Understanding User-Based Modifications to Information Quality in Response to Privacy and Trust Related Concerns in Online Social Networks

Brian P. Blake and Dr. Nitin Agarwal
University of Arkansas at Little Rock
Little Rock, Arkansas, USA
e-mail: bpblake@ualr.edu nxagarwal@ualr.edu

***Abstract*—As online social networks have surged in popularity, a new wave of privacy discussions are taking place as evolving technology influences perceptions and demands in regard to privacy. From a practitioner's perspective, there is a need to model, measure, and understand information quality in social networks and its relationship to data privacy and trust. From a user's perspective, there is a need to more fully understand both the trust aspects and the visibility and other privacy aspects of information shared online as well as implications from future use of that data. The goal of this research therefore is to model user based modifications to information quality due to data privacy and trust related concerns within online social networks in order to more fully explore the interrelationships and trade-offs between data privacy, trust, and information quality. This research focuses on: 1) development and validation of relationship matrices for data privacy, online social networks, information quality, and trust as a research framework, 2) development of syntax for a conceptual model of data privacy, trust, and information quality in online social networks, and 3) development of a structural equation model for understanding the trade-offs and influences between data privacy, trust, and information quality in online social networks. The greatest implications of this research come through development of integrated matrix frameworks, a privacy/trust/information quality modeling syntax, and structural equation scoring measures that will be applicable to future research efforts. The research will enhance methods of modeling and measuring data privacy, trust, and information quality within online social networks. In application to online social networks, it lends itself to a better understanding of the quality of shared information in given data privacy and trust scenarios. It provides future researchers with a formal framework for relating privacy, trust, and information quality as well as a formal way to understand information quality modification.**

Keywords-Information quality; privacy; trust; online social networks.

I. INTRODUCTION

Social media as communication media have surged in popularity over the past decade. Social networking websites such Facebook, MySpace, and Twitter have been the champions of this social phenomenon [1]. As the use of social media networks increases there are growing concerns about data privacy. A recent paper [2] noted that as information technology evolves it greatly influences perceptions and demands in regard to privacy. Because of this, developments in social computing are driving a new wave of privacy discussions. Government and corporate database privacy issues are often discussed and remain highly important, but according to Zittrain [3] these are "dwarfed by threats to privacy that do not fit the standard analytical template for addressing privacy issues". He used the term Privacy 2.0 to refer to this non-standard view. Zittrain argued that governments or corporations are not always the ones managing surveillance and that control of the transfer of personal information can be eliminated by peer-to-peer technologies.

Frederick Lane, when discussing privacy in a webbed world as part of American Privacy, declared that "information wants to be free" [4]. He continued that social network sites succeed because individuals crave community and will share personal information in order to build it. "Online social networks," he stated, "thrive because they enable us to share personal information more quickly and easily than ever before, creating the impression that we are all newsworthy now". Lane further noted that individuals make seemingly rational decisions to post information online in order to receive perceived benefits, but fully rational decisions require complete information and most individuals don't understand what little control they hold over information posted on social networking sites or personal websites. In a similar vein, Zittrain stated that "people might make rational decisions about sharing their personal information in the short term, but underestimate what might happen to information as it is indexed, reused, and repurposed by strangers" [3].

A. Research Focus

In research related to the general concepts of privacy, trust, and information quality (IQ) each is often addressed in a multi-faceted manner focusing on dimensions, aspects, and properties. To further this, trust, privacy, and information quality as areas of study are interrelated and overlapping in relation to online information disclosure, but how they interact with each other is not fully defined. This is especially true in relation to online social networks (OSNs). Previous research, such as Bertini [5], has noted that there is a direct relationship between privacy, trust, and an individual's willingness to share information of increasing quantity and quality. This creates an opportunity for research. From a practitioners' perspective, there is a need to model, measure, and understand social network information exchanges in regard to privacy, trust, and information quality trade-offs and modifications. From a users' perspective, there is a need to more fully understand both the trust aspects and the visibility of information shared online as well as implications from future use of that data. The goal of this research therefore is to apply an information quality perspective to the modeling of data privacy within social media networks in order to enable the exploration of the interrelationships and tradeoffs between data privacy, trust, and information quality.

This research will address two problem areas. First, a standard way to frame, model, and measure the relationship of the sub-aspects of data privacy, trust, and information quality to facilitate understanding does not exist. This limits research in relation to a comprehensive understanding and restricts cross-discipline communication. Second, a specific understanding of how information quality modification is used by members of online social networks as a reaction to privacy and trust related concerns has not been fully addressed by the information quality research field. This limits the understanding of outcomes based on existing research models in regard to both antecedent influence and behavioral intentions vs. actual behavior within online social networks from an information quality perspective. A greater understanding of these factors can facilitate online social network organization changes to encourage greater sharing while simultaneously giving a deeper insight into how information is shared from an information quality point of view.

B. *Research Implications*

The greatest implications of this research will come through development of integrated matrix frameworks, a privacy/trust/information quality modeling syntax, and structural equation scoring measures that will be applicable to future research efforts. The research can enhance methods of modeling and measuring data privacy at both the data element and entity levels. In application to online social networks, it may lend itself to raised awareness of data visibility in social media as well as a better understanding of the quality of shared information in given data privacy and trust scenarios.

C. *Structure*

The remainder of this paper is organized as follows. Section II describes background issues and related literature. Section III presents research methodologies. Section IV discusses initial results of the research. Section V considers challenges, limitations, and future research opportunities.

II. Background and Related Literature

A. *Privacy*

According to Daniel Solove in Understanding Privacy [6], nearly 120 years after "The Right to Privacy" by Warren and Brandeis was first published in the Harvard Law Review, current views in the field of privacy form a "sweeping concept" that includes "freedom of thought, control over one's body, solitude in one's home, control over personal information, freedom from surveillance, protection of one's reputation, and protection from searches and interrogations". He highlighted others who describe privacy as "exasperatingly vague", "infected with pernicious ambiguities", and "entangled in competing and contradictory dimensions". Helen Nissenbaum [7] noted that privacy is commonly characterized in literature as either a constraint on access or a form of control. As theorists conceptualize privacy, they are typically searching for a core common denominator that forms the essence of privacy, but Solove argued that privacy is not easily conceptualized in this manner. He stated that a common denominator approach broad enough to include the varied aspects of privacy is likely to be vague and overly inclusive, while narrower approaches risk being too exclusive and restrictive. Privacy conceptualizations in existing literature can therefore be grouped into targeted common core definitions and broader privacy frameworks.

Major privacy frameworks have been offered by Solove [6], Nissenbaum [7][8], Holtzman [9], and Rössler [10]. From a research perspective, these broader privacy frameworks have a strong structural relationship to the predominant multi-dimensional framework of information quality. Commonalities can be found across most of these privacy frameworks. The sub-components of the Solove and Rössler frameworks have a strong relationship to each other. Generally, sub-components of these frameworks, as Nissenbaum contended, focus around the twin concepts of access and control. In addition, varied determinations and combinations of these framework sub-components will form key aspects of the contextual norms on which Nissenbaum's contextual integrity framework is based.

