Chapter 13
Information Privacy on Online Social Networks: Illusion-in-Progress in the Age of Big Data?

Shwadhin Sharma and Babita Gupta

Abstract In the age of big data where vast amounts of data are collected, stored, and analyzed from all possible sources, the growth of social media and the culture of sharing personal information have created privacy and security related issues. Drawing on the prospect theory and rational apathy theory, we present a research model to investigate why people disclose personal information on Online Social Networks. This paper analyzes the impact of situational factors such as information control, ownership of personal information, and apathy towards privacy concern of users on Online Social Network. We describe the proposed research design for collecting our data and analysis using structural equation modeling to analyze the data. The findings and conclusions will be presented after the data is analyzed. This work contributes to the network analytics by developing new constructs using the Prospect Theory and the Rational Apathy theory from the fields of behavioral economics and social psychology respectively.

Keywords Online social networks • Big data • Information privacy • Information control • Ownership of personal information • Privacy apathy • Prospect theory • Information disclosure • Rational apathy

13.1 Introduction

The proliferation of social media and web 2.0 is enabling individuals and companies to engage with digital technologies at an unprecedented scale generating vast amounts of data, also referred to as “big data”. Big data is characterized by higher volume, velocity, and variety (the three V’s) of data that usually cannot be handled by traditional database management tools (Zikopoulos et al. 2012) and is often characterized as a massive volume of both structured and unstructured data that are

S. Sharma (✉) • B. Gupta
College of Business, California State University, Monterey Bay, Marina, CA, USA
e-mail: ssharma@csumb.edu; bgupta@csumb.edu

A.V. Deokar et al. (eds.), *Analytics and Data Science*, Annals of Information
Systems 21, DOI 10.1007/978-3-319-58097-5_13

generated at high velocity with veracity that adds value to the intended process (Demchenko et al. 2013; Kshetri 2014).

An Online Social Network (OSN) is an online platform that allows members to create public profiles within a bounded system, share texts, photos and videos, and other personal information, and thus, connect, develop, and maintain relationships (Boyd and Ellison 2007; Ellison et al. 2007). People may use OSNs for many different reasons including socialization, fun and enjoyment, usefulness in communicating and interacting with friends, bridging, bonding, and maintaining social capital. Use of OSNs have created humongous amount of structured and unstructured data. Indeed, the recent attention that big data is garnering can be largely credited to the rapid development of the Online Social Networks (OSNs). As OSNs have provided additional channels for interpersonal and business communication, huge volumes and variety of data are being generated for collection, storage, and aggregation from OSNs. These data can be used by the governments, business organizations, research agencies, marketing companies, etc. Manyika et al. (2011) estimated the value of big data for U.S. medical industry alone to be $300 billion. Companies in various industry sectors such as healthcare, retail, services market, supply chain and transportation, entertainment, and marketing and advertising have started to pay close attention to the big data phenomenon and thus, to OSNs as one of the primary sources of the big data (Tan et al. 2013). It is also important to note that despite several benefits of OSNs in the big data environment, the ability of organizations to collect, store, and analyze big data poses privacy and security related risks for the users.

The interaction of OSNs with users and the generation of big data on OSNs through these interactions are presented in Fig. 13.1 below. The interactions created in OSNs are accessed by many parties such as government, big organizations, third parties, and consumer and service firms. Figure 13.1 presents the simplistic view of how OSNs acts as a source of big data and thus, the source of privacy and security issues.

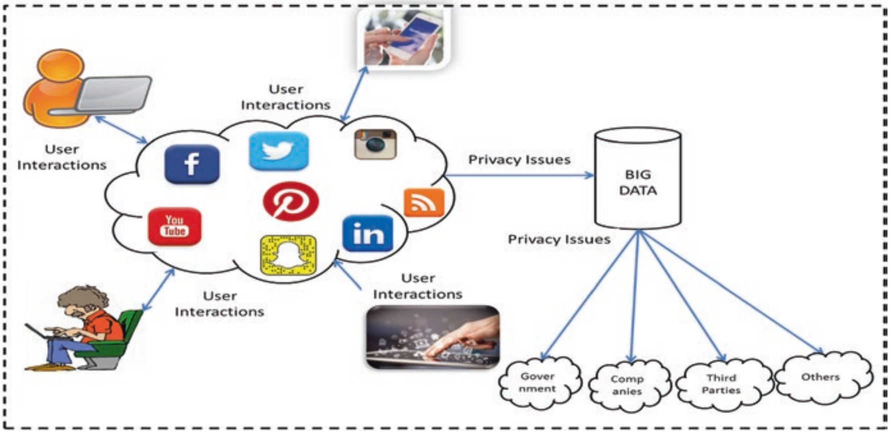

Fig. 13.1 OSN-big data-privacy

The growth of social media and the culture of sharing information have fueled the proliferation of OSNs such as Facebook, Instagram, Twitter, Google+, and Pinterest in individuals' daily life. These OSNs are becoming an important social platform for computer-mediated communication (Nadkarni and Hofmann 2012) at an exponential rate—be it for bonding, bridging, or maintaining social capital (Ellison et al. 2007), or using it as a medium of social interaction and exchanges (Boyd and Ellison 2007). With Facebook alone having more than one billion members (Sharma and Crossler 2014), it is no surprise to see that almost fourth-fifths of the Internet users use one or the other OSN (Conroy and Williams 2014). The exponential growth of OSNs has brought an intense focus on the privacy and security issues of its users. OSNs have been plagued with issues of privacy risks such as the surveillance, secondary use of Information, and collection of irrelevant information (Sharma and Crossler 2014). In the context of the big data, this already complex issue becomes further complicated.

