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

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

Abstract—**This research seeks to understand user-based modifications to information quality due to data privacy and trust related concerns within online social networks. It explores the interrelationships and trade-offs between data privacy, trust, and information quality. To this end, we present an extensive literature review to frame our research. The greatest implications of this research come through development of integrated research matrix frameworks, a privacy/trust/information quality modeling syntax, and forthcoming structural equation scoring measures that will be applicable to future research efforts. In application, the relationship matrices can be applied to the conceptual modeling syntax. Further, the results of the structural equation model will show the strength and directionality of the effects of related matrix aspects on one another. The research will enhance methods of modeling and measuring data privacy, trust, and information quality within online social networks. Regarding 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

This work is an extended version of a paper [1] previously presented at the Sixth International Conference on Social Media Technologies, Communication, and Informatics (SOTICS) in Rome, Italy on August 25, 2016.

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 [2]. As the use of social media networks increases there are growing concerns about data privacy. Borcea-Pfitzmann, Pfitzmann, and Berg [3] noted in 2011 that as information technology evolves it greatly influences perceptions and demands regarding 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 per Zittrain [4] 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" [5]. He continued that social network sites succeed because individuals crave community and will share personal information 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 to receive perceived benefits, but fully rational decisions require complete information and most individuals do not 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" [4].

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 [6], 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 regarding privacy, trust, and information quality trade-offs and modifications. From a users' perspective, there is a need to understand both the trust aspects and the visibility of information shared online more fully 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 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 regarding 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. Through these efforts, we hope to provide statistical models for advancing the understanding of privacy, trust, and information quality. 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 literature regarding privacy, trust, information quality, and online social networks. Section III presents a further review of literature as it bears on the interrelated aspects of this research. Section IV presents research methodologies and discusses initial results of the research. Section V summaries research, discusses challenges, and looks at future research opportunities.

II. Background

For better understanding, this section will highlight background literature regarding privacy, trust, information quality, and online social networks.

A. Privacy

According to Daniel Solove in Understanding Privacy [7], 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 [8] 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.

1) Privacy Common Core Conceptualizations

The six major common core conceptualizations reviewed by Solove [7] can be found in Table I and are presented in the following section. Privacy as the right to be let alone is closely tied to Warren and Brandeis as detailed above. Another common view is privacy as limited access to the self. According to Solove, this view is highlighted by Godkin, Bok, Gross, Van Den Haag, O'Brien, and Allan. As noted above, Godkin believed in privacy as the right to decide how much knowledge of personal thoughts and private doings the public at large should be allowed. Bok formulated privacy as protection from unwanted access by others. Van Den Haag, in turn, argued for exclusive access to a realm of one's own. A third common core conceptualization is privacy as secrecy. Posner presented privacy as the right to conceal information or facts about oneself. Similarly, Jourard defined privacy as an outcome of withholding certain knowledge from others. This sets up a dichotomy in which information is either hidden (private) or known (public) and once it is known it can no longer be considered private. Solove noted that this "fails to recognize that individuals may want to keep things private from some people but not others" [7], which is a truth highly relevant to information disclosure in online social network. The fourth conceptualization is privacy as control over personal information. Westin argued that privacy involves determining for oneself when, how, and to what extent information is shared. Miller viewed privacy as control of the circulation of information about oneself. Fried defined privacy not as the absence of information about us, but through the control of that information. This conceptualization is often the focus of privacy systems within online social media networks. Personhood and the right of individuality is the fifth conceptualization. Freund noted that certain attributes that are "irreductible" from self-identity. Protection of individuality and personal dignity is the core of privacy according to Bloustein. Likewise, Benn framed privacy as respect for individuals as choosers. A final common core conceptualization is privacy as intimacy. Gerstein argued that privacy is essential for the formation of intimate relationships. Privacy is extended beyond simple rational autonomy according to Farber. Finally, according to Inness, privacy deals with intimate information, access, and decisions.

TABLE I. PRIVACY COMMON CORE CONCEPTUALIZATIONS [7]

Common Core Conceptualizations		
Who	**What**	**Details**
Warren and Brandeis	The Right to be Let Alone	• The right of each individual to determine to what extent thoughts, sentiments, and emotions can be communicated to others. A general immunity of the person and the right to one's personality.
Godkin, Bok, Gross, Van Den Haag, O'Brien, Allan	Limited Access to the Self	• Right to decide how much knowledge of personal thoughts and private doings the public at large should be allowed (Godkin). • Protection from unwanted access by others (Bok). • Exclusive access to a realm of one's own (Van Den Haag).
Posner, Jourard	Secrecy	• The right to conceal information or facts about oneself (Posner). • Privacy as an outcome of withholding certain knowledge from others (Jourard).
Westin, Miller, Fried	Control over Personal Information	• Determining for oneself when, how, and to what extent information is shared (Westin). • Control of the circulation of information about oneself (Miller). • Not the absence of information about ourselves, but the control of that information (Fried).
Freund, Bloustein, Reiman, Benn	Personhood	• Attributes that are irreductible from oneself (Freund). • Protection of individuality and personal dignity (Bloustein). • Respect for individuals as choosers (Benn).
Farber, Gerstein, Inness	Intimacy	• Privacy as essential for intimate relationships (Gerstein). • Extends privacy beyond simple rational autonomy (Farber). • Privacy deals with intimate information, access, and decisions (Inness).

2) *Privacy Framework Conceptualizations*

Major privacy frameworks have been offered by Solove [7], Nissenbaum [8][9], Holtzman [10], and Rössler [11] (see Table II). 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.

TABLE II. PRIVACY FRAMEWORK CONCEPTUALIZATIONS

Privacy Frameworks Conceptualizations	
Daniel Solove [7] Multi-Dimensional Taxonomy of Privacy	• Privacy as "a cluster of many distinct yet related things." • Information Collection: Surveillance and Interrogation • Information Processing: Aggregation, Identification, Insecurity, Secondary Use, and Exclusion • Information Dissemination: Breach of confidentiality, Disclosure, Exposure, Increased accessibility, Blackmail, Appropriation, and Distortion • Invasions: Intrusion and Decisional interference
David Holtzman [10] The Seven Sins Against Privacy	• Basic Privacy Meanings: Seclusion (right to be hidden), Solitude (right to be left alone), self-determination (right to control information about oneself) • Seven Privacy Sins: intrusion, latency, deception, profiling, identity theft, outing, and loss of dignity • Privacy Torts: Appropriation, Intrusion, Private Facts, False Light
Helen Nissenbaum [8][9] Contextual Integrity	• "Contextual integrity ties adequate protection for privacy to norms of specific contexts, demanding that information gathering and dissemination be appropriate to that context and obey the governing norms of distribution within it." • "A right to live in a world in which our expectations about the flow of personal information are, for the most part, met; expectations that are shaped not only by force of habit and convention, but a general confidence in the mutual support these flows accord to key organizing principles of social life, including moral and political ones."
Beate Rössler [11] Characterization of Different Types of Privacy	• Informational Privacy: Limited access to information, confidentiality, secrecy, anonymity, and data protection • Physical Privacy: Limited access to persons, possessions, and personal property • Decisional Privacy: Decision-making about sex, families, religion, and health-care • Proprietary Privacy: Control over the attributes of personal identity

Daniel Solove is recognized as a global privacy expert with an extensive body of work on the subject. Solove [7] 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. These privacy areas include information collection, information processing, information dissemination, and invasions. Information collection encompasses surveillance and interrogation issues. Information processing encompasses aggregation, identification, insecurity, secondary use, and exclusion issues. Information dissemination encompasses breach of confidentiality, disclosure, exposure, increased accessibility, blackmail, appropriation, and distortion issues. Finally, Invasions encompasses intrusion and decisional interference issues. Further definition details for Solove’s privacy sub-areas can be found in Table III.

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.