Solove presented privacy as "a cluster of many distinct yet related things". His privacy framework conceptualization presented in Understanding Privacy organizes privacy into four areas containing related sub-aspects in which privacy concerns have been be historically raised (see Table I). His framework has a strong focus on the collection, processing, and dissemination of information. This aligns well with online social networks and standard information product flows. Solove's framework also aligns well with common multi-dimensional information quality concepts. Because of this, as well as his recognition as a privacy expert, Solove's privacy conceptualization is used as a basis for the privacy aspects of this research.

B. *Social Media Networks*

Social media is media designed to be disseminated through social interactions created using highly accessible and scalable publishing techniques. It uses internet and web-based technologies to transform broadcast media monologues (one to many) into social media dialogues (many to many). It supports the democratization of knowledge and information, transforming people from content consumers to content producers [11]. Social media networks have been growing in popularity in part due to the increased affordability and proliferation of internet enabled devices that bring social connectivity through personal computers, mobile devices, and internet tablets [12].

Boyd and Ellison [13] describe online social networks as services that enable individuals to “construct a public or semi-public profile within a bounded system”, to “articulate a list of other users with whom they share a connection”, and to “view and traverse their list of connections and those made by others within the system”. Aggarwal [12] states that social networks can be generalized as “information networks, in which the nodes could compromise either actors or entities, and the edges denote the relationship between them”. Online social networks are rich in data and provide unprecedented opportunities for knowledge discovery and data mining. From this perspective, there are two primary social network data types. The first type is linkage-based structural data and the second is content-based data. In relation to privacy, Aggarwal highlights three types of disclosure:

TABLE I. A TAXONOMY OF PRIVACY

A Taxonomy of Privacy	
Information Collection	
Surveillance	The watching, listening to, or recording of an individual’s activities
Interrogation	Various forms of questioning or probing for information
Information Processing	
Aggregation	The combination of various pieces of data about and individual
Identification	The linking of information to a particular individual
Insecurity	Carelessness in protecting stored information from leaks and improper access
Secondary Use	The use of collected information for a purpose different from the use for which it was collected without the data subject’s consent
Exclusion	The failure to allow data subjects to know about the data that others have about them and participate in its handling and use
Information Dissemination	
Breach of confidentiality	Breaking a promise to keep a person’s information confidential
Disclosure	The revelation of truthful information about a person that affects the way others judge his or her reputation
Exposure	Revealing another's nudity, grief, or bodily functions
Increased accessibility	Amplifying the accessibility of information
Blackmail	The threat to disclose personal information
Appropriation	The use of the data subject's identity to serve another's aims and interests
Distortion	Disseminating false or misleading information about individuals
Invasions	
Intrusion	Invasive acts that disturb one's tranquility or solitude
Decisional interference	Incursions into the data subject's decisions regarding her private affairs

> [S]ocial networks contain tremendous information about the individual in terms of their interests, demographic information, friendship link information, and other attributes. This can lead to disclosure of different kinds of information in the social network, such as identity disclosure, attribute disclosure, and linkage information disclosure. [12]

From a more structural perspective, Bruce Schneier [14] proposed that social network data can be divided into six categories (see Table II). Hart and Johnson [15] noted that Schneier’s taxonomy highlights three primary sources through which information can be disseminated: through the users themselves, through other individuals, or through inference. In regard to privacy, all three of these sources can lead to privacy compromises. A similar structured view of data is also shared by Facebook [16] in its published data use policy.

TABLE II. TYPES OF SOCIAL NETWORK DATA

Types of Social Network Data	
Service Data	Data users give to a social networking site in order to use it
Disclosed Data	What users post on their own pages
Entrusted Data	What users post on other people's pages
Incidental Data	What other people post about a user
Behavioral Data	Data the site collects about user habits by recording what users do and who users do it with
Derived Data	Information about users that is derived from all the other data

C. *Information Quality*

Information quality (also known as data quality) is a multidisciplinary field with research spanning a wide range of topics, but existing researchers are primarily operating in the disciplines of Management Information Systems and Computers Science [17]. Within quality literature, the concept of “fitness for use” has been widely adopted as a definition for data quality [5][17]-[20]. But in order to be applicable, this definition of fitness for use needs to be contextualized [5]. In this regard, previous writings and research have presented data quality as a multi-dimensional concept [17]-[21].

In 1996, Wang and Strong published a hierarchical framework to capture the multi-dimensional aspects of information quality that are most important to data consumers [19]. This research was presented in application by Strong, Lee and Wang in “Data Quality in Context” the following year [20]. Since that time, their framework has been widely cited in information quality literature. The Wang Strong Quality Framework [19] contains four categories of data quality: Intrinsic DQ, Contextual DQ, Representational DQ, and Accessibility DQ. These four categories contain fifteen data quality dimensions (see Table III).

TABLE III. WANG STRONG QUALITY FRAMEWORK

DQ Category	DQ Dimensions
Intrinsic DQ	Accuracy, Objectivity, Believability, Reputation
Accessibility DQ	Accessibility, Access Security
Contextual DQ	Relevancy, Value-Added, Timeliness, Completeness, Amount of Data
Representational DQ	Interpretability, Ease of Understanding, Concise Representation, Consistent Representation

D. *Trust*

Trust, like privacy and quality, is a widely studied concept across multiple disciplines. This has led to the development of a broad array of definitions and understandings of trust over time [22]-[26]. Marsh [22] highlighted that trust values have no units, but can still be measured by such notions as ‘worthwhileness’ and ‘intrinsic value’. At the same time, trust is an absolute medium in which one either trusts or does not trust. This implies that trust in application is based on threshold values above which or below which an entity is either trusted or not trusted as seen in Fig. 1. These thresholds will also vary with different entities and in different circumstances. In a similar manner, Kosa [27] noted that “[t]rust can be examined as a continuous measure, as in evaluation or reliability assessments, or a binary decision point when referring to a decision”.

Mayer, Davis, and Schoorman [28] strove to differentiate trust from other related constructs. They presented an integrative model of organizational trust. Within this research, they expanded upon the characteristics of a trustee and presented a concept of perceived trustworthiness. The identified characteristics, or primary factors, of perceived trustworthiness they presented are Ability, Benevolence, and Integrity. In this, Ability relates to the skills, characteristics, and competencies that enable someone to have influence with a specific domain. Benevolence is related to the level of goodwill a trustee is believed to have toward a trustor. Integrity relates to how a trustee is perceived to adhere to an acceptable set of principles. The authors proposed that “trust for a trustee will be a function of the trustee's perceived ability, benevolence, and integrity and of the trustor's propensity to trust”. They further noted that, while related, these three attributes are separable and may vary independently of one another.

Gefen [26] drew on concept of trustworthiness presented by Mayer, Davis, and Schoorman to develop a validated scale specifically related to online consumer trust. The results of his research showed that each of the aspects of trustworthiness as tested against online behavioral intentions is different. This may suggest that each of the three aspects of trustworthiness “affect different behavioral intentions because different beliefs affect different types of vulnerability”. Gefen’s research also illustrated the measurability of aspects such as trust in regard to interactions in an online domain. This is important to the research at hand.

