

2017

The Efficacy of Perceived Big Data Security, Trust, Perceived Leadership Competency, Information Sensitivity, Privacy Concern and Job Reward on Disclosing Personal Security Information Online

Iqbal Amiri

Nova Southeastern University, iamiri@gmail.com

This document is a product of extensive research conducted at the Nova Southeastern University [College of Engineering and Computing](#). For more information on research and degree programs at the NSU College of Engineering and Computing, please click [here](#).

Follow this and additional works at: https://nsuworks.nova.edu/gscis_etd

 Part of the [Computer Sciences Commons](#)

Share Feedback About This Item

NSUWorks Citation

Iqbal Amiri. 2017. *The Efficacy of Perceived Big Data Security, Trust, Perceived Leadership Competency, Information Sensitivity, Privacy Concern and Job Reward on Disclosing Personal Security Information Online*. Doctoral dissertation. Nova Southeastern University. Retrieved from NSUWorks, College of Engineering and Computing. (1024)
https://nsuworks.nova.edu/gscis_etd/1024.

This Dissertation is brought to you by the College of Engineering and Computing at NSUWorks. It has been accepted for inclusion in CEC Theses and Dissertations by an authorized administrator of NSUWorks. For more information, please contact nsuworks@nova.edu.

The Efficacy of Perceived Big Data Security, Trust, Perceived Leadership
Competency, Information Sensitivity, Privacy Concern and Job Reward on
Disclosing Personal Security Information Online

by

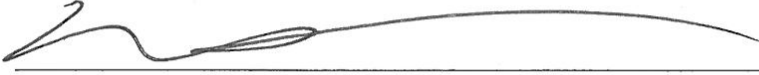
Iqbal Amiri

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in
Information Systems

College of Engineering and Computing
Nova Southeastern University

2017

We hereby certify that this dissertation, submitted by Iqbal Amiri, conforms to acceptable standards and is fully adequate in scope and quality to fulfill the dissertation requirements for the degree of Doctor of Philosophy.



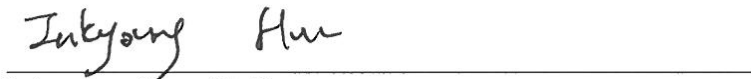
Ling Wang, Ph.D.
Chairperson of Dissertation Committee

12/11/2017
Date



Yair Levy, Ph.D.
Dissertation Committee Member

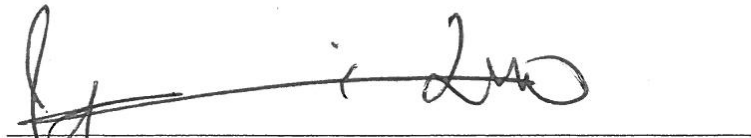
12/11/07
Date



Inkyoung Hur, Ph.D.
Dissertation Committee Member

12/11/07
Date

Approved:



Yong X. Tao, Ph.D., P.E., FASME
Dean, College of Engineering and Computing

12/11/07
Date

College of Engineering and Computing
Nova Southeastern University

2017

An Abstract of a Dissertation Submitted to Nova Southeastern University in Partial Fulfillment
of the Requirements for the Degree of Doctor of Philosophy

The Efficacy of Perceived Big Data Security, Trust, Perceived Leadership
Competency, Information Sensitivity, Privacy Concern and Job Reward on
Disclosing Personal Security Information Online

by

Iqbal Amiri

November 2017

Individuals' reluctance to provide sensitive personal information online could affect the US Governments' ability to hire and retain qualified personnel for sensitive cleared positions. The aim of this research study was to show how perceived big data security, trust, perceived leadership competency, information sensitivity, privacy concern and reward of a job play a significant role in limiting an individuals' willingness of disclosing sensitive personal information online. While a significant volume of research has examined information disclosure in the health care field, there has not been any published studies on the willingness of online disclosure in order to attain a US Government job. Therefore, this study was undertaken to address this gap, where the principles of Utility Theory were applied, which posits that people make choices by maximizing their utility function over multiple choices.

This study was a quantitative study that collected data through online survey using a 7-Point Likert Scale. Random sampling was used to collect data by sending the survey link through email and through Survey Monkey's participant outreach program to random participants. Partial Least Square Structural Equation Modeling (PLS-SEM) was used to analyze the data collected from a total of 206 responses received.

Based on the results, it was found that leadership competency, trust in website and job reward have a significant impact on an individual's willingness to disclose, while perceived big data security and privacy concern did not. It is recommended that the government thoroughly vet leaders in charge, as increase in perceived leadership competency has shown to have an increase in website trust, eventually leading to an individual's willingness to disclose. Of particular interest and contrary to previous studies on information disclosure, privacy concern did not show a significant influence on willingness to disclose information online. Similarly, from the three personality traits of extraversion, intellect and conscientiousness, only individuals with the conscientiousness trait, showed to have any significant impact on privacy concern. Finally, the aim of this study was to help the government understand online disclosure reluctance in order to hire and retain qualified personnel for cleared positions and contribute to the body of knowledge.

Acknowledgments

I would like to thank first and foremost my family – Uzma, my wife and my three kids –Sarina, Shiraz and Salman who have been very patient with me and given me space and time for me to concentrate on my PhD program. There have been lots of times where my wife had to take on extra responsibilities while I was dedicating my time on this great endeavor, therefore, I am very appreciative of her understanding.

I would like to thank my chair person, Dr. Wang, who has diligently helped me by guiding me and supporting me through this entire process. Through her knowledge and extensive experience in this field, I have gained much knowledge and insight into researching in the Information Systems field. Additionally, my study would have not been complete without the guidance and support of my committee members Dr. Levy and Dr. Hur and their valuable and insightful feedback. Dr. Hur has been instrumental in providing realistic and very helpful feedback to make this research possible.

Additionally, I would like to thank Dr. Levy, who has been a great professor and a mentor that has helped me throughout the PhD program and due to his extensive knowledge and passion for the Information Systems field, he has been a major source of motivation for me.

Table of Contents

Abstract iii

List of Tables vii

List of Figures viii

Chapters

1. Introduction 1

Problem Statement 3

Dissertation Goal 7

Research Questions and Hypotheses 9

Relevance and Significance 15

Barriers and Issues 16

Assumptions, Limitations and Delimitations 18

Summary 19

2. Review of the Literature 21

Overview 21

Foundational Literature 21

Theory 23

Criteria Justification 27

Past Literature Gaps 35

Analysis of Research Methods Used 36

Literature Synthesis 37

Summary 38

3. Methodology 40

Research Methods 40

Instrument Development 41

Validity and Reliability 43

Sample 45

Data Analysis 46

Results Format 47

Resource Requirements 48

Summary 50

4. Results 51

Pre-Analysis Data Screening 51

Data Analysis 54

Findings 59

5. Conclusions, Implications, Limitations and Summary 62

Conclusions 62

Implications 62

Limitations and Future Studies 65

Summary 66

Appendices

A. Survey Questionnaire 69

B. IRB Approval 77

C. Data Collected 78

D. Mahalanobis Distance and Box Plot 80

E. Rerun of Mahalanobis Distance and Box Plot after deleting Extremes 83

F. Normality and Scatter Plot 86

G. Descriptive Statistics 88

H. PLS Analysis 92

I. Model Fit, Validity, Reliability, Coefficient, Outer Loading 93

J. Rerun of PLS Analysis with PBDS1 and PBDS2 deleted 97

K. Model Fit, Validity, Reliability, Coefficient, Outer Loading 98

L. Significance with Bootstrapping 101

References 102

List of Tables

Tables

1. Three Personality traits 10
2. Big Data and the 3V's 32
3. Constructs and Instrument Source 42
4. Model Fit 55
5. PLS Factor Analysis 56
6. Construct Validity and Reliability 57
7. Discriminant Validity 58
8. Summary of Hypothesis Tests 61

List of Figures

Figures

1. Research Model: Willingness to Disclose 14
2. PLS Analysis Result for Willingness to Disclose 59

Chapter 1

Introduction

Cybersecurity breaches are a clear risk to the privacy of an individual's sensitive and personal data. As of July 2015, 888 cybersecurity breaches were reported involving some - 245.9 million records compromised worldwide for just that single year (Gemalto, 2015). Sensitive personal information is disclosed and shared online on numerous websites for shopping, mortgaging, banking as well as health and security clearance related transactions. As information disclosed on these online sites include personal identifiable information (PII) that can easily be shared to unwanted parties, it is likely for individuals to have concerns for cyber security, privacy, information protection as well as trust of online systems. As mentioned by Westin (2003), online users have a serious privacy concern about how their personal information is used, disclosed, and protected, and the degree of control they have over the dissemination of this information. Adding to this fear is the possible undesirable economic and social consequences resulting from the misuse of such information (Luck, Chang, Brown, & Lumpkin, 2006). Studies have mentioned that 88.2% of Internet users express concern about the privacy of their personal information (USC Annenberg School of Communication, 2004).

As mentioned by Beldad, Van Der Geest, Jong, and Steehouder (2012), the success of an online transaction with a government organization depends on citizen's willingness to share personal data relevant for the transaction. Furthermore, they added that disclosing personal data online is often times considered risky and that it has been substantiated in a number of empirical studies that perceptions of the risks involved in sharing personal data in the virtual environment

could hinder Internet users from engaging in online transactions that require the disclosure of personal data. Similarly, securing a job with the US Government require an individual to disclose their sensitive personal information online, through the Office of Personnel Management (OPM), which is the primary federal agency tasked with conducting and storing data related to the majority of federal background investigation used to gain security clearances (Sanger & Davis, 2015). Furthermore, in light of OPM's recent hacking and loss of data for 21.5 Million individuals from the OPM's servers, this uneasiness of providing their sensitive personal information online has exacerbated for individuals who were already concerned about their privacy (Castelluccio, 2015). On an article on Big Data and OPM hacking, Gertz (2016) highlighted that the ability to analyze Big Data tranches is now available with refined software, a development that has made the targeting of databases increasingly attractive to different nation states and as per Admiral Rogers this trend is likely to continue, "what you saw at OPM, my comment would be you are going to see a whole lot more" (para. 7). Additionally, Admiral Rogers called the compromise of 22 million records from the Office of Personnel Management, as well as millions of health care records in an earlier attack disclosed last year, a new form of cyber spying and stated that, "China's theft of millions of records on Americans was part of a big data spying program conducted by Beijing that has prompted the Pentagon to take new steps to secure large data concentrations" (para. 1).

Adding to the OPM hacking incident, Gallagher (2016) mentioned that the initial ongoing OPM attack was uncovered using the Department of Homeland Security's (DHS) – 'Einstein', which is the multi-billion-dollar intrusion detection and prevention system that stands guard over much of the federal government's Internet traffic. The author further added, that once baseline criminal attacks and network espionage tactics have been executed, Einstein may view the traffic

analysis as normal network traffic and only detect intrusions already in progress rather than prevent them from happening. By understanding the need of cyber security in order to protect sensitive big data of cleared individuals, it is the intent of this research study to understand the factors that affect an individuals' willingness to disclose sensitive personal information online in order to secure a cleared government job. There have been numerous studies that have separately researched big data cybersecurity, privacy concerns and previous online experience specially in the healthcare industry, but there has not been any published study found that has researched the effects perceived big data and cyber security, trust, perceived leadership competency, information sensitivity, privacy concern and job reward on the willingness of the user applying for sensitive positions online for secure government jobs. Additionally, the following research question was addressed by this study - *How does perceived big data security, information sensitivity, privacy concern, perceived leadership competency, trust of the online system and reward of getting a job affect an individuals' willingness of disclosing personal sensitive information online?*

Hence, this study will be important for not only practitioners but also to researchers in this field. Specifically, the US Government because individuals' reluctance to provide sensitive personal information online could affect the US Governments' ability to hire and retain qualified personnel for sensitive cleared positions.

Problem Statement

Individuals seeking jobs within the federal government, go through a set of background checks to meet the minimum clearable criteria. Working for the federal government requires an individual to go through some form of security clearance suitability process that can range from public trust to top secret. With the growth of big data and enormous amount of sensitive personal

identifiable information that is collected by the government, there is sometimes uneasiness about providing such personable information to any entity. Reluctance to provide sensitive personal information on online government clearance website could impede the success of online clearance process as well as slow down the ability of the government to find and hire cleared individuals for sensitive positions. Broken security clearance process has serious consequences, some of which include negative impact on those seeking to serve, and on the overall safety of our nation (Oversight of Government Management, 2005).

Prior to online systems being available to provide personal individual information, secure information exchange was accomplished through the use of paper forms. This process was slow and tedious as it required information submitted in paper form by individuals to be shared across several departments. Highly-skilled employees sat idly by for months, waiting for their security clearances to be finalized, while important national security work was not being done. Many people were dissuaded by the long process and looked-for opportunities elsewhere, thus denying the government of many hard working and smart people. Finally, government employees who already held security clearances were nevertheless faced with lengthy reinvestigations while seeking jobs in other agencies that required clearances (Oversight of Government Management, 2005). DOD's performance for completing the security clearance process is 75 days for an initial secret clearance, 120 days for an initial top secret, and 180 days for a reinvestigation of a top-secret clearance. Yet in fiscal year 2003, on average, it took 375 days for a security clearance to make its way through the whole process (Government Reform Committee, 2004). To alleviate, this delay and with the advent and improvement in Information System and IT Security, digital forms were developed and marketed for secure sensitive information submission. In an effort to improve the security clearance issuing process, in November 2003, Congress authorized a

proposed transfer of DOD's personnel security investigative functions and more than 1,800 investigative employees to the Office of Personnel Management [OPM] (Government Reform Committee, 2004). Shortly thereafter the OPM system for collection sensitive information was hacked, and as mentioned by Castellucio (2015), sensitive personal data of 21.5 million people was silently siphoned out of the agency's servers which added to an individuals' fear of having their personal information stolen from a government owned data repository.

As per Beldad et al. (2012), the success of an online transaction with a government organization depends on citizen's willingness to share personal data relevant for the transaction which is oftentimes considered risky. Privacy issues in the Internet age have received significant attention over the past few years. For example, allegations of governments spying on their citizens and new laws such as the "right to be forgotten" have opened up a whole range of debate. Internet users' perceptions of information privacy risks in an online environment can be attributed to the fact that users do not know exactly who is gathering the data and what is being done with them (Resnick & Montania, 2003). Perera, Ranjan, Wang, Khan, and Zomaya (2015) have raised the issue of government standardization and have mentioned that either the government or independent regulatory bodies must lead and enforce standardization and legal efforts. Standardization efforts should comprise both a certification process and a technology development process. But the question raised by these authors in their study which is extremely significant for this research is, '*what happens when the system that the government uses to secure, is itself compromised*'? OPM's data compromise is an example of such a scenario and to a larger extent the focus of this study in big data and cyber security. In spite of having a multi-billion dollars' intrusion detection system - "Einstein", that should have protected individuals PII

information from cyber security attacks, it detected intrusions in progress rather than prevent it (Gallagher, 2016).

As mentioned by Figueroa (2015), technology has progressed significantly in the last several decades, to the point where a system can be programmed and allowed to run with very little need for human interaction or supervision; nevertheless, the reality is that the human component in securing data remains the largest contributor to breach, loss, and theft. Furthermore, he added that these advancements have given rise to the threat of hacking organizations, corporate espionage, and state-sponsored government espionage.

The problem that this research addresses is that an individuals' reluctance in providing sensitive personal information online can affect the US Governments' ability to hire and retain qualified personnel for sensitive cleared positions. Therefore, to understand and address this reluctance by the individuals, the relationship and effects of these constructs was studied to understand their effect on the dependent variable – willingness to disclose personal information.

There have been numerous studies that have separately researched big data cybersecurity, privacy concerns and previous online experience, especially in the healthcare industry, but there has not been any published study found that has researched willingness to disclose information online to attain secure government jobs. Some examples of prior government clearance studies include studies in clearance process delay, discrimination on clearance awarding process or OPM security from the legal aspect but this was one of the first studies that researched big data and cyber security as well as the effects of trust, leadership, information sensitivity, privacy concern and job reward on the willingness of disclosing personal sensitive information online. Therefore, given the limited research in this area, it is the belief that there is a need for further research in this field, which can benefit researchers as well as practitioners of the US clearance processes.

Some relevant prior studies for this research includes studies in privacy by Osatuyi (2015), privacy, trust and e-government by Beldad et al. (2012) and privacy, personal dispositions and healthcare by Bansal, Zahedi, and Gefen (2010). These three studies were foundational for this research and extensive analysis and details on these studies have been highlighted in the literature review chapter. Based on these prior studies, this study was also a quantitative study that used surveys to collect user data. This study argues that an individuals' willingness to disclose sensitive information, which is the dependent variable, depends on perceived big data security, privacy concern, job reward, perceived leadership competency and trust of the online system, which are determined by personal dispositions such as personality traits and act as the independent variables in this study.

Finally, a relevant theory is important for the research to be based on, and as per Gregor (2016), to understand Information Systems, a theory is required that links the natural world, the social world, and the artificial world of human constructions. Earlier studies in this field, were based on various theories including but not limited to trust theory, leadership theory, social exchange theory, utility theory and competency theory but this study will be based on the foundations of utility theory and its application to choice theory, which details the fact that consumer preferences depend on personal characteristics. Utility Theory posits that people make choices by maximizing their utility function over multiple choices (or alternatives) (Ben-Akiva & Lerman, 1985; Luce, 1959). Further discussion on Utility Theory and its application to this study is detailed in the literature review chapter.

Dissertation Goal

The process of obtaining a security clearance is already a long process that includes background checks and investigations and it is not this research's intention to study the security

clearance process but rather focus on the factors that are relevant predictors of providing sensitive personal information to an online system that stores data for clearances. Specifically, the purpose of this research was to determine the effect the following independent variables of perceived big data security, trust, perceived leadership competency, information sensitivity, privacy concern and job reward will have on the dependent variable which is the individuals' willingness to disclose sensitive secure information online. This study looks into the various factors that contribute to an individual's reluctance in disclosing sensitive personal information while applying for a secure government job online. To measure these constructs, this study collected data using random sampling where all possible subsets of a population were given an equal probability of being selected. In this study, the unit of analysis was individuals and the study was not longitudinal, where data is collected over a long period during different times, but rather was collected one time only using the cross-sectional method. Through data collection and analysis, it was the intention of this research to look at the results and provide conclusive feedback that could be beneficial to the US Government to consider when attracting individuals applying for secure online jobs. It is also hoped that through this study the ability of the US Government to find and hire cleared individuals for sensitive nations security positions could possibly increase.

Another goal of this study was to research big data and how its security is vitally important for an individual applying for a secured job. It is the intent of this study to contribute to the existing body of knowledge on Big Data Security as well as aid in increasing knowledge on willingness of providing sensitive Personal Identifiable Information (PII) online for the scientific research community as well. As Personal Identifiable Information (PII) is collected from individuals for security clearance process, the amount of data collected can grow fairly easily.

This large set of data is categorized as Big Data and is differentiated from other forms of collected data only if it can satisfy the 3V's – Velocity (how fast data is retrieved), Variance (retrieved from different types of storage) and Volume (extremely large amount of data). In their research on Big Data, Perera et al. (2015) mentioned that big data has no clear definition, but it isn't wholly about size either. Therefore, to define and understand Big Data for this research, it is important to first, understand its characteristics which has been expanded in detail in chapter 2 literature review. Based on the goals of the study, some research questions and hypothesis have been developed for this research.

Research Questions

Based on the foundational elements as well as the constructs highlighted in this study, the following research questions were developed:

1. What is the effect of perceived big data security on an individual's privacy concern?
2. What is the effect of personality traits, such as that of, intellect, extraversion and conscientiousness on an individuals' privacy concern?
3. What is the effect of perceived leadership competency on an individual's trust of the online system?
4. How does privacy concern affect the willingness of individuals' disclosing sensitive information online?
5. How does perceived leadership competency affect an individuals' trust of the online system?
6. How does the reward of getting a job affect the relationship between trust and an individuals' willingness of applying online?

Based on the research questions, constructs and research model mentioned in this study, some hypothesis highlighted below, have been developed for this study.

There have been numerous studies that have researched personality traits and the health care system but one such study by Osatuyi (2015), examined the relationship between Big Five personality traits (Intellect, Extraversion, Agreeableness, Conscientiousness and Neuroticism) and privacy concern. Detailed definition of the three personality traits used in the study that are being collectively called Information Sensitivity, have been highlighted in Table 1.

Table 1

Three Personality Traits

Personality Trait	Definition
Extraversion	Extraversion is used to describe individuals that are full of life, energetic, dominant, gregarious, and outgoing and have been found to be interested in leadership positions (McCrae & Costa, 1991). Thus, these individuals were concerned about the erroneous uses of personal information.
Conscientiousness	Conscientiousness is the most widely studied personality trait of the big five traits and conscientiousness individual will sift through a variety of reputable information on privacy before submitting their information online (Osatuyi, 2015).
Intellect	Intellect is defined as openness to experience or an individual's propensity to try new things, to learn to be curious and intellectually challenged (McCrae & Costa, 1991). As per Osatuyi (2015), intellects are quick to learn the implications of sharing too little or too much information and as such, can become a reputable consult for information privacy practices.

The two other personality traits that are not included in this study as they are considered out of scope are Agreeableness and Neuroticism. Agreeableness is described as the act of being trusting, sympathetic, straightforward, and selfless (McCrae & Costa, 1987) which is in contrast to Neuroticism, which is described as the act of being anxious and angry (McCrae & Costa, 1991). Osatuyi (2015), in his research concluded that the refusal to share personal information

during online transactions is largely due to computer anxiety, which can be abated by understanding users concern for privacy. Similarly, for this study, to understand a citizens' reluctance in disclosing their secure information, it is important to also understand their personality traits as these traits might also impact how individuals disclose their personal information. Hence, the following hypothesis were formulated for this study,

H1a: An increase in the individuals' personality trait of intellect will lead to an increase on an individuals' privacy concern.

H1b: An increase in the individuals' personality trait of extraversion will lead to an increase on an individuals' privacy concern.

H1c: An increase in the individuals' personality trait of conscientiousness will lead to an increase on an individuals' privacy concern.

It is perceived that individuals with privacy concern issues might also have concerns of where their personal information might eventually end up when such information is provided into an online system. With OPM's hacking, Figueroa (2016) mentioned that the latest EINSTEIN intrusion detection iteration, EINSTEIN 3 Accelerated or "E3A" is purportedly capable of detecting the types of intrusions that occurred at OPM. In her study, she reports that while the system has been functional for a short while, Department of Homeland Security has been unsuccessful at securing its implementation across the federal agency network; the agency remains confident it will continue to expand the systems reach in order to detect future threats. The data collected on millions of individuals applying for security clearance is considered – "Big Data" as it satisfies the 3V's of - volume, variance and velocity. Detailed explanation and characteristics of the 3V's are highlighted in Table 2. As sensitive PII information is uploaded, stored and downloaded, it becomes imperative that not only this big data gets secured but also

the systems where the data resides are accessed from or reported from are equally secure as well. Therefore, for this study, it is more likely for individuals to have concerns of not only where their sensitive information resides but also of its security when providing their sensitive personal information online, hence, the following hypothesis is formulated,

H2: A decrease in perceived big data security will lead to an increase on an individual's privacy concern.