TABLE III. A TAXONOMY OF PRIVACY [7]

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

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 [12]. 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 [13]. In general, social media networks can be grouped into categories based on the nature of their social interactions (See Table IV). Examples of popular social network sites include Facebook, LinkedIn, Twitter, YouTube, Flickr, Instagram, and Pinterest. Apps, such as WhatsApp, could also fall under social signaling. With Facebook acquiring WhatsApp, it becomes quite non-trivial for users to understand the privacy aspects of the data sharing policies between WhatsApp and Facebook. More broadly speaking, the constant emergence of new social media apps and their acquisitions or mergers create a highly complex environment for users' awareness of the privacy policies that govern data capturing and sharing.

Boyd and Ellison [14] 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 [13] 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:

> [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. [13]

This research focuses primarily on attribute disclosure, but it may be possible in future research to extend it to the other two areas as well.

Several other classifications of social media data have also been published. Jeremiah Owyang [37] highlights seven types of social media data from a customer marketing perspective. These include demographic, product, psychographic, behavioral, referrals, location, and intention data. From a more structural perspective, Bruce Schneier [15] proposed that social network data can be divided into six categories (see Table IV). Hart and Johnson [16] 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. Regarding privacy, all three of these sources can lead to privacy compromises. Facebook [17] also shares a similarly structured view of data in its published data use policy.

TABLE IV. TYPES OF SOCIAL NETWORK DATA [15]

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

Because of the benefit of its structural divisions, Schneier's framework is used in this research as the foundation for social media network data classification. In addition, from a social media classification perspective, this research will focus on the friendship network aspects of the Social Signaling as illustrated in Table V. To further define the research scope, the modeling aspects of this research will focus on information shared by online social media users via disclosed data, entrusted data, and incidental data per Schneier's framework.

TABLE V. SOCIAL MEDIA CATEGORIES [12]

Social Media Categories	
Social Signaling	Blogs (Wordpress, Blogger), Microblogs (Twitter), Friendship networks (Facebook, MySpace, LinkedIn, Orkut), Snapchat
Social Bookmarking	Del.icio.us, StumbleUpon, Pocket
Media Sharing	Instagram, Flickr, Pinterest, Photobucket, YouTube, Megavideo, Justin.tv, Ustream
Social News	Digg, Reddit
Social Health	PatientsLikeMe, DailyStrength, CureTogether
Social Collaboration	Wikipedia, Wikiversity, Scholarpedia, AskDrWiki
Social Games	Pokémon Go, Foursquare, FarmVille, Second Life, EverQuest (Virtual Worlds)
Q & A	Quora, Yahoo! Answers

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 Computer Science [18]. Within quality literature, the concept of "fitness for use" has been widely adopted as a definition for data quality [6][18]-[21]. But to be applicable, this definition of fitness for use must be contextualized [6]. In this regard, previous writings and research have presented data quality as a multi-dimensional concept [18]-[22].

In 1996, Wang and Strong published an empirical framework to capture the multi-dimensional aspects of information quality that are most important to data consumers [20]. This research was presented in application by Strong, Lee and Wang in "Data Quality in Context" the following year [21]. Since that time, their framework has been widely cited in information quality literature. The Wang Strong Quality Framework [20] 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 VI).

TABLE VI. WANG STRONG QUALITY FRAMEWORK [20]

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

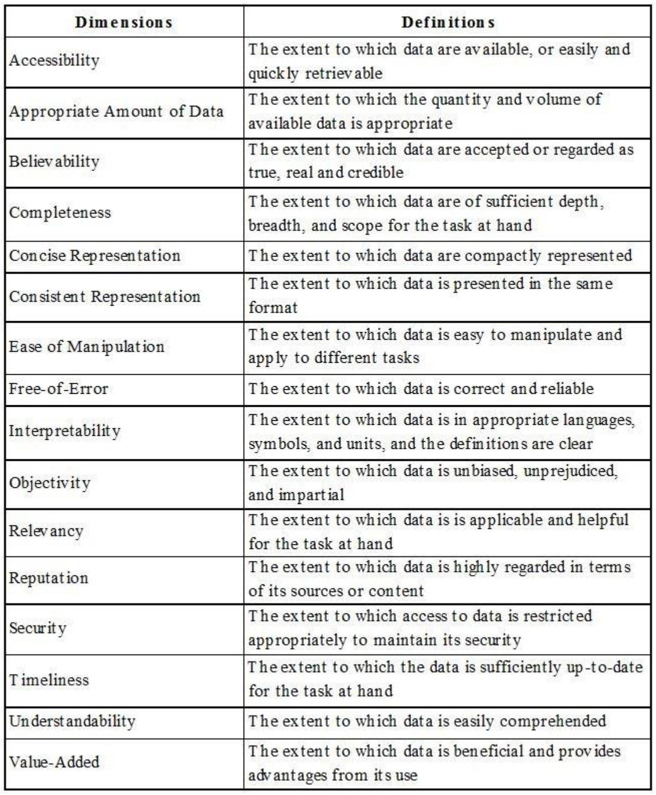
Intrinsic data quality includes the dimensions of Accuracy, Believability, Objectivity, and Reputation. Intrinsic dimensions “have quality in their own right” [20]. Fisher, Lauria, Chengalur-Smith, and Wang [19] describe these as non-contextual self-contained quality aspects.

Contextual data quality includes the dimensions of Value-Added, Relevancy, Timeliness, Completeness, and Amount of Data. Contextual dimensions “must be considered within the context of the task at hand” [20] and are “specifically tied to the particular use or user in order to determine quality” [19].

Representational data quality includes the dimensions of Interpretability, Ease of Understanding, Representational Consistency, Conciseness of Representation, and Manipulability. Representational dimensions relate to the format and meaning of the data [20] and focus on the importance of the presentation and usability of data [19].

Finally, Accessibility data quality includes the dimensions of Access and Security [20] and deal with the availability and protection of data [19]. Definitions of these data quality dimensions from Pipino, Lee, and Wang [22] can be found in Table VII.

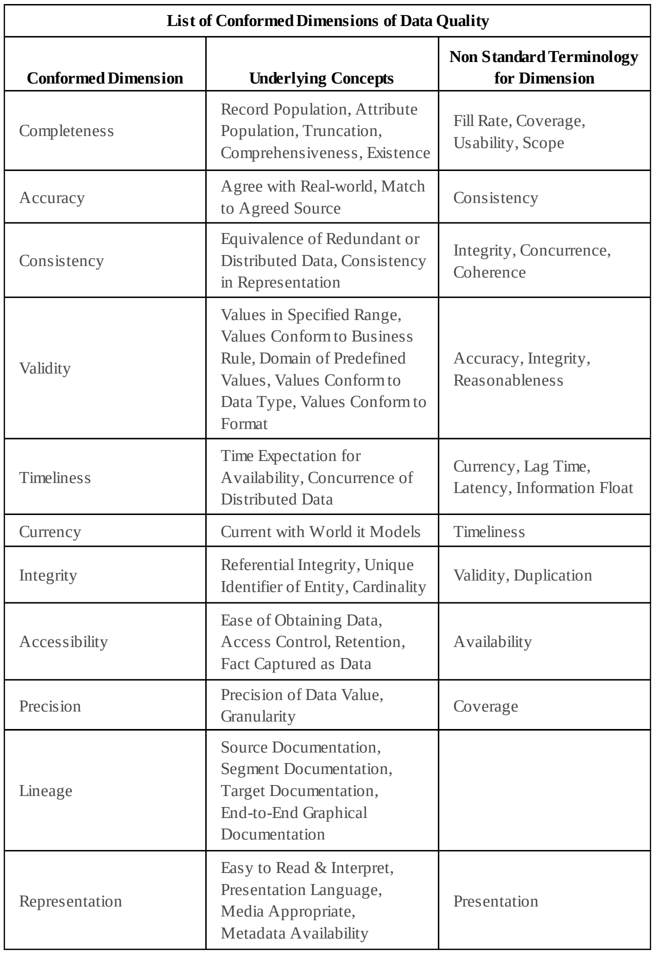
TABLE VII. DATA QUALITY DIMENSIONS [22]