An individual may feel a threat to their privacy when they lose control to their personal information. In an online environment, where users feel a certain amount of anonymity and the OSN providers have the freedom to aggregate and share the information easily, the issues of privacy and security may be more predominant. In a social network environment, information privacy may imply the level of identifiable information collected by the organization and the possible unauthorized uses of that information. These privacy concerns can range from information threats such as digital aggregation and improper access of personal data by third parties to dangers arising from the social environment such as online stalking, bullying, or leaking of private data to the world (Hogben 2007). The level of sophistication of technologies analyzing the big data generated by the OSNs has increased greatly over the last few years. In addition, cost-effective and innovative forms of collection and -processing of high volume, high velocity, and high-variety information assets has brought the privacy and security issues to the forefront (Kshetri 2014). As we start capturing life in digital reality in online social networks, it becomes easier for people and organizations with the right skills set to build an accurate portrait of our past, present, and future behavior, without our knowledge. Software such as Rapid Information Overlay Technology developed for the U.S. defense department (theguardian.com 2013) uses 'extreme-scale analytics' to gather information about individuals' online social network habits to predict their future behaviors. Internet giants like Google and Facebook (including Instagram) have been criticized for a long time for the lack of transparency on what's being done with the users' data they collect. An example of volume, velocity, and variety of data that Facebook stores and can retrieve is the Facebook Graph Search function that was launched in March 2013. This function can give answers to user's natural language queries by combining the big data acquired from its billions of members and external data into a search engine. These results can link Facebook activities such as pictures liked, relationship status, and comments made between a user's friends from the time they joined Facebook. Big data and it tools and techniques that are being used by many OSN companies are opaque, masked by the layers of technical, legal, physical design (Richards and King 2013), making the data being collected and used by these companies questionable. On top of it, there are several third-party applications on OSNs that also collect user information in real-time. As such, real-time structured and unstructured data provided and shared on OSNs such as Facebook, Instagram, Twitter, and Foursquare generally carry privacy risks.

However, even though social media is taking on the role of primary communication, people, especially the millennials, may be in a state of indifference when it comes to their privacy (Yoo et al. 2012). Some of these individuals using OSNs may not be aware of the risk associated with the release of personal information. Others may have experienced privacy invasion and thus, may not consider their information to be private anymore (Solove 2008) becoming apathetic towards their own privacy over time. Some people are comfortable giving up their privacy for patriotic reasons such as for national security while others believe that they have nothing to hide anyway as all of their information is already collected by big organizations like Google or the government (Goitein 2013).

Information Systems (IS) research has focused on privacy and its value. However, we do not yet fully understand why people, despite valuing privacy, still choose to freely share their personal information online. Thus, the concept of privacy on OSN is an interesting one to study as the value of privacy for each individual is situational in nature (Acquisti et al. 2013)—some users may modulate their privacy boundaries; for others, the definition of privacy in itself might vary from one situation/timeline to other. Using the prospect theory and rational apathy theory, this paper analyzes the impact of situational factors such as information control, ownership of personal information, benefits of information disclosure, and apathy towards user's privacy concern on OSN.

13.2 Literature Review

As OSNs are public platforms by design, any information shared on it carries a significant risk of being collected, stored and used without authorization as organizations and third parties such as advertising agents, employers, law enforcement agents, creditors, and tax authorities are increasingly seeking information shared and provided on OSNs users (Hogben 2007; Krasnova et al. 2012; Stieglitz et al. 2014). Privacy related issues can range from negative impact on personal and family lives, damages to reputation (Afroz et al. 2013), identity theft, and psychological pain such as embarrassment and addiction (Turel and Serenko 2012). With the increasing use of OSNs and big data, privacy concern construct has become one of the most widely used variables in IS research to predict the privacy-related behaviors (Dhami et al. 2013; Johnson et al. 2012; Xu et al. 2008). The findings of these privacy-related behavior studies have been often different from one another and sometimes, even contradictory. Some studies found that privacy concerns are more prevalent among the OSN users and negatively impact OSN usage behavior (O'Brien and Torres 2012; Xu et al. 2013). Other studies found that the OSNs users seem to be oblivious to privacy risks and thus, comfortable sharing their personal information on a social network (Hugl 2011; Rosenblum 2007). Despite the privacy risk, the users still use OSNs and share their personal information (Acquisti and Gross 2006; Tufekci 2008). This study explores how the introduction and rise of big data on OSNs would affect the perception of the users toward the privacy concerns and affect their OSN usage behavior.

It is important to study privacy in relation to big data as most of the hacking and privacy violations are now on bigger and broader terms. In 2010, Julian Assange used WikiLeaks to upload 90,000 documents related to Afghan War and started an unprecedented big data leak in the U.S. military history. Edward Snowden followed the trend by publishing 20 times as many documents. The data that was leaked provided a glimpse of how the U.S. government has been performing surveillance activities on its own citizens as well as leaders around the world such as Angela Merkel, Germany's chancellor. Recent big data breaches in Anthem Inc. and Ashley Madison are bringing a lot of attention to privacy violations as well. Big data has allowed people to extract implicit, previously unknown, and potentially personally identifiable information about the individuals.