With OPM's Big Data Security breach and the shortly thereafter resignation by the director of OPM, Katherine Archuleta (Castelluccio, 2015), it is believed that there might be some type of relationship that might exist between perceived leadership competency and an individual's trust of the online system. Thus, it is proposed in this study that lack of leadership competency and trust in leaders managing big data security inhibits the ability of individuals to trust these individuals, which raises trust issues of online systems, and consequently becomes detrimental to an agencies' ability to hire. Thus, it is proposed that 'leadership competency' inhibits the ability of citizens to trust these government leaders. This in return, raises trust issues of the online system. Hence, the following is hypothesis is formulated,

H3: An increase in perceived leader competency will lead to an increase in trust of the online system.

Bansal et al. (2010), in their research considered the impact of personal disposition and privacy concern on an individuals' intention of disclosing personal information for healthcare. In their research, they concluded that privacy concern has an effect on the possibility of individuals disclosing their healthcare information online. Similarly, as this study draws from the foundations of their study, it also focuses on privacy concern of individuals, and its effect on the willingness of the individual disclosing personal information online, therefore, it is important to

study these two constructs, their relationship and effect on one another, hence the following hypothesis is formulated,

H4: An increase in privacy concern will lead to a decrease in the likelihood of individuals disclosing sensitive information online.

The study by Beldad et al. (2012) looked into the risks and trust factors that individuals encounter when disclosing their personal information online and concluded that citizens' levels of trust in government organizations are instrumental in influencing their intention to supply personal data for online government services. Several studies on e-government have also indicated that trust is an essential ingredient for the acceptance and adoption of online government services (Belanger & Carter, 2008). Trust in organizations is also regarded as an important driver for Internet users' intention and willingness to disclose their personal data for computer-mediated transactions (Moore & Grover, 2010). Treating trust as an acceptance of and exposure to vulnerability (Doney, Cannon, & Mullen, 1998) and extending on the trust proposition further, it is the intention of the researcher to also look into how trust in the website will affect citizens' inclination to disclose their personal information online; hence the following hypothesis is formulated,

H5: An increase in trust of website will lead to an increase in the likelihood of individuals disclosing sensitive information online.

Dinev and Hart's (2006) extended privacy calculus model for e-commerce transactions identified trust and perceived risks as important factors driving people's willingness or disinclination to provide personal data. However, as per Beldad et al. (2012), the model did not consider the impact of expected benefits that can be derived from the data disclosure act as a potential determinant of the intention to disclose personal data.

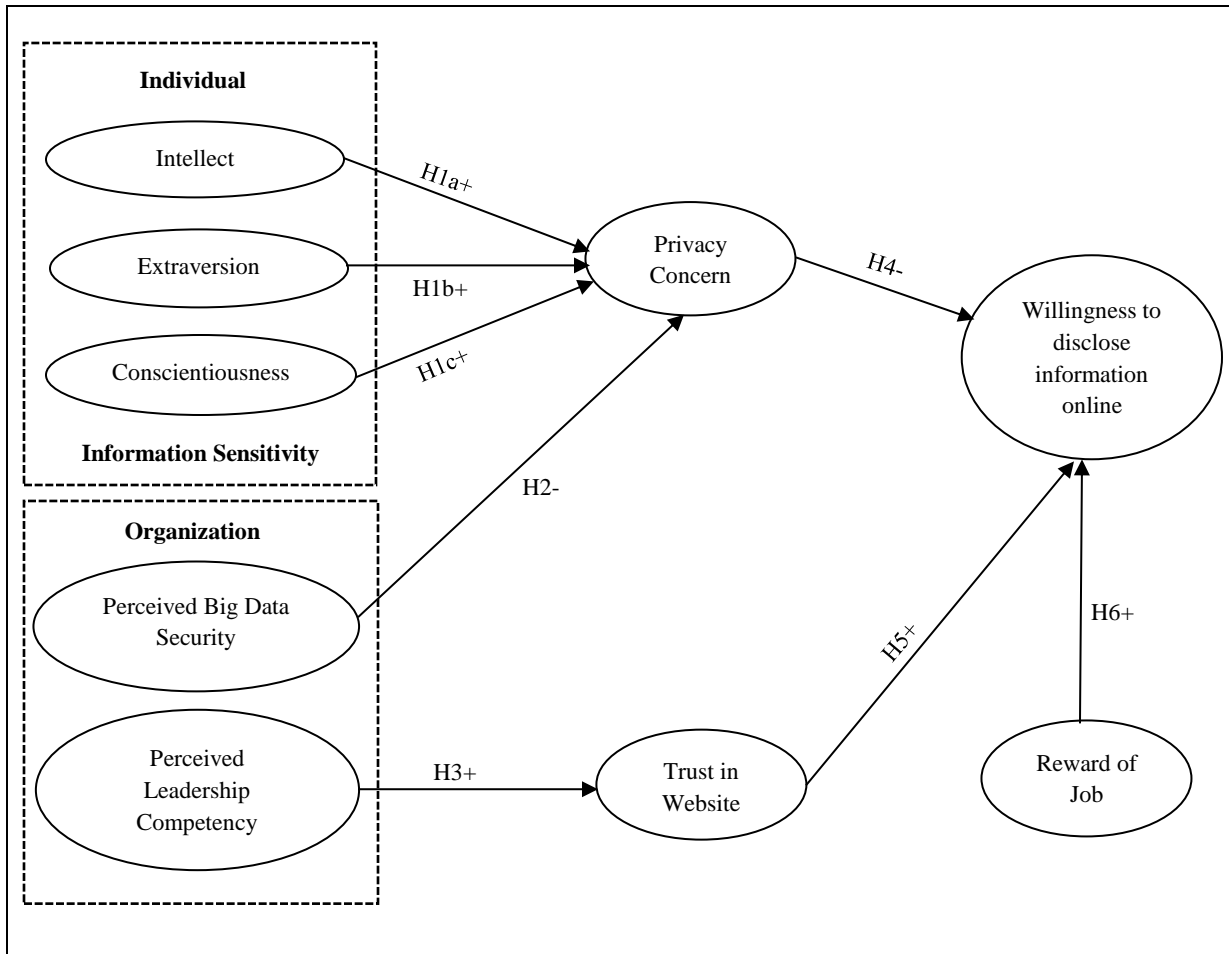


Figure 1. Research Model: Willingness to Disclose

Viewed from a calculus-based perspective (Laufer, & Wolfe, 1977), people's decision to share personal data online, despite the risks of having their online privacy compromised, is driven by expectations of tangible or intangible benefits (Berendt, Gunther, & Spiekermann, 2005). Thus, it is the proposition that the possibility of being able to obtain a stable and steady paying job (benefit) with the federal government will lead to an individual overlooking concerns of online system trust and lead to the individual providing sensitive information online. Hence, it is assumed that this moderating variable will affect both, trust of the website and the willingness of individuals disclosing their personal information, hence the following hypothesis is formulated,

H6: An increase in the ability of an individual of being rewarded with a job will lead to an increase in the willingness of the individual disclosing sensitive information online.

Relevance and Significance

According to Government Reform Committee (2004), delays in the clearance process cause major inefficiencies, which eventually lead to higher costs for taxpayers and ultimately harm national security. New programs, new technologies, and even new government agencies have been developed to deal with the threats appropriately. It is not surprising, then, that the demand for security clearances for both Government employees and industry personnel has dramatically increased over the last few years. Many defense contractor companies are unable to hire otherwise qualified employees because the security clearance process is requiring, on average, over a year to complete, with all signs pointing to continued increases if something does not change. Defense contractor companies often rely on hiring, almost at a premium already cleared employees from other firms, thus increasing contract costs, which are then passed on to the taxpayer. Ultimately, these backlogs hurt national security. When industry employees are hired to work in security programs but cannot work on projects while they are waiting to be cleared, the contracts are not being completed and national security is jeopardized (Government Reform Committee, 2004).

In addition to the existing background of challenges mentioned, there is also reluctance to provide sensitive personal information on online government clearance website that could impede the success of online clearance process as well as slow down the ability of the government to find and hire cleared individuals for sensitive positions. This in turn, can also lead to the government not being able to fill cleared positions which could eventually lead to agencies

either assigning extra work to current employees or possibly lead to agencies being less productive while potentially putting national security at risk.

Ever since the Office of Personal Management (OPM) was hacked and cleared individual's information was compromised, there is potentially uneasiness with individuals providing their sensitive personal information. Additionally, when resources cannot be put to work in these sensitive jobs the pending workload gets redistributed amongst current employees. The aim of this research would be to explore further the cause and effect of the willingness of disclosing sensitive information online so that there was less delay in hiring of qualified candidates and eventually help the nation with its national security needs.

This research is considered significant in the IS field due to the fact that there have been prior studies that have researched the effects of privacy, risk and trust or information sensitivity separately but there has not been any reported research found that has studied all of these constructs as well as big data security, job reward and leadership in one study. Additionally, this research will not only add to the existing knowledge base but possibly help future researchers in the IS and e-government field because it provides information on data security, trust of an online system, privacy concern, leadership, benefits and willingness of disclosing personal information online.

Barriers and Issues

Survey studies just like most empirical studies tend to base their results and conclusion based on the data that they collect. Therefore, data collection is not only an integral part of the research but definitely important as the results and conclusions was only as good as the instrument used to collect it. The instruments used to collect data have to undergo rigorous validation for accurate data collection and certainly will have to be valid and reliable for this

study. Survey Instrument and their validation is an important topic for “IS Positivist Researchers” and as mentioned by Straub et al. (2004), the argument for validation of instruments is based on the prior and primary need to validate instruments before such other crucial validities as internal validity and statistical conclusion validity are considered. The issue of whether IS positivist researchers were validating their instruments sufficiently was initially raised fifteen years ago by Straub, Boudreau, and Gefen, (2004) where they built upon four prior retrospectives of IS research and concluded that IS positivist researchers continue to face major barriers in instrument, statistical, and other forms of validation. As this study is a survey study, it is the intention to base this research from three follow-up studies by Boudreau et al. (2004), Boudreau et al. (2001), and Gefen et al. (2000) that suggest that the field is moving slowly but steadily toward more rigorous validation in positivist work. Additionally, it is the intention to base this study primarily on the foundational work of Straub et al. (2004), where he offered research heuristics for validation via content, construct, reliability, manipulation and statistical conclusion validity.

As responses to the survey in this study were also going to be from individuals working for the government using their personal computer and on their time, one of the challenges to this study was to find organizations and individuals that are cleared or are in the process of applying for security clearance. It was vitally important to make it clear to these respondents to not respond to the survey on either the governments’ time nor on its’ resources.

Even though the foundation of this study is based on previous research on the intention of information disclosure in the healthcare and e-government field, still this study does addresses a completely new problem in the big data and cybersecurity field for secure government clearance process with adequate scientific difficulty and warrants dissertation level work. By researching

and studying some previous theories, limitations, hypotheses, and constructs that have been vetted and validated in these scientific studies, it is rewarding to pursue and contribute to the scientific knowledge base as well as help practitioners in this field by undertaking scientific research.

Assumptions

For this study, data was collected randomly from the population including from organizations where the researcher had worked for or done business with companies in the past. Most, if not all of these facilities are secure facilities with some form of clearance required of their employees. As this study is a survey-based study, it was assumed that feedback on the survey was from individuals that already possess clearances, have some connection to cleared government work or have been exposed to some type of security in their jobs. It was beyond the scope of this study to check on the individual's clearance level or their ability of obtaining security clearance. Additionally, it was highlighted for respondents to respond to the survey on their time and using their private computers but to validate it was also beyond the scope of this study.

Limitations

As this study is based on security clearance and individuals that possess or are in the process of being granted a clearance, the limitation of this study is that its focus is on only a small percentage of the population. Including, but not limited to individuals that are in the process of applying or would like to apply for US Government cleared jobs. This study does not look at the cause and effect or the causality of the independent variables on the dependent variable – “likelihood of disclosing personal information online” for any e-government online sensitive information disclosure but rather focuses on only US Security Clearance related jobs only.

Delimitations

Random Sampling is used in this study to collect survey data to make the study generalizable, and to collect true values in the population in order for it to not suffer from sampling bias. All possible subsets of the sampling frame were given an equal probability of being selected, but as the surveys were online based only, there was some sampling bias as paper based surveys were not mailed. Furthermore, it is suggested that future studies research on data collected from different samples of the population from different parts of the world instead of from only the United States as in this study.

To limit the scope of this study, the construct of information sensitivity consisted of intellect, conscientiousness and extraversion. Agreeableness and neuroticism did not have data collected for, as it was considered out of scope for this research. Additionally, the process of obtaining a security clearance is already a long and tedious process that includes background checks and investigations and it was not this research's intention to study the security clearance process but rather focus on the cause and effect of providing sensitive personal information to an online system that stores sensitive PII data.

Summary

This introductory chapter of the study focused on defining a problem that exists within the Information Systems field, specifically, individuals' reluctance to provide sensitive personal information online and how it can affect the US Governments' ability to hire and retain qualified personnel for sensitive cleared positions. The aim of the introduction was to give a brief overview of how trust, privacy concern, information sensitivity, perceived big data security, perceived leadership competency and reward of a job play a significant role in limiting an individuals' willingness or likelihood of disclosing sensitive personal information online.

Background of the problem and the dissertation goal were highlighted and explained, followed by some research questions that this study would address.

This chapter also presented the constructs for this research including perceived big data security, privacy concern, trust, perceived leadership competency, job reward and the willingness of disclosing sensitive information online. Based on the research questions, this study highlighted and proposed some hypotheses as well as a conceptual model. Just like any other research, this study also had some barriers and issues that were encountered and therefore, have been duly noted in this chapter. Finally, assumptions, limitations, and delimitations of this research have been detailed to highlight the scope limitation of this research.

Chapter 2

Review of the Literature

Overview

This literature review chapter highlights the three-major foundational studies on which this research is based on and highlights their constructs, theories, hypothesis, research methodology, limitations, instrument reliability and validity as well as the results of the data analysis.

Additionally, several studies that separately researched leadership, trust, big data security and intention of disclosing personal information online have also been detailed in this chapter. Major sections of this chapter focus on the criteria justification of what and why a literature review is included, identification in past literature gaps, theories that have been used as well as the theory that this research was based on, an analysis of their research methods and how it might be beneficial for this research as well as detailing a new perspective on the literature. The constructs of this study are thoroughly detailed in the criteria justification section with extensive background on their inclusion based on past literature review as well. Finally, a summary section is included to summarize details and highlights of this chapter.

Foundational Literature

Three major foundational studies that this research is based on are the studies by Bansal et al. (2010), Beldad et al. (2012) and Osatuyi (2015). Bansal et al. (2010) looked into the process by which personal dispositions including risk and privacy concern impact individuals' trust and behavior intention to disclose personal information online for the healthcare industry. In their research, they concluded that both – trust and behavior intentions are impacted through

information sensitivity and privacy concern. Their recommendation was that these findings should enhance managers' understandings of the information needed to personalize and customize healthcare sites to address personal sensitive information. Additionally, their study was based on the utility theory. Similarly, as this study is based on the principles of Utility theory, detailed discussion on it, its foundation and its application to this research is further expanded in the theory section of this chapter.

Another foundational study that this research will focus on is by Osatuyi (2015), where the researcher focused on how personality traits affect information privacy concern for an individual. Personality traits are defined as an individual's dispositions or tendencies that lead to certain behavioral patterns across situations (Osatuyi, 2015). He based his research on the foundational work by McCrae and Costa (1987), where they developed a Big Five Model to consolidate important traits that were found to be reliable across domains. He mentioned that this model was groundbreaking as it organized an individual's personality trait into the following - extraversion, agreeableness, conscientiousness, neuroticism, and openness to experience (intellect). Osatuyi (2015), further examined the relationship between the Big Five personality traits and its' effect on an individual's privacy concern in order to provide personal information to online merchants and concluded that individuals have a growing concern for their privacy.

Lastly, the third foundational study by Beldad et al. (2012) looked into the factors influencing the behavioral intention to disclose personal data for online Government transactions. In their study, adoption is viewed in terms of citizens' willingness to disclose personal data for e-government services. Expanding on this, they mentioned that this perspective is predicated on the fact that the completion of an electronic form, which presses citizens to supply personal information, precedes the actual online transaction with a government organization. The

constructs used for their study were risks factors, trust of the site, benefits of services, adequacy of legal protection and previous online transaction experience. Data was collected through an online survey implemented by two research agencies in the Netherlands. Both research agencies randomly sent a link to the online questionnaire to approximately 3,500 members of their research panels, which are representative samples against the Dutch national census data. Through this data collection and analysis, they concluded that the perceived risks involved in an online sharing of personal data have been empirically proven to abate Internet users' inclination to engage in electronic exchanges and transactions compelling personal data disclosure. What is certain, as this study shows, is that low perceptions of risks and high levels of trust in government organizations have strong repercussions for Internet users' willingness to disclose personal data online. Trust in government organizations, in particular, has been found to play a very crucial role in augmenting disclosure intentions of users with and without e-government experience. Their study in disclosure of sensitive information to the government was extremely important for this research as it looks into the valid and reliable instruments as well as the research design that this study was based upon.

Theory

As mentioned by Bacharach (1989), a scientific theory is a system of constructs (concepts) and propositions (relationships between these constructs) that collectively presents a logical, systematic, and coherent explanation of a phenomenon of interest within some assumptions and boundary conditions. In simple terms, theories should explain why things happen, rather than just describe or predict. Gregor (2016) in her study, mentioned that theories are abstract entities that aim to describe, explain, and enhance understanding of the world and, in some cases, to provide predictions of what will happen in the future and to give a basis for intervention and action. She

researched the five different types of theories – Analyzing, Explaining, Predicting, Design and Action and Explaining and Predicting (EP) theories and explained that the goal of a theory should be consequently to Analyze, Explain, Predict and Prescribe, especially in the IS field. Expanding on her definition of theory, she mentioned that theories in different fields mean different things, as for example, theory in mathematics and music, mean different things, as knowledge is developed, specified, and used in different ways. As her research was focused on Information Systems and how theories are used in this field, she argued that the nature of theory in IS could differ from that found in other disciplinary areas and a characteristic that distinguishes IS from other fields is that it concerns the use of artifacts in human-machine systems. Most, if not all the studies that this research referenced have based their study on one or multiple theories including but not limited to leadership, competency, utility, social exchange, benefits and trust theory. Beldad et al. (2012) based justification of each hypothesis on a separate theory, including trust and intent to disclose on theory of reason action, risk and threat on protection motivation theory and social exchange theory for benefits. Disclosing personal data in consideration of the benefits that can be derived from the act could aptly be regarded as a form of social exchange. From a social exchange perspective, human behavior and social interaction is an exchange of both tangible and intangible goods (Homans, 1958, 1961). People engaged in exchanges consider what they are giving up as a cost and what they are about to receive as a reward or a benefit, and their behavior changes less as profits (rewards minus costs) are maximized (Homans, 1958). As mentioned by Beldad et al. (2012), tangible benefits for the disclosure of personal data online could be vouchers, cash, or gift items and for this study, the benefit or reward of getting a job.

In the study by Osatuyi (2015), concern for information privacy (CFIP) was researched to explain the effects of personality traits and privacy on the decision to disclose personal sensitive information to online vendors. The research based the study on the foundations of the CFIP theory and studied the effects of computer anxiety and behavioral intention. As similar past studies have been based on the foundations of well-established and vetted theories, similarly, it is the intent of this research to base this study on the principles of Utility Theory. Similar to the foundational study by Bansal et al. (2010) and their research on Utility Theory justification, this research will also justify the use of this theories measure of utility, disutility and desirable attributes. Lee (2001), mentioned that research in the information systems field examines more than just the technological system, or just the social system, or even the two side by side; in addition, it investigates the phenomena that emerge when the two interact. Therefore, to investigate the phenomena of interest that emerge when the constructs and the propositions of this study affect each other, this study will depend on the foundations of the Contemporary Utility Theory.

Individuals disclose sensitive personal information online to obtain security cleared jobs, for monetary gain, status or to improve their living situation and career. Disclosing personal sensitive information online, in order to get a government job is a decision that users have to make when using these online systems. Once the information is disclosed on the online systems, then that information can be leaked or sent to other agencies without requesting permission from the individual. Information sharing on such systems are legally accomplished, which might be highlighted in its' terms of use, but many a times this may occur through mistakes, negligence or even hacking. Once such hacking scenario occurred when the Office of the personnel management (OPM) website was hacked and sensitive personal information on 21.5 million

people was silently siphoned out of the agency's servers, which ultimately led to the resignation of the director of OPM, Katherine Archuleta (Castellucio, 2015). These potential undesirable outcomes are negative incentives, or disutility, in disclosing personal information that increase individuals' privacy concern (Bansal et al., 2010). Hui, Teo and Lee (2007) have noted that such disutility is due to people's desire to avoid unwanted disclosure. Furthermore, Bansal et al. (2010) added that given the usefulness of online services, individuals need to balance this potential disutility against the potential desirable outcomes, or utility, of disclosing their personal sensitive information online. Therefore, this study relies on the principles of Utility Theory where the constructs of this study are highlighted that play a role in individuals' decisions to disclose their personal sensitive information. As mentioned before, Contemporary Utility Theory posits that people make choices by maximizing their utility function over multiple choices (or alternatives) (Ben-Akiva & Lerman, 1985; Luce, 1959). These choices could have compensatory attributes where the utility of some attributes may compensate the disutility of other attributes. The established benefits (desirable attributes) individuals receive from disclosing their personal information online includes reward of a job, career growth and even social status.

As Utility Theory posits that individuals differ in their preferences, the utility of decisions and preferences are not uniform across individuals. Individuals' personal dispositions and circumstances (such as their personality traits, current job status, and privacy concern) influence their preferences and the extent of their judgment about the utility of choices available to them (Bansal et al., 2010). Furthermore, McFadden (2001) added that the expressed preferences of the consumers are functions of their taste template, experience, and personal characteristics, including both observed and unobserved components. In the absence of negative consequences, individuals should readily disclose their personal information in order to benefit in obtaining a

government job. This disutility can be categorized into two factors - reducers or enhancers of disutility. Since people in general regard their sensitive PII as private, privacy concern related to providing personal identifiable information is a disutility enhancer. The level of this disutility should depend on the particular job circumstance including an individual's current employment status and job satisfaction as well as on their personality traits (Extroversion, Conscientiousness and Intellect).

Finally, since it is not easy to measure benefit, satisfaction or happiness from a good or service (willingness to disclose personal information online), some of the ways devised to represent and measure utility include measuring the economic choices (reward of a job, social status or monetary gains). Hence, as represented in the conceptual model, we argue that the reward of job can influence an individual's willingness to disclose personal information online. Additionally, personal dispositions (information sensitivity) and perceived concern of security of the information provided (big data security) influence privacy concern, which also affects the willingness of the individual to disclose personal identifiable information.

Criteria Justification

As this research is a survey based research, there are several constructs that have been identified in this study namely – Perceived Big Data Security, Perceived Leadership Competency, Information Sensitivity, Trust, Privacy Concern, Job Reward and the Willingness of Disclosing Personal Information Online. In order to highlight, “*why*” and “*what*” is included and excluded in this study from prior literature, each construct of this study is detailed separately in order to justify the inclusion and or exclusion of not only the relevant prior literature associated with it but also the justification of the inclusion of the construct for this study as well.