Dimensions	Definitions
Accessibility	The extent to which data are available, or easily and quickly retrievable
Appropriate Amount of Data	The extent to which the quantity and volume of available data is appropriate
Believability	The extent to which data are accepted or regarded as true, real and credible
Completeness	The extent to which data are of sufficient depth, breadth, and scope for the task at hand
Concise Representation	The extent to which data are compactly represented
Consistent Representation	The extent to which data is presented in the same format
Ease of Manipulation	The extent to which data is easy to manipulate and apply to different tasks
Free-of-Error	The extent to which data is correct and reliable
Interpretability	The extent to which data is in appropriate languages, symbols, and units, and the definitions are clear
Objectivity	The extent to which data is unbiased, unprejudiced, and impartial
Relevancy	The extent to which data is is applicable and helpful for the task at hand
Reputation	The extent to which data is highly regarded in terms of its sources or content
Security	The extent to which access to data is restricted appropriately to maintain its security
Timeliness	The extent to which the data is sufficiently up-to-date for the task at hand
Understandability	The extent to which data is easily comprehended
Value-Added	The extent to which data is beneficial and provides advantages from its use

More recent research by Dan Myers [54][55] reviewed the major IQ dimension frameworks found in current literature and worked to conform them into a unified standard. This conformed standard is shown in Table VIII. Myers’ efforts are beneficial and as his conformed standard is further validated and accepted, our proposed framework matrices will likely be updated in future research to align with this standard. Initially, however, the Wang Strong Quality Framework will continue to be the information quality basis for our research framework.

TABLE VIII. CONFORMED DIMENSIONS OF DATA QUALITY [54]



List of Conformed Dimensions of Data Quality		
Conformed Dimension	Underlying Concepts	Non Standard Terminology for Dimension
Completeness	Record Population, Attribute Population, Truncation, Comprehensiveness, Existence	Fill Rate, Coverage, Usability, Scope
Accuracy	Agree with Real-world, Match to Agreed Source	Consistency
Consistency	Equivalence of Redundant or Distributed Data, Consistency in Representation	Integrity, Concurrence, Coherence
Validity	Values in Specified Range, Values Conform to Business Rule, Domain of Predefined Values, Values Conform to Data Type, Values Conform to Format	Accuracy, Integrity, Reasonableness
Timeliness	Time Expectation for Availability, Concurrence of Distributed Data	Currency, Lag Time, Latency, Information Float
Currency	Current with World it Models	Timeliness
Integrity	Referential Integrity, Unique Identifier of Entity, Cardinality	Validity, Duplication
Accessibility	Ease of Obtaining Data, Access Control, Retention, Fact Captured as Data	Availability
Precision	Precision of Data Value, Granularity	Coverage
Lineage	Source Documentation, Segment Documentation, Target Documentation, End-to-End Graphical Documentation	
Representation	Easy to Read & Interpret, Presentation Language, Media Appropriate, Metadata Availability	Presentation

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 [23]-[27]. Marsh [23] 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 [28] 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”.

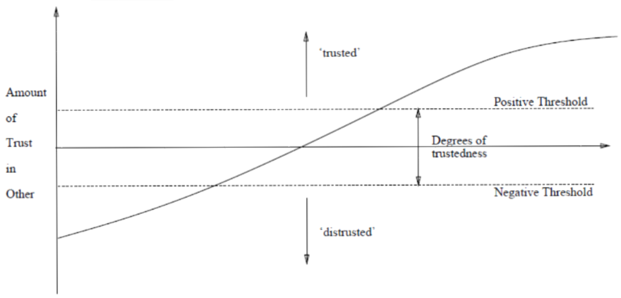


Figure 1. Positive and Negative Thresholds for Trust [23]

Gefen [27], citing Mayer, Davis, and Schoorman [29] defined trust as "a willingness to be vulnerable to the actions of another person or people". Continuing his review of trust literature, Gefen noted that trust is "an important component of many social and business relationships, determining the nature of the interactions and people's expectations of it". Highly relevant to the research being proposed is the role that trust plays in both online social and e-commerce interactions. In specific regard to trust in an online or data driven environment, Bertini [6], defined trust as "the willingness to assume the risks of disclosing data when benefits overcome concerns on the assumption that commitments undertaken by another part will be fulfilled".

Prior research has attempted to unify the disparate definitions and views of trust into various frameworks or models that show the multi-dimensionality of trust. Among these, McKnight and Chervany [25] defined four constructs of trust as well as ten measurable sub-constructs in an interdisciplinary conceptual typology of trust. Their four constructs include: Disposition to Trust meaning "the extent to which one displays a consistent tendency to be willing to depend on general others across a broad spectrum of situations and persons"; Institution Based Trust meaning "one believes the needed conditions are in place to enable one to anticipate a successful outcome in an endeavor or aspect of one's life"; Trusting Beliefs meaning "one believes (and feels confident in believing) that the other person has one or more traits desirable to one in a situation in which negative consequences are possible"; and Trusting Intention meaning "one is willing to depend on, or intends to depend on, the other person in a given task or situation with a feeling of relative security, even though negative consequences are possible."

Carsten D. Schultz [24] in his research presented a situational trust model. He related his work to the trust constructs of McKnight and Chervany [25] and built upon a communication model by Shannon and Weaver published in 1949. Schultz's situational trust model allows for trust to be stated as: "Specific trust is trust placed by a trustor in a trustee concerning a trust object in a trust environment" [24]. Subsequently, Schultz detailed the concept of trust transactions that show the progression cycle from initial trust to resulting trust as it passes through trustworthiness regarding intended behavior, trust in expectation of behavior, and evaluation of actual behavior. Finally, Schultz presents a trust equation that can supplement a given instance of his situational trust model with a reference to previous trust experiences.

Mayer, Davis, and Schoorman [29] 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 [27] 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" [27]. Gefen's research also illustrated the measurability of aspects such as trust regarding interactions in an online domain. This is important to the research at hand.

In specific regard to social networks, Adali et al. [30] 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. In their research, they referenced the concept of embeddedness and highlighted that trust may grow out of increased interactions between individuals. In this regard, they focused on behavioral trust, which they defined as "observed communication behavior in social networks". They further divide behavioral trust into the measurable components of conversational trust based on the communication between two nodes and propagation trust based on the sharing of received information. Other research by Zuo, Hu, & O-Keefe [36] focused on the transferability of trust in social networks through evaluation first of recommendation trust, which is a topical trust based on honest recommendations and second of attribute trust, which is an absolute trust based on general trustworthiness without regard to a specific topic.

E. Interdependencies

Prior research presented by Bertini [6] 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 [51], Hoffman, Novak, and Peralta [31], Neus [52], and Hui, Tan, and Goh [53], 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 [27], Paine et al. [60], and Hoffman, Novak, and Peralta [31], Bertini built on the concept that "[d]ata subjects' level of trust determine both the quantity and the quality of information they disclose" [6] 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 should 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 [28] 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 [23].

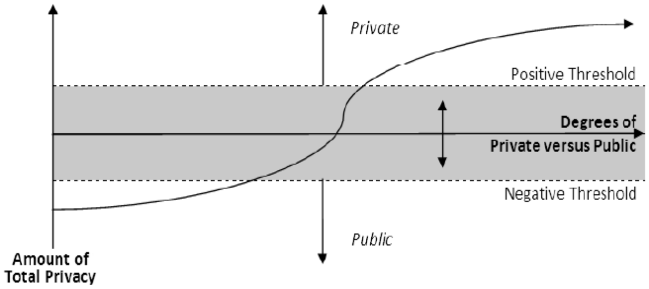


Figure 2. Proposed Thresholds for Privacy [28]

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 [6] 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. Literature Review

Literature has previously been highlighted in background overview of the four components related to this research: privacy, information quality, online social networks, and trust. This section will focus on the review of literature as it bears on the interrelated aspects of this research. Prior research focusing on online social media as it relates to privacy, trust, and quality can be grouped by topic area to include: analysis of user behaviors; privacy related application development; privacy scoring and privacy leakage; and privacy awareness, user control, and privacy visualization.

