Development of a Social Engineering eXposure Index (SEXI) using Open Source Personal Information

by

William Shawn Wilkerson

A dissertation proposal 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

2018

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

Development of a Social Engineering eXposure Index (SEXI) using Open Source Personal Information

By
William Shawn Wilkerson
April 2018

Millions of people willingly expose their lives via Internet technologies every day, and even the ones who refrain from the use of the Internet find themselves exposed through data breaches. Trillions of private information records are exposed through the Internet. Marketers gather personal preferences to influence shopping behavior. Providers gather personal information to provide enhanced services, and underground hacker networks provide repositories of immense data sets. Few users of Internet technologies have considered where their information is going or who has access to it. Even fewer are aware of how decisions made in their own lives expose significant pieces of information, which can be used by cyber hackers to harm the very organizations with whom they are affiliated. While this threat can affect any person holding any position at an organization, upper management poses a significantly higher risk due to their level of access to critical data and finances targeted by cybercrime

The goal of this research is to develop and validate a Social Engineering eXposure Index (SEXI) using open source personal information to assist in identifying and classifying social engineering vulnerabilities. This study will include a mixed methods approach combining an expert panel using the Delphi method, developmental research, and quantitative data collection. The expert panel will categorize and assess information privacy components into three identifiability groups, subsequently used to develop an algorithm that will form the basis for a Social Engineering eXposure Index (SEXI). Validation of the algorithm will use open source personal information found on the Internet for 50 executives of Fortune 500 organizations and 50 Hollywood celebrities. The exposure of each executive and persona will be quantified and the collected data will be evaluated, analyzed, and presented.

Table of Contents

List of Tables

Tables

List of Figures

Figures

Chapter 1

Introduction

Background

Cybersecurity issues are as ubiquitous as the Internet itself and can be observed in social engineering victims ranging from a child targeted by pedophiles to the Director of the U.S. Central Intelligence Agency (CIA) (Federal Bureau of Investigation, 2015b; Franceschi-Bicchierai, 2015). Cyber attackers can be anyone from teenagers to foreign government actors (Federal Bureau of Investigation, 2016; Kopan, 2015). Objectives are as diverse as embarrassment to murder, but usually takes the form of fraud with the loss for United States (U.S.) organizations averaging over $100,000 per incident in 2013 (Federal Bureau of Investigation, 2015a; Mouton, Leenen, & Venter, 2016).

Open source is defined herein as “publicly available print and digital/electronic data from unclassified, non-secret, and ‘grey literature’ sources,” not requiring credentials or special access, including data available through breaches, leaks, etc. (Fleisher, 2008, p. 853).Marketing (Culnan & Bies, 2003; Moon, 2000), personalization (Chellappa & Sin, 2005; Culnan, 1993; Kim & Pan, 2006), e-commerce (Dinev & Hart, 2006; Feijóo, Gómez-Barroso, & Voigt, 2014), self-surveillance (Kang, Shilton, Estrin, & Burke, 2011), surveys, contests, order forms, registrations (Federal Trade Commission, 2000), and social media (Acquisti, Brandimarte, & Loewenstein, 2015; Karaduman, 2013; Peer & Acquisti, 2016) are just a few ubiquitous open source repositories. Additionally, grey

literature is typically comprised of less-than-formal publications such as Websites and unpublished papers (Fleisher, 2008).

The exponential growth of personal information available online via open source technologies has exposed unsuspecting prey for social engineers to attack relentlessly (Acquisti et al., 2015; Mitnick & Simon, 2002). The open source personal information (OSPI) provided by social media and other platforms facilitate many successful SE attacks on potential victims (Krishnamurthy & Wills, 2009; Maynard, Greenwood, Roberts, Windsor, & Bontcheva, 2015). E-mail is another tool used to gain OSPI by disguising its origin and purpose, usually to appear as a trusted entity known by the intended victim (Almomani, Gupta, Atawneh, Meulenberg, & Almomani, 2013; Federal Bureau of Investigation, 2015a; Mouton et al., 2016). The increased availability of OSPI furnishes social engineers with a larger number of victims, with no end in sight (Acquisti et al., 2015; Mitnick & Simon, 2002).

Problem Statement

The research problem that this study will address is the proliferation of social engineering (SE) attacks due to publicly available OSPI (Heartfield & Loukas, 2015; Maynard et al., 2015; Mitnick & Simon, 2002). SE “is a combination of techniques used to manipulate victims into divulging confidential information or performing actions that compromise security” (Luo, Brody, Seazzu, & Burd, 2013, p. 2; Mitnick & Simon, 2002). Social engineers use deception and often use roleplaying to represent someone their intended targets are more susceptible to (Orgill, Romney, Bailey, & Orgill, 2004). Additionally, the use of pretense and persuasion is often noted in successful social

engineering attacks (Heartfield & Loukas, 2015; Mitnick & Simon, 2002). This behavior is consistent with the Theory of Mind (TOM) where an actor attempts to persuade another individual through pretense and deception, while remaining within the confines of the representation held by the other individual. TOM is defined as “the individual imputes mental states to himself and to others” (Premack & Woodruff, 1978, p. 515). A help desk may hold a representation of aiding those who request it. Several employees may hold a representation that an auditor is part of the IT department, if they are observing someone dressed in a manner acceptable for the role and who is appearing to perform functions that represent the expected activity (Krombholz, Hobel, Huber, & Weippl, 2013; Orgill et al., 2004). Social engineers are able to pretend and persuade even experts into behaving favorably for the attacker, even when they suspect something is wrong and are mandated as well as trained to take appropriate defensive action (Allen, 2006; Heartfield & Loukas, 2015).

Prior research has shown the information being used to execute SE attacks typically originates at the target or those closely associated with them (Heartfield & Loukas, 2015; Junger, Montoya, & Overink, 2017; Luo et al., 2013). Studies have also shown a significant increase of personal information exposed on social networking sites and the overall willingness to provide personal content by Americans (Acquisti et al., 2015; Boyd & Ellison, 2007; Hong & Thong, 2013). Olmstead and Smith (2017) stated that 64% of Americans had been exposed via a data breach.

According to Solove (2006), “Exposure involves the exposing to others of certain physical and emotional attributes about a person” (p. 533). Some studies suggested that

people willingly expose private information in exchange for content gratification, even after adjusting their settings for what they perceived as increased privacy (Sutanto, Palme, Tan, & Phang, 2013). Ku, Chen, and Zhang (2013) found that a positive association exists between the gratification of using social networking sites and the intention for continued usage. The availability of OSPI has grown substantially over recent years and looks to have exponential growth as more people gain access to the Web and service providers continually introduce innovative mechanisms for self-disclosure (Acquisti et al., 2015).

When Facebook, a social network site, first went public, it targeted the needs of business users to facilitate professional relationships and was later expanded to provide any user the ability to share far more personal information (Acquisti et al., 2015). Initially, the majority of information posted by Facebook users was related to business efforts providing very few self-identifying descriptive items, while also restricting the scope of people having access to the shared information (Acquisti et al., 2015; Pew Research Center, 2013). By 2014, the basic and extended profiles of a user's personally identifiable information (PII) were potentially shareable to anyone on the Internet with access to the original Facebook postings (Acquisti et al., 2015). Examples of PII may include name, email, postal address, phone or fax number (Federal Trade Commission, 2000). This availability of OSPI allows potential hackers to glean necessary information to successfully social engineer an exposed target via a myriad of attack vectors (Heartfield & Loukas, 2015; Luo et al., 2013). Acquisti et al. (2015) found that the number of Facebook categories of exposure increased from three (networks, genders, & names) to eight (networks, genders, names, friends, basic profile, extended profile, likes,

& pictures) between 2005 and 2014 beginning with text and progressively expanding to including live video content. Twitter microdata is another source of OSPI allowing indirect access to a user's identity (Singh, Bansal, & Sofat, 2014).

The literature typically describes PII as including any content that has the potential to identify an individual (McCallister, Grance, & Scarfone, 2010). Schwartz and Solove (2011) suggested another category of information: personally distinguishable information (PDI). They argue that PDI will *definitively* identify someone, whereas most PII only has the *potential* of identifying a specific individual (Schwartz & Solove, 2011). Additionally, a third category of PII is suggested, personally unidentifiable information (PUI), which has no chance to identify an individual on its own (McCallister et al., 2010; Schwartz & Solove, 2011). OSPI provides access to PDI, PII, and PUI making up the three primary categories of personal information, with PDI having the highest level of exposure, PII exhibiting the potential of exposure, and PUI offering no exposure by itself, however, combined with the prior two categories can add to the overall exposure of an individual (McCallister et al., 2010; Schwartz & Solove, 2011). PDI is any information which specifically distinguishes the individual on its own, slightly differing from PII in that the potential of exposure is absolute (Chellappa & Sin, 2005; Schwartz & Solove, 2011). PDI may include a digital photograph, video, social security number, Global Positioning System (GPS), passport number, credit card number, security clearance, bank account number, biometric data, date with the place of birth, mother's maiden name, criminal background, medical record, financial record, and educational transcript (42 U.S.C. § 200.82). PUI is any information which cannot solely be used to identify an individual (Chellappa & Sin, 2005; Schwartz & Solove, 2011). PUI may include age, date

of birth, gender, education, hobby, income, interest, the name of the software used, occupation, type of hardware in configuration, and Zip Code (Chellappa & Sin, 2005; Federal Trade Commission, 2000).

The threat to organizations with leaders having their PDI, PII, and PUI available via OSPI is easily translated into risk assessments. According to the U.S. Federal Bureau of Investigation (FBI) (2015a), Business Email Compromise (BEC) affected over 7000 organizations within the U.S. approaching $800 million in losses between October 2013 and August 2015. A substantial increase of over 270% in the number of BEC cases occurred during the opening months of 2015 indicating SE attacks are dramatically on the rise (Federal Bureau of Investigation, 2015a).

Phishing is another attack vector of SE, whereby the target is baited with a fake copy of a Web page or Website to solicit sensitive information or to inject malware onto the victim's computer or mobile devices (Heartfield & Loukas, 2015). Neupane, Rahman, Saxena, and Hirshfield (2015) conducted phishing research and found that the longer an individual looked at the content on a fake Web page, during each 10-second trial, the more likely they would accept it as being authentic. They also discovered the possibility of a successful phishing event significantly increased if the participant was distracted or sleep-deprived (Neupane et al., 2015). The growing availability of OSPI is providing the content used for successful SE attacks, and in the creation of effective phishing campaigns (Heartfield & Loukas, 2015; Neupane et al., 2015).

In 1994, a French social engineer called the FBI in Washington, D.C. and successfully persuaded someone to expose the information required to make phone calls

at the agency's expense (Allen, 2006; Schneier, 2000). In another example, the FBI described how the practices of a company facilitated a $737,000 transfer to an unauthorized recipient in China (Federal Bureau of Investigation, 2015a). In 2016, state-sponsored Iranians performed distributed denial of service (DDoS) attacks on U.S. financial institutions blocking hundreds of thousands of customers from accessing their bank accounts (Federal Bureau of Investigation, 2016). In 2018, 9 Iranians were indicted for the theft of over "31 terabytes of documents and data from more than 140 American universities, 30 American companies, five American government agencies", as well as "compromised approximately 8,000 professor email accounts across 144 U.S.-based universities" (U.S. Department of Justice, 2018, pp. 1,2) . Acquisti and Gross (2009) described the simplicity of predicting Social Security Numbers and the dangers of mass identity theft due to weaknesses in the U.S. identifier system.

To illustrate the effectiveness of SE against organizations, Orgill et al. (2004) described an unannounced security audit where 19 out 32 people gave their password to an unknown person walking through the facility with a name badge retrieved from a desk where an employee left it. While seven people supplied the username and password for another person's account with access elevated beyond their own, only four of the 32 employees asked for the auditor's identification (Orgill et al., 2004). Two days later the auditor returned and was able to find multiple company credit cards and a master key to the building within 30 seconds of beginning a general search near an executive's office (Orgill et al., 2004). Orgill et al. (2004) found that even organizations with a high awareness of data security and requirements to follow privacy standards are vulnerable to SE due to exposure.

The October 2015 BEC attack on the Director of the CIA provides an example where OSPI was used to gain access to a private email account. Teenagers were able to gather data from OSPI located across multiple online accounts belonging to the CIA Director, and use the information to pretext, another SE attack vector, customer service representatives via telephonic communication into exposing additional personal details (Franceschi-Bicchierai, 2015). Using the combined data, the attackers obtained the necessary information to access the personal email account of the CIA Director. Subsequently, the attackers released the PII of many of the CIA Director's associates and subordinates to WikiLeaks (Franceschi-Bicchierai, 2015; Kopan, 2015). The availability of OSPI allowed the successful targeting of the CIA Director by a group of high school students having no formal information security training (Franceschi-Bicchierai, 2015). Figure 1 represents the SE attack used on the CIA Director that may have been repeatable as the perpetrators had possession of the personal information of agents, contractors, and government personnel stored within the compromised e-mail account. The collected information could also be used in any number of other SE attacks as well.

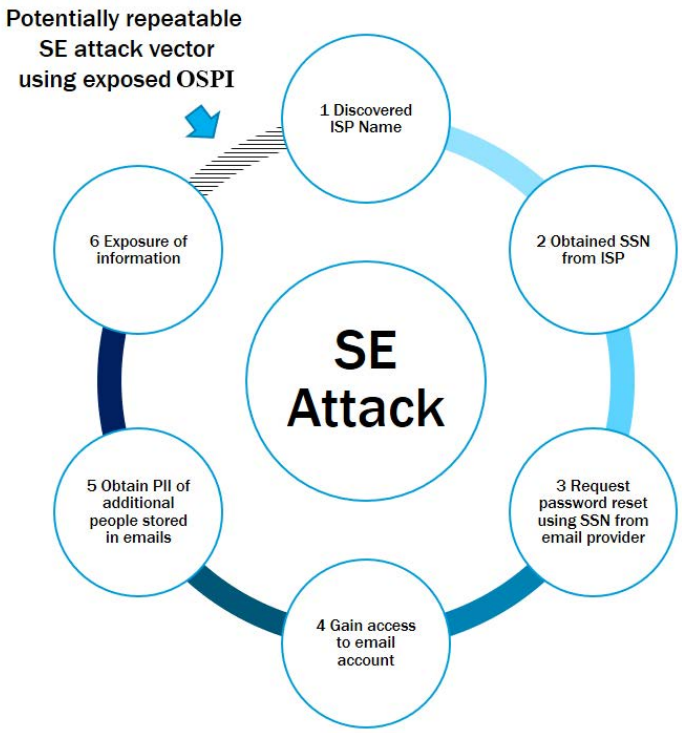

Figure 1: SE attack used against the CIA Director in 2015

Heartfield and Loukas (2015) found that familiarity with content, such as a logo, provides a substantial increase in employees mistaking a SE attack for an official request. Additionally, Acquisti et al. (2015) found that OSPI is readily accessible and increasingly available. According to the FBI, BEC attacks and the financial loss associated with them have significantly increased (Federal Bureau of Investigation, 2015a). The growth of BEC, SE, and OSPI indicate the current cybersecurity defense methodologies may not be sufficient to protect individuals or organizations from SE attacks (Tetri & Vuorinen, 2013). Thus, it appears additional research is warranted to assess and classify social engineering exposure of individuals, especially top executives of large organizations and key strategic personnel.

Dissertation Goal

The goal of this research is to develop and validate a Social Engineering eXposure Index (SEXI) using open source personal information (OSPI) to assist in identifying and classifying SE vulnerabilities. The index will be validated on 50 executives of Fortune 500 companies and 50 Hollywood personas. SEXI is proposed to provide a rating of the exposure to SE due to OSPI. The need for this research is demonstrated by the work of Mitnick and Simon (2002), Tetri and Vuorinen (2013), Heartfield and Loukas (2015), as well as Mouton et al. (2016) that acknowledged the progressive expansion of SE attack vectors, the lack of a predictive threat system, the availability of OSPI which circumvent organizational cybersecurity technologies, and the dearth of data on information gathering techniques for the successful execution of prior SE attacks.

Mouton et al. (2016) described the difficulty in performing SE research due to the lack of information provided in news articles, especially the method of attack and where the information was gathered to prosecute the intended target. Despite proposed SE attack templates, the effect of OSPI on target exposure is not a well-understood phenomenon, making it a viable and challenging research problem (Mouton et al., 2016). Mouton et al. (2016) reinforced the sentiment found by Mitnick and Simon (2002) that the human component is the weakest link for organizational security, as it serves both as a bypass to security technologies and as the fountain of information by which SE attacks occur. Additionally, Mouton et al. (2016) suggested that SE research is still in its infancy despite the rapid growth of information security research.

Heartfield and Loukas (2015) described the ineffectiveness of studying "semantic attacks" as it occurs after the damage is done and may be limited by a lens focused on a singular attack vector (p. 31). Of significance, for this proposed dissertation study, is the call for a prediction mechanism by Heartfield and Loukas (2015) for determining exposure in real time that is automatically updated with a rapid response window. The availability of OSPI used for SE attacks can also serve to determine SE exposure (Heartfield & Loukas, 2015; Tetri & Vuorinen, 2013). Armed with a SE prediction mechanism, executives can take an offensive stance in organizational security risk mitigation and likewise monitor the overall exposure of the organization in real-time by evaluating the availability of OSPI of key personnel – including themselves (Mouton et al., 2016).

This study builds on previous research by Bélanger and Crossler (2011), Tetri and Vuorinen (2013), Acquisti et al. (2015), as well as Heartfield and Loukas (2015). Bélanger and Crossler (2011) called for "the development of more (and easier to use) privacy protection tools for individuals, groups, organizations, and society" (p. 1035). Acquisti et al. (2015) described the exponential increase of OSPI via social networking sites while Tetri and Vuorinen (2013) found that its availability enabled as well as facilitated SE attackers across a broad spectrum of attack vectors. Current research and defense mechanisms tend to focus on a single attack vector or technique, thereby drastically limiting their actual benefit or significance to the security strategy (Tetri & Vuorinen, 2013). Specifically, Tetri and Vuorinen (2013) suggested that research might include an evaluation of where the information was obtained by attackers as well as how the SE attack vectors were possible in the first place (p. 1020). Heartfield and Loukas

(2015) called for the development of a formal framework that could profile the exposure of users to SE attacks. Schwartz and Solove (2011) argued that privacy must move beyond an ineffective legal system split between standard and rule towards an understanding of "identification in terms of risk level (p. 1979)" and realize "a standards-based approach can be made operational and predictable" (p. 1884). Ohm (2010) views the entire PII concept as broken and believes almost any information can be traced as well as used to identify an individual. This proposed study intends to develop and validate, using Subject Matter Experts (SMEs), a prototype tool to aid organizations is SE mitigation and an index of exposure to SE due to the availability of OSPI for 100 individuals and corporate executives.

While there have been many discussions in the literature concerning personal information, there is very little in the quantification and grouping of the components. The first specific goal of this research study is to gather the SME-approved components for an index of SE exposure by eliciting qualitative and quantitative feedback on personal information. The second specific goal of this research study is to assign categories to personal information components based on exposure. The third specific goal of this research study is to develop and validate, using SMEs, the components and hierarchical weights for SEXI via a Delphi method. The fourth specific goal of this research study is to apply the SEXI method to measure the OSPI exposure of 50 executives of Fortune 500 organizations and 50 Hollywood celebrities. The fifth specific goal of this research study is to assess and statistically test for significant mean differences of the SEXI of 100 individuals based on demographical indicators of age, gender, income, marital status, estimated worth, industry, organizational position, philanthropic contributions, and prior

military/police experience. The sixth specific goal of this research study is to compare the SEXI results from the set of US executives to those of Hollywood personas in an effort to uncover which group is more vulnerable to SE attack from an OSPI exposure perspective.

Research Questions

The main research question (RQ) that this study will address is: What are the expert-approved required components comprising an index of exposure to social engineering attacks due to OSPI? The specific research questions that this study will address are:

RQ1: What are the specific SMEs approved set of personal information components for an index of SE exposure?

RQ2: What are the specific SMEs approved categories for the identified set of personal information components?

RQ3: What are the specific SMEs identified weights of the personal information components and categories that enable a validated hierarchical aggregation to the Social Engineering eXposure Index (SEXI) benchmarking index?

RQ4: How are 100 individuals assessed and classified by SEXI using OSPI?

RQ5: Are there any statistically significant mean differences of SEXI based on demographical indicators of age, gender, income, marital status, estimated worth, industry, organizational position, philanthropic contributions, and prior military/police experience?

RQ6: Do SEXI results from the set of US executives and Hollywood personas indicate one group being more vulnerable to SE attack from their OSPI exposure perspective?

SE attacks are on the rise, and the OSPI used to perpetrate these crimes is far too readily available (Acquisti et al., 2015; Federal Bureau of Investigation, 2015a). Tetri and Vuorinen (2013) conducted a literature review of 40 journal articles and found them primarily explorative and descriptive with very few SE studies being empirical, thereby validating a knowledge gap in the literature. The merit of developing an exposure index is that it can assist in the prediction of the SE exposure of targets, the content of potential attacks, and possible attack vectors which current security structures may fail to detect or provide (Heartfield & Loukas, 2015; Mouton et al., 2016). Prior research indicates people readily expose themselves online (Acquisti et al., 2015; Pew Research Center, 2013; Smith, 2015) and that organizations can end up paying for their executives' in a myriad of ways (Federal Bureau of Investigation, 2015a; Mouton et al., 2016).

Relevance and Significance

Relevance

The privacy chain, defined as the flow of PII communication between two endpoints (Wilkerson, Levy, Kiper, & Snyder, 2017), appears to have no lack of supply (Mitnick & Simon, 2002; Tetri & Vuorinen, 2013) or demand (Federal Bureau of Investigation, 2012; Jasper, 2017). People continue to freely share PII even though they are aware of the consequences of doing so (Acquisti et al., 2015; Olmstead & Smith, 2017). The literature provides troubling insight into the primary creator of PII, the

subjects themselves. People continue self-disclosure even though 64% of Americans have experienced data breaches (Olmstead & Smith, 2017). Since 2015, the number of Facebook users has increased by 7%, bringing the total to 79% of Internet users using the service – 68% of American adults (Greenwood, Perrin, & Duggan, 2016). The alarming rate of PII released and subsequently available via OSPI is a continual threat to organizations (Mouton et al., 2016). Case in point, the successful attack on the Director of the CIA demonstrates how OSPI provided attackers access to a private email account of a key figure, which contained and provided PII of many CIA agents (Franceschi-Bicchierai, 2015; Kopan, 2015). Figure 2 provides the number of breaches and the number of records from 2005-2017. It should be noted that no correlation exists as to the number of data breaches and the number of records. A single data breach can exceed billions of records (Green, 2017), while others may contain no records at all (Privacy Rights Clearinghouse, 2018).

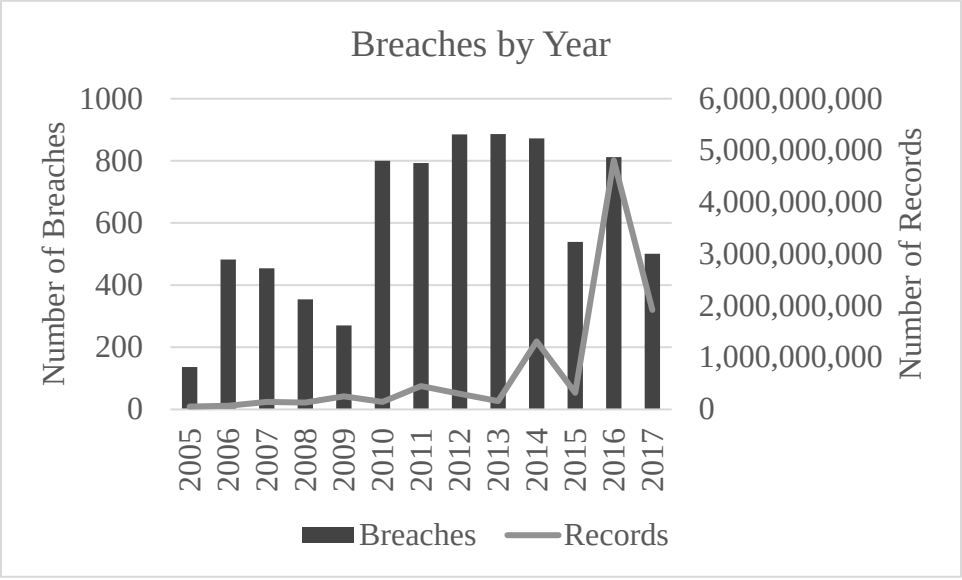

Figure 2: Reported data breaches from 2005 thru 2017. Adapted from “Data Breaches,” by the Privacy Rights Clearing House, 2017. Used with the permission of the Privacy Rights Clearinghouse, under a Creative Commons Attribution NonCommercial-ShareAlike 4.0 (CC BY-NC-SA 4.0).

The literature indicates that SE success often depends on the availability of PII (Junger et al., 2017). Combined with the exponential growth of PII available via open source technologies, an onslaught of effective SE attacks continues to plague organizations with a snowballing relentlessness (Acquisti et al., 2015; Bélanger & Crossler, 2011). In response, prior literature has assuaged the demand for security policies, training, and awareness efforts, but has shown limited effectiveness in curbing the crushing weight of potential PII-related threats (Mitnick & Simon, 2002; Mouton et al., 2016; Tetri & Vuorinen, 2013). Junger et al. (2017) found that people are typically ill prepared to make PII-related decisions, even with training and warnings. Additionally, research has shown that a direct connection and potential threat exists with the way people perceive the significance of PII between virtual and physical worlds (Junger et al., 2017).

SE attacks on organizations occur without the benefit of knowing what PII is available or from where the attack will come (Tetri & Vuorinen, 2013). In effect, organizations are largely ineffective in staving off SE attacks due to current security structures failing to predict PII exposure of organizational targets, the content of potential attacks, or possible attack vectors (Heartfield & Loukas, 2015). Given the documented exponential increase of the availability of PII, the relevance of this study is considerable.

Significance

The significance is the dramatic increase in the availability of OSPI due to the willingness of people to share on social networks and other media as well as trillions of records compromised via data breaches. The existence of hacker undergrounds where personal information and SE attack vectors are shared increases the exposure. Prior literature has documented the existence of OSPI as the precursor for many successful social engineering attacks (Heartfield & Loukas, 2015) . The significance is the development and validation of a SEXI using OSPI to assist in identifying and classifying SE exposure. Since privacy is highly subjective (Acquisti et al., 2015; Acquisti, Taylor, & Wagman, 2016; Moon, 2000) and traditionally understood understood through context (Heurix, Zimmermann, Neubauer, & Fenz, 2015; Hong & Thong, 2013) prior literature has called for a tool to serve as a predictor and determinant for potential SE attacks (Heartfield & Loukas, 2015; Mohaisen, Al-Ibrahim, Kamhoua, Kwiat, & Njilla, 2017) seeking the specificity of available information (Tetri & Vuorinen, 2013). The security training and policies implemented by organizations rely heavily on people (Mouton et al., 2016; Tetri & Vuorinen, 2013), which the literature indicates is the weakest defense point (Mitnick & Simon, 2002), the easiest to compromise (Neupane et al., 2015), and who associate their virtual openness to the current environment (Junger et al., 2017). While organizations implement security policies and training (Mouton et al., 2016), research has found that warnings issued to users may actually increase exposure of personal information (Junger et al., 2017). Zhang, Wu, Kang, Go, and Sundar (2014) found that even though users perceived a heightened online security threat, they tended to expose

even more personal information. Figure 3 provides the significance of unintended disclosures of datasets, which has grown during recent years.

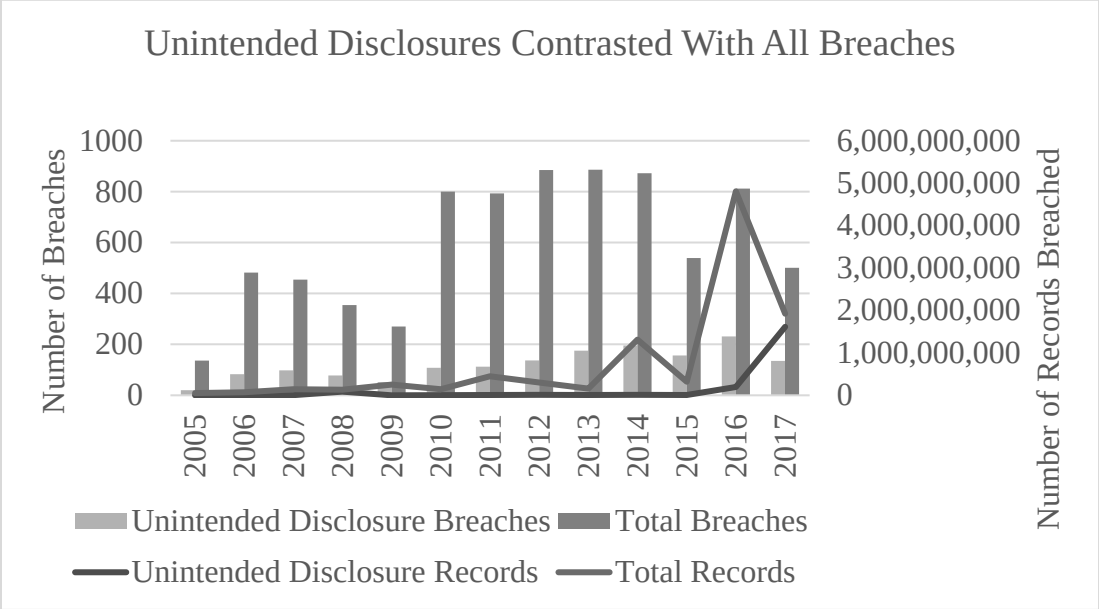

Figure 3: Unintended disclosures contrasted with all reported data breaches. Unintended disclosures accounted for nearly half of 2017's data breaches. Adapted from "Data Breaches," by the Privacy Rights Clearing House, 2017. Used with the permission of the Privacy Rights Clearinghouse, under a Creative Commons Attribution NonCommercial-ShareAlike 4.0 (CC BY-NC-SA 4.0).