Privacy Concern: Information privacy is defined as the “claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is

communicated to others”. Privacy concern, on the other hand, is the concern over the loss of privacy and the need for protection against unwarranted communication and use of personal information (Kim, Steinfield, & Lai, 2008). This construct was important for this study as the individuals are providing sensitive personal information online and there could be a possibility that these individuals might be concerned for the security of their privacy and unwarranted information leak to third parties. Information privacy concern is about being in control of information, security of information exchange, and whether the collector of this information behaved appropriately (Xu, Teo, & Tan, 2006). In their study, they outline the fact that personal disposition plays an important role in privacy concern which ultimately affects trust of the online system. This study will examine information privacy concern and its relationship to trust, and how personal disposition, comprised of personality traits and various measures of information sensitivity and experience, affects these concerns. As this study focuses on privacy concern, and its effect on the willingness of the individual disclosing personal information, therefore it was one of the constructs of this study.

Perceived Leadership Competency: To understand leadership competency, it is important to understand leadership and leadership styles. Leading an organization is vitally important and as mentioned by Kristen, Dyer, Hoopes, and Harris (2004), leadership has a direct impact on the way companies arrange knowledge, because leaders could set the example for employees. On the other hand, leadership competency is defined as, ‘*The ability to do something successfully or efficiently*’. Therefore, to be defined as having any specific leadership style, one must possess a series of dependent competencies associated with that style (Galvin, Gibbs, Sullivan, & Williams, 2014). Effective leaders are differentiated from other leaders through the exercise of a relatively small range of skill or competence areas (Dulewicz & Higgs, 2003). As projects in IT

services involve multiple stakeholders with different backgrounds and a variety of expertise, it is critical to select competent project managers (Lee, Park and Lee, 2013). In their study, Galvin et al. (2014) used the foundations of the competency theory to base their research upon and mentioned that this theory is extremely important in investigating the role of perceived leadership competency. They reviewed literature on leadership style and leadership competencies and from the three different leadership styles that it studied, fifteen competencies were outlined. The research by Lee, Park, and Lee (2013) on leadership competency concluded that leadership competencies of a project manager are critical for project success. They mentioned that as projects in IT services involve multiple stakeholders with different backgrounds and a variety of expertise working in a cross-disciplinary manner; it becomes critical to select competent IT managers. For this research, perceived leadership competency is an important construct as after OPM's Big Data Security breach and the shortly thereafter resignation by the director of OPM, Katherine Archuleta (Castelluccio, 2015), it is believed that there might be some type of relationship that might exist between perceived leadership competency and an individual's trust of the online system. Thus, it is proposed in this study that lack of perceived leadership competency and trust in leaders managing big data security inhibits the ability of individuals to trust these individuals, which raises trust issues of online systems, and consequently becomes detrimental to an agencies' ability to hire.

There has been significant research done in the field of leadership styles including transactional, transformational or passive-avoidance leadership in an organization as mentioned by Analoui, Clair, and Sambrook (2013). Additionally, as mentioned by Galvin et al. (2014), leaders lead organizations with varying types of leadership styles and as this study focuses on the effects of perceived leadership competency on trust, it is important to highlight that even though

leadership styles are important to precede perceived leadership competency, it is the intention of this research to not focus on leadership style but rather highlight perceived leadership competency in order to limit the scope of this study. In this study, it is believed that to understand an individual's trust of the online system, it is important to understand the antecedents of trust which is perceived leadership competency.

Trust: Definitions of trust are copious, although a universally accepted definition is still nonexistent (Barber, 1983; Das & Teng, 2004). As trust is a construct for this study, the study by Beldad et al. (2012) on trust is important for this research as they researched trust factors that individuals encounter when disclosing their personal information online and concluded that that high levels of trust and low perceptions of risk could propel the performance of a behavior specifically when sharing pieces of personal information for online transactions. Results of this study clearly indicated that regardless of whether citizens have any experience with online government transactions, their levels of trust in government organizations are instrumental in influencing their willingness to supply personal data for online government services. The risks involved in the sharing of personal data, such as selling or sharing those data to third parties, may propel internet users to critically assess the trustworthiness of a particular government agency.

Additional studies in the relationship of trust and leadership include one such study by Wang and Hsieh (2013), where they not only studied this relationship, but also looked into the turbulent and multitude of problems facing today's leaders. They mentioned that in today's global environment, organizations face rapid changes, enterprise experience ethical meltdowns as well as a multitude of challenging and turbulent problems. Sustainability of the organization is at stake due to the employees' lack of trust on leadership. Their study examined the effect of

leadership on employee engagement through employee trust. Additionally, they concluded that supervisors' authenticity is positively related to employee trust, and within authentic leadership it is a supervisor's consistency between words and actions that has the strongest influence. Also, trust was shown to be positively related to employee engagement in the organization. As mentioned by Covey and Merrill (2006), lack of trust in supervisors and the organization has been found to influence a lack of engagement by employees in their work or in this study possibly on disclosing PII information online. For this research, an individual's trust of the online system might also be affected by the leadership competency as highlighted before and thus for this study thus, it was important to study the relationship of perceived leadership competency, trust of the online website and the effect these have on the willingness or likelihood of disclosing personal identifiable information online.

Perceived Big Data Security: It is perceived that individuals with privacy concern issues might also have concerns on the security of their personal data submitted. As mentioned by Chen, Chiang, and Storey (2012), big data not only brings about challenges to information security, but also offers new opportunities for the development of cyber security mechanisms as well. As perceived big data and its security is a construct of this study, it is important for understand previous literature in this research area, its definition as well as its security vulnerability. In their study, Perera et al. (2015) mentioned that big data has no clear definition, but it isn't wholly about size. Rather, it's defined based on three primary characteristics, also known as the 3Vs: volume, variety, and velocity. Expanding on its definition, they explained that volume relates to the data's size, variety refers to the type of data and its source (sensors, devices, social networks, the Web, mobile phones, and so on) and velocity means how frequently the data is generated. These 3V's are what sets big data apart from any other regular data and have been detailed in

Table 2. As the data collected from millions of for the security clearance process satisfy the 3V's, the term Big Data and its Cyber Security has been included for this study.

There is an abundant amount of research done in the Big Data field including the works of Vera-Baquero, Colomo-Palacios, and Molloy (2013) where they looked into the study by Van der Aalst (2012), and highlighted the three types of business process analysis (BPA) - *validation*, *verification*, and *performance* - all of which require collecting and storing large volumes of process and event data. They also expanded on big data, its definition and its components including but not limited to Hadoop, Hive and HBase. Even though detailed information on big data as detailed in prior literatures are important for this study to highlight but as the focus of this study is on big data security and its vulnerability in relation to the willingness of individuals disclosing sensitive PII information, detailed explanation of its properties, its architecture, components and capabilities have not been emphasized in order to limit the scope of this study.

Table 2

Big Data and the 3V's

DEFINITION

<i>VOLUME</i>	Volume relates to the data's size or the amount of data that can be processed.
<i>VELOCITY</i>	Velocity refers to the speed of processing the data to meet the demands.
<i>VARIETY</i>	Variety refers to the type of data and its source. Some of the sources include data from sensors, devices, social networks, website, and mobile phones.

With the extremely fast development of data storage, networking and communicating capability along with the data collection capacity, big data is rapidly expanding in all science and engineering domains and not limited to only physical, biological and biomedical sciences. IDC proposes that if organizations are able to use big data solutions to the optimum in their business decisions, they will thrive and obtain a competitive advantage in the market (Villars, Eastwood,

& Olofson 2011). Additionally, another concept in the field of big data that has emerged recently is Internet of Things (IoT) which is defined as a network of networks in which a massive number of objects, sensors, or devices are connected through the ICT infrastructure to provide value-added services (Perera et al., 2015). The IoT connects people and things anytime, anyplace, with anything and anyone, ideally using any path or network and any service. The online system that is used by individuals to provide their sensitive information for security clearance process is also available in one of these easily accessible devices, thus generating several concerns over the security of this online system. One of the major concerns being that even though the device might be secure, still where it is accessed from and how wireless security could affect the ability of hackers to not only extract an individual's data but to also gain access to the online system where this sensitive big data resides. Adding to this dilemma of big data security is also the belief by Perera et al. (2015) that by 2020, 50 to 100 billion devices was connected to the Internet generating *big data* for analysis and knowledge extraction. For this research, this was an important concept to understand as individuals might be able to access and provide their sensitive information through these mobile devices that might not be secure, or their internet access might be vulnerable to security.

Chang, Hsu, and Wu (2015) mentioned that in recent years, big data has become a popular issue in the realm of IT. International Data Corporation (IDC) reported that 1773 zettabytes of information were created and replicated in 2011 and forecasted that worldwide information will grow to 7910 Exabyte in 2015. At the same time, the amount of information is growing quickly in enterprises (Gantz, Chute, Manfrediz, Minto, Reinsel, Schlichting, & Tocheva, 2008). Given the vast expansion of big data so quickly, its safety has drawn great attention of researchers. However, as per Chang et al. (2015), there is only limited research on the representation of

multi-source heterogeneous big data, measurement and semantic comprehension methods and that the future research in the field of big data security, including credibility, backup and recovery, completeness maintenance, and security should be further investigated. In another study by Wu, Zhu, Wu, and Ding (2014), they mentioned that information sharing is an ultimate goal for all systems involving multiple parties (Howe, Constanzo, Fey, Gojobori, & Hannick, 2008). While the motivation for sharing is clear, a real-world concern is that big data applications are related to sensitive information, such as banking transactions, security information and medical records. As this study's focus was on Big Data's Cyber Security, prior studies were referenced to measure the effect of Perceived Big Data Security construct on Trust of the Website.

Reward of Job: Finally, one of the aims of this research was to understand the role that “*Reward of a Job*” might have on an individual where the benefits of getting a job with the government in some instances might outweigh the concerns that these individuals might have, which in turn could lead to them overlooking the privacy and trust concerns. In this study, reward is being equated to benefit, where a substantial number of studies have researched benefits of disclosure. One such foundational study by Beldad et al. (2012) mentioned that it can be assumed that even with high perceptions of privacy risks and without trust, citizens would still opt to share personal data to avail a particular government service online if benefits can be expected. Viewed from a calculus-based perspective (Laufer, & Wolfe, 1977), people's decision to share personal data online, despite the risks of having their online privacy compromised, is driven by a belief that the benefits of data disclosure outweigh the estimated costs of the disclosure act (Culnan & Bies, 2003; Olivero & Lunt, 2004). Rewards in the form of monetary vouchers, for instance, have a positive impact on Internet users' willingness to provide accurate personal information (Xie,

Teo, & Wan, 2006). Similarly, it is assumed based on these prior benefits study there might be situations where benefits (obtaining a job with monetary gain) might outweigh trust of the online system. This variable could affect the likelihood and willingness of individuals providing sensitive information online; hence it was considered as one of the constructs for this research.

Past Literature Gaps

There have been numerous studies that have researched the effects of trust, risk, privacy concern and information sensitivity on willingness disclosing personal information online for the healthcare industry but there have not been any reported studies found that have also researched the effects of perceived big data security, perceived leadership, and job reward on the dependent variable. For example, some studies that researched causality and construct relationship particularly for health care, did not focus on the presence of vast amount of transactional data that was being disclosed by individuals online. Thus, the effect that data and its security might have on the dependent variable did not get researched.

Leadership Competency is another construct that has not been included in similar studies but for this research was extremely important because when OPM's website was hacked and sensitive data stolen for millions of individuals, shortly thereafter, the head in charge of OPM was forced to resign. Hence, leadership might have some relationship and causality in this study which was not researched in prior studies. As with prior studies, especially in the healthcare research, a similar construct to job reward would have been health advice reward and it would have been interesting to study how individuals would have disclosed their health information online while given the risk, privacy concern and trust of the online system.

Even though, these prior studies did not include all of the constructs mentioned for this study, it is a common practice for researchers to limit the number of constructs in their studies in

order to limit the scope of the research and most of these studies have clearly indicated it as such in their research.

Analysis of the Research Methods Used

There have been a wide variety of research methods and design utilized by prior studies. Some of the studies have been quantitative where they have used survey research, decision trees, correlational, causal-effect and experimental designs while others have been qualitative where the researchers have developed theories, used narratives, case studies and conducted interviews. The three major foundational studies that this research has been based on are the studies by Bansal et al. (2010), Beldad et al. (2012) and Osatuyi (2015) where these studies have used quantitative survey based research method. Most of these studies also used existing valid and reliable survey instruments and did not develop any new instrument for their research. Some studies have been based on prior studies that have either developed new theories or models such as the one by Osatuyi (2015) where his study was based on the foundational work by McCrae and Costa (1987), where they developed a Big Five Model to consolidate important traits that were found to be reliable across domains. To measure competency, some studies on leadership and competency used the works by Dulewicz and Higgs (2003) and Quinn and Rohrbaugh (1981) instrument for self-evaluation especially for the study by Shang and Wu (2013) on managerial competencies. For the foundational study by Beldad et al. (2012), they mentioned that they used structural equation modeling, a comprehensive statistical approach to test hypotheses about relations among observed and latent variables (Hoyle, 1995) and using AMOS 18.0, they were able to perform confirmatory factor analysis on the constructs of the study, to address the research hypotheses, and to test whether the research model fits the data. They mentioned that for their constructs, however, the items that they used, to measure ‘beliefs in the adequacy of legal

protection for on line transactions' were patterned after the statements used to measure 'legal framework' as a determinant of online trust in two studies (Cheung & Lee, 2006; Connolly & Bannister, 2007). Furthermore, they mentioned that these statements, however, were substantially reformulated to suit the context of their study and further detailed information was not provided. Similarly, no further detail was provided for the newly formulated constructs for 'behavioral intention to disclose personal data' that they developed for the study.

Majority of these studies also expanded on data analysis including use of descriptive statistics including mean, median and mode and inferential statistics including but not limited to t tests, regression analysis, good-fit, *p* and Cronbach Alpha and have done a very good job in outlining the results of the analysis in tables, charts and appendixes. Data captured for these studies also ranged from a couple of hundreds for the majority of these studies to a maximum of 3500 for the study in e-government disclosure by Beldad et al. (2012). As a result of the smaller sample size, most of the quantitative studies used PLS-SEM to analyze data such as in the study by Osatuyi (2015) where Smart PLS 2.0 was used.

Additionally, one common trend on most of the studies referenced for this research has been on limited research generalizability. Most if not all have indicated that their research has been based on local areas including but not limited to United States, South Korea, India, South Africa, Taiwan, Dubai and China and thus do not apply to the general population. None of these studies were longitudinal but rather cross-sectional which means that they were conducted one time only, instead of different period of times.

Literature Synthesis

The three major foundational studies that this research would be based on are the studies by Bansal et al. (2010), Beldad et al. (2012) and Osatuyi (2015). Bansal et al. (2010) looked into the

process by which personal dispositions impact individuals' trust and behavior intention to disclose personal information online for the healthcare industry whereas the research by Osatuyi (2015) focused on how personality traits affect information privacy concern for an individual. As the research by Bansal et al. (2010) targeted the health care industry, the research by Osatuyi (2015) was focused on online merchants where the big five personality traits and how these traits affect an individuals' privacy concern in order to provide personal information were researched. Additionally, the study by Beldad et al. (2012) researched the factors affecting individuals disclosing personal information for e-government including but not limited to trust, risk, adequacy of legal protection, prior online experience as well as the benefits of disclosing information and its implications.

This research not only combines both of these studies but also extends them by adding new constructs of perceived leadership competency, perceived big data security and job reward for the security cleared job market. As there've been numerous research done for online healthcare information disclosure field, there have not been any reported and published study found that have focused on the intention of providing sensitive information online in order to attain a secure and cleared US Government job. Even though individuals applying online might be aware of the impact that privacy concern, trust and prior hacking experience might have had in the past, it is proposed that the reward of getting a stable government job might outweigh these concerns.

Summary

This literature review chapter highlights literature from prior studies in the same field, limitations, assumptions and their research scope as well. Three major foundational studies that this research is based on are the studies by Bansal et al. (2010), Beldad et al. (2012) and Osatuyi (2015). Bansal et al. (2010) looked into the process by which personal dispositions including risk

and privacy concern impact individuals' trust and behavior intention to disclose personal information online for the healthcare industry. The research by Osatuyi (2015) focused on how personality traits affect information privacy concern for an individual and the study by Beldad et al. (2012) researched the factors affecting individuals disclosing personal information for e-government. Utility theory has been the basis of this research and details of the theory as well as the justification to base this research of, has been detailed in this chapter.

Additionally, this chapter details the constructs of the study including, perceived big data security, perceived leadership competency, privacy concern, trust and job reward and highlights where these constructs have been used and the research design and methods implemented in prior similar studies. Finally, there has also been additional emphasis on the shortcomings and gaps from prior research and the reasoning behind the perceived limitations as well and how this study will to add to the existing body of knowledge and how this study will help the reader gain a new perspective.

Chapter 3

Methodology

Research Method

A quantitative survey study was best suited for this research to analyze the effect of the independent variables – perceived big data security, trust, perceived leadership competency, information sensitivity, privacy concern and job reward on the dependent variable - willingness to disclose secure information online. Research Design included collection of data through random sampling, data analysis using descriptive and inferential statistics, interpretation of the data analyzed and support for the research question and proposed hypotheses. Also, as per Terrell (2016), for the reliability and validity of the research method, it is important to research if the results of the study were caused by an intervention and if they are generalizable to different locations or population. As this study is a survey based study, use of prior validated and reliable instruments was instrumental for this study.

The idea of causality, or the relation between cause and event, is central to many conceptions of theory. When theory is taken to involve explanation, it is intimately linked to ideas of causation (Gregor, 2006). As mentioned by Kim (1999), four prominent approaches to the analysis of event causation are regularity (nomological), counterfactual, probabilistic and manipulation (teleological) causal analysis. As this study is based on survey research, validation becomes very important as validation gives researchers, their peers, and society as a whole a high degree of confidence that positivist methods being selected are useful in the quest for scientific truth (Nunnally, 1978).

For this research, random sampling was used to collect survey data. Random sampling is generally used in quantitative studies where it becomes important to identify a sample that represents, as closely as possible, the population it was selected from (Terrell, 2016). It is important for the sample to be as representative of the characteristics in study as possible which is called the generalizability of the sample to the population. If the sample isn't generalizable, then the results based on the sample are likely not valid, and will not reflect the true values in the population which is called the sampling bias (Terrell, 2016). As the surveys, was online based only, there was some sampling bias as paper based surveys was not used.

In this study, the unit of analysis was the individual employees and the study was not longitudinal, where data is collected over a long period during different times, but rather was collected one time only using the cross-sectional method. After data collection and analysis, the proposed hypothesis was concluded to be either supported or not.

Descriptive statistics was used to provide description of the sample including mean, median mode as well as the standard deviation for the demographic data collected including age, gender, education level (highest level of educational degree completed), organizational position and years with the government. Research design included the use of inferential statistics tools such as t tests, ANOVAs and regression analysis that allowed the researcher to make decisions about the data collected and the hypotheses to be supported or rejected.

Instrument Development

As mentioned by DeVellis (2011) and Fink (2003), the key to selecting an appropriate instrument for a study is the type of data called for in the research questions and the hypotheses. As this research, drew from prior studies conducted in this field, it utilized vetted survey instrument as well as questions from those studies in order to not make sure that the constructs

and the study was valid and reliable. Overall, there were several survey instruments used in similar past studies that was extremely important for this research including Multifactor Leadership Questionnaire 5-S (MLQ) which is a survey to measure Transformational leadership, Organizational Leadership Questionnaire (OLQ) and International Personality Item Pool which is used to measure personality traits. To measure for competency, a survey tool designed by Quinn and Rohrbaugh (1981) was used to measure leadership competency. Similarly, to measure privacy concern as well, survey instruments was used from prior studies by Awad and Krishnan (2006).

Table 3

Constructs and Instrument Source

Construct	Definition	Source
<i>Willingness to Disclose</i>	Willingness to disclose sensitive personal information	(Chang et al., 2015) and (Shin, 2010)
<i>Organization</i>		
<i>Perceived Big Data Security</i>	Security of sensitive data collected	(Shin, 2010)
<i>Perceived Leadership Competency</i>	Competency of the Leadership	(Dulewicz & Higgs, 2005) and (Müller & Turner, 2010)
<i>Individual</i>		
<i>Information Sensitivity</i>	Personality traits of Extroversion, Intellect and Conscientiousness	(Osatuyi, 2015)
<i>Trust</i>	Trust in the sensitive information collecting website	(Hsu et al., 2014)
<i>Privacy</i>	Privacy concern of sensitive information	(Malhotra & Agrawal, 2004)
<i>Job Reward</i>	Reward or benefit of getting a job	(Chang et al., 2015)

Prior to testing the research hypotheses, construct reliability was determined by calculating the constructs' Cronbach's alpha scores. Alpha values above .70 indicate acceptable reliability (Hinton, 2008). Additionally, the reliability levels of the constructs for this study meeting the

proposed criterion was presented in a Table format as well. Additionally, Table 3 highlights a list of the constructs and their instrument source.

To accommodate for responses from individuals, a questionnaire, including previously vetted questions, was created. A 7-point Likert scale was designed to collect responses to questions, which varied from Strongly Agree to Strongly Disagree. The online survey was designed using Google Forms and later imported into Survey Monkey for data collection. To test for the reliability of the survey, a pilot test was performed including 15 participants. From the results of the pilot study, the survey was fixed and finalized, specifically when missing data for some of the questions was found, all of the questions were marked as required which fixed the issue of missing data. Some of these tools, research methods and techniques have been widely used in similar studies focusing on personal disposition, big data, trust and privacy and thus the above-mentioned research methods from similar past studies were also foundational for this study.

Validity and Reliability

Validity of a study is defined as the trustworthiness of the study's results (Gay, Mills & Airasian, 2012; Newton & Shaw, 2014). As per Bhattacharjee (2012), the quality of research designs can be defined in terms of four key design attributes: Internal, External, Construct and Statistical Conclusion Validity. Straub et al. (2004) mentioned that construct validity differs from internal validity in that it focuses on the measurement of individual constructs while internal validity focuses on alternative explanations of the strength of links between constructs. Internal Validity which is also called causality, examines whether the observed change in a dependent variable is indeed caused by a corresponding change in hypothesized independent variable, and not by variables extraneous to the research context (Bhattacharjee, 2012). As this study is a survey

based study, internal validity is poor for this research as they are not able to manipulate the independent variable (cause), and because cause and effect are measured at the same point in time which defeats temporal precedence making it equally likely that the expected effect might have influenced the expected cause rather than the reverse. Differentiating between external and construct validity, Bhattacharjee (2012) added furthermore, that external validity which is also known as generalizability refers to whether the observed associations can be generalized from the sample to the population (population validity), or to other people, organizations, contexts, or time (ecological validity). Whereas, he mentioned that construct validity defines how good a given measurement scale is measuring the theoretical construct that it is expected to measure. As per Terrell (2016), construct validity means investigating the degree to which an instrument measures what it claims to measure. For this research, to ensure construct validity, it was important to use items from existing scales whenever possible and to minimize the common method bias, where it was important to convert the items to semantic differential (0-10).