A. *Analysis of User Behaviors*

In many cases, prior research involved surveys of online social network users. Typically, these surveys focused on attitudes toward privacy, awareness of privacy issues, use of privacy controls, and disconnects between stated beliefs and actual online. Fogel and Nehmad [38] surveyed risking taking, trust, and privacy concerns in a small set of college students. Gross and Acquisti [39] analyzed patterns of information revelation and related privacy implications in a survey of more than 4,000 Carnegie Mellon University students. In further research, Acquisti and Gross [40] analyzed the impact of privacy concerns on behavior, compared stated and actual behavior, and documented behavior changes following exposure to privacy-related information. Hoadley, Xu, Lee, and Rosson [41] surveyed Facebook users soon after the introduction of Facebook's News Feed. This allowed them to explore how easier access to information and "illusory" loss of control can trigger privacy concerns in users. Madejski, Johnson, and Bellovin [42] presented an empirical evaluation based on a small subset of participants measuring privacy attitudes and intentions against actual privacy setting on Facebook. Dwyer, Hiltz, and Passerini [43] surveyed users of both Facebook and MySpace regarding perceptions of trust and privacy concerns as well as willingness to share information and develop new relationships. Andrew Boyd [44] presented a two-year longitudinal study of social media users to examine privacy attitudes and self-reported behaviors over time. He extended the Internet Users‘ Information Privacy Concern model (IUIPC) model for applicability within Social Networking sites to investigate influences on online attitudes and behaviors regarding privacy. Boyd found that with time privacy concerns and distrust increased while willingness to disclose personal information decreased. Another longitudinal study conducted by Dey, Jelveh, and Ross [45], used web crawling rather than user surveys to explore privacy trends for personal attributes available on public Facebook profile pages. They found that users had become dramatically more private between March 2010 and June 2011. They cited media attention and Facebook’s redesigned privacy page as key factors in this trend. Wisniewski, Knijnenburg, and Lipford [59] analyzed online social network users against 36 privacy behaviors and 20 feature awareness items to categorize users into six distinct privacy management strategies. Aspects of these strategies parallel our research well. This prior research generally focused on how privacy awareness affected the use of privacy controls and the overall disclosure of information. These aspects will be incorporated in the conceptual and structural models for this research, but this research will also expand on this by looking more fully at modification to the quality of information in the face of privacy awareness.

B. Privacy Related Application Development

Several types of privacy related applications are also presents in current literature. These include user interface concepts, APIs for controlling and/or visualizing privacy settings, and stand-alone privacy driven social network concepts. Concepts from this research may be extended into application development in the future, but it is beyond the scope of this current proposal.

C. Privacy Awareness, User Control, and Visualization

Current literature shows a strong focus on increasing awareness and understanding of privacy issues through visualization and user controls. Kolter and Pernul [46] presented a method for generating privacy preferences. They focused their research on awareness of what information websites and online services are seeking and the corresponding ability of users to minimize the amount of data they release as well as control and restrict how their disclosed data is used by the collecting service or passed on to third-party services. Krishnamurthy and Willis [47] highlighted the need for bit or data element level privacy controls noting that “[l]imiting access to just friends or those in a network is not fine-grained enough”. They proposed that each set of interactions in an online social network should indicate the bare minimum of private information required. This would allow users to set automated interaction thresholds based on their personal privacy thresholds as well as directly control access when additional information is requested. Acquisti and Gross [40] summarized that “the majority of [Facebook] members claim to know about ways to control visibility and searchability of their profiles, but a significant minority of members are unaware of those tools and options”. Hart and Johnson [16] noted that while users often disclose data directly, personal information can also be revealed accidently through aggregation of information, shared by service providers, or published by others. They further noted that users are often unaware of the impacts of their information disclosure or even when understood they do not want to expend the effort needed to utilize access control systems. Hart and Johnson proposed that a well-designed privacy preference system must achieve multiple goals: a) Allow users to specify viewers, b) Allow succinct polices to apply to large content collections, c) Utilize flexible access control policies, and d) Infer restricted privacy policies on new content. Offenhuber and Donath [48] developed ways to represent the individuality of nodes and links that comprise social networks. They focused on representing the actual activity and message exchanges between nodes to give context to generic high-level connections within a social network. Borcea-Pfitzmann, Pfitzmann, and Berg [3] proposed that privacy could only be preserved or regained through a combination of data minimization, user control, and contextual integrity. Finally, in more recent literature, Mármol, Pérez, and Pérez [56] discussed the user awareness and control aspect of reporting offensive content in social networks. They presented a reputation-based assessment approach to the flagging of content by users.

In extension of this prior research, the development of relationship matrices for data privacy, online social network data, trust, and information quality in this research will allow for more targeted awareness of privacy issues and specific focus areas for privacy controls. The development of syntax for conceptual modeling in turn lends itself, as Krishnamurthy and Willis highlighted, to better understanding a data element level view of information disclosure. Understanding gained through the development of a structured equation model will lend itself to measuring aspects of data minimization and user control in application within online social networks.

D. Privacy Scoring and Privacy Leakage

Becker and Chen [49] state that the prevention of information from going beyond its intended privacy boundaries is basic principle in computer science and that information escaping these boundaries is known as information leakage. Their research sought to measure and limit privacy risk attributed to friend connections within an online social network. The concept of risk attributed to online social network connections will be addressed in this research through the components of users’ privacy and trust in the conceptual model syntax. Irani, Webb, Pu, and Li [50] focused on the aggregation of information leakage across multiple networks, which they defined as the social footprint of an online identity. Through this, they developed measures of attribute leakage. In this research, information leakage through aggregation will be noted as a privacy concern in the overall relationship matrices, but it will lie outside the scope on the final stages of the research. Lui and Terzi [33] proposed a framework for computing user privacy scores that indicate the potential privacy risk due to social network participation. Their research utilizes concepts from Item Response Theory and their methodology incorporates both the sensitivity and visibility of individual data elements into the calculation of an aggregated privacy score. The concepts of Liu and Terzi strongly influenced the development of the conceptual syntax for this proposed research. Ros, Canelles, Pérez, Mármol, and Pérez [57] presented a method for optimized, delay-based posting in online social networks as a privacy protection against observed activity that may reveal time-sensitive details. Their paper can be related to this current research in that delay-based posted is a modification to the timeliness dimension of information quality as a method of privacy protection. Finally, Parra-Arnau, Rebollo-Monedero, and Forné [58] addressed privacy risks and proposed quantitative privacy measures of users' profiles.

IV. Methodology and Results

The research will contain three interconnected components. The first is 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 is 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. In application, the overlapping aspects of privacy, information quality, and trust in the relationship matrices can be applied to the expanded modeling syntax as illustrated in Fig. 6. Further, the results of the structural equation model will show the strength and directionality of the related matrix aspects' effects on one another. While these components can be generalized across multiple online social networks, for this research, when analyzing online social networks, Facebook will be used as the primary point of reference when talking about social media structures because of the size and activity levels of its user base.

A. Framework Matrices

This research focuses 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 regarding both online social networks and each other. This phase of the research hypothesizes that:

H1: The multi-faceted dimensions, aspects, and properties of trust, privacy, and information quality can be effectively overlaid within a series of related matrices.

H2: An understanding of intersections of these sub-aspects lends itself to a broader understanding of the relationship of these concepts.

H3: 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 [7] taxonomy of privacy, Schneier's [15] divisions of social network data, Wang and Strong's [20] multiple dimensions of information quality, and the trustworthiness characteristics of Ability, Benevolence, and Integrity as presented by Mayer, Davis, and Schoorman [29] and Gefen [27]. 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 [32], 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 will be undertaken. Their opinions in relation to the framework matrices will be gathered and reconciled. The framework matrix will be further validated in the proposed structured equation modeling phase of this research as the trade-offs between framework relationships are measured.

