SI: Networked Influence

SM+S
social media + society

Social Media + Society
July-September 2019: 1–14

DOI: 10.1177/2056305119866008
journals.sagepub.com/home/sms
SAGE

We Care About Different Things: Non-Elite Conceptualizations of Social Media Privacy

Kelly Quinn[1], Dmitry Epstein[2], and Brenda Moon[3]

Abstract

This study explores privacy from the perspective of the user. It leverages a "framing in thought" approach to capture how users make sense of privacy in their social media use. It builds on a unique dataset of privacy definitions collected from a representative sample of 608 US social media users. The data are analyzed using topic modeling and semantic network analysis to unpack the multidimensionality of social media privacy. These dimensions are further examined in relation to established demographic antecedents of privacy concerns and behaviors. Results indicate the dominance of frames related to horizontal dimensions of privacy, or privacy vis-à-vis peers, as compared with the vertical dimensions, or privacy vis-à-vis institutions. In addition, the findings suggest that user conceptualization of privacy reflects a cognate-based approach that emphasizes control and limits to information access. Implications for privacy research, policy, and technology design are discussed.

Keywords

privacy, user definitions, social media, framing

Privacy has become the de facto currency of the social media world. People routinely disclose information, which not too long ago was considered private, in exchange for digital tools and services. At the same time, the concept itself is surprisingly fluid (BeVier, 1995), its interpretation and enactment are highly contextual (Nissenbaum, 2010), and often do not align. While previous research has looked into the attitudinal and behavioral aspects of privacy (Acquisti, Brandimarte, & Loewenstein, 2015; Smith, Dinev, & Xu, 2011), or into the strategic deployment of its various meanings in policy deliberations (Epstein, Roth, & Baumer, 2014), there is limited inquiry into how social media users themselves perceive and interpret the idea of privacy. In other words, there is a need not merely to examine attitudes or concerns toward privacy, but also to explore how individuals understand privacy as an idea they may care about.

While previous research has examined privacy antecedents, attitudes, and behaviors in a plurality of contexts, the question of definition remains an unresolved debate in the field. Most empirical work drives on deductively constructed and measurable ideas of privacy, such as those mentioned above. As a result, privacy is typically treated primarily as a unidimensional construct, which limits both theoretical understanding of this phenomenon and the potential for policy interventions. Fewer studies have explicitly addressed the definition of privacy in an inductive fashion. In this article, we use recently collected survey data in an attempt to unpack how users of social media frame privacy. We chose to exclusively focus on social media both for the practical need to bound our inquiry and for the cultural and political significance of social media platforms in contemporary society. Our goal is to help conceptualize social media privacy in a participant-centric way, thus further enhancing efforts to theorize privacy, advance privacy research, and design privacy-sensitive tools and policy solutions.

Literature Review

In our effort to map the frames of references to privacy used by social media users, we draw on two main bodies of literature. First, we discuss framing literature, which offers an

[1]University of Illinois at Chicago, USA
[2]The Hebrew University of Jerusalem, Israel
[3]Queensland University of Technology, Australia

Corresponding Author:
Kelly Quinn, Department of Communication, University of Illinois at Chicago, 1007 W Harrison Street, BSB 1140, MC 132, Chicago, IL 60607, USA.
Email: kquinn8@uic.edu

established conceptual framework for understanding the importance of definitions for both behavioral and structural outcomes. Second, we discuss relevant privacy literature in order to situate our current work in earlier efforts to tease out the plurality and context-driven fluidity of privacy definitions. Finally, we review literature that has dealt explicitly with privacy framing to situate the unique contribution of this work.

Framing

Framing is a popular analytical framework, used in a variety of fields (e.g., psychology, political science, sociology, communication), across levels of analysis (e.g., individual, institutional), and with different epistemological approaches (Chong & Druckman, 2007; Scheufele, 1999). Goffman (1974) has originally referred to frames as primary "schemata of interpretation" that allow "its user to locate, perceive, identify, and label a seemingly infinite number of concrete occurrences defined in its terms" and within her personal context (p. 21). Gamson and Modigliani (1989) explained that frames present "a central organizing idea [. . .] for making sense of relevant events, suggesting what is at issue" and thus giving meaning to puzzles and ideas (p. 3).

Frames operate at both systemic (macro) and individual (micro) levels. At the macro level, competing social, political, and cultural actors engage in constructing, modifying, and disseminating frames both intentionally and unintentionally. In doing so, the actors operate with *frames in communication* as they engage in competitive behavior around frames of reference that align with their values, goals, and interests (Chong & Druckman, 2007; Scheufele, 1999). At the micro level, *frames in thought* delineate the realm of the possible or desirable by restricting and prioritizing a set of available considerations or by amending their subjective value and applicability in the eyes of an individual (Chong & Druckman, 2007; Nelson, Oxley, & Clawson, 1997). In other words, at the fundamental level, frames in communication are strategically constructed and deployed, typically by powerful social actors, in an attempt to influence audiences to adopt particular, interpretive frames in thought when making sense of complex issues (Epstein, Nisbet, & Gillespie, 2011).

The tension between frames in communication and frames in thought are at the core of framing research. Gamson and Modigliani (1989) have demonstrated how the gap between these two types of frames affected policy discourse and public understanding around nuclear power. Moreover, consistency between frames in communication and frames in thought contributes to the perceived strength of the communicated frame and may amplify its influence on attitudes and behaviors (Chong & Druckman, 2007; Sniderman & Theriault, 2004). Entman (2004) has proposed a cascading activation model, where issue frame construction and diffusion are linked through mass media portrayals. Framing is especially influential in areas lacking clear definitions, established frames of reference, or where mechanisms of influence are obscured from actors (Cobb, 2005; Gamson & Modigliani, 1989; Hart, 2011; Nisbet, Hart, Myers, & Ellithorpe, 2013; Scheufele & Lewenstein, 2005). Privacy is a prime example of one such area.

Privacy

Privacy, as Solove (2008) has described it, is a "concept in disarray" (p. 1) subject to an ongoing debate in the scholarly community. The debate surrounding privacy intertwines both conceptual and empirical aspects of the phenomenon. Conceptually, Westin (1967), in one of the early, hierarchical models of privacy, viewed it as nested spaces of political, sociocultural, and personal levels of analysis with each layer representing a different set of social structures that both constrain and enable privacy behaviors. Nissenbaum (2010) has further developed this idea by emphasizing the contextual integrity of information flows as underpinning both privacy expectations and privacy behaviors. Solove (2008) himself called for a pluralistic conceptualization of privacy as "a set of protections against a plurality of distinct, but related problems" (p. 171).

The rapid adoption of social media and the growing use of private information as de facto currency for digital (and increasingly physical) services have pushed the conceptualization of privacy to new frontiers with an emphasis on contextual nature of privacy and the plurality of meanings placed in the idea by different actors. Some argue for privacy context collapse, where multiple audiences of an actor collapse into one (Marwick & boyd, 2011; Vitak, 2012), which, in turn, requires a networked model of privacy determined through "constellation of audience dynamics, social norms, and technical functionality that affect the processes of information disclosure, concealment, obscurity, and interpretation within a networked public" (Marwick & boyd, 2014, p. 1063).

Within the conceptual privacy space, one can identify a number of more concrete frames of reference to privacy. Smith et al. (2011) roughly divide approaches to the study of privacy into value- and cognate-based conceptualizations. Central to the value-based conceptualizations is Warren and Brandeis' (1890) framing of privacy as a "right to be left alone". Critics claim, however, that this framing is rooted in the physical notion of privacy, which does not transfer well into the digital realm, where privacy is treated as a commodity or even as a luxury good (Acquisti et al., 2015; Papacharissi, 2010).