Lastly, Bhattacharjee (2012) mentioned that statistical conclusion validity examines the extent to which conclusions derived using a statistical procedure is valid. This means that whether the right statistical method was used for hypotheses testing, whether the variables used meet the assumptions of that statistical test such as the sample size or distributional requirements and so forth. As, it was extremely important to select the correct statistical method for the hypotheses testing, thus, it is equally important to look at previous such studies and review their selection process when selecting the appropriate method.

For this study to have reliable and valid constructs, it was important to use constructs that have already been used and validated for reliability in prior studies. Testing in studies need to have reliable instruments, which means that it should consistently measure what it's intended to

measure (Terrell, 2012). Additionally, to confirm reliability of the instrument for testing, it is important to compute a reliability coefficient, which in this study was the Cronbach's alpha, where the values of these coefficient range from zero (low reliability) to 1.0 (high reliability). In essence for a study, the higher the coefficient number the more reliable the test. As per Terrell (2012), the coefficient is used to look at the four types of instrument reliability; test-retest, equivalent forms, interrater, and split-half.

Finally, as positivist methods for data collection employ a deductive approach to research, starting with a theory and testing theoretical postulates using empirical data, similarly, this research was based on the foundations of the utility theory and the data collected used survey responses from individuals.

Sample

Sample collection, data analysis and reporting on the results are some of the important aspects of a survey study. Therefore, survey links were sent out randomly to various membership groups including university, residential and social networking groups that the researcher, friends and family members of the researcher has been in touch with in the past. Survey links were also sent to UMUC University students where the researcher teaches as well as places of prior and present employment. Survey Monkey was also used to collect random data through the participant outreach program. As random sampling, was used to collect responses to the survey, it was the intent of this research to collect data, where all possible subsets of a population or the sampling frame were given an equal probability of being selected. As the sampling frame is not subdivided or partitioned, the sample is unbiased, and the inferences are most generalizable amongst all probability sampling techniques. Approximately there were 700-800 emails sent and

a total of 206 responses received, including 80 responses that were collected through survey monkey's audience outreach program.

Data Analysis

Data analysis was performed on the data collected to inspect, clean, transform and model the data to extract useful information. Descriptive statistics was used to report mean, median, mode and standard deviation on the demographic data collected. Some of the descriptive statistics data collected included age, gender, education level, years of computer usage and prior experience with government websites usage. Pre-screening of the data was accomplished by using Mahalanobis Distance that was used to capture outliers. As mentioned by Mertler and Vannatta (2013), capturing outliers is extremely important when analyzing data where the Mahalanobis distance is calculated based on distance from the centroid (mean of all variables). For his study, Osatuyi (2015), used Partial Least Square Structural Equation Modeling (PLS-SEM) for data analysis instead of Covariance based Structural Equation Modeling (SEM) and cited the following reasons; 1) its more regression-based approach that minimizes the residual variances of the endogenous constructs, 2) it is more robust with fewer identifications issues, 3) it works well with much smaller as well as larger sample sizes, and 4) it readily incorporates formative as well as reflective constructs (Hair, Ringle, & Sarstedt, 2011). Similarly, for this study PLS-SEM was used for data analysis, specifically Smart PLS 3.0 was used as the tool for the reasons mentioned above. Additionally, it was hoped that the data collected for this study would be from at least 200 respondents, even though at a minimum 700-800 surveys were sent over to individuals for their feedback.

Additionally, various techniques were used including but not limited to factor analysis to make sure that the data collected was valid and reliable. Normality, linearity, Variance (AVE)

and reliability along with Cronbach Alpha which is used to determine the constructs internal consistency were also be measured. Finally, once the data was analyzed, data visualization techniques were used to help clearly and efficiently communicate the analysis of the data including scatter plots that are generated to indicate non-normal shapes to indicate normality and linearity and scree plots that are used to graphically represent the data for variance (Mertler & Vannatta, 2013).

Once the data collected was analyzed, the proposed hypothesis was justified to be either supported or not and results and conclusions of the research presented in chapter 4 of this study. Additionally, it was the intention of this research to provide conclusive feedback that would be beneficial to the US Government to consider when attracting individuals applying for jobs where US Government required a security clearance.

Results Format

All results from the data analysis was included in the dissertation report. Once the data was analyzed, results were presented in various formats including tables, bar charts to show the number of occurrences of a value in the data, figures as well as screen shots outputted from SPSS and Smart PLS instruments used to analyze the data. Most of the screen shots were included as Appendix towards the end of the study including but not limited to descriptive statistics, outliers, bootstrapping results, loading, variances, convergent and discriminant validity results as well as scatter plots. As the survey forms, was designed in Survey Monkey, the questionnaire form and screenshots of data collected were also presented in the Appendix section of this study.

Descriptive statistics including data on age, gender, profession and years of relevant experience with the government was presented in the Appendix section while the details of the data analysis including the mean, median and mode was presented in table format in the results

section. Some of the data from inferential statistics was presented in the Appendix section while some was highlighted in tables such as p and Cronbach Alpha which is used to support or reject the proposed hypotheses.

Resource Requirements

Resources for this study were not hard to arrange for as hardware such as computer and printer was easily accessible, and the software needed for data collection and analysis such as IBM's SPSS and R was available through the university. Surveys were initially created in google forms and then transferred into survey monkey tool which was also easily accessible. Smart PLS 3.0 was used for further data analysis through registering on their product website.

As this research drew from prior studies conducted in this field, this study utilized vetted survey instruments from those studies. Detailed information on the vetted survey instruments from prior studies has been highlighted in the instrument section of this chapter that includes but is not limited to Multifactor Leadership Questionnaire 5-S (MLQ), Organizational Leadership Questionnaire (OLQ), International Personality Item Pool that are ideal to measure personality traits as well as a leadership competency survey tool designed by Quinn and Rohrbaugh (1981).

To accommodate for responses from random individuals, questionnaire was created in survey monkey to collect the data. A 7-point Likert scale was used to collect responses to questions, where 7-Strongly Agree to 1-Strongly Disagree. A pilot test was performed including 15 participants. From the results of the pilot study, the survey was fixed and finalized after all of the questions were marked as required to accommodate for missing data. Some of these tools, research methods and techniques have been widely used in similar studies focusing on personal disposition, big data, trust and privacy and thus the above-mentioned research methods from similar past studies were foundational for this study.

Finally, one of the most important aspects of this research study was to collect data from particular target audience above the age of 18. As this study is based on security cleared individuals, it was ideal to also target individuals that have either gone through the clearance vetting process or would have been going through it soon in their career. Additionally, it was clearly highlighted that respondents use their own personal computer and respond on their own personal time instead of using government resources or time.

Summary

This Research Method chapter detailed the research design employed for this study and highlighted the survey instrument, data collection techniques, sample data used as well as the validity and reliability of the instruments to be used for this research. A quantitative survey study was considered best suited for this research to analyze the effect of the independent variables – perceived big data security, trust, perceived leadership competency, information sensitivity, privacy concern and job reward on the dependent variable - willingness to disclose secure information online. Some examples of the survey instruments used in similar prior studies were highlighted that helped this study to be extended further. Research Design included collection of data through random sampling and data analysis included the use of both descriptive (mean, median, and mode) and inferential statistics (*t* test, ANOVA and regression analysis). Descriptive Statistics included demographic data on age, gender, years of computer use, academic level achieved and prior experience with government websites. A pilot test involving 15 participants was performed and initially the surveys were created using Google Forms then transferred into Survey Monkey. Threat to validity and reliability of this study was detailed to understand if the results were caused by an intervention as well as to understand the generalizability of this study to different locations and populations.

Additionally, details on the proposed sample using random sample used for this study and the presentation of the results format was also detailed in this chapter. Finally, details on the interpretation of the data analyzed was highlighted where requirements involving either support or rejection for the hypotheses was detailed.

Chapter 4

Results

Pre-Analysis Data Screening

This study was a quantitative study that collected data through online survey (survey monkey) using a 7-Point Likert Scale (see Appendix A). A pilot test was performed including 15 participants to test the reliability of the online survey prior to data collection. The participants were friends, coworkers and neighbors of the researcher. During the pilot testing of the survey, some of the questions were missing data. As some of the questions were not marked required, this issue was corrected by making all of them required. Additionally, a final check through SPSS's frequency method was done to make sure that there were no further missing values. The final survey link was sent to coworkers, friends, relatives, previous employers, current managers of the researcher as well as University of Maryland University College students in the Information Systems Masters Level Program after obtaining IRB approval (see Appendix B).

Cross-sectional method was used to collect the data in the month of October 2017. This method is used to collect data only once, instead of at different intervals as per the longitudinal approach. Unit of analysis for this study was the individual. Random Sampling was used to collect the data where links to the survey were sent through email to approximately 700-800 individuals. A total of 206 responses were received, including 80 responses that were collected through survey monkey's audience outreach program (see Appendix C).

IBM's SPSS was used to analyze the outliers, normality, chart scatter plots, box plot and Q-Q Plot. Additionally, descriptive statistics was used to analyze the demographic data collected on

age, gender, academic level, years of computer use and if the individual has used the government website before which provided description of the sample including mean, median mode and standard deviation. Partial Least Square Structural Equation Modeling (PLS-SEM) was used for data analysis instead of Covariance based Structural Equation Modeling (SEM) due to the following reasons as cited by Osatuyi (2015), 1) its more regression-based approach that minimizes the residual variances of the endogenous constructs, 2) it is more robust with fewer identifications issues, 3) it works well with much smaller as well as larger sample sizes, and 4) it readily incorporates formative as well as reflective constructs (Hair, Ringle, & Sarstedt, 2011).

Mahalanobis Distance and Box Plot

Through SPSS analysis, a total of 16 outliers were found, specifically the values for the cases of 6, 72, 12, 123 and 29 were found to be above 82.60 (see Appendix D Mahalanobis Distance). Mahalanobis Distance was calculated from the critical value of chi-square at $p < .001$ with $df = 34$ which showed to the result of 59.773. The accepted criterion for outliers is a value for Mahalanobis distance that is significant beyond $p < .001$, determined by comparing the obtained value for Mahalanobis distance to the chi-square critical value (Mertler & Vannatta, 2013). Five of these extreme values were deleted and the Mahalanobis distance was run again (see Appendix E Rerun Mahalanobis Distance).

As per Mertler and Vannatta (2013), it is not always appropriate to drop the cases from analysis as there might be cases that might be interesting instead of being just simply bad. Therefore, out of 16 extreme cases, 5 were dropped that were identified as having the highest extreme values. After analyzing the data again with 201 cases, it was found that there were now 10 outliers. The following cases were shown to have extreme values – 127, 140, 68, 136, 125.

Mahalanobis Distance was calculated from the critical value of chi-square at $p < .001$ where $df=8$ showed to be at 26.125.

Normality and Scatter Plot

To test for the normality, all of the variables were aggregated into independent and dependent variables. Prior to deleting the 5 most extreme values shown through the box plot, the Skewness was at 1.524 and the Kurtosis at 2.736 (see Appendix F). Once these 5 outliers were deleted, there was a significant drop in the Skewness and Kurtosis, specifically, 1.116 for Skewness and .823 for Kurtosis, which only showed a slight peak in distribution as it was above the accepted value of 1.0 for Skewness. The accepted range is in between -1 to +1 (Hair et al., 2017). As a result of the deletion of the 5 extreme values, analysis showed that the data was normally distributed. Data visualization techniques are used to help clearly and efficiently communicate the analysis of the data including scatter plots that are generated to indicate non-normal shapes to indicate normality and linearity and scree plots that are used to graphically represent the data for variance (Mertler & Vannatta, 2013). In this study, the normality graph shows that the cases were very close to the diagonal line, values near the diagonal line are indicative of normality (see Appendix E Rerun Mahalanobis Distance). Additionally, through the scatter plot it was seen that the values were representative of a rectangular shape which is indicative of normal distribution (Mertler & Vannatta, 2013). Thus, it is understood that the data for this study is normally distributed.

Descriptive Statistics

Descriptive statistics was run on the data collected using SPSS through the frequency function in order to measure the mean, median, mode and standard deviation. Bar charts were presented to understand the ratio of male to female respondents, users with different academic

background, years of computer use and prior experience with government websites (see Appendix G). After deleting the 5 extreme outliers from the data collected (N=201), the number of male respondents were at 52.2 percent compared to female at 47.8%. Majority of respondents were between the ages of 35-50 at 40.8%, and, as required, there were no respondents under the age of 18. Highest academic level for the majority of respondents was at 33.8% for Bachelors followed by Masters at 28.9%. Majority of the respondents had more than 15 years of computer use experience at 77.1% and 74.6% had prior government website experience (see Appendix G).

Data Analysis

Smart PLS 3.0 software was used to analyze the data further including but not limited to model fit, convergent validity, factor loading, construct reliability and validity, and discriminant validity (see Appendix H). After running the PLS algorithm, it was noticed that the factor loadings were in the acceptable range, except for the latent variable PBDS, where PBDS1 was at -0.848 and PBDS2 was at -0.739. Model Fit's SRMR was at 0.093 which is above the accepted value of 0.08 (Hair et al., 2017) (see Appendix I).

As mentioned by Hair et. Al (2017), the SRMR is defined as the root mean square discrepancy between the observed correlations and the model implied correlations. Furthermore, as the SRMR is an absolute measure of fit, a value of zero indicates perfect fit. When applying CB-SEM, a value less than 0.08 is generally considered a good fit (Hu & Bentler, 1998). In this study the construct reliability for PBDS was at 0.008, which is not considered reliable as well.

After deleting PBDS1 and PBDS2 and rerunning PLS algorithm, PBDS has only one measure PBDS3. Therefore, the model fit's SRMR became 0.0773 which is under the 0.08 recommended value resulting in good fit (Hair et al., 2017) (see Appendix J and K). For a research model to be considered fit, Table 4 shows the values for saturated and estimated values.

Furthermore, deleting PBDS1 and PBDS2, the Cronbach Alpha for PBDS went from 0.008 to 1.0, which also is in the acceptable range (see Appendix K). As this study is based on survey research, validation becomes very important as validation gives researchers, their peers, and society as a whole a high degree of confidence that positivist methods being selected are useful in the quest for scientific truth (Nunnally, 1978). For this study, the validity and reliability of the model was rigorously tested to make sure that the model was both valid and reliable.

Table 4

Model Fit and Accepted Values

	Saturated Model	Estimated Model
SRMR	0.077	0.097
d_ ULS	3.149	5.017
d_ G1	2.125	2.207
d_ G2	1.313	1.392
Chi-Square	1,466.847	1,539.416
NFI	0.722	0.709

Convergent Validity

As per Lee, Park and Lee (2013), convergent validity refers to the degree to which a measure is correlated with other measures to which it is theoretically predicted to correlate. This implies that the measurement variables of each potential construct should be loaded with significant *t*-values. For this study, PLS Factor analysis was run to visualize the mean, median, loading, standard deviation, kurtosis and skewness of the variables.

Chin (1998), if the factor loading between the measurement item and the variable is 0.7 or more, the item is considered valid. In our study as shown in table 5 all relevant loadings were above 0.7 after deleting the values for PBDS1 (-0.848) and PBDS2 (-0.739). Prior to testing the research hypotheses, construct reliability was determined by calculating the constructs' Cronbach's alpha scores.

Table 5

PLS Factor Analysis

	Missing	Loading	Mean	Median	SD	Excess Kurtosis	Skewness
PBDS1	0.000	-0.848	4.353	5.000	1.806	-1.054	-0.150
PBDS2	0.000	-0.739	4.498	5.000	1.823	-0.845	-0.519
PBDS3	0.000	1.000	4.398	5.000	1.555	-0.386	-0.532
TW1	0.000	0.826	4.423	5.000	1.641	-0.931	-0.427
TW2	0.000	0.885	4.567	5.000	1.589	-0.684	-0.585
TW3	0.000	0.844	4.647	5.000	1.421	0.243	-0.552
TW4	0.000	0.902	4.612	5.000	1.558	-0.129	-0.743
PC1	0.000	0.762	5.189	6.000	1.802	-0.504	-0.851
PC2	0.000	0.812	4.502	5.000	1.533	-0.882	-0.259
PC3	0.000	0.787	4.716	5.000	1.703	-1.181	-0.250
PC4	0.000	0.609	5.448	6.000	1.400	-0.007	-0.814
JR1	0.000	0.842	3.512	4.000	1.400	-0.453	-0.091
JR2	0.000	0.944	3.537	4.000	1.624	-0.556	0.179
JR3	0.000	0.936	3.483	4.000	1.593	-0.828	0.129
JR4	0.000	0.893	3.328	4.000	1.587	-0.242	0.411
WDIO1	0.000	0.918	3.806	4.000	1.635	-1.055	-0.173
WDIO2	0.000	0.894	3.980	4.000	1.733	-1.092	-0.160
WDIO3	0.000	0.854	3.657	4.000	1.617	-1.108	-0.035
WDIO4	0.000	0.806	4.169	4.000	1.584	-0.862	-0.342
PLCI1	0.000	0.919	4.219	4.000	1.631	-0.628	-0.454
PLCI2	0.000	0.928	4.174	4.000	1.604	-0.535	-0.491
PLCI3	0.000	0.958	4.453	4.000	1.599	-0.150	-0.473
PLCI4	0.000	0.870	4.617	5.000	1.551	0.086	-0.672
PLCI5	0.000	0.923	4.403	4.000	1.687	-0.424	-0.526
ISE1	0.000	0.892	5.045	5.000	1.408	-0.363	-0.662
ISE2	0.000	0.934	4.846	5.000	1.293	-0.825	-0.446
ISE3	0.000	0.813	4.592	5.000	1.387	-0.508	-0.457
ISE4	0.000	0.557	4.378	4.000	1.475	-0.549	-0.233
ISC1	0.000	0.929	5.682	6.000	1.356	-0.195	-0.950
ISC2	0.000	0.303	5.736	6.000	1.113	0.499	-0.881
ISC3	0.000	0.915	5.075	5.000	1.184	-0.302	-0.526
ISI1	0.000	0.796	5.333	5.000	0.969	-0.203	-0.348
ISI2	0.000	0.902	5.259	5.000	1.089	0.184	-0.646
ISI3	0.000	0.802	5.388	6.000	1.213	0.242	-0.646

Table 6, shows the results of the construct reliability and validity which were shown to be acceptable, at a value of above 0.7 (Hair et al., 2017). As mentioned by Hinton (2008), Alpha

values above .70 indicate acceptable reliability. The internal consistency of a measurement item is shown by the values of average variance extracted (AVE), the composite reliability and Cronbach's Alpha. In other words, to evaluate convergent validity of reflective constructs, researchers consider the outer loading of the indicators and the average variance extracted (AVE) (Hair, Hult, Ringle & Sarstedt, 2017).

Table 6

Construct Reliability and Validity

	Cronbach's Alpha	rho_A	Composite Reliability	Average Variance Extracted (AVE)
<i>Conscientiousness</i>	0.727	0.734	0.792	0.597
<i>Extraversion</i>	0.830	0.946	0.882	0.660
<i>Intellect</i>	0.792	0.896	0.873	0.697
<i>Perceived Big Data Security</i>	-	-	-	-
<i>Perceived Leadership Competency</i>	0.954	0.962	0.965	0.847
<i>Privacy Concern</i>	0.739	0.758	0.833	0.558
<i>Reward of Job</i>	0.926	0.950	0.947	0.819
<i>Trust in Website</i>	0.888	0.904	0.922	0.748
<i>Willingness to Disclose</i>	0.892	0.903	0.925	0.755

Table 5 shows that Cronbach Alpha ranged from 0.727 – 1.0 and the composite reliability ranged from 0.792 – 1.0, both being above 0.7 which is the acceptable value. Also, AVE ranged from 0.558 – 1.0 which is also higher than the accepted value of 0.5 as mentioned by Hair et al. (2017). Therefore, these findings indicate that the measurement items in this study have convergent validity.

Discriminant Validity

As per Lee, Park and Lee (2013), discriminant validity refers to the low correlations that should exist between different measurements designed to measure different constructs. Additionally, they emphasized that the correlation coefficients of potential variables should show an appropriate pattern of factor loadings, and the measurement items should be highly loaded onto the allocated factors.

Table 7

Discriminant Validity

	Conscientiousness	Extraversion	Intellect	Perceived Big Data Security	Perceived Leadership Competency	Privacy Concern	Reward of Job	Trust in Website	Willingness to Disclose
<i>Conscientiousness</i>	0.773								
<i>Extraversion</i>	0.639	0.812							
<i>Intellect</i>	0.511	0.563	0.835						
<i>Perceived Big Data Security</i>	0.274	0.254	0.028	1.000					
<i>Perceived Leadership Competency</i>	0.111	0.199	0.273	-0.180	0.920				
<i>Privacy Concern</i>	0.369	0.308	0.251	0.215	0.155	0.747			
<i>Reward of Job</i>	-0.055	0.129	0.200	-0.018	0.277	0.290	0.905		
<i>Trust in Website</i>	0.115	0.194	0.274	-0.352	0.511	0.024	0.160	0.865	
<i>Willingness to Disclose</i>	0.010	0.184	0.229	-0.166	0.445	-0.005	0.370	0.485	0.869

As per Fornell and Larcker (1981), discriminant validity can be deemed adequate when the square roots of the AVE values are greater than the correlation coefficients between the variables. As per Hair et al. (2017), discriminant validity is the extent to which a construct is

truly distinct from other constructs by empirical standards, therefore, from Table 7, it is clear that the values in the diagonal direction are greater than other correlation coefficients. Therefore, for this study the measurement items have discriminant validity.

Findings

The hypotheses proposed in the study was tested using Smart PLS and the significance of all paths in the research model was tested using bootstrap procedure where re-sampling was selected to be at 500. Bootstrapping is used to show the significance levels in a structural model (see Appendix L). These independent constructs were found to show a variance in the dependent constructs with trust in website showing 26 percent explained by perceived leadership competency, privacy concern with 16 percent explained by information sensitivity and perceived big data security and 33 percent on willingness to disclose information online explained by trust in website, privacy concern and reward of job as shown in Figure 2.

From data analysis and as per Table 8, it is shown that willingness to disclose information online was influenced by trust of the website ($t=5.993$, $p=0.000$) and reward of a job ($t=5.178$, $p=0.000$). Perceived leadership competency was also shown to influence trust in website by the individual ($t=7.438$, $p=0.000$). From information sensitivity, only conscientiousness ($t=2.324$, $p=0.021$) showed to have significance on privacy concern. Perceived big data security initially showed significance to privacy concern when all the factors were loaded but once the two factors (PBDS1 and PBDS2) were deleted, the significance dropped as well, with $t=1.819$ and $p=0.070$. Surprisingly, privacy concern ($t=1.549$, $p=0.122$), was not shown to influence willingness to disclose information online. From PLS analysis, all of the data points are presented in Figure 2.