B. Syntax and Conceptual Modeling

Regarding 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 [33] 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. In support of this effort, the finalized syntax and conceptual model will be presented in an ontology language, such as OWL2, rather than in the simplified form presented here. This phase of the research hypothesizes that:

H4: 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.

H5: 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 can be implemented in a way that will 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 presented as a concept to be further developed in future research.

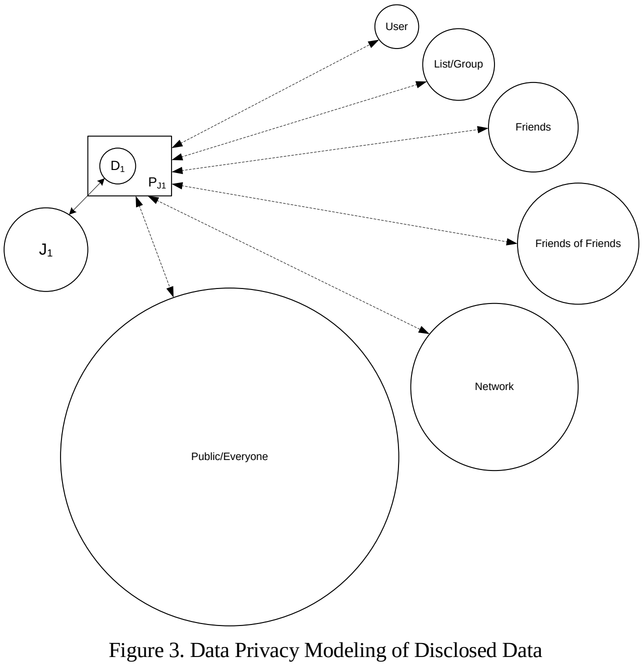


Figure 3. Data Privacy Modeling of Disclosed Data

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 user's determination of visibility related privacy settings. Disclosed data 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. This is shown in Fig. 3 with possible user and group related visibility settings illustrated.

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 terms of data visibility. Entrusted data 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. This is shown in Fig. 4 with possible user and group related visibility settings illustrated.

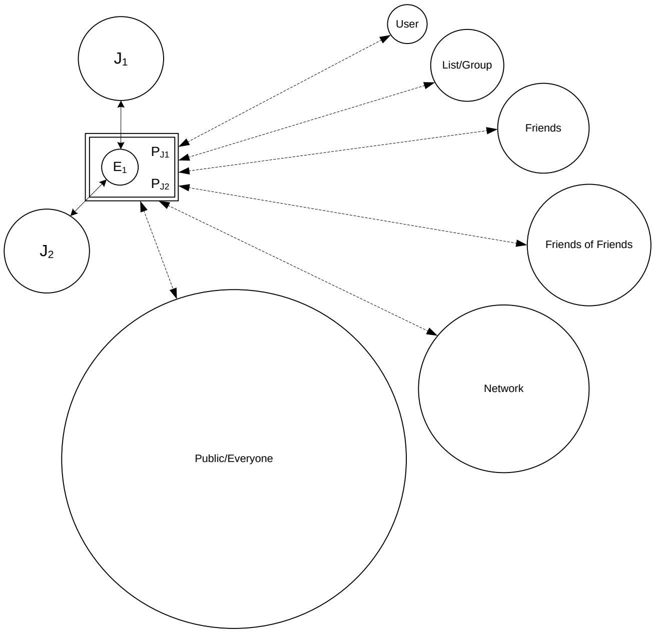


Figure 4. Data Privacy Modeling of Entrusted Data

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 that is the subject of a posted data element. Exclusion relates to the level of control and involvement a user has regarding information that is shared about or actions taken that affect him or her. Within online social networks, this relates to whether 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.

Incidental data 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. This is shown in Fig. 5 with possible user and group related visibility settings illustrated.

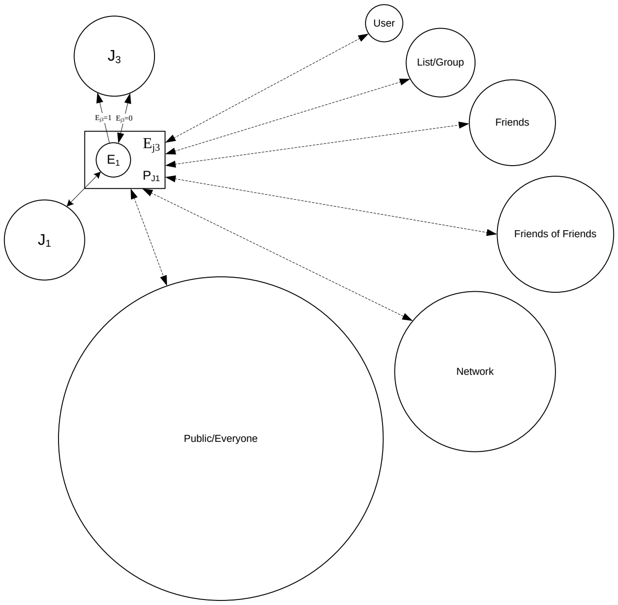


Figure 5. Data Privacy Modeling of Incidental Data

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 expanded 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. This expanded syntax is shown in Fig. 6.

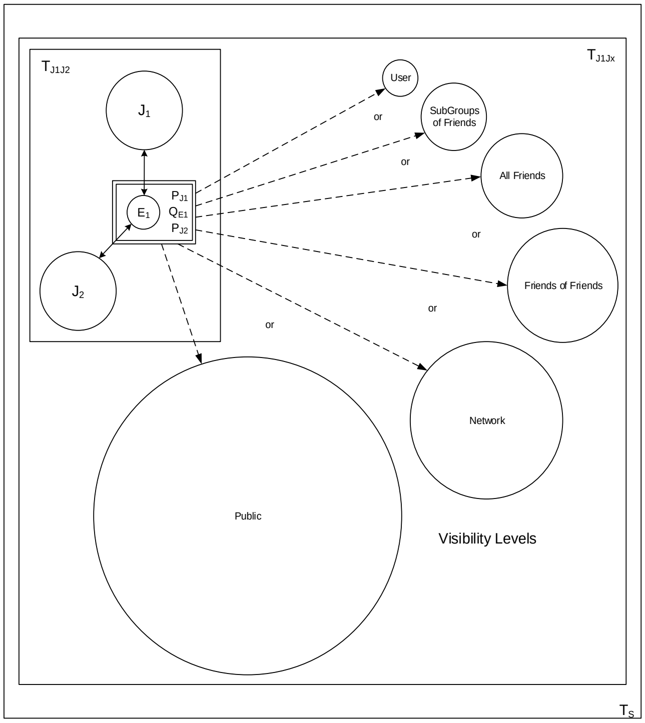
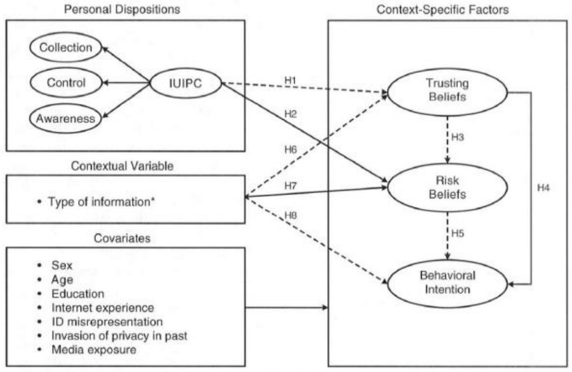


Figure 6. Data Privacy Modeling of Entrusted Data with Trust and Quality

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 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 has shown the benefit of structural equation models in the development and validation of the Internet Users' Information Privacy Concerns [34] and User Privacy Concerns and Identity in OSNs [35] constructs. This research will also use structural equation modeling to extend and build upon those concepts.