13.3 Theoretical Framework and Hypotheses

13.3.1 Prospect Theory

Prospect theory states that while making decisions, individuals appraise a set of decision alternatives based on personal heuristics, and then select the alternative that brings the highest satisfaction and outcome (Keith et al. 2012). However, such personal decision heuristics may demonstrate bounded rationality (Simon 1982) as the theory is based on the assumption that utility comes from the returns and not the value of assets. Thus, an individual's reference point would strongly affect the choice of their heuristics (Kahneman and Tversky 1979). This phenomenon has fascinating implications for individuals' decision to share their personal information on OSNs as users compare the utility derived from information sharing to the loss of information through privacy risk.

13.3.2 Rational Apathy Theory

In many cases, individuals are rationally apathetic towards a cause. When a voter feels that his vote would not have any real influence on the conclusion of an election or change the political scenario, he could develop apathy towards the election. Similarly, a rational shareholder would not put an extra effort to go through the length and complexity of proxy statements unless he feels that his effort will make a difference (Karuitha et al. 2013). Apathy is basically defined as a state of indifference or reasoned assessment where an individual has an absence of interest or concern to certain aspects of emotional, social or physical life often caused by "learned helplessness" (Sarfaraz et al. 2012). Similarly, individuals using OSNs may show a non-pathological lack of interest towards their privacy as they may not consider it important (Solmitz 2000) or may believe that their privacy is already too diluted by the companies collecting data to care about it anymore (Yoo et al. 2012). In a privacy context, a person with a reference point of complete information control may quickly travel to the point of privacy apathy.

Drawing from the Prospect theory which is an extension of expected utility hypothesis and from the Rational Apathy theory, we visualize our research model for this study in the Fig. 13.2. This research model has two dimensions to it: one is the privacy calculus that examines risk and benefits of information disclosure, and the other is the reference point for their privacy control, ownership, and belief. In this paper, we investigate how reference points such as information control, perceived ownership, and existing offline benefits affect user's protection belief, risk belief, their state of apathy towards privacy, and perceived benefits from using the OSNs. We further study how all these constructs would affect the user's tendency to disclose information on OSNs. Thus, this paper seeks to answer the following research questions:

1. How do the perceived ownership, perceived information control, and existing offline benefits affect protection belief, risk belief, perceived benefits and state of user's privacy apathy on OSNs?
2. How do privacy apathy, protection belief, risk belief, and benefits affect information disclosure on OSNs in the age of big data?

To control for an explanation of results due to extraneous factors, prior research on OSN and information systems identified a number of factors that may impact the actual behavior of respondents. Analyzing the impact of control variables is essen-

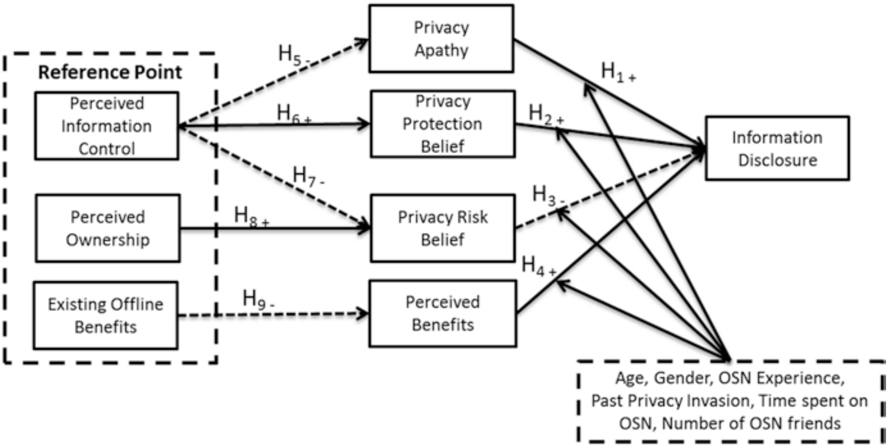

Fig. 13.2 Research model

tial for a research model as it removes any confounding variables (Ormond 2014). Thus, for this research model, age, gender, OSN experience, past privacy invasion, number of OSN friends, number of years of experience on Facebook, and time spent on OSN were included as the control variables to see if they impact the dependent variable.

13.4 Hypotheses Testing

While privacy apathy is a relatively newer concept in IS, the concept has been gaining momentum as a way to gauge the indifference of a user towards privacy concerns (Sharma and Crossler 2014). With big data being collected and stored by millions of web sites, applications, agencies, and third parties, individuals may believe that there is no such thing as privacy in the age of Web 2.0 technologies. As stated by Mark Zuckerberg, the co-founder of Facebook on January 2010, privacy is no more a "social norm". Similar sentiments were echoed by the United States Senate majority leader Harry Reid when he advised to "just calm down and understand that National Security Agency's (NSA) PRISM isn't anything that is brand new" (csmonitor.com 2013). Similarly, a recent survey showed that almost half of the Americans take NSA's PRISM program of data surveillance as "no big deal" as these people believe that "they're being tracked all over the Internet by companies like Google and Facebook" (csmonitor.com 2013). Thus, it is safe to hypothesize that users with privacy apathy put lower value and price to their personal information and thus, care less about information disclosure (Yoo et al. 2012).

H1: Privacy apathy would positively influence intention to disclose information on OSNs despite the threat of big data.

Privacy protection belief is the subjective possibility that consumers believe that their private information is protected as anticipated (Metzger 2004). In an online setting, users who exemplify higher protection beliefs are believed to have more control over their information and thus, are more in control over information disclosure and are more likely to disclose their personal information (Raschke et al. 2014). Thus, it is predicted that:

H2: Privacy protection belief would positively influence intention to disclose information on OSNs despite the threat of big data.