The cognate-based approach harbors the control and limited access paradigms of privacy. The control paradigm offers a broad conceptual continuum of what privacy may mean. For some, control over one's information equates with privacy itself. For others, it is viewed as a mediating factor in what constitutes "a dialectic and dynamic boundary regulation process" (Palen & Dourish, 2003). The limited access

paradigm treats privacy just as a state of limited access to a person or her information (Dhillon & Moores, 2001; Westin, 1967) and can also be viewed as a continuum, from absolute to minimal (Smith et al., 2011).

Empirically, the operationalization of privacy, derived from the conceptual foundations described above, has focused primarily on privacy-protecting behaviors and privacy concerns (Smith et al., 2011). The technological architecture of social media specifically set the stage for the plurality of privacy to be examined, especially because of significant opacity in the practices surrounding the collection, processing, and dissemination of user data. The focus on protections allows for measurement of implicit privacy-protecting practices, such as limiting profile visibility, reducing the size of one's network, or changing privacy settings from the default (e.g., Lankton, McKnight, & Tripp, 2017).

Research related to social media and privacy has tended to conceptualize privacy concerns as the driving force behind privacy-protecting behaviors, however, privacy concerns have been conceptualized broadly. Researchers have defined privacy concern to include trust in platform sponsors and risk of personal information loss (Wang, Min, & Han, 2016), as well as privacy orientation and disposition (Baruh, Secinti, & Cemalcilar, 2017). As a dependent variable, at the most fundamental level, past research points toward demographic differences as predictors of privacy concerns. For example, females, older, and more affluent individuals tend to be more concerned with their privacy compared to males, younger, and poorer individuals (Peluchette & Karl, 2009; Smith et al., 2011). As an independent variable, privacy concerns predict, albeit weakly, the use of social media and also the privacy management practices that are employed during use (Baruh et al., 2017). Concerns about privacy prompt less disclosure on social network sites (Dienlin & Metzger, 2016). They also evoke increased use of other privacy strategies, such as regulation of one's network size, self-censorship, and targeted disclosures—all these in addition to the technological features offered by these platforms to regulate privacy (Vitak & Kim, 2014).

Whether it is viewed through behavioral or attitudinal lenses, the lack of nuance in the interpretation of privacy has complicated efforts to examine privacy in the online environment. Gürses and Diaz (2013) referred to this lacuna as a unidimensional treatment of privacy, which limits the ability to unpack mediated relationships occurring in collapsing contexts. Instead, they refer to a distinction, frequently observed in social interactions online, between vertical (also known as institutional) privacy and horizontal (also known as social) privacy. Vertical privacy refers to the privacy relationship between an individual user of social media and institutions, such as schools, government, and corporate platform sponsors (Moorhouse, 2011). Horizontal privacy refers to the privacy relationships among individuals, or users of the social media platforms, building on Raynes-Goldie's (2010) conceptualization of social privacy. While a useful distinction, scholarly attention to the factors contributing to these orientations has been limited (Gürses & Diaz, 2013). For example, scholars have attended to user awareness of either the vertical aspects of privacy (e.g., Acquisti, John, & Loewenstein, 2013) or social privacy (e.g., Bartsch & Dienlin, 2016), while paying limited attention to the inherent tension between these two dimensions. Such gaps, in turn, limit the scope of available research puzzles and policy instruments (Gürses & Diaz, 2013).

The wide range of conceptualizations and operationalizations of privacy, suggests that framing of privacy itself is still very much in flux and subject to discursive contestation. Looking across the conceptual frameworks presented above emphasizes the need to unpack privacy perceptions and privacy discourse at the political, sociocultural, or personal levels, all while acknowledging the context dependency of said discourse, similarly to privacy itself (Epstein et al., 2014).

Framing Privacy

The rich conceptual debate about privacy is a fertile ground for framing research. Somewhat surprisingly, however, studies explicitly tackling the question of the framing of privacy are scarce (Fornaciari, 2014). Existing research focuses primarily on frames in communication of elite actors such as policymakers (Epstein et al., 2014), new media (Fornaciari, 2014), or technology designers (Obar & Oeldorf-Hirsch, 2017), who often view emphasizing privacy as a barrier to adoption. The elite status of these actors stems from their position of relative informational power in terms of the ability to influence the design of privacy policy or technology (Braman, 2009), as well as driving the agenda of public discourse about privacy. The lay public, while varying in terms of both privacy literacy and privacy efficacy, is typically lacking such instrumental abilities. In other words, while some users may appear more "elite" than the others, as individuals they fundamentally have limited or no ability to impact a structural change when it comes to digital privacy.

Fornaciari (2017), attempted to examine frames in communication of non-elite actors by studying privacy framing on Twitter. She identified eight distinct frames ranging from privacy and technology being the most frequent, to trading privacy being the least frequent frame. While we are not aware of research that has explicitly tackled the question of frames in thought of privacy among the non-elite actors, we do know from prior work that actors are mindful about the image they project and demonstrate substantial variance in how they think about their audiences. Studying college students on Facebook, Peluchette and Karl (2009), for example, demonstrated that those who considered family or potential employers among their imagined audience were less likely to post inappropriate content, compared with those who imagined their audience in more generic terms (they also show gender and age differences consistent with privacy literature). Later, Bernstein et al. (2013) and colleagues showed

Table 1. Sample Descriptives.

Characteristic	Sample descriptive	Scale of measurement
Mean age	47.8 (16.7) years	
Gender	53.1% female	
Median income	US$50,000–US$75,000	7 points, ranging from "Under US$25,000" to "Over US$200,000"
Median education	Some college	6 points, ranging from "Less than high school" to "Graduate school"
Race and ethnicity	77% Caucasian 8.9% African American 7.6% Latinx 4.9% Asian	
Most frequently used platform	77.2% Facebook 8.4% Twitter 7.2% Instagram	

that Facebook users severely underestimate the size of their audience when relying on limited markers, such as likes and comments, and on folk theories of how many of their connections might log in. Studying bloggers, Brake (2012) demonstrated that the way they envisioned their audiences (friends or strangers) and the kind of interaction they anticipated (one- or two-way or intrapersonal), would yield different communication styles and levels of disclosure.

Expanding this line of research, our current project asks to delve explicitly into privacy frames in thought of non-elite actors. Hence, we pursue a rather direct research question: how do users of social media frame social media privacy? We explore this question with original, cross-sectional survey data and by using two complementary approaches: topic modeling and semantic network analysis. Given our explicit focus on social media, we expect the frames in thought to reify or challenge more recent conceptual developments such as a networked view of privacy or the vertical versus horizontal distinction. To further understand, and potentially validate, our observations about the adapted frames in thought, we also explore the relationship between the articulated frames and sociodemographic variables, which were previously found to be related to privacy concerns and behaviors (Peluchette & Karl, 2009; Smith et al., 2011).

Methodology

Sample

Participants were recruited in the fall of 2017 using the Qualtrics panel service, and received a small incentive in the form of reward points by the survey platform sponsor. Respondents were matched by quota sampling to parameters of the 2015 US Census Bureau's American Community Survey data on age, income, and gender.[1] This study was approved by the Institutional Review Board at the University of Illinois at Chicago.

Data were collected in a self-administered, web-based survey that included questions on attitudes and behaviors related to social media and privacy. Included in these was the question, "With respect to [participant's most frequently used social media platform], what does privacy mean to you?" Participants were required to supply a definition consisting of a minimum of 135 characters. A list of most frequently used social media platforms was compiled based on the commonly used definitions mentioned in the works of Ellison and boyd (2013) and Obar and Wildman (2015). Common to those definitions are such components as a user profile with various degrees of privacy, the ability of users to connect with peers or form groups, the centrality of user-generated content, and a dynamic feed where users are exposed to said content. By casting this broad net, we wanted to capture the plurality of platforms used by our participants, as opposed to studying an individual platform. The resulting corpus included 608 individually generated definitions of privacy.