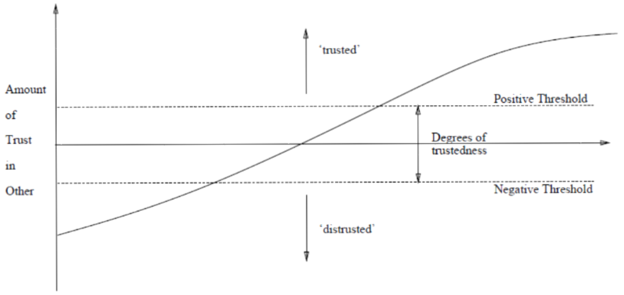


Figure 1 - Positive and Negative Thresholds for Trust [22]

In specific regard to social networks, Adali et al. [29] highlighted that trust also has a major role in the formation of social network communities, in assessing information quality and credibility, and in following how information moves within a network. They further noted the social mechanisms of trust formation in online communities are a new research area and there are many unknowns.

E. *Interdependencies*

Prior research presented by Bertini [5] begins to highlight the interdependencies between data privacy, trust, and information quality. If quality is defined as fitness for use and accuracy, reliability, and trustworthiness are key aspects of high quality data, then “high quality data require data subjects to disclose personal information raising some threat to their own privacy”. Bertini, citing Rose (2001), Hoffman et al. (1999), Neus (2000), and Hui et al. (2006), noted that “studies reveal that data subjects often provide incorrect information or withdraw from interaction when they consider the risks of disclosing personal data higher than the reward they can get from it”. As stated previously, control is a key aspect in several conceptualizations and definitions of privacy. Bertini emphasized that lack of control leads to increased concern over “unauthorized secondary use, excessive collection of data, improper access and processing or storing errors”. Citing research by Gefen (2002), Paine et al. (2006), and Hoffman et al. (1999), Bertini built on the concept that “[d]ata subjects’ level of trust determine both the quantity and the quality of information they disclose” by presenting the relationship between privacy and data quality as a trust mediated process. Bertini noted that the concept of benevolence as presented by Mayer, Davis, and Schoorman is a central trust factor in that both trustee and trustors need to believe that the other is sincere, otherwise data sharing processes breakdown or become cumbersome. He believed that giving users control and allowing them to interact with their data, especially dynamic data, will both increase trust and spontaneously improve data quality. Conversely, when privacy or control is threatened, it causes a loss of trust, which leads to an immediate decrease in the quality of data being disclosed.

Kosa [27] stated that "research on privacy and trust as linked phenomena remains scarce". She noted that the formalization of trust is much more mature than the formalization of privacy and proposed that because of their conceptual similarities formalization concepts developed in relation to trust could be utilized in the formalization of privacy. Kosa highlights that both trust and privacy are highly information type and sensitivity specific, relationship dependent, purpose driven, and measured on a continuous scale. In example of the application of trust formalizations to privacy, she diagramed, as seen in Fig. 2, proposed thresholds for privacy based on the trust threshold detailed by Marsh [22].

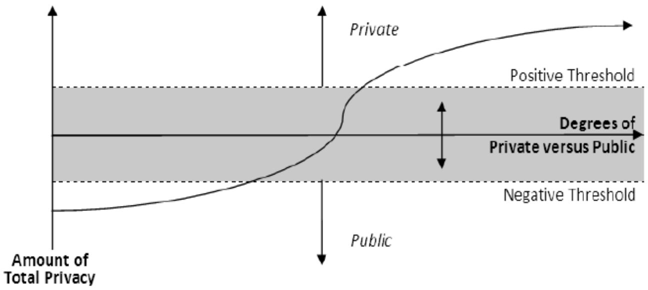


Figure 2 - Proposed Thresholds for Privacy [27]

Further, Kosa presented trust as positively correlated to privacy, but privacy as negatively related to trust. She stated that "Perceptions of trustworthiness may increase the tendency of people to share information willingly, thus giving up their privacy" but the "exercise of privacy may impede trust; if [one chooses] to withhold information, about for example, [his] identity the second party is less likely to trust [him] in the given exchange". This seems counter to the privacy/trust view presented by Bertini [5] above, but it is really a reflection on the relationship of different dimensions between trust and privacy.

For this research, the interdependency between trust, privacy, and information quality as well as the multi-dimensional nature of these concepts highlighted in this section are key foundations. These concepts will be extended in specific relation to online social network sites with a focus on modeling data privacy and measuring the corresponding trade-offs in information quality and/or trust.

III. Methodology

The research will contain three components that build upon each other. The first will be the development and validation of select relationship matrices for data privacy, online social network data, trust, and information quality as a research framework. The second will be the development of a syntax and conceptual model as a standard way to document the trust, privacy, and information quality aspects within online social networks. Finally, a structural equation model will be developed to measure and validate expected information quality modifications as a reaction to calculated privacy risks based on data elements of different data types, content sensitivity, and data visibility. While these components can be generalized across multiple online social networks, for this research, when analyzing online social networks, Facebook will be used as a primary reference because of the size and activity levels of its user base.

A. *Framework Matrix*

This research will focus on the general overlap of the multi-faceted dimensions, aspects, and properties of trust, privacy, information quality, and online social networks. It seeks to identify where these areas overlap both in regard to online social networks and to each other. This phase of the research hypothesizes that:

1) The multi-faceted dimensions, aspects, and properties of trust, privacy, and information quality can be effectively overlaid within a series of related matrices.

2) An understanding of intersections of these sub-aspects lends itself to a broader understanding of the relationship of these concepts.

3) An understanding of intersections of these sub-aspects lends itself to specific target areas for future research.

As a starting point for this research, a framework matrix has been developed to map the points of intersection between Solove's [6] taxonomy of privacy, Schneier's [14] divisions of social network data, Wang and Strong's [19] multiple dimensions of information quality, and the trustworthiness characteristics of Ability, Benevolence, and Integrity as presented by Mayer et al. [28] and Gefen [26]. As noted above, the development and validation of select relationship matrices for data privacy, online social networks, information quality, and trust as a research framework will be the first deliverable from this research. This will be accomplished in part through a validation in current literature. Hogben [31], for example, highlighted specific online social network privacy threats that include digital dossier aggregation, secondary data collection, recognition and identification, data permanence, infiltration of networks, profile squatting and ID theft related reputation slander, and cyberstalking/cyberbullying. These can be shown to align neatly with the proposed privacy components within the framework matrix. In addition, a select survey of information quality, online social network, and privacy related professionals and experts whose opinions will be gathered and reconciled.

B. *Syntax and Conceptual Modeling*

In regard to modeling privacy in social networks, one general approach is the mapping of entity level social graph connections of the network. This high level node and edge view is the most common social graph view. This approach visualizes the issue, but focuses on privacy at the level of overall connections. A second approach presented by Lui and Terzi [32] and others is the calculation of mathematical data element level and entity level privacy scores. This is a more detailed approach focused on the numeric scoring of data privacy. The concepts of Lui and Terzi were an early influence on the development of this syntax. This research gives the opportunity to blend previous research into an expanded approach. This is done by developing a method to model the data privacy of specific data elements that can then be incorporated in the future into trade-off scoring research. This method may also lend itself in future research to the creation of elemental data privacy social graphs which will allow for the visualization of actual data sharing, not just entity level connections.