Privacy risk belief implies the probability of potential loss because of disclosure of personal information (Malhotra et al. 2004). It is deemed to be the cost of privacy as disclosing information is often considered risky. Such cost and risks associated to OSN can range from unintended third parties receiving users' personal information to hacking of personal account based on information shared on OSN (Hogben 2007). Several studies have verified the negative effect of perceived privacy risk on people's intention to disclose personal information on online transactions and activities (Li et al. 2010; Malhotra et al. 2004).

H3: Privacy risk belief would negatively influence intention to disclose information on OSNs despite the threat of big data.

Perceived benefits refer to a user's overall expectation of positive outcomes from an OSN without any significant privacy threats (Bulgurcu 2012). Individuals are likely to give up a degree of privacy in return for potential benefits related to OSNs. In an OSN environment, the user's fear in the form of losing control of personal information is compensated by the several benefits such as information, enjoyment, and convenience (Hogben 2007). Thus, the following is hypothesized:

H4: Perceived benefit would positively influence intention to disclose information on OSNs despite the threat of big data.

Privacy and control has often been linked together in prior work (Westin 1967). The ability of people to control their information has been emphasized as critical in any concept of privacy (Wolfe and Laufer 1974). Thus, it is no surprise to see that there has been an outcry regarding how users have lost control of their information on OSNs (Boyd 2008; Hoadley et al. 2010). When people tend to share information on OSN, it is often broadcasted to the network of friends. Sometimes, such information is accessed by the third party applications installed by the users. An individual believing lower information control on OSNs would believe that such information has been collected and stored by OSNs and third parties and thus, would have higher privacy apathy. Similarly, a sense of higher information control would lead to a positive privacy protection belief and a lower privacy risk belief. Thus, we posit:

H5: Perceived information control would negatively influence privacy apathy.
H6: Perceived information control would positively influence privacy protection belief.
H7: Perceived information control would negatively influence privacy risk belief.

Perceived ownership implies a sense of possession and entitlement (Furby 1978). In the case of information and OSNs, perceived ownership implies the sense of entitlement, possession, and attachment towards the information shared on OSNs (Feuchtl and Kamleitner 2009; Sharma and Crossler 2014). When individuals believe that the information shared on OSN is their information and contains some level of attachment with their identity and privacy, it positively influences their privacy risk belief (Sharma and Crossler 2014). Thus, it is hypothesized that:

H8: Perceived ownership would positively influence privacy risk belief.

Away from OSN, there are tremendous opportunities for people to maintain a relationship, enjoy life, consume information, and develop an offline real-life image. Users of OSN will perceive lower benefits from OSN use when they are enjoying many of similar benefits offline in their real life. Thus, existing benefits decrease the perceived benefits of future disclosure (Keith et al. 2012). Thus, we propose:

H9: Existing offline benefits would negatively influence perceived benefits of OSNs.

13.5 Hypotheses Testing

The proposed conceptual model will be evaluated using survey design. An online questionnaire survey has been developed to collect data and perform empirical tests of the relationships proposed in the research model. The survey design technique fits the research phenomenon being studied in this research as the objective of this research is to explore user's information disclosure behavior on OSN. Also, a survey design provides the benefit of generalizability to the study as data could be collected from a wider range of respondents.

All the items used in the survey instrument are adapted from previous studies. The items are reflective likert-scale and have been adapted to fit the context of this study. The items for this study along with their respective original source/s have been presented in Table 13.1 below:

Although constructs adopted from earlier studies have been rigorously tested for reliability and validity, additional content validation using a multi-stage iterative procedure is recommended (Churchill 1979). Podsakoff et al. (2003) also have suggested using an ex-ante approach such as expert panel review and a pilot test to control Common Method Bias (CMV). Thus, a preliminary investigation consisting of expert panel review, pretest, and pilot test will be conducted to ensure measurement validity of the instrument. The changes suggested by the expert panel review and pre- and pilot tests such as revisions to wordings to improve clarity and precision, dropping items to make the survey fatigue-free, revision of items to make them unambiguous, etc. will be incorporated. This will ensure content validity of our survey instruments and also reduce CMV. Similarly, to reduce CMV we will keep our survey anonymous, optional, and relatively short. We will also assess the extent of common method variance with two statistical tests. First, we will perform Harman's single factor test by loading all of the items in a principal component factor analysis (Podsakoff et al. 2003). If the results show that there is more than a single factor that accounts for a majority of covariance, it would suggest absence of CMV in our study. However, as Harman's single factor test is increasingly contested for its ability to detect common method bias, we will also use Lindell and Whitney's (2001) test that uses a theoretically unrelated construct (termed a marker variable) to assess CMV. We will use "Perceived effectiveness of credit card guarantees" as our marker variable construct for this study and will use it to adjust the correlations among the principal constructs (Pavlou et al. 2007). The absence of high correlations among any of the items of the study's principal constructs and perceived effectiveness of credit card guarantees would indicate that the study doesn't have serious concerns about common method bias as the construct perceived effectiveness of credit card guarantees is expected to be weakly related to the study's principal constructs.

