20.014 *
		QUISC2	0.891	43.336 *
		QUISC3	0.849	25.873 *
		QUISC4	0.839	25.617 *
		QUISC5 (dropped)	*0.682*	9.171 *
		QUISC6	0.831	28.696 *
QUISD	0.943	QUISD1	0.879	40.814 *
		QUISD2	0.875	33.889 *
		QUISD3	0.923	74.408 *
		QUISD4	0.892	35.891 *
		QUISD5	0.840	23.173 *
		QUISD6	0.878	38.160 *
QUISE	0.826	QUISE1	0.757	14.793 *
		QUISE2	0.790	13.555 *
		QUISE3	0.835	21.381 *
		QUISE4	0.857	29.809 *
Information Quality (IQ)	0.893 (0.939**)	IQ1	0.896	39.604 *
		IQ2	0.910	42.168 *
		IQ3	0.894	45.796 *
		IQ4 (dropped)	*0.115*	0.804
		IQ5	0.840	19.685 *
		IQ6	0.859	30.486 *
		IQ7	0.855	30.350 *
Intention to Use (IU)	0.948	IU1	0.959	112.651 *
		IU2	0.966	139.282 *
		IU3	0.929	58.594 *
Satisfaction (SAT)	0.967	SAT1	0.957	109.488 *
		SAT2	0.947	91.718 *
		SAT3	0.960	119.791 *
		SAT4	0.951	100.368 *
Individual Benefit (IB)	0.966	IB1	0.955	75.428 *
		IB2	0.974	178.038 *
		IB3	0.972	174.841 *
Social Benefit (SB)	0.938	SB1	0.941	76.761 *
		SB2	0.951	90.757 *
		SB3	0.937	62.017 *

Notes: * Significant at the 0.01 level. ** Reliability value after a certain item is dropped.

Table 2 - Descriptive Statistics		
Construct	**Mean**	**Std. Dev.**
Information Quality	4.65	0.95
System Quality - QUISA. Overall Reactions to the System	4.53	1.05
System Quality - QUISB. Screen	4.72	1.10
System Quality - QUISC. Terminology and System Information	4.56	0.90
System Quality - QUISD. Learning	4.48	1.14
System Quality - QUISE. System Capabilities	4.79	0.74
Satisfaction	4.60	1.14
Intention to Use	4.73	1.29
Individual Benefit	4.62	1.22
Social Benefit	4.85	1.11

Table 3 - Validity Test Results[1]

Construct	Composite Reliability	AVE	IB	IQ	IU	QUISA	QUISB	QUISC	QUISD	QUISE	SAT	SB
IB	0.978	0.936	**0.967**									
IQ	0.952	0.767	0.787	**0.876**								
IU	0.966	0.905	0.908	0.785	**0.951**							
QUISA	0.968	0.833	0.815	0.840	0.815	**0.913**						
QUISB	0.941	0.798	0.776	0.761	0.772	0.861	**0.894**					
QUISC	0.931	0.729	0.746	0.744	0.770	0.789	0.836	**0.854**				
QUISD	0.954	0.777	0.787	0.729	0.790	0.817	0.813	0.777	**0.882**			
QUISE	0.884	0.657	0.692	0.641	0.727	0.706	0.707	0.778	0.767	**0.811**		
SAT	0.976	0.910	0.860	0.810	0.877	0.868	0.801	0.770	0.831	0.736	**0.954**	
SB	0.960	0.889	0.752	0.652	0.754	0.687	0.666	0.669	0.607	0.596	0.686	**0.943**

Notes: Diagonal elements in bold case are the square root of average variance extracted (AVE) by latent constructs from their indicators; off-diagonal elements are correlations among latent constructs.

[1] We acknowledge that some corrections are relatively high. However, the results indicate the support of both convergent and discriminant validity. We further conducted the variance inflation factors (VIF) tests for the inner model, and all VIF values were lower than the common threshold of 10 (Kleinbaum, Kupper, Muller, & Nizam, 1998; Lee & Xia, 2010; Seder & Gable, 2010).