As seen in Fig. 7, Malhotra, Kim, and Agarwal [34] 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 regarding Internet usage. This 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.



Notes. *Less sensitive information (0), more sensitive information (1), positive effect ⟶, negative effect - - -→.

Figure 7. Proposed Model by Malhotra, Kim, and Agarwal [34]

Krasnova, Günther, Spiekermann, and Koroleva [35] 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 high-level research model (see Fig. 8) 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.

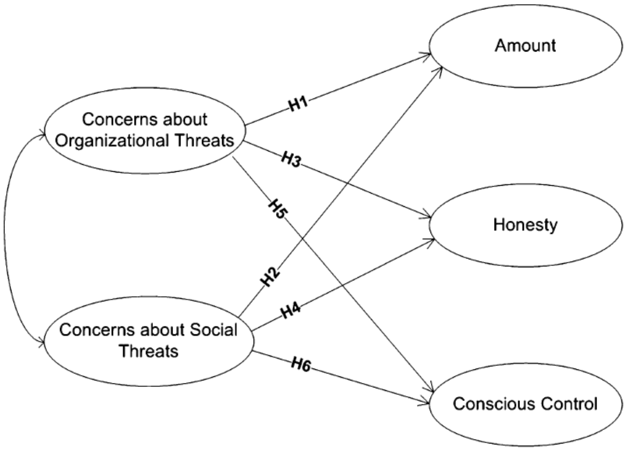

Amount
Concerns about Organizational Threats
Honesty
Concerns about Social Threats
Conscious Control
H1
H2
H3
H4
H5
H6

Figure 8. PCIOSN Research Model [35]

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 [35] to specific IQ dimensions, as well as incorporate other relevant IQ dimensions from the proposed framework matrix. Of additional research interest is whether the IUIPC and PCIOSN models can be incorporated into a single view through the modeling aspects of this research. This research hypothesizes that:

> **H6:** 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. 9).
>
> **H7:** Data element types (wall posts, photos, comments, shares, likes, check-ins, etc.) have measurably different thresholds for content sensitivity.
>
> **H8:** 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.

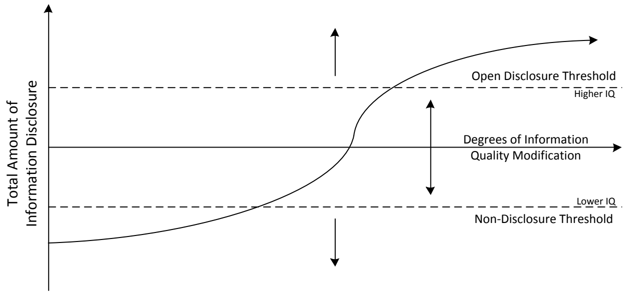


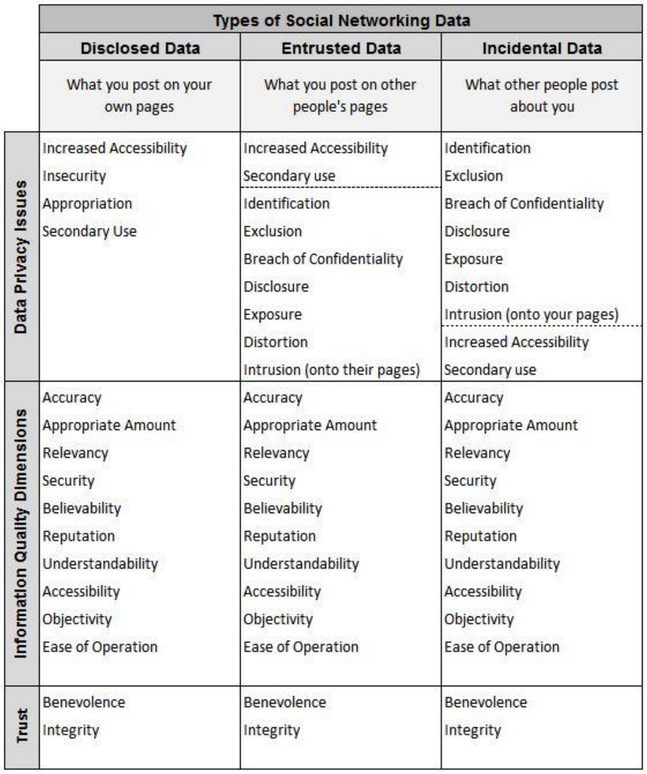
Figure 9. Initial Information Quality Modification Concept

Hypothesis 6 is an extension of Marsh’s Positive and Negative Thresholds for Trust [23] and Kosa’s Proposed Thresholds for Privacy [28] 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. 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 (SEM) will be utilized as a method for measuring the balance trade-offs present between specific trust characteristics, information quality dimensions and data privacy aspects. Structural Equation Modeling validation typically includes confirmatory factor analysis, as well as assessment of the internal consistency, convergent validity, and discriminant validity of the measured constructs. Multiple aspects of the research survey instrument needed to perform the SEM analysis will be based on results of the first two components of this research. As the framework matrix is validated and the syntax and conceptual model are designed, the survey instrument for SEM analysis will be finalized.

V. Conclusions, Challenges, and Opportunities

This paper presents an ongoing research effort. 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 and is being 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.

TABLE IX. 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

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 IX. This table 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 initially 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.

A. *Research Contributions and Implications*

To date, relationship matrices for data privacy, online social networks, information quality, and trust as a research framework has been developed and presented here. The framework is currently being validated via a survey of experts. We fully intended to include the results of our validation survey here, but those results have been delayed and will instead be presented in a forthcoming paper. An initial conceptual model and syntax for data privacy, trust, and information quality in online social networks has also been developed and shared. Furthermore, an Initial Information Quality Modification Concept has been presented in extension of Marsh’s Positive and Negative Thresholds for Trust and Kosa’s Proposed Thresholds for Privacy.

The greatest implication of this research is its applicability to future research efforts. This research could enhance methods of modeling and measuring privacy, trust, and information quality within online social networks. It will lend itself to a better understanding of the quality of shared information in given data privacy and trust scenarios. Finally, it will provide future researchers with a formal framework for relating privacy, information quality, and trust in online social networks as well as a method for understanding information quality modification.

B. *Limitations and Challenges*

First, while a broad framework matrix can be presented, the scope for validation and deeper research is 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, 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 must 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.

C. *Future Research Opportunities*

For the next phase of this research, a structural equation model for understanding the trade-offs and influences between data privacy, trust, and information quality in online social networks is being developed. A survey will be undertaken to validate the model. Results from these efforts will then be expressed in application via the presented conceptual model and syntax after it is formalized in an ontology language such as OWL2.

Future research is likely to include expanded validation of different areas of overlap within framework matrices. It would be of interest to explore application of this research beyond the user-controlled aspects such as Disclosed, Entrusted, and Incidental data to include Service, Behavioral, and Derived data within online social networks. Finally, updating the presented research framework matrices to fit new research as it develops, such as the Conformed Dimension of Data Quality, will keep this research applicable.

REFERENCES

[1] B. Blake and N. Agarwal, "Understanding User-Based Modifications to Information Quality in Response to Privacy and Trust Related Concerns in Online Social Networks," The Sixth International Conference on Social Media Technologies, Communication, and Informatics (SOTICS), pp. 18-28, 2016.

[2] 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), pp. 1257-1260, 2010.

[3] 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/ing/informatik/sya/ps/die-professur/beschaeftigte/kbo_de. 2017.05.29.

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

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

[6] 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/. 2017.05.29.

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

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

[9] 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. 2017.05.29.

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

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

[12] N. Agarwal, Types of Social Media, lecture presented for Social Media Mining and Analytics course at the University of Arkansas at Little Rock, 2016.

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

[14] 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.

[15] 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

[16] 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 2017.05.29.

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

[18] 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.

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

[20] 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.