Undergraduate students from different classes within a public university in California will be invited to complete the survey. The invitations will be sent through emails as well through classroom visits. As the age group of 18–25 years that are

Table 13.1 Pilot survey instrument

Survey instrument

Your Personal Information implies information that is related to you, can be used on its own or with other information to identify, contact, or locate you. Some of the examples of Personal Information can be address, location, race, relationship history, purchasing behavior, phone number, pictures (and tagging), etc.

OSN refers to Online Social Network that includes interactions on social media and web 2.0 technology platforms such as Facebook, Instagram, Snapchat, Vine, Youtube, Twitter, etc.

Construct	**Adapted item**	**Original source**
Perceived ownership (PO)	Information I share while on/to OSN is MY personal information.	Van Dyne and Pierce (2004)
	I sense that the information I provide on/to OSN is my own.	
	I feel a very high degree of personal ownership for the information I provide on/to OSN.	
	I sense that the information I provide on/to OSN is personal.	
	I believe that the information I disclose on/to OSN belongs to me.	
Privacy apathy (PA)	I have little interest in information privacy issues on information provided on/to OSN.	Yoo et al. (2012); Sharma and Crossler (2014)
	I care less about information privacy anymore on information provided on/to OSN.	
	I do not worry about privacy issues anymore on information provided on/to OSN.	
Privacy protection belief	I am confident that I know all the parties who would collect information that I share on/to OSN.	Li et al. (2011)
	I am aware of the exact nature of information that will be collected, stored, and used by OSN.	
	I believe there is an effective mechanism to address any violation of the information I provide on/to OSN	
	I am confident that I know all the parties who would collect information that I share on/to OSN.	
Privacy risk belief of information disclosure (PRB)	Sharing information on/to OSN would involve many unexpected problems.	Malhotra et al. (2004); Xu et al. (2009)
	It would be risky to disclose information on/to OSN.	
	There would be too much uncertainty with providing information on/to OSN	
	There would be high potential for loss in disclosing information on/to OSN.	

(continued)

Table 13.1 (continued)

Perceived information control	I believe I have control over the amount of your personal information collected on OSN.	Xu (2007)
	I believe I have control over who can get access to my personal information on OSN.	
	I believe I have control over my personal information that has been released on OSN.	
	I believe I have control over how my personal information is being used by OSN.	
	I believe I have control over my personal information that I provided on/to OSN.	
Perceived benefits	OSN is useful to exchange personal information with my friends.	Ellison et al. (2007); Krasnova et al. (2010)
	OSN is useful for me to monitor what others share about themselves.	
	Sharing personal information on OSN is fun.	
	By sharing personal information on OSN, I get more popular with my OSN-friends.	
	I share personal information via OSN because it's better than the alternatives.	
Existing offline benefits	I have more time to spend with my family and friends around me.	Self-developed
	Staying offline has several benefits than staying online.	
	I can build real relationships and stay happy and healthy when I am offline.	
	I have more time to pursue my hobbies and pursuits and form network with people I know.	
Behavioral intent to disclose information (BINT)	I am likely to provide my personal information on/to OSN.	Xu and Teo (2004)
	I plan to provide my personal information on/to OSN.	
	I intend to provide my personal information on/to OSN.	

educated and college students are the ones that use the OSNs the most (Lenhart et al. 2010), it is appropriate to have undergraduate students as the sample for this study.

A primary investigation consisting of reliability and validity testing, model fit test (i.e. goodness of fit), common method bias test, and t-test is conducted to ensure the validity of the structural model. We will use SmartPLS 2.0, SPSS along with AMOS for our instrument validation and testing of the structural model. SmartPLS uses a Partial Least Square (PLS) regression technique that employs a component-based approach for estimation and places minimal restrictions on sample size, measurement scales and residual distributions (Chin and Todd 1995) and it does not impose normality requirements on the data. We will also use AMOS which is a covariance based structured equation model that provides various overall goodness-of-fit indices for assessing model fit and method variance.

Before testing the hypothesized structural model, we evaluate the psychometric properties of the measures. All the constructs in this study are measured with multiple items. A PLS confirmatory analysis will be conducted to examine convergent validity, discriminant validity, and reliability using commonly accepted guidelines (Churchill 1979). Reliability for the constructs will be measured using composite reliability score and Cronbach's alpha. The Cronbach's alpha and composite reliability examine the internal consistency among the data. For all the constructs in our study, we will also perform descriptive statistics of all the constructs including means and standard deviations and the level of each item's contribution to the overall factor.

To further examine the validity of the measurement model, we will analyze how well the model fits the data with the help of model fit statistics available through AMOS (Anderson and Gerbing 1988). The goodness of fit index (GFI), comparative fit index (CFI), normed fit index (NFI) and incremental fit index (IFI) all assess the goodness of fit of the model with the data and should be above 0.90 to show model fit. Root mean square error of approximation (RMSEA) and Standardized Root Mean Square Residual (SRMR) which measures the "badness of fit" should both be below 0.05. Similarly, we will also assess if the relative chi-square (i.e. CMIN/df), which is also a "badness of fit", is below the threshold of 3 (Kline 1998) and is thus non-significant. Together, the result would indicate if our hypothesized measurement model is "fitting" the observed data.

The hypotheses and the relationships used for this study will be tested by examining the structural model of our study. A bootstrapping resampling procedure will be performed to assess the significance of the path coefficients within the structural model. The proposed hypotheses for this research model will be tested using t-statistics (p-value) for the standardized path coefficients. The t-statistics (p-value) provided by PLS structural model analysis would show us if the hypothesis is supported or not while the standardized path coefficients would determine the direction and strength of the relationship between exogenous and endogenous variables. The study will have a satisfactory and substantive model if the dependent factors have R-square (the variance explained by the independent variables) greater than 0.10 (Falk and Miller 1992). Thus, we will examine if the proposed paths were mostly significant and how well our model explained the variances in our endogenous variables. Also, we will analyze the effect of our control variables age, gender, education, experience in social networking, and experience on the Internet on the intention to disclose information on OSNs.