Structural Model Assessment

Figure 3 shows the research model testing results. Except for the relationship between satisfaction and social benefit, all causal paths in the model were statistically significant. Specifically, it was found that users' perceptions toward both the information quality and system quality of the large-scale, social media search system could significantly influence their intention to use the system, with path coefficients of 0.126 and 0.299, respectively. Thus, H1 and H2 were supported. In addition, both types of quality had significantly positive impacts on users' satisfaction on the system (with path coefficients of 0.246 and 0.685, respectively), in the support of H3 and H4. Users' satisfaction on this type of systems was found to have significantly positive impact on their intention to use the system, with the path coefficient of 0.509. Therefore, H5 was supported. The analysis results also showed that both users' intention to use the large-scale, social media search system, and their satisfaction on the system, could significantly influence their perceptions about the individual benefit that the system could provide, with path coefficients of 0.670 and 0.272, respectively. Thus, H6 and H7 were both supported. Further, users' intention to use the system was found to have a significantly positive impact on their perceptions about the social benefit associated with the system (with path coefficient of 0.658), in the support of H8. However, no statistically significant result was found about the impact of satisfaction on social benefit. Thus, H9 was not supported. In addition, the coefficients between the five first-order constructs (i.e., QUISA to QUISE) and the corresponding second-order construct of system quality also are shown in Figure 3. All relationships were statistically significant. This indicates that all five aspects related to the system, including the overall reactions, screen factors, terminology and system information, learning, and system capabilities, were important contributors to the overall quality of the system as perceived by the user.

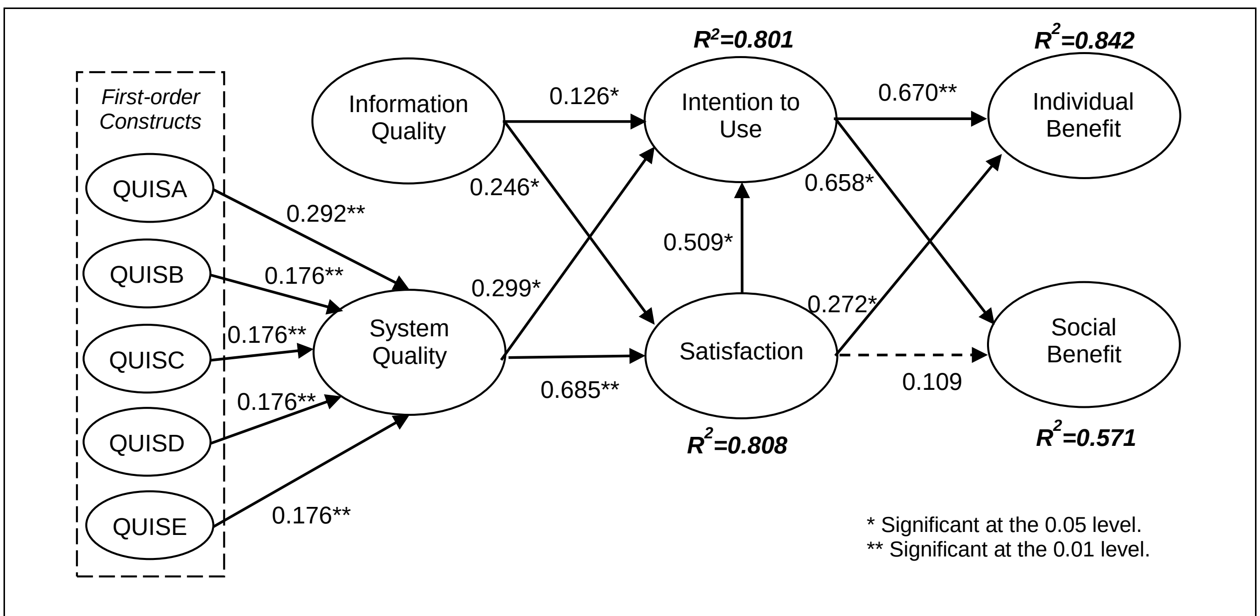

Figure 3 - Research Model Testing Results

The R-squared value associated with intention to use is 0.801, meaning that information quality, system quality, and satisfaction together explained 80.1 percent of its variance. The R-squared value associated with satisfaction is 0.808, indicating that both types of quality together explained 80.8 percent of the variance of satisfaction. In addition, the R-squared values associated individual benefit and social benefit are 0.842 and 0.571, respectively. These results indicate that the significant factor(s) associated with each of these dependent constructs explained 84.2 and 57.1 percent of the variance, respectively.