Research indicates that the majority of users do not read or understand privacy policies in their lives, because they appear unwilling to put forth any significant effort in managing the privacy they value (Acquisti et al., 2015; Hong & Thong, 2013). These same people make up the cyber defense of the organizations (Mouton et al., 2016). During late 2016, Yahoo announced one billion customer records had been stolen (Green,

2017). The Privacy Rights Clearinghouse (2018) has logged almost 10 trillion breached data records since 2005, with 18% (1.8 trillion) occurring in the first 10 months of 2017. According to Jasper (2017), often data from breaches are shared on the hacker underground marketplace within 72 hours, facilitating further successful attacks using the information. Public releases of stolen information are not uncommon, as is the case with the WikiLeaks release of CIA personnel PII instantly transforming the PII into OSPI (Franceschi-Bicchierai, 2015; Kopan, 2015). Public release of protected information serves as the foundation for SE attackers to mount attacks through unknown vectors using a massive amount accurate data to orchestrate a cacophony of SE attacks (Mouton et al., 2016; Tetri & Vuorinen, 2013). Given the documented increase in PII exposed via data breaches and the continual avalanche of successful SE attacks using OSPI, the significance of this proposed study is substantial. Armed with a SE prediction mechanism, executives can take an offensive stance in organizational security risk mitigation and likewise monitor the overall exposure of the organization in real-time by evaluating the availability of OSPI of key personnel – including themselves (Mouton et al., 2016).

Barriers and Issues

Limited discernable empirical literature appears to exist regarding exposure, personal information, and social engineering. In addition, it appears that there is limited literature with regards to exposure related to open source personal information. Hence, limited predictive literature indicates how to measure the exposure of individuals due to the availability of open source personal information. To resolve this, a new instrument is to be developed using Schwartz and Solove‘s (2011) privacy categories as well as

McCallister et al.'s (2010) privacy-related descriptions and definitions used to protect the confidentiality of PII. Reliability for the internal consistency of intercorrelated items of the SEXI instrument is one of the barriers that requires overcoming.

One potential barrier for this study is obtaining permission to measure TOM of the SMEs. IRB approval will be needed to use the SMEs as participants. Additionally, the SEXI instrument derived from the SMEs may indicate the existence of personal information and inadvertently create PII or PDI of the 50 executives of Fortune 500 organizations and 50 Hollywood celebrities. This study will not collect or retain any personal information. IRB approval will be obtained prior to the formation of the SMEs and data collection.

Exposure of the executives and their respective organizations is an issue in this study. This issue will be addressed by randomizing the list of fortune 500 companies and subsequently assigning each organization a nondescript identifier (F001, F002, etc.). Additionally, the executive position titles (e.g. CEO, CIO, CFO, etc.) will be randomized and given a title designation (e.g. C01, C02, C03, etc.), which will not directly indicate the position nor the executive. Efforts will be made to maintain the confidentiality of all Fortune 500 organizations and associated executives. A unique identifier will be applied to each executive, i.e. F023-C06, thereby obfuscating the organization and executives. The original designations will be stored in a separate system.

Exposure of the Hollywood personas is also an issue in this study. This will be addressed by randomizing a list of the 500 top grossing films of all time, filtered to exclude titles released before 1980, and assigning each movie a nondescript identifier (e.g. M001, M002, etc.). Hollywood personas will be randomly selected from the top 10

cast positions from each feature presentation according to the IMDB. Each persona will be obfuscated via a nondescript identifier (e.g. H01, H02, etc.). A unique identifier will be applied to each Hollywood persona, i.e. M081-H03 to maintain their confidentiality. The original designations will be stored in a separate system.

Another barrier that this research study must overcome is the requirement of validity. To address this barrier, a close-ended Delphi will be used with a pre-defined stop criterion. Content validity will be addressed by providing the findings of each Delphi round to the SMEs in aggregate form for them to evaluate (Linstone & Turoff, 1975). The responses of the SMEs solicited for participation in this study may not reach consensus or be constructive, thereby posing another issue. Therefore, to address this concern, each item will be individually assessed through multiple rounds. Items that do not reach consensus will be presented to the SEMs in a subsequent round for re-evaluation (von der Gracht, 2012).

TOM, the imputation of mental states to oneself and to others (Premack & Woodruff, 1978), within the SMEs is expected to be an issue due to their respective understanding of privacy. Mitnick and Simon (2002), McCallister et al. (2010), Schwartz and Solove (2011), Pavlou (2011), Junger et al. (2017) discussed the issue of privacy being contextual and thereby idiosyncratic. Therefore, to address this concern, the SMEs will be asked to answer a survey to better understand their respective experiences and conceptualization associated with privacy to provide a richer understanding of the panel composition and to ensure they meet the requirements. The survey will also present questions on organizational privacy policy and practices, as these may not necessarily be synonymous.

The sample represents a random selection of executives from U.S. organizations and Hollywood personas. The results may not be representative of all similar positions within U.S. organizations, entertainment industries, or those found in other countries. This research study will be performed on a fixed set of U.S. executives and Hollywood personas. To get a cross-section of executives, the sample will include individuals spread across randomly selected U.S. based organizations and positions from the list of Fortune 500 companies as of 2018. To get a cross-section of Hollywood personas, the sample will include individuals spread across the top 500 grossing films of all time, filtered to exclude titles before 1980.

Using the Delphi method is a potential barrier vis-à-vis over-simplification, suppression of uncertainty, and bias (Linstone & Turoff, 1975). This issue will be addressed by seeking SMEs from multiple industries having extensive professional privacy experience. Additionally, items of consensus and those discarded will be made available and discussed.

Assumptions, Limitations, and Delimitations

Assumptions

It is assumed that SMEs will be able to provide the required components and hierarchical weights as well as reach consensus required to develop the SEXI instrument. Additionally, it is assumed that the SMEs will provide honest and truthful responses as to their experience and expert opinion. An assumption is also made as to the availability and accessibility of personal information via open source.

Limitations

This research study will develop a new benchmarking instrument, the SEXI benchmarking index, based on the foundational literature, as well as the feedback, validation, and adjustments needed from the SMEs via the Delphi method. SMEs will be asked to provide feedback on the SE exposure candidate components found in the literature and provide additional relevant components that were not previously in the literature. The second limitation is the set of measures combined to form SEXI. Given that cyber attacks and SE attacks, in particular, are changing over time, the SEXI benchmarking index is based on the current SE threat vectors, techniques, or approaches. The SEXI instrument is envisioned to require more adjustments in the future in response to trends in SE, changes in social media security and privacy settings, as well as innovations that evolve the means by which identity theft occurs. The third limitation is the reliance on an American group of experts for the SME panel to establish the instrument. International participation of SMEs may represent broader population of SMEs, while providing more generalizability to the relative weights, criteria, and measures (Wilkerson et al., 2017). The fourth limitation is the group of executives from Fortune 500 companies as well as the Hollywood personas. Therefore, the results may not be generalizable to other populations.

The sixth limitation of this study may be response bias. The SMEs will be asked to describe their privacy experiences and organizational practices. A potential exists for response bias, acquiescence bias, or social desirability bias. To mitigate this limitation, the SMEs will be informed that their responses will not be attributable and will be

reported anonymously (with quotes sanitized, if necessary) or else reported in the aggregate.

Delimitations

First, a delimitation of this research study will be the convenience sampling of the experts recruited for the panel. Sekaran and Bougie (2013) defined convenience sampling as “the collection of information from members of the population who are conveniently available to provide it” (p. 252). The experts will be solicited from multiple professional associations.

The second delimitation is that each source will be validated to ensure that it correctly associates with the executive or Hollywood persona. A possibility exists for the returned data may be associated with another individual having the same identifier, such as name. Specific details will not be collected.

Data collection in this study will comprise a third delimitation, as it will depend on the existence of information at the point of the survey. The availability of personal information may be unpredictable as well as subject to technology implementation and limitations. The information may or may not exist when queried or on subsequent queries. The source of data may also change. To address this issue, each query will be timestamped, logged, and archived for analysis. A fourth delimitation will be that data will be collected during a specific period for the study. A fifth delimitation will be that all information items will be coded as either located (1) or not found (0), while the actual data will not be captured as it is not required for analysis or construction of the SEXI score. The sixth delimitation of this study will the restriction of the scope of this study to

validate the SEXI instrument on only 50 executives of Fortune 500 companies and 50 Hollywood personas.

Definition of Terms

The following represent terms and definitions.

Anonymous – "implies that the data cannot be manipulated or linked to identify an individual" (Sweeney, 1997, p. 100).

Anonymous information – "is defined as previously identifiable information that has been de-identified and for which a code or other association for re-identification no longer exists" (McCallister et al., 2010 p. 4-5).

Biometric – "A measurable, physical characteristic or personal behavioral trait used to recognize the identity, or verify the claimed identity, of an applicant. Facial images, fingerprints, and iris image samples are all examples of biometrics" (Ferraiolo, Cooper, Francomacaro, Mehta, & Sokol, 2013, p. 64).

Business email compromise – "the scammer skillfully impersonates a trusted entity, typically a colleague or vendor, asking the would-be victim to help perform a task… sending information or money" (Jakobsson, 2016, p. xiv).

Cognitive privacy link - the surmised private connection between an actor and a provider (Acquisti & Grossklags, 2005; Bandura, 2001).

Confidentiality – "preserving authorized restrictions on access and disclosure, including means for protecting personal privacy and proprietary information" (44 U.S.C. § 3552, p. 1).

Content validity – "the extent to which the questions on the instrument and the scores

from the questions are representative of all the possible questions that could be asked about the content or skills" (Creswell, 2012, p. 618).

Convenience sampling –"the collection of information from members of the population who are conveniently available to provide it" (Sekaran & Bougie, 2013, p. 252).

Deception – "manipulation of another person's thoughts—making someone believe something false" (Baron-Cohen, 1992, p. 1142).

Deidentified data – "all explicit identifiers, such as SSN, name, address, and telephone number, are removed, generalized, or replaced with a made-up alternative … does not guarantee that the result is anonymous" (Sweeney, 1997, p. 100).

Deidentified information – "is used to describe records that have had enough PII removed or obscured, also referred to as masked or obfuscated, such that the remaining information does not identify an individual and there is no reasonable basis to believe that the information can be used to identify an individual, [which] can be reidentified" (McCallister et al., 2010 p. 4-4).

Delphi method – " a method for structuring a group communication process so that the process is effective in allowing a group of individuals, as a whole, to deal with a complex problem" (Linstone & Turoff, 1975, p. 3).

Descriptive study – "often designed to collect data that describe the characteristics of persons, events, or situations (Sekaran & Bougie, 2013, p. 97).

Developmental research – "(i) supporting the development of prototypical products (including providing empirical evidence for their effectiveness), and (ii) generating methodological directions for the design and evaluation of such products" (Van den Akker, Branch, Gustafson, Nieveen, & Plomp, 2012, p. 4).

Distinguish – “is to identify an individual” (McCallister et al., 2010, p. 2-1).

Exploratory study – “used when not much is known about the situation at hand, or no information is available on how similar problems or research issues have been solved in the past” (Sekaran & Bougie, 2013, p. 96)

Exposure – “a measure of how well an object … can be observed … over a period of time” (Meguerdichian, Koushanfar, Qu, & Potkonjak, 2001, p. 139).

Grey literature – “is published material that is not indexed and often lacks data about the publisher” (Fleisher, 2008, p. 853).

Harm – “any adverse effects that would be experienced by an individual whose PII was the subject of a loss of confidentiality, as well as any adverse effects experienced by the organization that maintains the PII” (McCallister et al., 2010, p. ES-1).

Highly restricted personal information – “means an individual’s photograph or image, social security number, medical or disability information” (18 U.S.C. § 2725, p. 601).

Information – “Any communication or representation of knowledge such as facts, data, or opinions in any medium or form, including textual, numerical, graphic, cartographic, narrative, or audiovisual” (Ross, Viscuso, Guissanie, Dempsey, & Riddle, 2016, p. 22).

Information privacy – “the ability of the individual to personally control information about one's self” (Stone, Gueutal, Gardner, & McClure, 1983, p. 460).

Information security – “protecting information and information systems from unauthorized access, use, disclosure, disruption, modification, or destruction”(44 U.S.C. § 3552, p. 1).

Information type – “A specific category of information (e.g., privacy, medical, proprietary, financial, investigative, contractor sensitive, security management), defined

by an organization, or in some instances, by a specific law, Executive Order, directive, policy, or regulation" (FIPS 199, 2004).

Intimate self-disclosure – "are … those that contain high-risk (as opposed to low-risk) information that makes the discloser feel vulnerable in some way" (Moon, 2000, p. 323).

Intimate information exchanges – "as those involving risky, evaluative disclosures – tend to lead to resilient long-term relationships in which both parties experience strong feelings of commitment and loyalty" (Moon, 2000, p. 331).

Linkable information – "is information about or related to an individual for which there is a possibility of logical association with other information about the individual" (McCallister et al., 2010, p. 2-1).

Linked information – "is information about or related to an individual that is logically associated with other information about the individual" (McCallister et al., 2010, p. 2-1).

Measurement of the self – "a recording of an observation about the self, which may include the environment to which the self is exposed" (Kang et al., 2011, p. 814).

Mental states – "purpose or intention, as well as knowledge, belief, thinking, doubt, guessing, pretending, liking, and so forth" (Premack & Woodruff, 1978, p. 515).

Monetization – "often means parsing … data for behavioral targeting and advertising, in ways that the average user is unaware" (Kang et al., 2011, p. 824).

Obscured Data – "Data that has been distorted by cryptographic or other means to hide information. It is also referred to as being masked or obfuscated" (McCallister et al., 2010 p. E-1).

Open source – “publicly available print and digital/electronic data from unclassified, non-secret, and ‘grey literature’ sources,” not requiring credentials or special access, including data available through breaches, leaks, etc. (Fleisher, 2008, p. 853).

Open source personal information – personal information that is available openly to everyone who has access to the Internet (Fleisher, 2008)

Personal branding – “the process whereby people and their careers are marked as brands and it differs from reputation management and impression management with its purpose” (Karaduman, 2013, p. 465).

Personal information – “means information that identifies an individual, including an individual’s photograph, social security number, driver identification number, name, address (but not the 5-digit zip code), telephone number, and medical or disability information…” (18 U.S.C. § 2725, p. 601).

Personally distinguishable information – “any information about an individual maintained by an agency … that can be used to distinguish or trace an individual’s identity … and is linked or linkable to an individual” (McCallister et al., 2010, Section 2.1).

Personally identifiable information – “refers to information that can be used to identify or locate an individual” (Chellappa & Sin, 2005, p. 188).

Personally unidentifiable information – “information that, taken alone, cannot be used to identify or locate an individual” (Federal Trade Commission, 2000, p. 46).

Persuasion – “changing persons' mental states, usually as precursors to behavioral change” (O'keefe, 2002, p. 32).

Phishing – "is a criminal trick of stealing victims' personal information by sending them spoofed emails urging them to visit a forged webpage that looks like a true one" (Wenyin, Huang, Xiaoyue, Min, & Deng, 2005, p. 1060).

Pretending – "of 'acting as if' something is the case when it is not" (Leslie, 1987, p. 413).

Pretense – "deliberately distort reality" (Leslie, 1987, p. 412).

Pretext – "an imposter creates a setting designed to influence an intended victim to release sensitive information, pay money, or perform actions that compromise the confidentiality of information" (Workman, 2008, p. 3).

Privacy – "the degree to which an individual can control the collection, disclosure, and use of personal data" (Kang et al., 2011, p. 820).

Privacy chain – **"**the flow of PUI/PII/PDI [personal information] communication between two endpoints" (Wilkerson et al., 2017, p. 3).

Privacy Web the extent PUI/PII/PDI [personal information] is gathered and transferred in relation to an individual to heterogeneous systems (Acquisti et al., 2015; Braun et al., 2001; McCallister et al., 2010).

Publicly available information – "Information that has been published or broadcast for public consumption, is available on request to the public, is accessible on-line or otherwise to the public, is available to the public by subscription or purchase, could lawfully be seen or heard by any casual observer, is made available at a meeting open to the public, or is obtained by visiting any place or attending any vent that is open to the public" (Defense Intelligence Agency, 2011 p. GL-144).

Record – “means any item, collection, or grouping of information about an individual that is maintained by an agency [of the U.S. Federal Government], including, but not limited to, his education, financial transactions, medical history, and criminal or employment history and that contains his name, or the identifying number, symbol, or other identifying particular assigned to the individual, such as a finger or voice print or a photograph. (5 U.S.C. § 552a, p. 317).

Reidentification – “combines datasets that were meant to be kept apart, and in doing so, gains power through accretion: Every successful reidentification, even one that reveals seemingly nonsensitive data like movie ratings, abets future reidentification” (Ohm, 2010, p. 1705).

Representation – “to represent aspects of the world in an accurate, faithful, and literal way, in so far as this is possible for a given organism” (Leslie, 1987, p. 414).

Risk – “refers to uncertainty about and severity of the events and consequences (or outcomes) of an activity with respect to something that humans value” (Aven & Renn, 2009, p. 6).

Sanitization – “Process to remove information from media such that information recovery is not possible. It includes removing all labels, markings, and activity logs” (Ross, Katzke, & Johnson, 2006, p. 8).

Self –"a list of terms or features that have been derived from a lifetime of experience with personal data” (Rogers, Kuiper, & Kirker, 1977, p. 677).

Self-disclosure – “the act of revealing personal and sensitive information about oneself” (Moon, 2000; Peer & Acquisti, 2016, p. 429).

Self-surveillance – “a practice that measures, collects, and stores self-surveillance data” (Kang et al., 2011, p. 814).

Self-surveillance data – “are measurements of the individual self, initiated by the self, using sensors that are in one's control, for the primary purpose of measuring the self” (Kang et al., 2011, p. 814).

Semantic attack – “The manipulation of user-computer interfacing with the purpose to breach a computer system’s information security through user deception” (Heartfield & Loukas, 2015, p. 0:1).

Semantics – “the study of meaning and symbolization” (Heartfield & Loukas, 2015, p. 0:1).

SEXI – The social engineering exposure index is a logical and repeatable quantitative measure that indicates the level of personal exposure for an individual. It is also a data aggregation that provides a means for classifying personal information.

Social network sites – “web-based services that allow individuals to (1) construct a public or semi-public profile within a bounded system, (2) articulate a list of other users with whom they share a connection, and (3) view and traverse their list of connections and those made by others within the system” (Boyd & Ellison, 2007, p. 211).

Social engineering – “is a combination of techniques used to manipulate victims into divulging confidential information or performing actions that compromise security” (Luo et al., 2013, p. 2).

Subject matter experts – “define the curriculum universe which we then designate as the "content domain"” (Lawshe, 1975, p. 565).

Theory of mind – “the individual imputes mental states to himself and to others (either to conspecifics or to other species as well)” (Premack & Woodruff, 1978, p. 515).

Trace – “is to process sufficient information to make a determination about a specific aspect of an individual‘s activities or status” (McCallister et al., 2010, p. 2-1).

List of Acronyms

API – Application program interface

BEC – Business email compromise

CEO – Chief Executive Officer

CFO – Chief Finance Officer

CIA – Central Intelligence Agency

CIO – Chief Information Officer

CSO – Chief Security Officer

CVR -- Content Validity Ratio

DDoS-- Distributed Denial of Service

DNA – Does not apply

FBI – Federal Bureau of Investigation

FIPs – Fair Information Practices

GPS – Global Positioning System

IRB – Institutional review board

IS – Information Systems

OSPI – Open source personal information

PDI – Personally distinguishable information

PDIM – The measurement of personally distinguishable information

PICC – Personal Information Candidate Component

PII – Personally identifiable information

PIIM – The measurement of personally identifiable information

PUI – Personally unidentifiable information

PUIM – The measurement of personally unidentifiable information

SE – Social engineering

SEXI – Social engineering exposure index

TOM – Theory of mind

U.S. – United States

UNF – Unfamiliar (used during phase 1 of Delphi method)

Summary

The purpose of this section was to introduce the research study as well as to identify the research problem, barriers and issues, assumptions, limitations, and delimitations. A theoretical justification for the research study was also presented. The research problem that this study will address is the proliferation of SE attacks due to OSPI, which is increasing despite warnings, media exposure, laws, and data breaches. Supporting literature corroborates the research problem and the need for this study.

The literature demonstrates the exponential growth of personal information via open source repositories (Acquisti et al., 2015). Cybercrimes are also on the increase with little information as to where the SE attacks will come from or the composition used (Federal Bureau of Investigation, 2012; Mouton et al., 2016; Tetri & Vuorinen, 2013).

Consequently, the need to determine the availability of personal information as well as predicting potential SE attack vectors is significant to personal and organizational security (Mouton et al., 2016). The need for this work is demonstrated by the literature that acknowledged the progressive expansion of SE attack vectors (Mitnick & Simon, 2002), the lack of a predictive threat system (Tetri & Vuorinen, 2013), the availability of OSPI which circumvent organizational cybersecurity technologies (Heartfield & Loukas, 2015), and the dearth of data on information gathering techniques for the successful execution of prior SE attacks (Mouton et al., 2016).

The goal of this research is to develop and validate a SEXI using OSPI to assist in identifying and classifying SE vulnerabilities. The literature provides grounding for this research with the concept of categories of PII introduced by Schwartz and Solove (2011) and described by McCallister et al. (2010). The newly developed benchmarking index will be validated by measuring the SEXI of 50 Fortune 500 executives and 50 Hollywood personas. The collected data are to be analyzed to assess and statistically test for significant mean differences of the SEXI of 100 individuals and reported.

Multiple barriers need to be overcome to meet the requirements of this dissertation research. Given that limited discernible empirical literature appears to exist regarding how exposure of personal information to social engineering should be measured, rated, or summarized, an expert panel will be tasked with this purpose. The IRB process will address two associated issues in this study: the use of the SMEs to measure TOM, and the collection of publicly available personal information of 50 executives of Fortune 500 companies and 50 Hollywood personas. The SMEs will be informed that their responses will not be attributable and will be reported anonymously

(with quotes sanitized, if necessary) or else reported in the aggregate. The specific OSPI of the executives and Hollywood personas will be codified in a “found” / “not found” dichotomous scale to maintain confidentiality. IRB approval will be obtained before the Delphi method and data collection begins. Each Hollywood persona as well as executive and their respective organization will be coded into a concatenated identification label consisting of two random strings – the first denoting the organization or feature film with the remaining portion made up of a random identifier.

Another barrier that this research study must overcome is the requirement of validity in the weights, groupings, and rankings of the exposure of OSPI. To address this barrier, the SMEs must reach a consensus based on the literature. The resulting SEXI instrument will then be used to assess the exposure of 50 Fortune 500 executives and 50 Hollywood personas. The data are to be analyzed and subsequently reported. The literature will serve as the foundation for the benchmarking index development. The use of open-ended questions, Likert scales, and binary response will be used to facilitate the successful development of the SEXI benchmarking instrument. While the literature discussed taxonomies of SE attacks (Heartfield & Loukas, 2015), described SE (Mitnick & Simon, 2002), established privacy standards (McCallister et al., 2010), discussed privacy (Schwartz & Solove, 2011; Solove, 2006), and critiqued security policies (Wolff, 2016), the effort thus far appears to have fallen short. Mouton et al. (2016), Tetri and Vuorinen (2013), as well as Bélanger and Crossler (2011) have called for a predictive tool that can potentially facilitate organizational cyber-security by providing insight into possible SE attack vectors as well as potential personal information used in their

execution. Therefore, this proposed research will develop and validate the SEXI benchmarking index to measure the level of exposure of executives to SE due to OSPI.

Chapter 2

Review of Literature

Introduction

In this chapter, an overview of relevant literature is offered. Bhattacherjee (2012) described a three-fold purpose for literature review: survey, grounding, and gap identification. Hart (1998) described the obligation for researchers to have an exhaustive grasp of the literature in their area of interest, to provide a foundation for contribution. Ellis and Levy (2006) correlated significance and quality with the accuracy of the review of the literature. This interdisciplinary proposal involves an overview of the information systems (IS) literature using several databases from multiple fields: IS, psychology, law, and business.

Exposure

Meguerdichian et al. (2001) defined exposure as “a measure of how well an object … can be observed … over a period of time” (p. 139). In application, photographers manipulate exposure to control composition and context. Raskar, Agrawal, and Tumblin (2006) described how exposure could be manipulated to provide clarity – even to the most obscured subject by adjusting the amount of time it is viewed. The literature has discussed exposure in the areas of big data (Martin, 2015; Rosenbaum, 2015), biology (Kennedy, Eberhart, & Shi, 2001b; Maeterlinck, 1930), bring your own device (Garba, Armarego, Murray, & Kenworthy, 2015), information privacy (Acquisti et al., 2016;

Smith, Dinev, & Xu, 2011), law (Schwartz & Solove, 2011; Solove, 2006), mindfulness (Orlikowski & Baroudi, 1991; Shapiro, Carlson, Astin, & Freedman, 2006), persuasion (Johnston, Warkentin, & Siponen, 2015; Perloff, 2010), posttraumatic stress disorder (Keane et al., 1989; Youssef et al., 2013), SE (Conteh & Schmick, 2016; Mamonova & Koufaris, 2016), self-disclosure (Bélanger & Crossler, 2011; Culnan & Bies, 2003; Moon, 2000), smartphone (Boyd, 2014; Enck et al., 2014; Xu, Luo, Carroll, & Rosson, 2011), and social network sites (Boyd & Ellison, 2007; Choo, 2011; Minkus, Liu, & Ross, 2015).

There are many venues where people choose to expose their personal information, including social media, personalization, online forms, and smartphones (Acquisti et al., 2015; Falaki et al., 2010; Lee, Ahn, & Bang, 2011). Research suggests that people may be giving up on having privacy (Mamonova & Koufaris, 2016). Junger et al. (2017) found that in certain situations a warning may substantially increase disclosure – not always in accordance with assumptions of less personal information disclosed. Similarly, Wolff (2016) introduced a framework which included a measure to understand how and why humans unpredictably interact with technology in the context of information security. Zhang et al. (2014) found that even though users perceived a heightened online security threat, they tended to expose even more personal information.

The literature also indicates that news announcements of government privacy invasion, cyber threat warnings, and the number of Americans personally experiencing a data breach seem to adversely affect how participants control, protect, and even value their PII (Junger et al., 2017; Mamonova & Koufaris, 2016; Olmstead & Smith, 2017). Johnston et al. (2015) indicated that 40% of data breaches are due to organizational

insiders. Acquisti et al. (2015) described research that found when people are given enhanced control over their privacy they tend to increase the information shared – despite assumptions of researchers to the contrary. Chang, Krupka, Adar, and Acquisti (2016) found that the views and behaviors of people to share personal information become increasingly favorable after viewing images of scantily clad people. Exposure is of interest as the literature has shown that SE attacks usually comprise personal information originating at the target or from peripheral sources (Heartfield & Loukas, 2015; Junger et al., 2017; Luo et al., 2013). Little is known as to what personal information is available via OSPI or how it is specifically used in various SE attack vectors (Mouton et al., 2016; Tetri & Vuorinen, 2013).

McCallister et al. (2010) viewed exposure from the lens of the harm to individuals and organizations associated with the release of confidential information. Geletkanycz and Hambrick (1997) investigated the relations top executives have with external entities and how they are exposed to information as well as alternative understandings. Executive exposure has been the norm for top-level organizational leaders for many industries as a means to do business (Geletkanycz & Hambrick, 1997). This executive exposure was intended to facilitate daily operations and organizational stability (Coleman, 2000; Geletkanycz & Hambrick, 1997). This exposure has led to multiple SE attacks such as the one enacted by a penetration testing team hired by a company that used the voice and travel itinerary of a Chief Finance Officer (CFO) to access key systems (Granger, 2001). Executive exposure can occur in many forms, from shoulder surfing to dumpster diving (Granger, 2001; Mitnick & Simon, 2002). SE attacks via on-line technologies may intertwine email, postal mail, and other sources of readily available information each

providing inroads into the world of the executive via their personal information (Granger, 2001; Heartfield & Loukas, 2015; Luo et al., 2013; Mitnick & Simon, 2002; Mouton et al., 2016; Peltier, 2006). Just as photographers control the exposure of objects to gain clarity (Raskar et al., 2006), social engineers specialize in the collection of sensitive information and in the refactoring of exposed data into a treatise on potential executives, organizations, or other SE targets (Mitnick & Simon, 2002).

The 2015 BEC attack on the CIA Director illustrated how a single piece of information facilitated the exposure of the personal information of many people. Discovering the ISP of the CIA Director lead to a sequence of SE attacks on multiple organizations, each exposing additional personal information based on the prior discovered data. Eventually, the attackers were able to gain access to a personal email of the CIA Director, which in turn contained the personal information of agents and contractors (Franceschi-Bicchierai, 2015). Orgill et al. (2004) described the hazards of allowing extended exposure to the physical environment and employees of an organization resulting in the collection of usernames, passwords, and corporate credit cards. Tetri and Vuorinen (2013) stated, “Contrary to what the literature suggests, we believe that social engineers should get more credit for spotting organisational [sic] weaknesses from the outside rather than being celebrated as great persuaders” (p. 1019). Allen (2006) described how these outsiders expose weaknesses via SE by “gathering information, developing relationships, exploitation, and execution” – repeating the process with newly discovered information (p. 5). According to Mitnick and Simon (2002), exposure is the craft of SE, while organizations and key personnel form the playground.

In the book of Genesis, a story is told of Noah who became inebriated and naked, eventually passing out in the privacy of his own dwelling. An outsider, his son Canaan, found him naked and proceeded to expose that information to others (Genesis 1: 23, 24). Personal information exposure comes in many forms from voluntary disclosure (Bélanger & Crossler, 2011) to big data (Martin, 2015). For photographers, exposure facilitates composition and context (Raskar et al., 2006). For personal information, its exposure affects composition and context to third-parties, which RQ1 and RQ2 will quantitatively and qualitatively assessed (McCallister et al., 2010; Schwartz & Solove, 2011). A summary appears in Table 1 of the literature referenced in this section.