13.6 Conclusion

13.6.1 Study Summarization

The objective of this research is to explore the factors that may affect user's information disclosure behavior on online social networks. There has been limited research on why consumers choose to disclose personal information on OSN despite valuing in privacy. With big data analytics taking the center stage and OSNs becoming as the primary source of big data, privacy and security in social networks have become increasingly important. This research looks at factors that may explain individual's information disclosure behavior. We use prospect theory and rational apathy theory in our research model. As outlined in our research model, the information disclosure decision of the invidiual depends on variables such as the privacy risk, protection beliefs, and perceived benefits to disclose information on OSNs. The users information disclosure decisions would also be affected by their belief about who owns the information being provided and how the information that has already been collected and stored by social media companies is being used by these companies. Thus, this study may help us expand the concept of apathy, risk belief and privacy calculus in regard to the context of information disclosure behavior.

We will be using survey research as our research methodology as this helps in increasing the generalizability of this study. The survey instrument for this research will be hosted in Qualtrics. The main data will be collected from students as they represent the general demographic that use online social network the most. Prior to collecting the data, a preliminary investigation will include expert panel reviews, pre-test, and pilot studies to confirm the reliability and validity of the survey instrument. The loadings, cross-loadings, content and face validity, and reliability will also be examined during the pilot study. Then the structural model will be used to test the path coefficient and t-values for our hypotheses.

13.6.2 Key Findings

We will discuss the key findings based on our data analysis in the future publication.

13.6.3 Contribution and Implications

This paper has several theoretical and practical implications. First, this paper brings together the concept of big data and OSNs to analyze the privacy behavior of OSN users. Previous research on privacy and information disclosure has focused on internet transactions, eCommerce, and social networks (Dinev and Hart 2006; Keith et al. 2012) but the concept of big data and its impact on privacy behavior of OSN users has been studied by very few researchers. Second, this paper brings the concept of privacy apathy and perceived ownership into focus. Users of OSNs are believed to be worried about losing privacy in the age of big data. This paper seeks to study if some of the users would care less about privacy if they believe that they have already lost the ownership of their information on the internet. Third, we are also expanding the prospect theory and apathy theory and the concept of the reference point in guiding OSN users to decide about their privacy-related behavior.

IS research has regularly faced the criticism of lacking relevance to practice (Baskerville and Myers 2004; Benbasat and Zmud 1999). As such, this paper provides value to practitioners in many different ways. First, this study is helpful to OSN providers and third parties as these organizations would now understand how consumer's information disclosure behavior works. Second, this study helps us understand why people tend to disclose too much of their personal information on OSNs.

13.6.4 Limitations of this Study

McGrath (1995) stated that all research methods are inherently flawed, though each is flawed differently. Thus, the role of the researcher is always to minimize the flaws associated with the research by maximizing the three criteria of good research: generalizability, precision, and realism). This research is no exception to other research and thus, has its limitations. Some of the limitations of this study pertain to the generalizability of the study due to sample frame used for this study, theoretical constructs excluded from the study to achieve parsimonious research model, research method used for testing the proposed model, and use of self-reported scales. However, understanding these limitations also provides with the opportunities for future research. As this research is a research-in-progress, understanding these limitations will also provide us with the opportunities for strengthening the next steps in our research plan and our future research.

Biographies

Shwadhin Sharma is an Assistant Professor in the College of Business at California State University Monterey Bay. His research interests are in the areas of technical and behavioral aspects of big data analytics, privacy and security, electronic commerce and social commerce, role of dispositional factors in IT, and IT adoption and discontinuation. He has published his research in journals such as Journal of Computers Information Systems, Electronic Commerce Research and Applications and Computers & Security and academic conferences. He has served as reviewer for several reputed journals and conferences. He is currently serving on the editorial board of International Journal of the Internet of Things and Cyber-Assurance. He has also co-chaired as a SIGDSA mini-track on "Social Network Analytics in Big Data Environment" in AMCIS 2016.

Babita Gupta is Professor of Information Systems and the Director of AACSB Accreditation at the College of Business, California State University Monterey Bay. She teaches courses in information technology innovation strategies, business intelligence & analytics, database management, and information systems for decision making. Her research interests are in the areas of online security and privacy, business intelligence strategies, technology adoption, and the role of culture in IT. She has published in journals such as the *Communications of the Association for Information Systems (CAIS)*, *Journal of Electronic Commerce Research (JECR)*, the *Journal of Strategic Information Systems (JSIS)*, the *Communications of the ACM (CACM)*, the *Journal of Industrial Management and Data System*, the *Journal of Scientific and Industrial Research*, the *Journal of Information Technology Cases and Applications*, and the *Journal of Computing and Information Technology*.

She serves as an *Advisory Board Member* of the Teradata University Network, a non-profit division of Teradata.com. She was elected as the *Program-Chair-Elect Officer* in 2012 for the Special Interest Group on Decision Support and Analytics (SIGDSA) under the Association for information Systems (AIS) organization. She has also served as a Board Member of the California Coastal Rural Development Corporation for over a decade.