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

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

[23] S. P. Marsh, Formalising Trust as a Computational Concept, unpublished doctoral dissertation, University of Stirling, 1994. [Online]. Available from: https://dspace.stir.ac.uk/ 2017.05.29.

[24] 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, pp. 50:1-50:7, 2006.

[25] 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, p. 10, 2001.

[26] 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.

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

[28] 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.

[29] 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.

[30] 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.

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

[32] 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 2017.05.29

[33] 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.

[34] 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.

[35] 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.

[36] Y. Zuo, W. Hu, & T. O'Keefe. “Trust Computing for Social Networking,” Sixth International Conference on Information Technology: New Generations (ITNG '09), pp. 1534-1539, 2009, doi: 10.1109/ITNG.2009.278

[37] J. Owyang, “7 Types of Social Data that Help You Understand Consumers,” Lecture presented at Eleven Social Media Tips for 2011, Feb. 2011. [Online]. Available from: http://netbase11for11.com/

[38] J. Fogel and E. Nehmad, “Internet social network communities: Risk taking, trust, and privacy concerns,” Computers in Human Behavior, 25(1), pp. 153-160, 2009.

[39] R. Gross and A. Acquisti, “Information revelation and privacy in online social networks,” ACM Workshop on Privacy in the Electronic Society (WPES '05), pp. 71-80, 2005.

[40] A. Acquisti and R. Gross, “Imagined Communities: Awareness, Information Sharing and Privacy on The Facebook,” The 6th Workshop on Privacy Enhancing Technologies, pp. 1-22, 2006.

[41] C.M. Hoadley, H. Xu, J.J. Lee, and M.B. Rosson, “Privacy as Information Access and Illusory Control: The Case of the Facebook News Feed Privacy Outcry,” Electronic Commerce Research and Applications, 9(1), pp. 50-60, 2010.

[42] M. Madejski, M. Johnson, and S.M. Bellovin, The Failure of Online Social Network Privacy Settings, Available from: https://mice.cs.columbia.edu/getTechreport.php?techreportID=1459

[43] C. Dwyer, S. Hiltz, and K. Passerini, “Trust and Privacy Concern Within Social Networking Sites: A Comparison of Facebook and MySpace,” The Thirteenth Americas Conference on Information Systems, 2007. [Online]. Available from: csis.pace.edu/~dwyer/research/. 2017.05.29.

[44] A.W. Boyd, “A Longitudinal Study of Social Media Privacy Behavior,” ArXiv E-prints, pp. 1-10. [Online]. Available from http://arxiv.org/abs/1103.3174. 2017.05.29.

[45] R. Dey, Z. Jelveh, and K. Ross, “Facebook Users Have become Much More Private: A Large-Scale Study,” The 4th IEEE International Workshop on Security and Social Networking (SESOC), pp. 1-7, 2012. [Online]. Available from http://cis.poly.edu/~ratan/ 2017.05.29

[46] J. Kolter and G. Pernul, G. (2009). “Generating User-Understandable Privacy Preferences,” International Conference on Availability, Reliability and Security (AES ’09), pp. 299-306, 2009. doi: 10.1109/ARES.2009.89

[47] B. Krishnamurthy and C.E. Wills, “Characterizing Privacy in Online Social Networks,” The First Workshop on Online Social Networks (WOSN '08), pp. 37-42, 2008.

[48] D. Offenhuber and J. Donath, “Comment Flow: Visualizing Communication Along Network Path,” Interface Cultures: Artistic Aspects of Interaction, 2008. [Online]. Available from: medialab-prado.es/mmedia/1094 2017.05.29

[49] J. Becker and H. Chen, “Measuring Privacy Risk in Online Social Networks,” Web 2.0 Security and Privacy (W2SP 2009), 2009. [Online]. Available from http://w2spconf.com/ 2017.05.29

[50] D. Irani, S. Webb, C. Pu, and K. Li, “Modeling Unintended Personal-Information Leakage from Multiple Online Social Networks,” Internet Computing, 15(3), pp. 13-19, 2011. doi: 10.1109/MIC.2011.25

[51] E. Rose, “Balancing Internet Marketing Needs with Consumer Concerns: A Property Rights Framework,” ACM SIGCAS Computers and Society, 31(1), pp. 17-21, 2001.

[52] A. Neus, “The Quality of Online Registration Information: Factors Influencing User Decisions to Reveal Authentic Personal Information to Online Marketers as Part of a Perceived Barter,” MIT Conference on Information Quality (IQ 2000), 2000.

[53] K.L. Hui, B.C.Y. Tan, and C.Y. Goh, “Online Information Disclosure: Motivators and Measurements,” ACM Transactions on Internet Technology, 6(4), pp. 415-441, 2006.

[54] D. Myers, “Conformed Dimensions of Data Quality,” DQMatters, 2017. [Online]. Available from: http://dimensionsofdataquality.com 2017.05.29

[55] D. Myers, "The Value of Using the Dimensions of Data Quality." Information Management, Aug. 2013. [Online]. Available from: https://www.information-management.com/news/the-value-of-using-the-dimensions-of-data-quality 2017.05.29

[56] F.G. Marmol, M.G. Perez, and G.M. Perez, “Reporting Offensive Content in Social Networks: Toward a Reputation-Based Assessment Approach,” IEEE Internet Computing, 18(2), 32-40, 2014. doi:10.1109/mic.2013.132

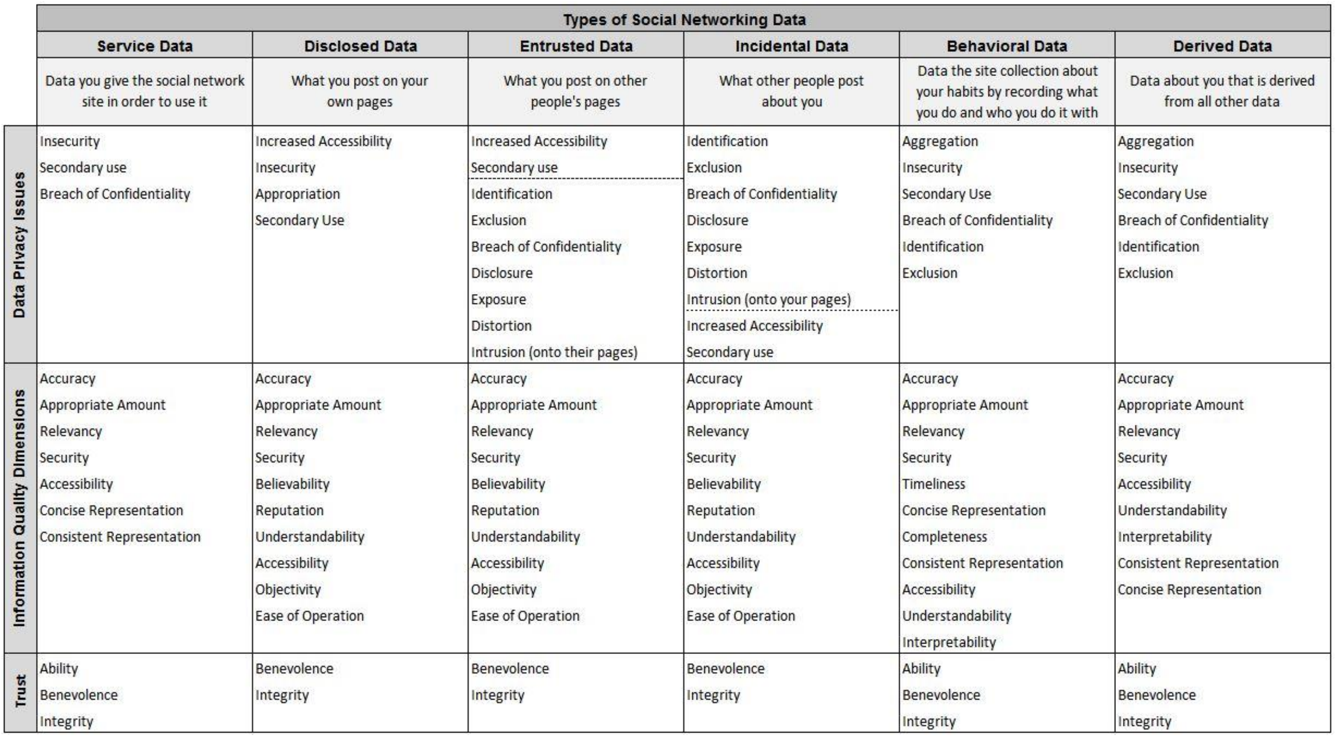
[57] S.P. Ros, A.P. Canelles, M.G. Pérez, F.G. Mármol, and G.M. Pérez, “Chasing Offensive Conduct in Social Networks,” ACM Transactions on Internet Technology, 15(4), 1-20, 2015. doi:10.1145/2797139