Table 1

Summary of Exposure Literature

Study	**Methodology**	**Sample**	**Instruments or Constructs**	**Main Finding or Contribution**
Acquisti et al. (2015)	Review		Literature streams: Context-Dependence Malleability and Influence Uncertainty	"Norms and behaviors regarding private and public realms greatly differ across cultures, within cultures, while varying dramatically for the same individual, and for societies, over time" (p. 513).
Acquisti et al. (2016)	Comprehensive Survey of Literature		Literature Streams:	"One of the themes emerging from this review is that both the sharing and the

			Consumers Unaware of Privacy threats Economic Theory Empirical Analysis of Privacy Exposure in Varying Scenarios Unifying Economic Theory of Privacy	protecting of personal data can have positive and negative consequences at both the individual and societal levels" (p. 483).
Allen (2006)	Descriptive		The Cycle: Information Gathering Developing Relationship Exploitation Execution	"[T]here will always be the possibility of the 'human factor' being influenced by a social, political and/or cultural event" (p. 9).
Bélanger and Crossler (2011)	Review	500 Articles 142 Journal Articles 102 Conference Proceedings	Framework of Theory Classifications Information Privacy Structural View of	Many topics Research focused largely on explaining and predicting Research is largely confined

			Information Privacy	to the U.S. and student contexts
Boyd and Ellison (2007)	Descriptive	Historical overview Serves as introduction of 7 Articles for a special issue	Exposure Signaling Theory	A formal definition of "social network sites" (p. 211). Overview of Social Network Sites and underlying methodology such as "friending".
Boyd (2014)	Survey	166 Formal, semistructured interviews of teens over three years	Audience Media Public	Insight into the minds of youth and their use of privacy-related technologies
Chang et al. (2016)	Experiment	Main study: 387 Turk Workers (105 female / 200 male) Study 2: 82 (38 female / 44 male)	Less Provocative Images Provocative Images	"Empirically identifying a key mechanism by which norm-shaping designs can change beliefs and subsequent disclosure behaviors" (p. 587).
Choo (2011)	Descriptive		Routine Activity Theory	The authors "explain how the Routine Activity Theory can help

				to inform and enhance cyber crime prevention strategies" (p.720).
Coleman (2000)	Longitudinal Study	4000 Students from public schools	Human Capital Social Capital	Demonstrated "the effect of social capital in the family and in the community in aiding the formation of human capital" (p. S118).
Conteh and Schmick (2016)	Review			"[W]hile technology has a role to play in reducing the impact of social engineering attacks, the vulnerability resides with human behaviour [sic], human impulses and psychological predispositions that can be influenced through education" (p. 37).
Culnan and Bies (2003)	Review		Fair Information Practices Justice theory	"[S]uggests new privacy rules are needed" (p. 335).

			Trust-gap	"[S]elf regulation is unlikely to work 100% of the time as there will always be bad actors or organizations who have implemented the formal trappings but not the substance of fair information practices creating a need for baseline privacy legislation" (p. 338).
Enck et al. (2014)	Descriptive		Taintdroid	"We have presented TaintDroid, an efficient, system-wide information-flow tracking tool that can simultaneously track multiple sources of sensitive data" (p. 5:25).
Falaki et al. (2010)	Field Study	Dataset 1: 33 Android users (16 high school students; 17 knowledge workers) Dataset 2: 222 Windows	Business Power User Life Power User Organizer Practical	"[W]e comprehensively characterized user activities and their impact on network and battery [and] quantify many hitherto unknown aspects of

		Mobile users (116 U.S.; 106 United Kingdom)	Social Communicator	smartphone usage" (p. 193).
Franceschi-Bicchierai (2015)	News Article			Describes the SE attack on CIA Director by teenagers.
Garba et al. (2015)	Review		Bring Your Own Device (BYOD) Information Security Mobile Computing Organizational Practices Privacy	"[A]ny attempt for organizations to adopt or implement BYOD without adequate attention to the security and privacy issues or challenges … may increase their risk of confidential information loss" (p. 52).
Geletkanycz and Hambrick (1997)	Descriptive	30 large publicly-traded firms in two industries: branded foods, computer	Performance Strategic conformity	"[G]reater understanding of interorganizational [sic] relations and the implications of external tie" (p. 673).

Granger (2001)	Descriptive			Provides real-world examples of SE.
Heartfield and Loukas (2015)	Taxonomy	Discusses research with 1900 malicious URLs, 308 users, and other	Deception Exploitation Execution Orchestration Vector	"It introduces a structured baseline for classifying semantic attacks by breaking them down into their components" (p. 0:31).
Keane et al. (1989)	Survey	362 male Vietnam-era veterans across three studies	Combat Exposure Scale	"[T]he three studies presented here confirms that the CES merits consideration for further use by clinicians and researchers" (p. 54).
Kennedy et al. (2001b)	Simulation using the De Jong Test Suit	P=20 or 100 N=20 or 100 (p. 306)	Emergent behavior (self-organization) Particles Swarm Theory	Interpretation and computer programs in relation to I. Minds are social. II. Particle swarms are a useful computational intelligence (soft computing) methodology (p. 395, 396).

Johnston et al. (2015)	Experiment	559 insiders of a Finland city government	Compliance Intention Formal Sanction Certainty Formal Sanction Severity Informal Sanction Certainty Informal Sanction Severity Intention to comply with recommended protective strategies Protection motivation theory Sanction Celerity Self-Efficacy Threat Severity	"This study develops and tests an enhanced fear appeal rhetorical framework that accounts for the distinction between threats to information assets and threats to human assets" (p. 130).

			Threat Susceptibility	
Junger et al. (2017)	Experiment	278 participants	Age Age Square Goals System Theory Priming Total Risk Warning	"This study found relatively high disclosure rates… Neither priming nor a warning influenced the degree of disclosure." (p. 85).
Lee et al. (2011)	Field Study	2 Firms	2 Price measures 3 consumer measures 3 consumer group measures for willingness to share personal information 4 Cost measures	"[S]trategic choices of privacy protection can work as a competition-mitigating mechanism in personalization…. A firm's privacy protection strategy under competition should be based on the investment cost of protection and the size of the personalization scope" (pp. 440-441).

		Personalization Scope	
		Game theory	
		Privacy calculus	
		Profit	
Luo et al. (2013)	Descriptive	Defenses	"in addition to advanced technologies counterattacking various security intrusions, human factors must be equally accounted for" (p. 7).
		Personality Traits	
		Psychological Aspects	
		Social Engineering	
		Techniques	
Maeterlinck (1930)	Exploratory	Ants	One of the earliest swarm behavior references in the literature.
		Precursor to Swarm Theory	
		Various other Swarming Species	

Mamonova and Koufaris (2016)	Experiment	Group 1: 222 technology users Group 2: 220 technology users	Government Intrusion Concerns Password Strength Privacy Concern Privacy Self-Efficacy	"[T]he exposure to government surveillance news led to the use of weaker passwords, suggesting that the exposure to government surveillance may trigger helplessness in relation to protecting privacy" (p. 64).
Martin (2015)	Exploratory		Aggregation Destructive Demand Downstream Uses Information Supply Chain Negative Externality os Surveillance Potential for Secondary Market Upstream Supplier	"[I]dentified the Big Data Industry as having both economic and ethical issues at the individual firm, supply chain and general industry level and has suggested associated solutions to preserve sustainable industry practices" (p.85).

			Use of Consumer-Level Data	
McCallister et al. (2010)	Descriptive		Defines key terms associated with privacy and personal information.	NIST 800-122 (Guide to Protecting the Confidentiality of Personally Identifiable Information (PII))
Meguerdichian et al. (2001)	Simulation and case studies	2 – 8 Sensors	Exposure Exposure-Based Coverage Model	"[W]e presented an efficient and effective algorithm for minimal exposure paths for any given distribution and characteristics of sensor networks" (p. 148).
Minkus et al. (2015)	Descriptive	2,383 Adult Facebook Users via shallow data mine limited to public posts Survey of 357 Adult Facebook Users 1,089 Instagram Users	Birthday Face Name Location Matched to Voter's registration for demographics	"We can therefore conclude that although a substantial percentage of parents are compromising the privacy of their children in their public Facebook pages, significantly more are doing so among Facebook friends" (p. 782).

Mitnick and Simon (2002)	Descriptive			Brought social engineering into the mainstream. Social engineering attack cycle
Moon (2000)	Experiment via interview using a computer as interviewer	30 participants	Reciprocity Self-disclosure Theory of social response	The wording and sequence of questions can successfully solicit intimate details from users via computer. Explicit reward is not required to successfully solicit personal information from a user.
Mouton et al. (2016)	Descriptive		Theory of Group Conformity SE attack -Compliance Principles -Goal -Medium -Social Engineer	Neither the literature or news media provide all the information concerning an attack. Usually little, if any, information is known about a potential attack. Little is known as to where the

			-Target	information is obtained for a SE attack.
			-Techniques SE Framework - Attack Formation	Little is known as to what information is available for a SE attack.
			-Debrief	
			-Develop Relationship	
			-Exploit Relationship	
			-Preparation Information Gathering	
Olmstead and Smith (2017)	Survey	1,014 adult-aged U.S. citizens	Demographics	64% of Americans have experienced a data breach. 12% use password management software.
Orgill et al. (2004)	Questionnaire	32 participants	Department	“This study demonstrated that even in a company where security is a

		-26 gave their username -19 gave their password -7 gave login credential information above their own access -4 asked for a name badge or identification	Number Surveyed Password Username	concern, these human traits [trust others, assist others, gain favor] can be ill-used if proper preventative measures are not taken … This study also shows the importance of assessing security effectiveness through means such as audits.… In order for an audit to be effective, the auditor has to be at least as thorough, through preliminary studying, planning, and execution as a potential social engineer would be" (p. 181). Some departments had more training and resisted the social engineer better.
Orlikowski and Baroudi (1991)	Review	155 Information systems articles	Epistemology Frequency Journal	"[R]esearchers should ensure that they adopt a perspective that is compatible with their own research interests and predispositions,

				while remaining open to the possibility of other assumptions and interests" (p. 24).
Peltier (2006)	Review			Magazine article describing SE to readers.
Perloff (2010)	Exploratory		Persuasion	Extensive discussion on persuasion, which is used in many SE attack vectors.
Raskar et al. (2006)	Descriptive	3 Cases	Coded blur Chops Flat blur	Demonstrated how manipulation of exposure increased clarity of the subject.
Rosenbaum (2015)	Survey *Dissertation	53 SMEs	Privacy Violation Scale	"[E]vidence strongly suggested that some practitioners were less willing to commit privacy violations than were other practitioners; this is based upon some practitioners identifications with various moral and computing Hallmark

				Features" (p. 115).
Schwartz and Solove (2011)	Exploratory		Identifiable Personal Information Identified Personal Information Unidentifiable Personal Information	"PII 2.0 protects information that relates either to an identified or identifiable person, and associates different legal interests with each category" (p. 1894).
Shapiro et al. (2006)	Exploratory		Attention Attitude Exposure Intention Mindfulness	"We have attempted to provide a first formulation of a model to describe how mindfulness might be fostering transformation and change" (pp. 384-385).
Smith et al. (2011)	Exploratory	Four decades of literature: 320 Privacy Articles 128 Books and Book Sections	Antecedents Outcomes Privacy Concerns	"[T]he overall [privacy] research stream has been suboptimized [sic] because of its disjointed nature" (p. 1008).

Solove (2006)	Exploratory	Information Collection Information Dissemination Information Processing Invasion	"I have attempted to provide a clearer and more robust account of privacy—one that provides us with a framework for understanding privacy problems" (p. 558).
Tetri and Vuorinen (2013)	Descriptive	Actor-Network Theory	Describes issues in SE research and suggests the theories from the psychology literature should only be applied to the persuasion component of SE.
Wolff (2016)	Exploratory	Classification of perverse effects Duality of technology Technology-Interaction perverse effects Theory of unintended consequences	"This classification scheme is intended as a step beyond simply warning defenders that they have to be careful when adding new security controls by giving them a framework for analyzing the different possible mechanisms by which those controls may interact with the

			User-Interaction perverse effects	system and its users to introduce new vulnerabilities and produce perverse effects" (p. 615).
Xu et al. (2011)	Exploratory		Covert vs. Overt Exchange Theory Interpersonal Differences Personalizatio n Privacy Calculus Purchase Intention Willingness to Share Personal Information in Location-Aware Marketing	"[T]he findings of this research have provided preliminary empirical evidence about how users strike a balance between value and risk" (p. 50).
Youssef et al. (2013)	Cross-Sectional Field Study	1,488 military personnel and veterans serving after	Beck Depression Inventory-Second Edition	"The study findings suggest that comprehensive assessment of both childhood

		September 2001	Beck Scale for Suicide Ideation Combat Exposure Scale Connor-Davidson Resilience Scale Davidson Trauma Scale Traumatic Life Events Questionnaire	trauma and resilience among military personnel and veterans can contribute to the understanding of their clinical status in terms of depression and suicidal ideation, and ultimately their clinical care" (p. 116).
Zhang et al. (2014)	Experiment	220 online U.S. resident adults	Attitude Behavioral Intention Instant Gratification cue Security cue Threat Trust	"[T] he security cue heightens perceived threat but also encourages greater disclosure of one's account and network strength on social media" (P. 113).

Personal Information

The term OSPI is sparingly used in the literature, though it is described extensively throughout the privacy literature as any personal information belonging to an individual extended to include any being publicly available (Federal Trade Commission, 2000; Schwartz & Solove, 2011). Rogers et al. (1977) fully integrated personal information with the definition of self, “as a list of terms or features that have been derived from a lifetime of experience with personal data” (p. 677). Furthermore, the literature appears to infer that the tendency of people to share information may be more of an attempt to process one’s respective life than an intentional self-disclosure (Rogers et al., 1977). Mitnick and Simon (2002) described open source information as “SEC filings and annual reports, marketing brochures, patent applications, press clippings, industry magazines, Web site content, and also dumpster diving” (p. 310). Maynard et al. (2015) described the accessibility of PII due to content associated with a social media service and application program interface (API), such as Twitter hashtags and posts. Oltmann (2010) described a continual degradation of the privacy of Facebook users sharing photos, data, and preferences. Sanders (2012) discussed the advent of credit reporting agencies using information collected from social networking sites. The Privacy Act of 1974 provides a broad understanding of personal information, when defining a record to include personal, medical, criminal, education, employment histories, etc., (5 U.S.C. § 552a). The literature also discusses the existence of underground hacker markets where

attack vectors, targets, and compromised PII are shared (Benjamin & Chen, 2012; Coleman & Golub, 2008; Jasper, 2017).

Following Schwartz and Solove (2011), OSPI is comprised of PUI (information which does not identify an individual), PII (information which can be used for identification), and PDI (information which explicitly identifies an individual). The primary source of OSPI is from people themselves (Acquisti et al., 2015), the organizations they work for (Federal Bureau of Investigation, 2015a), and social network sites (Acquisti et al., 2015; Federal Bureau of Investigation, 2012). Additional sources of OSPI such as data mining technologies can be used from command prompts on personal computers of any modern operating system (Russell, 2013), while credit reports and background checks can easily be requested even without consent (Sanders, 2012). Simple friend requests on social networks may reveal extreme amounts of PII and PDI (Boyd & Ellison, 2007; Maar, 2013; Mouton et al., 2016). The literature describes OSPI as personal information that is available openly to everyone who has access to the Internet (Fleisher, 2008).

The literature also discusses personalization, another exposure threat to privacy which may directly feed OSPI (Chellappa & Sin, 2005; Lee et al., 2011; Sutanto et al., 2013; Xu et al., 2011). Data brokers have formed entire supply chains (termed herein as privacy chains) of PDI, PII, and PUI pooled from a variety of sources and compiled into datasets, which are then repackaged and made available (Anthes, 2014; Kang et al., 2011). Similarly, FIPS 199 (2004) provided precedence for information categories including privacy, medical, financial, etc. Additionally, research has indicated that people are sharing an increasing amount of PII on social networks and continue to do so despite

being warned against it (Acquisti et al., 2015; Olmstead & Smith, 2017). Acquisti and Grossklags (2005) concluded “preliminary data show that privacy attitudes and behavior are complex but are also compatible with the explanation that time inconsistencies in discounting could lead to under-protection and overrelease [sic] of personal information” (p. 32). Krishnamurthy and Wills (2009) described the risk associated with exposure where specific identification of an American can be accomplished with only their date of birth, gender, and postal zip code. However, little is known in the literature about the role that OSPI play in SE attacks or even how much personal information is required for a successful attack (Krishnamurthy & Wills, 2009; Mouton et al., 2016; Tetri & Vuorinen, 2013).

Personal information is essentially the existence of an individual relegated to data points (Rogers et al., 1977). The literature described personal information as contextual (Culnan, 1993), having three levels of harm (McCallister et al., 2010), and three levels of exposure (Schwartz & Solove, 2011), which is the foundation of RQ1, RQ2, as well as RQ3. A summary appears in Table 2 of the literature referenced in this section.

Table 2

Summary of Personal Information Literature

Study	**Methodology**	**Sample**	**Instruments or Constructs**	**Main Finding or Contribution**
5 U.S.C. § 552a	Standard		Record	Defines various components of personal information, including medical, education, employment,

				criminal, and other histories.
Acquisti and Grossklags (2005)	Survey	119 Students 19–54 years old	General Privacy Concern Offline Identity Online Identity Personal Profile Professional Profile Sexual and Political Identity	"The evidence points to an alternation of awareness and unawareness from one scenario to the other" (p. 29). "Although respondents realize the risks associated with links between different pieces of personal data, they are not fully aware of how powerful those links are" (p. 30). "Even if individuals have access to complete information about their privacy risks and modes of protection, they might not be able to process vast amounts of data to formulate a rational privacy-sensitive decision" (p. 30).
Acquisti et al. (2015)	Literature review		Self-Disclosure	"Norms and behaviors regarding private and public realms greatly differ across cultures,

			Social Penetration Theory	within cultures, while varying dramatically for the same individual, and for societies, over time" (p. 513).
Anthes (2014)	Review			"Asking consumers to 'opt out' of data collection at myriad companies they have never heard of is unrealistic, and the existing online "notice and consent" forms—in which users "agree" to the collection and use of personal data—are ineffective because they are mostly ignored by consumers" (P. 30).
Benjamin and Chen (2012)	Exploratory	28,537 hackers 723,555 forum posts	Average Message Length Control Theory Number of Replies Number of Threads Involved Reputation	"Hackers that contributed to cognitive advance of their community or were considerably active had the highest reputations" (p. 6).

			Tenure Sum of Attachments Total Messages	
Boyd and Ellison (2007)	Descriptive	Historical overview Serves as an introduction of 7 Articles for a special issue	Exposure Signaling Theory	A formal definition of “social network sites” (p. 211). Overview of Social Network Sites and underlying methodology such as “friending”.
Chellappa and Sin (2005)	Empirical study	243 Consumers	Consumer Concern for Privacy Likelihood of Using Personalization Services Value of Online Personalization	“the consumers’ value for personalization is almost two times (0.59 vs. −0.34) more influential than the consumers’ concern for privacy in determining usage of personalization services” (p. 197).
Coleman and Golub (2008)	Descriptive		Liberalism, Anarchism, Hacker Ethics	“hacker practice makes visible socially relevant questions to those interested in the

		Political Theory	legal politics of information access" (p. 271).
Federal Trade Commission (2000)	Descriptive	Access Choice Notice Privacy Privacy Seal Security Self-regulation	FIPS (Privacy Online: Fair Information Practices in the Electronic Marketplace) "Because self-regulatory initiatives to date fall far short of broad-based implementation of self-regulatory programs, the Commission has concluded that such efforts alone cannot ensure that the online marketplace as a whole will follow the standards adopted by industry leaders" (p. ii).
FIPS 199 (2004)	Descriptive	Information Type Potential Impact	FIPS Publication 199 Standards for Security Categorization of Federal Information and Information Systems
Jasper (2017)	Review		Discusses Cyber Threat Intelligence Integration Center.

				"Therefore, the timely sharing of relevant and actionable cyber threat intelligence, in the context of cyber threat information and indicators, is imperative to reducing the impact of attacks" (p. 62).
Kang et al. (2011)	Descriptive		Personal Data Stream Personal Data Vault Privacy Rights Management	"Instead of direct behavioral regulation or blind faith in the market, our strategy is to modify indirectly the information ecosystem by introducing a new species, the [Personal Data Guardian]" (p. 847).
Krishnamurthy and Wills (2009)	Longitudinal Study	127 test data set sites 81 Web sites across nine categories	Company Acquisitions Cookies First-Party Content JavaScript Root Domain	"… users are being tracked by multiple entities when accessing a first-party site… [and] existing privacy protection techniques have limitations in preventing privacy diffusion" (p. 15).

			Subdomain Third-Party Content	
Lee et al. (2011)	Field Study	2 Firms	2 Price measures 3 consumer measures 3 consumer group measures for willingness to share personal information 4 Cost measures Personalization Scope Game theory Privacy calculus Profit	"[S]trategic choices of privacy protection can work as a competition-mitigating mechanism in personalization…. A firm's privacy protection strategy under competition should be based on the investment cost of protection and the size of the personalization scope" (pp. 440-441).
Maar (2013)	Survey	49 professional users of social	Benefit Deception Risks	"The study appears to indicate that the three concerns of privacy, deception, and security drive

		networking sites	Ease of Use Habit Linkage Ownership Permeability Personal Norm Privacy Risks Response Efficacy Security Risks Self-Efficacy Trust	the three factors of information protection, boundary permeability, linkage, and ownership respectively" (p. 268). "This study has found that perceived benefits of online social network may motivate users to commit personally to protecting its integrity, but may induce users to relax their vigilance and develop poor online habits" (pp. 268-269).
Martin (2015)	Exploratory		Aggregation Downstream Uses Information Supply Chain	"[I]dentified the Big Data Industry as having both economic and ethical issues at the individual firm, supply chain and general industry level and has suggested associated solutions to preserve

			Negative Externality	sustainable industry practices" (p.85).
			Potential for Secondary Market Destructive Demand	
			Upstream Supplier	
			Use of Consumer-Level Data	
Maynard et al. (2015)	Descriptive	1.8 million tweets, 42 political themes, 20 topics		Describes open source data mining (GATE) involving social networks.
Mitnick and Simon (2002)	Descriptive			Brought social engineering into the mainstream.
				Social engineering attack cycle
Mouton et al. (2016)	Descriptive		Theory of Group Conformity	Neither the literature or news media provide all the information concerning an attack.
			SE attack	
			-Compliance Principles	

			-Goal -Medium	Usually little, if any, information is known about a potential attack.
			-Social Engineer -Target -Techniques	Little is known as to where the information is obtained for a SE attack.
			SE Framework - Attack Formation -Debrief -Develop Relationship -Exploit Relationship -Preparation Information Gathering	Little is known as to what information is available for a SE attack.
Olmstead and Smith (2017)	Survey	1,014 adult-aged US citizens	Demographics	64% of Americans have experienced a data breach. 12% use password management software.

Oltmann (2010)	Exploratory			"If more users could be convinced to adjust their privacy settings, that could help preserve online privacy, which in turn might protect some of society's expectations for privacy in the broader offline world… [otherwise] our overall privacy will decrease" (p. 4).
Rogers et al. (1977)	Experiment	32 students - 16 Female -16 Male 40 adjectives	Self Self-reference	"In the realm of human information processing it is difficult to conceive of an encoding device that carries more potential for the rich embellishment of stimulus input than does self-reference" (p. 687).
Russell (2013)				A book providing tools and instructions for data mining popular social networking sites and online technologies.
Sanders (2012)	Descriptive			Discusses the use of Social Media by credit reporting agencies

	Non-peer-reviewed, non-journal			
Schwartz and Solove (2011)	Exploratory		Identifiable Personal Information Identified Personal Information Unidentifiable Personal Information	"PII 2.0 protects information that relates either to an identified or identifiable person, and associates different legal interests with each category" (p. 1894).
Sutanto et al. (2013)	Field experiment	193 participants	Information Boundary Theory User Gratification Theory	Users assume marketers are using their information, amidst advertisements,
Tetri and Vuorinen (2013)	Descriptive		Actor-Network Theory	Describes issues in SE research and suggests the theories from the psychology literature should only be applied to the persuasion component of SE.
Xu et al. (2011)	Exploratory		Covert vs. Overt	"[T]he findings of this research have provided preliminary empirical evidence about how users

	Exchange Theory	strike a balance between value and risk" (p. 50).
	Interpersonal Differences	
	Personalization	
	Privacy Calculus	
	Willingness to share personal information in location-aware marketing	
	Purchase Intention	

Personally Distinguishable Information

According to the literature, that not all personal information holds the same significance (Heurix et al., 2015; Hong & Thong, 2013; McCallister et al., 2010; Mitnick & Simon, 2002; Schwartz & Solove, 2011). Schwartz and Solove (2011) declared, "All current legal models for this concept are flawed" (p. 1835), while discussing the lack of consensus within the U.S. to define privacy legally and precisely. Additionally, Schwartz and Solove (2011) believed that there is no merit to whether data are identifiable to a specific person when focusing on whether or not information is PII versus non-PII. McCallister et al. (2010) discussed the concept of impact levels due to exposure, while describing confidentiality breaches.

McCallister et al. (2010) referred to the information used to identify an individual as being distinguishable, providing a subset of PII to separate ultimate exposure leading to definite identification from a generic catchall of potential exposure. Safety guides also warn users to not post GPS, social security number, security clearance, or information that can be used to answer security questions on Websites, on social media, etc. (Federal Bureau of Investigation, 2012, 2015b). PDI is defined as "any information about an individual maintained by an agency … that can be used to distinguish or trace an individual's identity … and is linked or linkable to an individual" (McCallister et al., 2010, Section 2.1). The primary difference between PII and PDI is the specificity of the information being directly connected to an individual's identity (e.g. a photograph or social security number) rather than only having the potential of identification (e.g. gender or zip code) (McCallister et al., 2010).

McCallister et al. (2010) integrated risk nomenclature to personal information, stating that some PII can prove "hazardous to both individuals and organizations" (p. ES-1) and that "unauthorized access, use, or disclosure of PII can seriously harm both individuals, by contributing to identity theft, blackmail, or embarrassment, and the organization, by reducing public trust in the organization or creating legal liability (p. 2-1). (Schwartz & Solove, 2011) argued "that the continuum of risk is different for these categories. The result is that the necessary legal protections should generally be different for identified and identifiable data" (p. 1818). The literature clearly makes the distinction that the exposure of specific personal information that makes an individual distinguishable is a higher risk and should be treated as such. In following the literature, the SMEs will be asked to categorize items

as PDI and to provide a weight to the PDI category. A summary appears in Table 3 of the literature referenced in this section.

Table 3

Summary of Personally Distinguishable Information Literature

Study	Methodology	Sample	Instruments or Constructs	Main Finding or Contribution
Federal Bureau of Investigation (2012)	Editorial			Provides the users of social media tips on how to mitigate the use of personal information in SE threats.
Federal Bureau of Investigation (2015b)	Editorial			Informs parents on how to discuss social media and its dangers with children.
Heurix et al. (2015)	Descriptive		Anonymity Behavior Cardinality Content Directionality Foundation Holder Identity	"[W]e have presented a taxonomy which covers common aspects of [privacy-enhancing technologies] across different application areas and demonstrated its applicability by applying it on several well-known approaches with different aims, including handling privacy issues with data-at-rest, data-in-motion, and cryptography-based approaches with diverse properties and purposes" (p. 14).

			Pseudonymity	
Hong and Thong (2013)	Empirical	4,000 Internet users	Awareness Control Information Management Interaction Management Internet Privacy Concerns Inter-Web-Personal Multidimensional Development Theory	Four theoretical IPC frameworks, six dimensions of measure, clarification of control in IPC, validation of a third-order factor structure, study of the effects of inconsistent wording in instruments
McCallister et al. (2010)	Descriptive			NIST 800-122
Mitnick and Simon (2002)	Descriptive			Brought social engineering into the mainstream.

			Social engineering attack cycle
Schwartz and Solove (2011)	Exploratory	Identifiable Personal Information Identified Personal Information Unidentifiable Personal Information	"PII 2.0 protects information that relates either to an identified or identifiable person, and associates different legal interests with each category" (p. 1894).

Personally Identifiable Information

Prosch (2008) described the use of accounting principles for protecting PII. The credit card industry self-regulates standards for the handling of PII for financial transactions (PCI Security Standards Council, 2016). Section 5131 of the Information Technology Management Reform Act of 1996 (Public Law 104-106) and the Federal Information Security Management Act of 2002 (Public Law 107-347) provided the foundation for the Federal Information Processing Standards for handling PII and other data: verification of personal identity of employees and contractors (Ferraiolo et al., 2013), verification of personal identity of employees and contractors (Ferraiolo et al., 2013), requirements of using cryptology for non-classified information (Dworkin et al., 2001), classification of all information and information systems (FIPS 199, 2004), minimum security for information and information systems (Ross et al., 2006), digital

signatures (Barker, 2013), secure hash standard (Dang, 2015), and a standard for the use of SHA-3 (Dworkin, 2015).

Ohm (2010) described PII as “an ever-expanding category” (p. 1742). Green (2017) stated, “Humanity produces 2.5 quintillion bytes of data daily.” Schwartz and Solove (2011) described current PII definitions within privacy law to be inconsistent and insufficient. PII “refers to information that can be used to identify or locate an individual” (Chellappa & Sin, 2005, p. 188). Regulators, lawmakers, and organizational policymakers typically view PII as the centroid of privacy issues (Schwartz & Solove, 2011). Peer and Acquisti (2016) discussed the extreme difficulty, if not an impossibility, of reversing the release of PII. The literature indicates that people feel an inability to control their PII (Culnan, 1993; Green, 2017; Palen & Dourish, 2003; Peer & Acquisti, 2016). Simpson (2016) reported that a large number of data breaches occurred, therein containing over a trillion PII via 4,600 data breaches. Privacy Rights Clearinghouse (2018) indicated over 1.9 trillion records had been exposed in 7,300 data breaches as of November 1, 2017. These studies appear to infer that eight trillion records were released in a single year. Though the literature provides details as to the type of breach and the number of affected records, little is known as to what information was released or what specific personal information has been exposed.

PII is the catch-all nomenclature for personal information in much of the literature, regulation, and U.S. law, giving little regard to levels exposure (Schwartz & Solove, 2011). McCallister et al. (2010) associated personal information to measures of risk and harm, thereby indicated that a one-size-fits-all understanding of PII may be ineffective. The elicited feedback from the SMEs for RQ1 and RQ2, should help quantify

PII as well as categorize it to produce a benchmarking prototype for measuring exposure for RQ3. A summary appears in Table 4 of the literature referenced in this section.