References

Acquisti A, Gross R (2006) Imagined communities: awareness, information sharing, and privacy on the Facebook, privacy enhancing technologies. Springer, Berlin, pp 36–58

Acquisti A, John LK, Loewenstein G (2013) What is privacy worth? J Leg Stud 42(2):249–274

Afroz S, Islam AC, Santell J, Chapin A, Greenstadt R (2013) How privacy flaws affect consumer perception. Socio-Technical Aspects in Security and Trust (STAST), 2013 Third Workshop on: IEEE, pp 10–17

Anderson JC, Gerbing DW (1988) Structural equation modeling in practice: a review and recommended two-step approach. Psychol Bull 103(3):411–423

Baskerville R, Myers MD (2004) Special issue on action research in information systems: making is research relevant to practice foreword. Manag Inf Syst Q 28(3):329–335

Benbasat I, Zmud RW (1999) Empirical research in information systems: the practice of relevance. Manag Inf Syst Q 23(1):3–16

Boyd D (2008) Facebook privacy trainwreck: exposure, invasion and social convergence. Convergence 14(1):13–20

Boyd DM, Ellison NB (2007) Social network sites: definition, history, and scholarship. J Comput-Mediat Commun 13(1):210–230

Bulgurcu B (2012) Understanding the information privacy-related perceptions and behaviors of an online social network user. University of British Columbia, Vancouver

Chin WW, Todd PA (1995) On the use, usefulness, and ease of use of structural equation modeling in MIS research: a note of caution. Manag Inf Syst Q 19(2):237–246

Churchill GA (1979) A paradigm for developing better measures of marketing constructs. J Mark Res 16(1):64–73

Conroy S, Williams A (2014) Use of internet, social networking sites, and mobile technology for volunteerism. AARP Office of Volunteerism and Service

csmonitor.com (2013) The danger of American apathy on NSA surveillance. The Christian Science Monitor. http://www.csmonitor.com/Commentary/Opinion/2013/0731/The-danger-of-American-apathy-on-NSA-surveillance. Accessed 19 Jan 2014

Demchenko Y, Grosso P, De Laat C, Membrey P (2013) Addressing big data issues in scientific data infrastructure. In: Proceedings of international conference on collaboration technologies and systems (CTS), San Diego, CA, pp 48–55

Dhami A, Agarwal N, Chakraborty TK, Singh, BP, Minj J (2013) Impact of trust, security and privacy concerns in social networking: an exploratory study to understand the pattern of information revelation in Facebook. In: Proceedings of 3rd international advance computing conference (IACC), pp 465–469

Dinev T, Hart P (2006) An extended privacy calculus model for E-commerce transactions. Inf Syst Res 17(1):61–80

Ellison NB, Steinfield C, Lampe C (2007) The benefits of Facebook "friends:" social capital and college students' use of online social network sites. J Comput-Mediat Commun 12(4):1143–1168

Falk RF, Miller NB (1992) A primer for soft modeling. University of Akron Press, Akron

Feuchtl S, Kamleitner B (2009) Mental ownership as important imagery content. Adv Consum Res 36(2):995–996

Furby L (1978) Possession in humans: an exploratory study of its meaning and motivation. Soc Behav Personal 6(1):49–65

Goitein E (2013) The danger of American apathy on NSA surveillance. The Christian Science Monitor. Available on November 3 from http://www.csmonitor.com/Commentary/Opinion/2013/0731/The-danger-of-American-apathy-on-NSA-surveillance. Accessed 31 Jul 2013

Hoadley MC, Xu H, Lee J, Rosson MB (2010) Privacy as information access and illusory control: the case of the Facebook news feed privacy outcry. Electron Commer Res Appl 9(1):50–60

Hogben G (2007) Security issues and recommendations for online social networks. Enisa Position Paper 1:1–36

Hugl U (2011) Reviewing person's value of privacy of online social networking. Internet Res 21(4):384–407

Johnson M, Egelman S, Bellovin SM (2012) Facebook and privacy: it's complicated. In: Proceedings of the eighth symposium on usable privacy and security, New York, USA, pp 9–15

Kahneman D, Tversky A (1979) Prospect theory: an analysis of decision under risk. Econometrica 47(2):263–291

Karuitha JK, Onyuma SO, Mugo R (2013) Do stock splits affect ownership concentration of firms listed at the Nairobi securities exchange? Res J Finan Acc 4(15):105–117

Keith MJ, Thompson SC, Hale J, Greer C (2012) Examining the rationality of information disclosure through mobile devices. In: Proceedings of 33rd international conference on information systems, Orlando, USA, pp 1–17

Kline RB (1998) Principles and practice of structural equation modeling. Guilford Press, New York

Krasnova H, Spiekermann S, Koroleva K, Hildebrand T (2010) Online social networks: why we disclose. J Inf Technol 25(2):109–125

Krasnova H, Veltri NF, Günther O (2012) Self-disclosure and privacy calculus on social networking sites: the role of culture. Bus Inf Syst Eng 4(3):127–135

Kshetri N (2014) Big data's impact on privacy, security and consumer welfare. Telecommun Policy 38(11):1134–1145

Lenhart A, Purcell K, Smith A, Zickuhr K (2010) Social media and mobile internet use among teens and young adults. Millennial. Pew Internet and American Life Project, Washington

Li H, Sarathy R, Xu H (2010) Understanding situational online information disclosure as privacy calculus. J Comput Inf Syst 51(1):62–71