The underlying sample is representative of the US population, based on 2010 US Census demographics, on characteristics of age, gender, and income as summarized in Table 1. Mean age was 47.8 years (SD = 16.7, range = 18–90, Mdn = 47.0) and gender was balanced (53.1% female, 46.2% male, 0.7% not reported). Racial/ethnic composition included: African American 8.9% (n = 54); Hispanic/Latino 7.6% (n = 46); Asian 4.9% (n = 30); Caucasian 77.0% (n = 468); multi-ethnic/other/undisclosed 1.2% (n = 7). Participants in the study were actively engaged with social media, with 90.8% reported having two or more social media profiles and 81.1% reported accessing their favored social media site at least once/day. Facebook was the most frequently used social media platform (n = 468, 77.2%), followed by Twitter (n = 51, 8.4%) and Instagram (n = 46, 7.6%).

Method

The analysis that follows is based on three main methodologies. First, we employ topic modeling to distill primary themes in the definitions. Topic modeling is a text-mining approach that uses statistical probabilities for discovering topics or themes in a collection of documents. It is based on

the notion that documents are collections of topics that reflect a thematic structure which can be inferred by examining the probability distribution of words appearing together (Steyvers & Griffiths, 2007). In each topic, different sets of terms have higher probabilities; topics can be visualized by listing and interpreting these terms (Blei, 2012b). It is a useful approach for analyzing unstructured texts to discover not only themes, but also how those themes may be connected to one another (Blei, 2012a). In contrast to traditional content analysis, topic modeling utilizes computer algorithms to identify patterns of word co-occurrence, and thus is useful for analyzing large datasets. In addition, because it can be used without a priori coding structures, topic modeling lends itself to inductive research.

ConText (Diesner, 2014) is an automated topic modeling tool for analyzing texts and networks that can be used to analyze a large volume of texts. ConText leverages the "MAchine Learning for LanguagE Toolkit" (MALLET, McCallum, 2002) to perform topic modeling, which is based on the Latent Dirichlet Allocation (LDA) model (Blei, Ng, & Jordan, 2003). LDA assumes documents are generated by drawing on fixed topic vocabularies that are composed of words with high probabilities; it then reverses this generative process to uncover the latent topics within the texts using probabilistic modeling (Blei, 2012a).

Prior studies on topic model evaluation have emphasized the importance of including real world evaluation mechanisms for validating topic models (Chang, Gerrish, Wang, & Blei, 2009), in addition to examining measures of topic coherence. In an effort to accommodate this form of support, as a second step we use semantic network analysis, a form of collocation analysis, to validate the topic modeling results (see Borge-Holthoefer & Arenas, 2010 for a review). These methods have been demonstrated to be robust for large and small corpora as well as corpora of differing quality (Bullinaria & Levy, 2007, 2012). To conduct this analysis, we employed the Python Natural Language Toolkit (NLTK; Bird, Klein, & Loper, 2009) to generate word pair co-occurrences and frequencies; then, using these as an undirected edge list, we performed clustering in Gephi (Bastian, Heymann, & Jacomy, 2009).

Finally, we employed binary logistic regression in an attempt to further unpack the dynamics of privacy framing of non-elites. Here, we examined whether traditional measures of socioeconomic status (e.g., income, educational attainment, and race) might be useful in predicting privacy orientation. In this last analysis, we specifically address the research question of, "Which sociodemographic factors contribute to a horizontal privacy orientation?"

Analysis

Data Cleaning. Initial cleaning of the data, preprocessing, was carried out in ConText. Our process included the removal of stopwords, or articles, prepositions, conjunctions, and transitive verbs that do not contribute to the meaning of the text (e.g., if, and, that, a, an, the, to, is, was, were). We applied a codebook to consolidate n-gram terms such as "social media" and "phone number." In addition, we removed the five most common words in the corpus; these included the words: information, privacy, means, people, and Facebook. Three of these appeared in the definition prompt for survey respondents; the remaining words, "people" and "information," appeared so frequently that they offered limited utility in differentiating among topics. The most relevant remaining term frequencies are summarized in Table 2. Finally, we used the native ConText algorithm to perform stemming and adjust tense and different forms of the same word into a unified morpheme.

Topic Modeling. To illustrate the themes found within the privacy definitions, we began by specifying 10 topics at 10 words per topic. We evaluated topic fit through a close reading of individual definitions and by assessing coherence of each topic solution. For topic coherence, we relied on a measure of perplexity (Blei et al., 2003)—the exponent of the log likelihood per token—with smaller values as an indication of better prediction. To ensure topic quality and coherence, three researchers performed a close interpretive reading of definitions with a Fit to Topic score of at least .90. The Fit to Topic score is the probability of an individual definition being classified in the topic by the model, and the close reading allowed for qualifying the topics with an emphasis on the vertical–horizontal distinction. We successively reduced the number of topics to improve the interpretability, while monitoring topic coherence and topic quality for each solution. Our final accepted model included two topics, with a perplexity score at an acceptable level of exp(–6.32623).

Semantic Network Analysis. To validate our observations derived from topic modeling, we conducted a collocation analysis of privacy definitions provided by survey respondents. We used NLTK (Bird, Klein, & Loper, 2009) to generate a list of co-occurring word pairs from the definitions, using a word-distance window of seven words. These were used to map an undirected network using Gephi (Bastian et al., 2009). To improve the clarity of the graph, the density was reduced by removing edges below a minimum weight. The edge weight is the frequency of the collocation in the corpus. Setting the minimum edge weight to 3 excludes any collocations that only appear once or twice in the corpus and improves the readability of the resulting graph. The removal of the low weight edges resulted in some nodes becoming disconnected, and these were also removed. We used Gephi's ForceAtlas2 algorithm to layout the graph followed by overlap reduction and label adjustment to improve readability. With this layout algorithm frequently co-occurring words appear closer together and the shape of the network reflects the association of the words within the corpus. The node and label sizes were set to represent the weighted degree of each

Table 2. Forty Most Relevant Term Frequencies.

Term	Freq.	TF × IDF	% of texts	Term	Freq.	TF × IDF	% of texts
post	197	0.023481	0.241776	business	37	0.009561	0.046053
share	194	0.023123	0.241776	put	40	0.009492	0.059211
personal	210	0.020882	0.305921	make	37	0.008958	0.055921
private	147	0.019822	0.200658	setting	36	0.008806	0.054276
friend	125	0.019154	0.161184	protect	36	0.008628	0.057566
thing	82	0.015081	0.111842	sell	31	0.008011	0.046053
profile	64	0.01311	0.087171	social	30	0.007844	0.044408
access	58	0.012466	0.077303	choose	29	0.00777	0.041118
public	57	0.011955	0.082237	view	29	0.007674	0.042763
give	58	0.011881	0.087171	set	28	0.007598	0.039474
safe	50	0.011218	0.069079	dont	29	0.007582	0.044408
control	52	0.010994	0.080592	find	29	0.007582	0.044408
info	45	0.010679	0.059211	person	29	0.007494	0.046053
life	46	0.010509	0.065789	address	27	0.007326	0.039474
family	46	0.010141	0.072368	hack	27	0.007326	0.039474
picture	44	0.010052	0.065789	site	27	0.007326	0.039474
account	44	0.010052	0.065789	ability	26	0.007245	0.036184
important	44	0.009872	0.069079	permission	28	0.007235	0.046053

TF: term frequency; IDF: inverse document frequency.

node; degree is the number of edges attached to a node, and weighted degree includes the frequency of the edge in the corpus. Finally, Gephi's native community detection algorithm (Blondel, Guillaume, Lambiotte, & Lefebvre, 2008) was applied with a resolution of 1.2 to identify any communities within the network. The resolution can vary around 1.0 with lower values increasing the number of communities detected and higher values identifying fewer, larger communities. Using a resolution of 1.2 resulted in five communities being identified. The color of each node was set to indicate the detected community (modularity class).