Discussion

This study makes several contributions. First, we empirically examined the large-scale, social media search system, which is a relatively new type of systems aiming at solving tasks related to data gathering and preprocessing in the "capture" stage of social media analytics. Such a type of systems can provide an organized and consistent way to efficiently and effectively store, present, and enable users' information searching across the massive amount of unstructured, multilingual, user-generated content over the Internet. It can also work as a technical infrastructure, upon which more advanced social media analytics tasks (i.e., those in the "understand" and "present" stages) can be better conducted. Our study is among one of the earliest to examine this type of systems. To do it, we leveraged the DeLone and McLean IS Success Model as our theoretical lens, and conducted a lab experiment using a social media search system with over 13 million postings in different languages that were collected from various social media sites. The data analysis results showed that, except for the path from satisfaction to social benefit, all causal relationships specified in the research model were statistically significant in the context of the large-scale, social media search system. Specifically, information quality and system quality were found to have significant impacts on satisfaction, the three of which then significantly influenced intention to use; both intention to use and satisfaction could significantly influence individual benefit; and intention to use was the significant, influential factor on social benefit.

Second, we tested and validated the DeLone and McLean IS Success Model in a new type of systems (i.e., the large-scale, social media search system), which hasn't been done in previous literature. Recent studies have applied and validated the DeLone and McLean IS Success Model in various types of systems, such as the e-commerce website (Hsu et al., 2014), the e-government system (Rana et al., 2015), the m-banking system (Tam & Oliveira, 2016), the digital library (Alzahrani et al., 2017), the nursing information system (Lin, 2017), and the e-learning system (Aparicio et al., 2017). This current study is among one of the earliest attempts to validate the model in the context of social media search systems. In general, we find that the DeLone and McLean IS Success Model still holds in this context.

Third, we measured the construct of system quality in a relatively comprehensive way. Specifically, we treated it as a multi-dimensional, second-order construct, and utilized the five major dimensions provided in the Questionnaire for User Interaction Satisfaction (QUIS 7.0), a widely acknowledged and adopted comprehensive assessment of system performance, as its first-order constructs. The data analysis results showed that all five dimensions were significant aspects to the overall quality of the system.

In addition, most of the existing literature on testing and validating the DeLone and McLean IS Success Model utilized "use" (either intention to use or actual use) as the dependent variable in their research models (Alzahrani et al., 2017; Hsu et al., 2014;

Rana et al., 2015). Occasionally, some research was seen to use the individual-level benefit as the dependent variable (Aparicio et al., 2017). Different from them, in this study, we particularly looked into two levels of benefit, individual benefit and social benefit, and included both of them in our research model as the dependent variables. Our data analysis results showed that users' intention to use the large-scale, social media search system significantly influenced their perceptions on both the individual benefit and social benefit that the system could provide, and their satisfaction toward this type of systems had a significant impact on their perceptions of the individual benefit associated with the system.

Future research can further improve and extend the current study in several ways. First, in this study, we utilized the DeLone and McLean IS Success Model to examine the large-scale, social media search system. Future studies can utilize other theoretical lenses to examine this relatively new type of systems from other perspectives. Future research may also consider developing and incorporating constructs that are specific to this new type of systems to see how they could influence user acceptance and adoption. Second, the updated DeLone and McLean IS Success Model has a feedback loop from net benefits to intention to use/use and satisfaction. However, to make the proposed research model less complicated, we excluded such a loop (in our case, that would be from individual benefit and social benefit to intention to use and satisfaction). Future research may consider adding back and investigating the loop(s). Third, the sample size of this study is relatively small (with 78 subjects). However, because one major advantage of the statistical method we use in this study (i.e., PLS) is that it can deal with small sample sizes, our sample size is large enough for conducting data analysis robustly with convincing results. Future research may consider validating the research model with a larger number of subjects. Fourth, in our empirical study, the social media search system was about homeland security. Future research can further assess social media search systems created with data sources in other areas, such as healthcare and businesses. In addition, in our study, participants were all novice users of this type of systems, without any prior usage experience. They got the chance to play with the system just during the time of the lab experiment. However, with the increased experience of using an information system, users' perceptions toward it (such as their perceived information quality and system quality) might change over time. Thus, it would be of interesting to conduct a longitudinal study to examine the difference (if any) between novice users vs. users with prior usage experience.