Table 4

Summary of Personally Identifiable Information Literature

Study	**Methodology**	**Sample**	**Instruments or Constructs**	**Main Finding or Contribution**
Barker (2013)	Descriptive		Digital Signature Algorithm RSA Digital Signature Elliptic Curve Digital Signature Algorithm	Digital Signature Standard (DSS) (FIPS 186-4)
Culnan (1993)	Survey	126 undergraduate students	Attitudes Toward Direct Mail Marketing Attitudes Toward Secondary Information Use Concern for Privacy Demographics	"60 percent or more of the participants hold negative attitudes toward … practices [involving] one or more of the following: acquisition and use of third-party information, use of financial information, profiling, and/or making inferences that some participants viewed as unwarranted

			or inappropriate" (p. 358).
Dang (2015)	Descriptive	SHA-1 SHA-224 SHA-56 SHA-384 SHA-512 SHA-512/224 SHA-512/256	Secure Hash Standard (SHS) (FIPS PUB 180-4)
Dworkin (2015)	Descriptive		SHA-3 Standard: Permutation-Based Hash and Extendable-Output Functions (FIPS PUB 202)
Ferraiolo et al. (2013)	Descriptive	Personal Identity Verification	FIPS PUB 201-2: Personal Identity Verification (PIV) of Federal Employees and Contractors
Green (2017)	Exploratory	Class Action Data Breach Standing	"Consumer data breach cases appear to satisfy both of these elements [injuries that are non-economic and non-physical], because the harm is broadly diffused throughout the

			economy and some of the injuries alleged are non-economic and non-physical" (p.316).
Ohm (2010)	Exploratory	Anonymization Deanonymize Reidentification	"Easy reidentification … undermines decades of assumptions about robust anonymization, assumptions that have charted the course for business relationships, individual choices, and government regulations…. This Article offers the difficult but necessary way forward: Regulators must use the factors provided to assess the risks of reidentification and carefully balance these risks against countervailing values" (p. 1776).
Palen and Dourish (2003)	Case studies	Disclosure Identity Privacy Publicity	"In offering both a framework and a vocabulary for talking about privacy and technology, our goal is to foster

			Temporality / Time	discussion between technology users, designers and analysts, and to encourage a more nuanced understanding of the impacts of technology on practice" (p. 8).
PCI Security Standards Council (2016)	Descriptive		Account Data Cardholder Data - Cardholder Name - Service Code - Expiration Date Sensitive Authentication Data - Full Track Data - CAV2, CVC2, CVV2, CID - Pin / Pin Block	Payment Card Industry Data Security Standard (PCI DSS)
Peer and Acquisti (2016)		716 adults from Amazon Mechanical Turk and a university pool	Perceived Intrusiveness Self-Disclosure Reversibility Irreversible	Participants disclose more when they are not warned. Perceived intrusiveness increased with the

			prior declaration of reversibility or irreversibility. Perceived intrusiveness rated differently before vs after answering.
Privacy Rights Clearinghouse (2018)	Descriptive	Breach year Eight types of breaches Seven types of organization breached	Tracks and categorizes data breaches
Prosch (2008)	Descriptive	Access Choice and consent Collection Disclosure to third-parties Management Monitoring and enforcement Notice Privacy Lifecycle Maturity Model Quality Security	AICPA Generally Accepted Privacy Principles [adapted from accounting]

		Use and retention	
Ross et al. (2006)	Descriptive		FIPS Publication 200: Minimum Security Requirements for Federal Information and Information Systems
Schwartz and Solove (2011)	Exploratory	Identifiable Personal Information Identified Personal Information Unidentifiable Personal Information	"PII 2.0 protects information that relates either to an identified or identifiable person, and associates different legal interests with each category" (p. 1894).
Simpson (2016)	Exploratory	Common Law of Torts Data Breach Elements of Personally Identifiable Data Importing EU Data Protection into American Law Regulatory Law Statutory Rights	"By adopting an improved definition of personally identifiable data, creating a new definition of data controllers and processors, and reforming statutory liability for data breaches, Americans can be protected, and protect themselves, from the serious risks posed by consumer data breaches both

	now and in the future" (p. 709).

Personally Unidentifiable Information

PUI is defined as "information that, taken alone, cannot be used to identify or locate an individual" (Chellappa & Sin, 2005, p. 188; Federal Trade Commission, 2000, p. 46). Schwartz and Solove (2011) warned that modern technologies make it increasingly difficult to keep PUI as deidentified information. Acquisti and Gross (2009) described an algorithm for predicting social security numbers as well as associated PUI. Additionally, four random pieces of deidentified data from credit card metadata were shown to reidentify 90% of people, with women being easier than men (de Montjoye, Radaelli, Singh, & Pentland, 2015). Kang et al. (2011) described the dangers of modern technologies that people use to surveil portions of their lives. The majority of PUI is intended to provide demographic and nonidentifying information (Schwartz & Solove, 2011). Sweeney (1997) demonstrated the ease of reidentification using only Zip Code, birth date, gender, and race – with only birth date and full ZIP Code required to identify 97% of voters. Benitez and Malin (2010) estimated the difficulty of reidentification when anonymized, classified as public-use, Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule protected data when combined with voter registration lists. Ohm (2010) declares anonymization and the concept of PUI a failure due to the literature showing adeptness in re-identifying individuals even using PUI as a starting point.

Though PUI is typically considered anonymous or deidentified information (Schwartz & Solove, 2011), the literature describes several methodologies for the

reidentification of an individual based on only a few pieces of demographic data (Sweeney, 1997). Rather than sidelining PUI as supposedly anonymous information, SMEs will be asked to assign weights to reflect the level of exposure each has in and of itself. A summary appears in Table 5 of the literature referenced in this section.

Table 5

Summary of Personally Unidentifiable Information

Study	**Methodology**	**Sample**	**Instruments or Constructs**	**Main Finding or Contribution**
Benitez and Malin (2010)	Mixed Methods -Survey of State's Elections Office -Quantitative Analysis	State populations segmented by combinations of County of Residence, Gender, Date of Birth, Race, and Birth Year	Estimated Proportion of a Population in a Group Expected number of Re-Identification General Attacker Monetary Cost of Re-Identification Voter Attacker	"This research provided a set of approaches for estimating the likelihood that de-identified information can be re-identified in the context of data sharing policies associated with the HIPAA Privacy Rule" (p. 177).
Chellappa and Sin (2005)	Empirical study	243 consumers	value of online personalization consumer concern for privacy likelihood of using	"the consumers' value for personalization is almost two times (0.59 vs. −0.34) more influential than the consumers' concern for privacy in determining usage of

			personalization services	personalization services" (p. 197).
de Montjoye et al. (2015)	Field Study	Credit card records of 1.1 million people in 10,000 shops over 3 months.	Price Resolution Risk of reidentification Spatial Resolution Temporal Resolution Unicity	"Our results render the concept of PII, on which the applicability of U.S. and European Union (EU) privacy laws depend, inadequate for metadata data sets…. our findings highlight the need to reform our data protection mechanisms beyond PII and anonymity and toward a more quantitative assessment of the likelihood of reidentification" (p. 539).
Federal Trade Commission (2000)	Descriptive		Access Choice Notice Privacy Privacy Seal Security	FIPS (Privacy Online: Fair Information Practices in the Electronic Marketplace) "Because self-regulatory initiatives to date fall far short of broad-based implementation of self-regulatory programs, the

		Self-regulation	Commission has concluded that such efforts alone cannot ensure that the online marketplace as a whole will follow the standards adopted by industry leaders" (p. ii).
Kang et al. (2011)	Exploratory	Personal Data Stream Personal Data Vault Privacy Rights Management	"Instead of direct behavioral regulation or blind faith in the market, our strategy is to modify indirectly the information ecosystem by introducing a new species, the [Personal Data Guardian]" (p. 847).
Ohm (2010)	Exploratory	Anonymization Deanonymize Reidentification	"Easy reidentification … undermines decades of assumptions about robust anonymization, assumptions that have charted the course for business relationships, individual choices, and government regulations…. This Article offers the difficult but necessary way forward: Regulators must use the factors provided to assess

				the risks of reidentification and carefully balance these risks against countervailing values" (p. 1776).
Schwartz and Solove (2011)	Exploratory		Identifiable Personal Information Identified Personal Information Unidentifiable Personal Information	"PII 2.0 protects information that relates either to an identified or identifiable person, and associates different legal interests with each category" (p. 1894).
Sweeney (1997)	Descriptive	53,033 Voter Records	µ-Argus System Datafly System Scrub System	"What is needed is a rational set of disclosure principles, based on comprehensive analysis of the fundamental issues, which are unlikely to evolve from piecemeal reactions to random incidents" (p. 108).

Social Engineering (SE)

SE "is a combination of techniques used to manipulate victims into divulging confidential information or performing actions that compromise security" (Luo et al., 2013, p. 2). It has been possible to group the SE literature into three main streams: attack vectors (Heartfield & Loukas, 2015; Hong, 2012; Jakobsson, 2016), defense (Conteh &

Schmick, 2016; Mouton et al., 2016; Tetri & Vuorinen, 2013), and the human component (Atkins & Huang, 2013; Krombholz et al., 2013; Luo et al., 2013; Mitnick & Simon, 2002; Workman, 2007). Supporting documentation provides statistics and information as to the number of reported events and the average cost to the victims (Federal Bureau of Investigation, 2015a, 2016). Occasionally, the details of a specific attack are released via the media providing insight into the phenomena (Federal Bureau of Investigation, 2016; Franceschi-Bicchierai, 2015), but this is not the norm as organizations are unwilling to share specifics (Mouton et al., 2016).

The literature typically associates persuasion, deception and exploitation with SE (Harl, 1997; Workman, 2007). Mitnick and Simon (2002) used the elaboration likelihood model to outline SE attack construction, whereas Allen (2006) contrasted the SE attack with the software development cycle. Krombholz et al. (2013) introduced a SE taxonomy. Mouton, Leenen, Malan, and Venter (2014) crafted a SE ontology and discussed the composition of a satisfactory definition, though limiting their own contribution to requiring computer technology. Mouton et al. (2016) provided templates that facilitate SE mitigation and assessment.

Greening (1996) conducted an experiment by which SE was used to obtain valid passwords from many of their 175 student participants. Orgill et al. (2004) used an "auditor" to determine the level of effort necessary to retrieve username and password information through a SE attack. The auditor did not work at the company, though he gained access, dressed similar to their computer department, found a name badge, and then collected usernames and passwords via conversations while walking through the building (Orgill et al., 2004). Hasle, Kristiansen, Kintel, and Snekkenes (2005) performed

a phishing experiment to determine the level of resistance to SE for each of the 120 participants using automation via the Web and email.

Allen (2006) introduced a four-step SE model: information gathering, relationship development, exploitation, execution. Peltier (2006) divided SE into two categories: technology-based and human-based, as well as applied social psychology to SE in the area of persuasion. Peltier (2006) found that gender played a significant role in the SE success. Workman (2007) conducted an empirical study of 588 participants using a questionnaire and observation grounded in threat assessment theory as well as the elaboration likelihood model. Workman (2007) found that trust, friendliness and perceived authority were contributing factors for successful SE attacks. Workman (2008) used cognitive dissonance theory and reactance theory to find how specific personality types can fall prey to SE attacks. Bilge, Strufe, Balzarotti, and Kirda (2009) described the use of automated systems for cloning a social network profile from a single social networking site or across multiple services. Bilge et al. (2009) defeated 215 CAPTCHAs and extracted information from approximately 6000 Web pages and 40,000 profiles each day – culminating with the contact information of five million people as well as the complete profile of 1.2 million people. Bilge et al. (2009) opted to stop their crawlers due to far surpassing their expectations, though they appear to have been able to continue indefinitely.

Chitrey, Singh, and Singh (2012) described the typical motivations for SE attacks: access to proprietary information (30%), financial gain (23%), competitive advantage (21%), enjoyment (11%), revenge (10%), and other (5%). Hong (2012) provided an overview of the composition and execution of phishing, an SE attack vector. Almomani et

al. (2013) provided a literature survey of how detection of phishing emails occurs. Atkins and Huang (2013) codified 100 phishing emails and 100 advanced-fee emails into persuasion categories as well as triggers. Atkins and Huang (2013) found the primary triggers and persuasion techniques used in SE were those that grabbed the attention of target or asked them to verify their account credentials.

Luo et al. (2013) described effective defenses against SE relying heavily on security policy as well as presented an argument that a correlation may exist between personality types and vulnerability to SE. Tetri and Vuorinen (2013) conducted a literature review of 40 journal articles, thereby suggested improvements in research quality and found that very few SE articles were empirical, while the majority were descriptive. Johnston et al. (2015) measured compliance with company policy via an enhanced fear appeal grounded on protection motivation theory and found that informal sanctions provide sufficient influence to raise awareness of security defensiveness. Neupane et al. (2015) conducted a three-dimensional study of phishing detection and warnings. Neupane et al. (2015) found personality types are significant to the success of SE phishing attacks and that people do not spend enough time looking at emails, subsequently failing to detect phishing attacks.

Conteh and Schmick (2016) provided a literature review of phishing research and suggested that repetition in training may improve detection of fake emails and Web sites. Heartfield and Loukas (2015) discussed semantic SE attacks, intentional manipulation of graphical representations to deceive the recipient, and provided a taxonomy to break an attack down to its base components to allow for faster defense through policy, training, and technology systems. Jakobsson (2016) described the compositing and execution of

BEC. Mouton et al. (2016) presented attack templates to provide a methodology to apply other frameworks to SE research.

Mitnick and Simon (2002) brought the human component of SE into a mainstream discussion between technical experts and decision makers with a collection of examples easily understood and communicated by both groups. The idea of the human component being the weakest link continues with researchers looking to internal characteristics and external influences contrasted against specific SE attack vectors (Fan, Lwakatare, & Rong, 2017). Heartfield and Loukas (2015) proposed the need for investigating methodologies to mitigate risk associated with user weakness as well as provided a mechanism to measure user susceptibility to SE, thereby extending the SE attack cycle put forth by Mitnick and Simon (2002). Additionally, the literature has found gender and psychological traits to have significance in successful SE attacks, which is of particular interest to RQ4 (Neupane et al., 2015; Peltier, 2006; Workman, 2008).

The SE literature is primarily explorative and descriptive with very few theoretical or empirical works (Tetri & Vuorinen, 2013). Much of the effort thus far, is a narrow after-the-fact examination of a specific SE attack vector, which may or may not generalize into further research or application (Luo et al., 2013; Mouton et al., 2016; Tetri & Vuorinen, 2013). The literature describes how the simple act of looking at a phishing email for more than a few seconds is enough for the user to accept it as authentic (Neupane et al., 2015; Wenyin et al., 2005). Systems such as BLADE, CANTINA+, and JSAND can be used to filter the harmful effects of phishing emails and BEC (Heartfield & Loukas, 2015), but have little effect on the face-to-face persuasions that the literature indicates people have trouble detecting (Perloff, 2010; Workman, 2007, 2008).

The SE domain has a few noted issues: generalizability (Heartfield & Loukas, 2015; Mouton et al., 2016), applicability (Neupane et al., 2015; Tetri & Vuorinen, 2013; Wenyin et al., 2005), and polarization (Conteh & Schmick, 2016; Junger et al., 2017; Luo et al., 2013; Mitnick & Simon, 2002). Generalization is a major issue in the SE literature in that little is known on who is conducting the SE attack (Heartfield & Loukas, 2015), where exactly the information was obtained (Mouton et al., 2016), or how many times the vector and information were successfully used (Jasper, 2017). The entire attack cycle is specific to the context defined by the persuasion, vector, and susceptibility of the target (Heartfield & Loukas, 2015; Mitnick & Simon, 2002).

The applicability of SE research may have limited effect between contexts as Neupane et al. (2015) noted the significance of personality type has on the success of an attack. Additionally, Tetri and Vuorinen (2013) suggested that functional dimensions of an SE attack are more important than the vector by which it occurred. Polarization within the SE domain is observed when contrasting the literature that stated there is no protection from SE (Conteh & Schmick, 2016; Junger et al., 2017) with those providing insight into the phenomena by providing a means to investigate and measure (Heartfield & Loukas, 2015; Luo et al., 2013; Mitnick & Simon, 2002; Mouton et al., 2016). The literature coalesces on the following assumption: SE attacks are continually increasing in number (Federal Bureau of Investigation, 2015a; Heartfield & Loukas, 2015; Hong, 2012; Tetri & Vuorinen, 2013; Workman, 2008) and the benefit of research has been minimal (Jasper, 2017; Junger et al., 2017; Luo et al., 2013; Mouton et al., 2016). A summary appears in Table 6 of the literature referenced in this section.

Table 6

Summary of Social Engineering Literature

Study	Methodology	Sample	Instruments or Constructs	Main Finding or Contribution
Allen (2006)	Descriptive		The Cycle: Information Gathering Developing Relationship Exploitation Execution	"[T]here will always be the possibility of the 'human factor' being influenced by a social, political and/or cultural event" (p. 9).
Almomani et al. (2013)	Survey		Authentication techniques Client-side tools and filters Network-level protection Server-side filters and classifiers User Education	"This survey improves the understanding of the phishing emails problem, the current solution space, and the future scope to filter phishing emails" (p. 2087).
Atkins and Huang (2013)		100 advanced-fee emails	Incentives	"[A]lert/warning/attention and account verification were the

		100 phishing emails	Persuasion techniques Triggers	two primary triggers used to raise the attention of e-mail recipients…. This study also discovered that social engineers have constructed statements in positive and negative manners to persuade readers to fall victim to their scams" (p. 30).
Bilge et al. (2009)	Descriptive	Used iCloner to clone the profiles of 5 people. The system then contacted 705 distinct people. This process continued until over 1 million people had their profiles completely exposed.	Captcha defeat iCloner Scoring system to determine if multiple accounts on a social media network belong to the same person.	"In this paper, we investigate how easy it would be for a potential attacker to launch automated crawling and identity theft (i.e., cloning) attacks against five popular social networking sites. We present and experimentally evaluate two identity theft attacks" (p. 560). A very high percentage of those contacted from cloned accounts click on "friend" requests.
Chitrey et al. (2012)	Questionnaire	90 responders located in India	Internet Security Awareness Program and Training	Provides data that infers that culturally, people in India have an elevated weakness level to SE attacks.

				New employees, customers, and IT professionals are the most likely targets of SE.
Conteh and Schmick (2016)	Review			"[W]hile technology has a role to play in reducing the impact of social engineering attacks, the vulnerability resides with human behaviour [sic], human impulses and psychological predispositions that can be influenced through education" (p. 37).
Fan et al. (2017)	Exploratory		I-E based model of human weakness for social engineering investigation Psychological states	"We captured two essential levels – [fourteen] internal characteristics of human nature and [nine] external circumstance influences - that shape the human weakness for social engineering" (p. 10).
Federal Bureau of Investigation (2015a)	Report	7,000 companies	business email compromise	"According to IC3, since the beginning of 2015 there has been a 270 percent increase in identified BEC victims" (p. 2).
Federal Bureau of	Report		State-sponsored actors	State-sponsored cyber threats (Iran).

Investigation (2016)				
Franceschi-Bicchierai (2015)	News Article			Describes the attack on CIA Director by teenagers.
Greening (1996)	Simulation of a large-scale SE attack	338 students over 16 days. 175 responded to a phishing e-mail 138 of the responses were valid passwords		Students continued to respond to the e-mail, even after the students were given a second email and formal announcement of the phishing exercise. Very few [61] people attempted to report the attack, and the majority [49] of those complaints were only curious.
Harl (1997)	Editorial		Early work describing SE and the human as the weakest link.	"Contrary to popular belief, it is often easier to hack people than [S]endmail. But it takes far less effort to have employees who can prevent and detect attempts at social engineering than it is to secure any [U]nix system" (p. 5).
Hasle et al. (2005)	Experiment	120 users separated into four groups of	Social Engineering Resistance Metric	"Our experiment shows that it is relatively cheap and easy to mount a large scale [sic] SE attack (or

		30 over 3 days. 59 people were active in that they completed a survey [31] or were presented with a login box [28].		experiment) with a high success rate" (p. 141).
Heartfield and Loukas (2015) 12 citations	Taxonomy	Discusses research with 1900 malicious URLs, 308 users, and other	Deception vector Exploitation Execution Orchestration	"It introduces a structured baseline for classifying semantic attacks by breaking them down into their components" (p. 0:31).
Hong (2012)	Descriptive			"Phishing also causes new problems for organizations, as they blur traditional security perimeters. One's lawyers and accountants may be attacked to surreptitiously gain access to documents. Facebook and other social media provide more contextual details

				that can be used for spear-phishing attacks. An employee falling for a phish in one context may cause a headache for your organization because of reused passwords" (pp. 6-7).
Jakobsson (2016)	Case-studies in chapter format			"The best way to develop and deploy ways to identify and measure the problem and how it changes is to identify not only what the scammers do, but also why…. Understanding why the scammers do what they do, we must also understand their intended victims, what they do—and fail to do" (p. 126).
Jasper (2017)	Review			Discusses Cyber Threat Intelligence Integration Center. "Therefore, the timely sharing of relevant and actionable cyber threat intelligence, in the context of cyber threat information and indicators, is imperative to reducing the impact of attacks" (p. 62).
Johnston et al. (2015)	Sequential mixed-methods	Potential: 2,475 insiders	Compliance intention	"We argue that the reason for these disappointing results [in

	Qualitative via interviews Quantitative via experimental design	Complete responses from 559 insiders of multiple organization within a city government Four organizational leaders were interviewed.	Conventional fear appeal Perceived threat Severity Perceived threat susceptibility Perceived self-efficacy Perceived response efficacy Fear Appears Formal sanction certainty Formal sanction severity Information sanction certainty Informal sanction severity Protection motivation theory Sanction celerity	fear appeals research in information security] from the inadequacy of the conventional fear appeal rhetorical framework and the misspecification of [protection motivation theory] within the information security literature… This study develops and tests an enhanced fear appeal rhetorical framework that accounts for the distinction between threats to information assets and threats to human assets" (p. 130). The enhanced fear appeal framework

Junger et al. (2017)	Experiment	278 participants	Age Age Square Goals System Theory Priming Total Risk Warning	"This study found relatively high disclosure rates… Neither priming nor a warning influenced the degree of disclosure." (p. 85).
Krombholz et al. (2013)	Taxonomy		Channel (How) Operator (What) Social Engineering Taxonomy Type	"[W]e introduced a comprehensive taxonomy to classify social engineering attacks with respect to the attack channel, the operator, different types of social engineering and specific attack scenarios" (p. 34).
Luo et al. (2013)	Descriptive		Personality traits Psychological aspects Social engineering	"in addition to advanced technologies counterattacking various security intrusions, human factors must be equally accounted for" (p. 7).

		Techniques Defenses	
Mitnick and Simon (2002)	Descriptive		Brought social engineering into the mainstream. Social engineering attack cycle
Mouton et al. (2016)	Descriptive	Theory of Group Conformity	Neither the literature or news media provide all the information concerning an attack.
		SE attack Goal Medium	Usually little, if any, information is known about a potential attack.
		Social engineer Target Compliance principles Techniques	Little is known as to where the information is obtained for a SE attack.
		SE framework Preparation Information gathering	Little is known as to what information is available for a SE attack.
		Attack formation Exploit relationship Develop relationship Debrief	Social engineering attack detection model

Neupane et al. (2015)	Experiment	25 participants Malware test (20 randomized trials) -10 warning -10 non-warning -Phishing detection (37 randomized trials) -13 real -12 fake -12 difficult fake	Gaze durations Number of fixations	"[O]ur results showed that users do not spend enough time looking at key phishing indicators and often fail at detecting these attacks, although they may be highly engaged in the task and subconsciously processing real sites differently than fake sites" (p. 489).
Orgill et al. (2004)	Questionnaire	32 participants -26 gave their username -19 gave their password	Department Number Surveyed Password Username	"This study demonstrated that even in a company where security is a concern, these human traits [trust others, assist others, gain favor] can be ill-used if proper preventative measures are not taken … This study also shows the importance of assessing security effectiveness through means such as

		-7 gave login credential information above their own access -4 asked for a name badge or identification		audits.… In order for an audit to be effective, the auditor has to be at least as thorough, through preliminary studying, planning, and execution as a potential social engineer would be" (p. 181). Some departments had more training and resisted the social engineer better.
Peltier (2006)	Review			Magazine article describing SE to readers.
Perloff (2010)	Exploratory		Persuasion	Extensive discussion on persuasion, which is used in many SE attack vectors.
Tetri and Vuorinen (2013)	Descriptive		Actor-Network Theory	Describes issues in SE research and suggests the theories from the psychology literature should only be applied to the persuasion component of SE.
Wenyin et al. (2005)	Exploratory	8 Phishing Web pages 6 Attacked true Web pages	Phishing Visual Similarity Between Two Web Pages	"Preliminary results show that our approach can successfully detect the phishing webpages [sic] with few false alarms for online use" (p. 1061).

		320 authentic home pages of banking institutions	-block level -layout -overall style Web page segmentation	
Workman (2007)	Field study - Questionnaire - Observation	588 participants from a single organization	Affective Commitment Continuance Commitment Normative Commitment Obedience Reactance Subjective Behaviors Threat Severity Trust Vulnerability	Elaboration likelihood model Threat assessment theory "[W]e found that people who are high in normative commitment feel obligated to reciprocate social engineering gestures and favors such as receiving free software or gift certificates by giving up company email addresses, employee identification numbers, financial and insurance data, and other confidential and sensitive information… people who are high in continuance commitment tend to provide information to escalating requests… High affective commitment was also found to contribute to successful social engineering" (pp. 327-328).

				Everyone is susceptible to SE to some degree.
Workman (2008)	Field Study	588 participants from a single U.S. organization	Control for Age, Gender, and Education	"Our investigation has attempted to bridge the theory that explains how people are persuaded through peripheral routes with the social engineering outcomes using an empirical field study in which we investigated whether the factors that account for how people are persuaded in marketing campaigns to make purchases may apply as well to social engineering to give up confidential information" (p. 10).

Theory of Mind (TOM)

The theoretical foundation for this research draws on the theory of mind (TOM). Herbsleb (2005) called for external theories to be used to bring greater understanding to computer science, specifically in software design research. While communicating complex concepts, software designers use anthropomorphic examples, which TOM research indicates is problematic for autistic people (Herbsleb, 2005). The context of the TOM is that an individual "imputes mental states to himself and others" (Premack & Woodruff, 1978, p. 515). Baron-Cohen, Leslie, and Frith (1985) state, "The ability to make inferences about what other people believe to be the case in a given situation allows

one to predict what they will do" (p. 39). Likewise, an individual does not have a TOM when he does not recognize the state of mind of another individual he is interacting with (Premack & Woodruff, 1978). For example, if two people are standing next to the water cooler and one tells a joke, the other person can only have TOM if they perceive the exchange as a joke (Baron-Cohen, 1997). TOM has been used to study chimpanzees (Premack & Woodruff, 1978), children (Baron-Cohen, 1997), autism (Leslie, 1987), and normal adults (Krombholz et al., 2013; Saxe, Schulz, & Jiang, 2006). TOM also offers multiple ingresses into this proposed research study: pretense, representation, pretending, and deception (Baron-Cohen, 1992; Leslie, 1987). Pretense is the intentional distortion of reality (Leslie, 1987), which is used in SE during phishing and other attacks (Mitnick & Simon, 2002). Representation is how an individual views the world (Leslie, 1987).

Workman (2007) described how an individual's representation of an actor might provide trust during a SE attack, even though facts do not fit the reality. Pretending occurs when someone acts as if one thing is real, when he knows that it isn't (Leslie, 1987), which can be observed in many SE attack vectors (Marczak & Paxson, 2017; Tetri & Vuorinen, 2013). Deception involves making someone believe an untruth (Baron-Cohen, 1992) and serves as the primary tool of SE and semantic attacks (Heartfield & Loukas, 2015; Mitnick & Simon, 2002).

Kennedy, Eberhart, and Shi (2001a) supposed that TOM might have inadvertently crept into academia when researchers superimpose their assumptions and abilities to their participants. In the literature, TOM is also used to describe an inability to understand the anthropomorphic descriptions used by software engineers to communicate complex abstract concepts during daily communication, such as a section of code "knowing,"

"seeing," or "dying" (Herbsleb, 2005). Kennedy et al. (2001a) warned that academics should not mistakenly assume a TOM with ordinary people, as not everyone has been trained to seek out explanations for phenomena methodically nor do they embody the expertise of the researcher. Krombholz et al. (2013) noted that the TOM of IS is not shared or even valued by SE attackers, while being used as a weapon against the knowledge workers themselves (Krombholz et al., 2013).

TOM literature endeavors to observe the mind with the understanding that mental states can allow the explanation and prediction of the behavior of others (Premack & Woodruff, 1978). While TOM tends to observe the persuasion (conviction, belief) of a subject, much of SE literature describes the use of persuasion (Mitnick & Simon, 2002; Mouton et al., 2016; Tetri & Vuorinen, 2013) in the commission of attacks. Both SE and TOM literature describe how poorly people detect deception (Krombholz, Hobel, Huber, & Weippl, 2015; Luo et al., 2013; Mitnick & Simon, 2002; Workman, 2008). For example, Saxe et al. (2006) empirically found that participants answering questions concerning a deceptive instrument demonstrated a slower response (mean 2.89 seconds) than false belief questions (mean 2.63 seconds).

The relevance of using TOM as a lens for SE research is supported by Luo et al. (2013), O'keefe (2002), and Peltier (2006). Luo et al. (2013) called for research to investigate how SE attacks can occur due to user participation with OSPI made readily available via social networking sites, thereby empowering deception. Peltier (2006) described the creation of a TOM so that all employees within an organization understand their significance in cyber defense. Keysar, Lin, and Barr (2003) argued that adults fail to associate the beliefs of someone and their actual behavior correctly. O'keefe (2002)

suggested that research move beyond linguistic persuasion and on to visual instruments, such as an instrument that measures exposure to SE, i.e., SEXI, as well as those seen in BEC, phishing, and other SE attacks. A summary appears in Table 7 of the literature referenced in this section.