The second key aspect of this research is to develop a syntax and conceptual model as a standard way to document the trust, privacy, and information quality aspects within online social networks. This phase of the research hypothesizes that:

1) Instances of trust, privacy, and information quality interactions can be expressed at the data element level in notation sets expressing element, users, privacy, trust, and quality components.

2) Instances of trust, privacy, and information quality interactions can be expressed at the data element level as a conceptual model.

A further research question, if these hypotheses hold true, is whether this be implemented in a way that can aggregate to an overall user level notation and conceptualization. This research will seek to validate these hypotheses through illustration of the conceptual model using synthetic and real world examples as well as validation by extension through structural equation modeling. To control for scope, this research will focus on the user controlled social sharing aspects of online social network information such as Disclosed, Entrusted, and Incidental data rather than organizational (system and third party) aspects such as Behavioral, Derived, and Service data. In this regard, the following syntax structures are being proposed as a concept to be further developed in this research.

For disclosed data elements that users post on their own pages, the most apparent privacy aspect is the visibility level of the data element set by the users' privacy settings. Visibility levels are typically set by users' overall privacy settings or by specific selection when posting a data element. One research question related to this is how trust and information quality are related to a users' determination of visibility related privacy settings. This syntax follows the form of Disclosed Data as D1(J1, PJ1) where D1 = Disclosed Data Element with a descriptive set of J1 = Posting Entity and PJ1 = User Privacy Factors.

For entrusted data elements that users post on other people's pages, there are two main privacy considerations related to the visibility level of the data element. The first is the posting entity's own privacy settings. The second is the receiving entity's privacy settings. Generally, the posting entity's privacy settings are the controlling factor in regard to data visibility. This syntax follows the form of Entrusted Data as E1(J1, J2, PJ1, PJ2) where E1 = Entrusted Data Element with a descriptive set of J1 = Posting Entity, J2 = Receiving Entity, PJ1 = Privacy Factors of the Posting Entity, and PJ2 = Privacy Factors of the Receiving Entity.

For incidental data elements that users post about others, there are also two main privacy considerations. As with entrusted data, the first consideration is the Posting Entity's own privacy settings. This most typically relates to the visibility of the data element. The second consideration is the exclusion factor of the Topic Entity. A Topic Entity is the person, group, or thing which is the subject of a posted data element. Exclusion relates to the level of control and involvement a user has in regard to information that is shared about or actions taken that affect him or her. Within online social networks, this relates to whether or not the incidental data element is directly linked, often through tagging, to the Topic Entity. Topic Entities can often reduce visibility of shared data by preventing tagging or removing tags on incidental data elements, but preventing tagging will increase a user's exclusion factor because the user will be less likely to be directly linked and therefore will not be notified when incidental data is posted. In addition, while a user can reduce visibility by blocking or removing user tags, he or she usually cannot prevent the comments or references themselves from being made by other users. Because of this lack of control, the trustworthiness characteristic of benevolence plays an important role in incidental data. This syntax follows the form of Incidental Data as I1(J1, J3, PJ1, EJ3) where I1 = Incidental Data Element with a descriptive set of J1 = Posting Entity, J3 = Topic Entity, PJ1 = Privacy Factors for the Posting Entity, and EJ3 = Exclusion factor of Topic Entity.

In expansion of this syntax, an important question to be addressed in this research is whether and how quality and trust components such as Q1 as Data Element Quality, TJ1J2/TJ1Jx as Relational Trust between Entities, and TS as System Trust can be incorporated directly into this model syntax. This will need to be developed to facilitate comparative measurement of trade-offs between data privacy, information quality, and trust. This syntax could follow the form of Entrusted Data with Trust and Quality as E1(J1, J2, PJ1, PJ2, TS, TJ1J2, TJ1Jx, QE1) where E1 = Entrusted Data Element with a descriptive set of J1 = Posting Entity, J2 = Receiving Entity, PJ1 = Privacy Factors for the Posting Entity, PJ2 = Privacy Factors for the Receiving Entity, TS = System Trust, TJ1J2 = Relational Trust between Posting and Receiving Entities (subset of TJ1Jx), TJ1Jx = Relational Trust between Connected Entities, and QE1 = Set of Data Element Information Quality Factors (see Fig. 3).

C. *Structural Equation Modeling*

The goal of the comparative scoring component of this research is to tie the conceptual modeling syntax back to information quality, trust, and data privacy relationships identified in the framework matrices in the first research

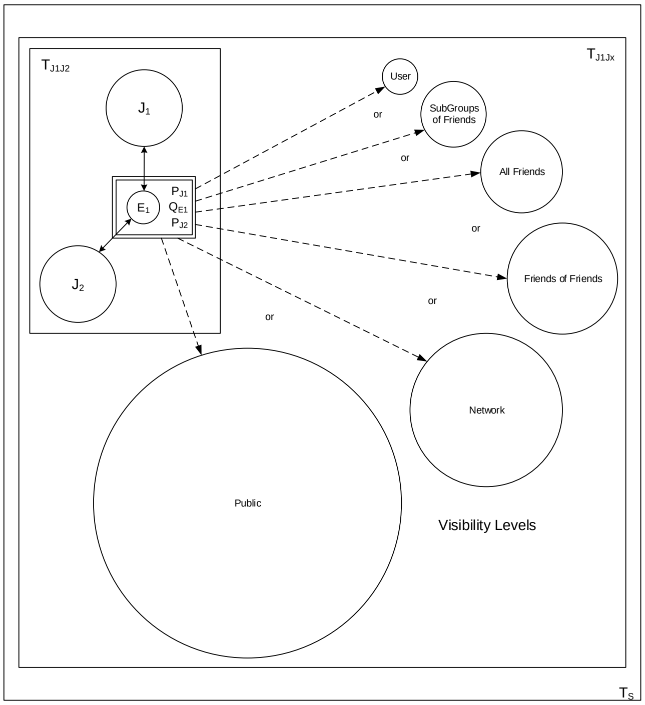


Figure 3 - Data Privacy Modeling of Entrusted Data with Trust and Quality

component. This will have a strong research impact through the creation of a comparative mathematical model of data privacy attributes, information quality dimensions, and trust characteristics. This research phase will develop a structural equation model to measure and validate expected information quality modifications as a reaction to calculated risks based on data elements of different data types, content sensitivity, and data visibility. Previous research showed the benefit of structural equation models in the development and validation of the Internet Users' Information Privacy Concerns [33] and User Privacy Concerns and Identity in OSNs [34] constructs. This research will also use structural equation modeling to extend and build upon those concepts.

Malhotra, Kim, and Agarwal [33] developed the Internet Users' Information Privacy Concerns (IUIPC) construct based on the extension of personal dispositions to data collection, privacy control, and privacy awareness to beliefs regarding trust and risk and how those beliefs affected behavioral intention in regard to Internet usage. This proposed research will extend the IUIPC casual model to online social network specific contextual variables of varied data element type and data sensitivity. It will also incorporate aspects of information quality modification rather than utilize the direct share/not share behavioral intention utilized by Malhotra, Kim, and Agarwal.