References

Akar, E., & Topçu, B. (2011). An Examination of the Factors Influencing Consumers' Attitudes Toward Social Media Marketing. *Journal of Internet Commerce, 10*(1), 35-67.

Alzahrani, A. I., Mahmud, I., Ramayah, T., Alfarraj, O., & Alalwan, N. (2017). Modelling digital library success using the DeLone and McLean information system success model. *Journal of Librarianship and Information Science*, https://doi.org/10.1177/0961000617726123, 1-16.

Aparicio, M., Bacao, F., & Oliveira, T. (2017). Grit in the path to e-learning success. *Computers in Human Behavior, 66*, 388-399.

Ashley, C., & Tuten, T. (2015). Creative Strategies in Social Media Marketing: An Exploratory Study of Branded Social Content and Consumer Engagement. *Psychology and Marketing, 32*(1), 15-27.

Au, N., Ngai, E., & Cheng, T. (2008). Extending the Understanding of End User Information Systems Satisfaction Formation: An Equitable Needs Fulfillment Model Approach. *MIS Quarterly, 32*(1), 43-66.

Batrinca, B., & Treleaven, P. C. (2015). Social media analytics: a survey of techniques, tools and platforms. *AI & SOCIETY, 30*(1), 89-116.

Bhattacherjee, A. (2001). Understanding information systems continuance: an expectation-confirmation model. *MIS Quarterly, 25*(3), 351-370.

Börner, K., Chen, C., & Boyack, K. W. (2003). Visualizing knowledge domains. *Annual Review of Information Science and Technology, 37*, 179-255.

Cao, G., Wang, S., Hwang, M., Padmanabhan, A., Zhang, Z., & Soltani, K. (2015). A scalable framework for spatiotemporal analysis of location-based social media data. *Computers, Environment and Urban Systems, 51*, 70-82.

Chau, M., & Xu, J. (2012). Business Intelligence in Blogs: Understanding Consumer Interactions and Communities. *MIS Quarterly, 36*(4), 1189-1216.

Chen, H., & Roco, M. C. (2008). *Mapping nanotechnology innovations and knowledge: Global, longitudinal patent and literature analysis*. Springer.

Chin, J. P., Diehl, V. A., & Norman, K. L. (1988). Development of an instrument measuring user satisfaction of the human-computer interface. *Paper presented at the Proceedings of SIGCHI '88, New York.*

Chin, W. W. (1998). Issues and Opinions on Structural Equation Modeling. *MIS Quarterly, 22*(1), 7-16.

Dang, Y., Zhang, Y., Hu, P. J. H., Brown, S. A., Ku, Y., Wang, J. H., & Chen, H. (2014). An Integrated Framework for Analyzing Multilingual Content in Web 2.0 Social Media. *Decision Support Systems, 61*(1), 126-135.

DeLone, W. H., & McLean, E. R. (1992). Information Systems Success: The Quest for the Dependent Variable. *Information Systems Research, 3*(1), 60-95.

DeLone, W. H., & McLean, E. R. (2003). The DeLone and McLean model of information systems success: A ten-year update. *Journal of Management Information Systems, 19*(4), 9-30.

Doll, W. J., & Torkzadeh, G. (1998). The measurementof end-user computing satisfaction. *MIS Quarterly, 12*(June), 259-273.

Fan, W., & Gordon, M. D. (2014). The Power of Social Media Analytics. *Communications of the ACM, 57*(6), 74-81.

Fang, X., & Zhan, J. (2015). Sentiment analysis using product review data. *Journal of Big Data, 2*(5), 1-14.

Gefen, D., Straub, D. W., & Boudreau, M. C. (2000). Structural Equation Modeling and Regression: Guidelines for Research Practice. *Communications of the AIS, 4*(7), 1-77.

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate Data Analysis*. Upper Saddle River, NJ: Prentice Hall.