Li H, Sarathy R, Xu H (2011) The role of affect and cognition on online consumers' decision to disclose personal information to unfamiliar online vendors. Decis Support Syst 51(3):434–445

Lindell MK, Whitney DJ (2001) Accounting for common method variance in cross-sectional research designs. J Appl Psychol 86(1):114–121

Malhotra NK, Kim SS, Agarwal J (2004) Internet users' information privacy concerns (IUIPC): the construct, the scale, and a causal model. Inf Syst Res 15(4):336–355

Manyika J, Chui M, Brown B, Bughin J, Dobbs R, Roxburgh C, Byers A (2011) Big data: the next frontier for innovation, competition, and productivity. McKinsey Global Institute

McGrath E (1995) Methodology matters: doing research in the behavioral and social sciences. In: Human-computer interaction. Morgan Kaufmann, San Francisco, pp 152–169

Metzger MJ (2004) Privacy, trust, and disclosure: exploring barriers to electronic commerce. J Comput-Mediat Commun 9(4)

Nadkarni A, Hofmann SG (2012) Why do people use Facebook? Personal Individ Differ 52(3):243–249

O'Brien D, Torres AM (2012) Social networking and online privacy: Facebook users' perceptions. Ir J Manag 31(2):63–97

Ormond DK (2014) The impact of affective flow on information security policy compliance. In: M.S. University (ed) pp 1–178

Pavlou PA, Liang H, Xue Y (2007) Understanding and mitigating uncertainty in online exchange relationships: a principal-agent perspective. Manag Inf Syst Q 31(1):105–136

Podsakoff PM, MacKenzie SB, Lee J-Y, Podsakoff NP (2003) Common method biases in behavioral research: a critical review of the literature and recommended remedies. J Appl Psychol 88(5):879–903

Raschke RL, Krishen AS, Kachroo P (2014) Understanding the components of information privacy threats for location-based services. J Inf Syst 28(1):227–242

Richards NM, King JH (2013) Three paradoxes of big data. Stanford Law Review Online 66(41):41–46

Rosenblum D (2007) What anyone can know: the privacy risks of social networking sites. IEEE Secur Priv 5(3):40–49

Sarfaraz A, Ahmed S, Khalid A, Ajmal MA (2012) Reasons for political interest and apathy among university students: a qualitative study. Pak J Soc Clin Psychol 10(1):61–67

Sharma S, Crossler RE (2014) Disclosing too much? Situational factors affecting information disclosure in social commerce environment. Electron Commer Res Appl 13(5):305–319

Simon HA (1982) Models of bounded rationality. MIT Press, Cambridge

Solmitz DO (2000) The roots of apathy and how schools can reduce apathy. Available from https://dwaynehoward.wordpress.com/2012/05/14/the-roots-of-apathy/

Solove DJ (2008) Understanding privacy. Harvard University Press, Cambridge

Stieglitz S, Dang-Xuan L, Bruns A, Neuberger C (2014) Social media analytics. Bus Inf Syst Eng 6(2):89–96

Tan W, Blake MB, Saleh I, Dustdar S (2013) Social-network-sourced big data analytics. IEEE Internet Comput 5:62–69

theguardian.com (2013) Software that tracks people on social media created by defence firm. Available from http://www.theguardian.com/world/2013/feb/10/software-tracks-social-media-defence

Tufekci Z (2008) Can you see me now? Audience and disclosure regulation in online social network sites. Bull Sci Technol Soc 28(1):20–36

Turel O, Serenko A (2012) The benefits and dangers of enjoyment with social networking websites. Eur J Inf Syst 21(5):512–528

Van Dyne L, Pierce JL (2004) Psychological ownership and feelings of possession: three field studies predicting employee attitudes and organizational citizenship behavior. J Organ Behav 25(4):439–459

Westin AF (1967) Privacy and freedom. Atheneum, New York

Wolfe M, Laufer RS (1974) The concept of privacy in childhood and adolescence. In: Margulis ST (ed) Privacy as a behavioral phenomenon, symposium presented at the meeting of the environmental design research association, Milwaukee

Xu H (2007) The effects of self-construal and perceived control on privacy concerns. In: Proceedings of international conference on information systems, Montreal, Canada, pp 1–14

Xu H, Teo HH (2004) Alleviating consumer's privacy concern in location-based services: a psychological control perspective. In: Proceedings of the twenty-fifth international conference on information systems, Charlottesville, Virginia, pp 793–806

Xu H, Dinev T, Smith HJ, Hart P (2008) Examining the formation of individual's privacy concerns: toward an integrative view. In: Proceedings of 29th international conference on information, Paris, France, pp 1–16

Xu H, Teo HH, Tan BC, Agarwal R (2009) The role of push-pull technology in privacy calculus: the case of location-based services. J Manag Inf Syst 26(3):135–174

Xu F, Michael K, Chen X (2013) Factors affecting privacy disclosure on social network sites: an integrated model. Electron Commer Res 13(2):151–168

Yoo CW, Ahn HJ, Rao HR (2012) An exploration of the impact of information privacy invasion. In: Proceeding of thirty third international conference on information systems, Orlando, Florida, pp 1–18

Zikopoulos PC, Eaton C, Deroos D, Deutsch T, Lapis G (2012) Understanding big data: analytics for enterprise class and streaming data. McGraw-Hills books. eBook, http://public.dhe.ibm.com/common/ssi/ecm/en/iml14296usen/IML14296USEN.PDF