Table 7

Summary of Theory of Mind Literature

Study	**Methodology**	**Sample**	**Instruments or Constructs**	**Main Finding or Contribution**
Baron-Cohen et al. (1985)	Experiment	61 Children -20 Autistic -14 Down Syndrome -27 Clinically normal	Wimmer and Perner's puppet play paradigm	"The fact that every single child taking part in the experiment correctly answered the control questions allows us to conclude that they all knew (and implicitly believed) that the marble was put somewhere else after Sally had left" (p. 42). "We therefore conclude that the autistic children did not appreciate the difference between their own and the doll's knowledge" (p. 43).

				The ability to know and believe something is separate from having a TOM.
Heartfield and Loukas (2015)	Taxonomy	Discusses research with 1900 malicious URLs, 308 users, and other	Deception Exploitation Execution Orchestration Vector	"It introduces a structured baseline for classifying semantic attacks by breaking them down into their components" (p. 0:31).
Herbsleb (2005)	Exploratory		Behavioral Science Computer Science Interdisciplinary Multidisciplinary	"As a field we have benefited enormously from our borrowings from behavioral science…. We need to continue in this strong interdisciplinary path, and … nurture our own theoretical tradition" (p. 26).
Kennedy et al. (2001a)	Review		Gestalt psychology Suggest an assumed theory of mind amongst researchers.	Provides an overview of the study of the mind.

Keysar et al. (2003)	Two Experiments	38 College students 40 College students (20 male / 20 female)	False belief Hidden object Ignorance	"[T]he ability to take the conceptual perspective of the other is an indispensable element in the fully-developed adult theory of mind. Our findings show that adults do not reliably consult this crucial knowledge about what others know when they interpret what others mean" (p. 37).
Krombholz et al. (2013)	Taxonomy		Social engineering taxonomy Channel (How) Operator (What) Type	"[W]e introduced a comprehensive taxonomy to classify social engineering attacks with respect to the attack channel, the operator, different types of social engineering and specific attack scenarios" (p. 34).
Leslie (1987)	Exploratory		Decoupling model of pretense Metarepresentational theory	"[T]he view advanced here offers for the first time a principled explanation

			Pretend Pretense Representation	for both the peculiarities of pretense and for the existence of these generalizations" (p. 424).
Luo et al. (2013)	Descriptive		Social engineering Psychological aspects Personality traits Techniques Defenses	"in addition to advanced technologies counterattacking various security intrusions, human factors must be equally accounted for" (p. 7).
Marczak and Paxson (2017)	Interviews	30 participants associated with the Middle East and Horn of Africa over two years	Government surveillance Perception of Risk	"Despite the availability of free online tools to check links and attachments, our subject population does not appear to widely use such resources" (p.162).
Mitnick and Simon (2002)	Descriptive			Brought social engineering into the mainstream.

			Social engineering attack cycle
Mouton et al. (2016)	Descriptive	Theory of Group Conformity SE attack -Compliance Principles -Goal -Medium -Social Engineer -Target -Techniques SE Framework - Attack Formation -Debrief -Develop Relationship -Exploit Relationship -Preparation Information Gathering	Neither the literature or news media provide all the information concerning an attack. Usually little, if any, information is known about a potential attack. Little is known as to where the information is obtained for a SE attack. Little is known as to what information is available for a SE attack.

O'keefe (2002)	Review		Attitudes Normative Considerations Self-Efficacy	"Systematic thought about processes of persuasion can be traced back to the ancient Greeks, but as these developments attest, the study of persuasion continues to be a locus of exciting theoretical, empirical, and methodological developments" (p. 40).
Peltier (2006)	Review			Magazine article describing SE to readers.
Premack and Woodruff (1978)	Experiment	Chimpanzee	Problem comprehension	"In assuming that other individuals want, think, believe, and the like, one infers states that are not directly observable and one uses these states anticipatorily, to predict the behavior of others as well as one's own. These inferences, which amount to a theory

				of mind, are, to our knowledge, universal in human adults" (p. 525).
Saxe et al. (2006)	Experiment	12 participants	fMRI brain scans Belief > Photo Stories (for TOM) Incompatible > Compatible response selection Overlap of TOM and Response	"Although they were given the same physical stimuli, and made the same correct responses, when subjects construed their task in terms of belief attribution, they responded faster, and selectively recruited an additional brain region than in the control task" (p. 294). "We found a striking lack of overlap in the brain regions implicated in executive control (specifically response selection and inhibition) and in ToM tasks" (p. 296). TOM (belief attribution) uses entirely different areas of the brain

				than response selection.
Tetri and Vuorinen (2013)	Descriptive		Actor-network theory	Describes issues in SE research and suggests the theories from the psychology literature should only be applied to the persuasion component of SE.
Workman (2007)	Field study -Questionnaire - Observation	588 participants from a single organization	Affective Commitment Continuance Commitment Normative Commitment Obedience Reactance Subjective Behaviors Threat Severity Trust Vulnerability	Elaboration likelihood model Threat assessment theory "[W]e found that people who are high in normative commitment feel obligated to reciprocate social engineering gestures and favors such as receiving free software or gift certificates by giving up company email addresses, employee identification numbers, financial and insurance data,

				and other confidential and sensitive information… people who are high in continuance commitment tend to provide information to escalating requests… High affective commitment was also found to contribute to successful social engineering" (pp. 327-328).
Workman (2008)	Field Study	588 participants from a single U.S. organization	Control for age, gender, and education	"Our investigation has attempted to bridge the theory that explains how people are persuaded through peripheral routes with the social engineering outcomes using an empirical field study in which we investigated whether the factors that account for how people are persuaded in marketing campaigns to make purchases may apply as

		well to social engineering to give up confidential information" (p. 10).

Summary of What is Known and Unknown

A review of various aspects of SE and personal information was conducted to provide a foundation for this study. Through this review of the literature, the constructs of exposure, personal information, and TOM were identified as they relate to social engineering. The literature review describes what is known and unknown about the constructs in this proposed research study. Research regarding SE extended across fields including IS, psychology, law, and business.

SE continues to plague organizations in increasingly alarming amounts (Acquisti et al., 2015; Bélanger & Crossler, 2011). Much of the research into the SE phenomena is primarily explorative and descriptive with limited theoretical or empirical works (Tetri & Vuorinen, 2013; Workman, 2007, 2008). Researchers described efforts thus far as narrow examination of limited details related to a specific SE attack vector, which may or may not generalize into further research or application (Luo et al., 2013; Mouton et al., 2016; Tetri & Vuorinen, 2013). SE literature has offered taxonomy (Heartfield & Loukas, 2015), templates (Mouton et al., 2016), examples of actual attacks (Dadkhah & Quliyeva, 2015; Federal Bureau of Investigation, 2015a, 2016; Krombholz et al., 2013) and occasional empirical research (Neupane et al., 2015; Workman, 2007, 2008).

SE and TOM literature describe how poorly people detect deception (Krombholz et al., 2015; Luo et al., 2013; Saxe et al., 2006; Workman, 2008). In response, the SE

literature has called for a mechanism to provide some level of insight into the available information, which can be weaponized into a cyber attack (Heartfield & Loukas, 2015; Mouton et al., 2016; Peer & Acquisti, 2016; Tetri & Vuorinen, 2013). The privacy literature describes the availability of OSPI via social networks (Acquisti et al., 2015; Greenwood et al., 2016), credit bureaus (Sanders, 2012), personalization (Chellappa & Sin, 2005; Xu et al., 2011), and simple mining programs (Russell, 2013). Similarly, Schwartz and Solove (2011) postulated the enhanced definition of PII to differentiate PUI and PDI would "provide different regimes of regulation for each … standard" (p. 1877) "by considering the applicability of FIPs [Fair Information Practices]" (p. 1879).

TOM is a theory from the psychology literature, which is used to observe the mind with the understanding that mental states can allow the explanation and prediction of the behavior of others (Leslie, 1987; Premack & Woodruff, 1978). Herbsleb (2005) described the unexpected properties of cognitive abilities within computer science where people can fumble through simple tasks while easily completing complicated ones. The literature also indicates that certain personality types (Workman, 2008) and genders are more susceptible to SE (Peltier, 2006). Neupane et al. (2015) found that the possibility of a successful phishing event significantly increased if the target was sleep deprived, distracted, or simply looked at the instrument too long.

The literature has called for an understanding of what information is available and how it can be weaponized into SE attack vectors (Heartfield & Loukas, 2015; Mouton et al., 2016). Disappointingly, the SE literature has not provided the return on the investment originally hoped for (Conteh & Schmick, 2016; Heartfield & Loukas, 2015). Little is known as to the availability of information used in SE or how said information is

obtained and weaponized into attack vectors (Luo et al., 2013; Mouton et al., 2016). Though researchers discussed security policy at length (Acquisti et al., 2016; Bishop & Gates, 2008; Parrish & Nicolas-Rocca, 2012), more study is required to understand the effect of organizational security training on the type and amount of OSPI shared by users in their personal lives (Anderson & Agarwal, 2010; Boss, Galletta, Benjamin Lowry, Moody, & Polak, 2015; Tetri & Vuorinen, 2013). Additionally, little is known as to the specificity of available OSPI (Heartfield & Loukas, 2015; Mouton et al., 2016) and the level of exposure that information poses (Oltmann, 2010). The effects of TOM on the exposure of personal information are also largely not understood (Herbsleb, 2005; Tetri & Vuorinen, 2013).

The constructs of exposure (Keane et al., 1989; Youssef et al., 2013), personal information (Schwartz & Solove, 2011), and TOM (Leslie, 1987) were identified as they relate to SE. Very limited research has explored these constructs within a single study. Therefore, additional research is warranted to examine exposure, personal information, and TOM to determine their contribution to SE.

This research will be used to assess the SE exposure of 100 individuals. The advent of social media, personalization and other technologies has facilitated the exponential increase of available personal information (Acquisti et al., 2015; Mitnick & Simon, 2002). Social engineers have access to OSPI, and a growing concern in SE literature is that the information is being weaponized into SE attack vectors (Heartfield & Loukas, 2015; Mouton et al., 2016; Tetri & Vuorinen, 2013). Because of this phenomenon, users may be exposing themselves and inadvertently the organization that

employs them. Therefore, assessing the exposure, personal information, and TOM of individuals may provide a better understanding of SE.

Chapter 3

Methodology

Introduction

The purpose of this chapter is to detail the research methods used in this study. This research study will be classified as developmental research and include a mixed methods approach. Richey and Klein (2005) stated, “It is not uncommon for a developmental research project to also utilize multiple research methodologies and designs, with different designs again being used for different phases of the project” (p. 31). This research study will comprise a literature review, expert panel feedback via the Delphi method, and quantitative data collection.

Ellis and Levy (2009) stated, “developmental research attempts to answer the question: How can researchers build a ‘thing’ to address the problem? It is especially applicable when there is not an adequate solution to even test for efficacy in addressing the problem...” (p. 328). Salkind (2012) stated that a benefit of developmental research is that it can “describe a particular phenomenon in a way that communicates the overall picture of whatever is being studied. Although these methods do not allow the luxury of implying any cause-and-effect relationship between variables, their use provides the tools needed to answer questions that are otherwise unanswerable” (p. 210). Richey and Klein (2005) stated, “Developmental research seeks to create knowledge grounded in data systematically derived from practice… In addition, it is a way to establish new procedures, techniques, and tools based upon a methodical analysis of specific cases (p. 24).

According to Ellis and Levy (2009), developmental research involves three components: 1) criteria establishment and validation, 2) formal development via accepted process, and 3) determination of criteria satisfaction. Richey and Klein (2005) maintained that developmental research is comprised of a literature review, a Delphi method, and instrument / tool validation. This research study follows the precedence of the body of knowledge with a literature review, Delphi method, and instrument / tool validation to satisfy the Ellis and Levy (2009) three components of developmental research. Figure 4 illustrates the design of this research study.

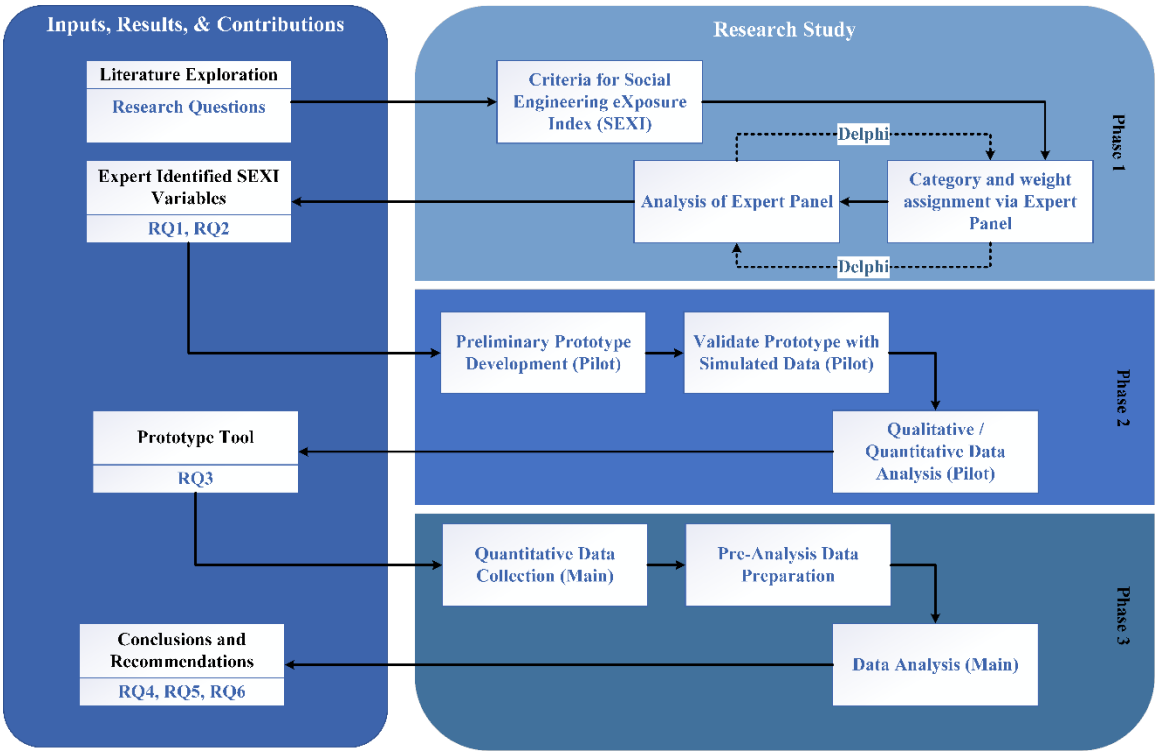

Figure 4: The Proposed Three Phase Development Research Design.

Prior research has utilized a literature review to better understand the information privacy literature (Pavlou, 2011), privacy in the digital age (Bélanger & Crossler, 2011),

and the privacy literature through an interdisciplinary lens (Smith et al., 2011). In this research, a literature review was performed to ascertain the candidate components of personal information as well as to determine a gap requiring further study. Richey and Klein (2005), stated “In developmental research the conceptual framework for the study may be found in literature from actual practice environments (for example, an evaluation report) as well as from traditional research literature directed toward theory construction” (p. 29). A gap was discovered, in that little is known as to the SE attack composition, available personal information, or potential attack vectors (Heartfield & Loukas, 2015; Luo et al., 2013; Mouton et al., 2016; Tetri & Vuorinen, 2013). For this research study, a literature review provided candidate components of personal information for consideration in the SEXI benchmarking instrument, named herein as personal information candidate components (PICCs). Table 8 illustrates the contextuality and ambiguity of personal information as described previously in the literature (Culnan, 1993; Solove, 2006). Table 8 also presents the PICCs categorized as PUI, PII, PDI or generalized in accordance with the respective source and provides the respective label that will be used for analysis..

Table 8

Personal Information Candidate Components by Source with Page Numbers[1]

Label	Item	PUI	PII	PDI	Generalized
PC001	Acceleration via personal tracking				Kang (814)
PC002	Account numbers		McCallister (ES-1)	McCallister (2-2)	PCI DSS (7)
PC003	Activities (daily life)		McCallister (ES-2)		
PC004	Age	Schwartz (1824)	McCallister (A-3)		
PC005	Agency seal / Organizational logo				FIPS 201 (29)
PC006	Alias		McCallister (ES-1)		
PC007	Area code		McCallister (ES-2)		
PC008	Audit log of user actions		McCallister (2-1)		
PC009	Biometric records (retina, iris, voice signature, facial geometry, facial recognition)		McCallister (ES-1)	McCallister (2-1)	FIPS 201 (44) Martin (68)
PC010	Bluetooth connections to other devices				Kang (816)
PC011	Calorie counting with images of food				Kang (815)

[1] Table 8 presents information based on source definition and usage. For example, Schwartz and Solove (2011) placed an item in multiple categories due to context, while McCallister et al. (2010) designated some items as capable of identifying a unique individual and others as not contributing to identification – while categorizing all as PII.

PC012	Cardholder name				PCI DSS (7)
PC013	Cell phone number		McCallister (2-2)		
PC014	Cell tower location				Kang (816)
PC015	Credit card account number		McCallister (ES-1)	Schwartz (1848) McCallister (2-2)	PCI DSS (7)
PC016	Credit card CAV2 / CVC2 / CVV2 / CID				PCI DSS (7)
PC017	Card expiration date				FIPS 201 (27) PCI DSS (7)
PC018	Credit card pin				PCI DSS (7)
PC019	Credit card service code				PCI DSS (7)
PC020	Credit score	McCallister (2-1)			
PC021	Criminal history		McCallister (B-1)		
PC022	Date of birth	Schwartz (1842)	McCallister (ES-2)	McCallister (2-1)	Acquisti (511)
PC023	Demographics	Sweeney (104)			HIPAA (89)
PC024	Driver's license [number]		McCallister (ES-1)	FIPS 201 (9) McCallister (2-2)	
PC025	Education information		McCallister (2-1) Schwartz (1822)		
PC026	Electricity usage				Kang (840) Martin (68)
PC027	Electronic facial image / Selfie		McCallister (ES-1)	McCallister (2-2)	FIPS 201 (39)
PC028	E-mail address		McCallister (ES-1) Schwartz (1857)		
PC029	Employee identification			McCallister (A-1)	
PC030	Employment history		McCallister (B-1)		

PC031	Employment information		McCallister (ES-2)		
PC032	Family income	Schwartz (1851)			
PC033	Favorite movies	Schwartz (1851)			
PC034	Favorite restaurants	Schwartz (1851)			
PC035	Favorite television shows	Schwartz (1851)			
PC036	Financial records / information, balances		McCallister (ES-2)	Schwartz (1882)	
PC037	Fingerprints		McCallister (ES-1)		FIPS 201 (6)
PC038	Fingerprints of two fingers				FIPS 201 (6)
PC039	Full name		McCallister (ES-1) Schwartz (1864)	McCallister (2-1) Schwartz (1848)	Schwartz (1830)
PC040	Full set of fingerprints				FIPS 201 (6)
PC041	Gender	Schwartz (1842)	McCallister (4-5)		Acquisti (513)
PC042	Genetic information	Schwartz (1845)			Kang (840)
PC043	Geographical indicators (location, i.e. city name, latitude, longitude, etc.)		McCallister (ES-2)		
PC044	Global Positioning Systems (GPS)				Kang (840) Martin (68)
PC045	Handwriting		McCallister (ES-1)		
PC046	High school name				Acquisti (511)
PC047	Holographic images (on identification)				FIPS 201 (23)
PC048	Host-specific persistent static identifier (system / host name, etc.)		McCallister (2-2)		
PC049	IP address (network location of network device; dynamic	Schwartz (1838)	McCallister (2-2) Schwartz (1839)		PCI DSS (12) Schwartz (1818)

	/ fixed)				
PC050	Laser etches (on identification)				FIPS 201 (23)
PC051	License plate				Martin (68)
PC052	MAC address (hardware ID of network device)		McCallister (2-2)		
PC053	Maiden name		McCallister (ES-1)	McCallister (2-2)	
PC054	Marital status	Schwartz (1851)			
PC055	Medical history		McCallister (2-2)		
PC056	Medical information	Schwartz (1845)	McCallister (ES-2)		
PC057	Medical test results		McCallister (2-2)		
PC058	Mental health	Schwartz (1824)			HIPAA (89)
PC059	Mother's maiden name		McCallister (ES-1)	McCallister (2-1)	
PC060	Nationality		McCallister (A-3)		
PC061	Newsletter subscription		McCallister (ES-3)		
PC062	Organization affiliation / membership	Schwartz (1851)			FIPS 201 (27)
PC063	Owned property (Mortgage, vehicle Registration, title)	Schwartz (1851)		Schwartz (1882)	
PC064	Parent's middle name		McCallister (3-3)		
PC065	Partner(s) Name		McCallister (3-3)		Acquisti (510)
PC066	Passport number		McCallister (ES-1)	FIPS 201 (9) McCallister (2-1)	
PC067	Password		McCallister (B-4)		PCI DSS (76)
PC068	Patient identification Number			McCallister (2-2)	
PC069	Payment for healthcare				HIPAA (89)

PC070	Persistent Identifier (customer number held in cookie, processor serial number, alphanumeric identifier)		Schwartz (1832)	Schwartz (1855)	
PC071	Personal heart-rate meter				Kang (814)
PC072	Photographic image		McCallister (2-2)		Acquisti (512)
PC073	Physical health				HIPAA (89)
PC074	Place of birth		McCallister (ES-2)	McCallister (2-1)	
PC075	Place of sensing moment				Kang (814)
PC076	Political views		McCallister (3-3)		Acquisti (510)
PC077	Professional title		McCallister (3-5)		
PC078	Provision of healthcare				HIPAA (89)
PC079	Race		McCallister (ES-2)		
PC080	Rank				FIPS 201 (28)
PC081	Recent purchases	Kang (825) Schwartz (1851)		Kang (825) Martin (71)	
PC082	Religion		McCallister (ES-2)		
PC083	Salary information		McCallister (2-2)		
PC084	Search engine query (miscellaneous to vanity)	Schwartz (1847)	Schwartz (1848)	Schwartz (1848)	Acquisti (510)
PC085	Sexual fantasy / behavior		McCallister (3-3)		Acquisti (513) Moon (336)
PC086	Sexual orientation		McCallister (3-3)		Acquisti (510)
PC087	Signature (digital)				FIPS 201 (40)

PC088	Signature (handwritten)				FIPS 201 (28)
PC089	Social media profile				Acquisti (509)
PC090	Social Security Number		McCallister (ES-1) Schwartz (1864)	FIPS 201 (9) McCallister (2-1) Schwartz (1824)	
PC091	Status updates		McCallister (2-1)		Kang (815)
PC092	Street address		McCallister (ES-1)		Schwartz (1830)
PC093	Tax records		McCallister (3-3)		
PC094	Taxpayer identification number		McCallister (ES-1)	McCallister (2-2)	
PC095	Telephone number		McCallister (2-2)		
PC096	Location / Time of sensing moment (self-surveillance via smartphone, fitness device)				Kang (814)
PC097	Timestamp of Web page visit		McCallister (3-3)		
PC098	Uniform Resource Locator (URL) of last Web page		McCallister (3-6)		
PC099	Unique health identifier				HIPAA (191)
PC100	User identification		McCallister (4-8)		
PC101	Web browser history	Schwartz (1858)			
PC102	Weight		McCallister (ES-2)		
PC103	Work phone		McCallister (2-2)		
PC104	X-Rays		McCallister (2-2)		
PC105	ZIP Code	Schwartz (1842)	McCallister (ES-3)		

1

The Delphi method has been used to bring clarification, definition, and enhanced understanding in complex problems, such as the one posed in this research study. The Delphi method has been used to refine a measure of resistance behavior (Rivard & Lapointe, 2012) and to identify information and communication technologies research issues (Lee, 2016). Dalkey and Helmer (1963) provided the following characteristics of the Delphi method, “Its object is to obtain the most reliable consensus of opinion of a group of experts. It attempts to achieve this by a series of intensive questionnaires interspersed with controlled opinion feedback” (p. 458). Delphi research typically consists of anonymity, iteration, controlled feedback, and an aggregated response (von der Gracht, 2012). This research follows the literature by soliciting cybersecurity experts to participate in a Delphi method involving multiple rounds of surveys, thereby providing feedback. Specifically, this research will assess the feedback of the SMEs for the purpose of designing the SEXI benchmarking instrument (Ramim & Lichvar, 2014).

For this research, the stop criteria for the Delphi study will trigger if over 95% of SME responses on PICC identification as one of the DNA, PUI, PII, or PDI across all items in a single round (von der Gracht, 2012). A second stop condition will trigger if there is 15% or less change in the categorization of all the PICCs between two consecutive rouds, thereby reaching stability (Dajani, Sincoff, & Talley, 1979; von der Gracht, 2012). Consensus for this Delphi study is defined as 75% for the PICC items presented to the SMEs with a 1 – 10 scale, as shown in Appendix C, and 80% for items presented to SMEs by exposure category, as shown in Appendix D (Diamond et al., 2014; von der Gracht, 2012).Following Fitch, Bernstein, Aguilar, Burnand, and LaCalle (2001), “the two-round process is designed to sort” the PICCs into three categories of exposure

(p. 5). Schwartz and Solove (2011) declared, “Despite the importance of the concept of PII to privacy law and regulation, there remains a lack of consensus in the United States about how to define it. All current legal models for this concept are flawed” (p. 1835). Therefore, this research will attempt to develop an instrument to measure exposure due to OSPI. Any PICC placed within the same personal information category: PDI, PII, or PUI by at least 75% of SMEs will be included in the SEXI benchmarking instrument. Items not reaching consensus will be accessed on an individual basis.

The main RQ that this study will address is: What are the expert-approved required components comprising an index of exposure to social engineering attacks due to OSPI? Richey and Klein (2005) stated, “research questions, rather than hypotheses, commonly serve as the organizing framework for developmental studies. This tactic is appropriate if there is not a firm base in the literature which one can use as a basis for formulating a hypothesis…, especially if the problem focuses on emerging technologies” (p. 27). This research study comprises six RQs, with RQ1, RQ2, and RQ3 seeking the development of the SEXI benchmarking instrument, while RQ4, RQ5, as well as RQ6 focus on validation. Figure 5 illustrates the primary steps of the Delphi method in this research.

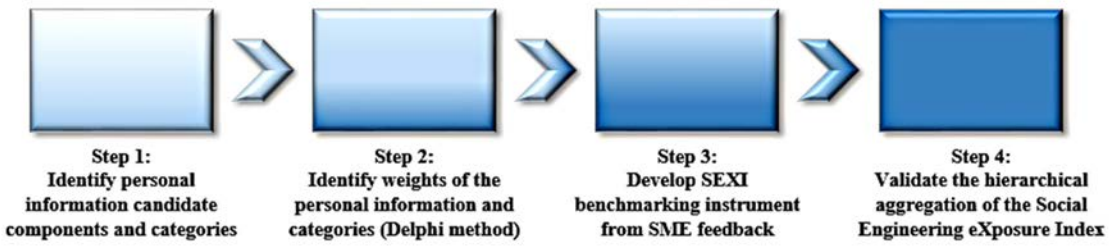

Figure 5: The Proposed Delphi Method Process Culminating in Instrument Validation.

To conduct the Delphi portion of this research, four steps will be required. The first step will involve a review of the literature to ascertain PICCs that are presented to SMEs for them to assess the level of exposure for each component. PICCs with the median SME score of < 1 are designated as not being personal information, those in the 1 – 3 range are categorized as PUI, those in the 4 – 8 range as PII, and those in the 9 – 10 range as PDI. Table 9 presents the classifications of each exposure category.

Table 9

Classification of Exposure Categories with 80% Consensus

Category	Exposure Level	Low Thresh hold	High Thresh hold
DNA	Does Not Apply	0	< 1
PUI	Unidentifiable	≥ 1	< 3
PII	Identifiable	≥ 3	< 9
PDI	Identified	≥ 9	10

The second step is to facilitate iterations of the Delphi method using Internet surveys presenting the PICCs to SMEs for assessment and feedback. Google Forms will host the surveys and function as the data collection platform, while providing the expert panel anonymity. Appendix C presents the first-round survey instrument to be administered to the panel of experts collecting information concerning the work environment, demographic information, and SEXI assessments from the SMEs.

The second survey, presented in Appendix D, will provide the results of the first survey to the SMEs seeking their agreement. This cycle should continue until a stop criterion has been triggered, thereby ending the Delphi process and Phase 1 (see Figure

4). Step 3 and Phase 2 will begin with the construction of the SEXI benchmarking instrument, based on feedback from the SMEs. The contributions of the SMEs will be assessed and reported to address RQ1, RQ2, and RQ3.

The second phase will operationalize a Social Engineering eXposure Index (SEXI) using open source personal information (OSPI). To answer RQ4, this study will attempt to measure the exposure of 50 Fortune 500 executives and 50 Hollywood personas to SE due to OSPI. Data collection will use the SME prescribed SEXI instrument to track the existence of each personal information indicator found, while not collecting any personal information. Appendix F illustrates the data collection instrument that will be used to measure exposure of the executives and personas. Once all data collection has been completed the second phase will conclude. The final phase will involve the analysis and reporting of the data to answer RQ4, RQ5, and RQ6.

The Delphi method allows this study to perform quantitative and qualitative assessments of the SEXI instrument (Creswell, 2012). However, little discernable literature existed at the time of this research addressing exposure to SE due to OSPI. This study will first seek to understand the phenomena. This study will be descriptive in that it endeavors to collect data that describes characteristics of personal exposure using candidate components of personal information, placed into three categories defined herein as PUI, PII, or PDI.

Research Methods

This study is a developmental research approach comprising three phases. Van den Akker et al. (2012) stated, developmental research involves the development of a

prototypical product and “generating methodological directions for the design and evaluation of such products” (p. 4). According to Ellis and Levy (2009), developmental research is “applicable when there is not an adequate solution to even test for efficacy in addressing the problem and presupposes that researchers don’t even know how to go about building a solution that can be tested” (p. 328). Ellis and Levy (2009) concluded that “developmental research attempts to answer the question: How can researchers build a ‘thing’ to address the problem?” (p. 328). Ellis and Levy (2009) described developmental research as consisting of three components, with the first, “establishing and validating criteria the product must meet” (p. 328). Reviewing and establishing the criteria of SEXI from the literature on this topic will meet this component. Second, “follow a formalized, accepted process for developing the product” (Ellis & Levy, 2009, p. 326). This second component will be satisfied by creating a set of criteria from literature to be used to develop the SEXI benchmarking instrument. The third component is “subjecting the product to a formalized, accepted process to determine if it satisfies the criteria” (Ellis & Levy, 2009, p. 326). The third component will be satisfied by the expert panel evaluating SEXI by way of assessing PICCs obtained from literature review and identify the significance of each criterion as PDI, PII, or PUI. The relative importance of each criterion within each measure, along with a relative importance of the measures, will be aggregated to develop the SEXI instrument.