Logistic Regression. To better understand the dimensions of privacy as they emerged from topic analysis, we also explored the relationship between sociodemographic variables and the adapted frames in thought. Using the Fit to Topic measures assigned in the accepted solution from the topic modeling step, definitions with a probability of .60 or higher for belonging to one of the topics were determined to have the privacy orientation of that topic. We then examined whether the sociodemographic characteristics of the definition's author were significant factors in predicting privacy orientation by using binary logistic regression.

Results

Topic Modeling

Results of the topic modeling are summarized in Table 3, with the words in each topic presented in order of their weight within that topic. Although the final solution includes only two topics, we note that topic coherence measures did not differ markedly between the 10-topic solution, perplexity = exp(–6.73921), and the two-topic solution, perplexity = exp(–6.32623). In contrast, the consistency of the definitions with the highest probability of being classified to the topic, as determined by consensus after close reading, was highest in the two-topic solution.

Topics are ordered by weight within the document corpus in Table 3. The first, most prominent, topic reflects horizontal perspective of privacy. The topic weight of 0.847 indicates the prevalence of this topic in the corpus and is the probability of each document in the corpus belonging to Topic 1. This topic is characterized by the highest probability terms for it; "post," "share," and "access,"; "personal," "private," and "public"; as well as "friend" and "profile." These terms suggest activities, elements, and actors typical of peer to peer communication on social media platforms. When viewed through the lens of Nissenbaum's (2010) contextual integrity theory, these elements reflect actors, types of information, mechanisms of data transmission, and social norms, all of which are part of the systemic context of horizontal social media use. Close reading of definitions with high (>.90) probability of being included in this topic also surfaces frequent references to friends, family, and even strangers as audiences for social media content. The example definitions for inclusion in this topic listed in Table 3 are top documents for the topic (i.e., documents with the highest probability of being assigned to the topic); some additional, typical comments include, ". . . privacy means that others, who are not already a part of my inner circle of friends and family, can not post or interact on my account/home page" and ". . . not everyone can see what I post and to ensure others don't post to my wall without permission."

Table 3. Topic Identification Using ConText, Along With Examples of Definitions With High Fit to Topic.

Topic	Topic weight	Members	Examples of complete definitions that belong to the topic	Fit to topic
1	0.846714	post–share–personal–private–friend–thing–profile–access–give–public	"Privacy means the ability to control who sees my information. I have several social circles in my life, e.g. my very Christian family and my pagan friends, and I need to be able to interact with one without it being immediatly posted to the other. I also require my private life to remain private. However, as far as non-personal metadata, or information without my name attatched, I would not be taking this survey if I was stingy with that data!" "In my opinion, the word privacy means no one can see anything of mine and that includes, but is not limited to; statuses, pictures, age, birthday, name, cover photo, what school I went to, where I work, where I have worked, my friends list, the groups I'm in, etc. This is only if you choose what you want to be private, but in my opinion, again, you should have the opportunity to pick everything you want to be private."	993661 993397
2	0.158515	personal–protect–secure–security–government–invade–web–include–free–business	"Not having my information put up for display. So that random people can come across details I'd rather have them not know. I need to be secure when I use the website, otherwise my identity online can be exposed and ruin future prospects due to my online history." "Privacy online means a personal outlet where I can communicate through a free medium like Facebook and securely express my views to people I care about without fear of reprisal."	947098 934894

The second topic, has a significantly lower weight in the corpus at 0.1585, and reflects concepts related to more vertical forms of privacy. This topic is again characterized not only by the word "personal" (the most dominant word remaining in the cleaned corpus) but also by words such as "protect," "secure," and "security," which indicate a substantively different set of contextual parameters, compared to Topic 1. The presence of terms such as "government" and "invade" helps to qualify it as a more vertical privacy systemic context and indicates awareness of a potential power imbalance in the privacy calculus. Close reading of the definitions with high probability of being assigned to this topic suggests frequent references to hacking, spyware, and identity threats. Some additional typical excerpts include statements such as, "Having to have my location on means no privacy" and "Privacy on Facebook means not intercepting or interfering with messages, profile, or altering them in any way."

Semantic Network Analysis

Semantic network analysis offers a complementary way to further examine the dimensionality privacy and provide additional clarity to the topics in the corpus of definitions. The initial network consisted of 532 nodes (words) and 2,095 edges (linking words that appear together within a seven-word window, weighted by the number of times the words appeared together in the corpus). This network was reduced to 205 nodes and 850 edges after filtering edges below a frequency of 3 in the corpus. We then apply Gephi's native community detection algorithm (Blondel et al., 2008) with a resolution of 1.2, which detects five communities as shown in Table 4 and Figure 1.

The first two communities in Table 4 are each much larger than the remaining three, and the weighted degree and degree of the top words for these large communities are much higher. This shows that the central terms in these two communities are much more important in the corpus. The two smallest communities only contain four and three words and both are strongly linked to the largest community. They are visible to the top right of Figure 1 with blue and gray nodes linking into the large community shown in green in Figure 1. The former captures a series of more mundane intrusions on privacy (observed, disturbed) and the later more criminal ones (steal, theft). The other small community is shown in red in Figure 1 to the right and bottom of the network and captures types of personal information (e.g., email, phone) and the extent of unwanted exposure (entire world, social media). It is again mainly linked to the large green community but does also have some links to the next largest community shown in purple in Figure 1. The two large communities (shown in green and purple in Figure 1) were extracted for further detailed analysis. The network layout step was repeated on each of these, but the node and label sizes were retained as indicating the weighted degree of the overall network, not the individual community.

The community marked in purple in Figure 1, which can be seen in more detail in Figure 2, depicts word pairs indicative of horizontal privacy that are similar to those in the first topic of the topic modeling analysis. The top terms for Community 2 by weighted degree are given in Table 4 and these appear to be types of information, such as "picture," "photos," and "profile," along with those whom such information might be willingly or unwillingly shared with, such as "friends," "family," and even "strangers." Note that "post" is both a central node in the network and one of the most

Table 4. Communities Detected by Gephi (Sorted by Size).

Community	Number of nodes	Top 10 terms (by weighted degree)			Color
		Term	Weighted degree	Degree	
1	121	personal	1,303	113	Green
		share	1,161	101	
		private	851	77	
		thing	431	49	
		public	289	30	
		access	279	33	
		safe	264	30	
		info	244	25	
		give	224	26	
		account	217	25	
2	63	post	1,225	100	Purple
		friend	749	74	
		profile	334	33	
		control	250	22	
		family	219	23	
		picture	210	21	
		page	207	26	
		setting	169	24	
		view	153	17	
		dont	130	14	
3	14	social	126	18	Red
		address	113	15	
		number	107	11	
		phone	89	9	
		media	82	12	
		email	66	8	
		world	61	11	
		sharing	31	4	
		talk	15	2	
		network	15	1	
4	4	free	23	3	Blue
		disturbed	11	4	
		condition	10	3	
		observed	9	3	
5	3	steal	34	2	Gray
		identity	51	4	
		theft	5	1	

frequent terms in the text, so nodes connected to it have higher relevance for interpreting this community. Linked to "post" are other prominent words like "control," "setting," and "view" combined with lower weight terms like "block" and "safety," which express the desire for control over who can see the shared information.