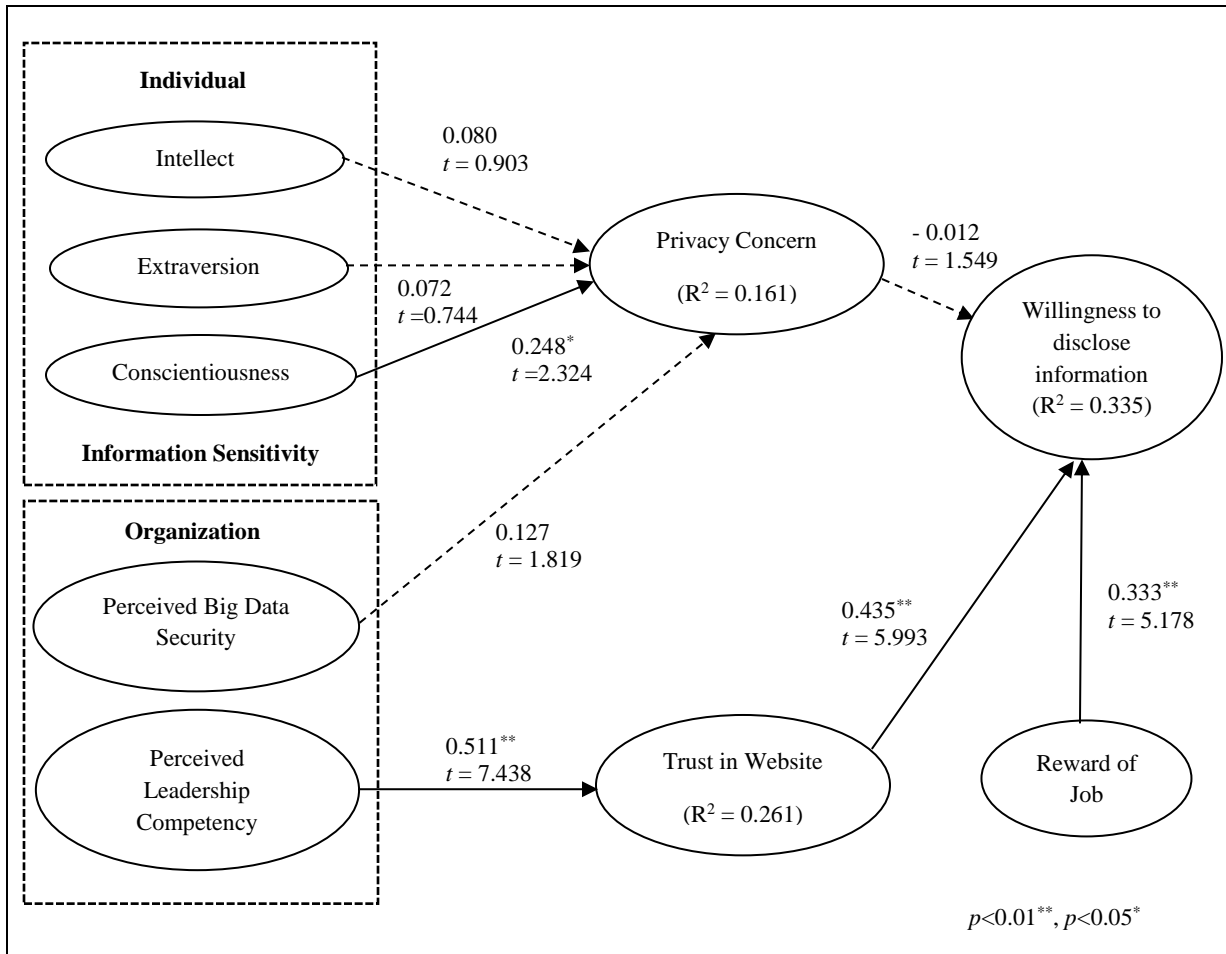


Figure 2. PLS Analysis Result for Willingness to Disclose

The path coefficients are the standardized beta coefficients. As expected Conscientiousness ($\beta=0.248$, $p<0.01$) displayed a significant and positive direct effect on privacy concern. Contrary to the hypotheses, Intellect ($\beta=0.080$, $p<0.01$), Extraversion ($\beta =0.072$, $p<0.01$) and perceived Big Data Security ($\beta=0.127$, $p<0.01$) did not. Thus, **H1c** was supported, **while H1a, H1b and H2** were not. Perceived leadership competency ($\beta=0.511$, $p<0.05$) showed a direct and positive effect on trust in website, thus **H3** was supported. While trust in website ($\beta=0.435$, $p<0.05$) and reward of job ($\beta=0.333$, $p<0.05$) had a direct and positive effect on willingness to disclose information online, privacy concern ($\beta=-0.012$, $p<0.01$) on the other hand did not. Thus, **H5** and **H6** are shown to be supported while **H4** is not supported.

Table 8

Summary of Hypothesis Tests

	Path Coefficient	t Value	p Value	Support
Conscientiousness -> Privacy Concern	0.248*	2.324	0.021	Yes
Extraversion -> Privacy Concern	0.072	0.744	0.457	No
Intellect -> Privacy Concern	0.080	0.903	0.367	No
Perceived Big Data Security -> Privacy Concern	0.127	1.819	0.070	No
Perceived Leadership Competency -> Trust in Website	0.511**	7.438	0.000	Yes
Privacy Concern -> Willingness to Disclose	-0.112	1.549	0.122	No
Reward of Job -> Willingness to Disclose	0.333**	5.178	0.000	Yes
Trust in Website -> Willingness to Disclose	0.435**	5.993	0.000	Yes

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

Chapter 5

Conclusions

Leadership competency is critical for an IT projects success. Based on the results, it is clear that competent leaders affect an individual's trust of the online system that collects personal sensitive information and eventually leads to individuals disclosing sensitive personal information online. As IT systems, such as the one that collects personal identifiable information for the purpose of providing US Governments jobs involves cutting edge data security and state of the art IT technology, it is important to have extremely knowledgeable, trained and certified leaders in charge of such projects. From data analysis it was noticed that willingness to disclose information online was also influenced by trust of the website and reward of a job. From information sensitivity, only conscientiousness showed to have significance on privacy concern. Perceived big data security showed significance to privacy concern when all the factors were loaded but once the two factors (PBDS1 and PBDS2) were deleted the significance dropped as well. Surprisingly and contrary to the hypothesis, privacy concern was not shown to influence willingness to disclose information online.

Implications

From the PLS Analysis and Figure 2, it was clear that an individuals' trust in the government website was explained by perceived leadership competency which was at 26 percent. Additionally, the relationship between these constructs showed the t -value to be at a very high value of 7.324. As mentioned by Hair et. al (1995), values for a study with a two-tailed test at a 5% significance level is acceptable when the t -value is greater than or equal to 1.96.

With such a high t -value, there was a strong relationship displayed between perceived leadership competency and trust in website. This means that it is important for leaders that comprise of project managers, program managers or directors to be competent for the overall success of a project, specifically IT projects that deal with cybersecurity, data and web site security as these IT systems contain personal sensitive information for millions of applicants. It is recommended that the government clearly and thoroughly vet these leaders as perceived leadership competency has shown an increase in the trust of the website and eventually willingness to provide sensitive information online. It is hoped that the US Government would consider the following 5 step process in regard to hiring of leaders:

1. Consider leaders with IT educational background that have shown years of learning in the Information System field.
2. Consider leaders with current critical industry certificates in cybersecurity, management and data security.
3. Consider leaders with relevant hands on years of experience in the Information Systems field.
4. With the fast-changing IT world, train leadership frequently and consistently.
5. Collect and maintain database on leadership, specifically on training, certification, education and experience and update the information regularly while notifying them of any current certificates or training required to be taken.

In this study, information sensitivity which comprised of three personality traits of conscientiousness, intellect and extroversion along with perceived big data security explained privacy concern of an individual at 16 percent. From information sensitivity, only the trait of conscientiousness measures the relationship with $t = 2.324$ whereas perceived big data security

comes close to the acceptable value above 1.96 with $t = 1.819$. Intellect and Extraversion both are way below the acceptable value, which implies that individuals with the personality trait of conscientiousness are definitely more likely to care for their privacy concern when providing their sensitive personal information online. This seems logical, as conscientiousness is the most widely studied personality trait of the big five traits and conscientiousness individual will sift through a variety of reputable information on privacy before submitting their information online (Osatuyi, 2015).

Lastly, willingness to disclose information online is explained by trust in website, privacy concern and reward of job at 33 percent. Trust in website has very high t -value of 5.993, privacy concern is very low with 1.549 which is well below the accepted value of 1.96 and above and reward of job is also high with 5.178. This implies that individuals that have trust in website are highly willing to disclose their personal information online. Job reward is also a significant factor to consider when attracting applicants for government and security related jobs. Thus, it is recommended that government agencies in charge of maintaining the website focus on increasing an individuals' trust in the website and decreasing their anxiety in order to get higher number of applicants. Some suggestions include, marketing and advertising clearly through the website the technical factors that show the website to be secure and most likely impenetrable to hacking. An example would be of providing a link or a page where individuals could understand how their information will be treated as well as who to contact in case of information leak. Additionally, clearly indicating steps taken to make the site and data secure by keeping up to date with the cybersecurity regulations provided by The National Institute of Standards and Technology (NIST) which provides guidance to ensure that sensitive federal information remains confidential when stored in federal and nonfederal information systems and organizations.

Reward of obtaining a government job includes job stability, monetary gain and prestige of working for the US Government, thus it is also suggested that the government highlight these benefits clearly to individuals' that are contemplating information disclosure. Some images and videos of current government employees received awards or enjoying a stable family time could encourage individuals as well.

Limitation and Future Studies

To limit the scope of this study, the construct of information sensitivity consisted of intellect, conscientiousness and extraversion. Agreeableness and neuroticism were not considered in scope for this research. It is hoped that future studies will include all of the personality traits. Even though for the Perceived Big Data Security construct there were three survey items, but due to very low factor loadings, two of the latent variables PBDS1 and PBDS2 had to be deleted which resulted in PBDS to be represented by only PBDS3. Surprisingly, this study revealed that privacy concern and perceived big data security were not considered substantial factors when disclosing personal information. Thus, it is suggested that future research focus on both privacy concern as well as perceived big data security and investigate these constructs further. Lastly, as this study collected data from United States only, therefore it is suggested that future studies research on data collected from different samples of the population and from different parts of the world instead of from only the United States.

Random sampling was used in this study to collect survey data to make the study generalizable, and to collect true values in the population in order for it to not suffer from sampling bias. All possible subsets of the sampling frame were given an equal probability of being selected, but as the surveys were online based only, there were some sampling bias as paper based surveys were not be mailed. Additionally, even though this study collected data

through random sampling, the scope and audience of this study was limited, meaning that it is only applicable to individuals that are considering or would consider providing personal sensitive information online in order to apply for a US Government cleared job only. Additionally, this study does not look at the cause and effect or the causality of the independent variables on the dependent variable – “willingness of disclosing personal information online” for any e-government online sensitive information disclosure but rather focuses on only US Clearance related jobs only. Therefore, one of the limitations of this study is that its focuses on only a small percentage of the population and probably future studies can look into expanding this study for different countries that have to go through the same rigorous information disclosure process.

Summary

This study first identified and defined a problem that exists within the Information Systems field, specifically, individuals’ reluctance to provide sensitive personal information online and how it can affect the US Governments’ ability to hire and retain qualified personnel for sensitive cleared positions. The aim of the introduction was to give a brief overview of how trust, privacy concern, information sensitivity, perceived big data security, perceived leadership competency and reward of a job play a significant role in limiting an individuals’ willingness of disclosing sensitive personal information online. Detailed description of prior literature and comparative studies were highlighted in this study. Some research questions were developed for this study and based on the research questions, this study highlighted and proposed some hypotheses as well as a conceptual model. Just like any other research, this study also had some barriers and issues that were encountered and therefore, have been duly noted in this study.

The literature review chapter highlighted literature from prior studies in the same field, limitations, assumptions and their research scope as well. Three major foundational studies that

this research is based on are the studies by Bansal et al. (2010), Beldad et al. (2012) and Osatuyi (2015). Bansal et al. (2010) looked into the process by which personal dispositions including risk and privacy concern impact individuals' trust and behavior intention to disclose personal information online for the healthcare industry. The research by Osatuyi (2015) focused on how personality traits affect information privacy concern for an individual and the study by Beldad et al. (2012) researched the factors affecting individuals disclosing personal information for e-government. Utility theory has been the basis of this research and details of the theory as well as the justification to base this research of were highlighted in the literature review chapter.

The Research Method chapter detailed the research design employed for this study and highlighted the survey instrument, data collection techniques, sample data used as well as the validity and reliability of the instruments to be used for this research. A quantitative survey study was considered best suited for this research to analyze the effect of the independent variables on the dependent variable. Research Design included collection of data through random sampling and data analysis included the use of both descriptive (mean, median, and mode) and inferential statistics (*t* test, ANOVA and factor analysis). Descriptive Statistics included demographic data on age, gender, years of computer use, academic level achieved and prior experience with government websites. A pilot test involving 15 participants was performed and initially the surveys were created using Google Forms then transferred into Survey Monkey. Threat to validity and reliability of this study was detailed to understand if the results were caused by an intervention as well as to understand the generalizability of this study to different locations and populations. Based on the results, details on the interpretation of the data analyzed was highlighted and consequently, evidence in support or rejection of the hypotheses was

detailed. In conclusion, implications of this study, limitations and recommendations for further studies was also highlighted.

Finally, given the limited research in this field and as mentioned by Levy and Ellis (2006), the main definitional component of research is the ability to add to the current body of knowledge, thus it is believed that this research will contribute to the body of knowledge on willingness to disclose information online. It is also hoped that this study and its implications will be beneficial to practitioners of the US Government clearance processes when attracting individuals applying for secure online jobs and the ability of the US Government to find and hire cleared individuals for sensitive nations security positions possibly increase.

Appendix

Appendix A:

Survey Questionnaire

Survey on willingness to disclose sensitive personal information on government website

Dear Participant,

I, Iqbal Amiri am a doctoral student from Nova Southeastern University pursuing a PhD in Information Systems and for my research, am seeking anonymous input to some survey questions.

The survey relates to understanding the factors that lead individuals' to disclose their personal information on government websites in order to obtain a security clearance and a job with the US Government.

Questions are based on your personality traits, privacy concern, trustworthiness and security of the website, competency of the leadership, your willingness to disclose information and the reward of getting a job after applying online.

Responses to the survey are completely anonymous, thus I will be neither collecting nor storing any personal identifiable information. As this survey is geared towards individuals that are working or looking for a job, therefore it is requested that only individuals above 18 years of age respond to this survey.

If you have any questions, you can reach me at: iamiri@sarinait.com
Thank you in advance for your participation in this survey.

Please rate the following questions using the following scale:

- 1 – Strongly disagree**
- 2 – Disagree**
- 3 – Somewhat disagree**
- 4 – Neither agree or disagree**
- 5 – Somewhat agree**
- 6 – Agree**
- 7 – Strongly agree**

Regards,
Iqbal Amiri

Demographic Information

* What is your Gender

- Male
 Female

* What is your Age

- 18 and under
 19-34
 35-50
 51-66
 67 or older

* Highest Academic Level Achieved

- High School
 Bachelors
 Masters
 PhD
 Other

* How many years have you been using Computers?

- Less than 5
 5-9
 10-14
 15 or More

* I have used a government website to do an online transaction

- Yes
 No

Appendix B:*IRB Approval*

To: Iqbal Amiri
College of Engineering and Computing

From: Nurit Sheinberg, Ed.D.
Chair, Institutional Review Board

Date: October 3, 2017

Subject: IRB Exempt Initial Approval Memo

TITLE: The Efficacy of Perceived Big Data Security, Trust, Perceived Leadership Competency, Information Sensitivity, Privacy Concern and Job Reward on Disclosing Personal Security Information Online— NSU IRB Protocol Number 2017-541

Dear Principal Investigator,

Your submission has been reviewed and approved by the Institutional Review Board on **September 14, 2017**. You may proceed with your study.

Please Note: If you receive stamped copies of consent, assent, and recruiting materials indicating approval date, these documents must be used when recruiting and consenting or assenting participants.

Level of Review: Exempt

Type of Approval: Initial Approval

Exempt Review Category: Exempt Category 2

Post-Approval Monitoring: The IRB Office conducts post-approval review and monitoring of all studies involving human participants under the purview of the NSU IRB. The Post-Approval Monitor may randomly select any active study for a Not-for-Cause Evaluation.

Final Report: You are required to notify the IRB Office within 30 days of the conclusion of the research that the study has ended using the IRB Closing Report Form.

Translated Documents: No

Please retain this document in your IRB correspondence file.

CC: Ling Wang, Ph.D.
Wei Li, Ph.D.

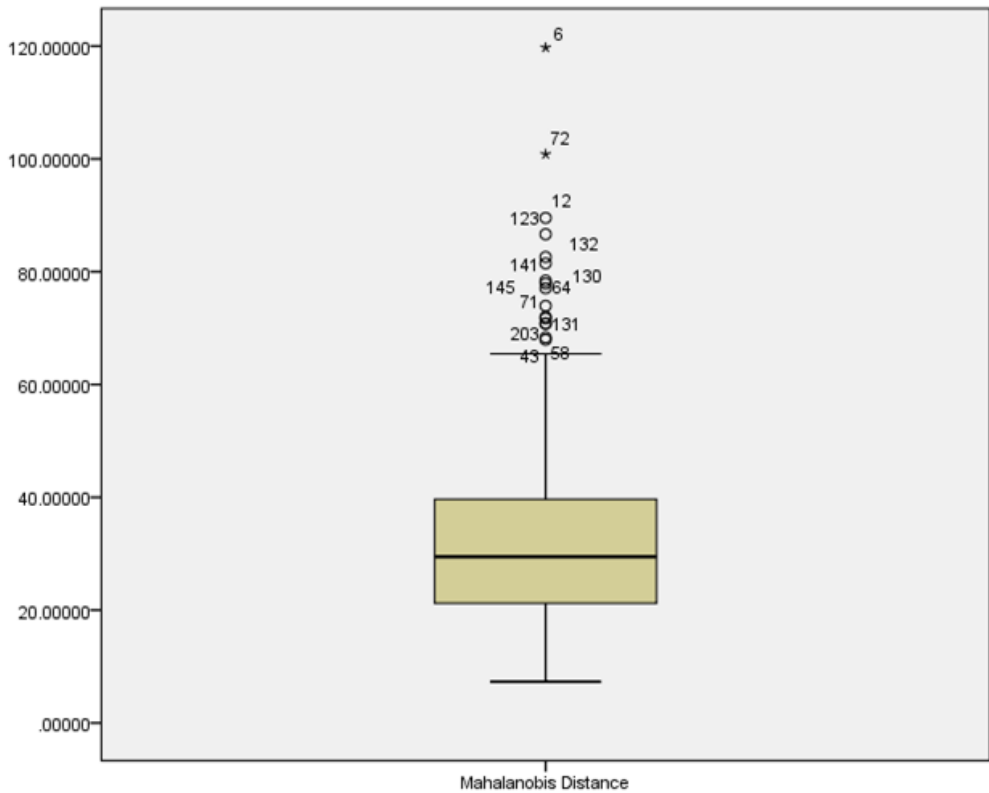
Appendix C:

Total of 206 Collected Data

The image shows a screenshot of an Excel spreadsheet titled "SMCleanedFinal - Excel". The spreadsheet contains a large table of data with columns labeled A through AN and rows numbered 1 through 206. The data is organized into a grid with various cell styles and colors. The ribbon at the top includes tabs for File, Home, Insert, Page Layout, Formulas, Data, Review, and View. The Home tab is active, showing options for font, paragraph, and styles. The Styles section shows a color palette with categories like Normal, Bad, Good, Neutral, Calculation, Input, and Note. The spreadsheet content is a dense grid of numbers and text, with some cells highlighted in green and others in yellow. The columns are labeled with letters A through AN, and the rows are numbered 1 through 206. The data appears to be a collection of numerical values, possibly representing collected data points, organized in a structured manner.

Appendix D:

Mahalanobis Distance and Stem & Leaf Plot



Mahalanobis Distance Stem-and-Leaf Plot

Frequency	Stem &	Leaf
1.00	0 .	7
27.00	1 .	023344444444444444444444444444
18.00	1 .	55555577888999999999
30.00	2 .	00000112222233333333333333344444
33.00	2 .	5555555666666666677788888999999999
30.00	3 .	0000011111111222333333444444444
16.00	3 .	555666678888899999
8.00	4 .	00112333
9.00	4 .	6777788899
5.00	5 .	01122
8.00	5 .	55566889
4.00	6 .	0224
1.00	6 .	5
16.00	Extremes	(>=68)

Stem width: 10.00000
 Each leaf: 1 case(s)

Extreme Values

		Case Number		Value
Mahalanobis Distance	Highest	1	6	119.73250
		2	72	100.83536
		3	12	89.54368
		4	123	86.63731
		5	29	82.60992
	Lowest	1	88	7.32105
		2	35	10.97143
		3	47	12.66929
		4	94	13.07484
		5	95	13.33153