The expert panel will be elicited from the official information security groups and organizations via official social media. Cybersecurity experts who take part in the study will be presented with OSPI properties (i.e. last name, social security number, etc.) and their suggested categories, obtained from the literature review from Acquisti et al. (2015),

Ferraiolo et al. (2013), “HIPAA” (1996), Kang et al. (2011), Martin (2015), McCallister et al. (2010), Moon (2000), Schwartz and Solove (2011), and Sweeney (1997) (see Table 8). The first survey will begin the information privacy iterations by presenting 105 PICCs from the literature to the SMEs. The expert panel will be asked to assign exposure ratings to each personal information indicator as well as exposure categories. The second survey will ask the SMEs to categorize the SME-suggested personal information indicators as well as evaluate those items designated during the first survey as not belonging to personal information. At the conclusion of phase one, phase two will begin with the development of the SEXI instrument based on SME feedback (Ellis & Levy, 2009).

The first phase (see Figure 4) will address RQ1 and RQ2, with the development and evaluation of the SEXI benchmarking instrument to be used to assess 50 executives of Fortune 500 companies and 50 Hollywood personas (a group under constant exposure) via an expert panel using the Delphi expert methodology. Clayton (1997) maintains that group size for Delphi panels should be between 15 – 30 for experts if they share a common discipline and 5 – 10 if they do not necessarily form a statistical population. The expert panel will be elicited from academia and practitioners holding industry certification.

The study will have two surveys. The first survey (see Appendix C) will help understand the composition of the panel of experts and present the initial PICCs as well as collect work environment, background, demographic information, while eliciting feedback on the PICCs from the SMEs. The second survey (see Appendix D) will present the results of the first survey to the SMEs eliciting their agreement with the assessments of the panel expert during the first round.

These surveys will ensure the requirements for the study are met. The first requirement is that each member of the panel of experts shares a TOM. This requirement will be met by evaluating the cybersecurity experience and work environment of the SMEs. The second requirement is an extensive background. This requirement will be met by ensuring respondents have a minimum of seven years of experience in information privacy. The third requirement of the study is that the participants fit within the context of U.S. privacy considerations. This requirement will be met by ensuring each SME has at least one industry-accepted certification. Responses for any panel member not meeting these requirements will be excluded.

In phase two of this research, RQ3 will be addressed with the development of the instrument based on the categorization and weight of PICCs feedback of the SME as well as data collection on a random selection of 50 Fortune 500 executives and 50 Hollywood personas. The SEXI benchmarking instrument will be used to collect data from OSPI sources on 100 individuals denoting the existence, not specifics, of personal information in publicly accessible venues (see Appendix F). Table 10 presents the collection of anonymized data indicating only if the specified information was found and an indicator of where it was found (i.e. FB = Facebook, LN = Linkedin, GS = Google Search).

Table 10

Proposed Data Collection Methodology of Personal Information Participant

Source	Identifier	DOB	Home Address	Postal Code	Picture	Gender
GS	F001-C3	0	1	1	0	0
FB	F001-C3	1	0	0	1	1
LN	F007-C1	1	1	0	1	1
GS	F002-C4	1	1	1	0	1

Phase three of this research study will include both the pre-analysis data screening and the data analysis from the data collected using the SEXI benchmarking instrument (see Figure 4). The results of the data analysis will be used to assess 100 individuals and develop the comparison reports addressing RQ4, RQ5, and RQ6. The comparison report will include graphical representation where appropriate, i.e. from the SEXI aggregation, etc. RQ6 may be of interest as it compares the SEXI of Hollywood persons, which are cataloged by bra size, favorite foods, and home address with the SEXI of executives of Fortune 500 companies with privacy, risk management, and cybersecurity implementations.

Instrument and Measures

Instruments

The proposed research study will follow the developmental methodology in pursuit of a SEXI. The proposed research will elicit responses from an expert panel to assess the validity of criteria content, identify measures, and establish weight allocations based on three sub-measures, each ranging from 0.0 to 1.0: the measurement of personally distinguishable information (PDIM), the measurement of personally identifiable information (PIIM), and the measurement of personally unidentifiable information (PUIM) (McCallister et al., 2010; Schwartz & Solove, 2011).

Two instruments used in this study are supported by literature via a review that found an excess of 105 PICCs in articles by Acquisti et al. (2015), Ferraiolo et al. (2013), “HIPAA” (1996), Kang et al. (2011), Martin (2015), McCallister et al. (2010), Moon (2000), Schwartz and Solove (2011), and Sweeney (1997) (see Table 8). To reduce the number of items presented to the SMEs, identical measures, i.e. demographics from any

source (Sweeney, 1997) and demographics created by or for a healthcare professional ("HIPAA", 1996), were consolidated as demographics. The set of PICCs offered to the SMEs totaled 105, which is presented in Table 8.

The first instrument will collect the assessments of 105 PICCs from a panel of experts via a Delphi method eliciting their opinion on the level of exposure of individual due to a particular PICC, in and of itself. The respective assessments of each SME will identify each PICC as PDI, PII, PUI, DNA, or UNF. In addition, the SMEs will be asked to suggest items that are currently not represented in the list of 105 PICCs. The aggregate assessments of the SMEs will provide the initial weights and categories of each PICC. Following Fitch et al. (2001), the SMEs will be presented each PICC on a scale of 1 to 10, where 1 means minimum exposure of an individual due to the item and 10 means maximum exposure of the individual as the item identifies them. A middle rating of 5 denotes a potential of identification in the PICC. The 1-10 scales will be treated as ordinal scales, and as such the median of the responses from the SMEs will be used rather than the mean (von der Gracht, 2012). This is primarily due to the inability to define distance between points (Linstone & Turoff, 1975). The SMEs will rate the PICCs at least twice via a Delphi method. Subsequent rounds will be added as necessary to reach consensus on each PICC. Linstone and Turoff (1975) discussed similar usage of the Delphi method "to identify and estimate linear weights for those aspects of experience, which they judged to be important in determining the quality of life or sense of well-being of an individual" (p. 383). This study differs from the S.C. Johnson Delphi study described by Linstone and Turoff (1975) as in that study the initial 200-300 components were based on feedback of SMEs, while this research presents 105 PICCs to SMEs from

literature review and elicits additional PICCs from the panel of experts. Following the S.C. Johnson study described by Linstone and Turoff (1975), this research seeks to cluster a large list of components into those having a similar trait (i.e. exposure level). The S.C. Johnson findings “indicated that group relative importance ratings produce reasonable ratio scales, and that the reliability. of such judgments across randomly selected groups is high” (Linstone & Turoff, 1975, p. 383).

The second instrument will present the aggregate groupings of the first instrument to the SMEs. The median values will be used to assign categories to the PICCs, as shown in Table 9. The items in the second instrument will be presented via nominal scale grouped by SME-identified categories (e.g., DNA, PDI, PII, PUI), thereby providing a mechanism for each expert to consider each PICC amongst items in the same category. The SME suggested items from the first instrument will be placed with the PICCs in the category suggested (e.g., DNA, PDI, PII, PUI) and presented to the SMEs. Appendix D provides the second instrument.

Measures

The intent of this research is to develop a single index value (SEXI) that is representative of the exposure to SE due to OSPI as measured by PUI, PII, PDI. Three primary measurements will be used to identify each, in and of itself, PICC: PDI – will *definitively* identify someone, PII – the *potential* of identifying a specific individual, and PUI – having no chance to identify an individual on its own. Two additional non-instrument measurements will be used: the first to designate items the SMEs identify for removal from the lists collected via literature review as not applying to personal information (DNA), as well as a second to designate items as not being familiar to the

respective expert panel member (UNF). A 1-10 scale will be used to assess exposure of each PICC, where 1 indicates minimum exposure and 10 represents maximum exposure. A middle rating of 5 will indicate the item has the potential to identify an individual. PICCs with the median SME score of 0 are designated as not being personal information, those in the 1 – 3 range are categorized as PUI, those in the 4 – 8 range as PII, and those in the 9 – 10 range as PDI (see Table 9). The SME-approved value for each PICC will serve to indicate its component weight. The measurement of each category (i.e. PDIM, PIIM, PUIM) will be the total of the sum of its components multiplied by the SME-identified category weight. Figure 6 illustrates the hierarchical structure from the three measures.

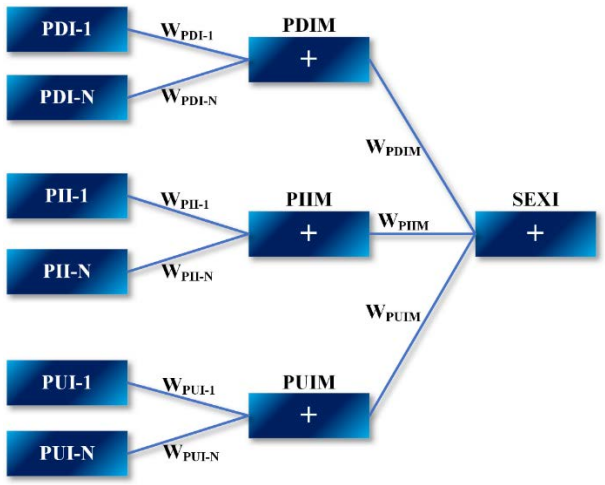

Figure 6: The SEXI Hierarchical Structure: Index, Measures, and Categories.

The SME responses will be used to assess 50 executives of Fortune 500 companies and 50 Hollywood personas by measuring the criteria established by the expert panel. Following Eom and Paek (2009), SEXI is calculated with an additive linear

model. The subsequent equations indicate the computations to be used in the constructs as well as the summation:

Equation 1 presents the PDIM where *i* = the number of PICCs categorized as PDI and PDIM is calculated by multiplying the SME-indicated weight by the existence of a PICC.

$$PDIM = \sum_{i=1}^{n} \mathrm{PDI_i} = \sum_{i=1}^{n} \mathrm{w_i PICC_i} \quad (1)$$

Equation 2 presents the PIIM where *i* = the number of PICCs categorized as PII and PIIM is calculated by multiplying the SME-indicated weight by the existence of a PICC.

$$PIIM = \sum_{i=1}^{n} \mathrm{PII_i} = \sum_{i=1}^{n} \mathrm{w_i PICC_i} \quad (2)$$

Equation 3 presents the PUIM where *i* = the number of PICCs categorized as PUI and PUIM is calculated by multiplying the SME-indicated weight by the existence of a PICC.

$$PUIM = \sum_{i=1}^{n} \mathrm{PUI_i} = \sum_{i=1}^{n} \mathrm{w_i PICC_i} \quad (3)$$

Equation 4 presents a single index value (SEXI) that is representative of the exposure to SE due to OSPI as measured by the sum of PDIM, PIIM, and PUIM each multiplied by their respective SME-indicated category weight.

$$\text{SEXI} = [(\text{W}_{\text{PDI}}PDIM) + (\text{W}_{\text{PII}}PDII) + (\text{W}_{\text{PUI}}PUIM)] \quad (4)$$

Validity and Reliability

An expert panel will evaluate the candidate components of SEXI, following a Delphi technique, derived from prior pertinent literature that described personal information where an individual is unidentifiable, identifiable, and identified (McCallister et al., 2010; Schwartz & Solove, 2011). The PICCs will be presented to the SMEs in a 10-point Likert scale, ranging from 1 (PUI) to 10 (PDI). Items identified as not applying to personal information (DNA) will be reported and removed from the SEXI benchmarking instrument. Feedback from an expert panel using the Delphi expert methodology will provide a weighted value to each item (Ramim & Lichvar, 2014). The instrument used to evaluate SEXI for each executive will utilize nominal scores indicating if exposure is found with a true or false status (Bhattacherjee, 2012; Cohen, 1960). Finally, the TOM of the SMEs will be assessed using nominal and Likert scales to evaluate the privacy practices implemented by SMEs to ensure each meets the requirements of the study (Anderson & Agarwal, 2010; Chellappa & Sin, 2005).

To avoid expert panel bias associated with the topic of privacy, the recruitment of SMEs will not be limited to a single type of industry or government. Privacy has existed in the literature for centuries (Pavlou, 2011) and preconceptions may have been formed by organizational policy (Mouton et al., 2016), legal mandates dictating behaviors and

activities of organizations (Culnan & Williams, 2009; FIPS 199, 2004; McCallister et al., 2010; Ross et al., 2006), as well as industry expectations (Barker, 2013; PCI Security Standards Council, 2016; Ryan & Loeffler, 2010). Tversky and Kahneman (1975) and Lewis (2017) discussed additional bias that may affect expert panels: significance assumed by familiarity, relative significance, imagined significance, and significance associated with frequency. To combat these potential expert panel bias, the list of construct items will be combined and alphabetized before their consideration.

Validity and reliability will be addressed in this proposed research by eliciting the feedback from an expert panel to verify and establish weights used for each item in the first instrument (Ramim & Lichvar, 2014). Mortality is due to participant attrition, subsequently changing the group composition before the study is completed (Salkind, 2012) and is a threat when Delphi expert methodology is used, so a minimum of 15 respondents are necessary for each survey (Clayton, 1997). Testing bias will not be a threat as no pre-test will be administered (Salkind, 2012; Sekaran & Bougie, 2013). To establish instrument validity for this study the content and constructs will be evaluated (Sekaran & Bougie, 2013) and feedback from the panel of experts will be solicited for ensuring SEXI is accurately measuring the exposure to SE (Sekaran & Bougie, 2013; Straub, Boudreau, & Gefen, 2004; Straub, 1989). External validity will be addressed in that the study is not using a contrived setting, thereby being increasingly generalizable (Bhattacherjee, 2012; Sekaran & Bougie, 2013).

Institutional Review Board (IRB) approval was obtained prior to any data collection or Delphi iteration. Appendix A presents the IRB approval letter. The SEXI benchmarking instrument has the potential to acquire PII for each participant via OSPI.

This research will not collect any such information. The purpose of this proposed study is not to collect personal information, but to evaluate the SEXI for each participant. Any personal information obtained through this study will be destroyed.

Proposed Sample

This proposed research is seeking the consensus of 35 SMEs, which satisfies the requirement of literature of 15 – 30 (Clayton, 1997). The resulting instrument will be used to assess the SEXI of 50 top executives of organizations from multiple industries and 50 Hollywood personas using convenience sampling from information gathered via OSPI. Creswell (2012) stated "in convenience sampling the researcher selects participants because they are willing and available to be studied" (p. 167). Sekaran and Bougie (2013) suggested for samples sizes to be between 30 and 500 for most research and noted that the sample size should be at least ten times the number of variables under investigation.

Pre-analysis Data Screening

Mertler and Reinhart (2013) stated pre-analysis screening is mandatory and should be conducted before statistical analysis. The survey questions will use an online research medium, while the SEXI expert survey will elicit binary responses. The results will be examined multiple times for accuracy via Statistical Package for the Social Sciences (SPSS®) (Mertler & Reinhart, 2013). The proper actions will be taken for outliers, missing data, and other anomalies (Mertler & Reinhart, 2013).

Data Analysis

Data analysis will be conducted on each data set. Four types of data analysis will be performed: factorial analysis of variance (ANOVA), factorial analysis of covariance

(ANCOVA), frequencies and percentages, and chi-square tests of independence (Mertler & Reinhart, 2013). Data aggregation will be addressed by providing each participant a unique identifier that will be used to validate the individual's entry.

Each Delphi round will be documented and reported (Linstone & Turoff, 1975). Consensus of the SMEs will be measured by the median of responses addressing RQ1 and RQ2 (Diamond et al., 2014). The level of SME agreement will be reported using the standard deviation and the mean of central tendency (Boone & Boone, 2012). Stability will be measured by comparing the results of two different rounds to evaluate consistency in the median of responses for each PICC (Dajani et al., 1979; von der Gracht, 2012). The significant mean differences of the exposure categories (e.g. PDI, PII, and PUI) will be evaluated by performing one-way ANOVA addressing RQ3 (Boone & Boone, 2012; Norman, 2010). RQ5 will be addressed by performing a ANOVA for each demographic group. RQ6 will be addressed by performing a *t*-test on the two groups: 50 executives of Fortune 500 organizations and 50 Hollywood celebrities (Norman, 2010).

Resources

Academics – to serve on the expert panel.

Computer – to analyze, communicate, and report on the dissertation.

Domain name – to host the project on.

Dissertation chair – to organize and direct the study as needed.

Dissertation committee – to evaluate the dissertation for authenticity and accuracy.

Email – to communicate with an expert panel, participants, dissertation chair/committee.

Executives – to requested participation as well as permission to evaluate their SEXI.

Expert Panel – to establish, validate, and aggregate SEXI criteria.

Industry experts – to serve on the expert panel.

Internet access – to communicate with participants, the university, the expert panel, and access OSPI to assess through SEXI.

IRB – to approve the study in that no people are to be harmed.

Postal Service – to communicate with the dissertation chair/committee.

OSPI – publicly available PII repositories by which to examine participants via SEXI.

Practitioners – to serve on the expert panel.

Secure Website – to conduct surveys and communicate with participants.

Software – IBM SPSS, Smart PLS 3, Microsoft Excel, Microsoft Word, Microsoft Visio, Adobe Photoshop, Email, etc.

SSL certificate – used to encrypt email, content, and Website traffic.

Survey – conducted via online venue.

Milestones

Complete Dissertation Idea Paper	August 2017
Complete Dissertation Proposal 1st draft	December 2017
Complete Dissertation Proposal Final draft	April 2018
Complete expert panel	May 2018
Complete data collection and analysis	August 2018
Complete and defend Dissertation Report	December 2018

Appendix A

Institutional Review Board Approval Letter

NOVA SOUTHEASTERN UNIVERSITY
Institutional Review Board

MEMORANDUM

To: **William Wilkerson**

From: **Ling Wang, Ph.D.,**
Center Representative, Institutional Review Board

Date: **December 8, 2017**

Re: **IRB #: 2017-700; Title, "Development of a Social Engineering eXposure Index (SEXI) using Open Source Personal Information"**

I have reviewed the above-referenced research protocol at the center level. Based on the information provided, I have determined that this study is exempt from further IRB review under **45 CFR 46.101(b) (Exempt Category 2)**. You may proceed with your study as described to the IRB. As principal investigator, you must adhere to the following requirements:

1) CONSENT: If recruitment procedures include consent forms, they must be obtained in such a manner that they are clearly understood by the subjects and the process affords subjects the opportunity to ask questions, obtain detailed answers from those directly involved in the research, and have sufficient time to consider their participation after they have been provided this information. The subjects must be given a copy of the signed consent document, and a copy must be placed in a secure file separate from de-identified participant information. Record of informed consent must be retained for a minimum of three years from the conclusion of the study.

2) ADVERSE EVENTS/UNANTICIPATED PROBLEMS: The principal investigator is required to notify the IRB chair and me (954-262-5369 and Ling Wang, Ph.D., respectively) of any adverse reactions or unanticipated events that may develop as a result of this study. Reactions or events may include, but are not limited to, injury, depression as a result of participation in the study, life-threatening situation, death, or loss of confidentiality/anonymity of subject. Approval may be withdrawn if the problem is serious.

3) AMENDMENTS: Any changes in the study (e.g., procedures, number or types of subjects, consent forms, investigators, etc.) must be approved by the IRB prior to implementation. Please be advised that changes in a study may require further review depending on the nature of the change. Please contact me with any questions regarding amendments or changes to your study.

The NSU IRB is in compliance with the requirements for the protection of human subjects prescribed in Part 46 of Title 45 of the Code of Federal Regulations (45 CFR 46) revised June 18, 1991.

Cc: Yair Levy, Ph.D.
Ling Wang, Ph.D.

3301 College Avenue • Fort Lauderdale, Florida 33314-7796
(954) 262-0000 • 800-672-7223, ext. 5369 • Email: *irb@nova.edu* • Web site: www.nova.edu/irb

Appendix B

Email to Expert Panel: Request for Participation

Dear cybersecurity expert,

We need your help in providing expert feedback on a framework for an upcoming doctoral research study. I am a PhD Candidate in Information Systems with a concentration in Information Security at the College of Engineering and Computing, Nova Southeastern University, working under the supervision of Dr. Yair Levy in the Levy CyLab (https://infosec.nova.edu/cylab/). My research is seeking to develop an index to measure if there is (or to what extent the magnitude exists) exposure to social engineering via publicly available personal information. To develop the index, I need assistance from professionals that have extensive experience dealing with personal information privacy activities, not limited to information security, information privacy, social engineering, law, medical, application development, etc.

You will be asked to complete two surveys. The first survey, should take approximately 20 minutes, will help me to understand your work environment, experience, and will be used to develop the Social Engineering eXposure Index (SEXI) benchmark instrument to assess the level of exposure to social engineering due to publicly available personal information. The second survey, should take approximately 10 minutes, will ask for your feedback on the expert panel aggregate responses from the first-round survey. Your expertise is being solicited to review the proposed measurement criteria for the documented privacy components and provide your expert opinion regarding their relative significance by assigning weights and categories to develop a novel privacy-related exposure measure.

The information provided will be used only for this research study and in aggregated form. Your personal information will not be collected. Your anonymity is assured, and no negative effect will accompany your truthful responses. If you are willing to participate, please click on the link below for access to the first-round survey, to be completed by TBD using password: PASSWORD.

[LinkToSurvey]

Thank you in advance for your consideration. I appreciate your assistance and contribution to this research study. Should you wish to receive the findings of the study, please send me an email, and I will be happy to provide you with information about the academic research publication(s) resulting from this study.

Regards,
W. Shawn Wilkerson, Ph.D. Candidate

E-mail: ww364@nova.edu
Information Systems with a concentration in Information Security
College of Engineering and Computing
Nova Southeastern University

Yair Levy, Ph.D.
E-mail: levyy@nova.edu
Professor of Information Systems and Cybersecurity
College of Engineering and Computing
Nova Southeastern University
Levy CyLab: https://infosec.nova.edu/cylab/

Appendix C

Round I Expert Panel Survey

Dear cybersecurity expert,

Thank you for taking time to participate in this expert panel survey on the exposure to social engineering due to publicly available personal information. In this phase, you will be asked to provide some background information and general demographics. The requested information helps me understand the composition of the expert panel. **It is imperative that your answers are as truthful and honest as possible.** The information provided will be used only for this research study and in aggregated form. **No personal information will be collected. Your anonymity is assured, and no negative effect will accompany your truthful responses.**

This expert panel survey is part of a Ph.D. doctoral dissertation research study that seeks to develop the Social Engineering eXposure Index (SEXI) benchmark instrument to measure exposure to social engineering due to publicly available information. Before this study can move towards the classification of personal information items, I must better understand the composition of experts taking part in the study.

Part 1 – Work Environment. Answer the following questions with the most appropriate answer.

BG01 [Policy] I work for an organization that has a well-defined privacy policy.

○	○	○	○	○	○	○
1 – Strongly Disagree	2 – Disagree	3 – Somewhat Disagree	4 – Neither Agree or Disagree	5 – Somewhat Agree	6 – Agree	7 – Strongly Agree

BG02 [TrainingPrivacy] I work for an organization that has mandatory training for privacy.

○	○	○	○	○	○	○
1 – Strongly Disagree	2 – Disagree	3 – Somewhat Disagree	4 – Neither Agree or Disagree	5 – Somewhat Agree	6 – Agree	7 – Strongly Agree

BG03 [Consequences] I work for an organization that has consequences for violating the privacy policy.

○	○	○	○	○	○	○
1 – Strongly Disagree	2 – Disagree	3 – Somewhat Disagree	4 – Neither Agree or Disagree	5 – Somewhat Agree	6 – Agree	7 – Strongly Agree

BG04 [TrainingSE] I work for an organization that has mandatory social engineering training.

○	○	○	○	○	○	○
1 – Strongly Disagree	2 – Disagree	3 – Somewhat Disagree	4 – Neither Agree or Disagree	5 – Somewhat Agree	6 – Agree	7 – Strongly Agree

BG05 [SecurityAudits] I work for an organization that has security audits.

○	○	○	○	○	○	○
1 – Strongly Disagree	2 – Disagree	3 – Somewhat Disagree	4 – Neither Agree or Disagree	5 – Somewhat Agree	6 – Agree	7 – Strongly Agree

BG06 [Pretending] I work for an organization that has experienced an attempt to gain access to unauthorized assets through someone pretending to be another individual.

○	○	○	○	○	○	○
1 – Strongly Disagree	2 – Disagree	3 – Somewhat Disagree	4 – Neither Agree or Disagree	5 – Somewhat Agree	6 – Agree	7 – Strongly Agree

BG07 [Persuasion] I work for an organization that has experienced an attempt to gain access to unauthorized assets at my organization through persuasion.

○	○	○	○	○	○	○
1 – Strongly Disagree	2 – Disagree	3 – Somewhat Disagree	4 – Neither Agree or Disagree	5 – Somewhat Agree	6 – Agree	7 – Strongly Agree

BG08 [AuthorityBypassPolicy] I work for an organization where someone has the authority to bypass policy on a case-by-case basis.

○	○	○	○	○	○	○

1 – Strongly Disagree	2 – Disagree	3 – Somewhat Disagree	4 – Neither Agree or Disagree	5 – Somewhat Agree	6 – Agree	7 – Strongly Agree

BG09 [UnauthorizedBypassPolicy] I work for an organization where an employee bypassed policy without authorization.

○	○	○	○	○	○	○
1 – Strongly Disagree	2 – Disagree	3 – Somewhat Disagree	4 – Neither Agree or Disagree	5 – Somewhat Agree	6 – Agree	7 – Strongly Agree

BG10 [Repercussion] I work for an organization where an employee bypassed policy without repercussion.

○	○	○	○	○	○	○
1 – Strongly Disagree	2 – Disagree	3 – Somewhat Disagree	4 – Neither Agree or Disagree	5 – Somewhat Agree	6 – Agree	7 – Strongly Agree

BG11 [PrivacyVsEfficiency] I work for an organization where employees feel like they must choose between privacy policy and efficiency.

○	○	○	○	○	○	○
1 – Strongly Disagree	2 – Disagree	3 – Somewhat Disagree	4 – Neither Agree or Disagree	5 – Somewhat Agree	6 – Agree	7 – Strongly Agree

BG12 [PrivacyCulture] I work for an organization where employees are shown ways to bypass policy by other employees.

○	○	○	○	○	○	○
1 – Strongly Disagree	2 – Disagree	3 – Somewhat Disagree	4 – Neither Agree or Disagree	5 – Somewhat Agree	6 – Agree	7 – Strongly Agree

BG13 [Consequence] I work for an organization where violating the privacy policy typically results in:

○	○	○	○	○	○	○

1 – No Consequence	2 – Informal Verbal Warning	3 – Formal Verbal Reprimand	4 – Written Reprimand	5 – Temporary Suspension of Duties	6 – Reassignment	7 – Termination / Legal Issues

Part 2 – Demographics

D01 [Gender] Gender:

1) Male
2) Female

D02 [Age] Age:

1) 19 – 24
2) 25 – 29
3) 30 – 34
4) 35 – 39
5) 40 – 44
6) 45 – 49
7) 50 – 54
8) 55 – 59
9) 60 – 64
10) 65+

D03 [Focus] How would you characterize your work focus?

1) Academia.
2) Mostly academic endeavors with occasional practitioner efforts.
3) Evenly between academic and practitioner efforts.
4) Practitioner.
5) Mostly practitioner endeavors with occasional academic efforts.
6) I am not affiliated with Information Security / Information Privacy.

DO4 [Educ] Please select the highest degree attained

1) Some college credit, no degree earned.
2) Trade/technical/vocational training

3) Associate
4) Bachelors
5) Masters
6) Doctorate

D05 [Certs] Which specialized industry certifications do you currently hold?

- □ [CAP] Certified Authorization Professional
- □ [CCENT] Cisco Certified Entry Networking Technician
- □ [CCEP] Certified Compliance & Ethics Professional
- □ [CCEP-I] Certified Compliance & Ethics Professional-International
- □ [CCFP] Certified Cyber Forensics Professional
- □ [CCSP] Certified Cloud Security Professional
- □ [CEH] Certified Ethical Hacker
- □ [CGEIT] Certified in the Governance of Enterprise IT
- □ [CHC] Certified in Healthcare Compliance
- □ [CHPC] Certified in Healthcare Privacy Compliance
- □ [CHRC] Certified in Healthcare Research Compliance
- □ [CIPM] Certified Information Privacy Manager
- □ [CIPP] Certified Information Privacy Professional
- □ [CIPT] Certified Information Privacy Technologist
- □ [CISA] Certified Information Systems Auditor
- □ [CISM] Certified Information Security Manager
- □ [CISSP] Certified Information Systems Security Professional
- □ [CRISC] Certified in Risk and Information Systems Control
- □ [CSSLP] Certified Secure Software Lifecycle Professional
- □ [CSX] Cybersecurity Nexus Certificate
- □ [CSX-P] Cybersecurity Nexus Certification
- □ [HCISPP] HealthCare Information Security and Privacy Practitioner
- □ [SSCP] Systems Security Certified Practitioner
- □ [OtherCert] Other: _____

D06 [CurrOcc] Current Occupation:

1) Chief Information Officer (CIO)
2) Chief Privacy Officer (CPO)
3) Chief Security Officer (CSO)
4) Chief Information Security Officer (CISO)
5) Consultant
6) IS/IT Professor
7) Law Enforcement
8) Law Professor
9) Privacy Lawyer
10) Privacy Specialist
11) Security Specialist
12) Other _____

D07 [CySecProYrs] Years as a Cybersecurity professional:

1) 1 – 3 Years
2) 4 – 6 Years
3) 7 – 9 Years
4) 10 – 12 Years
5) 13 – 15 Years
6) 16 – 18 Years
7) 19 – 21 Years
8) 22+ Years

D08 [Exp] Years working with information privacy:

1) 1 – 3 Years
2) 4 – 6 Years
3) 7 – 9 Years
4) 10 – 12 Years
5) 13 – 15 Years
6) 16 – 18 Years
7) 19 – 21 Years
8) 22+ Years

D09 [CurOccInd] Current Industry:

1) Banking & Finance
2) Consulting
3) Education
4) Energy
5) Healthcare
6) Government
7) Information Technology
8) Law Enforcement
9) Manufacturing
10) Retail
11) Telecommunication

12) Other

D10 [Mil] Have you ever served in the military?