Conversely, the community marked in green in Figure 1, comprises words indicative of vertical privacy awareness (Figure 3). The three most prominent terms by weighted degree—"private," "personal," and "share" (see Community 1 in Table 4 and Figure 3)—are also three of the most frequent words in the corpus (see Table 2), therefore, we look to the next largest nodes in this component for interpretation. References to platform sponsors, such as Twitter and Instagram (recall that the term "Facebook" was eliminated from the analysis due to its overwhelming presence in the corpus), along with terms such as "government," "sell," and "company" indicate a view of privacy that is oriented toward institutional actors. In addition, terms such as "secure," "safe," and "hack" indicate privacy that is influenced by structural or systemic elements. Of note in this component are references to norms that govern information flow in this context, including terms such as "consent," "authorize," and "safeguard."

Positioned between the communities are cognate-based approaches to privacy. Table 5 lists nodes with the highest positive E/I index values in the two largest clusters. The E/I index demonstrates interconnectivity between clusters (Krackhardt & Stern, 1988). Though many of the listed terms are somewhat neutral, terms such as "control," "setting," and "block" suggest that, with respect to social media, users conceptualize privacy as a boundary-control process whereby limits on access to information is prioritized. That is contrary to the value-based approaches, which view privacy as a right.

Logistic Regression

Finally, to better understand the underlying factors that contribute to the framing of privacy, we examined the sociodemographic characteristics of the survey respondents as they relate individually to the privacy definitions. Such inquiry is particularly intriguing given the dominance of the horizontal orientation toward privacy in our sample. We performed a binomial logistic regression to determine the effects of age (three categories: 18–44 years, 45–64 years, and 65+ years), income (seven levels: under US$25,000 to over US$200,000), gender, educational attainment (five levels, high school or less to graduate degree), and race on the likelihood that participants would have a horizontal orientation to privacy. These factors have been found important in predicting general privacy concerns and privacy-protecting behaviors (Smith et al., 2011), which makes them the first line of inquiry in unpacking the dimensions identified in this study. The model explained 6.7% (Nagelkerke R^2) of the variance and correctly classified 90.0% of the cases. Table 6 summarizes the results of this analysis.

Of the five predictor variables, only gender was statistically significant, demonstrating that females are three times (OR = 2.985, $p < .001$) more likely to have a horizontal orientation to privacy than males. While not statistically significant, the odds ratios for income also indicated a greater likelihood for a horizontal privacy orientation, that is those at higher income levels are more likely to frame privacy in horizontal terms. Caucasians, those in older age groups, and those with higher education were less likely to have a horizontal orientation to privacy.

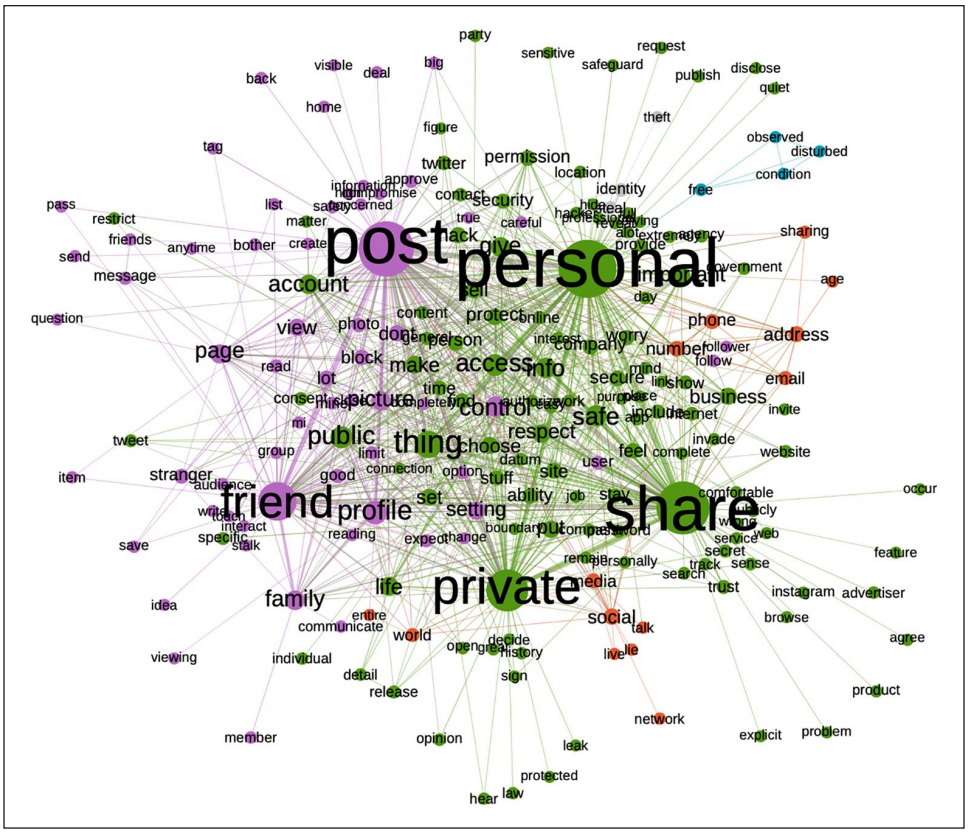

Figure 1. Collocations network.

Discussion and Conclusions

Before we engage with discussing the interpretation of our results, it is important to keep in mind that there is a number of limitations inherent to the current study. First, the scope of user-supplied definitions analyzed here was relatively limited and the level of detail among corpus definitions varied substantively. As data collection for these definitions also included user reports of privacy activities, future work might examine these definitions in the context of reported behaviors. Second, the automated methods we used draw on some foundational decisions, such as examining definitions using a seven-word window, and these may limit the interpretations. Third, the relatively small size of our sample, while sufficient for preliminary analysis presented here, falls short in affording robust examination of factors explaining a particular frame in thought. Fourth, it is important to remember that our observations are limited to the unique setting of the US-based population of social media users. Albeit large, active, and diverse, this group is not representative of the larger population of social media users worldwide. Similar criticism could be voiced toward our use of standard categories of gender and racial identities, limiting our ability to draw conclusions about nonconformist groups, even though this was not the aim of this study. Finally, our sample was skewed toward Facebook users, even though there were no significant differences in variables of interest across the users of different social media platforms in the sample. A larger sample, could potentially tease out perceptual differences between the users of different platforms. Future work may benefit from both a larger and a more diverse sample, and a refined prompt. Particularly, future research should explore data collected in a variety of cultural contexts outside of the United States and across a wider range of identities. With that, the current study offers an important stepping stone in efforts to both conceptualize and capture privacy in the increasingly mediated world.

Based on the topic modeling analysis and the resultant topic weights, our results suggest that users' conceptualizations emphasize dimensions of horizontal privacy (i.e., privacy between users of social media platforms), over conceptualizations of privacy that emphasize freedom from oversight, or vertical privacy. Term frequencies of three largest clusters in the semantic network analysis reinforce this interpretation. Such prioritization of the social aligns with ideas of networked privacy, where privacy is enforced by social norms, and may indicate that user framing of privacy is perhaps more focused on social aspects than what has been assumed by the research community, privacy activists, and especially policy makers. Vertical privacy, while not lost on users, seems to have lower levels of relevance across user definitions in our sample. These observations suggest three