Krasnova, Günther, Spiekermann, and Koroleva [34] developed a model for Privacy Concerns and Identity in Online Social Networks (PCIOSN). This cross-discipline research comes more from the social sciences and is developed through a social identity disclosure perspective. They argue that while IUIPC has been widely utilized these applications are lacking because "OSN members are subject to the specific privacy-related risks rooted in the public and social nature of OSNs". They further noted that in terms of primary privacy concerns individuals differentiate between online social network users and provider or third-party organizations. Their research model has a degree of overlap with the proposed framework matrix found in this research. It is based on specific privacy concerns affecting the amount, accuracy, and control aspects of shared information. This research will extend their model to directly map specific privacy and trust aspects from the framework matrix into the threat components of the PCIOSN model. The proposed research will also specifically map dimensions of individual self-disclosure [34] to specific IQ dimensions, as well as incorporate other relevant IQ dimensions, from the proposed framework matrix. Of additional research interest is whether or not the IUIPC and PCIOSN models can be incorporated into a single view through the modeling aspects of this research. This research hypothesizes that:

1) Behavioral intent to share information is not a simple binary response. Instead it is a degree based response that uses information quality modification to mitigate privacy and trust concerns between the thresholds of open disclosure and full non-disclosure (see Fig. 4).

2) Data element types (wall posts, photos, comments, shares, likes, check-ins, etc.) have measurably different thresholds for content sensitivity.

3) Completeness, Accuracy, Accessibility, Amount, Understandability, and similar quality dimensions of shared information are negatively related to calculated privacy and trust concerns as a modification control.

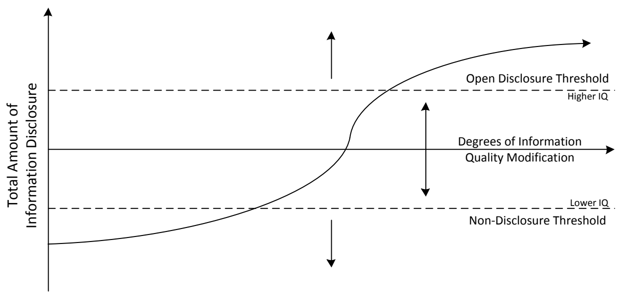


Figure 4 - Initial Information Quality Modification Concept

Hypothesis 1 is an extension of Marsh's Positive and Negative Thresholds for Trust [22] and Kosa's Proposed Thresholds for Privacy [27] as applied to information quality. It should also be noted that any modification of Accessibility IQ dimension mitigates privacy and trust concerns by changing the visibility of a given piece of information rather than changing the shared information itself. As with the second research component, this research will be confined to specific data elements within selected social network data types to control for scope. It will focus first on the user controlled social sharing aspect of Disclosed data, but may easily extend to Incidental and Entrusted data in future research. Specific trust characteristics, information quality dimensions and data privacy aspects will be selected.

ISBN: 978-1-61208-504-3

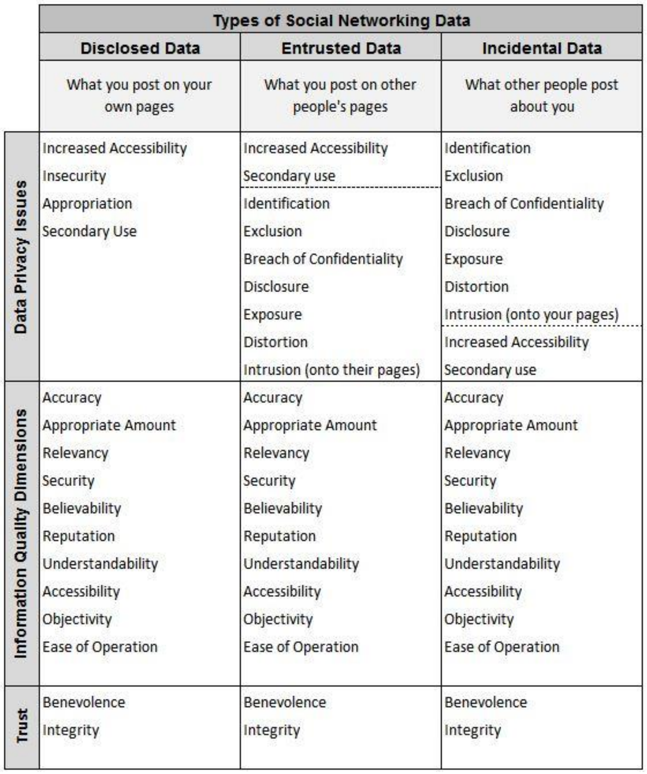
For these selected attributes, measurable indicators within online social networks will be identified and corresponding variables and questions for metrics and measurement will be determined. Structural equation modeling will be utilized as a method for measuring the balance trade-offs present between specific trust characteristics, information quality dimensions and data privacy aspects.

IV. CURRENT RESULTS

This paper presents in process doctoral dissertation research. To this point, the relationship matrices for data privacy, online social networks, information quality, and trust as a research framework have been developed and a corresponding validation survey has been created but not yet implemented. Furthermore, an initial syntax for conceptual modeling has been presented. Currently, elements of the proposed structural equation model and its required survey as a validation instrument are under development.

The developed framework matrices are presented in full in Appendices A-D, but as noted in the Section III, only syntax for conceptual modeling of Disclosed, Entrusted, and Incidental data has been developed. This framework matrix subset is presented in Table IV.

TABLE IV. FRAMEWORK MATRIX SUBSET



	Types of Social Networking Data		
	Disclosed Data	Entrusted Data	Incidental Data
	What you post on your own pages	What you post on other people's pages	What other people post about you
Data Privacy Issues	Increased Accessibility Insecurity Appropriation Secondary Use	Increased Accessibility Secondary use Identification Exclusion Breach of Confidentiality Disclosure Exposure Distortion Intrusion (onto their pages)	Identification Exclusion Breach of Confidentiality Disclosure Exposure Distortion Intrusion (onto your pages) Increased Accessibility Secondary use
Information Quality Dimensions	Accuracy Appropriate Amount Relevancy Security Believability Reputation Understandability Accessibility Objectivity Ease of Operation	Accuracy Appropriate Amount Relevancy Security Believability Reputation Understandability Accessibility Objectivity Ease of Operation	Accuracy Appropriate Amount Relevancy Security Believability Reputation Understandability Accessibility Objectivity Ease of Operation
Trust	Benevolence Integrity	Benevolence Integrity	Benevolence Integrity

Table IV illustrates several key factors. First, intersection points of the matrix may highlight different or similar aspects of privacy, trust, and information quality. Differentiations are shown for only data privacy issues in this subset, but they can be seen more readily in the full framework matrix presented in Appendix A. Second, related social sharing aspects of online social network information such as the user controlled areas of Disclosed, Entrusted, and Incidental data will be more similar to each other than to organizational (system and third party) aspects such as Behavioral, Derived, and Service data. It should also be noted that aspects as currently presented in the matrix intersection points are not in any specific rank order. Even when similar aspects are presented, those aspects may have different levels of importance based on the social networking data type being researched. Finally, the dotted lines found in the data privacy grids for Entrusted and Incidental data are there to indicate distinctions between data privacy violations that may happen to a user and data privacy violations that a user may cause to happen to others.