[58] J. Parra-Arnau, D. Rebollo-Monedero, and J. Forné, “Measuring the privacy of user profiles in personalized information systems,” Future Generation Computer Systems, 33, 53-63, 2014. doi:10.1016/j.future.2013.01.001

[59] P.J. Wisniewski, B.P. Knijnenburg, and H.R. Lipford, “Making privacy personal: Profiling social network users to inform privacy education and nudging,” International Journal of Human-Computer Studies, 98, 95-108, 2017. doi:10.1016/j.ijhcs. 2016.09.006

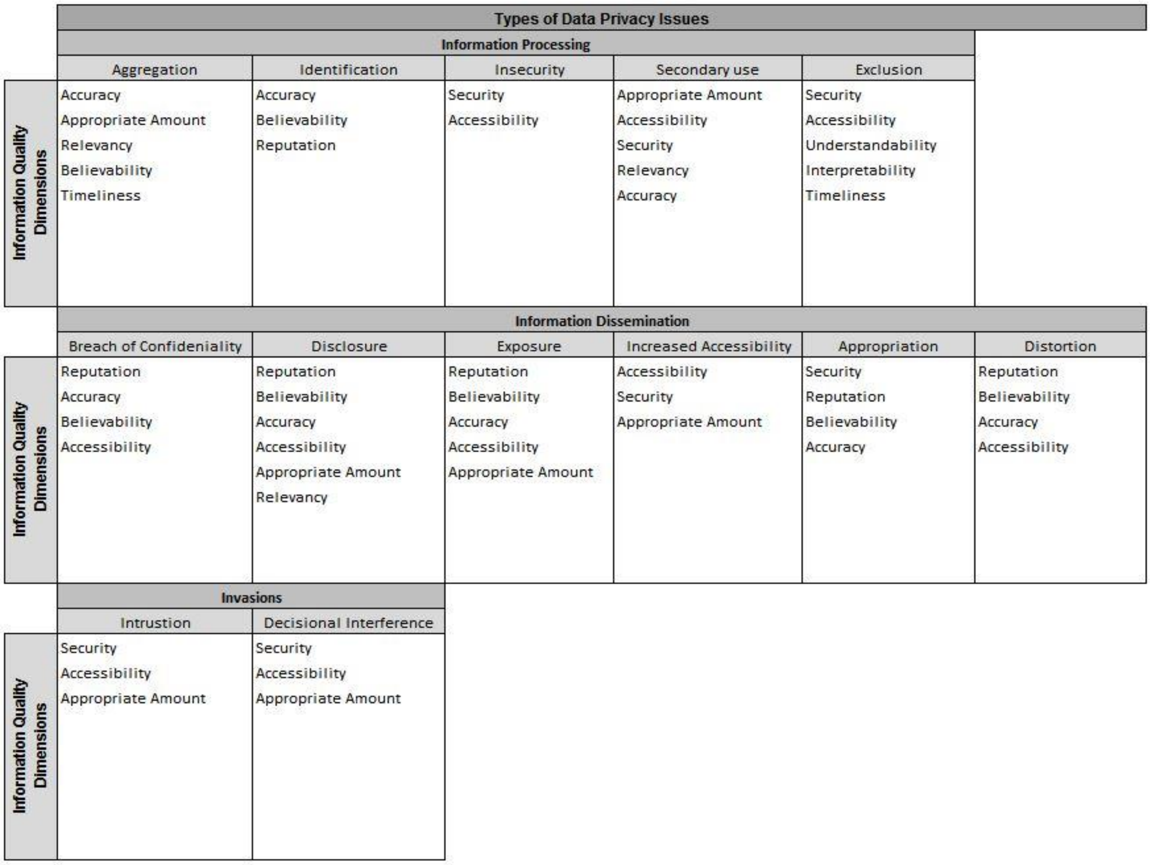
[60] C.B. Paine Schofield, A.N. Joinson, T. Buchanan, and U.-D. Reips, “Privacy and self-disclosure online,” Conference on Human Factors in Computing Systems, 2006.

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

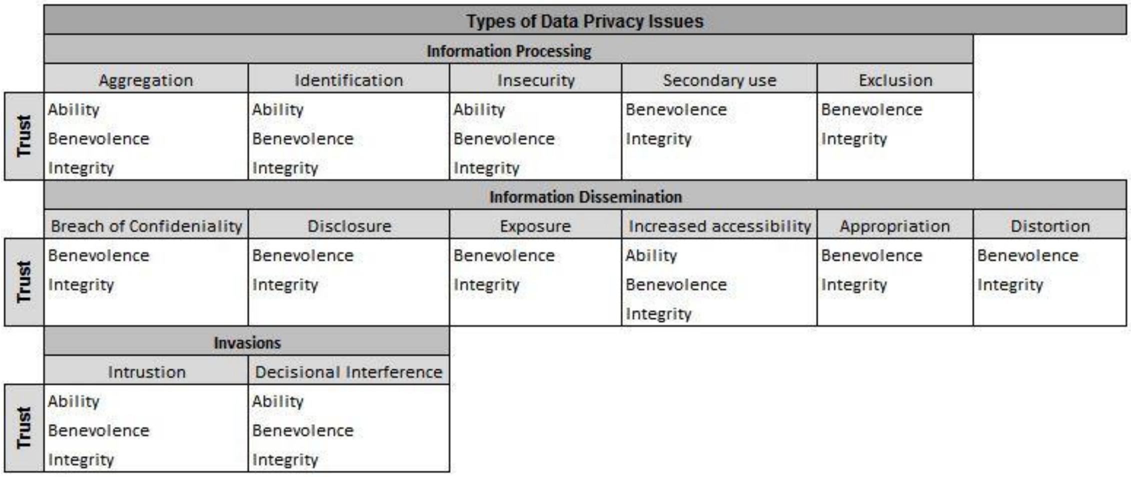


	Types of Data Privacy Issues				
	Information Processing				
	Aggregation	Identification	Insecurity	Secondary use	Exclusion
Information Quality Dimensions	Accuracy Appropriate Amount Relevancy Believability Timeliness	Accuracy Believability Reputation	Security Accessibility	Appropriate Amount Accessibility 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

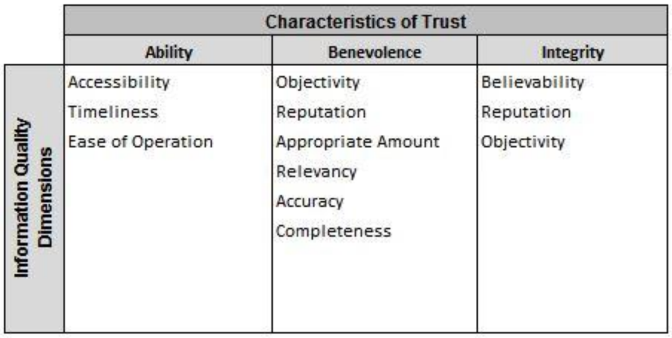
	Invasions	
	Intrustion	Decisional Interference
Information Quality Dimensions	Security Accessibility Appropriate Amount	Security Accessibility Appropriate Amount

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