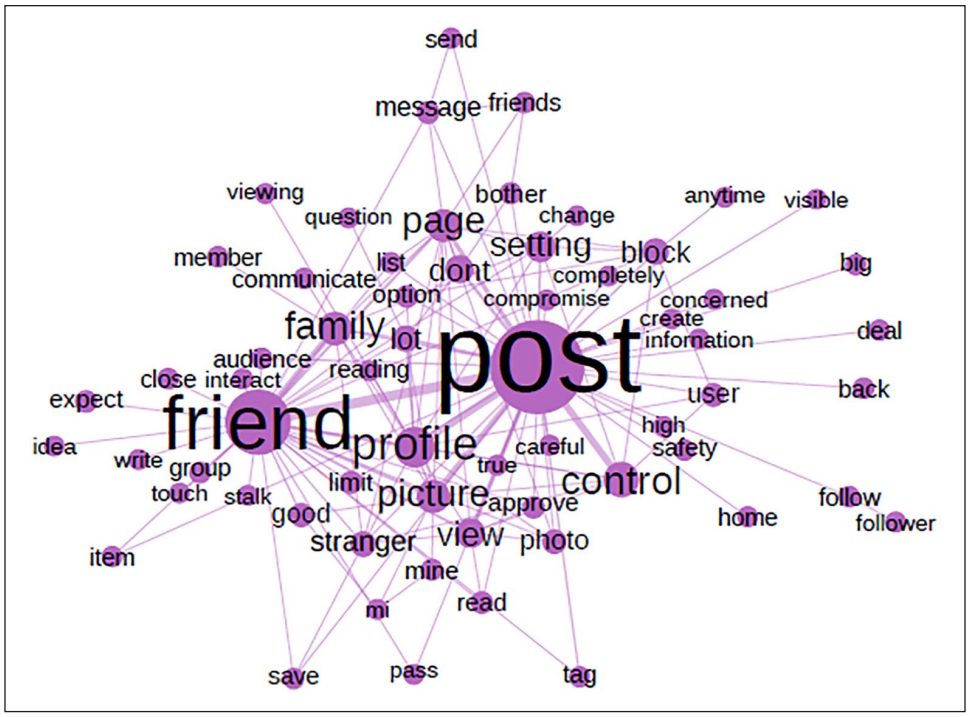

Figure 2. Horizontal privacy–Community 2.

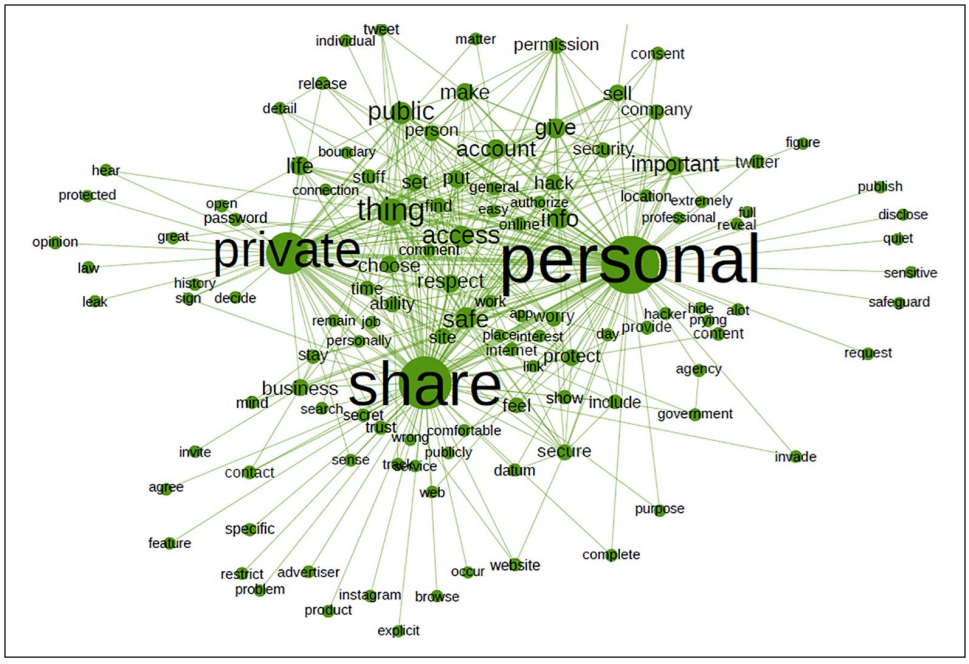

Figure 3. Vertical privacy—Community 1.

main potential implications to the study, design, regulation, and communication regarding social media privacy today.

First, our findings may suggest a gap between elites and non-elites when it comes to perceptions of privacy. While mostly consistent with prior work on privacy framing by non-elites (Fornaciari, 2017), these results demonstrate that users view their own social networks as their primary audiences, as opposed to platform sponsors or other institutions. This gap may have important implications. On one hand, it diminishes the perceived responsibility of platform providers

Table 5. Nodes With Positive E/I Index Values.

Term	Weighted degree	Community	E/I index
specific	31	2	0.217391
content	43	2	0.047619
close	21	1	0.411765
setting	169	1	0.308271
expect	38	1	0.250000
dont	130	1	0.238095
user	56	1	0.200000
post	1,225	1	0.198895
big	16	1	0.166667
control	250	1	0.164948
careful	7	1	0.142857
block	102	1	0.125000

Table 6. Factors Contributing to Horizontal Privacy Orientation.

Variable	OR	95% CI		*p*
		Lower	Upper	
Income	1.203	0.989	1.462	.064
Education	.707	0.432	1.155	.166
Female	2.985	1.617	5.510	.000
Caucasian	0.663	0.310	1.420	.290
Age	0.906	0.619	1.326	.610
Constant	7.346			.000

CI: confidence interval; OR: odds ratio.

for protecting their users' privacy from their own abuse as well as that of third parties, as illustrated in the recent Cambridge Analytica and Facebook scandal. Moreover, the focus on horizontal privacy by the users tends to steer privacy discourse to the dominance of technical solutions, such as privacy by design, whereby platform providers offer additional controls to regulate horizontal, but not necessarily vertical, privacy.

On the other hand, the existence of such a gap may render many policy or civic interventions ineffective, as both activists and regulators tend to focus on behaviors that fall outside the scope of those that constituents view as relevant. The early criticism of the recent implementation of the General Data Protection Regulation, for example, suggests that users do not exercise the levers offered by regulators to protect privacy, potentially because this is not the privacy they care about. Understanding, and then bridging, the elite/non-elite gap in privacy orientation may be the first necessary step for effectively enacting privacy regulation.

Although these implications are somewhat speculative, they are derived from the resonance of our findings with existing literature, which suggests that perceived or imagined audiences matter to how people communicate on social media (Bernstein et al., 2013; Brake, 2012; Peluchette & Karl, 2009). Thus, having institutional players largely absent from the users' imagined audiences is likely to impact how they perceive their privacy threats and potential solutions.

Second, from a more macroscopic perspective, our observation of the elite/non-elite gap in privacy perceptions may highlight mechanisms of power imbalance. The lack of attention to vertical privacy incursions may further reify existing power disparities, along the lines of what Braman (2009) describes as a "panspectron" society, a condition where information about an individual is collected continuously and where state (and increasingly large private players) know disproportionally more about an individual than that individual knows about the state and the large institutional players. If privacy is indeed viewed in such different terms by those who have the ability to abuse or, conversely, govern the privacy of others (either through regulation or design), then there is little incentive to adopt more ethical privacy practices. In light of this imbalance, the role of researchers can be critically important, especially with respect to unpacking the specific structures of power fueled by those distinct perceptions. Mechanisms to address this potential imbalance also include for designers to make data collection more explicit and transparent, and for policymakers to emphasize the importance of privacy education for users of social media.