Tests of Normality

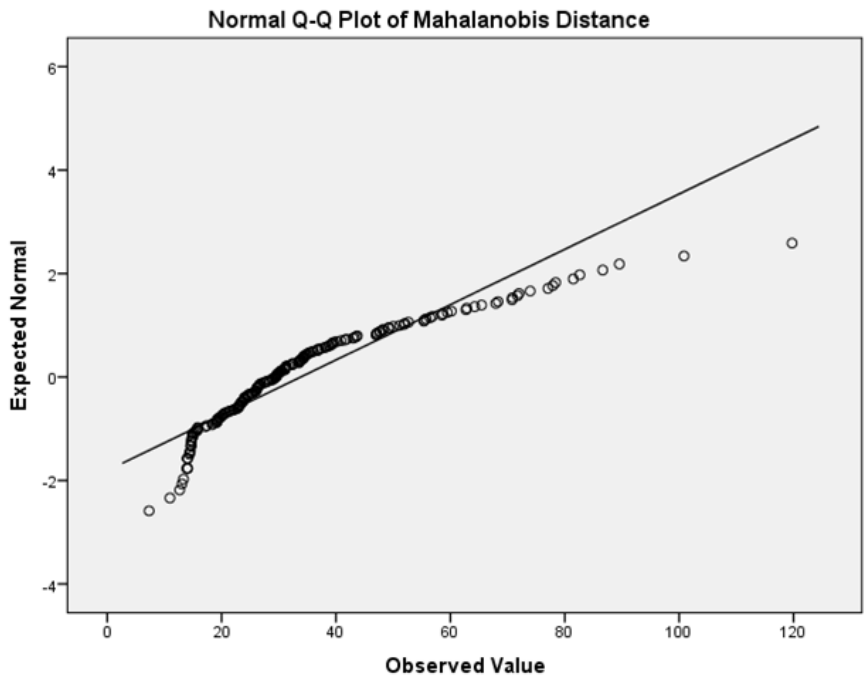
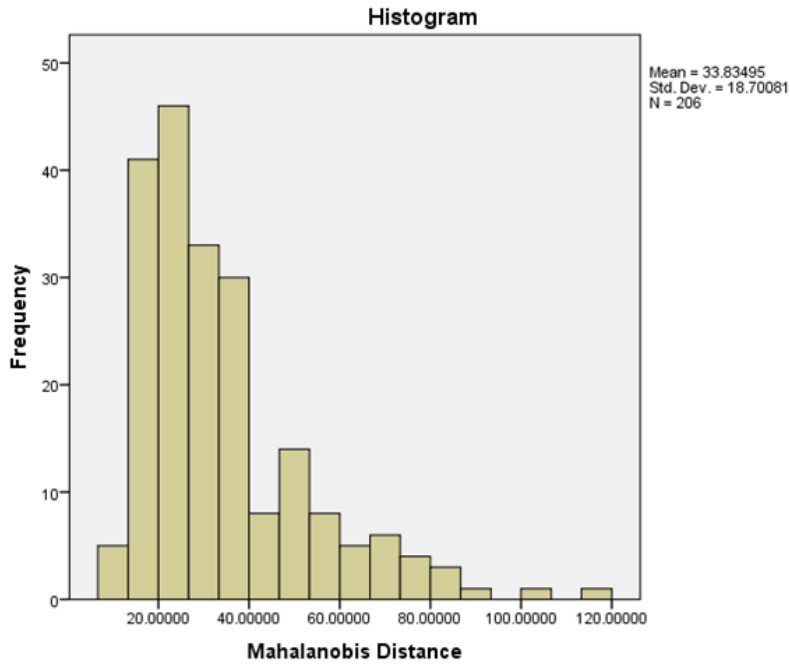
	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Mahalanobis Distance	.152	206	.000	.868	206	.000

a. Lilliefors Significance Correction

Descriptives

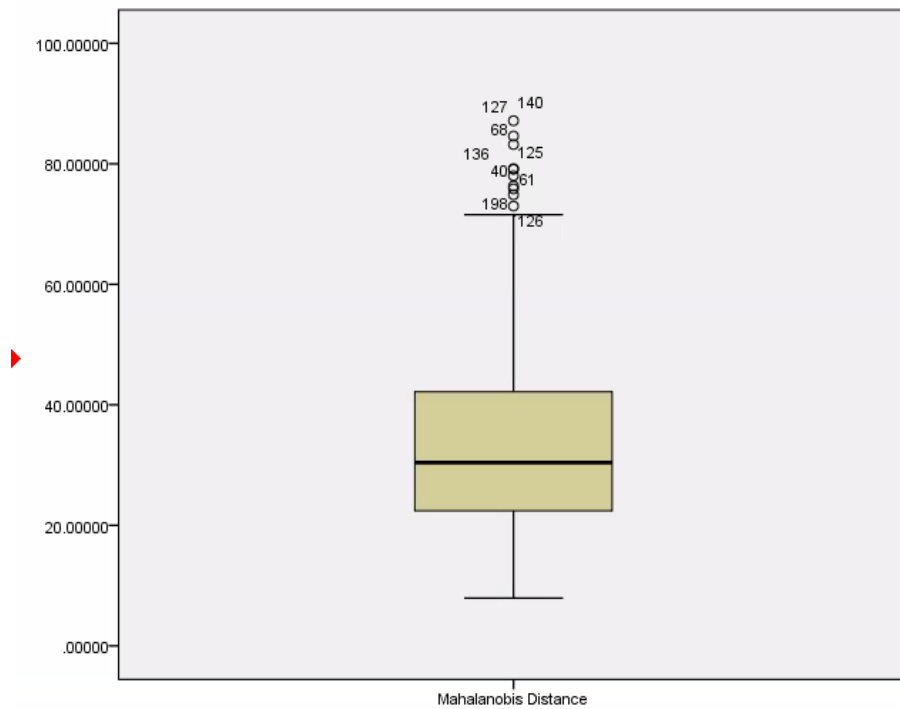
		Statistic	Std. Error	
Mahalanobis Distance	Mean	33.8349515	1.30294731	
	95% Confidence Interval for Mean	Lower Bound	31.2660560	
		Upper Bound	36.4038469	
	5% Trimmed Mean	32.0852649		
	Median	29.4529858		
	Variance	349.720		
	Std. Deviation	18.70081195		
	Minimum	7.32105		
	Maximum	119.73250		
	Range	112.41145		
	Interquartile Range	18.65066		
	Skewness	1.524	.169	
	Kurtosis	2.736	.337	

Mahalanobis Distance



Appendix E:

Rerun of Mahalanobis Distance and Stem & Leaf Plot after 5 extreme values deleted

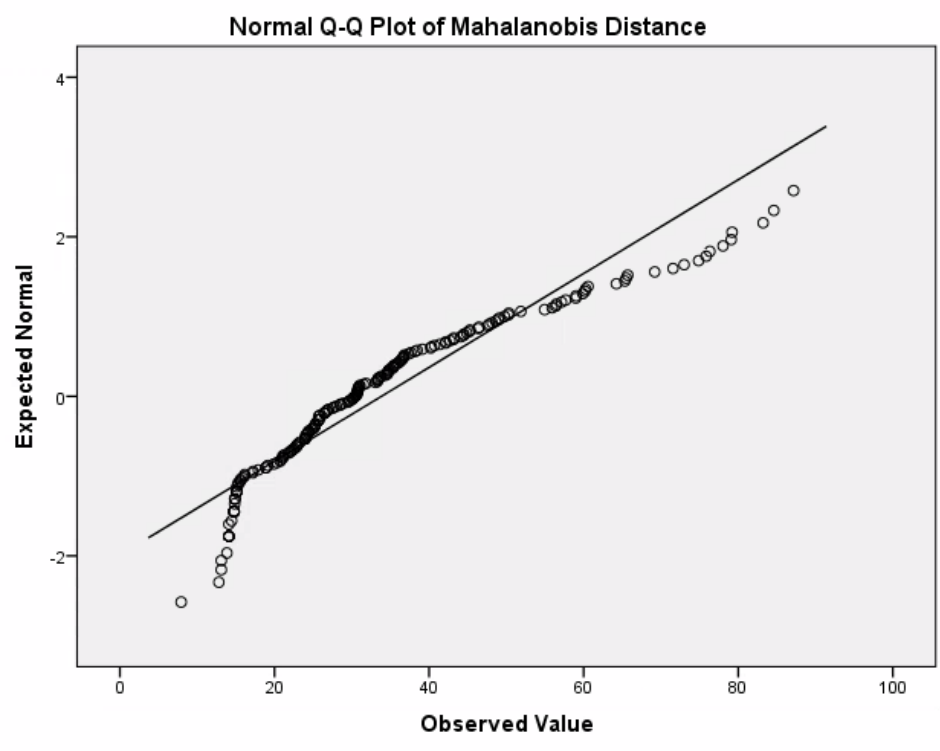
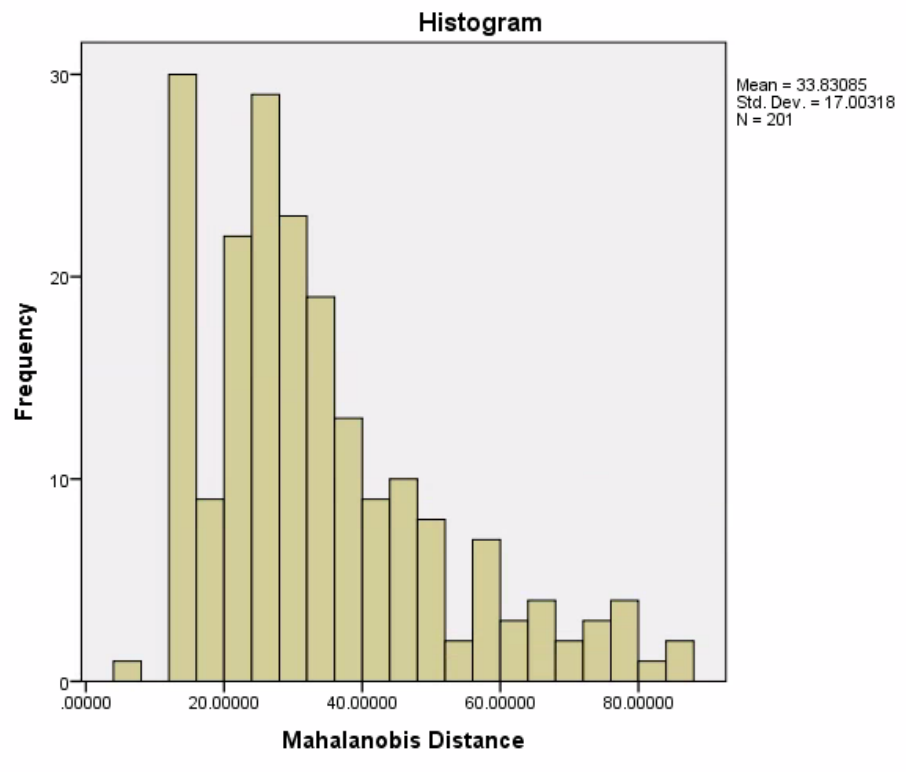


Mahalanobis Distance Stem-and-Leaf Plot

Frequency	Stem & Leaf
1.00	0 . 7
20.00	1 . 23334444444444444444
19.00	1 . 5555555555667778899
29.00	2 . 00011111222222233333334444444
28.00	2 . 555555555555666667777888999
31.00	3 . 00000000000000111133333334444444
18.00	3 . 555566666666677889
13.00	4 . 0001222334444
11.00	4 . 55667788899
4.00	5 . 0014
8.00	5 . 56677899
4.00	6 . 0004
4.00	6 . 5559
1.00	7 . 1
10.00	Extremes (>=73)

Stem width: 10.00000
Each leaf: 1 case(s)

Mahalanobis Distance



Extreme Values

			Case Number	Value
Mahalanobis Distance	Highest	1	127	87.15250
		2	140	84.60270
		3	68	83.21101
		4	136	79.20079
		5	125	79.09166
	Lowest	1	84	7.92354
		2	44	12.80731
		3	32	13.10870
		4	91	13.11293
		5	90	13.83522

Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Mahalanobis Distance	.131	201	.000	.903	201	.000

a. Lilliefors Significance Correction

Descriptives

		Statistic	Std. Error	
Mahalanobis Distance	Mean	33.8308458	1.19931152	
	95% Confidence Interval for Mean	Lower Bound	31.4659279	
		Upper Bound	36.1957636	
	5% Trimmed Mean	32.4666459		
	Median	30.4206624		
	Variance	289.108		
	Std. Deviation	17.00317536		
	Minimum	7.92354		
	Maximum	87.15250		
	Range	79.22897		
	Interquartile Range	19.89827		
	Skewness	1.116	.172	
	Kurtosis	.823	.341	

Appendix F:

Normality and Scatter Plot

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.616 ^a	.380	.354	1.15696

a. Predictors: (Constant), AGGRISI, AGGRPBDs, AGGRJR, AGGRPC, AGGRPLCI, AGGRISE, AGGRISC, AGGRTW

b. Dependent Variable: AGGRWDIO

ANOVA^a

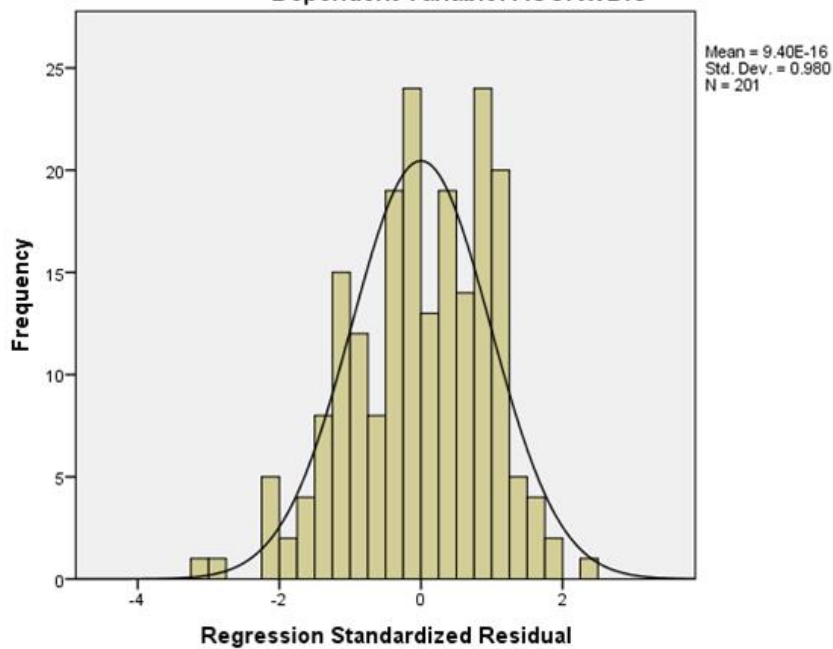
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	157.254	8	19.657	14.685	.000 ^b
	Residual	257.002	192	1.339		
	Total	414.256	200			

a. Dependent Variable: AGGRWDIO

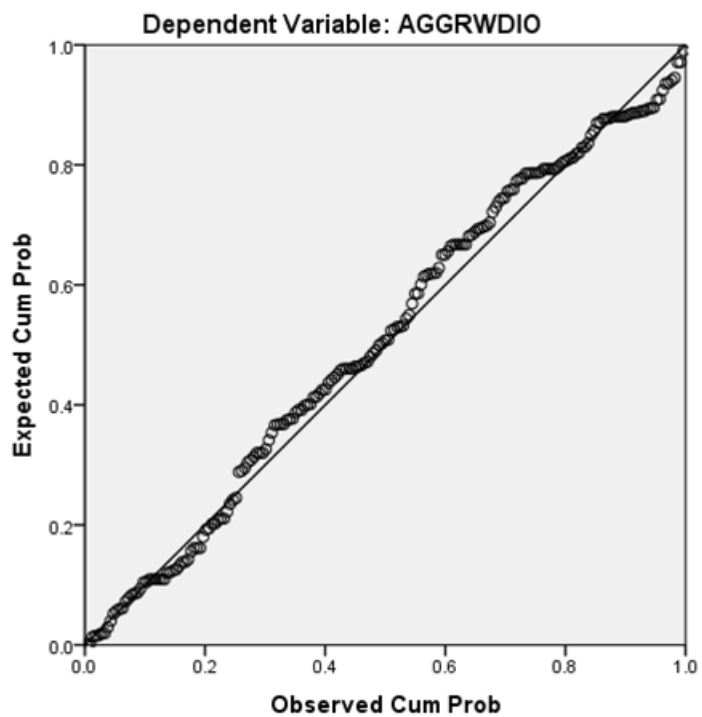
b. Predictors: (Constant), AGGRISI, AGGRPBDs, AGGRJR, AGGRPC, AGGRPLCI, AGGRISE, AGGRISC, AGGRTW

Histogram

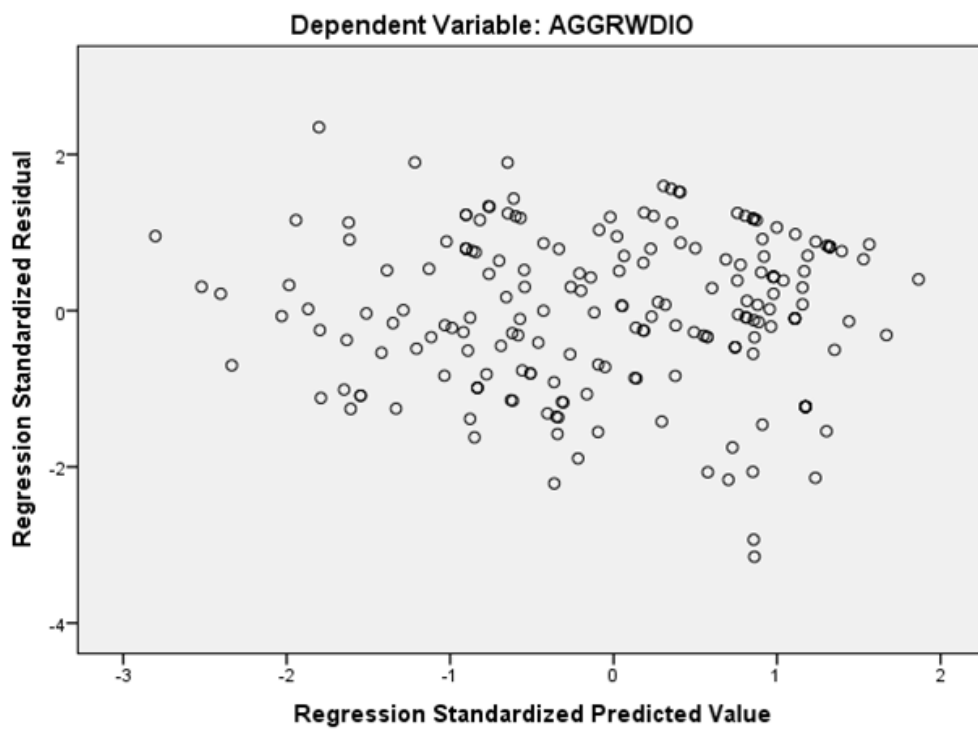
Dependent Variable: AGGRWDIO



Normal P-P Plot of Regression Standardized Residual



Scatterplot



Appendix G:

Descriptive Statistics

		Statistics				
		GNDR	AGE	HLA	CUY	UGC
N	Valid	201	201	201	201	201
	Missing	0	0	0	0	0
Mean		1.48	3.08	2.48	3.67	1.25
Median		1.00	3.00	2.00	4.00	1.00
Mode		1	3	2	4	1
Std. Deviation		.501	.902	1.141	.665	.436

Frequency Table

		GNDR			Cumulative
		Frequency	Percent	Valid Percent	Percent
Valid	1	105	52.2	52.2	52.2
	2	96	47.8	47.8	100.0
Total		201	100.0	100.0	

		AGE			Cumulative
		Frequency	Percent	Valid Percent	Percent
Valid	2	59	29.4	29.4	29.4
	3	82	40.8	40.8	70.1
	4	45	22.4	22.4	92.5
	5	15	7.5	7.5	100.0
Total		201	100.0	100.0	

		HLA			Cumulative
		Frequency	Percent	Valid Percent	Percent
Valid	1	42	20.9	20.9	20.9
	2	68	33.8	33.8	54.7
	3	58	28.9	28.9	83.6
	4	18	9.0	9.0	92.5
	5	15	7.5	7.5	100.0
Total		201	100.0	100.0	

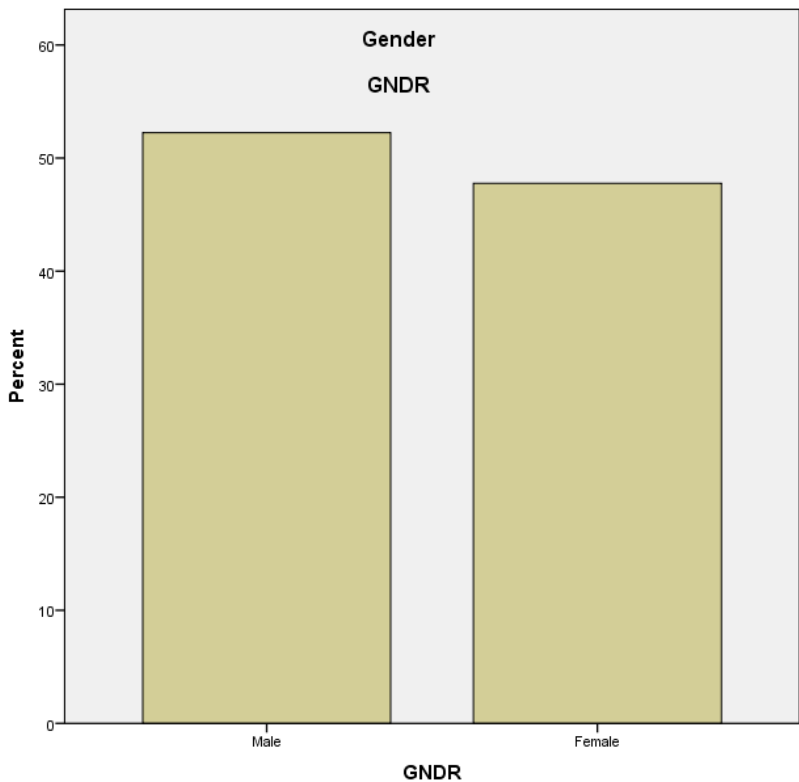
CUY

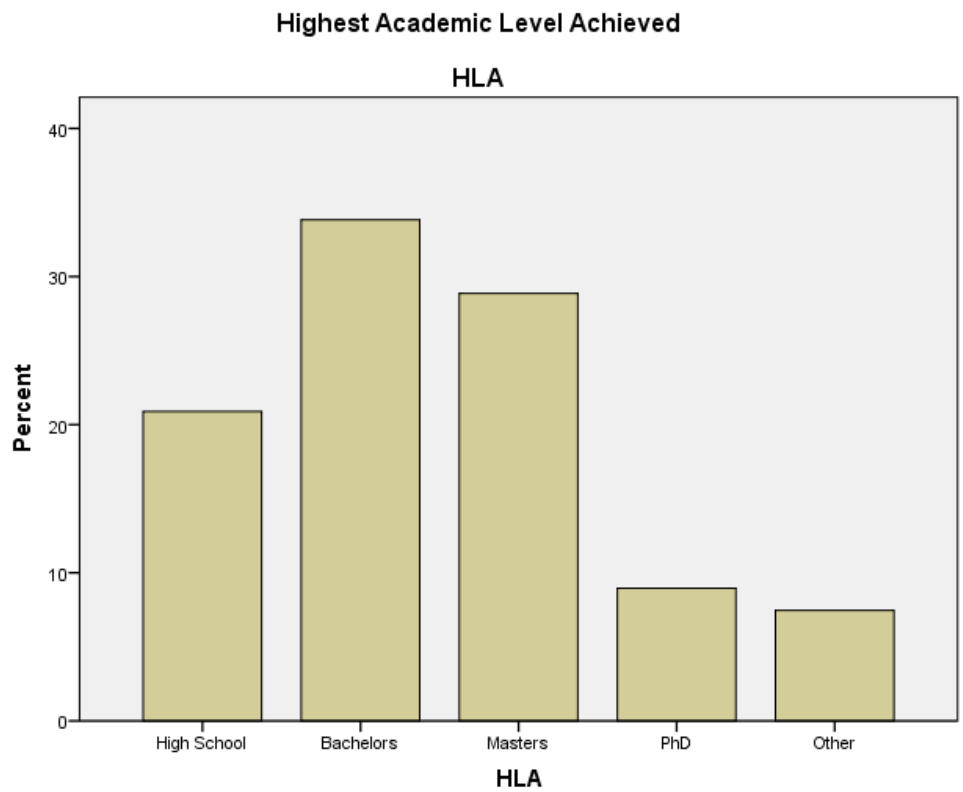
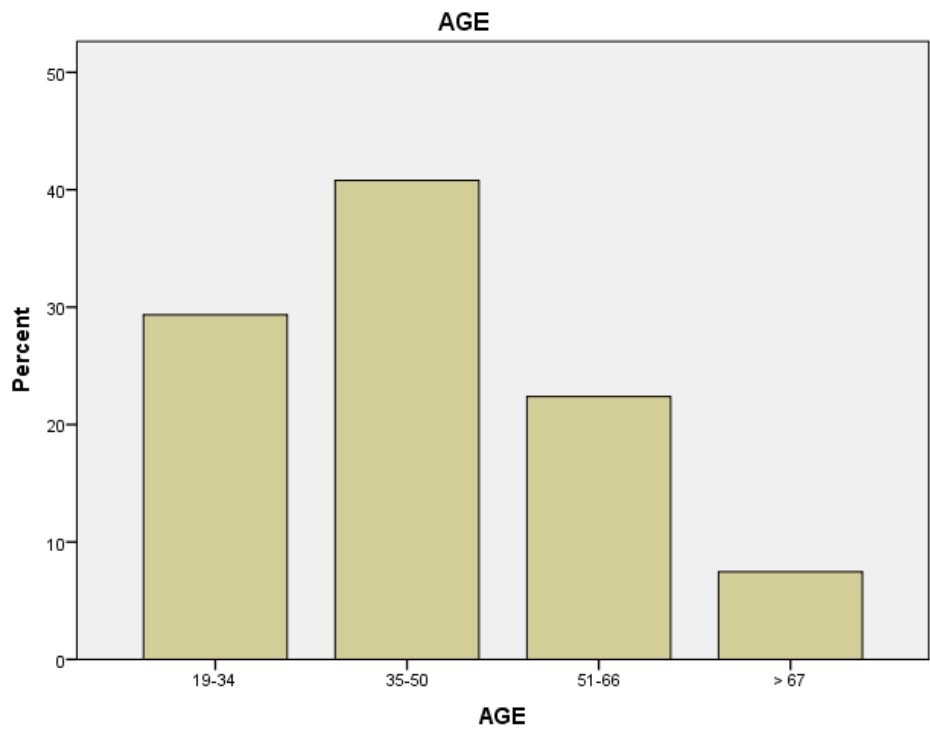
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	2	1.0	1.0	1.0
	2	16	8.0	8.0	9.0
	3	28	13.9	13.9	22.9
	4	155	77.1	77.1	100.0
	Total	201	100.0	100.0	