Hsu, M. H., Chang, C. M., Chu, K. K., & Lee, Y. J. (2014). Determinants of repurchase intention in online group-buying: The perspectives of DeLone & McLean IS success model and trust. *Computers in Human Behavior, 36*, 234-245.

IFL Science. (2017). How Much Data Does The World Generate Every Minute? Retrieved from http://www.iflscience.com/technology/how-much-data-does-the-world-generate-every-minute/ (Date Accessed January 13, 2018).

Kaipio, J., Lääveri, T., Hyppönen, H., Vainiomäki, S., Reponen, J., Kushniruk, A., Borycki, E. & Vänskäha, J. (2017). Usability problems do not heal by themselves: National survey on physicians' experiences with EHRs in Finland. *International Journal of Medical Informatics, 97*, 266-281.

Kaur, G. (2016). Social Media Marketing. *Asian Journal of Multidisciplinary Studies, 4*(7), 34-36.

Kleinbaum, D. G., Kupper, L. L., Muller, K. E., & Nizam, A. (1998). *Applied Regression Analysis and Other Multivariate Methods*. Belmont, California: Duxbury Press.

Lee, G., & Xia, W. (2010). Toward Agile: An Integrated Analysis of Quantitative and Qualitative Field Data on Software Development Agility. *MIS Quarterly, 34*(1), 87-114.

Li, N., & Wu, D. D. (2010). Using text mining and sentiment analysis for online forums hotspot detection and forecast. *Decision Support Systems, 48*(2), 354-368.

Lin, H. C. (2017). Nurses' Satisfaction With Using Nursing Information Systems From Technology Acceptance Model and Information Systems Success Model Perspectives: A Reductionist Approach. *CIN: Computers, Informatics, Nursing, 35*(2), 91-99.

Liu, B. (2007). *Web data mining*. Berlin Heidelberg: Springer.

Liu, D. R., Omar, H., Liou, C. H., Chi, H. C., & Hsu, C. H. (2015). Recommending blog articles based on popular event trend analysis. *Information Sciences, 305*(1), 302-319.

Luo, Q., & Zhong, D. (2015). Using social network analysis to explain communication characteristics of travel-related electronic word-of-mouth on social networking sites. *Tourism Management, 46*, 274-282.

Mathioudakis, M., & Koudas, N. (2010). Twittermonitor: trend detection over the twitter stream. In *Proceedings of the 2010 ACM SIGMOD International Conference on Management of data* (pp. 1155-1158). ACM.

Naeini, H. S., & Mostowfi, S. (2015). Using QUIS as a Measurement Tool for User Satisfaction Evaluation (Case Study: Vending Machine). *International Journal of Information Science, 5*(1), 14-23.

Otte, E., & Rousseau, R. (2002). Social network analysis: a powerful strategy, also for the information sciences. *Journal of Information Science, 28*(6), 441-453.

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval, 1*(1-2), 1-135.

Pang, B., Lee, L., & Vaithyanathain, S. (2002). Thumbs up? Sentiment classification using machine learning techniques. *Paper presented at the the ACL-02 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Philadelphia.

Pu, X., Wu, G., & Yuan, C. (2017). User-aware topic modeling of online reviews. *Multimedia Systems, https://doi.org/10.1007/s00530-017-0557-6*, 1-11.

Rai, A., Lang, S. S., & Welker, R. B. (2002). Assessing the Validity of IS Success Models: An Empirical Test and Theoretical Analysis. *Information Systems Research, 13*(1), 50-69.

Rana, N. P., Dwivedi, Y. K., Williams, M. D., & Weerakkody, V. (2015). Investigating success of an e-government initiative: Validation of an integrated IS success model. *Information Systems Frontiers, 17*(1), 127-142.

Resnik, P., Armstrong, W., Claudino, L., Nguyen, T., Nguyen, V. A., & Boyd-Graber, J. (2015). Beyond LDA: Exploring Supervised Topic Modeling for Depression-Related Language in Twitter. *Paper presented at the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality,* Denver, Colorado.

Ringle, C. M., Wende, S., & Will, A. (2005). *SmartPLS 2.0 (M3) beta*. Hamburg: http://www.smartpls.de.