Our study also suggests that disparities in privacy perception exist within the non-elites, along the lines of gender and affluence. On one hand, we observe in the current sample that the more affluent segments of the non-elites can "afford" focusing on horizontal privacy at expense of the vertical. This dynamic may reflect the structural power relationship whereby the poor, at least in the United States, tends to be subject to greater surveillance by the institutional players (Gilman, Madden, Levy, & Marwick, 2017). On the other hand, females are nearly three times more likely to have a horizontal orientation toward privacy. This relationship may reflect a heightened awareness of the voyeuristic potential of social media technologies that can be used as everyday surveillance technologies directed toward gendered bodies (Monahan, 2009). Both of these findings emphasize structural inequalities by which power is exercised differently toward different groups, but where the less powerful cannot afford to disengage from the power exercised toward them by other players. In the case of those less affluent, this would reflect governments, corporations, and platform operators. In the case of women, it reflects the power that others may hold over their bodies, in both a physical and informational forms. In both cases, privacy is viewed as a luxury commodity (Papacharissi, 2010), the luxury aspect lies in the ability to pay relatively more attention to one aspect of privacy, and less attention to the other.

Third, taken together, our observations further highlight the multidimensionality of privacy. This is perhaps the most important insight we would like to invite our readers to engage with. We argue that scholars, designers, and regulators need to pay more attention to the ways in which privacy is actually perceived and enacted by those on the receiving

end of both technology and regulation. The challenge here is both conceptual and empirical. On one hand, more qualitative, ground-up research is needed to better understand the nuances of how users perceive and enact privacy not only across various contexts of use, but also across distinct regulatory and cultural environments. On the other hand, those choosing to continue the trajectory of the current study may offer substantive contribution by unpacking privacy's multiple dimensions, and engaging with qualitative (not just orientational) differences among frames in thought about privacy. In other words, while the current work offers a stepping stone toward expanding both conceptual thinking and the empirical toolkit in privacy research, even our data lend itself to additional interpretations. Of particular interest is the perspective that equates privacy with data security—a view that was evident across our sample of definitions—as well as questions on how this narrative influences privacy orientation.

Conceptualizing and measuring privacy is becoming a fundamental task in unpacking social structures and power imbalances in the information society. Although privacy seems to grow in it importance in the public discourse, it remains a vague and intangible concept for most. Privacy definitions analyzed in the current study, in many ways, offer an illustration of that opaqueness. Given the fluidity of the idea of privacy and it is continuous change over time, it is increasingly important to think about systematic ways of not only tracking privacy perceptions, attitudes, and behaviors over time, but also to track these across technological and cultural contexts; both are fast moving targets in themselves. Yet, it is through attempts to bring clarity to such a muddy subject that we can both learn about the intangible power structures of our times and warn about the emerging power structures of the future.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Kelly Quinn https://orcid.org/0000-0001-9922-823X
Brenda Moon https://orcid.org/0000-0003-2571-0650

Note

1. US Census Bureau (2016).

References

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

Acquisti, A., John, L. K., & Loewenstein, G. (2013). What is privacy worth? *The Journal of Legal Studies*, *42*, 249–274. doi:10.1086/671754

Bartsch, M., & Dienlin, T. (2016). Control your Facebook: An analysis of online privacy literacy. *Computers in Human Behavior*, *56*, 147–154. doi:10.1016/j.chb.2015.11.022

Baruh, L., Secinti, E., & Cemalcilar, Z. (2017). Online privacy concerns and privacy management: A meta-analytical review. *Journal of Communication*, *67*, 26–53. doi:10.1111/jcom.12276

Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: An open source software for exploring and manipulating networks. In *Third International AAAI Conference on Weblogs and Social Media* (pp. 361–362). doi:10.1136/qshc.2004.010033

Bernstein, M. S., Bakshy, E., Burke, M., Karrer, B., & Park, M. (2013). Quantifying the invisible audience in social networks. In *SIGCHI Conference on Human Factors in Computing Systems* (pp. 21–30). Paris, France. doi:10.1145/2470654.2470658

BeVier, L. R. (1995). Information about individuals in the hands of government: Some reflections on mechanisms for privacy protection. *William & Mary Bill of Rights Journal*, *4*, 455–506. doi:10.1525/sp.2007.54.1.23

Bird, S., Klein, E., & Loper, E. (2009). NLTK book. Retrieved from https://www.nltk.org/book/

Blei, D. M. (2012a). Probabilistic topic models. *Communications of the ACM*, *55*(4), 77–84. doi:10.1145/2133806.2133826

Blei, D. M. (2012b). Topic modeling and digital humanities. *Journal of Digital Humanities*, *2*(1). Retrieved from http://journalofdigitalhumanities.org/2-1/topic-modeling-and-digital-humanities-by-david-m-blei/

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, *3*, 993–1022. doi:10.1162/jmlr.2003.3.4-5.993

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, *2008*(10), P10008. doi:10.1088/1742-5468/2008/10/P10008

Borge-Holthoefer, J., & Arenas, A. (2010). Semantic networks: Structure and dynamics. *Entropy*, *12*, 1264–1302. doi:10.3390/e12051264

Brake, D. R. (2012). Who do they think they're talking to? Framings of the audience by social media users. *International Journal of Communication*, *6*, 1056–1076.

Braman, S. (2009). *Change of state: Information, policy, and power*. Cambridge: The MIT Press.

Bullinaria, J. A., & Levy, J. P. (2007). Extracting semantic representations from word co-occurrence statistics: A computational study. *Behavior Research Methods*, *39*, 510–526. doi:10.3758/BF03193020

Bullinaria, J. A., & Levy, J. P. (2012). Extracting semantic representations from word co-occurrence statistics: Stop-lists, stemming, and SVD. *Behavior Research Methods*, *44*, 890–907. doi:10.3758/s13428-011-0183-8

Chang, J., Gerrish, S., Wang, C., & Blei, D. M. (2009). Reading tea leaves: How humans interpret topic models. *Advances in Neural Information Processing Systems*, *22*, 288–296. doi:10.1.1.100.1089

Chong, D., & Druckman, J. N. (2007). Framing theory. *Annual Review of Political Science*, *10*, 103–126. doi:10.1146/annurev.polisci.10.072805.103054

Cobb, M. D. (2005). Framing effects on public opinion about nanotechnology. *Science Communication*, *27*, 221–239. doi:10.1177/1075547005281473

Dhillon, G. S., & Moores, T. T. (2001). Internet privacy. *Information Resources Management Journal*, *14*(4), 33–37. doi:10.4018/irmj.2001100104

Dienlin, T., & Metzger, M. J. (2016). An extended privacy calculus model for SNSs: Analyzing self-disclosure and self-withdrawal in a representative U.S. sample. *Journal of Computer-Mediated Communication*, *21*, 368–383. doi:10.1111/jcc4.12163

Diesner, J. (2014). *ConText: Software for the integrated analysis of text data* and network data. Paper presented at the Social and Semantic Networks in Communication Research. Preconference at Conference of International Communication Association (ICA), Seattle, WA.

Ellison, N. B., & boyd d. (2013). Sociality through social network sites. In W. H. Dutton (Ed.), *The Oxford handbook of internet studies* (pp. 151–172). Oxford, UK: Oxford University Press.

Entman, R. M. (2004). *Projections of power: Framing news, public opinion, and US foreign policy*. Chicago: University of Chicago Press.

Epstein, D., Nisbet, E. C., & Gillespie, T. (2011). Who's responsible for the digital divide? Public perceptions and policy implications. *Information Society*, *27*, 92–104. doi:10.1080/01972243.2011.548695

Epstein, D., Roth, M. C., & Baumer, E. P. (2014). It's the definition, stupid! Framing of online privacy in the internet governance forum debates. *Journal of Information Policy*, *4*, 144–172. doi:10.5325/jinfopoli.4.2014.0144

Fornaciari, F. (2014). Mapping the territories of privacy: Textual analysis of privacy frames in American mainstream news. In *Annual Hawaii international conference on system sciences* (pp. 1823–1832). Waikoloa, HI. doi:10.1109/HICSS.2014.230

Fornaciari, F. (2017). iTweet about #privacy: Mapping privacy frames in Twitter conversation. In *ALLDATA: The Third international conference on big data, small data, linked data and open data* (pp. 70–73). Venice, Italy: IARIA.