UGC

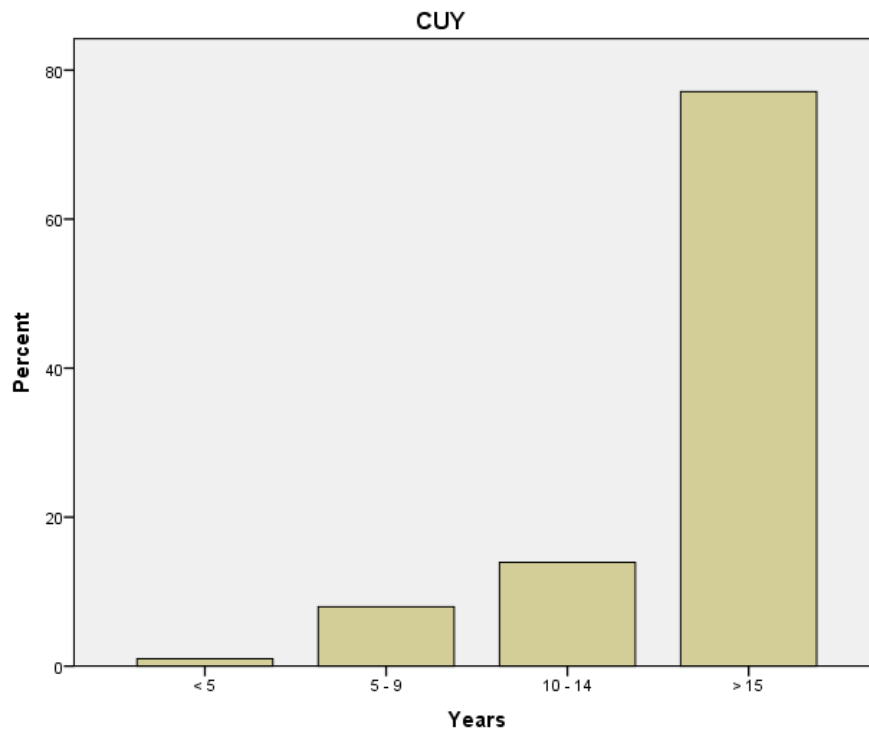
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	150	74.6	74.6	74.6
	2	51	25.4	25.4	100.0
	Total	201	100.0	100.0	

Bar Chart

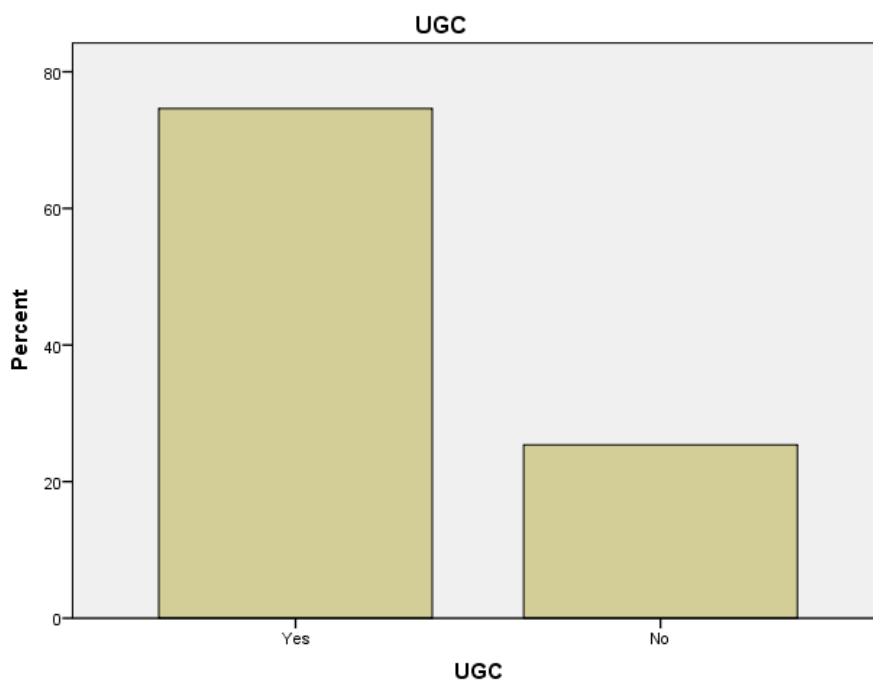




Years of Computer Use

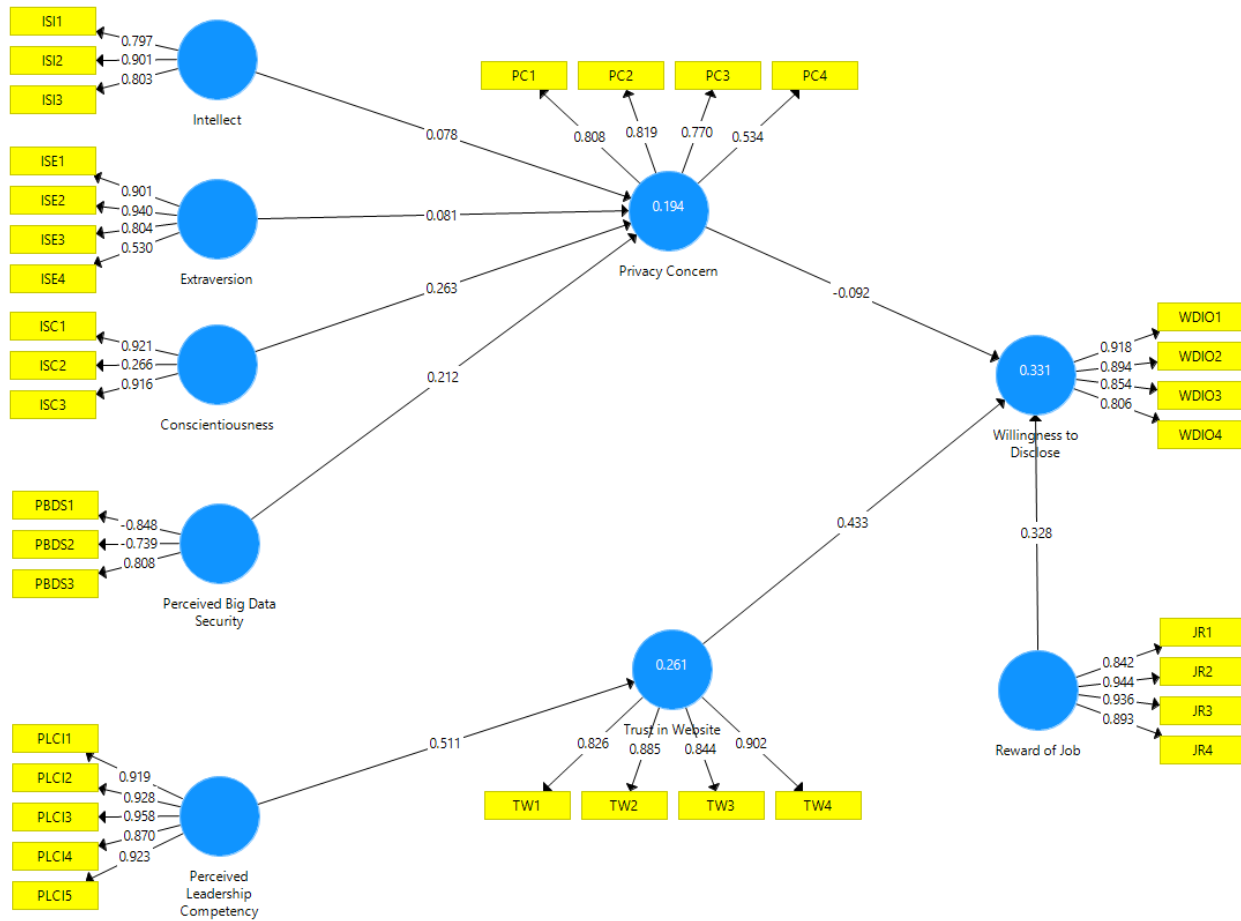


Prior Government Website Use



Appendix H:

PLS Analysis



Appendix I:

Model fit, Reliability, Validity, Coefficient and Outer Loading

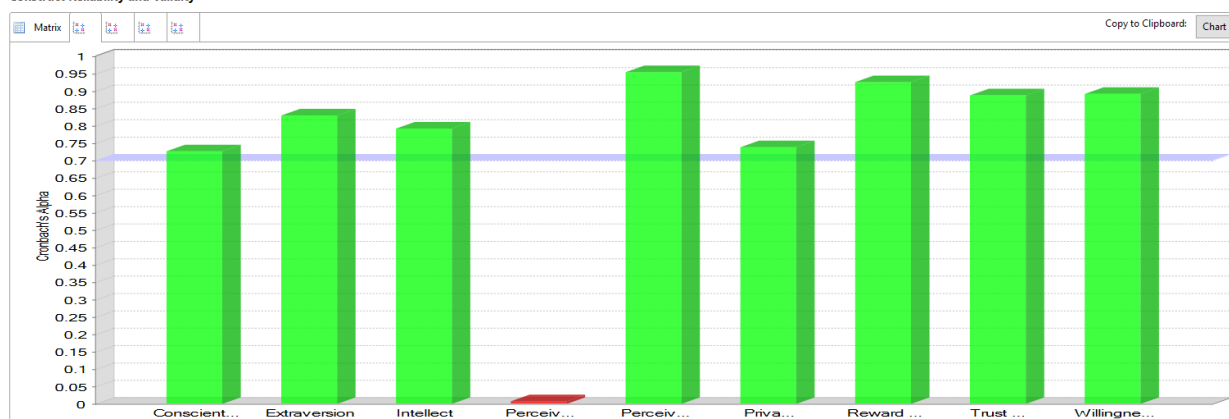
Model_Fit

Fit Summary		rms Theta	
	Saturated Model	Estimated Mo...	
SRMR	0.093	0.124	
d_ULS	5.189	9.168	
d_G1	2.837	3.052	
d_G2	1.812	2.036	
Chi-Square	1,915.341	2,070.501	
NFI	0.676	0.650	

Construct Reliability

	Cronbach's Alpha	rho_A	Composite Reliability	Average Variance Extracted (AVE)
Conscientiousness	0.727	0.669	0.781	0.586
Extraversion	0.830	0.971	0.880	0.655
Intellect	0.792	0.894	0.873	0.697
Perceived Big Data Security	0.008	0.745	0.359	0.639
Perceived Leadership Competency	0.954	0.962	0.965	0.847
Privacy Concern	0.739	0.810	0.827	0.550
Reward of Job	0.926	0.950	0.947	0.819
Trust in Website	0.888	0.904	0.922	0.748
Willingness to Disclose	0.892	0.904	0.925	0.755

Construct Reliability and Validity

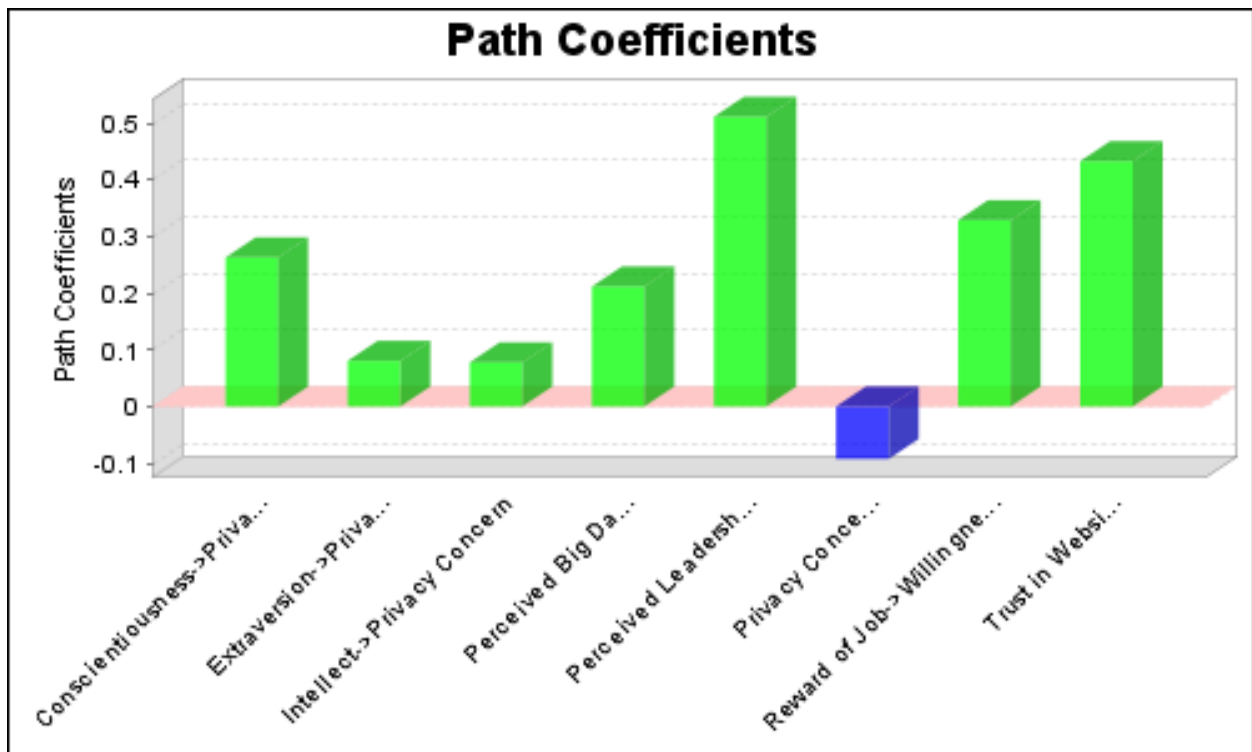


PLS Factor Analysis

	Missing	Loading	Mean	Median	SD	Excess Kurtosis	Skewness
PBDS1	0.000	-0.848	4.353	5.000	1.806	-1.054	-0.150
PBDS2	0.000	-0.739	4.498	5.000	1.823	-0.845	-0.519
PBDS3	0.000	1.000	4.398	5.000	1.555	-0.386	-0.532
TW1	0.000	0.826	4.423	5.000	1.641	-0.931	-0.427
TW2	0.000	0.885	4.567	5.000	1.589	-0.684	-0.585
TW3	0.000	0.844	4.647	5.000	1.421	0.243	-0.552
TW4	0.000	0.902	4.612	5.000	1.558	-0.129	-0.743
PC1	0.000	0.762	5.189	6.000	1.802	-0.504	-0.851
PC2	0.000	0.812	4.502	5.000	1.533	-0.882	-0.259
PC3	0.000	0.787	4.716	5.000	1.703	-1.181	-0.250
PC4	0.000	0.609	5.448	6.000	1.400	-0.007	-0.814
JR1	0.000	0.842	3.512	4.000	1.400	-0.453	-0.091
JR2	0.000	0.944	3.537	4.000	1.624	-0.556	0.179
JR3	0.000	0.936	3.483	4.000	1.593	-0.828	0.129
JR4	0.000	0.893	3.328	4.000	1.587	-0.242	0.411
WDIO1	0.000	0.918	3.806	4.000	1.635	-1.055	-0.173
WDIO2	0.000	0.894	3.980	4.000	1.733	-1.092	-0.160
WDIO3	0.000	0.854	3.657	4.000	1.617	-1.108	-0.035
WDIO4	0.000	0.806	4.169	4.000	1.584	-0.862	-0.342
PLCI1	0.000	0.919	4.219	4.000	1.631	-0.628	-0.454
PLCI2	0.000	0.928	4.174	4.000	1.604	-0.535	-0.491
PLCI3	0.000	0.958	4.453	4.000	1.599	-0.150	-0.473
PLCI4	0.000	0.870	4.617	5.000	1.551	0.086	-0.672
PLCI5	0.000	0.923	4.403	4.000	1.687	-0.424	-0.526
ISE1	0.000	0.892	5.045	5.000	1.408	-0.363	-0.662
ISE2	0.000	0.934	4.846	5.000	1.293	-0.825	-0.446
ISE3	0.000	0.813	4.592	5.000	1.387	-0.508	-0.457
ISE4	0.000	0.557	4.378	4.000	1.475	-0.549	-0.233
ISC1	0.000	0.929	5.682	6.000	1.356	-0.195	-0.950
ISC2	0.000	0.303	5.736	6.000	1.113	0.499	-0.881
ISC3	0.000	0.915	5.075	5.000	1.184	-0.302	-0.526
ISI1	0.000	0.796	5.333	5.000	0.969	-0.203	-0.348
ISI2	0.000	0.902	5.259	5.000	1.089	0.184	-0.646
ISI3	0.000	0.802	5.388	6.000	1.213	0.242	-0.646

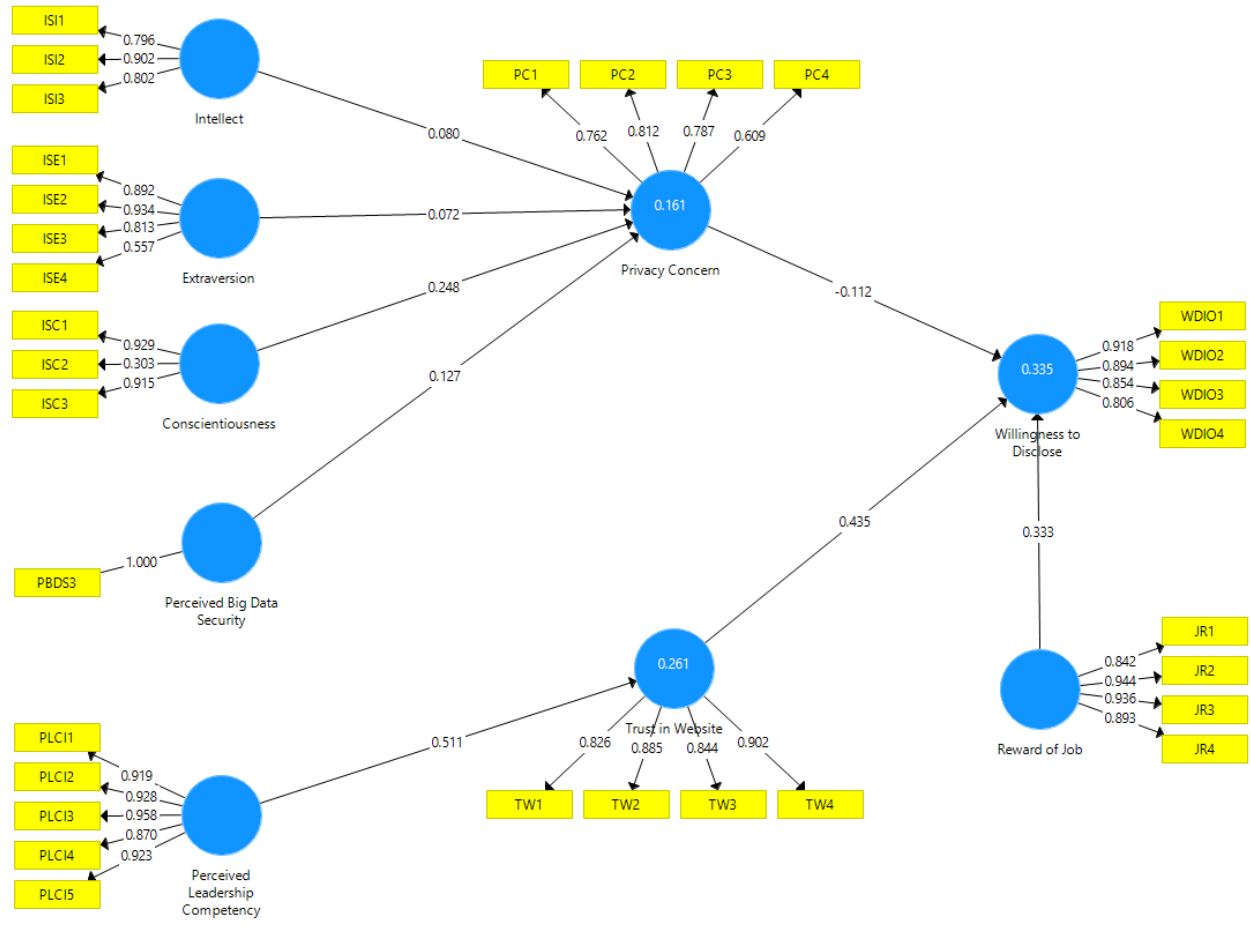
Path Coefficients

	Conscientious...	Extraversion	Intellect	Perceived Big ...	Perceived Lead...	Privacy Concern	Reward of Job	Trust in Website	Willingness to ...
Conscientiousness						0.263			
Extraversion						0.081			
Intellect						0.078			
Perceived Big Data Security						0.212			
Perceived Leadership Competency								0.511	
Privacy Concern									-0.092
Reward of Job									0.328
Trust in Website									0.433
Willingness to Disclose									



Appendix J:

PLS Analysis after deleting PBDS1 and PBDS2



Appendix K:

Model fit, Reliability, Validity, Coefficient and Outer Loading

Model_Fit

Fit Summary		rms Theta	
	Saturated Model	Estimated Model	
SRMR	0.077	0.097	
d_ULS	3.149	5.017	
d_G1	2.125	2.207	
d_G2	1.313	1.392	
Chi-Square	1,466.847	1,539.416	
NFI	0.722	0.709	

R Square

Matrix		R Square	R Square Adjusted
Privacy Concern		0.161	0.144
Trust in Website		0.261	0.258
Willingness to ...		0.335	0.325