Saif, H., He, Y., Fernandez, M., & Alani, H. (2016). Contextual semantics for sentiment analysis of Twitter. *Information Processing and Management, 52*(1), 5-19.

Samtani, S., Chinn, R., & Chen, H. (2015). Exploring hacker assets in underground forums. *Paper presented at the 2015 IEEE International Conference on Intelligence and Security Informatics* (ISI).

Scharl, A., Hubmann-Haidvogel, A., Jones, A., Fischl, D., Kamolov, R., Weichselbraun, A., & Rafelsberger, W. (2016). Analyzing the public discourse on works of fiction – Detection and visualization of emotion in online coverage about HBO's Game of Thrones. *Information Processing and Management, 52*(1), 129-138.

Seddon, P. B. (1997). A respecification and extension of the DeLone and McLean model of IS success. *Information Systems Research, 8*(3), 240-253.

Seder, D., & Gable, G. G. (2010). Knowledge Management Competence for Enterprise System Success. *The Journal of Strategic Information Systems, 19*(4), 296-306.

Shneiderman, B. (1996). The eyes have it: A task by data type taxonomy for information visualizations. In *Proceedings of the IEEE Symposium on Visual Languages*, Washington.

Stevens, J. (2017). Internet Stats & Facts for 2017. Retrieved from https://hostingfacts.com/internet-facts-stats-2016/ (Data Accessed January 13, 2018)

Tam, C., & Oliveira, T. (2016). Understanding the impact of m-banking on individual performance: DeLone & McLean and TTF perspective. *Computers in Human Behavior, 61*, 233-244.

Vargas, J. A. (2012). Spring Awakening: How an Egyptian Revolution Began on Facebook. *The New York Times*, retrieved from http://www.nytimes.com/2012/02/19/books/review/how-an-egyptian-revolution-began-on-facebook.html?pagewanted=all&_moc.semityn.www on January 19, 2018.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Towards a unified view. *MIS Quarterly, 27*(3), 425-478.

Wasserman, S., & Faust, K. (1994). *Social networks analysis: Methods and applications*. Cambridge: Cambridge University Press.

Wiebe, J., Wilson, T., Bruce, R., Bell, M., & Martin, M. (2004). Learning subjective language. *Computational Linguistics, 30*(3), 277-308.

Wixom, B. H., & Watson, H. J. (2001). An Empirical Investigation of the Factors Affecting Data Warehousing Success. *MIS Quarterly, 25*(1), 17-41.

Zhu, B., & Chen, H. (2005). Information visualization. *Annual Review of Information Science and Technology (ARIST), 39*, 139-178.

Appendix A. Measurement Items Used in the Study

Information Quality

1. Does the Dark Web Forum Portal provide the precise information you need?
2. Does the Dark Web Forum Portal provide output that is exactly what you need?
3. Does the Dark Web Forum Portal provide sufficient information to enable you to do your tasks?
4. Does the Dark Web Forum Portal have errors in the program that you must work around?
5. Are you satisfied with the accuracy of the Dark Web Forum Portal?
6. Are the output options (e.g. print types, page sizes allowed for, etc.) sufficient for your use?
7. Is the information provided helpful regarding your questions or problems?

System Quality – QUISA

1. Overall, I feel the Dark Web Forum Portal: terrible/wonderful
2. Overall, I feel the Dark Web Forum Portal: frustrating/satisfying
3. Overall, I feel the Dark Web Forum Portal: dull/stimulating
4. Overall, I feel the Dark Web Forum Portal: difficult/easy
5. Overall, I feel the Dark Web Forum Portal: inadequate power/adequate power
6. Overall, I feel the Dark Web Forum Portal: rigid/flexible

System Quality – QUISB

1. Overall, I feel: Characters on the computer screen: hard to read/easy to read
2. Overall, I feel: Highlighting on the screen simplifies task (not at all/very much)
3. Overall, I feel: Organization of information on screen (confusing/very clear)
4. Overall, I feel: Sequence of screens (confusing/very clear)