Gamson, W. A., & Modigliani, A. (1989). Media discourse and public opinion on nuclear power: A constructionist approach. *American Journal of Sociology*, *95*, 1–37. doi:10.1086/229213

Gilman, M., Madden, M., Levy, K., & Marwick, A. (2017). Privacy, poverty and big data: A matrix of vulnerabilities for poor Americans. *Washington University Law Review*, *95*, 53–125. Retrieved from http://openscholarship.wustl.edu/law_lawreview/vol95/iss1/6/

Goffman, E. (1974). *Frame analysis*. Cambridge, MA: Harvard University Press.

Gürses, S., & Diaz, C. (2013). Two tales of privacy in online social networks. *IEEE Security & Privacy*, *11*, 29–37.

Hart, P. S. (2011). One or many? The influence of episodic and thematic climate change frames on policy preferences and individual behavior change. *Science Communication*, *33*, 28–51. doi:10.1177/1075547010366400

Krackhardt, D., & Stern, R. N. (1988). Informal networks and organizational crises: An experimental simulation. *Social Psychology Quarterly*, *51*, 123–140. doi:10.2307/2786835

Lankton, N. K., McKnight, D. H., & Tripp, J. F. (2017). Facebook privacy management strategies: A cluster analysis of user privacy behaviors. *Computers in Human Behavior*, *76*, 149–163. doi:10.1016/j.chb.2017.07.015

Marwick, A. E., & boyd d. (2011). I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media & Society*, *13*, 114–133. doi:10.1177/1461444810365313

Marwick, A. E., & boyd d. (2014). Networked privacy: How teenagers negotiate context in social media. *New Media & Society*, *16*, 1051–1067. doi:10.1177/1461444814543995

McCallum, A. K. (2002). MALLETT: A machine learning for language toolkit. Retrieved from https://www.bibsonomy.org/bibtex/26dbb7b45a3a53997359a5e3c2677dc52/zeno

Monahan, T. (2009). Dreams of control at a distance: Gender, surveillance, and social control. *Cultural Studies ↔ Critical Methodologies*, *9*, 286–305. doi:10.1177/1532708608321481

Moorhouse, F. (2011). Beyond stigma: Musings on the sadness of privacy. *Griffith REVIEW*, *33*, 92–107.

Nelson, T. E., Oxley, Z. M., & Clawson, R. A. (1997). Toward a psychology of framing effects. *Political Behavior*, *19*, 221–246. doi:10.1023/A:1024834831093

Nisbet, E. C., Hart, P. S., Myers, T., & Ellithorpe, M. (2013). Attitude change in competitive framing environments? Open-/closed-mindedness, framing effects, and climate change. *Journal of Communication*, *63*, 766–785. doi:10.1111/jcom.12040

Nissenbaum, H. (2010). *Privacy in context: Technology, policy, and the integrity of social life*. Stanford, CA: Stanford University Press.

Obar, J. A., & Oeldorf-Hirsch, A. (2017). Clickwrap impact: Quick-join options and ignoring privacy and terms of service policies of social networking services. In *International conference on social media & society* (p. Article 50). Toronto, Ontario, Canada. doi:10.1145/3097286.3097336

Obar, J. A., & Wildman, S. (2015). Social media definition and the governance challenge: An introduction to the special issue. *Telecommunications Policy*, *39*, 745–750. doi:10.1016/j.telpol.2015.07.014

Palen, L., & Dourish, P. (2003). Unpacking "privacy" for a networked world. In *Proceedings of the conference on human factors in computing systems - CHI '03* (pp. 129–136). doi:10.1145/642633.642635

Papacharissi, Z. (2010). Privacy as a luxury commodity. *First Monday*, *15*(8). Retrieved from http://firstmonday.org/ojs/index.php/fm/article/view/3075

Peluchette, J., & Karl, K. (2009). Examining students' intended image on Facebook: "What were they thinking?!" *Journal of Education for Business*, *85*, 30–37. doi:10.1080/08832320903217606

Raynes-Goldie, K. (2010). Aliases, creeping, and wall cleaning: Understanding privacy in the age of Facebook. *First Monday*, *15*(1), 1–8. Available from http://firstmonday.org/

Scheufele, D. (1999). Framing as a theory of media effects. *Journal of Communication*, *49*, 103–122. doi:10.1111/j.1460-2466.1999.tb02784.x

Scheufele, D. A., & Lewenstein, B. V. (2005). The public and nanotechnology: How citizens make sense of emerging technologies. *Journal of Nanoparticleresearch*, *7*, 659–667. doi:10.1007/s11051-005-7526-2

Smith, H. J., Dinev, T., & Xu, H. (2011). Information privacy research: An interdisciplinary review. *MIS Quarterly*, *35*, 989–1016.

Sniderman, P. M., & Theriault, S. M. (2004). The structure of political argument and the logic of issue framing. In W. E. Saris &

P. M. Sniderman (Eds.), *Studies in public opinion* (pp. 133–165). Princeton, NJ: Princeton University Press.
Solove, D. J. (2008). *Understanding privacy*. Cambridge, MA: Harvard University Press.
Steyvers, M., & Griffiths, T. (2007). Probabilistic topic models. In T. Landauer, D. McNamara, S. Dennis, & W. Kintsch (Eds.), *Latent semantic analysis: A road to meaning* (pp. 427–448). Hillsdale, NJ: Laurence Erlbaum.
U.S. Census Bureau (2016). *One year public use microdata samples, 2011-2015 American Community Survey 5-year estimates*. Retrieved from https://factfinder.census.gov/faces/nav/jsf/pages/searchresults.xhtml?refresh=t
Vitak, J. (2012). The impact of context collapse and privacy on social network site disclosures. *Journal of Broadcasting & Electronic Media*, *56*, 451–470. doi:10.1080/08838151.2012.732140
Vitak, J., & Kim, J. (2014). You can't block people offline. In *Proceedings of CSCW '14* (pp. 461–474). New York, NY: ACM Press. doi:10.1145/2531602.2531672
Wang, Y., Min, Q., & Han, S. (2016). Understanding the effects of trust and risk on individual behavior toward social media platforms: A meta-analysis of the empirical evidence. *Computers in Human Behavior*, *56*, 34–44. doi:10.1016/j.chb.2015.11.011
Warren, S. D., & Brandeis, L. D. (1890). The right to privacy. *Harvard Educational Review*, *4*(5), 193–220.
Westin, A. F. (1967). *Privacy and freedom*. New York, NY: Athenum.

Author Biographies

Kelly Quinn (PhD, University of Illinois at Chicago) is a clinical assistant professor in the Department of Communication at the University of Illinois at Chicago. She has an interdisciplinary research focus on new media and how it intersects with such diverse areas as the life course, privacy, social capital, and friendship. Her recent studies have centered on how individuals conceptualize and navigate privacy online and the social and cognitive implications of social media use in older adults.

Dmitry Epstein (PhD, Cornell University) is an assistant professor in the Department of Communication and the Federmann School of Public Policy and Government at the Hebrew University of Jerusalem. His research program spans the topics of Internet governance, privacy, the digital divide, and online civic engagement in policy deliberation and decision-making. He currently serves as the chair of the Global Internet Governance Academic Network (GigaNet).

Brenda Moon (PhD, The Australian National University) is a data scientist at the Digital Media Research Centre at Queensland University of Technology in Brisbane, Australia. Her research uses interdisciplinary approaches to apply and develop digital methods. She has been investigating a range of approaches including looking at patterns in timeseries data, topic analysis, network analysis, image analysis, and working with large-scale social media data.