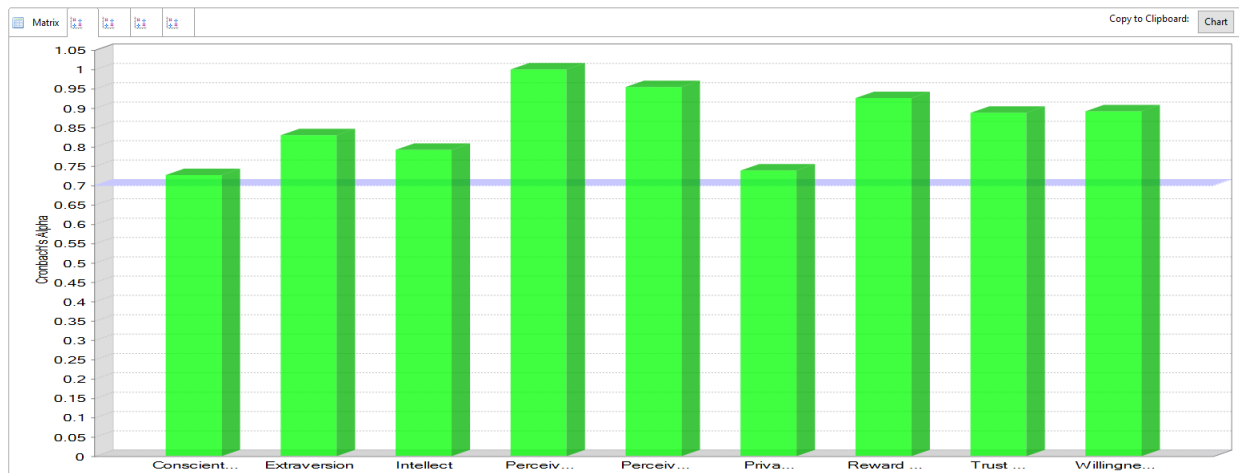
Discriminant Validity

Fornell-Larcker Criterion		Cross Loadings		Heterotrait-Monotrait Ratio (HTMT)					
	Conscientious...	Extraversion	Intellect	Perceived Big ...	Perceived Lead...	Privacy Concern	Reward of Job	Trust in Website	Willingness to ...
Conscientious...	0.773								
Extraversion	0.639	0.812							
Intellect	0.511	0.563	0.835						
Perceived Big ...	0.274	0.254	0.028	1.000					
Perceived Lead...	0.111	0.199	0.273	-0.180	0.920				
Privacy Concern	0.369	0.308	0.251	0.215	0.155	0.747			
Reward of Job	-0.055	0.129	0.200	-0.018	0.277	0.290	0.905		
Trust in Website	0.115	0.194	0.274	-0.352	0.511	0.024	0.160	0.865	
Willingness to ...	0.010	0.184	0.229	-0.166	0.445	-0.005	0.370	0.485	0.869

Construct Reliability and Validity

	Cronbach's Alpha	rho_A	Composite Reliability	Average Variance Extracted (AV...
Conscientiousness	0.727	0.734	0.792	0.597
Extraversion	0.830	0.946	0.882	0.660
Intellect	0.792	0.896	0.873	0.697
Perceived Big Data Security	1.000	1.000	1.000	1.000
Perceived Leadership Competency	0.954	0.962	0.965	0.847
Privacy Concern	0.739	0.758	0.833	0.558
Reward of Job	0.926	0.950	0.947	0.819
Trust in Website	0.888	0.904	0.922	0.748
Willingness to Disclose	0.892	0.903	0.925	0.755

Construct Reliability and Validity



Path Coefficients

	Original Sample (O)	Sample Mean (M)	Standard Deviation (STDEV)	T Statistics (O/STDEV)	P Values
Conscientiousness -> Privacy Concern	0.248	0.249	0.107	2.324	0.021
Extraversion -> Privacy Concern	0.072	0.074	0.097	0.744	0.457
Intellect -> Privacy Concern	0.080	0.095	0.088	0.903	0.367
Perceived Big Data Security -> Privacy Concern	0.127	0.122	0.070	1.819	0.070
Perceived Leadership Competency -> Trust in Website	0.511	0.512	0.069	7.438	0.000
Privacy Concern -> Willingness to Disclose	-0.112	-0.108	0.072	1.549	0.122
Reward of Job -> Willingness to Disclose	0.333	0.336	0.064	5.178	0.000
Trust in Website -> Willingness to Disclose	0.435	0.439	0.073	5.993	0.000

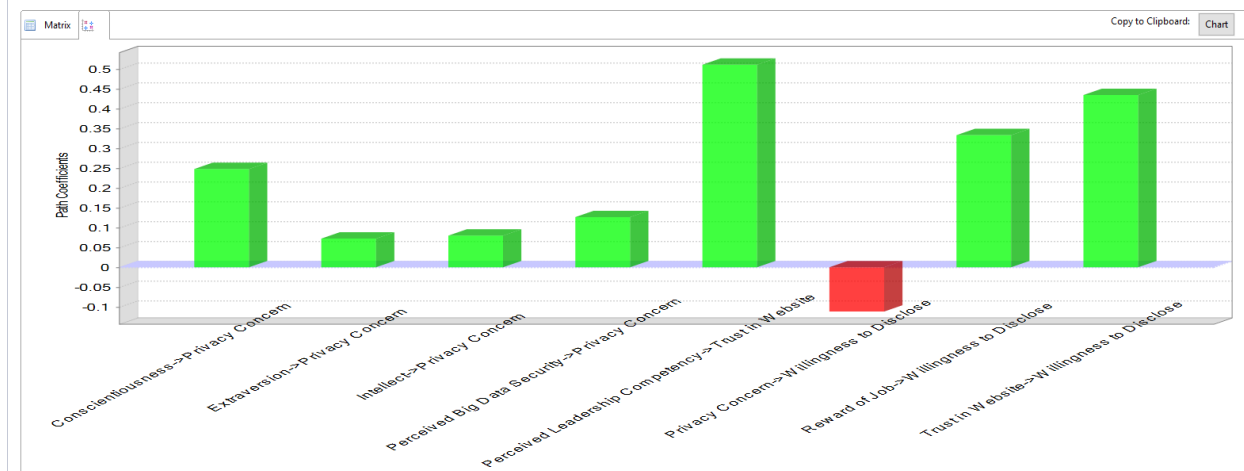
Outer Loadings

	Conscientiousness	Extraversion	Intellect	Perceived Big Data Security	Perceived Leadership Competency	Privacy Concern	Reward of Job	Trust in Website	Willingness to Disclose
ISC1	0.929								
ISC2	0.303								
ISC3	0.915								
ISE1		0.892							
ISE2		0.934							
ISE3		0.813							
ISE4		0.557							
ISI1			0.796						
ISI2			0.902						
ISI3			0.802						
JR1							0.842		
JR2							0.944		
JR3							0.936		
JR4							0.893		
PBDS3				1.000					
PC1						0.762			
PC2						0.812			
PC3						0.787			
PC4						0.609			
PLC11					0.919				
PLC12					0.928				
PLC13					0.958				
PLC14					0.870				
PLC15					0.923				
TW1								0.826	
TW2								0.885	
TW3								0.844	
TW4								0.902	
WDOI1									0.918
WDOI2									0.894
WDOI3									0.854
WDOI4									0.806

Path Coefficients

	Conscienti...	Extraversi...	Intellect	Perceived Bi...	Perceived Lead...	Privacy Concern	Reward of ...	Trust in Website	Willingness to ...
Conscientious...						0.248			
Extraversion						0.072			
Intellect						0.080			
Perceived Big ...						0.127			
Perceived Lead...								0.511	
Privacy Concern									-0.112
Reward of Job									0.333
Trust in Website									0.435
Willingness to ...									

Path Coefficients



References

- Access delayed fixing the security clearance process. Part II hearing before the Oversight of Government Management, the Federal Workforce, and the District of Columbia Subcommittee of the Committee on Homeland Security and Governmental Affairs, 109th Cong. 1* (2005).
- Analoui, B. D., Clair, H. D., & Sambrook, S. (2013). Leadership and knowledge management in UK ICT organisations. *The Journal of Management Development, 32*(1), 4-17.
- Avolio, B. J., & Gardner, W. L. (2005). Authentic leadership development: Getting to the root of positive forms of leadership. *Leadership Quarterly, 16*(3), 315-338.
- Awad, N.F., & Krishnan, M.S. (2006). The personalization privacy paradox: an empirical evaluation of information transparency and the willingness to be profiled online for personalization, *MIS Quarterly, 30*(1), 13–28.
- Bacharach, S. B. (1989). Organizational Theories: Some Criteria for Evaluation. *Academy of Management Review, 14*(4), 496-515.
- Bansal, G., & Gefen, D. (2010). The impact of personal dispositions on information sensitivity, privacy concern and trust in disclosing health information online. *Decision Support Systems, 49*(2), 138-150.
- Barber, B. (1983). *The logic and limits of trust*. New Brunswick, NJ: Rutgers University Press.
- Belanger, F., & Carter, L. (2008). Trust and risk in e-government adoption. *Journal of Strategic Information Systems, 17*, 165–176.
- Beldad, A., van der Geest, T., de Jong, M., & Steehouder, M. (2012). Shall I tell you where I live and who I am? Factors influencing the behavioral intention to disclose personal data for online government transactions. *International journal of human-computer interaction, 28*(3), 163-177.
- Ben-Akiva M., & Lerman S.R. (1985). *Discrete Choice Analysis: Theory and Application to Travel Demand*. Cambridge, MA: MIT Press.
- Berendt, B., Gunther, O., & Spiekermann, S. (2005). Privacy in e-commerce: Stated preferences vs. actual behavior. *Communications of the ACM, 48*, 101–106.
- Bhattacharjee, A. (2012). *Social Science Research: Principles, Methods, and Practices*, 2nd edition. Tampa, FL: Creative Commons Attribution 3.0 License.

- Boudreau, M., Gefen D., & Straub D. (2001). Validation in IS Research: A State-of-the-Art Assessment. *MIS Quarterly* (25)1, 1-23.
- Boudreau, M.-C., Ariyachandra T., Gefen D., & Straub D. (2004). Validating IS Positivist Instrumentation: 1997-20 in M. E. Whitman and A. B. Woszczyński (Eds.) *The Handbook of Information Systems Research*, Hershey, PA USA: Idea Group Publishing, 15-26.
- Bryman, A., Collinson, D. L., Grint, K., Jackson, B., & Uhl-Bien, M. (2011). *The Sage handbook of leadership*. Thousand Oaks, CA: Sage.
- Castelluccio, M. (2015). THE BIGGEST GOVERNMENT HACK YET. *Strategic Finance*, 97(2), 79.
- Chang, Y. W., Hsu, P. Y., & Wu, Z. Y. (2015). Exploring managers' intention to use business intelligence: the role of motivations. *Behaviour & Information Technology*, 34(3), 273-285.
- Chang, Y. W., Hsu, P. Y., & Shiao, W. L. (2014). An empirical study of managers' usage intention in BI. *Cognition, Technology & Work*, 16(2), 247-258.
- Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS quarterly*, 36(4), 1165-1188.
- Cheung, M. K., & Lee, M. K. O. (2006). Understanding consumer trust in Internet shopping: A multidisciplinary approach. *Journal of the American Society for Information and Technology*, 57, 479-492.
- Chin W.W. (1998). *The partial least squares approach for structural equation modeling*. Lawrence Erlbaum Associates Publishers, NY, 295-336.
- Connolly, R., & Bannister, F. (2007). Consumer trust in electronic commerce: social and technical antecedents. *Proceedings of World Academy of Science, Engineering, and Technology*, 25, 386-395.
- Cormode, G., & Srivastava, D. (2009, June). Anonymized data: generation, models, usage. In *Proceedings of the 2009 ACM SIGMOD International Conference on Management of data* (1015-1018).
- Culnan, M.J., & Bies, R.J. (2003). Consumer privacy: Balancing economic and justice considerations. *Journal of Social Issues*, 59, 323-342.
- Das, T. K. & Teng, B. S. (2004). The risk-based view of trust: A conceptual framework. *Journal of Business and Psychology*, 19, 85-116.

- DeVellis, R. (2011). *Scale Development: Theory and applications* (3rd ed.). Thousand Oaks, CA: Sage.
- Dinev, T., & Hart, P. (2006). An extended privacy calculus model for ecommerce transactions. *Information Systems Research*, *17*, 61–80.
- Doney, P. M., Cannon, J. P., & Mullen, M. R. (1998). Understanding the influence of national culture on the development of trust. *Academy of Management Review*, *23*, 601–620.
- Dulewicz, V., & Higgs, M.J. (2004). Assessing leadership styles and organizational context. *Journal of Managerial Psychology*, *20*, 105-123.
- Figuroa Z. (2016). Time to Rethink Cybersecurity Reform: The OPM Data Breach and the Case for Centralized Cybersecurity Infrastructure. *Catholic University Journal of Law*, Retrieved from <http://scholarship.law.edu/jlt/vol24/iss2/7>
- Fink, A. (2003). *The survey handbook* (2nd ed.). Thousand Oaks: Sage.
- Fornell C., & Larcker D.F. (1981). Structural Equation Models with unobservable variables and measurement error: Algebra and Statistics. *Journal of Marketing Research*, *18*(3), 382-388.
- Fraj, E., & Martinez E. (2006). Influence of personality on ecological consumer behavior. *Journal of Consumer Behavior*, *5*, 167–181.
- Gallagher, S. (2015). Security: Why the "biggest government hack ever" got past the feds. Retrieved from <http://arstechnica.com/security/2015/06/why-the-biggestgovernment-hack-ever-got-past-opm-dhs-and-nsa/>
- Galvin, T., Gibbs, M., Sullivan, J., & Williams, C. (2014). Leadership competencies of project managers: An empirical study of emotional, intellectual, and managerial dimensions. *Journal of Economic Development, Management, IT, Finance, and Marketing*, *6*(1), 35-60.
- Gantz, J. F., Chute, C., Manfrediz, A., Minton, S., Reinsel, D., Schlichting, W., & Toncheva, A. (2008). *The diverse and exploding digital universe* [IDC white paper]. Framingham, MA: International Data Corporation.
- Gay, L.R., Mills, G.E., & Airasian, P.W. (2012). *Educational research: Competencies for analysis and application* (10th ed). Boston, MA: Pearson.
- Gertz, B. (2016). Cybercom: OPM Hack Highlights China Big Data Spying, Pentagon moves to protect records from future attacks. Retrieved from <http://freebeacon.com/national-security/cybercom-opm-hack-highlights-china-big-data-spying/>

- Gefen, D., Karahanna, E., & Straub, D.W. (2003). Trust and TAM in online shopping: an integrated model. *MIS Quarterly*, 27(1), 51–90.
- Gefen, D., Straub, D., & Boudreau, M. C. (2000). Structural equation modeling and regression: Guidelines for research practice. *Communications of the association for information systems*, 4(1), 7.
- GEMALTO, INDEX (2015). FIRST HALF REVIEW: FINDINGS FROM THE BREACH LEVEL INDEX 3 (2015). Retrieved from: <http://bit.ly/244WHpj>
- Gregor, S. (2006). The Nature of Theory in Information Systems. *MIS Quarterly* (30)3, 611-642.
- Hair J.J., Anderson R., Tatham R., & Black W. (1995). *Multivariate data analysis with readings*. Upper Saddle River: NJ, Prentice-Hall Inc.
- Hair, J.F., Hult G.T.M., Ringle, C.M., & Sarstedt, M. (2017). *A primer on partial least squares structural equation modeling (PLS-SEM)*. Thousand Oaks, CA: Sage.
- Hair, J.F., Ringle, C.M., & Sarstedt, M. (2011). PLS-SEM: Indeed, a silver bullet, *The Journal of Marketing Theory and Practice*, 19(2), 139-152.
- Hinton, P.R. (2008). *Statistics explained* (2nd ed.). East Sussex, UK: Routledge.
- Homans, G. (1958). Social behaviour as exchange. *The American Journal of Sociology*, 63, 597–606.
- Homans, G. (1961). *Social behaviour: Its elementary forms*. London, UK: Routledge & Kegan Paul.
- Hoyle, R.H. (1995). *The structural equation modeling approach: Basic concepts and fundamental issues*. Thousand Oaks, CA: Sage.
- Howe, D., Costanzo, M., Fey, P., Gojobori, T., Hannick, L., Hide, W., & Rhee, S. Y. (2008). Big data: The future of biocuration. *Nature*, 455(7209), 47-50.
- Hu, L.T., Bentler, P.M. (1998). Fit Indices in covariance structure modeling: Sensitivity to under parameterized model misspecification. *Psychological Methods*, 3, 424-453.
- Hui, K.-L., Teo, H.H., & Lee S.T. (2007). The value of privacy assurance: an exploratory field experiment. *MIS Quarterly* 31(1), 19–33.
- 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.

- Kim, J. (1999). Causation. *The Cambridge Dictionary of Philosophy* (2nd ed.), R. Audi (ed.), Cambridge UK, 125-127.
- Kim, D. J., Steinfield, C., & Lai, Y. J. (2008). Revisiting the role of web assurance seals in business-to-consumer electronic commerce. *Decision Support Systems*, 44(4), 1000-1015.
- Kristen, B. D., Dyer, G., Hoopes, C., & Harris, S. (2004). Toward a model of effective knowledge management and directions for future research: Culture, leadership, and CKOs. *Journal of Leadership & Organizational Studies*, 10(4), 26-43.
- Larsson, G., & Eid, J. (2012). An idea paper on leadership theory integration. *Management Research Review*, 35(3), 177-191.
- Laufer, R. S., & Wolfe, M. (1977). Privacy as a concept and a social issue: A multidimensional developmental theory. *Journal of Social Issues*, 33(3), 22-42.
- Lee, A. S. (2001). Editorial, *MIS Quarterly* 25(1), iii-vii.
- Lee, H., Park, J., & Lee, J. W. (2013). Role of leadership competencies and team social capital in IT Services. *The Journal of Computer Information Systems*, 53(4), 1-11.
- Levy, Y. (2008). An empirical development of critical value factors (CVF) of online learning activities: An application of activity theory and cognitive value theory. *Computers & Education*, 51(4), 1664-1675.
- Levy Y., & Ellis T. (2006). A systems approach to conduct an effective literature review in support of information systems research. *Informing Science Journal*. 9, 181-208.
- Levy, Y., & Green, B. D. (2009). An empirical study of computer self-efficacy and the technology acceptance model in the military: A case of a U.S. navy combat information system. *Journal of Organizational and End User Computing*, 21(3), 11-13.
- Luce R.D. (1959). *Individual Choice Behavior: A Theoretical Analysis*. NY: Wiley.
- Luck J., Chang C., Brown E.R., Lumpkin J. (2006). Using local health information to promote public health, *Health Affairs*, 25(4), 979-991.
- Malhotra, N.K., Kim S.S., & Agarwal J. (2004). Internet users' information privacy concerns (IUIPC): the construct, the scale, and a causal model. *Information Systems Research*, 15(4), 336-355.
- McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of personality and social psychology*, 52(1), 81.
- McCrae, R. R., & Costa, P. T. (1991). Adding Liebe und Arbeit: The full five-factor model and well-being. *Personality and social psychology bulletin*, 17(2), 227-232.

- McFadden D.L. (2001). Economic choices. *American Economic Review* 91(3), 351–378.
- Mertler, C., & Vannatta, R. (2013). *Advanced and multivariate statistical methods: Practical application and interpretation* (Fifth ed.). Glendale, CA: Pyrczak Publishing.
- Müller, R., & Turner R. (2010). Leadership competency profiles of successful project managers. *International Journal of Project Management*, 28(5), 437-448.
- Newton, P., & Shaw, S. (2014). *Validity in educational and psychological assessment*. Thousand Oaks, CA: Sage.
- Nunnally, J. C. (1978). *Psychometric Theory*, 2nd edition. NY, USA: McGraw-Hill.
- Olivero, N., & Lunt, P. (2004). Privacy versus willingness to disclose in e-commerce exchanges: The effect of risk awareness on the relative role of trust and control. *Journal of Economic Psychology*, 25, 243–262.
- Osatuyi, B. (2015). Personality traits and information privacy concern on social media platforms. *The Journal of Computer Information Systems*, 55(4), 11.
- Perera, C., Ranjan, R., Wang, L., Khan, S. U., & Zomaya, A. Y. (2015). Big Data Privacy in the Internet of Things Era. *IT Professional*, 17(3), 32-39.
- Quinn, R. E., Faerman, S. R., Thompson, M. P., & McGrath, M. R. (2000). *Becoming a master manager: A competency framework*. New York: Wiley.
- Quinn, R. E., & Rohrbaugh, J. (1981). A competing values approach to organizational effectiveness. *Public Productivity Review*, 5(2), 122.
- Reid, M., & Levy, Y. (2008). Integrating trust and computer self-efficacy with TAM: An empirical assessment of customers' acceptance of banking information systems (BIS) in Jamaica. *Journal of Internet Banking and Commerce*, 12(3), 2008-12.
- Resnick, M. L., & Montania, R. (2003). Perceptions of customer service, information privacy, and product quality from semiotic design features in an online web store. *International journal of human-computer interaction*, 16(2), 211-234.
- Sanger D. E. & Davis J.H. (2015). Hacking Linked to China Exposes Millions of U.S. Workers, *N.Y. TIMES*. Retrieved from: <http://nyti.ms/1XsoDU6>
- Shin, D.-H. (2010). The effects of trust, security and privacy in social networking: A security-based approach to understand the pattern of adoption. *Interacting with Computers*, 22(5), 428-438.

- Shang, H., & Yu, W. (2013). Assessing Chinese managerial competencies from different perspectives. *Social Behavior and Personality*, 41(9), 1469-1485.
- Straub, D. W. (1989). Validating Instruments in MIS Research. *MIS Quarterly*, 13(2), 147-169.
- Straub, D. W. (1990). Effective IS Security: An Empirical Study. *Information Systems Research*, (1)3, 255-276.
- Straub, D., Boudreau M.-C., & Gefen D. (2004). Validation Guidelines for IS Positivist Research. *Communications of the Association for Information Systems* (14), 380-426.
- Terrell, S. R. (2016). Writing a proposal for your dissertation: Guidelines and Examples. New York: Guilford Press.
- USC Annenberg School of Communication-Center for the Digital Future (2004). *Surveying the digital future-year four*. Retrieved from http://www.digitalcenter.org/wpcontent/uploads/2013/02/2004_digital_future_report-year4.pdf
- Van Der Aalst, W.M.P. (2012). Process Mining, *Comm. ACM*, 55(8), 76–83.
- Vera-Baquero, A., Colomo-Palacios, R., & Molloy, O. (2013). Business process analytics using a big data approach. *IT Professional*, 15(6), 29-35.
- What's the hold up? A review of security clearance backlog and reciprocity issues plaguing today's government and private sector workforce. Hearing before the committee on Government Reform, House of Representatives, 108th Cong. 2 (2004).*
- Westin, A.F. (2003). Social and political dimensions of privacy. *Journal of Social Issues*, 59(2), 431–453.
- Wang, D., & Hsieh, C. (2013). The effect of authentic leadership on employee trust and employee engagement. *Social Behavior and Personality*, 41(4), 613-624.
- Wu, X., Zhu, X., Wu, G. Q., & Ding, W. (2014). Data mining with big data. *Knowledge and Data Engineering, IEEE Transactions on*, 26(1), 97-107.
- Xie, E., Teo. H. H., & Wan, W. (2006). Volunteering personal information on the internet: Effects of reputation, privacy notices, and rewards on online consumer behavior. *Marketing Letters*, 17, 61–74.
- Xu, H., Teo, H. H., & Tan, B. (2006). Information privacy in the digital era: an exploratory research framework. *AMCIS 2006 Proceedings*, 120.

Zimmer, J. C., Aarsal, R., Al-Marzouq, M., Moore, D., & Grover, V. (2010). Knowing your customers: Using a reciprocal relationship to enhance voluntary information disclosure. *Decision Support Systems, 48*, 395–406.