V. CHALLENGES, LIMITATIONS, AND LOOKING AHEAD

This research faces several challenges and limitations. First, while a broad framework matrix can be presented, the scope for validation and deeper research will be limited to social network data types that relate to user specific aspects of the framework matrix. The role of provider and third-party related online social network data types are highly noteworthy, but they will be addressed in only a limited manner, if at all, in this research. Second, to limit scope during the development of a syntax and conceptual model, not all variations of data element types and entity interactions will be addressed. Once again, in order to control research scope, the focus will be on select user specific aspects of the framework matrix as well as a targeted set of matrix overlays. This series of scope limitations is detailed more specifically within the Methodology section of this paper.

Challenges for this research may include determining and attracting a diverse set of respondents to create a representative population in phase three of this study. For measurements within structural equation modeling to be considered valid certain minimum respondent thresholds need to be met based on the number of components within the model. In addition, structural equation modeling analysis requires the identification of alternate models. Because of the dynamics of social networks, identifying all alternative models may be difficult. Further, finding field experts willing to participate in the framework matrix validation survey may also be difficult, but since only a small number are required it may be a challenge that is more easily overcome.

REFERENCES

[1] B. Blake, N. Agarwal, R. Wigand, and J. Wood, “Twitter Quo Vadis: Is Twitter Bitter or are Tweets Sweet?” The Seventh International Conference on Information Technology: New Generations (ITNG), 2010, pp. 1257-1260.

[2] K. Borcea-Pfitzmann, A. Pfitzmann, and M. Berg, “Privacy 3.0 := Data Minimization + User Control + Contextual Integrity,” it - Information Technology, vol. 53, no. 1, pp. 34-40, 2011. [Online]. Available from: https://tu-dresden.de/Members/katrin.borcea-pfitzmann/ 2016.07.16

[3] J. Zittrain, The Future of the Internet - And How to Stop it, New Haven, CT: Yale University Press, 2008.

[4] F. S. Lane, American Privacy: The 400-Year History of our Most Contested Right, Boston, MA: Beacon Press, 2009.

[5] P. Bertini, "Trust Me! Explaining the Relationship Between Privacy and Data Quality," Information Technology and Innovation Trend in Organization, 2010. [Online]. Available from http://www.cersi.it/itais2010/ 2016.07.16

[6] D. J. Solove, Understanding Privacy. Cambridge, MA: Harvard University Press, 2008.

[7] H. F. Nissenbaum, Privacy in Context: Technology, Policy, and the Integrity of Social Life. Stanford, CA: Stanford University Press, 2010.

[8] H. Nissenbaum, Privacy as contextual integrity. *Washington Law Review*, vol. *79*, no. 1, pp. 101-139, 2004. Available from http://www.nyu.edu/projects/nissenbaum/main_cv.html#pub 2016.07.16

[9] D. H. Holtzman, Privacy Lost: How Technology is Endangering Your Privacy, San Francisco: Jossey-Bass, 2006.

[10] B. Rössler (Ed.), Privacies: Philosophical Evaluations, Stanford, Calif: Stanford University Press, 2004.

[11] N. Agarwal, Types of Social Media, lecture presented for Social Media Mining and Analytics course at the University of Arkansas at Little Rock, Feb. 2011.

[12] C. C. Aggarwal, Social Network Data Analytics, New York: Springer, 2011.

[13] D. M. boyd and N. B. Ellison, "Social Network Sites: Definition, History, and Scholarship," Journal of Computer-Mediated Communication, vol. 13, no. 1, pp. 210-230, 2008.

[14] B. Schneier, "A Taxonomy of Social Networking Data," IEEE Security & Privacy Magazine, vol. 8, no. 4, p. 88, 2010, doi: 10.1109/MSP.2010.118

[15] M. Hart and R. Johnson, "Prevention and Reaction: Defending Privacy in the Web 2.0," 2010. [Online]. Available from: http://www.w3.org/2010/policy-ws/papers/04-Hart-stonybrook.pdf 2016.07.16

[16] Facebook, Data Policy, [Online]. Available from: https://www.facebook.com/about/privacy/your-info 2016.07.16

[17] S. E. Madnick, R. Y. Wang, Y. W. Lee, and H. Zhu, "Overview and Framework for Data and Information Quality Research," Journal of Data and Information Quality, vol. 1, pp. 2:1-2:22, 2009.

[18] C. Fisher, E. Lauria, S. Chengalur-Smith, R. Wang, Introduction to Information Quality, M.I.T. Information Quality Program, 2006

[19] R. Y. Wang and D. M. Strong, "Beyond Accuracy: What Data Quality Means to Data Consumers," Journal of Management Information Systems, vol. 12, no. 4, pp. 5-33, 1996.

[20] D. M. Strong, Y. W. Lee, and R. Y. Wang, "Data Quality in Context," Commun. ACM, vol. 40, pp. 103-110, May 1997.

[21] L. L. Pipino, Y. W. Lee, and R. Y. Wang, "Data Quality Assessment," Commun. ACM, vol. 45, pp. 211-218, Apr. 2002.

[22] S. P. Marsh, Formalising Trust as a Computational Concept, unpublished doctoral dissertation, University of Stirling, 1994. [Online]. Available from: https://dspace.stir.ac.uk/ 2016.07.16

[23] C. D. Schultz, "A Trust Framework Model for Situational Contexts," Proceedings of the 2006 International Conference on Privacy, Security and Trust: Bridge the Gap between PST Technologies and Business Services (PST '06), New York, NY, USA: ACM, 2006, pp. 50:1-50:7.

[24] D. McKnight and N. Chervany, "Conceptualizing Trust: A Typology and E-commerce Customer Relationships Model," Proceedings of the 34th Annual Hawaii International Conference on System Sciences, 2001, p. 10.

[25] A. Gutowska, Research in Online Trust: Trust Taxonomy as a Multi-Dimensional Model, Technical Report, School of Computing and Information Technology, University of Wolverhampton, 2007.

[26] D. Gefen, "Reflections on the Dimensions of Trust and Trustworthiness Among Online Consumers," SIGMIS Database, vol. 33, pp. 38-53, 2002.

[27] T. Kosa, "Vampire Bats: Trust in Privacy," Eighth Annual International Conference on Privacy Security and Trust (PST), 2010, pp. 96-102, doi: 10.1109/PST.2010.5593227.

[28] R. C. Mayer, J. H. Davis, and F. D. Schoorman, "An Integrative Model of Organizational Trust," The Academy of Management Review, vol. 20, no. 3, pp. 709-734, 1995.

[29] S. Adali, R. Escriva, M. K. Goldberg, M. Hayvanovych, M. Magdon-Ismail, B. K. Szymanski, and G. Williams, "Measuring Behavioral Trust in Social Networks," 2010 IEEE International Conference on Intelligence and Security Informatics (ISI), 2010, pp. 150-152.

[30] D. L. Hoffman, T. P. Novak, and M. Peralta, "Building Consumer Trust Online," Commun. ACM, vol. *42*, pp. 80-85, Apr. 1999.