- o Yes
- o No

D11 [Leo] Have you ever served in law enforcement?

- o Yes
- o No

Part 3 – Items contributing to identification via personal information

Personally identifiable information (PII) is typically thought of as including any personal information. In this section, you will be provided personal information candidate components that have been suggested or described by experts in leading journal articles, federal legislation, and in industry standards.

Read each item and select the best answer indicating where in the range of 1 (minimal exposure) to 10 (maximum exposure) the item, **can in and of itself**, identify a given individual. Select DNA for any item that you feel is not personal information. Select UNF for any item that you are unfamiliar with.

Definitions:

Does not Apply (DNA) – any information that is not personal information.
Unfamiliar (UNF) – any information that you are not familiar with.

	1 Minimum Exposure	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10 Maximum Exposure**	**DNA**	**UNF**
PC001 Acceleration via personal tracking	○	○	○	○	○	○	○	○	○	○	○	○
PC002 Account numbers	○	○	○	○	○	○	○	○	○	○	○	○
PC003 Activities (daily life)	○	○	○	○	○	○	○	○	○	○	○	○
PC004 Age	○	○	○	○	○	○	○	○	○	○	○	○

PC005 Agency seal / Organizational logo	○	○	○	○	○	○	○	○	○	○	○	○
PC006 Alias	○	○	○	○	○	○	○	○	○	○	○	○
PC007 Area code	○	○	○	○	○	○	○	○	○	○	○	○
PC008 Audit log of user actions	○	○	○	○	○	○	○	○	○	○	○	○
PC009 Biometric records (retina, iris, voice signature, facial geometry, facial recognition)	○	○	○	○	○	○	○	○	○	○	○	○
PC010 Bluetooth connections to other devices	○	○	○	○	○	○	○	○	○	○	○	○
PC011 Calorie counting with images of food	○	○	○	○	○	○	○	○	○	○	○	○
PC012 Cardholder name	○	○	○	○	○	○	○	○	○	○	○	○

1

	1 Minimum Exposure	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10 Maximum Exposure**	**DNA**	**UNF**
PC013 Cell phone number	○	○	○	○	○	○	○	○	○	○	○	○
PC014 Cell tower location	○	○	○	○	○	○	○	○	○	○	○	○
PC015 Credit card account number	○	○	○	○	○	○	○	○	○	○	○	○
PC016 Credit card CAV2 / CVC2 / CVV2 / CID	○	○	○	○	○	○	○	○	○	○	○	○
PC017 Card expiration date	○	○	○	○	○	○	○	○	○	○	○	○
PC018 Credit card pin	○	○	○	○	○	○	○	○	○	○	○	○
PC019 Credit card service code	○	○	○	○	○	○	○	○	○	○	○	○
PC020 Credit score	○	○	○	○	○	○	○	○	○	○	○	○

PC021 Criminal history	○	○	○	○	○	○	○	○	○	○	○	○
PC022 Date of birth	○	○	○	○	○	○	○	○	○	○	○	○
PC023 Demographics	○	○	○	○	○	○	○	○	○	○	○	○
PC024 Driver's license [number]	○	○	○	○	○	○	○	○	○	○	○	○

1

	1 Minimum Exposure	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10 Maximum Exposure**	**DNA**	**UNF**
PC025 Education information	○	○	○	○	○	○	○	○	○	○	○	○
PC026 Electricity usage	○	○	○	○	○	○	○	○	○	○	○	○
PC027 Electronic facial image / selfie	○	○	○	○	○	○	○	○	○	○	○	○
PC028 E-mail address	○	○	○	○	○	○	○	○	○	○	○	○
PC029 Employee identification	○	○	○	○	○	○	○	○	○	○	○	○
PC030 Employment history	○	○	○	○	○	○	○	○	○	○	○	○
PC031 Employment information	○	○	○	○	○	○	○	○	○	○	○	○
PC032 Family income	○	○	○	○	○	○	○	○	○	○	○	○
PC033 Favorite movies	○	○	○	○	○	○	○	○	○	○	○	○
PC034 Favorite restaurants	○	○	○	○	○	○	○	○	○	○	○	○
PC035 Favorite television shows	○	○	○	○	○	○	○	○	○	○	○	○
PC036 Financial records / information, balances	○	○	○	○	○	○	○	○	○	○	○	○

2

	1 Minimum Exposure	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10 Maximum Exposure**	**DNA**	**UNF**
PC037 Fingerprints	○	○	○	○	○	○	○	○	○	○	○	○

PC038 Fingerprints of two fingers	○	○	○	○	○	○	○	○	○	○	○	○
PC039 Full name	○	○	○	○	○	○	○	○	○	○	○	○
PC040 Full set of fingerprints	○	○	○	○	○	○	○	○	○	○	○	○
PC041 Gender	○	○	○	○	○	○	○	○	○	○	○	○
PC042 Genetic information	○	○	○	○	○	○	○	○	○	○	○	○
PC043 Geographical indicators (location, i.e. city name, latitude, longitude, etc.)	○	○	○	○	○	○	○	○	○	○	○	○
PC044 Global Positioning Systems (GPS)	○	○	○	○	○	○	○	○	○	○	○	○
PC045 Handwriting	○	○	○	○	○	○	○	○	○	○	○	○
PC046 High school name	○	○	○	○	○	○	○	○	○	○	○	○
PC047 Holographic images (on identification)	○	○	○	○	○	○	○	○	○	○	○	○
PC048 Host-specific persistent static identifier (system / hostname, etc.)	○	○	○	○	○	○	○	○	○	○	○	○

1

	1 Minimum Exposure	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10 Maximum Exposure**	**DNA**	**UNF**
PC049 IP address (network location of network device; dynamic / fixed)	○	○	○	○	○	○	○	○	○	○	○	○
PC050 Laser etches (on identification)	○	○	○	○	○	○	○	○	○	○	○	○
PC051 License plate	○	○	○	○	○	○	○	○	○	○	○	○
PC052 MAC address (hardware ID of network device)	○	○	○	○	○	○	○	○	○	○	○	○
PC053 Maiden name	○	○	○	○	○	○	○	○	○	○	○	○
PC054 Marital status	○	○	○	○	○	○	○	○	○	○	○	○
PC055 Medical history	○	○	○	○	○	○	○	○	○	○	○	○

PC056 Medical information	○	○	○	○	○	○	○	○	○	○	○	○
PC057 Medical test results	○	○	○	○	○	○	○	○	○	○	○	○
PC058 Mental health	○	○	○	○	○	○	○	○	○	○	○	○
PC059 Mother's maiden name	○	○	○	○	○	○	○	○	○	○	○	○
PC060 Nationality	○	○	○	○	○	○	○	○	○	○	○	○

1

	1 Minimum Exposure	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10 Maximum Exposure**	**DNA**	**UNF**
PC061 Newsletter subscription	○	○	○	○	○	○	○	○	○	○	○	○
PC062 Organization affiliation / membership	○	○	○	○	○	○	○	○	○	○	○	○
PC063 Owned property (mortgage, vehicle registration, title)	○	○	○	○	○	○	○	○	○	○	○	○
PC064 Parent's middle name	○	○	○	○	○	○	○	○	○	○	○	○
PC065 Partner(s) name	○	○	○	○	○	○	○	○	○	○	○	○
PC066 Passport number	○	○	○	○	○	○	○	○	○	○	○	○
PC067 Password	○	○	○	○	○	○	○	○	○	○	○	○
PC068 Patient identification number	○	○	○	○	○	○	○	○	○	○	○	○
PC069 Payment for health care	○	○	○	○	○	○	○	○	○	○	○	○
PC070 Persistent Identifier (customer number held in cookie, processor serial number, alphanumeric identifier)	○	○	○	○	○	○	○	○	○	○	○	○

2

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**D**	**U**

	Minimum Exposure									Maximum Exposure	NA	NF
PC071 Personal heart-rate meter	○	○	○	○	○	○	○	○	○	○	○	○
PC072 Photographic image	○	○	○	○	○	○	○	○	○	○	○	○
PC073 Physical health	○	○	○	○	○	○	○	○	○	○	○	○
PC074 Place of birth	○	○	○	○	○	○	○	○	○	○	○	○
PC075 Place of sensing moment	○	○	○	○	○	○	○	○	○	○	○	○
PC076 Political views	○	○	○	○	○	○	○	○	○	○	○	○
PC077 Professional title	○	○	○	○	○	○	○	○	○	○	○	○
PC078 Provision of health care	○	○	○	○	○	○	○	○	○	○	○	○
PC079 Race	○	○	○	○	○	○	○	○	○	○	○	○
PC080 Rank	○	○	○	○	○	○	○	○	○	○	○	○
PC081 Recent purchases	○	○	○	○	○	○	○	○	○	○	○	○
PC082 Religion	○	○	○	○	○	○	○	○	○	○	○	○

1

	1 Minimum Exposure	2	3	4	5	6	7	8	9	10 Maximum Exposure	DNA	UNF
PC083 Salary information	○	○	○	○	○	○	○	○	○	○	○	○
PC084 Search engine query (miscellaneous to vanity)	○	○	○	○	○	○	○	○	○	○	○	○
PC085 Sexual fantasy / behavior	○	○	○	○	○	○	○	○	○	○	○	○
PC086 Sexual orientation	○	○	○	○	○	○	○	○	○	○	○	○
PC087 Signature (digital)	○	○	○	○	○	○	○	○	○	○	○	○
PC088 Signature (handwritten)	○	○	○	○	○	○	○	○	○	○	○	○
PC089 Social media profile	○	○	○	○	○	○	○	○	○	○	○	○

PC090 Social Security Number	○	○	○	○	○	○	○	○	○	○	○	○
PC091 Status updates	○	○	○	○	○	○	○	○	○	○	○	○
PC092 Street address	○	○	○	○	○	○	○	○	○	○	○	○
PC093 Tax records	○	○	○	○	○	○	○	○	○	○	○	○
PC094 Taxpayer identification number	○	○	○	○	○	○	○	○	○	○	○	○
PC095 Telephone number	○	○	○	○	○	○	○	○	○	○	○	○
PC096 Location / Time of sensing moment (self-surveillance via smartphone, fitness device)												○

1

	1 Minimum Exposure	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10 Maximum Exposure**	**DNA**	**UNF**
PC097 Timestamp of Web page visit	○	○	○	○	○	○	○	○	○	○	○	○
PC098 Uniform Resource Locator (URL) of last Web page	○	○	○	○	○	○	○	○	○	○	○	○
PC099 Unique health identifier	○	○	○	○	○	○	○	○	○	○	○	○
PC100 User identification	○	○	○	○	○	○	○	○	○	○	○	○
PC101 Web browser history	○	○	○	○	○	○	○	○	○	○	○	○
PC102 Weight	○	○	○	○	○	○	○	○	○	○	○	○
PC103 Work phone	○	○	○	○	○	○	○	○	○	○	○	○
PC104 X-Rays	○	○	○	○	○	○	○	○	○	○	○	○
PC105 ZIP Code	○	○	○	○	○	○	○	○	○	○	○	○

2

3

4

Part 4 – Provide any suggestions for Personally Unidentifiable Information (PUI) not in the personal information candidate components above. If you have no additional items, please enter NA.

Part 5 – Provide any suggestions for Personally Identifiable Information (PII) not in the personal information candidate components above. If you have no additional items, please enter NA

Part 6 – Provide any suggestions for Personally Distinguishable Information (PDI) not in the personal information candidate components above. If you have no additional items, please enter NA.

Part 7 – Category Weight Assignment

The three proposed measures will be assessed based on the clusters of criteria identified by the expert panel. What should the importance of each category be relative to the other categories?

Please allocate from 1 -100 points in each of the Social Engineering eXposure Index (SEXI) categories (all 100 points should be used):

Personally unidentifiable information (PUI) – any information that *cannot identify* an individual by itself.	W_{PUI} [ ]
Personally identifiable information (PII) – any information that can *potentially identify* an individual by itself and not be PDI or PUI.	W_{PII} [ ]
Personally distinguishable information (PDI) – any information that can *definitely identify* an individual by itself.	W_{PDI} [ ]

Appendix D

Round II Expert Panel Survey

Dear cybersecurity expert,

Thank you for taking time to participate in this expert panel survey on the exposure to social engineering due to publicly available personal information. In this phase, you will be asked to provide feedback on the placement of the personal information components by a panel of experts. The information provided will be used only for this research study and in aggregated form. **No personal information will be collected. Your anonymity is assured, and no negative effect will accompany your truthful responses.**

This expert panel survey is part of a Ph.D. doctoral dissertation research study that seeks to develop the Social Engineering eXposure Index (SEXI) benchmark instrument to measure exposure to social engineering due to publicly available information.

Categories:

Personally unidentifiable information (PUI) – any information that *cannot identify* an individual by itself.

Personally identifiable information (PII) – any information that can *potentially identify* an individual by itself and not be PDI or PUI.

Personally distinguishable information (PDI) – any information that can *definitely identify* an individual by itself.

Does not Apply (DNA) – any information that is not personal information.

Please read over the following lists and indicate the group the personal information item belongs in

Part 1 – Items the expert panel selected as not being personal information.

	Does not Apply (DNA)	**Cannot Identify** (PUI)	**Potentially Identify** (PII)	**Definitely Identify** (PDI)
PC001 Acceleration via personal tracking	○	○	○	○
PC002 Account numbers	○	○	○	○
PC003 Activities (daily life)	○	○	○	○

PC004 Age	○	○	○	○
PC005 Agency seal / Organizational logo	○	○	○	○
PC006 Alias	○	○	○	○
PC007 Area code	○	○	○	○
PC008 Audit log of user actions	○	○	○	○
PC009 Biometric records (retina, iris, voice signature, facial geometry, facial recognition)	○	○	○	○
PC010 Bluetooth connections to other devices	○	○	○	○
PC011 Calorie counting with images of food	○	○	○	○
PC012 Cardholder name	○	○	○	○
PC013 Cell phone number	○	○	○	○
PC014 Cell tower location	○	○	○	○
PC015 Credit card account number	○	○	○	○
PC016 Credit card CAV2 / CVC2 / CVV2 / CID	○	○	○	○
PC017 Card expiration date	○	○	○	○
PC018 Credit card pin	○	○	○	○
PC019 Credit card service code	○	○	○	○
PC020 Credit score	○	○	○	○

Part 2 – Items the expert panel designated as personal information that *cannot identify* an individual by itself.

	Does not Apply (DNA)	**Cannot Identify** (PUI)	**Potentially Identify** (PII)	**Definitely Identify** (PDI)
PC021 Criminal history	○	○	○	○
PC022 Date of birth	○	○	○	○

PC023 Demographics	○	○	○	○
PC024 Driver's license [number]	○	○	○	○
PC025 Education information	○	○	○	○
PC026 Electricity usage	○	○	○	○
PC027 Electronic facial image / selfie	○	○	○	○
PC028 E-mail address	○	○	○	○
PC029 Employee identification	○	○	○	○
PC030 Employment history	○	○	○	○
PC031 Employment information	○	○	○	○
PC032 Family income	○	○	○	○
PC033 Favorite movies	○	○	○	○
PC034 Favorite restaurants	○	○	○	○
PC035 Favorite television shows	○	○	○	○
PC036 Financial records / information, balances	○	○	○	○
PC037 Fingerprints	○	○	○	○
PC038 Fingerprints of two fingers	○	○	○	○
PC039 Full name	○	○	○	○
PC040 Full set of fingerprints	○	○	○	○

Part 3 – Items the expert panel designated as personal information that has *the potential* to identify an individual by itself that are not PDI or PUI.

	Does not Apply	Cannot Identify	Potentially Identify	Definitely Identify

	(DNA)	(PUI)	(PII)	(PDI)
PC041 Gender	○	○	○	○
PC042 Genetic information	○	○	○	○
PC043 Geographical indicators location, i.e. city name, latitude, longitude, etc.)	○	○	○	○
PC044 Global Positioning Systems (GPS)	○	○	○	○
PC045 Handwriting	○	○	○	○
PC046 High school name	○	○	○	○
PC047 Holographic images (on identification)	○	○	○	○
PC048 Host-specific persistent static identifier (system / hostname, etc.)	○	○	○	○
PC049 IP address (network location of network device; dynamic / fixed)	○	○	○	○
PC050 Laser etches (on identification)	○	○	○	○
PC051 License plate	○	○	○	○
PC052 MAC address (hardware ID of network device)	○	○	○	○
PC053 Maiden name	○	○	○	○
PC054 Marital status	○	○	○	○
PC055 Medical history	○	○	○	○
PC056 Medical information	○	○	○	○
PC057 Medical test results	○	○	○	○
PC058 Mental health	○	○	○	○

PC059 Mother's maiden name	○	○	○	○
PC060 Nationality	○	○	○	○

Part 4 – Items the expert panel designated as personal information that *can definitely identify* an individual by itself.

	Does not Apply (DNA)	**Cannot Identify** (PUI)	**Potentially Identify** (PII)	**Definitely Identify** (PDI)
PC091 Status updates	○	○	○	○
PC092 Street address	○	○	○	○
PC093 Tax records	○	○	○	○
PC094 Taxpayer identification number	○	○	○	○
PC095 Telephone number	○	○	○	○
PC096 Location / Time of sensing moment (self-surveillance via smartphone, fitness device)	○	○	○	○
PC097 Timestamp of Web page visit	○	○	○	○
PC098 Uniform Resource Locator (URL) of last Web page	○	○	○	○
PC099 Unique health identifier	○	○	○	○
PC100 User identification	○	○	○	○
PC101 Web browser history	○	○	○	○
PC102 Weight	○	○	○	○
PC103 Work phone	○	○	○	○
PC104 X-Rays	○	○	○	○
PC105 ZIP Code	○	○	○	○

Part 5 – Expert Panel Suggested Items.

	Does not Apply (DNA)	**Cannot Identify** (PUI)	**Potentially Identify** (PII)	**Definitely Identify** (PDI)
SME001	○	○	○	○
SME002	○	○	○	○
SME003	○	○	○	○
SME004	○	○	○	○
SME005	○	○	○	○
SME006	○	○	○	○
SME007	○	○	○	○
SME008	○	○	○	○
SME009	○	○	○	○
SME010	○	○	○	○

Appendix E

Privacy Clearinghouse Data Breach Raw Data (2005-2017)[2]

	Payment Card Fraud	Disclosure	Hacking / Malware	Insider	Physical Loss	Portable Device	Stationary Device	Unknown	Total
2005	0	442,908	42,712,843	2,070,710	39,100	9,336,157	493,523	6,000	55,101,241
2006	1,440	4,604,453	3,110,964	1,437,024	2,263,910	33,641,957	3,929,175	19,591,788	68,580,711
2007	100	4,056,857	107,651,523	18,423,520	82,781	10,366,890	611,438	264,798	141,457,907
2008	1,529	84,874,221	6,863,468	18,224,157	67,809	19,050,275	958,093	738,348	130,777,900
2009	40,000	267,686	169,569,882	142,041	194,255	80,897,187	367,664	97,099	251,575,814
2010	25,444	2,645,814	22,198,121	1,429,042	5,426,302	8,646,522	313,564	100,236,082	140,920,891
2011	100,499	4,615,339	412,797,321	139,352	10,884,976	12,398,500	4,978,293	1,987,097	447,901,377
2012	7,027,673	9,543,749	208,923,746	577,885	1,307,772	2,806,846	147,243	68,431,841	298,766,755
2013	6,350	7,181,965	132,888,609	3,351,500	6,418,965	1,744,444	4,254,970	2,939,297	158,786,100
2014	0	11,343,897	1,292,186,207	821,010	8,625,255	107,090	181,014	372,779	1,313,637,252
2015	0	6,736,248	310,203,735	8,214	903,846	242,694	0	700,000	318,794,737
2016	0	194,278,871	4,579,253,932	28,321,442	2,175,247	6,405,395	955	100	4,810,435,942
2017	2,000,000	1,614,782,157	299,629,425	85,545	1,053,424	6,936	0	7,500,000	1,925,057,487
	9,203,035	1,945,374,165	7,587,989,776	75,031,442	39,443,642	185,650,893	16,235,932	202,865,229	10,061,794,114

[2] Data was retrieved and verified January 2018, from the source code of the respective Web page as it did not match display.

Appendix F

SEXI Data Collection Form

M081-03 (Nondescript identifier)

Label	Item	SRC1	SRC2	SRC3
PC001	Acceleration via personal tracking	□	□	□
PC002	Account numbers	□	□	□
PC003	Activities (daily life)	□	□	□
PC004	Age	□	□	□
PC005	Agency seal / Organizational logo	□	□	□
PC006	Alias	□	□	□
PC007	Area code	□	□	□
PC008	Audit log of user actions	□	□	□
PC009	Biometric records (retina, iris, voice signature, Facial geometry, facial recognition)	□	□	□
PC010	Bluetooth connections to devices	□	□	□
PC011	Calorie counting w/ images of food	□	□	□
PC012	Cardholder name	□	□	□
PC013	Cell phone number	□	□	□
PC014	Cell tower location	□	□	□
PC015	Credit card account number	□	□	□
PC016	Credit card CAV2 / CVC2 / CVV2 / CID	□	□	□
PC017	Card expiration date	□	□	□
PC018	Credit card pin	□	□	□
PC019	Credit card service code	□	□	□
PC020	Credit score	□	□	□
PC021	Criminal history	□	□	□
PC022	Date of birth	□	□	□
PC023	Demographics	□	□	□
PC024	Driver's license [number]	□	□	□
PC025	Education information	□	□	□
PC026	Electricity usage	□	□	□
PC027	Electronic facial image / Selfie	□	□	□
PC028	E-mail address	□	□	□
PC029	Employee identification	□	□	□
PC030	Employment history	□	□	□
PC031	Employment information	□	□	□

PC032	Family income	□	□	□
PC033	Favorite movies	□	□	□
PC034	Favorite restaurants	□	□	□
PC035	Favorite television shows	□	□	□
PC036	Financial records / information, balances	□	□	□
PC037	Fingerprints	□	□	□
PC038	Fingerprints of two fingers	□	□	□
PC039	Full name	□	□	□
PC040	Full set of fingerprints	□	□	□
PC041	Gender	□	□	□
PC042	Genetic information	□	□	□
PC043	Geographical indicators (location, i.e. city name, latitude, longitude, etc.)	□	□	□
PC044	GPS	□	□	□
PC045	Handwriting	□	□	□
PC046	High school name	□	□	□
PC047	Holographic images (on ID)	□	□	□
PC048	Host-specific persistent static identifier (system / hostname, etc.)	□	□	□
PC049	IP address	□	□	□
PC050	Laser etches (on ID)	□	□	□
PC051	License plate	□	□	□
PC052	MAC address	□	□	□
PC053	Maiden name	□	□	□
PC054	Marital status	□	□	□
PC055	Medical history	□	□	□
PC056	Medical information	□	□	□
PC057	Medical test results	□	□	□
PC058	Mental health	□	□	□
PC059	Mother's maiden name	□	□	□
PC060	Nationality	□	□	□
PC061	Newsletter subscription	□	□	□
PC062	Organization affiliation / membership	□	□	□
PC063	Owned property	□	□	□
PC064	Parent's middle name	□	□	□
PC065	Partner(s) Name	□	□	□
PC066	Passport number	□	□	□
PC067	Password	□	□	□
PC068	Patient identification Number	□	□	□
PC069	Payment for health care	□	□	□
PC070	Persistent Identifier (customer number held in cookie, processor serial number, alphanumeric	□	□	□

	identifier)			
PC071	Personal heart-rate meter	□	□	□
PC072	Photographic image	□	□	□
PC073	Physical health	□	□	□
PC074	Place of birth	□	□	□
PC075	Place of sensing moment	□	□	□
PC076	Political views	□	□	□
PC077	Professional title	□	□	□
PC078	Provision of health care	□	□	□
PC079	Race	□	□	□
PC080	Rank	□	□	□
PC081	Recent purchases	□	□	□
PC082	Religion	□	□	□
PC083	Salary information	□	□	□
PC084	Search engine query (miscellaneous to vanity)	□	□	□
PC085	Sexual fantasy / behavior	□	□	□
PC086	Sexual orientation	□	□	□
PC087	Signature (digital)	□	□	□
PC088	Signature (handwritten)	□	□	□
PC089	Social media profile	□	□	□
PC090	Social Security Number	□	□	□
PC091	Status updates	□	□	□
PC092	Street address	□	□	□
PC093	Tax records	□	□	□
PC094	Taxpayer identification number	□	□	□
PC095	Telephone number	□	□	□
PC096	Location / Time of sensing moment (self-surveillance via smartphone, fitness device)	□	□	□
PC097	Timestamp of Web page visit	□	□	□
PC098	Uniform Resource Locator (URL) of last Web page	□	□	□
PC099	Unique health identifier	□	□	□
PC100	User identification	□	□	□
PC101	Web browser history	□	□	□
PC102	Weight	□	□	□
PC103	Work phone	□	□	□
PC104	X-Rays	□	□	□
PC105	ZIP Code	□	□	□

1

References

5 U.S.C. § 552a.

18 U.S.C. § 2725.

42 U.S.C. § 200.82.

44 U.S.C. § 3552.

Acquisti, A., Brandimarte, L., & Loewenstein, G. (2015). Privacy and human behavior in the age of information. *Science, 347*(6221), 509-514. doi:10.1126/science.aaa1465

Acquisti, A., & Gross, R. (2009). Predicting social security numbers from public data. *Proceedings of the National Academy of Sciences, 106*(27), 10975-10980. doi:10.1073/pnas.0904891106

Acquisti, A., & Grossklags, J. (2005). Privacy and rationality in individual decision making. *IEEE Security & Privacy, 3*(1), 26-33. doi:10.1109/MSP.2005.22

Acquisti, A., Taylor, C., & Wagman, L. (2016). The economics of privacy. *Journal of Economic Literature, 54*(2), 442-492. doi:10.1257/jel.54.2.442

Allen, M. (2006). Social engineering: A means to violate a computer system. *SANS Institute, InfoSec Reading Room*.

Almomani, A., Gupta, B., Atawneh, S., Meulenberg, A., & Almomani, E. (2013). A survey of phishing email filtering techniques. *Communications Surveys & Tutorials, IEEE, 15*(4), 2070-2090. doi:10.1109/surv.2013.030713.00020

Anderson, C. L., & Agarwal, R. (2010). Practicing safe computing: A multimethod empirical examination of home computer user security behavioral intentions. *MIS Quarterly, 34*(3), 613-A615.

Anthes, G. (2014). Data brokers are watching you. *Communications of the ACM, 58*(1), 28-30. doi:10.1145/2686740

Atkins, B., & Huang, W. (2013). A study of social engineering in online frauds. *Open Journal of Social Sciences, 1*(03), 23.

Aven, T., & Renn, O. (2009). On risk defined as an event where the outcome is uncertain. *Journal of Risk Research, 12*(1), 1-11. doi:10.1080/13669870802488883

Bandura, A. (2001). Social cognitive theory of mass communication. *Media Psychology, 3*(3), 265-299.

Barker, E. B. (2013). *Digital signature standard (dss)*. (FIPS PUB 186-4). Washington, DC:National Institute of Standards and Technology (NIST) Retrieved from http://dx.doi.org/10.6028/NIST.FIPS.186-4.

Baron-Cohen, S. (1992). Out of sight or out of mind? Another look at deception in autism. *Journal of Child psychology and Psychiatry, 33*(7), 1141-1155.

Baron-Cohen, S. (1997). *Mindblindness: An essay on autism and theory of mind*: MIT press.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a “theory of mind”? *Cognition, 21*(1), 37-46. doi:10.1016/0010-0277(85)90022-8

Bélanger, F., & Crossler, R. E. (2011). Privacy in the digital age: A review of information privacy research in information systems. *MIS Quarterly, 35*(4), 1017-A1036. doi:10.2307/41409971

Benitez, K., & Malin, B. (2010). Evaluating re-identification risks with respect to the hipaa privacy rule. *Journal of the American Medical Informatics Association, 17*(2), 169-177. doi:10.1136/jamia.2009.000026

Benjamin, V., & Chen, H. (2012). Securing cyberspace: Identifying key actors in hacker communities. *2012 IEEE International Conference on Intelligence and Security Informatics*, 24-29. doi:10.1109/isi.2012.6283296

Bhattacherjee, A. (2012). *Social science research: Principles, methods, and practices.* Book 3, *Textbooks Collection.* Retrieved from http://scholarcommons.usf.edu/oa_textbooks/3

Bilge, L., Strufe, T., Balzarotti, D., & Kirda, E. (2009). All your contacts are belong to us: Automated identity theft attacks on social networks. *Proceedings of the 18th international conference on World wide web*, 551-560. doi:10.1145/1526709.1526784

Bishop, M., & Gates, C. (2008). *Defining the insider threat*. Paper presented at the Proceedings of the 4th annual workshop on Cyber security and informaiton intelligence research developing strategies to meet the cyber security and information intelligence challenges ahead - CSIIRW '08. doi:10.1145/1413140.1413158 doi:10.1145/1413140.1413158

Boone, H. N., & Boone, D. A. (2012). Analyzing likert data. *Journal of extension, 50*(2), 1-5.

Boss, S. R., Galletta, D. F., Benjamin Lowry, P., Moody, G. D., & Polak, P. (2015). What do systems users have to fear? Using fear appeals to engender threats and fear that motivate protective security behaviors. *MIS Quarterly, 39*(4), 837-864. doi:10.25300/misq/2015/39.4.5

Boyd, D. M. (2014). *It's complicated: The social lives of networked teens*. New Haven and London: Yale University Press.