System Quality – QUISC

1. Overall, I feel: Use of terms throughout system (inconsistent/consistent)
2. Overall, I feel: Computer terminology is related to the task you are doing (never/always)
3. Overall, I feel: Position of messages on screen (inconsistent/consistent)
4. Overall, I feel: Messages on screen which prompt user for input (confusing/clear)
5. Overall, I feel: Computer keeps you informed about what it is doing (never/always)
6. Overall, I feel: Error messages (unhelpful/helpful)

System Quality – QUISD

1. Overall, I feel: Learning to operate the system (difficult/easy)
2. Overall, I feel: Exploring new features by trial and error (difficult/easy)
3. Overall, I feel: Remembering the operations of using the system to finish tasks is (difficult/easy)
4. Overall, I feel: Tasks can be performed in a straight-forward manner (never/always)
5. Overall, I feel: Help messages on the screen (unhelpful/helpful)
6. Overall, I feel: Supplemental reference materials (confusing/clear)

System Quality – QUISE

1. Overall, I feel: System speed (too slow/fast enough)
2. Overall, I feel: System reliability (unreliable/reliable)
3. Overall, I feel: Correcting your mistakes (difficult/easy)
4. Overall, I feel: Experienced and inexperienced users' needs are taken into consideration (never/always).

Intention to Use

1. If given the opportunity, I would use Dark Web Forum Portal to search for forum messages in future.
2. I intend to use Dark Web Forum Portal rather than use any alternative means to search for forum messages in future.
3. If I could, I would like to avoid using Dark Web Forum Portal to search for forum messages in future. [reverse coding]

Satisfaction

1. Overall, my use of Dark Web Forum Portal has left me feeling: Very dissatisfied/very satisfied.
2. Overall, my use of Dark Web Forum Portal has left me feeling: Very displeased/very pleased.
3. Overall, my use of Dark Web Forum Portal has left me feeling: Very frustrated/very contented.
4. Overall, my use of Dark Web Forum Portal has left me feeling: Very terrible/very delighted.

Individual Benefit

1. Dark Web Forum Portal has changed my job significantly.
2. Dark Web Forum Portal has reduced the time it takes to support decision making to the end-user community.
3. Dark Web Forum Portal has reduced the effort it takes to support decision making to the end-user community.

Social Benefit

1. Dark Web Forum Portal has potential benefit for the security of our society.
2. Dark Web Forum Portal can help better monitor the security related content.
3. Dark Web Forum Portal is useful for supporting better decision-making related to societal security.

Appendix B. Tasks Used in the Lab Experiment

Q1: Please find one forum message from the Arabic forum “Alokab” that has the keyword Al-Qaeda.

Q2: Please find one forum message from the Arabic forum “Alokab” that has the keyword Hamas.

Q3: Please find one forum thread from the English forum “Islamic Awakening” that has the keyword “bomb” in its title.

Q4: Please find one forum message from the English forum “Islamic Awakening” that has the keyword “terrorist” and was posted by the forum user named Daniel.

About the Authors

Mandy Yan Dang is an Associate Professor of Information Systems in the W.A. Franke College of Business at Northern Arizona University. She received her Ph.D. in Management Information Systems from the University of Arizona. Her research interests include implementation and adoption of information systems, knowledge management, and computing education. Her research articles have appeared in Journal of Management Information

Systems, Decision Support Systems, Information Systems Frontiers, and etc.

Gavin Yulei Zhang is an Associate Professor of Information Systems in the W.A. Franke College of Business at Northern Arizona University. He received his Ph.D. in Management Information Systems from the University of Arizona. His research interests include social media analytics, Web and text mining, and computing education. His research articles have appeared in Journal of Management Information Systems, Decision Support Systems, Information Systems Frontiers, and etc.

Hsinchun Chen is University of Arizona Regents' Professor and Thomas R. Brown Chair in Management and Technology in the Management Information Systems Department and Professor of Entrepreneurship & Innovation in the McGuire Center for Entrepreneurship at the College of Management of the University of Arizona. He received the B.S. degree from the National Chiao-Tung University in Taiwan, the MBA degree from SUNY Buffalo, and the Ph.D. degree in Information Systems from the New York University. He is a Fellow of IEEE, ACM, and AAAS. He is author/editor of 20 books, 300 SCI journal articles, and 200 refereed conference articles covering digital library, data/text/web mining, business analytics, security informatics, and health informatics.