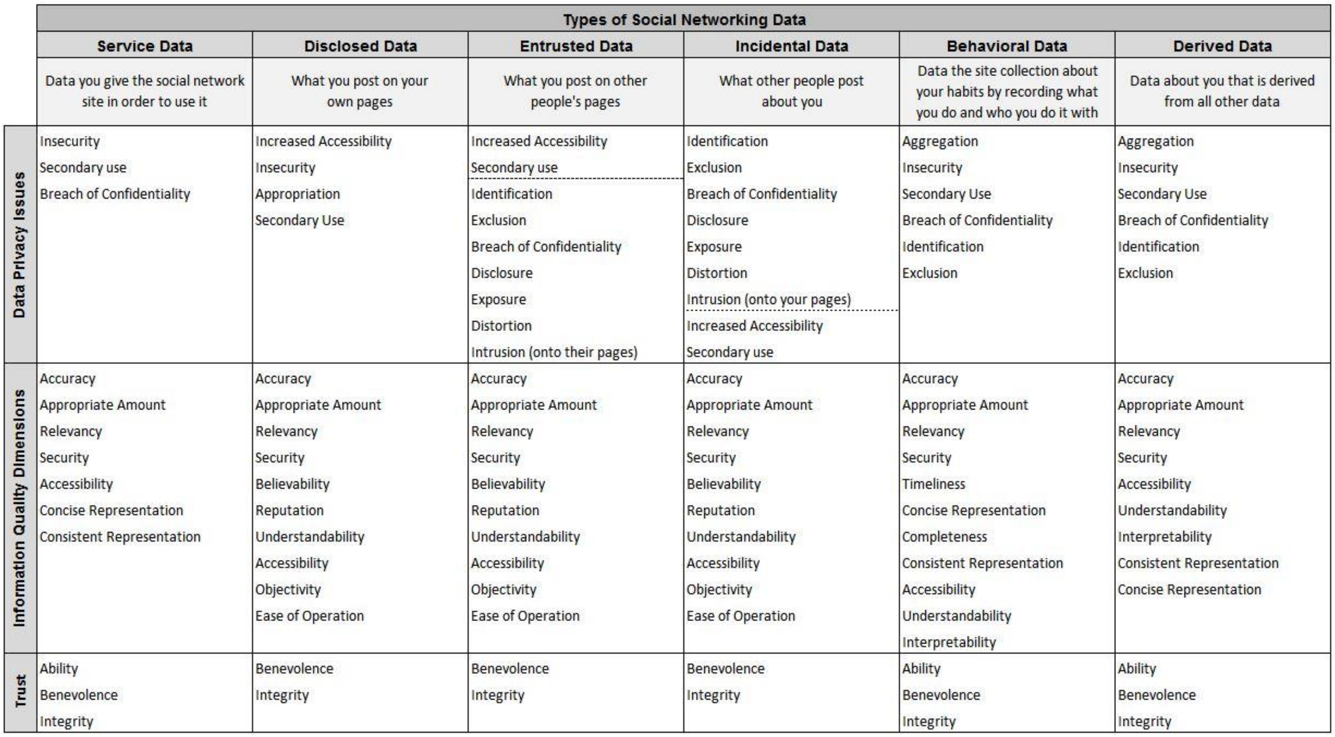
[31] G. Hogben (Ed.), ENISA Position Paper No. 1: Security Issues and Recommendations for Online Social Networks, European Network and Information Security Agency, Nov. 2007. [Online]. Available from: https://www.enisa.europa.eu/publications/archive/security-issues-and-recommendations-for-online-social-networks 2016.07.16

[32] K. Liu and E. Terzi, "A Framework for Computing the Privacy Scores of Users in Online Social Networks," Ninth IEEE International Conference on Data Mining (ICDM '09), 2009, pp. 288-297, doi: 10.1109/ICDM.2009.21.

[33] N. K. Malhotra, S. S. Kim, and J. Agarwal, "Internet Users' Information Privacy Concerns (IUIPC): The Construct, the Scale, and a Causal Model," Information Systems Research, vol. 15, no. 4, pp. 336-355, 2004. doi: 10.1287/isre.1040.0032.

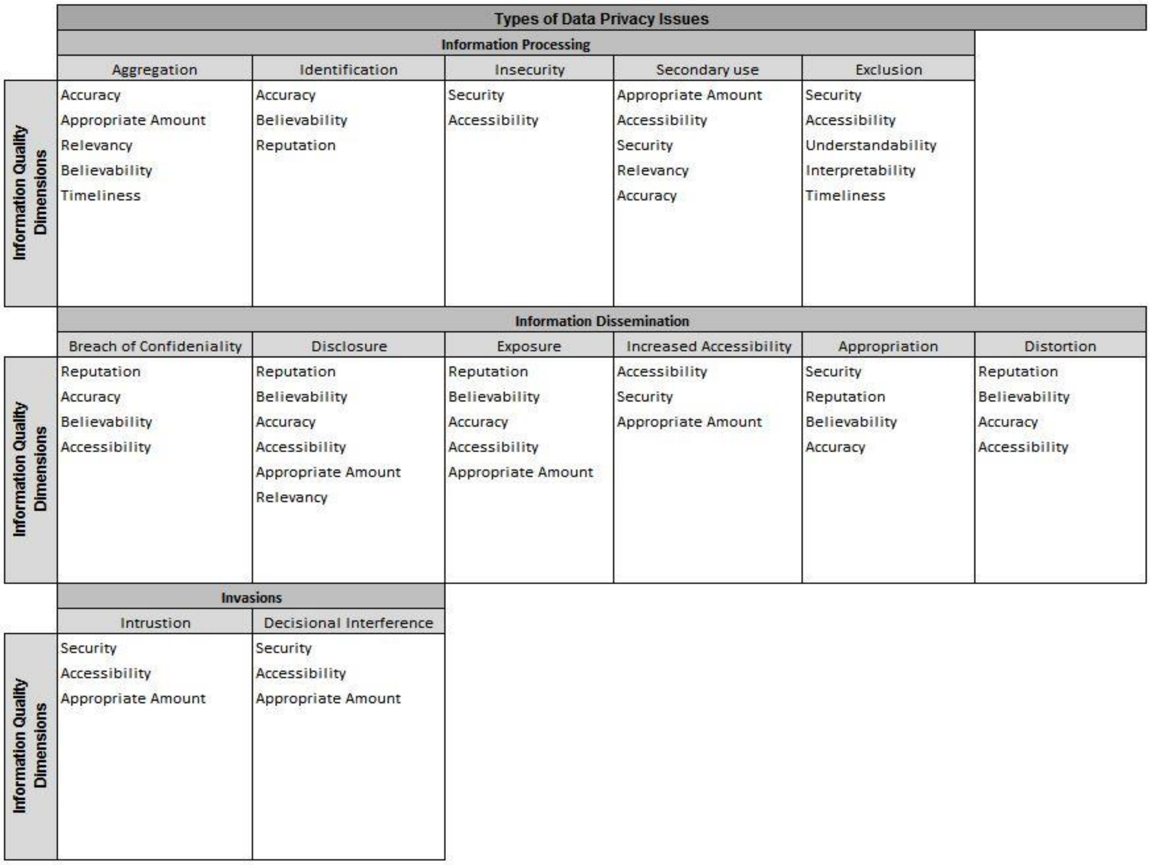
[34] H. Krasnova, O. Günther, S. Spiekermann, S., and K. Koroleva, "Privacy Concerns and Identity in Online Social Networks," Identity in the Information Society, vol. 2, no. 1, pp. 39-63, 2009, doi: DOI 10.1007/s12394-009.