Boyd, D. M., & Ellison, N. B. (2007). Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication, 13*(1), 210-230. doi:10.1111/j.1083-6101.2007.00393.x

Braun, T. D., Siegel, H. J., Beck, N., Bölöni, L. L., Maheswaran, M., Reuther, A. I., . . . Freund, R. F. (2001). A comparison of eleven static heuristics for mapping a class of independent tasks onto heterogeneous distributed computing systems. *Journal of Parallel and Distributed Computing, 61*(6), 810-837. doi:10.1006/jpdc.2000.1714

Chang, D., Krupka, E. L., Adar, E., & Acquisti, A. (2016). Engineering information disclosure: Norm shaping designs. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 587-597. doi:10.1145/2858036.2858346

Chellappa, R. K., & Sin, R. G. (2005). Personalization versus privacy: An empirical examination of the online consumer's dilemma. *Information Technology and Management, 6*(2-3), 181-202. doi:10.1007/s10799-005-5879-y

Chitrey, A., Singh, D., & Singh, V. (2012). A comprehensive study of social engineering based attacks in india to develop a conceptual model. *International Journal of Information and Network Security, 1*(2), 45. doi:10.11591/ijins.v1i2.426

Choo, K.-K. R. (2011). The cyber threat landscape: Challenges and future research directions. *Computers & Security, 30*(8), 719-731. doi:10.1016/j.cose.2011.08.004

Clayton, M. J. (1997). Delphi: A technique to harness expert opinion for critical decision‐making tasks in education. *Educational Psychology, 17*(4), 373-386. doi:10.1080/0144341970170401

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37-46. doi:10.1177/001316446002000104

Coleman, E. G., & Golub, A. (2008). Hacker practice moral genres and the cultural articulation of liberalism. *Anthropological Theory, 8*(3), 255-277. doi:10.1177/1463499608093814

Coleman, J. S. (2000). Social capital in the creation of human capital. *Knowledge and Social Capital*, 17-41. doi:10.1016/b978-0-7506-7222-1.50005-2

Conteh, N. Y., & Schmick, P. J. (2016). Cybersecurity: Risks, vulnerabilities and countermeasures to prevent social engineering attacks. *International Journal of Advanced Computer Research, 6*(23). doi:10.19101/IJACR.2016.623006

Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Boston, MA: Pearson Education.

Culnan, M. J. (1993). "How did they get my name?": An exploratory investigation of consumer attitudes toward secondary information use. *MIS Quarterly, 17*(3), 341-363. doi:10.2307/249775

Culnan, M. J., & Bies, R. J. (2003). Consumer privacy: Balancing economic and justice considerations. *Journal of Social Issues, 59*(2), 323-342. doi:10.1111/1540-4560.00067

Culnan, M. J., & Williams, C. C. (2009). How ethics can enhance organizational privacy: Lessons from the choicepoint and tjx data breaches. *MIS Quarterly, 33*(4), 673-687.

Dadkhah, M., & Quliyeva, A. (2015). Social engineering in academic world. *Journal of Contemporary Applied Mathematics-ISSN: 2222-5498, 4*(2).

Dajani, J. S., Sincoff, M. Z., & Talley, W. K. (1979). Stability and agreement criteria for the termination of delphi studies. *Technological Forecasting and Social Change, 13*(1), 83-90.

Dalkey, N., & Helmer, O. (1963). An experimental application of the delphi method to the use of experts. *Management Science, 9*(3), 458-467. doi:10.1287/mnsc.9.3.458

Dang, Q. H. (2015). *Secure hash standard*. Washington, DC:National Institute of Standards and Technology (NIST) Retrieved from http://dx.doi.org/10.6028/nist.fips.180-4.

de Montjoye, Y.-A., Radaelli, L., Singh, V. K., & Pentland, A. S. (2015). Unique in the shopping mall: On the reidentifiability of credit card metadata. *Science, 347*(6221), 536-539. doi:10.1126/science.1256297

Defense Intelligence Agency. (2011). *Terms & definitions of interest for dod counterintelligence professionals*. Retrieved from https://www.dni.gov/files/NCSC/documents/ci/CI_Glossary.pdf.

Diamond, I. R., Grant, R. C., Feldman, B. M., Pencharz, P. B., Ling, S. C., Moore, A. M., & Wales, P. W. (2014). Defining consensus: A systematic review recommends methodologic criteria for reporting of delphi studies. *Journal of Clinical Epidemiology, 67*(4), 401-409. doi:10.1016/j.jclinepi.2013.12.002

Dinev, T., & Hart, P. (2006). An extended privacy calculus model for e-commerce transactions. *Information Systems Research, 17*(1), 61-80. doi:10.1287/isre.l060.0080

Dworkin, M. J. (2015). *Sha-3 standard: Permutation-based hash and extendable-output functions*. (FIPS PUB 202). Washington, DC:National Institute of Standards and Technology (NIST) Retrieved from http://dx.doi.org/10.6028/nist.fips.202.

Dworkin, M. J., Barker, E. B., Nechvatal, J. R., Foti, J., Bassham, L. E., Roback, E., & Jr., J. F. D. (2001). *Advanced encryption standard (aes)*. (FIPS PUB 197). Washington, DC:National Institute of Standards and Technology (NIST) Retrieved from https://dx.doi.org/10.6028/NIST.FIPS.197.

Ellis, T. J., & Levy, Y. (2006). A systems approach to conduct an effective literature review in support of information systems research. *Informing Science: the International Journal of an Emerging Transdiscipline, 9*, 181+. doi:10.1.1.98.2369

Ellis, T. J., & Levy, Y. (2009). Towards a guide for novice researchers on research methodology: Review and proposed methods. *Issues in Informing Science and Information Technology, 6*, 323-337.

Enck, W., Gilbert, P., Han, S., Tendulkar, V., Chun, B.-G., Cox, L. P., . . . Sheth, A. N. (2014). Taintdroid: An information-flow tracking system for realtime privacy monitoring on smartphones. *ACM Transactions on Computer Systems, 32*(2). doi:10.1145/2619091

Eom, C. S., & Paek, J. H. (2009). Risk index model for minimizing environmental disputes in construction. *Journal of Construction Engineering and Management, 135*(1), 34-41. doi:doi:10.1061/(ASCE)0733-9364(2009)135:1(34)

Falaki, H., Mahajan, R., Kandula, S., Lymberopoulos, D., Govindan, R., & Estrin, D. (2010). Diversity in smartphone usage. *Proceedings of the 8th international conference on Mobile systems, applications, and services*, 179-194. doi:10.1145/1814433.1814453

Fan, W., Lwakatare, K., & Rong, R. (2017). Social engineering: Ie based model of human weakness for attack and defense investigations. *International Journal of Computer Network and Information Security, 9*(1), 1-11. doi:10.5815/ijcnis.2017.01.01

Federal Bureau of Investigation. (2012). Internet social networking risks. *Counterintelligence.*Retrieved from https://www.fbi.gov/file-repository/internet-social-networking-risks-1.pdf/view

Federal Bureau of Investigation. (2015a). Business e-mail compromise. *Stories.*Retrieved from https://www.fbi.gov/news/stories/2015/august/business-e-mail-compromise/business-e-mail-compromise

Federal Bureau of Investigation. (2015b). Social media safety. Retrieved from https://www.fbi.gov/audio-repository/news-podcasts-thisweek-social-media-safety.mp3/view

Federal Bureau of Investigation. (2016). Iranians charged with hacking u.S. Financial sector. Retrieved from https://www.fbi.gov/news/stories/iranians-charged-with-hacking-us-financial-sector/iranians-charged-with-hacking-us-financial-sector

Federal Trade Commission. (2000). Privacy online: Fair information practices in the electronic marketplace: A federal trade commission report to congress. *Washington DC: FTC*.

Feijóo, C., Gómez-Barroso, J. L., & Voigt, P. (2014). Exploring the economic value of personal information from firms' financial statements. *International Journal of Information Management*, *34*(2), 248-256. doi:10.1016/j.ijinfomgt.2013.12.005

Ferraiolo, H., Cooper, D. A., Francomacaro, S., Mehta, K. L., & Sokol, A. W. (2013). *Personal identity verification (piv) of federal employees and contractors*. (FIPS PUB 201-2). Washington, DC:National Institute of Standards and Technology (NIST) Retrieved from http://dx.doi.org/10.6028/nist.fips.201-2.

FIPS 199. (2004). *Standards for security categorization of federal information and information systems*. (FIPS PUB 199). Washington, DC:National Institute of Standards and Technology (NIST) Retrieved from http://dx.doi.org/10.6028/nist.fips.199.

Fitch, K., Bernstein, S. J., Aguilar, M. D., Burnand, B., & LaCalle, J. R. (2001). *The rand/ucla appropriateness method user's manual*. Retrieved from

Fleisher, C. S. (2008). Using open source data in developing competitive and marketing intelligence. *European Journal of Marketing*, *42*(7/8), 852-866. doi:doi:10.1108/03090560810877196

Franceschi-Bicchierai, L. (2015). Teen hackers: A '5-year-old' could have hacked into cia director's emails. Retrieved from https://motherboard.vice.com/read/teen-hackers-a-5-year-old-could-have-hacked-into-cia-directors-emails

Garba, A. B., Armarego, J., Murray, D., & Kenworthy, W. (2015). Review of the information security and privacy challenges in bring your own device (byod) environments. *Journal of Information Privacy & Security*, *11*(1), 38-54. doi:10.1080/15536548.2015.1010985

Geletkanycz, M. A., & Hambrick, D. C. (1997). The external ties of top executives: Implications for strategic choice and performance. *Administrative science quarterly, 42*(4), 654-681. doi:10.2307/2393653

Granger, S. (2001). Social engineering fundamentals, part i: Hacker tactics. *Security Focus, December, 18*.

Green, N. (2017). Standing in the future: The case for a substantial risk theory of "injury in fact" in consumer data breach class actions. *Boston College Law Review, 58*(1), 287-351.

Greening, T. (1996). Ask and ye shall receive: A study in "social engineering". *ACM SIGSAC, 14*(2), 8-14. doi:10.1145/228292.228295

Greenwood, S., Perrin, A., & Duggan, M. (2016, November 11, 2016). Social media update 2016. Retrieved from assets.pewresearch.org/wp-content/uploads/sites/14/2016/11/10132827/PI_2016.11.11_Social-Media-Update_FINAL.pdf

Harl, G. (1997). People hacking: The psychology of social engineering.

Hart, C. (1998). *Doing a literature review: Releasing the social science research imagination*. Thousand Oaks, CA: Sage Publications, Inc.

Hasle, H., Kristiansen, Y., Kintel, K., & Snekkenes, E. (2005). Measuring resistance to social engineering. *International Conference on Information Security Practice and Experience*, 132-143. doi:10.1007/978-3-540-31979-5_12

Heartfield, R., & Loukas, G. (2015). A taxonomy of attacks and a survey of defence mechanisms for semantic social engineering attacks. *ACM Computing Surveys, 48*(3), 1-39. doi:10.1145/2835375

Herbsleb, J. D. (2005, 15-21 May 2005). Beyond computer science. *Proceedings. 27th International Conference on Software Engineering, 2005. ICSE 2005.* IEEe. doi:10.1109/ICSE.2005.1553534

Heurix, J., Zimmermann, P., Neubauer, T., & Fenz, S. (2015). A taxonomy for privacy enhancing technologies. *Computers & Security, 53*, 1-17. doi:10.1016/j.cose.2015.05.002

Hong, J. (2012). The state of phishing attacks. *Communications of the ACM, 55*(1), 74-81. doi:10.1145/2063176.2063197

Hong, W., & Thong, J. Y. L. (2013). Internet privacy concerns: An integrated conceptualization and four empirical studies. *MIS Quarterly, 37*(1), 275-298. doi:10.25300/misq/2013/37.1.12

Jakobsson, M. (2016). Case study: Business email compromise. In M. Jakobsson (Ed.), *Understanding social engineering based scams* (pp. 115-122). New York, NY: Springer New York. doi:10.1007/978-1-4939-6457-4_11

Jasper, S. E. (2017). U.S. Cyber threat intelligence sharing frameworks. *International Journal of Intelligence and CounterIntelligence, 30*(1), 53-65. doi:10.1080/08850607.2016.1230701

Johnston, A. C., Warkentin, M., & Siponen, M. (2015). An enhanced fear appeal rhetorical framework: Leveraging threats to the human asset through sanctioning rhetoric. *MIS Quarterly, 39*(1), 113-A117. doi:10.25300/misq/2015/39.1.06

Junger, M., Montoya, L., & Overink, F. J. (2017). Priming and warnings are not effective to prevent social engineering attacks. *Computers in Human Behavior, 66*, 75-87. doi:10.1016/j.chb.2016.09.012

Kang, J., Shilton, K., Estrin, D., & Burke, J. (2011). Self-surveillance privacy. *Iowa Law Review, 97*, 809-848. doi:10.2139/ssrn.1729332

Karaduman, İ. (2013). The effect of social media on personal branding efforts of top level executives. *Procedia - Social and Behavioral Sciences, 99*, 465-473. doi:10.1016/j.sbspro.2013.10.515

Keane, T. M., Fairbank, J. A., Caddell, J. M., Zimering, R. T., Taylor, K. L., & Mora, C. A. (1989). Clinical evaluation of a measure to assess combat exposure. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 1*(1), 53-55. doi:10.1037/1040-3590.1.1.53

Kennedy, J., Eberhart, R. C., & Shi, Y. (2001a). Humans—actual, imagined, and implied *Swarm intelligence* (pp. 187-259). San Francisco: Morgan Kaufmann. doi:http://dx.doi.org/10.1016/B978-155860595-4/50005-X

Kennedy, J., Eberhart, R. C., & Shi, Y. (2001b). *Swarm intelligence*: Morgan Kaufmann Publishers.

Keysar, B., Lin, S., & Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition, 89*(1), 25-41. doi:10.1016/S0010-0277(03)00064-7

Kim, H.-W., & Pan, S. L. (2006). Towards a process model of information systems implementation: The case of customer relationship management (crm). *SIGMIS Database, 37*(1), 59-76. doi:10.1145/1120501.1120506

Kopan, T. (2015). Cia director john brennan 'outraged' by hack of his emails. Retrieved from http://www.cnn.com/2015/10/27/politics/john-brennan-email-hack-outrage/

Krishnamurthy, B., & Wills, C. E. (2009). On the leakage of personally identifiable information via online social networks. *Proceedings of the 2nd ACM workshop on Online social networks*, 7-12. doi:10.1145/1592665.1592668

Krombholz, K., Hobel, H., Huber, G. P., & Weippl, E. (2013). Social engineering attacks on the knowledge worker. *Proceedings of the 6th International Conference on Security of Information and Networks*, 28-35. doi:10.1145/2523514.2523596

Krombholz, K., Hobel, H., Huber, M., & Weippl, E. (2015). Advanced social engineering attacks. *Journal of Information Security and Applications, 22*, 113-122. doi:10.1016/j.jisa.2014.09.005

Ku, Y.-C., Chen, R., & Zhang, H. (2013). Why do users continue using social networking sites? An exploratory study of members in the united states and taiwan. *Information & Management, 50*(7), 571-581. doi:10.1016/j.im.2013.07.011

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology, 28*(4), 563-575.

Lee, D.-J., Ahn, J.-H., & Bang, Y. (2011). Managing consumer privacy concerns in personalization: A strategic analysis of privacy protection. *MIS Quarterly, 35*(2), 423-A428.

Lee, J. K. (2016). Invited commentary—reflections on ict-enabled bright society research. *Information Systems Research, 27*(1), 1-5. doi:10.1287/isre.2016.0627

Leslie, A. M. (1987). Pretense and representation: The origins of "theory of mind.". *Psychological Review, 94*(4), 412-426. doi:10.1037/0033-295X.94.4.412

Lewis, M. (2017). *The undoing project: A friendship that changed our minds*. New York: W.W. Norton & Company.

Linstone, H. A., & Turoff, M. (1975). *The delphi method: Techniques and applications* (Vol. 29): Addison-Wesley Reading, MA.

Luo, X. R., Brody, R., Seazzu, A., & Burd, S. (2013). Social engineering: The neglected human factor for information security management. *Managing Information Resources and Technology: Emerging Applications and Theories: Emerging Applications and Theories*. doi:10.4018/irmj.2011070101

Maar, M. C. (2013). *An examination of organizational information protection in the era of social media: A study of social network security and privacy protection.* (3556686 Ph.D.), Capella University, Ann Arbor. ProQuest Dissertations & Theses Global database.

Maeterlinck, M. (1930). Life of the white ant.

Mamonova, S., & Koufaris, M. (2016). The impact of exposure to news about electronic government surveillance on concerns about government intrusion, privacy self-efficacy, and privacy protective behavior. *Journal of Information Privacy & Security, 12*(2), 56-67. doi:10.1080/15536548.2016.1163026

Marczak, W. R., & Paxson, V. (2017). Social engineering attacks on government opponents: Target perspectives. *Proceedings on Privacy Enhancing Technologies, 2*, 152-164. doi:10.1515/popets-2017-0019

Martin, K. E. (2015). Ethical issues in the big data industry. *MIS Quarterly Executive, 14*(2).

Maynard, D., Greenwood, M. A., Roberts, I., Windsor, G., & Bontcheva, K. (2015). Real-time social media analytics through semantic annotation and linked open data. *Proceedings of the ACM Web Science Conference*, 1-2. doi:10.1145/2786451.2786500

McCallister, E., Grance, T., & Scarfone, K. (2010). *Guide to protecting the confidentiality of personally identifiable information (pii)*. (SP 800-122). Washington, DC:National Institute of Standards and Technology (NIST) Retrieved from http://doi.org/10.6028/NIST.SP.800-122.

Meguerdichian, S., Koushanfar, F., Qu, G., & Potkonjak, M. (2001). Exposure in wireless ad-hoc sensor networks. *Proceedings of the 7th annual international conference on Mobile computing and networking*, 139-150. doi:10.1145/381677.381691

Mertler, C. A., & Reinhart, R. V. (2013). *Advanced and multivariate statistical methods: Practical application and interpretation* (5th ed.). Glendale, CA: Pyrczak Publishing.

Minkus, T., Liu, K., & Ross, K. W. (2015). Children seen but not heard: When parents compromise children's online privacy. *Proceedings of the 24th International Conference on World Wide Web*, Florence, Italy. International World Wide Web Conferences Steering Committee. doi:10.1145/2736277.2741124

Mitnick, K. D., & Simon, W. L. (2002). *The art of deception: Controlling the human element of security*. New York, NY: John Wiley & Sons.

Mohaisen, A., Al-Ibrahim, O., Kamhoua, C., Kwiat, K., & Njilla, L. (2017). Rethinking information sharing for actionable threat intelligence. *arXiv preprint arXiv:1702.00548*.

Moon, Y. (2000). Intimate exchanges: Using computers to elicit self-disclosure from consumers. *Journal of Consumer Research, 26*(4), 323-339. doi:10.1086/209566

Mouton, F., Leenen, L., Malan, M. M., & Venter, H. (2014). Towards an ontological model defining the social engineering domain. *IFIP International Conference on Human Choice and Computers*. Springer.

Mouton, F., Leenen, L., & Venter, H. S. (2016). Social engineering attack examples, templates and scenarios. *Computers & Security, 59*, 186-209. doi:10.1016/j.cose.2016.03.004

Neupane, A., Rahman, M. L., Saxena, N., & Hirshfield, L. (2015). A multi-modal neuro-physiological study of phishing detection and malware warnings. *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*, 479-491. doi:10.1145/2810103.2813660

Norman, G. (2010). Likert scales, levels of measurement and the “laws” of statistics. *Advances in Health Sciences Education, 15*(5), 625-632. doi:10.1007/s10459-010-9222-y

O'keefe, D. J. (2002). *Persuasion: Theory and research* (Vol. 2). Thousand Oaks, CA: Sage Publications, Inc.

Ohm, P. (2010). Broken promises of privacy: Responding to the surprising failure of anonymization. *UCLA Law Review, 57*(6), 1701-1777.

Olmstead, K., & Smith, A. (2017). Americans and cybersecurity. *Pew Research Center.*Retrieved from assets.pewresearch.org/wp-content/uploads/sites/14/2017/01/26102016/Americans-and-Cyber-Security-final.pdf

Oltmann, S. M. (2010). Katz out of the bag: The broader privacy ramifications of using facebook. *Proceedings of the American Society for Information Science and Technology, 47*(1), 1-4. doi:10.1002/meet.14504701250

Orgill, G. L., Romney, G. W., Bailey, M. G., & Orgill, P. M. (2004). The urgency for effective user privacy-education to counter social engineering attacks on secure computer systems. *Proceedings of the 5th conference on Information technology education*, 177-181. doi:10.1145/1029533.1029577

Orlikowski, W. J., & Baroudi, J. J. (1991). Studying information technology in organizations: Research approaches and assumptions. *Information Systems Research, 2*(1), 1-28.

Palen, L., & Dourish, P. (2003). *Unpacking "privacy" for a networked world*. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Ft. Lauderdale, Florida, USA. http://www.dourish.com/publications/2003/chi2003-privacy.pdf doi:10.1145/642611.642635

Parrish, J. L., & Nicolas-Rocca, S. (2012). Toward better decisions with respect to is security: Integrating mindfulness into is security training. *Proceedings of the Seventh Pre-ICIS Workshop on Information Security and Privacy*.

Pavlou, P. A. (2011). State of the information privacy literature: Where are we and where should we go? *MIS Quarterly, 35*(4), 977-988.

PCI Security Standards Council. (2016). Payment card industry (pci) data security standard, v3.2. Retrieved from https://www.pcisecuritystandards.org/documents/PCI_DSS_v3-2.pdf

Peer, E., & Acquisti, A. (2016). The impact of reversibility on the decision to disclose personal information. *Journal of Consumer Marketing, 33*(6), 428-436. doi:10.1108/jcm-07-2015-1487

Peltier, T. R. (2006). Social engineering: Concepts and solutions. *Electronic data processing audit, control and security newsletter, 33*(8), 1-13. doi:10.1201/1079.07366981/45802.33.8.20060201/91956.1

Perloff, R. M. (2010). *The dynamics of persuasion: Communication and attitudes in the twenty-first century* (Second ed.): Routledge.

Pew Research Center. (2013). Social networking fact sheet. *Pew Research Center.*Retrieved from pewinternet.org/fact-sheets/social-networking-fact-sheet/

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1*(04). doi:10.1017/s0140525x00076512

Privacy Rights Clearinghouse. (2018). Data breaches. Retrieved from https://www.privacyrights.org/data-breaches

Prosch, M. (2008). Protecting personal information using generally accepted privacy principles (gapp) and continuous control monitoring to enhance corporate governance. *International Journal of Disclosure and Governance, 5*(2), 153-166. doi:10.1057/jdg.2008.7

Ramim, M. M., & Lichvar, B. T. (2014). Eliciting expert panel perspective on effective collaboration in system development projects. *Online Journal of Applied Knowledge Management, 2*(1), 122-136.

Raskar, R., Agrawal, A., & Tumblin, J. (2006). Coded exposure photography: Motion deblurring using fluttered shutter. *ACM Trans. Graph., 25*(3), 795-804. doi:10.1145/1141911.1141957

Richey, R. C., & Klein, J. D. (2005). Developmental research methods: Creating knowledge from instructional design and development practice. *Journal of Computing in Higher Education, 16*(2), 23-38. doi:10.1007/BF02961473

Rivard, S., & Lapointe, L. (2012). Information technology implementers' responses to user resistance: Nature and effects. *MIS Quarterly, 36*(3), 897-A895.

Rogers, T. B., Kuiper, N. A., & Kirker, W. S. (1977). Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology, 35*(9), 677-688. doi:10.1037/0022-3514.35.9.677

Rosenbaum, M. H. (2015). Identifying unethical personally identifiable information (pii) privacy violations committed by is/it practitioners: A comparison to computing moral exemplars.

Ross, R., Viscuso, P., Guissanie, G., Dempsey, K., & Riddle, M. (2016). *Protecting controlled unclassified information in nonfederal systems and organizations*. (SP 800-171).National Institute of Standards and Technology Retrieved from http://dx.doi.org/10.6028/nist.sp.800-171r1.

Ross, R. S., Katzke, S. W., & Johnson, L. A. (2006). *Minimum security requirements for federal information and information systems*. Washington, DC:National Institute of Standards and Technology (NIST) Retrieved from http://dx.doi.org/10.6028/nist.fips.200.

Russell, M. A. (2013). *Mining the social web* (Second Edition ed.). Sebastopol, CA: O'Reilly Media, Inc.

Ryan, W. M., & Loeffler, C. M. (2010). Insights into cloud computing. *Intellectual Property & Technology Law Journal, 22*(11), 22-28,21.

Salkind, N. J. (2012). *Exploring research* (8th ed.). Upper Saddle River, NJ: Pearson Education.

Sanders, S. D. (2012). Privacy is dead: The birth of social media background checks. *Southern University Law Review*(39), 243-264.

Saxe, R., Schulz, L. E., & Jiang, Y. V. (2006). Reading minds versus following rules: Dissociating theory of mind and executive control in the brain. *Social Neuroscience, 1*(3-4), 284-298. doi:10.1080/17470910601000446

Schneier, B. (2000). *Secret and lies*.

Schwartz, P. M., & Solove, D. J. (2011). The pii problem: Privacy and a new concept of personally identifiable information. *New York University Law Review, 86*(6), 1814-1894.

Sekaran, U., & Bougie, R. (2013). *Research methods for business: A skill-building approach* (6th ed.). West Sussex, United Kingdom: John Wiley & Sons LTD.

Shapiro, S. L., Carlson, L. E., Astin, J. A., & Freedman, B. (2006). Mechanisms of mindfulness. *Journal of Clinical Psychology, 62*(3), 373-386. doi:10.1002/jclp.20237

Simpson, M. D. (2016). All your data are belong to us: Consumer data breach rights and remedies in an electronic exchange economy. *University of Colorado Law Review, 87*, 669-709.

Singh, B., Bansal, D., & Sofat, S. (2014). An approach of privacy preserving based publishing in twitter. *Proceedings of the 7th International Conference on Security of Information and Networks*, 39-42. doi:10.1145/2659651.2659733

Smith, A. (2015, April 2015). U.S. Smartphone use in 2015. *Pew Research Center.*Retrieved from http://www.pewinternet.org/2015/04/01/us-smartphone-use-in-2015/

Smith, H. J., Dinev, T., & Xu, H. (2011). Information privacy research: An interdisciplinary review. *MIS Quarterly, 35*(4), 980-A927.

Solove, D. J. (2006). A taxonomy of privacy. *University of Pennsylvania Law Review, 154*(3), 477-560. doi:10.2307/40041279

Stone, E. F., Gueutal, H. G., Gardner, D. G., & McClure, S. (1983). A field experiment comparing information-privacy values, beliefs, and attitudes across several types of organizations. *Journal of Applied Psychology, 68*(3), 459-468. doi:10.1037/0021-9010.68.3.459

Straub, D., Boudreau, M.-C., & Gefen, D. (2004). Validation guidelines for is positivist research. *The Communications of the Association for Information Systems, 13*(1), 380-426.

Straub, D. W. (1989). Validating instruments in mis research. *MIS Quarterly, 13*(2), 147-169. doi:10.2307/248922

Sutanto, J., Palme, E., Tan, C.-H., & Phang, C. W. (2013). Addressing the personalization-privacy paradox: An empirical assessment from a field experiment on smartphone users. *MIS Quarterly, 37*(4), 1141-A1145. doi:10.25300/misq/2013/37.4.07

Sweeney, L. (1997). Weaving technology and policy together to maintain confidentiality. *The Journal Of Law, Medicine & Ethics: A Journal Of The American Society Of Law, Medicine & Ethics, 25*(2-3), 98-110. doi:10.1111/j.1748-720x.1997.tb01885.x

Tetri, P., & Vuorinen, J. (2013). Dissecting social engineering. *Behaviour & Information Technology, 32*(10), 1014-1023. doi:10.1080/0144929X.2013.763860

Tversky, A., & Kahneman, D. (1975). Judgment under uncertainty: Heuristics and biases *Utility, probability, and human decision making* (pp. 141-162): Springer. doi:10.1007/978-94-010-1834-0_8

U.S. Department of Justice. (2018). Nine iranians charged with conducting massive cyber theft campaign on behalf of the islamic revolutionary guard corps. Retrieved from https://www.justice.gov/opa/pr/nine-iranians-charged-conducting-massive-cyber-theft-campaign-behalf-islamic-revolutionary

Van den Akker, J., Branch, R. M., Gustafson, K., Nieveen, N., & Plomp, T. (2012). *Design approaches and tools in education and training*: Springer Science & Business Media.

von der Gracht, H. A. (2012). Consensus measurement in delphi studies: Review and implications for future quality assurance. *Technological Forecasting and Social Change, 79*(8), 1525-1536. doi:10.1016/j.techfore.2012.04.013

Wenyin, L., Huang, G., Xiaoyue, L., Min, Z., & Deng, X. (2005). *Detection of phishing webpages based on visual similarity*. Paper presented at the Special interest tracks and posters of the 14th international conference on World Wide Web, Chiba, Japan. doi:10.1145/1062745.1062868

Wilkerson, W. S., Levy, Y., Kiper, J. R., & Snyder, M. (2017). Towards a development of a social engineering exposure index (sexi) using publicly available personal information. *Conference on Cybersecurity Education, Research and Practice*, Kennesaw State University, USA.

Wolff, J. (2016). Perverse effects in defense of computer systems: When more is less. *Journal of Management Information Systems, 33*(2), 597-620. doi:10.1080/07421222.2016.1205934

Workman, M. (2007). Gaining access with social engineering: An empirical study of the threat. *Information Systems Security, 16*(6), 315-331. doi:10.1080/10658980701788165

Workman, M. (2008). Wisecrackers: A theory‐grounded investigation of phishing and pretext social engineering threats to information security. *Journal of the American Society for Information Science and Technology, 59*(4), 662-674. doi:10.1002/asi.20779

Xu, H., Luo, X., Carroll, J. M., & Rosson, M. B. (2011). The personalization privacy paradox: An exploratory study of decision making process for location-aware marketing. *Decision Support Systems, 51*(1), 42-52. doi:10.1016/j.dss.2010.11.017

Youssef, N. A., Green, K. T., Dedert, E. A., Hertzberg, J. S., Calhoun, P. S., Dennis, M. F., . . . Beckham, J. C. (2013). Exploration of the influence of childhood trauma, combat exposure, and the resilience construct on depression and suicidal ideation among u.S. Iraq/afghanistan era military personnel and veterans. *Archives of Suicide Research, 17*(2), 106-122. doi:10.1080/13811118.2013.776445

Zhang, B., Wu, M., Kang, H., Go, E., & Sundar, S. S. (2014). Effects of security warnings and instant gratification cues on attitudes toward mobile websites. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM. doi:10.1145/2556288.2557347