APPENDIX A - FRAMEWORK MATRIX: INFORMATION QUALITY, DATA PRIVACY, AND TRUST IN SOCIAL MEDIA NETWORKS



	Types of Social Networking Data					
	Service Data	Disclosed Data	Entrusted Data	Incidental Data	Behavioral Data	Derived Data
	Data you give the social network site in order to use it	What you post on your own pages	What you post on other people's pages	What other people post about you	Data the site collection about your habits by recording what you do and who you do it with	Data about you that is derived from all other data
Data Privacy Issues	Insecurity Secondary use Breach of Confidentiality	Increased Accessibility Insecurity Appropriation Secondary Use	Increased Accessibility Secondary use Identification Exclusion Breach of Confidentiality Disclosure Exposure Distortion Intrusion (onto their pages)	Identification Exclusion Breach of Confidentiality Disclosure Exposure Distortion Intrusion (onto your pages) Increased Accessibility Secondary use	Aggregation Insecurity Secondary Use Breach of Confidentiality Identification Exclusion	Aggregation Insecurity Secondary Use Breach of Confidentiality Identification Exclusion
Information Quality Dimensions	Accuracy Appropriate Amount Relevancy Security Accessibility Concise Representation Consistent Representation	Accuracy Appropriate Amount Relevancy Security Believability Reputation Understandability Accessibility Objectivity Ease of Operation	Accuracy Appropriate Amount Relevancy Security Believability Reputation Understandability Accessibility Objectivity Ease of Operation	Accuracy Appropriate Amount Relevancy Security Believability Reputation Understandability Accessibility Objectivity Ease of Operation	Accuracy Appropriate Amount Relevancy Security Timeliness Concise Representation Completeness Consistent Representation Accessibility Understandability Interpretability	Accuracy Appropriate Amount Relevancy Security Accessibility Understandability Interpretability Consistent Representation Concise Representation
Trust	Ability Benevolence Integrity	Benevolence Integrity	Benevolence Integrity	Benevolence Integrity	Ability Benevolence Integrity	Ability Benevolence Integrity

APPENDIX B - FRAMEWORK MATRIX: DATA PRIVACY AND INFORMATION QUALITY



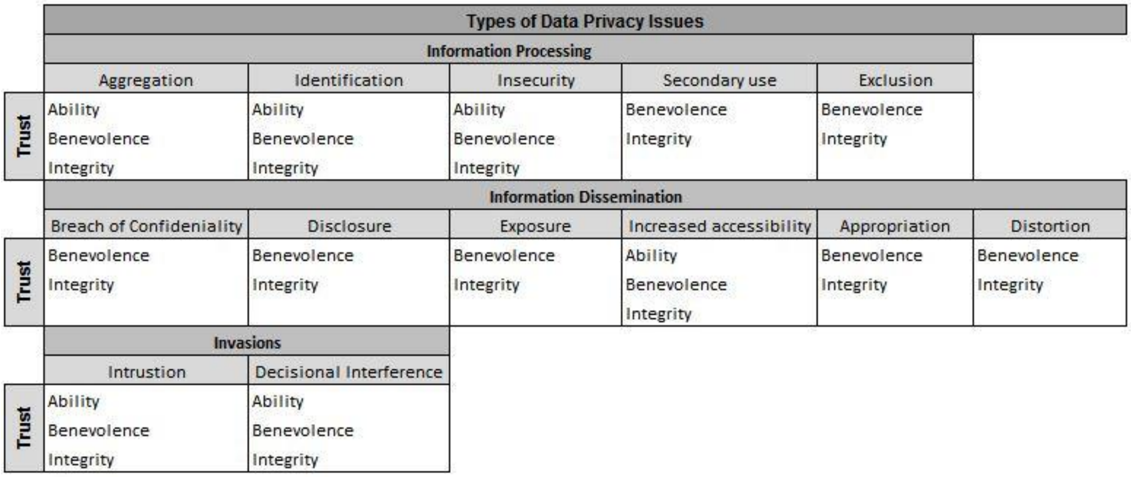
Types of Data Privacy Issues

	Information Processing				
	Aggregation	Identification	Insecurity	Secondary use	Exclusion
Information Quality Dimensions	Accuracy Appropriate Amount Relevancy Believability Timeliness	Accuracy Believability Reputation	Security Accessibility	Appropriate Amount Accessibility Security Relevancy Accuracy	Security Accessibility Understandability Interpretability Timeliness

	Information Dissemination					
	Breach of Confideniality	Disclosure	Exposure	Increased Accessibility	Appropriation	Distortion
Information Quality Dimensions	Reputation Accuracy Believability Accessibility	Reputation Believability Accuracy Accessibility Appropriate Amount Relevancy	Reputation Believability Accuracy Accessibility Appropriate Amount	Accessibility Security Appropriate Amount	Security Reputation Believability Accuracy	Reputation Believability Accuracy Accessibility

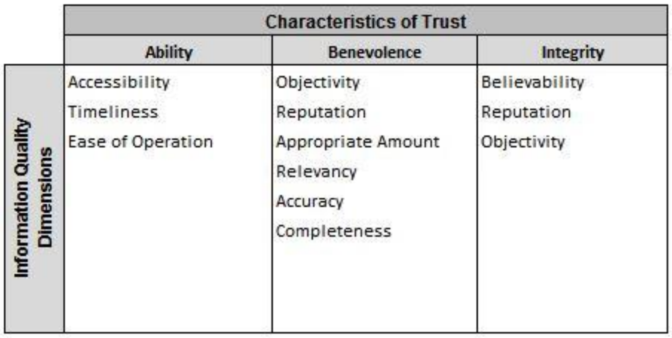
	Invasions	
	Intrustion	Decisional Interference
Information Quality Dimensions	Security Accessibility Appropriate Amount	Security Accessibility Appropriate Amount

APPENDIX C - FRAMEWORK MATRIX: DATA PRIVACY AND TRUST



	Types of Data Privacy Issues					
	Information Processing					
	Aggregation	Identification	Insecurity	Secondary use	Exclusion	
Trust	Ability Benevolence Integrity	Ability Benevolence Integrity	Ability Benevolence Integrity	Benevolence Integrity	Benevolence Integrity	
	Information Dissemination					
	Breach of Confideniality	Disclosure	Exposure	Increased accessibility	Appropriation	Distortion
Trust	Benevolence Integrity	Benevolence Integrity	Benevolence Integrity	Ability Benevolence Integrity	Benevolence Integrity	Benevolence Integrity
	Invasions					
	Intrustion	Decisional Interference				
Trust	Ability Benevolence Integrity	Ability Benevolence Integrity				

APPENDIX D - FRAMEWORK MATRIX: TRUST AND INFORMATION QUALITY



	Characteristics of Trust		
	Ability	Benevolence	Integrity
Information Quality Dimensions	Accessibility Timeliness Ease of Operation	Objectivity Reputation Appropriate Amount Relevancy Accuracy Completeness	Believability Reputation Objectivity